代数学讲义 (未定稿)

版本: 2024-10-20

李文威　著

https://www.wwli.asia

网络版

编译日期: 2024-10-20

版面: B5 (176×250mm)

本书将由北京大学出版社出版

李文威
个人主页: www.wwli.asia

目录

导言

大意

本书主题是初等意义上的代数, 面向数学和相关专业的低年级本科生或自学者. 内容属于现代数学及其应用的核心基础.

“代数” 的原义是解方程的技巧, 如今的代数更包括关于代数结构的一切研究; 对于何谓代数结构, 稍后将作初步解释. 两种涵义既有差异又密切相关; 其词义的演变史, 同时也是先贤们以具体问题为锚, 逐步锻造更简洁更有力的数学工具, 从而逐步接近数学实相的发展史.

本书将从代数的经典根源起步, 逐渐引入现代观点. 经典代数问题是解方程, 尤其是多项式方程组, 而其中相对容易的一类是线性方程组

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \vdots \qquad & \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m, \end{aligned}$$

此处 $X_1, \ldots, X_n$ 代表待解的变元. 应用中的许多问题都有此类表述. 对此, 经典的求解技术是 Gauss–Jordan 消元法和行列式理论, 它们是在称为矩阵的数学对象上操作的.

四则运算是表述并处理经典代数问题的必需. 现代意义的代数学探讨带有类似于加法或乘法等运算的集合, 称为代数结构, 以及这些集合之间保持运算的映射, 称为 “同态”. 代数学以集合论的语言表述. 例如 “域” 是带有四则运算的集合, 要求除法的分母非零, 并且服从于结合律, 交换律与分配律等种种性质; “环” 是舍弃除法运算和乘法交换律得到的结构. 域的初步实例是有理数域 $\mathbb{Q}$, 实数域 $\mathbb{R}$ 和复数域 $\mathbb{C}$, 它们都有四则运算. 环的初步实例是整数环 $\mathbb{Z}$, 其中的除法仅在整除情形才有意义; 尺寸为 $n \times n$ 的所有复矩阵也构成环 $\mathrm{M}_{n\times n}(\mathbb{C})$, 其乘法一般不满足交换律. 域和环的实例远不仅于此, 而代数结构也不仅限于域和环.

代数结构的另一则关键例子是某个域 F 上的向量空间, 其元素 (称为 “向量”) 可以彼此相加, 或者用 F 的元素来乘. 当域 F 给定, 向量空间之间的同态又称为线性映

射或线性变换. 向量空间理论是本书的重头戏, 而矩阵提供了具体操作向量空间和线性映射的工具.

从代数观点看, 线性方程组理论能够从向量空间和线性映射的视角来理解; 从几何观点看, 向量空间又能解释中学数学所熟知的平面和空间向量运算. 代数运算与几何直观由此得到关键的, 尽管还只是初步的综合.

除此之外, 常用的代数结构还有群 (对称性的体现), 模 (不妨视为一般的环上的向量空间), 它们对于经典问题都有精彩应用. 掌握了这些结构就能进一步涉足标准形, 张量积和 Witt 理论等较为复杂的概念. 当一切就绪, 读者便能有充足准备来探索代数学的进阶主题, 例如 [10] 或其它标准教材的内容.

方针

本书力图兼顾教材和自学的双重取向, 但由于底稿是讲义, 教材属性更加突出.

作为面向本科低年级, 主要是大一学生的教材, 本书对读者所要求的背景基本不超出高中理科数学的内容, 重叠部分亦不少; 之所以称 “基本”, 一则是因为高中教学内容因时而异, 各省市, 各校乃至于各人的状况又有不同; 二则是肯定人人皆有学习的良知良能, 体现为自见不足能努力追赶, 接触新知能虚心探索. 因此, 本书对于背景知识并非机械地取最大共约数, 而容许些许弹性; 目的是引人向上, 不是冰冷的门槛.

本科低年级的代数类课程一般是与数学分析同步学习的, 自学时也应当如此. 正如同数学分析的进阶内容涉及向量空间的语言 (例如线性微分方程), 随着本书的推进, 文中也会适度涉及一些数学分析的简单内容, 范围不超过定义及基本性质, 多数场合是作为例子或习题来运用.

在筹划本书的过程中, 笔者主要将代数学理解为一种教学单元, 它由学习者在特定阶段必须习得的一整套知识和技巧所构成, 这些主题皆与代数的词义有所联系. 其实, 各式教材所教授的 “代数” 多是一种标签, 或者说是整理知识的一套框架. 无论如何设想代数学的本质, 作为教学框架的 “代数学” 都不可能完全符应, 因为它是权衡诸多现实考量的产物; 一句话, 它属于教务范畴.

笔者以为教学框架的组织需要照顾至少四个点: (1) 概念本身的历史顺序. (2) 可读性. (3) 理论的规整, 流畅和优雅气韵, 业内人士谓之 “自然”. (4) 经济性, 体现为篇幅精简. 四点之间两两没有先天的一致性. 比方说, 行列式最早的表述方式既不易读, 也难言优雅或经济; 更自然的理论框架往往需要更长的预热; 经济和易读又趋于互斥, 除非思路有所创新. 本书侧重后三点, 而在难以兼顾时更倾向于第二和第三点. 至于代数类课程在传统上有何内容, 并非主要考量.

特别地, 笔者对现存的教学体系采取了批判地继承的态度. 教学方法与教学内容应当随时推移. 只要人类的数学事业不断前进, 学生掌握同等知识的时间便会不断提早, 而只要社会形式臻于良善, 则最优秀的教育资源必然会朝一切愿意学习的人平等敞开.

话虽如此, 必然仍有不少读者认为本书偏重经济性, 呈现为较高的学习坡度. 具体感受当然因人而异, 何况经济的考量虽然体现为冷酷外表, 内里却有真实的温柔, 因为

它悉心照顾人的有限性: 面对当代数学的宏伟大厦, 一生是过于短暂了.

本书对于数学史着墨不多, 特别是关于人的部分. 这方面的通俗书籍和网络资源颇为丰富, 顾及篇幅, 不必也不宜再添砖加瓦; 即便不知数学家的生平逸事, 损失的也只是谈资, 对于学习并无障碍.

最后对本书采用的术语和符号略作说明. 这两者都是社会共识的沉淀物, 出于约定俗成; 但术语是否达意, 以及符号是否明晰易懂, 运用中仍有客观的高下之分, 数学之外的明显例子是公制单位和英制单位的使用. 术语方面, 本书总体遵循 [9], 少数明显不妥处另译. 符号方面, 本书尽量遵循当前国际学界惯例, 这在多数情况下也是最明晰的写法.

提要

本书正文分成十六章, 书末有两则附录. 每一章末尾都附有若干习题, 基本按照各节内容来排序, 综合性的习题则不受此限. 多数习题带有提示. 原则上, 正文内容不依赖习题的结论, 但习题可能依赖于先前的其它习题.

各章和附录内容摘要如下, 所涉及的术语和概念将在正文中仔细解说.

▷ **第一章: 综观** 作为全书起点, 本章以解方程为线索, 简介代数学的经典渊源, 然后着力探讨线性方程组; 求解所用的 Gauss–Jordan 消元法简单高效, 后续将反复现身.

▷ **第二章: 集合, 映射与关系** 现代数学建立在集合论的语言上, 代数学尤其如此. 本章阐述全书所需的集合语言, 初学者需留意的重点包括集合的积和无交并, 等价关系与相应的商集, 以及基数的概念. 整数的算术和同余式虽是作为集合操作的示例而纳入本章, 但它们也是后续内容的重要线索.

▷ **第三章: 环, 域和多项式** 本章首先引入环和域的概念与实例, 它们是先前运用的整数集和各种数系的抽象化. 所谓的域, 既可以是由一些复数在四则运算之下生成的集合, 也可以是貌似更抽象的域, 譬如有限域; 后者有许多实际应用. 其次, 本章将严格地定义多项式及其算术, 并且明确作为抽象符号的多项式和多项式函数在一般的域上有何区别. 最后, 我们将介绍何谓域的特征.

▷ **第四章: 向量空间和线性映射** 向量空间在全书中扮演要角. 本章先从线性方程组的讨论入手, 提供代数的动机, 然后定义一个域上的向量空间, 基, 维数和线性映射. 矩阵虽然已在消元法的讨论中出现, 但在全书后续内容中, 它们更多地是作为具体操作线性映射的一种手段. 关于矩阵的大部分运算都有线性映射层次的对应物, 但矩阵技巧仍不可或缺.

▷ **第五章: 行列式** 就历史来看, 行列式也是由线性方程组求解所催生的概念, 其计算在高阶情形趋于复杂, 但仍有重要的理论价值. 本章将从置换的概念出发, 自

几何观点引入交错形式的概念, 再将行列式定义为 n 元交错形式在线性映射之下的缩放比例. 之后, 我们将以矩阵语言给出行列式的具体算法及一般性质. Cayley–Hamilton 定理和迹的讨论对于后续各章尤其重要.

▷ **第六章: 重访环和多项式** 本章回归多项式的讨论, 借助向量空间的理论作进一步的探究. 作为必要的准备, 本章开头将介绍环的理想, 相应的商环, 然后讨论整环的唯一分解性质; 这些内容属于代数结构的入门知识. 特别地, 本章末尾将以商环构造有 q 个元素的有限域, 其中 q 是某个素数 p 的幂.

▷ **第七章: 对角化** 大致上, 对角化是在共轭 (又称 "相似") 意义下将一个线性映射或 $n \times n$ 矩阵简化的技术. 本章先从线性递归数列的通项公式说明对角化的用处, 然后介绍何谓特征值和特征向量, 以此给出关于对角化的若干判准和算法. 并非所有 $n \times n$ 矩阵都能对角化, 然而本章末尾将证明在代数闭域 (譬如复数域) 上, 一切 $n \times n$ 矩阵都能上三角化.

▷ **第八章: 双线性形式** 双线性形式在数学及其应用中频繁出现, 它们也能以矩阵处理. 本章始于一般定义, 然后探讨非退化性质和线性映射的伴随, 这部分内容比先前各章稍加抽象. 在特征 $\neq 2$ 的域上, 对称和反对称双线性形式特别常见, 对称版本又称为二次型, 相关内容是本书的重点之一. 本章最后关于双重对偶和商的讨论也是相对抽象, 然而必须掌握的概念.

▷ **第九章: 实内积结构** 对于熟悉的平面和空间向量, 我们有长度和夹角的概念; 两者在一般的实向量空间中统合为称为内积的一类双线性形式. 本章讨论内积的基本概念和诸般性质, 然后相对于给定的内积深入探讨正交或自伴的线性映射. 后半部探讨的奇异值分解等主题皆关乎内积, 应用广泛; 由于实内积空间在数据科学等场景中自然地出现, 这是毫不意外的.

▷ **第十章: 复内积结构** 内积也可以在复向量空间上定义, 复内积不再是双线性形式, 而是所谓的半双线性形式. 它们和实内积有许多共性, 前一章的许多定理都有相应的复版本. 作为应用, 本章末尾将取道复数来推导实正交变换的标准形, 然后研究三维空间的旋转, 称为四元数的数学对象将在此发挥作用, 而不用环的概念便无法精准地理解四元数.

▷ **第十一章: 群的概念** 大而化之地说, 群是对称性的体现, 它在代数结构的谱系中占据比环, 域和向量空间更基本的地位. 经过之前各章的历练, 读者应当能迅速把握群的定义和基本性质; 特别地, 矩阵理论提供了关于群的大量实例. 本章后半部将从群论视角回首高次方程求解的问题, 接着确定空间中五种正多面体的对称群; 前者关乎群论的历史渊源, 后者是初等几何与代数技巧的二重奏.

▷ **第十二章: 模论入门** 域 F 上的向量空间可以设想为带有来自 F 的乘法运算的加法群. 若将 F 换成一般的环, 便有了模的概念. 模和向量空间既有通性, 也有许多

根本差异. 本章前半部着眼于模论的基本概念, 后半部则关注主理想环 (例如多项式环或整数环) 上的有限生成模, 这一特例对于矩阵或线性映射的标准形理论尤其有用, 能以矩阵来计算.

▷ **第十三章: 标准形** 标准形理论旨在判断两个 $n \times n$ 矩阵是否共轭, 或说是精确描述所有 $n \times n$ 矩阵的共轭类; 更加抽象地说, 其目的是分类 n 维向量空间上的单个线性映射. 本章将给出称为有理标准形和 Jordan 标准形的两种方案, 并解释其算法. 模论在此将显现其威力.

▷ **第十四章: 仿射空间与射影空间** 初等几何学中的向量能够以实向量空间的语言来理解, 虽然这摆脱了坐标系的桎梏, 但零向量依然表征了直观中一个选定的原点. 仿射空间可以大略地设想为不带原点的空间, 它们有纯粹代数的定义, 而且对许多几何问题更为自然; 例如在实仿射空间中可以自然地开展多面体和多面锥的严谨理论, 从代数观点看, 它们相当于线性不等式组的解集. 另一方面, 射影空间则可设想为向有限维仿射空间添加一系列 "无穷远点" 的产物. 这些概念无论在理论或实践方面都极有用.

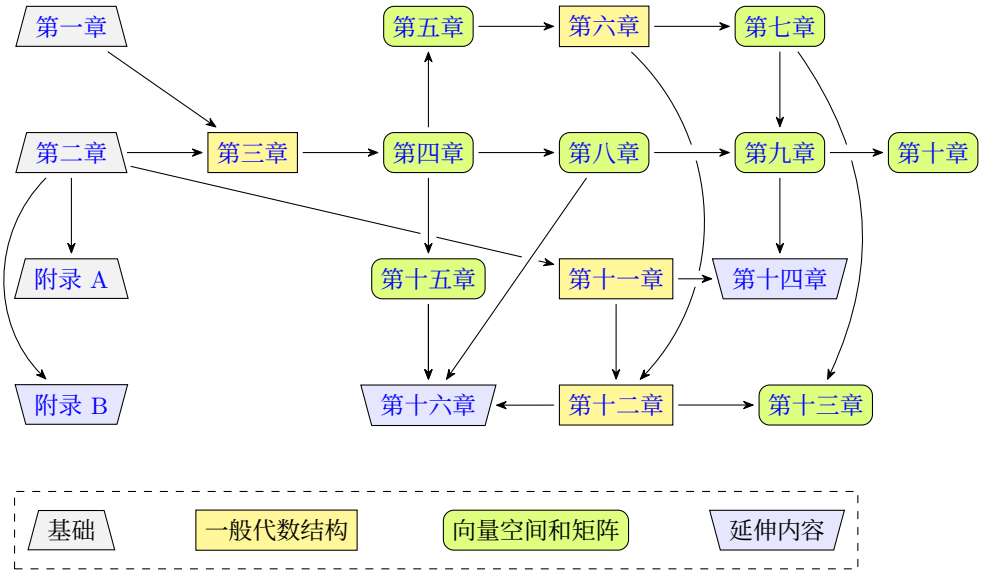
▷ **第十五章: 向量空间的张量积** 在选定的域上, 张量积是将两个向量空间 V 和 W 的元素形式地配对的一种手段, 由此得到新的向量空间 $V \otimes W$; 它的具体刻画适合以双线性形式的语言表述. 本章始于张量积的一般理论, 然后介绍对称代数与外代数, 及若干应用, 这些构造频繁用于几何等领域中.

▷ **第十六章: 二次型的 Witt 理论** 对于特征 $\neq 2$ 的域 F, 第八章已介绍过二次型的基本概念, 并在 F 是实数域或复数域的情形予以完整分类. 本章则借助更强的代数工具来处理二次型, 相关结论主要是 E. Witt 的贡献. 特别地, 我们将定义 F 上的 Witt 环和 Grothendieck–Witt 环, 这两种代数结构蕴藏关于 F 上的所有二次型的根本信息, 而且在 F 为实数域, 复数域或有限域的情形有简单描述. 许多结论都能进一步推及称为 Hermite 形式的结构.

▷ **附录 A: 集合论补遗** 内容包括从公理集合论严谨地构造非负整数集, 证明基数的一些简单性质, 以及介绍集合论中的 Zorn 引理. 代数学的一些底层事实依赖于 Zorn 引理, 例如它蕴涵每个向量空间都有基. 这些内容虽然基础, 但并非阅读正文所必需, 读者可酌情取用.

▷ **附录 B: 范畴引论** 范畴论是具有高度概括力与启发性的一套数学语言. 范畴和函子的定义仅需基本的集合论, 并且提供了理解一般代数结构的制高点; 但若没有处理具体结构的经验, 便不可能真正把握范畴的思想, 这也是本书置相关内容于附录的主因. 此处的材料只是引论, 并非系统性的介绍, 而且将频繁引用正文内容作为示例.

以下是各章和附录之间的依赖关系, 以及各章的大致属性, 由于涉及的有时只是每章之中的个别内容, 图示关系仅是概略的, 不必拘泥.



指引

对于学生 各章和附录之间的依赖关系已有图解, 请读者斟酌阅读顺序. 时间充足时, 循序阅读是最稳妥的. 前十章是代数学的初步事实, 应当扎实掌握. 后六章则相对进阶, 但总归属于代数学的基础部分.

第一章除 Gauss–Jordan 消元法之外的内容属于启发之用, 请放松阅读. 第十三章的标准形理论在一部分教材中占据核心地位, 它可以仅用矩阵语言来处理, 但本书选择以模论解释, 因此将相关内容后置. 第十四章和第十六章和其余部分关联较少, 应当在行有余力的前提下学习, 否则略去无妨.

两份附录的内容对于数学工作者皆属常识. 附录 B 未用于正文, 但考虑到范畴论的效用, 建议读者一旦对正文内容有相当掌握即可阅读, 或者跟随正文的学习进度反复查阅, 逐渐推进.

本书内容既有理论的或抽象的方面, 也有具体的一面. 数学的抽象方面让一部分初学者望而生怯, 对另一部分人又仿佛摇曳着诱人的光芒. 公允地说, 两方面对于代数学不可偏废. 一个人的抽象能力和具体能力就好比双手, 每人自有其惯用手, 左右开弓亦不乏其例, 但双手协作方能成事. 无论左撇子还是右撇子, 常人不会放任哪一只手萎缩, 更不至于自残; 然而在数学的学习中, 特别是一部分自学者群体, 倒是颇有些以思维能力残疾为荣的怪论. 所谓矫枉必须过正, 在当前的互联网时代, 问题更多在于对具体面向的忽视. 侈言对数学有单刀直入的顿悟, 却连本书为数不多的具体内容都无法掌握, 无有是处.

书中大部分抽象定义都附有相应实例, 算法或关于其思路的解释, 读者应力图全部掌握, 否则对定义的理解便不完整. 各章节中间穿插的练习也有助于巩固学习成果. 这

些练习或者是直接的验证, 或是例行的简单计算, 又或者有详细提示; 一部分简单练习对于后续内容还是必要的. 因此读者应当尽量在进入下一节之前完成所有练习.

具体性的一个重要面向是算法, 另外则是图像, 或谓几何直观. 不妨将图像的角色分成两类.

- ⋆ 问题或定义本身即基于图像, 例如实内积的基本性质, 三维空间的正交变换, 多面体等; 几何方法对此自是题中之义, 而问题本身的表述也应该伴以图像, 甚至于“以图为证”, 目击而道存.

- ⋆ 图像协助思维, 例如用来理解低维特例, 或作为证明的手段, 又或者提供理解问题的新视角.

无论对于哪种情形, 几何直观皆可谓船坚炮利, 但它并非自外部强加的; 倘若将图像, 尤其是呈现于纸张或显示设备上的影像 (相对于真实无妄的 “心像”) 执为解释数学概念的排他准则, 它们便不再是工具而是枷锁了. 在初学一般维数或一般域上的向量空间时, 尤其应当有此认知.

各章末尾的习题可能比练习费力, 然而难题多有提示, 部分题目有前后承接的关系, 建议读者尽力完成, 一时无法做完也不影响后续阅读. 习题包含一定数量的计算题, 特别是在早期各章. 不经手算则难以理解算法的本质, 读者勿等闲视之.

安排习题的目的是帮助读者掌握书中内容和拓展知识. 特别地, 本书习题不为应试, 然非不能应试; 只要读者充分掌握书中内容和习题, 取得高分不会有任何困难.

对于教师 首先是个人经验的方面. 笔者曾经基于本书内容讲授一学年的本科一年级基础课程, 每周 4 节课, 每学期包括考试在内计有 16 周. 授课对象主要来自北京大学数学科学学院, 以及信息科学学院, 元培学院等, 但不限于此.

虽然对象以数学专业的学生为主, 然而不属于实验班系列课程, 默认的背景知识仍不超过高中理科数学, 仅假定学生扎实掌握课内知识, 并且具有一定程度的探索精神.

粗分到章, 笔者能触及除了第十四章和两份附录以外的内容. 细分到节, 则有跳过者 (例如 §5.11, §6.5, §11.8, §11.11, §15.8, §15.9, 第十六章的大部分内容, 等等), 有略讲者 (§1.2, 第二章关于集合论的细部讨论, §3.6, §6.8, §§6.10–6.11, §11.10 等等). 其中一部分确实只能割舍, 一部分是考虑到多数学生已有涉猎, 另有一部分则由课后作业和助教予以补充.

授课时采用板书. 使用投影片能够提速, 但教学效果难免受影响. 笔者未制作配套的投影片课件, 如有同行愿意从事这方面的工作, 笔者无任欢迎.

在讲法方面, 巨细靡遗地讲授是不切实际的, 也不是最优解. 由于编撰时已经顾及自学需求, 只要教师充分了解学生情况, 留一些支线内容自学是合理的. 此外, 习题课在授课过程中起到关键作用. 按学院安排有两节习题课, 上学期每周, 下学期隔周; 助教们除解题和答疑外还补充了不少内容. 如果能在整个学年内安排每周至少三节的习题课, 效果想必更佳.

在课时受限或需要和别班同步的情况下, 可以考虑省略或延迟关于商空间的相关内容. 如果数学分析课程能覆盖集合论, 省去为佳. 如果学生学过初等数论, 则算术入门和同余式也可以省去或快速带过.

布局方面, 本书不刻意将代数学中的线性部分 (线性方程组, 矩阵, 向量空间) 与其它部分相区隔, 也不鼓励这种安排.

对于向量空间, 笔者认为无论就理论或应用考量, 都应该在一般的域上处理, 但选择从包括线性方程组在内的解方程问题切入. 考量之一是为环和域提供具体线索, 顺带衔接高中内容. 二则是因为若不掌握消元法, 则即便对于标准 n 维空间中的有限向量组, 都缺少具体算法来判定其是否线性相关, 或者是否生成全空间. 从教学的立场看, 在线性方程组之前引入向量空间未必合适.

如果课时充裕, 不妨在课程或习题课中讲解数学软件的使用. 如果教师选择介绍范畴论, 宜安排在课程最后.

习题方面, 本书的计算类题目相对较少, 尤其是在矩阵与行列式的部分; 有需求的教师不难以其它文献满足. 随着本书内容的深入, 计算题与证明题的边界将逐渐消失.

本书没有超纲题目. 相对困难的证明题常有提示; 对于专业人士, 提示几近于完整证明, 或者提供了证明的详细蓝图. 一部分习题是笔者认为重要的拓展知识, 另有部分习题则是代数学中广为人知的事实, 教师们应该不难找到参考材料.

致谢

本书的主体基于笔者于 2020–2024 年间在北京大学讲授的本科课程, 个别内容基于先前在中国科学院大学的授课材料. 谨向全体同学和助教们致谢, 没有他们的参与和反馈就不可能有这部讲义.

撰写过程中, 笔者广泛参考了既有的著作, 包括北京大学的系列教材, 以及许多从本科时期便给予笔者启发的经典名著. 书繁不及备载, 谨向众多先辈和同行们致敬. 珠玉在前, 笔者生硬缝补者有之, 擅出己意者有之, 祈望读者谅解.

此外, 也感谢北京大学出版社陈小红和曾琬婷两位编辑的专业工作和宝贵意见.

最后, 许多师友和读者对书稿提供了建议和指正, 无法一一备述, 记忆所及者包括: 丁一文, 高剑伟, 孙超超, 万铖睿, 薛江维 (按照姓名拼音排序). 在此致谢.

♣♣♣ 致谢待补: 更长名单.

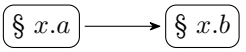
李文威

2024 年秋

于门头沟

体例

阅读顺序　为了帮助读者制定计划, 各章和附录的开头附有各节的阅读顺序, 其中图形



的意涵是 § $x.a$ 必须先于 § $x.b$ 阅读, 或者说 § $x.b$ 依赖于 § $x.a$. 由于具体的依赖程度有所区别, 图形只提示大致顺序, 并非绝对. 无论如何, 不等式 $a < b$ 在图中恒成立, 因此只要时间充裕, 按编号顺序阅读总是可行的方案.

标签　遵循数学教材的惯例, 本书将重要的陈述分成以下几类, 赋予标签和编号.

▷ **定义**　关于数学对象或数学概念的界定.

▷ **约定**　主要用于解释符号或习惯用语.

▷ **命题**　泛指各种数学命题的陈述.

▷ **定理**　较为重要或者具有总结性的数学命题, 通常是各章节的核心结论.

▷ **引理**　为了证明定理或命题而预备的辅助性结果, 其表述或证明往往比较复杂. 因此, 从技术的观点来看, 引理也同样包含数学的精华.

▷ **推论**　从先前陈述的命题或定理简单推得的结论.

▷ **例**　关于定义或命题的具体实例或简单应用, 经常附有简短的论证.

▷ **注记**　关于定义或命题的讨论或补充说明, 为了强调或方便参照而加以标明.

▷ **练习**　为了巩固读者对先前内容的了解而布置的练习, 夹杂在正文中间. 练习往往是简单的操演, 通常带有提示; 许多练习对于熟悉相关内容的读者应当是不证自明的. 一部分练习可能被之后内容所引用, 因此读者应当尽力在阅读途中完成所有练习.

▷ **算法**　对所论的数学问题提供适合于笔算或编程的详细步骤, 然而不涉及代码, 仅以自然语言表述. 算法和数学命题或其证明的界限有时是模糊的.

定义和数学命题经常可以混搭, 这是因为一则定义的合理性往往需要论证. 由此便产生了定义–定理, 定义–命题之类的名目.

命题, 定理, 引理和推论之下通常紧接着给出证明, 除非是特别明显的推论, 或一些较为深入或离题的证明, 在后一场合将给出参考书目.

证明的结尾以 □ 标记. 在一部分场合, 为了增进阅读体验, 证明也可能延后给出, 甚至并入附录.

编号 为了方便后续的参照, 书中的定义, 命题等内容都按照

$$z.j.s, \quad z = \text{章}, \ j = \text{节}, \ s = \text{顺序}$$

的格式连续编号, 例如 “定理 1.1.1”. 数学公式按照同样模式编号, 但是为了方便区分, 另加圆括号记为

$$(z.j.s)$$

的格式, 例如 “(1.1.1)”.

对于书中的各章, 节和附录, 我们将按照诸如

第一章, §1.1, §§A.1–A.3

的格式进行参照.

符号

习惯用语 表达式 $A := B$ 意谓 “A 被定义为 B”.

为了便利数学命题的表述和阅读, 当我们写下诸如 “设 A (或 B, 或 C) ..., 则 A' (或 B', 或 C') 成立” 的语句, 其意涵是: 设 A ... 则 A' 成立, 设 B ... 则 B' 成立, 设 C ... 则 C' 成立.

下述形式的写法也将频繁出现

$$f = \begin{cases} g, & \mathcal{P}, \\ h, & \mathcal{Q}; \end{cases}$$

它的意涵是当 $\mathcal{P}$ 成立时 $f = g$, 当 $\mathcal{Q}$ 成立时 $f = h$.

当我们说一个数学对象 (譬如一个数, 映射, 集合...) 是 “良定义” 的, 是指我们定义它的方法没有歧义, 不依赖定义过程中任何辅助资料的选取, 从而给出一个确定的数学对象.

逻辑符号 本书所使用的数学语言以基本的逻辑符号为底层, 主要是逻辑连接词和量词:

	连接词				量词	
符号	$p \wedge q$	$p \vee q$	$p \implies q$	$\neg p$	$\forall x$	$\exists x$
解读	p 而且 q	p 或 q	p 蕴涵 q	非 p	对所有 x	存在 x

对于 $p \implies q$, 常见的说法是 “p 仅当 q” 或者 “q 当 p”. 特别地, $p \iff q$ 相当于说 “p 当且仅当 q”.

本书并不会在严格规范下使用形式语言, 只是偶尔将 $\wedge$, $\vee$, $\forall$, $\exists$ 和 $\iff$ 等符号作为方便的简写来使用.

数学语句只能为真或为假, 不许两者兼具或皆非. 语句的真假可以通过真值表来剖析. 举例来说, 读者应当明白语句 "p 蕴涵 q" 的真假是按照下表规定的:

p	q	p 蕴涵 q
真	真	真
真	假	假
假	真	真
假	假	真

特别地, p 真方有可能触发 "p 蕴涵 q" 为假. 基于这一道理, 涉及全称量词的语句

$$\forall \text{ 满足性质 } \mathcal{P} \text{ 的 } x, \quad q \text{ 成立.}$$

应当被理解为

$$\forall x, \quad \big(x \text{ 满足性质 } \mathcal{P} \implies q \text{ 成立}\big),$$

它仅在确实存在满足性质 $\mathcal{P}$ 的 x 时才可能触发为假. 若不存在这般的 x, 或者说当全称量词 $\forall$ 取在空集上, 则此语句按规定为真.

集合 由元素 $a, b, c, \ldots$ 构成的集合记为 $\{a, b, c, \ldots\}$; 对于任意集合 S, 符号 $s \in S$ 代表 s 是 S 的元素, 而 S 中满足某个给定性质 $\mathcal{P}$ 的元素所成的子集表作

$$\big\{s \in S : s \text{ 满足 } \mathcal{P}\big\} \quad \text{或} \quad \big\{s \in S \,\big|\, s \text{ 满足 } \mathcal{P}\big\}.$$

符号 $S \subset T$ 代表集合 S 包含于 T, 容许相等[1)]; 若 $S \subset T$ 而 $S \neq T$, 则称 S 严格包含于 T, 或称 S 是 T 的真子集, 记为 $S \subsetneq T$.

集合 A 对 B 的差集记为 $A \setminus B := \{a \in A : a \notin B\}$.

数组的符号是 $\boldsymbol{x} = (x_1, \ldots, x_n)$ 或 $(x_i)_{i=1}^n$ 的形式; 计顺序, 也容许重复. 这种记法当然适用于 x_i 为其他数学对象的情形, 而下标 i 也可以遍历一般的集合 I 而不只是正整数, 记如 $(x_i)_{i\in I}$ 的形式, 或简记为 $(x_i)_i$. 数组中的 x_i 称为 $\boldsymbol{x}$ 的第 i 个分量或坐标.

[1)] 一些教材将这里的 $\subset$ 写作 $\subseteq$, 而以 $\subset$ 表示严格包含关系.

数系 熟知的几种数系记为

$$\begin{array}{ccccccc} \mathbb{Z} & \subset & \mathbb{Q} & \subset & \mathbb{R} & \subset & \mathbb{C} \\ \text{整数集} & & \text{有理数集} & & \text{实数集} & & \text{复数集.} \end{array}$$

非负整数集记为 $\mathbb{Z}_{\geq 0}$, 正整数集记为 $\mathbb{Z}_{\geq 1}$, 正实数集记为 $\mathbb{R}_{>0}$, 依此类推.

整数 a 和 b 的最大公因数记为 $\gcd(a,b)$, 最小公倍数记为 $\operatorname{lcm}(a,b)$.

对所有实数 x, 记不超过 x 的最大整数为 $\lfloor x \rfloor$, 换言之 $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$.

虚数单位记为 $\mathrm{i} \in \mathbb{C}$, 它满足 $\mathrm{i}^2 = -1$. 当 $D \in \mathbb{R}_{\leq 0}$ 时, 本书规定 $\sqrt{D} := \mathrm{i}\sqrt{|D|}$; 作为特例, $\mathrm{i} = \sqrt{-1}$.

复数 z 的实部记为 $\mathrm{Re}(z)$, 虚部记为 $\mathrm{Im}(z)$. 复数 z 的共轭记为 $\bar{z}$.

多项式 本书的惯例是以大写字母 $X, Y, Z, \ldots$ 代表多项式的变元, 亦即自变量. 在必须强调变元的场合, 我们也记以 $X, Y, Z, \ldots$ 为变元的多项式 f 为 $f(X, Y, Z, \ldots)$, 而对 f 代值 $X = x, Y = y, \ldots$ 的产物记为 $f(x, y, \ldots)$.

映射 从集合 A 到 B 的映射 f 常以箭头符号写作 $f : A \to B$, 其像集记为

$$\mathrm{im}(f) := \{f(a) : a \in A\} \subset B.$$

对任意子集 $A' \subset A$, 记 $f|_{A'} : A' \to B$ 为 f 在 A' 上的限制; 另外记 A' 在 f 下的像为 $f(A') := \mathrm{im}(f|_{A'})$. 在必须强调映射以 A 的元素为 "输入" 的场合, 我们也采取类似 $f(\cdot)$ 的记法.

映射 $A \xrightarrow{f} B$ 和 $B \xrightarrow{g} C$ 的合成记为 $g \circ f : A \to C$, 简记为 gf. 具体定义是

$$(gf)(a) = g(f(a)), \quad a \in A.$$

为了区别集合 A 和其元素在映射 $f : A \to B$ 下的像, 我们经常以符号 $f : a \mapsto b$ 或 $a \overset{f}{\mapsto} b$ 代表 $f(a) = b$. 在具体描述一个映射时, 我们将频繁使用诸如

$$\begin{aligned} f : A &\to B \\ a &\mapsto f(a) \end{aligned}$$

的写法.

任意集合 A 到自身的恒等映射 $a \mapsto a$ 记为 id_A, 不致混淆时也简记为 id.

谨介绍关于映射的几则标准术语和符号.

术语	定义条件	符号
单射 (或**嵌入**)	$a \neq a' \implies f(a) \neq f(a')$	$f : A \hookrightarrow B$
满射	对每个 $b \in B$ 都存在 $a \in A$ 使得 $f(a) = b$	$f : A \twoheadrightarrow B$
双射 (或**一一对应**)	既单又满	$f : A \xrightarrow{1:1} B$

所以 $f : A \to B$ 是满射当且仅当 $\mathrm{im}(f) = B$. 若 f 是双射, 其逆映射[2] f^{-1} 映 $f(a) \in B$ 为 $a \in A$.

对于一般的映射 $f : A \to B$ 和任意子集 $B' \subset B$, 记

$$f^{-1}(B') := \{a \in A : f(a) \in B'\},$$

称为 B' 对 f 的**原像**或**逆像**.

注意: 本书视 "映射" 与 "函数" 为同义词. 在其他语境中, "函数" 有时意指映至复数集的映射, 这种区别只是语言习惯.

连加与连乘 我们经常使用连加与连乘符号

$$\sum_{k=1}^{n} a_k = a_1 + \cdots + a_n, \quad \prod_{k=1}^{n} a_k = a_1 \cdots a_n,$$

及其种种变体. 当连加 (或连乘) 的下标集为空时, 将对应的空和 (或空积) 规定为 0 (或 1) 是很方便的.

作为连乘的特例, 阶乘定义为 $n! = \prod_{k=1}^{n} k$, 而 $0! = 1$. 本书将二项式系数记为

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}, \quad 0 \le k \le n.$$

矩阵 大致地说, 矩阵是以横行竖列的表格形式来表示的数组, 它们的完整定义和深入研究是本书正文的主题, 此处仅解释符号. 具有 m 行 n 列的矩阵称为 $m \times n$ 矩阵, 本书将这种矩阵写作

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{pmatrix} \begin{matrix} \\ \text{第 } i \text{ 行} \\ \\ \end{matrix}$$

第 j 列

的形式; 称 a_{ij} 为 $\boldsymbol{A}$ 的第 (i, j) 个矩阵元, 取在选定的数系 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 或更广泛的代数结构中; 这些矩阵所成的集合相应地记为 $\mathrm{M}_{m\times n}(\mathbb{Q})$, $\mathrm{M}_{m\times n}(\mathbb{R})$, $\mathrm{M}_{m\times n}(\mathbb{C})$, 依此类推. 有时对 $m \times n$ 矩阵也采取 $\boldsymbol{A} = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}}$ 的记法.

上述矩阵 $\boldsymbol{A}$ 的转置是交换行和列的角色所给出的 $n \times m$ 矩阵, 记作

$${}^{\mathrm{t}}\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} = (a_{ji})_{\substack{1\le i\le n \\ 1\le j\le m}}.$$

[2] 用函数的术语来说, 逆映射就是反函数.

本书经常将矩阵元为零的部分留白表示, 例如 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

行列式 行列式的严格定义同样是本书正文的主题. 在符号方面, 本书记 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})_{1 \leq i,j \leq n}$ 的行列式为

$$\det \boldsymbol{A} \quad \text{或} \quad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

的形式, 其中 $n \in \mathbb{Z}_{\geq 1}$. 在 $n = 0$ 时赋予行列式定义可以简化一些命题的表述: 在此场合, 规定相应的 “空行列式” 为 1.

单位 本书谈及角度时一律采取弧度制.

第一章 综观

从古至今, 代数一词的内涵历经种种变化. 本章目的是以方程求解为线索, 在中学数学的基础上简述代数的源与流. 在本章触及的各种代数问题中, 线性方程组和相应的 Gauss–Jordan 消元法 (§§1.3–1.4) 相对简单, 同时又是探究其他问题的基础, 它们将在全书的前半部分承担关键角色. 其余部分虽然和后续内容没有严格的逻辑先后关系, 但涉及的例子和思路仍将反复回响.

阅读顺序

§1.1 → §1.2　　§1.3 → §1.4

1.1 何谓代数

"代数" 一词源于公元 9 世纪左右波斯学者 al-Khwārizmī 的著作 *Al-Kitāb al-mukhtaṣar fī ḥisāb al-jabr wa'l-muqābala*, 其中的阿拉伯语名词 *al-jabr* 在拉丁文中被转写为 *algebra*, 亦即欧洲各地语言中的代数. 这部著作标志了代数学自西方数学传统中脱胎而出的第一步. 尽管代数学的内涵和应用范围嗣后大有扩张, 我们无妨先沿这条历史线索上溯, 特别是从词源入手, 来探索代数学关注的基本问题及其风格.

且看 al-Khwārizmī 的大作. 书名中的 al-jabr 和 al-muqabala 可以大略地翻译为解方程时的移项和相消. 他在书中考察的问题是一元一次和二次方程, 我们以现代语言简要地回顾.

⋆ 一次方程 $aX+b=0$, 其中 X 是变元, 而 a 和 b 是给定的系数, $a\neq 0$. 为了求解 X, 先作移项得到 $aX=-b$, 然后两边同除以 a, 得到唯一解 $X=\frac{-b}{a}$.

⋆ 二次方程 $aX^2+bX+c=0$; 仍然设 $a\neq 0$. 以配方法消去一次项, 将方程转化为

$$a\left(X+\frac{b}{2a}\right)^2-\frac{b^2}{4a}+c=0,$$

对之移项, 同除以 a, 开方再移项, 得到熟知的二次方程公式解

$$X=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

清代数学家李善兰将这些初等操作解释为 “补足相消之术”, 两种技巧在一次和二次方程的实例中已经悉数体现了. 之所以能如此列出一般的方程, 并讨论其通解, 前提是以抽象的变元 X, Y 等来代替具体之数. 在中国数学史上, 不晚于金, 元之际所出现的 “天元术”, 也正是以代表未知数的变元来布列方程的一种手段.

有鉴于此, 古典意义的代数学可以理解为以变元代替具体数字, 通过移项等运算来求解方程的一门技艺.

这个初步定义又引出一系列的问题: 何谓数字, 何谓方程, 以及更重要的是: 何谓技艺? 我们将从几个初步例子来审视这些问题.

数的概念 自从人猿相揖别, 在人类所发现或曰创造的各类 “数” 中, 最基本的是正整数 $1, 2, 3, \ldots$, 或称自然数. 数字感是人心的基本官能, 所有古文明都有各自的计数体系, 至于说计数在多大程度上是一种先天能力, 又是成熟于哪一个成长阶段, 这些问题最好留给发展心理学来回答.

考虑正整数之间的比例, 就顺理成章地得到了正有理数. 至于负数, 数学史家习惯将其溯源至欠账或财务亏损的表达法, 尽管负数因此具备了冷峻的实在感, 但它在西方文明登堂入室的时间要晚得多: 无论古希腊学者或 al-Khwārizmī 的著作都只论正系数的方程和正数解. 相反地, 中国汉代的《九章算术》则毫无心理负担地接受了负数, 并给出了相应的运算法则.

加, 减, 乘法在整数集 $\mathbb{Z}$ 上通行无阻; 加减乘除 (要求除数非零) 在有理数集 $\mathbb{Q}$ 上通行无阻. 由此观之, 可以说数的概念愈广, 操作就愈方便.

与计数同样基本的另一心理官能是度量, 包括长度, 面积和体积. 从给定的单位长度 1 出发, 有理数并不足以丈量生活中所有的几何对象, 比如等腰直角三角形的斜边长, 或圆面积等问题都避不开无理数. 古巴比伦人不担心这类问题: 以单位等腰直角三角形的斜边长为例, 他们直接取 $\sqrt{2}$ 在 60 进制下的近似值

$$1 + \frac{24}{60} + \frac{51}{60^2} + \frac{10}{60^3} = 1.41421296...$$

古希腊学者不能接受这种办法. 如果说万物皆数, 那么斜边的长度又该如何安放? Euclid 在《几何原本》中的权宜之计是将数字的运算化为关于线段长度的操作, 从而绕开了这个问题. 而按现今的观点, 线段的长度无非是正实数. Euclid 的进路影响深远, 但是依后见之明, 《几何原本》中几何化的代数操作既不自然也不方便, 尽管它符合希腊文明所推崇的严格性.

对于解方程, “数” 的界定不只是哲学问题, 它直接决定我们可以对方程进行哪些操作, 以及有哪些解是可行的. 以二次方程 $aX^2 + bX + c = 0$ 为例, 若判别式 $b^2 - 4ac < 0$ 则方程无实数解. 依此, 似乎可以说二次方程逼出了复数理论, 但这一理由显得牵强: 既然复数似乎不代表常人知觉所领纳的任何数量或度量, 直接规定这样的方程无解岂不干脆? 引入复数究竟有何好处, 又有多大必要? 完整的解释需要较为深入的数学知识. 比方说, 复数具有以下的重要性质.

定理 1.1.1 (代数基本定理) 设 $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ 为以 X 为变元的复系数 n 次多项式, 其中 $n \in \mathbb{Z}_{\geq 1}$, 则存在 $x_1, \ldots, x_n \in \mathbb{C}$ 使得

$$f = \prod_{k=1}^{n} (X - x_k).$$

这些 $x_1, \ldots, x_n$ 无非是多项式 f 的复根 (计入重数); 精确到重排, 它们是唯一的.

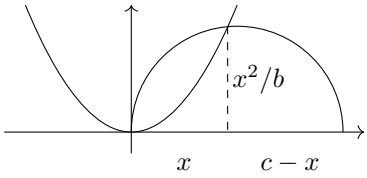
特别地, 不仅限于二次多项式, 任意非常数多项式都有复数根.

相对于眼下所探讨的初等代数, 代数基本定理是一则超纲的结果, 其所有证明或多或少都涉及数学分析, 读者可以参考 [11, §3.4 定理 6]. 我们还是暂且回归历史脉络. 引入复数的一个重要动力是三次方程的研究.

热身: 三次方程的 Cardano 公式 公元 11 世纪左右, 波斯学者 Omar Khayyam 发现了如何以圆锥曲线的交点表达三次方程的正根, 以方程

$$X^3 + b^2 X = b^2 c$$

为例, 其中 $b, c > 0$, 考虑抛物线 $X^2 = bY$ 和位于 Y 轴右侧, 直径为 c 而切原点的圆. 将两者在上半平面的交点表成 $(x, x^2/b)$, 如下图所示:



根据相似三角形的性质可知

$$\frac{x}{x^2/b} = \frac{x^2/b}{c - x},$$

整理后即是 $x^3 + b^2 x = b^2 c$.

对于其他种类的实系数三次方程, Khayyam 同样给出了求根的几何构造. 此法简则简矣, 却未能像一次或二次方程的情形一般给出明确公式. 它无论在理论或实践层面的价值都相当有限.

三次方程的第一个完整通解是意大利学者 G. Cardano 在 1545 年出版的著作 *Ars Magna* 中写下的. 考虑一元三次方程

$$X^3 + aX^2 + bX + c = 0.$$

其中的系数 a, b, c 可以是任意实数乃至于任意复数, 无关宏旨. 求解的第一步是命 $Y = X + \frac{a}{3}$, 将原方程化为

$$Y^3 + pY + q = 0$$

的形式. 由于今后仅考虑形式如上的三次方程, 不妨以 X 代 Y. 后续思路是寻求形如 $X=u+v$ 的解. 将此代入 $X^3+pX+q=0$, 给出

$$\left(u^3+v^3+q\right)+(u+v)(3uv+p)=0.$$

假若能找到 u 和 v 满足方程组

$$\begin{aligned} u^3+v^3+q&=0,\\ 3uv+p&=0, \end{aligned}$$

则 $u+v$ 便是原三次方程的解. 将第一式两边同乘以 u^3, 得出 $u^6+(uv)^3+qu^3=0$, 或改写成

$$u^6+qu^3-\left(\frac{p}{3}\right)^3=0. \tag{1.1.1}$$

若能解 (1.1.1) 和 $uv=-\frac{p}{3}$ 的联立, 便有望得到三次方程的解.

我们称 (1.1.1) 为原方程的辅助方程. 尽管辅助方程是 u 的 6 次方程, 它关于 u^3 却是二次的. 原三次方程的判别式定义为

$$D:=-4p^3-27q^2. \tag{1.1.2}$$

解辅助方程, 得

$$u^3=-\frac{q}{2}\pm\sqrt{\left(\frac{p}{3}\right)^3+\left(\frac{q}{2}\right)^2}=-\frac{q}{2}\pm\frac{\sqrt{-3D}}{2\cdot 3^2}.$$

由熟悉的代数运算可得

$$\begin{aligned} \left(-\frac{q}{2}+\frac{\sqrt{-3D}}{2\cdot 3^2}\right)\left(-\frac{q}{2}-\frac{\sqrt{-3D}}{2\cdot 3^2}\right)&=\frac{3^4q^2+3D}{2^2 3^4}\\ &=\frac{-3\cdot 4p^3}{2^2 3^4}=\left(-\frac{p}{3}\right)^3. \end{aligned} \tag{1.1.3}$$

所以只要取 u 和 v 分别为 $-\frac{q}{2}+\frac{\sqrt{-3D}}{2\cdot 3^2}$ 和 $-\frac{q}{2}-\frac{\sqrt{-3D}}{2\cdot 3^2}$ 的立方根, 则 u 和 v 都是辅助方程 (1.1.1) 的解. 我们断言:

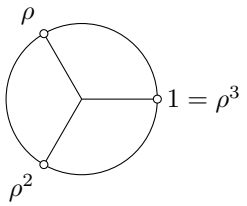
- ⋆ 立方根可以适当地取, 以确保 $uv=-\frac{p}{3}$;
- ⋆ 承上, 此时 $u+v$ 是三次方程 $X^3+pX+q=0$ 的解.

分成一般情形 $p\neq 0$ 和例外情形 $p=0$ 来讨论. 在 $p\neq 0$ 的前提下, 任意取立方根 u 和 v, 命 $\lambda:=uv/\left(-\frac{p}{3}\right)$, 则 (1.1.3) 表明 $\lambda^3=1$. 以 $\lambda^{-1}v$ 代 v 或以 $\lambda^{-1}u$ 代 u, 即可确保 $uv=-\frac{p}{3}$. 之前关于辅助方程的讨论已说明此时 $u+v$ 给出原三次方程的解.

对于 $p=0$ 的例外情形, 我们有 $-3D=3^4q^2$, 而 $-\frac{q}{2}\pm\frac{\sqrt{-3D}}{2\cdot 3^2}$ 中必有一者为 0. 所以此时 $uv=0=-\frac{p}{3}$ 自动成立, 而且 $u^3+v^3=-q$. 所以在例外情形下 $u+v$ 依然给出解. 断言于是得证.

为了得到 $X^3+pX+q=0$ 的所有解, 取

$$\rho := \frac{-1+\sqrt{-3}}{2}, \quad \text{在复数平面上：}$$



于是 $-\frac{q}{2}+\frac{\sqrt{-3D}}{2\cdot 3^2}$ 和 $-\frac{q}{2}-\frac{\sqrt{-3D}}{2\cdot 3^2}$ 的立方根分别是

$$u, \rho u, \rho^2 u \quad \text{和} \quad v, \rho v, \rho^2 v.$$

这些立方根可以适当地排序, 记为 u_1,u_2,u_3 和 v_1,v_2,v_3, 使得 $u_iv_i=-\frac{p}{3}$ 对 $i=1,2,3$ 皆成立. 这就给出了 X^3+pX+q 的三个根 (容许重复):

$$x_i = u_i + v_i, \quad i=1,2,3. \tag{1.1.4}$$

练习 1.1.2 说明 x_1, x_2, x_3 两两相异当且仅当判别式 $D\neq 0$.

提示〉 考虑先前论证中的立方根 u, v, 满足 $uv=-\frac{p}{3}$; 说明 $u+v, \rho u+\rho^2 v, \rho^2 u+\rho v$ 有所重复当且仅当 $u\in\{v,\rho v,\rho^2 v\}$. 若此条件成立, 则两边取立方给出

$$-\frac{q}{2}+\frac{\sqrt{-3D}}{2\cdot 3^2} = -\frac{q}{2}-\frac{\sqrt{-3D}}{2\cdot 3^2},$$

以此说明 $D=0$.

至此, 我们几乎已完成原方程的求解. 然而推导过程针对系数 p, q 排除了一些例外情形: 确切地说, 我们默认了 x_1,x_2,x_3 两两相异, 才能说它们穷尽了原方程的解. 这些都不是本质的困难. 事实上, 对所有 p 和 q 都可以按上述方法构造 u_i 和 v_i, 其中 $i=1,2,3$, 并且从根与系数的关系验证因式分解 (练习 1.1.3):

$$(X-(u_1+v_1))(X-(u_2+v_2))(X-(u_3+v_3)) = X^3+pX+q.$$

以上考虑的都是方程的复数解, 而在推导过程中论及平方根和立方根时, 也一律在复数域中操作, 否则无以穷尽 $X^3+pX+q=0$ 的通解. 这就产生了一些耐人寻味的观察.

1. 以方程 $X^3-15X-4=0$ 为例, 判别式 (1.1.2) 为 $D=13068$. 易见

$$X^3-15X-4 = (X-4)(X^2+4X+1);$$

由此知它有三个相异实根. 另一方面, Cardano 公式却给出形如

$$X = \sqrt[3]{2+\sqrt{-121}} + \sqrt[3]{2-\sqrt{-121}}$$

的通解, 其中 $\sqrt[3]{\cdots}$ 代表适当选取的立方根.

2. 推而广之, 若系数 $p,q\in\mathbb{R}$ 而 $D>0$, 则可以用初等方法说明 X^3+pX+q 的三个根全实. 尽管如此, 为了对一个现实的三次多项式写下同样现实的三个根, Cardano 公式中的 $\sqrt{-3D}$ 和 u_i, v_i 却无可避免地要容许为复数. 这和二次方程的情形明显不同.

正因如此, Cardano 公式是促使数学家们接受复数的重要推力之一.

练习 1.1.3 请回顾三次方程 $X^3+pX+q=0$.

(i) 尝试验证 $\prod_{i=1}^3(X-(u_i+v_i))=X^3+pX+q$. 换言之, Cardano 公式的确给出 X^3+pX+q 的所有根, 计入重数.

(ii) 设 $p,q\in\mathbb{R}$. 验证当 $D>0$ 时三根都是实数. 〈提示〉 令 $\rho:=\frac{-1+\sqrt{-3}}{2}$. 我们希望对 $i=1,2,3$ 证明 $\overline{u_i+v_i}=u_i+v_i$, 这里 $z\mapsto\bar z$ 表复数的共轭运算. 由 $\overline{u_i}^3=\overline{u_i^3}$ 可推得 $\overline{u_i}^3=v_i^3$, 因而存在 $k\in\{0,1,2\}$ 使得 $\overline{u_i}=\rho^k v_i$. 再取一次共轭得到 $\overline{v_i}=\rho^k u_i$. 又由于

$$\overline{u_iv_i}=-\frac{p}{3}=u_iv_i,$$

代入上述结果给出 $\rho^{2k}=1$. 配合 $\rho^3=1$ 可得 $k=0$.

展望高次方程 推而广之, 考虑形如

$$X^n+a_{n-1}X^{n-1}+\cdots+a_0=0$$

的 n 次方程, 其中 $a_0,a_1,\dots,a_{n-1}$ 是给定的系数. 有了 $n=1,2,3$ 时的经验, 自然的问题是对一般的 n 寻求公式解. 这里所谓的公式, 仅容许用到系数 $a_0,\dots,a_{n-1}$ 的四则运算 (当然, 分母非零) 和取 m 次根 $\sqrt[m]{\cdots}$ 的运算, 其中 $m\in\mathbb{Z}_{\geq1}$.

四次方程的公式解是 G. Cardano 及其学生 L. Ferrari 的工作, 同样见于 *Ars Magma*; $n\geq5$ 的情形则长期困扰着此后的数学家们, 直至 N. H. Abel, P. Ruffini 和 E. Galois 等人在 18 至 19 世纪之交的工作才彻底解答了这个问题: 五次及以上的方程无公式解. 其相关思想和技术成为近世代数学的滥觞.

完整解释 Galois 的结果需要较多的理论铺垫, 不属本书范围. 虽然这是一个否定性的结论, 历代学者对高次方程的研究并不是无用功, 关键在于深入剖析根的置换, 亦即根的重排. 关于置换的一般理论是后续章节的主题, 此外这还涉及对称多项式的理论. 在此之前, 我们不妨从三次的情形获取初步印象.

沿用先前关于三次方程 $X^3+pX+q=0$ 的符号, 记其三个复根为 x_1,x_2,x_3 (计入重数). 在公式 (1.1.4) 中, 我们通过一个六次辅助方程的解 u_1, u_2, u_3, v_1, v_2, v_3 来表达 x_1, x_2, x_3. 反过来, x_1, x_2, x_3 也能表达辅助方程的根: 适当地重排 u_1,u_2,u_3 后, 不妨假定

$$u_2=\rho u_1,\quad u_3=\rho^2u_1,$$

而因为 $u_iv_i=-\frac{p}{3}$, 相应地有

$$v_2=\rho^{-1}v_1=\rho^2v_1,\quad v_3=\rho^{-2}v_1=\rho v_1.$$

基于 $x_i = u_i + v_i$ 和 $1 + \rho + \rho^2 = 0$, 容易验证

$$\begin{aligned} u_1 &= (x_1 + \rho^2 x_2 + \rho x_3)/3, \\ u_2 &= (\rho x_1 + x_2 + \rho^2 x_3)/3, \\ u_3 &= (\rho^2 x_1 + \rho x_2 + x_3)/3, \\ v_1 &= (x_1 + \rho x_2 + \rho^2 x_3)/3, \\ v_2 &= (\rho^2 x_1 + x_2 + \rho x_3)/3, \\ v_3 &= (\rho x_1 + \rho^2 x_2 + x_3)/3. \end{aligned}$$

等式右边可以写作 $t(\rho, \tau) := \frac{1}{3}\sum_{k=1}^{3} \rho^{k-1} x_{\tau(k)}$, 其中 τ 遍历所有一对一映射

$$\tau : \{1, 2, 3\} \to \{1, 2, 3\};$$

换言之, 求和遍历 $(1, 2, 3)$ 的所有排列 $(\tau(1), \tau(2), \tau(3))$, 总共有 $3! = 6$ 种. 辅助方程 (1.1.1) 因而可以写作 $\prod_\tau (u - t(\rho, \tau)) = 0$.

现在我们将上述公式当作 $u_1, \ldots, v_3$ 的定义. 一旦能从辅助方程解出这六个数, 便能解原来的三次方程. 这是以下简单练习的内容.

练习 1.1.4 设 $X^3 + pX + q = (X - x_1)(X - x_2)(X - x_3)$. 以上述公式定义 $u_1, \ldots, v_3$. 说明 $x_i = u_i + v_i$ 对 $i = 1, 2, 3$ 成立.

提示〉 回忆到 $1+\rho+\rho^2 = 0$; 其次, 三次方程的 X^2 项系数为 0 导致 $x_1+x_2+x_3 = 0$.

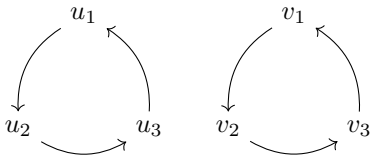
这些观察可以说是 Cardano 公式的实质. 推导的关键在于:

▷ **系数可表** 辅助方程的系数可以用 p 和 q 代数地表示, 只要辅助方程有公式解, 则 x_1, x_2, x_3 也随之有公式解;

▷ **方程可解** 辅助方程是 u^3 的二次方程, 因此它确实有公式解.

第一点既可以归结为冗长的计算, 也可以用以后将介绍的对称多项式理论来理解; 事实上, 这是我们取 $\prod_\tau (u - t(\rho, \tau))$ 的缘由.

第二点同样可以从根的置换来观照: 定义 $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ 为轮换, 由 $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$ 确定. 让轮换依下标重排 x_1, x_2, x_3. 则 $u_1, \ldots, v_3$ 在 σ 的作用下分为两个 "轨道":



周期性 $\rho^3 = 1$ 导致 $u_2 = \rho u_1$ 而 $u_3 = \rho u_2$; 基于恒等式

$$(X - Y)(X - \rho Y)(X - \rho^2 Y) = X^3 - Y^3,$$

我们推导出

$$(u-u_1)(u-u_2)(u-u_3)=(u-u_1)(u-\rho u_1)(u-\rho^2u_1)=u^3-u_1^3.$$

类似地, $v_2=\rho^2v_1$ 而 $v_3=\rho v_1$, 相同的论证导致

$$(u-v_1)(u-v_2)(u-v_3)=u^3-v_1^3.$$

这就为 Cardano 公式和辅助方程的取法提供了一个基于置换的解释, 本书在 §11.8 还会回到这个问题.

1.2 各种方程的求解

上一节从代数的历史引向解方程的问题, 进而讨论了一元多项式的求根方法. 解方程的问题不局限于一元多项式. 既可以考虑更一般的, 代数地定义的方程 (只涉及四则运算), 也可以对解的范围设限, 比如仅寻求非负实数解, 有理数解, 或整数解等.

求整数解的问题称为解**不定方程**. 虽然不定方程源远流长, 有最为初等的表述, 其求解或证明无解的过程却往往最为困难, 需要横跨数学各个领域的技术. 我们且来走马观花.

平方和问题 给定正整数 m, 试问 $X^2+Y^2=m$ 是否有整数解? 若有解, 能否确定解的个数?

在平面 $\mathbb{R}^2$ 上, $X^2+Y^2=m$ 描绘的无非是一个圆; 由于此处寻求的是整数解, 几何图像至多只能说明解的个数有限, 帮助极其有限. 代数工具则可以揭示更深层的结构: 定义复数集 $\mathbb{C}$ 的子集

$$\mathbb{Z}[\mathrm{i}]:=\{x+\mathrm{i}y\in\mathbb{C}:x,y\in\mathbb{Z}\}.$$

它包含 $\mathbb{Z}$, 并且对复数的乘法和加减法运算保持封闭, 换言之

$$x,y\in\mathbb{Z}[\mathrm{i}]\implies x\pm y,\ xy\in\mathbb{Z}[\mathrm{i}].$$

进一步, $\mathbb{Z}[\mathrm{i}]$ 对复共轭运算 $x+\mathrm{i}y\mapsto x-\mathrm{i}y$ 也保持封闭. 形如 $x+\mathrm{i}y$ 的复数 $(x,y\in\mathbb{Z})$ 也称为 Gauss 整数.

平方和问题依此改写为

$$z\overline{z}=m,\quad z=x+\mathrm{i}y\in\mathbb{Z}[\mathrm{i}].$$

由于共轭满足 $\overline{zz'}=\overline{z}\cdot\overline{z'}$, 从而 $z\overline{z}\cdot z'\overline{z'}=zz'\cdot\overline{zz'}$, 考虑 $\mathbb{Z}[\mathrm{i}]$ 的元素 $z=x+\mathrm{i}y$ 和 $z'=x'+\mathrm{i}y'$, 立见

$$x^2+y^2=m,\quad (x')^2+(y')^2=m'\\ \implies (xx'-yy')^2+(xy'+x'y)^2=mm'.$$

综上, $\mathbb{Z}[\mathrm{i}]$ 的乘性结构说明如何由平方和问题的既有解 “合成” 出新的解.

另一方面, 平方和问题对许多 m 是无解的. 为了演示一类典型例子, 以下运用整数的简单代数性质来说明

$$\text{当 } m \text{ 除以 } 4 \text{ 余 } 3 \text{ 时, 不定方程 } X^2+Y^2=m \text{ 无解.}$$

为此, 观察到对任何整数 m, 若 $m=2d$ 则 $m^2=4d^2$ 是 4 的倍数, 若 $m=2d+1$ 则 $m^2=4(d^2+d)+1$; 因此平方和 x^2+y^2 除以 4 的余数只能是 0 (若 x,y 皆偶), 1 (一奇一偶), 或 2 (皆奇). 关于余数的讨论是**同余**技巧的体现, 后续将有系统性的解说.

最后, 尚须确定的是 $X^2+Y^2=n$ 何时有解, 以及有几组解. 为此就有必要更详细地了解 $\mathbb{Z}[\mathrm{i}]$ 的代数性质, 或者略施奇技淫巧, 详见稍后习题.

勾股数 我们寻求 $X^2+Y^2=Z^2$ 的不全为 0 的整数解. 注意到 Z 不可能为 0. 等式两边同除以 Z^2 后, 问题等价于求解

$$X^2+Y^2=1, \quad X,Y\in\mathbb{Q}.$$

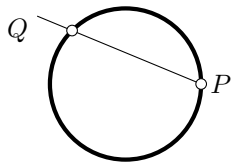
换言之, 我们寻求单位圆上的有理点. 与平方和问题不同, 几何直观在此可以起到帮助, 前提是要适当地融合几何与代数工具. 首先, 单位圆上有一个当然的有理点 $P=(1,0)$. 过 P 点的直线若不是过 P 的切线, 则总能表达成 $tX+Y=t$ 的形式, $t\in\mathbb{R}$ 是唯一确定的: $-t$ 无非是直线的斜率.

- ⋆ 若 $Q=(a,b)$ 是单位圆上的任一个有理点, $P\neq Q$, 作过 P 和 Q 的唯一直线 ℓ, 则对应的参数 t 是有理数 $\frac{-b}{a-1}$.
- ⋆ 反之, 考虑 $t\in\mathbb{Q}$ 和对应的直线 $\ell: tX+Y=t$, 则 ℓ 交单位圆于两点: 求交点相当于求解方程

  $$X^2+(t-tX)^2=1, \quad Y=t-tX.$$

  亦即解 $X^2-\frac{2t^2}{t^2+1}X+\frac{t^2-1}{t^2+1}=0$. 已知点 P 对应 $(X,Y)=(1,0)$, 故另一交点 Q 容易反解为

  $$(X,Y)=\left(\frac{t^2-1}{t^2+1},\frac{2t}{t^2+1}\right). \tag{1.2.1}$$



综上, 我们通过 $Q\leftrightarrow t$ 建立一一对应

$$\left\{ (a,b) \;\middle|\; \begin{array}{l} a^2+b^2=1,\ a,b\in\mathbb{Q} \\ (a,b)\neq(1,0) \end{array} \right\} \overset{1:1}{\longleftrightarrow} \mathbb{Q}.$$

对应的左侧欠缺美感, 因为它排除了 $P=(1,0)$. 解决方法倒也简单: 另外容许参数 t 为 ∞, 仅作为一个符号来理解, 而相应的直线 ℓ 是过 P 的切线; 合理地设想切线在 P 点交单位圆两次, 对应的解自然当是 $Q=P$ 了.

上述推导虽然和 Khayyam 关于三次方程的解法一样是基于几何, 结论 (1.2.1) 的精确性却无可比拟.

Fermat 方程 作为勾股数问题的自然延伸, 所谓的 Fermat 大定理断言不定方程

$$X^n+Y^n=Z^n,\quad n\geq 3$$

没有满足 $XYZ\neq 0$ 的整数解; 借由通分, 等价的说法是此方程没有满足 $XYZ\neq 0$ 的有理数解.

Fermat 大定理的完整证明是 R. Taylor 和 A. Wiles 在 1995 发表的工作. 几何观点在他们的工作中至关重要. 和之前的例子类似, 所谓几何并不是简单地描绘 $X^n+Y^n=Z^n$ 的图像, 因为实数解和有理数解的脾性完全不同. 我们需要的是一套能整合几何直观和代数技巧, 从而能处理不定方程的几何理论. 先前处理勾股数问题的技巧是代数与几何交融的最简单的例子, 所用的性质是圆锥曲线和直线一般而言交于两点, 但此法对于 3 次以上的曲线便不再适用.

练习 1.2.1 (D. Zagier) 设 p 为素数, p 除以 4 余 1. 按以下论证说明存在 $x,y\in\mathbb{Z}$ 使得 $x^2+y^2=p$: 定义有限集

$$S:=\left\{(x,y,z)\in\mathbb{Z}_{\geq 1}:x^2+4yz=p\right\}.$$

(i) 考虑映射 $f:S\to S$ 如下

$$f(x,y,z)=\begin{cases}(x+2z,z,y-x-z), & \text{若 } x<y-z,\\ (2y-x,y,x-y+z), & \text{若 } y-z<x<2y,\\ (x-2y,x-y+z,y), & \text{若 } x>2y.\end{cases}$$

验证此定义是合理的, 而且对所有 $(x,y,z)\in S$ 皆有 $f(f(x,y,z))=(x,y,z)$.

(ii) 说明 $f:S\to S$ 有唯一的不动点 (x,y,z), 亦即满足 $f(x,y,z)=(x,y,z)$ 的点. 由此说明 S 的元素个数是奇数.

(iii) 按 $g(x,y,z)=(x,z,y)$ 定义映射 $g:S\to S$, 说明 g 有不动点. 以此说明 $x^2+y^2=p$ 有整数解.

线性方程组 相对于不定方程和多元高次方程组, 求解线性方程组要简单得多. 线性方程意指一次方程. 线性方程组是形如

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \vdots \quad & \quad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m \end{aligned}$$

的方程组, 其中 a_{ij} 和 b_i 是给定的常数 $(1 \leq i \leq n, 1 \leq j \leq m)$, 而 $X_1, \ldots, X_n$ 是变元. 这样的方程可以在 $\mathbb{Q}$, $\mathbb{R}$ 或 $\mathbb{C}$ 上来考虑, 我们暂且不去明确所用的数系.

根本的问题是: 如何判定方程组是否有解? 若有解, 如何高效地求解? 所有解 $(X_1, \ldots, X_n)$ 构成的集合 (顺理成章地称为**解集**) 有怎样的结构?

对于 $n = m = 2$ 的情形, 解法想必是读者熟知的. 在二阶行列式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}$$

非零的前提下, 方程组有唯一解

$$X_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad X_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

依然设 $n = m$; 对于一般的 m 的情形, 我们也有称为 Cramer 法则的行列式解法, 详见 §5.7. 然而它涉及 $n+1$ 个 n 阶行列式, 非但需要行列式的一般理论, 所需的计算量也随 n 增长而暴增; Cramer 法则的用处在于理论层面, 不在计算层面. 我们行将介绍的 Gauss–Jordan 消元法则提供了一个简单快速的求解手段.

小结 现在进一步解释早先提出的问题: 何谓代数?

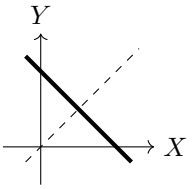
- ⋆ **何谓方程** — 来自经过有限次的加, 减, 乘, 除 (分母非零) 四则运算得到的表达式.
- ⋆ **何谓数** — 至少包括 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 这些常用的数系; 它们都具备四则运算, 但除法要求分母非零. 注意到 $\mathbb{Z}$ 不在列表中, 因为除法在 $\mathbb{Z}$ 中不能通行无阻.
- ⋆ **何谓求解的技艺** — 判定方程是否有解, 如何精确求解, 给出一套尽量高效的算法, 或者至少给出逼近方程的解的手段. 算法是一切应用的核心.

根据关于不定方程的讨论, 可知方程解集的性状和 "数" 的界定密切相关, 采用的技术也随之而千变万化.

我们关心的还有解集的结构. 数学中所谓的 “结构” 难以三言两语说清. 结构的一个重要面向是**对称性**. 在几何的经典场景中, 对称性意指图像在一族刚体变换作业下的不变性; 这里所谓的刚体变换包括平移, 旋转, 镜射, 后续章节将有完整讨论. 对称性也是美感经验的一大要素, 人类对于对称似乎有先天的敏感和爱好, 数学工作者尤其如此.

⋆ 方程 $X^2+Y^2=1$ 在平面 $\mathbb{R}^2$ 上的解集是单位圆, 对任何保持圆心不动的刚体变换都保持对称, 这些变换包括绕圆心的任意转动, 以及相对于 X 轴或 Y 轴的镜射等.

⋆ 考虑线性方程组 $X+Y=1$, 具体起见, 仍在平面 $\mathbb{R}^2$ 上求解. 其解集无非是直线:



直线上的点对于沿着 $(-1,1)$ 方向的所有平移都保持不变, 此外, 它相对于直线 $X=Y$ (上图虚线) 的镜射也具有对称性.

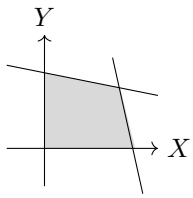
⋆ 除了基于图像的对称性, 还有针对抽象对象 (例如代数中使用的变元) 或其间的关系的对称性. 这种对称有时涉及特定集合或属性在对象的任意重排下, 或在按一定顺序轮换下的不变性, 但也可以涉及更广的变换. 这类对称性在三次方程的讨论中已经初露端倪.

上面给出了代数方程的几种典型例子, 那么何谓非代数方程? 典型的例子是涉及极限, 例如涉及无穷级数的操作, 这类问题是数学分析的主场. 另一类是涉及不等式, 例如形如

$$\begin{aligned} a_{11}X_1+\cdots+a_{1n}X_n &\le b_1\\ a_{21}X_1+\cdots+a_{2n}X_n &\le b_2\\ &\vdots \quad \vdots\\ a_{m1}X_1+\cdots+a_{mn}X_n &\le b_m \end{aligned}$$

的方程组; 这里取 a_{ij} 和 b_i 为实数, 所求的解 $X_1,\dots,X_n$ 亦然, 这是因为 $\mathbb{R}$ 相对于 $\mathbb{C}$ 具有额外的 “序结构”: 任两个实数可以合理地比大小.

涉及不等式的方程组也称为**半代数**的. 中学数学学过的线性规划便涉及这类方程组. 以 $n=2$ 情形为例, 解空间一般是由有限多条直线围出的区域, 譬如



的形式. 这类解空间也有特殊的 “结构”: 至少在有界的情形下, 它们是多边形 (高维情形: 多面体). 线性规划所求的是形如 $c_1X_1 + \cdots + c_nX_n$ 的函数在解空间上的极值. 当 $n = 2$ 时可以画图求解, $n = 3$ 时也可以发挥想象, 但实际应用场景中的 n 成千上万. 如何将低维的几何直觉可靠地推广到一般维数的线性规划? 尽管这是一个半代数的问题, 线性方法对此仍不可或缺.

若在线性规划的问题中另外要求 $X_1, \ldots, X_n$ 为整数, 对应的便是整数线性规划问题. 一如不定方程比求方程的复数解困难, 整数线性规划的难度也远高于线性规划. 按算法的术语说, 它是一个 NP 完全问题.

1.3 从线性方程组到 Gauss–Jordan 消元法

现在聚焦于最简单的一类方程, 即线性方程组. 具体地说, 考虑

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \vdots \quad & \quad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m \end{aligned} \tag{1.3.1}$$

除非另外声明, 对于变元个数 n 和方程个数 m 不加任何限制, 仅要求它们有限. 为了讨论方便, 且先假设是在 $\mathbb{C}$ 中求解.

线性方程组及其衍生结构将占据本书大半篇幅. 之所以选择它们作为学习代数的踏脚石, 大致可以点出几个原因:

- ⋆ 在所有代数方程组中, 线性方程组是充分广泛, 同时又相对简单的一类例子.
- ⋆ 实际应用中, 线性方程组常见而且重要.
- ⋆ 对于更复杂的方程和进阶的代数研究, 线性方法往往是必要的工具, 而线性方程组本身又自然地涉及高次多项式与抽象代数结构的思想. 两方面的技术因而是相互交融的.
- ⋆ 最后, 线性方程组的具体求解过程很能体现数学的算法特性.

既然拈出了算法作为一个必要, 有用而且可供操作的面向, 现在便来介绍称为**Gauss–Jordan 消元法**的求解手段. 大而化之地说, 消元法的思路是在同解的方程组之间过渡, 直至方程组化为一类可直接求解的形式为止. 何谓同解?

定义 1.3.1 如果以 $X_1,\dots,X_n$ 为变元的两个线性方程组有相同的解集, 则称它们是**同解**的.

且先看消元法的一个简单特例.

例 1.3.2 令 a,b,c 为给定的常数. 考虑

$$\begin{array}{rrrcll} X_1 & -X_2 & +X_3 & = & a & ① \\ X_1 & -X_2 & -X_3 & = & b & ② \\ 2X_1 & -2X_2 & -X_3 & = & c & ③ \end{array}$$

右列是方程的编号, 或谓 "行号". 将第一个方程两边乘以 -1 加到第二个方程; 类似地, 将第一个方程两边乘以 -2 加到第三个方程, 如是得到新的方程组

$$\begin{array}{rrrcll} X_1 & -X_2 & +X_3 & = & a & ① \\ & & -2X_3 & = & b-a & ②' := ② - ① \\ & & -3X_3 & = & c-2a & ③' := ③ - 2\cdot ① \end{array}$$

我们想反解 X_3, 所以将 $②'$ 乘以 $-\frac{1}{2}$, 然后用它消掉 $③'$ 的 X_3. 产物是

$$\begin{array}{rrrcll} X_1 & -X_2 & +X_3 & = & a & ① \\ & & X_3 & = & \frac{a-b}{2} & ②'' := -\frac{1}{2}②' \\ & & 0 & = & \frac{-a-3b+2c}{2} & ③'' := ③' + 3\cdot ②'' \end{array}$$

每一步都给出同解的方程组. 解集于是明朗了: 由下往上地解方程, 得到

- ⋆ 如果 $-a-3b+2c\neq 0$, 则方程无解, 因为 $③''$ 将是自相矛盾的;
- ⋆ 设 $-a-3b+2c=0$, 则 $②''$ 给出 X_3, 代入 ①, 得出方程的通解

$$\begin{aligned} X_1 &= a + X_2 - X_3 \\ &= \frac{a+b}{2} + X_2, \\ X_3 &= \frac{a-b}{2}. \end{aligned}$$

注意到 X_2 在通解中是**自由变元**, 它不受约束, 可以任取.

如果一切都取为实数, 将解集在三维空间中绘制, 则它或者是空集 (当 $-a-3b+2c\neq 0$), 或者是落在平面 $X_3=\frac{a-b}{2}$ 上的一条直线, 由坐标 X_2 参数化. 套用物理学的术语, 后一情形下可以说解集的**自由度**为 1, 因为它有一个可变参数.

这里的自由度只是一个权宜说词, 随着向量空间理论的渐次铺展, 对此将有更加精确的界定.

回到一般的线性方程组. 形如 (1.3.1) 的写法现在显得有些累赘了, 不如引进较为紧凑的符号.

定义 1.3.3 我们将 (1.3.1) 的方程组以称为**矩阵**的方式记为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

的形式. 去掉最右一列得到的矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称为该方程组的**系数矩阵**; 相对于此, 先前写下的带 $b_1, \ldots, b_m$ 的矩阵则称为**增广矩阵**.

矩阵行, 列的具体记法是

$$\begin{pmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{pmatrix} \begin{matrix} \\ \text{第 } i \text{ 行} \\ \\ \end{matrix}$$

第 j 列

每个矩阵元 a_{ij} 都等于 0 的矩阵称为**零矩阵**. 形如 a_{ii} 的矩阵元构成了 A 的**对角线**[1]. 为了排版考量, 有时也将 a_{ij} 写成 $a_{i,j}$. 本书称 m 行 n 列的矩阵为 $m \times n$ 矩阵.

由于方程组中的 $b_1, \ldots, b_m$ 角色毕竟不同于系数 a_{ij}, 有时在增广矩阵中区隔作:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

Gauss–Jordan 消元法的思路是用以下三种操作来简化方程组, 或者换句话说, 简化相应的矩阵.

[1] 更精确的术语是主对角线.

(A) 设 $1 \le i \ne k \le m$, 而 c 是任意常数. 将第 i 行乘以 c 的结果加到第 k 行, 其他的行保持不变:

$$\mathrm{A}(i,k,c):\quad \begin{array}{c} \\ i \\ \\ k \\ \\ \end{array}\!\!\begin{pmatrix} \vdots & \\ \cdots\ a_{ij}\ \cdots & \\ \vdots & \\ \cdots\ a_{kj}\ \cdots & \\ \vdots & \end{pmatrix} \xrightarrow{\cdot c} \begin{pmatrix} \vdots & \\ \cdots\ a_{ij}\ \cdots & \\ \vdots & \\ \cdots\ a_{kj}+ca_{ij}\ \cdots & \\ \vdots & \end{pmatrix}$$

(B) 设 $1 \le i < k \le m$. 交换第 i 行和第 k 行:

$$\mathrm{B}(i,k):\quad \begin{array}{c} \\ i \\ \\ k \\ \\ \end{array}\!\!\begin{pmatrix} \vdots & \\ \cdots\ a_{ij}\ \cdots & \\ \vdots & \\ \cdots\ a_{kj}\ \cdots & \\ \vdots & \end{pmatrix} \xrightarrow{\text{交换}} \begin{pmatrix} \vdots & \\ \cdots\ a_{kj}\ \cdots & \\ \vdots & \\ \cdots\ a_{ij}\ \cdots & \\ \vdots & \end{pmatrix}$$

(C) 设 $1 \le i \le m$ 而 c 是非零常数. 将第 i 行的每一项都乘以 c:

$$\mathrm{C}(i,c):\quad \begin{array}{c} \\ i \\ \\ \end{array}\!\!\begin{pmatrix} \vdots & \\ \cdots\ a_{ij}\ \cdots & \\ \vdots & \end{pmatrix} \curvearrowleft \cdot c \quad \Rightarrow \quad \begin{pmatrix} \vdots & \\ \cdots\ ca_{ij}\ \cdots & \\ \vdots & \end{pmatrix}$$

这些操作称为对矩阵的**初等行变换**. 我们仅容许交换行的顺序, 列的顺序不变, 所以变元 $X_1, \ldots, X_n$ 的顺序恒定. 如果 $(x_1, \ldots, x_n)$ 是原矩阵对应的方程组的解, 则相对于初等行变换之后的矩阵, 它仍是对应的方程组的解.

注意到上述每一种操作都可以被相应的逆操作撤销, 以回到原来的矩阵, 具体言之:

- ⋆ $\mathrm{A}(i,k,c)$ 的逆操作是 $\mathrm{A}(i,k,-c)$,
- ⋆ $\mathrm{B}(i,k)$ 的逆操作是 $\mathrm{B}(i,k)$ 自身,
- ⋆ $\mathrm{C}(i,c)$ 的逆操作是 $\mathrm{C}(i,1/c)$,

有请读者顺手检验. 综上:

矩阵的初等行变换给出同解的方程组.

也请注意到如果矩阵的某一列全为 0, 则无论如何作初等行变换, 该列依然为 0.

既然已表述了消元法涉及的初等行变换, 下一步自然是介绍消元法的目标, 称为行梯矩阵, 它们所对应的线性方程组易于求解.

定义 1.3.4 考虑 m 行 n 列的矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1.3.2}$$

当它的形式如下所示时, 称之为**行梯矩阵**:

其中左下空白部分的矩阵元皆为零, 涂灰的部分逐行向内严格缩进, 而且我们要求涂灰部分每一行的左端都是非零元, 它们称为此行梯矩阵的**主元**.

更加严格却不尽直观的定义如下:

- ⋆ 存在整数 $0 \leq r \leq m$ 使得第 i 行全为 0 当且仅当 $i > r$ (因此行梯矩阵中不全为 0 的行恰好是前 r 行);
- ⋆ 对于每个 $1 \leq k \leq r$ (非零行的编号), 取

$$j_k := \min\{j : a_{kj} \neq 0\},$$

则 $j_1 < j_2 < \cdots < j_r$ (相当于说涂灰部分逐行严格缩进).

现前提及的主元无非是 $a_{1,j_1}, \ldots, a_{r,j_r}$.

请读者沉思行梯矩阵的轮廓, 下述结果应该不言而喻.

练习 1.3.5 验证主元的个数 r 满足 $0 \leq r \leq \min\{n, m\}$. 无主元的行梯矩阵只能是零矩阵.

定义 1.3.6 在关于行梯矩阵的定义中, 倘若进一步对所有 $1 \leq k \leq r$ 要求:

- ⋆ $a_{k,j_k} = 1$, 换言之, 主元全为 1;
- ⋆ $i < k \implies a_{i,j_k} = 0$, 换言之, 落在主元以上的项全为 0;

则称此矩阵为**简化行梯矩阵**.

算法 1.3.7 (C. F. Gauss, W. Jordan) 对给定的矩阵如 (1.3.2), 按以下程序反复作初等行变换, 便能化之为行梯矩阵.

1. 如果矩阵的第一列全为 0, 则跳过第一列, 继续对下图框出的 “子矩阵” 作初等行变换:

$$\begin{pmatrix} 0 & \boxed{\begin{matrix} a_{12} & a_{13} & \cdots & a_{1n} \\ a_{22} & a_{23} & \cdots & a_{2n} \\ & \ddots & & \\ a_{m2} & a_{m3} & \cdots & a_{mn} \end{matrix}} \\ 0 & \\ \vdots & \\ 0 & \end{pmatrix}$$

   全为 0 的列不受后续变换的影响, 可以放心舍去.

2. 设矩阵第一列有非零元, 设之为 a_{k1}. 进行先前标为 $\mathrm{B}(k,1)$ 的变换以交换第 k 行和第 1 行, 可以化到 $k=1$, 亦即 $a_{11}\neq 0$ 的情形.

3. 设 $a_{11}\neq 0$. 接着对每个 $1<i\leq m$, 进行先前标为 $\mathrm{A}\left(1,i,-\frac{a_{i1}}{a_{11}}\right)$ 的变换, 将第一行乘以 $-\frac{a_{i1}}{a_{11}}$ 倍加到第 i 行. 如此的效果是将 a_{11} 以下的矩阵元全变为 0. 于是矩阵进一步化为

$$\begin{pmatrix} \boldsymbol{a}_{11} & a_{12} & \cdots & a_{1n} \\ 0 & \boxed{\begin{matrix} & & & \\ \cdots & a_{ij}-\frac{a_{1j}a_{i1}}{a_{11}} & \cdots & \\ & \vdots & & \end{matrix}} & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$$

   然后我们对框出的子矩阵继续操作.

之所以将 $\boldsymbol{a}_{11}$ 加黑, 是代表该矩阵元为主元. 而对于框出的子矩阵, 它或者处处是 0, 或者经过初等行变换还会给出新的主元. 依此类推, 算法必然在有限步内停止, 给出行梯矩阵.

目前仅用到 (A), (B) 两种类型的运算. 为了从行梯矩阵进一步得到简化行梯矩阵, 我们对给定的行梯矩阵继续以下操作, 这里需要 (C) 型操作.

4. 对每个主元 a_{k,j_k}, 作变换 $\mathrm{C}\left(k,a_{k,j_k}^{-1}\right)$ 以化约到行梯矩阵的主元全为 1 的情形.

5. 对每个主元 (取值为 1), 假设它位于第 k 行上, 对每个 $i<k$ 作变换 $\mathrm{A}\left(k,i,-a_{ij_k}\right)$; 换言之, 将第 k 行乘以 $-a_{i,j_k}$ 加到第 i 行上. 此操作将落在主元以上的矩阵元全化为 0.

显然, 最后得到的矩阵必然是简化行梯矩阵.

消元法的执行方式并不唯一. 比如步骤 2 涉及非零元 a_{k1} 的选法, 而且我们在选定非零元 a_{k1} 之前还可以进行若干次 (A) 或 (C) 型的初等行变换, 将各项化为更方便计算的形式. 当矩阵元全为整数时, 这种操作手法特别常见.

同一个矩阵可以通过初等行变换过渡到种种不同的行梯矩阵. 相对于此, 初等行变换给出的简化行梯矩阵则是唯一确定的. 本章习题部分将勾勒其证明.

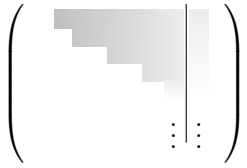
练习 1.3.8 假定读者对于编程有一定程度的了解. 令 $N := \max\{n, m\}$. 试说明随着 N 增大, Gauss–Jordan 消元法的涉及的操作次数的增长约略被 N^3 的某个常数倍控制.

1.4 关于线性方程组的总结

莫忘原初问题是解方程 (1.3.1). 我们对相应的增广矩阵

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

施行 Gauss–Jordan 消元法, 化之为简化行梯矩阵, 形如



相应的线性方程组与原方程组同解. 是故我们可以聚焦于简化行梯矩阵对应的线性方程组.

1. 如果简化行梯矩阵包含形如

   $$\left(\begin{array}{ccc|c} 0 & \cdots & 0 & 1 \end{array}\right)$$

   的行, 或者换句话说, 方程组包含等式 $0 = 1$, 则方程组无解.

2. 假设没有如上形式的行. 记简化行梯矩阵主元的个数为 r, 将主元出现的列号依次排开, 记为

   $$1 \leq j_1 < \cdots < j_r \leq n;$$

   称它们对应的列为**主列**. 则矩阵的第 k 行 $(1 \leq k \leq r)$ 对应到方程

   $$X_{j_k} + \sum_{j > j_k} a_{kj} X_j = b_k;$$

注意到简化行梯矩阵的定义确保 $X_{j_{k+1}},\dots,X_{j_r}$ 在左式中的系数为 0. 由此立可反解每个主列所对应的变元

$$X_{j_k} = b_k - \sum_{\substack{j>j_k\\ \text{非主列}}} a_{kj}X_j. \tag{1.4.1}$$

注意到方程组对非主列所对应的变元 X_j 没有约束 — 它们是 “自由变元”.

3. 因此 n 元线性方程组 (1.3.1) 或者无解, 或者它的解集依赖于 $n-r$ 个自由变化的参数 (亦即其 “自由度” 为 $n-r$), 其中 r 是 Gauss–Jordan 消元法给出的主元个数.

以上解方程 (1.3.1) 时对整个增广矩阵进行了消元. 包含 $b_1,\dots,b_m$ 的增广列当然是重要的, 它关系到方程组是否有解. 但只要方程组有解, 主元就不可能出现在增广列, 否则简化行梯矩阵将有形如 $(0\ \cdots\ 0 \mid 1)$ 的行. 这些讨论表明: 一旦方程组 (1.3.1) 有解, 则增广矩阵的主元无非是系数矩阵的主元.

尽管可以证明消元法给出的简化行梯矩阵是唯一的, 由于主元依赖于变元 $X_1,\dots,X_n$ 的排序, 主元相对于方程组本身仍显得是一个外部的, 不尽自然的概念. 另一方面, 上述讨论又表明一旦方程有解, 则主元个数 r 是一个内在于方程组本身的概念, 它连同变元个数 n 一并决定了解集的自由度 $n-r$. 为了理清这些问题, 有必要为线性方程组建立更深刻也更自然的理论框架. 这将涉及向量空间和线性变换的语言.

对于 Gauss–Jordan 消元法, 另一则重要观察是它只涉及矩阵元的四则运算, 和我们具体选取的数系 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 等并无关系. 然而若考虑系数在 $\mathbb{Z}$ 上的矩阵就会导致麻烦, 因为 Gauss–Jordan 消元法涉及除法. 这就启发我们进一步放宽数系的概念, 转向容许四则运算的抽象代数结构, 称为**域**. 由于代数结构的严谨定义离不开集合与映射, 首务是扼要地介绍集合论的相关概念, 这是下一章的主题.

习题

1. 说明 (或回忆) 如何用尺规作图来实现两个线段长度的相加和相减. 在给定单位长度线段的前提下, 用尺规作图实现线段长度的乘法和除法. 这些事实表明, 只要指定单位长度, 则所有能从尺规作图得到的线段长度对四则运算封闭, 因此它们也构成某种数系, 或者更精确地说, 构成一个 “域”.

请进一步说明 (或回忆) 如何用尺规作图构造平方根.

2. 设 $\zeta\in\mathbb{C}$ 是 5 次单位原根, 亦即 $\zeta^k=1 \iff 5\mid k$, 比如取 $\zeta=e^{2\pi\mathrm{i}/5}$ 即可. 设 $a,b\in\mathbb{Q}$, 定义

$$\alpha_i = \zeta^i\sqrt[5]{b+\sqrt{b^2-a^5}} + \zeta^{5-i}\sqrt[5]{b-\sqrt{b^2-a^5}}.$$

证明 $\alpha_0,\dots,\alpha_4$ 给出多项式

$$X^5-5aX^3+5a^2X-2b$$

的所有根, 记入重数.

提示〉 将题目中的 α_i 写作

$$\alpha_i=\zeta^i u+\zeta^{-i}v,$$

其中的复数 u,v 满足 $u^5v^5=a^5$ 而 $u^5+v^5=2b$. 用这些性质直接展开乘积来验证

$$\prod_{i=0}^{4}(X-\alpha_i)=X^5-5aX^3+5a^2X-2b.$$

因此 $\alpha_0,\dots,\alpha_4$ 确实给出原方程的根 (含重数). 这可谓是 Cardano 公式对 5 次方程的一种推广, 有兴趣的读者可参考 [5] 的讨论.

3. 以 Gauss–Jordan 算法解下列线性方程组.

(a)

$$\begin{aligned} X_1 &-3X_2 &-2X_3 &= 3\\ -2X_1 &+X_2 &-4X_3 &= -9\\ -X_1 &+4X_2 &-X_3 &= -7 \end{aligned}$$

(b)

$$\begin{aligned} X_1 &+3X_2 &+2X_3 &= 1\\ 2X_1 &+5X_2 &+5X_3 &= 7\\ 3X_1 &+7X_2 &+X_3 &= -8\\ -X_1 &-4X_2 &+X_3 &= 10 \end{aligned}$$

(c)

$$\begin{aligned} X_1 &-3X_2 &-2X_3 &-X_4 &= 6\\ 3X_1 &-8X_2 &+X_3 &+5X_4 &= 0\\ -2X_1 &+X_2 &-4X_3 &+X_4 &= -12\\ -X_1 &+4X_2 &-X_3 &-3X_4 &= 2 \end{aligned}$$

(d)

$$\begin{aligned} X_1 &+3X_2 &-7X_3 &= -8\\ 2X_1 &+5X_2 &+4X_3 &= 4\\ -3X_1 &-7X_2 &-2X_3 &= -3\\ X_1 &+4X_2 &-12X_3 &= -15 \end{aligned}$$

(e)

$$\begin{array}{rrrrcr} X_1 & -2X_2 & +3X_3 & -4X_4 & = & 4 \\ X_1 & +X_2 & -X_3 & +X_4 & = & -11 \\ X_1 & +3X_2 & & +X_4 & = & 1 \\ & -7X_2 & +3X_3 & +X_4 & = & -3 \end{array}$$

4. 确定关于复数 a, b 的条件, 使得以下方程组有解, 并具体将解用 a 和 b 来表达.

$$\begin{array}{rrrcr} aX_1 & +X_2 & +X_3 & = & 4 \\ X_1 & +bX_2 & +X_3 & = & 6 \\ X_1 & +2bX_2 & +X_3 & = & 9 \end{array}$$

5. 考虑两个大小相同的矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}.$$

如果可以通过一系列初等行变换从 $\boldsymbol{A}$ 过渡到 $\boldsymbol{B}$, 则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是**行等价**的.

(i) 取一列整数 $1 \leq c_1 < \cdots < c_h \leq n$. 从 $\boldsymbol{A}$ (或 $\boldsymbol{B}$) 删除第 $c_1, \ldots, c_h$ 列得到的矩阵记为 $\boldsymbol{A}'$ (或 $\boldsymbol{B}'$). 说明若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 行等价, 则 $\boldsymbol{A}'$ 和 $\boldsymbol{B}'$ 行等价.

(ii) 设所论矩阵形如

$$\boldsymbol{A} = \begin{pmatrix} 1 & & & x_1 \\ & \ddots & & \vdots \\ & & 1 & x_m \\ & & & \\ & & & \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & & & y_1 \\ & \ddots & & \vdots \\ & & 1 & y_m \\ & & & \\ & & & \end{pmatrix},$$

其中 $x_1, \ldots, x_m$ 和 $y_1, \ldots, y_m$ 是给定的数, 而矩阵中留白部分为 0. 说明若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 行等价, 则对所有 $1 \leq i \leq m$ 皆有 $x_i = y_i$.

提示〉 将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 视为 m 元线性方程组的增广矩阵, 解之.

(iii) 对于一般情形, 证明若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价的简化行梯矩阵, 则 $\boldsymbol{A} = \boldsymbol{B}$.

提示〉 设 $\boldsymbol{A} \neq \boldsymbol{B}$. 从左而右比较每一列, 设第一个相异的列为第 j 列. 从 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 删除所有 j 之后的列, 同时也删除第 j 列之前所有不含主元的列, 得到的矩阵分别记为 $\boldsymbol{A}'$ 和 $\boldsymbol{B}'$.

- ⋆ 论证第 j 列不可能包含主元.
- ⋆ 论证 $\boldsymbol{A}'$ 和 $\boldsymbol{B}'$ 必然是 (ii) 之中的形式. 配合 (i) 来推导 $\boldsymbol{A}' = \boldsymbol{B}'$, 从而导出矛盾.

6. (Leontief 投入–产出模型) 考虑一个理想化的经济体, 它有 n 个生产部门, 部门 i 只生产类型 i 的产品, 而生产过程投入的要素遵循固定的比例. 将这些产品统一以元计价. 设部门 i 产出价值 X_i 元的产品, 而部门 j 每生产 1 元的产品需要投入 a_{ij} 元的第 i 种产品, 其中 $1 \leq i, j \leq n$. 因此价值 X_i 元的类型 i 产品一部分供给其他部门 (称为中间产品需求), 剩下部分则供给消费者 (称为最终产品需求), 记后一部分的价值为 $d_i \in \mathbb{R}_{\geq 0}$. 列式得到

$$X_i = a_{i1}X_1 + \cdots + a_{in}X_n + d_i.$$

让 $1 \leq i \leq n$ 变动便得到 n 元线性方程组, 其中的系数 $a_{ij} \in \mathbb{R}_{\geq 0}$ 称为投入系数. 将这些系数作成 $n \times n$ 矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

说明之前的线性方程组能以增广矩阵的写法表作

$$\left(\begin{array}{c|c} \mathbf{1}_{n\times n} - \boldsymbol{A} & \boldsymbol{d} \end{array} \right),$$

其中 $\mathbf{1}_{n\times n}$ 是对角线上为 1, 其余位置全为 0 的 $n \times n$ 矩阵, 称为单位矩阵, $\mathbf{1}_{n\times n} - \boldsymbol{A}$ 意谓将两个矩阵逐项相减 (请写出它的大致样貌), 而 $\boldsymbol{d}$ 是只有一列的 $n \times 1$ 矩阵 (常称为 n 维列向量)

$$\boldsymbol{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}.$$

针对投入–产出的初步分析化为线性方程组或对应的矩阵的研究. 并非所有解都有现实意义; 比方说, 我们希望在所有 d_i 非负的前提下, 总存在使所有 X_i 非负的解. 为了确定是否恒有这种解[2)], 同时在 n 较大时高效地计算, 需要对矩阵有更加透彻的了解, 一句话, 需要更高段位的数学.

[2)] 常见的一种充要条件称为 Hawkins–Simon 条件, 又称 Kotelyanskiĭ 引理, 它要求 $\mathbf{1}_{n\times n} - \boldsymbol{A}$ 的某类子矩阵的行列式 (称为顺序主子式) 全为正. 矩阵的行列式是本书行将探讨的主题.

第二章 集合, 映射与关系

集合论是当代数学的底层语言. 数学中的一切陈述原则上都能够, 或者说应当能化约为集合的语言. 然而若仅凭对集合的朴素理解, 不加深思地操作, 便会面临种种矛盾, 其中最著名的当属 Russell 悖论 (注记 2.1.1). 为了避免矛盾, 也为了完整发挥集合论的效能, 公理化的表述实属必要. 公理集合论的进路不只一种, 本书介绍的是 Zermelo–Fraenkel 公理集合论 + 选择公理, 简称 ZFC.

本章的 §2.1 引入 ZFC 的各条公理, 我们尽量采取自然语言而非形式语言予以表述. 严守逻辑次序的理解方法也许符合一部分读者的禀性, 但公理集合论在数学实践中往往只是一个黑箱, 现阶段真正重要的毋宁说是:

1. 熟悉集合论的基本词汇, 包括集合与真类的区别, 以及相关符号;
2. 掌握集合论中的基本构造, 包括但不限于映射的运算, 一族集合的积与无交并, 以及集合上的序结构, 等价关系与商集的概念;
3. 了解无穷集之间如何比较大小, 亦即集合之间的等势或集合的基数概念.

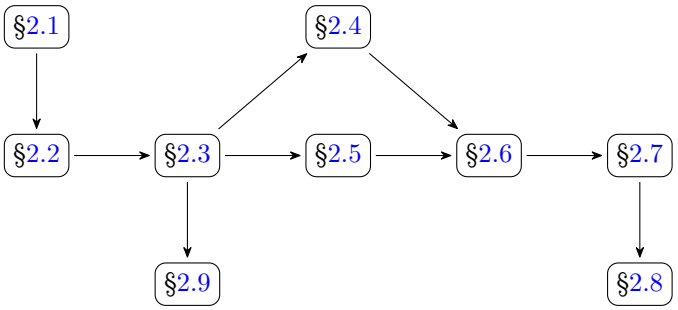
为了展示集合论的用法, 同时为后续章节打好基础, 本章后半部也将介绍

1. 如何以非负整数集 $\mathbb{Z}_{\geq 0}$ 为起点, 构造整数集 $\mathbb{Z}$ 连同其上的运算 (加, 减, 乘) 及大小关系;
2. 如何从 $\mathbb{Z}$ 构造有理数集 $\mathbb{Q}$ 连同其上的四则运算和大小关系;
3. 如何严谨地梳理整数集 $\mathbb{Z}$ 上的算术 (带余除法, 因式分解等);
4. 整数之间的同余关系和同余类, 它们分别是等价关系与商集的初步实例, 在数论中已有直接应用.

上述种种构造不但为初等数论奠定了严谨基础, 还能扩及更一般的代数结构, 这是本书后续部分将处理的课题.

本书附录包含关于非负整数集 $\mathbb{Z}_{\geq 0}$ 的公理化描述, 它在 ZFC 中的严格构造方法, 以及关于基数的若干基本性质, 供有需求的读者查阅.

阅读顺序



2.1 集合概论

近代数学的大厦以集合论为基石. 集合语言对于代数结构的研究尤其必要. 根据集合论创始人 G. Cantor 在 1895 年的界说, “集合意谓吾人感知或思想中一些确定的, 并且相互区别的对象汇集而成的整体, 这些对象称为该集合的元素.” 不妨就将集合 $S = \{a, b, \ldots\}$ 类比于收纳箱, 箱中包含不计顺序, 相互区别的元素 $a, b, \ldots$. 譬如:

$$\text{中国四大名园} = \big\{\text{承德避暑山庄}, \text{拙政园}, \text{颐和园}, \text{留园}\big\} \quad (\text{无排序}).$$

这一说法朴素而平易, 然而它缺乏数学标举的明晰性. 现代数学要求有清晰, 严格与形式化的基础. 一方面, 这是基于数学和逻辑学内部的考量, 因为滥造集合将导致矛盾; 另一方面它又影响应用, 因为若不采取谨守绳墨的形式语言, 就无法让计算机精确地处理. 职是之故, 我们希望初步地厘清以下问题:

1. 当我们谈论集合之时, 大致说着怎样的数学语言? 容许哪些符号和哪些句子?
2. 如何像平面几何学一般, 将集合论尽量精简地归结为几条基本公理?
3. 一旦回答了上述问题, 能否证明公理集合论在逻辑意义上是一致的, 换言之, 从中不会导出矛盾?

严格意义下的公理集合论是数学的一门专业领域. 作为基础, 它又是现代数学的黑箱; 在随后的内容中, 不得不打开集合论黑箱的场合是很稀少的. 虽是如此, 也不妨多说几句, 一是为了确定符号, 二是为了增广见闻. 由于本书不是集合论或数理逻辑的专著, 只能草草带过, 相关参考书籍包括但不限于 [13].

根据本书采纳的体系, 同时也是多数教材的共识, 集合论的所有对象都是集合; 特别地, 集合的元素仍是集合. 这好比说收纳箱里的物件也全是收纳箱, 或者说集合论不含 “原子”. 尽管略违直观, 如此建起的理论更加简洁, 而且已经足以构造正整数集, 从而构造数学中所须面对的一切集合[1)], 如 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 等.

[1)] 尽管无法囊括吾人感知或思想中的所有对象, 例如颐和园.

回顾 Cantor 对集合的原始界说, 从中至少能得到两点教益.

⋆ 所论的数学对象是确定的, 所以在集合论的语言中除了种种符号 $x, y, \ldots$ 以外, 还有一个相等符号 $=$. 它应该满足等号所需的一切性质; 当 x, y 给定, $x = y$ 或其否定 $x \neq y$ 恰有一者成立.

⋆ 集合由其全体元素确定. 所以集合语言需要一个符号 $\in$ 来指涉从属关系: $a \in A$ 意谓 a 是集合 A 的元素, 读作 "a 属于 A", 而 $a \notin A$ 则是其否定. 我们有时也用 $A \ni a$ 的写法.

进一步, 我们要求 $A = B$ 当且仅当 $\forall x,\ x \in A \iff x \in B$.

由此遂有派生的符号 $\subset$: 对任意集合 A 和 B, 表达式 $A \subset B$ 意谓对于任意 $x \in A$ 都有 $x \in B$; 这时称 A 是 B 的**子集**. 因此, 集合由其元素确定这一要求就体现为

▷ **外延公理** 对任意 A 和 B, 我们有

$$A = B \iff (A \subset B) \wedge (B \subset A).$$

当 $A \subset B$ 而 $A \neq B$ 时记之为 $A \subsetneq B$, 即严格包含, 此时也称 A 是 B 的真子集. 按自明的方式来解释 $A \supset B$ 和 $A \supsetneq B$.

表达一个集合的常见方法有两种, 一是枚举其元素, 二是列出其元素所需满足的条件. 两者都被提炼为集合论的公理.

▷ **配对公理** 对任意 a 和 b, 存在集合 $\{a, b\}$ 使得其元素恰好是 a 和 b. 基于外延公理, 这样的集合是唯一确定的. 特别地, 取 $a = b$ 便得到由 a 构成的**独点集** $\{a\}$.

留意到 $\{a, a\} = \{a\}$. 此外, $\{a, b\}$ 不区分 a 和 b 的次序, 但数学中常用的数组或元素组则是有序对 (a, b); 有序的涵义是 $(a, b) = (a', b')$ 当且仅当 $a = a'$ 且 $b = b'$. 尽管我们对有序对 (a, b) 已有清楚的认知, 但它们在集合论中的严格定义需要诉诸更基本的概念[2)], 具体办法是以配对公理命

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

这确实划定了 a 和 b 的顺序, 或者说是确定了 (a, b) 的第一个分量 a 和第二个分量 b.

▷ **分离公理模式** 对任意 A 和关于集合的性质 $\mathcal{P}$, 以 $\mathcal{P}(a)$ 代表集合 a 满足性质 $\mathcal{P}$, 则存在集合

$$\{a \in A : \mathcal{P}(a)\},$$

也写作 $\{a \in A \mid \mathcal{P}(a)\}$, 其元素正好是 A 中满足性质 $\mathcal{P}$ 的元素.

[2)] 以编程作类比, 这相当于要在一个只有集合的语言里实现数组 (或列表) 的数据结构. 然而许多编程语言中的情况正好相反.

分离公理模式之所以称为模式, 是因为它对每个性质 $\mathcal{P}$ 都产生一则公理. 作为简单应用, 我们定义任两个集合 A, B 的**差集**为

$$A \smallsetminus B := \{a \in A : a \notin B\}.$$

注记 2.1.1 (B. Russell) 分离公理模式是构造集合的主要媒介. 但在应用时需要将 a 局限在事先给定的集合 A 中. 这是为了避开下述的 **Russell 悖论**: 回忆到集合的元素仍是集合, 所以对任意集合 x 总能够问 $x \in x$ 是否成立. 考虑

$$\{\text{集合 } x : x \notin x\}.$$

倘若这确实构成集合, 记之为 R, 则无论 $R \in R$ 或 $R \notin R$ 都会导致矛盾. 分离公理模式不能用来取这般的 R: 全体集合并不构成集合.

如果非要为这般由一定的性质 $\mathcal{P}$ 确定的一族集合起名, 只能称之为**类**; 不成集合的类称为**真类**. 本节所介绍的公理体系不论真类, 仅视之为飘荡在非形式语言之中的一种说法: 当我们说 x 属于一个类 C 时, 实际是说 x 满足对应的性质 $\mathcal{P}$.

▷ **并集公理** 对任意集合 A, 存在集合 $\bigcup A$, 使得 $x \in \bigcup A$ 当且仅当存在 $a \in A$ 使得 $x \in a$.

方便起见, 今后这类陈述就简记为 $\bigcup A = \{x : \exists a \in A,\ x \in a\}$. 搭配先前的配对公理, 按此定义任两个集合 X 和 Y 的并为

$$X \cup Y := \bigcup \{X, Y\}.$$

对于给定的 $x, y, z, \ldots$ 可定义

$$\{x, y, z\} := \{x, y\} \cup \{z\}, \text{ 依此类推.}$$

▷ **幂集公理** 对任意集合 A, 它的所有子集也构成一个集合, 称为 A 的**幂集**, 记为

$$P(A) = \{B : B \subset A\},$$

有时也记之为 2^A.

▷ **无穷公理** 粗略地说: 存在无穷集.

无穷公理乍看比较费解, 它的功能在于保证我们能构造最简单的无穷集 $\mathbb{Z}_{\geq 0}$. 请感兴趣的读者参阅 §A.2.

▷ **替换公理模式** 设 A 为集合, 而 $\mathcal{F}$ 为一个定义在 A 上的函数[3], 映集合为集合 (也不妨设想为一族以 A 的元素为下标的集合 $X_a = \mathcal{F}(a)$), 则存在集合 $\mathcal{F}(A)$ 使得

[3] 此处的 $\mathcal{F}$ 并非 §2.2 将探讨的映射, 因为我们尚未指定一个集合作为它的值域; 值域集的存在性恰好是替换公理模式的目标. 较为严格的说法是将 $\mathcal{F}$ 理解为关于集合的某个性质 $\mathcal{P}$, 要求对所有 $a \in A$ 存在唯一的集合 b 使得有序对 (a, b) 满足 $\mathcal{P}$, 此 b 便是原表述中的 $\mathcal{F}(a)$.

对所有集合 b, 我们有

$$b \in \mathcal{F}(A) \iff \exists a \in A,\ b = \mathcal{F}(a);$$

集合 $\mathcal{F}(A)$ 经常简记为 $\{\mathcal{F}(a) : a \in A\}$.

定义 2.1.2 对于任何两个集合 A 和 B, 它们的 **Cartesius 积** $A \times B$ 也简称为**积**, 其元素是所有有序对 (a, b), 其中 $a \in A$ 而 $b \in B$. 换言之,

$$A \times B := \{(a, b) : a \in A,\ b \in B\}.$$

练习 2.1.3 说明 (a, b) 是 $P(P(A \cup B))$ 的元素. 因此 $A \times B$ 的定义是分离公理模式的应用: 其严格写法应是 $\{u \in P(P(A \cup B)) : \exists a \in A,\ \exists b \in B,\ u = (a, b)\}$.

推而广之, 有限个集合 $A, B, C, \ldots$ 的积便可以递归地定义为

$$A \times B \times C \times \cdots := \cdots((A \times B) \times C) \times \cdots,$$

其元素无非是所有有序元素组 $(a, b, c, \ldots) := (\cdots((a, b), c), \cdots)$. 特别地, 任意集合 A 都可以取 n 次幂:

$$A^n := \underbrace{A \times \cdots \times A}_{n\text{ 份}}, \quad n = 1, 2, 3, \ldots$$

譬如平面解析几何所用的坐标平面便是 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, 空间则是 $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

不含任何元素的集合称为**空集**, 记为 $\varnothing$. 一种严谨却诡异的构造是取道分离公理模式来定义

$$\varnothing := \{a \in A : a \neq a\},$$

其中的 A 是任意集合, 比如无穷公理所断言的某个无穷集. 空集是一切集合的子集, 而

$$P(\varnothing) = \{\varnothing\},$$

这是无中生有的非空集!

应用这些公理, 我们可以考虑一族集合 X_i, 其中 i 遍历某个下标集 I, 也看作定义在 I 上的函数, 对之定义**并**

$$\bigcup_{i \in I} X_i := \bigcup \{X_i : i \in I\}$$

和**交** (当 I 非空)

$$\bigcap_{i \in I} X_i := \{x : \forall i,\ x \in X_i\}.$$

替换公理模式确保 $\{X_i : i \in I\}$ 是集合, 故取并有意义; 至于交, 可任取 $i \in I$, 然后将交中的 x 限制在 X_i 里, 所以分离公理模式确保取交有意义.

有限多个集合 $X_1, X_2, \ldots$ 的并 (或交) 也记作 $X_1 \cup X_2 \cup \cdots$ (或 $X_1 \cap X_2 \cap \cdots$).

约定 2.1.4 对于下标集 $I=\varnothing$ 的极端情形, 相应的并规定为 $\varnothing$, 交则不予定义.

为何 $I=\varnothing$ 时不定义 $\bigcap_{i\in I}X_i$? 交的定义要求对于任意集合 x, 我们有

$$x\in\bigcap_{i\in I}X_i \iff \forall i\in I,\ x\in X_i,$$

右边在 $I=\varnothing$ 时成为空条件, 所以空交将给出所有集合 x 所构成的类, 这是真类.

类似于并集公理中的 $\bigcup A$, 一族集合之交又写作

$$\bigcap A:=\{x:\forall a\in A,\ x\in a\},\quad A:\text{ 任意非空集.}$$

练习 2.1.5 定义集合 X 和 Y 的**对称差**为 $X\Delta Y:=(X\smallsetminus Y)\cup(Y\smallsetminus X)$. 验证 $X\Delta Y=(X\cup Y)\smallsetminus(X\cap Y)$.

以上已经介绍了集合论最基本的几条公理. 后续两条公理相对复杂, 在此只能略述一二.

▷ **正则公理** 对任何非空集 A, 存在 $a\in A$ 使得对任意 $a'\in A$ 都有 $a'\notin a$; 等价的说法是 $a\cap A=\varnothing$.

▷ **选择公理** 设集合 A 的每个元素都是非空集, 此时存在函数 $g:A\to\bigcup A$ 使得对所有 $a\in A$ 都有 $g(a)\in a$.

练习 2.1.6 运用正则公理, 说明不存在一列集合 $x_1,x_2,\ldots$ (无穷延伸, 容许重复) 使得 $x_1\ni x_2\ni\cdots$. 特别地, 不存在满足 $x\in x$ 的集合 x. 〈提示〉 取 $A=\{x_1,x_2,\ldots\}$.

初等数学的绝大部分构造不需要正则公理, 它的主要意义在于集合论本身的研究, 读者可以略过. 至于选择公理, 其白话版本是说对于任何一族非空集 A, 总能从其中的每个集合 a 选出一个元素, 选法体现为公理中的选择函数 g. 尽管选法的存在性乍看是自明的, 但是对于一般的 A, 它会造成一些直观所难及的结论. 数学中许多存在性定理需要由选择公理来保证, 当前数学界的普遍共识是接受选择公理.

在假定其余公理满足一致性, 亦即它们不含矛盾的前提下, 无论正则公理或选择公理都独立于其余公理. 另一方面, 集合论的其余公理无法自证其一致性 (Gödel 第二不完全性定理). 为了说清何谓 “不含矛盾” 以及何谓 “独立”, 须进一步厘清**形式语言**与**证明**的实质, 这就超出本书范围, 而进入数理逻辑或计算机科学的基础面向了.

以上勾勒的体系称为 **Zermelo–Fraenkel 公理集合论**; 由于加入了选择公理, 对应的公理系统也简称为 ZFC, 这是当前的主流, 一些学者甚至声言: 所谓证明, 就是在 ZFC 之内作证明 [4]. 从纯粹形式化的观点, ZFC 是一套基于基本的逻辑符号, 并且以从属关系 $\in$ 为其原始概念的公理系统, 其中不仅不追问 “何谓集合” 这类问题, 也无法将 $\in$ 化约为更基本的概念. 由于 ZFC 确实有益于构筑一套简洁牢固的形式化基础, 本书依教奉行.

最后, 我们依然要强调: 对于初等代数的学习, 重点是如何讲好集合的语言, 而非钻研其文法.

2.2 映射的运算

我们先前已经约略谈过函数或映射的概念, 它可以大略地理解为一种 “法则”, 为定义域中的每个元素 (输入) 指派一个确定的值 (输出), 容许多对一, 不许一对多或一对无.

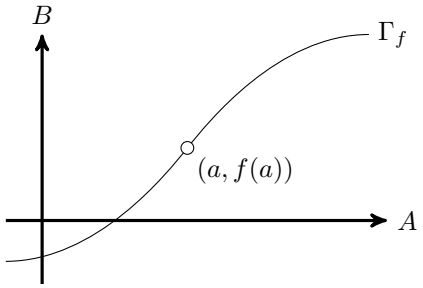
本书视 “映射” 与 “函数” 为同义词. 如前所述, 映射是一种对输入指派输出的法则. 如何以集合论的语言来描述这类法则? 更具体地说, 既然集合论探究的对象只有集合, 能否将映射定义为一种集合? 思路是熟悉的: 以函数图形来理解函数.

定义 2.2.1 设 A 和 B 为集合. 从 A 到 B 的映射写作 $f : A \to B$ 或 $A \xrightarrow{f} B$ 的形式. 用集合论的语言, 我们将映射 $f : A \to B$ 理解为 $A \times B$ 的一个子集, 记为 Γ_f, 它须满足以下条件: 对于每个 $a \in A$, 集合

$$\{b \in B : (a, b) \in \Gamma_f\}$$

是独点集, 其中的唯一元素记为 $f(a)$, 称为 a 在 f 下的像.

映射 $f : A \to B$ 所对应的 $\Gamma_f \subset A \times B$ 称为 f 的**图形**, 这是因为 $\Gamma_f = \{(a, f(a)) : a \in A\}$.



举例明之, 取 $A = B$, 则所谓对角子集 $\Delta_A = \{(a, a) : a \in A\}$ 显然满足定义中的条件; 以 Δ_A 为图形的映射称为 A 的**恒等映射**, 记为 $\mathrm{id}_A : A \to A$.

多变元映射可以用集合的积来料理, 不必另外定义, 比如将 $a_1 \in A_1$ 连同 $a_2 \in A_2$ 映至 $f(a_1, a_2) \in B$ 的二元函数便可以诠释为 $f : A_1 \times A_2 \to B$, 依此类推[4)].

定义容许 A 或 B 为空集的极端情形, 这是以下练习的内容.

练习 2.2.2 根据上述定义, 验证从空集 $\varnothing$ 到任何集合 B 都恰有一个映射. 对于任意非空集 A, 验证不存在从 A 到 $\varnothing$ 的映射.

提示〉 根据积的定义, 空集和任意集合的积仍是空集. 关于 Γ_f 的条件 (“对每个 a...”) 在 $A = \varnothing$ 时归于虚无.

[4)] 严格来说, 映射 $f : A_1 \times A_2 \to B$ 应当写作 $(a_1, a_2) \mapsto f((a_1, a_2))$ 而非 $(a_1, a_2) \mapsto f(a_1, a_2)$, 但这种执着毫无意义.

基于函数图形的观点, 映射 $f : A \to B$ 和 $g : B \to C$ 的合成 gf 对应到

$$\Gamma_{gf} = \{(a, c) \in A \times C : \exists b \in B,\ (a, b) \in \Gamma_f,\ (b, c) \in \Gamma_g\}.$$

映射的合成是一种抽象运算, 我们现在来探讨它的一些形式性质. 显然, 对于任意映射 $f : A \to B$ 都有

$$f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f.$$

此外, 映射的合成运算服从结合律: 对于任意映射

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$$

都有映射的等式

$$h(gf) = (hg)f : A \to D,$$

这是因为两边都映任意 $a \in A$ 为

$$h((gf)(a)) = h(g(f(a))) = (hg)(f(a)).$$

所以多个映射的合成可以省略括号, 写作 hgf 之类的形式.

回忆到双射 f 的逆映射映 $f(a) \in B$ 为 $a \in A$. 此概念有进一步的推广.

定义 2.2.3 *考虑一对映射 $A \underset{g}{\overset{f}{\rightleftarrows}} B$. 若 $gf = \mathrm{id}_A$, 则我们称 g 是 f 的***左逆***, 而 f 是 g 的***右逆***. 有左逆 (或右逆) 的映射称为左可逆 (或右可逆) 映射.*

练习 2.2.4 说明两个左可逆 (或右可逆) 映射的合成依然是左可逆 (或右可逆) 的.

左逆和右逆与映射在合成之下的消去律相关.

定义 2.2.5 *考虑映射 $f : A \to B$*

- ⋆ *若对于任意 C 和任一对映射 $g_1, g_2 : C \to A$ 都有*

$$fg_1 = fg_2 \implies g_1 = g_2,$$

则称 f 对映射合成具有左消去律.

- ⋆ *若对于任意 C 和任一对映射 $g_1, g_2 : B \to C$ 都有*

$$g_1 f = g_2 f \implies g_1 = g_2,$$

则称 f 对映射合成具有右消去律.

若 f 若有左逆 g, 则自动满足左消去律, 这是缘于

$$fg_1 = fg_2 \implies g(fg_1) = g(fg_2) \iff \underbrace{(gf)}_{=\mathrm{id}_A} g_1 = \underbrace{(gf)}_{=\mathrm{id}_A} g_2,$$

最后一个等式无非 $g_1 = g_2$. 完全类似地, f 若有右逆 g 则自动满足右消去律:

$$g_1 f = g_2 f \implies (g_1 f)g = (g_2 f)g \iff g_1 \underbrace{(fg)}_{=\mathrm{id}_B} = g_2 \underbrace{(fg)}_{=\mathrm{id}_B}.$$

命题 2.2.6 *对于映射 $f : A \to B$, 设 A 非空, 则以下性质等价: (i) f 是单射, (ii) f 有左逆, (iii) f 满足左消去律.*

类似地, 设 A 非空, 则以下性质等价: (i)′ f 是满射, (ii)′ f 有右逆, (iii)′ f 满足右消去律.

证明 这类论证有一种常见的模式. 以单射情形为例, 我们的目标是证 (i) $\implies$ (ii) $\implies$ (iii) $\implies$ (i). 设 (i) 成立. 任选 $a_0 \in A$, 然后定义 $g : B \to A$ 为

$$g(b) = \begin{cases} a, & \exists a \in A,\ b = f(a), \\ a_0, & b \notin \mathrm{im}(f); \end{cases}$$

由于 f 单, 定义确实是合理的. 由定义容易看出对所有 $a \in A$ 都有 $gf(a) = a$, 故 (ii) 成立.

先前已经说明 (ii) $\implies$ (iii), 接着设 (iii) 成立. 对于任意满足 $f(a_1) = f(a_2)$ 的 $a_1, a_2 \in A$, 对 $i \in \{1, 2\}$ 定义 $g_i : \{0\} \to A$ 使得 $g_i(0) = a_i$, 则 $fg_1 = fg_2 : \{0\} \to B$, 从而左消去律蕴涵 $g_1 = g_2$, 这也相当于说 $a_1 = a_2$. 因此 (iii) $\implies$ (i).

接着证明 (i)′ $\implies$ (ii)′ $\implies$ (iii)′ $\implies$ (i)′. 设 (i)′ 成立. 对每个 $b \in B$, 由于 f 满, 可取 $a \in A$ 使得 $f(a) = b$; 记如此选出的 a 为 $g(b)$, 然后让 b 变动, 则 $b \mapsto g(b)$ 确定的映射 $g : B \to A$ 满足 $fg = \mathrm{id}_B$, 是为右逆, 故 (ii)′ 成立. 留意到这里用到选择公理.

先前已经说明 (ii)′ $\implies$ (iii)′, 接着设 (iii)′ 成立. 定义 $g_1, g_2 : B \to \{0, 1\}$ 使得对所有 $b \in B$ 都有

$$g_1(b) = 0, \quad g_2(b) = \begin{cases} 0, & b \in \mathrm{im}(f), \\ 1, & b \notin \mathrm{im}(f). \end{cases}$$

于是 $g_1 f = g_2 f$ 自动成立; 右消去律蕴涵 $g_1 = g_2$, 由此可知 $\mathrm{im}(f) = B$, 故此时 (i)′ 成立. □

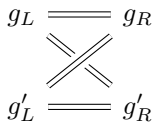
定义 2.2.7 *如果映射 f 左, 右皆可逆, 则称 f 是*可逆*映射. 此时存在唯一的 $f^{-1} : B \to A$ 使得 $f^{-1} \circ f = \mathrm{id}_A$ 而 $f \circ f^{-1} = \mathrm{id}_B$, 称之为 f 的逆.*

定义的后半段需要一些论证, 尽管这毫不困难. 首先设 f 可逆, g_L 为其左逆而 g_R 为其右逆. 映射合成的结合律导致

$$g_R = \mathrm{id}_A \circ g_R = (g_L \circ f) \circ g_R = g_L \circ (f \circ g_R) = g_L \circ \mathrm{id}_B = g_L.$$

这就说明左逆自动是右逆, 反之亦然. 如果 g_L 和 g_L' 都是 f 的左逆, g_R 和 g_R' 都是 f

的右逆, 则将左逆和右逆的四种组合代入上式, 可得



特别地, $g_L = g'_L$, $g_R = g'_R$.

综上, 在映射可逆的前提下, 左逆和右逆是一回事, 而且唯一, 可以合理地记为 f^{-1}.

为了强化印象, 我们再次重申 f 可逆的充要条件是存在反向映射 f^{-1} 使得 $f^{-1}f = \mathrm{id}_A$ 而 $ff^{-1} = \mathrm{id}_B$, 如此的 f^{-1} 若存在则唯一.

命题 2.2.8 *设映射 $f : A \to B$ 可逆, 则 $f^{-1} : B \to A$ 也可逆, 而且 $(f^{-1})^{-1} = f$.*

若映射 $f : A \to B$ 和 $g : B \to C$ 皆可逆, 则合成映射 $gf : A \to C$ 可逆, 而且 $(gf)^{-1} = f^{-1}g^{-1}$.

证明 第一个断言来自逆映射的刻画 $f^{-1}f = \mathrm{id}_A$ 和 $ff^{-1} = \mathrm{id}_B$: 将 f^{-1} 和 f 换位, 刻画不变.

对于第二个断言, 以合成的结合律直接验证 $(f^{-1}g^{-1})(gf) = f^{-1}(g^{-1}g)f = f^{-1}f = \mathrm{id}_A$ 和 $(gf)(f^{-1}g^{-1}) = g(ff^{-1})g^{-1} = gg^{-1} = \mathrm{id}_C$. □

在集合的世界中, 双射和可逆映射是一回事, 而双射的逆映射就是以上定义的 f^{-1}. 细说如下.

命题 2.2.9 *映射 f 是双射当且仅当 f 可逆, 此时它的逆映射正是之前定义的 f^{-1}.*

证明 设 f 是双射, 记 $f^{-1} : B \to A$ 为其逆映射, 映 $f(a)$ 为 a. 此时显然有 $f^{-1}f = \mathrm{id}_A$ 和 $ff^{-1} = \mathrm{id}_B$, 所以 f 可逆. 反之假定有 $f^{-1} : B \to A$ 满足 $f^{-1}f = \mathrm{id}_A$ 和 $ff^{-1} = \mathrm{id}_B$ 有逆映射, 则稍早的讨论说明 f 既单又满. □

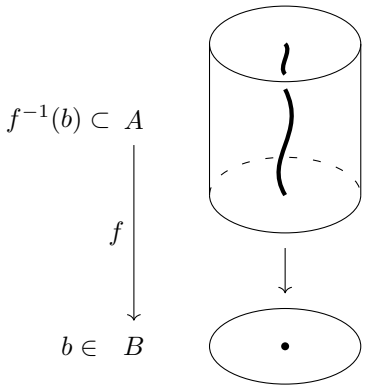
在上述定义和论证中, 我们尽量地避免指涉集合的元素和它们在映射下的像, 这是因为许多性质只关乎映射合成运算的形式性质. 这里隐约体现了一种现代意义的 "代数感", 它不再与解方程直接相关, 而是聚焦于定义在抽象对象上的运算 (例如映射的合成), 以及这些运算满足的规律 (例如合成的结合律, 左或右消去律).

回忆到符号 f^{-1} 在 $f : A \to B$ 非双射的情形也有意义: 对任意子集 $B' \subset B$, 符号 $f^{-1}(B')$ 代表 B' 相对于 B 的逆像, 这是 A 的一个子集.

定义 2.2.10 *对于映射 $f : A \to B$ 和 $b \in B$, 记*

$$f^{-1}(b) := f^{-1}(\{b\}) = \{a \in A : f(a) = b\}.$$

在和几何学相关的一些场景中, 大家偏好将 A 设想为竖在 B 上的空间, 而 f 类似于投影; 职是之故, $f^{-1}(b)$ 有时也称为 b 上的**纤维**, 可以按下图来想象:



相应地, A 分解为纤维的并

$$A = \bigcup_{b\in B} f^{-1}(b), \quad \text{两两无交.}$$

解方程可以看作是求映射纤维的问题. 回到 n 元线性方程组 (1.3.1) 具体地考察, 依然假设是在 $\mathbb{C}$ 上求解. 构造积集 $\mathbb{C}^n$ 和 $\mathbb{C}^m$, 并且定义映射

$$\begin{aligned} T : \mathbb{C}^n &\to \mathbb{C}^m \\ (x_i)_{i=1}^n &\mapsto \left(\sum_{i=1}^n a_{1i}x_i, \dots, \sum_{i=1}^n a_{mi}x_i \right). \end{aligned}$$

线性方程组的解集因而等于 $T^{-1}(b_1, \dots, b_m)$. 注意到 T 只依赖 (1.3.1) 的系数矩阵; 一旦系数矩阵给定, 求解方程组就相当于确定 T 的纤维.

当然, 问题不会因为抽象到这般程度就而自动解决. 在以上提纯过程中, 映射 T 的特殊性质未派上用场. 进一步的剖析将涉及向量空间的概念, 这是第四章的任务.

练习 2.2.11 考虑映射 $A \xrightarrow{f} B \xrightarrow{g} C$.

(i) 设 f 为满射, 说明 $\mathrm{im}(gf) = \mathrm{im}(g)$ 成立.

(ii) 设 g 为单射, 说明 $f^{-1}(b) = (gf)^{-1}(g(b))$ 对所有 $b \in B$ 成立.

练习 2.2.12 设 $f : A \to B$ 为映射, B_i (或 A_i) 为 B (或 A) 的一族子集, 其中下标 i 遍历某个非空集 I. 验证

$$\begin{aligned} f^{-1}(\bigcup_{i\in I} B_i) &= \bigcup_{i\in I} f^{-1}(B_i), \quad f^{-1}(\bigcap_{i\in I} B_i) = \bigcap_{i\in I} f^{-1}(B_i), \\ f(\bigcup_{i\in I} A_i) &= \bigcup_{i\in I} f(A_i), \quad f(\bigcap_{i\in I} A_i) \subset \bigcap_{i\in I} f(A_i). \end{aligned}$$

作为特例, 说明若 $b, b' \in B$ 相异, 则 $f^{-1}(b) \cap f^{-1}(b') = \varnothing$. 最后一式的 $\subset$ 是否有等号不成立的例子?

2.3 集合的积与无交并

我们已经见过如何对一列集合 $A_1,\dots,A_n$ 取积 $\prod_{i=1}^n A_i$, 或写作 $A_1\times\cdots\times A_n$, 其元素是 n 元组 $a=(a_1,\dots,a_n)$, 计顺序, 使得对每个 i 都有 $a_i\in A_i$. 这些元素也表作 $a=(a_i)_{i=1}^n$ 或 $(a_i)_i$ 的形式, 其中 a_i 称为 a 的第 i 个分量.

集合 $A_1\times(A_2\times A_3)$, $A_1\times A_2\times A_3$ 和 $(A_1\times A_2)\times A_3$ 之间有自然的一一对应

$$\begin{array}{ccccc} A_1\times(A_2\times A_3) & \xleftrightarrow{1:1} & A_1\times A_2\times A_3 & \xleftrightarrow{1:1} & (A_1\times A_2)\times A_3 \\ \cup & & \cup & & \cup \\ (a_1,(a_2,a_3)) & \longleftarrow\!\shortmid & (a_1,a_2,a_3) & \shortmid\!\longrightarrow & ((a_1,a_2),a_3). \end{array}$$

精确到这些双射, 集合的积满足结合律. 此外我们也有充当交换律的双射 $A\times B\xrightarrow{1:1}B\times A$, 方式当然是让 (a,b) 对应到 (b,a).

进一步, 对任意一族集合 $(A_i)_{i\in I}$ (容许重复) 也能定义积 $\prod_{i\in I}A_i$, 其中下标 i 遍历一个集合 I. 方法是利用映射的语言来定义

$$\prod_{i\in I}A_i:=\left\{\text{映射 } f:I\to\bigcup_{i\in I}A_i \;\middle|\; \forall i\in I,\ f(i)\in A_i\right\}. \tag{2.3.1}$$

这和先前的定义是兼容的: 设 $I=\{1,\dots,n\}$, 则 $f\in\prod_{i=1}^n A_i$ 对应到 $A_1\times\cdots\times A_n$ 的元素 $(f(1),\dots,f(n))$, 所做的不过是将下标转换成映射的变元. 今后不妨就将 $f(i)$ 记为 a_i, 于是 $\prod_{i\in I}A_i$ 的元素可以合理地记作

$$(a_i)_{i\in I}\in\prod_{i\in I}A_i.$$

对任意 $i\in I$ 都有由 $p_i\left((a_j)_{j\in I}\right)=a_i$ 确定的映射 $p_i:\prod_{j\in I}A_j\to A_i$, 称为第 i 个**投影映射**.

注记 2.3.1 由于指定 $\prod_{i\in I}A_i$ 的元素 $(a_i)_{i\in I}$ 相当于从每个 A_i 选出一个元素 a_i, 选择公理可以等价地表述为: 任意一族非空集的积仍然非空.

作为积的特例, 如果取每个 A_i 为同一个集合 A, 相应的积 $\prod_{i\in I}A$ 便化为从 I 到 A 的映射集, 以幂的符号记为

$$A^I:=\left\{\text{映射 } f:I\to A\right\}.$$

当 $I=\varnothing$ 时规定 $A^I:=\{\varnothing\}$, 理由见练习 2.2.2. 如果取 $n\in\mathbb{Z}_{\geq 1}$ 和 $I=\{0,\dots,n-1\}$, 则基于先前讨论不妨等同

$$A^{\{0,\dots,n-1\}}=\underbrace{A\times\cdots\times A}_{n\text{ 项}}\overset{\text{定义 2.1.2}}{=\!=}A^n;$$

因此也应当合理地定义 $A^0 := A^{\varnothing} = \{\varnothing\}$.

接着讨论并. 设集合 A 是一族子集 $(A_i)_{i\in I}$ 的并, 而且这些子集两两无交, 即

$$\forall i, j \in I,\ i \neq j \implies A_i \cap A_j = \varnothing,$$

这时称 A 为 $(A_i)_{i\in I}$ 的**无交并**, 或称 $(A_i)_{i\in I}$ 是 A 的一个**划分**, 写作

$$A = \bigsqcup_{i\in I} A_i.$$

可以说这是无交并的 "内在" 版本, 因为 A 已是一族子集的并, 而我们问这族子集在 A 之内是否无交.

相对于此, 另有 "外在" 的无交并: 任给一族集合 $(A_i)_{i\in I}$, 无论它们是否有交, 我们希望取适当的副本以将它们无交地取并, 由此构造一个更大的集合. 因为集合 A_i 之间可能有重叠, 我们替每一份 A_i 贴上一个标签 i 将其错开, 这一道工序是通过集合的积来实现的:

$$\bigsqcup_{i\in I} A_i := \left\{ (a, i) \in (\bigcup_{i\in I} A_i) \times I : a \in A_i \right\}.$$

对每个 $i \in I$, 我们有单射

$$\begin{aligned} \iota_i : A_i &\hookrightarrow \bigsqcup_{j\in I} A_j \\ a &\mapsto (a, i), \end{aligned}$$

称为第 i 个**嵌入映射**. 通过 ι_i 便能将 A_i 等同于元素与之一一对应的副本 $\mathrm{im}(\iota_i)$, 而 $\bigsqcup_{i\in I} A_i$ 是这些副本的 "内在" 版本无交并. 这就是为什么我们对内在和外在的无交并采用相同的符号; 由于罕见混淆, 今后也不再区分内在和外在版本.

对于有限多个集合 $A_1, \ldots, A_n$, 其无交并也记为 $A_1 \sqcup \cdots \sqcup A_n$; 这相当于 $I = \{1, \ldots, n\}$ 的特例.

约定 2.3.2 对于下标集 $I = \varnothing$ 的极端情形, 相应的无交并规定为 $\varnothing$, 而基于早先的讨论, 相应的积规定为独点集 $\{\varnothing\}$. 这些规定的合理性可以从所谓 "泛性质" 来说明, 建议读者在对代数有充分掌握之后参考例 B.5.5 和例 B.5.6.

仔细思考无交并的外在构造, 也许读者已察觉它直接依赖于副本的取法. 比方说, 为何不将 A_i 通过 $a \mapsto (a, i, i)$ 嵌入到 $(\bigcup_j A_j) \times I \times I$? 这种问题或许显得做作, 但是在关于结合律和交换律

$$A \sqcup (B \sqcup C) \overset{1:1}{\longleftrightarrow} (A \sqcup B) \sqcup C, \quad A \sqcup B \overset{1:1}{\longleftrightarrow} B \sqcup A$$

的思考中会碰到类似的, 尽管是容易处理的麻烦. 给定一族集合 $(A_i)_{i\in I}$, 命 $A := \bigsqcup_{i\in I} A_i$. 真正使无交并名副其实的不仅是集合 A 自身, 还涉及嵌入映射族 $(\iota_i : A_i \to A)_{i\in I}$ 的以下性质:

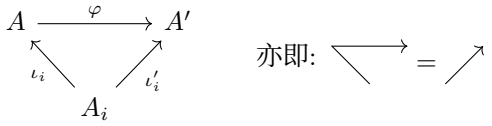
⋆ 每个 ι_i 都是单射,

⋆ A 是所有 $\mathrm{im}(\iota_i)$ 的并, 而且它们两两无交.

诚然, 如果另一个集合 A' 连同一族映射 $\iota'_i : A_i \hookrightarrow A'$ 也有上述性质, 则可以定义唯一的映射

$$\varphi : A \to A'$$

使得对所有 $i \in I$ 都有 $\varphi\iota_i = \iota'_i$; 或者用更清楚也更费纸张的方式来说, 存在唯一的 φ 使得对所有 $i \in I$, 下图的映射按两种方式合成是殊途同归的:

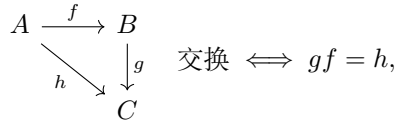


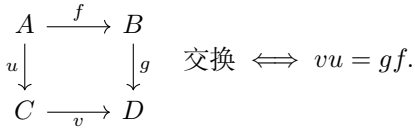
之所以如此, 是因为每个 $a \in A$ 都能写成 $\iota_i(a_i)$ 的形式, 其中 $i \in I$ 和 $a_i \in A_i$ 唯一确定, 故所求的 φ 能且仅能定义为 $\varphi(a) = \iota'_i(a_i)$. 由于 φ 的作用正是将每个 A_i 在 A 和 A' 中的副本一一对应起来, 它实际还是双射.

正是这些唯一的双射 φ 确保我们能无歧义地谈论无交并, 不必再管它实际上如何构造. 这是现代数学的一大特征, 以后还会面对各式各样的类似例子.

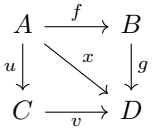
既然提到映射合成的图解, 我们顺势引进更广泛的定义.

约定 2.3.3 (交换图表) 将给定的一族集合标作图的节点, 映射标作箭头; 如果图表中的映射 (即箭头) 在合成运算下殊途同归, 则称之为交换图表. 从以下图例可以看得明白:





交换图表的一个特性是可拼贴. 举一简单情形为例, 若图表



的两个三角形部分交换 (亦即 $gf = x$, $vu = x$), 则整个外框交换 (亦即 $vu = gf$). 在处理较为复杂的合成操作时, 图解经常比文字占优. 以后会见到更多例子.

2.4 序结构

我们在初等数学中已经知道实数可以比大小; 在集合论中, 可以衡量两个集合之间有无包含关系 $\subset$; 在初等数论中, 整除关系也是一种用来比较两个正整数的尺度. 凡此种种, 都是偏序关系的例子. 何谓 "关系"? 这点容易用集合语言来解释.

定义 2.4.1 集合 A 和 B 之间的**二元关系**意指 $A\times B$ 的任意子集. 设 $R\subset A\times B$ 为二元关系, 则对于所有 $a\in A$ 和 $b\in B$, 我们以符号

$$aRb \quad \text{代表} \quad (a,b)\in R.$$

简便起见, 当 $A=B$ 时, 我们也称此为 A 上的二元关系.

举例明之, 等号 $=$ 是二元关系, 相应的子集 Δ_A 是 $A\times A$ 的**对角子集**:

$$\Delta_A=\{(a,a):a\in A\}. \tag{2.4.1}$$

尽管二元关系是一个近乎空泛的定义, 它却能将常识中模糊的关系概念转化为可操作, 可琢磨的数学对象.

举例明之, 映射 $f:A\to B$ 可以理解为 A 和 B 之间的一种二元关系, 方式是将映射视同其图形 $\Gamma_f=\{(a,f(a)):a\in A\}\subset A\times B$; 反之, 任意关系 $R\subset A\times B$ 对应到映射 $f:A\to B$ 当且仅当对于每个 $a\in A$, 存在唯一的 $b\in B$ 使得 aRb; 此时 $b=f(a)$. 这正是函数在集合论中的严格定义.

本节真正关心的是集合 A 上的一类特殊关系, 称为序, 它是种种顺序或大小关系的提纯, 在此习惯以符号 $\preceq$ 表示.

定义 2.4.2 设 $\preceq$ 是集合 A 上的二元关系. 当以下性质成立时, 称 $\preceq$ 为 A 上的预序, 相应地称资料 $(A,\preceq)$ 为**预序集**:

▷ **反身性** 对任意 $a\in A$ 都有 $a\preceq a$;

▷ **传递性** 对任意 $a,b,c\in A$, 若 $a\preceq b$ 而 $b\preceq c$, 则 $a\preceq c$.

如果它还满足下述条件, 则称 $\preceq$ 为 A 上的偏序, 相应地称 $(A,\preceq)$ 为**偏序集**:

▷ **反称性** 对任意 $a,b\in A$, 若 $a\preceq b$ 而且 $b\preceq a$, 则必有 $a=b$.

若偏序集 $(A,\preceq)$ 的任两个元素 a,b 皆可比大小, 换言之 $a\preceq b$ 或 $b\preceq a$ 至少有一者成立, 则称 $(A,\preceq)$ 为**全序集**或**链**.

在任意预序集 $(A,\preceq)$ 中, 习惯以符号 $a\prec b$ 代表 $a\preceq b$ 而 $a\neq b$. 不致混淆时, 我们也经常以 A 来代称预序集 $(A,\preceq)$.

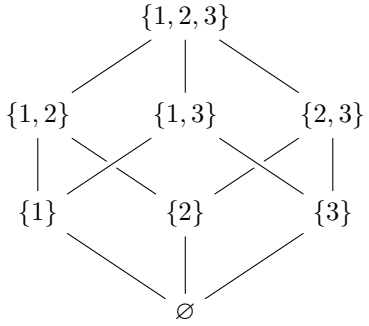
如果 $(A,\preceq)$ 是预序集 (或偏序集, 全序集), 而 $A'\subset A$ 是子集, 则相同的二元关系使得 $(A',\preceq)$ 也成为预序集 (或偏序集, 全序集).

- ⋆ 举例明之, $(\mathbb{R}, \leq)$ 或其子集 $(\mathbb{Z}, \leq)$ 等都是全序集. 如果考虑 $\mathbb{Z}_{\geq 1}$ 上的整除关系 $a|b \iff \frac{b}{a} \in \mathbb{Z}$, 则 $(\mathbb{Z}_{\geq 1}, |)$ 是偏序集而非全序集. 如果考虑 $\mathbb{Z} \smallsetminus \{0\}$ 而非 $\mathbb{Z}_{\geq 1}$ 上的整除关系, 这便仅是预序而非偏序了, 这是因为任意非零整数 x 和 y 相互整除当且仅当 $x = \pm y$.
- ⋆ 另一类例子是集合的包含关系. 赋予集合 S 的幂集 $P(S)$ 二元关系 $\subset$. 请读者迅速检验 $(P(S), \subset)$ 具有反身性, 传递性和反称性, 因此这给出偏序. 然而只要 S 有超过两个元素, S 便有互不包含的子集, 此时 $(P(S), \subset)$ 并非全序集.

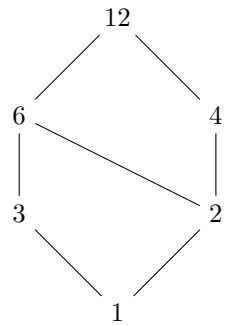
以下术语顺理成章.

定义 2.4.3 设 $f : A \to B$ 为预序集之间的映射. 如果 $a \preceq a' \implies f(a) \preceq f(a')$, 则称 f 是**保序**的. 如果 $a \preceq a' \iff f(a) \preceq f(a')$, 则称 f 是**严格保序**的.

对于有限偏序集, 一种特别方便的图解方式是将偏序集的元素标为顶点, 而如果 $a \prec b$ 而且不存在 c 使得 $a \prec c \prec b$, 则将 b 从上而下地通过边连接到 a. 这叫作 **Hasse 图**. 从下图的例子可以看得明白.



偏序集 $(P(\{1,2,3\}), \subset)$



偏序集 (12 的正因数, 整除性)

定义 2.4.4 设 $(A, \preceq)$ 为偏序集.

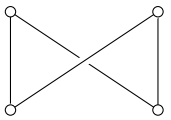
- ⋆ 满足以下性质的 $a_{\max} \in A$ 称为 A 的**极大元**: 不存在 $a \in A$ 使得 $a \succ a_{\max}$.
- ⋆ 满足以下性质的 $a_{\min} \in A$ 称为 A 的**极小元**: 不存在 $a \in A$ 使得 $a \prec a_{\min}$.

进一步设 A' 为 A 的子集.

- ⋆ 满足以下性质的 $a \in A$ 称为 A' 在 A 中的**上界**: 对所有 $a' \in A'$ 皆有 $a' \preceq a$.
- ⋆ 满足以下性质的 $a \in A$ 称为 A' 在 A 中的**下界**: 对所有 $a' \in A'$ 皆有 $a' \succeq a$.

练习 2.4.5 设 $(A, \preceq)$ 为全序集. 说明 A 的极大元 (或极小元) 和 A 在自身中的上界 (下界) 是一回事, 而且它们若存在则是唯一的.

就 Hasse 图来想象, 极大元可谓是其 “塔尖”, 极小元则相反. 在先前给出的两则图例中, 极大元 (或极小元) 存在, 唯一而且同时是整个集合的上界 (或下界); 一般情形下它们即便存在也不唯一, 而且未必等于整个集合的上界 (或下界). 比方说, Hasse 图完全可以形如下图.



请注意: 当谈论极大/极小元和上界/下界的存在性, 唯一性或其个数之时, 并不需要知道偏序集的元素具体是什么, 一切只依赖于元素之间抽象的序关系, 或者说只依赖 Hasse 图所呈现的结构.

练习 2.4.6 用 Hasse 图来分类所有至多有 4 个元素的偏序集, 精确到严格保序双射.

提示〉 共有 24 种 Hasse 图, 其中 3 个元素的有 5 种, 4 个元素的有 16 种.

定义 2.4.7 *若全序集 $(A,\preceq)$ 的每个非空子集 S 都有极小元, 则称 A 为*__良序集__.

不难看出, 良序集的子集仍是良序集.

良序集在集合论中占有特殊的地位. 根据我们纯净的直观, $\mathbb{Z}_{\geq 0}$ 是良序集这一事实似乎是不言而喻的, 它也可以从关于 $\mathbb{Z}_{\geq 0}$ 的 Peano 公理来推导, 感兴趣的读者可参阅命题 A.1.7.

现在将镜头拉远. 预序集是现代数学中所谓 “结构” 的一个例子, 因为我们不只看一个集合 A, 还考虑其上的二元关系 $\preceq$. 如果 $f: A \to B$ 是预序集之间的严格保序双射, 则称 f 为序结构之间的同构. 只要预序集 A 和 B 同构, 则一切能用序结构表达的性质对 A 和对 B 都是等价的, 它们可以通过同构在 A 和 B 之间相互翻译. 这是同构概念在本书中的第一个例子.

练习 2.4.8 设 $(A,\preceq)$ 是预序集 (或偏序集, 全序集), 而 $A' \subset A$ 是任意子集; 记包含映射为 $\iota: A' \hookrightarrow A$ (换言之, $\iota(a') = a' \in A$). 说明二元关系 $\preceq$ 可以自然地限制到 A' 上, 使得 $(A',\preceq)$ 成为预序集 (或偏序集, 全序集), 而且这是 A' 上使得 ι 严格保序的唯一预序. 我们也说 A' 从 A 继承了相应的序结构.

练习 2.4.9 设 $(A,\preceq)$ 为偏序集, A' 为 A 的子集.

- ⋆ 若 $a \in A$ 是 A' 在 A 中的上界, 而且对任何其他上界 $a_1 \in A$ 都有 $a \preceq a_1$, 则称 a 为 A' 的__上确界__.
- ⋆ 若 $a \in A$ 是 A' 在 A 中的下界, 而且对任何其他下界 $a_1 \in A$ 都有 $a \succeq a_1$, 则称 a 为 A' 的__下确界__.

说明 A' 的上确界 (或下确界) 若存在则是唯一的, 可以记为 $\sup A'$ (或 $\inf A'$). 试将数学分析中的上确界和下确界理解为一则特例.

2.5 等价关系与商集

和偏序针锋相对的另一种二元关系是等价关系, 常用符号是 $\sim$; 它指明一个集合 A 有哪些元素可以在一定意义下视作是相等的. 细说如下.

定义 2.5.1 集合 A 上的二元关系 $\sim$ 若满足以下性质, 则称为 A 上的**等价关系**.

▷ **反身性** 对所有 $a \in A$ 都有 $a \sim a$.

▷ **对称性** 设 $a, b \in A$, 则 $a \sim b$ 蕴涵 $b \sim a$;

▷ **传递性** 设 $a, b, c \in A$, 若 $a \sim b$ 而 $b \sim c$, 则 $a \sim c$.

既然等价关系的想法是为了将满足 $a \sim b$ 的元素 $a, b \in A$ 等量齐观, 我们自然会想将 A 的元素按照 $\sim$ 来划分. 这便导向等价类的概念.

定义 2.5.2 设 $\sim$ 是集合 A 上的等价关系. 若非空子集 $C \subset A$ 满足以下条件, 则称为 A 中的一个**等价类**:

⋆ C 中的元素相互等价, 亦即对任意 $x, y \in C$ 都有 $x \sim y$;

⋆ C 对 $\sim$ 封闭, 亦即对任意 $x \in C$ 和 $y \in A$, 我们有 $x \sim y$ 蕴涵 $y \in C$.

如果 C 是等价类而 $a \in C$, 则称 a 是 C 的一个**代表元**.

命题 2.5.3 设 $\sim$ 集合 A 上的等价关系, 则 A 是其中所有等价类的无交并.

证明 根据反身性, 任何元素 $a \in A$ 都属于

$$C_a := \{x \in A : x \sim a\}.$$

我们来验证 C_a 是一个等价类. 首先, 对称性和传递性蕴涵

$$x, y \in C_a \implies x \sim a \sim y \implies x \sim y,$$

同理有

$$(x \in C_a) \wedge (y \sim x) \implies y \sim x \sim a \implies y \in C_a,$$

所以 C_a 确实是含 a 的等价类, 从而 A 是等价类的并.

下一步是说明任两个等价类 $C_1, C_2 \subset A$ 或者无交, 或者相等. 设存在 $x \in C_1 \cap C_2$. 对任意 $x_2 \in C_2$, 从 $x_2 \sim x \in C_1$ 和 C_1 对 $\sim$ 的封闭性可见 $x_2 \in C_1$, 所以 $C_2 \subset C_1$. 由于场景对于 C_1 和 C_2 是对称的, 自然也有 $C_1 \subset C_2$, 故 $C_1 = C_2$. □

最细的等价关系当然是等号 $=$. 再看几个稍进一步的例子.

⋆ 给定任意映射 $f : A \to B$, 定义 A 上的二元关系 $\sim_f$ 使得 $a \sim_f a'$ 当且仅当 $f(a) = f(a')$. 请读者验证这是等价关系, 对应的等价类无非是 f 的纤维 $f^{-1}(b)$, 其中 $b \in \mathrm{im}(f)$.

⋆ 考虑所有形如 (1.3.1) 的线性方程组构成的集合, 它也可以等同于所有 m 行 $n+1$ 列矩阵构成的集合, 其中 n, m 固定. 以同解 (定义 1.3.1) 来定义这些方程组间的等价关系. 以 Gauss–Jordan 消元法解 (1.3.1) 的实质便是通过初等行变换在给定的同解等价类内移动, 直至抵达一类特别容易求解的代表元, 亦即对应到简化行梯矩阵的方程组.

等价关系的另一类重要例子是同余, 留待 §2.7 介绍.

不妨设想等价类中的元素 “精确到 $\sim$” 可视为相同. 商集的概念对此赋予了一个精确的数学意涵.

定义 2.5.4 设 $\sim$ 是非空集 A 上的等价关系, 相应的**商集**定义为幂集 $P(A)$ 的下述子集

$$A/\sim \ := \{C \subset A : \text{相对于 } \sim \text{ 的等价类}\}.$$

商集带有**商映射** $q : A \to A/\sim$, 映 $a \in A$ 为含 a 的唯一等价类.

由于等价类总有代表元, 商映射必然满. 商集连同商映射的内涵由下述性质阐明.

命题 2.5.5 设 $\sim$ 为集合 A 上的等价关系, $q : A \to A/\sim$ 为对应的商映射. 设映射 $f : A \to B$ 满足 $a \sim a' \implies f(a) = f(a')$, 则存在唯一的映射 $\bar{f} : (A/\sim) \to B$ 使得 $\bar{f} \circ q = f$.

证明 首先说明满足 $\bar{f} \circ q = f$ 的映射 $\bar{f}$ 至多仅一个. 这是因为 $A/\sim$ 的所有元素都能写作 $q(a)$ 的形式, 而 $\bar{f}(q(a)) = f(a)$. 另一种观点则是用满射 q 的右消去律推导 $\bar{f}_1 \circ q = \bar{f}_2 \circ q \implies \bar{f}_1 = \bar{f}_2$.

接着探讨 $\bar{f}$ 的存在性. 根据条件, 若 $q(a) = q(a')$ 则 $f(a) = f(a')$, 所以 $\bar{f} : q(a) \mapsto f(a)$ 确实定义了从 $A/\sim$ 到 B 的映射: 它仅依赖等价类 $q(a) \in A/\sim$ 而非 $a \in A$ 的选取. □

按照数学术语, 我们也说 $\bar{f} : q(a) \mapsto f(a)$ 是良定义的映射, 因为 $\bar{f}(q(a))$ 只依赖 $q(a)$, 不依赖辅助资料 a 的选取.

练习 2.5.6 固定 $n, m \in \mathbb{Z}_{\geq 1}$. 考虑所有形如

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

的矩阵构成的集合 $\mathrm{M}_{m\times n}$, 其中要求 a_{ij} 属于 $\mathbb{C}$ (或者属于任何一个选定的域, 这不影响论证). 证明第一章 习题中介绍的行等价是 $\mathrm{M}_{m\times n}$ 上的等价关系; 在 Gauss–Jordan

消元法的讨论中已经默认了这一则事实. 进一步用该节习题中的结果说明每个行等价类中存在唯一的简化行梯矩阵.

一些自然的数学对象往往能表为商集, 下面是一则简单例子.

例 2.5.7 考虑空间 $\mathbb{R}^3$. 在扣掉原点 $(0,0,0)$ 得到的子集 $\mathbb{R}^3 \smallsetminus \{(0,0,0)\}$ 上定义二元关系

$$(x,y,z) \sim (x',y',z') \iff \exists t \in \mathbb{R},\ t \neq 0,\ (x',y',z') = (tx,ty,tz).$$

请读者验证它是等价关系. 对应的商集有直观的描述:

$$\begin{array}{ccc} \left(\mathbb{R}^3 \smallsetminus \{(0,0,0)\}\right) / \sim & \xrightarrow{1:1} & \left\{\ell \subset \mathbb{R}^3 : \text{过原点的直线}\right\} \\ \cup & & \cup \\ (x,y,z) \text{ 的等价类} & \longmapsto & \text{过 } (x,y,z) \text{ 和原点的唯一直线}. \end{array}$$

如果在 $\sim$ 的定义中以 $t > 0$ 代替 $t \neq 0$, 则对应的商集和过原点的射线一一对应, 或者说和单位球面上的点一一对应, 方式是取射线和球面之交.

命题 2.5.8 对于任意映射 $f : A \to B$, 在 A 上定义等价关系 $\sim_f$ 使得 $a \sim_f a' \iff f(a) = f(a')$, 则命题 2.5.5 给出双射 $\bar{f} : (A/\sim_f) \xrightarrow{1:1} \mathrm{im}(f)$.

证明 命题 2.5.5 的前提对 f 和 $\sim_f$ 显然成立, 由此便得到 $\bar{f} : (A/\sim_f) \to B$ 使得 $\bar{f} \circ q = f$. 由于 q 满, 故 $\mathrm{im}(\bar{f}) = \mathrm{im}(f)$. 另一方面, 若 $\bar{f}(q(a)) = \bar{f}(q(a'))$ 则 $f(a) = f(a')$, 从而 $a \sim_f a'$; 这说明 $\bar{f}$ 也是单的. 证毕. □

以上命题虽然简单, 内涵却值得一提. 它说明尽管映射 f 的像集 $\mathrm{im}(f) \subset B$ 外在于 A, 却可以通过等价关系从 A 内在地构造. 在往后探讨环或向量空间等构造时, 这一主题还会反复回响.

2.6 从正整数集到有理数集

正整数集 $\mathbb{Z}_{\geq 0}$ 可以从集合论的公理来构造. 具体地说, 遵循万物皆集合的原则, 我们递归地定义

$$\begin{aligned} 0 &:= \varnothing, \quad 1 := \{\varnothing\}, \\ &\ldots, \quad n+1 := \{0,\ldots,n\}. \end{aligned} \tag{2.6.1}$$

严谨的构造依赖于公理集合论的语言. 对于眼下的主题, 要点仅只是

- ⋆ $\mathbb{Z}_{\geq 0}$ 带有加法和乘法运算, 满足结合律, 交换律, 分配律等熟悉的性质;
- ⋆ $\mathbb{Z}_{\geq 0}$ 相对于大小关系 $\leq$ 成为全序集, 其中每个非空子集都有极小元 (称为**良序原理**);
- ⋆ 在证明关于 $\mathbb{Z}_{\geq 0}$ 的种种论断时, 可以应用**数学归纳法** (换句话说, **递归**地证明).

详细说明可见 §§A.1–A.2. 本节视此为不言自明的事实, 而将重点置于如何从 $\mathbb{Z}_{\geq 0}$ 过渡到 $\mathbb{Z}$ 和 $\mathbb{Q}$, 以及如何在 $\mathbb{Z}$ 和 $\mathbb{Q}$ 上定义熟悉的代数运算. 之所以如此耐心地推演, 不单单是以严谨求心安, 更是因为这些技术将以更广泛的面目重现, 从而帮助我们认识崭新而自然的代数结构.

首先研究如何定义整数集 $\mathbb{Z}$ 和其上的代数运算[5]. 无论如何定义, 任意整数 x 都应当能表成 $m-n$ 的形式, 其中 $m,n\in\mathbb{Z}_{\geq 0}$. 减法运算是我们要构造的目标, 它不能作为立足点, 所以我们希望以数对 $(m,n)\in\mathbb{Z}_{\geq 0}^2:=\mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}$ 来表达 x; 尽管表法 $x=m-n$ 并非唯一, 对所有 $(m,n),(m',n')\in\mathbb{Z}_{\geq 0}^2$ 都应当有

$$m-n=m'-n' \iff m+n'=m'+n.$$

这就启发以下定义: 在 $\mathbb{Z}_{\geq 0}^2$ 上定义二元关系

$$(m,n)\sim(m',n') \iff m+n'=m'+n.$$

等价关系: 反身性和对称性不难看出, 唯一需要动笔验证的是传递性, 设 $(m,n)\sim(m',n')$ 而 $(m',n')\sim(m'',n'')$, 则从

$$m+n''+m'=m+n'+m''=m'+n+m''$$

两边消去 m' 可得 $m+n''=n+m''$, 亦即所求之 $(m,n)\sim(m'',n'')$. 这里用到了 $\mathbb{Z}_{\geq 0}$ 中的加法消去律.

定义 2.6.1 定义整数集 $\mathbb{Z}$ 为 $\mathbb{Z}_{\geq 0}^2$ 对 $\sim$ 的商. 暂记 $\mathbb{Z}_{\geq 0}^2$ 中包含 (m,n) 的等价类为 $[\![m,n]\!]$.

我们首先说明如何以单射将 $\mathbb{Z}_{\geq 0}$ 嵌入 $\mathbb{Z}$, 从而将 $\mathbb{Z}$ 视为 $\mathbb{Z}_{\geq 0}$ 的扩张. 想法是简单的: 既然 $m\in\mathbb{Z}_{\geq 0}$ 可以想象为 $m-0$, 自然应当取

$$\begin{aligned}\mathbb{Z}_{\geq 0}&\to\mathbb{Z}\\ m&\mapsto[\![m,0]\!].\end{aligned}$$

这是单射, 因为 $(m,0)\sim(m',0)$ 等价于 $m+0=m'+0$, 亦即 $m=m'$. 有鉴于此, 今后不加说明地将 $\mathbb{Z}_{\geq 0}$ 视同 $\mathbb{Z}$ 的子集.

既然已造出整数集, 下一步是定义其上的代数运算.

定义 2.6.2 对 $\mathbb{Z}$ 的任意元素 $[\![m,n]\!]$ 和 $[\![r,s]\!]$, 其和与积分别定义为

$$\begin{aligned}[\![m,n]\!]+[\![r,s]\!]&:=[\![m+r,n+s]\!],\\ [\![m,n]\!]\cdot[\![r,s]\!]&:=[\![mr+ns,nr+ms]\!].\end{aligned}$$

按惯例, 乘法 $x\cdot y$ 也常简写为 xy.

[5] 除了以下介绍的方法, 造 $\mathbb{Z}$ 的另一进路是以无交并向 $\mathbb{Z}_{\geq 0}$ 添入负整数. 这么做虽然避开了商集, 却会导致一些运算的定义变得做作, 而且难以推广到其他场景.

为了理解此定义, 仅须将 $[\![m,n]\!]$ 和 $[\![r,s]\!]$ 分别想象为 $m-n$ 和 $r-s$.

加法运算是良定义的: 设 $(m,n)\sim(m',n')$, 则由 $m+r+n'+s=m'+r+n+s$ 易得 $(m'+r,n'+s)\sim(m+r,n+s)$. 同理, 若 $(r,s)\sim(r',s')$ 则 $(m+r',n+s')\sim(m+r,n+s)$.

类似方法可以说明乘法良定义: 对 $m'+n=m+n'$ 两边先后乘以 r 和 s, 得到的等式相加给出

$$\begin{aligned} m'r+n's+nr+ms &= (m'+n)r+(m+n')s \\ &= (m+n')r+(m'+n)s \\ &= mr+ns+n'r+m's, \end{aligned}$$

从而 $(m'r+n's,n'r+m's)\sim(mr+ns,nr+ms)$. 同理有 $(mr'+ns',nr'+ms')\sim(mr+ns,nr+ms)$.

这些运算满足下列熟知的性质:

▷ **加法结合律** $x+(y+z)=(x+y)+z$.

▷ **加法交换律** $x+y=y+x$.

▷ **加法零元** $x+0=x$.

▷ **加法消去律** $x+z=y+z$ 蕴涵 $x=y$.

▷ **乘法结合律** $x(yz)=(xy)z$.

▷ **乘法对非零元的消去律** $xz=yz$ 而 $z\neq 0$ 蕴涵 $x=y$.

▷ **乘法幺元** $x\cdot 1=x$.

▷ **分配律** $(x+y)z=xz+yz$.

由此还可以进一步导出 $x\cdot 0=0$, 这是从 $x\cdot 0=x\cdot(0+0)=x\cdot 0+x\cdot 0$ 两边消去 $x\cdot 0$ 的结论. 上列诸性质可以化约到 $\mathbb{Z}_{\geq 0}$ 的相应版本, 路数照旧, 既不困难也不算特别复杂, 留作本章习题.

重点是 $\mathbb{Z}$ 与 $\mathbb{Z}_{\geq 0}$ 的不同之处, 即减法. 首先在 $\mathbb{Z}$ 上定义加法逆元运算

$$-[\![m,n]\!]:=[\![n,m]\!],\quad [\![m,n]\!]\in\mathbb{Z}.$$

这是良定义的: 若 $(m,n)\sim(m',n')$, 则由 $\sim$ 的定义立见 $(n,m)\sim(n',m')$. 我们有

▷ **加法逆元性质** $x+(-x)=0$.

诚然, $[\![m,n]\!]+[\![n,m]\!]=[\![m+n,m+n]\!]=[\![0,0]\!]=0$.

加法逆元性质抽象地刻画了从 $\mathbb{Z}$ 到 $\mathbb{Z}$ 的映射 $x\mapsto -x$, 换言之满足加法逆元性质的 $-x$ 由 x 唯一确定, 这是因为若 $x',x''\in\mathbb{Z}$ 满足 $x+x'=0=x+x''$, 则对两边消去 x 可得 $x'=x''$.

进一步, 暂且记 x 的加法逆元为 x', 则加法交换律导致加法逆元的刻画 $x+x'=0$ 对 x 和 x' 具有对称的形式; 回到惯用符号, 这就改写成负负得正

$$-(-x)=x.$$

今后简记 $x+(-y)$ 为 $x-y$.

运用乘法幺元性质和分配律, 从 $x+(-1)\cdot x=1\cdot x+(-1)\cdot x=(1-1)\cdot x=0$ 立即导出

$$(-1)\cdot x=-x.$$

注记 2.6.3 以商集构造具备所求性质的结构是代数学中的标准技术, 往后将反复运用. 基于相同思路, 附录部分的例 B.2.8 将推广从 $\mathbb{Z}_{\geq 0}$ 到 $\mathbb{Z}$ 的构造方法, 但不再考虑乘法; 该处的例 B.5.12 将说明这是 "添入加法逆元" 此一问题的唯一最优解, 因而是合理的.

继续引入 $\mathbb{Z}$ 上的序结构.

定义 2.6.4 在 $\mathbb{Z}$ 上定义全序 $\leq$ 如下

$$x\leq y \iff y-x\in\mathbb{Z}_{\geq 0}.$$

练习 2.6.5 针对上述定义, 验证:

(i) 设 $x=[\![m,n]\!]\in\mathbb{Z}$, 则 $x\geq 0 \iff m\geq n$ 而 $x\leq 0 \iff m\leq n$;

(ii) $(\mathbb{Z},\leq)$ 确实为全序集.

全序 $\leq$ 在加法和乘法之下具有熟悉的性质:

$$\begin{aligned} x\geq y &\implies x+z\geq y+z,\\ x\geq y\ \wedge\ z\geq 0 &\implies xz\geq yz,\\ x\geq 0 &\iff -x\leq 0. \end{aligned} \tag{2.6.2}$$

这些性质都可以从 $\leq$ 的定义和 $\mathbb{Z}_{\geq 0}$ 上的情形来推导, 细节留作本章习题.

给定 $x\in\mathbb{Z}$ 和 $y\in\mathbb{Z}\smallsetminus\{0\}$, 根据乘法对非零元 y 的消去律, 满足 $x=yq$ 的 $q\in\mathbb{Z}$ 若存在则是唯一的, 可以合理地记为 $\frac{x}{y}$; 这时我们说 y 整除 x, 或记为 $y\mid x$. 为了使分式[6)]表法 $\frac{y}{x}$ 或曰 "比例" 在非整除情形也有意义, 我们转向有理数集 $\mathbb{Q}$ 的严格构造.

思路是类似的, 我们希望用数对 $(r,s)\in\mathbb{Z}^2$ 来表达所欲构造的分式 $\frac{r}{s}$, 其中 $s\neq 0$; 所需的等价关系无非是分式的交叉相乘.

定义 2.6.6 定义有理数集 $\mathbb{Q}$ 为 $\mathbb{Z}\times(\mathbb{Z}\smallsetminus\{0\})$ 对以下等价关系的商集

$$(r,s)\sim(r',s') \iff rs'=r's.$$

暂记含 (r,s) 的等价类为 $[r,s]$. 通过 $x\mapsto[x,1]$, 我们将 $\mathbb{Z}$ 视同 $\mathbb{Q}$ 的子集.

[6)] 本书不区分分式和分数这两种术语.

应用乘法对非零元的消去律, 易见 $[r,s]=0$ 当且仅当 $r=0$.

练习 2.6.7 验证以上定义的 $\sim$ 确实是等价关系, 而 $\mathbb{Z}\to\mathbb{Q}$ 是单射.

通过在等价类中任取代表元, $\mathbb{Q}$ 上的加法和乘法运算定义为

$$[r_1,s_1]+[r_2,s_2]:=[r_1s_2+r_2s_1,s_1s_2],$$
$$[r_1,s_1]\cdot[r_2,s_2]:=[r_1r_2,s_1s_2].$$

这里用到了非零整数相乘依然非零这一性质.

定义 2.6.8 在 $\mathbb{Q}$ 上定义全序, 使得

$$[r,s]\geq 0 \iff rs\geq 0,\quad x\geq y \iff x-y\geq 0.$$

对任意 $x\in\mathbb{Q}$, 其绝对值 $|x|$ 按定义等于 x (当 $x\geq 0$) 或 $-x$ (当 $x<0$).

练习 2.6.9 验证上式确实给出 $\mathbb{Q}$ 上的全序, 而且此全序限制到 $\mathbb{Z}$ 上给出原有的序结构.

注意到在 s 整除 r 的情形, $[r,s]$ 正是之前定义的 $\frac{r}{s}$. 对于一般情形, 记

$$\frac{r}{s}:=[r,s]\in\mathbb{Q}$$

也是合理而方便的; r (或 s) 称为此表达式的分子 (或分母).

先前关于整数集 $\mathbb{Z}$, 其加法和乘法运算, 以及序结构的性质对有理数集 $\mathbb{Q}$ 依然成立, 而且 $\mathbb{Q}$ 上的加法和乘法限制到子集 $\mathbb{Z}$ 上给出原有的运算. 所需论证繁而不难, 请感兴趣的读者自行练习. 眼下的重点是 $\mathbb{Q}$ 和 $\mathbb{Z}$ 的不同之处: 任何非零元都有乘法逆元. 首先, 记

$$\mathbb{Q}^\times:=\mathbb{Q}\smallsetminus\{0\}.$$

命题 2.6.10 对于任意 $x\in\mathbb{Q}^\times$, 存在唯一的 $x^{-1}\in\mathbb{Q}^\times$ 使得 $xx^{-1}=1$.

证明 记 $x=\frac{r}{s}$, 其中 $r,s\in\mathbb{Z}\smallsetminus\{0\}$, 则 $x^{-1}:=\frac{s}{r}$ 满足 $xx^{-1}=\frac{rs}{rs}=1$. 至于唯一性, 设 $x',x''\in\mathbb{Q}$ 满足 $xx'=1=xx''$, 则有

$$x'=x'\cdot 1=x'xx''=1\cdot x''=x'';$$

这是我们之前见过的技巧. □

按惯例, x^{-1} 也记为 $\frac{1}{x}$. 从乘法逆元的刻画 $xx^{-1}=1$ 和乘法交换律容易看出

$$(x^{-1})^{-1}=x,\quad (xy)^{-1}=y^{-1}x^{-1}=x^{-1}y^{-1},$$

其中 $x,y\in\mathbb{Q}^\times$.

关于有理数集的构造留下了一个尾巴: 分式表达式 $x=\frac{r}{s}$ 相对于给定的 $x\in\mathbb{Q}$ 并非唯一. 能否为每个非零的有理数找到基本上唯一的, 而且最为经济的分式表法? 答案来自中学数学司空见惯的一则事实, 尽管不是人人都留心过它的证明.

定义–命题 2.6.11 设 $r,s\in\mathbb{Z}$, 其中 $s\neq 0$. 若 r 和 s 互素, 则称分式表达式 $\frac{r}{s}$ 是**既约**的. 每个有理数都有既约表达式, 而且两个既约分式满足

$$\frac{r_1}{s_1}=\frac{r_2}{s_2}$$

的充要条件是 $r_1=r_2$, $s_1=s_2$ 或 $r_1=-r_2$, $s_1=-s_2$.

既约分式的定义涉及一个尚未梳理的概念 — 互素, 它的证明也必须以整数的算术理论为先决要件. 欲知详情, 且待下节分解.

注记 2.6.12 从有理数集 $\mathbb{Q}$ 向实数集 $\mathbb{R}$ 的过渡或者需要 Dedekind 分割, 或者需要以 Cauchy 数列取完备化, 两者都不在代数学的传统范围, 所以本书不予处理. 请有兴趣的读者参考分析教材, 如 [7, §1.1].

2.7 算术入门

本节进一步勾勒整数集 $\mathbb{Z}$ 的基本代数性质. 它们多数是中学数学所介绍过的概念, 只是符号略有出入; 先前几节的例子中也曾不加说明地用过这些简单事实.

首先请回忆整除的定义. 对任意 $x\in\mathbb{Z}$ 定义 $x\mathbb{Z}:=\{xd:d\in\mathbb{Z}\}$, 它由 x 的所有倍数构成. 对于 $x,y\in\mathbb{Z}$, 如果 $y\in x\mathbb{Z}$ 则称 x 整除 y, 记为 $x\mid y$, 否则记为 $x\nmid y$. 当 $x\mid y$ 时, 我们称 x 为 y 的因数或因子.

命题 2.7.1 (整数的带余除法) 对于任意 $a,d\in\mathbb{Z}$, 若 $d\neq 0$, 则存在唯一的 $q,r\in\mathbb{Z}$ 使得 $0\leq r<|d|$ 而 $a=dq+r$.

证明 不妨假定 $d>0$. 考虑集合

$$R:=\{a-dq:q\in\mathbb{Z}\}\cap\mathbb{Z}_{\geq 0}.$$

这是 $\mathbb{Z}_{\geq 0}$ 的非空子集, 故有极小元, 记为 r; 相应地 $a=dq+r$. 必然有 $r<d$, 否则 $a=d(q+1)+(r-d)$ 将给出 $r-d\in R$ 使得 $0\leq r-d<r$, 与 r 的极小性质矛盾. 这就说明 q,r 的存在性.

至于唯一性, 设 $dq+r=dq'+r'$, 其中 $q,q'\in\mathbb{Z}$ 而 $0\leq r\leq r'<d$. 因为 $r'-r=d(q-q')$ 既被 d 整除, 又有 $0\leq r'-r\leq r'<d$, 易证唯一可能是 $r'=r$, 从而由 $d\neq 0$ 推得 $q=q'$. □

根据上述命题中的唯一性, 带余除法中的余数 $r=0$ 当且仅当 $d\mid a$.

引理 2.7.2 设 I 为 $\mathbb{Z}$ 的非空子集, 满足以下性质

$$\begin{aligned} x,y\in I &\implies x+y\in I,\\ a\in\mathbb{Z},\ x\in I &\implies ax\in I.\end{aligned}$$

此时存在唯一的 $g\in\mathbb{Z}_{\geq 0}$ 使得 $I=g\mathbb{Z}$.

证明 先讨论 g 的存在性. 不妨设 $I \neq \{0\}$, 否则唯一取法是 $g = 0$. 注意到 $g \in I \iff -g \in I$. 取 g 为 I 中的最小正整数. 包含关系 $I \supset g\mathbb{Z}$ 自明. 至于 $\subset$, 设 $m \in I$, 用带余除法表为 $m = gq + r$, 其中 $0 \leq r < g$. 于是 $r = m - gq$ 必为 0, 否则 r 将给出 I 中比 g 更小的正整数, 矛盾.

至于 g 的唯一性, 若正整数 g, g' 满足 $g\mathbb{Z} = g'\mathbb{Z}$, 则它们相互整除, 故唯一可能是 $g = g'$. □

以下选定 $n \in \mathbb{Z}_{\geq 1}$, 并考虑一族整数 $x_1, \ldots, x_n \in \mathbb{Z}$.

⋆ 记这族整数的最小公倍数为

$$\operatorname{lcm}(x_1, \ldots, x_n) := \begin{cases} \min\{m \in \mathbb{Z}_{\geq 1} : \forall 1 \leq i \leq n,\ x_i \mid m\}, & \forall i,\ x_i \neq 0, \\ 0, & \exists i,\ x_i = 0. \end{cases}$$

⋆ 记这族整数的最大公因数为

$$\gcd(x_1, \ldots, x_n) := \begin{cases} \max\{d \in \mathbb{Z}_{\geq 1} : \forall 1 \leq i \leq n,\ d \mid x_i\}, & \exists i,\ x_i \neq 0, \\ 0, & \forall i,\ x_i = 0. \end{cases}$$

定义在 $x_1 = \cdots = x_n = 0$ 时的合理性可以由稍后的命题 2.7.3 支持.

⋆ 若 $\gcd(x_1, \ldots, x_n) = 1$, 则称 $x_1, \ldots, x_n$ **互素**, 这也相当于说 $x_1, \ldots, x_n$ 没有 ± 1 之外的公因子.

对给定的 $x_1, \ldots, x_n \in \mathbb{Z}$, 定义 $\mathbb{Z}$ 的子集

$$\begin{aligned} \sum_{i=1}^{n} \mathbb{Z}x_i &= \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_n \\ &:= \left\{ \sum_{i=1}^{n} a_i x_i \in \mathbb{Z} : a_1, \ldots, a_n \in \mathbb{Z} \right\}; \end{aligned}$$

按照惯例, 极端情形 $n = 0$ (即 "空和") 按 $\sum_{i=1}^{0} \mathbb{Z}x_i := \{0\}$ 来解释.

命题 2.7.3 (É. Bézout) *设 $x_1, \ldots, x_n$ 为整数, 则*

$$\mathbb{Z}x_1 + \cdots + \mathbb{Z}x_n = \gcd(x_1, \ldots, x_n)\mathbb{Z}.$$

作为推论, $x_1, \ldots, x_n$ 互素的充要条件是存在 $a_1, \ldots, a_n \in \mathbb{Z}$ 使得 $a_1x_1 + \cdots + a_nx_n = 1$.

证明 先处理第一部分. 不妨设 $x_1, \ldots, x_n$ 不全为 0, 否则等式两边按定义同等于 0.

以引理 2.7.2 取 $g \in \mathbb{Z}_{\geq 1}$, 使得 $\mathbb{Z}x_1 + \cdots + \mathbb{Z}x_n = g\mathbb{Z}$. 请读者验证对于任意正整数 d, 我们有

$$(\forall 1 \leq i \leq n,\ d \mid x_i) \iff (\forall x \in \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_n,\ d \mid x) \iff d \mid g.$$

这就足以表明 $g = \gcd(x_1, \dots, x_n)$.

现在处理第二部分. 基于上述结果, $\gcd(x_1, \dots, x_n) = 1$ 蕴涵 1 可以表为 $a_1x_1 + \cdots + a_nx_n$ 的形式; 反之, 若存在 $a_1, \dots, a_n$ 使得 $1 = a_1x_1 + \cdots + a_nx_n$, 则 $x_1, \dots, x_n$ 当然互素. □

定义 2.7.4 设 $p \in \mathbb{Z} \smallsetminus \{0, \pm 1\}$. 如果 p 除了 ± 1 和 $\pm p$ 之外没有别的因数, 则称 p 为**素元**. 正的素元称为**素数**.

命题 2.7.5 (Euclid) 设 p 为素元, 若 $a, b \in \mathbb{Z}$ 使得 $p \mid ab$, 则必有 $p \mid a$ 或 $p \mid b$.

证明 不妨设 $p \nmid a$. 因为 p 是素元, 此时 a 和 p 必无 ± 1 以外的公因子, 亦即互素. 命题 2.7.3 蕴涵存在 $x, y \in \mathbb{Z}$ 使得 $1 = px + ay$. 于是 $p \mid pxb + aby = b$. □

定理 2.7.6 (算术基本定理) 任何非零整数 $n \in \mathbb{Z}$ 都有素因子分解

$$n = \pm p_1^{a_1} \cdots p_r^{a_r},$$

其中 $r \in \mathbb{Z}_{\geq 0}$ (当 $r = 0$ 时右式规定为 ± 1), $p_1, \dots, p_r$ 是相异素数, $a_1, \dots, a_r \in \mathbb{Z}_{\geq 1}$, 而且此分解不论顺序是唯一的.

证明 关于分解的存在性, 处理 $n \geq 1$ 情形即可. 我们寻求形如 $n = p_1^{a_1} \cdots p_r^{a_r}$ 的分解. 如果 n 既非 1 又非素数, 则分解为 $n = ab$, 其中 $1 < a, b < n$. 继续对 a 和 b 递归地操作, 最终可表 n 为若干个素数的乘积, 容许重复.

唯一性仍可简化到 $n \geq 1$ 情形. 设 $p_1^{a_1} \cdots p_r^{a_r} = q_1^{b_1} \cdots q_s^{b_s}$, 其中 $p_1, \dots, p_r$ 是相异素数, $q_1, \dots, q_s$ 也是相异素数, 而 $a_i, b_j \in \mathbb{Z}_{\geq 1}$. 注意到 $r = 0$ 当且仅当 $s = 0$, 此时两边都是 1. 故以下无妨设 $r, s \geq 1$.

由于 $p_1 \mid q_1^{b_1} \cdots q_s^{b_s}$, 反复应用命题 2.7.5 可知存在 $1 \leq j \leq s$ 使得 $p_1 \mid q_j$; 这进一步蕴涵 $p_1 = q_j$. 重排下标后不妨假设 $p_1 = q_1$, 必要时等号两边互换, 不妨假设 $a_1 \leq b_1$. 于是

$$p_2^{a_2} \cdots p_r^{a_r} = p_1^{b_1 - a_1} q_2^{b_2} \cdots q_s^{b_s}.$$

再次应用命题 2.7.5 可见 p_1 不整除左式, 故 $b_1 = a_1$. 按此递归地论证, 即得分解的唯一性. □

因此对于任何素数 p, 我们有 $p \mid n$ 当且仅当 p 在 n 的素因子分解中出现, 相应的指数 $a \in \mathbb{Z}_{\geq 1}$ 由以下性质唯一确定: $p^a \mid n$ 而 $p^{a+1} \nmid n$, 数论中的标准记法如下.

约定 2.7.7 设 p 为素数, 我们以符号 $p^a \parallel n$ 表达 $p^a \mid n$ 而 $p^{a+1} \nmid n$.

推论 2.7.8 考虑整数 $n = \pm \prod_{i=1}^r p_i^{a_i}$ 和 $m = \pm \prod_{i=1}^r p_i^{b_i}$, 其中 $p_1, \dots, p_r$ 是相异素数而 $a_i, b_i \in \mathbb{Z}_{\geq 0}$, 则

$$\gcd(n, m) = \prod_{i=1}^r p_i^{\min\{a_i, b_i\}}, \quad \operatorname{lcm}(n, m) = \prod_{i=1}^r p_i^{\max\{a_i, b_i\}}.$$

对于任意多个正整数的 gcd 和 lcm 也有类似结果.

练习 2.7.9 利用上述结果, 证明对所有正整数 n, m 皆有

$$nm = \gcd(n, m) \cdot \operatorname{lcm}(n, m).$$

练习 2.7.10 给出定义–命题 2.6.11 的完整证明.

提示〉 考虑分子非零的情形即可. 首先说明任何 $\frac{r}{s}$ 都能化为既约分式. 以算术基本定理 2.7.6 对 r 和 s 作素因子分解, 提出公因子即可.

其次, 考虑既约分式的等式 $\frac{r_1}{s_1} = \frac{r_2}{s_2}$. 从 $r_1 s_2 = r_2 s_1$ 和互素的条件证明 $s_1 \mid s_2$; 对称地, $s_2 \mid s_1$. 由此可得 $s_2 = \pm s_1$.

素数列 $2, 3, 5, 7, 11, \ldots$ 是数论关切的基本对象; 这方面最古老也是最基础的结果如下.

定理 2.7.11 (Euclid) *存在无穷多个素数.*

证明 对任意一列素数 $p_1 < \cdots < p_n$, 考虑

$$m := p_1 \cdots p_n + 1,$$

则 $m > 1$, 而且它不被 $p_1, \ldots, p_n$ 中任一个素数整除. 因此 m 的素因子分解中必有不同于 $p_1, \ldots, p_n$ 的素数. □

2.8 同余式

在 §2.5 和 §2.7 的基础上, 我们接着介绍整数的同余关系, 它的妙用在 §1.1 关于平方和问题的介绍中已显端倪.

定义 2.8.1 (同余式) *设 $N \in \mathbb{Z}$. 称 $a, b \in \mathbb{Z}$ 是 mod N **同余**的, 如果 $N \mid a - b$; 此关系也写作*

$$a \equiv b \pmod{N}.$$

易见 mod N 同余是 $\mathbb{Z}$ 上的等价关系: 诚然,

- ⋆ $a \equiv a \pmod{N}$ (因为 $N \mid 0$, 或者说 $0 \in N\mathbb{Z}$);
- ⋆ $a \equiv b \pmod{N}$ 等价于 $b \equiv a \pmod{N}$ (因为 $N \mid x$ 等价于 $N \mid -x$, 或者说 $N\mathbb{Z}$ 对运算 $x \mapsto -x$ 封闭);
- ⋆ $a \equiv b \pmod{N}$ 和 $b \equiv c \pmod{N}$ 蕴涵 $a \equiv c \pmod{N}$ (因为 $N\mathbb{Z}$ 对加法封闭).

定义 2.8.2 (同余类) *选定 $N \in \mathbb{Z}$, 记 $\mathbb{Z}$ 对等价关系 mod N 的商集 (定义 2.5.4) 为 $\mathbb{Z}/N\mathbb{Z}$, 或简记为 $\mathbb{Z}/N$; 其中的等价类也称为 mod N 同余类.*

给定 $x \in \mathbb{Z}$, 包含 x 的同余类可以具体地被描述为 $\mathbb{Z}$ 的子集

$$x + N\mathbb{Z} := \{a + Nd : d \in \mathbb{Z}\}.$$

以后我们也会使用更简短的符号如 $[x]$, $[x]_N$ 或 $x \bmod N$ 等来标记含 $x \in \mathbb{Z}$ 的 mod N 同余类, 便宜行事.

利用带余除法, 对每个 $a \in \mathbb{Z}$ 取其除以 N 的余数, 暂记为 $R_N(a) \in \{0, \dots, N-1\}$, 则有

$$a \equiv a' \pmod{N} \iff R_N(a) = R_N(a'),$$

此即 "同余" 之义. 代入命题 2.5.8 (取 $f = R_N$) 遂有双射

$$\overline{R_N} : \mathbb{Z}/N\mathbb{Z} \xrightarrow{1:1} \{0, \dots, N-1\}.$$

借此, mod N 的同余类和 $\{0, \dots, N-1\}$ 的元素一一对应, 这就将 $\mathbb{Z}$ 划分为 N 个 mod N 同余类, 以 $0, \dots, N-1$ 为具体的代表元; 这些代表元的取法当然不唯一, 我们同样可以取 $1, \dots, N$ 等. 取余数所呈现的经常是一种虚假的具体感; 多数场合下, 直接操作同余类更为简便.

练习 2.8.3 说明

$$a \equiv b \pmod{N} \implies \gcd(a, N) = \gcd(b, N);$$

特别地, 如果 $b > a \geq 1$ 而 b 用带余除法表作 $b = aq + r$, 其中 $0 \leq r < a$, 则

$$\gcd(a, b) = \gcd(r, a) \quad \begin{cases} = a, & \text{若 } r = 0, \\ \text{继续作带余除法}, & \text{若 } r \neq 0. \end{cases}$$

这便是以辗转相除法求最大公因数的实质.

练习 2.8.4 证明若 $x \equiv x' \pmod{N}$, $y \equiv y' \pmod{N}$, 则有

$$x + y \equiv x' + y' \pmod{N}, \quad xy \equiv x'y' \pmod{N}.$$

一言以蔽之, 同余关系兼容于加法和乘法. 我们在 §1.1 证明过 $m \equiv -1 \pmod{4}$ 时 $X^2 + Y^2 = m$ 无整数解, 当时运用的无非是关于同余的上述性质.

作为同余式的初步例子, 以下是称为 **Fermat 小定理**的著名结果. 它有简单的群论诠释, 行将给出的则是初等的迂回论证. 作为准备, 我们先来研究同余式 $xy \equiv 1 \pmod{N}$ 对哪些 $x \in \mathbb{Z}$ 有解; 当然, 答案仅依赖于 $x \bmod N$.

命题 2.8.5 设 $N \in \mathbb{Z}_{\geq 1}$. 对于任意 $x \in \mathbb{Z}$, 我们有

$$(\exists y \in \mathbb{Z},\ xy \equiv 1 \pmod{N}) \iff \gcd(N, x) = 1.$$

证明 左式有解相当于说存在 $y, z \in \mathbb{Z}$ 使得 $xy - Nz = 1$, 亦即 $x\mathbb{Z} + N\mathbb{Z} \ni 1$. 将此代入命题 2.7.3. □

定理 2.8.6 (P. Fermat) *设 p 为素数, 则对于所有 $x \in \mathbb{Z}$ 都有*

$$\gcd(p, x) = 1 \implies x^{p-1} \equiv 1 \pmod{p}. \tag{2.8.1}$$

作为推论, 所有 $x \in \mathbb{Z}$ 都满足 $x^p \equiv x \pmod{p}$.

证明 设 $p \nmid x$. 根据命题 2.8.5, 存在 $y \in \mathbb{Z}$ 使得 $xy \equiv 1 \pmod{p}$. 现在考虑 x 的所有整数倍. 如果 $k_1 x \equiv k_2 x \pmod{p}$, 利用练习 2.8.4 对两边同乘以 y, 可得 $k_1 \equiv k_2 \pmod{p}$. 另一方面, 素数的性质确保 $p \nmid k$ 时 $kx \not\equiv 0 \pmod{p}$. 这一切表明

$$kx, \quad k = 1, \dots, p-1$$

两两互不同余, 而且皆不 $\equiv 0 \pmod{p}$, 所以它们的同余类和 $1, \dots, p-1$ 的同余类仅差一个重排. 再次利用练习 2.8.4 可得

$$x^{p-1}(p-1)! = \underbrace{x \cdots ((p-1)x)}_{p-1 \text{ 项}} \equiv \underbrace{1 \cdots (p-1)}_{p-1 \text{ 项}} \equiv (p-1)! \pmod{p}.$$

因为 $p \nmid (p-1)!$, 仿照之前办法可从同余式两边消去 $(p-1)!$, 这就证明了第一部分.

对于一般的 $x \in \mathbb{Z}$, 或者 $\gcd(p, x) = 1$, 从而对 $x^{p-1} \equiv 1 \pmod{p}$ 两边同乘以 x 给出 $x^p \equiv x \pmod{p}$; 或者 $p \mid x$, 从而 $x^p \equiv x \pmod{p}$ 平凡地成立 (两边皆同余 0), 这就证明了第二部分. □

定理 2.8.6 的逆命题并不成立. 满足 (2.8.1) 而非素数的正整数 p 称为 Carmichael 数, 有无穷多个; 前五个 Carmichael 数是 $561, 1105, 1729, 2465, 2821$. 尽管存在这些反例, 性质 (2.8.1) 仍然在一些概率素性检测算法中扮演要角.

选定 $n \in \mathbb{Z}_{\geq 1}$. 回忆到一个整数与 n 互素与否仅依赖于它的 mod n 同余类. 细观命题 2.8.5 和定理 2.8.6 的证明, 可以发现与 n 互素的同余类在 $\mathbb{Z}/n\mathbb{Z}$ 中占有特别的地位.

定义 2.8.7 (Euler 函数) *设 $n \in \mathbb{Z}_{\geq 1}$, 定义 $\varphi(n)$ 为不超过 n 而与 n 互素的正整数个数.*

由于 $\mathbb{Z}/n\mathbb{Z}$ 可以通过代表元等同于 $\{1, \dots, n\}$, Euler 函数 $\varphi(n)$ 正是与 n 互素的 mod n 同余类个数. 注意到 $\varphi(1) = 1$.

练习 2.8.8 验证 Euler 函数 φ 的以下性质.

(i) 若 $n = p_1^{a_1} \cdots p_r^{a_r}$ 是素因子分解, 其中 $a_i \in \mathbb{Z}_{\geq 1}$, 则

$$\varphi(n) = \prod_{i=1}^{r} \left(p_i^{a_i} - p_i^{a_i - 1}\right) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

(ii) 若 $n, m \in \mathbb{Z}_{\geq 1}$ 互素, 则 $\varphi(nm) = \varphi(n)\varphi(m)$.

(iii) 证明 $\sum_{\substack{d \geq 1 \\ d|n}} \varphi(d) = n$.

(iv) 证明 $\lim_{n \to +\infty} \varphi(n) = +\infty$.

练习 2.8.9 (Möbius 函数) 定义 Möbius 函数 $\mu : \mathbb{Z}_{\geq 1} \to \{0, \pm 1\}$ 如下:

$$\mu(m) := \begin{cases} (-1)^{m \text{ 的素因子个数}}, & m \text{ 无平方因子}, \\ 0, & m \text{ 有平方因子}. \end{cases}$$

(i) 证明 Möbius 函数满足 $\gcd(a, b) = 1 \implies \mu(ab) = \mu(a)\mu(b)$.

(ii) 证明 $\varphi(n) = \sum_{d|n} \mu(d) \cdot \frac{n}{d}$; 当然, 此处的 $\sum_{d|n}$ 意谓对 n 的所有正因数求和.

2.9 集合的基数

集合的大小是数学中的一个基本概念. 显而易见, $\{0, \ldots, n-1\}$ 的大小应当是 n, 而空集的大小应当是 0. 对于一般的集合 A, 最简单的方法是分成有限和无穷两种情形. 为了代数学的研究, 还必须对无穷集的大小作进一步的细分.

从日常经验启发, 一种自然的思路是将 A 的元素排序, 以排序的结果来丈量集合的大小. 这可以类比于排队报数. 对于有限集 A, 选取 A 上的全序就相当于选取双射 $A \xrightarrow{1:1} \{0, \ldots, n-1\}$, 此时 n 正是 A 的大小, 它不依赖全序的选取.

对于无穷集的情形, 所需序结构的性质, 其存在性, 以及赖以丈量大小的标尺 (例如先前的全序集 $\{0, \ldots, n-1\}$) 都有待明确; 此外还要说明丈量结果和排序方式无关. 这一切终将导向 Cantor 的序数理论, 需要集合论的进阶知识. 本书退而求其次, 仅将集合的大小关系作为一种等价关系来理解.

定义 2.9.1 设 A 和 B 为集合, 若存在双射 $f : A \xrightarrow{1:1} B$, 则称 A 和 B **等势**, 或者说它们有相同的**基数**, 记作 $|A| = |B|$.

等势是集合之间的等价关系: 显然有

(i) $|A| = |A|$ (取 $f = \mathrm{id}_A$),

(ii) $|A| = |B|$ 蕴涵 $|B| = |A|$ (以 f^{-1} 代 f),

(iii) $|A| = |B|$ 连同 $|B| = |C|$ 蕴涵 $|A| = |C|$ (双射作合成仍是双射).

所以集合 A 的基数 $|A|$ 可以理解为一个等势类[7)].

定义 2.9.2 设 A 和 B 为集合, 若存在单射 $f : A \hookrightarrow B$, 则记作 $|A| \leq |B|$; 以 $|A| < |B|$ 表示 $|A| \leq |B|$ 而 $|A| \neq |B|$.

[7)] 所有集合的总体并不构成集合, 所以等势类确实只是注记 2.1.1 意义下的一个 "类". 若愿意改用序数的进路, 则可以具体将等势类或基数理解为一种特殊的集合.

一些简单观察: (i) 任何子集 $A' \subset A$ 都满足 $|A'| \leq |A|$ (取 f 为包含映射 $f(a') = a'$), (ii) 若 $|A| = |B|$ 则 $|A| \leq |B|$ (双射自然是单射), (iii) $|A''| \leq |A'|$ 连同 $|A'| \leq |A|$ 蕴涵 $|A''| \leq |A|$ (单射的合成依然单).

为了使以上定义的 $\leq$ 真正合乎我们对大小关系的要求, 以下性质也是必要的. 对证明感兴趣的读者请查阅 §A.3.

- ⋆ (引理 A.3.1) 若存在满射 $f : B \twoheadrightarrow A$, 则 $|A| \leq |B|$.
- ⋆ (Schröder–Bernstein 定理 A.3.2) 若 $|A| \leq |B|$ 而 $|B| \leq |A|$, 则 $|A| = |B|$.
- ⋆ (定理 A.3.3) 对任意集合 A 和 B, 必有 $|A| \leq |B|$ 或 $|B| \leq |A|$.

约定 2.9.3 设 $n \in \mathbb{Z}_{\geq 0}$, 如果 A 和 $\{0, \ldots, n-1\}$ 等势, 则记为 $|A| = n$; 特别地, $|A| = 0$ 等价于 $A = \varnothing$.

现在可以对有限集作一正式而精确的定义: 若存在 $n \in \mathbb{Z}_{\geq 0}$ 使得 $|A| = n$, 则称 A 有限, 否则称 A 无穷. 以下两条性质几乎是一目了然的, 严谨证明请见 §A.3.

命题 2.9.4 (抽屉原理) *设 A 和 B 是等势的有限集, 则任何单射 (或满射) $f : A \to B$ 自动是双射.*

命题 2.9.5 *集合 A 无穷当且仅当存在单射 $\mathbb{Z}_{\geq 0} \hookrightarrow A$.*

练习 2.9.6 设 A 是无穷集, 则存在单而非满 (或满而非单) 的映射 $f : A \to A$. 这是 R. Dedekind 对无穷集的刻画; 参看命题 2.9.4.

提示〉 以命题 2.9.5 化约到 $A = \mathbb{Z}_{\geq 0}$ 的情形.

我们可以将 $\mathbb{Z}_{\geq 0}$ 等同于有限基数, 而 $\mathbb{Z}_{\geq 0}$ 上的代数运算按以下方式扩展到任意基数.

- ▷ **加法** $|A| + |B| := |A \sqcup B|$ (无交并);
- ▷ **乘法** $|A| \cdot |B| := |A \times B|$ (积集);
- ▷ **指数** $|A|^{|B|} := \left|A^B\right|$ (映射集).

容易看出它们只和 $|A|$, $|B|$ 相关, 不依赖集合 A, B 的选法; 由于对无穷多个集合一样能作无交并或积, 加法和乘法也可以对无穷多个基数来操作.

作为特例, 回忆到 $P(A)$ 代表集合 A 的幂集, 再以 $\mathbf{1}_S : A \to \{0,1\}$ 代表在子集 $S \subset A$ 上取 1, 其外取 0 的函数, 则双射

$$\begin{array}{ccc} \{0,1\}^A & \xleftrightarrow{1:1} & P(A) \\ \cup & & \cup \\ f & \longmapsto & f^{-1}(1) \\ \mathbf{1}_S & \longmapsfrom & S \end{array} \qquad \text{(符号 } f^{-1}(1) \text{ 如定义 2.2.10)}$$

给出基数的等式

$$2^{|A|} \xlongequal{\text{定义}} \left|\{0,1\}^A\right| = |P(A)|.$$

命题 2.9.5 断言无穷集总是包含一份 $\mathbb{Z}_{\geq 0}$ 的副本, 所以 $|\mathbb{Z}_{\geq 0}|$ 是最小无穷基数, 它占有特别的地位.

定义 2.9.7 记 $\aleph_0 := |\mathbb{Z}_{\geq 0}|$. 满足 $|A| = \aleph_0$ 的集合称为**可数集**或**可列集**. 满足 $|A| \leq \aleph_0$ 的集合称为**至多可数集**, 这也相当于说 A 或者有限, 或者可数.

命题 2.9.8 有限多个可数集的并和积依然可数.

证明 说明两个可数集的无交并和积依然可数即可. 对于无交并的情形, 使用以下映射

$$\mathbb{Z}_{\geq 0} \sqcup \mathbb{Z}_{\geq 0} \xrightarrow{1:1} \mathbb{Z}_{\geq 0},$$

它映第一份 $\mathbb{Z}_{\geq 0}$ 之中的 x 为 $2x$, 映第二份 $\mathbb{Z}_{\geq 0}$ 之中的 x 为 $2x+1$.

对于积的情形, 使用以下映射:

$$\begin{aligned} \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} &\xrightarrow{1:1} \mathbb{Z}_{\geq 0} \\ (a,b) &\mapsto 2^a(2b+1) - 1. \end{aligned}$$

这相当于说 $(a,b) \mapsto 2^a(2b+1)$ 是从 $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ 到 $\mathbb{Z}_{\geq 1}$ 的双射, 双射性质是算术基本定理 2.7.6 的简单应用. 这些映射的取法当然不是唯一的. □

注记 2.9.9 推而广之, 以 κ 和 λ 代表任两个集合的基数, 其中至少有一者无穷, 则有 $\kappa \cdot \lambda = \max\{\kappa, \lambda\} = \kappa + \lambda$; 命题 2.9.8 不过是其特例 $\aleph_0 \cdot \aleph_0 = \aleph_0 = \aleph_0 + \aleph_0$. 一般情形的证明比较曲折, 详见 [10, 推论 1.4.9 (i)].

关于有限并的结论还可以强化.

推论 2.9.10 设 $(A_i)_{i\in I}$ 是一族可数集, 而下标集 I 本身也可数, 则 $\bigcup_{i\in I} A_i$ 可数.

证明 为每个 $i \in I$ 选定双射 $f_i : \mathbb{Z}_{\geq 0} \to A_i$. 以此定义映射

$$\begin{aligned} \varphi : I \times \mathbb{Z}_{\geq 0} &\to \bigcup_{i\in I} A_i \\ (i,n) &\mapsto f_i(n). \end{aligned}$$

由于 $\{i\} \times \mathbb{Z}_{\geq 0}$ 的像是 A_i, 映射 φ 显然满. 于是 $\left|\bigcup_{i\in I} A_i\right| \leq |I| \cdot \aleph_0 = \aleph_0 \cdot \aleph_0 = \aleph_0$ (命题 2.9.8). 另一方面, 每个 A_i 皆无穷, 故 $\bigcup_{i\in I} A_i$ 无穷. 综上, $\left|\bigcup_{i\in I} A_i\right| = \aleph_0$. □

如果将条件改成每个 A_i 都是至多可数的, 则 $\bigcup_{i\in I} A_i$ 也至多可数.

例 2.9.11 整数集 $\mathbb{Z}$ 是可数的, 因为它是 $\mathbb{Z}_{\geq 0}$ 和 $\mathbb{Z}_{\leq -1}$ 的并, 而 $\mathbb{Z}_{\leq -1}$ 当然和 $\mathbb{Z}_{\geq 0}$ 等势. 有理数集 $\mathbb{Q}$ 是可数的, 因为它是可数集 $\frac{1}{n}\mathbb{Z} := \left\{\frac{r}{n} : r \in \mathbb{Z}\right\}$ 的并, 其中 $n = 1, 2, \ldots$.

这些例子说明无穷集的大小关系和有限情形大有区别: 一个无穷集 (比如 $\mathbb{Z}$) 可以和它的真子集 (比如偶数子集或奇数子集) 等势. 这就造成了被许多科普书籍称为 "Hilbert 旅馆" 的奇谈怪论: 考虑一间有可数无穷多间客房的旅馆, 即使处在满房状态, 它也能够招待任何一位新客人, 办法是将 n 间的住客撵到第 $n+1$ 间, 其中 n 遍历 $\mathbb{Z}_{\geq 0}$, 从而腾出第 0 间房.

Hilbert 旅馆还能一次腾出可数多间房, 方法是应用命题 2.9.8 的双射 $\mathbb{Z}_{\geq 0} \xrightarrow{1:1} \mathbb{Z}_{\geq 0} \sqcup \mathbb{Z}_{\geq 0}$. 基于推论 2.9.10, 它甚至能招待可数个旅行团, 每团都有可数个人. 然而以下定理表明幂集旅行团 $P(\mathbb{Z}_{\geq 0})$ 可以轻易击穿其接待能力, 缘由是 $2^{\aleph_0} > \aleph_0$; 所以一旦无穷的概念得到细化, Hilbert 旅馆就不再显得突兀了.

定理 2.9.12 (G. Cantor) 对任意集合 A 都有 $2^{|A|} = |P(A)| > |A|$.

证明 映 $a \in A$ 为独点集 $\{a\}$ 给出单射 $A \hookrightarrow P(A)$, 因此 $|A| \leq |P(A)|$. 为了说明这是严格不等式, 以下证任何映射 $\phi : A \to P(A)$ 皆非满. 原因很简单: 考虑 A 的子集 $B := \{a \in A : a \notin \phi(a)\}$, 不存在 $a' \in A$ 使得 $B = \phi(a')$, 否则 $a' \in B$ 或 $a' \notin B$ 俱不成, 具体推导留给读者自娱. □

请注意到以上技巧和 Russell 悖论 (注记 2.1.1) 的相似性.

习题

1. 验证关于集合的以下等式, 其中 I 是非空集, 而 $(Y_i)_{i\in I}$ 代表一族以 $i \in I$ 为下标的集合:

(i) $X \cap (\bigcup_{i\in I} Y_i) = \bigcup_{i\in I}(X \cap Y_i)$;

(ii) $X \cup (\bigcap_{i\in I} Y_i) = \bigcap_{i\in I}(X \cup Y_i)$;

(iii) $X \smallsetminus \bigcup_{i\in I} Y_i = \bigcap_{i\in I}(X \smallsetminus Y_i)$;

(iv) $X \smallsetminus \bigcap_{i\in I} Y_i = \bigcup_{i\in I}(X \smallsetminus Y_i)$.

2. 设 A, B, C 为集合. 定义所谓的坐标投影映射

说明当映射 $f : A \to B$ 和 $g : B \to C$ 给定, $gf : A \to C$ 的函数图形 Γ_{gf} 是

$$\Gamma_{gf} = p_{13}\left(p_{12}^{-1}(\Gamma_f) \cap p_{23}^{-1}(\Gamma_g)\right).$$

3. 给出自然的双射 $A^{I\times J} \xrightarrow{1:1} (A^I)^J$, 其中 I, J, A 是任意集合.

4. 说明若在命题 2.2.6 中容许 A 为空集, 则只能得到等价 (i) $\iff$ (iii) 和 (i)' $\iff$ (iii)'.

5. 设 R 是集合 A 上的二元关系. 定义 A 上的二元关系 R^{op} 使得 $xRy \iff yR^{\mathrm{op}}x$.

(i) 证明 R 给出 A 上的预序 (或偏序, 全序) 结构当且仅当 R^{op} 亦然.

(ii) 对于任意集合 S, 给出从偏序集 $(P(S), \subset)$ 到 $(P(S), \subset^{\mathrm{op}})$ 的同构. 提示 对任意子集 $T \subset S$ 取 $S \smallsetminus T \subset S$.

6. 设 $R \subset A \times B$ 和 $S \subset B \times C$ 为二元关系. 定义它们的合成为

$$SR := \{(a, c) \in A \times C : \exists b \in B,\ aRb \wedge bSc\};$$

这仍是二元关系.

(a) 验证二元关系的合成满足结合律 $T(SR) = (TS)R$, 前提是合成有意义.

(b) 对二元关系 $R \subset A \times B$ 验证 $\Delta_B R = R = R\Delta_A$, 其中 Δ_A 和 Δ_B 如 (2.4.1).

(c) 说明若 R 来自映射 $f : A \to B$ 而 S 来自映射 $g : B \to C$, 则 SR 来自合成映射 $gf : A \to C$.

(d) 对于集合 A 上的二元关系 $R \subset A \times A$ 和 $n \in \mathbb{Z}_{\geq 1}$, 我们定义

$$R^n := \underbrace{R \cdots R}_{\text{合成 } n \text{ 份}}.$$

说明关系 R 满足传递性 (即: $a_1Ra_2 \wedge a_2Ra_3 \implies a_1Ra_3$) 当且仅当 $R^n \subset R$ 对所有 n 皆成立.

7. 设偏序集 $(P, \preceq)$ 的所有非空子集皆有极大元. 证明如果单射 $\theta : P \to P$ 满足 $x \preceq \theta(x)$, 则 $\theta = \mathrm{id}_P$.

提示 命 $S = \{x \in P : \theta(x) \neq x\}$. 若 S 非空则取其极大元 y, 从 $y \prec \theta(y)$ 推导 $\theta(y) \prec \theta(\theta(y))$, 继而推导矛盾.

8. 设 $(A, \preceq)$ 为预序集. 定义 A 上的二元关系 $\sim$ 如下:

$$a \sim a' \iff (a \preceq a') \wedge (a' \preceq a).$$

(i) 说明 $\sim$ 是 A 上的等价关系. 记包含 $a \in A$ 的等价类为 $[a]$.

(ii) 说明可以合理地定义 $A/\sim$ 上的偏序 $\preceq$, 使得 $[a] \preceq [a'] \iff a \preceq a'$ 对一切 $a, a' \in A$ 成立, 而且商映射 $q : A \to A/\sim$ 严格保序.

(iii) 对任意偏序集 $(B, \preceq)$, 明确地写下一一对应

$$\{f : A \to B \,|\, \text{保序映射}\} \overset{1:1}{\longleftrightarrow} \{\overline{f} : (A/\sim) \to B \,|\, \text{保序映射}\}.$$

提示 对应到 $\overline{f}$ 的映射是 $f := \overline{f} \circ q$.

9. 补全 §2.6 中关于下述性质的验证.

(i) 关于 $\mathbb{Z}$ 对加法和乘法的性质 (从加法结合律直到分配律).

(ii) 关于 $\mathbb{Z}$ 上的全序 $\leq$ 在加法和乘法之下的性质 (2.6.2).

10. 设 $(A,\preceq)$ 和 $(B,\preceq)$ 为偏序集. 在 $A\times B$ 上定义二元关系 $\preceq$ 如下: $(a,b)\preceq(a',b')$ 当且仅当

$$a\prec a',\quad \text{或 } (a=a')\wedge(b\preceq b').$$

(a) 说明这给出 $A\times B$ 上的偏序, 称为**字典序**.

(b) 证明若 A 和 B 是全序集, 则 $A\times B$ 亦然.

(c) 证明若 A 和 B 是良序集, 则 $A\times B$ 亦然.

11. 设 $x\in\mathbb{R}$ 而 $a,b\in\mathbb{Z}_{\geq 1}$. 证明 $\left\lfloor\frac{x}{ab}\right\rfloor=\left\lfloor\frac{\lfloor x/a\rfloor}{b}\right\rfloor$. 提示〉 以 ab 为除数作带余除法.

12. 设 p 为素数. 对所有非零整数 m 取唯一的 $v_p(m)\in\mathbb{Z}_{\geq 0}$ 使得 $p^{v_p(m)}\parallel m$ (约定 2.7.7).

(i) 设 $n\in\mathbb{Z}_{\geq 0}$. 证明 $v_p(n!)=\sum_{k=1}^{\infty}\left\lfloor\dfrac{n}{p^k}\right\rfloor$.

提示〉 将 $n!$ 写成 $p(2p)\cdots(\lfloor\frac{n}{p}\rfloor p)$ 乘以一个与 p 互素的数, 然后继续操作.

(ii) 作展开 $n=a_0+a_1p+\cdots+a_rp^r$, 其中 $0\leq a_i<p$. 基于 (i), 证明

$$v_p(n!)=\frac{n-\sum_{i=0}^{r}a_i}{p-1}.$$

13. 修改 Euclid 定理 2.7.11 的证明以推导

(i) 形如 $4n-1$ 的素数有无穷多个;

(ii) 形如 $6n-1$ 的素数有无穷多个.

提示〉 给定素数 p, 对 (i) 考虑 $(2^2357\cdots p)-1$, 对 (ii) 考虑 $(2357\cdots p)-1$, 讨论其素因子.

14. (Möbius 反演公式) 形如 $f:\mathbb{Z}_{\geq 1}\to\mathbb{C}$ 的映射称为算术函数; 如果 f 进一步满足 $\gcd(a,b)=1\implies f(ab)=f(a)f(b)$, 则称 f 为乘性算术函数. 对算术函数 f 和 g 定义 $f\star g:\mathbb{Z}_{\geq 1}\to\mathbb{C}$ 为

$$(f\star g)(n)=\sum_{d\mid n}f(d)g(n/d).$$

(i) 验证交换律 $f\star g=g\star f$ 与结合律 $(f\star g)\star h=f\star(g\star h)$.

(ii) 验证若 f 和 g 是乘性的, 则 $f\star g$ 亦然.

(iii) 定义算术函数 δ 使得 $\delta(1)=1$, 而 $n>1\implies\delta(n)=0$. 证明 δ 是乘性的, 而且 $\delta\star f=f$ 对所有 f 成立.

(iv) 考虑练习 2.8.9 介绍的 Möbius 函数 μ, 它是乘性的. 证明若 $g=\mathbf{1}\star f$ 则 $f=\mu\star g$, 其中 $\mathbf{1}$ 代表常值函数 1.

提示〉 将问题归结为证 $\mu\star\mathbf{1}=\delta$, 然后用乘性来简化.

Möbius 反演公式在局部有限偏序集中有所推广, 可参阅 [10, §5.4].

15. 对任意集合 A, 证明基数的等式

$$\aleph_0+|A|=\begin{cases}|A|, & A\text{ 无穷},\\ \aleph_0, & A\text{ 有限}.\end{cases}$$

16. 对于熟悉数学分析的读者, 请尝试证明区间 $[0,1]$ 的基数是 $2^{\aleph_0}$.

17. 对任意集合 S 定义

$$C_S := \{\text{映射 } f : S \to \{0,1\} \mid f^{-1}(1) \text{ 有限}\}.$$

试证当 S 有限时 $|C_S| = 2^{|S|}$, 当 S 无穷时 $|C_S| = |S|$. 〈提示〉 用注记 2.9.9 的性质.

第三章 环, 域和多项式

在回溯代数学的渊源时, 我们已经看到 “数” 的四则运算构成经典代数问题的舞台. 受此启发, 理应对 “数” 及其操作进行抽象. 所谓的环, 可理解为配备加法, 减法和乘法运算的集合 (例如整数环 $\mathbb{Z}$), 而域则是乘法有交换律, 包含至少两个元素, 而且当分母非零时能作除法的环 (例如有理数域 $\mathbb{Q}$). 譬如线性方程组的 Gauss–Jordan 消元法便只涉及四则运算, 因而能在任意域上表述并操作.

基于集合的语言, 环和域结构都有简明而严谨的公理化表述. 环 R 有乘法幺元 (或加法零元) 1_R (或 0_R), 简记为 1 (或 0). 从环 R 到环 R' 的映射 $f : R \to R'$ 若满足

$$f(x+y) = f(x) + f(y), \quad f(xy) = f(x)f(y), \quad f(1_R) = 1_{R'},$$

则称 f 为环同态; 兼为双射的环同态称为环同构. 相互同构的环本质上是相同的, 因为它们的元素及运算在双射之下相互匹配. 代数结构和其间的同态与同构是代数学中的基础概念, 而环同构不过是本书处理的第一则例子. 详细解说请见 §§3.1–3.2.

环的定义中仅要求加法交换, 乘法则未必; 乘法满足交换律的环称为交换环. 随着我们以后对线性方程组的理解不断深化, 非交换环将在关于矩阵与线性映射的讨论中自然地出现.

本章的另一个主题是 §§3.3–3.4 探讨的多项式. 一元多项式可以约略地理解为形如 $f = a_0 + a_1X + \cdots + a_nX^n$ 的符号, 其中 $n \in \mathbb{Z}_{\geq 0}$, X 是变元, 而 $a_0, \ldots, a_n$ 是一列系数; 这些系数在初等数学中默认为复数, 然而既然已有环和域的概念, 将系数取在任意域乃至于交换环中都是合理的. 一旦写下严格定义, 交换环 R 上的全体一元多项式对多项式的加法与乘法运算便成为环 $R[X]$; 多元的情形 $R[X, Y, \ldots]$ 当然也类似. 只要读者承认系数在 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 上的多项式是自然的数学对象, 则交换环 $\mathbb{Q}[X], \mathbb{R}[X], \mathbb{C}[X]$ 也同等地自然且必要.

对多项式代值给出多项式函数. 初等数学中的另一类自然对象是域 F 上的有理函数, 它们表作分式 $\frac{f}{g}$, 其中 $f \in F[X]$ 而 $g \in F[X]$, $g \neq 0$. 如果将多项式抽象地理解为一列系数, 而非对应的多项式函数, 则有理函数 $\frac{f}{g}$ 便不能也不必按字面理解为函数, 而应理解为由一对多项式 (f, g) 所表达的分式, 前提是 $g \neq 0$, 而且须将满足通分关系 $f_1g_2 = f_2g_1$ 的对 (f_1, g_1) 和 (f_2, g_2) 等量齐观. 这点将在 §3.5 通过等价关系与商集的语言来表述. 全体 $\frac{f}{g}$ 构成一元有理函数域 $F(X) = \mathrm{Frac}(F[X])$. 值得点出的是称为分式域的这一构造不仅适用于多项式环, 还适用于称为整环 (定义 3.1.11) 的一大类交换

环; 从 $\mathbb{Z}$ 到 $\mathbb{Q} = \mathrm{Frac}(\mathbb{Z})$ 的过渡不过是分式域的初步实例.

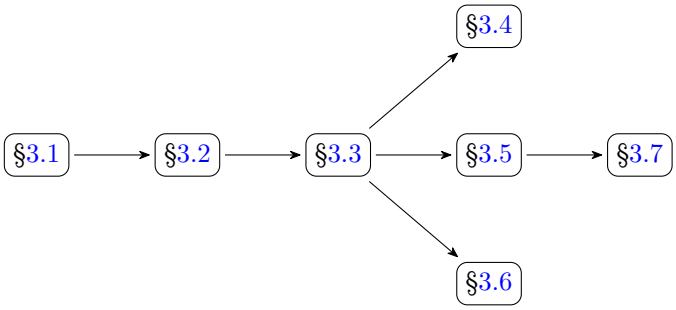
多项式与多项式函数在一般的域上需要严格区别; §3.6 将探讨两者的差异. 举例来说, 对所有素数 p, 整数的 $\bmod p$ 同余关系引出有限域 $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ (例 3.1.12), 而其上的多项式 $X^p - X$ 恒取零值 (例 3.3.5). 命题 3.6.2 将说明在包括无穷域在内的无穷整环上不会发生这种现象.

诸如 $\mathbb{F}_p$ 的有限域极为有用, 它们在基础数学与通信, 密码等应用领域中频繁出现; $\mathbb{F}_p$ 与 $\mathbb{C}$ 或其子域的一个重要差别是任何元素 $x \in \mathbb{F}_p$ 皆满足

$$\underbrace{x + \cdots + x}_{p \text{ 项}} = 0_{\mathbb{F}_p}.$$

由此便引出了关于域 F 的特征 $\mathrm{char}(F)$ 的重要概念, 这是 §3.7 将讨论的主题.

阅读顺序



3.1 环和域

自从 20 世纪以降, 数学家的目光开始从具体问题转向抽象集合上的抽象运算. 集合, 运算连同这些运算具备的基本性质, 一道构成了代数学中所谓的 “结构”. 大而化之地说, 非空集 S 上的 n 元运算 ($n \in \mathbb{Z}_{\geq 1}$) 无非是指一个映射 $S^n \to S$; 譬如加法 $+$ 和乘法 $\cdot$ 都是 $\mathbb{Z}$ 上的二元运算. 对于一般的二元运算

$$\star : S \times S \to S,$$

习惯的作法是将 $\star(s_1, s_2)$ 写成 $s_1 \star s_2$. 对于可以理解为某种乘法的运算, 通常以 $\cdot$ 标记; 简写 $s_1 s_2 = s_1 \cdot s_2$ 也是常用的.

我们在 §1.1 的综述中已经体认, 解方程的经典问题应该以具备四则运算的集合为舞台, 熟悉的例子有 $\mathbb{Q}$, $\mathbb{R}$ 和 $\mathbb{C}$. 行将介绍的 “域” 是这类结构的提纯. 我们将采取层层递进的方式, 首先引进具有加法和乘法的结构, 称为环, 要求乘法交换以得到交换环, 最后要求非零元皆可逆来抵达域的概念. 途中风景在应用中也有各自的位置, 譬如不定方

程必须在无除法的 $\mathbb{Z}$ 上求解, 而对于线性方程组的研究, 其中涉及的线性映射或矩阵将给出非交换环的自然例子.

现在进入正题. 环是配备加法 $+$ 和乘法 $\cdot$ 两种二元运算的集合, 而且乘法和加法分别具有相应的幺元 (又称单位元, 加法情形也称零元). 细说如下.

定义 3.1.1 一个**环**是指资料 $(R,+,\cdot,0_R,1_R)$, 其中 R 是集合, $0_R,1_R\in R$, 而 $+:R\times R\to R$ 和 $\cdot:R\times R\to R$ 都是二元运算, 使得以下条件成立.

1. 加法运算满足以下条件:
   - ▷ **结合律** $(x+y)+z=x+(y+z)$;
   - ▷ **零元性质** $x+0_R=x=0_R+x$.
   - ▷ **交换律** $x+y=y+x$.
   - ▷ **加法逆元** 对所有 x 皆存在 $-x$ 使得 $x+(-x)=0_R$.
2. 乘法运算 $x\cdot y$ 也简写为 xy, 它满足以下条件:
   - ▷ **结合律** $(xy)z=x(yz)$;
   - ▷ **幺元性质** $x\cdot 1_R=x=1_R\cdot x$;
3. 乘法对加法满足
   - ▷ **分配律** $(x+y)z=xz+yz,\quad z(x+y)=zx+zy$.

其中 x,y,z 代表 R 中的任意元素. 不致混淆时, 我们也把 0_R, 1_R 简记为 0, 1, 并以 R 总括资料 $(R,+,\cdot,0_R,1_R)$. 我们也将 $x+(-y)$ 写作 $x-y$.

以下介绍的几条运算性质都是定义的简单结论.

⋆ 结合律确保任意有限多个元素的加法和乘法可以不带括号地写作 $x+y+z$, xyz 等.

⋆ 分配律具有双边的版本:

$$a(x+y)b=(ax+ay)b=axb+ayb.$$

⋆ 加法和乘法幺元都由各自的幺元性质唯一确定. 何以故? 设 0_R 和 $0'_R$ 皆满足加法幺元性质, 1_R 和 $1'_R$ 皆满足乘法幺元性质, 则

$$0_R=0_R+0'_R=0'_R,\quad 1_R=1_R\cdot 1'_R=1'_R.$$

所以环的资料 $(R,+,\cdot,0_R,1_R)$ 中的 0_R 和 1_R 可以略去, 要求对加法或乘法都存在满足幺元性质的元素即可.

⋆ 加法满足消去律: 若 $x+y=x'+y$, 等式两边同加 $-y$, 应用加法结合律遂得 $x=x+y+(-y)=x'+y+(-y)=x'$.

- ⋆ 任何 x 的加法逆元 $-x$ 皆唯一, 这是因为若 $x+x'=0=x+x''$, 则加法消去律蕴涵 $x'=x''$. 因此取加法逆元 $x\mapsto -x$ 也可以视为 R 上的一元运算.
- ⋆ 从加法逆元的唯一性和 $x+(-x)=0=(-x)+x$ 立见 $-(-x)=x$.
- ⋆ 恒等式 $x\cdot 0=0=0\cdot x$ 成立. 以第一个等号为例, 缘由是 $x\cdot 0=x\cdot(0+0)=x\cdot 0+x\cdot 0$, 对两端应用消去律可得 $x\cdot 0=0$.
- ⋆ 恒等式 $(-x)y=-xy=x(-y)$ 成立, 这是缘于

  $$(-x)y+xy=(-x+x)y=0\cdot y=0,\quad x(-y)+xy=x(-y+y)=x\cdot 0=0$$

  和加法逆元的唯一性.
- ⋆ 作为上式的应用, 我们有 $(-1)\cdot y=-y$ 和 $-x=x\cdot(-1)$; 特别地, 代入 $x=-1$ 给出 $(-1)\cdot(-1)=1$.

注记 3.1.2 最平凡的环是**零环**: 这是只有单个元素 $1=0$ 的环, 从环论观点看是无趣的. 另一方面, 非零环必然满足 $1\neq 0$, 否则任何 x 都满足 $x=x\cdot 1=x\cdot 0=0$.

对于任意 $n\in\mathbb{Z}_{\geq 0}$ 和 $r\in R$, 我们引入自明的写法

$$\begin{gathered} n\cdot r=nr:=\underbrace{r+\cdots+r}_{n\text{ 项}},\quad n\geq 1,\\ 0\cdot r:=0,\quad (-n)\cdot r=(-n)r:=-(n\cdot r)\end{gathered}\tag{3.1.1}$$

请读者简单地自我说服

$$\begin{gathered} n(r+r')=nr+nr',\quad (n+m)r=nr+mr,\\ (nm)r=n(mr),\quad (nr)r'=n(rr'),\\ r(n\cdot 1_R)=nr=(n\cdot 1_R)r\end{gathered}\tag{3.1.2}$$

对所有 $n,m\in\mathbb{Z}$ 和 $r,r'\in R$ 皆成立. 以最后一个等号为例, 当 $n\geq 0$ 时 nr 等于

$$\underbrace{r+\cdots+r}_{n\text{ 项}}=1_R\cdot r+\cdots+1_R\cdot r\overset{\text{分配律}}{=\!=}\underbrace{(1_R+\cdots+1_R)}_{n\text{ 项}}r=(n\cdot 1_R)r,$$

而 $n<0$ 时则可通过取加法逆元来处理; 其他几条等式的论证同样简单. 这些运算规律表明 $\mathbb{Z}$ 虽然未必能视同环 R 的子集, (3.1.1) 的写法并不会导致混淆.

对于带有二元运算 $\star$ 的非空集 S 及其子集 S', 如果对所有 $s_1,s_2\in S'$ 都有 $s_1\star s_2\in S'$, 则我们顺理成章地说 S' 对运算 $\star$ **封闭**, 对于一般的 n 元运算当然也有类似的说法. 封闭性可以用来定义代数结构的子结构, 以下仍以环为例.

定义 3.1.3 如果 R 的子集 R_0 包含 $0_R,1_R$, 而且在加法, 乘法运算和加法取逆 $x\mapsto -x$ 之下封闭, 则 $(R_0,+,\cdot,0_R,1_R)$ 也是环, 称为 R 的**子环**.

例 3.1.4 环 R 的**中心**定义为

$$Z(R) := \{z \in R : \forall x \in R,\ zx = xz\}.$$

容易看出 $Z(R)$ 是 R 的子环.

定义 3.1.5 设 x 是环 R 的元素. 若存在 $y \in R$ 使得 $xy = 1$ (或 $yx = 1$), 则称 y 为 x 的右逆 (或左逆), 而 x 右可逆 (或左可逆). 若 x 左右皆可逆, 则称 x **可逆**. 由 R 的可逆元构成的子集记为 $R^\times$.

引理 3.1.6 如果环 R 的元素 x 可逆, 则 x 的左逆也必然是右逆, 而且存在唯一的 $x^{-1} \in R$ 使得 $x^{-1}x = 1 = xx^{-1}$; 此时 $(x^{-1})^{-1} = x$.

证明 论证和关于逆映射的版本 (见定义 2.2.7 之后的讨论) 如出一辙, 以幺元 1 代替该处的 id 便是, 不必重复. □

注意到 $R^\times$ 包含 1 (显然 $1^{-1} = 1$), 而且对乘法运算封闭: 从 $y^{-1}x^{-1}xy = 1 = xyy^{-1}x^{-1}$ 可得

$$(xy)^{-1} = y^{-1}x^{-1}, \quad x, y \in R^\times.$$

进一步, 性质 $(x^{-1})^{-1} = x$ 说明 $R^\times$ 对取逆运算 $x \mapsto x^{-1}$ 也封闭.

对于环中的元素 $r \in R$ 及其 $n \in \mathbb{Z}_{\geq 1}$, 我们记

$$r^n = \underbrace{r \cdots r}_{n \text{ 项}},$$

此外 $r^0 := 1$. 若 $r \in R^\times$, 则进一步记

$$r^{-n} := (r^n)^{-1} = (r^{-1})^n, \quad n \in \mathbb{Z}_{\geq 1}.$$

我们总有等式 $r^{m+n} = r^m r^n$; 当 r 可逆时, 此式对 m 或 n 为负的情形同样成立. 同理, $r^{mn} = (r^m)^n$.

练习 3.1.7 设 $u \in R^\times$, 证明 $r \in R$ 可逆等价于 ur 可逆, 也等价于 ru 可逆. 具体写下这些逆元之间的关系.

提示〉 我们有 $r = u^{-1}(ur) = (ru)u^{-1}$.

定义 3.1.8 如果环 R 的乘法满足交换律 $xy = yx$, 则称 R 为**交换环**.

因此 R 是交换环当且仅当 $Z(R) = R$.

现在万事俱备, 可以将域定义为能作除法的交换环, 前提是除数非零. 如果不要求乘法交换, 得到的概念则称为除环.

定义 3.1.9 满足 $R^\times = R \smallsetminus \{0\}$ (换言之: 零不可逆, 而非零元皆可逆) 的环称为**除环**. 交换除环称为**域**. 域的子环如果也构成域, 则称之为**子域**.

注意: 除环定义中关于零不可逆的条件是为了排除注记 3.1.2 的零环.

域是本书最常谈及的代数结构之一, 另一方面, 确实存在非交换的除环, 而且它们在许多研究中自然地出现, 最突出的例子是后续章节将介绍的四元数除环.

由于域的乘法顺序可换, 在域中可以合理地将 xy^{-1} 写作 x/y 或 $\frac{x}{y}$, 前提是 $y \neq 0$.

例 3.1.10 相对于寻常的乘法和加法运算, $\mathbb{C}$ 是域, 而 $\mathbb{R}$, $\mathbb{Q}$ 都是 $\mathbb{C}$ 的子域, 而子环 $\mathbb{Z}$ 不是域; 事实上 $\mathbb{Z}^\times = \{\pm 1\}$.

比域更宽松的概念是整环, 它以整数环 $\mathbb{Z}$ 为模板.

定义 3.1.11 非零交换环 R 若满足 $x, y \neq 0 \implies xy \neq 0$, 则称为**整环**.

整环的子环显然也是整环. 在整环中乘法对所有非零元都有消去律, 这是因为 $x \neq 0$ 和 $xy = xz$ 蕴涵 $x(y-z) = 0$, 因而蕴涵 $y = z$. 域自动是整环, 这是因为 $x \neq 0$ 和 $xy = 0$ 给出 $y = x^{-1}xy = x^{-1} \cdot 0 = 0$.

以上举出的域都是 $\mathbb{C}$ 的子域[1)]. 另一方面, 在数学其他领域和线性代数的应用中将不可避免地遇到**有限域**. 下一则例子将包含最简单的一类有限域, 其元素个数为素数.

例 3.1.12 (同余类构成的环) 设 $N \in \mathbb{Z}$. 定义 2.8.2 考察了 mod N 同余类构成的集合 $\mathbb{Z}/N\mathbb{Z}$, 或简记为 $\mathbb{Z}/N$. 以 $[x] \in \mathbb{Z}/N\mathbb{Z}$ 代表含 $x \in \mathbb{Z}$ 的同余类, 必要时也表作 $[x]_N$; 在 $\mathbb{Z}/N\mathbb{Z}$ 上定义加法和乘法运算如下

$$[x][y] := [xy], \quad [x] + [y] := [x+y],$$

其中 $x, y \in \mathbb{Z}$. 运算是良定义的, 也就是说运算产物仅依赖同余类 $[x]$ 和 $[y]$ 而不是 x 和 y 的具体取法, 这是练习 2.8.4 的内容. 取 $0_{\mathbb{Z}/N\mathbb{Z}} := [0]$, $1_{\mathbb{Z}/N\mathbb{Z}} := [1]$, 立见 $\mathbb{Z}/N\mathbb{Z}$ 对此运算成为交换环. 注意到 $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$ 而 $\mathbb{Z}/(-N)\mathbb{Z} = \mathbb{Z}/N\mathbb{Z}$, 因此以下不妨设 $N \in \mathbb{Z}_{\geq 1}$, 此时 $\mathbb{Z}/N\mathbb{Z}$ 恰有 N 个元素; 它是零环当且仅当 $N = 1$.

注意到 $[x] \in (\mathbb{Z}/N\mathbb{Z})^\times$ 相当于说同余式 $xy \equiv 1 \pmod{N}$ 有解 $y \in \mathbb{Z}$. 基于命题 2.7.3, 此式有解等价于 x 和 N 互素; 换言之,

$$(\mathbb{Z}/N\mathbb{Z})^\times = \{[x] : x \in \mathbb{Z},\ x, N \text{ 互素}\};$$

基于 Euler 函数 φ 的定义 2.8.7 及其后的讨论, 由此就得出 $|(\mathbb{Z}/N\mathbb{Z})^\times| = \varphi(N)$. 作为推论,

$$\mathbb{Z}/N\mathbb{Z} \text{ 为域} \iff \varphi(N) = N-1 \overset{\text{显然}}{\iff} N \text{ 为素数}.$$

本章习题将说明 $\mathbb{Z}/N\mathbb{Z}$ 为整环当且仅当它是域.

设 p 为素数. 域 $\mathbb{Z}/p\mathbb{Z}$ 是有限域的初步例子. 鉴于它的重要性, 我们另外引入符号

$$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}.$$

[1)] 一些教材将 $\mathbb{C}$ 的子域称为 "数域", 本书不从.

例 3.1.13 (环的直积) 取一族环 $(R_i)_{i\in I}$, 下标 i 遍历某个非空集 I. 我们在 §2.3 已经说明如何定义它们作为集合的积 $\prod_{i\in I} R_i$, 其成员表作元素组 $r=(r_i)_{i\in I}$, 其中 $r_i\in R_i$. 以下介绍如何赋之以自然的环结构, 称作 $(R_i)_{i\in I}$ 的**直积**, 这是从已有的环构造新环的众多途径之一.

在 $\prod_{i\in I} R_i$ 上逐分量地定义加法和乘法, 分别写作

$$\underbrace{(r_i)_i+(r'_i)_i}_{\prod_i R_i \text{ 的加法}} := (\underbrace{r_i+r'_i}_{R_i \text{ 的加法}})_{i\in I},\quad \underbrace{(r_i)_i\cdot(r'_i)_i}_{\prod_i R_i \text{ 的乘法}} := (\underbrace{r_i\cdot r'_i}_{R_i \text{ 的乘法}})_{i\in I}.$$

定义零元 0 为 $(0_i)_i$, 幺元 1 为 $(1_i)_i$, 下标 i 代表它们分别是 R_i 中的零元和幺元. 环论的公理都能化到每个 R_i 上来检验, 兹以加法结合律为例:

$$\begin{aligned}((r_i)_i+(r'_i)_i)+(r''_i)_i &= ((r_i+r'_i)+r''_i)_i\\ &= (r_i+(r'_i+r''_i))_i = (r_i)_i+((r'_i)_i+(r''_i)_i),\end{aligned}$$

其他情形也是类似的. 容易看出 $-(r_i)_i=(-r_i)_i$. 若 $I=\{1,\dots,n\}$, 对应的直积也写作 $R_1\times\cdots\times R_n$ 的形式.

接着考虑每个 R_i 都是同一个环 R 的特例, 这时 $\prod_{i\in I} R_i$ 化为映射集 $R^I=\{f: I\to R\}$ 相对于逐点或逐元素的运算

$$(f+g)(i):=f(i)+g(i),\quad (fg)(i):=f(i)g(i),\quad i\in I$$

所成的环, 方式是让 f 对应 $(f(i))_i\in\prod_{i\in I} R$; 特别地, 0_{R^I} 是常值映射 $i\mapsto 0_R$, 而 1_{R^I} 是常值映射 $i\mapsto 1_R$. 对于 $R=\mathbb{C}$ 或 $\mathbb{R}$ 的情形, 函数之间的逐点运算是数学分析中熟悉的主题.

迄今考虑的环主要是交换环, 非交换环的例子将在第四章自然地引入.

练习 3.1.14 设 $D\in\mathbb{Z}$ 是无平方因子的非零整数, $D\neq 1$.

(i) 验证
$$\mathbb{Q}\left(\sqrt{D}\right):=\left\{a+b\sqrt{D}\in\mathbb{C}: a,b\in\mathbb{Q}\right\}$$
是 $\mathbb{C}$ 的子域, 称为**二次域**. 具体说明非零元的求逆公式.

(ii) 说明 $\mathbb{Q}\left(\sqrt{D}\right)$ 的元素都能唯一地表作 $a+b\sqrt{D}$ 的形式.

(iii) 验证 $\mathbb{Z}[\sqrt{D}]:=\left\{a+b\sqrt{D}: a,b\in\mathbb{Z}\right\}$ 是整环. 验证当 $D\equiv 1 \pmod 4$ 时, 它包含于更大的整环
$$O_D:=\left\{\left.\frac{x+y\sqrt{D}}{2}\,\right|\, \begin{array}{l} x,y\in\mathbb{Z}\\ x\equiv y \pmod 2\end{array}\right\}.$$

3.2 同态和同构

设 R 和 R' 为环. 一如序结构的研究涉及保序映射, 在关于环的研究及其应用中, 我们所关心的映射 $f : R \to R'$ 也不是任意的, 它应该和环结构兼容; 或者更形象地说, f 应该将 R 的环论运算反映在 R' 中. 将这一想法严谨地表述, 便引出环同态的概念.

定义 3.2.1 设 $f : R \to R'$ 为环之间的映射. 当以下条件成立时, 称 f 为**环同态**:

- ⋆ $f(x+y) = f(x) + f(y)$,
- ⋆ $f(xy) = f(x)f(y)$,
- ⋆ $f(1_R) = 1_{R'}$,

其中 x, y 取遍 R 的元素. 从环 R 映到其自身的同态也称为 R 的**自同态**.

在介绍同态的实例之前, 先作几点简单观察:

▷ **保持零元** 性质 $f(0_R) = 0_{R'}$ 不在定义中, 因为这是自动的: 从 $f(0_R) = f(0_R + 0_R) = f(0_R) + f(0_R)$, 配合 R' 中的加法消去律, 即得 $f(0_R) = 0_{R'}$.

▷ **保持加法逆元** 性质 $f(-x) = -f(x)$ 自动成立, 这是 $0_{R'} = f(0_R) = f(x+(-x)) = f(x) + f(-x)$ 的推论.

▷ **保持乘法逆元** 若 $x \in R^\times$, 则 $f(x) \in (R')^\times$ 而 $f(x^{-1}) = f(x)^{-1}$, 这是因为 $1_{R'} = f(1_R) = f(xx^{-1}) = f(x)f(x^{-1})$.

▷ **恒等自同态** 任何环 R 到它自身的恒等映射 id_R 自动是环同态, 这是环同态的平凡例子.

▷ **同态的合成** 若 $f : R \to R'$ 和 $g : R' \to R''$ 为环同态, 则 $gf : R \to R''$ 也是环同态. 这是因为

$$gf(x+y) = g(f(x) + f(y)) = gf(x) + gf(y),$$
$$gf(xy) = g(f(x)f(y)) = gf(x)gf(y), \quad gf(1_R) = g(1_{R'}) = 1_{R''}.$$

▷ **像与子环** 对于环同态 $f : R \to R'$, 它的像 $f(R)$ 自然是 R' 的子环; 反过来说, 给定环 R' 及其子环 $R \subset R'$, 取 $\iota : R \hookrightarrow R'$ 为包含映射, 映 $r \in R$ 为 r, 则 ι 自然是环同态.

例 3.2.2 设 $N, M \in \mathbb{Z}$ 满足 $N \mid M$. 沿用例 3.1.12 关于同余类的符号, 考虑映射

$$p_N^M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$$
$$[x]_M \mapsto [x]_N,$$

这是良定义的: 对于任意 $x, y \in \mathbb{Z}$ 显然有

$$x \equiv y \pmod{M} \overset{\text{定义}}{\iff} M \mid x-y \overset{N\mid M}{\Longrightarrow} N \mid x-y \overset{\text{定义}}{\iff} x \equiv y \pmod{N}.$$

根据例 3.1.12 对环运算的定义, 显然也有

$$\begin{aligned} p_N^M\left([x]_M+[y]_M\right)&=p_N^M\left([x+y]_M\right)=[x+y]_N\\ &=p_N^M([x]_M)+p_N^M([y]_M), \end{aligned}$$

同样按部就班的推导给出

$$p_N^M\left([x]_M[y]_M\right)=p_N^M([x]_M)p_N^M([y]_M);$$

此外 $p_N^M([1]_M)=[1]_N$. 这就说明 p_N^M 是环同态, 其效果是抹掉 mod N 之外的信息.

例 3.2.3 设 $(R_i)_{i\in I}$ 为一族环, 其中 I 是非空集. 取例 3.1.13 所介绍的直积 $\prod_{i\in I}R_i$. 对任意 $i\in I$, 考虑集合的积所带有的投影映射

$$\begin{aligned} p_i:\prod_{j\in I}R_j&\to R_i\\ (r_j)_j&\mapsto r_i. \end{aligned}$$

容易验证这是环同态; 事实上, 例 3.1.13 赋予 $\prod_{i\in I}R_i$ 的是使得每个 p_i 皆为环同态的唯一环结构. 请读者就此仔细思忖, 直到一切都呈现为同义反复为止.

对于每个 R_i 都是同一个环 R 的特例, 投影同态化为求值同态 $\mathrm{ev}_i:R^I\to R$, 映 $f:I\to R$ 为 $f(i)\in R$.

在 §2.9 关于集合大小的讨论中, 我们将元素能一一对应的集合等量齐观. 在 §2.4 关于序结构的研究中, 以同构相对应的两个偏序集也被视为有相同的结构, 或者说它们具有相同形状的 Hasse 图. 此一思路适用于一般的数学结构, 称为同构. 对于眼下的环论情形, 其具体表述如下.

定义 3.2.4 *设 $f:R\to R'$ 为环同态. 如果存在环同态 $g:R'\to R$ 使得 $gf=\mathrm{id}_R$ 而 $fg=\mathrm{id}_{R'}$, 则称 f 为***环同构***, 而 g 为 f 的逆. 此时我们也说 R 和 R' 同构.*

条件 $gf=\mathrm{id}_R$ 和 $fg=\mathrm{id}_{R'}$ 表明 f 的逆无非是 f 作为映射的逆 $g=f^{-1}$. 反过来说, 容易证环同态 f 如果作为映射是双射, 那么它也是环同构.

命题 3.2.5 *设 $f:R\to R'$ 为环同态. 如果 f 是集合之间的双射, 则 f 是环同构.*

证明 问题归结为证 f 的逆映射 f^{-1} 也是环同态. 对 $f(1_R)=1_{R'}$ 两边取 f^{-1} 可得 $1_R=f^{-1}(1_{R'})$. 对 $f(x+y)=f(x)+f(y)$ 两边取 f^{-1}, 并且记 $u=f(x)$, $v=f(y)$, 可得 $f^{-1}(u)+f^{-1}(v)=f^{-1}(u+v)$. 同理可见 $f^{-1}(uv)=f^{-1}(u)f^{-1}(v)$. 由于所有 $u,v\in R'$ 都能表作 $u=f(x)$ 和 $v=f(y)$ 的形式, 综上可见 f^{-1} 确实是环同态. □

恒等映射 id_R 是同构最简单的例子. 此外, 两个同构 f 和 g 的合成 gf 依然是同构, 以 $f^{-1}g^{-1}$ 为逆.

约定 3.2.6 今后以符号 $f:R\xrightarrow{\sim}R'$ 代表映射 $f:R\to R'$ 是环同构; 在不必指明 f 的场合, 我们也以符号 $R\simeq R'$ 代表环 R 和 R' 同构.

类似记法也适用于以后将介绍的其他代数结构.

同构 $f : R \xrightarrow{\sim} R'$ 不但为集合 R 和 R' 建立了双射, 而且对应元素之间的一切环论运算 (加法, 乘法) 和幺元也在 f 之下相配对. 凡是以环论语言表述的一切性质, 对于同构的环 R 和 R' 都是等价的. 这是代数学中的一条基本原理.

练习 3.2.7 设 F 为域, R 为非零环, 而 $\varphi : F \to R$ 为环同态. 说明 φ 为单射.

提示〉 我们有 $\varphi(x) = \varphi(y) \iff \varphi(x-y) = 0$, 所以问题化为证 $x \neq 0 \implies \varphi(x) \neq 0$. 但是域 F 中的任意非零元都是可逆的, 而同态映可逆元为可逆元.

作为命题 3.2.5 的演示, 下面给出中国剩余定理的一种环论表述.

定理 3.2.8 (中国剩余定理 — 同构版本) 设 $N \in \mathbb{Z}_{\geq 1}$ 分解为 $n_1 \cdots n_k$, 其中 $n_1, \ldots, n_k$ 两两互素, 则有环同构

$$\varphi : \mathbb{Z}/N\mathbb{Z} \xrightarrow{\sim} \prod_{i=1}^{k} \mathbb{Z}/n_i\mathbb{Z}$$

$$[x]_N \longmapsto ([x]_{n_i})_{i=1}^{k},$$

其中关于同余类的符号如例 3.1.12.

证明 例 3.2.2 业已说明 $[x]_N \mapsto [x]_{n_i}$ 对所有 i 都给出同态 $\mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/n_i\mathbb{Z}$. 既然直积的环结构是逐分量定义的, φ 必保持环结构, 从而是同态.

此外, 映射两端作为集合都有 N 个元素, 基于抽屉原理 (命题 2.9.4), 证 φ 是单射即可. 互素条件在此派上用场: 设 $x, y \in \mathbb{Z}$ 满足 $\varphi([x]_N) = \varphi([y]_N)$, 则对所有下标 i 都有

$$n_i \mid x - y.$$

既然 $n_1, \ldots, n_k$ 两两互素, 故 $N \mid x - y$, 亦即 $[x]_N = [y]_N$. 单性得证. □

原初的中国剩余定理对应到 φ 的满性, 相当于说同余方程组 $X \equiv a_i \pmod{n_i}$ $(1 \leq i \leq k)$ 对所有数组 $(a_1, \ldots, a_k) \in \mathbb{Z}^k$ 皆有解; 以上仅以抽屉原理抽象地说明 φ 满, 未给出求解的具体算法. 由此观之, 上述证明的精神和宋代学者秦九韶的大衍求一术并不相通.

之后的定理 6.3.8 将给出中国剩余定理在主理想环上的一种推广, 其证明也具有算法的面向.

练习 3.2.9 说明 $\mathbb{Z}/4\mathbb{Z}$ 和 $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ 都是恰有 4 个元素的环, 但是两者并不同构.

提示〉 所有 $x \in \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ 都满足 $2x = 0$, 对 $\mathbb{Z}/4\mathbb{Z}$ 则不然.

练习 3.2.10 设 D_1 和 D_2 为无平方因子的非零整数, $D_1, D_2 \neq 1$. 证明练习 3.1.14 介绍的二次域 $\mathbb{Q}\left(\sqrt{D_1}\right)$ 和 $\mathbb{Q}\left(\sqrt{D_2}\right)$ 同构当且仅当 $D_1 = D_2$.

提示〉 问题在于 "仅当" 的方向. 设 $D \in \mathbb{Z} \smallsetminus \{0, 1\}$ 无平方因子, 具体剖析形如 $X^2 = D$ 的方程何时在 $\mathbb{Q}\left(\sqrt{D_i}\right)$ 中有解, 并且说明此性质只和域的同构类相关.

3.3 多项式环

一元多项式 $a_nX^n+\cdots+a_1X+a_0$ 是大家熟悉的老友. 环在此扮演两种角色.

1. 多项式的系数 $a_n,\ldots,a_0$ 可以容许取在一般的环 R, 而不只是熟悉的整系数, 复系数等. 这是因为在多项式的基本运算中, 起作用的只是系数的加法和乘法; 甚至不必要求乘法有交换律, 尽管交换环情形最为常用.

2. 一旦系数限制在给定的环里, 全体多项式相对于多项式的加, 减, 乘运算也构成一个环.

多项式环具有丰富的代数结构, 本节只论最基本的定义和性质, 多项式的算术则留待之后的章节处理.

定义 3.3.1 *设 R 为非零环. 以 X 为变元, 系数在 R 上的***多项式***定义为形如*

$$f=\sum_{n\geq 0}a_nX^n,\quad a_n\in R,\quad \text{至多有限个 } a_n \text{ 非零}$$

的形式和, $a_n=0$ 的项可以省去; 须凸显变元时也将 f 写作 $f(X)$. 所有这些多项式构成的集合记为 $R[X]$.

所谓形式和, 意指定义中的 $\sum_{n\geq 0}a_nX^n$ 仅被视为一种记法. 原教旨主义的诠释应当是取 $\mathbb{Z}_{\geq 0}$ 份 R 的积 $R^{\mathbb{Z}_{\geq 0}}$, 其元素按分量写作 $(a_n)_{n\geq 0}$ 的形式, 相应地 $R[X]$ 便是其子集

$$\left\{(a_n)_{n\geq 0}\in R^{\mathbb{Z}_{\geq 0}}:\text{仅有至多有限个 } n \text{ 使得 } a_n\neq 0\right\}.$$

如果按此观点, 则 X^n 在写法中 $\sum_{n\geq 0}a_nX^n$ 仅起到记录下标 n 的作用; 稍后定义多项式乘法时, 符号 X^n 的便利性就会凸显.

关于形式和的解释也表明

$$\sum_{n\geq 0}a_nX^n=\sum_{n\geq 0}b_nX^n\iff \forall n\in\mathbb{Z}_{\geq 0},\ a_n=b_n.$$

现在介绍关于多项式的标准术语, 以及几点简单注记. 设 $f=\sum_{n\geq 0}a_nX^n\in R[X]$.

术语	意义	符号
f 的 n 次项系数	系数 a_n	
f 的常数项	系数 a_0	
f 的首项	使系数 a_n 非零的最高次项 a_nX^n	
首一多项式	首项系数为 1 的多项式	
零多项式	系数全为 0 的多项式	0
非零多项式 f 的次数	$\max\{n \geq 0 : a_n \neq 0\}$	$\deg f$
常数多项式	除常数项以外系数均为零的多项式	

- ⋆ 注意到 R 自然地嵌入为 $R[X]$ 的子集, 方式是映 $r \in R$ 为相应的常数多项式, 仍记之为 r.
- ⋆ 当我们说一个多项式 $f \in R[X]$ 非零时, 确切意涵是 $f \in R[X] \smallsetminus \{0\}$, 而不是说它在某一点不取零, 也不是说它处处非零. 关于多项式的求值, 请见 (3.3.1) 和其后的讨论.
- ⋆ 一般不考虑零多项式的次数; 确实有需要时, 定义 $\deg 0 = -\infty$, 仅作为一个方便的符号来理解.

接着赋予 $R[X]$ 环结构. 多项式的加法定为逐项相加

$$\sum_{n\geq 0} a_nX^n + \sum_{n\geq 0} b_nX^n := \sum_{n\geq 0}(a_n + b_n)X^n,$$

乘法则定为

$$\left(\sum_{n\geq 0} a_nX^n\right) \cdot \left(\sum_{n\geq 0} b_nX^n\right) := \sum_{n\geq 0}\left(\sum_{\substack{h,k\geq 0,\\ h+k=n}} a_hb_k\right) X^n.$$

由此立见首一多项式的乘积显然还是首一多项式.

命题 3.3.2 *以上运算使得 $R[X]$ 成为环, 其中 $0_{R[X]}$ 是零多项式, 而 $1_{R[X]}$ 是对应于 1_R 的常数多项式; R 嵌入为 $R[X]$ 的子环. 若 R 是交换环, 则 $R[X]$ 亦然.*

证明 这些无非是定义的直接操演, 比如乘法的交换律归结为 R 的乘法交换律和 $X^i(X^jX^k) = X^{i+j+k} = (X^iX^j)X^k$. □

引理 3.3.3 设 R 为整环 (定义 3.1.11), 则对所有非零的 $f,g\in R[X]$ 都有 $\deg(fg)=\deg f+\deg g$, 此时 $R[X]$ 也是整环.

作为推论, 此时 $R[X]^\times=R^\times$.

证明 设 $f=a_nX^n+$低次项, $g=b_mX^m+$低次项, 其中 $a_n,b_m\neq 0$. 那么 $fg=a_nb_mX^{m+n}+$低次项, 而 $a_nb_m\neq 0$. 这就给出第一部分.

若 $f,g\in R[X]$ 满足 $fg=1$, 则第一部分表明 $\deg f=0=\deg g$, 换言之 f 和 g 可以视同 R 的非零元. 这就给出第二部分. □

举例明之, 整系数多项式构成整环 $\mathbb{Z}[X]$. 本书主要处理域上的多项式环.

同样的构造扩及多变元情形. 指定任何一族变元 $X,Y,\ldots$, 仅作为一族相互独立的符号来理解, 多元多项式环 $R[X,Y,\ldots]$ 的元素写作形式的有限和

$$\sum_{a,b,\ldots} c_{ab\ldots}X^aY^b\cdots,\quad c_{ab\ldots}\in R.$$

加法和乘法则按寻常的方式定义. 形如 $X^aY^b\cdots$ 的项称为**单项式**, 它们在加法和对 R 的乘法之下张成整个 $R[X,Y,\ldots]$.

简言之, $R[X,Y,\ldots]$ 的元素是从变元 $X,Y,\ldots$ 出发, 经过有限步形式地相加, 相乘和乘以 R 的元素 (和变元可交换), 所能得到的所有代数表达式. 数学的行话称变元 $X,Y,\ldots$ 是自由的, 在此按消极意义理解, 也就是说它们除了环论所要求的一般性质外, 不再服从任何额外的代数关系.

注意到我们对变元 X,Y 等有意地采取模棱两可的表法, 这是因为变元的数量容许有无穷个, 甚至不可数; 但因为我们仅容许有限步的运算, 或者说 $f=\sum_{a,b,\ldots}c_{a,b,\ldots}X^aY^b\cdots$ 要求是有限和, 每个 $f\in R[X,Y,\ldots]$ 都只涉及有限多个变元. 对于可数个变元的情形, 对应的多项式环遂表达为子环的渐增并

$$R[X,Y,Z,\ldots]=R[X]\cup R[X,Y]\cup R[X,Y,Z]\cup\cdots,$$
$$R[X]\subset R[X,Y]\subset R[X,Y,Z]\subset\cdots.$$

运算的有限性是代数学的本色之一.

定义 3.3.4 设 $f=\sum_{a_1,\ldots,a_n\geq 0}c_{a_1,\ldots,a_n}X_1^{a_1}\ldots X_n^{a_n}$ 是 $R[X_1,\ldots,X_n]$ 的元素. 若有 $N\in\mathbb{Z}_{\geq 0}$ 使得仅当 $a_1+\cdots+a_n=N$ 时才可能有 $c_{a_1,\ldots,a_n}\neq 0$, 则称 f 是 N 次**齐次**的.

由于无穷多个变元的多项式环可以写成有限变元子环的并, 齐次的概念也可以相应地推广, 兹不赘述.

举例明之, 二元二次齐次多项式是形如 $aX_1^2+bX_1X_2+cX_2^2$ 的多项式.

中学数学所见的多项式多数是作为多项式函数而出现的. 多项式之所以能给出函数, 缘由是可以将具体的数代入变元来求值. 简单起见, 以下设 R 为交换环. 对任何 $x,y,\ldots\in R$, 我们可以将 $X=x$, $Y=y$ 等代入 $f\in R[X,Y,\ldots]$ 进行求值, 给出

$$f=\sum_{a,b,\ldots}c_{ab\ldots}X^aY^b\cdots\mapsto\sum_{a,b,\ldots}c_{ab\ldots}x^ay^b\cdots=:f(x,y,\ldots).\tag{3.3.1}$$

按照多项式的加法和乘法的具体定义, 当下看出

$$\begin{aligned}(f+g)(x,y,\ldots)&=f(x,y,\ldots)+g(x,y,\ldots),\\(fg)(x,y,\ldots)&=f(x,y,\ldots)g(x,y,\ldots),\\(\text{常数多项式 } c)(x,y,\ldots)&=c.\end{aligned}\tag{3.3.2}$$

因此每个多项式 $f\in R[X,Y,\ldots]$ 都确定从 $R\times R\times\cdots$ (乘积项数 = 变元个数) 到 R 的映射, 这是多项式 f 所确定的**多项式函数**.

例 3.3.5 对于一般的交换环 R, 多项式未必由它对应的多项式函数确定. 一个例子是取 $R=\mathbb{F}_p:=\mathbb{Z}/p\mathbb{Z}$, 其中 p 是素数. 根据 Fermat 小定理 2.8.6, 单变元多项式

$$f(X)=X^p-X\in\mathbb{F}_p[X]$$

对所有 $x\in\mathbb{F}_p$ 都满足 $f(x)=0$, 所以尽管 X^p-X 并非零多项式, 它作为多项式函数却无异于零函数. 推而广之, 对于任意有限域 F, 非零多项式 $f(X):=\prod_{a\in F}(X-a)$ 在任何 $a\in F$ 上取值皆为 0.

有鉴于此, 对于一般的交换环, 必须区分作为一个代数表达式的多项式以及相应的函数或映射, 前者才是第一义的. 我们将在 §3.6 说明何时可以等同一个多项式及它所对应的函数.

定义 3.3.6 交换环 R 上的多项式也可以作合成. 给定 $n,m\in\mathbb{Z}_{\geq 1}$. 设有

$$f=\sum_{a_1,\ldots,a_n\geq 0}c_{a_1,\ldots,a_n}X_1^{a_1}\cdots X_n^{a_n}\ \in R[X_1,\ldots,X_n],$$
$$g_1,\ldots,g_n\in R[Y_1,\ldots,Y_m].$$

记 $g:=(g_1,\ldots,g_n)\in R[Y_1,\ldots,Y_m]^n$, 则可以定义

$$f\circ g:=\sum_{a_1,\ldots,a_n\geq 0}c_{a_1,\ldots,a_n}g_1^{a_1}\cdots g_n^{a_n}\ \in R[Y_1,\ldots,Y_m].$$

合成的定义和求值的定义 (3.3.1) 神似, 差别只是代入 $X_1,X_2,\ldots$ 之值未必是 R 的元素, 而属于比 R 更大的环 $R[Y_1,\ldots,Y_m]$. 于是合成可以视作求值的特例来处理, 它也满足类似的一般性质:

$$\begin{aligned}(f_1+f_2)\circ g&=f_1\circ g+f_2\circ g,\\(f_1f_2)\circ g&=(f_1\circ g)(f_2\circ g),\\(\text{常数多项式 } c)\circ g&=c.\end{aligned}$$

按合成的定义直接验证这些等式也毫不困难.

从多项式函数的观点, 合成的意义更加明白: 设 R 为交换环, 则 $g=(g_1,\ldots,g_n)$ 确定以下函数

$$\begin{aligned}\hat{g}:R^m&\longrightarrow R^n\\(y_1,\ldots,y_m)&\longmapsto\left(g_1(y_1,\ldots,y_m),\ldots,g_n(y_1,\ldots,y_m)\right),\end{aligned}$$

区别 g 和 $\hat{g}$ 只是为了避免混淆. 代入定义可见 $f\circ g\in R[Y_1,\dots,Y_m]$ 确定的函数 $R^m\to R$ 无非是合成函数

$$R^m \xrightarrow{\hat{g}} R^n \xrightarrow{\text{由 } f \text{ 确定}} R.$$

我们将在 §3.6 继续关于多项式函数的讨论.

最后, 读者应当知道多元多项式可以按照变元来集项, 例如在 $\mathbb{Q}[X,Y]$ 中

$$\sum_{a,b\geq 0} c_{ab}X^aY^b=\sum_{a\geq 0}\left(\sum_{b\geq 0}c_{ab}Y^b\right)X^a,$$

其中所有和都是有限和. 一旦系数放宽到一般的环, 这种写法就转译为一个简单的同构关系, 它在处理多元多项式环的结构时也是一个方便的技巧.

命题 3.3.7 *对于任意环 R, 存在自然的环同构 $R[X,Y]\simeq (R[X])[Y]$. 推而广之, 对任意 $n\geq 2$ 都有环同构*

$$R[X_1,\dots,X_n]\simeq R[X_1,\dots,X_{n-1}][X_n]\simeq\cdots\simeq R[X_1]\cdots[X_n].$$

证明 今后省略不必要的括号. 定义 $\varphi: R[X,Y]\to R[X][Y]$ 为以下映射

$$\sum_{a,b\geq 0}c_{ab}X^aY^b\mapsto\sum_{b\geq 0}\underbrace{\left(\sum_{a\geq 0}c_{ab}X^a\right)}_{\in R[X]}Y^b.$$

以下来验证 φ 是环同态, 而且作为映射是双射. 它保持加法是明显的, 至于乘法, 我们利用以下小技巧: 假设对 $R[X,Y]$ 的某个子集 $\mathcal{S}$, 已经证明了对任何 $s,t\in\mathcal{S}$ 都有 $\varphi(st)=\varphi(s)\varphi(t)$, 则当 s 和 t 是 $\mathcal{S}$ 中元素的有限和时, 同样有 $\varphi(st)=\varphi(s)\varphi(t)$. 这是乘法分配律和 φ 保持加法的直接结论: 设 $s_i,t_j\in\mathcal{S}$, 其中 i 和 j 分别遍历两个有限集, 则

$$\begin{aligned}\varphi\left(\sum_i s_i\sum_j t_j\right)&=\varphi\left(\sum_{i,j}s_it_j\right)=\sum_{i,j}\varphi(s_it_j)\\&\overset{\text{假设}}{=\!=}\sum_{i,j}\varphi(s_i)\varphi(t_j)=\left(\sum_i\varphi(s_i)\right)\left(\sum_j\varphi(t_j)\right)=\varphi\left(\sum_i s_i\right)\varphi\left(\sum_j t_j\right).\end{aligned}$$

现在取 $\mathcal{S}=\{cX^aY^b: a,b\geq 0,\ c\in R\}$, 那么 $\mathcal{S}$ 中元素的和穷尽 $R[X,Y]$. 另一方面, $\varphi(st)=\varphi(s)\varphi(t)$ 则是简单的: 若 $s=cX^aY^b$ 而 $t=c'X^{a'}Y^{b'}$, 则两边都等于 $(cc'X^{a+a'})Y^{b+b'}\in R[X][Y]$. 这就完成了环同态的验证.

为了说明 φ 是双射, 仅须具体写下逆映射 $\sum_b\left(\sum_a c_{ab}X^a\right)Y^b\mapsto\sum_{a,b}c_{ab}X^aY^b$ 并留意到所有和都是有限的. □

推论 3.3.8 设 R 为整环, 则任意多个变元的多项式环 $R[X,Y,\ldots]$ 也都是整环, 而且 $R[X,Y,\ldots]^\times = R^\times$.

证明 首先验证 $R[X,Y,\ldots]$ 为整环. 单变元 $R[X]$ 的情形无非是引理 3.3.3 的前半部. 有限多个变元 $R[X_1,\ldots,X_n]$ 的情形按照同构 $R[X_1,\ldots,X_n] \simeq R[X_1,\ldots,X_{n-1}][X_n]$ 逐步化约到 $n=1$. 对于无穷多个变元的情形, 回忆到任何给定的 $f,g \in R[X,Y,\ldots]$ 仅涉及有限多个变元, 因此两者的乘法可以在一个有限变元多项式子环中来考量.

至于 $R[X,Y,\ldots]^\times$ 的描述, 目的在说明若 $f,g \in R[X,Y,\ldots]$ 满足 $fg=1$, 则 $f,g \in R^\times$. 一旦 f 和 g 选定, 问题仅涉及有限多个变元, 由此将问题化为对所有 $n \geq 1$ 证

$$f,g \in R[X_1,\ldots,X_n],\ fg=1 \implies f,g \in R^\times.$$

这可以递归地论证. 当 $n=1$ 时, 此即引理 3.3.3 后半部. 当 $n>1$, 由 $R[X_1,\ldots,X_n] \simeq R[X_1,\ldots,X_{n-1}][X_n]$ 可得 $f,g \in R[X_1,\ldots,X_{n-1}]^\times = \cdots = R^\times$. □

3.4 一元多项式的带余除法与根

多项式的带余除法是中学数学的内容, 它是研究多项式理论的起点. 现在我们对一般的域上的情形作抽象的表述; 具体算法和中学情形完全相同, 为了严谨起见, 以下仍给出冗长而毫不困难的论证.

本节的 F 代表某个选定的域.

命题 3.4.1 (多项式的带余除法) 对于任意 $a,d \in F[X]$, 若 $d \neq 0$, 则存在唯一的 $q,r \in F[X]$ 使得 $\deg(r) < \deg(d)$ 而且 $a = dq + r$; 此处定义 $\deg(0) := -\infty$.

证明 考虑非空集 $\{a - dq : q \in F[X]\}$, 其中必有元素 $a-dq$ 使得 $\deg(a-dq)$ 极小 (容许为 $-\infty$); 以下说明 $r := a - dq$ 必满足 $\deg(r) < \deg(d)$, 这将给出 q 和 r 的存在性. 为此, 不妨假设 $r \neq 0$, 将这些多项式写作

$$r = \alpha_n X^n + \text{低次项}, \quad d = \beta_m X^m + \text{低次项}, \quad \alpha_n, \beta_m \neq 0.$$

假若 $\deg(r) \geq \deg(d)$, 亦即 $n \geq m$, 则

$$\deg\left(r - \frac{\alpha_n}{\beta_m} X^{n-m} d\right) < n = \deg(r),$$

从而 $r - \frac{\alpha_n}{\beta_m} X^{n-m} d = a - d\left(q + \frac{\alpha_n}{\beta_m} X^{n-m}\right)$ 的次数比 $a-dq$ 更低, 矛盾.

至于唯一性, 设 $dq_1 + r_1 = dq_2 + r_2$, 其中 $\deg(r_1), \deg(r_2) < \deg(d)$, 则比较等式

$$d(q_1 - q_2) = r_1 - r_2$$

两边的次数, 可知它成立的唯一可能是 $q_1 = q_2$ 而 $r_1 = r_2$. □

设 $a,d\in F[X]$. 如果存在 $q\in F[X]$ 使得 $a=dq$, 则以整除符号记为 $d\mid a$. 作为立即的推论, $d\mid a$ 的充要条件是带余除法中的余式 r 为 0.

练习 3.4.2 考虑一般的整环 R 和其上的多项式 $a,d\in R[X]$, 其中 $d\neq 0$. 说明若 d 的最高次项系数 (之前记为 β_m) 属于 $R^\times$, 则命题 3.4.1 的陈述仍成立.

定义 3.4.3 *若 $f\in F[X]$ 而 $a\in F$ 满足 $f(a)=0$, 则称 a 为 f 的***根**.

推而广之, 若 R 是交换环, $f\in R[X]$, $a\in R$ 而 $f(a)=0$, 则也称 a 是 f 的根, 或者更精确地称为它在 R 中的根.

关于根的基本观察是被称为**余式定理**的以下结果:

> 以带余除法将 f 表成 $(X-a)q+r$, 其中 r 是满足 $\deg r<1$ 的余式,
> 则 $r=f(a)$.

当然, 这无非是对 $f=(X-a)q+r$ 代入 $X=a$ 求值的结论. 由此推得 $f(a)=0$ 等价于 $(X-a)\mid f$.

命题 3.4.4 *设 $f\in F[X]\smallsetminus\{0\}$, 则 f 在 F 中至多只有 $\deg f$ 个相异的根.*

证明 对 $n:=\deg f$ 递归地论证; $n=0$ 情形不必多言. 设 $n\geq 1$. 若 f 有相异的根 $a_1,\dots,a_{n+1}\in F$, 则余式定理给出 $g\in F[X]\smallsetminus\{0\}$ 使得 $f=(X-a_{n+1})g$. 由于对每个 $1\leq i\leq n$ 皆有 $a_i-a_{n+1}\in F^\times$, 而

$$g(a_i)=\frac{f(a_i)}{a_i-a_{n+1}}=0,$$

故 g 有 n 个相异根, 与 $\deg g=n-1$ 矛盾. □

我们将在 §6.6 对多项式的根作更精细的讨论.

3.5 从整环的分式域到有理函数域

首先考虑经典的实系数情形. 在 $\mathbb{R}$ 上, 有理函数按定义是多项式的商

$$\frac{f}{g},\quad f,g\in\mathbb{R}[X],$$

其中要求 g 不是零多项式, 或者更简洁地说, g 不是环 $\mathbb{R}[X]$ 的零元. 这类函数在数学分析课程中作为习题成批地出现. 视为函数, 它在 $\mathbb{R}$ 上只是部分地定义的, 因为分母的零点必须排除. 全体实系数有理函数构成集合, 记为 $\mathbb{R}(X)$, 其中

$$\frac{f_1}{g_1}=\frac{f_1}{g_1}\text{ (作为有理函数)}\iff f_1g_2=f_2g_1\text{ (作为多项式)}.$$

集合 $\mathbb{R}[X]$ 上具有熟悉的四则运算, 如

$$\frac{f_1}{g_1}+\frac{f_2}{g_2}=\frac{f_1g_2+f_2g_1}{g_1g_2},\quad \frac{f_1}{g_1}\cdot\frac{f_2}{g_2}=\frac{f_1f_2}{g_1g_2},$$

服从于结合律, 交换律, 分配律等性质. 此外, 任何不恒为零的有理函数 $\frac{f}{g}$ 必满足 $f \neq 0$, 故有乘法逆 $\frac{g}{f}$. 换言之, $\mathbb{R}(X)$ 是域. 尽管四则运算可能改变有理函数的定义域, 但这不影响 $\mathbb{R}(X)$ 的上述代数性质.

我们希望将此拓展到更一般的情形. 由于在一般的域上, 非零多项式可能导致恒取零值的多项式函数 (例 3.3.5), 真正合适的观点是将有理函数视为多项式的 "形式商", 而多项式则如在 §3.3 一般视为形式和. 这样虽然摆脱了函数的桎梏, 却出现了新问题: 不同的分子和分母可以导致相同的商, 如何界定两个分式相等? 以下提供的方法和从 $\mathbb{Z}$ 造 $\mathbb{Q}$ 的手法完全相同, 它不仅适用于多项式, 还可以施于定义 3.1.11 所谓的整环.

设 R 为整环. 我们考虑集合

$$\mathrm{Ratio}(R) := \left\{(f, g) \in R^2 : g \neq 0\right\}.$$

在 $\mathrm{Ratio}(R)$ 上定义二元关系

$$(f_1, g_1) \sim (f_2, g_2) \iff f_1 g_2 = f_2 g_1.$$

以下验证这是等价关系. 唯一不显然的只有传递性: 设 $(f_1, g_1) \sim (f_2, g_2)$ 而 $(f_2, g_2) \sim (f_3, g_3)$, 则 R 的交换性导致

$$(f_1 g_2) g_3 = (f_2 g_1) g_3 = (f_2 g_3) g_1 = (f_3 g_2) g_1.$$

因为 R 是整环, 两边消去非零元 g_2 便得到 $f_1 g_3 = f_3 g_1$, 亦即 $(f_1, g_1) \sim (f_3, g_3)$.

定义商集 $\mathrm{Frac}(R) := \mathrm{Ratio}(R)/\sim$. 记 (f, g) 的等价类为 $[f, g] \in \mathrm{Frac}(R)$, 可设想为欲构造的分式 f/g. 显而易见,

$$[f, g] = [fh, gh], \quad h \in R \smallsetminus \{0\}. \tag{3.5.1}$$

接着来赋予 $\mathrm{Frac}(R)$ 环结构.

▷ **加法** 定义 $[f_1, g_1] + [f_2, g_2] = [f_1 g_2 + f_2 g_1, g_1 g_2]$. 注意到 $g_1 g_2 \neq 0$. 此定义只依赖 (f_i, g_i) 的等价类 $(i = 1, 2)$. 比方说若 $(f_1, g_1) \sim (\tilde{f}_1, \tilde{g}_1)$, 则有

$$\begin{aligned}(f_1 g_2 + f_2 g_1)\tilde{g}_1 g_2 &= f_1 \tilde{g}_1 g_2^2 + f_2 g_1 \tilde{g}_1 g_2 \\ &= \tilde{f}_1 g_1 g_2^2 + f_2 g_1 \tilde{g}_1 g_2 = (\tilde{f}_1 g_2 + f_2 \tilde{g}_1) g_1 g_2,\end{aligned}$$

故 $[\tilde{f}_1 g_2 + f_2 \tilde{g}_1, \tilde{g}_1 g_2] = [f_1 g_2 + f_2 g_1, g_1 g_2]$; 对于 $(f_2, g_2) \sim (\tilde{f}_2, \tilde{g}_2)$ 的情形也是类似料理.

进一步, 作加法时总可以按照 (3.5.1) 化约到 $g_1 = g_2$ 的情形, 这时加法公式化简为 $[f_1, g] + [f_2, g] = [f_1 + f_2, g]$.

▷ **乘法** 定义 $[f_1, g_1][f_2, g_2] = [f_1 f_2, g_1 g_2]$, 同样容易检查这是良定义的.

▷ **加法零元**　定义 $0_{\mathrm{Frac}(R)}=[0,g]$, 其中 $g\in R$ 是任意非零元, 其选取不影响等价类 $[0,g]$.

▷ **乘法幺元**　定义 $1_{\mathrm{Frac}(R)}=[1,1]$, 或者等价地定义为 $[g,g]$, 其中 $g\in R$ 是任意非零元.

以上显然是以熟知的分数运算规律为模板. 这使得 $\mathrm{Frac}(R)$ 成为交换环. 请读者逐条验证定义 3.1.1 的公理. 举例明之, 和加法相关的验证可以按 (3.5.1) 化约到分母 g 相同的简单情形, 例如有 $-[f,g]=[-f,g]$ 等; 一切都能还原到 R 本身的环论性质.

定义–命题 3.5.1 (分式域)　*以上定义使得 $\mathrm{Frac}(R)$ 成为域, 称为整环 R 的* **分式域**. *映射 $f\mapsto[f,1]$ 将 R 自然地嵌入为 $\mathrm{Frac}(R)$ 的子环.*

证明　对于任意 $f,g\in R\smallsetminus\{0\}$, 我们有

$$[g,f][f,g]=1_{\mathrm{Frac}(R)}=[f,g][g,f].$$

由于 R 是整环, 按定义可得 $[f,g]=0_{\mathrm{Frac}(R)}\iff f=0$, 这就表明 $\mathrm{Frac}(R)$ 的任何非零元都可逆.

其次考虑所断言的映射 $R\to\mathrm{Frac}(R)$. 按环结构的定义, 容易验证这是环同态. 由于 $[f_1,1]=[f_2,1]$ 按定义等价于 $f_1=f_2$, 它还是单的. 特别地, $\mathrm{Frac}(R)$ 不是零环, 因而是域. □

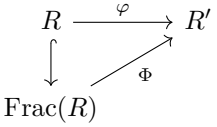
既然 $[f,g]\in\mathrm{Frac}(R)=[f,1][1,g]=[f,1][g,1]^{-1}$, 而 R 可以通过 $f\mapsto[f,1]$ 视同 $\mathrm{Frac}(R)$ 的子环, 我们今后不妨直接将 $\mathrm{Frac}(R)$ 的元素写成分式

$$\frac{f}{g}:=[f,g].$$

例 3.5.2　取 $R=\mathbb{Z}$, 则 §2.6 施行的构造正是说明 $\mathrm{Frac}(\mathbb{Z})=\mathbb{Q}$.

综上, 我们将整环 R 嵌入了一个域 $\mathrm{Frac}(R)$. 在后者作为商集的构造中, 既未加入多余的元素, 也没有施加任何不必要的等价关系. 这就提示了 $\mathrm{Frac}(R)$ 连同嵌入 $R\hookrightarrow\mathrm{Frac}(R)$ 应当是 “扩 R 为域” 这一问题的最优解. 何谓最优? 初学的读者对此可能仅有某种难以名状的确定感. 我们应用代数学中称为泛性质的思路 (详细讨论见诸本书附录), 将之拆分为一则命题和一则推论来作精确的界定.

命题 3.5.3　*设 R 为整环, R' 为交换环而 $\varphi:R\to R'$ 为环同态, 使得 $\varphi(R\smallsetminus\{0\})\subset(R')^\times$. 此时存在唯一的环同态 $\Phi:\mathrm{Frac}(R)\to R'$, 使得下图在约定 2.3.3 的意义下是交换图表:*



具体地说, Φ 必然映 f/g 为 $\varphi(f)\varphi(g)^{-1}$.

证明 如果这样的 Φ 存在, 则对任何 $g \in R \setminus \{0\}$ 必有交换环 R' 中的等式

$$1 = \Phi\left(\frac{g}{1} \cdot \frac{1}{g}\right) = \Phi\left(\frac{g}{1}\right)\Phi\left(\frac{1}{g}\right) = \varphi(g) \cdot \Phi\left(\frac{1}{g}\right),$$

从而 $\Phi(1/g) = \varphi(g)^{-1}$; 进一步, 对任何分式 $f/g \in \mathrm{Frac}(R)$ 必有

$$\Phi\left(\frac{f}{g}\right) = \Phi\left(\frac{f}{1}\right)\Phi\left(\frac{1}{g}\right) = \varphi(f)\varphi(g)^{-1}.$$

这就说明 Φ 的唯一性. 至于存在性, 我们定义

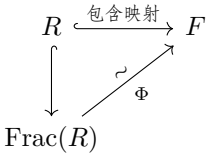
$$\tilde{\Phi} : \mathrm{Ratio}(R) \to R', \quad \tilde{\Phi}(f, g) = \varphi(f)\varphi(g)^{-1}.$$

若 $f_1 g_2 = f_2 g_1$, 则 $\varphi(f_1)\varphi(g_2) = \varphi(f_2)\varphi(g_1)$, 移项给出 $\tilde{\Phi}(f_1, g_1) = \tilde{\Phi}(f_2, g_2)$. 代入关于商集的命题 2.5.5 便得到映射

$$\Phi : \mathrm{Frac}(R) \to R', \quad \Phi\left(\frac{f}{g}\right) = \varphi(f)\varphi(g)^{-1}.$$

易见 Φ 限制到 R 上正好是 φ. 请读者按照 $\mathrm{Frac}(R)$ 的加法和乘法的定义说明 Φ 是环同态. □

推论 3.5.4 *设域 F 包含整环 R 作为子环, 而且 F 的所有元素都能表成 fg^{-1} 的形式, 其中 $f, g \in R$ 而 $g \neq 0$, 则存在唯一的环同构 $\Phi : \mathrm{Frac}(R) \to F$, 使得下图是交换图表:*



证明 因为 F 是域, 包含映射当然将 $R \setminus \{0\}$ 映入 $F \setminus \{0\} = F^\times$, 于是命题 3.5.3 给出唯一环同态 Φ 使图表交换, 问题在于证 Φ 既单又满.

设 $\Phi(f_1/g_1) = \Phi(f_2/g_2)$. 两边同乘以 $\Phi(g_1 g_2)$ 可得 $\Phi(g_2 f_1) = \Phi(g_1 f_2)$, 亦即 $g_2 f_1, g_1 f_2 \in R$ 在包含映射之下的像相同. 于是 $g_2 f_1 = g_1 f_2$ 而 $f_1/g_1 = f_2/g_2 \in \mathrm{Frac}(R)$. 单性得证.

任何 F 的元素都能写成 fg^{-1}, 其中 $f, g \in R$, $g \neq 0$. 然而 $f/g \in \mathrm{Frac}(R)$ 在 Φ 之下的像正是 $\Phi(f)\Phi(g)^{-1} = fg^{-1}$. 满性得证. □

回归有理函数的讨论. 推论 3.3.8 说明对于任何整环 A, 其上的多项式环 $A[X, Y, \ldots]$ (容许任意多个变元) 仍是整环, 因此可以代入分式域的构造.

定义 3.5.5 *设 F 为域, 则 $F[X, Y, \ldots]$ 的分式域称为以 $X, Y, \ldots$ 为变元的***有理函数域***, 记为 $F(X, Y, \ldots)$, 它包含 $F[X, Y, \ldots]$ 作为其子环. 有理函数域 $F(X, Y, \ldots)$ 的元素称为系数在 F 上, 以 $X, Y, \ldots$ 为变元的有理函数.*

正如同一般的域上的多项式区别于多项式函数, 以上定义的有理函数实非函数, 而是一些抽象符号的等价类. 一如多项式环的情形, $F(X,Y,\dots)$ 也可以设想为是由变元 $X,Y,\dots$ 连同 F 经过四则运算所自由地生成的域. 对于域 F 上由分式 f/g 代表的有理函数和 $x,y,\dots\in F$, 值 $f(x,y,\dots)/g(x,y,\dots)$ 只有在分母的值 $g(x,y,\dots)$ 非零时才有意义.

练习 3.5.6 定义 3.5.5 要求 F 是域. 对于一般的整环 R, 试给出同构 $\mathrm{Frac}(R[X])\simeq \mathrm{Frac}(R)(X)$; 说明多变元情形同样有 $\mathrm{Frac}(R[X,Y,\dots])\simeq \mathrm{Frac}(R)(X,Y,\dots)$.

提示〉 应用推论 3.5.4 对分式域的刻画.

次数的定义可以合理地从一元多项式扩展到一元有理函数上.

定义 3.5.7 (有理函数的次数) 设 F 为域. 对任意 $h=\frac{f}{g}\in F(X)$, 当 $h\neq 0$ 时定义 $\deg h:=\deg f-\deg g$. 另外规定 $\deg(0):=-\infty$.

注意到 $\frac{f_1}{g_1}=\frac{f_2}{g_2}$ 等价于 $F[X]$ 中的等式 $f_1g_2=f_2g_1$, 对两边取次数再移项, 可得

$$\deg f_1-\deg g_1=\deg f_2-\deg g_2,$$

所以有理函数的次数的确是良定义的. 对所有 $h_1,h_2\in F(X)$. 有理函数的次数满足:

$$\begin{aligned}\deg(h_1h_2)&=\deg h_1+\deg h_2,\\ \deg(h_1+h_2)&\leq \max\{\deg h_1,\deg h_2\}.\end{aligned}$$

第一条是定义 3.5.7 的直接结论. 对于第二条, 可以将分式 h_1,h_2 进行通分后计算加法, 将其归结为分子部分所满足的不等式.

3.6 多项式函数

在 §3.3 中, 我们说明了交换环 R 上的 n 元多项式

$$f=\sum_{i_1,\dots,i_n\geq 0} c_{i_1,\dots,i_n}X_1^{i_1}\cdots X_n^{i_n}\in R[X_1,\dots,X_n]\quad\text{(有限和)}$$

和它所对应的多项式函数

$$\begin{array}{ccc}R^n & \longrightarrow & R\\ (x_1,\dots,x_n) & \longmapsto & \sum_{i_1,\dots,i_n\geq 0} c_{i_1,\dots,i_n}x_1^{i_1}\cdots x_n^{i_n}\end{array}$$

之间并非一一对应: 例 3.3.5 已经给出一则反例.

这点和我们从中学数学或数学分析得到的印象是不同的. 将多项式看作函数并不能说是错误的观点, 关键在于厘清所需的条件.

考虑交换环 R 和 $n \in \mathbb{Z}_{\geq 1}$. 为了作明确区分, 姑且将从 n 元多项式到函数的映射记作

$$\mathrm{Fcn} : R[X_1, \ldots, X_n] \to \{\text{函数 } \phi : R^n \to R\}.$$

它映 $r \in R$ 为取值 r 的常值函数, 此外它保持加法和乘法:

$$\mathrm{Fcn}(f+g) = \mathrm{Fcn}(f) + \mathrm{Fcn}(g), \quad \mathrm{Fcn}(fg) = \mathrm{Fcn}(f)\mathrm{Fcn}(g),$$

其中函数的加法 (或乘法) 是逐点相加 (或相乘), 如 $(\phi_1 + \phi_2)(x_1, \ldots) = \phi_1(x_1, \ldots) + \phi_2(x_1, \cdots)$ 等. 这一简单事实不外是 (3.3.2) 的重述.

因此, 交换环 R 上的 n 元多项式可以视同函数的充要条件是 Fcn 是单射, 而因为 $\mathrm{Fcn}(f) = \mathrm{Fcn}(g)$ 等价于 $\mathrm{Fcn}(f-g)$ 为零函数, 这又相当于说 $\mathrm{Fcn}^{-1}(\text{零函数}) = \{0\}$.

往后真正关心的是域上的情形, 但容许整环并没有额外的困难. 关键在于命题 3.4.4 的以下推广.

引理 3.6.1 *设 R 为整环, $f \in R[X] \smallsetminus \{0\}$, 则 f 在 R 中至多只有 $\deg f$ 个相异的根.*

证明 将 R 嵌入为分式域 $F := \mathrm{Frac}(R)$ 的子环. 作为 $F[X]$ 的非零元, f 在 F 中至多只有 $\deg f$ 个相异根, 它在 R 中自然也是如此. □

命题 3.6.2 *设 R 为整环而 $n \in \mathbb{Z}_{\geq 1}$, 则 $\mathrm{Fcn} : R[X_1, \ldots, X_n] \to \{\text{函数 } R^n \to R\}$ 是单射当且仅当 R 有无穷多个元素.*

证明 设 R 有限, 则非零多项式 $\prod_{a \in R}(X_1 - a) \in R[X_1, \ldots, X_n]$ 在每个 $(x_1, \ldots, x_n) \in R^n$ 上取值皆为 0, 故 Fcn 非单射.

设 R 无穷, 而 $f \in R[X_1, \ldots, X_n]$ 满足 $\mathrm{Fcn}(f) = 0$. 以下递归地证明 $f = 0$.

设 $n = 1$ 而 $f \neq 0$, 则因为 R 无穷, $f \in R[X_1]$ 处处取零导致它至少有 $\deg f + 1$ 个相异根, 和引理 3.6.1 矛盾.

现在设 $n \geq 2$. 将 f 整理为有限和

$$f = \sum_{i_1 \geq 0} \left(\sum_{i_2, \ldots, i_n \geq 0} c_{i_1, i_2, \ldots, i_n} X_2^{i_2} \cdots X_n^{i_n} \right) X_1^{i_1}.$$

对每个 $(x_2, \ldots, x_n) \in R^{n-1}$, 一元多项式 $f(X_1, x_2, \ldots, x_n) \in R[X_1]$ 处处取值为 0, 因此 $n = 1$ 情形蕴涵它的所有系数皆为 0, 亦即

$$\forall i_1 \geq 0, \quad \sum_{i_2, \ldots, i_n \geq 0} c_{i_1, i_2, \ldots, i_n} x_2^{i_2} \cdots x_n^{i_n} = 0.$$

变动 $(x_2, \ldots, x_n)$ 并运用 $n-1$ 的情形, 这又导致 $R[X_2, \ldots, X_n]$ 中的等式

$$\forall i_1 \geq 0, \quad \sum_{i_2, \ldots, i_n \geq 0} c_{i_1, i_2, \ldots, i_n} X_2^{i_2} \cdots X_n^{i_n} = 0.$$

综上, $f = 0$. 明所欲证. □

定理 3.6.3 (代数等式的延拓原理) 设 R 为无穷整环, $f, g_1, \dots, g_m \in R[X_1, \dots, X_n]$, 其中 $g_1, \dots, g_m$ 皆非零元, 而且对于所有 $(x_1, \dots, x_n) \in R^n$ 皆有

$$(g_1(x_1, \dots, x_n) \neq 0 \ \wedge \ \dots \ \wedge \ g_m(x_1, \dots, x_n) \neq 0) \implies f(x_1, \dots, x_n) = 0,$$

则 $f = 0$.

证明 从条件可知 $f \prod_{i=1}^m g_i$ 代入所有 $(x_1, \dots, x_n) \in R^n$ 皆取零, 故命题 3.6.2 蕴涵 $f \prod_{i=1}^m g_i = 0$. 既然推论 3.3.8 确保 $R[X_1, \dots, X_n]$ 是整环, 这导致 $f = 0$. □

定理 3.6.3 相当于说一个代数方程 $f = 0$ 只要在 $g_1, \dots, g_m \neq 0$ 的一般情形成立, 则在某个 $g_i = 0$ 的例外情形也自动成立, 前提是 R 为无穷整环. 这种论证在一些文献中也称为**扰动法**.

练习 3.6.4 设 R 为交换环而 $f = a_1X_1 + \cdots + a_nX_n + b \in R[X_1, \dots, X_n]$. 说明系数 $a_1, \dots, a_n, b \in R$ 由多项式函数 $\mathrm{Fcn}(f)$ 唯一确定.

提示〉 代值读出系数.

3.7 域的特征

本节有必要以下标来区分环 R 的零元 0_R, 幺元 1_R 和作为整数的 0 和 1.

在 §3.1 关于域的讨论中, 我们已经发现有限域 $\mathbb{F}_p$ (其中 p 是素数) 和熟悉的 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 等域有一个重要差异: 设 n 为任意正整数. 对于域 $F \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, 按照 (3.1.1) 的符号, 我们有

$$n \cdot 1_F = 0_F \iff n = 0,$$

然而在域 $F = \mathbb{F}_p$ 中却有

$$p \cdot 1_F = 0_F;$$

基于 (3.1.2) 的运算规律, 后者又蕴涵对于任意 $x \in F$ 都有 $px = (p \cdot 1_F) \cdot x = 0_F$.

后一性质并非有限域独有. 以 $\mathbb{F}_p$ 上的有理函数域 $\mathbb{F}_p(X)$ 为例, 它有无穷多个元素, 但也满足 $p \cdot 1_{\mathbb{F}_p(X)} = 0_{\mathbb{F}_p(X)}$, 这点只须在其子域 $\mathbb{F}_p$ 里验证.

这就启发了域的特征的概念. 首先需要一则简单观察.

引理 3.7.1 对于任意环 R, 存在唯一的环同态 $\mathbb{Z} \to R$. 按照 (3.1.1) 的符号, 唯一可能的映法是 $n \mapsto n \cdot 1_R$.

证明 唯一性是明白的, 因为环同态必然映 1 为 1_R, 从而映 $n \geq 0$ 为 $\underbrace{1_R + \cdots 1_R}_{n \text{ 项}} = n \cdot 1_R$, 而在 $n < 0$ 时映 $n = -|n|$ 为 $-(|n| \cdot 1_R) = n \cdot 1_R$.

存在性问题则归结为检验 $n \mapsto n \cdot 1_R$ 确实是环同态. 除了显然的等式 $1 \cdot 1_R = 1_R$, 其余所需等式都已经由 (3.1.2) 提供了. □

因此我们可以考虑 $\mathbb{Z}$ 的子集 $K_R := \{n \in \mathbb{Z} : n \cdot 1_R = 0_R\}$, 它包含 0, 对加法封闭, 而且若 $n \in K_R$ 而 $m \in \mathbb{Z}$, 则 $mn \cdot 1_R = (m \cdot 1_R)(n \cdot 1_R) = 0_R$ 蕴涵 $mn \in K_R$. 基于这两种封闭性, 引理 2.7.2 遂说明存在唯一的 $g \in \mathbb{Z}_{\geq 0}$ 使得 $K_R = g\mathbb{Z}$. 当 R 是定义 3.1.11 的整环时, 对 g 还有更严格的限制.

定义–命题 3.7.2 (特征) *设 R 为整环. 存在唯一的 $\mathrm{char}(R) \in \mathbb{Z}_{\geq 0}$ 使得对所有 $n \in \mathbb{Z}$ 都有*

$$n \cdot 1_R = 0_R \iff \mathrm{char}(R) \mid n,$$

称之为整环 R 的特征; 它或者是 0, 或者是素数.

证明 按先前讨论的方式定义 K_R. 业已说明存在唯一的 $\mathrm{char}(R) \in \mathbb{Z}_{\geq 0}$ 使得 $K_R = \mathrm{char}(R)\mathbb{Z}$. 设 $\mathrm{char}(R) \neq 0$, 而且有因数分解 $\mathrm{char}(R) = ab$, 则因为 $n \mapsto n \cdot 1_R$ 是环同态, 故

$$\mathrm{char}(R) \cdot 1_R = (a \cdot 1_R)(b \cdot 1_R) = 0_R.$$

又因为 R 是整环, 必有 $a \in K_R$ 或 $b \in K_R$, 因此必有 $\mathrm{char}(R) \mid a$ 或 $\mathrm{char}(R) \mid b$; 留意到 $\mathrm{char}(R) \neq 1$ (否则将有 $1_R = 0_R$). 这足以说明 $\mathrm{char}(R)$ 若非零则必为素数. □

因此在特征为 $p > 0$ 的整环 R 中, 任意 $x \in R$ 的 p 倍必然为零: $px = (p \cdot 1_R)x = 0_R x = 0_R$.

练习 3.7.3 设 p 为素数, 而 R 为满足 $p \cdot 1_R = 0_R$ 的交换环 (例如特征 p 的整环). 证明对所有 $x, y \in R$ 皆有

$$(x + y)^p = x^p + y^p.$$

提示〉 展开左式, 问题归结为证二项式系数 $\binom{p}{k}$ 在 $0 < k < p$ 时总是 p 的倍数.

命题 3.7.4 *若 R_0 是整环 R 的子环, 则 $\mathrm{char}(R_0) = \mathrm{char}(R)$.*

证明 本书规定子环 R_0 必满足 $1_R = 1_{R_0}$, 所以等式 $n \cdot 1_R = 0_R$ 成立与否可以在子环 R_0 中判定. □

最常用的是 R 为域的情形. 本节开头的讨论相当于说

$$\mathrm{char}(\mathbb{Q}) = \mathrm{char}(\mathbb{R}) = \mathrm{char}(\mathbb{C}) = 0, \quad \mathrm{char}(\mathbb{F}_p) = \mathrm{char}(\mathbb{F}_p(X)) = p,$$

其中 p 是任意素数. 整环 R 的特征和它的分式域的特征是一回事: 诚然, 根据命题 3.7.4, 从 $R \subset \mathrm{Frac}(R)$ 可见 $\mathrm{char}(R) = \mathrm{char}(\mathrm{Frac}(R))$.

练习 3.7.5 设 E 和 F 为域, $\mathrm{char}(E) \neq \mathrm{char}(F)$, 说明不存在从 E 到 F 的环同态.
提示〉 可用练习 3.2.7.

因此, 不同特征的域无法直接沟通, 除非通过一个较大的整环相联系, 例如

$$\mathbb{F}_p \overset{\text{商映射}}{\twoheadleftarrow} \mathbb{Z} \overset{\text{包含}}{\hookrightarrow} \mathbb{Q},$$

或者是运用更复杂的代数或数论技术.

最后, 一个域 F 的特征取决于其中能否嵌入 $\mathbb{Q}$ 抑或 $\mathbb{F}_p$ 作为子域, 精确解释如下. 回忆到引理 3.7.1 给出唯一的环同态 $\mathbb{Z} \to F$, 而练习 3.2.7 说明域之间的同态必然单.

- ⋆ 如果 $\mathrm{char}(F) = 0$, 则同态 $\mathbb{Z} \to F$ 是单的, 这是因为 $m \cdot 1_F = n \cdot 1_F \iff (m-n) \cdot 1_F = 0 \iff m - n = 0$. 由于 F 的非零元皆可逆, 由此导出域之间的同态

$$\mathbb{Q} = \mathrm{Frac}(\mathbb{Z}) \longrightarrow F$$

$$\frac{a}{b} \longmapsto (b \cdot 1_F)^{-1}(a \cdot 1_F);$$

  这是命题 3.5.3 的一则应用, 尽管直接检验也毫无困难. 于是 $\mathbb{Q} \hookrightarrow F$.

- ⋆ 如果 $\mathrm{char}(F) = p > 0$, 则 $n \cdot 1_R$ 只和 n 的 mod p 同余类 $n + p\mathbb{Z}$ 相关. 这就给出良定义的映射

$$\mathbb{F}_p \longrightarrow F$$

$$n + p\mathbb{Z} \longmapsto n \cdot 1_F.$$

  容易看出映射保持加法和乘法, 由此得到域嵌入 $\mathbb{F}_p \hookrightarrow F$.

反过来说, 假若 $\mathbb{Q}$ (或 $\mathbb{F}_p$) 能嵌入为 F 的子域, 则命题 3.7.4 蕴涵 $\mathrm{char}(F) = \mathrm{char}(\mathbb{Q}) = 0$ (或 $\mathrm{char}(F) = \mathrm{char}(\mathbb{F}_p) = p$).

从上述论证也可以明白, 从域 F 中的 $0_F, 1_F$ 出发, 通过四则运算所能得到的最小子域或者是 $\mathbb{Q}$ 的一份副本 (特征 0 情形), 或者是 $\mathbb{F}_p$ (特征 $p > 0$ 情形) 的一份副本. 这一最小子域被称为 F 的**素域**.

练习 3.7.6 设 F 是特征为 p 的域, 其中 p 是素数或 0. 对所有不被 p 整除的整数 n 和 $x \in F$, 说明存在唯一的 $y \in F$ 使得 $ny = x$. 此元素可以合理地记为 $\frac{1}{n}x$.

提示〉 取 n 在素域中的像 $\overline{n}$, 则 $\overline{n} \neq 0_F$, 从而可以取 $y := \overline{n}^{-1}x$. 至于唯一性, $\overline{n} \neq 0_F$ 蕴涵 $\overline{n}y = \overline{n}y' \iff y = y'$.

习题

1. 设 A 和 B 为环, 考虑由 $f(a) = (a, 0)$ 确定的映射 $f : A \to A \times B$, 说明 f 保持加法, 乘法和零元, 然而 f 不是环同态.
2. 证明有限整环必为域. 提示〉 设 R 为有限整环. 若 $x \in R \smallsetminus \{0\}$, 则 x 给出的左乘映射 $\lambda_x : R \to R$ 和右乘映射 $\rho_x : R \to R$ 皆为单射. 应用抽屉原理 (命题 2.9.4) 推导 x 可逆.
3. (Wilson 定理) 设 p 为素数. 运用 $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ 为域的这一事实, 证明同余式

$$(p-1)! \equiv -1 \pmod{p}.$$

   提示〉 正整数 $1, \ldots, p-1$ 的 mod p 同余类遍历 $\mathbb{F}_p^\times$. 将 $\mathbb{F}_p^\times$ 的元素按 $x \leftrightarrow x^{-1}$ 配对. 分析有哪些元素自配对.

4. 在给定集合 S 的幂集 $P(S)$ 上定义两种二元运算

$$A+B:=(A\cup B)\smallsetminus(A\cap B),\quad A\cdot B=A\cap B.$$

证明 $P(S)$ 对此成为环. 〈提示〉 取 $0_{P(S)}$ 为 $\varnothing$, 取 $1_{P(S)}$ 为 S.

5. 设 R 为环, 而且所有 $r\in R$ 都满足 $r^2=r$. 证明 R 是交换环.

6. 设 R 为环, $x\in R^\times$. 定义映射

$$\begin{aligned}\mathrm{Ad}_x : R &\to R\\ r &\mapsto xrx^{-1}.\end{aligned}$$

证明 Ad_x 是从 R 到其自身的环同构 (简称自同构). 进一步验证 $\mathrm{Ad}_{xy}=\mathrm{Ad}_x\,\mathrm{Ad}_y$, 而 $(\mathrm{Ad}_x)^{-1}=\mathrm{Ad}_{x^{-1}}$. 这种自同构称为 R 的**内自同构**.

7. (N. Jacobson) 设 R 为环, $x,y\in R$. 证明若 $1-xy$ 可逆, 则 $1-yx$ 也可逆, 而且

$$(1-yx)^{-1}=1+y(1-xy)^{-1}x.$$

〈提示〉 既然逆的公式已经写下, 剩下只是计算. 关键在于如何想出这个公式? 数学分析的思路在此是有益的. 形式地操作无穷级数:

$$\begin{aligned}(1-yx)^{-1}&=1+yx+yxyx+yxyxyx+\cdots\\&=1+y(x+xyx+xyxyx+\cdots)\\&=1+y(1+xy+xyxy+\cdots)x\\&=1+y(1-xy)^{-1}x.\end{aligned}$$

为了使此式有意义, 环 R 必须具有一个合适而且完备的距离结构, 使得 xy 和 yx 到零点的距离都小于 1. 但对于一般的 R, 上式依然能协助我们猜出求逆的公式.

8. 取无交并 $\mathbb{R}\sqcup\{\infty\}$, 其中 ∞ 仅作为一个符号来理解. 在其上定义运算

$$x\oplus y:=\min\{x,y\},\quad x\odot y:=x+y,$$

其中涉及 ∞ ("正无穷大") 的运算按直观的方式理解.

(i) 证明 $(\mathbb{R}\cup\{\infty\},\oplus,\odot,\infty,0)$ 非环. 具体说明它缺乏哪一条性质.

(ii) 另一方面, 对所有 $n\in\mathbb{Z}_{\geq 1}$, 验证 $(x\oplus y)^{\odot n}:=(x\oplus y)\odot\cdots\odot(x\oplus y)$ 等于 $x^{\odot n}\oplus y^{\odot n}$.

9. 承上题, 考虑相对于 $\odot$ 运算的 n 元 "单项式" 函数, 形如

$$(x_1,\dots,x_n)\mapsto x_1^{i_1}\cdots x_n^{i_n}\in\mathbb{R},$$

其中 $i_1,\dots,i_n\in\mathbb{Z}_{\geq 0}$, 左式是 $\mathbb{R}^n$ 的元素, 右式则定义为 i_1 份 x_1, ..., i_n 份 x_n 按任意顺序以 $\odot$ 运算给出的结果, 不是寻常的乘法.

(i) 具体描述上述函数的样貌.

(ii) 推而广之, 描述 $\mathbb{R}^n$ 上的 "多项式" 函数

$$(x_1,\dots,x_n)\mapsto\left(a\odot x_1^{i_1}\cdots x_n^{i_n}\right)\oplus\left(b\odot x_1^{j_1}\cdots x_n^{j_n}\right)\oplus\cdots$$

(取有限和, $a,b,\dots\in\mathbb{R}$).

(iii) 说明若 p 是如上的 “多项式” 函数, 则

- ⋆ p 连续,
- ⋆ p 是分片线性的 (大致地解释其直观意义),
- ⋆ p 是凹函数, 换言之它满足

$$p\left(\frac{x+y}{2}\right) \geq \frac{p(x)+p(y)}{2}, \quad x, y \in \mathbb{R}^n.$$

10. 在平面上指定原点和规定方向的单位长度; 如果选定坐标系, 这也相当于指定 $\mathbb{R}^2$ 的点 $(0,0)$ 和 $(1,0)$. 说明若将 $\mathbb{R}^2$ 视同复平面 $\mathbb{C}$, 则从这两个给定的点出发, 通过平面上的尺规作图所能得到的所有点构成 $\mathbb{C}$ 的子域.

提示〉 以尺规作图实现复平面上的四则运算.

11. 设 p 为奇素数. 证明

(i) 同余式 $a^2 \equiv b^2 \pmod{p}$ 等价于 $a \equiv \pm b \pmod{p}$;

(ii) 存在 $a, b \in \mathbb{Z}$ 使得 $p \mid a^2 + b^2 + 1$. 提示〉 当 a 变化, a^2 除以 p 的余数取 $\frac{p+1}{2}$ 个值. 同理, $-1-b^2$ 除以 p 的余数也取 $\frac{p+1}{2}$ 个值.

12. 选定素数 p. 定义

$$\mathbb{Z}_{(p)} := \left\{ \frac{a}{b} \in \mathbb{Q} : a, b \in \mathbb{Z},\ p \nmid b \right\}.$$

(i) 证明 $\mathbb{Z}_{(p)}$ 是 $\mathbb{Q}$ 的子环.

(ii) 设 F 是域, $\mathrm{char}(F) \in \{p, 0\}$. 说明存在唯一的环同态 $\mathbb{Z}_{(p)} \to F$.

(iii) 承上, 对于 $n \in \mathbb{Z} \smallsetminus p\mathbb{Z}$ 和 $x \in F$, 说明练习 3.7.6 中的 $\frac{1}{n}x$ 等于 $\frac{1}{n} \in \mathbb{Z}_{(p)}$ 在 F 中的像和 x 的乘积.

13. (华罗庚) 设 R 为环, $a, b \in R^\times$, 而且 $1 - ab \in R^\times$. 证明

$$a - \left(a^{-1} + (b^{-1} - a)^{-1}\right)^{-1} = aba.$$

提示〉 观察到当 $x, 1-x \in R^\times$ 时 $\left(x^{-1} - 1\right)^{-1} = (1-x)^{-1} - 1$.

14. (Cartan–Brauer–华罗庚) 设 D 为除环, R 为 D 的子环, 而且 R 本身也是除环. 证明若对所有 $d \in D^\times$ 皆有 $dRd^{-1} \subset R$, 则必有 $R = D$ 或 $R \subset Z(D)$; 关于中心 $Z(D)$ 的定义请见例 3.1.4.

提示〉 验证以下等式: 若 $a, b \in D$ 满足 $ab \neq ba$, 则

$$a = \left(b - (a-1)^{-1}b(a-1)\right)\left(a^{-1}ba - (a-1)^{-1}b(a-1)\right)^{-1}.$$

由此推导 $D = \{a \in D : \forall b \in R,\ ab = ba\} \cup R$; 右式两项都是对加法及取加法逆元封闭的子集, 论证至少有一项是 D.

15. 举例说明存在环 R 和 $x \in R$, 使得 x 不可逆, 然而存在无穷多个 y 满足 $xy = 1$.

提示〉 一种构造是考虑 $S := \{\alpha := (a_n)_{n\geq 1} : \text{整数列}\}$, 定义 S 上的加法运算为逐项相加. 取 R 为所有满足 $\phi(\alpha+\beta) = \phi(\alpha) + \phi(\beta)$ 的映射 $\phi : S \to S$ 构成的集合. 说明 R 对加

法 $(\phi+\psi)(\alpha):=\phi(\alpha)+\psi(\alpha)$ 与函数合成运算 $\circ$ 成环, $1_R:=\mathrm{id}_S$. 取 $x\in R$ 为平移映射 $(a_1,a_2,\ldots)\mapsto(a_2,a_3,\ldots)$, 再对所有 $k\geq 1$ 考虑 $y_k\in R$, 使得 $(b_n)_{n\geq 1}:=y_k(\alpha)$ 满足

$$b_n=\begin{cases}a_{n-1}, & n\geq 2,\\ a_k, & n=1.\end{cases}$$

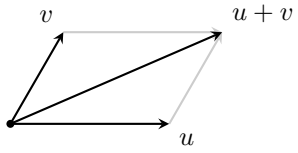
第四章 向量空间和线性映射

向量空间是最常用的代数结构之一. 它的根源至少有两方面.

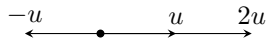
1. 线性方程组理论. 它将所求的解表作形如 $\boldsymbol{x} = (x_1, \dots, x_n)$ 的数组, 而数组可以逐个坐标地相加或伸缩:

$$(x_1, \dots, x_n) + (x_1', \dots, x_n') = (x_1 + x_1', \dots, x_n + x_n'),$$
$$t(x_1, \dots, x_n) = (tx_1, \dots, tx_n).$$

2. 初等几何学中的平面向量与空间向量理论. 这种意义的向量可以理解为带有方向和大小的量, 能按照平行四边形法则相加:



或以任意实数为比例作伸缩:



在建立坐标系后, 平面与空间向量也可以分别等同于 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中的数组.

基于数组的观点胜在操作简便, 而且容许任意大的 n, 但也有其局限. 在几何方面, 直观中的向量只在建立坐标系后才能等同于数组, 坐标系的选取有任意性, 而不依赖坐标的操作显然更优美, 也更贴近几何的本质. 在线性方程组的方面, 选定域 F (例如 $\mathbb{R}$ 或 $\mathbb{C}$), 并考虑系数在 F 上的方程组

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= 0, \\ \vdots \qquad & \quad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= 0. \end{aligned}$$

我们希望研究所有解 $\boldsymbol{x} = (x_1, \dots, x_n)$ 在 F^n 中构成的子集 $\mathcal{S}$. 尽管这是一个关于数组的问题, 却自然地引出更广的结构: 集合 $\mathcal{S}$ 包含 $\mathbf{0} = (0, \dots, 0)$, 而且有

$\boldsymbol{x},\boldsymbol{x}'\in\mathcal{S}\implies\boldsymbol{x}+\boldsymbol{x}'\in\mathcal{S}$ 和 $(t,\boldsymbol{x})\in F\times\mathcal{S}\implies t\boldsymbol{x}\in\mathcal{S}$. 这提示了解集 $\mathcal{S}$ 尽管未必是整个 F^n, 却和 F^n 或几何学中的向量具有类似的运算. 解线性方程组是再实际不过的问题, 研究并运用解集上的这种代数结构当然也是题中之义. 详细的说明和更多例证见诸 §4.1.

大致地说, 域 F 上的向量空间 (定义 4.2.1) 意谓一个集合 V, 连同两种运算

▷ **向量加法** $+:V\times V\to V$, 写作 $(v_1,v_2)\mapsto v_1+v_2$, 满足结合律, 交换律, 和逆元的存在性等性质;

▷ **纯量乘法** $\cdot:F\times V\to V$, 写作 $(t,v)\mapsto t\cdot v=tv$, 满足结合律与对加法的分配律等性质.

这一理论框架能统合先前介绍的所有情形, 此外还有源源不绝的实例. 向量空间的元素也简称为向量. 初等几何学中关于向量的直观在此依然起到指导作用. 详细的定义和例子是 §§4.2–4.3 的内容.

基与维数是关于向量空间的基本概念, 与之相关的概念还有生成系和线性无关子集, 都属于 §4.4 的内容. 基可以理解为向量空间的脚手架: 以有限维情形为例, 一旦选定 F-向量空间 V 的基, V 便能等同于数组构成的空间 F^n, 其中 $n=\dim V$ 是 V 的维数. 基存在但并不唯一; 选基与否, 如何选基往往是解决具体问题的关键.

本书主要探讨有限维向量空间. 许多一般结论能推及无穷维情形, 但涉及称为 Zorn 引理的集合论事实, 本书纳入附录处理. 无穷维空间在应用中也自然地出现, 最简单的实例是域 F 上的所有一元多项式构成 F-向量空间 $F[X]$.

如果向量空间之间的映射 $T:V\to W$ 满足恒等式

$$T(v_1+v_2)=T(v_1)+T(v_2),\quad T(tv)=tT(v),$$

则称 T 为线性映射. 一如我们在第三章曾见过的, 由这种保结构的映射自然地引申出同构的概念. 从 V 到 W 的全体线性映射构成集合 $\mathrm{Hom}(V,W)$, 它对逐点的运算 $(T+T')(v)=T(v)+T'(v)$ 和 $(tT)(v)=tT(v)$ 成为向量空间. 在有限维的情形, 一旦选基将 V (或 W) 等同于 F^n (或 F^m), 便能具体地以 $m\times n$ 矩阵描述线性映射. 我们在 §1.3 探讨 Gauss–Jordan 消元法时已经见过 $m\times n$ 矩阵, 记如

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix},$$

当时仅用以标记线性方程组中的常数, 然而矩阵的内涵远不止于此: 域 F 上的所有 $m\times n$ 矩阵对逐项运算构成 F-向量空间 $\mathrm{M}_{m\times n}(F)$, 除此之外, 矩阵之间还有乘法运算

$$\mathrm{M}_{m\times n}(F)\times\mathrm{M}_{n\times r}(F)\to\mathrm{M}_{m\times r}(F),\quad(\boldsymbol{A},\boldsymbol{B})\mapsto\boldsymbol{AB}.$$

矩阵的加法和乘法满足结合律, 分配律等性质, 但乘法一般不交换. 这些运算都能在线性映射的层次得到自然解释: 矩阵的加法与来自 F 的纯量乘法对应到 Hom 上相应的运算, 乘法则对应到线性映射的合成. 详细的定义与解释是 §§4.5–4.6 的主题.

线性映射与矩阵还有许多相互对应的概念与操作. 例子有 §4.7 介绍的转置, 这涉及向量空间的对偶空间; §4.8 介绍的核, 像与秩的概念, 它们与 Gauss–Jordan 消元法密切相关; §4.9 介绍的基的变换, 以及矩阵之间的共轭 (又称相似) 和相抵关系, 共轭的概念对于后续内容尤其重要. 事实上, 共轭的 $n \times n$ 矩阵相当于同一个线性映射在不同的基之下的表法.

在 §4.10, 我们将从内在和外在两种角度探讨向量空间的直和分解; 若一个线性映射兼容于给定的直和分解, 对应的矩阵便简化为分块对角的形式, 见 §4.11.

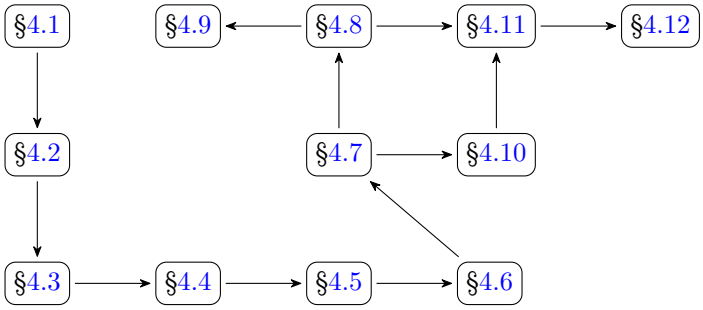
最后, §4.12 介绍的商空间相对而言更加抽象, 但仍是不可或缺的基本概念. 约略地说, 取 V 对子空间 U 的商相当于将 V 中满足 $v_1 - v_2 \in U$ 的元素 v_1 和 v_2 等量齐观, 这是 V 上的等价关系, 而对应的商集 V/U 仍有向量空间结构. 商空间可以用来描述任意线性映射 $T : V \to W$ 的像空间.

本章的内容只是向量空间理论的一个引子, 后续章节将有更深入的拓展.

阅读提示

若无另外说明, F 在本章均代表一个选定的域. 此外, 关于向量空间或矩阵的计算与抽象理论同等重要, 学习时应当同步掌握.

阅读顺序



4.1 引言: 回到线性方程组

回首 §1.3 考察的线性方程组

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \vdots \quad & \quad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m. \end{aligned} \tag{4.1.1}$$

当时我们默认系数和所求的解都在 $\mathbb{C}$ 中. 然而解方程所用的 Gauss–Jordan 消元法所需的仅是四则运算, 无关复数的一切其他性质, 所以现在可以假定 a_{ij} 和 b_i 都落在选定的域 F 中, 在 F 上求解. 我们有:

> 在任意域 F 上, 线性方程组 (4.1.1) 可以由 Gauss–Jordan 消元法判定是否有解; 若有解, 则消元法可以描述所有的解.

这是因为 §§1.3—1.4 介绍的行运算和 Gauss–Jordan 消元法理论可以一字不易地推广到任意域上. 在线性方程组的研究中, $b_1 = \ldots = b_m = 0$ 的情形占有特殊的地位.

定义 4.1.1 *考虑域 F 上形如 (4.1.1) 的 n 元线性方程组. 如果 $b_1 = \cdots = b_m = 0$, 则称此方程组为***齐次***的.*

消元法的基础是矩阵的初等行变换. 以下转换视角, 从映射的观点来理解域 F 上的线性方程组.

给定 $n, m \in \mathbb{Z}_{\geq 1}$ 和一族系数 $(a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$, 其中 $a_{ij} \in F$, 定义映射

$$\begin{aligned} T : F^n &\longrightarrow F^m \\ (x_j)_{j=1}^n &\longmapsto \left(\textstyle\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j\right). \end{aligned}$$

今后也将数组 $(x_1, \ldots, x_n)$ 写作 $\boldsymbol{x}$ 的形式. 对于给定的 $\boldsymbol{b} = (b_1, \ldots, b_m) \in F^m$. 解 (4.1.1) 相当于研究纤维 $T^{-1}(\boldsymbol{b})$ (定义 2.2.10). 对任何映射 $F^n \to F^m$ 都可以考虑同样的问题, 然而眼下问题的特点在于 T 是 "线性" 的. 更明确地说, 在 F^n 上定义

- ⋆ 两个数组的**加法**运算

$$(x_1, \ldots, x_n) + (x'_1, \ldots, x'_n) := (x_1 + x'_1, \ldots, x_n + x'_n),$$

  其中 $(x_1, \ldots, x_n), (x'_1, \ldots, x'_n) \in F^n$;

- ⋆ 域 F 之于数组的**纯量乘法**运算

$$t(x_1, \ldots, x_n) := (tx_1, \ldots, tx_n);$$

  其中 $t \in F$.

如此定义的加法满足结合律和交换律, 而纯量乘法对加法满足结合律和分配律:

$$(tt')\boldsymbol{x} = t(t'\boldsymbol{x}), \quad t(\boldsymbol{x}+\boldsymbol{y}) = t\boldsymbol{x} + t\boldsymbol{y}, \quad (t+t')\boldsymbol{x} = t\boldsymbol{x} + t'\boldsymbol{x};$$

我们不厌其烦地重申: $t, t' \in F$ 而 $\boldsymbol{x}, \boldsymbol{y} \in F^n$. 此外, $\mathbf{0} := (0, \dots, 0) \in F^n$ (必要时记为 $\mathbf{0}_n$) 对加法充当了零元的角色:

$$\boldsymbol{x} + \mathbf{0} = \boldsymbol{x} = \mathbf{0} + \boldsymbol{x}.$$

由于运算是逐分量定义的, 上述性质全都归结为域 F 的相应性质. 另记 $-\boldsymbol{x} := (-1)\boldsymbol{x}$ 和 $\boldsymbol{x}' - \boldsymbol{x} := \boldsymbol{x}' + (-\boldsymbol{x})$.

这对解方程的问题有何启发? 请读者验证映射 T 满足

$$T(\underbrace{\boldsymbol{x}+\boldsymbol{y}}_{F^n}) = \underbrace{T(\boldsymbol{x}) + T(\boldsymbol{y})}_{F^m}, \quad T(t\boldsymbol{x}) = t \cdot T(\boldsymbol{x}).$$

事实上, $T(\boldsymbol{x})$ 的每个分量 $\sum_{j=1}^{n} a_{ij}x_j \in F$ 都有这些性质 $(1 \le i \le m)$. 一个立即的结论是线性方程组的解可以线性叠加. 何谓线性叠加? 先固定方程组 (4.1.1) 的系数矩阵 $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$.

- ⋆ 若 $\boldsymbol{x} = (x_1, \dots, x_n)$ 是方程组对 $\boldsymbol{b} = (b_1, \dots, b_m)$ 的解, 而 $\boldsymbol{x}' = (x'_1, \dots, x'_n)$ 是对 $\boldsymbol{b}' = (b'_1, \dots, b'_m)$ 的解, 则 $\boldsymbol{x} + \boldsymbol{x}'$ 是方程组对 $\boldsymbol{b} + \boldsymbol{b}'$ 的解;
- ⋆ 若 $\boldsymbol{x} = (x_1, \dots, x_n)$ 是 (4.1.1) 对 $\boldsymbol{b} = (b_1, \dots, b_m)$ 的解, $t \in F$, 则 $t\boldsymbol{x}$ 是方程组对 $t\boldsymbol{b}$ 的解.

因此像集 $\mathrm{im}(T)$ 在 F^m 的加法和纯量乘法运算下封闭. 像集还是非空的, 因为 $\mathbf{0}_m = T(\mathbf{0}_n) \in \mathrm{im}(T)$.

作为推论, 若 $\boldsymbol{x}$ 和 $\boldsymbol{x}'$ 是方程组对同一个 $\boldsymbol{b} \in F^m$ 的解, 则 $\boldsymbol{y} := \boldsymbol{x} - \boldsymbol{x}'$ 是对应的齐次方程组的解. 反之, 若已知 $\boldsymbol{x}$ 是原方程组的解, 而 $\boldsymbol{y}$ 是对应的齐次方程组的解, 则 $\boldsymbol{x} + \boldsymbol{y}$ 仍是原方程组的解.

综之, 只要找出特解 $\boldsymbol{x} \in T^{-1}(\boldsymbol{b})$, 方程组 (4.1.1) 的通解便由下式描述

$$\begin{aligned} T^{-1}(\boldsymbol{b}) &= \boldsymbol{x} + T^{-1}(\mathbf{0}) \\ &:= \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{y} \in F^n,\ T(\boldsymbol{y}) = \mathbf{0}_m\}. \end{aligned}$$

这就引出了以下观察:

- ⋆ 线性方程组 (4.1.1) 有解当且仅当 $\boldsymbol{b} := (b_1, \dots, b_m) \in \mathrm{im}(T)$;
- ⋆ 一旦它有解, 则精确到用某个特解的 "平移", 解集 $T^{-1}(\boldsymbol{b})$ 的性状完全由对应的齐次方程组确定, 或者说由系数矩阵 $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$ 确定.

这表明对于解方程而言, F^n 的子集 $T^{-1}(\mathbf{0})$, 亦即对应的齐次方程组的解集和 $\mathrm{im}(T)$ 一样是重要的对象. 一如 $\mathrm{im}(T)$, 子集 $T^{-1}(\mathbf{0})$ 同样对 F^n 的加法和纯量乘法运算封闭, 这是因为

$$\begin{aligned} T(\boldsymbol{x}) = \mathbf{0}_m = T(\boldsymbol{x}') &\implies \\ T(\boldsymbol{x}+\boldsymbol{x}') = \mathbf{0}_m + \mathbf{0}_m &= \mathbf{0}_m, \\ t \in F, \quad T(\boldsymbol{x}) = \mathbf{0}_m &\implies \\ T(t\boldsymbol{x}) = tT(\boldsymbol{x}) = t\,\mathbf{0}_m &= \mathbf{0}_m. \end{aligned}$$

对 $T^{-1}(\mathbf{0}_m)$ 可以提许多问题, 比如说, 解集的参数化所指究竟为何? 如何确定所需的参数个数, 即先前所谓的 "自由度"?

定义 4.1.2 设 $T : F^n \to F^m$ 按上述方式对应到一个齐次线性方程组. 若 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h \in F^n$ 都是方程组的解, 而且所有解 $\boldsymbol{x} \in F^n$ 都可以通过加法和纯量乘法表作

$$\boldsymbol{x} = \sum_{i=1}^{h} t_i \boldsymbol{v}_i, \quad t_1, \ldots, t_h \in F,$$

其中 $(t_1, \ldots, t_h)$ 由 $\boldsymbol{x}$ 唯一确定, 则称 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h$ 是该齐次方程组的一组**基础解系**.

求出基础解系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h$ 相当于用 h 个参数来描述齐次线性方程组的解集.

命题 4.1.3 考虑形如 (4.1.1) 的 n 元齐次线性方程组, 其中 $\boldsymbol{b} = 0$. 设消元法给出的简化行梯矩阵有 r 个主元, 则对应的齐次方程组有基础解系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-r}$.

证明 沿用 §1.4 的记号. 将简化行梯矩阵中不含主元的列按编号枚举为

$$1 \le f_1 < \cdots < f_{n-r} \le n.$$

对公式 (1.4.1) 代入 $b_k = 0$ 给出齐次线性方程组的通解, 其中变元 $X_{f_1}, \ldots, X_{f_{n-r}}$ 可以任意赋值, 其余对应到主元的变元则由这些值唯一确定. 因此对每个 $1 \le i \le n-r$ 都可以取 $\boldsymbol{v}_i \in F^n$ 为齐次线性方程组的解, 使得

- ⋆ 它的第 f_i 个分量为 1,
- ⋆ 当 $1 \le j \le n-r$ 而 $j \ne i$ 时它的第 f_j 个分量为 0,
- ⋆ 其余分量由上述资料和 (1.4.1) 唯一确定.

若 $\boldsymbol{x} \in F^n$ 是齐次方程组的任意解, 则上述讨论表明 $\boldsymbol{x} = \sum_{i=1}^{n-r} t_i \boldsymbol{v}_i$, 其中的 $t_i \in F$ 是且仅能是 $\boldsymbol{x}$ 的第 f_i 个分量, 因此它们由 $\boldsymbol{x}$ 唯一确定. □

进一步, 我们还可以问解集 $T^{-1}(\mathbf{0})$ 的参数个数 $n-r$ 能否由 T 内在地刻画. 基于几何直观, 答案理应是肯定的. 我们现在有充分的动机来研究:

⋆ 带有加法和来自域 F 的纯量乘法这两种运算的代数结构, 例如 F^n, 或者先前提到的 $\mathrm{im}(T)$ 和 $T^{-1}(\mathbf{0})$;

⋆ 保持这种代数结构的映射, 例如之前考虑的 $T : F^n \to F^m$.

这就将向量空间的抽象理论推到了聚光灯下. 稍后展开的理论将会表明, 此一观点和 Gauss–Jordan 消元法是相互交融的.

4.2 向量空间

域 F 上的向量空间又称为线性空间. 这是具有加法, 纯量乘法[1)]运算以及零元的一种代数结构; 加法是空间上的二元运算, 纯量乘法则涉及域 F 的元素在 V 上的作用.

定义 4.2.1 域 F 上的 **F-向量空间**简称向量空间, 这是指资料 $(V, +, \cdot, 0_V)$, 其中 V 是集合, $0_V \in V$, 而 $+ : V \times V \to V$ 和 $\cdot : F \times V \to V$ 分别写作 $(u, v) \mapsto u + v$ 和 $(t, v) \mapsto t \cdot v$ 的形式, 使得以下条件成立.

1. 加法满足以下条件:
   - ▷ **结合律** $(u + v) + w = u + (v + w)$;
   - ▷ **幺元性质** $v + 0_V = v = 0_V + v$;
   - ▷ **交换律** $u + v = v + u$;
   - ▷ **加法逆元** 对所有 v 皆存在 $-v$ 使得 $v + (-v) = 0_V$.
2. 纯量乘法 $t \cdot v$ 也简写为 tv, 它满足以下条件:
   - ▷ **结合律** $s \cdot (t \cdot v) = (st) \cdot v$;
   - ▷ **幺元性质** $1 \cdot v = v$, 此处 1 代表 F 的乘法幺元.
3. 纯量乘法对加法满足:
   - ▷ **分配律之一** $(s + t) \cdot v = sv + tv$;
   - ▷ **分配律之二** $s \cdot (u + v) = su + sv$.

其中 u, v, w (或 s, t) 代表 V (或 F) 中的任意元素. 不致混淆时, 我们也简记 0_V 为 0, 将 $u + (-v)$ 写作 $u - v$, 并以 V 总括资料 $(V, +, \cdot, 0)$.

向量空间 V 的元素也称为其中的**向量**. 定义中的 0_V 又称为 V 的零元, 或**零向量**. 与向量相对, 域 F 的元素则称为**纯量**.

定义的模式和环的定义 3.1.1 明显相似. 由之推出的以下几条性质也是出于同样理路.

[1)] 在一些教材中, 纯量乘法又称为数乘或标量乘法.

- ⋆ 零元 0_V 由相应的幺元性质唯一确定, 所以资料中的 0_V 其实可以略去, 要求存在满足该性质的元素即可.
- ⋆ 加法满足消去律: $u+v=u'+v$ 蕴涵 $u=u'$. 作为推论, 向量 v 的加法逆元 $-v$ 唯一.
- ⋆ 逆元的唯一性也蕴涵 $-(-v)=v$.
- ⋆ 对于任何 $t\in F$ 都有 $t\cdot 0_V=0_V$, 这是对 $t\cdot(0_V)=t\cdot(0_V+0_V)=t\cdot 0_V+t\cdot 0_V$ 应用加法消去律的结论.
- ⋆ 我们有恒等式 $0\cdot v=0$; 等号左边的 $0\in F$, 而右边的 $0=0_V\in V$. 这是缘于 $0\cdot v=(0+0)\cdot v=0\cdot v+0\cdot v$ 和消去律.
- ⋆ 我们有 $(-1)\cdot v=-v$. 这是缘于 $(-1)\cdot v+v=(-1)\cdot v+1\cdot v=(-1+1)\cdot v=0\cdot v=0$.

定义 4.2.2 设 V 为 F-向量空间. 如果 V 的子集 V_0 包含 0, 而且在加法和纯量乘法运算下封闭, 则 $(V_0,+,\cdot,0)$ 也是 F-向量空间, 称之为 V 的**子空间**.

子空间的定义中未要求 V_0 对取逆 $v\mapsto -v$ 封闭, 这是因为 $-v=(-1)\cdot v$, 所以要求纯量乘法的封闭性已经足够.

练习 4.2.3 给定 F-向量空间 V, 请验证任意多个子空间的交仍然是 V 的子空间.

例 4.2.4 (零空间) 最平凡的向量空间是 $\{0\}$, 它是所有 F-向量空间的子空间.

向量空间理论也有直观的几何实例, 关乎中学数学里熟悉的平面与空间向量.

例 4.2.5 (向量几何) 平面向量对加法和纯量乘法构成 $\mathbb{R}$-向量空间, 空间向量亦同. 这是向量空间最直观的例子. 在未取定坐标的情况下, 向量的加法是按照平行四边形法则确定的, 而向量的纯量乘法则是带方向的伸缩, 负号代表反向; 相关的几何图像在本章开头已有回顾. 为了用代数语言表达角度和长度等几何概念, 还需要引入内积的概念, 这将是第九章的主题.

例 4.2.6 (域本身作为向量空间) 域 F 相对于域的加法和乘法 (在此扮演纯量乘法的角色) 构成 F-向量空间, 这是最简单的非零 F-向量空间, 它在后续章节还会反复现身.

例 4.2.7 设 $n\in\mathbb{Z}_{\geq 0}$. 按以下方式赋予 F^n 向量空间的结构

$$\begin{aligned}(x_1,\dots,x_n)+(y_1,\dots,y_n)&:=(x_1+y_1,\dots,x_n+y_n),\quad x_i,y_i\in F;\\ t(x_1,\dots,x_n)&:=(tx_1,\dots,tx_n),\quad t\in F.\end{aligned}$$

其中的零元取为 $\mathbf{0}:=(0,\dots,0)$. 这是 §4.1 讨论的例子. 定义所需的性质都是明显的. 注意到 F^0 规定为零空间, 而 F^1 无非是例 4.2.6 介绍的空间 F.

如果在 F^n 的定义中具体取 $F=\mathbb{R}$ 和 $n=2$ (或 $n=3$), 则 $\mathbb{R}^2$ (或 $\mathbb{R}^3$) 可以等同于坐标化的平面 (或空间). 中学数学所出现的平面 (或空间) 向量带有始点和终点, 写作诸如 $\overrightarrow{PQ}$ 的形式. 一旦在平面 (或空间) 中选定一点作为原点, 则总是可以将所有向量的始点平移到原点, 从而化约到例 4.2.5 的情境. 若进一步建立坐标系[2)], 便化约到此处讨论的 $\mathbb{R}^2$ 和 $\mathbb{R}^3$.

例 4.2.8 域 F 上的多项式环 $F[X]$ 相对于纯量乘法

$$t\cdot\sum_{n\geq 0}c_nX^n:=\sum_{n\geq 0}tc_nX^n,\quad t\in F$$

和 $F[X]$ 本身的加法构成 F-向量空间. 推而广之, 任意个变元的多项式环 $F[X,Y,\ldots]$ 也都是 F-向量空间.

以下介绍两种从已有的向量空间构造新空间的抽象途径.

例 4.2.9 (向量空间的直积) 首先回忆 (2.3.1) 的定义: 对任意一族集合 V_i, 其中下标 i 遍历给定的集合 I (暂时要求非空), 积集 $\prod_{i\in I}V_i$ 的元素是以 I 为下标的元素组 $(v_i)_{i\in I}$, 其中 $v_i\in V_i$. 如果每个 V_i 都带有 F-向量空间的结构, 则 $\prod_{i\in I}V_i$ 也有自然的 F-向量空间结构, 方法是逐分量地定义运算为

$$(v_i)_i+(w_i)_i=(v_i+w_i)_i,\quad t(v_i)_i=(tv_i)_i,$$

下标 i 遍历 I 而 $t\in F$. 我们称此向量空间 $\prod_{i\in I}V_i$ 为 $(V_i)_{i\in I}$ 的直积. 所需的结合律等诸般性质全部都能化到每个 V_i 上来验证. 思路和环的直积 (例 3.1.13) 并无二致.

如果取所有 V_i 都是同一个向量空间 V, 则和积集的情形类似, 我们将对应的 $\prod_{i\in I}V$ 记为 V^I. 回忆到在关于集合论的讨论中, 我们在 $I=\varnothing$ 时不无理由地规定 V^I 为独点集. 建立在独点集上的向量空间结构能且仅能是零空间, 因此我们规定当 $I=\varnothing$ 时 $V^I:=\{0\}$. 这一切可以视为例 4.2.7 的推广: 不难看出 F^n 即是 $F^{\{0,\ldots,n-1\}}$.

例 4.2.10 (向量空间的直和) 例 4.2.9 介绍的直积 $\prod_{i\in I}V_i$ 有一个更常用的子空间, 称为 $(V_i)_{i\in I}$ 的**直和**, 定义为

$$\bigoplus_{i\in I}V_i:=\left\{(v_i)_i\in\prod_{i\in I}V_i:\text{至多仅有有限个 }i\in I\text{ 使得 }v_i\neq 0\right\}.$$

留意到如果 $(v_i)_{i\in I}$ 和 $(w_i)_{i\in I}$ 都仅有至多有限个分量非零, 则 $(v_i+w_i)_{i\in I}$ 亦然, 因为两个有限集的并仍有限. 因此 $\bigoplus_{i\in I}V_i$ 对加法封闭. 此外它显然也对纯量乘法封闭, 并且包含 $0=(0)_{i\in I}$. 于是 $\bigoplus_{i\in I}V_i$ 确实是 $\prod_{i\in I}V_i$ 的子空间.

每个 V_i 都自然地嵌入为 $\bigoplus_{j\in I}V_j$ 作为子空间, 方式是映 $v\in V_i$ 为 $(v_j)_{j\in I}$, 其中 $v_i:=v$ 而 $j\neq i$ 时 $v_j:=0$. 考察各个分量, 立见在这些嵌入之下有

$$i\neq j\implies V_i\cap V_j=\{0\}.$$

[2)] 几何直观中的平面或空间既无指定的原点, 又无指定的坐标轴, 所以 $\mathbb{R}^n$ 对此是个不尽合身的模型, 以后讨论仿射空间时还会回到这个问题.

如果取所有 V_i 都是同一个向量空间 V, 则对应的直和表作 $V^{\oplus I}$, 以区别于直积 V^I. 当 I 有限时, 至多有限个分量非零的条件自动成立, 这时 $\prod_{i\in I} V_i = \bigoplus_{i\in I} V_i$. 本书涉及的直和的以这类情形居多.

约定 4.2.11 规定空直和与空直积 (相当于 $I=\varnothing$ 情形) 为零空间 $\{0\}$. 这与 §2.3 的规定兼容.

直和是向量空间理论中的一则基本构造. 关于直和的进一步讨论将在 §4.10 接续.

4.3 矩阵及其运算

矩阵在 §1.3 关于消元法的讨论中已经出现. 当时我们对矩阵元在哪个数系取值含糊其词. 现在既然有了域的概念, 便可以清楚地定义一般的矩阵.

定义 4.3.1 (域上的矩阵) 设 $m,n\in\mathbb{Z}_{\geq 1}$. 域 F 上的 $m\times n$ **矩阵**是指如下的资料

$$\boldsymbol{A} = (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{pmatrix} \begin{matrix} \\ \\ \text{第 } i \text{ 行} \\ \\ \\ \text{第 } j \text{ 列} \end{matrix}$$

其中 $a_{ij}\in F$, 称为矩阵 $\boldsymbol{A}$ 的第 (i,j) 个矩阵元或 (i,j)-项, 而 $n\times n$ 矩阵也称为 n 阶**方阵**.

今后我们将 F 上的所有 $m\times n$ 矩阵 (m 行 n 列) 构成的集合记为 $\mathrm{M}_{m\times n}(F)$.

按惯例, 我们经常将矩阵 $\boldsymbol{A}=(a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ 简记为 $(a_{ij})_{i,j}$. 严格地说, F 上的 $m\times n$ 矩阵是集合 $F^{\{1,\dots,m\}\times\{1,\dots,n\}}$ 的元素; 这种原教旨主义的解读当然是庸人自扰.

现在赋予集合 $\mathrm{M}_{m\times n}(F)$ 作为向量空间所需的运算.

▷ **加法** 对 $\mathrm{M}_{m\times n}(F)$ 的任两个元素 $\boldsymbol{A}=(a_{ij})_{i,j}$ 和 $\boldsymbol{B}=(b_{ij})_{i,j}$, 定义

$$\boldsymbol{A}+\boldsymbol{B} := (a_{ij}+b_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}.$$

▷ **纯量乘法** 对任意 $t\in F$ 和 $\boldsymbol{A}=(a_{ij})_{i,j}\in\mathrm{M}_{m\times n}(F)$, 定义

$$t\cdot\boldsymbol{A} = t\boldsymbol{A} := (ta_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}.$$

给定 m 和 n, 对应尺寸的**零矩阵**定义为

$$\mathbf{0}_{m\times n} := (0)_{i,j} \in \mathrm{M}_{m\times n}(F).$$

命题 4.3.2 这些运算使得 $\mathrm{M}_{m\times n}(F)$ 成为 F-向量空间. 其中的零元为零矩阵, 而矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}$ 的加法逆元 $-\boldsymbol{A}$ 为 $(-a_{ij})_{i,j}$.

证明 理应是明显的. 一切运算性质都对每个矩阵元化约到 F 中来验证. □

定义 $m\times n$ 矩阵的减法为 $\boldsymbol{A}-\boldsymbol{B}:=\boldsymbol{A}+(-\boldsymbol{B})$, 它的 (i,j) 矩阵元是 $a_{ij}-b_{ij}$.

约定 4.3.3 (行向量与列向量) 对于特例 $n=1$, 向量空间 $\mathrm{M}_{m\times 1}(F)$ 化约为例 4.2.7 讨论的 F^m; 出于直观的理由, $\mathrm{M}_{m\times 1}(F)$ 的元素也称为 m 维的**列向量**. 同样道理, 向量空间 $\mathrm{M}_{1\times n}(F)$ 可以等同于 F^n, 其元素称为 n 维的**行向量**. 当 $n=m=1$, 空间 $\mathrm{M}_{1\times 1}(F)$ 无非就是 F 自身.

作为向量空间, $\mathrm{M}_{m\times n}(F)$ 和 F^{mn} 似乎没有实质区别, 然而矩阵的特色在于除了加法和纯量乘法之外, 它们还具备乘法运算.

定义 4.3.4 (矩阵乘法) 矩阵乘法是按以下方式定义的映射

$$\begin{aligned}\mathrm{M}_{m\times n}(F)\times\mathrm{M}_{n\times r}(F)&\longrightarrow\mathrm{M}_{m\times r}(F)\\(\boldsymbol{A},\boldsymbol{B})&\longmapsto\boldsymbol{AB};\end{aligned}$$

若 $\boldsymbol{A}=(a_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$, $\boldsymbol{B}=(b_{jk})_{\substack{1\le j\le n\\1\le k\le r}}$, 则 $\boldsymbol{AB}=(c_{ik})_{\substack{1\le i\le m\\1\le k\le r}}$, 其中

$$c_{ik}:=\sum_{j=1}^{n}a_{ij}b_{jk}=\underbrace{\begin{pmatrix}a_{i1}&\cdots&a_{in}\end{pmatrix}}_{\text{第 } i \text{ 行}}\underbrace{\begin{pmatrix}b_{1k}\\\vdots\\b_{nk}\end{pmatrix}}_{\text{第 } k \text{ 列}}.$$

注意: 只有行数和列数合乎规格的矩阵才能相乘. 按惯例, $\boldsymbol{AB}$ 有时也写成 $\boldsymbol{A}\cdot\boldsymbol{B}$.

定义 n 阶**单位矩阵**为

$$\mathbf{1}_{n\times n}:=\begin{pmatrix}1&&\\&\ddots&\\&&1\end{pmatrix}\in\mathrm{M}_{n\times n}(F)$$

其中对角线矩阵元全为 1, 其余留白部分全为 0. 请读者循定义动手验证单位矩阵对所有 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$ 都有类似于乘法幺元的性质:

$$\mathbf{1}_{m\times m}\cdot\boldsymbol{A}=\boldsymbol{A}=\boldsymbol{A}\cdot 1_{n\times n}.$$

命题 4.3.5 矩阵乘法满足以下性质:

▷ **结合律** $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$;

▷ **分配律** $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$, $(\boldsymbol{B}+\boldsymbol{C})A=\boldsymbol{BA}+\boldsymbol{CA}$;

▷ **线性** $\boldsymbol{A}(t\boldsymbol{B})=t(\boldsymbol{AB})=(t\boldsymbol{A})\boldsymbol{B}$;

其中 $t\in F$ 和矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 是任意的, 前提是矩阵的行数和列数使运算有意义.

证明 对 $(\boldsymbol{AB})\boldsymbol{C}$ 和 $\boldsymbol{A}(\boldsymbol{BC})$ 的每个矩阵元进行考察, 可见结合律相当于说

$$\sum_k\left(\sum_j a_{ij}b_{jk}\right)c_{kl}=\sum_j a_{ij}\left(\sum_k b_{jk}c_{kl}\right),$$

下标 i, j, k, l 的范围取决于矩阵尺寸. 根据求和符号的基本操作, 交换 $\sum_k$ 和 $\sum_j$ 以后, 问题归结为证

$$(a_{ij}b_{jk})c_{kl}=a_{ij}(b_{jk}c_{kl}).$$

然而这无非是 F 本身的乘法结合律. 同理, 分配律归结为 F 本身的分配律

$$a_{ij}(b_{jk}+c_{jk})=a_{ij}b_{jk}+a_{ij}c_{jk}.$$

最后一则性质归结为 F 中的等式 $a_{ij}(tb_{jk})=t(a_{ij}b_{jk})=(ta_{ij})b_{jk}$, 而这又归结为 F 的乘法交换律. □

尽管推论 4.6.3 将包含下述结果, 眼下先就矩阵定义来验证是有益的, 细节也毫无困难.

练习 4.3.6 验证 $\mathrm{M}_{n\times n}(F)$ 相对于矩阵加法和乘法成环, 它以零矩阵 $\mathbf{0}_{n\times n}$ 为零元, 以单位矩阵 $\mathbf{1}_{n\times n}$ 为乘法幺元. 〈提示〉 前述诸性质的综合.

练习 4.3.7 矩阵给出非交换环的自然例子. 请动手验证以下等式以说明 $\mathrm{M}_{2\times 2}(F)$ 并非交换环.

$$\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}0&1\\1&1\end{pmatrix},$$
$$\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}.$$

注记 4.3.8 (环上的矩阵) 如果将域 F 换成一般的环 R, 则以上定义和大部分性质仍然通行无阻; 特别地, 我们有集合 $\mathrm{M}_{m\times n}(R)$ 以及矩阵之间的加法和乘法, 并且使 $\mathrm{M}_{n\times n}(R)$ 成环. 唯一成问题的是命题 4.3.5 的等式 $\boldsymbol{A}(t\boldsymbol{B})=t(\boldsymbol{AB})$ (其中 $t\in R$): 从证明不难看出, 障碍在于 $t\in R$ 对乘法未必能和 R 中的其他元素交换. 如果我们要求 R 是交换环, 则 $\boldsymbol{A}(t\boldsymbol{B})=t(\boldsymbol{AB})=(t\boldsymbol{A})\boldsymbol{B}$ 仍然成立.

此外, 在非交换环上必须区别 $t\in R$ 对矩阵的左乘 $t\boldsymbol{A}$ 和右乘 $\boldsymbol{A}t$.

一如 §1.3 所述, 线性方程组的布列和消元在一般的域 F 上也能用 F 上的矩阵来表述. 矩阵乘法还给出考量线性方程组的另一种视角: 对于取在 F 中的变元 $X_1,\dots,X_n$, 我们有

$$\begin{aligned} a_{11}X_1+\cdots+a_{1n}X_n &= b_1\\ a_{21}X_1+\cdots+a_{2n}X_n &= b_2\\ \vdots\quad &\quad\vdots\\ a_{m1}X_1+\cdots+a_{mn}X_n &= b_m \end{aligned} \iff \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} X_1\\ \vdots\\ X_n \end{pmatrix} = \begin{pmatrix} b_1\\ \vdots\\ b_m \end{pmatrix}.$$

取方程组的系数矩阵 $\boldsymbol{A}:=(a_{ij})_{i,j}\in \mathrm{M}_{m\times n}(F)$, 并且定义列向量

$$\boldsymbol{x}:=\begin{pmatrix} X_1\\ \vdots\\ X_n \end{pmatrix},\quad \boldsymbol{b}:=\begin{pmatrix} b_1\\ \vdots\\ b_m \end{pmatrix},$$

则原方程组等价于以 $\boldsymbol{x}$ 为变量的矩阵方程

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}.$$

这是矩阵乘法和线性方程组的直接联系, 矩阵写法显然胜在简洁, 而这也为矩阵乘法的定义提供了一部分的解释.

我们更愿意从线性映射的视角理解矩阵. 一旦在 §4.5 明确了矩阵和线性映射的关系, 则无论是命题 4.3.5 的矩阵乘法性质, 练习 4.3.6 或线性方程组的矩阵写法 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 都会得到自然的诠释. 在此之前, 有必要先引进基和维数的抽象概念.

4.4 基和维数

对于经典的平面向量或空间向量, 我们熟知一旦选定坐标系, 所有向量都能按坐标唯一地展开. 将此思路扩及一般的向量空间, 就引向了基的概念.

向量空间的基并不唯一, 通常也没有标准的取法. 在许多应用中[3], 适当选基, 自由换基还是解决问题的关键手段. 阐明基与基之间如何变换是稍后的重要任务.

严谨的解释需要一系列的准备工作. 设 S 为 F-向量空间 V 的子集, 或有限或无穷. 我们称形如

$$a_1v_1+\cdots+a_mv_m,\qquad \begin{aligned} m&\in\mathbb{Z}_{\geq 0},\\ a_i&\in F,\\ v_i&\in S \end{aligned} \tag{4.4.1}$$

[3] 例如中学物理力学中的静态平衡.

的向量为 S 中的元素的**线性组合**, 这是从 S 出发, 通过向量空间的所有运算所能得到的一切元素. 注意到上式取 $m=0$ 对应到 "空和", 约定为 0. 记

$$\langle S\rangle := \{S \text{ 中的元素的线性组合}\}.$$

按定义立见 $\langle S\rangle$ 是 V 的子空间. 事实上, $\langle S\rangle$ 是包含 S 的最小子空间 — 包含 S 的子空间必然也包含 $\langle S\rangle$. 这同样是线性组合定义的直接结论. 极端情形是 $\langle\varnothing\rangle=\{0\}$. 当 S 是有限集 $\{s_1,\dots,s_m\}$ 时, 我们也将 $\langle S\rangle$ 写成 $\langle s_1,\dots,s_m\rangle$.

职是之故, 我们也称 $\langle S\rangle$ 为 S 的线性包, 或 S 张成的子空间. 在探讨形如 (4.4.1) 的线性组合时, 更简洁的写法是让 S 的元素自为下标, 从而将 S 中的元素的线性组合写作

$$\sum_{s\in S} a_s s$$

的形式, 其中的系数 $a_s\in F$, 默认至多仅有限个非零, 因此上式总默认为有限和; 今后凡是写下诸如 $\sum_s a_s s$ 的公式时, 总是如此假设.

定义 4.4.1 设 S 为 F-向量空间 V 的子集.

- ⋆ 如果 $\langle S\rangle = V$, 则称 S **生成** V, 或称 S 为 V 的一族生成元, 或简称 S 为生成系.
- ⋆ 称形如
$$\sum_{s\in S} a_s s = 0$$
的等式为 S 中的线性关系; 如果所有系数 a_s 全为 0, 则称此关系平凡, 否则称为非平凡. 若 S 中存在非平凡的线性关系, 则称 S **线性相关**, 否则称 S **线性无关**.
- ⋆ 若 S 是线性无关的生成系, 则称 S 是 V 的一组**基**.

一种类似的说法是: 若 V 中的一列元素 $v_1,\dots,v_k$ 满足 $\sum_{i=1}^k a_i v_i = 0 \iff \forall i,\ a_i=0$, 则称 $v_1,\dots,v_k$ 线性无关, 否则称为线性相关. 唯一的差别在于此处的 $v_1,\dots.v_k$ 容许重复, 而有重复的情形当然是线性相关的.

例 4.4.2 以下性质都是定义的直接操演.

- ⋆ 空集按规定是线性无关的.
- ⋆ 若 S 含零向量 0, 则它自动线性相关.
- ⋆ 独点集 $\{v\}$ 线性无关当且仅当 $v\neq 0$.
- ⋆ 向量 v 和 w 线性相关当且仅当存在不全为 0 的常数 a,b 使得 $av+bw=0$, 当且仅当两者成比例: 存在 $c\in F$ 使得 $v=cw$ (取 $c=-b/a$) 或 $w=cv$ (取 $c=-a/b$), 视 a,b 何者非零而定.

找出向量空间 V 的一个生成系对于理解 V 显然有帮助. 为何考虑线性无关子集? 若子集 S 线性无关, 则任何线性组合 $v=\sum_{s\in S}a_ss$ 中的诸系数 a_s 都由 v 唯一确定, 这是因为

$$\sum_{s\in S}a_ss=\sum_{s\in S}b_ss\iff\sum_{s\in S}(a_s-b_s)s=0.$$

所以 S 是 V 的基当且仅当所有 v 都能够表成线性组合 $v=\sum_{s\in S}a_ss$, 而且系数 $(a_s)_{s\in S}$ 由 v 唯一确定. 这族系数可以合理地称为 v 在这组基下的坐标或分量.

向量空间 V 的所有生成族 (或线性无关子集) 可以按集合包含关系 $\subset$ 比较大小, 由此可以谈论其中的极大或极小元; 参见定义 2.4.4. 这就引向了基的另一种刻画.

引理 4.4.3 对于 F-向量空间 V 的任意子集 S, 我们有

$$S\text{ 是极小生成系}\iff S\text{ 是基}\iff S\text{ 是极大线性无关子集}.$$

证明 (第一个等价) 设 S 是极小生成系. 以下验证 S 线性无关, 从而说明 S 为基. 设 S 中存在非平凡线性关系 $a_1s_1+\cdots+a_ms_m=0$, 其中 $s_i\in S$ 相异; 不失一般性, 设 $a_1\neq0$. 那么

$$s_1=-a_1^{-1}\left(\sum_{i=2}^{m}a_is_i\right);$$

回忆线性组合的定义, 可知上式蕴涵 $S\smallsetminus\{s_1\}$ 和 S 生成同一个向量空间, 这同 S 的极小性矛盾.

反之, 设 S 为基, 以下来证明 S 是极小生成系. 设若不然, 则存在 $s\in S$ 使得 $S\smallsetminus\{s\}$ 也是生成系, 因此 s 可以表作线性组合 $\sum_{t\in S\smallsetminus\{s\}}a_tt$, 从而得到 S 中的非平凡线性关系 $s-\sum_{t\in S\smallsetminus\{s\}}a_tt=0$, 矛盾.

(第二个等价) 设 S 是极大线性无关子集, 以下验证 S 是生成系, 从而说明 S 为基. 对于任何 $v\in V$, 若 $v\in S$ 则当然有 $v\in\langle S\rangle$. 若 $v\notin S$ 则 $S\cup\{v\}$ 线性相关, 故有非平凡的线性关系

$$a_vv+\sum_{s\in S}a_ss=0.$$

此时必有 $a_v\neq0$, 否则将有 S 中的非平凡线性关系 $\sum_{s\in S}a_ss=0$, 矛盾. 于是 $v=-a_v^{-1}\sum_{s\in S}a_ss\in\langle S\rangle$. 综上可得 $V=\langle S\rangle$.

反之, 设 S 为基, 现证 S 是极大线性无关子集. 对于任意 $v\notin S$, 将它表为基的线性组合 $v=\sum_{s\in S}a_ss$. 非平凡线性关系 $v-\sum_{s\in S}a_ss=0$ 蕴涵 $S\cup\{v\}$ 线性相关. 由此可见 S 确实极大. □

约定 4.4.4 按以上定义, 有限维向量空间 V 的基是一类特殊的子集, 基的元素不计顺序. 如果进一步为基的元素排序, 譬如表作 $v_1,v_2,\ldots$ 等, 则称之为**有序基**以资区别.

尽管基不唯一, 对于某些特定的向量空间却有标准的选法.

例 4.4.5 (F^n 的标准基) 向量空间 F^n 具有标准有序基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$, 其中

$$\boldsymbol{e}_i := (0,\dots,\underbrace{1}_{i\text{-分量}},\dots,0),\quad i=1,\dots,n.$$

理由直截了当: 任何 $\boldsymbol{v}=(x_1,\dots,x_n)\in F^n$ 都可以表成

$$\begin{aligned}\boldsymbol{v}&=(x_1,0,\dots,0)+(0,x_2,0,\dots,0)+\cdots+(0,\dots,0,x_n)\\&=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n;\end{aligned}$$

反过来说, 如果 $\boldsymbol{v}$ 写作以上形式, 则 x_i 是 $\boldsymbol{v}$ 的第 i 个分量, 因此表法是唯一的.

例 4.4.6 ($\mathrm{M}_{m\times n}(F)$ 的标准基) 推而广之, 矩阵空间 $\mathrm{M}_{m\times n}(F)$ 具有标准基 $\{\boldsymbol{E}_{ij}\}_{i,j}$, 其中 $\boldsymbol{E}_{ij}$ 是第 (i,j) 个矩阵元为 1, 其他矩阵元全为 0 的 $m\times n$ 矩阵, 此处 $1\le i\le m$ 而 $1\le j\le n$. 任何矩阵 $\boldsymbol{A}$ 都能唯一地展开为

$$\boldsymbol{A}=\sum_{\substack{1\le i\le m\\1\le j\le n}}a_{ij}\boldsymbol{E}_{ij};$$

系数 a_{ij} 正是 $\boldsymbol{A}$ 的第 (i,j) 个矩阵元. 矩阵 $\boldsymbol{E}_{ij}$ 之间的乘法有简单的描述: 代入定义 4.3.4 仔细运算, 可见

$$\boldsymbol{E}_{ij}\boldsymbol{E}_{k\ell}=\begin{cases}\boldsymbol{0} & j\ne k,\\ \boldsymbol{E}_{i\ell} & j=k.\end{cases}$$

因为这些矩阵构成基, 只要承认命题 4.3.5 中关于矩阵乘法的基本性质, 则以上公式反过来确定了一般矩阵的乘法.

例 4.4.7 (单项式基) 由于 F 上的所有一元多项式都能唯一地表作有限和 $\sum_{n\ge0}a_nX^n$, 向量空间 $F[X]$ 有基 $1,X,X^2,\dots$, 称为单项式基; 注意此基有可数无穷多个元素. 推而广之, 多元多项式空间 $F[X,Y,\dots]$ 也有形如 $X^aY^b\cdots$ 的元素构成的单项式基.

算法 4.4.8 (计算极大线性无关子集) 对于 F-向量空间 F^n 中的元素 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_k$, 其中 $n,k\in\mathbb{Z}_{\ge1}$, Gauss–Jordan 消元法可以用来判断它们是否线性无关. 第一步是将每个 $\boldsymbol{v}_j$ 表为列向量

$$\boldsymbol{v}_j=\begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix},\quad j=1,\dots,k.$$

为了判定线性相关与否, 当然要研究等式 $\sum_{j=1}^k x_j\boldsymbol{v}_j=0$. 将等式按 n 个分量展开, 则这相当于判定齐次线性方程组

$$\begin{aligned}a_{11}X_1+\cdots+a_{1k}X_k&=0\\ &\vdots\quad\ \vdots\\ a_{n1}X_1+\cdots+a_{nk}X_k&=0\end{aligned}$$

是否有不全为 0 的解. 这点可以通过对系数矩阵

$$\boldsymbol{A} := (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq k}} \xlongequal{\text{按列向量展开}} \left(\boldsymbol{v}_1 \;\middle|\; \cdots \;\middle|\; \boldsymbol{v}_k\right) \in \mathrm{M}_{n\times k}(F)$$

消元来判定: 记消元法给出的简化行梯矩阵为 $\boldsymbol{A}'$, 则 S 线性无关当且仅当 $\boldsymbol{A}'$ 的主元个数 r 满足 $r = k$.

同理, 一个列向量 $\boldsymbol{w} \in F^n$ 能表成 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ 的线性组合当且仅当增广矩阵

$$\left(\boldsymbol{A} \;\middle|\; \boldsymbol{w}\right) = \left(\boldsymbol{v}_1 \;\middle|\; \cdots \;\middle|\; \boldsymbol{v}_k \;\middle|\; \boldsymbol{w}\right) \in \mathrm{M}_{n\times(k+1)}(F)$$

对应的线性方程组有解, 消元法可以判定它是否有解, 并且在有解的情形给出通解.

由此还可以萃取更多的信息. 将简化行梯矩阵 $\boldsymbol{A}'$ 的列向量记为 $\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots$, 再将其主列排序为

$$1 \leq j_1 < \cdots < j_r \leq k;$$

对应的列向量必然形如

$$\boldsymbol{v}'_{j_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ \boldsymbol{v}'_{j_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\ \boldsymbol{v}'_{j_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \ldots \quad (\text{共 } r \text{ 个}).$$

我们现在将信息转译回 $\boldsymbol{A}$, 来说明以下事实:

> $\boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_r}$ 线性无关, 而且给出 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ 的极大线性无关子集; 事实上它们是 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ 的基.

1. 首先说明它们线性无关. 诚然, $\boldsymbol{A}$ 行等价于 $\boldsymbol{A}'$, 而初等行变换是逐列操作的, 所以只看主列便有

$$\left(\boldsymbol{v}_{j_1} \;\middle|\; \cdots \;\middle|\; \boldsymbol{v}_{j_r}\right) \quad \text{行等价于} \quad \left(\boldsymbol{v}'_{j_1} \;\middle|\; \cdots \;\middle|\; \boldsymbol{v}'_{j_r}\right) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & \end{pmatrix}$$

   其中矩阵留白部分为零.

2. 其次说明 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle = \langle \boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_r} \rangle$. 诚然, 令 $1 \leq \ell \leq k$ 而 $\ell \notin \{j_1, \ldots, j_r\}$. 列向量 $\boldsymbol{v}'_\ell$ 能表为它左侧的主列向量 $\boldsymbol{v}'_{j_1}, \boldsymbol{v}'_{j_2}, \ldots$ 的线性组合, 这也相当于说增广

矩阵 $\left(\boldsymbol{v}'_{j_1}|\cdots|\boldsymbol{v}'_{j_r}|\boldsymbol{v}'_\ell\right)$ 对应的方程组有解. 然而和上一步相同的理由导致

$$(\boldsymbol{v}_{j_1}|\cdots|\boldsymbol{v}_{j_r}|\boldsymbol{v}_\ell) \quad \text{行等价于} \quad \left(\boldsymbol{v}'_{j_1}|\cdots|\boldsymbol{v}'_{j_r}|\boldsymbol{v}'_\ell\right) = \left(\begin{array}{ccc|c} 1 & & & * \\ & \ddots & & \vdots \\ & & 1 & * \\ & & & \end{array}\right),$$

其中 $*$ 代表 F 的某些元素[4]. 特别地, 两边对应的方程组同解. 这就说明了 $\boldsymbol{v}_\ell \in \langle \boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_r}\rangle$. 换句话说, $\boldsymbol{v}_\ell$ 总能表成 $\boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_r}$ 的线性组合.

3. 最后, 综合前两条可见 $\boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_r}$ 是 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\rangle$ 的基, 故是此空间的极大线性无关子集 (引理 4.4.3). 因此它在集合 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ 中自然也是极大线性无关子集.

顺带注意到这些性质对域扩张不敏感: 若 F 是域 E 的子域, 则这些问题无论置于 E 或 F 中考虑, 结论皆同, 这是由消元法的本质决定的.

现在回到抽象理论. 相对于例 4.2.10 介绍的直和操作, 基有简单的描述, 由无交并给出.

命题 4.4.9 *考虑一族 F-向量空间 V_i, 带有给定的基 B_i, 其中 i 遍历给定的集合 I. 将每个 V_i 按照例 4.2.10 所述方法嵌入为直和 $\bigoplus_{j\in I} V_j$ 的子空间, 相应地将 B_i 视同 $\bigoplus_{j\in I} V_j$ 的子集. 这些子集 B_i 两两无交, 而 $\bigsqcup_{i\in I} B_i$ 给出 $\bigoplus_{i\in I} V_i$ 的基.*

证明 由于 $i \neq j$ 蕴涵 $V_i \cap V_j = \{0\}$ 而基的元素总是非零, 故 $\{B_i\}_{i\in I}$ 在直和中两两无交; 记 B 为这些 B_i 在直和中的无交并.

按构造, 每个 $v \in \bigoplus_i V_i$ 都能够唯一地表作 $v = \sum_i v_i$, 其中 $v_i \in V_i$ 无非是 V 的第 i 个分量, 至多有限多个 v_i 非零. 通过将每个非零的 v_i 按基 B_i 展开, v 也能唯一地表作有限和

$$v = \sum_i \underbrace{\sum_{b\in B_i} c_b b}_{\in V_i} = \sum_{b\in B} c_b b$$

其中 $c_b \in F$, 仅对至多有限个 b 非零. □

留意到结论对直积并不成立; 直和中至多仅有限个分量非零的条件在论证中是必要的.

以下定理的证明虽然不难, 却需要一些集合论的背景知识, 此处述而不证. 感兴趣的读者可参考 [10, §6.4], 或见 §A.4.

[4] 计算机科学中所谓的通配符.

定义–命题 4.4.10 所有 F-向量空间 V 都有基; 事实上, 任意线性无关子集都可以扩充为基.

此外, V 的所有基都有相同的元素个数, 按 §2.9 介绍的基数来理解. 此一共同的基数称为 V 的**维数**, 记为 $\dim_F V$ 或 $\dim V$.

作为以上定义连同命题 4.4.9 的推论, 对任一族向量空间 $(V_i)_{i\in I}$ 皆有

$$\dim \bigoplus_{i\in I} V_i = \sum_{i\in I} \dim V_i, \tag{4.4.2}$$

右式应当理解为一族基数的和, 对应到集合 (基) 的无交并.

本书主要考虑的是以下定义的有限生成向量空间, 而定义–命题 4.4.10 的对应版本将在定义–命题 4.4.13 给出. 这时一切论证都将是构造性的, 具体操作归结为消元法.

定义 4.4.11 设 V 为 F-向量空间. 若存在有限的生成系, 则称 V 为**有限生成**的.

作为反面例子, 例 4.2.8 介绍的多项式空间 $F[X]$ 不是有限生成的, 因为有限多个多项式 $f_1,\dots,f_n \in F[X]$ 作线性组合后的次数绝不超过 $\max\{\deg f_1,\dots,\deg f_n\}$, 当然也就不可能生成 $F[X]$. 由于 $F[X]$ 有单项式基 $\{1, X, X^2, \dots\}$, 它的维数应当是可数无穷 $\aleph_0$.

从 V 的有限生成系出发, 不断移除多余的元素, 最终能得到极小生成系. 所以引理 4.4.3 说明有限生成向量空间总是存在有限基. 为了研究基的元素个数, 我们还需要以下精密的结果.

引理 4.4.12 设子集 $\{s_1,\dots,s_n\}$ 生成 F-向量空间 V, 则当 $m > n$ 时, 任意 m 个向量 $v_1,\dots,v_m \in V$ 都线性相关.

证明 思路是熟悉的消元法. 取系数 $a_{ij} \in F$, 其中 $1 \le i \le m$ 而 $1 \le j \le n$, 使得对所有 $1 \le i \le m$ 都有

$$v_i = a_{1i}s_1 + \cdots + a_{ni}s_n.$$

我们断言若 $(x_1,\dots,x_m) \in F^m$ 是 m 元齐次线性方程组

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1m}X_m &= 0 \\ \vdots \qquad & \quad \vdots \\ a_{n1}X_1 + \cdots + a_{nm}X_m &= 0 \end{aligned} \tag{4.4.3}$$

的一组解, 则 $x_1v_1 + \cdots + x_mv_m = 0$. 这是因为

$$\sum_{i=1}^{m} x_i v_i = \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n} a_{ji}s_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ji}x_i \right) s_j = 0;$$

这里用到换序求和法则 $\sum_i \sum_j = \sum_j \sum_i$. 于是问题化为证 $m > n$ 时方程组 (4.4.3) 有不全为 0 的解. 为此, 对其系数矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \qquad (m > n)$$

进行 Gauss–Jordan 消元法, 得出的简化行梯矩阵其主元个数 $\leq \min\{n, m\} < m$ (练习 1.3.5); 换言之, 解集至少包含一个自由参数. 这就说明 (4.4.3) 有不全为 0 的解. □

定义–命题 4.4.13 设 V 是有限生成的 F-向量空间, 则它有基. 事实上, V 的所有线性无关子集都能够扩充为基, 每组基的元素个数皆有限而且彼此相等. 此时任何一组基的元素个数称为 V 的**维数**, 记为 $\dim_F V$ 或 $\dim V$.

证明 既然 V 有一个有限生成系, 不妨设为 $w_1, \ldots, w_m$, 故引理 4.4.12 蕴涵线性无关子集的大小不能超过 m. 因此从 V 的任意线性无关子集 S 出发, 不断扩增, 最终必能得到 V 的极大无关子集 B 使得 $B \supset S$ 而 $|B| \leq m$. 引理 4.4.3 说明 B 是 V 的基.

设 B' 是 V 的任一组基, 则引理 4.4.12 说明 B' 有限, 而且 $|B'| \leq |B|$. 同理, $|B| \leq |B'|$. 因此每组基的元素个数皆相等. □

鉴于此, 有限生成向量空间更常被称为**有限维向量空间**, 本书今后也改用此术语.

注意到

$$\dim V = 0 \iff \varnothing \text{ 是 } V \text{ 的基} \iff V = \{0\}.$$

例 4.4.14 根据标准基的构造 (例 4.4.5, 4.4.6), 我们知道 F-向量空间 F^n 是 n 维的, 而 $\mathrm{M}_{m\times n}(F)$ 是 mn 维的.

推论 4.4.15 设 $\dim V = n$ 而 $v_1, \ldots, v_n \in V$. 当以下任一条件成立时, $\{v_1, \ldots, v_n\}$ 是 V 的基:

- ⋆ 它们线性无关;
- ⋆ 它们生成 V.

证明 应用引理 4.4.3. 对于线性无关的情形, $\{v_1, \ldots, v_n\}$ 有 n 个相异元素, 并且可以扩充为基; 对于生成系的情形, 它包含一组基作为子集. 然而任何基都恰有 n 个元素. □

推论 4.4.16 设 V 是有限维向量空间, 若 V_0 是 V 的子空间, 则 V_0 也是有限维的, 而且 $\dim V_0 \leq \dim V$, 等号成立当且仅当 $V_0 = V$.

证明 令 $n := \dim V$. 引理 4.4.12 说明在 V_0 中任何 $n+1$ 个向量都线性相关, 所以 V_0 必有极大线性无关子集, 即基, 其元素个数 $\leq n$.

若 $V = V_0$, 当然有 $\dim V_0 = \dim V$. 反之设 $\dim V_0 = n$ 成立, 取 $v_1, \ldots, v_n$ 为 V_0 的基. 根据推论 4.4.15, 它们既然线性无关, 自动是 V 的基, 故 $V = V_0$. □

4.5 线性映射

一如偏序集 (定义 2.4.3) 和环 (定义 3.2.1) 的情形所揭示的, 保结构的映射对于代数学的研究至关重要. 落实到向量空间情形, 结构是向量的加法和纯量乘法, 而线性映射便是保持这些运算的映射; 在代数学的框架中, 更合理的称呼兴许是 F-向量空间之间的同态.

定义 4.5.1 (线性映射) 设 V 和 W 是 F-向量空间. 若映射 $T : V \to W$ 满足

$$T(v_1 + v_2) = T(v_1) + T(v_2), \quad v_1, v_2 \in V,$$
$$T(tv) = tT(v), \quad t \in F,\ v \in V,$$

则称 T 为线性映射, 在一些场合也称线性变换或线性算子[5].

线性映射映零元为零元, 这是由于 $T(0) = T(0+0) = T(0) + T(0)$ 和加法的消去律蕴涵 $T(0) = 0$; 此处以 0 同时代表 V 和 W 的零元, 不致混淆. 类似地, $0 = T(0) = T(-v+v) = T(-v) + T(v)$ 蕴涵 $T(-v) = -T(v)$, 所以线性映射自动保持加法取逆的运算.

简便起见, 我们经常将 $T(v)$ 简写为 Tv.

线性映射的初步例子是空间 V 到自身的恒等映射 id_V. 另一个平凡的例子则是零映射 $0 : V \to W$, 映所有 $v \in V$ 为 0.

引理 4.5.2 设 $T : U \to V$ 和 $S : V \to W$ 都是线性映射, 则其合成 $ST : U \to W$ 也是线性映射.

证明 直接验证 $ST(v_1 + v_2) = S(Tv_1 + Tv_2) = STv_1 + STv_2$ 和 $ST(tv) = S(t \cdot Tv) = t \cdot ST(v)$. □

练习 4.5.3 设 V 和 W 是向量空间. 证明映射 $f : V \to W$ 是线性的当且仅当它的图形 Γ_f 是直和 $V \oplus W$ 的子空间.

以下概念和映射情形的定义 2.2.3 相似, 差别仅是此处的映射都要求为线性.

定义 4.5.4 考虑一对线性映射 $V \underset{S}{\overset{T}{\rightleftarrows}} W$. 若 $ST = \mathrm{id}_V$, 则称 S 是 T 的**左逆**, 称 T 是 S 的**右逆**. 有左逆 (或右逆) 的映射称为左可逆 (或右可逆) 线性映射.

定义 4.5.5 如果线性映射 $T : V \to W$ 左右皆可逆, 则称为**可逆**线性映射, 也称为**同构**. 这时存在唯一的线性映射 $T^{-1} : W \to V$ 使得 $T^{-1}T = \mathrm{id}_V$ 而 $TT^{-1} = \mathrm{id}_W$, 称为 T 的逆; 它同时是 T 的唯一左逆和唯一右逆.

[5] 算子一词更常用于物理或泛函分析.

对于可逆的 T, 其左逆/右逆的唯一性证都照搬集合的映射的情形 (定义 2.2.7) 进行论证. 如果 $T : V \to W$ 是同构, 这也写作 $T : V \xrightarrow{\sim} W$.

- ⋆ 如果 $U \xrightarrow{T} V \xrightarrow{S} W$, 则合成 ST 也是同构, $(ST)^{-1} = T^{-1}S^{-1}$;
- ⋆ 此外 T 可逆蕴涵 T^{-1} 可逆, $T = (T^{-1})^{-1}$.

论证既是熟悉的套话, 按下不表.

然而这里也出现了新问题: 如果一个线性映射作为集合的映射是可逆的, 它作为线性映射是否可逆? 答案是肯定的. 首先回忆到命题 2.2.9 表明对于集合之间的映射, 可逆和双射是一回事.

命题 4.5.6 设 $T : V \to W$ 为线性映射, 则 T 可逆当且仅当它既是线性映射又是双射.

证明 设 T 作为线性映射可逆, 则它作为集合之间的映射当然也可逆.

接着设 T 有集合意义下的逆映射 $T^{-1} : W \to V$. 今将说明 T^{-1} 必然是线性映射. 首先 T 是双射, 而对任何 $w_1, w_2 \in W$, 我们有

$$T\left(T^{-1}(w_1) + T^{-1}(w_2)\right) = TT^{-1}(w_1) + TT^{-1}(w_2) = w_1 + w_2;$$

这就相当于说 $T^{-1}(w_1) + T^{-1}(w_2) = T^{-1}(w_1 + w_2)$.

同理, 对任何 $t \in F$ 和 $w \in W$, 我们有

$$T\left(tT^{-1}(w)\right) = tTT^{-1}(w) = tw.$$

这就相当于说 $tT^{-1}(w) = T^{-1}(tw)$. 于是 T 作为线性映射是可逆的. □

练习 4.5.7 给定 F-向量空间. 若存在同构 $F : V \xrightarrow{\sim} W$, 则称 V 和 W 为同构的向量空间, 也记为 $V \simeq W$. 说明同构是向量空间之间的等价关系.

提示〉 反身性缘于 id_V 是同构, 对称性缘于 T 是同构蕴涵 T^{-1} 是同构, 传递性缘于 $(ST)^{-1} = T^{-1}S^{-1}$.

同构是代数学的核心概念之一. 如果 $T : V \xrightarrow{\sim} W$, 那么 V 和 W 不但作为集合能通过 T 建立一一对应, 而且此对应还保持所论的代数结构, 也就是加法与纯量乘法; 这表明只要所论的问题是以向量空间的语言来表述的, 则 V 和 W 可以通过 T 来等同. 举例来说: 设 $T : V \xrightarrow{\sim} W$ 为同构, 则

- ⋆ 任何子集 $S \subset V$ 是基 (或生成系, 线性无关) 当且仅当 $T(S) \subset W$ 是基 (或生成系, 线性无关);
- ⋆ $\dim V = \dim W$.

以下关于基的例子从另一个侧面阐释了同构的意涵.

例 4.5.8 设 $n \in \mathbb{Z}_{\geq 1}$ 而 V 为 n 维 F-向量空间, 则选定一组有序基 $v_1, \ldots, v_n$ 相当于将 V 等同于 F^n. 更明白地说, 这体现为集合之间的双射

$$\begin{aligned} \{v_1, \ldots, v_n : V \text{ 的有序基}\} &\xleftrightarrow{1:1} \{\varphi : F^n \to V, \text{ 向量空间的同构}\} \\ v_1, \ldots, v_n &\longmapsto [\varphi : (x_1, \ldots, x_n) \mapsto \textstyle\sum_{i=1}^n x_i v_i] \\ \varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n) &\longleftarrow\!\shortmid \varphi : \text{ 同构} \end{aligned}$$

其中 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \in F^n$ 是例 4.4.5 为 F^n 指定的标准基. 详细验证步骤如下.

- ⋆ 首先, 容易看出 $\to$ 方向映至的 φ 确实是从 F^n 到 V 的线性映射, 这是因为

$$\sum_{i=1}^n (x_i + y_i)v_i = \sum_{i=1}^n x_i v_i + \sum_{i=1}^n y_i v_i, \quad \sum_{i=1}^n (tx_i)v_i = t\sum_{i=1}^n x_i v_i.$$

  又由于所有 $v \in V$ 都能用基展开为 $\sum_{i=1}^n x_i v_i$ 的形式, 系数 $x_1, \ldots, x_n \in F$ 是被 v 唯一确定的, 因而 φ 是双射.

- ⋆ 其次看 $\leftarrow$ 方向. 基是用向量空间语言定义的概念, 不受同构影响, 因此 φ 映 F^n 的基为 V 的基. 映射 $\leftarrow$ 的定义合理.

- ⋆ 映射按 $\underset{\leftarrow}{\square}$ 方向的合成显然是恒等映射, 它将 $v_1, \ldots, v_n$ 映回自身 (请验证). 另一方面, 给定同构 $\varphi : F^n \to V$, 它向左映为有序基 $v_1 = \varphi(\boldsymbol{e}_1), \ldots, v_n = \varphi(\boldsymbol{e}_n)$, 再向右映回的同构 $F^n \to V$ 是

$$\begin{aligned} (x_1, \ldots, x_n) \mapsto \sum_{i=1}^n x_i v_i &= \sum_{i=1}^n x_i \varphi(\boldsymbol{e}_i) \\ &= \varphi\left(\sum_{i=1}^n x_i \boldsymbol{e}_i\right) = \varphi(x_1, \ldots, x_n). \end{aligned}$$

  所以映射按 $\underset{\rightarrow}{\square}$ 合成仍是恒等. 双向互为逆.

这一切表明精确到同构, 有限维向量空间和形如 F^n 的向量空间是一回事, 但同构并非标准的, 依赖于基的选择. 因此基的功能类似于脚手架, 它能让抽象的向量空间更具体而容易操作.

焦点转回一般的向量空间. 观察到

- ⋆ 如果 $T_1, T_2 : V \to W$ 都是线性映射, 则 $T_1 + T_2 : v \mapsto T_1(v) + T_2(v)$ 亦然;
- ⋆ 如果 $t \in F$ 而 $T : V \to W$ 是线性映射, 则 $tT : v \mapsto t \cdot T(v)$ 仍是线性映射.

这就启发我们将所有从 V 到 W 的线性映射作成向量空间, 其加法和纯量乘法由上述 "逐点" 或曰 "逐向量" 的加法和纯量乘法运算给出.

定义 4.5.9 设 V 和 W 为 F-向量空间. 定义 $\mathrm{Hom}(V,W)$ 为所有从 V 到 W 的线性映射所成的集合. 加法和纯量乘法分别按 $(T_1+T_2)(v)=T_1(v)+T_2(v)$ 和 $(tT)(v)=t\cdot Tv$ 确定. 空间 $\mathrm{Hom}(V,W)$ 的零元由零映射 $0:V\to W$ 给出.

请读者简单而迅速地论证 $\mathrm{Hom}(V,W)$ 对此确实成为向量空间. 线性映射的性质即刻给出关于映射合成的以下性质.

命题 4.5.10 设 U,V,W 为 F-向量空间. 映射的合成

$$\begin{aligned}\circ:\mathrm{Hom}(V,W)\times\mathrm{Hom}(U,V)&\to\mathrm{Hom}(U,W)\\(S,T)&\mapsto ST\end{aligned}$$

满足

$$(tS)\circ T=t(S\circ T)=S\circ(tT),\quad t\in F$$

和分配律

$$\begin{aligned}(S_1+S_2)\circ T&=S_1\circ T+S_2\circ T,\\S\circ(T_1+T_2)&=S\circ T_1+S\circ T_2.\end{aligned}$$

将线性映射对合成的分配律和 Hom 集是向量空间这一性质合并使用, 我们就得到环的新例子.

约定 4.5.11 设 V 为 F-向量空间, 记 $\mathrm{End}(V):=\mathrm{Hom}(V,V)$, 其元素称为 V 的**自同态**.

推论 4.5.12 设 V 为 F-向量空间, 则 $\mathrm{End}(V)$ 成为环, 其加法运算是线性映射的加法, 乘法是线性映射的合成 $(S,T)\mapsto ST$; 零元是零映射, 乘法幺元是恒等映射 id_V. 此外, $\mathrm{End}(V)$ 是零环当且仅当 $V=\{0\}$.

证明 成环的条件已经含藏于上述讨论. 注意到一个环 R 是零环当且仅当 $1_R=0_R$. 而对于向量空间 V, 不难看出 $\mathrm{id}_V=0$ 当且仅当 $V=\{0\}$. □

综上, 线性映射及其间的合成也自然地成为向量空间理论处理的对象.

4.6 从线性映射观矩阵

设 V 和 W 是域 F 上的向量空间. 本节的目标是尽量具体地了解并操作定义 4.5.9 的 $\mathrm{Hom}(V,W)$, 以及推论 4.5.12 拈出的特例 $\mathrm{End}(V)$. 关键是取基.

1. 首先假定对 V 已经选定了一组基 $\{v_j\}_{j\in J}$, 此处 J 是充当下标的某个集合. 任意 $v\in V$ 都能唯一地展开为线性组合 $v=\sum_{j\in J}x_jv_j$ (默认为有限和). 若 $T\in\mathrm{Hom}(V,W)$, 则
$$Tv=\sum_{j\in J}x_jT(v_j).$$
所以 T 由资料 $(Tv_j)_{j\in J}\in W^J$ 完全确定. 简言之:

线性映射由它在基上的作用完全确定.

反之, 假定已给定 W 中的一族向量 $(\underline{w}_j)_{j\in J}$, 定义

$$\begin{aligned} T : V &\longrightarrow W \\ \sum_{j\in J} x_j v_j &\longmapsto \sum_{j\in J} x_j \underline{w}_j. \end{aligned}$$

这是满足 $Tv_j = \underline{w}_j$ 的线性映射. 综上, 我们便得到双射

$$\begin{aligned} \mathrm{Hom}(V,W) &\xrightarrow{1:1} W^J \\ T &\longmapsto (Tv_j)_{j\in J}. \end{aligned}$$

2. 进一步选定 W 的基 $\{w_i\}_{i\in I}$, 此处 I 仍是某个充当下标的集合. 对每个 $j \in J$, 将 Tv_j 展开为
$$Tv_j = \sum_{i\in I} a_{ij} w_i$$
其中 $a_{ij} \in F$ 是由 T 唯一确定的一族系数, 而且当 j 固定, 它们至多仅对有限个 i 非零. 因此
$$\mathrm{Hom}(V,W) \xrightarrow{1:1} \left\{(a_{ij}) \in F^{I\times J} : \forall j,\ \exists \text{ 至多有限个 } i \text{ 使得 } a_{ij} \neq 0\right\}$$
其中 $(a_{ij})_{i,j}$ 对应的线性映射由下式刻画:
$$v_j \mapsto \sum_{i\in I} a_{ij} w_i, \quad j \in J. \tag{4.6.1}$$
对于一般的 $v \in V$, 若 $v = \sum_{j\in J} x_j v_j$, 则
$$\begin{aligned} Tv &= \sum_{j\in J} x_j T(v_j) = \sum_{j\in J} x_j \sum_{i\in I} a_{ij} w_i \\ &= \sum_{i\in I} \left(\sum_{j\in J} a_{ij} x_j \right) w_i; \end{aligned} \tag{4.6.2}$$
注意到每一步求和都只有有限多项非零.

3. 由此还能将 $\mathrm{Hom}(V,W)$ 的向量空间结构转译到 $F^{I\times J}$ 上:
   - ⋆ 若 $T, T' \in \mathrm{Hom}(V,W)$ 分别对应到 $(a_{ij})_{i,j}$ 和 $(a'_{ij})_{i,j}$, 则 $T + T'$ 对应到 $(a_{ij} + a'_{ij})_{i,j}$.
   - ⋆ 若 T 对应到 $(a_{ij})_{i,j}$ 而 $t \in F$, 则 tT 对应到 $(ta_{ij})_{i,j}$.

   这些断言都直接来自 (4.6.1) 的刻画.

4. 现在来探讨线性映射的合成. 给定 F-向量空间 U, V, W, 分别带有给定的基 $\{u_k\}_{k\in K}$, $\{v_j\}_{j\in J}$, $\{w_i\}_{i\in I}$. 考虑线性映射

$$\begin{aligned} S\in \mathrm{Hom}(V,W) &\quad\leftrightarrow\quad (a_{ij})_{i,j}\\ T\in \mathrm{Hom}(U,V) &\quad\leftrightarrow\quad (b_{jk})_{j,k}. \end{aligned}$$

为了描述 ST, 对每个 $k\in K$ 来计算

$$ST(u_k)=S\left(\sum_{j\in J} b_{jk}v_j\right)=\sum_{j\in J} b_{jk}S(v_j)=\sum_{j\in J}\sum_{i\in I} a_{ij}b_{jk}w_i;$$

每一步求和都只有有限多项非零. 因此线性映射 $ST: U\to W$ 对应的资料 $(c_{ik})_{i,k}$ 表作

$$c_{ik}=\sum_{j\in J} a_{ij}b_{jk},\quad i\in I,\ k\in K; \tag{4.6.3}$$

留意到上式的 $\sum_{j\in J}$ 对每个 (i,k) 都有限.

将上述讨论应用于有限维向量空间, 以下结果水到渠成.

定理 4.6.1 设 V 和 W 为有限维向量空间, 分别带选定的有序基 $v_1,\dots,v_n$ 和 $w_1,\dots,w_m$, 其中 $n,m\in\mathbb{Z}_{\geq 1}$. 此时有向量空间的同构

$$\begin{aligned} \boldsymbol{\mathcal{M}}: \mathrm{Hom}(V,W) &\xrightarrow{1:1} \mathrm{M}_{m\times n}(F)\\ T &\longmapsto (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \end{aligned}$$

其中 $(a_{ij})_{i,j}=\boldsymbol{\mathcal{M}}(T)$ 由以下性质刻画:

$$Tv_j=\sum_{i=1}^{m} a_{ij}w_i,\quad 1\leq j\leq n.$$

作为推论, 此时有 $\dim \mathrm{Hom}(V,W)=\dim V\dim W$.

证明 相当于在之前的讨论中取 $I=\{1,\dots,m\}$ 和 $J=\{1,\dots,n\}$. 关于 "存在至多有限个 i 使得 $a_{ij}\neq 0$" 的条件在此是多余的. 维数公式来自 $\dim \mathrm{M}_{m\times n}(F)=mn$. □

映射合成的描述也同样转译到矩阵空间上. 请先回忆矩阵乘法的定义 4.3.4 和关于交换图表的约定 2.3.3.

定理 4.6.2 设 U,V,W 为有限维向量空间, 分别带选定的有序基 $u_1,\dots,u_r$, $v_1,\dots,v_n$ 和 $w_1,\dots,w_m$. 以下图表交换:

$$\begin{array}{ccc} \mathrm{Hom}(V,W)\times\mathrm{Hom}(U,V) & \xrightarrow{\text{映射合成}} & \mathrm{Hom}(U,W)\\ \boldsymbol{\mathcal{M}}\downarrow \quad \downarrow\boldsymbol{\mathcal{M}} & & \downarrow\boldsymbol{\mathcal{M}}\\ \mathrm{M}_{m\times n}(F)\times \mathrm{M}_{n\times r}(F) & \xrightarrow[\text{矩阵乘法}]{} & \mathrm{M}_{m\times r}(F) \end{array}$$

换言之, $\boldsymbol{\mathcal{M}}(ST)=\boldsymbol{\mathcal{M}}(S)\boldsymbol{\mathcal{M}}(T)$, 左侧是映射合成, 右侧是矩阵乘法.

证明 比较矩阵乘法的定义和 (4.6.3), 其中代入 $K=\{1,\dots,r\}$. □

既然映射合成是自然的操作, 矩阵乘法的定义因而也是合理的.

基于上述结果, 足以用一句话说明矩阵乘法的交换律, 分配律等性质.

证明 (命题 4.3.5) 化约为线性映射相对于合成运算的种种相应性质, 如命题 4.5.10 所述. □

且看矩阵的两个特殊例子.

⋆ 无论如何选取 V 和 W 的基, 零映射 $0:V\to W$ 对应的矩阵都是零矩阵:

$$\mathbf{0}_{m\times n}:=\begin{pmatrix}0&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&0\end{pmatrix}\in \mathrm{M}_{m\times n}(F).$$

⋆ 任取 V 的有序基 $v_1,\dots,v_n$, 借以将 $\mathrm{End}(V):=\mathrm{Hom}(V,V)$ 等同于 $\mathrm{M}_{n\times n}(F)$. 此时恒等映射 id_V 由 $v_i\mapsto v_i$ 刻画, 因而对应到单位矩阵:

$$\mathbf{1}_{n\times n}:=\begin{pmatrix}1&&\\ &\ddots&\\ &&1\end{pmatrix}\in \mathrm{M}_{n\times n}(F).$$

推论 4.6.3 对所有 $n\in\mathbb{Z}_{\geq 1}$, 集合 $\mathrm{M}_{n\times n}(F)$ 对矩阵加法和乘法成为非零环, 其乘法幺元是 $\mathbf{1}_{n\times n}$, 零元是 $\mathbf{0}_{n\times n}$.

证明 这不外是推论 4.5.12 的转译. □

定义 4.6.4 称矩阵 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$ 是左可逆 (或右可逆) 的, 如果存在 $\boldsymbol{B}\in\mathrm{M}_{n\times m}(F)$ 使得 $\boldsymbol{BA}=\mathbf{1}_{n\times n}$ (或 $\boldsymbol{AB}=\mathbf{1}_{m\times m}$); 这样的 $\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的左逆 (或右逆).

如果 $m=n$ 而 $\boldsymbol{A}$ 左右皆可逆, 则称 $\boldsymbol{A}$ 是可逆 $n\times n$ 矩阵; 这时存在唯一的 $\boldsymbol{A}^{-1}\in\mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{A}^{-1}\boldsymbol{A}=\mathbf{1}_{n\times n}=\boldsymbol{A}\boldsymbol{A}^{-1}$; 它同时是 $\boldsymbol{A}$ 的唯一左逆和唯一右逆.

换言之, 可逆矩阵是环 $\mathrm{M}_{m\times m}(F)$ 的可逆元 (定义 3.1.5).

根据矩阵和线性映射的对应, 以上无非是重述线性映射 $T:F^n\to F^m$ 的左逆, 右逆和可逆性的定义 4.5.5. 为何对可逆映射要求 $m=n$? 从线性映射的观点看, 如果 $T:F^n\to F^m$ 可逆, 则因为同构保持维数, 当然要有 $n=m$.

无论在理论还是应用层面, 求逆都是矩阵论中的根本问题.

练习 4.6.5 设 $\boldsymbol{U} \in \mathrm{M}_{m\times m}(F)$ 可逆. 对于任意 $\boldsymbol{A} \in \mathrm{M}_{m\times m}(F)$, 说明 $\boldsymbol{A}$ 可逆等价于 $\boldsymbol{UA}$ 可逆, 也等价于 $\boldsymbol{AU}$ 可逆. 具体写下这些逆元之间的关系. 〈提示〉 对一般的环都成立, 见练习 3.1.7.

借助基的选取, 我们现已成功地为有限维向量空间建立线性映射与矩阵的一一对应, 并且说明了线性映射的合成对应到矩阵相乘. 另一方面, 我们也有必要对任意向量 $v \in V$ 探讨它在线性映射 $T : V \to W$ 之下的像 Tv. 如何以矩阵语言描述 Tv? 答案是容易的.

按惯例, 选定 V 的基 $v_1, \dots, v_n$ 和 W 的基 $w_1, \dots, w_m$, 计顺序. 设 $n, m \in \mathbb{Z}_{\geq 1}$ 以排除 V 或 W 为零空间的平凡例子. 给定 $T \in \mathrm{Hom}(V, W)$, 对应到矩阵 $\boldsymbol{A} := \boldsymbol{\mathcal{M}}(T) = (a_{ij})_{i,j} \in \mathrm{M}_{m\times n}(F)$. 按照 (4.6.2), 对于任意 $v = \sum_{j=1}^{n} x_j v_j \in V$, 我们有

$$Tv = \sum_{i=1}^{m} \underbrace{\sum_{j=1}^{n} a_{ij} x_j}_{\in F} w_i.$$

请回忆矩阵乘法的定义. 上式相当于说: 如果 Tv 用基展开为 $\sum_{i=1}^{m} y_i w_i$, 则其中的系数 $y_1, \dots, y_m$ 由 $\boldsymbol{A}$ 左乘列矩阵的结果确定:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix};$$

特别地, $\boldsymbol{A}$ 的第 j 列正是 Tv_j 按基 $w_1, \dots, w_m$ 展开的系数 $(j = 1, \dots, n)$.

综上, 一旦取定有限维向量空间的有序基, 矩阵便从两个面向具体地描述了线性映射.

1. 线性映射的合成对应矩阵乘法;

2. 线性映射的像由矩阵对列向量的左乘给出.

第二点也可以诠释为第一点的特例, 方法是将向量理解为一类特别的线性映射, 并且将线性映射的取值理解为映射合成的特例. 首先 F 自身也成为 F-向量空间 (例 4.2.6), 它是 1 维的, 以域的乘法幺元 $1 = 1_F$ 为当然的基. 对于任意 F-向量空间 V, 我们因而有向量空间的同构

$$\begin{aligned} V &\xrightarrow{\sim} \mathrm{Hom}(F, V) \\ v &\longmapsto [C_v : 1 \mapsto v] \\ C(1) &\longmapsto C. \end{aligned}$$

将 v 用基 $v_1,\dots,v_n$ 展开为 $\sum_{j=1}^n x_j v_j$, 则 $C_v \in \mathrm{Hom}(F,V)$ 相对于选定的基化为列向量

$$\boldsymbol{\mathcal{M}}(C_v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

相同的套路给出同构 $W \xrightarrow{\sim} \mathrm{Hom}(F,W)$, 仍记为 $w \mapsto C_w$. 所以原初的问题化为: 如何用 C_v 来表达 C_{Tv}? 观察到 $C_{Tv} = TC_v \in \mathrm{Hom}(F,W)$, 这是因为左右都映 $1 \in F$ 为 Tv. 于是

$$\begin{aligned} \boldsymbol{\mathcal{M}}(C_{Tv}) = \boldsymbol{\mathcal{M}}(TC_v) &= \boldsymbol{\mathcal{M}}(T)\boldsymbol{\mathcal{M}}(C_v) \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

一切殊途同归.

例 4.6.6 回到解方程的问题. 考虑熟悉的线性方程组

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m \end{aligned}$$

其中的系数 a_{ij} 和 b_i 都属于选定的域 F.

1. 在 §4.3 结尾, 我们已经说明原问题可以化为求解以列向量

$$\boldsymbol{x} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

为变量的矩阵方程 $\boldsymbol{Ax} = \boldsymbol{b}$, 其中

$$\boldsymbol{A} := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

2. 用向量空间的语言, 取 $V = F^n$, $W = F^m$, 它们各自配备标准的有序基

$$\begin{aligned} \boldsymbol{v}_1 &= (1,0,\dots,0), \quad \dots, \quad \boldsymbol{v}_n = (0,\dots,0,1); \\ \boldsymbol{w}_1 &= (1,0,\dots,0), \quad \dots, \quad \boldsymbol{w}_m = (0,\dots,0,1). \end{aligned}$$

定义 $T \in \mathrm{Hom}(V,W)$ 使得它映 $\boldsymbol{v}_j$ 为 $\sum_{i=1}^m a_{ij}\boldsymbol{w}_i$. 原问题又可以化为求解以 V 的元素 $\boldsymbol{v} = \sum_{j=1}^n x_j\boldsymbol{v}_j$ 为变量的方程

$$T\boldsymbol{v} = \boldsymbol{w}, \quad \boldsymbol{w} := \sum_{i=1}^m b_i\boldsymbol{w}_i.$$

因为 $\boldsymbol{\mathcal{M}}(T) = \boldsymbol{A}$, 这和前一种观点终归是一回事.

3. 反过来看, 对于任意有限维向量空间 V 和 W, 线性映射 $T \in \mathrm{Hom}(V,W)$, 以及 $w \in W$, 一旦选定了 V 和 W 的有序基, 求解 $Tv = w$ 也无非是求解对应的矩阵方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, 或者更具体地说, 无非是解相应的线性方程组.

   然而线性映射为原问题带来一个新的视角: 求解 $Tv = w$, 或者换句话说是研究 T 在 w 上的 "纤维" $T^{-1}(w)$, 其问题的本质与基无关. 不同的基可以导致形式迥异, 但实质相同的方程组. 为了阐明其中机制, 有必要仔细研究基的变换如何影响矩阵, 这是 §4.9 的内容.

最后, 我们试着从矩阵乘法或线性映射的角度来审视 Gauss–Jordan 消元法. 以下说明如何将三种初等行变换解释为某些 $m \times m$ 矩阵的左乘. 我们沿用 §1.3 的记号, 并考虑 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})_{i,j}$.

(A) 将 $\boldsymbol{A}$ 的第 i 行乘以 $c \in F$, 加到第 k 行 $(1 \leq i \neq k \leq m)$. 定义 $\mathrm{M}_{m\times m}(F)$ 的元素

$$\mathcal{A}(i,k,c) := \mathbf{1}_{m\times m} + c\boldsymbol{E}_{ki} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & c & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \text{第 } k \text{ 行} \end{matrix}$$

$$\text{第 } i \text{ 列}$$

其中对角线上是 1, 留白部分全为 0. 所论的初等行变换相当于将矩阵 $\boldsymbol{A}$ 代换为

$$\mathcal{A}(i,k,c)\boldsymbol{A} = \boldsymbol{A} + c\boldsymbol{E}_{ki}\boldsymbol{A} = \boldsymbol{A} + \begin{pmatrix} & & \\ ca_{i1} & \cdots & ca_{in} \\ & & \end{pmatrix} \text{第 } k \text{ 行}$$

留白部分的矩阵元依例全为 0.

(B) 交换 $\boldsymbol{A}$ 的第 i 行和第 k 行 $(1 \leq i < k \leq m)$. 定义 $\mathrm{M}_{m\times m}(F)$ 的元素

$$\mathcal{B}(i,k) := \sum_{\substack{1\leq j\leq m\\ j\neq i,k}} \boldsymbol{E}_{jj} + \boldsymbol{E}_{ik} + \boldsymbol{E}_{ki} = \begin{pmatrix} \ddots & & & & \\ & 0 & & 1 & \\ & & \ddots & & \\ & 1 & & 0 & \\ & & & & \ddots \end{pmatrix} \begin{matrix} \\ i \\ \\ k \\ \\ \end{matrix}$$

$$\begin{matrix} & & i & & k & \end{matrix}$$

其中省略号代表对角线上的 1. 所论的初等行变换相当于将矩阵 $\boldsymbol{A}$ 代换为 $\mathcal{B}(i,k)\boldsymbol{A}$.

(C) 将 $\boldsymbol{A}$ 的第 i 行的每一项都乘以 $c \in F^\times$. 定义 $\mathrm{M}_{m\times m}(F)$ 的元素

$$\mathcal{C}(i,c) := \sum_{\substack{1\leq j\leq m\\ j\neq i}} \boldsymbol{E}_{jj} + c\boldsymbol{E}_{ii} = \begin{pmatrix} \ddots & & \\ & c & \\ & & \ddots \end{pmatrix} \text{第 } i \text{ 行}$$

其中省略号代表对角线上的 1. 所论的初等行变换相当于将矩阵 $\boldsymbol{A}$ 代换为 $\mathcal{C}(i,c)\boldsymbol{A}$.

命题 4.6.7 以上出现的 $\mathcal{A}(i,k,c)$, $\mathcal{B}(i,k)$, $\mathcal{C}(i,c)$ 都是可逆矩阵; 它们的逆也对应于相同类型的初等行变换.

证明 首先以例 4.4.6 的乘法法则说明 $\mathcal{A}(i,k,c)^{-1} = \mathcal{A}(i,k,-c)$: 因为 $k \neq i$,

$$\begin{aligned}(\mathbf{1}_{m\times m} - c\boldsymbol{E}_{ki})(\mathbf{1}_{m\times m} + c\boldsymbol{E}_{ki}) &= \mathbf{1}_{m\times m} - c^2 \underbrace{\boldsymbol{E}_{ki}\boldsymbol{E}_{ki}}_{=\mathbf{0}} \\ &= (\mathbf{1}_{m\times m} + c\boldsymbol{E}_{ki})(\mathbf{1}_{m\times m} - c\boldsymbol{E}_{ki}).\end{aligned}$$

其次说明 $\mathcal{B}(i,k)^2 = \mathbf{1}_{m\times m}$. 最方便的看法是以 F^m 的标准有序基 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ 考量 $\mathcal{B}(i,k)$ 所对应的线性映射 $F^m \to F^m$. 逐列地考察矩阵, 可知此线性映射将 $\boldsymbol{e}_i$ 和 $\boldsymbol{e}_k$ 对调, 其余 $\boldsymbol{e}_j$ 不动. 既然线性映射由它在基上的作用确定, 这就表明此线性映射应用两次是恒等, 相应地 $\mathcal{B}(i,k)^2 = \mathbf{1}_{m\times m}$.

最后来验证 $\mathcal{C}(i,c)^{-1} = \mathcal{C}(i,c^{-1})$, 此处 $c \in F^\times$. 视同线性映射, $\mathcal{C}(i,c)$ 的效用是将标准基中的 $\boldsymbol{e}_i$ 映为 $c\boldsymbol{e}_i$, 其余 $\boldsymbol{e}_j$ 不动, 关于逆的描述因而是明白的. □

这三类可逆矩阵称为**初等矩阵**[6]. 无论对矩阵 $\boldsymbol{A}$ 施行何种初等行变换, 其效果都是以 $\boldsymbol{U}\boldsymbol{A}$ 代 $\boldsymbol{A}$, 其中 $\boldsymbol{U}$ 是 $m \times m$ 初等矩阵; 命题 4.6.7 对应于 §1.3 强调的初等行变换可以撤销的这一事实.

[6] 思考题: 何不称为初等行矩阵?

回到例 4.6.6 讨论的线性方程组, 表作矩阵方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的形式. 相应的增广矩阵按自明的方式标为 $(\boldsymbol{A}|\boldsymbol{b})\in \mathrm{M}_{m\times(n+1)}(F)$. 基于矩阵乘法的定义 4.3.4, 对应于初等矩阵 $\boldsymbol{U}$ 的初等行变换在增广矩阵上的效用是

$$\boldsymbol{U}\cdot(\boldsymbol{A}|\boldsymbol{b})=(\boldsymbol{U}\boldsymbol{A}|\boldsymbol{U}\boldsymbol{b})\ \in \mathrm{M}_{m\times(n+1)}(F),$$

右式对应到矩阵方程 $\boldsymbol{U}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{U}\boldsymbol{b}$.

对于任意可逆 $m\times m$ 矩阵 $\boldsymbol{U}$ 和任意 $\boldsymbol{x}\in \mathrm{M}_{m\times 1}(F)\simeq F^m$, 我们都有

$$\boldsymbol{U}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{U}\boldsymbol{b} \underset{\boldsymbol{U}}{\overset{\boldsymbol{U}^{-1}}{\rightleftarrows}} \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},$$

这就重新解释了初等行变换何以给出同解的线性方程组. 以此观之, 消元法是关于如何通过一列初等矩阵 $\boldsymbol{U}_1,\boldsymbol{U}_2,\dots$ 的左乘来简化系数矩阵 $\boldsymbol{A}$ 的一门技艺.

4.7 从矩阵的转置到对偶空间

将矩阵对主对角线作镜射, 或者说将行和列的下标交换角色, 便给出矩阵的转置运算.

本节前半部的操作适用于一般的环 R 上的矩阵, 见注记 4.3.8. 我们稍后将返回域 F 上的理论.

定义 4.7.1 考虑环 R 上的 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}$. 定义 $\boldsymbol{A}$ 的转置 ${}^{\mathrm{t}}\boldsymbol{A}$ 为 $n\times m$ 矩阵 $(a_{ji})_{i,j}$. 形象地说

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & & \ddots & \\ a_{m1} & \cdots & \cdots & a_{mn}\end{pmatrix} \Longrightarrow {}^{\mathrm{t}}\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1}\\ a_{12} & a_{22} & \cdots & a_{m2}\\ \vdots & & \ddots & \\ a_{1n} & \cdots & \cdots & a_{mn}\end{pmatrix}.$$

显然有

$${}^{\mathrm{t}}(\mathbf{1}_{m\times m})=\mathbf{1}_{m\times m},\quad {}^{\mathrm{t}}(\mathbf{0}_{m\times n})=\mathbf{0}_{n\times m}.$$

除此之外, 容易看出

$$\begin{aligned} {}^{\mathrm{t}}(\boldsymbol{A}+\boldsymbol{B})&={}^{\mathrm{t}}\boldsymbol{A}+{}^{\mathrm{t}}\boldsymbol{B},\\ {}^{\mathrm{t}}(s\boldsymbol{A})&=s\cdot{}^{\mathrm{t}}\boldsymbol{A},\\ {}^{\mathrm{t}\,\mathrm{t}}\boldsymbol{A}&=\boldsymbol{A},\end{aligned}$$

此处 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{m\times n}(R)$ 而 $s\in R$.

然而转置不保持乘法: 设 $\boldsymbol{C}=\boldsymbol{AB}$, 引入记号 ${}^{\mathrm{t}}a_{ij}=a_{ji}$ 等, 则 ${}^{\mathrm{t}}\boldsymbol{C}$ 的第 (i,k) 个矩阵元是 ${}^{\mathrm{t}}c_{ik}=c_{ki}$; 若 R 是交换环, 则它也等于

$$\sum_j a_{kj}b_{ji}=\sum_j {}^{\mathrm{t}}b_{ij}\,{}^{\mathrm{t}}a_{jk}.$$

这就说明了交换环 R 上的矩阵转置调换乘法顺序: 对于任意矩阵 $\boldsymbol{A}\in \mathrm{M}_{m\times n}(R)$ 和 $\boldsymbol{B}\in \mathrm{M}_{n\times r}(R)$, 我们有

$${}^{\mathrm{t}}(\boldsymbol{AB})={}^{\mathrm{t}}\boldsymbol{B}\,{}^{\mathrm{t}}\boldsymbol{A}\ \in \mathrm{M}_{r\times m}(R).$$

今起考虑域 F 上的矩阵及其转置运算. 转置给出 F-向量空间的同构

$$\mathrm{M}_{m\times n}(F)\xrightarrow{\sim}\mathrm{M}_{n\times m}(F).$$

命题 4.7.2　设 $\boldsymbol{A}\in \mathrm{M}_{m\times m}(F)$, 则 $\boldsymbol{A}$ 可逆当且仅当 ${}^{\mathrm{t}}\boldsymbol{A}$ 可逆, 而且此时 $({}^{\mathrm{t}}\boldsymbol{A})^{-1}={}^{\mathrm{t}}(\boldsymbol{A}^{-1})$.

证明　如果 $\boldsymbol{A}$ 有逆 $\boldsymbol{A}^{-1}$, 则对 $\boldsymbol{AA}^{-1}=\mathbf{1}_{m\times m}=\boldsymbol{A}^{-1}\boldsymbol{A}$ 两边取转置, 可得

$${}^{\mathrm{t}}(\boldsymbol{A}^{-1})\,{}^{\mathrm{t}}\boldsymbol{A}=\mathbf{1}_{m\times m}={}^{\mathrm{t}}\boldsymbol{A}\,{}^{\mathrm{t}}(\boldsymbol{A}^{-1}),$$

这相当于说 ${}^{\mathrm{t}}(\boldsymbol{A}^{-1})$ 是 ${}^{\mathrm{t}}\boldsymbol{A}$ 的逆. 由于 ${}^{\mathrm{t}\,\mathrm{t}}\boldsymbol{A}=\boldsymbol{A}$, 充要条件的另一方向也是如此. □

关于矩阵转置的这些公式完整说明了矩阵的行与列具有相互对称的角色. 以下仍以 Gauss–Jordan 消元法为例说明这一思路. 对于任意矩阵 $\boldsymbol{A}\in \mathrm{M}_{m\times n}(F)$, 我们有以下三种操作.

(A) $\mathrm{A}(j,k,c)_{\text{列}}$: 设 $1\leq j\neq k\leq n$ 而 $c\in F$, 将 $\boldsymbol{A}$ 的第 j 列乘以 c 加到第 k 列, 其他列保持不变.

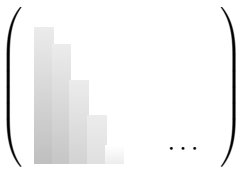
(B) $\mathrm{B}(j,k)_{\text{列}}$: 设 $1\leq j<k\leq n$, 交换 $\boldsymbol{A}$ 的第 j 列和第 k 列.

(C) $\mathrm{C}(j,c)_{\text{列}}$: 设 $1\leq j\leq n$ 而 $c\in F$ 非零, 将 $\boldsymbol{A}$ 的第 j 列的每一项都乘以 c.

这些是 §1.3 介绍的 $\mathrm{A}(i,k,c)$, $\mathrm{B}(i,k)$ 和 $\mathrm{C}(i,c)$ 三类初等行变换的列版本, 符号也是一一对应的, 理所当然地称为**初等列变换**. 对 $\boldsymbol{A}$ 作初等列变换相当于先将 $\boldsymbol{A}$ 行列对换取镜像, 在镜像世界里进行对应的初等行变换, 再将行列换回原状. 换言之,

$$\boldsymbol{A}\text{ 作初等列变换}={}^{\mathrm{t}}\big({}^{\mathrm{t}}\boldsymbol{A}\text{ 作相应的初等行变换}\big).\tag{4.7.1}$$

对于定义 1.3.4 的行梯矩阵和主元, 自然也有列的版本: **列梯矩阵**是每列严格向下缩进的矩阵, 形象地表作



其中右上空白部分全为 0. 如果进一步要求主元左边的项全为零, 则称之为**简化列梯矩阵**. 这当然是定义 1.3.6 的镜像. 行列对换的思路同样给出

$$\text{列梯矩阵} = {}^{\mathrm{t}}(\text{行梯矩阵}), \quad \text{简化列梯矩阵} = {}^{\mathrm{t}}(\text{简化行梯矩阵}).$$

注意到 §4.6 介绍的初等矩阵满足以下性质.

引理 4.7.3 *初等矩阵的转置仍是初等矩阵.*

证明 对 §4.6 定义的三族矩阵逐一考察: 我们有 ${}^{\mathrm{t}}\mathcal{A}(i,k,c) = \mathcal{A}(k,i,c)$, ${}^{\mathrm{t}}\mathcal{B}(i,k) = \mathcal{B}(i,k)$ 和 ${}^{\mathrm{t}}\mathcal{C}(i,c) = \mathcal{C}(i,c)$. □

对 ${}^{\mathrm{t}}\boldsymbol{A}$ 作初等行变换相当于左乘初等矩阵. 基于 (4.7.1), 对 $\boldsymbol{A}$ 作初等列变换便相当于右乘以初等矩阵. 将这一观察搭配引理 4.7.3, Gauss–Jordan 消元法的列版本因而可以总结如下.

命题 4.7.4 *任何 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 都可以通过一系列初等列变换化为简化列梯矩阵. 等价地说, 存在 $r \in \mathbb{Z}_{\geq 1}$ 和一列初等矩阵 $\boldsymbol{V}_1, \dots, \boldsymbol{V}_r$, 使得 $\boldsymbol{A}\boldsymbol{V}_1 \cdots \boldsymbol{V}_r$ 是简化列梯矩阵.*

迄今的全部讨论都以矩阵语言表述. 我们知道矩阵及其运算可以视为线性映射在取基之后的化身, 转置是否也有类似的诠释? 答案是肯定的, 这涉及对偶空间的概念.

回忆到定义 4.5.9 已经对所有形如 $\mathrm{Hom}(V,W)$ 的集合赋予了向量空间结构. 取 W 为 1 维空间 F (例 4.2.6), 则 $\mathrm{Hom}(V,F)$ 也自然地成为向量空间.

定义 4.7.5 *设 V 为 F-向量空间, 它的***对偶空间** $V^\vee$ *定义为 $\mathrm{Hom}(V,F)$.*

定义 4.7.6 *如果 $T : V \to W$ 是线性映射, 则我们可以定义 T 的***转置**映射为

$$\begin{aligned} {}^{\mathrm{t}}T : W^\vee &\longrightarrow V^\vee \\ \lambda &\longmapsto \lambda T, \end{aligned}$$

其中 λT 代表线性映射 $V \xrightarrow{T} W \xrightarrow{\lambda} F$ 的合成.

基于线性映射合成的一般性质 (命题 4.5.10), 对所有 $\lambda_1, \lambda_2 \in W^\vee$ 和 $s \in F$ 都有

$$\begin{aligned} {}^{\mathrm{t}}T(\lambda_1+\lambda_2) &:= (\lambda_1+\lambda_2)T \\ &= \lambda_1 T + \lambda_2 T, \\ &= {}^{\mathrm{t}}T(\lambda_1) + {}^{\mathrm{t}}T(\lambda_2), \\ {}^{\mathrm{t}}T(s\lambda) &:= (s\lambda)T \\ &= s\cdot \lambda T = s\cdot {}^{\mathrm{t}}T(\lambda), \end{aligned}$$

因此 ${}^{\mathrm{t}}T : W^\vee \to V^\vee$ 仍是线性映射.

请读者迅速检验转置保持加法和纯量乘法: 对于 $T_1, T_2 \in \mathrm{Hom}(V,W)$, 我们有 ${}^{\mathrm{t}}(T_1+T_2) = {}^{\mathrm{t}}T_1 + {}^{\mathrm{t}}T_2$, 此外 ${}^{\mathrm{t}}(sT) = s\cdot {}^{\mathrm{t}}T$.

另一方面, 转置倒转乘法顺序.

命题 4.7.7 *对于所有线性映射 $U \xrightarrow{T} V \xrightarrow{S} W$, 我们有*

$$ {}^{\mathrm{t}}(ST) = {}^{\mathrm{t}}T\ {}^{\mathrm{t}}S \ \in \mathrm{Hom}(W^\vee, U^\vee). $$

证明 应用线性映射的合成操作, 设 $\lambda \in W^\vee$, 则

$$ {}^{\mathrm{t}}T\left({}^{\mathrm{t}}S(\lambda)\right) = {}^{\mathrm{t}}T(\underbrace{\lambda S}_{\in V^\vee}) = \underbrace{\lambda ST}_{\in U^\vee}, $$

而末项又等于 $\lambda(ST) = {}^{\mathrm{t}}(ST)(\lambda)$. □

对于无穷维空间, $V^\vee$ 总是比 V 大得多, 请感兴趣的读者参照练习 4.10.11. 对于有限维的 V, 我们希望对 $V^\vee$ 得到易于上手的描述, 为此便需要一套更抽象的描述.

定义–命题 4.7.8 *相对于有限维向量空间 V 的基 $v_1,\dots,v_n$, 对每个 $1\le i\le n$ 定义*

$$ \check{v}_i \in V^\vee: \quad \check{v}_i\left(\sum_{j=1}^n x_j v_j\right) = x_i. $$

*这些 $\check{v}_1,\dots,\check{v}_n$ 构成 $V^\vee$ 的基, 称为 $v_1,\dots,v_n$ 的***对偶基***.*

作为推论, 当 V 有限维时总有 $\dim V^\vee = \dim V$.

证明 首先说明它们线性无关. 设 $\sum_{j=1}^n a_j\check{v}_j = 0$, 将两边作用在 v_i 上, 得到 $a_i = 0$; 由于 $1\le i\le n$ 可任取, 故 $a_1 = \cdots = a_n = 0$.

其次说明它们生成 $V^\vee$. 设 $\lambda \in V^\vee$, 定义 $x_i := \lambda(v_i) \in F$, 则 $\lambda - \sum_{j=1}^n x_j\check{v}_j$ 将每个 v_i 都映为 $\lambda(v_i) - x_i = 0$, 从而 $\lambda - \sum_{j=1}^n x_j\check{v}_j = 0$. □

对于特例 $V = F^n$, 我们有

$$ F^n = \mathrm{M}_{n\times 1}(F) = \{n \text{ 维列向量}\}, $$
$$ \text{标准有序基}: \ \boldsymbol{e}_1,\dots,\boldsymbol{e}_n. $$

于是 $(F^n)^\vee = \mathrm{Hom}(F^n, F) = \mathrm{Hom}(F^n, F^1)$ 自然地等同于 $\mathrm{M}_{1\times n}(F)$, 即 n 维行向量空间. 既然线性映射的取值体现为矩阵对列向量的左乘, 故对偶空间在向量上的求值映射

$$ \begin{aligned} V^\vee \times V &\to F \\ (\lambda, v) &\mapsto \lambda(v) \end{aligned} $$

在 $V = F^n$ 时等同于矩阵乘法

$$ \mathrm{M}_{1\times n}(F) \times \mathrm{M}_{n\times 1}(F) \to \mathrm{M}_{1\times 1}(F) = F. $$

更具体地说,

$$\boldsymbol{\lambda} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

$$\boldsymbol{\lambda}(\boldsymbol{v}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \left(\sum_{i=1}^{n} x_i y_i \right) \in F.$$

这也一并说明 F^n 的标准有序基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ 的对偶基 $\check{\boldsymbol{e}}_1, \dots, \check{\boldsymbol{e}}_n$ 正是行向量

$$\check{\boldsymbol{e}}_i = \begin{pmatrix} 0 & \cdots & \underset{i}{1} & \cdots & 0 \end{pmatrix} \quad (\text{第 } i \text{ 列之外全为零}), \quad 1 \leq i \leq n;$$

这是因为简单的矩阵计算给出 $\check{\boldsymbol{e}}_i \boldsymbol{e}_j$ 等于 0 (若 $i \neq j$) 或 1 (若 $i = j$).

现在便容易说明矩阵的转置和线性映射的转置之间的关系.

命题 4.7.9 设 V, W 为有限维向量空间, 各自带有选定的有序基 $v_1, \dots, v_n$ 和 $w_1, \dots, w_m$. 依此为 $W^\vee$ 和 $V^\vee$ 配备相应的对偶基. 设 $T : V \to W$ 对应的矩阵为 $\boldsymbol{A} := \boldsymbol{\mathcal{M}}(T) \in \mathrm{M}_{m \times n}(F)$, 则 ${}^{\mathrm{t}}T : W^\vee \to V^\vee$ 对应的矩阵 $\boldsymbol{\mathcal{M}}({}^{\mathrm{t}}T) \in \mathrm{M}_{n \times m}(F)$ 等于 ${}^{\mathrm{t}}\boldsymbol{A}$.

证明 请回忆定理 4.6.1: 矩阵 $\boldsymbol{A} = (a_{ij})_{i,j}$ 按 $Tv_j = \sum_{i=1}^{m} a_{ij} w_i$ 确定. 现在考虑 $V^\vee$ 的元素

$$({}^{\mathrm{t}}T)(\check{w}_i) : v_k \mapsto \check{w}_i(Tv_k) = \sum_{h=1}^{m} a_{hk} \check{w}_i(w_h) = a_{ik},$$

其中 $1 \leq k \leq n$ 任取; 然而右式也等于 $\sum_{j=1}^{n} a_{ij} \check{v}_j \in V^\vee$ 在 v_k 处的取值. 因此

$$({}^{\mathrm{t}}T)(\check{w}_i) = \sum_{j=1}^{n} a_{ij} \check{v}_j.$$

再次和定理 4.6.1 对 $\boldsymbol{\mathcal{M}}$ 的描述相比较, 但改为考虑对偶空间 $W^\vee$, $V^\vee$ 及它们的对偶基, 则上式说明 $\boldsymbol{\mathcal{M}}({}^{\mathrm{t}}T) = {}^{\mathrm{t}}\boldsymbol{A}$. □

以此观之, 转置矩阵的乘法性质 ${}^{\mathrm{t}}\boldsymbol{A}\ {}^{\mathrm{t}}\boldsymbol{B} = {}^{\mathrm{t}}(\boldsymbol{B}\boldsymbol{A})$ 无非是命题 4.7.7 在有限维情形的反映.

4.8 核, 像与消元法

选定域 F. 从给定的线性映射 T 出发, 至少可以定义两个重要的子空间.

定义 4.8.1 线性映射 $T : V \to W$ 的**核** (或称**零核**) 定义为

$$\ker T := \{v \in V : Tv = 0\} = T^{-1}(0).$$

它的**像**则是 $\operatorname{im} T := \{w \in W : \exists v \in V,\, Tv = w\}$.

命题 4.8.2 线性映射 $T : V \to W$ 的核是 V 的子空间, 像是 W 的子空间.

证明 按定义操演. 首先处理 $\ker T$. 由 $T(0) = 0$ 可见 $0 \in \ker T$. 若 $v_1, v_2 \in \ker T$, 则 $T(v_1+v_2) = Tv_1+Tv_2 = 0+0 = 0$. 若 $v \in \ker T$ 而 $t \in F$, 则 $T(tv) = t\cdot Tv = t\cdot 0 = 0$, 由此知 $\ker T$ 确实是子空间.

类似地, $T(0) = 0$ 蕴涵 $0 \in \operatorname{im} T$. 若 $w_1 = Tv_1$, $w_2 = Tv_2$, 则 $w_1+w_2 = T(v_1+v_2)$. 若 $w = Tv$ 而 $t \in F$, 则 $tw = t \cdot Tv = T(tv)$. 由此知 $\operatorname{im} T$ 也是子空间. □

核的重要性之一在于它描述了 T 的纤维. 对于所有 $v_1, v_2 \in V$, 我们有

$$\begin{aligned} Tv_1 = Tv_2 &\iff T(v_1 - v_2) = 0 \\ &\iff v_1 - v_2 \in \ker T. \end{aligned}$$

这就表明对于任意 $w \in W$, 或者 $w \notin \operatorname{im} T$, 亦即 $T^{-1}(w) = \varnothing$; 或者 $w \in \operatorname{im} T$, 此时任取 $v \in T^{-1}(w)$ 便给出

$$T^{-1}(w) = v + \ker T := \{v + u : u \in \ker T\}.$$

对于 $V = F^n$ 和 $W = F^m$ 的情形, 描述纤维 $T^{-1}(w)$ 就相当于解对应的 n 元线性方程组, 这正是 §4.1 所勾勒的内容, 而 $\ker(T)$ 的基正是定义 4.1.2 所谓的基础解系.

既有这些观察, 以下结论一望可知.

命题 4.8.3 线性映射 T 是单射当且仅当 $\ker T = \{0\}$.

此外, 对于线性映射 $U \xrightarrow{T} V \xrightarrow{S} W$, 我们有

$$S \text{ 是单射} \implies \ker(ST) = \ker(T),$$

这是因为 S 的单性蕴涵 $S(T(v)) = 0 \iff Tv = 0$.

如果 V 是有限维空间, 则 $\ker T$ 作为 V 的子空间也是有限维的 (推论 4.4.16); 像 $\operatorname{im} T$ 同样有限维, 这是因为若 $v_1, \ldots, v_n$ 生成 V, 则 $Tv_1, \ldots, Tv_n$ 生成 $\operatorname{im} T$. 线性映射的核与像空间的维数满足以下的基本等式.

定理 4.8.4 设 $T : V \to W$ 是向量空间之间的线性映射, V 是有限维的, 则

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T).$$

证明 任取 $\operatorname{im} T$ 的基 $w_1, \dots, w_m$ 和 $\ker T$ 的基 $u_1, \dots, u_k$. 另外取 $v_1, \dots, v_m \in V$ 使得对每个 $1 \le i \le m$ 都有 $Tv_i = w_i$. 证明 $u_1, \dots, u_k, v_1, \dots, v_m$ 是 V 的基即可.

首先说明它们线性无关. 若有一族系数 $a_1, \dots, a_m$ 和 $b_1, \dots, b_k$ 使得

$$\sum_{i=1}^{m} a_i v_i + \sum_{j=1}^{k} b_j u_j = 0,$$

则两边同取 T 给出 $\sum_{i=1}^{m} a_i w_i = 0$, 从而 a_i 全为 0; 代回原式给出 $\sum_{j=1}^{k} b_j u_j = 0$, 故 b_j 也全为 0.

接着说明它们生成整个 V. 给定 $v \in V$, 将 Tv 展开为 $\sum_{i=1}^{m} a_i w_i$, 则

$$T\left(v - \sum_{i=1}^{m} a_i v_i\right) = Tv - \sum_{i=1}^{m} a_i w_i = 0,$$

于是 $v - \sum_{i=1}^{m} a_i v_i \in \ker T$, 可以进一步表作 $\sum_{j=1}^{k} b_j u_j$; 综上, $v = \sum_{i=1}^{m} a_i v_i + \sum_{j=1}^{k} b_j u_j$. 这就验证了关于基的所有条件. □

命题 4.12.4 将以商空间的语言重新表述以上等式.

次一结果可以设想为抽屉原理 (命题 2.9.4) 的某种变体.

推论 4.8.5 设 $T : V \to W$ 是线性映射, $\dim V = \dim W$ 是有限的, 则

$$T \text{ 是同构} \iff T \text{ 单} \iff T \text{ 满}.$$

因此 T 可逆当且仅当它左可逆, 当且仅当它右可逆.

证明 我们有 T 单当且仅当 $\dim(\ker T) = 0$, 而 T 满当且仅当 $\dim(\operatorname{im} T) = \dim V$ (推论 4.4.16). 于是定理 4.8.4 确保单性和满性相互等价. 关于可逆性的断言留作简单练习. □

定义 4.8.6 (秩) 设 $T : V \to W$ 是有限维向量空间之间的线性映射, 其秩 $\operatorname{rk}(T)$ 定义为 $\dim(\operatorname{im} T)$.

矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 的秩 $\operatorname{rk}(\boldsymbol{A})$ 定义为 $\boldsymbol{A}$ 视作线性映射 $F^n \to F^m$ 的秩.

练习 4.8.7 证明 $\operatorname{rk}(\boldsymbol{A}+\boldsymbol{B}) \le \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{B})$ 恒成立.

例 4.8.8 (Sylvester 秩不等式) 定理 4.8.4 可以给出关于秩的一些有趣估计. 举例来说, 考虑线性映射 $U \xrightarrow{T} V \xrightarrow{S} W$, 设 V 有限维. 以下来说明

$$\operatorname{rk}(ST) \ge \operatorname{rk}(S) + \operatorname{rk}(T) - \dim V.$$

按定义, $\mathrm{rk}(ST)$ 等于 S 的限制 $S|_{\mathrm{im}(T)}:\mathrm{im}(T)\to W$ 的秩, 后者也等于

$$\dim(\mathrm{im}(T))-\dim(\ker(S)\cap\mathrm{im}(T))=\mathrm{rk}(T)-\dim(\ker(S)\cap\mathrm{im}(T)).$$

从 $\ker(S)\cap\mathrm{im}(T)\subset\ker(S)$ 知

$$\dim\left(\ker(S)\cap\mathrm{im}(T)\right)\le\dim\ker(S),$$

而另一方面 $\dim\ker(S)=\dim V-\mathrm{rk}(S)$, 组合起来就是所求的估计. 这种不等式当然也可以用矩阵语言表述.

练习 4.8.9 (Frobenius 秩不等式) 证明 $\mathrm{rk}(RST)\ge\mathrm{rk}(RS)+\mathrm{rk}(ST)-\mathrm{rk}(S)$, 前提是这些线性映射的合成有意义.

提示〉 手法类似例 4.8.8, 应用

$$\begin{aligned}\dim\left(\mathrm{im}(RST)\right)&=\dim\left(\mathrm{im}(ST)\right)-\dim\left(\mathrm{im}(ST)\cap\ker(R)\right)\\&\ge\dim\left(\mathrm{im}(ST)\right)-\dim\left(\mathrm{im}(S)\cap\ker(R)\right),\\\dim\left(\mathrm{im}(RS)\right)&=\dim\left(\mathrm{im}(S)\right)-\dim\left(\mathrm{im}(S)\cap\ker(R)\right).\end{aligned}$$

注记 4.8.10 (行秩和列秩) 将 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$ 分解为列向量 $\boldsymbol{c}_1,\dots,\boldsymbol{c}_n\in\mathrm{M}_{m\times1}(F)\simeq F^m$, 表作

$$\boldsymbol{A}=\left(\begin{array}{c|c|c}\boldsymbol{c}_1&\cdots&\boldsymbol{c}_n\end{array}\right).$$

请读者根据矩阵乘法的定义说明 $\boldsymbol{c}_i=\boldsymbol{A}\boldsymbol{e}_i$, 其中 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ 是 $F^n\simeq\mathrm{M}_{n\times1}(F)$ 的标准基. 所以若将 $\boldsymbol{A}$ 视同线性映射 $F^n\to F^m$, 则

$$\mathrm{im}(\boldsymbol{A})=\langle\boldsymbol{c}_1,\dots,\boldsymbol{c}_n\rangle,$$

从而 $\mathrm{rk}(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的列向量 $\boldsymbol{c}_1,\dots,\boldsymbol{c}_n\in F^m$ 所生成的子空间的维数. 有鉴于此, $\mathrm{rk}(\boldsymbol{A})$ 也称为 $\boldsymbol{A}$ 的**列秩**.

同理, 如将 $\boldsymbol{A}$ 分解为行向量, 表作

$$\boldsymbol{A}=\left(\begin{array}{c}\boldsymbol{r}_1\\\hline\vdots\\\hline\boldsymbol{r}_m\end{array}\right),$$

则 $\langle\boldsymbol{r}_1,\dots,\boldsymbol{r}_m\rangle\subset\mathrm{M}_{1\times n}(F)\simeq F^n$ 的维数自然当称为 $\boldsymbol{A}$ 的**行秩**. 应用定义 4.7.1 的转置矩阵来表述, 这相当于说

$$\boldsymbol{A}\text{ 的行秩}={}^{\mathrm{t}}\boldsymbol{A}\text{ 的列秩}.$$

定理 4.9.11 将证明任何矩阵的行秩都等于列秩, 这一事实并非显然.

注记 4.8.11 考虑有限维向量空间之间的线性映射 $U\xrightarrow{T}V\xrightarrow{S}W$. 从向量空间的抽象理论观之, 以下性质都是容易的.

- ⋆ $\mathrm{rk}(ST) \leq \min\{\mathrm{rk}(S), \mathrm{rk}(T)\}$, 这是因为 $\mathrm{im}(ST) = \mathrm{im}\,(S|_{\mathrm{im}\,T}) \subset \mathrm{im}\,S$, 而另一方面定理 4.8.4 又蕴涵 $\mathrm{im}(S|_{\mathrm{im}\,T})$ 的维数不超过 $\mathrm{im}\,T$ 的维数.
- ⋆ T 满 $\Longrightarrow$ $\mathrm{rk}(ST) = \mathrm{rk}(S)$, 这是因为此时 $\mathrm{im}(ST) = \mathrm{im}(S)$.
- ⋆ S 单 $\Longrightarrow$ $\mathrm{rk}(ST) = \mathrm{rk}(T)$, 这是因为此时 S 给出既单又满的线性映射 $\mathrm{im}\,T \to \mathrm{im}(ST)$, 因而为同构.

秩的定义不依赖基的选取, 但具体计算往往需要适当地取基, 化之为矩阵计算. 工具仍然是 Gauss–Jordan 消元法. 以下考虑 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$, 同时看作线性映射 $F^n \to F^m$. 我们现在知道对 $\boldsymbol{A}$ 作初等行变换相当于以 $\boldsymbol{U}\boldsymbol{A}$ 代 $\boldsymbol{A}$, 其中 $\boldsymbol{U} \in \mathrm{M}_{m\times m}(F)$ 是一系列初等矩阵的乘积, 因而可逆; 先前的注记说明 $\mathrm{rk}(\boldsymbol{U}\boldsymbol{A}) = \mathrm{rk}(\boldsymbol{A})$, 所以求秩的一种策略是将 $\boldsymbol{A}$ 通过一系列初等行变换化为简化行梯矩阵 $\boldsymbol{A}'$.

命题 4.8.12 *设 $\boldsymbol{A}'$ 为 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 通过初等行变换化成的简化行梯矩阵, 则 $\mathrm{rk}(\boldsymbol{A})$ 等于 $\boldsymbol{A}'$ 的主元个数 r.*

证明 按照注记 4.8.10 将 $\mathrm{rk}(\boldsymbol{A})$ 理解为列秩, 问题化为对列向量生成的子空间求基, 故对应的断言化约为算法 4.4.8. □

另一方面, 描述 $\ker(\boldsymbol{A})$ 等价于解以 $\boldsymbol{A}$ 为系数矩阵的齐次线性方程组, 而关于基础解系的命题 4.1.3 说明 $\ker(\boldsymbol{A})$ 是 $n-r$ 维空间. 更精确地说, 如果将 $\boldsymbol{A}'$ 中不含主元的列按编号枚举为

$$1 \leq f_1 < \cdots < f_{n-r} \leq n,$$

则线性方程组通解公式 (1.4.1) 所给出的无非是同构

$$\begin{aligned} F^n \supset \ker(\boldsymbol{A}') = \ker(\boldsymbol{A}) &\xrightarrow{\sim} F^{n-r} \\ (x_1, \ldots, x_n) &\longmapsto \left(x_{f_1}, \ldots, x_{f_{n-r}}\right). \end{aligned}$$

综上,

$$\begin{aligned} &\boldsymbol{A}' \text{ 的主元个数 } r = \mathrm{rk}(\boldsymbol{A}), \\ &n - r \overset{\text{消元法}}{=\!=} \dim \ker(\boldsymbol{A}') = \dim \ker(\boldsymbol{A}). \end{aligned}$$

以上也顺道借 Gauss–Jordan 消元法重新证明了定理 4.8.4, 而且论证是构造性的.

命题 4.8.13 *对于矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times m}(F)$, 以下性质相互等价.*

(i) *$\boldsymbol{A}$ 可逆;*

(ii) *对任意列向量 $\boldsymbol{v} \in F^m$, 我们有 $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0} \iff \boldsymbol{v} = \boldsymbol{0}$;*

(iii) *$\mathrm{rk}(\boldsymbol{A}) = m$;*

(iv) $\boldsymbol{A}$ 可以表为初等矩阵的乘积.

因此 $\boldsymbol{A}\in \mathrm{M}_{m\times m}(F)$ 可逆当且仅当它左可逆, 当且仅当它右可逆.

证明 性质 (i) — (iii) 的等价性不外是推论 4.8.5 的矩阵版本 (取 $V=W=F^m$), 而先前的讨论说明这点也可以从矩阵的消元法来推导.

设 (iii) 成立, 将 $\boldsymbol{A}$ 化为 $m\times m$ 简化行梯矩阵 $\boldsymbol{A}'$, 写作

$$\boldsymbol{U}_1\cdots\boldsymbol{U}_s\boldsymbol{A}=\boldsymbol{A}',\quad \boldsymbol{U}_i:\text{初等矩阵}.$$

然而 $\boldsymbol{A}'$ 的主元个数等于 $\mathrm{rk}(\boldsymbol{A})=m$, 唯一可能是 $\boldsymbol{A}'=\mathbf{1}_{m\times m}$, 故 $\boldsymbol{A}=\boldsymbol{U}_s^{-1}\cdots\boldsymbol{U}_1^{-1}$ 给出 (iv).

最后设 (iv) 成立, 因为初等矩阵皆可逆, 其乘积亦然, 故 (i) 成立. □

算法 4.8.14 命题 4.8.13 也指明了由 Gauss–Jordan 消元法求逆的途径. 设 $\boldsymbol{A}\in \mathrm{M}_{m\times m}(F)$, 将其增广为 $m\times 2m$ 矩阵 $(\boldsymbol{A}|\mathbf{1}_{m\times m})$. 对增广矩阵进行初等行变换相当于左乘以各种 $m\times m$ 初等矩阵 $\boldsymbol{U}$, 其效果可以分块描述: 设 $h\in\mathbb{Z}_{\geq 1}$, 则

$$\boldsymbol{U}\,(\boldsymbol{A}|\boldsymbol{B})=(\boldsymbol{U}\boldsymbol{A}|\boldsymbol{U}\boldsymbol{B}),\quad \boldsymbol{B}\in\mathrm{M}_{m\times h}(F).$$

我们在 §4.6 结尾已经用过上式的 $h=1$ 情形, 由于初等行变换 (或 $\boldsymbol{U}$ 的左乘) 可以对 $\boldsymbol{B}$ 逐列施行, 一般情形自然也对. 现在通过消元法化 $\boldsymbol{A}$ 为简化行梯矩阵 $\boldsymbol{A}'=\boldsymbol{U}_1\cdots\boldsymbol{U}_s\boldsymbol{A}$, 相应地

$$\boldsymbol{U}_1\cdots\boldsymbol{U}_s\,(\boldsymbol{A}|\mathbf{1}_{m\times m})=\left(\boldsymbol{A}'\middle|\boldsymbol{B}'\right),\quad \boldsymbol{B}':=\boldsymbol{U}_1\cdots\boldsymbol{U}_s.$$

若 $\boldsymbol{A}'\neq\mathbf{1}_{m\times m}$ 则主元个数 $<m$, 此时命题 4.8.12 和命题 4.8.13 蕴涵 $\boldsymbol{A}$ 不可逆. 若 $\boldsymbol{A}'=\mathbf{1}_{m\times m}$ 则 $\boldsymbol{A}^{-1}=\boldsymbol{U}_1\cdots\boldsymbol{U}_s$, 此即增广矩阵经过初等行变换以后的右栏 $\boldsymbol{B}'$.

这一求逆算法的复杂度类似于 Gauss–Jordan 消元法.

练习 4.8.15 定义 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$, 使得它的第 (i,j) 个矩阵元为 $\min\{i,j\}$. 以消元法说明 $\boldsymbol{A}$ 可逆, 并且具体将 $\boldsymbol{A}^{-1}$ 表为初等矩阵的乘积.

4.9 基的变换: 矩阵的共轭与相抵

设 V 为 n 维向量空间, W 为 m 维向量空间, $n,m\in\mathbb{Z}_{\geq 1}$. 我们在 §4.6 说明了一旦取定 V 的基 $v_1,\dots,v_n$ 和 W 的基 $w_1,\dots,w_m$ (记顺序), 则线性映射 $V\to W$ 可以由相应的 $m\times n$ 矩阵描写. 这体现为向量空间的同构

$$\boldsymbol{\mathcal{M}}:\mathrm{Hom}(V,W)\xrightarrow{\sim}\mathrm{M}_{m\times n}(F).$$

实际应用中往往涉及换基. 以下的目标是精确说明 $\boldsymbol{\mathcal{M}}$ 如何依赖基的选取. 我们将有序基 $v_1,\dots,v_n$ (或 $w_1,\dots,w_m$) 简记为 $\mathbf{v}$ (或 $\mathbf{w}$), 相应的映射 $\boldsymbol{\mathcal{M}}$ 记为 $\boldsymbol{\mathcal{M}}^{\mathbf{w}}_{\mathbf{v}}$. 我

们希望对 $\boldsymbol{\mathcal{M}}_{\mathbf{v}}^{\mathbf{w}}$ 有更清晰的理解. 为此, 回忆到例 4.5.8 说明给定有序基 $\mathbf{v}$ (或 $\mathbf{w}$) 相当于给定同构 $\varphi_{\mathbf{v}}: F^n \xrightarrow{\sim} V$ (或 $\varphi_{\mathbf{w}}: F^m \xrightarrow{\sim} W$). 将 $\mathrm{M}_{m\times n}(F)$ 的元素视同线性映射 $F^n \to F^m$, 我们断言下图在约定 2.3.3 的意义下交换:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle\varphi_{\mathbf{v}}}\uparrow\wr & & \wr\uparrow{\scriptstyle\varphi_{\mathbf{w}}} \\ F^n & \xrightarrow[\boldsymbol{\mathcal{M}}_{\mathbf{v}}^{\mathbf{w}}(T)]{} & F^m. \end{array} \tag{4.9.1}$$

为了验证交换性, 设 $\boldsymbol{\mathcal{M}}_{\mathbf{v}}^{\mathbf{w}}(T) = (a_{ij})_{i,j}$ 并且对 F^n 的标准基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ 研究每个元素 $\boldsymbol{e}_j$ 沿两路给出的像

$$\begin{array}{ccc} v_j & \longmapsto & Tv_j \\ \uparrow & & \\ \boldsymbol{e}_j & & \end{array} \quad \text{和} \quad \begin{array}{ccc} & & \sum_{i=1}^m a_{ij}w_i \\ & & \uparrow \\ \boldsymbol{e}_j & \longmapsto & \sum_{i=1}^m a_{ij}\boldsymbol{e}_i \end{array}$$

然而按照矩阵和线性映射的对应, $Tv_j = \sum_{i=1}^m a_{ij}w_i$, 故 (4.9.1) 的确交换.

现在另取 V 的基 $v_1', \dots, v_n'$ 和 W 的基 $w_1', \dots, w_m'$, 分别简记为 $\mathbf{v}'$ 和 $\mathbf{w}'$. 给定 $T \in \mathrm{Hom}(V, W)$, 眼下的问题是:

$$\text{如何以 } \boldsymbol{\mathcal{M}}_{\mathbf{v}}^{\mathbf{w}}(T) \text{ 来表示 } \boldsymbol{\mathcal{M}}_{\mathbf{v}'}^{\mathbf{w}'}(T) \text{ ?}$$

定义映射

$$\begin{aligned} \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}: F^n &\longrightarrow F^n \\ (x_1', \dots, x_n') &\longmapsto (x_1, \dots, x_n) \end{aligned}$$

其中 $(x_1, \dots, x_n)$ 由等式 $\sum_{i=1}^n x_i v_i = \sum_{i=1}^n x_i' v_i'$ 唯一确定; 换言之, $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 的功能是从 $\mathbf{v}'$ 到 $\mathbf{v}$ 的坐标转换. 敬请读者验证这点也可由交换图表来刻画为:

$$\begin{array}{ccc} & V & \\ {\scriptstyle\varphi_{\mathbf{v}'}}\nearrow & & \nwarrow{\scriptstyle\varphi_{\mathbf{v}}} \\ F^n & \xrightarrow[\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}]{} & F^n \end{array} \tag{4.9.2}$$

我们马上会说明 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 是线性的, 因而可视同 $n \times n$ 矩阵. 在此之前, 观察到映射 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 既然诠释为坐标转换, 对于 V 的任三组有序基 $\mathbf{v}$, $\mathbf{v}'$, $\mathbf{v}''$ 皆有

$$\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} \boldsymbol{P}_{\mathbf{v}''}^{\mathbf{v}'} = \boldsymbol{P}_{\mathbf{v}''}^{\mathbf{v}}, \quad \boldsymbol{P}_{\mathbf{v}}^{\mathbf{v}} = \mathrm{id}_{F^n}.$$

引理 4.9.1 以上的映射 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 是向量空间 F^n 的自同构, 其逆为 $\boldsymbol{P}_{\mathbf{v}}^{\mathbf{v}'}$.

证明 交换图表 (4.9.2) 表明 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} = \varphi_{\mathbf{v}}^{-1}\varphi_{\mathbf{v}'}$; 既然 $\varphi_{\mathbf{v}}$ 和 $\varphi_{\mathbf{v}'}$ 皆线性, $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 也是线性的. 最后, $\boldsymbol{P}_{\mathbf{v}}^{\mathbf{v}'} \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} = \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}'} = \mathrm{id}_{F^n}$ 和 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} \boldsymbol{P}_{\mathbf{v}}^{\mathbf{v}'} = \boldsymbol{P}_{\mathbf{v}}^{\mathbf{v}} = \mathrm{id}_{F^n}$ 说明两者互逆. □

注记 4.9.2 若将每个 v'_i 按照有序基 $\mathbf{v}$ 展成列向量, 记为 $\boldsymbol{v}'_i(\mathbf{v})$, 则 $\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}$ 视同 $n\times n$ 矩阵按列展开无非是

$$\left(\begin{array}{c|c|c} \boldsymbol{v}'_1(\mathbf{v}) & \cdots & \boldsymbol{v}'_n(\mathbf{v}) \end{array}\right).$$

我们称 $\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}$ 为从有序基 $\mathbf{v}'$ 到 $\mathbf{v}$ 的**转换矩阵**. 先前已见转换矩阵皆可逆. 反之也有以下事实:

所有可逆的 $\boldsymbol{P}\in \mathrm{M}_{n\times n}(F)$ 都是 F^n 的某两组有序基之间的转换矩阵.

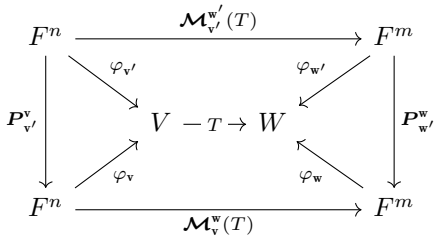
为了说明这点, 不妨就设 $\mathbf{v}$ 为 F^n 的标准基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$, 另外对 $i=1,\dots,n$ 令 $\boldsymbol{v}'_i := \boldsymbol{P}\boldsymbol{e}_i$, 则因为 $\boldsymbol{P}$ 是 F^n 的自同构, $\boldsymbol{v}'_1,\dots,\boldsymbol{v}'_n$ 也构成 F^n 的有序基 $\mathbf{v}'$, 而且它们正是 $\boldsymbol{P}$ 的列向量, 故 $\boldsymbol{P}=\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}$.

类似地, 我们也有 F^m 的自同构 $\boldsymbol{P}^{\mathbf{w}'}_{\mathbf{w}}$, 其逆为 $\boldsymbol{P}^{\mathbf{w}}_{\mathbf{w}'}$. 照例视同 $m\times m$ 矩阵.

定理 4.9.3 对于 V 的任意有序基 $\mathbf{v}$, $\mathbf{v}'$ 和 W 的任意有序基 $\mathbf{w}$ 和 $\mathbf{w}'$, 以及任意 $T\in\mathrm{Hom}(V,W)$, 我们有

$$\begin{aligned}\boldsymbol{\mathcal{M}}^{\mathbf{w}'}_{\mathbf{v}'}(T) &= \boldsymbol{P}^{\mathbf{w}'}_{\mathbf{w}}\boldsymbol{\mathcal{M}}^{\mathbf{w}}_{\mathbf{v}}(T)\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}\\ &= \left(\boldsymbol{P}^{\mathbf{w}}_{\mathbf{w}'}\right)^{-1}\boldsymbol{\mathcal{M}}^{\mathbf{w}}_{\mathbf{v}}(T)\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}.\end{aligned}$$

证明 原式等价于 $\boldsymbol{P}^{\mathbf{w}}_{\mathbf{w}'}\boldsymbol{\mathcal{M}}^{\mathbf{w}'}_{\mathbf{v}'}(T)=\boldsymbol{\mathcal{M}}^{\mathbf{w}}_{\mathbf{v}}(T)\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}$. 等式既可以按照坐标转换函数的刻画来验证, 也可以由交换图表一眼看穿. 虑及读者对交换图表可能欠熟悉, 我们按部就班地操作如下. 首先将 (4.9.1) 和 (4.9.2) 对两种基的情形合并为图表



其中的两个梯形和两个三角形都是交换图表. 问题相当于证箭头 ↘ 和 ↓→ 的合成相同. 为此, 我们在图表中腾挪来验证 $\mathrm{Hom}(F^n,W)$ 中的等式

$$\varphi_{\mathbf{w}}\boldsymbol{P}^{\mathbf{w}}_{\mathbf{w}'}\boldsymbol{\mathcal{M}}^{\mathbf{w}'}_{\mathbf{v}'}(T) \overset{\text{右三角}}{=\!=} \varphi_{\mathbf{w}'}\boldsymbol{\mathcal{M}}^{\mathbf{w}'}_{\mathbf{v}'}(T) \overset{\text{上梯形}}{=\!=} T\varphi_{\mathbf{v}'} \overset{\text{左三角}}{=\!=} T\varphi_{\mathbf{v}}\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'} \overset{\text{下梯形}}{=\!=} \varphi_{\mathbf{w}}\boldsymbol{\mathcal{M}}^{\mathbf{w}}_{\mathbf{v}}(T)\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}.$$

再从两边消去同构 $\varphi_{\mathbf{w}}$, 亦即左合成 $\varphi^{-1}_{\mathbf{w}}$, 即为所求. □

比起定理 4.9.3 的具体公式, 坐标转换的思想和推导技巧毋宁说是更为根本的.

重要的特例是 $V=W$, 也就是考虑自同态 $T\in\mathrm{End}(V)$ 的情形. 这时我们自然希望映射两边使用同一组基.

推论 4.9.4 对于 V 的任意有序基 $\mathbf{v}, \mathbf{v}'$ 以及 $T \in \mathrm{End}(V)$, 我们有

$$\boldsymbol{\mathcal{M}}_{\mathbf{v}'}^{\mathbf{v}'}(T) = \boldsymbol{P}^{-1}\boldsymbol{\mathcal{M}}_{\mathbf{v}}^{\mathbf{v}}(T)\boldsymbol{P},$$

其中 P 是按照注记 4.9.2 的符号给出的转换矩阵:

$$\boldsymbol{P} := \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} := \left(\boldsymbol{v}_1'(\mathbf{v}) \;\middle|\; \cdots \;\middle|\; \boldsymbol{v}_n'(\mathbf{v}) \right),$$
$$\boldsymbol{v}_j'(\mathbf{v}) := v_j' \text{ 按有序基 } \mathbf{v} \text{ 的展开 } \in F^n \quad (\text{列向量}).$$

既然可逆矩阵无非是某两组有序基之间的转换矩阵 (注记 4.9.2), 以下概念便是顺理成章的.

定义 4.9.5 (矩阵的共轭或相似) 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$, 若存在可逆的 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ **共轭**, 又称为**相似**.

共轭的方阵可以设想为同一个线性映射在不同有序基下的化身. 共轭不变的性质应当视为方阵的内蕴性质, 因为它不依赖基的选取.

命题 4.9.6 共轭是 $\mathrm{M}_{n\times n}(F)$ 上的等价关系.

证明 取 $\boldsymbol{P} = \mathbf{1}_{n\times n}$ 可见任何 $\boldsymbol{A}$ 都共轭于自身, 反身性成立. 若 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 则以乘法移项可得 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1} = (\boldsymbol{P}^{-1})^{-1}\boldsymbol{B}\boldsymbol{P}^{-1}$, 对称性成立. 设 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 而 $\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$, 则 $\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Q} = (\boldsymbol{P}\boldsymbol{Q})^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Q})$, 传递性成立. □

对于给定的向量空间同构 $\Phi : V' \xrightarrow{\sim} V$, 我们同样可以考虑类似的共轭运算

$$\begin{array}{ccc} \mathrm{End}(V) & \longrightarrow & \mathrm{End}(V') \\ T & \longmapsto & T' := \Phi^{-1}T\Phi \end{array}$$

而且 T 和 T' 的关系可以按约定 2.3.3 的方式写作交换图表

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ {\scriptstyle\Phi}\uparrow & & \uparrow{\scriptstyle\Phi} \\ V' & \xrightarrow[T']{} & V' \end{array}.$$

只要承认同构 Φ 将向量空间 V 和 V' 的结构丝毫不差地对应起来, 则共轭关系 $T' = \Phi^{-1}T\Phi$ 相当于说 T' 是通过同构 Φ 将 T "搬运" 到 V' 的结果, 它们所有内蕴的代数性质都是相同的.

共轭等价关系将是矩阵的对角化与标准形理论的主要课题.

练习 4.9.7 设 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$ 可逆, 说明 $\boldsymbol{A} \mapsto \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 给出环 $\mathrm{M}_{n\times n}(F)$ 的自同构, 它还是线性的: $\boldsymbol{P}^{-1}(t\boldsymbol{A})\boldsymbol{P} = t\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 对所有 $t \in F$ 成立. 对线性映射给出相应的陈述.

接着介绍一个同样和换基有关, 但比共轭简单得多的等价关系, 称为相抵.

定义 4.9.8 (矩阵的相抵) 称矩阵 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{m\times n}(F)$ 相抵, 如果存在 $\boldsymbol{Q}\in \mathrm{M}_{m\times m}(F)$ 和 $\boldsymbol{P}\in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 皆可逆, 而且

$$\boldsymbol{B}=\boldsymbol{Q}\boldsymbol{A}\boldsymbol{P}.$$

仿照先前的论证方式, 可以看出相抵是 $\mathrm{M}_{m\times n}(F)$ 上的等价关系. 由于命题 4.8.13 已说明可逆矩阵是初等矩阵的乘积, 两个矩阵相抵相当于说它们能够通过一系列初等行变换和列变换相互过渡.

命题 4.9.9 两个矩阵 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{m\times n}(F)$ 相抵的充要条件是 $\mathrm{rk}(\boldsymbol{A})=\mathrm{rk}(\boldsymbol{B})$.

证明 将 $\boldsymbol{A}\in \mathrm{M}_{m\times n}(F)$ 视同线性变换 $F^n\to F^m$. 记 $r:=\mathrm{rk}(\boldsymbol{A})$. 重拾定理 4.8.4 的证明, 取 $\operatorname{im}\boldsymbol{A}$ 的基 $\boldsymbol{w}_1,\dots,\boldsymbol{w}_r$ 和 $\ker\boldsymbol{A}$ 的基 $\boldsymbol{u}_1,\dots,\boldsymbol{u}_k$. 对每个 $1\leqslant i\leqslant r$, 任取 $\boldsymbol{v}_i\in F^n$ 使得 $\boldsymbol{A}\boldsymbol{v}_i=\boldsymbol{w}_i$. 如此则 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_r,\boldsymbol{u}_1,\dots,\boldsymbol{u}_k$ 是 F^n 的有序基, 简记为 $\mathbf{v}$.

另一方面, 将 $\boldsymbol{w}_1,\dots,\boldsymbol{w}_r$ 扩充为 F^m 的有序基 $\boldsymbol{w}_1,\dots,\boldsymbol{w}_m$, 简记为 $\mathbf{w}$. 相对于 $\mathbf{v}$ 和 $\mathbf{w}$, 线性映射 $\boldsymbol{A}$ 另有矩阵表法[7]为

$$\boldsymbol{D}_r:=\begin{array}{c} \\ \boldsymbol{w}_1\\ \vdots\\ \boldsymbol{w}_r\\ \vdots\\ \boldsymbol{w}_m\end{array}\begin{array}{c}\begin{array}{cccccc}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r & \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k\end{array}\\ \left(\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & \boldsymbol{0}_{r\times k} & \\ & & 1 & & & \\ \hline & \boldsymbol{0}_{(m-r)\times r} & & & \boldsymbol{0}_{(m-r)\times k} & \end{array}\right)\end{array} \tag{4.9.3}$$

其中留白部分全为 0. 定理 4.9.3 表明存在可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ (取为适当的转换矩阵), 使得 $\boldsymbol{D}_r=\boldsymbol{Q}\boldsymbol{A}\boldsymbol{P}$. 这就说明任意 $\boldsymbol{A}$ 皆和 $\boldsymbol{D}_{\mathrm{rk}(\boldsymbol{A})}$ 相抵. 因为相抵是等价关系, 由此知 $\mathrm{rk}(\boldsymbol{A})=\mathrm{rk}(\boldsymbol{B})$ 蕴涵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相抵. 反之若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相抵, 则注记 4.8.11 表明 $\mathrm{rk}(\boldsymbol{A})=\mathrm{rk}(\boldsymbol{B})$. 明所欲证. □

命题 4.9.9 的证明相当于说明每个相抵等价类都包含形如 $\boldsymbol{D}_r$ 的元素, 并且由于 $\mathrm{rk}(\boldsymbol{D}_r)=r$, 在 $r\neq s$ 时 $\boldsymbol{D}_r$ 和 $\boldsymbol{D}_s$ 互不相抵. 若矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{D}_r$ 相抵, 则称 $\boldsymbol{D}_r$ 为 $\boldsymbol{A}$ 的**相抵标准形**.

算法 4.9.10 这里勾勒相抵标准形的一种具体算法. 首先对矩阵 $\boldsymbol{A}$ 作初等行变换化为行梯矩阵, 再对后者作适当的列变换, 使得每个主元的上下左右所有元素变为 0; 这种矩阵可以经过适当的初等行变换或列变换 (重排行或列) 整理成 (4.9.3) 的形式. 这给出命题 4.9.9 的另一种证明.

[7] 这是 §4.11 行将介绍的分块矩阵表法, 但此处的涵义是明白的.

现在我们可以运用矩阵的语言证明行秩和列秩相等.

定理 4.9.11 任意矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 的行秩皆等于列秩.

证明 问题相当于证明 $\mathrm{rk}(\boldsymbol{A}) = \mathrm{rk}\,({}^{\mathrm{t}}\boldsymbol{A})$. 取可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 使得 $\boldsymbol{D}_r = \boldsymbol{Q}\boldsymbol{A}\boldsymbol{P}$, 其中 $r = \mathrm{rk}(\boldsymbol{A})$. 根据注记 4.8.11,

$$\begin{aligned}\mathrm{rk}(\boldsymbol{A}) &= \mathrm{rk}(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{P}) = \mathrm{rk}(\boldsymbol{D}_r)\\ &= r = \mathrm{rk}\,({}^{\mathrm{t}}\boldsymbol{D}_r) \quad (\text{从 (4.9.3) 一眼看穿})\\ &= \mathrm{rk}\left({}^{\mathrm{t}}(\boldsymbol{P}^{-1})\cdot {}^{\mathrm{t}}\boldsymbol{D}_r \cdot {}^{\mathrm{t}}(\boldsymbol{Q}^{-1})\right) = \mathrm{rk}\,({}^{\mathrm{t}}\boldsymbol{A})\,.\end{aligned}$$

明所欲证. □

以上论证是矩阵操作和线性映射的混搭, §8.9 将从对偶空间视角给出抽象证明.

留意到相抵标准形 $\boldsymbol{D}_r$ 在 (4.9.3) 中的形状已经表明 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 的秩满足

$$r := \mathrm{rk}(\boldsymbol{A}) \leq \min\{m, n\}.$$

⋆ 既然 r 也等于 $\boldsymbol{A}$ 的简化行梯矩阵的主元数 (命题 4.8.12), 这个不等式同时又是练习 1.3.5 的内容.

⋆ 如果采取抽象观点, 则对于有限维向量空间之间的线性映射 $T : V \to W$, 相应的不等式写作 $\mathrm{rk}(T) \leq \{\dim V, \dim W\}$. 这点既可以取基翻译为矩阵版本, 也可以从定理 4.8.4 毫不费力地推导, 敬请读者自行练习.

定义 4.9.12 若 $\mathrm{rk}(\boldsymbol{A}) = \min\{m, n\}$, 则称 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 是**满秩**的.

当 $m = n$ 时, 命题 4.8.13 业已说明满秩和可逆是一回事. 对于一般情形, 谨记录一则方便的而简单的引理.

引理 4.9.13 设矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 满秩. 若 $n \geq m$, 则可以从 $\boldsymbol{A}$ 删去 $n-m$ 列使得剩下的 $m \times m$ 矩阵可逆; 若 $m \geq n$, 则可以从 A 删去 $m-n$ 行使得剩下的 $n \times n$ 矩阵可逆.

证明 设 $n \geq m$. 既然 $\boldsymbol{A}$ 满秩, 而行秩等于列秩 (定理 4.9.11), 执行算法 4.4.8 便可从 $\boldsymbol{A}$ 取出 m 个线性无关的列, 它们构成的矩阵可逆. 至于 $n \leq m$ 的情形完全类似, 也可以通过转置化到前一情形. □

4.10 直和分解

选定域 F. 对于给定的一族 F-向量空间 $(V_i)_{i\in I}$, 例 4.2.10 定义了它们的直和 $\bigoplus_{i\in I} V_i$, 其元素表成向量组 $(v_i)_{i\in I}$ 的形式, 至多仅有有限个分量 v_i 非零; 每个 V_i 都自然地嵌入直和作为其子空间. 当 $I=\varnothing$ 时, 对应的直和规定为 $\{0\}$ (约定 4.2.11).

类比于 §2.3 探讨的无交并, 这种直和可谓 "外在" 的: 它从 $(V_i)_{i\in I}$ 通过向量组构造新的空间, 而这些空间是否有非零交, 甚至相等, 在此无关宏旨. 暂且称这种构造为外直和, 并且启用临时的记号 $\overset{\text{外}}{\bigoplus}_{i\in I} V_i$ 来标注.

我们也可以调换视角, 对于给定的 F-向量空间 V 及一族子空间 V_i, 下标 i 遍历某个集合 I, 我们想知道 V 能否等同于它们的直和. 何谓等同? 无论 $(V_i)_{i\in I}$ 如何取, 总是有从外直和映向 V 的线性映射

$$\begin{aligned} \sigma:\ \overset{\text{外}}{\bigoplus}_{i\in I} V_i &\longrightarrow V \\ (v_i)_{i\in I} &\longmapsto \textstyle\sum_{i\in I} v_i. \end{aligned} \tag{4.10.1}$$

由于至多有限个 v_i 非零, $\sum_{i\in I} v_i$ 总是 V 中的有限和, 故 σ 的定义合理. 先前的问题可以更精确地改述为:

何时能确保 σ 是向量空间的同构?

此一问题既是自然的, 也是实际的, 因为一旦 σ 是同构, 关于 V 的研究便有望在 V_i 上各个击破.

定义 4.10.1 给定 F-向量空间 V, 一族子空间 $(V_i)_{i\in I}$ 的和是

$$\sum_{i\in I} V_i := \left\{ \text{有限和 } \sum_i v_i \in V : \forall i,\ v_i \in V_i \right\}.$$

对于有限多个子空间的和, 诸如 $V_1+\cdots+V_n$ 的写法也是通行的; 对应于 $I=\varnothing$ 的空和规定为零空间 $\{0\}$.

不难看出 $\sum_{i\in I} V_i$ 总是对加法和纯量乘法封闭, 并且包含 0, 因而是 V 的子空间. 实际上, 它是 V 中包含每个 V_i 的最小子空间.

类似地, 对一族子空间 V_i 可以取交 $\bigcap_{i\in I} V_i$, 它包含 0 并且对加法和纯量乘法封闭, 因而是 V 的子空间; 事实上, $\bigcap_{i\in I} V_i$ 是包含于每个 V_i 的最大子空间[8].

回到 (4.10.1), 按定义立见 $\operatorname{im}(\sigma)=\sum_{i\in I} V_i$. 关键问题遂归结为: σ 何时给出从外直和到 $\sum_{i\in I} V_i$ 的同构?

[8] 尽管我们在讨论集合时禁止空交, 但在这里将子空间的空交 (对应于 $I=\varnothing$) 规定为 V 是合理的.

定义–命题 4.10.2 设 $(V_i)_{i\in I}$ 为 F-向量空间 V 的一族子空间, $I\neq\varnothing$. 若对每个 $i\in I$ 都有

$$V_i\cap\sum_{j\neq i}V_j=\{0\},$$

则我们将 $\sum_{i\in I}V_i$ 记为 $\bigoplus_{i\in I}V_i$, 称为 $(V_i)_{i\in I}$ 的内直和, 不致混淆时简称**直和**, 而每个 V_i 皆称为其中的**直和项**.

以上条件成立当且仅当 (4.10.1) 的 σ 给出从 $(V_i)_{i\in I}$ 的外直和到 $\sum_{i\in I}V_i$ 的同构.

证明 问题相当于问 σ 何时是单射. 从定义当下看出 σ 单当且仅当

$$\text{有限和 }\sum_{i\in I}v_i=\sum_{i\in I}v_i'\iff\forall i\in I,\ v_i=v_i',$$

其中 $v_i,v_i'\in V_i$, 而上式又可以改述为

$$\sum_{i\in I}v_i=0\iff\forall i\in I,\ v_i=0.$$

若以上等价不成立, 则存在线性关系 $\sum_{j\in I}v_j=0$ 连同某个 $i\in I$, 使得 $v_i\neq0$; 这又导致

$$0\neq v_i=-\sum_{j\neq i}v_j\in V_i\cap\sum_{j\neq i}V_j.$$

反之, 若 $V_i\cap\sum_{j\neq i}V_j\neq\{0\}$, 则任选其中的非零元 v_i 并且循 $\sum_{j\neq i}V_j$ 作展开, 又得到非平凡的线性关系. 至此, 我们业已探明了展开式 $v=\sum_{i\in I}v_i$ 的唯一性, 或者说是 σ 的单性的障碍所在. □

不妨这么看: 外直和从一族空间构造新空间, 内直和将给定的空间作分解.

对于任意一族向量空间 $(V_i)_{i\in I}$ 的外直和 $\overset{\text{外}}{\bigoplus_i}V_i$, 我们照例将每个 V_i 视同外直和的子空间, 仍称为其直和项. 对外直和中的每个向量 v 都可以提取其非零分量, 以唯一地表为有限和 $v=\sum_{i\in I}v_i$, 从而关于内直和的条件在此自动成立. 如此一来便产生了一种近乎绕口令的说法 — 外直和等于其直和项的内直和. 既然两种直和可以通过 σ 来等同, 今后直和不必再分内外, 也可以用 $\bigoplus_{i\in I}V_i$ 来代替别扭的符号 $\overset{\text{外}}{\bigoplus_{i\in I}}V_i$.

例 4.10.3 最常见的是两个子空间 $V_1,V_2\subset V$ 的直和 (即 $|I|=2$), 此时定义–命题 4.10.2 的条件简化为 $V_1\cap V_2=\{0\}$.

例 4.10.4 设 $(v_i)_{i\in I}$ 为 F-向量空间 V 的一组基, 则因为 V 的所有元素都能唯一地展开为 $\sum_{i\in I}c_iv_i$ 的形式, 故 $V=\bigoplus_{i\in I}Fv_i$.

练习 4.10.5 对于有限维向量空间 V 的任意子空间 V_1,V_2, 证明

$$\dim(V_1\cap V_2)+\dim(V_1+V_2)=\dim V_1+\dim V_2.$$

提示 外在地构造直和 $V_1\oplus V_2$ 及线性映射 $\partial:V_1\oplus V_2\to V$, 映 (v_1,v_2) 为 v_1-v_2. 证明 $\operatorname{im}(\partial)=V_1+V_2$ 而 $\ker(\partial)\simeq V_1\cap V_2$. 应用定理 4.8.4.

事实上, 向量空间的任何子空间都可以实现为直和项, 尽管具体实现方式取决于基的选择, 并非标准的.

命题 4.10.6 设 V 为 F-向量空间, $V_0 \subset V$ 为任意子空间, 则存在子空间 $V_1 \subset V$ 使得 $V = V_0 \oplus V_1$.

证明 我们承认 V_0 总是有一组基 $\{v_i\}_{i\in I_0}$, 并且可以扩充为 V 的一组基 $\{v_i\}_{i\in I}$, 其中 $I \supset I_0$ (定义–命题 4.4.10). 取 $V_1 := \sum_{i\in I\smallsetminus I_0} Fv_i$, 从而 $\{v_i\}_{i\in I\smallsetminus I_0}$ 是 V_1 的基. 由基的定义立见任意 $v \in V$ 都能唯一地写成 $v = v_0 + v_1$, 其中 $v_i \in V_i$. □

回到一般的直和 $\bigoplus_{j\in I} V_j$. 对每个 $i \in I$ 都有直和项的嵌入映射 ι_i 和投影映射 p_i:

$$V_i \xrightarrow{\text{嵌入 } \iota_i} \bigoplus_{j\in I} V_j \xrightarrow{\text{投影 } p_i} V_i$$

$$v_i \longmapsto \text{第 } i \text{ 个分量为 } v_i, \text{ 其余为 } 0$$

$$(v_j)_{j\in I} \longmapsto v_i.$$

两者都是线性的. 显然 ι_i 单而 p_i 满. 以上将 $\bigoplus_{j\in I} V_j$ 理解为外直和, 从而将其元素理解为向量组 $(v_j)_{j\in I}$; 若作内直和来理解, 则 ι_i 变为包含映射, 而 p_i 的效果是萃取 $v = \sum_{j\in I} v_j \in \bigoplus_{j\in I} V_j$ 的分量 v_i. 如前所述, 内外两种观点很容易自由切换.

练习 4.10.7 验证直和带有的嵌入和投影满足以下等式.

$$\begin{gathered} p_i\iota_i = \mathrm{id}_{V_i}, \\ i \neq j \implies p_j\iota_i = 0, \\ (\iota_i p_i)^2 = \iota_i p_i, \end{gathered}$$

而且当 I 有限时

$$\sum_{i\in I} \iota_i p_i = \mathrm{id}_{\bigoplus_{i\in I} V_i}.$$

练习 4.10.8 考虑向量空间 V 和 $P_1, \dots, P_s \in \mathrm{End}(V)$. 设

$$P_1 + \cdots + P_s = \mathrm{id}, \quad P_iP_j = \begin{cases} P_i, & i = j, \\ 0, & i \neq j. \end{cases}$$

对每个 i, 命 $V_i := \mathrm{im}(P_i) \subset V$. 证明

(i) $V = V_1 \oplus \cdots \oplus V_s$,

(ii) P_i 映 $v_1 + \cdots + v_s \in V$ 为 v_i, 其中 $v_j \in V_j$. 换言之, P_i 是向直和项 V_i 的投影.

(iii) 作为特例, 说明若 $P \in \mathrm{End}(V)$ 满足 $P^2 = P$, 则有直和分解

$$V = \mathrm{im}(P) \oplus \mathrm{im}(\mathrm{id} - P).$$

目光转向线性映射. 设 $T : V \to W$ 为线性映射, 而 V 和 W 各自带有有限直和分解

$$V = V_1 \oplus \cdots \oplus V_n, \quad W = W_1 \oplus \cdots \oplus W_m. \tag{4.10.2}$$

如何将 T 的描述相应地化到各个直和项 V_j, W_i 上?

1. 按照用矩阵表达线性映射的思路, 对每个 $1 \le i \le m$ 和 $1 \le j \le n$, 定义线性映射

$$\begin{aligned} T_{ij} := V_j \xrightarrow{\iota_j} V \xrightarrow{T} W \xrightarrow{p_i} W_i \quad \text{的合成} \\ = p_i T \iota_j \end{aligned}$$

   如此则 $T\iota_j(v_j) = (T_{1j}(v_j), \ldots, T_{mj}(v_j))$ 对所有 j 和 $v_j \in V_j$ 皆成立. 当 i 和 j 变动, 这说明映射族 $T_{ij} \in \mathrm{Hom}(V_j, W_i)$ 唯一地确定了 T.

2. 反过来说, 给定一族线性映射 $T_{ij} \in \mathrm{Hom}(V_j, W_i)$, 其中 $1 \le i \le m$ 而 $1 \le j \le n$, 总能够定义线性映射 $T : V \to W$ 使得它限制在直和项 V_j 上按坐标展开等于 $(T_{1j}, \ldots, T_{mj})$.

3. 双向操作显然是互逆的, 以上讨论总结为双射

$$\begin{aligned} \mathrm{Hom}(V, W) &\xrightarrow{1:1} \prod_{\substack{1\le i\le m \\ 1\le j\le n}} \mathrm{Hom}(V_j, W_i) \\ T &\longmapsto (T_{ij})_{i,j}. \end{aligned} \tag{4.10.3}$$

其次, 我们还希望知道线性映射的操作如何反映. 如果相对于同一组直和分解 (4.10.2), $T \in \mathrm{Hom}(V, W)$ 对应到 $(T_{ij})_{i,j}$, 而 $T' \in \mathrm{Hom}(V, W)$ 对应到 $(T'_{ij})_{i,j}$, 则 $T + T'$ 显然对应到逐项相加 $(T_{ij} + T'_{ij})_{i,j}$. 若 $t \in F$, 则纯量乘法 tT 显然对应到逐项纯量乘法 $(tT_{ij})_{i,j}$.

由此可知 (4.10.3) 是向量空间的同构, 其右式按照例 4.2.9 赋予向量空间结构; 由于此处仅涉及有限个 (i, j), 右式也等于向量空间的直和 $\bigoplus_{i,j} \mathrm{Hom}(V_j, W_i)$.

若在 (4.10.3) 中取 $W = F$, 相应地 $m = 1$, 则它给出

$$V^\vee \simeq V_1^\vee \oplus \cdots \oplus V_n^\vee.$$

这相当于说精确到自然同构, 有限直和与取对偶这两种操作可以交换顺序. 无穷直和的状况与此大不相同, 留作练习 4.10.11.

言归正传. 重点在于映射合成在 (4.10.3) 下的描述. 考虑向量空间 U, V, W, 各自带有直和分解

$$U = \bigoplus_{k\in K} U_k, \quad V = \bigoplus_{j\in J} V_j, \quad W = \bigoplus_{i\in I} W_i, \tag{4.10.4}$$

其中 I, J, K 是充当下标的有限集. 相应的嵌入和投影映射照例记为 ι_i, p_i 等. 考虑线性映射

$$U \xrightarrow{T} V \xrightarrow{S} W.$$

对任意 $(i,k)\in I\times K$, 应用直和项的嵌入和投影, 可以简便地计算

$$\begin{aligned}(ST)_{ik} &= p_i S T \iota_k \\ &= \sum_{j\in J} p_i S \iota_j p_j T \iota_k \qquad \because \sum_{j\in J}\iota_j p_j = \mathrm{id}_V \\ &= \sum_{j\in J} S_{ij}T_{jk}.\end{aligned} \tag{4.10.5}$$

这表明直和分解将线性映射的合成转译为类似于矩阵乘法的操作.

注记 4.10.9 如果 (4.10.2) 的直和分解都来自于基, 如 $V_j = Fv_j := \langle v_j\rangle$, $W_i = Fw_i := \langle w_i\rangle$ 等, 则基的选取给出 $\mathrm{Hom}(Fv_j, Fw_i)\simeq \mathrm{Hom}(F,F)\simeq F$, 从而 $T_{ij}\in \mathrm{Hom}(Fw_i, Fv_j)\simeq F$. 所以线性映射的矩阵表法是 (4.10.3) 的特例.

练习 4.10.10 设 U 为 $\mathrm{M}_{n\times n}(F)$ 的 $n+1$ 维子空间, 证明存在 $A\in U\smallsetminus\{\mathbf{0}_{n\times n}\}$ 使得 A 不可逆.

提示 一种方法是命 U' 为首列为零的 $n\times n$ 矩阵所成之子空间, 说明 $\dim U'\cap U>0$, 这一步需要练习 4.10.5. 另一种方法是任取非零列向量 $\boldsymbol{v}\in F^n$, 说明 $\boldsymbol{A}\mapsto \boldsymbol{Av}$ 是从 U 到 F^n 的线性映射, 故它的核非零.

练习 4.10.11 考虑两族 F-向量空间 $(V_j)_{j\in J}$ 和 $(W_i)_{i\in I}$, 试验证 F-向量空间的自然同构

$$\begin{aligned}\mathrm{Hom}\left(\textstyle\bigoplus_{j\in J}V_j, \prod_{i\in I}W_i\right) &\xrightarrow{\sim} \textstyle\prod_{(i,j)\in I\times J}\mathrm{Hom}(V_j,W_i)\\ T &\longmapsto (p_iT\iota_j)_{(i,j)\in I\times J}\end{aligned}$$

其中 $\iota_j: V_j\to\bigoplus_{j'}V_{j'}$ 是自然嵌入, $p_i:\prod_{i'}W_{i'}\to W_i$ 是自然投影. 作为特例, 说明直和的对偶空间满足以下同构:

$$\Big(\bigoplus_{j\in J}V_j\Big)^\vee \xrightarrow{\sim} \prod_{j\in J}V_j^\vee.$$

利用上述练习的最后一段, 能够说明当 V 维数无穷时, 对偶空间 $V^\vee$ 单单视作集合便有超过 $|V|$ 的基数, 其维数遂严格大于 $\dim V$. 为简单起见, 下面取 F 为仅有两个元素的有限域 $\mathbb{F}_2$ (见 §3.1) 来说明. 视 $\mathbb{F}_2$ 为 1 维 $\mathbb{F}_2$-向量空间, $\mathbb{F}_2^\vee\simeq\mathbb{F}_2$, 再以练习 4.10.11 和标准的符号写下同构

$$\underbrace{\left(\mathbb{F}_2^{\oplus\mathbb{Z}_{\geq 0}}\right)}_{\mathbb{Z}_{\geq 0}\text{ 份 }\mathbb{F}_2\text{ 的直和}}{}^{\vee} \simeq \left(\mathbb{F}_2^\vee\right)^{\mathbb{Z}_{\geq 0}} \simeq \underbrace{\left(\mathbb{F}_2\right)^{\mathbb{Z}_{\geq 0}}}_{\mathbb{Z}_{\geq 0}\text{ 份 }\mathbb{F}_2\text{ 的直积}}.$$

右式作为集合的基数是 $2^{|\mathbb{Z}_{\geq 0}|}=2^{\aleph_0}$. 另一方面, 因为直和只涉及有限和, 推论 2.9.10 确保可数并

$$\mathbb{F}_2^{\oplus\mathbb{Z}_{\geq 0}} = \bigcup_{n\geq 1}\underbrace{\mathbb{F}_2^{\{0,\dots,n-1\}}}_{\text{有限集}}$$

的基数必 $\leq \aleph_0$; 因为这是无穷集, 其基数也只能是 $\aleph_0$. 然而 Cantor 定理 2.9.12 断言 $\aleph_0 < 2^{\aleph_0}$.

4.11 分块矩阵运算

本节目的是将 §4.10 的讨论落实到矩阵计算. 设 V 和 W 为有限维 F-向量空间, 而且具备如 (4.10.2) 的直和分解; 我们对其中每个直和项 V_j 和 W_i 都选定有序基. 根据命题 4.4.9, 这些有序基的无交并分别给出 V 和 W 的有序基. 因此, $T \in \mathrm{Hom}(V,W)$ 和每个 $T_{ij} \in \mathrm{Hom}(V_j, W_i)$ 都有相应的矩阵表法, 按 §4.6 的成例, 记为 $\boldsymbol{\mathcal{M}}(T)$ 和 $\boldsymbol{\mathcal{M}}(T_{ij})$ 等.

矩阵 $\boldsymbol{A} := \boldsymbol{\mathcal{M}}(T)$ 按此分割为 mn 个子矩阵的并排, 形如

$$\boldsymbol{A} = \begin{array}{c} \\ W_1 \\ \vdots \\ W_m \end{array} \begin{array}{c} \begin{array}{ccc} V_1 & \cdots & V_n \end{array} \\ \left(\begin{array}{c|c|c} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1n} \\ \hline \vdots & \ddots & \vdots \\ \hline \boldsymbol{A}_{m1} & \cdots & \boldsymbol{A}_{mn} \end{array}\right) \end{array}, \qquad \boldsymbol{A}_{ij} := \boldsymbol{\mathcal{M}}(T_{ij}).$$

带有这般分割的矩阵 $\boldsymbol{A}$ 称为**分块矩阵**, 其中的 $\boldsymbol{A}_{ij}$ 称为它的 (i,j)-分块.

- ⋆ 线性映射的相加和纯量乘法对应到分块矩阵的逐块相加和纯量乘法; 当然, 相加时要求两个矩阵按照相同规格分块.
- ⋆ 对于线性映射 $S: V \to W$ 和 $T: U \to V$ 的合成, 公式 (4.10.5) 表明矩阵 $\boldsymbol{\mathcal{M}}(ST)$ 的 (i,k)-分块 $\boldsymbol{\mathcal{M}}((ST)_{ik})$ 等于

  $$\sum_{j\in J} \boldsymbol{\mathcal{M}}(S_{ij})\boldsymbol{\mathcal{M}}(T_{jk}),$$

  其中 $(i,k) \in I \times K$ 是任意的. 这相当于说分块矩阵相乘也服从矩阵乘法的规律, 然而这是将各个分块作为矩阵来相乘, 而非域 F 里的乘法; 分块之间的相乘一般不能任意调换顺序.
- ⋆ 分块矩阵相乘时须要求两个分块矩阵的规格兼容, 以确保分块乘法有意义. 按照向量空间观点, 关于分块规格的条件都源自 (4.10.4) 中各个直和项的维数.

例 4.11.1 将矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 按列 (或行) 分块

$$\boldsymbol{A} = \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{array}\right) = \left(\begin{array}{c} \boldsymbol{w}_1 \\ \hline \vdots \\ \hline \boldsymbol{w}_m \end{array}\right)$$

其中每个 $\boldsymbol{v}_j$ (或 $\boldsymbol{w}_i$) 都是 m 维列向量 (或 n 维行向量), 则对任何 $\boldsymbol{Q} \in \mathrm{M}_{r\times m}(F)$ (或 $\boldsymbol{P} \in \mathrm{M}_{n\times r}(F)$) 皆有

$$\boldsymbol{QA} = \left(\begin{array}{c|c|c} \boldsymbol{Q}\boldsymbol{v}_1 & \cdots & \boldsymbol{Q}\boldsymbol{v}_n \end{array} \right) \in \mathrm{M}_{r\times n}(F),$$

$$\boldsymbol{AP} = \left(\begin{array}{c} \boldsymbol{w}_1\boldsymbol{P} \\ \hline \vdots \\ \hline \boldsymbol{w}_m\boldsymbol{P} \end{array} \right) \in \mathrm{M}_{m\times r}(F).$$

这既是矩阵乘法定义的简单结论, 也是分块乘法的例子.

例 4.11.2 矩阵的增广也可以看作一种分块结构. 在求逆算法 4.8.14 中运用了形如

$$\boldsymbol{U}\,(\boldsymbol{A}|\boldsymbol{B}) = (\boldsymbol{U}\boldsymbol{A}|\boldsymbol{U}\boldsymbol{B})$$

的分块乘法, 其中 $\boldsymbol{U} \in \mathrm{M}_{m\times m}(F)$, $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$, $\boldsymbol{B} \in \mathrm{M}_{m\times h}(F)$

例 4.11.3 考虑形如

$$\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{C} & \boldsymbol{D} \end{array} \right), \quad \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D} \in \mathrm{M}_{n\times n}(F)$$

的 $2n \times 2n$ 分块矩阵, 其间的乘法满足

$$\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{C} & \boldsymbol{D} \end{array} \right) \left(\begin{array}{c|c} \boldsymbol{A}' & \boldsymbol{B}' \\ \hline \boldsymbol{C}' & \boldsymbol{D}' \end{array} \right) = \left(\begin{array}{c|c} \boldsymbol{A}\boldsymbol{A}' + \boldsymbol{B}\boldsymbol{C}' & \boldsymbol{A}\boldsymbol{B}' + \boldsymbol{B}\boldsymbol{D}' \\ \hline \boldsymbol{C}\boldsymbol{A}' + \boldsymbol{D}\boldsymbol{C}' & \boldsymbol{C}\boldsymbol{B}' + \boldsymbol{D}\boldsymbol{D}' \end{array} \right)$$

这和 2×2 矩阵乘法类似, 但须注意分块之间的乘法不能随意交换. 更多分块的情形准此可知。

练习 4.11.4 考虑如下的 $(n+1) \times (n+1)$ 分块矩阵

$$\left(\begin{array}{ccc|c} & & & \\ & \boldsymbol{A} & & \boldsymbol{v} \\ & & & \\ \hline 0 & \cdots & 0 & 1 \end{array} \right), \quad \boldsymbol{A} \in \mathrm{M}_{n\times n}(F),\ \boldsymbol{v} \in \mathrm{M}_{n\times 1}(F) \simeq F^n.$$

验证其间的乘法满足

$$\left(\begin{array}{ccc|c} & \boldsymbol{A} & & \boldsymbol{v} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right)\left(\begin{array}{ccc|c} & \boldsymbol{B} & & \boldsymbol{w} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right)=\left(\begin{array}{ccc|c} & \boldsymbol{AB} & & \boldsymbol{Aw}+\boldsymbol{v} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right),$$

其中 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$ 和列向量 $\boldsymbol{v},\boldsymbol{w}\in F^n$ 均为任意的.

另一方面, 对每组资料 $(\boldsymbol{A},\boldsymbol{v})$ 都能定义从列向量空间 F^n 到其自身的映射

$$\Phi_{\boldsymbol{A},\boldsymbol{v}}: F^n\to F^n,\quad \Phi_{\boldsymbol{A},\boldsymbol{v}}(\boldsymbol{u})=\boldsymbol{Au}+\boldsymbol{v}.$$

请验证映射的等式

$$\Phi_{\boldsymbol{A},\boldsymbol{v}}\Phi_{\boldsymbol{B},\boldsymbol{w}}=\Phi_{\boldsymbol{AB},\boldsymbol{Aw}+\boldsymbol{v}},$$

以及矩阵等式

$$\left(\begin{array}{ccc|c} & \boldsymbol{A} & & \boldsymbol{v} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right)\left(\begin{array}{c} \boldsymbol{u} \\ \hline 1 \end{array}\right)=\left(\begin{array}{c} \Phi_{\boldsymbol{A},\boldsymbol{v}}(\boldsymbol{u}) \\ \hline 1 \end{array}\right)$$

所以这种分块矩阵是非线性映射 $\Phi_{\boldsymbol{A},\boldsymbol{v}}$ 的一种矩阵化身, 使得映射合成反映为矩阵乘法. 以后探讨仿射空间时还会回到这个问题.

现在考虑有限维向量空间 V 以及 $T\in\mathrm{End}(V)$. 给定直和分解 $V=V_1\oplus\cdots\oplus V_n$, 我们将 T 按 (4.10.3) 表达为 $(T_{ij})_{1\le i,j\le n}$; 对每个 V_i 取有序基, 相应地将 T 对应的矩阵 $\boldsymbol{A}$ 表达为分块矩阵 $(\boldsymbol{A}_{ij})_{1\le i,j\le n}$.

定义 4.11.5 对于如上的 T (或 $\boldsymbol{A}$),

- ⋆ 若 $i\neq j\implies T_{ij}=0$ (或 $\boldsymbol{A}_{ij}=0$), 则称 T (或 $\boldsymbol{A}$) 相对于此直和分解 (或分块) 是**对角**的, 写作
$$T=\mathrm{diag}(T_{11},\ldots,T_{nn});$$
- ⋆ 若 $i>j\implies T_{ij}=0$ (或 $\boldsymbol{A}_{ij}=0$), 则称 T (或 $\boldsymbol{A}$) 相对于此直和分解 (或分块) 是**上三角**的;
- ⋆ 若 $i<j\implies T_{ij}=0$ (或 $\boldsymbol{A}_{ij}=0$), 则称 T (或 $\boldsymbol{A}$) 相对于此直和分解 (或分块) 是**下三角**的;

显然 $0_V = \mathrm{diag}(0_{V_1}, \dots, 0_{V_n})$ 而 $\mathrm{id}_V = \mathrm{diag}(\mathrm{id}_{V_1}, \dots, \mathrm{id}_{V_n})$.

按线性映射的观点, 定义 4.11.5 可作以下诠释: 相对于直和分解 $V = V_1 \oplus \cdots \oplus V_n$,

⋆ T 是对角的当且仅当

$$\forall 1 \le i \le n, \quad T(V_i) \subset V_i;$$

⋆ T 是上三角的当且仅当

$$\forall 1 \le i \le n, \quad T(V_i) \subset \bigoplus_{j \le i} V_j;$$

⋆ T 是下三角的当且仅当

$$\forall 1 \le i \le n, \quad T(V_i) \subset \bigoplus_{j \ge i} V_j.$$

由此也可以看出: 为了定义上三角的概念, 完整的直和分解 $V = V_1 \oplus \cdots \oplus V_n$ 并非必要. 所需的仅是 V 的一列子空间

$$\{0\} = \mathcal{V}_0 \subsetneq \mathcal{V}_1 \subsetneq \cdots \subsetneq \mathcal{V}_n = V,$$

使得 $T(\mathcal{V}_i) \subset \mathcal{V}_i$ 对所有 i 成立; 直和分解的情形对应到 $\mathcal{V}_i := \bigoplus_{j \le i} V_j$. 这样一列子空间称为 V 中长度为 n 的**旗**. 下三角的情形类似.

命题 4.11.6 相对于给定的直和分解 (或分块结构), 线性映射 (或分块矩阵) 的乘法满足

$$\text{对角} \cdot \text{对角} = \text{对角}, \quad \text{上三角} \cdot \text{上三角} = \text{上三角}, \quad \text{下三角} \cdot \text{下三角} = \text{下三角}.$$

无论上述哪一类线性映射 (或矩阵) 的乘法, 对角分块都由逐项相乘给出:

$$\forall 1 \le i \le n, \quad (ST)_{ii} = S_{ii}T_{ii}, \quad (\text{或 } (\boldsymbol{AB})_{ii} = \boldsymbol{A}_{ii}\boldsymbol{B}_{ii}).$$

证明 对于线性映射的版本, 这是分块乘法公式 (4.10.5) 的应用. 设 $S, T \in \mathrm{End}(V)$ 上三角, 则

$$(ST)_{ij} = \sum_{k=1}^{n} S_{ik}T_{kj};$$

为了使 $S_{ik}, T_{kj} \neq 0$, 必要条件是 $i \le k \le j$, 这就表明 ST 也是上三角的; 而且对于 $i = j$ 的对角元, 和式中只有 $i = k = j$ 一项非零, 此即 $(ST)_{ii} = S_{ii}T_{ii}$.

至于 S, T 同为对角或下三角的情形, 论证全然相似. □

推论 4.11.7 设相对于给定的直和分解 (或分块结构), 线性映射 T (或分块矩阵 $\boldsymbol{A}$) 是上三角的, 则对于任意多项式 $g = \sum_n a_n X^n \in F[X]$, 代值得到的线性映射 $g(T) := \sum_n a_n T^n$ (或分块矩阵 $g(\boldsymbol{A}) := \sum_n a_n \boldsymbol{A}^n$) 对之仍是上三角的.

以下三角或对角代替上三角, 结论亦同.

证明 加法和纯量乘法既然是逐分块操作, 当然保持上三角性质. 命题 4.11.6 说明乘法也保持上三角性质. 至于下三角或对角的情形, 理由相同. □

上三角或下三角矩阵的可逆性容易化约到对角线上来检验.

推论 4.11.8 若 $T \in \mathrm{End}(V)$ 相对于给定的直和分解 $V = \bigoplus_{i=1}^{n} V_i$ 是上三角 (或下三角, 对角) 的, 则有

$$T \text{ 可逆} \iff \forall 1 \leq i \leq n,\ T_{ii} \text{ 可逆},$$

而且此时 T^{-1} 也是上三角 (或下三角, 对角) 的, 而且其对角分块依序是 $T_{11}^{-1}, \ldots, T_{nn}^{-1}$. 如果 $T = \mathrm{diag}(T_{11}, \ldots, T_{nn})$ 可逆, 则 $T^{-1} = \mathrm{diag}\left(T_{11}^{-1}, \ldots, T_{nn}^{-1}\right)$.

证明 可逆性的刻画可以用第五章行将介绍的行列式理论来处理, 但以分块运算来验证也毫不困难. 对 T 借用分块矩阵表法. 首先, 基于 "块中有块" 的观察

$$T = \begin{pmatrix} \boxed{T_{11}} & \cdots & & \\ & \ddots & \vdots & \vdots \\ & & \boxed{T_{n-1,n-1}} & \\ & & & \boxed{T_{nn}} \end{pmatrix}$$

或者说是直和的简并 $V = (V_1 \oplus \cdots \oplus V_{n-1}) \oplus V_n$, 容易将问题简化到 $n = 2$ 的情形.

设 T 可逆, 将 T^{-1} 按照相同规格分块表成

$$T^{-1} = \left(\begin{array}{c|c} T'_{11} & T'_{12} \\ \hline T'_{21} & T'_{22} \end{array}\right);$$

按此作分块乘法, 则因为

$$\left(\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22} \end{array}\right)\left(\begin{array}{c|c} T'_{11} & T'_{12} \\ \hline T'_{21} & T'_{22} \end{array}\right) = \left(\begin{array}{c|c} * & * \\ \hline * & T_{22}T'_{22} \end{array}\right) = \mathrm{id}_V,$$

$$\left(\begin{array}{c|c} T'_{11} & T'_{12} \\ \hline T'_{21} & T'_{22} \end{array}\right)\left(\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22} \end{array}\right) = \left(\begin{array}{c|c} T'_{11}T_{11} & * \\ \hline * & * \end{array}\right) = \mathrm{id}_V,$$

比较两边可见 T_{22} 有右逆而 T_{11} 有左逆. 鉴于推论 4.8.5, 这表明 T_{11} 和 T_{22} 皆可逆.

反之, 设 T_{11} 和 T_{22} 皆可逆. 若 $B \in \mathrm{Hom}(V_2, V_1)$, 则

$$\left(\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22} \end{array}\right)\left(\begin{array}{c|c} T_{11}^{-1} & B \\ \hline 0 & T_{22}^{-1} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{id}_{V_1} & T_{11}B + T_{12}T_{22}^{-1} \\ \hline 0 & \mathrm{id}_{V_2} \end{array}\right),$$

$$\left(\begin{array}{c|c} T_{11}^{-1} & B \\ \hline 0 & T_{22}^{-1} \end{array}\right)\left(\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{id}_{V_1} & T_{11}^{-1}T_{12} + BT_{22} \\ \hline 0 & \mathrm{id}_{V_2} \end{array}\right).$$

这表明 T 的逆可以取作

$$T^{-1}=\left(\begin{array}{c|c} T_{11}^{-1} & -T_{11}^{-1}T_{12}T_{22}^{-1} \\ \hline 0 & T_{22}^{-1} \end{array}\right).$$

最后, 关于 $T=\operatorname{diag}(T_{11},\dots,T_{nn})$ 的求逆公式已经包含于上述论证, 办法仍是化到 $n=2$ 情形. □

以上论证实际对分块上三角 (或下三角, 对角) 可逆映射给出了求 T^{-1} 的具体算法, 但是公式随着 n 增长而复杂化.

当直和分解来自 V 的有序基 $v_1,\dots,v_n$ 时 (参见注记 4.10.9), 对应的分块都是 1×1 矩阵, 亦即域 F 的元素. 此时定义 4.11.5 化作更单纯的形式.

定义 4.11.9 称矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}\in \mathrm{M}_{n\times n}(F)$ 是

- ⋆ **对角**的, 如果 $i\neq j \implies a_{ij}=0$;
- ⋆ **上三角**的, 如果 $i>j \implies a_{ij}=0$;
- ⋆ **下三角**的, 如果 $i<j \implies a_{ij}=0$.

显而易见, 转置将上三角矩阵和下三角矩阵互换. 矩阵 $\boldsymbol{A}$ 为上三角的当且仅当

$$\begin{aligned} \boldsymbol{A}\boldsymbol{e}_1 &\subset \langle \boldsymbol{e}_1\rangle, \\ \boldsymbol{A}\boldsymbol{e}_2 &\subset \langle \boldsymbol{e}_1,\boldsymbol{e}_2\rangle, \\ &\vdots \\ \boldsymbol{A}\boldsymbol{e}_{n-1} &\subset \langle \boldsymbol{e}_1,\dots,\boldsymbol{e}_{n-1}\rangle. \end{aligned}$$

4.12 商空间

我们在 §2.5 学到如何将一个集合 A 对某个等价关系 $\sim$ 取商, 得到商集 $A/\sim$ 连同商映射 $q: A\to A/\sim$. 本节关注向量空间框架下的相应问题: 对于 F-向量空间 V 上的哪一类等价关系, 能确保商集具有自然的 F-向量空间结构, 使得商映射成为线性映射? 商集上的向量空间结构又有怎样的刻画?

这类问题在实践中经常用于探讨线性映射的像空间. 以下便从这一线索切入, 以界定所需探讨的等价关系.

给定线性映射 $T: V\to W$, 任意 $v_1,v_2\in V$ 在 W 中的像相同当且仅当它们相差一个 $\ker T$ 的元素. 命 $U:=\ker(T)$, 这启发我们在 V 上引进一个等价关系

$$v_1\sim_U v_2 \iff v_1-v_2\in U.$$

换言之, $v_1\sim_U v_2$ 当且仅当 $Tv_1=Tv_2$.

上述等价关系 $\sim_U$ 对 V 的任何子空间 U 都能定义, 不必先有线性映射 T. 等价关系的条件逐一验证如下.

▷ **反身性** $v \sim_U v$, 这是因为 $v - v = 0 \in U$.

▷ **传递性** 若 $v_1 \sim_U v_2$ 而 $v_2 \sim_U v_3$, 则 $v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in U$, 从而 $v_1 \sim_U v_3$.

▷ **对称性** 若 $v_1 \sim_U v_2$, 则 $v_2 - v_1 = -(v_1 - v_2) \in U$, 从而 $v_2 \sim_U v_1$.

约定 4.12.1 对于任意 F-向量空间 V 及其子空间 U, 可以合理地将 $\sim_U$ 的等价类写作

$$v + U := \{v + u : u \in U\}$$

的形式, 其中 $v \in V$. 形如 $v + U$ 的子集称为 U 在 V 中的**陪集**, 而 v 称为该陪集的代表元. 我们有 $v + U = v' + U$ 当且仅当 $v - v' \in U$.

回到线性映射 $T : V \to W$ 和 $U = \ker(T)$ 的情形, 理应提出一个问题: 像空间 $\mathrm{im}(T)$ 是否可以仅从 V 及其子空间 U 来描述? 命题 2.5.8 在集合层次提供了一个答案: 存在从商集 $V/\sim_U$ 到 $\mathrm{im}(T)$ 的双射 $\overline{T}$, 它映 $v + U$ 为 Tv. 问题在于能否赋予 $V/\sim_U$ 自然的向量空间结构, 使得 $\overline{T}$ 是向量空间的同构? 此结构应当由 V 和 U 完全确定, 和外铄的资料 T 无关.

此外, 我们希望商映射 $q : V \to V/\sim_U$ 本身也是线性的, 如此则 T 拆解为两个线性映射的合成

$$V \xrightarrow{q} V/\sim_U \xrightarrow{\overline{T}} V.$$

这些考量引出了商空间的概念.

定义 4.12.2 (商空间) 设 U 是 F-向量空间 V 的子空间. 定义

$$\begin{aligned} V/U &:= \{U \text{ 的陪集 } v + U : v \in V\} \\ &= V/\sim_U . \end{aligned}$$

在 V/U 上可以对陪集合理地定义以下运算.

▷ **加法** $(v_1 + U) + (v_2 + U) := v_1 + v_2 + U$, 其中 $v_1, v_2 \in V$;

▷ **纯量乘法** $t(v + U) := tv + U$, 其中 $t \in F$ 而 $v \in V$;

▷ **零元** $0_{V/U} := U = 0_V + U$ (作为 U 的陪集).

由此得到向量空间 $(V/U, +, \cdot, 0_{V/U})$, 称为 V 对 U 的商空间.

关键在于说明这些定义仅依赖陪集, 不依赖于陪集中代表元的选取. 先看加法, 若将每个代表元 v_i 换为 $v_i' = v_i + u_i$, 其中 $u_i \in U$, 则

$$v_1' + v_2' + U = v_1 + v_2 + \overbrace{\underbrace{u_1 + u_2}_{\in U} + U}^{=U} = v_1 + v_2 + U.$$

接着考虑纯量乘法. 若将代表元 v 换为 $v' = v + u$, 其中 $u \in U$, 则由 $tu \in U$ 可知 $tv' + U = tv + tu + U = tv + U$.

其次, 我们必须对 $(V/U, +, \cdot, 0_{V/U})$ 验证向量空间的公理 (定义 4.2.1).

1. 首先是加法部分. 既然知道陪集的加法无关代表元的选取, 结合律遂归结为 V 中的结合律

$$\begin{aligned}((v_1 + U) + (v_2 + U)) + (v_3 + U) &= v_1 + v_2 + v_3 + U \\ &= (v_1 + U) + ((v_2 + U) + (v_3 + U)).\end{aligned}$$

   同样应用 V 的性质, 零元的性质归结为 $(v+U)+(0_V+U) = (v+0_V)+U = v+U$, 同理 $(0_V + U) + (v + U) = v + U$.

   类似地, V/U 的加法交换律归结为 V 的加法交换律.

2. 陪集 $v + U$ 的加法逆元是

$$-(v + U) := (-v) + U.$$

   验证方法相同: $(v + U) + (-v + U) = (v + (-v)) + U = U$.

3. 对于纯量乘法, 结合律, 幺元性质和对加法的分配律, 按同样手法化到 V 的相应性质.

对等价关系 $\sim_U$ 应用定义 2.5.4 给出商映射 $q : V \to V/U$, 它映 v 为 $v + U$. 按照上述讨论,

$$\begin{aligned}q(v_1 + v_2) &= (v_1 + v_2) + U = (v_1 + U) + (v_2 + U), \\ q(tv_1) &= tv_1 + U = t(v_1 + U),\end{aligned}$$

于是 q 是线性映射. 它当然是满的, 核为

$$\begin{aligned}\ker(q) &= \{v : v + U = U\} \\ &= \{v : v \sim_U 0_V\} = U.\end{aligned}$$

练习 4.12.3 说明 $V/\{0\}$ 可以等同于 V, 而 V/V 可以等同于零空间. 进一步, 设 U 为 V 的子空间, 则 V/U 是零空间当且仅当 $U = V$.

命题 4.12.4 设 V 是有限维向量空间, U 是其子空间, 则

$$\dim U + \dim(V/U) = \dim V.$$

证明 因为 $\mathrm{im}(q) \simeq V/U$ 而 $\ker(q) = U$, 这化为定理 4.8.4. □

下述构造既是商空间的简单例子, 也是向量空间理论中的一个重要概念. 它的地位将在 §8.9 得到进一步的说明.

定义 4.12.5 (余核) 线性映射 $T : V \to W$ 的余核定义为 W 对 $\mathrm{im}(T)$ 的商空间

$$\mathrm{coker}(T) := W/\mathrm{im}(T).$$

因此 T 的余核自然地和商映射 $q : W \twoheadrightarrow \mathrm{coker}(T)$ 相伴出现. 它虽然是通过 $\mathrm{im}(T)$ 来定义的, 但 $\mathrm{im}(T)$ 也可以反过来用 $\mathrm{coker}(T)$ 和它带有的商映射描述为

$$\mathrm{im}(T) = \ker\left[W \xrightarrow{\text{商映射}} \mathrm{coker}(T)\right]. \tag{4.12.1}$$

以下性质可谓是命题 4.8.3 的对偶版本.

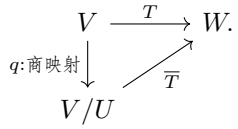
命题 4.12.6 线性映射 $T : V \to W$ 是满的当且仅当 $\mathrm{coker}(T) = \{0\}$.

证明 练习 4.12.3 的直接应用. □

对于线性映射 $T : V \to W$, 商空间 $V/\ker(T)$ 和像空间 $\mathrm{im}(T)$ 的关联可以用约定 2.3.3 介绍的交换图表来总结.

命题 4.12.7 设 U 是向量空间 V 的子空间, $T : V \to W$ 是线性映射.

(i) 若 $U \subset \ker(T)$, 则存在唯一的线性映射 $\overline{T} : V/U \to W$ 使得下图交换:



具体地说, $\overline{T}$ 映陪集 $v + U$ 为 Tv.

(ii) 若 $U = \ker(T)$ 而 $W = \mathrm{im}(T)$, 则 $\overline{T}$ 是向量空间的同构.

证明 既然 q 满, 使 (i) 的图表交换的唯一选择是命 $\overline{T}(v + U) = Tv$. 这确实是良定义的, 因为 $U \subset \ker(T)$ 蕴涵 Tv 只和陪集 $v + U$ 相关. 它同时也是线性映射, 这是因为

$$\begin{aligned}
\overline{T}((v_1 + U) + (v_2 + U)) &= \overline{T}(v_1 + v_2 + U) = T(v_1 + v_2) \\
&= Tv_1 + Tv_2 = \overline{T}(v_1 + U) + \overline{T}(v_2 + U), \\
\overline{T}(t(v + U)) &= \overline{T}(tv + U) = T(tv) = tT(v) \\
&= t\overline{T}(v + U).
\end{aligned}$$

若进一步要求 $U = \ker(T)$ 而 $W = \mathrm{im}(T)$, 则 $\overline{T}q = T$ 说明 $\mathrm{im}(\overline{T}) = W$. 此外

$$\overline{T}(v + U) = 0 \iff T(v) = 0 \iff v \in U \iff v + U = U,$$

因此 $\ker(\overline{T}) = \{0\}$, 从而 $\overline{T}$ 是同构. □

命题 4.12.7 (ii) 的同构 $\overline{T}: V/\ker(T) \xrightarrow{\sim} \mathrm{im}(T)$ 是以后常用的工具. 它说明线性满射和取商是一回事. 商空间 $V/\ker(T)$ 在一些场合也称为 T 的**余像**.

现在容易解决本节开头的问题. 设在向量空间 V 上有等价关系 $\sim$, 而商集 $V/\sim$ 具有向量空间结构, 使得商映射 $q: V \twoheadrightarrow V/\sim$ 是线性的. 记 $U := \ker(q)$, 则 $v_1 \sim v_2 \iff q(v_1) = q(v_2) \iff v_1 - v_2 \in U$ 导致 $\sim$ 便是先前定义的 $\sim_U$. 既然 q 满, 命题 4.12.7 给出同构 $\overline{q}: V/U \xrightarrow{\sim} V/\sim$, 使得 $\overline{q}(v+U) = q(v)$; 既然 $\sim = \sim_U$, 这般的 $\overline{q}$ 其实是恒等. 这表明我们寻求的等价关系 $\sim$ 和 $V/\sim$ 上的结构无非是对某个子空间取商.

线性映射自然地确定商空间之间的线性映射, 前提是映射必须和给定的子空间相容, 精确的表述如下.

推论 4.12.8 设 $T: V \to W$ 是线性映射, $U \subset V$ 和 $U' \subset W$ 是子空间, 而且 $T(U) \subset U'$, 则存在唯一的线性映射 $\overline{T}: V/U \to W/U'$ 使得下图交换:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q'} \\ V/U & \xrightarrow[\overline{T}]{} & W/U' \end{array}$$

图中的 q 和 q' 分别是从 V 和 W 到各自商空间的商映射. 具体地说, $\overline{T}$ 映 $v+U$ 为 $Tv+U'$.

证明 合成映射 $q'T: V \to W/U'$ 仍是线性的, 而且映 U 为 0, 所以命题 4.12.7 说明存在唯一的 $\overline{T}$ 使得 $\overline{T}q = q'T$, 它映 $v+U$ 为 $q'(T(v)) = Tv+U'$. □

这里的符号 $\overline{T}$ 和命题 4.12.7 (i) 并不冲突: 若取 $U' = \{0\}$, 则条件化为 $U \subset \ker(T)$, 从而命题 4.12.7 化为上述结果的一则特例.

映射 $\overline{T}$ 称为 $T \in \mathrm{Hom}(V,W)$ 所 "诱导" 的线性映射, 其构造具有以下性质.

- ⋆ 由图表或具体公式 $\overline{T}(v+U) = Tv+U'$ 立见 $t\overline{T} = \overline{tT}$, 其中 $t \in F$.
- ⋆ 对于一对线性映射 $T_1, T_2: V \to W$, 在 $T_1(U), T_2(U) \subset U'$ 的前提下
$$\overline{T_1+T_2} = \overline{T_1} + \overline{T_2}.$$
- ⋆ 若有线性映射 $V_1 \xrightarrow{T} V_2 \xrightarrow{S} V_3$ 和子空间 $U_i \subset V_i$ (其中 $i = 1,2,3$), 使得 $T(U_1) \subset U_2$ 而 $S(U_2) \subset U_3$, 则交换图表
$$\begin{array}{ccccc} V_1 & \xrightarrow{T} & V_2 & \xrightarrow{S} & V_3 \\ {\scriptstyle q_1}\downarrow & & \downarrow{\scriptstyle q_2} & & \downarrow{\scriptstyle q_3} \\ V_1/U_1 & \xrightarrow[\overline{T}]{} & V_2/U_2 & \xrightarrow[\overline{S}]{} & V_3/U_3 \end{array}$$
蕴涵 $q_3ST = \overline{S}q_2T = \overline{S}\,\overline{T}\,q_1$, 故推论 4.12.8 的刻画给出 $\overline{ST} = \overline{S}\,\overline{T}$. 这点自然也容易就 $\overline{S}$, $\overline{T}$ 和 $\overline{ST}$ 的具体公式来检验: 两边都映 v_1+U_1 为 STv_1+U_3.

接着来比较商空间 V/U 的子空间与 V 的子空间, 以及相应的商.

命题 4.12.9 设 U 是 V 的子空间. 记 $\overline{V} := V/U$, 而商映射仍记为 $q : V \to \overline{V}$. 我们有双射

$$\{W \subset V : \text{子空间},\ W \supset U\} \xleftrightarrow{1:1} \left\{\overline{W} \subset \overline{V} : \text{子空间}\right\}$$

$$W \longmapsto \overline{W} := q(W)$$

$$W := q^{-1}(\overline{W}) \longmapsfrom \overline{W}.$$

此双射具有下述性质.

- ⋆ 它是严格保序的: $\overline{W}_1 \supset \overline{W}_2$ 当且仅当 $W_1 \supset W_2$.
- ⋆ 若 W 对应到 $\overline{W}$, 则有自然同构

$$\begin{aligned} V/W &\xrightarrow{\sim} \overline{V}/\overline{W} \\ v + W &\mapsto q(v) + \overline{W}; \end{aligned}$$

将 $\overline{W} := q(W)$ 具体写作商空间 W/U, 则同构也可以表为 "分母相消" 的形式 $V/W \xrightarrow{\sim} (V/U)/(W/U)$.

证明 先验证双向箭头互逆. 给定包含 U 的子空间 W, 我们断言 $q^{-1}(q(W)) = W$. 首先 $\supset$ 是同义反复; 至于 $\subset$, 设 $v \in V$ 满足 $q(v) \in q(W)$, 则存在 $w \in W$ 使得 $q(v) = q(w)$, 亦即 $v \in w + \ker(q) = w + U$, 因而从 $U \subset W$ 可见 $v \in W$.

对于另一个方向, 给定 $\overline{V}$ 的子空间 $\overline{W}$, 我们断言 $q(q^{-1}(\overline{W})) = \overline{W}$. 一如先前情况, $\subset$ 是同义反复. 至于 $\supset$ 则缘于 q 的满性.

由于双向映射各自满足 $W_1 \supset W_2 \implies \overline{W}_1 \supset \overline{W}_2$ 和 $\overline{W}_1 \supset \overline{W}_2 \implies W_1 \supset W_2$, 这对互逆映射必然是严格保序的.

最后, 定义线性映射 $T : V \to \overline{V}/\overline{W}$ 为 q 和商映射 $\overline{V} \to \overline{V}/\overline{W}$ 的合成, 映 v 为 $q(v) + \overline{W}$, 它作为满射的合成依然满, 核则是 $W := q^{-1}(\overline{W})$. 因此 T 诱导同构 $\overline{T} : V/W \xrightarrow{\sim} \overline{V}/\overline{W}$, 映 $v + W$ 为 $q(v) + \overline{W}$. □

由于取商和线性满射是一回事, 上述结果对任意线性满射 $q : V \twoheadrightarrow \overline{V}$ 同样有效.

命题 4.12.10 设 V, W 为向量空间 $\mathcal{V}$ 的子空间, 则有同构

$$V/(V \cap W) \xrightarrow{\sim} (V + W)/W$$

$$v + (V \cap W) \longmapsto v + W.$$

证明 定义 $T : V \to (V + W)/W$ 为包含映射 $V \hookrightarrow V + W$ 与商映射 $V + W \to (V + W)/W$ 的合成, 这些映射都是线性的; 具体地说, $T(v) = v + W$. 因此 $T(v) = 0$ 当且仅当 $v \in V \cap W$. 另一方面, W 在 $V + W$ 中的所有陪集都形如 $v + w + W = v + W$, 其中 $v \in V$ 而 $w \in W$, 故 T 满.

对 T 应用命题 4.12.7 (ii), 便得到同构 $\overline{T}: V/(V\cap W)\xrightarrow{\sim}(V+W)/W$, 映法是 $\overline{T}(v+(V\cap W))=T(v)=v+W$. □

由此便容易阐明直和与商的联系.

推论 4.12.11 设 $V=U\oplus W$, 则相应的商映射 $q: V\to V/U$ 限制为同构 $q|_W: W\xrightarrow{\sim} V/U$.

证明 直接验证 $q|_W$ 为双射并不困难, 以下进路则基于命题 4.12.10. 将 U 和 W 视同 V 的子空间, 则 $V=U+W$ 而 $U\cap W=\{0\}$, 于是有同构 $W\simeq W/(U\cap W)\xrightarrow{\sim}(U+W)/U=V/U$, 它映 w 为 $w+U\in V/U$, 但这正是 $q|_W$ 的描述. □

注意: 尽管命题 4.10.6 确保对 V 的任何子空间 U 都有 W 使得 $V=U\oplus W$, 但取法并非典范的, 所以商空间并不能在概念上等同于直和项.

最后, 我们回顾推论 4.12.8 构造的诱导映射 $\overline{T}: V/U\to W/U'$. 以下设 V 和 W 都是有限维空间, 就矩阵观点考察 T 和 $\overline{T}$ 的关系. 取

$$\begin{aligned} u_1,\dots,u_k &: U \text{ 的有序基},\\ \overline{v}_1,\dots,\overline{v}_m &: V/U \text{ 的有序基},\\ u'_1,\dots,u'_{k'} &: U' \text{ 的有序基},\\ \overline{w}_1,\dots,\overline{w}_{m'} &: W/U' \text{ 的有序基}. \end{aligned}$$

再对所有 i,j 任取 $v_i\in q^{-1}(\overline{v}_i)$ 和 $w_j\in(q')^{-1}(\overline{w}_j)$. 已知 $u_1,\dots,u_k,v_1,\dots,v_m$ 构成 V 的有序基; 详见定理 4.8.4 证明. 同样地, $u'_1,\dots,u'_{k'},w_1,\dots,w_{m'}$ 也构成 W 的有序基.

命 $\boldsymbol{A}\in \mathrm{M}_{(k'+m')\times(k+m)}(F)$ 为 T 相对于这些有序基的矩阵. 条件 $T(U)\subset U'$ 表明存在一族系数 α_{ij} 使得 $Tu_j=\sum_{i=1}^{k'}\alpha_{ij}u'_i$. 记 $\boldsymbol{A}_U:=(\alpha_{ij})_{\substack{1\le i\le k'\\ 1\le j\le k}}$. 这是 $\boldsymbol{A}$ 的左上分块, 它同时又是 $T|_U: U\to U'$ 对应的矩阵. 此外, $\boldsymbol{A}$ 的左下 $m'\times k$ 分块则是零矩阵.

另一方面, 存在一族系数 β_{ij} 和 $\overline{\alpha}_{ij}$, 使得 $Tv_j=\sum_{i=1}^{k'}\beta_{ij}u'_i+\sum_{i=1}^{m'}\overline{\alpha}_{ij}w_i$. 基于 $\overline{T}q=q'T$, 我们立刻得到

$$\overline{T}(\overline{v}_j)=\sum_{i=1}^{m'}\overline{\alpha}_{ij}\overline{w}_i.$$

记 $\boldsymbol{A}_{V/U}=(\overline{\alpha}_{ij})_{\substack{1\le i\le m'\\ 1\le j\le m}}$. 这是 $\boldsymbol{A}$ 的右下分块, 同时也是 $\overline{T}$ 对应的矩阵. 这些观察即刻导向以下结论.

命题 4.12.12 在上述情境中, T 对应的矩阵 $\boldsymbol{A}$ 有分块表法

$$\boldsymbol{A}=\left(\begin{array}{c|c} \boldsymbol{A}_U & \star\\ \hline \boldsymbol{0}_{m'\times k} & \boldsymbol{A}_{V/U}\end{array}\right),$$

其中 $\boldsymbol{A}_U$ 是 $T|_U: U\to U'$ 对应的矩阵, $\boldsymbol{A}_{V/U}$ 是 $\overline{T}: V/U\to W/U'$ 对应的矩阵.

特别地, $\boldsymbol{A}_U$ (或 $\boldsymbol{A}_{V/U}$) 仅取决于 U 和 U' (或 V/U 和 W/U') 的有序基如何选取, 无关 v_i 和 w_j 的选法; 后者仅影响标为 $\star$ 的右上分块.

练习 4.12.13 推而广之, 设 V 和 W 分别带有长度为 d 的一列子空间

$$\{0\} = U_0 \subsetneq \cdots \subsetneq U_d = V, \quad \{0\} = U'_0 \subsetneq \cdots \subsetneq U'_d = W,$$

而且 $T : V \to W$ 对所有 $0 \leq h \leq d$ 皆满足 $T(U_h) \subset U'_h$. 对每个 U_h/U_{h-1} 和 U'_h/U'_{h-1} 选定有序基 $(1 \leq h \leq d)$. 说明如何将其提升为 V 的有序基, 使得 T 对应的矩阵 A 具有分块上三角的表法

$$\boldsymbol{A} = \left(\begin{array}{c|c|c|c} \boldsymbol{A}_1 & \star & \cdots & \star \\ \hline & \ddots & \cdots & \vdots \\ \hline & & \ddots & \star \\ \hline & & & \boldsymbol{A}_d \end{array}\right),$$

其中留白部分默认为零, 而矩阵 $\boldsymbol{A}_h$ 对应到 $T|_{U_h}$ 诱导的线性映射 $\overline{T}_h : U_h/U_{h-1} \to U'_h/U'_{h-1}$. 留意到 $\overline{T}_1$ 无非是 $T|_{U_1} : U_1 \to U'_1$.

数学家经常将一列子空间 $\{0\} = U_0 \subsetneq \cdots \subsetneq U_d = V$ 称为 V 的**滤过**. 若 T 在 $\forall h\ T(U_h) \subset U'_h$ 的意义下保持滤过, 则 $\overline{T}$ 对应到分块上三角矩阵, 其对角分块体现 T 在每一段所 "析出" 的线性映射 $\overline{T}_h$; 它们比原映射简单, 却保留许多有用信息. 这是练习 4.12.13 的实质.

习题

1. 设 V 是任意域 F 上的向量空间. 满足 $\{0\} \neq V' \subsetneq V$ 的子空间 V' 称为非平凡真子空间.

(i) 设 V_1, V_2 为非平凡真子空间, 证明 $V_1 \cup V_2 \neq V$.

(ii) 给出 $V = V_1 \cup V_2 \cup V_3$ 的例子, 其中每个 V_i 都是非平凡真子空间. 提示 取 F 为仅有两个元素的有限域 $\mathbb{F}_2$.

(iii) 证明当 $F = \mathbb{C}$ 时不存在非平凡真子空间 $V_1, \ldots, V_k$ 使得 $k \geq 2$ 而 $V = V_1 \cup \cdots \cup V_k$. 能否将此推及 F 为无穷域的情形?

2. 验证以下的矩阵等式

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1y_1 & \cdots & x_1y_n \\ \vdots & \ddots & \vdots \\ x_my_1 & \cdots & x_my_n \end{pmatrix},$$

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} \lambda_1a_{11} & \cdots & \lambda_1a_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_na_{n1} & \cdots & \lambda_na_{nm} \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_m \end{pmatrix} = \begin{pmatrix} \mu_1a_{11} & \cdots & \mu_ma_{1m} \\ \vdots & \ddots & \vdots \\ \mu_1a_{n1} & \cdots & \mu_ma_{nm} \end{pmatrix},$$

其中留白部分的矩阵元皆为 0.

3. 计算以下 $n \times n$ 矩阵的 n 次幂:

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

其中留白部分的矩阵元皆为 0.

4. 计算

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^n, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n$$

其中 $n \in \mathbb{Z}_{\geq 1}$.

5. 为了熟悉矩阵的计算, 请对所有 $t \in F^\times$ 验证等式

$$\begin{pmatrix} 0 & t \\ -t^{-1} & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -t^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} = \begin{pmatrix} 0 & t \\ -t^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

6. (Sherman–Morrison–Woodbury 公式) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 和 $\boldsymbol{C} \in \mathrm{M}_{m\times m}(F)$ 皆可逆, 而 $\boldsymbol{U} \in \mathrm{M}_{n\times m}(F)$, $\boldsymbol{V} \in \mathrm{M}_{m\times n}(F)$ 使得 $\boldsymbol{C}^{-1} - \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}$ 可逆. 证明 $\boldsymbol{A} - \boldsymbol{U}\boldsymbol{C}\boldsymbol{V}$ 可逆而且

$$(\boldsymbol{A} - \boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1} - \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}.$$

此式在一些优化算法中有所应用.

提示 先考虑 $\boldsymbol{A}=\mathbf{1}_{n\times n}$ 而 $\boldsymbol{C}=\mathbf{1}_{m\times m}$ 的特例, 欲证等式化为 $(\mathbf{1}_{n\times n}-\boldsymbol{U}\boldsymbol{V})^{-1}=\mathbf{1}_{n\times n}+\boldsymbol{U}(\mathbf{1}_{m\times m}-\boldsymbol{V}\boldsymbol{U})^{-1}\boldsymbol{V}$. 这点可以直接验证; 事实上, $n=m$ 的特例已包含于第三章的一则习题 (归于 N. Jacobson). 对于一般情形, 用 $\boldsymbol{A}^{-1}\boldsymbol{U}$ 和 $\boldsymbol{C}\boldsymbol{V}$ 替换上式的 $\boldsymbol{U}$ 和 $\boldsymbol{V}$.

7. (纯量限制) 设 F 为域 E 的子域.

(i) 设 V 为 E-向量空间, 说明若将纯量乘法运算 $E\times V\to V$ 限制到 $F\times V$ 上, 向量加法保持不变, 则 V 成为 F-向量空间. 若 $T:V_1\to V_2$ 是 E-向量空间之间的线性映射, 则 T 也是对应的 F-向量空间之间的线性映射.

(ii) 作为上述构造的特例, 说明 E 成为 F-向量空间, 使得向量的加法是域 E 中的加法, 而纯量乘法 $F\times E\to E$ 是域 E 中的乘法.

(iii) 若 V 是 $\mathbb{C}$-向量空间, 有基 $v_1,\dots,v_n$, 则它作为 $\mathbb{R}$-向量空间有基 $v_1,\mathrm{i}v_1,\dots,v_n,\mathrm{i}v_n$; 特别地, $\dim_{\mathbb{R}}V=2\dim_{\mathbb{C}}V$.

(iv) 作为上述构造的特例, 说明 $\mathbb{C}$ 是 2 维 $\mathbb{R}$-向量空间, 以 $1,\mathrm{i}$ 为基.

8. 证明 $\mathbb{R}$ 作为 $\mathbb{Q}$-向量空间是无穷维的.

9. 定义从 $\mathbb{C}$ 到 $\mathrm{M}_{2\times2}(\mathbb{R})$ 的单射 φ 如下

$$\varphi(x+\mathrm{i}y)=\begin{pmatrix}x & -y\\ y & x\end{pmatrix},\quad x,y\in\mathbb{R}.$$

验证 $\varphi(1)=\mathbf{1}_{2\times2}$, $\varphi(z_1+z_2)=\varphi(z_1)+\varphi(z_2)$ 和 $\varphi(z_1z_2)=\varphi(z_1)\varphi(z_2)$; 换言之 φ 是环同态. 这表明复数及其运算能由 2×2 实矩阵实现.

10. 确定所有使以下向量在 $\mathbb{R}^3$ 中线性无关的 $t\in\mathbb{R}$:

$$(t,1,0),\quad (1,t,1),\quad (0,1,t).$$

11. 在 $\mathbb{Q}^4$ 中, 判断以下的向量 $\boldsymbol{w}$ 能否表成 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ 的线性组合; 如果能, 写下具体的表达式.

(i) $\boldsymbol{v}_1=\begin{pmatrix}-1\\3\\0\\-5\end{pmatrix}$, $\boldsymbol{v}_2=\begin{pmatrix}2\\0\\7\\-3\end{pmatrix}$, $\boldsymbol{v}_3=\begin{pmatrix}-4\\1\\-2\\6\end{pmatrix}$, $\boldsymbol{w}=\begin{pmatrix}8\\3\\-1\\-25\end{pmatrix}$;

(ii) $\boldsymbol{v}_1=\begin{pmatrix}3\\-5\\2\\-4\end{pmatrix}$, $\boldsymbol{v}_2=\begin{pmatrix}-1\\7\\-3\\6\end{pmatrix}$, $\boldsymbol{v}_3=\begin{pmatrix}3\\11\\-5\\10\end{pmatrix}$, $\boldsymbol{w}=\begin{pmatrix}2\\-30\\13\\-26\end{pmatrix}$

12. 设 $v_1,\dots,v_n$ 是向量空间 V 中线性无关的元素. 考虑

$$\begin{array}{rcccccc} v_1' &=& a_{11}v_1 &+& a_{12}v_2 &+\cdots+& a_{1n}v_n,\\ v_2' &=& & & a_{22}v_2 &+\cdots+& a_{2n}v_n,\\ \vdots & & & & & & \vdots\\ v_n' &=& & & & & a_{nn}v_n, \end{array}$$

其中 $a_{ij}\in F$, 而且每个 a_{ii} 皆非零. 证明 $v_1',\dots,v_n'$ 也线性无关, 并且 $\langle v_1,\dots,v_n\rangle=\langle v_1',\dots,v_n'\rangle$.

13. 设 F 是域, $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$. 记 $Z(\boldsymbol{A}):=\{\boldsymbol{B}\in \mathrm{M}_{n\times n}(F):\boldsymbol{AB}=\boldsymbol{BA}\}$.

(i) 简要地说明 $Z(\boldsymbol{A})$ 是 $\mathrm{M}_{n\times n}(F)$ 的子空间, 同时也是子环.

(ii) 对以下特例给出 $Z(\boldsymbol{A})$ 的一组基

$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\0&1&0\\1&1&1\end{pmatrix}.$$

(iii) 证明 $Z(\boldsymbol{A})=\mathrm{M}_{n\times n}(F)$ 的充要条件是

$$\boldsymbol{A}=\lambda\cdot\mathbf{1}_{n\times n},\quad \lambda\in F.$$

14. 设 $\boldsymbol{A}=(a_{ij})_{i,j}\in \mathrm{M}_{m\times m}(F)$ 是对角元全为零的上三角 (或下三角) 矩阵. 证明 $\boldsymbol{A}^m=\mathbf{0}_{m\times m}$.

15. 设 X 和 Y 为向量空间 V 的两组基, 而 $y\in Y\smallsetminus X$. 证明存在 $x\in X\smallsetminus Y$ 使得

$$(Y\smallsetminus\{y\})\cup\{x\}\text{ 仍是 }V\text{ 的基.}$$

16. 对于下列的 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_5\in\mathbb{Q}^4$, 求出一个极大线性无关子集.

$$\boldsymbol{v}_1=\begin{pmatrix}1\\-1\\2\\4\end{pmatrix},\ \boldsymbol{v}_2=\begin{pmatrix}0\\3\\1\\2\end{pmatrix},\ \boldsymbol{v}_3=\begin{pmatrix}3\\0\\7\\14\end{pmatrix},\ \boldsymbol{v}_4=\begin{pmatrix}1\\-1\\2\\0\end{pmatrix},\ \boldsymbol{v}_5=\begin{pmatrix}2\\1\\5\\6\end{pmatrix}.$$

17. 设 F 为有 q 个元素的有限域. 请将以下问题的答案用 q 来表达.

(i) 确定有限集 $\left\{\begin{array}{l|l}\text{向量组 }(\boldsymbol{v}_1,\dots,\boldsymbol{v}_m)\\ \in \underbrace{F^n\times\cdots\times F^n}_{m\text{ 份}}\end{array}\right.$ 线性无关 $\left.\right\}$ 的元素个数, 其中 $1\le m\le n$.

提示 说明 $\boldsymbol{v}_1$ 有 q^n-1 种选法, 而 $\boldsymbol{v}_1$ 选定后 $\boldsymbol{v}_2$ 有 q^n-q 种选法, 依此类推.

(ii) 确定 F 上的 $n\times n$ 可逆矩阵的个数.

提示 对 (i) 取 $m=n$.

(iii) 确定 F^n 有几个 m 维子空间, 其中 $1 \leq m \leq n$.

提示〉 任何 m 维子空间都有有序基 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m \in F^n$, 而且有序基的个数由 (i) 确定. 另一种观点: 该子空间的有序基和 $m \times m$ 可逆矩阵一样多, 因为 $(\boldsymbol{w}_i)_{i=1}^m, (\boldsymbol{w}'_i)_{i=1}^m$ 确定相同的子空间当且仅当它们通过一个 $m \times m$ 可逆矩阵来转换, 此矩阵的取法是唯一的.

18. 记 $\mathbb{C}[X]$ 为复系数多项式对于加法和乘法构成的 $\mathbb{C}$-向量空间, X 代表多项式的变元. 给定 $a \in \mathbb{C}$. 定义 $T : \mathbb{C}[X] \to \mathbb{C}[X]$ 为映射 $(Tf)(X) = f(X + a)$, 其中 $f \in \mathbb{C}[X]$. 试判定 T 是不是线性映射.

19. 将取值在域 F 上的所有数列 $(a_n)_{n\geq 0}$ 作成一个 F-向量空间 $\mathcal{S}$, 方式是

$$(a_n)_n + (b_n)_n = (a_n + b_n)_n, \quad t(a_n)_n = (ta_n)_n.$$

给定 $k \in \mathbb{Z}_{\geq 1}$ 和 $c_1, \ldots, c_k \in F$. 证明递归方程

$$a_{n+k} = c_k a_{n+k-1} + \cdots + c_1 a_n, \quad n \in \mathbb{Z}_{\geq 0}$$

的所有解 $(a_n)_n$ 构成 $\mathcal{S}$ 的子空间. 说明这个子空间同构于 F^k.

提示〉 映 $(a_n)_{n\geq 0}$ 为 $(a_0, \ldots, a_{k-1}) \in F^k$.

20. 设向量空间 V 和 W 分别带有序基 $v_1, \ldots, v_n$ 和 $w_1, \ldots, w_m$. 线性变换 $T : V \to W$ 和 $m \times n$ 矩阵 $\boldsymbol{\mathcal{M}}(T)$ 的对应在一些教材中陈述为等式:

$$\begin{pmatrix} w_1 & \cdots & w_m \end{pmatrix} \boldsymbol{\mathcal{M}}(T) = \begin{pmatrix} Tv_1 & \cdots & Tv_n \end{pmatrix};$$

留意到 w_i 和 Tv_j 只是抽象的向量, 未必是列向量, 所以上式需要适当的解读. 试说明此式涵义, 以及它如何刻画矩阵 $\boldsymbol{\mathcal{M}}(T)$.

21. 说明注记 4.9.2 介绍的转换矩阵 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}$ 也可以依照上一题的思路刻画为

$$\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}} = \begin{pmatrix} v'_1 & \cdots & v'_n \end{pmatrix}.$$

22. 试证若 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 满足 $\boldsymbol{A}^k = \boldsymbol{0}_{n\times n}$, 其中 k 是正整数, 则 $\boldsymbol{1}_{n\times n} - \boldsymbol{A}$ 可逆; 试给出 $(\boldsymbol{1}_{n\times n} - \boldsymbol{A})^{-1}$ 的简单表达式.

提示〉 用等比级数公式.

23. 设 V 和 W 为 F-向量空间, $T : V \to W$ 为线性映射. 用定义–命题 4.4.10) 证明

(i) T 是单射当且仅当它有线性的左逆;

(ii) T 是满射当且仅当它有线性的右逆.

24. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 证明

(i) $\boldsymbol{A}$ 的秩 ≤ 1 当且仅当它能表成

$$\boldsymbol{A} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}$$

的形式, $x_i, y_i \in F$;

(ii) 若 $\boldsymbol{A}$ 的秩 ≤ 1, 则存在 $k \in F$ 使得 $\boldsymbol{A}^2 = k\boldsymbol{A}$.

25. 考虑实数域 $\mathbb{R}$ 上的 $n \times n$ 矩阵 $\boldsymbol{P} = (p_{ij})_{1\leq i,j\leq n}$. 如果 $\boldsymbol{P}$ 满足

$$p_{ij} \geq 0, \quad \sum_{i=1}^{n} p_{ij} = 1,$$

则称之为 Markov 矩阵. 如果 $\mathbb{R}$ 上的 n 维列向量 $\boldsymbol{x} = (x_i)_{i=1}^n$ 满足 $x_i \geq 0$ 和 $\sum_{i=1}^n x_i = 1$, 则称 $\boldsymbol{x}$ 为概率向量.

(i) 证明 $\boldsymbol{P}$ 是 Markov 矩阵当且仅当

$$\boldsymbol{x} \text{ 是概率向量} \implies \boldsymbol{P}\boldsymbol{x} \text{ 是概率向量}.$$

(ii) 证明如果 $\boldsymbol{x}$ 是概率向量, $\boldsymbol{P}$ 是 Markov 矩阵, 而且对所有 i, j 都有 $p_{ij} > 0$, 则 $\boldsymbol{P}\boldsymbol{x}$ 是各个坐标皆 > 0 的概率向量.

(iii) 证明 Markov 矩阵的乘积仍然是 Markov 矩阵.

(iv) 设 $\boldsymbol{P}$ 是 Markov 矩阵. 证明存在列向量 $\boldsymbol{x} \neq \boldsymbol{0}_{n\times 1}$ 使得 $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{x}$. 提示〉 等价于证明 $\boldsymbol{P} - \boldsymbol{1}_{n\times n}$ 不可逆.

26. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别是域 F 上的 $m\times n$ 和 $n\times \ell$ 矩阵. 证明 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{m\times \ell}$ 蕴涵 $\mathrm{rk}(\boldsymbol{A})+\mathrm{rk}(\boldsymbol{B}) \leq n$. 提示〉 一种方法是将矩阵视同线性映射, 从而 $\mathrm{im}(\boldsymbol{B}) \subset \ker(\boldsymbol{A})$.

27. 设 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$, $\boldsymbol{B} \in \mathrm{M}_{m\times \ell}(F)$, 证明矩阵方程

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}, \quad \boldsymbol{X} \in \mathrm{M}_{n\times \ell}(F)$$

有解 $\boldsymbol{X}$ 的充要条件是分块矩阵 $(\boldsymbol{A}|\boldsymbol{B}) \in \mathrm{M}_{m\times(n+\ell)}(F)$ 满足 $\mathrm{rk}(\boldsymbol{A}|\boldsymbol{B}) = \mathrm{rk}(\boldsymbol{A})$. 提示〉 关于秩的条件等价于说 $\boldsymbol{B}$ 的每列都能写成 $\boldsymbol{A}$ 的列的线性组合.

28. 回忆到任意多项式 $f \in F[X]$ 总能代入矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 来求值 $f(\boldsymbol{A}) \in \mathrm{M}_{n\times n}(F)$. 说明若 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$ 可逆, 则 $f(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}) = \boldsymbol{P}^{-1}f(\boldsymbol{A})\boldsymbol{P}$.

29. 选定 $n_1, \ldots, n_r \in \mathbb{Z}_{\geq 1}$, 命 $n = n_1 + \cdots + n_r$. 以下矩阵的留白部分默认为零.

(i) 考虑 $n \times n$ 分块对角矩阵

$$\boldsymbol{A} = \begin{pmatrix} \boxed{\lambda_1 \boldsymbol{1}_{n_1\times n_1}} & & \\ & \ddots & \\ & & \boxed{\lambda_r \boldsymbol{1}_{n_r\times n_r}} \end{pmatrix}$$

其中 $\lambda_1, \ldots, \lambda_r \in F$ 两两相异. 试确定所有满足

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$$

的 $n \times n$ 矩阵 $\boldsymbol{B}$.

提示〉 答案是 $\boldsymbol{B}$ 为分块对角矩阵, 对角分块的尺寸依序是 $n_1 \times n_1, \ldots, n_r \times n_r$.

(ii) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 以类似的计算说明 $\boldsymbol{AB} = \boldsymbol{BA}$ 对于所有规格如下的分块对角矩阵

$$\boldsymbol{B} = \begin{pmatrix} \boxed{\boldsymbol{B}_1} & & \\ & \ddots & \\ & & \boxed{\boldsymbol{B}_r} \end{pmatrix}, \quad B_i \in \mathrm{M}_{n_i\times n_i}(F)$$

都成立的充要条件是存在 $\lambda_1, \dots, \lambda_r \in F$ 使得

$$\boldsymbol{A} = \begin{pmatrix} \boxed{\lambda_1 \mathbf{1}_{n_1\times n_1}} & & \\ & \ddots & \\ & & \boxed{\lambda_r \mathbf{1}_{n_r\times n_r}} \end{pmatrix}.$$

30. 将命题 4.12.4 扩及一般的向量空间, 前提是将维数的相加诠释为基数加法, 见 §2.9.

31. 证明若 $U^\flat, U, V$ 是某向量空间 W 的子空间, $U^\flat \subset U$, 则

$$U^\flat + (V \cap U) = (U^\flat + V) \cap U.$$

32. 证明对任意有限维 F-向量空间 V, 线性映射 $T : V \to V$ 和 $n \in \mathbb{Z}_{\geq 1}$ 都有

$$\dim\left(\mathrm{im}(T^{n-1}) \cap \ker(T)\right) = \dim\ker(T^n) - \dim\ker(T^{n-1}).$$

33. 设 V 是有限维 F-向量空间, $T \in \mathrm{End}(V)$. 证明 $\dim\ker(T) = \dim\mathrm{coker}(T)$. 注意到对于数学分析中出现的一些无穷维空间, 即使 $\ker(T)$ 和 $\mathrm{coker}(T)$ 皆有限维, 两者也未必相等; 这牵涉到称为指标的一类重要不变量.

34. (链复形) 考虑 F-向量空间之间的一列线性映射

$$\cdots \to V_{n+1} \xrightarrow{T_{n+1}} V_n \xrightarrow{T_n} V_{n-1} \xrightarrow{T_{n-1}} V_{n-2} \to \cdots$$

下标 n 既可以遍历整数集, 也容许单边或双边有界. 若 $T_{n-1}T_n = 0$ (换言之, $\mathrm{im}(T_n) \subset \ker(T_{n-1})$) 对所有 n 皆成立, 则称这样一列线性映射为**链复形**, 而将空间 V_n 称为其 n 次项; 线性映射 $(T_n)_n$ 经常在符号中省略. 链复形的第 n 个**同调**定义为商空间

$$\mathrm{H}_n := \ker(T_{n-1}) \big/ \mathrm{im}(T_n).$$

若 H_n 对所有 n 都是零空间, 则称此链复形**正合**, 或者称此列线性映射为**正合列**. 若链复形的 $> n$ (或 $< n$) 次项全为零空间, 则我们也将它简记为 $0 \to V_n \to V_{n-1} \to \cdots$ (或 $\cdots \to V_{n+1} \to V_n \to 0$) 的形式.

(i) 说明 $0 \to V_1 \xrightarrow{T_1} V_0$ 正合当且仅当 T_1 单.

(ii) 说明 $V_1 \xrightarrow{T_1} V_0 \to 0$ 正合当且仅当 T_1 满.

(iii) 说明 $0 \to V_1 \xrightarrow{T_1} V_0 \to 0$ 正合当且仅当 T_1 为同构.

(iv) 说明若 $0 \to V_1 \to V_2 \to V_3 \to 0$ 是正合列, 则 V_1 可视同 V_2 的子空间, 而 V_3 可视同商空间 V_2/V_1. 明确 "视同" 的意义.

提示〉 要点在于明确所用的同构, 它们必须在某种精确意义下为 "自然" 的, 而不是任选的. 应用命题 4.12.7 (ii).

(v) (Euler–Poincaré 原理) 考虑形如

$$0 \to V_n \to \cdots \to V_0 \to 0$$

的链复形. 观察到当 $k < 0$ 或 $k > n$ 时 H_k 为零. 现在假定每个 V_k 都是有限维向量空间, 证明

$$\sum_{k=0}^{n} (-1)^k \dim \mathrm{H}_k = \sum_{k=0}^{n} (-1)^k \dim V_k.$$

(vi) 将命题 4.12.4 诠释为 Euler–Poincaré 原理的一则特例.

35. (线性码) 通信领域的一大主题是编码. 给定一个有限集 Σ (字母), 我们将信息理解为 Σ^k 的元素 (字); 通信的基本机制呈现为一对映射, 其中 $k, n \in \mathbb{Z}_{\geq 1}$:

$$\Sigma^k \xrightarrow[\text{编码}]{E} \Sigma^n \xrightarrow[\text{解码}]{D} \Sigma^k;$$

编码与解码分别由发送端和接收端执行, 编码后的信息 $\boldsymbol{z} \in \Sigma^n$ 由信道传递. 基本的要求是 $DE = \mathrm{id}$. 然而实际的信道总有衰减或杂讯, 接收端收到的 $\boldsymbol{z}$ 未必等于发出的 $E(\boldsymbol{a})$. 若需要一定的纠错能力, 必须要求 $\boldsymbol{a} \neq \boldsymbol{b}$ 时 $E(\boldsymbol{a})$ 和 $E(\boldsymbol{b})$ 应该有较大的区隔. 为了量化, 设 Σ 为任意集合, 对 Σ^n 的元素 $\boldsymbol{a} = (a_i)_i$ 和 $\boldsymbol{b} = (b_i)_i$ 定义其间的 Hamming 距离为

$$d(\boldsymbol{a}, \boldsymbol{b}) := |\{1 \leq i \leq n : a_i \neq b_i\}|;$$

它满足 $d(\boldsymbol{a}, \boldsymbol{c}) \leq d(\boldsymbol{a}, \boldsymbol{b}) + d(\boldsymbol{b}, \boldsymbol{c})$ 和 $d(\boldsymbol{a}, \boldsymbol{b}) = 0 \iff \boldsymbol{a} = \boldsymbol{b}$. 设 $c \in \mathbb{R}_{\geq 0}$. 如果对所有 $\boldsymbol{a} \in \Sigma^k$ 和 $\boldsymbol{z} \in \Sigma^n$ 皆有

$$d(\boldsymbol{z}, E(\boldsymbol{a})) \leq c \implies D(\boldsymbol{z}) = \boldsymbol{a},$$

则称此码能纠正至多 c 位错误. 实践中我们希望让 k/n 和 c/n 尽可能大, 而映射 E 与 D 应当能高效地计算. 这是编码学的核心主题.

(i) 命 $C := \mathrm{im}(E)$. 定义 $d(C) := \min\{d(\boldsymbol{z}, \boldsymbol{w}) : \boldsymbol{z}, \boldsymbol{w} \in C,\ \boldsymbol{z} \neq \boldsymbol{w}\}$. 说明能够适当地定义解码映射 D, 使得此码能纠正至多 $\frac{d(C)-1}{2}$ 位错误.

提示〉 对任意 $\boldsymbol{z}$, 存在至多一个 $\boldsymbol{w} \in C$ 使得 $d(\boldsymbol{w}, \boldsymbol{z}) < \frac{d(C)}{2}$.

(ii) 常用的一类编码方式为线性码: 取 F 为有限域, $E : F^k \to F^n$ 为线性单射, 此时 C 是 F^n 的子空间. 说明存在 $\boldsymbol{H} \in \mathrm{M}_{(n-k)\times n}(F)$, 称为校验矩阵, 使得 $\boldsymbol{H}$ 满秩而 $C = \{\boldsymbol{z} \in F^n : \boldsymbol{H}\boldsymbol{z} = \boldsymbol{0}\}$.

(iii) 承上, 证明 $d(C) = \min\limits_{\substack{\boldsymbol{z} \in C \\ \boldsymbol{z} \neq \boldsymbol{0}}} d(\boldsymbol{0}, \boldsymbol{z})$.

(iv) 承上, 记 $\boldsymbol{H}$ 的第 i 列为 $\boldsymbol{h}_i$. 证明

$$d(C)=\min\left\{\ 1\le t\le n-k\ \middle|\ \begin{array}{l}\exists 1\le i_1<\cdots<i_t\le n\\ \boldsymbol{h}_{i_1},\dots,\boldsymbol{h}_{i_t}\text{ 线性相关}\end{array}\ \right\}.$$

提示〉 设 $\boldsymbol{z}$ 的各分量为 $z_1,\dots,z_n$, 则 $\boldsymbol{H}\boldsymbol{z}=\boldsymbol{0}$ 等价于 $\sum_{i=1}^n z_i\boldsymbol{h}_i=\boldsymbol{0}$, 而 $d(\boldsymbol{0},\boldsymbol{z})=|\{i:z_i\neq 0\}|$.

36. (Hamming 码) 承接上一题的讨论. 取仅有两个元素的有限域 $\mathbb{F}_2$. 对所有 $k\in\mathbb{Z}_{\ge 1}$ 定义矩阵 $\boldsymbol{H}_k\in \mathrm{M}_{k\times(2^k-1)}(\mathbb{F}_2)$, 使得它的第 i 列是 i 的二进制表法, 例如

$$\boldsymbol{H}_3=\begin{pmatrix}0&0&0&1&1&1&1\\0&1&1&0&0&1&1\\1&0&1&0&1&0&1\end{pmatrix}.$$

(i) 证明 $\mathrm{rk}(\boldsymbol{H}_k)=k$.

(ii) 命 $C:=\left\{\boldsymbol{z}\in\mathbb{F}_2^{2^k-1}:\boldsymbol{H}_k\boldsymbol{z}=\boldsymbol{0}\right\}$. 证明 $d(C)=3$.

(iii) 勾勒以 $\boldsymbol{H}_k$ 为校验矩阵的线性码如何操作.

37. (纠删码) 与纠错码密切相关的概念是纠删码. 给定编码映射 $E:\Sigma^k\to\Sigma^n$. 设 $c\in\mathbb{R}$. 如果存在一个算法, 使得当 $\boldsymbol{a}\in\Sigma^k$ 以及一列正整数 $1\le i_1<\cdots\le i_d\le n$ 给定, 其中 $d\le c$, 记从 $E(\boldsymbol{a})$ 删除第 $i_1,\dots,i_d$ 位得到的信息记为 $E(\boldsymbol{a})'\in\Sigma^{n-d}$, 仍然能从 $E(\boldsymbol{a})'$ 反解 $\boldsymbol{a}$, 则称此码能纠正至多 c 位删除. 纠删码广泛应用于数据存储领域.

(i) 给定编码映射 E, 按先前习题的方法定义 $C:=\mathrm{im}(E)$ 和 $d(C)$. 证明此码能纠正至多 $d(C)-1$ 位删除.

(ii) 纠删和纠错这两个概念有何区别? 提示〉 纠删码论及的位 $i_1<\cdots<i_d$ 给定; 纠错码中的出错位不明.

(iii) 最简单也最古老的纠删码是多复本储存 $E:\Sigma^k\to\Sigma^{hk}$, 其中 h 是复本个数, 其优点是运算量少, 缺点是 h 较大时浪费空间. 试分析它能纠正至多几位的删除.

第五章 行列式

泛泛而论, n 阶行列式是由 $n \times n$ 矩阵确定的一个特定的量. 在探讨行列式具体是什么之前, 不妨先解释背后动机. 按照历史的观点, 行列式的根源和向量空间一样可以追溯到两条线索.

1. 代数方面, 在给定的域 F 上, 人们希望判定线性方程组

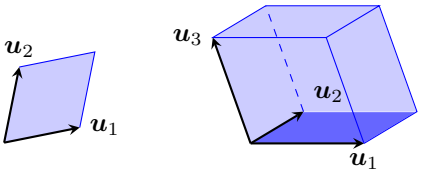
$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \vdots \qquad & \quad \vdots \\ a_{n1}X_1 + \cdots + a_{nn}X_n &= b_n. \end{aligned}$$

   是否有唯一解, 并寻求精确的公式解. Cramer 法则 (推论 5.7.6) 说明方程组有唯一解当且仅当系数矩阵 $(a_{ij})_{1\le i,j\le n}$ 的行列式非零, 并且给出了基于行列式的求解公式.

2. 几何方面, 人们关心 $\mathbb{R}^n$ 中 n 个向量 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ 张成的子集

$$\diamondsuit(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) := \left\{ \sum_{i=1}^{n} t_i \boldsymbol{u}_i : \forall i,\ t_i \in [0,1] \right\}$$

   的体积; 在 $n = 2$ (或 $n = 3$) 时, 此即平面上的平行四边形 (或空间中的平行六面体), 体积有直观意涵, 如下图:



   而高维情形则需要抽象地处理. 一旦将每个 $\boldsymbol{u}_i$ 按照直角坐标系展开为 $(a_{1i}, \ldots, a_{ni})$, 则所求的体积终归能以 $(a_{ij})_{1\le i,j\le n}$ 的行列式表达.

事实上, 和行列式直接相关的并非 $\diamondsuit(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ 的体积, 而是其 "有向体积", 这是将向量 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ 的方向和次序记入考量的一个量, 容许为负数. 详细讨论见诸 §5.2 和例 5.4.2.

无论在 Cramer 法则还是体积公式中, 所谓行列式都是关于矩阵 $\boldsymbol{A} = (a_{ij})_{1\le i,j\le n}$ 的 n 个列向量的一类特殊映射, 其值记为

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad \text{或} \quad \det \boldsymbol{A}.$$

依循几何的线索, §5.3 将从我们所寻求的行列式或有向体积映射 $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \mapsto \det(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)$ 抽象出一些必要性质, 具备这些性质的映射称为 n 维向量空间上的 n 元交错形式, 可以在任何域 F 上加以考量. 我们将证明非零的 n 元交错形式确实存在, 而且两两成比例; 论证的过程将自然地引出行列式的具体公式

$$\begin{aligned} \det \boldsymbol{A} &= \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}; \end{aligned}$$

由此便能严格地定义并刻画矩阵的行列式 $\det \boldsymbol{A}$. 上式中的置换 σ, 全体置换集 $\mathfrak{S}_n$ 和置换的符号 $\operatorname{sgn}(\sigma) \in \{\pm 1\}$ 在关于 n 元交错形式的脉络中既是自然的, 也是必要的, §5.1 为此提供了必要的准备.

矩阵是线性映射在取基之后的化身. 依照现代观点, 更应当关注的是线性映射 $T : V \to V$ 的行列式, 其中 V 是 n 维向量空间; 它不再是 V 上某个特定的 n 元交错形式, 而是所有这些交错形式在 T 的作用下确定的一个比例常数 $\det T \in F$; 如果用之前的几何观点来启发, 取 $V = \mathbb{R}^n$, 这等价于说 $\det T$ 是 T 对有向体积的伸缩比例:

$$\diamondsuit(T\boldsymbol{u}_1, \dots, T\boldsymbol{u}_n) \text{ 的有向体积 } = (\det T) \cdot \big(\diamondsuit(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) \text{ 的有向体积}\big);$$

对于了解微积分的读者, 应当注意到这正是行列式在重积分换元公式中的角色.

从比例的观点出发, 很容易推导行列式的乘性:

$$\det(\mathrm{id}_V) = 1, \quad \det(ST) = \det S \det T, \quad S, T : \text{线性映射 } V \to V.$$

关于线性映射的行列式的完整定义和一般性质将在 §5.4 和盘托出. 其后的 §5.5–5.6 则有进一步的算例和方法.

尽管基于矩阵的观点与算法是计算行列式的重要手段, 但只有当 V 的基选定, $\det T$ 才能视同交错形式 $V^n \to F$, 具体体现为相应矩阵的行列式 $\det \boldsymbol{A}$; 如果不能清楚地区隔作为比例常数的 $\det T$ 和 V 上的 n 元交错形式, 将导致对行列式有错误的理解.

在 §5.7, 我们主要介绍行列式对线性方程组的应用, 即先前提及的 Cramer 法则; 对于后续章节, 更重要的是此节涉及的行列式与可逆性的关系 (命题 5.7.1), 以及经典伴随矩阵 $\boldsymbol{A}^\vee$ (定义 5.7.3).

在行列式的基础上, §5.8 将对 n 维向量空间 V 上的线性映射 $T : V \to V$ (或相应矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$) 定义特征多项式 Char_T (或 $\mathrm{Char}_{\boldsymbol{A}}$), 展开后写作

$$X^n + c_{n-1}X^{n-1} + \cdots + c_0, \quad c_0, \ldots, c_{n-1} \in F.$$

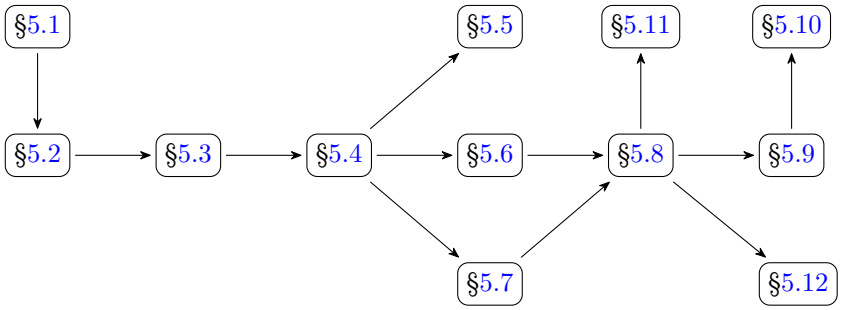
该节的核心结论是 Cayley–Hamilton 定理 5.8.8, 它断言将 T (或 $\boldsymbol{A}$) 代入其特征多项式给出零映射 (或零矩阵): $T^n + c_{n-1}T^{n-1} + \cdots + c_0 = 0_V$. 本章取道 $\boldsymbol{A}^\vee$ 的性质, 直接证明定理的矩阵版本. 后续章节还会给出各种应用和不同的证法.

特征多项式的各项系数 $c_0, \ldots, c_{n-1}$ 蕴藏 T 或 $\boldsymbol{A}$ 的重要信息, 譬如 $c_0 = (-1)^n \det T$, 而 §5.9 探讨的量是 $\mathrm{Tr}(T) = -c_{n-1}$, 称为 T 的迹, 它的矩阵版本比行列式简单得多: $\mathrm{Tr}(\boldsymbol{A}) = \sum_{i=1}^n a_{ii}$.

在 §5.10, 我们引入不变子空间的概念, 配合商空间的构造来诠释特征多项式的分块算法; 不变子空间也是尔后将反复出现的概念. 在 §5.11, 我们探讨如何对 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 和 $\boldsymbol{B} \in \mathrm{M}_{n\times m}(F)$ 写下 $\det(\boldsymbol{AB})$ 的公式, 容许 $m \neq n$, 相应的结果称为 Cauchy–Binet 定理 5.11.4; 本章习题将介绍此公式在图论中的应用 (矩阵-树定理).

最后的 §5.12 将对一般的交换环 R 及其上的矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 探讨行列式. 假若照搬线性映射的抽象方法, 则必然涉及称为模的代数结构 (第十二章); 为了避免离题, 该节的进路是直接以置换写下 $\det \boldsymbol{A}$ 的定义. 在证明 R 上行列式的性质时, 我们运用将矩阵元 "泛化" 的技巧, 化问题到 R 为域, 甚至是特征零的域的已知情形; 这种技术以后还会反复使用.

阅读顺序



5.1 置换概论

集合 X 上的**置换**是指从 X 到其自身的双射. 我们在关于矩阵初等行 (或列) 变换的研究中已经见过某些置换, 具体地说是矩阵任两行 (或列) 的互换. 有限集上的置换理论是研究行列式的必要准备, 而置换本身也是饶富兴味且用途广泛的数学对象.

定义 5.1.1 设 X 为非空集, 其上的置换构成集合

$$\mathfrak{S}_X := \{\text{双射 } \sigma : X \to X\}.$$

它包含恒等映射 $\mathrm{id} = \mathrm{id}_X \in \mathfrak{S}_X$. 这些置换可以作为映射来作合成 $(\sigma, \sigma') \mapsto \sigma\sigma'$ 或取逆 $\sigma \mapsto \sigma^{-1}$, 产物依然属于 $\mathfrak{S}_X$.

若 $n \in \mathbb{Z}_{\geq 1}$ 而 $X = \{1, \ldots, n\}$, 此时也记 $\mathfrak{S}_n := \mathfrak{S}_{\{1,\ldots,n\}}$.

设 $\sigma, \tau \in \mathfrak{S}_X$, 则 $\sigma\tau\sigma^{-1} \in \mathfrak{S}_X$ 称为 τ 对 σ 的**共轭**, 其映法是

$$\sigma(x) \xmapsto{\sigma^{-1}} x \xmapsto{\tau} \tau(x) \xmapsto{\sigma} \sigma(\tau(x)). \tag{5.1.1}$$

若 Y 是 X 的非空子集, 记 $\mathfrak{S}_{X,Y} \subset \mathfrak{S}_X$ 为所有在 Y 之外不动, 亦即满足 $\forall x \in X \smallsetminus Y$, $\sigma(x) = x$ 的置换 $\sigma \in \mathfrak{S}_X$ 所成的子集, 于是自然地得到双射

$$\begin{gathered} b : \mathfrak{S}_Y \xrightarrow{1:1} \mathfrak{S}_{X,Y} \subset \mathfrak{S}_X, \\ b(\sigma)|_Y = \sigma, \quad b(\sigma)|_{X\smallsetminus Y} = \mathrm{id}_{X\smallsetminus Y}. \end{gathered} \tag{5.1.2}$$

易见此双射保持乘法和取逆运算:

$$b(\mathrm{id}_Y) = \mathrm{id}_X, \quad b(\sigma\sigma') = b(\sigma)b(\sigma'), \quad b(\sigma^{-1}) = b(\sigma)^{-1}.$$

基于 (5.1.2), 我们从 $\{1\} \subset \{1, 2\} \subset \cdots$ 得到一列保持乘法的嵌入

$$\{\mathrm{id}\} = \mathfrak{S}_1 \hookrightarrow \mathfrak{S}_2 \hookrightarrow \mathfrak{S}_3 \hookrightarrow \cdots.$$

若非空有限集 X 有 n 个元素, 则 $\mathfrak{S}_X$ 和 $\mathfrak{S}_n$ 的研究本质上是一回事. 简单的计数表明 $\mathfrak{S}_n$ 有 $n!$ 个元素. 置换 $\sigma \in \mathfrak{S}_n$ 可以设想为 $1, \ldots, n$ 的排列 $\sigma(1), \ldots, \sigma(n)$, 或者更明确地表达成两行的矩阵形式

$$\sigma = \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}; \tag{5.1.3}$$

譬如 $1, \ldots, n$ 上的**轮换**便可以按此记为

$$\lambda = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}, \quad \lambda^{-1} = \begin{pmatrix} 1 & 2 & \cdots & n \\ n & 1 & \cdots & n-1 \end{pmatrix}.$$

定义 5.1.2 选定 $n \in \mathbb{Z}_{\geq 1}$. 设 $1 \leq i \neq j \leq n$, 相应的**对换** $(i\ j) \in \mathfrak{S}_n$ 定义为以下置换:

$$(i\ j) : k \mapsto \begin{cases} j, & k = i, \\ i, & k = j, \\ k, & k \neq i, j; \end{cases}$$

换言之, $(i\ j)$ 对调 i 和 j, 保持其余元素不动.

对换 $(i\ j)$ 按两行形式表为

$$\begin{pmatrix} \cdots & i & \cdots & j & \cdots \\ \cdots & j & \cdots & i & \cdots \end{pmatrix}.$$

显然 $(i\ j)^2 = \mathrm{id}$. 它们重要性在于所有置换都能写成一列对换的乘积. 事实上, 为了表达所有置换, 需要的仅是如下的**单对换**

$$s_i := (i \quad i+1) \in \mathfrak{S}_n, \quad 1 \le i \le n-1.$$

这是稍后的命题 5.1.8 的内容.

注记 5.1.3 对换两两共轭: 给定 $i \neq j$ 和 $i' \neq j'$, 命 $\tau := (i\ j)$. 任取 $\sigma \in \mathfrak{S}_n$ 使得 $\sigma(i) = i'$ 且 $\sigma(j) = j'$, 则 (5.1.1) 表明 $\tau' := \sigma\tau\sigma^{-1}$ 满足

$$\tau'(i') = j', \quad \tau'(j') = i', \quad \text{其余元素不动},$$

换言之, $\sigma(i\ j)\sigma^{-1} = (i'\ j')$.

练习 5.1.4 验证以下关于单对换的等式.

$$\begin{gathered} s_i^2 = \mathrm{id}, \\ s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}, \\ |i-j| > 1 \implies s_i s_j = s_j s_i. \end{gathered}$$

为了更深入地了解置换如何分解为单对换, 我们需要以下概念.

定义 5.1.5 设 $\sigma \in \mathfrak{S}_n$, 以下集合的元素称为 σ 的**逆序**:

$$\mathrm{Inv}_\sigma := \left\{(i,j) \in \mathbb{Z}^2 : 1 \le i < j \le n,\ \sigma(i) > \sigma(j)\right\}.$$

定义 σ 的**逆序数**为 $\ell(\sigma) := |\mathrm{Inv}_\sigma|$.

引理 5.1.6 我们有 $\sigma = \mathrm{id}$ 当且仅当 $\ell(\sigma) = 0$.

证明 条件 $\ell(\sigma) = 0$ 相当于说 $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ 是严格递增映射, 但这样的映射必然是恒等. □

练习 5.1.7 证明逆序数 $\ell(\sigma)$ 还有下述性质.

(i) $\ell(\sigma) = \ell(\sigma^{-1})$;

提示〉 写下 Inv_σ 和 $\mathrm{Inv}_{\sigma^{-1}}$ 之间的双射.

(ii) 若 s 是单对换, 则 $\ell(\sigma s) - \ell(\sigma) = \pm 1$, $\ell(s\sigma) - \ell(\sigma) = \pm 1$;

提示〉 设 $s = s_i$. 对于 $\ell(\sigma s)$ 和 $\ell(\sigma)$ 的比较, 可能增加或减少的逆序只有 $(i, i+1)$.

命题 5.1.8 设 $\sigma \in \mathfrak{S}_n$, 则存在 $\ell \in \mathbb{Z}_{\geq 0}$ 和一族单对换 $\tau_1, \dots, \tau_\ell \in \mathfrak{S}_n$ 使得

$$\sigma = \tau_1 \cdots \tau_\ell;$$

当 $\ell = 0$ 时, 右式的乘积理解为 id. 我们称 ℓ 为上述分解的长度. 在 σ 的所有单对换分解中, 最短可能的长度为 $\ell(\sigma)$.

证明 将置换 σ 看作 $1, \dots, n$ 的重排, 则证明思路类似于算法理论中最简单的 "冒泡排序" 法. 第一步是说明所有 σ 都有长度 $\leq \ell(\sigma)$ 的单对换分解. 我们将对 $n + \ell(\sigma)$ 递归地论证.

当 $n = 1$ 时 $\sigma = \mathrm{id}$ 而 $\ell(\sigma) = 0$, 此时天下太平.

今起假定 $n \geq 2$. 若 $\sigma(1) = 1$, 则 σ 可以等同于 $\{2, \dots, n\}$ 上的置换; 在 $\{2, \dots, n\}$ 上计算的逆序数仍是 $\ell(\sigma)$, 这就将问题递归地化到 $\mathfrak{S}_{n-1}$ 上.

以下设 $i := \sigma^{-1}(1) > 1$. 定义 $\sigma' := \sigma s_{i-1}$. 因此

$$\sigma = \begin{pmatrix} \cdots & i-1 & i & \cdots \\ \cdots & \sigma(i-1) & 1 & \cdots \end{pmatrix}, \quad \sigma' = \begin{pmatrix} \cdots & i-1 & i & \cdots \\ \cdots & 1 & \sigma(i-1) & \cdots \end{pmatrix}$$

其中的省略部分无异. 于是 $\ell(\sigma') \leq \ell(\sigma) - 1$. 递归可知存在单对换 $\tau_1, \dots, \tau_{k'}$ 使得 $\sigma' = \tau_1 \cdots \tau_{k'}$ 而 $k' \leq \ell(\sigma')$, 于是 $\sigma = \tau_1 \cdots \tau_{k'} s_{i-1}$ 给出长度 $\leq \ell(\sigma)$ 的分解.

第二步是证明 $\ell(\sigma)$ 是最短可能的长度. 设有 $\sigma = \tau_1 \cdots \tau_\ell$, 其中每个 τ_i 都是单对换. 反复应用练习 5.1.7 (ii) 可得

$$\ell(\sigma) = \ell(\tau_1 \cdots \tau_\ell) \leq \ell(\tau_1 \cdots \tau_{\ell-1}) + 1 \leq \cdots \leq \ell(\mathrm{id}) + \ell = \ell.$$

明所欲证. □

长度为 $\ell(\sigma)$ 的单对换分解称为 σ 的**既约表法**. 同一个 σ 可以有多种既约表法, 例如练习 5.1.4 揭示的 $s_1 s_2 s_1 = s_2 s_1 s_2$ 等.

综上, 逆序数可以设想为置换的某种 "长度", 这类长度函数在置换群及其推广的研究中至关重要.

例 5.1.9 作为简单然而常用的特例, 我们选定 $2 \leq m \leq n$, 研究

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & m & m+1 & \cdots & n \\ m & 1 & \cdots & m-1 & m+1 & \cdots & n \end{pmatrix} \in \mathfrak{S}_n$$

如何分解为单对换的乘积. 注意到 σ 的效果是让 $1, \dots, m-1$ 作为一个整体和 m 调换, 或者说是让 $1, \dots, m$ 向右轮换, 保持 $> m$ 的部分不动; 它可以通过逐步将 m 左移而得, 写法是

$$\sigma = s_1 \cdots s_{m-1}.$$

同时也注意到 σ 的逆序是 $(1,2), \dots, (1,m)$, 故 $\ell(\sigma) = m - 1$, 而上式为既约表法.

练习 5.1.10 证明 $\mathfrak{S}_n$ 中存在唯一的 σ 使得 $\ell(\sigma)$ 极大; 具体写下它的形式.

提示〉 为了使逆序数极大, 请思量从 $\{1,\dots,n\}$ 到 $\{1,\dots,n\}$ 的严格递减函数.

定义–命题 5.1.11 *存在唯一的映射 $\operatorname{sgn}:\mathfrak{S}_n\to\{\pm1\}$ 使下述性质成立:*

(i) *对所有 $\sigma,\xi\in\mathfrak{S}_n$ 皆有 $\operatorname{sgn}(\sigma\xi)=\operatorname{sgn}(\sigma)\operatorname{sgn}(\xi)$,*

(ii) *若 $\tau\in\mathfrak{S}_n$ 为对换, 则 $\operatorname{sgn}(\tau)=-1$.*

这样的 sgn 必然满足 $\operatorname{sgn}(\mathrm{id})=1$ 和 $\operatorname{sgn}(\sigma^{-1})=\operatorname{sgn}(\sigma)^{-1}$. 它可以用逆序数表达为

$$\operatorname{sgn}(\sigma)=(-1)^{\ell(\sigma)}.$$

证明 首先说明唯一性. 因为任何 σ 都能写成对换的乘积 $\tau_1\cdots\tau_\ell$, 故 (i) 和 (ii) 蕴涵 $\operatorname{sgn}(\sigma)=(-1)^{\ell}$. 注意到尽管分解 $\sigma=\tau_1\cdots\tau_\ell$ 不唯一, 但只要所求的映射 sgn 存在, 则上述论证不仅说明 sgn 唯一, 还说明 ℓ 的奇偶性只依赖于 σ.

对条件 (i) 代入 $\sigma=\xi=\mathrm{id}$ 可得 $\sigma(\mathrm{id})=1$. 再代入 $\xi=\sigma^{-1}$ 可得 $\operatorname{sgn}(\sigma^{-1})=\operatorname{sgn}(\sigma)^{-1}=\operatorname{sgn}(\sigma)$.

存在性需要一些准备工作. 考虑函数

$$\begin{aligned}\Delta:\mathbb{Z}^n &\longrightarrow \mathbb{Z}\\ \vec{x}=(x_1,\dots,x_n) &\longmapsto \prod_{1\le i<j\le n}(x_i-x_j).\end{aligned}$$

对所有 $\vec{x}\in\mathbb{Z}^n$, 记

$$\sigma\vec{x}:=\left(x_{\sigma^{-1}(1)},\dots,x_{\sigma^{-1}(n)}\right)\in\mathbb{Z}^n.$$

对于任意 $i<j$, 或者 $\sigma^{-1}(i)<\sigma^{-1}(j)$, 此时 $(h,k):=(\sigma^{-1}(i),\sigma^{-1}(j))\notin\mathrm{Inv}_\sigma$; 又或者 $\sigma^{-1}(i)>\sigma^{-1}(j)$, 此时 $(h,k):=(\sigma^{-1}(j),\sigma^{-1}(i))\in\mathrm{Inv}_\sigma$. 所有满足 $1\le h<k\le n$ 的 $(h,k)\in\mathbb{Z}^2$ 都是如此得到的, 按此重新整理连乘积便得到

$$\begin{aligned}\Delta(\sigma\vec{x}) &= \prod_{\substack{h<k\\(h,k)\notin\mathrm{Inv}_\sigma}}(x_h-x_k)\prod_{\substack{h<k\\(h,k)\in\mathrm{Inv}_\sigma}}(x_k-x_h)\\ &=(-1)^{\ell(\sigma)}\Delta(\vec{x}).\end{aligned}$$

另一方面, 我们断言对所有 $\sigma,\xi\in\mathfrak{S}_n$ 和 $\vec{x}\in\mathbb{Z}^n$ 都有

$$(\sigma\xi)\vec{x}=\left(x_{\xi^{-1}\sigma^{-1}(1)},\dots,x_{\xi^{-1}\sigma^{-1}(n)}\right)=\sigma(\xi\vec{x}).$$

为此, 不妨视 $\vec{x}$ 为映射 $\{1,\dots,n\}\to\mathbb{Z}$, 映 i 为 $\vec{x}(i)=x_i$; 按定义, $\sigma\vec{x}$ 对应到映射的合成 $\vec{x}\circ\sigma^{-1}$, 而 $(\sigma\xi)\vec{x}=\vec{x}\circ(\sigma\xi)^{-1}=(\vec{x}\circ\xi^{-1})\circ\sigma^{-1}=\sigma(\xi\vec{x})$. 断言得证.

综上可得

$$\Delta\left((\sigma\xi)\vec{x}\right) = (-1)^{\ell(\sigma)}\Delta(\xi\vec{x}) = (-1)^{\ell(\sigma)}(-1)^{\ell(\xi)}\Delta(\vec{x});$$

另一方面 $\Delta\left((\sigma\xi)\vec{x}\right) = (-1)^{\ell(\sigma\xi)}\Delta(\vec{x})$. 取 $\vec{x}$ 使其坐标各异, 则 $\Delta(\vec{x}) \neq 0$, 由此知 $(-1)^{\ell(\sigma\xi)} = (-1)^{\ell(\sigma)}(-1)^{\ell(\xi)}$.

定义 $\operatorname{sgn}(\sigma) := (-1)^{\ell(\sigma)}$, 于是上一段说明 (i) 成立. 对于 (ii), 注意到对所有 $1 \leq i \leq n-1$ 皆有 $\ell(s_i) = 1$, 从而 $\operatorname{sgn}(s_i) = -1$. 对于一般的对换 $\tau \in \mathfrak{S}_n$, 注记 5.1.3 说明存在 σ 使得 $\tau = \sigma s_1 \sigma^{-1}$, 从而

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(s_1)\operatorname{sgn}(\sigma)^{-1} = \operatorname{sgn}(s_1) = -1.$$

明所欲证. □

定义 5.1.12 设 $\sigma \in \mathfrak{S}_n$. 若存在一族对换 $\tau_1, \ldots, \tau_\ell$ 使得 $\sigma = \tau_1 \cdots \tau_\ell$, 其中 $\ell \in \mathbb{Z}_{\geq 0}$ 为偶数 (或奇数), 则称 σ 为**偶置换** (或**奇置换**).

定义—命题 5.1.11 证明第一段说明任意对换分解 $\sigma = \tau_1 \cdots \tau_\ell$ 中的 ℓ 的奇偶性仅由 σ 决定; 更精确地说, 而 $\ell \equiv \ell(\sigma) \pmod 2$. 因此 σ 是偶置换当且仅当 $\operatorname{sgn}(\sigma) = 1$. 这些正负号将直接体现在行列式的操作中.

注记 5.1.13 尽管上述定义针对的是 $\{1, \ldots, n\}$ 上的置换, 但一个元素 $\tau \in \mathfrak{S}_n$ 是否是对换并不依赖 $1, \ldots, n$ 的排序. 因此, 对于任意有限非空集 X 上的置换都可以谈论其奇偶性. 作为应用, 回忆 (5.1.2) 对非空子集 $Y \subset X$ 给出的嵌入 $b : \mathfrak{S}_Y \hookrightarrow \mathfrak{S}_X$; 按定义, b 显然映对换为对换, 同时又保持乘法, 因而嵌入 b 也自动保持置换的奇偶性.

5.2 几何动机: 有向体积

读者或许已经见过 2 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}$$

和 3 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32};$$

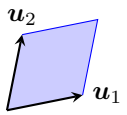
中学数学中习惯取 a_{ij} 为有理数, 实数或复数, 但这些代数表达式在一般的域 F 上同样有意义. 就代数学的观点, 这些量和 2 元或 3 元线性方程组的解直接相关. 在 $F = \mathbb{R}$ 的情形, 它们则有基于面积 (或体积) 的几何诠释.

以 $\mathbb{R}$ 上的 2 阶行列式为例, 记

$$\boldsymbol{u}_1 := (a_{11}, a_{21}), \quad \boldsymbol{u}_2 := (a_{12}, a_{22}), \quad \boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^2,$$

并且引入符号

$$\begin{aligned} \diamondsuit(\boldsymbol{u}_1, \boldsymbol{u}_2) := \mathbb{R}^2 \text{ 中的平行四边形} \quad & \\ = \{t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 : t_1, t_2 \in [0,1]\}. & \end{aligned}$$



平面解析几何学的一则标准事实是

$$|a_{11}a_{22} - a_{12}a_{21}| = \diamondsuit(\boldsymbol{u}_1, \boldsymbol{u}_2) \text{ 的面积}.$$

左式加绝对值是必要的, 因为面积总是非负实数. 但我们也可以转换思路, 将行列式 $a_{11}a_{22}-a_{12}a_{21}$ 看作是一种带正负号的面积, 称为有向面积. 可以验证 $a_{11}a_{22}-a_{12}a_{21} > 0$ 当且仅当向量组 $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ 线性无关, 并且是正向的, 也就是说从 $\boldsymbol{u}_1$ 到 $\boldsymbol{u}_2$ 是逆时针走向[1)]. 若 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 线性相关, 则 $\diamondsuit(\boldsymbol{u}_1, \boldsymbol{u}_2)$ 的面积为 0, 免论正负号.

类似地, 三阶行列式可以诠释为 $\mathbb{R}^3$ 中三个向量 $\boldsymbol{u}_j := (a_{1j}, a_{2j}, a_{3j})$ 张成的平行六面体的有向体积 $(j = 1, 2, 3)$, 其中正向的概念以右手定则确定. 假如读者愿意想象高维度的体积, 则行将定义的 n 阶行列式都可以按此类推.

关于平行四边形和平行六面体体积的推导是初等的, 本节目的不在详述这些几何内容, 而是从有向体积的直观概念出发, 说明它们所须具备的一般性质. 这些推导非但不以行列式理论为前提, 还为 n 阶行列式的一般定义提供了必要的线索.

设 $n \in \mathbb{Z}_{\geq 1}$, 而 V 为 n 维 $\mathbb{R}$-向量空间. 以下假定:

1. 在 V 上可以合理地定义体积和向量的长度, 夹角的概念, 具备和 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 情形类似的性质. 精确的描述涉及第九章将介绍的内积结构, 在此先按直观行事.

2. 在 V 上已经赋予了一个定向. 这相当于给定一个无交并分解

$$\{(u_1, \ldots, u_n) : V \text{ 的有序基}\} = \{\text{正向}\} \sqcup \{\text{负向}\}.$$

定向的精确定义需要以行列式理论为前提, 留待练习 5.4.11 来处理. 本节暂且将定向领会为一个 "自然如此" 的观念; 在定向所需的一切性质中, 后续讨论只涉及最直观的两条:

▷ **变号反向** 若 $(u_1, \ldots, u_n)$ 是正向的, $1 \leq i \leq n$, 则 $(u_1, \ldots, -u_i, \ldots, u_n)$ 是负向的;

[1)] 一种方法是验证旋转 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 不改变行列式, 在此前提下, 问题化到 $\boldsymbol{u}_1 = (1, 0)$ 的简单情形.

▷ **微扰不变** 若在所有有序基构成的空间中, 自 $(u_1,\dots,u_n)$ 可以连续地过渡到 $(u_1',\dots,u_n')$, 则两者的定向相同.

对任意向量组 $(u_1,\dots,u_n)\in V^n$, 记它们张成的几何对象为

$$\diamondsuit(u_1,\dots,u_n):=\left\{\sum_{i=1}^n t_iu_i:\forall i,\ t_i\in[0,1]\right\};$$

这是平行四边形或平行六面体的高维版本. 留意到当 $u_1,\dots,u_n$ 线性相关时, $\diamondsuit(u_1,\dots,u_n)$ 是一个至多仅有 $n-1$ 个自由度的几何对象, 它作为 V 的闭子集应当不占体积. 于是有向体积理应取以下定义.

定义 5.2.1 设 V 上已有取定的定向等概念. 定义映射 $D:V^n\to\mathbb{R}$ 如下:

$$D(u_1,\dots,u_n)=\begin{cases}0, & u_1,\dots,u_n \text{ 线性相关},\\ \diamondsuit(u_1,\dots,u_n)\text{ 的体积}, & \text{线性无关, 正向},\\ -(\diamondsuit(u_1,\dots,u_n)\text{ 的体积}), & \text{线性无关, 负向}.\end{cases}$$

我们称 $D(u_1,\dots,u_n)$ 是由向量组 $u_1,\dots,u_n\in V$ (记顺序) 确定的**有向体积**.

有向体积之 "向" 即是有序基的定向, 体现为正负号.
尽管目前对 D 仍一无所知, 但以下性质可以从基本原理直接推得.

▷ **退化** 若存在 $1\le i\ne j\le n$ 使得 $u_i=u_j$, 则 $D(u_1,\dots,u_n)=0$.

这是因为此时 $u_1,\dots,u_n$ 线性相关.

▷ **纯量乘法** 设 $1\le i\le n$ 而 $t\in\mathbb{R}$, 则

$$D(u_1,\dots,tu_i,\dots,u_n)=tD(u_1,\dots,u_n).$$

为此, 无妨假设 $u_1,\dots,u_n$ 是基, 否则两边出现的向量组皆线性相关, 同样给出 0.

- ⋆ 当 $t>0$ 时, 将 u_i 伸缩为 tu_i 不改变向量组的定向 (利用微扰不变性: 让 u_i 连续地变化到 tu_i), 而 $\diamondsuit(\dots,tu_i,\dots)$ 在 u_i 的方向较原来伸缩了 t 倍, 在其余 $n-1$ 个方向不变, 因此体积也应当相应地伸缩 t 倍.
- ⋆ 若 $t<0$, 则

$$\begin{aligned}D(\dots,tu_i,\dots)=D(\dots,-|t|u_i,\dots)&\xlongequal{\text{定向倒转}}-D(\dots,|t|u_i,\dots)\\&\xlongequal{\text{上一步}}-|t|D(\dots,u_i,\dots)=tD(\dots,u_i,\dots).\end{aligned}$$

最后, $t=0$ 的情形是容易的, 因为 $u_1,\dots,0,\dots,u_n$ 线性相关.

▷ **加法** 对任意 $u \in V$ 和 $1 \le i \le n$, 我们有

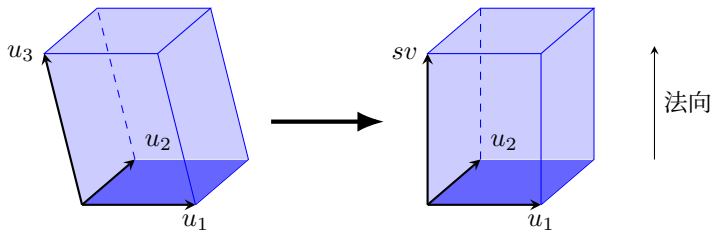
$$D(u_1, \dots, u_i + u, \dots, u_n) = D(u_1, \dots, u_i, \dots, u_n) + D(u_1, \dots, u, \dots, u_n).$$

注意到若 $\{u_j : j \ne i\}$ 线性相关, 则两侧涉及的向量组也全都线性相关. 因此处理 $\{u_j : j \ne i\}$ 线性无关的情形即可.

任取 $v \in V$ 使得 $u_1, \dots, \underbrace{v}_{\text{第 } i \text{ 项}}, \dots u_n$ 为基. 于是存在系数 s, t 和 $(a_j)_j, (b_j)_j$ 使得

$$u_i = sv + \sum_{j \ne i} a_j u_j, \quad u = tv + \sum_{j \ne i} b_j u_j.$$

观察到 $\diamondsuit(\dots, u_i, \dots)$ 和 $\diamondsuit(\dots, sv, \dots)$ 互为 "斜推", 三维情形如下图所示 (取 $i = n = 3$):



两者体积相同, 这是因为斜推不改变平行六面体的底和高, 或者说因为两者沿法向的每个水平切片都仅差一个平移, 故切片等面积. 高维情形依此发挥想象力.

请读者简单地验证 $u_1, \dots, u_i, \dots, u_n$ 线性无关当且仅当 $s \ne 0$. 当 $s \ne 0$ 时, 斜推不改变定向, 这是微扰不变性的应用 (因为可以 "连续地" 斜推到位), 于是

$$D(\dots, u_i, \dots) = D(\dots, sv, \dots) \overset{\text{纯量乘法性质}}{=\!=} sD(\dots, v, \dots).$$

上式对 $s = 0$ 或者说 $u_1, \dots, u_n$ 线性相关的情形也平凡地成立. 同理,

$$D(\dots, u, \dots) = tD(\dots, v, \dots),$$
$$D(\dots, u_i + u, \dots) = (s + t)D(\dots, v, \dots),$$

故 D 的加法性质确实成立.

以上对有向体积 D 提炼出的三条性质乍看仿佛不劳而获, 然而 §5.3 将说明它们几乎唯一刻画了 D, 至多差一个比例常数. 满足此三条性质的映射 D 因此是合理的研究对象. 又因为这些性质的表述并不涉及实数域 $\mathbb{R}$ 所特有的性质, 故可扩及一般的域 F, 由此便为行列式的一般定义铺平了道路.

5.3 一类交错形式的刻画

选定域 F. 作为下节内容的预告, 我们即将从两种观点对一般的 $n \in \mathbb{Z}_{\geq 1}$ 探讨 n 阶行列式. 考虑矩阵 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$.

1. 定义–命题 5.4.6 将以具体的公式定义 $\boldsymbol{A}$ 的行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

   第一种观点是视此为以 n 个列向量 $\boldsymbol{v}_i := (a_{1i}, \ldots, a_{ni}) \in F^n$ 为变元的函数 $(i = 1, \ldots, n)$, 取值在域 F 中. 暂且将此函数记作 D. 受 §5.2 关于有向体积的讨论启发, 我们期望 D 应当具有以下性质:

   - ⋆ 它对每个变元都是线性的, 换言之, 对每个 $1 \leq i \leq n$, 我们有

$$\begin{aligned} D(\ldots, v_i + v_i', \ldots) &= D(\ldots, v_i, \ldots) + D(\ldots, v_i', \ldots), \\ D(\ldots, tv_i, \ldots) &= tD(\ldots, v_i, \ldots), \end{aligned}$$

     此处 $t \in F$, 省略号代表其余 $n-1$ 个变元, 它们是选定不动的列向量.

   - ⋆ 任两列相等导致行列式为 0: 我们有 $D(\ldots, v, \ldots, v, \ldots) = 0$.

2. 将 $\boldsymbol{A}$ 通过 F^n 的标准基视同 F^n 到自身的线性变换, 则 $\boldsymbol{A}$ 对应的 n 阶行列式又记为 $\det \boldsymbol{A}$. 第二种观点是将行列式看成由线性变换确定的一个量. 定义 5.4.1 将给出 $\det \boldsymbol{A}$ 的抽象定义, 它无关基的选取.

关于行列式的这两种观点密切相关, 但它们的理论角色需要区别: 我们将对 n 维 F-向量空间 V 研究一类特殊的映射 $D : V^n \to F$, 它们具有第一种观点提及的性质, 然后将任意线性映射 $T \in \mathrm{End}(V)$ 的行列式 $\det T$ 抽象地理解为一个相关的**比例常数**; 在选定有序基的前提下, 命题 5.4.5 将阐明两种观点如何联系. 其余一切性质都是水到渠成的.

定义 5.3.1 (交错形式) 设 V 为 F-向量空间, 对所有 $m \in \mathbb{Z}_{\geq 1}$, 记

$$\mathcal{D}_{V,m} := \{D : V^m \to F, \text{ 满足以下性质 } \mathbf{D.1}, \mathbf{D.2}\}.$$

D.1 对每个 $1 \leq i \leq m$,

$$\begin{aligned} D(\ldots, v_i + v_i', \ldots) &= D(\ldots, v_i, \ldots) + D(\ldots, v_i', \ldots), \\ D(\ldots, tv_i, \ldots) &= tD(\ldots, v_i, \ldots), \end{aligned}$$

此处 $v_i, v_i' \in V$ 而 $t \in F$, 省略号代表的其余 $m-1$ 个变元选定不动.

D.2 若存在 $1 \le i < j \le m$ 使得 $v_i = v_j$, 则 $D(v_1, \dots, v_m) = 0$.

属于 $\mathcal{D}_{V,m}$ 的映射也称为 V 上的 m 元交错形式.

对于有限维 F-向量空间 V, 记 $n := \dim V$ 并且定义

$$\mathcal{D}_V := \begin{cases} \mathcal{D}_{V,n}, & n \ge 1, \\ F, & n = 0. \end{cases}$$

留意到 $\mathcal{D}_{V,m}$ 非空, 它至少包含零映射 $0 : V^m \to F$, 映一切 $(v_1, \dots, v_m)$ 为 0. 一如许多映射空间的例子, $\mathcal{D}_{V,m}$ 对以下的 "逐点" 运算也构成向量空间:

$$\begin{aligned} (D + D')(v_1, \dots, v_m) &:= D(v_1, \dots, v_m) + D'(v_1, \dots, v_m), \\ (tD)(v_1, \dots, v_m) &:= t \cdot D(v_1, \dots, v_m). \end{aligned}$$

注意: $m = 1$ 时条件 **D.2** 是多余的, 这时 $\mathcal{D}_{V,1} = \mathrm{Hom}(V, F) = V^\vee$.

从 **D.1** 和 **D.2** 能够推导更多重要的运算规律, 其表述涉及 §5.1 介绍的置换.

引理 5.3.2 设 $m \in \mathbb{Z}_{\ge 1}$, $D \in \mathcal{D}_{V,m}$, 而 $\sigma \in \mathfrak{S}_m$, 则

$$D\left(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(m)}\right) = \mathrm{sgn}(\sigma) D(v_1, \dots, v_m),$$

其中 $\mathrm{sgn} : \mathfrak{S}_m \to \{\pm 1\}$ 如定义–命题 5.1.11. 作为特例, 取 σ 为对换 $(i\ j)$ 则有

$$D(\dots, v_j, \dots, v_i, \dots) = -D(\dots, v_i, \dots, v_j, \dots).$$

证明 记 $\vec{v} = (v_1, \dots, v_m) \in V^m$. 对所有 $\vec{v} \in V^m$ 和 $\sigma \in \mathfrak{S}_m$ 定义

$$\sigma\vec{v} = \left(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(m)}\right).$$

熟悉的验证方式给出

$$(\sigma\xi)\vec{v} = \sigma(\xi\vec{v}), \quad \sigma, \xi \in \mathfrak{S}_m,\ \vec{v} \in V^m.$$

原问题相当于证 $D(\sigma\vec{v}) = \mathrm{sgn}(\sigma) D(\vec{v})$. 从对换 $\sigma = (i\ j)$ 的情形起步, 照定义推导

$$\begin{aligned} 0 &= D(\dots, \underbrace{v_i + v_j}_{\text{变元 } i}, \dots, \underbrace{v_i + v_j}_{\text{变元 } j}, \dots) \\ &\overset{\textbf{D.1}}{=\!=} D(\dots, v_i, \dots, v_i, \dots) + D(\dots, v_i, \dots, v_j, \dots) \\ &\qquad + D(\dots, v_j, \dots, v_i, \dots) + D(\dots, v_j, \dots, v_j, \dots) \\ &\overset{\textbf{D.2}}{=\!=} D(\dots, v_i, \dots, v_j, \dots) + D(\dots, v_j, \dots, v_i, \dots) \\ &= D(\vec{v}) + D(\sigma\vec{v}). \end{aligned}$$

至于一般情形, 以命题 5.1.8 将 σ 表为对换的积 $\tau_1 \cdots \tau_\ell$. 上一步表明

$$D(\sigma\vec{v}) = D(\tau_1 \cdots \tau_\ell \vec{v}) = -D(\tau_2 \cdots \tau_\ell \vec{v}) = \cdots = (-1)^\ell D(\vec{v}).$$

然而定义–命题 5.1.11 表明 $(-1)^\ell = \mathrm{sgn}(\sigma)$. 证毕. □

注记 5.3.3 如果承认 $D : V^m \to F$ 具有交换第 i 和 j 个变元导致变号的性质, 则代入 $v_i = v = v_j$ 立刻导致

$$2 \cdot D(\ldots, \underbrace{v}_{\text{变元 } i}, \ldots, \underbrace{v}_{\text{变元 } j}, \ldots) = 0.$$

由于一般的域 F 中未必有 $2x = 0 \implies x = 0$, 上式不足以导出 $D(\ldots, v, \ldots, v, \ldots) = 0$; 详见 §3.7 关于域特征的讨论. 因此定义 5.3.1 的 **D.2** 强于变号条件.

交错形式的另一个性质是对于任意 $D \in \mathcal{D}_{V,m}$ 和 $1 \leq i \neq j \leq m$, 我们有

$$D(\ldots, v_i, \ldots, v_j, \ldots) = D(\ldots, v_i + cv_j, \ldots, v_j, \ldots),$$

其中 $v_1, \ldots, v_m \in V$ 和 $c \in F$ 任取. 这是因为两边的差等于

$$cD(\cdots, v_j, \ldots, v_j, \ldots) = 0.$$

由上一则性质可以推知

$$v_1, \ldots, v_m \in V \quad \text{线性相关} \implies D(v_1, \ldots, v_m) = 0. \tag{5.3.1}$$

诚然, 不失一般性可假设 $v_1 = c_2 v_2 + \cdots + c_m v_m$, 其中 $c_2, \ldots, c_m \in F$; 将上一条性质反复运用, 给出

$$D(v_1, \ldots, v_m) = D\left(v_1 - \sum_{i=2}^{m} c_i v_i, v_2, \ldots, v_m\right) = D(0, v_2, \ldots, v_m).$$

由于 D 对第一个变元是线性映射, 故最右式为 0.

对于 n 维向量空间 V (设 $n \in \mathbb{Z}_{\geq 1}$), 以上性质表明 $m = n$ 情形的交错形式地位特殊, 因为根据引理 4.4.12, 一旦 $m > n$ 便不复有非零的 $D \in \mathcal{D}_{V,m}$.

现在聚焦于 $\mathcal{D}_V := \mathcal{D}_{V,n}$ 的结构, 这与我们行将定义的行列式直接相关. 它是否包含非零元? 它的维数几何? 为了进一步处理这些问题, 我们选取 V 的有序基 $e_1, \ldots, e_n$. 将任意 $(v_1, \ldots, v_n) \in V^n$ 按此展开如

$$v_i = \sum_{j=1}^{n} a_{i,j} e_j, \quad i = 1, \ldots, n, \tag{5.3.2}$$

其中的系数 $a_{i,j} \in F$. 设 $D \in \mathcal{D}_V$, 我们有

$$\begin{aligned} D(v_1, \ldots, v_n) &= \sum_{j_1=1}^{n} a_{1,j_1} D(e_{j_1}, v_2, \ldots, v_n) \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1,j_1} a_{2,j_2} D(e_{j_1}, e_{j_2}, v_3, \ldots, v_n) \\ &= \cdots \\ &= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1,j_1} \cdots a_{n,j_n} D(e_{j_1}, \ldots, e_{j_n}). \end{aligned}$$

若存在 $i \neq k$ 使得 $j_i = j_k$, 则 $D(e_{j_1}, \ldots, e_{j_n}) = 0$. 所以对上式有贡献的 $(j_1, \ldots, j_n)$ 只能是 $1, \ldots, n$ 的排列, 既不遗漏也不重复. 将这些排列诠释为 $\{1, \ldots, n\}$ 到自身的双射 $\sigma : i \mapsto j_i$, 上式遂改写为

$$\sum_{\sigma \in \mathfrak{S}_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} D\left(e_{\sigma(1)}, \ldots, e_{\sigma(n)}\right).$$

应用引理 5.3.2 和定义–命题 5.1.11 包含的性质 $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, 进一步导出

$$D(v_1, \ldots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} D(e_1, \ldots, e_n). \tag{5.3.3}$$

这就说明 $\mathcal{D}_V$ 的元素至多只依赖一个参数, 即它在 $(e_1, \ldots, e_n)$ 处的取值.

引理 5.3.4　*设 V 为有限维 F-向量空间, 则 $\dim \mathcal{D}_V \leq 1$.*

证明　零空间情形按定义有 $\mathcal{D}_V = F$, 故可设 $n := \dim V \geq 1$. 选定 V 的有序基 $e_1, \ldots, e_n$. 考虑映射

$$\begin{aligned} \mathcal{D}_V &\longrightarrow F \\ D &\longmapsto D(e_1, \ldots, e_n). \end{aligned}$$

这显然是线性映射, 而 (5.3.3) 说明这还是单射, 因此 $\dim \mathcal{D}_V \leq \dim F = 1$. □

定理 5.3.5　*设 V 为有限维向量空间, 则 $\dim \mathcal{D}_V = 1$. 若 $e_1, \ldots, e_n$ 是 V 的有序基 $(n \geq 1)$, 记为 $\mathbf{e}$, 则存在唯一的 $D_{\mathbf{e}} \in \mathcal{D}_V$ 使得 $D_{\mathbf{e}}(e_1, \ldots, e_n) = 1$.*

证明　仍然假定 $n := \dim V \geq 1$, 否则结论是平凡的. 首先说明 $\dim \mathcal{D}_V = 1$. 由于引理 5.3.4 已说明 $\dim \mathcal{D}_V \leq 1$, 问题在于具体地构造 $D \in \mathcal{D}_V \smallsetminus \{0\}$. 由于 (5.3.3) 的启发, 我们选定有序基 $e_1, \ldots, e_n$ 并且定义

$$D(v_1, \ldots, v_n) := \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

其中 a_{ij} 是诸 v_i 在 $\mathbf{e}$ 下的坐标, 按 (5.3.2) 确定. 以下验证 $D \in \mathcal{D}_V$.

首先验证 **D.1**. 固定 $1 \leq i \leq n$. 将 v_i 换作 $v_i + v_i'$ 相当于对所有 j 将 $a_{i,j}$ 换作 $a_{i,j} + a_{i,j}'$, 其余不动; 此处取展开 $v_i' = \sum_{j=1}^n a_{i,j}' e_j$. 因此 $D(\ldots, v_i + v_i', \ldots)$ 等于

$$\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)} \cdots a_{n,\sigma(n)} + \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)}' \cdots a_{n,\sigma(n)},$$

亦即 $D(\ldots, v_i, \ldots) + D(\ldots, v_i', \ldots)$. 类似地, 将 v_i 换作 tv_i 相当于对所有 j 将 $a_{i,j}$ 换作 $ta_{i,j}$, 由此见得 $D(\ldots, tv_i, \ldots) = tD(\ldots, v_i, \ldots)$.

其次设 $n \geq 2$ 并验证 **D.2**. 设 $v_1, \ldots, v_n \in V$ 满足 $v_i = v_j$, 其中 $1 \leq i < j \leq n$. 于是有

$$a_{i,h} = a_{j,h}, \quad h = 1, \ldots, n.$$

命 τ 为对换 $(i\ j)\in\mathfrak{S}_n$. 对所有 $\sigma\in\mathfrak{S}_n$, 这导致

$$\begin{aligned} a_{i,\sigma(i)} &= a_{j,\sigma(i)} = a_{j,\sigma\tau(j)},\\ a_{j,\sigma(j)} &= a_{i,\sigma(j)} = a_{i,\sigma\tau(i)},\\ k\notin\{i,j\} &\implies a_{k,\sigma(k)} = a_{k,\sigma\tau(k)}. \end{aligned} \tag{5.3.4}$$

现在将 $\mathfrak{S}_n$ 的元素按照

$$\sigma \leftrightarrow \sigma\tau$$

进行配对. 因为 $(\sigma\tau)\tau=\sigma\tau^2=\sigma$, 配对是双向的; 此外元素不能自配对, 因为 $\sigma\tau=\sigma$ 两边左合成 σ^{-1} 给出 $\tau=\mathrm{id}$, 矛盾. 在 D 的和式中, 我们可先将每一对 $\sigma\leftrightarrow\sigma\tau$ 中的两项相加, 再沿着所有子集 $\{\sigma,\sigma\tau\}$ 加总, 其产物是

$$\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sgn}(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} = \sum_{\{\sigma,\sigma\tau\}\subset\mathfrak{S}_n}\left(\operatorname{sgn}(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}+\operatorname{sgn}(\sigma\tau)a_{1,\sigma\tau(1)}\cdots a_{n,\sigma\tau(n)}\right).$$

然而 (5.3.4) 导致

$$a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} = a_{1,\sigma\tau(1)}\cdots a_{n,\sigma\tau(n)},$$

另一方面 $\operatorname{sgn}(\tau)=-1$ 蕴涵 $\operatorname{sgn}(\sigma)+\operatorname{sgn}(\sigma\tau)=0$. 综上可得 $D(v_1,\dots,v_n)=0$.

将以上构造的 $\mathcal{D}_V$ 的元素记为 $D_{\mathbf{e}}$. 以下验证 $D_{\mathbf{e}}(e_1,\dots,e_n)=1$. 代入 $v_i=e_i$ 可见 $a_{i,i}=1$, 而 $i\neq j$ 时 $a_{i,j}=0$. 于是使 $a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}$ 非零的唯一可能是 $\sigma=\mathrm{id}$, 此时 $a_{1,1}\cdots a_{n,n}=1$ 而 $\operatorname{sgn}(\mathrm{id})=1$. 因此 $D_{\mathbf{e}}(e_1,\dots,e_n)=1$; 特别地 $D_{\mathbf{e}}$ 不恒为 0.

一般的 $D\in\mathcal{D}_V$ 依 (5.3.3) 由它在 $(e_1,\dots,e_n)$ 处的取值唯一确定. 这就说明满足 $D_{\mathbf{e}}(e_1,\dots,e_n)=1$ 的 $D_{\mathbf{e}}$ 唯一. 至此证完所有断言. □

上述论证不仅说明了 $\dim\mathcal{D}_V=1$, 还对 V 的每个有序基 $\mathbf{e}$ 具体构造了非零元 $D_{\mathbf{e}}\in\mathcal{D}_V$; 这些交错形式 $D_{\mathbf{e}}$ 将是行列式理论的重要一环.

例 5.3.6 取 V 为 n 维 $\mathbb{R}$-向量空间. 只要在 V 上具有合理的定向等概念, 则 §5.2 探讨的有向体积 $D:V^n\to\mathbb{R}$ 总是 $\mathcal{D}_V$ 的元素; 不论 D 的实际定义为何, 定理 5.3.5 表明它精确到一个比例常数是唯一的.

5.4 行列式的定义和基本性质

定义 5.3.1 已经对所有 V 和 $m\geq 1$ 定义了向量空间 $\mathcal{D}_{V,m}$. 现在考虑有限维的 V, W 连同线性映射 $T\in\operatorname{Hom}(V,W)$. 由 T 诱导相应的映射

$$\begin{aligned} T^*:\mathcal{D}_{W,m} &\longrightarrow \mathcal{D}_{V,m}\\ D &\longmapsto [T^*D:(v_1,\dots,v_m)\mapsto D(Tv_1,\dots,Tv_m)]; \end{aligned}$$

它的效用是将交错形式从 W^m 拖回 V^m, 走向和 T 正好相反.

为了说明定义合理, 必须验证对于每个 $D \in \mathcal{D}_{W,m}$, 右边的映射 T^*D 满足 **D.1** 和 **D.2**; 因为 T 是线性映射, 这毫不困难.

进一步, T^* 还是从 $\mathcal{D}_{W,m}$ 到 $\mathcal{D}_{V,m}$ 的线性映射. 这也是定义的直接结论.

现在取 $V = W$ 和 $m = n := \dim V$, 于是 T^* 是从 1 维空间 $\mathcal{D}_V$ (定理 5.3.5) 到其自身的线性映射, 这般映射必然是伸缩 $D \mapsto cD$, 其中 $c \in F$. 一句话, T 的行列式就是此一比例常数 c.

定义 5.4.1 (线性映射的行列式) 设 V 是 n 维向量空间, $n \in \mathbb{Z}_{\geq 1}$. 对每个 $T \in \mathrm{End}(V)$ 定义 $\det T \in F$ 为使得下式成立的唯一元素:

$$T^*(D) = (\det T) \cdot D, \quad D \in \mathcal{D}_V;$$

等价的说法是: 对于每个 $D \in \mathcal{D}_V$ 和 $(v_1, \dots, v_n) \in V^n$,

$$D(Tv_1, \dots, Tv_n) = \det T \cdot D(v_1, \dots, v_n).$$

对于零空间的情形 ($n = 0$), 我们对 $T = 0_V = \mathrm{id}_V$ 规定 $\det(T) := 1$.

由于 $\dim \mathcal{D}_V = 1$, 上述条件对任何一个 $D \in \mathcal{D}_V \smallsetminus \{0\}$ 验证即可.

例 5.4.2 接续例 5.3.6 的讨论, 取 V 为 n 维 $\mathbb{R}$-向量空间, 然后将 $D \in \mathcal{D}_V \smallsetminus \{0\}$ 设想为有向体积映射, 则定义 5.4.1 相当于说 $T \in \mathrm{End}(V)$ 将有向体积按照比例 $\det T$ 作伸缩. 这是行列式的一种几何诠释, 它在多重积分的变量替换公式中自然地出现.

定理 5.4.3 (行列式的乘性) 行列式具有以下性质.

(i) $\det(\mathrm{id}_V) = 1$.

(ii) 设 $S, T \in \mathrm{End}(V)$, 则 $\det(ST) = \det(S)\det(T)$.

(iii) 当 T 可逆时 $\det T \in F$ 也可逆, 并且 $\det(T^{-1}) = (\det T)^{-1}$.

证明 零空间的情形是平凡的, 故以下假设 $n := \dim V \geq 1$. 首先, 行列式作为比例的定义即刻导致 $\det(\mathrm{id}_V) = 1$. 给定 $S, T \in \mathrm{End}(V)$, 按定义操演可得

$$\begin{aligned} \det(ST)D(v_1, \dots, v_n) &= D(S(Tv_1), \dots, S(Tv_n)) \\ &= (\det S)D(Tv_1, \dots, Tv_n) = (\det S)(\det T)D(v_1, \dots, v_n), \end{aligned}$$

其中 $D \in \mathcal{D}_V$ 和 $(v_1, \dots, v_n) \in V^n$ 任取, 这就给出 $\det(ST) = \det(S)\det(T)$. 代入 $S = T^{-1}$ 立得 $\det(T^{-1})\det T = 1$. □

命题 5.4.4 (行列式的共轭不变性) 设 $T \in \mathrm{End}(V)$ 而 $S : V \xrightarrow{\sim} W$ 是有限维向量空间的同构, 则 $STS^{-1} \in \mathrm{End}(W)$ 满足 $\det(STS^{-1}) = \det T$.

证明 不妨设 $n := \dim V \geq 1$. 所需论证是一个直截了当的 "结构搬运", 细说如下. 由 S 可得线性映射

$$\begin{aligned} \mathcal{D}_W &\longrightarrow \mathcal{D}_V \\ D' &\longmapsto [D : (v_1, \dots, v_n) \mapsto D'(Sv_1, \dots, Sv_n)], \end{aligned}$$

亦即先前定义的 S^*. 因为 S 是同构, $D' \mapsto D$ 自然也是同构. 现在按定义计算

$$\det(STS^{-1})D'(w_1, \dots, w_n) = D'\left(STS^{-1}w_1, \dots, STS^{-1}w_n\right),$$

其中 $w_1, \dots, w_n \in W$, $D' \in \mathcal{D}_W$. 记 $v_i = S^{-1}w_i$, 并且定义 D 如上, 则前一式化为

$$\det(STS^{-1})D(v_1, \dots, v_n) = D(Tv_1, \dots, Tv_n).$$

因为 S 是同构, 这里的 $v_1, \dots, v_n$ (或 D) 实则可以取遍 V (或 $\mathcal{D}_V$), 上式遂说明 $\det(STS^{-1}) = \det(T)$. □

行列式 $\det T$ 的计算和 $D \in \mathcal{D}_V$ 的计算紧密相关, 这是缘于以下结果.

命题 5.4.5 取定 n 维向量空间 V (设 $n \geq 1$) 的有序基 $e_1, \dots, e_n$, 记为 $\mathbf{e}$, 并且取定理 5.3.5 给出的 $D_{\mathbf{e}} \in \mathcal{D}_V$, 则

$$\det T = D_{\mathbf{e}}\left(Te_1, \dots, Te_n\right).$$

证明 按定义, $\det T = D_{\mathbf{e}}(Te_1, \dots, Te_n)/D_{\mathbf{e}}(e_1, \dots, e_n)$, 分母等于 1. □

一旦取基, 线性变换的行列式便落实为 $n \times n$ 矩阵的行列式. 以下为矩阵的行列式给出具体定义.

定义–命题 5.4.6 (矩阵的行列式) 取 $\mathbf{e}$ 为 F^n 的标准有序基, $n \in \mathbb{Z}_{\geq 1}$. 依此将每个 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$ 等同于 $\mathrm{End}(F^n)$ 的元素. 定义矩阵 $\boldsymbol{A}$ 的行列式为

$$\begin{aligned} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} &:= \det \boldsymbol{A} = D_{\mathbf{e}}\left(\boldsymbol{A}e_1, \dots, \boldsymbol{A}e_n\right) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \end{aligned}$$

证明 在 $\det \boldsymbol{A}$ 之后的第一个等式不外是复述命题 5.4.5. 第二个等式是基于 $\boldsymbol{A}e_i = \sum_{j=1}^n a_{ji}e_j$ 和 (5.3.3). 至于第三个等式, 留意到

$$a_{\sigma(1),1} \cdots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)},$$

而 $\sigma \mapsto \sigma^{-1}$ 是 $\mathfrak{S}_n$ 到自身的双射, 这是由于 $\sigma = (\sigma^{-1})^{-1}$. 综之,

$$\sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1}\cdots a_{\sigma(n),n} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma^{-1}) a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.$$

既然 $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, 由此便得到第三个等式. □

因此, 在选定有序基的前提下, 行列式既是一种比例常数, 又是 F^n 上的一种特定的交错形式.

留意到根据定义 5.4.1, 零空间 $F^0 := \{0\}$ 对应的 0 阶行列式或曰 "空行列式" 应规定为 1.

练习 5.4.7　验证 $n = 2, 3$ 时定义–命题 5.4.6 给出熟知的 2 阶和 3 阶行列式. 验证 $n = 1$ 时定义化简为 $\det(a) = a$, 其中 $a \in F$.

我们即将介绍行列式按行或列的展开, 为此需要余子式的概念.

定义 5.4.8　设 $1 \leq i, j \leq n$, 其中 $n \in \mathbb{Z}_{\geq 1}$. 定义矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 的第 (i,j) 个**余子式**为从 $\boldsymbol{A}$ 删去第 i 行与第 j 列, 余下的 $(n-1)\times(n-1)$ 矩阵 $\boldsymbol{\mathcal{M}}_{ij}$ 的行列式, 记为 $M_{ij} := \det \boldsymbol{\mathcal{M}}_{\boldsymbol{ij}} \in F$.

当 $n = 1$ 时, $\boldsymbol{A}$ 的唯一余子式按定义是空行列式, 已经规定为 1.

举 $n = 3$ 的情形为例, 第 $(2,2)$ 个余子式的取法如下所示.

$$\boldsymbol{A} = \begin{matrix} \\ i=2 \\ \\ \\ \end{matrix}\!\!\begin{pmatrix} a_{11} & \overset{\displaystyle\vdots}{a_{12}} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & \underset{\displaystyle\vdots}{a_{32}} & a_{33} \end{pmatrix}\!\!, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}.$$
$$\qquad\qquad j=2$$

定理 5.4.9　设 $n \in \mathbb{Z}_{\geq 1}$ 选定, $n \times n$ 矩阵的行列式具有以下性质.

▷ **在单位矩阵的取值**　我们有 $\det(\mathbf{1}_{n\times n}) = 1$.

▷ **转置不变性**　设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 则

$$\det \boldsymbol{A} = \det\left({}^{\mathrm{t}}\boldsymbol{A}\right).$$

▷ **行/列的置换**　设 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$. 若 $\sigma \in \mathfrak{S}_n$ 而 $\boldsymbol{B} = (b_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$ 由下式确定:

$$b_{i,j} = a_{\sigma(i),j} \quad (\text{行的置换}),$$

或

$$b_{i,j} = a_{i,\sigma(j)} \quad (\text{列的置换}),$$

则 $\det \boldsymbol{B} = \operatorname{sgn}(\sigma) \det \boldsymbol{A}$.

▷ **对每一行/列的线性** 对所有 $1 \le i \le n$, 以下性质成立:

$$\begin{vmatrix} \vdots & & \vdots \\ a_{i1}+a'_{i1} & \cdots & a_{in}+a'_{in} \\ \vdots & & \vdots \end{vmatrix} = \begin{vmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{vmatrix} + \begin{vmatrix} \vdots & & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & & \vdots \end{vmatrix},$$

$$\begin{vmatrix} \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{in} \\ \vdots & & \vdots \end{vmatrix} = c \begin{vmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{vmatrix}$$

等号两边除第 i 行以外的矩阵元皆相同, $c \in F$.

类似地, 对所有 $1 \le j \le n$,

$$\begin{vmatrix} \cdots & a_{1j}+a'_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj}+a'_{nj} & \cdots \end{vmatrix} = \begin{vmatrix} \cdots & a_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj} & \cdots \end{vmatrix} + \begin{vmatrix} \cdots & a'_{1j} & \cdots \\ & \vdots & \\ \cdots & a'_{nj} & \cdots \end{vmatrix},$$

$$\begin{vmatrix} \cdots & ca_{1j} & \cdots \\ & \vdots & \\ \cdots & ca_{nj} & \cdots \end{vmatrix} = c \begin{vmatrix} \cdots & a_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj} & \cdots \end{vmatrix}$$

等号两边除第 j 列以外的矩阵元皆相同, $c \in F$.

▷ **交错性** 若 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 有两行 (或列) 相同, 则 $\det \boldsymbol{A} = 0$.

▷ **按行/列和余子式展开** 对任意 $1 \le i, j \le n$, 任意 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 的行列式可以按第 i 行作展开

$$\det \boldsymbol{A} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik}$$

或者按第 j 列作展开

$$\det \boldsymbol{A} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} M_{kj},$$

其中的 M_{ik} 如定义 5.4.8.

证明 依定义, 行列式在单位矩阵的取值等于 $D_{\mathrm{e}}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = 1$.

对转置的不变性无非是定义–命题 5.4.6 最后一则等式的内容. 基于这一对称性, 以下性质仅须对行或列的情形择一验证即可.

列的置换性质是 $\det \boldsymbol{A} = D_{\mathrm{e}}(\boldsymbol{A}\boldsymbol{e}_1, \ldots, \boldsymbol{A}\boldsymbol{e}_n)$ 和引理 5.3.2 的直接结论. 同理, 对每一列的线性性质 (或交错性质) 转译为定义 5.3.1 的条件 **D.1** (或 **D.2**).

接着处理余子式展开. 空行列式等于 1, 所以 $n=1$ 情形是平凡的. 以下均假定 $n\geqslant 2$.

首务是确立对第一列的展开. 我们有

$$D_{\mathbf{e}}(\boldsymbol{A}\boldsymbol{e}_1,\dots,\boldsymbol{A}\boldsymbol{e}_n)=\sum_{i=1}^{n}a_{i1}\underbrace{D_{\mathbf{e}}(\boldsymbol{e}_i,\boldsymbol{A}\boldsymbol{e}_2,\dots,\boldsymbol{A}\boldsymbol{e}_n)}_{=:D_i}.$$

且先固定 $1\leqslant i\leqslant n$. 为了计算 D_i, 同样将向量 $\boldsymbol{A}\boldsymbol{e}_2,\dots,\boldsymbol{A}\boldsymbol{e}_n$ 按有序基 $\mathbf{e}$ 展开, 但排头的 $\boldsymbol{e}_i$ 导致这些展开式中含 $\boldsymbol{e}_i$ 的部分对 D_i 无贡献. 考虑 $n-1$ 个线性无关向量

$$\boldsymbol{e}_1,\dots,\widehat{\boldsymbol{e}_i},\dots\boldsymbol{e}_n,$$

符号 $\widehat{\boldsymbol{e}_i}$ 代表此项略去, 记以它们为基的子空间为 V_i, 并且记 V_i 的此组有序基为 $\mathbf{e}[i]$. 对每个 $2\leqslant j\leqslant n$, 命

$$\boldsymbol{f}_j:=\boldsymbol{A}\boldsymbol{e}_j-a_{ij}\boldsymbol{e}_i\ \in V_i.$$

因此 $D_i=D_{\mathbf{e}}(\boldsymbol{e}_i,\boldsymbol{f}_2,\dots,\boldsymbol{f}_n)$.

现在考虑映射

$$\begin{aligned}E_i:V_i^{n-1}&\longrightarrow F\\(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-1})&\longmapsto D_{\mathbf{e}}(\boldsymbol{e}_i,\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-1}),\end{aligned}$$

留意到 E_i 从 $D_{\mathbf{e}}$ 继承定义 5.3.1 的条件 **D.1** 和 **D.2**, 故 $E_i\in\mathcal{D}_{V_i}$. 基于例 5.1.9 和引理 5.3.2, 我们有

$$\begin{aligned}E_i(\boldsymbol{e}_1,\dots,\widehat{\boldsymbol{e}_i},\dots,\boldsymbol{e}_n)&=D_{\mathbf{e}}\left(\boldsymbol{e}_i,\boldsymbol{e}_1,\dots,\widehat{\boldsymbol{e}_i},\dots,\boldsymbol{e}_n\right)\\&=(-1)^{i+1}D_{\mathbf{e}}(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)=(-1)^{i+1},\end{aligned}$$

符号 $\widehat{\boldsymbol{e}_i}$ 仍代表该项略去. 配合定理 5.3.5 便得到 $E_i=(-1)^{i+1}D_{\mathbf{e}[i]}$. 代入之前的计算, 得到

$$\begin{aligned}D_i=E_i(\boldsymbol{f}_2,\dots,\boldsymbol{f}_n)=(-1)^{i+1}\det\boldsymbol{\mathcal{M}}_{i1}=(-1)^{i+1}M_{i1},\\\det\boldsymbol{A}=\sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{i1},\end{aligned}$$

这是因为相对于有序基 $\mathbf{e}[i]$, 列向量 $\boldsymbol{f}_2,\dots,\boldsymbol{f}_n$ 构成的矩阵正是定义 5.4.8 中的 $\boldsymbol{\mathcal{M}}_{i1}$. 此即行列式对第一列的展开.

接着推导对第 j 列的展开, $1<j\leqslant n$. 办法是先以形如

$$\sigma=\begin{pmatrix}1&2&\cdots&j&j+1&\cdots&n\\j&1&\cdots&j-1&j+1&\cdots&n\end{pmatrix}\in\mathfrak{S}_n$$

的置换将第 j 列搬到第 1 列, 得到的矩阵记为 $\boldsymbol{A}'$, 从而由 $\operatorname{sgn}(\sigma)=(-1)^{j+1}$ (例 5.1.9) 可得 $\det \boldsymbol{A}'=(-1)^{j+1}\det \boldsymbol{A}$. 根据上一步,

$$\det \boldsymbol{A}=(-1)^{j+1}\det \boldsymbol{A}' \overset{\text{对第一列展开}}{=\!=} \sum_{k=1}^{n}(-1)^{k+1+j+1}a'_{k1}M'_{k1}.$$

然而 $a'_{k1}=a_{kj}$ 而 $M'_{k1}=M_{kj}$ (置换 σ 不改变这些余子式的列序), 此即所求展开. □

按行或列展开对于计算包含较多零元的行列式往往是很方便的.

推论 5.4.10 对任意 $1\leq i\neq j\leq n$ 和 $\boldsymbol{A}=(a_{ij})_{i,j}\in \mathrm{M}_{n\times n}(F)$, 皆有

$$\sum_{k=1}^{n}(-1)^{k+j}a_{ik}M_{jk}=0,$$
$$\sum_{k=1}^{n}(-1)^{k+j}a_{ki}M_{kj}=0.$$

证明 基于行和列的对称性, 以下仅证第二条等式. 注意到 $\sum_{k=1}^{n}(-1)^{k+j}a_{ki}M_{kj}$ 无关 $\boldsymbol{A}$ 的第 j 列. 定义 $\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{B}$ 的第 j 列等于 $\boldsymbol{A}$ 的第 i 列, 其余诸列和 $\boldsymbol{A}$ 相同. 将 $\det \boldsymbol{B}$ 按第 j 列展开, 便是

$$\det \boldsymbol{B}=\sum_{k=1}^{n}(-1)^{k+j}a_{ki}M_{kj}.$$

然而 $\boldsymbol{B}$ 有两列相同, 交错性质遂表明 $\det \boldsymbol{B}=0$. 证毕. □

练习 5.4.11 (定向) 现在可以简单地处理 §5.3 中尚未定义的定向. 设 V 为 $\mathbb{R}$-向量空间, $n:=\dim V\in \mathbb{Z}_{\geq 1}$. 在集合

$$\mathrm{OrdBs}:=\{(v_1,\dots,v_n):V \text{ 的有序基}\}$$

上定义二元关系

$$(v_1,\dots,v_n)\sim(v'_1,\dots,v'_n)\iff \det \boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}>0,$$

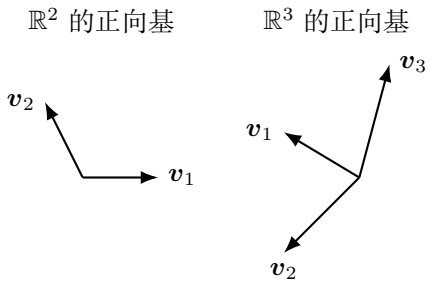
其中 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}\in \mathrm{M}_{n\times n}(\mathbb{R})$ 是注记 4.9.2 中的转换矩阵.

(i) 验证 $\sim$ 是 OrdBs 上的等价关系.

(ii) 说明 OrdBs 被 $\sim$ 划分为两个等价类.

(iii) 任何一个 $\sim$ 等价类 $P\subset \mathrm{OrdBs}$ 都称为 V 的一个**定向**; 我们约定属于 P 的有序基为正向, 反之为负向. 说明如此定义的定向确实具有 §5.2 要求的性质.

作为推论, 任何有序基 $v_1,\dots,v_n$ 都唯一确定 V 的一个定向, 使得 $(v_1,\dots,v_n)$ 为正向. 对于 $V=\mathbb{R}^n$ 及其标准基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$, 相应的定向称为 $\mathbb{R}^n$ 的标准定向; 以 $\mathbb{R}^2$ 为例, 其中的正向基是逆时针排列的基, 而 $\mathbb{R}^3$ 的正向基则由右手定则刻画, 如下图所示.



5.5 一些特殊行列式

若 $\sigma \in \mathfrak{S}_n$, 而矩阵 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$ 由

$$a_{ij} = \begin{cases} 1, & i = \sigma(j), \\ 0, & i \neq \sigma(j) \end{cases}$$

确定, 则称之为 σ 对应的**置换矩阵**, 记作 $\boldsymbol{P}_\sigma$. 它的每一行 (或每一列) 都恰好有一个非零矩阵元, 即 1. 作为从 F^n 到 F^n 的线性映射, 它在标准基上的作用是

$$\boldsymbol{P}_\sigma : \boldsymbol{e}_j \mapsto \boldsymbol{e}_{\sigma(j)}.$$

既然 $i = \sigma(j)$ 等价于 $j = \sigma^{-1}(i)$, 取转置就得到 ${}^{\mathrm{t}}(\boldsymbol{P}_\sigma) = \boldsymbol{P}_{\sigma^{-1}}$.
从上述定义即刻看出 $\boldsymbol{P}_{\mathrm{id}} = \boldsymbol{1}_{n\times n}$. 此外, 置换的合成反映为置换矩阵的乘法.

命题 5.5.1 *若 $\sigma, \tau \in \mathfrak{S}_n$, 则 $\boldsymbol{P}_{\sigma\tau} = \boldsymbol{P}_\sigma \boldsymbol{P}_\tau$.*

证明 依照线性映射的视角, 对所有 $1 \leq j \leq n$ 都有 $\boldsymbol{P}_\sigma \boldsymbol{P}_\tau \boldsymbol{e}_j = \boldsymbol{P}_\sigma \boldsymbol{e}_{\tau(j)} = \boldsymbol{e}_{\sigma\tau(j)}$, 而 $\boldsymbol{P}_{\sigma\tau}(\boldsymbol{e}_j) = \boldsymbol{e}_{\sigma\tau(j)}$. 既然线性映射由基上的取值确定, 故 $\boldsymbol{P}_{\sigma\tau} = \boldsymbol{P}_\sigma \boldsymbol{P}_\tau$. □

特别地, $(\boldsymbol{P}_\sigma)^{-1} = \boldsymbol{P}_{\sigma^{-1}}$. 现在来确定置换矩阵的行列式.

命题 5.5.2 *设 $\sigma \in \mathfrak{S}_n$, 则 $\det \boldsymbol{P}_\sigma = \mathrm{sgn}(\sigma)$.*

证明 按照定义–命题 5.4.6 和以上描述,

$$\det \boldsymbol{P}_\sigma = D_{\mathbf{e}}(\boldsymbol{P}_\sigma \boldsymbol{e}_1, \ldots, \boldsymbol{P}_\sigma \boldsymbol{e}_n) = D_{\mathbf{e}}\left(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)}\right).$$

代入引理 5.3.2 可得

$$\det \boldsymbol{P}_\sigma = \mathrm{sgn}(\sigma^{-1}) D_{\mathbf{e}}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = \mathrm{sgn}(\sigma^{-1}).$$

然而 $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$. □

命题 5.5.3 设 $\boldsymbol{A}$ 是上三角或下三角矩阵 (定义 4.11.9), 则

$$\det \boldsymbol{A} = \prod_{i=1}^{n} a_{ii}.$$

证明 既然转置不改变行列式, 处理上三角情形即可. 由于 $a_{21} = \cdots = a_{n1} = 0$, 将 $\det \boldsymbol{A}$ 按第一列展开可得

$$\det \boldsymbol{A} = a_{11} M_{11}.$$

然而去掉第一行和第一列得到的子矩阵 $\boldsymbol{\mathcal{M}_{11}}$ 仍是上三角, 以 $a_{22}, \ldots, a_{nn}$ 为对角元, 因此可以递归地处理.

另一种方法则是直接代公式 $\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$. 兹断言对于上三角的 $\boldsymbol{A}$, 使乘积 $a_{1\sigma(1)} \cdots a_{n\sigma(n)} \neq 0$ 的唯一可能是 $\sigma = \mathrm{id}$. 缘由如下: 为使乘积非零, 必然对所有 i 皆有 $\sigma(i) \geq i$. 代入 $i = n$ 得到 $\sigma(n) = n$; 代入 $i = n-1$ 则有 $\sigma(n-1) \in \{n-1, n\}$, 但值 n 已取过, 故 $\sigma(n-1) = n-1$. 依此类推可见 $\sigma = \mathrm{id}$. □

算法 5.5.4 一般 $n \times n$ 矩阵的行列式通过 Gauss–Jordan 消元法化约到上三角情形计算. 具体地说, 我们在 §1.3 将初等行变换分为三类:

- ⋆ 记为 $\mathrm{A}(i,k,c)$ 的变换将第 i 行乘以 c 加到第 k 行, 这不改变行列式;
- ⋆ 记为 $\mathrm{B}(i,k)$ 的变换将第 i 行和第 k 行对换, 这将行列式变号;
- ⋆ 记为 $\mathrm{C}(i,c)$ 的变换将第 i 行乘以 c, 这将行列式乘以 c.

依此可以在 Gauss–Jordan 消元法的每一步追踪行列式的变化. 最终得到的行梯矩阵 $\boldsymbol{A}'$ 必然是上三角的, 因为它逐行缩进. 如果行梯矩阵的主元个数 $< n$, 则必有对角元为 0, 故 $\det \boldsymbol{A}' = \det \boldsymbol{A} = 0$. 如果主元有 n 个, 则 $\det \boldsymbol{A}' = 1$.

另一方面, 由命题 5.5.2 和 5.5.3 容易推得三种初等行变换对应的初等矩阵分别以 1, -1 和 c 为其行列式. 考虑到初等行变换对应到初等矩阵的左乘, 以及行列式的乘法性质 (定理 5.4.3), 这就以另一种方式说明了初等行变换如何影响行列式.

以 Gauss–Jordan 消元法计算 n 阶行列式所需的操作次数大致按 n^3 增长 (练习 1.3.8). 相比之下, 定义–命题 5.4.6 的公式则有 $n!$ 项. 计算效率高下立判.

另一类常用的特殊行列式是 Vandermonde 行列式.

命题 5.5.5 (Vandermonde 行列式) 设 $x_1, \ldots, x_n \in F$, 则

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq j < i \leq n} (x_i - x_j)$$

证明 注意到 $n = 1$ 的特例对应到一阶行列式 $|1| = 1$, 而右式的 "空积" 按惯例律定为 1 (域 F 的乘法幺元).

对于 $n>1$ 的情形, 我们将第 $n-1$ 行乘以 $-x_1$ 加到第 n 行上, 再将第 $n-2$ 行乘以 $-x_1$ 加到第 $n-1$ 行, 依此类推, 得到原行列式等于

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & x_3-x_1 & \cdots & x_n-x_1 \\ 0 & x_2^2-x_1x_2 & x_3^2-x_1x_3 & \cdots & x_n^2-x_1x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_2^{n-1}-x_1x_2^{n-2} & x_3^{n-1}-x_1x_3^{n-2} & \cdots & x_n^{n-1}-x_1x_n^{n-2} \end{vmatrix}$$

按第一列展开, 再从余子式 M_{11} 的第 $i-1$ 列提出 x_i-x_1, 其中 $i=2,\dots,n$, 可见上式等于

$$\begin{vmatrix} x_2-x_1 & \cdots & x_n-x_1 \\ x_2^2-x_1x_2 & \cdots & x_n^2-x_1x_n \\ \vdots & \ddots & \vdots \\ x_2^{n-1}-x_1x_2^{n-2} & \cdots & x_n^{n-1}-x_1x_n^{n-2} \end{vmatrix} = \prod_{i=2}^{n}(x_i-x_1)\cdot \begin{vmatrix} 1 & \cdots & 1 \\ x_2 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_2^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}.$$

右端的 $n-1$ 阶行列式无非是 $x_2,\dots,x_n$ 给出的 Vandermonde 行列式. 因此递归地得出

$$\prod_{i=2}^{n}(x_i-x_1)\prod_{2\leq j<i\leq n}(x_i-x_j)=\prod_{1\leq j<i\leq n}(x_i-x_j).$$

明所欲证. □

5.6 分块行列式

设 V 是带有直和分解 $V=V_1\oplus\cdots\oplus V_n$ 的有限维向量空间, 则任何 $T\in\mathrm{End}(V)$ 都按 (4.10.3) 的方式得到相应的分解 $(T_{ij})_{1\leq i,j\leq n}$, 由矩阵观点看便是分块. 我们希望研究 $\det T$ 和各个分块之间的关系.

回忆到 $T_{ij}\in\mathrm{Hom}(V_j,V_i)$.

命题 5.6.1 设 $T\in\mathrm{End}(V)$ 相对于 V 的直和分解是上三角或下三角线性变换, 则

$$\det T=\prod_{i=1}^{n}\det T_{ii}.$$

证明 这一性质以矩阵观点处理是比较简便的. 由于转置不改变行列式, 处理上三角情

形即可. 我们的目标相当于证明

$$\det\left(\begin{array}{c|c|c} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1n} \\ \hline \boldsymbol{0} & \ddots & \vdots \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{nn} \end{array}\right) = \prod_{i=1}^{n} \det \boldsymbol{A}_{ii},$$

其中每个 $\boldsymbol{A}_{ii}$ 都是方阵. 基于 "块中有块" 的观察 (请回忆推论 4.11.8 证明), 问题可以逐步化到 $n=2$. 今后不妨设所论的分块上三角矩阵为

$$\begin{aligned} \boldsymbol{A} &= (a_{ij})_{1\leq i,j\leq n_1+n_2} \\ &= \left(\begin{array}{c|c} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \hline \boldsymbol{0} & \boldsymbol{A}_{22} \end{array}\right), \\ \boldsymbol{A}_{11} &\in \mathrm{M}_{n_1\times n_1}(F),\ \boldsymbol{A}_{22} \in \mathrm{M}_{n_2\times n_2}(F). \end{aligned}$$

按定义, $\det \boldsymbol{A} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n_1+n_2} a_{i,\sigma(i)}$. 对于给定的 σ, 为了使对应的连乘积非零, $\boldsymbol{A}$ 的样貌说明必要条件是对于每个 i,

$$n_1 < i \leq n_1+n_2 \implies n_1 < \sigma(i) \leq n_1+n_2.$$

但是这等价于说 σ 限制为 $\{n_1+1,\dots,n_1+n_2\}$ 上的置换 (记为 σ_2), 从而也限制为 $\{1,\dots,n_1\}$ 上的置换 (记为 σ_1).

利用平移给出的双射 $\{1,\dots,n_2\} \xrightarrow{1:1} \{n_1+1,\dots,n_1+n_2\}$, 可将 σ_2 等同于 $\mathfrak{S}_{n_2}$ 的元素. 现在观察到若 σ 按此方式对应到 $(\sigma_1,\sigma_2)\in\mathfrak{S}_{n_1}\times\mathfrak{S}_{n_2}$, 则

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2),$$

这是因为 σ 的逆序发生在两个无交子集的内部, 不相干涉. 综上可得

$$\begin{aligned} \det \boldsymbol{A} &= \sum_{(\sigma_1,\sigma_2)\in\mathfrak{S}_{n_1}\times\mathfrak{S}_{n_2}} \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2) \prod_{i=1}^{n_1} a_{i,\sigma_1(i)} \prod_{i=1}^{n_2} a_{n_1+i,n_1+\sigma_2(i)} \\ &= \left(\sum_{\sigma_1} \operatorname{sgn}(\sigma_1) \prod_{i=1}^{n_1} a_{i,\sigma_1(i)}\right)\left(\sum_{\sigma_2} \operatorname{sgn}(\sigma_2) \prod_{i=1}^{n_2} a_{n_1+i,n_1+\sigma_2(i)}\right). \end{aligned}$$

此即 $\det \boldsymbol{A}_{11} \det \boldsymbol{A}_{22}$. □

兹介绍分块行列式计算的一些经典例子.

命题 5.6.2 设 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$, $\boldsymbol{B}\in\mathrm{M}_{n\times m}(F)$, 则

$$\det\left(\begin{array}{c|c} \boldsymbol{1}_{n\times n} & \boldsymbol{B} \\ \hline \boldsymbol{A} & \boldsymbol{1}_{m\times m} \end{array}\right) = \det\left(\boldsymbol{1}_{m\times m} - \boldsymbol{A}\boldsymbol{B}\right).$$

证明 分块矩阵的乘法给出

$$\left(\begin{array}{c|c} \mathbf{1}_{n\times n} & \mathbf{0}_{n\times m} \\ \hline -\boldsymbol{A} & \mathbf{1}_{m\times m} \end{array}\right)\left(\begin{array}{c|c} \mathbf{1}_{n\times n} & \boldsymbol{B} \\ \hline \boldsymbol{A} & \mathbf{1}_{m\times m} \end{array}\right)=\left(\begin{array}{c|c} \mathbf{1}_{n\times n} & \boldsymbol{B} \\ \hline \mathbf{0}_{m\times n} & \mathbf{1}_{m\times m}-\boldsymbol{AB} \end{array}\right);$$

这也可以看作将断言中的矩阵作分块初等行变换 (将第一行左乘以 $-\boldsymbol{A}$ 加至第二行). 对两边同取行列式, 再运用乘性定理 5.4.3 即可. □

推论 5.6.3 设 $\boldsymbol{A}\in \mathrm{M}_{m\times n}(F)$, $\boldsymbol{B}\in \mathrm{M}_{n\times m}(F)$, 则

$$\det\left(\mathbf{1}_{m\times m}-\boldsymbol{AB}\right)=\det\left(\mathbf{1}_{n\times n}-\boldsymbol{BA}\right).$$

证明 在命题 5.6.2 中调换 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的角色, 可得

$$\det\left(\begin{array}{c|c} \mathbf{1}_{m\times m} & \boldsymbol{A} \\ \hline \boldsymbol{B} & \mathbf{1}_{n\times n} \end{array}\right)=\det\left(\mathbf{1}_{n\times n}-\boldsymbol{BA}\right).$$

然而若按线性映射的语言, 将 F^{n+m} 的标准有序基从

$$\boldsymbol{e}_1,\dots,\boldsymbol{e}_n,\boldsymbol{e}_{n+1},\dots\boldsymbol{e}_{n+m}$$

重排为

$$\boldsymbol{e}_{m+1},\dots,\boldsymbol{e}_{m+n},\boldsymbol{e}_1,\dots,\boldsymbol{e}_m,$$

则 $\left(\begin{array}{c|c} \mathbf{1}_{m\times m} & \boldsymbol{A} \\ \hline \boldsymbol{B} & \mathbf{1}_{n\times n} \end{array}\right)$ 所代表的线性映射对新的有序基便表为

$$\left(\begin{array}{c|c} \mathbf{1}_{n\times n} & \boldsymbol{B} \\ \hline \boldsymbol{A} & \mathbf{1}_{m\times m} \end{array}\right),$$

其行列式已知是 $\det\left(\mathbf{1}_{m\times m}-\boldsymbol{AB}\right)$. 对不同的有序基计算线性映射的行列式给出相等结果, 所求等式因此成立. □

5.7 Cramer 法则

行列式的主要应用之一是求矩阵的逆. 我们首先介绍一则抽象的结果: 在域 F 上, 矩阵或线性变换的可逆性可以由行列式判断.

命题 5.7.1 设 V 为有限维 F-向量空间, 则 $T\in\mathrm{End}(V)$ 可逆当且仅当 $\det T\in F^\times$.

证明 设 T 可逆, 则命题 5.4.3 蕴涵 $\det T$ 可逆. 反之设 T 不可逆, 则 T 的秩严格小于 $n:=\dim V$. 以下排除 $n=0$ 的平凡情形. 对于任意 n 个向量 $v_1,\dots,v_n\in V$, 必有

$$Tv_1,\dots,Tv_n\ \in\mathrm{im}(T)\quad \text{线性相关}.$$

根据交错形式的一般性质 (5.3.1), 对任意 $D \in \mathcal{D}_V$ 都有

$$D(Tv_1, \dots, Tv_n) = 0.$$

根据行列式作为比例常数的定义, 上式蕴涵 $\det T = 0$. □

推论 5.7.2 设 $v_1, \dots, v_n$ 是 n 维 F-向量空间 V 的元素 ($n \in \mathbb{Z}_{\geq 1}$), 则以下性质等价:

(i) $v_1, \dots, v_n$ 线性相关;

(ii) 对所有 $D \in \mathcal{D}_V$ 皆有 $D(v_1, \dots, v_n) = 0$;

(iii) 存在 V 的有序基 $e_1, \dots, e_n$, 记为 $\mathbf{e}$, 使得 $D_{\mathbf{e}}(v_1, \dots, v_n) = 0$.

证明 (i) $\Longrightarrow$ (ii) 是交错形式的一般性质 (5.3.1). (ii) $\Longrightarrow$ (iii) 属显然. 至于 (iii) $\Longrightarrow$ (i), 定义 $T \in \mathrm{End}(V)$ 使得 $Te_i = v_i$, 则

$$\det T = D_{\mathbf{e}}(Te_1, \dots, Te_n) = D_{\mathbf{e}}(v_1, \dots, v_n) = 0,$$

因此 T 不可逆. 推论 4.8.5 蕴涵 T 非单, 所以存在不全为 0 的 $x_1, \dots, x_n \in F$ 使得 $\sum_{i=1}^n x_i v_i = 0$. □

若矩阵或线性变换可逆, 如何写出求逆的精确公式? 为了具体地回答这一问题, 以下仅针对矩阵进行计算. 答案实质上就是线性方程组理论中的 Cramer 法则. 今起选定 $n \in \mathbb{Z}_{\geq 1}$.

定义 5.7.3 设 $n \in \mathbb{Z}_{\geq 1}$. 对于 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$, 其**经典伴随矩阵**定义为

$$\boldsymbol{A}^\vee = (A_{ji})_{i,j} \in \mathrm{M}_{n\times n}(F),$$

其中

$$A_{ij} := (-1)^{i+j} M_{ij}, \quad 1 \leq i, j \leq n,$$

而 M_{ij} 是定义 5.4.8 中的 $\boldsymbol{A}$ 的余子式.

当 $n = 1$ 时, $\boldsymbol{A}^\vee$ 的定义退化为 $\boldsymbol{A}^\vee = 1$, 这是因为 0 阶行列式规定为 1. 其次, $n = 2$ 的情形具体写作

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\vee = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

推论 5.8.10 将说明如何为线性映射 T 内蕴地定义 $T^\vee$, 现阶段只考虑矩阵.

定理 5.7.4 对任意 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 皆有

$$\boldsymbol{A}\boldsymbol{A}^\vee = \det \boldsymbol{A} \cdot \mathbf{1}_{n\times n} = \boldsymbol{A}^\vee \boldsymbol{A}.$$

证明 按定义计算 $\boldsymbol{A}\boldsymbol{A}^\vee$ 的 (i,j) 矩阵元为

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \sum_{k=1}^{n} a_{ik}(-1)^{k+j}M_{jk}.$$

综合定理 5.4.9 和推论 5.4.10 可得

$$\sum_{k=1}^{n}(-1)^{k+j}a_{ik}M_{jk} = \begin{cases} \det \boldsymbol{A}, & i=j, \\ 0, & i\neq j. \end{cases}$$

因此 $\boldsymbol{A}\boldsymbol{A}^\vee = \det \boldsymbol{A}\cdot \mathbf{1}_{n\times n}$.

以类似方法计算 $\boldsymbol{A}^\vee\boldsymbol{A}$ 的 (j,i) 矩阵元, 结果是

$$\begin{aligned}\sum_{k=1}^{n} A_{kj}a_{ki} &= \sum_{k=1}^{n}(-1)^{k+j}M_{kj}a_{ki} \\ &= \begin{cases} \det \boldsymbol{A}, & i=j, \\ 0, & i\neq j. \end{cases}\end{aligned}$$

因此 $\boldsymbol{A}^\vee\boldsymbol{A} = \det \boldsymbol{A}\cdot \mathbf{1}_{n\times n}$. □

命题 5.7.1 说明域 F 上的矩阵可逆等价于行列式非零, 因此定理 5.7.4 即刻给出以下的精确求逆公式.

推论 5.7.5 *若 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 满足 $\det \boldsymbol{A}\in F^\times$, 则命*

$$\boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1}\boldsymbol{A}^\vee.$$

在 $n=2$ 情形, 上式具体写作

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

推论 5.7.6 (Cramer 法则) *考虑域 F 上的 n 元线性方程组*

$$\begin{aligned} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ &\vdots \qquad\quad \vdots \\ a_{n1}X_1 + \cdots + a_{nn}X_n &= b_n. \end{aligned}$$

将其系数矩阵 $\boldsymbol{A} = (a_{ij})_{i,j}\in \mathrm{M}_{n\times n}(F)$ 视作从 F^n 到 F^n 的线性映射.

(i) *如果方程组齐次, 亦即 $b_1 = \cdots = b_n = 0$, 则其解集等于 $\ker \boldsymbol{A}$ (见定义 4.8.1). 特别地, 此时方程组有不全为零的解当且仅当 $\det \boldsymbol{A} = 0$.*

(ii) 当 $(b_1,\dots,b_n)\in F^n$ 给定, 方程组或者无解, 或者其解集形如

$$(x_1,\dots,x_n)+\ker \boldsymbol{A},$$

其中 $(x_1,\dots,x_n)\in F^n$ 是方程的任何一组解.

(iii) 设 $\det \boldsymbol{A}\in F^\times$, 则对任何 $(b_1,\dots,b_n)\in F^n$, 方程组都有唯一解 $(x_1,\dots,x_n)$, 其中

$$x_j=\frac{\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n}\\ a_{21} & \cdots & b_2 & \cdots & a_{2n}\\ \vdots & & \vdots & & \vdots\\ a_{n1} & \cdots & b_n & \cdots & a_{nn}\end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n}\\ a_{21} & \cdots & a_{2i} & \cdots & a_{2n}\\ \vdots & & \vdots & & \vdots\\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn}\end{vmatrix}}\overset{\text{第 } j \text{ 列}}{},\quad j=1,\dots,n.$$

证明 鉴于此前的一系列讨论, 并且回忆到 $\ker(\boldsymbol{A})=\{0\}$ 等价于 $\boldsymbol{A}$ 可逆 (命题 4.8.13), 断言 (i) 和 (ii) 不外是复述 §4.1 和 §4.8 的基本内容. 对于 (iii), 我们有

$$(x_1,\dots,x_n)\text{ 是解}\iff \boldsymbol{A}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\iff \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\boldsymbol{A}^{-1}\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}.$$

然而 $\boldsymbol{A}^{-1}=(\det \boldsymbol{A})^{-1}\boldsymbol{A}^\vee$ (推论 5.7.5) 表明列向量

$$\boldsymbol{A}^{-1}\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$$

的第 j 个坐标等于

$$(\det \boldsymbol{A})^{-1}\sum_{k=1}^{n}A_{kj}b_k=(\det \boldsymbol{A})^{-1}\sum_{k=1}^{n}(-1)^{k+j}b_kM_{kj},$$

而定理 5.4.9 说明 $\sum_{k=1}^{n}(-1)^{k+j}b_kM_{kj}$ 是按第 j 列展开行列式

$$\begin{vmatrix} a_{11} & \cdots & \overset{\text{第 } j \text{ 列}}{b_1} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}$$

的产物. 明所欲证. □

尽管 Cramer 法则在系数矩阵可逆的情况下给出线性方程组的精确解. 但实用中行列式的计算量将随 n 而迅速增长. 纵使有时能高效地计算这些行列式, Cramer 法则在数值运算上依然不稳定: 因为它涉及取商, 当分母 $\det \boldsymbol{A}$ 接近 0 时, 一点微小的扰动都可能造成 $x_1,\dots,x_n$ 的大幅改变. 因此 Cramer 法则的价值主要在于理论层面.

练习 5.7.7 设 $n \geq 2$ 而 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 证明

$$\operatorname{rk}(\boldsymbol{A}^\vee) = \begin{cases} n, & \operatorname{rk}(\boldsymbol{A}) = n, \\ 1, & \operatorname{rk}(\boldsymbol{A}) = n-1, \\ 0, & \operatorname{rk}(\boldsymbol{A}) < n-1. \end{cases}$$

提示 〉 可逆情形 $\operatorname{rk}(\boldsymbol{A}) = n$ 是容易的. 当 $\operatorname{rk}(\boldsymbol{A}) < n-1$ 时, 每个余子式里的列向量都线性相关, 因而余子式全为 0. 当 $\operatorname{rk}(\boldsymbol{A}) = n-1$ 时, 先以算法 4.4.8 取出 $n-1$ 个线性无关的列, 再以引理 4.9.13 说明存在非零余子式, 从而 $\operatorname{rk}(\boldsymbol{A}^\vee) \geq 1$; 为了说明此时 $\operatorname{rk}(\boldsymbol{A}^\vee) = 1$, 代入例 4.8.8 来推导

$$\operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{A}^\vee) \leq n.$$

练习 5.7.8 证明 $\det(\boldsymbol{A}^\vee) = (\det \boldsymbol{A})^{n-1}$.

练习 5.7.9 证明转置与经典伴随可以交换顺序: $({}^{\mathrm{t}}\boldsymbol{A})^\vee = {}^{\mathrm{t}}(\boldsymbol{A}^\vee)$.

提示 〉 可尝试从定义证明, 但这也是稍后的定理 5.8.9 的直接结论.

5.8 特征多项式和 Cayley–Hamilton 定理

以下需要任意域 F 上的多项式环 $F[X]$ 和有理函数域 $F(X)$ 的语言, 符号 X 代表变元. 详见 §§3.3—3.5.

首先介绍如何将线性映射代入多项式来求值. 设 V 为 F-向量空间, $T \in \operatorname{End}(V)$. 回忆到 $\operatorname{End}(V)$ 成环, 因此对所有 $k \in \mathbb{Z}_{\geq 0}$ 皆可定义 $T^k \in \operatorname{End}(V)$; 特别地, $T^0 := \mathrm{id}_V$. 对于任意多项式

$$f = \sum_{k=0}^{m} a_k X^k \in F[X],$$

代入 T 的产物是

$$f(T) := \sum_{k=0}^{m} a_k T^k \ \in \mathrm{End}(V). \tag{5.8.1}$$

根据多项式加法和乘法的定义, 易见

$$(fg)(T) = f(T)g(T), \quad (f+g)(T) = f(T)+g(T), \quad (tf)(T) = t\cdot f(T),$$
$$f, g \in F[X],\ t \in F.$$

现在假设 V 是有限维的, $n := \dim V$. 且从对 T 求逆的问题入手, 考虑 $\mathrm{End}(V)$ 的 n^2+1 个元素

$$T^0, \dots, T^{n^2}.$$

因为 $\dim \mathrm{End}(V) = \dim \mathrm{M}_{n\times n}(F) = n^2$, 它们必然线性相关. 换言之存在 $c_0, \dots, c_{n^2} \in F$, 不全为 0, 使得

$$\sum_{k=0}^{n^2} c_k T^k = 0_V.$$

特别地, 我们得知存在非零多项式 $f \in F[X]$ 使得

$$\deg f \le n^2, \quad f(T) = 0_V.$$

若 T 可逆, 则我们可进一步取 f 使得 $c_0 \neq 0$, 原因如下. 设若 $c_0 = 0$, 则从 $T\left(\sum_{k=1}^{m} c_k T^{k-1}\right) = 0_V$ 两边消去 T, 可得 $c_1 \mathrm{id}_V + \cdots + c_m T^{m-1} = 0_V$, 如是降次直到常数项非零为止. 这便导致如下的存在性结果.

命题 5.8.1 设 $T \in \mathrm{End}(V)$ 可逆, 则存在非零多项式 $g \in F[X]$ 使得 $T^{-1} = g(T)$.

证明 根据先前讨论, 取 $f = \sum_{k=0}^{m} c_0 X^k$ 使得 $f(T) = 0_V$ 而 $c_0 \neq 0$. 于是

$$T\left(\sum_{k=1}^{m} c_k T^{k-1}\right) = -c_0 \cdot \mathrm{id}_V = \left(\sum_{k=1}^{m} c_k T^{k-1}\right) T.$$

令 $g := -c_0^{-1} \sum_{k=1}^{m} c_k X^{k-1} \in F[X]$ 即有 $g(T) = T^{-1}$. □

尽管命题 5.8.1 的结论显得抽象, 它经常能给出实用的信息. 举例明之, 设 T 相对于给定的直和分解 $V = V_1 \oplus \cdots \oplus V_n$ 是分块上三角 (或下三角, 对角) 的, 而且 T 可逆, 则 T^{-1} 也必然是分块上三角 (或下三角, 对角) 的. 原因如下: 取 $g \in F[X]$ 使得 $T^{-1} = g(T)$, 则推论 4.11.7 蕴涵 $g(T)$ 和 T 同样是分块上三角 (或下三角, 对角) 的. 这就重新证明了推论 4.11.8 的后半部.

事实上, 命题 5.8.1 的论证还蕴涵可取 g 使得 $\deg g \le n^2$. 这一论证适用于更广泛的一类代数结构, 而对于眼下的情形, 对 g 可以有更精确也更经济的取法. 具体起见, 今后设 $V = F^n$, 将问题化到 $n\times n$ 矩阵的情形[2)].

[2)]之所以不直接对 $T \in \mathrm{End}(V)$ 定义特征多项式, 根本原因是我们暂时说不清如何将 T 变为一个 $F(X)$-线性映射, 这点需要张量积的理论; 矩阵情形则无此困扰, 因为一切都是具体的.

定义 5.8.2 (特征多项式) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 我们将 F 嵌入为有理函数域 $F(X)$ 的子域, 从而构造矩阵 $X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}\in\mathrm{M}_{n\times n}(F(X))$. 定义

$$\mathrm{Char}_{\boldsymbol{A}} := \det\left(X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}\right);$$

根据行列式的具体公式 (参考定义–定理 5.4.6), 可见 $\mathrm{Char}_{\boldsymbol{A}}\in F[X]$, 称之为 $\boldsymbol{A}$ 的特征多项式.

引入称为 "Kronecker 的 δ 符号" 的记法

$$\delta_{i,j} := \begin{cases} 1, & i=j,\\ 0, & i\neq j.\end{cases}$$

行列式的具体公式给出

$$\det\left(X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}\right)=\sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)\prod_{i=1}^{n}\left(\delta_{i,\sigma(i)}X-a_{i,\sigma(i)}\right),$$

它对 X 的最高次项来自 $\sigma=\mathrm{id}$ 的贡献, 给出 X^n, 故 $\deg\mathrm{Char}_{\boldsymbol{A}}=n$.

另一方面, 从每一行或每一列提出 -1 可见 $\det(-\boldsymbol{A})=(-1)^n\det\boldsymbol{A}$, 故特征多项式和行列式的关系为

$$\mathrm{Char}_{\boldsymbol{A}}(0)=\det(-\boldsymbol{A})=(-1)^n\det(\boldsymbol{A}).$$

今后将 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 的特征多项式表成

$$\mathrm{Char}_{\boldsymbol{A}}=X^n+c_{n-1}X^{n-1}+\cdots+c_0.$$

已知 $c_0=(-1)^n\det\boldsymbol{A}$. 事实上, 共轭的矩阵有相同的特征多项式, 故每个系数 $c_0,\ldots,c_{n-1}$ 都给出 $\boldsymbol{A}$ 的共轭不变量. 这是以下命题的内容.

命题 5.8.3 设 $\boldsymbol{P}\in\mathrm{M}_{n\times n}(F)$ 可逆, 则 $\mathrm{Char}_{\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}}=\mathrm{Char}_{\boldsymbol{A}}$.

证明 这是因为在 $\mathrm{M}_{n\times n}(F(X))$ 中有等式

$$X\cdot\mathbf{1}_{n\times n}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\left(X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}\right)\boldsymbol{P},$$

故两边的行列式相同. □

有鉴于此, 对任意有限维 F-向量空间 V 和 $T\in\mathrm{End}(V)$ 都能定义特征多项式 Char_T, 方法是取有序基 $v_1,v_2,\ldots$ 将 T 表成矩阵 $\boldsymbol{A}$, 然后命 $\mathrm{Char}_T:=\mathrm{Char}_{\boldsymbol{A}}$. 不同的有序基对应的矩阵差一个共轭, 故共轭不变性说明 Char_T 不依赖有序基的取法.

按照我们对行列式的规定, 零空间 $V=\{0\}$ 上的 $T=0_V=\mathrm{id}_V$ 满足 $\mathrm{Char}_T=1$; 不必过度操心这种极端情形.

先来介绍特征多项式的一些基本性质.

命题 5.8.4 转置不改变特征多项式: 对一切 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 都有 $\mathrm{Char}_{\boldsymbol{A}} = \mathrm{Char}_{{}^{\mathrm{t}}\boldsymbol{A}}$.

证明 运用 $X\cdot \mathbf{1}_{n\times n} - {}^{\mathrm{t}}\boldsymbol{A} = {}^{\mathrm{t}}(X\cdot \mathbf{1}_{n\times n} - \boldsymbol{A})$ 和行列式的转置不变性即可. □

命题 5.8.5 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 是分块上三角矩阵

$$\boldsymbol{A} = \left(\begin{array}{c|c|c} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1m} \\ \hline & \ddots & \vdots \\ \hline & & \boldsymbol{A}_{mm} \end{array}\right),$$

其中 $\boldsymbol{A}_{ii} \in \mathrm{M}_{n_i\times n_i}(F)$ 而 $n_1 + \cdots + n_m = n$. 此时

$$\mathrm{Char}_{\boldsymbol{A}} = \prod_{i=1}^{m} \mathrm{Char}_{\boldsymbol{A}_{ii}}.$$

对分块下三角矩阵同样有相应的性质.

证明 注意到 $X\cdot \mathbf{1}_{n\times n} - \boldsymbol{A}$ 也是分块上三角的, 以 $X\cdot \mathbf{1}_{n_i\times n_i} - \boldsymbol{A}_{ii}$ 为其对角分块, $i = 1,\ldots,m$. 于是可按命题 5.6.1 计算 $X\cdot \mathbf{1}_{n\times n} - \boldsymbol{A}$ 的行列式. 分块下三角情形的论证完全相同, 也可以取转置以相互过渡. □

因此, 对于带直和分解 $V = \bigoplus_{i=1}^{m} V_i$ 的有限维 F-向量空间 V 和分块上三角或下三角线性映射 $T \in \mathrm{End}(V)$, 我们也有 $\mathrm{Char}_T = \prod_{i=1}^{m} \mathrm{Char}_{T_{ii}}$.

命题 5.8.6 对任意 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 和 $\boldsymbol{B} \in \mathrm{M}_{n\times m}(F)$ 都有 $F[X]$ 中的等式

$$X^n \mathrm{Char}_{\boldsymbol{AB}} = X^m \mathrm{Char}_{\boldsymbol{BA}}.$$

特别地, 当 $m = n$ 时 $\mathrm{Char}_{\boldsymbol{AB}} = \mathrm{Char}_{\boldsymbol{BA}}$.

证明 在有理函数域 $F(X)$ 中操作. 对 $\boldsymbol{A}$ 和 $X^{-1}\boldsymbol{B}$ 应用推论 5.6.3, 可得 $F(X)$ 中的等式

$$\begin{aligned} X^{-m}\det\left(X\cdot \mathbf{1}_{m\times m} - \boldsymbol{AB}\right) &= \det\left(\mathbf{1}_{m\times m} - X^{-1}\boldsymbol{AB}\right) \\ &= \det\left(\mathbf{1}_{n\times n} - X^{-1}\boldsymbol{BA}\right) \\ &= X^{-n}\det\left(X\cdot \mathbf{1}_{n\times n} - \boldsymbol{BA}\right). \end{aligned}$$

两端同乘以 X^{m+n} 便是. □

当 $m = n$ 时, 命题结论化为 $\mathrm{Char}_{\boldsymbol{AB}} = \mathrm{Char}_{\boldsymbol{BA}}$, 本章习题将对此介绍另一种证法.

例 5.8.7 (友矩阵) 设 $n \in \mathbb{Z}_{\geq 1}$. 给定 $c_0, \dots, c_{n-1} \in F$, 定义矩阵

$$\boldsymbol{C} := \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix} \in \mathrm{M}_{n\times n}(F),$$

称之为多项式 $f := c_0 + \cdots + c_{n-1}X^{n-1} + X^n$ 的友矩阵. 最低阶的特例是

$$n = 1: \quad \boldsymbol{C} = \begin{pmatrix} -c_0 \end{pmatrix}, \qquad n = 2: \quad \boldsymbol{C} = \begin{pmatrix} 0 & -c_0 \\ 1 & -c_1 \end{pmatrix}.$$

以下来验证 $\mathrm{Char}_{\boldsymbol{C}} = f$. 首先, $n = 1$ 时显然 $\mathrm{Char}_{\boldsymbol{C}} = X + c_0$.

对于一般的 $n \geq 2$, 将行列式 $\det(X \cdot \mathbf{1}_{n\times n} - \boldsymbol{C})$ 按第一行展开可得

$$\begin{vmatrix} X & 0 & \cdots & 0 & c_0 \\ -1 & X & \cdots & 0 & c_1 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & X + c_{n-1} \end{vmatrix} =$$

$$X \begin{vmatrix} X & \cdots & 0 & c_1 \\ -1 & X & \cdots & c_2 \\ & \ddots & & \vdots \\ 0 & \cdots & -1 & X + c_{n-1} \end{vmatrix} + (-1)^{n-1} c_0 \begin{vmatrix} -1 & X & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & -1 \end{vmatrix}$$

第一项可以递归地化为 $X(c_1 + \cdots + c_{n-1}X^{n-2})$, 第二项则按上三角行列式的公式 (命题 5.5.3) 写作 $(-1)^{n-1}(-1)^{n-1}c_0 = c_0$, 相加即 f.

多项式 $\mathrm{Char}_{\boldsymbol{A}} \in F[X]$ 可以代入 $\boldsymbol{A}$ 求值, 给出 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) \in \mathrm{M}_{n\times n}(F)$. 类似地, 对有限维 F-向量空间 V 和 $T \in \mathrm{End}(V)$ 同样有 $\mathrm{Char}_T(T) \in \mathrm{End}(V)$.

定理 5.8.8 (A. Cayley, W. R. Hamilton) 设 $n \in \mathbb{Z}_{\geq 1}$. 对一切 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 皆有 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \mathbf{0}_{n\times n}$. 类似地, 对一切有限维 F-向量空间 V 和 $T \in \mathrm{End}(V)$ 皆有 $\mathrm{Char}_T(T) = 0_V$.

证明 处理矩阵版本即可. 这一著名结果有许多证明. 有些理路自然, 但需要进一步的域论知识. 有些涉及张量代数, 可以扩及其他的数学对象. 另一些初等证明则比较取巧. 以下选取的论证属于初等而取巧的类型, 优势在于它和经典伴随矩阵相联系.

我们需要次一定理 5.8.9, 它将蕴涵

$$\begin{aligned}
&\boldsymbol{A}^n + c_{n-1}\boldsymbol{A}^{n-1} + \cdots + c_0\mathbf{1}_{n\times n} \\
&\quad = \boldsymbol{A}(\underbrace{\boldsymbol{A}^{n-1} + c_{n-1}\boldsymbol{A}^{n-2} + \cdots + c_1\mathbf{1}_{n\times n}}_{=(-1)^{n-1}\boldsymbol{A}^\vee\ \text{(待证)}}) + (-1)^n \det\boldsymbol{A}\cdot\mathbf{1}_{n\times n} \\
&\quad = (-1)^n\left(\det\boldsymbol{A}\cdot\mathbf{1}_{n\times n} - \boldsymbol{A}\boldsymbol{A}^\vee\right) \overset{\text{定理 5.7.4}}{=\!=} \mathbf{0}_{n\times n}.
\end{aligned}$$

由之便完成定理的证明. □

证明将由以下定理补全, 它本身也是一则有趣的结果.

定理 5.8.9 设 $n \geq 1$. 记 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 的特征多项式为 $X^n + c_{n-1}X^{n-1} + \cdots + c_0$, 则

$$(-1)^{n-1}\boldsymbol{A}^\vee = c_1\cdot\mathbf{1}_{n\times n} + \cdots + c_{n-1}\boldsymbol{A}^{n-2} + \boldsymbol{A}^{n-1},$$

右式在 $n = 1$ 时理解为 $\boldsymbol{A}^0 := 1 \in F = \mathrm{M}_{1\times 1}(F)$.

证明 注意到当 $n = 1$ 时, $\boldsymbol{A}^\vee$ 按定义 5.7.3 规定为 1, 断言平凡地成立.

以下对 $n \geq 2$ 情形在有理函数域 $F(X)$ 上计算 $X\cdot\mathbf{1}_{n\times n} - \boldsymbol{A}$ 的经典伴随矩阵. 注意到变元 X 在各个余子式的每一行或列都至多出现一次, 所以各个余子式作为 X 的多项式至多是 $n-1$ 次的. 将经典伴随矩阵的各个矩阵元按 X 的次数整理, 写作

$$(X\cdot\mathbf{1}_{n\times n} - \boldsymbol{A})^\vee = \sum_{k=0}^{n-1} X^k\boldsymbol{D}_k, \quad \boldsymbol{D}_0,\ldots,\boldsymbol{D}_{n-1} \in \mathrm{M}_{n\times n}(F).$$

这种展开式是唯一的.

今考虑 $\mathrm{M}_{n\times n}(F(X))$ 中的等式

$$\begin{aligned}
\left(c_0 + \cdots + c_{n-1}X^{n-1} + X^n\right)\mathbf{1}_{n\times n} &\overset{\text{定理 5.7.4}}{=\!=} (X\cdot\mathbf{1}_{n\times n} - \boldsymbol{A})(X\cdot\mathbf{1}_{n\times n} - \boldsymbol{A})^\vee \\
&= (X\cdot\mathbf{1}_{n\times n} - \boldsymbol{A})\sum_{k=0}^{n-1} X^k\boldsymbol{D}_k.
\end{aligned}$$

等式两边的矩阵元都是多项式. 将两边按次数整理, 逐项比较每个 X^k 的系数, 便得到

$$\begin{aligned}
\boldsymbol{D}_{n-1} &= \mathbf{1}_{n\times n}, \\
-\boldsymbol{A}\boldsymbol{D}_{n-1} + \boldsymbol{D}_{n-2} &= c_{n-1}\cdot\mathbf{1}_{n\times n}, \\
&\vdots \\
-\boldsymbol{A}\boldsymbol{D}_1 + \boldsymbol{D}_0 &= c_1\cdot\mathbf{1}_{n\times n}, \\
-\boldsymbol{A}\boldsymbol{D}_0 &= c_0\cdot\mathbf{1}_{n\times n}.
\end{aligned} \tag{5.8.2}$$

将前 n 条等式依序左乘以 $\boldsymbol{A}^{n-1},\ldots,\boldsymbol{A}^0 = \mathbf{1}_{n\times n}$, 其右式相加给出 $c_1\cdot\mathbf{1}_{n\times n} + \cdots + c_{n-1}\boldsymbol{A}^{n-2} + \boldsymbol{A}^{n-1}$, 而左式前后相消, 于是乎

$$\boldsymbol{D}_0 = c_1\cdot\mathbf{1}_{n\times n} + \cdots + c_{n-1}\boldsymbol{A}^{n-2} + \boldsymbol{A}^{n-1}.$$

然而 $\boldsymbol{D}_0$ 按定义即是 $(-\boldsymbol{A})^\vee = (-1)^{n-1}\boldsymbol{A}^\vee$. 明所欲证. □

在定理 5.8.9 的证明中如果将 (5.8.2) 的每个等式依序左乘以 $\boldsymbol{A}^n, \boldsymbol{A}^{n-1}, \dots, \boldsymbol{A}^0$ 然后相加, 则可直接给出 Cayley–Hamilton 定理 $c_0 \cdot \mathbf{1}_{n\times n} + \cdots + \boldsymbol{A}^n = \mathbf{0}_{n\times n}$, 但获得的信息也稍打折扣.

推论 5.8.10 设 $n \geq 1$. 对任意 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 和任意可逆矩阵 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$, 皆有 $(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^\vee = \boldsymbol{P}^{-1}\boldsymbol{A}^\vee\boldsymbol{P}$.

证明 因为特征多项式的系数 $c_0, \dots, c_{n-1}$ 只依赖 $\boldsymbol{A}$ 的共轭类, 或者说特征多项式是内蕴的, 故

$$\begin{aligned}(-1)^{n-1}\boldsymbol{P}^{-1}\boldsymbol{A}^\vee\boldsymbol{P} &= c_1 \cdot \mathbf{1}_{n\times n} + \cdots + c_{n-1}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^{n-2} + (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^{n-1}\\ &= (-1)^{n-1}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^\vee.\end{aligned}$$

明所欲证. □

既然特征多项式是内蕴的, 对任意 n 维非零 F-向量空间 V 及线性映射 $T \in \mathrm{End}(V)$ 也都可以内蕴地定义

$$T^\vee := (-1)^{n-1}\left(c_1 \cdot \mathrm{id}_V + \cdots + c_{n-1}T^{n-2} + T^{n-1}\right) \in \mathrm{End}(V)$$

右式在 $n = 1$ 时理解为 $T^0 = \mathrm{id}_V$, 使得它在一切有序基下的矩阵 $\boldsymbol{\mathcal{M}}(T^\vee)$ 都满足 $\boldsymbol{\mathcal{M}}(T^\vee) = \boldsymbol{\mathcal{M}}(T)^\vee$.

特征多项式的深入研究和应用需要以多项式的进阶理论为前提, 这是第六章的主题.

练习 5.8.11 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 可逆, $\mathrm{Char}_{\boldsymbol{A}} = X^n + c_{n-1}X^{n-1} + \cdots + c_0$. 证明

$$\mathrm{Char}_{\boldsymbol{A}^{-1}} = X^n + \frac{c_1}{c_0} \cdot X^{n-1} + \cdots + \frac{c_{n-1}}{c_0} \cdot X + \frac{1}{c_0}.$$

提示 观察 $\det \boldsymbol{A} \cdot \det\left(X \cdot \mathbf{1}_{n\times n} - \boldsymbol{A}^{-1}\right) = (-X)^n \det\left(X^{-1} \cdot \mathbf{1}_{n\times n} - \boldsymbol{A}\right)$.

5.9 线性映射的迹

考虑 n 维 F-向量空间 V 和线性映射 $T \in \mathrm{End}(V)$, 其中 $n \in \mathbb{Z}_{\geq 1}$. 特征多项式 $\mathrm{Char}_T = X^n + c_{n-1}X^{n-1} + \cdots + c_0$ 的每个系数 $c_i \in F$ 都蕴藏关于 T 的重要信息, 例如 $c_0 = (-1)^n \det T$. 在这一列系数中, 相对容易操作的是 c_{n-1}. 为了方便计算, 今后选定空间的基, 将问题化为矩阵的特征多项式来处理.

引理 5.9.1 设 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$, 特征多项式 $\mathrm{Char}_{\boldsymbol{A}} \in F[X]$ 写作 $X^n + c_{n-1}X^{n-1} + \cdots + c_0$, 则

$$-c_{n-1} = \sum_{i=1}^{n} a_{ii}.$$

证明 按照行列式的公式, $\mathrm{Char}_{\boldsymbol{A}} = \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(\sigma)\prod_{i=1}^n \alpha_{i,\sigma(i)}$, 其中

$$\alpha_{ij} := X\cdot\delta_{i,j} - a_{ij}.$$

对于和式中的每个 $\sigma\in\mathfrak{S}_n$, 为了使对应项贡献一个次数 $\geq n-1$ 的多项式, 连乘积中的 $(i,\sigma(i))$ 必须至少有 $n-1$ 项落在对角线上, 这是因为变元 X 仅在对角线上出现. 换言之, 映射 $\sigma:\{1,\dots,n\}\to\{1,\dots,n\}$ 至少要有 $n-1$ 个不动点. 然而 σ 既是双射, 剩下的点也不动, 这就表明 $\sigma=\mathrm{id}$.

既然 $\mathrm{sgn}(\mathrm{id})=1$, 综上可见

$$\mathrm{Char}_{\boldsymbol{A}} = \prod_{i=1}^n (X-a_{ii}) + (\text{次数小于 } n-1 \text{ 的多项式}),$$

因此 c_{n-1} 等于 X^{n-1} 在 $\prod_{i=1}^n(X-a_{ii})$ 中的系数, 众所周知即 $-\sum_{i=1}^n a_{ii}$. □

定义 5.9.2 (迹) 对于矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}\in\mathrm{M}_{n\times n}(F)$, 其迹定义为

$$\mathrm{Tr}(\boldsymbol{A}) := \sum_{i=1}^n a_{ii}.$$

由于 $-\mathrm{Tr}(\boldsymbol{A})$ 等于 $\mathrm{Char}_{\boldsymbol{A}}$ 的 $n-1$ 次项系数 (引理 5.9.1), 而这些系数是共轭不变量, 我们有

$$\mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}) = \mathrm{Tr}(\boldsymbol{A}),\quad \boldsymbol{P}\in\mathrm{M}_{n\times n}(F)\text{ 可逆}.$$

依此可以对任何 n 维 F-向量空间 V 定义 $T\in\mathrm{End}(V)$ 的迹为 $\mathrm{Tr}(T):=\mathrm{Tr}(\boldsymbol{A})$, 其中 $\boldsymbol{A}=\boldsymbol{\mathcal{M}}(T)$ 是 T 在某组基下的矩阵.

等价地说, $\mathrm{Tr}(T)=-c_{n-1}$, 其中 $X^n+c_{n-1}X^{n-1}+\cdots+c_0=\mathrm{Char}_T$.

命题 5.9.3 设 V 是带有直和分解 $V=\bigoplus_{i=1}^n V_i$ 的有限维向量空间, $T\in\mathrm{End}(V)$ 相应地写成 $(T_{ij})_{1\leq i,j\leq n}$ 的分块形式, 则

$$\mathrm{Tr}(T) = \sum_{i=1}^n \mathrm{Tr}(T_{ii}).$$

证明 将问题转译为矩阵版本, 则因为迹定义为对角元之和, 它相对于分块的加性是明显的. □

迹还具备以下的初步性质, 且以矩阵语言来表述.

▷ **线性** 取迹给出线性映射 $\mathrm{Tr}:\mathrm{M}_{n\times n}(F)\to F$. 这是因为定义显然地导致

$$\mathrm{Tr}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{Tr}(\boldsymbol{A})+\mathrm{Tr}(\boldsymbol{B}),\quad \mathrm{Tr}(t\boldsymbol{A})=t\,\mathrm{Tr}(\boldsymbol{A}),$$
$$\boldsymbol{A},\boldsymbol{B}\in\mathrm{M}_{n\times n}(F),\ t\in F.$$

▷ **对称性** 对任何 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$ 皆有 $\mathrm{Tr}(\boldsymbol{AB})=\mathrm{Tr}(\boldsymbol{BA})$. 这是因为按定义

$$\begin{aligned}\mathrm{Tr}(\boldsymbol{AB}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}\\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \mathrm{Tr}(\boldsymbol{BA}).\end{aligned}$$

▷ **共轭不变性** 若 $\boldsymbol{P}\in \mathrm{M}_{n\times n}(F)$ 可逆, 则 $\mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{AP})=\mathrm{Tr}(\boldsymbol{A})$. 这是定义中业已提及的事实, 但也可以按对称性论证如下:

$$\mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{AP}) = \mathrm{Tr}\left((\boldsymbol{P}^{-1}\boldsymbol{A})\boldsymbol{P}\right) = \mathrm{Tr}\left(\boldsymbol{PP}^{-1}\boldsymbol{A}\right) = \mathrm{Tr}(\boldsymbol{A}).$$

迹的对称性可以进一步推广为轮换对称性. 对于任一列矩阵 $\boldsymbol{A}_1,\dots,\boldsymbol{A}_k\in \mathrm{M}_{n\times n}(F)$, 对称性给出

$$\mathrm{Tr}\left(\boldsymbol{A}_1\cdots\boldsymbol{A}_k\right) = \mathrm{Tr}\left(\boldsymbol{A}_1(\boldsymbol{A}_2\cdots\boldsymbol{A}_k)\right) = \mathrm{Tr}\left(\boldsymbol{A}_2\cdots\boldsymbol{A}_k\boldsymbol{A}_1\right).$$

反复运用此性质, 可以得到 $\mathrm{Tr}\left(\boldsymbol{A}_3\cdots\boldsymbol{A}_k\boldsymbol{A}_1\boldsymbol{A}_2\right)$, $\mathrm{Tr}\left(\boldsymbol{A}_4\cdots\boldsymbol{A}_k\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\right)$ 等, 共有 k 种形式. 这相当于说 $\mathrm{Tr}(\boldsymbol{A}_1\cdots\boldsymbol{A}_k)$ 在下标 $1,\dots,k$ 的轮换之下保持不变. 对于一般的置换 $\sigma\in\mathfrak{S}_k$ 则无此性质.

5.10 不变子空间

设 V 是 F-向量空间. 以下概念在向量空间理论中是简单而基本的.

定义 5.10.1 (不变子空间) 给定线性映射 $T\in\mathrm{End}(V)$, 如果子空间 $U\subset V$ 满足 $T(U)\subset U$, 则称 U 为 V 的 T-不变子空间, 或简称不变子空间.

如果 $U\subset V$ 是 T-不变子空间, 则对于任意多项式 $f\in F[X]$, 它也是 $f(T)$-不变子空间.

对任意 T-不变子空间 $U\subset V$, 我们有 §4.12 定义的商空间 V/U, 而 $T(U)\subset U$ 和推论 4.12.8 给出线性映射 $\overline{T}: V/U\to V/U$, 映陪集 $v+U$ 为 $Tv+U$. 我们的目的是更具体地描述 $\overline{T}$ 和 T 的关系.

以下设 V 有限维. 基于命题 4.12.12, 一旦为 U 和 V/U 取基, 再为 V/U 的基的每个元素在 V 中任选原像, 则 T 对相应的直和分解便有分块上三角的形式:

$$\left(\begin{array}{c|c} T|_U & * \\ \hline 0 & \overline{T} \end{array}\right). \tag{5.10.1}$$

由此容易对特征多项式作相应的分解.

命题 5.10.2 设 V 是有限维 F-向量空间, $T \in \mathrm{End}(V)$ 而 U 是 V 的 T-不变子空间, 则推论 4.12.8 给出的线性映射 $\overline{T} \in \mathrm{End}(V/U)$ 满足

$$\mathrm{Char}_{T|_U} \cdot \mathrm{Char}_{\overline{T}} = \mathrm{Char}_T .$$

证明 将分块表达式 (5.10.1) 代入命题 5.8.5 来计算 Char_T. □

既然特征多项式相对于不变子空间分解为乘积, 行列式和迹作为其系数自然有相应的分解. 直接证明也是毫不困难的.

命题 5.10.3 设 V 是有限维向量空间, $T \in \mathrm{End}(V)$ 而 U 是 V 的 T-不变子空间, 则对应的 $\overline{T} \in \mathrm{End}(V/U)$ 满足

$$\begin{aligned} \det(T|_U) \cdot \det(\overline{T}) &= \det(T), \\ \mathrm{Tr}(T|_U) + \mathrm{Tr}(\overline{T}) &= \mathrm{Tr}(T). \end{aligned}$$

证明 作形如 (5.10.1) 的分解, 则关于 $\det(T)$ 和 $\mathrm{Tr}(T)$ 的断言分别化约为命题 5.6.1 和命题 5.9.3 的公式. □

上述等式也可以从命题 5.10.2 来推导, 留给读者练习.

5.11 子式与 Cauchy–Binet 公式

设 V 是有限维 F-向量空间. 定理 5.4.3 断言行列式具备乘性 $\det(ST) = \det(S)\det(T)$, 其中 $S, T \in \mathrm{End}(V)$. 我们自然也可以问: 对于有限维 F-向量空间 V, W 和线性映射 $V \xrightarrow{T} W \xrightarrow{S} V$, 行列式 $\det(ST)$ 可以如何表达? 这一问题适合以矩阵的语言来表述. 今起假设 $m, n \in \mathbb{Z}_{\geq 1}$.

定义 5.11.1 考虑非空子集 $I \subset \{1, \dots m\}$ 和 $J \subset \{1, \dots, n\}$, 具体写成 $I = \{i_1, \dots, i_r\}$ 和 $J = \{j_1, \dots, j_s\}$ 之形, 其中 $i_1 < \cdots < i_r$ 而 $j_1 < \cdots < j_s$. 矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 的 (I, J)-**子矩阵**定义为

$$\begin{aligned} \boldsymbol{A}\binom{I}{J} &= \boldsymbol{A}\binom{i_1 \cdots i_r}{j_1 \cdots j_s} \\ &:= \left(a_{i_p, j_q}\right)_{\substack{1\leq p\leq r \\ 1\leq q\leq s}} \in \mathrm{M}_{r\times s}(F). \end{aligned}$$

换言之, 这是从 $\boldsymbol{A}$ 中按序抽取属于 I 的行和属于 J 的列所给出的矩阵. 对于 $|I| = |J|$ 的情形, $\det \boldsymbol{A}\binom{I}{J}$ 称为 (I, J) 确定的**子式**; 形如 $\det \boldsymbol{A}\binom{I}{I}$ 的子式称为 I 确定的**主子式**.

练习 5.11.2 对 $n = m$ 的情形, 说明定义 5.4.8 的余子式是哪些 (I, J) 确定的子式.

首先说明如何以主子式表达多项式 $\det(X \cdot \mathbf{1}_{n\times n} + \boldsymbol{C})$ 的系数, 其中 $\boldsymbol{C} \in \mathrm{M}_{n\times n}(F)$; 它们也是 $-\boldsymbol{C}$ 的特征多项式的系数.

引理 5.11.3 设 $\boldsymbol{C}\in \mathrm{M}_{n\times n}(F)$ 而 $1\leq k\leq n$, 则多项式 $\det(X\cdot \mathbf{1}_{n\times n}+\boldsymbol{C})\in F[X]$ 中 X^{n-k} 项的系数等于 $\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=k}}\det \boldsymbol{C}\binom{I}{I}$.

证明 应用公式

$$\det(X\cdot \mathbf{1}_{n\times n}+\boldsymbol{C})=\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}(X\cdot\delta_{i,\sigma(i)}+c_{i,\sigma(i)})$$

其中的 $\delta_{i,j}$ 如定义 5.8.2 之下所述. 对每个 $\sigma\in\mathfrak{S}_n$, 连乘积 $\operatorname{sgn}(\sigma)\prod_{i=1}^{n}(\cdots)$ 对 X^{n-k} 系数的贡献是

$$\operatorname{sgn}(\sigma)\sum_{I'}\prod_{i\notin I'}c_{i,\sigma(i)}$$

其中 I' 遍历所有满足 $|I'|=n-k$ 和 $\forall i\in I',\ \sigma(i)=i$ 的子集 $I'\subset\{1,\dots,n\}$. 上式对 $\sigma\in\mathfrak{S}_n$ 求和便是所论的系数.

我们也可以换序, 先对所有含 k 个元素的子集 $I\subset\{1,\dots,n\}$ 求和, 对应于先前的 $\{1,\dots,n\}\smallsetminus I'$, 然后再对所有在 I 之外不动的置换 $\sigma\in\mathfrak{S}_{n,I}$ 求和, 此处符号如 (5.1.2). 这给出

$$\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=k}}\ \sum_{\sigma\in\mathfrak{S}_{n,I}}\operatorname{sgn}(\sigma)\prod_{i\in I}c_{i,\sigma(i)}.$$

将每个 I 保序地等同于 $\{1,\dots,k\}$ 并且按 (5.1.2) 将 $\mathfrak{S}_n$ 的子集 $\mathfrak{S}_{n,I}$ 等同于 $\mathfrak{S}_k$, 这不影响置换的奇偶性 (注记 5.1.13). 于是上式进一步化为

$$\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=k}}\ \sum_{\tau\in\mathfrak{S}_k}\operatorname{sgn}(\tau)\prod_{j=1}^{k}c\binom{I}{I}_{j,\tau(j)},$$

其中 $\left(c\binom{I}{I}_{j,l}\right)_{j,l}=\boldsymbol{C}\binom{I}{I}$. 然而内层的和无非是 $\det \boldsymbol{C}\binom{I}{I}$. □

定理 5.11.4 (A.-L. Cauchy, J. P. M. Binet) 设 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$, $\boldsymbol{B}\in\mathrm{M}_{n\times m}(F)$.

(i) 若 $m>n$, 则 $\det(\boldsymbol{AB})=0$.

(ii) 若 $m\leq n$, 则

$$\det(\boldsymbol{AB})=\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=m}}\det \boldsymbol{A}\begin{pmatrix}1&\cdots&m\\ I_1&&I_m\end{pmatrix}\det \boldsymbol{B}\begin{pmatrix}I_1&&I_m\\ 1&\cdots&m\end{pmatrix}.$$

证明 若 $m>n$, 则注记 4.8.11 蕴涵 $\operatorname{rk}(\boldsymbol{AB})\leq\min\{\operatorname{rk}(\boldsymbol{A}),\operatorname{rk}(\boldsymbol{B})\}<m$, 故 $\boldsymbol{AB}\in\mathrm{M}_{m\times m}(F)$ 不可逆, $\det(\boldsymbol{AB})=0$.

以下设 $m\leq n$. 在命题 5.8.6 中代入 $\boldsymbol{A}$ 和 $-\boldsymbol{B}$ 可得

$$X^{n-m}\det(X\cdot\mathbf{1}_{m\times m}+\boldsymbol{AB})=\det(X\cdot\mathbf{1}_{n\times n}+\boldsymbol{BA}).$$

比较 X^{n-m} 在两边的系数. 左式系数是 $\det(X\cdot\mathbf{1}_{m\times m}+\boldsymbol{AB})$ 的常数项, 亦即 $\det(\boldsymbol{AB})$. 右式系数按引理 5.11.3 展开为

$$\sum_{\substack{I\subset\{1,\dots,n\}\\|I|=m}}\det\left((\boldsymbol{BA})\binom{I}{I}\right).$$

根据 $m\le n$ 和矩阵乘法的定义, 不难看出

$$(\boldsymbol{BA})\binom{I}{I}=\boldsymbol{B}\binom{I}{1\cdots m}\boldsymbol{A}\binom{1\cdots m}{I},$$

这是因为左式来自于从 B 取下标属于 I 的行, 从 A 取下标属于 I 的列, 然后相乘. 对右侧应用行列式的乘性 (定理 5.4.3), 然后对 I 求和, 便是 Cauchy–Binet 公式. □

Cauchy–Binet 的一则精彩应用是证明图论中的矩阵-树定理, 相关内容留作本章习题, 供感兴趣的读者尝试.

练习 5.11.5 设 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(\mathbb{R})$.

(i) 试以 Cauchy–Binet 公式说明 ${}^{\mathrm{t}}\boldsymbol{AA}$ 的行列式非负.

(ii) 进一步说明 ${}^{\mathrm{t}}\boldsymbol{AA}$ 的所有主子式皆非负.

提示〉 给定 $\{1,\dots,n\}$ 的子集 I, 从 $\boldsymbol{A}$ 萃取属于 I 的列得到子矩阵 $\boldsymbol{B}$, 然后考虑 ${}^{\mathrm{t}}\boldsymbol{BB}$.

5.12 交换环上的行列式

注记 4.3.8 已说明矩阵及其基本运算 (加减法, 乘法) 能定义在一般的环上. 自然地, 我们也想将矩阵的行列式从域推及一般的环. 然而乘法交换律在行列式理论中扮演关键角色, 为了得到完整的类比, 本节旨在对交换环 R 定义 n 阶行列式并探讨其性质. 这些理论将在本书后续章节用到.

在此前的讨论中, 我们主要从域 F 上的有限维向量空间和线性映射的观点来定义行列式, 而定义–命题 5.4.6 的具体公式不过是派生的. 对于交换环 R, 类似的进路虽然存在, 却涉及较复杂的代数结构[3], 建议感兴趣的读者参考 [10, §7.8]. 现阶段最合适的方式还是从具体公式出发, 直接定义 R 上的 n 阶行列式如下.

定义–命题 5.12.1 *设 R 为交换环, $n\in\mathbb{Z}_{\ge1}$, 而 $\boldsymbol{A}=(a_{ij})_{1\le i,j\le n}\in\mathrm{M}_{n\times n}(R)$. 定义 $\boldsymbol{A}$ 的行列式为 R 的以下元素*

$$\begin{vmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&a_{nn}\end{vmatrix}:=\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sgn}(\sigma)a_{\sigma(1),1}\cdots a_{\sigma(n),n};$$

[3] 参见关于模论与外代数的后续章节及习题.

又记为 $\det \boldsymbol{A}$. 这也可以改写为

$$\det \boldsymbol{A} = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

我们另外规定 $n = 0$ 时的“空行列式”为 1.

证明 唯一需要说明的是

$$\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

论证和定义–命题 5.4.6 一字不差, 需要的仅是环 R 的乘法交换律. □

定义是简单的. 重点在于说明行列式的种种一般性质在交换环 R 上依然成立. 我们先给出一部分陈述, 证明兴许比定理本身更有教益.

定理 5.12.2 设 R 为交换环, $n \in \mathbb{Z}_{\geq 1}$ 选定, 则 R 上的 n 阶行列式具有以下性质.

▷ **在单位矩阵的取值** 我们有 $\det(\mathbf{1}_{n\times n}) = 1$.

▷ **转置不变性** 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$, 则

$$\det \boldsymbol{A} = \det\left({}^{\mathrm{t}}\boldsymbol{A}\right).$$

▷ **乘性** 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(R)$, 则

$$\det(\boldsymbol{A}\boldsymbol{B}) = \det \boldsymbol{A} \det \boldsymbol{B}.$$

▷ **行/列的置换** 设 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(R)$. 若 $\sigma \in \mathfrak{S}_n$ 而 $\boldsymbol{B} = (b_{ij})_{i,j} \in \mathrm{M}_{n\times n}(R)$ 由下式确定:

$$b_{i,j} = a_{\sigma(i),j} \quad (\text{行的置换}),$$

或

$$b_{i,j} = a_{i,\sigma(j)} \quad (\text{列的置换}),$$

则 $\det \boldsymbol{B} = \operatorname{sgn}(\sigma) \det \boldsymbol{A}$.

▷ **对每一行/列的线性** 对所有 $1 \leq i \leq n$, 以下性质成立:

$$\begin{vmatrix} \vdots & & \vdots \\ a_{i1} + a'_{i1} & \cdots & a_{in} + a'_{in} \\ \vdots & & \vdots \end{vmatrix} = \begin{vmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{vmatrix} + \begin{vmatrix} \vdots & & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & & \vdots \end{vmatrix},$$

$$\begin{vmatrix} \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{in} \\ \vdots & & \vdots \end{vmatrix} = c \begin{vmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{vmatrix}$$

等号两边除第 i 行以外的矩阵元皆相同, $c \in R$.

类似地, 对所有 $1 \leq j \leq n$,

$$\begin{vmatrix} \cdots & a_{1j}+a'_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj}+a'_{nj} & \cdots \end{vmatrix} = \begin{vmatrix} \cdots & a_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj} & \cdots \end{vmatrix} + \begin{vmatrix} \cdots & a'_{1j} & \cdots \\ & \vdots & \\ \cdots & a'_{nj} & \cdots \end{vmatrix},$$

$$\begin{vmatrix} \cdots & ca_{1j} & \cdots \\ & \vdots & \\ \cdots & ca_{nj} & \cdots \end{vmatrix} = c \begin{vmatrix} \cdots & a_{1j} & \cdots \\ & \vdots & \\ \cdots & a_{nj} & \cdots \end{vmatrix}$$

等号两边除第 j 列以外的矩阵元皆相同, $c \in R$.

▷ **交错性** 若 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 有两行 (或列) 相同, 则 $\det \boldsymbol{A} = 0$.

▷ **按行/列和余子式展开** 对任意 $1 \leq i, j \leq n$, 任意 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 的行列式可以按第 i 行作展开

$$\det \boldsymbol{A} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik}$$

或者按第 j 列作展开

$$\det \boldsymbol{A} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} M_{kj},$$

其中 $\boldsymbol{A}$ 的余子式 M_{ij} 如定义 5.4.8, 但此处在 R 上操作.

▷ **分块上三角或下三角情形** 对于分块上三角矩阵, 我们有

$$\det \left(\begin{array}{c|c|c} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1n} \\ \hline \boldsymbol{0} & \ddots & \vdots \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{nn} \end{array}\right) = \prod_{i=1}^{n} \det \boldsymbol{A}_{ii};$$

分块下三角情形亦同.

证明 按定义直接计算 $\mathbf{1}_{n\times n}$ 的行列式, 公式中仅有 $\sigma = \mathrm{id}$ 有贡献. 其次, 转置不变性已包含于定义–命题 5.12.1.

其余性质在域上的版本是通过线性映射语言处理的. 在一般的环 R 上, 如果以定义–命题 5.12.1 为出发点, 则我们必须直面具体公式来验证. 这是可行的, 尽管某些场合 (例如乘性的证明) 显得相对复杂.

以下介绍另一种方法, 它能将问题化约到域上的已知情形. 具体起见, 我们举乘性为例:

$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \det \boldsymbol{B}, \quad \boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(R).$$

欲证的等式两边都是关于矩阵元 a_{ij} 和 b_{kl} 的多项式, 系数为整数 (更严格地说, 系数属于唯一环同态 $\mathbb{Z}\to R$ 的像, 见引理 3.7.1). 回想矩阵乘法与行列式的定义, 可见这些多项式的系数与环 R 无关. 我们按此将欲证的等式等价地表成

$$f(a_{ij},b_{kl})=g(a_{ij},b_{kl})$$

的形式, 其中 f 和 g 都是 $\mathbb{Z}$ 上的 $2n^2$ 元多项式, 而 i,j,k,l 各自遍历 $1,\dots,n$.

1. 首先观察到若 R 是交换环 S 的子环, 而所求的等式 $\det(\boldsymbol{AB})=\det\boldsymbol{A}\det\boldsymbol{B}$ 或等价的多项式表述对环 S 成立, 则它对子环 R 当然也成立.

2. 其次, 给定环同态 $\phi:R\to R'$, 命

$$\phi(\boldsymbol{A}):=(\phi(a_{ij}))_{1\leq i,j\leq n},\quad \phi(\boldsymbol{B}):=(\phi(b_{ij}))_{1\leq i,j\leq n}.$$

   兹断言若所求等式对 R 上的 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 成立, 则它对 R' 上的 $\phi(\boldsymbol{A})$ 和 $\phi(\boldsymbol{B})$ 成立. 从多项式等式的观点看, 论证再容易不过: 由于 f 与 g 为整系数多项式, 反复运用 ϕ 保持加法与乘法这一性质, 可得

$$f(\phi(a_{ij}),\phi(b_{kl}))=\phi(f(a_{ij},b_{kl})),\quad g(\phi(a_{ij}),\phi(b_{kl}))=\phi(g(a_{ij},b_{kl})),$$

   因此 R 中的等式 $f(a_{ij},b_{kl})=g(a_{ij},b_{kl})$ 蕴涵 R' 中的等式 $f(\phi(a_{ij}),\phi(b_{kl}))=g(\phi(a_{ij}),\phi(b_{kl}))$, 而这也相当于说

$$\det(\phi(\boldsymbol{A})\phi(\boldsymbol{B}))=\det\phi(\boldsymbol{A})\det\phi(\boldsymbol{B}).$$

3. 回到原问题. 给定 R, 考虑整系数 $2n^2$ 元多项式环

$$\tilde{R}:=\mathbb{Z}[X_{ij},Y_{kl}:1\leq i,j,k,l\leq n],$$

   以及其上的 "泛矩阵"

$$\widetilde{\boldsymbol{A}}:=(X_{ij})_{i,j},\quad \widetilde{\boldsymbol{B}}:=(Y_{kl})_{k,l}.$$

   这般操作的好处是对每个 X_{ij} (或 Y_{kl}) 代入 a_{ij} (或 b_{kl}) 给出的求值映射 $\phi:\tilde{R}\to R$ 为同态 (参见 (3.3.2)), 而且

$$\phi(\widetilde{\boldsymbol{A}})=\boldsymbol{A},\quad \phi(\widetilde{\boldsymbol{B}})=\boldsymbol{B}.$$

   根据先前第二点观察, 如能对 $\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{B}}\in\mathrm{M}_{n\times n}(\tilde{R})$ 证明相应的等式, 则取 ϕ 便给出 $\boldsymbol{A},\boldsymbol{B}\in\mathrm{M}_{n\times n}(R)$ 的版本.

4. 因此问题化约到 $R=\mathbb{Z}[X_{ij},Y_{kl}:1\leq i,j,k,l\leq n]$ 情形. 将整环 R 嵌入为分式域 $S:=\mathrm{Frac}(R)$ 的子环 (见 §3.5). 根据先前第一点观察, 若等式对 S 成立, 则对 R 也成立. 然而 S 是域 (事实上 S 是有理函数域 $\mathbb{Q}(X_{ij},Y_{kl})$), 问题终归化到域上的已知情形.

行列式的其他性质也可以按此化到域上处理. 论证写法机械化, 不必重复. □

以上手法相当于将所论的矩阵元 “变元化”. 简言之, 矩阵乘法与行列式的定义未涉及交换环 R 的任何特殊性质, 欲证的等式实际是以所有矩阵元为变元的一则或一族多项式等式 — 换句话说, 我们面对的其实是关于泛矩阵的等式. 矩阵论中的许多基本公式皆是如此.

类似地, 在一般的交换环 R 上能定义 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 的经典伴随矩阵

$$\boldsymbol{A}^\vee = (A_{ji})_{i,j} \in \mathrm{M}_{n\times n}(R), \quad A_{ij} := (-1)^{i+j} M_{ij},$$

以及特征多项式

$$\mathrm{Char}_{\boldsymbol{A}} := \det\left(X \cdot \mathbf{1}_{n\times n} - \boldsymbol{A}\right).$$

定理 5.12.3 设 R 为交换环, $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$. 我们有 $\boldsymbol{A}\boldsymbol{A}^\vee = \det \boldsymbol{A} \cdot \mathbf{1}_{n\times n} = \boldsymbol{A}^\vee \boldsymbol{A}$.

证明 这同样是关于矩阵元的一族整系数多项式等式, 故可按先前手法化约到域上的定理 5.7.4. □

Cayley–Hamilton 定理也有所推广.

定理 5.12.4 设 R 为交换环. 对一切 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 皆有 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \mathbf{0}_{n\times n}$.

证明 依然是化约到域上的定理 5.8.8. □

练习 5.12.5 对交换环 R 和 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 定义 $\boldsymbol{A}$ 的迹为 $\mathrm{Tr}(\boldsymbol{A}) := \sum_{i=1}^{n} a_{ii}$. 证明 $-\mathrm{Tr}(\boldsymbol{A})$ 是 $\mathrm{Char}_{\boldsymbol{A}}$ 的 $n-1$ 次项系数, 而且当 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(R)$ 时

$$\mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{Tr}(\boldsymbol{B}\boldsymbol{A}).$$

自然地, 我们还希望将关于矩阵取逆的推论 5.7.5 推及交换环 R. 在此之前, 必须明确何谓 R 上的可逆矩阵.

定义 5.12.6 设 R 为任意环. 若 $\boldsymbol{A}$ 是矩阵环 $\mathrm{M}_{n\times n}(R)$ 的可逆元 (定义 3.1.5), 换言之若存在 $\boldsymbol{B} \in \mathrm{M}_{n\times n}(R)$ 使得 $\boldsymbol{A}\boldsymbol{B} = \mathbf{1}_{n\times n} = \boldsymbol{B}\boldsymbol{A}$, 则称 $\boldsymbol{A}$ 为 R 上的可逆矩阵, 而称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆, 另记为 $\boldsymbol{A}^{-1}$.

根据环论, 定义中的 $\boldsymbol{B}$ 若存在则唯一, 故 $\boldsymbol{A}^{-1}$ 的记法合理.

可逆性关乎环的选取. 作为例证, 请读者按定义验证 $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ 在 $\mathbb{Q}$ 上可逆, 在 $\mathbb{Z}$ 上不可逆.

推论 5.12.7 设 R 为交换环, 则 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(R)$ 可逆的充要条件是 $\det \boldsymbol{A} \in R^\times$, 此时 $\boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1} \boldsymbol{A}^\vee$.

证明 设 $\boldsymbol{A}$ 可逆. 对 $\boldsymbol{A}\boldsymbol{A}^{-1} = \mathbf{1}_{n\times n}$ 两边同取行列式, 可得 $\det(\boldsymbol{A}) \det(\boldsymbol{A}^{-1}) = 1$, 故 $\det(\boldsymbol{A}) \in R^\times$.

至于另一个方向, 设 $\det \boldsymbol{A} \in R^\times$, 则定理 5.12.3 表明 $(\det \boldsymbol{A})^{-1}\boldsymbol{A}^\vee$ 确实具有 $\boldsymbol{A}^{-1}$ 所需的性质. □

习题

1. 说明若 $\sigma, \sigma' \in \mathfrak{S}_n$ 满足 $\mathrm{Inv}_\sigma = \mathrm{Inv}_{\sigma'}$, 则 $\sigma = \sigma'$.

2. 选定正整数 n. 若对所有 $1 \le i \le n$ 都有 $\sigma(i) \neq i$, 则称 $\sigma \in \mathfrak{S}_n$ 为错排. 记错排的个数为 $C(n)$, 另外规定 $C(0) = 1$. 证明

$$C(n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k!.$$

3. 置换 $\sigma \in \mathfrak{S}_n$ 的下降集定义为

$$X(\sigma) := \{1 \le i < n : \sigma(i) > \sigma(i+1)\}.$$

给定一列整数 $1 \le s_1 < \cdots < s_k < n$, 证明满足 $X(\sigma) = \{s_1, \ldots, s_k\}$ 的置换个数是

$$\sum_{j=0}^{k} \sum_{1 \le i_1 < \cdots < i_j \le k} (-1)^{k-j} \binom{n}{s_{i_1}, s_{i_2} - s_{i_1}, \ldots, n - s_{i_j}}$$

此处定义

$$\binom{n}{a_1, \ldots, a_j} := \frac{n!}{a_1! \cdots a_j!}, \quad a_1 + \cdots + a_j = n,$$

而在 $j = 0$ 时对应的 $\binom{n}{\cdots}$ 定义为 1.

提示〉 给定 $1 \le t_1 < \cdots < t_j < n$. 为了确定一个满足 $X(\sigma) \subset \{t_1, \ldots, t_j\}$ 的置换 σ, 先指定 $\sigma(1) < \cdots < \sigma(t_1)$, 再指定 $\sigma(t_1 + 1) < \cdots < \sigma(t_2)$, 依此类推. 说明这种置换共有 $\binom{n}{t_1, t_2 - t_1, \ldots, n - t_j}$ 个.

4. 对第一列作展开以证明

$$\begin{vmatrix} x_{0,1} & x_{0,2} & \cdots & x_{0,k} & x_{0,k+1} \\ 1 & x_{1,2} & \cdots & x_{1,k} & x_{1,k+1} \\ 0 & 1 & x_{2,3} & \cdots & x_{2,k+1} \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & x_{k,k+1} \end{vmatrix} = \sum_{j=0}^{k} \sum_{1 \le i_1 < \cdots < i_j \le k} (-1)^{k-j} x_{0,i_1} x_{i_1,i_2} \cdots x_{i_j,k+1}.$$

5. 承上两题, 证明满足 $X(\sigma) = \{s_1, \ldots, s_k\}$ 的置换 $\sigma \in \mathfrak{S}_n$ 个数是 $n! \det \boldsymbol{A}$, 其中 $\boldsymbol{A}$ 的定义是 $j + 1 \ge i$ 时 $a_{ij} := \dfrac{1}{(s_{j+1} - s_i)!}$, 否则 $a_{ij} := 0$.

提示〉 当 $j > i$ 时取 $x_{i,j} := \dfrac{1}{(s_j - s_i)!}$, 规定 $s_0 = 0$ 而 $s_{k+1} = n$.

6. 设 $a_1,\dots,a_n$ 皆非零, $n\geq 2$. 计算

$$\begin{vmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix} = (-1)^{n-1}a_1\cdots a_n\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right).$$

7. 证明所有可逆之 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 都可以表成 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{P}_\sigma\boldsymbol{V}$ 的形式, 其中 $\sigma\in\mathfrak{S}_n$ 给出置换矩阵 $\boldsymbol{P}_\sigma$, 而 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 都是上三角可逆矩阵.

这种分解在代数群理论中称为 Bruhat 分解, 可以证明其中的 σ 是唯一的.

提示〉 取 $\boldsymbol{A}$ 第一列最后一个非零元, 作适当的初等行/列变换化之为 1, 化同一行和同一列的其他矩阵元为 0. 说明这一操作可由将 $\boldsymbol{A}$ 两边乘以上三角可逆矩阵来实现. 对剩下的 $n-1$ 级方阵继续操作, 说明最后得到的是置换矩阵.

8. 证明对任何 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 和 $t\in F$ 都有 $\det(t\boldsymbol{A})=t^n\det(\boldsymbol{A})$. 从行列式作为比例常数的观点给出另一种说明.

9. 以简单的公式表达以下 n 阶行列式 ($n\geq 2$), 并给出证明:

$$\begin{vmatrix} \cos x & 1 & & & & & \\ 1 & 2\cos x & 1 & & & & \\ & 1 & 2\cos x & \ddots & & & \\ & & 1 & \ddots & \ddots & & \\ & & & \ddots & \ddots & 1 & \\ & & & & \ddots & 2\cos x & 1 \\ & & & & & 1 & 2\cos x \end{vmatrix}$$

其中留白部分为 0.

10. 对所有 $n\in\mathbb{Z}_{\geq 1}$ 定义 $\mathbb{Q}$ 上的 $n\times n$ 矩阵

$$\boldsymbol{A}_n=\begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix},$$

例如 $\boldsymbol{A}_1=(2)$ 而 $\boldsymbol{A}_2=\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$.

(i) 将行列式按行展开以证明当 $n \geq 3$ 时 $\det \boldsymbol{A}_n = 2\det \boldsymbol{A}_{n-1} - \det \boldsymbol{A}_{n-2}$.

提示〉 按最后一行展开 (做两次).

(ii) 证明 $\det \boldsymbol{A}_n = n+1$.

(iii) 对于 $n \geq 2$, 定义 $n \times n$ 矩阵 $\boldsymbol{C}_n$ 使得它的第 $(1,2)$ 个矩阵元为 -2, 其余矩阵元和 $\boldsymbol{A}_n$ 相同. 试证 $\det \boldsymbol{C}_n = 2$.

提示〉 同样按行展开两次得到 $n \geq 4$ 时 $\det \boldsymbol{C}_n = 2\det \boldsymbol{C}_{n-1} - \det \boldsymbol{C}_{n-2}$, 此外 $\det \boldsymbol{C}_2 = \det \boldsymbol{C}_3 = 2$.

11. 设 $f_{ij} : \mathbb{R} \to \mathbb{R}$ 为一族可导函数 $(1 \leq i, j \leq n)$. 证明

$$\begin{vmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & & \vdots \\ f_{n1} & \cdots & f_{nn} \end{vmatrix}' = \sum_{k=1}^{n} \begin{vmatrix} & \vdots & \\ f'_{k1} & \cdots & f'_{kn} \\ & \vdots & \end{vmatrix}_{\text{第 } k \text{ 行}} \quad (\text{其余元素仍是 } f_{ij}).$$

12. 考虑取值在域 F 中的所有数列 $(a_N)_{N\geq 0}$ 构成的向量空间, 其间的加法和纯量乘法都按逐项的方式定义:

$$(a_N)_{N\geq 0} + (b_N)_{N\geq 0} = (a_N + b_N)_{N\geq 0},$$
$$t(a_N)_{N\geq 0} = (ta_N)_{N\geq 0}.$$

(i) 设 $x_1, \ldots, x_n$ 为域 F 的相异元, 以 Vandermonde 行列式说明它们确定的等比数列 $(x_i^N)_{N\geq 0}$ (其中 $1 \leq i \leq N$) 是线性无关的; 这里约定 $0^0 = 1$.

(ii) 考虑次数从任意正整数 k 起步的等比数列 $(x_i^N)_{N\geq k}$, 陈述并证明相应的结果.

13. 设 F 为任意域. 考虑 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$ 和分块矩阵

$$\boldsymbol{X} := \left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{B} & \boldsymbol{A} \end{array}\right) \in \mathrm{M}_{2n\times 2n}(F).$$

证明 $\det \boldsymbol{X} = \det(\boldsymbol{A}+\boldsymbol{B}) \det(\boldsymbol{A}-\boldsymbol{B})$.

提示〉 除了行列式的标准操作外, 另一种思路是先要求 $\mathrm{char}(F) \neq 2$, 用标准基将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 看成线性映射 $F^n \to F^n$. 考虑 F^n 的两份副本, 其标准有序基分别记为 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ 和 $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$. 取 $F^{2n} \simeq F^n \oplus F^n$ 的有序基 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$, 依此将 $\boldsymbol{X}$ 看成 F^{2n} 到自身的线性映射, 然后用新基 $\{\boldsymbol{e}_i \pm \boldsymbol{f}_i\}_{i=1}^n$ 计算 $\det \boldsymbol{X}$. 特别地, 可以取 F 为 $\mathbb{Q}$ 上的 $2n^2$ 元有理函数域, 而 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的矩阵元是 $2n^2$ 个变元; 等式两边都是整系数多项式, 代值即可处理 F 为任意域或交换环的情形.

14. 记 $(a,b) := \gcd(a,b)$. 回忆 Euler 函数 φ 的定义 2.8.7. 证明

$$\begin{vmatrix} (1,1) & (1,2) & \cdots & (1,n) \\ (2,1) & (2,2) & \cdots & (2,n) \\ \vdots & \vdots & & \vdots \\ (n,1) & (n,2) & \cdots & (n,n) \end{vmatrix} = \varphi(1)\varphi(2)\cdots\varphi(n).$$

提示〉 一种思路是尝试逐步将第 k 行化到

$$\dots, \underbrace{\varphi(k)}_{\text{第 } k \text{ 项}}, \dots, \underbrace{\varphi(k)}_{\text{第 } 2k \text{ 项}}, \dots, \underbrace{\varphi(k)}_{\text{第 } 3k \text{ 项}}, \dots$$

的形式, 省略项皆为 0, 其中 $k = 1, 2, 3, \dots$. 可以运用练习 2.8.8 (iii) 的等式.

15. 选定 $n \in \mathbb{Z}_{\geq 1}$. 在 $\mathbb{C}$ 上计算 n 阶行列式

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & & & \vdots \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{vmatrix}.$$

提示〉 考虑对应的矩阵在列向量

$$\begin{pmatrix} 1 \\ \zeta \\ \vdots \\ \zeta^{n-1} \end{pmatrix}, \quad \zeta \in \mathbb{C},\ \zeta^n = 1.$$

上的作用, ζ 有 n 种选法.

16. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 其每行与每列之和皆为零. 记 $\boldsymbol{A}_0$ 为从 $\boldsymbol{A}$ 删除最后一行和最后一列得到的子矩阵, 证明 X 在 $\mathrm{Char}_{\boldsymbol{A}}$ 中的系数是 $(-1)^{n-1} n \det(\boldsymbol{A}_0)$, 而常数项是 0.

提示〉 将 $X \cdot \mathbf{1}_{n\times n} - \boldsymbol{A}$ 的前 $n-1$ 行加到第 n 行, 提出 X, 得到形如 $\mathrm{Char}_{\boldsymbol{A}}(X) = X \det \boldsymbol{B}(X)$ 的等式; 接着将 $\boldsymbol{B}(0)$ 的前 $n-1$ 列加到第 n 列.

17. 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$.

(a) 在 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 可逆的前提下, 不用推论 5.6.3 直接证明命题 5.8.6 的等式 $\mathrm{Char}_{\boldsymbol{AB}} = \mathrm{Char}_{\boldsymbol{BA}}$.

(b) 承上, 证明 $\mathrm{Char}_{\boldsymbol{AB}} = \mathrm{Char}_{\boldsymbol{BA}}$ 在不可逆情形也成立, 此处甚至可以容许矩阵元取在某个交换环 R 上.

提示〉 当域 F 有无穷多个元素时, 应用代数等式的延拓原理 (定理 3.6.3); 当 F 是有限域, 可将其扩张为无穷域来处理, 例如取有理函数域 $F(t)$. 对于一般的交换环 R, 试以 §5.12 的方法化到有理函数域 $\mathbb{Q}(X_{ij}, Y_{kl})$ 的情形.

18. (A.-L. Cauchy) 设 F 为域, $x_i, y_j \in F$ 为 $2n$ 个两两相异的元素 $(1 \leq i, j \leq n)$.

(i) 考虑 n 阶方阵 $\boldsymbol{C}_n = \left(\frac{1}{x_i - y_j}\right)_{1\leq i,j\leq n}$, 证明

$$\det \boldsymbol{C}_n = \prod_{1\leq i,j\leq n} (x_i - y_j)^{-1} \prod_{1\leq i<j\leq n} (x_j - x_i) \prod_{1\leq i<j\leq n} (y_i - y_j).$$

(ii) 大致地写下 $\boldsymbol{C}_n$ 的逆.

提示〉 对 (i) 采用递归论证. 先从第 $1,\dots,n-1$ 列减掉第 n 列, 从每行每列提出公因式, 再从第 $1,\dots,n-1$ 行减掉第 n 行, 然后继续提出公因式. 最后的产物是

$$\frac{\prod_{r=1}^{n-1}(y_r-y_n)}{\prod_{r=1}^{n}(x_r-y_n)}\frac{\prod_{r=1}^{n-1}(x_n-x_r)}{\prod_{r=1}^{n-1}(x_n-y_r)}\det \boldsymbol{C}_{n-1}.$$

对于 (ii), 代入 Cramer 法则可知 $\boldsymbol{C}_n^{-1}$ 的 (p,q) 矩阵元是

$$(-1)^{p+q}\prod_{1\leq i,j\leq n}{}'(x_i-y_j)\prod_{1\leq i<j\leq n}{}'(x_j-x_i)^{-1}\prod_{1\leq i<j\leq n}{}'(y_i-y_j)^{-1},$$

此处 $\prod'$ 代表连乘积中只取涉及 $i=q$ 或 $j=p$ 的项.

19. 证明若 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 是上三角 (或下三角) 矩阵, 则 $\boldsymbol{A}^\vee$ 也是上三角 (或下三角) 矩阵.

20. 设 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$.

(i) 设 $F=\mathbb{C}$. 证明若 $\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{A}$ 则 $\boldsymbol{A}$ 不可逆. 能否推及更一般的域 F?

(ii) 给出域 F 和 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{2\times 2}(F)$ 的例子, 使得 $\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{A}$ 而 $\boldsymbol{A}$ 可逆. 提示〉 取 F 使得 $\mathrm{char}(F)=2$.

21. 对所有向量空间 V 和 $A,B\in \mathrm{End}(V)$, 定义 $\mathrm{End}(V)$ 的元素

$$[A,B]:=AB-BA.$$

(a) 证明若 $\mathrm{char}(F)=0$ (例如 $F=\mathbb{C}$), 则当 V 有限维时 $[A,B]=\mathrm{id}$ 不可能成立.

(b) 给出无穷维向量空间 V 和 $A,B\in \mathrm{End}(V)$ 的例子, 使得 $[A,B]=\mathrm{id}$.

(c) 给出域 F, 有限维 F-向量空间 V 和 $A,B\in \mathrm{End}(V)$ 的例子, 使得 $[A,B]=\mathrm{id}$.

22. 设 F 是满足 $\mathrm{char}(F)=0$ 的域, V 是有限维 F-向量空间, $P_1,\dots,P_s\in \mathrm{End}(V)$ 满足 $P_1+\cdots+P_s=\mathrm{id}_V$. 证明下列三条陈述等价:

(i) 对每个 i 都有 $P_i^2=P_i$.

(ii) $\sum_{i=1}^s \mathrm{rk}(P_i)=\dim V$.

(iii) 对于任意 $i\neq j$ 皆有 $P_iP_j=0=P_jP_i$.

提示〉 (i) $\Longrightarrow$ (ii): 条件和练习 4.10.8 说明 P_i 是相对于直和分解 $V=V_i\oplus V_i'$ 的投影, 其中 $V_i=\mathrm{im}(P_i)$, $V_i'=\mathrm{im}(\mathrm{id}_V-P_i)$. 因此

$$\mathrm{Tr}(P_i)=\dim V_i=\mathrm{rk}(P_i).$$

于是有 $\sum_{i=1}^s \mathrm{rk}(P_i)=\sum_{i=1}^s \mathrm{Tr}(P_i)=\mathrm{Tr}(P_1+\cdots+P_s)=\mathrm{Tr}(\mathrm{id}_V)=\dim V$.

(ii) $\Longrightarrow$ (iii): 先说明线性映射

$$\sigma:\bigoplus_{i=1}^s \mathrm{im}(P_i)\to V,\quad (v_i)_{i=1}^s\mapsto\sum_{i=1}^s v_i$$

为满; 比较维数可知 σ 实则是同构. 综上, 每个 $v\in V$ 都有唯一表法 $v=\sum_{i=1}^s v_i$, 其中 $v_i\in \mathrm{im}(P_i)$, 由此推导 $P_iv=v_i$. 如此 $P_iP_j=0=P_jP_i$ 便属显然.

(iii) $\Longrightarrow$ (i): 不失一般性设 $i=1$. 我们有 $0=P_1(P_2+\cdots+P_s)$.

23. 设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}\in\mathbb{R}^3$, 视同列向量. 对分块矩阵

$$\boldsymbol{A}:=\begin{pmatrix}{}^{\mathrm{t}}\boldsymbol{a}\\ \hline {}^{\mathrm{t}}\boldsymbol{c}\end{pmatrix},\quad \boldsymbol{B}:=\left(\begin{array}{c|c}\boldsymbol{b} & \boldsymbol{d}\end{array}\right)$$

的乘积 $\boldsymbol{AB}$ 应用 Cauchy–Binet 定理 5.11.4, 以证明空间解析几何中的 Lagrange 恒等式

$$\begin{vmatrix}\boldsymbol{a}\cdot\boldsymbol{b} & \boldsymbol{a}\cdot\boldsymbol{d}\\ \boldsymbol{c}\cdot\boldsymbol{b} & \boldsymbol{c}\cdot\boldsymbol{d}\end{vmatrix}=(\boldsymbol{a}\times\boldsymbol{c})\cdot(\boldsymbol{b}\times\boldsymbol{d});$$

当然地, 以上的 $\cdot$ (或 $\times$) 代表空间向量的内积 (或叉积).

24. 设 $n\in\mathbb{Z}_{\geq 1}$. 证明 Cauchy 恒等式

$$\left(\sum_{i=1}^n a_ic_i\right)\left(\sum_{j=1}^n b_jd_j\right)-\left(\sum_{i=1}^n a_id_i\right)\left(\sum_{j=1}^n b_jc_j\right)=\sum_{1\leq i<j\leq n}(a_ib_j-a_jb_i)(c_id_j-c_jd_i).$$

这是 Lagrange 恒等式 ($n=3$) 的一种推广.

提示〉 直接展开比较, 或者对 Cauchy–Binet 定理 5.11.4 代入

$$\boldsymbol{A}=\begin{pmatrix}a_1 & \cdots & a_n\\ b_1 & \cdots & b_n\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}c_1 & d_1\\ \vdots & \vdots\\ c_n & d_n\end{pmatrix}.$$

25. 考虑交换环 R. 证明 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(R)$ 可逆当且仅当它左可逆, 当且仅当它右可逆. 交换环的条件是必须的, 否则在 $n=1$ 时已经有左逆非右逆的例子, 见第三章习题. 提示〉 应用推论 5.12.7 关于可逆性的判准, 以及行列式的乘性.

26. 说明 Cauchy–Binet 定理 5.11.4 在一般的交换环 R 上依然成立.

27. (有限图与矩阵) 所谓**无向图** (此处简称为图) 意指结构 $G=(V,E,r)$, 其中集合 V 的元素称为 G 的顶点, E 的元素称为 G 的边, 映射 r 映每个 $e\in E$ 为 V 的子集, 形如 $r(e)=\{v,v'\}$, 其元素称为 e 的端点, 不计顺序; 此处要求 $v\neq v'$, 换言之排除形如 $\overset{v}{\circlearrowleft}$ 的边. 按显然的方式对图定义连通性和子图的概念. 若 V 和 E 皆有限, 则称 G 为有限图. 假如对每个边 $e\in E$ 都指定 $r(e)$ 的一个元素为边的始点, 另一个为终点, 则称此为图 G 的一个定向.

若存在 $e\in E$ 使得 $r(e)=\{v,v'\}$, 则称顶点 v 和 v' 相邻; 顶点 v 的次数 $\deg(v)$ 定义为与 v 相邻的顶点个数. 若有一列相异的边 $v_0\overset{e_1}{\text{—}}v_1\overset{e_2}{\text{—}}\cdots\overset{e_n}{\text{—}}v_n=v_0$, 使得顶点 $v_0,\ldots,v_{n-1}$ 也相异 ($n\geq 2$), 则称之为 G 中的环路.

以下皆假定 G 为有限图, 并且将 V (或 E) 的元素标号为 $v_1,\ldots,v_p$ (或 $e_1,\ldots,e_q$).

(i) 无环路的连通有限图称为**树**, 因此树不能有形如 $v_0\overset{=}{\text{—}}v_1$ 的多重边. 证明以下陈述等价: (a) G 是树, (b) G 连通且 $q=p-1$, (c) G 无环路且 $q=p-1$.

以此说明任意有限图 G 中的 $p-1$ 条边或者含环路, 或者是某个树的边集; 在后一情形, 这些边的端点穷尽 G 的所有顶点.

(ii) 分别定义 G 的邻接矩阵 $\boldsymbol{A} \in \mathrm{M}_{p\times p}(\mathbb{Q})$, 关联矩阵 $\boldsymbol{B} \in \mathrm{M}_{p\times q}(\mathbb{Q})$ (在指定定向的情形), 和 Laplace 矩阵 $\boldsymbol{L} \in \mathrm{M}_{p\times p}(\mathbb{Q})$ 如下. 它们的矩阵元是

$$a_{ij} := \big|\{e \in E : e \text{ 以 } i,j \text{ 为端点}\}\big|, \quad b_{ij} := \begin{cases} 1, & e_j \text{ 始点为 } v_i, \\ -1, & e_j \text{ 终点为 } v_i, \\ 0, & \text{其他情形}, \end{cases}$$

$$\ell_{ij} := \begin{cases} -a_{ij}, & i \neq j, \\ \deg(v_i), & i = j. \end{cases}$$

因此 ${}^{\mathrm{t}}\boldsymbol{A} = \boldsymbol{A}$. 证明 $\boldsymbol{B}\,{}^{\mathrm{t}}\boldsymbol{B} = \boldsymbol{L}$.

(iii) 指定 G 的一个定向, 并且设 G 连通. 定义 $\boldsymbol{B}_0 \in \mathrm{M}_{(p-1)\times q}(\mathbb{Q})$ 为从 $\boldsymbol{B}$ 删除最后一行 (对应顶点 v_p) 得到的矩阵. 设 $S \subset E$ 为 $p-1$ 条边构成的子集, 也视同 $\{1,\dots,q\}$ 的子集. 沿用定义 5.11.1 的子矩阵符号, 证明

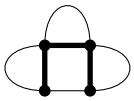
$$\det \boldsymbol{B}_0 \begin{pmatrix} 1,\dots,p-1 \\ S \end{pmatrix} = \begin{cases} 0, & S \text{ 包含环路}, \\ \pm 1, & S \text{ 是树的边集}. \end{cases}$$

鉴于 (i), 这穷尽 S 的所有可能性.

提示〉 首先设 S 包含环路 C, 沿 C 的边 $e_{i_1},\cdots,e_{i_k}$ 绕行, 走向可能和某些边的定向相反, 但调整定向只让行列式差 ±1, 故可假设一路顺行. 说明 $\boldsymbol{B}$ 的第 $i_1,\dots,i_k$ 个列向量的和为零向量.

其次设 S 是树的边集; 无妨设 $p \geq 2$. 用 (i) 取边 $e_k \in S$ 使得 v_p 是其端点. 说明 $\boldsymbol{B}_0\binom{1,\cdots,p-1}{S}$ 相应的列恰有一个非零元, 为 ±1. 令 $\boldsymbol{B}_0'$ 为从 $\boldsymbol{B}_0\binom{1,\cdots,p-1}{S}$ 删除含此非零元的行和列所得到的 $(p-2)\times(p-2)$ 矩阵; 按列展开行列式以说明 $\det \boldsymbol{B}_0\binom{1,\cdots,p-1}{S} = \pm \det \boldsymbol{B}_0'$. 在 S 中将 e_k 收缩为一点 u, 得到树 S', 说明 $\boldsymbol{B}_0'$ 是从 S' 的关联矩阵 $\boldsymbol{B}'$ 删除顶点 u 对应的列的产物, 于是递归可得 $\det \boldsymbol{B}_0' = \pm1$.

28. (Kirchhoff 矩阵-树定理) 沿用上一题的术语, 设 $G = (V,E,r)$ 为有限连通图 (无向). 若子图 H 本身是树, 而且和 G 有相同的顶点集, 则称 H 为 G 的一个**生成树**. 譬如下图给出满足 $(p,q) = (4,7)$ 的图, 加粗部分是其生成树.



生成树的研究是图论及其应用中的重要课题.

(i) 记 $\kappa(G)$ 为 G 的生成树个数. 记 $\boldsymbol{L}_0$ 为从 G 的 Laplace 矩阵 $\boldsymbol{L}$ 删除最后一行和最后一列 (对应顶点 v_p) 得到的子矩阵. 证明

$$\det \boldsymbol{L}_0 = \kappa(G).$$

提示〉 根据上一道习题的 (i) 及其符号, 指定生成树相当于指定 $E \simeq \{1,\dots,q\}$ 的 $p-1$ 元子集, 使得 S 不含环路. 基于 Cauchy–Binet 定理 5.11.4 推导

$$\det \boldsymbol{L}_0 = \sum_{\substack{S\subset\{1,\dots,q\} \\ |S|=p-1}} \left(\det \boldsymbol{B}_0 \begin{pmatrix} 1,\dots,p-1 \\ S \end{pmatrix}\right)^2$$

并且运用上一道习题的 (iii).

(ii) 在上述等式中, 顶点 v_p 的选择纯属人为, 用 $\boldsymbol{L}$ 的任何 $(n-1)\times(n-1)$ 主子式代替 $\det \boldsymbol{L}_0$ 也能得到 $\kappa(G)$. 证明 $\kappa(G)$ 还有无关任何选取的表法如下:

$$\mathrm{Char}_{\boldsymbol{L}}(X) = (-1)^{p-1} p\kappa(G) X + \text{高次项}.$$

提示〉 说明 $\boldsymbol{L}$ 每行与每列之和皆为零, 然后用此前的一道习题来表达 $\det \boldsymbol{L}_0$.

矩阵-树定理是 G. R. Kirchhoff 于 1847 年在关于直流电路的研究中发现的.

29. (A. Cayley) 对所有 $n \geq 2$ 定义无向图 K_n, 使其恰有 n 个顶点, 而且任两个相异顶点之间恰有一条边. 沿用上一题符号, 证明 $\kappa(K_n) = n^{n-2}$.

第六章 重访环和多项式

我们在第三章已经初识环, 域和多项式的基本概念. 本章回归这一基调, 而风景自殊. 原因一则是我们在前两章的学习中已经累积了许多经验, 实例, 与百尺竿头更进一步的想望. 二则是为了深化对向量空间和线性映射, 乃至对其它代数结构的理解, 环与多项式的进阶理论必不可少.

本章从关于理想和商环的 §6.1 起步. “理想” 是一个环的一类特殊子集, 由之可以在环上定义等价关系, 取商得到新的环, 思路与向量空间的商一脉相承, 属于代数学的基本操作. 这部分内容适用于一般的环.

此后的内容聚焦于交换环, 尤其是整环. 我们在 §6.2 先介绍整环中的整除关系和素元与不可约元的概念, 然后引进唯一分解环的定义. 素元与不可约元在唯一分解环中相等价. 我们将以辗转相除法证明域 F 上的多项式环 $F[X]$ 是唯一分解环, 从而揭示 $F[X]$ 与 $\mathbb{Z}$ 共享许多算术性质.

实际上, 辗转相除法证明了 $F[X]$ 和 $\mathbb{Z}$ 皆是主理想环, 亦即每个理想都能由单个元素生成. 主理想环总是唯一分解环. 详细的讨论是 §6.3 的主题.

关于多项式的进一步讨论见诸 §§6.4–6.6, 包括形式求导, 与之相关的根与重因式判准, 以及作为一则有趣应用的 Mason–Stothers 定理 6.5.3, 由之可以推导 Fermat 大定理的多项式版本.

在 §6.7 探讨的所谓 n 元对称多项式, 是指对所有置换 $\sigma \in \mathfrak{S}_n$ 皆满足

$$f(X_{\sigma(1)}, \dots, X_{\sigma(n)}) = f(X_1, \dots, X_n)$$

的 $f \in F[X_1, \dots, X_n]$. 对称多项式基本定理 6.7.7 说明对称多项式总能表为 $g(e_1, \dots, e_n)$, 其中 $g \in F[X_1, \dots, X_n]$ 而 $e_1, \dots, e_n$ 是 n 元初等对称多项式

$$e_k := \sum_{1 \le i_1 < \cdots < i_k \le n} X_{i_1} \cdots X_{i_k};$$

定理 6.7.8 还说明此 g 唯一. 由于一元首一多项式的系数总能表作根的初等对称多项式, 见 (6.7.1), 对称多项式提供了探究高次方程公式解的经典工具.

在 §6.8 探讨的结式是多项式理论中的另一经典概念, 它能判断 $F[X]$ 的两个元素有无次数 > 1 的公因式; 这一问题当然也能以辗转相除法高效地计算, 然而结式的优点是它基于一个明确的公式, 由关于多项式系数的某个行列式给出. 结式也能用来表达判别式, 见推论 6.8.6.

最后, §§6.9–6.11 围绕不可约多项式进行讨论. 不可约多项式在 $\mathbb{Q}=\mathrm{Frac}(\mathbb{Z})$ 上的情形具有数论的意涵, 相关内容包括著名的 Gauss 引理 (引理 6.9.4) 和 Eisenstein 判准 (定理 6.9.8) 在 $\mathbb{Z}$ 上的情形; 在这些思路的延长线上, 本章习题将勾勒如何证明唯一分解环上的多项式环仍是唯一分解环. 作为推论, 域上的 n 元多项式环是唯一分解环, 尽管它在 $n\geq 2$ 时不再是主理想环.

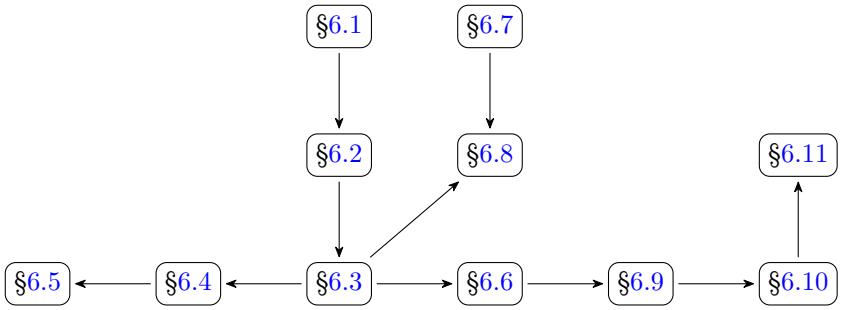
不可约多项式的另一个用法是构造扩域, 这相当于形式地向域添入多项式的根, 得到域的扩张 (推论 6.10.5); 具体办法涉及对理想取商. 借此方法, 命题 6.11.1 和 6.11.5 将说明存在有 q 个元素的有限域当且仅当 q 是某个素数 p 的幂. 这就在 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$ 之外构造了更多的有限域; 习题将介绍它们在编码学中的应用.

最后这部分讨论自然地触及扩域, 扩张次数和分裂域的概念, 本书浅尝辄止. 对于扩域的系统性讨论, 以及关于有限域的进阶研究 (例如有 $q=p^m$ 个元素的有限域 $\mathbb{F}_q$ 的唯一性), 可以参阅 [10, 第八和第九章] 或其它教材.

阅读提示

若无另外申明, 本章考虑的环都默认为非零环. 除了 §6.1, 本章只讨论交换环.

阅读顺序



6.1 理想和商环

对于环 R 的任意非空子集 S 和 $r\in R$, 引入方便的记法

$$rS:=\{rs:s\in S\},$$
$$Sr:=\{sr:s\in S\}.$$

定义 6.1.1 设 I 为环 R 的非空子集. 当以下条件成立时, 称 I 为 R 的**理想**[1].

▷ **加法封闭性** 若 $x,y\in I$ 则 $x+y\in I$.

[1] 这一术语源自 19 世纪 Kummer, Dedekind 等数学家对数论中的 “理想数” 的研究.

▷ **乘法双边封闭性** 对任意 $r \in R$ 皆有 $rI \subset I$ 和 $Ir \subset I$.

由于定义中不要求 R 交换, 所以乘法封闭性必须对双边来陈述; 对于非交换环, 我们也经常将上述定义中的理想称为 R 的**双边理想**. 如果将乘法封闭性放宽到 $rI \subset I$ (或 $Ir \subset I$), 便相应地得到左理想 (或右理想) 的概念, 但这不是本章关心的重点.

理想自动对加法逆元封闭: 若 $x \in I$ 则 $-x \in I$, 这是基于环论的等式 $-x = (-1)\cdot x$ 和理想的乘法封闭性. 理想的平凡例子有 $I = \{0\}$ (零理想) 和 $I = R$. 满足 $I \neq R$ 的理想 I 称为真理想.

命题 6.1.2 对于环 R 的任意理想 I, 我们有

$$I = R \iff 1 \in I.$$

证明 方向 $\Longrightarrow$ 显然, 至于 $\Longleftarrow$, 留意到 $1 \in I$ 蕴涵对所有 $r \in R$ 皆有 $r = 1\cdot r \in I$.□

真理想不可能是 R 的子环, 因为它不含乘法幺元 1.

例 6.1.3 (主理想) 设 R 为交换环, $x \in R$, 则 $(x) := xR = \{xr : r \in R\}$ 给出 R 的理想, 称为 x 确定的主理想.

举例明之, 引理 2.7.2 表明环 $\mathbb{Z}$ 的所有理想都是主理想.

推而广之, 对于交换环 R 的任意子集 S, 可以定义 S 生成的理想为

$$\langle S\rangle := \left\{ \sum_{s\in S} r_s s \in R : r_s \in R, \text{ 至多有限个 } r_s \text{ 非零} \right\},$$

它是包含 S 的最小理想; 我们规定 $\langle \varnothing\rangle := \{0\}$. 对于一列元素 $r_1, r_2, \ldots$, 记

$$\langle r_1, r_2, \ldots\rangle := \langle \{r_1, r_2, \ldots\}\rangle,$$

特别地, 先前定义的主理想 (x) 等于此处的 $\langle x\rangle$.

理想的另一类关键例子来自环同态的核, 细说如下.

定义–命题 6.1.4 (环同态的核) 设 $f : R \to R'$ 为环同态, 其核 (又称零核) 定义为

$$\ker(f) := f^{-1}(0) = \{x \in R : f(x) = 0\}.$$

这是 R 的理想

证明 首先验证加法封闭: 若 $x, y \in \ker(f)$, 则 $f(x+y) = f(x) + f(y) = 0 + 0 = 0$, 故 $x + y \in \ker(f)$. 其次验证乘法双边封闭. 若 $x \in \ker(f)$ 而 $r \in R$, 则

$$f(xr) = f(x)f(r) = 0 \cdot f(r) = 0 = f(r) \cdot 0 = f(r)f(x) = f(rx),$$

因此 $xr, rx \in \ker(f)$. □

以下选定环同态 $f : R \to R'$. 和向量空间和线性映射的情形类似, 对任意 $x, y \in R$ 皆有 $f(x) = f(y) \iff f(x-y) = 0 \iff x - y \in \ker(f)$.

回忆到 $\mathrm{im}(f)$ 是 R' 的子环, 自然可以问: 环 $\mathrm{im}(f)$ 是否可以仅由 R 及其理想 $I := \ker(f)$ 来内在地描述? 对于向量空间和线性映射的情形, §4.12 已经处理过相似的问题, 此一类比将为我们指明方向.

我们仍从集合的层次入手, 逐步加上代数结构. 首先, 在 R 上定义二元关系 $\sim_I$ 如下:

$$x \sim_I y \iff x - y \in I.$$

对于 $I = \ker(f)$ 的例子, $x \sim_I y$ 无非是说 $f(x-y) = 0$, 亦即 $f(x) = f(y)$; 于是

$$f \text{ 是单同态} \iff \ker(f) = \{0\}.$$

但是上述定义也可以施于任何理想 I: 由于 I 对加法和取加法逆元封闭, $\sim_I$ 确实是等价关系, 其论证和 §4.12 开头部分并无二致. 这就导向了约定 4.12.1 的环论版本.

约定 6.1.5 对任意环 R 及其理想 I, 相对于 $\sim_I$ 的等价类可以表成

$$x + I = \{x + r : r \in I\}$$

的形式, R 中的这种子集称为 I 的**陪集**, 而 x 称为该陪集的代表元.

回到 $I = \ker(f)$ 的情形. 在集合的层次上, 命题 2.5.8 给出从商集 $R/\sim_I$ 到 $\mathrm{im}(f)$ 的自然双射 $\overline{f}$, 映陪集 $x + I$ 为 $f(x)$. 一如向量空间的情形, 眼下的问题在于如何赋予 $R/\sim_I$ 环结构, 使得 $\overline{f}$ 成为环同构. 我们期望环结构的定义不再涉及 f, 而仅由 R 和理想 I 完全确定.

此外, 我们还希望商映射 $q : R \to R/\sim_I$ 本身也是环同态, 从而 f 便拆分为两个环同态 q 和 $R/\sim_I \xrightarrow{\sim} \mathrm{im}(f)$ 的合成. 这一要求为商集上的运算定义指明了道路: 陪集之间的代数运算只能定义为

$$\begin{aligned}(x+I)+(y+I) &= (x+y)+I \quad (\text{对应到 } q(x)+q(y) = q(x+y)),\\ (x+I)\cdot(y+I) &= xy+I \quad (\text{对应到 } q(x)q(y) = q(xy))\end{aligned}$$

而 $1_{R/\sim_I}$ 只能取为 $1_R + I$ (对应到 $q(1_R) = 1_{R/\sim_I}$). 关键在于验证这是 $R/\sim_I$ 上良定义的运算. 一旦确立这点, 今后便可以合理地将商集带有的这一环结构另记为 R/I.

定义 6.1.6 (商环) 设 I 是环 R 的理想. 定义

$$\begin{aligned}R/I &:= \{I \text{ 的陪集 } x + I : x \in R\}\\ &= R/\sim_I .\end{aligned}$$

在 R/I 上可以对陪集合理地定义

▷ **加法** $(x+I)+(y+I) := x+y+I$;

▷ **乘法** $(x+I)\cdot(y+I):=xy+I$;

▷ **零元** $0_{R/I}:=I=0_R+I$ (作为 I 的陪集);

▷ **幺元** $1_{R/I}:=1_R+I$.

由此得到的环 $(R/I,+,\cdot,0_{R/I},1_{R/I})$ 称为 R 对理想 I 的**商环**.

注记 6.1.7 易见 $x+I=I$ 当且仅当 $x\in I$. 按定义, R/I 是注记 3.1.2 的零环当且仅当对所有 x 都有 $x\in I$, 亦即 $I=R$. 有鉴于此, 在关于商环的讨论中通常只论真理想.

一如向量空间的商空间的定义 4.12.2, 上述运算只依赖陪集: 先看加法, 若 $x'=x+u$, $y'=y+v$, 其中 $u,v\in I$, 则理想的加法封闭性和 R 的加法交换律蕴涵

$$x'+y'=x+u+y+v\in x+y+\underbrace{u+v}_{\in I}+I=x+y+I,$$

因此 $x'+y'+I=x+y+I$; 同理, 理想的乘法封闭性蕴涵

$$x'y'=xy+\underbrace{xv+uy+uv}_{\in I}\in xy+I,$$

因此 $x'y'+I=xy+I$.

其次, 还必须对 R/I 验证环的公理, 这也毫不困难.

1. 关于加法运算的验证和商空间情形完全类似. 结合律, 交换律和零元的所需性质都化到 R 上去检验, 而陪集 $x+I$ 的加法逆元无非是 $-x+I$.
2. 关于乘法的性质也是类似的. 至于乘法对加法的分配律, 我们同样有

$$\begin{aligned}((x+I)+(y+I))(z+I)=(x+y+I)(z+I)&=(x+y)z+I\\ \xlongequal{R\text{ 的分配律}}(xz+yz)+I&=(xz+I)+(yz+I)\\ &=(x+I)(z+I)+(y+I)(z+I).\end{aligned}$$

同理可证另一侧的分配律. 一句话, 一切都回归 R 的性质.

集合层次的商映射 $q:R\to R/I$ 映 x 为陪集 $x+I$. 按定义 6.1.6 显见它保持加法, 乘法和幺元, 因此 q 是环同态, 称为**商同态**; 由于 $x+I=I$ 等价于 $x\in I$, 故 $\ker(q)=I$.

命题 6.1.8 如果 R 是交换环, $I\subset R$ 是理想, 则 R/I 仍是交换环.

证明 原因是 $(x+I)(y+I)=xy+I=yx+I=(y+I)(x+I)$. □

例 6.1.9 取 $R=\mathbb{Z}$. 我们从引理 2.7.2 知道它的所有理想都形如 $n\mathbb{Z}$, 其中 $n\in\mathbb{Z}$, 而对应的陪集无非是 mod n 的同余类 (定义 2.8.2). 对照 §2.7, 可见对应的商环正是 $\mathbb{Z}/n\mathbb{Z}$. 这是同余运算的环论诠释.

下述性质是命题 4.12.7 的自然类比.

命题 6.1.10 设 I 是环 R 的理想, 而 $f: R \to R'$ 是环同态.

(i) 若 $I \subset \ker(f)$, 则存在唯一的环同态 $\overline{f}: R/I \to R'$ 使得下图交换:

$$\begin{array}{ccc} R & \xrightarrow{f} & R' \\ {\scriptstyle q}\downarrow & \nearrow_{\overline{f}} & \\ R/I & & \end{array} \qquad \text{换言之 } \overline{f}q = f;$$

具体地说, $\overline{f}(x+I) = f(x)$.

(ii) 若 $I = \ker(f)$ 而 $R' = \mathrm{im}(f)$, 则 $\overline{f}$ 是环同构.

证明 既然 q 是满的, 使 (i) 的图表交换的唯一选择是命 $\overline{f}(x+I) = f(x)$. 之所以良定义, 是因为 $I \subset \ker(f)$ 蕴涵 $f(x)$ 只依赖陪集 $x+I$. 环同态的性质归为

$$\begin{aligned} \overline{f}((x+I)+(y+I)) &= \overline{f}(x+y+I) = f(x+y) \\ &= f(x)+f(y) = \overline{f}(x+I)+\overline{f}(y+I), \\ \overline{f}((x+I)(y+I)) &= \overline{f}(xy+I) = f(xy) = f(x)f(y) \\ &= \overline{f}(x+I)\overline{f}(y+I), \\ \overline{f}(1_R+I) &= f(1_R) = 1_{R'}. \end{aligned}$$

若进一步要求 $I = \ker(f)$ 而 $R' = \mathrm{im}(f)$, 则 $f = \overline{f}q$ 说明 $\overline{f}$ 满. 此外

$$\overline{f}(x+I) = 0 \iff f(x) = 0 \iff x \in I,$$

因此 $\overline{f}$ 也是单射, 从而是同构. □

推论 6.1.11 设 $f: R_1 \to R_2$ 是环同态, $I_1 \subset R_1$ 和 $I_2 \subset R_2$ 是理想, 而且 $f(I_1) \subset I_2$, 则存在唯一的环同态 $\overline{f}: R_1/I_1 \to R_2/I_2$ 使得下图交换:

$$\begin{array}{ccc} R_1 & \xrightarrow{f} & R_2 \\ {\scriptstyle q_1}\downarrow & & \downarrow{\scriptstyle q_2} \\ R_1/I_1 & \xrightarrow[\overline{f}]{} & R_2/I_2 \end{array} \qquad \text{换言之 } q_2 f = \overline{f} q_1,$$

图中的 q_1 和 q_2 分别是从 R_1 和 R_2 到各自商环的商同态. 具体地说, $\overline{f}(x+I_1) = f(x)+I_2$.

证明 合成同态 $q_2 f: R_1 \to R_2/I_2$ 映 I_1 为 0, 所以命题 6.1.10 说明存在唯一的 $\overline{f}$ 使得 $q_2 f = \overline{f} q_1$, 它确实映 $x+I = q_1(x)$ 为 $q_2 f(x) = f(x)+I_2$. □

若有环同态 $R_1 \xrightarrow{f} R_2 \xrightarrow{g} R_3$ 和理想 $I_i \subset R_i$ (其中 $i=1,2,3$), 使得 $f(I_1)\subset I_2$ 而 $g(I_2)\subset I_3$, 则交换图表

$$\begin{array}{ccccc} R_1 & \xrightarrow{f} & R_2 & \xrightarrow{g} & R_3 \\ {\scriptstyle q_1}\downarrow & & \downarrow{\scriptstyle q_2} & & \downarrow{\scriptstyle q_3} \\ R_1/I_1 & \xrightarrow[\overline{f}]{} & R_2/I_2 & \xrightarrow[\overline{g}]{} & R_3/I_3 \end{array}$$

和推论 6.1.11 对 $\overline{gf}$ 的刻画给出 $\overline{gf}=\overline{g}\circ\overline{f}$: 事实上, 两边都映 x_1+I_1 为 $g(f(x_1))+I_3$. 这是 §4.12 探讨商空间时见过的套路.

同样地, 可以比较商环 R/I 的理想与 R 的理想, 以及相应的商.

命题 6.1.12 设 I 是 R 的理想, $\overline{R}:=R/I$, 而商映射仍记为 $q:R\to\overline{R}$. 我们有双射

$$\begin{aligned} \{J\subset R:\text{理想},\ J\supset I\} &\xleftrightarrow{1:1} \{\overline{J}\subset\overline{R}:\text{理想}\} \\ J &\longmapsto \overline{J}:=q(J) \\ J:=q^{-1}(\overline{J}) &\longmapsfrom \overline{J}. \end{aligned}$$

此双射具有下述性质.

⋆ 它是严格保序的: $\overline{J}_1\supset\overline{J}_2$ 当且仅当 $J_1\supset J_2$.

⋆ 若 J 对应到 $\overline{J}$, 则有自然的环同构

$$\begin{aligned} R/J &\xrightarrow{\sim} \overline{R}/\overline{J} \\ x+J &\mapsto q(x)+\overline{J}. \end{aligned}$$

证明 和向量空间版本的命题 4.12.9 如出一辙. 先证双向箭头互逆. 给定 J, 我们断言 $q^{-1}(q(J))=J$. 首先 $\supset$ 是同义反复; 至于 $\subset$, 设 $x\in R$ 满足 $q(x)\in q(J)$, 则存在 $y\in J$ 使得 $q(y)=q(x)$, 亦即 $x\in y+\ker(q)=y+I$, 因而从 $I\subset J$ 可见 $x\in J$.

另一个方向相当于证 $q(q^{-1}(\overline{J}))=\overline{J}$. 一如既往 $\subset$ 是同义反复, 至于 $\supset$ 则缘于 q 的满性.

由于双向映射都保持理想的包含关系, 这对互逆映射严格保序.

最后, 定义同态 $f:R\to\overline{R}/\overline{J}$ 为 q 和商映射 $\overline{R}\to\overline{R}/\overline{J}$ 的合成, 映 x 为 $q(x)+\overline{J}$, 它作为满射的合成依然满, 核则是 $J:=q^{-1}(\overline{J})$. 因此 f 诱导同构 $\overline{f}:R/J\xrightarrow{\sim}\overline{R}/\overline{J}$, 映 $x+J$ 为 $q(x)+\overline{J}$. □

由于取商和满同态是一回事, 上述结果对环的任意满同态 $q:R\twoheadrightarrow\overline{R}$ 同样有效.

练习 6.1.13 设 R 是非零交换环.

(i) 说明 R 是域当且仅当它没有 $\{0\}$ 和 R 之外的理想. 〈提示〉 任何非零理想都包含形如 (x) 的理想, $x \neq 0$, 而 x 可逆等价于 $(x) = R$.

(ii) 设 F 是域, 证明任何环同态 $\varphi : F \to R$ 都是单的. 〈提示〉 考虑 $\ker(\varphi)$.

6.2 多项式的唯一分解性质

域上的一元多项式环有许多性质和整数环 $\mathbb{Z}$ 相近. 为了使概念清晰, 我们先从交换环中的整除关系入手, 继而探讨整环 (定义 3.1.11) 的面向, 层层递进.

首先设 R 为交换环, $x, y \in R$. 我们以符号 $x \mid y$ 代表存在 $d \in R$ 使得 $y = dx$; 读作 "x 整除 y".

交换环中的整除关系自然是以 $R = \mathbb{Z}$ 的情形为模板, 见 §2.7. 在关于 $\mathbb{Z}$ 的算术的研究中, 读者应该已经察觉到, 对于整除关系和因数分解的问题, 相差一个正负号并无区别. 注意到 $\mathbb{Z}$ 是整环而 $\mathbb{Z}^\times = \{\pm 1\}$; 外推到一般整环, 便引出如下的记法.

约定 6.2.1 设 R 为整环, $x, y \in R$. 若存在 $r \in R^\times$ 使得 $x = ry$, 则记为 $x \sim y$.

这是 R 上的等价关系. 诚然, 若存在 $r \in R^\times$ 使得 $x = ry$, 则 $y = r^{-1}x$, 故对称性成立; 从 $x = 1 \cdot x$ 和 $1 \in R^\times$ 可见反身性成立; 若 $x = ry$ 而 $y = sz$, 其中 $r, s \in R^\times$, 则 $x = rsz$ 而 $rs \in R^\times$, 故传递性成立.

此外, 易见 $x \sim 1$ 当且仅当 $x \in R^\times$.

从元素 x 过渡到它对 $\sim$ 的等价类, 相当于只看它生成的主理想 (x) (例 6.1.3), 而整除关系只和 $\sim$ 等价类或对应的主理想相关. 这是以下结果的内涵.

引理 6.2.2 *设 R 为整环, $x, y \in R$, 则*

- *$x \mid y$ 当且仅当 $(x) \supset (y)$;*
- *$x \sim y$ 当且仅当 x 和 y 相互整除, 当且仅当 $(x) = (y)$.*

证明 先处理第一条. 设 $y = dx$, 其中 $d \in R$, 则 $(y) = dxR \subset xR = (x)$. 反之, 若 $(y) \subset (x)$, 则 $y \in (y) \subset (x)$ 蕴涵存在 d 使得 $y = dx$.

其次处理第二条的前半段. 设 $x \sim y$, 写作 $y = rx$, 其中 $r \in R^\times$, 则 $x \mid y$. 基于 $\sim$ 的对称性, 同样有 $y \mid x$. 故两者相互整除.

设 x 和 y 相互整除, 则存在 $a, b \in R$ 使得 $y = ax$ 而 $x = by$. 于是 $abx = x$. 分两种情形讨论: 若 $x \neq 0$, 则因为 R 是整环, 必有 $ab = 1$, 于是 $a, b \in R^\times$ 而 $x \sim y$; 若 $x = 0$, 则 $y = ax = 0$, 此时 $x \sim y$ 仍然成立.

最后, 证明第一步的结果表明 x 和 y 相互整除相当于 $(x) = (y)$, 此即第二条的后半段. □

今后我们主要对 R 的非零元考虑它们的 $\sim$ 等价类.

定义 6.2.3 设 p 为整环 R 的非零元, $p \notin R^\times$.

- ⋆ 若 p 满足 $p \mid ab \iff (p \mid a) \vee (p \mid b)$, 则称 p 为**素元**.
- ⋆ 若 p 满足 $a \mid p \iff (a \sim p) \vee (a \sim 1)$, 则称 p 为**不可约元**.

不同于整数情形, 素元和不可约元在一般整环中并非等价的概念, 然而容易证明素元必不可约.

引理 6.2.4 设 p 为整环 R 的素元, 则 p 是不可约元.

证明 若 $a \in R$ 满足 $a \mid p$, 写作 $p = ab$, 则 $p \mid ab$. 若 $p \mid a$, 则 a 和 p 相互整除, 故引理 6.2.2 蕴涵 $a \sim p$. 若 $p \nmid a$, 则 $p \mid b$, 同样论证继而给出 $p \sim b$; 换言之, 存在 $r \in R^\times$ 使得

$$ab = p = rb.$$

因为素元 $p \neq 0$, 故 $b \neq 0$. 从上式两端消去 b 遂得到 $a = r \in R^\times$, 亦即 $a \sim 1$. 证毕. □

算术基本定理 2.7.6 的陈述在一般整环中可以提炼为以下形式.

定义 6.2.5 若对于整环 R 中的所有非零元 r, 都存在 $n \in \mathbb{Z}_{\geq 0}$ 和不可约元 $p_1, \cdots, p_n \in R$ 使得

$$r \sim p_1 \cdots p_n,$$

而且 $p_1, \ldots, p_n$ (计重数) 的 $\sim$ 等价类是由 r 唯一确定的, 至多差一个重排, 则称 R 是**唯一分解环**. 按惯例, $n = 0$ 时应将上式解读为 $r \sim 1$.

对于唯一分解环 R 中的任两个非零元 r, s, 取其唯一分解

$$r \sim \prod_i p_i^{a_i}, \quad s \sim \prod_i p_i^{b_i}, \quad a_i, b_i \in \mathbb{Z}_{\geq 0},$$

其中 $p_1, p_2, \ldots$ 是相对于 $\sim$ 互不等价的不可约元. 类似于整数的情形, 由此可以分别定义它们的最大公因数和最小公倍数为

$$\gcd(r, s) \sim \prod_{i=1}^{r} p_i^{\min\{a_i, b_i\}}, \quad \operatorname{lcm}(r, s) \sim \prod_{i=1}^{r} p_i^{\max\{a_i, b_i\}}.$$

上文所谓之 "数", 实际是 R 对 $\sim$ 的一个等价类, 而 "最大" 和 "最小" 自然是相对于整除而论. 这些定义也适用于两个以上, 甚至无穷多个元素. 若 $r_1, \ldots, r_n$ 的最大公因数 ~ 1, 则称它们**互素**.

和整数环之于有理数域的情形相似, R 的唯一分解性导致分式域 $\operatorname{Frac}(R)$ 的非零元具有本质上唯一的既约表法.

定义–命题 6.2.6 设 R 为唯一分解环, $h \in \mathrm{Frac}(R) \smallsetminus \{0\}$, 则存在 $f, g \in R$ 使得 $g \neq 0$ 而 f 和 g 互素, 而且 $h = \frac{f}{g}$. 这般分式 $\frac{f}{g}$ 称为**既约分式**.

倘若 $f_1, g_1 \in R$ 也满足 $g_1 \neq 0$ 而 $h = \frac{f_1}{g_1}$, 则必有 $f \mid f_1$ 和 $g \mid g_1$.

论证方式和整数情形是完全相同的. 作为推论, 既约分式表法中的资料 (f, g) 精确到 $R^\times$ 是由 h 唯一确定的.

唯一分解环的当然例子是 $\mathbb{Z}$. 基于平凡的理由, 域也是唯一分解环. 更进一步的例子则来自于多项式环.

简单起见, 以下仅探讨域 F 上的一元多项式环 $F[X]$; 引理 3.3.3 确保 $F[X]$ 是整环, 而且 $F[X]^\times = F^\times$.

注记 6.2.7 由 $F[X]^\times = F^\times$ 可见 $F[X]$ 上的关系 $\sim$ 再简单不过: $f \sim g$ 当且仅当它们相差一个非零常数倍. 因此从任何非零 $\sim$ 等价类中都可以挑出唯一一个首一多项式, 以充当等价类的标准代表元.

下一则结果是引理 2.7.2 的多项式版本, 证明方法也是类似的, 其要点在于多项式的带余除法 (命题 3.4.1).

引理 6.2.8 整环 $F[X]$ 的所有理想 I 都是主理想.

证明 若 $I = \{0\}$ 则毋须证明. 以下不妨设 $I \neq \{0\}$, 取 $f \in I$ 使得 $f \neq 0$ 而 $\deg f$ 极小, 兹说明 $I = (f)$: 基于理想的乘法封闭性, 包含关系 $(f) \subset I$ 是容易的. 对于另一方向, 任给 $g \in I$, 以带余除法表之为 $g = df + r$, 则加法封闭性蕴涵 $r = g - df \in I$, 然而 $\deg(r) < \deg(f)$, 故唯一可能是 $r = 0$, 从而 $g \in (f)$. □

引理 6.2.9 整环 $F[X]$ 的所有不可约元都是素元.

证明 和整数情形的命题 2.7.5 为同一思路: 设 p 是 $F[X]$ 的不可约元, $p \mid ab$. 考虑 p 和 a 在 $F[X]$ 中生成的理想 $\langle p, a\rangle$. 以引理 6.2.8 取 $f \in F[X]$ 使得 $(f) = \langle p, a\rangle$. 观察到

$$(f) \supset (a) \implies f \mid a, \quad (f) \supset (p) \implies f \mid p;$$

因为 p 不可约, 第二式进一步蕴涵 $f \sim p$ 或 $f \sim 1$.

- ⋆ 若 $f \sim p$ 则 $f \mid a$ 导致 $p \mid a$.
- ⋆ 若 $f \sim 1$ 则 $\langle p, a\rangle = F[X]$, 故存在 x, y 使得 $px + ay = 1$, 两边同乘以 b 得到 $p \mid pxb + aby = b$.

这便对 p 验证了素元所需的条件. □

配合引理 6.2.4, 我们便推知素元和不可约元在 $F[X]$ 中是一回事. 此一事实是证明唯一分解性的关键.

定理 6.2.10 (多项式环的算术基本定理) 整环 $F[X]$ 是唯一分解环.

证明　首先说明任意 $f \in F[X] \smallsetminus \{0\}$ 都能分解为不可约元的乘积. 如果 $f \in F[X]^\times = F^\times$ 或 f 不可约, 则无事可作. 在这两种状况以外, 存在 $a, b \in F[X]$ 使得 $f = ab$, 而且 $a, b \notin F^\times$. 此时 $\deg a$ 和 $\deg b$ 皆严格小于 $\deg f$. 由此递归地将 a 和 b 分解为不可约元的乘积即可.

其次说明分解唯一. 将不可约分解表作 $f \sim p_1^{a_1} \cdots p_r^{a_r}$ 的形式, 其中 $p_1, \ldots, p_r$ 是相对于 $\sim$ 互不等价的不可约元, $a_1, \ldots, a_r \in \mathbb{Z}_{\geq 1}$. 给定 f 的两个不可约分解

$$p_1^{a_1} \cdots p_r^{a_r} \sim f \sim q_1^{b_1} \cdots q_s^{b_s}.$$

如果 $r = 0$, 则左式按规定是 1, 故右式可逆, 因而其中出现的每个因子都可逆, 唯一可能是 $s = 0$; 根据对称性, 同样有 $s = 0$ 蕴涵 $r = 0$. 故以下不妨设 $r, s \geq 1$.

考虑 p_1, 引理 6.2.9 说明它也是素元, 并且整除右式, 于是 p_1 必然整除某一个 q_j, 从而 $p_1 \sim q_j$. 适当重排右式下标, 不妨设 $p_1 = q_1$; 必要时左右两边互换, 不妨假设 $a_1 \leq b_1$. 两头消去 $p_1^{a_1}$ (根据整环和 $\sim$ 的定义, 这是合理的), 可得

$$p_2^{a_2} \cdots p_r^{a_r} \sim p_1^{b_1 - a_1} q_2^{b_2} \cdots q_s^{b_s}.$$

因为 $p_1 \not\sim p_2, \ldots, p_r$ 而且 p_1 是素元, 它不整除左式, 由此可得 $b_1 = a_1$. 递归地论证便有不可约分解的唯一性. □

约定 6.2.11　若多项式 f, g 满足 $f \mid g$, 则称 f 是 g 的**因式**. 多项式环 $F[X]$ 中的不可约元 (等价地说, 素元) 称为**不可约多项式**, 定理 6.2.10 的分解称为多项式的素因式分解.

多项式环也有一些不共于 $\mathbb{Z}$ 的算术性质. 数学分析中常用的部分分式分解便是一则简单例子, 陈述如下.

定理 6.2.12 (部分分式分解)　设 $f, g \in F[X], g \neq 0$ 而且 $g = g_1 \cdots g_n$, 其中 $g_1, \ldots, g_n \in F[X]$ 两两互素, 则存在唯一的 $q, h_1, \ldots, h_n \in F[X]$ 使得

- $h_i \in F[X]$ 满足 $\deg h_i < \deg g_i$, 其中 $1 \leq i \leq n$;
- 我们有 $F(X)$ 中的等式
$$\frac{f}{g} = q + \sum_{i=1}^{n} \frac{h_i}{g_i}.$$

证明　首先处理存在性. 两两互素的条件蕴涵

$$\hat{g}_1 := g_2 \cdots g_n, \quad \hat{g}_2 := g_1 g_3 \cdots g_n, \quad \ldots, \quad \hat{g}_n := g_1 \cdots g_{n-1}$$

互素, 尽管它们不是两两互素; 这点可以从唯一分解性读出. 存在 $k_1, \ldots, k_n \in F[X]$ 使 $\sum_{i=1}^n k_i \hat{g}_i = 1$. 于是

$$\frac{f}{g} = \frac{f}{g} \cdot \sum_{i=1}^{n} k_i \hat{g}_i = \sum_{i=1}^{n} \frac{f k_i}{g_i}.$$

将每个 fk_i 以带余除法写成 $q_ig_i + h_i$ 的形式, 其中 $q_i, h_i \in F[X]$, $\deg h_i < \deg g_i$, 便得到所求的分解

$$\frac{f}{g} = \underbrace{\sum_{i=1}^{n} q_i}_{=:q} + \sum_{i=1}^{n} \frac{h_i}{g_i}.$$

至于唯一性, 设有分解 $\frac{f}{g} = q + \sum_{i=1}^{n} \frac{h_i}{g_i}$. 取定 $k_1, \dots, k_n$ 如上, 则对于每个 $1 \le i \le n$ 皆有

$$\begin{aligned} fk_i &= gqk_i + \sum_{j=1}^{n} h_j \hat{g}_j k_i \\ &= g_i \cdot \text{多项式} + h_i k_i \hat{g}_i \\ &= g_i \cdot \text{多项式} + h_i \left(1 - \sum_{j \ne i} k_j \hat{g}_j\right) \\ &= g_i \cdot \text{另一个多项式} + h_i. \end{aligned}$$

这便确定 h_i 为 fk_i 除以 g_i 的余式. 故 $q = \frac{f}{g} - \sum_{i=1}^{n} \frac{h_i}{g_i}$ 也是唯一确定的. □

如果定理 6.2.12 中的 g_i 形如 $(X - a_i)^{d_i}$, 则通过将 h_i 表成 $\sum_{k=0}^{d_i-1} c_k (X - a_i)^k$, 分解的第 i 项可以进一步表成

$$\frac{h_i}{g_i} = \frac{c_{d_i-1}}{X - a_i} + \cdots + \frac{c_0}{(X - a_i)^{d_i}}.$$

这是部分分式分解在应用中所常见的形式. 证明中关于 h_i 的描述仅是理论性的, 并未给出最高效的算法. 在具体计算时, 以待定系数法求 $c_0, \dots, c_{d_i-1}$ 通常更方便.

最后, 我们在 §3.1 讨论过商环 $\mathbb{Z}/N\mathbb{Z}$ 的结构. 基于 $F[X]$ 与 $\mathbb{Z}$ 的类比, 对商环 $F[X]/(f)$ 也能提出种种类似的问题, 比如说, 整数环上的中国剩余定理 3.2.8 有无多项式环的版本? 答案是肯定的. 这些问题更适合在主理想环的抽象框架下处理, 这是下一节的主旨.

6.3 简单推广: 主理想环的唯一分解性

整数和多项式的算术可以放在更广泛的环论背景下来考察. 这非但不增加论证难度, 反而有助于理清思路.

定义 6.3.1 *设 R 为整环. 若 R 的所有理想都是例 6.1.3 所谓的主理想, 则称 R 为* **主理想环**.

整数环 $\mathbb{Z}$ (引理 2.7.2) 和域 F 上的多项式环 $F[X]$ (引理 6.2.8) 都是主理想环的实例.

在证明多项式环的算术基本定理 6.2.10 时, 实际运用的只是 $F[X]$ 的抽象环论性质, 不涉及多项式的具体操作. 现将所需的环论性质提炼为两条.

命题 6.3.2 整环 R 是唯一分解环当且仅当以下条件成立.

- ⋆ 所有 $r \in R \setminus \{0\}$ 都能写成不可约元的乘积.
- ⋆ 所有不可约元都是素元.

证明 对于 "仅当" 方向, 不可约分解的存在性直接来自唯一分解环定义, 下面说明不可约元必为素元. 设 p 不可约, $p \mid ab$, 取不可约分解 $a = q_1 \cdots q_m$ 和 $b = r_1 \cdots r_n$, 则 $q_1 \cdots q_m r_1 \cdots r_n$ 是 ab 的不可约分解. 另一方面, 将 $\frac{ab}{p}$ 的不可约分解乘上 p, 即得 ab 的另一个不可约分解, 故唯一分解性质确保 p 必须出现在 $q_1, \ldots, q_m, r_1, \cdots r_n$ 之中, 精确到 $\sim$. 这便说明 $p \mid a$ 或 $p \mid b$.

至于 "当" 的方向, 请细观定理 6.2.10 的证明. 不可约分解的存在性只用到第一条, 唯一性只用到第二条, 后者在 $F[X]$ 的情形来自引理 6.2.9. □

以下的结果称为主理想环的 Noether 性质, 它是环论的基本概念之一.

引理 6.3.3 设 R 为主理想环, 而 $(I_n)_{n=1}^{\infty}$ 是 R 的一列理想, 满足 $I_1 \subset I_2 \subset I_3 \subset \cdots$. 此时对充分大的 $n \in \mathbb{Z}_{\geq 1}$ 必有 $I_n = I_{n+1} = \cdots$

证明 命 $I := \bigcup_{n=1}^{\infty} I_n$, 它是 R 的理想, 具体验证如下: 若 $x \in I_a$ 而 $y \in I_b$, 则 $x + y \in I_{\max\{a,b\}}$, 而且对所有 $r \in R$ 皆有 $rx \in I_a$.

因此可取 $h \in R$ 使得 $I = (h)$. 既然 I 是并, 必然有 n 使得 $h \in I_n$, 从而 $I \subset I_n$. 按定义又有 $I_n \subset I_{n+1} \subset \cdots \subset I$, 综之 $I_n = I_{n+1} = \cdots = I$. □

上述性质相当于说主理想环中的理想升链总会停止.

定理 6.3.4 若 R 是主理想环, 则 R 是唯一分解环.

证明 首先以反证法说明所有 $r \in R \setminus \{0\}$ 都能写成不可约元的乘积. 命 $r_0 := r$. 设 r_0 无不可约分解, 则 r_0 必可约, 故有 $r_0 = r_1 s_1$, 其中 $r_1, s_1 \not\sim r_0$. 引理 6.2.2 蕴涵 $(r_1) \supsetneq (r_0)$ 和 $(s_1) \supsetneq (r_0)$.

若 r_1 和 s_1 都有不可约分解, 则 r_0 亦然, 矛盾. 故不失一般性可设 r_1 无不可约分解, r_1 必可约, 故进一步有 $r_1 = r_2 s_2$, 其中 $r_2, s_2 \not\sim r_1$. 于是 $(r_2) \supsetneq (r_1)$ 而 $(s_2) \supsetneq (r_1)$. 同样地, 不失一般性可设 r_2 无不可约分解, 然后继续操作.

综上可见若 r 无不可约分解, 则有主理想构成的严格升链

$$(r_0) \subsetneq (r_1) \subsetneq (r_2) \subsetneq \cdots,$$

这与引理 6.3.3 矛盾.

为了说明 R 的不可约元都是素元, 论证与多项式情形的引理 6.2.9 毫无差别, 需要的只是主理想环的定义.

将上述结果代入命题 6.3.2, 即知 R 是唯一分解环. □

命题 2.7.3 在一般的主理想环中也有相应的推广.

命题 6.3.5 设 R 为主理想环. 元素 $r_1,\dots,r_n$ 互素当且仅当 $\langle r_1,\dots,r_n\rangle = R$, 或者等价地说, 存在 $s_1,\dots,s_n \in R$ 使得 $\sum_{i=1}^n r_i s_i = 1$.

证明 取 $h \in R$ 使得 $\langle r_1,\dots,r_n\rangle = (h)$. 设 $r_1,\dots,r_n$ 互素, 则由 $r_i \in (h)$ 知 $h \mid r_i$ 对所有 i 成立, 从而互素性质导致 $h \sim 1$, 换言之 $(h) = R$.

反之设 $(h) = R$, 则存在 $s_1,\dots,s_n \in R$ 使得 $\sum_{i=1}^n r_i s_i = 1$. 于是 $r_1,\dots,r_n$ 的任意公因子 a 也整除 1, 故 $a \sim 1$, 由此推知 $r_1,\dots,r_n$ 互素. □

不妨将 $\langle r_1,\dots,r_n\rangle = R$ 当作是互素在主理想环中的定义, 这样做的一个细微好处是它直接适用于 $r_1,\dots,r_n$ 包含零元的情形.

接着探讨主理想环的商环.

命题 6.3.6 设 R 为主理想环, $t \in R \setminus \{0\}$ 而且 $t \notin R^\times$. 以下性质等价:

(i) $R/(t)$ 是域;

(ii) $R/(t)$ 是整环;

(iii) t 是素元, 或者等价地说 t 不可约 (见命题 6.3.2).

证明 (i) $\Longrightarrow$ (ii) 是平凡的.

(ii) $\Longrightarrow$ (iii): 设 $R/(t)$ 是整环而 $t = fg$, 则 $R/(t)$ 中的等式

$$(f+(t))(g+(t)) = fg+(t) = 0+(t)$$

蕴涵 $t \mid f$ 或 $t \mid g$ 至少有一者成立; 另一方面 f 和 g 皆整除 t, 这就说明或者 $t \sim f$, 或者 $t \sim g$. 因此 t 是素元.

(iii) $\Longrightarrow$ (i). 设 t 不可约, 故 $R \neq (t)$. 考虑 $R/(t)$ 的非零元 $f+(t)$, 换言之 $f \in R$ 而 $t \nmid f$. 此时 t 和 f 互素, 故命题 6.3.5 给出 $h,k \in R$ 使得

$$hf + kt = 1.$$

因此 $(h+(t))(f+(t)) = 1+(t)$. 综上, $R/(t)$ 的非零元皆可逆. □

练习 6.3.7 原封不动地照搬命题 2.7.3 的论证, 进一步证明在主理想环 R 中

$$\langle r_1,\dots,r_n\rangle = \gcd(r_1,\dots,r_n)R;$$

此式也可以当作 $\gcd(r_1,\dots,r_n)$ 在主理想环中的定义.

我们也得到中国剩余定理 3.2.8 的一则推广, 它适用于所有主理想环, 包括域上的多项式环 $F[X]$. 更广的版本可见 [10, 定理 5.5.2].

定理 6.3.8 (主理想环的中国剩余定理) 设 R 为主理想环, $a_1,\dots,a_n \in R\setminus\{0\}$ 两两互素, $a := a_1\cdots a_n$, 则有环同构

$$\begin{aligned}\varphi : R/(a) &\longrightarrow \textstyle\prod_{i=1}^n R/(a_i)\\ r+(a) &\longmapsto (r+(a_i))_{i=1}^n .\end{aligned}$$

证明 首先处理 $n=2$ 情形. 为证 φ 单, 观察到 $\varphi(r+(a))=(0,0)$ 等价于 $a_1 \mid r$ 而且 $a_2 \mid r$, 而互素条件说明这等价于 $a = a_1a_2 \mid r$.

为证 φ 满, 取 $x_1 \in (a_1)$ 和 $x_2 \in (a_2)$ 使得 $x_1+x_2=1$, 则对于所有 $r\in R$, 从 $rx_1+rx_2=r$ 可推得 $\varphi(rx_1+(a))=(0,r+(a_2))$ 和 $\varphi(rx_2+(a))=(r+(a_1),0)$ 成立, 从而推得满性.

对于 $n\ge 3$ 情形, 按照 $R/(a) \xrightarrow{\sim} R/(a_1\cdots a_{n-1})\times R/(a_n) \xrightarrow{\sim}\cdots$ 逐步化约即足, 请读者验证所需的细节. □

6.4 形式求导

实多项式的求导运算起源于数学分析, 由于它的公式是代数地给出的, 可以对任意域 F 上的多项式来定义, 并且具有和分析学中相类似的性质.

定义 6.4.1 (形式求导) 设 $f=\sum_{n\ge 0} a_nX^n \in F[X]$, 定义 f 的形式导数为

$$f' := \sum_{n\ge 1} na_nX^{n-1} \ \in F[X],$$

其中 $na_n \in F$ 按照 (3.1.1) 的方式理解. 递归地按照

$$f^{(0)}=f, \quad f^{(m)} := (f^{(m-1)})' \quad (m\in\mathbb{Z}_{\ge 1})$$

定义 f 的任意阶导数; 我们也记 $f''=(f')'$, 依此类推.

数学分析中的求导规律在此同样适用. 对于所有 $f,g\in F[X]$, 我们有

$$\begin{aligned}(f+g)' &= f'+g',\\ (fg)' &= f'g+fg',\\ c' &= 0, \quad c\in F,\\ (cf)' &= cf', \quad c\in F.\end{aligned}$$

第一条性质是定义的直接结论, 第三条和第四条性质同样是明显的. 第二条性质称为 Leibniz 律, 仍可按定义直接验证, 但以下手法兴许更方便. 首先注意到如果

$f=\sum_{i=1}^{a} f_i$, $g=\sum_{j=1}^{b} g_j$, 而且 $(f_ig_j)'=f_i'g_j+f_ig_j'$ 对所有 i,j 都成立, 则

$$(fg)'=(\sum_{i,j} f_ig_j)'=\sum_{i,j}(f_ig_j)'=\sum_{i,j} f_i'g_j+\sum_{i,j} f_ig_j'$$
$$=\sum_j\left(\sum_i f_i'\right)g_j+\sum_i f_i\left(\sum_j g_j'\right)=f'g+fg'.$$

按此观察, 容易将 Leibniz 律化约到 $f=X^h$ 而 $g=X^k$ 的简单情形来验证.

以下公式是递归地应用 Leibniz 律的产物, 同样对应到数学分析的标准结果.

$$\left(f^k\right)'=kf^{k-1}f',\quad k\in\mathbb{Z}_{\geq 1},\ f\in F[X]. \tag{6.4.1}$$

推论 6.4.2 *求导映射 $f\mapsto f'$ 是从 F-向量空间 $F[X]$ 到其自身的线性映射.*

需要格外留意的是 $c'=0$ 未必蕴涵 $c\in F$, 这取决于域的特征.

命题 6.4.3 *性质 $f'=0\iff f\in F$ 在 $F[X]$ 中成立的充要条件是 $\mathrm{char}(F)=0$.*

证明 设 $f=\sum_{n\geq 0}a_nX^n$, 则 $f'=0$ 等价于 $na_n=0$ 对所有 $n\geq 1$ 成立. 若 $\mathrm{char}(F)=0$, 则等式左边可以消去 $n\cdot 1_F\in F^\times$, 给出 $n\geq 1\implies a_n=0$, 亦即 $f\in F$.

另一方面, 若 $\mathrm{char}(F)=p>0$, 则形如 $f=\sum_{n\geq 0}a_nX^{pn}$ 的多项式都满足 $f'=\sum_{n\geq 1}pna_nX^{pn-1}=0$. □

练习 6.4.4 *进一步说明当 $\mathrm{char}(F)=p>0$ 时, 满足 $f'=0$ 的多项式正好是形如 $f(X)=\sum_{n\geq 0}a_nX^{pn}=g(X^p)$ 的多项式, 其中 $g:=\sum_{n\geq 0}a_nX^n$.*

形式求导可以从多项式环 $F[X]$ 延拓到有理函数域 $F(X)$ 上. 从数学分析的观点, 这是毫不意外的.

定义–命题 6.4.5 *存在唯一的映射*

$$F(X)\to F(X)$$
$$f\mapsto f'$$

使得它限制在 $F[X]$ 上给出多项式的形式求导, 而且对所有 $f,g\in F(X)$ 皆有

$$(f+g)'=f'+g',\quad (fg)'=f'g+fg'.$$

事实上, 它可以精确地由下式给出:

$$\left(\frac{f}{g}\right)'=\frac{f'g-fg'}{g^2},\quad f,g\in F[X],\ g\neq 0.$$

证明 首先说明唯一性. 对任意非零之 $g\in F[X]$, 从

$$0=1'=\left(g\cdot\frac{1}{g}\right)'=\frac{g'}{g}+g\left(\frac{1}{g}\right)'$$

可得 $\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$. 进一步, 对于任意 $f \in F[X]$, 我们有

$$\left(\frac{f}{g}\right)' = \left(f \cdot \frac{1}{g}\right)' = \frac{f'}{g} - \frac{fg'}{g^2} == \frac{f'g - fg'}{g^2}.$$

问题在于反过来说明按上式定义的 $\left(\frac{f}{g}\right)'$ 只和 $\frac{f}{g}$ 相关, 不依赖 f, g 的选取, 而且它满足列出的所有性质. 基于 §3.5 对分式域的定义, 这些验证完全是例行公事, 请读者应用对多项式形式求导的性质予以验证. □

例 6.4.6 (对数求导) 数学分析或复分析经常考虑的一类导函数形如 $\frac{f'}{f}$, 可将其设想为 $\log(f)'$, 其中 log 是自然对数. 尽管取 log 在代数框架下颇成问题, 但对于非零的 $f \in F[X]$, 考虑对数求导 $\frac{f'}{f}$ 是完全合理的, 而且它仍然有良好的性质.

- ⋆ 化乘为加: $\frac{(fg)'}{fg} = \frac{f'}{f} + \frac{g'}{g}$, 这是 Leibniz 律的直接结论.
- ⋆ 若 f 分解成 $p_1^{a_1} \cdots p_r^{a_r}$, 其中 $p_1, \ldots, p_r \in F[X]$ 非零而 $a_1, \ldots, a_r \in \mathbb{Z}_{\geq 1}$, 则

$$\frac{f'}{f} = \sum_{i=1}^{r} a_i \cdot \frac{p_i'}{p_i}. \tag{6.4.2}$$

诚然, 对数求导化乘为加, 所以问题立即简化到 $r = 1$, $a_1 = 1$ 的情形, 进而化为同义反复.

练习 6.4.7 证明高次形式求导具有如下的 Leibniz 律:

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

形式求导自然地推及多变元情形, 类似于数学分析中的求偏导运算.

定义 6.4.8 (形式偏导数) 对于表作有限和的 n 元多项式

$$f = \sum_{i_1, \ldots, i_n \geq 0} c_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_n^{i_n} \in F[X_1, \ldots, X_n]$$

和 $1 \leq k \leq n$, 我们借用分析学的符号来定义

$$\frac{\partial f}{\partial X_k} := \sum_{\substack{i_1, \ldots, i_n \geq 0 \\ i_k \neq 0}} i_k c_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_k^{i_k - 1} \cdots X_n^{i_n}.$$

练习 6.4.9 (Euler 恒等式) 设 $m \in \mathbb{Z}_{\geq 0}$ 而 $f \in F[X_1, \ldots, X_n]$ 是 m 次齐次的, 换言之

$$f = \sum_{i_1 + \cdots + i_n = m} c_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_n^{i_n}.$$

验证

$$\sum_{k=1}^{n} X_k \frac{\partial f}{\partial X_k} = mf.$$

6.5 应用: Mason–Stothers 定理

令 F 为域. 先前已提及整数环 $\mathbb{Z}$ 和多项式环 $F[X]$ 有许多相似的算术性质. 此一类比意义深远. 一般而言, 多项式环上的算术问题往往比整数版本容易, 这是因为多项式具有更多的操作, 例如求导. 本节介绍的 Mason–Stothers 定理是这方面的一个初等结果, 由此可以容易地推导 Fermat 大定理的多项式版本: 当 $n \geq 3$ 时, 等式 $u^n + v^n = w^n$ 没有非平凡的多项式解, 前提是 u, v, w 互素而且形式导数非零.

定义 6.5.1 对于非零多项式 $f \in F[X]$, 定义 f 的**根基**为

$$\operatorname{rad}(f) := \prod_{\substack{p \mid f \\ \text{不可约首一多项式}}} p;$$

当 $f \in F^\times$ 时, 右式理解为 1.

引理 6.5.2 设 $f, g \in F[X]$ 互素, 而且 $f' \neq 0$, $g' \neq 0$, 则 $\frac{f'}{f} \neq \frac{g'}{g}$.

证明 若 $\frac{f'}{f} = \frac{g'}{g}$, 则由 $f \mid f'g - fg' = 0$ 和 f, g 互素可见 $f \mid f'$. 然而 $f' \neq 0$ 而 $\deg f' < \deg f$, 此无可能. □

留意到当 $\operatorname{char}(F) = 0$ 时, 命题 6.4.3 确保 $f' \neq 0 \iff f \notin F$.

定理 6.5.3 (R. C. Mason, W. W. Stothers, 1981) 取 $a, b, c \in F[X]$ 为两两互素的多项式, a', b', c' 皆非零, 并且满足 $a + b + c = 0$, 则

$$\max\{\deg a, \deg b, \deg c\} \leq \deg \operatorname{rad}(abc) - 1.$$

证明 形式求导得到 $a' + b' + c' = 0$. 因此有理函数域 $F(X)$ 上的齐次线性方程组

$$\begin{aligned} 1 \cdot X &+ 1 \cdot Y &+ 1 \cdot Z &= 0 \\ \tfrac{a'}{a} \cdot X &+ \tfrac{b'}{b} \cdot Y &+ \tfrac{c'}{c} \cdot Z &= 0 \end{aligned}$$

有非零解 (a, b, c). 为了求出其通解, 我们对系数矩阵进行简单的 Gauss–Jordan 消元法.

首先, 注意到系数矩阵的两行不可能成比例, 否则必有 $\frac{a'}{a} = \frac{b'}{b} = \frac{c'}{c}$, 这与两两互素的条件相矛盾, 见引理 6.5.2. 原方程的解空间因而是 1 维的. 消元法表明解空间由非零向量

$$\left(\frac{c'}{c} - \frac{b'}{b},\ \frac{a'}{a} - \frac{c'}{c},\ \frac{b'}{b} - \frac{a'}{a}\right)$$

生成; 细节请读者操作. 和解 (a, b, c) 比较, 可知存在 $\lambda \in F(X)^\times$ 使

$$\lambda \cdot (a, b, c) = \left(\frac{c'}{c} - \frac{b'}{b},\ \frac{a'}{a} - \frac{c'}{c},\ \frac{b'}{b} - \frac{a'}{a}\right). \tag{6.5.1}$$

按定义 3.5.7 及其后所述的性质比较有理函数的次数, 可见

$$\forall f \in \{a,b,c\}, \quad \deg\lambda + \deg f \leq -1. \tag{6.5.2}$$

第二点观察是 (6.4.2) 蕴涵 $\frac{a'}{a}$ 表作 $\frac{p_i'}{p_i}$ 的整系数线性组合, 其中 $p_1,p_2,\ldots$ 遍历 a 的素因子. 对 $\frac{b'}{b}$ 和 $\frac{c'}{c}$ 也是如此, 由此可见

$$\mathrm{rad}(abc)\lambda\cdot(a,b,c) = \mathrm{rad}(abc)\cdot\big(\text{(6.5.1) 的右式}\big) \in F[X]^3.$$

鉴于唯一分解性, 上式蕴涵 $\mathrm{rad}(abc)\lambda$ 的既约分式的分母同时整除 a, b, c. 然而 a,b,c 互素, 故 $\mathrm{rad}(abc)\lambda \in F[X]$. 而这又说明

$$\deg\mathrm{rad}(abc) \geq -\deg\lambda.$$

代入 (6.5.2) 便是所求的不等式. □

推论 6.5.4 (Fermat 大定理的多项式版本) *设 $n\in\mathbb{Z}_{\geq 1}$, 而且 $\mathrm{char}(F)\nmid n$. 若存在两两互素的多项式 $u,v,w\in F[X]$ 使得 u',v',w' 皆非零, 而且满足 $u^n+v^n=w^n$, 则必有 $n<3$.*

证明 在定理 6.5.3 中代入 $(a,b,c)=(u^n,v^n,-w^n)$; 根据 (6.4.1) 和关于 $\mathrm{char}(F)$ 的条件, 它们的形式导数皆非零, 由此可得

$$n\cdot\max\{\deg u,\deg v,\deg w\} \leq \deg\mathrm{rad}\,(u^nv^n(-w)^n) - 1.$$

根据根基的定义, 上式右侧等于 $\deg\mathrm{rad}(uvw)-1$. 另一方面, 假若 $n\geq 3$, 则

$$\deg\mathrm{rad}(uvw) \leq \deg(uvw) \leq n\cdot\max\{\deg u,\deg v,\deg w\},$$

显然矛盾. □

当 $\mathrm{char}(F)=0$ 时, 导数非零的条件也可以改写为 $u,v,w\notin F$.

练习 6.5.5 试说明导数非零和 $\mathrm{char}(F)\nmid n$ 的必要性.

提示 用练习 3.7.3: 设 p 为任意素数, 在满足 $\mathrm{char}(F)=p$ 的域上有 $u^p+v^p=(u+v)^p$.

6.6 根和重因式

仍然固定域 F 并考虑其上的多项式环 $F[X]$. 本节旨在细化 §3.4 关于根的讨论.

我们现在知道 $F[X]$ 是唯一分解整环 (定理 6.2.10). 基于次数理由, 一次多项式 $X-a$ 当然不可约; 因此 $X-a$ 与 $h\in F[X]$ 互素当且仅当 $X-a\nmid h$, 当且仅当 $h(a)\neq 0$ (余式定理). 当 $a\in F$ 给定, 任意非零的 $f\in F[X]$ 都能表成

$$f = (X-a)^{m_a}h, \quad m_a\in\mathbb{Z}_{\geq 0},\ h\in F[X], \tag{6.6.1}$$

其中 h 与 $X-a$ 互素, 亦即 $h(a)\neq 0$. 指数 m_a 由 f 和 a 唯一确定; $m_a>0$ 当且仅当 a 是 f 的根.

定义 6.6.1 分解式 (6.6.1) 中的 m_a 称为 a 在非零多项式 f 中的**重数**.

若 (6.6.1) 中的 h 还有一次因式, 换言之它在 F 中还有根, 则可以继续类似的操作, 最后得到分解

$$f=(X-a_1)\cdots(X-a_m)r,$$

其中 $m\in\mathbb{Z}_{\geq 0}$, $a_1,\ldots,a_m\in F$, 容许重复, 而 $r\in F[X]$ 在 F 中无根. 若 $r\notin F$, 则还能将 r 进一步分解, 但它的不可约因式至少是二次的.

综上, 若能确定 f 在 $F[X]$ 中的不可约分解, 或者至少是其中涉及一次因式的部分, 便能够确定 f 在 F 中的所有根及其重数. 我们在谈论多项式的根时经常将重数计入, 比如 $(X-1)^2$ 的根是 $1,1$ (计重数).

根和域的选择大有关系. 一个多项式可以在一个域 F 上无根, 而在足够大的扩域 $E\supset F$ 上有根. 所有根都落在 F 中的多项式是特别容易处理的.

定义 6.6.2 如果域 F 上的多项式 f 在 $F[X]$ 中分解为一次因式的积, 则称 f **分裂**.

命题 6.6.3 设非零多项式 f 在 F 中的根为 $a_1,\ldots,a_m$ (计重数), 则 $0\leq m\leq\deg f$; 我们有 $m=\deg f$ 当且仅当 f 或者是常数多项式, 或者分裂.

证明 注意到 f 是常数的情形对应到 $m=0=\deg f$. 对于 $\deg f\geq 1$ 的情形, 不失一般性可设 f 首一; 取素因子分解

$$f=(X-a_1)\cdots(X-a_m)h_1\cdots h_k,$$

其中 $k\in\mathbb{Z}_{\geq 0}$, 每个 h_i 都是至少二次的不可约多项式. 那么

$$\deg f=m+\sum_{i=1}^{k}\deg h_i\geq m,$$

等号成立当且仅当 $k=0$, 亦即 f 的素因子都是一次的. □

这细化了命题 3.4.4 的陈述.

练习 6.6.4 证明奇数次实多项式必有实根. 〈提示〉 设 $f\in\mathbb{R}[X]$ 而 $\lambda\in\mathbb{C}$, 则 $f(\overline{\lambda})=\overline{f(\lambda)}$, 因此 f 的非实根能成对地从 f 提出.

对于代数学的研究而言, 能使每个非常数多项式都有根的域是再好不过的. 这种域称为代数闭域.

定义 6.6.5 (代数闭域) 如果域 F 上的每个非常数多项式皆分裂, 则称 F 是代数闭域.

例 6.6.6 代数基本定理 1.1.1 断言复数域 $\mathbb{C}$ 是代数闭域.

代数基本定理本质上是分析学的结果, 而一则纯粹代数的事实则是所有域 F 都能嵌入一个代数闭域 $\overline{F}$, 其构造可以粗略地设想为不断向 F 添入所有多项式的根; 本书对此不拟细说, 详阅 [10, §8.2]. 至于单个非常数多项式 f 的情形, §6.10 将说明如何沿添根的思路来构造域的嵌入 $F \hookrightarrow E$, 使得 f 在 E 上分裂.

例 6.6.7 (Lagrange 插值法) 多项式插值问题常见于各种应用场景. 给定两两相异的 $x_1, \dots, x_n \in F$ 和任意的 $y_1, \dots, y_n \in F$, 问题是求多项式 $f \in F[X]$ 使得

$$\deg f \leq n-1, \quad \forall 1 \leq i \leq n,\ f(x_i) = y_i.$$

所有次数 $\leq n-1$ 的多项式构成 F-向量空间 $F[X]_{\leq n-1}$, 它以 $1, \dots, X^{n-1}$ 为有序基. 从 $x_1, \dots, x_n$ 可定义线性映射

$$\begin{aligned} \epsilon : F[X]_{\leq n-1} &\longrightarrow F^n \\ f &\longmapsto (f(x_1), \dots, f(x_n)). \end{aligned}$$

所以问题相当于研究 $\epsilon^{-1}(y_1, \dots, y_n)$, 或者说是解对应的线性方程组.

假如 $\epsilon(f) = (0, \dots, 0)$, 则 f 是有 n 个相异根 $x_1, \dots, x_n$ 而次数 $\leq n-1$ 的多项式, 故 $f = 0$. 这说明 $\ker(\epsilon) = \{0\}$. 又由于两边维数相同, ϵ 实际还是同构. 这就给出了插值问题的抽象解答: 对所有 i 都满足 $f(x_i) = y_i$ 的 $f \in F[X]_{\leq n-1}$ 存在且唯一. 问题在于求具体解. J.-L. Lagrange 提供的答案是

$$f = \sum_{i=1}^{n} y_i \mathcal{L}_{n,i}, \quad \mathcal{L}_{n,i} := \prod_{\substack{1 \leq k \leq n \\ k \neq i}} \frac{X - x_k}{x_i - x_k} \in F[X]_{\leq n-1}.$$

直接代值可见 $\mathcal{L}_{n,i}(x_i) = 1$ 而当 $j \neq i$ 时 $\mathcal{L}_{n,i}(x_j) = 0$, 所以上式确实成立.

线性方程组的理论也可以为 Lagrange 插值多项式给出一个自然的推导. 设所求之 f 为 $c_0 + c_1 X + \cdots + c_{n-1} X^{n-1}$, 则关于系数 $(c_i)_{i=0}^{n-1}$ 的条件化为线性方程组

$$\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

每个 c_{j-1} 都可以用推论 5.7.6 的 Cramer 法则表出. 精确到转置, Cramer 公式中的分母是命题 5.5.5 的 Vandermonde 行列式, 值为 $\prod_{j<i}(x_i - x_j) \neq 0$. 分子则可以按行列式的第 j 列展开, 继而调整为更小的 Vandermonde 行列式. 可以验证这确实给出 Lagrange 的公式 (当然如此), 细节留给感兴趣的读者.

从数学分析的观点来看, 实多项式 f 在 $a \in \mathbb{R}$ 处有重根相当于说 f 在 a 处与 X 坐标轴相切, 或者说 $f'(a) = f(a) = 0$. 此一观察可以推广到一般域上以及重因式的情形.

设 $f \in F[X]$ 非零. 若存在不可约多项式 p 使得 $p^2 \mid f$, 则称 f 有**重因式**. 重因式的存在性可以通过不可约分解来剖析. 首先, 适当乘以 $F^\times$ 的元素后, 不妨假设 f 是首一多项式, 带有分解

$$f = p_1^{e_1} \cdots p_n^{e_n}, \quad e_i \in \mathbb{Z}_{\geq 1}, \tag{6.6.2}$$

其中 $p_1, \ldots, p_n$ 两两相异, 首一不可约. 因此 f 无重因式当且仅当 $e_1 = \cdots = e_n = 1$. 形式求导的 Leibniz 律蕴涵

$$f' = \sum_{i=1}^{n} e_i p_i^{e_i - 1} p_i' \prod_{j \neq i} p_j^{e_j}.$$

命题 6.6.8 设 $f \in F[X]$ 非零.

(i) 若 f 和 f' 互素, 则 f 无重因式.

(ii) 若 f 无重因式, 而且它的每个素因式 p 都满足 $p' \neq 0$ (这在 $\mathrm{char}(F) = 0$ 时是自动的), 则 f 和 f' 互素.

证明 不妨设 f 是首一多项式, 并且取 (6.6.2) 的分解. 既然 $p_1, \ldots, p_n$ 不可约而且相异, f' 的公式表明对每个 i 都有

$$p_i \mid f' \iff p_i \mid e_i p_i^{e_i - 1} p_i'.$$

设 f 和 f' 互素, 则上述观察表明 $e_i - 1 = 0$ 恒成立, 故 f 无重因式.

设 f 无重因式, 而且 $p_1', \ldots, p_n' \neq 0$. 对每个 i 都有 $e_i = 1$, 而且次数的理由导致 $p_i \nmid p_i'$, 因此上述观察表明 $p_i \nmid f'$. 由此知 f 和 f' 互素. □

在所有素因式 p 都是一次的情形, (ii) 包含的前提 $p' \neq 0$ 自动成立, 这就导致以下结果.

推论 6.6.9 设 $f \in F[X]$ 分裂, 则 f 无重根当且仅当 f 和 f' 互素.

练习 6.6.10 尽管根的概念和所考察的域直接相关, 但 f 和 f' 互素与否却不依赖于域. 确切地说, 设 $f_1, \ldots, f_n \in F[X]$, 而 F 是域 E 的子域. 请证明 $f_1, \ldots, f_n$ 在 $F[X]$ 中互素当且仅当它们在 $E[X]$ 中互素.

提示〉 在 F 和在 E 中作辗转相除法是一回事.

6.7 对称多项式

选定域 F 和 $n \in \mathbb{Z}_{\geq 1}$. 对于 n 元多项式 $f \in F[X_1, \ldots, X_n]$ 和置换 $\sigma \in \mathfrak{S}_n$, 用 σ 重排变元 $X_1, \ldots, X_n$ 的产物记为

$$\sigma f := f(X_{\sigma(1)}, \ldots, X_{\sigma(n)}) \in F[X_1, \ldots, X_n].$$

尽管以下性质暂时不用, 其验证是毫无困难的:

$$\begin{aligned} \mathrm{id}\, f &= f, \\ \sigma(\tau f) &= (\sigma\tau) f, \quad \sigma, \tau \in \mathfrak{S}_n. \end{aligned}$$

在任何重排下都不改变的多项式具有特殊的用处与兴味. 它们对于求根问题尤其重要.

定义 6.7.1 (对称多项式) 设 $f \in F[X_1, \ldots, X_n]$. 若 $\sigma f = f$ 对所有 $\sigma \in \mathfrak{S}_n$ 成立, 则称 f 为对称多项式.

练习 6.7.2 说明若 $f, g \in F[X_1, \ldots, X_n]$ 都是对称多项式, 则 fg 和 $\alpha f + \beta g$ 亦然, 其中 $\alpha, \beta \in F$.

例 6.7.3 对所有 $k \in \mathbb{Z}_{\geq 0}$ 定义 $p_k := \sum_{i=0}^{n} X_i^k$, 这是对称多项式, 称为 k 次**幂和**. 本章习题介绍的 Newton 公式能够将 p_k 以行将定义的初等对称多项式递归地表出. 更详细的讨论可见 [10, 定理 5.8.7]. 留意到 $k = 0$ 对应的幂和是常数多项式, 这是最简单的对称多项式.

所有 n 元对称多项式构成 $F[X_1, \ldots, X_n]$ 的子环, 它同时也是 $F[X_1, \ldots, X_n]$ 作为 F-向量空间的子空间.

定义 6.7.4 对于 $1 \leq k \leq n$, 定义 n 元多项式

$$e_k := \sum_{1 \leq i_1 < \cdots < i_k \leq n} X_{i_1} \cdots X_{i_k};$$

容易看出它是 k 次齐次的对称多项式, 称为第 k 个**初等对称多项式**. 对于 $k = 0$ 的情形, 方便的规定是 $e_0 := 1$.

初等对称多项式和求根问题直接相关: 引进变元 Y, 则有

$$\prod_{i=1}^{n} (Y + X_i) = \sum_{i=0}^{n} e_i Y^{n-i},$$

或者换个写法:

$$\prod_{i=1}^{n} (Y - X_i) = Y^n - e_1 Y^{n-1} + \cdots + (-1)^n e_n. \tag{6.7.1}$$

上式无非是多项式根与系数关系的改写, 又称 **Vieta 公式**. 特别地, $e_n = X_1 \cdots X_n$.

引理 6.7.5 设 f 是 n 元对称多项式, 则代入 $X_n = 0$ 的产物 $f(X_1, \dots, X_{n-1}, 0)$ 为零的充要条件是 $e_n \mid f$.

证明 显然 $e_n(X_1, \dots, X_{n-1}, 0) = 0$, 所以充分性显然. 至于必要性, 设

$$f = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} X_1^{i_1} \cdots X_n^{i_n}.$$

对给定的下标 $i_1, \dots, i_n$, 条件 $f(X_1, \dots, X_{n-1}, 0) = 0$ 相当于说 $c_{i_1, \dots, i_n} \neq 0 \implies i_n \geq 1$; 基于 f 的对称性, 对于每个 $1 \leq k \leq n$, 对 f 代入 $X_k = 0$ 的产物也都是零, 于是同理可得 $c_{i_1, \dots, i_n} \neq 0 \implies i_k \geq 1$. 这便说明可从 f 提出 $e_n = X_1 \cdots X_n$. □

既然对称多项式成环, $g(e_1, \dots, e_n)$ 对于任何 $g \in F[X_1, \dots, X_n]$ 也依然是对称的. 对称多项式基本定理 6.7.7 断言所有对称多项式都来自于这一构造. 作为准备, 我们先将次数的概念方便地拓展到 n 元多项式.

约定 6.7.6 (多项式的齐次部分) 设 $f = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} X_1^{i_1} \cdots X_n^{i_n} \in F[X_1, \dots, X_n]$. 对所有 $d \in \mathbb{Z}_{\geq 0}$, 定义 f 的 d 次齐次部分为

$$f_d := \sum_{i_1 + \cdots + i_n = d} c_{i_1, \dots, i_n} X_1^{i_1} \cdots X_n^{i_n};$$

这使得 $f = \sum_d f_d$. 因此满足 $f = f_d$ 的多项式正是 d 次齐次多项式. 另定义

$$\deg f := \begin{cases} \max\{d \in \mathbb{Z}_{\geq 0} : f_d \neq 0\}, & f \neq 0, \\ -\infty, & f = 0. \end{cases} \tag{6.7.2}$$

定理 6.7.7 设 f 为 n 元对称多项式, 则存在 $g \in F[X_1, \dots, X_n]$ 使得

$$f = g(e_1, \dots, e_n).$$

证明 注意到 f_d 对每个 $d \geq 0$ 都是对称多项式. 问题按此立刻化约到 f 是 d 次齐次对称多项式的情形.

对于所有 n 元多项式 g, 定义其权重为

$$\mathrm{wt}(g) := \begin{cases} \max\left\{\sum_{k=1}^n k i_k : X_1^{i_1} \cdots X_n^{i_n} \text{ 在 } g \text{ 中系数非零}\right\}, & g \neq 0, \\ -\infty, & g = 0. \end{cases}$$

今将证明: 若 f 是 d 次齐次的, 则断言中的 g 不仅存在, 还能取到 $\mathrm{wt}(g) \leq d$. 我们将对 $n + d$ 递归地论证.

当 $d = 0$, 亦即 $f \in F$ 时一切是平凡的. 以下设 $d \geq 1$.

对所有 n 元对称多项式 h, 记 $h^\flat := h(X_1, \dots, X_{n-1}, 0)$; 观察到 $h^\flat$ 是 $n-1$ 元对称多项式, $n = 1$ 时理解为常数. 按此得到

- $f^\flat$, 它依然是 d 次齐次的;

⋆ $e_1^\flat, \dots, e_{n-1}^\flat$, 它们正好是第 $1, \dots, n-1$ 个初等 $n-1$ 元对称多项式.

我们知道存在 $g_1 \in F[X_1, \dots, X_{n-1}]$ 使得 $f^\flat = g_1(e_1^\flat, \dots, e_{n-1}^\flat)$, 而且 $\mathrm{wt}(g_1) \le d$. 由此可见

$$\deg g_1(e_1, \dots, e_{n-1}) \le d$$

成立, 从而

$$f_1 := f - g_1(e_1, \dots, e_{n-1})$$

是次数 $\le d$ 的 n 元对称多项式, 并且 $f_1^\flat = 0$. 引理 6.7.5 蕴涵 $e_n \mid f_1$. 命

$$f_2 := \frac{f_1}{e_n} \in F[X_1, \dots, X_n].$$

从 f_1 和 e_n 的对称性, 易见 f_2 也是对称多项式, 且 $\deg f_2 \le d-n$ (包括 $f_2 = 0$ 的情况). 将 f_2 分解为齐次部分的和, 并应用递归假设, 可知存在 g_2 使得

$$f_2 = g_2(e_1, \dots, e_n), \quad \mathrm{wt}(g_2) \le d-n,$$

从而

$$\begin{aligned} f = f_1 + g_1(e_1, \dots, e_{n-1}) &= e_n f_2 + g_1(e_1, \dots, e_{n-1}) \\ &= e_n g_2(e_1, \dots, e_n) + g_1(e_1, \dots, e_{n-1}). \end{aligned}$$

综上应取 $g := X_n g_2 + g_1$, 它显然满足 $\mathrm{wt}(g) \le d$. □

上述证明具有算法特性, 它可以用来具体地求 g.

定理 6.7.7 断言了表法 $f = g(e_1, \dots, e_n)$ 的存在性, 然而 g 是否由 f 唯一确定? 答案是肯定的: 若 $g(e_1, \dots, e_n) = h(e_1, \dots, e_n)$, 则 $(g-h)(e_1, \dots, e_n) = 0$, 故唯一性归结为以下陈述, 称为 $e_1, \dots, e_n$ 的**代数无关性**.

定理 6.7.8 设 $g \in F[X_1, \dots, X_n]$. 若 $g(e_1, \dots, e_n) = 0$, 则 $g = 0$.

定理 6.7.8 的证明不难, 然而需要一些后续的结果, 留作本章习题. 对称多项式还有许多值得称道的性质, 有意深究的读者不妨参考 [10, §5.8], 该处对定理 6.7.7 和 6.7.8 也提供了不同的证明.

练习 6.7.9 (判别式) 考虑首一多项式

$$f = X^n + c_{n-1}X^{n-1} + \cdots + c_0 = \prod_{i=1}^{n}(X - \alpha_i).$$

(i) 对所有置换 $\sigma \in \mathfrak{S}_n$, 说明

$$\prod_{i<j}(\alpha_{\sigma(i)} - \alpha_{\sigma(j)}) := \mathrm{sgn}(\sigma) \prod_{i<j}(\alpha_i - \alpha_j).$$

(ii) 定义 f 的判别式为

$$\operatorname{disc}(f)=\prod_{i<j}(\alpha_i-\alpha_j)^2,$$

说明若将 $\alpha_1,\dots,\alpha_n$ 视作变元, 则 $\operatorname{disc}(f)$ 是关于 $\alpha_1,\dots,\alpha_n$ 的对称多项式.

(iii) 定理 6.7.7 确保 $\operatorname{disc}(f)$ 可以表作系数 $c_0,\dots,c_{n-1}$ 的多项式, 定理 6.7.8 还确保表法唯一; 依此, $\operatorname{disc}(f)$ 能对任何首一多项式 f 来定义, 不必要求 f 在给定的域 F 上分裂. 对于 $n=2,3$ 情形具体写下 $\operatorname{disc}(f)$, 并将之与 §1.1 介绍的二次和三次方程理论相对比.

举例明之, X^3+pX+q 的判别式为 $-4p^3-27q^2$, 和 (1.1.2) 一致.

6.8 结式

我们已经在 §6.2 谈及关于多项式互素的概念. 在具体例子中, 辗转相除法可以高效地算出两个一元多项式的最大公因式, 但是出于理论的需求, 我们也时常寻觅一个精确的量来判定它们互素与否, 这个量应该由多项式的系数直接确定. 尽管问题初看并非线性的, 线性方程组与行列式的理论依然能提供一个优雅的答案.

本节仍然选定域 F, 并且约定以降幂方式表达多项式.

定义 6.8.1 (结式) 设 $n,m\in\mathbb{Z}_{\geq 1}$. 考如虑 $F[X]$ 的元素

$$\begin{aligned} f&=v_0X^n+\cdots+v_n,\\ g&=w_0X^m+\cdots+w_m, \end{aligned}$$

其中 $v_i,w_j\in F$. 定义 f 和 g 的结式为

$$\operatorname{Res}(f,g):=\begin{vmatrix} v_0 & \cdots & \cdots & v_n & & & \\ & \ddots & & & \ddots & & \\ & & v_0 & \cdots & \cdots & v_n & \\ w_0 & \cdots & \cdots & w_m & & & \\ & w_0 & \cdots & \cdots & w_m & & \\ & & \ddots & & & \ddots & \\ & & & w_0 & \cdots & \cdots & w_m \end{vmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ m\text{ 行} \\ \\ \\ n\text{ 行} \end{matrix}$$

这是 $n+m$ 阶行列式, 空白部分均为零.

观察到若将系数 $v_0,\dots,w_m$ 视作变元, 而不代入具体数值, 则 $\operatorname{Res}(f,g)$ 展开后是关于 $v_0,\dots,w_m$ 的整系数 $n+m+2$ 元多项式, 系数和 F 无关.

练习 6.8.2 运用行列式的性质, 迅速地说明

1. $\mathrm{Res}(g,f)=(-1)^{nm}\mathrm{Res}(f,g)$;

2. 对所有 $t\in F$ 皆有

$$\mathrm{Res}(tf,g)=t^m\mathrm{Res}(f,g),\quad \mathrm{Res}(f,tg)=t^n\mathrm{Res}(f,g);$$

3. 在 $\mathrm{Res}(f,g)$ 的展开式中, $v_0^m w_m^n$ 的系数为 1.

引理 6.8.3 选定 $n,m\in\mathbb{Z}_{\geq 1}$. 设 $f,g\in F[X]$ 具有如定义 6.8.1 的表法, 则 $\mathrm{Res}(f,g)=0$ 当且仅当存在 $f_1,g_1\in F[X]$, 不全为零, 使得

$$fg_1+gf_1=0,\quad \deg f_1<n,\quad \deg g_1<m.$$

证明 将 f_1 和 g_1 表作

$$\begin{aligned} f_1&=a_1X^{n-1}+\cdots+a_n,\\ g_1&=b_1X^{m-1}+\cdots+b_m, \end{aligned}$$

然后将 $fg_1+gf_1=0$ 展开为关于

$$b_1,\ldots,b_m,a_1,\ldots,a_n$$

的 $n+m$ 元线性方程组 (注意顺序). 依序考察 $X^{n+m-1},\ldots,1$ 的系数, 可见线性方程组的系数矩阵是以下矩阵的转置

$$\begin{pmatrix} v_0 & \cdots & v_n & & & \\ & v_0 & \cdots & v_n & & \\ & & \ddots & & \ddots & \\ & & & v_0 & \cdots & v_n \\ w_0 & \cdots & w_m & & & \\ & w_0 & \cdots & w_m & & \\ & & \ddots & & \ddots & \\ & & & w_0 & \cdots & w_m \end{pmatrix}.$$

既然转置不影响行列式, 根据 Cramer 法则 (推论 5.7.6), 方程组有不全为零的解当且仅当 $\mathrm{Res}(f,g)=0$. □

定理 6.8.4 选定 $n,m\in\mathbb{Z}_{\geq 1}$. 设 $f,g\in F[X]$ 具有如定义 6.8.1 的表法, 则 $\mathrm{Res}(f,g)=0$ 当且仅当以下任何一个条件成立: $v_0=0=w_0$, 或者 f 和 g 有次数 >0 的公因式.

证明 先处理 “当” 的方向. 若 $v_0 = 0 = w_0$, 则 $\mathrm{Res}(f,g)$ 是一个首列全为 0 的行列式, 故 $\mathrm{Res}(f,g) = 0$. 若 f 和 g 有次数 > 0 的公因式 d, 则可取多项式 $f_1 := -f/d$ 和 $g_1 := g/d$, 它们不全为零, 而且

$$fg_1 + gf_1 = 0, \quad \deg f_1 < \deg f \leq n, \quad \deg g_1 < \deg g \leq m.$$

代入引理 6.8.3 便得到 $\mathrm{Res}(f,g) = 0$.

现在处理 “仅当” 方向. 设若 f, g 当中有一者为零, 不失一般性可设 $f = 0$, 则:

- ⋆ 或者 $\deg g > 0$, 此时 g 本身便是寻求的公因式;
- ⋆ 或者 g 为常数多项式, 因为 $m \geq 1$, 这将导致 $w_0 = 0 = v_0$;

无论哪种情形, 断言都成立. 因此以下可以合理地假定 f, g 皆非零. 进一步还能设 v_0, w_0 不全为零, 不妨便设 $w_0 \neq 0$. 我们的目标是证明此时 f 和 g 不可能互素.

基于 $\mathrm{Res}(f,g) = 0$, 按引理 6.8.3 的方式取不全为零的 $f_1, g_1 \in F[X]$. 它们满足

$$fg_1 + gf_1 = 0, \quad \deg g_1 < m = \deg g.$$

上式相当于 $F(X)$ 中的等式 $\frac{f}{g} = \frac{-f_1}{g_1}$. 假若 f 和 g 互素, 则关于既约分式表法的定义–命题 6.2.6 将蕴涵 $g \mid g_1$, 这与 $\deg g_1 < \deg g$ 矛盾. 明所欲证. □

作为一则立即的推论, 若 f 和 g 有公共根 α, 亦即 $X - \alpha$ 是它们的公因式, 则 $\mathrm{Res}(f,g) = 0$.

结式的历史渊源之一是高次方程组的求解. 本章习题将有进一步的介绍.

我们接着对 f 和 g 分裂的情形说明结式何以为 “结”.

定理 6.8.5 *选定 $n, m \in \mathbb{Z}_{\geq 1}$. 设 $f, g \in F[X]$ 有以下表法*

$$f = a\prod_{i=1}^{n}(X - \alpha_i),$$
$$g = b\prod_{j=1}^{m}(X - \beta_j),$$

其中 a, b 和 α_i, β_j 都是 F 的元素, 则

$$\begin{aligned}\mathrm{Res}(f,g) &= a^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{nm} b^n \prod_{j=1}^{m} f(\beta_j) \\ &= a^m b^n \prod_{i,j}(\alpha_i - \beta_j).\end{aligned}$$

证明 既然 $\mathrm{Res}(f,g) = (-1)^{nm}\mathrm{Res}(g,f)$, 证 $\mathrm{Res}(f,g) = a^m \prod_{i=1}^{n} g(\alpha_i)$ 即可. 进一步, 问题还容易化约到 $a = 1$ 的情形.

我们先处理 $g(\alpha_1),\dots,g(\alpha_n)$ 两两相异的情形. 引入变元 Y 并在 $F(Y)$ 上考虑 $\operatorname{Res}(f,g-Y)$. 这相当于将 g 的常数项 b_m 改成 b_m-Y, 给出一个关于 Y 的多项式. 就结式的定义思忖半响, 可以得出降幂表法

$$\operatorname{Res}(f,g-Y)=(-1)^nY^n+\cdots+\operatorname{Res}(f,g).$$

对所有 $1\leqslant i\leqslant n$, 既然 α_i 是 f 和 $g-g(\alpha_i)$ 的公根, 故 $\operatorname{Res}(f,g-g(\alpha_i))=0$, 从而 $g(\alpha_i)-Y\mid\operatorname{Res}(f,g-Y)$. 因为 $g(\alpha_1),\dots,g(\alpha_n)$ 两两相异, 我们得出

$$\prod_{i=1}^{n}(g(\alpha_i)-Y)\mid\operatorname{Res}(f,g-Y).$$

两边都是 Y 的 n 次多项式, 比较 Y^n 的系数可见上式实则是等式. 代入 $Y=0$ 便是 $\operatorname{Res}(f,g)=\prod_{i=1}^{n}g(\alpha_i)$.

现在考虑一般情形. 以下采取的技巧是以 n 个变元来代替 $\alpha_1,\dots,\alpha_n$, 另外记为 $Z_1,\dots,Z_n$, 并相应地用 $\tilde f=\prod_{i=1}^{n}(X-Z_i)$ 代替 f. 我们先着手在环 $F[Z_1,\dots,Z_n]$ 中证明

$$\operatorname{Res}(\tilde f,g)=\prod_{i=1}^{n}g(Z_i).\tag{6.8.1}$$

当然, 我们可设 $b\neq0$, 否则等式是平凡的. 于是 $g(Z_1),\dots,g(Z_n)$ 显然两两相异, 所求等式因而是先前处理的情况在有理函数域 $F(Z_1,\dots,Z_n)$ 上的应用.

现在对等式 (6.8.1) 代入 $Z_i=\alpha_i$. 既然结式是关于多项式系数的整系数多项式, 而且这些整系数与域的选取无关, 故左式化为 $\operatorname{Res}(f,g)$, 右式则化为 $\prod_{i=1}^{n}g(\alpha_i)$. 证毕. □

证明中关于一般情形的论证也能以用定理 3.6.3 的延拓原理处理, 为此需先将 F 扩张为无穷域, 留给有兴趣的读者练习.

作为一则推论, 现在可以明确结式与练习 6.7.9 介绍的判别式有何联系, 这涉及 §6.4 的形式求导. 首先将判别式的定义推广到非首一的情形: 对

$$f=a\prod_{i=1}^{n}(X-\alpha_i),\quad a,\alpha_1,\dots,\alpha_n\in F,\quad a\neq0$$

定义其判别式为

$$\operatorname{disc}(f):=a^{2n-2}\prod_{i<j}(\alpha_i-\alpha_j)^2.\tag{6.8.2}$$

推论 6.8.6 对于如上的 f, 我们有 $a\cdot\operatorname{disc}(f)=(-1)^{\frac{n(n-1)}{2}}\operatorname{Res}(f,f')$; 右式对应于在结式定义中取 $m=n-1$ 和 $g=f'$.

证明 定理 6.8.5 给出

$$\operatorname{Res}(f,f')=a^{n-1}\prod_{i=1}^{n}f'(\alpha_i).$$

因为 $f'=a\sum_{k=1}^{n}\prod_{j\neq k}(X-\alpha_j)$, 对所有 $1\leq i\leq n$ 都有

$$f'(\alpha_i)=a\prod_{j\neq i}(\alpha_i-\alpha_j),$$

因之

$$\mathrm{Res}(f,f')=a^{2n-1}\prod_{i=1}^{n}\prod_{j\neq i}(\alpha_i-\alpha_j)=a^{2n-1}(-1)^{\frac{n(n-1)}{2}}\prod_{i<j}(\alpha_i-\alpha_j)^2.$$

右式整理为 $(-1)^{\frac{n(n-1)}{2}}a$ 乘上先前所定义的 $\mathrm{disc}(f)$. □

若 $f=a_0X^n+\cdots+a_n$, 则 $\mathrm{Res}(f,f')$ 涉及的行列式表作 $a_0,\ldots,a_n$ 的整系数多项式, 而且从行列式首列能提出 $a=a_0$, 因此 $(-1)^{n(n-1)/2}a^{-1}\mathrm{Res}(f,f')$ 表作 $a_0,\ldots,a_n$ 的整系数多项式. 这说明 $\mathrm{disc}(f)$ 在 f 非首一时的定义确实合理.

6.9 不可约多项式初探

我们在 §6.2 初步介绍过不可约多项式的概念. 一般而言, 判断一个多项式可约与否是颇为棘手的问题. 本节仅探讨一些简单然而基础的面向.

当然, 问题和域 F 的选取密切相关.

例 6.9.1 域 F 是定义 6.6.5 所谓的代数闭域当且仅当 F 上的不可约多项式都是一次多项式, 细说如下. 设 $f\in F[X]\smallsetminus F$:

- ⋆ 若 F 代数闭, 则因为 f 分解为一次因式的积, 使其不可约的唯一可能是 f 本身便是一次的;
- ⋆ 反之, 若 F 上的不可约多项式都是一次的, 则将 f 分解为不可约因式便是分解为一次因式的积.

因此复数域 $\mathbb{C}$ 上的不可约多项式都是一次多项式.

例 6.9.2 实数域 $\mathbb{R}$ 上的不可约多项式分成两类. 不失一般性, 考虑首一不可约多项式即可:

- ▷ **一次情形** $X-a$;
- ▷ **二次情形** X^2+bX+c, 要求判别式 $b^2-4c<0$.

为了说明这点, 设 $f\in\mathbb{R}[X]$ 首一. 先将 f 在 $\mathbb{C}[X]$ 中分解为 $f=\prod_{i=1}^{n}(X-\lambda_i)$, 其中满足 $\lambda_i\in\mathbb{R}$ 的因式 $X-\lambda_i$ 已经属于 $\mathbb{R}[X]$, 其余则按复共轭合并为

$$(X-\lambda_i)(X-\overline{\lambda_i})=X^2-(\lambda_i+\overline{\lambda_i})X+\lambda_i\overline{\lambda_i}\ \in\mathbb{R}[X];$$

综上, $\mathbb{R}[X]$ 的不可约元只能是一次或二次的.

尚须确定哪些实二次多项式不可约. 在任意域 F 上, 多项式 X^2+bX+c 可约等价于它在 F 中有根; 当 $F=\mathbb{R}$, 有根等价于 $b^2-4c\geq 0$. 这就完成了所求的分类.

下一个考量的情形是有理数域 $\mathbb{Q}$. 观察到对任意 $f\in\mathbb{Q}[X]$, 将系数通分给出 $f=\frac{g}{t}$, 其中 $t\in\mathbb{Z}$ 非零而 $g\in\mathbb{Z}[X]$. 于是 f 不可约当且仅当 g 不可约. 进一步从 g 的系数提出最大公因数, 便将多项式的可约性化约到各项系数互素的整系数多项式来考量. 这一类多项式具有特别的兴味.

定义 6.9.3 设 $f=\sum_{k=0}^{n}a_kX^k\in\mathbb{Z}[X]$ 非零. 命

$$c(f):=\ a_0,\dots,a_n \text{ 的最大公因数}.$$

若 $c(f)=1$, 则称 f 为**本原多项式**.

任何 $f\in\mathbb{Z}[X]\smallsetminus\{0\}$ 都能写成 $f=c(f)f_0$ 的形式, 其中的 f_0 是本原多项式.

以下的重要结果被称为 **Gauss 引理**.

引理 6.9.4 (C. F. Gauss) 设 $g,h\in\mathbb{Z}[X]$ 为本原多项式, 则 gh 亦然.

证明 将 g 按升幂展开为 $a_0+\cdots+a_nX^n$, 类似地, $h=b_0+\cdots+b_mX^m$, 其中 $a_n,b_m\neq 0$. 方便起见, 不妨对整数 $i\notin[0,n]$ (或 $j\notin[0,m]$) 命 $a_i=0$ (或 $b_j=0$).

对于任意素数 p, 由于 p 不整除所有 $a_0,\dots,a_n$, 可取 $0\leq r\leq n$ 使得

$$i>r\implies p\mid a_i,\quad p\nmid a_r.$$

同理可取 $0\leq s\leq m$ 使得

$$j>s\implies p\mid b_j,\quad p\nmid b_s.$$

由于 gh 的 k 次项系数是 $\sum_{i+j=k}a_ib_j$, 取 $k=r+s$ 则有

$$p\nmid a_rb_s+\underbrace{\sum_{i>r}a_ib_{r+s-i}+\sum_{j>s}a_{r+s-j}b_j}_{\text{被 } p \text{ 整除}},$$

因此 gh 有不被 p 整除的系数. 既然 p 可任选, 这表明 gh 是本原多项式. □

引理 6.9.5 对于所有非零的 $g,h\in\mathbb{Z}[X]$, 我们有 $c(gh)=c(g)c(h)$.

证明 按定义可将 g (或 h) 写作 $g=c(g)g_0$ (或 $h=c(h)h_0$), 其中 g_0,h_0 是本原多项式. 于是 $gh=c(g)c(h)g_0h_0$, 而引理 6.9.4 蕴涵 g_0h_0 是本原多项式. 这就表明 $c(gh)=c(g)c(h)$. □

定理 6.9.6 设 $f\in\mathbb{Z}[X]$ 为本原多项式, $\deg f>0$. 以下陈述等价:

(a) f 作为 $\mathbb{Q}[X]$ 的元素是不可约的;

(b) 不存在次数皆 >0 而且满足 $f=gh$ 的多项式 $g,h\in\mathbb{Z}[X]$.

此外, 若本原多项式 f 能在 $\mathbb{Q}[X]$ 中分解为 gh, 则将 g 和 h 分别乘以某个 $\alpha\in\mathbb{Q}^\times$ 以及 α^{-1} 后, 可以确保 g 和 h 都是本原多项式.

证明 以下论证 (a) 和 (b) 两者的否定相互等价. 选定本原多项式 $f\in\mathbb{Z}[X]$, $\deg f>0$.

若有次数 >0 的 $g,h\in\mathbb{Z}[X]$ 使得 $f=gh$, 将此式放在 $\mathbb{Q}[X]$ 中考量, 便表明 f 作为 $\mathbb{Q}[X]$ 的元素可约.

反之设 f 在 $\mathbb{Q}[X]$ 中可约, 分解为 $f=gh$, 其中 $g,h\in\mathbb{Q}[X]$ 的次数皆 >0. 取 $u,v\in\mathbb{Z}_{\geq 1}$ 使得 $ug,vh\in\mathbb{Z}[X]$, 则 $uvf=(ug)(vh)$ 而

$$uv=c(uvf)=c(ugvh)\xlongequal{\text{引理 6.9.5}}c(ug)c(vh),$$

于是

$$f=\frac{ug}{c(ug)}\cdot\frac{vh}{c(vh)}.$$

命 $g_1:=\frac{ug}{c(ug)}$ 和 $h_1:=\frac{vh}{c(vh)}$, 它们都是本原多项式, 而且 $f=g_1h_1$. 这就证明了 (a) 等价于 (b), 也顺带证明了最后一则断言在 $\deg f>0$ 时的情形, 然而该断言在 $\deg f=0$ 时是平凡的. □

这套技术足以分类整环 $\mathbb{Z}[X]$ 的所有不可约元 (定义 6.2.3), 从而证明 $\mathbb{Z}[X]$ 是唯一分解环 (定义 6.2.5).

定理 6.9.7 整环 $\mathbb{Z}[X]$ 的不可约元分成两类:

▷ **第一类** $\mathbb{Z}$ 的不可约元.

▷ **第二类** 次数 >0 并满足定理 6.9.6 的等价条件 (a) 或 (b) 的本原多项式 f.

进一步, $\mathbb{Z}[X]$ 是唯一分解环.

证明 先分类不可约元. 首先容易看出 $\mathbb{Z}[X]$ 中满足 $\deg f=0$ 的不可约元必然是 $\mathbb{Z}$ 的不可约元, 这是第一类.

其次设 $f\in\mathbb{Z}[X]$ 满足 $\deg f>0$. 若 f 在 $\mathbb{Z}[X]$ 中不可约, 则 f 必为本原多项式, 否则能从系数中提取 $\mathbb{Z}$ 的不可约元, 故以下可设 f 为本原多项式. 若有 $\mathbb{Z}[X]$ 中的分解 $f=gh$, 而且 $g,h\notin\mathbb{Z}^\times=\mathbb{Z}[X]^\times$, 则必有 $\deg f,\deg g>0$; 因此 f 在 $\mathbb{Z}[X]$ 中不可约当且仅当它满足定理 6.9.6 的条件 (b), 这是第二类.

接着证明 $\mathbb{Z}[X]$ 是唯一分解环. 给定 $f\in\mathbb{Z}[X]\smallsetminus\{0\}$, 先将 $c(f)$ 在 $\mathbb{Z}$ 中作不可约分解, 然后将本原多项式 $f_0:=\frac{f}{c(f)}$ 在 $\mathbb{Q}[X]$ 中作不可约分解; 注意到根据定理 6.9.6 的第二部分, 从 f_0 析出的所有不可约因式皆可取为本原多项式, 因而满足定理 6.9.6 的等价条件 (a), 给出 $\mathbb{Z}[X]$ 的第二类不可约元. 综上, f 表作 $\mathbb{Z}[X]$ 的不可约元的乘积.

最后说明 $\mathbb{Z}[X]$ 的不可约分解唯一: 对整环 $\mathbb{Z}[X]$ 和 $\mathbb{Z}$ 沿用约定 6.2.1 的等价关系 $\sim$. 回忆到 $\mathbb{Z}^\times = \mathbb{Z}[X]^\times$, 故两个环上的 $\sim$ 不致混淆. 设有

$$a_1 \cdots a_m p_1 \cdots p_n \sim b_1 \cdots b_r q_1 \cdots q_s,$$

其中 a_i, b_j (或 p_i, q_j) 是 $\mathbb{Z}[X]$ 的第一类 (或第二类) 不可约元, 容许重复. 两边同取 $c(\cdot)$ 得到 $a_1 \cdots a_m \sim b_1 \cdots b_r$, 根据 $\mathbb{Z}$ 的唯一分解性遂有 $m = r$ 而 $a_1, \ldots, a_m$ 和 $b_1, \ldots, b_r$ 仅差一个重排和 $\sim$.

于是 $p_1 \cdots p_n \sim q_1 \cdots q_s$. 根据 $\mathbb{Q}[X]$ 的唯一分解性, $n = s$ 而两边仅差一个重排和来自 $\mathbb{Q}^\times = \mathbb{Q}[X]^\times$ 的倍数. 讨论 $\mathbb{Q}^\times$ 中元素的既约分式, 易见两个本原多项式相差一个 $\mathbb{Q}^\times$ 的倍数当且仅当它们相差 $\mathbb{Z}^\times$ 的倍数. 这就证明了分解的唯一性. □

定理 6.9.6 的好处在于 f 的不可约性是对 $\mathbb{Q}[X]$ 而论, 但它能否分解为次数 > 0 的本原多项式乘积则是一个整系数的问题. 关于 $\mathbb{Q}$ 上不可约多项式的研究因此带有数论意涵. 以下是称为 Eisenstein 判准的一则著名应用.

定理 6.9.8 (G. Eisenstein) *设 $n \in \mathbb{Z}_{\geq 1}$ 而 $f \in a_0 + a_1 X + \cdots + a_n X^n \in \mathbb{Z}[X]$. 若有素数 p 满足*

$$\begin{gathered} p \nmid a_n, \\ 0 \leq i < n \implies p \mid a_i, \\ p^2 \nmid a_0, \end{gathered}$$

则 f 作为 $\mathbb{Q}[X]$ 的元素不可约.

证明 首先化约到 f 是本原多项式的情形. 条件 $p \nmid a_n$ 说明 $p \nmid c(f)$, 因此 $\frac{f}{c(f)} \in \mathbb{Z}[X]$ 的系数对 p 满足相同的条件. 以 $\frac{f}{c(f)}$ 代 f 可确保 f 是本原的.

以下采取反证法. 假若 f 可约, 则定理 6.9.6 确保有次数 > 0 的本原多项式

$$g = b_0 + \cdots + b_m X^m, \quad h = c_0 + \cdots + c_l X^l, \quad b_m, c_l \neq 0$$

使得 $f = gh$. 观察到 $l + m = n$, 而 $a_n = b_m c_l$, $a_0 = b_0 c_0$. 于是有

$$p \nmid b_m, \quad p \nmid c_l, \quad p \mid a_0 = b_0 c_0.$$

不失一般性可以假定 $p \mid b_0$. 现在取 $0 < k \leq m$ 使得

$$0 \leq i < k \implies p \mid b_i, \quad p \nmid b_k.$$

我们有同余式

$$a_k = \sum_{i=0}^{k} b_i c_{k-i} \equiv b_k c_0 \pmod{p};$$

另一方面 $k \leq m < n$ 蕴涵 $a_k \equiv 0 \pmod{p}$, 故 $b_k c_0 \equiv 0 \pmod{p}$. 既然 $p \nmid b_k$, 必有 $p \mid c_0$, 然而这与 $p^2 \nmid a_0 = b_0 c_0$ 矛盾. □

例 6.9.9 取 p 为素数. Eisenstein 判准的标准应用是证多项式

$$\Phi_p := \frac{X^p-1}{X-1} = 1 + X + \cdots + X^{p-1}$$

不可约. 证明的第一步是以二项式展开将 Φ_p 改写为

$$\frac{((X-1)+1)^p-1}{X-1} = \frac{\sum_{k=1}^{p}\binom{p}{k}(X-1)^k}{X-1} = \sum_{k=1}^{p}\binom{p}{k}(X-1)^{k-1}.$$

第二步是以下观察: 在任意域 F 上, 若 $h = \sum_n a_n X^n \in F[X]$ 不可约, 则 $h(X-1) = \sum_n a_n (X-1)^n$ 也不可约. 何以故? 映射 $h \mapsto h(X-1)$ 是环 $F[X]$ 的自同构, 以 $h \mapsto h(X+1)$ 为逆, 所以不可约性也相应地在 h 和 $h(X-1)$ 之间转译.

现将第二步的观察用于 $F = \mathbb{Q}$ 和 $h := \sum_{k=1}^{p}\binom{p}{k}X^{k-1}$. 第一步说明 $h(X-1) = \Phi_p$, 问题遂归结为证 h 不可约. 观察到 h 是首一整系数多项式, 常数项是 p; 基于 Eisenstein 判准, 问题进一步化作证明

$$1 \le k \le p-1 \implies p \mid \binom{p}{k}.$$

然而 $\binom{p}{k} = \frac{p(p-1)\cdots(p-k+1)}{k!}$ 已知是整数, 当 $1 \le k \le p-1$ 时分子含 p 而分母和 p 互素, 这就表明 $p \mid \binom{p}{k}$. 明所欲证.

例 6.9.9 现身的 Φ_p 是分圆多项式的一员. 本章习题将介绍分圆多项式的一般定义.

练习 6.9.10 设 $f = \sum_n a_n X^n \in \mathbb{Z}[X]$ 是本原多项式. 设 p 为素数, 记 $[a_n]_p$ 为 a_n 在 $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ 中的像, 亦即同余类 $a_n + p\mathbb{Z}$. 证明若 f 的最高次系数和 p 互素, 而且 $\mathbb{F}_p[X]$ 的元素 $f_p := \sum_n [a_n]_p X^n$ 不可约, 则 f 不可约.

提示 若 f 可约, 则必可写成次数 >0 的本原多项式 g 和 h 的积, 而且 g 和 h 的最高次系数都和 p 互素. 从 $f_p = g_p h_p$ 说明此时 f_p 可约.

对于一般的整系数多项式, 有以下的简单算法来确定它的一次因式, 即根. 由于不超过三次的多项式可约当且仅当有根, 它们的可约性便容易按此判定.

命题 6.9.11 (一次因式检验法) 设有理数 $\alpha = \frac{u}{v}$ 是整系数多项式 $a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{Z}[X]$ 的根, 其中 u 和 v 互素, 则有 $v \mid a_n$ 和 $u \mid a_0$.

证明 对 $a_n\alpha^n + a_{n-1}\alpha^{n-1}\cdots + a_0 = 0$ 两边同乘以 v^n, 得到

$$a_n u^n + a_{n-1}u^{n-1}v + \cdots + a_0 v^n = 0;$$

因此左式被 u 和 v 整除. 利用整数的唯一分解性质, 关于互素的条件便蕴涵 $v \mid a_n$ 和 $u \mid a_0$. □

推论 6.9.12 首一整系数多项式的有理根必然是整数.

证明 在一次因式检验法中代入 $a_n = 1$. □

读者可能已经察觉, 在本节关于整系数多项式的论证中, 起关键作用的是整环 $\mathbb{Z}$ 的唯一分解性. 关于 $\mathbb{Z}$ 上的多项式的这些结果容易推广到任意唯一分解环 R 及其分式域 K 上. 譬如设 $f = \sum_n a_n X^n \in R[X] \smallsetminus \{0\}$, 则当 $c(f) := \gcd(a_0, \dots, a_n) \sim 1$ 时称 f 为本原多项式, 而定理 6.9.7 则改述如下:

$$
\begin{aligned}
f \text{ 在 } R[X] \text{ 中不可约} \iff & \\
& \text{或者 } f \text{ 是 } R \text{ 的不可约元,} \\
& \text{或者 } \deg f > 0,\ f \text{ 本原, 而且在 } K[X] \text{ 中不可约;}
\end{aligned}
$$

此外 $R[X]$ 仍是唯一分解环. 论证是完全相同的. 感兴趣的读者可以进一步参考 [10, §5.7] 或本章习题.

算法 6.9.13 为了判断高次整系数多项式的可约性, 最明智的选择是让计算机代劳. 为了展示这确实是可计算的, 下面介绍称为 Kronecker 法的一种初等手段.

给定本原多项式 $f \in \mathbb{Z}[X]$ 和 $k \in \mathbb{Z}_{\geq 1}$, 为了判定是否存在 $g, h \in \mathbb{Z}[X]$ 使得 $f = gh$ 而且 $\deg g \leq k$, 选取相异的 $x_0, \dots, x_k \in \mathbb{Z}$, 避开 f 的零点; 因为整系数多项式在整点取整值, 故

$$
f = gh \implies \forall j = 0, \dots, k, \quad g(x_j) \mid f(x_j).
$$

由于每个 $f(x_j)$ 仅有有限多个因数, $(g(x_j))_{j=0}^k$ 也仅有有限多种可能. 然而所有数组 $(y_j)_{j=0}^k$ 都以插值法 (例 6.6.7) 确定唯一的 g 使得 $\deg g \leq k$ 而 $g(x_j) = y_j$ 对 $j = 0, \dots, k$ 成立. 综上, 因式被限制在一个有限而且可算的集合中; 对其中每个 g, 带余除法能判定它是否整除 f.

代入 $k = \lfloor \frac{\deg f}{2} \rfloor$. 若上述算法的结果是不存在这般的整系数因式 g 使得 $\deg g > 0$, 则 f 不可约; 如果 f 可约, 则算法可迭代给出 f 的不可约分解.

Kronecker 法只在 f 的次数和系数较小的时候容易实行, 现代数学软件另有更曲折的, 涉及 $\bmod p$ 运算的算法.

6.10 从不可约多项式构造扩域

域 F 上的多项式未必有根, 自古以来这就是困扰学者们的一则麻烦. 本节的目的是说明如何抽象地向多项式添根, 方式是将 F 嵌入足够大的域 L. 从环同态的观点看, 这相当于取适当的单同态 $\xi : F \to L$. 自 L 观之, F 嵌入为子域; 自 F 观之, L 则称为其**扩域**或**扩张**. 扩域的概念当然是相对于 ξ 而言, 但在不致混淆的前提下, 我们惯常省略 ξ, 而将 F 直接视同 L 的子域, 以节约笔墨.

构造扩域的关键一步是取 $F[X]$ 的商. 引理 6.2.8 说明考虑形如 (f) 的理想即可. 留意到 $F[X]/(f)$ 是零环当且仅当 $(f) = F[X]$, 当且仅当 $f \in F^\times$. 另一方面, 取 $f = 0$ 则有 $F[X]/(0) = F[X]$. 因此以下不妨设 $f \in F[X] \smallsetminus F$.

引理 6.10.1 设 $f \in F[X] \smallsetminus F$, 则

$$\begin{aligned} \iota : F &\longrightarrow F[X]/(f) \\ a &\longmapsto a + (f) \end{aligned}$$

是环之间的单同态.

证明 显见 ι 是标准嵌入 $F \hookrightarrow F[X]$ 和商映射 $F[X] \twoheadrightarrow F[X]/(f)$ 的合成, 既然两者都是同态, ι 亦然. 其次, 从域到非零环的环同态总是单的 (练习 6.1.13), 但这点也可以直接验证: 如果 $f \mid a$ 而 $a \in F$, 则因为 $f \notin F$, 唯一可能是 $a = 0$. □

对于任意环 L, 一旦选定了同态 $\xi : F \to L$, 便可以赋 L 以 F-向量空间的结构, 方法是取加法为环 L 的加法, 而 $t \in F$ 的纯量乘法映 $x \in L$ 为 $\xi(t)x \in L$.

将上述思路代入引理 6.10.1 的情境, 便使得 $F[X]/(f)$ 通过同态 ι 成为 F-向量空间, 其加法是商环的加法, 而 $t \in F$ 的纯量乘法映 $h + (f)$ 为 $\iota(t)(h + (f)) = th + (f)$. 这也正是 $F[X]$ 对子空间 (f) 取商得到的向量空间结构.

接着确定 $F[X]/(f)$ 的维数.

引理 6.10.2 设 $f \in F[X] \smallsetminus F$, 则 $1, X, \dots, X^{\deg f - 1}$ 对 (f) 的陪集给出 $F[X]/(f)$ 的基.

证明 带余除法说明任何 $h \in F[X]$ 都属于形如 $r + (f)$ 的陪集, 其中 $\deg r < \deg f$, 而这般的 r 是由 h 唯一确定的. □

设 $f \in F[X] \smallsetminus F$. 对商环 $F[X]/(f)$ 应用命题 6.3.6 可得

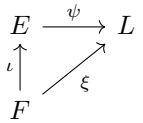
$$F[X]/(f) \text{ 是域} \iff F[X]/(f) \text{ 是整环} \iff f \text{ 不可约}.$$

万事俱备, 现在可以严格地表述向不可约多项式添根的构造, 并且进一步说明此构造方式是最经济的. 先介绍一则符号: 给定环同态 $\xi : F \to L$ 和 $f = \sum_{n \geq 0} a_n X^n \in F[X]$ (有限和), 记 $f^\xi := \sum_{n \geq 0} \xi(a_n) X^n \in L[X]$. 容易看出 $f \mapsto f^\xi$ 给出同态 $F[X] \to L[X]$.

命题 6.10.3 设 $f = \sum_{n \geq 0} a_n X^n \in F[X]$ 不可约.

(i) 相对于引理 6.10.1 的域嵌入 $\iota : F \hookrightarrow F[X]/(f) =: E$, 记 $\alpha := X + (f) \in E$, 则多项式 $f^\iota \in E[X]$ 满足 $f^\iota(\alpha) = 0$.

(ii) 若 L 是交换环, $\xi : F \to L$ 是环同态, 而且 $\beta \in L$ 满足 $f^\xi(\beta) = 0$, 则存在唯一的环同态 $\psi : E \to L$ 使得 $\psi(\alpha) = \beta$, 并且使下图交换:



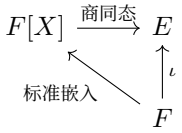
这也等价于说相对于 E (或 L) 上由 ι (或 ξ) 赋予的 F-向量空间结构, ψ 是线性映射.

证明 先证 (i). 直接求 $f^\iota(\alpha)$ 的值为

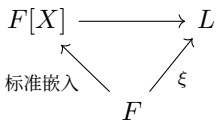
$$\begin{aligned}\sum_{n\geq 0}\iota(a_n)(X+(f))^n &= \sum_{n\geq 0}(a_n+(f))(X^n+(f))\\ &= \sum_{n\geq 0}a_nX^n+(f)\\ &= f+(f)=(f).\end{aligned}$$

此即 $E:=F[X]/(f)$ 的零元.

对于 (ii), 首先假设断言中的同态 $\psi: E\to L$ 存在, 则将交换图表



和断言中的交换图表拼接, 可得交换图表



记上图的水平箭头为 Ψ, 则 $\Psi(X)=\psi(\alpha)=\beta$. 因此对于一般的多项式 $g=\sum_n b_nX^n$, 交换图表给出

$$\Psi(g)=\sum_n\Psi(b_n)\Psi(X)^n=\sum_n\xi(b_n)\beta^n=g^\xi(\beta);$$

换言之, $\Psi: F[X]\to L$ 是求值同态 $g\mapsto g^\xi(\beta)$, 而 ψ 由 $\psi(g+(f))=\Psi(g)$ 完全确定. 唯一性得证.

这也为 ψ 的存在性指明了道路. 何以故? 从 $F[X]$ 到 L 的求值同态 $\Psi: g\mapsto g^\xi(\beta)$ 总是能定义, 由 $\Psi(f)=f^\xi(\beta)=0$ 可见 Ψ 映理想 (f) 为零, 因此命题 6.1.10 将 Ψ 唯一地分解为 $F[X]\xrightarrow{\text{商同态}}E\xrightarrow{\psi}L$, 其中 ψ 映 $g+(f)$ 为 $g^\xi(\beta)$. 特别地, 对所有 $a\in F$ 都有 $\psi(\iota(a))=\xi(a)$, 这便给出所需的交换图表. □

我们更愿意将 F 通过 ι 视同 E 的子域, 从而将 $F[X]$ 视同 $E[X]$ 的子环, 故今后将不加说明地省略 f^ι 的上标 ι, 而直接将 f 看作 $E[X]$ 的元素. 实践表明这在大部分情境下不会导致混淆.

例 6.10.4 取 $F=\mathbb{R}$ 以及其上的不可约多项式 X^2+1. 添根的构造给出嵌入

$$\mathbb{R}\hookrightarrow\mathbb{R}[X]/(X^2+1).$$

在域 $\mathbb{R}[X]/(X^2+1)$ 上, -1 自动有平方根 $\alpha := X + (X^2+1)$; 事实上 $\mathbb{R}[X]/(X^2+1)$ 按两种方式自然地同构于 $\mathbb{C}$. 具体方法是在命题 6.10.3 (ii) 中考虑

$$L := \mathbb{C}, \quad \xi : \mathbb{R} \xrightarrow{\text{标准嵌入}} \mathbb{C}, \quad \beta^{\pm} := \pm \mathrm{i}.$$

既然 $\pm\mathrm{i}$ 都是 $f := X^2+1$ 的根, 我们得到同态 $\psi^{\pm}$ 连同交换图表

$$\begin{array}{ccc} \mathbb{R}[X]/(X^2+1) & \xrightarrow{\psi^{\pm}} & \mathbb{C} \\ \uparrow & \nearrow & \\ \mathbb{R} & & \end{array}$$

使得 $\psi^{\pm}(\alpha) = \pm\mathrm{i}$. 既然 $\mathbb{R}[X]/(X^2+1)$ 是域, $\psi^{\pm}$ 必然单 (练习 6.1.13). 此外它还是满的: 这点既可以从维数看出, 也可以从 $\mathbb{C}$ 的所有元素都能写成 $\pm\mathrm{i}$ 的多项式这点直接推导 — 事实上, 所有复数都能写成 $a \pm b\mathrm{i}$ 的形式.

上述同构 $\psi^{\pm} : \mathbb{R}[X]/(X^2+1) \xrightarrow{\sim} \mathbb{C}$ 并非唯一而是 "唯二" 的, 取决于如何指定 -1 的平方根. 从代数的立场, 这两个平方根无法区别, 也毋须区别.

推论 6.10.5 *对任意域 F 和任意非常数多项式 $f \in F[X]$, 存在域的嵌入 $F \hookrightarrow E_f$ 使得 f 在 E_f 上分裂 (定义 6.6.2).*

证明 对次数 $\deg f$ 递归地论证, 但 F 是可变的. 若 $\deg f = 1$ 则 f 业已分裂. 以下设 $\deg f > 1$.

任取 f 的不可约因式 p. 以命题 6.10.3 构造域嵌入 $F \hookrightarrow E_1$ 使得 p 在 E_1 上有根, 记此根为 α, 则 f 可以置于 $E_1[X]$ 中写成 $f = (X-\alpha)g$ 的形式, $g \in E_1[X]$ 而 $\deg g = \deg f - 1$. 对 $g \in E_1[X]$ 继续递归地操作, 以得到一系列的域嵌入

$$F \hookrightarrow E_1 \hookrightarrow E_2 \hookrightarrow \cdots \hookrightarrow E_r,$$

使得 f 在 E_r 上分裂; 留意到 $r \leq \deg f$. □

在实际应用中, 我们不仅寻求扩域 E 使给定的多项式分裂, 还关心 E 相对于 F 的大小. 为了阐明这个概念, 且先回到本节开头的情境, 并作如下定义.

定义 6.10.6 *考虑域 F, 环 L 连同给定的同态 $F \to L$, 这些资料使得 L 成为 F-向量空间. 定义 L 相对于 F 的*次数*为*

$$[L : F] := \dim_F L.$$

特别地, 对于 F 的任意扩域 E 都能讨论其次数 $[E : F]$.

以引理 6.10.1 的嵌入 $F \hookrightarrow F[X]/(f) =: L$ 为例, 其中 $f \in F[X] \smallsetminus F$. 引理 6.10.2 表明 $[L : F] = \deg f$.

回忆到推论 6.10.5 给出的 $F \hookrightarrow E_f$ 实际源于扩域塔 $F \hookrightarrow E_1 \hookrightarrow \cdots$. 以下结果说明扩域的次数可以在塔中逐层计算.

命题 6.10.7 设 E 是 F 的扩域, L 是任意环, 带有给定的同态 $E \to L$, 从而使 L 成为 E-向量空间; 如将纯量乘法限制在子域 F 上, 则 L 也是 F-向量空间. 它们的次数满足塔性质:

$$[L : F] = [L : E][E : F]$$

证明 不失一般性, 可设 L 不是零环. 既然 $E \to L$ 总是单射 (练习 6.1.13), 今后不妨将 E 和 F 视同 L 的子环, 以简化符号.

取 L 作为 E-向量空间的基 $B \subset L$, 再取 E 作为 F-向量空间的基 $C \subset E$. 兹断言 $\{cb : (b, c) \in B \times C\}$ 是 L 作为 F-向量空间的基.

首先, 任意 $x \in L$ 都能表成有限和 $x = \sum_{b\in B} t_b b$, 而每个系数 t_b 又能在 E 中表成有限和 $\sum_{c\in C} t_{b,c} c$, 其中 $t_{b,c} \in F$. 因此 $x = \sum_{(b,c)\in B\times C} t_{b,c} cb$.

兹断言若有系数在 F 上的线性关系式 $\sum_{(b,c)\in B\times C} t_{b,c} cb = 0$ (有限和), 则对所有 (b, c) 皆有 $t_{b,c} = 0$. 办法仍然是集项: $\sum_b \left(\sum_c t_{b,c} c\right) b = 0$ 蕴含每个 $\sum_c t_{b,c} c \in E$ 皆为 0, 从而每个 $t_{b,c} \in F$ 皆为 0.

特别地, 这还蕴涵当 (b, c) 变动, cb 在 L 中必须两两相异, 因此以上构造的基恰有 $|B| \cdot |C| = [L : E][E : F]$ 个元素. □

推论 6.10.8 设 $f \in F[X]$ 满足 $n := \deg f \geq 1$, 则推论 6.10.5 中的扩域 E_f 可以适当地选取, 使得 $[E_f : F] \leq n!$.

证明 回到推论 6.10.5 中的递归构造 $F \hookrightarrow E_1 \hookrightarrow \cdots \hookrightarrow E_r =: E_f$, 第一层满足 $E_1 \simeq F[X]/(p)$, 其中 p 是 f 的某个不可约因子, 因此

$$[E_1 : F] = \deg p \leq \deg f = n.$$

其后各层是对 $n - 1$ 次多项式 $g := f/(X - \alpha) \in E_1[X]$ 递归操作的产物 ($\alpha \in E_1$ 是 p 的一个根). 综上, $[E_f : F] \leq n(n - 1) \cdots = n!$. □

6.11 应用: 构造有限域

我们已经在一些场合提到有限域, 然而除了在例 3.1.12 通过同余关系定义的 $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ (其中 p 是素数) 之外, 迄今还缺乏更多例子. 有限域不只在代数学的研究中举足轻重, 在密码学等领域中也极重要. 为了增进对有限域的了解, 我们先来确定它们的特征.

对于任意域 F, 根据 §3.7 结尾的讨论:

- ⋆ 若 $\mathrm{char}(F) = 0$, 则有自然的域嵌入 $\mathbb{Q} \hookrightarrow F$;
- ⋆ 若 $\mathrm{char}(F) = p$ 是素数, 则有自然的域嵌入 $\mathbb{F}_p \hookrightarrow F$.

若 F 是有限域, 显然只有第二种情形能发生. 因此可以按定义 6.10.6 来谈论扩域的次数 $[F:\mathbb{F}_p]$; 既然 F 有限, $[F:\mathbb{F}_p]$ 当然是一个正整数.

本节后续的讨论都将固定域的特征 p. 我们接着来考察有限域所能具有的元素个数.

命题 6.11.1 *设 p 为素数, F 为特征 p 的有限域, 则 $|F|=p^n$, 其中 $n:=[F:\mathbb{F}_p]$.*

证明 作为 $\mathbb{F}_p$-向量空间, F 同构于 n 份 $\mathbb{F}_p$ 的直和, 作为集合来看也就是 n 份 $\mathbb{F}_p$ 的积, 故 $|F|=|\mathbb{F}_p|^n=p^n$. □

对于所有 $n\in\mathbb{Z}_{\geq 1}$, 确实存在有 p^n 个元素的有限域. 抽象构造并不困难, 但需要引入一些定义.

首先, 称为 Frobenius 自同态的一类映射在特征 p 的世界中格外有用. 我们给出交换环版本的一般定义.

定义–命题 6.11.2 (Frobenius 自同态) *设 p 为素数, R 为交换环, 并且满足 $p\cdot 1_R=0_R$. 对于所有 $q=p^n$, 其中 $n\in\mathbb{Z}_{\geq 1}$, 映射*

$$\begin{aligned}\mathrm{Fr}_q : R &\to R\\ x &\mapsto x^q\end{aligned}$$

是环 R 的自同态, 称为 Frobenius 自同态.

证明 显然 Fr_q 保持 0_R 和 1_R 不动, 此外由于 R 交换, $\mathrm{Fr}_q(xy)=\mathrm{Fr}_q(x)\,\mathrm{Fr}_q(y)$ 也是显然的. 至于加法, 练习 3.7.3 说明 $(x+y)^p=x^p+y^p$, 迭代 n 次便是 $\mathrm{Fr}_q(x+y)=\mathrm{Fr}_q(x)+\mathrm{Fr}_q(y)$. □

留意到 $(\mathrm{Fr}_q)^m=\mathrm{Fr}_{q^m}$.

练习 6.11.3 说明若 F 是特征 p 的有限域, 则 $\mathrm{Fr}_p:F\to F$ 是自同构. 另一方面, 具体举出一个特征 p 的域, 使得 Fr_p 并非同构.

提示 域上的 Fr_p 总是单射. 非同构的例子可取有理函数域 $\mathbb{F}_p(X)$.

引理 6.11.4 *设 $\phi:R\to R$ 是环 R 的自同态, 则不动点集 $\{x\in R:\phi(x)=x\}$ 是 R 的子环. 若 R 是域, 则不动点集还是子域.*

证明 同态的定义确保 $0_R,1_R$ 都是不动点. 若 $\phi(x)=x$ 且 $\phi(y)=y$, 则 $\phi(x+y)=\phi(x)+\phi(y)=x+y$ 而 $\phi(xy)=\phi(x)\phi(y)=xy$, 此外还有 $\phi(-x)=-\phi(x)=-x$. 至此验证了子环所需的条件.

进一步设 R 为域, 而 $x\neq 0$ 满足 $\phi(x)=x$, 则我们有 $\phi(x^{-1})=\phi(x)^{-1}=x^{-1}$, 因而不动点集的非零元对取逆也封闭. □

命题 6.11.5 *设 p 为素数, $n\in\mathbb{Z}_{\geq 1}$, 则存在恰有 $q:=p^n$ 个元素的有限域.*

证明 考虑多项式 $f := X^q - X \in \mathbb{F}_p[X]$. 推论 6.10.5 给出 $\mathbb{F}_p$ 的某个扩域 L, 使得 f 在 L 上分裂. 定义

$$F := \{x \in L : x^q = x\};$$

既可以将 F 理解为 f 在 L 中的根集, 亦可理解为 Frobenius 自同态 $\mathrm{Fr}_q : L \to L$ 的不动点集. 引理 6.11.4 蕴涵 F 是 L 的子域.

其次, $p \mid q$ 蕴涵 $\mathbb{F}_p[X]$ 中的等式

$$f' = qX^{q-1} - 1 = -1.$$

特别地 f 和 f' 互素, 命题 6.6.8 (i) 遂确保 f 无重根. 既然 f 在 L 上分裂, 这相当于说 $|F| = q$.

综上可知 F 是恰有 q 个元素的域. 明所欲证. □

一旦掌握更多域论工具, 便能证明有 q 个元素的域不仅存在, 而且彼此同构, 尽管这种同构并不是唯一的 (比方说, 考虑练习 6.11.3 的 Frobenius 自同构). 细节超出本书范围, 可参考 [10, §9.3]. 在不致混淆的情形, 一般将有 q 个元素的有限域记为 $\mathbb{F}_q$.

习题

1. 设 F 为域. 说明 $\mathrm{M}_{n\times n}(F)$ 的双边理想只有 $\{0\}$ 和 $\mathrm{M}_{n\times n}(F)$ 本身.

2. 我们在 §6.1 定义了任意环 R 中的双边理想. 现在仍假设非空子集 $I \subset R$ 满足加法封闭性, 但考虑乘法封闭性放宽后的两种版本:

- ⋆ 如果对所有 $r \in R$ 皆有 $rI \subset I$, 则称 I 为**左理想**;
- ⋆ 如果对所有 $r \in R$ 皆有 $Ir \subset I$, 则称 I 为**右理想**.

设 F 为域, $n \in \mathbb{Z}_{\geq 1}$. 试确定 $\mathrm{M}_{n\times n}(F)$ 的所有左理想和右理想.

提示〉 对于左理想的情形, 考虑矩阵的列向量. 对右理想则考虑行向量.

3. 证明 a, b 互素当且仅当 $ab, a+b$ 互素, 这里的 a 和 b 可以是任何唯一分解环的元素.

4. 设 R 为交换环, $a \in R$ 而 $m, n \in \mathbb{Z}_{\geq 1}$. 证明理想的等式

$$\langle a^m - 1, a^n - 1\rangle = \left\langle a^{\gcd(m,n)} - 1 \right\rangle.$$

提示〉 辗转相除.

5. 运用上一题证明若 R 是主理想环, 则 $\gcd(a^m - 1, a^n - 1) \sim a^{\gcd(m,n)} - 1$.

6. 设 $D \in \mathbb{Z} \smallsetminus \{0, 1\}$ 无平方因子. 按照练习 3.1.14 的方式定义整环 $\mathbb{Z}[\sqrt{D}]$. 对所有 $x = a + b\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$, 记 $\mathrm{N}(x) := (a + b\sqrt{D})(a - b\sqrt{D}) = a^2 - Db^2 \in \mathbb{Z}$.

(i) 验证 $\mathrm{N}(1) = 1$ 和 $\mathrm{N}(xy) = \mathrm{N}(x)\,\mathrm{N}(y)$.

(ii) 证明 $x \in \mathbb{Z}[\sqrt{D}]^\times$ 当且仅当 $\mathrm{N}(x) = \pm 1$.

(iii) 证明若 $\mathrm{N}(x)$ 是 $\mathbb{Z}$ 中的素数, 则 x 是 $\mathbb{Z}[\sqrt{D}]$ 的不可约元.

(iv) 说明 $\mathbb{Z}[\sqrt{D}]$ 的任何素元 $\mathfrak{p}$ 都整除一个唯一的素数 p.

(v) 证明 $a + b\sqrt{D} \mapsto a - b\sqrt{D}$ 是环 $\mathbb{Z}[\sqrt{D}]$ 的自同构. 〈提示〉 当 $D < 0$ 时此映射来自复共轭, 因而自动保持加法和乘法结构, 但 $D > 0$ 的情形仍然需要论证.

本题的一些操作能够扩及更广泛的二次代数整数环, 不过基于教学的考量, 此处只看 $\mathbb{Z}[\sqrt{D}]$ 的特例.

7. 设 $D \in \mathbb{Z}_{\geq 1}$, 将目光转向整环 $\mathbb{Z}[\sqrt{-D}]$.

(i) 验证 $\mathrm{N}(x) = x\bar{x}$, 其中 $\bar{x}$ 表 x 的复共轭. 因此 $\mathrm{N}(x) \in \mathbb{Z}_{\geq 0}$.

(ii) 具体确定有哪些 D 使得 $\mathbb{Z}[\sqrt{-D}]^\times \supsetneq \{\pm 1\}$.

(iii) 证明 $\mathbb{Z}[\sqrt{-D}]$ 的非零元 x 或者可逆, 或者分解为不可约元的乘积. 〈提示〉 对 $\mathrm{N}(x)$ 递归地论证.

(iv) 代入 $D = 14$. 说明 3 和 5 作为 $\mathbb{Z}[\sqrt{-14}]$ 的元素皆不可约.
〈提示〉 以 3 为例, 若 $3 = xy$ 则 $9 = \mathrm{N}(x)\,\mathrm{N}(y)$; 如果 x, y 皆不可逆, 唯一可能是 $\mathrm{N}(x) = \mathrm{N}(y) = 3$, 然而方程 $a^2 + 14b^2 = 3$ 无整数解.

(v) 进一步说明 $1 \pm \sqrt{-14}$ 也不可约, 然后从 $3 \cdot 5 = (1 + \sqrt{-14})(1 - \sqrt{-14})$ 说明 $\mathbb{Z}[\sqrt{-14}]$ 不是唯一分解环.
〈提示〉 若 $1 \pm \sqrt{-14} = xy$ 则 $15 = \mathrm{N}(x)\,\mathrm{N}(y)$; 在 x, y 皆不可逆的前提下得到 $\mathrm{N}(x) = 3$ 或 $\mathrm{N}(y) = 3$, 和之前结果矛盾.

8. 整环 $\mathbb{Z}[\sqrt{-1}]$ 称为 Gauss 整数环.

(i) 证明带余除法: 对于任意 $a, b \in \mathbb{Z}[\sqrt{-1}]$, 若 $b \neq 0$, 则存在 $q \in \mathbb{Z}[\sqrt{-1}]$ 使得 $\mathrm{N}(a - qb) < \mathrm{N}(b)$.

(ii) 证明 $\mathbb{Z}[\sqrt{-1}]^\times = \{\pm 1, \pm \mathrm{i}\}$.

(iii) 仿照 §2.7 的进路, 证明 $\mathbb{Z}[\sqrt{-1}]$ 是唯一分解环.

(iv) 具体给出 2 在 $\mathbb{Z}[\sqrt{-1}]$ 中的不可约分解.

(v) 设 p 为素数, $p \equiv 1 \pmod 4$. 证明存在 $\mathbb{Z}[\sqrt{-1}]$ 的不可约元 $\mathfrak{p}$, 使得 $p = \mathfrak{p}\bar{\mathfrak{p}}$. 〈提示〉 一种方法是从练习 1.2.1 的结论出发.

(vi) 设 p 为素数, $p \equiv 3 \pmod 4$. 证明 p 在 $\mathbb{Z}[\sqrt{-1}]$ 中仍是不可约元. 〈提示〉 若可约, 则因子可以按照复共轭配对.

(vii) 描述 $\mathbb{Z}[\sqrt{-1}]$ 的所有不可约元, 按照它们所整除的素数来分类.

(viii) 试完整地描述 $X^2 + Y^2 = n$ 对哪些 n 有整数解.

相关讨论亦可参见 [10, 定理 5.7.9].

9. 设域 F 满足 $\mathrm{char}(F) = 0$, 而 $h \in F(X)$ 满足 $h' = 0$. 证明 $h \in F$. 〈提示〉 表 h 为既约分式.

10. 简单地证明多项式

$$f_n := 1 + \frac{X}{1!} + \cdots + \frac{X^n}{n!}, \quad n \in \mathbb{Z}_{\geq 1}$$

在 $\mathbb{C}$ 中无重根. 〈提示〉 探讨 f_n 和 f_n' 的最大公因式.

I. Schur 证明了形如 $1 + \sum_{m=1}^{n-1} c_m \frac{X^m}{m!} \pm \frac{X^n}{n!}$ 的多项式 $(c_0, \ldots, c_{n-1} \in \mathbb{Z})$ 在 $\mathbb{Q}[X]$ 中是不可约的. 证明需要一些数论的结论, 请感兴趣的读者参阅 [1].

11. 设 F 为域, 而 $f, g \in F[X]$ 满足 $\deg f < n$ 和 $g = \prod_{i=1}^{n}(X - x_i)$, 其中 $x_1, \ldots, x_n \in F$ 两两相异. 证明

$$\frac{f}{g} = \sum_{i=1}^{n} \frac{f(x_i)}{g'(x_i)(X - x_i)}.$$

以此重新推导例 6.6.7 的 Lagrange 插值公式. 〈提示〉 确定部分分式分解 $\frac{f}{g} = \sum_{i=1}^{n} \frac{a_i}{X - x_i}$ 中的系数 $a_1, \ldots, a_n$.

12. 考虑 $\mathbb{R}$ 上的多项式 $f = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$. 设 $a_0 > 0$ 而 $a_n \neq 0$. 说明 f 有偶数个正实根 (计重数) 当且仅当 $a_n > 0$.

13. (Descartes 符号律) 承上题, 在相同条件下, 证明若实数列 $a_0, \ldots, a_n$ 的变号数为 α, 则

(i) f 的正实根 (计重数) 个数 $\leq \alpha$,

(ii) f 的正实根 (计重数) 个数与 α 的奇偶性相同.

变号按如下方式计数: 给定一列非零实数 $c_0, \ldots, c_m$, 如果 $c_i c_{i+1} < 0$ 则称此数列在 i 和 $i+1$ 之间变号; 对于一般的实数列 $a_0, \ldots, a_n$, 定义变号数 α 为将零元剔除后的数列的变号数.

关于多项式的实根还有许多可说, 读者可以参考 [8, §7.3].

〈提示〉 涉及求导和数学分析中的 Rolle 定理.

14. 按照以下步骤完成初等对称多项式 $e_1, \ldots, e_n$ 的代数无关性 (定理 6.7.8) 之证明. 设 F 为域, $g \in F[X_1, \ldots, X_n]$ 给定. 目标是证 $g \neq 0 \implies g(e_1, \ldots, e_n) \neq 0$.

(i) 说明可以用 F 的任意扩域 L 代替 F, 将 g 视为 $L[X_1, \ldots, X_n]$ 的元素来处理.

(ii) 按此将问题化到 F 是无穷域的情形. 〈提示〉 将 F 嵌入有理函数域 $F(Y)$.

(iii) 在 F 无穷的前提下, 说明若 $g \neq 0$ 则存在 $(y_1, \ldots, y_n) \in F^n$ 使得 $g(y_1, \ldots, y_n) \neq 0$. 〈提示〉 代入命题 3.6.2.

(iv) 承上, 说明存在 F 的扩域 L 以及 $(x_1, \ldots, x_n) \in L^n$, 使得 $e_k(x_1, \ldots, x_n) = y_k$ 对 $1 \leq k \leq n$ 成立. 〈提示〉 以推论 6.10.5 取 L 使得多项式 $X^n - y_1 X^{n-1} + \cdots + (-1)^n y_n$ 分裂.

(v) 证明 $g \neq 0 \implies g(e_1, \ldots, e_n) \neq 0$. 〈提示〉 按以上方法取扩域, 然后说明对 $g(e_1, \ldots, e_n)$ 代入 $(x_1, \ldots, x_n)$ 取值非零.

15. 将以下关于 X, Y, Z 的对称多项式用初等对称多项式表出, 系数在 $\mathbb{Q}$ 上.

(i) $X^3Y + XY^3 + X^3Z + XZ^3 + Y^3Z + YZ^3$;

(ii) $X^4 + Y^4 + Z^4$;

(iii) $(XY+Z^2)(YZ+X^2)(XZ+Y^2)$.

16. 写下 $X^n-a\in\mathbb{C}[X]$ 的判别式. 〈提示〉 答案是 $(-a)^{n-1}(-1)^{\frac{n(n-1)}{2}}n^n$. 可考虑 $(X^n-1)'$ 在 ζ^j 处的取值以辅助计算, $\zeta:=e^{2\pi\mathrm{i}/n}$.

17. 证明结式满足 $\mathrm{Res}(fg,h)=\mathrm{Res}(f,h)\mathrm{Res}(g,h)$; 明确各项所涉及的 n,m 等参数.
〈提示〉 取适当扩域后, 不妨假设这些多项式皆分裂.

18. 设 f 和 g 分别为 n 次和 m 次多项式, 从上一题结果推导判别式 (6.8.2) 的乘积公式

$$\mathrm{disc}(fg)=\mathrm{disc}(f)\mathrm{disc}(g)\mathrm{Res}(f,g)^2.$$

19. (Newton 公式) 证明例 6.7.3 的 n 元幂和 $p_1,p_2,\dots$ 在 $1\leq k\leq n$ 时满足

$$p_k-e_1p_{k-1}+e_2p_{k-2}+\cdots+(-1)^{k-1}e_{k-1}p_1+(-1)^k ke_k=0,$$

而在 $k>n$ 时满足

$$p_k-e_1p_{k-1}+e_2p_{k-2}+\cdots+(-1)^n e_np_{k-n}=0.$$

以此给出将 p_k 用初等对称多项式展开的递归算法. 〈提示〉 必要时可参考 [10, 定理 5.8.7].

20. 分别在 $\mathbb{R}$ 和 $\mathbb{C}$ 上将 X^n-1 分解为不可约多项式.

21. 写下域 $\mathbb{F}_2:=\mathbb{Z}/2\mathbb{Z}$ 上所有次数 ≤ 4 的不可约首一多项式. 〈提示〉 次数为 3 的有 2 个, 次数为 4 的有 3 个.

22. 考虑素数 p 和有限域 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$ 上的有理函数域 $F:=\mathbb{F}_p(t)$. 证明 $f:=X^p-t\in F[X]$ 是不可约多项式, 而 $f'=0$. 〈提示〉 不可约性涉及整环 $\mathbb{F}_p[t]$ 上的 Gauss 引理和 Eisenstein 判别法; 见 §6.9 接近结尾处的讨论.

23. (分圆多项式) 对所有 $n\in\mathbb{Z}_{\geq 1}$ 在有理函数域 $\mathbb{Q}(X)$ 中定义

$$\Phi_n:=\prod_{d|n}(X^d-1)^{\mu(n/d)};$$

此处的 μ 是练习 2.8.9 介绍的 Möbius 函数.

(i) 证明 $\prod_{d|n}\Phi_d=X^n-1$.

〈提示〉 归结为对所有 $m\in\mathbb{Z}_{\geq 1}$ 证 $\sum_{h|m}\mu(h)=\begin{cases}1, & m=1,\\ 0, & m>1\end{cases}$. 取 m 的因数分解, 将问题化到 m 为素数幂的简单情形.

(ii) 证明 $\Phi_n\in\mathbb{Z}[X]$ 是 $\varphi(n)$ 次首一多项式, 称为第 n 个**分圆多项式**; 关于 Euler 函数 $\varphi(n)$ 请见定义 2.8.7.

〈提示〉 对 n 递归地论证, 并且应用练习 2.8.8 的 $\sum_{d|n}\varphi(d)=n$, 或者取道 (iii).

(iii) 说明在复数域上有分解 $\Phi_n=\prod_{\substack{1\leq a\leq n\\ \gcd(a,n)=1}}(X-\zeta^a)$, 其中 $\zeta:=e^{2\pi\mathrm{i}/n}$.

(iv) 证明当 $n>1$ 时 Φ_n 的常数项是 1.

(v) 对所有 $n \in \mathbb{Z}_{\geq 1}$ 和素数 p, 证明

$$\Phi_{pn}(X) = \begin{cases} \Phi_n(X^p), & p \mid n, \\ \Phi_n(X^p)/\Phi_n(X), & p \nmid n; \end{cases}$$

按此计算 $\Phi_1, \dots, \Phi_{10}$.

分圆多项式 Φ_n 对所有 $n \geq 1$ 皆是 $\mathbb{Q}[X]$ 中的不可约多项式, 详见 [10, 定理 9.4.5].

24. 证明分圆多项式满足 $\Phi_n(X) = \Phi_{\mathrm{rad}(n)}\left(X^{n/\mathrm{rad}(n)}\right)$, 其中 $\mathrm{rad}(n)$ 定义为 n 的素因子的乘积.

25. 证明 $f := \prod_{k=1}^{18}(X-k) + 23 \in \mathbb{Q}[X]$ 不可约.

提示 设有 $\mathbb{Z}[X]$ 中的分解 $f = gh$ 满足 $1 \leq \deg g \leq 9$. 说明存在 $q \in \mathbb{Z}[X]$, 整数 $1 \leq a_1 < \cdots < a_5 \leq 18$ 和 $r \in \{\pm 1, \pm 23\}$ 使得 $g = \prod_{i=1}^{5}(X-a_i)q + r$. 说明存在整数 $a \in [1,18] \smallsetminus \{a_1, \dots, a_5\}$ 使得 $g(a) \neq r$, 从而推得

$$\prod_{i=1}^{5}(a-a_i)q(a) = g(a) - g(a_1) \in \{\pm 2, \pm 22, \pm 24, \pm 46\},$$

进而导出矛盾.

26. 设 R 为唯一分解环, $K := \mathrm{Frac}(R)$. 依照 §6.9 结尾的线索, 说明 $R[X_1, \dots, X_n]$ 也是唯一分解环, 并且进一步分类 $R[X_1, \dots, X_n]$ 的不可约元.

▷ **第一类** R 中的不可约元,

▷ **第二类** 满足以下条件的非常数多项式 $f = \sum_{i_1,\dots,i_n} c_{i_1,\dots,i_n} X_1^{i_1} \cdots X_n^{i_n}$:

- ⋆ $c(f) = c_R(f) := \gcd\{c_{i_1,\dots,i_n} : i_1, \dots, i_n \geq 0\} \sim 1$,
- ⋆ f 在 $K[X_1, \dots, X_n]$ 中不可约.

提示 一种策略是先处理 $n = 1$ 情形, 再以 $R[X_1, \dots, X_n] \simeq R[X_1] \cdots [X_n]$ 递归地推导 $R[X_1, \dots, X_n]$ 是唯一分解环. 关于非常数不可约元的分类在 $n > 1$ 时稍费工夫. 可以命 $D := R[X_1, \dots, X_{n-1}]$, 并将 f 视为 $D[X_n]$ 的元素以定义 $c_D(f)$.

- ⋆ 先说明非常数的 f 不可约等价于 $c_D(f) \sim 1$ 而 f 是 $K(X_1, \dots, X_{n-1})[X_n]$ 的不可约元;
- ⋆ 其次, 从 $n = 1$ 情形论证上述条件也等价于 $c(f) \sim 1$ 而 f 是 $K[X_1, \dots, X_{n-1}][X_n]$ 的不可约元.

另一种可能的策略是证明某种多元版本的 Gauss 引理.

27. 说明当 F 是域而 $n > 1$ 时 $F[X_1, \dots, X_n]$ 是唯一分解环, 却非主理想环.

28. 设 F 为域. 考虑方程

$$f(x,y) = g(x,y) = 0,$$
$$f, g \in F[X,Y] \smallsetminus \{0\}, \quad (x,y) \in F^2.$$

集项将 f 和 g 视作以 Y 为变元, 系数在 $F[X]$ 上的一元多项式, 次数分别为 n 和 m. 对此取结式

$$R := \mathrm{Res}_Y(f,g) \ \in F[X].$$

(i) 设 $(x,y)\in F^2$ 满足 $f(x,y)=g(x,y)=0$, 说明 $R(x)=0$.

(ii) 按此给出求解 $f(x,y)=g(x,y)=0$ 的算法.

提示〉 当 $R\neq 0$ 时, 对 R 的每个根 x 求 $f(x,Y)$ 和 $g(x,Y)$ 在 $F[Y]$ 中的最大公因式. 当 $R=0$ 时, 说明 f 和 g 在 $F[X,Y]$ 中有公因式 h, 其中 h 对 Y 的次数 ≥ 1; 这部分可能需要先前习题的结果.

(iii) 说明当 $R\neq 0$ 时, 或者 f 和 g 有形如 $X-x_0$ 的公因式 ($x_0\in F$), 或者 $f(x,y)=g(x,y)=0$ 的解 (x,y) 个数有限.

(iv) 对以下两个特例确定所有解 $(x,y)\in\mathbb{C}^2$:

$$f=X^2Y^2-25X^2+9,\quad g=4X+Y,$$
$$f=5X^2-6XY+5Y^2-16,\quad g=2X^2-XY+Y^2-X-Y-4.$$

对于更复杂的高次方程组, 结式求解将变得不切实际.

结式的另一个相关用法是研究射影平面 $\mathbb{P}^2$ 上两条代数曲线的相交数, 这点宜留给代数几何的教材来解说.

29. (Reed–Solomon 码) 回顾第四章习题最后部分提及的编码理论, 以及 Hamming 距离的概念. 以下介绍的 Reed–Solomon 算法是纠错码的重要例子, 涉及有限域上的线性方程组与多项式理论.

(i) 设 $n\geq k\geq 1$. 给定域 F 的一列相异元素 $\alpha_1,\dots,\alpha_n$. 对所有满足 $\deg f\leq k-1$ 的 $f\in F[X]$ 定义 $\boldsymbol{e}_f:=(f(\alpha_i))_{i=1}^n\in F^n$. 说明若 $f_1\neq f_2$, 则 F^n 上的 Hamming 距离满足 $d(\boldsymbol{e}_{f_1},\boldsymbol{e}_{f_2})>n-k$.

(ii) 考虑如上的 $(\alpha_i)_{i=1}^n$, 另取 $z_1,\dots,z_n\in F$. 说明存在

$$g,h\in F[X],\quad \deg g\leq\frac{n+k-2}{2},\quad \deg h\leq\frac{n-k}{2},$$

使得 h 非零, 而且 $g(\alpha_i)=z_ih(\alpha_i)$ 对所有 $1\leq i\leq n$ 成立.

提示〉 解齐次线性方程组.

(iii) 承上, 说明若 $f\in F[X]$ 满足 $\deg f\leq k-1$, 而且存在子集 $\mathcal{E}\subset\{1,\dots,n\}$ 使得

- $|\mathcal{E}|\leq\frac{n-k}{2}$,
- $f(\alpha_i)=z_i$ 对所有 $i\notin\mathcal{E}$ 成立,

则 $g=fh$.

提示〉 说明 $\deg(g-fh)<\frac{n+k}{2}$, 而且 $(g-fh)(\alpha_i)=0$ 对所有 $i\notin\mathcal{E}$ 成立.

(iv) 取 Σ 为至少有 n 个元素的有限域 F, 取相异元素 $\alpha_1,\dots,\alpha_n\in F$.

- 编码映射 E 定义为 $E(a_0,\dots,a_{k-1})=\boldsymbol{e}_f$, 其中 $f:=\sum_{i=0}^{k-1}a_iX^i\in F[X]$; 注意到这是线性映射.
- 解码映射 D (非唯一) 定义如下: 给定 $(z_1,\dots,z_n)\in F^n$, 按之前的方法计算 $g,h\in F[X]$, 然后选取相异元素 $x_1,\dots,x_k$ 使得 $h(x_i)\neq 0$, 解出唯一的 f 使得 $f(x_i)=g(x_i)/h(x_i)$.

说明此码能纠正至多 $\frac{n-k}{2}$ 位错误, 能纠正至多 $n-k$ 位删除.

(v) 承上, 尽可能分析 E 和 D 在算法面向的时间复杂度, 说明可以只用到关于 n,k 的多项式.

第七章 对角化

设 V 为域 F 上的有限维向量空间, $T : V \to V$ 为线性映射. 如果 $\lambda \in F$ 和 $v \in V \smallsetminus \{0\}$ 满足 $Tv = \lambda v$, 则称 λ 为 T 的特征值, 而 v 称为相应的特征向量. 特征值能等价地描述为特征多项式 Char_T 在 F 中的根. 这一切当然也有矩阵表述.

特征值和特征向量在物理学, 工程和信息科学中有广泛的应用, 在基础数学中更是无处不在. 就代数学本身而言, 特征值和特征向量与对角化问题直接相关. 何谓对角化?

- ⋆ 对于先前的线性映射 T, 对角化相当于寻求 V 的基 $v_1, \dots, v_n$ 连同 $\lambda_1, \dots, \lambda_n \in F$, 使得 $Tv_i = \lambda_i v_i$ 对 $i = 1, \dots, n$ 成立.
- ⋆ 用矩阵的语言来说, 将 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 对角化相当于使其共轭于对角矩阵, 亦即寻求可逆矩阵 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$ 使得

$$\boldsymbol{D} := \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} \text{ 为对角矩阵 } \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

我们将在 §7.1 详细解释这些定义以及其间的等价性. 一般的 T 或 $\boldsymbol{A}$ 未必能对角化; 若可对角化, 则涉及的 $\lambda_1, \dots, \lambda_n$ 恰好是特征多项式的 n 个根, 计重数不计顺序. 一旦解出特征多项式的所有根, 则关于可对角化与否, 以及 $v_1, \dots, v_n$ (或 $\boldsymbol{P}$) 的计算都能以 Gauss–Jordan 消元法料理. 由于涉及求根, 处理对角化问题时往往要求 F 是定义 6.6.5 所谓的代数闭域, 例如 $F = \mathbb{C}$, 或至少要求特征多项式在 F 上分裂 (定义 6.6.2).

对角矩阵 $\boldsymbol{D}$ 的运算特别简单. 若 $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 则不难证明 $\boldsymbol{A}^N = \boldsymbol{P}\boldsymbol{D}^N\boldsymbol{P}^{-1}$, 从而简化 $\boldsymbol{A}^N$ 的计算 ($N \in \mathbb{Z}_{\geq 0}$). 作为对角化问题的动机之一, 例 7.1.1 将从这一观察探讨线性递归数列的通解, 而例 7.1.9 将演示 Fibonacci 数列的特例.

在 §7.2, 我们介绍线性映射 T (或矩阵 $\boldsymbol{A}$) 的极小多项式 Min_T (或 $\mathrm{Min}_{\boldsymbol{A}}$); 它是特征多项式的首一因式, 而且两者在 F 上有相同的根集. 由此可推得关于对角化问题的另一种判准. 这部分依赖于多项式环的性质, 见第六章.

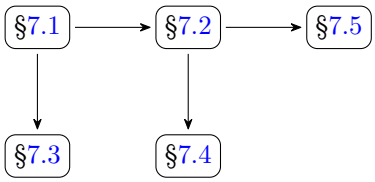
在 §7.3, 我们将对线性映射或矩阵探讨上三角化. 可上三角化的充要条件是特征多项式分裂, 例如 F 为代数闭域时总是如此. 尽管条件宽松, 上三角化有助于处理许多问题, 例如它能给出 Cayley–Hamilton 定理 5.8.8 的另一种证明, 见注记 7.3.7.

广义特征子空间和相应的直和分解 (7.4.1) 是 §7.4 的主题, 仅需要特征多项式分裂的条件. 在此基础上, 定义 7.4.4 介绍特征值 λ 的几何重数与代数重数; 代数重数总大于或等于几何重数, 而可对角化等价于两种重数对所有 λ 皆相等 (推论 7.4.6).

广义特征子空间理论还给出称为 Jordan–Chevalley 分解的实用工具, 分成加性和乘性两种版本 (定理 7.4.8, 7.4.9). 这些内容全都涉及关于多项式的知识.

最后的 §7.5 介绍同步对角化的概念, 这相当于寻求基 $v_1,\dots,v_n$ (或可逆矩阵 $\boldsymbol{P}$), 借之将一族线性映射 (或矩阵) $\mathcal{S}$ 的成员同时对角化. 定理 7.5.3 说明 $\mathcal{S}$ 可同步对角化的充要条件是每个 $T\in\mathcal{S}$ 皆可对角化, 而且它们对乘法两两交换.

阅读顺序



7.1 特征值与特征向量

选定域 F 并考虑有限维 F-向量空间 V. 对于从 V 到其自身的线性映射 T, 或者更具体的 $m\times m$ 矩阵, 此前的一系列实例和习题已经充分表明它们的乘法运算在高维时可以极端复杂. 为了克服这一困难, 我们从两方面探寻线索.

⋆ 如果 V 带有合适的直和分解 $V=V_1\oplus\cdots\oplus V_n$, 则分块对角的线性变换

$$T=\operatorname{diag}(T_1,\dots,T_n)\overset{\text{表为矩阵}}{=\!=}\begin{array}{c}\begin{array}{ccc}V_1&\cdots&V_n\end{array}\\ \left(\begin{array}{c|c|c} T_1 & & \\ \hline & \ddots & \\ \hline & & T_n \end{array}\right)\end{array}\begin{array}{c}V_1\\ \vdots\\ V_n\end{array}$$

特别容易处理: 运算可以逐块操作

$$\begin{aligned}\operatorname{diag}(S_1,\dots,S_n)+\operatorname{diag}(T_1,\dots,T_n)&=\operatorname{diag}(S_1+T_1,\dots,S_n+T_n),\\ \operatorname{diag}(S_1,\dots,S_n)\operatorname{diag}(T_1,\dots,T_n)&=\operatorname{diag}(S_1T_1,\dots,S_nT_n).\end{aligned}$$

这就提示我们寻求适当的直和分解, 使得给定的线性变换 T 具有简单的分块形式, 例如对角分块.

⋆ 线性映射 T 的矩阵表达式 $\boldsymbol{A}=\boldsymbol{\mathcal{M}}(T)$ 依赖于基的选取. 基的变换在 §4.9 已有讨论, 其中最重要的结论是推论 4.9.4: 换基的效果相当于将 $\boldsymbol{A}$ 换作其共轭

$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 其中 $\boldsymbol{P}$ 是不同基之间的坐标转换矩阵. 将先前的讨论翻译为具体的矩阵语言, 则问题在于寻求合适的可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 具有简单的分块形式.

⋆ 在探讨多个线性映射或矩阵之间的运算时, 上述手续应当在相同的直和分解或转换矩阵 $\boldsymbol{P}$ 下操作, 方能和运算兼容. 以矩阵情形为例,

$$\begin{aligned}\boldsymbol{P}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{P}&=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}+\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P},\\ \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{B}\boldsymbol{P}&=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\cdot\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}.\end{aligned}\tag{7.1.1}$$

最简单的莫过于每个分块都是 1×1 的情形, 这相当于 $V=V_1\oplus\cdots\oplus V_n$ 的每个直和项都是 1 维的; 此时对每个 $1\leq i\leq n$ 取 $v_i\in V$ 使得 $V_i=\langle v_i\rangle$, 则 $v_1,\dots,v_n$ 成为 V 的基. 反过来说, V 的任意有序基 $v_1,\dots,v_n$ 都通过 $V_i:=\langle v_i\rangle$ 给出 V 的直和分解. 由于 T 是分块对角的相当于 $T(V_i)\subset V_i$ 恒成立, 因此在 $V_i=\langle v_i\rangle$ 的前提下, 这也导致存在唯一的 $\lambda_i\in F$ 使得 $Tv_i=\lambda_i v_i$.

在最优的情境下, 我们的愿望因而相当于寻求 V 的基 $v_1,\dots,v_n$ 连同 $\lambda_1,\dots,\lambda_n\in F$, 使得

$$Tv_i=\lambda_i v_i,\quad i=1,\dots,n.\tag{7.1.2}$$

采取矩阵语言, 则这也相当于寻求可逆矩阵 $\boldsymbol{P}$ 和 $\lambda_1,\dots,\lambda_n\in F$ 使得

$$\boldsymbol{D}:=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\text{ 为对角矩阵 }\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}.\tag{7.1.3}$$

上式称为矩阵的**对角化**问题. 两种表述 (7.1.2) 和 (7.1.3) 的等价性也可以如下理解: 将 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 视同线性映射 $F^n\to F^n$, 则给定可逆矩阵 $\boldsymbol{P}$ 相当于给定 F^n 的有序基 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$, 其中 $\boldsymbol{v}_i=\boldsymbol{P}\boldsymbol{e}_i$ 是 $\boldsymbol{P}$ 的第 i 列, 而 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}$ 或等价的 $\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{D}$ 便相当于说

$$\boldsymbol{A}\boldsymbol{v}_i=\boldsymbol{A}\boldsymbol{P}\boldsymbol{e}_i=\boldsymbol{P}\boldsymbol{D}\boldsymbol{e}_i=\lambda_i\boldsymbol{v}_i,\quad i=1,\dots,n,$$

而这无非是 (7.1.2) 的内涵.

例 7.1.1 (线性递归数列) 选定 $n\in\mathbb{Z}_{\geq 1}$ 和 $c_0,\dots,c_{n-1}\in F$. 考虑满足

$$R_N+c_{n-1}R_{N-1}+\cdots+c_0R_{N-n}=0$$

的列 $(R_N)_{N\in\mathbb{Z}_{\geq 0}}$, 要求 $R_N\in F$. 这样的列完全由初值 $(R_0,\dots,R_{n-1})\in F^n$ 和上述递归关系式确定. 因为递归关系式是关于 $R_N,\dots,R_{N-n}$ 的线性关系式, 这也称为 n 阶线性递归数列, 在计算机科学中用处不少. 它们可以利用矩阵工具来处理. 关键是将递归关系式改写为矩阵等式

$$(R_{N-n+1}\ \cdots\ R_N)=(R_{N-n}\ \cdots\ R_{N-1})\boldsymbol{C},\quad N\in\mathbb{Z}_{\geq n},$$

其中的 $n\times n$ 矩阵

$$\boldsymbol{C}:=\begin{pmatrix}0&0&\cdots&0&-c_0\\1&0&\cdots&0&-c_1\\0&1&\cdots&0&-c_2\\\vdots&&\ddots&&\vdots\\0&0&\cdots&1&-c_{n-1}\end{pmatrix}$$

是我们在例 5.8.7 打过照面的友矩阵. 反复迭代给出

$$(R_N\ \cdots\ R_{N+n-1})=(R_0\ \cdots\ R_{n-1})\boldsymbol{C}^N,\quad N\in\mathbb{Z}_{\geq 0}.$$

线性递归数列的通解因之化为 $\boldsymbol{C}^N$ 的计算. 倘若能明确地施行对角化 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P}$, 则反复应用 (7.1.1) 可见

$$\boldsymbol{C}^N=(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^N=\boldsymbol{P}\boldsymbol{D}^N\boldsymbol{P}^{-1}$$

对所有 N 皆成立, 而 $\boldsymbol{D}^N$ 是容易计算的.

本章的习题部分将给出求线性递归数列通解的另一种进路.

对于给定的矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 或相应的线性映射, 对角化拆分为以下几个问题:

- ⋆ 如何判定 $\boldsymbol{A}$ 是否共轭于一个对角矩阵 $\boldsymbol{D}$?
- ⋆ 如果答案是肯定的, $\boldsymbol{D}$ 在何种意义下由 $\boldsymbol{A}$ 确定?
- ⋆ 如何具体地求出 $\boldsymbol{P}$ 和 $\boldsymbol{D}$ 的对角元 $\lambda_1,\dots,\lambda_n$?

以下主要采取线性映射的视角, 逐一地梳理这些问题.

定义 7.1.2 (特征值和特征向量) 设 $T\in\mathrm{End}(V)$ 而 $\lambda\in F$.

- ⋆ 称子空间 $V_\lambda:=\ker(\lambda\cdot\mathrm{id}_V-T)$ 为 T 的 λ-**特征子空间**, 若 $V_\lambda\neq\{0\}$ 则称 λ 为 T 的**特征值**.
- ⋆ 若 $v\in V_\lambda$ (换言之 $Tv=\lambda v$), 而且 $v\neq 0$, 则称 v 是 T 的一个**特征向量**, 以 λ 为其特征值.

由于特征向量 v 按定义必须非零, 对应的特征值 λ 由 $Tv=\lambda v$ 唯一确定. 以上定义暂未要求 V 有限维.

下述定义总结了先前关于对角化的讨论, 特别是 (7.1.2) 和 (7.1.3).

定义 7.1.3 (可对角化线性变换和矩阵) 设 $n\in\mathbb{Z}_{\geq 1}$. 如果 n 维向量空间 V 有基 $v_1,\dots,v_n$ 使得每个 v_i 都是 T 的特征向量, 则称 T 在 F 上是**可对角化的**.

若将矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 看作线性映射 $F^n\to F^n$, 则 $\boldsymbol{A}$ 在 F 上可对角化相当于说存在可逆的 $\boldsymbol{P}\in\mathrm{M}_{n\times n}(F)$, 使得 $\boldsymbol{D}:=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角矩阵.

在矩阵 $\boldsymbol{A}$ 可对角化的情形, 设 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n \in F^n$ 对应的特征值依序是 $\lambda_1,\dots,\lambda_n$, 则本节开头的讨论说明 $\boldsymbol{P}$ 和 $\boldsymbol{D}$ 可以分别取为

$$\boldsymbol{P} = \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{array} \right), \quad \boldsymbol{D} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

研究对角化的第一步是确定特征值. 根据古早推论 4.8.5 和命题 5.7.1,

$$V_\lambda \neq \{0\} \iff \det(\lambda \cdot \mathrm{id}_V - T) = 0.$$

然而右式无非特征多项式 $\mathrm{Char}_T \in F[X]$ (定义 5.8.2) 在 λ 处的值. 因此

$$\lambda \text{ 是 } T \text{ 的特征值} \iff \mathrm{Char}_T(\lambda) = 0. \tag{7.1.4}$$

注记 7.1.4 设 Char_T 分裂 (见定义 6.6.2), 对其应用根与系数的关系, 可见 T 的行列式 (或迹) 无非是特征值之乘积 (或和), 记入重数; 此处所谓特征值的重数, 是作为 Char_T 的根而言的, 这也正是定义 7.4.4 将要介绍的代数重数. 要求 Char_T 分裂是为了获取所有根; 一般情形下, 推论 6.10.5 说明总是可以过渡到 F 的扩域来确保 Char_T 分裂.

命题 7.1.5 命 $n := \dim V \in \mathbb{Z}_{\geq 1}$. 若 $T \in \mathrm{End}(V)$ 在 F 上可对角化, 换言之若存在 V 的基 $v_1,\dots,v_n$ 和 $\lambda_1,\dots,\lambda_n \in F$ 使得 $Tv_i = \lambda_i v_i$ 恒成立, 则

$$\mathrm{Char}_T = \prod_{i=1}^{n} (X - \lambda_i),$$

而且对于每个 $\lambda \in F$, 我们有

$$\dim V_\lambda = |\{1 \leq i \leq n : \lambda_i = \lambda\}| = \lambda \text{ 作为 } \mathrm{Char}_T \text{ 的根的重数}.$$

作为推论, $\lambda_1,\dots,\lambda_n$ 精确到重排是唯一由 T 确定的.

证明 在有序基 $v_1,\dots,v_n$ 之下计算特征多项式, 得到

$$\mathrm{Char}_T = \begin{vmatrix} X-\lambda_1 & & \\ & \ddots & \\ & & X-\lambda_n \end{vmatrix} = \prod_{i=1}^{n} (X - \lambda_i).$$

这就给出第一部分. 对于第二部分, 我们将基 $v_1,\dots,v_n$ 适当地重排, 使得对应的特征值为

$$\underbrace{\lambda_1,\dots,\lambda_{n-d}}_{\text{皆} \neq \lambda}, \underbrace{\lambda_{n-d+1},\dots,\lambda_n}_{\text{皆} = \lambda}, \quad 0 \leq d \leq n.$$

于是 $d = |\{1 \leq i \leq n : \lambda_i = \lambda\}|$; 另一方面, T 相对于此有序基的矩阵 $\boldsymbol{A}$ 满足

$$\lambda \cdot \mathbf{1}_{n\times n} - \boldsymbol{A} = \left(\begin{array}{c|c} \boldsymbol{B} & \\ \hline & \mathbf{0}_{d\times d} \end{array}\right), \quad \boldsymbol{B} := \begin{pmatrix} \lambda - \lambda_1 & & \\ & \ddots & \\ & & \lambda - \lambda_{n-d} \end{pmatrix}.$$

这已经是简化行梯矩阵, 以此算出 $\lambda \cdot \mathrm{id}_V - T$ 的秩是 $n - d$, 故 $d = \dim V_\lambda$.

最后, 既然多项式的根 (计重数) 精确到重排是唯一确定的, 资料 $\lambda_1, \ldots, \lambda_n$ 亦然.□

由此可见 Char_T 分裂是 T 可对角化的必要条件, 尽管它远非充分的.

引理 7.1.6 设 $\lambda_1, \ldots, \lambda_m \in F$ 两两相异, 而 $v_i \in V_{\lambda_i}$ (其中 $i = 1, \ldots, m$) 满足 $\sum_{i=1}^m v_i = 0$, 则 $v_1 = \cdots = v_m = 0$.

作为推论, 如果 $v_1, \ldots, v_m \in V$ 是依序对应到相异特征值 $\lambda_1, \ldots, \lambda_m$ 的特征向量, 则它们线性无关.

证明 对于所有满足 $v_i \in V_{\lambda_i}$ 的线性关系式 $\sum_{i=1}^m v_i = 0$, 记

$$J := \{1 \leq i \leq m : v_i \neq 0\}.$$

我们的目标是证明 $J = \varnothing$ 恒成立. 设若不然, 存在至少一个线性关系式 $\sum_{i=1}^m v_i = 0$ 使得对应的 J 非空, 但 J 的元素个数尽可能少. 此时必然有 $|J| > 1$, 否则存在 i 使得 $J = \{i\}$, 然而这蕴涵 $v_i = 0$, 矛盾.

因此可取相异的 $j, j' \in J$. 我们有

$$\sum_{i=1}^m \lambda_i v_i = T\left(\sum_{i=1}^m v_i\right) = 0.$$

将上式与 $\lambda_j \sum_{i=1}^m v_i = 0$ 相减, 得到

$$\sum_{\substack{1\leq i\leq m \\ i\neq j}} (\lambda_i - \lambda_j) v_i = 0.$$

此线性关系式的非零项严格少于 $|J|$ 个, 因为扣掉了第 j 项; 另一方面, $(\lambda_{j'} - \lambda_j) v_{j'} \neq 0$, 所以仍有非零项. 这与 J 的选法矛盾. □

现在可以给出可对角化线性映射的第一种刻画.

定理 7.1.7 对于 $T \in \mathrm{End}(V)$, 以下性质相互等价:

(i) T 或对应的矩阵在 F 上可对角化;

(ii) $\sum_{\lambda\in F} \dim V_\lambda = \dim V$;

(iii) $\bigoplus_{\substack{\lambda\in F\\ T\text{ 的特征值}}} V_\lambda = V$.

证明 首先留意到仅当 λ 是 T 的特征值时才有 $V_\lambda \neq \{0\}$, 既然 Char_T 的根数有限, (ii) 和 (iii) 的左式都是有限和. 记 T 的相异特征值为 $\lambda_1,\dots,\lambda_m$, 则引理 7.1.6 说明对应特征子空间的和是直和

$$V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_m}\subset V; \tag{7.1.5}$$

读者如有疑惑, 请回顾定义–命题 4.10.2 之前的讨论及其论证.

设 (i) 成立. 命题 7.1.5 说明 $\dim V_{\lambda_i}$ 是 λ_i 作为 Char_T 的根的重数, 而且 Char_T 分裂. 因此 $\sum_{i=1}^m \dim V_{\lambda_i}$ 相当于将根的重数加总, 给出 $\deg \mathrm{Char}_T$, 即 (ii) 所要求的 $\dim V$.

设 (ii) 成立. 比较维数知 (7.1.5) 化为等式, 此即 (iii).

设 (iii) 成立. 对每个 V_{λ_i} 取基, 它们的并给出 V 的一组基, 由特征向量组成. 此即 (i). □

推论 7.1.8 命 $n := \dim V$. 如果 T 有 n 个相异的特征值 $\lambda_1,\dots,\lambda_n\in F$, 则 T 在 F 上可对角化.

证明 因为 $V_{\lambda_1},\dots,V_{\lambda_n}$ 皆非零, 故 $\sum_{i=1}^n \dim V_{\lambda_i}\geq n=\dim V$. 结合 (7.1.5) 即见定理 7.1.7 (ii) 的等式成立, T 在 F 上可对角化. □

例 7.1.9 (Fibonacci 数列) 回到例 7.1.1 的线性递归数列. 取特例 $F=\mathbb{C}$, $n=2$ 和

$$R_N - R_{N-1} - R_{N-2} = 0, \quad N\in\mathbb{Z}_{\geq 2}.$$

对应的矩阵是

$$\boldsymbol{C} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

为了求 R_N 在初值 (R_0,R_1) 给定时的公式, 我们希望将 $\boldsymbol{C}$ 对角化以计算 $\boldsymbol{C}^N$. 首先

$$\mathrm{Char}_{\boldsymbol{C}} = \begin{vmatrix} X & -1 \\ -1 & X-1 \end{vmatrix} = X^2 - X - 1.$$

有两根 $\frac{1\pm\sqrt{5}}{2}$. 解相应的特征向量相当于解齐次线性方程组

$$\begin{pmatrix} \frac{1+\sqrt{5}}{2} & -1 \\ -1 & \frac{-1+\sqrt{5}}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$\begin{pmatrix} \frac{1-\sqrt{5}}{2} & -1 \\ -1 & \frac{-1-\sqrt{5}}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

两者的解空间都是 1 维的, 生成元不妨取为列向量

$$v_1 = \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

因此

$$\boldsymbol{D} := \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} = \boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P}, \quad \boldsymbol{P} := \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

以此求 $\boldsymbol{C}^N$, 进而求 R_N, 不过如此得到的通解可能略显繁琐. 取巧的方法是观察到 $\boldsymbol{C}^N = \boldsymbol{P}\boldsymbol{D}^N\boldsymbol{P}^{-1}$ 的矩阵元总是 $\left(\frac{1+\sqrt{5}}{2}\right)^N$ 和 $\left(\frac{1-\sqrt{5}}{2}\right)^N$ 的线性组合, 其系数和 N 无关. 因此存在无关 N 的常数 $a, b \in \mathbb{R}$ 使得

$$R_N = a\left(\frac{1+\sqrt{5}}{2}\right)^N + b\left(\frac{1-\sqrt{5}}{2}\right)^N, \quad N \in \mathbb{Z}_{\geq 0}.$$

根据第五章的一道简单习题 (涉及 Vandermonde 行列式), 系数 a, b 是唯一确定的. 代入初值 R_0, R_1 可以容易地反解 a, b, 不必计算特征向量.

算法 7.1.10 总结上述讨论, 对角化可以分成为三步:

1. 确定线性映射或矩阵的特征多项式, 找出它的所有根, 亦即特征值;
2. 对每个特征值 λ 确定相应的特征子空间 V_λ, 并找出一组基.
3. 判断这些 V_λ 是否张成整个空间 V.

第一步是多项式求根的问题, 第二步本质上是解线性方程组, 第三步则关乎维数的估计.

为了确立更完整的理论, 我们需要进一步的铺垫.

7.2 极小多项式

本节将用到多项式环 $F[X]$ 的一些代数性质, 包括主理想环性质以及唯一分解性. 见 §6.2.

对任意 F-向量空间 V, 线性映射 $T \in \mathrm{End}(V)$ 和多项式 $h \in F[X]$, 按照 (5.8.1) 的方法向 T 代值, 得到线性映射 $h(T) \in \mathrm{End}(V)$; 以此定义子空间

$$\begin{aligned} V[h] &:= \ker(h(T)) \\ &= \{v \in V : h(T)v = 0\}. \end{aligned} \tag{7.2.1}$$

注意到 $T(V[h]) \subset V[h]$, 这是因为 T 和自身的任意次幂相交换, 故

$$Th(T) = h(T)T, \quad h(T)v = 0 \implies h(T)Tv = Th(T)v = 0;$$

换言之, $V[h]$ 是定义 5.10.1 所谓的 T-不变子空间.

如果取 $h = X - \lambda$, 则 $V[h]$ 无非是定义 7.1.2 介绍的特征子空间 V_λ, 因而 V_λ 也是 T-不变的.

引理 7.2.1 *设 $f \in F[X]$ 分解为 $f = gh$, 其中 $g, h \in F[X]$ 互素. 当 $T \in \mathrm{End}(V)$ 给定, 我们有直和分解*

$$V[f] = V[g] \oplus V[h].$$

证明 因为 g, h 互素, 命题 6.3.5 蕴涵存在 $a, b \in F[X]$ 使得 $ag + bh = 1$. 对于任意 $v \in V$, 我们有

$$v = 1 \cdot v = (ag + bh)(T) \cdot v = a(T)g(T)v + b(T)h(T)v.$$

然而当 $v \in V[f]$ 时 $g(T)v \in V[h]$ 而 $h(T)v \in V[g]$, 而且 $V[g]$ 和 $V[h]$ 都是 T-不变子空间. 综上可见

$$V[f] = V[g] + V[h].$$

其次, 若 $v \in V[g] \cap V[h]$, 则同样由 $v = a(T)g(T)v + b(T)h(T)v$ 可得 $v = 0$. 这就验证了直和分解的全部条件. □

引理 7.2.1 的证明实际给出了将任意 $v \in V[f]$ 按 $V[g] \oplus V[h]$ 分解的具体手段: 一旦 g, h 给定, 多项式的辗转相除法能够高效地判定它们是否互素, 并且在互素的情形给出 a, b 使得 $ag + bh = 1$. 取 $v = a(T)g(T)v + b(T)h(T)v$ 便是所求的分解.

今起要求 V 是取定的有限维 F-向量空间. 以下定义涉及理想的语言, 见 §6.1.

定义–命题 7.2.2 *设 $T \in \mathrm{End}(V)$. 考虑集合*

$$I := \{h \in F[X] : h(T) = 0\},$$

则其中存在唯一的首一多项式 $\mathrm{Min}_T \in F[X]$ 使得 $\deg \mathrm{Min}_T$ 极小, 而且 $I = (\mathrm{Min}_T)$ (符号如例 6.1.3). 如是之 Min_T 称为 T 的 **极小多项式**.

证明 根据 §5.8 开头的讨论, $I \neq \{0\}$. 若 $h_1, h_2 \in I$, 则 $(h_1+h_2)(T) = h_1(T)+h_2(T) = 0$; 若 $h \in I$ 而 $k \in F[X]$, 则 $(hk)(T) = h(T)k(T) = 0$. 综上可见 I 是 $F[X]$ 的理想, $I \neq \{0\}$. 代入 6.2.8 可见存在 $f \in F[X] \smallsetminus \{0\}$ 使得 $I = (f)$. 仅是精确到约定 6.2.1 的 $\sim$ 等价关系, f 才是唯一确定的; 然而注记 6.2.7 说明等价类有唯一的首一代表元. □

对于 $V = \{0\}$ 的极端情形, 我们有 $I = F[X]$, 此时 $\mathrm{Min}_T = 1$. 当 $V \neq \{0\}$ 时, 我们有 $T = 0_V$ 当且仅当 $\mathrm{Min}_T = X$.

命题 7.2.3 *任何 $T \in \mathrm{End}(V)$ 的极小多项式 Min_T 总是整除特征多项式 Char_T. 特别地, $\deg \mathrm{Min}_T \leq \dim V$.*

证明 根据 Min_T 的定义, 这等价于 Cayley–Hamilton 定理 5.8.8 断言的 $\mathrm{Char}_T(T) = 0_V$. □

下一则结果的证明用到一个简单观察: 设 $Tv=\lambda v$, 其中 $\lambda\in F$, 则对所有 $k\geq 0$ 都有 $T^kv=\lambda^kv$, 从而对所有 $f=\sum_k a_kX^k\in F[X]$ 都有

$$f(T)v=\sum_k a_kT^kv=\sum_k a_k\lambda^kv=f(\lambda)v.$$

命题 7.2.4 设 $\lambda\in F$, 则 λ 是 T 的特征值当且仅当 $\mathrm{Min}_T(\lambda)=0$. 作为推论, Min_T 和 Char_T 在 F 上有相同的根集 (不计重数).

证明 设 $v\in V\smallsetminus\{0\}$ 是以 λ 为特征值的特征向量, 则 $0=\mathrm{Min}_T(T)v=\mathrm{Min}_T(\lambda)v$ 导致 $\mathrm{Min}_T(\lambda)=0$.

反之设 λ 非特征值, 则 $T-\lambda\cdot\mathrm{id}_V$ 可逆; 因此 $X-\lambda$ 不能整除 Min_T, 否则对 $\mathrm{Min}_T(T)=0_V$ 两边作用 $(T-\lambda\cdot\mathrm{id}_V)^{-1}$ 将给出 $\frac{\mathrm{Min}_T}{X-\lambda}(T)=0_V$, 与极小性矛盾. □

留意到上述论证并不依赖 Cayley–Hamilton 定理 5.8.8.

对于矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$, 我们将 $\boldsymbol{A}$ 视作 $\mathrm{End}(F^n)$ 的元素来定义极小多项式 $\mathrm{Min}_{\boldsymbol{A}}$. 一如特征多项式的情形, 域的扩张不影响极小多项式.

命题 7.2.5 设 F 是域 E 的子域, $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$. 记 $\boldsymbol{A}$ 作为 F 上的矩阵的极小多项式为 $\mathrm{Min}_{\boldsymbol{A},F}$, 作为 E 上的矩阵的极小多项式为 $\mathrm{Min}_{\boldsymbol{A},E}$, 则 $\mathrm{Min}_{\boldsymbol{A},F}=\mathrm{Min}_{\boldsymbol{A},E}$.

证明 令 $d:=\deg\mathrm{Min}_{\boldsymbol{A},F}$. 显然 $\mathrm{Min}_{\boldsymbol{A},E}\mid\mathrm{Min}_{\boldsymbol{A},F}$; 由于极小多项式按定义是首一的, $\mathrm{Min}_{\boldsymbol{A},E}=\mathrm{Min}_{\boldsymbol{A},F}$ 等价于以下论断: 对所有 $a_0,\dots,a_{d-1}\in E$,

$$a_0\boldsymbol{A}^0+\cdots+a_{d-1}\boldsymbol{A}^{d-1}=\boldsymbol{0}_{n\times n}\iff a_0=\cdots=a_{d-1}=0.$$

然而左式是系数在 F 上的 d 元齐次线性方程组, 它有无非平凡解是由消元法确定的, 不受扩域影响. □

我们接着聚焦于极小多项式的因式分解.

引理 7.2.6 考虑 $T\in\mathrm{End}(V)$. 设有直和分解 $V=V_1\oplus\cdots\oplus V_k$ 使得每个 V_i 都是 T-不变子空间 (定义 5.10.1). 对每个 $1\leq i\leq k$ 记 $T_i:=T|_{V_i}\in\mathrm{End}(V_i)$, 则 Min_T 是 $\mathrm{Min}_{T_1},\dots,\mathrm{Min}_{T_k}$ 的最小公倍式.

证明 设 $f\in F[X]$. 由于 T-不变子空间 $V_1,\dots,V_k$ 生成 V, 我们有 $f(T)=0_V$ 当且仅当 $f(T)|_{V_i}=f(T_i)$ 对每个 i 皆等于 0_{V_i}. 然而根据极小多项式的性质, 后者又等价于 $\mathrm{Min}_{T_i}\mid f$. 综上可见 $\mathrm{Min}_T\mid f$ 当且仅当 f 是 $\mathrm{Min}_{T_1},\dots,\mathrm{Min}_{T_k}$ 的公倍式. □

例 7.2.7 设 $T\in\mathrm{End}(V)$ 在 F 上可对角化, 其特征值记为 $\lambda_1,\dots,\lambda_m\in F$ (不计重数), 则

$$\mathrm{Min}_T=\prod_{i=1}^m(X-\lambda_i).$$

验证是容易的: 定理 7.1.7 说明 V 是特征子空间的直和 $V=\bigoplus_{i=1}^m V_{\lambda_i}$. 既然 T 限制在非零子空间 V_{λ_i} 上等于 $\lambda_i\cdot \mathrm{id}_{V_{\lambda_i}}$, 相应的极小多项式不外乎 $\mathrm{Min}_{T_i}=X-\lambda_i$. 基于引理 7.2.6, 它们的最小公倍式 $\prod_{i=1}^m (X-\lambda_i)$ 即 Min_T.

一般而言, Min_T 在 $F[X]$ 中分解为乘积

$$\mathrm{Min}_T = p_1^{e_1}\cdots p_m^{e_m},$$

其中 p_i 是两两相异的首一不可约多项式, $e_i\in\mathbb{Z}_{\geq 1}$, 而且此分解精确到重排是唯一的. 此处同样有套话: 当 $m=0$ 时右式理解为常数 1, 对应于 $V=\{0\}$ 的极端情形.

对此分解反复运用引理 7.2.1, 便将 V 分解为 T-不变子空间的直和

$$V = V[p_1^{e_1}]\oplus\cdots\oplus V[p_m^{e_m}]. \tag{7.2.2}$$

引理 7.2.8 *考虑 $T\in\mathrm{End}(V)$. 设 $\mathrm{Min}_T=gh$, 其中 $g,h\in F[X]$ 互素. 记 $T':=T|_{V[g]}$, $T'':=T|_{V[h]}$, 则*

$$\mathrm{Min}_{T'}=g,\quad \mathrm{Min}_{T''}=h.$$

作为上式的特例, 考虑 (7.2.2) 的分解, 记 T_i 为 T 在不变子空间 $V[p_i^{e_i}]$ 上的限制, 则对所有 $1\leq i\leq m$ 皆有 $\mathrm{Min}_{T_i}=p_i^{e_i}$.

证明 根据 $V[g]$ 的定义, $g(T')=g(T)|_{V[g]}=0_{V[g]}$. 同理有 $g(T'')=0_{V[h]}$, 所以极小多项式的性质导致 $\mathrm{Min}_{T'}\mid g$ 和 $\mathrm{Min}_{T''}\mid h$. 另一方面, $\mathrm{Min}_T=gh$ 又是 $\mathrm{Min}_{T'}$ 和 $\mathrm{Min}_{T''}$ 的最小公倍式. 唯一可能是 $\mathrm{Min}_{T'}=g$ 而 $\mathrm{Min}_{T''}=h$.

对 $\mathrm{Min}_T=p_1^{e_1}\cdots p_m^{e_m}$ 反复运用此分解, 便得到 $\mathrm{Min}_{T_i}=p_i^{e_i}$ 恒成立. □

现在得到可对角化的第二种刻画, 它借由极小多项式表述.

定理 7.2.9 *线性映射 $T\in\mathrm{End}(V)$ 在 F 上可对角化的充要条件是极小多项式 $\mathrm{Min}_T\in F[X]$ 分裂而且无重根. 当上述条件成立时,*

$$\mathrm{Min}_T=\prod_{i=1}^m (X-\lambda_i),$$

其中 $\lambda_1,\dots,\lambda_m$ 遍历 T 的特征值, 不计重数.

证明 以下排除 $V=\{0\}$ 的极端情形. 当 T 可对角化时, 例 7.2.7 已经说明 $\mathrm{Min}_T=\prod_{i=1}^m(X-\lambda_i)$, 它分裂无重根.

反之, 设 $\mathrm{Min}_T=\prod_{i=1}^m(X-\lambda_i)$, 其中 $\lambda_1,\dots,\lambda_m\in F$ 两两相异. 此时 (7.2.2) 化为

$$V = V[X-\lambda_1]\oplus\cdots\oplus V[X-\lambda_m].$$

然而 $V[X-\lambda_i]=V_{\lambda_i}$, 而且 T 再无其他的非零特征子空间, 这就回到可对角化的判准, 即定理 7.1.7 (iii). □

推论 7.2.10 设 $T \in \mathrm{End}(V)$ 可对角化, 而 $V_0 \subset V$ 是 T-不变子空间, 则 $T|_{V_0} \in \mathrm{End}(V_0)$ 和推论 4.12.8 给出的 $\overline{T} \in \mathrm{End}(V/V_0)$ 皆可对角化.

证明 基于定理 7.2.9, 说 T 可对角化相当于说 Min_T 分裂无重根. 容易看出 $\mathrm{Min}_T(T|_{V_0}) = 0_{V_0}$ 而 $\mathrm{Min}_T(\overline{T}) = 0_{V/V_0}$, 所以 $T|_{V_0}$ 和 $\overline{T}$ 的极小多项式都整除 Min_T, 从而分裂无重根. □

推论 7.2.10 的逆并不成立. 以 $V = F^2$ 为例, 取 $V_0 = F\boldsymbol{e}_1$ 和表作矩阵 $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ 的 T, 则 $T|_{V_0}$ 和 $\overline{T}$ 都是一维空间上的恒等映射, 对应到两个对角元 1, 然而读者可以验证 $x \neq 0$ 时 T 不可对角化.

7.3 上三角化

一般的矩阵未必能对角化. 退而求其次, 我们寻求将矩阵通过共轭化为上三角的形式. 这是一个简单得多的问题: 以下将说明只要特征多项式在域 F 上分裂, 上三角化总是可行的. 作为特例, 在 F 代数闭的前提下 (定义 6.6.5), 所有线性映射都能上三角化. 相关推导不依赖 §7.2 的结果.

我们首先给出明确的定义. 以下的 V 都是有限维 F-向量空间, $n := \dim V$.

定义 7.3.1 (旗) 向量空间 V 的旗意谓 V 的一列子空间

$$\{0\} = V_0 \subsetneq \cdots \subsetneq V_m = V,$$

称 m 为此旗的长度. 若 $m = \dim V$, 则称之为完备旗.

对于给定的旗, 若线性映射 $T \in \mathrm{End}(V)$ 对所有 $0 \leq i \leq m$ 皆满足 $T(V_i) \subset V_i$, 则称 T 保持此旗.

因为 $0 = \dim V_0 < \cdots < \dim V_m = n$, 完备旗必然对所有 i 满足 $\dim V_i = i$.

例 7.3.2 取 V 的有序基 $v_1, \ldots, v_n$, 由之可以构造完备旗

$$V_0 := \{0\}, \quad V_i := \langle v_1, \ldots, v_i \rangle, \quad 1 \leq i \leq n.$$

反过来说, 对于任意完备旗 $V_0 \subsetneq \cdots \subsetneq V_n$, 因为 $\dim V_i = i$, 从 V_i 过渡到 V_{i+1} 只要多添一个生成元. 我们可以从 $i = 1$ 起步, 逐步选取 $v_1, \ldots, v_n$ (精确到一个比例常数) 使得 $V_i = \langle v_1, \ldots, v_i \rangle$. 这使 $v_1, \ldots, v_n$ 成为 V 的基.

定义 7.3.3 (上三角化–线性映射版本) 若 $T \in \mathrm{End}(V)$ 保持 V 的一组完备旗, 则称 T 在 F 上可以上三角化.

定义 7.3.4 (上三角化–矩阵版本) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 若存在可逆矩阵 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是上三角的, 则称 $\boldsymbol{A}$ 在 F 上可上三角化.

两套概念相互等价: T 可上三角化当且仅当它在某个有序基下的矩阵 $\boldsymbol{A}$ 可上三角化. 这是 §4.11 结尾解释过的事实.

定理 7.3.5 线性映射 $T \in \mathrm{End}(V)$ 在 F 上可上三角化的充要条件是特征多项式 $\mathrm{Char}_T \in F[X]$ 分裂.

证明 首先设 T 在 F 上可上三角化. 取 V 的有序基, 使得 T 对应的矩阵形如

$$\begin{pmatrix} \lambda_1 & * & * \\ & \ddots & \vdots \\ & & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \in F,$$

留白部分照例默认为 0. 按照分块来计算特征多项式 (命题 5.8.5), 可得 $\mathrm{Char}_T = \prod_{i=1}^n (X - \lambda_i)$.

反之设 $\mathrm{Char}_T = \prod_{i=1}^n (X - \lambda_i)$, 其中 $\lambda_1, \dots, \lambda_n \in F$. 以下对 $n = \dim V$ 递归地论证 T 在 F 上可上三角化. 当 $n = 0$ 时无事可作. 以下设 $n \geq 1$. 既然 Char_T 分裂, 总是存在特征值 $\lambda_1 \in F$ 和对应的特征向量 $v_1 \neq 0$. 定义 1 维子空间 $V_1 := \langle v_1 \rangle$, 另外记 $\overline{V} := V/V_1$, 则因为 $T(V_1) \subset V_1$, 推论 4.12.8 (代入 $W = V$ 和 $U = U' = V_1$) 给出 $\overline{T} \in \mathrm{End}(\overline{V})$. 代入命题 5.10.2 可得

$$\mathrm{Char}_{\overline{T}} = \frac{\mathrm{Char}_T}{X - \lambda_1} = \prod_{i=2}^n (X - \lambda_i).$$

由于 $\dim \overline{V} = n - 1$, 我们知道 $\overline{T}$ 可上三角化: 它保持某个完备旗

$$\{0\} = \overline{V}_1 \subsetneq \cdots \subsetneq \overline{V}_n = \overline{V}.$$

以命题 4.12.9 取它们相对于商映射 $q : V \to \overline{V}$ 的原像 $V_1 \subsetneq \ldots \subsetneq V_n$; 因为 $\{0\}$ 的原像确实是 V_1, 记法是合理的. 由此得到 V 的完备旗

$$\{0\} =: V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V.$$

以上的旗也被 T 保持, 缘由如下: $v \in V_i$ 蕴涵 $qT(v) = \overline{T}q(v) \in \overline{T}(\overline{V_i}) \subset \overline{V_i}$, 从而 $Tv \in V_i$. 这就说明 T 可上三角化. □

以后介绍 Jordan 标准形时将对上三角化给出更精密的版本. 然而上三角化定理 7.3.5 很容易证明, 应用中又往往已经足够方便, 以下是一些范例.

推论 7.3.6 若特征多项式 Char_T 分裂为 $\prod_{i=1}^n (X - \lambda_i)$, 则对于任意 $m \in \mathbb{Z}_{\geq 0}$, 我们有 $\mathrm{Char}_{T^m} = \prod_{i=1}^n (X - \lambda_i^m)$.

证明 以定理 7.3.5 化到 T 对应于上三角矩阵的情形, 以命题 4.11.6 描述 T^m 的对角元, 然后按命题 5.8.5 比较 T 和 T^m 的特征多项式. □

注记 7.3.7 上三角化的论证技巧给出 Cayley–Hamilton 定理 5.8.8 的第二种证明. 且先假设 Char_T 已在 F 上分解为 $\prod_{i=1}^{n}(X-\lambda_i)$, 以解释证明的思路.

论证同样基于递归, 不妨设 $n \geq 1$. 回忆定理 7.3.5 的证明, 其第一步是取特征值为 λ_1 的特征向量 v_1 以及 $V_1 := \langle v_1 \rangle$, 然后考虑商空间 $\overline{V}$ 层次的线性映射 $\overline{T}$. 递归地假定

$$\prod_{i=2}^{n}\left(\overline{T}-\lambda_i \cdot \mathrm{id}_{\overline{V}}\right) = \mathrm{Char}_{\overline{T}}(\overline{T}) = 0_{\overline{V}}.$$

拉回 V 上观照, 这相当于说对每个 $v \in V$ 皆有

$$\prod_{i=2}^{n}\left(T-\lambda_i \mathrm{id}_V\right)(v) \ \in V_1.$$

两边同取 $T-\lambda_1 \cdot \mathrm{id}_V$ 的像便给出 $\mathrm{Char}_T(T)(v)=0$. 明所欲证.

对于一般的域, 考虑 Cayley–Hamilton 定理的矩阵表述. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 关键在于将 F 嵌入足够大的域 E, 以确保 $\mathrm{Char}_{\boldsymbol{A}}$ 在 E 上分裂; 推论 6.10.5 确保这般的域嵌入总是存在. 如果 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$ 在 $\mathrm{M}_{n\times n}(E)$ 中成立, 则在 $\mathrm{M}_{n\times n}(F)$ 中同样成立. 明所欲证.

练习 7.3.8 设 V 为 n 维 $\mathbb{C}$-向量空间, $T \in \mathrm{End}(V)$. 通过纯量限制将 V 视同 $2n$ 维向量空间, 并以 $\det_{\mathbb{R}} T$ 代表 T 作为 $\mathbb{R}$-线性映射的行列式. 试证 $\det_{\mathbb{R}} T = |\det T|^2$.

提示〉 以 $\mathbb{C}$ 上的三角化定理 7.3.5 化到 $n=1$ 情形具体计算.

7.4 广义特征子空间

本节继续选定有限维向量空间 V. 一般的线性映射 $T \in \mathrm{End}(V)$ 未必能对角化, 但是在假设 Char_T 分裂的前提下, 总能够将 V 分解为广义特征子空间的直和.

定义 7.4.1 设 $T \in \mathrm{End}(V)$ 而 $\lambda \in F$. 沿用 (7.2.1) 的记法, 命

$$V_{[\lambda],N} := V\left[(X-\lambda)^N\right], \quad N \in \mathbb{Z}_{\geq 0}$$
$$V_{[\lambda]} := \bigcup_{N\geq 1} V_{[\lambda],N}.$$

我们称 $V_{[\lambda]}$ 是 T 相对于 λ 的**广义特征子空间**.

我们有 $V_{\lambda,[0]} = \{0\}$ 而 $V_\lambda = V_{[\lambda],1} \subset V_{[\lambda],2} \subset \cdots$. 关于 $V_{[\lambda]}$ 还需要一些说明. 尽管子空间的并通常不再是子空间, 但此处的递增性质确保 $V_{[\lambda]}$ 确实是子空间; 举加法封闭性为例, 若 $v, v' \in V_{[\lambda]}$, 可取 N 和 N' 使得 $v \in V_{[\lambda],N}$ 而 $v' \in V_{[\lambda],N'}$, 于是 $v+v' \in V_{[\lambda],\max\{N,N'\}} \subset V_{[\lambda]}$. 纯量乘法封闭性则是容易的.

引理 7.4.2 我们有 $V_{[\lambda]} \neq \{0\}$ 当且仅当 λ 是 T 的特征值.

证明 一个方向是容易的: 若 λ 是特征值, 则 $\{0\} \neq V_\lambda \subset V_{[\lambda]}$.

对于另一个方向, 取 $v \in V_{[\lambda]} \smallsetminus \{0\}$, 再取最小的 $N \geq 1$ 使得 $(T - \lambda \cdot \mathrm{id}_V)^N v = 0$, 则 $(T - \lambda \cdot \mathrm{id}_V)^{N-1}(v)$ 是 V_λ 的非零元. □

留意到定义 7.4.1 和引理 7.4.2 并不要求 V 有限维, 但本节后续的论证只适用于有限维的 V. 命 $n := \dim V$.

以下设 Char_T 分裂, 具体表作

$$\mathrm{Char}_T = \prod_{i=1}^{m} (X - \lambda_i)^{a_i},$$

其中 $\lambda_1, \ldots, \lambda_m$ 两两相异, $a_i \in \mathbb{Z}_{\geq 1}$ 满足 $a_1 + \cdots + a_m = n$. 鉴于命题 7.2.4, 极小多项式 Min_T 也有相同的根集, 而 $\mathrm{Min}_T \mid \mathrm{Char}_T$, 故

$$\mathrm{Min}_T = \prod_{i=1}^{m} (X - \lambda_i)^{b_i},$$
$$1 \leq b_i \leq a_i.$$

引理 7.4.3 设 Char_T 分裂并且定义 λ_i, b_i 如上. 对每个 $1 \leq i \leq m$ 皆有

$$V_{[\lambda_i]} = V\left[(X - \lambda_i)^{b_i}\right] = V_{[\lambda_i], \dim V}.$$

证明 先处理第一个等号. 定义显然导致 $V\left[(X - \lambda_i)^{b_i}\right] \subset V_{[\lambda_i]}$. 对于另一方向的包含关系, 设 $v \in V_{[\lambda_i]}$, 再取 $N \geq 1$ 使得

$$(T - \lambda_i \cdot \mathrm{id}_V)^N (v) = 0.$$

若 $h \in F[X]$ 是 $(X - \lambda_i)^N$ 和 Min_T 的最大公因式, 则因为 λ_i 是 Min_T 的 b_i 重根, h 整除 $(X - \lambda_i)^{b_i}$; 另一方面, 存在 $f, g \in F[X]$ 使得

$$h = (X - \lambda_i)^N f + \mathrm{Min}_T\, g,$$
$$h(T)v = f(T)(T - \lambda_i \cdot \mathrm{id}_V)^N v = 0.$$

综上遂有 $(T - \lambda_i \cdot \mathrm{id}_V)^{b_i} v = 0$.

考虑第二个等号. 由于 $b_i \leq a_i \leq n$, 故 $V_{[\lambda_i], b_i} \subset V_{[\lambda_i], n} \subset V_{[\lambda_i]}$. 配合上一步的结果可知每个包含都是等号. □

由此得到适用于 V 在 T 作用下的**广义特征子空间分解**:

$$V \overset{(7.2.2)}{=\!=} \bigoplus_{i=1}^{m} V\left[(X - \lambda_i)^{b_i}\right] = \bigoplus_{i=1}^{m} V_{[\lambda_i]}. \tag{7.4.1}$$

记 $T_i := T|_{V_{[\lambda_i]}}$, 则

$$\begin{gathered}\mathrm{Min}_{T_i} \mid (X-\lambda_i)^{b_i},\\ \mathrm{Char}_{T_i} \text{ 和 } \mathrm{Min}_{T_i} \text{ 有相同的根集 (命题 7.2.4)},\\ \prod_{i=1}^{m}\mathrm{Char}_{T_i}=\mathrm{Char}_T=\prod_{i=1}^{m}(X-\lambda_i)^{a_i}.\end{gathered}$$

既然 $\lambda_1,\dots,\lambda_m$ 相异, 综上立见

$$\begin{aligned}\mathrm{Char}_{T_i}&=(X-\lambda_i)^{a_i},\\ \dim V_{[\lambda_i]}&=\deg\mathrm{Char}_{T_i}=a_i.\end{aligned}\tag{7.4.2}$$

定义 7.4.4 设 Char_T 分裂, $\lambda\in F$. 定义 λ 的**几何重数**为 $\dim V_\lambda$, 其**代数重数**为 λ 作为 Char_T 的根的重数.

鉴于引理 7.4.2, 无论几何重数或代数重数, 它们非零的充要条件都是 λ 为 T 的特征值.

定理 7.4.5 任意 λ 的代数重数皆大于或等于几何重数. 两者相等当且仅当 $V_{[\lambda]}=V_\lambda$.

证明 可设 λ 是特征值, 否则所论的重数和空间全为零. 鉴于 (7.4.2), 代数重数等于 $\dim V_{[\lambda]}$, 显然大于等于几何重数 $\dim V_\lambda$, 而且等号成立的充要条件是 $V_{[\lambda]}=V_\lambda$. □

目光转回对角化的问题. 命题 7.1.5 表明特征多项式分裂是可对角化的必要条件, 因此以下关于可对角化的第三种判准也只论分裂情形.

推论 7.4.6 考虑线性映射 $T\in\mathrm{End}(V)$. 假设 Char_T 分裂, 则 T 可对角化当且仅当每个特征值的代数重数皆等于几何重数.

证明 可对角化的充要条件是 $V=\bigoplus_{i=1}^m V_{\lambda_i}$. 将此与 V 已有的分解 (7.4.1) 比较, 可见 T 可对角化当且仅当 $V_{\lambda_i}=V_{[\lambda_i]}$ 恒成立. 将此代入定理 7.4.5. □

关乎广义特征子空间的一个标准结论是矩阵或线性映射的 **Jordan–Chevalley 分解**, 有两种版本. 先引进一则常用概念.

定义 7.4.7 选定有限维 F-向量空间 V (或 $n\in\mathbb{Z}_{\geq 1}$). 设 $N\in\mathrm{End}(V)$ (或 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$). 若存在 $k\in\mathbb{Z}_{\geq 1}$ 使得 $N^k=0_V$ (或 $\boldsymbol{A}^k=\boldsymbol{0}_{n\times n}$), 则称 N (或 $\boldsymbol{A}$) 为**幂零**的.

定理 7.4.8 (加性 Jordan–Chevalley 分解) 考虑 $T\in\mathrm{End}(V)$ 并假设 Char_T 分裂. 存在唯一一对 $S,N\in\mathrm{End}(V)$ 使得

- S 可对角化,
- N 幂零,
- $T=S+N$ 而 $SN=NS$.

此外, 存在 $f,g\in F[X]$ 使得 $S=f(T)$ 而 $N=g(T)$. 当 T 可逆时 S 亦可逆.

证明 先取相对于 T 的广义特征子空间分解 (7.4.1), 记 $T_i:=T|_{V_{[\lambda_i]}}$. 命 $n=\dim V$.

关于存在性, 定义 S 使得它在每个 $V_{[\lambda_i]}$ 上限制为 $\lambda_i\mathrm{id}$. 其次定义 N 使得它在每个 $V_{[\lambda_i]}$ 上限制为 $T_i-\lambda_i\mathrm{id}$. 从 $V_{[\lambda_i]}=V_{[\lambda_i],n}$ 可见 N^n 限制在每个 $V_{[\lambda_i]}$ 上皆为零映射, 从而 $N^n=0_V$.

极易在每个 $V_{[\lambda_i]}$ 上检验 $T=S+N$ 和 $SN=NS$, 这就给出所求的资料. 观察到当 T 可逆时 $\lambda_i\in F^\times$, 故以上构造的 S 也可逆.

至于唯一性, 先记 S 的相异特征值为 $\mu_1,\dots,\mu_l$, 相应地分解 V 为特征子空间直和 $\bigoplus_{j=1}^l V_j$. 对所有 j, 从 $SN=NS$ 得出 V_j 是 N-不变子空间, 而 $(T_j-\mu_j\mathrm{id})^k=N^k|_{V_j}=0_{V_j}$ 蕴涵 T_j 在 V_j 上的唯一特征值为 μ_j. 这就说明

- $l=m$ 而 $\mu_j=\lambda_j$ 在重排下标后对所有 j 成立,
- 承上, $V_j\subset V_{[\lambda_j]}$ 对所有 j 成立.

与广义特征子空间分解对照, 立见 $V_j=V_{[\lambda_j]}$. 因此 S 由 $S|_{V_{[\lambda_i]}}=\lambda_i\mathrm{id}$ 唯一确定, 而 $N=T-S$.

最后讨论 f 的取法. 对主理想环 $F[X]$ 应用中国剩余定理 6.3.8, 得到 f 使得对所有 $1\leq i\leq m$ 皆有

$$f\equiv\lambda_i\pmod{(X-\lambda_i)^n};$$

故 $f(T)|_{V_{[\lambda_i]}}=\lambda_i\mathrm{id}$, 回顾唯一性论证得到 $f(T)=S$. 取 $g=X-f$. □

接着从加性版本推导乘性版本.

定理 7.4.9 (乘性 Jordan–Chevalley 分解) 设 $T\in\mathrm{End}(V)$ 可逆而 Char_T 分裂. 存在唯一一对可逆之 $S,U\in\mathrm{End}(V)$ 使得

- S 可对角化,
- $U-\mathrm{id}_V$ 幂零,
- $SU=T=US$.

此外, 存在 $f,g\in F[X]$ 使得 $S=f(T)$ 而 $U=g(T)$.

证明 对于存在性, 先取加性 Jordan–Chevalley 分解 $T=S+N$, 其中 S 已知可逆. 命 $U:=\mathrm{id}_V+S^{-1}N$, 则 $SU=T$. 此外从 $SN=NS$ 推得 $U-\mathrm{id}_V=S^{-1}N$ 幂零, 而且 $SU=US$.

对于唯一性, 将分解写作 $T=S+S(U-\mathrm{id}_V)$. 命 $N=S(U-\mathrm{id}_V)$, 则从 $SU=US$ 和 $U-\mathrm{id}_V$ 幂零推出 $SN=NS$ 以及 N 幂零. 于是加性版本说明 S 由 T 唯一确定, U 随之也唯一确定.

最后, 加性版本给出 $f \in F[X]$ 使得 $S = f(T)$. 兹断言 f 与 Char_T 互素: 设若不然, 则存在 T 的特征值 λ 使得 $X - \lambda \mid f$, 从而相应的特征向量 v 满足 $Sv = 0$, 矛盾. 由断言知存在 $g, h \in F[X]$ 使得 $gf + h\,\mathrm{Char}_T = X$, 故 $g(T)f(T) = T$, $g(T) = TS^{-1} = U$. 明所欲证. □

7.5 同步对角化

本节设 V 是有限维 F-向量空间. 前几节已经从种种面向说明如何判定 $T \in \mathrm{End}(V)$ 能否对角化. 然而许多实际场景中面对的不是单个线性映射, 而是一族 $T_1, \dots, T_k \in \mathrm{End}(V)$, 甚至是无穷多个线性映射. 这就引出以下定义.

定义 7.5.1 设 $\mathcal{S}$ 是 $\mathrm{End}(V)$ 的子集. 如果存在 V 的基 $v_1, \dots, v_n$ 使得每个 v_i 都是所有 $T \in \mathcal{S}$ 共同的特征向量, 则称 $\mathcal{S}$ 在 F 上可以**同步对角化**.

用矩阵语言来说, 子集 $\mathcal{S} \subset \mathrm{M}_{n\times n}(F)$ 可以同步对角化相当于说存在可逆之 $\boldsymbol{P}$ 使得矩阵 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 对所有 $\boldsymbol{A} \in \mathcal{S}$ 都是对角的. 定义 7.1.3 相当于 $\mathcal{S}$ 是独点集 $\{T\}$ 的特例.

即使每个 $T \in \mathcal{S}$ 都可以对角化, $\mathcal{S}$ 也未必能同步对角化, 这是因为定义要求 $\mathcal{S}$ 有一族共同的特征向量作为 V 的基; 或者以矩阵语言改述, 定义要求可逆矩阵 $\boldsymbol{P}$ 的选取不依赖于 $\boldsymbol{A} \in \mathcal{S}$. 关键条件是乘法交换性.

引理 7.5.2 设 $S, T \in \mathrm{End}(V)$ 满足 $ST = TS$, 则对于所有 $\lambda \in F$, 相对于 T 的特征子空间 $V_\lambda = \{v \in V : Tv = \lambda v\}$ 都是 S-不变子空间.

证明 设 $v \in V_\lambda$, 则

$$TSv = STv = S(\lambda v) = \lambda Sv,$$

因此 $Sv \in V_\lambda$. □

定理 7.5.3 设 $\mathcal{S}$ 是 $\mathrm{End}(V)$ 的子集, 则 $\mathcal{S}$ 在 F 上可以同步对角化的充要条件是以下两则性质成立.

- ⋆ 每个 $T \in \mathcal{S}$ 在 F 上皆可对角化;
- ⋆ 对所有 $T, T' \in \mathcal{S}$ 皆有 $TT' = T'T$.

证明 首先说明条件的必要性. 设 $v_1, \dots, v_n$ 是 V 的基, 满足

$$Tv_i = \lambda_i(T)v_i, \quad T \in \mathcal{S}, \quad 1 \le i \le n,$$

其中 $\lambda_i(T) \in F$ 是对应的特征值, 依赖于 T, 则对任意 $T, T' \in \mathcal{S}$ 和所有 i 皆有

$$\begin{aligned} TT'v_i &= T\left(\lambda_i(T')v_i\right) = \lambda_i(T)\lambda_i(T')v_i, \\ T'Tv_i &= T'\left(\lambda_i(T)v_i\right) = \lambda_i(T')\lambda_i(T)v_i, \end{aligned}$$

由此立见 $TT'=T'T$.

接着说明充分性. 取 $\mathcal{S}$ 作为 $\mathrm{End}(V)$ 的子集的极大线性无关子集 $\{T_1,\dots,T_k\}$. 由于 $\mathcal{S}$ 的所有元素都能写成线性组合 $\sum_{i=1}^k a_iT_k$, 问题简化到 $\mathcal{S}$ 是有限集 $\{T_1,\dots,T_k\}$ 的情形. 以下对 k 递归地论证.

当 $k=1$ 时, T_1 本身按条件已经可对角化.

设 $k\geqslant 2$. 对 T_1 应用定理 7.1.7 可得相对于 T_1 的特征子空间分解 $V=\bigoplus_\lambda V_\lambda$, 下标 $\lambda\in F$ 遍历 T_1 的特征值. 引理 7.5.2 蕴涵 V_λ 对所有 i 都是 T_i-不变子空间. 现在对每个 $\lambda\in F$ 定义

$$T_i^\lambda:=T_i|_{V_\lambda}\in\mathrm{End}(V_\lambda),\quad 1\leqslant i\leqslant k.$$

一旦 $T_1^\lambda,\dots,T_k^\lambda$ 对所有 λ 都可以同步对角化, 则取直和可见 $T_1,\dots,T_k$ 亦然. 所以同步对角化的问题可以化约到每个 V_λ 上来处理. 接着选定 λ 并观察到

- ⋆ 每个 T_i^λ 皆可对角化, 这是基于推论 7.2.10;
- ⋆ $T_i^\lambda T_j^\lambda=T_j^\lambda T_i^\lambda$ 恒成立, 这是将等式 $T_iT_j=T_jT_i$ 限制到 V_λ 上的产物.

然而 $T_1^\lambda=\lambda\cdot\mathrm{id}_{V_\lambda}$, 所以 V_λ 上的同步对角化问题实则只涉及 $T_2^\lambda,\dots,T_k^\lambda$. 这就化约到 $k-1$ 的递归假设. □

习题

1. 设 V 是有限维 F-向量空间, $T\in\mathrm{End}(V)$. 证明若 V 的每个非零元都是 T 的特征向量, 则存在 $\lambda\in F$ 使得 $T=\lambda\cdot\mathrm{id}_V$.

2. 设 $\boldsymbol{A}\in\mathrm{M}_{2\times 2}(\mathbb{C})$. 证明 $\boldsymbol{A}$ 或者可对角化, 或者共轭于形如 $\begin{pmatrix}a&1\\0&a\end{pmatrix}$ 的矩阵, 两种情况是互斥的.

3. 证明矩阵 $\boldsymbol{A}=\begin{pmatrix}2&0&0\\u&2&0\\v&w&-1\end{pmatrix}$ 可对角化的充要条件是 $u=0$.

4. 在复数域 $\mathbb{C}$ 上, 求下列矩阵 $\boldsymbol{A}$ 的所有特征值和特征向量, 并且判断哪些矩阵可以对角化. 在可对角化的情形, 确定可逆矩阵 $\boldsymbol{P}$ 和对角矩阵 $\boldsymbol{D}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}$.

$$\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix},\quad\begin{pmatrix}2&3&2\\1&8&2\\-2&-14&-3\end{pmatrix},\quad\begin{pmatrix}6&2&4\\2&3&2\\4&2&6\end{pmatrix}.$$

5. 满足 $\boldsymbol{A}^2=\boldsymbol{A}$ 的 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 称为**幂等矩阵**. 证明幂等矩阵总能对角化, 并且其特征值都属于 $\{0,1\}$.

提示〉 任何 $\boldsymbol{v}\in F^n$ 皆可表为 $\boldsymbol{A}\boldsymbol{v}+(\boldsymbol{v}-\boldsymbol{A}\boldsymbol{v})$. 或者应用关于极小多项式的知识.

6. 设 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$. 证明若 $\boldsymbol{AB}=\boldsymbol{BA}$, 而且 $\boldsymbol{A}$ 在 F 上有 n 个相异特征值, 则存在多项式 $f\in F[X]$ 使得 $\boldsymbol{B}=f(\boldsymbol{A})$.

7. 对所有列向量 $\boldsymbol{v},\boldsymbol{w}\in F^n$ 确定 $n\times n$ 矩阵 $\boldsymbol{v}\cdot{}^{\mathrm{t}}\boldsymbol{w}$ 的特征多项式, 讨论它可否对角化. 提示〉 回忆 $\mathrm{Char}_{\boldsymbol{AB}}$ 和 $\mathrm{Char}_{\boldsymbol{BA}}$ 的比较.

8. 对以下复矩阵, 确定它们的特征多项式, 并且对每个特征值确定几何重数和代数重数, 从而判断它们是否可对角化.

$$\begin{pmatrix}4&-3&0\\2&-1&0\\1&-1&1\end{pmatrix},\quad \begin{pmatrix}1&1&0\\0&1&0\\0&0&2\end{pmatrix}.$$

9. 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 满足 $\boldsymbol{A}^N=\mathbf{1}_{n\times n}$, 其中 N 是正整数, 说明 $\boldsymbol{A}$ 可对角化. 另一方面, 举出 $\boldsymbol{A}\in \mathrm{M}_{2\times 2}(\mathbb{F}_2)$ 的例子, 使得 $\boldsymbol{A}^2=\mathbf{1}_{2\times 2}$ 而 $\boldsymbol{A}$ 在 $\mathbb{F}_2$ 的任何扩域上都无法对角化.

10. 设域 F 满足 $\mathrm{char}(F)=0$. 记由域 F 中的所有数列 $(a_0,a_1,\ldots)=(a_n)_{n=0}^{\infty}$ 构成的 F-向量空间为 $\mathcal{S}$, 其加法和纯量乘法逐项地定义. 对所有 $x\in F$ 和 $k\in\mathbb{Z}_{\geq 0}$, 定义 $\mathcal{S}$ 的元素如下.

- ⋆ 若 $x\neq 0$, 定义
$$ {}^{k}\mathbf{x}:=\left(n^k x^n\right)_{n=0}^{\infty},\quad (0^0:=1);$$
- ⋆ 若 $x=0$, 定义 ${}^{k}\mathbf{x}$ 使得它的第 k 项是 1, 其余为 0.

(i) 定义线性映射 $T:\mathcal{S}\to\mathcal{S}$, 使之映 $(a_n)_{n=0}^{\infty}$ 为 $(a_{n+1})_{n=0}^{\infty}$. 说明
$$(T-x\cdot\mathrm{id}_{\mathcal{S}})^{k+1}\left({}^{k}\mathbf{x}\right)=0,\quad (T-x\cdot\mathrm{id}_{\mathcal{S}})^{k}\left({}^{k}\mathbf{x}\right)\neq 0.$$

(ii) 证明 ${}^{0}\mathbf{x},{}^{1}\mathbf{x},{}^{2}\mathbf{x},\ldots$ 在 $\mathcal{S}$ 中线性无关.

(iii) 运用上述结果来确定满足
$$a_{n+d}=c_{d-1}a_{n+d-1}+\cdots+c_0a_n,\quad n\in\mathbb{Z}_{\geq 0}$$
的线性递归数列 $(a_n)_{n=0}^{\infty}\in\mathcal{S}$ 的通解, 其中 $d\geq 1$ 和 $c_0,\ldots,c_{d-1}\in F$ 是给定的.
提示〉 令 $f=X^d-c_{d-1}X^{d-1}-\cdots-c_0$, 问题相当于求 $f(T)\in\mathrm{End}(\mathcal{S})$ 的核. 设 x 是 f 的根, 重数为 $e(x)$, 验证 ${}^{0}\mathbf{x},\ldots{}^{e(x)-1}\mathbf{x}$ 都落在核里, 然后说明当根 x 变动, 得到 d 个线性无关的解.

11. 设 V 为有限维 $\mathbb{R}$-向量空间, $V\neq\{0\}$, 而 $T\in\mathrm{End}(V)$. 证明或者 T 有特征向量, 或者存在 2 维的 T-不变子空间. 提示〉 用极小多项式.

12. 给定有限维向量空间 V 和 $T\in\mathrm{End}(V)$, 考虑从 $\mathrm{End}(V)$ 到其自身的线性映射
$$\begin{aligned}L_T:\mathrm{End}(V)&\to\mathrm{End}(V)\\ S&\mapsto TS.\end{aligned}$$

(i) 证明 L_T 和 T 有相同的极小多项式.

(ii) 分别用 $\det T$ 和 $\mathrm{Tr}(T)$ 来描述 $\det L_T$ 和 $\mathrm{Tr}(L_T)$.

提示〉 从矩阵观点看, $\mathrm{End}(V) \simeq \mathrm{M}_{n\times n}(F)$ 分解为列空间的直和.

13. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 说明若 $\mathrm{Tr}(\boldsymbol{A}) = 0$, 而且 $\mathrm{char}(F) = 0$, 则存在 $\boldsymbol{A}' = (a'_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$ 使得

$$\boldsymbol{A} \text{ 共轭于 } \boldsymbol{A}', \quad a'_{11} = \cdots = a'_{nn} = 0.$$

提示〉 可设 $n \geq 2$. 取 $\boldsymbol{v} \in F^n \smallsetminus \{0\}$ 使得 $\boldsymbol{v}$ 和 $\boldsymbol{A}\boldsymbol{v}$ 不成比例, 将 $\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}$ 扩充为 F^n 的有序基 $\mathbf{v}$, 考虑 $\boldsymbol{A}$ 作为线性映射 $F^n \to F^n$ 对 $\mathbf{v}$ 的矩阵表法, 然后递归地化约到 $n-1$ 情形.

14. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 幂零 (定义 7.4.7), 证明

(i) $\boldsymbol{A}$ 的唯一特征值是 0,

(ii) $\mathrm{Char}_{\boldsymbol{A}} = X^n$,

(iii) $\boldsymbol{A}$ 可对角化当且仅当 $\boldsymbol{A} = \mathbf{0}_{n\times n}$.

15. 将定理 7.3.5 的充分条件部分强化为以下版本: 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 满足 $\mathrm{Char}_{\boldsymbol{A}} = \prod_{i=1}^{n}(X - \lambda_i)$, 证明存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角元依序为 $\lambda_1, \ldots, \lambda_n$ 的上三角矩阵.

16. 证明 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 幂零当且仅当存在可逆的 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角线全为 0 的上三角矩阵.

17. 设 $\boldsymbol{A} = \left(\begin{smallmatrix} a & \\ & b \end{smallmatrix}\right) \in \mathrm{M}_{2\times 2}(\mathbb{R})$, 其中 $a \neq b$ 而且 $a < 0$. 证明不存在 $\boldsymbol{B} \in \mathrm{M}_{2\times 2}(\mathbb{R})$ 使得 $\boldsymbol{B}^2 = \boldsymbol{A}$. 提示〉 一种策略是应用推论 7.3.6.

18. 回忆到 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 的经典伴随矩阵记为 $\boldsymbol{A}^\vee$ (定义 5.7.3). 证明若 $\mathrm{Char}_{\boldsymbol{A}} = \prod_{i=1}^{n}(X - \lambda_i)$, 则 $\mathrm{Char}_{\boldsymbol{A}^\vee} = \prod_{i=1}^{n}(X - \Lambda_i)$, 其中

$$\Lambda_i := \lambda_1 \cdots \widehat{\lambda_i} \cdots \lambda_n,$$

符号 $\widehat{\cdots}$ 代表在连乘积中省略该项.

提示〉 一种方法是先说明问题只和 $\boldsymbol{A}$ 的共轭类相关, 从而化到 $\boldsymbol{A}$ 是上三角矩阵的情形, 然后计算 $\boldsymbol{A}^\vee$ 的对角元.

19. 设 $n \geq 2$. 定义 $n \times n$ 矩阵

$$\boldsymbol{A} := \begin{pmatrix} 0 & 0 & 0 & \cdots & & 0 \\ 1 & 0 & 0 & \cdots & & 0 \\ 0 & \ddots & & & & \\ \vdots & & \ddots & & & \vdots \\ 0 & \cdots & & 1 & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}$$

换言之它在 "次对角线" 上为 1, 其余为 0. 试说明 $\boldsymbol{A}$ 不可能共轭于以下形式的分块对角矩阵

$$\left(\begin{array}{c|c} \boldsymbol{C} & \mathbf{0}_{c\times d} \\ \hline \mathbf{0}_{d\times c} & \boldsymbol{D} \end{array}\right), \quad \boldsymbol{C} \in \mathrm{M}_{c\times c}(F), \quad \boldsymbol{D} \in \mathrm{M}_{d\times d}(F)$$

其中 $c, d \geq 1$, $c + d = n$.

20. 设 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$, $\boldsymbol{B} \in \mathrm{M}_{n\times m}(F)$. 证明 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 有相同的非零特征值.

提示〉 应用命题 5.8.6, 或者作以下观察: 设 $\boldsymbol{v}$ 是 $\boldsymbol{AB}$ 的特征值为 $\lambda \neq 0$ 的特征向量, 则 $\boldsymbol{ABv} = \lambda\boldsymbol{v}$ 蕴涵 $\boldsymbol{Bv} \neq \boldsymbol{0}$ 和 $\boldsymbol{BA}(\boldsymbol{Bv}) = \lambda\boldsymbol{Bv}$.

21. 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$. 证明如果它们的特征多项式皆分裂, 而且 $\boldsymbol{AB} = \boldsymbol{BA}$, 则它们可以同步上三角化: 存在可逆之 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 和 $\boldsymbol{P}^{-1}\boldsymbol{BP}$ 皆为上三角的.

条件 $\boldsymbol{AB} = \boldsymbol{BA}$ 对于同步上三角化是充分而非必要的. 精确的充要条件略为复杂, 适合从 Lie 理论来观照, 在此不论.

22. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为代数闭域 F 上的 $n \times n$ 矩阵, $\boldsymbol{AB} = \boldsymbol{BA}$ 而 $\boldsymbol{B}^n = \boldsymbol{0}_{n\times n}$. 证明 $\boldsymbol{A}$ 和 $\boldsymbol{A} + \boldsymbol{B}$ 有相同的特征多项式.

提示〉 将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 同步上三角化.

23. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 其中 F 是任意域. 请在不用 Cayley–Hamilton 定理 5.8.8 的前提下证明以下陈述.

(i) 若 $\boldsymbol{A}$ 在 F 的某个扩域上可对角化, 则 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$.

(ii) 若 $\mathrm{Char}_{\boldsymbol{A}}$ 的判别式 (练习 6.7.9) 非零, 则 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$.

提示〉 应用 (i) 和推论 7.1.8.

(iii) 由此说明 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$ 总是成立. 这给出 Cayley–Hamilton 定理的另一证明.

提示〉 所求等式可以表述为 $f_{ij} = 0$, 其中 $1 \leq i, j \leq n$ 而 f_{ij} 是关于 $\boldsymbol{A}$ 的矩阵元的多项式. 当 F 无穷时应用 (ii) 与代数等式的延拓原理 (定理 3.6.3) 来证明. 若 F 有限, 则取无穷扩域, 例如 F 上的有理函数域.

第八章 双线性形式

设 V 和 W 为域 F 上的向量空间, 如果映射 $B : V \times W \to F$ 单独对每个变元都是线性的, 则称之为双线性形式; 将 F 替换成一般的向量空间 X, 便得到双线性映射的概念 (定义 8.1.1). 双线性形式的实例有几何与分析学中的内积, 理论力学中的辛形式, 概率与数据科学中的协方差等, 不一而足. 当 $V = F^m$ 而 $W = F^n$ 时, 双线性形式 $B : V \times W \to F$ 也能用矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 表达为

$$B(\boldsymbol{v}, \boldsymbol{w}) = {}^{\mathrm{t}}\boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{w}.$$

本章的 §§8.1–8.3 表述双线性映射与双线性形式的基本概念与实例, 并且在有限维的情形给出矩阵描述. 其中格外重要的是非退化双线性形式的概念 (定义 8.2.1), 等价于对应的矩阵可逆. 给定有限维向量空间 V_i, V_i' 连同双线性形式 $B : V_i \times V_i' \to F$, 其中 $i = 1, 2$. 在 B_1 非退化的前提下, 定义–命题 8.2.8 将

- ⋆ 对线性映射 $T : V_1 \to V_2$ 定义其右伴随 $T^* : V_2' \to V_1'$ 使得

$$B_2(Tv_1, v_2') = B_1\left(v_1, T^* v_2'\right) \quad \text{恒成立,}$$

- ⋆ 对线性映射 $T : V_1' \to V_2'$ 定义其左伴随 ${}^*T : V_2 \to V_1$ 使得

$$B_2(v_2, Tv_1') = B_1\left({}^*T v_2, v_1'\right) \quad \text{恒成立.}$$

这些跷跷板似的伴随关系是代数学中的常见思路. 最根本的是 $V_1 = V_2$, $V_1' = V_2'$ 而 $B_1 = B_2$ 或者对称 (亦即 $B_i(v, w) = B_i(w, v)$) 或者反对称 (亦即 $B_i(v, w) = -B_i(w, v)$) 的情形, 此时伴随不必再分左右, 由此可得自伴或反自伴线性映射的概念 (定义 8.2.15). 一切都能用矩阵具体地表达.

我们关心形如 $B : V \times V \to F$ 的双线性形式, 简记为 (V, B), 其中 V 是有限维的; 一如我们迄今见识的所有代数结构, 这些资料之间也有同构的概念. 给定 (V_1, B_1) 和 (V_2, B_2), 若有向量空间的同构 $\varphi : V_1 \xrightarrow{\sim} V_2$ 使得 $B_2(\varphi(v_1), \varphi(v_1')) = B_1(v_1, v_1')$ 对所有 $v_1, v_1' \in V_1$ 成立, 则称 (V_1, B_1) 和 (V_2, B_2) 同构, 它们本质上是同一个结构. 分类所有 (V, B) 的同构类是研究双线性形式的基本目标. 用矩阵的语言来说, 这相当于研究 $\mathrm{M}_{n\times n}(F)$ 对等价关系

$$\boldsymbol{A} \sim {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}, \quad \boldsymbol{C} \in \mathrm{M}_{n\times n}(F) \text{ 可逆}$$

的商集; 上述等价关系也称为矩阵的合同关系 (定义 8.3.8).

分类问题在 B 对称和反对称的情形差异鲜明, 此外问题还和 $\mathrm{char}(F)$ 相关. 本书在这部分总假设 $\mathrm{char}(F) \neq 2$.

对称情形由 §§8.4–8.6 处理. 记 $n = \dim V$, 关于资料 (V, B) 的分类也可以等价地表述为 n 元齐次二次多项式 (简称 n 元二次型)

$$f = \sum_{1 \leq i,j \leq n} a_{ij} X_i X_j, \quad a_{ij} = a_{ji}$$

在线性坐标变换下的分类问题; 二次型的分类在代数闭域 (例如 $\mathbb{C}$) 上近乎平凡, 在 $\mathbb{R}$ 上则由 Sylvester 惯性定理 8.6.6 给出完整的答案. 尽管问题乍看是二次的, 依然落在线性工具的射程之内.

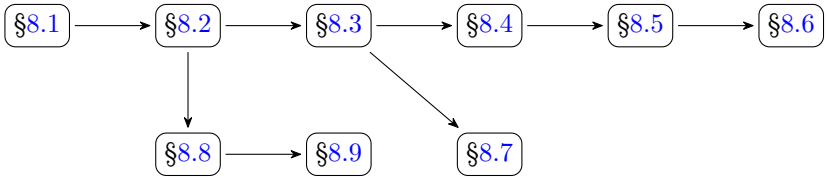
反对称情形由 §8.7 处理. 非退化且反对称的 (V, B) 称为辛空间, 其分类在所有域 F 上都是一致的: 同构类由 $\dim V$ 完全确定, 而且 $\dim V$ 必为偶数 (推论 8.7.5). 证明诀窍是研究 V 的 Lagrange 子空间, 关键结论是定理 8.7.4.

最后的 §8.8 和 §8.9 分别研究向量空间 V 的双重对偶 $V^{\vee\vee}$ 和商空间的对偶, 需要本章前几节的若干结论. 我们将证明线性映射 T 的单性及满性, 以及核 $\ker(T)$ 及余核 $\mathrm{coker}(T)$ 都是相互 “对偶” 的概念, 像 $\mathrm{im}(T)$ 则是 “自对偶” 的, 对此将有严格的数学解释. 基于这些抽象结论, 推论 8.9.6 将重新证明矩阵的列秩等于行秩, 并将等式升级为一个同构.

阅读提示

若无另外说明, F 代表一个选定的域. 在 §§8.4–8.7 默认 $\mathrm{char}(F) \neq 2$.

阅读顺序



8.1 双线性形式

双线性形式是双线性映射的特例. 我们先从后者的定义入手.

定义 8.1.1 (双线性映射和双线性形式) 设 V,W,X 为 F-向量空间, 若映射

$$B : V \times W \to X$$

对每个变元都是线性的, 亦即

$$\begin{aligned} B(v_1+v_2,w) &= B(v_1,w)+B(v_2,w),\\ B(v,w_1+w_2) &= B(v,w_1)+B(v,w_2),\\ B(tv,w) = tB(v,w) &= B(v,tw), \quad t\in F, \end{aligned}$$

则称 B 为双线性映射. 这些双线性映射构成的集合记为 $\mathrm{Bil}(V,W;X)$

取 $X=F$ 的特例, 则双线性映射 $B:V\times W\to F$ 也称为 $V\times W$ 上的双线性形式, 或称双线性型.

一如线性映射的情形, 集合 $\mathrm{Bil}(V,W;X)$ 也对逐点运算

$$\begin{aligned} (B_1+B_2)(v,w) &:= B_1(v,w)+B_2(v,w),\\ (tB)(v,w) &:= t\cdot B(v,w) \end{aligned}$$

成为向量空间, 以零映射 $B(\cdot,\cdot)=0$ 为零元; 一切所需性质都可以化到单个变元来检验.

今后的重点在于双线性形式的情形. 双线性形式 $V\times W\to F$ 有时也形象地称为 V 和 W 的**配对**.

例 8.1.2 矩阵乘法 $\mathrm{M}_{m\times n}(F)\times \mathrm{M}_{n\times l}(F)\to \mathrm{M}_{m\times l}(F)$ 是双线性映射.

例 8.1.3 取特例 $F=\mathbb{R}$. 在直角坐标系下, 平面向量的内积 $(x_1,x_2)\cdot(y_1,y_2) := \sum_{i=1}^{2} x_iy_i$ 给出 $\mathbb{R}^2\times\mathbb{R}^2$ 上的双线性形式. 类似地, 空间向量的内积 $(x_1,x_2,x_3)\cdot(y_1,y_2,y_3) := \sum_{i=1}^{3} x_iy_i$ 给出 $\mathbb{R}^3\times\mathbb{R}^3$ 上的双线性形式. 实向量空间上的内积结构是第九章的主题.

例 8.1.4 (典范配对) 取特例 $W=V^\vee$ (回忆定义 4.7.5), 定义 $V^\vee$ 和 V 之间的典范配对为

$$\begin{aligned} \langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_V : V^\vee\times V &\to F\\ (\lambda,v) &\mapsto \langle\lambda,v\rangle := \lambda(v). \end{aligned}$$

对偶空间 $V^\vee$ 的定义使得 $\langle\cdot,\cdot\rangle$ 对两个变元都是线性的. 线性映射 $T:V\to W$ 的转置 ${}^{\mathrm{t}}T : W^\vee\to V^\vee$ 的定义 4.7.6 依此改写为

$$\langle {}^{\mathrm{t}}T(\lambda),v\rangle = \lambda(Tv) = \langle\lambda,Tv\rangle\,,$$

其中 $v\in V$ 而 $\lambda\in W^\vee$.

按照 §4.7 介绍的方式, 将 F^n 视同列向量空间 $\mathrm{M}_{n\times 1}(F)$, 将 $(F^n)^\vee$ 视同行向量空间 $\mathrm{M}_{1\times n}(F)$, 典范配对 $\langle\cdot,\cdot\rangle$ 便化为矩阵的乘法

$$\begin{aligned}\mathrm{M}_{1\times n}(F)\times \mathrm{M}_{n\times 1}(F) &\longrightarrow \mathrm{M}_{1\times 1}(F)=F\\ ({}^{\mathrm{t}}\boldsymbol{w},\boldsymbol{v}) &\longmapsto {}^{\mathrm{t}}\boldsymbol{w}\cdot\boldsymbol{v},\end{aligned}$$

而例 8.1.4 关于转置映射的写法则以矩阵语言表为

$$\langle \boldsymbol{w},\boldsymbol{A}\boldsymbol{v}\rangle = {}^{\mathrm{t}}\boldsymbol{w}\cdot\boldsymbol{A}\boldsymbol{v} = {}^{\mathrm{t}}({}^{\mathrm{t}}\boldsymbol{A}\cdot\boldsymbol{w})\cdot\boldsymbol{v} = \langle {}^{\mathrm{t}}\boldsymbol{A}\cdot\boldsymbol{w},\boldsymbol{v}\rangle,$$

其中 $\boldsymbol{v}\in F^n$, $\boldsymbol{w}\in F^m$ 而 $\boldsymbol{A}\in \mathrm{M}_{m\times n}(F)$.

现在留意到对于任意集合 $\mathcal{A},\mathcal{B},\mathcal{C}$, 记 $\mathcal{C}^{\mathcal{A}} := \{\text{所有映射 } \mathcal{A}\to\mathcal{C}\}$, 则有自然的双射[1)]

$$\begin{aligned}\left(\mathcal{C}^{\mathcal{A}}\right)^{\mathcal{B}} &\xrightarrow{\text{1:1}} \mathcal{C}^{\mathcal{A}\times\mathcal{B}} \xleftarrow{\text{1:1}} \left(\mathcal{C}^{\mathcal{B}}\right)^{\mathcal{A}}\\ \varphi &\longmapsto [(a,b)\mapsto \varphi(b)(a)]\\ &\qquad [(a,b)\mapsto \psi(a)(b)] \longmapsto \psi.\end{aligned}$$

若将集合代换成向量空间, 类似现象在线性的框架下同样成立. 为了简化讨论, 且将以上的集合 $\mathcal{C}$ 代换为 1 维向量空间 F.

给定双线性形式 $B: V\times W\to F$. 当 $v\in V$ 固定, $B(v,\cdot): W\to F$ 确定 $W^\vee$ 的元素, 记为 $\psi(v)$; 同理, 当 $w\in W$ 固定, $B(\cdot,w)$ 确定 $V^\vee$ 的元素, 记为 $\varphi(w)$. 通过这一观察, 所有双线性形式都可以通过典范配对来表达, 细说如下.

命题 8.1.5 *对于任意 F-向量空间 V 和 W, 我们有向量空间的同构*

$$\begin{aligned}\mathrm{Hom}(W,V^\vee) &\xrightarrow{\sim} \mathrm{Bil}(V,W;F) \xleftarrow{\sim} \mathrm{Hom}(V,W^\vee)\\ [w\mapsto B(\cdot,w)] &\longleftarrow B \longmapsto [v\mapsto B(v,\cdot)]\\ \varphi &\longmapsto [B(v,w):=\langle\varphi(w),v\rangle]\\ &\qquad [B(v,w):=\langle\psi(v),w\rangle] \longmapsto \psi\end{aligned}$$

证明 首先, 注意到图中所有映射皆线性. 问题在于逐一验证两个方向来回合成的产物都是恒等.

先论 $\mathrm{Hom}(W,V^\vee)\leftrightarrow\mathrm{Bil}(V,W;F)$. 给定 $\varphi\in\mathrm{Hom}(W,V^\vee)$, 它向右映为 $B(v,w)=\langle\varphi(w),v\rangle$, 而 B 再向左映为 $\mathrm{Hom}(W,V^\vee)$ 的如下元素: $w\mapsto B(\cdot,w)=\langle\varphi(w),\cdot\rangle=\varphi(w)$. 因此 $\reflectbox{$\rightarrow$}$ 合成为 id.

另一方面, 给定双线性形式 B, 它向左映为 $\varphi: w\mapsto B(\cdot,w)$, 后者再向右映为双线性形式 $(v,w)\mapsto\langle\varphi(w),v\rangle$, 然而 $\langle\varphi(w),v\rangle=\varphi(w)(v)=B(v,w)$, 因此 $\rightarrow$ 也合成为 id.

至于 $\mathrm{Bil}(V,W;F)\leftrightarrow\mathrm{Hom}(V,W^\vee)$, 论证完全是对称的, 不必重复. □

[1)] 这种技巧在计算机科学或数理逻辑中称为 Curry 化.

如果考虑更一般的 $\mathrm{Bil}(V,W;X)$, 则命题 8.1.5 中的 $V^\vee$ (或 $W^\vee$) 应替换为

$$\mathrm{Hom}(V,X) \quad (\text{或 } \mathrm{Hom}(W,X));$$

论证无异, 只是符号更冗长.

命题 8.1.6 (以矩阵表达双线性形式) 设 $m,n\in\mathbb{Z}_{\geq 1}$. 将 F^n 的元素视同列向量, 则有向量空间的同构

$$\begin{aligned}\mathrm{M}_{m\times n}(F) &\xrightarrow{\sim} \mathrm{Bil}(F^m,F^n;F)\\ \boldsymbol{A} &\longmapsto [B(\boldsymbol{v},\boldsymbol{w}) := {}^{\mathrm{t}}\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{w}].\end{aligned}$$

进一步, 对所有 $(x_i)_i\in F^m$ 和 $(y_i)_i\in F^n$ 皆有

$$B\left(\sum_{i=1}^m x_i\boldsymbol{e}_i,\sum_{j=1}^n y_j\boldsymbol{e}_j\right)=\sum_{i=1}^m\sum_{j=1}^n a_{ij}x_iy_j.$$

证明 断言中的同构有至少两种论证. 先介绍略微抽象的第一种方法. 命题 8.1.5 的同构 $\mathrm{Hom}(W,V^\vee)\xrightarrow{\sim}\mathrm{Bil}(V,W;F)$ 中代入 $V=F^m$ 和 $W=F^n$. 另外将行向量空间 $(F^m)^\vee$ 和列向量空间 F^m 通过转置等同, 那么 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$ 可以等同于线性映射 $\varphi:F^n\to(F^m)^\vee$, 映 $\boldsymbol{w}$ 为 $\varphi(\boldsymbol{w})={}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{w})$. 将典范配对表达成行向量对列向量的乘法, φ 对应的双线性形式 B 遂表为

$$\begin{aligned}B(\boldsymbol{v},\boldsymbol{w})&=\langle\varphi(\boldsymbol{w}),\boldsymbol{v}\rangle={}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{w})\cdot\boldsymbol{v}\\&={}^{\mathrm{t}}\boldsymbol{w}\cdot{}^{\mathrm{t}}\boldsymbol{A}\cdot\boldsymbol{v}\\&={}^{\mathrm{t}}\boldsymbol{v}\boldsymbol{A}\boldsymbol{w};\end{aligned}$$

最后一个等式成立的缘由是将两边作为 1×1 矩阵取转置.

接着验证 $B\left(\sum_{i=1}^m x_i\boldsymbol{e}_i,\sum_{j=1}^n y_j\boldsymbol{e}_j\right)$ 的公式. 双线性导致

$$B\left(\sum_i x_i\boldsymbol{e}_i,\sum_j y_j\boldsymbol{e}_j\right)=\sum_i\sum_j x_iy_jB(\boldsymbol{e}_i,\boldsymbol{e}_j),$$

问题化为证 $B(\boldsymbol{e}_i,\boldsymbol{e}_j)=a_{ij}$. 回忆到 $\boldsymbol{A}\boldsymbol{e}_j$ 是列向量 $\sum_{k=1}^m a_{kj}\boldsymbol{e}_k$. 因此 ${}^{\mathrm{t}}\boldsymbol{e}_i\cdot(\boldsymbol{A}\boldsymbol{e}_j)$ 取出此列向量的第 i 个分量 a_{ij}. 此即 $B(\boldsymbol{e}_i,\boldsymbol{e}_j)=a_{ij}$.

至于第二种方法, 扼要地说, $\mathrm{M}_{m\times n}(F)\to\mathrm{Bil}(F^m,F^n;F)$ 显然是线性映射, 姑且记为 Θ. 由于双线性导致 B 由资料 $(B(\boldsymbol{e}_i,\boldsymbol{e}_j))_{i,j}$ 唯一确定, 故映射 Θ 的右侧维数 $\leq mn$, 而左侧已知是 mn 维的. 然而上一段的结论 $B(\boldsymbol{e}_i,\boldsymbol{e}_j)=a_{ij}$ 表明 $\boldsymbol{A}$ 被 $B=\Theta(\boldsymbol{A})$ 确定, 故 Θ 单. 比较维数遂知 Θ 是同构. □

注记 8.1.7 若取 $n=m$, 并且设双线性形式 B 对应于 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$, 则

$$B(\boldsymbol{w},\boldsymbol{v})={}^{\mathrm{t}}\boldsymbol{w}\cdot\boldsymbol{A}\cdot\boldsymbol{v}={}^{\mathrm{t}}({}^{\mathrm{t}}\boldsymbol{w}\cdot\boldsymbol{A}\boldsymbol{v})={}^{\mathrm{t}}\boldsymbol{v}\cdot{}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{w},$$

所以在命题 8.1.6 的对应下, 对调双线性形式的变元相当于将对应的矩阵作转置.

定义 8.1.8 设 $B_1 : V_1 \times V_1' \to F$ 和 $B_2 : V_2 \times V_2' \to F$ 为双线性形式. 定义

$$B : (V_1 \oplus V_2) \times (V_1' \oplus V_2') \to F$$
$$B((v_1, v_2), (v_1', v_2')) = B_1(v_1, v_1') + B_2(v_2, v_2').$$

易见 B 也是双线性形式, 称为 B_1 和 B_2 的**直和**, 记为 $B_1 \oplus B_2$.

具体矩阵的角度看, 设 $V_i = F^{n_i}$, $V_i' = F^{n_i'}$ 而 B_i 对应到矩阵 $\boldsymbol{A}_i \in \mathrm{M}_{n_i \times n_i'}(F)$, 其中 $i = 1, 2$, 则直和 $B_1 \oplus B_2$ 对应于分块对角矩阵

$$\boldsymbol{A} := \left(\begin{array}{c|c} \boldsymbol{A}_1 & \\ \hline & \boldsymbol{A}_2 \end{array}\right) \in \mathrm{M}_{(n_1+n_2)\times(n_1'+n_2')}(F).$$

接着探讨两个变元取在同一个空间的双线性形式.

定义 8.1.9 设 V 为向量空间, $B : V \times V \to F$ 为双线性形式.

- ⋆ 若对所有 $v_1, v_2 \in V$ 皆有 $B(v_1, v_2) = B(v_2, v_1)$, 则称 B 为**对称**的;
- ⋆ 若对所有 $v_1, v_2 \in V$ 皆有 $B(v_1, v_2) = -B(v_2, v_1)$, 则称 B 为**反对称**的.

如果 B_1 和 B_2 皆对称 (或反对称), 则其直和亦然.
双线性形式的对称性或反对称性在矩阵语言下有直接了当的描述.

定义 8.1.10 若矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 满足 ${}^{\mathrm{t}}\boldsymbol{A} = \boldsymbol{A}$ (或 ${}^{\mathrm{t}}\boldsymbol{A} = -\boldsymbol{A}$), 则称 $\boldsymbol{A}$ 为对称 (或反对称) 矩阵.

引理 8.1.11 设双线性形式 $B : F^n \times F^n \to F$ 按命题 8.1.6 的方式对应到 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 则 B 是对称 (或反对称) 的当且仅当 $\boldsymbol{A}$ 是对称 (或反对称) 矩阵.

证明 注记 8.1.7 的直接结论. □

对称和反对称这两种双线性形式将是我们未来探讨的重点.

定义 8.1.12 给定双线性形式 $B : V \times W \to F$, 可以定义任何子空间 $W_0 \subset W$ 的正交空间为

$${}^{\perp}W_0 := \{v \in V : \forall w_0 \in W_0,\ B(v, w_0) = 0\}.$$

类似地, 对任何子空间 $V_0 \subset V$ 可以定义

$$V_0^{\perp} := \{w \in W : \forall v_0 \in V_0,\ B(v_0, w) = 0\}.$$

由双线性形式的性质可见 ${}^{\perp}W_0$ 确实是 V 的子空间, 而 $V_0^{\perp}$ 是 W 的子空间. 此外, 近乎同义反复的论证表明

$$W_0 \subset ({}^{\perp}W_0)^{\perp}, \quad V_0 \subset {}^{\perp}(V_0^{\perp}).$$

对于一般的双线性形式 B, 上述包含关系可以是严格的; 以极端情形 $B=0$ 为例, 所有 V_0 都有相同的正交空间 $V_0^{\perp}=W$. 我们将在 §8.2 对非退化双线性形式来考察正交空间所承载的信息.

注记 8.1.13 在 $V=W$ 的场合, 对子空间 V_0 取 ${}^{\perp}V_0$ 和 $V_0^{\perp}$ 的产物未必相同. 在本书稍后所考虑的情境中, B 总是对称或反对称的, 这时左右两种正交空间相同, 不妨统一标为 $V_0^{\perp}$.

8.2 非退化形式与伴随映射

继续关于双线性形式的讨论.

定义 8.2.1 *考虑任意向量空间 V, W 和双线性形式 $B\in \mathrm{Bil}(V,W;F)$.*

- ⋆ *定义 B 的*__左根__*为 V 的子空间 ${}^{\perp}W=\{v\in V : B(v,\cdot)=0\}$.*
- ⋆ *定义 B 的*__右根__*为 W 的子空间 $V^{\perp}=\{w\in W : B(\cdot,w)=0\}$.*
- ⋆ *设 V 和 W 都是有限维的. 若 B 的左根和右根都是零空间, 则称 B* __非退化__*.*

假如 $V=W$ 而 B 是定义 8.1.9 中的对称或反对称双线性形式, 则 B 的左根自动等于右根.

例 8.2.2 设 V 有限维. 例 8.1.4 定义的典范配对 $\langle\cdot,\cdot\rangle : V^{\vee}\times V\to F$ 是非退化的.

- ⋆ 首先考虑它的左根: 若 $\lambda\in V^{\vee}$ 满足 $\langle\lambda,v\rangle=\lambda(v)=0$ 对所有 $v\in V$ 成立, 则按定义 $\lambda=0$.
- ⋆ 其次考虑右根. 设 $v\in V$ 非零. 将 $v=v_1$ 扩充为基 $v_1,\dots,v_n$, 则其对偶基 $\check{v}_1,\dots,\check{v}_n$ 中的 $\check{v}_1$ 满足 $\langle\check{v}_1,v\rangle=1$, 故 v 不属于右根. 换言之, 右根也是零空间.

引理 8.2.3 *设 $B\in\mathrm{Bil}(V,W;F)$ 按命题 8.1.5 的同构对应到 $\psi\in\mathrm{Hom}(V,W^{\vee})$ 和 $\varphi\in\mathrm{Hom}(W,V^{\vee})$, 则*

$$B\text{ 的左根 }=\ker(\psi),\quad B\text{ 的右根 }=\ker(\varphi).$$

证明 这是定义和命题 8.1.5 的具体构造的立即结论. □

作一则简单的观察: 若 V 和 W 皆为有限维向量空间, 则仅当 $\dim V=\dim W$ 时才可能存在非退化双线性形式, 解释如下. 设两者维数不同, 不妨就假设 $\dim V>\dim W=\dim(W^{\vee})$, 则必有 $\dim\ker(\psi)=\dim V-\mathrm{rk}(\psi)>0$, 此时 B 退化.

基于上述事实, 今后探讨有限维向量空间上的非退化双线性形式 $B : V\times W\to F$ 时, 一律默认 $\dim V=\dim W$.

命题 8.2.4 设 V 和 W 是有限维 F-向量空间, $\dim V = \dim W$, 则对于任意 $B \in \mathrm{Bil}(V,W;F)$, 以下性质相互等价:

(i) B 非退化,

(ii) B 的左根为 $\{0\}$,

(iii) B 的右根为 $\{0\}$,

当以上任一条件成立时, B 对应的 $\varphi : W \to V^\vee$ 和 $\psi : V \to W^\vee$ 都是同构.

证明 按非退化双线性型的定义立见 (i) 蕴涵 (ii) 和 (iii). 以下说明 (ii) $\Longrightarrow$ (i). 设 (ii) 成立, 则 $\psi : V \to W^\vee$ 是单射; 由于 $\dim V = \dim W = \dim W^\vee$, 它自动是同构. 我们的目的是说明 B 的右根也是 $\{0\}$. 设 $w \in W$ 属于右根, 亦即对所有 $v \in V$ 皆有 $B(v,w) = 0$, 然而这也相当于说 $\psi(v)$ 总是属于 $W^\vee$ 的子空间

$$ {}^{\perp}\langle w\rangle := \{\lambda \in W^\vee : \lambda(w) = 0\}. $$

假若 $w \neq 0$, 则 ${}^{\perp}\langle w\rangle \subsetneq W^\vee$, 这是因为 w 可扩充为 W 的基 $w_1, \dots, w_n$ 使得 $w_1 = w$; 取其对偶基 $\check{w}_1, \dots, \check{w}_n \in W^\vee$, 则 $\check{w}_1 \notin \langle w\rangle^{\perp}$. 于是 $\mathrm{im}(\psi)$ 包含于 $W^\vee$ 的真子空间, 与 ψ 为同构矛盾. 综上, B 非退化.

(iii) $\Longrightarrow$ (i) 的论证是完全相同的, 此时同样有 $\varphi : W \xrightarrow{\sim} V^\vee$. □

例 8.2.5 取 $V = W = F^n$, 等同于列向量空间, 并将其对偶空间等同于行向量空间. 对于 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 依命题 8.1.6 的范式来考虑双线性形式

$$ \begin{aligned} B : F^n \times F^n &\longrightarrow F \\ (\boldsymbol{v}, \boldsymbol{w}) &\longmapsto {}^{\mathrm{t}}\boldsymbol{v}\,\boldsymbol{A}\,\boldsymbol{w}. \end{aligned} $$

它对应的线性映射 $\varphi : F^n \to (F^n)^\vee$ 已知是 $\boldsymbol{w} \mapsto {}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{w})$, 映列向量为行向量. 因此 B 非退化当且仅当 $\boldsymbol{A}$ 可逆.

例 8.2.6 (迹形式) 设 V 为有限维 F-向量空间, 则 $\mathrm{End}(V)$ 也是有限维的. 运用 §5.9 介绍的迹映射, 可以在 $\mathrm{End}(V)$ 上定义自然的双线性形式

$$ \begin{aligned} \mathrm{End}(V) \times \mathrm{End}(V) &\longrightarrow F \\ (S,T) &\longmapsto \mathrm{Tr}(ST), \end{aligned} $$

称为迹形式. 因为 $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$ 恒成立, 这是定义 8.1.9 所谓的对称双线性形式. 现在说明它还是非退化的. 为此, 仅须对所有 S 证

$$ \forall T \in \mathrm{End}(V),\ \mathrm{Tr}(ST) = 0 \implies S = 0. $$

取 V 的基以将 $\mathrm{End}(V)$ 等同于矩阵环 $\mathrm{M}_{n\times n}(F)$. 设 S 对应于 $\boldsymbol{A}=(a_{ij})_{i,j}\in\mathrm{M}_{n\times n}(F)$. 代入 $\mathrm{M}_{n\times n}(F)$ 的标准基的元素 $\boldsymbol{B}=\boldsymbol{E}_{ij}$ 可得

$$\boldsymbol{A}\boldsymbol{E}_{ij}=\begin{pmatrix} 0 & \cdots & a_{1i} & \cdots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \cdots & a_{ni} & \cdots & 0 \end{pmatrix}\begin{matrix}\\ \\ \leftarrow \text{第 } j \text{ 列}\\ \\ \\ \end{matrix},\quad \mathrm{Tr}(\boldsymbol{A}\boldsymbol{E}_{ij})=a_{ji}.$$

根据假设, $\mathrm{Tr}(\boldsymbol{A}\boldsymbol{E}_{ij})=0$; 既然其中的 (i,j) 可以任取, 由此立得 $\boldsymbol{A}=\boldsymbol{0}_{n\times n}$.

以下练习给出命题 8.2.4 的强化版本, 它涉及定义 4.12.2 介绍的商空间.

练习 8.2.7　设 V 和 W 是任意 F-向量空间, $B\in\mathrm{Bil}(V,W;F)$. 分别记 B 的左根和右根为 L 和 R.

(i) 证明 B 自然地诱导出非退化双线性映射 $\overline{B}:(V/L)\times(W/R)\to F$, 由下式给出

$$\overline{B}(v+L,w+R)=B(v,w).$$

(ii) 应用 (i) 来说明 $\dim L-\dim R=\dim V-\dim W$.

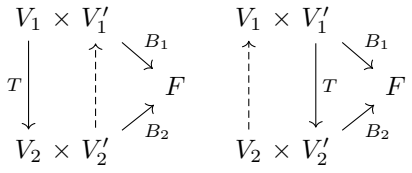
结合命题 8.1.5 和命题 8.2.4 可知, 对于有限维 F-向量空间 V 和 W,

$$\text{指定非退化双线性形式 } V\times W\xrightarrow{B}F \iff \text{指定同构 } \varphi:W\xrightarrow{\sim}V^\vee \iff \text{指定同构 } \psi:V\xrightarrow{\sim}W^\vee,$$

而 B 在这些同构之下对应到 V 或 W 上的典范配对. 命题 8.2.4 的刻画还表明非退化双线性形式 B 可以用来表示对偶空间的所有元素, 更精确地说:

- ⋆ 对于任意 $\lambda\in V^\vee$, 存在唯一的 $w\in W$ 使得 $\lambda=\varphi(w):=B(\cdot,w)$;
- ⋆ 对于任意 $\eta\in W^\vee$, 存在唯一的 $v\in V$ 使得 $\eta=\psi(v):=B(v,\cdot)$.

作为应用, 现在考虑一个貌似复杂 (实则不然) 的情境: 给定有限维向量空间上的两个双线性形式 B_1 和 B_2, 设 B_1 非退化. 对于以下两种情况



其中 T 是给定的线性映射, 我们都寻求虚线所示的线性映射, 它和 T 之间须具备某种翘板似的 “伴随关系”.

定义–命题 8.2.8 (伴随映射) 考虑双线性形式 $B_i \in \mathrm{Bil}(V_i, V_i'; F)$, 其中 $i = 1, 2$, 其中所有向量空间均为有限维, 并且假设 B_1 非退化.

(i) 存在唯一的线性映射

$$\begin{aligned} \mathrm{Hom}(V_1, V_2) &\longrightarrow \mathrm{Hom}(V_2', V_1') \\ T &\longmapsto T^*, \end{aligned}$$

其中的 T^* 称为 T 相对于 B_1 和 B_2 的**右伴随**, 由下式刻画:

$$B_2(Tv_1, v_2') = B_1(v_1, T^* v_2'), \quad v_1 \in V_1,\ v_2' \in V_2'.$$

(ii) 存在唯一的线性映射

$$\begin{aligned} \mathrm{Hom}(V_1', V_2') &\longrightarrow \mathrm{Hom}(V_2, V_1) \\ T &\longmapsto {}^*T, \end{aligned}$$

其中的 ${}^*T$ 称为 T 相对于 B_1 和 B_2 的**左伴随**, 由下式刻画:

$$B_2(v_2, Tv_1') = B_1({}^*Tv_2, v_1'), \quad v_1' \in V_1', \quad v_2 \in V_2.$$

对于 $V_1 = V_2$, $V_1' = V_2'$ 而 $B_1 = B_2 =: B$ 的情况, 称 T^* (或 ${}^*T$) 为 T 相对于 B 的右 (或左) 伴随.

证明 讨论 (i) 即足. 思路是清楚的: 固定 $v_2' \in V_2'$, 则映射

$$B_2(T(\cdot), v_2') : v_1 \mapsto B_2(Tv_1, v_2')$$

给出 $V_1^\vee$ 的元素; 因为 B_1 非退化, 先前的观察说明存在唯一的 $v_1' \in V_1'$ 使得 $B_2(T(\cdot), v_2') = B_1(\cdot, v_1')$. 所求的 T^* 的映法必然是 $v_2' \mapsto v_1'$, 然而须说明这般定义的 T^* 确实是线性映射, 而且 $T \mapsto T^*$ 也是线性的. 最后这两步没有本质上的困难, 请读者基于 v_1' 的刻画自行检验. □

当 B_1 和 B_2 都非退化时, 关于左伴随和右伴随的刻画即刻给出

$$({}^*T)^* = T = {}^*(T^*); \tag{8.2.1}$$

这几乎是同义反复.

例 8.2.9 在定义–命题 8.2.8 中取特例 $V_1 = V_2 =: V$, $V_1' = V_2' =: V'$ 和 $T = \mathrm{id}_V$, 可见对于

⋆ 给定的非退化双线性形式 $B_1 \in \mathrm{Bil}(V, V'; F)$,

⋆ 任意双线性形式 $B_2 \in \mathrm{Bil}(V, V'; F)$,

存在唯一的映射 $S \in \mathrm{End}(V')$ 使得 $B_2(v, v') = B_1(v, Sv')$ 对所有 $v, v' \in V$ 成立; 事实上 $S = \mathrm{id}^*$. 换言之, 任何 B_2 和给定的非退化双线性形式之间都仅相差一个线性映射.

如果给定 $B_i \in \mathrm{Bil}(V_i, V_i'; F)$ (其中 $i = 1, 2, 3$) 和线性映射 $V_1 \xrightarrow{T} V_2 \xrightarrow{S} V_3$, 并假设 B_1 和 B_2 非退化, 则

$$(ST)^* = T^*\, S^*.$$

这是因为

$$B_3\left(STv_1, v_3'\right) = B_2\left(Tv_1, S^* v_3'\right) = B_1\left(v_1, T^*\, S^* v_3'\right).$$

左伴随的情况当然是类似的. 在 B_1 和 B_2 皆非退化的前提下, 这顺带说明若 T 可逆, 则 T^* (或 ${}^*T$) 也可逆, 而

$$(T^*)^{-1} = \left(T^{-1}\right)^*, \quad ({}^*T)^{-1} = {}^*\left(T^{-1}\right). \tag{8.2.2}$$

伴随映射的行为显然和先前介绍的转置有些类似, 它们之间有简单的关系.

练习 8.2.10 设上述的双线性形式 B_i 按命题 8.1.5 对应到 $\varphi_i : V_i' \to V_i^\vee$, 其中 $i = 1, 2$, 而 φ_1 可逆. 验证 $\mathrm{Hom}(V_2', V_1^\vee)$ 中的等式 $\varphi_1 \circ T^* = {}^tT \circ \varphi_2$. 换言之, 应用约定 2.3.3 的交换图表, $T : V_1 \to V_2$ 的右伴随 T^* 和转置 ${}^tT$ 的关系可以总结为交换图表

$$\begin{array}{ccc} V_2' & \xrightarrow{T^*} & V_1' \\ {\scriptstyle\varphi_2}\downarrow & & \downarrow{\scriptstyle\varphi_1} \\ V_2^\vee & \xrightarrow[{}^tT]{} & V_1^\vee \end{array}.$$

以 ψ_i 代替 φ_i 给出左伴随的版本. 提示〉 以例 8.1.4 的典范配对来验证

$$\begin{aligned} B_1\left(v_1, T^* v_2'\right) &= \langle \varphi_1(T^* v_2'), v_1 \rangle_{V_1}, \\ B_2\left(Tv_1, v_2'\right) &= \langle \varphi_2(v_2'), Tv_1 \rangle_{V_2} = \langle {}^tT\left(\varphi_2(v_2')\right), v_1 \rangle_{V_1}. \end{aligned}$$

练习 8.2.11 基于上一道练习的结果, 证明当 B_1 和 B_2 皆非退化时 $\mathrm{rk}(T^*) = \mathrm{rk}(T) = \mathrm{rk}({}^*T)$.

命题 8.2.12 取定非退化双线性形式 $B_i : V_i \times V_i \to F$, 其中 $i = 1, 2$. 若 B_1 和 B_2 都是对称的, 或者都是反对称的, 则对所有 $T \in \mathrm{Hom}(V_1, V_2)$ 都有 $T^* = {}^*T$; 换言之, 此时 $B_2(v_2, Tv_1) = B_1\left(T^* v_2, v_1\right)$ 对所有 v_1, v_2 皆成立.

作为推论, 此时有 $(T^*)^* = T = {}^*({}^*T)$.

证明 当 B_1 和 B_2 都对称时命 $\epsilon = 1$, 否则命 $\epsilon = -1$. 此时

$$\begin{aligned} B_2\left(v_2, Tv_1\right) &= \epsilon B_2\left(Tv_1, v_2\right) = \epsilon B_1\left(v_1, T^* v_2\right) \\ &= \epsilon^2 B_1\left(T^* v_2, v_1\right) = B_1\left(T^* v_2, v_1\right). \end{aligned}$$

这就说明 ${}^*T = T^*$, 而第二部分是 (8.2.1) 的改写. □

以下术语因之是合理的.

约定 8.2.13 在 B_1 和 B_2 同为对称或反对称非退化双线性形式的情形, 我们将 T 的左伴随和右伴随统一称为伴随, 记为 T^*.

最常用的还是 $(V_1,B_1)=(V_2,B_2)$ 的特例. 对于矩阵情形, 一切都是具体的.

例 8.2.14 取 $V_i=V_i':=F^{n_i}$, 等同于列向量空间 $(i=1,2)$. 考虑 $\boldsymbol{A}_i\in \mathrm{M}_{n_i\times n_i}(F)$ 确定的双线性形式

$$\begin{aligned} B_i: F^{n_i}\times F^{n_i} &\longrightarrow F \\ (\boldsymbol{v},\boldsymbol{w}) &\longmapsto {}^{\mathrm{t}}\boldsymbol{v}\,\boldsymbol{A}_i\,\boldsymbol{w}. \end{aligned}$$

设 $\boldsymbol{A}_1$ 可逆; 根据例 8.2.5, 这相当于说 B_1 非退化.

给定矩阵 $\boldsymbol{T}\in \mathrm{M}_{n_2\times n_1}(F)$, 视同线性映射 $F^{n_1}\to F^{n_2}$, 如何明确它的右伴随? 兹断言

$$\boldsymbol{T}^* \overset{\text{等同于}}{=\!=} \boldsymbol{A}_1^{-1}\cdot {}^{\mathrm{t}}\boldsymbol{T}\cdot \boldsymbol{A}_2 \ \in \mathrm{M}_{n_1\times n_2}(F).$$

这点可以由练习 8.2.10 推导, 但更省心的方法或许是直接用矩阵来计算

$$\begin{aligned} B_2\left(\boldsymbol{T}\boldsymbol{v}_1,\boldsymbol{v}_2\right) &= {}^{\mathrm{t}}(\boldsymbol{T}\boldsymbol{v}_1)\,\boldsymbol{A}_2\,\boldsymbol{v}_2 = ({}^{\mathrm{t}}\boldsymbol{v}_1)\,{}^{\mathrm{t}}\boldsymbol{T}\,\boldsymbol{A}_2\boldsymbol{v}_2 = ({}^{\mathrm{t}}\boldsymbol{v}_1)\boldsymbol{A}_1\boldsymbol{A}_1^{-1}\cdot {}^{\mathrm{t}}\boldsymbol{T}\,\boldsymbol{A}_2\boldsymbol{v}_2 \\ &= ({}^{\mathrm{t}}\boldsymbol{v}_1)\boldsymbol{A}_1\left(\boldsymbol{A}_1^{-1}\cdot {}^{\mathrm{t}}\boldsymbol{T}\cdot \boldsymbol{A}_2\right)\boldsymbol{v}_2 = B_1\left(\boldsymbol{v}_1,(\boldsymbol{A}_1^{-1}\cdot {}^{\mathrm{t}}\boldsymbol{T}\cdot \boldsymbol{A}_2)\boldsymbol{v}_2\right); \end{aligned}$$

这相当于说 $\boldsymbol{A}_1^{-1}\cdot {}^{\mathrm{t}}\boldsymbol{T}\cdot \boldsymbol{A}_2$ 也具有刻画 $\boldsymbol{T}^*$ 的性质. 证毕.

定义 8.2.15 设 $B: V\times V\to F$ 是非退化双线性形式. 考虑线性映射 $T\in \mathrm{End}(V)$. 以下概念都是相对于 B 而言的.

- ⋆ 若 $T^*=T$, 则称 T 是**自伴**的.
- ⋆ 若 $T^*=-T$, 则称 T 是**反自伴**的.

严格来说, 应该按照 ${}^*T$ 和 T^* 分为左右两组版本. 由于今后主要考量对称或反对称的 B, 此处不深究.

例 8.2.16 考虑 $(\boldsymbol{v},\boldsymbol{w})\mapsto {}^{\mathrm{t}}\boldsymbol{w}\cdot\boldsymbol{v}$ 确定的对称非退化双线性形式 $F^n\times F^n\to F$; 这相当于在例 8.2.14 中取 $\boldsymbol{A}_1=\boldsymbol{A}_2=\mathbf{1}_{n\times n}$. 将 $\boldsymbol{T}\in \mathrm{M}_{n\times n}(F)$ 视同 $\mathrm{End}(F^n)$ 的元素, 从例 8.2.14 即刻导出

$$\text{伴随映射 } \boldsymbol{T}^* \overset{\text{等同于}}{=\!=} \text{作为矩阵取转置 } {}^{\mathrm{t}}\boldsymbol{T},$$

以及

$$\begin{aligned} \boldsymbol{T}\text{ 自伴} &\iff \underbrace{{}^{\mathrm{t}}\boldsymbol{T}=\boldsymbol{T}}_{\text{作为矩阵}}, \\ \boldsymbol{T}\text{ 反自伴} &\iff \underbrace{{}^{\mathrm{t}}\boldsymbol{T}=-\boldsymbol{T}}_{\text{作为矩阵}}. \end{aligned}$$

这又接上了定义 8.1.10 介绍的对称和反对称矩阵, 差异在于矩阵在 §8.1 承担了双线性形式的角色, 在此却是线性映射的化身.

最后接续 §8.1 结尾遗留的问题. 我们取 V 和 W 为有限维向量空间, $B : V \times W \to F$ 为双线性形式. 对给定的子空间 $W_0 \subset W$ 和 $V_0 \subset V$ 考虑定义 8.1.12 中的子空间 ${}^{\perp}W_0$ 和 $V_0^{\perp}$.

命题 8.2.17 设 $B : V \times W \to F$ 非退化, 则

$$\dim V_0^{\perp} + \dim V_0 = \dim V, \quad \dim {}^{\perp}W_0 + \dim W_0 = \dim W.$$

证明 两式的论证全然相似, 以下仅给出第一式的论证. 记 $d := \dim V_0$, $n := \dim V = \dim W$. $\psi : V \xrightarrow{\sim} W^{\vee}$ 为对应到 B 的同构. 考虑 V_0 在 ψ 之下的像, 取其基 $\check{w}_1, \dots, \check{w}_d$, 所以 $V_0^{\perp} \subset W$ 是由以下等式定义的子空间

$$\langle \check{w}_1, \cdot \rangle = \cdots = \langle \check{w}_d, \cdot \rangle = 0. \tag{8.2.3}$$

根据即将证明的维数引理 8.2.18, 可见 $V_0^{\perp}$ 的维数确实是 $n - d$. □

请注意: 命题 8.2.17 仅是确定了 $V_0^{\perp}$ 的维数. 对于一般的 $V_0 \subset V$ 和非退化双线性形式 $B : V \times V \to F$, 未必有直和分解 $V = V_0 \oplus V_0^{\perp}$.

为了补全上述证明, 现在来确立一则关于维数的一般性结果, 证明方法也颇有趣.

引理 8.2.18 设 W 为有限维向量空间, $d \in \mathbb{Z}_{\geq 1}$ 而 $\check{w}_1, \dots, \check{w}_d \in W^{\vee}$ 线性无关, 则 W 的子空间

$${}^{\perp}\langle \check{w}_1, \dots, \check{w}_d \rangle := \{ w \in W : \langle \check{w}_1, w \rangle = \cdots = \langle \check{w}_d, w \rangle = 0 \}$$

的维数是 $\dim W - d$.

证明 从两个极端状况入手. 命 $n := \dim W$. 先设 $d = n$, 则 $\langle \check{w}_1, \dots, \check{w}_n \rangle^{\perp}$ 便是 $\{ w \in W : \forall \check{w}, \langle \check{w}, w \rangle = 0 \}$. 例 8.2.2 已经说明这是零空间.

其次设 $d = 1$. 已知 $\check{w}_1 \neq 0$, 故 $\check{w}_1$ 作为线性映射 $W \to F$ 必然满, 从而 ${}^{\perp}\langle \check{w}_1 \rangle = \ker(\check{w}_1)$ 的维数是 $n - 1$ (定理 4.8.4).

对于一般情形, 将 $\check{w}_1, \dots, \check{w}_d$ 扩充为 $W^{\vee}$ 的基 $\check{w}_1, \dots, \check{w}_n$ 并考虑 W 的子空间列

$${}^{\perp}\langle \check{w}_1 \rangle \supset \cdots \supset {}^{\perp}\langle \check{w}_1, \dots, \check{w}_n \rangle . \tag{8.2.4}$$

一般而言, 设 $W_1, W_2 \subset W$ 为子空间, 则有

$$\begin{aligned} \dim(W_1 \cap W_2) &= \dim W_1 + \dim W_2 - \dim(W_1 + W_2) \\ &\geq \dim W_1 + \dim W_2 - n; \end{aligned}$$

详见练习 4.10.5. 对所有 $1 \leq k < n$, 基于先前解释过的 ${}^{\perp}\langle \check{w}_{k+1} \rangle = n - 1$, 我们得到

$$\begin{aligned} \dim {}^{\perp}\langle \check{w}_1, \dots, \check{w}_{k+1} \rangle &= \dim {}^{\perp}\langle \check{w}_1, \dots, \check{w}_k \rangle \cap {}^{\perp}\langle \check{w}_{k+1} \rangle \\ &\geq \dim {}^{\perp}\langle \check{w}_1, \dots, \check{w}_k \rangle - 1. \end{aligned}$$

然而 (8.2.4) 的首项是 $n-1$ 维, 末项是 0 维, 故它每步恰好降 1 维. 特别地, $\langle \check{w}_1, \dots, \check{w}_d \rangle^\perp$ 是 $n-d$ 维的. □

现在可以简单地对非退化双线性形式描述双重正交子空间.

定理 8.2.19 设 $B : V \times W \to F$ 非退化, 则对所有子空间 $V_0 \subset V$ 和 $W_0 \subset W$ 皆有

$$ {}^\perp(V_0^\perp) = V_0, \quad ({}^\perp W_0)^\perp = W_0. $$

对于 $V = W$ 和 B 对称或反对称的情形, 上式也可以简记为 $V_0^{\perp\perp} = V_0$.

证明 基于平凡的理由, $V_0 \subset {}^\perp(V_0^\perp)$. 然而 $\dim V = \dim W$ 和命题 8.2.17 (应用两次) 又蕴涵两边维数相等, 于是 $V_0 = {}^\perp(V_0^\perp)$. 关于 W_0 的论证完全类似. □

8.3 分类问题的提出

先前着重讨论了形如 $B : V \times V \to F$ 的双线性形式. 我们希望进一步地了解, 甚至是在合适条件下分类这种数学对象. 本节不解决分类问题, 而是为问题的提出作好准备.

一如熟悉的向量空间情形, 当我们谈及分类, 真正重要的不是资料 (V, B) 本身, 而是它们的同构类. 何谓同构?

定义 8.3.1 考虑资料 (V_1, B_1) 和 (V_2, B_2), 其中 V_i 是 F-向量空间而 $B_i : V_i \times V_i \to F$ 是双线性形式 $(i = 1, 2)$. 从 (V_1, B_1) 到 (V_2, B_2) 的**同构**意谓满足以下条件的线性映射 $\varphi \in \mathrm{Hom}(V_1, V_2)$:

- ⋆ φ 是向量空间的同构;
- ⋆ $B_2(\varphi(v_1), \varphi(v_1')) = B_1(v_1, v_1')$ 对所有 $v_1, v_1' \in V_1$ 成立.

上述同构也记为 $\varphi : (V_1, B_1) \xrightarrow{\sim} (V_2, B_2)$. 若存在这样的同构, 则称资料 (V_1, B_1) 和 (V_2, B_2) 是同构[2)]的, 简记为 $(V_1, B_1) \simeq (V_2, B_2)$.

双线性形式的同构是等价关系:

- ▷ **反身性** id_V 当然地给出同构 $(V, B) \xrightarrow{\sim} (V, B)$;
- ▷ **对称性** 若 $\varphi : (V_1, B_1) \xrightarrow{\sim} (V_2, B_2)$, 则逆映射给出 $\varphi^{-1} : (V_2, B_2) \xrightarrow{\sim} (V_1, B_1)$, 这是因为易见原条件等价于

$$ B_2(v_2, v_2') = B_1(\varphi^{-1}(v_2), \varphi^{-1}(v_2')), \quad v_2, v_2' \in V_2; $$

[2)] 一些文献称之为等价.

▷ **传递性** 若 $\varphi:(V_1,B_1)\xrightarrow{\sim}(V_2,B_2)$ 而 $\psi:(V_2,B_2)\xrightarrow{\sim}(V_3,B_3)$, 则 $\psi\varphi:(V_1,B_1)\xrightarrow{\sim}(V_3,B_3)$. 验证也毫无困难.

在上述定义中, 不仅底层的向量空间 V_1 和 V_2 以 φ 相互等同, 其上的双线性形式也通过 φ 相互匹配. 因此同构的双线性形式可以设想为实质相同的. 我们已经对不少代数结构 (域, 环, 向量空间...) 谈过同构的概念, 此处不过是搬演同样的套路.

关于双线性形式的一切性质都可以通过同构来转译, 以下不过是略举其简单面向. 证明大概是多余的.

命题 8.3.2 *设 (V_i,B_i) 为双线性形式 $(i=1,2)$, 而 $\varphi:(V_1,B_1)\xrightarrow{\sim}(V_2,B_2)$ 为同构. 我们有:*

(i) $\dim V_1=\dim V_2$;

(ii) *φ 映 B_1 的左根 (或右根) 为 B_2 的左根 (或右根);*

(iii) *B_1 非退化当且仅当 B_2 亦然;*

(iv) *B_1 对称 (或反对称) 当且仅当 B_2 亦然.*

在后续针对对称或反对称双线性形式的研究中, 根基的概念将会频繁出现.

定义 8.3.3 当 $B:V\times V\to F$ 是对称或反对称双线性形式时, 其左根等于右根, 统一称为 (V,B) 的**根基**, 视情况另记为 $R(V)$, $R(B)$ 或 $R(V,B)$.

定义 8.3.4 对于资料 (V,B) 和 V 的子空间 V_0, 将双线性映射 B 限制在 $V_0\times V_0$ 上给出资料 (V_0,B), 称之 B 在 V_0 上的**限制**. 若 B 是对称 (或反对称) 的, 则其限制亦然.

从已有的双线性形式构造新货的一种手段是直和. 我们顺带引入相关符号.

定义 8.3.5 资料 (V_1,B_1) 和 (V_2,B_2) 的**直和** $(V_1,B_1)\oplus(V_2,B_2):=(V_1\oplus V_2,B_1\oplus B_2)$ 按定义 8.1.8 的方式来构造. 由于 V_1 和 V_2 的元素在 $B_1\oplus B_2$ 配对下总是取零, 这种直和也常被称为**正交直和**.

若有 $\varphi_i:(V_i,B_i)\xrightarrow{\sim}(V_i',B_i')$, 其中 $i=1,2$, 则自然也有 $(V_1,B_1)\oplus(V_2,B_2)\xrightarrow{\sim}(V_1',B_1')\oplus(V_2',B_2')$.

同样地, 可以定义任意份资料的 (正交) 直和 $\bigoplus_i(V_i,B_i)$.

直和的另一种用法是将给定的 (V,B) 简化. 比方说它可以将 (V,B) 的研究化约到非退化情形, 细说如下.

命题 8.3.6 *考虑资料 (V,B). 存在子空间 K 使得 $V=R(V)\oplus K$, 而且 B 在 K 上的限制是非退化双线性形式 $K\times K\to F$. 因此我们得到直和分解*

$$(V,B)=(R(V),0)\oplus(K,B),$$

其中 0 代表恒取零的平凡双线性形式

证明 命题 4.10.6 确保存在子空间 K 使得 $V = R(V) \oplus K$. 兹说明 (K, B) 非退化. 设 $v \in K$ 使得对所有 $v' \in K$ 皆有 $B(v, v') = 0$, 则因为条件对于 $v' \in R(V)$ 显然也成立, 故 $B(v, \cdot)$ 恒为零, $v \in R(V) \cap K = \{0\}$. 其余断言都是平凡的. □

练习 8.3.7 从商映射限制而来的同构 $K \xrightarrow{\sim} V/R(V)$ (推论 4.12.11) 等同 $B : K \times K \to F$ 与商空间上诱导的双线性形式 $\overline{B} : V/R(V) \times V/R(V) \to F$ (练习 8.2.7). 试按定义严格地检验这些断言.

泛泛而论双线性形式 (V, B) 的分类几乎是没有内容的. 在后续研究中, 我们将对双线性形式的对称/反对称性, 非退化性, 维数以及域 F 本身加上种种限定. 比方说, 我们将在大部分场合假定 V 是有限维的. 不同场景下的结果和技术都具有各自的鲜明风格.

最后将理论落实到 $V = F^n$ 的具体场景, 说明如何在矩阵的层次诠释资料 (F^n, B) 与 (F^n, B') 之间的同构. 这将涉及以下概念.

定义 8.3.8 (矩阵的合同关系) 设 $\boldsymbol{A}, \boldsymbol{A}' \in \mathrm{M}_{n\times n}(F)$. 若存在可逆的 $\boldsymbol{C} \in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{A} = {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C}$, 则称 $\boldsymbol{A}$ 和 $\boldsymbol{A}'$ **合同**.

练习 8.3.9 运用转置的性质 ${}^{\mathrm{t}}(\boldsymbol{C}_1\boldsymbol{C}_2) = {}^{\mathrm{t}}\boldsymbol{C}_2\,{}^{\mathrm{t}}\boldsymbol{C}_1$ 和 $({}^{\mathrm{t}}\boldsymbol{C})^{-1} = {}^{\mathrm{t}}(\boldsymbol{C}^{-1})$, 说明合同是 $\mathrm{M}_{n\times n}(F)$ 上的等价关系. 验证若 $\boldsymbol{A}$ 和 $\boldsymbol{A}'$ 合同, 则 $\boldsymbol{A}$ 对称 (或反对称) 当且仅当 $\boldsymbol{A}'$ 亦然.

此外,

$$\begin{aligned} {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C} = \boldsymbol{A} &\iff \forall \boldsymbol{v}_1, \boldsymbol{v}_2, \quad {}^{\mathrm{t}}\boldsymbol{v}_1 \left({}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C}\right) \boldsymbol{v}_2 = {}^{\mathrm{t}}\boldsymbol{v}_1\, \boldsymbol{A}\, \boldsymbol{v}_2 \\ &\iff \forall \boldsymbol{v}_1, \boldsymbol{v}_2, \quad {}^{\mathrm{t}}(\boldsymbol{C}\boldsymbol{v}_1)\, \boldsymbol{A}'\, (\boldsymbol{C}\boldsymbol{v}_2) = {}^{\mathrm{t}}\boldsymbol{v}_1\, \boldsymbol{A}\, \boldsymbol{v}_2, \end{aligned}$$

其中 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 遍历 F^n 的元素, 视同列向量. 若考虑 $\boldsymbol{A}$ (或 $\boldsymbol{A}'$) 所对应的双线性形式 B (或 B'), 将可逆矩阵 $\boldsymbol{C}$ 视同向量空间的同构 $F^n \xrightarrow{\sim} F^n$, 则最后一则条件也相当于说

$$B'(\boldsymbol{C}\boldsymbol{v}_1, \boldsymbol{C}\boldsymbol{v}_2) = B(\boldsymbol{v}_1, \boldsymbol{v}_2),$$

而这正是定义 8.3.1 所谓的同构. 这些观察总结如下.

命题 8.3.10 设 B 和 B' 为双线性形式 $F^n \times F^n \to F$, 分别对应到 $n \times n$ 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}'$, 则有双射

$$\begin{aligned} \left\{\boldsymbol{C} \in \mathrm{M}_{n\times n}(F) : \text{可逆}, {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C} = \boldsymbol{A}\right\} &\xrightarrow{1:1} \left\{\varphi : (F^n, B) \xrightarrow{\sim} (F^n, B')\right\} \\ \boldsymbol{C} &\longmapsto \text{相应的同构 } F^n \xrightarrow{\sim} F^n. \end{aligned}$$

因此, 有限维向量空间上的双线性形式之间的同构和矩阵的合同是一回事.

8.4 二次型的基本概念

本节考量有限维 F-向量空间上的对称双线性形式 (V, B). 以下设 F 是满足 $\mathrm{char}(F) \neq 2$ 的域 (定义 3.7.2), 这是因为相关操作中经常需要作 "除以 2" 的运算, 见练习 3.7.6 的讨论.

命 $n := \dim V$. 选定 V 的有序基以得到同构 $\varphi : V \xrightarrow{\sim} F^n$. 于是按照 §8.3 的语言, $\varphi : (V, B) \xrightarrow{\sim} (F^n, B')$, 其中的 $B' : F^n \times F^n \to F$ 由

$$B'(\varphi(v_1), \varphi(v_2)) = B(v_1, v_2), \quad v_1, v_2 \in V$$

确定. 综之, n 维对称双线性形式的分类化约到 $V = F^n$ 的情形. 根据命题 8.1.6 和引理 8.1.11, 它们进一步与 F 上的 $n \times n$ 对称矩阵一一对应.

现在将这些资料用更经典而具体的多项式语言改述. 选定 $n \in \mathbb{Z}_{\geq 1}$.

定义 8.4.1 (n 元二次型) 域 F 上的 n 元齐次二次多项式 (定义 3.3.4) 称为 n 元二次型[3].

按定义, n 元二次型可以表为

$$f = \sum_{i=1}^{n} a_{ii} X_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j$$

的形式, 其中的 $a_{ij} \in F$ 是由 f 唯一确定的系数 $(1 \leq i \leq j \leq n)$. 我们也可以另外对 $i \geq j$ 的情形命 $a_{ij} := a_{ji}$. 尽管这么做显得冗余, 但好处是 f 可以更简练地表作 $\sum_{1 \leq i,j \leq n} a_{ij} X_i X_j$, 而且由此得到定义 8.1.10 意义下的对称矩阵

$$\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n \times n}(F).$$

反过来说, 任何对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n \times n}(F)$ 都按 $f = \sum_{i,j} a_{ij} X_i X_j$ 唯一地确定了 n 元二次型. 综上, 我们得到双射

$$\begin{array}{ccccc} \{n \text{ 元二次型}\} & \xleftrightarrow{1:1} & \left\{ \begin{array}{l} \boldsymbol{A} \in \mathrm{M}_{n \times n}(F) \\ \text{对称矩阵} \end{array} \right\} & \xleftrightarrow[\text{命题 8.1.6}]{1:1} & \left\{ \begin{array}{l} B : F^n \times F^n \to F \\ \text{对称双线性形式} \end{array} \right\} \\ \cup & & \cup & & \cup \\ f = \sum_{1 \leq i,j \leq n} a_{ij} X_i X_j & & \boldsymbol{A} = (a_{ij})_{i,j} & & B(\boldsymbol{v}_1, \boldsymbol{v}_2) = {}^{\mathrm{t}}\boldsymbol{v}_1 \, \boldsymbol{A} \, \boldsymbol{v}_2 \end{array} \tag{8.4.1}$$

按惯例, 矩阵运算中将 F^n 的元素 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 视同列向量; 此外它们还能代入 n 元多项式来求值.

因此, 对于二次型至少有三种观点: 作为多项式, 对称矩阵或对称双线性形式. 它们各有优点.

[3] 在经典文献中, "型" 往往是一类齐次多项式的代称.

命题 8.4.2 设 n 元二次型 f 按 (8.4.1) 对应到对称双线性形式 $B : F^n \times F^n \to F$, 则

$$f(\boldsymbol{v}) = B(\boldsymbol{v}, \boldsymbol{v}), \quad B(\boldsymbol{v}_1, \boldsymbol{v}_2) = \frac{1}{2}\left(f(\boldsymbol{v}_1 + \boldsymbol{v}_2) - f(\boldsymbol{v}_1) - f(\boldsymbol{v}_2)\right).$$

证明 第一式直接来自 $B(\boldsymbol{v}, \boldsymbol{v}) = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j$, 其中 $\boldsymbol{v} = (x_1, \ldots, x_n)$. 第二式则是来自于

$$\begin{aligned} f(\boldsymbol{v}_1 + \boldsymbol{v}_2) &= B(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2) \\ &= B(\boldsymbol{v}_1, \boldsymbol{v}_1) + B(\boldsymbol{v}_2, \boldsymbol{v}_2) + B(\boldsymbol{v}_1, \boldsymbol{v}_2) + B(\boldsymbol{v}_1, \boldsymbol{v}_2) \\ &= f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2) + 2B(\boldsymbol{v}_1, \boldsymbol{v}_2). \end{aligned}$$

其中 $\boldsymbol{v}_1 = (x_1, \ldots, x_n)$ 而 $\boldsymbol{v}_2 = (y_1, \ldots, y_n)$; 此处 B 的对称性是关键. □

留意到如果 B 对应的是一般的矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 则代入 $\boldsymbol{v} = (x_1, \ldots, x_n)$ 给出

$$B(\boldsymbol{v}, \boldsymbol{v}) = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} (a_{ij} + a_{ji}) x_i x_j,$$

这和对应到 $\frac{1}{2}(\boldsymbol{A} + {}^{\mathrm{t}}\boldsymbol{A})$ 的对称双线性形式给出一样的结果.

练习 8.4.3 对应 (8.4.1) 和命题 8.4.2 顺道说明尽管 n 元二次型 f 是 $F[X_1, \ldots, X_n]$ 的元素, 它却由对应的多项式函数 $F^n \to F$ 完全确定. 请尝试直接证明: 若域 F 满足 $\mathrm{char}(F) \neq 2$ 而 $f = \sum_{i \leq j} a_{ij} X_i X_j + \sum_i b_i X_i + c \in F[X_1, \ldots, X_n]$, 则 f 由对应的多项式函数唯一确定.

定义 8.4.4 (向量空间上的二次型) 考虑资料 (V, B), 其中 V 是有限维 F-向量空间, $B : V \times V \to F$ 是对称双线性形式. 此时我们简称 (V, B) 为域 F 上的二次型, 实现在空间 V 上. 基于命题 8.4.2, 一个 n 元二次型 (V, B) 也可以等价地用资料 (V, f) 来刻画, 其中 $f = B(v, v)$.

鉴于 (8.4.1), 实现在 F^n 上的二次型无非是定义 8.4.1 所谓的 n 元二次型. 多项式的观点经常是比较具体的, 但基于 (V, B) 的定义在操作中更具弹性, 例如可以由定义 8.1.8 直接说明何谓二次型的直和.

命题 8.3.10 以矩阵的合同等价来诠释双线性形式之间的同构. 施之于 F^n 上的二次型 B 和 B', 便推得 $(F^n, B) \simeq (F^n, B')$ 当且仅当对应的 $n \times n$ 对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}'$ 合同.

约定 8.4.5 当上述条件成立时, 我们也称相应的 n 元二次型 f 和 f' 同构.

从多项式的角度观照, n 元二次型 f 和 f' 之间的同构是由线性而且可逆的变量替换

$$\boldsymbol{Y} := \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \boldsymbol{C} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \boldsymbol{C}\boldsymbol{X}, \quad \boldsymbol{C} = (c_{ij})_{1 \leq i,j \leq n} \text{ 可逆}$$

来实现的: 设 f 对应到对称矩阵 $\boldsymbol{A}$ 而 f' 对应到对称矩阵 $\boldsymbol{A}'$, 并且有合同关系 $\boldsymbol{A} = {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C}$, 则

$$f(X_1, \dots, X_n) = {}^{\mathrm{t}}\boldsymbol{X}\,{}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}'\boldsymbol{C}\boldsymbol{X} = {}^{\mathrm{t}}\boldsymbol{Y}\boldsymbol{A}'\boldsymbol{Y} = f'(Y_1, \dots, Y_n).$$

由于 Y_i 具体展开为多项式 $\sum_{j=1}^{n} c_{ij}X_j$, 因此 f 和 f' 的关系也可以简单地表作多项式的合成

$$f = f' \circ (Y_1, \dots, Y_n) \ \in F[X_1, \dots, X_n],$$

来理解 (定义 3.3.6).

综上, 二次型的分类问题转译为对称矩阵在合同意义下的分类. 处理此问题的第一步是在合同意义下进行对角化, 从多项式的角度来看便是配方.

8.5 配方法

仍然设 F 是域, $\mathrm{char}(F) \neq 2$. 首先说明任何 n 元二次型都可以用同构化到 $a_1X_1^2 + \cdots + a_nX_n^2$ 的形式, $a_i \in F$. 这种二次型对应到对角矩阵, 因而也称是为对角的.

命题 8.5.1 (二次型的对角化) *任何 n 元二次型都同构于形如 $a_1X_1^2 + \cdots + a_nX_n^2$ 的对角二次型.*

证明 对 n 递归地论证: $n = 1$ 情形平凡. 设 $n \geq 2$. 将所论的二次型写成 $f = \sum_{1 \leq i,j \leq n} a_{ij}X_iX_j$. 无妨假设系数 a_{ij} 不全为零, 否则无事可作. 分两种情形讨论.

如果存在 $1 \leq i \leq n$ 使得 $a_{ii} \neq 0$, 适当重排变元后不妨假设 $i = 1$, 则

$$f = a_{11}\left(X_1 + \frac{1}{a_{11}}\sum_{j=2}^{n} a_{1j}X_j\right)^2 + \sum_{2 \leq i,j \leq n} a_{ij}X_iX_j - \frac{1}{a_{11}}\left(\sum_{j=2}^{n} a_{1j}X_j\right)^2.$$

考虑变量替换

$$Y_i := \begin{cases} X_1 + \frac{1}{a_{11}}\sum_{j=2}^{n} a_{1j}X_j, & i = 1, \\ X_i, & 2 \leq i \leq n. \end{cases}$$

这是线性而且可逆的, 由 $X_1 = Y_1 - \frac{1}{a_{11}}\sum_{j=2}^{n} a_{1j}Y_j$ 反解. 二次型在这些变量下改写成 $a_{11}Y_1^2 + g(Y_2, \dots, Y_n)$, 其中

$$g = \sum_{2 \leq i \leq j \leq n} a_{ij}Y_iY_j - \frac{1}{a_{11}}\left(\sum_{j=2}^{n} a_{1j}Y_j\right)^2$$

是 $n-1$ 元二次型, 因此 g 可以继续对角化.

如果所有 a_{ii} 全为 0, 则存在 $i < j$ 使得 $a_{ij} \neq 0$. 施行变量替换

$$Y_k := \begin{cases} X_i - X_j & k = i, \\ X_k, & k \neq i. \end{cases}$$

由 $X_i = Y_i + Y_j$ 可见这仍是线性而且可逆的. 按此将 $\frac{1}{2}f = \sum_{k<h} a_{kh}X_kX_h$ 表示为

$$\frac{1}{2}f = \sum_{\substack{1\le k<h\le n\\ k,h\neq i}} a_{kh}Y_kY_h + \sum_{i<h\le n} a_{ih}(Y_i+Y_j)Y_h + \sum_{1\le k<i} a_{ki}Y_k(Y_i+Y_j).$$

上式中 Y_j^2 的系数为 $a_{ij} \neq 0$. 按此化约到前一情形. □

证明过程具有算法的特性. 事实上, 处理关键的 $a_{11} \neq 0$ 情形的手段正是**配方**.

按照直和的观点 (定义 8.3.5), 对角化也相当于在同构意义下将二次型 f 表为 n 个形如 aX^2 的一元二次型的直和.

算法 8.5.2 任何可逆矩阵皆可写成初等矩阵的积 $\boldsymbol{U}_1\cdots\boldsymbol{U}_k$. 考虑对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 命题 8.5.1 的对角化手续相当于寻求一列初等矩阵 $\boldsymbol{U}_1,\dots,\boldsymbol{U}_k$ 使得

$$^{\mathrm{t}}\boldsymbol{U}_k\cdots{}^{\mathrm{t}}\boldsymbol{U}_1\,\boldsymbol{A}\,\boldsymbol{U}_1\cdots\boldsymbol{U}_k \quad \text{为对角矩阵.}$$

左式相当于对 $\boldsymbol{A}$ 依序施行 $\boldsymbol{U}_i$ 对应的初等行变换以及列变换 (两者顺序可换), 其中 $i = 1,\dots,k$. 对应的 $\boldsymbol{C} = \boldsymbol{U}_1\cdots\boldsymbol{U}_k$ 则是对单位矩阵 $\mathbf{1}_{n\times n}$ 依序施行 $\boldsymbol{U}_1,\dots,\boldsymbol{U}_k$ 对应的初等列变换的成果.

成对地施行这些初等行, 列变换, 总可以将 A 的非对角线元素零化. 这是命题 8.5.1 证明过程的矩阵表述. 请读者尝试用矩阵语言重证命题 8.5.1.

尽管 n 元二次型能够化到 $\sum_{i=1}^n a_iX_i^2$ 的形式, 其中 $a_1,\dots,a_n \in F$, 但不同的 $a_1,\dots,a_n$ 给出的二次型可以相互同构. 比如命 $Y_i = tX_i$, 其中 $t \in F^\times$, 则 $a_iX_i^2 = t^{-2}a_iY_i^2$. 但至少二次型的根基 (即: 对应的对称双线性形式的根基) 容易从对角化读出.

引理 8.5.3 *若 n 元二次型 f 写作 $\sum_{i=1}^r a_iX_i^2$, 其中 $0 \le r \le n$ 而 $a_1,\dots,a_r \in F^\times$, 则 f 的根基等于 $\langle \boldsymbol{e}_{r+1},\dots,\boldsymbol{e}_n\rangle$.*

证明 由于 f 对应的对称矩阵 $\boldsymbol{A}$ 以 $a_1,\dots,a_r,0,\dots,0$ 为对角元, 其他位置为 0, 回忆引理 8.2.3 的内容可知子空间

$$\begin{aligned}\ker(\boldsymbol{A}) &= \{(x_1,\dots,x_n) \in F^n : a_1x_1 = \cdots = a_rx_r = 0\}\\ &= \{(0,\dots,0,x_{r+1},\dots,x_n) \in F^n : x_{r+1},\dots,x_n \in F\}\\ &= \langle \boldsymbol{e}_{r+1},\dots,\boldsymbol{e}_n\rangle\end{aligned}$$

给出 f 的根基. □

于是 n 元二次型 f 非退化当且仅当它对角化之后的系数 $a_1,\dots,a_n$ 全部非零.

定义 8.5.4 *二次型 f 的**秩** r 定义为变元个数 n 减去 f 的根基的维数.*

同构的二次型有相同的秩. 非退化 n 元二次型正是秩为 n 的二次型. 若二次型 f 对应到对称矩阵 $\boldsymbol{A}$, 则 f 的秩等于 $\mathrm{rk}(\boldsymbol{A})$, 这是定理 4.8.4 的结论.

例 8.5.5 (代数闭域的情形) 取 F 为定义 6.6.5 所谓的代数闭域, 例如 $F=\mathbb{C}$. 任何 n 元二次型都能化到 $a_1X_1^2+\cdots+a_rX_r^2$ 的形式, 其中 $r\leq n$ 而 $a_1,\ldots,a_r\in F^\times$. 以 b_iX_i 代替 X_i, 其中 $1\leq i\leq r$ 而 $b_i\in F$ 满足 $b_i^2=a_i$, 便将二次型进一步化为 $X_1^2+\cdots+X_r^2$ 的形式, 其中的 r 无非是二次型的秩, 它仅依赖二次型的同构类. 综上, 此时有双射

$$\{n\text{ 元二次型的同构类}\}\xrightarrow{1:1}\{0,1,\ldots,n\}$$
$$f\text{ 的同构类}\longmapsto f\text{ 的秩}.$$

换言之, 此时二次型完全由秩来分类. 代数闭的条件在此其实过强, 真正用到的条件是 F 的所有元素都有平方根.

8.6 实二次型的分类

在 §8.5 已经看到复二次型容易分类, 本节转向 $F=\mathbb{R}$ 情形. 配方法可以化任意 n 元实二次型 f 为 $a_1X_1^2+\cdots+a_rX_r^2$ 的形式, 其中 r 是 f 的秩, 而 $a_1,\ldots,a_r\in\mathbb{R}^\times$. 由于 a_i 在 $\mathbb{R}$ 中未必有平方根. 所以在将变元适当重排后, 我们只能将实二次型 f 在同构意义下化到

$$X_1^2+\cdots+X_p^2-X_{p+1}^2-\cdots-X_r^2,$$

其中 $0\leq p\leq r$. 姑且称以上形式的实 n 元二次型或对应的对称双线性形式为规范形.

规范形中的数字 p 如何诠释? 我们先引入若干概念. 对称双线性形式的语言在此更为方便.

定义 8.6.1 设 V 是 $\mathbb{R}$-向量空间, $B:V\times V\to\mathbb{R}$ 是对称双线性形式.

- ⋆ 若 $B(v,v)\geq 0$ 对所有 $v\in V$ 成立, 则称 B 是**半正定**的;
- ⋆ 若 B 半正定, 而且 $v\neq 0$ 蕴涵 $B(v,v)>0$, 则称之为**正定**的.

如果 $-B$ 半正定 (或正定), 则称 B **半负定** (或**负定**). 若以上皆非, 则称 B **不定**.

若进一步要求 V 有限维, 便能为定义 8.4.4 介绍的二次型 (V,B) 在 $F=\mathbb{R}$ 时探讨何谓半正定, 正定, 半负定, 负定, 不定. 这些概念只和 (V,B) 的同构类相关.

练习 8.6.2 说明正定或负定的二次型 (V,B) 必然非退化. 〈提示〉 若 v 属于 B 的根基, 则 $B(v,v)=0$.

进一步取 $V=\mathbb{R}^n$, 再用命题 8.4.2 的对应, 便能够对作为多项式的 n 元实二次型 f 谈论正定性等种种概念, 它们同样只依赖 f 的同构类.

命题 8.6.3 设 n 元实二次型 f 同构于规范形 $X_1^2+\cdots+X_p^2-X_{p+1}^2-\cdots-X_r^2$, 其中 r 是 f 的秩, 则

⋆ f 半正定当且仅当 $p = r$;

⋆ f 正定当且仅当 $p = n$.

对于半负定和负定性也有类似的刻画, 代 f 为 $-f$ 即可相互过渡.

证明 要点在于非零实数的平方和总为正数. 另一方面, 倘若规范形包含 $X_p^2 - X_{p+1}^2 = (X_p + X_{p+1})(X_p - X_{p+1})$ 的部分, 则显然对任何 $c \in \mathbb{R}$ 皆可找到 $\boldsymbol{v} \neq \mathbf{0}$ 使 $f(\boldsymbol{v}) = c$.□

现已初见规范形和正定性或半正定性的关系. 一般情形下, 实二次型的同构类中的规范形是否唯一? 答案是肯定的, 由此可以推导实二次型的完整分类.

约定 8.6.4 对于任何二次型 (V, B) 和 V 的子空间 V_0. 按照定义 8.3.4 考虑 B 在 V_0 上的限制. 若此限制为正定 (或负定) 的二次型, 则称 V_0 是正定 (或负定) 子空间. 类似地定义何谓半正定 (或半负定) 子空间.

引理 8.6.5 设二次型 (V, B) 同构于规范型 $X_1^2 + \cdots + X_p^2 - X_{p+1}^2 - \cdots - X_r^2$, 其中 $0 \leq p \leq r$, 则:

(i) 存在 V 的 p 维正定子空间;

(ii) 任何维数 $> p$ 的子空间都不是正定的.

证明 命 $n := \dim V$. 由于断言中的性质只和二次型的同构类相关, 不妨就假设 $V = \mathbb{R}^n$ 而 B 对应到多项式 $f = X_1^2 + \cdots + X_p^2 - X_{p+1}^2 - \cdots - X_r^2$. 取 p 维子空间

$$V_+ := \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_p \rangle .$$

命题 8.6.3 说明 V_+ 正定, (i) 得证.

命 $N := \langle \boldsymbol{e}_{p+1}, \ldots, \boldsymbol{e}_n \rangle$, 命题 8.6.3 说明 N 半负定. 设 V 的子空间 V' 满足 $\dim V' > p$, 则

$$\begin{aligned} \dim(V' \cap N) &= \dim V' + \dim N - \dim(V' + N) \\ &\geq \dim V' + \dim N - n \\ &= \dim V' + n - p - n > 0. \end{aligned}$$

故存在 $V' \cap N$ 的非零元 $\boldsymbol{v}$. 从 $\boldsymbol{v} \in N$ 得 $f(\boldsymbol{v}) \leq 0$. 这表明 V' 非正定, (ii) 得证. □

定理 8.6.6 (惯性定理) 选定 $n \in \mathbb{Z}_{\geq 1}$, 则 n 元实二次型的每个同构类中有唯一的规范形.

证明 考虑二次型 (V, B), (V', B') 和 $\varphi : (V, B) \xrightarrow{\sim} (V', B')$. 记 n (或 r) 为 (V, B) 和 (V', B') 共有的维数 (或秩). 取 V 的基 $v_1, \ldots, v_n$ 和 V' 的基 $v_1', \ldots, v_n'$, 使得 B 和 B'

分别对应于对角元为

$$\underbrace{1,\dots,1}_{p\text{ 项}},\underbrace{-1,\dots,-1}_{r-p\text{ 项}},\underbrace{0,\dots,0}_{n-r\text{ 项}},$$
$$\underbrace{1,\dots,1}_{p'\text{ 项}},\underbrace{-1,\dots,-1}_{r-p'\text{ 项}},\underbrace{0,\dots,0}_{n-r\text{ 项}}$$

的对角矩阵. 目标是证 $p=p'$.

使用反证法. 设 $p'>p$. 对 (V',B') 应用引理 8.6.5 (i) 可得 p' 维正定子空间 $V'_+\subset V'$, 以同构搬运便得到 V 的正定子空间 $\varphi^{-1}(V'_+)$. 然而对 (V,B) 应用引理 8.6.5 (ii) 可知任何 p' 维子空间均不可能正定. 矛盾. □

这一结果由 J. J. Sylvester 发现并命名, 他以 "惯性" 譬喻坐标变换下的不变性.

鉴于定理 8.6.6, 以下定义是合理的.

定义 8.6.7 *在实二次性 f 的同构类中取规范形, 其中的 p 称为 f 的*__正惯性指数__, $q:=r-p$ *称为*__负惯性指数__, *而两者的差 $p-q=2p-r$ 称为 f 的*__符号差__.

正负惯性指数的和等于秩, 因此两个 n 元实二次型同构当且仅当它们有相同的秩和符号差. 两个非退化 n 元实二次型同构当且仅当它们的符号差相同.

基于以上一系列说明, 实二次型的秩, 惯性指数或符号差可以通过配方法和观察正负号来确定. 往后的定理 9.7.4 将另外给出一种以行列式判断正定性的方法.

8.7 反对称双线性形式: 辛空间

之前侧重于有限维向量空间上的对称双线性形式. 现在我们将目光转向反对称情形, 它们的分类是比较简单的. 本节继续默认域 F 满足 $\mathrm{char}(F)\neq 2$. 这就导致反对称双线性形式 B 总是满足 $B(v,v)=0$.

首先聚焦于带有非退化反对称双线性形式的有限维向量空间. 它们在几何学, 数学物理与计算数学等领域中有精彩应用, 值得一个特殊的名字.

定义 8.7.1 *设 V 为有限维向量空间, 其上的非退化反对称双线性形式 $B:V\times V\to F$ 称为*__辛形式__, *而资料 (V,B) 也称为*__辛空间__, *往往简记为 V.*

按此便可以谈论辛空间的直和, 其间的同构, 以及辛空间的分类问题. 先看辛空间的一则具体例子. 考虑带有有序基 p,q 的二维空间 $Fp\oplus Fq$, 取其上的反对称双线性形式 B 使得

$$B(p,q)=1=-B(q,p);$$

因为反对称性蕴涵 $B(p,p)=0=B(q,q)$, 上式完全刻画了 B. 若以有序基 p,q 将此空

间等同于 F^2, 则 B 按照命题 8.1.6 的方式对应到反对称矩阵

$$\boldsymbol{J}_1 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix};$$

这种二维辛空间也称为**双曲辛平面**.

定义 8.7.2 设 (V, B) 为辛空间.

- ⋆ 设 V_0 为 V 的子空间, 若 $B|_{V_0 \times V_0}$ 恒为零, 则称 V_0 是**全迷向**的;
- ⋆ 极大全迷向子空间 L 称为 V 的 **Lagrange 子空间**, 这相当于说 L 全迷向, 而任何严格包含 L 的子空间都不是全迷向的.

基于维数的理由, 任何全迷向子空间总是包含于某个 Lagrange 子空间; 特别地, Lagrange 子空间总存在, 它们是理解辛空间结构的钥匙.

引理 8.7.3 设 (V, B) 为辛空间, L 为 Lagrange 子空间, 则有 $L^\perp = L$ 和 $\dim V = 2 \dim L$.

证明 先证 $L^\perp = L$. 全迷向性质等价于 $L \subset L^\perp$. 假若存在 $v \in L^\perp \smallsetminus L$, 则 $L + \langle v \rangle$ 仍是全迷向的, 而且严格包含 L, 与极大性质矛盾.

其次, 命题 8.2.17 说明 $\dim L^\perp = \dim V - \dim L$, 代入 $L^\perp = L$ 立得 $\dim V = 2 \dim L$. □

作为推论, 辛空间必是偶数维的.

定理 8.7.4 (J.-G. Darboux) 设 (V, B) 为辛空间, L 为 Lagrange 子空间, 则 L 的任何有序基 $p_1, \dots, p_n$ 都能扩充为 V 的有序基

$$p_1, \dots, p_n, q_1, \dots, q_n,$$

使得对一切 $1 \leq i, j \leq n$ 皆有

$$B(p_i, p_j) = B(q_i, q_j) = 0,$$
$$B(p_i, q_j) = -B(q_j, p_i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

满足上述条件的有序基 $p_1, \dots, q_n$ 也称为辛空间 V 的**辛基**.

证明 以下提供一则初等论证, 对偶空间理论能给出更简短的证明, 详见本章习题的勾勒.

对 $i = 1, \dots, n$ 定义 $n-1$ 维子空间 $L_i := \langle p_1, \dots, \widehat{p_i}, \dots, p_n \rangle$, 其中 $\widehat{\cdot}$ 代表省略该项. 从 $L_i \subset L$ 连同命题 8.2.17 得到

$$L_i^\perp \supset L^\perp = L, \quad \dim L_i^\perp = 2n - \dim L_i = n + 1.$$

现在逐步构造辛基. 兹断言对所有 $1 \le k \le n$, 存在 $q_1, \dots, q_k \in V$ 使得

$$1 \le i, j \le k, \implies B(q_i, q_j) = 0,$$
$$1 \le i \le n,\ 1 \le j \le k \implies B(p_i, q_j) = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases}$$

首先是 $k = 1$ 情形. 存在 $q_1 \in L_1^\perp \smallsetminus L$, 这相当于说 $B(p_1, q_1) \ne 0$ 而 $i > 1 \implies B(p_i, q_1) = 0$, 适当伸缩后可确保 $B(p_1, q_1) = 1$. 条件满足.

今假定 $k < n$, 而且已有 $q_1, \dots, q_k$ 使条件成立. 基于 $L_{k+1}^\perp \supsetneq L$, 依然可取 q'_{k+1} 使得

$$B(p_i, q'_{k+1}) = \begin{cases} 1, & i = k+1, \\ 0, & i \ne k+1 \end{cases}$$

对所有 $1 \le i \le n$ 成立. 若将 q'_{k+1} 代换为

$$q_{k+1} := q'_{k+1} + a_1 p_1 + \cdots + a_k p_k, \quad a_1, \dots, a_k \in F$$

上述等式不受影响; 取 $a_i := B(q_i, q'_{k+1})$ 便能确保 $B(q_1, q_{k+1}) = \cdots = B(q_{k+1}, q_{k+1}) = 0$. 这就完成了第 $k+1$ 步构造.

取 $k = n$ 给出满足所需等式的 $q_1, \dots, q_n$. 最后来验证 $p_1, \dots, q_n$ 线性无关, 因而给出 V 的基. 设 $\sum_{i=1}^n a_i p_i + \sum_{i=1}^n b_i q_i = 0$. 对所有 i, 对等式两边取 $B(\cdot, q_i)$ 得到 $a_i = 0$, 取 $B(p_i, \cdot)$ 得到 $b_i = 0$. 明所欲证. □

辛基的存在性相当于断言所有辛空间 (V, B) 都同构于 $n := \frac{\dim V}{2}$ 份双曲辛平面的直和, 第 i 份对应到子空间 $\langle p_i, q_i \rangle$. 这就导向下述结果.

推论 8.7.5 *对于所有 $n \in \mathbb{Z}_{\ge 0}$, 存在唯一的 $2n$ 维辛空间, 精确到同构.*

练习 8.7.6 设 (V, B) 为辛空间. 满足 $W \cap W^\perp = \{0\}$ 的子空间 W 称为非退化的, 也称为**辛子空间**. 试说明

(i) W 是辛子空间当且仅当 B 在 W 上的限制非退化.

(ii) 若 W 是辛子空间, 则有直和分解 $V = W \oplus W^\perp$, 而且 $W^\perp$ 也是辛子空间.

实践表明, 将辛基排序为

$$p_1, p_2, \dots, p_n, q_n, \dots, q_2, q_1$$

往往更加顺手; 此时对应的辛形式便对应到 $2n \times 2n$ 反对称矩阵

$$\boldsymbol{J}_{2n} := \begin{pmatrix} & & & & & & 1 \\ & & & & & \iddots & \\ & & & & 1 & & \\ & & & -1 & & & \\ & & \iddots & & & & \\ -1 & & & & & & \end{pmatrix}. \tag{8.7.1}$$

现在放下非退化的条件, 回到本节开头的问题框架.

推论 8.7.7 *考虑资料 (V_i, B_i), 其中 B_i 反对称 $(i = 1, 2)$, 则 (V_1, B_1) 同构于 (V_2, B_2) 当且仅当*

$$\dim V_1 = \dim V_2, \quad \dim R(V_1) = \dim R(V_2).$$

证明 若 (V_1, B_1) 同构于 (V_2, B_2), 则当然有向量空间的同构 $V_1 \simeq V_2$ 以及 $R(V_1) \simeq R(V_2)$. 重点在 "仅当" 方向.

设断言中的条件成立, 命题 8.3.6 表明存在直和分解

$$(V_i, B_i) = (R(V_i), 0) \oplus \underbrace{(K_i, B_i)}_{\text{非退化}}, \quad i = 1, 2.$$

从 $\dim R(V_1) \simeq \dim R(V_2)$ 立得 $(R(V_1), 0) \simeq (R(V_2), 0)$. 问题遂化约为证辛空间的同构 $(K_1, B_1) \simeq (K_2, B_2)$. 既然 $\dim K_i = \dim V_i - \dim R(V_i)$, 从而有 $\dim K_1 = \dim K_2$, 应用推论 8.7.5 便完成证明. □

8.8 双重对偶

对于一般的域 F 上的向量空间 V, 对偶空间 $V^\vee$ 和 V 本身难以在不取基的前提下建立自然的联系, 但 V 和双重对偶 $(V^\vee)^\vee$ 却可以自然地加以比较. 这一现象和双线性形式的语言不无关系.

首先, $V^\vee$ 的元素既然是从 V 到 F 的映射, 自然可以在任意 $v \in V$ 处求值. 当 v 固定, 这便给出求值映射

$$\begin{aligned} \mathrm{ev}_v = \langle \cdot, v \rangle &: V^\vee \to F \\ \mathrm{ev}_v(\check{v}) &= \langle \check{v}, v \rangle, \end{aligned}$$

其中 $\langle \cdot, \cdot \rangle : V^\vee \times V \to F$ 是例 8.1.4 引进的典范配对. 既然 $\langle \cdot, \cdot \rangle$ 是双线性的, 故 ev_v 是线性映射, 而且

$$\mathrm{ev}_{v_1+v_2} = \mathrm{ev}_{v_1} + \mathrm{ev}_{v_2}, \quad \mathrm{ev}_{tv} = t\,\mathrm{ev}_v,$$

这些都是 $(V^\vee)^\vee = \mathrm{Hom}(V^\vee, F)$ 中的等式, 其中 $t \in F$, $v_1, v_2 \in V$.

命题 8.8.1 求值运算按上述方式给出自然的线性映射

$$\begin{aligned} \mathrm{ev}_V : V &\longrightarrow (V^\vee)^\vee \\ v &\longmapsto \mathrm{ev}_v. \end{aligned}$$

若 V 有限维, 则 ev_V 为同构. 对于一般的 V, 它总是单射.

证明 此前的讨论已经说明 ev_V 的确是线性映射. 设 $v \in V$ 非零. 当 V 有限维时, 存在基 $v_1, \dots, v_n$ 使得 $v_1 = v$. 取对偶基 $\check{v}_1, \dots, \check{v}_n$, 则 $\mathrm{ev}_v \in (V^\vee)^\vee$ 在 $\check{v}_1 \in V^\vee$ 上的取值是 $\langle \check{v}_1, v\rangle = 1$. 这就表明 $\ker(\mathrm{ev}_V) = \{0\}$. 既然 $\dim V^\vee = \dim V$, 故 ev_V 必为同构.

上一段不过是复述例 8.2.2 的论证, 但证明 $\ker(\mathrm{ev}_V) = \{0\}$ 的策略同样适用于一般的 V, 前提是述而未证的定义–命题 4.4.10: 任何非零的 $v \in V$ 都包含于 V 的一组基 B. 尽管无穷维空间不再有对偶基, 但总是能定义 $\check{v} \in V^\vee$ 使得对一切有限和 $\sum_{b\in B} c_b b \in V$ 皆有

$$\left\langle \check{v}, \sum_{b\in B} c_b b \right\rangle = c_v.$$

这就说明 $\mathrm{ev}_V(v) = \mathrm{ev}_v$ 在 $\check{v}$ 上取值为 1. 明所欲证. □

在 §4.10 结尾已经说明当 V 无穷维时, 对偶空间 $V^\vee$ 总是比 V 大得多, 因而 ev_V 不可能是同构. 尽管如此, 我们可以探讨嵌入 ev_V 和关于双重对偶运算的种种是否兼容, 例如线性映射 T 的双重转置 ${}^{\mathrm{t}}({}^{\mathrm{t}}T)$. 答案不出所料: 应容尽容. 详述如下.

命题 8.8.2 对任意线性映射 $T : V \to W$, 下图是约定 2.3.3 意义下的交换图表:

$$\begin{array}{ccc} V & \xrightarrow{\ T\ } & W \\ {\scriptstyle \mathrm{ev}_V}\downarrow & & \downarrow{\scriptstyle \mathrm{ev}_W} \\ V^{\vee\vee} & \xrightarrow[{}^{\mathrm{t}}({}^{\mathrm{t}}T)]{} & W^{\vee\vee}. \end{array}$$

证明 设 $v \in V$ 而 $\check{w} \in W^\vee$, 小心翼翼地展开定义, 得到

$$\left\langle \underbrace{{}^{\mathrm{t}}({}^{\mathrm{t}}T)\,(\mathrm{ev}_V(v))}_{\in W^{\vee\vee}}, \underbrace{\check{w}}_{\in W^\vee} \right\rangle = \left\langle \underbrace{\mathrm{ev}_V(v)}_{\in V^{\vee\vee}}, \underbrace{{}^{\mathrm{t}}T(\check{w})}_{\in V^\vee} \right\rangle$$
$$= \langle {}^{\mathrm{t}}T(\check{w}), v\rangle = \langle \check{w}, Tv\rangle = \langle \mathrm{ev}_W(Tv), \check{w}\rangle .$$

由于 $\check{w}$ 可任取, ${}^{\mathrm{t}}({}^{\mathrm{t}}T)(\mathrm{ev}_V(v)) = \mathrm{ev}_W(Tv)$. 由于 v 可任取, ${}^{\mathrm{t}}({}^{\mathrm{t}}T) \circ \mathrm{ev}_V = \mathrm{ev}_W \circ T$. 图表交换性得证. □

其次, 对有限维向量空间的有序基也可以取双重对偶. 它们同样有简单的描述.

命题 8.8.3 设 V 是有限维向量空间, $v_1,\dots,v_n \in V$ 为有序基, 其对偶基记为 $\check{v}_1,\dots,\check{v}_n \in V^\vee$, 则 $V^{\vee\vee}$ 的元素

$$\mathrm{ev}_V(v_1),\dots,\mathrm{ev}_V(v_n)$$

是 $\check{v}_1,\dots,\check{v}_n$ 的对偶基.

证明 回忆到有限维情形有 $\dim V = \dim V^\vee = \dim V^{\vee\vee}$. 对所有 $1 \le i,j \le n$,

$$\left\langle \underbrace{\mathrm{ev}_V(v_i)}_{\in V^{\vee\vee}},\ \underbrace{\check{v}_j}_{\in V^\vee} \right\rangle = \langle \check{v}_j, v_i\rangle = \begin{cases} 1, & i=j,\\ 0, & i\neq j. \end{cases}$$

此即对偶基在 $V^{\vee\vee}$ 中的刻画. □

以下性质适用于任意向量空间及其对偶, 证明也不外是按定义逐步验证.

命题 8.8.4 设 V 和 W 为向量空间. 考虑命题 8.1.5 中的同构

$$\mathrm{Hom}(V,W^\vee) \xrightarrow{\sim} \mathrm{Bil}(V,W;F) \xrightarrow{\sim} \mathrm{Hom}(W,V^\vee),$$

则它们的合成映 $\psi \in \mathrm{Hom}(V,W^\vee)$ 为 $W \xrightarrow{\mathrm{ev}_W} (W^\vee)^\vee \xrightarrow{{}^{\mathrm{t}}\psi} V^\vee$ 的合成 ${}^{\mathrm{t}}\psi \circ \mathrm{ev}_W$. 类似地,

$$\mathrm{Hom}(W,V^\vee) \xrightarrow{\sim} \mathrm{Bil}(V,W;F) \xrightarrow{\sim} \mathrm{Hom}(V,W^\vee)$$

的合成映 $\varphi \in \mathrm{Hom}(W,V^\vee)$ 为 ${}^{\mathrm{t}}\varphi \circ \mathrm{ev}_V$.

证明 说明 $\varphi \in \mathrm{Hom}(W,V^\vee)$ 的情形即可. 按定义, φ 在 $\mathrm{Bil}(V,W;F)$ 中的像是 $B(v,w) = \langle \varphi(w), v\rangle$, 而后者在 $\mathrm{Hom}(V,W^\vee)$ 中的像映任意 $v \in V$ 为 $B(v,\cdot) = \langle \varphi(\cdot), v\rangle$.

现在考虑 ${}^{\mathrm{t}}\varphi \circ \mathrm{ev}_V$. 展开定义, 它映任意 $v \in V$ 为这样一个从 W 到 F 的线性映射:

$$w \mapsto \left\langle \underbrace{\mathrm{ev}_V(v)}_{\in (V^\vee)^\vee}, \underbrace{\varphi(w)}_{\in V^\vee} \right\rangle, \quad w \in W.$$

然而按照 $\mathrm{ev}_V(v)$ 的定义, 此即 $w \mapsto \langle \varphi(w), v\rangle$. 证毕. □

设 V 为有限维 F-向量空间. 我们已经知道它和对偶空间之间虽然存在同构 $V \simeq V^\vee$, 但同构依赖于有序基的选取; 基的选取既然任意, 同构 $V \simeq V^\vee$ 没有自然的选法, 除非对 V 再指定其他资料.

另一方面, 关于双重对偶的同构 $\mathrm{ev}_V : V \simeq (V^\vee)^\vee$ 却是**自然**的, 或者称为**典范**的. 这个形容词在代数学中处处可见, 至少能按两种方式理解. 且以 ev_V 为例.

1. 它不依赖基的选取; 事实上, 我们直接写下了它的公式 $\mathrm{ev}_V(v) = \langle \cdot, v\rangle : V^\vee \to F$, 不需要诸如有序基之类的任何辅助资料.

2. 它和线性映射是兼容的, 或者用数学工作者的术语来说, 它满足 “函子性”, 明确意义通过命题 8.8.2 的交换图表表述.

两种理解方式内涵相异却彼此相关. 在代数学的研究中, 一条经验法则是: 只要能合理地, 自然地[4]定义代数结构之间的一类映射, 使得定义不依赖带有任意性的辅助资料, 则这些映射通常也具备体现为种种交换图表的函子性 — 至少通常如此.

另外一个例子则是命题 8.1.5, 其中的所有同构都是典范的, 不依赖基的选取.

基于这些考量, 我们尔后将继续使用自然映射, 典范映射之类的术语, 并且在必要时写下所涉及的交换图表.

8.9 对偶与商

现在可以结合先前介绍的种种概念, 以求更透彻地理解向量空间的对偶. 稍加具体地说, 眼下的主题是比较线性映射 $T: V \to W$ 及其转置 ${}^{\mathrm{t}}T: W^\vee \to V^\vee$ 的性质.

引理 8.9.1 (单及满相对偶) *设 $T: V \to W$ 为线性映射. 若 T 单则 ${}^{\mathrm{t}}T$ 满, 若 T 满则 ${}^{\mathrm{t}}T$ 单.*

证明 设 T 单. 任取 V 的基 $\mathcal{X}$, 则 $T(\mathcal{X})$ 是 W 的线性无关子集. 用定义–命题 4.4.10 将 $T(\mathcal{X})$ 扩充为 W 的基 $\mathcal{Y}$. 由于任意 $w \in W$ 都有唯一的有限和展开 $\sum_{y\in\mathcal{Y}} c_y y$, 其中 $c_y \in F$, 对给定的 $\mu \in V^\vee$ 可定义 $\tilde{\mu}: W \to F$ 为

$$\tilde{\mu}\left(\sum_{y\in\mathcal{Y}} c_y y\right) = \mu\left(\sum_{x\in\mathcal{X}} c_{T(x)} x\right).$$

显然 $\tilde{\mu} \in W^\vee$ 而 ${}^{\mathrm{t}}T(\tilde{\mu})$ 映 $x \in \mathcal{X}$ 为 $\tilde{\mu}(T(x)) = \mu(x)$. 由此见 ${}^{\mathrm{t}}T(\tilde{\mu}) = \mu$. 既然 μ 可任取, 故 ${}^{\mathrm{t}}T$ 满.

设 T 满而 $\tilde{\mu} \in W^\vee$ 满足 ${}^{\mathrm{t}}T(\tilde{\mu}) := \tilde{\mu}T = 0$, 则 T 的满性蕴涵 $\tilde{\mu} = 0$, 故 ${}^{\mathrm{t}}T$ 单. □

引理 8.9.2 *设线性映射 $T: V \to W$ 满, 而 $\mu \in V^\vee$, 则 $\mu|_{\ker(T)} = 0$ 当且仅当存在 $\tilde{\mu} \in W^\vee$ 使得 ${}^{\mathrm{t}}T(\tilde{\mu}) = \mu$.*

证明 对于 “仅当” 方向, 关键是建立以下交换图表:

$$\begin{array}{ccc} V & \xrightarrow{\ T\ } & W \\ {\scriptstyle\mu}\downarrow & \searrow^{s} & \wr\uparrow{\scriptstyle\overline{T}} \\ F & \xleftarrow[\ \overline{\mu}\ ]{} & V/\ker(T) \end{array} \qquad s:\ \text{商映射}.$$

由于 $\mathrm{im}(T) = W$, 上半部的交换三角和同构 $\overline{T}$ 来自命题 4.12.7 (ii). 由于 $\mu|_{\ker(T)} = 0$, 下半部的交换三角和 $\overline{\mu}$ 来自命题 4.12.7 (i).

[4] 按日常语言理解.

现在取 $\tilde{\mu} := \overline{\mu} \circ \overline{T}^{-1} : W \to F$, 交换图表遂给出

$$\tilde{\mu} T = \overline{\mu} \circ \overline{T}^{-1} T = \overline{\mu} s = \mu.$$

至于 "当" 的方向, ${}^{\mathrm{t}}T(\tilde{\mu}) = \tilde{\mu} T$, 而 $(\tilde{\mu} T)|_{\ker(T)} = 0$ 实属显然. □

回忆到除了 V 的子空间 $\ker(T)$, 称为 T 的核, 从 T 还能构造 W 的商空间 $\operatorname{coker}(T)$, 称为 T 的余核 (定义 4.12.5). 它们带有的包含映射和商映射分别记为

$$i : \ker(T) \hookrightarrow V, \quad q : W \twoheadrightarrow \operatorname{coker}(T);$$

在对偶空间方面, 我们便得到 ${}^{\mathrm{t}}i$ 和 ${}^{\mathrm{t}}q$.

定理 8.9.3 (核及余核相对偶) *给定线性映射 $T : V \to W$, 考虑相应的*

$$\begin{array}{ccccc} \operatorname{coker}(T)^\vee & \xrightarrow{{}^{\mathrm{t}}q} & W^\vee & \xrightarrow{{}^{\mathrm{t}}T} & V^\vee & \xrightarrow{{}^{\mathrm{t}}i} & \ker(T)^\vee \\ \lambda & \longmapsto & \lambda q & & & & \\ & & \tilde{\mu} & \longmapsto & \tilde{\mu} T & & \\ & & & & \mu & \longmapsto & \mu i. \end{array}$$

借此, $\operatorname{coker}(T)^\vee$ 等同于 $\ker({}^{\mathrm{t}}T)$, 而 $\ker(T)^\vee$ 等同于 $\operatorname{coker}({}^{\mathrm{t}}T)$. 更精确地说, 我们有交换图表

$$\begin{array}{ccc} \operatorname{coker}(T)^\vee & \xrightarrow{{}^{\mathrm{t}}q} & W^\vee \\ \downarrow \wr & \nearrow & \\ \ker({}^{\mathrm{t}}T) & & \end{array} \qquad \begin{array}{ccc} V^\vee & \xrightarrow{{}^{\mathrm{t}}i} & \ker(T)^\vee \\ & \searrow & \uparrow \wr \\ & & \operatorname{coker}({}^{\mathrm{t}}T) \end{array} \tag{8.9.1}$$

其中垂直方向的同构由左右两个交换图表各自唯一确定.

证明 因为 q 满而 i 单, ${}^{\mathrm{t}}q$ 的单性和 ${}^{\mathrm{t}}i$ 的满性皆来自引理 8.9.1.

关于垂直方向同构的存在性, 我们从 T 是单射或满射的简单情形入手.

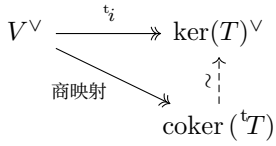
首先假定 T 是单射, 此时 $\ker(T) = \{0\}$ 而引理 8.9.1 蕴涵 ${}^{\mathrm{t}}T$ 满. 右边图表化为废话; 对于左边图表, 唯一待证的是 $\operatorname{coker}(T)^\vee \hookrightarrow W^\vee$ 的像正好是 $\ker({}^{\mathrm{t}}T)$. 考虑 $\lambda \in \operatorname{coker}(T)^\vee$, 从 $qT = 0$ 可得 ${}^{\mathrm{t}}T(\lambda q) = \lambda q T = 0$, 这给出一边的包含关系. 至于另一边, 设 $\tilde{\mu} \in W^\vee$ 满足 ${}^{\mathrm{t}}T(\tilde{\mu}) = \tilde{\mu} T = 0$, 则命题 4.12.7 (i) 给出线性映射

$$\begin{aligned} \lambda : \operatorname{coker}(T) = W / \operatorname{im}(T) &\to F \\ w + \operatorname{im}(T) &\mapsto \tilde{\mu}(w), \end{aligned}$$

而这正好说明 $\lambda \in \operatorname{coker}(T)^\vee$ 满足 $\lambda q = \tilde{\mu}$.

其次假定 T 是满射, 此时 $\operatorname{coker}(T) = \{0\}$ 而引理 8.9.1 蕴涵 ${}^{\mathrm{t}}T$ 单. 左边图表化为废话; 对于右边图表, 唯一待证的是 $V^\vee \twoheadrightarrow \ker(T)^\vee$ 可以按断言的方式等同于

$\operatorname{coker}({}^{\mathrm{t}}T)$. 请端详下图实线部分:



(8.9.2)

引理 8.9.2 相当于说 $\ker({}^{\mathrm{t}}i) = \operatorname{im}({}^{\mathrm{t}}T)$. 代入命题 4.12.7 (ii) 可见确实存在唯一的虚线所示同构, 使得全图交换. 这就说明 $V^\vee \twoheadrightarrow \ker(T)^\vee$ 按断言的方式等同于 $\operatorname{coker}({}^{\mathrm{t}}T)$.

现在考虑一般的 T, 将 T 作满–单分解 $V \xrightarrow{T'} U \xhookrightarrow{T''} W$, 例如取 $U = \operatorname{im}(T)$ 而 T'' 是包含映射. 命题 4.7.7 和引理 8.9.1 相应地给出 ${}^{\mathrm{t}}T$ 的满–单分解 $W^\vee \xrightarrow{{}^{\mathrm{t}}(T'')} U^\vee \xhookrightarrow{{}^{\mathrm{t}}(T')} V^\vee$. 如此则

$$\begin{aligned} \ker(T) = \ker(T'), \quad \operatorname{coker}(T) = \operatorname{coker}(T''), \\ \ker({}^{\mathrm{t}}T) = \ker({}^{\mathrm{t}}(T'')), \quad \operatorname{coker}({}^{\mathrm{t}}T) = \operatorname{coker}({}^{\mathrm{t}}(T')). \end{aligned}$$

由此容易化约到先前处理的特例.

最后, 观察到使 (8.9.1) 的两个图表交换的垂直箭头都是唯一的 (存在性业已说明): 这分别来自单射和满射对映射合成的消去律. □

对于一切 F-向量空间 W 及其子空间 V, 定义 $W^\vee$ 的子空间

$$^\perp V := \{\check{w} \in W^\vee : \forall v \in V,\ \langle \check{w}, v\rangle = 0\};$$

这也是定义 8.1.12 对典范配对 $\langle\cdot,\cdot\rangle : W^\vee \times W \to F$ 的应用.

推论 8.9.4 *设 V 是 W 的子空间, 则有典范同构*

$$\begin{aligned} W^\vee/{}^\perp V &\xrightarrow{\sim} V^\vee & (W/V)^\vee &\xrightarrow{\sim} {}^\perp V \\ \tilde{\mu} + {}^\perp V &\longmapsto \tilde{\mu}|_V. & \lambda &\longmapsto [w \mapsto \lambda(w+V)] \end{aligned}$$

证明 命 $T : V \hookrightarrow W$ 为包含映射, 则 ${}^{\mathrm{t}}T : W^\vee \twoheadrightarrow V^\vee$ 映 $\tilde{\mu}$ 为 $\tilde{\mu}|_V$, 而且这是满射. 代入之前的定义, 立见 $\ker({}^{\mathrm{t}}T) = {}^\perp V$. 应用命题 4.12.7 (ii) 可知 $\tilde{\mu} + {}^\perp V \mapsto \tilde{\mu}|_V$ 给出同构 $W^\vee/{}^\perp V \xrightarrow{\sim} V^\vee$. 此即第一式.

至于第二式, 商映射 $q : W \to \operatorname{coker}(T) = W/V$ 映 w 为 $w+V$, 故定理 8.9.3 的映射列的左段写作

$$\begin{aligned} (W/V)^\vee &\xhookrightarrow{{}^{\mathrm{t}}q} W^\vee \\ \lambda &\longmapsto [w \mapsto \lambda(w+V)] \end{aligned}$$

而且它的像正是 $\ker({}^{\mathrm{t}}T) = {}^\perp V$. 这就说明 ${}^{\mathrm{t}}q$ 给出同构 $(W/V)^\vee \xrightarrow{\sim} {}^\perp V \subset W^\vee$. □

定理 8.9.3 可用来比较线性映射 T 及 ${}^{\mathrm{t}}T$ 的像.

定理 8.9.5 (像自对偶) 设 $T : V \to W$ 是线性映射, 则有典范同构

$$\operatorname{im}(T)^\vee \xrightarrow{\sim} \operatorname{im}({}^{\mathrm{t}}T).$$

证明 根据 (4.12.1), 我们有 $\operatorname{im}(T) = \ker\left[q : W \xrightarrow{\text{商映射}} \operatorname{coker}(T)\right]$. 先后对 q 和 T 应用定理 8.9.3 可得

$$\begin{aligned}\operatorname{im}(T)^\vee &\simeq \operatorname{coker}\left[{}^{\mathrm{t}}q : \operatorname{coker}(T)^\vee \to W^\vee\right]\\ &= \operatorname{coker}\left[\ker({}^{\mathrm{t}}T) \hookrightarrow W^\vee\right]\\ &= W^\vee / \ker({}^{\mathrm{t}}T).\end{aligned}$$

然而根据命题 4.12.7 (ii), 末项典范地同构于 $\operatorname{im}({}^{\mathrm{t}}T)$. □

本节迄今得到的结论适用于所有向量空间. 但最常用的是有限维的情形.

推论 8.9.6 设 $T : V \to W$ 是有限维向量空间之间的线性映射, 则 $\dim \operatorname{im}(T) = \dim \operatorname{im}({}^{\mathrm{t}}T)$.

证明 此时 $\operatorname{im}(T)$ 是有限维的. 对定理 8.9.5 的同构两边取维数, 得 $\dim \operatorname{im}(T)^\vee = \dim \operatorname{im}({}^{\mathrm{t}}T)$, 然而定义–命题 4.7.8 表明 $\dim \operatorname{im}(T)^\vee = \dim \operatorname{im}(T)$. □

如果取定 V 和 W 的有序基, 将 T 等同于矩阵, 则 $\dim \operatorname{im}(T)$ 无非是矩阵的列秩, $\dim \operatorname{im}({}^{\mathrm{t}}T)$ 则是其转置的列秩, 即行秩. 这就重新证明了定理 4.9.11:

$$\text{矩阵的列秩} = \text{行秩}.$$

以上论证不涉及任何具体的矩阵操作, 因此也不依赖于定理 4.9.11 的原证. 尽管优美与否见仁见智, 但定理 8.9.5 确实具有比 "行秩 = 列秩" 严格高出一个档次的威力, 因为它不仅断言两个数相等, 而是给出两个向量空间之间的典范同构, 前者不过是后者取 dim 之后的简单结论.

习题

1. 设 V 是有限维 F-向量空间, $B : V \times V \to F$ 是非退化双线性形式. 命 $n := \dim V$. 证明 $v_1, \dots, v_n \in V$ 线性无关当且仅当 $(B(v_i, v_j))_{1 \le i,j \le n}$ 是可逆 $n \times n$ 矩阵.
2. 设 V 和 W 是 n 维 F-向量空间, $B : V \times W \to F$ 是非退化双线性形式. 证明对于任意有序基 $w_1, \dots, w_n \in W$, 存在唯一的有序基 $v_1, \dots, v_n \in V$ 使得

   $$B(v_i, w_j) = \begin{cases} 1, & i = j,\\ 0, & i \ne j \end{cases}$$

   说明这与定义–命题 4.7.8 的对偶基有何联系.

3. 参照 §8.5 介绍的配方法, 将以下 $\mathbb{Q}$ 上的二次型对角化, 并具体写下所用的变换.

(i) $f = X_1^2 - 2X_2^2 + X_3^2 - 2X_1X_2 + 4X_2X_3$.

(ii) $f = X_1X_2 - 2X_2X_3 + 3X_3X_4$.

(iii) $f = X_1X_2 + X_1X_3 + X_2X_3$

4. 设 F 为域, $\mathrm{char}(F) \neq 2$, 而 $a, b \in F$ 满足 $a + b \neq 0$. 证明 F 上的二元二次型

$$aX_1^2 + bX_2^2 \quad \text{和} \quad (a+b)X_1^2 + ab(a+b)X_2^2$$

相互同构.

5. 说明有限域 $\mathbb{F}_5 := \mathbb{Z}/5\mathbb{Z}$ 上的二元二次型 X_1X_2 和 $X_1^2 + X_2^2$ 同构.

6. 在实数域 $\mathbb{R}$ 上, 将 $n \times n$ 对称矩阵

$$\boldsymbol{A} = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} \quad \text{(留白部分为零)}$$

所对应的二次型化到规范形, 并确定其符号差.

7. 设 $\boldsymbol{A}$ 是 $n \times n$ 实对称矩阵. 证明若有列向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathbb{R}^n$ 使得 ${}^{\mathrm{t}}\boldsymbol{\alpha}_1\boldsymbol{A}\boldsymbol{\alpha}_1 > 0$ 而 ${}^{\mathrm{t}}\boldsymbol{\alpha}_2\boldsymbol{A}\boldsymbol{\alpha}_2 < 0$, 则存在非零之 $\boldsymbol{\alpha}_3 \in \mathbb{R}^n$ 使得 ${}^{\mathrm{t}}\boldsymbol{\alpha}_3\boldsymbol{A}\boldsymbol{\alpha}_3 = 0$.

8. 对于域 F 上的向量空间 V, 线性映射 $L : V \to F$ 也称为 V 上的一次型或线性型. 在 $\mathrm{char}(F) \neq 2$ 的前提下, 说明若 L, M 是 V 上的一次型, 则 $LM : v \mapsto L(v)M(v)$ 给出 V 上的二次型, 并尽可能精确地给出对称双线性形式 $B : V \times V \to F$ 使得 $L(v)M(v) = B(v, v)$.

9. 证明一个不恒为零的 n 元实二次型 $f \in \mathbb{R}[X_1, \ldots, X_n]$ 分解成两个一次型 (见上题) 的乘积当且仅当以下任一条件成立:

- ⋆ 对应的对称矩阵秩为 2, 并且二次型的符号差为 0;
- ⋆ 对应的对称矩阵秩为 1.

10. 考虑 $\mathbb{R}$-向量空间 $V = \mathbb{R}^n$ 及其上的二次型 f.

(i) 设 f 可以表成 $f_+ - f_-$, 其中 $f_\pm$ 都是半正定二次型, f_+ 的正惯性指数为 p 而 f_- 的正惯性指数为 q. 证明 f 的正惯性指数 $\leq p$, 负惯性指数 $\leq q$.

提示〉 记 f 的正惯性指数为 p'. 存在 p' 维子空间 $U' \subset V$ 使得 f 在 U' 上正定. 存在 $n - p$ 维子空间 U 使得 f_+ 在 U 上半负定. 假若 $p' > p$ 则

$$\dim U \cap U' = \dim U + \dim U' - \dim(U + U') > p + (n - p) - n = 0;$$

取其中的非零向量 v 便有 $0 < f(v) \leq f_+(v) \leq 0$, 矛盾.

(ii) 证明若二次型 f 可以表成

$$f = L_1^2 + \cdots + L_p^2 - L_{p+1}^2 - \cdots - L_{p+q}^2$$

其中 $L_1, \ldots, L_{p+q}$ 是 $\mathbb{R}^n$ 上的一次型, $p, q \in \mathbb{Z}_{\geq 0}$, 则 f 的正惯性指数 $\leq p$, 负惯性指数 $\leq q$.

11. 设对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 在 $\mathbb{R}^n$ 上给出的二次型是正定的 (称之为对称正定矩阵). 证明 n 元二次型

$$f(X_1, \ldots, X_n) = \det \left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{x} \\ \hline {}^{\mathrm{t}}\boldsymbol{x} & 0 \end{array}\right)_{(n+1)\times(n+1)}, \quad x = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

是负定的. 提示〉 设法化到 $\boldsymbol{A}$ 为规范形的情形.

12. 设 $\boldsymbol{A} = (a_{ij})_{1\leq i,j\leq n} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 为正定对称矩阵. 证明

$$\det \boldsymbol{A} \leq a_{11} \cdots a_{nn}.$$

此外, 试明确等号成立的充要条件.

提示〉 设 $n \geq 2$. 命 $\boldsymbol{A}^\flat := (a_{ij})_{1\leq i,j\leq n-1}$, 留意到 $\boldsymbol{A}^\flat$ 仍是正定对称矩阵, 而且

$$\det \boldsymbol{A} = a_{nn} \det \boldsymbol{A}^\flat + \det \left(\begin{array}{ccc|c} & & & a_{1n} \\ & \boldsymbol{A}^\flat & & a_{2n} \\ & & & \vdots \\ \hline a_{n1} & a_{n2} & \cdots & 0 \end{array}\right)$$

然后配合上一题推得 $\det \boldsymbol{A} \leq a_{nn} \det \boldsymbol{A}^\flat$, 递归地论证.

13. 设 L_1, L_2 是辛空间 (V, B) 的 Lagrange 子空间. 说明存在 (V, B) 的自同构 φ (换言之, φ 是满足 $B(\varphi(v), \varphi(v')) = B(v, v')$ 的可逆线性映射 $V \xrightarrow{\sim} V$), 使得 $L_2 = \varphi(L_1)$.

提示〉 对 $i = 1, 2$ 取 L_i 的基 $p_1^{(i)}, \ldots, p_n^{(i)}$, 扩充为辛基 $p_1^{(i)}, \ldots, q_n^{(i)}$. 定义 $\varphi \in \mathrm{End}(V)$ 使得 $\varphi\left(p_j^{(1)}\right) = p_j^{(2)}$ 而 $\varphi\left(q_j^{(1)}\right) = q_j^{(2)}$; 说明 φ 是 (V, B) 的自同构.

14. 设 (V, B) 为辛空间而 V_0 为全迷向子空间 (定义 8.7.2). 说明 B 自然地诱导 $V_0^\perp/V_0$ 上的辛形式.

15. 试以推论 8.9.4 为定理 8.7.4 给出更简短的证明.

提示〉 考虑 $v \mapsto B(\cdot, v)$ 给出的同构 $V \xrightarrow{\sim} V^\vee$. 验证 L 被映为 ${}^\perp L$, 而推论 8.9.4 的同构 $V^\vee/{}^\perp L \xrightarrow{\sim} L^\vee$ 是将每个 $B(\cdot, v) \in V^\vee$ 限制到 L 上. 现在取 $p_1, \ldots, p_n$ 的对偶基 $\check{p}_1, \ldots, \check{p}_n$, 再取 $q_1, \ldots, q_n$ 使得 $B(\cdot, q_i) \in V^\vee$ 映至 $\check{p}_i$. 最后逐步用 L 调整以确保 $B(q_i, q_j) = 0$ 恒成立.

第九章 实内积结构

向量的长度和夹角是几何学的基本概念. 在向量空间的抽象语言中, 两者统合为一个简单的代数结构, 称为内积. 内积分为实与复两种版本, 本章专注于实数域 $\mathbb{R}$ 上的情形.

实向量空间 V 上的内积是满足正定性的对称双线性形式 $V \times V \to \mathbb{R}$, 本书惯例记为 $(\cdot|\cdot)$; 对称意谓 $(v_1|v_2) = (v_2|v_2)$, 而正定意谓 $(v|v) \geq 0$ 且 $(v|v) = 0 \iff v = 0$. 实向量空间 V 配上内积 $(\cdot|\cdot)$ 所成的结构称为实内积空间, 本章简称为内积空间.

内积的标准范例来自 $\mathbb{R}^n$ 上的双线性形式 $(\boldsymbol{x}|\boldsymbol{y}) := \sum_{i=1}^n x_i y_i$, 也可以用列向量的观点表为 ${}^{\mathrm{t}}\boldsymbol{x}\,\boldsymbol{y}$. 向量 $\boldsymbol{x}$ 的长度以内积表作

$$\|\boldsymbol{x}\| := \sqrt{(\boldsymbol{x}|\boldsymbol{x})} = \sqrt{\sum_{i=1}^n x_i^2},$$

而非零向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的夹角 $\angle(\boldsymbol{x}, \boldsymbol{y}) \in [0, \pi]$ 则由

$$(\boldsymbol{x}|\boldsymbol{y}) = \|x\|\|y\| \cos\angle(\boldsymbol{x}, \boldsymbol{y})$$

刻画; 对于平面或空间向量, 以上两个等式分别是勾股定理和余弦定理的应用, 而在一般的内积空间中, 我们反过来以这些公式来定义向量的长度和夹角. 长度为 1 的向量称为单位向量, 内积为零的一对向量称为是正交的.

在 §§9.1–9.2, 我们将从 $\mathbb{R}^n$ 的标准内积出发, 引入内积空间的一般定义, 并从定义推导内积空间中的勾股定理 (命题 9.2.2), Cauchy–Bunyakovsky–Schwarz 不等式 (定理 9.2.3) 与三角不等式 (推论 9.2.4), 它们的几何意涵明朗, 但证明纯粹是代数的. 我们还将引入内积空间之间的保距线性映射和同构的概念.

在 §9.3, 我们将定义内积空间中的正交向量族与单位正交向量族, 然后证明 Gram–Schmidt 正交化定理 9.3.5; 它蕴涵任何有限维内积空间都有单位正交基, 从而同构于标准内积空间 $(\mathbb{R}^n, (\cdot|\cdot))$, 尽管同构的选法非唯一. 对有限维子空间的正交投影和正交直和分解也是该节的重头戏.

有限维内积空间之间的线性映射是 §§9.4–9.5 的主题. 基于伴随映射的概念, 我们将介绍何谓正交变换和自伴线性映射 (又称自伴算子), 并且提供矩阵诠释. 这部分最重要的结论是自伴算子或实对称矩阵的正交对角化 (定理 9.5.2), 以及称为主轴定理的应用 (定理 9.5.5); 后者涉及二次型, 由之容易得到 $\mathbb{R}^n$ 中的二次超曲面的完整分类.

本章后半部的 §§9.6–9.10 介绍内积空间上的的种种经典结论, 应用极为广泛, 其内容具体包括:

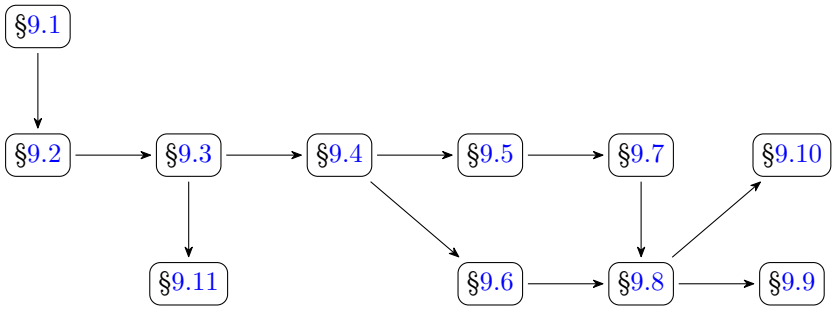
- ⋆ 线性方程组的最小二乘解 (定义 9.6.1) 及其刻画 (定理 9.6.3),
- ⋆ 关于正定二次型的顺序主子式判准 (定理 9.7.4), 半正定线性映射的平方根 (定义–命题 9.7.7) 和可逆线性映射的极分解 (定理 9.7.8),
- ⋆ 数据科学中常用的奇异值分解 (定理 9.8.1) 及其矩阵形式 (9.8.1),
- ⋆ Moore–Penrose 广义逆 (定义 9.9.2) 的存在和唯一性 (定理 9.9.3), 与基于奇异值分解的算法 (命题 9.9.4),
- ⋆ 极小化极大原理在有限维内积空间上的情形 (定理 9.10.3).

本章最后的 §9.11 只需要 §§9.1–9.3 的基础, 但技术性相对突出. 该节探讨非负矩阵的特征值和特征向量, 这类问题在应用场合自然地出现. 正文部分的主要结论是定理 9.11.7, 所需条件可以用有向图的概念来解释; 适用范围更广的定理 9.11.8 仅述而不证, 但本章习题将勾勒证明的步骤, 并提供提示.

阅读提示

若无另外说明, 本章的向量空间都默认为 $\mathbb{R}$-向量空间. 除了 §§9.1–9.3 和本章习题的一部分内容, 其余部分只处理有限维向量空间; §9.11 只涉及矩阵语言.

阅读顺序

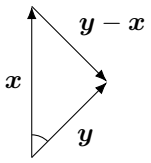


9.1 引言: 标准内积

在初等平面几何学中, 一旦选定原点和直角坐标系, 则可以将平面方便地等同于 $\mathbb{R}^2$. 平面向量 $\boldsymbol{x}=(x_1,x_2)$ 的长度可以写作

$$\|\boldsymbol{x}\| := \sqrt{x_1^2+x_2^2}.$$

现在考虑两个平面向量 $\boldsymbol{x},\boldsymbol{y}$. 暂且假设它们非零. 记其夹角为 $\angle(\boldsymbol{x},\boldsymbol{y})\in[0,\pi]$. 考虑它们的差 $\boldsymbol{y}-\boldsymbol{x}$: 若将这些向量适当平移, 使其起点前后衔接, 则构成如下图所示的三角形



当 $\boldsymbol{x},\boldsymbol{y}$ 皆非零时, 夹角和长度的关系由余弦定理确定:

$$\|\boldsymbol{y}-\boldsymbol{x}\|^2 = \|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2-2\|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\angle(\boldsymbol{x},\boldsymbol{y}).$$

这就启发我们考虑以下的量

$$\|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\angle(\boldsymbol{x},\boldsymbol{y}) = \frac{-1}{2}\left(\|\boldsymbol{y}-\boldsymbol{x}\|^2-\|\boldsymbol{x}\|^2-\|\boldsymbol{y}\|^2\right).$$

在上式中以 $-\boldsymbol{x}$ 代 $\boldsymbol{x}$, 则因为 $\|\boldsymbol{x}\|=\|-\boldsymbol{x}\|$, 而余弦函数的定义导致 $\cos\angle(-\boldsymbol{x},\boldsymbol{y}) = -\cos\angle(\boldsymbol{x},\boldsymbol{y})$, 整理便得到 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积, 其定义是

$$\begin{aligned}(\boldsymbol{x}|\boldsymbol{y}) &:= \|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\angle(\boldsymbol{x},\boldsymbol{y})\\ &= \frac{1}{2}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2-\|\boldsymbol{x}\|^2-\|\boldsymbol{y}\|^2\right).\end{aligned}$$

注意到 $(\boldsymbol{x}|\boldsymbol{y})$ 对所有 $\boldsymbol{x},\boldsymbol{y}$ 都有定义; 相对于坐标表法 $\boldsymbol{x}=(x_1,x_2)$, $\boldsymbol{y}=(y_1,y_2)$, 它有简单的代数表达式

$$\begin{aligned}(\boldsymbol{x}|\boldsymbol{y}) &= \frac{1}{2}\left((x_1+y_1)^2+(x_2+y_2)^2-(x_1^2+x_2^2)-(y_1^2+y_2^2)\right)\\ &= x_1y_1+x_2y_2.\end{aligned}$$

内积定义涉及向量的长度, 反过来, 内积又能统合长度和夹角这两种几何概念:

$$\begin{aligned}\|\boldsymbol{x}\|^2 &= (\boldsymbol{x}|\boldsymbol{x})\\ \cos\angle(\boldsymbol{x},\boldsymbol{y}) &= \frac{(\boldsymbol{x}|\boldsymbol{y})}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|},\quad \boldsymbol{x},\boldsymbol{y}\neq\boldsymbol{0}.\end{aligned}$$

推而广之, 赋予 n 维空间 (不妨想象 $n=3$) 直角坐标系, 将之等同于 $\mathbb{R}^n$. 在 Euclid 几何的框架下, 向量 $\boldsymbol{x}=(x_1,\dots,x_n)$ 的长度应当合理地定为

$$\|\boldsymbol{x}\|=\sqrt{\sum_{i=1}^{n}x_i^2}.$$

当 $n\geq 2$ 时, 任两个非零向量 $\boldsymbol{x},\boldsymbol{y}$ 都同时包含于一个二维子空间. 在此平面上用余弦定理计算向量的加法, 长度和夹角, 则同样的论证给出

$$\begin{aligned}\|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\angle(\boldsymbol{x},\boldsymbol{y})&=\frac{1}{2}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2-\|\boldsymbol{x}\|^2-\|\boldsymbol{y}\|^2\right)\\&=:(\boldsymbol{x}|\boldsymbol{y}).\end{aligned}$$

注意到这一公式对于 $n=1$ 的情形同样成立. 相对于坐标 $\boldsymbol{x}=(x_1,\dots,x_n)$ 和 $\boldsymbol{y}=(y_1,\dots,y_n)$ 计算 $\|\boldsymbol{x}+\boldsymbol{y}\|^2-\|\boldsymbol{x}\|^2-\|\boldsymbol{y}\|^2$, 同样得到

$$(\boldsymbol{x}|\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i.$$

由此容易看出以下性质.

▷ **双线性** $(\cdot|\cdot):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ 是双线性形式.

▷ **对称性** 对所有 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ 皆有 $(\boldsymbol{x}|\boldsymbol{y})=(\boldsymbol{y}|\boldsymbol{x})$.

▷ **正定性** 我们有 $(\boldsymbol{x}|\boldsymbol{x})\geq 0$, 等号成立当且仅当 $\boldsymbol{x}=\boldsymbol{0}$.

定义 9.1.1 以上对 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ 定义的 $(\boldsymbol{x}|\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i$ 称为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的**标准内积**. 我们称 $\|\boldsymbol{x}\|$ 为向量 $\boldsymbol{x}$ 的 **Euclid 长度**, 称 $\|\boldsymbol{y}-\boldsymbol{x}\|$ 为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的 **Euclid 距离**.

至今所有定义皆直接依赖直角坐标系的选取. 举例来说, 若将坐标沿着某个轴作伸缩, 则标准内积, 长度和角度都会受到扭曲. 这促使我们思考以下问题:

⋆ 内积是否有不依赖直角坐标系的表述方式?

⋆ 坐标变换如何影响内积?

⋆ 哪些坐标变换保持内积不变, 或者至多差一个伸缩 (例如图像的相似变换)?

如果内积的情形得到解答, 则关于长度和角度的问题也能随之解决.

9.2 内积空间

今后主要考虑实数域 $\mathbb{R}$ 上的向量空间, 简称实向量空间.

定义 9.2.1 实向量空间 V 上的**内积**意指正定对称双线性形式 $(\cdot|\cdot): V\times V\to\mathbb{R}$; 见定义 8.6.1. 这般资料 $(V,(\cdot|\cdot))$ 统称为**实内积空间**, 本章简称为**内积空间**.

为了帮助读者回忆, 现在将内积所需的性质复述如下.

▷ **双线性** $(\cdot|\cdot): V\times V\to\mathbb{R}$ 是双线性形式 (定义 8.1.1).

▷ **对称性** 对所有 $v,w\in V$ 皆有 $(v|w)=(w|v)$.

▷ **正定性** 我们有 $(v|v)\geqslant 0$, 等号成立当且仅当 $v=0$.

注意到当 V 有限维时, 正定性蕴涵内积是定义 8.2.1 所谓的非退化线性形式. 正定是一个显著强于非退化性质的条件.

和标准内积的情形类似, 对于内积空间 $(V,(\cdot|\cdot))$,

⋆ 记向量 $v\in V$ 的**长度**为 $\|x\|:=\sqrt{(x|x)}$, 因此 $\|tx\|=|t|\cdot\|x\|$, 其中 $t\in\mathbb{R}$.

⋆ 满足 $\|v\|=1$ 的向量称为**单位向量**. 任何非零向量 v 都可以通过伸缩化为单位向量: 取 $v/\|v\|$ 便是.

⋆ 若向量 $v,w\in V$ 满足 $(v|w)=0$, 则称它们**正交**, 也写作 $v\perp w$.

⋆ 推而广之, 若子空间 $V_1,V_2\subset V$ 满足

$$v_1\in V_1,\quad v_2\in V_2\implies v_1\perp v_2,$$

则称 V_1 与 V_2 正交, 记为 $V_1\perp V_2$.

一如命题 8.4.2, 长度也反过来通过称为**配极化**的手法确定内积:

$$(v|w)=\frac{1}{2}\left(\|v+w\|^2-\|v\|^2-\|w\|^2\right). \tag{9.2.1}$$

这是内积的双线性和对称性的简单应用. 等式 (9.2.1) 直接导致以下结果.

命题 9.2.2 (内积空间的勾股定理) 设 $(V,(\cdot|\cdot))$ 为内积空间, v 和 w 在其中正交, 则

$$\|v+w\|^2=\|v\|^2+\|w\|^2.$$

设 $n:=\dim V\in\mathbb{Z}_{\geqslant 1}$. 若选定 V 的有序基将 V 等同于 $\mathbb{R}^n$, 将 $(\cdot|\cdot)$ 作为双线性形式用矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$ 来表示, 亦即

$$(\boldsymbol{v}|\boldsymbol{w})={}^{\mathrm{t}}\boldsymbol{v}\boldsymbol{A}\boldsymbol{w},\quad \text{如命题 8.1.6,}$$

则对称性 $(\boldsymbol{v}|\boldsymbol{w})=(\boldsymbol{w}|\boldsymbol{v})$ 相当于说 $\boldsymbol{A}={}^{\mathrm{t}}\boldsymbol{A}$. 矩阵 $\boldsymbol{A}$ 又称为内积 $(\cdot|\cdot)$ 对此一有序基的 **Gram 矩阵**.

在进一步刻画正定性质之前, 先来介绍它的一些重要推论.

定理 9.2.3 (Cauchy–Bunyakovsky–Schwarz 不等式) 设 $(V,(\cdot|\cdot))$ 为内积空间, 则对任何 $v,w\in V$ 皆有

$$(v|w)^2\leq(v|v)(w|w);$$

等式成立当且仅当 v 和 w 线性相关.

证明 线性相关相当于说存在 $t\in\mathbb{R}$ 使得 $v=tw$ 或 $w=tv$, 两种情况下 $(v|w)^2=(v|v)(w|w)$ 都是自明的. 以下假设 v 和 w 线性无关, 如此则 $w\neq0$, 而且 $v+tw$ 对任何 $t\in\mathbb{R}$ 皆非零, 于是正定性导致

$$0<(v+tw|v+tw)=t^2(w|w)+2t(v|w)+(v|v),\quad t\in\mathbb{R}.$$

视此为关于 t 的二次多项式, 无实根, 故判别式满足

$$4(v|w)^2-4(v|v)(w|w)<0,$$

此即所求的不等式. □

之后证明复内积版本 (定理 10.3.6) 时, 我们将为上述不等式提供另一种论证.

推论 9.2.4 (内积空间的三角不等式) 设 $(V,(\cdot|\cdot))$ 为内积空间, 则对任何 $v,w\in V$ 皆有

$$\|v+w\|\leq\|v\|+\|w\|,$$

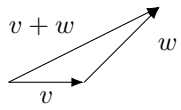
等号成立当且仅当存在 $t\geq0$ 使得 $v=tw$ 或 $w=tv$.

证明 对定理 9.2.3 的两边开平方根, 可得

$$\begin{aligned}\|v+w\|^2&=\|v\|^2+\|w\|^2+2(v|w)\\&\leq\|v\|^2+\|w\|^2+2|(v|w)|\\&\leq\|v\|^2+\|w\|^2+2\|v\|\cdot\|w\|=(\|v\|+\|w\|)^2.\end{aligned}$$

取到等号的充要条件是 v 和 w 线性相关而且 $(v|w)\geq0$, 这也等价于 v 和 w 差一个非负的比例常数. □

如果形象地考虑向量 v, w 和 $v+w$ 组成的三角形



则推论 9.2.4 相当于说三角形两边和大于等于第三边, 等号成立当且仅当三角形是压扁的.

定义 9.2.5 设 $(V,(\cdot|\cdot))$ 为内积空间. 相应的**距离函数**定义为

$$\begin{aligned} d : V \times V &\to \mathbb{R}_{\geq 0} \\ (v,w) &\mapsto \|w-v\|. \end{aligned}$$

因为任意 $x \in V$ 都满足 $\|x\| = \|-x\|$, 我们有 $d(v,w) = d(w,v)$. 推论 9.2.4 相当于说

$$d(u,v) + d(v,w) \leq d(u,w),$$

而内积的正定性相当于 $d(v,w) = 0$ 当且仅当 $v = w$. 综上, 资料 (V,d) 构成数学分析所探讨的度量空间, 其上可以探讨数列的收敛性和极限, 以及函数的连续性等.

我们至今见识过种种的代数结构; 对每一种结构, 都应该考虑保持结构的映射. 内积空间是叠架在向量空间上的结构, 以下定义不出所料.

定义 9.2.6 设 $(V,(\cdot|\cdot)_V)$ 和 $(W,(\cdot|\cdot)_W)$ 为内积空间. 线性映射 $\varphi : V \to W$ 若对所有 $v \in V$ 皆满足 $\|\varphi(v)\|_W = \|v\|_V$, 则称 φ 是**保距**的.

因为 $\varphi(v) = 0$ 等价于 $\|\varphi(v)\|_W = 0$, 保距线性映射必然单. 又由于内积可以通过 (9.2.1) 以长度来表达, 保距线性映射 $T : V \to W$ 也自动保内积

$$(Tv_1|Tv_2)_W = (v_1|v_2)_V.$$

定义 9.2.7 设 $(V,(\cdot|\cdot)_V)$ 和 $(W,(\cdot|\cdot)_W)$ 为内积空间, 若

$$V \underset{S}{\overset{T}{\rightleftarrows}} W$$

是一对保距线性映射, 使得 $TS = \mathrm{id}_W$, $ST = \mathrm{id}_V$, 则称 S 和 T 为内积空间的互逆同构.

注意到如果 $T : V \to W$ 保距, 而且是向量空间的同构, 则 T 自动是内积空间的同构. 缘由是简单的: 因为 T 是双射, 对 T 的保距条件 $\|Tv\|_W = \|v\|_V$ 代入 $w = Tv$, 则它自动蕴涵 T^{-1} 的版本 $\|T^{-1}w\|_V = \|w\|_W$.

9.3 Gram–Schmidt 正交化

首先考虑一般的实内积空间.

定义 9.3.1 若 V 中的一族非零元素两两正交, 则称之为**正交向量族**; 若进一步要求它们都是单位向量, 则称之为**单位正交向量族**.

给定任何正交向量族, 对其元素进行伸缩 $v \rightsquigarrow v/\|v\|$, 都能够得到单位正交向量族.

引理 9.3.2 正交向量族必然线性无关.

证明 设在正交向量族 S 中有线性关系 $\sum_{s\in S} a_s s = 0$, 要求仅有至多有限个系数非零. 对选定的 $t \in S$, 等式两边同时对 t 取内积, 得到

$$\left(\sum_{s\in S} a_s s \,\middle|\, t\right) = \sum_{s\in S} a_s (s|t) = a_t (t|t).$$

由 $(t|t) > 0$ 即得 $a_t = 0$. □

定义 9.3.3 由单位正交向量族给出的基称为**单位正交基**.

现在考虑有限维内积空间 $(V, (\cdot|\cdot))$. 命 $n := \dim V$. 引理 9.3.2 确保 n 个单位正交向量必然给出单位正交基. 设 $v_1, \dots, v_n \in V$ 是单位正交基, 则对任意元素 $v \in V$ 的展开式

$$v = \sum_{i=1}^{n} a_i v_i$$

两边同取内积 $(\cdot|v_i)$, 可以确定每一项的系数

$$a_i = (v|v_i), \quad i = 1, \dots, n.$$

这就导致下面的简单结论.

命题 9.3.4 若 $v_1, \dots, v_n$ 是有限维内积空间 $(V, (\cdot|\cdot))$ 的单位正交基 (计顺序), 则它所确定的线性映射

$$\begin{aligned} \mathbb{R}^n &\to V \\ (a_1, \dots, a_n) &\mapsto \sum_{i=1}^{n} a_i v_i \end{aligned}$$

给出从 $(\mathbb{R}^n$, 标准内积$)$ 到 $(V, (\cdot|\cdot))$ 的同构.

证明 既然 $v_1, \dots, v_n$ 是基, 此映射已经是向量空间的同构. 问题在于匹配两边的内积. 我们有

$$\left(\sum_{i=1}^{n} a_i v_i \,\middle|\, \sum_{j=1}^{n} b_j v_j\right) = \sum_{i,j} a_i b_j (v_i|v_j) = \sum_{i=1}^{n} a_i b_i.$$

因此这确实对应 $\mathbb{R}^n$ 的标准内积. □

问题在于单位正交基是否存在, 以及如何构造, 这是后续的重点. 许多相关结果也适用于无穷维的内积空间.

定理 9.3.5 (Gram–Schmidt 正交化) 设 V 中的一族向量 $v_1, v_2, \dots$ 线性无关 (容许个数有限或可数无穷). 递归地定义

$$\begin{aligned} w_1 &:= v_1, \\ w_k &:= v_k - \sum_{i=1}^{k-1} \frac{(w_i|v_k)}{(w_i|w_i)} \cdot w_i, \quad k \in \mathbb{Z}_{\geq 2}. \end{aligned}$$

则 $w_1, w_2, \dots$ 是正交向量族. 进一步命 $u_k := w_k/\|w_k\|$, 则 $u_1, u_2, \dots$ 是单位正交向量族, 而且对所有 $k \geq 1$ 皆有

$$\langle u_1, \dots, u_k \rangle = \langle w_1, \dots, w_k \rangle = \langle v_1, \dots, v_k \rangle .$$

证明 所求的性质基本蕴藏在定义中. 观察到 $w_1, \dots, w_k$ 的构造仅涉及 $v_1, \dots, v_k$. 以下递归地对所有 $k \geq 1$ 证明 $w_1, \dots, w_k$ 是正交向量族, 而且 $\langle w_1, \dots, w_k \rangle = \langle v_1, \dots, v_k \rangle$. 因为 u_k 是 w_k 的伸缩, 当然也有 $\langle u_1, \dots, u_k \rangle = \langle w_1, \dots, w_k \rangle$.

按定义, $k = 1$ 的情形是自明的. 设 $k \geq 2$, 由于 $w_k \in v_k + \langle w_1, \dots, w_{k-1} \rangle$, 故

$$\begin{aligned} \langle w_1, \dots, w_k \rangle = \langle w_1, \dots, w_{k-1} \rangle + \langle w_k \rangle &= \langle v_1, \dots, v_{k-1} \rangle + \langle w_k \rangle \\ &= \langle v_1, \dots, v_{k-1} \rangle + \langle v_k \rangle = \langle v_1, \dots, v_k \rangle . \end{aligned}$$

此外, $w_1, \dots, w_{k-1}$ 已知正交, 而当 $j < k$ 时

$$\begin{aligned} (w_k \mid w_j) &= (v_k|w_j) - \sum_{i=1}^{k-1} \frac{(v_k|w_i)}{(w_i|w_i)} \cdot (w_i|w_j) \\ &= (v_k|w_j) - \frac{(v_k|w_j)}{(w_j|w_j)} \cdot (w_j|w_j) = 0. \end{aligned}$$

最后留意到 $w_k \neq 0$, 否则 $v_k \in \langle w_1, \dots, w_{k-1} \rangle = \langle v_1, \dots, v_{k-1} \rangle$, 和线性无关的条件矛盾. 综上可见 $w_1, \dots, w_k$ 是正交向量族. □

练习 9.3.6 研究 Gram–Schmidt 正交化在 $v_1, v_2, \dots$ 线性相关时的产物; w_k 仍然按相同公式递归地定义.

提示〉 产物可能含零向量, 其余非零向量仍然正交.

推论 9.3.7 任何有限维内积空间都有单位正交基[1)].

证明 以定理 9.3.5 将任意基正交化, 再作伸缩即可. 由于内积正定, 这点也同样可由命题 8.6.3 导出. □

单位正交基的存在性还能进一步精细化.

练习 9.3.8 说明若 Gram–Schmidt 正交化中的 $v_1, v_2, \dots$ 是正交向量族, 则 $w_k = v_k$.

推论 9.3.9 任何有限维内积空间中的单位正交向量族都能扩充为单位正交基.

证明 设 $v_1, \dots, v_k$ 是有限维内积空间 V 中的单位正交向量族. 引理 9.3.2 说明它们线性无关, 故可扩充为基 $v_1, \dots, v_k, \dots, v_n$. 对此基施行正交化, 得到 V 的单位正交基 $u_1, \dots, u_n$; 按构造, 前 k 项既然已经单位正交, 它们不受正交化影响, 故 $1 \leq i \leq k$ 时 $u_i = v_i$. □

[1)] Gram–Schmidt 正交化还说明所有可数维内积空间都有单位正交基, 然而这类空间的应用不多, 无穷维情形更常用的概念是 Hilbert 空间.

定义 9.3.10 (正交补) 设 S 为内积空间 V 的任意子集, 命

$$S^{\perp} := \{v \in V : \forall s \in S,\ (s|v) = 0\}.$$

我们称 $S^{\perp}$ 为 S 的正交补.

因为 $(\cdot|\cdot)$ 是双线性的, $S^{\perp}$ 自动对加法和纯量乘法封闭, 而且包含 0, 因而是子空间. 从定义还容易推导 $S^{\perp} = \langle S\rangle^{\perp}$, 所以今后将着眼于 $S = V_0$ 是子空间的情形, 这也是定义 8.1.12 的特例. 特别地, 定理 8.2.19 蕴涵 $V_0^{\perp\perp} = V_0$.

命题 9.3.11 设 $(V, (\cdot|\cdot))$ 为内积空间, V_0 为 V 的子空间, 则 V_0 连同 $(\cdot|\cdot)$ 在 $V_0 \times V_0$ 上的限制仍然是内积空间.

证明 所需的双线性, 对称性和正定性条件显然可以被 V_0 继承. 顺带一提, 有限维向量空间上的对称非退化形式限制后仍是对称的, 却未必非退化. □

定义 9.3.12 给定内积空间 $(V, (\cdot|\cdot))$ 和一族子空间 $(V_i)_{i\in I}$, 如果

- ⋆ 向量空间的直和分解 $V = \bigoplus_{i\in I} V_i$ 成立,
- ⋆ 当 $i \neq j$ 时 $V_i \perp V_j$,

则称 $V = \bigoplus_{i\in I} V_i$ 为 V 的**正交直和分解**.

命题 9.3.13 设 $(V, (\cdot|\cdot))$ 为内积空间, 子空间 V_0 是有限维的, 则此时有正交直和分解

$$V = V_0 \oplus V_0^{\perp}.$$

具体地说, 若 $v_1, \ldots, v_m$ 是 V_0 的单位正交基, 则任意 $v \in V$ 有唯一的分解

$$v = \underbrace{\sum_{i=1}^{m} (v_i|v)v_i}_{\in V_0} + \underbrace{v - \sum_{i=1}^{m} (v|v_i)v_i}_{\in V_0^{\perp}}.$$

因此 $\sum_{i=1}^{m} (v_i|v)v_i$ 可谓是 v 在 V_0 上的**正交投影**.

证明 给定 $v \in V$, 首先显然有 $\sum_{i=1}^{m} (v_i|v)v_i \in V_0$. 其次, 对每个 $1 \leq j \leq m$, 从单位正交基的定义可知

$$\begin{aligned}\left(v - \sum_{i=1}^{m} (v_i|v)v_i \,\middle|\, v_j\right) &= (v|v_j) - \sum_{i=1}^{m} (v|v_i)(v_i|v_j) \\ &= (v|v_j) - (v|v_j) = 0.\end{aligned}$$

因此左式属于 $V_0^{\perp}$. 综上, $V = V_0 + V_0^{\perp}$.

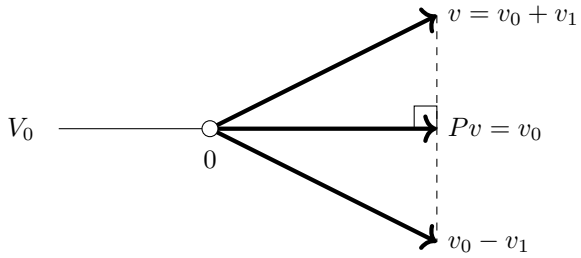
为了证明这是直和, 注意到若 $v \in V_0 \cap V_0^{\perp}$, 则 $\|v\|^2 = (v|v) = 0$, 故 $v = 0$. □

对于任意向量空间 V, 其任意子空间 V_0 总能实现为直和项, 这是命题 4.10.6 的内容; 一般来说, 一旦有向量空间的直和分解 $V = V_0 \oplus V_1$, 便有相应的投影映射

$$\begin{aligned} P &: V \to V_0 \\ P(v_0 + v_1) &= v_0, \quad v_i \in V_i, \end{aligned}$$

问题在于只有 V_0 是给定的, 直和分解中的 V_1 并无标准的或曰 "典范" 的取法.

另一方面, 借助于内积结构, V_1 在命题 9.3.13 的情境下有典范的取法 $V_0^\perp$, 相应的 $P : V \to V_0$ 从而也是典范地定义的. 由于数学物理或数学分析的研究中习惯将线性映射称为线性算子, 故 P 也经常被称为向子空间 V_0 的**正交投影算子**. 几何图像如下.



举例明之, Gram–Schmidt 正交化 (定理 9.3.5) 中构造的

$$w_k := v_k - \sum_{i=1}^{k-1} \frac{(w_i|v_k)}{(w_i|w_i)} \cdot w_i = v_k - \sum_{i=1}^{k-1} (u_i|v_k) u_i.$$

正是 v_k 减去它在 $\langle v_1, \dots, v_{k-1} \rangle$ 上的正交投影.

由几何图像易知 $v_0 - v_1$ 应当是 $v = v_0 + v_1$ 对 V_0 的镜像, 这一操作也容易用正交投影算子表达.

定义 9.3.14 记 P 为向子空间 V_0 的正交投影算子, 视同 $\mathrm{End}(V)$ 的元素, 则对 V_0 的**镜射**定义为 $2P - \mathrm{id}_V \in \mathrm{End}(V)$; 它映 $v_0 + v_1$ 为 $v_0 - v_1$, 其中 $v_0 \in V_0$ 而 $v_1 \in V_0^\perp$.

一如我们所熟悉的低维情形, 正交投影在一般的内积空间中也与一个向量到 V_0 的最短距离密切相关.

命题 9.3.15 设 $(V, (\cdot|\cdot))$ 为内积空间, V_0 为有限维子空间, 则对于任意 $v \in V$, 距离

$$\|u - v\|, \quad u \in V_0$$

取到极小值当且仅当 u 是 v 在 V_0 上的正交投影.

证明 作分解 $v = v_0 + v_1$, 其中 $v_0 \in V_0$ 而 $v_1 \in V_0^\perp$, 则内积空间的勾股定理 (命题 9.2.2) 对 $u \in V_0$ 给出

$$\|u - v\|^2 = \|(u - v_0) - v_1\|^2 = \|u - v_0\|^2 + \|v_1\|^2 \geq \|v_1\|^2,$$

等号成立的充要条件是 $u = v_0$. □

例 9.3.16 (Legendre 多项式) 将 $\mathbb{R}$ 上的所有多项式作成 $\mathbb{R}$-向量空间 $\mathbb{R}[X]$. 将多项式视同 $\mathbb{R}$ 上的多项式函数, 对所有 $f,g\in\mathbb{R}[X]$ 定义

$$(f|g):=\int_{-1}^{1}f(x)g(x)\,\mathrm{d}x,\quad f,g\in\mathbb{R}[X].$$

基于数学分析的基础知识 (例如 $\int_{-1}^{1}f(x)^2\,\mathrm{d}x=0\iff f|_{[-1,1]}=0\iff f=0$), 可见 $(\mathbb{R}[X],(\cdot|\cdot))$ 构成内积空间. 另一方面, $\mathbb{R}[X]$ 又有一组自然的基

$$1,X,X^2,X^3,\dots.$$

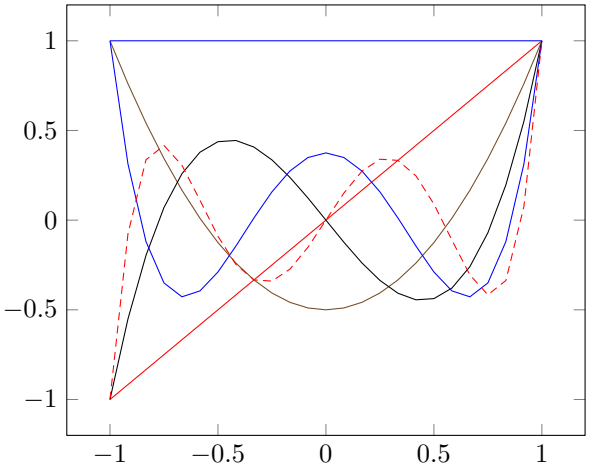
对它们施行 Gram–Schmidt 正交化的产物是颇富兴味的数学对象. 尊重传统, 此处将向量的下标按 $n=0,1,\dots$ 排列, 不难看出对 $v_n:=X^n$ 作正交化给出的 $w_0,w_1,\dots$ 由以下性质唯一刻画:

$$\begin{gathered}w_0=1,\\ w_n=X^n+\text{低次项},\\ i\neq j\implies(w_i|w_j)=0.\end{gathered}$$

事实上有 $w_n(1)\neq0$, 这当然是需要证明的, 细节留至本章习题处理; 承认这一事实, 则 $w_0,w_1,\dots$ 也可以适当地伸缩, 给出由以下条件唯一刻画的 $P_0,P_1,\dots\in\mathbb{R}[X]$, 称为 **Legendre 多项式**:

$$\begin{gathered}\deg P_n=n,\\ P_n(1)=1,\\ i\neq j\implies(P_i|P_j)=0.\end{gathered}$$

以下是 $P_0,\dots,P_5$ 及其函数图形的描绘.



$$P_0=1,$$

$$P_1=X,$$

$$P_2=\tfrac{1}{2}(3X^2-1),$$

$$P_3=\tfrac{1}{2}(5X^3-3X),$$

$$P_4=\tfrac{1}{8}(35X^4-30X^2+3),$$

$$P_5=\tfrac{1}{8}(63X^5-70X^3+15X).$$

这些资料初步提示了一些规律, 比如函数 P_n 的奇偶性与 n 相同, 以及 P_n 在 $[-1,1]$ 上有 n 个根. 本章习题对此将有较深入的探究, 比方说, 我们将建立以下的 **Rodrigues 公式**, 这对研究 Legendre 多项式是特别方便的:

$$P_n = \frac{1}{2^n n!} \underbrace{\left((X^2-1)^n\right)^{(n)}}_{n\text{ 次导数}}, \quad n \in \mathbb{Z}_{\geq 1}.$$

此外, 习题也会介绍另一类称为 Chebyshev 多项式的对象, 它们同样与正交性密切相关.

9.4 内积空间上的伴随映射和正交变换

定义 9.2.6 已介绍了保距线性映射的概念. 本节聚焦于从内积空间 V 到自身的保距线性映射.

选定有限维内积空间 $(V,(\cdot|\cdot)_V)$ 和 $(W,(\cdot|\cdot)_W)$. 内积非退化而且对称, 于是根据定义–命题 8.2.8, 所有线性映射 $T: V \to W$ 皆有伴随[2] $T^*: W \to V$, 这是使得等式

$$(Tv \mid w)_W = (v|T^*w)_V, \quad \text{或等价地} \quad (T^*w \mid v)_V = (w|Tv)_W$$

对所有 $v \in V$ 和 $w \in W$ 成立的唯一线性映射 T^* 回忆到命题 8.2.12 断言

$$(ST)^* = T^*S^*, \quad (T^*)^* = T,$$

前提是映射合成有意义.

命题 9.4.1 取 $(V,(\cdot|\cdot)_V)$ 和 $(W,(\cdot|\cdot)_W)$ 如上. 线性映射 $T: V \to W$ 是内积空间的同构 (定义 9.2.7) 当且仅当 $T^* = T^{-1}$.

证明 设 T 是内积空间的同构, 则

$$(Tv \mid w)_W = \left(T^{-1}Tv \,\middle|\, T^{-1}w\right)_V = \left(v \,\middle|\, T^{-1}w\right)_V, \quad v \in V,\ w \in W,$$

这表明 $T^* = T^{-1}$. 反之设 $T^* = T^{-1}$, 则 $(Tv_1 \mid Tv_2)_W = (v_1 \mid T^*Tv_2)_V = (v_1|v_2)_V$, 而 $v_1, v_2 \in V$ 可任选, 故 T 是保距的线性同构. □

命题 9.4.2 设 $T: V \to W$ 是有限维内积空间之间的线性映射, $v_1, \dots, v_n$ 是 V 的单位正交基 (定义 9.3.3), 则线性映射 T 是内积空间的同构当且仅当 $Tv_1, \dots, Tv_n$ 是 W 的单位正交基.

[2] 所以改记为 $^*T$ 也完全合理.

证明 若 T 是内积空间的同构, 则它保持一切关于内积的性质, 故 $Tv_1,\dots,Tv_n$ 仍是单位正交基. 反之设 $Tv_1,\dots,Tv_n$ 是单位正交基, 则

$$\left\|\sum_{i=1}^n a_i v_i\right\|_V^2 = \sum_{i=1}^n a_i^2,$$
$$\left\|T\left(\sum_{i=1}^n a_i v_i\right)\right\|_W^2 = \left\|\sum_{i=1}^n a_i Tv_i\right\|_W^2 = \sum_{i=1}^n a_i^2,$$

从而 T 保距. 此外 $\dim V = n = \dim W$, 故 T 是同构. □

今后主要针对 $(V,(\cdot|\cdot)_V)=(W,(\cdot|\cdot)_W)$ 的情形探讨伴随映射.

定义 9.4.3 有限维内积空间 $(V,(\cdot|\cdot))_V$ 的自同构称为 V 上的**正交变换**.

由于保距线性映射自动是单的, 所以 V 上的正交变换相当于是 V 到其自身的保距线性映射. 命题 9.4.1 相当于说 $T\in\mathrm{End}(V)$ 是正交变换当且仅当 $T^*=T^{-1}$.

现在取 $V=\mathbb{R}^n$, 配备标准内积. 我们着手以矩阵来描述正交变换.

定义–命题 9.4.4 对于矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$, 以下性质等价.

(i) ${}^{\mathrm{t}}\boldsymbol{A}\,\boldsymbol{A}=\mathbf{1}_{n\times n}$;

(ii) $\boldsymbol{A}\,{}^{\mathrm{t}}\boldsymbol{A}=\mathbf{1}_{n\times n}$;

(iii) 视为从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性映射, $\boldsymbol{A}$ 相对于标准内积是正交变换.

具有这种性质的矩阵称为实**正交矩阵**.

证明 性质 (i) 和 (ii) 的等价性是关于矩阵的一般事实, 见命题 4.8.13 末段. 将 $\mathbb{R}^n$ 的元素看作列向量, 则标准内积表为

$$(\boldsymbol{v}|\boldsymbol{w}) = {}^{\mathrm{t}}\boldsymbol{w}\,\boldsymbol{v}\in \mathrm{M}_{1\times 1}(\mathbb{R})=\mathbb{R}.$$

将 $\boldsymbol{A}$ 视同线性映射 $\mathbb{R}^n\to\mathbb{R}^n$, 则它相对于标准内积的伴随映射是作为矩阵取转置 ${}^{\mathrm{t}}\boldsymbol{A}$ (例 8.2.16). 于是 (iii) 和 (i) 或 (ii) 的等价性归结为命题 9.4.1. □

因此实正交矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$ 的等价刻画是 ${}^{\mathrm{t}}\boldsymbol{A}=\boldsymbol{A}^{-1}$.

命题 9.4.5 以下断言适用于 $n\times n$ 实矩阵.

(i) 单位矩阵 $\mathbf{1}_{n\times n}$ 是正交矩阵.

(ii) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是正交矩阵, 则 $\boldsymbol{AB}$ 亦然.

(iii) 若 $\boldsymbol{A}$ 是正交矩阵, 则 $\boldsymbol{A}^{-1}$ 亦然.

(iv) 正交矩阵的行列式必为 ± 1.

证明　性质 (i) — (iii) 不外是将同构的一般性质从内积空间转译到矩阵: 无论对任何结构, 恒等映射总是同构, 而同构的合成和逆仍然是同构. 至于 (iv), 对 ${}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{A} = \mathbf{1}_{n\times n}$ 两边取行列式, 便得到 $(\det \boldsymbol{A})^2 = \det {}^{\mathrm{t}}\boldsymbol{A} \det \boldsymbol{A} = 1$. □

练习 9.4.6　设 $v_1, \dots, v_n$ 为内积空间 V 的单位正交基, 记这组资料为 $\mathbf{v}$. 考虑 V 的有序基 $v'_1, \dots, v'_n$, 记之为 $\mathbf{v}'$. 注记 4.9.2 引入的转换矩阵 $\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 记录了如何将 $v'_1, \dots, v'_n$ 用 $\mathbf{v}$ 展开. 说明

$$ {}^{\mathrm{t}}(\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'})\, \boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'} = \big((v'_i | v'_j)\big)_{1\le i,j\le n}; $$

因此 $\mathbf{v}'$ 是单位正交基当且仅当 $\boldsymbol{P}^{\mathbf{v}}_{\mathbf{v}'}$ 是正交矩阵.

定理 9.3.5 介绍的 Gram–Schmidt 正交化给出一种实用的矩阵分解.

推论 9.4.7 ($\boldsymbol{QR}$ 分解)　任何可逆矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 都能表作 $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$, 其中 $\boldsymbol{Q}$ 是正交矩阵, $\boldsymbol{R}$ 是可逆上三角矩阵.

证明　将 $\boldsymbol{A}$ 按列表成

$$ \boldsymbol{A} = \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{array} \right), $$

其中每个 $\boldsymbol{v}_i$ 视同内积空间 $(\mathbb{R}^n, \cdot)$ 的元素. 根据命题 9.4.2, 所求的分解等价于说存在上三角矩阵 $\boldsymbol{R}' \in \mathrm{M}_{n\times n}(\mathbb{R})$, 相当于断言中的 $\boldsymbol{R}$ 的逆, 使得

$$ \boldsymbol{Q} = \left(\begin{array}{c|c|c} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{array} \right) := \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{array} \right) \boldsymbol{R}' \quad \text{是正交矩阵,} \tag{9.4.1} $$

或者等价地说, 使得 $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \mathbb{R}^n$ 相对于标准内积成为单位正交基. 注意到 $\boldsymbol{R}'$ 自动可逆, 因为 $\boldsymbol{A}$ 和 $\boldsymbol{Q}$ 皆可逆; 它记录着如何按 $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ 来展开 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$.

既然 $\boldsymbol{A}$ 可逆, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ 线性无关, 故 Gram–Schmidt 正交化定理 9.3.5 给出单位正交基 $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$, 使得对每个 $1 \le j \le n$ 皆有 $\boldsymbol{u}_j \in \langle \boldsymbol{v}_1, \dots, \boldsymbol{v}_j \rangle$; 换言之存在一族系数 $a_{ij} \in \mathbb{R}$, 其中 $1 \le i \le j \le n$, 使得

$$ \boldsymbol{u}_j = \sum_{i=1}^{j} a_{ij} \boldsymbol{v}_i, \quad j = 1, \dots, n. $$

另对 $i > j$ 定义 $a_{ij} := 0$, 则上三角矩阵 $\boldsymbol{R}' := (a_{ij})_{i,j}$ 使 (9.4.1) 成立. □

Gram–Schmidt 正交化实际还给出了构造 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的具体算法.

练习 9.4.8　说明推论 9.4.7 中的 $\boldsymbol{R}$ 可以适当选取, 使得对角线上都是正数. 提示〉 等价于说明证明中的 $a_{ii} > 0$. 为此, 请回顾 Gram–Schmidt 正交化对 $\boldsymbol{w}_k$ 的定义, 以及 $\boldsymbol{u}_k = \frac{\boldsymbol{w}_k}{\|\boldsymbol{w}_k\|}$.

作为应用, 我们回到 §9.3 介绍的正交投影算子 $P : V \to V_0$, 在此 $(V, (\cdot|\cdot))$ 取作有限维内积空间, 而 V_0 是 V 的子空间. 既然 $V_0 \subset V$, 也可以将 P 视为 $\mathrm{End}(V)$ 的元素. 这些算子有简明的刻画.

命题 9.4.9 给定有限维内积空间 $(V,(\cdot|\cdot))$, 则 $P\in\mathrm{End}(V)$ 是向某个子空间 V_0 的正交投影算子当且仅当

$$P^*=P,\quad P^2=P;$$

此时 $V_0=\mathrm{im}(P)$.

证明 设 P 是向 V_0 的正交投影算子. 将任意 $v\in V$ 唯一地表成 v_0+v_1, 其中 $v_0\in V_0$ 而 $v_1\in V_0^\perp$. 于是 $Pv=v_0$ 而 $P^2v=P(v_0)=v_0$, 故 $P^2=P$. 若另有 $v'\in V$ 和类似的分解 $v'=v_0'+v_1'$, 则

$$\begin{aligned}(Pv|v')&=(v_0|v')=(v_0|v_0')\\&=(v|v_0')=(v|Pv'),\end{aligned}$$

这说明 $P^*=P$. 最后 $\mathrm{im}(P)=V_0$ 是当然的.

反之设 $P\in\mathrm{End}(V)$ 满足 $P^*=P$ 和 $P^2=P$. 取 $V_0:=\mathrm{im}(P)$. 将任意 $v\in V$ 按先前方式作唯一分解 $v=Pu+v_1$, 其中 $u\in V$ 而 $v_1\in V_0^\perp$, 则

$$Pv=P^2u+Pv_1=Pu+Pv_1.$$

由于 $v_1\in V_0^\perp$, 从

$$(Pv_1|Pv_1)=(P^*Pv_1|v_1)=(\underbrace{P^2v_1}_{\in V_0}|v_1)=0$$

推知 $Pv_1=0$. 因此 $Pv=Pu$. 这便说明 P 是向 V_0 的正交投影算子. □

推论 9.4.10 取 $(V,(\cdot|\cdot))$ 和 V_0 如上, 则对 V_0 的镜射 $2P-\mathrm{id}_V$ (定义 9.3.14) 是正交变换.

证明 直接从镜射的观点验证并不困难, 以下则从 P 的性质来论证. 只须观察到 $P^*=P$ 和 $P^2=P$ 导致

$$\begin{aligned}(2P-\mathrm{id}_V)^*(2P-\mathrm{id}_V)&=(2P-\mathrm{id}_V)(2P-\mathrm{id}_V)\\&=4P^2-4P+\mathrm{id}_V=\mathrm{id}_V;\end{aligned}$$

同理也有 $(2P-\mathrm{id}_V)(2P-\mathrm{id}_V)^*=\mathrm{id}_V$. □

既然在有限维内积空间 $(V,(\cdot|\cdot))$ 中指定子空间 V_0 和指定相应的正交投影 $P\in\mathrm{End}(V)$ 是一回事 (通过 $\mathrm{im}(P)=V_0$ 对应), 定义 9.3.12 的正交分解也可以用投影算子的性质来刻画如下.

命题 9.4.11 设 $(V,(\cdot|\cdot))$ 为有限维内积空间, $V_1,\dots,V_s$ 为子空间, 依序对应到投影算子 $P_1,\dots,P_s\in\mathrm{End}(V)$. 以下陈述等价:

(i) 有正交直和分解 $V=V_1\oplus\cdots\oplus V_s$

(ii) $\sum_{i=1}^{s} P_i = \mathrm{id}_V$ 而 $i \neq j \implies P_i P_j = 0$.

证明 (i) $\implies$ (ii). 根据条件, 任何 $v \in V$ 都能写成 $\sum_i v_i$, 其中 $v_i \in V_i$. 既然 $V_1, \dots, V_s$ 两两正交, 易知 $V_i^{\perp} = \bigoplus_{j \neq i} V_j$, 从而有 $P_i v = v_i$. 这就蕴涵 $\sum_{i=1}^{s} P_i v = \sum_{i=1}^{s} v_i = v$. 此外, 这也说明当 $i \neq j$ 时 $P_i P_j(v) = P_i v_j = 0$.

(ii) $\implies$ (i). 根据条件, 任何 $v \in V$ 都等于 $\sum_{i=1}^{s} P_i v$, 从而属于 $\sum_{i=1}^{s} V_i$. 为了说明这是正交直和分解, 对所有 $i \neq j$ 证 $V_i \perp V_j$ 即可. 设 $v_i \in V_i$ 而 $v_j \in V_j$, 则

$$\begin{aligned}(v_i | v_j) &= (P_i v_i | P_j v_j) = (v_i | P_i^* P_j v_j) \\ &= (v_i | P_i P_j v_j) = 0.\end{aligned}$$

明所欲证. □

回忆到根据定义 8.2.15, 满足 $T^* = T$ 的线性映射 $T \in \mathrm{End}(V)$ 称为自伴的, 又称自伴算子. 上述论证已经初步示范了自伴性质的妙用, 下一节将有更深入的说明.

9.5 自伴算子的正交对角化

取定 $\mathbb{R}$ 上的有限维内积空间 $(V, (\cdot|\cdot))$. 以下论及的伴随映射 T^* 都是相对于此内积而言的.

引理 9.5.1 设 $T \in \mathrm{End}(V)$ 而 V_0 是 V 的 T-不变子空间 (定义 5.10.1), 则 $V_0^{\perp}$ 是 T^*-不变子空间.

证明 操练定义. 对于任意 $w \in V_0^{\perp}$ 和 $v \in V_0$, 我们有 $(T^* w \mid v) = (w \mid Tv)$, 然而条件蕴涵 $Tv \in V_0$, 故此内积为 0. 这就表明 $T^*(V_0^{\perp}) \subset V_0^{\perp}$. □

现在已有足够工具证明自伴线性映射或自伴算子可对角化.

定理 9.5.2 (正交对角化: 自伴情形) 设 $T \in \mathrm{End}(V)$ 自伴, 则存在 V 的单位正交基 $v_1, \dots, v_n$ 使得每个 v_i 都是 T 的特征向量.

证明 对 $n = \dim V$ 递归地论证. 当 $n = 0, 1$ 时无事可作. 以下设 $n \geq 2$.

关键的第一步是说明 T 有特征值 $\lambda_1 \in \mathbb{R}$. 由于 $(V, (\cdot|\cdot))$ 同构于标准内积空间 $(\mathbb{R}^n, \cdot)$, 不妨将 T 等同于矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$, 满足 ${}^{\mathrm{t}}\boldsymbol{A} = \boldsymbol{A}$. 稍后的引理 9.5.4 (或本章习题) 将给出 $\boldsymbol{A}$ 的特征值 $\lambda_1 \in \mathbb{R}$.

承认这一性质, 取 $v_1 \in V$ 非零使得 $Tv_1 = \lambda_1 v_1$. 适当伸缩后不妨假设 $\|v_1\| = 1$. 考虑正交直和分解

$$V = \langle v_1 \rangle \oplus \langle v_1 \rangle^{\perp}.$$

留意到 $\langle v_1 \rangle^{\perp}$ 相对于 $(\cdot|\cdot)$ 仍是内积空间. 既然 $\langle v_1 \rangle$ 是 T-不变子空间, 引理 9.5.1 蕴涵 $\langle v_1 \rangle^{\perp}$ 亦然. 恒等式 $(Tv|w) = (v|Tw)$ 限制到 $v, w \in \langle v_1 \rangle^{\perp}$ 上依然成立, 所以 T 限制在 $\langle v_1 \rangle^{\perp}$ 上仍然自伴.

根据递归假设, 存在 $\langle v_1\rangle^\perp$ 的单位正交基 $v_2,\dots,v_n$ 连同 $\lambda_2,\dots,\lambda_n\in\mathbb{R}$, 使得 $Tv_i=\lambda_i v_i$ 对 $i=2,\dots,n$ 成立. 于是 $v_1,\dots,v_n$ 给出 V 的单位正交基, 使得每个 v_i 都是以 λ_i 为特征值的特征向量 $(i=1,\dots,n)$. □

定理 9.5.2 也可以用矩阵语言表述. 不失一般性, 设 $(V,(\cdot|\cdot))$ 为标准内积空间 $(\mathbb{R}^n,\cdot)$, 将 T 等同于矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$, 满足 ${}^{\mathrm{t}}\boldsymbol{A}=\boldsymbol{A}$. 取 $\boldsymbol{P}$ 为以单位正交基 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$ 为列向量的矩阵, 则 $\boldsymbol{P}$ 是正交矩阵, 而定理的结论相当于给出矩阵对角化

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},\quad \lambda_1,\dots,\lambda_n\in\mathbb{R}.$$

由于 ${}^{\mathrm{t}}\boldsymbol{P}=\boldsymbol{P}^{-1}$, 上式既是矩阵的共轭关系, 同时又是定义 8.3.8 所谓的合同关系.

为了补全定理 9.5.2 的证明, 最简单的途径是在复数域上操作. 本章习题将另外给出一种只涉及实数的证明方案.

约定 9.5.3 对任意复矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}\in \mathrm{M}_{m\times n}(\mathbb{C})$, 定义

$$\overline{\boldsymbol{A}}:=(\overline{a_{ij}})_{i,j}\in \mathrm{M}_{m\times n}(\mathbb{C}).$$

显然,

$$\overline{\boldsymbol{A}+\boldsymbol{B}}=\overline{\boldsymbol{A}}+\overline{\boldsymbol{B}},\quad \overline{\boldsymbol{A}\cdot\boldsymbol{B}}=\overline{\boldsymbol{A}}\cdot\overline{\boldsymbol{B}},$$

前提是矩阵的加法和乘法在此有定义. 特别地, 对任意 $\boldsymbol{v}=(x_1,\dots,x_n)\in\mathbb{C}^n$ 也可以定义 $\overline{\boldsymbol{v}}=(\overline{x_1},\dots,\overline{x_n})$. 我们另外定义

$${}^{\dagger}\boldsymbol{A}:={}^{\mathrm{t}}\overline{\boldsymbol{A}}={}^{\mathrm{t}}\overline{\boldsymbol{A}}.$$

这使得 ${}^{\dagger}({}^{\dagger}\boldsymbol{A})=\boldsymbol{A}$, ${}^{\dagger}(\lambda\boldsymbol{A})=\overline{\lambda}\,{}^{\dagger}\boldsymbol{A}$ (其中 $\lambda\in\mathbb{C}$), 而且 ${}^{\dagger}(\boldsymbol{A}\boldsymbol{B})={}^{\dagger}\boldsymbol{B}\,{}^{\dagger}\boldsymbol{A}$ 对所有 $\boldsymbol{A},\boldsymbol{B}$ 成立, 前提是乘法有定义.

取 $\boldsymbol{v}=(z_1,\dots,z_n)\in\mathbb{C}^n$, 视同列向量, 则

$${}^{\dagger}\boldsymbol{v}\cdot\boldsymbol{v}=\sum_{i=1}^{n}|z_i|^2\ \in\mathbb{R}_{\geq 0};\quad {}^{\dagger}\boldsymbol{v}\cdot\boldsymbol{v}=0\iff \boldsymbol{v}=0.$$

引理 9.5.4 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 满足 ${}^{\dagger}\boldsymbol{A}=\epsilon\boldsymbol{A}$, 其中 $\epsilon\in\mathbb{C}$, 则 $\boldsymbol{A}$ 的所有特征值都满足 $\overline{\lambda}=\epsilon\lambda$.

证明 设 $\boldsymbol{v}\in\mathbb{C}^n$ 是以 $\lambda\in\mathbb{C}$ 为特征值的特征向量. 考虑

$${}^{\dagger}\boldsymbol{v}\cdot(\boldsymbol{A}\boldsymbol{v})=\lambda\cdot({}^{\dagger}\boldsymbol{v}\,\boldsymbol{v}).$$

对两边取 ${}^{\dagger}(\cdots)$. 左边给出 ${}^{\dagger}\boldsymbol{v}\,{}^{\dagger}\boldsymbol{A}\,\boldsymbol{v}=\epsilon\cdot{}^{\dagger}\boldsymbol{v}\,\boldsymbol{A}\,\boldsymbol{v}=\epsilon\lambda\,({}^{\dagger}\boldsymbol{v}\,\boldsymbol{v})$, 右边则给出 $\overline{\lambda}\,({}^{\dagger}\boldsymbol{v}\,\boldsymbol{v})$. 两边消去 ${}^{\dagger}\boldsymbol{v}\,\boldsymbol{v}\in\mathbb{R}_{>0}$ 即是 $\overline{\lambda}=\epsilon\lambda$. □

反自伴和正交矩阵的对角化必须取道复数域, 这是定理 10.4.1 的内容, 届时也将一并讨论一些在实数域上的应用.

最后探讨正交对角化对于实二次型的初步应用. 基于双线性形式的定义 8.4.4 对此是颇为方便的.

定理 9.5.5 *设 $(V,(\cdot|\cdot))$ 是有限维实内积空间, (V,B) 是二次型, 则存在 V 的单位正交基 $v_1,\dots,v_n$ 使得 B 所对应的对称矩阵 $(B(v_i,v_j))_{1\leq i,j\leq n}$ 是对角矩阵.*

证明 不妨假设 $V=\mathbb{R}^n$ 而 $(\cdot|\cdot)$ 是标准内积, 将 B 等同于对称矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$. 根据正交对角化定理 9.5.2 的矩阵形式, 存在正交矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角的. 依旧记 $\mathbb{R}^n$ 的标准基为 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$, 视同列向量, 并将 $\boldsymbol{P}$ 按列展开为 $(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)$, 则由 $\boldsymbol{v}_i=\boldsymbol{P}\boldsymbol{e}_i$ 和 $\boldsymbol{P}^{-1}={}^{\mathrm{t}}\boldsymbol{P}$ 可知

$$\begin{aligned} B(\boldsymbol{v}_i,\boldsymbol{v}_j) &= {}^{\mathrm{t}}\boldsymbol{v}_i\,\boldsymbol{A}\,\boldsymbol{v}_j = {}^{\mathrm{t}}\boldsymbol{e}_i\,{}^{\mathrm{t}}\boldsymbol{P}\,\boldsymbol{A}\,\boldsymbol{P}\,\boldsymbol{e}_j \\ &= \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} \text{ 的 } (i,j) \text{ 矩阵元.} \end{aligned}$$

综上, $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$ 即所求的单位正交基. □

总结上述讨论, 在内积空间 V 上给定二次型, 总能找到单位正交基 $v_1,\dots,v_n$, 使得 V 借此等同于 $\mathbb{R}^n$, 而二次型相应地化为

$$\lambda_1 X_1^2+\cdots\lambda_n X_n^2$$

的形式. 在 $V=\mathbb{R}^n$ 的标准情形, 这与关于 $\mathbb{R}^n$ 中的二次超曲面

$$\sum_{1\leq i,j\leq n} a_{ij}X_iX_j=0,\quad a_{ij}\in\mathbb{R}$$

的初等研究直接相关. 我们以后将说明变换单位正交基相当于在空间中施行刚体变换 (旋转, 镜射...), 而上述结论相当于说给定二次超曲面, 总能适当地施行刚体变换, 使得其方程不再包含交叉项; 举熟悉的 $n=2$ 情形为例, 不含交叉项的形式分为三类:

$$\begin{aligned} \text{椭圆} &\quad \lambda_1\lambda_2>0 \\ \text{双曲线} &\quad \lambda_1\lambda_2<0 \\ \text{抛物线} &\quad \lambda_1\lambda_2=0 \end{aligned}$$

考量定理 9.5.5 证明中的 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$, 则 $\boldsymbol{v}_i$ 是以 λ_i 为特征值的特征向量. 在特征值两两相异的一般情形下, 直线 $\mathbb{R}\boldsymbol{v}_1,\dots,\mathbb{R}\boldsymbol{v}_n$ 是唯一确定的, 几何直观上便是椭圆或双曲线 (或其高维推广) 的轴. 出于这一考量, 定理 9.5.5 也常被称为**主轴定理**.

我们将在 §9.10 继续研究主轴定理中的 $\lambda_1,\dots,\lambda_n$.

9.6 应用: 最小二乘解

考虑有限维向量空间之间的线性映射 $T : V \to W$. 给定 $w \in W$, 求解

$$Tv = w, \quad v \in V$$

相当于解线性方程组, 前提是取定 V 和 W 的基. 一般而言, 方程 $Tv = w$ 未必有解 v, 应用中寻求的往往也不是精确解, 而是它的某种逼近. 何谓逼近, 如何逼近? 当 V 和 W 都是实数域 $\mathbb{R}$ 上的内积空间时, 相应的度量结构导向了以下概念. 记 V (或 W) 的内积为 $(\cdot|\cdot)_V$ (或 $(\cdot|\cdot)_W$), 而 $\|v\|_V := \sqrt{(v|v)_V}$, $\|w\|_W := \sqrt{(w|w)_W}$.

定义 9.6.1 (最小二乘解) 给定 $w \in W$, 使得 $\|Tv - w\|_W$ 取到极小值的 $v \in V$ 称为方程 $Tv = w$ 的最小二乘解.

一旦取定单位正交基, 则 $\|\cdot\|_W^2$ 可以表示成坐标的平方和, 这是最小二乘解一词的由来. 至于极小值能否取到, 以及相应的 $v \in V$ 有何刻画, 则是以下引理的内容.

引理 9.6.2 给定 $w \in W$, 将它按命题 9.3.13 的正交直和分解唯一地表成

$$w = w' + w'', \quad w' \in \mathrm{im}(T), \quad w'' \in \mathrm{im}(T)^\perp,$$

则 $v \in V$ 是最小二乘解当且仅当 $Tv = w'$; 这般的 v 总是存在, 等价的条件是 $Tv - w \in \mathrm{im}(T)^\perp$.

证明 求 $\|Tv - w\|_W$ 的极小值相当于求向量 w 到子空间 $\mathrm{im}(T)$ 的距离, 我们知道极小值恰好在 Tv 等于 w 向 $\mathrm{im}(T)$ 的正交投影时取到 (命题 9.3.15); 这一条件无非是 $Tv = w'$, 也等价于 $Tv - w \in \mathrm{im}(T)^\perp$. □

留意到定义最小二乘解时仅在 W 上考虑了向量长度的极值, 对应的 v 并非唯一的. 如果进一步取长度极小的 v, 便导向了 Moore–Penrose 广义逆的概念, 这是 §9.9 的主题.

运用 T 的伴随 $T^* : W \to V$, 最小二乘解可以具体刻画为另一线性方程组的解.

定理 9.6.3 给定 $w \in W$, 则方程 $Tv = w$ 的最小二乘解 $v \in V$ 正好是 $T^*Tv = T^*w$ 的解.

证明 回忆到 v 是最小二乘解当且仅当 $Tv - w \in \mathrm{im}(T)^\perp$, 换言之

$$\forall v' \in V,\ (Tv - w \mid Tv')_W = 0.$$

然而 $(Tv - w \mid Tv')_W = (T^*Tv - T^*w \mid v')_V$, 而 $v' \in V$ 可任取, 故上式等价于

$$T^*Tv = T^*w.$$

明所欲证. □

观察到 $T^*T \in \mathrm{End}(V)$ 和 $TT^* \in \mathrm{End}(W)$ 皆自伴, 这是因为在内积空间的框架下, 任意线性映射取双重伴随回到自身; 见命题 8.2.12. 这两种映射将在往后的论证中反复出现. 关于最小二乘解的讨论顺带给出下述结果.

推论 9.6.4 对有限维实内积空间 V, W 和线性映射 $T: V \to W$, 我们有

$$\begin{aligned} \mathrm{im}\,(T^*T) &= \mathrm{im}\,(T^*), \quad \mathrm{im}\,(TT^*) = \mathrm{im}\,(T), \\ \ker(T^*T) &= \ker(T), \quad \ker(TT^*) = \ker(T^*), \\ \mathrm{rk}\,(T^*T) &= \mathrm{rk}(T), \quad \mathrm{rk}\,(TT^*) = \mathrm{rk}\,(T^*). \end{aligned}$$

证明 先说明 $\mathrm{im}\,(T^*T) = \mathrm{im}\,(T^*)$. 包含关系 $\subset$ 属显然. 至于 $\supset$, 给定 $T^*w \in \mathrm{im}\,(T^*)$, 取方程 $Tv = w$ 的最小二乘解 $v \in V$ 给出 $T^*w = T^*Tv \in \mathrm{im}\,(T^*T)$.

其次说明 $\ker(T^*T) = \ker(T)$. 包含关系 $\supset$ 属显然. 至于 $\subset$, 若 $T^*Tv = 0$ 则 $(Tv|Tv)_W = (T^*Tv \mid v)_V = 0$ 蕴涵 $Tv = 0$.

等式 $\mathrm{rk}\,(T^*T) = \mathrm{rk}(T)$ 是上一则等式和定理 4.8.4 的综合.

以上证明了左侧的三个等式. 鉴于 $T = (T^*)^*$ (命题 8.2.12), 将上述结果施于 T^* 便给出右侧三个等式. □

练习 9.6.5 沿用先前符号, 说明 $\ker(T^*) = \mathrm{im}(T)^\perp$, 因此 T^* 单等价于 T 满.

提示〉 等式 $T^*w = 0$ 等价于 $(Tv|w)_W = 0$ 对所有 $v \in V$ 成立.

9.7 对于正定二次型的应用

本节的第一部分探讨如何以行列式判断正定性, 这在维数较低时是很方便的手段, 其证明需要正交对角化.

定义 9.7.1 (顺序主子式) 设 $\boldsymbol{A} = (a_{ij})_{1\le i,j\le n}$ 为任意域上的 $n \times n$ 矩阵. 以下 n 个行列式

$$|a_{11}|, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ldots, \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

称为 $\boldsymbol{A}$ 的顺序主子式.

顺序主子式的定义依赖于有序基的选择. 相关定理的陈述和论证也因而需要使用矩阵语言.

定义 9.7.2 设实对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 对应到 n 元实二次型 f. 若 f 正定 (或半正定), 则称 $\boldsymbol{A}$ 是正定 (或半正定) 实对称矩阵.

定理 9.7.3 实对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 正定 (或半正定) 当且仅当 $\boldsymbol{A}$ 的所有特征值皆正 (或皆非负).

证明 记 $\boldsymbol{A}$ 对应的 n 元实二次型为 f. 正交对角化定理 9.5.2 给出

$$ {}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \qquad \begin{aligned} &\boldsymbol{C} \in \mathrm{M}_{n\times n}(\mathbb{R}),\ {}^{\mathrm{t}}\boldsymbol{C} = \boldsymbol{C}^{-1}, \\ &\lambda_1, \dots, \lambda_n \in \mathbb{R}:\ \boldsymbol{A} \text{ 的特征值, 记重数.} \end{aligned} $$

这相当于通过 $\boldsymbol{C}$ 作变量代换, 化 f 为 $\sum_{i=1}^{n} \lambda_i X_i^2$; 我们继续对满足 $\lambda_i \neq 0$ 的项作代换 $Y_i := \sqrt{|\lambda_i|}X_i$, 便能将二次型的系数化到 $\{-1, 0, 1\}$ 中. 然而这便足以说明 f 正定 (或半正定) 当且仅当 $\lambda_1, \dots, \lambda_n$ 全为正 (或非负). □

定理 9.7.4 (Sylvester 判准) 实对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 正定当且仅当 $\boldsymbol{A}$ 的所有顺序主子式都是正数.

证明 记 $\boldsymbol{A}$ 对应的 n 元实二次型为 f. 首先假设 f 是正定二次型. 惯性定理 8.6.6 蕴涵存在可逆的 $\boldsymbol{C} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 使得 ${}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{1}_{n\times n}$, 因此 $\det \boldsymbol{A} = (\det \boldsymbol{C})^{-2} > 0$.

其次, 注意到对于每个 $1 \leq m \leq r$, 对称矩阵

$$ \boldsymbol{A}_m := \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} $$

来自于 m 元二次型 $f_m(X_1, \dots, X_m) := f(X_1, \dots, X_m, 0, \dots, 0)$, 而 f_m 当然也正定. 将第一段的观察施于 $f_1, \dots, f_n$, 立见每个顺序主子式皆正.

现在论证另一方向. 注意到 $n = 1$ 情形是平凡的, 我们可以对 n 递归地论证, 以下皆设 $n \geq 2$. 假设 $\boldsymbol{A}$ 的每个顺序主子式皆正. 鉴于定理 9.7.3, 说明 $\boldsymbol{A}$ 的特征值皆正即可.

将 $\boldsymbol{A}$ 的特征值列作 $\lambda_1, \dots, \lambda_n$, 计重数; 等式

$$ \lambda_1 \cdots \lambda_n = \det \boldsymbol{A} > 0 $$

蕴涵负特征值若存在, 必然成对出现. 假定 $\boldsymbol{A}$ 有负特征值, 不妨设其为 λ_1 和 λ_2. 根据正交对角化定理 9.5.2, 相应的特征向量 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 可以取为 $\mathbb{R}^n$ 中相对于标准内积的正交单位向量. 对任意 $\alpha, \beta \in \mathbb{R}$,

$$ \begin{aligned} f(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) &= {}^{\mathrm{t}}(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2)\,\boldsymbol{A}\,(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) \\ &= \alpha^2\,({}^{\mathrm{t}}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_1) + \beta^2\,({}^{\mathrm{t}}\boldsymbol{v}_2\,\boldsymbol{A}\,\boldsymbol{v}_2) + 2\alpha\beta\,({}^{\mathrm{t}}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_2) \\ &= \alpha^2\lambda_1 + \beta^2\lambda_2 + 2\alpha\beta\lambda_2\,({}^{\mathrm{t}}\boldsymbol{v}_1\,\boldsymbol{v}_2) \\ &= \alpha^2\lambda_1 + \beta^2\lambda_2 \leq 0, \end{aligned} $$

最后的不等式仅在 $\alpha=0=\beta$ 时取等号.

易见存在不全为零的 α,β 使得 $\alpha\boldsymbol{v}_1+\beta\boldsymbol{v}_2$ 作为列向量的第 n 个坐标为 0, 从而它可以视同 $\mathbb{R}^{n-1}\simeq\mathbb{R}^{n-1}\times\{0\}\subset\mathbb{R}^n$ 的元素 $\boldsymbol{w}$. 按照证明第一步的符号, 我们有 $f(\alpha\boldsymbol{v}_1+\beta\boldsymbol{v}_2)=f_{n-1}(\boldsymbol{w})\leq 0$.

另一方面, $\boldsymbol{A}$ 的前 $n-1$ 个顺序主子式皆正, 所以 f_{n-1} 正定. 综上可得 $f_{n-1}(\boldsymbol{w})=0$, 从而 $\boldsymbol{w}=0$. 这和 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 线性无关相矛盾. □

本节的第二部分探讨平方根的矩阵版本, 这需要正定或半正定的前提. 以下采取线性映射的语言. 若无另外申明, 内积空间都是有限维的; V 上的内积记为 $(\cdot|\cdot)_V$ 或 $(\cdot|\cdot)$.

定义 9.7.5　设 $(V,(\cdot|\cdot))$ 为有限维内积空间. 若 $T\in\mathrm{End}(V)$ 自伴, 而且 $(v_1,v_2)\mapsto(Tv_1|v_2)$ 是 V 上的正定 (或半正定) 二次型, 则称 T 是正定 (或半正定) 的.

对于标准内积空间 $\mathbb{R}^n$ (列向量) 及对应于实对称矩阵 $\boldsymbol{A}$ 的自伴算子, 定义中考虑的二次型无非是 $(\boldsymbol{v}_1,\boldsymbol{v}_2)\mapsto{}^{\mathrm{t}}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_2$. 所以上述的正定 (或半正定) 概念和定义 9.7.2 是兼容的.

因此正定 (或半正定) 的 T 总能正交对角化, 而且其特征值皆正 (或非负). 兹介绍正定和半正定线性映射的一类基本例子.

引理 9.7.6　设 $T:V\to W$ 为内积空间之间的线性映射, 则 $T^*T\in\mathrm{End}(V)$ (或 $TT^*\in\mathrm{End}(W)$) 是半正定的; 若 T 单 (或 T^* 单, 等价说法是 T 满), 则它们是正定的.

证明　处理 T^*T 版本即可, 因为以 T^* 代 T 可得另一版本.

自伴性质已知. 半正定性是 $(T^*Tv\mid v)_V=(Tv|Tv)_W$ 的结论. 此外这还蕴涵 $(T^*Tv\mid v)_V=0\iff Tv=0$; 若 T 单则推得 $(T^*Tv\mid v)_V=0\iff v=0$, 故此时 T^*T 正定. □

定义–命题 9.7.7　设 $T\in\mathrm{End}(V)$ 正定 (或半正定), 则存在唯一的 $S\in\mathrm{End}(V)$ 使得 S 正定 (或半正定) 而且 $S^2=T$. 因此, 这样的 S 可以合理地记为 $\sqrt{T}$.

证明　存在性比较容易. 取由 T 的特征向量组成的单位正交基 $v_1,\dots,v_n$, 对应的特征值记为 $\lambda_1,\dots,\lambda_n$, 它们皆正 (或非负). 定义 S 使得 $Sv_i=\sqrt{\lambda_i}v_i$ 对 $i=1,\dots,n$ 成立, 则易见 S 正定 (或半正定), 并且 $S^2=T$.

现在考虑唯一性. 对于具备所需性质的 S, 特征值 (不记重数) 排列为 $\mu_1>\cdots>\mu_m\in\mathbb{R}_{\geq 0}$, 则 V 相应地分解为特征子空间的直和

$$V=\bigoplus_{i=1}^{m}V_{\mu_i}.$$

由于 $T=S^2$, 这说明 T 的特征值是 $\mu_1^2>\cdots>\mu_m^2$ (不记重数), 而 μ_i^2 对应的特征子空间正是上述 V_{μ_i}.

这反过来说明如何从 T 来确定 S: 将 T 的特征值 (不记重数) 排列为 $\lambda_1>\cdots>\lambda_m$, 则 S 在 T 的 λ_i-特征子空间上必然等于 $\sqrt{\lambda_i}\cdot\mathrm{id}$. □

作为平方根构造的应用, 本节最后给出线性映射的极分解定理.

定理 9.7.8 (极分解) 仍然设 $(V,(\cdot|\cdot))$ 为内积空间, 而 $T\in\mathrm{End}(V)$ 可逆, 则存在唯一一对 $R,U\in\mathrm{End}(V)$, 使得 R 正定 (定义 9.7.5), U 是正交变换, 而且 $T=RU$.

证明 首先说明 (R,U) 唯一. 我们有 $TT^*=RUU^*R=R^2$. 因为 T 可逆, 引理 9.7.6 确保 TT^* 和 R 同样正定, 从而 $R=\sqrt{TT^*}$; 因为 R 可逆, 必有 $U=R^{-1}T$.

下面处理存在性. 线索已经明朗: 必须取 $R:=\sqrt{TT^*}$, 这是正定的, 而 $U:=R^{-1}T$. 关键在于证 $U^*=U^{-1}$. 从 $\left(R^{-1}\right)^*=(R^*)^{-1}=R^{-1}$ (请回忆 (8.2.2)) 和 $R^2=TT^*$ 得到

$$\begin{aligned}U^*&=T^*\left(R^{-1}\right)^*=T^*R^{-1}\\&=T^{-1}TT^*(R^2)^{-1}R\\&=T^{-1}R=U^{-1}.\end{aligned}$$

明所欲证. □

极分解可设想为非零复数的极分解 $z=re^{\mathrm{i}\theta}$ 的矩阵类比, 其中 $r=|z|=\sqrt{z\bar z}$ 而 θ 是 z 的幅角.

9.8 奇异值分解

本节取 V 和 W 为 $\mathbb{R}$ 上的有限维内积空间, 仍将其内积分别记为 $(\cdot|\cdot)_V$ 和 $(\cdot|\cdot)_W$. 记 $m:=\dim V$, $n:=\dim W$.

对给定的线性映射 $T:V\to W$, 我们将需要推论 9.6.4 的内容:

$$\ker(T)=\ker\left(T^*T\right),\quad \mathrm{rk}(T)=\mathrm{rk}\left(T^*T\right);$$

此外, 引理 9.7.6 表明 $T^*T\in\mathrm{End}(V)$ 和 $TT^*\in\mathrm{End}(W)$ 都是自伴而且半正定的.

定理 9.8.1 (奇异值分解) 对于任意线性映射 $T:V\to W$, 记 $p:=\min\{m,n\}$, 则存在

- ⋆ V 的单位正交基 $v_1,\dots,v_m$,
- ⋆ W 的单位正交基 $w_1,\dots,w_n$,
- ⋆ 非负实数 $\sigma_1\geq\cdots\geq\sigma_p$,

使得

$$Tv_i=\begin{cases}\sigma_iw_i, & 1\leq i\leq p,\\ 0, & i>p.\end{cases}$$

此处的 $\sigma_1\geq\cdots\geq\sigma_p$ 由 T 唯一确定, 称为 T 的**奇异值**.

证明 先来说明 $\sigma_1\geq\cdots\geq\sigma_p$ 的唯一性. 首先验证

$$T^*w_j=\begin{cases}\sigma_jv_j, & 1\leq j\leq p,\\ 0, & j>p;\end{cases}$$

诚然, 易证上式确定的线性映射 T^* 确实符合伴随映射的条件 $(v_i|T^*w_j)_V = (Tv_i|w_j)_W$. 由此可见 $v_1,\dots,v_m$ 是 T^*T 的特征向量, 对应的特征值为 $\sigma_1^2,\dots\sigma_p^2,0,\dots,0$ (计重数). 按此唯一地确定 $\sigma_1 \geq \cdots \geq \sigma_p$.

这也启发存在性的证明. 给定 T, 对 T^*T 作正交对角化以得到单位正交基 $v_1,\dots,v_m \in V$ 和对应的特征值 $\lambda_1,\dots,\lambda_m \in \mathbb{R}_{\geq 0}$, 适当重排以确保 $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$. 注意到 $r := \mathrm{rk}(T) = \mathrm{rk}(T^*T) \leq p$, 非零项正是前 r 项. 对所有 $1 \leq i \leq p$ 定义 $\sigma_i := \sqrt{\lambda_i}$, 另外定义

$$w_i := \sigma_i^{-1} T v_i, \quad 1 \leq i \leq r.$$

兹断言 $w_1,\dots,w_r$ 是 W 中的单位正交向量族. 缘由在于

$$\begin{aligned}(w_i|w_j)_W &= (\sigma_i\sigma_j)^{-1}\,(Tv_i \mid Tv_j)_W \\ &= (\sigma_i\sigma_j)^{-1}\,(v_i \mid T^*Tv_j)_V \\ &= (\sigma_i\sigma_j)^{-1}\lambda_j (v_i|v_j)_V \\ &= \begin{cases} 1, & i=j, \\ 0, & i\neq j. \end{cases}\end{aligned}$$

依推论 9.3.9 将 $w_1,\dots,w_r$ 扩充为 W 的单位正交基 $w_1,\dots,w_n$. 观察到 $Tv_i = \sigma_i w_i$ 在 $1 \leq i \leq r$ 时按定义自动成立; 在 $i > r$ 时 $T^*Tv_i = 0$ 蕴涵 $Tv_i = 0$, 而 $r < i \leq p$ 时 $\sigma_i = 0$. 将这些信息组合起来便是所求等式. □

现在设 $V = \mathbb{R}^m$, $W = \mathbb{R}^n$, 各自配备标准内积, 并且将 T 等同于矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times m}(\mathbb{R})$. 对于定理 9.8.1 中的单位正交基, 以列向量定义正交矩阵

$$\boldsymbol{P} := (\boldsymbol{v}_1|\cdots|\boldsymbol{v}_m) \in \mathrm{M}_{m\times m}(\mathbb{R}), \quad \boldsymbol{Q} := (\boldsymbol{w}_1|\cdots|\boldsymbol{w}_n) \in \mathrm{M}_{n\times n}(\mathbb{R}),$$

另外用奇异值定义

$$\boldsymbol{\Sigma} := \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \boxed{\text{补 } 0} \end{pmatrix} \in \mathrm{M}_{n\times m}(\mathbb{R}),$$

其中补 0 的项仅在 $n > m$ 时出现. 奇异值分解化为矩阵等式

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{\Sigma},$$

亦即

$$^{\mathrm{t}}\boldsymbol{Q}\,\boldsymbol{A}\,\boldsymbol{P} = \boldsymbol{\Sigma} \quad \text{或} \quad \boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{\Sigma}\,{}^{\mathrm{t}}\boldsymbol{P}. \tag{9.8.1}$$

这是应用中常见的表达方式, 它所涉及的 $\boldsymbol{P}$, $\boldsymbol{Q}$ 和 $\boldsymbol{\Sigma}$ (亦即 A 的奇异值) 有稳定且高效的数值算法.

在实际应用场景中, $\boldsymbol{A}$ 代表混沌而庞大的数据集. 奇异值 $\sigma_1 \geq \sigma_2 \geq \cdots$ 经常能从 $\boldsymbol{A}$ 提取我们所关心的信息; 它们可以设想为 $\boldsymbol{A}$ 蕴藏的某种本质特征, 按其重要性由大到小排列. 比方说, 对于所有 $k \leq \mathrm{rk}(\boldsymbol{A})$, 秩 k 矩阵

$$\boldsymbol{A}_k := \boldsymbol{Q} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_k & \\ & & & \boxed{\text{补 } 0} \end{pmatrix} {}^{\mathrm{t}}\boldsymbol{P}$$

仅保留了 $\boldsymbol{A}$ 的前 k 个奇异值; $\boldsymbol{A}_k$ 可以设想为 $\boldsymbol{A}$ 的秩 k 逼近, 在应用中, 这往往体现了对 $\boldsymbol{A}$ 的信息进行压缩或降噪的一种手段. 实际上, 可以证明 $\boldsymbol{A}_k$ 在某种严格意义下是 $\boldsymbol{A}$ 的最优秩 k 逼近; 由于这已是数值分析的主题, 就此打住.

9.9 Moore–Penrose 广义逆

对于域 F 上的有限维向量空间 V 和 W, 尽管线性映射 $T : V \to W$ 一般不可逆, 应用中往往需要考虑 T^{-1} 的某种代替品, 其具体的定义和构造视用途而定. 一个最简单的版本是考虑线性映射 $S : W \to V$, 使得 $TS|_{\mathrm{im}(T)} = \mathrm{id}_{\mathrm{im}(T)}$; 等价地说, 我们要求 $TST = T$.

定义–命题 9.9.1 (广义逆) *设 V 和 W 是有限维 F-向量空间. 对任意线性映射 $T : V \to W$, 存在线性映射 $S : W \to V$ 使得 $TST = T$. 具此性质的 S 称为 T 的一个广义逆. 广义逆唯一当且仅当 T 可逆.*

证明 起手式是熟悉的, 取 V 的基 $v_1, \ldots, v_m$ 和 W 的基 $w_1, \ldots, w_n$, 使得

$$\begin{gathered} w_1, \ldots, w_r \text{ 是 } \mathrm{im}(T) \text{ 的基}, \quad r := \mathrm{rk}(T), \\ Tv_i = w_i, \quad 1 \leq i \leq r, \\ v_{r+1}, \ldots, v_m \text{ 是 } \ker(T) \text{ 的基}. \end{gathered}$$

广义逆的条件相当于 $T(Sw_i) = w_i$, 其中 $1 \leq i \leq r$, 这也相当于要求 $Sw_i \in v_i + \ker(T)$, 而当 $i > r$ 时 Sw_i 可任意指定. 满足上述条件的 S 当然存在, 而且其唯一性等价于 $\ker(T) = \{0\}$ 而且 $r = n$, 换言之, 等价于 T 可逆. □

当 T 可逆时, 广义逆的唯一选法是 T^{-1}.

对于 $F = \mathbb{R}$ 而 V 和 W 是内积空间的情形, 广义逆理论有更丰富的内涵. 以下将 V 和 W 上的内积统一记为 $(\cdot|\cdot)$, 以简化符号.

定义 9.9.2 (Moore–Penrose 广义逆) *相对于给定的 $T : V \to W$, 满足下述条件的线性映射 $S : W \to V$ 称为 T 的 Moore–Penrose 广义逆.*

MP.1 $TST = T$,

MP.2 $STS = S$,

MP.3 $TS = (TS)^*$,

MP.4 $ST = (ST)^*$.

留意到上述条件对于 S 和 T 是对称的. 应用中往往只考虑一部分的条件, 或者同时再另加性质; 只满足 **MP.1** 的 S 无非是定义–命题 9.9.1 所谓的广义逆. Moore–Penrose 广义逆的优点在于它存在且唯一.

构造的第一条线索是 §9.6 对方程 $Tv = w$ 定义的最小二乘解 v. 但我们不能简单地取 $Sw = v$, 因为最小二乘解 v 并不唯一; 然而陪集 $v + \ker(T)$ 是良定义的, 从中可取长度最短的向量; 根据勾股定理 (命题 9.2.2) 可知此向量是唯一的, 等于 v 向 $\ker(T)^\perp$ 的正交投影. 详述如下.

定理 9.9.3 给定 $T : V \to W$, Moore–Penrose 广义逆 $S : W \to V$ 存在而且唯一, 由以下构造给出: 对任意 $v \in V$ 和 $w \in W$, 我们有唯一分解

$$\begin{aligned} v &= v' + v'', \quad v' \in \ker(T), \quad v'' \in \ker(T)^\perp, \\ w &= w' + w'', \quad w' \in \operatorname{im}(T), \quad w'' \in \operatorname{im}(T)^\perp. \end{aligned}$$

对给定的 $w \in W$ 任取 $v \in T^{-1}(w')$, 则 $Sw = v''$.

特别地, Sw 是定义 9.6.1 介绍的一个最小二乘解 (相对于方程 $Tv = w$).

证明 首先验证以上给出的 S 是 Moore–Penrose 广义逆. 取 v 和 w 如上. 观察到 v'' 只依赖于陪集 $v + \ker(T) = T^{-1}(w')$, 所以 Sw 是良定义的. 请读者简单地检验 S 对 w 也是线性的.

对任意 $v \in V$, 在上述构造中代入 $w = Tv = w'$ 可得 $ST(v) = v''$; 另一方面, $TS(w) = Tv'' = Tv = w'$. 于是对所有 v 和 w 都有

$$\begin{aligned} TST(v) &= T(v'') = T(v), \\ STS(w) &= S(w') = S(w), \end{aligned}$$

这分别验证了 **MP.1** 和 **MP.2**.

前述观察也说明 TS 是正交投影 $W \to \operatorname{im}(T)$, 而 ST 是正交投影 $V \to \ker(T)^\perp$; 既然正交投影是自伴的, **MP.3** 和 **MP.4** 得证.

至于唯一性, 设 S 和 R 都是 Moore–Penrose 广义逆, 则

$$\begin{aligned} S \overset{\textbf{MP.2}}{=\!=} STS \overset{\textbf{MP.3}}{=\!=} S\,(TS)^* = S\,S^*\,T^* \overset{\textbf{MP.1}}{=\!=} S\,S^*\,(TRT)^* \\ = S\,S^*\,T^*\,R^*\,T^* = S\,(TS)^*\,(TR)^* \overset{\textbf{MP.3}}{=\!=} STSTR \overset{\textbf{MP.1}}{=\!=} STR. \end{aligned}$$

完全对称的论证给出 $R \overset{\textbf{MP.2}}{=\!=} RTR \overset{\textbf{MP.4}}{=\!=} (RT)^*\,R = \cdots = STR$, 因此 $S = R$. 明所欲证. □

在 §9.8 介绍的奇异值分解是计算 Moore–Penrose 广义逆的一种方便手段. 依旧考虑线性映射 $T : V \to W$, 并且记 $m := \dim V$, $n := \dim W$. 取奇异值分解 (定理 9.8.1) 包含的资料 $(v_i)_{i=1}^m$, $(w_j)_{j=1}^n$ 和奇异值 $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$, 其中 $p := \min\{m, n\}$.

命题 9.9.4 取定上述资料, 命 $r := \mathrm{rk}(T)$. 定义线性映射 $S : W \to V$ 使得

$$Sw_j = \begin{cases} \sigma_j^{-1} v_j, & 1 \leq j \leq r, \\ 0, & r < j \leq n. \end{cases}$$

则 S 是 T 的 Moore–Penrose 广义逆.

证明 给定 $w = \sum_{j=1}^n a_j w_j \in W$, 它在 $\mathrm{im}(T) = \langle w_1, \ldots, w_r \rangle$ 上的正交投影是 $w' = \sum_{j=1}^r a_j w_j$. 因此 $T^{-1}(w')$ 的元素皆形如

$$v = \sum_{i=1}^r a_i \sigma_i^{-1} v_i + \sum_{i=r+1}^m b_i v_i, \quad b_{r+1}, \ldots, b_m \in \mathbb{R}.$$

它在 $\ker(T)^\perp = \langle v_{r+1}, \ldots, v_m \rangle^\perp$ 上的正交投影无非是 $\sum_{j=1}^r a_j \sigma_j^{-1} v_j = Sw$. 明所欲证. □

以矩阵的语言表述如下. 取 $V = \mathbb{R}^m$ 和 $W = \mathbb{R}^n$, 赋予标准内积. 若 T 视为矩阵带有如 (9.8.1) 的奇异值分解

$$\boldsymbol{T} = \boldsymbol{Q} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \boxed{\text{补 } 0} \end{pmatrix}_{n \times m} {}^{\mathrm{t}}\boldsymbol{P},$$

则 $\boldsymbol{T}$ 的 Moore–Penrose 广义逆写作

$$\boldsymbol{S} = \boldsymbol{P} \begin{pmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & \boxed{\text{补 } 0} \end{pmatrix}_{m \times n} {}^{\mathrm{t}}\boldsymbol{Q}.$$

以上讨论限于实矩阵, 然而这些理论对于 §10.3 行将介绍的酉空间都有复版本, 唯一差别在于须以 Hermite 内积取代实内积, §10.5 将予以简要的说明.

9.10 极小化极大原理

设 $(V,(\cdot|\cdot))$ 是 n 维内积空间, $B:V\times V\to\mathbb{R}$ 是对称双线性形式. 一旦取定 V 相对于 $(\cdot|\cdot)$ 的单位正交基, 则不妨假定 $(V,(\cdot|\cdot))$ 是配备标准内积的 $\mathbb{R}^n$, 而 B 对应于对称矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$. 将 $\boldsymbol{A}$ 的 n 个特征值 (计重数) 递降列出:

$$\lambda_1\geqslant\cdots\geqslant\lambda_n.$$

这些特征值是内禀的量: 由于 $(\cdot|\cdot)$ 非退化, 例 8.2.9 表明存在唯一的 $S\in\mathrm{End}(V)$ 使得 B 可以通过内积写成

$$B(v_1,v_2)=(v_1\mid Sv_2),\quad v_1,v_2\in V$$

的形式. 既然 B 和 $(\cdot|\cdot)$ 都是对称的, $(v_1\mid Sv_2)=(Sv_1\mid v_2)$, 故 S 必然自伴, 而 $\boldsymbol{A}$ 就是 S 在一组单位正交基之下对应的矩阵.

命题 9.10.1 对于 (V,B) 如上, S 的极大特征值 λ_1 等于 $\max_{\|v\|=1}B(v,v)$, 极小特征值 λ_n 等于 $\min_{\|v\|=1}B(v,v)$.

证明 根据定理 9.5.2, 可取 V 的单位正交基 $v_1,\dots,v_n$ 使得 $Sv_i=\lambda_iv_i$ 对每个 $1\leqslant i\leqslant n$ 成立. 命 $v=\sum_{i=1}^n a_iv_i$. 若 $\|v\|=1$ 则

$$\lambda_n=\lambda_n\sum_{i=1}^n a_i^2\leqslant\left(B(v,v)=\sum_{i=1}^n a_i^2\lambda_i\right)\leqslant\lambda_1\sum_{i=1}^n a_i^2=\lambda_1.$$

左右两个不等式分别在 $a_n=1$ 和 $a_1=1$ 时取到等号. □

用数学分析的术语来说, $\|v\|=1$ 是 V 的紧子集, 而 $v\mapsto B(v,v)$ 既然是由二次多项式确定的, 自然是 V 上的连续函数, 所以所论的极大值和极小值确实存在.

由于 $\frac{B(v,v)}{\|v\|^2}=B\left(\frac{v}{\|v\|},\frac{v}{\|v\|}\right)$, 以上讨论的 $\max_{\|v\|=1}B(v,v)$ 可以等价地写作 $\max\limits_{v\neq0}\dfrac{B(v,v)}{\|v\|^2}$, 其中的商称为 **Rayleigh 商**.

注记 9.10.2 当 B 正定时, 上述极大值和极小值也可以直观地理解为嵌球问题. 命 $\mathcal{S}:=\{v\in V:B(v,v)\leqslant1\}$, 这是 V 中包含原点 0 的有界闭区域; 对于 2 维 (或 3 维) 情形, 它是椭圆 (或椭球). 对于任意 $r>0$, 我们有

$$\max_{\|v\|=1}B(v,v)\leqslant r\iff\max_{\|v\|=1}B\left(\frac{v}{\sqrt{r}},\frac{v}{\sqrt{r}}\right)\leqslant1\iff\left\{w\in V:\|w\|=\frac{1}{\sqrt{r}}\right\}\subset\mathcal{S}.$$

于是 $\mathcal{S}$ 所能包含的最大球面以 $\lambda_1^{-1/2}$ 为半径; 同理, 包含 $\mathcal{S}$ 的最小球面以 $\lambda_n^{-1/2}$ 为半径. 此处的球面都假设以 0 为球心.

其余特征值需要更精确的技术来确定, 称为有限维内积空间上的 Courant–Fischer **极小化极大原理**.

定理 9.10.3 (R. Courant, E. Fischer) 对于 (V,B) 如上, 对每个 $1\leq k\leq n$ 皆有

$$\lambda_k = \min_{\substack{U\subset V\\ (n-k+1)\text{-维子空间}}} \max_{\substack{v\in U\\ \|v\|=1}} B(v,v)$$

以及

$$\lambda_k = \max_{\substack{U\subset V\\ k\text{-维子空间}}} \min_{\substack{v\in U\\ \|v\|=1}} B(v,v).$$

证明 仍然按定理 9.5.2 取 V 的单位正交基 $v_1,\dots,v_n$ 使得 $Sv_i=\lambda_i v_i$.

先选定 k 来探讨第一个等式. 若 $\dim U = n-k+1$, 基于维数的理由, $U\cap\langle v_1,\dots,v_k\rangle$ 包含非零向量, 记为 $v=\sum_{i=1}^k a_i v_i$, 伸缩后不妨假设 $\|v\|=1$. 我们有

$$B(v,v)=(v|Sv)=\sum_{i=1}^k \lambda_i a_i^2 \geq \lambda_k \sum_{i=1}^k a_i^2 = \lambda_k.$$

因此 $\max\limits_{\substack{v\in U\\ \|v\|=1}} B(v,v)\geq\lambda_k$, 从而

$$\inf_{\dim U=n-k+1} \max_{\substack{v\in U\\ \|v\|=1}} B(v,v)\geq\lambda_k. \tag{9.10.1}$$

然而若取 $U=\langle v_k,\dots,v_n\rangle$, 则对于满足 $\|v\|=1$ 的 $v=\sum_{i=k}^n a_i v_i\in U$, 类似的论证导致

$$B(v,v)=\sum_{i=k}^n \lambda_i a_i^2 \leq \lambda_k \sum_{i=1}^k a_i^2 = \lambda_k,$$

于是对此 U 有 $\max\limits_{\substack{v\in U\\ \|v\|=1}} B(v,v)\leq\lambda_k$.

综上可知 (9.10.1) 实则是等号, 而且它左侧的 inf 确实被 $U=\langle v_k,\dots,v_n\rangle$ 取到, 故可合理地改写成 min.

第二个等式的论证全然类似, 更直截了当的方法则是以 $-B$ 代 B 来相互过渡. □

极小化极大原理有种种不同场景下的推广, 它还可以用来求任意线性映射 $T:V\to W$ 的奇异值, 其中 V 和 W 是有限维内积空间; 见定理 9.8.1. 这是因为 T 的奇异值 $\sigma_1,\dots,\sigma_p$ 无非是自伴线性映射 $T^*T:V\to V$ 的前 p 个特征值的平方根.

练习 9.10.4 设 $T:V\to W$ 为有限维内积空间之间的线性映射, 它的奇异值记为 $\sigma_1,\dots,\sigma_p$, 其中 $p:=\min\{\dim V,\dim W\}$. 为了使陈述干净, 将奇异值序列补零扩展为 $\sigma_1\geq\ldots\geq\sigma_{\dim V}$. 这使得奇异值分解中的 $Tv_i=\sigma_i w_i$ 对所有 i 成立. 证明

$$\sigma_1^2=\max_{v\neq0}\frac{\|Tv\|}{\|v\|},\quad \sigma_{\dim V}^2=\min_{v\neq0}\frac{\|Tv\|}{\|v\|}.$$

提示 在先前讨论中代入 $S=T^*T$, 亦即取 $B(v,v')=(Tv|Tv')$.

9.11 Perron–Frobenius 定理

本节探讨的主题是一类非负矩阵的特征值和特征向量, 这些结果源于基础数学中的连分数理论, 其应用却遍及经济学, 网页搜索乃至流行病学等领域; 这类应用场景中出现的矩阵元经常都是正实数, 或者至少是非负的.

本节涉及的 m, n 均默认为正整数.

约定 9.11.1 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{m\times n}(\mathbb{R})$. 符号 $\boldsymbol{A} \geq \boldsymbol{B}$ (或 $\boldsymbol{A} > \boldsymbol{B}$) 意谓 $a_{ij} \geq b_{ij}$ (或 $a_{ij} > b_{ij}$) 对所有 (i,j) 皆成立. 关系式 $\boldsymbol{A} \geq \boldsymbol{0}_{m\times n}$ (或 $\boldsymbol{A} > \boldsymbol{0}_{m\times n}$) 也写作 $\boldsymbol{A} \geq 0$ (或 $\boldsymbol{A} > 0$).

注意: 上述符号不应和一些书籍中关于正定矩阵的符号混淆.

今后将 $\mathbb{C}^n$ 的元素等同于列向量, 依此谈论 $\mathbb{R}^n \subset \mathbb{C}^n$ 的元素是否 ≥ 0 或 > 0. 对于任意 $\boldsymbol{v} \in \mathbb{C}^n$, 记其第 i 个分量为 $v_i \in \mathbb{C}$.

引理 9.11.2 设 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(\mathbb{R})$ 而 $\boldsymbol{x} \in \mathbb{R}^n$. 若 $\boldsymbol{A} > 0$ 而 $\boldsymbol{x} \geq 0$ 而且 $\boldsymbol{x} \neq \boldsymbol{0}$, 则 $\boldsymbol{A}\boldsymbol{x} > 0$. 若仅要求 $\boldsymbol{A} \geq 0$, 则 $\boldsymbol{A}\boldsymbol{x} \geq 0$.

证明 将 $\boldsymbol{A}\boldsymbol{x}$ 的第 i 个分量表作 $\sum_{j=1}^{n} a_{ij}x_j$. 按假设, $a_{ij} > 0$, $x_j \geq 0$, 而且必有某个 j 使得 $x_j > 0$, 故 $\sum_{j=1}^{n} a_{ij}x_j > 0$. 至于 $\boldsymbol{A} \geq 0$ 的情形也是类似的. □

今后谈及特征值时, 均在复数域中考虑.

定义 9.11.3 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$, 命

$$\rho(\boldsymbol{A}) := \max\{|\lambda| : \lambda \in \mathbb{C} \text{ 是 } \boldsymbol{A} \text{ 的特征值}\},$$

称之为 $\boldsymbol{A}$ 的**谱半径**.

引理 9.11.4 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$, $\boldsymbol{A} > 0$, 则

(i) 存在 $\rho > 0$ 和 $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v} > 0$ 使得 $\boldsymbol{A}\boldsymbol{v} = \rho\boldsymbol{v}$;

(ii) 承上, 这样的 ρ 可以取为谱半径 $\rho(\boldsymbol{A})$.

特别地, $\rho(\boldsymbol{A}) > 0$.

证明 命 $S := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| = 1,\ \boldsymbol{x} \geq 0\}$, 这是 $\mathbb{R}^n$ 的紧子集. 定义

$$\begin{aligned} \mathcal{L} : S &\to \mathbb{R}_{>0} \\ x &\mapsto \min\left\{\frac{\boldsymbol{A}\boldsymbol{x} \text{ 的第 } i \text{ 个分量}}{x_i} : 1 \leq i \leq n,\ x_i \neq 0\right\}, \end{aligned}$$

这是连续映射; 特别地, 可以谈论 $\mathcal{L}$ 在 S 上的极值. 命 $\rho \in \mathbb{R}_{>0}$ 为 $\mathcal{L}$ 的极大值, 它被某个 $\boldsymbol{v} \in S$ 取到. 兹断言 $\boldsymbol{A}\boldsymbol{v} = \rho\boldsymbol{v}$ 而且 $\boldsymbol{v} > 0$.

首先证明 $\boldsymbol{A}\boldsymbol{v}=\rho\boldsymbol{v}$. 对满足 $v_i\neq 0$ 的 $1\leq i\leq n$ 比较 $\boldsymbol{A}\boldsymbol{v}$ 的第 i 个分量和 v_i, 由 $\mathcal{L}(\boldsymbol{v})=\rho$ 和 $\mathcal{L}$ 的定义可见 $\boldsymbol{A}\boldsymbol{v}\geq\rho\boldsymbol{v}$. 假若 $\boldsymbol{A}\boldsymbol{v}\neq\rho\boldsymbol{v}$, 则引理 9.11.2 确保 $\boldsymbol{A}(\boldsymbol{A}\boldsymbol{v}-\rho\boldsymbol{v})>0$. 因此可取充分小的正实数 ϵ 使得 $\boldsymbol{A}(\boldsymbol{A}\boldsymbol{v}-\rho\boldsymbol{v})>\epsilon\boldsymbol{A}\boldsymbol{v}$. 另一方面, 引理 9.11.2 同样确保 $\boldsymbol{A}\boldsymbol{v}>0$, 因此存在 $t>0$ 使得 $\boldsymbol{w}:=t\boldsymbol{A}\boldsymbol{v}\in S$.

将上述不等式改写为 $\boldsymbol{A}(\boldsymbol{A}\boldsymbol{v})>(\rho+\epsilon)\boldsymbol{A}\boldsymbol{v}$, 再调整为 $\boldsymbol{A}\boldsymbol{w}>(\rho+\epsilon)\boldsymbol{w}$, 由此可得 $\mathcal{L}(\boldsymbol{w})>(\rho+\epsilon)$, 与 ρ 的取法矛盾. 因此必有 $\boldsymbol{A}\boldsymbol{v}=\rho\boldsymbol{v}$. 既然已说明 $\boldsymbol{A}\boldsymbol{v}>0$, 故 $\boldsymbol{v}=\rho^{-1}\boldsymbol{A}\boldsymbol{v}>0$. 综上得到 (i).

谱半径的定义直接蕴涵 (i) 中的 ρ 满足 $\rho\leq\rho(\boldsymbol{A})$, 特别地, $\rho(\boldsymbol{A})>0$. 下面对此 ρ 证明 $\rho=\rho(\boldsymbol{A})$, 从而完成 (ii) 的证明.

设 $\mu\in\mathbb{C}$ 和 $\boldsymbol{w}\in\mathbb{C}^n\smallsetminus\{0\}$ 满足 $\boldsymbol{A}\boldsymbol{w}=\mu\boldsymbol{w}$. 对每个 $1\leq i\leq n$ 都有

$$|\mu|\cdot|w_i|=|(\mu\boldsymbol{w})_i|=\left|\sum_{j=1}^{n}a_{ij}w_j\right|\leq\sum_{j=1}^{n}a_{ij}|w_j|.$$

命 $\boldsymbol{w}':=(|w_1|,\dots,|w_n|)\in\mathbb{R}^n$, 则上式表明 $\boldsymbol{A}\boldsymbol{w}'\geq|\mu|\cdot\boldsymbol{w}'$. 将特征向量 $\boldsymbol{w}$ 适当伸缩后不妨假设 $\|\boldsymbol{w}'\|=1$; 此时 $\boldsymbol{w}'\in S$, 而上式说明 $\mathcal{L}(\boldsymbol{w}')\geq|\mu|$. 根据之前 ρ 的取法, 必然有 $|\mu|\leq\rho$. 这足以说明 $\rho(\boldsymbol{A})\leq\rho$, 从而 $\rho(\boldsymbol{A})=\rho$. □

注记 9.11.5 上述证明中途获得的公式 $\rho(\boldsymbol{A})=\max_{\boldsymbol{v}\in S}\mathcal{L}(\boldsymbol{v})$ 称为 Collatz–Wielandt 公式; 我们实际说明了若 $\boldsymbol{w}\in S$ 使 $\mathcal{L}(\boldsymbol{w})$ 取极大值, 则 $\boldsymbol{A}\boldsymbol{w}=\rho(\boldsymbol{A})\boldsymbol{w}$. 这是一则很有用的结果.

定理 9.11.6 (O. Perron, 1907) 设 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$, $\boldsymbol{A}>0$. 以下性质成立:

(i) $\rho(\boldsymbol{A})>0$, 而且存在 $\boldsymbol{v}\in\mathbb{R}^n$ 使得 $\boldsymbol{v}>0$ 而 $\boldsymbol{A}\boldsymbol{v}=\rho(\boldsymbol{A})\boldsymbol{v}$;

(ii) 若 $\mu\in\mathbb{C}$ 是 A 的特征值, $\mu\neq\rho(\boldsymbol{A})$, 则 $|\mu|<\rho(\boldsymbol{A})$;

(iii) 对应 $\rho(\boldsymbol{A})$ 的特征子空间是 1 维的;

(iv) $\rho(\boldsymbol{A})$ 是特征多项式 $\mathrm{Char}_{\boldsymbol{A}}$ 的单根.

证明 断言 (i) 是引理 9.11.4 的内容. 下面证明 (ii). 设 $\boldsymbol{w}\in\mathbb{C}^n$ 是对应 μ 的特征向量, $\|\boldsymbol{w}\|=1$, 并且假设 $|\mu|=\rho(\boldsymbol{A})$, 我们的目的是证 $\mu=\rho(\boldsymbol{A})$. 在引理 9.11.4 证明中, 我们实际论证了 $\boldsymbol{w}':=(|w_1|,\dots,|w_n|)\in S$ 必然满足

$$\rho(\boldsymbol{A})=\max_{v\in S}\mathcal{L}(\boldsymbol{v})\geq\mathcal{L}(\boldsymbol{w}')\geq|\mu|=\rho(\boldsymbol{A}),$$

从而 $\mathcal{L}$ 在 $\boldsymbol{w}'$ 处取极大值, 而 $\boldsymbol{A}\boldsymbol{w}'=\rho(\boldsymbol{A})\boldsymbol{w}'$. 这就说明了对所有 $1\leq i\leq n$ 皆有

$$\sum_{j=1}^{n}a_{ij}|w_j|=\rho(\boldsymbol{A})|w_i|=|\mu w_i|=\left|\sum_{j=1}^{n}a_{ij}w_j\right|.$$

使复平面上的三角不等式 $|z+z'| \leq |z|+|z'|$ 为等号的充要条件是 z, z' 落在同一条发自原点的射线上. 反复运用此性质, 可知 $w_1, \dots, w_n$ 也落在同一条射线上. 取 $c \in \mathbb{C}$ 为此射线上的任意非零元, 则 $c^{-1}\boldsymbol{w} \in \mathbb{R}^n$ 非零, 而 $c^{-1}\boldsymbol{w} \geq 0$. 但这就表明对应的特征值 μ 必为正实数. 故 (ii) 得证.

接着讨论 (iii). 设 $\boldsymbol{v}, \boldsymbol{v}' \in \mathbb{R}^n$ 都是以 $\rho(\boldsymbol{A})$ 为特征值的特征向量, 而且 $\boldsymbol{v} > 0$. 今将证明 $\boldsymbol{v}'$ 必然是 $\boldsymbol{v}$ 的倍数. 不失一般性, 可设 $\boldsymbol{v}'$ 至少有一个分量为正数. 假若 $\boldsymbol{v}'$ 和 $\boldsymbol{v}$ 不成比例, 则可取 $\epsilon > 0$ 使得 $\boldsymbol{v} - \epsilon\boldsymbol{v}' \geq 0$, 其中至少有一个分量为 0, 但 $\boldsymbol{v} - \epsilon\boldsymbol{v}' \neq \boldsymbol{0}$. 然而引理 9.11.2 蕴涵

$$\boldsymbol{v} - \epsilon\boldsymbol{v}' = \rho(\boldsymbol{A})^{-1}\boldsymbol{A}\,(\boldsymbol{v} - \epsilon\boldsymbol{v}') > 0,$$

这和 ϵ 的取法矛盾.

最后证明 (iv). 不妨假设 $n \geq 2$, 否则问题是平凡的. 取定对应 $\rho(\boldsymbol{A})$ 的特征向量 $\boldsymbol{v} > 0$. 因为 $\operatorname{Char}_{\boldsymbol{A}} = \operatorname{Char}_{{}^{\mathrm{t}}\boldsymbol{A}}$, 从而 $\rho(\boldsymbol{A}) = \rho({}^{\mathrm{t}}\boldsymbol{A})$; 此外 ${}^{\mathrm{t}}\boldsymbol{A} > 0$, 所以存在 $\boldsymbol{u} \in \mathbb{R}^n$ 使得 $\boldsymbol{u} > 0$ 而 ${}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{u} = \rho(\boldsymbol{A})\boldsymbol{u}$. 容易看出

$$\langle \boldsymbol{u} \rangle^{\perp} = \{\boldsymbol{x} \in \mathbb{R}^n : {}^{\mathrm{t}}\boldsymbol{u} \cdot \boldsymbol{x} = 0\}$$

是 $n-1$ 维 $\boldsymbol{A}$-不变子空间, 而且 ${}^{\mathrm{t}}\boldsymbol{u} \cdot \boldsymbol{v} > 0$ 故 $\boldsymbol{v} \notin \langle \boldsymbol{u} \rangle^{\perp}$, 所以 $\mathbb{R}^n = \mathbb{R}\boldsymbol{v} \oplus \langle \boldsymbol{u} \rangle^{\perp}$ 是 $\boldsymbol{A}$-不变直和分解. 按线性映射的观点定义

$$\boldsymbol{B} := \boldsymbol{A}|_{\langle \boldsymbol{u} \rangle^{\perp}}, \quad \text{取基后可视同 } \mathrm{M}_{(n-1)\times(n-1)}(\mathbb{R}) \text{ 的元素.}$$

我们有 $\operatorname{Char}_{\boldsymbol{A}} = (X - \rho(\boldsymbol{A}))\operatorname{Char}_{\boldsymbol{B}}$. 假如 $\rho(\boldsymbol{A})$ 是 $\operatorname{Char}_{\boldsymbol{A}}$ 的重根, 则 $\rho(\boldsymbol{A})$ 也是 $\boldsymbol{B}$ 的特征值, 相应地存在特征向量 $\boldsymbol{v}'$. 但根据上述直和分解, $\boldsymbol{v}'$ 也是 $\boldsymbol{A}$ 的特征向量, 并且与 $\boldsymbol{v}$ 线性无关, 这同 (iii) 矛盾. 定理得证. □

可以证明所有满足 $\boldsymbol{v} \geq \boldsymbol{0}$ 的特征向量 $\boldsymbol{v}$ 都以 $\rho(\boldsymbol{A})$ 为特征值, 一般性的陈述请见练习 9.11.10.

现在来尝试将条件 $\boldsymbol{A} > 0$ 放宽为 $\boldsymbol{A} \geq 0$.

定理 9.11.7 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 满足 $\boldsymbol{A} \geq 0$, 而且存在 $m \geq 1$ 使得 $\boldsymbol{A}^m > 0$, 则 Perron 定理 9.11.6 的断言 (i) — (iv) 对 $\boldsymbol{A}$ 仍然成立.

证明 枚举 $\boldsymbol{A}$ 的特征值 $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, 计重数. 不妨要求 $|\lambda_1| = \rho(\boldsymbol{A})$. 于是 $\boldsymbol{A}^m$ 的特征值是 $\lambda_1^m, \dots, \lambda_n^m$ (见推论 7.3.6). 因为 $|\lambda_1^m| = \rho(\boldsymbol{A}^m)$, Perron 定理 9.11.6 于是蕴涵 $\lambda_1^m > 0$; 因此也有 $\rho(\boldsymbol{A}) > 0$. 这是断言 (i) 的第一部分.

其次, 取 $\boldsymbol{v} \in \mathbb{C}^n \smallsetminus \{0\}$ 使得 $\boldsymbol{A}\boldsymbol{v} = \lambda_1\boldsymbol{v}$, 则从 $\boldsymbol{A}^m\boldsymbol{v} = \lambda_1^m\boldsymbol{v}$ 可知 $\boldsymbol{v}$ 和某个 > 0 的向量成比例, 适当伸缩后不妨就设 $\boldsymbol{v} > 0$. 于是 $\lambda_1\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} \geq 0$ 连同 $|\lambda_1| = \rho(\boldsymbol{A})$ 确保 $\lambda_1 = \rho(\boldsymbol{A})$.

Perron 定理也蕴涵 $i > 1$ 时 $|\lambda_i^m| < |\lambda_1^m|$, 所以 $|\lambda_i| < \rho(\boldsymbol{A})$, 故 (ii) 和 (iv) 成立. 根据定理 7.4.5, (iv) 又蕴涵 (iii). 至此证出所需的全部性质. □

定理 9.11.7 的条件和称为**有向图**的结构相关. 有向图意谓资料 (V, E, s, t):

⋆ 集合 V 的元素称为图的顶点;

⋆ 集合 E 的元素称为图的边;

⋆ s 和 t 都是从 E 到 V 的映射, 对于每个边 $e \in E$, 称 $s(e)$ (或 $t(e)$) 为 e 的起点 (或终点).

为了理解这何以称为有向图, 宜将 V 的元素设想为点, 将 E 的元素设想为其间的箭头 $s(e) \xrightarrow{e} t(e)$; 注意到这里容许有满足 $s(e) = t(e)$ 的边 $e \in E$.

⋆ 顶点和边个数有限的有向图称为**有限**的.

⋆ 若对于所有 $(i, j) \in V^2$, 存在一列边

$$i \xrightarrow{e_1} \cdots \xrightarrow{e_k} j, \quad k \in \mathbb{Z}_{\geq 1},$$

则称此有向图**连通**.

给定 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$, $\boldsymbol{A} \geq 0$. 定义有限有向图 $G(\boldsymbol{A}) := (V, E, s, t)$ 如下:

$$\begin{gathered} V := \{1, \ldots, n\}, \\ E := \{(i, j) : a_{ij} > 0\}, \\ \forall (i, j) \in E, \quad s(i, j) := i, \quad t(i, j) = j. \end{gathered}$$

观察到 $G(\boldsymbol{A})$ 的任两个顶点 v, v' 之间至多仅有一条边 $v \xrightarrow{e} v'$. 例如:

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \implies G(\boldsymbol{A}) = \left[1 \rightleftarrows 2 \circlearrowright \right].$$

由于 $\boldsymbol{A}^m$ 的 (i, k) 矩阵元是 $\sum_{j_1,\ldots,j_{m-1}} a_{ij_1} a_{j_1 j_2} \cdots a_{j_{m-1}k}$, 定理 9.11.7 中的条件 $\boldsymbol{A}^m > 0$ 相当于说 $G(\boldsymbol{A})$ 的任两个顶点 i, k 都由 m 条边相连: $i \to j_1 \to \cdots \to j_{m-1} \to k$. 这蕴涵 $G(\boldsymbol{A})$ 连通, 然而 m 不依赖 (i, k), 所以条件 $\boldsymbol{A}^m > 0$ 强于 $G(\boldsymbol{A})$ 的连通性. 对于仅要求连通性的情形, 定理 9.11.7 有如下版本.

定理 9.11.8 (G. Frobenius, 1912) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 满足 $\boldsymbol{A} \geq 0$, 而且对应的有向图 $G(\boldsymbol{A})$ 连通. 以下性质成立:

(i) $\rho(\boldsymbol{A}) > 0$;

(ii) 存在 $\boldsymbol{v} \in \mathbb{R}^n$ 使得 $\boldsymbol{v} > 0$ 而 $\boldsymbol{A}\boldsymbol{v} = \rho(\boldsymbol{A})\boldsymbol{v}$;

(iii) 对应 $\rho(\boldsymbol{A})$ 的特征子空间是 1 维的;

(iv) $\rho(\boldsymbol{A})$ 是特征多项式 $\mathrm{Char}_{\boldsymbol{A}}$ 的单根.

定理证明具有一定的技巧性, 然而终究是初等的, 留作本章习题供读者赏玩; 习题也将对 $\boldsymbol{A}$ 的特征值给出更多信息.

练习 9.11.9 验证以下矩阵所对应的 $G(\boldsymbol{A})$ 连通, 然而不存在 m 使得 $\boldsymbol{A}^m > 0$.

$$\boldsymbol{A} = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix}$$

对之计算 $\rho(\boldsymbol{A})$ 和对应的特征向量.

练习 9.11.10 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$. 证明若 ${}^{\mathrm{t}}\boldsymbol{A}$ 有特征值记为 ρ 的特征向量 $\boldsymbol{u} > 0$, 则 $\boldsymbol{A}$ 的所有满足 $\boldsymbol{v} \geq 0$ 的特征向量 $\boldsymbol{v}$ 都以 ρ 为特征值. 以此结果强化本节的定理.

提示〉 设 $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. 从 $\rho\, {}^{\mathrm{t}}\boldsymbol{u}\,\boldsymbol{v} = {}^{\mathrm{t}}\boldsymbol{u}\,\boldsymbol{A}\,\boldsymbol{v} = \lambda\, {}^{\mathrm{t}}\boldsymbol{u}\,\boldsymbol{v}$ 和 ${}^{\mathrm{t}}\boldsymbol{u}\,\boldsymbol{v} > 0$ 推导 $\lambda = \rho$.

习题

1. 考虑配备标准内积的 $\mathbb{R}^n$. 对有序基 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$ 取其 Gram 矩阵 $\boldsymbol{A}$ (见命题 9.2.2 之后的讨论). 证明 $\left\{\sum_{i=1}^n t_i\boldsymbol{v}_i : \forall i,\ t_i \in [0,1]\right\}$ 的体积是 $\sqrt{\det \boldsymbol{A}}$.
   提示〉 命 $\boldsymbol{P} = \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{array}\right)$, 则 $\boldsymbol{A} = {}^{\mathrm{t}}\boldsymbol{P}\,\boldsymbol{P}$ 而 $\det \boldsymbol{A} = (\det \boldsymbol{P})^2$.
2. 在配备标准内积的空间 $\mathbb{R}^4$ 中, 设
   $$\boldsymbol{v}_1 = (1,1,0,0), \quad \boldsymbol{v}_2 = (1,0,1,0), \quad \boldsymbol{v}_3 = (1,0,0,-1).$$
   以 Gram–Schmidt 正交化为它们生成的子空间求出正交基, 不必化为单位向量.
3. 证明实对称正定矩阵 (定义 9.7.2) 的逆仍然是实对称正定矩阵.
4. 证明 $n \times n$ 实对称矩阵 $\boldsymbol{A}$ 半正定 (或正定) 的充要条件是存在实对称 (或可逆实对称) 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A} = \boldsymbol{C}^2$.
5. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 的第 (i,j) 个矩阵元为 $\min\{i,j\}$. 以定理 9.7.4 说明 $\boldsymbol{A}$ 正定.
6. 对所有 $n \in \mathbb{Z}_{\geq 1}$ 考虑实对称 $n \times n$ 矩阵
   $$\boldsymbol{A}_n = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$
   证明 $\boldsymbol{A}_n$ 正定.

7. 设 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\in\mathbb{R}^n$, 视同列向量. 命 $\boldsymbol{A}:=(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)\in \mathrm{M}_{n\times n}(\mathbb{R})$, 证明

$$|\det \boldsymbol{A}|\leq \prod_{i=1}^{n}\|\boldsymbol{v}_i\|,$$

其中 $\|\cdot\|$ 是相对于 $\mathbb{R}^n$ 的标准内积的长度. 试明确等号成立的充要条件.

提示〉 不妨设 $\boldsymbol{A}$ 可逆. 一种思路是应用 Gram–Schmidt 正交化定理 9.3.5. 另一种思路是对 $\boldsymbol{B}:=\boldsymbol{A}^*\boldsymbol{A}$ 应用第八章习题的不等式 $\det \boldsymbol{B}\leq b_{11}\cdots b_{nn}$.

8. (Rodrigues 公式) 定义 $\mathbb{R}[X]$ 的一列元素

$$P_0:=1,\quad P_n:=\frac{1}{2^n n!}\left((X^2-1)^n\right)^{(n)},\quad n\in\mathbb{Z}_{\geq 1}.$$

(i) 证明 $P_n=\dfrac{(2n)!}{2^n(n!)^2}\cdot X^n+$低次项.

(ii) 证明 $P_n(1)=1$.

提示〉 应用练习 6.4.7 的 Leibniz 律对 $(X^2-1)^n$ 求导 n 次.

(iii) 证明当 $0\leq k<n$ 时 $\int_{-1}^{1}t^kP_n(t)\,\mathrm{d}t=0$. 由此说明当 $i\neq j$ 时 $\int_{-1}^{1}P_i(t)P_j(t)\,\mathrm{d}t=0$.

提示〉 之前的方法指明当 $m<n$ 时对 $(X^2-1)^n$ 求导 m 次给出 X^2-1 的倍式, 从而在 ±1 处皆取零. 按此分部积分.

(iv) 基于上述结果, 说明 $P_n=\dfrac{(2n)!}{2^n(n!)^2}\cdot w_n$, 其中 $w_0,w_1,\dots$ 是例 9.3.16 介绍的 Legendre 多项式. 由此得出 $w_n(1)\neq 0$.

(v) 说明 $P_n(-t)=(-1)^nP_n(t)$.

(vi) 证明 $\int_{-1}^{1}P_n(t)^2\,\mathrm{d}t=\dfrac{2}{2n+1}$.

提示〉 以下事实可能有帮助:

$$\int_0^1\left(1-t^2\right)^n\mathrm{d}t=\int_0^{\pi/2}\sin(t)^{2n+1}\,\mathrm{d}t=\frac{2n(2n-2)\cdots 2}{(2n+1)(2n-1)\cdots 1}.$$

9. 承上题, 所有无穷次可微的实值函数构成无穷维 $\mathbb{R}$-向量空间 $\mathcal{C}$. 考虑线性映射

$$S:\mathcal{C}\to\mathcal{C},$$
$$(Sf)(t)=(t^2-1)f''(t)+2tf'(t).$$

(i) 说明 S 相对于内积 $(f|g):=\int_{-1}^{1}f(t)g(t)\,\mathrm{d}t$ 具有自伴性质 $(Sf|g)=(f|Sg)$.

(ii) 说明 S 映多项式函数 P_n 为 $n(n+1)P_n$.

提示〉 对 $(X^2-1)\cdot\left((X^2-1)^n\right)'=2nX(X^2-1)^n$ 两边求导 $n+1$ 次.

10. 证明递归关系式 $(n+1)P_{n+1}=(2n+1)XP_n-nP_{n-1}$, 其中 $n\in\mathbb{Z}_{\geq 1}$.

11. 继续关于实多项式 $P_0,P_1,\dots$ 的讨论. 对所有 $1\leq k\leq n$, 命 $f_k:=\left((X^2-1)^n\right)^{(k)}$.

(i) 应用 Rolle 定理说明 f_1 在开区间 $(-1,1)$ 中有根 c, 事实上可取 $c=0$.

(ii) 继续说明 f_2 在 $(-1,c)$ 和 $(c,1)$ 中分别有根 d 和 e. 然后对 $f_3,\dots,f_n$ 依此类推. 以此说明 P_n 有 n 个相异实根, 全部落在区间 $(-1,1)$. 提示〉 复用之前习题证明的 $(X^2-1)\mid f_k$, 其中 $k<n$.

Legendre 多项式的这些根在数值分析中起作用, 它们是 Gauss–Legendre 求积公式的关键要素.

12. 对所有无穷次可微的实值函数 f,g, 定义

$$[\![f|g]\!] := \int_{-1}^{1} \frac{f(t)g(t)}{\sqrt{1-t^2}}\,\mathrm{d}t.$$

验证积分收敛, 并且在 $f,g\in\mathbb{R}[X]$ 的情形给出 $\mathbb{R}[X]$ 上的内积. 这是先前的 $(f|g)$ 的加权版本. 提示〉 作三角代换 $t=\cos x$.

13. 递归地定义**第一类 Chebyshev 多项式**为 $\mathbb{R}[X]$ 的元素

$$\begin{aligned} T_0 &= 1, \quad T_1 = X, \\ T_{n+1} &= 2XT_n - T_{n-1}, \quad n \geq 1. \end{aligned}$$

(i) 具体写下 T_2, T_3, T_4, T_5. 证明 $T_n(-X)=(-1)^nT_n(X)$, 以及当 $n\in\mathbb{Z}_{\geq 1}$ 时

$$T_n = 2^{n-1}X^n + \text{低次项}.$$

(ii) 证明 $T_n(\cos x)=\cos(nx)$.

(iii) 证明当 $t\in(-1,1)$ 时

$$T_n(t) = \frac{(-2)^n n!}{(2n)!}\sqrt{1-t^2}\left((1-t^2)^{n-\frac{1}{2}}\right)^{(n)}.$$

(iv) 基于上一道习题的符号, 证明

$$[\![T_i|T_j]\!] = \begin{cases} 0, & i\neq j, \\ \pi/2, & i=j\neq 0, \\ \pi, & i=j=0. \end{cases}$$

依此说明 $T_0,T_1,T_2,\dots$ 与对 $1,X,X^2,\dots$ 和内积 $[\![\cdot|\cdot]\!]$ 作 Gram–Schmidt 正交化的产物有何联系. 提示〉 作三角代换 $t=\cos x$ 并应用 $T_n(\cos x)=\cos(nx)$.

(v) 证明 P_n 在开区间 $(-1,1)$ 中有 n 个相异实根.

14. 考虑从无穷次可微的实值函数空间 $\mathcal{C}$ 到自身的线性映射

$$(Rf)(t) = (t^2-1)f''(t) + tf'(t).$$

(i) 说明 R 相对于 $[\![\cdot|\cdot]\!]$ 具有自伴性质.

(ii) 说明 R 映多项式函数 T_n 为 n^2T_n.

另有一族被称为第二类 Chebyshev 多项式的 $U_0, U_1, \ldots \in \mathbb{R}[X]$, 满足

$$U_n(\cos x) = \frac{\sin(n+1)x}{\sin x},$$
$$(t^2-1)U_n''(t) + 3tU_n'(t) = n(n+2)U_n(t).$$

两类 Chebyshev 多项式有简单的关系 $T_n = U_n - XU_{n-1}$. 这些性质可以无穷尽地写下去, 不再详述. 它们和 Legendre 多项式所满足的常微分方程是 Sturm–Liouville 方程的特例.

15. 引入变元 q, 证明 P_n 和 T_n 满足

$$\sum_{n=0}^{\infty} P_n(t)q^n = \frac{1}{\sqrt{1-2qt+q^2}},$$
$$\sum_{n=0}^{\infty} T_n(t)q^n = \frac{1-qt}{1-2qt+q^2}.$$

按组合学的术语, 等号右边分别称为 P_n 和 T_n 的生成函数. 左边的无穷级数连同变元 q, t 只起形式作用, 收敛性并非重点; 倘若读者偏好收敛级数, 也可以要求 $t \in [-1,1]$ 而 $|q|$ 充分小.

16. 证明若 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 是正交矩阵, 则 $X^n \operatorname{Char}_{\boldsymbol{A}}(1/X) = (-1)^n \det(\boldsymbol{A}) \operatorname{Char}_{\boldsymbol{A}}(X)$.

17. 试说明不可逆的 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 也有类似于推论 9.4.7 的分解 $\boldsymbol{A} = \boldsymbol{QR}$, 差别仅在于上三角矩阵 $\boldsymbol{R}$ 在此不要求可逆.

18. 设 $(V, (\cdot|\cdot))$ 为有限维内积空间, V_0 为 V 的子空间. 证明对 V_0 的镜射 (定义 9.3.14) 其行列式为 $(-1)^{\dim V - \dim V_0}$.

19. 设 $n \geq 2$. 所谓 Jacobi 矩阵是具有以下形式的 $n \times n$ 实矩阵:

$$\boldsymbol{A} = \begin{pmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & b_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{pmatrix}, \quad \forall i,\ b_i c_i > 0.$$

(i) 证明 Jacobi 矩阵 $\boldsymbol{A}$ 总能在 $\mathbb{R}$ 上对角化. 〈提示〉 适当地取 $\boldsymbol{P}$ 使得 $\boldsymbol{PAP}^{-1}$ 对称.

(ii) 证明 Jacobi 矩阵 $\boldsymbol{A}$ 有 n 个相异实特征值. 〈提示〉 具体描述每个特征子空间.

20. 考虑带标准内积结构的 $\mathbb{R}^n$, 对以下线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 求一组最小二乘解 $\boldsymbol{x}$.

(i) $\boldsymbol{A} = \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$.

(ii) $\boldsymbol{A}=\begin{pmatrix}1&1&0\\1&0&-1\\0&1&1\\-1&1&-1\end{pmatrix}$, $\boldsymbol{b}=\begin{pmatrix}2\\5\\6\\6\end{pmatrix}$.

21. 将上一题的矩阵 $\boldsymbol{A}$ 视同线性映射, 计算它们的 Moore–Penrose 广义逆 (定义 9.9.2).

22. 求以下实矩阵 $\boldsymbol{A}$ 的奇异值分解 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{\Sigma}\,{}^{\mathrm{t}}\boldsymbol{P}$:

$$\boldsymbol{A}=\begin{pmatrix}-4&-6\\3&-8\end{pmatrix}.$$

提示〉 答案是

$$\boldsymbol{\Sigma}=\begin{pmatrix}10&0\\0&5\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \boldsymbol{Q}=\frac{1}{5}\begin{pmatrix}-3&-4\\-4&3\end{pmatrix}.$$

23. 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$ 为对应到正定二次型的对称矩阵. 说明存在对角元全为正的上三角矩阵 $\boldsymbol{R}$ 使得 $\boldsymbol{A}={}^{\mathrm{t}}\boldsymbol{R}\boldsymbol{R}$, 这称为 $\boldsymbol{A}$ 的 **Cholesky 分解**.

提示〉 将 $\boldsymbol{A}$ 写成 ${}^{\mathrm{t}}\boldsymbol{C}\boldsymbol{C}$, 对 $\boldsymbol{C}$ 作推论 9.4.7 的 $\boldsymbol{Q}\boldsymbol{R}$ 分解.

24. (可逆矩阵的 $\boldsymbol{L}\boldsymbol{U}$ 分解) 设 F 为任意域而 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$. 证明 $\boldsymbol{A}$ 的所有顺序主子式 (定义 9.7.1) 皆非零当且仅当存在 $\boldsymbol{L},\boldsymbol{U}\in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{L}$ 下三角可逆, $\boldsymbol{U}$ 上三角可逆, 而且 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$.

提示〉 对于 "仅当" 方向, 基于 $a_{11}\neq 0$, 以初等行和列变换将第一行和第一列的其他矩阵元全消为零, 说明当 $n>1$ 时对剩下的 $(n-1)\times(n-1)$ 矩阵可递归地操作.

25. 设 $T:V\to W$ 为有限维向量空间之间的线性映射. 证明当 $w\in W$ 给定, 存在 $v\in V$ 使得 $Tv=w$ 当且仅当存在广义逆 $S:W\to V$ (定义 9.9.1) 使得 $w=TSw$, 而且当上述条件成立时, 对每个 $v\in T^{-1}(w)$ 都存在广义逆 S 使得 $v=Sw$.

26. 设 $T:V\to W$ 为有限维内积空间之间的线性映射. 记 $C(t):=T^*T+t\cdot \mathrm{id}_V$. 说明 T 的 Moore–Penrose 广义逆 $S:W\to V$ 由下式确定

$$S=\lim_{\substack{t\to 0\\ \det C(t)\neq 0}} C(t)^{-1}T^*.$$

极限的严格意义可以放在矩阵空间中理解. 提示〉 运用奇异值分解.

27. 设 V 为有限维实内积空间, $T\in \mathrm{End}(V)$ 自伴. 对于熟悉数学分析的读者. 试在不用正交对角化定理 9.5.2 的情形下, 直接说明映射 $v\mapsto (Tv|v)$ 在 $\{v\in V:\|v\|=1\}$ 上取到极大值 $\lambda_{\max}$ 和极小值 $\lambda_{\min}$, 而且它们都是 T 的特征值.

提示〉 使用 Lagrange 乘子法. 这反过来为正交对角化定理提供了一个不用复数的证明.

28. 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$, $\boldsymbol{A}>0$. 以注记 9.11.5 的 Collatz–Wielandt 公式证明谱半径满足 $\rho(\boldsymbol{A})\geq \max\{a_{11},\dots,a_{nn}\}$.

29. 回顾定理 9.11.8 之前的讨论. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 满足 $\boldsymbol{A} \geq 0$. 若 $G(\boldsymbol{A})$ 连通则称 A 为不可约的; 若存在 $m \in \mathbb{Z}_{\geq 1}$ 使得 $\boldsymbol{A}^m > 0$ 则称 $\boldsymbol{A}$ 为本原的. 本原条件强于不可约.

(i) 说明若 $\boldsymbol{A}$ 不可约, 则 $\boldsymbol{A}$ 没有全为零的行或列. 由此说明若 $m \in \mathbb{Z}_{\geq 1}$ 使得 $\boldsymbol{A}^m > 0$, 则 $k \geq m \implies \boldsymbol{A}^k > 0$.

(ii) 称有向图 $G(\boldsymbol{A})$ 中一列形如 $i \xrightarrow{e_1} \cdots \xrightarrow{e_m} i$ 的边为长度 m 的环路. 设 $\boldsymbol{A}$ 没有全为零的行或列, 定义 $\boldsymbol{A}$ 的周期为 $G(\boldsymbol{A})$ 中所有环路长度的最大公因数. 证明 $\boldsymbol{A}$ 是本原的当且仅当 $\boldsymbol{A}$ 不可约而且周期为 1.

提示 若正整数 $x_1, \ldots, x_n$ 互素, 则初等数论说明充分大的正整数总能写成 $\sum_{i=1}^n a_i x_i$ 之形, 其中 $x_i \in \mathbb{Z}_{\geq 0}$.

(iii) 设 $\boldsymbol{A}$ 不可约, 周期为 p. 任取 $G(\boldsymbol{A})$ 的顶点 v, 对每个 $i = 1, \ldots, p$ 定义

$$C(i) := \left\{ \text{顶点 } u : \text{存在 } v \xrightarrow{e_1} \cdots \xrightarrow{e_k} u, \text{ 使得 } k \equiv i \pmod{p} \right\}.$$

说明这将 $G(\boldsymbol{A})$ 的顶点集划分为 p 个子集, 而且若有边 $C(i) \xrightarrow{e} C(j)$, 则 $j \equiv i+1 \pmod{p}$.

提示 相当于说明对任意顶点 u 和 v, 从 v 到 u 的步长 $\bmod p$ 是唯一确定的. 观察到若从 v 到 u 有步长为 k 的一列边, 则从 u 到 v 的任一列边的步长必 $\equiv -k \bmod p$.

(iv) 基于 (iii), 说明若 $\boldsymbol{A}$ 不可约且周期为 p, 则存在置换矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} & \boldsymbol{A}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_{p-1} \\ \boldsymbol{A}_p & & & \end{pmatrix}$$

上式为分块表法, 其中 $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p \geq 0$, 但它们未必是方阵.

(v) 承上, 验证 $(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^p$ 是分块对角矩阵, 每个对角分块都有相同的特征值.

(vi) 对不可约的 $\boldsymbol{A}$ 完成定理 9.11.8 的证明. 进一步说明下述性质.

(a) 若 p 是 $\boldsymbol{A}$ 的周期, $\zeta \in \mathbb{C}$ 是 p 次单位根 (亦即 $\zeta^p = 1$), 则 λ 是 $\boldsymbol{A}$ 的特征值当且仅当 $\zeta\lambda$ 亦然.

(b) 若 $\boldsymbol{A}$ 的特征值 $\mu \in \mathbb{C}$ 满足 $|\mu| = \rho(\boldsymbol{A})$, 则存在 p 次单位根 ζ 使得 $\mu = \zeta\rho(\boldsymbol{A})$.

提示 不妨设 $\boldsymbol{A} = \begin{pmatrix} & \boldsymbol{A}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_{p-1} \\ \boldsymbol{A}_p & & & \end{pmatrix}$. 说明方阵 $\boldsymbol{A}_1 \cdots \boldsymbol{A}_p$ 是 $\boldsymbol{A}^p$ 的一个对角分块, 并且是本原的 (应用本题 (ii) 的结果), 谱半径为 $\rho(\boldsymbol{A})^p$. 取列向量 $\boldsymbol{w} > 0$

使得 $\boldsymbol{A}_1\cdots\boldsymbol{A}_p\boldsymbol{w}=\rho(\boldsymbol{A})^p\boldsymbol{w}$, 然后考虑

$$\boldsymbol{v}:=\begin{pmatrix}\rho(\boldsymbol{A})^{p-1}\boldsymbol{w}\\ \boldsymbol{A}_2\cdots\boldsymbol{A}_p\boldsymbol{w}\\ \rho(\boldsymbol{A})\boldsymbol{A}_3\cdots\boldsymbol{A}_p\boldsymbol{w}\\ \vdots\\ \rho(\boldsymbol{A})^{p-2}\boldsymbol{A}_p\boldsymbol{w}\end{pmatrix}.$$

30. 试以上题的结果处理一般的 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})\smallsetminus\{\boldsymbol{0}_{n\times n}\}$, $\boldsymbol{A}\geq 0$.

提示〉 在 $\{1,\dots,n\}$ 上引入等价关系: $i\sim j$ 当且仅当 $i=j$ 或者存在边 $i\to\cdots\to j$ 和 $j\to\cdots\to i$. 命 $\mathcal{Q}$ 为相应的商集, 边的走向赋予 $\mathcal{Q}$ 偏序. 顺此思路来说明存在置换矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 成为分块上三角矩阵, 每个对角分块或者是零矩阵, 或者不可约.

第十章 复内积结构

本章是第九章的延续. 考虑复向量空间 V, W, X. 在 §10.1, 半双线性映射 $B : V \times W \to X$ 被定义为满足恒等式

$$\begin{aligned} B(v_1 + v_2, w) &= B(v_1, w) + B(v_2, w), \\ B(tv, w) &= \bar{t}B(v, w), \\ B(v, w_1 + w_2) &= B(v, w_1) + B(v, w_2), \\ B(v, tw) &= tB(v, w) \end{aligned}$$

的映射, 换言之要求 B 对第二个变元是线性的, 对第一个变元则是 "半线性" 的. 许多教材规定第二个变元为半线性变元, 这纯粹是惯例的选择, 不影响理论实质.

对应到 $X = \mathbb{C}$ 的半双线性映射称为半双线性形式. 此前关于双线性形式的许多结论都能移植到半双线性情形; 譬如当 $V = \mathbb{C}^m$ 而 $W = \mathbb{C}^n$ 时, 半双线性形式 B 一一对应于复矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(\mathbb{C})$, 刻画为

$$B(\boldsymbol{v}, \boldsymbol{w}) = {}^{\dagger}\boldsymbol{v}\, \boldsymbol{A}\, \boldsymbol{w}, \quad {}^{\dagger}\boldsymbol{v} := {}^{\mathrm{t}}\overline{\boldsymbol{v}},$$

此处 ${}^{\dagger}\boldsymbol{v}$ 的含义见约定 9.5.3. 此外, 非退化形式与伴随映射的概念在有限维情形依然适用.

从 §10.2 起, 考虑的半双线性形式形如 $B : V \times V \to \mathbb{C}$. 若 $B(v_1, v_2) = \overline{B(v_2, v_1)}$ (或 $B(v_1, v_2) = -\overline{B(v_2, v_1)}$) 恒成立, 则称 B 为 Hermite (或反 Hermite) 形式. 有限维复向量空间上的 Hermite 形式有简单的分类 (定义–定理 10.2.8), 形似实二次型的惯性定理.

有了这些准备, 我们在 §10.3 简洁地定义复内积为满足正定性的 Hermite 形式 $(\cdot|\cdot) : V \times V \to \mathbb{C}$. 复向量空间与复内积的搭配称为复内积空间, 又称为酉空间; 标准范例是 $\mathbb{C}^n$ 上的 $(\boldsymbol{x}|\boldsymbol{y}) = \sum_{i=1}^{n} \overline{x_i} y_i = {}^{\dagger}\boldsymbol{x}\, \boldsymbol{y}$. 尽管复内积不如实内积直观, 这种结构在数学分析与物理学等应用中自然地出现.

从实内积衍生的许多概念都能推及复内积, 包括向量的长度 $\|v\| := \sqrt{(v|v)}$, 正交性, 勾股定理 (命题 10.3.5), Cauchy–Bunyakovsky–Schwarz 不等式 (定理 10.3.6), 三角不等式 (推论 10.3.7), Gram–Schmidt 正交化, 对有限维子空间的正交投影与分解, 等等. 特别地, 任何有限维复内积空间都同构于标准的 $(\mathbb{C}^n, (\cdot|\cdot))$.

选定有限维复内积空间 $(V,(\cdot|\cdot))$. 满足 $T^*T=TT^*$ 的 $T\in\mathrm{End}(V)$ 称为正规的. 在 §10.4, 我们将证明酉对角化定理 10.4.1: 存在 V 的单位正交基 $v_1,\dots,v_n$ 将 T 对角化的充要条件是 T 正规. 此结论又称为正规算子的谱分解, 是本章的核心之一.

在 §10.5, 我们将 §§9.6–9.10 的许多结论推广到复的情形.

以复数域上的酉对角化定理为工具, 实数域上的正交变换构成 §10.6 以下的主题. 我们先从二维情形起步, 确定 2 维旋转矩阵 $\boldsymbol{R}(\theta)$ 的样貌. 之后的定理 10.6.6 完全描述了高维数情形的正交变换: 就几何来看, 它们在合适的单位正交基下呈现为 $\pm\mathrm{id}$ 和多轴旋转的组合.

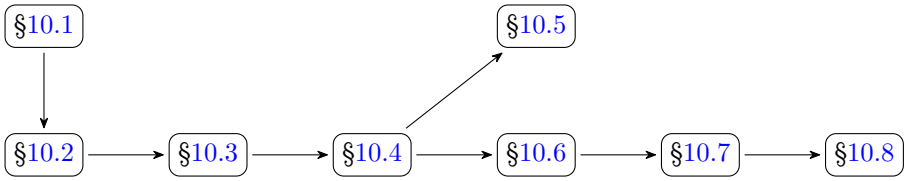
三维情形的正交变换是 §10.7 的主题, 它们的几何图像格外具体, 而且是飞行器控制和游戏引擎等应用领域的基础. 这些变换或者是绕一个定轴的旋转, 或者是旋转与镜射的合成, 两者按行列式来区分. 该节将介绍 Euler 角的概念, 借以将三维空间中的旋转参数化; 相应的分解有内在 (定理 10.7.4) 和外在 (定理 10.7.6) 两种形式, 各有用途.

在实践和理论两方面, Euler 角都有其缺陷; 借助这一契机, §10.8 将介绍 Hamilton 发现的四元数. 就代数而言, 四元数是带有四个实坐标的 “数”; 它们是复数的拓展, 具有四则运算, 但乘法不满足交换律, 从而给出非交换除环的自然例子. 就几何而言, 四元数为三维旋转提供了一种自然的参数化. 四元数环 $\mathbb{H}$ 能嵌入为 $\mathrm{M}_{2\times2}(\mathbb{C})$ 的子环 (命题 10.8.7), 因此 $\mathbb{H}$ 的定义从代数视角来看并不神秘.

阅读提示

除了 §§10.6–10.8 和相应的习题, 本章的向量空间都是复的. 在 §§10.4–10.8 仅考虑有限维向量空间. 由于前半部许多内容是实内积情形的简单类比, 叙述步调将适度加快.

阅读顺序



10.1 半双线性形式

在定义半双线性形式之前, 有必要先引入半线性映射的概念. 按惯例, 复数 z 的复共轭记为 $\bar{z}$.

定义 10.1.1 复向量空间之间的映射 $T : V \to W$ 若满足

$$T(v_1 + v_2) = T(v_1) + T(v_2),$$
$$T(tv) = \bar{t}T(v),$$

其中 $v, v_1, v_2 \in V$ 而 $t \in \mathbb{C}$, 则称为**半线性映射**.

半线性映射自动是 $\mathbb{R}$-线性的. 线性与半线性映射有许多相似的性质, 证明通常可以照搬. 为了在两者间系统性地建立联系, 引入向量空间的复共轭是一劳永逸的办法.

定义 10.1.2 设 V 是 $\mathbb{C}$-向量空间, 它的复共轭 $\overline{V}$ 是按以下方式确定的 $\mathbb{C}$-向量空间 $(V, +, \odot)$.

- ⋆ 集合 V 和加法运算 $+$ 与原空间相同.
- ⋆ 纯量乘法 $\odot : \mathbb{C} \times V \to V$ 定义为 $t \odot v := \bar{t}v$.

显然, $\overline{\overline{V}} = V$.

从定理当下看出: 从 V 到 W 的半线性映射无非是线性映射 $\overline{V} \to W$, 也可以等价地说是线性映射 $V \to \overline{W}$.

练习 10.1.3 验证关于复共轭的以下性质, 其中的 V_i 和 W 都是 $\mathbb{C}$-向量空间.

(i) 映射 $z \mapsto \bar{z}$ 给出 $\mathbb{C}$-向量空间的同构 $\overline{\mathbb{C}} \xrightarrow{\sim} \mathbb{C}$.

(ii) 验证 $\mathbb{C}$-向量空间的等式 $\overline{V_1} \oplus \overline{V_2} = \overline{V_1 \oplus V_2}$ 和 $\mathrm{Hom}(V_1, V_2) = \overline{\mathrm{Hom}(\overline{V}_1, \overline{V}_2)}$.

(iii) 直接验证 $\mathbb{C}$-向量空间的同构 $\overline{W}^\vee \xrightarrow{\sim} \overline{W^\vee}$, 它映 $\lambda \in \overline{W}^\vee$ 为 $\overline{\lambda} : w \mapsto \overline{\lambda(w)}$.

以下是定义 8.1.1 在 $\mathbb{C}$ 上的简单变奏, 差别仅在于一个复共轭.

定义 10.1.4 (半双线性映射和半双线性形式) 设 V, W, X 为 $\mathbb{C}$-向量空间, 所谓从 $V \times W$ 到 X 的半双线性映射, 是指满足以下条件的映射

$$B : V \times W \to X$$

- ⋆ 它对第一个变元是半线性的:

$$B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w),$$
$$B(tv, w) = \bar{t}B(v, w).$$

⋆ 它对第二个变元是线性的:

$$B(v,w_1+w_2)=B(v,w_1)+B(v,w_2),$$
$$B(v,tw)=tB(v,w).$$

全体半双线性映射 $V\times W\to X$ 对加法

$$(B_1+B_2)(v,w)=B_1(v,w)+B_2(v,w)$$

和纯量乘法

$$(tB)(v,w)=tB(v,w)$$

构成 $\mathbb{C}$-向量空间, 记为 $\mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V,W;X)$.

对于 $X=\mathbb{C}$ 的特例, 半双线性映射 $B:V\times W\to\mathbb{C}$ 也称为 $V\times W$ 上的半双线性形式, 它们构成的 $\mathbb{C}$-向量空间记为 $\mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V,W)$.

比较定义 10.1.4 和定义 10.1.2, 立得 $\mathbb{C}$-向量空间的等式

$$\mathrm{Bil}(\overline{V},W;X)=\mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V,W;X). \tag{10.1.1}$$

有鉴于此, 关于双线性映射的大部分观念与性质都可以简单地移植到半双线性映射上.

本书今后仅讨论半双线性形式, 而非更广义的半双线性映射.

基于 (10.1.1) 的观察, 指定半双线性形式 $B:V\times W\to\mathbb{C}$ 等价于指定 $\varphi\in\mathrm{Hom}(W,\overline{V}^\vee)$, 也等价于指定 $\psi\in\mathrm{Hom}(\overline{V},W^\vee)$. 下述定义是 (10.1.1) 的另一则应用.

定义 10.1.5 对于有限维 $\mathbb{C}$-向量空间 V 和 W, 以及半双线性形式 $B:V\times W\to\mathbb{C}$,

⋆ 定义 B 的**左根**为 $\{v\in V:B(v,\cdot)=0\}$.

⋆ 定义 B 的**右根**为 $\{w\in W:B(\cdot,w)=0\}$.

⋆ 设 V 和 W 都是有限维的. 若 B 的左根和右根都是 $\{0\}$, 则称 B **非退化**.

一如双线性情形, 左根等于 $\ker(\psi)$ 而右根等于 $\ker(\varphi)$, 它们分别是 V 和 W 的子空间 ($\overline{V}$ 和 V 的子空间是一回事). 类似地, 仅在 $\dim V=\dim W$ 皆有限时方可能有非退化的半双线性形式 $V\times W\to\mathbb{C}$.

下述结果无非是命题 8.2.4 对 $\overline{V}$ 和 W 的应用.

命题 10.1.6 设 V 和 W 是有限维 $\mathbb{C}$-向量空间, $\dim V=\dim W$, 则对于任意 $B\in\mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V,W)$, 以下性质相互等价:

(i) B 非退化,

(ii) B 的左根为零,

(iii) B 的右根为零.

当以上任一条件成立时, B 对应的 $\varphi: W \to \overline{V}^\vee$ 和 $\psi: \overline{V} \to W^\vee$ 都是同构.

一如双线性形式的情形, 半双线性形式也可以由矩阵表达. 约定 9.5.3 关于矩阵的符号 ${}^\dagger\boldsymbol{A} := {}^{\mathrm{t}}\overline{\boldsymbol{A}}$ 在此是称手的. 第一步是以下观察: 在一般域 F 上的双线性形式的研究中, 我们将 F^n 的元素等同于列向量, $(F^n)^\vee$ 的元素等同于行向量, 其间以 $\boldsymbol{v} \mapsto {}^{\mathrm{t}}\boldsymbol{v}$ 相互过渡. 对于半双线性形式, 我们需要在 $\mathbb{C}^n$ 和 $(\overline{\mathbb{C}^n})^\vee$ 之间过渡, 其方式自然地和先前差了一个复共轭: 必须改取 $\boldsymbol{v} \mapsto {}^\dagger\boldsymbol{v}$.

命题 10.1.7 (以矩阵表达半双线性形式) 设 $m, n \in \mathbb{Z}_{\geq 1}$. 将 $\mathbb{C}^m$ 和 $\mathbb{C}^n$ 的元素视同列向量, 则有向量空间的同构

$$\begin{aligned} \mathrm{M}_{m\times n}(\mathbb{C}) &\xrightarrow{\sim} \mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(\mathbb{C}^m, \mathbb{C}^n) \\ \boldsymbol{A} &\longmapsto [B(\boldsymbol{v}, \boldsymbol{w}) := {}^\dagger\boldsymbol{v}\,\boldsymbol{A}\,\boldsymbol{w}]; \end{aligned}$$

进一步,

$$B\left(\sum_{i=1}^m x_i\boldsymbol{e}_i, \sum_{j=1}^n y_j\boldsymbol{e}_j\right) = \sum_{i=1}^m\sum_{j=1}^n a_{ij}\overline{x_i}y_j,$$

而 B 非退化的充要条件是对应的矩阵 $\boldsymbol{A}$ 可逆.

证明 和命题 8.1.6 的双线性版本基本相同. 以该处的抽象方法为例, B 对应到 $\varphi \in \mathrm{Hom}\left(\mathbb{C}^n, (\overline{\mathbb{C}^m})^\vee\right)$, 然而我们有 $\mathbb{C}$-向量空间的同构

$$\mathbb{C}^m \xrightarrow{\sim} (\overline{\mathbb{C}^m})^\vee = \left\{\text{半线性映射 } \mathbb{C}^m \to \mathbb{C}\right\},$$

方法是映列向量 $\boldsymbol{v}$ 为半线性映射 $\boldsymbol{v}_1 \mapsto {}^{\mathrm{t}}\boldsymbol{v}\,\overline{\boldsymbol{v}_1}$. 按此让 φ 对应到矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(\mathbb{C})$, 从而 $\varphi(\boldsymbol{w})$ 是半线性映射 $\boldsymbol{v} \mapsto {}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{w})\,\overline{\boldsymbol{v}}$, 故

$$B(\boldsymbol{v}, \boldsymbol{w}) = \langle\varphi(\boldsymbol{w}), \boldsymbol{v}\rangle = {}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{w})\,\overline{\boldsymbol{v}} = {}^{\mathrm{t}}\boldsymbol{w}\ {}^{\mathrm{t}}\boldsymbol{A}\,\overline{\boldsymbol{v}}.$$

取转置可见末项也等于 ${}^\dagger\boldsymbol{v}\,\boldsymbol{A}\,\boldsymbol{w}$.

展开矩阵乘积便得到 B 的具体公式. 由于先前已将 φ 等同于 $\boldsymbol{A}$ 对应的线性映射 $\mathbb{C}^n \to \mathbb{C}^m$, 而 B 的右根无非 $\ker\varphi$, 故非退化等价于 $m = n$ 而 A 可逆. □

许多文献对半双线性形式的定义是它对第二个变元是半线性的; 这相当于在矩阵表达式中改取 $B(\boldsymbol{v}, \boldsymbol{w}) = {}^{\mathrm{t}}\boldsymbol{v}\boldsymbol{A}\overline{\boldsymbol{w}}$.

循着双线性形式的线索, 接着来探讨 $V = W$ 的场景.

定义 10.1.8 设 V 是 $\mathbb{C}$-向量空间, $\epsilon \in \{\pm 1\}$. 若半双线性形式 $B: V \times V \to \mathbb{C}$ 满足

$$B(v, w) = \epsilon\overline{B(w, v)},$$

则称 B 是 ϵ-Hermite 形式.

我们将 $(+1)$-Hermite 形式简称为 V 上的 **Hermite 形式**, 将 (-1)-Hermite 形式简称为 V 上的**反 Hermite 形式**.

对于 Hermite 或反 Hermite 形式 B, 定义直接导致 B 的左根和右根是一回事, 称之为 B 的**根基**.

我们主要关切有限维的 V. 对于 $V=\mathbb{C}^n$ 的情形, 命题 10.1.7 对 B 的矩阵表达式表明 ϵ-半双线性形式对应到以下定义的 ϵ-Hermite 矩阵.

定义 10.1.9 设 $\epsilon\in\{1,-1\}$. 满足 ${}^\dagger\boldsymbol{A}=\epsilon\boldsymbol{A}$ 的 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 称为 ϵ-Hermite 矩阵. 我们将 $(+1)$-Hermite 矩阵简称为 **Hermite 矩阵**, 将 (-1)-Hermite 矩阵简称为**反 Hermite 矩阵**.

半双线性形式的 ϵ-Hermite 性质可以化到基上来检验: $B:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$ 是 ϵ-Hermite 的当且仅当 $B(\boldsymbol{e}_i,\boldsymbol{e}_j)=\epsilon\overline{B(\boldsymbol{e}_j,\boldsymbol{e}_i)}$ 对所有 i,j 成立, 而这又等价于 $a_{ij}=\epsilon\overline{a_{ji}}$. 这点从命题 10.1.7 的公式看也是明白的.

定义–命题 10.1.10 给定有限维 $\mathbb{C}$-向量空间 V_i, V_i' 和 $B_i\in \mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V_i,V_i')$, 其中 $i=1,2$, 并且假设 B_1 非退化.

(i) 存在唯一的半线性映射

$$\begin{aligned}\mathrm{Hom}(V_1,V_2)&\longrightarrow \mathrm{Hom}(V_2',V_1')\\ T&\longmapsto T^*,\end{aligned}$$

其中的 T^* 称为 T 相对于 B_1 和 B_2 的**右伴随**, 由下式刻画:

$$B_2(Tv_1,v_2')=B_1(v_1,T^*v_2'),\quad v_1\in V_1,\ v_2'\in V_2'.$$

(ii) 存在唯一的半线性映射

$$\begin{aligned}\mathrm{Hom}(V_1',V_2')&\longrightarrow \mathrm{Hom}(V_2,V_1)\\ T&\longmapsto {}^*T,\end{aligned}$$

其中的 ${}^*T$ 称为 T 相对于 B_1 和 B_2 的**左伴随**, 由下式刻画:

$$B_2(v_2,Tv_1')=B_1({}^*Tv_2,v_1'),\quad v_1'\in V_1',\quad v_2\in V_2.$$

证明 若将 $\mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V_i,V_i')$ 视同 $\mathrm{Bil}(\overline{V_i},V_i';\mathbb{C})$, 则这几乎就是定义–命题 8.2.8 的内容. 差别在于定义–命题 8.2.8 对此给出线性映射

$$\mathrm{Hom}(\overline{V_1},\overline{V_2})\to\mathrm{Hom}(V_2',V_1'),\quad T\mapsto T^*,$$

但 $\mathrm{Hom}(\overline{V_1},\overline{V_2})=\overline{\mathrm{Hom}(V_1,V_2)}$, 故上述产物也相当于从 $\mathrm{Hom}(V_1,V_2)$ 到 $\mathrm{Hom}(V_2',V_1')$ 的半线性映射. □

一如 (8.2.1), 当 B_1 和 B_2 皆非退化时, 关于左伴随和右伴随的刻画即刻给出

$$({}^*T)^* = T = {}^*(T^*).$$

给定 $B_i \in \mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V_i, V_i')$, 其中 $i = 1, 2, 3$, 和线性映射 $V_1 \xrightarrow{T} V_2 \xrightarrow{S} V_3$, 并假设 B_1 和 B_2 非退化, 则

$$(ST)^* = T^*\, S^*;$$

左伴随的情形当然也类似.

例 10.1.11 假设 $V_i = V_i' = \mathbb{C}^{n_i}$, 等同于列向量空间 $(i = 1, 2)$. 考虑 $\boldsymbol{A}_i \in \mathrm{M}_{n_i \times n_i}(\mathbb{C})$ 确定的半双线性形式 B_i; 假设 $\boldsymbol{A}_1$ 可逆, 亦即 B_1 非退化. 以下给定矩阵 $\boldsymbol{T} \in \mathrm{M}_{n_2 \times n_1}(\mathbb{C})$, 视同线性映射 $\mathbb{C}^{n_1} \to \mathbb{C}^{n_2}$. 如何确定其伴随? 一如例 8.2.14, 答案是

$$\boldsymbol{T}^* \overset{\text{等同于}}{=\!=} \boldsymbol{A}_1^{-1}\, {}^\dagger\boldsymbol{T}\, \boldsymbol{A}_2 \ \in \mathrm{M}_{n_1 \times n_2}(\mathbb{C}).$$

这点既可以化约到例 8.2.14 的版本, 也可以直接代入矩阵表达式来检验等式

$${}^\dagger(\boldsymbol{T}\boldsymbol{v}_1)\, \boldsymbol{A}_2\, \boldsymbol{v}_2 = {}^\dagger\boldsymbol{v}_1\, \boldsymbol{A}_1\, \left(\boldsymbol{A}_1^{-1}\, {}^\dagger\boldsymbol{T}\, \boldsymbol{A}_2\right) \boldsymbol{v}_2$$

确实对所有列向量 $\boldsymbol{v}_i \in \mathbb{C}^{n_i}$ 成立 $(i = 1, 2)$, 而这刻画了 $\boldsymbol{T}^*$.

此外我们也有对应于命题 8.2.12 的以下结果. 所涉论证和双线性形式情形毫无差别, 不必重复.

命题 10.1.12 *取定非退化的 $B_i \in \mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V_i, V_i)$, 其中 $i = 1, 2$. 若存在 $\epsilon \in \{\pm 1\}$ 使得 B_1 和 B_2 都是 ϵ-Hermite 的, 则 ${}^*T = T^*$; 换言之, 此时 $B_2(v_2, Tv_1) = B_1(T^*v_2, v_1)$ 对所有 v_1, v_2 皆成立.*

作为推论, 此时有 $(T^*)^* = T = {}^*({}^*T)$.

以下概念对应到实数情形的定义 8.2.15.

定义 10.1.13 *设 $B \in \mathrm{Sesq}_{\mathbb{C}|\mathbb{R}}(V, V)$ 非退化. 考虑线性映射 $T \in \mathrm{End}(V)$. 以下概念都是相对于 B 而言的.*

- *若 $T^* = T$, 则称 T **自伴**.*
- *若 $T^* = -T$, 则称 T **反自伴**.*

严格来说, 此处应该对左伴随 ${}^*T = \pm T$ 给出相应的定义. 由于之后考虑的都是 B 为 ϵ-Hermite 形式的情形, 此处不必过多着墨.

对于 $V = \mathbb{C}^n$ 而 B 对应于可逆矩阵 $\boldsymbol{A} \in \mathrm{M}_{n \times n}(\mathbb{C})$ 的情形, 例 10.1.11 即刻给出

$$\begin{aligned} \boldsymbol{T} \text{ 自伴} &\iff \boldsymbol{A}^{-1}\, {}^\dagger\boldsymbol{T}\, \boldsymbol{A} = \boldsymbol{T}, \\ \boldsymbol{T} \text{ 反自伴} &\iff \boldsymbol{A}^{-1}\, {}^\dagger\boldsymbol{T}\, \boldsymbol{A} = -\boldsymbol{T}, \end{aligned}$$

等价右侧的等式都是作为矩阵来理解的. 和双线性的情况不同, 自伴和反自伴线性映射在此容易相互过渡: 设 c 是非零的纯虚数, 则 T 自伴当且仅当 cT 反自伴.

定义 10.1.14 (正规线性映射) 给定 V 和非退化 ϵ-Hermite 形式 $B: V\times V\to\mathbb{C}$, 满足 $T^*T=TT^*$ 的线性映射 $T: V\to V$ 称为正规的.

正规线性映射也常被称为正规变换或正规算子. 自伴和反自伴线性映射当然都是正规的, 一般的 T 则唯一地分解为这两类映射的和.

命题 10.1.15 给定 V 和非退化 ϵ-Hermite 形式 $B: V\times V\to\mathbb{C}$. 对任意 $T\in\mathrm{End}(V)$, 存在唯一的自伴线性映射 T' 和反自伴线性映射 T'' 使得 $T=T'+T''$. 若进一步要求 T 正规, 则 $T'T''=T''T'$.

证明 首先说明分解的唯一性, 若有两个分解 $T_1'+T_1''=T_2'+T_2''$, 则 $T_1'-T_2'=T_2''-T_1''$ 既自伴又反自伴, 唯一可能是 $T_1'=T_2'$ 而 $T_1''=T_2''$.

至于存在性, 取 $T':=\frac{1}{2}(T+T^*)$ 和 $T'':=\frac{1}{2}(T-T^*)$ 即是; 此处关键是运用性质 $(T^*)^*=T$.

若 $T^*T=TT^*$, 则以上定义的 T' 和 T'' 也满足 $T'T''=T''T'$. 至此完成证明. □

在介绍复内积 (定义 10.3.1) 的概念之后, §10.4 将对复内积空间上包括自伴算子在内的所有正规算子证明谱定理. 自伴算子在数学物理和相关领域中频繁出现.

10.2 Hermite 形式的分类

我们在 §8.4 介绍过 n 元二次型的概念. 当 $\epsilon\in\{1,-1\}$ 取定, 它的半双线性版本是形如

$$f(x_1,\dots,x_n)=\sum_{1\leq i,j\leq n}a_{ij}\overline{x_i}x_j$$

的映射 $f:\mathbb{C}^n\to\mathbb{C}$, 其中要求 $a_{ij}=\epsilon\overline{a_{ij}}$ 对所有 i,j 成立.

练习 10.2.1 仅用涉及 $i\leq j$ 的系数 a_{ij} 来表达 f, 然后证明 $\epsilon=1$ 时 f 取值总为实数, 当 $\epsilon=-1$ 时 f 取值总为纯虚数.

具有上述性质的映射 f 称为 n 元 ϵ-Hermite 型. 由于涉及复共轭, 它们不再是 n 元多项式, 然而先前的技术依然管用: 仍有双射

$$\begin{array}{ccccc}
\{n\text{ 元 }\epsilon\text{-Hermite 型}\} & \xleftrightarrow{1:1} & \left\{\begin{array}{l}\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{C})\\ \epsilon\text{-Hermite 矩阵}\end{array}\right\} & \xleftrightarrow{1:1} & \left\{\begin{array}{l}B:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}\\ \epsilon\text{-Hermite 形式}\end{array}\right\}\\
\cup & & \cup & & \cup\\
f=\sum_{1\leq i,j\leq n}a_{ij}\overline{x_i}x_j & & \boldsymbol{A}=(a_{ij})_{i,j} & & B(\boldsymbol{v}_1,\boldsymbol{v}_2)={}^{\dagger}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_2
\end{array}\tag{10.2.1}$$

具体地说, 映射 f 可在给定的列向量 $\boldsymbol{v}\in\mathbb{C}^n$ 上求值, 而这些对应即刻给出

$$f(\boldsymbol{v})={}^{\dagger}\boldsymbol{v}\boldsymbol{A}\boldsymbol{v}=B(\boldsymbol{v},\boldsymbol{v}).$$

涉及的论证和二次型是类似的. 举例来说, 按半双线性展开 $f(\boldsymbol{v}+\boldsymbol{w}) = B(\boldsymbol{v}+\boldsymbol{w}, \boldsymbol{v}+\boldsymbol{w})$ 可得

$$f(\boldsymbol{v}+\boldsymbol{w}) = \begin{cases} f(\boldsymbol{v}) + f(\boldsymbol{w}) + 2\operatorname{Re} B(\boldsymbol{v},\boldsymbol{w}), & \epsilon = +1, \\ f(\boldsymbol{v}) + f(\boldsymbol{w}) + 2\mathrm{i}\operatorname{Im} B(\boldsymbol{v},\boldsymbol{w}), & \epsilon = -1; \end{cases} \tag{10.2.2}$$

上式不只从 f 确定 $\operatorname{Re} B$ 或 $\operatorname{Im} B$, 还确定 B 本身, 这是因为 $\operatorname{Im} B(\boldsymbol{v},\boldsymbol{w}) = \operatorname{Re} B(\mathrm{i}\boldsymbol{v},\boldsymbol{w})$. 从而 $\boldsymbol{A}$ 也由 f 唯一确定.

遵循双线性形式的成例, 我们进一步考虑所有资料 (V,B) 的分类问题, 其中

- ⋆ V 是有限维 $\mathbb{C}$-向量空间,
- ⋆ B 是 V 上的 ϵ-Hermite 形式; $\epsilon \in \{-1,1\}$ 是选定的.

从 (V,B) 到 (V',B') 的**同构**意谓满足以下条件的线性同构 $\varphi : V \xrightarrow{\sim} V'$

$$B(v_1,v_2) = B'(\varphi(v_1),\varphi(v_2)), \quad v_1,v_2 \in V.$$

而 (V_1,B_1) 和 (V_2,B_2) 的**直和** $(V_1\oplus V_2, B_1\oplus B_2)$ 定为

$$(B_1\oplus B_2)((v_1,v_2),(v_1',v_2')) = B_1(v_1,v_1') + B_2(v_2,v_2').$$

注记 10.2.2 正负号 ϵ 在定义中总是固定, Hermite 形式与反 Hermite 形式不相混杂. 尽管如此, Hermite 和反 Hermite 形式的基本性质完全类似, 这是基于以下简单观察:

$$\begin{aligned} \boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C}) \text{ 是 Hermite 的 } &\iff \mathrm{i}\boldsymbol{A} \text{ 是反 Hermite 的,} \\ B : V\times V \to \mathbb{C} \text{ 是 Hermite 的 } &\iff \mathrm{i}B \text{ 是反 Hermite 的.} \end{aligned}$$

本节仅有最后的分类定理须区分 Hermite 与反 Hermite 情形, 在其余部分容许一个可变的 ϵ 并不增加任何难度.

由于有限维向量空间总有基, 不失一般性可以在资料 (V,B) 中具体取 $V=\mathbb{C}^n$, 然后取对应 B 的 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 与 $f:\mathbb{C}^n\to\mathbb{C}$. 若 B_i 对应到 ϵ-Hermite 矩阵 $\boldsymbol{A}_i$, 其中 $i=1,2$, 则取直和 $B := B_1\oplus B_2$ 便对应于取分块对角矩阵

$$\boldsymbol{A} := \left(\begin{array}{c|c} \boldsymbol{A}_1 & \\ \hline & \boldsymbol{A}_2 \end{array}\right).$$

另一方面, 同构定义中的 $\varphi:\mathbb{C}^n \xrightarrow{\sim} \mathbb{C}^n$ 转译为可逆矩阵 $\boldsymbol{C} = (c_{ij})_{i,j} \in \mathrm{M}_{n\times n}(\mathbb{C})$, 关于 B 和 B' 同构的条件相应地化为

- ▷ **矩阵观点** $\boldsymbol{A},\boldsymbol{A}' \in \mathrm{M}_{n\times n}(\mathbb{C})$ 为 ϵ-Hermite 矩阵, $\boldsymbol{A} = {}^{\dagger}\boldsymbol{C}\,\boldsymbol{A}'\,\boldsymbol{C}$;
- ▷ **映射观点** $f,f' : \mathbb{C}^n \to \mathbb{C}$ 为 n 元 ϵ-Hermite 型, $f = f'\circ(Y_1,\dots,Y_n)$, 其中 $(Y_1,\dots,Y_n)$ 是从 $\mathbb{C}^n$ 到 $\mathbb{C}^n$ 的映射, 其第 i 个坐标是 $Y_i := \sum_{j=1}^n c_{ij}X_j$.

所以同构可以按照三种方式理解:

⋆ 可逆线性变量代换 $(Y_1,\dots,Y_n)$ (对于 n 元 ϵ-Hermite 型 f),
⋆ 由 $\boldsymbol{A}\sim {}^{\dagger}\boldsymbol{C}\,\boldsymbol{A}\,\boldsymbol{C}$ 确定的等价关系, 其中 $\boldsymbol{C}$ 可逆 (对于 ϵ-Hermite 矩阵 $\boldsymbol{A}$),
⋆ 保持半双线性形式的线性同构 $\varphi: V\xrightarrow{\sim} V'$ (对于资料 (V,B)).

我们关心的是 ϵ-Hermite 形式在同构意义下的分类. 三种观点殊途同归, 各有长处, 例如 n 元 ϵ-Hermite 型容易用配方法来实施对角化.

命题 10.2.3 (ϵ-Hermite 型的对角化) 任何 n 元 ϵ-Hermite 型 f 都同构于形如

$$(x_1,\dots,x_n)\mapsto a_1|x_1|^2+\cdots+a_n|x_n|^2$$

的对角型, 其中 $a_1,\dots,a_n$ 在 $\epsilon=+1$ 时为实数, 在 $\epsilon=-1$ 时为纯虚数.

证明 与命题 8.5.1 全然类似. 以该处证明的第一步为例, 不妨设 $n\geq 2$ 而

$$f(x_1,\dots,x_n)=\sum_{1\leq i,j\leq n} a_{ij}\overline{x_i}x_j,$$

其中 $\boldsymbol{A}=(a_{ij})_{i,j}$ 是 ϵ-Hermite 矩阵, 并且设 $a_{11}\neq 0$. 基于 $a_{1j}=\epsilon\overline{a_{j1}}$ 不难检验

$$f(x_1,\dots,x_n)=a_{11}\left|x_1+\frac{1}{a_{11}}\sum_{j=2}^{n}a_{1j}x_j\right|^2+\text{不含 } x_1 \text{ 的部分}.$$

作变量代换 $y_1:=x_1+\frac{1}{a_{11}}\sum_{j=2}^{n}a_{1j}x_j$ 和 $y_2=x_2,\dots,y_n=x_n$, 遂可递归地实施配方. 对于 $a_{11}=\cdots=a_{nn}=0$ 的情形则需要一些调整来制造对角系数, 技巧依然如命题 8.5.1. □

定义 10.2.4 设 f 为 n 元 ϵ-Hermite 型. 定义其**根基**为对应的 ϵ-Hermite 形式的根基. 定义其**秩**为 n 减去根基的维数.

因此非退化等价于根基为零空间, 等价于秩 n.

引理 10.2.5 若 $f(x_1,\dots,x_n)=\sum_{i=1}^{r}a_i|x_i|^2$, 其中 $a_1,\dots,a_r\neq 0$, 则 f 的根基等于 $\mathbb{C}^n$ 的子空间 $\langle \boldsymbol{e}_{r+1},\dots,\boldsymbol{e}_n\rangle$, 而秩等于 r. 特别地, 一个 ϵ-Hermite 型非退化当且仅当它对角化之后的系数全部非零.

证明 对应的 ϵ-Hermite 矩阵是以 $a_1,\dots,a_r,0,\dots,0$ 为对角元的对角矩阵, 而根基是该矩阵作为线性映射的核, 即 $\langle \boldsymbol{e}_{r+1},\dots,\boldsymbol{e}_n\rangle$; 秩按定义是 $n-(n-r)=r$. □

这些陈述自然都能扩及更一般的资料 (V,B). 相对地, 以下介绍的概念只适用于 Hermite 形式.

定义 10.2.6 考虑 $\mathbb{C}$-向量空间 V 上的 Hermite 形式 B. 如果 $B(v,v)\geq 0$ (或 ≤ 0) 对所有 v 成立, 则称 (V,B) 为半正定 (或半负定) 的; 如果进一步要求 $B(v,v)=0\iff v=0$, 则称 (V,B) 为正定 (或负定) 的. 若以上皆非, 则称 B 不定.

现已看出任何 n 元 Hermite 型 f 都能通过同构化到 $\sum_{i=1}^{r} a_i|x_i|^2$ 的形式, 其中 r 是秩, $1 \leq i \leq r$ 而 $a_i \in \mathbb{R} \smallsetminus \{0\}$; 进一步以 $y_i := \sqrt{|a_i|}x_i$ 代换, 便能看出 f 同构于形如

$$|y_1|^2 + \cdots + |y_p|^2 - |y_{p+1}|^2 - \cdots - |y_r|^2$$

的 Hermite 型, 其中 $0 \leq p \leq r$. 姑且称样貌如上的 n 元 Hermite 型为规范形. 我们的目的是借助规范形以分类 Hermite 型. 首先是命题 8.6.3 的类比.

命题 10.2.7 *设 n 元 Hermite 型 f 表为规范形 $|y_1|^2 + \cdots + |y_p|^2 - |y_{p+1}|^2 - \cdots - |y_r|^2$, 其中 r 是 f 的秩, 则*

- *f 半正定当且仅当 $p = r$;*
- *f 正定当且仅当 $p = n$.*

以 $-f$ 代 f 便可对负定和半负定性得到类似刻画.

证明 要点在于非零复数的绝对值平方总为正数. 另一方面, 倘若规范形包含 $|x_p|^2 - |x_{p+1}|^2 = (|x_p| + |x_{p+1}|)(|x_p| - |x_{p+1}|)$ 的部分, 则显然对任何 $c \in \mathbb{R}$ 皆可找到 $\boldsymbol{v} \neq \mathbf{0}$ 使 $f(\boldsymbol{v}) = c$. □

现在已有充足的准备来陈述惯性定理 8.6.6 的复版本.

定义–定理 10.2.8 *选定 $n \in \mathbb{Z}_{\geq 1}$, 则 n 元 Hermite 型的每个同构类中有唯一的规范形. 对于同构类中的任何 f, 我们称规范形中的 p 为 f 的***正惯性指数***, 称 $q := r - p$ 为***负惯性指数***, 而 $p - q = 2p - r$ 称为 f 的***符号差***.*

证明 与定理 8.6.6 的证明思路丝毫不差. 关键在于以 Hermite 形式内禀的性质来刻画规范形中的 p: 存在 p 维正定子空间 $V_+ \subset \mathbb{C}^n$, 而任何维数 $> p$ 的子空间皆非正定. 请参考引理 8.6.5 的论证; 命题 10.2.7 也是必要的. □

因此两个 n 元 Hermite 型同构当且仅当它们有相同的秩和符号差. 两个非退化 n 元 Hermite 型同构当且仅当它们的符号差相同.

对于 n 元反 Hermite 型 f, 考虑 Hermite 型 $\mathrm{i}f$ 便有对应的分类. 然而这种手段不尽自然, 因为 -1 的平方根 $\pm\mathrm{i}$ 依代数观点并无自然的选法 (它们 “唯二”), 在此就不多讨论了.

10.3 复内积空间和酉变换

本节的内容和 §9.2 处理的实内积空间几乎是平行的. 由于有了处理实内积空间的经验, 步调会适当加快.

定义 10.3.1 复向量空间 V 上的 **Hermite 内积** (又称**复内积**) 意指满足以下性质的映射 $(\cdot|\cdot): V\times V\to\mathbb{C}$.

- $(\cdot|\cdot)$ 是 V 上的 Hermite 形式 (定义 10.1.8).
- $(\cdot|\cdot)$ 正定 (定义 10.2.6).

设 V 为复向量空间, $(\cdot|\cdot)$ 为 V 上的内积, 则资料 $(V,(\cdot|\cdot))$ 统称为 **复内积空间**, **Hermite 空间**, 或**酉空间**.

对于复内积空间 V 中的向量, 定义其长度为

$$\|v\|:=\sqrt{(v|v)}.$$

因此 $\|tv\|=|t|\cdot\|v\|$ 对所有 $t\in\mathbb{C}$ 成立. 满足 $(v|w)=0$ 的向量 $v,w\in V$ 称为**正交**的. 满足 $\|v\|=1$ 的向量称为**单位向量**. 两两正交的一族非零向量称为**正交向量族**, 如果进一步要求其元素都是单位向量, 则称之为**单位正交向量族**. 由单位正交向量组成的基称为**单位正交基**.

正交向量族必然线性无关. 原因和内积空间相同: 设 $v_1,\dots,v_n\in V$ 是正交向量族, 则

$$v=\sum_{i=1}^n a_iv_n\implies\forall 1\leq i\leq n,\ a_i=\frac{(v_i|v)}{(v_i|v_i)}.$$

这里要留意[1)]的是内积必须取 $(v_i|v)$, 调换顺序得到的将是 $\overline{a_i}$.

推而广之, 若 V_0,V_1 是 V 的子空间, 而且对任意 $v_0\in V_0$ 和 $v_1\in V_1$ 皆有 $(v_0|v_1)=0$, 则称 V_0 和 V_1 正交, 记作 $V_0\perp V_1$.

定义 10.3.2 给定复内积空间 $(V,(\cdot|\cdot))$ 和一族子空间 $(V_i)_{i\in I}$, 如果

- 向量空间的直和分解 $V=\bigoplus_{i\in I}V_i$ 成立,
- 当 $i\neq j$ 时 $V_i\perp V_j$,

则称 $V=\bigoplus_{i\in I}V_i$ 为 V 的**正交直和分解**.

[1)] 如果像一些其他教材要求 $(\cdot|\cdot)$ 对第二个变元是半双线性的, 则应改取 $(v|v_i)$.

例 10.3.3 ($\mathbb{C}^n$ 上的标准 Hermite 内积) 对 $\mathbb{C}^n$ 的任意元素 $x=(x_1,\dots,x_n)$ 和 $y=(y_1,\dots,y_n)$, 定义**标准 Hermite 内积**为

$$(x|y):=\sum_{i=1}^{n}\overline{x_i}y_i,$$

这使 $(\mathbb{C}^n,(\cdot|\cdot))$ 成为复内积空间, 称为 n 维**标准复内积空间**. 对于任意矩阵 $\boldsymbol{T}\in \mathrm{M}_{n_2\times n_1}(\mathbb{C})$, 视同线性映射 $\mathbb{C}^{n_1}\to\mathbb{C}^{n_2}$, 则相对于标准 Hermite 内积, 其伴随映射 $\boldsymbol{T}^*$ 由矩阵 ${}^\dagger\boldsymbol{T}$ 给出, 这是在例 10.1.11 中代入 $\boldsymbol{A}_i=\mathbf{1}_{n_i\times n_i}$ 的结论.

以下观察将大有用处.

引理 10.3.4 设 $(V,(\cdot|\cdot))$ 是复内积空间, 则对所有 $v,w\in V$ 皆有

$$\|v+w\|^2=\|v\|^2+2\,\mathrm{Re}(v|w)+\|w\|^2.$$

证明 直接应用 (10.2.2), 代入 $\epsilon=1$ 和 $f(v)=\|v\|^2$. □

下面的结果因而是一望可知的.

命题 10.3.5 (复内积空间的勾股定理) 设 v,w 是复内积空间 $(V,(\cdot|\cdot))$ 中的正交向量, 则

$$\|v+w\|^2=\|v\|^2+\|w\|^2.$$

定理 10.3.6 (复内积空间的 Cauchy–Bunyakovsky–Schwarz 不等式) 设 $(V,(\cdot|\cdot))$ 是复内积空间, 则对所有 $v,w\in V$ 皆有

$$|(v|w)|^2\le(v|v)(w|w),$$

等式成立当且仅当 v, w 线性相关.

证明 若 $w=0$ 则无事可作, 故以下设 $w\neq 0$. 对任意 $t\in\mathbb{C}$, 我们有

$$0\le\|v+tw\|^2=\|v\|^2+2\,\mathrm{Re}\,(t(v|w))+|t|^2\|w\|^2.$$

代入 $t:=-\frac{(w|v)}{\|w\|^2}$ 可得 $0\le\|v\|^2-\frac{|(v|w)|^2}{\|w\|^2}$; 若等式成立则 $v+tw=0$.

反过来说, 若 v, w 线性相关, 亦即成比例, 则等式显然成立[2)]. □

推论 10.3.7 (复内积空间的三角不等式) 设 $(V,(\cdot|\cdot))$ 是复内积空间, 则对所有 $v,w\in V$ 皆有

$$\|v+w\|\le\|v\|+\|w\|,$$

等式成立当且仅当存在 $t\in\mathbb{R}_{\ge 0}$ 使得 $v=tw$ 或 $w=tv$.

[2)] 一些教材对 Cauchy–Bunyakovsky–Schwarz 不等式给出了基于正交分解的几何诠释, 这和上述论证实质上是一回事.

证明 以命题 10.3.6 推导

$$\begin{aligned}\|v+w\|^2 &= \|v\|^2 + 2\,\mathrm{Re}(v|w) + \|w\|^2\\ &\le \|v\|^2 + 2|(v|w)| + \|w\|^2\\ &\le \|v\|^2 + 2\|v\|\cdot\|w\| + \|w\|^2 = (\|v\|+\|w\|)^2.\end{aligned}$$

此即所求的不等式. 使等号成立的条件是

- ⋆ 首先 v 和 w 线性相关, 亦即存在 $t\in\mathbb{C}$ 使得 $v=tw$ 或 $w=tv$;
- ⋆ 其次, $\mathrm{Re}(v|w)=|(v|w)|$.

将第一则条件的 $v=tw$ 或 $w=tv$ 代入第二则条件, 立见 $t\in\mathbb{R}_{\ge 0}$. □

内积空间的大多数基本性质都可以移植到复内积空间上. 举例说明如下.

▷ Gram–Schmidt **正交化** 设 V 中的一族向量 $v_1, v_2, \dots$ (或有限或无穷) 线性无关, 递归地定义

$$\begin{aligned}w_1 &:= v_1,\\ w_k &:= v_k - \sum_{i=1}^{k-1} \frac{(w_i|v_k)}{(w_i|w_i)}\cdot w_i, \quad k>1,\end{aligned}$$

则 $w_1, w_2, \dots$ 是正交向量族, 而且对所有 $k\ge 1$ 皆有 $\langle v_1,\dots,v_k\rangle = \langle w_1,\dots,w_k\rangle$.

▷ **单位正交基的扩充** 任何有限维复内积空间中的单位正交向量族皆可扩充为单位正交基.

▷ **单位正交基的存在性** 任何有限维复内积空间皆有单位正交基.

▷ **正交补** 对任意子空间 $V_0\subset V$, 定义

$$V_0^{\perp} := \{v\in V : \forall w\in V_0,\ (v|w)=0\};$$

这是 V 的子空间. 当 V_0 有限维时我们有正交直和分解

$$V = V_0 \oplus V_0^{\perp}.$$

作为 Gram–Schmidt 正交化的应用, 推论 9.4.7 的 $\boldsymbol{QR}$ 分解也适用于复矩阵, 留作本章的简单习题.

定义 10.3.8 设 $(V,(\cdot|\cdot)_V)$ 和 $(W,(\cdot|\cdot)_W)$ 为复内积空间.

(i) 若线性映射 $T: V\to W$ 对所有 $v\in V$ 皆满足 $\|Tv\|_W = \|v\|_V$, 则称为**保距**的.

(ii) 若 $V \underset{S}{\overset{T}{\rightleftarrows}} W$ 是一对保距线性映射, 使得 $TS = \mathrm{id}_W$, $ST = \mathrm{id}_V$, 则称 S 和 T 为复内积空间的互逆**同构**.

以下几点注记和实内积空间的情形完全类似.

⋆ 保距等价于保内积: 若 $(Tv_1 \mid Tv_2)_W = (v_1|v_2)_V$ 对所有 $v_1, v_2 \in V$ 成立, 则代入 $v_1 = v_2$ 可见 T 保距. 反之设 T 保距, 则由

$$\begin{aligned} \|T(v_1+v_2)\|^2 &= \|Tv_1\|^2 + 2\,\mathrm{Re}\,(Tv_1 \mid Tv_2)_W + \|Tv_2\|^2, \\ \|v_1+v_2\|^2 &= \|v_1\|^2 + 2\,\mathrm{Re}(v_1|v_2) + \|v_2\|^2 \end{aligned}$$

可见 T 至少保持内积的实部 $\mathrm{Re}(\cdot|\cdot)$. 然而 T 是线性映射, 而

$$\mathrm{Im}(v_1|v_2) = \mathrm{Re}(\mathrm{i}v_1|v_2),$$

所以 T 也保持内积的虚部 $\mathrm{Im}(\cdot|\cdot)$.

⋆ 若 $T : V \to W$ 是保距线性映射, 而且它作为线性映射是可逆的, 则 T^{-1} 也保距, 因而 T 是复内积空间的同构.

⋆ 设 $n := \dim V \in \mathbb{Z}_{\geq 0}$, 则 V 的有序单位正交基一一对应于同构

$$(V, (\cdot|\cdot)) \xrightarrow{\sim} (\mathbb{C}^n, \text{标准 Hermite 内积});$$

设 $v_1, \ldots, v_n$ 是单位正交基, 计顺序, 则对应的同构映 v_i 为 $\boldsymbol{e}_i$, 其中 $1 \leq i \leq n$.

由于 V 总有单位正交基, 这说明任何有限维复内积空间都同构于标准复内积空间 $\mathbb{C}^n$.

定义 10.3.9 从复内积空间 $(V, (\cdot|\cdot))$ 到其自身的同构称为 V 上的**酉变换**.

注意到当复内积空间 V 维数有限时, 正定条件蕴涵 $(\cdot|\cdot) : V \times V \to \mathbb{C}$ 在定义 10.1.5 意义下非退化. 以下命题因而有意义.

命题 10.3.10 设 $(V, (\cdot|\cdot))$ 是有限维复内积空间. 线性映射 $T \in \mathrm{End}(V)$ 是酉变换的充要条件是 $T^* = T^{-1}$.

证明 由于此时的酉变换无非是保持 Hermite 内积的线性变换, 论证和命题 9.4.1 无异. □

命题 9.4.2 同样具备以下的酉版本, 证明一字不易. 设 $T : V \to W$ 是有限维复内积空间之间的线性映射, $v_1, \ldots, v_n$ 是 V 的单位正交基, 则 T 是复内积空间的同构当且仅当 $Tv_1, \ldots, Tv_n$ 是 W 的单位正交基.

取 $(V,(\cdot|\cdot))$ 为例 10.3.3 的标准复内积空间, 我们推知 $\boldsymbol{T}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 对应的线性映射是酉变换当且仅当 ${}^{\dagger}\boldsymbol{T}=\boldsymbol{T}^{-1}$. 由于

$$\boldsymbol{T} \text{ 是酉变换} \iff \boldsymbol{T}\boldsymbol{e}_1,\dots,\boldsymbol{T}\boldsymbol{e}_n \text{ 是 } \mathbb{C}^n \text{ 的单位正交基},$$

酉变换也可以等价地刻画为形如 $(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)$ 的 $n\times n$ 矩阵, 其中的列向量 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$ 构成标准复内积空间 $\mathbb{C}^n$ 的单位正交基.

定义 10.3.11 满足 ${}^{\dagger}\boldsymbol{P}=\boldsymbol{P}^{-1}$ 的矩阵 $\boldsymbol{P}\in \mathrm{M}_{n\times n}(\mathbb{C})$ 称为**酉矩阵**.

之前的讨论表明 $n\times n$ 酉矩阵无非是标准复内积空间 $\mathbb{C}^n$ 上的酉变换.

下述结果是推论 9.6.4 的酉版本, 陈述及证明皆无异.

命题 10.3.12 对有限维复内积空间 V, W 和线性映射 $T:V\to W$, 线性映射 $T^*T\in \mathrm{End}(V)$ 和 $TT^*\in \mathrm{End}(W)$ 皆自伴, 而且

$$\begin{aligned} \mathrm{im}\,(T^*T)&=\mathrm{im}\,(T^*), \quad \mathrm{im}\,(TT^*)=\mathrm{im}\,(T),\\ \ker(T^*T)&=\ker(T), \quad \ker(TT^*)=\ker(T^*),\\ \mathrm{rk}\,(T^*T)&=\mathrm{rk}(T), \quad \mathrm{rk}\,(TT^*)=\mathrm{rk}\,(T^*). \end{aligned}$$

10.4 正规算子的酉对角化

本节研究从选定的有限维复内积空间 $(V,(\cdot|\cdot))$ 到其自身的正规算子, 见定义 10.1.14. 我们的目标是对正规算子证明所谓的谱定理或谱分解, 具体陈述如下.

定理 10.4.1 (正规算子的谱分解) 设 $T\in \mathrm{End}(V)$, 则以下两者等价:

(i) 存在单位正交基 $v_1,\dots,v_n\in V$ 和 $\lambda_1,\dots,\lambda_n\in\mathbb{C}$, 使得

$$Tv_i=\lambda_i v_i,\quad 1\le i\le n.$$

(ii) T 是正规算子.

若将 $(V,(\cdot|\cdot))$ 具体地取为标准复内积空间 $\mathbb{C}^n$, 其元素视同列向量, 并将 $T\in \mathrm{End}(V)$ 等同于矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{C})$, 则 (i) 相当于说 $\boldsymbol{A}$ 能作" 酉对角化", 亦即

$$\exists\, \boldsymbol{P}=(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)\in \mathrm{M}_{n\times n}(\mathbb{C}),\quad \exists\,\lambda_1,\dots,\lambda_n\in\mathbb{C},$$

$${}^{\dagger}\boldsymbol{P}=\boldsymbol{P}^{-1},\quad \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

另一方面, (ii) 相当于说 ${}^{\dagger}\boldsymbol{A}\,\boldsymbol{A} = \boldsymbol{A}\,{}^{\dagger}\boldsymbol{A}$.

难点在于 (ii) $\Longrightarrow$ (i). 我们需要借一系列的简单引理来推进. 回忆到取伴随映射的运算满足 $(t_1N_1 + t_2N_2)^* = \overline{t_1}\,N_1{}^* + \overline{t_2}\,N_2{}^*$ 和 $(N_1N_2)^* = N_2{}^*\,N_1{}^*$.

引理 10.4.2 设 $N \in \mathrm{End}(V)$ 是正规算子, $f \in \mathbb{C}[X]$, 则 $f(N) \in \mathrm{End}(V)$ 也是正规的.

证明 易见 $\left(\sum_{i=0}^m a_iN^i\right)^* = \sum_{i=0}^m \overline{a_i}\,(N^*)^i$, 其中 $a_0, \dots, a_m \in \mathbb{C}$, 这和 N 的任何多项式都交换. □

引理 10.4.3 设 $N \in \mathrm{End}(V)$ 是正规算子, $\lambda \in \mathbb{C}$, 而 $v \in V$. 若 $Nv = \lambda v$ 则 $N^*v = \overline{\lambda}v$.

证明 取 $M := N - \lambda \cdot \mathrm{id}_V$, 这仍是正规的, 而 $M^* = N^* - \overline{\lambda} \cdot \mathrm{id}_V$. 条件 $Nv = \lambda v$ 相当于 $v \in \ker(M)$, 而命题 10.3.12 连同 M 的正规性所蕴涵的 $\ker(M) = \ker(M^*\,M) = \ker(MM^*) = \ker(M^*)$ 又导致 $N^*v = \overline{\lambda}v$. □

引理 10.4.4 设 $N \in \mathrm{End}(V)$ 是正规算子, 若存在 $k \in \mathbb{Z}_{\geq 1}$ 使得 $N^k = 0$, 则 $N = 0$.

证明 首先假设 N 自伴; 于是 N 的任意幂依然自伴. 先处理 $k = 2$ 的情形: 对任意 $v \in V$, 我们有 $(Nv|Nv) = (N^2v|v) = 0$, 故 $N = 0$ 成立.

对于一般的 $k \geq 2$, 当 k 是偶数时 $(N^{k/2})^2 = N^k = 0$ 导致 $N^{k/2} = 0$, 而当 k 是奇数时 $(N^{(k+1)/2})^2 = N^{k+1} = 0$; 按此减小 k 直到导出 $N = 0$, 完事.

现在去除 N 自伴的假设. 从正规性和 $N^k = 0$ 可得 $(N^*\,N)^k = (N^*)^k\,N^k = 0$. 已知 $N^*\,N$ 自伴, 故 $N^*\,N = 0$, 继而由命题 10.3.12 推得 $N = 0$. □

引理 10.4.5 设 $N \in \mathrm{End}(V)$ 是正规算子, 对应到 $\lambda, \mu \in \mathbb{C}$ 的特征子空间照例分别记为 V_λ 和 V_μ, 则当 $\lambda \neq \mu$ 时 $V_\lambda \perp V_\mu$.

证明 任取 $v \in V_\lambda$ 和 $w \in V_\mu$, 引理 10.4.3 蕴涵

$$(v|Nw) = (v|\mu w) = \mu(v|w),$$
$$(N^*v|w) = (\overline{\lambda}v|w) = \lambda(v|w).$$

由于 $\lambda \neq \mu$, 两式相等导致 $(v|w) = 0$. □

证明 (定理 10.4.1) 先检验 (i) $\Longrightarrow$ (ii). 对于所示的 $\lambda_1, \dots, \lambda_n$ 和单位正交基 $v_1, \dots, v_n$, 定义线性映射

$$S : V \to V, \quad Sv_i = \overline{\lambda_i}v_i.$$

兹断言 $S = T^*$. 首先, 对所有 $1 \leq i, j \leq n$ 皆有 $(v_i|Sv_j) = (Tv_i|v_j)$; 既然 $(\cdot|S(\cdot))$ 和 $(T(\cdot)|\cdot)$ 都是半双线性形式, 由此立见 $(v|Sw) = (Tv|w)$ 恒成立.

此外显然有 $ST = TS$, 这就表明 T 正规.

接着检验 (ii) $\Longrightarrow$ (i). 将特征多项式 Char_T 在 $\mathbb{C}[X]$ 中分解成一次因式

$$\mathrm{Char}_T = \prod_{i=1}^k (X - \mu_i)^{a_i},$$

其中 $\mu_1,\dots,\mu_k\in\mathbb{C}$ 两两相异, 而 $a_i\in\mathbb{Z}_{\geq 1}$. 定义

$$m:=\prod_{i=1}^{k}(X-\mu_i)\ \in\mathbb{C}[X],$$

取 $a:=\max\{a_1,\dots,a_k\}$, 则 $\mathrm{Char}_T\mid m^a$, 故 $m(T)^a=0$. 然而 $m(T)$ 正规, 故引理 10.4.4 蕴涵 $m(T)=0$.

于是 T 的极小多项式 Min_T 也整除 m. 既然 m 无重根, Min_T 亦无重根. 将此代入定理 7.2.9 立见 T 可对角化; 换言之, 我们有直和分解

$$V=V_{\mu_1}\oplus\cdots\oplus V_{\mu_k}.$$

引理 10.4.5 说明上式为正交直和分解. 从每个特征子空间 V_{μ_i} 任取单位正交基, 取并给出 V 的单位正交基 $v_1,\dots,v_n$, 其中每个 v_i 都是 T 的特征向量. 明所欲证. □

注记 10.4.6 正规算子 $T\in\mathrm{End}(V)$ 的谱分解定理 10.4.1 也经常改写作

$$T=\sum_{i=1}^{k}\mu_iP_i,$$

的形式, 其中

- ⋆ $\mu_1,\dots,\mu_k\in\mathbb{C}$,
- ⋆ P_i 是向某个子空间 V_i 的正交投影 $(i=1,\dots,k)$,
- ⋆ 有正交直和分解 $V=V_1\oplus\cdots\oplus V_k$;

事实上, $\mu_1,\dots,\mu_k$ 正是证明 (ii) $\Longrightarrow$ (i) 时考虑的相异特征值, 而 $V_i=V_{\mu_i}$.

基于酉对角化定理, 本章习题将说明正规算子 N 的伴随 N^* 必为 N 的多项式.

推论 10.4.7 设 $T\in\mathrm{End}(V)$ 是正规算子, 则:

- ⋆ T 自伴当且仅当所有特征值 λ 都满足 $\lambda\in\mathbb{R}$.
- ⋆ T 反自伴当且仅当所有特征值 λ 都满足 $\lambda\in\mathrm{i}\mathbb{R}$.
- ⋆ T 是酉变换当且仅当所有特征值 λ 都满足 $|\lambda|=1$.

证明 基于定理 10.4.1, 可以选取单位正交基以化约到 $V=\mathbb{C}^n$ 配备标准 Hermite 内积, 而

$$T\text{ 对应到矩阵 }\boldsymbol{A}=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}$$

的情形; 对角元 $\lambda_1,\dots,\lambda_n$ 无非是 T 的所有特征值, 记入重数. 此时

$$T^* \text{ 对应到矩阵 } {}^\dagger\boldsymbol{A} = \begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix}.$$

关于 T 得自伴, 反自伴和酉变换性质分别转译为 ${}^\dagger\boldsymbol{A} = \boldsymbol{A}$, ${}^\dagger\boldsymbol{A} = -\boldsymbol{A}$ 和 ${}^\dagger\boldsymbol{A} = \boldsymbol{A}^{-1}$. 由此容易完成证明. □

算法 10.4.8 谱分解或者说矩阵的酉对角化的计算与一般的对角化算法 7.1.10 类似.

1. 首先为有限维复内积空间上的正规算子 $T \in \mathrm{End}(V)$ 找出所有特征值.
2. 其次, 对每个特征值 λ 确定相应的特征子空间 V_λ, 写下它的基.
3. 接着施行 Gram–Schmidt 正交化, 以得到每个 V_λ 的单位正交基 B_λ.

由于谱分解定理 10.4.1 一方面确保 $V = \bigoplus_\lambda V_\lambda$, 另一方面又确保 $\lambda \neq \mu \implies V_\lambda \perp V_\mu$ (后者也是引理 10.4.5 的内容), 所以 $\bigsqcup_\lambda B_\lambda =: \{v_1,\dots,v_n\}$ 给出 V 的单位正交基, 连同 $\lambda_1,\dots,\lambda_n \in \mathbb{C}$, 使得 $Tv_i = \lambda_i v_i$ 恒成立.

10.5 实定理的复推广

我们在第九章介绍了许多和实内积空间相关的基本定理. 它们大部分都有复内积空间上的类比. 我们先从定理 9.7.3 的类比入手.

定理 10.5.1 设 f 为 n 元 Hermite 型, 对应的 Hermite 矩阵记为 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$, 则 f 正定 (或半正定) 当且仅当 $\boldsymbol{A}$ 的所有特征值皆正 (或非负).

证明 对 Hermite 矩阵 $\boldsymbol{A}$ 应用定理 10.4.1 可得

$${}^\dagger\boldsymbol{C}\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \qquad \begin{aligned} &\boldsymbol{C} \in \mathrm{M}_{n\times n}(\mathbb{C}),\ {}^\dagger\boldsymbol{C} = \boldsymbol{C}^{-1},\\ &\lambda_1,\dots,\lambda_n \in \mathbb{R}:\ \boldsymbol{A} \text{ 的特征值, 记重数.} \end{aligned}$$

这相当于通过 $\boldsymbol{C}$ 作变量代换, 化 f 为 $\sum_{i=1}^n \lambda_i |x_i|^2$; 继续对满足 $\lambda_i \neq 0$ 的 i 作代换 $y_i := \sqrt{|\lambda_i|}x_i$, 便将 Hermite 型的系数化到 $\{-1,0,1\}$ 中.

然而上式连同命题 10.2.7 足以说明 f 正定 (或半正定) 当且仅当 $\lambda_1,\dots,\lambda_n$ 全为正 (或非负). □

读者应该已经察觉这几乎是原封不动地照搬定理 9.7.3 的论证, 以下几则定义和定理也莫不如此. 证明既相似, 不如节约笔墨.

定义 10.5.2 设 $(V,(\cdot|\cdot))$ 是有限维复内积空间. 若 $T\in\mathrm{End}(V)$ 自伴, 而且 $(v_1,v_2)\mapsto(Tv_1|v_2)$ 是 V 上的正定 (或半正定) Hermite 形式, 则称 T 是正定 (或半正定) 的.

练习 10.5.3 对应于实版本的练习 9.6.5, 请用相同方法对有限维复内积空间之间的线性映射 $T:V\to W$ 证明 $\ker(T^*)=\mathrm{im}(T)^\perp$. 由此可知 T^* 单等价于 T 满.

引理 10.5.4 设 $T:V\to W$ 为有限维复内积空间之间的线性映射, 则 T^*T (或 TT^*) 是半正定的; 若 T 单 (或 T^* 单, 等价说法是 T 满), 则它们正定.

证明 与引理 9.7.6 全同. 以 T^*T 情形为例, 这是等式 $(T^*Tv\mid v)_V=(Tv|Tv)_W$ 的直接结论. □

定义–命题 10.5.5 设 $T\in\mathrm{End}(V)$ 正定 (或半正定), 则存在唯一的 $S\in\mathrm{End}(V)$ 使得 S 正定 (或半正定) 而且 $S^2=T$. 因此, 这样的 S 可以合理地记为 $\sqrt{T}$.

证明 鉴于正定 (或半正定) 变换的特征值全为正 (或非负) 实数这一事实, 论证和定义–命题 9.7.7 全同. □

以下是极分解定理的酉版本.

定理 10.5.6 (极分解) 设 $(V,(\cdot|\cdot))$ 为有限维复内积空间, $T\in\mathrm{End}(V)$ 可逆, 则存在唯一一对 $R,U\in\mathrm{End}(V)$, 使得 R 正定, U 是酉变换, 而且 $T=RU$.

证明 与定理 9.7.8 全同, 重点是取 $R=\sqrt{TT^*}$. □

定理 10.5.7 (奇异值分解) 对于任意复内积空间之间的线性映射 $T:V\to W$, 记 $p:=\min\{m,n\}$, 则存在

- V 的单位正交基 $v_1,\dots,v_m$,
- W 的单位正交基 $w_1,\dots,w_n$,
- 非负实数 $\sigma_1\geq\cdots\geq\sigma_p$,

使得

$$Tv_i=\begin{cases}\sigma_i w_i, & 1\leq i\leq p,\\ 0, & i>p.\end{cases}$$

此处的 $\sigma_1\geq\cdots\geq\sigma_p$ 由 T 唯一确定, 称为 T 的**奇异值**.

证明 与定理 9.8.1 全同. 对于心存疑问的读者, 请注意到关键是对半正定的 T^*T 作酉对角化; 对 T^*T 的特征值取平方根, 降序排列即是奇异值. □

为了得到奇异值分解的矩阵版本, 现在设 $V=\mathbb{C}^m$, $W=\mathbb{C}^n$, 各自配备标准 Hermite 内积, 将 T 等同于矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times m}(\mathbb{C})$. 将定理 10.5.7 中的单位正交基视同列向量, 以之定义酉矩阵

$$\boldsymbol{P}:=(\boldsymbol{v}_1|\cdots|\boldsymbol{v}_m)\in\mathrm{M}_{m\times m}(\mathbb{C}),\quad \boldsymbol{Q}:=(\boldsymbol{w}_1|\cdots|\boldsymbol{w}_n)\in\mathrm{M}_{n\times n}(\mathbb{C}),$$

另外用奇异值定义

$$\boldsymbol{\Sigma} := \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \boxed{\text{补 } 0} \end{pmatrix} \in \mathrm{M}_{n\times m}(\mathbb{R}),$$

奇异值分解遂化为矩阵等式

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{\Sigma},$$

亦即

$$^{\dagger}\boldsymbol{Q}\,\boldsymbol{A}\,\boldsymbol{P} = \boldsymbol{\Sigma} \quad \text{或} \quad \boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{\Sigma}\,{}^{\dagger}\boldsymbol{P}.$$

以下接续 §9.9 关于 Moore–Penrose 的讨论, 但考量的是有限维复内积空间 V 和 W, 以及其间的线性映射.

定义–定理 10.5.8 (Moore–Penrose 广义逆) 相对于给定的 $T : V \to W$, 满足下述条件的线性映射 $S : W \to V$ 称为 T 的 Moore–Penrose 广义逆.

(MP.1) $TST = T$,

(MP.2) $STS = S$,

(MP.3) $TS = (TS)^*$,

(MP.4) $ST = (ST)^*$.

对于给定的 T, Moore–Penrose 广义逆存在而且唯一.

证明 存在和唯一性的论证与定理 9.9.3 相同, 关键在于运用伴随 $(\cdots)^*$ 的运算规律以及正交投影的性质, 这点在内积空间和在复内积空间中并无二致. □

为了联系奇异值分解和 Moore–Penrose 广义逆, 以下沿用定理 10.5.7 中关于 $T : V \to W$ 的奇异值分解的相关符号, 命 $r := \mathrm{rk}(T)$.

命题 10.5.9 取定上述资料, 定义线性映射 $S : W \to V$ 使得

$$\begin{aligned} S &: W \to V \\ w_j &\mapsto \begin{cases} \sigma_j^{-1} v_j, & 1 \leq j \leq r, \\ 0, & r < j \leq n. \end{cases} \end{aligned}$$

则 S 是 T 的 Moore–Penrose 广义逆.

证明 与命题 9.9.4 全同, 同样是基于正交投影的性质. □

取 $V=\mathbb{C}^m$ 和 $W=\mathbb{C}^n$, 赋予标准 Hermite 内积, 并将 T 等同于矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times m}(\mathbb{C})$. 若有奇异值分解

$$\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \boxed{\text{补 }0}\end{pmatrix}_{n\times m}{}^{\dagger}\boldsymbol{P},$$

则 T 的 Moore–Penrose 广义逆对应于矩阵

$$\boldsymbol{P}\begin{pmatrix}\sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & \boxed{\text{补 }0}\end{pmatrix}_{m\times n}{}^{\dagger}\boldsymbol{Q}.$$

练习 10.5.10 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$ 对称. 试尽量直接从定义出发, 证明 $\mathbb{R}^n$ 上的对称双线性形式 $(\boldsymbol{v}_1,\boldsymbol{v}_2)\mapsto {}^{\mathrm{t}}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_2$ 正定 (或半正定) 当且仅当 $\mathbb{C}^n$ 上的 Hermite 形式 $(\boldsymbol{v}_1,\boldsymbol{v}_2)\mapsto {}^{\dagger}\boldsymbol{v}_1\,\boldsymbol{A}\,\boldsymbol{v}_2$ 正定 (或半正定).

练习 10.5.11 在前一道练习的基础上, 考虑定理 10.5.1 — 定理 10.5.6 以及定义–定理 10.5.8 的矩阵表述, 说明如何由之推导相应的实版本.

提示〉 以极分解为例, 为了从复版本推导实版本, 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(\mathbb{R})$ 可逆; 复版本给出唯一分解 $\boldsymbol{A}=\boldsymbol{R}\boldsymbol{U}$, 其中 $\boldsymbol{R}$ 正定而 $\boldsymbol{U}$ 是酉矩阵. 取矩阵的复共轭 (约定 9.5.3) 得 $\boldsymbol{A}=\overline{\boldsymbol{A}}=\overline{\boldsymbol{R}}\,\overline{\boldsymbol{U}}$. 说明 $\overline{\boldsymbol{R}}$ (或 $\overline{\boldsymbol{U}}$) 仍是正定 (或酉) 的, 从而 $\overline{\boldsymbol{R}}=\boldsymbol{R}$ 而 $\overline{\boldsymbol{U}}=\boldsymbol{U}$, 故两者皆属于 $\mathrm{M}_{n\times n}(\mathbb{R})$.

上述练习之所以限于矩阵, 是为了避开有限维实向量空间及其内积如何 “复化” 的问题. 对于 $\mathbb{R}^n$ 及其标准内积, 复化的选择当然是 $\mathbb{C}^n$ 及其标准 Hermite 内积, 此即矩阵表述的实质. 对于一般的空间, 我们需要的是一种典范的, 毋须选基的构造方式. 工序不算棘手, 但最好在一个更自然的时机来探讨这一问题.

练习 10.5.12 对有限维复内积空间 $(V,(\cdot|\cdot))$ 上的 Hermite 形式 B, 表述并证明 Courant–Fischer 定理 9.10.3 的相应版本.

10.6 实正交变换的标准形

本节取 $(V,(\cdot|\cdot))$ 为有限维实内积空间. 回忆到 $T\in\mathrm{End}(V)$ 有相对于实内积的伴随 T^*, 不必取分左右. 对此, 定义 10.1.14 的实版本表述如下.

定义 10.6.1 若线性映射 $T\in\mathrm{End}(V)$ 满足 $T^*T=TT^*$, 则称为**正规**的.

若选取 V 的单位正交基将 T 等同于矩阵 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{R})$, 则条件相当于

$$^{\mathrm{t}}\boldsymbol{A}\,\boldsymbol{A}=\boldsymbol{A}\,{}^{\mathrm{t}}\boldsymbol{A}.$$

因为 $^{\mathrm{t}}\boldsymbol{A}={}^{\dagger}\boldsymbol{A}$, 这相当于说 $\boldsymbol{A}$ 作为 $\mathrm{M}_{n\times n}(\mathbb{C})$ 的元素对应到标准复内积空间 $\mathbb{C}^n$ 上的正规变换.

若 $T\in\mathrm{End}(V)$ 正规, $f\in\mathbb{R}[X]$, 则 $f(T)$ 仍是正规变换; 从 V 到自身的自伴, 反自伴和正交线性映射皆正规. 这一切都是复版本的重新搬演, 也可以通过矩阵表法归结到复的情形.

引理 10.6.2 设正规线性映射 $T\in\mathrm{End}(V)$ 满足 $T^k=0$, 其中 $k\in\mathbb{Z}_{\geq 1}$, 则 $T=0$.

证明 按先前的方法, 取基化约到关于矩阵的酉版本, 即引理 10.4.4, 或者照搬该处的证明. □

本节的重点是 V 上的正交变换, 我们将用上述工具来确立正交变换的标准形. 这并非唯一途径, 可能也不是最短途径, 却有助于揭示更多技巧.

首先, $\dim V=1$ 的情形是容易的: 正交变换只能是 $\pm\mathrm{id}_V$. 其次是 $\dim V=2$ 的情形: 选取单位正交基以化约到 $V=\mathbb{R}^2$ 而 $(\cdot|\cdot)$ 是标准内积的情形. 正交变换 $T\in\mathrm{End}(V)$ 等同于由下述条件刻画的 2×2 正交矩阵

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix},\qquad \begin{aligned}\alpha^2+\gamma^2=1=\beta^2+\delta^2\\ \alpha\beta+\gamma\delta=0.\end{aligned}$$

我们有 $\alpha\delta-\beta\gamma=\det T=\pm 1$. 既然 (α,γ) 和 (β,δ) 都在单位圆上, 可取 $\theta\in\mathbb{R}$ 使得 $(\alpha,\gamma)=(\cos\theta,\sin\theta)$; 从几何来看, 与 (α,γ) 正交的单位向量 (β,δ) 恰有两个, 从代数观点分别对应到线性方程

$$\beta\cos\theta+\delta\sin\theta=0$$

唯二的单位向量解 $(\beta,\delta)=\pm(-\sin\theta,\cos\theta)$, 此处的符号 $\pm$ 正是 $\det T$. 举例明之, 若取 $\theta=0$ 而符号 $\pm$ 取为负, 相应的正交变换便是

$$\begin{pmatrix}1 & \\ & -1\end{pmatrix}:\quad \text{对 } X \text{ 轴作镜射}$$

另一方面, $\det T = 1$ 的情形则相当于在上述公式中取符号 $\pm$ 为正, 相应的矩阵称为旋转矩阵.

定义 10.6.3 (旋转矩阵) 对所有 $\theta \in \mathbb{R}$, 定义实 2×2 矩阵

$$\boldsymbol{R}(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

相对于平面 $\mathbb{R}^2$ 的标准直角坐标系, $\boldsymbol{R}(\theta)$ 无非是按逆时针方向旋转 θ 的线性变换.

根据关于旋转的几何直观, 或者三角函数的公式, 容易得到

$$\boldsymbol{R}(\theta_1+\theta_2) = \boldsymbol{R}(\theta_1)\boldsymbol{R}(\theta_2), \quad \boldsymbol{R}(0) = \boldsymbol{R}(2\pi) = \mathbf{1}_{2\times 2}, \quad \boldsymbol{R}(\pi) = -\mathbf{1}_{2\times 2}.$$

对于 $\theta_1, \theta_2 \in \mathbb{R}$, 若 $\theta_1 - \theta_2 \in 2\pi\mathbb{Z}$, 则记为 $\theta_1 \equiv \theta_2 \pmod{2\pi}$. 这是 $\mathbb{R}$ 上的等价关系; 含 θ 的等价类也记为 $\theta \bmod 2\pi$. 我们有

$$\boldsymbol{R}(\theta_1) = \boldsymbol{R}(\theta_2) \iff e^{\mathrm{i}\theta_1} = e^{\mathrm{i}\theta_2} \iff \theta_1 \equiv \theta_2 \pmod{2\pi}.$$

平面上所有以原点为圆心的旋转因之得到了参数化. 严格来说, 转角 θ 依赖于 $\mathbb{R}^2$ 的单位正交基的选取: 换基可能导致转角相差一个负号.

引理 10.6.4 符号如上. 对正交矩阵 $\boldsymbol{P} \in \mathrm{M}_{2\times 2}(\mathbb{R})$ 记 $\epsilon := \det \boldsymbol{P}$, 则 $\boldsymbol{P}^{-1}\boldsymbol{R}(\theta)\boldsymbol{P} = \boldsymbol{R}(\epsilon\theta)$.

证明 若 $\epsilon = 1$, 则之前的推导表明存在 $\psi \in \mathbb{R}$ 使得 $\boldsymbol{P} = \boldsymbol{R}(\psi)$, 此时 $\boldsymbol{P}^{-1}\boldsymbol{R}(\theta)\boldsymbol{P} = \boldsymbol{R}(-\psi)\boldsymbol{R}(\theta)\boldsymbol{R}(\psi) = \boldsymbol{R}(-\psi+\theta+\psi) = \boldsymbol{R}(\theta)$.

若 $\epsilon = -1$, 则因为正交矩阵的乘积依然正交, 故 $\boldsymbol{P}\left(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right)$ 是行列式为 1 的正交矩阵, 形如 $\boldsymbol{R}(\psi)$. 注意到 $\left(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right)^2 = \mathbf{1}_{2\times 2}$, 因此 $\boldsymbol{P} = \boldsymbol{R}(\psi)\left(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right)$. 此时

$$\boldsymbol{P}^{-1}\boldsymbol{R}(\theta)\boldsymbol{P} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \boldsymbol{R}(\theta) \begin{pmatrix} 1 & \\ & -1 \end{pmatrix},$$

简单的计算表明产物正是 $\boldsymbol{R}(-\theta)$. □

高维情形的实正交变换仍是通过旋转变换来分类的. 为此还需要一些准备.

引理 10.6.5 设 $T \in \mathrm{End}(V)$ 是正交变换, 则 $T + T^{-1} \in \mathrm{End}(V)$ 自伴, 其所有特征值 λ 皆满足 $|\lambda| \leq 2$.

证明 以 $T^* = T^{-1}$ 直接计算得 $\left(T+T^{-1}\right)^* = T^{-1} + T$, 故 $T + T^{-1}$ 自伴. 取单位正交基将 T 等同于 $\mathrm{M}_{n\times n}(\mathbb{R})$ 中的正交矩阵, 它也是 $\mathrm{M}_{n\times n}(\mathbb{C})$ 中的酉矩阵, 从而可以在 $\mathbb{C}$ 上对角化[3]. 于是 $T + T^{-1}$ 的特征值皆形如 $\lambda = \mu + \mu^{-1}$, 其中 $\mu \in \mathbb{C}$ 是 T 的特征值. 但 $|\mu| = 1$, 故 $|\lambda| \leq |\mu| + |\mu^{-1}| = 2$. □

[3] 对于此处的论证, 上三角化已经足够了.

定理 10.6.6 (正交变换的标准形)　记 $n := \dim V$. 设 $T \in \mathrm{End}(V)$ 为正交变换, 则存在 V 的单位正交基, 使得 T 在此基之下写作分块对角矩阵

$$\begin{pmatrix} \boxed{\mathbf{1}_{a\times a}} & & & & \\ & \boxed{-\mathbf{1}_{b\times b}} & & & \\ & & \boxed{\boldsymbol{R}(\theta_1)} & & \\ & & & \ddots & \\ & & & & \boxed{\boldsymbol{R}(\theta_m)} \end{pmatrix},$$

其中 $a+b+2m = n$ 而 $\theta_1, \dots \theta_m \in \mathbb{R}$ 满足 $\theta_i \notin \mathbb{Z}\pi$. 资料 (a,b,m) 和 $(\pm\theta_i \mod 2\pi)_{1\le i\le m}$ (计重数, 精确到重排) 由 T 唯一地确定.

证明　命 $S := T + T^{-1}$, 引理 10.6.5 说明它自伴. 于是定理 9.5.2 蕴涵 V 是 S 的所有特征子空间 V_λ 的正交直和. 留意到 $ST = TS$ 蕴涵 V_λ 是 T-不变的, 见引理 7.5.2. 所求的分解因之化到 $V = V_\lambda$ 的情形, 其中 $\lambda \in \mathbb{R}$. 引理 10.6.5 确保 $|\lambda| \le 2$.

将 $T + T^{-1} = \lambda \cdot \mathrm{id}_V$ 整理为 $T^2 - \lambda T + \mathrm{id}_V = 0$. 若 $\lambda = \pm 2$ 则 $(T \mp \mathrm{id}_V)^2 = 0$, 配合引理 10.6.2 可得 $T = \pm \mathrm{id}_V$. 这相当于断言中的 ± 1 分块.

设 $\lambda \neq \pm 2$, 则从 $|\lambda| < 2$ 知 $X^2 - \lambda X + 1$ 无实根, 不可约, 于是它必然是 T 的极小多项式. 特别地, T 无实特征值. 这就导致

$$\forall v \in V,\ v \neq 0 \implies v, Tv \text{ 线性无关}.$$

选定 $v \in V \smallsetminus \{0\}$, 定义子空间 $W := \langle v, Tv\rangle$, 则 T 满足的二次方程说明 W 是 T-不变的. 作直和分解 $V = W \oplus W^\perp$; 注意到 $W^\perp$ 在 $T^* = T^{-1}$ 作用下也不变 (引理 9.5.1), 既然 $T = (T^{-1})^{-1}$ 是 T^{-1} 的多项式, 故 $W^\perp$ 是 $n-2$ 维 T-不变子空间. 以 W 代 V 将原问题进一步化约到 $\dim V = 2$, $T^2 - \lambda T + \mathrm{id}_V = 0$ 且 $|\lambda| < 2$ 的情形.

回忆到 $X^2 - \lambda X + 1 \in \mathbb{R}[X]$ 此时不可约; 因为 $\dim V = 2$, 它同时是 T 的极小多项式和特征多项式. 特别地, $\det T = 1$. 我们因此回到之前介绍的旋转矩阵, 它便是断言中的 2×2 分块 $\boldsymbol{R}(\theta)$; 注意到此处不可能有 $\theta \in \pi\mathbb{Z}$, 否则 T 将有特征值 ± 1.

结合上述讨论, 便对一般的 T 得到所求的分块表达式. 适当换基可将任意分块 $\boldsymbol{R}(\theta_i)$ 化为 $\boldsymbol{R}(-\theta_i)$, 这是引理 10.6.4 确保的.

最后探讨资料 (a,b,m) 和 $(\pm\theta_i \mod 2\pi)_i$ 的唯一性. 回忆到 T 在 $\mathbb{C}$ 上可对角化; 分块表达式中 a 和 b 分别对应到特征值 1 和 -1 的重数, 而每个 $\boldsymbol{R}(\theta_i)$ 在 $\mathbb{C}$ 上都贡献一对复共轭的特征值 $e^{\pm \mathrm{i}\theta_i} \notin \{\pm 1\}$, 这又对应到 $\mathrm{Char}_T \in \mathbb{R}[X]$ 的二次不可约因式. 综上, 资料 (a,b,m) 和 $(\pm\theta_i \mod 2\pi)_i$ (计重数, 精确到重排) 确实由 T 唯一地确定. □

注意到 $\det \boldsymbol{R}(\theta) = 1$, 因此分块表达式中的 b 按 $\det T = (-1)^b$ 确定了正交变换的行列式.

例 10.6.7 (三维空间中的旋转) 考虑配备标准内积的 $\mathbb{R}^3$, 以及行列式为 1 的正交变换 $T:\mathbb{R}^3\to\mathbb{R}^3$. 从定理 10.6.6 容易推得存在 V 的单位正交基 v_1,v_2,v_3, 使得 T 在此基下表作矩阵

$$\begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}, \quad 0\le\theta<2\pi.$$

注意到这里不排除 $\theta\in\{0,\pi\}$, 它们分别对应到 $\boldsymbol{R}(0)=\mathbf{1}_{2\times2}$ 和 $\boldsymbol{R}(\pi)=-\mathbf{1}_{2\times2}$.

当 $\theta\neq0$ 时, 这般的 T 总是一个旋转, 以 $\langle v_1\rangle$ 为轴, 以 θ 为相对于 v_1 方向的转角. 另一方面, 对任意过原点的平面作镜射也是正交变换, 其行列式为 -1; 如果 T 是行列式 -1 的正交变换, 则 T 右合成一个镜射便是行列式 1 的正交变换, 即旋转. 综上,

$$\mathbb{R}^3 \text{ 上的正交变换} = \begin{cases} \text{旋转}, & \det=1, \\ (\text{旋转}\circ\text{镜射}), & \det=-1. \end{cases}$$

对于维数 n 任意高的情形, 定理 10.6.6 也给出类似的刻画, 此时的正交变换将涉及多轴旋转, 以及对 $n-1$ 维子空间作镜射.

正交变换有助于理解不同的单位正交基如何联系. 我们引进一则常用术语.

定义 10.6.8 设 $(V,(\cdot|\cdot))$ 为 n 维内积空间. 若 $v_1,\dots,v_n$ 构成 V 的单位正交基, 则称 $(v_1,\dots,v_n)\in V^n$ 为 V 的**正交标架**.

显然地, 正交变换 $T\in\mathrm{End}(V)$ 映正交标架为正交标架.

另一方面, 给定正交标架 $(v_1,\dots,v_n)$ 和 $(v_1',\dots,v_n')$, 分别简记为 $\mathbf{v}$ 和 $\mathbf{v}'$. 练习 9.4.6 表明 $\mathbf{v}'$ 是从 $\mathbf{v}$ 作一个唯一的正交变换 T 得到的; 在有序基 $\mathbf{v}$ 之下对应于 T 的矩阵正是转换矩阵 $\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}\in\mathrm{M}_{n\times n}(\mathbb{R})$.

综上, 我们对选定的正交标架 $\mathbf{v}$ 得到双射

$$\begin{aligned} \{T\in\mathrm{End}(V):\text{正交变换}\} &\xrightarrow{1:1} \{\mathbf{v}':\text{正交标架}\} \\ T &\mapsto (Tv_1,\dots,Tv_n). \end{aligned} \tag{10.6.1}$$

以下不过是练习 5.4.11 在正交标架情形的复述.

定义 10.6.9 设 $\mathbf{v}$ 和 $\mathbf{v}'$ 为 V 的正交标架. 如果 $\det\boldsymbol{P}_{\mathbf{v}'}^{\mathbf{v}}=1$, 则称 $\mathbf{v}$ 和 $\mathbf{v}'$ 同定向.

练习 10.6.10 试从定理 10.6.6 对正交变换的描述说明上述定义符合几何直观.

10.7　三维空间中的旋转与 Euler 角

在 §10.6 确定了三维实内积空间的正交变换, 它们或者是对某个轴的旋转, 或者是旋转与镜射的合成; 两者分别对应 $\det = 1$ 和 $\det = -1$ 的情形. 本节关注的是旋转. 确定一个非恒等的旋转相当于确定它的

- ⋆ 带方向的转轴, 这相当于过原点的一条射线, 或者说对应于单位球面上的点;
- ⋆ 转角, 相当于一个实数 $\bmod 2\pi$ 的等价类.

因此, 初步的直观表明旋转应当仅由 3 个参数决定. 我们希望更直接地用参数写下对应的矩阵; 进一步, 我们还希望尽可能地用参数来描述旋转的合成.

为了对旋转得到明确的参数化, 行将介绍的第一种方法是经典的, 涉及所谓的 **Euler 角**. 考虑带有标准内积的 $\mathbb{R}^3$.

约定 10.7.1　若 $\mathbb{R}^3$ 的正交标架 $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ 与 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 同定向, 则称为正向正交标架, 否则称为负向.

引理 10.7.2　基于以上符号, 我们有双射

$$\begin{aligned} \{\text{旋转 } T : \mathbb{R}^3 \to \mathbb{R}^3\} &\xrightarrow{1:1} \{\text{正向正交标架 } (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3) \in \mathbb{R}^3\} \\ T &\longmapsto (T\boldsymbol{e}_1, T\boldsymbol{e}_2, T\boldsymbol{e}_3). \end{aligned}$$

证明　所求的映射是上一节末尾的双射 (10.6.1) 之限制 (取 $\mathbf{v}$ 为标准正交标架), 该处的讨论还说明左侧的 T 为旋转等价于右侧的 $\mathbf{v}'$ 为正向. □

旋转的参数化因而翻译为正向正交标架的参数化.

今后谈论转轴时皆记入方向, 转角皆逆时针计算; 对于所有单位向量 $\boldsymbol{u} \in \mathbb{R}^3$, 记以 $\boldsymbol{u}$ 为转轴, 转角为 θ 的旋转所对应之矩阵为 $\boldsymbol{R}_{\boldsymbol{u}}(\theta) \in \mathrm{M}_{3\times 3}(\mathbb{R})$. 譬如

$$\boldsymbol{R}_{\boldsymbol{e}_1}(\theta) = \begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix},$$

$$\boldsymbol{R}_{\boldsymbol{e}_2}(\theta) = \begin{pmatrix} \cos\theta & & \sin\theta \\ & 1 & \\ -\sin\theta & & \cos\theta \end{pmatrix},$$

$$\boldsymbol{R}_{\boldsymbol{e}_3}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & \\ \sin\theta & \cos\theta & \\ & & 1 \end{pmatrix}.$$

练习 10.7.3　说明上式正负号有何道理.　〈提示〉探讨定向.

为了明确地表出从 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$ 过渡到任意指定的正交标架 $(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3)$ 的旋转 T, 首先定义单位向量 $\boldsymbol{f}_2$ 如下.

- ⋆ 若 $\boldsymbol{e}_3=\pm\boldsymbol{u}_3$, 命 $\boldsymbol{f}_2:=\boldsymbol{e}_2$.

- ⋆ 若 $\boldsymbol{e}_3$ 和 $\boldsymbol{u}_3$ 线性无关, 定义 $\boldsymbol{f}_2$ 为与 $\boldsymbol{e}_3$ 和 $\boldsymbol{u}_3$ 皆正交, 而且使正交标架 $(\boldsymbol{e}_3,\boldsymbol{u}_3,\boldsymbol{f}_2)$ 为正向的唯一单位向量; 以三维向量的叉积运算 $\times$ 表示即为

$$\boldsymbol{f}_2:=\boldsymbol{e}_3\times\boldsymbol{u}_3.$$

  此时直线 $\mathbb{R}\boldsymbol{f}_2$ 也称为旋转 T 的**节线**, 其中 T 是映 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$ 为 $(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3)$ 的唯一旋转. 请读者迅速验证节线等于平面 $\langle\boldsymbol{e}_1,\boldsymbol{e}_2\rangle$ 和 $\langle T\boldsymbol{e}_1,T\boldsymbol{e}_2\rangle$ 之交.

接着分三步操作.

顺序	转轴	目的	转角	映法
第一步	$\boldsymbol{e}_3$	将 $\boldsymbol{e}_2$ 旋转至 $\boldsymbol{f}_2$	ψ	$\boldsymbol{R}_{\boldsymbol{e}_3}(\psi):(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\mapsto(\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{e}_3)$
第二步	$\boldsymbol{f}_2$	将 $\boldsymbol{e}_3$ 旋转至 $\boldsymbol{u}_3$	θ	$\boldsymbol{R}_{\boldsymbol{f}_2}(\theta):(\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{e}_3)\to(\boldsymbol{g}_1,\boldsymbol{f}_2,\boldsymbol{u}_3)$
第三步	$\boldsymbol{u}_3$	将 $\boldsymbol{g}_1$ 旋转至 $\boldsymbol{u}_1$	φ	$\boldsymbol{R}_{\boldsymbol{u}_3}(\varphi):(\boldsymbol{g}_1,\boldsymbol{f}_2,\boldsymbol{u}_3)\to(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3)$

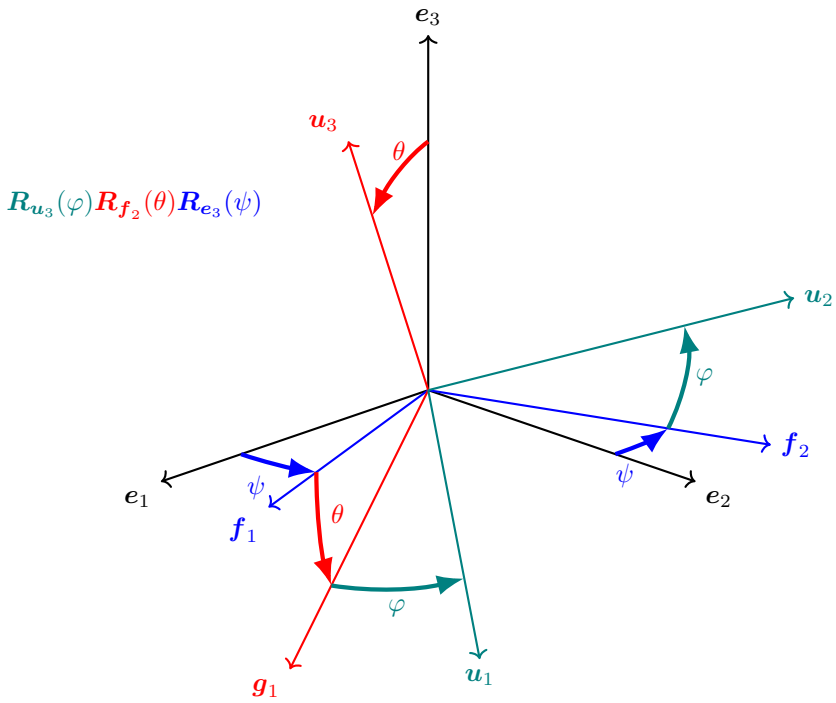
由于 $\boldsymbol{f}_2$ 总正交于 $\boldsymbol{e}_3$ 和 $\boldsymbol{u}_3$, 前两步旋转确实可行, 而中间产物 $\boldsymbol{f}_1$ 和 $\boldsymbol{g}_1$ 也是唯一确定的. 第三步的产物应当先写作 $(\boldsymbol{u}_1,\boldsymbol{g}_2,\boldsymbol{u}_3)$ 的形式, 其中 $\boldsymbol{g}_2$ 待定; 然而 $(\boldsymbol{u}_1,\boldsymbol{g}_2,\boldsymbol{u}_3)$ 是正向正交标架, 故唯一可能是 $\boldsymbol{g}_2=\boldsymbol{u}_2$. 结论如下.

定理 10.7.4 记 T 为映 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$ 为正交标架 $(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3)$ 的唯一旋转. 上述构造说明 T 对应到矩阵

$$\boldsymbol{R}_{\boldsymbol{u}_3}(\varphi)\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi);$$

我们称旋转 T 是由 **Euler 角** (φ,θ,ψ) 描述的.

三个旋转的合成示意如下, 图片原作者为 Dorian Depriester.



Euler 角提供了将旋转参数化的一种方法. 由于三步分别是以正交标架 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, $(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{e}_3)$ 和 $(\boldsymbol{g}_1, \boldsymbol{f}_2, \boldsymbol{u}_3)$ 的第三个, 第二个和第三个向量为轴的旋转, 而它们表征了正交标架在旋转中途的状态, 故定理 10.7.4 的分解可谓动态的, 但我们也可以改从静态的正交标架 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 观照, 答案同样简单.

引理 10.7.5　设 $\boldsymbol{P} : \mathbb{R}^3 \to \mathbb{R}^3$ 为正交变换, $\epsilon := \det \boldsymbol{P}$, 而 $\boldsymbol{R}_{\boldsymbol{u}}(\theta)$ 为给定的旋转, 则 $\boldsymbol{R}_{\boldsymbol{Pu}}(\epsilon\theta) = \boldsymbol{P}\boldsymbol{R}_{\boldsymbol{u}}(\theta)\boldsymbol{P}^{-1}$.

证明　扩展 $\boldsymbol{u}$ 为正向正交标架 $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$, 则 $(\boldsymbol{Pu}, \boldsymbol{Pv}, \epsilon\boldsymbol{Pw})$ 也是正交标架, 而且和 $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ 同定向. 请读者由此验证 $\boldsymbol{R}_{\boldsymbol{Pu}}(\epsilon\theta)\boldsymbol{P}$ 的映法是

$$\boldsymbol{u} \mapsto \boldsymbol{Pu}, \quad \boldsymbol{v} \mapsto \cos(\epsilon\theta)\boldsymbol{Pv} + \sin(\epsilon\theta)\epsilon\boldsymbol{Pw}, \quad \epsilon\boldsymbol{w} \mapsto -\sin(\epsilon\theta)\boldsymbol{Pv} + \cos(\epsilon\theta)\epsilon\boldsymbol{Pw}.$$

另一方面 $\boldsymbol{P}\boldsymbol{R}_u(\theta)$ 的映法是

$$\boldsymbol{u} \mapsto \boldsymbol{Pu}, \quad \boldsymbol{v} \mapsto \cos\theta\boldsymbol{Pv} + \sin\theta\boldsymbol{Pw}, \quad \boldsymbol{w} \mapsto -\sin\theta\boldsymbol{Pv} + \cos\theta\boldsymbol{Pw}.$$

由于 $\epsilon \in \{\pm 1\}$, 讨论三角函数的奇偶性易见 $\boldsymbol{R}_{\boldsymbol{Pu}}(\epsilon\theta)\boldsymbol{P} = \boldsymbol{P}\boldsymbol{R}_{\boldsymbol{u}}(\theta)$. □

定理 10.7.6 设旋转 T 由 Euler 角 (φ,θ,ψ) 描述, 则 T 对应到矩阵

$$\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi).$$

证明 因为 $\boldsymbol{f}_2=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{e}_2$, 引理 10.7.5 表明 $\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)^{-1}$, 故

$$\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta).$$

同理, $\boldsymbol{u}_3=\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{e}_3=\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{e}_3$ 导致

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{u}_3}(\varphi)&=\left(\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\right)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi)\left(\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\right)^{-1}\\&=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)^{-1}\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)^{-1}.\end{aligned}$$

综上,

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{u}_3}(\varphi)\boldsymbol{R}_{\boldsymbol{f}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)&=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)^{-1}\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)^{-1}\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\\&=\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_2}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi).\end{aligned}$$

代入定理 10.7.4 便完成证明. □

观察转轴和转角顺序, 可见定理 10.7.6 与定理 10.7.4 的分解正好颠倒. 静态与动态旋转的倒序关系不限于三个旋转的合成, 可以推及正交标架的任意一系列旋转, 在此将一般情形留给读者表述.

基于 Euler 角的动态和静态分解各有优点. 就矩阵表达式的观点, 定理 10.7.6 的静态版本直接为旋转 T 给出具体的矩阵表法

$$\begin{pmatrix}\cos\psi & -\sin\psi & \\ \sin\psi & \cos\psi & \\ & & 1\end{pmatrix}\begin{pmatrix}\cos\theta & & \sin\theta \\ & 1 & \\ -\sin\theta & & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\varphi & -\sin\varphi & \\ \sin\varphi & \cos\varphi & \\ & & 1\end{pmatrix}.$$

另一方面, 定理 10.7.4 的动态版本对于飞行器或航天器姿态控制而言则是更自然的选项.

注记 10.7.7 在以上推导中, 坐标轴 $\boldsymbol{e}_2$ 和 $\boldsymbol{e}_3$ 的选取纯属人为. 一些文献中采取了诸如 $\boldsymbol{R}_{\boldsymbol{e}_3}(\psi)\boldsymbol{R}_{\boldsymbol{e}_1}(\theta)\boldsymbol{R}_{\boldsymbol{e}_3}(\varphi)$ 之类的表法, 论证并无不同.

Euler 角尽管具体而直观, 实用中也有其不便. 譬如它依赖于转轴的人为选取, 旋转的合成难以用 Euler 角描述, 难以插补, 此外还有称为万向节锁的现象[4]. 下一节将介绍的四元数是代数上更为简洁的一种方案, 同时又免除了应用中的若干缺陷.

[4] 有可能局部上连续地变动 Euler 角 (φ,θ,ψ), 使对应的旋转不改变. 标架的旋转在工程上经常以称为万向节的装置实现, 这相当于说在特定位置沿特定方向变化参数时, 装置呈现锁定不动的现象. 这为许多应用造成了麻烦. 此现象有拓扑的解释, 但也能以简单计算来阐明. 例如 $(\varphi,0,\psi)$ 对应的旋转只和 $\varphi+\psi$ 相关.

10.8　四元数与旋转

从实数到复数是人类数学史上的一次跨越. 抽象的代数结构一旦进入视野, 寻求比复数更广, 但仍具有四则运算的 “数” 并给出其应用, 便成为自然的追求. W. R. Hamilton 在 1843 年灵光一闪得到了答案: 他给出了包含 $\mathbb{C}$ 作为子环的除环 $\mathbb{H}$, 其元素可以用四个实坐标来展开, 因此 $\mathbb{H}$ 的元素被称为**四元数**.

现在直接给出 $\mathbb{H}$ 的构造.

定义–命题 10.8.1 (W. R. Hamilton) *命 $\mathbb{H}$ 为带有指定的基 $1,\mathrm{i},\mathrm{j},\mathrm{k}$ 的 4 维 $\mathbb{R}$-向量空间, 它由此带有加法运算. 在 $\mathbb{H}$ 上定义以乘法符号表示的 $\mathbb{R}$-双线性映射*

$$\cdot : \mathbb{H}\times\mathbb{H}\to\mathbb{H},\quad (x,y)\mapsto x\cdot y,$$

照例简记 $xy=x\cdot y$, 方法是在基上规定其像为

$$\forall x\in\{1,\mathrm{i},\mathrm{j},\mathrm{k}\},\quad 1\cdot x=x=x\cdot 1,$$
$$\mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=-1,$$
$$\mathrm{ij}=\mathrm{k}=-\mathrm{ji},\quad \mathrm{jk}=\mathrm{i}=-\mathrm{kj},\quad \mathrm{ki}=\mathrm{j}=-\mathrm{ik}.$$

这使 $(\mathbb{H},+,\cdot,0,1)$ 成环, 其中的 0 和 1 分别是向量空间的零元和基中给定的元素.

证明　验证环的定义 3.1.1. 关于加法的条件直接来自向量空间结构, 乘法对加法的分配律来自于乘法 $\mathbb{H}\times\mathbb{H}\to\mathbb{H}$ 的双线性条件. 至于乘法结合律, 根据分配律可将问题简化为验证

$$\forall x,y,z\in\{1,\mathrm{i},\mathrm{j},\mathrm{k}\},\quad x(yz)=(xy)z;$$

具体计算没有本质困难. 最后, 1 作为乘法幺元的性质同样简化为定义包含的

$$\forall x\in\{1,\mathrm{i},\mathrm{j},\mathrm{k}\},\quad 1\cdot x=x=x\cdot 1.$$

明所欲证. □

定义中的乘法性质有冗余: 一旦要求 $\mathrm{ij}=\mathrm{k}=-\mathrm{ji}$, 则涉及 k 的其他等式都是使得结合律成立的必要条件.

今后将把 $a\cdot 1\in\mathbb{H}$ 简写为 a, 其中 $a\in\mathbb{R}$. 我们相应地将 $\mathbb{H}$ 的元素表作

$$q=a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$$

的形式, 其中 $a,b,c,d\in\mathbb{R}$ 是由 $q\in\mathbb{H}$ 唯一确定的.

命题 10.8.2　*暂时地将 $\mathrm{i}\in\mathbb{C}$ 写作 $\mathbf{i}$, 以和 $\mathrm{i}\in\mathbb{H}$ 区别. 考虑映射*

$$\iota:\mathbb{C}\to\mathbb{H}$$
$$a+b\mathbf{i}\mapsto a+b\mathrm{i}.$$

这将 $\mathbb{C}$ 等同于 $\mathbb{H}$ 的子环.

证明 显然 ι 保持加法和 1. 从定义可见 $\iota(xy)=\iota(x)\iota(y)$ 对 $x,y\in\{1,\mathbf{i}\}$ 皆成立. 由于 $\mathbb{C}$ 和 $\mathbb{H}$ 的乘法都是 $\mathbb{R}$-双线性的, 故 ι 也保持乘法. □

类似地, 映 $a+b\mathbf{i}\in\mathbb{C}$ 为 $a+b\mathrm{j}$ 或 $a+b\mathrm{k}$ 同样是环嵌入, 论证无异.

我们今后总将 $\mathbb{C}$ 通过 ι 嵌入为 $\mathbb{H}$ 的子环, 符号 $\mathbf{i}$ 和 i 不必再区分. 相应地, $\mathbb{R}$ 也嵌入为 $\mathbb{H}$ 的子环; 向量空间 $\mathbb{H}$ 的纯量乘法

$$t\cdot(a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k})=ta+(tb)\mathrm{i}+(tc)\mathrm{j}+(td)\mathrm{k}$$

按此等同于环的乘法, 既是左乘也是右乘. 进一步, 请读者检验以下简单事实.

练习 10.8.3 验证 $Z(\mathbb{H}):=\{x\in\mathbb{H}:\forall y\in\mathbb{H},\ xy=yx\}$ 等于 $\mathbb{R}$.

提示〉 因为 $\mathbb{H}$ 的乘法是 $\mathbb{R}$-双线性的, $\supset$ 成立. 对于 $\subset$ 方向, 对 $y\in\{\mathrm{i},\mathrm{j},\mathrm{k}\}$ 检验条件即可.

练习 10.8.4 验证所有 $q\in\mathbb{H}$ 都能唯一地表作 $q=z+\mathrm{j}w$, 其中 $z,w\in\mathbb{C}$; 此外, $\mathrm{j}z=\overline{z}\mathrm{j}$ 对所有 $z\in\mathbb{C}$ 成立.

定义–命题 10.8.5 对 $q=a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}\in\mathbb{H}$ 定义其

▷ **共轭** $\overline{q}:=a-b\mathrm{i}-c\mathrm{j}-d\mathrm{k}$,

▷ **迹** $\mathrm{Tr}(q):=q+\overline{q}$,

▷ **范数** $\mathrm{N}(q):=q\overline{q}$.

这些操作具备以下性质

(i) $\overline{\overline{q}}=q$;

(ii) $\overline{q_1+q_2}=\overline{q_1}+\overline{q_2}$ 和 $\overline{q_1q_2}=\overline{q_2}\,\overline{q_1}$ 对所有 $q_1,q_2\in\mathbb{H}$ 成立;

(iii) $\mathrm{Tr}(q)=2a$, 而且 $\mathrm{Tr}:\mathbb{H}\to\mathbb{R}$ 是 $\mathbb{R}$-线性映射;

(iv) $\mathrm{N}(q)=a^2+b^2+c^2+d^2\in\mathbb{R}$, 而且 $\mathrm{N}(1)=1$ 而 $\mathrm{N}(q_1q_2)=\mathrm{N}(q_1)\,\mathrm{N}(q_2)$.

证明 注意到 $q\mapsto\overline{q}$ 是从 $\mathbb{H}$ 到自身的 $\mathbb{R}$-线性映射. 断言 (i) 自明. 对于 (ii), 唯一须说明的是 $\overline{q_1q_2}=\overline{q_2}\,\overline{q_1}$. 由于两侧都是关于 (q_1,q_2) 的双线性映射, 等式立刻化约到 $q_1,q_2\in\{1,\mathrm{i},\mathrm{j},\mathrm{k}\}$ 的情形直接验证.

断言 (iii) 是明白的.

直接计算断言 (iv) 的等式 $\mathrm{N}(q)=a^2+b^2+c^2+d^2\in\mathbb{R}$ 并不复杂, 请读者自行展开, 而 $\mathrm{N}(1)=1$ 则属自明. 最后, (ii) 蕴涵 $\mathrm{N}(q_1q_2)=q_1q_2\overline{q_2}\,\overline{q_1}$, 又因为 $q_2\overline{q_2}\in\mathbb{R}$ 可以和 $\mathbb{H}$ 的所有元素交换乘法顺序, 故 $\mathrm{N}(q_1q_2)=q_1\overline{q_1}q_2\overline{q_2}=\mathrm{N}(q_1)\,\mathrm{N}(q_2)$. □

顺带留意到 $\mathbb{R}=\{q\in\mathbb{H}:q=\overline{q}\}$, 而 $\mathbb{H}$ 的共轭限制在 $\mathbb{C}$ 上无非是复共轭. 现在可以简单地说明 $\mathbb{H}$ 确实是除环, 方法和 $\mathbb{C}$ 上类似, 都涉及共轭.

定理 10.8.6 环 $\mathbb{H}$ 是除环. 更精确地说, 任何 $q \in \mathbb{H} \smallsetminus \{0\}$ 皆满足 $\mathrm{N}(q) \neq 0$, 而且 $q^{-1} = \mathrm{N}(q)^{-1}\overline{q}$.

证明 设 $q = a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$, 其中 $a, b, c, d \in \mathbb{R}$ 不全为零, 则 $\mathrm{N}(q) = a^2+b^2+c^2+d^2 \neq 0$. 我们有

$$q \cdot \mathrm{N}(q)^{-1}\overline{q} = \mathrm{N}(q)^{-1} q\overline{q} = \mathrm{N}(q)^{-1}\,\mathrm{N}(q) = 1.$$

因此 $q \in \mathbb{H}^\times$. □

综上所述, 我们有一列除环

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H},$$

每一步都拓展了数的概念, 同时又付出代价: 从 $\mathbb{R}$ 到 $\mathbb{C}$ 失去序结构, 从 $\mathbb{C}$ 到 $\mathbb{H}$ 则失去乘法交换律. 这穷尽了所有可能的扩展, 因为 Frobenius 的一则定理断言: 若除环 D 包含 $\mathbb{R}$ 作为子环, 而且相对于环乘法给出的 $\mathbb{R}$-向量空间结构, D 有限维而且乘法是双线性映射, 则 D 必然同构于 $\mathbb{R}$, $\mathbb{C}$ 或 $\mathbb{H}$; 证明可参阅 [10, 定理 7.2.9].

四元数环 $\mathbb{H}$ 也可以嵌入为 $\mathrm{M}_{2\times 2}(\mathbb{C})$ 的子环, 这将提供看待 $\mathbb{H}$ 的另一种视角.

命题 10.8.7 从 $\mathbb{H}$ 到 $\mathrm{M}_{2\times 2}(\mathbb{C})$ 的映射

$$\boldsymbol{\Phi} : z + \mathrm{j}w \mapsto \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}, \quad z, w \in \mathbb{C}$$

给出环的嵌入 $\boldsymbol{\Phi} : \mathbb{H} \hookrightarrow \mathrm{M}_{2\times 2}(\mathbb{C})$.

证明 论证基于简单的练习 10.8.4. 由于所有 $q \in \mathbb{H}$ 都有唯一表法 $q = z + \mathrm{j}w$, 映射 $\boldsymbol{\Phi}$ 是单的, 并且映 1 为 $\mathbf{1}_{2\times 2}$. 加性 $\boldsymbol{\Phi}(q_1 + q_2) = \boldsymbol{\Phi}(q_1) + \boldsymbol{\Phi}(q_2)$ 自明. 最后, 为了验证 $\boldsymbol{\Phi}(q_1q_2) = \boldsymbol{\Phi}(q_1)\boldsymbol{\Phi}(q_2)$, 只需运用 $\mathrm{j}z = \overline{z}\mathrm{j}$ 和 $\mathrm{j}^2 = -1$ 作按部就班的验证. □

容易看出 $\mathrm{N}(q) = \det \boldsymbol{\Phi}(q)$, $\mathrm{Tr}(q) = \mathrm{Tr}\,\boldsymbol{\Phi}(q)$, 而 $q^{-1} = \mathrm{N}(q)^{-1}\overline{q}$ 则对应到矩阵求逆的 Cramer 法则.

将焦点转回三维空间中的旋转. 回忆到 $\mathbb{H}$ 也是 $\mathbb{R}$-向量空间. 将 $\mathbb{R}^3$ 等同于 $\mathbb{H}$ 的子空间 $\mathbb{H}_0 := \mathbb{R}\mathrm{i} \oplus \mathbb{R}\mathrm{j} \oplus \mathbb{R}\mathrm{j}$, 其元素又称为纯四元数; 更为内在的, 不直接涉及基的描述则是

$$\begin{aligned} \mathbb{H}_0 &= \{q \in \mathbb{H} : \mathrm{Tr}(q) = 0\} \\ &= \{q \in \mathbb{H} : \overline{q} = -q\}. \end{aligned} \tag{10.8.1}$$

相应地, $\mathbb{R}^3$ 的标准内积 $(\cdot|\cdot)$ 搬运到 $\mathbb{H}_0$ 上, 对应的正定二次型也有内在的描述:

$$\|q\|^2 = (q|q) = \mathrm{N}(q), \quad q \in \mathbb{H}_0.$$

引理 10.8.8 设 $x \in \mathbb{H}^\times = \mathbb{H} \smallsetminus \{0\}$.

(i) 对所有 $q \in \mathbb{H}$ 皆有 $\mathrm{N}(xqx^{-1}) = \mathrm{N}(q)$;

(ii) $q \mapsto xqx^{-1}$ 给出内积空间 $\mathbb{H}_0$ 的自同构, 记为 R_x;

(iii) 承上, $\det R_x = 1$.

证明 我们需要定义–命题 10.8.5 列举的性质以及以下观察: $\mathrm{N}(x^{-1})\,\mathrm{N}(x) = \mathrm{N}(1) = 1$, 故 $\mathrm{N}(x^{-1}) = \mathrm{N}(x)^{-1}$; 类似观察给出 $\overline{x^{-1}} = \overline{x}^{-1}$.

断言 (i) 因而是 $\mathrm{N}(xqx^{-1}) = \mathrm{N}(x)\,\mathrm{N}(q)\,\mathrm{N}(x)^{-1} = \mathrm{N}(q)\,\mathrm{N}(x)\,\mathrm{N}(x)^{-1} = \mathrm{N}(q)$ 的结论; 此处只涉及 $\mathbb{R}$ 中的乘法.

对于断言 (ii), 暂且先设 $\mathrm{N}(x) = 1$. 于是 $\overline{x} = x^{-1}$, 配合 $q \in \mathbb{H}_0$ 可得

$$\begin{aligned}\overline{xqx^{-1}} &= \overline{x}^{-1} \cdot \overline{q} \cdot \overline{x} \\ &= -\overline{x}^{-1} \cdot q \cdot \overline{x} \\ &= -xqx^{-1}.\end{aligned}$$

因此 $xqx^{-1} \in \mathbb{H}_0$. 一般的 x 可以写成 $x = ry$, 其中 $r = \sqrt{\mathrm{N}(x)} \in \mathbb{R}_{>0}$ 而 $\mathrm{N}(y) = 1$. 因为 $xqx^{-1} = ryqy^{-1}r^{-1} = yqy^{-1}$, 对应的等式化到 y 的情形.

于是 $q \mapsto xqx^{-1}$ 确实是 $\mathbb{H}_0$ 到自身的映射, 它的 $\mathbb{R}$-线性性质也是明白的. 断言 (i) 说明它保距, 故为有限维内积空间的同构.

最后以连续性论证处理 (iii). 如将 $\mathbb{H}^\times$ 等同于 $\mathbb{R}^4 \smallsetminus \{0\}$, 则 $R_x \in \mathrm{End}(\mathbb{H}_0)$ 的每个矩阵元都是关于 x 的连续函数, 故 $\det R_x \in \{\pm 1\}$ 也对 x 连续; 另一方面, $\mathbb{R}^4 \smallsetminus \{0\}$ 连通而 $\{\pm 1\}$ 离散, 故 $\det R_x$ 实际是常值函数. 于是对任一点 x 验证等式 $\det R_x = 1$ 即可. 最简单的取法是 $x = 1$, 此时 $R_x = \mathrm{id}$. □

上述结果表明每个 $x \in \mathbb{H}^\times$ 都给出三维空间的旋转 R_x. 易见

$$R_1 = \mathrm{id}, \quad R_{x_1 x_2} = R_{x_1} R_{x_2},$$

因此遂有 $R_{x^{-1}} = R_x^{-1}$. 是否所有旋转都能表作 R_x 的形式? 答案是肯定的. 由于伸缩 x 不改变 R_x, 不妨就聚焦于 $\mathrm{N}(x) = 1$ 的情形.

定理 10.8.9 *设 T 为三维空间 $\mathbb{H}_0$ 中的旋转, 则存在 $x \in \mathbb{H}^\times$ 使得 $\mathrm{N}(x) = 1$ 而 $T = R_x$; 这般的 x 精确到乘以 ± 1 是唯一的.*

证明 首先处理存在性. 以有序基 $\mathrm{i}, \mathrm{j}, \mathrm{k}$ 将 $\mathbb{H}_0$ 等同于 $\mathbb{R}^3$. 由于旋转 T 具有如定理 10.7.6 的静态 Euler 角表法, 不妨就假定 T 是相对于三个坐标轴中任一者的旋转.

取 $x = \cos\theta + \sin\theta\,\mathrm{i}$, 于是 $x^{-1} = \cos\theta - \sin\theta\,\mathrm{i}$ 而

$$\begin{aligned}x\mathrm{j}x^{-1} &= (\cos\theta\,\mathrm{j} + \sin\theta\,\mathrm{k})(\cos\theta - \sin\theta\,\mathrm{i}) \\ &= \big((\cos\theta)^2 - (\sin\theta)^2\big)\,\mathrm{j} + (2\cos\theta\sin\theta)\,\mathrm{k} \\ &= \cos 2\theta\,\mathrm{j} + \sin 2\theta\,\mathrm{k}, \\ x\mathrm{k}x^{-1} &= (\cos\theta\,\mathrm{k} - \sin\theta\,\mathrm{j})(\cos\theta - \sin\theta\,\mathrm{i}) \\ &= -(2\cos\theta\sin\theta)\,\mathrm{j} + \big((\cos\theta)^2 - (\sin\theta)^2\big)\,\mathrm{k} \\ &= -\sin 2\theta\,\mathrm{j} + \cos 2\theta\,\mathrm{k}.\end{aligned}$$

此外由于 $\mathbb{C}$ 是 $\mathbb{H}$ 的交换子环, $x,\mathrm{i}\in\mathbb{C}$, 故 $x\mathrm{i}x^{-1}=xx^{-1}\mathrm{i}=\mathrm{i}$. 综上可见 R_x 表作转角为 2θ 的旋转矩阵

$$\begin{pmatrix}1&&\\&\cos 2\theta&-\sin 2\theta\\&\sin 2\theta&\cos 2\theta\end{pmatrix}.$$

对于 $x=\cos\theta+\sin\theta\mathrm{j}$ 和 $x=\cos\theta+\sin\theta\mathrm{k}$ 也有本质上相同的计算, 转角总是两倍. 既然 θ 可任选, 对三个坐标轴的旋转确实都能写作 R_x 之形.

对于唯一性的部分, $R_x=R_y$ 等价于 $R_{xy^{-1}}=\mathrm{id}$, 故问题归结为证 $R_x=\mathrm{id}\implies x=\pm1$. 然而 $\mathrm{i},\mathrm{j},\mathrm{k}\in\mathbb{H}_0$, 故 $R_x=\mathrm{id}$ 蕴涵

$$x\mathrm{i}=\mathrm{i}x,\quad x\mathrm{j}=\mathrm{j}x,\quad x\mathrm{k}=\mathrm{k}x,$$

因而也蕴涵 $x\in\mathbb{R}$, 故由 $\mathrm{N}(x)=1$ 得 $x=\pm1$. 明所欲证. □

我们继续将 $\mathbb{H}_0$ 等同于 $\mathbb{R}^3$, 按照 §10.7 的方式来谈论旋转的转轴和转角等概念, 只是以下改用线性映射的记法而非矩阵. 定理 10.8.9 的陈述可按此进一步细化.

推论 10.8.10 *设 $u\in\mathbb{H}_0$ 满足 $\mathrm{N}(u)=1$. 以 u 为转轴 (计方向), 转角为 θ 的旋转可以实现为 $R_x:q\mapsto xqx^{-1}$, 其中*

$$x=\cos\left(\frac{\theta}{2}\right)+\sin\left(\frac{\theta}{2}\right)u.$$

证明 易见 $\mathrm{N}(x)=\cos\left(\frac{\theta}{2}\right)^2+\sin\left(\frac{\theta}{2}\right)^2\mathrm{N}(u)=1$. 存在三维空间的旋转 P 使得 $P(\mathrm{i})=u$. 基于 §10.7 对旋转的记法, 引理 10.7.5 蕴涵

$$R_u(\theta)=PR_{\mathrm{i}}(\theta)P^{-1}.$$

定理 10.8.9 给出 $y\in\mathbb{H}^\times$ 使得 $R_y=P$; 此外, 其证明还说明

$$R_{\mathrm{i}}(\theta)=R_{\cos\left(\frac{\theta}{2}\right)+\sin\left(\frac{\theta}{2}\right)\mathrm{i}}.$$

所论的旋转 $R_u(\theta)$ 遂等于 R_x, 其中

$$\begin{aligned}x&=y\left(\cos\left(\frac{\theta}{2}\right)+\sin\left(\frac{\theta}{2}\right)\mathrm{i}\right)y^{-1}\\&=\cos\left(\frac{\theta}{2}\right)+\sin\left(\frac{\theta}{2}\right)y\mathrm{i}y^{-1},\end{aligned}$$

然而 $y\mathrm{i}y^{-1}=R_y(\mathrm{i})=P(\mathrm{i})=u$. 证毕. □

练习 10.8.11 (范数 1 的四元数与酉矩阵) 基于命题 10.8.7 对四元数的矩阵诠释, 满足 $\mathrm{N}(x)=1$ 的 $x\in\mathbb{H}$ 对应到

$$\boldsymbol{\Phi}(x)=\begin{pmatrix}z&-\overline{w}\\w&\overline{z}\end{pmatrix},\quad |z|^2+|w|^2=1.$$

说明这种矩阵正是满足 $\det \boldsymbol{A} = 1$ 的酉矩阵 $\boldsymbol{A} \in \mathrm{M}_{2\times 2}(\mathbb{C})$.

提示 记 $\boldsymbol{A}$ 的第一列为 (z, w), 则 $|z|^2 + |w|^2 = 1$, 而第二列是与第一列正交的单位向量, 由 $\det \boldsymbol{A}$ 唯一确定.

习题

1. (有限维情形的 Bessel 不等式) 设 $(V, (\cdot|\cdot))$ 为复内积空间 (或内积空间), $v_1, \dots, v_k$ 是其中的一族单位正交向量 ($k \in \mathbb{Z}_{\geq 1}$). 对所有 $v \in V$, 证明

$$\sum_{i=1}^{k} |(v_i|v)|^2 \leq \|v\|^2,$$

等号成立当且仅当 $v \in \langle v_1, \dots, v_k\rangle$, 而且此时 $v = \sum_{i=1}^{k} (v_i|v) v_i$.

2. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$ 满足 ${}^{\dagger}\boldsymbol{A} = -\boldsymbol{A}$. 简记 $\mathbf{1} := \mathbf{1}_{n\times n}$. 说明 $\mathbf{1} \pm \boldsymbol{A}$ 皆可逆, 而且 $(\mathbf{1} - \boldsymbol{A})(\mathbf{1} + \boldsymbol{A})^{-1}$ 是酉矩阵.

3. (酉上三角化) 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$. 证明存在酉矩阵 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(\mathbb{C})$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为上三角矩阵.

提示 上三角化定理 7.3.5 给出的 $\boldsymbol{P}$ 未必是酉矩阵; 对其列向量作 Gram–Schmidt 正交化.

4. 对标准复内积空间 $\mathbb{C}^n$ 表述并证明推论 9.4.7 和练习 9.4.8 的相应版本.

5. 对 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(\mathbb{C})$ 验证 $\mathrm{Tr}\left({}^{\dagger}\boldsymbol{A}\,\boldsymbol{A}\right) = \sum_{i,j} |a_{ij}|^2$, 而且 $(\boldsymbol{A}, \boldsymbol{B}) \mapsto \mathrm{Tr}\left({}^{\dagger}\boldsymbol{A}\,\boldsymbol{B}\right)$ 使 $\mathrm{M}_{m\times n}(\mathbb{C})$ 成为复内积空间; 请对有限维复内积空间之间的线性映射 $T \in \mathrm{Hom}(V, W)$ 陈述相应的版本. 这种内积又称有限维情形的 **Hilbert–Schmidt 内积**, 记为 $(\cdot|\cdot)_{\mathrm{HS}}$.

6. 设 V 和 W 为有限维复内积空间, $T \in \mathrm{Hom}(V, W)$. 定义

$$\|T\| := \max_{v:\|v\|_V = 1} \|Tv\|_W,$$

称为 T 的**算子范数**. 若取单位正交基将 V 和 W 等同于标准复内积空间, 则可见 $\{v : \|v\|_V = 1\}$ 是紧集而 $v \mapsto \|Tv\|_W$ 是连续函数, 所以极值确实被取到.

(i) 说明 $\|tT\| = |t| \cdot \|T\|$ (其中 $t \in \mathbb{C}$), $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$ 以及 $\|T\| = 0 \iff T = 0$.

(ii) 说明若 $V \xrightarrow{T} W \xrightarrow{S} U$ 为有限维复内积空间之间的线性映射, 则 $\|ST\| \leq \|S\| \cdot \|T\|$; 此外, $\|\mathrm{id}_V\| = 1$.

(iii) 用上一道习题介绍的 $(\cdot|\cdot)_{\mathrm{HS}}$ 定义 Hilbert–Schmidt 范数 $\|T\|_{\mathrm{HS}} := \sqrt{(T|T)_{\mathrm{HS}}} = \sqrt{\mathrm{Tr}\,(T^* T)}$. 说明对所有 $T \in \mathrm{Hom}(V, W)$ 皆有

$$\frac{1}{\sqrt{\dim V}} \|T\|_{\mathrm{HS}} \leq \|T\| \leq \|T\|_{\mathrm{HS}}.$$

提示 我们有 $\|T\|^2 = \max_{\|v\|_V=1} (Tv|Tv)_W = \max_{\|v\|_V=1} (T^* Tv|v)_V$; 以命题 9.10.1 的复版本确定极大值, 用特征值表达.

(iv) 对所有 $T, T' \in \mathrm{Hom}(V, W)$ 定义

$$d_1(T, T') := \|T - T'\|, \quad d_2(T, T') := \|T - T'\|_{\mathrm{HS}},$$

说明两者都使 $\mathrm{Hom}(V, W)$ 成为度量空间, 而且 $\mathrm{Hom}(V, W)$ 的点列相对于 d_1 和 d_2 有相互等价的收敛性和极限概念. 说明 Cauchy 点列的概念对两种度量也是等价的.

(v) 承上, 说明若取 $V = \mathbb{C}^n$ 和 $W = \mathbb{C}^m$ 为标准复内积空间, 则 d_2 对应到 $\mathrm{M}_{m\times n}(\mathbb{C}) = \mathbb{C}^{mn}$ 上的标准度量结构. 由此易见 $\mathrm{Hom}(V, W)$ 完备, 亦即所有 Cauchy 点列皆收敛.

7. 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$. 说明 $\rho(\boldsymbol{A}) \leq \left\|\boldsymbol{A}^k\right\|^{1/k}$ 对所有 $k \in \mathbb{Z}_{\geq 1}$ 成立, $\rho(\boldsymbol{A})$ 是谱半径, $\|\boldsymbol{A}\|$ 是先前习题定义的算子范数 (相对于 $\mathbb{C}^n$ 的标准复内积).

8. (矩阵指数映射) 设 V 为有限维复内积空间.

(i) 对所有 $T \in \mathrm{End}(V)$, 说明 $e^T := \sum_{k=0}^{\infty} \frac{T^k}{k!}$ 在 $\mathrm{End}(V)$ 中收敛; 此处以上一道习题引入的 d_1 (或等价的 d_2) 赋予 $\mathrm{End}(V)$ 度量空间结构.

提示〉 用 d_1 来考虑, 并且用 $\|T^n\| \leq \|T\|^n$ 将问题化约到 $\sum_k \frac{\|T\|^k}{k!}$ 在 $\mathbb{R}$ 中的收敛性.

(ii) 验证若 $T_1 T_2 = T_2 T_1$ 则 $e^{T_1+T_2} = e^{T_1} e^{T_2}$. 由此说明 $e^{-T} = (e^T)^{-1}$.

(iii) 设 W 为有限维复内积空间, $S : V \xrightarrow{\sim} W$ 为复向量空间的同构, 验证 $e^{STS^{-1}} = Se^T S^{-1}$.

(iv) 证明 $\det(e^T) = e^{\mathrm{Tr}(T)}$. 提示〉 上三角化.

(v) 取 $V = \mathbb{C}^n$ 为标准复内积空间, 按此便能够定义关于矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{C})$ 的指数函数 $e^{\boldsymbol{A}}$. 设 $t \in \mathbb{R}$. 说明 $e^{t\boldsymbol{A}}$ 的各个矩阵元都对 t 可微, 并且对各个矩阵元求导给出

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{t\boldsymbol{A}} = \boldsymbol{A} e^{t\boldsymbol{A}}.$$

指数映射的构造对有限维实内积空间当然也成立; 若 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 则 $e^{\boldsymbol{A}} \in \mathrm{M}_{n\times n}(\mathbb{R})$.

9. 设 V 为实 (或复) 内积空间, 对可逆线性映射 $T \in \mathrm{End}(V)$ 作极分解 $T = RU$, 其中 R 是对应到正定双线性 (或半双线性) 形式的自伴算子, U 是正交 (或酉) 算子. 证明 T 是正规算子当且仅当 $RU = UR$.

提示〉 可能需要以下辅助: 论证 R 可以表作 R^2 的多项式.

10. 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(\mathbb{R})$.

(i) 设 $\boldsymbol{S} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 可逆而且有极分解 $\boldsymbol{S} = \boldsymbol{R}\boldsymbol{V}$, 其中 $\boldsymbol{R}$ 对称正定, $\boldsymbol{V}$ 正交, 使得 $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{B}$ 而 ${}^{\mathrm{t}}\boldsymbol{S}\,\boldsymbol{A} = \boldsymbol{B}\,{}^{\mathrm{t}}\boldsymbol{S}$. 证明 $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{B}$.

提示〉 证 $\boldsymbol{A}\boldsymbol{R}^2 = \boldsymbol{R}^2\boldsymbol{A}$ 并回忆到 $\boldsymbol{R}$ 是 $\boldsymbol{R}^2 = \boldsymbol{S}\,{}^{\mathrm{t}}\boldsymbol{S}$ 的多项式.

(ii) 说明若存在酉矩阵 $\boldsymbol{U} \in \mathrm{M}_{n\times n}(\mathbb{C})$ 使得 $\boldsymbol{B} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$, 则存在正交矩阵 $\boldsymbol{V} \in \mathrm{M}_{n\times n}(\mathbb{R})$ 使得 $\boldsymbol{B} = \boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}$.

提示〉 首先推导 $\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{B}$ 和 ${}^{\mathrm{t}}\boldsymbol{U}\,\boldsymbol{A} = \boldsymbol{B}\,{}^{\mathrm{t}}\boldsymbol{U}$. 将 $\boldsymbol{U}$ 按实部和虚部分解为 $\boldsymbol{U} = \boldsymbol{C} + \mathrm{i}\boldsymbol{D}$. 对 $s \in \mathbb{R}$ 命 $\boldsymbol{S} = \boldsymbol{C} + s\boldsymbol{D} \in \mathrm{M}_{n\times n}(\mathbb{R})$, 说明除了有限个 s 它总是可逆, 代入 (i) 以得出 $\boldsymbol{V}$.

11. 设 $(V,(\cdot|\cdot))$ 为有限维复内积空间, $A,N\in\mathrm{End}(V)$ 而 N 为正规算子. 证明若 $AN=NA$ 则 $A\,N^*=N^*\,A$.

提示〉 先将问题归结为证 $B:=A\,N^*-N^*\,A$ 满足 $\mathrm{Tr}(B\,B^*)=0$. 应用条件来论证

$$\begin{aligned}B\,B^*=(A\,N^*-N^*\,A)(N\,A^*-A^*\,N)&=(A\,N^*-N^*\,A)N\,A^*-(A\,N^*-N^*\,A)\,A^*\,N\\&=N\,((A\,N^*-N^*\,A)A^*)-((A\,N^*-N^*\,A)A^*)\,N.\end{aligned}$$

12. 承上题, 说明存在 $f\in\mathbb{C}[X]$ 使得 $N^*=f(N)$.

提示〉 选取单位正交基将 N 化为对角矩阵, 所求断言此时可以直接验证, 或者应用上题结果连同第四章关于对易关系 "$AB=BA$" 的一道习题.

13. 设 V 为实内积空间, $T\in\mathrm{End}(V)$ 为正规算子. 说明存在单位正交基使 T 表作分块对角实矩阵, 其对角分块或者是 1×1 的, 或者形如 $\begin{pmatrix}a&-b\\b&a\end{pmatrix}$; 说明后者作用在 $\mathbb{R}^2$ 上的几何意义是 "旋转再伸缩".

14. 考虑复内积空间 V 和 $T\in\mathrm{End}(V)$.

(i) 说明若 V_0 是 V 的 T-不变子空间, 则 $V_0^\perp$ 是 T^*-不变子空间. 这是引理 9.5.1 的简单类比.

(ii) 仿照定理 9.5.2 的方法, 直接证明若 $T^*=\pm T$, 则 T 可酉对角化; 换言之, 存在由特征向量构成的单位正交基.

(iii) 证明若 $T',T''\in\mathrm{End}(V)$ 皆可酉对角化, 而且 $T'T''=T''T'$, 则它们可以在同一个单位正交基之下对角化.

提示〉 这是定理 7.5.3 的酉版本.

(iv) 由此重新证明正规算子的谱分解定理 10.4.1.

提示〉 用命题 10.1.15 化到 $T^*=\pm T$ 的情形.

15. (平方根的连续性) 设 $(V,(\cdot|\cdot))$ 为有限维复内积空间.

(i) 设 $S,T\in\mathrm{End}(V)$ 皆为正定自伴的. 设 $v\in V$ 是 $\sqrt{S}-\sqrt{T}$ 的特征向量, 特征值为 μ. 证明

$$((S-T)v\mid v)=\mu\left((\sqrt{S}+\sqrt{T})v\,\middle|\,v\right).$$

提示〉 利用自伴性质与 $S-T=(\sqrt{S}-\sqrt{T})\sqrt{S}+\sqrt{T}(\sqrt{S}-\sqrt{T})$.

(ii) 取 $\lambda>0$ 使得 S 和 T 的极小特征值都 $\geq\lambda$. 证明算子范数的不等式 $\|\sqrt{S}-\sqrt{T}\|\leq\frac{\|S-T\|}{2\sqrt{\lambda}}$. 提示〉 以特征值描述自伴算子的算子范数.

(iii) 承上, 说明正定线性映射的平方根映射 $T\mapsto\sqrt{T}$ 是连续的. 此处以 $\mathrm{End}(V)$ 上的算子范数或等价的 Hilbert–Schmidt 范数探讨连续性.

16. (极分解的连续性) 基于上一题的结果, 证明可逆线性映射的极分解 $T=RU$ (定理 10.5.6) 中的 R 和 U 对 T 都是连续的. 类似结论当然也适用于实内积空间.

17. 说明若 $\boldsymbol{A}\in\mathrm{M}_{2\times2}(\mathbb{R})$ 为正交矩阵, $\det\boldsymbol{A}=-1$, 则 $\boldsymbol{A}$ 是相对于 $\mathbb{R}^2$ 中某条过原点直线的镜射; 试由 $\boldsymbol{A}$ 具体地确定此直线.

提示〉 这是定理 10.6.6 的简单应用, 但请尝试从基本定义来证明.

18. 命 $\mathbb{H}_0$ 为纯四元数构成的空间. 对 $q_1, q_2 \in \mathbb{H}_0$ 定义

$$q_1 \cdot q_2 := \frac{-1}{2} \cdot (q_1 q_2 + q_2 q_1), \quad q_1 \times q_2 := \frac{1}{2}(q_1 q_2 - q_2 q_1).$$

说明相对于标准同构 $\mathbb{H}_0 \simeq \mathbb{R}^3$, 它们分别对应到 $\mathbb{R}^3$ 的标准内积 $\cdot$ 和叉积 $\times$.

19. (Rodrigues 旋转公式) 设 $\boldsymbol{u} = (u_1, u_2, u_3)$ 为标准内积空间 $\mathbb{R}^3$ 中的单位向量. 说明以 $\boldsymbol{u}$ 为转轴 (计方向), 转角为 θ 的旋转可以表作矩阵

$$\mathbf{1}_{3\times 3} + (\sin\theta)\boldsymbol{A} + (1-\cos\theta)\boldsymbol{A}^2, \quad \boldsymbol{A} := \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}.$$

20. 记 $\mathrm{SO}(n)$ 为行列式为 1 的 $n \times n$ 正交实矩阵所成集合, 另记

$$\mathfrak{so}(n) := \{\boldsymbol{A} \in \mathrm{M}_{n\times n}(\mathbb{R}) : {}^{\mathrm{t}}\boldsymbol{A} + \boldsymbol{A} = \mathbf{0}_{n\times n}\}.$$

(i) 证明先前习题定义的矩阵指数映射 $\boldsymbol{A} \mapsto e^{\boldsymbol{A}}$ 限制为满射 $\mathfrak{so}(n) \twoheadrightarrow \mathrm{SO}(n)$.

(ii) 试阐明 (i) 与 Rodrigues 旋转公式的关联.

21. 记 $\mathrm{O}(n)$ 为 $\mathbb{R}^n$ 上的所有正交变换所成集合, 它是 $\mathrm{M}_{n\times n}(\mathbb{R})$ 的子集, 因此也具有度量空间的结构. 对于所有 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{O}(n)$, 说明以下陈述等价:

(i) 存在连续映射 $\gamma : [0,1] \to \mathrm{O}(n)$ 使得 $\gamma(0) = \boldsymbol{A}$ 而 $\gamma(1) = \boldsymbol{B}$;

(ii) $\det \boldsymbol{A} = \det \boldsymbol{B}$.

按照数学分析的语言, 这相当于说 $\mathrm{O}(n)$ 有两个连通分支, 由行列式区分.

提示〉 重点是说明若 $\det \boldsymbol{A} = 1$, 则存在 γ 使得 $\gamma(0) = \boldsymbol{A}$ 而 $\gamma(1) = \mathbf{1}_{n\times n}$. 定理 10.6.6 将此化到 $n = 2$ 情形.

22. 接续上题理路, 记 $\mathrm{GL}(n,\mathbb{R})$ 为所有 $n\times n$ 可逆矩阵所成集合. 对于所有 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{GL}(n,\mathbb{R})$, 说明以下陈述等价:

(i) 存在连续映射 $\gamma : [0,1] \to \mathrm{O}(n)$ 使得 $\gamma(0) = \boldsymbol{A}$ 而 $\gamma(1) = \boldsymbol{B}$;

(ii) $\mathrm{sgn}(\det \boldsymbol{A}) = \mathrm{sgn}(\det \boldsymbol{B})$.

试由此说明练习 5.4.11 对定向的定义合理: 两组有序基同定向当且仅当它们可以连续地相互过渡.

提示〉 可应用推论 9.4.7 和练习 9.4.8 化到 $\mathrm{O}(n)$ 的情形.

23. (A. Hurwitz) 定义四元数环 $\mathbb{H}$ 的子集

$$\begin{aligned} O &:= \left\{ \frac{a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}}{2} \in \mathbb{H} : a,b,c,d \in \mathbb{Z} \text{ 同奇或同偶} \right\} \\ &= \mathbb{Z}\left(\frac{1+\mathrm{i}+\mathrm{j}+\mathrm{k}}{2}\right) + \mathbb{Z}\mathrm{i} + \mathbb{Z}\mathrm{j} + \mathbb{Z}\mathrm{k}. \end{aligned}$$

验证 O 是 $\mathbb{H}$ 的子环, 而且 $q \in O \implies \bar{q} \in O$.

24. 说明 $q \in O$ 蕴涵 $\mathrm{N}(q) := q\overline{q} \in \mathbb{Z}$, 而且 $q \in O^\times$ 当且仅当 $\mathrm{N}(q) = 1$.

25. (环 O 的左带余除法) 设 $a, b \in O$ 而 $b \neq 0$. 说明存在 $q \in O$ 使得 $\mathrm{N}(a - qb) < \mathrm{N}(b)$.

提示〉 先处理 $b \in \mathbb{Z}_{\geq 1}$ 的情形. 对于一般情形, 说明存在 $q \in O$ 使得 $\mathrm{N}(a\overline{b} - qb\overline{b}) < \mathrm{N}(b\overline{b})$, 由此推导 $\mathrm{N}(a - qb) < \mathrm{N}(b)$.

26. 设 p 为奇素数. 记 $Op := \{qp : q \in O\}$. 证明存在 $x \in O$ 使得 $x \notin Op$ 而 $p \mid \mathrm{N}(x)$.

提示〉 一道第三章的习题说明存在 $a, b \in \mathbb{Z}$ 使得 $p \mid a^2 + b^2 + 1$.

27. 若子集 $M \subset O$ 非空, 对加法封闭, 而且 $q \in O$ 和 $x \in M$ 蕴涵 $qx \in M$, 则称 M 为 O 的左理想.

(i) 说明所有左理想都形如 Oy, 其中 $y \in O$.

(ii) 设 p 为奇素数, 以先前的结果说明存在左理想 M 使得 $Op \subsetneq M \subsetneq O$.

提示〉 取 $x \notin Op$ 使得 $p \mid \mathrm{N}(x)$. 取 $M := Op + Ox = \{q_1 p + q_2 x : q_1, q_2 \in O\}$, 然后用范数和简单的同余来说明 $1 \notin M$.

(iii) 仍设 p 为奇素数. 说明存在 $r, s \in O$, 它们不是环 O 的可逆元, 而且 $p = rs$.

提示〉 将 (ii) 的 M 表为 Or 的形式.

28. (Lagrange 四平方和定理)

(i) 证明所有奇素数 p 都能表为 $x^2 + y^2 + z^2 + w^2$ 的形式, 其中 $x, y, z, w \in \mathbb{Z}$.

提示〉 取 $r, s \in O \setminus O^\times$ 使得 $p = rs$. 先说明 $\mathrm{N}(r) = \mathrm{N}(s) = p$, 因此存在同奇或同偶的整数 a, b, c, d 使得 $4p = a^2 + b^2 + c^2 + d^2$. 若它们同偶则收工, 否则考虑 $\dfrac{a \pm b}{2}$ 和 $\dfrac{c \pm d}{2}$ 将 $2p$ 表为四平方和, 然后继续讨论奇偶性来表达 p.

(ii) 证明所有非负整数都能表为 $x^2 + y^2 + z^2 + w^2$ 的形式, 其中 $x, y, z, w \in \mathbb{Z}$.

第十一章 群的概念

所谓的群 (定义 11.1.1), 简言之是带有记为乘法的二元运算 $G \times G \to G$ 的非空集 G, 条件是 (a) 结合律 $(xy)z = x(yz)$ 成立, (b) 存在幺元 1_G 或简记为 1, 而且 (c) 每个 $g \in G$ 都有乘法逆元 g^{-1}. 对于先前接触过的种种数学结构, 其间的 "自同构" 在合成运算之下总是成群. 以下是一部分重要例子.

- ⋆ 设 V 为域 F 上的向量空间, 则 V 的全体自同构成群 $\mathrm{GL}(V)$, 称为 V 上的一般线性群; 基于矩阵和线性变换之间的对应, 取 $V = F^n$ 可见 F 上的全体 $n \times n$ 可逆矩阵对乘法也成群, 记为 $\mathrm{GL}(n, F)$.
- ⋆ 标准实内积空间 $\mathbb{R}^n$ 上的正交变换构成 $\mathrm{GL}(n, \mathbb{R})$ 的子群 $\mathrm{O}(n)$, 称为正交群; 其中行列式为 1 的元素构成子群 $\mathrm{SO}(n)$, 称为特殊正交群. 对于 $n = 3$ 的情形, $\mathrm{SO}(3)$ 便是空间中的旋转变换群, 它有清晰的几何意义.
- ⋆ 考虑不带额外结构的非空有限集, 不妨取作 $\{1, \ldots, n\}$, 则它的 "自同构" 是行列式理论中见过的 n 个元素的置换, 对应的群称为对称群或置换群, 记为 $\mathfrak{S}_n$.

因此, 群的见地捕捉了结构的 "对称性", 提供理解许多数学问题的钥匙. 由这些例子可以看出群的二元运算一般并不交换; 在交换的情形, 常将群的运算记为加法.

除了结构的对称性, 代数结构本身往往也搭建在群上, 例如环和向量空间的元素对加法成群, 域 F 的非零元对乘法成群 $F^\times$. 由此观之, 群是比环和向量空间更基础的代数结构; 如将定义进一步放宽, 还能得到幺半群和半群的概念 (定义 11.1.5). 详细的定义和讨论见诸 §11.1.

我们在 §11.2 继续研究保持群结构的映射. 若映射 $f : G \to G'$ 满足恒等式 $f(xy) = f(x)f(y)$, 则称之为群同态; 可以证明同态自动保持幺元和逆元. 可逆的群同态称为群同构. 相互同构的群具有完全相同的群论性质, 可以等量齐观.

最简单的群是 §11.3 探讨的循环群, 无穷循环群同构于整数的加法群 $(\mathbb{Z}, +)$, 有限循环群则同构于 $\bmod n$ 同余类的加法群 $(\mathbb{Z}/n\mathbb{Z}, +)$. 循环群及其子群有完整分类.

群 G 对子群 H 的左或右陪集是群论中另一个基本概念, 见 §11.4, 它们自然地将 G 分解为等价类; 关于陪集个数 $(G : H)$ 的 Lagrange 定理 11.4.6 是算术性质在有限群论中的初步应用. 由之易推得元素 σ 的阶 $\mathrm{ord}(\sigma)$ (定义 11.4.9) 恒整除 $|G|$.

陪集分解是群作用下的轨道分解的特例, 详阅 §§11.5–11.6. 如同之前所见, 群的许多实例都是通过它们在其它结构 (例如向量空间或有限集) 上的作用呈现的. 群作用之

下的轨道 (定义 11.5.3) 等同于稳定化子群的陪集; 轨道分解因而与一些计数问题相联系, 如 (11.5.1) 与命题 11.5.10. 反过来看, 作用也有助于澄清群本身的结构; 给定素数 p, 著名的应用包括断言 p-群中心非平凡的命题 11.6.4, 断言 p 阶元存在的 Cauchy 定理 11.6.5, 以及关于 Sylow p-子群 (定义 11.6.6) 的三个 Sylow 定理. 本书对 Sylow 三定理述而不证, 仅提供配套的习题和参考文献.

为了从实例阐释轨道分解, 考虑置换 $\sigma \in \mathfrak{S}_n$ 生成的子群 $\langle\sigma\rangle$, 我们将在 §11.7 明确 $\{1,\dots,n\}$ 在 $\langle\sigma\rangle$ 作用下的轨道, 从而推导 σ 的循环分解, 这是研究置换的有力工具. 历史上, 置换的早期研究与高次方程理论息息相关. 基于群论语言, §11.8 将粗略介绍称为 Lagrange 预解式的经典技术, 由之能推导不超过四次方程的公式解. 预解式在现代域论中仍有作用.

高次方程的研究是群论初创时期的主要背景, 在其先驱 É. Galois 引入的许多基本概念中, 正规子群是荦荦大者; 虽然这一概念受方程理论启发, 但 §11.9 选择从抽象视角引入正规子群的定义 11.9.1, 然后将它与商群的构造 (定义–命题 11.9.10) 绑定. 这些思路与向量空间 (或环) 对子空间 (或理想) 取商有相通之处. 该节后半部将介绍关于群的几个同态定理.

除了 G 本身和 $\{1\}$ 再无其它正规子群的非平凡群 G 称为单群 (定义 11.9.2); 不难证明交换单群只有素数阶循环群 $\mathbb{Z}/p\mathbb{Z}$, 非交换情形则远为复杂. 本章习题有更多关于单群的例子.

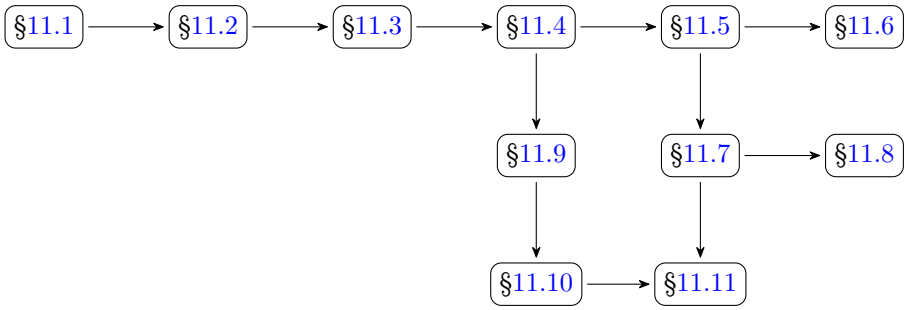
从既有的群 N 和 H 构造新群, 或反向分解大群的一种技术是 §11.10 介绍的半直积 $N \rtimes_\varphi H$, 而直积 $N \times H$ 不过是其特例. 半直积的实例是例 11.10.5 介绍的二面体群 D_{2n}, 它们是平面上正 n 边形的对称群.

再向上一维, 空间中的正多面体及其对称性是 §11.11 的主题. 我们将回顾或勾勒五种正多面体的分类, 然后在定理 11.11.6 确定其旋转对称群. 在此基础上, 定理 11.11.7 将进一步分类 SO(3) 的有限子群: 精确到共轭, 它们是有限循环群, 二面体群, 以及正多面体的三种旋转对称群. 几何与代数在此体现了优美的交融.

阅读提示

本章是基于群论在全书中的角色而设计的, 和其它代数类教材相比不无遗漏: 例如本章对有限群着墨甚少, 而群的许多重要性质也被移入习题. 对此可以参考 [10, 第四章] 或任一本关于抽象代数的教材来补足. 此外, 本章 §11.11 关于正多面体的讨论基于几何直观, 许多论证只作勾勒, 以避免非必要的麻烦; 严谨理论可参考第十四章的相关内容.

阅读顺序



11.1 群的基本定义

从我们提出解方程, 特别是解线性方程组的问题以来, 经过矩阵, 向量空间, 行列式, 内积空间乃至于四元数的一系列移步换影, 种种数学结构的对称性渐次呈现. 对称性的突出例子包括:

- ⋆ 向量空间 V 的自同构;
- ⋆ 集合 X 上的置换;
- ⋆ 实内积空间 $(V,(\cdot|\cdot))$ 的正交变换;
- ⋆ 复内积空间 $(W,(\cdot|\cdot))$ 的酉变换;
- ⋆ 三维空间 $\mathbb{R}^3$ 的旋转.

这些操作尽管面貌各异, 总归都具有乘法与取逆的运算, 并且带有对乘法不起作用的幺元.

在上述例子中, 探讨的总是某个集合到自身的特定双射, 而乘法操作无非是映射作合成, 取逆则是对双射取逆. 另一方面, 我们也遇见不少带有类似操作, 却不与映射直接相关的数学对象, 例如:

- ⋆ 环 R 的可逆元;
- ⋆ 域 F 上的 $n\times n$ 可逆矩阵;
- ⋆ 满足 $\mathrm{N}(q)=1$ 的四元数 q;

对应的运算取作各自的乘法, 仍服从结合律等诸般性质. 尽管这些数学对象有时可以诠释为双射 (例如可逆矩阵作为线性双射), 但是在推导中起作用的经常只是形式化的运算规则, 执着于映射观点反而容易自缚手脚.

在这一切实例中, 共通的核心是二元运算, 幺元和取逆. 一旦熟识这些数学结构, 关于 "群" 的概念便已箭在弦上, 不得不发了.

定义 11.1.1 (群) 所谓的群, 是指资料 $(G,\cdot)$, 其中 G 是非空集, 而 $\cdot: G\times G\to G$ 是满足以下条件的二元运算, 称为群的乘法.

▷ **结合律** $(x\cdot y)\cdot z = x\cdot(y\cdot z)$ 对所有 $x,y,z\in G$ 成立, 故今后乘法可以省略括号.

▷ **幺元** 存在 $1_G\in G$ 使得 $1_G\cdot x = x = x\cdot 1_G$ 对所有 $x\in G$ 成立.

▷ **逆元** 对所有 $x\in G$, 存在 $x^{-1}\in G$ 使得 $x\cdot x^{-1} = 1_G = x^{-1}\cdot x$.

不致混淆时, 我们也将群的资料 $(G,\cdot)$ 简写为 G.

如果群 G 的乘法满足交换律 $xy = yx$, 则称 G 为**交换群**或 **Abel 群**.

二元运算 $\cdot$ 经常称为 G 的乘法, 按 $xy = x\cdot y$ 的方式简记; 1_G 称为 G 的幺元, x^{-1} 则称为元素 x 的逆.

⋆ 上述定义中的幺元 1_G 是唯一的: 设若 $1_G, 1'_G\in G$ 都满足幺元的性质, 则

$$1_G = 1_G 1'_G = 1'_G.$$

不致混淆时, G 的幺元 1_G 常简写为 1.

⋆ 逆元的存在性立刻导致乘法消去律: 对所有 $x,y,z\in G$ 皆有

$$\begin{aligned} xy = xz &\implies y = z,\\ yx = zx &\implies y = z. \end{aligned}$$

以第一式为例, 两边同时左乘以 x 的某个逆元 x^{-1} 并应用结合律可得

$$y = 1_G y = x^{-1}xy = x^{-1}xz = 1_G z = z.$$

对于第二式, 改为右乘以 x^{-1} 即可.

⋆ 每个元素 $x\in G$ 的逆元 x^{-1} 也是唯一的: 设 y 和 z 都具有 x^{-1} 所需的性质, 则对 $xy = 1_G = xz$ 应用乘法消去律可得 $y = z$.

⋆ 若 $x,y\in G$, 则有

$$(xy)^{-1} = y^{-1}x^{-1},$$

这是因为 $y^{-1}x^{-1}xy = y^{-1}y = 1_G$ 而 $xyy^{-1}x^{-1} = xx^{-1} = 1_G$, 故 $y^{-1}x^{-1}$ 确实符合逆元的条件.

⋆ 最简单的群是仅有一个元素 1 的群, 其乘法当然是 $1\cdot 1 = 1$, 称之为**平凡群**.

类似技巧在 §3.1 探讨环和 §4.2 探讨向量空间时已经出现, 此处的形式更为纯粹, 因为群只涉及一种二元运算, 而环或向量空间涉及两种 (加法与乘法).

定义 11.1.2 *群 G 作为集合的基数 $|G|$ 称为其***阶数***.*

基于熟悉的思路, 可以并且应当考虑在群运算之下保持封闭的子结构.

定义 11.1.3 *设 G 为群, 而 H 为 G 的子集. 如果*

(i) $1_G \in H$,

(ii) *对乘法封闭: 若 $x, y \in H$ 则 $xy \in H$,*

(iii) *对取逆封闭: 若 $x \in H$ 则 $x^{-1} \in H$,*

*则称 H 为 G 的***子群***; 此时 $(H, \cdot)$ 也是群, 以 $1_H := 1_G$ 为其幺元.*

群 G 有当然的子群 G 和 $\{1\}$, 我们称 $\{1\}$ 为**平凡子群**; 满足 $H \subsetneq G$ 的子群称为**真子群**. 任意一族子群的交仍是子群.

例 11.1.4 (群的中心) 对任意群 G, 定义

$$Z_G := \{z \in G : \forall g \in G,\ zg = gz\}.$$

这是 G 的子群: 显然 $1_G \in Z_G$, 其乘法封闭性来自 $z_1, z_2 \in Z_G \implies z_1 z_2 g = z_1 g z_2 = g z_1 z_2$; 最后, $zg = gz$ 等价于 $gz^{-1} = z^{-1}g$, 由此可得取逆封闭性. 我们称 Z_G 为 G 的中心; G 交换等价于 $Z_G = G$.

定义 11.1.5 (半群与幺半群) *倘若在定义 11.1.1 中只要求结合律成立, 则对应的代数结构称为半群; 倘若只要求结合律和幺元存在, 则对应的代数结构称为幺半群. 乘法运算满足交换律的半群或幺半群仍称为交换的.*

基于熟悉的论证, 幺半群的幺元依然唯一. 因此幺半群是有幺元的半群, 而群是所有元素皆有逆的幺半群. 不难猜出子幺半群的合理定义: 若 M 为幺半群, 则子幺半群意谓包含幺元的乘法封闭子集.

在幺半群中, 我们可以对任意元素 x 和 $n \in \mathbb{Z}_{\geq 0}$ 定义

$$x^n := \underbrace{x \cdots x}_{n \text{ 项}}, \quad \text{特例 } x^0 := 1;$$

在群中则可以进一步对 $n \in \mathbb{Z}_{<0}$ 定义

$$x^n := \left(x^{|n|}\right)^{-1} = \left(x^{-1}\right)^{|n|}.$$

无论哪种情况, 上述记法都满足

$$x^{m+n} = x^m x^n, \quad x^{mn} = (x^m)^n, \quad x^{-n} = (x^{-1})^n = (x^n)^{-1}.$$

练习 11.1.6 设 G 为群, 定义二元运算 $\cdot^{\mathrm{op}} : G \times G \to G$ 为 $x \cdot^{\mathrm{op}} y := y \cdot x$. 说明 G 对 $\cdot^{\mathrm{op}}$ 仍是群, 称为 G 的**相反群**, 对应的结构记为 G^{op}. 说明 G 交换等价于 $G = G^{\mathrm{op}}$. 这套构造对幺半群或半群同样适用.

提示〉 群的公理左右对称.

练习 11.1.7 设 H 为群 G 的子群, $g \in G$. 验证

$$gHg^{-1} := \{ghg^{-1} : h \in H\}$$

是 G 的子群, 称为 H 对 g 的共轭子群.

例 11.1.8 (对称群) 非空集 X 上的全体置换 (定义 5.1.1) 构成对称群 $\mathfrak{S}_X$, 又称置换群: 它的二元运算是双射的合成 $(\sigma_1, \sigma_2) \mapsto \sigma_1\sigma_2 := \sigma_1 \circ \sigma_2$, 幺元是恒等映射 id_X, 而 $\sigma \in \mathfrak{S}_X$ 的逆元无非是它作为双射的逆映射. 所需的结合律等性质全部归结为映射的基本操作.

对于 $n \in \mathbb{Z}_{\geq 1}$, 群 $\mathfrak{S}_n := \mathfrak{S}_{\{1,\dots,n\}}$ 也称为 n 元对称群或置换群, 它有 $n!$ 个元素. 行列式的研究中已经广泛运用 $\mathfrak{S}_n$ 的群结构.

例 11.1.9 (交错群) 全体偶置换 (见定义 5.1.12 及其后讨论) 构成 $\mathfrak{S}_n$ 的子集, 记为 $\mathfrak{A}_n$, 它是 $\mathfrak{S}_n$ 的子群, 称为 n 元交错群: 首先, 恒等 id 当然是偶置换; 若 $\sigma_1, \sigma_2 \in \mathfrak{A}_n$ 则 $\mathrm{sgn}(\sigma_1\sigma_2) = \mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2) = 1$ 蕴涵乘法封闭性; 最后, $\sigma \in \mathfrak{A}_n$ 蕴涵 $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)^{-1} = 1$.

例 11.1.10 (一般线性群) 设 V 为域 F 上的向量空间, 则 V 的所有自同构 $T : V \xrightarrow{\sim} V$ 构成集合 $\mathrm{GL}(V)$, 它对线性映射的合成运算成群. 这是合理的, 因为同构的合成仍是同构, 结合律一眼可见. 群 $\mathrm{GL}(V)$ 以恒等自同构为 id_V 为幺元, 而 $T \in \mathrm{GL}(V)$ 的逆元是逆同构 T^{-1}.

矩阵情形自然是类似的, 相当于在上述讨论中取 $V = F^n$. 所有 $n \times n$ 可逆矩阵构成集合 $\mathrm{GL}(n, F)$, 它对矩阵乘法成群, 以 $\mathbf{1}_{n\times n}$ 为幺元, 而 $\boldsymbol{A} \in \mathrm{GL}(n, F)$ 的逆元是逆矩阵 $\boldsymbol{A}^{-1}$. 一切都是矩阵运算的简单推论. 这些群称为域 F 上的一般线性群.

例 11.1.11 (特殊线性群) 承继上一则例子, 设 V 是有限维的, 命

$$\mathrm{SL}(V) := \{T \in \mathrm{GL}(V) : \det T = 1\},$$

则 $\mathrm{SL}(V)$ 是 $\mathrm{GL}(V)$ 的子群, 这是因为 $\det(\mathrm{id}_V) = 1$ 而 $\det(T_1T_2) = \det(T_1)\det(T_2)$, $\det(T^{-1}) = \det(T)^{-1}$.

与此类似, $\mathrm{SL}(n, F) := \{\boldsymbol{A} \in \mathrm{GL}(n, F) : \det \boldsymbol{A} = 1\}$ 也是 $\mathrm{GL}(n, F)$ 的子群. 这些群称为域 F 上的特殊线性群.

例 11.1.12 (正交群与特殊正交群) 有限维实内积空间 $(V, (\cdot|\cdot))$ 的所有自同构成群, 称为 $(V, (\cdot|\cdot))$ 的正交群, 这是 $\mathrm{GL}(V)$ 的子群, 记为 $\mathrm{O}(V)$. 另外记 $\mathrm{SO}(V) := \mathrm{O}(V) \cap \mathrm{SL}(V)$, 称之为特殊正交群.

对于标准实内积空间 $\mathbb{R}^n$, 习惯将对应的群记为 $\mathrm{O}(n)$ 与 $\mathrm{SO}(n)$, 它们都是 $\mathrm{GL}(n,\mathbb{R})$ 的子群. 留意到 $\mathrm{SO}(2)$ 由平面上的旋转组成, 这是交换群. 另一方面 $\mathrm{SO}(3)$ 则由空间中的旋转组成, 它的结构相对复杂得多.

例 11.1.13 (辛群) 设 F 为满足 $\mathrm{char}(F)\neq 2$ 的域, $n\in\mathbb{Z}_{\geq 1}$. 对 F 上的辛空间 (V,B) (即: V 是有限维 F-向量空间, $B:V\times V\to F$ 是非退化反对称双线性形式), 命

$$\mathrm{Sp}(V)=\mathrm{Sp}(V,B):=\{g\in\mathrm{GL}(V):\forall x,y\in V,\ B(gx,gy)=B(x,y)\};$$

熟悉的论证表明 $\mathrm{Sp}(V)$ 是 $\mathrm{GL}(V)$ 的子群, 称为辛群.

例 11.1.14 (酉群与特殊酉群) 基于完全类似的思路, 有限维复内积空间 $(W,(\cdot|\cdot))$ 的所有自同构也成群, 称为 $(W,(\cdot|\cdot))$ 的酉群 $\mathrm{U}(W)$, 这是 $\mathrm{GL}(W)$ 的子群. 另外记 $\mathrm{SU}(W):=\mathrm{U}(W)\cap\mathrm{SL}(W)$, 称之为特殊酉群. 对于标准复内积空间 $\mathbb{C}^n$, 习惯将对应的群记为 $\mathrm{U}(n)$ 和 $\mathrm{SU}(n)$, 它们是 $\mathrm{GL}(n,\mathbb{C})$ 的子群.

对于以上几种群, 我们早已身经百战, 明白乘法一般并不交换. 另一方面, 交换群同样是数学的常客.

例 11.1.15 (整数加法群和同余加法群) 整数集 $\mathbb{Z}$ 对加法运算 $+:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$ 成交换群, 小学算术表明它以 0 为其幺元, 而 $n\in\mathbb{Z}$ 对加法的逆元是 $-n$.

进一步, 对任意整数 n, 定义 2.8.2 介绍的

$$\mathbb{Z}/n\mathbb{Z}:=\{\text{整数 }\bmod n\text{ 的同余类}\}$$

对同余类加法也成为群; 若 $n=0$, 此群无非是 $\mathbb{Z}$, 否则它是 $|n|$ 阶有限群.

事实上, 我们知道 $\mathbb{Z}/n\mathbb{Z}$ 具有环结构, 前一则例子因而是以下一般构造的特例.

例 11.1.16 (环的加法群) 设 R 为环, 则它对环的加法运算 $+:R\times R\to R$ 成交换群, 以环的零元 0_R 为其幺元, 而 $r\in R$ 对加法的逆元是 $-r$.

例 11.1.17 (环的可逆元群) 环 R 对乘法不构成群, 它只是幺半群, 以 1_R 为幺元, 这是因为非零环总有不可逆元 (例如零元 0_R). 若只看可逆元子集 $R^\times$, 则 $R^\times$ 对乘法成群. 譬如矩阵环 $\mathrm{M}_{n\times n}(F)$ 的可逆元群正是 $\mathrm{GL}(n,F)$.

由此观之, 尽管幺半群和群在本书中出现的顺序晚于环, 它们却是比环更基础的结构, 而环是搭建在同一个集合 R 上的交换群结构 $(R,+)$ 与幺半群结构 $(R,\cdot)$ 的综合, 以乘法对加法的分配律为纽带.

受以上两种加法群的例子启发, 以下术语显得自然而合理.

约定 11.1.18 (加法群) 对于交换群 A, 有时会将 A 的二元运算记为加法 $+$, 将 A 的幺元记为 0_A 或 0, 将元素 $a\in A$ 对加法的逆元记为 $-a$. 这时我们称 A 为加法群以资区分.

在加法群中, 当 $n \in \mathbb{Z}_{\geq 0}$ 时记 $na := \underbrace{a + \cdots + a}_{n\text{ 项}}$, 记 $(-n)a := -(na)$; 这些规定隐含 $0 \cdot a := 0_A$.

加法群的倍数运算也具有以下标准性质

$$(m+n)a = ma + na, \quad (-n)a = -(na) = n(-a), \quad m(na) = (mn)a;$$

此外 $m(a_1 + a_2) = ma_1 + ma_2$, 这是加法交换性的体现.

注记 11.1.19 容易确定加法群 $(\mathbb{Z}, +)$ 的所有子群: 以下说明它们都能写作 $d\mathbb{Z}$ 的形式, $d \in \mathbb{Z}_{\geq 0}$. 这是 $\mathbb{Z}$ 特有的性质.

首先, 子群 $A \subset \mathbb{Z}$ 的条件一方面要求 A 对加法封闭, 因而所有 $a \in A$ 都满足

$$n \in \mathbb{Z}_{\geq 0} \implies na = a + \cdots + a \in A.$$

另一方面, 取逆封闭性蕴涵 $-a \in A$, 故

$$n \in \mathbb{Z}_{<0} \implies na = (-|n|)a = |n|(-a).$$

因此 A 对取任意倍数封闭. 代入引理 2.7.2 可知存在唯一的 $d \in \mathbb{Z}_{\geq 0}$ 使得 $A = d\mathbb{Z}$.

关于群的另一则重要观念是任意子集在运算之下生成的子结构.

定义 11.1.20 设 G 为群, S 为 G 的子集. 定义 $\langle S \rangle$ 为所有形如 $s_1^{a_1} \cdots s_m^{a_m}$ 的元素构成的子集, 其中 $m \in \mathbb{Z}_{\geq 0}$, $s_i \in S$ 而 $a_i \in \mathbb{Z}$; 对应到 $m = 0$ 的"空乘积"理解为 1.

上述定义导致 $\langle S \rangle$ 含 1 并且对乘法和取逆封闭, 从而是 G 的子群, 称之为 S 生成的子群. 特例 $\langle \varnothing \rangle$ 应当理解为平凡子群 $\{1\}$.

容易看出 $\langle S \rangle$ 是包含 S 的最小子群: G 的任何子群若包含 S, 则包含所有形如 $s_1^{a_1} \cdots s_m^{a_m}$ 的元素, 故包含 $\langle S \rangle$. 特别地, $\langle \varnothing \rangle := \{1\}$ 是完全合理的.

对于 S 为有限集 $\{s_1, \ldots, s_m\}$ 的情形, 我们沿用 $\langle S \rangle = \langle s_1, \ldots, s_m \rangle$ 的方便记法. 若 $G = \langle S \rangle$ 则称 S 生成 G, 或称 S 是 G 的一族**生成元**. 如果 G 能由某个有限子集生成, 则称 G 为**有限生成**的.

在 §4.4 关于线性组合的讨论中, 对向量空间的子集张成的子空间采用了相同符号, 它们的理路相通.

例 11.1.21 对称群 $\mathfrak{S}_n$ 由单对换 $s_1, \ldots, s_{n-1}$ 生成, 此处 $s_i := (i \quad i+1)$; 这是命题 5.1.8 的一部分内容.

从已有的群构造新群的一种方法是取它们作为集合的积, 然后赋予它自然的群结构.

定义–命题 11.1.22 (群的直积) 设 $(G_i)_{i \in I}$ 为一族群, 其中 I 是非空集. 在集合的积 $\prod_{i \in I} G_i$ (见 (2.3.1)) 上定义二元运算

$$(x_i)_{i \in I} \cdot (y_i)_{i \in I} := (x_i y_i)_{i \in I}.$$

这使得 $\prod_{i\in I} G_i$ 成为群, 其幺元是 $(1_{G_i})_{i\in I}$, 而 $(x_i)_{i\in I}^{-1} = \left(x_i^{-1}\right)_{i\in I}$.

对于所有 G_i 取为同一个群 G 的情形, 相应的直积记为 G^I.

群论公理的验证毫无困难; 简言之, 一切都化到每个 G_i 上的分量来检查. 有限多个群的乘积也写作诸如 $G_1 \times G_2 \times \cdots$ 或 G^n 的形式. 不难想象, 直积的顺序应该不影响群结构, 亦即 $G_1 \times (G_2 \times G_3)$ 和 $(G_1 \times G_2) \times G_3$ 应当等同; 这种 "等同" 的严格意涵需要同构的概念来表述, 待 §11.2 处理.

练习 11.1.23 以 (5.1.3) 的表法定义 $\mathfrak{S}_4$ 的元素

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad \tau := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

计算置换 σ^2, τ^2, $\sigma\tau$ 和 $\tau\sigma$, 并说明

$$V := \{\mathrm{id}, \sigma, \tau, \sigma\tau\}$$

是 $\mathfrak{S}_4$ 的交换子群. 群 $V = \langle \sigma, \tau \rangle$ 又称为 **Klein 4-群**. 具体写下 V 的 4×4 乘法表.

11.2 同态与同构

群是一种代数结构. 保结构的映射是代数学的主角之一, 譬如环同态之于环, 或者线性映射之于向量空间. 对于群的例子, 相应的映射称为群同态.

定义 11.2.1 设 $f : G \to G'$ 为群之间的映射. 当以下条件成立时, 称 f 为**群同态**, 或简称为同态:

$$f(xy) = f(x)f(y), \quad x, y \in G.$$

群同态的合成显然仍是群同态.

这一切与环同态的定义 3.2.1 颇为相似, 但又有些微差异.

- ⋆ 定义未要求 f 保持幺元, 因为这是自动的: 我们有 $f(1_G)1_{G'} = f(1_G) = f(1_G 1_G) = f(1_G)f(1_G)$, 对此应用群 G' 的乘法消去律, 即得 $f(1_G) = 1_{G'}$.
- ⋆ 由此可知同态 f 也自动保逆: 我们有

$$f(x^{-1})f(x) = f(x^{-1}x) = f(1_G) = f(xx^{-1}) = f(x)f(x^{-1}),$$

而且已知 $f(1_G) = 1_{G'}$, 故 $f(x^{-1}) = f(x)^{-1}$.

注记 11.2.2 对于定义 11.1.5 介绍的半群和幺半群也有同态的概念.

- ⋆ 如果 $f : S \to S'$ 是半群之间的映射, 满足恒等式 $f(xy) = f(x)f(y)$, 则称 f 为半群同态.

⋆ 如果 $f : M \to M'$ 是幺半群之间的映射, 满足恒等式 $f(xy) = f(x)f(y)$ 和 $f(1_M) = 1_{M'}$, 其中 1_M 和 $1_{M'}$ 分别是 M 和 M' 的幺元, 则称 f 为幺半群同态.

相较于群的版本, 条件 $f(1_M) = 1_{M'}$ 对于幺半群必需另加, 这是因为幺半群未必有乘法消去律, 此前关于 $f(1_G) = 1_{G'}$ 的论证不再适用.

例 11.2.3 设 H 为 G 的子群, 则由 $\iota(h) = h$ 确定的包含映射 $\iota : H \to G$ 是群同态.

对于一般的群同态 $f : G \to G'$, 像集 $\mathrm{im}(f)$ 是 G' 的子群: 这是群同态保持乘法, 幺元和逆元的直接结论. 当 f 是单射时, 由于两边的群结构匹配, 将 G 通过 f 视同 G' 的子群往往是方便的, f 则相应地等同于包含映射.

另一方面, 若 $H' \subset G'$ 为子群, 则也容易验证逆像 $f^{-1}(H')$ 为 G 的子群.

例 11.2.4 集合 $\{\pm 1\}$ 对乘法成群. 设 $n \in \mathbb{Z}_{\geq 1}$. 定义–命题 5.1.11 的映射 $\mathrm{sgn} : \mathfrak{S}_n \to \{\pm 1\}$ 是群同态.

例 11.2.5 设 F 为域, $n \in \mathbb{Z}_{\geq 1}$. 考虑群 F 的可逆元群 $F^\times$. 行列式的乘性说明 $\det : \mathrm{GL}(n, F) \to F^\times$ 是群同态. 如果容许不可逆元, 则 $\det : \mathrm{M}_{n\times n}(F) \to F$ 是幺半群同态, 此处幺半群的运算仍取为乘法.

例 11.2.6 考虑一族群 $(G_i)_{i\in I}$, 其中 I 是非空集. 取群的直积 $\prod_{i\in I} G_i$. 对所有 $j \in I$ 定义映射

$$\begin{aligned} p_j : \prod_{i\in I} G_i &\to G_j \\ (g_i)_{i\in I} &\mapsto g_j. \end{aligned}$$

从群直积的定义可见 p_j 保乘法 (请验证), 因而是群同态, 称为直积的第 j 个**投影同态**. 反观之, 容易证明 $\prod_{i\in I} G_i$ 上的群结构是使每个映射 p_j 都成为群同态的唯一选择.

定义 11.2.7 设 $f : G \to G'$ 为群同态. 如果存在群同态 $g : G' \to G$ 使得 $gf = \mathrm{id}_G$ 而 $fg = \mathrm{id}_{G'}$, 则称 f 为**群同构**, 简称同构, 而 g 为 f 的逆. 此时我们也说 G 和 G' 同构.

群 G 的恒等映射 id_G 是同构的平凡例子. 同构的合成依然是同构. 从 G 映到其自身的同构称为 G 的自同构.

条件 $gf = \mathrm{id}_G$ 和 $fg = \mathrm{id}_{G'}$ 表明 g 作为映射是 f 的逆, 因而 f 和 g 都是双射. 我们有以下的事实.

命题 11.2.8 设 $f : G \to G'$ 为群同态. 如果 f 是集合之间的双射, 则 f 是群同构.

证明 类似的陈述已在环和向量空间的场合见过. 群的情形更简单. 关键是证 f 的逆映射 f^{-1} 也是同态. 对 $f(xy) = f(x)f(y)$ 两边取 f^{-1}, 并且记 $u = f(x)$, $v = f(y)$, 即得 $f^{-1}(u)f^{-1}(v) = f^{-1}(uv)$. □

现在可以顾名思义: 两个群同构相当于说两者的元素能通过一个双射来对应, 而且群的运算也在双射之下对应. 它们的元素名异实同, 是结构相同的群.

约定 11.2.9 设 G 和 G' 为群. 今后以符号 $f : G \xrightarrow{\sim} G'$ 代表映射 $f : G \to G'$ 是群同构. 在不必指明 f 的场合, 我们也以符号 $G \simeq G'$ 代表群 G 和 G' 同构.

无需赘言, 对幺半群或半群也有同构的概念, 而且上述结果依然成立, 惯例依然适用.

练习 11.2.10 验证 $\mathrm{GL}(2,\mathbb{Q})$ 的子集

$$V' := \left\{ \mathbf{1}_{2\times 2}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

是子群. 写下乘法表以说明 V' 和练习 11.1.23 的群 V 以及 $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ 相互同构. 试确定它们是否和 $\mathbb{Z}/4\mathbb{Z}$ 同构.

练习 11.2.11 设 G_1, G_2, G_3 为群. 验证以下两个双射

$$\begin{array}{ccccc} (G_1 \times G_2) \times G_3 & \xrightarrow{1:1} & G_1 \times G_2 \times G_3 & \xleftarrow{1:1} & G_1 \times (G_2 \times G_3) \\ ((g_1, g_2), g_3) & \longmapsto & (g_1, g_2, g_3) & \longmapsfrom & (g_1, (g_2, g_3)) \end{array}$$

皆是群同构.

练习 11.2.12 对群 G 和 $g \in G$, 以 $\mathrm{Ad}_g(x) = gxg^{-1}$ 定义映射 $\mathrm{Ad}_g : G \to G$. 说明:

(i) Ad_g 是群 G 的自同构;

(ii) $\mathrm{Ad}_g = \mathrm{id}_G$ 等价于 $gh = hg$ 对所有 $h \in G$ 成立, 等价于 $g \in Z_G$, 因此 G 是交换群当且仅当 $\mathrm{Ad}_g = \mathrm{id}_G$ 恒成立;

(iii) $\mathrm{Ad}_g \mathrm{Ad}_h = \mathrm{Ad}_{gh}$ 而 $(\mathrm{Ad}_g)^{-1} = \mathrm{Ad}_{g^{-1}}$.

这种自同构称为 G 的**内自同构**, 类似构造在矩阵或线性映射的研究中早已司空见惯. 若将 G 的所有自同构作成群 $\mathrm{Aut}(G)$, 二元运算取作同构的合成, 则 $g \mapsto \mathrm{Ad}_g$ 给出群同态 $\mathrm{Ad} : G \to \mathrm{Aut}(G)$.

推而广之, 一切代数结构 (群, 环, 向量空间) X 的自同构都成群, 记为 $\mathrm{Aut}(X)$; 对于无结构的集合 X, 对应的自同构便是惯常记为 $\mathfrak{S}_X$ 的对称群.

11.3 循环群

本节介绍最为简单的一类群, 称为循环群, 并且在同构意义下予以分类.

定义 11.3.1 能由单个元素生成的群称为**循环群**.

换言之, G 是循环群当且仅当存在 $\sigma \in G$ 使得 $G = \langle \sigma \rangle$, 或者说 G 的一切元素都能写作 σ^k 之形, 其中 $k \in \mathbb{Z}$. 循环群的初步例子是 $\mathbb{Z}/n\mathbb{Z}$, 其中 $n \in \mathbb{Z}_{\geq 0}$; 留意到 $n = 0$ 给出 $\mathbb{Z}$. 它们实际上穷尽了所有循环群.

命题 11.3.2 *设 G 为循环群, 由元素 σ 生成.*

- ⋆ *若 G 无穷, 则有同构 $\mathbb{Z} \xrightarrow{\sim} G$, 映 k 为 σ^k;*
- ⋆ *若 G 有限, 记 $n := |G|$, 则有同构 $\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} G$, 映 $k + n\mathbb{Z}$ 为 σ^k.*

在有限的情形, $|G| = \min\{k \in \mathbb{Z}_{\geq 1} : \sigma^k = 1\}$.

证明 无论哪种情形, 总有满射 $\mathbb{Z} \to G$, 映 k 为 σ^k. 由于 $\sigma^{k+k'} = \sigma^k \sigma^{k'}$, 这还是从加法群 $\mathbb{Z}$ 到 G 的满同态. 以下分两种情况讨论.

若 $\mathbb{Z} \to G$ 是单的, 则它是群同构; 特别地, 此时 G 无穷.

以下设 $\mathbb{Z} \to G$ 非单, 因此存在 $i < j$ 使得 $\sigma^i = \sigma^j$, 从而 $\sigma^{j-i} = 1$. 现在可以合理地定义

$$n := \min\left\{k \in \mathbb{Z}_{\geq 1} : \sigma^k = 1\right\}.$$

因为 $\sigma^n = 1$, 元素 $g = \sigma^k$ 只和 k 的 $\bmod n$ 同余类 $k + n\mathbb{Z}$ 相关. 现在说明 $k + n\mathbb{Z}$ 由 g 唯一确定: 若有 $0 \leq k \leq h < n$ 使得 $\sigma^k = \sigma^h$, 则 $\sigma^{h-k} = 1$ 而 $0 \leq h - k < n$, 故 n 的选法蕴涵 $h = k$.

综上可得双射 $\mathbb{Z}/n\mathbb{Z} \to G$, 映 $k + n\mathbb{Z}$ 为 σ^k. 由于 $\sigma^{k+k'} = \sigma^k \sigma^{k'}$ 而 $(k + k') + n\mathbb{Z} = (k + n\mathbb{Z}) + (k' + n\mathbb{Z})$. 它实则是群同构. 特别地, 此时 $|G| = n$ 有限. □

进一步, 不同的 $n \in \mathbb{Z}_{\geq 0}$ 给出的循环群 $\mathbb{Z}/n\mathbb{Z}$ 互不同构, 缘由很简单: 它们作为集合的大小不同.

循环群的研究因此化约到 $\mathbb{Z}/n\mathbb{Z}$ 的研究, $n \in \mathbb{Z}_{\geq 0}$. 我们可以进一步明确这些群的生成元.

命题 11.3.3 *设 n 为非零整数, 则同余类 $a + n\mathbb{Z}$ 生成加法群 $\mathbb{Z}/n\mathbb{Z}$ 当且仅当整数 a 与 n 互素.*

证明 已知 $1 + n\mathbb{Z}$ 生成 $\mathbb{Z}/n\mathbb{Z}$, 因此 $a + n\mathbb{Z}$ 生成 $\mathbb{Z}/n\mathbb{Z}$ 当且仅当 $1 + n\mathbb{Z} \in \langle a + n\mathbb{Z} \rangle$, 而这又等价于存在 $x \in \mathbb{Z}$ 使得 $ax \equiv 1 \pmod{n}$. 然而此同余方程有解的充要条件已知是 $\gcd(a, n) = 1$. □

推论 11.3.4 设 G 是有限循环群, $n := |G|$, 则 G 的生成元个数为 $\varphi(n)$, 其中 φ 是 Euler 函数.

证明 不失一般性, 可设 $G = \mathbb{Z}/n\mathbb{Z}$. 回忆到 $\varphi(n)$ 是集合 $\{1 \leq a \leq n : \gcd(a,n) = 1\}$ 的元素个数, 而 $\{1,\dots,n\}$ 又是 $\mathbb{Z}/n\mathbb{Z}$ 在 $\mathbb{Z}$ 中的一族代表元. 代入命题 11.3.3. □

11.4 陪集分解

考虑群 G 的子群 H 和 $x, y \in G$.

- ⋆ 若存在 $h \in H$ 使得 $x = hy$, 则记为 $x \sim_{左} y$;
- ⋆ 若存在 $h \in H$ 使得 $x = yh$, 则记为 $x \sim_{右} y$.

引理 11.4.1 以上定义的 $\sim_{左}$ 和 $\sim_{右}$ 都是 G 上的等价关系.

证明 左右两种版本的论证相同, 以下只论 $\sim_{左}$. 首先, 反身性 $x \sim_{左} x$ 源于 $x = 1 \cdot x$ 和 $1 \in H$. 其次, 若存在 $h \in H$ 使得 $x = hy$, 则 $y = h^{-1}hy = h^{-1}x$ 而 $h^{-1} \in H$, 故对称性成立. 设 $x = hy$ 而 $y = h'z$, 其中 $h, h' \in H$, 则 $x = hh'z$ 而 $hh' \in H$, 故传递性成立. □

任何等价关系都将集合划分为等价类. 对所有 $g \in G$, 定义 G 的子集

$$Hg := \{hg : h \in H\}, \quad gH := \{gh : h \in H\}.$$

由定义立见 Hg 和 gH 分别是包含 g 的 $\sim_{左}$ 和 $\sim_{右}$ 等价类. 它们在群论中拥有特殊的名字.

定义 11.4.2 设 H 为群 G 的子群. 相应的**右陪集**是形如 Hg 的子集, **左陪集**是形如 gH 的子集, 其中 $g \in G$; 不致混淆时简称为**陪集**.

既然引理 11.4.1 将左右两种陪集分别诠释为等价类, 任两个右陪集 Hx 和 Hy (或左陪集 xH 和 yH) 或者相等, 或者无交; G 分解为相异右陪集 (或左陪集) 的无交并. 这些等价关系给出的商集也拥有特殊的符号.

定义 11.4.3 设 H 为群 G 的子群, 命

$$H\backslash G := \{\text{右陪集 } Hg : g \in G\}, \quad G/H := \{\text{左陪集 } gH : g \in G\}.$$

对任意子集 $A \subset G$, 记 $A^{-1} := \{a^{-1} : a \in A\} \subset G$.

引理 11.4.4 设 $g \in G$, 我们有 $(Hg)^{-1} = g^{-1}H$.

证明 归结为 $H^{-1} = H$. □

定义–命题 11.4.5 (子群的指数) 设 H 为群 G 的子群. 映射 $Hg \mapsto (Hg)^{-1}$ 给出从 $H\backslash G$ 到 H/G 的双射. 因此 $H\backslash G$ 和 G/H 作为集合有相同的基数, 记之为 $(G:H)$, 称为 H 在 G 中的指数.

证明 引理 11.4.4 表明 $Hg \mapsto (Hg)^{-1}$ 确实映 H 在 G 中的右陪集为左陪集, 因而给出映射 $H\backslash G \to G/H$; 同理, $gH \mapsto (gH)^{-1}$ 给出映射 $G/H \to H\backslash G$. 双向的映射显然互逆. □

若 G 是有限群, 则 H 也有限, 而 $(G:H)$ 自然也是一个正整数, 而不只是集合的基数.

定理 11.4.6 (J.-L. Lagrange) 若 H 是群 G 的子群, 则有 $|H| \cdot (G:H) = |G|$. 当 G 无穷时, 等式的左侧作为基数乘法来理解; 见 §2.9.

证明 要点是给出双射 $H \times (H\backslash G) \xrightarrow{1:1} G$. 对每个右陪集 C (换言之 $C \in H\backslash G$) 选取 $x_C \in G$ 使得 $C = Hx_C$. 定义

$$\begin{aligned} H \times (H\backslash G) &\to G \\ (h, C) &\mapsto hx_C. \end{aligned}$$

这是满射: 任何 $g \in G$ 都属于某个右陪集 C, 因而能表作 $g = hx_C$ 的形式, 其中 $h \in H$.

这也是单射: 若 $hx_C = h'x_{C'}$, 其中 C 和 C' 是右陪集而 $h, h' \in H$, 则 C 和 C' 有交, 故 $C = C'$, 继而有 $x_C = x_{C'}$. 群的乘法消去律进一步确保 $h = h'$. 证毕. □

论证也可以按下述方式理解, 实质不变: G 分解为 $(G:H)$ 个右陪集的无交并, 而每个右陪集 Hg 都和 H 等势, 方法是映 $h \in H$ 为 $hg \in Hg$; 消去律说明这是双射.

推论 11.4.7 若 G 是有限群, 则所有子群 H 的阶数 $|H|$ 都整除 $|G|$.

练习 11.4.8 设 $H, K \subset G$ 为子群, 而且 $|H|$ 和 $|K|$ 是互素的正整数, 证明 $H \cap K = \{1\}$.

定义 11.4.9 (元素的阶) 设 σ 为群 G 的元素. 定义 $\mathrm{ord}(\sigma) := |\langle\sigma\rangle|$, 称为 σ 的**阶**.

按定义, $\langle\sigma\rangle$ 是循环群. 命题 11.3.2 蕴涵

$$\mathrm{ord}(\sigma) = \min\left\{k \in \mathbb{Z}_{\geq 1} : \sigma^k = 1_G\right\},$$

右式在 $\langle\sigma\rangle \simeq \mathbb{Z}$ 的情况下当然地理解为 ∞. 阶数因而体现为最小周期:

$$\sigma^k = 1_G \iff \mathrm{ord}(\sigma) \mid k. \tag{11.4.1}$$

特别地, $\mathrm{ord}(\sigma) = 1$ 当且仅当 $\sigma = 1_G$.

推论 11.4.10 设 G 为有限群, $n := |G|$. 对任何 $\sigma \in G$, 总有 $\mathrm{ord}(\sigma) \mid n$.

证明 因为 $\langle\sigma\rangle$ 是 G 的子群, 代入定理 11.4.6 便是. □

推论 11.4.11 *若 G 是有限群, $|G|$ 是素数, 则 G 是循环群.*

证明 记 $p := |G|$. 任取 $g \in G \smallsetminus \{1_G\}$, 则 $\mathrm{ord}(g) \mid p$ 而 $\mathrm{ord}(g) \neq 1$, 由此可知 $\langle g\rangle = G$, 因为两边元素个数同为 p. □

练习 11.4.12 设 H 为 G 的子群, K 为 H 的子群. 试证 $(G:K) = (G:H)(H:K)$.

提示〉 设有陪集分解 $G = \bigsqcup_{i\in I} Hg_i$ 和 $H = \bigsqcup_{j\in J} Kh_j$, 其中 I 和 J 是适当的下标集. 说明 $G = \bigsqcup_{(i,j)\in I\times J} Kh_jg_i$.

练习 11.4.13 设 R 为交换环, $n \in \mathbb{Z}_{\geq 1}$, 验证

$$\mu_n(R) = \left\{r \in R^\times : r^n = 1\right\}$$

是群 $R^\times$ 的子群, 其元素称为 R 中的 n 次**单位根**. 说明交换环之间的同态 $R \to R'$ 给出群同态 $\mu_n(R) \to \mu_n(R')$.

若 F 是域则 $|\mu_n(F)| \leq n$. 本章习题将说明 $\mu_n(F)$ 总是循环群. 因此若 $r \in \mu_n(F)$ 满足 $\mathrm{ord}(r) = n$, 则 $\mu_n(F) = \langle r\rangle \simeq \mathbb{Z}/n\mathbb{Z}$, 此时 r 也称为 F 中的 n 次**单位原根**. 在 $F = \mathbb{C}$ 时一切清晰: $\mu_n(\mathbb{C}) = \left\langle e^{2\pi \mathrm{i}/n}\right\rangle$, 而 $e^{2\pi \mathrm{i}k/n}$ 是单位原根等价于 $\gcd(k,n) = 1$.

11.5 群作用

引入群的概念时, 我们谈及许多群具体是由某个集合到自身的一类双射组成的, 典型例子是集合 X 上的对称群 $\mathfrak{S}_X$; 每个置换 $\sigma \in \mathfrak{S}_X$ 可以作用在 X 的元素 x 上, 给出 $\sigma(x) \in X$. 这种作用当然地满足 $\mathrm{id}(x) = x$ 和 $(\sigma\tau)(x) = \sigma(\tau(x))$, 其中 $\sigma, \tau \in \mathfrak{S}_X$.

既然有了群的抽象定义, 不妨反过来考察一个群 G 如何作用在集合上.

定义 11.5.1 *群 G 在集合 X 上的***左作用***意谓满足下述性质的映射 $a : G \times X \to X$,*

$$\begin{aligned} a(g_1g_2, x) &= a(g_1, a(g_2, x)),\\ a(1_G, x) &= x, \end{aligned}$$

其中 $g_1, g_2 \in G$ 而 $x \in X$. 我们经常将 $a(g,x)$ 简记为乘法 gx 或 $g\cdot x$, 因此条件也相当于说 $(g_1g_2)x = g_1(g_2x)$ 和 $1_Gx = x$.

类似地, G 在集合 X 上的右作用意谓满足下述性质的映射 $a' : X \times G \to X$, 同样将 $a'(x,g)$ 简记为乘法 xg 或 $x\cdot g$:

$$x(g_1g_2) = (xg_1)g_2, \quad x1_G = x.$$

若无其他说明, 群的作用在本书中默认为左作用.

练习 11.5.2 给定群 G 与集合 X, 说明指定 G 的左作用 $a : G \times X \to X$ 和指定相反群 G^{op} 的右作用 $a' : X \times G^{\mathrm{op}} \to X$ 是一回事, 方法是通过 $a'(x, g) = a(g, x)$ 相互对应. 因此左作用和右作用的一般性质相通, 今后只须处理左作用的情形.

举例来说, 对称群 $\mathfrak{S}_X$ 当然地通过 $a(\sigma, x) = \sigma(x)$ 作用在 X 上. 不妨设想 X 的元素在 G 的作用下移动, 这就导向一些形象化的术语.

定义 11.5.3 设群 G 作用在 X 上. 对所有 $x \in X$, 定义其

▷ **轨道** $Gx := \{gx : g \in G\} \subset X$, 又称 G-轨道;

▷ **稳定化子** $\mathrm{Stab}_G(x) := \{g \in G : gx = x\} \subset G$.

对于右作用同样有类似定义.

稳定化子 $\mathrm{Stab}_G(x)$ 自动是 G 的子群: $1_G x = x$ 蕴涵它非空, $g_1, g_2 \in \mathrm{Stab}_G(x) \implies (g_1 g_2)x = g_1(g_2 x) = g_1 x = x$ 蕴涵它对乘法封闭, 而 $g \in \mathrm{Stab}_G(x) \implies x = g^{-1}gx = g^{-1}x$ 蕴涵它对取逆封闭.

按相同方式也能定义幺半群在集合 X 上的左作用和右作用, 此时每个 $x \in X$ 的稳定化子都是子幺半群. 另一方面, 轨道的概念通常仅对群来考虑, 这是因为以下结果的证明涉及取逆.

定义–命题 11.5.4 设群 G 作用在集合 X 上. 在 X 上定义二元关系 $\sim_G$, 使得 $x \sim_G y$ 当且仅当存在 $g \in G$ 使得 $y = gx$, 则:

(i) $\sim_G$ 是等价关系,

(ii) 对 $\sim_G$ 的等价类正是 X 中的 G-轨道.

对应于的商集记为 $G\backslash X$. 对于右作用也有相应的结果, 对应的商集记为 X/G.

证明 先处理 (i). 由 $x = 1_G x$ 可见反身性成立. 若 $y = gx$ 则 $x = g^{-1}y$, 故对称性成立. 若 $y = gx$ 而 $z = hy$, 其中 $g, h \in G$, 则 $z = h(gx) = (hg)x$ 表明传递性成立.

接着考虑 (ii). 给定 $x \in X$, 定义表明 $y \sim_G x$ 当且仅当 $y \in Gx$, 这就说明含 x 的等价类是 x 的轨道. □

推论 11.5.5 (轨道分解) 设群 G 作用在集合 X 上, 则 X 是其所有 G-轨道的无交并.

以下术语是标准的, 陈述左作用的版本即足.

定义 11.5.6 继续假定群 G 作用在 X 上.

⋆ 若 $\bigcap_{x \in X} \mathrm{Stab}_G(x) = \{1_G\}$, 则称此作用**忠实**.

⋆ 若对所有 x 皆有 $\mathrm{Stab}_G(x) = \{1_G\}$, 则称此作用**自由**.

⋆ 若 X 仅有一个轨道, 换言之对所有 $x, y \in X$ 皆存在 $g \in G$ 使得 $y = gx$, 则称此作用**传递**.

现在进一步明确各个轨道的构造.

引理 11.5.7 设群 G 左作用于 X. 对 $x \in X$, 命 $H := \mathrm{Stab}_G(x)$, 则有双射

$$\begin{aligned} G/H &\xrightarrow{1:1} Gx \\ gH &\mapsto gx. \end{aligned}$$

对于右作用, 相应的双射则写作 $H\backslash G \to xG$.

证明 首先观察到若 $h \in H$, 则 $(gh)x = g(hx) = gx$, 所以 gx 只和陪集 gH 相关, 写下的映射是良定义的. 轨道的定义即刻导致映射为满. 为了证明它还是单射, 观察到 $g_1 x = g_2 x$ 等价于 $g_1^{-1} g_2 x = x$, 等价于 $h := g_1^{-1} g_2 \in H$, 然而这又相当于说存在 $h \in H$ 使得 $g_2 = g_1 h$, 亦即 $g_1 H = g_2 H$. □

练习 11.5.8 证明若 $g \in G$, 则 $\mathrm{Stab}_G(gx) = g\,\mathrm{Stab}_G(x) g^{-1}$.

练习 11.5.9 设 H 为 G 的子群.

(i) 验证群的乘法 $G \times H \to G$ (或 $H \times G \to G$) 给出 H 在 G 上的右 (或左) 作用, 对应的轨道是对 H 的陪集, 而商集则是 G/H (或 $H\backslash G$). 因此, 关于群作用的轨道和商集的符号和陪集版本兼容.

(ii) 定义映射 $a : G \times (G/H) \to G/H$, 使得 $a(g, xH) = gxH$. 验证 a 使 G 左作用在 G/H 上. 验证引理 11.5.7 的双射 $G/H \to Gx$ 让两边的 G-作用相互对应.

这就表明对于带传递 G-作用的 X, 只要任选 $x \in X$ 并且命 $H := \mathrm{Stab}_G(x)$, 则能将 X 连同其 G-作用等同于 G/H.

(iii) 对 (ii) 给出左右交换后的版本.

若 X 是有限集, 则它仅有有限多个 G-轨道, 标作 $Gx_1, \ldots, Gx_n$. 记 $H_i := \mathrm{Stab}_G(x_i)$, 于是轨道分解 $X = \bigsqcup_{i=1}^n Gx_i$ 和引理 11.5.7 即刻导致

$$|X| = \sum_{i=1}^{n} (G : H_i). \tag{11.5.1}$$

等式 (11.5.1) 有助于处理和群作用有关的计数问题. 以下是另一则关乎计数的有用结果, 常被称为 **Burnside 引理**, 论证方法同样基于轨道分解. 本章习题将介绍其用法.

命题 11.5.10 设有限群 G 作用于有限集 X. 对所有 $g \in G$ 定义 $X^g := \{x \in X : gx = x\}$, 则

$$|G\backslash X| \cdot |G| = \sum_{g \in G} |X^g|\,.$$

证明 先从传递作用的情形起步. 选定 $x_0 \in X$. 既然 $\mathrm{Stab}_G(gx_0) = g\,\mathrm{Stab}_G(x_0)g^{-1}$ 导致 $|\mathrm{Stab}_G(gx_0)| = |\mathrm{Stab}_G(x_0)|$, 而 G 的作用是传递的, 应用 Lagrange 定理 11.4.6 和引理 11.5.7 遂得

$$\sum_{g\in G}|X^g| = \sum_{x\in X}|\mathrm{Stab}_G(x)| = (G:\mathrm{Stab}_G(x_0))|\mathrm{Stab}_G(x_0)| = |G|.$$

此外 $|G\backslash X| = 1$, 故断言成立.

对于一般的情形, 将 X 分解为 G-轨道 $X_1,\dots,X_n$ 的无交并, 每个 X_i 都具有传递 G-作用, 因此

$$\sum_{g\in G}|X^g| = \sum_{i=1}^{n}\sum_{g\in G}|X_i^g| = n|G|,$$

而 $n = |G\backslash X|$, 证毕. □

无妨将上述等式理解为 G 的每个元素 "平均而言" 有 $|G\backslash X|$ 个不动点.

注记 11.5.11 以下说明指定群作用 $a: G\times X\to X$ 和指定群同态 $A: G\to \mathfrak{S}_X$ 是等价的. 注意到指定映射 $a: G\times X\to X$ 相当于指定映射 $A: G\to\{\text{映射 } X\to X\}$, 两者之间的对应由下式确定.

$$A(g)(x) = a(g,x),\quad g\in G,\ x\in X.$$

容易看出

$$\begin{aligned} a(g_1g_2,x) = a(g_1,a(g_2,x)) &\iff A(g_1g_2)(x) = A(g_1)\left(A(g_2)(x)\right),\\ a(1_G,x) = x &\iff A(1_G)(x) = x. \end{aligned}$$

左侧是群作用的条件. 当 $x\in X$ 变动, 右侧的条件也等价于说 $A(g_1g_2) = A(g_1)A(g_2)$ 而 $A(1_G) = \mathrm{id}_X$; 这还蕴涵 $A(g)A(g^{-1}) = \mathrm{id}_X$, 所以此时 $A(g)\in\mathfrak{S}_X$. 综之, 右侧条件相当于说 $A: G\to\mathfrak{S}_X$ 是群同态.

如果考虑 G 在 X 上的右作用, 相应的同态应当是 $G^{\mathrm{op}}\to\mathfrak{S}_X$. 作为简单练习, 读者可以检验 $\mathfrak{S}_X$ 在 X 上的置换作用对应到 $\mathrm{id}: \mathfrak{S}_X\to\mathfrak{S}_X$.

通过群作用的概念, 可以将抽象的群嵌入具体的对称群, 群的元素由此也实现为具体的置换.

引理 11.5.12 *群 G 在 X 上的作用忠实 (定义 11.5.6) 当且仅当对应的同态 $A: G\to\mathfrak{S}_X$ 为单.*

证明 同态 A 为单相当于说对于所有 $g_1,g_2\in G$,

$$(\forall x,\ g_1x = g_2x) \iff g_1 = g_2.$$

然而 $g_1x = g_2x \iff g_1^{-1}g_2\in\mathrm{Stab}_G(x)$, 所以左式等价于 $g_1^{-1}g_2\in\bigcap_x\mathrm{Stab}_G(x)$. 另一方面右式等价于 $g_1^{-1}g_2 = 1_G$. □

定理 11.5.13 (A. Cayley) *任何群 G 都能嵌入为 $\mathfrak{S}_G$ 的子群.*

证明 让 G 通过群的乘法作用于 G 本身; 消去律给出 $g_1x = g_2x \iff g_1 = g_2$, 故作用自由, 因而也是忠实的. 由注记 11.5.11 和引理 11.5.12 即得嵌入 $A : G \hookrightarrow \mathfrak{S}_G$. □

11.6 轨道分解的几则应用

在许多场景中, 我们的真正兴致在于群作用的不动点. 以下的 G 都默认为群.

定义 11.6.1 *设 G 作用在 X 上, 记 $X^G := \{x \in X : \forall g,\ gx = x\}$, 其元素称为 X 在 G 作用下的*__不动点__.

不动点也等于只有一个元素的轨道 $\{x\}$.

定义 11.6.2 *设 p 为素数. 若 $|G| = p^m$, 其中 $m \in \mathbb{Z}_{\geq 0}$, 则称 G 为 p-*__群__.

命题 11.6.3 *设 p-群 G 作用在有限集 X 上, 则*

$$|X| \equiv |X^G| \pmod{p}.$$

证明 将 X 分解为轨道 $Gx_1, \ldots, Gx_n$, 然后记 $H_i := \mathrm{Stab}_G(x_i)$. 不动点对应到满足 $|Gx_i| = 1$ 的轨道, 亦即满足 $H_i = G$ 的轨道. 于是 (11.5.1) 化为

$$|X| = |X^G| + \sum_{\substack{1 \leq i \leq n \\ H_i \neq G_i}} (G : H_i).$$

然而 Lagrange 定理 11.4.6 说明 $|H_i|$ 整除 $|G| = p^m$, 从而当 $H_i \neq G_i$ 时 $(G : H_i)$ 是 p 的某个正幂次, 故上式蕴涵 $|X| \equiv |X^G| \pmod{p}$. □

命题 11.6.3 的第一则应用涉及例 11.1.4 介绍的中心 Z_G.

命题 11.6.4 *设 G 为非平凡的 p-群, 则 $Z_G \neq \{1\}$.*

证明 定义映射 $a : G \times G \to G$ 如下

$$a(g, x) = gxg^{-1},$$

或者用练习 11.2.12 的符号表作 $a(g, x) = \mathrm{Ad}_g(x)$. 这是 G 对其自身的左作用. 对于 $g, x \in G$, 等式 $gxg^{-1} = x$ 等价于 $gx = xg$, 因此作用下的不动点集是 Z_G. 从 $|Z_G| \equiv |G| \equiv 0 \pmod{p}$ 立见 $Z_G \neq \{1\}$. □

命题 11.6.3 的第二则应用是称为 Cauchy 定理的基本结果. 它的方向与推论 11.4.10 正好相反.

定理 11.6.5 (A.-L. Cauchy) *设 G 为有限群, p 为 $|G|$ 的素因数, 则存在 $g \in G$ 使得 $\mathrm{ord}(g) = p$.*

证明 让循环群 $\mathbb{Z}/p\mathbb{Z}$ 通过轮换作用于下述集合

$$X := \{(g_1, \dots, g_p) \in G^p : g_1 \cdots g_p = 1\};$$

精确地说, 不妨将 $(g_i)_{i=1}^p \in X$ 的下标 i 视同 $\mathbb{Z}/p\mathbb{Z}$ 的元素, 然后定义群作用所对应的 $a : (\mathbb{Z}/p\mathbb{Z}) \times X \to X$ 为

$$a\,(k + p\mathbb{Z}, (g_i)_i) = (g_{i+k})_i;$$

观察到若 $(g_i)_i \in X$, 则

$$g_2 \cdots g_p g_1 = g_1^{-1}(g_1 \cdots g_p)g_1 = g_1^{-1} g_1 = 1,$$

故 $(g_2, \dots, g_p, g_1) \in X$. 同理 $(g_3, \dots, g_p, g_1, g_2) \in X$, 依此类推. 因此 a 确实取值在 X 中.

进一步, $X^{\mathbb{Z}/p\mathbb{Z}} = \{(g, \dots, g) \in G^p : g^p = 1\}$. 从 (11.4.1) 可见 $g^p = 1$ 蕴涵 $\operatorname{ord}(g) \in \{1, p\}$, 而 $\operatorname{ord}(g) = 1$ 无非是说 $g = 1$. 综上, 问题归结为证明 $X^{\mathbb{Z}/p\mathbb{Z}}$ 含有 $(1, \dots, 1)$ 之外的元素.

指定 $(g_1, \dots, g_p) \in X$ 相当于指定 $(g_1, \dots, g_{p-1}) \in G^{p-1}$, 因为剩下的 g_p 由等式 $g_1 \cdots g_p = 1$ 唯一确定. 于是 $|X| = |G|^{p-1}$ 被 p 整除. 现在对 $\mathbb{Z}/p\mathbb{Z}$ 在 X 上的作用应用命题 11.6.3, 可得

$$\left|X^{\mathbb{Z}/p\mathbb{Z}}\right| \equiv |X| \equiv 0 \pmod{p}.$$

于是 $X^{\mathbb{Z}/p\mathbb{Z}} \supsetneq \{(1, \dots, 1)\}$, 明所欲证. □

鉴于推论 11.4.11, Cauchy 定理等价于说若 p 是 $|G|$ 的素因数, 则存在 G 的 p 阶子群. 推而广之, 我们也可以问具有其他指定阶数的子群 H 是否存在. Lagrange 定理 11.4.6 对此提供了必要条件: $|H|$ 必须整除 $|G|$. 至于充分性, 特别常用的是下述情形.

定义 11.6.6 (Sylow p-子群) *设 G 为有限群, p 为素数. 满足 $|P| = p^a$ 的子群 P 称为 G 的 p-子群, 其中 $a \in \mathbb{Z}_{\geq 0}$; 若进一步有 $p^a \parallel |G|$, 则称 P 为 G 的* Sylow p-子群.

因此 Sylow p-子群是在 Lagrange 定理的约束下所能存在的最大 p-子群. 给定 G 和素数 p 如上, 关于 Sylow p-子群有以下基本结论.

▷ Sylow **第一定理** 存在 Sylow p-子群.

▷ Sylow **第二定理** 任两个 Sylow p-子群 $P, P' \subset G$ 皆共轭: 存在 $g \in G$ 使得 $P' = gPg^{-1}$.

▷ Sylow **第三定理** 令 N_p 为 G 中的 Sylow p-子群个数, 则 $N_p \equiv 1 \pmod{p}$.

虽然 Sylow 定理只需要轨道分解和简单的数论论证, 在此给出证明则岔题太远, 细节可参阅 [10, §4.5] 或任何一本群论教材. 在许多实际场景中, 重点是对具体的群写下一个 Sylow p-子群, 而 Sylow 第二定理说明其他 Sylow p-子群是它的共轭; 本章习题将有这方面的例子.

11.7 应用: 置换的循环分解

在非空集 X 上的所有置换中, 轮换是特别常见的; 顾名思义, 它的效果是将 X 中的一列元素轮流地调换, 周而复始. 我们给出正式的定义.

定义 11.7.1 设 X 为非空集, $a_1,\dots,a_m$ 是 X 中相异的元素 ($m\in\mathbb{Z}_{\geq 1}$). 方便起见, 取同余类将 a_i 的下标 i 视同 $\mathbb{Z}/m\mathbb{Z}$ 的元素. 具有以下形式的置换

$$\begin{aligned}\sigma &: X\to X\\ \sigma(a_i)&=a_{i+1},\quad i\in\mathbb{Z}/m\mathbb{Z},\\ \sigma(x)&=x,\quad x\in X\smallsetminus\{a_1,\dots,a_m\},\end{aligned}$$

称为 m 中的 m-**循环**或**轮换**, 也记为 $\sigma=(a_1\ \cdots\ a_m)\in\mathfrak{S}_X$. 对于 $m=2$ 的情形, 相应的 $(a_1\ a_2)$ 称为元素 a_1 和 a_2 的**对换**.

符号 $(a_1\ \cdots\ a_m)$ 和 $(a_2\ \cdots\ a_m\ a_1)$ 因而代表相同的循环或轮换, 依此类推. 对于 $X=\{1,\dots,n\}$ 的情形, 此处定义的对换以及相应的符号和定义 5.1.2 兼容. 注意到 1-循环无非是 id_X.

练习 11.7.2 扼要地说明 m-循环 $\xi:=(a_1\ \cdots\ a_m)$ 满足 $\mathrm{ord}(\xi)=m$.

定义 11.7.3 给定集合 X 上的循环 $(a_1\ \cdots\ a_n)$ 和 $(b_1\ \cdots\ b_m)$, 若 $\{a_1,\dots,a_n\}\cap\{b_1,\dots,b_m\}=\varnothing$, 则称这两个循环不交.

引理 11.7.4 设循环 $\sigma,\tau\in\mathfrak{S}_X$ 不交, 则 $\sigma\tau=\tau\sigma$.

证明 由于 σ 和 τ 分别在两个无交的子集上作用, 且记为 $A,B\subset X$, 合成 $\sigma\tau$ 和 $\tau\sigma$ 的作用因而相同: 它们都映 $a\in A$ 为 $\sigma(a)$, 映 $b\in B$ 为 $\tau(b)$, 对其他元素则是恒等. □

今后将聚焦于 X 为非空有限集的情形. 不妨就设 $X=\{1,\dots,n\}$, 其中 $n\in\mathbb{Z}_{\geq 1}$, 并考虑相应的对称群 $\mathfrak{S}_n$.

命题 11.7.5 (循环分解) 设 $n\in\mathbb{Z}_{\geq 1}$ 而 $\sigma\in\mathfrak{S}_n$, 则 σ 分解为两两不交的循环的乘积

$$\sigma=(a_{1,1}\ \cdots\ a_{1,\ell_1})(a_{2,1}\ \cdots\ a_{2,\ell_2})\cdots(a_{m,1}\ \cdots\ a_{m,\ell_m}),$$

满足 $\sum_{i=1}^m\ell_i=n$, 而且这 m 个循环是唯一的, 至多差一个重排.

证明 有限子群 $\langle\sigma\rangle$ 作用在 $\{1,\dots,n\}$ 上, 因此推论 11.5.5 给出 $\langle\sigma\rangle$-轨道的无交并

$$\{1,\dots,n\}=C_1\sqcup\cdots\sqcup C_m.$$

对每个 $1\leq i\leq m$ 任取 $a_{i,1}\in C_i$, 再取最小的 $\ell_i\in\mathbb{Z}_{\geq 1}$ 使得 $\sigma^{\ell_i}(a_{i,1})=a_{i,1}$. 命 $a_{i,k}:=\sigma^{k-1}(a_{i,1})$, 则 σ 在 C_i 上的映法是

$$a_{i,1}\overset{\sigma}{\longmapsto}a_{i,2}\overset{\sigma}{\longmapsto}\cdots\overset{\sigma}{\longmapsto}a_{i,\ell_i}\overset{\sigma}{\longmapsto}a_{i,1},$$

而且当 $1 \le h < k \le \ell_i$ 时 $a_{i,h} \neq a_{i,k}$, 否则将有 $\sigma^{k-h}(a_{i,1}) = a_{i,1}$, 与 ℓ_i 的选法矛盾. 因此 $|C_i| = \ell_i$.

综上, σ 在子集 C_i 上按照循环 $(a_{i,1}\ \cdots\ a_{i,\ell_i})$ 的方式作用. 既然不同轨道无交, 对应的循环当然也不交. 综上即得所求的循环分解, 而且 $n = \sum_{i=1}^{m} |C_i| = \sum_{i=1}^{m} \ell_i$.

反过来说, 若 σ 有如是的循环分解, 命 $C_i := \{a_{i,1}, \dots, a_{i,\ell_i}\}$, 则从循环的定义可见 $\langle\sigma\rangle$ 的轨道确实就是 $C_1, \dots, C_m$, 而 $\sigma|_{C_i}$ 则等于循环 $(a_{i,1}\ \cdots\ a_{i,\ell_i})$. 因此循环分解的唯一性归结为轨道分解的唯一性, 精确到轨道的重排. □

既然 1-循环是恒等, 我们经常在循环分解中省略 $\ell_i = 1$ 的部分.

置换的奇偶性和阶数都容易从循环分解读出.

命题 11.7.6 设 σ 有如命题 11.7.5 的循环分解, 则

$$\begin{aligned} \operatorname{sgn}(\sigma) &= (-1)^{\sum_{i=1}^{m}(\ell_i - 1)}, \\ \operatorname{ord}(\sigma) &= \operatorname{lcm}(\ell_1, \dots, \ell_m). \end{aligned}$$

证明 命 $\xi_i := (a_{i,1}\ \cdots\ a_{i,\ell_i})$. 对于 $\operatorname{sgn}(\sigma)$ 的部分, $\operatorname{sgn}(\sigma) = \prod_{i=1}^{m} \operatorname{sgn}(\xi_i)$ 将问题化为对所有 ℓ-循环 $(a_1\ \cdots\ a_\ell)$ 证 $\operatorname{sgn}((a_1\ \cdots\ a_\ell)) = (-1)^{\ell-1}$. 容易验证 $\ell \ge 2$ 时

$$(a_1\ \cdots\ a_\ell) = (a_1\ a_\ell)(a_1\ \cdots\ a_{\ell-1}),$$

由此递归地推导 $\operatorname{sgn}((a_1\ \cdots\ a_\ell)) = (-1)\cdot\operatorname{sgn}((a_1\ \cdots\ a_{\ell-1})) = \cdots = (-1)^{\ell-1}$.

对于 $\operatorname{ord}(\sigma)$ 的部分, 考虑 $k \in \mathbb{Z}$, 则 $\xi_1^k, \dots, \xi_m^k$ 各自在无交的子集 $C_i \subset \{1, \dots, n\}$ 上作用, 因此

$$\sigma^k = \xi_1^k \cdots \xi_m^k = \mathrm{id} \iff \forall i,\ \xi_i^k = \mathrm{id}.$$

问题化为证 $\operatorname{ord}(\xi_i) = \ell_i$, 这正是简单的练习 11.7.2. □

例 11.7.7 练习 11.1.23 介绍的子群 $V \subset \mathfrak{S}_4$ 用循环分解简洁地表为

$$V = \{\mathrm{id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\};$$

除了 id, 所有元素的阶数都是 2. 事实上, $V \smallsetminus \{\mathrm{id}\}$ 穷尽了 $\mathfrak{S}_4$ 中所有循环分解形如 $(\star\star)(\star\star)$ 的置换.

运用命题 11.7.6 可见 V 的所有元素都是偶置换, 亦即 $V \subset \mathfrak{A}_4$; 这是真子群, 因为 $(1\ 2\ 3) \in \mathfrak{A}_4 \smallsetminus V$.

练习 11.7.8 (共轭类) 说明在任何群 G 上, 可以按照 $x \sim y \iff \exists g,\ y = gxg^{-1}$ 定义等价关系 $\sim$, 称为**共轭**, 对应的等价类称为**共轭类**. 接着考虑特例 $G = \mathfrak{S}_X$, 其中 X 是任意集合. 说明对任意 $\tau \in \mathfrak{S}_X$ 皆有

$$\tau(a_1 \cdots a_m)\tau^{-1} = (\tau(a_1)\ \cdots\ \tau(a_m)).$$

由此对所有 $n \in \mathbb{Z}_{\ge 1}$ 分类 $\mathfrak{S}_n$ 中的共轭类.

11.8 回首高次方程

回到本书开头探讨过的高次方程求解问题. 将所论的多项式表达为

$$f = X^n + a_{n-1}X^{n-1} + \cdots + a_0 = \prod_{k=1}^{n}(X - x_k),$$

其中 $a_0, \ldots, a_{n-1}$ 属于选定的域 F, 而 $x_1, \ldots, x_n$ 是所求的根.

对于 $n \leq 4$ 或某些特定的高次情形, 一切基于代数学的公式解法都涉及另一个依赖于 f 的多项式 $g \in F[X]$, 其根 $y_1, \ldots, y_m$ 可以化到次数较低的情形求解, 然后再设法将 $x_1, \ldots, x_n$ 用 $a_0, \ldots, a_n$ 和 $y_1, \ldots, y_m$ 代数地表示. 根的置换在每一步都扮演关键角色.

由于这些论证总会涉及一些单位根, 和以正整数为分母的除法运算, 为了简化陈述, 本节在复数域 $\mathbb{C}$ 上操作.

我们先介绍 Lagrange 的技术, 这可以对低次方程的公式解提供统一的解释. 考虑满足 $\omega^n = 1$ 的 $\omega \in \mathbb{C}^\times$. 存在 k 使得 $\omega = \zeta_n^k$, 其中

$$\zeta_n := e^{2\pi \mathrm{i}/n} = \cos\left(\frac{2\pi}{n}\right) + \mathrm{i}\sin\left(\frac{2\pi}{n}\right);$$

这只和 k 的同余类 $k + n\mathbb{Z}$ 相关, 而且 $k + n\mathbb{Z}$ 由 ω 唯一确定.

对于 ω 和 $\sigma \in \mathfrak{S}_n$, 定义相应的 **Lagrange 预解式**为

$$t(\omega, \sigma) := x_{\sigma(1)} + \omega x_{\sigma(2)} + \cdots + \omega^{n-1}x_{\sigma(n)}. \tag{11.8.1}$$

引理 11.8.1 设 $1 \leq k \leq n$ 而 $\sigma \in \mathfrak{S}_n$, 则

$$x_{\sigma(k)} = \frac{1}{n}\sum_{\substack{\omega \in \mathbb{C}^\times \\ \omega^n = 1}} \omega^{-k+1} t(\omega, \sigma).$$

证明 对于每个 $1 \leq h \leq n$, 右式中 $x_{\sigma(h)}$ 的系数是 $\frac{1}{n}\sum_{\omega^n=1}\omega^{h-k}$. 对 $\sum_{\omega^n=1}\omega^{h-k}$ 按两种顺序进行求和:

$$\omega = 1, \zeta_n, \ldots, \zeta_n^{n-1}, \quad \omega = \zeta_n, \ldots, \zeta_n^{n-1}, \underbrace{\zeta_n^n}_{=1},$$

两者差一个比例 ζ_n^{h-k}. 故当 $h \not\equiv k \pmod{n}$ 时和必为 0, 而 $h \equiv k \pmod{n}$ 时和显然是 n. □

这表明只要对所有 ω 和 σ 求出 $t(\omega, \sigma)$, 便容易反解 f 的根 $x_1, \ldots, x_n$.

引理 11.8.2 考虑 n-循环 $\xi := (1\ 2\ \cdots\ n) \in \mathfrak{S}_n$, 则

$$t(\omega, \sigma\xi) = \omega^{-1} t(\omega, \sigma).$$

证明 根据 ξ 的映法, 我们将 (11.8.1) 与

$$t(\omega, \sigma\xi) = \omega^{n-1} x_{\sigma(1)} + x_{\sigma(2)} + \cdots + \omega^{n-2} x_{\sigma(n)}.$$

作比较, 其余是 $\omega^n = 1$ 的立即结论. □

以下按照 ω 在群 $\mathbb{C}^\times$ 中的阶数作讨论. 设 $h \mid n$ 而 $\mathrm{ord}(\omega) = h$; 换言之, ω 是 h 次单位原根. 将 $\{1, \dots, n\}$ 分成 $C_1, \dots, C_h$ 的无交并, 其中

$$C_k := \left\{ k, h+k, \dots, h\left(\frac{n}{h} - 1\right) + k \right\}, \quad |C_k| = \frac{n}{h}.$$

它们在 ξ 之下的像容易描述:

$$\xi(C_1) = C_2, \quad \dots, \quad \xi(C_{h-1}) = C_h, \quad \xi(C_h) = C_1.$$

实际上, 这些子集正是 $\{1, \dots, n\}$ 在 $\langle \xi^h \rangle$ 作用下的轨道.

引理 11.8.3 设 $h := \mathrm{ord}(\omega) \mid n$ 如上. 命

$$K_h := \{ \tau \in \mathfrak{S}_n : \forall k,\ \tau(C_k) = C_k \},$$

则 K_h 是 $\mathfrak{S}_n$ 的子群, $|K_h| = ((n/h)!)^h$, 而且

$$t(\omega, \sigma\tau) = t(\omega, \sigma), \quad \sigma \in \mathfrak{S}_n,\ \tau \in K_h.$$

证明 关于 K_h 成子群的论断以及 $|K_h|$ 的计算是容易的, 事实上 $K_h \simeq (\mathfrak{S}_{n/h})^h$. 若 i, j 属于同一个 C_k, 则因为 $\omega^h = 1$, 根 $x_{\sigma(i)}$ 和 $x_{\sigma(j)}$ 在 $t(\omega, \sigma)$ 中的系数相同, 这导致 $\tau \in K_h$ 时 $t(\omega, \sigma\tau) = t(\omega, \sigma)$. □

因此 Lagrange 预解式 $t(\omega, \sigma)$ 只依赖 σ 的陪集 σK_h.

练习 11.8.4 验证 $\xi K_h \xi^{-1} = K_h$. 作为推论, 对于 $\sigma, \sigma' \in \mathfrak{S}_n$:

(i) 陪集 $\sigma\xi K_h$ 由 σK_h 确定; 〉提示〈 $\sigma\xi K_h = \sigma K_h \xi$

(ii) $\sigma\xi K_h = \sigma'\xi K_h$ 当且仅当 $\sigma K_h = \sigma' K_h$. 〉提示〈 消去律

继续设 $h := \mathrm{ord}(\omega) \mid n$. 为了解 $t(\sigma, \omega)$, 我们按引理 11.8.3 构造

$$P_\omega := \prod_{\sigma K_h \in \mathfrak{S}_n / K_h} (X - t(\omega, \sigma)), \quad \deg P_\omega = \frac{n!}{((n/h)!)^h}.$$

暂且将 $x_1, \dots, x_n$ 看作独立的变元. 连乘积中的 $t(\omega, \sigma)$ 相当于将它们等分成 h 堆, 按系数 $1, \dots, \omega^{h-1}$ 加总, 再将 $X - t(\omega, \sigma)$ 按所有等分方式相乘便是 P_ω. 由此可见 P_ω

不依赖 $x_1,\dots,x_n$ 的排序, 故对称多项式基本定理说明 P_ω 的系数能以 $x_1,\dots,x_n$ 的初等对称多项式, 亦即 $a_0,\dots,a_{n-1}$ 来表达, 系数依赖于 ω.

基于练习 11.8.4, P_ω 又可以拆解为若干个形如

$$\prod_{a=0}^{h-1}(X-t(\omega,\sigma\xi^a)) \overset{\text{引理 11.8.2}}{=\!=} \prod_{a=0}^{h-1}\big(X-\omega^{-a}t(\omega,\sigma)\big) = X^h - t(\omega,\sigma)^h$$

的多项式之积. 综上, P_ω 是 X^h 的 $\frac{n!}{((n/h)!)^h h}$ 次多项式; 特别地, $h=n$ 时 P_ω 是 X^n 的 $(n-1)!$ 次多项式; $h=1$ 时 $\deg P_\omega = 1$, 事实上此时 $t(1,\sigma) = -a_{n-1}$.

这些观察为求解的问题带来一丝希望. 回忆到我们希望对所有 ω 和 σ 求解 $t(\omega,\sigma)$.

▷ $n=2$ 对应的 $h := \operatorname{ord}(\omega)$ 或者是 1, 或者 $h=2$ 而 $t(-1,\sigma)^2 = (x_1-x_2)^2$ 即熟悉的判别式, 这给出二次方程的公式解.

▷ $n=3$ 对应的 h 或者是 1, 又或者 $h=3$, 此时 $t(\omega,\sigma)^3$ 能通过一个 2 次方程求解. 请读者对 $a_2=0$ 的情形验证此即三次方程的 Cardano 公式.

▷ $n=4$ 对应的 h 或者是 1, 或者 $h=2$, 此时 $t(-1,\sigma)^2$ 能通过一个 $\frac{4!}{2^2 2}=3$ 次方程来求解, 又或者 $h=4$, 此时 $t(\omega,\sigma)^4$ 涉及一个 $3!=6$ 次方程. 然而从 $t(1,\sigma)$ 和 $t(\pm1,\sigma)$ 已经足以求根, 缘由是线性方程组

$$\begin{aligned} t(1,\mathrm{id}) &= x_1+x_2+x_3+x_4,\\ t(-1,\mathrm{id}) &= x_1-x_2+x_3-x_4,\\ t(-1,(1\ 2\ 3)) &= x_1+x_2-x_3-x_4,\\ t(-1,(3\ 4)) &= x_1-x_2-x_3+x_4; \end{aligned} \tag{11.8.2}$$

有唯一解 $(x_1,\dots,x_4)$, 例如 x_1 和 x_3 来自

$$\begin{aligned} x_1+x_3 &= \frac{1}{2}(t(1,\mathrm{id})+t(-1,\mathrm{id})),\\ x_1-x_3 &= \frac{1}{2}(t(-1,(1\ 2\ 3)))+t(-1,(3\ 4)), \end{aligned}$$

如将右式的相加改成相减, 便能用类似方法反解 x_2 和 x_4. 于是我们得到 4 次方程的公式解.

对于 $n=4$ 的情形, Lagrange 预解式并非唯一进路. 另一种方法是引入以下三次多项式

$$Q := (X-\underbrace{(x_1x_2+x_3x_4)}_{=:t})(X-\underbrace{(x_1x_3+x_2x_4)}_{=:t'})(X-\underbrace{(x_1x_4+x_2x_3)}_{=:t''}).$$

容易看出 Q 和 $x_1,\dots,x_4$ 的排序无关, 故其系数能以 $a_0,\dots,a_3$ 来表达. 基于 Cardano 公式, t, t' 和 t'' 有公式解, 它们承担预解式的功能.

接着反解 $\alpha := x_1 + x_2$ 和 $\beta := x_3 + x_4$. 注意到 $\alpha + \beta = -a_3$ 而 $\alpha\beta = t' + t''$, 所以这相当于解二次方程.

同理可解 $x_1 + x_3$, $x_2 + x_4$, $x_1 + x_4$ 和 $x_2 + x_3$, 情况完全是对称的. 由此解线性方程组易得 $x_1, \dots, x_4$. 最终产物是 L. Ferrari 对四次方程的求根公式.

练习 11.8.5 回忆到 $\mathfrak{S}_4$ 左作用在 $\mathbb{C}[X_1, \dots, X_4]$ 上. 试明确 $X_1X_2 + X_3X_4$ 的稳定化子, 说明它是 $\mathfrak{S}_4$ 的 8 阶子群, 由此解释 Q 的性质.

Lagrange 尝试将此思路拓及更高次的方程. 然而 $n \geq 5$ 时奇迹不再出现. 虽然一种方法的失败未必蕴涵原问题无解, 但是这至少提示了一条线索: 相较于低次情形, 五次或超过五次的方程的求解具有本质的困难, 能否求解取决于根 $x_1, \dots, x_n$ 的置换, 与对称群 $\mathfrak{S}_n$ 的结构紧密相关.

问题的完整解答要留待 Galois 的工作, 域的自同构将替代根的置换, 其现代表述涉及扩域的自同构群, 以及它们的正规子群; 这类子群在上述讨论中对应到预解式的对称性. 这些理论当属 Galois 理论的内容, 读者可以参考 [10, §9.7]; 值得一提的是 Lagrange 预解式在一般理论中仍有用处.

就群论本身的观点, 正规子群的概念也能由同态与商结构的研究自然引出, 这是下一节的主题.

11.9 正规子群与商群

在以下讨论中选定群 G, 其幺元 1_G 按惯例时常简记为 1. 对于任意子集 $A \subset G$ 和 $x, y \in G$, 我们采取自明的记法 $xAy^{-1} := \{xay^{-1} : a \in A\}$.

定义 11.9.1 *若 G 的子群 H 满足以下条件*

$$\forall g \in G,\ gHg^{-1} = H,$$

*则称 H 为 G 的***正规子群***, 也记为 $H \lhd G$.*

- ⋆ 正规性等价于要求 $gHg^{-1} \subset H$ 对所有 g 成立, 这是因为以 g^{-1} 代 g 可得 $g^{-1}Hg \subset H$, 亦即反向包含 $H \subset gHg^{-1}$.
- ⋆ 移项可知正规性也等价于对所有 g 皆有 $gH = Hg$, 或者按照上述方式放宽为 $gH \subset Hg$ (或 $Hg \subset gH$). 因此正规子群的陪集不必分左右.

我们再看一些初步例子.

- ⋆ 如果 G 是交换群, 则因为 $gHg^{-1} = gg^{-1}H = H$, 所有子群 H 皆正规.
- ⋆ 例 11.1.4 介绍的中心 Z_G 是正规子群, 道理类似: $gZ_Gg^{-1} = gg^{-1}Z_G = Z_G$.
- ⋆ 如果 $N \lhd G$ 而 H 是 G 的任意子群, 则可按定义直接检验 $N \cap H \lhd H$.

⋆ 若 $(G:H)=2$, 则 H 必然正规. 为了说明这点, 对 g 分两种情况讨论:

- 若 $g\in H$, 则 $Hg=H=gH$;
- 若 $g\notin H$, 则 $Hg\neq H$, 故 $(G:H)=2$ 导致陪集分解写作 $G=H\sqcup Hg$, 从而 $Hg=G\smallsetminus H$, 但同样理由也导致 $gH=G\smallsetminus H$, 故 $Hg=gH$.

⋆ 最后, G 总有当然的正规子群 $\{1\}$ 和 G.

定义 11.9.2 *若 G 不是平凡群, 而且 G 没有除 $\{1\}$ 和 G 之外的正规子群, 则称 G 为* **单群**.

单群的研究是群论中饶富兴味的一类问题. 除了以下初步例子, 本章习题将有更多讨论.

例 11.9.3 若 p 是素数, $|G|=p$, 则 G 必然是单群, 这是因为任何子群 H 的元素个数必整除 p, 故只能是 1 (对应 $H=\{1\}$) 或 p (对应 $H=G$).

事实上, $|G|=p$ 蕴涵 $G\simeq\mathbb{Z}/p\mathbb{Z}$, 这是因为素数阶的群必是循环群 (推论 11.4.11).

正规子群的特色之一是可以和任意子群相乘给出更大的子群. 对任意子集 $A,B\subset G$ 引入记号 $AB:=\{ab\in G:a\in A,\ b\in B\}$.

引理 11.9.4 *设 H 和 K 为 G 的子群, 则 $HK=KH$ 当且仅当 HK 为 G 的子群.*

证明 先说明 $HK=KH$ 蕴涵 HK 为子群. 显然 $1=1\cdot 1\in HK$.

其次, 若 $h_1,h_2\in H$ 而 $k_1,k_2\in K$, 可取 $h'\in H$ 和 $k'\in K$ 使得 $k_1h_2=h'k'$, 故 $h_1k_1h_2k_2=h_1h'k'k_2\in HK$, 乘法封闭性得证.

设 $h\in H$ 而 $k\in K$, 则 $(hk)^{-1}=k^{-1}h^{-1}\in KH=HK$, 取逆封闭性得证.

接着说明 HK 为子群蕴涵 $HK=KH$. 对任意子集 $A\subset G$ 沿用稍早的记号 $A^{-1}:=\{a^{-1}:a\in A\}$, 则 $HK=(HK)^{-1}=K^{-1}H^{-1}=KH$. □

命题 11.9.5 *设 H 为 G 的子群, K 为 G 的正规子群, 则 $HK=KH$, 而且 HK 是 G 的子群.*

证明 对 $h\in H$ 和 $k\in K$, 由 K 的正规性可得 $hk=(hkh^{-1})h\in KH$ 和 $kh=h(h^{-1}kh)\in HK$, 故 $HK=KH$. 关于 HK 构成子群的断言来自引理 11.9.4. □

对任意子群 $K\subset G$, 定义其**中心化子**为 G 的子群

$$Z_G(K):=\left\{g\in G:\forall x\in K,\ gxg^{-1}=x\right\},$$

定义其**正规化子**为 G 的子群

$$N_G(K):=\left\{g\in G:gKg^{-1}=K\right\}.$$

易见 $K\lhd N_G(K)\supset Z_G(K)$, 而 $K\lhd G$ 等价于 $N_G(K)=G$. 命题 11.9.5 关于 HK 成群的充分条件可以放宽为 $H\subset N_G(K)$, 论证无异.

练习 11.9.6 对所有子群 $H,K\subset G$, 证明 $|HK|\cdot|H\cap K|=|H|\cdot|K|$; 此式可以理解为基数的等式, 但也无妨假设 HK 是有限集, 将其视为正整数的等式.

提示〉 群的乘法给出满射 $m:H\times K\twoheadrightarrow HK$. 对所有 $x=hk\in HK$, 试给出从 $m^{-1}(x)$ 到 $H\cap K$ 的双射.

练习 11.9.7 验证练习 11.1.23 中的群 V 满足 $V\lhd\mathfrak{S}_4$. 由此说明 $\mathfrak{S}_4$ 和 $\mathfrak{A}_4$ 皆非单群.
提示〉 参考例 11.7.7 以及其后的练习.

正规子群的典型例子是同态的核.

定义–命题 11.9.8 (群同态的核) *设 $f:G\to G'$ 为群同态, 记*

$$\ker(f):=\{g\in G:f(g)=1_{G'}\},$$

*则 $\ker(f)\lhd G$, 称之为群同态 f 的***核**.

证明 首先说明 $\ker(f)$ 是 G 的子群. 由 $f(1_G)=1_{G'}$ 立见 $1_G\in\ker(f)$. 若 $x,y\in\ker(f)$, 则 $f(xy)=f(x)f(y)=1_{G'}$ 导致 $xy\in\ker(f)$. 若 $x\in\ker(f)$, 则 $f(x^{-1})=f(x)^{-1}=1_{G'}$ 导致 $x^{-1}\in\ker(f)$. 综上知 $\ker(f)$ 为子群.

接着说明 $\ker(f)$ 正规. 设 $x\in\ker(f)$ 而 $g\in G$, 则 $f(gxg^{-1})=f(g)f(x)f(g)^{-1}=f(g)f(g)^{-1}=1_{G'}$, 故 $gxg^{-1}\in\ker(f)$. 这相当于验证了 $g\ker(f)g^{-1}\subset\ker(f)$ 恒成立, 故 $\ker(f)$ 正规. □

同态 f 的核能够描述 f 的纤维, 它们是核的陪集.

命题 11.9.9 *设 $f:G\to G'$ 为群同态, $N:=\ker(f)$. 若 $x,y\in G$, 则 $f(x)=f(y)$ 当且仅当 $xN=yN$. 特别地, f 是单同态当且仅当 $N=\{1_G\}$.*

证明 我们有 $f(x)=f(y)$ 当且仅当 $f(xy^{-1})=1_{G'}$, 当且仅当 $xy^{-1}\in N$; 但这又相当于说 $x\in yN$, 亦即 $xN=yN$; 此处用到陪集无交的性质. 关于单同态的刻画是 $|xN|=|N|$ 的直接结论. □

本节的后续主题是说明如何对给定的 $N\lhd G$ 在集合 $G/N=\{\text{陪集 }xN:x\in G\}$ 上赋予自然的群结构, 使得商映射 $q:G\to G/N$ 为同态, 这是群论中最基础的构造之一.

定义–命题 11.9.10 (商群) *设 $N\lhd G$. 在 G/N 上定义记作乘法的二元运算*

$$xN\cdot yN=xyN,$$

这只与陪集 $xN,yN\in G/N$ 相关, 无关元素 x,y 的选取.

(i) *此运算使得 G/N 成群, 其幺元和取逆运算分别是*

$$1_{G/N}=N=1_GN,\quad (xN)^{-1}=x^{-1}N.$$

(ii) *映 x 为 xN 的商映射 $q:G\to G/N$ 是群同态, 而且 $\ker(q)=N$.*

群 G/N 称为 G 对 N 的商群.

证明 首先说明 xyN 只和 xN, yN 有关. 设 $xN = x'N$, $yN = y'N$, 则存在 $u, v \in N$ 使得 $x' = xu$ 而 $y' = yv$, 故确实有

$$x'y'N = xuyvN = xy\underbrace{y^{-1}uy}_{\in N}vN = xyN.$$

接着说明 G/N 对此成群. 乘法的结合律化约为 G 的结合律:

$$(xNyN)(zN) = (xyN)(zN) = xyzN = (xN)(yzN) = (xN)(yNzN).$$

关于 $1_GN = N$ 的幺元性质也化到 G 上检验:

$$(1_GN)(xN) = (1_Gx)N = xN = (x\,1_G)N = (xN)(1_GN).$$

关于 $x^{-1}N$ 的逆元性质同样化到 G 上, 不必重复.

最后, 商映射 q 之所以为同态不外是乘法定义 $(xN)(yN) = xyN$ 的改述, 因为左式为 $q(x)q(y)$ 而右式为 $q(xy)$; 核 $\ker(q)$ 的描述归结为 $xN = N \iff x \in N$. □

上述论证实际说明了 G/N 上的群结构是使商映射 $q : G \to G/N$ 成同态的唯一选择. 严格来说, q 也是商群结构的一员, 单提 G/N 还不足以彰显它作为商群的角色.

例 11.9.11 设 R 为环而 I 为理想, 因此 I 也是加法群 $(R, +)$ 的正规子群. 环论中的商环 R/I 事实上是搭建在加法商群 $(R/I, +)$ 上的乘法构造.

与此类似, 向量空间 V 对子空间 U 的商空间 V/U 是搭建在加法商群 $(V/U, +)$ 上的纯量乘法构造. 它们都是商群构造的分化.

命题 11.9.12 *设 $f : G \to G'$ 为群同态, 而 $N \lhd G$ 满足 $N \subset \ker(f)$, 则存在唯一的同态 $\overline{f} : G/N \to G'$ 使得 $f = \overline{f}q$, 或以交换图表写作*

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ {\scriptstyle q}\downarrow & \nearrow_{\overline{f}} & \\ G/N & & \end{array} \qquad \text{交换.}$$

这般的 $\overline{f}$ 称为 f 所诱导的同态.

证明 条件 $f = \overline{f}q$ 相当于要求 $\overline{f}(xN) = f(x)$, 因此这般的 $\overline{f}$ 若存在则唯一. 问题在于存在性.

设 $xN = x'N$, 亦即存在 $u \in N$ 使得 $x' = xu$, 则 $N \subset \ker(f)$ 蕴涵 $f(x') = f(x)f(u) = f(x)$. 这就说明 $\overline{f}(xN) := f(x)$ 确实给出良定义的映射 $G/N \to G'$. 此外,

$$\begin{aligned} \overline{f}(xN\,yN) &= \overline{f}(xyN) = f(xy) \\ &= f(x)f(y) = \overline{f}(xN)\overline{f}(yN), \end{aligned}$$

故 $\overline{f}$ 是同态. 明所欲证. □

留意到诱导同态的刻画 $f = \overline{f}q$ 也蕴涵 $\mathrm{im}(\overline{f}) = \mathrm{im}(f)$.

推论 11.9.13 设 $f : G \to G'$ 为群同态, $N \lhd G$, $N' \lhd G'$ 而 $f(N) \subset N'$, 则存在唯一的群同态 $\overline{f} : G/N \to G'/N'$ 使下图交换

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q'} \\ G/N & \xrightarrow[\overline{f}]{} & G'/N' \end{array}$$

此处 q 和 q' 代表商同态.

证明 图表交换相当于说 $\overline{f}q = q'f$. 考虑同态 $q'f : G \to G'/N'$, 条件 $f(N) \subset N'$ 导致 $q'f(N) = \{1\}$, 故命题 11.9.12 确保所求之 $\overline{f}$ 存在且唯一. □

推论中的 $\overline{f}$ 同样称为 f 所诱导的同态, 而命题 11.9.12 的版本则归约为 $N' = \{1\}$ 的特例.

对命题 11.9.12 代入特例 $N := \ker(f)$, 我们便得到诱导同态 $\overline{f} : G/\ker(f) \to G'$.

命题 11.9.14 设 $f : G \to G'$ 为群同态, 则诱导同态 $\overline{f} : G/\ker(f) \to G'$ 给出群同构 $G/\ker(f) \xrightarrow{\sim} \mathrm{im}(f)$. 作为推论, $(G : \ker(f)) = |\mathrm{im}(f)|$.

证明 记 $N := \ker(f)$. 只要说明 $\ker(\overline{f}) = \{1\}$, 则命题 11.2.8 便蕴涵 $\overline{f} : G/N \xrightarrow{\sim} \mathrm{im}(\overline{f}) = \mathrm{im}(f)$. 验证不难: $\overline{f}(xN) = 1$ 相当于 $f(x) = 1$, 相当于 $x \in N$, 然而这又等价于 $xN = 1_{G/N}$. □

因此, 商群的构造为满同态 $f : G \to G'$ 给出了一种只涉及 G 的内在描述. 任何满同态本质上都是取商.

例 11.9.15 考虑同态 $\mathrm{sgn} : \mathfrak{S} \to \{\pm 1\}$, 它在 $n \geq 2$ 时是满的, 而 $\ker(\mathrm{sgn}) = \mathfrak{A}_n$, 故此时 $(\mathfrak{S}_n : \mathfrak{A}_n) = 2$. 这表明 $n \geq 2$ 时 $\mathfrak{S}_n$ 中的奇偶置换各占一半.

下一则例子涉及例 11.1.12 与例 11.1.14 介绍的正交群和酉群, 以及 §10.8 介绍的四元数.

例 11.9.16 全体满足 $\mathrm{N}(q) = 1$ 的四元数 q 对乘法构成群 $\mathbb{H}^1$. 定理 10.8.9 和其上的讨论给出满的群同态

$$R : \mathbb{H}^1 \twoheadrightarrow \mathrm{SO}(3), \quad \ker(R) = \{\pm 1\}.$$

另一方面, 简单的练习 10.8.11 给出群同构 $\Phi : \mathbb{H}^1 \xrightarrow{\sim} \mathrm{SU}(2)$, 使得 $\Phi(-1) = -\mathbf{1}_{2\times 2}$. 综上, 命题 11.9.14 给出

$$\mathrm{SU}(2)/\{\pm\mathbf{1}_{2\times 2}\} \xleftarrow[\text{由 } \Phi \text{ 诱导}]{\sim} \mathbb{H}^1/\{\pm 1\} \xrightarrow[\text{由 } R \text{ 诱导}]{\sim} \mathrm{SO}(3).$$

下一步是探讨满同态两边的子群有何关联.

命题 11.9.17 设 $f:G\to G'$ 为满同态, 则有双射

$$\begin{array}{ccc}\{\text{子群 } H'\subset G'\} & \xleftrightarrow{1:1} & \{\text{子群 } H\subset G : H\supset \ker(f)\}\\ \cup & & \cup\\ \{\text{正规子群 } H'\lhd G'\} & \xleftrightarrow{1:1} & \{\text{正规子群 } H\lhd G : H\supset \ker(f)\}\end{array}$$

$$H' \longmapsto f^{-1}(H')$$

$$f(H) \longmapsfrom H.$$

此双射满足以下性质:

- $H_1'\subset H_2' \iff f^{-1}(H_1')\subset f^{-1}(H_2')$,
- 若 $H'\lhd G'$ 对应到 $H\lhd G$, 则合成同态 $G\xrightarrow{f} G'\xrightarrow{\text{商}} G'/H'$ 诱导出群同构 $G/H\xrightarrow{\sim} G'/H'$,
- 当 f 取为对某个 $N\lhd G$ 的商同态 $G\to G/N$ 时, 上述群同构进一步改写为“分母相消”的形式
$$G/H\xrightarrow{\sim}(G/N)\big/(H/N),$$
其中要求 $N\subset H\lhd G$.

证明 包含关系 $H\subset f^{-1}(f(H))$ 不需要任何前提, 而从 $H\supset\ker(f)$ 容易推得反向的 $\supset$, 故 $H=f^{-1}(f(H))$; 另一方面, $H'=f(f^{-1}(H'))$ 则是满射的一般性质. 综上得到互逆双射. 显然 f 和 f^{-1} 都保持子群的包含关系.

关于正规子群的对应则是因为 $f(gHg^{-1})=f(g)f(H)f(g)^{-1}$ 而 f 满, 故上述双射导致

$$\forall g\in G,\ gHg^{-1}=H \iff \forall \bar g\in G',\ \bar g f(H)\bar g^{-1}=f(H).$$

对于 $H'\lhd G'$, 同构 $G/H\xrightarrow{\sim}G'/H'$ 可以从命题 11.9.14 推导, 这是因为 $G\twoheadrightarrow G'\twoheadrightarrow G'/H'$ 合成依然满, 而它的核显然是 $H:=f^{-1}(H')$. 事实上, 同构具体写为 $gH\mapsto f(g)H'$.

最后, 考虑 $N\lhd G$ 和商同态 $f:G\to G/N=:G'$, 相应地取 H 为 G 中包含 N 的子群, 此时 $N\lhd H$. 对应于 H 有 $H':=f(H)=H/N$. 上一段得到的同构因而改写为 $G/H\xrightarrow{\sim}(G/N)\big/(H/N)$. □

例 11.9.18 有限循环群 $\mathbb{Z}/n\mathbb{Z}$ 不外是无穷循环群 $\mathbb{Z}$ 对子群 $n\mathbb{Z}$ 的商群. 注记 11.1.19 已说明 $\mathbb{Z}$ 的子群皆形如 $m\mathbb{Z}$, 将此代入命题 11.9.17 (取 f 为商同态 $q:\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$), 便可知 $\mathbb{Z}/n\mathbb{Z}$ 的子群皆形如

$$m\mathbb{Z}/n\mathbb{Z}=q(m\mathbb{Z}),$$

其中要求 $m\mathbb{Z}\supset\ker(q)=n\mathbb{Z}$, 亦即 $m\mid n$. 循环群的子群依此完全确定.

接着研究子群 $H,K\subset G$ 的乘积 HK 和商的关系; 为了确保 HK 成群, 要求 $H\subset N_G(K)$.

命题 11.9.19 设 $H,K\subset G$ 为子群, $H\subset N_G(K)$, 则 $K\lhd HK$, $H\cap K\lhd H$, 而且此时有群同构

$$H/(H\cap K)\xrightarrow{\sim} HK/K$$

$$h(H\cap K)\longmapsto hK$$

其中 $h\in H$.

证明 以下均假定 $h\in H$, $k\in K$. 从 $H\subset N_G(K)$ 可得 $(hk)K(hk)^{-1}=hkKk^{-1}h^{-1}=hKh^{-1}=K$, 故 $K\lhd HK$. 另一方面, 若 $x\in H\cap K$, 则 hxh^{-1} 一方面是子群 H 中的乘积, 故仍属于 H, 另一方面它又属于 K, 综上得 $H\cap K\lhd H$.

接着处理群同构. 定义同态 $f:H\to HK/K$ 为包含同态 $H\hookrightarrow HK$ 与商同态 $HK\twoheadrightarrow HK/K$ 的合成; 具体地说 $f(h)=hK$.

观察到 HK/K 的元素都可以表成 $hkK=hK$ 的形式, 故 f 满. 此外还有 $\ker(f)=H\cap K$, 这是因为 $f(h)=1\iff hK=K\iff h\in H\cap K$.

代入命题 11.9.14 即知 f 诱导同构 $\overline{f}:H/(H\cap K)\xrightarrow{\sim}\mathrm{im}(f)=HK/K$; 然而 $\overline{f}$ 按定义映 $h(H\cap K)$ 为 $f(h)=hK$, 此即断言中的映法. □

11.10 群的半直积

我们在 §11.1 介绍过如何定义两个群 G_1 和 G_2 的直积 $G_1\times G_2$, 这是搭建在积集上的群结构. 环和向量空间等代数结构也有类似构造, 但对于群的情形, 直积另有一种由自同构 “扭曲” 的版本, 这是从既有的群构造新群, 或将大群分解为小群的重要手段.

回忆到任意群 N 的自同构成群 $\mathrm{Aut}(N)$, 以同构的合成为乘法运算.

定义–命题 11.10.1 (半直积) 设 H 和 N 为群, $\varphi:H\to\mathrm{Aut}(N)$ 为群同态; 记 $h\in H$ 对 φ 的像为 $\varphi_h:N\to N$. 在积集 $N\times H$ 上定义二元运算

$$(n,h)(n',h'):=(n\varphi_h(n'),hh').$$

这给出群结构, 称为 H 和 N 相对于 φ 的半直积, 记为 $N\rtimes_\varphi H$. 它满足

$$1_{N\rtimes H}=(1_N,1_H),\quad (n,h)^{-1}=\left(\varphi_{h^{-1}}(n^{-1}),h^{-1}\right).$$

群 N 和 H 分别通过 $n\mapsto(n,1_H)$ 和 $h\mapsto(1_N,h)$ 嵌入为 $N\rtimes H$ 的子群. 进一步, $N\lhd N\rtimes_\varphi H$; 事实上,

$$(1_N,h)(n,1_H)(1_N,h)^{-1}=(\varphi_h(n),1_H).$$

证明 先说明定义的动机. 我们的思路是将 H 和 N 嵌入一个更大的群 G, 使得 $N \lhd G$ 而且 G 的所有元素都能唯一地表为 nh, 其中 $n \in N$ 而 $h \in H$. 熟悉的写法

$$nh \cdot n'h' = n \underbrace{hn'h^{-1}}_{\in N} hh'$$

表明只要能对所有 h 和 n' 描述 $\mathrm{Ad}_h(n') = hn'h^{-1}$ (练习 11.2.12 的符号), 则 G 的乘法便唯一从 N 和 H 的乘法确定. 定义中的 φ_h 正是此处的 $\mathrm{Ad}_h \in \mathrm{Aut}(N)$. 关于幺元和逆元的描述也都可以按此理解.

严格论证将反其道而行, 从 N, H 和 φ 构造这般的群 G, 而自然的思路是在积集 $N \times H$ 上建群.

现在开始严格证明. 首先是乘法结合律:

$$\begin{aligned}((n,h)(n',h'))(n'',h'') &= (n\varphi_h(n'), hh')(n'',h'') = (n\varphi_h(n')\varphi_{hh'}(n''), hh'h''),\\ (n,h)((n',h')(n'',h'')) &= (n,h)(n'\varphi_{h'}(n''), h'h'') = (n\varphi_h(n'\varphi_{h'}(n'')), hh'h'').\end{aligned}$$

问题化为证 $\varphi_h(n')\varphi_{hh'}(n'') = \varphi_h(n'\varphi_{h'}(n''))$; 因为 φ_h 是同态, 目标进一步化为 $\varphi_{hh'} = \varphi_h\varphi_{h'}$, 然而 $\varphi : H \to \mathrm{Aut}(N)$ 也是同态, 故结合律得证.

幺元 $(1_N, 1_H)$ 的性质容易归结为 $\varphi_{1_H} = \mathrm{id}_N$ 和 $\varphi_h(1_N) = 1_N$. 至于逆元的性质, 我们有

$$\begin{aligned}(n,h)\left(\varphi_{h^{-1}}(n^{-1}), h^{-1}\right) &= \left(n\varphi_h\left(\varphi_{h^{-1}}(n^{-1})\right), hh^{-1}\right)\\ &= (nn^{-1}, hh^{-1}) = (1_N, 1_H),\\ \left(\varphi_{h^{-1}}(n^{-1}), h^{-1}\right)(n,h) &= \left(\varphi_{h^{-1}}(n^{-1})\varphi_{h^{-1}}(n), h^{-1}h\right)\\ &= \left(\varphi_{h^{-1}}(n^{-1}n), h^{-1}h\right) = (1_N, 1_H).\end{aligned}$$

最后, 嵌入 $H \hookrightarrow N \rtimes_\varphi H$ 和 $N \hookrightarrow N \rtimes_\varphi H$ 的同态性质是毫无困难的. 按此计算 $(1_N,h)(n,1_H)(1_N,h)^{-1} = (\varphi_h(n),h)(1_N,h^{-1}) = (\varphi_h(n),1_H)$; 此处用到了 $\varphi_h(1_N) = 1_N$. □

作为特例, 如果取 $\varphi : H \to \mathrm{Aut}(N)$ 为平凡同态, 亦即 $\forall h$, $\varphi_h = \mathrm{id}_N$, 则 $N \rtimes H$ 便化为直积 $N \times H$.

留意到若将 N 和 H 等同于 $N \rtimes_\varphi H$ 的子群, 则它们的交为平凡子群, 而且 $N \rtimes_\varphi H = NH$; 这蕴涵 $N \rtimes_\varphi H$ 都能唯一地写成 nh 的形式.

练习 11.10.2 验证 $(n,h) \mapsto h$ 给出满同态 $N \rtimes_\varphi H \to H$, 而且它诱导同构 $(N \rtimes_\varphi H)/N \xrightarrow{\sim} H$. 〈提示〉 代入命题 11.9.14.

上述构造皆属 "外在": 我们从群 H 和 N 出发构造新群 $N \rtimes_\varphi H$, 两者的关系仅是给定的 φ, 然后将它们嵌入为 $N \rtimes_\varphi H$ 的子群. 以下探讨的则是 "内在" 视角: 给定群 G 及其子群 H, N, 我们寻求将 G 等同于半直积 $N \rtimes_\varphi H$ 的条件. 思路在定义–命题 11.10.1 证明的开头已经和盘托出了.

定义–命题 11.10.3 (半直积的内在版本) 设 H 和 N 为群 G 的子群, 满足下述条件

$$N \lhd G, \quad G = NH, \quad N \cap H = \{1\}.$$

考虑由 $\mathrm{Ad}_h(n) = hnh^{-1}$ 给出的同态 $\mathrm{Ad} : H \to \mathrm{Aut}(N)$, 则有群同构

$$\begin{aligned} \Phi : N \rtimes_{\mathrm{Ad}} H &\xrightarrow{\sim} G \\ (n, h) &\longmapsto nh. \end{aligned}$$

此时也称 G 是子群 H 和正规子群 N 的半直积, 合理地记为 $G = N \rtimes H$.

证明 验证 Φ 是群同态: $(n,h)(n',h') = (n\,\mathrm{Ad}_h(n'), hh')$ 被映为 $n\,\mathrm{Ad}_h(n')hh'$, 然而后者即 $nhn'h^{-1}hh' = (nh)(n'h')$.

条件 $G = NH$ 相当于说 Φ 满. 最后证明 Φ 单: 若 $\Phi(n,h) = 1$, 则 $n = h^{-1} \in N \cap H = \{1\}$. 因此 Φ 是同构. □

内在与外在的辩证与向量空间直和的讨论是类似的.

练习 11.10.4 设 $n \in \mathbb{Z}_{\geq 2}$. 任取对换 $\tau := (i\ j) \in \mathfrak{S}_n$.

(i) 说明 $\mathfrak{S}_n = \mathfrak{A}_n \rtimes \langle \tau \rangle$.

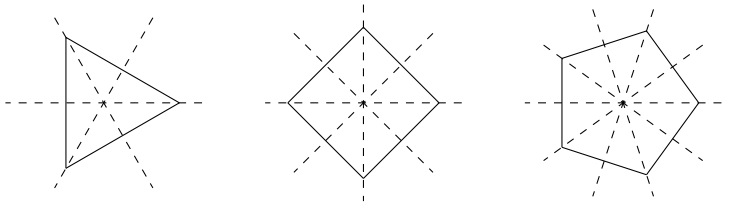
(ii) 说明当 $n \geq 3$ 时这并非直积.

例 11.10.5 (二面体群) 考虑平面上的正 n 边形 ($n \in \mathbb{Z}_{\geq 3}$). 为明确起见, 将平面 $\mathbb{R}^2$ 等同于 $\mathbb{C}$, 并且设正 n 边形的顶点为 $1, \zeta, \ldots, \zeta^{n-1}$, 其中 $\zeta := e^{2\pi \mathrm{i}/n}$.

保持正 n 边形不变的所有正交变换构成正交群 $\mathrm{O}(2)$ 的子群, 记为 D_{2n}. 为了理解 D_{2n} 的结构, 我们写下其中的两类元素.

▷ **旋转** 设 σ 为转角 $2\pi/n$ 的旋转, 它轮换正 n 边形的顶点, 或以复数表示为 $\sigma(\zeta^a) = \zeta^{a+1}$; 于是 σ 生成 D_{2n} 的循环子群 $\langle \sigma \rangle \simeq \mathbb{Z}/n\mathbb{Z}$.

▷ **镜射** 我们也可以对过原点的某类直线作镜射 (下图标为虚线), 使得正 n 边形保持不变:



$n = 3$ $n = 4$ $n = 5$

我们即将用代数工具说明 D_{2n} 共含 n 个镜射. 现阶段不妨先取 τ 为对水平轴的镜射, 亦即复共轭 $\mathbb{C} \to \mathbb{C}$, 于是 $\tau \in D_{2n}$ 而 $\mathrm{ord}(\tau) = 2$.

以下来说明

$$D_{2n} = \langle\sigma\rangle \rtimes \langle\tau\rangle .$$

首先, 考虑角度可知属于 D_{2n} 的所有旋转都能写成 σ^k 的形式. 由此推知 $\langle\sigma\rangle \lhd D_{2n}$: 元素 $g \in D_{2n}$ 是旋转相当于说 $\det g = 1$, 然而这也蕴涵 $\det(hgh^{-1}) = 1$ 对所有 $h \in D_{2n}$ 成立, 故 hgh^{-1} 仍是旋转.

其次, 由于 τ 非旋转, 条件 $\langle\sigma\rangle \cap \langle\tau\rangle = \{\mathrm{id}\}$ 成立.

剩下的任务是证明 D_{2n} 的元素 g 皆形如 σ^k 或 $\sigma^k\tau$. 设 $g(1) = \zeta^k$, 则以 $\sigma^{-k}g$ 代 g 可以假定 g 保持顶点 1 不动. 然而保持 1 不动的正交变换或者是恒等, 或者是对过 0 和 1 两点的直线作镜射, 后者即 τ. 验证完毕.

留意到 $\tau, \sigma\tau, \dots, \sigma^{n-1}\tau$ 正是 D_{2n} 包含的 n 个镜射. 讨论 $e^{\mathrm{i}\theta}$ 的像 $(0 \le \theta < 2\pi)$, 便能验证它们分别保持 $\pm 1, \pm e^{\pi\mathrm{i}/n}, \dots, \pm e^{(n-1)\pi\mathrm{i}/n}$ (不全是顶点) 不动.

容易看出 $\tau\sigma\tau^{-1} = \sigma^{-1}$, 所以半直积对应的同态 $\langle\tau\rangle \to \mathrm{Aut}(\langle\sigma\rangle)$ 映 τ 为 $\langle\sigma\rangle$ 的自同构 $\sigma^k \mapsto \sigma^{-k}$. 综上,

$$D_{2n} \simeq \mathbb{Z}/n\mathbb{Z} \rtimes_\varphi \mathbb{Z}/2\mathbb{Z},$$

其中 $\varphi : \mathbb{Z}/2\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/n\mathbb{Z})$ 映非零元为 $\mathbb{Z}/n\mathbb{Z}$ 的自同构 $x \mapsto -x$.

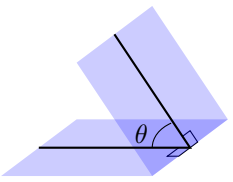
既然 φ 对所有 $n \in \mathbb{Z}_{\ge 0}$ 都有意义, 我们也可以纯代数地对所有 $n \ge 0$ 重新定义二面体群为 $\mathbb{Z}/n\mathbb{Z} \rtimes_\varphi \mathbb{Z}/2\mathbb{Z}$. 特别地, $D_4 = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

前一则例子阐明了平面上正 n 边形的对称性是群 D_{2n}. 至于空间中的正多面体, 其分类和对称性都需要更深入的讨论, 这是 §11.11 的任务.

11.11 正多面体的对称群

本节所谓 "空间" 特指三维空间, 建立坐标系后等同于 $\mathbb{R}^3$. 空间中的多面体均默认为凸多面体. 凸性与多面体的严谨定义分别是 §14.6 与 §14.7 的任务. 由于本节只考虑空间中的情形, 图像可以直观地把握, 这也是以下将采取的进路.

多面体的两个面或者交于一个共同顶点, 或者交于一整个棱, 对于后一情形, 我们称两面相邻, 此时两面之间的**二面角**如下图所示.



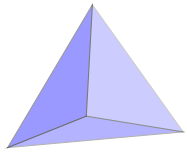
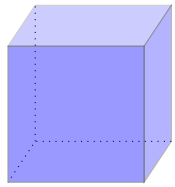
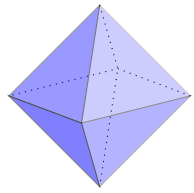
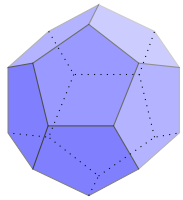
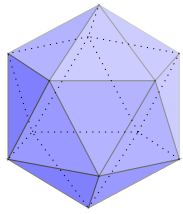
$(0 < \theta < \pi)$

定义 11.11.1 各个面都是相互全等的正多边形, 而且环绕每个顶点的面数都相同的有界多面体称为**正多面体**.

全等性质导致正多面体的每个棱都有相同的长度.

以上论及平面上多边形的全等. 如果两个多面体在空间中仅相差一个平移和绕某个轴的旋转, 则称为全等的. 从几何的视角, 全等的两个多面体应视为相同. 正多面体在此意义下的分类是一则古老而辉煌的定理.

定理 11.11.2 精确到全等, 棱长为 1 的正多面体分成以下五类.

名称	图像	顶点数	棱数	面数
正四面体		4	6	4
正立方体		8	12	6
正八面体		6	12	8
正十二面体		20	30	12
正二十面体		12	30	20

此外, 这些正多面体具有如下性质:

- ⋆ 每个正多面体的所有二面角皆相同;

⋆ 每个正多面体都内接于一个球面;

⋆ 对任两个顶点 x, y, 都存在一系列旋转 $R_1, \dots, R_k$, 转轴都通过球心, 使得 $T := R_1 \cdots R_k$ 映正多面体为自身, 而 $T(x) = y$.

证明 (勾勒)　若正多面体的每个面都是正 p 边形, 每个顶点都被 q 个面环绕, 则称此正多面体为 $\{p, q\}$ 型的 (所谓 Schläfli 符号). 当顶点给定, 每个面在该顶点处的内角为 $(1 - \frac{2}{p})\pi$. 将顶点处的 q 个角在平面上摊平, 加总后根据多面体的凸性得到 $q\left(1 - \frac{2}{p}\right)\pi < 2\pi$; 这也等价于 $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$, 又等价于 $(p-2)(q-2) < 4$. 穷举得

$$(p, q) \in \{(3,3), (4,3), (3,4), (3,5), (5,3)\}.$$

行文至此, 读者应已察觉 $\{p, q\}$ 决定正多面体在每个顶点周围的局部样貌; 既然正多面体的顶点个数有限, 这表明当棱长指定, $\{p, q\}$ 型正多面体若存在则是唯一的, 精确到全等. 所以关键是对上式列出的五种 $\{p, q\}$ 说明确实存在同型的正多面体, 而且它们符合断言中的性质. 顶点和棱的个数可由简单的计数 (两面共一棱, q 面共一顶点) 求得. 以下给出的构造是具体的, 但论证将省略一些细节[1)].

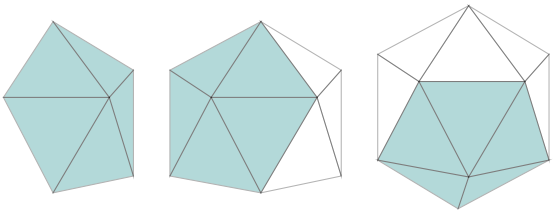
对应 $\{3, 3\}$ 的是正四面体, 构造如下. 先在平面上构造边长为 1 的正三角形 $\triangle ABC$, 从 $\triangle ABC$ 的重心向上作垂直于平面的直线; 在直线上取唯一的点 D 使得它到 A, B, C 的距离同为 1, 按此得到以 A, B, C, D 为顶点的正四面体. 旋转可见环绕 D 的三个面有相同二面角.

留意到如果从 $\triangle ACD$ 等为出发点, 构造产物完全相同, 由此便看出二面角全同. 空间中不共面的四点总落在唯一球面上. 以此球心为原点, 分别对过四个面的重心的轴作角度 $2\pi/3$ 的旋转, 即可按需搬运顶点.

对应 $\{4, 3\}$ 的是正立方体[2)], 构造和相应的性质都是熟知的.

对应 $\{3, 4\}$ 的是正八面体. 在给定的球中取三条相互正交的直径, 它们交球面于六点; 将这些点适当地连接, 得到八个正三角形, 围成正八面体. 考虑以这三条直径为轴, 角度为 $\pi/2$ 的旋转, 可见任两个顶点和任两个棱都能相互搬运; 这也蕴涵二面角全同.

对应 $\{3, 5\}$ 的是正二十面体. 直观的构造是先取边长为 1 的正五边形, 它内接于平面上的圆; 从圆心作平面的垂线, 然后在垂线上取点, 使得它和五边形的每个顶点距离皆为 1. 这就给出下图左侧由五个正三角形拼成的图像. 旋转对称性表明其二面角全同.



[1)] 理解正多面体构造的最佳方式是动手组装模型.

[2)] 或称正六面体.

重复相同构造, 然后利用二面角相等的性质, 可将这两个配件沿两个正三角形黏合, 如上图中间所示, 其中的二面角依然全同.

再构造相同的图像, 基于二面角的等同, 可按上图右侧方式与上一步的产物黏合, 得到 10 个正三角形构成的 "碗". 它的二面角全同, 而且碗沿 (六个棱) 的每个内角均等于正五边形的内角.

现在构造另一个碗. 基于二面角和碗沿内角的相等, 可以论证两者精准扣合, 给出正二十面体, 而且其二面角仍然全同. 从任两个相邻面的重心作垂线, 必有唯一交点 O, 它和两个面的 4 个顶点等距. 利用二面角相等的性质, 易推得 O 到所有顶点皆等距, 这就使得正二十面体内接于球.

留意到第一步造出的配件可以作角度为 $\frac{2\pi}{5}$ 的旋转, 相应的正二十面体保持不变. 由于对环绕每个顶点的五个面都可以如是旋转, 任两个面能通过以 O 为原点的一系列旋转相互搬运. 此外对每个面也能作角度为 $\frac{2\pi}{3}$ 的旋转, 保持正二十面体不变, 故任两个顶点也都能相互搬运.

对应 $\{5,3\}$ 的是正十二面体, 它可以从正二十面体构造如下[3)]. 对正二十面体的每个顶点, 将环绕它的五个正三角形的重心循序连接, 给出正五边形. 按此方法得到十二个正五边形, 其中每个顶点 (对应于原正二十面体的面) 都被三个面环绕 (对应于原面的三个顶点). 这些顶点依然和 O 等距, 故内接于较小的球面, 适当缩放可确保棱长为 1.

正二十面体的对称性蕴涵正十二面体的二面角全部相同. 用旋转挪动正二十面体的面相当于挪动正十二面体的顶点, 而由先前讨论可知任两个顶点皆可相互搬运. □

每个正多面体 P 所内接的球是唯一确定的, 其球心称为 P 的对称中心. 不妨将所论的空间等同于 $\mathbb{R}^3$, 使得对称中心即原点. 今后依循例 11.1.12 的符号, 考虑 $\mathrm{O}(3)$ 及其正规子群 $\mathrm{SO}(3)$ 在 $\mathbb{R}^3$ 上的作用.

定义 11.11.3 *设正多面体 P 的对称中心为原点. 定义 P 的对称群为*

$$\mathrm{symm}(P) := \{T \in \mathrm{O}(3) : T(P) = P\},$$

其旋转对称群为

$$\mathrm{symm}^+(P) := \{T \in \mathrm{SO}(3) : T(P) = P\};$$

它们分别是 $\mathrm{O}(3)$ 和 $\mathrm{SO}(3)$ 的子群. 群 $\mathrm{symm}(P)$ 的元素也称为 P 的对称性.

不难想见 P 的每个对称性都映顶点 (或棱, 面) 为顶点 (或棱, 面); 这点也可以从 §14.7 的理论得到严谨的解释.

[3)] 正十二面体的 20 个顶点坐标能够从构造直接写下, 它们适当伸缩后是 $(\pm1,\pm1,\pm1)$, $(0,\pm\phi,\pm\phi^{-1})$, $(\pm\phi,\pm\phi^{-1},0)$, $(\pm\phi^{-1},0,\pm\phi)$, 其中 $\phi := \frac{1+\sqrt{5}}{2}$ 是黄金比例. 正二十面体的 12 个顶点也有类似的描述: $(\pm\phi,\pm1,0)$, $(\pm1,0,\pm\phi)$, $(0,\pm\phi,\pm1)$. 这种描述便于制图.

若 P 的顶点数为 m, 将顶点排序, 则 T 在顶点集上的作用给出群同态 $\nu : \mathrm{symm}(P) \to \mathfrak{S}_m$. 由于 P 的体积为正, 顶点集必然包含 $\mathbb{R}^3$ 的基, 故 ν 是单同态. 因此 $\mathrm{symm}(P)$ (或 $\mathrm{symm}^+(P)$) 总是 $\mathrm{O}(3)$ (或 $\mathrm{SO}(3)$) 的有限子群.

本节仅论旋转对称群, 对称群 $\mathrm{symm}(P)$ 留待习题处理. 现在将目光放广, 考虑 $\mathrm{SO}(3)$ 的所有有限子群.

任何 $T \in \mathrm{SO}(3) \setminus \{\mathrm{id}\}$ 都有唯一确定的转轴, 交单位球面于两点, 称之为 T 的**极点**. 若 $T, U \in \mathrm{SO}(3)$ 而 x 是 $T \neq \mathrm{id}$ 的极点, 则 Ux 是 UTU^{-1} 的极点.

今起设 G 为 $\mathrm{SO}(3)$ 的有限子群. 对单位球面上的所有点 x, 记 $G_x := \mathrm{Stab}_G(x)$, 记 $n_x := |G_x|$.

命题 11.11.4 *群 G_x 总是有限循环群, 它能由某个角度为 $\frac{2\pi}{n_x}$ 的旋转生成.*

证明 不妨设 G_x 非平凡. 将旋转的角度 θ 取在 $-\pi < \theta \leq \pi$, 考虑 G_x 中角度最小然而非零的旋转 T. 请读者以类似带余除法的论证说明 $G_x = \langle T \rangle$. 由此容易进一步取到具有所需转角的生成元. □

另外记

$$\begin{aligned} \mathcal{S}_G &:= \{x : n_x > 1\}, \quad \text{带 } G \text{ 的左作用}, \\ \Omega_G &:= \big\{(T, x) : T \in G,\ T \neq \mathrm{id},\ x \text{ 是 } T \text{ 的极点}\big\}. \end{aligned}$$

因为每个 $T \in G \setminus \{\mathrm{id}\}$ 恰有两个极点, 故 $|\Omega_G| = 2(|G| - 1)$. 另一方面, 每个 $x \in \mathcal{S}_G$ 对应到 $n_x - 1$ 个以之为极点的 T, 故又有 $|\Omega_G| = \sum_{x \in \mathcal{S}_G} (n_x - 1)$. 对 $\mathcal{S}_G$ 的所有 G-轨道取代表元 $x_1, \ldots, x_k$, 并且记 $n_i := n_{x_i} \leq |G|$, 遂有

$$2|G| - 2 = \sum_{x \in \mathcal{S}_G} (n_x - 1) = \sum_{i=1}^{k} (G : G_{x_i})(n_i - 1) = \sum_{i=1}^{k} |G| \left(1 - \frac{1}{n_i}\right),$$

亦即 $2 - \dfrac{2}{|G|} = \displaystyle\sum_{i=1}^{k} \left(1 - \frac{1}{n_i}\right)$. 此外 (11.5.1) 还蕴涵 $|\mathcal{S}_G| = \sum_{i=1}^{k} \frac{|G|}{n_i}$.

观察到 $n_i \geq 2$. 由上式容易推得 $|G| \geq 2$ 时 $k \in \{2, 3\}$.

引理 11.11.5 *设 G 为 $\mathrm{SO}(3)$ 的非平凡有限子群, 则以上数值的可能性仅有*

$\lvert G\rvert$	k	$\lvert\mathcal{S}_G\rvert$	n_1	n_2	n_3
n	2	2	n	n	无
$2n$	3	$2n+2$	2	2	n
12	3	14	2	3	3
24	3	26	2	3	4
60	3	62	2	3	5

其中 $n\in\mathbb{Z}_{\geq 2}$, 第一行对应到 G 为 n 阶循环群的情形, 第二行对应到 G 同构于二面体群 D_{2n} 的情形 (例 11.10.5).

证明 设 $k=2$, 则资料的唯一选法是 $n_1=n_2=|G|$, 从而 $|\mathcal{S}_G|=2$. 因此 G 绕单个轴旋转, 是循环群.

设 $k=3$ 并且不失一般性设 $n_1\leq n_2\leq n_3$. 容易验证不可能有 $n_1\geq 3$, 故 $n_1=2$. 此外还也容易排除 $n_2\geq 4$ 的情形, 故 $n_2\in\{2,3\}$.

若 $n_2=2$ 则必有 $n_3=\frac{|G|}{2}=:n\geq 2$. 此时 $\frac{|G|}{n_1}=\frac{|G|}{n_2}=n$ 而 $\frac{|G|}{n_3}=2$. 于是:

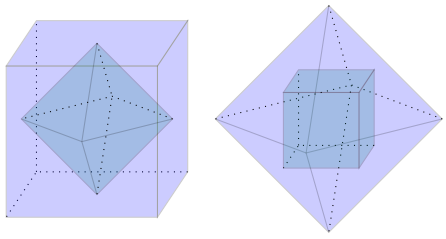
- $|\mathcal{S}_G|=2n+2$;
- G 包含绕某个轴的 n 阶旋转群 $C_n:=G_{x_3}$ (亦即正 n 边形的旋转群), 极点用掉 $\mathcal{S}_G$ 的两个元素, 此外 $(G:C_n)=2$ 蕴涵 $C_n\lhd G$;
- 剩下的 n 个旋转恰有 $2n$ 个极点, 与 C_n 的极点无交, 而且它们正规化 C_n, 故保持 C_n 的转轴. 因此这 n 个旋转必翻转 C_n 的转轴, 对应到以例 11.10.5 图像的 n 个虚线为轴作角度 π 的旋转, 限制在纸面上就是镜射.

综上立见 $k=3$ 而 $(n_1,n_2)=(2,2)$ 的情形给出二面体群 D_{2n}. 留意到当 $n=2$ 时上述论证仍然适用, 给出 $D_4\simeq(\mathbb{Z}/2\mathbb{Z})^2$, 由对相互正交的两轴作角度 π 的旋转所生成.

剩下的是 $(n_1,n_2)=(2,3)$ 的情形, 留给读者穷举. □

问题在于研究引理 11.11.5 列表的后三行. 为此, 且将目光转回正多面体的旋转对称群.

基于定理 11.11.2 证明中的构造, 正立方体与正八面体相互内接, 正十二面体与正二十面体相互内接, 诀窍是 “顶点内接于面的重心”; 以下是正立方体与正八面体的示意图.



(11.11.1)

所以正多面体的旋转对称群其实只有三种, 记为

$$\mathbf{T}:=\mathrm{symm}^+(\text{正四面体}),\quad \mathbf{O}:=\mathrm{symm}^+(\text{正八面体}),\quad \mathbf{I}:=\mathrm{symm}^+(\text{正二十面体}).$$

此处为每种正多面体在全等类中选定一个标准代表元, 以原点为对称中心, 按此得到确定的群. 留意到将多面体对原点缩放不改变旋转对称群.

定理 11.11.6 对于以上定义的三种旋转对称群, 存在典范同构

$$\begin{aligned}\mathbf{T} &\simeq \mathfrak{A}_4,\\ \mathbf{O} &\simeq \mathfrak{S}_4,\\ \mathbf{I} &\simeq \mathfrak{A}_5;\end{aligned}$$

此处的 $\mathfrak{A}_n$ 是例 11.1.9 介绍的交错群.

证明 (勾勒) 对于 $\mathbf{T}$, 所论的同构将取为 $\nu : \mathbf{T} \hookrightarrow \mathfrak{S}_4$, 细说如下.

连接每个顶点及其对面重心为轴 (4 种选法), 作角度为 $\pm\frac{2\pi}{3}$ 的旋转给出群 $\mathbf{T}$ 的元素; 取任一对无交的棱 (不计顺序有 3 对), 连接其中点为轴作角度 π 的旋转, 也给出 $\mathbf{T}$ 的元素. 不难看出这穷尽 $\mathbf{T} \smallsetminus \{\mathrm{id}\}$ 的所有元素. 对应的极点有 $2(4+3) = 14$ 个.

将此代入引理 11.11.5 的列表. 注意到表中的资料 $(n_i)_{i=1}^k$ 决定群中所有的旋转角度, 即 $\frac{2\pi\mathbb{Z}}{n_i}$; 见命题 11.11.4. 上一段关于旋转的描述说明 $\mathbf{T}$ 不可能同构于循环群或二面体群 D_6, 唯一可能是第三行的群, $|\mathbf{T}| = 12$. 此外, 这些描述也表明 $\mathbf{T}$ 的元素诱导顶点的偶置换, 故 $\nu(\mathbf{T}) \subset \mathfrak{A}_4$. 然而 $|\mathfrak{A}_4| = 12$.

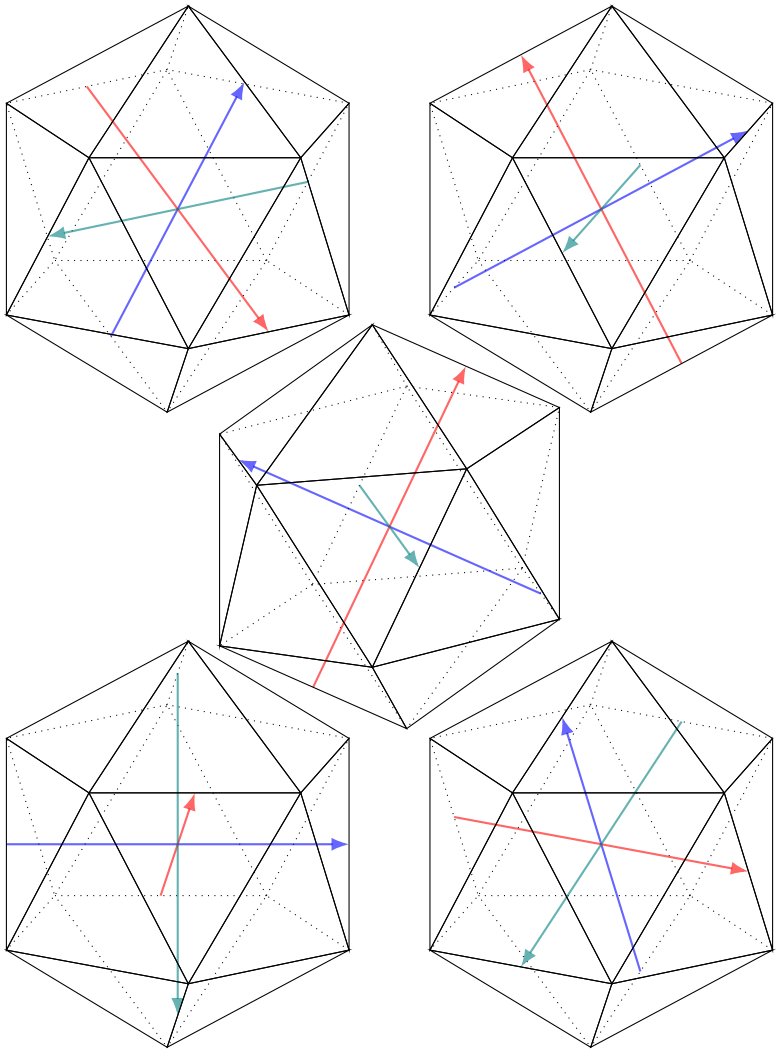
接着考虑群 $\mathbf{O}$. 它显然包含角度为 $\frac{\pi}{2}$ 的旋转, 可选转轴至少三个, 故 $\mathbf{O}$ 不可能同构于循环群或 D_8. 代入引理 11.11.5 的列表, 可知是第四行的群, $|\mathbf{O}| = 24$.

现考虑正八面体任两个无交的面 (不记顺序有 4 对), 连接其重心, 则 $\mathbf{O}$ 的作用置换这 4 个直线 $\ell_1, \dots, \ell_4$, 给出群同态 $\mu : \mathbf{O} \to \mathfrak{S}_4$. 我们有 $\ker(\mu) = \{\mathrm{id}\}$, 这是因为 $T \in \ker(\mu)$ 作为 $\mathbb{R}^3$ 的线性自同态有 3 个线性无关的特征向量, 若 $T \neq \mathrm{id}$, 则这种旋转的特征值必为 $-1, -1, 1$. 几何上, 这相当于绕某个 ℓ_i 作角度 π 的旋转, 然而这种旋转不可能是正八面体的对称性: 角度只容许为 $\frac{2\pi}{3}$ 的倍数. 由于 $|\mathfrak{S}_4| = 24$, 综上可知 μ 是同构.

最后考虑 $\mathbf{I}$. 首先回顾定理 11.11.2 证明中对正十二面体的构造, 可知 $\mathbf{I}$ 包含角度 $\frac{2\pi}{5}$ 的旋转, 可选转轴至少有 12 个. 于是 $\mathbf{I}$ 不能同构于循环群或 D_{10}. 代入引理 11.11.5 的列表, 可知是第五行的群, $|\mathbf{I}| = 60$.

以下构造群同态 $\omega : \mathbf{I} \to \mathfrak{S}_5$. 取正二十面体的对棱中点连线 (过原点), 共有 15 条. 可以证明[4]其中两两正交的三元组 (不计顺序) 共有 5 个, 如下图所示:

[4] 读者可尝试给出几何论证, 更朴素的方法是用坐标来计算. 一旦知道各棱的顶点, 计算机便能代劳.

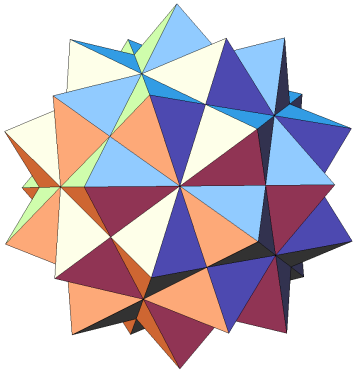


其中的箭头总是指向可见 (非虚线) 的棱. 群 $\mathbf{I}$ 置换这些三元组, 给出 ω.

观图见得 ω 非平凡. 本章习题及其提示将说明 $\mathbf{I}$ 是定义 11.9.2 所谓单群[5]. 由此可得 ω 是单同态. 本章的另一道习题 (基于 $\mathfrak{A}_5$ 的单性) 则说明 $\mathfrak{S}_5$ 的 60 阶子群只有 $\mathfrak{A}_5$, 由此即得 $\omega : \mathbf{I} \xrightarrow{\sim} \mathfrak{A}_5$. 尚有针对正十二面体来操作的几何论证, 可参见 [12, §2.5]. □

关于证明中考虑的 5 个三元组, 等价的说法是正二十面体自然地内接 5 个正八面体, 而 $\mathbf{I}$ 置换之. 这 5 个正八面体的复合体是饶富兴味的几何对象, 由 Edmund Hess 首先发现, 如下图.

[5] 代数学中一则广为人知的事实 (也是本章习题之一) 是 60 阶单群必和 $\mathfrak{A}_5$ 同构, 从此便能推导 $\mathbf{I} \simeq \mathfrak{A}_5$, 而不必取道 5 个正八面体的复合体; 然而这般得到的同构并非典范的.



图源: Wikimedia Commons.

现在回头处理 SO(3) 的有限子群, 并给出完整分类.

定理 11.11.7　精确到共轭, SO(3) 的有限子群 G 仅有以下五类: 旋转生成的有限循环群 $\mathbb{Z}/n\mathbb{Z}$ (其中 $n \in \mathbb{Z}_{\geq 1}$), 二面体群 D_{2n} (其中 $n \in \mathbb{Z}_{\geq 2}$), 以及正多面体的旋转对称群 $\mathbf{T}$, $\mathbf{O}$, $\mathbf{I}$.

证明　不难验证嵌入于 SO(3) 的所有 $\mathbb{Z}/n\mathbb{Z}$ 和 D_{2n} 都相互共轭, 故讨论引理 11.11.5 列表中的后三种情况即可. 定理 11.11.6 表明它们分别包括 $\mathbf{T}$, $\mathbf{O}$, $\mathbf{I}$ 作为特例. 关键在说明这三种特例的共轭穷尽所有情形.

策略是从极点集 $\mathcal{S}_G$ 中的 G-轨道构造正多面体. 从 $(n_1, n_2, n_3) = (2, 3, 3)$ 情形起步. 考虑 n_3 在 $\mathcal{S}_G$ 中对应的 G-轨道, 写作 $\{x_1, \ldots, x_4\}$. 非平凡旋转不可能固定球面上三个相异点, 而 $|G_{x_1}| = 3$, 故 G_{x_1} 轮换 x_2, x_3, x_4, 它们和 x_1 等距; 同理可见 $x_1, \ldots, x_4$ 两两等距. 对照正四面体的构造可见它们张成[6]正四面体 P.

剩余的论证是容易的: P 的对称中心被 G 的所有元素固定. 然而引理 11.11.5 的列表说明 G 至少有两个不同转轴, 故对称中心便是 $\mathbb{R}^3$ 的原点. 综上可知 $G \subset \mathrm{symm}^+(P)$. 基于正多面体的分类, 可取 $g \in \mathrm{SO}(3)$ 使得 $g(P)$ 等于标准的正四面体, 相应地 $gGg^{-1} \subset \mathbf{T}$. 比较阶数立得 $gGg^{-1} = \mathbf{T}$.

对于 $(n_1, n_2, n_3) = (2, 3, 4)$ 的情形, n_3 在 $\mathcal{S}_G$ 中对应的 G-轨道有 6 个元素 $x_1, \ldots, x_6$. 群 G_{x_1} 由角度 $\frac{\pi}{2}$ 的旋转生成, 类似的几何理由表明它轮换 $x_2, \ldots, x_6$ 中的 4 个元素, 剩下者不动. 按此便将 $x_1, \ldots, x_6$ 的元素分成 3 对: x_i 和 x_j 配对当且仅当存在保持两者皆不动的 4 阶旋转, 当且仅当 $x_i = -x_j$.

给定配对的 x_i 和 x_j, 剩下四个元素都落在与其连线正交的某个平面上, 因此 3 对 (x_i, x_j) 给出相互正交的直线, 皆过原点. 既然 G 传递地作用, $x_1, \ldots, x_6$ 和原点等距. 将此与定理 11.11.2 中的构造比较, 可知它们张成正八面体. 其余论证和先前相同.

对于 $(n_1, n_2, n_3) = (2, 3, 5)$ 的情形, n_3 在 $\mathcal{S}_G$ 中对应的 G-轨道有 12 个元素

[6]用 §14.7 的精确语言来说, 这四个点的凸包是正四面体.

$x_1,\dots,x_{12}$. 同样的论证说明 G_{x_1} 由某个角度 $\frac{2\pi}{5}$ 的旋转 S 生成, 保持另一个元素 (设为 x_{12}) 不动, 然后将剩余的 10 个元素等分成两个轨道, 每个轨道的元素被 S 轮换.

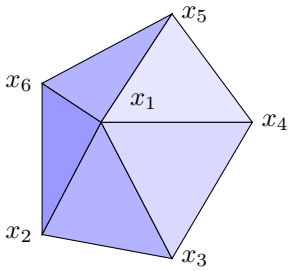
这也说明 $x_{12}=-x_1$, 它们确定 S 的转轴; 推而广之, $\{x_1,\dots,x_{12}\}$ 对 $x\mapsto -x$ 保持不变. 调整下标后不妨 $\langle S\rangle$ 的两个非平凡轨道是

$$x_2,\dots,x_6,\quad x_7,\dots,x_{11},$$

而且满足 $x_3=Sx_2$ 和 $x_8=Sx_7$, 依此类推. 两个轨道的地位迄今是平等的, 但稍后将规定 x_2 的取法.

既然是 $x_2,\dots,x_6$ 是 S 作用下的轨道, 它们在与 S 转轴正交的某个平面上依序构成边长记为 ℓ 的正五边形, 各顶点和 x_1 等距, 也和 $x_{12}=-x_1$ 等距. 此外, 正五边形的图像表明 $-x_2\notin\{x_2,\dots,x_6\}$. 故另一个轨道表为 $\{-x_2,\dots,-x_6\}$, 在 S 旋转下仍构成边长 ℓ 的正五边形.

取 $U\in G$ 使得 $U(x_1)\notin\{x_1,x_{12}\}$, 然后取 $T=U^{-1}SU$, 则 T 是角度同为 $\frac{2\pi}{5}$ 的旋转, 而且考虑 S 的极点和特征值可见 $T(x_1)\notin\{x_1,x_{12}\}$. 不失一般性 (调整下标), 无妨设 $T(x_1)=x_2$. 基于对称考量 (用 U 搬运), x_1 到 $x_2=T(x_1)$ 的距离也等于 ℓ. 综上, $x_1,\dots,x_6$ 和定理 11.11.2 证明中的正二十面体基本构件全等, 如下图:



继续类似构造, 直到穷尽 12 个顶点. 这表明 $x_1,\dots,x_{12}$ 张成正二十面体. 其余论证和先前相同. □

习题

1. 验证 $\{(x,y)\in\mathbb{R}^2 : x\neq 0\}$ 对二元运算 $(a,b)(c,d):=(ac,ad+b)$ 成群.

2. 将对称群 $\mathfrak{S}_n$ 嵌入为 $\mathrm{GL}(n,F)$ 的子群, 其中 F 是任意域. 〈提示〉 考虑置换矩阵.

3. 设 G 为群.

(i) 证明若 H,K 为 G 的子群, $G=H\cup K$, 则必有 $G=H$ 或 $G=K$.

(ii) 给定一族递增子群 $G_1\subset G_2\subset G_3\subset\cdots$, 证明 $\bigcup_{k=1}^{\infty}G_k$ 仍是 G 的子群.

4. 设 G 是有限半群, 而且满足左消去律 $xy=xz\implies y=z$ 和右消去律 $yx=zx\implies y=z$. 证明 G 是群.

5. 说明对于所有群 G, 映射 $g \mapsto g^{-1}$ 给出群同构 $G \xrightarrow{\sim} G^{\mathrm{op}}$.

6. 设 a, b 属于群 G.

(i) 证明若 ab 是有限阶元素, 则 ba 亦然, 此时 $\mathrm{ord}(ab) = \mathrm{ord}(ba)$.

(ii) 举例说明当 a, b 都有限阶时, ab 未必有限阶.

7. 对幺半群 M 及 $x \in M$, 若存在 $x^{-1} \in M$ 使得 $xx^{-1} = 1_M = x^{-1}x$, 则称 x 可逆. 记 M 的所有可逆元构成的子集为 $M^\times$.

(i) 说明以上的 x^{-1} 若存在则唯一, 故可合理地称为 x 的逆.

(ii) 说明 $M^\times$ 对 M 的乘法构成群, 包含映射 $M^\times \to M$ 是幺半群之间的同态.

(iii) 对于环 R, 将环论中定义的可逆元集 $R^\times$ 诠释为上述构造的特例.

(iv) 设 G 为群而 $\varphi : G \to M$ 为幺半群之间的同态, 亦即 $\varphi(1_G) = 1_M$ 而 $\varphi(xy) = \varphi(x)\varphi(y)$ 恒成立. 说明 φ 的像落在 $M^\times$ 中.

8. 设群 $G \neq \{1\}$. 说明 G 没有非平凡真子群的必要条件是 G 同构于素数阶循环群; 此条件也是充分的, 见例 11.9.3.

9. 证明对称群 $\mathfrak{S}_n$ 的中心在 $n \geq 3$ 时是平凡的.

10. 证明 $\mathfrak{S}_n$ 由 $(1\ 2)$ 和 $(1\ 2\ \cdots\ n)$ 生成.

11. 证明若 G/Z_G 是循环群, 则群 G 交换.

12. 设 p 为素数. 证明满足 $|G| = p^2$ 的群 G 皆交换. 〉提示〈 结合上一题与命题 11.6.4.

13. 证明若 G 有真子群 H 使得 $(G : H)$ 有限, 则存在 $N \lhd G$ 使得 $N \neq G$ 而 $(G : N)$ 有限.

14. (L. Euler) 设 $n \in \mathbb{Z}_{\geq 1}$. 运用推论 11.4.10 证明若 $x \in \mathbb{Z}$ 与 n 互素, 则

$$x^{\varphi(n)} \equiv 1 \pmod{n},$$

其中 φ 是 Euler 函数. 进一步取 n 等于某个素数 p, 由此推导 Fermet 小定理.

〉提示〈 取群 G 为 $(\mathbb{Z}/n\mathbb{Z})^\times$, 确定其元素个数.

15. 考虑可逆四元数群 $\mathbb{H}^\times$ 的子群 $Q_8 := \{\pm 1, \pm \mathrm{i}, \pm \mathrm{j}, \pm \mathrm{k}\}$ (共 8 个元素).

(i) 证明 Q_8 的所有子群皆正规, Q_8 非交换.

(ii) 记 Q_8 的中心为 Z. 证明 Q_8/Z 同构于 $(\mathbb{Z}/2\mathbb{Z})^2$.

(iii) 写下 Q_8 中的所有共轭类 (共五个).

16. 设有限群 G 满足以下条件: 对所有 $d \in \mathbb{Z}_{\geq 1}$, 集合 $\{x \in G : x^d = 1\}$ 至多只有 d 个元素. 记 $n := |G|$.

(i) 说明对于所有 $d \mid n$, 集合 $\{g \in G : \mathrm{ord}(g) = d\}$ 或者是空集, 或者有 $\varphi(d)$ 个元素.
〉提示〈 论证 $\mathrm{ord}(g) = d$ 蕴涵 $\langle g \rangle = \{x \in G : x^d = 1\}$, 再对 $\langle g \rangle$ 应用推论 11.3.4.

(ii) 基于 (i), 说明 G 是 n 阶循环群.
〉提示〈 对所有 $d \mid n$, 记 $m_d := |\{g \in G : \mathrm{ord}(g) = d\}|$, 然后注意到 $\sum_{d|n} m_d = n = \sum_{d|n} \varphi(d)$.

17. 设 F 为域. 利用上一题的结果, 证明 $F^\times$ 的所有有限子群皆为循环群. 作为推论, 若 F 是有限域则 $F^\times$ 是循环群.

18. (原根) 考虑环 $\mathbb{Z}/n\mathbb{Z}$, 要求 $n \in \mathbb{Z}_{\geq 2}$. 证明 $(\mathbb{Z}/n\mathbb{Z})^\times$ 是循环群当且仅当 $n = 2, 4, p^k, 2p^k$, 其中 p 是奇素数而 $k \in \mathbb{Z}_{\geq 1}$. 若 $a + n\mathbb{Z}$ 是 $(\mathbb{Z}/n\mathbb{Z})^\times$ 的生成元, 则称 a 为 $\bmod n$ **原根**; 原根的分布和算法是数论及其应用的重要问题.

提示〉 上一则习题可处理 n 为素数的情形. 一般情形的推导基于素数情形和中国剩余定理, 然而一些数论技巧仍属必要. 详见各种初等数论教材.

19. 设群 G 的自同构只有恒等. 证明 $|G| \leq 2$. 提示〉 考虑内自同构可见 G 交换, 不妨将群运算写作加法. 接着考虑自同构 $x \mapsto -x$, 将 G 作成 $\mathbb{F}_2$-向量空间.

20. (交换环上的一般线性群) 说明任意交换环 R 上的所有 $n \times n$ 可逆矩阵对乘法也构成群, 记为 $\mathrm{GL}(n, R)$, 它有子群 $\mathrm{SL}(n, R) := \{\boldsymbol{A} \in \mathrm{M}_{n\times n}(R) : \det \boldsymbol{A} = 1\}$.

21. 对任意域 F, 证明形如

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix}, \quad x, y \in F$$

的矩阵构成群 $\mathrm{SL}(2, F)$ 的一族生成元. 在此基础上, 试以类似方法给出 $\mathrm{SL}(n, F)$ 的生成元.

提示〉 运用以下等式:

$$\begin{pmatrix} 0 & t \\ -t^{-1} & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -t^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} = \begin{pmatrix} 0 & t \\ -t^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

22. 承上题, 证明上述矩阵在 $x, y \in \mathbb{Z}$ 的情形也是群 $\mathrm{SL}(2, \mathbb{Z})$ 的一族生成元.

提示〉 首先注意到上述矩阵 (取 $x, y \in \mathbb{Z}$) 足以实现下述初等行 (或列) 变换, 采用 §1.3 的符号:

- ⋆ $\mathrm{A}(i, k, c)$, 其中 $c \in \mathbb{Z}$;
- ⋆ 先做 $\mathrm{B}(i, k)$, 再将某行 (或列) 乘以 -1;
- ⋆ 将矩阵乘以 $-\mathbf{1}_{2\times 2}$.

给定 $\gamma \in \mathrm{SL}(2, \mathbb{Z})$, 对之实行上述变换所得的所有矩阵构成集合 $\mathcal{C}$, 证 $\mathbf{1}_{2\times 2} \in \mathcal{C}$ 即可. 考虑 $\mathcal{C}$ 的所有元素的所有矩阵元中非零而绝对值最小者, 然后操作带余除法.

23. 证明交换单群必然同构于 $\mathbb{Z}/p\mathbb{Z}$, 其中 p 是素数. 提示〉 先将问题化到循环群的情形考虑.

24. (双陪集) 设 $H, K \subset G$ 为子群.

(i) 让 H 通过左乘左作用在 G/K 上. 证明陪集 gK 的稳定化子是 $gKg^{-1} \cap H$.

(ii) 让 K 通过右乘右作用在 $H\backslash G$ 上. 证明陪集 Hg 的稳定化子是 $g^{-1}Hg \cap K$.

(iii) 让 $H \times K^{\mathrm{op}}$ 按照 $(h, k) \cdot g := hgk$ 左作用在 G 上. 说明这确实是群作用, 而且对应的商集既等同于 $H\backslash(G/K)$, 也等于 $(H\backslash G)/K$.

相对于 $H \times K^{\mathrm{op}}$ 对 G 的左作用, 其轨道称为 G 对 H 和 K 的**双陪集**, 具体表作 $HgK := \{hgk : h \in H,\ k \in K\}$ 的形式, 其中 $g \in G$.

25. 将命题 11.9.17 的 $G/H \xrightarrow{\sim} G'/H'$ (其中 $H' = f(H)$) 推及 $N \lhd G$ 而 H 是包含 N 的任意子群的情形. 此时 G/H 和 G'/H' 未必有群结构, 而 $G/H \to G'/H'$ 只是集合之间的双射.

提示〉 直接验证映射 $gH \mapsto f(g)H'$ 良定义, 满而且单.

26. 设 $p > q$ 为两个素数, 群 G 满足 $|G| = pq$.

(a) 证明存在正规子群 P 使得 $(G : P) = q$. 提示〉 取子群 H 使得 $(G : H) = q$. 让 G 以左乘作用于 G/H, 对应的群同态 $A : G \to \mathfrak{S}_q$ 必满足 $|\ker A| = p$.

(b) 证明若 $q \nmid p - 1$ 则 G 是循环群. 提示〉 将 G 分解为半直积, $\mathrm{Aut}(\mathbb{Z}/p\mathbb{Z}) \simeq (\mathbb{Z}/p\mathbb{Z})^\times \simeq \mathbb{Z}/(p-1)\mathbb{Z}$.

(c) 设 G_1 和 G_2 是非交换群, $|G_1| = pq = |G_2|$, 证明 $G_1 \simeq G_2$.
提示〉 半直积 $(\mathbb{Z}/p\mathbb{Z}) \rtimes (\mathbb{Z}/q\mathbb{Z})$ 可由 $\{a \in (\mathbb{Z}/p\mathbb{Z})^\times : a^q = 1\} \simeq \mathbb{Z}/q\mathbb{Z}$ 的元素描述. 研究 $\mathbb{Z}/q\mathbb{Z}$ 的自同构如何影响 a.

(d) 设 p 为奇素数. 证明 $2p$ 阶非交换群必同构于 D_{2p}.

27. 描述循环群 $\mathbb{Z}/n\mathbb{Z}$ 的所有自同态, 其中 $n \geq 0$.

28. 描述二面体群 D_{2n} 的自同构群 $\mathrm{Aut}(D_{2n})$.

29. 如果群 G 的子群 H 对于所有自同构 $\varphi : G \xrightarrow{\sim} G$ 都满足 $\varphi(H) = H$, 则称之为**特征子群**.

(i) 证明特征子群总是正规子群, 特征子群的特征子群仍是特征子群.

(ii) 说明 G 的中心 Z_G 是特征子群, 而且群 G 的**导出子群**

$$G_{\mathrm{der}} := \left\langle aba^{-1}b^{-1} : a, b \in G \right\rangle$$

也是特征子群.

(iii) 证明 $\mathrm{GL}(n, F)_{\mathrm{der}} = \mathrm{SL}(n, F)$, 其中 F 是包含至少 3 个元素的域.
提示〉 对于 $n = 2$ 的情形, 观察到

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & (a-1)x \\ 0 & 1 \end{pmatrix}.$$

30. 群 G 的**交换化**定义为 $G_{\mathrm{ab}} := G/G_{\mathrm{der}}$.

(i) 说明 G_{ab} 交换, 而且对于任何交换群 A 和同态 $f : G \to A$, 存在唯一的同态 $\overline{f} : G_{\mathrm{ab}} \to A$ 使得 f 分解为 $G \xrightarrow{\text{商}} G_{\mathrm{ab}} \xrightarrow{\overline{f}} A$ 的合成. 用 §B.5 的术语来说, 这是交换化的泛性质.

(ii) 就 $\mathfrak{S}_3$, Q_8, D_{2n} 和 $\mathrm{GL}(n, F)$ 这几个例子确定其交换化, 其中 F 是任意域.

(iii) 对于群 $G = \mathrm{GL}(n, F)$, 其中 F 是含至少三个元素的域, 试将商同态 $G \to G_{\mathrm{ab}}$ 等同于行列式.

31. 设群 G 作用在集合 X 上, $n \in \mathbb{Z}_{\geq 1}$. 让 G 按照 $g(x_1,\dots,x_n)=(gx_1,\dots,gx_n)$ 作用于 X^n 及其子集

$$X^{[n]} := \{(x_1,\dots,x_n) \in X^n : i \neq j \implies x_i \neq x_j\}.$$

设 $|X| \geq n$. 若 G 对 $X^{[n]}$ 的作用传递, 则称 G 对 X 的作用为 **n-传递**的.

(i) 说明 $\mathfrak{S}_n$ 在 $\{1,\dots,n\}$ 上的标准作用是 n-传递的.

(ii) 说明若 $m \leq n$, 则 n-传递蕴涵 m-传递.

(iii) 试证若 G 和 X 有限, 而且作用是 2-传递的, 则 $\sum_{g\in G}|X^g|^2 = 2|G|$, 其中 $X^g := \{x \in X : gx = x\}$.

提示〉 对所有 $x \in X$, 对 $\mathrm{Stab}_G(x)$ 在 $X \smallsetminus \{x\}$ 上的传递作用应用命题 11.5.10, 推得 $\sum_{g\in \mathrm{Stab}_G(x)}|X^g| = 2|\mathrm{Stab}_G(x)|$, 然后对 x 加总, 右侧需以 G 的传递性处理.

32. 设群 G 作用于集合 X, 从而作用于 X^2; 对所有 $x \in X$ 记 $G_x := \mathrm{Stab}_G(x)$. 证明 (x,gx) 和 $(x,g'x)$ 在 X^2 中属于同一个 G-轨道当且仅当 $g' \in G_x g G_x$.

33. 接续上一题, 进一步要求 $|X| \geq 2$ 而 G 对 X 的作用为 2-传递的.

(i) 证明 G_x 总是 G 的极大真子群: 换言之, 包含 G_x 的子群只有 G 和 G_x 本身.
提示〉 对所有 $g \in G \smallsetminus G_x$ 皆有 $G = G_x \cup G_x g G_x$.

(ii) 说明任何正规子群 $N \lhd G$ 在 X 上的作用或者平凡, 或者传递.
提示〉 取定 $y \neq y'$. 若存在 x 和 n 使得 $nx \neq x$, 则取 g 使得 $gx = y$ 而 $g(nx) = y'$, 按此得 $gng^{-1}y = y'$.

(iii) (岩泽健吉) 假设 G 作用忠实, $G = G_{\mathrm{der}}$, 而且存在 x 使得 G_x 有正规交换子群 U, 而 U 在 G 中的所有共轭生成 G. 证明 G 为单群.
提示〉 取定 $N \lhd G$ 和 x, 讨论 $NG_x = G_x$ 和 $NG_x = G$ 两种情形. 若 $NG_x = G_x$ (亦即 $N \subset G_x$) 则 N 的作用平凡. 若 $NG_x = G$ 则从 $NU \lhd NG_x = G$ 易见 $NU = G$, 由此推导 G/N 交换.

34. 对所有域 F 验证群 $\mathrm{SL}(n,F)$ 的中心等于 $\{\lambda \cdot \mathbf{1}_{n\times n} : \lambda \in F,\ \lambda^n = 1\}$.

35. ($\mathrm{PSL}(2,F)$ 的单性, $|F| \geq 4$) 设 F 为域, 命 Z 为 $\mathrm{SL}(n,F)$ 的中心. 按照以下步骤证明 $n = 2$ 而 $|F| \geq 4$ 时 $\mathrm{PSL}(2,F) := \mathrm{SL}(2,F)/Z$ 为单群.

(i) 说明 $|F| \geq 4$ 时 $\mathrm{SL}(2,F) = \mathrm{SL}(2,F)_{\mathrm{der}}$. 提示〉 将先前证明 $\mathrm{GL}(2,F)_{\mathrm{der}} = \mathrm{SL}(2,F)$ 时取的 $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ 改为 $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$.

(ii) 让 $\mathrm{PSL}(2,F)$ 以显然方式作用在 $\mathbb{P}^1(F) := \{F^2$ 的 1 维子空间$\}$ 上. 记 $(x,y) \in F^2 \smallsetminus \{\mathbf{0}\}$ 生成的子空间为 $(x:y)$. 说明这是 2-传递作用, 然后写下 $(1:0)$ 的稳定化子群 H.

(iii) 代入岩泽健吉的判准, 推导 $|F| \geq 4$ 时 $\mathrm{PSL}(2,F)$ 为单群.
提示〉 取 H 的子群

$$U := \left\{ \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} Z \right\}.$$

36. (PSL(n,F) 的单性, $n\geq 3$) 承上题, 设 F 为任意域. 按以下步骤证明 $n\geq 3$ 时 $\mathrm{PSL}(n,F):=\mathrm{SL}(n,F)/Z$ 为单群.

(i) 说明 $\mathrm{SL}(n,F)=\mathrm{SL}(n,F)_{\mathrm{der}}$. 〈提示〉 以 $n=3$ 为例, 计算 $ghg^{-1}h^{-1}$, 其中

$$g:=\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix},\quad h:=\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}.$$

(ii) 让 $\mathrm{PSL}(n,F)$ 以显然方式作用在 $\mathbb{P}^{n-1}(F):=\{F^n \text{ 的 1 维子空间}\}$ 上. 说明这是 2-传递作用, 然后写下 $(1:0:\cdots:0)$ 的稳定化子群 H.

(iii) 代入岩泽健吉的判准, 推导 $\mathrm{PSL}(n,F)$ 单.

〈提示〉 取 H 的子群

$$U:=\left\{\left(\begin{array}{c|c}\mathbf{1}_{1\times 1}&*\\\hline 0&\mathbf{1}_{(n-1)\times(n-1)}\end{array}\right)Z\right\}.$$

在 §14.4 将对 $\mathbb{P}^{n-1}(F)$ 作进一步的介绍.

37. (圈积) 设 D 和 Q 为群, 非空集 Ω 带左 Q-作用. 取群的直积 D^Ω, 定义群同态 $\varphi: Q\to \mathrm{Aut}(D^\Omega)$ 使得对所有 $q\in Q$ 都有 $\varphi_q\left((d_\omega)_{\omega\in\Omega}\right)=\left(d_{q^{-1}\omega}\right)_{\omega\in\Omega}$. 定义相应的**圈积**为

$$D\wr_\Omega Q:=D^\Omega\rtimes_\varphi Q.$$

(i) 设 D 左作用在非空集 X 上, 说明下式定义 $D\wr_\Omega Q$ 在 $X\times\Omega$ 上的左作用.

$$\left((d_\omega)_{\omega\in\Omega},q\right)\cdot(x,\eta):=\left(d_{q\eta}x,q\eta\right),\quad x\in X,\ \eta\in\Omega.$$

(ii) 承上题, 说明如果 D 和 Q 在 X 和 Ω 上的作用皆忠实, 则 $D\wr_\Omega Q$ 在 $X\times\Omega$ 上的作用亦然.

38. 考虑有 $q:=p^n$ 个元素的有限域 $\mathbb{F}_q$. 设 $m\in\mathbb{Z}_{\geq 1}$. 记 $U\subset\mathrm{GL}(m,\mathbb{F}_q)$ 为对角线全为 1 的上三角矩阵所成子群. 证明 U 是 $\mathrm{GL}(m,\mathbb{F}_q)$ 的 Sylow p-子群 (定义 11.6.6).

39. (L. A. Kaluznin) 以下运用圈积探讨对称群的 Sylow p-子群, p 是选定的素数.

(i) 对所有 $m\in\mathbb{Z}_{\geq 1}$ 构造 $\mathfrak{S}_{p^m}$ 的 Sylow p-子群 P_m, 使得 $P_1\simeq\mathbb{Z}/p\mathbb{Z}$ 而 $P_{m+1}\simeq P_m\wr_\Omega(\mathbb{Z}/p\mathbb{Z})$; 此处让 $\mathbb{Z}/p\mathbb{Z}$ 以加法作用于 $\Omega:=\mathbb{Z}/p\mathbb{Z}$.

〈提示〉 当 $m=1$ 时显然. 设已有 Sylow p-子群 $P_m\subset\mathfrak{S}_{p^m}$, 适当地将 $P_m\wr_\Omega(\mathbb{Z}/p\mathbb{Z})$ 嵌入 $\mathfrak{S}_{p^{m+1}}$ 并计数.

(ii) 对一般的 $n\in\mathbb{Z}_{\geq 1}$ 构造 $\mathfrak{S}_n$ 的 Sylow p-子群.

〈提示〉 作展开 $n=a_0+a_1p+\cdots+a_rp^r$, 其中 $0\leq a_i<p$. 将 $\{1,\ldots,n\}$ 划分为 $\sum_{i=0}^r a_i$ 个子集, 其中元素个数为 p^i 的子集有 a_i 个. 在这些子集各自的对称群里取 Sylow p-子群, 再将它们的直积嵌入 $\mathfrak{S}_n$.

40. 设 F 为满足 $\mathrm{char}(F)\neq 2$ 的域, 对辛空间 (V,B) 考虑例 11.1.13 定义的辛群 $\mathrm{Sp}(V)$. 证明 $Z_{\mathrm{Sp}(V)}=\{\pm\mathrm{id}_V\}$.

提示〉 给定 $z\in Z_{\mathrm{Sp}(V)}$, 取辛基 $p_1,q_1,\dots,p_n,q_n$ (计顺序). 于是辛群包含形式如下的矩阵

$$\left(\begin{array}{c|c} a\mathbf{1}_{2\times 2} & \\ \hline & b\mathbf{1}_{(n-2)\times(n-2)} \end{array}\right),\quad a,b\in\{1,-1\},$$

故 z 也对应到相同规格的分块对角矩阵, 其描述因而化约到 $n=1$ 情形, 此时辛群 $\simeq \mathrm{SL}(2,F)$ 的中心容易确定. 因此一般情形下 z 是分块皆为 $\pm\mathbf{1}_{2\times 2}$ 的分块对角矩阵. 考虑置换辛基所给出的辛群元素, 可见符号必然全正或全负.

41. 设 (V,B) 为域 F 上的辛空间, $\mathrm{char}(F)\neq 2$. 在集合 $V\times F$ 上定义二元运算

$$(v,t)\left(v',t'\right)=\left(v+v',t+t'+B(v,v')\right).$$

(i) 说明 $V\times F$ 对此运算成为群, 记为 $H(V,B)$, 称为 **Heisenberg 群**[7)].

(ii) 说明 F 作为加法群按照 $t\mapsto(0,t)$ 嵌入为 $H(V,B)$ 的子群, 而且 $Z_{H(V,B)}=F$.

(iii) 设 $\varphi:(V,B)\xrightarrow{\sim}(V',B')$ 为辛空间的同构, 说明 $(v,t)\mapsto(\varphi(v),t)$ 给出群同构 $H(V,B)\xrightarrow{\sim}H(V',B')$.

(iv) 考虑二维辛空间 (V_2,B_2), 带有分解 $V_2=Fp\oplus Fq$ 使得 $B_2(p,q)=1$, 因此 $V^\natural:=V_2\oplus V$ 对 $B^\natural:=B_2\oplus B$ 成为辛空间. 对所有 $(v,t)\in V\times F$ 按下式定义 $V^\natural$ 的线性自同态

$$\begin{aligned}\nu(v,t)(p)&=p+v+tq,\\ \nu(v,t)(w)&=w+B(w,v)q,\quad w\in V,\\ \nu(v,t)(q)&=q.\end{aligned}$$

验证这给出群的嵌入 $\nu:H(V,-B)\hookrightarrow\mathrm{Sp}(V^\natural)$.

(v) 验证 $\mathrm{Sp}(V)$ 在 $\mathrm{Sp}(V^\natural)$ 中的像包含于 $N_{\mathrm{Sp}(V^\natural)}(\mathrm{im}(\nu))$, 并描述 $\mathrm{Sp}(V)$ 如何按此作用于 $H(V,-B)$.

42. 应用命题 11.5.10 证明以下事实.

(i) 用 n 种颜色将一个正立方体的每个面着色. 两种着色如果差一个空间中的旋转, 则视为相同. 说明共有 $\frac{1}{24}\left(n^6+3n^4+12n^3+8n^2\right)$ 种着色方式.

提示〉 所有着色方式构成一个集合 X, 有 n^6 个元素. 正立方体的旋转对称群 G 作用于 X; 此群即定理 11.11.6 描述的 24 阶群 $\mathbf{O}$.

(ii) 用 n 种颜色将一个正三角形的每条边着色. 两种着色如果差一个平面上的旋转或镜射, 则视为相同. 说明共有 $\frac{1}{6}\left(n^3+3n^2+2n\right)$ 种着色方式.

提示〉 此时 $|X|=n^3$ 而 $|G|=6$; 群 G 正是二面体群 D_6.

43. (É. Galois) 试证 $n\geq 5$ 时 $\mathfrak{A}_n$ 是单群. 提示〉 参阅 [10, 定理 4.9.7].

[7)] 得名于量子力学中的基本对易关系 (W. Heisenberg, M. Born, P. Jordan). 一些文献中考虑的版本是 $H\left(V,\frac{1}{2}B\right)$.

44. 证明若 $n \geq 5$ 而子群 $H \subset \mathfrak{S}_n$ 满足 $(\mathfrak{S}_5 : H) = 2$, 则 $H = \mathfrak{A}_n$.

提示〉 依照 $H \lhd \mathfrak{S}_n$ 和 $\mathfrak{A}_n$ 的单性讨论 $H \cap \mathfrak{A}_n$.

45. (60 阶单群的唯一性) 设 G 为单群, $|G| = 60$.

(i) 说明不存在满足 $(G : K) \leq 4$ 的真子群 K.

提示〉 考虑 G 对 G/K 左乘作用诱导的同态 $G \to \mathfrak{S}_{(G:K)}$.

(ii) 基于 §11.6 介绍的 Sylow 定理, 证明 G 必有满足 $|H| = 12$ 的子群 H.

提示〉 取 G 的 Sylow 2-子群 S. 所有 Sylow 2-子群的个数 N_2 也等于 $(G : N_G(S))$, 故 $N_2 \in \{1, 3, 5, 15\}$. 先用 (i) 排除 1, 3; 若 $N_2 = 5$ 则取 $H = N_G(S)$.

考虑 $N_2 = 15$ 的情形. 用 (i) 讨论 Sylow 5-子群的个数 N_5, 通过计数说明存在 Sylow 2-子群 $S_1 \neq S_2$ 使得 $S_1 \cap S_2 \neq \{1\}$. 论证 $|Z_G(S_1 \cap S_2)| \in \{12, 20\}$.

(iii) 从 (i) 推导 $G \simeq \mathfrak{A}_5$.

提示〉 考虑 G 对 G/H 左乘诱导的同态 $G \to \mathfrak{S}_5$.

(iv) 作为推论, 证明 $\mathrm{PSL}(2, \mathbb{F}_5) \simeq \mathfrak{A}_5$.

事实上, 基于 Sylow 定理可以证明阶数 < 60 的单群必为素数阶循环群.

46. 记空间 $\mathbb{R}^3$ 中的正二十面体的旋转对称群为 $\mathbf{I}$ (见 §11.11). 如果球面上两点 x 和 y 落在同一条直径上, 则称它们互为对径点.

(i) 说明 $\mathbf{I}$ 是以下 5 个共轭类的并.

转轴	转角	元素个数
任意	0	1
对棱中点连线	π	15
对面重心连线	$\frac{\pm 2\pi}{3}$	20
顶点及其对径点连线	$\frac{\pm 2\pi}{5}$	12
	$\frac{\pm 4\pi}{5}$	12

末两个共轭类区分如下: 第一类的元素将某些面映到相邻面, 第二类则无此现象.

(ii) 以此证明 $\mathbf{I}$ 是单群.

提示〉 若 $N \lhd \mathbf{I}$, 则 $|N|$ 整除 60, 而且 N 是一些共轭类的并.

47. 设正多面体 P 的对称中心是原点, 按以下方式描述其对称群 $\mathrm{symm}(P)$:

▷ **正四面体** 同构于 $\mathfrak{S}_4 \simeq \mathbf{T} \rtimes (\mathbb{Z}/2\mathbb{Z})$ (试写下相应的同态 $\mathbb{Z}/2\mathbb{Z} \to \mathrm{Aut}(\mathbf{T})$),

▷ **立方体/正八面体** 同构于 $\mathbf{O} \times (\mathbb{Z}/2\mathbb{Z})$,

▷ **正十二面体/正二十面体** 同构于 $\mathbf{I} \times (\mathbb{Z}/2\mathbb{Z})$.

以此分类 $\mathrm{O}(3)$ 的有限子群, 精确到共轭.

提示〉 观察到 $(\mathrm{symm}(P) : \mathrm{symm}^+(P)) \in \{1, 2\}$. 当 P 为正四面体时, 直接构造保持 P 的镜射并验证 $\mathrm{symm}(P) \simeq \mathbf{T} \rtimes (\mathbb{Z}/2\mathbb{Z}) \simeq \mathfrak{S}_4$. 对于立方体, 易见取对径点给出 P 的对称性, 由此得到 $\mathbf{O} \times (\mathbb{Z}/2\mathbb{Z})$. 对于正十二面体或正二十面体, 取对径点仍给出 P 的对称性; 对此, 一种不甚优雅的办法是用 §11.11 脚注给出的顶点坐标.

48. 证明任何非空集 S 上都存在至少一种交换群结构.

提示〉 当 S 无穷时, 用集合论部分的最后一道习题说明 S 与某个 $\mathbb{F}_2$-向量空间等势.

第十二章 模论入门

模是向量空间的直接推广: 一个 F-向量空间对加法成群, 另有来自域 F 的纯量乘法, 而 R-模的概念则是在加法群的基础上考虑来自一个环 R 的纯量乘法, 依然要求有结合律与分配律. 模的定义和基本运算法则是 §12.1 的主题, 和向量空间大致类似, 但由于环 R 未必交换, 因此依照乘法的作用方向产生了左模, 右模乃至于本章习题介绍的双模等种种概念. 以下若无另外说明, 所论的 R-模 M 默认为左模, 带有加法 $M \times M \to M$ 和纯量乘法映射 $R \times M \to M$.

一如其它代数结构, §12.2 将对 R-模探讨同态和同构的概念. 此外模还可以对任意子模取商, 得到的商模和群论情形一样满足若干同态定理; 然而模的加法运算交换, 故相应的陈述比群简单. 向量空间的商空间是商模的特例. 关于直和与直积的定义也与向量空间的办法类似, 详见 §12.3.

模的实例包括交换群 (取 $R = \mathbb{Z}$), 域 F 上的向量空间 (取 $R = F$) 或者向量空间配上一个自同态 (取 $R = F[X]$); 最后一个例子是标准形理论的基石, 在第十三章将有深入讨论. 由此可见模论应用范围之广.

向量空间的许多基本概念 (如基和维数) 都能推及除环上的模, 但在除环之外则大有不同. 基的概念在模论中涉及 §12.4 探讨的自由模. 外在地看, 集合 X 上的自由 R-模可以定义为 X 份 R 的直和 (定义 12.4.1); 内在地看, 一个模 M 是否同构于自由模也可以从是否有基来刻画 (定义–命题 12.4.4). 并非所有模都是自由模. 交换环上的自由 R-模有良定义的秩 (定义–命题 12.4.9), 这是维数的推广.

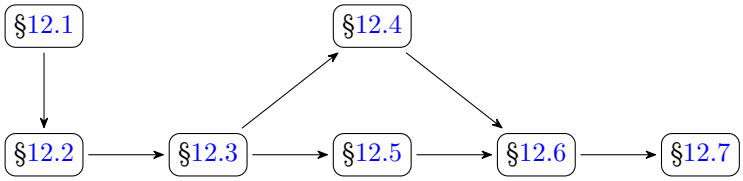
本章后续部分以主理想环上的有限生成模分类定理 12.6.3 为指归, 仅涉及交换环. 在 §12.5 的热身之后; §12.6 将对主理想环 R 上的有限生成模给出直和分解. 更确切地说, 分解有基于不变因子和初等因子的两种版本, 两族因子仅依赖于模的同构类, 容易相互过渡. 这部分内容主要是作为标准形理论的铺垫, 但是它也能用以分类有限生成交换群 (推论 12.6.5).

本章最后的 §12.7 从矩阵视角审视之前的理论, 主要结论是关于 $\mathbb{Z}$ 或 $F[X]$ 上的矩阵的 Smith 标准形定理 12.7.6, 其证明具有算法特质, 对于第十三章涉及的计算尤其实用. 事实上, §12.6 的主要结论可以用矩阵语言证明, 不过本书未选择这一进路.

阅读提示

由于本章主要为标准形理论服务, 关于模的介绍难免失之片面, 有心深究的读者应当参考 [10, 第六章] 或其它覆盖模论内容的代数类教材.

阅读顺序



12.1 模的基本定义

模是带有加法和来自某个环 R 的纯量乘法的一种代数结构. 由于已有群的概念, 更精确地说是约定 11.1.18 所谓的加法群, 在模的定义中可以节省力气.

定义 12.1.1 (模) 所谓左 R-模, 是指下述资料:

- 加法群 $(M,+)$,
- 映射 $R\times M\to M$, 以乘法符号记为 $(r,m)\mapsto r\cdot m=rm$, 也称为模的纯量乘法,

条件是要求以下性质成立:

$$\begin{aligned} r(m_1+m_2)&=rm_1+rm_2,\\ (r_1+r_2)m&=r_1m+r_2m,\\ (r_1r_2)m&=r_1(r_2m),\\ 1_Rm&=m, \end{aligned}$$

其中 $r,r_i\in R$ 而 $m,m_i\in M$. 按照惯例, 这组资料也常简记为 M.

类似手法可以定义右 R-模 M, 差别是纯量乘法此时是 $M\times R\to M$, 记为右乘 $(m,r)\mapsto m\cdot r=mr$, 而左模的条件 $(r_1r_2)m=r_1(r_2m)$ 相应地改写为

$$m(r_1r_2)=(mr_1)r_2.$$

若无其他说明, R-模在本书中默认为左 R-模.

注意到左 (或右) 模的最后两条公理相当于说纯量乘法给出幺半群 $(R,\cdot)$ 对 M 的左 (或右) 作用. 我们经常将 M 作为加法群的零元简记为 0, 这不会造成混淆.

练习 12.1.2 定义环 R 的相反环 R^{op}, 方法是修改 R 的乘法为 $\cdot^{\mathrm{op}}$, 使得 $x\cdot^{\mathrm{op}}y:=yx$. 因此 R 交换等价于 $R=R^{\mathrm{op}}$. 说明左 R-模和右 R^{op}-模是一回事.

因此在表述模的一般性质时考虑左模即可. 交换环上的模不必区分左右.

例 12.1.3 (零模) 最无聊的模是零模 $\{0\}$, 经常也简写为 0, 对应的加法群即零群 (或称平凡群), 纯量乘法当然是 $r \cdot 0 = 0$.

例 12.1.4 环 R 本身可以作成左 R-模, 其加法取为环的加法, 纯量乘法 $R \times R \to R$ 取为环的乘法; 模论的一切公理在此都由环的性质来保证. 同样地, R 本身也可以作成右 R-模. 对于 R 交换的情形, 请读者验证 R 的 R-子模无非是理想.

练习 12.1.5 推而广之, 考虑环同态 $\phi : R \to R'$, 验证 R' 按此自然地成为左 (或右) R-模, 其加法是环 R' 的加法, 纯量乘法以环 R' 的乘法表作 $(r, r') \mapsto \phi(r)r'$ (或 $(r', r) \mapsto r'\phi(r)$).

例 12.1.6 (向量空间) 设 F 为域, 关于 F-模的公理直接回归到 F-向量空间的公理, 因此模是向量空间这一概念的直接推广.

一如向量空间的情形, 模的纯量乘法具有以下性质:

$$\begin{aligned} 0_R \cdot m &= m, \\ (-1_R) \cdot m &= -m, \\ (n \cdot 1_R) \cdot m &= nm, \\ (-n \cdot 1_R) \cdot m &= -nm, \end{aligned}$$

其中 $m \in M$; 后两式的 n 属于 $\mathbb{Z}$, 而 $\pm nm \in M$ 则是置于加法群 $(M, +)$ 中理解的倍数运算. 相关论证与向量空间或环论的版本完全相同, 读者应该已有充足的经验与识见.

例 12.1.7 (交换群作为 $\mathbb{Z}$-模) 设 A 为交换群, 群运算表作加法. 已知 A 具有自然的整数倍运算

$$\mathbb{Z} \times A \to A, \quad (n, a) \mapsto na,$$

满足 $n(a_1 + a_2) = na_1 + na_2$, $(n_1 + n_2)a = n_1 a + n_2 a$, $(n_1 n_2)a = n_1(n_2 a)$ 和 $1 \cdot a = a$ 等标准性质, 这就说明此运算使 A 成 $\mathbb{Z}$-模.

反之设 A 为 $\mathbb{Z}$-模, 则对任何 $n \in \mathbb{Z}_{\geq 0}$ 皆有

$$na = (\underbrace{1 + \cdots + 1}_{n \text{ 份}})a = \underbrace{a + \cdots + a}_{n \text{ 份}}$$

和 $(-n)a = -(na)$, 这说明 A 的纯量乘法能且只能是它作为加法群的整数倍运算.

综上, 模论的解释能力又一次体现为:

$$\mathbb{Z}\text{-模} = \text{交换群}.$$

一如我们见过的种种代数结构, 模也有对应的子结构.

定义 12.1.8 (子模) 设 M 为 R-模. 如果 N 是 M 的加法子群, 而且对纯量乘法有如下的封闭性

$$r \in R,\ x \in N \implies rx \in N,$$

则称 N 为 M 的 R-子模, 简称子模. 类似定义对右模同样适用.

子模 $N \subset M$ 既然是 M 作为加法群的子群, 对之当然有陪集的概念, 用加法符号写作 $x+N$ 的形式 ($x \in M$).

若 $(M_i)_{i\in I}$ 是 M 的一族子模 ($I \neq \varnothing$), 则 $\bigcap_{i\in I} M_i$ 也是子模. 另一方面, 定义

$$\sum_{i\in I} M_i := \left\{ \text{有限和 } \sum_i x_i \in M : \forall i \in I,\ x_i \in M_i \right\};$$

一如向量空间的情形, 有限和意谓: 存在有限子集 $I_0 \subset I$, 使得 $i \notin I_0$ 时 $x_i = 0$. 这是 M 的子模. 当 $I = \varnothing$ 时, 方便地规定空和为零子模 $\{0\}$.

我们也同样能探讨一个子集 $S \subset M$ 生成的子模, 照例记为

$$\langle S\rangle := \left\{ \text{有限和 } \sum_{s\in S} r_s s : \forall s \in S,\ r_s \in R \right\},$$

另对有限子集的情形定义 $\langle s_1, \dots, s_n\rangle := \langle \{s_1, \dots, s_n\}\rangle$. 作为特例, 任意 $x \in M$ 皆生成子模

$$\langle x\rangle = Rx := \{rx \in M : r \in R\}.$$

至于一般的 S, 它生成的子模也写作

$$\langle S\rangle = \sum_{s\in S} Rs, \quad \langle s_1, \dots, s_n\rangle = Rs_1 + \cdots + Rs_n.$$

子模 $\langle S\rangle$ 是 M 中包含 S 的极小子模: 对任何子模 N 都有 $N \supset S \implies N \supset \langle S\rangle$. 于是

$$\langle S\rangle = \bigcap_{\text{子模 } N\supset S} N.$$

定义 12.1.9 设 M 为 R-模. 若存在 M 的有限子集 S 使得 $M = \langle S\rangle$, 则称 M 为**有限生成**的或**有限型**的.

顺势引入另一则定义, 它将在标准形的研究中扮演要角.

定义 12.1.10 若 R-模 M 能由单个元素生成, 亦即存在 $x \in M$ 使得 $M = Rx$, 则称 M 为**循环模**. 类似定义也适用于右 R-模.

举例明之, R 本身作为 R-模是循环模 (取 $x = 1_R$). 对于 $R = \mathbb{Z}$ 的情形, 循环 $\mathbb{Z}$-模在例 12.1.7 之下对应的正是循环群.

12.2 模的同态, 同构与商

本节选定环 R, 进一步探究保持 R-模结构的映射, 称为 R-模同态, 不致混淆时简称模同态或同态.

定义 12.2.1 (模同态) 设 M 和 M' 为 R-模, 所谓 R-模同态, 意谓满足下述条件的映射 $f : M \to M'$

▷ **保加法** f 是从加法群 $(M,+)$ 到 $(M',+)$ 的群同态, 亦即 $f(x+y) = f(x) + f(y)$ 对所有 $x, y \in M$ 成立;

▷ **保纯量乘法** $f(rx) = rf(x)$ 对所有 $r \in R$ 和 $x \in M$ 成立.

对右 R-模也有相似的定义.

同态的合成仍是同态. 取常值 0 的映射 $M \to M'$ 也是同态, 称为零同态.

定义 12.2.2 设 $f : M \to M'$ 为模同态. 如果存在模同态 $g : M' \to M$ 使得 $gf = \mathrm{id}_M$ 而 $fg = \mathrm{id}_{M'}$, 则称 f 为**模同构**, 简称同构, 而 g 为 f 的逆. 此时我们也说 M 和 M' 同构.

恒等映射 id_M 是同构. 同构的合成依然是同构.

关于逆的条件表明 g 作为映射是 f 的逆, 因而 f 和 g 都是双射. 和向量空间, 环与群的情形类似, 我们仍有以下事实.

命题 12.2.3 设 $f : M \to M'$ 为模同态. 如果 f 是集合之间的双射, 则 f 是同构.

证明 论证是熟悉的, 证 f 的逆映射 f^{-1} 也是同态即可. 首先, 加法群的情形 (命题 11.2.8) 确保 f^{-1} 已是群同态; 其次, 对 $f(rx) = rf(x)$ 两边取 f^{-1}, 并且记 $y := f(x)$, 可得 $rf^{-1}(y) = f^{-1}(ry)$. 由于所有 $y \in M'$ 都能写成 $f(x)$, 这表明 f^{-1} 也保纯量乘法. □

定义–命题 12.2.4 设 M 和 M' 为 R-模, 记所有同态 $f : M \to M'$ 构成的集合为 $\mathrm{Hom}_R(M, M')$ 或 $\mathrm{Hom}(M, M')$. 它具有加法群结构, 以零同态为零元, 具体方法是对 $f, g \in \mathrm{Hom}(M, M')$ 定义

$$(f+g)(x) = f(x) + g(x), \quad x \in M.$$

另外定义 $\mathrm{End}_R(M) := \mathrm{Hom}_R(M, M)$, 也记为 $\mathrm{End}(M)$; 它具有环结构, 加法是同态的加法, 乘法则是同态的合成, 以 id_M 为环的幺元. 右 R-模的情况完全相同.

证明 从 $(f+g)(x+y) = f(x+y) + g(x+y) = f(x) + f(y) + g(x) + g(y) = (f+g)(x) + (f+g)(y)$ 可见 $f+g$ 仍是同态; 结合律化到 M' 的层次来验证, 零同态作为零

元的性质则属显然. 同态 f 的加法逆元无非是 $(-f)(x) = -f(x)$, 请读者迅速验证 $-f$ 仍是同态.

关于 $\mathrm{End}(M)$ 成环的论断同样容易, 譬如分配律 $(f+g)h = fh + gh$ 的两边都映 $x \in M$ 为 $f(h(x)) + g(h(x))$. □

注记 12.2.5 (Hom-模) 若 R 是交换环, 则 $\mathrm{Hom}(M, M')$ 的加法群结构升级为 R-模, 方法是对所有 $r \in R$ 和 $f \in \mathrm{Hom}(M, M')$ 定义映射 $rf : M \to M'$ 如下

$$(rf)(x) := rf(x), \quad x \in M.$$

确实有 $rf \in \mathrm{Hom}(M, M')$, 原因是 $(rf)(x+y) = r(f(x)+f(y)) = rf(x) + rf(y)$ 和 $(rf)(r'x) = rr'f(x) = r'(rf)(x)$; 纯量乘法所需的性质易化到 M' 的 R-模结构. 应当注意到第二式基于 R 的交换性; 若 R 非交换, 则 $\mathrm{Hom}(M, M')$ 没有自然的 R-模结构.

对于 R 为域的情形, 向量空间之间的 Hom 因而有自然的向量空间结构, 这种构造是熟知的.

核的概念在模论中继续扮演重要角色.

定义 12.2.6 (模同态的核) 同态 $f : M \to M'$ 的**核**定义为

$$\ker(f) := \{x \in M : f(x) = 0\}.$$

容易看出 $\ker(f)$ 是 M 的子模, $\mathrm{im}(f)$ 是 M' 的子模. 既然模同态也是加法群的同态, 关于群的命题 11.9.9 立刻导致以下结果.

命题 12.2.7 设 $f : M \to M'$ 为模同态, 则 $f(x) = f(y)$ 当且仅当 $x - y \in \ker(f)$ (等价地说, 陪集 $x + \ker(f)$ 等于 $y + \ker(f)$). 因此 f 是单射当且仅当 $\ker(f) = \{0\}$.

模也有商的概念, 而且由于模论中不需要类似于正规子群的概念, 商模的构造比商群简单不少.

定义–命题 12.2.8 (商模) 设 N 为 M 的子模. 在加法商群 M/N 上定义纯量乘法映射

$$\begin{aligned} R \times (M/N) &\to M/N \\ (r, x+N) &\mapsto rx + N, \end{aligned}$$

这只与 r 和陪集 $x+N$ 相关, 无关元素 x 的选取.

(i) 此运算使加法群 M/N 成模.

(ii) 商映射 $q : M \to M/N$ 是模同态, $\ker(q) = N$.

模 M/N 称为 M 对 N 的商模.

证明　设 $x+N=x'+N$, 则存在 $y\in N$ 使得 $x'=x+y$, 于是 N 对纯量乘法和加法封闭导致

$$rx'+N=rx+\underbrace{ry}_{\in N}+N=rx+N.$$

纯量乘法所需的结合律等性质按此化到 M 的层次处理. 综上, M/N 成模. 根据商群的构造, q 是加法群的同态, 它作为群同态的核为 N, 而 q 保纯量乘法这一断言相当于说

$$q(rx)=rx+N=r(x+N)=rq(x),\quad r\in R,\quad x\in M.$$

然而上式正由 M/N 的纯量乘法定义确保. □

以上论证也说明 M/N 的模结构是使 q 成为模同态的唯一选择, 而 q 也应当和 M/N 一道视为商模结构的一员.

由于模同态 f 的像仍是子模, 对之取商便给出核的对偶版本, 这是群论场景所无的新操作.

定义 12.2.9 (模同态的余核)　*模同态 $f:M\to M'$ 的余核定义为商模 $M'/\operatorname{im}(f)$, 记为* $\operatorname{coker}(f)$.

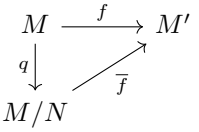
核能判定 f 的单性, 余核则能判定满性.

命题 12.2.10　*模同态 $f:M\to M'$ 为满当且仅当* $\operatorname{coker}(f)=\{0\}$.

证明　满性相当于说 $\operatorname{im}(f)=M'$, 这也等价于说 $\operatorname{im}(f)$ 只有一个陪集 $0+\operatorname{im}(f)=\operatorname{im}(f)$, 等价于 $\operatorname{coker}(f)$ 是零模. □

以下结果和 §11.9 平行, 而且陈述或论证都有所简化, 所以只做简单勾勒. 继续选定环 R, 同态均指模同态.

命题 12.2.11　*设 $f:M\to M'$ 为同态, M 的子模 N 包含于* $\ker(f)$, *则存在唯一的同态 $\overline{f}:M/N\to M'$ 使得 $f=\overline{f}q$, 或以交换图表写作*



交换.

这般的 $\overline{f}$ 称为 f 所诱导的同态. 它满足 $\operatorname{im}(\overline{f})=\operatorname{im}(f)$.

证明　在加法群的层次, 群同态 $\overline{f}$ 的存在性, 唯一性以及 $\operatorname{im}(\overline{f})=\operatorname{im}(f)$ 都来自命题 11.9.12; 留意到加法群的子群总是正规的.

问题因此归结为证明 $\overline{f}(x+N)=f(x)$ 描述的 $\overline{f}$ 满足 $\overline{f}(rx+N)=r\overline{f}(x+N)$, 然而这不过是说 $f(rx)=rf(x)$. 证毕. □

推论 12.2.12 设 $f: M \to M'$ 为同态, $N \subset M$ 和 $N' \subset M'$ 是子模, 而 $f(N) \subset N'$. 存在唯一的同态 $\overline{f}: M/N \to M'/N'$ 使下图交换

$$\begin{array}{ccc} M & \xrightarrow{f} & M' \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q'} \\ M/N & \xrightarrow[\overline{f}]{} & M'/N' \end{array}$$

此处 q 和 q' 代表商同态. 具体地说, $\overline{f}(x+N) = f(x) + N'$.

证明 由于条件导致 $q'f(N) = \{0\}$, 对 $q'f: M \to M'/N'$ 应用前一则结果即可得到 $\overline{f}$ 的存在性和唯一性. 具体描述从 $q'f = \overline{f}q$ 推得. □

代入特例 $N := \ker(f)$, 便得到诱导同态 $\overline{f}: M/\ker(f) \to M'$.
诱导同态具有以下简单性质.

- ⋆ 对于同态 $f, g: M \to M'$, 在 $f(N), g(N) \subset N'$ 的前提下有 $\mathrm{Hom}(M/N, M'/N')$ 中的等式
$$\overline{f+g} = \overline{f} + \overline{g};$$
两边都映 $x+N$ 为 $f(x)+g(x)+N'$.

- ⋆ 若有同态 $L \xrightarrow{g} M \xrightarrow{f} N$ 和各自的子模 L', M' 和 N', 使得 $g(L') \subset M'$ 而 $f(M') \subset N'$, 则 $\overline{fg} = \overline{f} \cdot \overline{g}$. 这点既可以从之前刻画诱导同态的交换图表来检验, 也可以直接从 $\overline{f}$ 和 $\overline{g}$ 的具体映法得见.

命题 12.2.13 设 $f: M \to M'$ 为群同态, 则诱导同态 $\overline{f}: M/\ker(f) \to M'$ 给出同构 $M/\ker(f) \xrightarrow{\sim} \mathrm{im}(f)$.

证明 在加法群层次, 命题 11.9.14 给出群同构 $\overline{f}: M/\ker(f) \xrightarrow{\sim} \mathrm{im}(f)$; 特别地, 这是双射. 然而已知 $\overline{f}$ 是模同态, 故 $\overline{f}: M/\ker(f) \xrightarrow{\sim} \mathrm{im}(f)$ 在模的层次同样成立. □

商模的构造因而为满同态 $f: M \twoheadrightarrow M'$ 给出了一种只涉及 M 的内在描述. 任何满同态本质上都是取商.

命题 12.2.14 设 $f: M \to M'$ 为满同态, 则有双射

$$\{\text{子模 } N' \subset M'\} \xleftrightarrow{1:1} \{\text{子模 } N \subset M : N \supset \ker(f)\}$$

$$N' \longmapsto f^{-1}(N')$$

$$f(N) \longmapsfrom N.$$

此双射满足以下性质:

- ⋆ $N_1' \subset N_2' \iff f^{-1}(N_1') \subset f^{-1}(N_2')$,
- ⋆ 若 $N' \subset M'$ 对应到 $N \subset M$, 则合成同态 $M \xrightarrow{f} M' \xrightarrow{\text{商}} M'/N'$ 诱导出同构 $M/N \xrightarrow{\sim} M'/N'$,
- ⋆ 当 f 取为对某个 $N \subset M$ 的商同态 $M \to M/N$ 时, 上述同构进一步改写为“分母相消”的形式

$$M/L \xrightarrow{\sim} (M/N)\big/(L/N),$$

其中 $N \subset L \subset M$.

证明 在加法群的层次, 所求的对应无非是命题 11.9.17 内容. 为了加进 R-模结构, 只须说明:

- ⋆ 若 N' 是 M' 的子模, 则 $f^{-1}(N')$ 是 M 的子模;
- ⋆ 若 N 是 M 的子模, 则 $f(N)$ 是 M' 的子模.

验证当然是容易的. 此外, 断言中涉及的 $M/N \xrightarrow{\sim} M'/N'$ 和 $M/L \xrightarrow{\sim} (M/N)\big/(L/N)$ 在加法群层次同样已知. 它们都由模同态所诱导, 因此自动地升级为模的同构. □

接着考虑一个给定的模 $\mathcal{M}$, 探讨其子模 $M, N \subset \mathcal{M}$ 的和 $M+N$ 与商的关系. 相较于群的版本, 此处不需要正规性条件.

命题 12.2.15 设 $M, N \subset \mathcal{M}$ 为子模, 则有同构

$$\begin{aligned} M/(M\cap N) &\xrightarrow{\sim} (M+N)/N \\ x+(M\cap N) &\longmapsto x+N. \end{aligned}$$

证明 定义同态 $f : M \to (M+N)/N$ 为包含同态 $M \hookrightarrow M+N$ 与商同态 $M+N \to (M+N)/N$ 的合成; 具体地说, $f(x) = x+N$, 其中 $x \in M$. 它限制在 $M \cap N$ 上为零同态, 故诱导 $\overline{f} : M/(M\cap N) \to (M+N)/N$, 映 $x+(M\cap N)$ 为 $x+N$.

在加法群的层次, 命题 11.9.19 说明 $\overline{f}$ 是群同构; 特别地, $\overline{f}$ 是双射. 这就说明 $\overline{f}$ 也是模的同构. □

对于 R 为域的情形, 本节的结论在讨论向量空间的商空间时已有介绍, 它们对向量空间理论的妙用历历在目.

练习 12.2.16 设 R 为交换环, 将其视同 R-模, 而 $I, I' \subset R$ 为理想. 说明存在模同构 $R/I \simeq R/I'$ 当且仅当 $I = I'$.

提示〉 对所有 R-模 M 定义 $\operatorname{ann}(M) := \{t \in R : \forall x \in M,\ tx = 0\}$. 论证 $\operatorname{ann}(R/I) = I$.

12.3 直和分解

本节选定环 R, 所论的模默认为左 R-模, 但一切结果都有右模版本.

首先, 回忆到根据定义–命题 11.1.22, 一族交换群 $(A_i)_{i\in I}$ 的积集 $\prod_{i\in I} A_i$ 具有自然的群结构, 称为它们的直积; 这依然是交换群, 它们的运算可写作加法.

定义 12.3.1 设 $(M_i)_{i\in I}$ 为一族 R-模, 在它们作为加法群的直积 $\prod_{i\in I} M_i$ 上定义纯量乘法

$$\begin{aligned} R\times\textstyle\prod_{i\in I} M_i &\longrightarrow \textstyle\prod_{i\in I} M_i,\\ (r,(x_i)_{i\in I}) &\longmapsto (rx_i)_{i\in I}, \end{aligned}$$

这使得 $\prod_{i\in I} M_i$ 成为 R-模, 称为 $(M_i)_{i\in I}$ 的**直积**.

定义 $(M_i)_{i\in I}$ 的**直和**为 $\prod_{i\in I} M_i$ 的如下子模

$$\bigoplus_{i\in I} M_i := \left\{ (x_i)_{i\in I} \in \prod_{i\in I} M_i : \text{仅有至多有限个 } i \text{ 使得 } x_i\neq 0 \right\}.$$

所需的模论公理容易验证, 不在话下.

当 $I=\varnothing$ 时, 规定直积与直和为零模 $\{0\}$ 是方便的. 观察到当 I 有限时直和等于直积. 有限个模 $M_1, M_2, \dots$ 的直和也写作 $M_1\oplus M_2\oplus\cdots$ 之形. 直和与直积依然有表为模同构的结合律, 譬如

$$M_1\oplus(M_2\oplus M_3)\simeq M_1\oplus M_2\oplus M_3\simeq (M_1\oplus M_2)\oplus M_3,$$

方法当然是等同 $(x_1,(x_2,x_3))$, (x_1,x_2,x_3) 与 $((x_1,x_2),x_3)$.

约定 12.3.2 若在定义中取每个 M_i 为同一个模 M, 对应的直积 (或直和) 将记为 M^I (或 $M^{\oplus I}$). 对 $n\in\mathbb{Z}_{\geq 0}$ 记 $M^{\oplus n} := \underbrace{M\oplus\cdots\oplus M}_{n\text{ 份}}$.

定义 12.3.3 对每个 $i\in I$, 从 M_i 到直和 $\bigoplus_{j\in I} M_i$ 的包含或嵌入定义为单同态

$$\begin{aligned} M_i &\xrightarrow{\iota_i} \textstyle\bigoplus_{j\in I} M_j\\ x &\longmapsto (x_j)_{j\in I}, \end{aligned} \qquad x_j := \begin{cases} x, & j=i,\\ 0, & j\neq i; \end{cases}$$

从直积 $\prod_{j\in I} M_j$ 到 M_i 的投影则定义为满同态

$$\begin{aligned} \textstyle\prod_{j\in I} M_j &\xrightarrow{p_i} M_i\\ (x_j)_{j\in I} &\longmapsto x_i. \end{aligned}$$

因此我们可以通过 ι_i 将每个 M_i 等同于直和 $\bigoplus_{j\in I} M_j$ 的子模.

观察到直和 $\bigoplus_{i\in I} M_i$ 是所有子模 M_i 的和, 直积则不然, 除非 I 有限. 后续的讨论将专注于直和, 因为它比直积更常用.

以上从一族模 $(M_i)_{i\in I}$ 出发, 抽象地构造直和 $M := \bigoplus_{i\in I} M_i$ 并将 M_i 嵌入其中, 这种构造是外在的. 从内而观之, 我们也可以从给定的模 M 和它的一族子模 $(M_i)_{i\in I}$ 出发, 探讨 M 能否分解为这些子模的直和, 从而简化关于 M 的研究. 何谓分解为子模的直和? 思路和向量空间的情形是一贯的.

首先, 给定 M 的一族子模如上, 可以定义同态 $\sigma : \bigoplus_{i\in I} M_i \to M$, 使得 $\sigma((x_i)_{i\in I}) = \sum_i x_i$; 注意到右边总是有限和. 这是沟通内外的桥梁.

定义–命题 12.3.4 设 $(M_i)_{i\in I}$ 为模 M 的一族子模. 以下陈述等价:

(i) $\sum_{i\in I} M_i = M$, 而且对每个 $i \in I$ 都有

$$M_i \cap \sum_{j\neq i} M_j = \{0\};$$

(ii) 每个 $x \in M$ 都能唯一地写成 $x = \sum_{i\in I} x_i$, 其中 $x_i \in M_i$, 至多有限项非零;

(iii) σ 给出同构 $\bigoplus_{i\in I} M_i \xrightarrow{\sim} M$.

当以上任一条件成立时, 我们也称 M 为其子模族 $(M_i)_{i\in I}$ 的直和, 写作 $M = \bigoplus_{i\in I} M_i$, 而每个 M_i 皆称为其中的**直和项**.

证明 (i) $\Longrightarrow$ (ii). 由 $M = \sum_{i\in I} M_i$ 可知每个 $x \in M$ 都能表为有限和 $\sum_{i\in I} x_i$. 唯一性等价于说 $\sum_{i\in I} x_i = 0 \iff \forall i\ x_i = 0$, 然而 $\sum_{j\in I} x_j = 0$ 蕴涵

$$x_i = -\sum_{j\neq i} x_j \in M_i \cap \sum_{j\neq i} M_j = \{0\}$$

对所有 $i \in I$ 成立, 故表法确实唯一.

(ii) $\iff$ (iii). 每个 x 都能表为 $\sum_{i\in I} x_i$ 相当于说 σ 满, 表法中的 $(x_i)_{i\in I}$ 唯一相当于说 σ 单.

(ii) $\Longrightarrow$ (i). 将 $x \in M$ 表为 $x = \sigma((x_i)_i) = \sum_i x_i$ 可见 $M = \sum_i M_i$. 现在固定 $i \in I$ 和 $x_i \in M_i$, 若有 $x_i \in \sum_{j\neq i} M_j$, 则存在 $(x_j)_{j\neq i} \in \bigoplus_{j\neq i} M_j$ 使得 $x_i = \sum_{j\neq i} x_j$, 这种表法的唯一性蕴涵 $x_i = 0$, 因为等式右侧不含 i 项. □

例 12.3.5 给定一族 R-模 $(M_i)_{i\in I}$. 若将每个 M_i 通过 ι_i 等同于外在地构造的 $\bigoplus_{j\in I} M_j$ 的子模, 则条件 (ii) 对 $M := \bigoplus_{i\in I} M_i$ 的这族子群显然成立, 所以 M_i 实现为 $\bigoplus_{j\in I} M_j$ 的直和项. 此时的 σ 化为 $\bigoplus_i M_i$ 上的恒等映射.

严格的说法应当是将 $\bigoplus_{i\in I} M_i$ 称为一族模 $(M_i)_{i\in I}$ 的外直和, 而将模 M 连同满足定义–命题 12.3.4 的等价条件的一族子模 $(M_i)_{i\in I}$ 所给出的分解称为内直和. 基于前一段的讨论和 (iii) 的同构, 内外两种直和一般不必再区分, 符号的重叠也不会造成混淆.

相对于直和项的商模有简单明了的描述.

命题 12.3.6 设 L 和 N 为 M 的子模, 使得 $M = N \oplus L$, 则商同态 $q : M \to M/N$ 限制为同构 $q|_L : L \xrightarrow{\sim} M/N$.

证明 我们有 $M = L + N$ 和 $L \cap N = \{0\}$. 于是命题 12.2.15 给出同构

$$\begin{aligned} L = L/(L \cap N) &\xrightarrow{\sim} (L+N)/N = M/N \\ x &\mapsto x + N, \end{aligned}$$

其中 $x \in L$, 然而 $x \mapsto x + N$ 也正是 $q|_L$ 的映法. □

并非对所有子模 $N \subset M$ 都能找到 L 使得 $M = N \oplus L$; 例如可取 $R = \mathbb{Z}$, $M = \mathbb{Z}/4\mathbb{Z}$ 和 $N = 2\mathbb{Z}/4\mathbb{Z}$. 换言之, 并非所有子模都能实现为直和项. 这点和向量空间的情形迥异.

最后, 从直和到直积的 Hom 集有和向量空间情形相同的描述, 亦即

$$\mathrm{Hom}(\bigoplus_j M_j, \prod_i N_i) \simeq \prod_{i,j} \mathrm{Hom}(M_j, N_i);$$

论证无异, 谨留作本章的简单习题.

练习 12.3.7 考虑一族模 $(M_i)_{i\in I}$, 对每个 $i \in I$ 设有 M_i 的子模 N_i. 明确写下并验证模同构

$$\bigoplus_{i\in I} M_i \Big/ \bigoplus_{i\in I} N_i \xrightarrow{\sim} \bigoplus_{i\in I} M_i/N_i.$$

12.4 自由模

直和在本书中衍生的第一个概念是自由模. 本节选定环 R 并考虑左 R-模. 回忆到 R 本身通过环的乘法成为左 R-模.

定义 12.4.1 设 X 为集合, 其上的**自由模**定义为直和 $R^{\oplus X}$. 我们可以将 X 自然地嵌入为 $R^{\oplus X}$ 的子集, 方法是让 $x \in X$ 对应到 $(r_y)_{y\in X} \in R^{\oplus X}$, 其中

$$r_y := \begin{cases} 1, & y = x, \\ 0, & y \neq x. \end{cases}$$

按此将 X 视同自由模 $R^{\oplus X}$ 的子集, 则 Rx 便是第 x 份 R 给出的直和项, 而 $R^{\oplus X} = \bigoplus_{x\in X} Rx$. 每个 $m \in R^{\oplus X}$ 都有唯一表法 $m = \sum_{x\in X} r_x x$ (有限和).

按照关于直和的规定, $X = \varnothing$ 对应的自由模等于零模 $\{0\}$.

命题 12.4.2 设 N 为任意 R-模, 则有双射

$$\begin{aligned}\mathrm{Hom}(R^{\oplus X}, N) &\xleftrightarrow{1:1} \{\text{映射 } f: X \to N\}\\ \varphi &\longmapsto \varphi|_X\\ \left[\textstyle\sum_{x\in X} r_x x \mapsto \sum_{x\in X} r_x f(x)\right] &\longmapsfrom f\end{aligned}$$

证明 由于 $R^{\oplus X} = \bigoplus_{x\in X} Rx$ 的所有元素都能唯一地表为 $\sum_x r_x x$, 其中 $r_x \in R$, 至多有限项非零, 故 $\sum_{x\in X} r_x x \mapsto \sum_{x\in X} r_x f(x)$ 确实定义了模同态. 剩下工作是验证双向互逆, 全是例行公事. □

上述构造从给定的集合 X 构造自由模 $R^{\oplus X}$ 连同集合的嵌入 $X \hookrightarrow R^{\oplus X}$. 遵循熟悉的思路, 这种构造可谓是外在的, 而内在观点则是从给定的模 M 连同子集 $X \subset M$ 出发, 寻求将 M 等同于 $R^{\oplus X}$ 的条件. 答案是容易的.

定义 12.4.3 设 M 为 R-模, S 为 M 的子集. 如果有限和 $\sum_{s\in S} r_s s$ (其中 $r_s \in R$) 为零当且仅当所有 r_s 全为 0, 则称 S 为线性无关子集.

类似地, 如果 $x_1, \dots, x_n \in M$ 满足 $\sum_{i=1}^n r_i x_i = 0 \iff r_1 = \cdots = r_n = 0$, 则称列 $x_1, \dots, x_n \in M$ 线性无关.

现在可以给出自由模的内在描述. 对于任意子集 $X \subset M$, 在命题 12.4.2 双射的右侧取包含映射 $f: X \to M$, 相应地便得到模同态 $\varphi: R^{\oplus X} \to M$, 它限制到 $X \subset R^{\oplus X}$ 上无非是 f.

定义–命题 12.4.4 设 X 为模 M 的子集. 以下陈述等价:

(i) X 生成 M, 而且线性无关;

(ii) 每个 $m \in M$ 都能唯一地表为有限和 $m = \sum_{x\in X} r_x x$, 其中 $r_x \in R$;

(iii) $\varphi: R^{\oplus X} \to M$ 为同构.

当以上任一条件成立时, 我们也称 M 是以 X 为**基**的自由模.

证明 (i) $\iff$ (ii). 生成条件相当于说每个 m 都能表为有限和 $\sum_{x\in X} r_x x$, 相减可知表法唯一相当于说 $\left(\sum_{x\in X} r_x x = 0 \iff \forall x\ r_x = 0\right)$, 然而这也等价于 X 线性无关.

(ii) $\iff$ (iii). 映射 φ 满相当于说每个 m 都能表为 $\sum_{x\in X} r_x x$, 单相当于表法中的系数 r_x 唯一. □

上述条件 (ii) 对于外在地构造的 $X \hookrightarrow R^{\oplus X}$ 显然是成立的, 对应的 $\varphi: R^{\oplus X} \to R^{\oplus X}$ 化为恒等. 所以关于自由模与基的内在和外在版本不相混淆.

若模 M 含有某个线性无关的生成集 X, 则我们也迳称 M 为自由模.

练习 12.4.5 考虑一族自由模 $(M_i)_{i\in I}$, 每个 M_i 都带有给定的基 X_i, 说明无交并 $\bigsqcup_{i\in I} X_i$ 给出 $\bigoplus_{i\in I} M_i$ 的基, 使之自由. 〈提示〉 参照向量空间的情形.

练习 12.4.6 说明自由模 $R^{\oplus X}$ 是有限生成的当且仅当 X 有限. 〈提示〉 每个生成元都只有有限多个分量非零.

例 12.4.7 对于 R 为域, 亦即向量空间的情形, 关于生成元, 线性无关子集和基的定义即刻化为熟悉的定义. 向量空间总有基, 但这点对于一般的环 R, 甚至对于整环上的模都不再成立. 只要忆及向量空间情形的证明依赖于除法, 便可以发现难点所在.

例 12.4.8 设 R 为定义 3.1.11 所谓的整环, M 为 R-模. 若 $x \in M$ 满足 $rx = 0 \iff r = 0$, 则称 x **无挠**, 否则称 x 为**挠元**. 自由模 $M = R^{\oplus X}$ 没有非零的挠元, 这是因为对于所有 $r \in R$ 和 $\sum_x r_x x \in R^{\oplus X}$, 我们有

$$\sum_{x\in X} rr_x x = 0 \iff \forall x \in X,\ rr_x = 0,$$

而整环性质导致右式仅在 r_x 全为 0 或者 $r = 0$ 方能成立. 这点可以用来给出许多非自由模.

- ⋆ 如果 $M = R/I$, 其中 I 是非零真理想, 则 $1 + I \in M$ 是非零挠元, 因为它被 I 的所有元素零化; 于是 R/I 非自由. 推而广之, 若 M 包含同构于 R/I 的子模, 便不是自由模.
- ⋆ 非零元无挠不能反推自由. 举例明之, $\mathbb{Q}$ 构成 $\mathbb{Z}$-模, 以 $\mathbb{Q}$ 中的乘法为纯量乘法, 它不含非零挠元, 但不是自由的: 所有 $q \in \mathbb{Q}$ 都能写成 $q = 2q'$ 的形式, 但形如 $\mathbb{Z}^{\oplus X}$ 的自由 $\mathbb{Z}$-模 (X 是非空集) 则无此性质: $m = \sum_{x\in X} r_x x$ 可写成 $m = 2m'$ 的形式当且仅当 r_x 全为偶数. 所以 $\mathbb{Q} \not\simeq \mathbb{Z}^{\oplus X}$.

  尽管如此, $\mathbb{Q}$ 可以写成一列自由子 $\mathbb{Z}$-模的递增并, 例如

$$\mathbb{Q} = \mathbb{Z} \cup \frac{1}{2!}\mathbb{Z} \cup \cdots \cup \underbrace{\frac{1}{k!}\mathbb{Z}}_{\simeq\mathbb{Z}} \cup \cdots .$$

定义–命题 12.4.9 设 R 为交换环, M 为自由模, 则 M 的任两组基 X 和 Y 都满足 $|X| = |Y|$. 这些基的共同基数称为 M 的**秩**, 记为 $\mathrm{rk}(M)$.

立即的结论是两个自由模同构当且仅当它们等秩. 定义–命题 12.4.9 在一般情形的证明参见 [10, 命题 6.3.8]. 以下只对整环上的有限生成自由模给出论证.

证明 (R 为整环, M 有限生成的情形) 首先观察到 M 的基必然有限, 这是练习 12.4.6 的内容.

以下说明 $|Y| \le |X|$. 命 $K := \mathrm{Frac}(R)$ 和 $n := |X| \in \mathbb{Z}_{\ge 0}$, 将 M 通过基 X 等同于 $R^{\oplus n}$, 再将 $R^{\oplus n}$ 嵌入 K^n. 由 K-向量空间的理论可知对于任何 $y_1, \ldots y_{n+1} \in M \subset K^n$, 存在不全为零的 $a_1, \ldots, a_{n+1} \in K$ 使得 $\sum_{i=1}^{n+1} a_i y_i = 0$, 通分后不妨假设 $a_i \in R$. 于是 M 中任何 $n+1$ 个元素皆线性相关, 这说明 $|Y| \le |X|$.

根据问题的对称性, 同样有 $|X| \le |Y|$, 证毕. □

注记 12.4.10 (自由模之间的同态与矩阵) 一如向量空间的情形, 有限秩自由模之间的同态可以由 R 上的矩阵来表述, 右模版本对此是比较方便的. 具体地说, 将右 R-模 $R^{\oplus n}$ 的元素写作列向量, 其 n 个分量都是 R 的元素, 则从 $R^{\oplus n}$ 到 $R^{\oplus m}$ 的模同态可以等同于矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}\in \mathrm{M}_{m\times n}(R)$, 同态的映法是用 $\boldsymbol{A}$ 左乘

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \boldsymbol{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1+\cdots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n \end{pmatrix}.$$

这确实是右 R-模同态. 详细推导没有本质的困难, 留作本章的简单习题.

12.5 基于挠子模的分解

本节介绍的一类直和分解与挠子模的概念相关. 以下默认 R 为交换环, R-模不必区分左右. 相关结论将在 §12.6 用到.

定义 12.5.1 *对任意理想 $I\subset R$ 和 R-模 M, 定义 M 的 I-挠子模为*

$$M[I]:=\{x\in M:\forall t\in I,\ tx=0\}.$$

对于主理想 $I=(h)$ 的情形 ($h\in R$), 另记

$$M[h]:=M[(h)]=\{x\in M:hx=0\},$$

称之为 M 的 h-挠子模.

因此 $M[I]=\bigcap_{h\in I}M[h]$. 下列性质毫无困难, 交由读者练习.

$$\begin{aligned} I\supset J &\implies M[I]\subset M[J],\\ M[I]\cap M[J]&=M[I+J],\\ (\bigoplus_i M_i)[I]&=\bigoplus_i (M_i[I]),\\ M[0]=M,&\quad M[1]=\{0\},\\ h,k\in R,\ h\mid k &\implies M[h]\subset M[k]. \end{aligned}$$

在 R 为主理想环的情形, $(h)+(k)=(\gcd(h,k))$, 其中 gcd 代表 R 中的最大公因数, 精确到 $R^\times$; 关于交的性质化为

$$M[h]\cap M[k]=M[\gcd(h,k)].$$

命题 12.5.2 *设 R 为主理想环, M 为 R-模, 而且存在 $t\in R\smallsetminus\{0\}$ 使得 $M=M[t]$. 将 t 作不可约分解 $p_1^{a_1}\cdots p_n^{a_n}$, 其中 $p_1,\dots,p_n$ 是互不等价的素元, 则有直和分解*

$$M=\bigoplus_{i=1}^{n} M[p_i^{a_i}].$$

证明 由于 $M = M[t]$, 问题递归地化为证: 若 $a, b \in R \smallsetminus \{0\}$ 互素, 则 $M[ab] = M[a] \oplus M[b]$. 基于互素性质在主理想环中的刻画, 存在 $u, v \in R$ 使得 $ua + vb = 1$. 若 $x \in M[ab]$, 则 $x = uax + vbx \in M[b] + M[a]$. 另一方面, $\gcd(a, b) \sim 1$ 则导致 $M[a] \cap M[b] = M[1] = \{0\}$. □

研究线性映射的极小多项式时已见过类似的思路. 一般而言, 设 p 为主理想环 R 的素元, 则从 $M[p] \subset M[p^2] \subset \cdots$ 易见

$$M[p^\infty] := \bigcup_{n=1}^{\infty} M[p^n]$$

是 M 的子模. 在相同前提下, 命题 12.5.2 更简洁地表为

$$M = \bigoplus_{p:\text{素元}\ /\sim} M[p^\infty]. \tag{12.5.1}$$

这是因为当 $n \gg 0$ 时 $M[p^n] = M[p^n] \cap M[t] = M[\gcd(p^n, t)] = M[p^a]$, 其中 $p^a \parallel t$. 子模 $M[p^\infty]$ 也称为 M 的 **p-准素部分**.

接着对一些基本情形确定 $M[p^\infty]$.

引理 12.5.3 *设 R 为主理想环, $t \in R \smallsetminus \{0\}$, 而 p 是 R 的素元. 考虑 $a \in \mathbb{Z}_{\geq 0}$.*

(i) *若 $p^a \mid t$ 则 $(R/(t))[p^a] \simeq R/(p^a)$;*

(ii) *若 $p^a \parallel t$ 则 $(R/(t))[p^\infty] \simeq R/(p^a)$.*

证明 设 $p^a \mid t$ 而 $x \in R$. 命 $s := t/p^a$. 条件 $x + (t) \in R/(t)[p^a]$ 等价于存在 $y \in R$ 使得 $p^a x = ty$; 两边消去 p^a 可知这等价于存在 y 使得 $x = sy$. 综上, $(R/(t))[p^a] = (s)/(t)$.

考虑从 R 到 $(s)/(t)$ 的模同态 $y \mapsto sy + (t)$, 它显然满, 核是 $\{y : t \mid sy\} = (p^a)$. 于是得到模同构 $R/(p^a) \xrightarrow{\sim} (s)/(t)$, 此即 (i).

既然 $(R/(t))[t] = R/(t)$, 先前已论证 $p^a \parallel t$ 蕴涵 $(R/(t))[p^\infty] = (R/(t))[p^a]$, 故 (ii) 成立. □

最后引进两则适用任意整环的概念.

定义 12.5.4 *设 R 为任意整环, M 为 R-模, 记 M 的所有挠元所成子集为 M_{tor}, 它是 M 的子模, 称为 M 的* **挠子模**. *商模 $M_{\mathrm{tf}} := M/M_{\mathrm{tor}}$ 称为 M 的* **无挠商**.

注意到 $M_{\mathrm{tor}} = \bigcup_{t \neq 0} M[t]$. 为了说明 M_{tor} 为子模, 设 $x, y \in M_{\mathrm{tor}}$, 取 $s, t \in R \smallsetminus \{0\}$ 使得 $sx = 0 = ty$, 则 $st \neq 0$ 而 $st(x + y) = 0$; 此外, 若 $r \in R$ 则 $s(rx) = r(sx) = 0$. 因此 M_{tor} 确实对加法和纯量乘法封闭. 无挠商一词的解释如下.

练习 12.5.5 证明 M_{tf} 无非零挠元.

提示〉 若 $x \in M$ 的像 $\bar{x} \in M_{\mathrm{tf}}$ 满足 $t\bar{x} = 0$, 则存在 $s \in R \smallsetminus \{0\}$ 使得 $stx = 0$.

命题 12.5.6 将整环 R 视同 R-模. 设 $x \in M \smallsetminus M_{\mathrm{tor}}$, 则有 R-模同构 $R \xrightarrow{\sim} Rx$.

证明 考虑由 $\varphi(r) = rx$ 确定的 R-模同态 $\varphi : R \to Rx$. 它当然满, 而 $r \in \ker(\varphi) \iff x \in M[r]$, 故 $x \notin M_{\mathrm{tor}}$ 蕴涵 $\ker(\varphi) = \{0\}$. □

以下是 (12.5.1) 的简单推广.

推论 12.5.7 设 R 为主理想环, R-模 M 满足 $M = M_{\mathrm{tor}}$, 则

(i) 对所有有限生成子模 $M_0 \subset M$ 皆存在 $t \in R \smallsetminus \{0\}$ 使得 $M_0 = M_0[t]$;

(ii) M 分解为准素部分的直和:

$$M = \bigoplus_{p:\text{素元}/\sim} M[p^\infty].$$

证明 对于 (i), 设子模 M_0 由 $x_1, \dots, x_n$ 生成, 对每个 i 取 $t_i \in R \smallsetminus \{0\}$ 使得 $t_i x_i = 0$, 再取 $t = t_1 \cdots t_n \neq 0$, 则 $tx_i = 0$ 对所有 $1 \leq i \leq n$ 成立. 故 $M_0 = M_0[t]$.

对于 (ii), 运用定义–命题 12.3.4 对直和的刻画. 给定 $x \in M$, 命 $M_0 = Rx$, 则 (i) 和 (12.5.1) 说明

$$x \in M_0 = \sum_p M_0[p^\infty] \subset \sum_p M[p^\infty].$$

另一方面, 设有互不等价的素元 $p_1, \dots, p_n$ 和等式 $\sum_{i=1}^n r_i x_i = 0$, 其中 $r_i \in R$ 而 $x_i \in M[p_i^\infty]$. 命 $M_0 = \sum_{i=1}^n Rx_i$, 于是 $x_i \in M_0[p_i^\infty]$, 从而 (i) 配合 (12.5.1) 对 M_0 给出的直和分解表明必有 $r_1 = \cdots = r_n = 0$. 综上知 $M = \bigoplus_{p:\text{素元}/\sim} M[p^\infty]$. □

12.6 主理想环上的有限生成模

本节的 R 默认为主理想环. 主角是有限生成 R-模, 但一些结论也适用于非有限生成模, 详阅 [10, §6.7].

引理 12.6.1 设 E 为秩 $n \in \mathbb{Z}_{\geq 0}$ 的自由 R-模, 则 E 的任意子模 N 仍是自由 R-模, 其秩 $\leq n$.

证明 对 n 递归. 由于 $n = 0$ 的情形平凡, 以下设 $n \geq 1$. 首先观察到若视 R 为 R-模, 则对于任意 R-模 M 和同态 $\lambda : M \to R$, 其像 $\lambda(M)$ 总是 R 的理想.

任取 E 的基 $f_1, \dots, f_n$, 定义 $E' := \bigoplus_{i=2}^n Rf_i$. 定义同态 $p_1 : E \to R$ 为第一个坐标投影 $\sum_{i=1}^n r_i f_i \mapsto r_1$, 则理想 $p_1(N)$ 可以表作 (d) 的形式, $d \in R$. 若 $d = 0$ 则 $N \subset E'$, 问题化到秩 $n-1$ 的已知情形.

以下假设 $d \neq 0$. 取 $x \in N$ 使得 $p_1(x) = d$. 对于任意 $y \in N$, 将其展开为 $y = \sum_{i=1}^n r_i f_i$, 则 $d \mid r_1$ 而

$$y = \frac{r_1}{d} \cdot x + \left(y - \frac{r_1}{d} \cdot x \right).$$

记 $N' := E' \cap N$, 上式表明 $N = Rx + N'$. 另一方面, 由于 f_1 在 x 中的系数 d 非零, 而 R 是整环, 故 $Rx \cap N' \subset Rx \cap E' = \{0\}$. 综上可得 $N = Rx \oplus N'$.

对 $N' \subset E'$ 应用秩 $n-1$ 的已知情形可得 N' 自由, 秩 $\leq n-1$. 又由 $x \neq 0$ 可知 x 无挠, 故 $Rx \simeq R$ (命题 12.5.6) 而 N 是秩 $\leq n$ 的自由模. □

引理 12.6.2 设 R 为主理想环, $\mathcal{S}$ 为 R 中的一族理想, $\mathcal{S} \neq \varnothing$. 按照理想的包含关系赋予 $\mathcal{S}$ 偏序, 则 $\mathcal{S}$ 必有极大元.

证明 若 $\mathcal{S}$ 无极大元, 则可以从 $\mathcal{S}$ 逐步选出一列严格递增的理想 $I_1 \subsetneq I_2 \subsetneq I_2 \subsetneq \cdots$, 然而主理想环的理想升链必然在有限步内停止, 矛盾. □

定理 12.6.3 (主理想环上的有限生成模分类) 设 R 为主理想环, M 为有限生成 R-模, 则有同构

$$M \simeq R/I_1 \oplus \cdots \oplus R/I_k \oplus E,$$

其中

- $k \in \mathbb{Z}_{\geq 0}$ 而 $I_1 \supset \cdots \supset I_k$ 是 R 中的一列非零真理想,
- E 是有限生成自由 R-模.

它们具有唯一性: 若有如上的分解

$$M \simeq R/I_1 \oplus \cdots R/I_k \oplus E, \quad M' \simeq R/I_1' \oplus \cdots R/I_{k'}' \oplus E',$$

则 $M \simeq M'$ 蕴涵 $\mathrm{rk}(E) = \mathrm{rk}(E')$, $k = k'$, 以及 $I_i = I_i'$ 对所有 i 成立.

上述分解中的真理想链

$$I_1 \supset \cdots \supset I_k \supset \underbrace{\{0\} = \cdots = \{0\}}_{\mathrm{rk}(E) \text{ 份}}$$

称为 M 的**不变因子**[1], 它们由 M 的同构类唯一确定.

证明比较曲折, 我们先处理存在性的部分, 论证的核心是以下结果, 它相当于说有限秩自由模 E 的子模 N 不仅自由, 还可以适当取基将 N 与 E 对齐.

引理 12.6.4 设 E 为主理想环 R 上的自由模, 秩为 $n \in \mathbb{Z}_{\geq 0}$, 而 N 为其子模, 则存在 E 的基 $f_1, \ldots, f_n$ 以及 R 的一列元素

$$d_1 \mid \cdots \mid d_n,$$

使得若取 $0 \leq r \leq n$ 使 $d_j = 0 \iff j > r$, 则元素 $d_1 f_1, \ldots, d_r f_r$ 构成 N 的基.

证明 不妨设 $n \geq 1$, 否则问题平凡. 论证分为三步.

[1] 包括 [10] 在内的部分书籍称之为初等因子, 但这容易和稍后要讨论的 $M[p^\infty]$ 的不变因子混淆.

▷ **析出**　存在 $f_1 \in E \smallsetminus \{0\}$ 和 $d_1 \in R$ 连同直和分解

$$E = Rf_1 \oplus E', \quad N = Rd_1 f_1 \oplus N',$$

其中 $N' := E' \cap N$;

▷ **递归**　若 $E' = \{0\}$ 则操作停止, 否则在前一步能够以 $N' \subset E'$ 代 $N \subset E$, 继续析出 $f_2 \in E'$ 和 $d_2 \in R$ 使得 $d_1 \mid d_2$, 依此类推;

▷ **终止**　上述操作在有限步之内停止, 给出所求的资料.

先处理析出步骤. 引理 12.6.2 确保理想族 $\{\lambda(N) : \lambda \in \mathrm{Hom}(E,R)\}$ 必有相对于 $\subset$ 的极大元, 记为 $\lambda_1(N)$; 取 $d_1 \in R$ 使得 $\lambda_1(N) = (d_1)$.

当 $d_1 = 0$ 时, 任取 E 的基 $f_1, \dots, f_n$, 对应的坐标投影 $p_1, \dots, p_n \in \mathrm{Hom}(E,R)$ 限制在 N 上全为零, 故此时 $N = \{0\}$. 取 $E' = \bigoplus_{i=2}^n Rf_i$, 所求的直和分解平凡地成立.

以下设 $d_1 \neq 0$, 取 $x_1 \in N$ 使得 $\lambda_1(x_1) = d_1$. 兹说明

$$\forall \lambda \in \mathrm{Hom}(E,R), \quad d_1 \mid \lambda(x_1). \tag{12.6.1}$$

这是因为理想 $(b) := (\lambda(x_1)) + (d_1)$ 包含 (d_1), 而且存在 $u, v \in R$ 使得

$$b = u\lambda(x_1) + vd_1 = u\lambda(x_1) + v\lambda_1(x_1) = \underbrace{(u\lambda + v\lambda_1)}_{\in \mathrm{Hom}(E,R)}(x_1),$$

故 $(u\lambda + v\lambda_1)(N) \supset (b) \supset (d_1)$, 而 (d_1) 的选法导致这些包含关系全为等号; 从 $(b) = (d_1)$ 推得 $(\lambda(x_1)) \subset (d_1)$.

现在让 λ 遍历 E 相对于某个基的坐标投影, 性质 (12.6.1) 蕴涵 x_1 的每个坐标都被 d_1 整除, 故存在 $f_1 \in E$ 使得 $x_1 = d_1 f_1$. 由于 $d_1 \neq 0$, 从 $d_1 = \lambda_1(x_1) = d_1\lambda_1(f_1)$ 推得 $\lambda_1(f_1) = 1$; 特别地 $f_1 \neq 0$. 以下验证

$$E = Rf_1 \oplus E', \quad E' := \ker(\lambda_1); \tag{12.6.2}$$
$$N = Rx_1 \oplus N', \quad N' := N \cap E'. \tag{12.6.3}$$

由 $\lambda_1(f_1) = 1$ 易见 $Rf_1 \cap E' = \{0\}$, 从而 $Rx_1 \cap N' = \{0\}$. 此外所有 $f \in E$ 都能表成

$$f = \lambda_1(f)f_1 + (f - \lambda_1(f)f_1) \in Rf_1 + E',$$

故 (12.6.2) 得证. 当 $f \in N$ 时 $\lambda_1(f)f_1 \in Rd_1f_1 = Rx_1 \subset N$, 上一步给出 $f \in Rx_1 + N'$, 故 (12.6.3) 得证.

接着处理递归步骤. 引理 12.6.1 说明 E' 是自由模. 在 $E' \neq \{0\}$ 的前提下, 同样操作可对 $N' \subset E'$ 迭代, 给出 $(d_2) = \lambda_2(N')$, 其中 $\lambda_2 \in \mathrm{Hom}(E', R)$. 兹断言

$$d_1 \mid d_2. \tag{12.6.4}$$

取 $x_2 \in N'$ 使得 $\lambda_2(x_2) = d_2$. 可将 λ_2 按直和 (12.6.2) 延拓为 $\tilde{\lambda}_2 \in \mathrm{Hom}(E, R)$, 使得 $\tilde{\lambda}_2(f_1) = 0$. 取 $u, v \in R$ 使得 $(ud_1 + vd_2) = (d_1) + (d_2)$. 不难验证

$$(u\lambda_1 + v\tilde{\lambda}_2)(x_1 + x_2) = ud_1 + vd_2.$$

因此 $(u\lambda_1 + v\tilde{\lambda}_2)(N) \supset (d_1) + (d_2) \supset (d_1)$, 而 (d_1) 的选法导致包含关系全是等号, 故 (12.6.4) 成立.

最后说明递归必终止. 这是因为 $f_1 \neq 0$ 蕴涵 f_1 无挠, $Rf_1 \simeq R$ (命题 12.5.6), 从而 (12.6.2) 蕴涵 $\mathrm{rk}(E') = \mathrm{rk}(E) - 1$, 故有限步之内止于零模. □

证明 (定理 12.6.3 的存在性部分) 不妨设 $M \neq \{0\}$. 选定 M 的生成元 $x_1, \ldots, x_n$ 和相应的满同态

$$E := R^{\oplus n} \twoheadrightarrow M, \quad (r_i)_{i=1}^n \mapsto \sum_{i=1}^n r_i x_i.$$

记其核为 N, 因此 $E/N \xrightarrow{\sim} M$. 对 $N \subset E$ 应用引理 12.6.4 得到 E 的基 $f_1, \ldots, f_n$ 和 $d_1 \mid \cdots \mid d_n$ 和 $0 \leq r \leq n$, 满足 $d_j = 0 \iff j > r$, 而且 $N = \bigoplus_{i=1}^r Rd_i f_i$. 由之推得

$$M \simeq E/N = \bigoplus_{i=1}^n Rf_i \Big/ \bigoplus_{i=1}^r Rd_i f_i \xrightarrow[\sim]{\text{练习 12.3.7}} \bigoplus_{i=1}^r (R/(d_i)) \oplus \underbrace{R^{\oplus n-r}}_{\text{自由}}.$$

舍去 $d_i \in R^\times$ 的项, 将剩下的主理想 (d_i) 标为 $I_1 \supset \cdots \supset I_k$ 便是. □

唯一性的证明需要 §12.5 的符号和结论.

证明 (定理 12.6.3 的唯一性部分) 第一步是化到 $M = M_{\mathrm{tor}}$ 情形, 第二步是化到 $M = M[p^\infty]$ 情形, 其中 p 是 R 的素元.

给定的直和分解限制为 $M_{\mathrm{tor}} \simeq R/I_1 \oplus \cdots \oplus R/I_k$, 从而命题 12.3.6 导致

$$M_{\mathrm{tf}} = (M_{\mathrm{tor}} \oplus E)/M_{\mathrm{tor}} \simeq E.$$

设有同构 $\varphi : M \xrightarrow{\sim} M'$, 而且 M' 带有类似的直和分解, 则 φ 限制为 $M_{\mathrm{tor}} \xrightarrow{\sim} M'_{\mathrm{tor}}$, 其诱导同构 $M_{\mathrm{tf}} \xrightarrow{\sim} M'_{\mathrm{tf}}$ 给出 $E \simeq E'$.

唯一性因此化简到 $M = M_{\mathrm{tor}}$ 的情形. 推论 12.5.7 (ii) 给出

$$M = \bigoplus_{p:\text{素元}/\sim} M[p^\infty].$$

将每个 I_i 写作 (t_i). 对于选定的素元 p (精确到等价) 和 $1 \leq i \leq k$, 取 $b_i = b_i(p) \in \mathbb{Z}_{\geq 0}$ 使得 $p^{b_i} \parallel t_i$. 于是

$$b_1 \leq b_2 \leq \cdots. \tag{12.6.5}$$

对每个 i 应用引理 12.5.3 (ii) 遂有

$$M[p^\infty] \simeq \bigoplus_{i=1}^k R/(t_i)[p^\infty] \simeq \bigoplus_{i=1}^k R/\left(p^{b_i}\right).$$

设有同构 $M \simeq M'$, 则它对每个 p 皆限制为 $M[p^\infty] \simeq M'[p^\infty]$, 而右式同样有分解

$$M'[p^\infty] \simeq \bigoplus_{i=1}^{k'} R/\left(p^{b'_i}\right).$$

留意到 b_i 和 b'_i 可能为零. 为了方便比较, 不妨在 M 或 M' 的直和分解中补零, 相应地从 (12.6.5) 的左端补零, 使得 $k = k'$. 为了证明 $I_i = I'_i$ 恒成立, 从而证明唯一性, 目标化为对每个 p 证明

$$b_i = b'_i, \quad 1 \leq i \leq k.$$

我们的目的是对于选定的素元 p, 将 $b_1, \ldots, b_k$ 诠释为 M 的同构不变量[2)], 尽管它们的初始定义依赖于直和分解 $M[p^\infty] \simeq \bigoplus_{i=1}^{k} R/(p^{b_i})$.

不失一般性, 今后可设 $M = M[p^\infty]$. 对引理 12.5.3 (i) 代入 $a = 1$ 可得

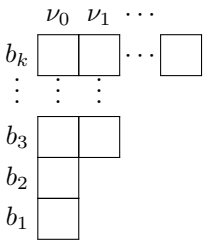
$$M[p] \simeq \bigoplus_i \left(R/(p^{b_i})\right)[p] \simeq \bigoplus_{i:b_i>0} R/(p).$$

由于 p 在这些 R-模上的纯量乘法皆为零, 这实则是 $R/(p)$-模的同构. 然而在本书稍早关于主理想环的讨论中已说明 p 是素元确保 $R/(p)$ 为域, 同构遂给出 $\nu_0 := |\{i : b_i > 0\}| = \dim_{R/(p)} M[p]$, 这是 M 的同构不变量.

推而广之, 对于所有 $j \geq 0$, 易见

$$p^j M \simeq \bigoplus_{i:b_i \geq j} p^j R/\left(p^{b_i}\right) \simeq \bigoplus_{i:b_i \geq j} R/\left(p^{b_i - j}\right),$$
$$\nu_j := |\{i : b_i > j\}| = \dim_{R/(p)} (p^j M)[p].$$

数列 $\nu_0 \geq \nu_1 \geq \cdots$ 是 M 的同构不变量, 又唯一确定 $b_1 \leq b_2 \leq \cdots$. 何以故? 设想在虚空之中从下而上逐行地放置 $b_1, b_2, \ldots$ 个方块, 如下图所示:



从左向右逐列扫描也能够重建图表, 读出的方块个数正是 $\nu_0, \nu_1, \ldots$. 明所欲证. □

上述论证提示我们: 为了解一个有限生成 R-模 $M = M_{\mathrm{tor}}$ 的结构, 既可以直接代入定理 12.6.3 按不变因子作分解, 也可以先将 M 分解为 p-准素部分 $M[p^\infty]$ 的直和,

[2)] 如果一个量 $I(M)$ 是由模 M 唯一确定的, 而且 $M \simeq M'$ 蕴涵 $I(M) = I(M')$, 则称 $M \mapsto I(M)$ 为模的同构不变量; 这相当于说 $I(M)$ 由 M 的内在结构决定.

再对每个 $M[p^\infty]$ 按不变因子作分解, 由此得到的资料

$$p^{b_1(p)}, p^{b_2(p)}, \dots, \quad p: \text{素元} / \sim, \\ b_1(p) \le b_2(p) \le \cdots$$

同样描述了 M 的等价类, 称为 M 的**初等因子**. 虽然两种分解写法不同, 却容易相互过渡.

为了阐明这点, 考虑素元 $p, q \in R$, 设 $p \not\sim q$. 取 R-模

$$M = \underbrace{\big(R/(p) \oplus R/(p^2)\big)}_{=M[p^\infty]} \oplus \underbrace{R/(q)}_{=M[q^\infty]}.$$

引理 12.5.3 (ii) 或主理想环上的中国剩余定理给出 $R/(p^2q) \simeq R/(p^2) \oplus R/(q)$, 故 $M \simeq R/(p) \oplus R/(p^2q)$, 这就转化为定理 12.6.3 的分解. 一般情形依此可知.

由于整数环 $\mathbb{Z}$ 是主理想环, 我们立刻得到下述应用.

推论 12.6.5 (有限生成交换群的结构定理) 设 A 为有限生成交换群, 群运算表作加法, 则有同构

$$A \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_k\mathbb{Z} \oplus \mathbb{Z}^{\oplus m},$$

其中 $k, m \in \mathbb{Z}_{\ge 0}$ 而 $d_1 \mid \cdots \mid d_k$ 是一列正整数, $d_1 > 1$.

它们具有唯一性: 若有如上的分解

$$A \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \mathbb{Z}/d_k\mathbb{Z} \oplus \mathbb{Z}^{\oplus m}, \quad A' \simeq \mathbb{Z}/d_1'\mathbb{Z} \oplus \cdots \mathbb{Z}/d_{k'}'\mathbb{Z} \oplus \mathbb{Z}^{\oplus m'},$$

则 $A \simeq A'$ 蕴涵 $m = m'$, $k = k'$, 以及 $d_i = d_i'$ 对所有 i 成立.

证明 以例 12.1.7 将交换群等同于 $\mathbb{Z}$-模, 然后在定理 12.6.3 中代入 $R = \mathbb{Z}$. □

定理 12.6.3 的唯一性陈述及其冗长论证形成了明显反差. 模的张量积理论能为唯一性提供直截了当的证明, 有兴趣的读者请参考 [10, §7] 习题.

12.7 基于矩阵的算法

结构定理 12.6.3 的存在性部分以引理 12.6.4 为基础. 为了更好地理解其构造, 同时为一类主理想环 R 提供算法, 本节采用矩阵语言加以改述.

约定 12.7.1 取定交换环 R 和 $n, m \in \mathbb{Z}_{\ge 1}$. 对于任意 R-模 M 中的两列元素

$$e_1, \dots, e_n, \quad x_1, \dots, x_m$$

以及 R 上的矩阵 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times m}(R)$, 引入形式的记法

$$\begin{pmatrix} x_1 \,\middle|\, \cdots \,\middle|\, x_m \end{pmatrix} = \begin{pmatrix} e_1 \,\middle|\, \cdots \,\middle|\, e_n \end{pmatrix} \boldsymbol{A} \tag{12.7.1}$$

来代表 $x_j = \sum_{i=1}^n a_{ij} e_i$ 对所有 $1 \le j \le m$ 成立[3)].

[3)] 尽管交换环上的模不必分左右, 此处更适合将 M 理解为右 R-模.

⋆ 虽然 x_j 和 e_i 一般而言并非 R 上的列向量, 但矩阵运算规律仍成立, 比如

$$\begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}(\boldsymbol{AB}) = \left(\begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}\boldsymbol{A}\right)\boldsymbol{B};$$

展开定义, 可见这不过是反映模的纯量乘法的结合律和分配律.

⋆ 此外, 若 $e_1,\dots,e_n$ 线性无关 (定义 12.4.3), 则 (12.7.1) 中的 $\boldsymbol{A}$ 是由 $(e_i)_{i=1}^n$ 和 $(x_j)_{j=1}^m$ 唯一确定的.

以下内容需要交换环上的行列式理论.

定义–命题 12.7.2 设 R 为交换环. 命 $\mathrm{GL}(n,R)$ 为所有满足 $\det \boldsymbol{P} \in R^\times$ 的 $\boldsymbol{P} \in \mathrm{M}_{n\times n}(R)$ 所成集合, 则有

$$\begin{aligned}\mathrm{GL}(n,R) &= \{\boldsymbol{P} \in \mathrm{M}_{n\times n}(R) : \exists \boldsymbol{Q},\ \boldsymbol{PQ} = \mathbf{1}_{n\times n} = \boldsymbol{QP}\}\\ &= \mathrm{M}_{n\times n}(R)^\times.\end{aligned}$$

作为推论, $\mathrm{GL}(n,R)$ 对矩阵乘法成群.

证明 设 $\boldsymbol{P},\boldsymbol{Q} \in \mathrm{M}_{n\times n}(R)$ 满足 $\boldsymbol{PQ} = \mathbf{1}_{n\times n} = \boldsymbol{QP}$, 则 $1 = \det(\boldsymbol{PQ}) = \det \boldsymbol{P} \det \boldsymbol{Q}$ 蕴涵 $\boldsymbol{P},\boldsymbol{Q} \in \mathrm{GL}(n,R)$.

反之设 $\boldsymbol{P} \in \mathrm{GL}(n,R)$. 取 $\boldsymbol{Q} := (\det \boldsymbol{P})^{-1}\boldsymbol{P}^\vee$ 即有 $\boldsymbol{PQ} = \mathbf{1}_{n\times n} = \boldsymbol{QP}$. □

引理 12.7.3 设 $y_1,\dots,y_m$ 为 R-模 M 的元素, $\boldsymbol{P} \in \mathrm{GL}(m,R)$, 则由

$$\begin{pmatrix} y_1' & \cdots & y_m' \end{pmatrix} = \begin{pmatrix} y_1 & \cdots & y_m \end{pmatrix}\boldsymbol{P}$$

确定的 $y_1',\dots,y_m'$ 生成的子模等于 $\sum_{j=1}^m Ry_j$.

证明 显然 $\sum_j Ry_j' \subset \sum_j Ry_j$. 另一方面

$$\begin{pmatrix} y_1 & \cdots & y_m \end{pmatrix} = \begin{pmatrix} y_1' & \cdots & y_m' \end{pmatrix}\boldsymbol{P}^{-1},$$

故同理可见 $\sum_j Ry_j \subset \sum_j Ry_j'$. □

引理 12.7.4 设 $e_1,\dots,e_n$ 为自由 R-模 E 的基. 对任意 $e_1',\dots,e_n' \in E$, 取唯一矩阵 $\boldsymbol{Q} \in \mathrm{M}_{n\times n}(R)$ 使得

$$\begin{pmatrix} e_1' & \cdots & e_n' \end{pmatrix} = \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}\boldsymbol{Q},$$

则 $e_1',\dots,e_n'$ 是 E 的基当且仅当 $\boldsymbol{Q} \in \mathrm{GL}(n,R)$.

证明 如果 $e_1',\dots,e_n'$ 是基, 则它们也可以唯一地通过某个矩阵 $\boldsymbol{Q}'$ 的右乘表出 $e_1,\dots,e_n$. 于是有

$$\begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} = \begin{pmatrix} e_1' & \cdots & e_n' \end{pmatrix}\boldsymbol{Q}' = \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}\boldsymbol{QQ}',$$

于是 $\boldsymbol{Q}\boldsymbol{Q}' = \mathbf{1}_{n\times n}$; 同理有 $\boldsymbol{Q}'\boldsymbol{Q} = \mathbf{1}_{n\times n}$. 故 $\boldsymbol{Q} \in \mathrm{GL}(n, R)$.

反之若 $\boldsymbol{Q} \in \mathrm{GL}(n, R)$, 则引理 12.7.3 (代入 $y_i = e_i$ 和 $y_i' = e_i'$) 说明 $e_1', \dots, e_n'$ 也生成 E, 而线性关系式 $\sum_{i=1}^n a_i e_i' = 0$ 等价于

$$\begin{pmatrix} e_1' \,\middle|\, \cdots \,\middle|\, e_n' \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} e_1 \,\middle|\, \cdots \,\middle|\, e_n \end{pmatrix} \boldsymbol{Q} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0,$$

这进一步等价于 $a_1 = \cdots = a_n = 0$, 故 $e_1', \dots, e_n'$ 线性无关. □

现在考虑交换环 R 上的秩 n 自由 R-模 E 及其有限生成子模 N. 选定

- E 的基 $e_1, \dots, e_n$,
- N 的生成元 $x_1, \dots, x_m$.

它们按照 (12.7.1) 唯一地确定了 $\boldsymbol{A} \in \mathrm{M}_{n\times m}(R)$. 基于引理 12.7.3 和引理 12.7.4, 矩阵 $\boldsymbol{A}$ 可以从两边调整:

- 将 $\boldsymbol{A}$ 右乘以 $\mathrm{GL}(m, R)$ 的元素不过是调整 N 的生成元 $x_1, \dots, x_m$;
- 将 $\boldsymbol{A}$ 左乘以 $\mathrm{GL}(n, R)$ 的元素相当于改变 E 的基.

所以它在 $\mathrm{GL}(m, R)$ 和 $\mathrm{GL}(n, R)$ 双边乘法作用下的轨道更贴近 $N \subset E$ 的本质结构. 这是以下处理的主题.

定义 12.7.5 我们针对域的情形定义过三类初等矩阵 $\mathcal{A}(i, k, c)$, $\mathcal{B}(i, k)$ 和 $\mathcal{C}(i, c)$; 其公式同样适用于一般的交换环 R 和 $c \in R$, 唯一差别是 $\mathcal{C}(i, c)$ 情形要求 $c \in R^\times$, 而不只是 $c \neq 0$. 对应的三类矩阵称为交换环 R 上的初等矩阵.

简单计算表明 R 上的初等矩阵必属于 $\mathrm{GL}(n, R)$. 对 $\boldsymbol{A} \in \mathrm{M}_{n\times m}(R)$ 用 $n \times n$ (或 $m \times m$) 初等矩阵左乘 (或右乘) 给出的变换仍称为初等行变换 (或初等列变换).

以下结果经常被称为 **Smith 标准形**, 它推广了域上的相抵标准形, 而内涵比域的情形丰富得多.

定理 12.7.6 (H. J. S. Smith) 设 R 或者是 $\mathbb{Z}$, 或者是某个域上的一元多项式环. 给定 $\boldsymbol{A} \in \mathrm{M}_{n\times m}(R)$, 存在 $\boldsymbol{P} \in \mathrm{GL}(m, R)$ 和 $\boldsymbol{Q} \in \mathrm{GL}(n, R)$ 使得

$$\boldsymbol{A} = \boldsymbol{Q} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix} \boldsymbol{P},$$

其中 $d_1 \mid d_2 \mid \cdots$ 落在对角线上, $d_i \in R$, 其余空白部分为零.

事实上, 对 $\boldsymbol{P}$, $\boldsymbol{Q}$ 和 $d_1, d_2, \dots$ 都有算法, 而 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 能进一步取为初等矩阵的乘积.

证明　以下的论证实际适用于任何 Euclid 整环 R, 定义详见 [10, 引理 5.7.6], 但此处仅讨论 $\mathbb{Z}$ 和域上一元多项式环的特例. 重点在于利用带余除法.

对 $m+n$ 递归地论证. 定义映射 $\mathbf{N}: R\setminus\{0\}\to\mathbb{Z}_{\geq 0}$ 如下

$$\mathbf{N}(a)=\begin{cases}|a|, & R=\mathbb{Z},\\ \deg a, & R \text{ 为域上的一元多项式环.}\end{cases}$$

若 $\boldsymbol{A}=\mathbf{0}_{n\times m}$ 则无事可作, 否则命 $D(\boldsymbol{A}):=\min\{\mathbf{N}(a_{ij}):a_{ij}\neq 0\}$. 取 (i_0,j_0) 使得 $\mathbf{N}(a_{i_0,j_0})=D(\boldsymbol{A})$.

1. 通过初等行和列变换调换行列顺序, 不失一般性可设 $(i_0,j_0)=(1,1)$.

2. 若 a_{11} 不整除所有的 a_{1j} 和 a_{i1}, 其中 $i,j\geq 2$, 则可以通过适当的初等行或列变换作带余除法, 得到 $\mathbf{N}$ 值 $<D(\boldsymbol{A})$ 的非零矩阵元.

3. 若 a_{11} 整除所有 a_{1j} 和 a_{i1}, 则继续以初等矩阵实现除法, 将 $\boldsymbol{A}$ 化为分块矩阵

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & \boldsymbol{B} & \\ 0 & & & \end{pmatrix} \tag{12.7.2}$$

的情形; 注意到 $n=1$ 或 $m=1$ 时无 $\boldsymbol{B}$. 如果 $\boldsymbol{B}$ 的某个矩阵元不被 a_{11} 整除, 则可用初等行变换将它复制到第一行, 不触及 a_{11}, 然后用 a_{11} 作带余除法, 可以得到 $\mathbf{N}$ 值 $<D(\boldsymbol{A})$ 的非零矩阵元.

综上, 存在初等矩阵的乘积 $\boldsymbol{P}_1$ 和 $\boldsymbol{Q}_1$ 使得 $\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{A}'\boldsymbol{P}_1$, 其中:

(a) 或者 $D(\boldsymbol{A}')<D(\boldsymbol{A})$,

(b) 或者 $\boldsymbol{A}'$ 形如 (12.7.2), 而 $n=1$ 或 $m=1$,

(c) 又或者 $\boldsymbol{A}'$ 形如 (12.7.2), $n,m\geq 2$ 并且存在 $\boldsymbol{A}^\flat\in \mathrm{M}_{(n-1)\times(m-1)}(R)$ 使得 $\boldsymbol{B}=a_{11}\boldsymbol{A}^\flat$.

对于 (a) 的情形, 继续先前操作, 由于 D 值递减, 有限步内必然化到 (b) 或 (c). 命 $d_1:=a_{11}$. 在 (b) 的情形构造完结. 在 (c) 的情形递归可得

$$\boldsymbol{A}^\flat=\boldsymbol{Q}^\flat\begin{pmatrix} d_2' & & \\ & d_3' & \\ & & \ddots \end{pmatrix}\boldsymbol{P}^\flat,\quad d_2'\mid d_3'\mid\cdots$$

而且 $\boldsymbol{P}^\flat$ 和 $\boldsymbol{Q}^\flat$ 都是初等矩阵的乘积. 对所有 $i \geq 2$ 命 $d_i := d_1 d_i'$. 于是

$$\boldsymbol{A} = \boldsymbol{Q}_1 \begin{pmatrix} 1 & \\ & \boldsymbol{Q}^\flat \end{pmatrix} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix} \begin{pmatrix} 1 & \\ & \boldsymbol{P}^\flat \end{pmatrix} \boldsymbol{P}_1.$$

明所欲证. □

定理 12.7.6 中的 $d_1, d_2, \ldots$ 可以各自用 $R^\times$ 调整, 因为这些倍数容易并入 $\boldsymbol{P}$ 或 $\boldsymbol{Q}$.

练习 12.7.7 证明若 R 是 $\mathbb{Z}$ 或域上的一元多项式环, 则 $\boldsymbol{P} \in \mathrm{GL}(n, R)$ 当且仅当 $\boldsymbol{P}$ 是初等矩阵的乘积.

现以定理 12.7.6 重新审视有限生成 R-模结构定理所依赖的引理 12.6.4.

推论 12.7.8 设 R 为 $\mathbb{Z}$ 或域上的一元多项式环. 设 E 为秩 n 自由 R-模, N 为其子模. 任选 E 的基 $e_1, \ldots, e_n$ 和 N 的生成元 $x_1, \ldots, x_m$, 按 (12.7.1) 定义 $\boldsymbol{A} \in \mathrm{M}_{n\times m}(R)$, 然后代入定理 12.7.6 得到 $d_1 \mid d_2 \mid \cdots$. 取 $0 \leq r \leq n$ 使得 $d_j = 0 \iff j > r$, 则:

- ⋆ 此法给出 E 的基 $f_1, \ldots, f_n$ 使得 $N = \bigoplus_{i=1}^r Rd_i f_i$;
- ⋆ $E/N \simeq \bigoplus_{i=1}^n R/(d_i)$, 其中 $d_i \in R^\times$ 的直和项可舍去.

此处涉及的一切对象都有算法; 事实上, 若取 $\boldsymbol{Q} \in \mathrm{GL}(n, R)$ 如定理 12.7.6, 则 $f_1, \ldots f_n$ 可取为下式确定的基:

$$\begin{pmatrix} f_1 \,\middle|\, \cdots \,\middle|\, f_n \end{pmatrix} = \begin{pmatrix} e_1 \,\middle|\, \cdots \,\middle|\, e_n \end{pmatrix} \boldsymbol{Q}. \tag{12.7.3}$$

证明 我们有

$$\begin{pmatrix} x_1 \,\middle|\, \cdots \,\middle|\, x_m \end{pmatrix} = \begin{pmatrix} e_1 \,\middle|\, \cdots \,\middle|\, e_n \end{pmatrix} \boldsymbol{Q} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix} \boldsymbol{P},$$

亦即

$$\begin{aligned} \begin{pmatrix} x_1 \,\middle|\, \cdots \,\middle|\, x_m \end{pmatrix} \boldsymbol{P}^{-1} &= \begin{pmatrix} e_1 \,\middle|\, \cdots \,\middle|\, e_n \end{pmatrix} \boldsymbol{Q} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix} \\ &= \begin{pmatrix} f_1 \,\middle|\, \cdots \,\middle|\, f_n \end{pmatrix} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix}. \end{aligned}$$

一如先前所见, 配合 r 的定义推得 N 由 $d_1f_1,\dots,d_rf_r$ 生成. 按直和取商便是 E/N 的描述. □

推论 12.7.9 Smith 标准形定理 12.7.6 中的 $d_1,d_2,\dots$ 精确到 $R^\times$ 是唯一确定的.

证明 命 $E:=R^{\oplus n}$ 而 N 为 $\boldsymbol{A}$ 的列向量生成的子模. 定理 12.6.3 的唯一性部分说明 $E/N\simeq R/(d_1)\oplus R/(d_2)\oplus\cdots$ 中的 $(d_1)\supset(d_2)\supset\dots$ 由 E/N 唯一确定; 该定理仅论及 $(d_i)\neq R$ 的部分, 但从这些资料和 $\boldsymbol{A}$ 的列数容易解出有几个 i 使得 $(d_i)=R$. □

计算的实例将在 §13.4 呈上.

注记 12.7.10 对于一般的主理想环 R, 定理 12.7.6 仍有相应版本, 由之同样能推导结构定理 12.6.3 的存在性部分, 但是相应的 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 未必初等, 也不再有具体算法. 本章习题有所勾勒.

习题

1. (双模) 设 A 和 B 为环. 所谓 (A,B)-双模, 意谓一个加法群 $(M,+)$ 连同写作左乘的映射 $A\times M\to M$ 和写作右乘的映射 $M\times B\to M$, 满足

- ⋆ M 对左乘成为左 A-模, 对右乘成为右 B-模;
- ⋆ $(am)b=a(mb)$ 对所有 $a\in A$, $b\in B$ 和 $m\in M$ 成立.

(i) 扼要地对双模表述子模, 同态, 同构和商模的概念, 简单说明 §12.2 的基本性质同样适用于双模.

(ii) 说明左 A-模等同于 $(A,\mathbb{Z})$-双模, 右 B-模等同于 $(\mathbb{Z},B)$-双模.

(iii) 说明若 A 为交换环, 则对任何左 A-模 M 都能定义右乘 $xa:=ax$, 使得 M 成为 (A,A)-双模.

(iv) 考虑环 A, B, C. 设 M 为 (A,B)-双模, N 为 (A,C)-双模. 考虑左 A-模的同态群 $\mathrm{Hom}_A(M,N)$. 说明它不只是加法群, 还带有 (B,C)-双模结构如下: 对所有 $f\in\mathrm{Hom}_A(M,N)$ 和 $b\in B$, $c\in C$, 定义

$$bf:x\mapsto f(xb),\quad fc:x\mapsto f(x)c,\quad x\in M.$$

类似地, 对于 (A,B)-双模 M 和 (C,B)-双模 N, 试赋予 $\mathrm{Hom}_B(M,N)$ 自然的 (C,A)-双模结构.

2. 说明 $\mathbb{Q}$ 作为 $\mathbb{Z}$-模无法分解为两个非零子模的直和.

3. 考虑环 R_1 和 R_2 的直积 $R_1\times R_2$. 对于 $i\in\{1,2\}$, 任何左 R_i-模 M_i 皆可按 $(r_1,r_2)m=r_im$ 作成左 $R_1\times R_2$-模.

(i) 给定 $R_1\times R_2$-模 M, 说明子集 $(1,0)M\subset M$ 是子模, 而且可按 $r_1\cdot(1,0)m:=(r_1,0)m$ 赋予 $(1,0)M$ 左 R_1-模结构, 使得它的 $R_1\times R_2$-模结构按照上述方式从 R_1-模结构而来. 同理, 子模 $(0,1)M$ 来自左 R_2-模.

提示〉 缘由是 $(r_1,r_2)(1,0)m=(r_1,0)m=(1,0)(r_1,0)m$.

(ii) 将左 R_i-模 M_i 按照上述方式作成左 $R_1\times R_2$-模 $(i=1,2)$. 依此构造左 $R_1\times R_2$-模 $M:=M_1\oplus M_2$. 说明 $(1,0)M=M_1$ 而 $(0,1)M=M_2$.

(iii) 证明 $M=(1,0)M\oplus(0,1)M$ 对所有左 $R_1\times R_2$-模成立.

提示〉 所有 $m\in M$ 都能写成 $(1,1)m=(1,0)m+(0,1)m$; 从 $(1,0)(0,1)=(0,0)=(0,1)(1,0)$ 说明 $(1,0)M\cap(0,1)M=\{0\}$.

(iv) 给定 M 和 N, 有自然双射

$$\begin{aligned}\mathrm{Hom}_{R_1\times R_2}(M,N)&\xrightarrow{1:1}\mathrm{Hom}_{R_1}((1,0)M,(1,0)N)\times\mathrm{Hom}_{R_2}((0,1)M,(0,1)N)\\ f&\mapsto\left(f|_{(1,0)M},f|_{(0,1)M}\right),\end{aligned}$$

(下标代表对何种模结构考虑同态). 双射与同态的合成与加法运算兼容.

有鉴于此, 直积上的左模及同态可以化到各个分量来研究, 精确的说法则涉及范畴等价的概念. 右模的情形完全相同.

4. 给定两族左或右 R-模 $(M_j)_{j\in J}$ 和 $(N_i)_{i\in I}$, 验证加法群的同构

$$\begin{aligned}\mathrm{Hom}\left(\textstyle\bigoplus_{j\in J}M_j,\prod_{i\in I}N_i\right)&\xrightarrow{\sim}\textstyle\prod_{(i,j)\in I\times J}\mathrm{Hom}(M_j,N_i)\\ f&\longmapsto(p_if\iota_j)_{(i,j)\in I\times J}\end{aligned}$$

其中 $\iota_j:M_j\hookrightarrow\bigoplus_{j'}M_{j'}$ 是包含同态, $p_i:\prod_{i'}N_{i'}\twoheadrightarrow N_i$ 是投影同态. 当 R 交换时, 这还是 R-模的同构.

5. 设 R 为环, 将 R 以环乘法作成左或右 R-模.

(i) 验证环同构 $\lambda:R\xrightarrow{\sim}\mathrm{End}_{\text{右 }R\text{-模}}(R)$, 元素 $r\in R$ 对应左乘映射 $\lambda_r:x\mapsto rx$. 提示〉 自同态由 1_R 的像唯一确定.

(ii) 验证环同构 $\rho:R^{\mathrm{op}}\xrightarrow{\sim}\mathrm{End}_{\text{左 }R\text{-模}}(R)$, 其中 R^{op} 是 R 的相反环 (练习 12.1.2), 元素 $r\in R$ 对应右乘映射 $\rho_r:x\mapsto xr$.

(iii) 对所有 $n,m\in\mathbb{Z}_{\geq 0}$, 说明 $\mathrm{Hom}_{\text{右 }R\text{-模}}\left(R^{\oplus n},R^{\oplus m}\right)$ 按照和向量空间情形一样的方式等同于 $\mathrm{M}_{m\times n}(R)$, 使得态射加法对应到矩阵加法, 态射合成对应到矩阵乘法, 而且在 R 交换时这还是 R-模同构. 提示〉 基于 (i) 和前一道习题, 照搬向量空间版本即可.

6. 设 M, M', M'' 为主理想环 R 上的有限生成模. 证明:

(i) 若 $M\oplus M'\simeq M\oplus M''$ 则 $M'\simeq M''$;

(ii) 若存在 $n\in\mathbb{Z}_{\geq 1}$ 使得 $M^{\oplus n}\simeq(M')^{\oplus n}$, 则 $M\simeq M'$.

7. 分类所有阶为 100000 的交换群; 共有 49 个同构类.

提示〉 分别讨论阶为 2^5 和 5^5 的交换群.

8. 分类所有阶为 540 的交换群; 共有 6 个同构类.

9. 写下交换群或 $\mathbb{Z}$-模 $\mathbb{Z}/5\mathbb{Z}\oplus\mathbb{Z}/15\mathbb{Z}\oplus\mathbb{Z}/25\mathbb{Z}\oplus\mathbb{Z}/54\mathbb{Z}$ 的不变因子和初等因子.

10. 设 R 为交换环而 M 为秩 n 自由 R-模, $n\in\mathbb{Z}_{\geq 1}$. 证明若元素 $x_1,\dots,x_n\in M$ 生成 M, 则它们是 M 的基. 提示〉 按 (12.7.1) 的写法, 将 $x_1,\dots,x_n$ 用给定的基 $e_1,\dots,e_n$ 表出.

11. 对 Smith 标准形定理 12.7.6 证明以下的主理想环版本: 设 R 为主理想环, $\boldsymbol{A}\in \mathrm{M}_{n\times m}(R)$, 存在 $\boldsymbol{P}\in\mathrm{GL}(m,R),\boldsymbol{Q}\in\mathrm{GL}(n,R)$ 和 $d_1,d_2,\dots\in R$ 使得

$$\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \end{pmatrix}\boldsymbol{P},\quad d_1\mid d_2\mid\cdots,$$

而且 $d_1,d_2,\dots$ 精确到 $R^\times$ 由 $\boldsymbol{A}$ 唯一确定.

提示〉 在所有形如 $\boldsymbol{A}'=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}^{-1}$ 的矩阵中, 取之使得 $a'_{11}\neq 0$ 非零而且素因子分解 $a'_{11}\sim p_1^{e_1}\cdots p_m^{e_m}$ 中的 $e_1+\cdots+e_m$ 尽可能小. 先来论证 a'_{11} 整除所有 a'_{i1} 和 a'_{1j}, 其中 $i,j\neq 1$: 以前者为例, 若 $ua'_{11}+va'_{i1}=\gcd(a'_{11},a'_{i1})$, 则 $u,v\in R$ 必互素, 取形如

$$\begin{pmatrix} u & v \\ r & s \end{pmatrix},\quad ur+vs=1$$

的矩阵, 适当补零后对 $\boldsymbol{A}'$ 左乘以制造矩阵元 $\gcd(a'_{11},a'_{i1})$. 说明 $a'_{11}\nmid a'_{i1}$ 引致矛盾.

基于上述观察, 不失一般性可设 $\boldsymbol{A}'$ 形如

$$\begin{pmatrix} a'_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \boldsymbol{B}' & \\ 0 & & & \end{pmatrix};$$

说明在 $n,m\geq 2$ 的前提下, $\boldsymbol{B}'$ 的所有矩阵元必被 a'_{11} 整除: 设若不然, 可左乘适当矩阵将该矩阵元加到第一行, 再应用上一步的手法. 按此递归地操作.

12. 承上题, 证明可以取到 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 使得 $\det\boldsymbol{P}=1=\det\boldsymbol{Q}$.

13. 设 D 为除环. 受 D 为域的情形启发, 将左 (或右) D-模另称为左 (或右) D-向量空间. 试尽可能地将向量空间中的线性无关性, 基, 维数, 消元法, 线性映射等概念从域上的情形推广到除环 (比方说, 线性映射无非是模同态). 注意到行列式理论无法直接推广到除环上.

第十三章 标准形

本章旨在解决以下问题: 给定域 F 上的两个 $n\times n$ 矩阵, 如何判断它们是否共轭? 如何描述矩阵的所有共轭类? 问题也可以用线性映射的语言表述. 我们期望的解答是在每个共轭类中确定称为标准形的唯一元素, 同时为标准形提供可行的算法.

关于对角化的研究已经部分地回答了上述问题: 任何可对角化矩阵都共轭于一个具体可算的对角矩阵, 两个对角矩阵共轭当且仅当其对角元至多差一个置换. 由此可见对角矩阵扮演了近乎标准形的角色. 然而并非所有矩阵都能对角化. 为了处理一般情形, 更深入的代数工具不可或缺, 而第十二章已为此作好了准备.

我们在 §§13.1–13.2 先将问题以 $F[X]$-模的语言重新表述. 对于 $n\times n$ 矩阵或 n 维空间上的线性变换, 研究共轭类相当于分类 F-维数为 n 的 $F[X]$-模, 精确到模同构.

本章介绍的标准形理论有两种. 首先是 §13.3 介绍的有理标准形理论, 对应的标准形是由称为友矩阵的一列

$$\boldsymbol{C}_{f_i}=\begin{pmatrix}0&0&\cdots&0&-c_{i,0}\\1&0&\cdots&0&-c_{i,1}\\0&1&\cdots&0&-c_{i,2}\\\vdots&&\ddots&&\vdots\\0&0&\cdots&1&-c_{i,n_i-1}\end{pmatrix},\qquad \begin{aligned}f_i&=c_{i,0}+\cdots+c_{i,n_i-1}X^{n_i-1}+X^{n_i}\\&\in F[X]\end{aligned}$$

组成的分块对角矩阵 $(1\le i\le k)$, 要求 $f_1\mid\cdots\mid f_k$ 而 $\sum_{i=1}^k n_i=n$. 证明是有限生成 $F[X]$-模分类定理 12.6.3 的简单应用. 该节将为有理标准形给出两种形式及其应用. 在 §13.4 将证明求 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 的有理标准形等价于求 $X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}\in \mathrm{M}_{n\times n}(F[X])$ 的 Smith 标准形, 算法基于多项式的带余除法; 我们将给出例子.

第二种是 §13.5 介绍的 Jordan 标准形理论. 它要求 F 是代数闭域 (例如 $F=\mathbb{C}$), 或者至少要求所论的矩阵有分裂的特征多项式. 对应的标准形矩阵是由一列称为 $d\times d$

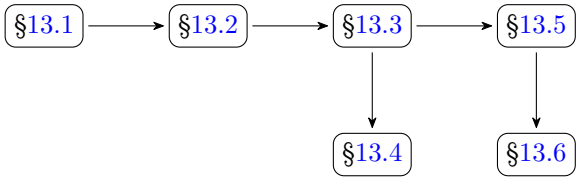
Jordan 块的上三角矩阵

$$\boldsymbol{J}_d(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} = \lambda \cdot \mathbf{1}_{d\times d} + \boldsymbol{J}_d(0), \quad \lambda \in F$$

组成的分块对角矩阵, 精确到置换; 一些文献考虑的是下三角 Jordan 块 ${}^{\mathrm{t}}\boldsymbol{J}_d(\lambda)$, 结论不变. Jordan 标准形可以视为有理标准形在特征多项式分裂情形的一种变体; 实际上 ${}^{\mathrm{t}}\boldsymbol{J}_d(0)$ 正是多项式 X^d 确定的友矩阵, 相应的分解也能从有理标准形来推导, 大致思路是化约到幂零情形 (定义 13.5.1).

Jordan 标准形虽然需要较多的条件, 但确有其优点. 例如它们与特征值的关系更明朗, 而且对角化是 Jordan 标准形的特例: 对角矩阵无非是 1×1 Jordan 块组成的标准形. 相关算法可见 §13.6; 由于计算过程涉及对特征多项式求根, 因此不如有理标准形简单.

阅读顺序



13.1 线性映射和模结构

本节选定域 F, 将它等同于多项式环 $F[X]$ 的子环.

对于任何 $F[X]$-模 V, 将纯量乘法从 $F[X]$ 限制到 F, 便得到 V 上的 F-向量空间结构. 反过来说, 若 F-向量空间 V 兼具 $F[X]$-模的结构, 使得纯量乘法限制到 F 上给出原有向量空间, 则称 V 按此 “升级” 为 $F[X]$-模.

引理 13.1.1 *设 V 为 F-向量空间. 指定 $T\in \mathrm{End}(V)$ 等价于将 V 升级为 $F[X]$-模, 方式是要求 $F[X]$ 的纯量乘法满足*

$$X\cdot v = T(v), \quad v\in V.$$

进一步, 若 $F[X]$-模 V 和 V' 分别对应到 $T\in\mathrm{End}(V)$ 和 $T'\in\mathrm{End}(V')$, 则

$$\{F[X]\text{-模同态 }\varphi:V\to V'\}=\{\varphi\in\mathrm{Hom}(V,V'):T'\varphi=\varphi T\}\,.$$

证明 以下论证服务于偏好细节的读者. 将 F-向量空间 V 升级为 $F[X]$-模相当于对所有 $f\in F[X]$ 指定映射 $\rho_f:V\to V$, 对应到 f 的纯量乘法, 条件是:

- ⋆ $\rho_f(v_1+v_2)=\rho_f(v_1)+\rho_f(v_2)$,
- ⋆ $\rho_{f+g}(v)=\rho_f(v)+\rho_g(v)$,
- ⋆ $\rho_{fg}(v)=\rho_f(\rho_g(v))$,
- ⋆ 当 $c\in F$ 时 $\rho_c(v)=cv$ (这是 "升级" 的意涵), 特别地 $\rho_1=\mathrm{id}_V$.

设 V 已升级为 $F[X]$-模. 命 $T:=\rho_X:V\to V$, 则 T 保加法, 此外 $\rho_cT=\rho_{cX}=\rho_{Xc}=T\rho_c$ (取 $c\in F$) 导致 T 保持来自 F 的纯量乘法, 故 $T\in\mathrm{End}(V)$. 对于一般的 $f=a_nX^n+\cdots+a_0$, 对应的 ρ_f 被 T 确定为

$$\begin{aligned}\rho_f&=a_n\rho_{X^n}+\cdots+a_0\rho_1\\&=a_n(\rho_X)^n+\cdots+a_0\cdot\mathrm{id}_V\\&=a_nT^n+\cdots+a_0\cdot\mathrm{id}_V=f(T).\end{aligned}$$

反之给定 $T\in\mathrm{End}(V)$, 对所有 $f\in F[X]$ 定义 $\rho_f=f(T):V\to V$, 我们希望说明这使 V 升级为 $F[X]$-模. 所列条件缘自已知的 $f(T)\in\mathrm{End}(V)$, $(f+g)(T)=f(T)+g(T)$, $(fg)(T)=f(T)g(T)$ 和

$$c\in F\implies c(T)=c\cdot\mathrm{id}_V.$$

双向的构造显然互为逆.

最后, 考虑 $F[X]$-模之间的映射 $\varphi:V\to V'$, 则 φ 是模同态等价于它保持加法, 而且 $\varphi f(T)=f(T')\varphi$ 对所有 $f\in F[X]$ 成立. 展开 $f(T)$ 和 $f(T')$ 可知这又相当于说 φ 是 F-线性的, 而且 $(T')^n\circ\varphi=\varphi\circ T^n$ 对所有 $n\geq 0$ 成立; 后者进一步等价于 $T'\varphi=\varphi T$. □

练习 13.1.2 进一步说明从 V 到 V' 的 $F[X]$-模同构无非是满足 $T'\varphi=\varphi T$ 的向量空间同构 $\varphi:V\xrightarrow{\sim}V'$.

约定 13.1.3 对于任意 $F[X]$-模 V, 今后将 V 作为 F-向量空间的维数简称为 V 的维数, 依此定义何谓有限维 $F[X]$-模.

命题 13.1.4 有限维 $F[X]$-模必然是有限生成 $F[X]$-模.

证明 对任意 $F[X]$-模, 它作为向量空间的生成元当然也是作为模的生成元. □

练习 13.1.5 设 M 为有限维 $F[X]$-模, 对应到 F-向量空间 V 连同 $T\in\mathrm{End}(V)$. 说明命题 12.5.2 的分解和讨论极小多项式时用过的分解 $V[f]=V[g]\oplus V[h]$ 本质相同 ($f=gh\in F[X]$ 而 g,h 互素).

13.2 问题的表述

仍然选定域 F 和 $n \in \mathbb{Z}_{\geq 1}$. 就具体的矩阵视角, 研究标准形的动机是为了判断两个矩阵 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$ 是否共轭. 回忆到共轭意谓:

$$\text{存在可逆之 } \boldsymbol{P} \in \mathrm{M}_{n\times n}(F) \text{ 使得 } \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}.$$

我们希望选出一族特殊的 $n \times n$ 矩阵, 称为标准形, 使得每个矩阵都共轭于一个基本上唯一的标准形, 并提供算法. 这将为共轭问题给出具体的解答. 对于可对角化矩阵的情形, 对角矩阵扮演的角色类似于标准形, 共轭类的分类问题对此归结为特征值的计算. 然而并非所有 $n \times n$ 矩阵都能对角化, 是以需要更广泛的理论.

现在表述三种分类问题. 选定 $n \in \mathbb{Z}_{\geq 0}$.

▷ **矩阵版本** 将矩阵 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$ 的共轭关系写作 $\boldsymbol{A} \sim \boldsymbol{B}$. 上述矩阵共轭问题相当于研究商集 $\mathrm{M}_{n\times n}(F)/\sim$.

▷ **线性映射版本** 考虑形如 (V, T) 的资料, 其中 V 是 n 维 F-向量空间, 而 $T \in \mathrm{End}(V)$. 对于资料 (V, T) 和 (V', T'), 若存在向量空间的同构 $\varphi : V \xrightarrow{\sim} V'$ 使得下图交换

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & V' \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T'} \\ V & \xrightarrow[\varphi]{} & V' \end{array} \quad \text{亦即 } T'\varphi = \varphi T,$$

则记此为 $(V, T) \approx (V', T')$. 易见 $\approx$ 是等价关系. 自然的问题是分类这些资料, 亦即研究商集

$$\{\text{资料 } (V, T),\ \dim V = n\}/\approx .$$

▷ **模论版本** 在同构意义下分类 n-维 $F[X]$-模, 亦即研究商集 $\{n\text{-维 } F[X]\text{-模}\}/\simeq$.

命题 13.2.1 上述三种分类问题相互等价: 我们有以下双射

$$\begin{array}{ccc} \mathrm{M}_{n\times n}(F)/\sim & & \boldsymbol{A} \\ {\scriptstyle 1:1}\downarrow & & \downarrow \\ \{\text{资料 } (V,T),\ \dim V = n\}/\approx & (F^n, \boldsymbol{A}) & (V, T) \\ {\scriptstyle 1:1}\downarrow & & \downarrow \\ \{n \text{ 维 } F[X]\text{-模}\}/\simeq & & \text{将 } V \text{ 用 } T \text{ 升级为 } F[X]\text{-模}. \end{array}$$

严格来说, 右侧所示的映法是在等价类上操作的.

证明 第二步的升级手续来自引理 13.1.1, 该处同样说明了第二步确实是双射.

对于第一步, 注意到若 $\boldsymbol{A},\boldsymbol{B}\in \mathrm{M}_{n\times n}(F)$, 皆视同线性映射 $F^n\to F^n$, 则 $(F^n,\boldsymbol{A})\approx(F^n,\boldsymbol{B})$ 等价于存在可逆之 $\boldsymbol{P}\in \mathrm{M}_{n\times n}(F)$ 使得 $\boldsymbol{BP}=\boldsymbol{PA}$, 亦即 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{BP}$. 这就说明 $\boldsymbol{A}\mapsto(F^n,\boldsymbol{A})$ 在商集层次也是良定义的, 并给出单射.

为了说明第一步的箭头为满, 对给定的资料 (V,T) 取定同构 $\varphi: V\xrightarrow{\sim} F^n$, 亦即选定 V 的有序基, 相应地命 $\boldsymbol{A}:=\varphi T\varphi^{-1}\in \mathrm{End}(F^n)$. 于是 $\boldsymbol{A}$ 视作 $\mathrm{M}_{n\times n}(F)$ 的元素是 T 相对于有序基的矩阵表法, 而且由 φ 立得 $(V,T)\approx(F^n,\boldsymbol{A})$. 证毕. □

综上, 研究矩阵的标准形理论相当于分类有限维 $F[X]$-模的同构类, 并且在每个同构类中指定代表元.

分类问题之间的等价性也与直和兼容.

命题 13.2.2 给定一族矩阵 $\boldsymbol{A}_i\in \mathrm{M}_{n_i\times n_i}(F)$, 其中 $i=1,\dots,k$. 设 $\boldsymbol{A}_i$ 对应到 n_i 维 $F[X]$-模 M_i, 而 $n:=\sum_{i=1}^k n_i$, 则分块对角矩阵 $\mathrm{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_k)$ 对应得 n 维 $F[X]$-模是 $M_1\oplus\cdots\oplus M_k$.

证明 设 X 的纯量乘法对每个 $1\leq i\leq k$ 给出 F-线性映射 $T_i: M_i\to M_i$, 在选定的有序基之下对应到矩阵 $\boldsymbol{A}_i$, 则 X 的纯量乘法在 $M_1\oplus\cdots\oplus M_k$ 上给出线性映射 $(x_1,\dots,x_k)\mapsto(T_1x_1,\dots,T_kx_k)$; 这在矩阵层面对应的正是 $\mathrm{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_k)$. □

已知 $F[X]$ 是主理想环, 故 §12.6 的理论立刻派上用场. 行将介绍的标准形理论分成两种: §13.3 的有理标准形适用一般的域, 而在不对矩阵设限的前提下, §13.5 的 Jordan 标准形需要 F 为代数闭域, 但结果更简明. 我们将从有理标准形入手, Jordan 标准形容易由之推导.

13.3 有理标准形

既然命题 13.2.1 提到 n 维 $F[X]$-模 M 对应到矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$, 精确到共轭, 现在考虑非零循环 $F[X]$-模

$$M:=F[X]/(f),\quad f\in F[X]\smallsetminus F.$$

我们有 $n:=\deg f=\dim_F M$. 事实上, M 作为 F-向量空间的有序基可取为

$$1, X, \dots, X^{n-1} \text{ 对 } (f) \text{ 的陪集.}$$

相对于上述有序基, $\boldsymbol{A}$ 对应到线性映射 $T: g+(f)\mapsto Xg+(f)$, 其中 $g\in F[X]$.

不失一般性可设 f 首一, 写作 $f=c_0+\cdots+c_{n-1}X^{n-1}+X^n$. 当 $i<n-1$ 时有陪集的等式 $T\left(X^i+(f)\right)=X^{i+1}+(f)$, 而对 $i=n-1$ 则有

$$\begin{aligned}T\left(X^{n-1}+(f)\right)&=X^n+(f)\\&=-c_0-c_1X-\cdots-c_{n-1}X^{n-1}+(f).\end{aligned}$$

以下结论水到渠成.

命题 13.3.1 设 $f = c_0 + \cdots + c_{n-1}X^{n-1} + X^n \in F[X]$, 其中 $n \in \mathbb{Z}_{\geq 1}$. 对于 $F[X]$-模 $F[X]/(f)$, 取其有序基为 $1, X, \ldots, X^{n-1}$ 对 (f) 的陪集, 则对应的 $n \times n$ 矩阵是友矩阵

$$\boldsymbol{C}_f = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}.$$

推而广之, 给定首一多项式 $f_1, \ldots, f_k \in F[X]$, 按上述方式选定每个 $F[X]/(f_i)$ 的有序基, 则 $F[X]$-模 $\bigoplus_{i=1}^k F[X]/(f_i)$ 对应分块对角矩阵 $\operatorname{diag}(\boldsymbol{C}_{f_1}, \ldots, \boldsymbol{C}_{f_k})$.

证明 先前已说明 $F[X]/(f)$ 的情形, 一般情形则是命题 13.2.2 的应用. □

定理 13.3.2 (有理标准形) 设 $n \in \mathbb{Z}_{\geq 1}$ 而 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 存在唯一一列前后整除的非常数首一多项式

$$f_1 \mid \cdots \mid f_k, \quad k \in \mathbb{Z}_{\geq 1}, \quad f_i \in F[X],$$

使得 $\sum_{i=1}^k \deg f_i = n$ 而 $\boldsymbol{A}$ 共轭于分块对角矩阵

$$\operatorname{diag}(\boldsymbol{C}_{f_1}, \ldots, \boldsymbol{C}_{f_k}) = \begin{pmatrix} \boxed{\boldsymbol{C}_{f_1}} & & \\ & \ddots & \\ & & \boxed{\boldsymbol{C}_{f_k}} \end{pmatrix}.$$

上述矩阵称为 $\boldsymbol{A}$ 的**有理标准形**, 而 $f_1, \ldots, f_k$ 称为 $\boldsymbol{A}$ 的不变因子.

证明 基于命题 13.2.1 的对应, 记 $\boldsymbol{A}$ 对应的 n 维 $F[X]$-模为 M. 将 M 代入关于主理想环 $F[X]$ 上的有限生成模的结构定理 12.6.3, 分解中的自由部分必然是 0, 这是因为 M 有限维, 而 $F[X]$ 则不然; 此外 $M \neq \{0\}$. 因此我们得到唯一的 $k \in \mathbb{Z}_{\geq 1}$ 和一列非零真理想 $I_1 \supset \cdots \supset I_k$, 使得

$$M \simeq F[X]/I_1 \oplus \cdots \oplus F[X]/I_k.$$

进一步对每个 I_i 取唯一的首一多项式 f_i 使得 $I_i = (f_i)$, 故 $M \simeq \bigoplus_{i=1}^k F[X]/(f_i)$. 命题 13.3.1 遂表明 $\boldsymbol{A}$ 共轭于 $\operatorname{diag}(\boldsymbol{C}_{f_1}, \ldots, \boldsymbol{C}_{f_k})$.

反之, 若 $\boldsymbol{A}$ 共轭于 $\operatorname{diag}(\boldsymbol{C}_{f_1}, \ldots, \boldsymbol{C}_{f_k})$, 则命题 13.2.1 和命题 13.3.1 表明 $M \simeq \bigoplus_i F[X]/(f_i)$, 故 $f_1 \mid \cdots \mid f_k$ 的唯一性归结为有限生成 $F[X]$-模的不变因子的唯一性. □

因此矩阵的共轭类完全由不变因子来决定, 此即 “标准形” 的内涵.

推论 13.3.3 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 的不变因子为 $f_1 \mid \cdots \mid f_k$, 则:

(i) $\boldsymbol{A}$ 的极小多项式 $\mathrm{Min}_{\boldsymbol{A}}$ 等于 f_k;

(ii) $\boldsymbol{A}$ 的特征多项式 $\mathrm{Char}_{\boldsymbol{A}}$ 等于 $\prod_{i=1}^{k} f_i$.

证明 仍记 $\boldsymbol{A}$ 对应的 n 维 $F[X]$-模为 M. 对于 (i), 以模论语言将 $\mathrm{Min}_{\boldsymbol{A}}$ 刻画为满足下式的首一多项式

$$(\mathrm{Min}_{\boldsymbol{A}}) = \{t \in F[X] : \forall x \in M,\ tx = 0\}.$$

然而从分解 $M \simeq \bigoplus_i F[X]/(f_i)$ 和前后整除的条件立见右式即 (f_k).

对于 (ii), 按分块计算可知 $\mathrm{Char}_{\boldsymbol{A}} = \prod_{i=1}^{k} \mathrm{Char}_{\boldsymbol{C}_{f_i}}$. 对每个 $1 \leq i \leq k$, 本书在讨论特征多项式时已直接按定义算出 $\mathrm{Char}_{\boldsymbol{C}_{f_i}} = f_i$. □

在有理标准形的基础上, 前述论证只用到极小多项式的刻画, 特征多项式的分块分解和友矩阵的计算, 不涉及特征多项式的进阶内容, 而由此已经足以推得 $\mathrm{Min}_{\boldsymbol{A}} \mid \mathrm{Char}_{\boldsymbol{A}}$, 换言之 $\mathrm{Char}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$. 这给出 Cayley–Hamilton 定理的另证, 它只依赖有理标准形定理 13.3.2, 而后者不过是主理想环上的有限生成模结构定理 12.6.3 的直接应用.

在证明定理 12.6.3 的唯一性部分时, 我们曾运用分解 $M = \bigoplus_p M[p^\infty]$. 在矩阵的场景, 这直接给出有理标准形的第二种形式.

推论 13.3.4 设 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$, 取 $\mathrm{Min}_{\boldsymbol{A}}$ 的不可约分解 $p_1^{e_1}\cdots p_h^{e_h}$, 其中 $p_1, \ldots, p_h$ 是相异的不可约首一多项式, 则 $\boldsymbol{A}$ 共轭于形如 $\mathrm{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_h)$ 的分块对角矩阵, 其中对每个 $1 \leq j \leq h$ 皆有

$$\boldsymbol{A}_j = \mathrm{diag}\left(\boldsymbol{C}_{p_j^{b_{1,j}}}, \ldots, \boldsymbol{C}_{p_j^{b_{r_j,j}}}\right), \quad r_j \in \mathbb{Z}_{\geq 1}, \quad 1 \leq b_{1,j} \leq \cdots \leq b_{r_j,j}.$$

一旦 $p_1, \ldots, p_h$ 的顺序选定, 资料 $(b_{i,j})_{i,j}$ 也唯一确定.

证明 写下 $\boldsymbol{A}$ 的不变因子 $f_1 \mid \cdots \mid f_k$, 则 $p_1, \ldots, p_h$ 便是 f_k 的所有不可约素因子. 记 $\boldsymbol{A}$ 对应的 $F[X]$-模为 M. 断言中的 $\boldsymbol{A}_j$ 无非是 $M[p_j^\infty]$ 所对应的有理标准形. 在定理 12.6.3 的唯一性证明中已经说明了当 $p_1, \ldots, p_h$ 顺序取定, 资料 $(b_{i,j})_{i,j}$ 由 M 的同构类唯一确定, 亦即由 $\boldsymbol{A}$ 的共轭类唯一确定. □

观察到 $\mathrm{Char}_{\boldsymbol{A}_j} = p_j^{b_{1,j}+\cdots}$ 正是 $\mathrm{Char}_{\boldsymbol{A}}$ 中被 p_j 整除的部分.

注记 13.3.5 (转置版本) 在许多教材中, 有理标准形涉及的友矩阵写作

$$\begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & \cdots & -c_{n-1} \end{pmatrix}$$

亦即先前定义的 $\boldsymbol{C}_f$ 的转置. 因为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 共轭当且仅当 ${}^{\mathrm{t}}\boldsymbol{A}$ 和 ${}^{\mathrm{t}}\boldsymbol{B}$ 共轭, 有理标准形定理 13.3.2 的转置版本和原版相互等价, 其一切推论也相等价.

这一切当然也能够用线性映射的语言表述.

定理 13.3.6 (有理标准形: 线性映射版本) 设 V 为 n 维 F-向量空间, $T \in \mathrm{End}(V)$.

▷ **第一种形式** 存在 V 的有序基和一列非常数首一多项式 $f_1 \mid \cdots \mid f_k$, 使得 $\sum_{i=1}^{k} \deg f_i = n$, 而 T 表作矩阵 $\mathrm{diag}(\boldsymbol{C}_{f_1}, \ldots, \boldsymbol{C}_{f_k})$.

▷ **第二种形式** 取 Min_T 的不可约分解 $p_1^{e_1} \cdots p_h^{e_h}$, 则存在 V 的有序基, 使得 T 的矩阵为 $\mathrm{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_h)$, 其中

$$\boldsymbol{A}_j = \mathrm{diag}\left(\boldsymbol{C}_{p_j^{b_{1,j}}}, \ldots, \boldsymbol{C}_{p_j^{b_{r_j,j}}}\right), \quad r_j \in \mathbb{Z}_{\geq 1}, \quad 1 \leq b_{1,j} \leq \cdots \leq b_{r_j,j}.$$

第一种形式中的资料 $f_1 \mid \cdots \mid f_k$ (称为 T 的不变因子) 由 T 唯一确定. 在第二种形式中, 一旦选定 $p_1, \ldots, p_h$ 的顺序, 则资料 $(b_{i,j})_{i,j}$ 也由 T 唯一确定; 资料 $p_j^{b_{i,j}}$ 称为 T 的初等因子.

证明 由于矩阵是线性映射在有序基之下的表象, 而换基相当于将对应的矩阵取共轭, 这不过是定理 13.3.2 和推论 13.3.4 的简单改述. □

一如推论 13.3.3, 由此对 $T \in \mathrm{End}(V)$ 推得[1)]

$$\mathrm{Min}_T = f_k, \quad \mathrm{Char}_T = \prod_{i=1}^{k} f_i. \tag{13.3.1}$$

有理标准形有时也称为 Frobenius 标准形. 之所以称为有理标准形, 是因为当域 F 给定 (譬如 $F = \mathbb{Q}$), 一切都在 F 中操作, 不必过渡到 F 的扩域 (譬如 $\mathbb{C}$) 以确保特征多项式分裂. 这点与即将讨论的 Jordan 标准形理论形成反差.

[1)] 在某些教材中, 模论的顺序早于向量空间, 特征多项式则直接定义为不变因子的乘积.

13.4 有理标准形的计算

有理标准形的实质是有限生成 $F[X]$-模的结构定理 12.6.3. 在该定理的证明中, 分解的存在性是通过将 M 表作有限秩自由模的商来处理的. 本节稍微改变符号, 选定 $n \in \mathbb{Z}_{\geq 1}$, 考虑 n 维 F-向量空间 V 配上 $T \in \mathrm{End}(V)$, 以 T 将 V 升级为 $F[X]$-模.

取 V 的基 $v_1, \ldots, v_n$. 定义秩 n 自由 $F[X]$-模 $F[X]^{\oplus n}$ 及其标准基 $e_1, \ldots, e_n$. 现在考虑 $F[X]$-模的同态

$$\begin{aligned} \varphi : F[X]^{\oplus n} &\to V \\ \sum_{i=1}^{n} r_i e_i &\mapsto \sum_{i=1}^{n} r_i(T) v_i. \end{aligned}$$

同态 φ 显然满. 另一方面, 设 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{n\times n}(F)$ 是 T 对应的矩阵, 定义 $F[X]^{\oplus n}$ 的一列元素

$$x_j := Xe_j - \sum_{i=1}^{n} a_{ij} e_i, \quad 1 \leq j \leq n,$$

以及它们生成的 $F[X]$-子模

$$N := \langle x_1, \ldots, x_n \rangle .$$

引理 13.4.1 同态 φ 的核等于 N.

证明 按照定义,

$$\varphi(x_j) = \varphi\left(Xe_j - \sum_{i=1}^{n} a_{ij} e_i \right) = Tv_j - \sum_{i=1}^{n} a_{ij} v_i = 0$$

对所有 j 成立, 故 $N \subset \ker(\varphi)$. 另一方面, 由于 Xe_j 属于陪集 $\sum_i a_{ij} e_i + N$, 对每个 j 反复运用此等式 (不断用 X 左乘), 可以对一般的 $x \in F[X]^{\oplus n}$ 将陪集 $x + N$ 表如

$$x + N = \sum_{j=1}^{n} c_j e_j + N, \quad c_j \in F.$$

所以 $\varphi(x) = 0$ 当且仅当 $\sum_{j=1}^{n} c_j v_j = 0$, 当且仅当 $c_1 = \cdots = c_n = 0$; 然而这便蕴涵 $x \in N$. □

因此从 $T \in \mathrm{End}(V)$ 和 V 的基 $v_1, \ldots, v_n$ 起步, 我们抵达 $F[X]$-模的同构 $\overline{\varphi} : F[X]^{\oplus n}/N \xrightarrow{\sim} V$. 按照 (12.7.1) 的记法,

$$\left(\begin{array}{c|c|c|c} x_1 & \cdots & x_m \end{array} \right) = \left(\begin{array}{c|c|c|c} e_1 & \cdots & e_n \end{array} \right) \begin{pmatrix} X - a_{11} & -a_{12} & -a_{13} & \cdots \\ -a_{21} & X - a_{22} & -a_{23} & \cdots \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X - a_{nn} \end{pmatrix};$$

右式的 $n\times n$ 矩阵即 $X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}$, 是 $\mathrm{M}_{n\times n}(F[X])$ 的元素[2)], 故 §12.7 的方法可资应用.

命题 13.4.2 按照定理 12.7.6 取 $\boldsymbol{P},\boldsymbol{Q}\in\mathrm{GL}(n,F[X])$ 和 $F[X]$ 的元素 $d_1\mid d_2\mid\cdots$ 使得

$$X\cdot\mathbf{1}_{n\times n}-\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix}d_1&&\\&\ddots&\\&&d_n\end{pmatrix}\boldsymbol{P},$$

则 d_i 皆可取为首一的, 而且从 $(d_i)_{i=1}^n$ 删除 $d_i=1$ 的项后, 产物是 T 的有理标准形中的不变因子.

证明 根据推论 12.7.8,

$$V\simeq F[X]^{\oplus n}/N\simeq\bigoplus_{i=1}^{n}F[X]/(d_i).$$

由于 $\dim V=n$ 有限, 对所有 i 皆有 $d_i\neq 0$, 故可设 d_i 首一. 根据有理标准形与结构定理的关系 (请回忆 §13.3), 不变因子正是 $d_1,\dots,d_n$ 中 $\neq 1$ 的项. □

注记 13.4.3 沿用上述符号. 命 $g_i\in F[X]^{\oplus n}$ 为 $\boldsymbol{Q}$ 的第 i 列, 则

$$\underbrace{\begin{pmatrix}x_1&\big|&\cdots&\big|&x_m\end{pmatrix}\boldsymbol{P}^{-1}}_{\text{生成 }N}=\underbrace{\begin{pmatrix}g_1&\big|&\cdots&\big|&g_n\end{pmatrix}}_{F[X]^{\oplus n}\text{ 的基}}\begin{pmatrix}d_1&&\\&\ddots&\\&&d_n\end{pmatrix}.$$

鉴于模同构

$$\begin{array}{ccccc}\bigoplus_{i=1}^n F[X]/(d_i)&\xrightarrow{\sim}&F[X]^{\oplus n}/N&\xrightarrow[\sim]{\overline{\varphi}}&V\\(r_i+(d_i))_{i=1}^n&\longmapsto&\sum_{i=1}^n r_ig_i+N&\longmapsto&\sum_{i=1}^n r_i(T)\varphi(g_i),\end{array}$$

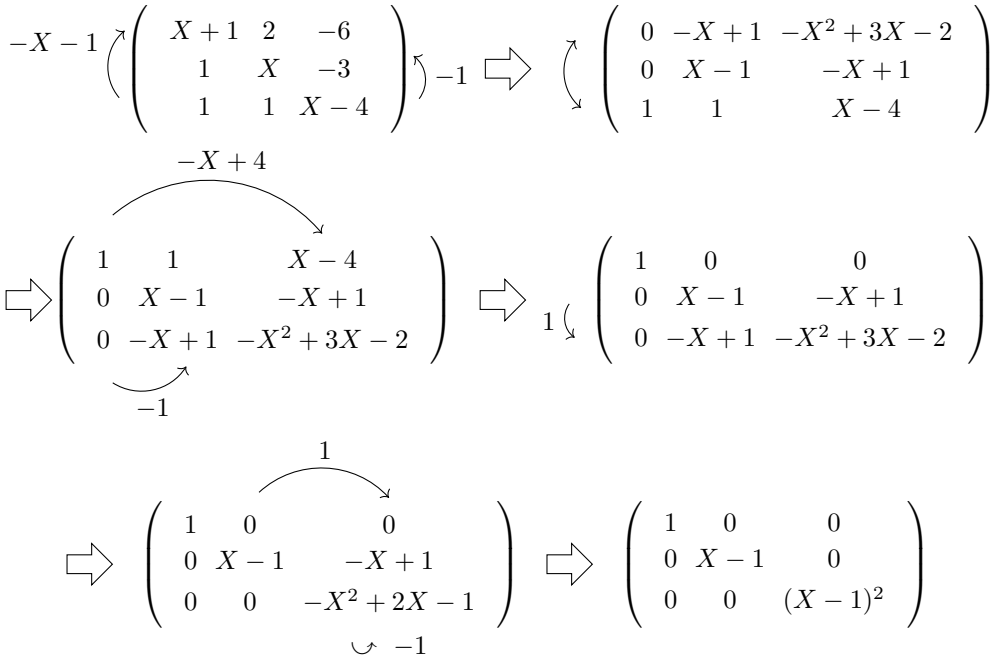
我们知道在 $d_i\neq 1$ 的前提下, $\varphi(g_i)$ 正是有理标准形中对应于 d_i 的分块的首个基向量.

有理标准形的计算按此化约为 Smith 标准形定理 12.7.6 包含的算法. 它只涉及多项式带余除法, 不必求根或作因式分解. 以下考量一则例子.

例 13.4.4 取 $V=F^3$ 而 T 对应到矩阵 $\boldsymbol{A}=\begin{pmatrix}-1&-2&6\\-1&0&3\\-1&-1&4\end{pmatrix}$. 为了将 $X\cdot\mathbf{1}_{3\times 3}-\boldsymbol{A}$ 化

[2)] 在一些教材中, 变元 X 被记为 λ, 相应的矩阵则称为 λ-矩阵.

为所需形式, 我们进行初等行和列变换:



由此读出 $\boldsymbol{A}$ 的不变因子 $d_2 = X-1$ 和 $d_3 = (X-1)^2$.

基于此前结果, 容易按 $F[X]$ 上的相抵关系来理解矩阵的共轭关系. 首先仿照域上的思路, 对于任意交换环 R, 设 $\boldsymbol{L}, \boldsymbol{M} \in \mathrm{M}_{m\times n}(R)$. 若存在 $\boldsymbol{P} \in \mathrm{GL}(n,R)$ 和 $\boldsymbol{Q} \in \mathrm{GL}(m,R)$ 使得 $\boldsymbol{L} = \boldsymbol{QMP}$, 则称 $\boldsymbol{L}$ 与 $\boldsymbol{M}$ 相抵; 这是 $\mathrm{M}_{m\times n}(R)$ 上的等价关系.

命题 13.4.5 对所有 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$, 以下陈述等价:

(i) $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 在 $\mathrm{M}_{n\times n}(F)$ 中共轭;

(ii) $X\cdot \mathbf{1}_{n\times n} - \boldsymbol{A}$ 和 $X\cdot \mathbf{1}_{n\times n} - \boldsymbol{B}$ 在 $\mathrm{M}_{n\times n}(F[X])$ 中相抵.

证明 将 $X\cdot \mathbf{1}_{n\times n} - \boldsymbol{A}$ 简记为 $\Lambda_{\boldsymbol{A}}$. 向 $\boldsymbol{A}$ 的有理标准形中的不变因子插入 1, 可得 $F[X]$ 的一列首一元素 $d_1 \mid \cdots \mid d_n$. 同理, 对 $\boldsymbol{B}$ 有 $\Lambda_{\boldsymbol{B}}$ 和 $d_1' \mid \cdots \mid d_n'$.

(i) $\Longrightarrow$ (ii). 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 共轭, 则不变因子的唯一性蕴涵 $d_i = d_i'$ 恒成立. 代入命题 13.4.2 可知 $\Lambda_{\boldsymbol{A}}$ 和 $\Lambda_{\boldsymbol{B}}$ 都相抵于 $\mathrm{diag}(d_1,\dots,d_n)$.

(ii) $\Longrightarrow$ (i). 若 $\Lambda_{\boldsymbol{A}}$ 和 $\Lambda_{\boldsymbol{B}}$ 相抵, 则命题 13.4.2 蕴涵

$$\mathrm{diag}(d_1,\dots,d_n) \text{ 和 } \mathrm{diag}(d_1',\dots,d_n') \text{ 相抵.}$$

应用推论 12.7.9 推得 $d_i = d_i'$ 恒成立, 故 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 共轭. □

13.5 Jordan 标准形

本节伊始, 考虑的仍是任意域 F. 继续选定 $n \in \mathbb{Z}_{\geq 1}$. Jordan 标准形的出发点是幂零矩阵. 幂零元的概念可在任意环中定义.

定义 13.5.1 设 R 为环, $r \in R$. 若存在 $d \geq 1$ 使得 $r^d = 0_R$, 则称 r 为**幂零**的. 条件中最小可能的 d 称为 r 的**幂零指数**.

本节仅将上述定义应用于 $R = \mathrm{M}_{n\times n}(F)$ (矩阵) 或 $R = \mathrm{End}(V)$ (线性映射) 两种情形, 其中 V 是有限维 F-向量空间.

约定 13.5.2 为简化符号, 今后对所有 $\lambda \in F$ 将 $\lambda \cdot \mathbf{1}_{n\times n}$ (或 $\lambda \cdot \mathrm{id}_V$) 简记为 λ.

当 $T \in \mathrm{End}(V)$ 给定, $V_{[\lambda]} := \ker\left((T-\lambda)^n\right)$ 即 $\lambda \in F$ 在 V 中对应的广义特征子空间 ($n = \dim V$), 它非零当且仅当 λ 是 T 的特征值, 此外 V 有直和分解 $V = \bigoplus_\lambda V_{[\lambda]}$.

命题 13.5.3 设 $\dim V = n \in \mathbb{Z}_{\geq 1}$, 而 $T \in \mathrm{End}(V)$. 以下陈述相互等价:

(i) T 幂零;

(ii) 存在 $k \in \mathbb{Z}_{\geq 1}$ 使得 $\mathrm{Min}_T = X^k$;

(iii) $\mathrm{Char}_T = X^n$;

(iv) $V = V_{[0]}$.

事实上, (ii) 中的 k 正是 T 的幂零指数, 而 $k \leq n$.

证明 (i) $\Longrightarrow$ (ii): 若 $T^d = 0_V$, 则 $\mathrm{Min}_T \mid X^d$ 导致存在 $1 \leq k \leq d$ 使得 $\mathrm{Min}_T = X^k$.

(ii) $\Longrightarrow$ (iii): 一种方法是代入 (13.3.1).

(iii) $\Longrightarrow$ (iv): 此时唯一特征值是 $\lambda = 0$, 故 V 的广义特征子空间直和分解化为 $V = V_{[0]}$.

(iv) $\Longrightarrow$ (i): 按定义 $T - \lambda$ 在 $V_{[\lambda]}$ 上是幂零的; 代入 $\lambda = 0$.

关于幂零指数的断言来自 Min_T 的定义. □

定义 13.5.4 (Jordan 块) 设 $\lambda \in F$ 而 $d \in \mathbb{Z}_{\geq 1}$.

⋆ 特征值为 λ 的 $d\times d$ 上三角 Jordan 块 (简称 Jordan 块) 意谓以下矩阵

$$\boldsymbol{J}_d(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} \in \mathrm{M}_{d\times d}(F);$$

⋆ 特征值为 λ 的 $d\times d$ 下三角 Jordan 块意谓

$$\boldsymbol{J}_d^{\text{下}}(\lambda) := {}^{\mathrm{t}}\boldsymbol{J}_d(\lambda).$$

此处规定 $\boldsymbol{J}_1(\lambda) = \boldsymbol{J}_1^{\text{下}}(\lambda) = \lambda \in F = \mathrm{M}_{1\times 1}(F)$.

鉴于 $\boldsymbol{J}_d(\lambda) = \lambda\cdot \mathbf{1}_{d\times d} + \boldsymbol{J}_d(0)$, 特征值为 0 的 Jordan 块是更基本的对象. 下三角情形亦同, 而且它正是多项式 X^d 的友矩阵:

$$\boldsymbol{J}_d^{\text{下}}(0) = {}^{\mathrm{t}}\boldsymbol{J}_d(0) = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & 0 \end{pmatrix} = \boldsymbol{C}_{X^d}.$$

回顾 §13.3 的简单计算, $\mathrm{Min}_{\boldsymbol{J}_d^{\text{下}}(0)} = X^d = \mathrm{Char}_{\boldsymbol{J}_d^{\text{下}}(0)}$. 因此 $\boldsymbol{J}_d^{\text{下}}(0)$ 是幂零指数为 d 的幂零矩阵. 用 $\lambda\cdot\mathbf{1}_{d\times d}$ 平移立得

$$\mathrm{Min}_{\boldsymbol{J}_d^{\text{下}}(\lambda)} = (X-\lambda)^d = \mathrm{Char}_{\boldsymbol{J}_d^{\text{下}}(\lambda)}; \tag{13.5.1}$$

再取转置便有

$$\mathrm{Min}_{\boldsymbol{J}_d(\lambda)} = (X-\lambda)^d = \mathrm{Char}_{\boldsymbol{J}_d}(\lambda). \tag{13.5.2}$$

事实上, 任何幂零矩阵都能唯一地分解为特征值为 0 的 Jordan 块, 精确到共轭.

引理 13.5.5 设 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 幂零, 则

⋆ 存在唯一的正整数列 $1\le b_1\le\cdots\le b_r$ 使得 $\sum_{i=1}^r b_i = n$ 而 $\boldsymbol{A}$ 共轭于分块对角矩阵

$$\begin{pmatrix} \boxed{\boldsymbol{J}_{b_1}(0)} & & \\ & \ddots & \\ & & \boxed{\boldsymbol{J}_{b_r}(0)} \end{pmatrix};$$

⋆ $\boldsymbol{A}$ 的幂零指数为 b_r;

⋆ 若在以上陈述中以 $\boldsymbol{J}_{b_i}^{\text{下}}(0)$ 代替 $\boldsymbol{J}_{b_i}(0)$, 论断依然成立.

类似地, 设 V 是 n 维 F-向量空间而 $T\in\mathrm{End}(V)$ 幂零, 则存在 V 的有序基和 $1\leq b_1\leq\cdots\leq b_r$, 使得 T 表作上述分块对角矩阵, 而 $(b_j)_{j=1}^r$ 由 T 完全确定, T 的幂零指数为 b_r.

证明 处理矩阵版本即足. 先来证明涉及 $\boldsymbol{J}_{b_i}^{\text{下}}(0)$ 的版本.

记 $\boldsymbol{A}$ 的幂零指数为 d. 取 $\boldsymbol{A}$ 的有理标准形 $\mathrm{diag}(\boldsymbol{C}_{f_1},\dots,\boldsymbol{C}_{f_r})$, 其中的不变因子 $f_1\mid\cdots\mid f_r$ 满足 $f_r=\mathrm{Min}_{\boldsymbol{A}}=X^d$ (命题 13.5.3). 于是可将 f_i 写作 X^{b_i}, 其中 $1\leq b_1\leq\cdots\leq b_r=d$; 相应地, $\boldsymbol{C}_{f_i}=\boldsymbol{J}_{b_i}^{\text{下}}(0)$. 综上, $\boldsymbol{A}$ 共轭于 $\mathrm{diag}\left(\boldsymbol{J}_{b_1}^{\text{下}}(0),\dots,\boldsymbol{J}_{b_r}^{\text{下}}(0)\right)$.

基于表法 $\mathrm{diag}\left(\boldsymbol{J}_{b_1}^{\text{下}}(0),\dots\right)$ 与有理标准形的联系, 资料 $(b_i)_{i=1}^r$ 对 $\boldsymbol{A}$ 的唯一性归结为不变因子 X^{b_i} 的唯一性.

至于 $\boldsymbol{J}_{b_i}(0)$ 的版本, 观察到取转置既不改变幂零性质和幂零指数, 也不影响矩阵的共轭关系, 所以对 ${}^{\text{t}}\boldsymbol{A}$ 应用上一步的结论即足. □

从以上证明可见下三角 Jordan 块就本书的进路而言更为自然. 后续讨论中采用上三角 Jordan 块的唯一理由是习俗.

行将介绍的 Jordan 标准形仅适用于特征多项式在 F 上分裂的情形; 见定义 6.6.2. 我们以线性映射的语言表述.

定理 13.5.6 (Jordan 标准形) 设 V 为 n 维 F-向量空间, $T\in\mathrm{End}(V)$. 设 Char_T 在 F 上分裂, 记其相异根为 $\lambda_1,\dots,\lambda_m\in F$. 存在 V 的有序基, 使得 T 表为分块对角矩阵 $\mathrm{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_m)$, 其中

$$\boldsymbol{A}_j:=\begin{pmatrix}\boxed{\boldsymbol{J}_{b_{1,j}}(\lambda_j)}&&\\&\ddots&\\&&\boxed{\boldsymbol{J}_{b_{r_j,j}}(\lambda_j)}\end{pmatrix},\quad 1\leq j\leq m,$$

而每个 j 对应的正整数列

$$b_{1,j}\leq\cdots\leq b_{r_j,j},\quad r_j\in\mathbb{Z}_{\geq 1}.$$

由 T 唯一确定. 上述矩阵 $\mathrm{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_m)$ 称为 T 的 **Jordan 标准形**.

若在以上陈述中以 $\boldsymbol{J}_{b_{i,j}}^{\text{下}}(\lambda_j)$ 代替 $\boldsymbol{J}_{b_{i,j}}(\lambda_j)$, 论断依然成立.

证明 取对 T 的广义特征子空间分解 $V=V_{[\lambda_1]}\oplus\cdots\oplus V_{[\lambda_m]}$. 每个 $V_{[\lambda_j]}$ 都是 T-不变子空间, 而 $d_j:=\dim V_{[\lambda_j]}>0$. 记 T 在 $V_{[\lambda_j]}$ 上的限制为 T_j.

选定 $1\leq j\leq m$. 已知 $T_j-\lambda_j$ 幂零, 故引理 13.5.6 给出有序基和 $1\leq b_{1,j}\leq\cdots\leq b_{r_j,j}$, 使得 $T_j-\lambda_j$ 表为矩阵 $\mathrm{diag}\left(\boldsymbol{J}_{b_{1,j}}(0),\dots\right)$. 用 λ_j 平移可知 T_j 表为矩阵 $\boldsymbol{A}_j:=\mathrm{diag}\left(\boldsymbol{J}_{b_{1,j}}(\lambda_j),\dots\right)$. 至此证得 Jordan 标准形的存在性.

至于唯一性, 设 T 已有如是矩阵表法. 选定 j. 注意到 $\boldsymbol{A}_j-\lambda_j$ 幂零而 $\lambda_1,\dots,\lambda_m$ 相异, 由此可推得分块 $\boldsymbol{A}_j$ 必对应 T_j. 关于 $b_{1,j}\le\cdots\le b_{1,r_j}$ 的唯一性因而化到引理 13.5.5 处理过的幂零情形. □

证明表明分块 $\boldsymbol{A}_j$ 对应到 T 在广义特征子空间 $V_{[\lambda_j]}$ 上的作用. 综之, Jordan 标准形不过是有理标准形的第二种形式在 Char_T 分裂时的简单变形.

在 Jordan 标准形的表述中, 相应的有序基又称为 Jordan 基. 关于特征多项式分裂的前提是必要的, 这是因为 Jordan 块的特征多项式分裂. 若假定 F 是代数闭域, 例如 $F=\mathbb{C}$, 则分裂条件总是成立.

练习 13.5.7 设 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 而且 $\mathrm{Char}_{\boldsymbol{A}}$ 分裂.

(i) 说明 $\boldsymbol{A}$ 可对角化当且仅当其所有 Jordan 块都是 1×1 的.

(ii) 说明 $\mathrm{Char}_{\boldsymbol{A}}=\prod_{j=1}^m(X-\lambda_j)^{a_j}$ 而 $\mathrm{Min}_{\boldsymbol{A}}=\prod_{j=1}^m(X-\lambda_j)^{b_{r_j,j}}$, 其中 $a_j:=b_{1,j}+\cdots+b_{r_j,j}$.

共轭于 $\boldsymbol{J}_n(0)$, 或等价地说幂零指数为 n 的幂零矩阵称为 $\mathrm{M}_{n\times n}(F)$ 中的**主幂零元**, 以下练习 (取特例 $\lambda=0$) 予以刻画.

练习 13.5.8 设 $n\in\mathbb{Z}_{\ge 2}$, $\lambda\in F$ 而 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$ 形如

$$\begin{pmatrix}\lambda & & & \\ t_1 & \ddots & & \\ \vdots & \ddots & \ddots & \\ * & \cdots & t_{n-1} & \lambda\end{pmatrix}\quad\text{或}\quad\begin{pmatrix}\lambda & t_1 & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & t_{n-1} \\ & & & \lambda\end{pmatrix}$$

其中 $t_1,\dots,t_{n-1}\in F$, 留白部分为 0. 证明在两种情况下, $\boldsymbol{A}$ 共轭于 $\boldsymbol{J}_n(\lambda)$ 的充要条件都是 $\prod_{i=1}^{n-1}t_i\in F^\times$.

提示〉 处理上三角版本即足. 首先用一个平移化约到 $\lambda=0$ 情形, 此时 $\boldsymbol{A}$ 幂零. 其次观察到 $\boldsymbol{A}$ 共轭于 $\boldsymbol{J}_n(0)$ 当且仅当 $\boldsymbol{A}$ 的幂零指数为 n. 以

$$\boldsymbol{A}\boldsymbol{e}_i=\begin{cases}t_{i-1}\boldsymbol{e}_{i-1}+\langle\boldsymbol{e}_{i-2},\dots\rangle, & i>2,\\ t_{i-1}\boldsymbol{e}_{i-1}, & i=2,\\ \boldsymbol{0}, & i=1,\end{cases}$$

研究 $\boldsymbol{A},\boldsymbol{A}^2,\dots$ 在基上的作用, 可给出 $\boldsymbol{A}$ 的幂零指数为 n 的充要条件.

13.6 Jordan 标准形的计算

由于 Jordan 标准形不外是有理标准形的简单改写, 在特征多项式分裂的前提下, 只要能求根, §13.4 的方法同样足以确定 Jordan 标准形.

然而 Jordan 标准形也可以通过计算秩来确定, 这是本节所要介绍的方法. 依旧从幂零情形入手.

引理 13.6.1 设 V 是 n 维 F-向量空间, $T \in \mathrm{End}(V)$ 幂零.

(i) 在 T 的 Jordan 标准形中, Jordan 块的总数为 $n - \mathrm{rk}(T)$;

(ii) 对于每个 $d \geq 1$, 标准形中的 $d \times d$ Jordan 块的个数 $N(d)$ 满足

$$N(d) = \mathrm{rk}\left(T^{d+1}\right) + \mathrm{rk}\left(T^{d-1}\right) - 2\,\mathrm{rk}\left(T^{d}\right).$$

证明 对于 (i), 设 $d \geq 1$. 由于 $\boldsymbol{J}_d(0)$ 已是简化行梯矩阵, 秩为 $d-1$, 故 $\mathrm{rk}(T)$ 等于 n 减去 Jordan 块的总数.

对于 (ii), 将 T 的 Jordan 块尺寸依序标为 $b_1 \leq \cdots \leq b_r$. 给定 $b \geq 1$ 和 $k \geq 0$, 简单地计算 $\boldsymbol{J}_b(0)^k$ (参见练习 13.5.8) 可见

$$\mathrm{rk}\left(\boldsymbol{J}_b(0)^k\right) = \begin{cases} 0, & k > b, \\ b-k, & k \leq b. \end{cases}$$

由此知当 $d \geq 1$ 时

$$\begin{aligned} \mathrm{rk}\left(T^{d+1}\right) - \mathrm{rk}\left(T^{d}\right) &= \sum_{j:b_j \geq d+1} (b_j - d - 1) - \sum_{j:b_j \geq d} (b_j - d) \\ &= \sum_{j:b_j \geq d+1} (-1), \\ \mathrm{rk}\left(T^{d}\right) - \mathrm{rk}\left(T^{d-1}\right) &= \sum_{j:b_j \geq d} (-1). \end{aligned}$$

两式相减即 $\sum_{j:b_j=d} 1 = N(d)$. □

现在给出对一般的 T 计算 Jordan 块个数的一种方法. 仍采用约定 13.5.2.

定理 13.6.2 设 V 是 n 维 F-向量空间, $T \in \mathrm{End}(V)$ 满足 Char_T 分裂. 记 T 的相异特征值为 $\lambda_1, \ldots, \lambda_m$. 选定 $1 \leq j \leq m$.

(i) 在 T 的 Jordan 标准形中, 特征值为 λ_j 的 Jordan 块的总数为 $n - \mathrm{rk}(T - \lambda_j)$;

(ii) 对于每个 $d \geq 1$, 特征值为 λ_j 的 $d \times d$ Jordan 块的个数 $N_j(d)$ 满足

$$N_j(d) = \mathrm{rk}\left((T-\lambda_j)^{d+1}\right) + \mathrm{rk}\left((T-\lambda_j)^{d-1}\right) - 2\,\mathrm{rk}\left((T-\lambda_j)^{d}\right).$$

证明 将 T 的 Jordan 标准形写作 $\operatorname{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_m)$, 使得 $\boldsymbol{A}_k$ 对应到 T 在 $V_{[\lambda_k]}$ 上的限制; 命 $d_k := \dim V_{[\lambda_k]}$.

对于 (i), 观察到

$$n-\operatorname{rk}(T-\lambda_j)=\sum_{k=1}^{m}\left(d_k-\operatorname{rk}\left(\boldsymbol{A}_k-\lambda_j\right)\right).$$

由于 $\boldsymbol{A}_k$ 的特征值只有 λ_k, 故 $k\neq j$ 时 $\boldsymbol{A}_k-\lambda_j$ 可逆, 秩为 d_k. 于是上式仅有 $k=j$ 的项; 将 $\boldsymbol{A}_j-\lambda_j$ 代入引理 13.6.1 (i) 便得到特征值 λ_j 的 Jordan 块总数.

至于 (ii), 对所有 $d\geq 1$ 将断言之等式右边写成

$$\sum_{k=1}^{m}\left[\operatorname{rk}\left(\left(\boldsymbol{A}_k-\lambda_j\right)^{d+1}\right)+\operatorname{rk}\left(\left(\boldsymbol{A}_k-\lambda_j\right)^{d-1}\right)-2\operatorname{rk}\left(\left(\boldsymbol{A}_k-\lambda_j\right)^{d}\right)\right],$$

则同理可知除 $k=j$ 外的项全为零. 由引理 13.6.1 (ii) 即得 $N_j(d)$. □

一旦将 T 具体表作矩阵, 并且确定了 λ_j, 公式中的秩便能以消元法计算. 且看一则例子.

例 13.6.3 取 $F=\mathbb{Q}$ 并考虑矩阵

$$\boldsymbol{A}=\begin{pmatrix}2&3&2\\1&8&2\\-2&-14&-3\end{pmatrix}.$$

计算 $\operatorname{Char}_{\boldsymbol{A}}=(X-1)(X-3)^2$ 可知 1 在 $\boldsymbol{A}$ 的 Jordan 标准形的对角线上恰好出现一次, 对应的块是 1×1 的; 具体计算 $\ker(\boldsymbol{A}-1)$ 可得一个特征向量

$$\boldsymbol{v}_1=\begin{pmatrix}2\\0\\-1\end{pmatrix}.$$

接着计算 $\operatorname{rk}(\boldsymbol{A}-3)=2$, 故它对应到一个 2×2 的 Jordan 块; 块的第一列对应到特征值为 2 的特征向量, 解得

$$\boldsymbol{v}_2=\begin{pmatrix}1\\-1\\2\end{pmatrix},$$

块的第二列则对应到 $(\boldsymbol{A}-3)\boldsymbol{v}_3=\boldsymbol{v}_2$ 的任意解 $\boldsymbol{v}_3$, 具体计算给出特解

$$\boldsymbol{v}_3=\begin{pmatrix}-1\\0\\0\end{pmatrix}.$$

综上得出 Jordan 基 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 使得

$$\boldsymbol{A}\left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{array}\right) = \left(\begin{array}{c|c|c} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{array}\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

观察到上式是关于 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ (或其坐标) 的线性方程组. 只要通过秩的计算确定 A 的 Jordan 标准形, 则也可以按此列式解出 Jordan 基.

练习 13.6.4 Jordan 标准形有助于计算矩阵的高次幂. 试以例 13.6.3 的矩阵 $\boldsymbol{A}$ 为例, 给出计算 $\boldsymbol{A}^{1898}$ 的具体方法.

习题

1. 使用有理标准形说明若 F 是域 E 的子域, 则 $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{M}_{n\times n}(F)$ 共轭当且仅当它们在 $\mathrm{M}_{n\times n}(E)$ 中共轭.

提示〉 说明扩域不影响不变因子即可. 将 F^n (或 E^n) 用 $\boldsymbol{A}$ 升级为 $F[X]$-模 (或 $E[X]$-模). 说明若有 $F[X]$-模同构 $\bigoplus_{i=1}^{k} F[X]/(f_i) \xrightarrow{\sim} F^n$, 其中 $f_1 \mid \cdots \mid f_k$, 便有 $E[X]$-模同构 $\bigoplus_{i=1}^{k} E[X]/(f_i) \xrightarrow{\sim} E^n$.

2. 说明 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 共轭于某个友矩阵 $\boldsymbol{C}_f$ 的充要条件是 $\mathrm{Min}_{\boldsymbol{A}} = \mathrm{Char}_{\boldsymbol{A}}$.

3. 说明 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$ 总和它的转置 ${}^{\mathrm{t}}\boldsymbol{A}$ 共轭.

提示〉 一种方法是以有理标准形化到 $\boldsymbol{A} = \boldsymbol{C}_f$ 情形, 然后运用前一题的结论. 另一种方法是在充分大的扩域上取 Jordan 标准形, 然后直接证明 $\boldsymbol{J}_d(0)$ 共轭于 ${}^{\mathrm{t}}\boldsymbol{J}_d(0)$.

4. 设 $\boldsymbol{A} \in \mathrm{M}_{5\times 5}(\mathbb{C})$ 满足 $\mathrm{Char}_{\boldsymbol{A}} = (X-3)^4(X+2)$ 而 $\mathrm{rk}(\boldsymbol{A} - 3\cdot \boldsymbol{1}_{5\times 5}) = 2$. 试确定 $\boldsymbol{A}$ 的 Jordan 标准形.

5. 判断以下矩阵在 $\mathbb{Q}$ 上有无 Jordan 标准形; 如有, 写下其 Jordan 标准形.

(a) $\begin{pmatrix} -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$, (c) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & 1 \end{pmatrix}$.

6. 求以下矩阵的 Jordan 标准形.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

7. 具体对所有 $d,k\geq 1$ 计算 $\boldsymbol{J}_d(0)^k$, 确定它的秩并写下 Jordan 标准形.

8. 设 $k\in\mathbb{Z}_{\geq 1}$ 在域 F 中的像非零, 而 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$ 的特征值全为 1. 证明 $\boldsymbol{A}^k$ 与 $\boldsymbol{A}$ 共轭.

提示〉 先化约到 $\boldsymbol{A}=\boldsymbol{J}_n(1)$ 的情形. 算出 $\boldsymbol{A}^k$ 的次对角线是 $k,\dots,k$, 然后代入练习 13.5.8 的结论.

9. 设 F 为满足 $\mathrm{char}(F)\neq 2$ 的域, $\lambda\in F^\times$ 在 F 中有平方根. 证明 $\boldsymbol{J}_n(\lambda)$ 在 $\mathrm{M}_{n\times n}(F)$ 中有平方根. 由此证明任何可逆的 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{C})$ 都有平方根. 试探讨更高次的情形.

10. 设 λ 是域 F 的非零元, 求 $\boldsymbol{J}_d(\lambda)^2$ 的 Jordan 标准形 ($d\geq 1$). 提示〉 按照 2λ 是否为 0 分开讨论.

11. 证明复矩阵指数映射

$$\exp:\mathrm{M}_{n\times n}(\mathbb{C})\to \mathrm{GL}(n,\mathbb{C}),\quad \exp(\boldsymbol{A})=\sum_{k=0}^{\infty}\frac{\boldsymbol{A}^k}{k!}$$

对所有 n 皆满.

提示〉 为了说明所有 $\boldsymbol{B}\in\mathrm{GL}(n,\mathbb{C})$ 都能写成 $\exp(\boldsymbol{A})$, 基于乘性 Jordan–Chevalley 分解, 处理 $\boldsymbol{B}\in\mathbb{C}^\times\boldsymbol{1}_{n\times n}$ 和 $\boldsymbol{B}-\boldsymbol{1}_{n\times n}$ 幂零这两种情形即可. 前者相当于 $n=1$ 的情形, 后者则通过 $\exp(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{-1}\exp(\boldsymbol{A})\boldsymbol{P}$ 归结为特例 $\boldsymbol{B}=\boldsymbol{J}_n(1)$. 从练习 13.5.8 推导 $\exp(\boldsymbol{J}_n(0))$ 共轭于 $\boldsymbol{J}_n(1)$.

12. 假定读者了解数学分析中关于闭包的概念. 将 $\mathrm{M}_{n\times n}(\mathbb{C})$ 等同于 $\mathbb{C}^{n^2}$, 以谈论其中的极限与闭包. 证明以下关于 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(\mathbb{C})$ 的陈述等价.

(i) $\boldsymbol{A}$ 幂零;

(ii) 存在可逆之 $\boldsymbol{P}$ 使得 $\lim_{m\to+\infty}\boldsymbol{P}^m\boldsymbol{A}\boldsymbol{P}^{-m}=\boldsymbol{0}_{n\times n}$;

(iii) $\boldsymbol{0}_{n\times n}$ 属于共轭类 $\left\{\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}:\boldsymbol{P}\text{ 可逆}\right\}$ 在 $\mathrm{M}_{n\times n}(\mathbb{C})$ 中的闭包;

(iv) $\mathrm{Char}_{\boldsymbol{A}}=X^n$.

13. (P. Deligne) 设 V 为某个域上的向量空间, $N\in\mathrm{End}(V)$, 而且存在 $d\geq 0$ 使得 $N^{d+1}=0_V$. 通过对 d 递归, 证明存在唯一一列子空间

$$\cdots\subset V_{i-1}\subset V_i\subset V_{i+1}\subset\cdots,\quad V_i\subset V\quad(i\in\mathbb{Z}),$$

使得

- ⋆ 当 i 充分小 (或充分大) 时 $V_i=\{0\}$ (或 $V_i=V$);
- ⋆ $N(V_i)\subset V_{i-2}$ 对所有 i 成立;
- ⋆ 记 $\mathrm{gr}_i V:=V_i/V_{i-1}$, 则对所有 $k\geq 0$ 皆有线性同构

$$\begin{aligned}\overline{N^k}:\mathrm{gr}_k V&\xrightarrow{\sim}\mathrm{gr}_{-k}V\\ x+V_{k-1}&\mapsto N^k(x)+V_{-k-1}.\end{aligned}$$

提示〉 先处理存在性. 当 $d=0$ 时 $N=0_V$, 取 $V_{-1}=\{0\}$ 而 $V_0=V$. 设 $d\geq 1$, 取

$$(V_{-d-1}=\{0\})\subset\Big(V_{-d}=\mathrm{im}(N^d)\Big)\subset\Big(V_{d-1}=\ker(N^d)\Big)\subset(V_d=V).$$

于是 N^d 诱导 $\mathrm{gr}_d V = V/\ker(N^d) \xrightarrow{\sim} \mathrm{im}(N^d) = \mathrm{gr}_{-d} V$. 由于 $N^{2d} = 0_V$ 而 N^d 在 $\ker(N^d)/\mathrm{im}(N^d)$ 上诱导零映射, 可在 $\ker(N^d)/\mathrm{im}(N^d)$ 上递归地构造剩下的子空间列. 改造以上论证以说明子空间列 $(V_i)_{i\in\mathbb{Z}}$ 唯一.

14. 承上题, 进一步对所有 i 考虑 N 诱导的线性映射 $\overline{N}_i : \mathrm{gr}_i(V) \to \mathrm{gr}_{i-2}(V)$. 称 $P_i := \ker(\overline{N}_i)$ 为 $\mathrm{gr}_i(V)$ 的本原部分. 这种结构常见于代数几何及相关领域, 亦可由 Lie 代数 $\mathfrak{sl}(2)$ 的表示理论解释.

(i) 说明 $i > 0$ 时 $\overline{N}_i$ 单, 从而 $P_i = \{0\}$.

提示〉 用 N^i 诱导之 $\mathrm{gr}_i(V) \xrightarrow{\sim} \mathrm{gr}_{-i}(V)$.

(ii) 设 $i \geq 0$, 证明 $\mathrm{gr}_{-i}(V)$ 是 P_{-i} 和 $\mathrm{im}\left[\mathrm{gr}_{i+2}(V) \hookrightarrow \mathrm{gr}_{-i}(V)\right]$ (由 N^{i+1} 诱导) 的直和.

提示〉 合成映射 $\mathrm{gr}_{i+2}(V) \to \mathrm{gr}_{-i}(V) \xrightarrow{\overline{N}_{-i}} \mathrm{gr}_{-i-2}(V)$ 是同构.

(iii) 对所有 i 给出典范同构 $\mathrm{gr}_i(V) \simeq \bigoplus_{\substack{j\geq|i|\\ j\equiv i \pmod 2}} P_{-j}$.

(iv) 证明 $\overline{N}_i$ 在此同构之下化为直和之间自明的包含映射 (当 $i > 1$), 投影映射 (当 $i < 1$), 或恒等 ($i = 1$).

(v) 证明 $N(V_i) = \mathrm{im}(N) \cap V_{i-2}$.

提示〉 要点在于证 $\supset$. 基于 (iv), 在 $i \leq 1$ 的情形证明 $N(V_i) = V_{i-2}$; 在 $i \geq 0$ 的情形证明 N 诱导单射 $V/V_i \to V/V_{i-2}$, 从而 $N^{-1}(V_{i-2}) \subset V_i$.

(vi) 说明若将向量空间替换为某个环 R 上的左模 (或右模), 则包括上一题在内的先前所有论断依然成立.

15. 承上题, 取 V 为有限维 F-向量空间. 基于 Jordan 标准形, 之前研究的对象可化约到 $V = F^{d+1}$ 而 N 对应于下三角 Jordan 块 $\boldsymbol{J}^{\text{下}}_{d+1}(0)$ 的情形. 方便起见, 将 F^{d+1} 的标准有序基标号为

$$\boldsymbol{e}_d,\ \boldsymbol{e}_{d-2},\ \ldots,\ \boldsymbol{e}_{-d+2},\ \boldsymbol{e}_{-d}.$$

证明 $V_i = \bigoplus_{j\leq i} F\boldsymbol{e}_j$, 然后描述 P_i.

第十四章 仿射空间与射影空间

在引入向量空间的概念时, 我们曾经从关于平面或空间向量的直观来启发, 这种类比其实是片面的. 几何直观中的向量是从一点到另一点的方向, 带有大小, 此概念不依赖坐标系或原点的选择, 而且只有头尾相接的向量才能相加. 向量空间的抽象理论虽然摆脱了坐标的桎梏, 但仍然以零向量作为指定的原点, 所以它们对于几何学并非完全贴切的模型. 本章 §§14.1–14.2 介绍的仿射空间为此给出一套合理的解答.

给定域 F 上的向量空间 V, 相应地有加法群 $(V,+)$. 所谓 V 作用下的仿射空间, 是指带有自由且传递的 $(V,+)$-作用的非空集 E (定义 14.1.1); 本章将作用写作 $(x,v)\mapsto x+v$ 之形 ($x\in E$, $v\in V$). 对于 $x,y\in E$, 几何直观中的向量 $\vec{xy}$ 无非是由 $x+v=y$ 所唯一确定的 $v\in V$, 也写作 $v:=y-x$. 另将 E 的维数定义为 V 的维数.

举例来说, 从 V 本身的加法可得仿射空间 $(V,+)$, 而一旦在仿射空间 E 中指定一个基点 x_0, 便能通过 $v\mapsto x_0+v$ 将 E 等同于 V. 尽管仿射空间在选定基点后等同于向量空间, 但关于仿射空间的一切概念都需要以无关基点的方式来表达. 以下是一部分对比.

仿射空间中的概念	参考	向量空间的类比
仿射组合	定义 14.1.4	线性组合
仿射基	定义 14.1.7	基
仿射子空间	定义 14.1.10	子空间
仿射线性映射	定义–命题 14.2.1	线性映射
仿射变换群	定义–命题 14.2.6	一般线性群

在实数域 $\mathbb{R}$ 上, 以上概念都有清晰的几何含义, 譬如求两点 x 和 y 的仿射组合相当于求线段 xy 上的定比分点.

进一步, 对于实内积空间作用下的仿射空间 E, 还能定义两点 x 和 y 的距离为 $d(x,y):=\|y-x\|$, 使之成为度量空间. 保距自映射又称为刚体运动, 这是 §14.3 的主题. 定理 14.3.3 将分类 E 上的刚体运动: 一旦选定基点 $o\in E$, 则它们唯一地分解为平移与相对于 o 的保距线性映射; 特别地, 刚体运动总是仿射线性的.

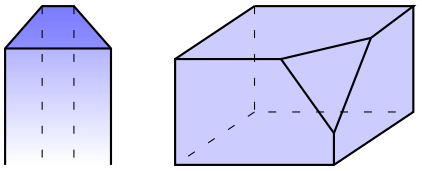
随后的 §§14.4–14.5 介绍射影空间和射影变换. 就历史来看, 射影空间起源于透视法绘画的研究, 如今它的身影已遍及数学各领域. 简言之, 非零 F-向量空间 V 中的所有 1 维子空间构成射影空间 $\mathbb{P}(V)$, 另记

$$\begin{aligned}\mathbb{P}^n &:= \mathbb{P}(F^{n+1}) \\ &= \big(F^{n+1} \smallsetminus \{\mathbf{0}\}\big) \Big/ (x_0,\dots,x_n) \sim (tx_0,\dots,tx_n),\ t \in F^\times.\end{aligned}$$

由 $(x_0,\dots,x_n)$ 在 $\mathbb{P}^n$ 中确定的等价类记为 $(x_0 : \cdots : x_n)$, 此记法称为齐次坐标, 而 $\mathbb{P}^n$ 能表成 $n+1$ 份 n 维仿射空间之并, 或者表成一个 n 维仿射空间与一系列 “无穷远点” 之并; 这是 “仿射” 一词的来由.

就数学来看, $\mathbb{P}^n$ 的许多几何性质比 n 维仿射空间更容易处理. 射影几何具有丰富的内容, 这两节仅触及基本概念. 此外, 定义–命题 14.5.6 还将介绍称为交比的一种经典射影不变量.

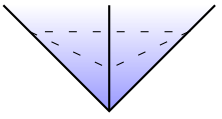
实仿射空间为多面体理论提供了一套方便而不失直观的理论框架. 我们先在 §14.6 介绍实仿射空间 E 中的凸集和实向量空间中的凸锥, 其中涉及的凸组合, 凸包与半空间是尔后将频繁登场的概念. 随后, §14.7 探讨有限维实仿射空间中的凸多面体, 它们是由有限多个半空间 $\alpha \geq 0$ 取交截出的非空子集, 其中 $\alpha : E \to \mathbb{R}$ 为非常值仿射线性映射; 若取定坐标系, 这也相当于是由有限多个线性不等式 (容许常数项非零) 截出的非空区域, 示意如下.



在多面体理论中, 低维的几何直观不仅起到敲门砖的作用, 还是理解后续许多定义和证明的钥匙, 这种现象在全书中是少见的.

当然, 针对多面体的一切论证仍有适用于任意维数的代数写法, 代价是某些论证将显得异常复杂. 根植于图像又超越图像的这种严格性恰好是现实应用所需; 就以 §14.8 简介的线性规划问题为例, 除了中学课本以外, 变量个数 ≤ 3 的模型是不切实际的.

类似地, §14.9 探讨的多面锥是有限维实向量空间 V 中由有限多个线性不等式 $\lambda \geq 0$ 截出的区域, 其中 $\lambda : V \to \mathbb{R}$ 是不恒为零的线性映射. 示意如下.



多面锥同样出现于许多基础与应用领域. Fourier–Motzkin 定理 14.9.4 说明 V 中的多面锥等价于形如 $\sum_{i=1}^m \mathbb{R}_{\geq 0} x_i$ 的子集; 定理 14.9.11 将在 C 严格凸的前提下说明对于 $x_1,\dots,x_m$ 最经济的选法恰好是 C 的 “端射线” (定义 14.6.7).

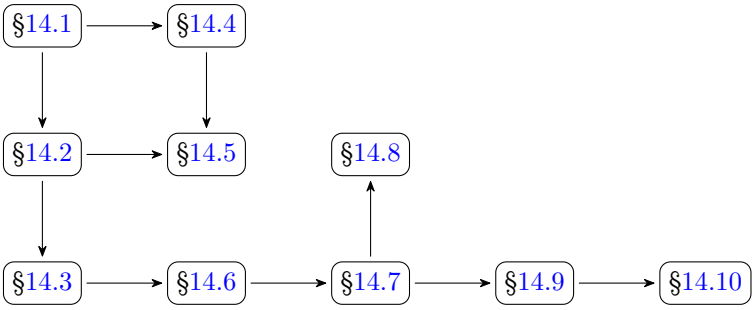
多面锥和多面体理论有许多相通之处. 有界多面体称为多胞体, 例如三维空间中的正多面体便是多胞体. 我们在 §14.10 将以多面锥理论为工具, 进一步明确多胞体的结构. Minkowski 的多胞体基本定理 14.10.4 断言有限维实仿射空间 E 中的多胞体无非是 E 中有限多个点的凸包; 尽管这一陈述在直观上几乎无可辩驳, 然而完整的证明需要不少迂回.

读者不应被多面体的初等外表所迷惑, 它们有极丰富的几何与组合学结构, 值得用一部专著来处理. 本章仅介绍其中最基本的部分.

阅读提示

本章内容不在书中其它部分用到, 请读者按需阅读; 习题部分另有许多延伸内容可供参考. 此外, 本章 §14.6 关于凸性的内容也常见于泛函分析或运筹学的教材, 不过该节目标是为后续内容作铺垫, 并非关于凸性的详尽介绍.

阅读顺序



14.1 仿射空间

向量的语言对初等几何学的研究是极有力的工具, 然而向量空间毕竟和直观的几何空间有异, 差别在于:

- ⋆ 与向量空间带有零元这一事实相反, 直观的空间并没有一个特定的基点; 尽管基点无妨任意指定, 但选法毕竟不自然.
- ⋆ 向量空间有加法, 但直观的几何空间中任两点 x 和 y 并不能相加; 具有意义的毋宁说是它们确定的向量 $\overrightarrow{xy}$.

以抽象的数学语言梳理, 这就引申出仿射空间的概念, 它是 “空间中两点确定一个向量” 这一思路的提纯.

定义 14.1.1 设 V 为域 F 上的向量空间. 所谓 V 作用下的**仿射空间**, 是指资料 $(E,+)$, 其中 E 是非空集, $+:E\times V\to E$ 是映射, 写作 $(x,v)\mapsto x+v$ 的形式. 它们服从于下

述条件:

⋆ 对任意 $x \in E$ 和 $v, w \in V$, 有结合律

$$(x + v) + w = x + (v + w);$$

⋆ 对任意 $x \in E$, 有 $x + 0 = x$;

⋆ 对任意 $x, y \in E$, 存在唯一的 $v \in V$ 使得 $y = x + v$, 此 v 也写作 $y - x$.

定义仿射空间 E 的维数为向量空间 V 的维数.

根据结合律, 可以无歧义地将 $(x + v) + w$ 写成 $x + v + w$. 此外, 不难从定义推导

$$\begin{aligned}(y + v) - (x + v) &= y - x,\\ z - x &= (y - x) + (z - y).\end{aligned}$$

不妨直观地将集合 E 理解为空间中的点, 而 $x + v$ 是以 x 为起点, 向 v 方向平移的结果. 以上定义的 $y - x$ 相当于初等几何学中考虑的向量 $\overrightarrow{xy}$.

例 14.1.2 取定向量空间 V, 取 $+$ 为 V 的加法, 使 V 成为它自身作用下的仿射空间, 这是仿射空间的平凡例子.

反过来说, 给定 V 作用下的仿射空间 $(E, +)$ 和 $o \in E$, 它们确定互逆的双射

$$\begin{aligned}V &\xleftrightarrow{1:1} E\\ v &\longmapsto o + v\\ x - o &\longmapsfrom x.\end{aligned}$$

此双射保持两边作为仿射空间的加法. 具体地说, 对于任意 $v \in V$ (看作仿射空间 V 的点) 和 $w \in V$ (看作向量空间 V 的元素),

$$v + w \mapsto o + (v + w) = (o + v) + w.$$

这就说明双射匹配两边的仿射空间结构. 这般选定的 $o \in E$ 称为 E 的一个基点. 一句话,

$$\text{仿射空间} + \text{基点} = \text{向量空间}.$$

练习 14.1.3 对于熟悉群作用的读者, 请说明 V 作用下的仿射空间相当于带有加法群 $(V, +)$ 作用的集合 E, 使得该作用自由且传递 (定义 11.5.6).

尽管仿射空间的元素不能像向量一般作线性组合, 但它们有称为仿射组合的基本操作.

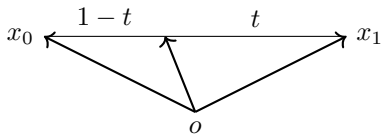
定义 14.1.4 (仿射组合) 设 E 是向量空间 V 作用下的仿射空间. 设 $n \in \mathbb{Z}_{\geq 1}$, 考虑 $x_0, \ldots, x_n \in E$ 和 $t_0, \ldots, t_n \in F$, 要求 $\sum_{i=0}^n t_i = 1$. 任选基点 o 以定义

$$\sum_{i=0}^n t_i x_i := o + \sum_{i=0}^n t_i (x_i - o) \ \in E.$$

上述定义无关 o 的选取: 若 $o' = o + v$, 则

$$\begin{aligned} o' + \sum_{i=0}^n t_i (x_i - o') &= o + \left(v + \sum_{i=0}^n t_i (x_i - o) - \sum_{i=0}^n t_i v \right) \\ &= o + \sum_{i=0}^n t_i (x_i - o). \end{aligned}$$

为了具体理解仿射组合的几何意涵, 不妨取 $F = \mathbb{R}$ 和 $n = 1$, 设 $t \in [0,1]$. 平面向量几何的知识说明 $tx_0 + (1-t)x_1$ 落在连接 x_0 与 x_1 的线段上, 与两端的距离比为 $1-t : t$, 图示如下.



如果容许 $t < 0$ (或 $t > 1$), 便能得到直线上位于 x_1 右侧 (或 x_0 左侧) 的点. 一般而言, 至少在 $F = \mathbb{R}$ 而 $t_i \geq 0$ 的具体情形, 不妨设想 x_i 为质量为 t_i 的质点, 则仿射组合 $\sum_{i=0}^n t_i x_i$ 便是系统的重心[1)], 它不依赖坐标系或原点的选取.

命题 14.1.5 *给定仿射空间 E 中的一族点 $(x_i)_{i \in I}$, 其中 I 是非空有限集, 具备无交并分解 $I = I_1 \sqcup \cdots \sqcup I_r$. 设 $(t_i)_{i \in I} \in F^I$ 满足 $\sum_{i \in I} t_i = 1$. 在 $s_j := \sum_{k \in I_j} t_k \neq 0$ 对所有 $1 \leq j \leq r$ 成立的前提下, 对 $t_i \in I_j$ 定义 $t_i' := \frac{t_i}{s_j}$, 则*

$$\sum_{i \in I} t_i x_i = \sum_{j=1}^r s_j \left(\sum_{i \in I_j} t_i' x_i \right),$$

右式视为二重的仿射组合.

证明 留意到 $\sum_{i \in I_j} t_i' = 1$ 而 $\sum_{j=1}^r s_j = 1$, 故二重仿射组合有意义. 选定基点后, 问题化为向量空间中平凡的等式 $\sum_i t_i v_i = \sum_j s_j \sum_{i \in I_j} \frac{t_i}{s_j} v_i$. □

练习 14.1.6 运用命题 14.1.5 来说明:

⋆ 有限多个点的仿射组合可逐步化约到两点的仿射组合,

[1)] 物理学上更正确的说法是质心.

⋆ 仿射组合的仿射组合可以 “摊开” 为仿射组合.

基于仿射组合的操作, 有限维仿射空间也有基的概念.

定义 14.1.7 (仿射基) 设 $(E,+)$ 是有限维向量空间 V 作用下的仿射空间, $n\in\mathbb{Z}_{\geq 0}$, 而 $x_0,\dots,x_n\in E$.

⋆ 若 $\sum_{i=0}^n t_ix_i=\sum_{i=0}^n t_i'x_i$ 当且仅当 $t_i=t_i'$ 对所有 $0\leq i\leq n$ 成立, 则称 $x_0,\dots,x_n$ 仿射无关, 否则称之为仿射相关.

⋆ 若所有 $x\in E$ 都能表为仿射组合 $x=\sum_{i=0}^n t_ix_i$, 则称 $x_0,\dots,x_n$ 生成 E.

⋆ 既仿射无关又生成 E 的 $x_0,\dots,x_n$ 称为**仿射基**; 若 $x=\sum_{i=0}^n t_ix_i\in E$, 则称 $(t_0,\dots,t_n)$ 为 x 的**重心坐标**.

命题 14.1.8 给定 $x_0,\dots,x_n$ 如上, 则它们仿射无关 (或生成 E) 当且仅当 $x_1-x_0,\dots,x_n-x_0$ 在 V 中线性无关 (或生成 V).

作为推论, 仿射基 $x_0,\dots,x_n$ 总是存在, 并且 $n=\dim V$ 总成立.

证明 在仿射组合的定义中取 $o:=x_0$, 记 $y_i:=x_i-x_0\in V$, 则 $x_0,\dots,x_n$ 仿射无关等价于以下陈述: $\sum_{i=1}^n t_iy_i=\sum_{i=1}^n t_i'y_i$ 当且仅当 $t_i=t_i'$ 对所有 $1\leq i\leq n$ 成立. 它们生成 E 等价于所有 $x_0+y\in E$ 皆能写成 $x_0+\sum_{i=1}^n t_iy_i$, 亦即 $y=\sum_{i=1}^n t_iy_i$. □

练习 14.1.9 说明仿射基等价于极大仿射无关子集, 也等价于极小仿射生成集.

继续介绍仿射空间的子结构.

定义 14.1.10 设 E 是向量空间 V 作用下的仿射空间, 而 V' 是 V 的子空间. 形如

$$E'=o+V':=\{o+v'\in E:v'\in V'\}$$

的非空子集称为 E 的**仿射子空间**, 其中 $o\in E$. 在 V 有限维的情形, 若 $\dim V'=\dim V-1$, 则称 E' 为**超平面**.

对于给定的 E', 定义中的 o 并非唯一, 任两者仅差一个 V' 的平移. 另一方面, V' 由 E' 唯一确定, 因为它由所有 $e_1'-e_2'$ 张成, 其中 $e_1',e_2'\in E'$; 于是可以合理地称 V' 为 E' 的**向量部分**.

例 14.1.11 考虑线性映射 $T:V\to W$ 和 $w\in\mathrm{im}(T)$; 任选 $v\in T^{-1}(w)$, 则 $T^{-1}(w)=v+\ker(T)$. 因此 $T^{-1}(w)$ 是 V 作为仿射空间的仿射子空间, 向量部分为 $\ker(T)$. 探讨线性方程组的通解时已见过这种结构. 稍后的命题 14.2.5 将给出更广的版本.

引理 14.1.12 设 E' 为仿射空间 E 的非空子集. 以下陈述等价:

(i) E' 是 E 的仿射子空间;

(ii) 对于所有 $n \in \mathbb{Z}_{\geq 1}$ 和 $x_0, \dots, x_n \in E'$, 这些点的所有仿射组合仍属于 E'.

证明 (i) $\Longrightarrow$ (ii) 是 $E' = o + V'$ 和仿射组合定义的简单结论. 以下解释 (ii) $\Longrightarrow$ (i). 命 V' 为所有 $e'_1 - e'_2$ 张成的子空间, 其中 $e'_1, e'_2 \in E'$. 任选 $o \in E'$. 以下说明 $E' = o + V'$.

先说明 $\subset$: 对所有 $x \in E'$ 皆有 $x - o \in V'$, 故 $x = o + (x - o) \in o + V'$.

其次说明 $\supset$. 设 $x = o + \sum_{i=1}^{m} t_i(x_i - y_i)$, 其中 $x_i, y_i \in E'$ 而 $t_i \in F$. 此时 x 是 $o, x_1, y_1, \dots, x_m, y_m$ 的仿射组合, 对应的系数依序取为 $1, t, -t, \dots, t, -t$, 故 $x \in E'$. 明所欲证. □

基于引理 14.1.12 可知对于 E 的任一族仿射子空间 $(E_i)_{i \in I}$, 其中 $I \neq \varnothing$, 若其交 $\bigcap_i E_i$ 非空, 则交是 E 的仿射子空间.

定义 14.1.13 设 S 为仿射空间 E 的非空子集, 定义其**仿射包**为

$$\operatorname{aff}(S) := \bigcap_{\substack{E' : E \text{ 的仿射子空间} \\ E' \supset S}} E'.$$

此交当然非空 (它包含 S), 因而是仿射子空间.

仿射包有以下更明确的描述.

练习 14.1.14 记 S 中任意有限多个点的所有仿射组合所成集合为 E_S. 说明 $E_S = \operatorname{aff}(S)$.

提示〉 练习 14.1.6 说明对 E_S 的元素取仿射组合不产生新的点; 代入引理 14.1.12.

练习 14.1.15 设 A 和 B 分别是 n 维仿射空间 E 的 p 维和 q 维仿射子空间. 说明如果 $A \cap B \neq \varnothing$, 则 $\dim(A \cap B) \geq p + q - n$.

提示〉 任取 $o \in A \cap B$, 将问题化为向量空间的版本.

14.2 仿射线性映射

接着来介绍保持仿射空间结构的一类映射. 继续选定域 F.

定义–命题 14.2.1 设 $(E_i, +)$ 为 V_i 作用下的仿射空间 $(i = 1, 2)$, 而 $A : E_1 \to E_2$ 为映射. 如果对所有 $o_1 \in E_1$ 都存在 $o_2 \in E_2$ 和 $T \in \operatorname{Hom}(V_1, V_2)$, 使得

$$A(o_1 + v_1) = o_2 + Tv_1$$

对所有 $v_1 \in V_1$ 成立, 则称 A 为**仿射线性映射**. 事实上,

- ⋆ 当 o_1 选定, 则 o_2 和 T 是唯一确定的;
- ⋆ 只要所述恒等式对某个 $o_1 \in E_1$ 成立, 则自动对所有 o_1 成立;

⋆ 不同的 o_1 总对应同一个 $T \in \mathrm{Hom}(V_1, V_2)$, 它可以合理地称为仿射线性映射 A 的**向量部分**.

证明 待验证的是列出的三条性质. 选定的 o_1, 代入 $v_1 = 0$ 可见 $o_2 = A(o_1)$, 而 $Tv_1 = o_2 - A(o_1 + v_1)$. 故 o_2 和 T 由 o_1 唯一确定.

接着考虑任意 $o_1' \in E_1$. 存在 $u \in V_1$ 使得 $o_1' = o_1 + u$. 对所有 v_1 皆有

$$A(o_1' + v_1) = A(o_1 + (u + v_1)) = o_2 + (Tu + Tv_1) = (o_2 + Tu) + Tv_1.$$

这表明所示条件对 o_1' 亦成立, 而且在 $\mathrm{Hom}(V_1, V_2)$ 中对应的元素仍是 T, 对应的 o_2' 则是 $o_2 + Tu$. □

仿射线性映射的合成依然是仿射线性的. 对此的一种视角是通过选取基点, 将一切化约到向量空间上讨论, 细说如下.

例 14.2.2 对 $i = 1, 2$ 取 E_i 为向量空间 V_i 对应的仿射空间 (例 14.1.2), 再取 o_1 为向量空间 V_1 的零元, 则仿射线性映射 $E_1 \to E_2$ 的描述化为以下形式:

$$\begin{aligned} \Phi_{T,v_2} : V_1 &\to V_2 \\ v_1 &\mapsto Tv_1 + v_2, \quad T \in \mathrm{Hom}(V_1, V_2), \quad v_2 \in V_2. \end{aligned}$$

观察到 $\Phi_{T,v_2} = \Phi_{\mathrm{id},v_2}\Phi_{T,0}$, 所以向量空间之间的仿射线性映射无非是一个线性映射 (形如 $\Phi_{T,0}$) 配上平移 (形如 Φ_{id,v_2}).

对于一列线性映射 $V_1 \xrightarrow{T} V_2 \xrightarrow{S} V_3$ 和 $v_2 \in V_2$, $v_3 \in V_3$, 仿射线性映射的合成规律是

$$\Phi_{S,v_3}\Phi_{T,v_2} = \Phi_{ST,Sv_2+v_3}; \tag{14.2.1}$$

诚然, 左侧先映 v_1 为 $Tv_1 + v_2$, 再映为 $STv_1 + Sv_2 + v_3$.

练习 14.2.3 沿用上述符号, 并且设 $T : V_1 \to V_2$ 为向量空间的同构. 验证

$$\begin{aligned} \Phi_{T,0}\Phi_{\mathrm{id},v_1}\Phi_{T,0}^{-1} &= \Phi_{\mathrm{id},Tv_1}, \\ \Phi_{T,v_2}^{-1} &= \Phi_{T^{-1},-T^{-1}v_2}. \end{aligned}$$

以下结果提供看待仿射线性映射的另一种等价视角: 它们是保持仿射组合 (定义 14.1.4) 的映射.

命题 14.2.4 设 E_1 和 E_2 分别是 V_1 和 V_2 作用下的仿射空间. 映射 $A : E_1 \to E_2$ 是仿射线性的当且仅当它具有如下性质: 对所有 $n \in \mathbb{Z}_{\geq 1}$ 和 $x_0, \ldots, x_n \in E_1$, $t_0, \ldots, t_n \in F$, 在 $\sum_{i=0}^n t_i = 1$ 的前提下皆有

$$A\left(\sum_{i=0}^n t_i x_i\right) = \sum_{i=0}^n t_i A(x_i).$$

证明 先看 “仅当” 方向. 选定基点将 A 表作 $A(o_1+v_1)=o_2+Tv_1$ 之形, 则

$$\begin{aligned} A\left(\sum_{i=0}^{n} t_i x_i\right) &= A\left(o_1+\sum_{i=0}^{n} t_i(x_i-o_1)\right) \\ &= o_2+\sum_{i=0}^{n} t_i T\left(x_i-o_1\right) \\ &= o_2+\sum_{i=0}^{n} t_i\left(A(x_i)-o_2\right)=\sum_{i=0}^{n} t_i A(x_i). \end{aligned}$$

对于 “当” 的方向, 选定 $o_1\in E_1$, 命 $o_2:=A(o_1)$, 须证 $T: v_1\mapsto A(o_1+v_1)-o_2$ 属于 $\mathrm{Hom}(V_1,V_2)$. 为了证明 T 保加法, 将 $o_1+(v_1+v_1')$ 表成仿射组合

$$t\left(o_1+v_1\right)+t'\left(o_1+v_1'\right)+t''o_1, \quad (t,t',t'):=(1,1,-1).$$

它对 A 的像是

$$t\left(o_2+Tv_1\right)+t'\left(o_2+Tv_1'\right)+t''o_2=o_2+(Tv_1+Tv_1').$$

故 $T(v_1+v_1')=Tv_1+Tv_1'$.

为了证明 T 保纯量乘法, 给定 $t\in F$, 将 o_1+tv 表成仿射组合 $t(o_1+v)+(1-t)o_1$, 然后类似地论证. □

命题 14.2.5 设 $A: E_1\to E_2$ 为仿射线性映射, 其向量部分为 $T: V_1\to V_2$.

(i) 若 E_1' 是 E_1 的仿射子空间, 向量部分为 V_1', 则 $A(E_1')$ 是 E_2 的仿射子空间, 向量部分为 $T(V_1')$.

(ii) 若 E_2' 是 E_2 的仿射子空间, 向量部分为 V_2', 则 $A^{-1}(E_2')$ 或者空, 或者是 E_1 的仿射子空间, 以 $T^{-1}(V_2')$ 为向量部分.

证明 两者都是仿射子空间定义的简单结论. 对于 (i), 任取 $o_1\in E_1'$, 命 $o_2:=A(o_1)$, 则 $A(E_1')=o_2+T(V_1')$. 对于 (ii), 设 $A^{-1}(E_2')$ 非空, 选取其元素 o_1 并且命 $o_2:=A(o_1)$, 则从 $E_2'=o_2+V_2'$ 可得

$$A(o_1+v_1)=o_2+T(v_1)\in E_2' \iff T(v_1)\in V_2',$$

换言之 $A^{-1}(E_2')=o_1+T^{-1}(V_2')$. 证毕. □

定义–命题 14.2.6 设 E_i 为 V_i 作用下的仿射空间 $(i=1,2)$. 对于仿射线性映射 $A: E_1\to E_2$, 以下性质成立.

(i) 若 A 是双射, 则它的逆映射仍是仿射线性的, 并且 A 和 A^{-1} 的向量部分互逆.

(ii) 反过来说, 若 A 的向量部分可逆, 则 A 是双射.

满足以上任一条件的 A 称为仿射空间的同构.

上述性质表明仿射空间 E 的所有自同构对映射的合成成群, 记为 $\mathrm{Aff}(E)$, 也称为 E 上的**仿射变换群**.

证明 以下选定基点, 将 A 表作 $A(o_1+v_1)=o_2+Tv_1$ 之形.

考虑 (i). 设 A 为双射, 则从 A 单立见 T 单. 另一方面, 对所有 $v_2\in V_2$, 存在 $v_1\in V_1$ 使得 $A(o_1+v_1)=o_2+v_2$, 然而这也等价于 $Tv_1=v_2$. 综上, T 为双射.

此外, 给定 $o_2+v_2\in E_2$, 则有 $A\left(o_1+T^{-1}(v_2)\right)=o_2+v_2$, 这就说明 A^{-1} 的映法是 $o_2+v_2\mapsto o_1+T^{-1}(v_2)$, 这是向量部分为 T^{-1} 的仿射线性映射.

考虑 (ii). 设 T 可逆, 定义仿射线性映射 $o_2+v_2\mapsto o_1+T^{-1}(v_2)$, 易见它和 A 互逆. □

下述练习将明确 $\mathrm{Aff}(E)$ 的群结构.

练习 14.2.7 对所有 $v\in V$, 对应的平移 $R_v\in\mathrm{Aff}(E)$ 映 x 为 $x+v$. 单同态 $v\mapsto R_v$ 将加法群 $(V,+)$ 嵌为 $\mathrm{Aff}(E)$ 的子群.

(i) 验证 $V\lhd\mathrm{Aff}(E)$.

(ii) 选定 $o\in E$, 对所有线性自同构 $T\in\mathrm{GL}(V)$ 可以定义 $\Phi_{T,o}\in\mathrm{Aff}(E)$ 使得 $\Phi_{T,o}(o+w)=o+Tw$. 验证 $T\mapsto\Phi_{T,o}$ 将群 $\mathrm{GL}(V)$ 嵌入 $\mathrm{Aff}(E)$.

(iii) 承上, 证明半直积分解 $\mathrm{Aff}(E)=V\rtimes\mathrm{GL}(V)$ (定义–命题 11.10.3), 并明确半直积所涉及的群同态 $\mathrm{GL}(V)\to\mathrm{Aut}(V)$.

14.3 刚体运动

本节取 $F=\mathbb{R}$, 并且设 $(V,(\cdot|\cdot))$ 为实内积空间. 对于 V 作用下的仿射空间 E, 从 V 的内积空间结构得到 E 上的距离函数

$$\begin{aligned} d:E\times E&\to\mathbb{R}_{\geq 0}\\ (x,y)&\mapsto\|y-x\|. \end{aligned}$$

它使得 (E,d) 成为分析学中所谓的度量空间. 换言之, d 具备下列性质:

$$\begin{gathered} d(x,y)=d(y,x),\\ d(x,y)=0\iff x=y,\\ d(x,y)+d(y,z)\geq d(x,z)\quad(\text{三角不等式}). \end{gathered}$$

第一则性质缘于 $\|v\|=\|-v\|$. 第二则性质缘于内积的正定性. 至于最后一则性质, 由于 $z-x=(y-x)+(z-y)$, 它直接来自内积空间的三角不等式. 当 E 来自于向量空间 V 时, 这种度量已经在内积空间的相关章节中讨论过了.

定义 14.3.1 设 $(E,+)$ 是内积空间 V 作用下的仿射空间. 映射 $R:E\to E$ 若对所有 $x,y\in E$ 皆满足 $d(R(y),R(x))=d(y,x)$, 则称 R 为 E 上的**刚体运动**.

刚体运动之所以 "刚", 在于它不改变两点的距离. 刚体运动的合成还是刚体运动. 且来打量一些具体例子.

▷ **平移** 设 $v\in V$, 则 $R_v:x\mapsto x+v$ 是刚体运动. 对任意 $v,w\in V$, 显然有 $R_vR_w=R_{v+w}$, 而且 $R_0=\mathrm{id}$. 作为推论, $R_v^{-1}=R_{-v}$.

▷ **保距线性变换** 设 $T\in\mathrm{End}(V)$ 保距, $o\in E$ 是选定的基点. 命

$$\begin{aligned}R_{T,o}:E&\to E\\ o+v&\mapsto o+Tv.\end{aligned}$$

它为何是刚体运动? 按定义作计算: 设 $x=o+v$, $y=o+w$, 则由 T 保距可见

$$\begin{aligned}d\left(R_{T,o}(y),R_{T,o}(x)\right)&=d\left(o+Tw,o+Tv\right)=\|T(w-v)\|\\&=\|w-v\|=d(y,x).\end{aligned}$$

定义即刻给出 $R_{S,o}R_{T,o}=R_{ST,o}$ 和 $R_{\mathrm{id},o}=\mathrm{id}_E$; 当 T 可逆时, 按此得到 $R_{T,o}^{-1}=R_{T^{-1},o}$.

留意到 $R_{T,o}$ 和 R_v 都是仿射线性映射, 而且不难验证当 $o'=o+u$ 时

$$R_{T,o'}=R_{u-Tu}R_{T,o}.$$

如果进一步以基点 o 将 E 等同于 V, 则 $R_{T,o}$ (或 R_v) 正是先前定义过的仿射线性映射 $\Phi_{T,0}$ (或 $\Phi_{\mathrm{id},v}$).

现在着手来刻画所有刚体运动. 第一步是说明刚体运动保持内积.

引理 14.3.2 设 $R:E\to E$ 为刚体运动, 则对于任意 $x,y,y'\in E$ 皆有

$$(R(y)-R(x)\mid R(y')-R(x))=(y-x\mid y'-x)\,.$$

证明 基本工具是对任何内积空间皆成立的恒等式

$$(v|w)=\frac{\|v-w\|^2-\|v\|^2-\|w\|^2}{-2}.$$

命 $v:=R(y)-R(x)$ 和 $w:=R(y')-R(x)$, 则 $v-w=R(y)-R(y')$; 又因为 R 是刚体运动, 遂有

$$\begin{aligned}(v|w)=\frac{\|y-y'\|^2-\|y-x\|^2-\|y'-x\|^2}{-2}\\ =\frac{\|(y-x)-(y'-x)\|^2-\|y-x\|^2-\|y'-x\|^2}{-2},\end{aligned}$$

而这又等于 $(y-x|y'-x)$. 证毕. □

定理 14.3.3 设 $(E,+)$ 为内积空间 V 作用下的仿射空间. 选定 $o\in E$. 设 $R:E\to E$ 为刚体运动, 则存在唯一的 $v\in V$ 和保距线性映射 $T:V\to V$ 使得

$$R=R_vR_{T,o}.$$

特别地, R 是仿射线性映射.

当 E 维数有限时, R 总是可逆, 而且 R^{-1} 仍是刚体运动; 更精确地说, 此时 $(R_vR_{T,o})^{-1}=R_{T^{-1},o}R_{-v}$.

证明 首先由于 $R_vR_{T,o}$ 是映 $o+u$ 为 $(o+v)+Tu$ 的仿射线性映射, T 和 v 的唯一性已包含于定义–命题 14.2.1.

关键在于说明所有 R 都有此表法. 取 $v=R(o)-o$, 并以 $R_{-v}R$ 代 R, 容易化约到 $R(o)=o$ 的情形, 目标是证存在保距的 $T\in\mathrm{End}(V)$ 使得 $R=R_{T,o}$. 首先, 按 $o+T(v)=R(o+v)$ 定义映射 $T:V\to V$, 它满足 $T(0)=0$.

⋆ 兹断言 T 保持内积:

$$(T(v)|T(v'))=(v|v'),\quad v,v'\in V.$$

诚然, 这是在引理 14.3.2 中代入 $y=o+v$, $y'=o+v'$, $x=o$ 的直接结论.

⋆ 其次 T 保持纯量乘法, 亦即 $T(tv)=tT(v)$ 对所有 $v\in V$ 和 $t\in\mathbb{R}$ 成立. 为此, 请端详

$$\begin{aligned}&(T(tv)-tT(v)\mid T(tv)-tT(v))=\\&\quad(T(tv)\mid T(tv))-t\,(T(tv)\mid T(v))-t\,(T(v)\mid T(tv))+t^2\,(T(v)\mid T(v))\\&\qquad=(tv|tv)-t\,(tv|v)-t\,(v|tv)+t^2\,(v|v)=0;\end{aligned}$$

此处用到 T 保内积. 内积正定遂导致 $T(tv)-tT(v)=0$.

⋆ 最后, $T(v_1+v_2)=Tv_1+Tv_2$ 对所有 $v_1,v_2\in V$ 成立. 方法类似, 考虑

$$\begin{aligned}&(T(v_1+v_2)-Tv_1-Tv_2\mid T(v_1+v_2)-Tv_1-Tv_2)=\\&\qquad(T(v_1+v_2)\mid T(v_1+v_2))+(Tv_1\mid Tv_1)+(Tv_2\mid Tv_2)\\&\qquad\quad-2\,(T(v_1+v_2)|\,Tv_1)-2\,(T(v_1+v_2)|\,Tv_2)+2\,(Tv_1\mid Tv_2)\end{aligned}$$

由于 T 保内积, 这等于

$$\begin{aligned}&(v_1+v_2\mid v_1+v_2)+(v_1|v_1)+(v_2|v_2)\\&\qquad-2\,(v_1+v_2\mid v_1)-2\,(v_1+v_2\mid v_2)+2\,(v_1\mid v_2)\\&\qquad\qquad=((v_1+v_2)-v_1-v_2\mid(v_1+v_2)-v_1-v_2)=0.\end{aligned}$$

因此 $T(v_1+v_2)=Tv_1+Tv_2$.

综上, $T : V \to V$ 是保距线性映射.

对于最后一部分, 考虑刚体运动 $R = R_v R_{T,o}$. 回忆到当 V 有限维时, T 自动是内积空间的同构, 亦即正交变换, 于是写法 $R_{T^{-1},o} R_{-v}$ 有意义, 而且容易看出它是 $R_v R_{T,o}$ 的逆. □

推论 14.3.4 *设 $(E, +)$ 为有限维内积空间 V 作用下的仿射空间, 则所有刚体运动 $R : E \to E$ 对映射的合成构成群.*

证明 刚体运动的合成仍是刚体运动, 恒等映射 id_E 是刚体运动, 故所需验证的仅是逆元的存在性, 然而这正是定理 14.3.3 的内容. □

一类有趣的刚体运动来自镜射.

定义–命题 14.3.5 *设 $(E, +)$ 为有限维内积空间 V 作用下的仿射空间. 给定 E 的仿射子空间 E_0, 其向量部分记为 V_0; 记从 V 向 V_0 的正交投影算子为 P.*

任选 $o \in E_0$, 定义映射

$$\begin{aligned} \rho_{E_0} : E &\to E \\ o + v &\mapsto o + (2Pv - v). \end{aligned}$$

*此定义无关 $o \in E_0$ 的选取, 并且给出 E 上的刚体运动, 称为对 E_0 的***镜射***.*

证明 对所有 $o' \in E_0$, 取 $v' \in V_0$ 使得 $o' = o + v'$, 则 $o + v = o' + (v - v')$, 而

$$\begin{aligned} o' + (2P(v - v') - (v - v')) &= o' + (2Pv - v - v') \\ &= o + (2Pv - v), \end{aligned}$$

这里用到了正交投影的性质 $Pv' = v'$. 因此 ρ_{E_0} 不依赖 o 的选取.

为了说明 ρ_{E_0} 是刚体运动, 留意到 $2P - \mathrm{id}_V \in \mathrm{End}(V)$ 是正交变换, 而 $\rho_{E_0} = R_{o,2P-\mathrm{id}}$. □

留意到 $\rho_{E_0}|_{E_0} = \mathrm{id}_{E_0}$.

练习 14.3.6 设 F 为域, E 为有限维 F-向量空间 V 作用下的仿射空间. 证明子集 $H \subset E$ 是 E 中的超平面 (定义 14.1.10) 当且仅当存在非常值仿射线性映射 $\alpha : E \to F$ 使得

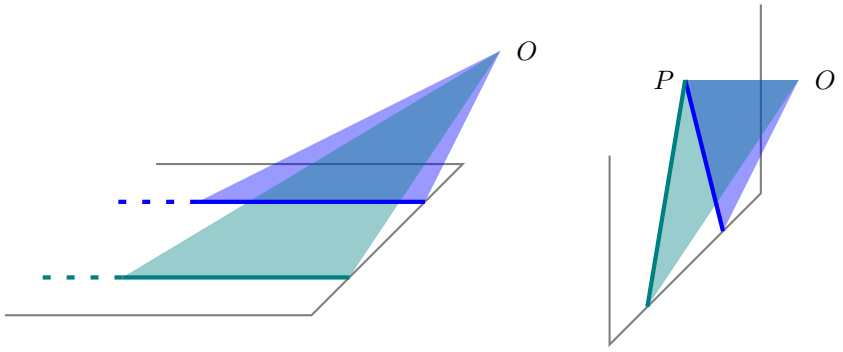
$$H = \{x \in E : \alpha(x) = 0\};$$

对于 $F = \mathbb{R}$ 而 V 为内积空间的情形, 选定与 H 的向量部分正交的 $\nu \in V \smallsetminus \{0\}$, 亦即 H 的法向量. 说明只要适当伸缩 ν 或 α, 总有 $\rho_H(x) = x - 2\alpha(x)\nu$; 试明确所需的条件.

对超平面的镜射足以在合成之下生成所有刚体变换. 玩过万花筒的读者应能体会镜射的合成能有多复杂, 本章习题将有进一步的探究.

14.4 射影空间

射影几何可以追溯到文艺复兴时期对透视法绘画的研究, 其抽象理论则始于 J.-V. Poncelet, K. von Staudt 和 J. Steiner 在 19 世纪前半叶的工作. 考虑地面上两条向远处无穷延伸的平行线, 如下图左侧所示, 画家的眼睛位置标为 O; 右侧竖立的平面代表画家眼前的画布或取景框, 大地上两条平行线在画布上投射的图像仍是两条直线, 但它们不再平行而是有唯一交点 P. 连线 OP 平行于地面, 故 P 不再对应到大地上的任何一点, 但在画家目中可设想为这对平行线在无穷远处共同的 “消没点”.



平行线交于无穷远处在视觉经验中习以为常. 通过适当地设置消没点, 透视法能在绘画中逼真地展现物体远近. 就画家的立场, 无论左图的大地或右图的立面都是某种画布, 起关键作用的毋宁说是目光, 亦即过 O 点的直线. 画布可以是空间中任何一个不含 O 的平面, 而目光与画布的唯一交点便呈现为图画. 每幅画布都有不可及的无穷远点: 以先前的讨论为例, 大地上画不出平行线的交点, 而竖直的画布也无法描绘画家脚下.

这就启发以下定义: 将 O 等同于 $\mathbb{R}^3$ 的零点, 实射影平面 $\mathbb{P}^2$ 定义为 $\mathbb{R}^3$ 中全体过零点的直线, 亦即 $\mathbb{R}^3$ 的所有 1 维子空间所成集合. 于是之前所谓的大地便对应到 $\mathbb{P}^2$ 中所有不与 XY 平面平行的直线, 这般的直线有唯一的形如 $(x, y, 1)$ 的生成元.

以上表述只涉及向量空间的操作, 在代数学中具有更广泛的形式. 今后选定域 F.

定义 14.4.1 (射影空间及其维数) 设 V 为非零 F-向量空间, 相应的射影空间定义为

$$\mathbb{P}(V) := \{\ell \subset V : \text{1 维子空间}\};$$

这也可以视同 $V \smallsetminus \{0\}$ 对等价关系

$$v \sim v' \iff \exists t \in F^\times,\ v' = tv$$

的商集, 方法是让 v 的等价类对应到 $\ell = Fv$. 如果 $\dim V = n+1$, 其中 $n \in \mathbb{Z}_{\geq 0}$, 则称 $\mathbb{P}(V)$ 为 n 维射影空间.

对于 $V=F^{n+1}$ 的特例, 我们记[2)]

$$\mathbb{P}^n := \mathbb{P}(F^{n+1}).$$

非零向量空间之间的任何线性单射 $\varphi: V_0 \to V$ 皆诱导嵌入 $\mathbb{P}(V_0) \hookrightarrow \mathbb{P}(V)$, 它映 1 维子空间 ℓ 为 $\varphi(\ell)$. 这就引出射影子空间的概念.

定义 14.4.2 (射影子空间) 在 $\mathbb{P}(V)$ 中形如 $\mathbb{P}(V_0)$ 的子集称为 $\mathbb{P}(V)$ 的射影子空间, 其中 V_0 是 V 的非零子空间.

- ⋆ 0 维射影子空间无非是 $\mathbb{P}(V)$ 的元素, 简称为 $\mathbb{P}(V)$ 中的点,
- ⋆ 1 维射影子空间简称为 $\mathbb{P}(V)$ 中的线,
- ⋆ 2 维射影子空间简称为 $\mathbb{P}(V)$ 中的面.

非零 F-向量空间 V 的某一类仿射子空间也能自然地嵌入 $\mathbb{P}(V)$, 方式如下. 设 E 为 V 中的仿射子空间, 要求 $0 \notin E$. 此时有嵌入

$$\begin{aligned} E &\hookrightarrow \mathbb{P}(V) \\ x &\mapsto Fx. \end{aligned} \tag{14.4.1}$$

上式之所以单, 是因为若 $x, y \in E$ 相异而且在 V 中成比例, 则 $y-x$ 同样和 x 成比例, 并且属于 E 的向量部分, 由此易推得 $0 \in E$, 矛盾.

练习 14.4.3 以后的情景中, (14.4.1) 涉及的 E 都形如 $x+V_0$, 要求 V 有限维而子空间 V_0 满足 $\dim V_0 = \dim V - 1$. 证明此时 (14.4.1) 的像是 $\mathbb{P}(V) \smallsetminus \mathbb{P}(V_0)$.

目光转向更具体的 $\mathbb{P}^n$.

定义 14.4.4 (齐次坐标) 记 $(x_0, \dots, x_n) \in F^{n+1} \smallsetminus \{\mathbf{0}\}$ 在 $\mathbb{P}^n$ 中确定的点为

$$(x_0 : \cdots : x_n).$$

因此 $\mathbb{P}^n$ 的所有点 x 都能表为 $(x_0 : \cdots : x_n)$ 的形式, 其中 $(x_0, \dots, x_n) \in F^{n+1} \smallsetminus \{\mathbf{0}\}$ 精确到用 $F^\times$ 伸缩由 x 唯一确定, 称之为点 x 的齐次坐标.

最简单的射影空间是独点集 $\mathbb{P}^0$. 另一方面 $\mathbb{P}^1 = \{(1:x) : x \in F\} \sqcup \{(0:1)\}$, 亦即仿射直线和一个无穷远点 $\infty := (0:1)$ 的无交并. 当 $F=\mathbb{R}$ 时, $\mathbb{P}^2$ 即是本节开头介绍的实射影平面.

现在以 $X_0, \dots, X_n$ 代表 F^{n+1} 的标准坐标. 对每个 $0 \le i \le n$ 取 $X_i = 1$ 在 F^{n+1} 中截出的 n 维仿射子空间, 它显然地同构于 F^n; 记其在 (14.4.1) 之下对应的像为

$$\begin{aligned} U_i &= \{(x_0 : \cdots : x_n) \in \mathbb{P}^n : x_i = 1\} \\ &= \{(x_0 : \cdots : x_n) \in \mathbb{P}^n : x_i \neq 0\}. \end{aligned}$$

[2)] 符号 $\mathbb{P}^n(F)$ 也是合理的. 另有一些教材记此为 $F\mathrm{P}^n$.

任何 $x \in \mathbb{P}^n$ 的齐次坐标总有非零分量, 故

$$\mathbb{P}^n = U_0 \cup \cdots \cup U_n,$$

这表明 n 维射影空间总能写成 $n+1$ 个 n 维仿射空间的并; 本章习题将说明 $n+1$ 是最小取法.

另一种观照 $\mathbb{P}^n$ 的方法是在 $n \geq 1$ 时注意到

$$\mathbb{P}^n \smallsetminus \underbrace{U_0}_{\simeq F^n} = \{(0 : x_1 : \cdots : x_n) : (x_i)_{i=1}^n \in F^n \smallsetminus \{\mathbf{0}\}\} \simeq \mathbb{P}^{n-1};$$

这表明 $\mathbb{P}^n$ 在 U_0 (对应到本节开头的 "画布") 之外的点构成低一维的射影空间. 当 $n \geq 2$ 时继续处理 $\mathbb{P}^{n-1}$, 最终能将 $\mathbb{P}^n$ 分解为仿射空间的无交并:

$$\mathbb{P}^n \simeq F^n \sqcup F^{n-1} \sqcup \cdots \sqcup F^0.$$

泛泛而论, 射影几何的宗旨是研究射影空间中的点与线, 乃至于各种射影子空间之间的交集和包含关系. 平面与空间上的射影几何学已有大量经典结论, 对圆锥曲线的研究也大有裨益, 然而本章只介绍射影几何中直接关乎代数学的部分.

注记 14.4.5 射影子空间在齐次坐标下总能写作线性方程组 $\sum_{j=0}^n a_{ij} X_j = 0$ 的非 $\mathbf{0}$ 解集在 $\mathbb{P}^n$ 中的像 (其中 $1 \leq i \leq m$), 因为 F^{n+1} 的子空间可按此描述. 推而广之, 考虑一族齐次多项式 $P_1, \ldots, P_m \in F[X_0, \ldots, X_n]$; 齐次性质确保 $P_1 = \cdots = P_m = 0$ 的非 $\mathbf{0}$ 解集对 $F^\times$ 的伸缩不变, 从而在齐次坐标下给出 $\mathbb{P}^n$ 的子集. 粗略地说, 按此描述的子集称为**射影代数簇**或简称**射影簇**, 其深入研究是代数几何学的主题.

有限维射影空间虽然包含同维数的仿射空间作为子集 (譬如先前的 $\mathbb{P}^n \supset U_i \simeq F^n$), 但一些几何性质在射影空间中更便于操作. 以下是关于相交的一则例子.

命题 14.4.6 *在 n 维射影空间 $\mathbb{P}^n$ 中, 任何 a 维和 b 维子射影空间的交是至少 $a+b-n$ 维的子射影空间.*

证明 设所论的射影空间为 V, 子射影空间对应到 $V_0, V_1 \subset V$, 其中 $\dim V = n+1$, $\dim V_0 = a+1$ 而 $\dim V_1 = b+1$. 向量空间的常识表明

$$\dim V_0 \cap V_1 = \dim V_0 + \dim V_1 - \dim(V_0 + V_1) \geq a + b - n + 1;$$

然而 $\mathbb{P}(V_0) \cap \mathbb{P}(V_1) = \mathbb{P}(V_0 \cap V_1)$. □

作为特例, 射影平面上任两条线或者相等, 或者交于唯一一点. 与此相反, 仿射平面的两条线可以无交.

射影几何学的另一个特色是对偶性. 以射影平面为例, 在任何一个关于点, 线及其包含关系的陈述中, 如果将 "点" 和 "线" 二字全部对调, 并且将包含关系倒转, 则新陈述和原陈述相等价. 为了从代数视角理解, 关键是在对偶空间确定的射影空间中考量新陈述, 记 F-向量空间 V 的对偶空间为 $V^\vee$.

命题 14.4.7 (射影空间的对偶原理) 设 V 为 $n+1$ 维 F-向量空间, $n \in \mathbb{Z}_{\geq 0}$. 对所有 $0 \leq m < n$, 我们有以下双射[3)]

$$\begin{aligned} \{\mathbb{P}(V) \text{ 的 } m \text{ 维射影子空间}\} &\xrightarrow{1:1} \{\mathbb{P}(V^\vee) \text{ 的 } n-m-1 \text{ 维射影子空间}\} \\ \mathbb{P}(V_0) &\longmapsto \mathbb{P}(V_0^\perp) \end{aligned}$$

其中 $V_0^\perp := \{\lambda \in V^\vee : \lambda|_{V_0} = 0\}$, 它具有反序性质: $\mathbb{P}(V_0) \subset \mathbb{P}(V_1)$ 当且仅当 $\mathbb{P}(V_0^\perp) \supset \mathbb{P}(V_1^\perp)$.

证明 考虑满足 $\dim V_0 = m+1$ 的子空间 $V_0 \subset V$. 对偶空间的一般性质说明 $\dim V_0^\perp = (n+1) - (m+1) = (n-m-1)+1$, 特别地 $V_0^\perp \neq \{0\}$, 故映射是良定义的. 对偶空间的常识还说明 V 自然地等同于 $V^{\vee\vee}$, 相应地 V_0 等同于 $(V_0^\perp)^\perp$, 故映射单. 同一套常识说明所有子空间 $W_0 \subset V^\perp$ 都等同于 $(W_0^\perp)^\perp$, 其中 $W_0^\perp \subset V$ 定义为 $\{v \in V : \forall \lambda \in W_0,\ \lambda(v) = 0\}$, 故映射满.

显然有 $V_0 \subset V_1 \implies V_0^\perp \supset V_1^\perp$. 反之若 $V_0^\perp \supset V_1^\perp$, 则同理有 $V^{\vee\vee}$ 中的包含关系 $(V_0^\perp)^\perp \subset (V_1^\perp)^\perp$, 亦即 V 中的 $V_0 \subset V_1$. 明所欲证. □

对于射影平面, 亦即 $n=2$ 的情形, 双射确实将 $\mathbb{P}(V)$ 中的线 (或点) 映至 $\mathbb{P}(V^\vee)$ 中的点 (或线), 而且反转线和点的包含关系.

14.5 射影变换与交比

选定域 F 和有限维非零 F-向量空间 V, 由此得到一般线性群 $\mathrm{GL}(V)$ (例 11.1.10).

定义 14.5.1 对应于 V 的**射影线性群**定义为商群

$$\mathrm{PGL}(V) := \mathrm{GL}(V) \big/ F^\times \mathrm{id}_V.$$

对于 $V = F^n$ 的情形, 记 $\mathrm{PGL}(n, F) := \mathrm{PGL}(F^n)$.

每个 $T \in \mathrm{GL}(V)$ 都诱导 $\mathbb{P}(V)$ 到自身的双射, 记为 $[T]$; 这般双射称为 $\mathbb{P}(V)$ 上的**射影变换**. 对所有 $c \in F^\times$, 易见 $[cT] = [T]$. 因此 $[T]$ 仅依赖 T 在 $\mathrm{PGL}(V)$ 中的像.

命题 14.5.2 上述构造给出双射

$$\mathrm{PGL}(V) \xrightarrow{1:1} \{\mathbb{P}(V) \text{ 上的射影变换}\}, \quad T\text{的陪集} \mapsto [T],$$

而且有 $[ST] = [S][T]$ 和 $[\mathrm{id}_V] = \mathrm{id}_{\mathbb{P}(V)}$. 作为推论, 全体射影变换对映射合成成群, 同构于 $\mathrm{PGL}(V)$.

[3)] 一些文献将 $\mathbb{P}(V)$ 的点定义为 V 的 $\dim V - 1$ 维子空间, 而非 1 维子空间. 此双射说明两种定义方式仅差一个对偶.

证明 此前的讨论说明 $T \to [T]$ 确实给出映射 $\mathrm{PGL}(V) \to \{\mathbb{P}(V)$ 上的射影变换$\}$, 而且射影变换的定义说明这是满射. 性质 $[ST] = [S][T]$ 和 $[\mathrm{id}_V] = \mathrm{id}_{\mathbb{P}(V)}$ 明白, 由此推得 $[T^{-1}] = [T]^{-1}$.

为了证明映射为单, 设 $[R] = [S]$, 则 $[RS^{-1}] = \mathrm{id}_{\mathbb{P}(V)}$. 命 $T = RS^{-1}$ 则 $[T] = \mathrm{id}_{\mathbb{P}(V)}$, 每个 $v \in V \smallsetminus \{0\}$ 都是 T 的特征向量; 根据特征值与特征向量的常识, 这蕴涵 $T = \lambda \cdot \mathrm{id}_V$, 其中 λ 是所有 $v \neq 0$ 共有的特征值. □

群 $\mathrm{PGL}(V)$ 以射影变换将射影子空间映为同维度的射影子空间, 保持其间的包含关系与相交关系. 射影变换群是射影几何中的 "对称性", 在射影变换下不变的性质正是射影几何学所关心的内容.

例 14.5.3 (仿射变换作为射影变换) 将仿射空间 F^n 以 $(x_1, \dots, x_n) \mapsto (1 : x_1 : \cdots : x_n)$ 嵌入 $\mathbb{P}^n$. 将仿射变换群 $\mathrm{Aff}(F^n)$ (定义–命题 14.2.6) 的元素表作 $\Phi_{\boldsymbol{A},\boldsymbol{v}}$ 之形, $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F) = \mathrm{End}(F^n)$ 而 $\boldsymbol{v} \in F^n$ 视同列向量. 易得群的嵌入

$$\mathrm{Aff}(F^n) \hookrightarrow \mathrm{PGL}(n+1, F) \overset{\text{等同于}}{=\!=} \{\text{射影变换 } \mathbb{P}^n \to \mathbb{P}^n\}$$

$$\Phi_{\boldsymbol{A},\boldsymbol{v}} \longmapsto \left[\begin{array}{ccc|c} & & & \\ & \boldsymbol{A} & & \boldsymbol{v} \\ & & & \\ \hline 0 & \cdots & 0 & 1 \end{array}\right]$$

符号 $[\cdots]$ 代表 $\mathrm{PGL}(n+1, F)$ 的元素. 坐标无关的版本 $\mathrm{Aff}(W) \hookrightarrow \mathrm{PGL}(F \oplus W)$ 留给读者表述.

练习 14.5.4 设 σ 为 $\mathbb{P}^1$ 上的射影变换. 证明若 σ 在 $\mathbb{P}^1$ 上有至少 3 个不动点, 则 $\sigma = \mathrm{id}_{\mathbb{P}^n}$.

提示〉 设 σ 来自 $T \in \mathrm{GL}(2, F)$, 则不动点对应到 T 的特征向量.

练习 14.5.5 点与线的包含关系是射影几何学的经典主题. 设 $x, y, z \in \mathbb{P}(V)$ 共线而两两相异, $x', y', z' \in \mathbb{P}(V)$ 也共线而两两相异. 证明存在唯一的 $\sigma \in \mathrm{PGL}(V)$ 使得 $(\sigma(x), \sigma(y), \sigma(z)) = (x', y', z')$.

提示〉 关于存在性, 由于 V 中任两个 2 维子空间都能以 $\mathrm{GL}(V)$ 相互搬运, 故由射影变换可简化到 $\mathbb{P}^1$ 情形. 其次再以射影变换简化到 $x = x' = F\boldsymbol{e}_0$ 和 $z = z' = F\boldsymbol{e}_1$ 的情形, 其中 $\boldsymbol{e}_0$ 和 $\boldsymbol{e}_1$ 是 F^2 的标准基, 然后考虑适当的对角矩阵作用. 唯一性由上一道练习解决.

相比之下, $\mathbb{P}(V)$ 的共线相异四点[4] (计顺序) 的 $\mathrm{PGL}(V)$-轨道则不只一个, 其描述

[4] 当 $|F| = 2$ 时不存在共线相异四点, 以下讨论均排除此情形.

涉及以下构造. 设 $\tilde{x},\tilde{y}$ 和 $\tilde{z},\tilde{w}$ 皆是 2 维 F-向量空间 V_0 的基. 设

$$\begin{aligned}\tilde{x} &= a\tilde{z} + b\tilde{w}\\ \tilde{y} &= c\tilde{z} + d\tilde{w}\end{aligned}$$

其中 $a,b,c,d\in F$. 记

$$\left(\frac{\tilde{x},\tilde{y}}{\tilde{z},\tilde{w}}\right) := \begin{vmatrix} a & b\\ c & d\end{vmatrix} \in F^\times.$$

以下在 $x\in\mathbb{P}(V)$ 所对应的 1 维子空间中任取非零向量 $\tilde{x}\in V$, 依此类推.

定义–命题 14.5.6 (交比) 设 $x,y,z,w\in\mathbb{P}(V)$ 相异, 而且同属于 $\mathbb{P}(V_0)$, 其中 $V_0\subset V$ 是 2 维子空间; 因此按先前的符号有

$$\langle\tilde{x},\tilde{z}\rangle = \langle\tilde{x},\tilde{w}\rangle = \langle\tilde{y},\tilde{z}\rangle = \langle\tilde{y},\tilde{w}\rangle = V_0.$$

定义此四点的交比为 F 的元素

$$(x,y;z,w) := \left(\frac{\tilde{x},\tilde{z}}{\tilde{x},\tilde{w}}\right)\cdot\left(\frac{\tilde{y},\tilde{z}}{\tilde{y},\tilde{w}}\right)^{-1},$$

它只与 x,y,z,w (计顺序) 相关.

证明 选定 $t\in F^\times$. 设 $\tilde{x}$ 被换为 $t\tilde{x}$, 从

$$\begin{aligned}\tilde{x} &= 1\cdot\tilde{x} + 0\cdot\tilde{w}\\ \tilde{z} &= c\cdot\tilde{x} + d\cdot\tilde{w}\end{aligned}$$

得到

$$\begin{aligned}t\tilde{x} &= 1\cdot t\tilde{x} + 0\cdot\tilde{w}\\ \tilde{z} &= t^{-1}c\cdot t\tilde{x} + d\cdot\tilde{w}\end{aligned}$$

故有 $\left(\frac{\tilde{x},\tilde{z}}{\tilde{x},\tilde{w}}\right) = \left(\frac{t\tilde{x},\tilde{z}}{t\tilde{x},\tilde{w}}\right)$, 交比不变. 同样地, 将 $\tilde{y}$ 换成 $t\tilde{y}$ 不改变交比.

设 $\tilde{z}$ 被换为 $t\tilde{z}$, 则有

$$\begin{aligned}\tilde{x} &= 1\cdot\tilde{x} + 0\cdot\tilde{w}\\ t\tilde{z} &= tc\cdot\tilde{x} + td\cdot\tilde{w}\end{aligned}$$

故 $\left(\frac{\tilde{x},t\tilde{z}}{\tilde{x},\tilde{w}}\right) = t\left(\frac{\tilde{x},\tilde{z}}{\tilde{x},\tilde{w}}\right)$; 类似地 $\left(\frac{\tilde{y},t\tilde{z}}{\tilde{y},\tilde{w}}\right) = t\left(\frac{\tilde{y},\tilde{z}}{\tilde{y},\tilde{w}}\right)$, 交比仍不变. 同理可见将 $\tilde{w}$ 换成 $t\tilde{w}$ 不改变交比. □

留意到定义中的 V_0 由 x,y,z,w 唯一确定. 交比的上述定义纯以向量空间语言表述, 故它被向量空间的同构所搬运; 特别地, 对所有 $\sigma\in\mathrm{PGL}(V)$ 皆有交比等式

$$(\sigma(x),\sigma(y);\sigma(z),\sigma(w))=(x,y;z,w). \tag{14.5.1}$$

因此我们称交比是共线相异四点的 $\mathrm{PGL}(V)$-不变量.

接着来寻求交比的具体公式.

命题 14.5.7 设 $x_1,x_2,x_3,x_4\in\mathbb{P}^1$ 相异, 用齐次坐标表作

$$x_i=(\alpha_i:\beta_i),\quad 1\leq i\leq 4,$$

则有

$$(x_1,x_2;x_3,x_4)=\frac{\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_3&\beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_2&\beta_2\\ \alpha_4&\beta_4\end{vmatrix}}{\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_4&\beta_4\end{vmatrix}\cdot\begin{vmatrix}\alpha_2&\beta_2\\ \alpha_3&\beta_3\end{vmatrix}};$$

如果进一步要求 $\alpha_i\neq 0$ 对所有 i 成立, 并且命 $t_i:=\frac{\beta_i}{\alpha_i}$, 则

$$(x_1,x_2;x_3,x_4)=\frac{(t_3-t_1)(t_4-t_2)}{(t_4-t_1)(t_3-t_2)}.$$

证明 沿用定义–命题 14.5.6 的记法. 取 $\tilde{x}_i=(\alpha_i,\beta_i)\in F^2$, 则从 $\tilde{x}_1=\tilde{x}_1$ 和 $\tilde{x}_3=c\tilde{x}_1+d\tilde{x}_4$ 得到 $\left(\frac{\tilde{x}_1,\tilde{x}_3}{\tilde{x}_1,\tilde{x}_4}\right)=d$ 以及

$$\alpha_3=c\alpha_1+d\alpha_4,\quad \beta_3=c\beta_1+d\beta_4.$$

按此算出

$$\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_3&\beta_3\end{vmatrix}=d\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_4&\beta_4\end{vmatrix},$$

$$\left(\frac{\tilde{x}_1,\tilde{x}_3}{\tilde{x}_1,\tilde{x}_4}\right)=d=\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_3&\beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_1&\beta_1\\ \alpha_4&\beta_4\end{vmatrix}^{-1}.$$

交换 $\tilde{x}_1$ 和 $\tilde{x}_2$ 的角色可得

$$\left(\frac{\tilde{x}_2,\tilde{x}_3}{\tilde{x}_2,\tilde{x}_4}\right)=\begin{vmatrix}\alpha_2&\beta_2\\ \alpha_3&\beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_2&\beta_2\\ \alpha_4&\beta_4\end{vmatrix}^{-1}.$$

代入交比定义即得 $(x_1,x_2;x_3,x_4)$ 的第一个公式.

因为交比仅依赖齐次坐标, 在 $\alpha_i\neq 0$ 的前提下, 不妨进一步设 $\alpha_i=1$, 从而 $t_i=\beta_i$; 按此计算行列式即得第二个公式. □

作为例子, 且在 $\mathbb{P}^1$ 中取相异的点

$$x_1=(1:\lambda),\quad x_2=(1:1),\quad x_3=(1:0),\quad x_4=(0:1),$$

此处 $\lambda\in F\smallsetminus\{0,1\}$. 基于常用的写法 $\mathbb{P}^1=U_0\sqcup\{(0:1)\}\simeq F\sqcup\{\infty\}$, 请读者代入命题 14.5.7 计算

$$(x_1,x_2;x_3,x_4)=\lambda,\quad \text{或简练地记作 } (\lambda,1;0,\infty)=\lambda. \tag{14.5.2}$$

定理 14.5.8 设 x,y,z,w 为 $\mathbb{P}(V)$ 中的共线相异四点, $x',y',z',w'\in\mathbb{P}(V)$ 亦然, 则存在 $\sigma\in\mathrm{PGL}(V)$ 使得 $(\sigma(x),\sigma(y),\sigma(z),\sigma(w))=(x',y',z',w')$ 的充要条件是交比的等式

$$(x,y;z,w)=(x',y';z',w').$$

此外, 交比能取到的所有值正好是 $F\smallsetminus\{0,1\}$.

证明 条件的必要性来自 (14.5.1). 以下处理充分性. 根据练习 14.5.5, 存在 $\tau,\tau'\in\mathrm{PGL}(V)$ 使得

$$(\tau(y),\tau(z),\tau(w))=(\tau'(y'),\tau'(z'),\tau'(w'));$$

事实上, 若将它们所共之线等同于 $\mathbb{P}^1$, 则可以进一步设等式两边都等于 $(1,0,\infty)$. 然而 τ 和 τ' 不改变交比, 故

$$(\tau(x),1;0,\infty)=(x,y;z,w)=(x',y',z',w')=(\tau'(x'),1;0,\infty).$$

代入 (14.5.2) 得 $\tau(x)=\tau'(x')$, 换言之 $\sigma:=\tau^{-1}\tau'$ 映 (x',y',z',w') 为 (x,y,z,w).

以上论证也说明如何将交比化到 $(\lambda,1;0,\infty)$ 的情形来计算, 从而说明交比的取值能且仅能是 $F\smallsetminus\{0,1\}$ 的元素. □

上述结果表明交比作为共线相异四点的 $\mathrm{PGL}(V)$-不变量是完备的, 换言之, 它完全描述了相应的 $\mathrm{PGL}(V)$-轨道:

$$\begin{aligned}\mathrm{PGL}(V)\backslash\big\{(x,y,z,w)\in\mathbb{P}(V)^4:\text{共线且相异}\big\}&\xrightarrow{1:1}F\smallsetminus\{0,1\}\\ \mathrm{PGL}(V)(x,y,z,w)\text{ 的轨道}&\longmapsto(x,y;z,w).\end{aligned}$$

须留意交比 $(x,y;z,w)$ 关乎四个点的顺序, 本章习题将给出具体的变换公式.

14.6 仿射空间的凸子集

本节聚焦于凸性, 而仿射空间的语言为此提供了一个自然的舞台. 以下考虑的向量空间均为实向量空间, 仿射空间因此是在实向量空间作用下而论的.

定义 14.6.1 设 $x_0,\dots,x_n$ 为仿射空间 E 的点 ($n\in\mathbb{Z}_{\geq 1}$), 它们的**凸组合**定义为形如

$$\sum_{i=0}^{n} t_i x_i,\quad t_0,\dots,t_n\in\mathbb{R}_{\geq 0},\quad t_0+\cdots+t_n=1$$

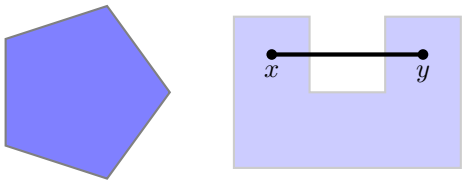
的仿射组合. 对任意 $x_0,x_1\in E$, 其间的**线段**定义为

$$[x_0,x_1]:=\{x_0,x_1\text{ 的所有凸组合}\}\subset E.$$

对于非空子集 $K\subset E$, 如有 $x,y\in K\implies [x,y]\subset K$, 则称 K 为 E 的**凸子集**.

从定义可知任意一族凸子集的交若非空, 则仍是凸子集.

考虑到 §14.1 对仿射组合的几何诠释, 可知 $[x,y]$ 确实能理解为 x 和 y 之间的 "线段", 包含端点. 凸子集的几何意涵因此显然. 譬如以下左图着色部分是平面的凸子集, 右图的着色部分则非凸, 因为其中标为 $[x,y]$ 的线段不完全包含于该子集.



凸组合既然是仿射组合的特例, §14.1 关于仿射组合的大部分结论都有凸版本. 例如练习 14.1.6 的凸版本 (请读者迅速地证明) 说明一般的凸组合可化约到两个点的凸组合, 而凸组合的凸组合仍是凸组合.

此外, 定义 14.1.13 及后续讨论也有以下版本.

定义 14.6.2 设 S 为仿射空间 E 的非空子集. 定义它的**凸包**为

$$\operatorname{conv}(S):=\bigcap_{\substack{K:E\text{ 的凸子集}\\ K\supset S}} K.$$

事实上, $\operatorname{conv}(S)$ 可以更明确地表为 S 中所有点的凸组合所成之集合.

因此 $\operatorname{conv}(S)\subset\operatorname{aff}(S)$, 而 S 是凸子集当且仅当 $\operatorname{conv}(S)=S$.

例 14.6.3 在平面 $\mathbb{R}^2$ 上考虑不共线的三点 x,y,z, 则 $S:=\{x,y,z\}$ 的仿射包 $\operatorname{aff}(S)$ 是整个平面, 凸包 $\operatorname{conv}(S)$ 则是以 S 为顶点集的三角形.

定义 14.6.4 对于 E 中的非空凸子集 K 和 $x\in K$, 如果 $K\smallsetminus\{x\}$ 仍是凸的, 则称 x 为 K 的一个**端点**; 等价的说法是若 $a,b\in K$ 皆不等于 x, 则 $x\notin[a,b]$.

端点在凸集的一般理论中扮演要角; 推论 14.7.10 将对多面体的特例进一步刻画端点.

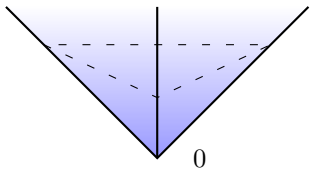
另一类重要的凸集是凸锥, 它们是在向量空间, 亦即带基点的仿射空间中定义的.

定义 14.6.5 设 C 为向量空间 V 的非空子集. 如果对于所有 $x \in C$ 皆有

$$\mathbb{R}_{\geq 0}x := \{tx : t \in \mathbb{R}_{\geq 0}\} \subset C,$$

则称 C 为 V 中的**锥**. 如果锥 C 本身还是 V 的凸子集, 则称之为**凸锥**.

因此锥总是包含原点 0. 形如 $\mathbb{R}_{\geq 0}x$ 的子集 (要求 $x \neq 0$) 称为 V 中的**射线**. 下图是凸锥的一个典型例子; 虚线部分标出凸锥的一个截面.



(14.6.1)

定义–命题 14.6.6 设 C 为 V 的子集, 在 V 的对偶空间 $V^\vee$ 中定义子集

$$C^* := \{\lambda \in V^\vee : \forall x \in C,\ \lambda(x) \geq 0\},$$

则 C^* 是 $V^\vee$ 中的凸锥, 称为 C 的**对偶锥**. 典范嵌入 $V \hookrightarrow V^{\vee\vee}$ 限制为 $C \hookrightarrow C^{**}$.

证明 容易看出 C^* 为凸锥. 设 $x \in C$ 而 $\lambda \in C^*$, 则 x 在 $V^{\vee\vee}$ 中的像映 λ 为 $\lambda(x) \geq 0$, 这就说明 $C \hookrightarrow C^{**}$. □

显然 $C \subset D$ 蕴涵 $D^* \subset C^*$. 对于 $\dim V$ 有限的情形, 引理 14.9.3 将刻画满足 $C \xrightarrow{\sim} C^{**}$ 的子集 C.

定义 14.6.7 设 C 为向量空间 V 中的凸锥, ρ 为包含于 C 的射线. 如果对于任意相异之 $a, b \in C$ 皆有

$$(\exists\, x \in [a,b] \smallsetminus \{a,b\},\ x \in \rho) \implies [a,b] \subset \rho,$$

则称 ρ 为 C 的**端射线**.

练习 14.6.8 设 $\mathbb{R}_{\geq 0}x$ 是凸锥 C 的端射线, 而且 $x = \sum_{i=1}^m x_i$, 其中 $x_1, \ldots, x_m \in C$, 证明 $x_1, \ldots, x_m \in \mathbb{R}_{\geq 0}x$.

提示〉 先化约到 $m = 2$ 情形. 不失一般性可设 $x_1, x_2 \neq 0$. 从 $x \in [x_1, x_1 + 2x_2]$ 推导 $x_1 \in \mathbb{R}_{\geq 0}x$; 同理有 $x_2 \in \mathbb{R}_{\geq 0}x$.

端射线的定义思路和端点类似, 请读者思考其几何图像. 稍后在 §14.9 将对多面锥及其端射线作进一步的考察.

以下进一步要求 E 是有限维的. 若选取 E 的基点和 V 的基以将 E 等同于 $\mathbb{R}^m$, 则可以按照数学分析中熟悉的方式谈论 E 的子集是否开, 闭或有界; 这些概念和基点 (仅差一个平移) 与基 (仅差一个可逆线性变换) 的选取无关. 仿射子空间显然是闭子集.

定义 14.6.9 对于 E 中的凸子集 K, 它的**相对内部** $\mathrm{r.int}(K) \subset K$ 定义为 K 作为仿射空间 $\mathrm{aff}(K)$ 的子集的内部, 它的**边界** ∂K 定义为 $\overline{K} \smallsetminus \mathrm{r.int}(K)$, 其中 $\overline{K}$ 代表 K 在 E 中 (等价地说, 在 $\mathrm{aff}(K)$ 中) 的闭包.

为了将闭包与内部的概念施于 $\mathrm{aff}(K)$ 的子集, 同样须选定同构 $\mathrm{aff}(K) \simeq \mathbb{R}^d$, 其中 $d := \dim \mathrm{aff}(K)$, 然而这些概念同样不依赖同构的选取.

引理 14.6.10 设 K 为 E 的非空凸子集, 则 $\mathrm{r.int}(K) \neq \varnothing$.

证明 所有 $y \in \mathrm{aff}(K)$ 皆能表作 K 中元素的仿射组合. 现在让 y 遍历 $\mathrm{aff}(K)$ 的某个仿射基, 运用练习 14.1.6 即可从 K 中取到 $\mathrm{aff}(K)$ 的生成元. 以练习 14.1.9 从中萃取 $\mathrm{aff}(K)$ 的仿射基 $x_0, \dots, x_d \in K$, 以此将 $\mathrm{aff}(K)$ 等同于 $\mathbb{R}^d$. 现在容易看出系数全正的凸组合 $\sum_{i=0}^{d} t_i x_i$ 必属于 $\mathrm{r.int}(P)$, 证毕. □

直观上, 这说明 K 作为 $\mathrm{aff}(K)$ 的子集总是 "有厚度" 的.

引理 14.6.11 设 K 为 E 的凸子集, $x \in \mathrm{r.int}(K)$ 而 $y \in K$, 则 $[x, y] \smallsetminus \{y\} \subset \mathrm{r.int}(K)$.

证明 不失一般性可设 $E = \mathrm{aff}(K)$. 设 $z = tx + (1-t)y$, $t \neq 0$. 列式可见当 $t \neq 0$ 和 y 固定, x 和 z 可以用连续映射相互反解. 按条件 x 有包含于 K 的邻域 $\mathcal{U}$, 而对所有 $x' \in \mathcal{U}$ 皆有 $[x', y] \subset K$, 相应地 z 便有包含于 K 的邻域 $\mathcal{V}$, 由此推得 $z \in \mathrm{r.int}(K)$. □

现在介绍三类简单而重要的凸子集.

约定 14.6.12 以下的 E 均取为有限维实仿射空间. 设 $\alpha : E \to \mathbb{R}$ 为非常值仿射线性函数, 其中 $\mathbb{R}$ 带有标准的仿射空间结构; α 必然满. 从 α 得到 E 的以下子集:

▷ **超平面** $H_\alpha := \{x \in E : \alpha(x) = 0\}$,

▷ **正半空间** $H_\alpha^{\geq 0} := \{x \in E : \alpha(x) \geq 0\}$,

▷ **负半空间** $H_\alpha^{\leq 0} := \{x \in E : \alpha(x) \leq 0\}$.

显然 H_α, $H_\alpha^{\leq 0}$ 和 $H_\alpha^{\geq 0}$ 都是 E 的闭凸子集; 半空间也有开的版本 $H_\alpha^{>0}$ 和 $H_\alpha^{<0}$. 回忆到所有超平面 $H \subset E$ 都形如 $H = H_\alpha$ (练习 14.3.6).

定义 14.6.13 设 K 为 E 的闭凸子集. 如果超平面 $H = H_\alpha$ 与 K 有交, 而且 $K \subset H_\alpha^{\geq 0}$ 或 $K \subset H_\alpha^{\leq 0}$ 二居其一, 则称 H 为 K 的一个**支持超平面**.

多面体的支持超平面具有丰富的组合学结构. 这是下一节探讨的主题.

练习 14.6.14 说明 H_α 和 $H_\alpha^{\geq 0}$ (或 $H_\alpha^{>0}$) 分别能在多大程度上确定 $\alpha : E \to \mathbb{R}$. 提示〉 答案是精确到 $\mathbb{R}^\times$ (或 $\mathbb{R}^\times_{>0}$) 的乘法作用.

14.7 多面体

本节延续 §14.6 关于实仿射空间的假设与定义, 特别是约定 14.6.12.

定义 14.7.1 若非空子集 $P \subset E$ 是有限多个半空间的交, 则称 P 为 E 中的**多面体**, 并且定义 $\dim P := \dim \mathrm{aff}(P)$; 有界多面体称为**多胞体**.

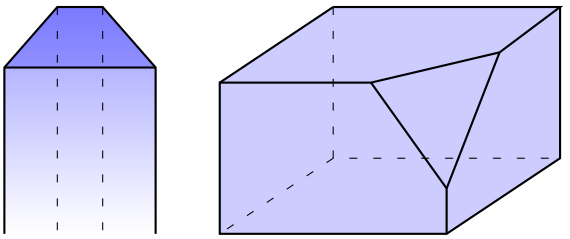
多面体总是闭凸子集; 2 维多胞体又称多边形. 如果取 $E = \mathbb{R}^m$, 则其中的多面体无非是由有限多个线性不等式 $a_{i1}X_1 + \cdots + a_{im}X_m \leq b_i$ 截出的区域, 其中 $1 \leq i \leq n$ 而 $a_{ij}, b_i \in \mathbb{R}$. 留意到定义中容许 $n = 0$, 相应地 $P = E$.

在 $E = \mathbb{R}^3$ 的情形, §11.11 介绍的五种正多面体都是多胞体. 当维数更高时, 多面体可以有各种维数的 "面", 都需要纳入考虑, 术语因此有所改动.

定义 14.7.2 设 P 为 E 中的多面体. 如果 $F = P$ 或 $F = H \cap P$, 其中 H 是 P 的某个支持超平面, 则称 F 为 P 的一个**面**; 注意到多面体 (或多胞体) 的面仍是多面体 (或多胞体). 严格包含于 P 的面称为**真面**, 维数为 $\dim P - 1$ 的面称为 P 的**台面**, 维数为 1 的面称为**边**, 维数为 0 的面称为**顶点**.

按照上述定义, 面不能是空集. 若 F 为 P 的真面, 则截出 F 的支持超平面 H 不可能包含 $\mathrm{aff}(P)$, 从而 $\mathrm{aff}(H \cap P) \subset H \cap \mathrm{aff(P)}$ 导致 $\dim F < \dim P$, 这与几何直观相符.

下图描绘 $\mathbb{R}^3$ 中两个并不算十分复杂的多面体, 右侧是多胞体. 请读者描述它们的面, 并且为每个面指出一个对应的支持超平面.



以下给出的诸般性质在维数 ≤ 3 时都有自明的几何图像, 然而高维情形需要严格论证. 读者初次阅读时可以先跳过证明.

引理 14.7.3 设 P 为 E 中的多面体, F 为其面. 若 $x, y \in P$, 而且存在 $z \in [x, y] \smallsetminus \{x, y\}$ 使得 $z \in F$, 则 $[x, y] \subset F$.

证明 可设 F 为真面. 取 α 使得 $F = H_\alpha \cap P$ 而且 $P \subset H_\alpha^{\geq 0}$. 闭区间 $[0, 1]$ 上的函数 $t \mapsto \alpha(tx + (1-t)y) = t\alpha(x) + (1-t)\alpha(y)$ 非负, 它或者是常值, 或者严格单调, 而且在某个 $t \neq 0, 1$ 处 (对应到 z) 为 0, 故它恒为 0. 因此 $[x, y] \subset H_\alpha \cap P = F$. □

引理 14.7.4 设 P 为 E 中的多面体, H 为 E 中的超平面, 满足 $H \cap P \neq \varnothing$, 则

$$H \text{ 是支持超平面} \wedge H \not\supset P \iff H \cap P \subset \partial P.$$

证明 先处理 $\Longrightarrow$. 设 H 为 P 的支持超平面. 若存在 $x \in H \cap \mathrm{r.int}(P)$, 则可在 $\mathrm{aff}(P)$ 中选定以 x 为原点的坐标系, 再取足够小的球 B, 以 x 为球心, 使得 $B \subset P$. 对 B 的每条直径应用引理 14.7.3 可知每条直径都包含于面 $H \cap P$, 故 B 亦然. 显然 $\mathrm{aff}(B) = \mathrm{aff}(P)$, 故此时 $H \supset \mathrm{aff}(B) \supset P$.

其次处理 $\Longleftarrow$. 取 α 使得 $H = H_\alpha$. 因为 $\mathrm{r.int}(P) \not\subset H$ 且 $\mathrm{r.int}(P) \neq \varnothing$ (引理 14.6.10), 故 $H_\alpha^{<0}$ 和 $H_\alpha^{>0}$ 必有一者交 $\mathrm{r.int}(P)$. 不妨设 $x \in \mathrm{r.int}(P) \cap H_\alpha^{>0}$. 假若 H 非支持超平面, 则必存在 $y \in P \cap H_\alpha^{<0}$. 引理 14.6.11 表明 $[x, y] \smallsetminus \{y\} \subset \mathrm{r.int}(P)$; 基于连续性, 存在 $z \in [x, y]$ 使得 $\alpha(z) = 0$, 而 $z \neq y$, 故 $H \cap \mathrm{r.int}(P) \neq \varnothing$, 矛盾. □

如果从多面体 P 的表法 $P = H_{\alpha_1}^{\geqslant 0} \cap \cdots \cap H_{\alpha_k}^{\geqslant 0}$ 中删除任何一项 $H_{\alpha_i}^{\geqslant 0}$ 皆给出严格包含 P 的集合, 则称此表法**无赘**.

引理 14.7.5 *设 $P = H_{\alpha_1}^{\geqslant 0} \cap \cdots \cap H_{\alpha_k}^{\geqslant 0}$ 为 E 中的多面体, 则*

(i) *若 E 的凸子集 K 满足 $K \subset \partial P$, 必存在 $1 \leqslant i \leqslant k$ 使得 $K \subset H_{\alpha_i}$;*

(ii) *P 的所有台面皆形如 $H_{\alpha_i} \cap P$, 其中 $i = 1, \ldots, k$;*

(iii) *当上述表法无赘, 而且 $\mathrm{aff}(P) = E$ 时,*

- *每个 $H_{\alpha_i} \cap P$ 都是 P 的台面,*
- *半空间 $H_{\alpha_1}^{\geqslant 0}, \ldots, H_{\alpha_k}^{\geqslant 0}$ 两两相异, 精确到重排由 P 唯一确定.*

证明 以反证法处理 (i). 设若不然, 对每个 i 选取 $x_i \in K \smallsetminus H_{\alpha_i}$, 凸组合 $x := \frac{1}{n}(x_1 + \cdots + x_k)$ 属于 $K \cap \bigcap_{i=1}^k H_{\alpha_i}^{>0}$; 由于 $\bigcap_i H_{\alpha_i}^{>0}$ 是 E 的开子集, 而且包含于 P, 故 $x \in \mathrm{r.int}(P)$. 矛盾.

接着处理 (ii). 以 $\mathrm{aff}(P)$ 代 E, 相应地以 $\alpha_i|_{\mathrm{aff}(P)}$ 代 α_i, 但扣除常值的项, 问题便化到 $E = \mathrm{aff}(P)$ 的特例. 设 P 的台面 F 由支持超平面 H 截出. 引理 14.7.4 蕴涵 $H \cap P \subset \partial P$, 由 (i) 遂有 i 使得 $F \subset H_{\alpha_i}$. 分两种情形讨论.

- 若 $H = H_{\alpha_i}$ 则推得 $F = H_{\alpha_i} \cap P$.
- 若 $H \neq H_{\alpha_i}$, 则 $\mathrm{aff}(F) \subset H \cap H_{\alpha_i}$. 然而 $H \cap H_{\alpha_i} \neq \varnothing$ 确保 $\alpha_i|_H$ 非常值, 故 $\dim H \cap H_{\alpha_i} \leqslant \dim E - 2$, 与 $\dim F = \dim E - 1$ 矛盾.

以下处理 (iii) 的前半段. 为了证明所有 $H_{\alpha_i} \cap P$ 皆是 P 的台面, 不妨设 $i = 1$ 而 $k \geqslant 2$. 命 $P' := \bigcap_{j=2}^k H_{\alpha_j}^{\geqslant 0}$. 由于 H_{α_1} 交 P' 却非 P' 的支持超平面 (用无赘条件), 引理 14.7.4 蕴涵 H_{α_1} 交 $\mathrm{r.int}(P')$. 由于 $\mathrm{aff}(P') = E$, 这表明 $H_{\alpha_1} \cap P' = H_{\alpha_1} \cap P$ 包含 H_{α_1} 的某个非空开子集, 由此可知 $\dim H_{\alpha_1} \cap P = \dim H_{\alpha_1} = \dim P - 1$.

断言 (ii) 的后半段是先前结果的直接应用: 基于维数理由, 从台面 $F = H \cap P$ 可反解 $H = \mathrm{aff}(F)$, 对应的半空间则按 P 落在 H 的哪一侧来确定. 这也说明不同的 i, j 确定不同台面, 否则 $H_{\alpha_i}^{\geqslant 0} = H_{\alpha_j}^{\geqslant 0}$. □

推论 14.7.6 设 P 为 E 中的多面体, 则 ∂P 是 P 的所有台面之并, 而 P 的每个真面都包含于某个台面.

证明 不失一般性可设 $\operatorname{aff}(P)=E$. 取无赘表法 $P=\bigcap_{i=1}^{k} H_{\alpha_i}^{\geqslant 0}$. 引理 14.7.4 说明每个真面都包含于 ∂P; 引理 14.7.5 (i) 说明若 $x\in\partial P$, 则存在 i 使得 $\alpha_i(x)=0$, 从而 (iii) 说明 x 属于台面 $H_{\alpha_i}\cap P$, 第一则断言得证.

其次, 将 P 给定的真面表作 $H\cap P$, 则 $H\not\supset P$, 故 $H\cap P\subset\partial P$ (引理 14.7.4). 对引理 14.7.5 (i) 代入 $K=H\cap P$ 即得第二则断言. □

如果多面体 P 的子集 A 是 E 的仿射子空间, 而且 P 不含严格包含 A 的仿射子空间, 则称 A 为 P 的极大仿射子空间.

命题 14.7.7 设 $P=H_{\alpha_1}^{\geqslant 0}\cap\cdots\cap H_{\alpha_k}^{\geqslant 0}$ 为 E 中的多面体. 记 W 为所有 H_{α_i} 的向量部分之交, 则

$$\{A : P \text{ 的极大仿射子空间}\}=\{x+W : x\in P\}.$$

证明 设 A 为包含于 P 的仿射子空间, 记 W' 为其向量部分. 取 $x\in A$. 对所有 $w'\in W'$, 易见 $x+\mathbb{R}w'\subset P$ 导致 $w'\in W$, 因此 $W'\subset W$ 而 $A\subset x+W$. 反之, 形如 $x+W$ 的仿射子空间 ($x\in P$) 显然包含于 P. □

定理 14.7.8 (多面体的面结构) 设 P 为 E 中的多面体.

(i) 若 F 是 P 的面, 而 G 是 F 的面, 则 G 是 P 的面.

(ii) 若 $F,G\subset P$ 为有交的两个面, 则 $F\cap G\subset P$ 为面.

(iii) 所有真面 F 都能表为台面之交; 特别地, P 的面数有限.

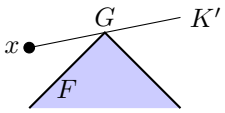
(iv) 考虑面的链 $F_0\subsetneq\cdots\subsetneq F_m=P$, 要求它是极大的, 亦即链中无法再插项, 则 F_0 是 P 的极大仿射子空间, 而 $\dim F_{i+1}=\dim F_i+1$ 对所有 $0\leqslant i<m$ 成立.

(v) 对所有 $x\in P$, 存在唯一的面 F 使得 $x\in\operatorname{r.int}(F)$; 相对于包含关系, F 是包含 x 的极小面.

证明 考虑 (i). 可设 F 为真面. 推论 14.7.6 说明 F 包含于某个台面 F'; 递归可知 G 是 F' 的面. 以 F' 代 F 便将问题化约到 G 为台面 F 的真面的情形.

不失一般性可设 $\operatorname{aff}(P)=E$. 在 $\operatorname{aff}(F)$ 中取 F 的支持超平面 K 使得 $K\cap F=G$; 我们寻求取 $x\in E\smallsetminus\operatorname{aff}(F)$ 将 K 扩充为 $K':=\operatorname{aff}(K\cup\{x\})$; 基于维数理由, K' 是 E 中的超平面, $K'\cap\operatorname{aff}(F)=K$, 但需要适当选择 x 以确保它是 P 的支持超平面, 而且满足 $K'\cap P=G$.

为此, 遵循引理 14.7.5 的符号, 先取定 P 的无赘表法, 再取 E 中截出 F 的 H_{α_i}. 存在 $x\in H_{\alpha_i}^{<0}\cap\bigcap_{j\neq i}H_{\alpha_j}^{>0}$. 命 $K':=\operatorname{aff}(K\cup\{x\})$. 低维几何图像如下.



先来验证 $K'\cap P=G$: 首先 $\supset$ 是容易的, 另一方面若 $y\in K'\cap P$ 表作凸组合 $tx+(1-t)k$, 其中 $k\in K$, 则 $\alpha_i(y)=t\alpha_i(x)+(1-t)\alpha_i(k)=t\alpha_i(x)\geq 0$ 蕴涵 $t=0$, 亦即 $y\in K\cap P=K\cap\operatorname{aff}(F)\cap P=K\cap F=G$.

其次验证 P 的任两点总落在 K' 的同侧. 设有 $a,b\in P\smallsetminus K'$ 相异使得 $[a,b]\cap K'\neq\varnothing$, 则存在 $c\in[a,b]\smallsetminus\{a,b\}$ 使得 $c\in K'\cap P=G\subset F$; 应用引理 14.7.3 遂有 $[a,b]\subset F$. 因此 $[a,b]$ 落在 K 的同侧 (在 $\operatorname{aff}(F)$ 中), 故落在 K' 的同侧 (在 E 中). 矛盾.

考虑 (ii). 可设 F 为真面. 设 H 为 P 的支持超平面, $H\cap P=F$, 则因为 $F\cap G$ 非空, H 对于 E 中的多面体 G 也是支持超平面, 并且 $H\cap G=F\cap G$. 于是 $F\cap G$ 是 G 的面, 从 (i) 知它也是 P 的面.

考虑 (iii). 对 $c:=\dim P-\dim F$ 递归论证; $c=1$ 情形平凡, 以下设 $c\geq 2$. 取无赘表法 $P=\bigcap_{i=1}^{k}H_{\alpha_i}^{\geq 0}$. 引理 14.7.5 (ii) 和推论 14.7.6 说明不失一般性可设 $F\subset H_{\alpha_1}\cap P=:P'$, 而 $\dim P'=\dim P-1$. 记 $\alpha_i':=\alpha_i|_{H_{\alpha_1}}$; 于是 $P'=\bigcap_{\substack{2\leq i\leq k\\ \alpha_i'\text{ 非常值}}}H_{\alpha_i'}^{\geq 0}$. 因此 F 作为 P' 的面或者是 P', 或者可写成 $\bigcap_{i\in I}H_{\alpha_i'}\cap P'$, 其中 $I\subset\{2,\dots,k\}$.

综上, $F=\bigcap_{i\in I\cup\{1\}}H_{\alpha_i}\cap P$, 引理 14.7.5 (iii) 说明此为台面之交.

考虑 (iv). 当 $m\geq 1$ 时, F_{m-1} 包含于 P 的某个台面, 故极大条件蕴涵 F_{m-1} 是台面, $\dim F_{m-1}=\dim P-1$. 鉴于 (i) 可知 $F_0\subsetneq\cdots\subsetneq F_{m-1}$ 相对于多面体 F_{m-1} 依然是极大链, 对 $\dim P$ 递归地论证便有维数关系式.

容易证明无真面的多面体是仿射子空间, 故 F_0 可表为 $x+W'$ 的形式. 考虑 P 中包含 F_0 的极大仿射子空间, 则命题 14.7.7 及其符号蕴涵 $W'\subset W$. 另一方面 F_0 是若干个台面之交, 考虑向量部分之交易见 $W'\supset W$. 故命题 14.7.7 说明 $F_0=x+W$ 是极大仿射子空间.

最后考虑 (v). 基于 (ii) 和 (iii), 所有含 x 的面取交仍是面, 因此含 x 的极小面是唯一的, 记为 F. 对 F 应用推论 14.7.6 可得 $x\in F\smallsetminus\partial F=\operatorname{r.int}(F)$. 若 G 是另一个含 x 的面, 则 $F\subsetneq G$ 蕴涵 $F\subset\partial G$, 故 $x\notin\operatorname{r.int}(G)$. □

推论 14.7.9 *多面体 P 的每个极小面都是 P 的极大仿射子空间. 如果 P 包含至少一个有界的面, 则 P 的极小面无非是 P 的顶点.*

证明 设 F_0 是极小面, 将它扩充为面的极大链 $F_0\subsetneq\cdots\subsetneq F_m=P$, 则定理 14.7.8 (iv) 蕴涵 F_0 是 P 的极大仿射子空间.

顶点当然是极小面. 反之若面 F 有界, 将它双向扩充为极大链 $F_0\subsetneq\cdots\subsetneq F_m=P$, 则 F_0 是有界仿射子空间, 从而是独点集. 基于命题 14.7.7 可得 $W=\{0\}$, 所有极大仿射空间都是独点集, 从而极小面都是顶点. □

推论 14.7.10 *多面体 P 的顶点无非是定义 14.6.4 所谓的端点.*

证明 从引理 14.7.3 推得任何顶点 x 都是 P 的端点.

反之设 x 是端点. 按照定理 14.7.8 (v) 取面 F 使得 $x \in \mathrm{r.int}(F)$. 若 $F = \{x\}$ 则 x 是顶点; 否则 $\dim \mathrm{aff}(F) \geq 1$, 容易在 $\mathrm{aff}(F)$ 中将 x 实现为某个线段 $[a,b]$ 的元素, 其中 $a, b \in \mathrm{r.int}(F) \smallsetminus \{x\}$, 矛盾. □

14.8 关于极值问题

线性规划问题和多面体理论直接相关. 抽象地说, 显式地给定

- ⋆ 有限维实仿射空间 E 中的多面体 $P = \bigcap_{i=1}^{m} H_{\alpha_i}^{\leq 0}$ (此处不取 $H_{\alpha_i}^{\geq 0}$ 是尊重应用中的惯例),
- ⋆ 仿射线性映射 $\gamma : E \to \mathbb{R}$,

线性规划相当于问

- ⋆ $\max_{x \in P} \gamma(x)$ 是否存在?
- ⋆ 如果存在, 求 $x \in P$ 使得 $\gamma(x)$ 取到极大值.

应用数学与计算机科学中的许多问题都能转化为线性规划问题. 一个基本观察如下.

命题 14.8.1 *设 P 有顶点. 仿射线性映射 γ 在 P 上若有极大值, 则必然被某个顶点取到.*

证明 设 $\gamma|_P$ 有极大值 M, 则 $P \subset H_{M-\gamma}^{\geq 0}$, 而且 $H_{M-\gamma} \cap P \neq \varnothing$, 这就说明 $H_{M-\gamma}$ 是 P 的一个支持超平面. 故极大值 M 在 P 的某个面 F 上取到.

现在取面的极大链 $F_0 \subsetneq \cdots \subsetneq F_m = P$, 使得某个 F_i 等于 F. 推论 14.7.9 蕴涵 F_0 必为 P 的顶点, 记为 $\{x\}$. 于是 $\gamma(x) = M$. □

命 $n := \dim E$. 如何具体描述 P 的顶点? 不妨设 $n \geq 1$, 否则问题平凡. 若 x 是顶点, 则定理 14.7.8 (iv) 确保 x 包含于某个台面 F; 应用引理 14.7.5 (ii) 可知存在 i_1 使得 $H_{\alpha_{i_1}} \cap P = F$. 对 $n-1$ 维仿射空间 $H_{\alpha_{i_1}}$ 和 $x \in F$ 继续递归地操作, 然后将下标排序, 便可取到序列

$$1 \leq i_1 < \cdots < i_n \leq m$$

使得

$$\bigcap_{k=1}^{n} H_{\alpha_{i_k}} = \{x\}.$$

进一步设 E 为 $\mathbb{R}$-向量空间 V 作用下的仿射空间. 兹断言当上式成立时 $\alpha_{i_1}, \ldots, \alpha_{i_n}$ 的向量部分构成 $V^\vee$ 的基. 为此, 以 x 为基点将 E 等同于 V, 则断言化为向量空间理论的已知事实.

有鉴于此, 处理线性规划的一个初步思路是搜寻 P 的顶点, 或者更具体地搜寻 $\{1,\dots,m\}$ 的子集 I, 要求所有 α_i 的向量部分构成 $V^\vee$ 的基 ($i\in I$), 借以判断 γ 在哪里取到极大值. 然而穷举顶点过于费力; 从给定的初始顶点出发, 我们需要一套更高效的算法在顶点之间过渡, 使得 γ 的值单调递增, 并且判断何时取到极大值, 以及 γ 何时无极大值. 往细处说, 我们还需要照顾到 P 无顶点的情形, 并且给出构造初始顶点的手段.

对此, 目前广泛采用的算法是 G. Dantzig 提出的**单纯形法**; 几何思路很简单: 从给定的顶点 x 出发,

- ⋆ 如果对于每一条以 x 为顶点的棱, γ 沿着该方向皆为常值或递减, 则输出 x 作为极值点;
- ⋆ 否则可以适当选择一条使 γ 严格递增的棱,
  - 若这条棱无穷延伸, 则判定 γ 无极大值,
  - 不然沿这条棱过渡到下一个顶点 x', 迭代操作.

这些操作实际是针对 I 进行的, 一切计算都能用矩阵实现.

单纯形法缘于初等的几何直观, 也和本章的其他结果一样能建基于严格的数学基础, 但是完整的说明与论证并非本书主题, 请感兴趣的读者参阅 [2, Chapter 4].

14.9 多面锥

本节设 V 为有限维实向量空间, 它也自然是自身作用下的仿射空间. 回忆到 $V^\vee := \mathrm{Hom}(V,\mathbb{R})$.

定义 14.9.1 在 V 中形如 $C = H^{\geqslant 0}_{\lambda_1}\cap\cdots\cap H^{\geqslant 0}_{\lambda_k}$ 的子集 C 称为 V 中的**多面锥**, 其中 $\lambda_1,\dots,\lambda_k\in V^\vee\smallsetminus\{0\}$.

定义中容许 $k=0$, 此时的 C 规定为 V. 多面锥显然是定义 14.6.5 所谓的凸锥, 它们也和 §14.7 的多面体理论自然相关.

引理 14.9.2 对于 V 中的非空集 C, 以下等价:

(i) C 是多面锥,

(ii) C 是多面体, 而且它的每个支持超平面都包含 0.

作为推论, 多面锥 C 的面依然是多面锥.

证明 (i) $\Longrightarrow$ (ii). 由于 $C=\bigcap_{i=1}^k H^{\geqslant 0}_{\lambda_i}$, 它当然是多面体. 以下说明它的每个支持超平面 H_α 都包含 0, 其中 α 表作非零线性映射 $\lambda\in V^\vee$ 和常数 $c\in\mathbb{R}$ 的和. 设有

$x \in H_\alpha \cap C$ 而且 $C \subset H_\alpha^{\geq 0}$, 则 $\alpha(tx) \geq 0$, 亦即 $t\lambda(x) \geq -c$ 对所有 $t \in \mathbb{R}_{>0}$ 成立; 此外又有 $\lambda(x) = -c$. 唯一可能是 $c = 0$.

(ii) $\Longrightarrow$ (i). 取多面体 C 的无赘表法 $\bigcap_{i=1}^k H_{\alpha_i}^{\geq 0}$, 则每个 H_{α_i} 都是支持超平面. 既然 $0 \in H_{\alpha_i}$, 故 α_i 实则是线性映射 $V \to \mathbb{R}$.

最后, 注意到多面锥 C 的面总由某个支持超平面 $H_\mu = \ker(\mu)$ 截出 ($\mu \in V^\vee \smallsetminus \{0\}$), 故仍然形如 $\bigcap_i H_{\lambda_i'}$, 其中 $\lambda_i' := \lambda_i|_{\ker(\mu)}$, 这是 $\ker(\mu)$ 中的多面锥. □

以下性质涉及定义–命题 14.6.6 引入的对偶锥 C^* 和双重对偶 C^{**}. 由于 V 可等同于 $V^{\vee\vee}$, 以下将 C^{**} 视同 V 的子集.

引理 14.9.3 设 C 是 V 的子集, 则 $C = C^{**}$ 当且仅当 C 形如 $\bigcap_{\lambda\in\Lambda} H_\lambda^{\geq 0}$, 其中 Λ 是 $V^\vee \smallsetminus \{0\}$ 的子集 (容许无穷, 空交理解为 V).

证明 定义蕴涵 $C^{**} = \bigcap_{\substack{\lambda\in C^*\\ \lambda\neq 0}} H_\lambda^{\geq 0}$, 故 "仅当" 方向成立. 反之设 $C = \bigcap_{\lambda\in\Lambda} H_\lambda^{\geq 0}$, 则 $\Lambda \subset C^*$, 从而 $C^{**} \subset \bigcap_{\lambda\in\Lambda} H_\lambda^{\geq 0} = C$. 配合已知的 $C \subset C^{**}$ 立得等式. □

对所有 $m \in \mathbb{Z}_{\geq 0}$ 和 $x_1, \dots, x_m \in V$, 定义

$$\sum_{i=1}^m \mathbb{R}_{\geq 0} x_i := \left\{ \sum_{i=1}^m t_i x_i \in V : \forall i,\ t_i \in \mathbb{R}_{\geq 0} \right\}.$$

这是 V 中的凸锥; 当 $m = 0$ 时将此理解为 $\{0\}$. 我们也称上述的 $x_1, \dots, x_m$ 为此凸锥的一族**生成元**. 定义容易推广到无穷多个生成元的情形.

定理 14.9.4 (J. Fourier, T. Motzkin) 设 C 为 V 中的凸锥.

(i) C 是多面锥当且仅当存在 $m \in \mathbb{Z}_{\geq 0}$ 和 $x_1, \dots, x_m \in C$ 使得 $C = \sum_{i=1}^m \mathbb{R}_{\geq 0} x_i$.

(ii) 设 C 是多面锥, $\lambda_1, \dots, \lambda_m \in V^\vee \smallsetminus \{0\}$, 则 $C = \bigcap_{i=1}^m H_{\lambda_i}^{\geq 0}$ 等价于 $C^* = \sum_{i=1}^m \mathbb{R}_{\geq 0} \lambda_i$.

(iii) 若 C 是多面锥, 则 C^* 亦然.

证明 先处理 (i) 的 "当" 方向. 对 m 递归地论证. 由于 $m = 0$ 情形显然, 以下设 $m \geq 1$.

设有 $\kappa_1, \dots, \kappa_r \in V^\vee \smallsetminus \{0\}$ 使得

$$C' := \sum_{i=1}^{m-1} \mathbb{R}_{\geq 0} x_i = \bigcap_{h=1}^r H_{\kappa_h}^{\geq 0}.$$

命 $x = x_m$. 适当重排下标 h, 可以取 $0 \leq p \leq q \leq r$ 使得

$$\kappa_h(x) \begin{cases} = 0, & 1 \leq h \leq p, \\ > 0, & p < h \leq q, \\ < 0, & q < h \leq r. \end{cases}$$

对所有 $p<i\le q$ 和 $q<j\le r$, 命

$$\mu_{ij}:=\underbrace{\kappa_i(x)}_{>0}\kappa_j-\underbrace{\kappa_j(x)}_{<0}\kappa_i.$$

于是 $\mu_{ij}(x_1),\dots,\mu_{ij}(x_{m-1})\ge 0$ 而 $\mu_{ij}(x)=0$. 命

$$D:=H^{\ge0}_{\kappa_1}\cap\cdots\cap H^{\ge0}_{\kappa_q}\cap\bigcap_{i,j}H^{\ge0}_{\mu_{ij}};$$

当满足条件的 (i,j) 不存在时 $\bigcap_{i,j}$ 项规定为 V.

今将往证 $C=D$. 包含关系 $\subset$ 容易, 因为按构造有 $x_1,\dots,x_{m-1},x_m=x\in D$.

反之给定 $y\in D$, 我们寻求表法 $y=z+tx$, 其中 $t\ge0$ 而 $z\in C'$. 换言之, 目标是证明 $y-\mathbb{R}_{\ge0}x$ 与 C' 有交. 讨论如下.

h	$\kappa_h(y)$	$\kappa_h(x)$	$\kappa_h(y-tx)\ge0$ 的充要条件
$1\le h\le p$	≥0	$=0$	恒成立
$p<h\le q$	≥0	>0	$t\le a_h:=\dfrac{\kappa_h(y)}{\kappa_h(x)}$
$q<h\le r$		<0	$t\ge b_h:=\dfrac{\kappa_h(y)}{\kappa_h(x)}$

由于 $a_i\ge0$ 对所有 $p<i\le q$ 成立, 而且

$$p<i\le q<j\le r\overset{y\in D}{\Longrightarrow}\mu_{ij}(y)\ge0\implies a_i\ge b_j.$$

故以上讨论说明确实存在 $t\ge0$ 使得 $y-tx\in C'$.

接着证明 (ii). 先设 $C=\bigcap_{i=1}^m H^{\ge0}_{\lambda_i}$, 命 $\hat C:=\sum_{i=1}^m\mathbb{R}_{\ge0}\lambda_i$; 由上一步已知 $\hat C$ 是 $V^\vee$ 中的多面锥, 而且显然 $\hat C\subset C^*$ 而 $C=\hat C^*$. 引理 14.9.3 遂蕴涵 $\hat C=\hat C^{**}=C^*$.

其次设 C 是多面锥而 $C^*=\sum_{i=1}^m\mathbb{R}_{\ge0}\lambda_i$. 同理有 $C=C^{**}$, 然而易见 $C^{**}=\bigcap_{i=1}^m H^{\ge0}_{\lambda_i}$. 综上 (ii) 得证.

回头证明 (i) 的 "仅当" 的方向. 设多面锥 C 表作 $\bigcap_{j=1}^k H^{\ge0}_{\lambda_j}$, 则 (ii) 说明 $C^*=\sum_{j=1}^k\mathbb{R}_{\ge0}\lambda_j$. 对 $C^*\subset V^\vee$ 应用 (i) 的已知方向, 可将 C^* 表为 $\bigcap_{i=1}^m H^{\ge0}_{x_i}$ 的形式, 其中 $x_i\in V$. 对 C^* 应用 (ii) 遂有 $C=C^{**}=\sum_{i=1}^m\mathbb{R}_{\ge0}x_i$.

最后, (iii) 是 (i) 和 (ii) 的简单结论. □

针对定理 14.9.4 (i) 的论证给出了从生成元求多面锥表法的一套算法, 这也称为 **Fourier–Motzkin 消元法**.

练习 14.9.5 以下结论是 **Farkas 引理**的一种形式, 应用甚广. 设 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(\mathbb{R})$ 而 $\boldsymbol{b}\in\mathbb{R}^m$ (列向量). 证明以下两条陈述恰有一者成立:

(a) 存在 $\boldsymbol{x}\in\mathbb{R}^n$ 使得 $\boldsymbol{Ax}=\boldsymbol{b}$ 而 $\boldsymbol{x}\ge\boldsymbol{0}$;

(b) 存在 $\boldsymbol{z}\in\mathbb{R}^m$ 使得 ${}^{\mathrm{t}}\boldsymbol{z}\,\boldsymbol{A}\geq\boldsymbol{0}$ 而 ${}^{\mathrm{t}}\boldsymbol{z}\,\boldsymbol{b}<0$.

符号 $\boldsymbol{u}\geq\boldsymbol{0}$ 意谓向量 $\boldsymbol{u}$ 的每个分量皆非负.

提示 命 $C:=\{\boldsymbol{A}\boldsymbol{x}:\boldsymbol{x}\geq\boldsymbol{0}\}$, 这是由 $\boldsymbol{A}$ 的列向量生成的凸锥. 陈述 (a) 等价于 $\boldsymbol{b}\in C$, 而 (b) 等价于某个过 $\boldsymbol{0}$ 的超平面能区隔 C 和 $\boldsymbol{b}$. 按定理 14.9.4 (i) 表 C 为 $\bigcap_{i=1}^{k}\{\boldsymbol{y}:{}^{\mathrm{t}}\boldsymbol{z}_i\,\boldsymbol{y}\geq 0\}$. 若 (a) 不成立, 取 $\boldsymbol{z}=\boldsymbol{z}_i$ 使得 ${}^{\mathrm{t}}\boldsymbol{z}_i\,\boldsymbol{y}<0$.

基于定理 14.7.8 的多面锥情形和定理 14.9.4 (i), 容易联系 C 和 C^* 的面结构.

定理 14.9.6 *设 C 是 V 中的多面锥. 对 C 的所有面 F, 命*

$$\langle F\rangle:=F\text{ 生成的子空间},\quad F_C^*:=\langle F\rangle^{\perp}\cap C^*,$$

则 F_C^* 是 C^* 的面, 而且这给出相对于包含关系的反序双射

$$\begin{array}{ccc}\{C\text{ 的面}\} & \xrightarrow{1:1} & \{C^*\text{ 的面}\}\\ F & \longmapsto & F_C^*\\ \check{F}_{C^*}^* & \longmapsfrom & \check{F}\end{array}\qquad(\because\ C^{**}=C).$$

证明 以定理 14.9.4 (i) 将 F 写成 $\sum_{i=1}^{m}\mathbb{R}_{\geq 0}x_i$. 对所有 i 要求 $x_i\neq 0$, 则 $(\mathbb{R}x_i)^{\perp}$ 是 C^* 的支持超平面, 故 $C^*\cap(\mathbb{R}x_i)^{\perp}$ 是 C^* 的面, 而这 m 个面的交 F_C^* 仍是面. 反序性质 $G\subset F\implies F_C^*\subset G_C^*$ 显然成立.

接着说明 $F\mapsto F_C^*$ 为单射. 给定 C 的面 F, 选取 $\lambda\in C^*\smallsetminus\{0\}$ 使得 $F=C\cap H_\lambda$. 对于 C 的所有面 G, 我们有

$$G\subset F\iff G\subset H_\lambda\iff\lambda\in G_C^*.$$

特别地, $G_C^*=F_C^*$ 将蕴涵 $G\subset F$, 根据对称性也蕴涵 $F\subset G$, 故 $F=G$. 单性得证.

以 C^* 代 C 仍有上述结论. 从 C 的面集到其自身的映射 $\theta:F\mapsto(F_C^*)_{C^*}^*$ 是满足 $F\subset\theta(F)$ (应用 $\langle F_C^*\rangle\subset\langle F\rangle^{\perp}$) 的单射. 由于面集有限, 不难从上述性质推导 $\theta=\mathrm{id}$. 明所欲证. □

本章习题将继续探讨面的对偶性. 最后, 我们来引入严格凸性的概念.

引理 14.9.7 *设 C 为 V 中的多面锥, 则 C 有唯一的极小面, 它等于 C 所包含的极大向量子空间.*

证明 极小面总是极大仿射子空间 (推论 14.7.9), 它们作为面必然含 0 (基于引理 14.9.2), 因而是 V 的向量子空间; 又因为极大仿射子空间有共同的向量部分 W (命题 14.7.7), 所以极小面唯一并且等于 W.

如果 W' 是包含于 C 的向量空间, 则 W' 包含于 C 的某个极大仿射子空间 A; 从 $0\in W'\subset A$ 知 $A=0+W=W$, 从而有 $W'\subset W$. 这说明 W 是 C 的极大向量子空间. □

定义–命题 14.9.8 设 C 为 V 中的多面锥. 以下条件等价:

(i) $C \cap (-C) = \{0\}$,

(ii) C 不含 V 的非零向量子空间,

(iii) 0 是 C 作为多面体的唯一顶点,

(iv) 存在非零的 $\lambda \in V^\vee$ 使得 $H_\lambda \cap C = \{0\}$ 而且 $C \subset H_\lambda^{\geq 0}$.

满足上述任一条件的多面锥称为**严格凸**的.

证明 (i) $\Longrightarrow$ (ii). 若 W_0 是包含于 C 的向量子空间, 则 $W_0 = W_0 \cap (-W_0) \subset C \cap (-C) = \{0\}$.

(ii) $\Longrightarrow$ (iii). 引理 14.9.7 已说明极小面即 C 的极大向量子空间, 但 C 的向量子空间只能是 $\{0\}$. 根据推论 14.7.9 对极小面和顶点的描述, 立见 0 是唯一顶点.

(iii) $\Longrightarrow$ (iv). 取从 C 截出顶点 0 的支持超平面.

(iv) $\Longrightarrow$ (i). 若 $v \in C \cap (-C)$, 则 $C \subset H_\lambda^{\geq 0}$ 导致 $\lambda(v) = 0$, 从而 $v \in H_\lambda \cap C = \{0\}$. 证毕. □

推论 14.9.9 设 C 为 V 中的多面锥, 则 C 严格凸当且仅当 $\langle C^* \rangle = V^\vee$.

证明 由定理 14.9.4 (iii) 已知 C^* 为多面锥, 而 C 非严格凸等价于存在 $x \neq 0$ 使得 $C \supset \mathbb{R}x$, 等价于存在这般 x 使得 $C^* \subset (\mathbb{R}x)^* = \{\lambda \in V^\vee : \lambda(x) = 0\}$; 既然 $0 \in C^*$, 后者又等价于 C^* 包含于某个超平面, 等价于 $\langle C^* \rangle \neq V^\vee$. □

顶点是多面体的端点 (定义 14.6.4, 推论 14.7.10). 扣除 $C = \{0\}$ 的平凡情形, 严格凸锥总有棱, 以下说明它们恰好是定义 14.6.7 所谓的端射线.

命题 14.9.10 对于 V 中的严格凸多面锥 C, 其棱无非是其中的端射线. 特别地, 端射线的个数有限.

证明 给定 C 的棱 ρ, 它必然包含唯一顶点 0, 故包含一条射线 $\mathbb{R}_{\geq 0}x$; 易见 $\rho \in \{\mathbb{R}_{\geq 0}x, \mathbb{R}x\}$, 然而 0 是端点故后一种情形可排除. 现在可比较引理 14.7.3 与端射线的定义, 得出 ρ 是端射线.

给定 C 的端射线 ρ. 取 $x \in \rho \smallsetminus \{0\}$, 以定理 14.7.8 (v) 取唯一的面 F 使得 $x \in \mathrm{r.int}(F)$. 必有 $\dim F = 1$, 否则容易建立以 x 为原点的坐标系并找出相异之 $a, b \in \mathrm{r.int}(F) \smallsetminus \rho$, 使得 $[a, b] \cap \rho = \{x\}$. □

定理 14.9.4 (i) 已经说明多面锥总有一族有限生成元 $x_1, \ldots, x_m$; 如果从中移除任何一个元素便不再生成 C, 则称之为无赘的. 无赘生成元当然存在, 而且可以重排或将每个 x_i 用 $\mathbb{R}_{>0}$ 伸缩. 以下对严格凸的情形说明 C 的端射线给出唯一的无赘生成元, 精确到重排和伸缩.

定理 14.9.11 设 C 是 V 中严格凸的多面锥, 则精确到顺序和伸缩, 存在 C 的唯一一族无赘生成元 $x_1,\dots,x_m$, 其元素生成的射线恰好是 C 的所有端射线.

证明 推论 14.9.9 说明 $\langle C^*\rangle = V^\vee$. 于是引理 14.7.5 (iii) 蕴涵多面体的无赘表法

$$C^* = H_{x_1}^{\geq 0} \cap \cdots \cap H_{x_m}^{\geq 0}$$

中的 $x_1,\dots,x_m \in V \simeq V^{\vee\vee}$ 是唯一确定的, 精确到重排和用 $\mathbb{R}_{>0}$ 伸缩. 对 C^* 应用定理 14.9.4 (ii) 即推得多面锥的无赘表法 $C = \sum_{i=1}^m \mathbb{R}_{\geq 0} x_i$ 唯一.

从练习 14.6.8 易得每个端射线都是某个 $\mathbb{R}_{\geq 0}x_i$. 以下说明每个 $\mathbb{R}_{\geq 0}x_i$ 都是端射线, 亦即 C 的棱 (命题 14.9.10). 设 $m \geq 1$. 取 $\lambda \in \mathrm{r.int}(H_{x_i} \cap C^*)$, 则 $C \subset H_\lambda^{\geq 0}$, 而且 $\lambda(x_i) = 0$ 而对所有 $j \neq i$ 皆有 $\lambda(x_j) > 0$. 对于 $y = \sum_{j=1}^m t_j x_j$, 其中 $t_j \in \mathbb{R}_{\geq 0}$, 我们有

$$\lambda(y) = 0 \iff \forall j \neq i,\ t_j = 0,$$

因此 H_λ 是截出 $\mathbb{R}_{\geq 0}x_i$ 的支持超平面. 证毕. □

14.10 多胞体基本定理

本节第一部分介绍如何在严格凸多面锥和多面体之间相互过渡. 对实向量空间 V 的任意子集 A, 记 $\mathbb{R}_{\geq 0}A = \{ta : t \in \mathbb{R}_{\geq 0},\ a \in A\}$. 以下皆设 V 为有限维的.

引理 14.10.1 设 C 为 V 中的严格凸多面锥, 取 $\lambda \in V^\vee \smallsetminus \{0\}$ 使得 $C \subset H_\lambda^{\geq 0}$ 而 $H_\lambda \cap C = \{0\}$. 对所有 $a \in \mathbb{R}_{>0}$, 以仿射线性映射 $\lambda - a$ 定义

$$P := H_{\lambda - a} \cap C,$$

则 $\mathbb{R}_{\geq 0}P = C$ 而且 P 是多胞体.

证明 对所有 $x \in C \smallsetminus \{0\}$ 皆有 $\frac{a}{\lambda(x)}x \in P$, 故 $\mathbb{R}_{\geq 0}P = C$. 选定 C 的一族非零生成元 $x_1,\dots,x_m$ (定理 14.9.4). 任意 $x \in C$ 皆能表作 $x = \sum_{i=1}^m t_i x_i$, 其中 $t_i \geq 0$, 而

$$x \in P \iff \sum_{i=1}^m t_i \underbrace{\lambda(x_i)}_{>0} = a \implies \forall i,\ 0 \leq t_i \leq \frac{a}{\lambda(x_i)}.$$

这足以说明 P 在 V 中有界. □

直观来看, P 是严格凸多面锥 C 的一个截面, 如 (14.6.1). 接着介绍反向构造.

定义 14.10.2 设 P 为 V 中的多面体, $\mathrm{aff}(P) = V$. 以引理 14.7.5 (iii) 取其唯一的无赘表法

$$P = \bigcap_{i=1}^k H_{\alpha_i}^{\geq 0}, \quad \alpha_i = \lambda_i + c_i,\ \lambda_i \in V^\vee,\ c_i \in \mathbb{R}.$$

考虑向量空间 $V\oplus\mathbb{R}$, 定义 $h:V\oplus\mathbb{R}\to\mathbb{R}$ 为向最后一个坐标的投影, 定义 $V\oplus\mathbb{R}$ 中的多面锥

$$C(P):=\bigcap_{i=1}^{k}H_{\overline{\alpha_i}}^{\geqslant 0}\cap H_h^{\geqslant 0},$$

其中 $\overline{\alpha_i}\in(V\oplus\mathbb{R})^\vee$ 映 (v,a) 为 $\lambda_i(v)+c_ia$.

以下观察是容易的:

- $C(P)\cap H_{h-1}=P\times\{1\}$, 推而广之, 当 $a>0$ 时 $C(P)\cap H_{h-a}=aP\times\{a\}$;
- $C(P)\cap H_h=\mathrm{rec}(P)\times\{0\}$, 其中

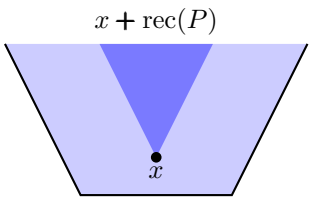
$$\mathrm{rec}(P):=\bigcap_{i=1}^{k}H_{\lambda_i}^{\geqslant 0};$$

- 当 $\mathrm{rec}(P)=\{0\}$ 时 $C(P)$ 严格凸, 因为此时 H_h 是截出顶点 0 的支持超平面.

所以 $C(P)$ 可以看作从 P 到 $\mathrm{rec}(P)$ 的 "形变". 多面锥 $\mathrm{rec}(P)$ 称为 P 的**回收锥**.

练习 14.10.3 说明 $\mathrm{rec}(P)\subset V$ 有内在的刻画如下: 任选 $x\in P$, 则对所有 $v\in V$, 我们有 $v\in\mathrm{rec}(P)$ 当且仅当 $x+\mathbb{R}_{\geq 0}v\subset P$.

因此多面锥 $\mathrm{rec}(P)$ 记录了 P 的无穷远方向, 几何图像如下.



定理 14.10.4 (H. Minkowski) 设 E 为有限维实仿射空间, P 为其中的非空子集. 以下陈述等价:

(i) P 是多胞体;

(ii) P 是多面体, 而且 P 等于其顶点集的凸包;

(iii) P 是 E 中有限多个点的凸包.

证明 问题只和 P 在 E 中张成的仿射子空间相关, 不妨设 $E=\mathrm{aff}(P)$. 任选基点将 E 等同于实向量空间 V.

(i) $\Longrightarrow$ (ii). 多胞体 P 按定义自然是多面体, 记其端点集为 $\{x_1,\dots,x_m\}$, 其中 $m\geq 1$. 因为 P 有界, 由 $\mathrm{rec}(P)$ 的几何诠释 (练习 14.10.3) 知 $\mathrm{rec}(P)=\{0\}$, 从而 $C(P)$

严格凸, $C(P)\cap H_h=\{0\}$. 定理 14.9.11 遂说明 $C(P)$ 由端射线生成. 兹断言 $C(P)$ 的端射线是 $\mathbb{R}_{\geq 0}(x_i,1)$, 其中 $i=1,\dots,m$.

设 ρ 是 $C(P)$ 中的射线, 它不可能包含于 H_h, 故形如 $\mathbb{R}_{\geq 0}(x,1)$, 其中 $x\in P$ 唯一确定. 需要以下观察: 若有不成比例之 $a,b\in C(P)$, 使得

$$\exists t\in\mathbb{R},\quad 0<t<1,\ ta+(1-t)b\in\rho,$$

则将 a 和 b 伸缩后可进一步取到相异之 $a,b\in P\times\{1\}$ 使上式成立. 为了说明这点, 取 $s,u>0$ 使得 $sa,ub\in P\times\{1\}$. 我们寻求和 $ta+(1-t)b$ 成比例的凸组合 $t'sa+(1-t')ub$, 亦即寻求

$$\frac{t's}{t}=\frac{(1-t')u}{1-t},\quad 0<t'<1$$

的解 t'. 上式整理成 $t'\left(\frac{s}{t}+\frac{u}{1-t}\right)=\frac{u}{1-t}$, 故确实有解.

基于前述观察, 将端射线的定义 14.6.7 和端点的定义 14.6.4 相对照, 立见 ρ 是端射线等价于 ρ 包含某个 $(x_i,1)$. 断言得证.

于是 P 的元素皆能写作 $\sum_{i=1}^m t_ix_i$, 其中 $t_i\in\mathbb{R}_{\geq 0}$ 而 $\sum_{i=1}^m t_i=1$, 这便说明 P 是其顶点集的凸包.

(ii) $\Longrightarrow$ (iii). 这是因为 P 的顶点个数有限.

(iii) $\Longrightarrow$ (i). 设 P 是 $x_1,\dots,x_m\in E$ 的凸包. 考虑 $V\oplus\mathbb{R}$ 中的凸锥 $C:=\sum_{i=1}^m\mathbb{R}_{\geq 0}(x_i,1)$. 定理 14.9.4 (i) 说明 C 是多面锥. 因此 $C\cap(V\times\{1\})$ 给出仿射空间 $V\times\{1\}$ 中的多面体.

若将 V 以 $v\mapsto(v,1)$ 等同于 $V\times\{1\}$, 则 $C\cap(V\times\{1\})$ 无非是凸包 P, 故 P 是多面体. 有限多个点的凸包是有界子集, 因此 P 还是多胞体. 证毕. □

定理 14.10.4 有时也称为多胞体基本定理, 它说明多胞体有两种等价描述: 一是有限多个半空间交出的有界子集, 二是有限多个点的凸包. 对于更一般的多面体, 相应的刻画称为 Motzkin 定理, 详见本章习题.

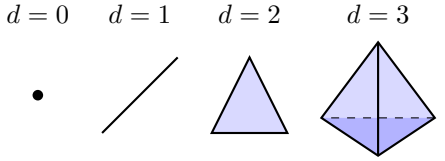
例 14.10.5 (单纯形) 一个 d 维多胞体如能表成 $d+1$ 个点的凸包, 则称之为 d-单纯形 ($d\in\mathbb{Z}_{\geq 0}$). 譬如标准 d-单纯形定义为

$$v_0:=(1,0,\dots,0),\ v_1:=(0,1,\dots,0),\ \dots,\ v_d:=(0,\dots,0,1)$$

的凸包, 它们也可以在 d 维仿射空间 $X_0+\cdots+X_d=1$ 中表成半空间

$$X_0\geq 0,\ \dots,\ X_d\geq 0$$

的交, 表法显然无赘. 低维的几何图像如下.



留意到 $d+1$ 个点 $y_0,\dots,y_d$ 的凸包 P 是 d 维的当且仅当它们是 $\mathrm{aff}(P)$ 的仿射基 (定义 14.1.7): “当” 的方向来自定义, “仅当” 方向则基于练习 14.1.9. 所以任两个 d-单纯形 P 和 P' 都通过仿射空间的某个同构 $\varphi:\mathrm{aff}(P)\xrightarrow{\sim}\mathrm{aff}(P')$ 相互等同: 选定生成元, 再要求 $\varphi(y_i)=y_i'$ 即可 $(0\leq i\leq d)$.

单纯形的面结构格外单纯. 考虑标准 d-单纯形, 对所有 $0\leq h\leq k$, 它的 h 维面一一对应到 $\{0,\dots,d\}$ 的 $h+1$ 元子集 $\{i_0,\dots,i_h\}$, 表为 $v_{i_0},\dots,v_{i_h}$ 的凸包, 或者等价地表成超平面 $X_j=0$ 在 $X_0+\cdots+X_d=1$ 中的交, 其中 j 遍历 $\{i_0,\dots,i_h\}$ 的补集. 特别地, 它的 h 维面都是 h-单纯形.

推论 14.10.6 *设 $\varphi:E\to E'$ 为仿射线性映射. 若 P 是 E 中的多胞体, 则 $\varphi(P)$ 是 E' 中的多胞体.*

证明 若 P 是 $x_1,\dots,x_m$ 的凸包, 则 $\varphi(E)$ 是 $\varphi(x_1),\dots,\varphi(x_m)$ 的凸包. □

特别地, 我们得出空间 $\mathbb{R}^n$ 中的多胞体向任何一个超平面的投影仍是多胞体; 一旦确定原多胞体的顶点, 它们投影后的凸包便是所求投影.

定义–命题 14.10.7 *设 P 和 Q 为有限维实向量空间 V 中的多胞体, 则*

$$P+Q:=\{x+y:x\in P,\ y\in Q\}$$

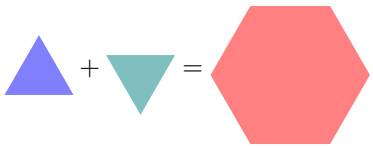
也是 V 中的多胞体, 称为 P 和 Q 的 **Minkowski 和**.

证明 考虑 $P\times Q\subset V\oplus V$, 从多胞体的定义直接看出 $P\times Q$ 是多胞体, 而 $P+Q$ 是它在加法映射 $V\oplus V\to V$ 之下的像. 加法映射是线性的, 对之应用推论 14.10.6. □

练习 14.10.8 基于上述结论, 进一步说明如果 P (或 Q) 是 $x_1,\dots,x_m$ (或 $y_1,\dots,y_n$) 的凸包, 则 $P+Q$ 是所有 x_i+y_j 的凸包.

提示〉 设 $x\in P$ 而 $y\in Q$. 基于端点的定义 14.6.4, 说明若 x (或 y) 非端点, 则 $x+y$ 非 $P+Q$ 的端点.

练习 14.10.9 试诠释关于 Minkowski 和的以下图像等式



并给出证明.

习题

1. 设 E 为有限维实内积空间 V 作用下的仿射空间, $n := \dim V$.

(i) 具有相同向量部分的仿射子空间称为平行的. 证明若超平面 $E_1, E_2 \subset E$ 平行, 则 E 上的刚体运动 $\rho_{E_1}\rho_{E_2}$ 是平移. 明确写下平移所对应的向量.

(ii) 承上题, 说明 E 上的所有平移都能写作 $\rho_{E_1}\rho_{E_2}$ 的形式.

(iii) 说明 E 上的所有刚体运动 R 都能写作若干个相对于超平面的镜射的合成. 估计至多需要几个镜射.

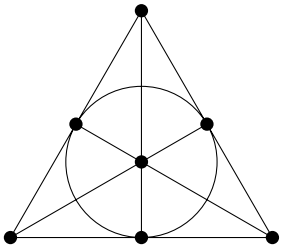
提示〉 取定基点 o, 将问题化约到 $R(o) = o$ 的情形, 从而化约为说明 V 的所有正交变换 $T \in \mathrm{End}(V)$ 都能写成 V 对有限多个 $n-1$ 维子空间的镜射的合成. 以正交变换的标准形定理将问题进一步化约到 $V = \mathbb{R}^2$ 的特例.

2. 在 $\mathbb{R}^2$ 上, 具体描述

(i) 对直线 $X + Y = 0$ 的镜射,

(ii) 对直线 $X + Y = 1$ 的镜射.

3. (Fano 平面) 证明域 $\mathbb{F}_2$ 上的射影平面 $\mathbb{P}^2(\mathbb{F}_2)$ 由 7 条线和 7 个点构成, 每条线有 3 个点, 过每点有 3 条线; 解释其间的相交关系如何由下图形象地描述 (图中的圆也视为一条“线”).

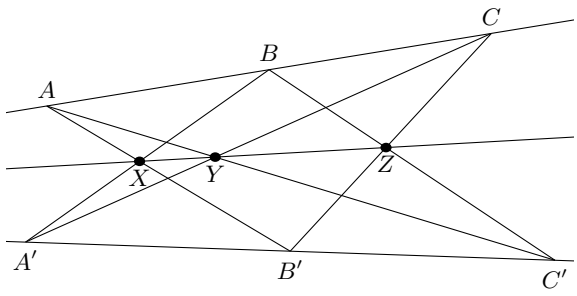


4. 以下的范例演示对偶原理 (命题 14.4.7) 和射影变换对平面几何学的应用. 试证:

▷ Pappus **定理** 设平面上相异三点 A, B, C 共线, 相异三点 A', B', C' 共另一线. 取 X (或 Y, Z) 为线的交点 $AB' \cap A'B$ (或 $AC' \cap A'C$, $BC' \cap B'C$), 则 X, Y, Z 三点共线.

▷ **对偶** Pappus **定理** 设平面上相异三线 α, β, γ 共点, 相异三线 α', β', γ' 共另一点. 取 ξ (或 η, ζ) 为点的连线 $(\alpha \cap \beta')(\alpha' \cap \beta)$ (或 $(\alpha \cap \gamma')(\alpha' \cap \gamma)$, $(\beta \cap \gamma')(\beta' \cap \gamma)$), 则 ξ, η, ζ 三线共点.

所谓 “平面” 意指实射影平面, AB (或 $\alpha \cap \beta$) 代表相异两点 A 和 B 的唯一连线 (或相异两线 α 和 β 的唯一交点).; 用仿射平面截取便有相应的仿射版本. 下面是 Pappus 定理在仿射平面上的图像.



提示〉 以对偶版本为例. 以适当的射影变换将所共两点搬运到无穷远, 化到一个相对容易的仿射版本.

5. 承上题, 说明 Pappus 定理及其对偶版本也适用于一般的域上的射影平面.

6. 设 V 为 F-向量空间, $\dim V = n+1 \in \mathbb{Z}_{\geq 1}$, 而 $E_0, \dots, E_k \subset V$ 为不含 0 的仿射子空间. 证明 $k < n$ 时 $\mathbb{P}(V)$ 无法被 $E_0, \dots, E_k$ 在 (14.4.1) 之下的像覆盖.

提示〉 不失一般性可设每个 E_i 皆为 n 维的, 由方程 $\lambda_i = c_i$ 截出 ($\lambda_i \in V^\vee \smallsetminus \{0\}$, $c_i \in F^\times$). 用练习 14.4.3 描述 E_i 的像, 然后讨论 $\dim \bigcap_i \ker(\lambda_i)$.

7. 设 V 为 n 维 F-向量空间, 而 $\mu_n(F) := \{t \in F^\times : t^n = 1\}$ 嵌入为 $\mathrm{GL}(V)$ 的子群, 定义 $\mathrm{PSL}(V) := \mathrm{SL}(V)/\mu_n(F)$. 包含同态 $\mathrm{SL}(V) \hookrightarrow \mathrm{GL}(V)$ 诱导同态 $\iota : \mathrm{PSL}(V) \to \mathrm{PGL}(V)$.

(i) 证明 $\mu_n(F)$ 等于群 $\mathrm{SL}(V)$ 的中心.

(ii) 说明 ι 总是单同态.

(iii) 举例说明 ι 可以非满, 尝试给出 ι 为同构的充要条件.

8. 将域 F 上的 $\mathbb{P}^1$ 等同于 $F \sqcup \{\infty\}$, 其中 $\infty = (0:1)$ 而 F 通过 $t \mapsto (1:t)$ 嵌入 $\mathbb{P}^1$. 验证 $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ 给出的射影变换 σ 限制在 F 上表为

$$t \mapsto \frac{at+b}{ct+d}$$

的形式; 为了得到 σ 的完整描述, 请说明如何在 $t = \infty$ 或 $ct+d = 0$ 时诠释右式, 这和数学分析中处理 ∞ 的习惯应当是相容的. 表作上述形式的映射也称为 $F \sqcup \{\infty\}$ 上的**线性分式变换**.

9. 设 V 为有限维非零 F-向量空间, 而 $x, y, z, w \in \mathbb{P}(V)$ 为共线相异四点. 将定义–命题 14.5.6 中的交比 $(x, y; z, w)$ 记为 λ. 证明:

$$\begin{aligned} (z, w; x, y) = (y, x; w, z) &= \lambda, \\ (y, x; z, w) &= \lambda^{-1}, \\ (x, z; y, w) &= 1 - \lambda. \end{aligned}$$

由此说明在四点的 24 种排列下, 交比仅取 6 个值 $\lambda^{\pm 1}, (1-\lambda)^{\pm 1}, \left(\frac{\lambda}{1-\lambda}\right)^{\pm 1}$. 相应地描述 $\mathfrak{S}_4$ 的 6 阶商群.

提示〉 保持交比的子群正是练习 11.1.23 的 $V \lhd \mathfrak{S}_4$. 商群 $\mathfrak{S}_4/V$ 同构于 $\mathfrak{S}_3$.

10. 证明若双射 $\sigma:\mathbb{P}^1\to\mathbb{P}^1$ 保持交比不变, 则它必来自 $\mathrm{PGL}(2,F)$.

提示〉 化简到 σ 保持 $(0,1,\infty)$ 的特例, 再应用 (14.5.2).

11. 证明 $\mathbb{R}^n$ 的子集有界当且仅当它包含于有限多个点的凸包. 因此有界性能够以仿射几何的语言来刻画.

12. (C. Carathéodory) 设 S 为 m 维实仿射空间 E 的子集 ($m\in\mathbb{Z}_{\geq 1}$). 按以下提示证明对每个 $x\in\mathrm{conv}(S)$ 都存在 $x_0,\dots,x_m\in S$, 使得 x 是它们的凸组合.

注意到 $x_0,\dots,x_m$ 可能随 x 而变. 例如矩形 $S=[0,1]\times[0,1]\subset\mathbb{R}^2$ 的元素总能表为四个顶点中某三个点的凸组合, 但 S 无法写成三个点的凸包.

提示〉 设 x 能表为 S 中 $q+1$ 个元素 ($q>m$) 的凸组合 $\sum_{i=0}^q t_ix_i$, 系数全正. 目标是表 x 为更短的凸组合. 任选 $o\in E$. 存在非平凡线性关系 $\sum_{i=1}^q s_i(x_i-x_0)=0$, 整理为 $\sum_{i=0}^q r_i(x_i-o)=0$, 其中 $r_i\in\mathbb{R}$ 不全为 0 而 $\sum_{i=0}^q r_i=0$. 命

$$\alpha:=\max_{0\leq i\leq q}\left\{-\frac{t_i}{r_i}:r_i>0\right\}<0<\min_{0\leq i\leq q}\left\{-\frac{t_i}{r_i}:r_i<0\right\}=:\beta,$$
$$A:=\{a\in\mathbb{R}:\forall\,0\leq i\leq q,\ r_i\neq 0\implies t_i+ar_i\geq 0\}=[\alpha,\beta].$$

存在 j 使得 $t_j+\alpha r_j=0$; 配合 $\sum_{i=0}^q r_i=0$ 推得

$$\begin{aligned}x&=\sum_{i=0}^q t_ix_i=o+\sum_{i=0}^q t_i(x_i-o)\\&=o+\sum_{i=0}^q(t_i+\alpha r_i)(x_i-o)\\&=\sum_{\substack{0\leq i\leq q\\ i\neq j}}(t_i+\alpha r_i)x_i.\end{aligned}$$

利用 $\alpha\in A$ 和 $\sum_{i=0}^q r_i=0$ 说明最后一式是凸组合.

13. 证明在有限维实仿射空间中, 紧子集的凸包依然紧. 提示〉 应用上一题的 Carathéodory 定理.

14. 设 E 为有限维实内积空间作用下的仿射空间, $\dim E\geq 2$. 试为 E 中的多面体严格地定义二面角的概念 (见 §11.11), 并说明刚体运动保持二面角.

15. (极值的局部性) 设 E 为实仿射空间, 若函数 $\alpha:E\to\mathbb{R}$ 对所有 $x,y\in E$ 和 $t\in[0,1]$ 皆满足

$$\alpha(tx+(1-t)y)\geq t\alpha(x)+(1-t)\alpha(y),$$

则称 α 为上凸函数; 若 $-\alpha$ 上凸则称 α 为下凸函数. 举例明之, 仿射线性函数既是上凸的也是下凸的.

现在设 E 有限维, C 为 E 的凸子集, $\alpha:E\to\mathbb{R}$ 为上 (或下) 凸函数. 证明若 $x\in C$, 而且存在开子集 $U\subset E$ 使得 $x\in U$ 而 $\alpha|_{U\cap C}$ 在 x 处取极大 (或极小) 值, 则 $\alpha|_C$ 在 x 处取极大 (或极小) 值. 直观地说明 C 的凸性是必要的. 这一观察在运筹学中起重大作用.

16. 设 A 为仿射空间 E 的仿射子空间, $x_0 \in A$. 说明存在仿射子空间 A' 使得 $A \cap A' = \{x_0\}$ 而且映射

$$\begin{aligned} A \times A' &\to E \\ (x, x') &\mapsto x + (x' - x_0) \end{aligned}$$

是双射; 此时称 A 与 A' 互补.

17. 设 P 为有限维实仿射空间 E 中的多面体, A 为其极大仿射子空间 (命题 14.7.7). 证明之前习题的 $A \times A' \to E$ 限制为双射 $A \times P' \to P$, 其中 $P' := P \cap A'$ 是有顶点的多面体.

提示〉 设 w 属于 A 的向量部分, $x = x_0 + w \in A$, 而 $x' \in A'$. 考虑 $y = x + (x' - x_0) = x' + w$. 于是 $y \in P \iff x' \in P'$. 应用 A 的极大性证明 P' 有顶点.

18. 设 C 为有限维实向量空间 V 中的多面锥, W 为 C 的极大向量子空间. 任取 V 的子空间 W' 使得 $V = W \oplus W'$, 定义 $C' := C \cap W'$. 证明向量加法诱导双射 $W \times C' \to C$, 而且 C' 是严格凸多面锥. 提示〉 和前一题类似.

19. 说明在多面体 P 的所有面连同 $\varnothing$ 对 $\subset$ 所构成的偏序集中, 任两个元素 F, G 都有上确界与下确界. 提示〉 上确界可以取为所有包含 $F \cup G$ 的面之交.

20. 证明如果 Q 是多面体 P 的面, $\dim P - \dim Q = 2$, 则满足 $Q \subsetneq F \subsetneq P$ 的面 F 恰有两个.

提示〉 取 P 在 $\mathrm{aff}(P)$ 中的无赘表法 $\bigcap_i H_{\alpha_i}^{\geq 0}$, 所求的 F 表为 $H_{\alpha_i} \cap P$. 适当地将问题化到 $\dim P = 2$ 的情形.

21. 在定理 14.9.4 的表述中, 方便起见, 选基等同 V 与 $\mathbb{R}^n$, 再按标准方式等同 $(\mathbb{R}^n)^\vee$ 与 $\mathbb{R}^n$. 说明若多面锥 C 的生成元 $x_1, \ldots, x_m$ 能取自 $\mathbb{Q}^n$, 则 $C = \bigcap_{i=1}^k H_{\lambda_i}^{\geq 0}$ 中的 $\lambda_1, \ldots, \lambda_k$ 能取自 $\mathbb{Q}^n \smallsetminus \{0\}$, 而且反之亦然.

22. 证明定理 14.9.6 中的双射满足 $\dim F + \dim F_C^* = \dim V$.

提示〉 先说明 $\dim C + \dim C_C^* = \dim V$, 再用双射的反序性质和定理 14.7.8 (iv).

23. 证明定义 14.10.2 中的多面锥 $C(P) \subset V \oplus \mathbb{R}$ 是 $\sum_{y \in P} \mathbb{R}_{\geq 0}(y, 1)$ 的闭包.

24. (T. Motzkin) 设 P 是有限维实向量空间 V 的子集. 证明以下陈述等价:

(i) P 是多面体;

(ii) 存在 V 中的多胞体 Q 和多面锥 C 使得 $P = Q + C := \{q + c : q \in Q,\ c \in C\}$.

提示〉 (i) $\Longrightarrow$ (ii). 构造 $V \oplus \mathbb{R}$ 中的多面锥 $C(P)$, 命 $h : V \oplus \mathbb{R} \to \mathbb{R}$ 为投影映射. 取 $C(P)$ 的生成元集 $X_0 \sqcup X_1$, 使得 $X_i \subset h^{-1}(i)$ 对 $i = 0, 1$ 成立. 取 Q 为 $h(X_1)$ 的凸包, 取 $C = \sum_{x \in X_0} \mathbb{R}_{\geq 0} x$.

(ii) $\Longrightarrow$ (i). 取 C 的生成元集 X, 记 Q 的顶点集为 $\mathrm{vert}(Q)$. 在 $V \oplus \mathbb{R}$ 中考虑由 $X \times \{0\}$ 和 $\mathrm{vert}(Q) \times \{1\}$ 生成的多面锥, 取它和 $V \times \{1\}$ 之交.

25. 设 $\varphi : E \to E'$ 是有限维实仿射空间之间的仿射线性映射, P 和 P' 分别是其中的多面体. 证明 $\varphi(P)$ 和 $\varphi^{-1}(P')$ 仍是多面体. 提示〉 对于 $\varphi(P)$ 可用上一题的结论.

26. 设 P, Q 为有限维实向量空间 V 中的多面体. 证明 $P + Q$ 仍是多面体.

27. 设 P 为有限维实仿射空间 E 中的多胞体, 记其顶点为 $x_1, \ldots, x_m$. 定义凸组合 $z_P := \sum_{i=1}^m \frac{1}{m} x_i$.

(i) 验证 $z_{T(P)} = T(z_P)$ 对所有 $T \in \mathrm{Aff}(E)$ 成立.

(ii) 取 $E = \mathbb{R}^3$ 而 P 为内接于某个球的正多面体, 见 §11.11. 证明 z_P 即球心. 作为推论, 若球心为原点, 则所有满足 $R(P) = P$ 的刚体运动 R 都属于 $\mathrm{O}(3)$.

28. (对偶多胞体) 设 P 是有限维实向量空间 V 中的多胞体, 定义 P 的对偶为

$$P^\triangle := \left\{\lambda \in V^\vee : \forall x \in P,\ \lambda(x) \leq 1\right\}.$$

(i) 说明 $P^\triangle$ 是多面体, 而且 $P^\triangle$ 包含 0 在 $V^\vee$ 中的某个开邻域.

提示〉 用定理 14.10.4 (iii) 对 P 的描述.

(ii) 验证 $(-P)^\triangle = -(P^\triangle)$.

(iii) 设 P 包含 0 在 V 中的某个开邻域, 证明 $P^\triangle$ 是多胞体.

(iv) 延续 (iii) 的前提, 证明 $P^{\triangle\triangle} = P$; 进一步证明此时有反序双射

$$\begin{aligned}\{P \text{ 的真面}\} &\to \{P^\triangle \text{ 的真面}\}\\ F &\mapsto F_P^\triangle := \{\lambda \in P^\triangle : \forall x \in F,\ \lambda(x) = 1\},\end{aligned}$$

其逆映射是 $\check{F} \mapsto \check{F}^\triangle_{P^\triangle}$, 而且 $\dim F + \dim F_P^\triangle = \dim V - 1$.

提示〉 尝试化到多面锥版本.

(v) 取 $V = \mathbb{R}^3$. 说明正立方体与正八面体互为对偶, 正四面体的对偶仍是正四面体.

29. 设 P 是 $\mathbb{R}^3$ 中的多胞体, 包含 $\mathbb{R}^3$ 的某个开子集. 分别记 P 的顶点, 边和面 (亦即台面) 的个数为 V, E 和 F. 按照下述图论方法证明 Euler 示性数公式 $V - E + F = 2$.

(i) 构造 $\mathbb{R}^3$ 中的多胞体 P' 使得 P' 仍包含某个开子集, P' 的顶点 (或边, 或面) 一一对应于 P 的面 (或边, 或顶点); P' 的顶点是对应面的重心, 而 P' 的边是对应边区隔的两个面的重心连线.

提示〉 参考 (11.11.1) 的图像.

(ii) 取半径足够小的球面 $\mathbb{S}$ 使得 $\mathbb{S} \subset P, P'$. 从球心出发作射线, 即可将 P (或 P') 的顶点和边投影到 $\mathbb{S}$ 上, 从而得到实现在 $\mathbb{S}$ 上的连通无向图 G (或 G'). 无向图的严格表述见诸第五章习题.

设 e 和 e' 分别是 P 和 P' 的边. 证明它们在 $\mathbb{S}$ 上的投影或者无交, 或者 e 和 e' 依照 (i) 的构造相互对应, 两者在后一情形有唯一交点 (非端点).

(iii) 任取 G 的生成树 H, 这是包含 G 的所有顶点而无环路的连通子图, 见第五章习题. 说明 H 有 $V - 1$ 条边. 对于剩下的每一条边, 取它们在 G' 中对应的边, 得到 G' 的子图 H'; 证明 H' 有 F 个顶点, 而且 H 与 H' 在 $\mathbb{S}$ 上无交.

(iv) 承上, 证明 H' 也是树, 从而是 G' 的生成树, 故有 $F - 1$ 条边.

提示〉 连通性缘于 H 无环路. 另一方面, 若 H' 含环路, 则环路分球面为两个连通成分, 但两者内部包含能由 H 相连接的顶点, 矛盾.

(v) 基于 P 的边或者来自 H, 或者来自 H' 这一事实, 证明 $V - E + F = 2$.

上述论证来自 K. G. C. von Staudt.

第十五章 向量空间的张量积

略而言之, 在选定的域 F 上, 张量积是将两个向量空间 V 和 W 的元素形式地 "配对", 借以构造新的向量空间 $V \otimes W$ 的手段; 这种构造不仅用于数学, 还在量子力学等应用领域中频繁现身. 记任意 $v \in V$ 和 $w \in W$ 配对的产物为 $v \otimes w \in V \otimes W$, 我们期望:

- ⋆ 配对给出双线性映射 $V \times W \xrightarrow{(v,w)\mapsto v\otimes w} V \otimes W$;
- ⋆ 空间 $V \otimes W$ 连同双线性映射 $(v, w) \mapsto v \otimes w$ 应当尽可能地 "泛", 详见下一段的说明;
- ⋆ 构造还应当是自然的: 一如迄今所见的所有基本操作, 它不应依赖基的选取.

定义张量积的关键是泛性质. 形式上, V 和 W 的张量积意谓一个向量空间 L_{univ} 连同双线性映射 $B_{\mathrm{univ}} : V \times W \to L_{\mathrm{univ}}$ (按先前符号即 $L_{\mathrm{univ}} = V \otimes W$ 和 $B_{\mathrm{univ}}(v, w) = v \otimes w$), 使得对所有向量空间 L 和双线性映射 $B : V \times W \to L$ 皆存在唯一的线性映射 $\varphi : L_{\mathrm{univ}} \to L$ 使得 $B = \varphi B_{\mathrm{univ}}$, 换言之使得下图交换:

$$\begin{array}{ccc} V \times W & \xrightarrow{B_{\mathrm{univ}}} & L_{\mathrm{univ}} \\ & \searrow^{B} & \downarrow{\scriptstyle \varphi} \\ & & L. \end{array}$$

这相当于说资料 $(L_{\mathrm{univ}}, B_{\mathrm{univ}})$ 在所有出自 $V \times W$ 的双线性映射中最 "泛", 其它 (L, B) 都唯一地由之推出. 张量积的严谨构造因此分成两步 (命题 15.1.1):

1. 说明确实存在满足上述性质的资料 $(L_{\mathrm{univ}}, B_{\mathrm{univ}})$;
2. 说明对于任两个满足上述性质的 $(L_{\mathrm{univ}}, B_{\mathrm{univ}})$ 和 $(L'_{\mathrm{univ}}, B'_{\mathrm{univ}})$, 存在唯一同构 $\varphi : L_{\mathrm{univ}} \xrightarrow{\sim} L'_{\mathrm{univ}}$ 使得 $B'_{\mathrm{univ}} = \varphi B_{\mathrm{univ}}$.

第一步构造用到商空间, 第二步则是形式的论证. 一旦知悉这两点, 便能合理地将张量积记为 $(V \otimes W, (v, w) \mapsto v \otimes w)$, 尽管它作为集合论的对象并非唯一, 却有**精确到一个唯一同构** φ 的唯一性.

在绝大多数场合, 张量积具体 “是什么” 无关宏旨; 真正起作用的是张量积具有的泛性质, 上述唯一性已然足够. 以泛性质刻画数学对象是一套标准技术, 也体现了范畴论的思路, 见附录 B. 读者应当借机培养相关的觉察力.

除了最后一节, 正文内容分成三部分. 第一部分的 §§15.1–15.2 给出向量积的定义, 然后推及多元张量积 $V_1 \otimes \cdots \otimes V_n$. 我们将建立关于张量积的下述基本性质:

- ▷ **函子性** 任何一族线性映射 $f_i : V_i \to W_i$ (其中 $1 \leq i \leq n$) 都诱导张量积之间的线性映射 $f_1 \otimes \cdots \otimes f_n : V_1 \otimes \cdots \otimes V_n \to W_1 \otimes \cdots \otimes W_n$, 由适当的交换图表刻画 (定义–命题 15.1.6);
- ▷ **结合约束** 存在自然同构 $V_1 \otimes (V_2 \otimes V_3) \simeq V_1 \otimes V_2 \otimes V_3 \simeq (V_1 \otimes V_2) \otimes V_3$ (命题 15.2.2);
- ▷ **幺约束** 将 F 本身视为向量空间, 则有自然同构 $F \otimes V \simeq V \simeq V \otimes F$ (命题 15.2.3);
- ▷ **交换约束** 存在自然同构 $c(V, W) : V \otimes W \xrightarrow{\sim} W \otimes V$ (命题 15.2.4).

之所以称之为结合约束而非结合律等, 是因为它们体现为自然同构而非严格等式. 关于函子性与自然同构的进一步解释将涉及范畴的概念.

我们也将证明张量积保持直和 (命题 15.2.5), 由此易知若 V 和 W 分别有基 $(v_i)_{i\in I}$ 和 $(w_j)_{j\in J}$, 则 $V \otimes W$ 有基 $(v_i \otimes w_j)_{(i,j)\in I\times J}$ (推论 15.2.6); 线性映射的张量积 $f_1 \otimes f_2$ 依此在矩阵层次体现为例 15.2.7 的 Kronecker 积 $\boldsymbol{A}_1 \otimes \boldsymbol{A}_2$.

张量积和对偶空间与 Hom 空间的关系是 §15.3 的主题. 命题 15.3.1 在 V 或 W 有限维时给出自然同构 $V^\vee \otimes W \xrightarrow{\sim} \mathrm{Hom}(V, W)$. 当 $V_1, \ldots, V_n$ 皆有限维时命题 15.3.5 给出自然同构 $V_1^\vee \otimes \cdots \otimes V_n^\vee \xrightarrow{\sim} (V_1 \otimes \cdots \otimes V_n)^\vee$.

定义 V 的张量幂

$$V^{\otimes m} := \underbrace{V \otimes \cdots \otimes V}_{m \text{ 份}}, \quad V^{\otimes 0} := F.$$

我们称 $V^{\otimes p} \otimes (V^\vee)^{\otimes q}$ 的元素为 (p, q)-型张量, 它们足以表达 V 的元素 ($(p, q) = (1, 0)$), 自同态 ($(p, q) = (1, 1)$) 和双线性形式 ($(p, q) = (0, 2)$) 等种种对象, 这是张量积在几何等领域中的初步面貌.

作为张量积的一则应用, §15.4 将对任意扩域 $E \supset F$ 给出将 F-向量空间拓展为 E-向量空间的一种手段, 它由泛性质刻画, 不依赖基的选取.

第二部分的 §§15.5–15.6 介绍何谓域 F 上的 “代数”, 又称 F-代数; 按定义, 这是搭建在 F-向量空间上的环结构, 使得乘法是双线性的. 除了熟悉的矩阵代数 $\mathrm{M}_{n\times n}(F)$, 定义–命题 15.5.6 将赋 $T(V) := \bigoplus_{m\geq 0} V^{\otimes m}$ 以 F-代数的结构, 称为 V 的张量代数. 通过对适当的理想取商, 定义 15.6.3 进一步定义 V 的对称代数 $\mathrm{Sym}(V) = \bigoplus_{m\geq 0} \mathrm{Sym}^m(V)$ 和外代数 $\bigwedge(V) = \bigoplus_{m\geq 0} \bigwedge^m(V)$, 其 m 次直和项分别通过泛性质联系于

▷ **对称 m 重线性映射** $C : \underbrace{V \times \cdots \times V}_{m\text{ 份}} \to M$, 条件是

$$C(\ldots, x, y, \ldots) = C(\ldots, y, x, \ldots), \quad x, y \in V;$$

▷ **交错 m 重线性映射** $C : \underbrace{V \times \cdots \times V}_{m\text{ 份}} \to M$, 条件是

$$C(\ldots, x, x, \ldots) = 0, \quad x \in V.$$

详见定义 15.6.1. 这两种代数是数学家的日常工具, 尤其常用于几何学. 取值在 $M = F$ 的交错 m 重线性映射称为 m 元交错形式; 我们在行列式理论中已经和交错形式打过照面, 外代数和行列式的关系 (推论 15.6.8) 自是题中之义.

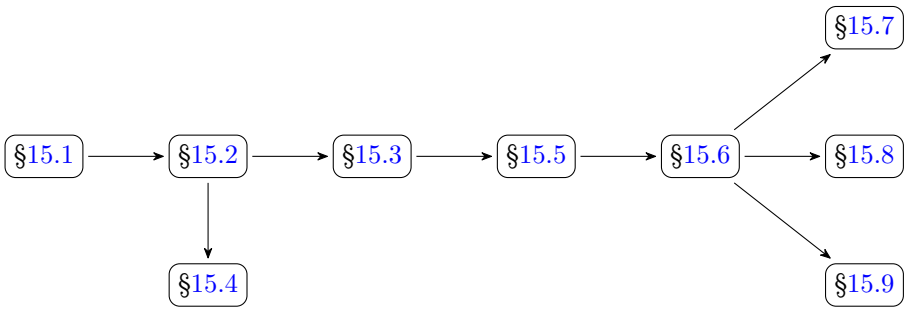
第三部分的 §§15.7–15.8 介绍外代数的两则运用, 技术性比较突出. 设 $\boldsymbol{A}$ 为 $2n \times 2n$ 交错矩阵, 定义 15.7.2 引入了称为 Pfaff 型的纯量 $\mathrm{Pf}(\boldsymbol{A})$, 定理 15.7.6 则断言 $\det \boldsymbol{A} = \mathrm{Pf}(\boldsymbol{A})^2$. Amitsur–Levitzki 定理 15.8.4 给出了代入所有 $n \times n$ 矩阵 $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_{2n}$ 皆成立的多项式恒等式. 两则结论都是化到特征为 0 的域上证明的, 涉及将矩阵元 "泛化" 的技巧.

最后的 §15.9 属于补充, 意在解释在 $\mathrm{char}(F) = 0$ 时如何将 $\mathrm{Sym}(V)$ (或 $\bigwedge(V)$) 嵌入为 $T(V)$ 的子空间而非商空间, 并将其乘法诠释为 $T(V)$ 中乘法的对称化 (或交错化). 这是某些文献采取的定义, 然而仅适用于特征零的域.

阅读提示

张量积的定义能推及任意环上的模或双模; 简单起见, 本书仅论向量空间情形, 感兴趣的读者宜另外参考 [10, §6.5] 或本章习题. 关于泛性质的部分可搭配附录 B 阅读.

阅读顺序



15.1 以泛性质定义张量积

向量空间 V 和 W 的张量积是将两者元素进行配对, 给出新的向量空间的一种标准方法. 所谓配对, 理解为映向某个向量空间 L 的双线性映射 $B : V \times W \to L$, 而张量积则是最 "泛" 的配对. 以下先从泛性质给出抽象而严格的表述, 随后介绍具体面向. 关于泛性质的深入讨论需要范畴的概念, 详见 §B.5.

选定任意域 F.

命题 15.1.1 设 V 和 W 为 F-向量空间.

(i) 存在 F-向量空间 L_{univ} 连同双线性映射 $B_{\mathrm{univ}} : V \times W \to L_{\mathrm{univ}}$, 使得对所有 F-向量空间 L 和双线性映射 $B : V \times W \to L$, 存在唯一的线性映射 $\varphi : L_{\mathrm{univ}} \to L$ 使下图交换:

$$\begin{array}{ccc} V \times W & \xrightarrow{B_{\mathrm{univ}}} & L_{\mathrm{univ}} \\ & \searrow^{B} & \downarrow{\scriptstyle \varphi} \\ & & L \end{array}$$

(ii) 若资料 $(L_{\mathrm{univ}}, B_{\mathrm{univ}})$ 和 $(L'_{\mathrm{univ}}, B'_{\mathrm{univ}})$ 皆具有 (i) 的性质, 则存在唯一的向量空间同构 $\varphi : L_{\mathrm{univ}} \xrightarrow{\sim} L'_{\mathrm{univ}}$ 使得下图交换:

$$\begin{array}{ccc} V \times W & \xrightarrow{B_{\mathrm{univ}}} & L_{\mathrm{univ}} \\ & \searrow^{B'_{\mathrm{univ}}} & \wr\downarrow{\scriptstyle \varphi} \\ & & L'_{\mathrm{univ}} \end{array}$$

证明 对于存在性 (i), 构造以集合 $V \times W$ 为基的 F-向量空间 $F^{\oplus(V\times W)}$, 其元素表作有限线性组合 $\sum_i c_i(v_i, w_i)$, 其中 $(v_i, w_i) \in V \times W$. 定义 $\mathcal{N}$ 为以下元素生成的子空间

$$\begin{gathered} (v + v', w) - (v, w) - (v', w), \\ (v, w + w') - (v, w) - (v, w'), \\ (tv, w) - t(v, w), \\ (v, tw) - t(v, w), \end{gathered}$$

其中 $v, v' \in V$, $w, w' \in W$ 而 $t \in F$. 定义商空间

$$L_{\mathrm{univ}} := F^{\oplus(V\times W)} \big/ \mathcal{N},$$

连同映射 $B_{\mathrm{univ}} : V \times W \to L_{\mathrm{univ}}$ 如下

$$B_{\mathrm{univ}}(v, w) = (v, w) + \mathcal{N}.$$

映射 B_{univ} 是双线性的, 原因是

$$\begin{aligned}
B_{\mathrm{univ}}(v+v',w)-B_{\mathrm{univ}}(v,w)-B_{\mathrm{univ}}(v',w)&=(v+v',w)-(v,w)-(v',w)+\mathcal{N},\\
B_{\mathrm{univ}}(v,w+w')-B_{\mathrm{univ}}(v,w)-B_{\mathrm{univ}}(v,w')&=(v,w+w')-(v,w)-(v,w')+\mathcal{N},\\
B_{\mathrm{univ}}(tv,w)-tB_{\mathrm{univ}}(v,w)&=(tv,w)-t(v,w)+\mathcal{N},\\
B_{\mathrm{univ}}(v,tw)-tB_{\mathrm{univ}}(v,w)&=(v,tw)-t(v,w)+\mathcal{N},
\end{aligned}$$

而 $\mathcal{N}$ 的定义正是为了使右侧的陪集全部等于 $\mathcal{N}$.

现在考虑任意双线性映射 $B:V\times W\to L$. 它确定唯一的线性映射 $\Phi:F^{\oplus(V\times W)}\to L$ 使得 $\Phi((v,w))=B(v,w)$. 然而

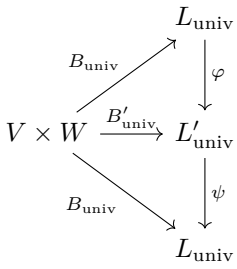
$$\Phi((v+v',w)-(v,w)-(v',w))=B(v+v',w)-B(v,w)-B(v',w)=0,$$

同理可见 Φ 映 $\mathcal{N}$ 的所有生成元为零, 故诱导唯一线性映射 $\varphi:L_{\mathrm{univ}}\to L$ 使得 $\varphi(x+\mathcal{N})=\Phi(x)$ 对所有 $x\in F^{\oplus(V\times W)}$ 成立; 然而在 $F^{\oplus(V\times W)}$ 的基 $V\times W$ 上检验可见 φ 的刻画等价于

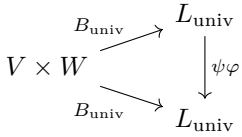
$$\varphi(B_{\mathrm{univ}}(v,w))=B(v,w),\quad (v,w)\in V\times W,$$

而此即所求的交换图表.

唯一性 (ii) 的证明是形式化的, 细说如下. 应用 (i) 两次, 可得唯一线性映射 φ 和 ψ, 分别使下图的两个三角部分皆交换:



我们希望证明 φ 和 ψ 互逆. 注意到上图的外框也是交换图表:



然而若将 $\psi\varphi$ 替换为 id, 图表当然也交换; 对资料 $(L_{\mathrm{univ}},B_{\mathrm{univ}})$ 和它本身应用 (i) 的唯一性部分, 遂见 $\psi\varphi=\mathrm{id}$. 基于对称性, 同理可见 $\varphi\psi=\mathrm{id}$. 综上得到所求的唯一同构. □

命题 15.1.1 (i) 的性质称为资料 $(L_{\mathrm{univ}},B_{\mathrm{univ}})$ 的**泛性质**, 而 (ii) 中的唯一同构是 (i) 的线性映射 φ 在 $(L,B)=(L'_{\mathrm{univ}},B'_{\mathrm{univ}})$ 时的特例.

定义 15.1.2 (向量空间的张量积) 满足上述泛性质的资料 $B_{\mathrm{univ}}: V\times W\to L_{\mathrm{univ}}$ 也记为

$$\begin{aligned} V\times W &\to V\otimes W(=L_{\mathrm{univ}})\\ (v,w)&\mapsto v\otimes w; \end{aligned}$$

不致混淆时也省去资料中的双线性映射, 直接记作 $V\otimes W$. 在需要强调域 F 的场合, 我们也采取 $V\underset{F}{\otimes}W$ 的记法.

此向量空间连同双线性映射 $(v,w)\mapsto v\otimes w$ 称为 V 和 W 的**张量积**.

泛性质因此表作双射

$$\begin{aligned} \mathrm{Hom}\,(V\otimes W,L)&\xrightarrow{\sim}\mathrm{Bil}(V,W;L)=\{\text{双线性映射 } V\times W\to L\}\\ \varphi&\mapsto \varphi B_{\mathrm{univ}}, \end{aligned}$$

此双射显然还是线性的. 这族同构也反过来确定 B_{univ}, 它是代入 $L=V\otimes W$ 和 $\varphi=\mathrm{id}_{V\otimes W}$ 的产物.

命题 15.1.1 证明中以商空间构造了张量积, 由之易见 $\{v\otimes w:(v,w)\in V\times W\}$ 是 $V\otimes W$ 的一族生成元. 读者也可以尝试按泛性质予以论证.

尽管 $V\otimes W$ 有标准的构造, 在多数场景中起作用的是且仅是它的泛性质, 而对于双线性映射的讨论而言, 这种 "精确到唯一同构" 的唯一性是自然且充分的. 今后我们不再关注张量积具体 "是什么", 而只关心它的泛性质.

上述定义容易推及多变元的情形, 进路依旧基于泛性质.

定义 15.1.3 (多重线性映射) 设 $V_1,\dots,V_n$ 和 M 为 F-向量空间 $(n\in\mathbb{Z}_{\geq 1})$. 若映射

$$C: V_1\times\cdots\times V_n\to M$$

对每个变元都是线性的, 则称 C 为 n 重 (或泛称为多重) 线性映射. 对于 $M=F$ 的特例, 对应的 n 重线性映射也称为 n 重 (或泛称为多重) 线性形式或 n 重线性型.

取 $n=1$ 回归线性映射, 取 $n=2$ 即双线性映射, 而此前讨论行列式时引入的 m 元交错形式则是一类特殊的 m 重线性形式. 以下性质及符号都是重要的.

练习 15.1.4 记多重线性映射 $V_1\times\cdots\times V_n\to M$ 构成的集合为 $\mathrm{Mul}(V_1,\dots,V_n;M)$. 验证它对运算

$$\begin{aligned} (C+C')(v_1,\dots,v_n)&=C(v_1,\dots,v_n)+C'(v_1,\dots,v_n),\\ (tC)(v_1,\dots,v_n)&=tC(v_1,\dots,v_n) \end{aligned}$$

成为向量空间.

定义–命题 15.1.5 (多元张量积) 给定 F-向量空间 $V_1,\dots,V_n$.

(i) 存在 n 重线性映射 $C_{\mathrm{univ}}: V_1\times\cdots\times V_n\to M_{\mathrm{univ}}$, 使得对所有 n 重线性映射 $C: V_1\times\cdots\times V_n\to M$ 皆存在唯一的线性映射 φ 使下图交换:

$$\begin{array}{ccc} V_1\times\cdots\times V_n & \xrightarrow{C_{\mathrm{univ}}} & M_{\mathrm{univ}} \\ & \searrow^{C} & \downarrow{\scriptstyle \varphi} \\ & & M \end{array}$$

(ii) 若资料 $(M_{\mathrm{univ}}, C_{\mathrm{univ}})$ 和 $(M'_{\mathrm{univ}}, C'_{\mathrm{univ}})$ 皆具有 (i) 的性质, 则存在唯一的向量空间同构 $\varphi: M_{\mathrm{univ}}\xrightarrow{\sim} M'_{\mathrm{univ}}$ 使下图交换:

$$\begin{array}{ccc} V_1\times\cdots\times V_n & \xrightarrow{C_{\mathrm{univ}}} & M_{\mathrm{univ}} \\ & \searrow^{C'_{\mathrm{univ}}} & \wr\downarrow{\scriptstyle \varphi} \\ & & M'_{\mathrm{univ}} \end{array}$$

基于 (ii) 揭示的唯一性, 满足 (i) 的泛性质的资料也记为

$$\begin{aligned} V_1\times\cdots\times V_n &\to V_1\otimes\cdots\otimes V_n (= M_{\mathrm{univ}}) \\ (v_1,\ldots,v_n) &\mapsto v_1\otimes\cdots\otimes v_n, \end{aligned}$$

称为 $V_1,\ldots,V_n$ 的**张量积**. 在需要强调域 F 的场合, 也采用 $V_1\underset{F}{\otimes}\cdots\underset{F}{\otimes}V_n$ 的记法.

证明 论证和命题 15.1.1 无异, 比如一种具体构造是

$$M_{\mathrm{univ}} = F^{\oplus(V_1\times\cdots\times V_n)}/\mathcal{N},$$

而子空间 $\mathcal{N}$ 由形如

$$\begin{gathered} (\ldots, v_i+v_i', \ldots) - (\ldots, v_i, \ldots) - (\ldots, v_i', \ldots), \\ (\ldots, tv_i, \ldots) - t(\ldots, v_i, \ldots) \end{gathered}$$

的元素生成, 其中 $1\le i\le n$, $v_i\in V_i$ 而 $t\in F$. 细节不赘. □

泛性质也表作

$$\begin{aligned} \mathrm{Hom}\,(V_1\otimes\cdots\otimes V_n, M) &\xrightarrow{\sim} \mathrm{Mul}(V_1,\ldots,V_n;M) \\ \varphi &\mapsto \varphi C_{\mathrm{univ}}, \end{aligned}$$

而 $C_{\mathrm{univ}}: V_1\times\cdots\times V_n\to V_1\otimes\cdots\otimes V_n$ 的多重线性性质体现为

$$\begin{aligned} \cdots\otimes(v_i+v_i')\otimes\cdots &= \cdots\otimes v_i\otimes\cdots + \cdots\otimes v_i'\otimes\cdots, \\ \cdots\otimes(tv_i)\otimes\cdots &= t\,(\cdots\otimes v_i\otimes\cdots)\,. \end{aligned}$$

这些记法中的第 i 个位置 $(1\le i\le n)$ 常称为张量积的第 i 个张量位或槽.

定义–命题 15.1.6 给定一族线性映射 $f_i : V_i \to W_i$, 其中 $i = 1, \dots, n$. 存在唯一的线性映射 $f_1 \otimes \cdots \otimes f_n$ 使下图交换

$$\begin{array}{ccc} V_1 \times \cdots \times V_n & \xrightarrow{(f_1,\dots,f_n)} & W_1 \times \cdots \times W_n \\ \downarrow & & \downarrow \\ V_1 \otimes \cdots \otimes V_n & \xrightarrow[f_1 \otimes \cdots \otimes f_n]{} & W_1 \otimes \cdots \otimes W_n, \end{array}$$

亦即 $(f_1 \otimes \cdots f_n)(v_1 \otimes \cdots \otimes v_n) = f_1(v_1) \otimes \cdots \otimes f_n(v_n)$ 恒成立. 今后称此线性映射为 $f_1, \dots, f_n$ 所诱导的映射.

证明 合成 $V_1 \times \cdots \times V_n \to W_1 \times \cdots \times W_n \to W_1 \otimes \cdots \otimes W_n$ 显然是多重线性的, 故存在唯一的 $f_1 \otimes \cdots \otimes f_n$ 使得图表交换. □

练习 15.1.7 说明若 $C \in \mathrm{Mul}(W_1, \dots, W_n; M)$ 对应到 $\varphi \in \mathrm{Hom}(W_1 \otimes \cdots \otimes W_n, M)$, 其中 M 是 F-向量空间, 则多重线性映射 $C \circ (f_1 \times \cdots \times f_n)$ 对应到 $\varphi \circ (f_1 \otimes \cdots \otimes f_n)$.

提示〉 考虑以下形式的交换图表即可:

$$\begin{array}{ccccc} V_1 \times \cdots \times V_n & \xrightarrow{f_1 \times \cdots \times f_n} & W_1 \times \cdots \times W_n & & \\ \downarrow & & \downarrow & \searrow^{C} & \\ V_1 \otimes \cdots \otimes V_n & \xrightarrow[f_1 \otimes \cdots \otimes f_n]{} & W_1 \otimes \cdots \otimes W_n & \xrightarrow[\varphi]{} & M. \end{array}$$

从交换图表的刻画确立诱导映射的性质是标准技巧. 作为补充, 请读者验证以下性质.

- ⋆ 诱导映射与合成兼容: 设有一族线性映射 $U_i \xrightarrow{g_i} V_i \xrightarrow{f_i} W_i$, 则

$$(f_1 \otimes \cdots \otimes f_n)(g_1 \otimes \cdots \otimes g_n) = f_1 g_1 \otimes \cdots f_n g_n.$$

- ⋆ 恒等诱导恒等: $\mathrm{id}_{V_1} \otimes \cdots \otimes \mathrm{id}_{V_n} = \mathrm{id}_{V_1 \otimes \cdots \otimes V_n}$.
- ⋆ 诱导映射与线性组合兼容: 给定 $1 \le i \le n$, $a, a' \in F$ 和 $f_i, f_i' : V_i \to W_i$, 则

$$\cdots \otimes (a f_i + a' f_i') \otimes \cdots = a(\cdots \otimes f_i \otimes \cdots) + a'(\cdots \otimes f_i' \otimes \cdots).$$

练习 15.1.8 说明若 $V_1, \dots, V_n$ 之中任一者为零空间, 则 $V_1 \otimes \cdots \otimes V_n$ 为零空间.

提示〉 从 $\mathrm{Mul}(V_1, \dots, V_n; M) = \{0\}$ (所有 M) 推导 $\mathrm{End}(V_1 \otimes \cdots \otimes V_n) = \{0\}$.

15.2 张量积的基本性质

多个向量空间可按不同的结合顺序取张量积, 例如 $V_1\otimes(V_2\otimes V_3)$ 和 $(V_1\otimes V_2)\otimes V_3$, 我们期望将其等同. 因为张量积是通过泛性质描述的, 谈论它们作为集合是否相等无甚意义, 关键在于将体现为严格等号的结合律放宽为同构, 这种同构称为结合约束. 先记录一则简单的引理.

引理 15.2.1 设 $V_1,\dots,V_n$ 和 M 为 F-向量空间, 而 $\varphi,\psi\in\mathrm{Hom}(V_1\otimes\cdots\otimes V_n,M)$. 若 $\varphi(v_1\otimes\cdots\otimes v_n)=\psi(v_1\otimes\cdots\otimes v_n)$ 对所有 $v_i\in V_i$ 成立 $(1\leq i\leq n)$, 则 $\varphi=\psi$.

证明 条件等价于 φ 和 ψ 对应同一个 n 重线性映射 $V_1\times\cdots\times V_n\to M$. □

命题 15.2.2 (结合约束) 设 V_1,V_2,V_3 为 F-向量空间, 则有同构

$$\begin{array}{ccccc} V_1\otimes(V_2\otimes V_3) & \xleftarrow{\sim} & V_1\otimes V_2\otimes V_3 & \xleftarrow{\sim} & (V_1\otimes V_2)\otimes V_3 \\ v_1\otimes(v_2\otimes v_3) & \longmapsto & v_1\otimes v_2\otimes v_3 & \longmapsto & (v_1\otimes v_2)\otimes v_3. \end{array}$$

证明 基于问题的对称性, 确立 $(V_1\otimes V_2)\otimes V_3\simeq V_1\otimes V_2\otimes V_3$ 即足. 关键在于说明存在线性映射如下

$$\begin{array}{ccccc} V_1\otimes(V_2\otimes V_3) & \xrightarrow{\alpha} & V_1\otimes V_2\otimes V_3 & \xrightarrow{\beta} & V_1\otimes(V_2\otimes V_3) \\ v_1\otimes(v_2\otimes v_3) & \longmapsto & v_1\otimes v_2\otimes v_3 & \longmapsto & v_1\otimes(v_2\otimes v_3). \end{array}$$

为了定义 α, 考虑映射

$$\begin{aligned} B:V_1\times(V_2\times V_3)&\to V_1\otimes V_2\otimes V_3\\ (v_1,(v_2,v_3))&\mapsto v_1\otimes v_2\otimes v_3. \end{aligned}$$

当 $v_1\in V_1$ 固定, $B(v_1,\cdot):V_2\times V_3\to V_1\otimes V_2\otimes V_3$ 是双线性映射, 按泛性质对应到

$$\varphi_{v_1}\in\mathrm{Hom}(V_2\otimes V_3,V_1\otimes V_2\otimes V_3).$$

于是有映射

$$\begin{aligned} B':V_1\times(V_2\otimes V_3)&\to V_1\otimes V_2\otimes V_3\\ (v_1,w)&\mapsto\varphi_{v_1}(w)\\ (v_1,v_2\otimes v_3)&\mapsto v_1\otimes(v_2\otimes v_3) \end{aligned}$$

因为 φ_{v_1} 是线性映射, 上式对 $V_2\otimes V_3$ 变元是线性的; 此外 $B(v_1,\cdot)\in\mathrm{Bil}(V_2,V_3;V_1\otimes V_2\otimes V_3)$ 线性地依赖于 v_1, 故 φ_{v_1} 亦然, 从而 B' 对 V_1 变元也是线性的. 综上, B' 是双线性映射, 故进一步给出线性映射

$$\begin{aligned} \alpha:V_1\otimes(V_2\otimes V_3)&\to V_1\otimes V_2\otimes V_3\\ v_1\otimes(v_2\otimes v_3)&\mapsto v_1\otimes v_2\otimes v_3, \end{aligned}$$

至于 β 的构造更简单: 易见映射

$$\begin{aligned} C : V_1 \times V_2 \times V_3 &\to V_1 \otimes (V_2 \otimes V_3) \\ (v_1, v_2, v_3) &\mapsto v_1 \otimes (v_2 \otimes v_3) \end{aligned}$$

是三重线性的, 故有相应的线性映射 $\beta : V_1 \otimes V_2 \otimes V_3 \to V_1 \otimes (V_2 \otimes V_3)$.

考虑合成映射 $\alpha\beta$. 它映所有 $v_1 \otimes v_2 \otimes v_3$ 为 $v_1 \otimes v_2 \otimes v_3$, 故应用引理 15.2.1 可得 $\alpha\beta = \mathrm{id}$.

其次, $\beta\alpha$ 映所有 $v_1 \otimes (v_2 \otimes v_3)$ 为 $v_1 \otimes (v_2 \otimes v_3)$. 先固定 v_1, 应用引理 15.2.1 可得 $\beta\alpha$ 映所有 $v_1 \otimes w$ 为 $v_1 \otimes w$, 其中 $w \in V_2 \otimes V_3$. 再次应用引理 15.2.1 可得 $\beta\alpha = \mathrm{id}$. 于是 α 和 β 互逆.

以上论证基于泛性质. 基于商空间的构造也可以给出相对冗长的证明. □

同理可对一般的 n 元张量积按类似方法来插入或移除括号.

以下说明 F 作为 1 维向量空间充当了张量积的某种幺元. 幺元的性质在此同样体现为同构, 称为幺约束. 证明理路类似命题 15.2.2.

命题 15.2.3 (幺约束) *设 V 为 F-向量空间, 则有同构*

$$\begin{array}{ccccc} F \otimes V & \xrightarrow{\sim} & V & \xleftarrow{\sim} & V \otimes F \\ t \otimes v & \longmapsto & tv & \longmapsto & v \otimes t. \end{array}$$

证明 处理 $F \otimes V \simeq V$ 即足. 首先定义 $m : F \times V \to V$ 为 $m(t, v) = tv$. 这是双线性映射, 故诱导线性映射 $\lambda_V : F \otimes V \to V$ 使得 $\lambda_V(t \otimes v) = tv$.

另一方面, 定义 $\nu_V : V \to F \otimes V$, 映 v 为 $1 \otimes v$. 易见 ν_V 是线性映射 (譬如有 $1 \otimes tv = t \otimes v$). 由这些描述立见 $\lambda_V \nu_V = \mathrm{id}$. 另一方面 $\nu_V \lambda_V : F \otimes V \to F \otimes V$ 映 $t \otimes v$ 为 $t \otimes v$, 故引理 15.2.1 说明 $\nu_V \lambda_V = \mathrm{id}$. □

张量积还具有体现为同构的交换性.

命题 15.2.4 (交换约束) *设 V 和 W 为 F-向量空间, 则有同构*

$$\begin{aligned} c(V, W) : V \otimes W &\xrightarrow{\sim} W \otimes V \\ v \otimes w &\mapsto w \otimes v, \end{aligned}$$

它们满足 $c(W, V)c(V, W) = \mathrm{id}_{V \otimes W}$.

证明 定义映射 $B : V \times W \to W \otimes V$ 为 $B(v, w) = w \otimes v$. 它和 $W \times V \to W \otimes V$ 仅差一个换位, 仍是双线性的. 于是从 B 得到断言中的线性映射.

同理可定义 $c(W, V) : W \otimes V \to V \otimes W$, 刻画为 $c(W, V)(w \otimes v) = v \otimes w$.

鉴于引理 15.2.1, 这些描述立刻蕴涵 $c(W, V)c(V, W) = \mathrm{id}_{V \otimes W}$. □

以下继续运用泛性质来说明张量积保直和.

命题 15.2.5 设 V 和 W 为 F-向量空间, V 带有直和分解 $V=\bigoplus_{i\in I}V_i$, 则有同构

$$\begin{aligned} V\otimes W &\xrightarrow{\sim} \bigoplus_{i\in I}(V_i\otimes W)\\ \left(\sum_{i\in I}v_i\right)\otimes w &\longmapsto (v_i\otimes w)_{i\in I}\end{aligned}$$

上式默认至多只有有限个 $v_i\in V_i$ 非零, 并且将 V_i 视同 V 的子空间.

类似地 $W\otimes V\xrightarrow{\sim}\bigoplus_{i\in I}W\otimes V_i$; 对多元张量积亦同.

证明 处理第一个同构即可. 设 L 为任意 F-向量空间. 指定双线性映射 $V\times W\to L$ 相当于对每个 i 指定双线性映射 $V_i\times W\to L$, 这给出

$$\begin{aligned}\mathrm{Bil}(V,W;L)&\xrightarrow{\sim}\prod_{i\in I}\mathrm{Bil}(V_i,W;L)\\ B&\mapsto (B|_{V_i\times W})_{i\in I}\,.\end{aligned}$$

对每个 i 记张量积自带的双线性映射为 $B_{i,\mathrm{univ}}:V_i\times W\to V_i\otimes W$, 则

$$\begin{aligned}\mathrm{Hom}(V_i\otimes W,L)&\xrightarrow{\sim}\mathrm{Bil}(V_i,W;L)\\ \varphi_i&\mapsto \varphi_i B_{i,\mathrm{univ}}.\end{aligned}$$

另一方面, 将 $V_i\otimes W$ 视同 $\bigoplus_{j\in I}(V_j\otimes W)$ 的子空间, 则有

$$\begin{aligned}\mathrm{Hom}\left(\bigoplus_{i\in I}(V_i\otimes W),L\right)&\xrightarrow{\sim}\prod_{i\in I}\mathrm{Hom}(V_i\otimes W,L)\\ \varphi&\mapsto(\varphi|_{V_i\otimes W})_{i\in I}\,.\end{aligned}$$

综上遂有同构

$$\mathrm{Hom}\left(\bigoplus_{i\in I}(V_i\otimes W),L\right)\xrightarrow{\sim}\mathrm{Bil}(V,W;L);\tag{15.2.1}$$

其具体映法如下. 给定左边的元素 φ, 命 $\varphi_i:=\varphi|_{V_i\otimes W}$, 则 φ 的像 $B:V\times W\to L$ 在 $V_i\times W$ 上的限制是

$$(v_i,w)\mapsto\varphi_iB_{i,\mathrm{univ}}(v_i,w)=\varphi_i(v_i\otimes w).$$

现在定义双线性映射 $B'_{\mathrm{univ}}:V\times W\to\bigoplus_{i\in I}(V_i\otimes W)$, 使得它映 $(\sum_i v_i,w)$ 为 $(v_i\otimes w)_i$. 同构 (15.2.1) 遂等于 $\varphi\mapsto\varphi B'_{\mathrm{univ}}$. 至此, 我们得知 $\bigoplus_{i\in I}(V_i\otimes W)$ 连同 B'_{univ} 满足张量积的泛性质, 故存在唯一的同构连同交换图表

$$\begin{array}{ccc} V\times W & \xrightarrow{(v,w)\mapsto v\otimes w} & V\otimes W\\ & \searrow^{B'_{\mathrm{univ}}} & \downarrow\wr\\ & & \bigoplus_{i\in I}(V_i\otimes W)\end{array}$$

观察定义可见同构必映 $(\sum_i v_i)\otimes w$ 为 $(v_i\otimes w)_i$. 证毕. □

推论 15.2.6 设 V 有基 $(v_i)_{i\in I}$ 而 W 有基 $(w_j)_{j\in J}$, 则 $(v_i\otimes w_j)_{(i,j)\in I\times J}$ 是 $V\otimes W$ 的基. 多元张量积的基也有类似描述.

证明 基于命题 15.2.5,

$$\begin{aligned} V\otimes W &= \left(\bigoplus_{i\in I} Fv_i\right)\otimes\left(\bigoplus_{j\in J} Fw_j\right) \\ &\simeq \bigoplus_{i\in I}\left(Fv_i\otimes\bigoplus_{j\in J} Fw_j\right) \simeq \bigoplus_{i\in I}\bigoplus_{j\in J} Fv_i\otimes Fw_j \\ &= \bigoplus_{(i,j)\in I\times J} Fv_i\otimes Fw_j. \end{aligned}$$

对所有 (i,j), 命题 15.2.3 蕴涵

$$Fv_i\otimes Fw_j\simeq F\otimes F\simeq F;$$

故左边是 1 维空间, 它作为 $V\otimes W$ 的子空间由 $v_i\otimes w_j$ 生成. 综上, $V\otimes W=\bigoplus_{i,j}F(v_i\otimes w_j)$. □

作为推论, 如果 $V_1,\dots,V_n$ 皆是有限维向量空间, 则

$$\dim(V_1\otimes\cdots\otimes V_n)=\prod_{i=1}^{n}\dim V_i.$$

例 15.2.7 (矩阵的 Kronecker 积) 考虑线性映射 $f:V\to V'$ 和 $g:W\to W'$. 设 V (或 V') 有基 $v_1,\dots,v_n$ (或 $v'_1,\dots,v'_{n'}$), 而 W (或 W') 有基 $w_1,\dots,w_m$ (或 $w'_1,\dots,w'_{m'}$). 命 $e_{ij}:=v_i\otimes w_j$ 和 $e'_{ij}:=v'_i\otimes w'_j$. 于是

$$\begin{aligned} &e_{11},e_{12},\dots,e_{1m},e_{21},e_{22},\dots,e_{nm} \text{ 是 } V\otimes W \text{ 的基,}\\ &e'_{11},e'_{12},\dots,e'_{1m'},e'_{21},e'_{22},\dots,e'_{n'm'} \text{ 是 } V'\otimes W' \text{ 的基.}\end{aligned}$$

设 f 对应到矩阵 $\boldsymbol{A}\in\mathrm{M}_{n'\times n}(F)$ 而 g 对应到矩阵 $\boldsymbol{B}\in\mathrm{M}_{m'\times m}(F)$. 我们有

$$\begin{aligned}(f\otimes g)(e_{ij}) &= \left(\sum_{k=1}^{n'} a_{ki}v'_k\right)\otimes\left(\sum_{\ell=1}^{m'} b_{\ell j}w'_\ell\right)\\ &= a_{1i}b_{1j}e'_{11}+a_{1i}b_{2j}e'_{12}+\cdots+a_{2i}b_{1j}e'_{21}+a_{2i}b_{2j}e'_{22}+\cdots.\end{aligned}$$

相对于上述有序基, $f\otimes g:V\otimes W\to V'\otimes W'$ 遂对应到 $m'n'\times mn$ 分块矩阵

$$\boldsymbol{A}\otimes\boldsymbol{B}:=\left(\begin{array}{c|c|c} a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B}\\ \hline \vdots & \ddots & \vdots\\ \hline a_{n'1}\boldsymbol{B} & \cdots & a_{n'n}\boldsymbol{B}\end{array}\right);$$

矩阵 $\boldsymbol{A}\otimes\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的 Kronecker 积. 关于 $f\otimes g$ 的许多性质都能翻译为 Kronecker 积在矩阵层次的性质, 譬如

$$\begin{aligned}(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})&=\boldsymbol{AC}\otimes\boldsymbol{BD},\\ \boldsymbol{A}\otimes(\boldsymbol{P}+\boldsymbol{Q})&=\boldsymbol{A}\otimes\boldsymbol{P}+\boldsymbol{A}\otimes\boldsymbol{Q},\\ (\boldsymbol{P}+\boldsymbol{Q})\otimes\boldsymbol{B}&=\boldsymbol{P}\otimes\boldsymbol{B}+\boldsymbol{Q}\otimes\boldsymbol{B}.\end{aligned}$$

练习 15.2.8 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 皆为上三角 (或下三角, 对角) 矩阵, 证明 Kronecker 积 $\boldsymbol{A}\otimes\boldsymbol{B}$ 也是上三角 (或下三角, 对角) 矩阵, 并描述其对角元.

命题 15.2.9 *给定一族线性映射 $f_i: V_i\to W_i$, 其中 $1\leq i\leq n$.*

(i) *若每个 f_i 皆满, 则 $f_1\otimes\cdots\otimes f_n$ 满.*

(ii) *若每个 f_i 皆单, 则 $f_1\otimes\cdots\otimes f_n$ 单.*

证明 对于 (i), 先将 $W_1\otimes\cdots\otimes W_n$ 的元素 y 表为有限线性组合 $y=\sum_i c_i w_{i1}\otimes\cdots\otimes w_{in}$, 其中 $w_{ij}\in W_j$. 对每个 (i,j) 取 $v_{ij}\in V_j$ 使得 $f_j(v_{ij})=w_{ij}$, 则 $\sum_i c_i v_{i1}\otimes\cdots\otimes v_{in}$ 被映为 y.

对于 (ii), 不失一般性[1]可设 $n=2$, 而 f_1 是某个子空间 $V_1'\subset V_1$ 的包含映射, $f_2=\mathrm{id}_{V_2}$. 任取直和分解 $V_1=V_1'\oplus V_1''$. 于是命题 15.2.5 导致

$$V_1\otimes V_2\xrightarrow{\sim}(V_1'\otimes V_2)\oplus(V_1''\otimes V_2),$$

而 $f_1\otimes f_2: V_1'\otimes V_2\to V_1\otimes V_2$ 对应的无非是直和项的包含, 故单. □

本章习题将继续讨论张量积与线性映射的余核, 商以及核的关系.

15.3 张量积与对偶空间

本节的第一个目标是以张量积和对偶空间的概念诠释 Hom 空间. 回忆到 F-向量空间 V 的对偶空间记为 $V^\vee:=\mathrm{Hom}(V,F)$; 在 $V^\vee$ 和 V 之间有典范配对

$$\begin{aligned}V^\vee\times V&\to F\\ (\check{v},v)&\mapsto\langle\check{v},v\rangle:=\check{v}(v).\end{aligned}$$

设 $T\in\mathrm{Hom}(V,W)$. 若 $\mathrm{im}(T)$ 有限维, 则称 T 为有限秩的.

命题 15.3.1 *设 V 和 W 为 F-向量空间. 存在线性映射如下*

$$\begin{aligned}V^\vee\otimes W&\xrightarrow{\Theta_{V,W}}\mathrm{Hom}(V,W)\\ \lambda\otimes w&\longmapsto[v\mapsto\langle\lambda,v\rangle\,w].\end{aligned}$$

[1] 详细的化约方法请读者补全. 针对特例的论证也能直接推及一般情形, 只是符号将显得复杂.

映射 $\Theta_{V,W}$ 总是单射, 而 $\mathrm{im}(\Theta_{V,W}) = \{T \in \mathrm{Hom}(V,W) : \text{有限秩}\}$. 当 V 或 W 有限维时, $\Theta_{V,W}$ 是同构.

证明 给定 λ 和 w, 总能定义线性映射 $[v \mapsto \langle \lambda, v\rangle\, w]$. 这给出双线性映射 $V^\vee \times W \to \mathrm{Hom}(V,W)$, 从而按泛性质确定所需的 $\Theta_{V,W}$.

为了说明单性, 考虑 $\sum_{i=1}^n \lambda_i \otimes w_i \in \ker(\Theta_{V,W})$. 适当整理后, 不妨设 $w_1, \dots, w_n \in V^\vee$ 线性无关. 对所有 $v \in V$ 皆有

$$\langle \lambda_1, v\rangle\, w_1 + \cdots + \langle \lambda_n, v\rangle\, w_n = 0;$$

线性无关条件蕴涵 $\langle \lambda_1, v\rangle = \cdots = \langle \lambda_n, v\rangle = 0$. 既然 v 是任意的, $\lambda_1 = \cdots = \lambda_n = 0$.

现在来刻画 $\Theta_{V,W}$ 的像. 按定义, $\Theta_{V,W}(\sum_{i=1}^n \lambda_i \otimes w_i)$ 的像包含于 $\sum_{i=1}^n Fw_i$, 因而是有限秩的.

反之设 $T \in \mathrm{Hom}(V,W)$ 有限秩, 取 $\mathrm{im}(T)$ 的基 $w_1, \dots, w_n$. 对所有 $v \in V$, 将 $T(v)$ 唯一地展开为 $c_1 w_1 + \cdots + c_n w_n$, 则每个系数 $c_i \in F$ 对 v 都是线性的, 故可表为 $c_i = \langle \lambda_i, v\rangle$, 其中的 $\lambda_i \in V^\vee$ 与 v 无关. 由此验证 $T = \Theta_{V,W}(\sum_i \lambda_i \otimes w_i)$.

综上可知 $\mathrm{im}(\Theta_{V,W})$ 由所有有限秩的 $T \in \mathrm{Hom}(V,W)$ 构成. 当 V 或 W 有限维时, 所有 T 都是有限秩的. 证毕. □

例 15.3.2 对于 $V = W$ 有限维的情形, 任选 V 的基 $v_1, \dots, v_n$ 及其对偶基 $\check{v}_1, \dots, \check{v}_n \in V^\vee$, 则

$$\Theta_{V,V}\left(\sum_{i=1}^n \check{v}_i \otimes v_i\right) = \mathrm{id}_V;$$

为此, 仅须留意到左侧映每个 v_j 为 $\sum_{i=1}^n \langle \check{v}_i, v_j\rangle\, v_j = v_j$.

对 $V = F^n$ 而 $W = F^m$ 的特例, 将 $V^\vee$ (或 W) 等同于 $\mathrm{M}_{1\times n}(F)$ (或 $\mathrm{M}_{m\times 1}(F)$), 不难验证

$$\boldsymbol{\lambda} \in \mathrm{M}_{1\times n}(F),\ \boldsymbol{w} \in \mathrm{M}_{m\times 1}(F) \implies \Theta_{F^n,F^m}(\boldsymbol{\lambda} \otimes \boldsymbol{w}) = \boldsymbol{w}\boldsymbol{\lambda} \quad \text{(矩阵乘法)}.$$

本节的第二个目标是介绍张量积的缩并操作, 常用于微分几何与物理学.

定义 15.3.3 (缩并) 给定 F-向量空间 V, 通过张量积的泛性质对应到双线性形式

$$\begin{aligned} V^\vee \times V &\to F \\ (\check{v}, v) &\mapsto \langle \check{v}, v\rangle \end{aligned}$$

的线性映射 $V^\vee \otimes V \to F$ 称为缩并. 具体地说, $\sum_i \check{v}_i \otimes v_i \in V^\vee \otimes V$ 的缩并是 $\sum_i \langle \check{v}_i, v_i\rangle$.

练习 15.3.4 在 V 有限维的情形, 验证缩并运算等于 $V^\vee \otimes V \xrightarrow{\Theta_{V,V}} \mathrm{End}(V) \xrightarrow{\mathrm{Tr}} F$ 的合成. 〈提示〉 归结为证 $\langle \lambda, \cdot\rangle\, w$ 的迹是 $\langle \lambda, w\rangle$.

本节的第三个目标是研究张量积的对偶空间.

命题 15.3.5 设 $V_1,\dots,V_n$ 为 F-向量空间, 则有典范的线性映射

$$\Psi_{V_1,\dots,V_n}:V_1^\vee\otimes\cdots\otimes V_n^\vee\to(V_1\otimes\cdots\otimes V_n)^\vee,$$

它通过张量积的泛性质对应到以下 n 重线性映射:

$$\begin{aligned}V_1^\vee\times\cdots\times V_n^\vee&\to(V_1\otimes\cdots\otimes V_n)^\vee\\(\check v_1,\dots,\check v_n)&\mapsto\left[v_1\otimes\ldots\otimes v_n\mapsto\prod_{i=1}^n\langle\check v_i,v_i\rangle\right].\end{aligned}$$

当每个 V_i 皆为有限维时, $\Psi_{V_1,\dots,V_n}$ 为同构.

证明 首务是观察到当 $\check v_1,\dots,\check v_n$ 固定, $v_1\otimes\ldots\otimes v_n\mapsto\prod_{i=1}^n\langle\check v_i,v_i\rangle$ 是良定义的: 它对应到 n 重线性映射

$$\begin{aligned}V_1\times\cdots\times V_n&\to F\\(v_1,\dots,v_n)&\mapsto\prod_{i=1}^n\langle\check v_i,v_i\rangle.\end{aligned}$$

此外, 当 $(v_1,\dots,v_n)$ 固定, $\prod_{i=1}^n\langle\check v_i,v_i\rangle$ 对每个 $\check v_i$ 都是线性的, 因此这给出 n 重线性映射 $V_1^\vee\times\cdots\times V_n^\vee\to(V_1\otimes\cdots\otimes V_n)^\vee$, 相应地有 $\Psi_{V_1,\dots,V_n}$.

以下设每个 V_i 都是有限维的. 先证 $\Psi_{V_1,\dots,V_n}$ 单. 对每个 $1\le h\le n$ 为 V_h 取基 $(v_{h,i})_{i\in I_h}$, 相应地有 $V_h^\vee$ 的对偶基 $(\check v_{h,i})_{i\in I_h}$. 推论 15.2.6 说明

$$\check v_{1,i_1}\otimes\cdots\otimes\check v_{n,i_n},\quad\forall h,\ i_h\in I_h$$

构成 $V_1^\vee\otimes\cdots\otimes V_n^\vee$ 的基.

将每个 $\lambda\in V_1^\vee\otimes\cdots\otimes V_n^\vee$ 以上述基展开, 取 $\Psi_{V_1,\dots,V_n}(\lambda)$ 在 $v_{i,i_1}\otimes\cdots\otimes v_{n,i_n}$ 处的值便是 $\check v_{i,i_1}\otimes\cdots\otimes\check v_{n,i_n}$ 在 λ 中的系数. 由此可见 $\Psi_{V_1,\dots,V_n}$ 单.

最后, 比较维数即由 $\Psi_{V_1,\dots,V_n}$ 的单性推得同构. □

本节的最后一个目标是引入关于张量幂的符号.

约定 15.3.6 (张量幂) 给定 F-向量空间 V 和 $n\in\mathbb{Z}_{\ge1}$, 对应的张量幂定义为

$$V^{\otimes n}:=V\otimes\cdots\otimes V\quad(n\text{ 份});$$

受幺约束的启发, 另规定 $V^{\otimes0}:=F$.

结合约束与幺约束导致 $(V^{\otimes a})^{\otimes b}\simeq V^{\otimes ab}$ 和 $V^{\otimes a}\otimes V^{\otimes b}\simeq V^{\otimes(a+b)}$, 所以记法确实是方便的. 此外交换约束还导致 $(V\otimes W)^{\otimes a}\simeq V^{\otimes a}\otimes W^{\otimes a}$.

基于张量幂的符号, 命题 15.3.1 和命题 15.3.5 表明从有限维 F-向量空间 V 出发, 反复应用 $\otimes$, Hom 和取对偶空间这三种操作, 所得产物都同构于形如

$$V^{\otimes p}\otimes (V^\vee)^{\otimes q},\quad p,q\in\mathbb{Z}_{\geq 0}$$

的向量空间; 上述向量空间的元素也称为 V 上的 (p,q)**-型张量**.

例 15.3.7 空间 V 的元素是 $(1,0)$-型张量, $V^\vee$ 的元素是 $(0,1)$-型张量, $\mathrm{End}(V)$ 的元素等同于 $(1,1)$-型张量. 双线性形式 $V\times V\to F$ 等同于 $\mathrm{Hom}(V\otimes V,F)\simeq (V^\vee)^{\otimes 2}$ 的元素 (基于泛性质和命题 15.3.5), 因而等同于 $(0,2)$-型张量.

此外, 当 $p,q\geq 1$ 时, 交换约束和缩并运算给出线性映射

$$V^{\otimes p}\otimes (V^\vee)^{\otimes q}\xrightarrow[\sim]{\text{分别取出第一个张量位}}(V^\vee\otimes V)\otimes V^{\otimes(p-1)}\otimes (V^\vee)^{\otimes(q-1)}\xrightarrow{\text{缩并}}V^{\otimes(p-1)}\otimes (V^\vee)^{\otimes(q-1)},$$

其合成由下式刻画 (回忆引理 15.2.1):

$$v_1\otimes\cdots\otimes v_p\otimes \check{v}_1\otimes\cdots\otimes\check{v}_q\mapsto \langle \check{v}_1,v_1\rangle\, v_2\otimes\cdots\otimes v_p\otimes\check{v}_2\otimes\cdots\otimes\check{v}_q.$$

取其他张量位作缩并同样可行, 但定义时必须清楚指定.

张量积还有许多值得一提的基础性质, 而且这些性质往往能从域推广到一般的环上. 篇幅所限, 不再深究. 有需求的读者可以参考 [10, §6.5], 或者尝试本章的相关习题.

15.4 应用: 域的变换

设 E 为域 F 的扩域. 对于任何 E-向量空间 $\widetilde{V}$, 总是能将纯量乘法从 E 限制到 F 上, 得到 F-向量空间. 本节关注的是反向操作: 给定 F-向量空间 V, 如何自然地, 或曰典范地将 V 扩展为 E-向量空间?

对于 $V=F^n$, 自然的取法是 E^n. 然而将 n 维向量空间 V 等同于 F^n 相当于选定 V 的有序基, 所以这种构造依赖基的选法. 与此相反, 我们所期待的构造应当是由泛性质刻画的, 它应当具有精确到唯一同构的唯一性, 不涉及基的选取, 而在 $V=F^n$ 的情形它应当回到 E^n.

线索来自张量积. 以下将 E 视为 F-向量空间, 纯量乘法来自域的乘法. 为了强调域 F 的角色, 将 F-向量空间的张量积记为 $\underset{F}{\otimes}$. 乘法映射 $(x,y)\mapsto xy$ 显然是 F-双线性映射 $E\times E\to E$, 从而对应到 F-线性映射

$$\begin{aligned}\mu: E\underset{F}{\otimes}E&\to E\\ x\otimes y&\mapsto xy.\end{aligned}$$

对于 F-向量空间 V, 我们有线性映射

$$E \underset{F}{\otimes} (E \underset{F}{\otimes} V) \xrightarrow{\sim} (E \underset{F}{\otimes} E) \underset{F}{\otimes} V \xrightarrow{\mu\otimes \mathrm{id}_V} E \underset{F}{\otimes} V$$

$$x \otimes (y \otimes v) \longmapsto (x \otimes y) \otimes v \longmapsto xy \otimes v,$$

相应的双线性映射 $E \times (E \underset{F}{\otimes} V) \to E \underset{F}{\otimes} V$ 其具体映法是

$$\left(x, \sum_i y_i \otimes v_i\right) \mapsto \sum_i xy_i \otimes v_i, \quad x, y_i \in E,\ v_i \in V.$$

引理 15.4.1 以上运算使得 $E \underset{F}{\otimes} V$ 成为 E-向量空间.

证明 纯量乘法所需的结合律, 分配律等性质即刻归结为域 E 本身的性质, 以及 $E \times V \to E \underset{F}{\otimes} V$ 的双线性性质. □

现在我们拥有从 F-向量空间构造 E-向量空间的一种手段. 同样构造也可以在线性映射的层次操作.

引理 15.4.2 设 $f \in \mathrm{Hom}_F(V, V')$, 则 $\mathrm{id}_E \otimes f \in \mathrm{Hom}_E\left(E \underset{F}{\otimes} V, E \underset{F}{\otimes} V'\right)$. 此处 Hom_E 和 Hom_F 分别代表作为 E-向量空间和 F-向量空间的 Hom.

证明 已知 $\mathrm{id}_E \otimes f$ 是 F-线性的. 设 $t \in E$, 它对 $E \underset{F}{\otimes} V$ (或 $E \underset{F}{\otimes} V'$) 的纯量乘法给出自同态 m_t (或 m'_t). 对所有 $x \in E$ 和 $v \in V$, 映射 $m'_t(\mathrm{id}_E \otimes f)$ 和 $(\mathrm{id}_E \otimes f)m_t$ 都映 $x \otimes v$ 为 $tx \otimes f(v)$, 因此 $m'_t(\mathrm{id}_E \otimes f) = (\mathrm{id}_E \otimes f)m_t$ (引理 15.2.1). □

例 15.4.3 若 V 有基 $(v_i)_{i\in I}$, 则 $E \underset{F}{\otimes} V = \bigoplus_{i\in I}(E \underset{F}{\otimes} Fv_i)$; 然而 $E \underset{F}{\otimes} Fv_i \simeq E \underset{F}{\otimes} F \simeq E$, 这是 1 维 E-向量空间, 由 $\tilde{v}_i := 1 \otimes v_i$ 生成. 于是 $(\tilde{v}_i)_{i\in I}$ 给出 $E \underset{F}{\otimes} V$ 作为 E-向量空间的基. 特别地, 若 $V = F^n$, 则 $E \underset{F}{\otimes} V$ 通过 $\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n$ 自然地等同于 E^n.

在线性映射层次, 设 $f \in \mathrm{Hom}_F(V, V')$, 而 V (或 V') 有选定的基 $v_1, \ldots, v_n$ (或 $v'_1, \ldots, v'_m$), 依此得到 f 的矩阵 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$, 并且对 $E \underset{F}{\otimes} V$ 和 $E \underset{F}{\otimes} V'$ 取对应的基, 则

$$(\mathrm{id}_E \otimes f)(\tilde{v}_j) = \sum_i a_{ij}(1 \otimes v'_i) = \sum_i a_{ij}\tilde{v}'_i$$

表明 E-线性映射 $\mathrm{id}_E \otimes f$ 的矩阵无非是将 $\boldsymbol{A}$ 置于 $\mathrm{M}_{m\times n}(E)$ 中考量. 这是我们再熟悉不过的操作.

构造 $V \mapsto E \underset{F}{\otimes} V$ 的精髓在于它的泛性质. 为此, 必须将 $E \underset{F}{\otimes} V$ 连同典范映射

$$\begin{aligned} \iota : V &\to E \underset{F}{\otimes} V \\ v &\mapsto 1 \otimes v \end{aligned} \tag{15.4.1}$$

一并纳入考量.

引理 15.4.4 上述映射 ι 是 F-线性的单射.

证明 将 ι 表作 $V \xrightarrow{\sim} F \underset{F}{\otimes} V \to E \underset{F}{\otimes} V$ 的合成, 第一段是命题 15.2.3 的应用, 映 v 为 $1 \otimes v$, 第二段则由包含映射 $F \to E$ 诱导. 两段都是 F-线性的, 故 ι 亦然. 此外, 命题 15.2.9 (ii) 确保第二段是单射, 故 ι 亦然. □

因此, $E \underset{F}{\otimes} V$ 确实是 V 的某种扩展. 对于特例 $V = F^n$, 因为 $\iota(\boldsymbol{e}_i) = 1 \otimes \boldsymbol{e}_i = \tilde{\boldsymbol{e}}_i$, 映射 ι 等同于包含映射 $F^n \to E^n$.

命题 15.4.5 设 $\widetilde{V}$ 为 E-向量空间, 则对任何 F-线性映射 $f : V \to \widetilde{V}$ 都存在唯一的 E-线性映射 $\tilde{f} : E \underset{F}{\otimes} V \to \widetilde{V}$ 使得 $f = \tilde{f}\iota$. 换言之, 有同构

$$\begin{aligned} \mathrm{Hom}_E\left(E \underset{F}{\otimes} V, \widetilde{V}\right) &\xrightarrow{\sim} \mathrm{Hom}_F\left(V, \widetilde{V}\right) \\ \tilde{f} &\longmapsto \tilde{f}\iota. \end{aligned}$$

证明 单性的缘由是对所有 $x \in E$ 和 $v \in V$ 易见

$$\tilde{f}(x \otimes v) = x\tilde{f}(\iota(v)).$$

至于满性, 设 $f \in \mathrm{Hom}_F(V, \widetilde{V})$, 则有 F-双线性映射

$$\begin{aligned} E \times V &\to \widetilde{V} \\ (x, v) &\mapsto xf(v). \end{aligned}$$

它对应的 F-线性映射 $\tilde{f} : E \underset{F}{\otimes} V \to \widetilde{V}$ 满足 $\tilde{f}(x \otimes v) = xf(v)$. 然而从 $\tilde{f}(xy \otimes v) = xyf(v) = x(\tilde{f}(y \otimes v))$ 可说明 $\tilde{f}$ 实则为 E-线性的. 它满足 $\tilde{f}\iota = f$. □

命题 15.4.5 是泛性质的又一实例. 何以故? 考虑形如 $(\widetilde{V}, f)$ 的资料, 其中 $\widetilde{V}$ 是 E-向量空间, 而 $f : V \to \widetilde{V}$ 是 F-线性映射. 设资料 $(\widetilde{V}_{\mathrm{univ}}, f_{\mathrm{univ}})$ 满足以下性质: 对所有 $(\widetilde{V}, f)$, 存在唯一的 E-线性映射 $\tilde{f}$ 使下图交换

$$\begin{array}{ccc} V & \xrightarrow{f_{\mathrm{univ}}} & \widetilde{V}_{\mathrm{univ}} \\ & \searrow^{f} & \downarrow{\scriptstyle \tilde{f}} \\ & & \widetilde{V} \end{array}$$

则这种资料精确到唯一同构是唯一的. 一切和刻画张量积的命题 15.1.1 异曲同工, 问题的实质仅是 $(\widetilde{V}_{\mathrm{univ}}, f_{\mathrm{univ}})$ 的存在性, 命题 15.4.5 表明可取之为 $(E \underset{F}{\otimes} V, \iota)$.

15.5 域上的代数

设 F 为域. 所谓 F-代数, 可以理解为兼具 F-向量空间结构的环.

定义 15.5.1 (域上的代数) 设 A 为环, 同时又有 F-向量空间的结构, 使得环的加法等于向量空间加法, 而环的乘法 $A\times A\to A$ 是双线性映射, 则称 A 为 F-代数.

若 F-代数 A 作为环是交换的, 则称之为交换 F-代数.

设 A 和 A' 为 F-代数. 如果环同态 $f:A\to A'$ 同时也是 F-线性映射, 则称之为 F-代数的同态. 按照类似方法定义 F-代数的同构.

因此 F-代数 A 的乘法也可以等价地表为线性映射 $A\otimes A\to A$. 若 F-代数 A 的子环 A_0 同时也是子空间, 则称 A_0 为 A 的**子代数**.

例 15.5.2 域 F 上的多项式环 $F[X_1,X_2,\dots]$ 是 F-代数; 全体 n 元对称多项式构成 $F[X_1,\dots,X_n]$ 的子代数.

例 15.5.3 对所有 F-向量空间 V, 环 $\mathrm{End}(V)$ 自然是 F-代数; 类似地, $\mathrm{M}_{n\times n}(F)$ 是 F-代数. 此外, 全体上三角 (或下三角, 对角) 矩阵构成 $\mathrm{M}_{n\times n}(F)$ 的子代数.

例 15.5.4 设域 E 是 F 的扩域, 视之为 F-向量空间, 则 E 也是 F-代数: 乘法的双线性 $x(ty)=t(xy)=(tx)y$ 不过是乘法交换律的反映, 其中 $t\in F$ 而 $x,y\in E$.

推而广之, 回忆到环 A 的中心定义为子环 $Z(A):=\{z\in A:\forall a\in A,\ za=az\}$. 若有环同态 $\varphi:F\to Z(A)$, 则可赋 A 以 F-向量空间的结构

$$ta:=\varphi(t)a,\quad t\in F,\ a\in A,$$

条件 $\mathrm{im}(\varphi)\subset Z(A)$ 确保 A 的乘法满足 $x(ty)=t(xy)=(tx)y$, 因而成为 F-代数.

练习 15.5.5 说明反向的构造: 给定 F-代数 A 和 $t\in F$, 定义 $\varphi(t)=t\cdot 1_A$ 将给出环同态 $\varphi:F\to Z(A)$. 按此说明将环 A 升级为 F-代数和指定环同态 $F\to Z(A)$ 是一回事.

对于本章的研究, 最重要的例子是 F-向量空间所对应的张量代数. 沿用约定 15.3.6.

定义–命题 15.5.6 (张量代数) 对 F-向量空间 V, 定义 $T(V):=\bigoplus_{n\geq 0}V^{\otimes n}$. 在 $T(V)$ 上存在 F-代数结构, 使得 $V^{\otimes a}\cdot V^{\otimes b}\subset V^{\otimes(a+b)}$ 对所有 $a,b\in\mathbb{Z}_{\geq 0}$ 成立, 而且对应的线性映射 $V^{\otimes a}\otimes V^{\otimes b}\to V^{\otimes(a+b)}$ 刻画为

$$(v_1\otimes\cdots\otimes v_a)\otimes(v_1'\otimes\cdots\otimes v_b')\mapsto v_1\otimes\cdots\otimes v_a\otimes v_1'\otimes\cdots\otimes v_b'.$$

它的幺元是 $1\in F=V^{\otimes 0}$. 这一结构称为 V 的张量代数.

证明 对于给定的 $a,b\in\mathbb{Z}_{\geq 0}$, 所示的线性映射 $V^{\otimes a}\otimes V^{\otimes b}\to V^{\otimes(a+b)}$ 无非是结合约束 (命题 15.2.2) 或其多元版本的直接应用, 它按照直和唯一地延拓为 $T(V)\otimes T(V)\to T(V)$, 而乘法所需的结合律等诸般性质显然. □

原空间 V 嵌入为 $T(V)$ 的子空间 $V^{\otimes 1}$. 依照定义, 若 $V=\{0\}$ 则 $T(V)=F$. 留意到 $T(V)$ 一般不交换, 这是因为一般而言 $x\otimes y\neq y\otimes x$.

今后将 $T(V)$ 的乘法记为 $\otimes$.

命题 15.5.7 设 $(v_i)_{i\in I}$ 是 V 的基, 则所有 $v_{i_1}\otimes\cdots\otimes v_{i_m}$ 构成 $T(V)$ 的基, 其中 m 遍历 $\mathbb{Z}_{\geq 0}$ 而 $(i_1,\dots,i_m)$ 遍历 I^m; 当 $m=0$ 时将 $v_{i_1}\otimes\cdots\otimes v_{i_m}$ 理解为 $1\in F=V^{\otimes 0}$.

证明 推论 15.2.6 在多元情形的应用. □

定义–命题 15.5.8 任何线性映射 $\psi:V\to W$ 都诱导唯一的 F-代数同态 $T(\psi):T(V)\to T(W)$, 使得它限制为 $\psi:V=V^{\otimes 1}\to W^{\otimes 1}=W$.

证明 对每个 $m\in\mathbb{Z}_{\geq 0}$, 定义–命题 15.1.6 给出 $\psi^{\otimes m}:=\psi\otimes\cdots\otimes\psi:V^{\otimes m}\to W^{\otimes m}$, 它在 $m=1$ 时化为 ψ, 而在 $m=0$ 时规定为 $\mathrm{id}_F:F\to F$. 取直和得到线性映射 $T(\psi):T(V)\to T(W)$. 为了说明 $T(\psi)$ 是 F-代数同态, 基于直和分解, 问题化为对所有 $a,b\geq 0$ 证明

$$\begin{array}{ccc} V^{\otimes a}\otimes V^{\otimes b} & \xrightarrow{\sim} & V^{\otimes(a+b)} \\ \downarrow{\scriptstyle \psi^{\otimes a}\otimes\psi^{\otimes b}} & & \downarrow{\scriptstyle \psi^{\otimes(a+b)}} \\ W^{\otimes a}\otimes W^{\otimes b} & \xrightarrow[\sim]{} & W^{\otimes(a+b)} \end{array}$$

是交换图表, 其中横向同构是结合约束, 同时也是张量代数的乘法.

左上空间的元素 $(x_1\otimes\cdots\otimes x_a)\otimes(y_1\otimes\cdots\otimes y_b)$ 按两种方式都映为 $\psi(x_1)\otimes\cdots\otimes\psi(x_a)\otimes\psi(y_1)\otimes\cdots\otimes\psi(y_b)$. 这些元素生成整个空间, 图表交换性得证.

由于 $V^{\otimes m}$ 由形如 $x_1\otimes\cdots\otimes x_m$ 的元素生成 ($x_i\in V$), 任何 F-代数的同态 $T(V)\to T(W)$ 都由它在 $V=V^{\otimes 1}$ 上的限制所确定. 所求同态的唯一性得证. □

在 §15.6 将需要以下的简单构造.

练习 15.5.9 设 A 为 F-代数而 I 为 A 的理想. 说明 I 也是子空间, 因而商环 A/I 同时也是对 I 的商空间. 说明 A/I 按此具有 F-代数的结构, 称为 A 对 I 的**商代数**.

提示〉 设 $t\in F$ 而 $x\in I$, 从 $tx=t(1_A\cdot x)=(t\cdot 1_A)x$ 可见 I 对纯量乘法封闭.

上述所有定义和性质都能从域 F 推及一般的交换环, 但本书正文只论域上的张量积, 对应的代数便也限制在域上.

15.6 对称代数与外代数

考虑 F-向量空间 V, M 和 m 重线性映射 $C \in \mathrm{Mul}(V,\dots,V;M)$.

定义 15.6.1 若上述之 C 满足性质

$$C(\dots,x,y,\dots) = C(\dots,y,x,\dots)$$

则称之为**对称**的. 若 C 满足性质

$$C(\dots,x,x,\dots) = 0$$

则称之为**交错**的.

对称多重线性映射相当于说调换任两个相邻变元 x 和 y 不改变 C 的取值; 这些对换生成对称群 $\mathfrak{S}_m$ (例 11.1.8), 对称性因而等价于

$$C(x_1,\dots,x_m) = C(x_{\sigma(1)},\dots,x_{\sigma(m)}), \quad \sigma \in \mathfrak{S}_m.$$

另一方面, 交错 m 重线性映射必满足

$$0 = C(\dots,x+y,x+y,\dots) = C(\dots,x,y,\dots) + C(\dots,y,x,\dots),$$

因此调换任两个相邻变元导致 C 的取值变号. 这就蕴涵

$$C(x_1,\dots,x_m) = \mathrm{sgn}(\sigma)C(x_{\sigma(1)},\dots,x_{\sigma(m)}), \quad \sigma \in \mathfrak{S}_m;$$

上式连同交错性质还进一步导致 C 在任两个变元相等时零化:

$$C(\dots,x,\dots,x,\dots) = 0.$$

注记 15.6.2 若 m 重线性映射 $C : V \times \cdots \times V \to M$ 对所有 $\sigma \in \mathfrak{S}_m$ 皆满足

$$C(x_1,\dots,x_m) = \mathrm{sgn}(\sigma)C\left(x_{\sigma(1)},\dots,x_{\sigma(m)}\right),$$

则称 C 为反对称的, 这也等价于 $C(\dots,x,y,\dots) = -C(\dots,y,x,\dots)$ 恒成立. 先前的讨论说明交错蕴涵反对称. 另一方面, 若 C 反对称则 $2C(\dots,x,x,\dots) = 0$, 由此立见当 $\mathrm{char}(F) \neq 2$ 时交错和反对称是等价的.

当 $M = F$ 时, 对应的对称 (或交错) m 重线性映射也称为 m 重对称 (或交错) 形式. 交错形式在行列式的定义中早已现身. 本节旨在从张量积的角度理解对称和交错线性映射.

定义 15.6.3 (对称代数与外代数) 设 V 为 F-向量空间, 定义张量代数 $T(V)$ 的理想如下:

$$\begin{aligned} I_{\mathrm{Sym}} &:= \text{形如 } x\otimes y - y\otimes x \text{ 的元素生成的理想},\\ I_{\wedge} &:= \text{形如 } x\otimes x \text{ 的元素生成的理想}, \end{aligned}$$

其中 $x, y \in V$. 相应的商代数定义为

$$\begin{aligned} \mathrm{Sym}(V) &:= T(V)\big/I_{\mathrm{Sym}},\\ \bigwedge(V) &:= T(V)/I_{\wedge}; \end{aligned}$$

我们称 $\mathrm{Sym}(V)$ 为 V 的对称代数, 称 $\bigwedge(V)$ 为 V 的外代数; 外代数又称为 Grassmann 代数.

具体地说, I_{Sym} 的元素是形如

$$u\otimes x\otimes y\otimes v - u\otimes y\otimes x\otimes v, \quad u\in V^{\otimes a}, \quad v\in V^{\otimes b}$$

的元素之线性组合; 同理, $I_{\wedge}$ 的元素是形如

$$u\otimes x\otimes x\otimes v$$

的元素之线性组合, 要求 $a, b \in \mathbb{Z}_{\geq 0}$ 而 $x, y \in V$. 上述元素都落在 $T(V)$ 的某个直和项中, 因此有

$$\begin{aligned} I_{\mathrm{Sym}} &= \bigoplus_{m\geq 0} I^m_{\mathrm{Sym}}, & I^m_{\mathrm{Sym}} &:= I_{\mathrm{Sym}}\cap V^{\otimes m},\\ I_{\wedge} &= \bigoplus_{m\geq 0} I^m_{\wedge}, & I^m_{\wedge} &:= I_{\wedge}\cap V^{\otimes m}. \end{aligned}$$

按此逐项取商: 命 $\mathrm{Sym}^m(V) = V^{\otimes m}/I^m_{\mathrm{Sym}}$ 而 $\bigwedge^m(V) = V^{\otimes m}/I^m_{\wedge}$, 则可以等同

$$\mathrm{Sym}(V) = \bigoplus_{m\geq 0}\mathrm{Sym}^m(V), \quad \bigwedge(V) = \bigoplus_{m\geq 0}\bigwedge{}^m(V).$$

易见 I^m_{Sym} 和 $I^m_{\wedge}$ 仅在 $m\geq 2$ 时方可能非零, 因此

$$\mathrm{Sym}^0(V) = F = \bigwedge{}^0(V), \quad \mathrm{Sym}^1(V) = V = \bigwedge{}^1(V).$$

约定 15.6.4 设 $x_1, \ldots, x_m \in V$. 今后将 $x_1\otimes\cdots\otimes x_m \in V^{\otimes m}$ 在 $\mathrm{Sym}^m(V)$ 中的像写作 $x_1\cdots x_m$, 将它在 $\bigwedge^m(V)$ 中的像写作 $x_1\wedge\cdots\wedge x_m$. 相应地, $\mathrm{Sym}(V)$ 的乘法用通常的乘法符号表示, 而 $\bigwedge(V)$ 的乘法用 $\wedge$ 表示.

根据理想 I_{Sym} (或 $I_{\wedge}$) 的定义, $(x_1, \ldots, x_m) \mapsto x_1\cdots x_m$ (或 $x_1\wedge\cdots\wedge x_m$) 是取值在 $\mathrm{Sym}^m(V)$ (或 $\bigwedge^m(V)$) 的对称 (或交错) m 重线性映射. 现在解释它们的泛性质.

命题 15.6.5 设 M 为 F-向量空间, $m \in \mathbb{Z}_{\geq 1}$, 则 $\mathrm{Mul}(V,\dots,V;M) \simeq \mathrm{Hom}\,(V^{\otimes m}, M)$ 限制为

$$\{C \in \mathrm{Mul}(V,\dots,V;M) : \text{对称}\} \simeq \mathrm{Hom}\,(\mathrm{Sym}^m V, M),$$
$$\{C \in \mathrm{Mul}(V,\dots,V;M) : \text{交错}\} \simeq \mathrm{Hom}\left(\bigwedge^m V, M\right),$$

此处将 $\mathrm{Hom}(\mathrm{Sym}^m V, M)$ 和 $\mathrm{Hom}(\bigwedge^m V, M)$ 嵌入为 $\mathrm{Hom}(V^{\otimes m}, M)$ 的子空间, 分别对应到在 I^m_{Sym} 和 $I^m_{\wedge}$ 上为零的线性映射.

证明 设 C 对应到 $\varphi \in \mathrm{Hom}(V^{\otimes m}, M)$. 对称性相当于

$$\varphi((\cdots \otimes x \otimes y \otimes \cdots) - (\cdots \otimes y \otimes x \otimes \cdots)) = 0,$$

交错性则相当于

$$\varphi(\cdots \otimes x \otimes x \otimes \cdots) = 0,$$

其中 $x, y \in V$. 问题因此回归 I_{Sym} 和 $I_{\wedge}$ 的定义. □

练习 15.6.6 验证对称代数 $\mathrm{Sym}(V)$ 满足乘法交换律 $xy = yx$, 而外代数 $\bigwedge(V)$ 的乘法则具有以下的反交换性

$$x \in \bigwedge^a(V),\ y \in \bigwedge^b(V) \implies x \wedge y = (-1)^{ab} yx.$$

提示 从 $x, y \in V$ 的特例来推导.

以下也将 $I_{\mathrm{Sym}}, I_{\wedge}$ 记为 $I_{\mathrm{Sym}}(V), I_{\wedge}(V)$ 以强调 V 的角色.

任给线性映射 $\psi : V \to W$, 定义–命题 15.5.8 的 F-代数同态 $T(\psi) : T(V) \to T(W)$ 显然限制为 $I_{\mathrm{Sym}}(V) \to I_{\mathrm{Sym}}(W)$ 和 $I_{\wedge}(V) \to I_{\wedge}(W)$, 因此 $T(\psi)$ 诱导商代数之间的同态, 具体写作

$$\begin{aligned} \mathrm{Sym}(\psi) : \mathrm{Sym}(V) &\to \mathrm{Sym}(W) \\ x_1 \cdots x_m &\mapsto \psi(x_1) \cdots \psi(x_m), \\ \bigwedge(\psi) : \bigwedge(V) &\to \bigwedge(W) \\ x_1 \wedge \cdots \wedge x_m &\mapsto \psi(x_1) \wedge \cdots \wedge \psi(x_m), \end{aligned}$$

其中 $x_i \in V$ 而 $m \in \mathbb{Z}_{\geq 0}$. 今后以 $\mathrm{Sym}^m(\psi)$ (或 $\bigwedge^m(\psi)$) 代表 $\mathrm{Sym}(\psi)$ (或 $\bigwedge(\psi)$) 在直和项 $\mathrm{Sym}^m(V)$ (或 $\bigwedge^m(V)$) 上的限制.

现在万事具备, 可以确定外代数的结构.

定理 15.6.7 设 V 为 n 维向量空间 ($n \in \mathbb{Z}_{\geq 0}$).

(i) 当 $m > n$ 时 $\bigwedge^m(V) = \{0\}$.

(ii) 当 $1 \leq m \leq n$ 时 $\dim \bigwedge^m(V) = \binom{n}{m}$; 更精确地说, 取定 V 的基 $v_1, \ldots, v_n$, 则

$$v_{i_1} \wedge \cdots \wedge v_{i_m}, \quad 1 \leq i_1 < \cdots < i_m \leq n$$

构成 $\bigwedge^m(V)$ 的基.

(iii) 我们有 $\dim \bigwedge(V) = 2^n$.

证明 取 V 的基 $v_1, \ldots, v_n$. 向量空间 $\bigwedge^m(V)$ 由形如 $v_{i_1} \wedge \cdots \wedge v_{i_m}$ 的元素生成. 按交错性质整理后可设 $1 \leq i_1 < \cdots < i_m \leq n$. 然而当 $m > n$ 时不存在这般的递增列, 故 (i) 成立.

设 $m \leq n$. 断言 (ii) 归结为证明 $v_{i_1} \wedge \cdots \wedge v_{i_m}$ 在 $\bigwedge^m(V)$ 中线性无关, 其中 $1 \leq i_1 < \cdots < i_m \leq n$. 这将给出所求的基和维数公式.

考虑特例 $m = n$. 此时 $\mathrm{Hom}(\bigwedge^n V, F)$ 同构于 n 元交错形式构成的空间; 探讨行列式时已证明此空间是 1 维的. 这也说明 $v_1 \wedge \cdots \wedge v_n$ 是 $\bigwedge^n(V)$ 的基.

对于一般的 $m \leq n$, 设有 $\bigwedge^m(V)$ 中的线性关系式

$$\sum_{i'_1 < \cdots < i'_m} c_{i'_1, \ldots, i'_m} v_{i'_1} \wedge \cdots \wedge v_{i'_m} = 0. \tag{15.6.1}$$

选定 $1 \leq i_1 < \cdots < i_m \leq n$. 定义 V 的 m 维子空间 $V' := \sum_{k=1}^m F v_{i_k}$ 和线性映射 $\psi : V \to V'$, 使得

$$\psi(v_a) = \begin{cases} v_a, & a \in \{i_1, \ldots, i_m\}, \\ 0, & a \notin \{i_1, \ldots, i_m\}. \end{cases}$$

线性映射 $\bigwedge^m(\psi)$ 映 $v_{i_1} \wedge \cdots \wedge v_{i_m}$ 为 $\bigwedge^m(V')$ 中同样形式的元素, 其余 $v_{i'_1} \wedge \cdots \wedge v_{i'_m}$ 必被映为 0. 依此对 (15.6.1) 取 $\bigwedge^m(\psi)$ 并运用 $\bigwedge^m(V') = F v_{i_1} \wedge \cdots \wedge v_{i_m}$ 维数为 1 这一事实, 可得 $c_{i_1, \ldots, i_m} = 0$. 变动 $i_1 < \cdots < i_m$ 遂知所有系数全为 0, 证得线性无关.

基于二项式定理和 $\bigwedge^0(V) = F$, (iii) 是 (i) 和 (ii) 的立即结论. □

以上采用的证明用到行列式理论的一部分结果, 而行列式也能反过来从外代数的视角来理解.

推论 15.6.8 设 V 为 n 维 F-向量空间, $n \in \mathbb{Z}_{\geq 0}$, 而 $\psi \in \mathrm{End}(V)$, 则 $\det(\psi) \in F$ 由下式刻画:

$$\bigwedge^n(\psi) = \det(\psi)\, \mathrm{id}_{\bigwedge^n(V)}.$$

证明 根据定义 5.4.1, 行列式 $\det(\psi)$ 的刻画是使得对所有 n 元交错形式 $D : V \times \cdots \times V \to F$ 都有

$$D(\psi(x_1), \ldots, \psi(x_n)) = \det(\psi) D(x_1, \ldots, x_n).$$

然而若 D 对应到 $\varphi \in \mathrm{Hom}(\bigwedge^n V, F)$, 则左式的交错形式对应到 $\varphi \circ \bigwedge^n(\psi)$, 右式对应到 $\det(\psi)\varphi$. 从 $\dim \bigwedge^n(V) = 1$ 可知存在 $\mu \in F$ 使得 $\bigwedge^n(\psi) = \mu \cdot \mathrm{id}_{\bigwedge^n(V)}$. 取不恒为零的 D (等价地, φ), 比较可见 $\mu = \det \psi$.

另一种方法是取 V 的基 $v_1,\dots,v_n$, 将 ψ 表为矩阵 $\boldsymbol{A}$, 然后直接写下

$$\bigwedge^n(\psi)(v_1\wedge\cdots\wedge v_n)=\sum_{1\le i_1,\dots,i_n\le n} a_{i_1,1}\cdots a_{i_n,n}v_{i_1}\wedge\cdots v_{i_n};$$

若 $i_1,\dots,i_n$ 包含重复项, 则 $v_{i_1}\wedge\cdots\wedge v_{i_n}=0$; 若 $i_1,\dots,i_n$ 是 $1,\dots,n$ 的重排, 记为 $\sigma\in\mathfrak{S}_n$, 则对应的项整理为

$$\operatorname{sgn}(\sigma)a_{\sigma(1),1}\cdots a_{\sigma(n),n}v_1\wedge\cdots\wedge v_n,$$

求和的产物正是 $\det\boldsymbol{A}$ 的熟知公式. □

有限维向量空间的对称代数同构于多项式代数, 同构依赖于有序基的选取. 以下仅陈述结果, 细节列入本章习题.

定理 15.6.9 *设 V 为 n 维空间 ($n\in\mathbb{Z}_{\ge0}$). 取定 V 的基 $v_1,\dots,v_n$, 则有 F-代数的同构*

$$\operatorname{Sym}(V)\xrightarrow{\sim}F[X_1,\dots,X_n];$$

在生成元的层次, 此同构映 v_i 为 X_i. 当 $n=0$ 时同构的右侧理解为 F.

注记 15.6.10 如果考虑张量代数 $T(V)$, 将 v_i 在其中的像记为 Y_i, 则对应的代数结构可以理解为 F 上的 n 元非交换多项式代数, 因为在 $T(V)$ 中不再有交换律 $Y_iY_j=Y_jY_i$; 强迫交换律成立相当于对所有 $Y_iY_j-Y_jY_i$ 在 $T(V)$ 中生成的理想取商, 亦即取商代数 $T(V)/I_{\mathrm{Sym}}$, 产物正是 $\operatorname{Sym}(V)\simeq F[X_1,\dots,X_n]$, 其中 X_i 对应到 Y_i 的像.

最后将理论落实到双线性形式. 回忆到对称双线性形式对应于对称矩阵. 以下阐明二元交错形式所对应的矩阵, 其定义适用于所有交换环.

定义 15.6.11 (交错矩阵) *设 R 为交换环, $\boldsymbol{A}\in\mathrm{M}_{n\times n}(R)$. 若 $a_{ij}=-a_{ji}$ 而且 $a_{ii}=0$ 对所有 i,j 成立, 则称 $\boldsymbol{A}$ 为交错矩阵.*

约定 15.6.12 *若整数 x 在环 R 中的像可逆, 则简记为 $x\in R^\times$.*

在满足 $2\in R^\times$ 的交换环 R 上, 交错矩阵和反对称矩阵 (亦即满足 ${}^{\mathrm{t}}\boldsymbol{A}=-\boldsymbol{A}$ 的矩阵) 是相同的概念.

命题 15.6.13 *设 $\boldsymbol{A}\in\mathrm{M}_{n\times n}(F)$, 相应的双线性形式 $B\in\mathrm{Bil}(F^n,F^n;F)$ 为 $B(\boldsymbol{x},\boldsymbol{y})={}^{\mathrm{t}}\boldsymbol{x}\,\boldsymbol{A}\,\boldsymbol{y}$, 则 B 是交错形式当且仅当 $\boldsymbol{A}$ 是交错矩阵.*

证明 回忆到 $a_{ij}={}^{\mathrm{t}}\boldsymbol{e}_i\,\boldsymbol{A}\,\boldsymbol{e}_j=B(\boldsymbol{e}_i,\boldsymbol{e}_j)$. 若 B 是交错形式, 则 $a_{ij}=B(\boldsymbol{e}_i,\boldsymbol{e}_j)=-B(\boldsymbol{e}_j,\boldsymbol{e}_i)=-a_{ji}$ 而 $a_{ii}=B(\boldsymbol{e}_i,\boldsymbol{e}_i)=0$, 故 $\boldsymbol{A}$ 是交错矩阵.

若 $\boldsymbol{A}$ 是交错矩阵, 则对所有 $\boldsymbol{x}=\sum_{i=1}^n x_i\boldsymbol{e}_i\in F^n$ 皆有

$$\begin{aligned}B(\boldsymbol{x},\boldsymbol{x})&=\sum_{1\le i,j\le n}x_ia_{ij}x_j\\&=\sum_{i<j}(a_{ij}+a_{ji})x_ix_j+\sum_i a_{ii}x_i^2=0,\end{aligned}$$

因此 B 是交错形式. □

练习 15.6.14 说明若 R 是交换环, $\boldsymbol{A},\boldsymbol{Q}\in \mathrm{M}_{n\times n}(R)$ 而 $\boldsymbol{A}$ 是交错矩阵, 则 ${}^{\mathrm{t}}\boldsymbol{Q}\,\boldsymbol{A}\,\boldsymbol{Q}$ 也是交错矩阵.

15.7 Pfaff 型与交错矩阵的行列式

设 F 为域. 选定 $n\in\mathbb{Z}_{\geq 1}$, 照例记 F^{2n} 的标准基为 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_{2n}$. 考虑 $\bigwedge^2(F^{2n})$ 的元素

$$\eta=\sum_{1\leq i<j\leq 2n} a_{ij}\boldsymbol{e}_i\wedge\boldsymbol{e}_j.$$

本节重点是计算 $\eta^n\in\bigwedge^{2n}(F^{2n})$. 先将 η^n 展开为如下表达式的和

$$a_{h_1k_1}\cdots a_{h_nk_n}\boldsymbol{e}_{h_1}\wedge\boldsymbol{e}_{k_1}\wedge\cdots\wedge\boldsymbol{e}_{h_n}\wedge\boldsymbol{e}_{k_n},\tag{15.7.1}$$

其中 $1\leq h_i<k_i\leq 2n$. 为了得到非零项, $h_1,k_1,\dots,h_n,k_n$ 必须两两相异, 换言之这列元素必须对应到 $\sigma\in\mathfrak{S}_{2n}$, 使得对所有 $1\leq i\leq n$ 皆有

$$h_i:=\sigma(2i-1)<\sigma(2i)=:k_i.$$

对于这般之 σ, 我们有

$$\boldsymbol{e}_{h_1}\wedge\boldsymbol{e}_{k_1}\wedge\cdots\wedge\boldsymbol{e}_{h_n}\wedge\boldsymbol{e}_{k_n}=\operatorname{sgn}(\sigma)\boldsymbol{e}_1\wedge\boldsymbol{e}_2\wedge\cdots\wedge\boldsymbol{e}_{2n-1}\wedge\boldsymbol{e}_{2n}.$$

对调 (h_i,k_i) 和 (h_{i+1},k_{i+1}) 相当于将 σ 换为

$$\sigma(2i-1\ 2i+1)(2i\ 2i+2),$$

这不改变 $\operatorname{sgn}(\sigma)$. 于是 (15.7.1) 仅和数组集 $\{(h_1,k_1),\dots,(h_n,k_n)\}$ 相关. 令 $\mathcal{S}$ 为所有相互无交而且满足 $h_i<k_i$ 的数组集 $\mathbf{s}=\{(h_i,k_i)\}_{i=1}^n$ 所成集合, 然后对 $\mathbf{s}$ 任取相应的 $\sigma\in\mathfrak{S}_{2n}$ (共有 $n!$ 种选法) 以定义

$$\operatorname{sgn}(\mathbf{s}):=\operatorname{sgn}(\sigma).$$

综上可得

$$\begin{aligned}\eta^n&=\sum_{(h_1,k_1),\dots,(h_n,k_n)}\left(\prod_{i=1}^n a_{h_ik_i}\right)\boldsymbol{e}_{h_1}\wedge\boldsymbol{e}_{k_1}\wedge\cdots\wedge\boldsymbol{e}_{h_n}\wedge\boldsymbol{e}_{k_n}\\&=n!\sum_{\mathbf{s}\in\mathcal{S}}\operatorname{sgn}(\mathbf{s})\left(\prod_{(h,k)\in\mathbf{s}}a_{hk}\right)\boldsymbol{e}_1\wedge\boldsymbol{e}_2\wedge\cdots\wedge\boldsymbol{e}_{2n-1}\wedge\boldsymbol{e}_{2n}.\end{aligned}$$

上式出现的 $\sum_{\mathbf{s}}\operatorname{sgn}(\mathbf{s})\prod_{(h,k)\in\mathbf{s}}a_{hk}$ 是关于诸 a_{hk} 的整系数多项式. 为了突出这点, 现将这些系数取为自由变元.

定义 15.7.1 引入 $n(2n-1)$ 个变元 X_{hk}, 其中 $1 \leq h < k \leq 2n$. 定义 $\mathbb{Z}[X_{hk} : 1 \leq h < k \leq 2n]$ 的元素

$$\mathrm{Pf} := \sum_{\mathrm{s} \in \mathcal{S}} \mathrm{sgn}(\mathrm{s}) \prod_{(h,k) \in \mathrm{s}} X_{hk}.$$

在 Pf 中代值, 便引出以下定义.

定义 15.7.2 (J. F. Pfaff) 设 R 为交换环, $\boldsymbol{A} \in \mathrm{M}_{2n \times 2n}(R)$ 为交错矩阵. 定义 R 的元素

$$\mathrm{Pf}(\boldsymbol{A}) := \mathrm{Pf}\left((a_{hk})_{h<k}\right).$$

我们称 $\mathrm{Pf}(\boldsymbol{A})$ 为 $2n \times 2n$ 交错矩阵 $\boldsymbol{A}$ 的 **Pfaff 型**.

在 F 上, 指定交错矩阵 $\boldsymbol{A}$ 相当于指定 $\bigwedge^2(F^{2n})$ 的元素 $\eta = \sum_{h<k} a_{hk} \boldsymbol{e}_h \wedge \boldsymbol{e}_k$. 综上所述, 它们满足

$$n! \mathrm{Pf}(\boldsymbol{A}) = \eta^n. \tag{15.7.2}$$

例 15.7.3 取 $n = 1$ 并考虑 $\boldsymbol{J}_1 := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$. 直接代入定义可得 $\mathrm{Pf}(\boldsymbol{J}_1) = 1$, 这也等于 $\det \boldsymbol{J}_1$.

Pfaff 型相对于直和具有乘性. 以下继续在交换环 R 上讨论.

引理 15.7.4 给定交错矩阵 $\boldsymbol{A} \in \mathrm{M}_{2a \times 2a}(R)$ 和 $\boldsymbol{B} \in \mathrm{M}_{2b \times 2b}(R)$, 命 $n = a + b$, 则相应的分块对角矩阵满足 $\mathrm{Pf}(\mathrm{diag}(\boldsymbol{A}, \boldsymbol{B})) = \mathrm{Pf}(\boldsymbol{A})\mathrm{Pf}(\boldsymbol{B})$.

证明 在 $\mathrm{Pf}(\mathrm{diag}(\boldsymbol{A}, \boldsymbol{B}))$ 的定义中, 只有 $1 \leq h, k \leq 2a$ 或 $2a+1 \leq h, k \leq 2n$ 的项方有贡献; 对于由这种数组构成的 s, 相应地 $\mathrm{sgn}(\mathrm{s})$ 也分解为两部分乘积. □

命题 15.7.5 对于定义 15.7.2 中的 $\boldsymbol{A}$ 以及任意 $\boldsymbol{P} \in \mathrm{M}_{2n \times 2n}(R)$, 我们有

$$\mathrm{Pf}(\boldsymbol{P}\, \boldsymbol{A}\, {}^{\mathrm{t}}\boldsymbol{P}) = (\det \boldsymbol{P}) \mathrm{Pf}(\boldsymbol{A}).$$

证明 先假定 R 为域, 另记为 F, 而且 $2, n! \in F^\times$, 则

$$\eta := \sum_{1 \leq i < j \leq 2n} a_{ij} \boldsymbol{e}_i \wedge \boldsymbol{e}_j = \frac{1}{2} \sum_{1 \leq i,j \leq 2n} a_{ij} \boldsymbol{e}_i \wedge \boldsymbol{e}_j.$$

视 $\boldsymbol{P}$ 为线性映射 $F^{2n} \to F^{2n}$, 相应地有 $\bigwedge^2 \boldsymbol{P} : \bigwedge^2(F^{2n}) \to \bigwedge^2(F^{2n})$. 作计算

$$\begin{aligned} \left(\bigwedge^2 \boldsymbol{P}\right)(\eta) &= \frac{1}{2} \sum_{i,j,k,l} a_{ij} p_{ki} p_{lj} \boldsymbol{e}_k \wedge \boldsymbol{e}_l \\ &= \frac{1}{2} \sum_{1 \leq k,l \leq 2n} \left(\sum_{i,j} p_{ki} a_{ij} p_{lj} \right) \boldsymbol{e}_k \wedge \boldsymbol{e}_l. \end{aligned}$$

因此若让交错矩阵 $\boldsymbol{P}\,\boldsymbol{A}\,{}^{\mathrm{t}}\boldsymbol{P}$ 对应到 $\eta' \in \bigwedge^2(F^{2n})$, 则 $\eta' = \bigwedge^2(\boldsymbol{P})(\eta)$. 既然 $\bigwedge(\boldsymbol{P})$ 是 F-代数 $\bigwedge(F^{2n})$ 的自同态,

$$(\eta')^n = \bigwedge^2(\boldsymbol{P})(\eta) \wedge \cdots \wedge \bigwedge^2(\boldsymbol{P})(\eta) = \bigwedge^{2n}(\boldsymbol{P})(\eta^n)$$
$$\xlongequal{\text{推论 15.6.8}} (\det \boldsymbol{P})\eta^n.$$

代入 (15.7.2) 遂有 $n!\mathrm{Pf}(\boldsymbol{P}\,\boldsymbol{A}\,{}^{\mathrm{t}}\boldsymbol{P}) = n!(\det \boldsymbol{P})\mathrm{Pf}(\boldsymbol{A})$, 两边可同步消去 $n!$.

为了放宽 R 的条件, 现将问题 "泛化". 取 $\mathcal{R}$ 为关于 X_{hk} 和 Y_{ab} 的整系数多元多项式环, 其中 $1 \leq h < k \leq 2n$ 而 $1 \leq a, b \leq 2n$. 取

$$\boldsymbol{\mathcal{A}} := \begin{pmatrix} 0 & X_{12} & X_{13} & \cdots & X_{1n} \\ -X_{12} & 0 & X_{23} & \cdots & X_{2n} \\ -X_{13} & -X_{23} & 0 & \cdots & X_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -X_{1n} & -X_{2n} & -X_{3n} & \cdots & 0 \end{pmatrix}, \quad \boldsymbol{\mathcal{P}} := (Y_{ab})_{1\leq a,b\leq n}.$$

注意到 det 和 Pf 都是关于矩阵元的整系数多项式. 我们欲验证 $\mathcal{R}$ 中等式

$$\mathrm{Pf}(\boldsymbol{\mathcal{P}}\,\boldsymbol{\mathcal{A}}\,{}^{\mathrm{t}}\boldsymbol{\mathcal{P}}) = (\det \boldsymbol{\mathcal{P}})\mathrm{Pf}(\boldsymbol{\mathcal{A}}). \tag{15.7.3}$$

这点可以置于 $\mathcal{R}$ 的分式域 F 中处理; 然而 $\mathbb{Q} \subset F$, 证明第一步的结果可资运用.

最后, 对于一般的 R 和 $\boldsymbol{A}, \boldsymbol{P}$, 在环 $\mathcal{R}$ 中的等式 (15.7.3) 中代值 $X_{hk} = a_{hk}$ 和 $Y_{ab} = p_{ab}$ 即可. □

定理 15.7.6 设 R 为交换环, $\boldsymbol{A} \in \mathrm{M}_{2n\times 2n}(R)$ 为交错矩阵, 则 $\det \boldsymbol{A} = \mathrm{Pf}(\boldsymbol{A})^2$.

证明 先假设 R 为域, $2 \in R^\times$ 而且 $\det \boldsymbol{A} \neq 0$. 此时 $\boldsymbol{A}$ 对应到辛形式, 从而辛形式的分类蕴涵有可逆的 $\boldsymbol{Q}$ 使得

$${}^{\mathrm{t}}\boldsymbol{Q}\,\boldsymbol{A}\,\boldsymbol{Q} = \left(\begin{array}{c|c|c} \boldsymbol{J}_1 & & \\ \hline & \ddots & \\ \hline & & \boldsymbol{J}_1 \end{array}\right), \quad \boldsymbol{J}_1 := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}.$$

根据命题 15.7.5, 等号左侧的 Pfaff 型是 $(\det \boldsymbol{Q})\mathrm{Pf}(\boldsymbol{A})$, 行列式是 $(\det \boldsymbol{Q})^2(\det \boldsymbol{A})$. 例 15.7.3 和引理 15.7.4 说明右侧的行列式和 Pfaff 型同为 1. 按此比较即得 $\det \boldsymbol{A} = (\det \boldsymbol{Q})^{-2} = \mathrm{Pf}(\boldsymbol{A})^2$.

对于一般情形, 观察到 $\det \boldsymbol{A}$ 和 $\mathrm{Pf}(\boldsymbol{A})^2$ 都是关于 $\boldsymbol{A}$ 的矩阵元的整系数多项式, 于是可仿照命题 15.7.5 的证明模式 "泛化" 到 $\boldsymbol{A}$ 的矩阵元为变元 X_{ij}, 而环为 $R = \mathbb{Z}[X_{ij} : i < j]$ 的情形. 尚须说明 $\det \boldsymbol{A} \neq 0$. 为此, 说明对 X_{ij} 适当代入整数值可得整系数可逆矩阵即足. 然而 $\mathrm{diag}(\boldsymbol{J}_1, \ldots, \boldsymbol{J}_1)$ 显然可以代值得到. 证毕. □

推论 15.7.7 设 $\boldsymbol{A}$ 为交换环 R 上的 $2n\times 2n$ 可逆交错矩阵, 而 $\boldsymbol{Q}\in \mathrm{M}_{2n\times 2n}(R)$ 和 $r\in R$ 满足 ${}^{\mathrm{t}}\boldsymbol{Q}\,\boldsymbol{A}\,\boldsymbol{Q}=r\boldsymbol{A}$, 则 $\det \boldsymbol{Q}=r^n$.

证明 因为 $\boldsymbol{A}$ 可逆, $\det \boldsymbol{A}\in R^\times$, 故定理 15.7.6 蕴涵 $\mathrm{Pf}(\boldsymbol{A})\in R^\times$. 对 ${}^{\mathrm{t}}\boldsymbol{Q}$ 和 $\boldsymbol{A}$ 应用命题 15.7.5, 并且留意到 $\mathrm{Pf}(\boldsymbol{A})$ 是关于矩阵元的 n 次齐次多项式, $\mathrm{Pf}(r\boldsymbol{A})=r^n\mathrm{Pf}(\boldsymbol{A})$, 即得 $\det \boldsymbol{Q}=r^n$. □

注记 15.7.8 定理 15.7.6 描述了偶数阶交错矩阵的行列式. 对于 $(2n+1)\times(2n+1)$ 的交错矩阵 $\boldsymbol{A}$, 相应的问题是平凡的: ${}^{\mathrm{t}}\boldsymbol{A}=-\boldsymbol{A}$ 蕴涵 $\det(\boldsymbol{A})=(-1)^{2n+1}\det(\boldsymbol{A})=-\det(\boldsymbol{A})$, 从而在 $2\in R^\times$ 的前提下 $\det \boldsymbol{A}=0$; 对于一般的 R, 先前证明中的泛化和代值技术同样能说明 $\det \boldsymbol{A}=0$.

练习 15.7.9 设 (V,B) 为域 F 上的 $2n$ 维辛空间, $\mathrm{char}(F)\neq 2$. 定义

$$\mathrm{GSp}(V):=\left\{g\in \mathrm{GL}(V):\exists t\in F^\times,\ \forall x,y\in V,\ B(gx,gy)=tB(x,y)\right\}.$$

(i) 说明 $\mathrm{GSp}(V)$ 是 $\mathrm{GL}(V)$ 的子群, 而定义中的 t 由 g 唯一确定, 称为 $g\in \mathrm{GSp}(V)$ 的**相似比**; 说明 $g\mapsto t$ 给出群同态 $\nu:\mathrm{GSp}(V)\to F^\times$.

(ii) 说明 $\ker(\nu)$ 等于例 11.1.13 介绍的辛群 $\mathrm{Sp}(V)$, 而 ν 满.

(iii) 说明若 $g\in \mathrm{GSp}(V)$, 则 $\det g=\nu(g)^n$. 特别地, $\mathrm{Sp}(V)\subset \mathrm{SL}(V)$.

提示〉 应用推论 15.7.7.

群 $\mathrm{GSp}(V)$ 的元素也称为 V 上的**辛相似变换**.

15.8 Amitsur–Levitzki 定理

设 $\mathcal{A}$ 为域 F 上的代数. 所谓 $\mathcal{A}$ 上的多项式恒等式, 意谓 F 上的一个 m 元非交换多项式 $f\neq 0$, 按照注记 15.6.10 的方式理解 ($m\in\mathbb{Z}_{\geq 1}$), 使得对 f 的变元 $Y_1,\dots,Y_m$ 代入任何 $a_1,\dots,a_m\in\mathcal{A}$ 均为 0. 初步例子:

- ⋆ 若 R 为交换 F 代数, 可取 $f=Y_1Y_2-Y_2Y_1$;
- ⋆ 若 R 为 $\mathrm{M}_{n\times n}(F)$ 的上三角子代数, 可取 $f=(Y_1Y_2-Y_2Y_1)^n$ (请证明).

选定 $n\in\mathbb{Z}_{\geq 1}$. 本节旨在为 $\mathcal{A}=\mathrm{M}_{n\times n}(F)$ 具体地写下一个多项式恒等式, 此处的 F 还能进一步取为任何交换环. 行将介绍的两则引理既是证明所需, 在其他场合也多有应用.

引理 15.8.1 设 R 为交换环, $\boldsymbol{A}\in \mathrm{M}_{n\times n}(R)$. 将 $\mathrm{Char}_{\boldsymbol{A}}\in R[X]$ 展开成

$$\mathrm{Char}_{\boldsymbol{A}}=(-1)^n b_n+(-1)^{n-1}b_{n-1}X+\cdots-b_1X^{n-1}+X^n,$$

另规定 $b_0=1$, 则对所有 $1\leq k\leq n$ 皆有

$$\operatorname{Tr}(\boldsymbol{A}^k)-b_1\operatorname{Tr}(\boldsymbol{A}^{k-1})+b_2\operatorname{Tr}(\boldsymbol{A}^{k-2})+\cdots+(-1)^{k-1}b_{k-1}\operatorname{Tr}(\boldsymbol{A})+(-1)^k kb_k=0.$$

证明 每个 b_i 都是关于 $\boldsymbol{A}$ 的矩阵元的多项式, 于是所求断言是关于 $\boldsymbol{A}$ 的矩阵元的多项式等式, 而且涉及的系数都是无关 R 的整数. 照搬命题 15.7.5 证明的 "泛化" 手法, 原断言便化约到 R 为某个域 $F\supset\mathbb{Q}$ 的情形.

适当扩张域 F, 不妨假定 $\operatorname{Char}_{\boldsymbol{A}}$ 分裂, 从而 $\boldsymbol{A}$ 共轭于某个上三角矩阵, 对角元记为 $\lambda_1,\ldots,\lambda_n$. 于是 $b_1,\ldots,b_n$ 能以 n 元初等对称多项式 $e_1,\ldots,e_n$ 表作 $b_i=e_i(\lambda_1,\ldots,\lambda_n)$. 规定 $e_0=1$.

另一方面, 称为幂和的对称多项式

$$p_k=\sum_{i=1}^n X_i^k,\quad 1\leq k\leq n;$$

满足 $\operatorname{Tr}(\boldsymbol{A}^k)=p_k(\lambda_1,\ldots,\lambda_n)$. Newton 公式 (见 [10, 定理 5.8.7] 或本书关于对称多项式的习题) 说明

$$p_k-e_1p_{k-1}+e_2p_{k-2}+\cdots+(-1)^{k-1}e_{k-1}p_1+(-1)^k ke_k=0$$

对所有 $1\leq k\leq n$ 成立. 将此式在 $(\lambda_1,\ldots,\lambda_n)$ 处代值即完成证明. □

引理 15.8.2 设 R 为满足 $n!\in R^\times$ 的交换环, $\boldsymbol{A}\in\mathrm{M}_{n\times n}(R)$, 而且 $\operatorname{Tr}(\boldsymbol{A}^k)=0$ 对所有 $1\leq k\leq n$ 成立, 则 $\boldsymbol{A}^n=\boldsymbol{0}_{n\times n}$.

证明 由于对每个 $1\leq k\leq n$ 皆有 $k\in R^\times$, 引理 15.8.1 配合条件 $\operatorname{Tr}(\boldsymbol{A}^1)=\cdots=\operatorname{Tr}(\boldsymbol{A}^n)=0$ 即刻给出 $b_k=0$. 这就说明 $\operatorname{Char}_{\boldsymbol{A}}=X^n$, 从而 Cayley–Hamilton 定理蕴涵 $\boldsymbol{A}^n=\boldsymbol{0}_{n\times n}$. □

继续设 R 为交换环. 对所有 $r,n\in\mathbb{Z}_{\geq 1}$ 和 $\boldsymbol{B}_1,\ldots,\boldsymbol{B}_r\in\mathrm{M}_{n\times n}(R)$, 命

$$S(\boldsymbol{B}_1,\ldots,\boldsymbol{B}_r):=\sum_{\sigma\in\mathfrak{S}_r}\operatorname{sgn}(\sigma)\boldsymbol{B}_{\sigma(1)}\cdots\boldsymbol{B}_{\sigma(r)}.$$

引理 15.8.3 当 r 为偶数时 $2\operatorname{Tr}S(\boldsymbol{B}_1,\ldots,\boldsymbol{B}_r)=0$.

证明 考虑轮换 $\xi:=(2\cdots r\,1)$. 对所有 $\sigma\in\mathfrak{S}_r$ 皆有

$$\begin{aligned}\operatorname{Tr}\left(\boldsymbol{B}_{\sigma(1)}\cdots\boldsymbol{B}_{\sigma(r)}\right)&=\operatorname{Tr}\left(\boldsymbol{B}_{\sigma(2)}\cdots\boldsymbol{B}_{\sigma(r)}\boldsymbol{B}_{\sigma(1)}\right)\\&=\operatorname{Tr}\left(\boldsymbol{B}_{\sigma\xi(1)}\cdots\boldsymbol{B}_{\sigma\xi(r)}\right).\end{aligned}$$

另一方面, r 为偶数蕴涵 $\operatorname{sgn}(\sigma\xi)=-\operatorname{sgn}(\sigma)$. □

定理 15.8.4 (S. A. Amitsur, J. Levitzki) 设 R 为交换环, 则对所有 $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_{2n}\in\mathrm{M}_{n\times n}(R)$ 皆有 $S(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_{2n})=\boldsymbol{0}_{n\times n}$, 亦即

$$\sum_{\sigma\in\mathfrak{S}_{2n}}\operatorname{sgn}(\sigma)\boldsymbol{A}_{\sigma(1)}\cdots\boldsymbol{A}_{\sigma(2n)}=\boldsymbol{0}_{n\times n}.$$

证明 (S. Rosset) 所求断言是关于 $2n^3$ 个矩阵元的一族多项式等式, 系数是无关 R 的整数, 因此照搬命题 15.7.5 证明的手法, 同样可将问题化到 R 等于某个域 $F \supset \mathbb{Q}$ 的情形.

将 F^{2n} 嵌入为外代数 $\Lambda := \bigwedge(F^{2n})$ 的子空间. 以 Λ 上的矩阵运算定义

$$\boldsymbol{A} := \boldsymbol{A}_1\boldsymbol{e}_1 + \cdots + \boldsymbol{A}_{2n}\boldsymbol{e}_{2n} \in \mathrm{M}_{n\times n}(\Lambda).$$

展开计算可见对所有 $k \geq 1$ 皆有

$$\boldsymbol{A}^k = \sum_{i_1<\cdots<i_k} \underbrace{S(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_k})}_{\in \mathrm{M}_{n\times n}(F)} \boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{e}_{i_k};$$

作为特例,

$$S(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_{2n})\boldsymbol{e}_1\wedge\cdots\wedge\boldsymbol{e}_{2n} = \boldsymbol{A}^{2n} = (\boldsymbol{A}^2)^n.$$

考虑 Λ 的交换子代数 $C := \bigoplus_{2|d}\bigwedge^d(F^{2n})$; 由乘法的交错性, 实际可推得 $C \subset Z(\Lambda)$ (见练习 15.6.6). 注意到 $\boldsymbol{A}^2 \in \mathrm{M}_{n\times n}(C)$. 应用引理 15.8.3 可得

$$\underbrace{\operatorname{Tr}\left((\boldsymbol{A}^2)^s\right)}_{\text{在 } C \text{ 上取}} = \sum_{i_1<\cdots<i_{2s}} \underbrace{\operatorname{Tr} S(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_{2s}})}_{\text{在 } F \text{ 上取, } =0} \boldsymbol{e}_{i_1}\wedge\cdots\wedge\boldsymbol{e}_{i_{2s}} = 0$$

对所有 $s \geq 1$ 成立, 从而引理 15.8.2 (环取为 C) 蕴涵 $(\boldsymbol{A}^2)^n = \boldsymbol{0}_{n\times n}$. 配合定理 15.6.7 (ii) 进一步推得 $S(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_{2n}) = \boldsymbol{0}_{n\times n}$. 明所欲证. □

按本节开头的说法, Amitsur–Levitzki 定理 15.8.4 表明 $n\times n$ 矩阵代数满足以下非交换多项式所确定的多项式恒等式.

$$f = \sum_{\sigma\in\mathfrak{S}_{2n}} \operatorname{sgn}(\sigma) Y_{\sigma(1)}\cdots Y_{\sigma(2n)}.$$

15.9 特征零的情形

在一些旧式教材中, 对称代数与外代数并非 $T(V)$ 的商空间, 而是 $T(V)$ 的子空间, 乘法的描述也不同. 基于子空间的构造仅适用于 $\operatorname{char}(F) = 0$ 的情形. 以下会通这两种观点, 以方便读者阅读文献, 本书不需要相关结论.

给定 F-向量空间 V, 对所有 $n \in \mathbb{Z}_{\geq 1}$, 运用交换约束 (命题 15.2.4) 让 $\mathfrak{S}_n$ 通过置换 n 个张量位作用于 $V^{\otimes n}$, 具体刻画如下: 设 $\sigma \in \mathfrak{S}_n$,

$$\begin{aligned} \sigma(ax+by) &= a(\sigma x) + b(\sigma y), \quad a,b\in F,\ x,y\in V^{\otimes n},\\ \sigma\left(v_1\otimes\cdots\otimes v_n\right) &= v_{\sigma^{-1}(1)}\otimes\cdots\otimes v_{\sigma^{-1}(n)}, \quad v_1,\ldots,v_n\in V. \end{aligned}$$

请读者验证这确实是 $\mathfrak{S}_n$ 的左作用.

定义 15.9.1 对所有 $n \in \mathbb{Z}_{\geq 1}$ 定义

$$\begin{aligned}\left(V^{\otimes n}\right)_{\mathrm{Sym}} &:= \left\{x \in V^{\otimes n} : \forall \sigma \in \mathfrak{S}_n,\ \sigma x = x\right\}\\ \left(V^{\otimes n}\right)_{\wedge} &:= \left\{x \in V^{\otimes n} : \forall \sigma \in \mathfrak{S}_n,\ \sigma x = \mathrm{sgn}(\sigma) x\right\},\end{aligned}$$

另规定 $(V^{\otimes 0})_{\mathrm{Sym}} = (V^{\otimes 0})_{\wedge} := F$. 定义 $T(V)$ 的子空间

$$\begin{aligned}T_{\mathrm{Sym}}(V) &:= \bigoplus_{n\geq 0} \left(V^{\otimes n}\right)_{\mathrm{Sym}},\\ T_{\wedge}(V) &:= \bigoplus_{n\geq 0} \left(V^{\otimes n}\right)_{\wedge}.\end{aligned}$$

另一方面, 按照对称代数和外代数的定义 15.6.3, 我们有商映射

$$\begin{aligned}q_{\mathrm{Sym}} &: T(V) \twoheadrightarrow \mathrm{Sym}(V),\\ q_{\wedge} &: T(V) \twoheadrightarrow \bigwedge(V);\end{aligned}$$

对所有 n, 它们给出从 $T^n(V)$ 到 $\mathrm{Sym}^n(V)$ 和 $\bigwedge^n(V)$ 的满射, 分别记为 q^n_{Sym} 和 $q^n_{\wedge}$. 沿用约定 15.6.12.

定理 15.9.2 设 $n \in \mathbb{Z}_{\geq 0}$. 若 $n! \in F^\times$, 则 q^n_{Sym} 和 $q^n_{\wedge}$ 分别限制为向量空间的同构

$$\left(V^{\otimes n}\right)_{\mathrm{Sym}} \xrightarrow{\sim} \mathrm{Sym}^n(V), \quad \left(V^{\otimes n}\right)_{\wedge} \xrightarrow{\sim} \bigwedge^n(V).$$

作为推论, 当 $\mathrm{char}(F) = 0$ 时有向量空间的同构

$$q_{\mathrm{Sym}} : T_{\mathrm{Sym}}(V) \xrightarrow{\sim} \mathrm{Sym}(V), \quad q_{\wedge} : T_{\wedge}(V) \xrightarrow{\sim} \bigwedge(V).$$

证明 见 [10, 定理 7.6.10]. 留意到 $n = 0$ 情形是平凡的. □

以下阐明 Sym^n 和 $\bigwedge^n$ 与对偶空间的联系.

推论 15.9.3 设 V 是有限维 F-向量空间, $n \in \mathbb{Z}_{\geq 0}$ 满足 $n! \in F^\times$, 则有自然的同构

$$\mathrm{Sym}^n(V^\vee) \simeq (\mathrm{Sym}^n V)^\vee, \quad \bigwedge^n(V^\vee) \simeq (\bigwedge^n V)^\vee.$$

证明 不妨设 $n \geq 2$, 否则问题平凡. 命题 15.3.5 和张量积的泛性质给出同构

$$\begin{array}{ccccc}(V^\vee)^{\otimes n} & \xrightarrow[\Psi]{\sim} & (V^{\otimes n})^\vee & \xrightarrow[\Phi]{\sim} & \mathrm{Mul}(V,\ldots,V;F)\\ \check{v}_1 \otimes \cdots \otimes \check{v}_n & & \xrightarrow{\quad\Phi\Psi\quad} & & [C(v_1,\ldots,v_n) = \prod_{i=1}^n \langle \check{v}_i, v_i\rangle],\end{array}$$

其中 $\check{v}_i \in V^\vee$ 而 $v_i \in V$. 群 $\mathfrak{S}_n$ 在左端已有作用, 它在右端的作用如下: $\sigma \in \mathfrak{S}_n$ 映 C 为

$${}^{\sigma}C(v_1,\ldots,v_n) = C\left(v_{\sigma(1)},\ldots,v_{\sigma(n)}\right).$$

请读者检验同构 $\Phi\Psi$ 保持 $\mathfrak{S}_n$-作用.

注记 15.6.2 和 $2 \in F^\times$ 确保交错等价于反对称. 比较 $\mathfrak{S}_n$ 的作用可见 $\Phi\Psi$ 限制为 $(V^\vee)^{\otimes n}_{\mathrm{Sym}} \xrightarrow{\sim} \{C : \text{对称}\}$ 和 $(V^\vee)^{\otimes n}_{\wedge} \xrightarrow{\sim} \{C : \text{交错}\}$. 然而命题 15.6.5 (代入 $M = F$) 表明 $(\mathrm{Sym}^n V)^\vee$ 和 $(\bigwedge^n V)^\vee$ 嵌入 $(V^{\otimes n})^\vee$ 之后分别通过 Φ 对应到对称和交错形式子空间. 综上可见 Ψ 诱导

$$(V^\vee)^{\otimes n}_{\mathrm{Sym}} \simeq (\mathrm{Sym}^n V)^\vee, \quad (V^\vee)^{\otimes n}_{\wedge} \xrightarrow{\sim} (\bigwedge^n V)^\vee;$$

再应用定理 15.9.2 即证毕. □

以上给出的均是向量空间同构, 未论及乘法. 为了在 $T(V)$ 的子空间 $T_{\mathrm{Sym}}(V)$ (或 $T_\wedge(V)$ 上反映 $\mathrm{Sym}(V)$ (或 $\bigwedge(V)$) 的乘法, 关键在于以下操作.

定义–命题 15.9.4 设 $n \in \mathbb{Z}_{\geq 1}$ 而 $n! \in F^\times$. 定义 $V^{\otimes n}$ 的线性自同态 e^n_{Sym} 和 $e^n_\wedge$ 如下:

$$e^n_{\mathrm{Sym}}(x) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sigma x,$$
$$e^n_\wedge(x) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma)\sigma x.$$

它们具有以下性质:

(i) e^n_{Sym} (或 $e^n_\wedge$) 的像等于 $(V^{\otimes n})_{\mathrm{Sym}}$ (或 $(V^{\otimes n})_\wedge$);

(ii) e^n_{Sym} (或 $e^n_\wedge$) 限制在 $(V^{\otimes n})_{\mathrm{Sym}}$ (或 $(V^{\otimes n})_\wedge$) 上是恒等;

(iii) $q^n_{\mathrm{Sym}}\left(e^n_{\mathrm{Sym}}(x)\right) = q^n_{\mathrm{Sym}}(x)$ (或 $q^n_\wedge\left(e^n_\wedge(x)\right) = q^n_\wedge(x)$) 对所有 $x \in V^{\otimes n}$ 成立.

证明 性质 (i) 和 (ii) 实际是 [10, 定理 7.6.10] 证明前的准备工作. 以 $\wedge$ 的版本为例, 回忆到 $\mathrm{sgn} : \mathfrak{S}_n \to \{\pm 1\}$ 是群同态和 $(-1)^2 = 1$ 可知对所有 $\tau \in \mathfrak{S}_n$,

$$\tau\left(\frac{1}{n!}\sum_\sigma \mathrm{sgn}(\sigma)\sigma x\right) = \frac{1}{n!}\sum_\sigma \mathrm{sgn}(\tau)\,\mathrm{sgn}(\tau\sigma)(\tau\sigma)x$$
$$\overset{\sigma' := \tau\sigma}{=\!=} \mathrm{sgn}(\tau)\frac{1}{n!}\sum_{\sigma'} \mathrm{sgn}(\sigma')\sigma' x,$$

故 $\mathrm{im}(e^n_\wedge) \subset (V^{\otimes n})_\wedge$, 而当 $x \in (V^{\otimes n})_\wedge$ 时

$$e^n_\wedge(x) = \frac{1}{n!}\sum_\sigma \mathrm{sgn}(\sigma)\sigma x = \frac{1}{n!}\sum_\sigma x = x.$$

这就一道确立了 (i) 和 (ii). 对于 (iii) 的 $\wedge$ 版本, 交错形式的性质导致 $q^n_\wedge(\sigma x) = \mathrm{sgn}(\sigma) q^n_\wedge(x)$, 从而 $q^n_\wedge(e^n_\wedge(x)) = \frac{1}{n!}\sum_\sigma \mathrm{sgn}(\sigma) q^n_\wedge(\sigma x) = q^n_\wedge(x)$.

类推可证 Sym 的版本, 而且更简单. □

鉴于定理 15.9.2, 记 $q^n_{\mathrm{Sym}} : (V^{\otimes n})_{\mathrm{Sym}} \xrightarrow{\sim} \mathrm{Sym}^n(V)$ (或 $q^n_\wedge : (V^{\otimes n})_\wedge \xrightarrow{\sim} \bigwedge^n(V)$) 的逆映射为 p^n_{Sym} (或 $p^n_\wedge$). 它们将用于以下结论.

命题 15.9.5 以下设 $a, b \geq 1$, 而 $(a+b)! \in F^\times$ (例如 $\mathrm{char}(F) = 0$ 的情形).

- ⋆ 对所有 $x \in (V^{\otimes a})_{\mathrm{Sym}}$ 和 $y \in (V^{\otimes b})_{\mathrm{Sym}}$,

$$p^{a+b}_{\mathrm{Sym}}\left(q^a_{\mathrm{Sym}}(x) q^b_{\mathrm{Sym}}(y)\right) = \frac{a!b!}{(a+b)!} \sum_{\sigma \in \mathfrak{S}_{a+b}/(\mathfrak{S}_a \times \mathfrak{S}_b)} \sigma(xy);$$

  和式中的 σ 实际代表 $\mathfrak{S}_a \times \mathfrak{S}_b$ 陪集中的任意代表元, $\mathfrak{S}_a \times \mathfrak{S}_b$ 按自明方式嵌入为 $\mathfrak{S}_{a+b}$ 的子群.

- ⋆ 类似地, 对所有 $x \in (V^{\otimes a})_\wedge$ 和 $y \in (V^{\otimes b})_\wedge$,

$$p^{a+b}_\wedge\left(q^a_\wedge(x) q^b_\wedge(y)\right) = \frac{a!b!}{(a+b)!} \sum_{\sigma \in \mathfrak{S}_{a+b}/(\mathfrak{S}_a \times \mathfrak{S}_b)} \mathrm{sgn}(\sigma)\sigma(xy).$$

证明 仍以 $\wedge$ 的版本为例. 由于 $q_\wedge : T(V) \to \bigwedge(V)$ 是同态, $q^a_\wedge(x) q^b_\wedge(y) = q^{a+b}_\wedge(xy)$. 定义–命题 15.9.4 (iii) 蕴涵

$$q^{a+b}_\wedge(xy) = q^{a+b}_\wedge\left(e^{a+b}_\wedge(xy)\right) = q^{a+b}_\wedge\left(\frac{1}{(a+b)!} \sum_{\sigma \in \mathfrak{S}_{a+b}} \mathrm{sgn}(\sigma)\sigma(xy)\right),$$

而且右侧括号内是 $\left(V^{\otimes(a+b)}\right)_\wedge$ 的元素. 若 $(\tau, \eta) \in \mathfrak{S}_a \times \mathfrak{S}_b \subset \mathfrak{S}_{a+b}$, 则

$$(\tau,\eta)(xy) = \tau(x)\eta(y) = \mathrm{sgn}(\tau)\,\mathrm{sgn}(\eta)xy = \mathrm{sgn}((\tau,\eta))xy,$$

故 $\mathrm{sgn}(\sigma)\sigma(xy)$ 只依赖陪集 $\sigma(\mathfrak{S}_a \times \mathfrak{S}_b)$. 每个陪集有 $a!b!$ 个元素, 将右侧括号按陪集求和, 得到

$$q^a_\wedge(x) q^b_\wedge(y) = q^{a+b}_\wedge\left(\frac{a!b!}{(a+b)!} \sum_{\sigma \in \mathfrak{S}_{a+b}/(\mathfrak{S}_a \times \mathfrak{S}_b)} \mathrm{sgn}(\sigma)\sigma(xy)\right).$$

对其两边同取 $p^{a+b}_\wedge$ 即可. □

命题 15.9.5 排除了 $a = 0$ 或 $b = 0$ 情形, 然而来自 $(V^{\otimes 0})_{\mathrm{Sym}} = (V^{\otimes 0})_\wedge = F$ 的乘法即纯量乘法. 因此在 $\mathrm{char}(F) = 0$ 的前提下, 定理 15.9.2 连同命题 15.9.5 在 $T(V)$ 的子空间 $T_{\mathrm{Sym}}(V)$ (或 $T_\wedge(V)$) 上描述了 $\mathrm{Sym}(V)$ (或 $\bigwedge(V)$) 的乘法, 它是对 $T(V)$ 的乘法实施对称化 (或反对称化) 的产物.

习题

1. 设 X 和 Y 为集合, F 为域. 所有映射 $f : X \to F$ 构成 F-向量空间 $C(X)$; 同理有 $C(Y)$.

(i) 对所有 $f \in C(X)$ 和 $g \in C(Y)$, 有 $C(X \times Y)$ 的元素 $(x, y) \mapsto f(x)g(y)$. 证明这给出双线性映射 $C(X) \times C(Y) \to C(X \times Y)$, 从而按张量积的泛性质确定线性映射 $\tau : C(X) \otimes C(Y) \to C(X \times Y)$.

(ii) 证明 τ 是单射.

提示〉 设 $\sum_{i=1}^{m} f_i \otimes g_i$ 被映为零, 不失一般性可设 $g_1, \dots, g_m$ 在 $C(Y)$ 中线性无关. 从恒等式 $\sum_{i=1}^{m} f_i(x)g_i(y) = 0$ 论证对所有 x 都有 $f_1(x) = \cdots = f_m(x) = 0$.

(iii) 证明 τ 是满射当且仅当 X 和 Y 其中之一是有限集.

提示〉 对于 "仅当" 方向, 留意到若 $h \in \mathrm{im}(\tau)$, 则存在 $C(Y)$ 的有限维子空间 V 使得 $h(x, \cdot) \in V$ 对所有 $x \in X$ 成立. 在 X 和 Y 皆无穷的情形构造不满足上述性质的 $h \in C(X \times Y)$.

(iv) 具体写下一个连续函数 $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, 使得 h 无法表如 $(x, y) \mapsto \sum_{i=1}^{m} f_i(x)g_i(y)$ 之形, 其中 f_i, g_i 是 $\mathbb{R}$ 上的连续函数.

2. 若在张量积的泛性质 (命题 15.1.1 (i)) 中改用图表

$$\begin{array}{ccc} V \times W & \xrightarrow{B^{\mathrm{univ}}} & L^{\mathrm{univ}} \\ & \searrow^{B} & \uparrow{\scriptstyle \exists!\, \varphi} \\ & & L \end{array}$$

对应的资料 $(L^{\mathrm{univ}}, B^{\mathrm{univ}})$ 是否存在? 有何描述?

3. 试证张量积保余核 (定义 4.12.5). 更精确地说, 设 $f : V \to V'$ 为线性映射, 证明对所有 F-向量空间 W 皆有同构

$$\begin{aligned} \mathrm{coker}(f) \otimes W &\xrightarrow{\sim} \mathrm{coker}\,(f \otimes \mathrm{id}_W) \\ (v' + \mathrm{im}(f)) \otimes w &\longmapsto v' \otimes w + \mathrm{im}(f \otimes \mathrm{id}_W) \end{aligned}$$

此处 $f \otimes \mathrm{id}_W$ 是 f 诱导的线性映射 $V \otimes W \to V' \otimes W$. 对于 $\mathrm{id}_W \otimes f$ 或多元张量积也有类似的同构.

提示〉 最简单的方法是检验双向都是良定义的线性映射, 而且互逆.

4. 说明张量积保商. 更精确地说, 设 V_0 为 V 的子空间, 以命题 15.2.9 (ii) 将 $V_0 \otimes W$ 嵌入为 $V \otimes W$ 的子空间, 则有同构

$$\begin{aligned} (V/V_0) \otimes W &\xrightarrow{\sim} (V \otimes W)/(V_0 \otimes W) \\ (v + V_0) \otimes w &\mapsto v \otimes w + (V_0 \otimes W). \end{aligned}$$

提示〉 在前一道习题中取 f 为包含映射 $V_0 \hookrightarrow V$.

5. 说明域上的张量积保核. 更精确地说, 设 $f : V \to V'$ 为线性映射, 则 $\ker(f) \otimes W \xrightarrow{\sim} \ker(f \otimes \mathrm{id}_W)$, 映法写作 $v \otimes w \mapsto v \otimes w$.

提示〉 先以 $\mathrm{im}(f)$ 代 V', 以命题 15.2.9 (ii) 将问题化到 f 满的情形, 记 j 为包含映射 $\ker(f) \hookrightarrow V$, 此时 f 可等同于 $V \twoheadrightarrow \mathrm{coker}(j)$. 用上一题将 $f \otimes \mathrm{id}_W$ 等同于 $V \otimes W \twoheadrightarrow \mathrm{coker}(j \otimes \mathrm{id}_W)$, 以推导 $\ker(f \otimes \mathrm{id}_W) = \mathrm{im}(j \otimes \mathrm{id}_W)$.

6. 设 V 和 W 为有限维向量空间, 维数都大于 1. 说明

$$\{v\otimes w : v\in V,\ w\in W\}\neq V\otimes W.$$

提示〉 取基 $(e_i)_i$ 和 $(f_j)_j$. 一种方法是注意到 $\sum_{i,j}x_{ij}e_i\otimes f_j:=(\sum_i a_ie_i)\otimes(\sum_j b_jf_j)$ 的系数是 $x_{ij}=a_ib_j$, 故此时 $x_{ij}x_{ji}=x_{ii}x_{jj}$.

7. 设 $\boldsymbol{A}\in \mathrm{M}_{p\times p}(F)$ 而 $\boldsymbol{B}\in \mathrm{M}_{q\times q}(F)$. 试证 $\det(\boldsymbol{A}\otimes\boldsymbol{B})=(\det\boldsymbol{A})^q(\det\boldsymbol{B})^p$.

8. 设 A 为域 F 上的有限维代数. 对 $a\in A$ 和 $f=\sum_n c_nX^n\in F[X]$ 定义 $f(a):=\sum_n c_na^n\in A$. 以下要求 A 作为 F-向量空间是有限维的.

(i) 说明对每个 $a\in A$ 都存在 $f\in F[X]\smallsetminus\{0\}$ 使得 $f(a)=0$.

(ii) 仿照线性映射的情形, 应用 (i) 对 $a\in A$ 定义其极小多项式.

(iii) 说明 a 可逆当且仅当它有左逆, 当且仅当它有右逆, 而且此时存在 $f\in F[X]$ 使得 $a^{-1}=f(a)$.

提示〉 若 a 有左逆或右逆, 则极小多项式的常数项必非零, 由此能以 a 的多项式表出 a^{-1}.

9. 设 R 为环. 对右 R-模 M, 左 R-模 N 和交换群 A (群运算写作加法), 满足以下性质的映射 $B:M\times N\to A$ 称为**平衡积**:

- $B(x+x',y)=B(x,y)+B(x',y)$,
- $B(x,y+y')=B(x,y)+B(x,y')$,
- $B(xr,y)=B(x,ry)$ 对所有 $r\in R$ 成立.

记所有平衡积构成的集合为 $\mathrm{Bil}(M,N;A)$; 它自然地对加法成群.

(i) 仿照向量空间的张量积, 说明存在交换群 $M\underset{R}{\otimes}N$ 连同平衡积 $B_{\mathrm{univ}}:M\times N\to M\underset{R}{\otimes}N$, 使得对所有 A 皆有群同构

$$\left\{\text{群同态 } M\underset{R}{\otimes}N\to A\right\}\xrightarrow{\sim}\mathrm{Bil}(M,N;A)$$
$$\varphi\mapsto\varphi B_{\mathrm{univ}}.$$

(ii) 说明上述泛性质唯一地刻画了 $M\underset{R}{\otimes}N$ 连同 B_{univ}, 精确到唯一同构. 今后称 $M\underset{R}{\otimes}N$ 为 M 和 N 的张量积, 并且记 $B_{\mathrm{univ}}(x,y)=x\otimes y$.

(iii) 说明若 M 是 (Q,R)-双模, N 是 (R,S)-双模 (见第十二章习题), 则 $M\underset{R}{\otimes}N$ 具有自然的 (Q,S)-双模结构使得

$$q(x\otimes y)s=qx\otimes ys,\quad q\in Q,\ s\in S.$$

(iv) 在 R 为交换环的情形, 说明 R-模的张量积仍是 R-模. 提示〉 任何 R-模皆可视为 (R,R)-双模, 方法是定义 $rxr':=(rr')x$.

(v) 试将结合约束 (命题 15.2.2), 幺约束 (命题 15.2.3), 保直和性质 (命题 15.2.5), 基的描述 (推论 15.2.6) 和保满射性质 (命题 15.2.9 (i)) 推广至任意环 R 上的张量积.

(vi) 对于交换环 R 上的模, 试推广张量积的交换约束 (命题 15.2.4), 并将对称代数和外代数的定义推广至 R-模.

这部分内容也可以参考 [10, §6.5, §7.5].

10. 以下基本性质称为 $\otimes$ 与 Hom 的伴随关系.

(i) 设 F 为域而 M, N, A 为 F-向量空间. 具体写下 F-向量空间的典范同构

$$\begin{aligned}\operatorname{Hom}(M \otimes N, A) &\simeq \operatorname{Hom}(N, \operatorname{Hom}(M, A)) \\ &\simeq \operatorname{Hom}(M, \operatorname{Hom}(N, A))\end{aligned}$$

其中的 Hom 意谓 F-线性映射所成集合, 它们本身也是 F-向量空间.

提示 各项都可以明确而典范地等同于 $\operatorname{Bil}(M, N; A)$.

(ii) 按照前一道习题的内容, 考虑双模的张量积. 设 Q, R, S 为环, 对

$$M : (Q, R)\text{-双模}, \quad N : (R, S)\text{-双模}, \quad A : (Q, S)\text{-双模}$$

具体写下加法群的典范同构

$$\begin{aligned}\operatorname{Hom}_{(Q,S)\text{-双模}}\left(M \underset{R}{\otimes} N, A\right) &\simeq \operatorname{Hom}_{(R,S)\text{-双模}}\left(N, \operatorname{Hom}_{\text{左 } Q\text{-模}}(M, A)\right) \\ &\simeq \operatorname{Hom}_{(Q,R)\text{-双模}}\left(M, \operatorname{Hom}_{\text{右 } S\text{-模}}(N, A)\right),\end{aligned}$$

其下标指明 Hom 是何种意义下的同态集; 这些 Hom 集具有自然的模结构, 请见第十二章习题.

提示 映法和 (i) 类似, 也可以参考 [10, 定理 6.6.5].

(iii) 试将 (i) 诠释为 (ii) 的特例.

11. 若 $f : R \to R'$ 是环同态, 则可以通过 f 使 R' 自然地成为 (R', R)-双模. 说明若 M 是左 R-模, 则 $R' \underset{R}{\otimes} M$ 按此具有自然的 R'-模结构.

12. 证明若 I, J 是交换环 R 的理想, 则有 R-模同构

$$(R/I) \underset{R}{\otimes} (R/J) \simeq R/(I + J).$$

13. 设 F 是域 E 的子域. 对所有 F-向量空间 V, 依 §15.4 定义 E-向量空间 $V_E := E \underset{F}{\otimes} V$. 对于所有 F-向量空间 $V_1, \ldots, V_n$, 试给出自然的 E-线性同构

$$(V_1 \underset{F}{\otimes} \cdots \underset{F}{\otimes} V_n)_E \simeq V_{1,E} \underset{E}{\otimes} \cdots \underset{E}{\otimes} V_{n,E}.$$

14. 设 A 为域 F 上的代数. 对所有 $a \in A$ 记 $L_a \in \operatorname{End}(A)$ 为 F-线性变换 $x \mapsto ax$. 证明 $a \mapsto L_a$ 给出 F-代数的嵌入 $L : A \hookrightarrow \operatorname{End}(A)$.

15. 将 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 视同线性映射 $F^n \to F^m$, 试以子行列式对所有 $k \leq \min\{m, n\}$ 描述 $\bigwedge^k \boldsymbol{A} : \bigwedge^k(F^n) \to \bigwedge^k(F^m)$.

16. 设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in\mathbb{R}^3$, 证明

$$x\boldsymbol{e}_2\wedge\boldsymbol{e}_3+y\boldsymbol{e}_3\wedge\boldsymbol{e}_1+z\boldsymbol{e}_1\wedge\boldsymbol{e}_2\mapsto(x,y,z)$$

定义的同构 $\bigwedge^2(\mathbb{R}^3)\xrightarrow{\sim}\mathbb{R}^3$ 映 $\boldsymbol{a}\wedge\boldsymbol{b}$ 为向量的叉积 $\boldsymbol{a}\times\boldsymbol{b}$. 其次, 证明关于纯量三重积的等式

$$\boldsymbol{a}\wedge\boldsymbol{b}\wedge\boldsymbol{c}=\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})\boldsymbol{e}_1\wedge\boldsymbol{e}_2\wedge\boldsymbol{e}_3.$$

17. 按下述步骤完成定理 15.6.9 的证明.

(i) 设 V 为 F-向量空间. 验证交换 F-代数 $\mathrm{Sym}(V)$ 连同线性映射 $\iota:V=\mathrm{Sym}^1(V)\to\mathrm{Sym}(V)$ 具有以下泛性质: 对所有资料 (A,ι_A), 其中 A 是交换 F-代数而 $\iota_A:V\to A$ 是线性映射, 存在唯一的 F-代数同态 f 使下图交换

$$\begin{array}{ccc} V & \xrightarrow{\ \iota\ } & \mathrm{Sym}(V) \\ & {\scriptstyle\iota_A}\searrow & \downarrow{\scriptstyle f} \\ & & A. \end{array}$$

(ii) 说明泛性质刻画了资料 $(\mathrm{Sym}(V),\iota)$, 精确到唯一的同构.

(iii) 现在设 $\dim V=n\in\mathbb{Z}_{\geq1}$, 选定 V 的基 $v_1,\dots,v_n$ 并定义线性映射 $\iota':V\to F[X_1,\dots,X_n]$ 使得 $\iota'(v_i)=X_i$. 说明 $(F[X_1,\dots,X_n],\iota')$ 也满足上述泛性质.

(iv) 依据上述结果补全定理 15.6.9 的证明.

对称代数的上述泛性质并非最精确的, 因为它未涉及直和分解 $\bigoplus_m\mathrm{Sym}^m(V)$. 相关的细化请见 [10, 定理 7.6.6].

18. 承上题, 对张量代数 $T(V)$ 连同 $\iota:V=V^{\otimes1}\to T(V)$ 给出类似的泛性质刻画.

19. 设 V 是域 F 上的 n 维向量空间, $T\in\mathrm{End}(V)$. 证明

$$\det(\lambda\cdot\mathrm{id}-T)=\sum_{k=0}^{n}(-1)^k\,\mathrm{Tr}\left(\bigwedge^k T\right)\lambda^{n-k}.$$

此处规定 $\bigwedge^0(V)=F$ 而 $\bigwedge^0T=\mathrm{id}_F$.

20. 选定 $\omega\in\bigwedge^p(V)\smallsetminus\{0\}$, 其中 $1\leq p\leq n:=\dim V$. 定义 $\mathrm{ann}(\omega):=\{v\in V:\omega\wedge v=0\}$. 若存在 $v_1,\dots,v_p\in V$ 使得

$$\omega=v_1\wedge\cdots\wedge v_p,$$

则称 ω 是可分解的.

(i) 说明 $\mathrm{ann}(\omega)$ 是 V 的子空间. 设 $x_1,\dots,x_p\in V$ 线性无关, 证明 $\mathrm{ann}(x_1\wedge\cdots\wedge x_p)=\langle x_1,\dots,x_p\rangle$.

(ii) 证明若 $\mathrm{ann}(\omega)$ 有基 $v_1,\dots,v_r$, 则 $r\leq p$ 而且存在 $\eta\in\bigwedge^{p-r}(V)$ 使得 $\omega=v_1\wedge\cdots\wedge v_r\wedge\eta$.

(iii) 承上题, 说明 $r=p$ 当且仅当 ω 可分解.

21. 承上题, 证明所有 $\omega\in\bigwedge^{n-1}(V)\smallsetminus\{0\}$ 都是可分解的.

提示〉 选基 $v_1,\dots,v_n$, 则 $\omega\wedge v=c(v)v_1\wedge\cdots\wedge v_n$, 其中 $c:V\to F$ 是线性的.

22. 设 U 和 W 为 V 的 p 维子空间, U 有基 $x_1,\dots,x_p$ 而 W 有基 $y_1,\dots,y_p$. 证明 $U=W$ 当且仅当 $x_1\wedge\cdots\wedge x_p$ 和 $y_1\wedge\cdots\wedge y_p$ 成比例.

23. 默认 $\operatorname{char}(F)\neq 2$. 设 $\dim V\in\mathbb{Z}_{\geq 2}$ 而 $\omega\in\bigwedge^2(V)\smallsetminus\{0\}$. 证明 ω 可分解当且仅当 $\omega\wedge\omega=0$.

24. 按照推论 15.6.8 的方式, 以外代数理解线性映射的行列式, 按此给出 Cauchy–Binet 定理的另证.

25. 考虑 F-向量空间 V, 要求 $\operatorname{char}(F)\neq 2$.

(i) 证明 $V^{\otimes 2}=\left(V^{\otimes 2}\right)_{\mathrm{Sym}}\oplus\left(V^{\otimes 2}\right)_{\wedge}$, 符号如定义 15.9.1.

(ii) 对所有 F-向量空间 L 证明

$$\operatorname{Bil}(V,V;L)=\left\{B:V\times V\to L\;\middle|\;\text{对称}\right\}\oplus\left\{B:V\times V\to L\;\middle|\;\text{交错}\right\}.$$

(iii) 基于定理 15.9.2, 试联系 (i) 与 (ii) 的直和分解.

第十六章　二次型的 Witt 理论

设域 F 满足 $\mathrm{char}(F) \neq 2$. 在此前关于双线性形式的讨论中, F 上的二次型的基本定义及分类问题已经得到表述. 然而除了 F 为代数闭域或 $F = \mathbb{R}$ 的情形, 该处对分类问题并未取得其它实质进展.

如今既然拥有更深入的代数工具, 本章将复归二次型的主题, 在一般的域 F 上进一步梳理其结构. 这部分的主要结论归功于 E. Witt [6], 包括搬运定理 16.1.12, 消去定理 16.2.4, 分解定理 16.2.10 和等价链定理 16.3.6.

Witt 的创新之处还在于他引入了域 F 的 Witt 群 $\mathcal{W}(F)$ (定义–命题 16.3.3), 此群蕴藏关于 F 上的二次型理论的根本信息. 约略言之, 这是用 F 上的所有非退化二次型以及直和运算作成的加法群, 但必须对双曲平面 (定义 16.2.5) 取商; 等价的说法则是对所谓的 Witt 等价 (定义 16.3.1) 取商.

各节内容具体介绍如下. 我们先在 §16.1 介绍域 F 上的二次型及其自同构群, 亦即正交群, 然后证明 Witt 搬运定理. 随后的 §16.2 先为二次型的对角化 (所谓 “配方”) 提供简短证明, 然后给出关键的 Witt 消去定理与分解定理. 作为示例, 例 16.2.12 简洁地重证关于实二次型分类的惯性定理. 我们在 §16.3 定义 F 的 Witt 群 $\mathcal{W}(F)$ 与 Grothendieck–Witt 群 $\widehat{\mathcal{W}}(F)$.

此后, §16.4 引入全迷向子空间的概念. 非退化二次型的极大全迷向子空间格外重要; 定理 16.4.6 将说明它们可以用正交群彼此搬运, 由此便能得到 Witt 搬运定理的一则推广 (推论 16.4.7).

在前一节的基础上, §16.5 将证明 Cartan–Dieudonné 定理 16.5.3, 它断言对于非退化 n 维二次型 (V, q), 正交群 $\mathrm{O}(V, q)$ 的所有元素都能写成不超过 n 个镜射的积; 这一部分具有较强的技术性.

我们接着在 §16.6 引入二次型的张量积, 它实现在对应的向量空间的张量积上. 这一运算赋予 $\mathcal{W}(F)$ 和 $\widehat{\mathcal{W}}(F)$ 乘法, 使之成环. 此外, 张量积还给出将二次型从 F 变换到任意扩域 E 上的自然方法, 不涉及基的选取. 关于这些环在 $F = \mathbb{R}, \mathbb{C}$ 和有限域情形的精确描述是 §16.7 的主题.

本章的 §§16.8–16.9 涉及域上的 Hermite 形式. 复向量空间上的 Hermite 形式已经在复内积的章节谈过, 此处的起点则是更一般的域 E 及满足 $\tau^2 = \mathrm{id}_E$ 的自同构 $\tau : E \xrightarrow{\sim} E$, 依旧要求 $\mathrm{char}(E) \neq 2$; 它们取代了早先的 $\mathbb{C}$ 及复共轭映射 $z \mapsto \bar{z}$ 的角色. 在此框架下, 仍可定义何谓半双线性映射与 ϵ-Hermite 形式 ($\epsilon = \pm 1$), 理路与 $\mathbb{C}$ 上

类似. 如果取 $\tau_E = \mathrm{id}_E$ 而 $\epsilon = 1$ (或 $\epsilon = -1$), 便能将 E 上的二次型 (或辛形式) 纳入 ϵ-Hermite 形式的框架.

我们最关心的自然是 $\tau \neq \mathrm{id}_E$ 的情形. 关于二次型的许多结论都能推及此时的 Hermite 形式, 详见定理 16.9.5; 由于论证总体类似, 该处仅作简单勾勒.

最后, §16.10 对于更一般的环 R 和称为对合的一类自映射 $\tau : R \to R$ 勾勒相应的 ϵ-Hermite 形式理论, 但要求 $2 \in R^\times$. 由于 R 未必是域, 甚至未必是除环, 先前的有限维 E-向量空间需替换为定义 16.10.5 所谓的有限生成投射 R-模, 由此引出正则 ϵ-Hermite 模的概念 (定义 16.10.8). 我们将扼要地讨论两种简单例子:

- ▷ **四元数情形** $R = \mathbb{H}$ 而 $\tau(q) = \bar{q}$ (例 16.10.11), 在 $\epsilon = 1$ 时容易证明四元数版本的惯性定理 16.10.12;
- ▷ **直积情形** $R = F \times F$, 其中 F 是域, 而 $\tau(x, y) = (y, x)$ (例 16.10.13), 该处将解释如何将正则 ϵ-Hermite 模理论化约为有限维 F-向量空间理论.

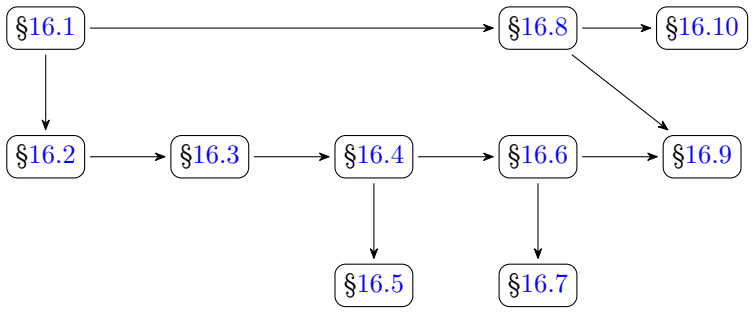
本章 §§16.9–16.10 的证明将比较简略.

在一般的域或环上研究 Hermite 形式并非只为追逐抽象, 它们之于数论或表示论领域的许多问题实属必要. 此外, $\mathbb{Q}$ 或其有限扩域上的二次型理论也有丰富的算术内涵. 由于这些面向需要更多数论背景, 最低限度也需要 p-进数的知识, 本书就此打住.

阅读提示

本章默认域的特征不等于 2, 在 §16.10 则相应地要求环 R 满足 $2 \in R^\times$. 在更一般的域或环上依然能开展二次型或 Hermite 形式的理论, 也确实有所应用, 但技术上复杂不少, 故本章未予讨论. 感兴趣的读者宜参考专著, 如 [3].

阅读顺序



16.1 二次型与正交群

今后默认 F 是满足 $\mathrm{char}(F)\neq 2$ 的域. 根据定义 8.4.4, 所谓二次型, 意谓资料 (V,B), 其中 V 是有限维 F-向量空间, 而 $B:V\times V\to F$ 是对称双线性形式.

- ⋆ 此二次型的维数定义为向量空间 V 的维数.
- ⋆ 若 $x,y\in V$ 满足 $B(x,y)=0$, 则称它们正交, 写作 $x\perp y$;
- ⋆ 给定子空间 $U_1,U_2\subset V$, 若 $B|_{U_1\times U_2}=0$ 则称 U_1 和 U_2 正交, 写作 $U_1\perp U_2$;
- ⋆ 对任意子空间 $U\subset V$, 定义 $U^\perp:=\{v\in V:\forall u\in U,\ B(u,v)=0\}$, 因此 B 的根基等于 $V^\perp$, 而 (V,B) 非退化当且仅当 $V^\perp=\{0\}$.

一族二次型 $(V_1,B_1),\dots,(V_n,B_n)$ 的直和 $(V,B):=(\bigoplus_i V_i,\bigoplus_i B_i)$ 也是二次型. 由于 B 的刻画是 $B|_{V_i\times V_i}=B_i$ 和 $i\neq j\implies V_j\perp V_j$, 二次型的直和又称为正交直和. 易见 (V,B) 非退化当且仅当所有 (V_i,B_i) 皆非退化.

定义–命题 16.1.1 *设 (V,B) 为二次型, U 为 V 的子空间, 则 U 上的二次型 $B_U:=B|_{U\times U}$ 的根基等于 $U^\perp\cap U$. 若 (U,B_U) 非退化, 则 $V=U\oplus U^\perp$, 此时我们称 U 为非退化子空间.*

证明 关于根基的断言近乎同义反复. 现在设 (U,B_U) 非退化, 故 $U\cap U^\perp=\{0\}$. 取 U 的基 $u_1,\dots,u_m$. 可取 $u_1',\dots,u_m'\in U$ 使得

$$B(u_i,u_j')=\begin{cases}1, & i=j,\\ 0, & i\neq j;\end{cases}$$

事实上, 它们是对偶基 $\check{u}_1,\dots,\check{u}_m\in U^\vee$ 在 B_U 对应的同构 $U\xrightarrow{\sim}U^\vee$ 之下的像. 于是对任意 $v\in V$, 考虑 $B(u_i,\cdot)$ 可得

$$v-\sum_{i=1}^m B(u_i,v)u_i'\in U^\perp,$$

故 $V=U+U^\perp$. 分解得证. □

若 (V,B) 和 (U,B_U) 皆非退化, 则定义–命题 16.1.1 之前的讨论说明 $(U^\perp,B_{U^\perp})$ 亦然.

约定 16.1.2 指定 B 相当于指定由 $q(v)=B(v,v)$ 确定的映射 $q:V\to F$, 是故资料 (V,q) 或 q 本身也经常用以指代二次型, 视场合而定, 此时 $B(x,y)=\frac{1}{2}(q(x+y)-q(x)-q(y))$.

二次型的直和也依此简记为 $q_1\oplus\cdots\oplus q_n$ 的形式.

定义 16.1.3 设 (V_1, B_1) 和 (V_2, B_2) 为二次型.

- ⋆ 若线性映射 $\varphi : V_1 \to V_2$ 对所有 $x, y \in V_1$ 皆满足 $B_2(\varphi(x), \varphi(y)) = B_1(x, y)$, 则称 φ 为等距; 等价的刻画是 $q_2\varphi = q_1$.
- ⋆ 若 $\varphi : V_1 \to V_2$ 既是等距又是线性同构, 则称 φ 为从 (V_1, B_1) 到 (V_2, B_2) 的同构, 或更简洁地说是从 q_1 到 q_2 的同构. 此时 φ^{-1} 也是等距.

今后将同构记为 $(V_1, B_1) \simeq (V_2, B_2)$, $(V_1, q_1) \simeq (V_2, q_2)$, 或者更简洁的 $q_1 \simeq q_2$.

二次型也有行列式的概念. 选定 V 的有序基, 将其等同于 F^n, 从而 $B : V \times V \to F$ 等同于对称矩阵 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(F)$. 二次型非退化当且仅当 $\boldsymbol{A}$ 可逆. 换基相当于将 $\boldsymbol{A}$ 代换为 $\boldsymbol{A}' = {}^{\mathrm{t}}\boldsymbol{P}\,\boldsymbol{A}\boldsymbol{P}$, 其中 $\boldsymbol{P} \in \mathrm{GL}(n, F)$, 于是

$$\det \boldsymbol{A}' = \det \boldsymbol{A} (\det \boldsymbol{P})^2.$$

定义群 $F^\times$ 的子群 $F^{\times 2} := \left\{ t^2 : t \in F^\times \right\}$. 因此行列式在 $F^{\times 2}$ 对 F 的乘法作用下的轨道

$$\det \boldsymbol{A} \cdot F^{\times 2} \ \in F/F^{\times 2}$$

对换基不变, 从而是二次型的同构不变量.

定义 16.1.4 以上定义的 $F^{\times 2}$-轨道称为二次型 (V, B) 或 (V, q) 的行列式, 简记为 $\det q$. 规定零维二次型的行列式为轨道 $F^{\times 2}$.

一如我们迄今学习到的所有代数结构, 自同构或曰对称性在二次型理论中扮演关键角色.

定义 16.1.5 (正交群) 二次型 (V, q) 的正交群是

$$\mathrm{O}(V, q) := \{q \text{ 的自同构}\};$$

这是 V 的线性自同构群 $\mathrm{GL}(V)$ 的子群, 也简记为 $\mathrm{O}(V)$. 另记

$$\mathrm{SO}(V, q) := \mathrm{O}(V, q) \cap \mathrm{SL}(V).$$

对于 $V = \{0\}$ 的情形, 规定 $\mathrm{O}(V, q)$ 为平凡群.

注记 16.1.6 对于任意 $t \in F^\times$, 定义 tq 为 V 上的二次型 $v \mapsto t \cdot q(v)$. 尽管一般而言 $tq \not\simeq q$, 但 $\mathrm{O}(V, q) = \mathrm{O}(V, tq)$.

显然 $\pm\mathrm{id} \in \mathrm{O}(V, q)$. 正交群 $\mathrm{O}(V, q)$ 中的一类重要元素是镜射, 它们由非迷向向量给出, 具体定义如下.

定义 16.1.7 给定二次型 (V, q). 若 $x \in V$ 满足 $q(x) = 0$ 则称 x 是**迷向**的, 否则称 x 为**非迷向**的.

定义–命题 16.1.8 设 $x \in V$ 非迷向. 对所有 $v \in V$ 定义

$$r_x(v) = r_x^V(v) := v - \frac{2B(x,v)}{q(x)}x,$$

则 $r_x \in \mathrm{O}(V,q)$. 我们称 r_x 为以直线 Fx 为法向的**镜射**, 它满足 $r_x^2 = \mathrm{id}$.

证明 显然 $r_x \in \mathrm{End}(V)$. 既然 x 非迷向, 便有正交分解 $V = Fx \oplus (Fx)^\perp$. 从定义看出若 $v = ax + y$, 其中 $a \in F$ 而 $y \perp x$, 则 $r_x(v) = -ax + y$. 由此立见 $r_x \in \mathrm{O}(V,q)$ 而 $r_x^2 = \mathrm{id}$. □

注记 16.1.9 镜射 r_x 在 $V = Fx \oplus (Fx)^\perp$ 之下的描述立即给出 $\det r_x = -1$.

镜射只依赖于直线 Fx. 当 $F = \mathbb{R}$ 而 B 是内积时, $\frac{B(x,v)}{q(x)}x$ 等于 v 在 $\mathbb{R}x$ 上的正交投影, 因此这与定义 9.3.14 的镜射是兼容的: 该处的 V_0 对应到此处的 $(\mathbb{R}x)^\perp$.

以下两则练习不过是操演定义, 请读者完成.

练习 16.1.10 设 (V,q) 是二次型 $(V_1,q_1),\dots(V_n,q_n)$ 的直和. 说明有群的单同态

$$\begin{aligned}\prod_{i=1}^n \mathrm{O}(V_i,q_i) &\to \mathrm{O}(V,q)\\ (g_i)_{i=1}^n &\mapsto \mathrm{diag}(g_1,\dots,g_n).\end{aligned}$$

练习 16.1.11 作为上一则练习的特例, 若 (V,q) 是二次型而 U 是非退化子空间 (定义–命题 16.1.1), 则 $\mathrm{O}(U,q|_U)$ 通过 $V = U \oplus U^\perp$ 自然地嵌入为 $\mathrm{O}(V,q)$ 的子群. 验证:

(i) 若 $x \in U$ 非迷向, 则它作为 V 的元素也是非迷向的;

(ii) 设 $x \in U$ 非迷向, 相应地有镜射 $r_x^U \in \mathrm{O}(U,q|_U)$, 则 r_x^U 在 $\mathrm{O}(V,q)$ 中的像是 r_x^V.

从镜射的作用便能推导以下的 Witt 搬运定理.

定理 16.1.12 (E. Witt) 考虑二次型 (V,q). 若 $q(x) = q(y) \neq 0$, 则存在镜射 $r \in \mathrm{O}(V,q)$ 使得 $r(x) = \pm y$; 特别地, 此时总存在 $g \in \mathrm{O}(V,q)$ 使得 $g(x) = y$.

证明 易证 $x + y \perp x - y$. 此外, 初等的公式

$$q(x+y) + q(x-y) = 2q(x) + 2q(y)$$

表明 $x \pm y$ 不可能同时迷向.

设 $q(x-y) \neq 0$, 此时由镜射的描述易得

$$r_{x-y}(x-y) = -x + y, \quad r_{x-y}(x+y) = x + y,$$

于是 $r_{x-y}(x) = y$. 设若 $q(x+y) \neq 0$, 则同理有 $r_{x+y}(x) = -y$. 相应地取 $g = r_{x-y}$ 或 $g = -r_{x+y}$ 即是. □

推论 16.4.7 将给出定理 16.1.12 的进一步推广.

练习 16.1.13 设 (V,q) 为二次型, $\dim V$ 为奇数, 证明有群的互逆同构

$$\begin{aligned} \mathrm{SO}(V,q)\times\{\pm1\} &\underset{\sim}{\overset{\sim}{\rightleftarrows}} \mathrm{O}(V,q)\\ (g,\pm1) &\longmapsto \pm g\\ ((\det h)h, \det h) &\longmapsto h. \end{aligned}$$

练习 16.1.14 给定二次型 (V,q), 其**正交相似变换群**定义为

$$\mathrm{GO}(V,q) := \big\{(g,t)\in \mathrm{GL}(V)\in F^\times : \forall v\in V,\ q(g(v)) = tq(v)\big\};$$

条件也可以等价地写为 $B(gx,gy) = tB(x,y)$ 对所有 $x,y\in V$ 成立, 其中 $B(x,y) := \frac{q(x+y)-q(x)-q(y)}{2}$.

(i) 说明 $\mathrm{GO}(V,q)$ 确实是 $\mathrm{GL}(V)\times F^\times$ 的子群; 当 q 不恒为零时, 它也能通过 $(g,t)\mapsto g$ 视同 $\mathrm{GL}(V)$ 的子群.

(ii) 说明 $(g,t)\mapsto t$ 给出群同态 $\nu:\mathrm{GO}(V,q)\to F^\times$, 称为相似比同态; $\ker(\nu)$ 等同于 $\mathrm{O}(V,q)$.

(iii) 说明 $g\in\mathrm{GO}(V,q)$ 蕴涵 $(\det g)^2 = \nu(g)^{\dim V}$ (请与练习 15.7.9 (iii) 对比).

(iv) 像 $\mathrm{im}(\nu)$ 总是包含 $F^{\times2}$; 举例说明 ν 未必满. 提示 可考虑 $F=\mathbb{R}$ 而 q 正定的例子.

16.2 消去定理与分解定理

本节从关于对角二次型的定义起步.

定义 16.2.1 设 $a_1,\dots,a_n\in F$. 记 n 元二次齐次多项式 $\sum_{i=1}^n a_iX_i^2$ 在 F^n 上给出的二次型为 $\langle a_1,\dots,a_n\rangle$.

因此 $\langle a_1,\dots,a_n\rangle = \langle a_1\rangle\oplus\cdots\oplus\langle a_n\rangle$, 对应的对称矩阵是对角矩阵 $\mathrm{diag}(a_1,\dots,a_n)$, 而

$$\det\langle a_1,\dots,a_n\rangle = \prod_{i=1}^n a_i\cdot F^{\times2}.$$

由此可见 $\langle a_1,\dots,a_n\rangle$ 非退化的充要条件是 $a_1,\dots,a_n\in F^\times$.

我们熟悉如何以配方法将二次型化到 $\langle a_1,\dots,a_n\rangle$ 的形式. 配方具有算法的意味, 但对角化也有简单的抽象证明, 而且结论有所细化.

命题 16.2.2 任何二次型 (V,q) 都同构于形如 $\langle a_1,\dots,a_n\rangle$ 的二次型, 其中 $n=\dim V$. 进一步, 若 $x\in V$ 非迷向, 则 a_1 可以取为 $q(x)$.

证明 对 n 递归地论证. 若 q 恒等于 0 则 $q \simeq \langle 0,\dots,0\rangle$, 构造结束. 否则可取非迷向的 $x \in V \smallsetminus \{0\}$, 对之命 $a_1 := q(x) \in F^\times$. 观察到 Fx 非退化, 故定义–命题表明 $V = Fx \oplus (Fx)^\perp$.

若 $n = 1$ 则 $V = Fx$ 而 $q \simeq \langle a_1\rangle$, 构造结束, 否则对 q 在 $(Fx)^\perp$ 上的限制递归地操作, 可得 $q \simeq \langle a_1\rangle \oplus \langle a_2,\dots\rangle = \langle a_1, a_2,\dots\rangle$. □

以下例子说明对角化涉及的系数并非唯一.

例 16.2.3 设 $a \in F$ 而 $b \in F^\times$, 显见

$$\langle a\rangle \simeq \langle ab^2\rangle;$$

若要求 $a + b \neq 0$, 则

$$\langle a, b\rangle \simeq \langle a+b, ab(a+b)\rangle.$$

论证如下. 将 $\langle a, b\rangle$ 实现在 F^2 上, 取标准基 $\boldsymbol{e}_1, \boldsymbol{e}_2$. 易见 $\boldsymbol{e}_1 + \boldsymbol{e}_2$ 和 $b\boldsymbol{e}_1 - a\boldsymbol{e}_2$ 仍是基, 相互正交, 并且有 $q(\boldsymbol{e}_1 + \boldsymbol{e}_2) = a + b$ 和 $q(b\boldsymbol{e}_1 - a\boldsymbol{e}_2) = ab(a+b)$.

以下结论称为 Witt 的消去定理.

定理 16.2.4 (E. Witt) 设二次型 (V_i, q_i) (其中 $i = 0, 1, 2$) 满足 $q_0 \oplus q_1 \simeq q_0 \oplus q_2$, 而且 q_0 非退化, 则 $q_1 \simeq q_2$.

证明 不妨设 $q_0 = \langle a_1,\dots,a_n\rangle$. 对 n 递归地论证, 可将问题化到 $q_0 = \langle a\rangle$ 的情形 ($a \in F^\times$). 命 $(V, q) := (F, \langle a\rangle) \oplus (V_1, q_1)$ 和 $(V', q') := (F, \langle a\rangle) \oplus (V_2, q_2)$. 按照条件, 存在同构 $\varphi : (V', q') \xrightarrow{\sim} (V, q)$. 定义 V 的元素

$$x := (1, \underbrace{0}_{V_1}), \quad y := \varphi(1, \underbrace{0}_{V_2}); \quad q(x) = a = q(y).$$

于是有

$$V_2 \xrightarrow{\varphi|_{V_2}}_{\sim} (Fy)^\perp \subset V \supset (Fx)^\perp = V_1.$$

综上, 原问题归结为证明给定二次型 (V, q) 和 $x, y \in V$, 若 $q(x) = a = q(y) \in F^\times$, 则必有 $g \in \mathrm{O}(V, q)$ 使得 $g((Fx)^\perp) = (Fy)^\perp$. 然而定理 16.1.12 给出 $g \in \mathrm{O}(V, q)$ 使得 $g(x) = y$, 故 $g((Fx)^\perp) = (Fy)^\perp$. □

定理 16.2.4 中出现了非退化的条件. 就二次型的分类而论, 处理非退化情形已然足够. 何以故? 给定二次型 (V, q), 记其根基为 $R(V)$; 任取子空间 K 使得 $V = R(V) \oplus K$, 相应地有二次型的直和分解

$$(V, q) = (R(V), 0) \oplus (K, q|_K),$$

而且从根基的定义和 $K \cap R(V) = \{0\}$ 易知 $q|_K$ 非退化; 见命题 8.3.6. 事实上, B 诱导 $V/R(V)$ 上的对称双线性形式 $\overline{B} : V/R(V) \times V/R(V) \to F$, 使得同构 $K \xrightarrow{\sim} V/R(V)$

实际是二次型的同构 (这是简单的练习 8.3.7); 因此上述分解中的 $(K, q|_K)$ 的同构类由 (V, q) 唯一确定. 这就将二次型的分类问题简化到非退化情形.

现在着手对非退化二次型作进一步的分解, 为此需要引进一种简单而特殊的二次型.

定义 16.2.5 二元二次多项式 X_1X_2 在 F^2 上给出的非退化二次型称为**双曲平面**[1], 记为 $\mathcal{H}$; 对应的 $q : F^2 \to F$ 为 $q(x_1, x_2) = x_1x_2$, 对应的对称矩阵则是 $\begin{pmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{pmatrix}$.

另以 $m\mathcal{H}$ 代表 $\mathcal{H}^{\oplus m} := \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ (共 m 份), 当 $m = 0$ 时将此理解为零维空间上的二次型, 简称零二次型.

定义 16.2.6 给定非退化二次型 (V, q) 和 V 的子空间 U. 若存在迷向的 $u \in U \smallsetminus \{0\}$, 则称 U 为迷向子空间, 否则称之为非迷向子空间.

若 V 本身是迷向子空间, 则称二次型 (V, q) 或 q 迷向, 否则称之为非迷向二次型.

按照上述定义, 零二次型非迷向.

命题 16.2.7 所有迷向 2 维非退化二次型 (V, q) 皆同构于 $\mathcal{H}$. 更精确地说, 若 $x \in V \smallsetminus \{0\}$ 满足 $q(x) = 0$, 则存在同构 $\varphi : (V, q) \xrightarrow{\sim} \mathcal{H}$ 使得 $\varphi(x) = (1, 0) \in F^2$.

证明 将断言中的 x 扩充为基 x, y, 则必有 $B(y, x) \neq 0$, 否则 x 将属于 B 的根基. 将 y 伸缩后可假设 $B(y, x) = 1$. 命 $b := q(y)$, 则

$$z := \frac{-b}{4}x + \frac{1}{2}y \ \in V$$

满足 $q(z) = 0$ 和 $B(z, x) = \frac{1}{2}$. 易证 $\varphi : rx + sz \mapsto (r, s) \in F^2$ 给出同构 $(V, q) \xrightarrow{\sim} \mathcal{H}$. □

推论 16.2.8 对所有 $a \in F^\times$ 皆有 $\mathcal{H} \simeq \langle a, -a\rangle$.

证明 基于命题 16.2.7, 说明 F^2 上的非退化二次型 $\langle a, -a\rangle$ 迷向即足: 它显然映 $(1, 1)$ 为 0. □

引理 16.2.9 若非退化二次型 q 迷向, 则存在 q' 使得 $q \simeq \mathcal{H} \oplus q'$.

证明 取 $x \neq 0$ 使得 $q(x) = 0$; 按照非退化性质, 存在 $y \notin Fx$ 使得 $B(x, y) \neq 0$, 不难验证 $W := Fx \oplus Fy$ 是非退化子空间, 此外它还是迷向的. 基于命题 16.2.7 可知 $q|_W \simeq \mathcal{H}$. 取 $q' := q|_{W^\perp}$. □

现在可以表述 Witt 的分解定理.

定理 16.2.10 (E. Witt) 所有非退化二次型 q 都有分解 $q \simeq q_{\mathrm{ani}} \oplus m\mathcal{H}$, 其中 q_{ani} 是非迷向的, 而 q_{ani} 的同构类和 $m \in \mathbb{Z}_{\geq 0}$ 由 q 的同构类唯一确定.

[1] 切莫和双曲几何学中的双曲平面混淆.

证明 应用引理 16.2.9 逐步从 q 分离出 $\mathcal{H}$, 直到抵达非迷向二次型为止, 由此得到分解的存在性.

至于唯一性, 设有 $q_{\mathrm{ani}} \oplus m\mathcal{H} \simeq q'_{\mathrm{ani}} \oplus m'\mathcal{H}$, 不失一般性可设 $m \geq m'$. 消去定理 16.2.4 蕴涵 $q_{\mathrm{ani}} \oplus (m-m')\mathcal{H} \simeq q'_{\mathrm{ani}}$, 而右侧非迷向, 故 $m = m'$, 从而 $q_{\mathrm{ani}} \simeq q'_{\mathrm{ani}}$. □

定义 16.2.11 给定非退化二次型 q. 定理 16.2.10 中的分解称为 q 的 Witt 分解, q_{ani} 的同构类称为 q 的**非迷向核**, m 称为 q 的 **Witt 指数**.

例 16.2.12 取 $F = \mathbb{R}$. 因为非零实数的平方和总是正数, 一个非退化二次型非迷向的充要条件是它形如 $\langle 1, \dots, 1\rangle$ 或 $\langle -1, \dots, -1\rangle$. 因为 $\mathcal{H} \simeq \langle 1, -1\rangle$,

$$\underbrace{\langle 1, \dots, 1\rangle}_{a\text{ 份}} \oplus m\mathcal{H} \simeq \underbrace{\langle 1, \dots, 1\rangle}_{a+m\text{ 份}} \oplus \underbrace{\langle -1, \dots, -1\rangle}_{m\text{ 份}},$$
$$\underbrace{\langle -1, \dots, -1\rangle}_{b\text{ 份}} \oplus m\mathcal{H} \simeq \underbrace{\langle 1, \dots, 1\rangle}_{m\text{ 份}} \oplus \underbrace{\langle -1, \dots, -1\rangle}_{b+m\text{ 份}}.$$

于是 Witt 分解定理蕴涵实二次型的惯性定理的非退化版本, 而实二次型的符号差等于 m 或 $-m$, 取决于 q_{ani} 正定或负定.

在 §16.3 将需要以下的技术性结论.

引理 16.2.13 设 q 和 q' 为非迷向非退化二次型, $q \oplus (-q')$ 的 Witt 指数 $\geq n$, 其中 $n \in \mathbb{Z}_{\geq 1}$, 则存在非退化二次型 r, q_1, q'_1 使得

$$\dim r = n, \quad q \simeq r \oplus q_1, \quad q' \simeq r \oplus q'_1.$$

证明 先从 $n = 1$ 情形起步. 此时 $q \oplus (-q')$ 迷向, 故存在 x 和 x' 使得 $q(x) = q'(x') \neq 0$. 此时命题 16.2.2 给出所求的 $r = \langle q(x)\rangle = \langle q'(x')\rangle$ 和 q_1, q'_1.

接着设 $n > 1$. 用前一步结果得到分解 $q \simeq \langle a\rangle \oplus q_2$ 和 $q' \simeq \langle a\rangle \oplus q'_2$. 留意到 q_2 和 q'_2 仍非迷向. 从

$$q \oplus (-q') \simeq \langle a, -a\rangle \oplus q_2 \oplus (-q'_2) \simeq \mathcal{H} \oplus q_2 \oplus (-q'_2)$$

可见 $q_2 \oplus (-q'_2)$ 的 Witt 指数为 $q \oplus (-q')$ 的 Witt 指数减 1. 递归操作. □

16.3 Witt 群

Witt 分解定理 16.2.10 表明非迷向二次型是理解一般的非退化二次型的钥匙, 以下讨论均以此为基础. 回忆到 q_{ani} 代表非退化二次型 q 的非迷向核.

定义 16.3.1 (Witt 等价) 设 q 和 q' 为非退化二次型. 若 $q_{\mathrm{ani}} \simeq q'_{\mathrm{ani}}$, 则称 q 和 q' 为 Witt 等价的, 记为 $q \sim q'$.

举例明之, $m\mathcal{H}$ (见定义 16.2.5) 按此等价于零二次型.

引理 16.3.2 非退化二次型 q 和 q' 满足 $q \sim q'$ 的充要条件是存在 $m, m' \in \mathbb{Z}_{\geq 0}$ 使得 $q \oplus m'\mathcal{H} \simeq q' \oplus m\mathcal{H}$.

证明 取 Witt 分解 $q \simeq q_{\mathrm{ani}} \oplus n\mathcal{H}$ 和 $q' \simeq q'_{\mathrm{ani}} \oplus n'\mathcal{H}$.

若 $q \oplus m'\mathcal{H} \simeq q' \oplus m\mathcal{H}$, 则比较两边的 Witt 分解可得 $q_{\mathrm{ani}} \simeq q'_{\mathrm{ani}}$, 即 $q \sim q'$.

反之若 $q \sim q'$ 则 $q_{\mathrm{ani}} \simeq q'_{\mathrm{ani}}$, 故取 $m = n$ 和 $m' = n'$ 满足所求. □

今后记非退化二次型 q 的 Witt 等价类为 $[q]$.

定义–命题 16.3.3 (Witt 群) 记 $\mathcal{W}(F)$ 为 F 上的所有非退化二次型对 Witt 等价 $\sim$ 的商集[2], 则 $[q_1] + [q_2] := [q_1 \oplus q_2]$ 在 $\mathcal{W}(F)$ 上给出良定义的二元运算, 使之成为交换群, 以零二次型为零元. 我们称 $\mathcal{W}(F)$ 为域 F 的 Witt 群. 它满足 $[-q] = -[q]$.

证明 设 $q_1 \sim q'_1$, 按引理 16.3.2 取 m 和 m' 使得 $q_1 \oplus m'\mathcal{H} \simeq q'_1 \oplus m\mathcal{H}$, 则

$$q_1 \oplus q_2 \oplus m'\mathcal{H} \simeq q'_1 \oplus q_2 \oplus m\mathcal{H},$$

而这又说明 $q_1 \oplus q_2 \sim q'_1 \oplus q_2$. 同理可证若 $q_2 \sim q'_2$ 则 $q_1 \oplus q_2 \sim q_1 \oplus q'_2$.

综上可知 $[q_1] + [q_2] := [q_1 \oplus q_2]$ 是良定义的二元运算. 这使 $\mathcal{W}(F)$ 成为交换幺半群: 结合律, 交换律与零元的性质全部化到二次型与直和的层次来验证. 为了说明这是群, 所需的性质只是 $[-q] + [q] = 0$.

作对角化 $q \simeq \langle a_1, \ldots, a_n \rangle$, 则 $-q \simeq \langle -a_1, \ldots, -a_n \rangle$. 推论 16.2.8 说明 $\langle a_i, -a_i \rangle \simeq \mathcal{H}$ 对所有 $1 \leq i \leq n$ 成立, 由此知 $q \oplus (-q) \simeq n\mathcal{H}$, 故 $[q] + [-q] = 0$. □

推论 16.3.4 非退化二次型 q_1 和 q_2 同构的充要条件是 $[q_1] = [q_2]$ 而且 $\dim q_1 = \dim q_2$.

证明 只须说明充分性. 从 $[q_1] = [q_2]$ 知存在非迷向的 q_{ani} 和 $m_1, m_2 \in \mathbb{Z}_{\geq 0}$ 使得 $q_i \simeq q_{\mathrm{ani}} \oplus m_i\mathcal{H}$ $(i = 1, 2)$. 进一步比较维数可得 $m_1 = m_2$, 故 $q_1 \simeq q_2$. □

练习 16.3.5 说明 $[q_1] = [q_2]$ 蕴涵 $\dim q_1 \equiv \dim q_2 \pmod 2$.

Witt 群也有助于非退化二次型的对角化究竟在何种程度上是唯一的. 以下结论又称为 Witt 等价链定理.

定理 16.3.6 (E. Witt) 设 $n \in \mathbb{Z}_{\geq 1}$. 若非退化二次型 $\langle a_1, \ldots, a_n \rangle$ 和 $\langle b_1, \ldots, b_n \rangle$ 相互同构, 则同构可以通过例 16.2.3 介绍的两类基本同构

$$\left\langle ab^2 \right\rangle \simeq \langle a \rangle, \quad \langle a, b \rangle \simeq \langle a + b, ab(a + b) \rangle$$

的一连串迭代来实现, 每次只涉及 $a_1, \ldots, a_n$ 中的单项或两个相邻项.

[2] 一些读者可能会问: F 上的全体非退化二次型是否成为一个集合? 此处是否在对一个真类取商? 为了排除顾虑, 请留意到同构的二次型必然 Witt 等价, 而且有限维向量空间必有基, 故探讨 Witt 等价类时只需考虑实现在 F^n 上的非退化二次型, n 遍历 $\mathbb{Z}_{\geq 0}$, 而所有这些二次型构成一个集合, 对之取商毫无问题.

证明 对 n 递归地论证. 回忆到 $a_i, b_i \in F^\times$. 当 $n=1$ 时, 容易证明 $\langle a_1\rangle \simeq \langle b_1\rangle$ 当且仅当 $a_1 b_1^{-1} \in F^{\times 2}$.

当 $n=2$ 时, 同构蕴涵有 $x_1, x_2 \in F$ 使得 $b_1 = a_1 x_1^2 + a_2 x_2^2$. 若 $x_2 = 0$ 则 $x_1 \neq 0$ 而 $\langle b_1\rangle = \langle a_1 x_1^2\rangle \simeq \langle a_1\rangle$; 消去定理 16.2.4 又导致 $\langle b_2\rangle \simeq \langle a_2\rangle$, 从而归结为 $n=1$ 的情形. 若 $x_1 = 0$, 按同样方法处理.

接着设 $x_1, x_2 \in F^\times$. 将 a_1 和 a_2 用 $F^{\times 2}$ 调整后, 不妨就取 $x_1 = 1 = x_2$, 亦即 $b_1 = a_1 + a_2$. 我们有

$$\langle a_1, a_2\rangle \simeq \langle a_1 + a_2, a_1 a_2 (a_1 + a_2)\rangle ;$$

在同构 $\langle a_1 + a_2, a_1 a_2 (a_1 + a_2)\rangle \simeq \langle b_1, b_2\rangle$ 中消去 $\langle a_1 + a_2\rangle$, 得 $\langle b_2\rangle \simeq \langle a_1 a_2 (a_1 + a_2)\rangle$, 亦即 b_2 可用 $F^{\times 2}$ 的元素调整为 $a_1 a_2 (a_1 + a_2)$.

现在设 $n \geqslant 3$. 命 $q := \langle a_1, \dots, a_{n-1}\rangle$ 和 $q' := \langle b_1, \dots, b_{n-1}\rangle$, 则 $q \oplus (-q')$ 在 $\mathcal{W}(F)$ 中的像等于 $\langle b_n, -a_n\rangle$ 的像; 因此 $q \oplus (-q')$ 的 Witt 指数 $\geqslant n-2$. 代入引理 16.2.13 可得一个 $n-2$ 维二次型, 不妨记为 $\langle c_1, \dots, c_{n-2}\rangle$, 连同 $a, b \in F^\times$, 使得

$$q \simeq \langle c_1, \dots, c_{n-2}\rangle \oplus \langle a\rangle , \quad q' \simeq \langle c_1, \dots, c_{n-2}\rangle \oplus \langle b\rangle .$$

上式两边分别与 $\langle a_n\rangle$ 和 $\langle b_n\rangle$ 作直和给出同构的二次型; 两边消去 $\langle c_1, \dots, c_{n-2}\rangle$, 便有 $\langle a, a_n\rangle \simeq \langle b, b_n\rangle$.

基于递归假设和 $n=2$ 情形, 通过两类基本同构足以:

- ⋆ 从 $q = \langle a_1, \dots, a_{n-1}\rangle$ 过渡到 $\langle c_1, \dots, c_{n-2}, a\rangle$,
- ⋆ 从 $\langle a, a_n\rangle$ 过渡到 $\langle b, b_n\rangle$,
- ⋆ 从 $\langle c_1, \dots, c_{n-2}, b\rangle$ 过渡到 $q' = \langle b_1, \dots, b_{n-1}\rangle$.

然而这便指明了如何以两类基本同构从 $\langle a_1, \dots, a_n\rangle$ 过渡到 $\langle b_1, \dots, b_n\rangle$, 每次只变动单项或两个相邻项. 证毕. □

与 Witt 群密切相关的另一种群结构是 Grothendieck–Witt 群. 回忆到群运算表作加法的交换群无非是 $\mathbb{Z}$-模.

定义 16.3.7 (Grothendieck–Witt 群) 记 $\mathcal{M}(F)$ 为以 F 上的所有非退化二次型的同构类为基[3)]的自由 $\mathbb{Z}$-模, 其元素唯一地写作有限和 $m_1[\![q_1]\!] + \cdots + m_k[\![q_k]\!]$, 其中 $k \in \mathbb{Z}_{\geqslant 0}$, $m_i \in \mathbb{Z}$ 而 $[\![q_i]\!]$ 表非退化二次型 q_i 的同构类. 定义商模

$$\widehat{\mathcal{W}}(F) := \mathcal{M}(F) \big/ \langle [\![q \oplus q']\!] - [\![q]\!] - [\![q']\!]\rangle .$$

基于对角化, 无论 $\widehat{\mathcal{W}}(F)$ 或 $\mathcal{W}(F)$ 作为交换群都由诸 $\langle a\rangle$ 的等价类生成, 其中 a 遍历 $F^\times$. 问题在于描述这些生成元在 $\widehat{\mathcal{W}}(F)$ 和 $\mathcal{W}(F)$ 中满足哪些关系式. 为此需要一些准备工作.

[3)] 如同定义–命题 16.3.3 的脚注所述, 所有这些同构类确实构成一个集合.

由于 $\dim(q \oplus' q) = \dim q + \dim q'$, 取维数给出 $\mathbb{Z}$-模 (亦即交换群) 的满同态

$$\begin{aligned} \dim : \widehat{\mathcal{W}}(F) &\twoheadrightarrow \mathbb{Z} \\ \textstyle\sum_i m_i[\![q_i]\!] &\longmapsto \textstyle\sum_i m_i \dim q_i. \end{aligned}$$

此外, 还能定义 $\mathbb{Z}$-模的满同态

$$\begin{aligned} \pi : \widehat{\mathcal{W}}(F) &\twoheadrightarrow \mathcal{W}(F) \\ \textstyle\sum_i m_i[\![q_i]\!] &\longmapsto \textstyle\sum_i m_i[q_i]. \end{aligned}$$

诚然, 鉴于 $\mathcal{W}(F)$ 中的等式 $[q \oplus q'] - [q] - [q'] = 0$, 同态 π 良定义; 又由于形如 $[q]$ 的元素生成 $\mathcal{W}(F)$, 同态为满.

引理 16.3.8 上述满同态 π 的核为 $\mathbb{Z}[\![\mathcal{H}]\!] \simeq \mathbb{Z}$.

证明 显然 $\pi([\![\mathcal{H}]\!]) = 0$. 接着考虑 $\sum_i m_i[\![q_i]\!] - \sum_j n_j[\![r_j]\!] \in \widehat{\mathcal{W}}(F)$, 其中 $m_i, n_j \in \mathbb{Z}_{\geq 1}$. 若此元素属于 $\ker(\pi)$, 则 $\sum_i m_i[q_i] = \sum_j n_j[r_j]$, 进而存在 $m, m' \in \mathbb{Z}_{\geq 0}$ 使得

$$\bigoplus_i q_i^{\oplus m_i} \oplus m'\mathcal{H} \simeq \bigoplus_j r_j^{\oplus n_j} \oplus m\mathcal{H}.$$

然而这便说明

$$\sum_i m_i[\![q_i]\!] - \sum_j n_j[\![r_j]\!] = (m - m')[\![\mathcal{H}]\!].$$

综上, $\ker(\pi) = \mathbb{Z}[\![\mathcal{H}]\!]$. 由于 $\dim[\![\mathcal{H}]\!] = 2$, 必然有 $\mathbb{Z}[\![\mathcal{H}]\!] \simeq \mathbb{Z}$. 明所欲证. □

有鉴于此, $\mathcal{W}(F)$ 的另一种等价描述是作为商 $\widehat{\mathcal{W}}(F)/\mathbb{Z}[\![\mathcal{H}]\!]$.

以下是等价链定理 16.3.6 的应用.

推论 16.3.9 定义 $\mathcal{N}(F)$ 为以 $F^\times$ 为基的自由 $\mathbb{Z}$-模; 将每个 $x \in F^\times$ 对应的元素写作 $\langle x \rangle$, 则 $\mathcal{N}(F)$ 的元素唯一地表作有限和 $m_1\langle x_1 \rangle + \cdots + m_k\langle x_k \rangle$, 其中 $k \in \mathbb{Z}_{\geq 0}$, $m_i \in \mathbb{Z}$ 而 $x_i \in F^\times$.

(i) 命 $\widehat{R}$ 为由诸

$$\begin{gathered} \langle ab^2 \rangle - \langle a \rangle \quad (a, b \in F^\times), \\ \langle a \rangle + \langle b \rangle - \langle a + b \rangle - \langle ab(a + b) \rangle \quad (a, b \in F^\times,\ a + b \neq 0) \end{gathered}$$

在 $\mathcal{N}(F)$ 中生成的子模. 我们有 $\mathbb{Z}$-模同构

$$\begin{aligned} \mathcal{N}(F)/\widehat{R} &\xrightarrow{\sim} \widehat{\mathcal{W}}(F) \\ \langle x \rangle + \widehat{R} &\mapsto [\![\langle x \rangle]\!]. \end{aligned}$$

(ii) 命 R 为 $\mathcal{N}(F)$ 的子模 $\widehat{R}+\sum_{a\in F^\times}\mathbb{Z}(\langle a\rangle+\langle -a\rangle)$. 我们有 $\mathbb{Z}$-模同构

$$\begin{aligned}\mathcal{N}(F)/R &\xrightarrow{\sim} \mathcal{W}(F)\\ \langle x\rangle + R &\mapsto [\langle x\rangle].\end{aligned}$$

证明 对于 (i), 问题相当于将形如 $\sum_i a_i[\![\langle x_i\rangle]\!]$ 的元素在 $\widehat{\mathcal{W}}(F)$ 中所能满足的一切关系式通过 $\widehat{R}$ 中的关系来实现. 读者思考半晌, 应当能够明白定理 16.3.6 足以确保这点.

至于 (ii), 引理 16.3.8 表明 $\mathcal{W}(F)\simeq\widehat{\mathcal{W}}(F)/\mathbb{Z}[\![\mathcal{H}]\!]$, 故一切归结为 (i) 与推论 16.2.8 的应用. □

稍后的 §16.6 将赋予 $\mathcal{W}(F)$ 和 $\widehat{\mathcal{W}}(F)$ 环结构, 之后 §16.7 将介绍若干例子.

16.4 全迷向子空间

本节应用 Witt 分解与 Witt 等价研究非退化二次型的全迷向子空间. 首先给出定义.

定义 16.4.1 设 (V,q) 为二次型, U 为 V 的子空间. 若对所有 $u\in U$ 皆有 $q(u)=0$, 则称 U 为**全迷向**的.

子空间 U 全迷向等价于 $U\subset U^\perp$.

引理 16.4.2 给定非退化二次型 (V,q) 的全迷向子空间 U, 二次型 q (等价地看, 对称双线性形式 B) 可以限制到 $U^\perp/U$ 上, 记为 $\overline{q}$ (相应地, $\overline{B}$), 它是非退化的; 由此得到二次型 $(U^\perp/U,\overline{q})$.

证明 从 $U^\perp$ 的定义立见有良定义的对称双线性形式

$$\begin{aligned}\overline{B}: U^\perp/U\times U^\perp/U &\longrightarrow F\\ (x+U,y+U) &\longmapsto B(x,y);\end{aligned}$$

相应地, $\overline{q}(x+U)=\overline{B}(x+U,x+U)=q(x)$.

设陪集 $x+U$ 属于 $\overline{B}$ 的根基, 其中 $x\in U^\perp$, 则上述定义和 B 的非退化条件导致 $x\in(U^\perp)^\perp=U$. 因此 $\overline{B}$ 非退化. □

我们称 $\overline{q}$ 是从 q 诱导而来的. 下述结果是引理 16.2.9 的细化. 回忆到二次型 $\mathcal{H}$ 底层的向量空间是 F^2.

引理 16.4.3 给定非退化二次型 (V,q) 和非零全迷向子空间 $U\subset V$, 记 q 在 $U^\perp/U$ 上诱导的二次型为 $\overline{q}$. 给定 U 的有序基 $x_1,\dots,x_m$, 存在二次型的同构 $\psi: q\xrightarrow{\sim} m\mathcal{H}\oplus\overline{q}$, 使得 $\psi(x_i)$ 是第 i 份 $\mathcal{H}$ 中的 $(1,0)$, 其中 $i=1,\dots,m$.

证明 对 m 递归地操作. 留意到 $(Fx_2 \oplus \cdots \oplus Fx_m)^\perp \supsetneq (Fx_1 \oplus \cdots \oplus Fx_m)^\perp$, 故存在 $y \in (Fx_2 \oplus \cdots \oplus Fx_m)^\perp$ 使得 $B(y, x_1) \neq 0$, 伸缩后可设 $B(y, x_1) = 1$. 按此定义 V 的非退化子空间 $W := Fx_1 \oplus Fy$. 兹断言

$$(Fx_1)^\perp = Fx_1 \oplus W^\perp, \tag{16.4.1}$$

$$U^\perp = Fx_1 \oplus (W^\perp \cap U^\perp), \tag{16.4.2}$$

$$U = Fx_1 \oplus \underbrace{(W^\perp \cap U)}_{=Fx_2 \oplus \cdots \oplus Fx_m}. \tag{16.4.3}$$

诚然, $B(y, x_1) = 1$ 导致 $Fx_1 \cap W^\perp = \{0\}$. 另一方面

$$z \in (Fx_1)^\perp \implies z = B(y,z)x_1 + (z - B(y,z)x_1) \in Fx_1 + W^\perp,$$

综上可得 (16.4.1). 如果上式中要求 $z \in U^\perp \subset (Fx_1)^\perp$, 则易见 $i > 1$ 时也有 $z - B(y,z)x_1 \in (Fx_i)^\perp$, 故 $z - B(y,z)x_1 \in U^\perp$. 由此推得 (16.4.2). 最后, 若进一步要求 $z \in U$, 则 $z - B(y,z)x_1 \in U$, 故推得 (16.4.3).

从 W 非退化可知 $W^\perp$ 亦非退化 (定义–命题 16.1.1). 记 $q_1 := q|_{W^\perp}$. 命题 16.2.7 给出同构 $q|_W \xrightarrow{\sim} \mathcal{H}$ 使得 x_1 映至 $(1,0)$. 于是 $V = W \oplus W^\perp$ 导致有同构 $\psi_1 : q = q|_W \oplus q_1 \xrightarrow{\sim} \mathcal{H} \oplus q_1$, 满足 $\psi_1(x_1) = (1,0) \in F^2$.

若 $m = 1$ 则 (16.4.1) 蕴涵 $q_1 \simeq \overline{q}$, 构造完结. 以下设 $m \geq 2$. 定义 $W^\perp$ 的子空间

$$\overline{U} := W^\perp \cap U = Fx_2 \oplus \cdots \oplus Fx_m,$$

它对 q_1 仍是全迷向子空间; 递归可得同构

$$\psi_2 : q_1 \xrightarrow{\sim} (m-1)\mathcal{H} \oplus q_2$$

其中 q_2 是 q_1 在 $\overline{U}^\perp/\overline{U}$ 上诱导的二次型, 而 $i > 1$ 时 $\psi_2(x_i)$ 是第 $i-1$ 份 $\mathcal{H}$ 中的 $(1,0)$. 取

$$\psi := (\mathrm{id}_{\mathcal{H}}, \psi_2)\psi_1 : q \xrightarrow{\sim} \mathcal{H} \oplus (m-1)\mathcal{H} \oplus q_2 = m\mathcal{H} \oplus q_2.$$

剩余工作是证 $q_2 \simeq \overline{q}$. 从 $x_1 \in W$ 易见 $\overline{U}^\perp = W^\perp \cap U^\perp$. 基于 (16.4.1), (16.4.2) 和 (16.4.3) 按直和取商可得

$$\overline{U}^\perp/\overline{U} = (W^\perp \cap U^\perp)/(W^\perp \cap U) \simeq U^\perp/U, \quad v + \overline{U} \mapsto v + U.$$

这给出二次型的同构 $q_2 \xrightarrow{\sim} \overline{q}$. 明所欲证. □

命题 16.4.4 *给定非退化二次型 (V,q) 的全迷向子空间 U, 则 q 在 $U^\perp/U$ 上诱导的二次型 $\overline{q}$ 和 q 是 Witt 等价的.*

证明 引理 16.3.2 和引理 16.4.3 的直接应用. □

继续假定 (V,q) 是非退化二次型. 所谓 V 的极大全迷向子空间, 是指不被任何全迷向子空间严格包含的全迷向子空间. 基于维数的理由, 它们当然存在.

命题 16.4.5 符号同上. 设 U 是 V 的极大全迷向子空间, 则:

(i) q 在 $U^\perp/U$ 上诱导的二次型 $\overline{q}$ 是非迷向的, 它同构于 q 的非迷向核 q_{ani};

(ii) $\dim U = \frac{1}{2}(\dim q - \dim q_{\mathrm{ani}})$, 这也等于 q 的 Witt 指数.

证明 不可能存在非零之 $\overline{x} \in U^\perp/U$ 使得 $\overline{q}(\overline{x}) = 0$, 否则将 $\overline{x}$ 写作陪集 $x + U$, 便有更大的全迷向子空间 $U + Fx$. 因此 $\overline{q}$ 是非迷向二次型.

命题 16.4.4 给出 Witt 等价 $\overline{q} \sim q$, 而另一方面 $q \sim q_{\mathrm{ani}}$, 故 $\overline{q} \sim q_{\mathrm{ani}}$; Witt 等价的定义 16.3.1 遂表明 $\overline{q} \simeq q_{\mathrm{ani}}$. 断言 (i) 得证.

对于 (ii), 我们有 $\dim U^\perp + \dim U = \dim V$ 和 $\dim U^\perp - \dim U = \dim q_{\mathrm{ani}}$. 两式联立解出 $2\dim U = \dim V - \dim q_{\mathrm{ani}}$. 关于 Witt 指数的断言是 Witt 分解定理 16.2.10 的直接结论. □

定理 16.4.6 设 (V, q) 为非退化二次型, U 和 U' 为 V 的极大全迷向子空间. 选定 U (或 U') 的基 $x_1, \dots, x_m$ (或 $x'_1, \dots, x'_m$), 则存在 $g \in \mathrm{O}(V, q)$ 使得

$$\begin{aligned} &g(U) = U', \\ &\forall i,\ g(x_i) = x'_i. \end{aligned}$$

证明 对 U 的基 $x_1, \dots, x_m$ 应用引理 16.4.3, 搭配命题 16.4.5 得到同构

$$\psi : q \xrightarrow{\sim} m\mathcal{H} \oplus q_{\mathrm{ani}},$$

使得 $\psi(x_i)$ 是第 i 份 $\mathcal{H}$ 中的 $(1, 0)$. 类似地, 考虑 U' 的基 $x'_1, \dots, x'_m$ 也有同构 $\psi' : q \xrightarrow{\sim} m\mathcal{H} \oplus q_{\mathrm{ani}}$ 使得 $\psi'(x'_i)$ 是第 i 份 $\mathcal{H}$ 中的 $(1, 0)$. 现在取

$$g := (\psi')^{-1}\psi \in \mathrm{O}(V, q)$$

即所求. □

以下是定理 16.1.12 的推广.

推论 16.4.7 (E. Witt) 设 (V, q) 为非退化二次型, U_1 和 U_2 是 V 的子空间, 而 $f : (U_1, q|_{U_1}) \xrightarrow{\sim} (U_2, q|_{U_2})$ 是二次型 (容许退化) 的同构. 此时必存在 $g \in \mathrm{O}(V, q)$ 使得 $g|_{U_1} = f$.

证明 分两种情况讨论. 首先设 U_1 不是全迷向子空间, 则存在 $u_1 \in U_1$ 使得 $q(u_1) \neq 0$; 命 $u_2 := f(u_1)$. 定理 16.1.12 说明存在 $k \in \mathrm{O}(V, q)$ 使得 $k(u_1) = u_2$. 以 $U'_1 := k(U_1) \ni u_2$ 和 $f' := fk^{-1}|_{U'_1}$ 代替 $U_1 \ni u_1$ 和 f, 不妨在原问题中假定 $U_1 \cap U_2$ 包含一个非迷向向量 $u \neq 0$ 使得 $f(u) = u$. 基于分解

$$V = Fu \oplus (Fu)^\perp, \quad U_1 = Fu \oplus ((Fu)^\perp \cap U_1), \quad U_2 = Fu \oplus ((Fu)^\perp \cap U_2),$$

并且注意到 f 给出 $(Fu)^\perp \cap U_1 \xrightarrow{\sim} (Fu)^\perp \cap U_2$, 问题进一步化到 $(Fu)^\perp$ 上. 递归处理.

其次设 U_1 是全迷向子空间, 则 U_2 亦然. 取 U_1 的基 $x_1,\dots,x_r$, 相应地 U_2 有基 $x'_1=f(x_1),\dots,x'_r=f(x_r)$, 然后将它们分别扩充为极大全迷向子空间的基 $x_1,\dots,x_m$ 和 $x'_1,\dots,x'_m$. 定理 16.4.5 给出 $g\in \mathrm{O}(V,q)$ 使得 $g(x_i)=x'_i$ 恒成立; 特别地, $g|_{U_1}=f$. 明所欲证. □

16.5 Cartan–Dieudonné 定理

取 (V,q) 为非退化二次型, $V\neq\{0\}$, 相应地有双线性形式 $B(x,y)=\frac{1}{2}(q(x+y)-q(x)-q(y))$. 本节的主定理 16.5.3 将说明正交群 $\mathrm{O}(V,q)$ 能由镜射生成.

恒等映射 id_V 在本节频繁出现, 简记为 1, 零映射 0_V 则简记为 0.

引理 16.5.1 设 $g\in\mathrm{O}(V,q)$, 我们有

(i) $\ker(g-1)=\mathrm{im}(g-1)^\perp$;

(ii) $\ker(g-1)^\perp=\mathrm{im}(g-1)$;

(iii) $(g-1)^2=0$ 当且仅当 $\mathrm{im}(g-1)$ 是 V 的全迷向子空间.

(iv) 对所有 $w\in V$ 皆有

$$B((g-1)w,(g-1)w)=-2B((g-1)w,w)=2B(g(w),(g-1)(w)).$$

证明 考虑 (i). 若 $g(v)=v$ 则对所有 w 皆有 $B(v,g(w)-w)=B(v,g(w))-B(v,w)=B(g^{-1}(v),w)-B(v,w)=0$, 故 $\subset$ 成立. 另一方面,

$$\dim\mathrm{im}(g-1)=\dim V-\dim\ker(g-1),$$

故 $\mathrm{im}(g-1)^\perp$ 和 $\ker(g-1)$ 同维数, $\subset$ 改进为等号.

断言 (ii) 是对 (i) 的两边取 $(\cdots)^\perp$ 的产物.

考虑 (iii). 等式 $(g-1)^2=0$ 等价于 $\mathrm{im}(g-1)\subset\ker(g-1)$, 右边因 (i) 等于 $\mathrm{im}(g-1)^\perp$. 这就说明 $(g-1)^2=0$ 等价于 $\mathrm{im}(g-1)$ 全迷向.

对于 (iv), 观察到

$$\begin{aligned}B\left((g-1)(w),(g-1)(w)\right)&=B(g(w),g(w))-2B(g(w),w)+B(w,w)\\&=2(B(w,w)-B(g(w),w)).\end{aligned}$$

我们有 $B(w,w)-B(g(w),w)=-B((g-1)(w),w)$; 另一方面这也等于 $B(g(w),g(w))-B(g(w),w)=B(g(w),(g-1)(w))$. 证毕. □

引理 16.5.2 假设 $(g-1)^2\neq 0$.

(i) 存在非迷向的 $w\in V\smallsetminus\{0\}$ 使得 $x:=(g-1)(w)$ 或者是 0, 或者非迷向.

(ii) 若 (i) 中的 $x \neq 0$, 则 $r_x g(w) = w$, 其中 $r_x \in \mathrm{O}(V,q)$ 是定义–命题 16.1.8 给出的镜射.

证明 以反证法处理 (i). 设其不成立, 则对所有非迷向的 $w \in V \smallsetminus \{0\}$ 皆有 $x := (g-1)(w)$ 非零且迷向; 特别地, w 和 x 线性无关. 引理 16.5.1 (iv) 蕴涵 $B(x,w) = 0$. 由此可见二次型限制在子空间 $Fw \oplus Fx$ 上退化, 故 $\dim V \geq 3$.

兹断言

$$\forall w \in V, \quad B(g(w), w) = B(w, w). \tag{16.5.1}$$

为此不妨设 $w \neq 0$. 若 w 非迷向, 前一段论证已给出 (16.5.1). 以下设 w 迷向. 由 $\dim V \geq 3$ 可知 $(Fw)^\perp \supsetneq Fw$. 引理 16.4.2 表明 q 在 $(Fw)^\perp/Fw$ 上诱导非退化二次型; 在其中任取非迷向的非零向量, 则它在 $(Fw)^\perp \smallsetminus \{0\}$ 中的原像 y 是与 w 正交的非迷向向量; 于是 (16.5.1) 的已知情形说明 $B(g(y), y) = B(y, y)$.

继续命 $u := w + ty$, 其中 $t \in F^\times$; 从

$$q(u) = B(w+ty, w+ty) = t^2 B(y,y) \neq 0$$

可得 $u \in V \smallsetminus \{0\}$ 非迷向. 于是 $B(g(u), u) = B(u,u)$, 亦即

$$\forall t \in F^\times, \quad B(g(w) + tg(y), w+ty) - B(w+ty, w+ty) = 0.$$

等式左侧整理为

$$(B(g(y),y) - B(y,y))t^2 + (\cdots)t + B(g(w),w) - B(w,w).$$

已知 t^2 的系数为零, 又因此多项式有 $|F^\times| \geq 2$ 个根 (因为 $\mathrm{char}(F) \neq 2$), 常数项也必然为零, (16.5.1) 得证.

结合 (16.5.1) 与引理 16.5.1 (iii) 和 (iv) 推得 $(g-1)^2 = 0$, 与假设矛盾.

以下处理 (ii). 问题相当于证 $gw - \frac{2B(x,g(w))}{B(x,x)}x = w$, 亦即

$$x = \frac{2B(x,g(w))}{B(x,x)}x;$$

然而上式已由引理 16.5.1 (iv) 料理. □

定理 16.5.3 (É. Cartan, J. Dieudonné) 设 (V,q) 为非退化二次型, 则所有 $g \in \mathrm{O}(V,q)$ 都能写成至多 $\dim V$ 个镜射的乘积.

证明 记 $n := \dim V$. 对 n 递归地论证. 当 $n = 1$ 时 $\mathrm{O}(V,q) = \{\pm 1\}$ 而 -1 是唯一的镜射. 以下设 $n > 1$. 兹断言

$$(\exists w \in \ker(g-1),\ q(w) \neq 0) \implies g \text{ 是至多 } n-1 \text{ 个镜射的乘积.} \tag{16.5.2}$$

诚然, 此时 g 来自子群 $\mathrm{O}((Fw)^\perp)$, 故 g 可在 $\mathrm{O}((Fw)^\perp)$ 中写成至多 $n-1$ 个镜射的乘积, 参看练习 16.1.11.

接着分情况讨论. 首先假定 $(g-1)^2 \neq 0$. 取引理 16.5.2 (i) 中的 $w \neq 0$ 和 $x = (g-1)(w)$.

⋆ 假若 $x = 0$, 则 (16.5.2) 表明 g 是至多 $n-1$ 个镜射的乘积.

⋆ 假若 $x \neq 0$, 则 x 非迷向, 而引理 16.5.2 (ii) 表明 $r_x g(w) = w$, 故 (16.5.2) 将 $r_x g$ 表为至多 $n-1$ 个镜射的乘积, 于是 $g = r_x(r_x g)$ 是至多 n 个镜射的乘积.

以下处理 $(g-1)^2 = 0$ 的情况. 假如存在非迷向的 $w \in \ker(g-1) \smallsetminus \{0\}$, 则 (16.5.2) 可资应用. 因此以下进一步假定 $\ker(g-1)$ 全迷向; 这也相当于

$$\ker(g-1) \subset \ker(g-1)^{\perp} \overset{\text{引理 16.5.1 (ii)}}{=\!=} \operatorname{im}(g-1).$$

又因为 $(g-1)^2 = 0$ 蕴涵 $\operatorname{im}(g-1) \subset \ker(g-1)$, 遂有 $\ker(g-1) = \ker(g-1)^{\perp}$, 所以 $n = 2\dim\ker(g-1)$.

由于 g 在 $\ker(g-1)$ 上是恒等, 而且它在 $V/\ker(g-1) = V/\operatorname{im}(g-1)$ 上诱导的线性映射也是恒等, 故此时 $\det g = 1$.

任取镜射 s, 则 $\det(sg) = -1$ 说明 sg 属于此前处理过的其他情况, 从而能表为至多 n 个镜射的乘积. 于是 $g = s(sg)$ 表为至多 $n+1$ 个镜射的乘积; 因为 $\det g = 1$ 而 $n \in 2\mathbb{Z}$, 它不可能是 $n+1$ 个镜射的乘积. 明所欲证. □

Cartan–Dieudonné 定理 16.5.3 中的镜射个数是最优的, 原因如下. 设 $1 \leq m < n := \dim V$ 并考虑 m 个镜射 $r_1, \dots, r_m$; 每个 r_i 都在某个 $n-1$ 维子空间 V_i 上为恒等, 故 $r_1 \cdots r_m$ 在维数至少为 $n - m \geq 1$ 的子空间 $\bigcap_{i=1}^{m} V_i$ 上为恒等. 所以任何不含特征值 1 的 $g \in \mathrm{O}(V,q)$ (例如 $-\mathrm{id}$) 都无法表为 $r_1 \cdots r_m$.

16.6 环结构: 二次型的张量积

设 (V,q) 和 (V',q') 为二次型, 对应的双线性形式仍记为 B 和 B'. 相应地有线性映射

$$\begin{aligned} \alpha : V \otimes V &\to F & \qquad \alpha' : V' \otimes V' &\to F \\ x \otimes y &\mapsto B(x,y), & x' \otimes y' &\mapsto B'(x',y'). \end{aligned}$$

基于张量积的结合约束 (命题 15.2.2), 交换约束 (命题 15.2.4) 和幺约束 (命题

15.2.3) 可得

$$\begin{array}{ccc}
(V\otimes V')\otimes(V\otimes V') & & (x\otimes x')\otimes(y\otimes y') \\
\wr\downarrow & & \downarrow \\
(V\otimes V)\otimes(V'\otimes V') & & (x\otimes y)\otimes(x'\otimes y') \\
{\scriptstyle \alpha\otimes\alpha'}\downarrow & & \downarrow \\
F\otimes F & & B(x,y)\otimes B'(x',y') \\
\wr\downarrow & & \downarrow \\
F & & B(x,y)B'(x',y').
\end{array} \tag{16.6.1}$$

上式的合成对应到双线性映射

$$\begin{aligned}(V\otimes V')\times(V\otimes V') &\to F\\ (x\otimes x', y\otimes y') &\mapsto B(x,y)B'(x',y'),\end{aligned}$$

记之为 $B\otimes B'$. 它仍然是对称的.

定义–命题 16.6.1 将上述对称双线性形式 $B\otimes B'$ 所对应的二次型记为 $q\otimes q'$: $V\otimes V'\to F$, 它满足

$$\begin{aligned}(q\otimes q')(x\otimes x') &= B(x,x)B'(x',x')\\ &= q(x)q'(x').\end{aligned}$$

此外, $q\otimes q'$ 非退化的充要条件是 q 和 q' 皆非退化.

证明 唯一待证的是关于非退化性质的部分. 不失一般性可设 $q=\langle a_1,\dots,a_m\rangle$ 而 $q'=\langle b_1,\dots,b_n\rangle$. 易见 $B\otimes B'$ 对应于对角元为 $\prod_{i,j}a_ib_j$ 的 $mn\times mn$ 对角矩阵. □

容易看出

$$\begin{aligned}q\otimes(q'\otimes q'') &\simeq (q\otimes q')\otimes q'',\\ q\otimes q' &\simeq q'\otimes q,\\ q\otimes(q_1'\oplus q_2') &\simeq (q\otimes q_1')\oplus q\otimes q_2',\\ (q_1\otimes q_2)\otimes q' &\simeq (q_1\otimes q')\oplus(q_2\otimes q').\end{aligned}$$

在向量空间层次, 上述典范同构全是 §15.2 探讨的内容. 剩下的只是说明这些同构和二次型匹配, 毫无困难. 此外,

$$q\otimes\langle t\rangle\simeq tq\simeq\langle t\rangle\otimes q,\quad t\in F.$$

引理 16.6.2 给定二次型 (V,q) 和 (V',q'). 设 U (或 U') 为 V (或 V') 的全迷向子空间, 则 $U\otimes V'$ (或 $V\otimes U'$) 嵌入为 $V\otimes V'$ 的全迷向子空间.

证明 考虑 U 情形即可. 命题 15.2.9 (ii) 说明 $U\hookrightarrow V$ 诱导的线性映射 $U\otimes V'\to V\otimes V'$ 为单, 给出所求嵌入. 全迷向性质直接来自定义–命题 16.6.1 的公式. □

定义–命题 16.6.3 运算 $\otimes$ 可以定义在 Witt 等价类的层次, 使得 $\mathcal{W}(F)$ 对已有的加法和 $\otimes$ 成为交换环, 以 $\langle 1\rangle$ 的等价类为乘法幺元. 此结构称为域 F 的 **Witt 环**[4].

类似地, $\widehat{\mathcal{W}}(F)$ 对已有的加法和 $\otimes$ 也成为交换环, 称为域 F 的 **Grothendieck–Witt 环**, 而 $\widehat{\mathcal{W}}(F)\to\mathcal{W}(F)$ 是环的满同态, 以 $\mathbb{Z}[\![\mathcal{H}]\!]$ 为核.

证明 关于 $\widehat{\mathcal{W}}(F)$ 的情形是容易的. 我们定义 $[\![q_1]\!]\otimes[\![q_2]\!]:=[\![q_1\otimes q_2]\!]$, 然后按照对每个变元的加性将二元运算 $\otimes$ 延拓到 $\widehat{\mathcal{W}}(F)$ 的所有元素. 乘法运算所需的结合律, 分配律, 交换律全部化到二次型的层次来检验.

引理 16.3.8 将群 $\mathcal{W}(F)$ 等同于群 $\widehat{\mathcal{W}}(F)$ 对 $\mathbb{Z}[\![\mathcal{H}]\!]$ 的商. 为了将交换环的结构从 $\widehat{\mathcal{W}}(F)$ 下降到 $\mathcal{W}(F)$, 证明 $\mathbb{Z}[\![\mathcal{H}]\!]$ 是 $\widehat{\mathcal{W}}(F)$ 的理想即可.

考虑二次型 $q\otimes m\mathcal{H}$, 其中 $m\in\mathbb{Z}_{\geq 1}$. 由于 $m\mathcal{H}$ 有 m 维全迷向子空间, 引理 16.6.2 蕴涵 $q\otimes m\mathcal{H}$ 有 $m\dim q$ 维全迷向子空间. 将之扩充为极大全迷向子空间, 然后代入命题 16.4.5; 比较维数可知 $q\otimes m\mathcal{H}$ 的 Witt 分解必然形如 $(m\dim q)\mathcal{H}$. 明所欲证. □

练习 16.6.4 说明可按 $\overline{\dim}[q]:=\dim q \mod 2$ 定义映射 $\overline{\dim}:\mathcal{W}(F)\to\mathbb{Z}/2\mathbb{Z}$. 它是环同态, 而

$$I_F:=\ker(\overline{\dim})=\sum_{a\in F^\times}\mathbb{Z}\left[\langle 1,-a\rangle\right].$$

提示 显然诸 $\langle a,b\rangle$ 生成 I_F, 然而 $\langle a,b\rangle$ 与 $\langle 1,a\rangle\oplus(-\langle 1,-b\rangle)$ Witt 等价.

张量积的另一个功能是变换二次型的域, 这部分涉及 §15.4 的构造.

定义–命题 16.6.5 设 F 为域 E 的子域. 给定有限维 F-向量空间 V 和双线性形式 $B:V\times V\to F$, 在 E-向量空间 $V_E:=E\underset{F}{\otimes}V$ 上存在唯一的双线性形式 B_E 使得

$$B_E(s\otimes x,t\otimes y)=stB(x,y),\quad s,t\in E,\ x,y\in V.$$

进一步, B 非退化 (或对称, 反对称) 当且仅当 B_E 亦然. 对于对应到非退化双线性形式 B 的二次型 (V,q), 我们因而得到域 E 上的二次型 (V_E,q_E), 它对应到 B_E 并满足

$$q_E(t\otimes x)=t^2q(x),\quad t\in E,\ x\in V.$$

证明 以 $(s,x,t,y)\mapsto stB(x,y)$ 定义四重线性映射

$$E\times V\times E\times V\to E.$$

它对应到 F-线性映射 $(E\underset{F}{\otimes}V)\underset{F}{\otimes}(E\underset{F}{\otimes}V)\to E$, 由此又得到 F-双线性映射

$$\begin{aligned}B_E:V_E\times V_E&\to E\\(s\otimes x,t\otimes y)&\mapsto stB(x,y).\end{aligned}$$

[4] 莫和 Witt 向量构成的环 [10, §10.9] 混淆, 两者无关.

容易看出 B_E 还是 E-双线性的, 而且上式唯一确定了 B_E.

现在选取 V 的基 $v_1,\dots,v_n$ 将 B 等同于矩阵 $\boldsymbol{A}\in \mathrm{M}_{n\times n}(F)$, 则 $a_{ij}=B(v_i,v_j)$. 配合上式立见 B_E 对应到 $\boldsymbol{A}$ 在 $\mathrm{M}_{n\times n}(E)$ 中的像. 因此 B 非退化 (或对称, 反对称) 当且仅当 B_E 亦然. 关于 $q_E(t\otimes x)=B_E(t\otimes x,t\otimes x)$ 的公式是自明的. □

显然 $q\simeq q' \implies q_E\simeq q'_E$. 如果 $q\simeq\langle a_1,\dots,a_n\rangle$, 则在 E 的层次也有 $q_E\simeq\langle a_1,\dots,a_n\rangle$. 取 $q=\mathcal{H}$ 则容易看出 $\mathcal{H}_E$ 等同于 E 上的双曲平面.

非迷向的 q 可能给出迷向的 q_E (例如考虑 $E=\mathbb{C}$ 和 $F=\mathbb{R}$). 所以 $q\mapsto q_E$ 在 Witt 环的层次应当取道 Grothendieck–Witt 环来考量. 重点是以下结论.

命题 16.6.6 *设 F 为域 E 的子域, 则有环同态如下:*

$$\begin{aligned}\widehat{\mathcal{W}}(F)&\to\widehat{\mathcal{W}}(E) & \mathcal{W}(F)&\to\mathcal{W}(E)\\ [\![q]\!]&\mapsto[\![q_E]\!], & [q]&\mapsto[q_E].\end{aligned}$$

证明 第一个同态源自 Grothendieck–Witt 环的定义和明显的性质

$$(q\oplus q')_E\simeq q_E\oplus q'_E,\quad (q\otimes q')_E\simeq q_E\otimes q'_E,\quad \langle 1\rangle_E\simeq\langle 1\rangle.$$

第二个同态则是因为 $\mathcal{H}_E$ 是 E 上的双曲平面, 而 Witt 环是 Grothendieck–Witt 环对 $\mathbb{Z}[\![\mathcal{H}]\!]$ 的商环. 这一描述不涉及二次型的非迷向核. □

16.7 具体实例

鉴于 Witt 分解定理 16.2.10, Witt 环的结构足以确定 F 上的非退化二次型以及其间的直和与张量积运算, 相关研究的内涵丰富, 牵涉的技术也同样广泛, 本节仅介绍最简单的几则特例.

例 16.7.1 因为 $\mathbb{R}$ 上的非迷向非退化二次形只有 $\langle 1,\dots,1\rangle$ 和 $\langle -1,\dots,-1\rangle$, 故 $\mathcal{W}(\mathbb{R})$ 是由 $\langle 1\rangle$ 的 Witt 等价类生成的循环群, $\mathcal{W}(\mathbb{R})\simeq\mathbb{Z}$; 由于 $\langle 1\rangle\otimes\langle 1\rangle\simeq\langle 1\rangle$, 实际上这还是环同构.

例 16.7.2 对于 F 代数闭的情形, 例如 $F=\mathbb{C}$, 所有非退化二次型都能对角化为 $\langle 1,\dots,1\rangle$, 故 $\mathcal{W}(F)$ 仍是由 $\langle 1\rangle$ 的 Witt 等价类生成的循环群, 而且基于维数奇偶性的缘由, $[\langle 1\rangle]\neq 0$. 然而 $\langle 1,1\rangle\simeq\langle 1,-1\rangle\simeq\mathcal{H}$, 故 $\mathcal{W}(F)\simeq\mathbb{Z}/2\mathbb{Z}$; 从 $\langle 1\rangle\otimes\langle 1\rangle\simeq\langle 1\rangle$ 同样可见这是环同构.

接着探讨有限域的 Witt 环, 需要一些简单准备. 回忆到对任意域 E 及其子域 F, 通过域的乘法可将 E 作成 F-向量空间, 按此定义扩域次数 $[E:F]$ 为 $\dim_F E$.

若有限域 F 满足 $\mathrm{char}(F)\neq 2$, 则 $|F|$ 是奇数. 这属于域论常识, 解释如下.

1. 首先 $\mathrm{char}(F)\neq 0$, 否则 $\mathbb{Q}$ 嵌入为 F 的子域, 与 F 有限相矛盾.

2. 因此 $p := \mathrm{char}(F)$ 是奇素数, 而 $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ 嵌入为 F 的子域; 将 F 视为 $\mathbb{F}_p$-向量空间, 计数可得 $|F| = p^{[F:\mathbb{F}_p]}$, 特别地 $|F|$ 是奇数.

接着确定 $F^\times$ 中的 $F^{\times 2}$-轨道.

引理 16.7.3 *设 F 为有限域, $\mathrm{char}(F) \neq 2$, 则 $|F^\times/F^{\times 2}| = 2$, 写作*

$$F^\times/F^{\times 2} = \left\{1 \cdot F^{\times 2}, a \cdot F^\times\right\},$$

其中

- *当 $|F| \equiv 3 \pmod 4$ 时取 $a = -1$,*
- *当 $|F| \equiv 1 \pmod 4$ 时任取 $a \in F^\times \smallsetminus F^{\times 2}$, 此时有 $-1 \in F^{\times 2}$.*

证明 有限域的乘法可逆元群总是循环群 (第十一章带提示的习题), 故 $F^\times \simeq \mathbb{Z}/(|F|-1)\mathbb{Z}$. 既然 $|F|-1$ 是偶数, 按此得到群同构

$$F^\times/F^{\times 2} \simeq \frac{\mathbb{Z}/(|F|-1)\mathbb{Z}}{2\mathbb{Z}/(|F|-1)\mathbb{Z}} \simeq \mathbb{Z}/2\mathbb{Z},$$

见命题 11.9.17. 于是有 $F^\times/F^{\times 2} = \left\{1 \cdot F^{\times 2}, a \cdot F^\times\right\}$, 其中 $a \in F^\times \smallsetminus F^{\times 2}$.

剩下的问题是判定何时能取 $a = -1$. 观察到 $X^2 - 1 = 0$ 在 F 中恰有两根 ± 1, 故 -1 是群 $F^\times$ 中唯一的 2 阶元. 于是 $-1 \in F^{\times 2}$ 当且仅当群 $F^\times$ 有 4 阶元; 由于 $F^\times \simeq \mathbb{Z}/(|F|-1)\mathbb{Z}$, 这也等价于 4 整除 $|F|-1$. 明所欲证. □

定义 16.7.4 *设 (V, q) 为任意域 F (满足 $\mathrm{char}(F) \neq 2$) 上的二次型, $d \in F^\times$. 若存在 $x \in V$ 使得 $q(x) = d$, 则称 q 能够表出 d.*

由于对所有 $t \in F$ 皆有 $q(tx) = t^2 q(x)$, 能否用 q 表出只和 d 的 $F^{\times 2}$-轨道相关. 举例明之, 双曲平面 $\mathcal{H}$ 对应的二次型 $(x_1, x_2) \mapsto x_1 x_2$ 显然能表出 $F^\times$ 的所有元素.

引理 16.7.5 *考虑满足 $\mathrm{char}(F) \neq 2$ 的有限域 F.*

(i) *任何满足 $\dim q = 2$ 的非退化二次型 q 都能表出 $F^\times$ 的所有元素.*

(ii) *任何满足 $\dim q \geq 3$ 的非退化二次型皆迷向.*

证明 考虑断言 (i). 基于对角化 (命题 16.2.2), 引理 16.7.3 及其符号, q 的同构类至多只有三种选法:

$$q_1 = \langle 1, 1\rangle, \quad q_2 = \langle a, a\rangle, \quad q_3 = \langle 1, a\rangle.$$

显然 q_3 既能表出 1 又能表出 a, 故能表出 $F^\times$ 的所有元素. 由于 q_1 和 q_2 仅差一个伸缩, 而 q_1 显然能表出 1, 故说明 q_1 能表出 a 即可.

若 $-1 \in F^{\times 2}$, 则 $q_1 \simeq \langle 1, -1\rangle \simeq \mathcal{H}$ 能表出所有元素.

若 $-1\notin F^{\times 2}$, 考虑 F 的子集 $F^{\times 2}$ 和 $1+F^{\times 2}$; 由于 $1\notin 1+(F^{\times 2})$, 两者不等而基数相同, 故存在 $1+x^2\in(1+F^{\times 2})\smallsetminus F^{\times 2}$. 注意到 $-1\notin F^{\times 2}$ 确保 $1+x^2\neq 0$, 故存在 $b\in F^\times$ 使得 $a=b^2(1+x^2)=b^2+(bx)^2$.

对于断言 (ii), 在 $\dim q\geq 3$ 时通过对角化将 q 写成 $q'\oplus q''$, 其中 $\dim q'=2$. 任取 $y''\neq 0$ 使得 $q''(y'')\neq 0$, 再由 (i) 取 y' 使得 $q'(y')=-q''(y'')$, 便由 $q((y',y''))=0$ 知 q 迷向. □

命题 16.7.6 沿用引理 16.7.3 关于有限域 F 的符号. 另将加法群 $\mathbb{Z}/2\mathbb{Z}$ 表作 $\{\overline{0},\overline{1}\}$, 其中 $\overline{k}:=k+2\mathbb{Z}$.

- ⋆ 当 $|F|\equiv 3\pmod 4$ 时有环同构 $\mathcal{W}(F)\xrightarrow{\sim}\mathbb{Z}/4\mathbb{Z}$, 映 $\langle 1\rangle$ 的类为同余类 $1+4\mathbb{Z}$.
- ⋆ 当 $|F|\equiv 1\pmod 4$ 时有环同构 $\mathcal{W}(F)\xrightarrow{\sim}\mathbb{F}_2[\mathbb{Z}/2\mathbb{Z}]$, 其中

$$\mathbb{F}_2[\mathbb{Z}/2\mathbb{Z}]:=\text{以 }\mathbb{Z}/2\mathbb{Z}\text{ 为基的 }\mathbb{F}_2\text{-向量空间},$$
$$\text{环的乘法}:\ (x\overline{0}+y\overline{1})(x'\overline{0}+y'\overline{1})=(xx'+yy')\overline{0}+(xy'+x'y)\overline{1},$$

而同构映 $\langle 1\rangle$ (或 $\langle a\rangle$) 的类为 $\overline{0}$ (或 $\overline{1}$).

证明 在 $|F|\equiv 3\pmod 4$ 的情况, 观察到 $\langle 1\rangle\not\simeq\langle -1\rangle$, 而且 $\langle 1,1\rangle\not\simeq\mathcal{H}$ (否则由消去定理 16.2.4 将有 $\langle 1\rangle\simeq\langle -1\rangle$). 兹断言 F 上的非迷向非退化二次型仅有

$$0,\quad \langle 1\rangle,\quad \langle -1\rangle,\quad \langle 1,1\rangle.$$

诚然, 引理 16.7.5 (ii) 说明考虑不超过 2 维的二次型即足, 而上述选项之外仅有 $\langle 1,-1\rangle$ 和 $\langle -1,-1\rangle$. 前者同构于 $\mathcal{H}$ 故迷向. 兹断言后者同构于 $\langle 1,1\rangle$. 这是因为引理 16.7.5 (i) 说明 $\langle -1,-1\rangle$ 能表出 1, 故存在 $b\in F^\times$ 使得 $\langle -1,-1\rangle\simeq\langle 1,b\rangle$ (命题 16.2.2 后半部分), 比较行列式立见 $b\in F^{\times 2}$.

我们顺带得到 $\langle 1,1,1,1\rangle\simeq\langle 1,1\rangle\oplus\langle -1,-1\rangle\simeq 2\mathcal{H}$, 故 $\langle 1,1,1\rangle\sim\langle -1\rangle$.

综合以上描述, 易得 $\langle 1\rangle\mapsto 1+4\mathbb{Z}$ 确定环同构 $\mathcal{W}(F)\xrightarrow{\sim}\mathbb{Z}/4\mathbb{Z}$.

接着考虑 $|F|\equiv 1\pmod 4$ 的情况. 此时有 $\langle -1\rangle\simeq\langle 1\rangle\not\simeq\langle a\rangle$ 和 $\langle 1,a\rangle\not\simeq\mathcal{H}$ (用消去定理), 而且 F 上的非迷向非退化二次型仅有

$$0,\quad \langle 1\rangle,\quad \langle a\rangle,\quad \langle 1,a\rangle.$$

论证比前一情形更简单, 因为 $-1\in F^{\times 2}$ 确保 $\langle 1,1\rangle\simeq\mathcal{H}\simeq\langle a,a\rangle$.

综上易从 $\langle 1\rangle\mapsto\overline{0}$ 和 $\langle a\rangle\mapsto\overline{1}$ 确定群同构 $\mathcal{W}(F)\xrightarrow{\sim}\mathbb{F}_2[\mathbb{Z}/2\mathbb{Z}]$. 为了给出乘法在 $\mathbb{F}_2[\mathbb{Z}/2\mathbb{Z}]$ 上的反映, 留意到 $\langle a\rangle\otimes\langle a\rangle\simeq\langle a^2\rangle\simeq\langle 1\rangle\simeq\langle 1\rangle\otimes\langle 1\rangle$ 即足. □

关于 $\mathbb{F}_2[\mathbb{Z}/2\mathbb{Z}]$ 的构造是所谓群代数的一则特例, 详见 [10, 定义 5.6.3].

16.8 域上的 Hermite 形式

本节旨在说明如何将先前的二次型理论推及 Hermite 或反 Hermite 形式的情形. 此处的 Hermite 形式不再限于从前考虑的复向量空间, 而容许在更广的域 E 上操作; 这种 Hermite 形式在数论等应用中频繁而自然地出现. 有必要先引入一系列概念.

以下设 E 为域, $\mathrm{char}(E)\neq 2$, 而 $\tau: E\xrightarrow{\sim} E$ 是满足 $\tau^2=\mathrm{id}$ 的域自同构. 定义

$$F:=\{x\in E:\tau(x)=x\}.$$

引理 16.8.1 在上述前提下, F 是 E 的子域, 而且扩张次数 $[E:F]$ (可参考 §16.7 的回顾) 或者等于 1 (此时 $\tau=\mathrm{id}$ 而 $E=F$), 或者等于 2 (此时 $\tau\neq\mathrm{id}$).

证明 由 τ 是同构这一性质容易看出 F 是 E 的子环; 若 $x\in F\smallsetminus\{0\}$, 则对 $x^{-1}x=1=xx^{-1}$ 两边同取 τ 可得 $\tau(x^{-1})x=1=x\tau(x^{-1})$, 从而逆元唯一性蕴涵 $\tau(x^{-1})=x^{-1}$, 亦即 $x^{-1}\in F$. 综上, F 是子域.

由于 $\mathrm{char}(E)\neq 2$ 而 $\tau: E\to E$ 是 F-线性映射, 从 $\tau^2=\mathrm{id}$ 可得 F-向量空间的直和分解

$$E=E_1\oplus E_{-1},\quad E_{\pm 1}:=\{x\in E:\tau(x)=\pm x\},$$

分解具体写作 $y=\frac{y+\tau(y)}{2}+\frac{y-\tau(y)}{2}$; 留意到 $E_1=F$.

如果 $E=F$, 则 $\tau=\mathrm{id}$ 而 $[E:F]=1$. 如果 $E\neq F$, 任取 $a\in E_{-1}\smallsetminus\{0\}$, 则由 $\tau(x)=-x\implies xa^{-1}\in F$ 可知 $E_{-1}=Fa$; 此时 $\tau\neq\mathrm{id}$ 而 $E=F\oplus Fa$ 蕴涵 $[E:F]=2$. □

经典的例子是取 $E=\mathbb{C}$ 而 τ 为复共轭 $a+b\mathrm{i}\mapsto a-b\mathrm{i}$; 此时 $F=\mathbb{R}$.

尽管上述定义中的 F 是从 E 和 τ 派生的对象, 今后将以扩域的符号 $E|F$ 来总括上述资料. 本章习题将进一步明确扩域 $E\supset F$ 与 τ 的关系.

以下记 $\otimes:=\underset{E}{\otimes}$.

定义 16.8.2 (相对于 $E|F$ 的半双线性映射和半双线性形式) 选定 $E|F$ 如上. 设有 E-向量空间 V,W,X. 所谓从 $V\times W$ 到 X 的**半双线性映射**, 是指满足以下条件的映射

$$B: V\times W\to X$$

⋆ 它对第一个变元是半线性的, 亦即:

$$B(v_1+v_2,w)=B(v_1,w)+B(v_2,w),$$
$$B(tv,w)=\tau(t)B(v,w).$$

⋆ 它对第二个变元是线性的:

$$B(v,w_1+w_2)=B(v,w_1)+B(v,w_2),$$
$$B(v,tw)=tB(v,w).$$

以上默认 $t \in E$. 全体半双线性映射 $V \times W \to X$ 对逐点的加法和纯量乘法构成 E-向量空间, 记为 $\mathrm{Sesq}_{E|F}(V, W; X)$.

对于 $X = E$ 的特例, 半双线性映射 $B : V \times W \to E$ 也称为 $V \times W$ 上的**半双线性形式**, 它们构成的 E-向量空间记为 $\mathrm{Sesq}_{E|F}(V, W)$.

今后仅关注半双线性形式, 而非更广的半双线性映射.

一如 $E = \mathbb{C}$ 的情形, 通过引入 V 的 τ-扭曲版本 $\overline{V}$, 使其底层的加法群 V 不变而纯量乘法 $\odot : E \times V \to V$ 定义为 $t \odot v := \tau(t)v$, 可以将半双线性形式 $V \times W \to E$ 等同于 E-双线性形式 $\overline{V} \times W \to E$, 或等同于 E-线性映射 $\overline{V} \otimes W \to E$.

定义 16.8.3 设 V 是 E-向量空间, $\epsilon \in \{\pm 1\}$. 若 $B \in \mathrm{Sesq}_{E|F}(V, V)$ 满足

$$B(x, y) = \epsilon\tau(B(y, x)), \quad x, y \in V,$$

则称 B 是相对于 $E|F$ 的 ϵ-Hermite 形式. 我们将 $(+1)$-Hermite 形式简称为 V 上的 **Hermite 形式**, 将 (-1)-Hermite 形式简称为 V 上的**反 Hermite 形式**.

给定 $\epsilon \in \{\pm 1\}$. 对于一列 ϵ-Hermite 形式 $(V_1, B_1), \ldots, (V_k, B_k)$, 能在 $\bigoplus_{i=1}^{k} V_i$ 上构造它们的直和 (或称正交直和); 在 ϵ-Hermite 形式之间也有同构的概念. 定义方式都是自明的.

例 16.8.4 当 $E = F$ 时, Hermite 形式 (或反 Hermite 形式) 回归于对称 (或反对称) 双线性形式, 两者的性状颇为不同. 与此相反, 若 $[E : F] = 2$ 而 $a \in E^\times$ 满足 $\tau(a) = -a$, 则 B 是 Hermite 形式当且仅当 aB 是反 Hermite 形式, 两种概念只差一个伸缩.

例 16.8.5 当 $E = \mathbb{C}$ 而 τ 为复共轭时 $F = \mathbb{R}$, 上述定义回归 §10.1 介绍过的概念.

今后聚焦于有限维 E-向量空间上的半双线性形式. 具体取 $V = E^m$, $W = E^n$, 其元素视同列向量. 对 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{m\times n}(E)$, 记

$$^\dagger\boldsymbol{A} = (\tau(a_{ji}))_{i,j} \in \mathrm{M}_{n\times m}(E).$$

一如 §10.1, 我们有向量空间的同构

$$\begin{aligned} \mathrm{M}_{m\times n}(E) &\xrightarrow{\sim} \mathrm{Sesq}_{E|F}(E^m, E^n) \\ \boldsymbol{A} &\longmapsto [B(\boldsymbol{v}, \boldsymbol{w}) := {}^\dagger\boldsymbol{v}\,\boldsymbol{A}\,\boldsymbol{w}] \\ (B(\boldsymbol{e}_i, \boldsymbol{e}_j))_{i,j} &\longmapsfrom B \end{aligned}$$

使得 ϵ-Hermite 形式对应到满足 $^\dagger\boldsymbol{A} = \epsilon\boldsymbol{A}$ 的 $n \times n$ 矩阵. 直和对应于从满足 $^\dagger\boldsymbol{A}_i = \epsilon\boldsymbol{A}_i$ 的一列矩阵 $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ 构造分块对角矩阵 $\mathrm{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k)$.

对于半双线性形式 B 也能够定义何谓左根, 右根和非退化 (有限维情形); B 非退化的充要条件是取基后对应的矩阵 $\boldsymbol{A}$ 可逆, 而 ϵ-Hermite 形式的左根等于右根, 统称为**根基**.

实现在零空间上的半双线性形式称为**零形式**; 按照定义, 零形式非退化, 而且它既是 Hermite 又是反 Hermite 的.

定义–命题 16.8.6 命 $\mathrm{N}_{E|F}: E^\times \to F^\times$ 为 $x \mapsto x\tau(x)$ 定义的群同态[5]. 给定 ϵ-Hermite 形式 (V,B), 使得 $n := \dim V \in \mathbb{Z}_{\geq 1}$. 取基后将 B 等同于 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(E)$, 则 $\det \boldsymbol{A}$ 的 $\mathrm{N}_{E|F}(E^\times)$-轨道无关基的选取, 可称之为 (V,B) 的行列式; (V,B) 非退化当且仅当其行列式非零.

另规定零形式的行列式为轨道 $\mathrm{N}_{E|F}(E^\times)$.

证明 和二次型情形类似, 换基相当于将 $\boldsymbol{A}$ 换作 ${}^\dagger\boldsymbol{C}\,\boldsymbol{A}\,\boldsymbol{C}$, 其中 $\boldsymbol{C} \in \mathrm{GL}(n,E)$, 相应地

$$\det({}^\dagger\boldsymbol{C}\,\boldsymbol{A}\,\boldsymbol{C}) = \tau(\det \boldsymbol{C}) \det \boldsymbol{A} \det \boldsymbol{C} = \mathrm{N}_{E|F}(\det \boldsymbol{C}) \det \boldsymbol{A}.$$

关于非退化性质的断言已包含于先前的讨论. □

定义 16.8.7 对于 ϵ-Hermite 形式 (V,B), 其中的 V 今后默认为有限维的, 我们定义 (V,B) 的维数为 $\dim V$, 并且按照 §16.1 的方式定义

- ⋆ V 的元素 x,y (或子空间 U_1,U_2) 之间的正交关系 $x \perp y$ (或 $U_1 \perp U_2$);
- ⋆ 子空间 U 的正交空间 $U^\perp := \{x \in V : \forall u \in U,\ B(x,u) = 0\}$;
- ⋆ 非退化子空间 U 的概念, 这相当于说 $(U, B|_{U\times U})$ 非退化;
- ⋆ 群 $\mathrm{U}(V,B) := \{g \in \mathrm{GL}(V) : \forall x,y \in V,\ B(gx,gy) = B(x,y)\}$ 称为 (V,B) 的**酉群**. 若选基将 V 等同于 E^n, 将 B 等同于 $\boldsymbol{A} \in \mathrm{M}_{n\times n}(E)$, 则

$$\mathrm{U}(V,B) \overset{\text{等同于}}{=\!=} \{\boldsymbol{C} \in \mathrm{GL}(n,E) : {}^\dagger\boldsymbol{C}\,\boldsymbol{A}\,\boldsymbol{C} = \boldsymbol{A}\};$$

这是 $\mathrm{GL}(n,E)$ 的子群.

一如二次型的情形, 若 U 非退化则有 $V = U \oplus U^\perp$. 当 $E = F$ 而 $\epsilon = 1$ 时, $\mathrm{U}(V,B)$ 即是正交群.

约定 16.8.8 给定 V 上的 ϵ-Hermite 形式 B, 命 $h(x) := B(x,x)$, 则 $B(x,y)+B(y,x) = h(x+y) - h(x) - h(y)$.

- ⋆ 在 $E = F$ 而 $\epsilon = 1$ 的情形, h 按此确定 B;

[5] 事实上是乘法幺半群的同态 $E \to F$.

⋆ 在 $[E:F]=2$ 的情形, 任取 $a\in E\smallsetminus F$, 则从

$$\begin{aligned} B(x,y) &+ \epsilon\tau(B(x,y)) &&= h(x+y)-h(x)-h(y)\\ \tau(a)B(x,y) &+ \epsilon a\tau(B(x,y)) &&= h(ax+y)-h(ax)-h(y)\end{aligned} \tag{16.8.1}$$

可以反解 $B(x,y)$.

综上, 在这两类情形下以 (V,h) 或 h 来标记 ϵ-Hermite 形式是合理而且方便的. 对应的酉群表为

$$\mathrm{U}(V,h)=\{g\in\mathrm{GL}(V):\forall x\in V,\ h(g(x))=h(x)\};$$

另外定义

$$\mathrm{SU}(V,h):=\mathrm{U}(V,h)\cap\mathrm{SL}(V).$$

注意到对所有满足 $\tau(a)\in\{a,-a\}$ 的 $a\in E^\times$ 皆有 $\mathrm{U}(V,h)=\mathrm{U}(V,ah)$, 因此对于酉群的研究而言, Hermite 形式和反 Hermite 形式并无差别.

练习 16.8.9 仿照例 16.1.14 的模式, 定义**酉相似变换群** $\mathrm{GU}(V,h)$ 及相似比同态 $\nu:\mathrm{GU}(V,h)\to F^\times$, 并且推导类似的性质, 例如 $\mathrm{N}_{E|F}(E^\times)\subset\mathrm{Im}(\nu)$.

16.9 Hermite 形式的 Witt 理论

上一节介绍了关于 Hermite 形式的基本概念. 本节将介绍如何相应地推广关于二次型的许多结论. 我们继续选定域 E 和自同构 τ, 详细条件和 §16.8 相同; 特别地, $\mathrm{char}(E)\neq 2$, 而 E 的 τ-不动点构成子域 F, 满足 $[E:F]\leq 2$. 取 $\epsilon\in\{\pm1\}$. 继续默认 ϵ-Hermite 形式都是有限维的.

定义 16.9.1 给定 ϵ-Hermite 形式 (V,B) 和对应的 $h:V\to E$. 若 $x\in V$ 满足 $h(x)=0$ 则称 x 是**迷向**的, 否则称 x 为**非迷向**的.

注意: $\tau(h(x))=\epsilon h(x)$ 总是成立; 当 $E=F$ 而 $\epsilon=-1$ 时所有 x 皆迷向, 故下一陈述可排除此情况.

定义–命题 16.9.2 给定 ϵ-Hermite 形式 (V,B) 和对应的 h. 设 $x\in V$ 非迷向, 而 $u\in E^\times$ 满足 $\mathrm{N}_{E|F}(u)=1$. 对所有 $v\in V$ 定义

$$r_{x,u}(v)=r^V_{x,u}(v):=v+(u-1)\frac{B(x,v)}{h(x)}x,$$

则 $r_{x,u}\in\mathrm{U}(V,h)$. 我们称 $r_{x,u}$ 为以直线 Ex 为法向的 u-**镜射**.

证明 显然 $r_{x,u}\in\mathrm{End}(V)$. 若 $v=ax+y$, 其中 $a\in E$ 而 $y\in(Ex)^\perp$, 则有 $r_{x,u}(v)=uax+y$, 由此立见 $r_{x,u}\in\mathrm{U}(V,h)$. □

留意到 $r_{x,u}$ 只和 Ex 相关; $u=1$ 给出 id_V, 而 $u=-1$ 在 $E=F$ 时给出定义–命题 16.1.8 的变换.

接着将对角二次型的定义推及 ϵ-Hermite 形式.

定义 16.9.3 设 $a_1,\dots,a_n\in E$ 满足 $\tau(a_i)=\epsilon a_i$. 对角矩阵 $\mathrm{diag}(a_1,\dots,a_n)$ 在 F^n 上给出的 ϵ-Hermite 形式记为 $\langle a_1,\dots,a_n\rangle$, 它也等于 $\langle a_1\rangle\oplus\cdots\oplus\langle a_n\rangle$.

因此 $\langle a_1,\dots,a_n\rangle$ 非退化当且仅当 $a_1,\dots,a_n\in E^\times$ (定义–命题 16.8.6).

定义–命题 16.9.4 包含非零迷向向量的 2 维非退化 ϵ-Hermite 形式存在, 而且彼此同构; 在同构类中选定由矩阵 $\boldsymbol{A}=\begin{pmatrix}&1/2\\ \epsilon/2&\end{pmatrix}$ 确定的代表元 $\mathcal{H}$, 称之为**双曲平面**.

证明 首先观察到矩阵 $\boldsymbol{A}$ 确实在 E^2 上给出非退化 ϵ-Hermite 形式, 使得 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 皆迷向. 这就解决了存在性.

至于唯一性, 处理 $[E:F]=2$ 情形即可, 这是因为当 $E=F$ 时, $\epsilon=1$ 情形已知, 而 $\epsilon=-1$ 情形包含于辛形式的分类.

还可以进一步设 $\epsilon=1$, 这是因为当 $\epsilon=-1$ 时可取 $a\in E^\times$ 使得 $\tau(a)=-a$, 然后以 a 伸缩将问题化到 $\epsilon=1$ 情形.

现在照搬命题 16.2.7 论证, 即可证明唯一性. □

定理 16.9.5 以下结论适用于 ϵ-Hermite 形式 (V,h), 其中要求或者 $[E:F]=2$, 或者 $E=F$ 而 $\epsilon=1$.

▷ **非迷向向量的搬运定理** 设 $h(x)=h(y)\neq 0$, 则存在 $g\in\mathrm{U}(V,h)$ 使得 $g(x)=y$ (参照定理 16.1.12).

▷ **对角化** 存在 $a_1,\dots,a_n$, 其中 $n:=\dim V$, 使得 $(V,h)\simeq\langle a_1,\dots,a_n\rangle$, 而且若 $x\in V$ 非迷向则可取 $a_1=h(x)$ (参照命题 16.2.2).

▷ Witt **消去定理** 设有 $h_0\oplus h_1\simeq h_0\oplus h_2$ 而 h_0 非退化, 则 $h_1\simeq h_2$ (参照定理 16.2.4).

▷ Witt **分解定理** 设 h 非退化, 则有分解 $h\simeq h_{\mathrm{ani}}\oplus m\mathcal{H}$, 其中 h_{ani} 非迷向 (亦即其迷向向量只有 0), $m\mathcal{H}:=\mathcal{H}^{\oplus m}$, 而 h_{ani} 和 $m\in\mathbb{Z}_{\geq 0}$ 由 h 的同构类唯一确定 (参照定理 16.2.10).

上述分解称为 h 的 Witt 分解, h_{ani} 的同构类称为 h 的**非迷向核**, m 称为 h 的 **Witt 指数**.

证明 只需考虑 $[E:F]=2$ 的情形, 而且 $\epsilon=-1$ 的情形还能进一步以伸缩化到 $\epsilon=1$ 的情形处理. 论证基本与二次型情形无异, 只是有时需要调整.

以下举非迷向向量的搬运为例, 其论证较定理 16.1.12 复杂. 首先从 $h(x+y)+h(x-y)=2h(x)+2h(y)$ 可见 $x+y$ 和 $x-y$ 必有一者非迷向; 因为 $-\mathrm{id}\in \mathrm{U}(V,h)$, 必要时以 $-y$ 代 y, 不妨设 $x-y$ 非迷向. 取分解 $y=ax+y'$, 其中 $a\in E$ 而 $y'\perp x$, 则

$$h(x)=h(y)=\mathrm{N}_{E|F}(a)h(x)+h(y'),$$

由此得到

$$\begin{aligned}0\neq h(x-y)&=h((1-a)x-y')=h((1-a)x)+h(y')\\&=\mathrm{N}_{E|F}(1-a)h(x)+(1-\mathrm{N}_{E|F}(a))h(x)=(2-a-\tau(a))h(x);\end{aligned}$$

于是 $a\neq 1$, 而

$$B(x-y,x)=B((1-a)x,x)=(1-\tau(a))h(x)\neq 0.$$

命 $u=1-\frac{h(x-y)}{B(x-y,x)}$. 将 $h(x-y)$ 和 $B(x-y,x)$ 用 $h(x)$ 表达, 可得

$$u=\frac{a-1}{1-\tau(a)}\in E^\times,\quad \tau(u)=u^{-1}.$$

因此有 u-镜射 $r_{x-y,u}\in \mathrm{U}(V,h)$ (定义–命题 16.9.2). 易见 $r_{x-y,u}(x)=y$. 非迷向向量的搬运得证. □

注记 16.9.6 Witt 消去定理和分解定理对 $E=F$ 而 $\epsilon=-1$ 的情形同样适用, 此时所有向量皆迷向, 先前的非迷向核 h_{ani} 应当替换为零形式, 而相应的消去与分解定理都能由辛形式的分类定理处理.

对于 $E=\mathbb{C}$, $F=\mathbb{R}$ 而 $\epsilon=1$ 的情形, Witt 分解蕴涵 Hermite 形式的惯性定理, 解释方法与例 16.2.12 如出一辙.

接着将 Witt 群推及 ϵ-Hermite 形式. 既然有 Witt 分解在手, 我们可以推广定义 16.3.1 的 Witt 等价, 从而得到以下概念.

定义–命题 16.9.7 *设 $[E:F]=2$ 或者 $E=F$ 而 $\epsilon=1$. 记 $\mathcal{W}^\epsilon(E|F)$ 为相对于 $E|F$ 的所有非退化 ϵ-Hermite 形式对 Witt 等价 $\sim$ 的商集, $[h_1]+[h_2]:=[h_1\oplus h_2]$ 在 $\mathcal{W}^\epsilon(E|F)$ 上给出良定义的二元运算, 使之成为交换群, 以零形式为零元. 它满足 $[-h]=-[h]$.*

定义 16.3.7 的 Grothendieck–Witt 群也有相应版本 $\widehat{\mathcal{W}}^\epsilon(E|F)$, 而 $\mathcal{W}^\epsilon(E|F)$ 同构于 $\widehat{\mathcal{W}}^\epsilon(E|F)$ 对 $\mathbb{Z}[\![\mathcal{H}]\!]$ 的商, 这是引理 16.3.8 的推广. 当 $[E:F]=2$ 时, 以满足 $\tau(a)=-a$ 的 $a\in E^\times$ 伸缩给出群同构 $\mathcal{W}^\epsilon(E|F)\simeq\mathcal{W}^{-\epsilon}(E|F)$ 和 $\widehat{\mathcal{W}}^\epsilon(E|F)\simeq\widehat{\mathcal{W}}^{-\epsilon}(E|F)$.

练习 16.9.8 上述讨论排除了 $E=F$ 而 $\epsilon=-1$ 的情形, 但它实则最为简单. 依循注记 16.9.6 的思路, 试从辛形式的分类定理说明此时的 Grothendieck–Witt 群 $\widehat{\mathcal{W}}^-(F|F)$ 同构于 $\mathbb{Z}$, 而 Witt 群 $\mathcal{W}^-(F|F)$ 应当取为平凡群.

全迷向子空间的概念 (定义 16.4.1) 也能直接照搬到非退化 ϵ-Hermite 形式.

定理 16.9.9 设 $[E:F]=2$ 或者 $E=F$ 而 $\epsilon=1$. 设 (V,h) 是非退化 ϵ-Hermite 形式, 而 U 是 V 的极大全迷向子空间, 则:

(i) h 在 $U^\perp/U$ 上诱导的 ϵ-Hermite 形式 $\overline{h}$ 同构于 h_{ani};

(ii) $\dim U=\frac{1}{2}(\dim h-\dim h_{\mathrm{ani}})$, 这也等于 h 的 Witt 指数.

此外, 给定极大全迷向子空间 U (或 U') 及其基 $x_1,\dots,x_m$ (或 $x'_1,\dots,x'_m$), 存在 $g\in \mathrm{U}(V,h)$ 使得 $g(U)=U'$ 而且 $g(x_i)=x'_i$ 对所有 $1\leq i\leq m$ 成立.

证明 和命题 16.4.5 与定理 16.4.6 相同. □

留意到在 $E=F$ 而 $\epsilon=-1$ 的情形, 上述陈述中的 h_{ani} 应当替换为零形式, 而一切断言都化为关于辛空间的 Lagrange 子空间的已知性质.

推论 16.9.10 在定理 16.9.9 的场景中, 设 U_1 和 U_2 是 V 的子空间, 而 $f:(U_1,h|_{U_1})\xrightarrow{\sim}(U_2,h|_{U_2})$ 是 ϵ-Hermite 形式 (容许退化) 的同构. 此时必存在 $g\in \mathrm{U}(V,h)$ 使得 $g|_{U_1}=f$.

证明 基于此前的结论, 论证和推论 16.4.7 无异. □

其实推论 16.9.10 也适用于 $E=F$ 而 $\epsilon=-1$ 的情形, 相当于在辛空间 (V,B) 中讨论, 细节留作本章习题.

关于 ϵ-Hermite 形式的乘法结构与二次型情形略有差异. 给定半双线性形式 (V,B) 和 (V',B'), 先前的构造 (16.6.1) 对此改写为

$$\begin{array}{ccc}(\overline{V}\otimes\overline{V'})\otimes(V\otimes V') & & (x\otimes x')\otimes(y\otimes y')\\ \downarrow & & \downarrow\\ E & & B(x,y)B'(x',y').\end{array}\tag{16.9.1}$$

引理 16.9.11 有 E-向量空间的典范同构 $\overline{V\otimes V'}\simeq\overline{V}\otimes\overline{V'}$, 其刻画为 $x\otimes x'\mapsto x\otimes x'$, 其中 $x\in V$ 而 $x'\in V'$.

证明 取道泛性质验证如下. 对所有 E-向量空间 L, 易见 $\mathrm{Bil}(\overline{V},\overline{V'};L)=\overline{\mathrm{Bil}(V,V';\overline{L})}$, 从而有

$$\begin{aligned}\mathrm{Hom}(\overline{V\otimes V'},L)\simeq\overline{\mathrm{Hom}(V\otimes V',\overline{L})}\simeq\overline{\mathrm{Bil}(V,V';\overline{L})}\\ =\mathrm{Bil}(\overline{V},\overline{V'};L)\simeq\mathrm{Hom}(\overline{V}\otimes\overline{V'},L).\end{aligned}$$

其余验证都是例行公事. □

基于引理 16.9.11 和张量积的泛性质, (16.9.1) 便对应于半双线性形式

$$\begin{aligned}B\otimes B':(V\otimes V')\times(V\otimes V')&\to E\\ (x\otimes x',y\otimes y')&\mapsto B(x,y)B'(x',y').\end{aligned}$$

因为 E 的乘法交换, $(V\otimes V',B\otimes B')\simeq(V'\otimes V,B'\otimes B)$.

命题 16.9.12 设 $\epsilon,\epsilon'\in\{\pm1\}$ 而 B (或 B') 为 ϵ-Hermite (或 ϵ'-Hermite) 形式, 照例默认为有限维的, 则上述构造给出的半双线性形式 $B\otimes B'$ 是 $\epsilon\epsilon'$-Hermite 形式; 在 $[E:F]=2$ 或者 $E=F$ 而 $\epsilon=\epsilon'=1$ 的前提下, 它也能合理地记为 $h\otimes h'$.

证明 从 $(B\otimes B')(x\otimes x',y\otimes y')=B(x,y)B'(x',y')$ 验证. □

从 $h\otimes h'$ 的表法也容易看出 $\langle a_1,\dots,a_m\rangle\otimes\langle b_1,\dots,b_n\rangle$ 同构于所有 $\langle a_ib_j\rangle$ 的正交直和; 特别地, $\det(h\otimes h')=\det(h)\det(h')$, 而 $h\otimes h'$ 非退化的充要条件是 h 和 h' 皆非退化.

引理 16.9.13 给定 ϵ-Hermite 形式 (V,B) 和 ϵ'-Hermite 形式 (V',B'). 设 U (或 U') 为 V (或 V') 的全迷向子空间, 则 $U\otimes V'$ (或 $V\otimes U'$) 嵌入为 $V\otimes V'$ 的全迷向子空间.

证明 和引理 16.6.2 的论证全然相同. □

命题 16.9.14 张量积运算 $\otimes$ 给出映射

$$\mathcal{W}^{\epsilon}(E|F)\times\mathcal{W}^{\epsilon'}(E|F)\to\mathcal{W}^{\epsilon\epsilon'}(E|F),\quad \widehat{\mathcal{W}}^{\epsilon}(E|F)\times\widehat{\mathcal{W}}^{\epsilon'}(E|F)\to\widehat{\mathcal{W}}^{\epsilon\epsilon'}(E|F),$$

这些运算对每个变元都是加性的 (或称具有 “双加性”).

证明 张量积对每个变元都保持正交直和, 由此得到 Grothendieck–Witt 群上的张量积运算, 具有双加性.

为了将张量积运算从 $\widehat{\mathcal{W}}^{\epsilon}(E|F)$ 下降到 $\mathcal{W}^{\epsilon}(E|F)$, 照搬定义–命题 16.6.3 的论证即可, 差别仅在于用引理 16.9.13 替代原来的引理 16.6.2. □

特别地, $\widehat{\mathcal{W}}^{+}(E|F)\oplus\widehat{\mathcal{W}}^{-}(E|F)$ 和 $\mathcal{W}^{+}(E|F)\oplus\mathcal{W}^{-}(E|F)$ 对 $\otimes$ 成环, 此处的 $\oplus$ 是 $\mathbb{Z}$-模亦即加法群的直和, 而 Hermite 形式 $\langle1\rangle$ 充当了乘法幺元的角色..

练习 16.9.15 具体写下同构 $\mathcal{W}^{+}(\mathbb{C}|\mathbb{R})\simeq\mathbb{Z}\simeq\mathcal{W}^{-}(\mathbb{C}|\mathbb{R})$ 并描述乘法运算.

16.10 环上的 Hermite 形式概观

本节旨在说明如何将 §16.8 的理论推及更一般的环, 只勾勒理论概貌, 不追求完整的推导和示例. 读者如果对相关理论感兴趣, 应当参考专著如 [3, Chapter 7]; 该处的理论还适用于 2 在环中的像不可逆的情形, 使得本章的许多基本定理在适当修改后依然成立.

定义 16.10.1 设 R 为环, 如果映射 $\tau:R\to R$ 满足 $\tau^2=\mathrm{id}_E$, $\tau(1)=1$ 和

$$\tau(x+y)=\tau(x)+\tau(y),\quad \tau(xy)=\tau(y)\tau(x),\quad x,y\in R,$$

则称 τ 为 R 的**对合**.

例 16.10.2 上述定义不要求 R 交换. 对于 $\mathbb{R}$ 上的四元数代数 $\mathbb{H}$, 共轭运算 $a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}\mapsto a-b\mathrm{i}-c\mathrm{j}-d\mathrm{k}$ 给出对合; 对于交换环 C 上的矩阵环 $\mathrm{M}_{n\times n}(C)$, 转置运算 $\boldsymbol{A}\mapsto {}^{\mathrm{t}}\boldsymbol{A}$ 给出对合.

对合 τ 给出从 R 到相反环 R^{op} 的同构. 因此任何左 (或右) R-模 M 都通过 $mr:=\tau(r)m$ (或 $rm=m\tau(r)$) 成为右 (或左) R-模, 另记为 $\overline{M}$.

另一方面, 左 (或右) R-模同态构成的加法群 $\mathrm{Hom}(M,R)$ 通过 $(fr)(m)=f(m)\cdot r$ (或 $(rf)(m)=r\cdot f(m)$) 成为右 (或左) R-模, 其中 $r\in R$ 而 $f\in\mathrm{Hom}(M,A)$. 记此模为 $M^\vee$.

因此 $M\mapsto\overline{M}$ 和 $M\mapsto M^\vee$ 提供左右切换的两种途径, 前者涉及对合 τ. 实际操作时经常用到以下的构造.

定义–命题 16.10.3 *给定 R 的对合 τ. 设 M 为右 R-模, 则有右 R-模的同构*

$$\begin{aligned}\overline{M}^\vee &\xrightarrow{\sim}\overline{M^\vee}\\ f&\mapsto[m\mapsto\tau(f(m))]\,.\end{aligned}$$

记 $M^*:=\overline{M^\vee}$, 则有自然的右 R-模同态

$$\varpi_M:M\to(M^*)^*,$$

它映 $m\in M$ 为 $f\mapsto\tau(f(m))$, 其中 $f\in\overline{M^\vee}$.

证明 操演定义; 逆映射 $\overline{M^\vee}\to\overline{M}^\vee$ 的写法同样是 $g\mapsto[m\mapsto\tau(g(m))]$. □

定义–命题 16.10.4 *给定 R 的对合 τ, 右 R-模 M 上的*半双线性形式[6]*意谓映射*

$$B:M\times M\to R,$$

要求 B 对每个变元都具有加性, 而且

$$B(xr,ys)=\tau(r)B(x,y)s,\quad r,s\in R,\ x,y\in M.$$

指定 M 上的半双线性映射也相当于指定 R-模同态 $\psi:M\to M^*$, 刻画为

$$\psi(x)(y)=B(x,y).$$

证明 用 $M^*=\overline{M^\vee}$ 的定义来检验 B 和 ψ 之间的对应. □

以下概念是有限维向量空间在一般环上的合适类比; 我们仅陈述右模版本, 左模情形是完全类似的.

[6] 鉴于左右切换的操作, 更科学的方法或许是先对左 R-模 N 和右 R-模 M 定义 $N\times M$ 上的双线性形式为满足双加性和 $B(rx,ys)=rB(x,y)s$ 的映射 $B:N\times M\to R$. 引入对合 τ, 则 M 上的半双线性形式无非是 $\overline{M}\times M$ 上的双线性形式.

定义 16.10.5 设 R 为任意环, M 为右 R-模. 如果存在有限秩自由右 R-模 L, 使得 M 能嵌入为 L 的某个直和项, 则称 M 为**有限生成投射模**.

当 R 是除环, 有限生成右 R-模总是自由的, 换言之总有基, 此外所有基的元素个数皆有限且相等, 这点和有限维向量空间类似. 因此除环上的有限生成模总是有限生成投射模, 它们和域上的向量空间一样具有维数 (或称为秩) 的概念.

为了简化, 后续定义仅在 2 可逆的前提下表述. 回忆到 $Z(R)$ 代表环 R 的中心.

定义 16.10.6 设 $2 \in R^\times$. 给定 R 的对合 τ 和满足 $\epsilon\tau(\epsilon) = 1$ 的 $\epsilon \in Z(R)$, 如果右 R-模 M 上的半双线性形式 B 满足恒等式

$$B(x, y) = \epsilon\tau(B(y, x)), \quad x, y \in M,$$

则称之为 M 上的 ϵ-Hermite 形式. 若进一步要求 M 是有限生成投射模, 则称上述资料 (M, B) 为 **ϵ-Hermite 模**; 若 $\epsilon = 1$ (或 $\epsilon = -1$), 也称此为 Hermite (或反 Hermite) 模.

对 ϵ-Hermite 模可以定义正交直和, 同构与自同构的概念. 此外还能定义 $\mathrm{U}(M, B)$, 它是 M 作为 R-模的自同构群的子群.

练习 16.10.7 对于 $M = R^{\oplus n}$ 的情形 ($n \in \mathbb{Z}_{\geq 1}$), 以 R 上的矩阵具体描述 $\mathrm{U}(M, B)$.

提示 参照定义 16.8.7 的描述; 将非交换环 R 上的 $n \times n$ 可逆矩阵定义为矩阵环 $\mathrm{M}_{n\times n}(R)$ 的可逆元, 由此定义群 $\mathrm{GL}(n, R)$. 为此, 需要先将右 R-模 $R^{\otimes n}$ 的自同态等同于矩阵左乘.

定义 16.10.8 设 $2 \in R^\times$ 而 (M, B) 是 ϵ-Hermite 模. 如果对应的 $\psi : M \to M^*$ 是同构 (或单射), 则称 (M, B) 为**正则** (或**非退化**) 的.

读者应当完成以下简单练习.

练习 16.10.9 沿用定义 16.10.6 的符号.

(i) 证明 $\tau(Z(R)) = Z(R)$.

(ii) 设 $a \in Z(R)^\times$, 则 $\mu := \frac{a}{\tau(a)}$ 满足 $\mu\tau(\mu) = 1$. 证明 (M, B) 是 ϵ-Hermite 的当且仅当 (M, aB) 是 $\epsilon\mu$-Hermite 的.

(iii) 设 $Z(R)$ 为域. 给定 Hermite 形式 (M, B), 命 $h(x) = B(x, x)$; 适当修改 (16.8.1) (代入 $\epsilon = 1$) 以说明在 B 由 h 唯一确定.

例 16.10.10 (域的情形) 取 R 为满足 $\mathrm{char}(E) \neq 2$ 的域 E; 对合相当于域 E 的自同构 τ, 要求 $\tau^2 = \mathrm{id}$, 因此上述定义悉数化约为 §16.8 已有的理论; 特别地, 比较维数立见正则等价于非退化. 差别仅在于此处仅要求关系式

$$B(x, y) = \epsilon\tau(B(y, x)), \quad x, y \in V,$$

中的 $\epsilon \in E^\times$ 满足 $\epsilon\tau(\epsilon) = 1$ (这是 B 不恒为零的必要条件). 容许一般的 ϵ 并不带来新理论, 原因是当 $E = F$ 时 $\epsilon = \pm 1$, 而当 $[E : F] = 2$ 时域论中的 Hilbert 第 90 定理 [10, 定理 9.6.9] 给出 $a \in E^\times$ 使得 $\epsilon^{-1} = \frac{a}{\tau(a)}$, 从而练习 16.10.9 说明 aB 是 Hermite 形式.

例 16.10.11 (四元数情形) 取例 16.10.2 的除环 $R = \mathbb{H}$ 连同对合 τ. 因为 τ 在 $Z(\mathbb{H}) = \mathbb{R}$ 上作用平凡, 以下考虑 $\epsilon \in \{\pm 1\}$ 即可.

首先取 $\epsilon = -1$. 正则反 Hermite 模 (M, B) 的分类简单明了: (M, B) 的同构类完全由 M 的维数确定, 详见 [3, Chapter 10, 3.7 Theorem].

对于 $\epsilon = 1$ 的情形, 正则 Hermite 模的分类与 $\mathbb{R}$ 上的二次型分类具有相似的样貌. 以下给出详细的表述与证明.

定理 16.10.12 *考虑四元数除环 $\mathbb{H}$ 及共轭运算给出的对合 τ. 对所有 $a \in \mathbb{R}$ 定义 $\mathbb{H}$ 上的 Hermite 形式*

$$B_a(x, y) = \tau(x)ay,$$

并且将 $(\mathbb{H}, B_a)$ 简记为 $\langle a\rangle$, 然后命 $\langle a_1, \dots, a_n\rangle = \langle a_1\rangle \oplus \cdots \oplus \langle a_n\rangle$. 我们有双射

$$\begin{aligned} \left\{(p, q) \in \mathbb{Z}_{\geq 0}^n\right\} &\xrightarrow{1:1} \{\text{相对于 } (\mathbb{H}, \tau) \text{ 的正则 Hermite 模}\} / \simeq \\ (p, q) &\mapsto \left\langle \underbrace{1, \dots, 1}_{p \text{ 份}}, \underbrace{-1, \dots, -1}_{q \text{ 份}} \right\rangle \text{ 的同构类.} \end{aligned}$$

因此正则 Hermite 模由维数 $p + q$ 和符号差 $p - q$ 完全确定, 精确到同构.

证明 考虑正则 Hermite 模 (M, B). 由于 $\{q \in \mathbb{H} : \tau(q) = q\} = \mathbb{R}$, 而 B 由 $h(x) = B(x, x)$ 给出的映射 $h : M \to \mathbb{R}$ 唯一确定, 和二次型完全类似的论证说明存在 $a_1, \dots, a_n \in \mathbb{R}$ 使得 $(M, B) \simeq \langle a_1, \dots, a_n\rangle$, 此外不难说明 (M, B) 正则当且仅当 $a_1, \dots, a_n \in \mathbb{R}^\times$, 详见定义–命题 16.1.1 与命题 16.2.2 的论证.

在正则的情形, 可伸缩 $a_1, \dots, a_n$ 使得它们都属于 $\{\pm 1\}$, 因此断言中的映射满. 为了说明它单, 记 $\mathrm{Tr} : \mathbb{H} \to \mathbb{R}$ 为迹映射 $q \mapsto q + \tau(q)$. 若 (M, B) 是相对于 $(\mathbb{H}, \tau)$ 的 Hermite 形式, 则 $(M, \mathrm{Tr} \circ B)$ 是 $\mathbb{R}$ 上的二次型, 而且 $(M, B) \simeq (M', B')$ 蕴涵 $(M, \mathrm{Tr} \circ B) \simeq (M', \mathrm{Tr} \circ B')$.

设 $a \in \mathbb{R}$, 由 $\tau(q)q$ 的公式可知对 $\langle a\rangle$ 取迹给出 $\mathbb{R}$ 上的二次型 $\langle a, a, a, a\rangle$. 今设有正则 Hermite 模

$$(M, B) \simeq \left\langle \underbrace{1, \dots, 1}_{p \text{ 份}}, \underbrace{-1, \dots, -1}_{q \text{ 份}} \right\rangle.$$

对此取迹给出非退化二次型, 其正 (或负) 惯性指数为 $4p$ (或 $4q$). 实二次型的惯性定理表明 (p, q) 由 $(M, \mathrm{Tr} \circ B)$ 的同构类唯一确定, 因而由 (M, B) 的同构类唯一确定. 明所欲证. □

最后探讨关于 (R, τ) 的另一类简单取法, 以及相应的 ϵ-Hermite 模, 其中的 R 交换但不再是域.

例 16.10.13 (直积情形) 设 F 为域, $\mathrm{char}(F)\neq 2$. 取交换环 $R=F\times F$, 则 $\tau:(a,b)\mapsto(b,a)$ 是 R 的对合. 任何右 R-模 M 都能典范地分解成 $M=M_1\oplus M_2$, 其中 M_1 和 M_2 是 F-向量空间, R 通过第 i 个分量作用于 M_i; 事实上 M_1 (或 M_2) 可以等同于 $M(1,0)$ (或 $M(0,1)$), 这是第十二章的一道简单习题.

关于 M 为有限生成投射模的条件等价于 M_1 和 M_2 都是有限维的. 同样基于维数理由, 正则与非退化 ϵ-Hermite 模的概念在此等价. 此时的 ϵ 必形如 $(a,a^{-1})=\frac{\tau(1,a)}{(1,a)}$, 故根据练习 16.10.9 (ii) 不妨设 $\epsilon=1$. 今起考虑 R 上的 Hermite 模 (M,B). 因为 $(1,0)(0,1)=0$ 而 R 交换, 对所有 $x,y\in M$ 皆有

$$\begin{aligned}
&B(x(1,0),y(1,0))=(0,1)B(x,y)(1,0)=0,\\
&B(x(0,1),y(0,1))=(1,0)B(x,y)(0,1)=0,\\
&B(x(1,0),y(0,1))=(0,1)B(x,y)(0,1)=(0,1)B(x,y),\\
&B(x(0,1),y(1,0))=(1,0)B(x,y)(1,0)=(1,0)B(x,y).
\end{aligned}$$

记 $M_i^\vee$ 为 M_i 作为 F-向量空间的对偶. 不难由此推得 B 对应的 ψ 限制为

$$\psi_1:M_1\to M_2^\vee,\quad \psi_2:M_2\to M_1^\vee,$$

而 B 的 Hermite 性质说明 ψ_1 和 ψ_2 按下图相互确定:

$$\begin{array}{ccc}
 & \text{第二个分量} & \\
\psi_1(x)(y) & \longmapsto & \underbrace{B(x,y)}_{\in\{0\}\times F\simeq F} \\
\| & & \downarrow\tau \\
\psi_2(y)(x) & \longmapsto & \underbrace{B(y,x)}_{\in F\times\{0\}\simeq F} \\
 & \text{第一个分量} &
\end{array}\qquad x\in M_1,\ y\in M_2.$$

要求 (M,B) 为正则 Hermite 模相当于要求 ψ_1 和 ψ_2 皆为同构.

综上, R 上的正则 Hermite 模 (M,B) 典范地同构于 $M=V\oplus V^\vee$ 之形, 其中 V 是有限维 F-向量空间, R 通过第一个 (或第二个) 分量作用于 V (或 $V^\vee$), 而 B 用典范配对 $\langle\cdot,\cdot\rangle:V^\vee\times V\to F$ 描述如下 (相当于取 $\psi_2=\mathrm{id}_{V^\vee}$):

$$\begin{aligned}
&B(v,v')=(0,0)=B(\check v,\check v'),\quad v,v'\in V,\ \check v,\check v'\in V^\vee,\\
&B(v,\check v)=\left(0,\langle\check v,v\rangle\right),\quad B(\check v,v)=\left(\langle\check v,v\rangle,0\right).
\end{aligned}$$

此时 M 总是有限秩自由 R-模, (M,B) 的同构类完全由 M 的秩确定.

这说明精确到典范同构, R 上的正则 ϵ-Hermite 模理论不过是有限维 F-向量空间理论, 这点也可以写成范畴的等价 (§B.4).

练习 16.10.14 在例 16.10.13 的场景中考虑正则 Hermite 模 (M,B), 作分解 $M=V\oplus V^\vee$. 试给出同构 $\mathrm{U}(M,B)\simeq\mathrm{GL}(V)$.

习题

1. 考虑域 F 上的 2 维非退化二次型 q. 说明 $q \simeq \mathcal{H}$ 当且仅当 $\det q = (-1)F^{\times 2}$.

2. 设 $a, b \in F^\times$ 而 $\langle a, b, ab\rangle$ 迷向. 证明 $\langle 1, a, b, ab\rangle \simeq 2\mathcal{H}$. 提示 考虑 $\langle a, b, ab\rangle$ 的行列式.

3. 在练习 16.1.14 的场景中, 设 (V, q) 非退化, $\dim V$ 为偶数, 从而 $(\det g)^2 = \nu(g)^{\dim V}$ 蕴涵 $\det g = \pm\nu(g)^{\dim V/2}$, 其中 $g \in \mathrm{GO}(V, q)$.

(i) 说明所有满足 $\det g = \nu(g)^{\dim V/2}$ 的 g 构成子群, 一些文献记此为 $\mathrm{GSO}(V, q)$.

(ii) 举例说明负号确实可能出现.

4. 对域 F 上的非退化非零二次型 (V, q), 定义

$$D(q) := \left\{d \in F^\times : \exists x \in V,\ q(x) = d\right\}.$$

(i) 说明若 q 迷向 (定义 16.2.6), 则 $D(q) = F^\times$.

(ii) 对于 $d \in F^\times$, 说明 $d \in D(q)$ 当且仅当 $q \oplus \langle -d\rangle$ 迷向.

(iii) 设 q_1 和 q_2 为非退化非零二次型, 证明 $q_1 \oplus (-q_2)$ 迷向当且仅当 $D(q_1) \cap D(q_2) \neq \varnothing$.

5. 承上题, 给定 $r \in \mathbb{Z}_{\geq 1}$ 和域 F (满足 $\mathrm{char}(F) \neq 2$), 证明以下陈述等价:

(i) 对所有 r 维非退化二次型 q 皆有 $D(q) = F^\times$;

(ii) 所有 $r+1$ 维非退化二次型皆迷向.

6. 设 (V, q) 为非退化二次型, $n := \dim q$. 定义

$$d_\pm q := (-1)^{\frac{n(n-1)}{2}} \det q \ \in F^\times/F^{\times 2}.$$

记 $n := \dim q$, $n' := \dim q'$. 验证 $d_\pm(q \oplus q') = d_\pm q \cdot d_\pm q' \cdot (-1)^{nn'}$. 另外验证

$$d_\pm(q \oplus \mathcal{H}) = d_\pm q;$$

这表明 $d_\pm$ 诱导映射 $\mathcal{W}(F) \to F^\times/F^{\times 2}$.

7. 域 F 上的所有非退化二次型的同构类构成一个集合 $\mathcal{Q}(F)$, 它对 $\oplus$ 成为交换幺半群, 以零二次型为零元. 假定读者了解交换幺半群的群化 (例 B.2.8), 试证群化 $\mathcal{Q}(F)^{\mathrm{grp}}$ 同构于定义 16.3.7 的群 $\widehat{\mathcal{W}}(F)$.

提示 写下自然的幺半群同态 $\xi : \mathcal{Q}(F) \to \widehat{\mathcal{W}}(F)$, 证明 ξ 满足例 B.5.12 揭示的泛性质: 对所有交换群 A 和同态 $f : \mathcal{Q}(F) \to A$, 存在唯一的同态 $\overline{f} : \widehat{\mathcal{W}}(F) \to A$ 使得 $\overline{f}\xi = f$.

8. 设 (V, q) 为域上的非退化二次型, 证明以下陈述等价.

(i) 存在 $m \in \mathbb{Z}_{\geq 0}$ 使得 $q \simeq m\mathcal{H}$;

(ii) 存在 V 的子空间 L 使得 $L = L^\perp$ (基于和辛空间的类比, 这种子空间又称 V 的 Lagrange 子空间);

(iii) 存在有限维 F-向量空间 L 使得 (q,V) 同构于 $L\oplus L^\vee$ 连同以下线性映射 $\psi: L\oplus L^\vee\to(L\oplus L^\vee)^\vee$ 确定的二次型: 等同 $(L\oplus L^\vee)^\vee$ 与 $L^\vee\oplus L^{\vee\vee}$, 则

- ⋆ ψ 在直和项 L 上是典范同构 $L\xrightarrow{\sim}L^{\vee\vee}$,
- ⋆ ψ 在直和项 $L^\vee$ 上是恒等 $L^\vee\xrightarrow{\mathrm{id}}L^\vee$.

也请尝试对 §16.8 介绍的 ϵ-Hermite 形式给出相应的版本.

9. (内积空间的张量积) 设 V_i 为实 (或复) 向量空间, $(\cdot|\cdot)_i$ 是 V_i 上的实 (或复) 内积, $i=1,2$. 按照 (16.6.1) (或 (16.9.1)) 的方法在 $V_1\otimes V_2$ 上定义双线性 (或 Hermite) 形式 $(\cdot|\cdot)_1\otimes(\cdot|\cdot)_2$. 说明这使 $V_1\otimes V_2$ 成为实 (或复) 内积空间.

10. 设 F 为域, $\mathrm{char}(F)\neq 2$. 在 F-向量空间 $\mathrm{M}_{2\times2}(F)$ 上定义双线性形式 $T(\boldsymbol{A},\boldsymbol{B})=\mathrm{Tr}(\boldsymbol{AB})$. 考虑子空间 $V:=\{\boldsymbol{A}\in\mathrm{M}_{2\times2}(F):\mathrm{Tr}(\boldsymbol{A})=0\}$.

(i) 说明 T 在 $\mathrm{M}_{2\times2}(F)$ 和 V 上都是非退化的; 事实上, $V=(F\cdot\mathbf{1}_{2\times2})^\perp$. 记 V 上所对应的非退化二次型为 q. 说明 (V,q) 迷向.

(ii) 对所有 $g\in\mathrm{GL}(2,F)$ 定义 $\pi(g)\in\mathrm{GL}(V)$ 为 $\boldsymbol{A}\mapsto g\boldsymbol{A}g^{-1}$. 证明 $\pi(g)\in\mathrm{SO}(V,q)$.

提示 易证 $\pi(g)\in\mathrm{O}(V,q)$; 难点是证 $\det\pi(g)=1$. 直接展开或作如下论证: 不妨设 F 是无穷域, 以将多项式等同于多项式函数; 已知 $\det\pi(g)\in\{1,-1\}$, 而 $\det\pi(g)$ 是关于 g 的矩阵元和 $\det(g)^{\pm1}$ 的多项式, 按此说明它是常值.

(iii) 证明 (ii) 给出的 $\pi:\mathrm{GL}(2,F)\to\mathrm{SO}(V,q)$ 是满足 $\ker(\pi)=F^\times\mathbf{1}_{2\times2}$ 的群同态.

(iv) 证明 π 满, 从而诱导群同构 $\mathrm{PGL}(2,F)\xrightarrow{\sim}\mathrm{SO}(V,q)$.

提示 基于 Cartan–Dieudonné 定理 16.5.3, 证明对 V 的所有非迷向元素 $\boldsymbol{A}$ 皆有 $-r_{\boldsymbol{A}}\in\mathrm{im}(\pi)$ 即可. 这相当于寻求 $g\in\mathrm{GL}(2,F)$ 使得 $g^2\in F^\times\mathbf{1}_{2\times2}$, $g\notin F^\times\mathbf{1}_{2\times2}$ 而 $g\boldsymbol{A}g^{-1}=\boldsymbol{A}$. 分成 $\boldsymbol{A}$ 可对角化或 $\mathrm{Char}_{\boldsymbol{A}}$ 不可约两种情形讨论.

(v) 证明对所有迷向的 3 维非退化二次型 (V',q') 皆有同构 $\mathrm{PGL}(2,F)\simeq\mathrm{SO}(V',q')$.

提示 这种二次型精确到同构和伸缩是唯一的.

11. 给定满足 $\mathrm{char}(F)\neq2$ 的域 F, 设二次型 (V,q) 同构于双曲平面 $\mathcal{H}$, 证明

$$\mathrm{O}(V,q)\simeq F^\times\rtimes_f(\mathbb{Z}/2\mathbb{Z}),$$

半直积中的 $f:\mathbb{Z}/2\mathbb{Z}\to\mathrm{Aut}(F^\times)$ 映 $1+2\mathbb{Z}$ 为取逆; 借此确定 $\mathrm{O}(V,q)$ 的中心 (例 11.1.4).

提示 直接计算矩阵, 并注意到 $|F|=3\iff f(1+2\mathbb{Z})=\mathrm{id}_{F^\times}$.

12. 设 (V,q) 是非退化二次型, $n:=\dim V\in\mathbb{Z}_{\geq1}$.

(i) 设 $x\in V\smallsetminus\{0\}$ 非迷向, 证明 $g\in\mathrm{O}(V,q)$ 满足 $r_xg=gr_x$ 当且仅当 x 是 g 的特征向量.

(ii) 证明当 $n\neq2$ 或 $|F|>3$ 时, $\mathrm{O}(V,q)$ 的中心是 $\{\pm\mathrm{id}\}$.

提示 在 $n=2$ 而 $|F|>3$ 时应用先前习题的结论. 以下设 $n\geq3$ 而 $q=\langle a_1,\ldots,a_n\rangle$. 由 (i) 可知非迷向向量总是中心的元素 z 的特征向量; 特别地, z 对应到对角矩阵. 为了说明对角元相等, 可设 $n=3$ 而 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_3$ 特征值不同. 取 t 使得 $\boldsymbol{e}_1+t\boldsymbol{e}_2+\boldsymbol{e}_3$ 非迷向以推导矛盾.

(iii) 证明当 $n \neq 2$ 时, $\mathrm{SO}(V,q)$ 的中心是 $\{\pm\mathrm{id}\}$ (若 n 是偶数) 或平凡子群 (若 n 是奇数).

提示〉 若 n 是奇数则问题相对容易. 设 $n \geq 4$ 为偶数, $q = \langle a_1, \ldots, a_n\rangle$. 由于 $\mathrm{SO}(V,q)$ 包含所有对角元属于 $\{\pm 1\}$ 而且其乘积为 1 的对角矩阵, 由此论证中心 z 的元素必形如

$$\begin{pmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \boldsymbol{A} & \\ 0 & & & \end{pmatrix}$$

从而 $zr_{\boldsymbol{e}_1} = r_{\boldsymbol{e}_1} z$, 故 z 实则属于 $\mathrm{O}(V,q)$ 的中心.

13. 设 F 为任意域, V 为 F-向量空间. 记对偶空间 $V^\vee$ 和 V 的典范配对为 $\langle\cdot,\cdot\rangle : V^\vee \times V \to F$. 典范嵌入 $\varpi_+ : V \to V^{\vee\vee}$ 映 v 为 $\lambda \mapsto \langle \lambda, v\rangle$.

(i) 指定双线性形式 $B : V \times V \to F$ 相当于指定线性映射 $\psi : V \to V^\vee$, 两者之间的关系是 $B(v,w) = \langle \psi(v), w\rangle$. 证明 B 对称当且仅当下图交换:

$$\begin{array}{ccc} (V^\vee)^\vee & \xrightarrow{\psi^\vee} & V^\vee. \\ {\scriptstyle\varpi_+}\uparrow & \nearrow_{\psi} & \\ V & & \end{array}$$

(ii) 说明若在上述图表中以 $\varpi_- := -\varpi_+ : V \to V^{\vee\vee}$ 代替 ϖ_+, 则 B 反对称当且仅当图表交换.

(iii) 给定满足 $\mathrm{char}(E) \neq 2$ 的域 E 和满足 $\tau^2 = \mathrm{id}_E$ 的自同构 $\tau : E \xrightarrow{\sim} E$, 对 Hermite 形式和反 Hermite 形式 (定义 16.8.3) 证明和 (i), (ii) 相应的性质.

提示〉 将 $V^\vee$ 改为 $V^* := \overline{V^\vee}$, 相应地定义 ψ^* 并写下典范的 E-线性映射 $\varpi_\pm : V \to V^{**}$.

14. 设 F 为域 E 的子域, $[E:F] = 2$ 而 $\mathrm{char}(E) \neq 2$. 证明满足 $\tau|_F = \mathrm{id}_F$ 而 $\tau \neq \mathrm{id}_E$ 的域自同构 $\tau : E \xrightarrow{\sim} E$ 恰有一个.

提示〉 取 $a \in E \smallsetminus F$ 则有 F-向量空间的直和分解 $E = F \oplus Fa$; 根据配方法, 可取到 a 使得 $b := a^2 \in F^\times \smallsetminus F^{\times 2}$; 以下不妨记 $\sqrt{b} := a$. 注意到 F-代数的同构

$$\begin{aligned} F[X]/(X^2 - b) &\xrightarrow{\sim} E \\ f + (X^2 - b) &\mapsto f(\sqrt{b}). \end{aligned}$$

记法 $E = F(\sqrt{b})$ 是常用的. 按此说明所求的 τ 能且只能是 $\tau(\sqrt{b}) = -\sqrt{b}$.

15. 定义–命题 16.6.5 解释了如何对二次型作域的变换. 状况对于 ϵ-Hermite 形式稍有不同. 对 §16.8 中给定的 E, τ 和 F, 考虑相对于 $E|F$ 的非退化 ϵ-Hermite 形式 (V,B). 另给定域嵌入 $F \hookrightarrow F'$.

(i) 说明 $E' := E \underset{F}{\otimes} F'$ 可以按照

$$(x\otimes y)(x'\otimes y') = xx'\otimes yy', \quad \tau'(x\otimes y) = \tau(x)\otimes y$$

作成带有自同构 τ' 的 F'-代数, 使得 $(\tau')^2 = \mathrm{id}$ 而 $\{z\in E' : \tau'(z) = z\}$ 等同于 F'.

(ii) 试赋予 $V' := V \underset{F}{\otimes} F'$ 自然的 E'-模结构, 然后在 E' 为域的前提下说明

$$B'(x\otimes r, y\otimes s) = rB(x,y)s, \quad x,y\in V, \quad r,s\in F'$$

在 V' 上给出相对于 $E'|F'$ 的非退化 ϵ-Hermite 形式 B'.

提示〉 定义 $x\otimes y\in E'$ 在 V' 上的乘法为 $v\otimes r\mapsto xv\otimes ry$.

(iii) 证明若 E' 非域, 则有 F'-代数的同构 $E'\simeq F'\times F'$ 使得 τ' 对应于 $(r_1,r_2)\mapsto(r_2,r_1)$. 此时的 (V',B') 按照例 16.10.13 的方式对应到 F'-向量空间.

提示〉 按照前一道习题的提示记 $E = F(\sqrt{b}) \simeq F[X]/(X^2-b)$, 由张量积保商 (第十五章的简单习题) 可见 $E'\simeq F'[X]/(X^2-b)$, 这是域当且仅当 $b\notin (F')^{\times 2}$; 当 $b\in (F')^{\times 2}$ 时中国剩余定理给出同构 $F'[X]/(X^2-b)\simeq F'\times F'$; 验证 (V',B') 此时为 §16.10 所谓的非退化 ϵ-Hermite 模.

16. 证明推论 16.9.10 的下述辛版本: 设 (V,B) 为辛空间, U_1 和 U_2 是 V 的子空间, 而向量空间的同构 $f: U_1 \xrightarrow{\sim} U_2$ 保持辛形式, 换言之 $B(f(x),f(y)) = B(x,y)$ 对所有 $x,y\in U_1$ 皆成立. 证明存在 $g\in \mathrm{Sp}(V,B)$ 使得 $g|_{U_1} = f$.

提示〉 尝试修改推论 16.4.7 的证明, 但在 U_1 非全迷向的第一种情况, 需从 U_1 取出的不再是非迷向向量 u_1, 而是 U_1 的 2 维子空间 $Fu_1\oplus Fu_1'$, 满足 $B(u_1,u_1') = 1$.

17. 证明定义–命题 16.10.3 的同态 $\varpi_M : M\to (M^*)^*$ 在 M 是有限生成投射模时是同构.

提示〉 论证 $\varpi_{M_1\oplus M_2}$ 是同构当且仅当 ϖ_{M_1} 和 ϖ_{M_2} 都是同构, 然后化到 $M = R$ 的特例.

18. 对所有 $a,b\in\mathbb{Z}_{\geq 0}$, 记实二次型 $q_{a,b} := \underbrace{\langle 1,\dots,1\rangle}_{a\text{ 份}} \oplus \underbrace{\langle -1,\dots,-1\rangle}_{b\text{ 份}}$ 的正交群为 $\mathrm{O}(a,b)$, 行列式为 1 的子群记为 $\mathrm{SO}(a,b)$. 以下要求读者对点集拓扑中的连通性有所了解.

(i) 证明 $\mathrm{O}(n) = \mathrm{O}(n,0)$ 恰有两个连通分支 $\{\det = 1\}$ 和 $\{\det = -1\}$.

提示〉 关键在于证 $\mathrm{SO}(n)$ 道路连通. 一种方法是用正交变换的标准形, 另一种方法则是用 Cartan–Dieudonné 定理 16.5.3, 通过变动镜射的法向量, 将问题归结为 $\mathbb{R}^n\smallsetminus\{\mathbf{0}\}$ 的连通性.

(ii) 形如 $\mathrm{O}(1,n)$ 的群常见于物理学; 特例 $\mathrm{O}(1,3)$ 又称 Lorentz 群. 记与 $q_{1,n}$ 对应的双线性形式为 $B_{1,n}$, 记 $\mathbb{R}^{1+n}$ 的标准基为 $\boldsymbol{e}_0,\dots,\boldsymbol{e}_n$. 证明对所有 $g\in\mathrm{O}(1,n)$ 皆有 $|B_{1,n}(\boldsymbol{e}_0, g\boldsymbol{e}_0)|\geq 1$, 并且证明 $g\mapsto \operatorname{sgn} B_{1,n}(\boldsymbol{e}_0,g\boldsymbol{e}_0)$ 是从 $\mathrm{O}(1,n)$ 到 $\{\pm 1\}$ 的群同态.

提示〉 从 $g\in\mathrm{O}(1,3)$ 的矩阵表法

$$g = \left(\begin{array}{c|c} a & {}^{\mathrm{t}}\boldsymbol{w} \\ \hline \boldsymbol{v} & \boldsymbol{A} \end{array}\right), \quad a\in\mathbb{R}, \quad \boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^3, \quad \boldsymbol{A}\in \mathrm{M}_{3\times 3}(\mathbb{R})$$

推导 $a^2 = 1 + \|\boldsymbol{v}\|^2_{\mathbb{R}^3}$; 另一方面, 将 g^{-1} 用 g 的转置表出, 可见 $a^2 = 1 + \|\boldsymbol{w}\|^2_{\mathbb{R}^3}$; 按此论证 $g \mapsto \operatorname{sgn} B_{1,n}(\boldsymbol{e}_0, g\boldsymbol{e}_0)$ 为同态.

(iii) 对所有非迷向向量 $\boldsymbol{x}$, 按照 $q_{1,n}(\boldsymbol{x}) > 0$ 或 $q_{1,n}(\boldsymbol{x}) < 0$ 描述 $B_{1,n}(\boldsymbol{e}_0, r_{\boldsymbol{x}}\boldsymbol{e}_0)$ 的正负号. 由此说明有群的满同态

$$\begin{aligned} \alpha : \mathrm{O}(1,n) &\to \{\pm 1\} \times \{\pm 1\} \\ g &\mapsto (\det g,\ \operatorname{sgn} B_{1,n}(\boldsymbol{e}_0, g\boldsymbol{e}_0)) . \end{aligned}$$

(iv) 承上题, 说明 α 的纤维正好是 $\mathrm{O}(1,n)$ 的连通分支, 共有 4 个.

提示〉 由 (iii) 知 α 的每个纤维都是 $\mathrm{O}(1,n)$ 的连通分支的并. 问题归结为证明 $\ker(\alpha)$ 道路连通. 一种方法是运用 Cartan–Dieudonné 定理 16.5.3.

推而广之, 当 $a, b \geq 1$ 时 $\mathrm{O}(a,b)$ 也恰有 4 个连通分支. 这点的证明需要一些 Lie 群的知识, 故本书不论.

19. 记对称正定 $n \times n$ 实矩阵构成的集合为 S_n^+. 极分解给出双射

$$S_n^+ \times \mathrm{O}(n) \xrightarrow{1:1} \mathrm{GL}(n,\mathbb{R}), \quad (\boldsymbol{R}, \boldsymbol{U}) \mapsto \boldsymbol{R}\boldsymbol{U}.$$

承继上一题的符号, 命 $n = a + b$. 证明上式限制为双射

$$(S_n^+ \cap \mathrm{O}(a,b)) \times (\mathrm{O}(a) \times \mathrm{O}(b)) \xrightarrow{1:1} \mathrm{O}(a,b).$$

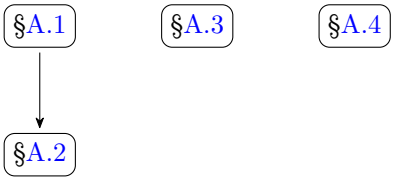
提示〉 可将 $\mathrm{O}(n)$ 理解为 $\mathrm{GL}(n,\mathbb{R})$ 的自同构 $\boldsymbol{A} \mapsto ({}^{\mathrm{t}}\boldsymbol{A})^{-1}$ 的不动点集, 类似方法可诠释 $\mathrm{O}(a,b)$, 然而这两个自同构相互交换.

附录 A　集合论补遗

这份附录旨在补全正文遗漏或省略的若干集合论常识, 分成三个独立的板块. 第一部分 (§§A.1–A.2) 是关于 Peano 算术的公理, 以及 von Neumann 构造非负整数集的方法. 第二部分 (§A.3) 证明关于基数的一些性质, 和正文的关系较为紧密. 第三部分 (§A.4) 介绍关于偏序集的 Zorn 引理, 此一结论等价于选择公理, 它在代数学的后续研究中占有重要地位. Zorn 引理可以用来说明任何向量空间都有基, 任何线性无关子集都能延拓为基, 这些是正文所默认的事实.

此处介绍的关于基数的知识只是皮毛, 有需求的读者应当继续参考其他教材, 如 [13] 或 [10, 第一章].

阅读顺序



A.1　Peano 算术

非负整数的算术几乎是一切数学所共享的地基. 算术的形式化因而也是关于数学基础的任何纲领的第一块试金石. 在这方面, 目前通行的是 G. Peano 在 1889 年的著作 *Arithmetices principia, nova methodo exposita* 提出的公理系统, 现称为 **Peano 算术**.

按现代观点, Peano 算术所意图描述的对象是非负整数, 除了标准的逻辑符号和等号 $=$, 它只涉及一个常数 0 (代表零) 和一个运算 s (代表后继元), 映 x 为 $\mathsf{s}(x)$; 因此 Peano 算术其实可以建立在形式语言和一系列公理的基础上, 未必要涉及集合. 为了简化一些逻辑上的技术细节, 本书仅讨论 Peano 算术在 ZFC 公理集合论里的表述.

定义 A.1.1 (G. Peano)　设 N 是集合, $0 \in N$ 是给定的元素, 而 $\mathsf{s} : N \to N$ 是映射. 若以下条件成立, 则称 $(N, 0, \mathsf{s})$ 满足 Peano 公理.

▷ **零非后继**　不存在 $x \in N$ 使得 $0 = \mathsf{s}(x)$;

▷ **后继元运算是单射** 若 $\mathsf{s}(x)=\mathsf{s}(y)$, 则 $x=y$;

▷ **数学归纳法原理** 设 $I\subset N$ 为包含 0 的子集, 而且对于任何 $x\in N$ 都有 $x\in I\implies \mathsf{s}(x)\in I$, 则 $I=N$.

熟悉的正整数按此在 N 中诠释为 $1=\mathsf{s}(0)$, $2=\mathsf{s}(1)$ 等, 依此类推.

引理 A.1.2 设 $(N,0,\mathsf{s})$ 满足 Peano 公理, $x\in N$ 而 $x\neq 0$, 则必存在 $y\in N$ 使得 $x\neq y$ 而 $x=\mathsf{s}(y)$.

证明 设若不然, 则 $I:=N\smallsetminus\{x\}$ 满足数学归纳法的所有条件, 从而 $I=N$, 矛盾. □

非负整数的加法和乘法可以用 Peano 公理在 N 上递归地定义为

$$\begin{aligned} x+0&:=x,\quad & x+\mathsf{s}(y)&:=\mathsf{s}(x+y),\\ x\cdot 0&:=0,\quad & x\cdot\mathsf{s}(y)&:=(x\cdot y)+x. \end{aligned}$$

特别地, $\mathsf{s}(x)=x+1$. 为了严谨地完成这些定义, 以下工具必不可少, 然而证明是比较迂回的.

定理 A.1.3 (递归原理) 设 $(N,0,\mathsf{s})$ 满足 Peano 公理, Y 是任意集合, $\mathsf{t}:Y\to Y$ 是映射, 而 $y_0\in Y$. 此时存在唯一的映射 $f:N\to Y$ 使得 $f(0)=y_0$ 而且

$$\forall x\in N,\ f(\mathsf{s}(x))=\mathsf{t}(f(x)).$$

证明 首先处理相对简单的唯一性. 设 $f_1,f_2:N\to Y$ 皆具上述性质, 考虑 $\Delta:=\{x\in N:f_1(x)=f_2(x)\}$, 则按条件立见 $0\in\Delta$, 而且 $x\in\Delta$ 蕴涵 $\mathsf{s}(x)\in\Delta$. 于是 $\Delta=N$, 亦即 $f_1=f_2$.

接着证明 f 的存在性. 请考虑满足下列性质的子集 $\Gamma\subset N\times Y$:

⋆ $(0,y_0)\in\Gamma$,

⋆ 若 $(x,y)\in\Gamma$, 则 $(\mathsf{s}(x),\mathsf{t}(y))\in\Gamma$.

令 $\mathcal{U}$ 为这些子集 Γ 所成的集合; 显然有 $N\times Y\in\mathcal{U}$, 因此 $\mathcal{U}$ 非空, 取交 $\Gamma_{\min}:=\bigcap_{\Gamma\in\mathcal{U}}\Gamma$ 有意义; 它依然属于 $\mathcal{U}$. 以下将说明 $\Gamma_{\min}$ 是某个映射 $f:N\to Y$ 的图形, 而 f 即所求的映射.

为此, 需要验证的是所有 $x\in N$ 都有如下两条性质.

$$\begin{aligned} \mathcal{P}(x):&\quad \forall y,y'\in Y,\ (x,y),(x,y')\in\Gamma_{\min}\implies y=y',\\ \mathcal{Q}(x):&\quad \exists y\in Y,\ (x,y)\in\Gamma_{\min}. \end{aligned}$$

考虑集合 $I:=\{x\in N:\mathcal{P}(x)\wedge\mathcal{Q}(x)\}$. 我们断言 $0\in I$. 显然 $\mathcal{Q}(0)$ 按 $\mathcal{U}$ 的定义自动成立, 至于 $\mathcal{P}(0)$, 命

$$\Gamma_0:=\{(0,y_0)\}\cup((N\smallsetminus\{0\})\times Y)\,;$$

注意到 $\Gamma_0 \in \mathcal{U}$, 而 $y \neq y_0$ 时 $(0,y) \notin \Gamma_0$, 故此时 $(0,y) \notin \Gamma_{\min}$.

接着说明若 $x \in I$ 则 $\mathsf{s}(x) \in I$. 根据前提, 存在唯一的 y 使得 $(x,y) \in \Gamma_{\min}$, 由此可得 $(\mathsf{s}(x), \mathsf{t}(y)) \in \Gamma_{\min}$. 现在设 $y' \in Y$ 满足 $y' \neq \mathsf{t}(y)$ 并考虑

$$\Gamma_1 := \Gamma_{\min} \smallsetminus \{(\mathsf{s}(x), y')\}.$$

显然 $(0,y_0) \in \Gamma_1$; 其次, $\Gamma_{\min}$ 中第一个分量为 x 的元素只有 (x,y), 而 s 单, 所以从 $\Gamma_{\min}$ 扣掉 $(\mathsf{s}(x), y')$ 并不影响定义 $\mathcal{U}$ 的第二条性质. 综上可见 $\Gamma_1 \in \mathcal{U}$. 于是 $\Gamma_{\min} = \Gamma_1$, 亦即 $(\mathsf{s}(x), y') \notin \Gamma_{\min}$. 故 $\mathcal{P}(\mathsf{s}(x))$ 和 $\mathcal{Q}(\mathsf{s}(x))$ 皆成立.

应用数学归纳法可得 $I = N$, 证毕. □

加法和乘法运算满足熟知的代数性质.

▷ **加法结合律** $x + (y+z) = (x+y) + z$.

▷ **加法交换律** $x + y = y + x$.

▷ **加法零元** $x + 0 = x$.

▷ **加法消去律** $x + z = y + z$ 蕴涵 $x = y$.

▷ **乘法结合律** $x(yz) = (xy)z$.

▷ **乘法对非零元的消去律** $xz = yz$ 而 $z \neq 0$ 蕴涵 $x = y$.

▷ **乘法幺元** $x \cdot 1 = x$.

▷ **分配律** $(x+y)z = xz + yz$.

以加法结合律为例, 当 $z = 0$ 时, 等式显然对所有 x, y 成立; 现在设 $z \in N$ 给定, 使得等式对所有 x, y 成立, 则

$$\begin{aligned} x + (y + \mathsf{s}(z)) &= x + \mathsf{s}(y+z) = \mathsf{s}(x + (y+z)) \\ &= \mathsf{s}((x+y)+z) = (x+y) + \mathsf{s}(z). \end{aligned}$$

故 $\mathsf{s}(z)$ 也有相同的性质. 设 I 为对所有 x, y 都满足加法结合律的 z 所成集合, 应用数学归纳法可得 $I = N$, 故加法结合律恒成立.

从这些性质可以推出 $x \cdot 0 = x \cdot (0+0) = x \cdot 0 + x \cdot 0$, 两边消去 $x \cdot 0$, 便得到

$$x \cdot 0 = 0.$$

练习 A.1.4 完整证明以上几条代数性质.

指数运算 x^y 在 N 上也有类似的递归定义. 尽管分析学家或有异议, 此处规定 $0^0 = 1$.

练习 A.1.5 使用递归原理严谨地定义阶乘函数 $x \mapsto x!$.

定义–命题 A.1.6 设 $(N, 0, \mathsf{s})$ 满足 Peano 公理. 在 N 上定义二元关系 $\leq$ 使得

$$x \leq y \iff \exists d \in N,\ y = x + d.$$

这使 $(N, \leq)$ 成为全序集. 进一步, 对所有 $x, y, z \in N$ 都有

$$\begin{aligned} x \geq y &\implies x + z \geq y + z, \\ x \geq y &\implies xz \geq yz. \end{aligned}$$

证明 先说明 $\leq$ 是偏序. 由 $x = x + 0$ 可得反身性 $x \leq x$. 设 $x \leq y$ 而 $y \leq z$, 表作 $y = x + d$ 和 $z = y + e$, 其中 $d, e \in N$, 则加法结合律给出

$$z = x + (d + e),$$

于是 $x \leq z$, 是为传递性. 下面验证反称性: 设 $x \leq y$ 而 $y \leq x$, 亦即存在 $d, e \in N$ 使得 $y = x + d$ 和 $x = y + e$. 由此可得 $x = x + (d + e)$, 再用加法消去律可得 $d + e = 0$. 兹断言 $d = e = 0$; 设若不然, 则根据引理 A.1.2, 在 d 或 e 中必有一者是某个元素的后继, 从而根据加法的递归定义, 0 也随之是某元素的后继, 矛盾.

下一步是验证全序性质, 亦即任两个 $x, y \in N$ 皆可比大小. 论证依然是递归的: $y = 0$ 是 N 的下界, 当然可与任何 $x \in N$ 比大小. 现在设 $y \in N$ 可与任何 x 比大小, 然后考虑 $\mathsf{s}(y)$: 给定 x,

- ⋆ 或者 $x \leq y$, 此时 $x \leq \mathsf{s}(y)$;
- ⋆ 或者 $x > y$, 写作 $x = y + d$ 的形式 $(d \neq 0)$, 以引理 A.1.2 表 $d = \mathsf{s}(e)$ 并且运用加法的性质, 可得 $x = \mathsf{s}(y) + e$, 故此时 $x \geq \mathsf{s}(y)$.

由数学归纳法遂知所有 y 都能和所有 x 比大小.

最后来证明关于 $x, y, z \in N$ 的不等式. 设 $x \geq y$, 写作 $x = y + d$. 于是 $x + z = (y + z) + d \geq y + z$; 另一方面, $xz = yz + dz \geq yz$. 至此完成所有证明. □

命题 A.1.7 (良序原理) 设 $(N, 0, \mathsf{s})$ 满足 Peano 公理, 则 $(N, \leq)$ 是良序集; 换言之, 任何非空子集 $S \subset N$ 相对于全序 $\leq$ 都有极小元.

证明 令 $S^\flat := \{x \in N : \forall y \in S,\ x \leq y\}$; 换言之, $S^\flat$ 由 S 在 N 中的所有下界构成. 显然 $0 \in S^\flat$. 另一方面, S 非空蕴涵 $S^\flat \neq N$; 比方说任取 $y \in S$, 则 $y < \mathsf{s}(y)$ 就导致 $\mathsf{s}(y) \notin S^\flat$.

必然存在 $x \in S^\flat$ 使得 $\mathsf{s}(x) \notin S^\flat$, 否则数学归纳法将导致 $S^\flat = N$. 兹断言 x 即所求之极小元. 首先, 已知 $x \leq y$ 对所有 $y \in S$ 成立. 其次, $x \in S$ 必然成立, 否则前述的 $x \leq y$ 可以改进为 $x < y$ (其中 y 遍历 S), 由此易推得 $\mathsf{s}(x) \in S^\flat$, 与 x 的选法矛盾. □

基于良序原理, 很容易导出数学归纳法的第二种形式.

命题 A.1.8 设 $(N, 0, \mathsf{s})$ 满足 Peano 公理, I 是 N 的子集. 假设 $0 \in I$, 而且对于给定的 $x \in N$, 若对所有 $y < x$ 皆有 $y \in I$, 则也有 $x \in I$. 此时 $I = N$.

证明 若 $N \smallsetminus I$ 非空, 则命题 A.1.7 确保 $N \smallsetminus I$ 有极小元 x. 此极小性质表明对所有 $y < x$ 皆有 $y \in I$, 于是 $x \in I$, 矛盾. □

注意: 迄今只是从 Peano 公理作推导, 尚未说明满足公理的资料 $(N, 0, \mathsf{s})$ 确实存在, 实际构造是 §A.2 的主题.

A.2 构造非负整数集

除了空集 $\varnothing$ 和由之衍生的 $\{\varnothing\}$, $\{\{\varnothing\}, \varnothing\}$ 等, 公理集合论不直接给予任何具体的集合. 然而此一立足点已经足以支起 Peano 算术. 以下介绍的构造是 von Neumann 提出的, 基本想法是取

$$\mathbf{0} := \varnothing, \quad \mathbf{1} := \{\varnothing\},$$
$$\ldots, \quad \mathbf{n}+1 := \{\mathbf{0}, \ldots, \mathbf{n}\}.$$

之所以将 $\mathbf{n}$ 加粗标记, 是为了暂时地和日常语言中的 n 区隔. 这一定义看来简洁又合理, 而且 $\mathbf{n}$ 作为一个集合, 其元素个数正好是它所要对应的数 n. 我们希望定义非负整数集 $\mathbb{Z}_{\geq 0}$ 为 $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\}$, 将全体非负整数作为一个集合来把握.

如何严格地定义 $\mathbb{Z}_{\geq 0}$? 问题隐藏在上述公式中的省略号 $\ldots$, 它需要无穷多步. 这是否意味着省略号是某种 "依此类推"? 如果这种解读有严谨的意义, 那么它只能够依着 $0, 1, 2, \ldots$ 来类推, 也就是沿着我们所求的集合 $\mathbb{Z}_{\geq 0}$ 来类推, 这又绕回了出发点.

为了掌握定义的实质, 我们暂且承认这么构造 $\mathbb{Z}_{\geq 0}$ 是合理的. 于是对于任意非负整数 n 和 m, 当下看出 (严谨证明见引理 A.2.5 和练习 A.2.7):

$$n < m \iff \mathbf{n} \in \mathbf{m} \implies \mathbf{n} \subset \mathbf{m}.$$

上式的 $\implies$ 不能倒转; 比方说 $\{\mathbf{0}, \mathbf{2}\}$ 是 $\mathbf{3} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}\}$ 的子集, 却非其元素, 也不对应于任何非负整数. 何以故? 障碍显然在于 "断链": $\mathbf{2}$ 的元素 $\mathbf{1}$ 不在 $\{\mathbf{0}, \mathbf{2}\}$ 之中. 这就启示我们作以下定义.

定义 A.2.1 设 α 为集合, 如果 α 的每个元素都是 α 的子集, 则称 α 为**传递集**.

请读者验证上述定义相当于说对于任意 β 和 γ, 若 $\beta \in \alpha$ 而 $\gamma \in \beta$, 则 $\gamma \in \alpha$. 这是传递性的意涵.

引理 A.2.2 若 α 是传递集, 则 $\alpha \cup \{\alpha\}$ 也是传递集.

证明 设 $x \in \alpha \cup \{\alpha\}$. 若 $x \in \alpha$, 则 α 的传递集性质确保 $x \subset \alpha \subset \alpha \cup \{\alpha\}$. 若 $x \in \{\alpha\}$ 则 $x = \alpha \subset \alpha \cup \{\alpha\}$. □

基于平凡的理由, $\varnothing$ 是传递集. 于是 von Neumann 构造之下的 $\mathbf{0},\mathbf{1},\ldots$ 可以设想为从空集出发, 不断地从 α 过渡到 $\alpha\cup\{\alpha\}$ 所能得到的所有集合 (需要 “依此类推”), 它们理应是传递集.

定义 A.2.3 *满足以下条件的集合 x 称为*归纳集:

- $\varnothing\in x$;
- *若* $\alpha\in x$ *则* $\alpha\cup\{\alpha\}\in x$.

于是 $\{\mathbf{0},\mathbf{1},\ldots\}$ 理应是最小可能的归纳集. 何谓最小? 易见对一族归纳集取交仍得到归纳集, 所以最小可能的归纳集无非是所有归纳集的交. 然而空交是被禁止的, 因此 $\mathbb{Z}_{\geq 0}$ 的构造便归结为

▷ **无穷公理** 存在归纳集.

这是 §2.1 提及的无穷公理的精确版本, 它纯以公理集合论的语言表述, 不再涉及任何 “依此类推” 的话术.

定义 A.2.4 (J. von Neumann) *定义 $\mathbb{Z}_{\geq 0}$ 为所有归纳集之交; 它也是归纳集, 因而可定义后继元映射 $\mathsf{s}:\mathbb{Z}_{\geq 0}\to\mathbb{Z}_{\geq 0}$ 为 $\mathsf{s}(x):=x\cup\{x\}$, 并且定义 $\mathbf{0}:=\varnothing\in\mathbb{Z}_{\geq 0}$.*

引理 A.2.5 *所有元素 $\mathbf{n}\in\mathbb{Z}_{\geq 0}$ 都是传递集.*

证明 设 x 为任意归纳集, 考虑其子集 $x':=\{\alpha\in x:\alpha\text{ 是传递集}\}$. 我们有 $\varnothing\in x'$. 此外若 $\alpha\in x'$, 则 $\alpha\cup\{\alpha\}\in x$ 仍是传递集 (引理 A.2.2), 故 $\alpha\cup\{\alpha\}\in x'$. 综上可见 x' 仍是归纳集. 现在取 $x:=\mathbb{Z}_{\geq 0}$; 既然它是最小的归纳集, 自然有 $x=x'$. □

定理 A.2.6 *定义 A.2.4 的资料 $(\mathbb{Z}_{\geq 0},\mathbf{0},\mathsf{s})$ 满足定义 A.1.1 的 Peano 公理.*

证明 注意到 $x\in\mathsf{s}(x):=x\cup\{x\}$ 总是成立. 首先验证 $\mathbf{0}$ 非后继. 若存在集合 x 使得 $\mathsf{s}(x)=x\cup\{x\}=\mathbf{0}:=\varnothing$, 则 $x\in\varnothing$ 悖理.

其次验证 s 是单射. 设 $x\cup\{x\}=y\cup\{y\}$. 从 $x\in x\cup\{x\}=y\cup\{y\}$ 可知

- 或者 $x\in\{y\}$, 这导致 $x=y$;
- 又或者 $x\in y$, 因为引理 A.2.5 确保 y 是传递集, 这导致 $x\subset y$.

无论哪种情况都有 $x\subset y$. 同理有 $y\subset x$, 故 $x=y$.

最后验证数学归纳法原理. 设子集 $I\subset\mathbb{Z}_{\geq 0}$ 含 $\mathbf{0}$ 并且对 s 保持封闭. 按定义, I 是归纳集. 既然 $\mathbb{Z}_{\geq 0}$ 定义为最小的归纳集, 故 $I=\mathbb{Z}_{\geq 0}$. 明所欲证. □

练习 A.2.7 由定义–定理 A.1.6 赋予 $\mathbb{Z}_{\geq 0}$ 全序 $\leq$. 说明

$$\mathbf{n}<\mathbf{m}\iff\mathbf{n}\in\mathbf{m}.$$

提示〉 设 $\mathbf{m}=\mathbf{n}+\mathbf{1}+\mathbf{d}>\mathbf{n}$, 对 $\mathbf{d}\in\mathbb{Z}_{\geq 0}$ 行数学归纳法来说明 $\mathbf{n}\in\mathbf{m}$.

反之设 $\mathbf{n}\in\mathbf{m}$. 使用反证法: 若 $\mathbf{n}\geq\mathbf{m}$ 则有 $\mathbf{m}\in\mathbf{n}$ 或 $\mathbf{m}=\mathbf{n}$, 说明两者都导致 $\mathbf{m}\in\mathbf{m}$, 进而导致 $\mathsf{s}(\mathbf{m})=\mathbf{m}$; 配合引理 A.1.2 推导矛盾.

留意到上述构造不涉及 §2.1 介绍的正则公理, 尽管它可以用来缩短某些论证.

既已完成 $\mathbb{Z}_{\geq 0}$ 的构造, 我们顺带说明满足 Peano 公理的资料 $(N,0,\mathsf{s})$ 具有适当的唯一性, 作为本节之收尾.

命题 A.2.8 设 $(N,0,\mathsf{s})$ 和 $(N',0',\mathsf{s}')$ 皆满足 Peano 公理, 则存在唯一的映射 $f:N\to N'$ 使得 $f(0)=0'$ 而 $f\mathsf{s}=\mathsf{s}'f$; 这般的 f 必然是双射.

证明 在递归原理 (定理 A.1.3) 中取 $Y=N'$, $\mathsf{t}=\mathsf{s}'$ 和 $y_0=0'$, 可见有唯一的 $f:N\to N'$ 使得 $f(0)=0'$ 而 $f\mathsf{s}=\mathsf{s}'f$.

接着调换 N 和 N' 的角色, 可得唯一的 $g:N'\to N$ 使得 $g(0')=0$ 而 $g\mathsf{s}'=\mathsf{s}g$. 现在考虑 $gf:N\to N$ 和 $fg:N'\to N'$. 我们有

$$\begin{aligned} gf(0)=g(0')=0,&\quad fg(0')=f(0)=0',\\ gf\mathsf{s}=g\mathsf{s}'f=\mathsf{s}gf,&\quad fg\mathsf{s}'=f\mathsf{s}g=\mathsf{s}'fg. \end{aligned}$$

因为 $\mathrm{id}_N(0)=0$ 而 $\mathrm{id}_N\circ\mathsf{s}=\mathsf{s}\circ\mathrm{id}_N$ 显然成立, 将之前关于映射唯一性的结果应用于 $N'=N$ 的特例, 两相比较可见 $gf=\mathrm{id}_N$. 同理, $fg=\mathrm{id}_{N'}$. 这就说明 f 和 g 互为逆映射. □

由于序结构, 加法与乘法都是从 Peano 公理唯一确定的, 命题 A.2.8 的双射 $f:N\to N'$ 不只保持零元与后继元, 还自动使两边的全序, 加法和乘法相匹配. 结合定理 A.2.6 的实际构造, 这就说明 Peano 算术的模型不仅存在, 本质上还是唯一的. 就使用者的角度, 我们遂可以心安理得地操作算术, 不必再管 $\mathbb{Z}_{\geq 0}$ 实际是什么.

A.3 基数补遗

本节的目的是补充 §2.9 省略的一些细节. 这些论证只是显得相对曲折, 没有本质的困难.

引理 A.3.1 设 $f:A\to B$ 是集合之间的满射, 则 $|B|\leq|A|$.

证明 不妨设 A 非空. 根据命题 2.2.6 (依赖于选择公理), 存在 f 的右逆 $g:B\to A$, 亦即 $fg=\mathrm{id}_B$. 这也说明 g 有左逆 f, 故 g 为单射而 $|B|\leq|A|$. □

定理 A.3.2 (Schröder–Bernstein) 若集合 A 和 B 满足 $|A|\leq|B|\leq|A|$, 则 $|A|=|B|$.

证明 根据条件, 存在单射 $f:A\to B$ 和 $g:B\to A$. 在断言中以 $g(B)$ 代替 B, 不妨假设 $B\subset A$ 而 g 是包含映射, 即 $f(A)\subset B\subset A$. 置 $A_0:=A$, $B_0:=B$, 递归地对所有

$n \in \mathbb{Z}_{\geq 0}$ 定义

$$\begin{aligned} A_{n+1} &:= f(A_n), \\ B_{n+1} &:= f(B_n). \end{aligned}$$

容易递归地证明 $B_n \subset A_n$ 而 $A_{n+1} \subset B_n$ (从 $n = 0$ 起, 反复取 f 的像). 于是

$$\underbrace{A_0}_{=A} \supset \underbrace{B_0}_{=B} \supset \cdots \supset A_n \supset B_n \supset A_{n+1} \supset \cdots .$$

以此将 A 分解为以下两列子集的无交并

$$\begin{array}{ccc} (A_0 \smallsetminus B_0) & \sqcup & (B_0 \smallsetminus A_1) \\ {\scriptstyle 1:1}\downarrow{\scriptstyle f} & & \\ (A_1 \smallsetminus B_1) & \sqcup & (B_1 \smallsetminus A_2) \\ {\scriptstyle 1:1}\downarrow{\scriptstyle f} & & \\ \vdots & & \vdots \end{array}$$

并且注意到 $A_0 \smallsetminus B_0$ 之外的部分取并为 B. 定义映射 $\phi : A \to B$ 为

$$\phi(a) = \begin{cases} f(a), & \text{如果 } \exists n \geq 0,\ x \in A_n \smallsetminus B_n, \\ a, & \text{其他情形.} \end{cases}$$

容易验证 ϕ 是双射. □

定理 A.3.3 *对于任意集合 A 和 B, 或者存在单射 $A \hookrightarrow B$, 或者存在单射 $B \hookrightarrow A$.*

证明 这个结果的证明或者需要序数理论, 或者需要 Zorn 引理. 两者都依赖选择公理. 基于序数的进路可见 [10, 命题 1.4.5]. □

回忆到与 $\{0, \ldots, n-1\}$ 等势的集合 A 称为有限集, 其中 $n \in \mathbb{Z}_{\geq 0}$; 此时记 $|A| = n$. 非有限的集合称为无穷集.

命题 A.3.4 (= 命题 2.9.4) *设 A 和 B 是等势的有限集, 则任何单射 (或满射) $f : A \to B$ 自动是双射.*

证明 首先考虑 f 是单射的情形. 不失一般性, 设 $n := |A| \geq 1$, 对 n 递归地论证. 首先 $n = 1$ 情形是平凡的. 接着设 $n > 1$, 任选 $a \in A$, 定义 $A' := A \smallsetminus \{a\}$ 和 $B' := B \smallsetminus \{f(a)\}$, 则容易说明 $|A'| = |A| - 1$ 和 $|B'| = |B| - 1$, 而且 $f|_{A'} : A' \to B'$ 仍然是单射, 从而是双射. 这就说明 $f : A \to B$ 是双射.

其次设 f 是满射. 仍可设 A 非空. 根据命题 2.2.6, 存在 $g : B \to A$ 使得 $fg = \mathrm{id}_B$, 从而 g 是单射. 根据前一步, g 实际是可逆映射, 因此 $f = (fg)g^{-1} = g^{-1}$ 也可逆. □

命题 A.3.5 (= 命题 2.9.5) *集合 A 无穷当且仅当存在单射 $\mathbb{Z}_{\geq 0} \hookrightarrow A$.*

证明 思路是从无穷集 A 不断地挑出元素, 以构造 $\mathbb{Z}_{\geq 0}$ 的副本. 容易验证从 A 扣掉任何元素仍是无穷集, 从而扣掉任意有限多个元素亦然. 以选择公理从 A 的每个非空子集挑出一个元素, 选法具体表作映射

$$g : P(A) \smallsetminus \{\varnothing\} \to A, \quad \forall S,\ g(S) \in S.$$

对所有 $n \in \mathbb{Z}_{\geq 0}$, 考虑以下性质 $\mathcal{P}(n)$: 存在映射 $f_n : \{0, \dots, n\} \to A$, 使得对所有 $0 \leq k < n$ 皆有

$$f_n(k) = g(\underbrace{A \smallsetminus \{f_n(0), \dots f_n(k-1)\}}_{\neq \varnothing}).$$

根据对右式的标准理解, 自动有 $f_n(0) = g(A)$, 而 $f_n(k)$ 按选择函数 g 的定义必不属于 $\{f_n(0), \dots, f_n(k-1)\}$, 由此可见 f_n 是单射.

基于数学归纳法原理, 容易说明 $\mathcal{P}(n)$ 对所有 n 成立. 该原理还进一步说明每个 f_n 都唯一由 g 确定. 这点确保映射族 $(f_n)_{n\geq 0}$ 黏合为 $f : \mathbb{Z}_{\geq 0} \to A$, 使得 $f|_{\{0,\dots,n-1\}} = f_n$. 既然每个 f_n 皆单, 故 f 单. □

A.4 Zorn 引理与基的存在性

先回顾本书正文曾出现过的一些概念.

设 $(A, \leq)$ 为偏序集. 设 $a_{\max} \in A$, 若不存在 $a' \in A$ 使得 $a' > a_{\max}$, 则称 $a_{\max}$ 为 A 的一个极大元. 若 $a \in A$ 满足 $\forall c \in C,\ c \leq a$, 则称 a 为 C 在 A 中的一个上界.

若子集 $C \subset A$ 对 $\leq$ 成为全序集, 则称 C 为 A 中的链.

命题 A.4.1 (Zorn 引理) *设 $(A, \leq)$ 为非空偏序集, 而且其中的每个链 C 在 A 中都有上界, 则在 A 中存在极大元 $a_{\max}$.*

从 Zorn 引理推导选择公理是相对容易的. 选择公理断言对于任何一族非空集 A, 存在映射 $g : A \to \bigcup A$ 使得 $g(a) \in a$ 对所有 $a \in A$ 成立; 换言之, 存在从 A 中的每个集合 a 选取元素 $g(a)$ 的一种方法, 体现为 "选择映射" g. 为了说明选择映射存在, 定义偏序集 $(\mathcal{P}, \preceq)$ 如下.

- $\star$ $\mathcal{P}$ 的元素是资料 (A', g'), 其中 $A' \subset A$ 而 $g' : A' \to \bigcup A'$ 使得 $g'(a) \in a$ 对所有 $a \in A'$ 成立.
- $\star$ $(A', g') \preceq (A'', g'')$ 意谓 $A' \subset A''$ 而 $g''|_{A'} = g'$.

容易验证 $(\mathcal{P}, \preceq)$ 为偏序集, 而且它非空: 事实上, 我们知道如何理解从空集出发的映射, 见 §2.2, 因此取 $A' = \varnothing$ 毫无问题. 设 C 为 $(\mathcal{P}, \preceq)$ 中的链. 定义 $A_0 := \bigcup_{(A',g')\in C} A'$, 对所有 $a \in A_0$, 任取 $(A', g') \in C$ 使得 $a \in A'$, 再命 $g_0(a) := g'(a) \in A' \subset A_0$. 容易验证:

- $\star$ $g_0(a)$ 不依赖 (A', g') 的选取 (此处需要 C 作为链的性质);

⋆ $(A_0, g_0) \in \mathcal{P}$;

⋆ $(A', g') \preceq (A_0, g_0)$ 对所有 $(A', g') \in C$ 成立.

代入 Zorn 引理遂知 $(\mathcal{P}, \preceq)$ 有极大元 $(A_{\max}, g_{\max})$. 若能证明 $A_{\max} = A$, 则 $g_{\max}$ 即所求的选择映射. 设若不然, 存在 $a \in A \smallsetminus A_{\max}$, 则可任取 $x \in a$, 再定义

$$\tilde{g} : A_{\max} \sqcup \{a\} \to (\bigcup A_{\max}) \cup a$$

使得 $\tilde{g}|_{A_{\max}} = g_{\max}$ 而 $\tilde{g}(a) = x$. 这与 $(A_{\max}, g_{\max})$ 对 $\preceq$ 的极大性质矛盾.

从选择公理推导 Zorn 引理较为复杂. 有兴趣的读者可以参考 [10, 定理 1.3.6] 或其他集合论相关教材.

基于 Zorn 引理, 现在可以容易地说明向量空间总有基. 回忆到空子集按定义是线性无关子集.

命题 A.4.2 (= 定义–命题 4.4.10 的前半部) *设 F 为域, V 为 F-向量空间, 而 S 为 V 的线性无关子集, 则 S 可以扩充为 V 的基. 特别地, 取 $S = \varnothing$ 可知 V 总是有基.*

证明 定义集合

$$\mathcal{P} := \big\{T \subset V : \text{线性无关子集}, T \supset S\big\}.$$

以包含关系 $\subset$ 赋予 $\mathcal{P}$ 偏序. 如能证明 $\mathcal{P}$ 有极大元 B, 则 B 必是极大线性无关子集, 故给出包含 S 的基.

为此, 设 C 为 $\mathcal{P}$ 中的链. 定义 $T' := \bigcup_{T \in C} T$. 它当然也包含 S, 以下说明 T' 线性无关: 任何线性关系式 $\sum_{t \in T'} a_t t$ 只涉及有限个非零系数, 设其为 $t_1, \dots, t_n$; 对每个 $1 \leq i \leq n$ 取 $T_i \in C$ 使得 $t_i \in T_i$. 基于全序性质, 重排后不妨设 $T_1 \subset \cdots \subset T_n$, 则 $\sum_{i=1}^{n} a_{t_i} t_i = 0$ 是 T_n 中的线性关系, 故系数必然全为 0.

这就说明 $\mathcal{P}$ 是每个链 C 都有上界 T' 的偏序集; 又由 $S \in \mathcal{P}$ 可知它非空, 故 Zorn 引理确保 $\mathcal{P}$ 有极大元. 明所欲证. □

以上仅是定义–命题 4.4.10 的前半部内容; 后半部关乎维数的唯一性, 由于这在无穷维情形需要一些基数的性质, 而且用处相对少, 在此不予处理.

此外, 这些结论也能够从域上的向量空间推广到除环上的模, 论证完全相同. 请参见例 12.4.7 的讨论.

附录 B　范畴引论

范畴论作为一套数学语言具有高度的概括力与启发性, 它的应用不限于代数, 然而在代数学中体现得最为直接. 以逻辑顺序而言, 范畴论可以置于一切代数结构之前; 然而就学习而言, 若无法对代数结构掌握足够的例子和经验, 就无法切实理解范畴论中的构造, 因此本书将相关内容置于附录.

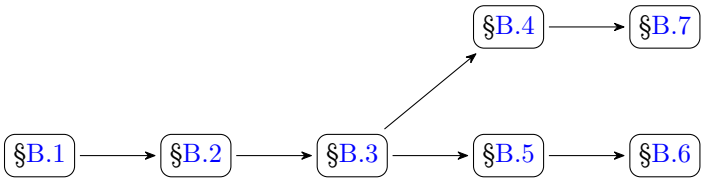
实例至关重要. 本篇附录所举例子将涉及本书正文中处理的各种结构, 预设一定的背景知识. 最低限度应了解集合的操作以及域上的向量空间; 其次, 关于序结构和群论的常识也有益处, 如果了解模论则更佳. 此外, 尽力完成练习对于熟悉相关概念是必要的.

本篇附录仅是引论, 既未涵盖范畴论的所有基本概念 (如伴随性, 极限, 完备性等), 在所涵盖的内容中也不及实践所需的深度. 读者若希望进一步了解范畴论或在其他领域中予以应用, 应当进一步研读其他书籍, 例如 [10, 第二章, 第三章].

阅读提示

本书在定义范畴时采取的集合论假设和 [10] 稍有不同, 详见注记 B.1.16. 倘若读者只关心泛性质 (见 §B.5), 则这些细节毫无影响.

阅读顺序



B.1 范畴

本书探讨了形形色色的数学结构, 包括但不限于集合, 环, 域, 向量空间, 群. 单从形式操作的层面观之, 每一种理论都涉及一类指定的 "对象" (例如集合, 环), 对象 X 和 Y 之间的种种 "箭头" $f : X \to Y$ (例如集合之间的映射, 环之间的环同态), 以及头尾相接的箭头之间的 "合成" 运算. 尽管箭头对于每种结构都有不同意义, 但合成运算有诸多共性, 譬如任何对象 X 到自身有恒等 id_X, 箭头的合成服从结合律 $(fg)h = f(gh)$, 而恒等满足 $f\,\mathrm{id}_X = f = \mathrm{id}_Y\, f$.

进一步, 无论在哪种理论中, 关于箭头的左可逆性, 右可逆性和可逆性的概念总是按相同方式定义的, 而当 $f : X \to Y$ 可逆时, 关于 f^{-1} 唯一性的证明无论在哪种情形都有同样的格式. 因此, "同构" 的概念也总是具有相同的格式; 譬如对于集合论的情形, 所谓同构无非是集合等势. 本书探讨的许多课题都相当于对同构类的精细研究.

除此之外, 本书所频繁运用的交换图表也仅涉及对象及箭头的形式操作, 并不关心它们具体 "是什么". 对于掌握本书内容的读者, 交换图表的这一特色及随之而来的便利应当是不言而喻的.

基于这些经验教训, 下述定义自然而然.

定义 B.1.1 (范畴) 一个范畴 $\mathcal{C}$ 意谓以下资料:

- ⋆ 一个类 $\mathrm{Ob}(\mathcal{C})$, 其元素称作 $\mathcal{C}$ 的**对象**.
- ⋆ 对所有 $X, Y \in \mathrm{Ob}(\mathcal{C})$ 指定一个集合 $\mathrm{Hom}_{\mathcal{C}}(X, Y)$; 此集合称为从 X 到 Y 的 Hom-集, 也简记为 $\mathrm{Hom}(X, Y)$, 其元素称作 $\mathcal{C}$ 中从 X 到 Y 的**态射**.
- ⋆ 对每个 $X \in \mathrm{Ob}(\mathcal{C})$ 指定元素 $\mathrm{id}_X \in \mathrm{Hom}_{\mathcal{C}}(X, X)$, 称为 X 到自身的**恒等态射**.
- ⋆ 对于任意 $X, Y, Z \in \mathrm{Ob}(\mathcal{C})$, 指定态射间的**合成映射**

$$\begin{aligned} \circ : \mathrm{Hom}_{\mathcal{C}}(Y, Z) \times \mathrm{Hom}_{\mathcal{C}}(X, Y) &\longrightarrow \mathrm{Hom}_{\mathcal{C}}(X, Z) \\ (g, f) &\longmapsto g \circ f, \end{aligned}$$

不致混淆时也将 $g \circ f$ 简记为 gf. 它满足:

▷ **结合律** 对 $\mathcal{C}$ 中的所有态射 h, g, f, 若合成 $h(gf)$ 和 $(hg)f$ 都有定义, 则

$$h(gf) = (hg)f.$$

故两边可以同写为 $h \circ g \circ f$ 或 hgf;

▷ **恒等态射的性质** 对所有对象 X, Y 和态射 $f \in \mathrm{Hom}_{\mathcal{C}}(X, Y)$, 我们有

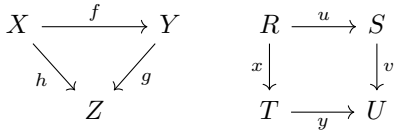
$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f.$$

恒等态射 id_X 由其性质唯一确定, 这是因为如果 id_X 和 id'_X 都满足恒等映射的性质, 则

$$\mathrm{id}_X = \mathrm{id}_X \mathrm{id}'_X = \mathrm{id}'_X.$$

一般也将 $f \in \mathrm{Hom}_{\mathcal{C}}(X,Y)$ 写作 $f : X \to Y$ 或 $X \xrightarrow{f} Y$, 故态射又称为箭头. 态射的合成诠释为箭头的头尾相接; 当态射的名称 f 自明或不重要时, 常从图中略去.

注记 B.1.2 (范畴中的交换图表) 基于箭头的图解表法是范畴论的基本语言. 其中最常用的是交换图表. 所谓 “交换”, 意谓箭头合成殊途同归, 例如图表



的交换性分别相当于说 $gf = h$ 和 $vu = yx$. 本书已经在种种具体场合中使用交换图表, 此处不过是在范畴论的视野下给出一般解释.

如前所述, 种种数学结构自然地构成范畴. 以下例子涉及的层面相对广, 假定读者至少对其中一部分结构有所了解.

例 B.1.3 (集合范畴) 所有集合构成一个范畴, 记为 Set. 它的对象是所有集合构成的类. 对任两个集合 X 和 Y, 态射集 $\mathrm{Hom}_{\mathsf{Set}}(X,Y)$ 是映射集 $Y^X = \{f : X \to Y \text{ 映射}\}$, 态射 $f : X \to Y$ 和 $g : Y \to Z$ 的合成 $gf : X \to Z$ 是映射的合成. 恒等 id_X 是集合 X 到自身的恒等映射, 这在 $X = \varnothing$ 时也有定义. 范畴定义中的结合律和恒等态射的性质都化为集合论的常识.

注意到 $\mathrm{Ob}\,(\mathsf{Set})$ 确实必须是一个类, 因为全体集合并不构成一个集合, 而是所谓 “真类”.

例 B.1.4 (代数结构给出的范畴) 每一种类型的代数结构都自然地构成一个范畴. 例如环范畴 Ring 的定义是 $\mathrm{Ob}(\mathsf{Ring}) = \{R : \text{环}\}$ (这依然是类), 而对所有环 R 和 R' 定义 $\mathrm{Hom}_{\mathsf{Ring}}(R,R') = \{\text{环同态 } \varphi : R \to R'\}$. 其他代数结构当然也按类似格式成为范畴, 下表仅是一部分例子.

范畴	对象	态射	合成运算	恒等态射
Ring	环	环同态	映射合成	恒等映射
Field	域	域嵌入		
Vect(F)	域 F 上的向量空间	线性映射		
$\mathrm{Vect}_{\mathrm{f}}(F)$	域 F 上的有限维向量空间			
Mon	幺半群	幺半群同态		
Grp	群	群同态		
Ab	交换群			
R-Mod	环 R 上的左模	模同态		
Mod-R	环 R 上的右模	模同态		
F-Alg	域 F 上的代数	代数同态		

留意到以上的 Ring 含零环, 这与 [10] 的规定有所出入, 但终究只是定义问题. 关于幺半群同态和群同态的比较请见注记 11.2.2.

在 F-向量空间范畴 Vect(F) 的例子中, 我们知道任两个对象 V, V' 之间的 $\mathrm{Hom}(V, V')$ 不只是集合, 还自然地成为 F-向量空间, 而态射的合成

$$\mathrm{Hom}(V', V'') \times \mathrm{Hom}(V, V') \to \mathrm{Hom}(V, V'')$$

对此是双线性的. 附加于 Hom 集的这类额外结构称为范畴的**充实**, 实践中频繁出现. 相关理论可参考 [10, §3.4].

例 B.1.5 (其他数学结构给出的范畴) 除了代数结构之外, 数学中出现的其他结构往往也构成范畴. 下表只是较为初等而具有代表性的几则例子.

范畴	对象	态射	合成运算	恒等态射
Preordset	预序集	保序映射	映射合成	恒等映射
Poset	偏序集			
Metr	度量空间	保距映射		
Rel	集合	二元关系	二元关系合成	恒等关系

因此范畴语言对于分析或几何学的研究也是有益的.

关于 Rel 的例子可能需要解释. 回忆到集合 A 和 B (计顺序) 之间的二元关系定义为任意子集 $R \subset A \times B$; 定义二元关系 $R \subset A \times B$ 和 $S \subset B \times C$ 的合成为

$$SR := \{(a,c) \in A \times C : \exists b \in B,\ (a,b) \in R \wedge (b,c) \in S\},$$

而 $A \times A$ 上的恒等关系定义为对角子集 $\{(a,a) : a \in A\}$, 亦即 $=$ 给出的二元关系. 关于 Rel 的范畴论公理不过是按步就班的简单验证, 已在相关章节的练习中做过.

例 B.1.6 (序结构作为范畴) 任选一个集合 $\star$. 设 $(A, \preceq)$ 是预序集. 定义相应的范畴 $\mathcal{A}$ 如下:

$$\begin{aligned} \mathrm{Ob}(\mathcal{A}) &= A, \\ \mathrm{Hom}_{\mathcal{A}}(X,Y) &= \begin{cases} \text{独点集 } \{\star\}, & X \preceq Y, \\ \varnothing, & \text{其他情形}, \end{cases} \\ \mathrm{id}_X &= \text{唯一元素 } \star \in \mathrm{Hom}_{\mathcal{A}}(X,X). \end{aligned}$$

态射 $X \to Y$ 和 $Y \to Z$ 的合成只在 $X \preceq Y$ 而 $Y \preceq Z$ 时才有意义; 由于 $\preceq$ 的传递性确保 $X \preceq Z$ 也成立, 此时的 Hom 集都是独点集, 合成运算能且只能定义为 $\star \circ \star = \star$. 合成的结合律与恒等态射的性质因而都是平凡的.

特别地, 任何偏序集 $(P, \preceq)$ 都给出范畴 $\mathcal{P}$. 与此前介绍的例子不同, 预序集确定的范畴是 "小" 的: 它们的所有对象构成集合, 而不仅只是类.

例 B.1.7 作为前一则例子的应用, 我们得到:

▷ **空范畴 0** 对应到空集给出的偏序集, 这是满足 $\mathrm{Ob}(\mathcal{C}) = \varnothing$ 的唯一范畴 $\mathcal{C}$.

▷ **范畴 1** 对应到仅有一个元素的偏序集, 这是恰有一个对象 (记为 0) 及一个态射 id_0 的范畴, 合成运算由 $\mathrm{id}_0\mathrm{id}_0 = \mathrm{id}_0$ 道尽.

▷ **范畴 $\mathbf{n}$** 推而广之, 对任意 $n \in \mathbb{Z}_{\geq 0}$, 全序集 $\{0, \ldots, n-1\}$ 对应的范畴记为 $\mathbf{n}$, 其态射不外是恒等态射, 下图的所有箭头

$$0 \to 1 \to \cdots \to n-1,$$

及其合成.

▷ **离散范畴** 对任意集合 S 皆可赋予偏序 $\preceq$ 使得

$$s \preceq s' \iff s = s';$$

此偏序集对应的范畴 $\mathsf{Disc}(S)$ 称为由 S 确定的离散范畴, 它满足 $\mathrm{Ob}\left(\mathsf{Disc}(S)\right) = S$ 而

$$\mathrm{Hom}_{\mathsf{Disc}(S)}(s,s') = \begin{cases} \{\mathrm{id}_s\}, & s = s', \\ \varnothing, & s \neq s', \end{cases}$$

态射的合成运算由此唯一确定. 观察到 $\mathrm{Disc}(\varnothing) = \mathbf{0}$ 而 $\mathrm{Disc}(\text{独点集}) = \mathbf{1}$.

回到范畴的一般理论. 下述概念在集合与向量空间等具体情境中早已熟悉.

定义 B.1.8 给定范畴 $\mathcal{C}$. 对于其中的态射 $f: X \to Y$, 若存在 $g: Y \to X$ 使得 $fg = \mathrm{id}_Y$ 而 $gf = \mathrm{id}_X$, 则称 f 是**同构** (或称可逆, 写作 $f: X \xrightarrow{\sim} Y$), 称 g 为 f 的**逆**, 记为 f^{-1}.

如果 f 和 g 满足 $fg = \mathrm{id}_Y$, 则称 f 是 g 的一个左逆, g 是 f 的一个右逆.

记法 f^{-1} 合理, 这是因为逆若存在则唯一: 若 g, g' 都是态射 f 的逆, 则有 $g = g\,\mathrm{id}_Y = g(fg') = (gf)g' = \mathrm{id}_X g' = g'$. 与此相反, f 的左逆 (或右逆) 未必唯一. 有左逆 (或右逆) 的态射称为**左可逆** (或**右可逆**) 的.

以集合范畴 Set 为例, 其中的左可逆 (或右可逆) 态射等价于单射 (或满射), 同构等价于双射. 这些事实在关于集合的章节中已有详细解释.

命题 B.1.9 考虑范畴 $\mathcal{C}$ 中的态射 $f: X \to Y$.

(i) 若 f 左可逆 (或右可逆), 则它对合成满足左 (或右) 消去律: $fg = fg' \iff g = g'$ (或 $gf = g'f \iff g = g'$) 恒成立, 其中 $g, g': Z \to X$ (或 $g, g': Y \to Z$) 是态射, $Z \in \mathrm{Ob}(\mathcal{C})$.

(ii) 态射 f 可逆当且仅当它既有左逆也有右逆, 此时左逆和右逆皆唯一, 并且皆等于 f^{-1}.

(iii) 若 f 是同构, 则 f^{-1} 亦然, 而且 $(f^{-1})^{-1} = f$.

(iv) 恒等态射 id_X 是同构, 它自为逆. 设 $f: X \to Y$ 与 $g: Y \to Z$ 皆为同构, 则 gf 也是同构, 而且 $(gf)^{-1} = f^{-1}g^{-1}$.

证明 与集合的场景丝毫不差, 请读者回忆或自行练习. □

注记 B.1.10 对合成满足上述左 (或右) 消去律的态射称为单 (或满) 态射.

如果存在从对象 X 到 Y 的同构 f, 则记为 $X \simeq Y$. 命题 B.1.9 (iv) 说明 $\simeq$ 是 $\mathrm{Ob}(\mathcal{C})$ 上的等价关系.

约定 B.1.11 对于 $X \in \mathrm{Ob}(\mathcal{C})$, 记

$$\begin{aligned} \mathrm{End}_{\mathcal{C}}(X) &:= \mathrm{Hom}_{\mathcal{C}}(X, X), \\ \mathrm{Aut}_{\mathcal{C}}(X) &:= \{f \in \mathrm{End}_{\mathcal{C}}(X) : \text{ 可逆}\}. \end{aligned}$$

上述集合对二元运算 $\circ$ 封闭, $\mathrm{End}_{\mathcal{C}}(X)$ 是幺半群, 而 $\mathrm{Aut}_{\mathcal{C}}(X)$ 是群, 分别称为 X 的自同态幺半群和自同构群. 这种构造也是我们司空见惯的.

例 B.1.12 范畴 Set 的对象 $\{1, \ldots, n\}$ 的自同构群是对称群 $\mathfrak{S}_n$. 设 F 为域, 则 $\mathsf{Vect}(F)$ 的对象 V 的自同构群是一般线性群 $\mathrm{GL}(V)$, 自同态幺半群则是 $\mathrm{End}(V)$. 若 $\mathrm{char}(F) \neq 2$, 命 $\mathsf{Quad}(F)$ 为 F 上的全体二次型所成范畴 (见 §16.1), 以二次型之间的同构为态射, 则 $\mathsf{Quad}(F)$ 的对象 (V, q) 的自同构群是正交群 $\mathrm{O}(V, q)$.

练习 B.1.13 设 M 为幺半群. 说明可定义范畴 $\mathcal{M}$ 使得它仅有一个对象 $\star$, 而且有幺半群的等式 $\mathrm{End}_{\mathcal{M}}(\star) = M$.

范畴既然也是一种数学结构, 对之便有子结构的概念, 但定义方法和代数结构的子结构略有不同.

定义 B.1.14 设 $\mathcal{C}$ 和 $\mathcal{C}'$ 为范畴. 当以下性质成立时, 称 $\mathcal{C}'$ 是 $\mathcal{C}$ 的**子范畴**:

(i) $\mathrm{Ob}(\mathcal{C}') \subset \mathrm{Ob}(\mathcal{C})$;

(ii) 对所有 $X, Y \in \mathrm{Ob}(\mathcal{C}')$ 皆有 $\mathrm{Hom}_{\mathcal{C}'}(X,Y) \subset \mathrm{Hom}_{\mathcal{C}}(X,Y)$;

(iii) 对所有 $X \in \mathrm{Ob}(\mathcal{C}')$, 它相对于 $\mathcal{C}'$ 和 $\mathcal{C}$ 的恒等态射相同;

(iv) 态射在 $\mathcal{C}'$ 中的合成运算是由 $\mathcal{C}$ 限制而来的.

如果进一步对所有 $X, Y \in \mathrm{Ob}(\mathcal{C}')$ 要求 $\mathrm{Hom}_{\mathcal{C}'}(X,Y) = \mathrm{Hom}_{\mathcal{C}}(X,Y)$, 则称 $\mathcal{C}'$ 是**全子范畴**.

例 B.1.15 借用集合论的符号 $\mathcal{C}' \subset \mathcal{C}$ 来代表 $\mathcal{C}'$ 是 $\mathcal{C}$ 的全子范畴. 在例 B.1.4 介绍的例子中,

$$\mathsf{Field} \subset \mathsf{Ring}, \quad \mathsf{Vect}_{\mathrm{f}}(F) \subset \mathsf{Vect}(F), \quad \mathsf{Ab} \subset \mathsf{Grp} \subset \mathsf{Mon}.$$

在例 B.1.5 介绍的例子中, $\mathsf{Poset} \subset \mathsf{Preordset}$; 此外, 倘若读者记得映射 $f : A \to B$ 在集合论中被定义为一类特殊的二元关系 $\Gamma_f \subset A \times B$, 则不难看出 Set 是 Rel 的子范畴, 但不是全子范畴.

最后考虑例 B.1.6 从序结构得到的范畴. 如果 A' 是预序集 $(A, \preceq)$ 的子集, 赋予 A' 限制而来的序结构, 则相应地有 $\mathcal{A}' \subset \mathcal{A}$. 另一方面, 若忘却 $(A, \preceq)$ 的序结构, 视之为集合再取离散范畴 $\mathsf{Disc}(A)$, 则 $\mathsf{Disc}(A)$ 是 $\mathcal{A}$ 的子范畴, 但一般不是全子范畴.

注记 B.1.16 (范畴与集合论) 本书在定义 B.1.1 中涉及了 "类" 的概念, 并且举例说明其必要性, 这或许会让一些 ZFC 集合论的用户感到忧虑. 在一些文献中, 范畴定义中的 $\mathrm{Ob}(\mathcal{C})$ 与每个 $\mathrm{Hom}_{\mathcal{C}}(X,Y)$ 均要求是集合, 而 Hom 集的大小则往往有所限制.

例如 [10] 便采取了 Grothendieck 宇宙 $\mathcal{U}$ 的概念, 这是一族可视为 "小" 的集合, 详见 [10, 定义 1.5.1]; 定义 B.1.1 中的类 $\mathrm{Ob}(\mathcal{C})$ 在该书中对应于一般的集合, 而集合 $\mathrm{Hom}_{\mathcal{C}}(X,Y)$ 在该书中则对应于和 $\mathcal{U}$ 中的某个集合等势的集合 (称为 $\mathcal{U}$-小集). 综之, 本书的范畴对应于 [10, 定义 2.1.3] 所谓的 $\mathcal{U}$-范畴[1], 而例 B.1.3—B.1.5 在此框架下需要适当改写[2], 以避免真类.

在关于范畴的进阶操作中[3], 类或者集合的大小问题需要严肃面对, 本书不涉及这些面向, 唯一例外是定理 B.4.4 证明的脚注.

[1] 另一些文献则称之为局部小范畴.

[2] 更具体地说, 需要将所论的集合, 环等结构建立在属于 $\mathcal{U}$ 的集合上, 换言之建立在合理地 "小" 的集合上, 使得这些结构的全体构成集合而非真类.

[3] 譬如构造 "范畴的范畴".

B.2 函子

一如映射之于集合, 同态之于代数结构, 范畴之间也有称为函子的概念作为桥梁. 函子是分别定义在对象与态射层次的一族映射, 使得范畴的基本操作 (态射合成, 恒等态射) 得到保持.

定义 B.2.1 (函子) 设 $\mathcal{C}', \mathcal{C}$ 为范畴. 从 $\mathcal{C}'$ 到 $\mathcal{C}$ 的函子 F 意谓以下资料:

(i) 为 $\mathcal{C}'$ 的每个对象 X 唯一地指定 $\mathcal{C}$ 的对象 FX, 也用映射的符号写作 $F : \mathrm{Ob}(\mathcal{C}') \to \mathrm{Ob}(\mathcal{C})$,

(ii) 为 $\mathcal{C}'$ 的所有对象 X, Y 指定 Hom-集之间以相同符号标记的映射

$$F : \mathrm{Hom}_{\mathcal{C}'}(X, Y) \to \mathrm{Hom}_{\mathcal{C}}(FX, FY),$$

条件是要求

$$F(gf) = F(g)F(f), \quad F(\mathrm{id}_X) = \mathrm{id}_{FX}$$

对 $\mathcal{C}'$ 中所有可合成的态射 g, f 以及所有对象 X 皆成立.

我们习惯将上述函子 F 表作 $F : \mathcal{C}' \to \mathcal{C}$. 给定函子 $F : \mathcal{C}_1 \to \mathcal{C}_2$ 和 $G : \mathcal{C}_2 \to \mathcal{C}_3$, 合成函子 $GF = G \circ F : \mathcal{C}_1 \to \mathcal{C}_3$ 的定义是在对象层次取合成映射

$$\mathrm{Ob}(\mathcal{C}_1) \xrightarrow{F} \mathrm{Ob}(\mathcal{C}_2) \xrightarrow{G} \mathrm{Ob}(\mathcal{C}_3),$$

而在态射层次取合成映射

$$\mathrm{Hom}_{\mathcal{C}_1}(X, Y) \xrightarrow{F} \mathrm{Hom}_{\mathcal{C}_2}(FX, FY) \xrightarrow{G} \mathrm{Hom}_{\mathcal{C}_3}(GF(X), GF(Y)).$$

注意到定义中采用了惯常的简写 $FX = F(X)$ 和 $Ff = F(f)$. 函子的合成当然地服从结合律 $(HG)F = H(GF)$.

以下概念可以粗略地设想为映射的单性和满性对于函子的类比.

定义 B.2.2 设 $F : \mathcal{C}' \to \mathcal{C}$ 为函子.

- ⋆ 如果对所有 $X, Y \in \mathrm{Ob}(\mathcal{C}')$, 映射 $F : \mathrm{Hom}_{\mathcal{C}'}(X, Y) \to \mathrm{Hom}_{\mathcal{C}}(FX, FY)$ 皆为单射 (或满射, 双射), 则称 F 为**忠实** (或**全**, **全忠实**) 的.
- ⋆ 如果对所有 $T \in \mathrm{Ob}(\mathcal{C})$ 都存在 $X \in \mathrm{Ob}(\mathcal{C}')$ 使得 $T \simeq FX$, 则称 F 为**本质满**的.

例 B.2.3 设 $\mathcal{C}'$ 为 $\mathcal{C}$ 的子范畴, 则有忠实的包含函子 $\iota : \mathcal{C}' \to \mathcal{C}$, 它在对象和态射层次都是包含映射; 它是全忠实的当且仅当 $\mathcal{C}'$ 是 $\mathcal{C}$ 的全子范畴.

定义 B.2.4 范畴 $\mathcal{C}$ 的**反范畴** $\mathcal{C}^{\mathrm{op}}$ 定义如下:

- ⋆ $\mathrm{Ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{Ob}(\mathcal{C})$,
- ⋆ 对所有对象 X, Y, $\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(X, Y) := \mathrm{Hom}_{\mathcal{C}}(Y, X)$,
- ⋆ 态射 $f \in \mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(Y, Z)$ 与 $g \in \mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(X, Y)$ 在 $\mathcal{C}^{\mathrm{op}}$ 中的合成 $f \circ^{\mathrm{op}} g$ 定义为 $\mathcal{C}$ 中的反向合成 $g \circ f$,
- ⋆ 恒等态射的定义与 $\mathcal{C}$ 中相同.

容易看出 $\mathcal{C}^{\mathrm{op}}$ 确实满足范畴的定义, 而且 $(\mathcal{C}^{\mathrm{op}})^{\mathrm{op}} = \mathcal{C}$.

简言之, $\mathcal{C}^{\mathrm{op}}$ 的功效是反转箭头. 如果我们能对所有范畴 $\mathcal{C}$ 证明某性质 $\mathbf{P}(\mathcal{C})$, 则将 $\mathbf{P}(\mathcal{C})$ 涉及的所有箭头反转后的性质依然对 $\mathcal{C}$ 成立, 因为反转后的陈述等价于 $\mathbf{P}(\mathcal{C}^{\mathrm{op}})$. 这是范畴论中一种简单而方便的对偶原理.

举命题 B.1.9 (i) 的内容为例. 假定我们已经证明了 $\mathcal{C}$ 中的左可逆态射满足左消去律. 为了说明右可逆态射满足右消去律, 我们或者可以重复论证, 或者含糊其词地诉诸左右对称性, 而更严格的方式则是注意到 $\mathcal{C}$ 的右可逆态射相当于 $\mathcal{C}^{\mathrm{op}}$ 的左可逆态射 (箭头反转导致合成反向), $\mathcal{C}$ 中的右消去律转译为 $\mathcal{C}^{\mathrm{op}}$ 的左消去律; 综之, 通过 $\mathcal{C} \leftrightarrow \mathcal{C}^{\mathrm{op}}$ 作切换, 左右两种版本择一证明即可.

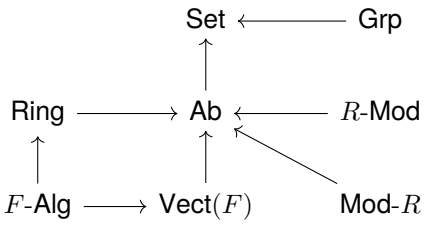
注记 B.2.5 在一些旧文献中, 定义 B.2.1 中的函子称为共变函子, 与此相对的则是反变函子 $F : \mathcal{C}' \to \mathcal{C}$, 其定义在对象层次照旧, 在态射层次则改为

$$F : \mathrm{Hom}_{\mathcal{C}'}(X, Y) \to \mathrm{Hom}_{\mathcal{C}}(FY, FX),$$

并且 $F(gf) = F(f)F(g)$, $F(\mathrm{id}_X) = \mathrm{id}_{FX}$; 换言之, 要求 F 反转箭头方向. 用定义 B.2.4 的语言来说, 从 $\mathcal{C}'$ 到 $\mathcal{C}$ 的反变函子无非是共变函子 $\mathcal{C}' \to \mathcal{C}^{\mathrm{op}}$ 或 $(\mathcal{C}')^{\mathrm{op}} \to \mathcal{C}$, 不必另起炉灶.

函子在实践中自然地出现. 以下同样从熟知的代数结构中举例, 要求读者至少对其中一部分的实例有所了解.

例 B.2.6 (忘却函子) 代数结构通常是叠架定义的, 譬如忘却环的乘法, 便能将环视为加法群; 忘却所有结构便得到底层的集合. 这种考量给出所谓的忘却函子. 回顾例 B.1.3 与例 B.1.4 出现的一些范畴, 我们有下图描述的种种忘却函子



其中 R (或 F) 是环 (或域). 在对象层次, 这些函子忘却一部分的代数结构 (环的乘法, 向量空间的纯量乘法...). 在态射层次, 这些函子化为 Hom-集之间的包含映射; 譬如对

于任两个环 R, R', 映向 Ab 的忘却在态射集上是

$$\mathrm{Hom}_{\mathsf{Ring}}(R,R') = \{\text{环同态 } R\to R'\} \overset{\text{包含}}{\hookrightarrow} \{\text{加法群同态 } R\to R'\} = \mathrm{Hom}_{\mathsf{Ab}}(R,R').$$

特别地, 这些忘却函子都是忠实的, 然而并非全忠实.

例 B.2.7 (对偶空间函子) 选定域 F. 对 F-向量空间 V 定义其对偶空间为 F-向量空间 $V^\vee = \mathrm{Hom}(V,F)$. 这给出函子

$$D : \mathsf{Vect}(F)^{\mathrm{op}} \to \mathsf{Vect}(F).$$

它在对象层次是 $V\mapsto D(V) := V^\vee$; 在态射层次, 线性映射 $T: V\to W$ 的像 $D(T)$ 取为

$$\begin{aligned} {}^{\mathrm{t}}T : W^\vee &\to V^\vee \\ \lambda &\mapsto \lambda T. \end{aligned}$$

基于向量空间理论中熟知的 ${}^{\mathrm{t}}(ST) = {}^{\mathrm{t}}T\ {}^{\mathrm{t}}S$ 和 ${}^{\mathrm{t}}(\mathrm{id}_V) = \mathrm{id}_{V^\vee}$, 易见 D 确实是函子. 上述构造也能写作 $\mathsf{Vect}(F)\to\mathsf{Vect}(F)^{\mathrm{op}}$ 的形式, 另记为 D^{op} 以资区别.

例 B.2.8 (交换幺半群的群化) 记 CMon 为交换幺半群所成范畴, 其对象 M 上的二元运算 (或幺元, 亦即零元) 以加法符号记作 $+$ (或 0). 在集合 $M\times M$ 上有等价关系

$$(a,b)\sim(a',b') \iff \exists c\in M,\ a+b'+c = a'+b+c;$$

记 (a,b) 的等价类为 $[\![a,b]\!]$. 赋予集合 $M^{\mathrm{grp}} := (M\times M)/\sim$ 二元运算

$$[\![a,b]\!] + [\![c,d]\!] = [\![a+c, b+d]\!].$$

相信读者能证明这一切都是良定义的, 而且 M^{grp} 对 $+$ 成为交换群, $-[\![a,b]\!] = [\![b,a]\!]$, 零元是等价类 $[\![0,0]\!]$. 进一步, 我们有幺半群的同态

$$\begin{aligned} \eta_M : M &\to M^{\mathrm{grp}} \\ a &\mapsto [\![a,0]\!], \end{aligned}$$

它满足 $[\![a,b]\!] = [\![a,0]\!] + [\![0,b]\!] = \eta_M(a) - \eta_M(b)$.

此外, 给定幺半群同态 $f: M_1\to M_2$, 存在唯一的群同态 f^{grp} 使下图交换:

$$\begin{array}{ccc} M_1 & \xrightarrow{\ f\ } & M_2 \\ {\scriptstyle \eta_{M_1}}\downarrow & & \downarrow{\scriptstyle \eta_{M_2}} \\ M_1^{\mathrm{grp}} & \xrightarrow[f^{\mathrm{grp}}]{} & M_2^{\mathrm{grp}} \end{array}$$

具体描述是 $f^{\mathrm{grp}}([\![a,b]\!]) = [\![f(a),f(b)]\!]$; 特别地, $\mathrm{id}_M^{\mathrm{grp}} = \mathrm{id}_{M^{\mathrm{grp}}}$. 所以 $(\cdot)^{\mathrm{grp}}$ 给出函子 $\mathsf{CMon}\to\mathsf{Ab}$. 这是对交换幺半群进行 "群化" 的自然方式.

取特例 $M = \mathbb{Z}_{\geq 0}$, 则因为其加法有消去律, 等价关系 $\sim$ 简化为 $[\![a,b]\!]\sim[\![a',b']\!] \iff a+b' = a'+b$. 一切化约为从 $\mathbb{Z}_{\geq 0}$ 造 $\mathbb{Z}$ 的标准构造, 给出 $(\mathbb{Z}_{\geq 0})^{\mathrm{grp}} = \mathbb{Z}$.

下面介绍的 Hom 函子属于更一般的构造.

定义 B.2.9 (Hom 函子) 给定范畴 $\mathcal{C}$ 及其对象 X, 我们有函子

$$\mathrm{Hom}_{\mathcal{C}}(X,\cdot):\mathcal{C}\to\mathsf{Set},$$

它映对象 Y 为集合 $\mathrm{Hom}_{\mathcal{C}}(X,Y)$, 映态射 $f:Y\to Z$ 为映射

$$\begin{aligned} f_*:\mathrm{Hom}_{\mathcal{C}}(X,Y)&\to\mathrm{Hom}_{\mathcal{C}}(X,Z)\\ h&\mapsto fh. \end{aligned}$$

请读者验证这确实是函子, 所需性质归结为 $\mathcal{C}$ 中态射的结合律与恒等态射的性质.

我们也可以将 X 置于 Hom 的第二个位置. 由此得到的函子是反变的, 写作

$$\mathrm{Hom}_{\mathcal{C}}(\cdot,X):\mathcal{C}^{\mathrm{op}}\to\mathsf{Set},$$

它映对象 Y 为 $\mathrm{Hom}_{\mathcal{C}}(Y,X)$, 映态射 $f:Y\to Z$ 为

$$\begin{aligned} f^*:\mathrm{Hom}_{\mathcal{C}}(Z,X)&\to\mathrm{Hom}_{\mathcal{C}}(Y,X)\\ g&\mapsto gf, \end{aligned}$$

其原理依旧.

以上构造能进一步对 Hom 的两个位置同时操作, 给出函子 $\mathrm{Hom}_{\mathcal{C}}:\mathcal{C}^{\mathrm{op}}\times\mathcal{C}\to\mathsf{Set}$. 严格表述与验证并不难, 见 [10, 例 2.3.3] 或请读者练习, 重点是需要积范畴的以下的定义.

定义 B.2.10 (积范畴) 给定范畴 $\mathcal{C}_1$ 和 $\mathcal{C}_2$, 定义积范畴 $\mathcal{C}_1\times\mathcal{C}_2$ 如下.

▷ **对象** 命 $\mathrm{Ob}(\mathcal{C}_1\times\mathcal{C}_2)=\mathrm{Ob}(\mathcal{C}_1)\times\mathrm{Ob}(\mathcal{C}_2)$; 换言之 $\mathcal{C}_1\times\mathcal{C}_2$ 的对象是形如 (X_1,X_2) 的对, X_i 是 $\mathcal{C}_i$ 的对象 $(i=1,2)$.

▷ **态射** 命 $\mathrm{Hom}_{\mathcal{C}_1\times\mathcal{C}_2}((X_1,X_2),(Y_1,Y_2))=\mathrm{Hom}_{\mathcal{C}_1}(X_1,Y_1)\times\mathrm{Hom}_{\mathcal{C}_2}(X_2,Y_2)$.

态射的合成取为

$$\begin{aligned} \mathrm{Hom}((Y_1,Y_2),(Z_1,Z_2))\times\mathrm{Hom}((X_1,X_2),(Y_1,Y_2))&\to\mathrm{Hom}((X_1,X_2),(Z_1,Z_2))\\ ((g_1,g_2),(f_1,f_2))&\mapsto(g_1f_1,g_2f_2), \end{aligned}$$

而恒等态射 $\mathrm{id}_{(X_1,X_2)}$ 取为 $(\mathrm{id}_{X_1},\mathrm{id}_{X_2})$. 换言之, 一切都是逐分量地定义的, 所需性质也按此验证.

推而广之, 还能定义 $\mathcal{C}_1\times\mathcal{C}_2\times\mathcal{C}_3$, 依此类推.

B.3 自然变换

除了范畴以及范畴之间的函子, 还有必要考量函子之间的态射, 又称为函子之间的自然变换.

定义 B.3.1 (自然变换或函子间的态射) 考虑从 $\mathcal{C}'$ 到 $\mathcal{C}$ 的一对函子 F,G. 从 F 到 G 的自然变换 (或称从 F 到 G 的态射) 意谓一族态射 $\theta=(\theta_X)_X$, 其中

$$\theta_X\in\mathrm{Hom}_{\mathcal{C}}(FX,GX),\quad X\in\mathrm{Ob}(\mathcal{C}'),$$

要求下图对 $\mathcal{C}'$ 中的所有态射 $f:X\to Y$ 皆交换:

$$\begin{array}{ccc} FX & \xrightarrow{\theta_X} & GX \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle Gf} \\ FY & \xrightarrow[\theta_Y]{} & GY. \end{array}$$

上述自然变换写作 $\theta:F\to G$ 或 $\theta:F\Rightarrow G$ (以区别于对象之间的态射); 另一种常见的图解方式为

$$\begin{array}{ccc} & \overset{F}{\curvearrowright} & \\ \mathcal{C}' & \Downarrow\theta & \mathcal{C}. \\ & \underset{G}{\curvearrowright} & \end{array}$$

显然 $\theta:F\to G$ 和 $\psi:G\to H$ 可以合成为 $\psi\theta:F\to H$, 使得 $(\psi\theta)_X=\psi_X\theta_X$. 合成运算满足结合律.

进一步, 函子之间也有同构的概念. 记函子 F 到自身的恒等变换为 $\mathrm{id}_F=(\mathrm{id}_{FX})_X$.

定义 B.3.2 如果 θ_X 对所有 X 皆为同构, 则称 θ 是同构, 其逆 $\theta^{-1}:G\to F$ 按照 $(\theta^{-1})_X=(\theta_X)^{-1}$ 来定义. 它由 $\theta^{-1}\theta=\mathrm{id}_F$ 和 $\theta\theta^{-1}=\mathrm{id}_G$ 所刻画.

练习 B.3.3 说明自然变换 $\theta:F_1\to F_2$ 诱导 $\theta G=(\theta_{GY})_Y:F_1G\to F_2G$ 和 $G'\theta=(G'(\theta_X))_X:G'F_1\to G'F_2$, 前提是函子的合成 F_iG 和 $G'F_i$ 有意义, 此外还有结合律 $(G'\theta)G=G'(\theta G)$. 说明如果 θ 是同构, 则 θG 和 $G'\theta$ 亦然.

代数学中不时提及 "典范" 或 "自然" 的同态. 尽管我们在实践中已经获得感性认知, 自然变换的概念能赋之以更精确的意义.

约定 B.3.4 我们将自然变换 $\theta:F\to G$ 称为从函子 F 到 G 的态射. 实用中经常会省略理论框架, 只说态射 $\theta_X:FX\to GX$ 对于变元 X 是**自然**的, **典范**的, 或称其具备**函子性**.

例 B.3.5 (双重对偶与求值) 选定域 F. 考虑例 B.2.7 的函子 $D:\mathsf{Vect}(F)^{\mathrm{op}}\to\mathsf{Vect}(F)$ 和 $D^{\mathrm{op}}:\mathsf{Vect}(F)\to\mathsf{Vect}(F)^{\mathrm{op}}$. 今将给出自然变换

$$\mathrm{ev}:\mathrm{id}_{\mathsf{Vect}(F)}\to DD^{\mathrm{op}};$$

对每个 F-向量空间 V, 对应的态射定为

$$\begin{aligned} \mathrm{ev}_V : V &\to (V^\vee)^\vee \\ v &\mapsto [\mathrm{ev}_v : \underbrace{\lambda}_{\in V^\vee} \mapsto \underbrace{\lambda(v)}_{\in F}]. \end{aligned}$$

每个 ev_V 都是良定义的线性单射, 而自然变换所需的交换图表即

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \mathrm{ev}_V}\downarrow & & \downarrow{\scriptstyle \mathrm{ev}_W} \\ V^{\vee\vee} & \xrightarrow[{}^{\mathrm{t}}(^{\mathrm{t}}T)]{} & W^{\vee\vee} \end{array} \qquad T \in \mathrm{Hom}(V, W).$$

这些性质已在关于双重对偶的章节中处理过, 对于熟悉向量空间理论的读者, 展开定义直接验证也毫无困难. 最后, 完全相同的写法给出 $\mathrm{ev} : \mathrm{id}_{\mathsf{Vect}(F)^{\mathrm{op}}} \to D^{\mathrm{op}} D$.

例 B.3.6 按照下述方式定义幂集函子 $P : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Set}$.

- ⋆ 在对象层次, 它映集合 S 为其幂集 $P(S)$;
- ⋆ 在态射层次, 它映 $f : S \to T$ 为反向的映射 $P(f) : P(T) \to P(S)$ 如下:

$$P(f)(T_0) := f^{-1}(T_0) \subset S, \quad T_0 \subset T : \text{ 任意子集.}$$

函子所需的性质不难照定义检验. 现在考虑定义 B.2.9 的 Hom 函子

$$\mathrm{Hom}(\cdot, \{0,1\}) : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Set},$$

定义一对自然变换 $\mathrm{Hom}(\cdot, \{0,1\}) \underset{\psi}{\overset{\theta}{\rightleftarrows}} P$ 如下: 对所有集合 S,

$$\begin{array}{ccc} \mathrm{Hom}(S, \{0,1\}) & \underset{\psi_S}{\overset{\theta_S}{\rightleftarrows}} & P(S) \\ e & \longmapsto & e^{-1}(1) \\ \mathbf{1}_{S_0} & \longmapsfrom & S_0 \end{array}$$

其中 $\mathbf{1}_{S_0}$ 在 S_0 上取 1, 其外取 0. 作为例行练习, 请读者验证它们确实给出互逆的自然变换. 这就说明函子 $\mathrm{Hom}(\cdot, \{0,1\})$ 与 P 相互同构.

范畴论的建立始于 S. Eilenberg 与 S. MacLane 的工作. 从历史的观点看, 澄清自然变换的意涵是他们的初始动机, 由此启发了函子以及范畴的定义, 这与概念铺陈的次序正好相反. 这些动机关乎拓扑学的一些基本问题, 并非纯粹的玄想.

关乎函子与自然变换的另一个重要概念是范畴的等价, 在应用中同样重要, 这是 §B.4 的主题.

练习 B.3.7 给定范畴 $\mathcal{C}$ 的对象 S 和函子 $F:\mathcal{C}\to\mathsf{Set}$. 考虑函子 $k_{\mathcal{C}}(S):=\mathrm{Hom}_{\mathcal{C}}(S,\cdot):\mathcal{C}\to\mathsf{Set}$. 说明有以下的双射

$$\begin{aligned}\{\text{自然变换 } k_{\mathcal{C}}(S)\to F\} &\xrightarrow{1:1} F(S)\\ \theta &\mapsto \theta_S(\mathrm{id}_S).\end{aligned}$$

这一事实包含于范畴论中的**米田引理** [10, 定理 2.5.1], 但已足以体现米田引理及其证明的实质.

提示 为了构造逆映射, 对给定的 $u\in F(S)$ 和 $\mathcal{C}$ 的任意对象 T, 命 $\theta_T:\mathrm{Hom}_{\mathcal{C}}(S,T)\to F(T)$ 为映射 $f\mapsto (Ff)(u)$. 从定义说明这给出满足 $u=\theta_S(\mathrm{id}_S)$ 的自然变换 $\theta=(\theta_T)_T$, 而且是唯一的取法.

B.4 范畴等价

等价的概念旨在说明两个范畴如何能通过函子来等同. 等价的范畴应当具有同样的范畴论性质.

定义 B.4.1 (等价) 考虑一对函子 $\mathcal{C}_1 \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{C}_2$, 如果存在函子之间的同构 (定义 B.3.2)

$$\theta: FG\xrightarrow{\sim}\mathrm{id}_{\mathcal{C}_2},\quad \psi: GF\xrightarrow{\sim}\mathrm{id}_{\mathcal{C}_1},$$

则称 G 是 F 的**拟逆函子**, 并且称 F 是从 $\mathcal{C}_1$ 到 $\mathcal{C}_2$ 的**等价**.

恒等函子显然是等价, 这是最平凡的例子.

练习 B.4.2 验证以下简单性质:

(i) 等价 F_1 和 F_2 的复合 F_1F_2 仍是等价, 而且若 G_i 是 F_i 的逆拟, 则 F_1F_2 的拟逆可取为 G_2G_1;

(ii) 设 F 是等价而 F' 同构于 F, 则 F' 也是等价.

读者可能想将范畴的等价与代数结构的同构作比较. 两种概念有所差异, 因为范畴等价仅要求一对函子精确到同构互逆, 而不是严格的互逆; 由于范畴论真正关心的是精确到同构的性质, 根据后见之明, 这点并不让人意外.

注记 B.4.3 若进一步在定义 B.4.1 中要求严格等式 $FG=\mathrm{id}_{\mathcal{C}_2}$ 和 $GF=\mathrm{id}_{\mathcal{C}_1}$, 则称 F 是范畴间的同构, 而 G 是 F 的逆. 实践表明范畴的同构不如等价实用; 将严格等式放宽为对象之间的同构是范畴论的基本思路之一.

实践中经常以下述结论来识别范畴的等价.

定理 B.4.4 对于函子 $F:\mathcal{C}_1\to\mathcal{C}_2$, 以下陈述等价:

(i) F 是范畴等价, 换言之它有某个拟逆函子 $G:\mathcal{C}_2\to\mathcal{C}_1$;

(ii) F 是全忠实本质满函子 (定义 B.2.2).

证明 先解释 (i) $\Longrightarrow$ (ii). 设有定义 B.4.1 中的 F, G, θ 和 ψ. 给定 $X,Y\in\mathrm{Ob}(\mathcal{C}_1)$, 兹断言映射列

$$\begin{array}{ccccccc}\mathrm{Hom}(X,Y) & \xrightarrow{F} & \mathrm{Hom}(FX,FY) & \xrightarrow{G} & \mathrm{Hom}(GF(X),GF(Y)) & \xrightarrow{1:1} & \mathrm{Hom}(X,Y)\\ & & \cup & & & & \cup\\ & & h & \longmapsto & & & \psi_Y h\psi_X^{-1}\end{array}$$

的合成是恒等. 这相当于 $\psi_Y GF(f)=f\psi_X$ 对所有 $f\in\mathrm{Hom}(X,Y)$ 恒成立, 也相当于说图表

$$\begin{array}{ccc}GF(X) & \xrightarrow{\psi_X} & X\\ {\scriptstyle GF(f)}\downarrow & & \downarrow{\scriptstyle f}\\ GF(Y) & \xrightarrow[\psi_Y]{} & Y\end{array}$$

恒交换, 因此断言归结为自然变换 ψ 的自然性.

由此推得 F (或 G) 在态射集上给出的映射有左 (或右) 逆. 交换 F 和 G 的角色可知另一个方向也有逆, 故为双射. 综上, F 是全忠实函子.

其次, 对所有 $Z\in\mathrm{Ob}(\mathcal{C}_2)$ 皆有同构 $\theta_Z:F(G(Z))\xrightarrow{\sim}Z$, 由此知 F 本质满.

对于 (ii) $\Longrightarrow$ (i), 要点是先以本质满条件对 $\mathcal{C}_2$ 的每个对象 Z 选取[4]对象 $G(Z)\in\mathrm{Ob}(\mathcal{C}_1)$ 连同同构 $\theta_Z:F(G(Z))\xrightarrow{\sim}Z$. 以下步骤足以将 $Z\mapsto G(Z)$ 升级为拟逆函子.

1. 对所有态射 $f:Z\to Z'$, 下图中存在唯一的虚线箭头使图表交换:

$$\begin{array}{ccc}F(G(Z)) & \xrightarrow[\sim]{\theta_Z} & Z\\ \vdots & & \downarrow{\scriptstyle f}\\ F(G(Z')) & \xrightarrow[\theta_{Z'}]{\sim} & Z'.\end{array}$$

2. 因为 F 全忠实, 上图的虚线箭头来自唯一的态射 $G(f):G(Z)\to G(Z')$. 进一步证明形如 $G(ff')=G(f)G(f')$ 和 $G(\mathrm{id}_Z)=\mathrm{id}_{G(Z)}$ 的恒等式 (提示: 比较两边对 F 的像), 便将 G 升级为函子, 并且使 $(\theta_Z)_Z$ 给出同构 $\theta:FG\xrightarrow{\sim}\mathrm{id}_{\mathcal{C}_2}$.

3. 为了得到反向的 $\psi:GF\xrightarrow{\sim}\mathrm{id}_{\mathcal{C}_1}$, 对所有 $X\in\mathrm{Ob}(\mathcal{C}_1)$ 考虑 $\theta_{FX}:FGF(X)\xrightarrow{\sim}F(X)$; 由 F 全忠实可推得

[4]定义 $Z\mapsto G(Z)$ 需要一种适用于真类的选择公理. 严格的方案是或者在 (ii) 的陈述中施加更多条件, 或者如 [10] 一般采用 Grothendieck 宇宙或类似的操作, 抑或改用其他集合论 (如 NBG + 整体选择公理). 这个问题的严重程度见仁见智.

- ⋆ 存在唯一的 $\psi_X : GF(X) \to X$ 使得 $\theta_{FX} = F\psi_X$,
- ⋆ 此外, ψ_X 还是同构.

为了验证 $(\psi_X)_X$ 给出自然变换 ψ, 考虑态射 $h : X \to X'$. 由于 F 忠实, 说明下图交换即可

$$\begin{array}{ccc} FGF(X) & \xrightarrow{F\psi_X} & F(X) \\ {\scriptstyle FGFh}\downarrow & & \downarrow{\scriptstyle Fh} \\ FGF(X') & \xrightarrow[F\psi_{X'}]{} & F(X'). \end{array}$$

然而这不过是 $\theta : FG \to \mathrm{id}$ 的自然性的应用.

详细推导留给读者, 此外也可以参考 [10, 定理 2.2.13] 的进路. □

练习 B.4.5 设 $F : \mathcal{C}_1 \to \mathcal{C}_2$ 是范畴等价, 说明 F 将两个范畴的对象的同构类一一对应.

推论 B.4.6 *若函子 $F : \mathcal{C}_1 \to \mathcal{C}_2$ 是全忠实的, 则存在 $\mathcal{C}_2$ 的全子范畴 $\mathcal{C}_2'$ 连同范畴的等价 $F' : \mathcal{C}_1 \to \mathcal{C}_2'$, 使得 F 同构于 $\iota F'$, 其中 $\iota : \mathcal{C}_2' \to \mathcal{C}_2$ 是包含函子.*

证明 定义 $\mathcal{C}_2'$ 为以所有 FX 为对象的全子范畴 ($X \in \mathrm{Ob}(\mathcal{C}_1)$), 则 F 分解为 $\mathcal{C}_1 \xrightarrow{F'} \mathcal{C}_2' \xrightarrow{\iota} \mathcal{C}_2$. 按照 $\mathcal{C}_2'$ 的定义, F' 自动是本质满而且全忠实的, 对其运用定理 B.4.4. □

有鉴于此, 全忠实函子可视为范畴的 "嵌入".

例 B.4.7 接续例 B.3.5 的讨论. 有限维 F-向量空间的对偶仍是有限维, 故对偶空间函子限制为 $D : \mathsf{Vect}_{\mathrm{f}}(F)^{\mathrm{op}} \to \mathsf{Vect}_{\mathrm{f}}(F)$; 相应地也有 $D^{\mathrm{op}} : \mathsf{Vect}_{\mathrm{f}}(F) \to \mathsf{Vect}_{\mathrm{f}}(F)^{\mathrm{op}}$. 例 B.3.5 的自然变换 ev 因而限制为

$$\mathrm{ev} : \mathrm{id}_{\mathsf{Vect}_{\mathrm{f}}(F)} \to DD^{\mathrm{op}}.$$

在关于向量空间的理论中, 我们已证明当 V 有限维时 $\mathrm{ev}_V : V \to (V^\vee)^\vee$ 是同构[5)], 因此自然变换 ev 是同构. 同理, $\mathrm{ev} : \mathrm{id}_{\mathsf{Vect}_{\mathrm{f}}(F)^{\mathrm{op}}} \to D^{\mathrm{op}}D$ 也是同构.

有鉴于此, 范畴 $\mathsf{Vect}_{\mathrm{f}}(F)$ 和 $\mathsf{Vect}_{\mathrm{f}}(F)^{\mathrm{op}}$ 相互等价, 以 D 和 D^{op} 为一对拟逆函子.

例 B.4.8 继续选定域 F. 定义范畴 $\mathsf{Mat}(F)$ 如下:

▷ **对象** 非负整数 $0, 1, 2 \ldots$;

▷ **态射** 从 n 到 m 的态射集取为 $\mathrm{M}_{m\times n}(F)$, 当 $n = 0$ 或 $m = 0$ 时将此理解为 $\{0\}$.

态射的合成取为矩阵乘法, 取恒等态射 id_n 为 $\mathbf{1}_{n\times n}$. 矩阵乘法的结合律与单位矩阵的性质说明这确实是范畴; 这些定义在 $n = 0$ 或 $m = 0$ 的极端情形依然适用.

[5)] 回忆证明要领: ev_V 总是单射, 而取对偶基可见 $\dim V = \dim V^\vee$.

现在定义函子 $\Phi : \mathsf{Mat}(F) \to \mathsf{Vect}_{\mathrm{f}}(F)$, 使得 Φ 映对象 n 为 F^n, 映态射 $\boldsymbol{A} \in \mathrm{M}_{m\times n}(F)$ 为它在标准基之下对应的线性映射 $F^n \to F^m$.

向量空间的常识说明 $\mathrm{M}_{m\times n}(F) \to \mathrm{Hom}(F^n, F^m)$ 是双射, 使矩阵乘法对应到线性映射的合成, 恒等矩阵对应到恒等映射, 因此 Φ 是全忠实函子. 有限维向量空间都有基, 因而同构于某个 F^n, 因此 Φ 本质满. 代入定理 B.4.4 遂知 Φ 是等价.

这便从范畴论视角解释了为何在处理关于有限维 F-向量空间的许多问题时, 可以不失一般性地假定空间为 F^n, 然后以矩阵来操作.

注意到 Φ 的拟逆没有标准的取法.

例 B.4.9 对于任何环 R, 取其相反环 R^{op}, 则有 R-Mod 和 Mod-R^{op} 之间的等价; 事实上, 两者还是同构的 (注记 B.4.3). 这是练习 12.1.2 的内容.

交换群范畴 Ab 等价于 $\mathbb{Z}$-Mod, 它实际也是同构. 所需构造已包含于例 12.1.7 的讨论中.

B.5 泛性质

代数学中的许多构造是由泛性质刻画的, 给出的产物是精确到唯一同构的对象, 这一技术能在范畴论中予以合理的说明. 我们先介绍何谓范畴中的始对象和终对象, 然后以此诠释泛性质, 再解释实例.

定义 B.5.1 *考虑范畴 $\mathcal{C}$ 及其对象 S. 若对于所有对象 X, 存在唯一的态射 $S \to X$ (或 $X \to S$), 则称 S 是 $\mathcal{C}$ 的***始对象*** (或***终对象***); 兼为始终对象的 S 称为***零对象***.*

始对象或终对象未必存在, 但它们若存在则具有如下的唯一性.

命题 B.5.2 *设 S, S' 同为 $\mathcal{C}$ 的始对象 (或终对象), 则存在唯一的同构 $S \xrightarrow{\sim} S'$.*

证明 考虑始对象的版本. 根据定义, 存在唯一的态射 $f : S \to S'$ 和 $g : S' \to S$; 特别地, 所求同构若存在则必为 f. 取合成得到 $gf : S \to S$ 和 $fg : S' \to S'$; 再次用始对象定义中的唯一性部分, 可得 $gf = \mathrm{id}_S$ 和 $fg = \mathrm{id}_{S'}$, 因此它们是互逆的同构.

对于终对象, 可重复相同论证, 或留意到 $\mathcal{C}$ 的终对象无非是 $\mathcal{C}^{\mathrm{op}}$ 的始对象. □

练习 B.5.3 验证以下事实. 空集是 Set 的始对象, 独点集是 Set 的终对象; 环 $\mathbb{Z}$ 是 Ring 的始对象, 零环是 Ring 的终对象; 平凡群是 Grp 的零对象, 零空间是 $\mathsf{Vect}(F)$ 的零对象, 零模是 R-Mod (或 Mod-R) 的零对象.

练习 B.5.4 假定读者了解范畴等价的概念. 设 $F : \mathcal{C}_1 \to \mathcal{C}_2$ 是范畴的等价, 证明 S 是 $\mathcal{C}_1$ 的始对象 (或终对象) 当且仅当 FS 是 $\mathcal{C}_2$ 的始对象 (或终对象). 这是等价保持范畴论性质的一则简单体现.

如果一种数学构造能在适当的范畴中刻画为始对象或终对象, 则称此刻画为该构造的***泛性质***. 以下是一些例子.

例 B.5.5 (积) 给定范畴 $\mathcal{C}$ 的一族对象 $(X_i)_{i\in I}$, 其中 I 是任意集合, 定义范畴 $\mathcal{D}$ 如下:

▷ **对象** 资料 $(S,(f_i)_{i\in I})$, 其中 $S\in \mathrm{Ob}(\mathcal{C})$ 而 $f_i\in \mathrm{Hom}_{\mathcal{C}}(S,X_i)$;

▷ **态射** 从 $(S,(f_i)_{i\in I})$ 到 $(S',(f'_i)_{i\in I})$ 的态射取为 $\varphi\in \mathrm{Hom}_{\mathcal{C}}(S,S')$, 要求 $f'_i\varphi=f_i$ 对所有 i 成立. 态射的合成和恒等态射与 $\mathcal{C}$ 中相同.

如果范畴 $\mathcal{D}$ 有终对象, 则称之为 $(X_i)_{i\in I}$ 的积, 相应的资料记为 $\left(\prod_{i\in I}X_i,(p_i)_{i\in I}\right)$. 终对象的性质在 $\mathcal{C}$ 中表述为: 对所有对象 S 连同一族态射 $f_i:S\to X_i$, 存在唯一的 $\varphi:S\to\prod_{i\in I}X_i$ 使下图对所有 $i\in I$ 交换

$$\begin{array}{ccc} S & \xrightarrow{\varphi} & \prod_{j\in I}X_j \\ & \searrow^{f_i} & \downarrow{\scriptstyle p_i} \\ & & X_i. \end{array}$$

命题 B.5.2 说明积 $(\prod_i X_i,(p_i)_i)$ 若存在则唯一, 精确到 $\mathcal{D}$ 中的唯一同构.

积在 $I=\varnothing$ 时也有意义, 此时泛性质中的 $f_i:S\to X_i$ 和对应的条件不复存在, 故空积等于 $\mathcal{C}$ 的终对象.

在范畴 **Set** 中, 积 $\prod_i X_i$ 存在并由集合的 Cartesius 积给出, 投影映射给出 p_i, 空积是独点集. 在范畴 **Ring**, **Vect**(F), **Grp**, R-**Mod** 或 **Mod**-R 中 (但不限于此), 积由代数结构的直积给出, 而 p_i 仍是相应的投影同态; 这些代数结构的直积仍是实现在 Cartesius 积上的. 相关验证没有本质的困难.

注意到 **Ring** 中的空积对应到零环. 若在定义中排除零环, **Ring** 便只有非空积; 举例明之, $\mathbb{Z}/2\mathbb{Z}$ 和 $\mathbb{Z}/3\mathbb{Z}$ 无法用同态映向同一个非零环.

例 B.5.6 (余积) 仍给定 $\mathcal{C}$ 的一族对象 $(X_i)_{i\in I}$, 定义范畴 $\mathcal{E}$ 如下:

▷ **对象** 资料 $(T,(g_i)_{i\in I})$, 其中 $T\in \mathrm{Ob}(\mathcal{C})$ 而 $g_i\in \mathrm{Hom}_{\mathcal{C}}(X_i,T)$;

▷ **态射** 从 $(T,(g_i)_{i\in I})$ 到 $(T',(g'_i)_{i\in I})$ 的态射取为 $\psi\in \mathrm{Hom}_{\mathcal{C}}(T,T')$, 要求 $\psi g_i=g'_i$ 对所有 i 成立. 态射的合成和恒等态射与 $\mathcal{C}$ 中相同.

如果范畴 $\mathcal{E}$ 有始对象, 则称之为 $(X_i)_{i\in I}$ 的余积, 记为 $\left(\coprod_{i\in I}X_i,(\iota_i)_{i\in I}\right)$. 这相当于要求对所有对象 T 连同一族态射 $g_i:X_i\to T$, 存在唯一的 $\psi:\coprod_{i\in I}X_i\to T$ 使下图对所有 i 交换

$$\begin{array}{ccc} T & \xleftarrow{\psi} & \coprod_{j\in I}X_j \\ & \nwarrow^{g_i} & \uparrow{\scriptstyle \iota_i} \\ & & X_i. \end{array}$$

读者由此当意识到余积是积的对偶版本[6], 相当于反转定义中的箭头; 特别地, $\mathcal{C}$ 中的余积相当于 $\mathcal{C}^{\mathrm{op}}$ 的积. 空余积相当于 $\mathcal{C}$ 的始对象.

[6] 在数学中, 前缀 "余" 字指涉这种对偶性.

在范畴 Set 中, 余积由集合的无交并给出, 包含映射给出 ι_i, 空余积是空集. 在 Vect(F), R-Mod 或 Mod-R 中, 余积由向量空间或模的直和给出, 直和项的包含映射给出 ι_i. 对于 Ring (容许零环) 和 Grp, 余积存在但相对复杂, 此处不多谈.

对于 Vect(F) 或更广的 R-Mod 或 Mod-R, 我们知道当 I 有限时直和与直积是一回事, 此时 $(X_i)_{i\in I}$ 的余积和积建立在同一个对象 $\bigoplus_i X_i$ 上. 这一现象并非偶然, 请参阅 [10, 定理 3.4.9].

当 I 有限, 范畴中的积 (或余积) 若存在则经常写作 $X_1\times\cdots\times X_n$ (或 $X_1\sqcup\cdots\sqcup X_n$) 之形. 严格来说, 它们只有和态射族 $(p_i)_i$ (或 $(\iota_i)_i$) 一道考虑才有意义.

练习 B.5.7 设两族对象 $(X_i)_{i\in I}$ 和 $(Y_i)_{i\in I}$ 的积皆存在. 给定态射族 $f_i : X_i \to Y_i$, 说明存在唯一的态射 $\prod_i f_i : \prod_i X_i \to \prod_i Y_i$ 使得下图对所有 $i \in I$ 皆交换:

$$\begin{array}{ccc} \prod_{j\in I} X_j & \xrightarrow{p_i^X} & X_i \\ {\scriptstyle \prod_i f_i}\downarrow & & \downarrow{\scriptstyle f_i} \\ \prod_{j\in I} Y_j & \xrightarrow[p_i^Y]{} & Y_i; \end{array}$$

由此进一步说明有形如 $(\prod_i g_i)(\prod_i f_i) = \prod_i (g_i f_i)$ 和 $\prod_i \mathrm{id} = \mathrm{id}$ 的性质. 这是积的函子性, 余积的情形可类推.

练习 B.5.8 选定范畴 $\mathcal{C}$. 试从泛性质说明:

(i) 积具有称为结合约束 (或交换约束) 的典范同构 $X\times(Y\times Z)\simeq X\times Y\times Z\simeq (X\times Y)\times Z$ (或 $X\times Y\simeq Y\times X$), 前提是这些积在 $\mathcal{C}$ 中存在.

提示 一般情形见诸 [10, 引理 2.7.11, 2.7.12].

(ii) 设 $\star$ 为 $\mathcal{C}$ 的终对象, 则对所有对象 X 皆存在积 $X\times\star$ 和 $\star\times X$, 对应的对象是 X 自身, 使得

$$\star\times X \xrightarrow[\sim]{p_2} X \xleftarrow[\sim]{p_1} X\times\star.$$

(iii) 设积 $X\times Y$ 存在, 则有典范同构 $c(X,Y) : X\times Y \xrightarrow{\sim} Y\times X$, 其刻画是使下图交换:

$$\begin{array}{ccccc} & & X\times Y & & \\ & \overset{p_1}{\swarrow} & \downarrow{\scriptstyle c(X,Y)} & \overset{p_2}{\searrow} & \\ X & \underset{p_2'}{\nwarrow} & Y\times X & \underset{p_1'}{\nearrow} & Y. \end{array}$$

对于以上两道练习, 考虑 $\mathcal{C}^{\mathrm{op}}$ 便有余积的版本.

例 B.5.9 本书早在从整数构造有理数, 以及从多项式构造有理函数时就已经用到分式域的构造. 按照范畴论的观点看, 对给定的整环 R 考虑范畴 $\mathcal{F}$ 如下:

▷ **对象** 资料 (A,φ), 其中 A 是交换环, $\varphi: R\to A$ 是环同态, 使得 $\varphi(R\smallsetminus\{0\})\subset A^\times$;

▷ **态射** 从 (A,φ) 到 (B,ψ) 的态射取为环同态 $f: A\to B$, 要求 $\psi = f\varphi$. 态射的合成与恒等态射都在 Ring 中取.

范畴 $\mathcal{F}$ 有始对象, 它正是 R 的分式域 $\mathrm{Frac}(R)$ 连同自然嵌入 $R\hookrightarrow \mathrm{Frac}(R)$. 所需的泛性质已在环论的章节介绍过. 此构造还有称为局部化的推广, 详见 [10, 命题 5.3.9].

读者也许更偏好从集合观点理解先前的一系列例子, 然而以下讨论的张量积却绕不开泛性质.

例 B.5.10 (张量积) 选定域 F 和 F-向量空间 V, W. 定义范畴 $\mathcal{B}_{V,W}$ 如下:

▷ **对象** 资料 (L,B), 其中 L 是 F-向量空间而 $B: V\times W\to L$ 是双线性映射;

▷ **态射** 从 (L,B) 到 (L',B') 的态射取为使下图交换的线性映射 $\varphi: L\to L'$

$$\begin{array}{ccc} V\times W & \xrightarrow{\ B\ } & L \\ & \searrow^{B'} & \downarrow{\scriptstyle \varphi} \\ & & L'. \end{array}$$

如果范畴 $\mathcal{B}_{V,W}$ 有始对象 $(L_{\mathrm{univ}}, B_{\mathrm{univ}})$, 则称之为 V 和 W 的张量积. 此资料是唯一的, 精确到 $\mathcal{B}_{V,W}$ 中的唯一同构, 也写作 $(V\otimes W, (v,w)\mapsto v\otimes w)$ 之形.

以上无非是以范畴语言改述张量积的定义 15.1.2, 而命题 15.1.1 证明了 $\mathcal{B}_{V,W}$ 确实有始对象. 多元张量积 (定义–命题 15.1.5) 的范畴论表述是完全类似的.

命题 15.1.1 对张量积的具体构造涉及两步, 一是构造以给定集合为基的向量空间, 二是对子空间取商, 两者都有基于泛性质的刻画. 我们将在 §B.6 讨论商空间, 以下先讨论第一步构造.

例 B.5.11 给定域 F 和集合 X. 定义范畴 $\mathcal{C}_X(F)$ 如下:

▷ **对象** 资料 (V,i), 其中 V 是 F-向量空间, i 是集合之间的映射 $X\to V$.

▷ **态射** 从 (V,i) 到 (V',i') 的态射是满足 $i' = \alpha i$ 的线性映射 $\alpha: V\to V'$.

态射的合成与恒等态射和 $\mathsf{Vect}(F)$ 相同. 以下说明范畴 $\mathcal{C}_X(F)$ 有始对象, 它由 X 份 F 的直和 $F^{\oplus X}$ (视同以 X 为基的 F-向量空间), 连同自然嵌入 $\iota: X\hookrightarrow F^{\oplus X}$ 给出. 验证始对象的条件相当于对所有 F-向量空间 V 验证双射

$$\begin{aligned} \mathrm{Hom}\left(F^{\oplus X}, V\right) &\xrightarrow{1:1} \{\text{集合的映射 } i: X\to V\} \\ \alpha &\longmapsto \alpha\iota; \end{aligned}$$

这当然是成立的: 指定 α 相当于指定它在基 X 上的值, 也相当于指定集合的映射 $i: X\to V$. 此泛性质刻画 $F^{\oplus X}$, 精确到 $\mathcal{C}_X(F)$ 中的唯一同构.

请读者快速说明 $X \mapsto F^{\oplus X}$ 升级为函子 $\mathsf{Set} \to \mathsf{Vect}(F)$, 暂记为 L. 记忘却函子 $\mathsf{Vect}(F) \to \mathsf{Set}$ 为 U. 先前的双射遂改写为

$$\mathrm{Hom}_{\mathsf{Vect}(F)}(L(X), V) \overset{1:1}{\longleftrightarrow} \mathrm{Hom}_{\mathsf{Set}}(X, U(V)),$$

而且当 (X, V) 变动, 两边分别给出函子 $\mathsf{Set}^{\mathrm{op}} \times \mathsf{Vect}(F) \to \mathsf{Set}$. 从构造不难检验这族双射是定义 B.3.2 所谓的自然同构.

若将 F-向量空间换成一般的 R-模, 仍有类似结论.

推而广之, 若有一对函子 $F : \mathcal{C} \rightleftarrows \mathcal{D} : G$ 连同自然双射

$$\mathrm{Hom}_{\mathcal{D}}(F(\cdot), \cdot) \xrightarrow{\sim} \mathrm{Hom}_{\mathcal{C}}(\cdot, G(\cdot)),$$

则称此资料为**伴随对**, 称 F 为 G 的左伴随, G 为 F 的右伴随. 伴随性是范畴论的基本概念之一, 有心深入的读者应当参阅相关教材或 [10, §2.6].

为了说明词源, 留意到若将伴随对定义中的范畴代换为有限维向量空间, 将函子 F 和 G 代换为线性映射, 将 Hom 函子代换为非退化双线性形式 (譬如内积), 将自然双射代换为等号, 则函子的左/右伴随可与线性映射的左/右伴随映射相比较, 两者的刻画具有相似的样貌.

例 B.5.12 回顾例 B.2.8 的群化函子 $(\cdot)^{\mathrm{grp}} : \mathsf{CMon} \to \mathsf{Ab}$. 对于所有交换幺半群 M 与交换群 A, 二元运算皆表作加法, 请读者验证以下的互逆映射

$$\begin{aligned} \mathrm{Hom}_{\mathsf{Ab}}(M^{\mathrm{grp}}, A) &\overset{1:1}{\rightleftarrows} \mathrm{Hom}_{\mathsf{CMon}}(M, A) \\ \overline{f} &\longmapsto \overline{f}\eta_M \\ [[a, b]] \mapsto f(a) - f(b)] &\longmapsfrom f \end{aligned}$$

其中 $[\![a, b]\!] \in M^{\mathrm{grp}}$. 记 $\iota : \mathsf{Ab} \to \mathsf{CMon}$ 为包含函子, 上述双射及其函子性表明

$$(\cdot)^{\mathrm{grp}} : \mathsf{CMon} \rightleftarrows \mathsf{Ab} : \iota$$

是伴随对. 另一方面, 双射也可以理解为刻画 M^{grp} 连同 $\eta_M : M \to M^{\mathrm{grp}}$ 的泛性质: 在全体从 M 映向某个交换群的幺半群同态所成范畴中, η_M 是始对象.

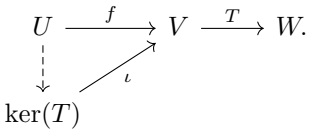
练习 B.5.13 设 E 为域 F 的扩域. 验证 §15.4 介绍的域的变换给出函子 $E \underset{F}{\otimes} (\cdot) : \mathsf{Vect}(F) \to \mathsf{Vect}(E)$, 而且它是纯量限制函子 $\mathsf{Vect}(E) \to \mathsf{Vect}(F)$ 的左伴随. 具体写下所需双射.

提示 参考命题 15.4.5.

B.6 等化子及余等化子

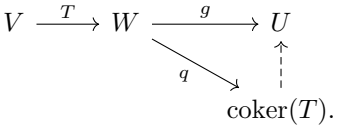
本节是 §B.5 的延续. 我们将从向量空间理论中熟知的核及余核入手, 然后触及更广泛的范畴论构造, 它们都由泛性质刻画.

例 B.6.1 (核, 余核及相关概念) 选定域 F. 向量空间之间的线性映射 $T : V \to W$ 的核 $\ker(T)$ 连同包含映射 $\iota : \ker(T) \to V$ 由以下泛性质刻画: 对所有向量空间 U 连同满足 $Tf = 0$ 的线性映射 $f : U \to V$, 存在唯一的虚线所示线性映射[7]使得下图交换:



不难看出 $(\ker(T), \iota)$ 是所有 (U, f) 所成范畴 (态射定义自明) 的终对象.

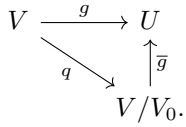
类似地, 余核 $\operatorname{coker}(T) = W/\operatorname{im}(T)$ 连同商映射 $q : W \to \operatorname{coker}(T)$ 由以下泛性质刻画: 对所有向量空间 U 连同满足 $gT = 0$ 的线性映射 $g : W \to U$, 存在唯一的虚线所示线性映射[8]使得下图交换:



换言之, $(\operatorname{coker}(T), q)$ 是所有 (U, g) 所成范畴的始对象.

线性映射 T 的像可以间接地用 $\ker[W \xrightarrow{q} \operatorname{coker}(T)]$ 或 $\operatorname{coker}[\ker(T) \xrightarrow{\iota} V]$ 来刻画, 两者同构. 因此 $\operatorname{im}(T)$ 也具有范畴论的解释.

现在考虑子空间 V_0 的包含映射 $\iota : V_0 \hookrightarrow V$. 商空间 V/V_0 等于 $\operatorname{coker}(\iota)$. 根据上述讨论, V/V_0 连同商映射 $q : V \to V/V_0$ 遂由以下的泛性质刻画: 对于所有向量空间 U 连同满足 $g|_{V_0} = 0$ 的线性映射 $g : V \to U$, 存在唯一的 $\overline{g} : V/V_0 \to U$ 使得下图交换:



以上仅对 F-向量空间讨论, 然而对于任何环 R, 关于 R-模的商与 R-模同态的核, 余核及像的泛性质诠释是完全相同的.

练习 B.6.2 按照类似的模式, 以泛性质刻画

[7] 原因是 f 的像包含于 $\ker(T)$.

[8] 虚线箭头能且仅能是 $w + \operatorname{im}(T) \mapsto g(w)$.

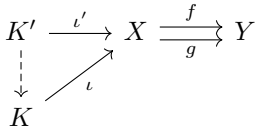
(i) 群同态的核, 群对正规子群的商群;

(ii) 环对理想的商环.

核及余核进一步由以下概念统摄.

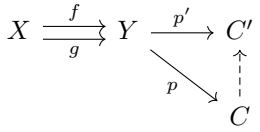
定义 B.6.3 (等化子及余等化子) 设 $f,g:X\rightrightarrows Y$ 为范畴 $\mathcal{C}$ 中的一对态射.

⋆ 设态射 $\iota:K\to X$ 满足 $f\iota=g\iota$. 如果对所有满足 $f\iota'=g\iota'$ 的态射 $\iota':K'\to X$, 存在唯一的虚线所示态射使下图交换



则称 $K\xrightarrow{\iota}X$ 为 f 和 g 的等化子, 也记为 $\ker(f,g)\to X$.

⋆ 设态射 $p:Y\to C$ 满足 $pf=pg$. 如果对所有满足 $p'f=p'g$ 的态射 $p':Y\to C'$, 存在唯一的虚线所示态射使下图交换



则称 $Y\xrightarrow{p}C$ 为 f 和 g 的余等化子, 也记为 $Y\to\operatorname{coker}(f,g)$.

容易将 $\ker(f,g)$ (或 $\operatorname{coker}(f,g)$) 解释为全体 (K',ι') (或 (C',p')) 所成范畴的终对象 (或始对象), 因此它若存在则唯一, 精确到该范畴中的唯一同构.

读者应已察觉到等化子和余等化子仅差一个箭头反转, 因此 $\mathcal{C}$ 中的余等化子无非是同一对态射在 $\mathcal{C}^{\mathrm{op}}$ 中的等化子 (如果存在). 不致混淆时, 资料中的 ι 和 p 经常省略.

例 B.6.4 在范畴 $\mathsf{Vect}(F)$ 中考虑一对态射 $T_1,T_2:V\to W$, 比较泛性质可见

$$\ker(T_1-T_2)=\ker(T_1,T_2),\quad \operatorname{coker}(T_1-T_2)=\operatorname{coker}(T_1,T_2);$$

取 $T_2=0$ 便可将核 (或余核) 视为等化子 (或余等化子) 的特例. 相应的结论在模范畴 R-Mod 中同样成立. 关键在于这些范畴的 Hom 集具有加法群结构, 使得态射合成是双加性的.

等化子, 余等化子连同 §B.5 介绍的积, 余积都能以范畴论中的极限概念来解释, 见 [10, §2.7].

练习 B.6.5 扼要地说明范畴间的等价保持等化子和余等化子.

练习 B.6.6 在范畴 Set 中考虑一对态射 $f,g : A \rightrightarrows B$. 验证 $\ker(f,g)$ 是 A 的子集 $\{a \in A : f(a) = g(a)\}$. 至于 $\operatorname{coker}(f,g)$, 定义 $\sim$ 为 B 上满足 $\forall a\ f(a) \sim g(a)$ 的最细等价关系 (请说明其意涵), 然后验证商集 $B/\sim$ 给出 $\operatorname{coker}(f,g)$.

练习 B.6.7 设 $R \subset A \times A$ 给出集合 A 上的等价关系. 命 $p_1^R, p_2^R : R \rightrightarrows A$ 为向两个分量的投影映射. 证明在范畴 Set 中的 $\operatorname{coker}\left(p_1^R, p_2^R\right)$ 由商集 A/R 给出.

B.7 幺半范畴一瞥

幺半群是具有乘法二元运算的集合, 使得乘法具有结合律和幺元. 幺半范畴是这一思路在范畴层次的实现, 乘法在这种范畴 $\mathcal{V}$ 上体现为函子 $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$ (积范畴 $\mathcal{V} \times \mathcal{V}$ 见诸定义 B.2.10), 幺元则体现为一个指定的对象. 带有这种乘法运算的范畴在应用中频繁出现. 本节仅简单勾勒基本定义, 更详细的讨论请见 [10, 第三章].

在代数学中, 幺半范畴的典型例子是 F-向量空间的张量积, F 是选定的域; 张量积给出函子 $\otimes : \mathsf{Vect}(F) \times \mathsf{Vect}(F) \to \mathsf{Vect}(F)$. 然而 $\otimes$ 没有严格的结合律, 取而代之的是结合约束 (命题 15.2.2), 体现为一族同构

$$a(X,Y,Z) : (X \otimes Y) \otimes Z \xrightarrow{\sim} X \otimes (Y \otimes Z);$$

同理, F 本身作为 F-向量空间对 $\otimes$ 扮演了类似幺元的角色, 相应的幺约束 (命题 15.2.3) 体现为两族同构

$$F \otimes X \xrightarrow[\sim]{\lambda_X} X \xleftarrow[\sim]{\rho_X} X \otimes F.$$

以下定义受此启发, 我们借用张量积的符号, 将所论的函子写作 $\otimes$.

定义 B.7.1 (幺半范畴) 一个幺半范畴意谓一组资料 $(\mathcal{V}, \otimes, a, \mathbf{1}, \lambda, \rho)$, 其中

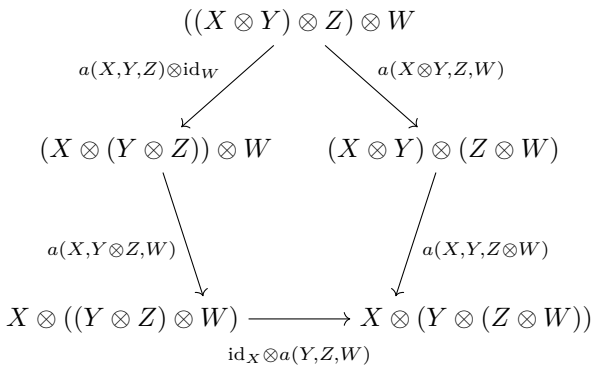
(i) $\mathcal{V}$ 是一个范畴;

(ii) $\otimes : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ 是函子, 在对象和态射上的映法分别记为 $(X,Y) \mapsto X \otimes Y$ 和 $(f,g) \mapsto f \otimes g$;

(iii) $a = (a(X,Y,Z))_{X,Y,Z \in \mathrm{Ob}(\mathcal{V})}$ 是函子的同构

$$a : ((\cdot \otimes \cdot) \otimes \cdot) \xrightarrow{\sim} (\cdot \otimes (\cdot \otimes \cdot)),$$

使得对所有对象 X,Y,Z,W, 下图交换:

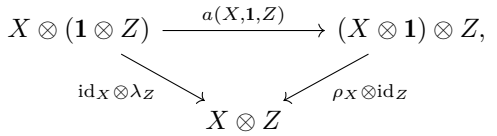


(iv) 对象 $\mathbf{1}\in \mathrm{Ob}(\mathcal{V})$ 称为幺元, λ (或 ρ) 是自然同构 $\mathbf{1}\otimes(\cdot)\xrightarrow{\sim}\mathrm{id}_{\mathcal{V}}$ (或 $(\cdot)\otimes\mathbf{1}\xrightarrow{\sim}\mathrm{id}_{\mathcal{V}}$), 具体写作

$$\lambda_X:\mathbf{1}\otimes X\xrightarrow{\sim}X,\quad \rho_X:X\otimes\mathbf{1}\xrightarrow{\sim}X,$$

其中 X 遍历 $\mathcal{V}$ 的对象, 要求

- ⋆ 下图对所有 X 和 Z 交换



- ⋆ 态射之间的等式 $\lambda_{\mathbf{1}}=\rho_{\mathbf{1}}:\mathbf{1}\otimes\mathbf{1}\xrightarrow{\sim}\mathbf{1}$.

若无混淆之虞, 我们经常将一个幺半范畴用 $\mathcal{V}$ 表示, 略去其他构件.

留意到 $\mathbf{1}\otimes(\cdot)$ 和 $(\cdot)\otimes\mathbf{1}$ 都是幺半范畴 $\mathcal{V}$ 到自身的等价, 因为它们皆同构于恒等函子.

注记 B.7.2　对照于 [10, §3.1] 关于幺半范畴的定义, 该书未将 ρ 和 λ 纳入定义, 只是

- ⋆ 要求 $\mathbf{1}\otimes(\cdot)$ 和 $(\cdot)\otimes\mathbf{1}$ 都是等价,
- ⋆ 指定了相当于此处的 $\lambda_{\mathbf{1}}=\rho_{\mathbf{1}}$ 的同构 $\iota:\mathbf{1}\otimes\mathbf{1}\xrightarrow{\sim}\mathbf{1}$.

因此该处的定义乍看较宽松; 然而 [10, 引理 3.1.5] 说明两种定义其实等价.

例 B.7.3　一如本节开头所述, 当域 F 选定, $\mathsf{Vect}(F)$ 连同张量积运算 $\otimes$ 给出幺半范畴. 在此逐一验证定义 B.7.1 的条件:

- ⋆ 条件 (ii): 定义–命题 15.1.6 及后续讨论足以说明 $\otimes$ 是函子.
- ⋆ 条件 (iii): 结合约束 a 来自命题 15.2.2, 五边形交换图表按照 $a(X,Y,Z)$ 的具体映法直接检验.

⋆ 条件 (iv): 取幺元为 F, 则幺约束 λ 和 ρ 来自命题 15.2.3, 对应的交换图表和 $\lambda_F = \rho_F$ 仍然按具体映法检验.

严格来说, 第十五章并未明确解释 a, λ, ρ 是自然变换, 然而从具体映法来看则是明了的.

如果读者了解如何对交换环 R 上的模定义张量积 $\otimes$, 上述论断依然适用.

幺半范畴的另一类例子来自范畴中的积和余积. 细节留作以下练习.

练习 B.7.4 设范畴 $\mathcal{C}$ 中任意有限多个对象皆有积, 包括空积亦即终对象 (记为 $\star$). 证明 $\mathcal{C}$ 具有幺半范畴的结构, 使得乘法函子映 $(X, Y) \in \mathrm{Ob}(\mathcal{C} \times \mathcal{C})$ 为积 $X \times Y$, 而幺元为 $\star$.

将积换成余积, 终对象换成始对象, 同样论断依然适用.

提示〉 练习 B.5.7 说明取积确实给出函子 $\mathcal{C} \times \mathcal{C} \to \mathcal{C}$. 以练习 B.5.8 的结论验证幺半范畴的公理. 过渡到 $\mathcal{C}^{\mathrm{op}}$ 便是余积版本.

如果 $\mathcal{V}$ 是幺半范畴, 而且定义 B.7.1 中的 a 和 λ, ρ 都是严格等式而不仅是同构, 则称 $\mathcal{V}$ 为严格幺半范畴; 此时定义中的 (iii) 和 (iv) 处处是等号, 故平凡地成立. 尽管严格幺半范畴的操作更加简单且直观, 张量积的例子已表明日常生活中的幺半范畴通常并非严格. 读者兴许会问: 如果一则范畴论性质适用于严格幺半范畴, 它是否也自动适用于一般的幺半范畴? 答案是肯定的. MacLane 的融贯定理 [10, 定理 3.2.2] 说明所有幺半范畴都等价于一个严格幺半范畴, 使得两者的幺半结构在等价之下相匹配. 这表明定义 B.7.1 的公理尽管精简, 却足以蕴涵我们对乘法运算的所有期望.

另一个问题是交换性. 满足乘法交换律 $ab = ba$ 的幺半群称为交换幺半群. 在幺半范畴 $\mathcal{V}$ 的层次, 交换律应当放宽为自然同构 $c(X, Y) : X \otimes Y \xrightarrow{\sim} Y \otimes X$, 使得它与 a, λ, ρ 满足一系列的兼容性条件; 这般的 $c = (c(X, Y))_{X,Y \in \mathrm{Ob}(\mathcal{C})}$ 被称为 $\mathcal{V}$ 上的辫结构, 相应地称 $(\mathcal{V}, c)$ 为辫幺半范畴, 详见 [10, §3.3].

实践中出现的辫结构有时满足 $c(Y, X)c(X, Y) = \mathrm{id}_{X \otimes Y}$, 有时则不然; 满足上述等式的辫幺半范畴称为对称幺半范畴. 例 B.7.3 和练习 B.7.4 的幺半范畴都有对称辫结构 (前者请见命题 15.2.4), 而在数学物理与拓扑学等领域中则自然地面临非对称辫结构.

凡此种种, 都是从集合 (幺半群) 上升到范畴 (幺半范畴) 所伴生的现象. 不难设想, 随着范畴论朝更高阶推进, 幺半结构的结合律和交换律将呈现出更为缤纷的谱系, 它们都能应用于实际的数学问题.

参考文献

[1] Keith Conrad. “Irreducibility of truncated exponentials”. URL: https://kconrad.math.uconn.edu/blurbs/gradnumthy/schurtheorem.pdf (引用于 p. 273).

[2] Michael Joswig and Thorsten Theobald. *Polyhedral and algebraic methods in computational geometry*. Universitext. Springer, London, 2013, pp. x+250. ISBN: 978-1-4471-4816-6; 978-1-4471-4817-3. DOI: 10.1007/978-1-4471-4817-3 (引用于 p. 546).

[3] Winfried Scharlau. *Quadratic and Hermitian forms*. Vol. 270. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1985, pp. x+421. ISBN: 3-540-13724-6. DOI: 10.1007/978-3-642-69971-9 (引用于 pp. 602, 631, 634).

[4] Saharon Shelah. “The future of set theory”. 刊于: *Set theory of the reals (Ramat Gan, 1991)*. Vol. 6. Israel Math. Conf. Proc. Bar-Ilan Univ., Ramat Gan, 1993, pp. 1–12 (引用于 p. 44).

[5] Blair K. Spearman and Kenneth S. Williams. “Characterization of solvable quintics $x^5 + ax + b$”. 刊于: *Amer. Math. Monthly* 101.10 (1994), pp. 986–992. ISSN: 0002-9890,1930-0972. DOI: 10.2307/2975165 (引用于 p. 35).

[6] Ernst Witt. “Theorie der quadratischen Formen in beliebigen Körpern”. 刊于: *J. Reine Angew. Math.* 176 (1937), pp. 31–44. ISSN: 0075-4102,1435-5345. DOI: 10.1515/crll.1937.176.31 (引用于 p. 601).

[7] 伍胜健. 数学分析 (第一册). 北京: 北京大学出版社, 2009. ISBN: 978-7-301-15685-8 (引用于 p. 63).

[8] 席南华. 基础代数 (第一卷). 北京: 科学出版社, 2016. ISBN: 978-7-03-049843-4 (引用于 p. 273).

[9] 张鸿林, 葛显良. 英汉数学词汇 (第二版). 北京: 清华大学出版社, 2010 (引用于 p. 3).

[10] 李文威. 代数学方法（第一卷）. Vol. 67.1. 现代数学基础丛书. 北京: 高等教育出版社, 2019. ISBN: 978-7-04-050725-6 (引用于 pp. 2, 71, 74, 122, 218, 232, 244, 251, 253, 255, 265, 271, 272, 274, 275, 407, 416, 434, 440, 464, 468, 480, 483, 484, 488, 491, 563, 576, 590, 592, 593, 597, 598, 620, 623, 634, 641, 648, 650, 651, 654, 657, 661, 664–666, 669–671, 673–676).

[11] 伍胜健 谭小江. 复变函数简明教程. 北京: 北京大学出版社, 2006. ISBN: 978-7-30-108530-3 (引用于 p. 17).

[12] 赵春来, 徐明曜. 抽象代数 *I*. 北京: 北京大学出版社, 2008. ISBN: 978-7-301-14168-7 (引用于 p. 456).

[13] 郝兆宽, 杨跃. 集合论：对无穷概念的探索. 逻辑与形而上学教科书系列. 上海: 复旦大学出版社, 2014. ISBN: 978-7-309-10710-4 (引用于 pp. 40, 641).

符号索引

名词索引暨英译

中文术语按汉语拼音排序.

W

X