

代数学讲义

版本: 2025-04-29

李文威 著

<https://www.wvli.asia>

网络版

编译日期: 2025-04-29

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



本作品采用知识共享署名 4.0 国际许可协议进行许可. 访问 <http://creativecommons.org/licenses/by/4.0/> 查看该许可协议.

目录

上册	1
导言	3
第一章 综观	17
1.1 何谓代数	17
1.2 各种方程的求解	24
1.3 从线性方程组到 Gauss-Jordan 消元法	29
1.4 关于线性方程组的总结	35
习题	36
第二章 集合, 映射与关系	41
2.1 集合概论	42
2.2 映射的运算	47
2.3 集合的积与无交并	52
2.4 序结构	55
2.5 等价关系与商集	58
2.6 从非负整数集到有理数集	60
2.7 算术入门	65
2.8 同余式	68
2.9 集合的基数	71
习题	74
第三章 环, 域和多项式	79
3.1 环和域	80
3.2 同态和同构	86
3.3 多项式环	89

3.4	一元多项式的带余除法与根	94
3.5	从整环的分式域到有理函数域	95
3.6	多项式函数	100
3.7	域的特征	102
	习题	104
第四章	向量空间和线性映射	109
4.1	引言: 回到线性方程组	112
4.2	向量空间	115
4.3	矩阵及其运算	118
4.4	基和维数	121
4.5	线性映射	129
4.6	从线性映射观矩阵	132
4.7	从矩阵的转置到对偶空间	140
4.8	核, 像与消元法	145
4.9	基的变换: 矩阵的共轭与相抵	149
4.10	直和分解	155
4.11	分块矩阵运算	160
4.12	商空间	166
	习题	173
第五章	行列式	181
5.1	置换概论	184
5.2	几何动机: 有向体积	189
5.3	一类交错形式的刻画	192
5.4	行列式的定义和基本性质	197
5.5	一些特殊行列式	204
5.6	分块行列式	206
5.7	Cramer 法则	208
5.8	特征多项式和 Cayley–Hamilton 定理	212
5.9	线性映射的迹	218
5.10	不变子空间	220
5.11	子式与 Cauchy–Binet 公式	221
5.12	交换环上的行列式	223
	习题	228

第六章	重访环和多项式	237
6.1	理想和商环	239
6.2	多项式的唯一分解性质	244
6.3	简单推广: 主理想环的唯一分解性	248
6.4	形式求导	251
6.5	应用: Mason–Stothers 定理	254
6.6	根和重因式	255
6.7	对称多项式	259
6.8	结式	262
6.9	不可约多项式初探	266
6.10	从不可约多项式构造扩域	271
6.11	应用: 构造有限域	275
	习题	277
第七章	对角化	283
7.1	特征值与特征向量	284
7.2	极小多项式	290
7.3	上三角化	294
7.4	广义特征子空间	296
7.5	同步对角化	300
	习题	301
第八章	双线性形式	305
8.1	双线性形式	307
8.2	非退化形式与伴随映射	311
8.3	分类问题的提出	318
8.4	二次型的基本概念	321
8.5	配方法	323
8.6	实二次型的分类	325
8.7	反对称双线性形式: 辛空间	327
8.8	双重对偶	330
8.9	对偶与商	333
	习题	336
第九章	实内积结构	339
9.1	引言: 标准内积	341
9.2	内积空间	343

9.3	Gram–Schmidt 正交化	345
9.4	正交补与正交直和分解	349
9.5	内积空间上的伴随映射和正交变换	351
9.6	自伴算子的正交对角化	355
9.7	应用: 最小二乘解	358
9.8	对于正定二次型的应用	359
9.9	奇异值分解	362
9.10	Moore–Penrose 广义逆	364
9.11	极小化极大原理	367
9.12	Perron–Frobenius 定理	369
	习题	373
第十章	复内积结构	381
10.1	半双线性形式	383
10.2	Hermite 形式的分类	388
10.3	复内积空间和酉变换	392
10.4	正规算子的酉对角化	396
10.5	实定理的复推广	399
10.6	实正交变换的标准形	403
10.7	三维空间中的旋转与 Euler 角	407
10.8	四元数与旋转	410
	习题	416
	下册	421
第十一章	群的概念	423
11.1	群的基本定义	425
11.2	同态与同构	431
11.3	循环群	434
11.4	陪集分解	435
11.5	群作用	437
11.6	轨道分解的几则应用	441
11.7	应用: 置换的循环分解	443
11.8	回首高次方程	445
11.9	正规子群与商群	449
11.10	群的半直积	455
11.11	正多面体的对称群	458

习题	467
第十二章 模论入门	477
12.1 模的基本定义	478
12.2 模的同态, 同构与商	481
12.3 直和分解	486
12.4 自由模	488
12.5 基于挠子模的分解	491
12.6 主理想环上的有限生成模	493
12.7 基于矩阵的算法	498
习题	503
第十三章 标准形	507
13.1 线性映射和模结构	508
13.2 问题的表述	510
13.3 有理标准形	511
13.4 有理标准形的计算	515
13.5 Jordan 标准形	518
13.6 Jordan 标准形的计算	522
习题	524
第十四章 仿射空间与射影空间	527
14.1 仿射空间	529
14.2 仿射线性映射	533
14.3 刚体运动	536
14.4 射影空间	540
14.5 射影变换与交比	543
14.6 仿射空间的凸子集	548
14.7 多面体	551
14.8 关于极值问题	555
14.9 多面锥	557
14.10 多胞体基本定理	562
习题	565
第十五章 向量空间的张量积	571
15.1 以泛性质定义张量积	574
15.2 张量积的基本性质	579

15.3	张量积与对偶空间	583
15.4	应用: 域的变换	586
15.5	域上的代数	589
15.6	对称代数与外代数	591
15.7	Pfaff 型与交错矩阵的行列式	596
15.8	Amitsur–Levitzki 定理	599
15.9	特征零的情形	601
	习题	604
第十六章 二次型的 Witt 理论		611
16.1	二次型与正交群	613
16.2	消去定理与分解定理	616
16.3	Witt 群	619
16.4	全迷向子空间	623
16.5	Cartan–Dieudonné 定理	626
16.6	环结构: 二次型的张量积	628
16.7	具体实例	631
16.8	域上的 Hermite 形式	634
16.9	Hermite 形式的 Witt 理论	637
16.10	环上的 Hermite 形式概观	642
	习题	646
附录 A 集合论补遗		651
A.1	Peano 算术	651
A.2	构造非负整数集	655
A.3	基数补遗	657
A.4	Zorn 引理与基的存在性	659
附录 B 范畴引论		661
B.1	范畴	662
B.2	函子	668
B.3	自然变换	672
B.4	范畴等价	674
B.5	泛性质	677
B.6	等化子及余等化子	682
B.7	么半范畴一瞥	684

参考文献	687
符号索引	689
名词索引暨英译	691

上册

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, 如今的代数还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, 而其中相对容易而常见的一类是线性方程组

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

此处 X_1, \dots, X_n 代表待解的变元. 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 $M_{n \times n}(\mathbb{C})$, 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; 目的是激励, 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简单的数学分析, 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”都不可能完全符合, 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点: (一) 概念本身的历史顺序. (二) 可读性. (三) 理论的规整, 流畅和优雅气韵, 业内人士谓之“自然”. (四) 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, 仍会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. 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; 纵不知数学家的生平逸事, 损失的只是谈资, 对学习不成障碍.

最后对本书采用的术语和符号略作说明. 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用, 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差别. 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. 每章末尾附有若干习题, 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, 但习题可能依赖于更早的习题.

各章和附录内容摘要如下, 所涉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; 不使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; 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. 模论在此显现威力.

- ▷ **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.

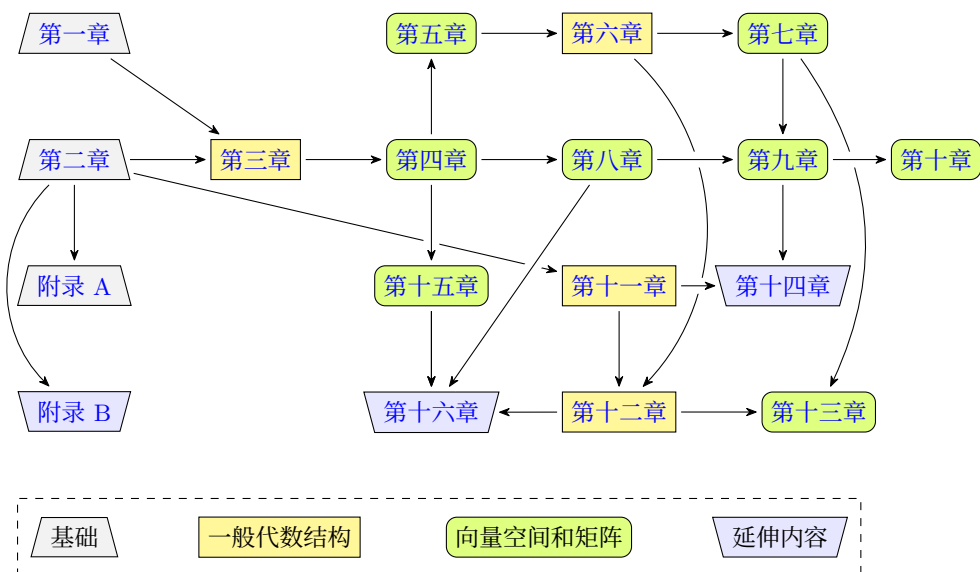
- ▷ **第十五章: 向量空间的张量积** 在选定的域上, 张量积是将两个向量空间 V 和 W 的元素形式地配对的一种手段, 由此得到新的向量空间 $V \otimes W$; 它的具体刻画适合以双线性形式的语言表述. 本章始于张量积的一般理论, 然后介绍对称代数与外代数, 及若干应用, 这些构造频繁用于几何等领域中.

- ▷ **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, 本章将定义 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.



索引

致学生 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，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，但两手协作方能成事。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，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, 此外则是图像, 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. 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.

致教师 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. 授课对象多数来自北京大学数学科学学院, 但不限于此.

虽然对象以数学专业的学生为主, 然而该课程不属于实验班, 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, 则其中有跳过者 (例如 §5.11, §5.12, §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, §8.9, §11.10, 第十五章, 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.

根本上, 笔者认为教科书作者的使命是提供齐全可靠而且有条理的素材, 剪裁变化之妙则在于讲者, 否则要教师何用? 因此本书既不划分选学内容, 也不按课时来切分章节. 但身为广大教师群体的一员, 请容笔者从上述个人经验提出几点建议.

首先, 板书是授课正道. 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.

巨细靡遗地讲授全书是不切实际的, 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, 可留一些支线内容自学. 习题课的安排多多益善.

在课时受限或需要对齐进度的情况下, 可以考虑省略或推迟关于商空间与环论的内容. 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. 随着内容的深入, 计算题与证明题的边界将逐渐消弭.

最后, 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.

鸣谢

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积累的材料. 谨向全体同学和助教们致谢, 没有他们的积极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, 记忆所及且有署名者包括: 曹连谦, 陈泽如, 褚浩云, 丁一文, 高剑伟, 蓝青, 马致远, 孙超超, 万铖睿, 王宇腾, 温焯暘, 吴家驹, 薛江维, 尹梓僮, 杨家铭, 朱奕 (按照姓名拼音排序). 在此致谢.

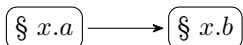
李文威

2024 年秋

谨识于门头沟

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 $\S x.a$ 必须先于 $\S x.b$ 阅读, 或者说 $\S x.b$ 依赖于 $\S x.a$. 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 \square 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)”. 图片按章编号, 例如“图 2-2”.

对于书中的各章, 节和附录, 我们将按照诸如

$$\text{第一章}, \quad \S 1.1, \quad \S\S A.1\text{--}A.3$$

的格式进行参照.

符号

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 \Longleftrightarrow 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 满足性质 \mathcal{P} 的 x , q 成立.

应当被理解为

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

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, \dots 构成的集合记为 $\{a, b, c, \dots\}$; 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

$\{s \in S : s \text{ 满足 } \mathcal{P}\}$ 或 $\{s \in S \mid s \text{ 满足 } \mathcal{P}\}.$

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 $\mathbf{x} = (x_1, \dots, x_n)$ 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 \mathbf{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 $\text{lcm}(a, b)$. 如果非零整数 n 不被 1 以外的任何完全平方数整除, 则称 n 无平方因子.

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

复数 z 的实部记为 $\text{Re}(z)$, 虚部记为 $\text{Im}(z)$. 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, \dots 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, \dots 为变元的多项式 f 为 $f(X, Y, Z, \dots)$, 而对 f 代值 $X = x, Y = y, \dots$ 的产物记为 $f(x, y, \dots)$.

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

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

对任意子集 $A' \subset A$, 记 $f|_{A'} : A' \rightarrow B$ 为 f 在 A' 上的限制; 另外记 A' 在 f 下的像为 $f(A') := \text{im}(f|_{A'})$. 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 $g \circ f : A \rightarrow C$, 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 $f : A \rightarrow B$ 下的像, 我们经常以符号 $f : a \mapsto b$ 或 $a \xrightarrow{f} b$ 代表 $f(a) = b$. 在具体描述一个映射时, 我们将频繁使用诸如

$$\begin{array}{l} f : A \rightarrow B \\ a \mapsto f(a) \end{array}$$

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 $f : A \rightarrow B$ 是满射当且仅当 $\text{im}(f) = B$. 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 $f : A \rightarrow B$ 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

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

并且在 $n < k$ 时规定 $\binom{n}{k} = 0$.

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

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

²⁾用函数的术语来说, 逆映射就是反函数.

的形式; 称 a_{ij} 为 \mathbf{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 $M_{m \times n}(\mathbb{Q}), M_{m \times n}(\mathbb{R}), M_{m \times n}(\mathbb{C})$, 依此类推. 有时对 $m \times n$ 矩阵也采取 $\mathbf{A} = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ 的记法.

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

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

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 $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq n}$ 的行列式为

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

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 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, 我们不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-muqābala* 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 $1, 2, 3, \dots$ ，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

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

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 $x_1, \dots, x_n \in \mathbb{C}$ 使得

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

这些 x_1, \dots, 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.

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

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

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.

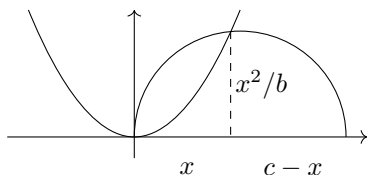


图 1-1

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 $x^3 + b^2x = b^2c$.

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. 考虑关于变元 \mathcal{X} 的一元三次方程

$$\mathcal{X}^3 + a\mathcal{X}^2 + b\mathcal{X} + c = 0.$$

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

$$X^3 + pX + q = 0$$

的形式, 今后仅考虑此类的三次方程. 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

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

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

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

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

$$D := -4p^3 - 27q^2. \quad (1.1.2)$$

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

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

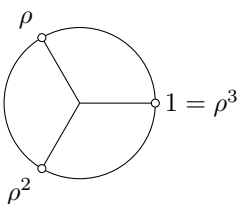
所以只要取 u 和 v 分别为 $-\frac{q}{2} + \frac{\sqrt{-3D}}{2 \cdot 3^2}$ 和 $-\frac{q}{2} - \frac{\sqrt{-3D}}{2 \cdot 3^2}$ 的立方根, 则 u 和 v 都是辅助方程 (1.1.1) 的解. 我们断言:

- ★ 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 $\lambda := uv / (-\frac{p}{3})$, 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 $-3D = 3^4 q^2$, 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 $u_i v_i = -\frac{p}{3}$ 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):

$$x_i = u_i + v_i, \quad i = 1, 2, 3. \quad (1.1.4)$$

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, 也一律在复数系里操作, 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.

(i) 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 $\sqrt[3]{\dots}$ 代表适当选取的立方根.

- (ii) 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

$$\overline{u_i v_i} = -\frac{p}{3} = u_i v_i,$$

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 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 $m \in \mathbb{Z}_{\geq 1}$.

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*; $n \geq 5$ 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

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

而因为 $u_i v_i = -\frac{p}{3}$, 相应地有

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

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

$$u_1 = (x_1 + \rho^2 x_2 + \rho x_3)/3,$$

$$u_2 = (\rho x_1 + x_2 + \rho^2 x_3)/3,$$

$$u_3 = (\rho^2 x_1 + \rho x_2 + x_3)/3,$$

$$v_1 = (x_1 + \rho x_2 + \rho^2 x_3)/3,$$

$$v_2 = (\rho^2 x_1 + x_2 + \rho x_3)/3,$$

$$v_3 = (\rho x_1 + \rho^2 x_2 + x_3)/3.$$

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

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

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, \dots, 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, \dots, 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 $x_1 + x_2 + x_3 = 0$.

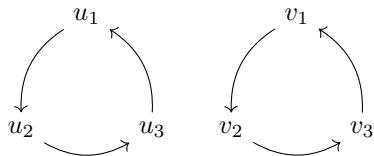
这些观察可以说是 Cardano 公式的实质. 推导的关键在于:

▷ **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 $\sigma : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ 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, \dots, 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

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

类似地, $v_2 = \rho^2 v_1$ 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

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

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

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

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

平方和问题依此改写为

$$z\bar{z} = m, \quad z = x + iy \in \mathbb{Z}[i].$$

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

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

综上, $\mathbb{Z}[i]$ 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

当 m 除以 4 余 3 时, 不定方程 $X^2 + Y^2 = m$ 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 $X^2 + Y^2 = n$ 何时解, 以及有几组解. 为此就有必要更详细地了解 $\mathbb{Z}[i]$ 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: 留意到 $t^2 + 1 \neq 0$, 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

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

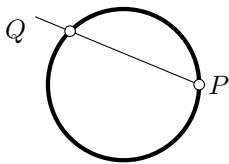


图 1-2

综上, 我们通过 $Q \leftrightarrow t$ 建立一一对应

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

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

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

(i) 考虑映射 $f: S \rightarrow S$ 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 $f: S \rightarrow S$ 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 $g: S \rightarrow S$, 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

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

的方程组, 其中 a_{ij} 和 b_i 是给定的常数 ($1 \leq i \leq m, 1 \leq j \leq n$), 而 X_1, \dots, 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, \dots, 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, 可知方程解集的性质和“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:

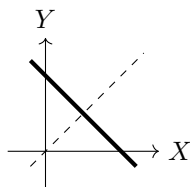


图 1-3

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之下, 或在按一定顺序轮换之下的不变性, 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

$$\begin{cases} a_{11}X_1 + \cdots + a_{1n}X_n \leq b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n \leq b_2 \\ \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n \leq b_m \end{cases}$$

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

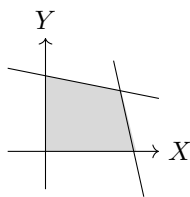


图 1-4

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, \dots, 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

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

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

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

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

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

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

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

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 $\textcircled{3}''$ 将是自相矛盾的;
- ★ 设 $-a - 3b + 2c = 0$, 则 $\textcircled{2}''$ 给出 X_3 , 代入 $\textcircled{1}$, 得出方程的通解

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

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, \dots, b_m 的矩阵则称为**增广矩阵**.

矩阵行, 列的具体记法是

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

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, \dots, b_m 角色毕竟不同于系数 a_{ij} , 有时在增广矩阵中作区隔:

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

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 $1 \leq i \neq k \leq m$, 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:

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

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

$$B(i, k): \begin{matrix} i \\ \vdots \\ \cdots & a_{ij} & \cdots \\ \vdots \\ k \\ \cdots & a_{kj} & \cdots \\ \vdots \end{matrix} \xrightarrow{\text{交换}} \begin{pmatrix} \vdots \\ \cdots & a_{kj} & \cdots \\ \vdots \\ \cdots & a_{ij} & \cdots \\ \vdots \end{pmatrix}$$

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

$$C(i, c): \begin{matrix} i \\ \vdots \\ \cdots & a_{ij} & \cdots \\ \vdots \end{matrix} \xrightarrow{\cdot c} \begin{pmatrix} \vdots \\ \cdots & ca_{ij} & \cdots \\ \vdots \end{pmatrix}$$

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, \dots, X_n 的顺序恒定. 如果 (x_1, \dots, 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:

- ★ $A(i, k, c)$ 的逆操作是 $A(i, k, -c)$,
- ★ $B(i, k)$ 的逆操作是 $B(i, k)$ 自身,
- ★ $C(i, c)$ 的逆操作是 $C(i, 1/c)$,

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

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

当它的形式如下所示时, 称之为**行梯矩阵**:

$$\begin{pmatrix} \text{---} & & & \\ & \text{---} & & \\ & & \text{---} & \\ & & & \ddots \\ & & & & \vdots \end{pmatrix}$$

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 $a_{1,j_1}, \dots, a_{r,j_r}$.

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.

- (i) 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:

$$\left(\begin{array}{c|cccc} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & \ddots & \\ 0 & a_{m2} & a_{m3} & \cdots & a_{mn} \end{array} \right)$$

全为 0 的列不受后续变换的影响, 可以放心舍去.

- (ii) 设矩阵第一列有非零元, 设之为 a_{k1} . 进行先前标为 $B(k, 1)$ 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.
- (iii) 设 $a_{11} \neq 0$. 接着对每个 $1 < i \leq m$, 进行先前标为 $A\left(1, i, -\frac{a_{i1}}{a_{11}}\right)$ 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

$$\left(\begin{array}{c|cccc} \mathbf{a_{11}} & a_{12} & & \cdots & a_{1n} \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \begin{array}{ccc} & & \\ \cdots & a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}} & \cdots \\ & \vdots & \end{array} \right)$$

然后我们对框出的子矩阵继续操作.

之所以将 $\mathbf{a_{11}}$ 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.

- (iv) 对每个主元 a_{k,j_k} , 作变换 $C\left(k, a_{k,j_k}^{-1}\right)$ 以化约到行梯矩阵的主元全为 1 的情形.
- (v) 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 $A(k, i, -a_{ij_k})$; 换言之, 将第 k 行乘以 $-a_{ij_k}$ 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.

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

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

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”.

- (iii) 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 $\zeta = e^{2\pi i/5}$ 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 $u^5 v^5 = a^5$ 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.

(i)

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

(ii)

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

(iii)

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

(iv)

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

(v)

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

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.

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

5. 考虑两个大小相同的矩阵

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

如果可以通过一系列初等行变换从 A 过渡到 B , 则称 A 和 B 是行等价的.

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

(ii) 设所论矩阵形如

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

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

提示 将 A 和 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 A 和 B 是行等价的简化行梯矩阵, 则 $A = B$.

提示 设 $A \neq B$. 从左而右比较每一列, 设第一个相异的列为第 j 列. 从 A 和 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 A' 和 B' .

★ 论证第 j 列不可能包含主元.

★ 论证 A' 和 B' 必然是 (ii) 之中的形式. 配合 (i) 来推导 $A' = B'$, 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

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

说明之前的线性方程组能以增广矩阵的写法表作

$$\left(\mathbf{1}_{n \times n} - \mathbf{A} \mid \mathbf{d} \right),$$

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, $\mathbf{1}_{n \times n} - \mathbf{A}$ 意谓将两个矩阵逐项相减 (请写出它的大致样貌), 而 \mathbf{d} 是只有一列的 $n \times 1$ 矩阵 (常称为 n 维列向量)

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

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 $\mathbf{1}_{n \times n} - \mathbf{A}$ 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.

