线性代数讲义

上册

李 忠华

同济大学　数学科学学院

理论是灰色的, 生命之树常青!

歌德《浮士德》

代数几何熔一炉,
乾坤万物坐标书,
图形百态方程绘,
变换有规矩阵筹.

李尚志

前言

代数学是研究数, 数量, 关系与结构的一个数学分支, 其又分为古典代数和近世(抽象)代数. 古典代数主要研究代数方程的解, 比如研究一元二次方程

$$ax^2 + bx + c = 0 \quad (a \neq 0) \tag{0.1}$$

的解; 而近世代数则研究代数结构, 比如研究群, 环和域等.

线性代数研究线性空间及其之间的线性变换, 矩阵是线性代数的一个重要代数工具, 线性方程组则是一个重要的研究对象. 线性方程组的解的研究属于古典代数学的范畴, 而线性空间则是最基本的代数结构. 在线性代数中我们还需要多项式这一工具.

下面简单介绍一下上面出现的一些名词, 它们是本讲义的主角.

我们在中学就学过了(一元)多项式, 比如$ax + b$ $(a \neq 0)$是一次多项式, 而

$$ax^2 + bx + c \quad (a \neq 0)$$

是二次多项式. 在中学我们也学过方程, 比如一元二次方程(0.1)是中学数学的一个重要研究对象. 我们称形如

$$ax + by + c = 0$$

的一次方程为(二元)线性方程, 而若干个线性方程组成的方程组则称为线性方程组. 看下面的经典例子.

例(鸡兔同笼《孙子算经》). *今有雉兔同笼, 上有三十五头, 下有九十四足, 问雉兔各几何?*

☞ **解.** 可以如下求解:

$$\begin{aligned} \text{兔子数} &= \frac{94 - 35 \times 2}{2} = 12, \\ \text{鸡数} &= 35 - 12 = 23. \end{aligned}$$

也可以假设鸡有x只, 兔有y只, 则得到方程组

$$\begin{cases} x + y = 35, \\ 2x + 4y = 94. \end{cases}$$

上面这个方程组就是一个线性方程组, 可如下求解.

$$\begin{cases} x + y = 35, & (1) \\ 2x + 4y = 94. & (2) \end{cases} \xrightarrow{(2)-2\times(1)} \begin{cases} x + y = 35, & (1) \\ \quad 2y = 24. & (2) \end{cases}$$

$$\xrightarrow{\frac{1}{2}\times(2)}\begin{cases} x+y=35, & (1)\\ \quad\ \ y=12. & (2)\end{cases}\xrightarrow{(1)-(2)}\begin{cases} x=23,\\ y=12.\end{cases}$$

上面求解线性方程组的方法在中学也学过, 这里我们称为**Gauss消元法**. □

我们称矩形(长方形)数表为矩阵. 例如上面鸡兔同笼问题得到的线性方程组中, 未知数x和y不是本质的(我们可以用其它字母做未知数), 本质的是未知数的系数和常数项, 于是我们得到下面的矩阵

$$\begin{pmatrix} 1 & 1 & 35\\ 2 & 4 & 94\end{pmatrix},$$

它是对应的线性方程组的增广矩阵. 上面解方程组的Gauss消元法就对应到增广矩阵的化简(矩阵的初等变换)

$$\begin{pmatrix} 1 & 1 & 35\\ 2 & 4 & 94\end{pmatrix}\xrightarrow{r_2-2r_1}\begin{pmatrix} 1 & 1 & 35\\ 0 & 2 & 24\end{pmatrix}\xrightarrow{\frac{1}{2}r_2}\begin{pmatrix} 1 & 1 & 35\\ 0 & 1 & 12\end{pmatrix}\xrightarrow{r_1-r_2}\begin{pmatrix} 1 & 0 & 23\\ 0 & 1 & 12\end{pmatrix}.$$

矩阵是数的推广, 数有四则运算, 类似的我们将定义矩阵的四则运算: 加, 减, 乘和"除"; 我们还将如上面Gauss消元法一样研究矩阵的化简及其应用.

线性空间是定义有"好的"加法和数乘(数量乘法)的非空集合, 最经典的例子就是我们生活的三维空间. 当取定了(直角)坐标系后, 三维空间的每一点都可以唯一用一个有序的三元数组表示, 于是三维空间可以表示为集合

$$\mathbb{R}^3=\{(a_1,a_2,a_3)\mid a_1,a_2,a_3\in\mathbb{R}\},$$

其中$\mathbb{R}$表示实数的集合, 见图0.1. 对应到三维向量的加法和数乘, 这个集合可以定义如下运算:

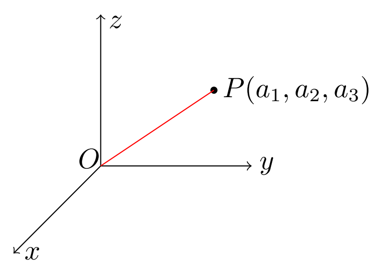


图 0.1: 3维空间

$$\begin{aligned}&\text{加法}:(a_1,a_2,a_3)+(b_1,b_2,b_3):=(a_1+b_1,a_2+b_2,a_3+b_3),\\ &\text{数乘}:a(a_1,a_2,a_3):=(aa_1,aa_2,aa_3).\qquad (a\in\mathbb{R})\end{aligned}$$

在这两个运算下, $\mathbb{R}^3$成为一个线性空间. 读者应该用这个具体的几何例子来理解本讲义中的许多抽象概念.

线性变换是线性空间之间"好的"映射. 和线性空间一样, 线性变换也是一个几何对象, 它们是本讲义的重点.

严格地讲, 多项式理论并不属于线性代数的范畴, 但是一般的处理方法都是和线性代数合在一起学习的[1]. 由于多项式理论从感觉上和线性代数很不一样, 所以为了不破坏线性代数的完整性, 我

[1]当然, 也有不同的处理办法. 比如, 笔者当年多项式理论是在初等数论这门课程中, 学完整数的素分解理论后, 再学习的.

们先讲多项式理论, 然后再进入线性代数. 对于线性代数, 我们先讲古典理论, 即线性方程组和矩阵, 再讲抽象理论, 即线性空间和线性变换. 本讲义的前半部分可以以线性方程组的求解为主线, 我们先介绍了线性方程组求解的Gauss消元法, 由此引出矩阵以及矩阵的初等变换, 从而得到线性方程组解的个数的判别方法和求解算法. 由于矩阵这一角色的重要性, 我们马上介绍了矩阵的四则运算. 求解线性方程组除了Gauss消元法外, 还有所谓的加减消元法, 这种消元法可以求解一类特殊的线性方程组, 从而得到一定条件下的线性方程组的求解公式, 这一公式中出现了方阵的行列式. 于是在矩阵的运算后, 我们介绍了行列式的概念以及用加减消元法得到的一定条件下线性方程组的求解公式(Cramer法则). 最后, 我们来数线性方程组中(真正)方程的个数, 以及有解的线性方程组解的(真正)个数, 这就需要秩的概念. 至此, 线性方程组的故事就结束了.

本讲义的后半部分是线性代数的抽象(几何)理论. 对此, 我们建议读者把后半部分看成是高维的解析几何. 如前面, 在我们生活的三维空间建立(直角)坐标系, 那么空间中的每个点(向量)就唯一对应到有序的三元数组, 而点(向量)之间的运算就可以翻译成三元有序数组之间的运算. 三维空间是向量组成的空间, 而线性空间则可以想成广义向量组成的空间. 我们在介绍完了线性空间以及子空间等概念后, 一个主要的目标就是想在线性空间中建立坐标系. 当建立了坐标系后, 可以看到, n维(维数可以看成坐标系中坐标轴的个数)线性空间本质上等同于n元有序数组组成的空间. 在三维空间中, 我们常常需要研究变换, 于是我们也希望研究一般线性空间中好的变换—线性变换. 用解析几何来解决几何问题时, 问题解决的难易常常与坐标系的选取有关. 例如, 我们在平面上画个半径为1的圆. 如果如图0.2(a)中将坐标原点放在圆心的话, 则这个圆在这个坐标系中的方程就是

$$x^2 + y^2 = 1.$$

但是如果像图0.2(b)中建立坐标系的话, 则这个圆在这个坐标系中的方程会很复杂. 类似的, 做解析几何, 建立坐标系后, 可以看到线性变换就对应到矩阵. 取不同的坐标系, 线性变换的表现形式也不一样. 一个基本问题就是找最合适的坐标系, 使得线性变换的表现形式最简单, 如此在这个坐标系下就更容易解决问题. 这个基本问题对应到矩阵在相应的等价关系下的分类问题. 三维空间中通常取直角坐标系, 于是我们接着讨论了可以有直角坐标系的线性空间—内积空间, 以及这个空间中的线性变换的基本问题: 找最合适的直角坐标系, 使得线性变换的表现形式最简单. 当然, 这一问题也对应到矩阵的分类. 在微积分中, 考虑了一元函数后, 自然考虑多元函数. 类似地, 我们最后讨论了线性空间上的特殊的二元函数—双线性函数, 这可以看成是内积空间中内积的一般化. 做解析几何, 建立坐标系后, 双线性函数也对应到矩阵. 我们想找最合适的坐标系, 使得双线性函数在这个坐标系下的表达形式最简单. 这一基本问题也对应到矩阵的某种分类. 所以本讲义的后半部分, 从几何上看讲的是找最合适的坐标系, 使得线性空间上的线性变换等在这个坐标系中的表现形式最简单; 而从代数上看, 则是讲矩阵在各种等价关系下的分类.

笔者的线性代数是李尚志老师教的, 有幸从他那学到了许多的矩阵技巧和对线性代数的理解[2]; 后来当宋光天老师的线性代数课程的助教时, 又学到了许多几何讨论的方法, 比如一般数域上的Jordan标准形理论. 进入同济大学数学系后, 又从同事那里学到了许多线性代数课程的处理方法. 本讲义既是两位老师传授的一些知识和观点的传承, 也是笔者个人讲授线性代数课程的总结, 还主要参考了经典教材[10, 11, 12, 14, 15], 特别是当年笔者念的教材[11]. 我们从求解线性方程组入手, 一

[2]有学长曾经说, 李尚志老师讲线性代数其实是在讲艺术.

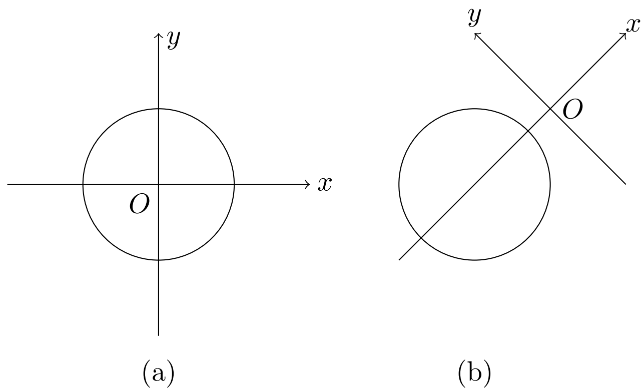


图 0.2: 坐标系的选取

步一步进入线性代数的多彩世界, 有下面特点:

(i) 我们尽量做到讲义中的每一个概念和结果都是自然和必要的. 为此, 我们尽量用“白话”, 而不用“黑话”, 给每个概念和结果的出现一个理由. 作为一门完备的数学课程, 除了将来在其它数学分支或者其它科学领域的应用外, 线性代数也有其关心的问题和发展历程. 除了掌握线性代数的一些基本概念, 基本结果和基本方法外, 我们更希望读者多问为什么, 能对线性代数在干什么给出自己的理解, 并发现线性代数的美. 在课程中我们尽量给出对这门课程的理解, 所以在这门课程中我们在讲故事, 她有支线故事, 有前传, 有开头, 有高潮, 也有结尾. 我们希望整个课程是一些有关联的小故事: 多项式的整数世界历险记, 线性方程组的求解之旅, 寻找坐标系的故事等等;

(ii) 空间为体, 矩阵为用. 我们强调几何(空间)和代数(矩阵)的对应和联系. 许多问题, 我们同时给出了几何和代数处理, 希望读者可以体会这两种不同处理的特点. 我们既强调传自华罗庚先生的矩阵打洞技巧(初等变换), 也强调抽象的理论力量;

(iii) 习题是讲义的重要组成部分, 我们花了很大力气收集习题, 并且自编了部分习题. 有些习题在后面的正文中要用到, 有些习题则需要其它一些习题的结果. 按照我们的经验, 我们将习题分成三类: A类, B类和C类, 其中A类是基本题, B类题需要一定的思考, 而C类题则有一定的难度. 要学好线性代数, 除了天才外, 在理解了课程内容后, 都需要做一定数量的习题. 我们希望在学完本讲义后, 大家可以做到“老谋深算”, 不但会算, 还会证明.

我们使用下面的数集记号:

$$
\begin{aligned}
&\mathbb{N} := \{1, 2, 3, 4, \ldots\} = \{\text{正整数}\},\\
&\mathbb{Z}_{\geqslant 0} := \{0, 1, 2, 3, \ldots\} = \{\text{非负整数}\},\\
&\mathbb{Z} := \{0, \pm 1, \pm 2, \pm 3, \ldots\} = \{\text{整数}\},\\
&\mathbb{Q} := \left\{\frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{N}\right\} = \{\text{有理数}\},\\
&\mathbb{R} := \{\text{实数}\},
\end{aligned}
$$

$$\mathbb{C} := \{a + bi \mid a, b \in \mathbb{R}\} = \{\text{复数}\},$$

其中$i = \sqrt{-1}$是虚数单位. 显然有下面的包含关系

$$\mathbb{N} \subsetneqq \mathbb{Z}_{\geqslant 0} \subsetneqq \mathbb{Z} \subsetneqq \mathbb{Q} \subsetneqq \mathbb{R} \subsetneqq \mathbb{C}.$$

一般地, 我们用$\mathbb{F}$表示一个数域, 用小写字母a, b, c等表示$\mathbb{F}$中的数, 有时也用λ表示$\mathbb{F}$中的数. 我们用大写字母A, B, C等表示矩阵, 用希腊字母α, β, γ等表示向量. 最后, 对于一个有限集合A, 我们用$\sharp A$或者$|A|$表示A中所含元素的个数.

笔者要感谢一路走来给予了我谆谆教诲和无私帮助的师长们, 以及我所念过和看过的线性代数教材的作者们; 要感谢同济大学数学科学学院各位老师在线性代数教学上的帮助, 特别是叶家琛教授和靳全勤教授; 还要感谢同济大学数学科学学院主动或被动念过这本讲义初稿的同学们, 特别是2016级的赵乾宇和2018级的赵志鹏, 他们提供了一些习题的精妙解法; 当然, 还要感谢我的妻子的理解和默默付出.

2020年9月

目录

第一章　多项式 **1**

1.1 数域 1
1.2 一元多项式 2
1.3 带余除法 6
1.4 整除性 11
1.5 最大公因式 14
1.6 多项式互素 21
1.7 唯一分解定理 23
1.8 多项式的根 29
1.9 重因式 32
1.10 复系数多项式的因式分解 38
1.11 实系数多项式的因式分解 44
1.12 有理系数多项式的因式分解 46
1.13 多元多项式 54
1.14 对称多项式 60
1.15 附录: 代数基本定理的几个证明 69
1.16 附录: 一元三次和一元四次多项式的求根公式 74
1.17 附录: Schur定理的一个证明 76
1.18 附录: 判别式和结式 78
1.19 附录: 群, 环, 域和主理想整环 91
1.20 补充题 98

第二章　线性方程组与矩阵 **101**

2.1 线性方程组的初等变换 101
2.2 矩阵和线性方程组 105
2.3 矩阵的初等变换 108
2.4 线性方程组的求解 113

第三章 矩阵的运算 **120**

3.1 矩阵的运算 120
3.2 可逆矩阵的定义和性质 132
3.3 矩阵的分块 137
3.4 初等阵与初等变换 145
3.5 可逆矩阵求逆 150
3.6 分块矩阵求逆 161
3.7 补充题 166

第四章 行列式 **167**

4.1 线性方程组与二, 三阶行列式 167
4.2 排列及其符号 172
4.3 行列式的定义 175
4.4 初等变换和行列式 181
4.5 分块矩阵的初等变换和行列式 192
4.6 行列式计算之例一 194
4.7 行列式的按行(列)展开 203
4.8 行列式计算之例二 211
4.9 Laplace展开 219
4.10 Binet-Cauchy公式 228
4.11 Cramer法则 236
4.12 附录: 行列式的几何意义 239
4.13 补充题 241

第五章 向量组与矩阵的秩 **246**

5.1 向量与向量组 246
5.2 向量组的线性组合 252
5.3 向量组的线性相关性 254
5.4 向量组间的等价 262
5.5 向量组的秩 265
5.6 矩阵的秩—定义与计算 270
5.7 矩阵的秩—性质与应用 281
5.8 向量空间 295
5.9 线性方程组解的结构 297
5.10 附录: 线性方程组和初等变换标准形 307
5.11 附录: 矩阵的广义逆 311
5.12 补充题 318

第六章 线性空间 **321**
6.1 线性空间 321
6.2 子空间 325
6.3 线性相关性 330
6.4 基与维数 334
6.5 基变换与坐标变换 345
6.6 同构 351
6.7 直和 360
6.8 商空间 367
6.9 附录: 线性空间的无限基的势 371
6.10 附录: Fibonacci数列和幻方 372
6.11 附录: 模 375
6.12 补充题 383

第一章　多项式

多项式在讨论某些线性代数问题中是重要的工具, 本章主要介绍了一元多项式的基本理论. 多项式世界和整数世界可以看成在某种程度上相平行的两个世界, 所以本章可以看成是多项式在整数世界的历险故事. 每个整数可以唯一分解成素数的乘积, 本章的主要目的是介绍多项式世界的对应结果—因式分解存在唯一定理. 为此, 我们需要介绍整除理论, 最大共因式以及多项式世界的素数—不可约多项式. 本章最后, 我们还简单介绍了多元多项式, 特别是对称多项式的基本理论.

这部分内容严格讲不属于线性代数, 所以不是本讲义的重点. 但是多项式是代数学中一个重要的对象, 不但许多代数结构的例子可以用多项式构造出来, 而且多项式在表示论和代数几何等学科中将不断出现.

§ 1.1　数域

我们要对数进行加, 减, 乘和除四则运算, 当然希望运算出来的数我们还认识, 也就是希望在对四则运算封闭的数集中考虑问题. 例如, 所有复数组成的集合[1]$\mathbb{C}$当然对四则运算封闭, 所以我们可以在$\mathbb{C}$中考虑问题. 但有时候$\mathbb{C}$太大了, 所以很自然地引入下面的概念.

定义1.1. 称子集$\mathbb{F} \subset \mathbb{C}$是数域, 如果满足

(i) $0, 1 \in \mathbb{F}$;

(ii) $\mathbb{F}$对数的加, 减, 乘, 除(除数非零)封闭, 即对任意$a, b \in \mathbb{F}$, 有

$$a + b,\ a - b,\ a \cdot b,\ \frac{a}{b}(b \neq 0) \in \mathbb{F}.$$

事实上, 如果一个非空集合可以定义满足好的运算律的四则运算, 则该集合成为一种特殊的代数结构—域(参见本章附录). 数域是一种特殊的域, 线性代数包括多项式的许多理论都可以毫无困难地推广到一般的域上.

例1.1. **(1)** $\mathbb{Q}$, $\mathbb{R}$和$\mathbb{C}$都是数域, 分别称为有理数域, 实数域和复数域. 本讲义中的数域$\mathbb{F}$通常为这三个数域;

(2) 整数集合$\mathbb{Z}$不是数域, 它对除法不封闭(事实上, 容易证明(见本节习题A1), 如果$\mathbb{F}$是一个数域, 那么有$\mathbb{Q} \subset \mathbb{F}$.).

[1]这是我们中学学过的最大数集.

我们最后看一个不那么明显的数域的例子[2].

例1.2. 证明: 集合

$$\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$$

是数域.

☞ **证明.** 首先我们有

$$0 = 0 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2}), \quad 1 = 1 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

其次对于任意$a_1 + b_1\sqrt{2}, a_2 + b_2\sqrt{2} \in \mathbb{Q}(\sqrt{2})$, 其中$a_i, b_i \in \mathbb{Q}$, 我们有

$$\begin{aligned}(a_1 + b_1\sqrt{2}) \pm (a_2 + b_2\sqrt{2}) &= (a_1 \pm a_2) + (b_1 \pm b_2)\sqrt{2} \in \mathbb{Q}(\sqrt{2}),\\ (a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2}) &= (a_1a_2 + 2b_1b_2) + (a_1b_2 + a_2b_1)\sqrt{2} \in \mathbb{Q}(\sqrt{2}).\end{aligned}$$

再假设$a_2 + b_2\sqrt{2} \neq 0$, 则$a_2 - b_2\sqrt{2} \neq 0$ (否则, $a_2 = b_2\sqrt{2}$. 如果$b_2 = 0$, 则$a_2 = 0$, 这得到$a_2 + b_2\sqrt{2} = 0$, 矛盾. 因此$b_2 \neq 0$, 进而$\sqrt{2} = a_2/b_2 \in \mathbb{Q}$, 矛盾于$\sqrt{2}$是无理数.). 于是

$$\frac{a_1 + b_1\sqrt{2}}{a_2 + b_2\sqrt{2}} = \frac{(a_1 + b_1\sqrt{2})(a_2 - b_2\sqrt{2})}{(a_2 + b_2\sqrt{2})(a_2 - b_2\sqrt{2})} = \frac{a_1a_2 - 2b_1b_2}{a_2^2 - 2b_2^2} + \frac{a_2b_1 - a_1b_2}{a_2^2 - 2b_2^2}\sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

这就证明了$\mathbb{Q}(\sqrt{2})$是数域. □

本书中, 我们总用$\mathbb{F}$表示一个数域, 而$\mathbb{F}^\times$表示$\mathbb{F}$中非零数构成的集合, 即

$$\mathbb{F}^\times := \mathbb{F} - \{0\} = \{c \in \mathbb{F} \mid c \neq 0\}.$$

习题1.1

A1. 设$\mathbb{F}$是数域, 证明: $\mathbb{Q} \subset \mathbb{F}$.

A2. 设整数$d > 1$, 且d无平方因子[3], 证明: 数集

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$$

是数域.

B1. 叙述域的定义, 并证明数域是域, 再举一个是域而不是数域的例子.

§ 1.2 一元多项式

下面我们开始多项式的故事. 首先要有一个固定的数域$\mathbb{F}$. 再设x是一个不定元[4](形式符号), 下面引入本章故事的主角—关于不定元x的以$\mathbb{F}$为系数的多项式, 对此读者并不陌生, 其实这些是中学所学的一个复习.

[2]在代数数论中可以看到更多数域的例子.

[3]称正整数n无平方因子, 如果对n的任意大于1的因子d, d都不是完全平方数.

[4]这里不称x为未知数, 因为将来x不一定只代入数, 它可以代入其它合适的对象, 例如矩阵.

定义1.2. 称形式表达式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

为$\mathbb{F}$上的一个关于不定元x的一元多项式, 也称为一个$\mathbb{F}$-系数一元多项式. 这里, $a_0, a_1, \ldots, a_n \in \mathbb{F}$是多项式$f(x)$的系数, 而$n \in \mathbb{Z}_{\geqslant 0}$是非负整数. 有时也将$f(x)$简记为[5]$f$.

下面介绍一些常用的名词和记号, 可以把这个想象成介绍主角的外貌特征和衣着等.

(i) 用求和记号, $f(x)$也可以记为

$$f(x) = \sum_{i=0}^{n} a_i x^i.$$

称$a_i x^i$为i次项, 称a_i为i次项系数. 此时对任意整数$j > n$, 也称$f(x)$的j次项系数为零. 即对任意$j > n$, 定义$a_j = 0$, 则有

$$f(x) = \sum_{i \geqslant 0} a_i x^i = \sum_{i=0}^{\infty} a_i x^i;$$

(ii) 当$a_n \neq 0$时, 称$a_n x^n$为首项, 又称a_n为首项系数. 此时, n称为$f(x)$的次数, 也称$f(x)$是一个n次多项式, 记为

$$\deg(f(x)) = n;$$

(iii) 称a_0为常数项;

(iv) 首项系数为1的多项式称为首一多项式;

(v) 所有系数都为零的多项式称为零多项式, 记为0. 约定零多项式的次数为$-\infty$, 即

$$\deg 0 = -\infty.$$

于是零多项式没有首项, 不是首一多项式;

(vi) 数域$\mathbb{F}$上所有一元多项式做成的集合记为$\mathbb{F}[x]$, 即

$$\mathbb{F}[x] := \{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \mid n \in \mathbb{Z}_{\geqslant 0}, a_0, a_1, \ldots, a_n \in \mathbb{F}\}.$$

本章的主要目的就是研究这个集合.

由上面的定义, 在多项式$f(x)$的表达式中未出现的项, 都认为其系数为零. 例如,

$$ax^2 + bx + c = 0 \cdot x^3 + ax^2 + bx + c = 0 \cdot x^4 + 0 \cdot x^3 + ax^2 + bx + c = \cdots.$$

取每个多项式的次数, 得到一个函数

$$\deg : \mathbb{F}[x] \longrightarrow \mathbb{Z}_{\geqslant 0} \cup \{-\infty\}; f(x) \mapsto \deg(f(x)).$$

[5]将$f(x)$简记为f是把$f(x)$看成$\mathbb{F}$上的一个函数

$$f : \mathbb{F} \longrightarrow \mathbb{F}; a \mapsto f(a).$$

利用次数, 可以判别一个多项式是否非零: 对任意$f(x) \in \mathbb{F}[x]$, 有

$$f(x) = 0 \Longleftrightarrow \deg(f(x)) = -\infty.$$

还要注意区分零多项式和零次多项式. 由定义, 零次多项式是$\mathbb{F}$中的非零数, 于是零次多项式组成的集合为$\mathbb{F}^{\times}$. 可以将零多项式与$\mathbb{F}$中的零等同, 于是$\mathbb{F} \subset \mathbb{F}[x]$.

最后我们再给出两个多项式相等的如下自然的定义.

定义1.3. 设$f(x), g(x) \in \mathbb{F}[x]$, 称$f(x)$和$g(x)$相等, 如果它们的任意同次项的系数都相等. 记为$f(x) = g(x)$.

至此, 我们知道了名叫多项式的这样一个角色. 它们来到了整数$\mathbb{Z}$的世界, 为这个世界的丰富多彩而感动. 比如, 它们看到整数可以做加法, 减法, 乘法和除法("带余除法"). 于是, 它们也希望可以做类似的运算. 怎么办呢? 多项式有两部分, 系数和不定元. 我们看到, 系数域$\mathbb{F}$是可以做运算的. 利用$\mathbb{F}$中的加法, 减法和乘法, 它们想到了下面自然的加法, 减法和乘法运算的定义, 这也是我们在中学时学到的.

设有多项式

$$f(x) = \sum_{i=0}^{n} a_i x^i,\ g(x) = \sum_{j=0}^{m} b_j x^j \in \mathbb{F}[x].$$

定义1.4 (加法)**.** 将多项式$f(x)$和$g(x)$的同次项系数相加, 得到的多项式记为$f(x) + g(x)$, 即

$$f(x) + g(x) := (a_0 + b_0) + (a_1 + b_1)x + \cdots = \sum_{i \geqslant 0} (a_i + b_i) x^i \in \mathbb{F}[x].$$

事实上, 多项式相加就是我们在中学学过的合并同次项. 例如, 我们有

$$(x^2 + 2x + 1) + (3x^3 - 2x^2 - 4x + 5) = 3x^3 - x^2 - 2x + 6.$$

由定义, 多项式相加, 其次数满足[6]

$$\deg(f(x) + g(x)) \leqslant \max\{\deg(f(x)), \deg(g(x))\}.$$

下面的性质也容易证明

(A1) (交换律) $f(x) + g(x) = g(x) + f(x)$;

(A2) (结合律) $(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$;

(A3) (零元存在) $f(x) + 0 = f(x) = 0 + f(x)$;

[6] max为取大函数, 即对于有限的有序集合A, $\max A$取出A中的最大元, 特别有

$$\max\{m, n\} = \begin{cases} m, & \text{如果 } m \geqslant n, \\ n, & \text{如果 } m < n. \end{cases}$$

这里我们约定对任意$n \in \mathbb{Z}_{\geqslant 0}$, 有$-\infty < n$.

(A4) (负元存在) $\forall f(x)=\sum_{i=0}^{n} a_i x^i$, 令$-f(x)=\sum_{i=0}^{n}(-a_i)x^i$, 则

$$f(x)+(-f(x))=0.$$

定义1.5 (减法)**.** 将多项式$f(x)$和$g(x)$的同次项系数相减, 得到的多项式记为$f(x)-g(x)$, 即

$$f(x)-g(x):=(a_0-b_0)+(a_1-b_1)x+\cdots=\sum_{i\geqslant 0}(a_i-b_i)x^i\in\mathbb{F}[x].$$

用加法表示, 就是

$$f(x)-g(x)=f(x)+(-g(x)).$$

本节最后定义多项式的乘法.

定义1.6 (乘法)**.** 将多项式$f(x)$和$g(x)$代数相乘, 得到的多项式记为$f(x)g(x)$, 即

$$\begin{aligned}f(x)g(x)=&(a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0)(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0)\\:=&a_nb_mx^{n+m}+(a_nb_{m-1}+a_{n-1}b_m)x^{n+m-1}+\cdots+a_0b_0\\=&\sum_{k=0}^{n+m}\left(\sum_{\substack{i+j=k\\ i,j\geqslant 0}}a_ib_j\right)x^k\in\mathbb{F}[x].\end{aligned}$$

事实上, 多项式相乘就是我们中学学过的去括号然后合并同次项. 例如, 我们有

$$(x-1)(x+2)=x^2+2x-x-2=x^2+x-2.$$

由定义, 我们可以得到下面非常有用的事实[7].

(i) $\mathbb{F}[x]$中无零因子, 即若$f(x)\neq 0, g(x)\neq 0$, 则有$f(x)g(x)\neq 0$. 进而$\mathbb{F}[x]$中(乘法)消去律成立, 即

$$f(x)g(x)=0,\ f(x)\neq 0\Longrightarrow g(x)=0;$$

(ii) $\deg(f(x)g(x))=\deg(f(x))+\deg(g(x))$.

还可以证明下面的性质.

(M1) (交换律) $f(x)g(x)=g(x)f(x)$;

(M2) (结合律) $(f(x)g(x))h(x)=f(x)(g(x)h(x))$;

(M3) (单位元存在) $1\cdot f(x)=f(x)$;

(D) (分配律) $(f(x)+g(x))h(x)=f(x)h(x)+g(x)h(x)$.

[7] 约定

$$(-\infty)+(-\infty)=-\infty,\quad (-\infty)+n=n+(-\infty)=-\infty,\quad (\forall n\in\mathbb{Z}_{\geqslant 0}).$$

于是, $\mathbb{F}[x]$在多项式的加法和乘法下成为有单位元的交换环[8], 称其为(一元)多项式环.

所有整数的集合$\mathbb{Z}$在数的加法和乘法下也成为有单位元的交换环, 称其为整数环. 可见, 如上定义的多项式运算至少符合多项式们的期望—像整数那样运算. 下节中, 我们继续看多项式如何学习整数的带余除法.

习题1.2

A1. 设$f(x)$和$g(x)$是非零多项式, 证明: $f(x)g(x) \neq 0$. 进而证明: $\mathbb{F}[x]$中(乘法)消去律成立.

A2. 设$f(x) \in \mathbb{F}[x]$, 如果存在$g(x) \in \mathbb{F}[x]$, 使得$f(x)g(x) = 1$, 则称$f(x)$是可逆元. 多项式环$\mathbb{F}[x]$中所有可逆元的全体做成的集合记为$\mathbb{F}[x]^{\times}$, 证明: $\mathbb{F}[x]^{\times} = \mathbb{F}^{\times}$.

B1. 设非零的实系数多项式$f(x)$(即系数都是实数的多项式)满足$f(f(x)) = (f(x))^k$, 其中k是给定的正整数. 求多项式$f(x)$.

B2. 设$k \geqslant 2$是正整数, $f(x)$是非零的正次数实系数多项式, 满足$f(x^k) = (f(x))^k$. 求多项式$f(x)$.

B3. 叙述有单位元的交换环的定义和环中的可逆元的定义. 将整数环$\mathbb{Z}$中所有可逆元做成的集合记为$\mathbb{Z}^{\times}$, 证明: $\mathbb{Z}^{\times} = \{\pm 1\}$.

B4. 设$\mathbb{F}$是数域, R是$\mathbb{F}$的一个扩环(不一定交换). 任意给定$t \in R$, 设t和$\mathbb{F}$中任意元乘法可交换. 定义映射$\sigma_t : \mathbb{F}[x] \to R$, 使得对任意$f(x) = \sum\limits_{i=0}^{n} a_i x^i \in \mathbb{F}[x]$, 有

$$\sigma_t(f(x)) = f(t) := \sum_{i=0}^{n} a_i t^i.$$

证明: σ_t保持加法和乘法运算, 即如果在$\mathbb{F}[x]$中, 有

$$f(x) + g(x) = h(x), \quad f(x)g(x) = p(x),$$

那么在R中有

$$f(t) + g(t) = h(t), \quad f(t)g(t) = p(t).$$

映射σ_t称为x用t代入.

§ 1.3 带余除法

前面像整数一样, 多项式们定义了它们之间的加法, 减法和乘法, 下面来看多项式如何学习整数世界中的除法. 我们知道, 通常情况下, 两个整数相除所得结果不一定是整数, 而应该是

$$\text{整数} \div \text{整数} = \text{商} + \text{余数}.$$

例如, 由算式

$$\begin{array}{r} 2 \\ 5 \overline{)\,13} \\ \underline{10} \\ 3 \end{array}$$

[8]参看本章附录.

得到13除以5的商是2, 余数是3. 注意到余数满足$3 < 5$, 所以不能再往下除了. 上面的除式在有理数的世界为

$$\frac{13}{5} = 2 + \frac{3}{5},$$

在整数世界则可表示为

$$13 = 2 \times 5 + 3.$$

一般的, 可以证明整数世界$\mathbb{Z}$中的带余除法定理: 对任意$a, b \in \mathbb{Z}$, 其中$b > 0$, 存在唯一的$q, r \in \mathbb{Z}$, 使得

$$a = qb + r, \quad 0 \leqslant r < b.$$

称q是b除a的商, 而r是余数.

多项式们也可以做类似的事, 例如, 由我们学过的多项式的长除法, 有下面的算式

$$
\begin{array}{rllll}
 & & x & -1 & \\ \cline{2-5}
x+1 \big) & x^2 & & & +1 \\
 & x^2 & +x & & \\ \cline{2-3}
 & & -x & +1 & \\
 & & -x & -1 & \\ \cline{3-5}
 & & & 2 &
\end{array}
$$

上面的结果在$\mathbb{F}[x]$中写成等式就是

$$x^2 + 1 = (x - 1)(x + 1) + 2,$$

即$x + 1$除$x^2 + 1$的商是$x - 1$, 而余式是2. 观察上面的长除法, 可以发现实际上每次都是把剩下的多项式的首项用$x + 1$"干掉", 而当首项次数不小于1时, 这很容易做到. 直到最后剩下的2的次数已经小于$x + 1$的次数, 我们不能再往下除了, 除法结束.

一般的, 设$f(x), g(x) \in \mathbb{F}[x]$且$g(x) \neq 0$. 当$\deg(f(x)) < \deg(g(x))$时, 此时不能除. 当$\deg(f(x)) \geqslant \deg(g(x))$时, 设

$$
\begin{aligned}
f(x) = & a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_n \neq 0, \\
g(x) = & b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0, \quad b_m \neq 0,
\end{aligned}
$$

则$n \geqslant m$. 如果$n = 0$, 则$m = 0$, 即$f(x)$和$g(x)$都为非零数, 此时做除法是平凡的. 设$n > 0$, 我们要把$f(x)$的首项$a_n x^n$干掉, 可以进行如下的长除法

$$
\begin{array}{rl}
 & \qquad\qquad \frac{a_n}{b_m} x^{n-m} \\ \cline{2-2}
b_m x^m + b_{m-1} x^{m-1} + \cdots \big) & \qquad a_n x^n + a_{n-1} x^{n-1} + \cdots \\
 & \qquad a_n x^n + \frac{a_n}{b_m} b_{m-1} x^{n-1} + \cdots \\ \cline{2-2}
 & f_1(x) = \left(a_{n-1} - \frac{a_n}{b_m} b_{m-1} \right) x^{n-1} + \cdots
\end{array}
$$

此时, 有$\deg(f_1(x)) \leqslant n - 1$. 如果$\deg(f_1(x)) < m$, 则不能除了, 除法结束; 否则继续用$g(x)$去除$f_1(x)$. 该过程有限步后必结束. 于是, 我们有关于多项式的带余除法定理.

定理1.1 (带余除法). 设$f(x), g(x) \in \mathbb{F}[x]$且$g(x) \neq 0$, 则存在唯一的$q(x), r(x) \in \mathbb{F}[x]$, 使得

$$f(x) = q(x)g(x) + r(x), \quad \deg(r(x)) < \deg(g(x)).$$

我们称$g(x)$除$f(x)$的商是$q(x)$, 而余式是$r(x)$.

☞ **证明.** 上面其实给出了存在性的证明(想法), 这里我们重新用数学语言写一下[9]. 当$\deg(f(x)) < \deg(g(x))$时, 由于

$$f(x) = 0 \cdot g(x) + f(x),$$

所以可取$q(x) = 0$, $r(x) = f(x)$. 下设$\deg(f(x)) \geqslant \deg(g(x))$, 我们对$\deg(f(x)) \geqslant 0$归纳证明[10]. 设

$$\begin{aligned} f(x) =& a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad a_n \neq 0, \\ g(x) =& b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, \quad b_m \neq 0, \end{aligned}$$

则$n \geqslant m$. 如果$n = 0$, 则$m = 0$. 此时有

$$f(x) = a_0 \neq 0, \quad g(x) = b_0 \neq 0.$$

由于

$$f(x) = a_0 = (a_0 b_0^{-1})b_0 + 0 = (a_0 b_0^{-1})g(x) + 0,$$

所以可取$g(x) = a_0 b_0^{-1}$, $r(x) = 0$.

下设$n > 0$, 且结论对次数小于n的多项式成立. 令

$$\begin{aligned} f_1(x) =& f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \\ =& \left(a_{n-1} - \frac{a_n}{b_m} b_{m-1}\right) x^{n-1} + \cdots \in \mathbb{F}[x], \end{aligned}$$

则

$$\deg(f_1(x)) \leqslant n - 1 < n.$$

当$\deg(f_1(x)) < \deg(g(x))$时, 由

$$f(x) = \frac{a_n}{b_m} x^{n-m} g(x) + f_1(x),$$

可取$q(x) = \frac{a_n}{b_m} x^{n-m}$和$r(x) = f_1(x)$. 当$\deg(f_1(x)) \geqslant \deg(g(x))$时, 对$f_1(x)$用归纳假设, 存在$q_1(x), r_1(x) \in \mathbb{F}[x]$, 使得

$$f_1(x) = q_1(x)g(x) + r_1(x), \quad \deg(r_1(x)) < \deg(g(x)).$$

于是有

$$f(x) = \left(\frac{a_n}{b_m} x^{n-m} + q_1(x)\right) g(x) + r_1(x),$$

所以可取$q(x) = \frac{a_n}{b_m} x^{n-m} + q_1(x)$, $r(x) = r_1(x)$. 存在性得证.

---

[9] 请读者体会如何把脑袋里的想法写成漂亮的数学证明.

[10] 设$P(n)$是与整数$n(n \geqslant n_0)$有关的一个命题. 则第一数学归纳法为: 如果

(i) $P(n_0)$成立;

(ii) 假设当$n = k(k \geqslant n_0)$时命题成立, 可证$P(k+1)$成立,

则对任意整数$n \geqslant n_0$, $P(n)$成立. 而第二数学归纳法为: 如果

(i) $P(n_0)$成立;

(ii)$'$ 假设当$n_0 \leqslant n \leqslant k(k \geqslant n_0)$时命题成立, 可证$P(k+1)$成立,

则对任意整数$n \geqslant n_0$, $P(n)$成立.

再证唯一性. 假设还有$\widetilde{q(x)}, \widetilde{r(x)} \in \mathbb{F}[x]$, 使得

$$f(x) = \widetilde{q(x)}g(x) + \widetilde{r(x)}, \quad \deg(\widetilde{r(x)}) < \deg(g(x)).$$

于是得到

$$q(x)g(x) + r(x) = \widetilde{q(x)}g(x) + \widetilde{r(x)},$$

即

$$(q(x) - \widetilde{q(x)})g(x) = \widetilde{r(x)} - r(x).$$

如果$q(x) \neq \widetilde{q(x)}$, 则有$\deg(q(x) - \widetilde{q(x)}) \geqslant 0$. 于是

$$\deg(\widetilde{r(x)} - r(x)) = \deg(q(x) - \widetilde{q(x)}) + \deg(g(x)) \geqslant \deg(g(x)).$$

但是由余式满足的条件, 我们又有

$$\deg(\widetilde{r(x)} - r(x)) \leqslant \max\{\deg(\widetilde{r(x)}), \deg(r(x))\} < \deg(g(x)),$$

这得到矛盾. 于是$q(x) = \widetilde{q(x)}$, 进而有$r(x) = \widetilde{r(x)}$. 唯一性得证. □

例1.3. 设$f(x) = x^3 + 2x^2 + 3x + 4$, $g(x) = 2x^2 + 3x - 4$, 用$g(x)$除$f(x)$, 求商$q(x)$和余式$r(x)$.

☞ **解.** 由于

$$\begin{array}{r}
\frac{1}{2}x+\frac{1}{4} \\
2x^2+3x-4\,\big)\overline{\;x^3+2x^2+3x+4} \\
x^3+\frac{3}{2}x^2-2x \quad \\ \hline
\frac{1}{2}x^2+5x+4 \\
\frac{1}{2}x^2+\frac{3}{4}x-1 \\ \hline
\frac{17}{4}x+5
\end{array}$$

所以$q(x) = \frac{1}{2}x + \frac{1}{4}$, 而$r(x) = \frac{17}{4}x + 5$. □

当除式$g(x) = x - c$是一次首一多项式时, 商和余式可以用一种特殊的方法得到. 事实上, 设

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = q(x)(x - c) + r,$$

其中$n \geqslant 1$, $r \in \mathbb{F}$, 而

$$q(x) = b_{n-1}x^{n-1} + \cdots + b_1 x + b_0,$$

则由于

$$\begin{aligned}
q(x)(x-c) + r &= (x-c)(b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) + r \\
&= b_{n-1}x^n + (b_{n-2} - cb_{n-1})x^{n-1} + \cdots + (b_0 - cb_1)x + (r - cb_0),
\end{aligned}$$

所以有

$$\begin{cases}
b_{n-1} = a_n, \\
b_{n-2} = a_{n-1} + cb_{n-1}, \\
\qquad \vdots \\
b_1 = a_2 + cb_2, \\
b_0 = a_1 + cb_1, \\
r = a_0 + cb_0.
\end{cases}$$

这个结果可以用下面的算式表示

$$\begin{array}{r|ccccccc} c & a_n & a_{n-1} & \cdots & a_i & \cdots & a_1 & a_0 \\ & & cb_{n-1} & \cdots & cb_i & \cdots & cb_1 & cb_0 \\ \hline & b_{n-1} & b_{n-2} & \cdots & b_{i-1} & \cdots & b_0 & r \end{array}$$

其中, 竖线右边第一行为$f(x)$的各项系数按照次数从高到低排列; 第三行为上两行的和; 第二行左边第一位为空白, 看成零, 而其它的每个数等于c乘以第三行对应位置往左边一位的数. 于是, 我们可以依次写出第三行第一位, 第二行第二位, 第三行第二位, 等等. 最后的商和余式就可以从第三行得到. 这种得到多项式除以首一一次多项式的商和余式的方法称为综合除法.

例1.4. 设$f(x)=2x^4-3x+4$, $g(x)=x+3$, 用$g(x)$除$f(x)$, 求商$q(x)$和余式$r(x)$.

☞ **解.** 由于

$$\begin{array}{r|rrrrr} -3 & 2 & 0 & 0 & -3 & 4 \\ & & -6 & 18 & -54 & 171 \\ \hline & 2 & -6 & 18 & -57 & 175 \end{array}$$

所以有$q(x)=2x^3-6x^2+18x-57$和$r(x)=175$. □

最后举一个带余除法应用的例子: 带余除法可以降低次数.

例1.5. 设多项式$f(x)=x^5-4x^4+3x^3-2x^2+x-1$, 计算$f(1+\sqrt{3})$, 这里$f(c)$表示将不定元$x$用$c$替代后所得值.

☞ **解.** 记$c=1+\sqrt{3}$, 则$(c-1)^2=3$, 即$c^2-2c-2=0$. 于是令$g(x)=x^2-2x-2$, 有$g(c)=0$. 做带余除法

$$\begin{array}{r|rrrrrr}
\multicolumn{1}{r}{} & & & x^3 & -2x^2 & +x & -4 \\ \cline{2-7}
x^2-2x-2 & x^5 & -4x^4 & +3x^3 & -2x^2 & +x & -1 \\
 & x^5 & -2x^4 & -2x^3 & & & \\ \cline{2-4}
 & & -2x^4 & +5x^3 & -2x^2 & +x & -1 \\
 & & -2x^4 & +4x^3 & +4x^2 & & \\ \cline{3-5}
 & & & x^3 & -6x^2 & +x & -1 \\
 & & & x^3 & -2x^2 & -2x & \\ \cline{4-6}
 & & & & -4x^2 & +3x & -1 \\
 & & & & -4x^2 & +8x & +8 \\ \cline{5-7}
 & & & & & -5x & -9
\end{array}$$

于是得到

$$f(x)=(x^3-2x^2+x-4)g(x)+(-5x-9).$$

上面等式中令$x=c$, 就得到

$$f(c)=(c^3-2c^2+c-4)g(c)+(-5c-9)=-5c-9=-14-5\sqrt{3}.$$

即所求为$-14-5\sqrt{3}$. □

本讲义中有许多定理, 读者应该理解每个定理的条件和结论, 反复咀嚼其意义, 并且抓住其本质的地方. 多项式的带余除法的一大利点是余式的次数严格小于除式的次数, 于是可以用带余除法降次. 而正如上例一样, 低次多项式总是相对容易处理的.

习题1.3

A1. 用$g(x)$除$f(x)$, 求商$q(x)$和余式$r(x)$, 其中
(1) $f(x)=2x^4+3x^3-5x+6$, $g(x)=x^2+x+1$;
(2) $f(x)=x^3-4x-2$, $g(x)=2x^2-3x+1$.
A2. 利用综合除法求$g(x)$除$f(x)$的商$q(x)$和余式$r(x)$, 其中
(1) $f(x)=3x^4-6x^2+1$, $g(x)=x+2$;
(2) $f(x)=x^3+x^2+1$, $g(x)=x+i$.
A3 设$f(x)=x^4-2x^3+ax^2-4x+b$, $g(x)=x^2+3x-1$. 如果$g(x)$除$f(x)$的余式是$-66x+24$, 求a和b.
A4. 将$f(x)=x^4+x^3+x^2+x+1$表示成$x-2$的方幂和的形式: $f(x)=b_0+b_1(x-2)+b_2(x-2)^2+b_3(x-2)^3+b_4(x-2)^4$.
A5. 设$f(x)=x^4-3x^3+2x^2-2x+1$, 记$f(1+\sqrt{2})=a+b\sqrt{2}$, 其中$a,b\in\mathbb{Z}$, 求$a$和$b$.
B1. 证明$\mathbb{Z}$中的带余除法, 即对任意$a,b\in\mathbb{Z}$, 其中$b>0$, 证明: 存在唯一的$q,r\in\mathbb{Z}$, 使得

$$a=qb+r,\quad 0\leqslant r<b.$$

§ 1.4 整除性

我们知道, 两个整数做除法不一定可以除的尽, 通常有非零的余数. 如果整数a除以b的余数为零(这等价于$a/b\in\mathbb{Z}$), 则称b整除a, 记为$b\mid a$; 否则称b不整除a, 记为$b\nmid a$. 例如, 我们有

$$2\mid 10,\quad 2\nmid 11.$$

由带余除法, 容易知道$b\mid a$的充分必要条件是存在$q\in\mathbb{Z}$, 使得$a=qb$.

多项式们当然也有整除的概念, 满足类似的性质.

定义1.7. *设$f(x),g(x)\in\mathbb{F}[x]$.*

(1) *如果存在$q(x)\in\mathbb{F}[x]$, 使得$f(x)=q(x)g(x)$, 则称$g(x)$(在$\mathbb{F}[x]$中)整除$f(x)$, 记作$g(x)\mid f(x)$; 否则称$g(x)$不整除$f(x)$, 记为$g(x)\nmid f(x)$;*

(2) *当$g(x)$整除$f(x)$时, 称$g(x)$是$f(x)$的一个因式, 而$f(x)$是$g(x)$的一个倍式.*

由定义, 当$g(x)=0$时, $g(x)$的倍式必为0. 而且可以得到

命题1.2. *设$f(x),g(x)\in\mathbb{F}[x]$, 且$g(x)\neq 0$, 则*

$$g(x)\mid f(x)\Longleftrightarrow g(x)\text{ 除 }f(x)\text{ 的余式为 }0.$$

☞ **证明.** “$\Longleftarrow$” 由条件, 存在商$q(x)\in\mathbb{F}[x]$, 使得

$$f(x)=q(x)g(x)+0=q(x)g(x).$$

于是$g(x)\mid f(x)$.

“$\Longrightarrow$” 由于$g(x) \mid f(x)$, 所以存在$q(x) \in \mathbb{F}[x]$, 使得

$$f(x) = q(x)g(x) = q(x)g(x) + 0.$$

由带余除法中的唯一性, 得余式$r(x) = 0$. □

例1.6. 证明: $x - 1 \mid x^3 - 1$, 而$x + 1 \nmid x^3 - 1$.

☞ **证明.** 由综合除法

$$\begin{array}{r|rrrr} 1 & 1 & 0 & 0 & -1 \\ & & 1 & 1 & 1 \\ \hline & 1 & 1 & 1 & 0 \end{array} \qquad \begin{array}{r|rrrr} -1 & 1 & 0 & 0 & -1 \\ & & -1 & 1 & -1 \\ \hline & 1 & -1 & 1 & -2 \end{array}$$

得

$$x^3 - 1 = (x-1)(x^2+x+1), \quad x^3 - 1 = (x+1)(x^2-x+1) - 2,$$

所以$x - 1 \mid x^3 - 1$, 而$x + 1 \nmid x^3 - 1$. □

例1.7. 设$f(x) = x^4 - x^3 + 4x^2 + ax + b$, $g(x) = x^2 + 2x - 3$. 求$g(x) \mid f(x)$的充分必要条件.

☞ **解.** 由

$$\begin{array}{r}
x^2 \; -3x + 13 \\
x^2 + 2x - 3 \overline{\big)\; x^4 - x^3 + 4x^2 + ax + b} \\
\underline{x^4 + 2x^2 \; -3x^2} \qquad\qquad \\
-3x^3 \; +7x^2 \; +ax + b \\
\underline{-3x^3 \; -6x^2 \; +9x \qquad} \\
13x^2 \; +(a-9)x + b \\
\underline{13x^2 \quad +26x \quad -39} \\
(a-35)x + (b+39)
\end{array}$$

得$g(x)$去除$f(x)$的余式是$(a-35)x + (b+39)$. 于是

$$g(x) \mid f(x) \Longleftrightarrow (a-35)x + (b+39) = 0 \Longleftrightarrow a = 35, b = -39.$$

解毕. □

多项式的整除满足下面简单的性质.

命题1.3. 设$f(x), g(x), h(x), f_1(x), \ldots, f_r(x) \in \mathbb{F}[x]$, 其中$r \in \mathbb{N}$. 下面成立:

(1) $g(x) \mid g(x)$, $g(x) \mid 0$, $a \mid g(x)$ $(\forall a \in \mathbb{F}^\times)$;

(2) $g(x) \mid f(x)$且$f(x) \neq 0 \Longrightarrow \deg(g(x)) \leqslant \deg(f(x))$;

(3) $f(x) \mid g(x)$且$g(x) \mid f(x) \Longleftrightarrow \exists c \in \mathbb{F}^\times$, 使得$f(x) = cg(x)$.

此时, 称$f(x)$和$g(x)$相伴, 记为$f(x) \sim g(x)$;

(4) (传递性) $f(x) \mid g(x)$, $g(x) \mid h(x) \Longrightarrow f(x) \mid h(x)$;

(5) 如果$g(x) \mid f_i(x)$, $i=1,2,\ldots,r$, 则对任意$u_1(x),u_2(x),\ldots,u_r(x)\in\mathbb{F}[x]$, 有

$$g(x) \mid (u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_r(x)f_r(x)).$$

☞ **证明.** (3) "$\Longleftarrow$" 如果$f(x)=cg(x)$, $c\neq 0$, 则$g(x)=c^{-1}f(x)$. 于是$g(x)\mid f(x)$, $f(x)\mid g(x)$.
"$\Longrightarrow$" 设$f(x)\mid g(x)$且$g(x)\mid f(x)$, 则存在$q_1(x),q_2(x)\in\mathbb{F}[x]$, 使得

$$g(x)=q_1(x)f(x),\quad f(x)=q_2(x)g(x).$$

于是有

$$f(x)=q_2(x)q_1(x)f(x)\Longrightarrow (q_2(x)q_1(x)-1)f(x)=0.$$

由此得到

$$f(x)=0\quad \text{或者}\quad q_2(x)q_1(x)=1.$$

如果$f(x)=0$, 则$g(x)=0$, 此时$f(x)=1\cdot g(x)$; 如果$q_2(x)q_1(x)=1$, 由

$$\deg(q_1(x))+\deg(q_2(x))=\deg 1=0$$

得$\deg(q_1(x))=\deg(q_2(x))=0$. 于是$q_2(x)\in\mathbb{F}^{\times}$.
(4) 由条件, 存在$q_1(x),q_2(x)\in\mathbb{F}[x]$, 使得

$$g(x)=q_1(x)f(x),\quad h(x)=q_2(x)g(x).$$

于是$h(x)=q_2(x)q_1(x)f(x)$, 这得到$f(x)\mid h(x)$.
(5) 由条件, 存在$q_i(x)\in\mathbb{F}[x]$, 使得

$$f_i(x)=q_i(x)g(x),\quad i=1,2,\ldots,r.$$

所以有

$$u_1(x)f_1(x)+\cdots+u_r(x)f_r(x)=(u_1(x)q_1(x)+\cdots+u_r(x)q_r(x))g(x),$$

这得到$g(x)\mid(u_1(x)f_1(x)+\cdots+u_r(x)f_r(x))$. □

由上面的性质, 在讨论整除问题时, 常常可以只考虑首一多项式.

上面关于整除性的讨论都是在$\mathbb{F}[x]$中进行的. 事实上, 多项式之间的整除关系不会因为系数域$\mathbb{F}$的扩大而改变.

例1.8. 设$\mathbb{F}_1$和$\mathbb{F}_2$都是数域, 且$\mathbb{F}_1\subset\mathbb{F}_2$. 设$f(x),g(x)\in\mathbb{F}_1[x](\subset\mathbb{F}_2[x])$, 则

$$g(x)\mid f(x)\ in\ \mathbb{F}_1[x]\Longleftrightarrow g(x)\mid f(x)\ in\ \mathbb{F}_2[x].$$

☞ **证明.** "$\Longrightarrow$" 显然.
"$\Longleftarrow$" 设在$\mathbb{F}_2[x]$中, 有$g(x)\mid f(x)$. 如果$g(x)=0$, 则$f(x)=0$. 此时, 在$\mathbb{F}_1[x]$中也有$g(x)\mid f(x)$.

下设$g(x)\neq 0$, 在$\mathbb{F}_1[x]$中做带余除法, 则存在$q(x),r(x)\in\mathbb{F}_1[x]$, 使得

$$f(x)=q(x)g(x)+r(x),\quad \deg(r(x))<\deg(g(x)).$$

由于$f(x),g(x),q(x),r(x)\in\mathbb{F}_2[x]$, 所以上式也可看成$\mathbb{F}_2[x]$中的带余除法. 但在$\mathbb{F}_2[x]$中有$g(x)\mid f(x)$, 这得到$r(x)=0$. 从而在$\mathbb{F}_1[x]$中有$g(x)\mid f(x)$. □

习题1.4

A1. 求$g(x) \mid f(x)$的充分必要条件, 其中
(1) $f(x) = x^4 - 3x^3 + ax + b$, $g(x) = x^2 - 3x - 1$;
(2) $f(x) = x^4 + 3x^2 + ax + b$, $g(x) = x^2 - 2ax + 2$.
A2. 设$f(x) = x^3 + px^2 + qx + r$, $g(x) = ax^2 + bx + c$, 且$abc \neq 0$, 证明: $g(x) \mid f(x)$的充分必要条件是
$$\frac{ap-b}{a} = \frac{aq-c}{b} = \frac{ar}{c}.$$
A3. 设$x^2 + mx - 1$整除$x^3 + px + q$, 求m, p, q满足的条件.
A4. (1) p, q, m满足什么条件时, $(x-m)^2$整除$x^3 + px + q$;
(2) p, q满足什么条件时, 存在m使得$(x-m)^2$整除$x^3 + px + q$.
B1. 设$a \in \mathbb{F}^\times$, $d, n \in \mathbb{N}$, 证明: 在域$\mathbb{F}$上, $x^d - a^d \mid x^n - a^n$当且仅当$d \mid n$.
B2. 设$d, n \in \mathbb{N}$, 证明:
$$a^d - 1 \mid a^n - 1 \quad (\forall a \in \mathbb{Z}) \Longleftrightarrow d \mid n.$$
B3. 设整数$m, n, l \geqslant 0$, 证明: $x^2 + x + 1 \mid x^{3m} + x^{3n+1} + x^{3l+2}$.
B4. 写出并证明命题1.3在整数世界对应的结论.

§ 1.5 最大公因式

前面我们定义了多项式的因式和倍式, 如果有几个多项式在一起, 自然可以考虑它们的公共因式和公共倍式. 我们处理两个多项式的公共因式, 而把多个多项式的公共因式和公共倍式作为习题留给读者(请注意, 这些在后面课程中都要用到).

先看整数世界如何研究两个整数的公共因子. 设$a, b \in \mathbb{Z}$, 它们的公共因子称为公因子. 通常, 公因子有很多, 如何把握它们呢? 看一个简单的例子: 设$a = 8$和$b = 12$, 则它们的公因子有
$$\pm 1, \quad \pm 2, \quad \pm 4,$$
这些公因子恰好是4的所有因子, 而4则是a和b的所有公因子中最大者. 研究发现这是一般规律. 设整数a和b不全为零, 则它们的公因子只有有限多个(非零整数的因子只有有限多个). 此时, 存在a和b的最大公因子, 记为(a, b). 可以知道(a, b)是正整数, 且对$d \in \mathbb{N}$, 有$d = (a, b)$的充分必要条件是

(i) $d \mid a$, $d \mid b$; (d是公因子)

(ii) $\forall d_1 \in \mathbb{Z}$, $d_1 \mid a$, $d_2 \mid b \Longrightarrow d_1 \mid d$. ($d$是任意公因子的倍数)

由此可以看出, 公因子本质上由最大公因子所控制: 所有的公因子都是最大公因子的因子.

类比于整数世界, 多项式们当然希望可以得到类似的结果, 即两个多项式的公因式可以由所谓的最大公因式决定. 那么如何定义两个多项式的最大公因式呢? 一种方法是类比于整数的世界, 定义两个多项式的公因式中的"最大者"为最大公因式. 这就需要比较多项式的大小, 我们发现次数是比较好的一个选择, 即定义次数最大的公因式为最大公因式. 然后证明所有的公因式都是最大公因式的因式, 即多项式世界也有上面类似的事实成立. 我们将这种定义方法留给读者, 而采取另一种等价的定义方法, 即用希望最大公因式能满足的性质来定义最大公因式. 请大家注意这种处理方法: 我们想

定义对象X, 使其满足性质P; 那么我们就定义满足性质P的所有对象就是我们的X, 然后再证明存在性.

定义1.8. 设$f(x), g(x) \in \mathbb{F}[x]$.

(1) 称$\varphi(x) \in \mathbb{F}[x]$是$f(x)$和$g(x)$的一个公因式, 如果满足

$$\varphi(x) \mid f(x), \quad \varphi(x) \mid g(x);$$

(2) 设$f(x), g(x)$不全为零, $d(x) \in \mathbb{F}[x]$. 称$d(x)$是$f(x)$和$g(x)$的一个最大公因式, 如果满足

(i) $d(x) \mid f(x)$, $d(x) \mid g(x)$;

(ii) $\forall \varphi(x) \in \mathbb{F}[x]$, $\varphi(x) \mid f(x)$, $\varphi(x) \mid g(x) \Longrightarrow \varphi(x) \mid d(x)$.

注意到不全为零的$f(x)$和$g(x)$的公因式一定非零, 而且如下可证: 如果存在最大公因式, 那么最大公因式在相伴意义下唯一, 且确实是所有公因式中“最大”的.

引理1.4. 设$f(x), g(x) \in \mathbb{F}[x]$不全为零, $d(x) \in \mathbb{F}[x]$.

(1) 如果$d(x)$是$f(x)$和$g(x)$的最大公因式, $d_1(x) \in \mathbb{F}[x]$, 则

$$d_1(x)\text{也为最大公因式} \Longleftrightarrow d_1(x) \sim d(x);$$

(2) 若$f(x)$和$g(x)$有最大公因式, 则$d(x)$为最大公因式的充分且必要条件是

(i) $d(x) \mid f(x)$, $d(x) \mid g(x)$;

(ii)′ $\forall \varphi(x) \in \mathbb{F}[x]$, $\varphi(x) \mid f(x)$, $\varphi(x) \mid g(x) \Longrightarrow \deg(\varphi(x)) \leqslant \deg(d(x))$.

证明. (1) “$\Longrightarrow$” 设$d_1(x)$也是最大公因式, 则由定义得到

$$d_1(x) \mid d(x), \quad d(x) \mid d_1(x),$$

即它们相伴.

“$\Longleftarrow$” 设$d_1(x) \sim d(x)$, 则

$$d_1(x) \mid d(x), \quad d(x) \mid d_1(x).$$

于是由整除的传递性, 容易知道$d_1(x)$满足最大公因式定义中的(i)和(ii), 即$d_1(x)$也是最大公因式.

(2) 假设$d(x)$是最大公因式, 则显然(i)成立, 进而对$f(x)$和$g(x)$的任意公因式$\varphi(x)$, 有$\varphi(x) \mid d(x)$. 因为$d(x) \neq 0$, 于是$\deg(d(x)) \geqslant \deg(\varphi(x))$.

反之, 假设$d(x)$满足(i)和(ii)′. 任取$f(x)$和$g(x)$的一个最大公因式$d_1(x)$. 由(i), $d(x)$是公因式, 因此有$d(x) \mid d_1(x)$. 这得到

$$d_1(x) = q(x)d(x), \quad (\exists q(x) \in \mathbb{F}[x]),$$

进而有

$$\deg(d_1(x)) = \deg(q(x)) + \deg(d(x)).$$

由于$d_1(x)$也是公因式, 于是由(ii)′, 有

$$\deg(d(x)) \geqslant \deg(d_1(x)).$$

而公因式$d(x), d_1(x) \neq 0$, 进而$q(x) \neq 0$, 也就有

$$\deg(d(x)), \deg(d_1(x)), \deg(q(x)) \geqslant 0.$$

于是一定有

$$\deg(q(x)) = 0,$$

即$q(x) \in \mathbb{F}^{\times}$. 由(1)得$d(x)$是最大公因式. □

上面引理的(1)表明, 最大公因式一定是两两相伴的, 于是本质上可以考虑首一的最大公因式.

定义1.9. (1) 设$f(x), g(x) \in \mathbb{F}[x]$不全为零, 定义

$$(f(x), g(x)) := \text{“首一最大公因式”};$$

(2) 约定0和0的最大公因式为0, 且$(0,0) := 0$.

下面讨论最大公因式的存在性. 先看零多项式和其它多项式的最大公因式.

例1.9. 设$f(x) \in \mathbb{F}[x]$, 求$(f(x), 0)$.

☞ **解.** 当$f(x) = 0$时, $(f(x), 0) = 0$. 下设$f(x) \neq 0$. 由于$f(x) \mid f(x)$, $f(x) \mid 0$, 且$f(x)$和0的任意公因式$\varphi(x)$满足$\varphi(x) \mid f(x)$, 所以$f(x)$是$f(x)$和0的最大公因式. 这得到

$$(f(x), 0) = a^{-1} f(x),$$

其中a是$f(x)$的首项系数. □

下设$f(x), g(x) \in \mathbb{F}[x]$, 我们要求$f(x)$和$g(x)$的最大公因式. 由上面的例子, 可以不妨设$f(x) \neq 0$, $g(x) \neq 0$. 我们要在$f(x)$和$g(x)$的所有公因式中找最大者, 这似乎是一个艰巨的任务. 一个自然的想法是能不能将问题转化为上面的例子, 即使得$f(x)$和$g(x)$中有一个为零. 如何把一个多项式搞成零呢? 还好我们知道的不是很多, 因式和除法有关, 而除法我们只知道带余除法, 带余除法可以降次, 而零多项式是次数最低者. 设$f(x) = q(x)g(x) + r(x)$, 其中$q(x), r(x) \in \mathbb{F}[x]$, 则$r(x) = f(x) - q(x)g(x)$. 于是对任意$\varphi(x) \in \mathbb{F}[x]$, 容易得到

$$\varphi(x) \mid f(x),\ \varphi(x) \mid g(x) \Longleftrightarrow \varphi(x) \mid g(x),\ \varphi(x) \mid r(x).$$

即

引理1.5. 如果在$\mathbb{F}[x]$中有等式

$$f(x) = q(x)g(x) + r(x),$$

则$f(x), g(x)$和$g(x), r(x)$有相同的公因式集.

上面的引理告诉我们, 要研究$f(x)$和$g(x)$的公因式集, 只要研究$g(x)$和$r(x)$的公因式集, 而后者比前者简单—多项式$r(x)$的次数小于$g(x)$的次数. 于是, 反复用这个转化办法, 每次我们都可以讨论更低次数的多项式的公因式集; 有限步后, 必然有一个多项式为零多项式, 此时就回到了上面的例子, 自然就求出了最大公因式.

具体地，设$f(x),g(x)\in\mathbb{F}[x]$，我们来求$f(x)$和$g(x)$的(一个)最大公因式$d(x)$. 如果$f(x)=0$，则可取$d(x)=g(x)$；如果$g(x)=0$，则$d(x)=f(x)$. 下设$f(x)\neq 0$且$g(x)\neq 0$. 做带余除法:

$$\begin{cases} f(x)=q_1(x)g(x)+r_1(x), & \deg(r_1(x))<\deg(g(x)),\\ g(x)=q_2(x)r_1(x)+r_2(x), & \deg(r_2(x))<\deg(r_1(x)),\\ r_1(x)=q_3(x)r_2(x)+r_3(x), & \deg(r_3(x))<\deg(r_2(x)),\\ \qquad\vdots & \qquad\vdots\\ r_{s-3}(x)=q_{s-1}(x)r_{s-2}(x)+r_{s-1}(x), & \deg(r_{s-1}(x))<\deg(r_{s-2}(x)),\\ r_{s-2}(x)=q_s(x)r_{s-1}(x)+r_s(x), & \deg(r_s(x))<\deg(r_{s-1}(x)),\\ r_{s-1}(x)=q_{s+1}(x)r_s(x)+0. \end{cases}$$

由余式的次数所满足的不等式，该过程有限步后必停止. 现在反复利用引理1.5，可得下面的多项式对都有相同的公因式集:

$$\begin{array}{l} f(x),g(x);\quad g(x),r_1(x);\quad r_1(x),r_2(x);\quad r_2(x),r_3(x);\quad \cdots\\ r_{s-3}(x),r_{s-2}(x);\quad r_{s-2}(x),r_{s-1}(x);\quad r_{s-1}(x),r_s(x);\quad r_s(x),0. \end{array}$$

特别地，$f(x),g(x)$与$r_s(x),0$有相同的公因式集. 而$r_s(x),0$有最大公因式$r_s(x)$，所以$r_s(x)$也是$f(x)$，$g(x)$的最大公因式，即可取$d(x)=r_s(x)$.

这种通过不断做带余除法直到整除停止得到最大公因式的方法称为Euclid辗转相除法.

定理1.6 (存在性和辗转相除法). *设$f(x),g(x)\in\mathbb{F}[x]$，则存在$f(x)$和$g(x)$的最大公因式$d(x)\in\mathbb{F}[x]$，且存在$u(x)$，$v(x)\in\mathbb{F}[x]$，使得*

$$d(x)=u(x)f(x)+v(x)g(x).$$

*上面的等式称为***Bezout等式**[11].

证明. 存在性已经证明，下面证明Bezout等式成立. 由于最大公因式两两相伴，所以只要对任意一个最大公式证明Bezout等式成立即可. 我们取最大公因式为上面存在性讨论中的$d(x)$.

当$f(x)=0$时，$d(x)=g(x)$，此时有等式

$$g(x)=0\cdot f(x)+1\cdot g(x).$$

当$g(x)=0$时，$d(x)=f(x)$，有等式

$$f(x)=1\cdot f(x)+0\cdot g(x).$$

当$f(x),g(x)\neq 0$时，$d(x)=r_s(x)$. 上面的带余除法等式可以改写为

$$\begin{cases} r_1(x)=f(x)-q_1(x)g(x),\\ r_2(x)=g(x)-q_2(x)r_1(x),\\ r_3(x)=r_1(x)-q_3(x)r_2(x),\\ \qquad\vdots\\ r_{s-1}(x)=r_{s-3}(x)-q_{s-1}(x)r_{s-2}(x),\\ r_s(x)=r_{s-2}(x)-q_s(x)r_{s-1}(x). \end{cases}$$

[11] 用交换环$\mathbb{F}[x]$的语言，Bezout等式本质上说明由$f(x)$和$g(x)$生成的$\mathbb{F}[x]$的理想是主理想，且有生成元$d(x)$.

将上面第$s-1$式代入第s式, 可以消去$r_{s-1}(x)$, 得到

$$r_s(x) = u_{s-1}(x)r_{s-2}(x) + v_{s-1}(x)r_{s-3}(x), \quad (\exists u_{s-1}(x), v_{s-1}(x) \in \mathbb{F}[x]).$$

将第$s-2$式代入上式, 可以消去$r_{s-2}(x)$, 得到

$$r_s(x) = u_{s-2}(x)r_{s-3}(x) + v_{s-2}(x)r_{s-4}(x), \quad (\exists u_{s-2}(x), v_{s-2}(x) \in \mathbb{F}[x]).$$

继续代入, 直到将第1式代入, 可以消去$r_1(x)$, 最后得到

$$r_s(x) = u_1(x)f(x) + v_1(x)g(x), \quad (\exists u_1(x), v_1(x) \in \mathbb{F}[x]).$$

取$u(x) = u_1(x)$, $v(x) = v_1(x)$即可.

学过λ矩阵理论后, 也可以如下证明(得到)Bezout等式. 将带余除法改写成矩阵形式, 得到

$$\begin{aligned}
\begin{pmatrix} f(x) \\ g(x) \end{pmatrix} &= \begin{pmatrix} q_1(x) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} g(x) \\ r_1(x) \end{pmatrix}, \\
\begin{pmatrix} g(x) \\ r_1(x) \end{pmatrix} &= \begin{pmatrix} q_2(x) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix}, \\
\begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix} &= \begin{pmatrix} q_3(x) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_2(x) \\ r_3(x) \end{pmatrix}, \\
&\vdots \\
\begin{pmatrix} r_{s-2}(x) \\ r_{s-1}(x) \end{pmatrix} &= \begin{pmatrix} q_s(x) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_{s-1}(x) \\ r_s(x) \end{pmatrix}.
\end{aligned}$$

将下一个矩阵等式代入前一个, 得到

$$\begin{pmatrix} f(x) \\ g(x) \end{pmatrix} = A(x) \begin{pmatrix} r_{s-1}(x) \\ r_s(x) \end{pmatrix},$$

其中

$$A(x) = \begin{pmatrix} q_1(x) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} q_2(x) & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} q_s(x) & 1 \\ 1 & 0 \end{pmatrix}.$$

由于$\det(A(x)) = (-1)^s$, 所以$A(x)$可逆, 这得到

$$\begin{pmatrix} r_{s-1}(x) \\ r_s(x) \end{pmatrix} = A^{-1}(x) \begin{pmatrix} f(x) \\ g(x) \end{pmatrix},$$

进而有

$$r_s(x) = (0,1)A^{-1}(x) \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}.$$

令

$$(u(x), v(x)) = (0,1)A^{-1}(x)$$

即可. □

例1.10. 设$f(x) = x^3 + 5x^2 + 7x + 3$, $g(x) = x^3 - 7x + 6$, 求$(f(x), g(x))$. 并求$u(x), v(x)$, 使得

$$(f(x), g(x)) = u(x)f(x) + v(x)g(x).$$

☞ **解.** 做辗转相除, 可列下面算式

$q_1(x)=1$	$f(x)=x^3+5x^2+7x+3$	$g(x)=x^3-7x+6$	$q_2(x)=\frac{1}{5}x-\frac{14}{25}$
	x^3-7x+6	$x^3+\frac{14}{5}x^2-\frac{3}{5}x$	
$q_3(x)=\frac{125}{36}x-\frac{25}{36}$	$r_1(x)=\quad 5x^2+14x-3$	$-\frac{14}{5}x^2-\frac{32}{5}x+6$	
	$5x^2+15x$	$-\frac{14}{5}x^2-\frac{196}{25}x+\frac{42}{25}$	
	$-x-3$	$r_2(x)=\quad \frac{36}{25}x+\frac{108}{25}$	
	$-x-3$		
	0		

这等价于等式

$$
\begin{aligned}
&f(x)=1\cdot g(x)+(5x^2+14x-3),\\
&g(x)=\left(\frac{1}{5}x-\frac{14}{25}\right)(5x^2+14x-3)+\left(\frac{36}{25}x+\frac{108}{25}\right),\\
&5x^2+14x-3=\left(\frac{125}{36}x-\frac{25}{36}\right)\left(\frac{36}{25}x+\frac{108}{25}\right).
\end{aligned}
$$

由于

$$\frac{36}{25}x+\frac{108}{25}=\frac{36}{25}(x+3),$$

所以有$(f(x),g(x))=x+3$. 而

$$
\begin{aligned}
\frac{36}{25}x+\frac{108}{25}=&g(x)-\left(\frac{1}{5}x-\frac{14}{25}\right)(5x^2+14x-3)\\
=&g(x)-\left(\frac{1}{5}x-\frac{14}{25}\right)(f(x)-g(x))\\
=&-\left(\frac{1}{5}x-\frac{14}{25}\right)f(x)+\left(\frac{1}{5}x+\frac{11}{25}\right)g(x),
\end{aligned}
$$

所以有

$$(f(x),g(x))=x+3=\left(-\frac{5}{36}x+\frac{7}{18}\right)f(x)+\left(\frac{5}{36}x+\frac{11}{36}\right)g(x).$$

于是可取

$$u(x)=-\frac{5}{36}x+\frac{7}{18},\quad v(x)=\frac{5}{36}x+\frac{11}{36}.$$

请注意这里$u(x),v(x)$不唯一. □

最后, 多项式的最大公因式不因系数域的扩大而改变.

习题1.5

A1. 求最大公因式$(f(x),g(x))$, 其中

(1) $f(x)=x^4+x^3-3x^2-4x-1$, $g(x)=x^3+x^2-x-1$;

(2) $f(x)=x^4-4x^3+1$, $g(x)=x^3-3x^2+1$.

A2. 求$(f(x),g(x))$, 并求多项式$u(x)$和$v(x)$, 使得$u(x)f(x)+v(x)g(x)=(f(x),g(x))$, 其中

(1) $f(x)=x^4-x^3-4x^2+4x+1$, $g(x)=x^2-x-1$;

(2) $f(x)=4x^4-2x^3-16x^2+5x+9$, $g(x)=2x^3-x^2-5x+4$.

A3. 设$d(x)$是$f(x)$和$g(x)$的公因式, 且存在多项式$u(x)$和$v(x)$, 使得$u(x)f(x)+v(x)g(x)=d(x)$, 证明: $d(x)$是$f(x)$和$g(x)$的一个最大公因式. 如果去掉$d(x)$是$f(x)$和$g(x)$的公因式的条件, 结论是否还成立? 说明理由.

A4. 设$f(x), g(x), h(x) \in \mathbb{F}[x]$, 其中$h(x)$是首一的, 证明: $(f(x)h(x), g(x)h(x)) = (f(x), g(x))h(x)$.
A5. 设$f_1(x)$, $f_2(x)$, …, $f_s(x) \in \mathbb{F}[x]$不全为零, 如果$d(x) \in \mathbb{F}[x]$是$f_1(x)$, $f_2(x)$, …, $f_s(x)$的一个公因式, 且$f_1(x)$, $f_2(x)$, …, $f_s(x)$的任一公因式都整除$d(x)$, 则称$d(x)$是$f_1(x)$, $f_2(x)$, …, $f_s(x)$的一个最大公因式. 约定全为零的s个多项式的最大公因式只有零多项式. 证明: 如果$f_1(x)$, $f_2(x)$, …, $f_{s-1}(x)$的最大公因式存在, 则$f_1(x)$, $f_2(x)$, …, $f_s(x)$的最大公因式也存在, 而且

$$(f_1(x), f_2(x), \ldots, f_{s-1}(x), f_s(x)) = ((f_1(x), f_2(x), \ldots, f_{s-1}(x)), f_s(x)).$$

这里, $(f_1(x), f_2(x), \ldots, f_s(x))$表示$f_1(x)$, $f_2(x)$, …, $f_s(x)$的首一最大公因式(当$f_1(x) = \cdots = f_s(x) = 0$时, 约定$(f_1(x), f_2(x), \ldots, f_s(x)) = 0$). 再利用所证表达式证明: 存在多项式$u_1(x), u_2(x), \ldots, u_s(x)$, 使得

$$u_1(x)f_1(x) + u_2(x)f_2(x) + \cdots + u_s(x)f_s(x) = (f_1(x), f_2(x), \ldots, f_s(x)).$$

B1. 设$f(x), g(x) \in \mathbb{F}[x]$, 如果多项式$m(x) \in \mathbb{F}[x]$满足条件

(i) $m(x)$是$f(x)$和$g(x)$的公倍式;

(ii) $f(x)$与$g(x)$的任意公倍式都是$m(x)$的倍式,

则称$m(x)$为$f(x)$与$g(x)$的一个最小公倍式.
(1) 证明: 零多项式是$f(x)$和$g(x)$的最小公倍式的充分必要条件是$f(x)$和$g(x)$中至少有一个是零多项式;
(2) 证明: $\mathbb{F}[x]$中任意两个多项式都有最小公倍式, 并且$f(x)$和$g(x)$的最小公倍式在相伴意义下唯一;
(3) 当$f(x)$和$g(x)$都不为零时, 用$[f(x), g(x)]$表示它们的首一最小公倍式(如果$f(x) = 0$或者$g(x) = 0$, 约定$[f(x), g(x)] = 0$). 证明: 如果非零多项式$f(x)$和$g(x)$都是首一的, 则

$$f(x)g(x) = [f(x), g(x)](f(x), g(x)).$$

B2. 设$f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$, 如果$m(x)$是$f_1(x)$, $f_2(x)$, …, $f_s(x)$的公倍式, 且整除$f_1(x)$, $f_2(x)$, …, $f_s(x)$的任意公倍式, 则称$m(x)$是$f_1(x)$, $f_2(x)$, …, $f_s(x)$的一个最小公倍式. 证明: 如果$f_1(x)$, $f_2(x)$, …, $f_{s-1}(x)$的最小公倍式存在, 则$f_1(x)$, $f_2(x)$, …, $f_s(x)$的最小公倍式也存在, 而且

$$[f_1(x), f_2(x), \ldots, f_{s-1}(x), f_s(x)] = [[f_1(x), f_2(x), \ldots, f_{s-1}(x)], f_s(x)].$$

这里, $[f_1(x), f_2(x), \ldots, f_s(x)]$表示$f_1(x)$, $f_2(x)$, …, $f_s(x)$的首一最小公倍式(当$f_1(x)$, $f_2(x)$, …, $f_s(x)$有零多项式时, 约定$[f_1(x), f_2(x), \ldots, f_s(x)] = 0$).
B3. 设a, b是不全为零的两个整数, $d, m \in \mathbb{N}$, 证明:
(1) $d = (a, b)$的充分必要条件是

(i) $d \mid a$, $d \mid b$;

(ii) $\forall d_1 \in \mathbb{Z}$, $d_1 \mid a$, $d_2 \mid b \Longrightarrow d_1 \mid d$;

(2) 存在$u, v \in \mathbb{Z}$, 使得

$$ua + vb = (a, b);$$

(3) $m(a, b) = (ma, mb)$.

B4. 设m和n是正整数, 证明:

$$(x^m - 1, x^n - 1) = x^{(m,n)} - 1.$$

§ 1.6 多项式互素

设有两个多项式相遇, 我们来看它们在整除下的关系. 如果某个多项式整除另一个多项式, 则这两个多项式有密切的关系. 除了这个极端关系外, 还可能发生另一种极端情形—它们极端不互相整除. 此时, 它们的公因式只有非零数, 即最大公因式为1. 在这种情形下, 可以认为它们没有关系, 我们称这种没有关系的关系为互素. 可以想象, 没有关系的多项式在讨论除法时, 它们互相不干扰. 在整数世界有类似的现象, 本节也是整数世界的互素讨论在多项式世界的再现.

定义1.10. *设$f(x), g(x) \in \mathbb{F}[x]$, 如果$(f(x), g(x)) = 1$, 则称$f(x)$与$g(x)$互素.*

由定义可以得到, $f(x)$和$g(x)$互素的充分必要条件是$f(x)$和$g(x)$的公因式只有$\mathbb{F}$中非零数. 一方面, 我们可以通过辗转相除法求出两个多项式的最大公因式是非零常数得到互素, 另一方面也可以从Bezout等式得到互素, 而后者常常非常方便.

定理1.7. *设$f(x), g(x) \in \mathbb{F}[x]$, 则$f(x)$与$g(x)$互素的充分必要条件是存在$u(x), v(x) \in \mathbb{F}[x]$, 使得*

$$u(x)f(x) + v(x)g(x) = 1.$$

☞ **证明.** 必要性已证(定理1.6), 于是只要证明充分性. 假设

$$u(x)f(x) + v(x)g(x) = 1, \quad (\exists u(x), v(x) \in \mathbb{F}[x]).$$

记$d(x) = (f(x), g(x))$, 则由$d(x) \mid f(x)$, $d(x) \mid g(x)$得到

$$d(x) \mid u(x)f(x) + v(x)g(x) = 1.$$

由于$d(x)$首一, 所以$d(x) = 1$. □

从下面性质可以看出互素多项式在整除讨论中的互不相干性, 我们在后面多项式的唯一分解讨论中需要用到这些性质.

命题1.8. *设$f(x), g(x), h(x) \in \mathbb{F}[x]$, 其中$f(x)$和$g(x)$互素.*

(1) *如果$f(x) \mid g(x)h(x)$, 则$f(x) \mid h(x)$;*

(2) *如果$f(x) \mid h(x)$, $g(x) \mid h(x)$, 则$f(x)g(x) \mid h(x)$.*

☞ **证明.** (1) 由Bezout等式, 存在$u(x), v(x) \in \mathbb{F}[x]$, 使得

$$u(x)f(x) + v(x)g(x) = 1.$$

于是
$$h(x)=u(x)f(x)h(x)+v(x)g(x)h(x).$$
因为
$$f(x)\mid f(x),\quad f(x)\mid g(x)h(x),$$
所以就得到
$$f(x)\mid h(x).$$
(2) 由于$f(x)\mid h(x)$, 所以
$$h(x)=q(x)f(x),\quad (\exists q(x)\in\mathbb{F}[x]).$$
又
$$g(x)\mid h(x)=q(x)f(x),$$
利用(1)就得$g(x)\mid q(x)$. 所以就有
$$f(x)g(x)\mid f(x)q(x)=h(x),$$
证毕. □

最后举一个例子.

例1.11. 试将$\frac{1}{\sqrt[3]{4}+2\sqrt[3]{2}+3}$的分子分母乘以适当的根式将分母有理化.

解. 记$c=\sqrt[3]{2}$, $f(x)=x^2+2x+3$, 则所给需化简数的分母为$f(c)$. 而$c^3=2$, 所以$g(c)=0$, 其中$g(x)=x^3-2$. 做辗转相除

$x-2$	x^2+2x+3	$x^3 \qquad -2$	$x-2$
	x^2+4x	x^3+2x^2+3x	
	$-2x+3$	$-2x^2-3x-2$	
	$-2x-8$	$-2x^2-4x-6$	
	11	$x+4$	

于是11是$f(x)$和$g(x)$的最大公因式. 我们有
$$\begin{aligned}11=&f(x)-(x-2)(x+4)=f(x)-(x-2)(g(x)-(x-2)f(x))\\=&[1+(x-2)^2]f(x)+(-x+2)g(x)=(x^2-4x+5)f(x)+(-x+2)g(x).\end{aligned}$$
将$x=c$代入, 就得到
$$11=(c^2-4c+5)f(c).$$
所以可如下分母有理化
$$\frac{1}{\sqrt[3]{4}+2\sqrt[3]{2}+3}=\frac{\sqrt[3]{4}-4\sqrt[3]{2}+5}{(\sqrt[3]{4}+2\sqrt[3]{2}+3)(\sqrt[3]{4}-4\sqrt[3]{2}+5)}=\frac{\sqrt[3]{4}-4\sqrt[3]{2}+5}{11}.$$
我们在第七章中利用线性变换将给出一个新的解法. □

习题1.6

A1. 如果多项式$f(x),g(x)$不全为零, 证明: $\left(\frac{f(x)}{(f(x),g(x))},\frac{g(x)}{(f(x),g(x))}\right)=1$.
A2. 设多项式$f(x),g(x)$不全为零, 且存在多项式$u(x)$和$v(x)$使得$(f(x),g(x))=u(x)f(x)+v(x)g(x)$, 证明: $(u(x),v(x))=1$.

A3. 设$m \in \mathbb{N}$, 多项式$f(x)$和$g(x)$互素, 证明: $f(x^m)$和$g(x^m)$也互素.

A4. 设多项式$f(x), g(x)$互素, 且$\deg(f(x)) > 0, \deg(g(x)) > 0$. 证明: 存在唯一一组$u(x)$, $v(x) \in \mathbb{F}[x]$, 使得

$$u(x)f(x) + v(x)g(x) = 1,$$

且$\deg(u(x)) < \deg(g(x))$, $\deg(v(x)) < \deg(f(x))$.

A5. 求次数最低的多项式$u(x)$和$v(x)$, 使得$u(x)f(x) + v(x)g(x) = 1$, 其中

(1) $f(x) = x^3$, $g(x) = (x-3)^2$;

(2) $f(x) = x^3 - 3$, $g(x) = x^2 - 2x + 3$.

A6. 将分数$\frac{1}{\sqrt[3]{9}-2\sqrt[3]{3}+3}$的分子分母同乘以适当的数, 将分母化成有理数.

A7. 设$f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$, 如果$(f_1(x), f_2(x), \ldots, f_s(x)) = 1$, 则称$f_1(x)$, $f_2(x)$, $\ldots$, $f_s(x)$互素. 证明: $f_1(x), f_2(x), \ldots, f_s(x)$互素的充分和必要条件是, 存在$u_1(x)$, $u_2(x)$, $\ldots$, $u_s(x) \in \mathbb{F}[x]$, 使得

$$u_1(x)f_1(x) + u_2(x)f_2(x) + \cdots + u_s(x)f_s(x) = 1.$$

B1. 证明: 如果$(f(x), g(x)) = 1$, $(f(x), h(x)) = 1$, 则$(f(x), g(x)h(x)) = 1$.

B2. 证明: 如果$(f(x), g(x)) = 1$, 则$(f(x)g(x), f(x) + g(x)) = 1$.

B3. 设$a, b, c \in \mathbb{Z}$, 且a和b不全为零.

(1) 证明: a和b互素的充分和必要条件是, 存在$u, v \in \mathbb{Z}$, 使得$ua + vb = 1$;

(2) 假设a和b互素. 证明: 如果$a \mid bc$, 那么$a \mid c$;

(3) 假设a和b互素. 证明: 如果$a \mid c$且$b \mid c$, 那么$ab \mid c$.

B4. (中国剩余定理)设$m_1, m_1, \ldots, m_s \in \mathbb{Z}$两两互素[12], $b_1, b_2, \ldots, b_s \in \mathbb{Z}$, 证明: 存在$a \in \mathbb{Z}$, 使得[13] $a \equiv b_i \pmod{m_i}$, $i = 1, 2, \ldots, s$.

B5. 今有物不知其数, 三三数之剩二, 五五数之剩三, 七七数之剩二, 问物几何?

B6. (中国剩余定理)设$g_1(x), g_2(x), \ldots, g_s(x) \in \mathbb{F}[x]$两两互素, $f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$, 证明:

(1) 存在唯一的$f(x) \in \mathbb{F}[x]$, 使得[14] $f(x) \equiv f_i(x) \pmod{g_i(x)}$, $i = 1, 2, \ldots, s$, 且$\deg(f(x)) < \sum\limits_{i=1}^{s} \deg(g_i(x))$;

(2) 集合$\{g(x) \in \mathbb{F}[x] \mid g(x) \equiv f_i(x) \pmod{g_i(x)}, i = 1, 2, \ldots, s\}$为

$$\{f(x) + q(x)g_1(x)\cdots g_s(x) \mid q(x) \in \mathbb{F}[x]\}.$$

B7. 求次数最低的多项式$f(x)$, 使它被x^3除的余式为$x^2 + 2x + 3$, 被$(x-3)^2$除的余式为$3x - 7$.

§ 1.7 唯一分解定理

数学研究有许多基本的想法, 其中常常用到的一个是: 先简单再复杂, 先特殊再一般. 比如, 在整

[12] 设a和b是整数, 如果$(a, b) = 1$, 则称a和b互素.

[13] 对于$a, b, m \in \mathbb{Z}$, 如果$m \mid a - b$, 则称a和b模m同余, 记为$a \equiv b \pmod{m}$.

[14] 对于多项式$f(x), g(x), m(x)$, 如果$m(x) \mid f(x) - g(x)$, 则称$f(x)$和$g(x)$模$m(x)$同余, 记为$f(x) \equiv g(x) \pmod{m(x)}$.

数的世界里, 我们知道可以将整数分解为更小的整数的乘积. 例如, 我们有

$$24 = 2^3 \times 3.$$

这里2和3已经不能再非平凡分解, 而只有等式

$$2 = 1 \times 2 = (-1) \times (-2).$$

这样的正整数称为素数. 具体地, 对于$p \in \mathbb{N}$, 如果$p \neq 1$, 且p的因子只有± 1和$\pm p$, 则称p是素数; 否则, 称p是合数. 这里我们不把1看成素数, 是为了素分解的唯一性. 可以证明算术基本定理: 对任意的$n \in \mathbb{N}$, $n > 1$, n可以唯一地写成素数的乘积.

由算术基本定理, 素数是构成整数世界的基石, 搞清楚素数的性质有重大意义. 许多关于整数的问题, 特别是与乘法有关的问题, 考虑素分解常常可以提供很大的帮助.

下面我们考虑多项式世界的素分解, 即将多项式分解为更简单的多项式的乘积, 这其实就是我们在中学学过的因式分解. 比如, 我们知道有分解

$$x^2 - 1 = (x+1)(x-1),$$

而且感觉上这是最好的分解, 即不能再分解了. 我们希望给出一般的多项式分解的理论, 至少要说明每个多项式都可以这样地分解. 于是我们首先需要定义什么叫不能再分解的多项式, 即要定义多项式世界的“素数”. 在整数世界, 一个大于1的整数, 如果不能非平凡分解则是素数, 这等价于该整数没有非平凡的因子. 类比于此, 我们有下面的定义[15].

定义1.11. *设$p(x) \in \mathbb{F}[x]$且$\deg(p(x)) \geqslant 1$. 如果$p(x)$满足下面的等价条件之一:*

(i) *$p(x)$不能写成$\mathbb{F}$上两个次数比$p(x)$小的多项式的乘积;*

(ii) *$p(x)$的因式只有*

$$c \in \mathbb{F}^\times \quad \text{和} \quad cp(x),\ c \in \mathbb{F}^\times,$$

则称$p(x)$是$\mathbb{F}$上的(或者: $\mathbb{F}[x]$中的)不可约多项式. 否则, 称$p(x)$是可约多项式.

由定义, $p(x)$是不可约多项式的充分必要条件是$cp(x)$是不可约多项式, 其中$c \in \mathbb{F}^\times$. 也就是说, 如果两个多项式相伴, 那么它们是否可约是一致的. 另一方面, 集合$\mathbb{F}[x]$可以写成如下的无交并[16]:

$$\mathbb{F}[x] = \{0\} \cup \mathbb{F}^\times \cup \{\text{不可约多项式}\} \cup \{\text{可约多项式}\}.$$

要注意的是, 和整除以及最大公因式不同, 多项式的可约性与系数域$\mathbb{F}$有关. 例如, x^2-2在$\mathbb{Q}[x]$中不可约, 而在$\mathbb{R}[x]$中可约. 后者由于

$$x^2 - 2 = (x+\sqrt{2})(x-\sqrt{2});$$

而前者可以如下说明: 如果x^2-2在$\mathbb{Q}[x]$中可约, 则x^2-2在$\mathbb{Q}[x]$中有首一一次因式, 这只能为$x \pm \sqrt{2}$; 这得到$\sqrt{2} \in \mathbb{Q}$, 矛盾.

[15]整数环$\mathbb{Z}$ 中的$\{\pm 1\}$对应到多项式环$\mathbb{F}[x]$中的$\mathbb{F}^\times$. 事实上, 它们分别是对应环中乘法可逆元的全体所成子集.

[16]这对应到无交并

$$\mathbb{Z} = \{0\} \cup \{\pm 1\} \cup \{\pm\text{素数}\} \cup \{\pm\text{合数}\}.$$

例1.12. 设$a,b\in\mathbb{F}$且$a\neq 0$, 则$ax+b$是$\mathbb{F}$上的不可约多项式. 即一次因式必不可约.

不可约多项式有下面重要性质.

命题1.9. 设$p(x)\in\mathbb{F}[x]$是不可约多项式.

(1) 对任意$f(x)\in\mathbb{F}[x]$, 有

$$(p(x),f(x))=1\quad\text{或者}\quad p(x)\mid f(x);$$

(2) 设$f(x),g(x)\in\mathbb{F}[x]$, 如果$p(x)\mid f(x)g(x)$, 则$p(x)\mid f(x)$或者$p(x)\mid g(x)$. 进而, 如果

$$p(x)\mid f_1(x)\cdots f_s(x),\quad (f_1(x),\ldots,f_s(x)\in\mathbb{F}[x]),$$

则存在$i(1\leqslant i\leqslant s)$, 使得$p(x)\mid f_i(x)$.

☞ **证明.** (1) 设$(p(x),f(x))=d(x)$, 则$d(x)\mid p(x)$且$d(x)\mid f(x)$. 由于$p(x)$不可约, 所以有

$$d(x)=1\quad\text{或者}\quad d(x)=cp(x),\ (c\in\mathbb{F}^{\times}).$$

当$d(x)=1$时, 有$(p(x),f(x))=1$; 当$d(x)=cp(x)$时, 有$p(x)\mid f(x)$.
(2) 如果$p(x)\nmid f(x)$, 则由(1)有$(p(x),f(x))=1$. 再由$p(x)\mid f(x)g(x)$和命题1.8, 就得到$p(x)\mid g(x)$. □

有了不可分解多项式的概念, 就可以讨论因式分解了. 设$f(x)\in\mathbb{F}[x]$, 且$\deg(f(x))\geqslant 1$. 如果$f(x)$已经不可约, 则$f(x)=f(x)$是其因式分解; 否则存在次数更小的多项式$f_1(x),f_2(x)$, 使得$f(x)=f_1(x)f_2(x)$. 再继续分解$f_1(x)$和$f_2(x)$, 就可以将$f(x)$分解为不可约多项式的乘积. 于是我们有多项式世界的因式分解存在(唯一)定理.

定理1.10 (因式分解存在唯一性定理)**.** 设$f(x)\in\mathbb{F}[x]$且$\deg(f(x))\geqslant 1$, 则$f(x)$可以唯一分解成$\mathbb{F}$上有限个不可约多项式的乘积. 这里唯一性指的是: 如果有两个分解式

$$\begin{aligned}f(x)=&p_1(x)p_2(x)\cdots p_s(x),\\ =&q_1(x)q_2(x)\cdots q_t(x),\end{aligned}$$

其中$p_1(x),\ldots,p_s(x)$和$q_1(x),\ldots,q_t(x)$都是$\mathbb{F}$上的不可约多项式, 则$s=t$, 且适当排列$q_j(x)$的次序后有$p_i(x)\sim q_i(x)$, $i=1,2,\ldots,s$.

☞ **证明.** "分解存在性" 上面已经说明了分解存在性, 这里再写出严格的证明. 对$\deg(f(x))$归纳. 当$\deg(f(x))=1$时, $f(x)$是一次多项式, 不可约, 因此有不可约分解$f(x)=f(x)$. 下设$\deg(f(x))>1$. 如果$f(x)$不可约, 则$f(x)=f(x)$就是不可约分解. 否则$f(x)$可约, 即存在$f_1(x),f_2(x)\in\mathbb{F}[x]$, 使得

$$f(x)=f_1(x)f_2(x),\quad \deg(f_1(x))<\deg(f(x)),\ \deg(f_2(x))<\deg(f(x)).$$

由此可得$\deg(f_1(x)),\deg(f_2(x))\geqslant 1$, 因此可以对$f_1(x)$和$f_2(x)$用归纳假设, 得到$f_1(x)$和$f_2(x)$都可以写成有限个不可约多项式的乘积, 进而$f(x)$亦然.
"分解唯一性" 设有不可约分解

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots q_t(x).$$

我们对s归纳证明. 当$s=1$时, 有$f(x)=p_1(x)$不可约. 于是$t=1$, 且$p_1(x)=q_1(x)$. 下设$s>1$. 由于

$$p_1(x) \mid q_1(x)q_2(x)\cdots q_t(x),$$

且$p_1(x)$不可约, 所以命题1.9导出

$$p_1(x) \mid q_i(x), \quad (\exists i, 1 \leqslant i \leqslant t).$$

适当重排, 可以不妨假设$i=1$. 而$q_1(x)$不可约, 所以必有$p_1(x)$和$q_1(x)$相伴, 即

$$q_1(x)=cp_1(x), \quad (\exists c \in \mathbb{F}^{\times}).$$

这得到

$$p_2(x)\cdots p_s(x)=cq_2(x)\cdots q_t(x).$$

上面也是某一多项式的两个不可约分解, 由归纳假设就得到$s-1=t-1$, 即$s=t$; 以及适当重排后有$p_i(x)\sim q_i(x)$相伴, $i=2,\dots,s$. □

上面证明中未给出分解多项式的具体有效的方法. 事实上, 具体对多项式进行不可约分解是很困难的任务.

设$f(x)\in\mathbb{F}[x]$且$\deg(f(x))\geqslant 1$, 将相伴的不可约多项式放在一起, $f(x)$有下面的不可约分解:

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

其中$c\in\mathbb{F}^{\times}$是$f(x)$的首项系数, 而$p_1(x),p_2(x),\dots,p_s(x)$是两两互不相同的首一不可约多项式, 它们两两互素, 而指数$r_1,r_2,\dots,r_s\in\mathbb{N}$. 称这种分解式为$f(x)$的标准分解式. 称$p_1(x)$, $p_2(x)$, $\dots$, $p_s(x)$是$f(x)$的不可约因式.

利用不可约分解的存在唯一性定理, 也可以重新证明整除的相关结果, 因为我们有下面的结论.

命题1.11. **(1)** 设有标准分解式

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

而$0\neq g(x)\in\mathbb{F}[x]$的首项系数为$d$, 则$g(x)\mid f(x)$的充分必要条件是

$$g(x)=dp_1^{l_1}(x)p_2^{l_2}(x)\cdots p_s^{l_s}(x), \quad 0\leqslant l_i\leqslant r_i,\ i=1,2,\dots,s;$$

(2) 设$0\neq f(x),g(x)\in\mathbb{F}[x]$有分解式

$$f(x)=cp_1^{a_1}(x)p_2^{a_2}(x)\cdots p_t^{a_t}(x), \quad g(x)=dp_1^{b_1}(x)p_2^{b_2}(x)\cdots p_t^{b_t}(x),$$

其中c和d是首项系数, $p_1(x),p_2(x),\dots,p_t(x)$是两两互不相同的首一不可约多项式, 而$a_1$, $\dots$, a_t, b_1, $\dots$, $b_t\in\mathbb{Z}_{\geqslant 0}$, 则

$$(f(x),g(x))=p_1^{r_1}(x)p_2^{r_2}(x)\cdots p_t^{r_t}(x),$$

其中[17]

$$r_i=\min\{a_i,b_i\}, \quad (i=1,2,\dots,t).$$

特别地, $(f(x),g(x))=1$的充分必要条件是它们无相同的不可约因式.

☞ **证明.** (1) 充分性显然, 下证必要性. 设$g(x) \mid f(x)$, 则

$$f(x) = g(x)h(x), \quad (\exists h(x) \in \mathbb{F}[x]).$$

设有分解

$$g(x) = dp_1^{l_1}(x)\cdots p_s^{l_s}(x)q_1^{m_1}(x)\cdots q_k^{m_k}(x),$$
$$h(x) = ep_1^{j_1}(x)\cdots p_s^{j_s}(x)q_1^{n_1}(x)\cdots q_k^{n_k}(x),$$

其中$q_1(x),\ldots,q_k(x)$是两两互不相同的不可约多项式, 且也与$p_1(x),\ldots,p_s(x)$互不相同, 而

$$l_1,\ldots,l_s,m_1,\ldots,m_k,j_1,\ldots,j_s,n_1,\ldots,n_k \in \mathbb{Z}_{\geqslant 0}.$$

于是有

$$cp_1^{r_1}(x)\cdots p_s^{r_s}(x) = dep_1^{l_1+j_1}(x)\cdots p_s^{l_s+j_s}(x)q_1^{m_1+n_1}(x)\cdots q_k^{m_k+n_k}(x).$$

由分解的唯一性得

$$\begin{cases} r_i = l_i + j_i, & i = 1,\ldots,s, \\ 0 = m_i + n_i, & i = 1,\ldots,k, \end{cases}$$

这导出

$$\begin{cases} r_i \geqslant l_i, & i = 1,\ldots,s, \\ m_i = n_i = 0, & i = 1,\ldots,k, \end{cases}$$

即(1)的必要性成立.

(2) 记$d(x) = p_1^{r_1}(x)p_2^{r_2}(x)\cdots p_t^{r_t}(x)$, 则$d(x)$是首一多项式. 因为

$$r_i \leqslant a_i,\ r_i \leqslant b_i, \quad i = 1,2,\ldots,t,$$

所以由(1)得

$$d(x) \mid f(x), \quad d(x) \mid g(x).$$

任取$f(x)$和$g(x)$的公因式$\varphi(x)$, 由(1)可设

$$\varphi(x) = ep_1^{l_1}(x)p_2^{l_2}(x)\cdots p_t^{l_t}(x), \quad 0 \leqslant l_i \leqslant \begin{matrix} a_i \\ b_i \end{matrix},\ i = 1,2,\ldots,t.$$

于是

$$0 \leqslant l_i \leqslant \min\{a_i, b_i\} = r_i, \quad i = 1,2,\ldots,t,$$

再用(1)有$\varphi(x) \mid d(x)$. 所以$(f(x), g(x)) = d(x)$. □

习题1.7

A1. 证明: 下面多项式在实数域和有理数域上都不可约:

(1) x^2+1; (2) x^2+x+1.

A2. 求下面多项式的标准分解式(系数域分别为$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$):

[17] min为取小函数, 即对于有限的有序集合A, $\min A$取出A中的最小元, 特别有

$$\min\{a,b\} = \begin{cases} a, & 如果\ a \leqslant b, \\ b, & 如果\ a > b. \end{cases}$$

(1) x^3-2x^2-2x+1;　　(2) x^4+4.

A3. 设$\mathbb{F}_1$和$\mathbb{F}_2$是数域, 且$\mathbb{F}_1\subset\mathbb{F}_2$. 设$p(x)\in\mathbb{F}_1[x]\subset\mathbb{F}_2[x]$, 证明: 如果$p(x)$在$\mathbb{F}_2$上不可约, 则$p(x)$也在$\mathbb{F}_1$上不可约.

A4. 设$p(x)\in\mathbb{F}[x]$且$\deg(p(x))\geqslant 1$, 证明: $p(x)$是$\mathbb{F}$上的不可约多项式的充分和必要条件是, 对任意$f(x),g(x)\in\mathbb{F}[x]$, 如果$p(x)\mid f(x)g(x)$, 则$p(x)\mid f(x)$或者$p(x)\mid g(x)$.

A5. 设$p(x)\in\mathbb{F}[x]$, $n=\deg(p(x))\geqslant 1$. 证明:

(1) 对于$a\in\mathbb{F}$, $p(x)$是数域$\mathbb{F}$上不可约多项式的充分必要条件是$p(x+a)$是数域$\mathbb{F}$上不可约多项式;

(2) 若$p(0)\neq 0$, 则$p(x)$是数域$\mathbb{F}$上不可约多项式的充分必要条件是$x^np\left(\frac{1}{x}\right)$是数域$\mathbb{F}$上不可约多项式.

A6. 设$f(x)\in\mathbb{F}[x]$, 且$\deg(f(x))>0$, 证明: 下面命题等价

(i) 存在$\mathbb{F}[x]$中的不可约多项式$p(x)$和$m\in\mathbb{N}$, 使得$f(x)\sim p^m(x)$;

(ii) 对任意$g(x)\in\mathbb{F}[x]$, 有$(f(x),g(x))=1$, 或者存在$m\in\mathbb{N}$, 使得$f(x)\mid g^m(x)$;

(iii) 对任意$g(x),h(x)\in\mathbb{F}[x]$, 如果$f(x)\mid g(x)h(x)$, 那么$f(x)\mid g(x)$, 或者存在$m\in\mathbb{N}$, 使得$f(x)\mid h^m(x)$.

A7. 设$m\in\mathbb{N}$, $f(x),g(x)\in\mathbb{F}[x]$, 证明:

$$g^m(x)\mid f^m(x)\Longleftrightarrow g(x)\mid f(x).$$

A8. 设$m\in\mathbb{N}$, $f(x),g(x)\in\mathbb{F}[x]$, 证明:

$$(f^m(x),g^m(x))=(f(x),g(x))^m,\quad [f^m(x),g^m(x)]=[f(x),g(x)]^m.$$

特别地, 如果$f(x)$和$g(x)$互素, 那么$f^m(x)$和$g^m(x)$也互素.

A9. 设非零多项式$f_1(x),f_2(x),\ldots,f_s(x)\in\mathbb{F}[x]$有分解式

$$f_i(x)=c_ip_1^{a_{i1}}(x)p_2^{a_{i2}}(x)\cdots p_t^{a_{it}}(x),\quad i=1,2,\ldots,s,$$

其中$c_i\in\mathbb{F}$, $p_1(x),p_2(x),\ldots,p_t(x)\in\mathbb{F}[x]$是两两互不相同的首一不可约多项式, 而$a_{ij}$是非负整数. 证明:

$$(f_1(x),f_2(x),\ldots,f_s(x))=p_1^{m_1}(x)p_2^{m_2}(x)\cdots p_t^{m_t}(x),$$
$$[f_1(x),f_2(x),\ldots,f_s(x)]=p_1^{n_1}(x)p_2^{n_2}(x)\cdots p_t^{n_t}(x),$$

其中

$$m_j=\min\{a_{ij}\mid i=1,2,\ldots,s\},\quad n_j=\max\{a_{ij}\mid i=1,2,\ldots,s\},\quad j=1,2,\ldots,t.$$

B1. (1) 设p是素数, $a\in\mathbb{Z}$, 证明: $p\mid a$或者p与a互素;

(2) 设p是素数, $a,b\in\mathbb{Z}$, 且$p\mid ab$, 证明: $p\mid a$或者$p\mid b$;

(3) 证明算术基本定理: 对任意的$n\in\mathbb{N}$, $n>1$, n可以唯一地写成素数的乘积.

B2. 设$u\in\mathbb{C}$, 如果存在非零有理系数多项式$f(x)$, 使得$f(u)=0$, 则称u是一个代数数.

(1) 证明: $\sqrt{2}$, $\sqrt{2}+\sqrt{3}$都是代数数;

(2) 如果u是代数数, 则存在有理数域上的次数最小的首一多项式$g(x)$, 使得$g(u)=0$. 多项式$g(x)$由u唯一确定, 称$g(x)$为u的极小多项式. 证明: $g(x)$是有理数域上的不可约多项式, 且对于$f(x)\in\mathbb{Q}[x]$, 成立

$$f(u)=0\Longleftrightarrow g(x)\mid f(x);$$

(3) 求代数数$\sqrt{2}+\sqrt{3}$的极小多项式.

B3. 设m是正整数, $f(x)=x^4+m$.

(1) $f(x)$是否在$\mathbb{R}[x]$中可约? 如果可约, 写出它的标准分解式;

(2) 求$f(x)$在$\mathbb{Q}[x]$中可约的充分必要条件. 当$f(x)$在$\mathbb{Q}[x]$中可约时, 写出它的标准分解式.

B4. 在有理数域上因式分解:

(1) $x^{15}-1$; (2) $x^{18}+x^{15}+x^{12}+x^9+x^6+x^3+1$.

§ 1.8 多项式的根

前面我们从理论证明了, 和整数一样, 多项式可以唯一地分解为不能再分解的多项式的乘积, 即给了中学所学的多项式因式分解的一般理论. 下面很自然就要问, 给定一个多项式, 如何具体地进行因式分解? 我们回到前面不可约分解的存在唯一性定理的存在性的证明. 设多项式$f(x)\in\mathbb{F}[x]$, 且$\deg(f(x))\geqslant 1$. 如果$f(x)$不可约, 那么$f(x)=f(x)$就是所要求的因式分解; 否则存在$f(x)$的(次数最小的)不可约因式$p(x)$, 做带余除法可得$f(x)=p(x)f_1(x)$. 再继续分解$f_1(x)$, 有限步后停止, 就得到$f(x)$的因式分解. 从这个过程可以看出, 我们需要知道如何判断多项式是否可约和知道如何找多项式的不可约因式, 而这似乎是个困难的任务, 毕竟多项式是否可约还和数域有关系. 能否从整数世界得到一些启发和提示呢? 不可约多项式对应到素数, 于是对应到整数世界的问题是如何判断一个正整数是否是素数和如何找一个正整数的素因子. 这些问题在整数世界并不容易, 所以似乎多项式们从整数那儿已经学不到更多东西了. 于是我们要考虑一些多项式所具有的特殊性质, 比如多项式可以看成函数, 这是整数世界所没有的.

设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{F}[x],$$

则得到多项式函数$f:\mathbb{F}\to\mathbb{F}$, 使得$f$在$x=c\in\mathbb{F}$的值是[18]

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0\in\mathbb{F}.$$

如果还有$g(x)\in\mathbb{F}[x]$, 而

$$u(x)=f(x)+g(x),\quad v(x)=f(x)g(x),$$

则对任意的$c\in\mathbb{F}$, 显然成立[19]

$$u(c)=f(c)+g(c),\quad v(c)=f(c)g(c).$$

进而, 多项式代数等式中令$x=c$后等式仍然成立.

多项式在某点处的值还可以如下得到. 假设$\mathbb{F}=\mathbb{R}$, 考虑$x=c$处多项式函数f的性质, 可以在$x=c$处Taylor展开

$$f(x)=f(c)+\frac{f'(c)}{1!}(x-c)+\frac{f''(c)}{2!}(x-c)^2+\cdots.$$

[18]后面还可以将多项式的不定元x代入其它对象, 比如矩阵.

[19]参看习题1.2.

将上式改写为

$$f(x)=f(c)+(x-c)\left(\frac{f'(c)}{1!}+\frac{f''(c)}{2!}(x-c)+\cdots\right),$$

从多项式的观点看就是$f(x)$除以首一多项式$x-c$的带余除法等式. 于是多项式在某点处的值可以用带余除法(综合除法)得到, 即有下面的余式定理.

命题1.12 (余式定理)**.** 设$f(x)\in\mathbb{F}[x]$, $c\in\mathbb{F}$, 则

$$f(c)=\text{“}f(x)\text{除以}x-c\text{的余式”}.$$

证明. 这里给出一个代数证明. 设

$$f(x)=q(x)(x-c)+r(x),$$

其中$q(x),r(x)\in\mathbb{F}[x]$, 且$\deg(r(x))<\deg(x-c)=1$. 于是$\deg(r(x))=0$或$-\infty$, 即$r(x)=a\in\mathbb{F}$. 将$x=c$代入上面多项式等式, 就得到

$$f(c)=q(c)(c-c)+a=a,$$

得证. □

现在多项式有两个身份, 即可以看成多项式, 又可以看成多项式函数. 那么这两个身份是否一致? 即两个多项式相等是否等价于它们作为多项式函数相等? 两个函数相等, 当且仅当它们在定义域的每一点上的函数值都相等. 所以如果两个多项式$f(x)$和$g(x)$作为多项式相等, 那么作为多项式函数, 它们自然也相等. 我们考察逆命题是否成立. 假设$f(x)$和$g(x)$作为多项式函数相等, 我们要看它们作为多项式是否相等. 考虑$f(x)-g(x)$, 这个问题等价于: 设多项式$f(x)$是零函数, 那么$f(x)$是否是零多项式? 再看其逆否命题: 设$f(x)$是非零多项式, 那么$f(x)$是否不是零函数? 所谓零函数, 就是在定义域每一点上的值都为零的函数. 不是零函数, 就是定义域中有一点的函数值不为零. 于是, 我们需要考虑函数值等于0的点, 它们就是所谓的零点.

定义1.12. 设$f(x)\in\mathbb{F}[x]$, $c\in\mathbb{F}$, 如果$f(c)=0$, 则称c为$f(x)$(在$\mathbb{F}$上)的一个根(或零点).

结合余式定理, 立刻发现根对应到多项式的一次因式, 即有下面的因式定理.

命题1.13 (因式定理)**.** 设$f(x)\in\mathbb{F}[x]$, $c\in\mathbb{F}$, 则

$$c\text{ 为 }f(x)\text{ 的根}\iff x-c\mid f(x).$$

设$f(x)\in\mathbb{F}[x]$, $c\in\mathbb{F}$, 可以假设

$$f(x)=(x-c)^k g(x),\quad x-c\nmid g(x),\ k\in\mathbb{Z}_{\geqslant 0}.$$

由因式定理得

$$c\text{为 }f(x)\text{ 的根}\iff k>0.$$

当$k>0$时, 我们称k为$f(x)$的根c的重数, 也称c为$f(x)$的k重根. 如果$k=1$, 则称c是$f(x)$的单根; 如果$k>1$, 则称c为重根.

例1.13. 设$f(x)=x^4-4$, 我们有下面的标准分解

$$\begin{aligned}f(x)=&(x^2+2)(x^2-2) \qquad (\mathbb{Q}[x]\text{中})\\ =&(x^2+2)(x+\sqrt{2})(x-\sqrt{2}) \qquad (\mathbb{R}[x]\text{中})\\ =&(x+\sqrt{2}i)(x-\sqrt{2}i)(x+\sqrt{2})(x-\sqrt{2}) \qquad (\mathbb{C}[x]\text{中})\end{aligned}$$

于是$f(x)$在$\mathbb{Q}$中有零个根, 在$\mathbb{R}$中有2个根, 在$\mathbb{C}$中有4个根, 且根都是单根.

例1.14. 设$f(x)=x^3-3x+2$, 由于

$$f(x)=(x-1)^2(x+2),$$

所以1是$f(x)$的2重根, -2是$f(x)$的单根, 进而$f(x)$在$\mathbb{Q}$(或者$\mathbb{R}$, 或者$\mathbb{C}$)中有3个根(重根计算重数).

回到上面问题, 给定非零多项式$f(x)$, 我们希望存在$\mathbb{F}$上的点不是$f(x)$的根. 或者说我们希望$f(x)$的根不能太多, 而根对应到一次因式, $f(x)$只有有限个一次因式, 所以确实$f(x)$不能有太多的根. 具体地, 我们有下面一般性的结论.

定理1.14. 设$0\neq f(x)\in\mathbb{F}[x]$, $\deg(f(x))=n\in\mathbb{Z}_{\geqslant 0}$, 则$f(x)$在$\mathbb{F}$中最多有$n$个根(重根计算重数).

☞ **证明.** 当$n=0$时, $f(x)\in\mathbb{F}^\times$, 此时有零个根. 当$n>0$时, 由唯一分解定理, $f(x)$可以分解为

$$f(x)=(x-c_1)^{k_1}(x-c_2)^{k_2}\cdots(x-c_s)^{k_s}g(x),$$

其中, $c_1,c_2,\ldots,c_s\in\mathbb{F}$互不相同, $k_1,k_2,\ldots,k_s\in\mathbb{N}$, 而$g(x)\in\mathbb{F}[x]$无一次因式. 于是取次数, 得到

$$n=\deg(f(x))=k_1+k_2+\cdots+k_s+\deg(g(x))\geqslant k_1+k_2+\cdots+k_s,$$

即$f(x)$在$\mathbb{F}$中根的个数$k_1+k_2+\cdots+k_s$不超过n. □

最后就可得非零多项式一定是非零函数, 即两个多项式如果作为多项式函数相等, 则作为多项式也相等.

推论1.15. 设$f(x),g(x)\in\mathbb{F}[x]$满足

$$\deg(f(x)),\deg(g(x))\leqslant n\in\mathbb{N},$$

又设$a_1,a_2,\ldots,a_{n+1}\in\mathbb{F}$两两互不相同, 满足

$$f(a_i)=g(a_i),\quad i=1,2,\ldots,n+1,$$

则有$f(x)=g(x)$.

☞ **证明.** 令$h(x)=f(x)-g(x)\in\mathbb{F}[x]$, 则$\deg(h(x))\leqslant n$. 而

$$h(a_i)=f(a_i)-g(a_i)=0,\quad i=1,2,\ldots,n+1,$$

所以$h(x)$在$\mathbb{F}$中至少有$n+1$个根. 由定理1.14, 得$h(x)=0$, 即$f(x)=g(x)$. □

该推论还告诉我们, 要确定一个n次多项式, 只需要知道其在$n+1$个点上的取值, 参看本节习题.

习题1.8

A1. 设$f(x), g(x) \in \mathbb{F}[x]$, 且$\deg(f(x)) > 0$, $\deg(g(x)) > 0$. 证明: $f(x)$和$g(x)$有公共复根的充分必要条件是它们不互素.

A2. 下面有理系数多项式$f(x)$与$g(x)$有无公共复根, 如果有, 把它们求出来, 其中

$$f(x) = x^3 + 2x^2 + 2x + 1, \quad g(x) = x^4 + x^3 + 2x^2 + x + 1.$$

A3. 设$f(x), g(x) \in \mathbb{F}[x]$, 其中$\deg(g(x)) \leqslant 1$, 证明: $g(x) \mid f^2(x)$的充分必要条件是$g(x) \mid f(x)$.

A4. 证明:

(1) $x(x+1)(2x+1) \mid (x+1)^{2n} - x^{2n} - 2x - 1$;

(2) $x^4 + x^3 + x^2 + x + 1 \mid x^5 + x^{11} + x^{17} + x^{23} + x^{29}$.

A5. 设$a \in \mathbb{F}$, $f(x) \in \mathbb{F}[x]$, $n \in \mathbb{N}$. 证明: 如果$(x-a) \mid f(x^n)$, 则$(x^n - a^n) \mid f(x^n)$.

A6. 设$f(x) \in \mathbb{F}[x]$且$\deg(f(x)) \geqslant 2$.

(1) 证明: 如果$f(x)$是$\mathbb{F}[x]$中的不可约多项式, 则$f(x)$在$\mathbb{F}$中没有根;

(2) 如果不要求$\deg(f(x)) \geqslant 2$, (1)是否还成立? 说明理由;

(3) (1)的逆命题是否成立? 说明理由.

B1. 设$f_1(x), f_2(x) \in \mathbb{Q}[x]$, 证明: 如果$(x^2+x+1) \mid (f_1(x^3) + xf_2(x^3))$, 则1是$f_i(x)$的根, $i = 1, 2$.

B2. 设$f_1(x), f_2(x) \in \mathbb{Q}[x]$, 证明: 如果$(x^2+1) \mid (f_1(x^4) + xf_2(x^4))$, 则1是$f_i(x)$的根, $i = 1, 2$.

B3. 设$f(x) \in \mathbb{F}[x]$, 且$\deg(f(x)) > 0$, $n \in \mathbb{N}$. 证明: 如果$f(x) \mid f(x^n)$, 则$f(x)$的复根只能是零或者单位根[20].

B4. 证明Lagrange插值定理: 设$a_1, a_2, \ldots, a_n$是$\mathbb{F}$中n个两两不同的数, $b_1, b_2, \ldots, b_n$是$\mathbb{F}$中任意n个数, 则存在唯一的$f(x) \in \mathbb{F}[x]$, 满足$\deg(f(x)) \leqslant n-1$, 且

$$f(a_i) = b_i, \quad i = 1, 2, \ldots, n.$$

并求出这个多项式.

B5. 设$f(x)$是一个n次多项式, 且当$k = 0, 1, \ldots, n$时, 有$f(k) = \frac{k}{k+1}$. 求$f(n+1)$.

§ 1.9 重因式

设$f(x) \in \mathbb{F}[x]$且$\deg(f(x)) \geqslant 1$, 我们要求$f(x)$的标准分解式

$$f(x) = cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x).$$

这等价于两部分: 确定$f(x)$的所有两两不同的不可约因式$p_1(x), \ldots, p_s(x)$, 确定每个不可约因式$p_i(x)$在$f(x)$的因式分解中出现的次数r_i. 本节先讨论后者. 我们看到整数r_i满足

$$p_i^{r_i}(x) \mid f(x), \quad p_i^{r_i+1}(x) \nmid f(x),$$

称r_i为$f(x)$的不可约因式$p_i(x)$的重数.

[20]复数a是单位根, 如果存在$m \in \mathbb{N}$, 使得$a^m = 1$.

定义1.13. 设$p(x)\in\mathbb{F}[x]$是首一不可约多项式, $k\in\mathbb{N}$, 称$p(x)$是$f(x)\in\mathbb{F}[x]$的k重因式, 如果

$$p^k(x)\mid f(x),\quad p^{k+1}(x)\nmid f(x).$$

当$k=1$时, 称$p(x)$是$f(x)$的单因式; 当$k>1$时, 称$p(x)$是$f(x)$的重因式. 为了方便, 如果$p(x)\nmid f(x)$, 也称$p(x)$为$f(x)$的0重因式.

由该定义, $p(x)$是$f(x)$的k重因式当且仅当$p(x)$在$f(x)$的不可约分解中恰好出现k次:

$$f(x)=p^k(x)g(x),\quad p(x)\nmid g(x).$$

设$f(x)$的标准分解式是

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

则$p_i(x)$是$f(x)$的r_i重因式, $i=1,2,\dots,s$. 而其它的首一不可约多项式都不是$f(x)$的因式.

重因式的概念是重根概念的推广. 设$c\in\mathbb{F}$, $f(x)\in\mathbb{F}[x]$, 则可知

$$c\text{ 是 }f(x)\text{ 的 }k\text{ 重根}\Longleftrightarrow x-c\text{ 是 }f(x)\text{ 的 }k\text{ 重因式}.$$

例1.15. 设$f(x)=-8(x+1)^3(x-5)^2(x^2+x+1)(x^2+1)^2\in\mathbb{R}[x]$, 则

- $x+1$是$f(x)$的3重因式, -1是$f(x)$的3重根;
- $x-5$是$f(x)$的2重因式, 5是$f(x)$的2重根;
- x^2+x+1是$f(x)$的单因式;
- x^2+1是$f(x)$的2重因式.

如何求一个不可约因式的重数? 设$\mathbb{F}=\mathbb{R}$, 多项式$f(x)$有根$x=1$, 我们看如何求其重数. 此时可设

$$f(x)=(x-1)^kg(x),\quad x-1\nmid g(x).$$

将多项式函数$g(x)$在$x=1$处Taylor展开, 得到

$$g(x)=g(1)+\frac{g'(1)}{1!}(x-1)+\cdots,\quad g(1)\neq 0.$$

于是

$$f(x)=(x-1)^kg(1)+\frac{g'(1)}{1!}(x-1)^{k+1}+\cdots,$$

这也是$f(x)$在$x=1$处的Taylor展开. 比较可得

$$f(1)=f'(1)=\cdots=f^{(k-1)}(1)=0,\quad f^{(k)}(1)=k!g(1)\neq 0,$$

上面就是$x=1$是$f(x)$的k重根的充分必要条件, 也就是不可约因式$x-1$的重数是k的充分必要条件.

一般地, 多项式的不可约因式的重数应该和这个多项式的导数有关. 当数域不是$\mathbb{R}$或$\mathbb{C}$时, 可能不好讲极限, 但是我们只需要多项式的导数, 所以可如下定义任意数域上多项式的导数.

定义1.14. 设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{F}[x]$, 定义$f(x)$的导数为

$$f'(x)=f^{(1)}(x):=na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+2a_2x+a_1\in\mathbb{F}[x].$$

对于$k\in\mathbb{Z}_{>1}$, 归纳定义k阶导数为

$$f^{(k)}(x):=\left(f^{(k-1)}(x)\right)'.$$

也记$f^{(2)}(x)=f''(x)$, $f^{(3)}(x)=f'''(x)$.

例1.16. 设$f(x)=2x^4+3x^3-4x^2+5x-6$, 则

$$\begin{aligned}&f'(x)=8x^3+9x^2-8x+5, \qquad f''(x)=24x^2+18x-8,\\ &f'''(x)=48x+18, \qquad f^{(4)}(x)=48, \qquad f^{(k)}(x)=0, \quad (\forall k\geqslant 5).\end{aligned}$$

设$f(x),g(x)\in\mathbb{F}[x]$, 容易证明下面的性质成立:

(1) $(f(x)+g(x))'=f'(x)+g'(x)$;

(2) $(cf(x))'=cf'(x), \quad (c\in\mathbb{F})$;

(3) $(f(x)g(x))'=f'(x)g(x)+f(x)g'(x)$;

(4) $(f^m(x))'=mf^{m-1}(x)f'(x), \quad (m\in\mathbb{N})$.

有了多项式导数的概念, 下面我们看如何求多项式的不可约因式的重数. 设$f(x)\in\mathbb{F}[x]$, $p(x)$是$f(x)$的k重因式, 其中$k\in\mathbb{N}$. 则

$$f(x)=p^k(x)g(x), \quad (\exists g(x)\in\mathbb{F}[x],\ p(x)\nmid g(x)).$$

求导得到

$$\begin{aligned}f'(x)=&(p^k(x))'g(x)+p^k(x)g'(x)\\ =&kp^{k-1}(x)p'(x)g(x)+p^k(x)g'(x)\\ =&p^{k-1}(x)(kp'(x)g(x)+p(x)g'(x)).\end{aligned}$$

由于$\deg(p'(x))<\deg(p(x))$且$p'(x)\neq 0$, 所以$p(x)\nmid p'(x)$. 又因为$k>0$, $p(x)\nmid g(x)$和$p(x)$不可约, 所以

$$p(x)\nmid kp'(x)g(x).$$

而$p(x)\mid p(x)g'(x)$, 所以

$$p(x)\nmid kp'(x)g(x)+p(x)g'(x),$$

即有$p(x)$是$f'(x)$的$k-1$重因式.

于是我们得到

定理1.16. 设$p(x), f(x) \in \mathbb{F}[x]$, 其中$p(x)$是$f(x)$首一不可约因式, 设$k \in \mathbb{N}$, 则

$$p(x) \text{ 为 } f(x) \text{ 的 } k \text{ 重因式 } \Longleftrightarrow p(x) \text{ 是 } f'(x) \text{ 的 } k-1 \text{ 重因式}.$$

特别地, $f(x)$的单因式不再是$f'(x)$的因式.

我们看一个简单的例子.

例1.17. 设$f(x) = x^3 - 3x + 2 \in \mathbb{R}[x]$, 则

$$f'(x) = 3x^2 - 3 = 3(x+1)(x-1).$$

由于$f(1) = 0$, $f(-1) = 4$, 所以$x-1$是$f(x)$的因式, 而$x+1$不是. 由定理1.16, $x-1$是$f(x)$的2重因式, $f(x)$还有一个不同于$x \pm 1$的单因式. 该单因式为

$$\frac{f(x)}{(x-1)^2} = x + 2.$$

于是

$$f(x) = (x-1)^2(x+2).$$

注意到$(f(x), f'(x)) = x - 1$.

定理1.16有下面的推论.

推论1.17. 设$k \in \mathbb{N}$, $p(x), f(x) \in \mathbb{F}[x]$, 其中$p(x)$为首一不可约多项式.

(1) 若$p(x)$是$f(x)$的k重因式, 则$p(x)$分别是$f'(x), f''(x), \ldots, f^{(k-1)}(x)$的$k-1$, $k-2$, $\ldots$, 1重因式, 而不是$f^{(k)}(x)$ 的因式;

(2) $p(x)$是$f(x)$的k重因式当且仅当

$$p(x) \mid f(x), f'(x), \ldots, f^{(k-1)}(x), \quad \text{但是} \quad p(x) \nmid p^{(k)}(x);$$

(3) $p(x)$为$f(x)$的重因式$\Longleftrightarrow p(x) \mid (f(x), f'(x))$.

此时, 如果$p(x) \mid (f(x), f'(x))$为k重因式, 则$p(x)$为$f(x)$的$k+1$重因式;

(4) $f(x)$无重因式$\Longleftrightarrow (f(x), f'(x)) = 1$.

☞ **证明.** (1) 反复应用定理1.16.
(2) 必要性由(1). 反之, 由于

$$p(x) \mid f^{(k-1)}(x), \quad p(x) \nmid f^{(k)}(x),$$

所以由定理1.16, 可得$p(x)$是$f^{(k-1)}(x)$的单因式. 再由$p(x) \mid f^{(k-2)}(x)$和定理1.16, 可知$p(x)$是$f^{(k-2)}(x)$的2重因式. 继续这个讨论, 就得到$p(x)$是$f(x)$的k重因式.
(3) 由(2).
(4) 由(3). □

例1.18. 证明: $f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} \in \mathbb{Q}[x]$无重因式.

证明. 我们有
$$f'(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!} = f(x) - \frac{x^n}{n!},$$
于是有
$$(f(x), f'(x)) = \left(f(x), f(x) - \frac{x^n}{n!}\right) = \left(f(x), \frac{x^n}{n!}\right) = 1.$$
这得到$f(x)$无重因式. □

利用推论1.17可以得到一个多项式有重根的充分和必要条件.

例1.19. 设$f(x) \in \mathbb{F}[x]$, $c \in \mathbb{C}$, $k \in \mathbb{N}$. 证明:

(1) c是$f(x)$的重根$\Longleftrightarrow f(c) = f'(c) = 0$;

(2) c是$f(x)$的k重根$\Longleftrightarrow f(c) = f'(c) = \cdots = f^{(k-1)}(c) = 0$, 而$f^{(k)}(c) \neq 0$;

(3) $f(x)$在$\mathbb{F}$中有重根$\Longleftrightarrow (f(x), f'(x))$在$\mathbb{F}[x]$中有一次因式.

证明. 利用推论1.17, 因式定理和事实: c是$f(x)$的k重根$\Longleftrightarrow x - c$是$f(x)$的$k$重因式. □

如果$f(x) \in \mathbb{F}[x]$在$\mathbb{F}$中有重根, 那么$f(x)$在$\mathbb{F}[x]$中有重因式. 但是, 如果$f(x)$在$\mathbb{F}[x]$中有重因式, 不一定可以得到$f(x)$在$\mathbb{F}[x]$中有重根. 例如$f(x) = (x^2-2)^2 \in \mathbb{Q}[x]$有二重因式$x^2 - 2$, 但是$f(x)$在$\mathbb{Q}$中没有根, 当然更没有重根了. 要注意的是, 当$\mathbb{F} = \mathbb{C}$时, $\mathbb{C}[x]$中的多项式$f(x)$在$\mathbb{C}$中有重根当且仅当$f(x)$有重因式.

例1.20. 设有有理系数多项式$f(x) = x^3 - 3x^2 + ax + 4$, $g(x) = x^4 + x^3 + bx + c$.

(1) 求a的值, 使得$f(x)$在$\mathbb{Q}$中有重根, 并求出重根和相应的重数;

(2) 设-1是$g(x)$的二重根, 求b, c.

解. (1) 我们有$f'(x) = 3x^2 - 6x + a$. 由带余除法, 得
$$f(x) = \frac{1}{3}(x-1)f'(x) + \frac{2}{3}(a-3)x + \frac{a+12}{3}.$$
于是, 当$a = 3$时, $(f(x), f'(x)) = 1$, 此时$f(x)$无重根. 下设$a \neq 3$, 继续做带余除法
$$f'(x) = \left(\frac{9}{2(a-3)}x - \frac{45a}{4(a-3)^2}\right)\left(\frac{2}{3}(a-3)x + \frac{a+12}{3}\right) + \frac{a(4a^2 - 9a + 216)}{4(a-3)^2}.$$
所以当$a \neq 3$时, $(f(x), f'(x)) \neq 1$的充分和必要条件是
$$a(4a^2 - 9a + 216) = 0,$$
且此时
$$(f(x), f'(x)) = x + \frac{a+12}{2a-6}.$$
而$a \in \mathbb{Q}$, 所以得到
$$(f(x), f'(x)) \neq 1 \Longleftrightarrow a = 0.$$
此时, $(f(x), f'(x)) = x - 2$. 于是$f(x)$在$\mathbb{Q}$中有重根的充分必要条件是$a = 0$, 且此时2是$f(x)$的二重根.
(2) 我们有
$$g'(x) = 4x^3 + 3x^2 + b, \quad g''(x) = 12x^2 + 6x.$$

于是-1是$g(x)$的二重根当且仅当

$$g(-1) = g'(-1) = 0, \quad g''(-1) \neq 0,$$

即

$$-b + c = 0, \quad -1 + b = 0, \quad 6 \neq 0.$$

这得到$b = c = 1$. □

由推论1.17, 多项式$f(x)$的重因式恰为$(f(x), f'(x))$的因式, 而且重数恰好差1. 进而, 如果令$f(x)$的标准分解式为

$$f(x) = cp_1^{r_1}(x) \cdots p_s^{r_s}(x),$$

则由定理1.16可知

$$(f(x), f'(x)) = p_1^{r_1-1}(x) \cdots p_s^{r_s-1}(x).$$

于是得到

$$g(x) := \frac{f(x)}{(f(x), f'(x))} = cp_1(x) \cdots p_s(x),$$

即$g(x)$无重因式, 且与$f(x)$有相同的不可约因式(首项系数也相同). 由$(f(x), f'(x))$以及$g(x)$的因式分解可得$f(x)$的因式分解.

例1.21. 求有理系数多项式$f(x) = x^4 + x^3 - 3x^2 - 5x - 2$的重因式及标准分解式, 并求首一多项式$g(x)$, 使得$g(x)$无重因式, 且与$f(x)$有相同的不可约因式.

☞ **解.** 我们有

$$f'(x) = 4x^3 + 3x^2 - 6x - 5.$$

由辗转相除

$$\begin{array}{r|r|r|r}
\frac{1}{4}x + \frac{1}{16} & x^4 + x^3 - 3x^2 - 5x - 2 & 4x^3 + 3x^2 - 6x - 5 & -\frac{64}{27}x + \frac{80}{27} \\
 & x^4 + \frac{3}{4}x^3 - \frac{3}{2}x^2 - \frac{5}{4}x & 4x^3 + 8x^2 + 4x & \\
\cline{2-3}
 & \frac{1}{4}x^3 - \frac{3}{2}x^2 - \frac{15}{4}x - 2 & -5x^2 - 10x - 5 & \\
 & \frac{1}{4}x^3 + \frac{3}{16}x^2 - \frac{3}{8}x - \frac{5}{16} & -5x^2 - 10x - 5 & \\
\hline
 & -\frac{27}{16}x^2 - \frac{27}{8}x - \frac{27}{16} & 0 &
\end{array}$$

和

$$-\frac{27}{16}x^2 - \frac{27}{8}x - \frac{27}{16} = -\frac{27}{16}(x^2 + 2x + 1)$$

得到

$$(f(x), f'(x)) = x^2 + 2x + 1 = (x+1)^2.$$

于是$f(x)$有3重因式$x + 1$. 而

$$g(x) = \frac{f(x)}{(f(x), f'(x))} = x^2 - x - 2 = (x+1)(x-2),$$

进而得到

$$f(x) = g(x)(f(x), f'(x)) = (x+1)^3(x-2),$$

此即标准分解式. □

习题1.9

A1. 判断有理系数多项式$f(x)$有无重因式, 如果有, 试求出它的重数. 并求首一多项式$g(x)$, 使得$g(x)$无重因式, 且与$f(x)$有相同的不可约因式, 其中
(1) $f(x)=x^5-5x^4+7x^3-2x^2+4x-8$;
(2) $f(x)=x^4+4x^2-4x-3$.
A2. 求t值使$f(x)=x^3-3x^2+tx-1$有重因式.
A3. 设$f(x)\in\mathbb{F}[x]$且$\deg(f(x))>0$, $a\in\mathbb{F}$, 而$g(x)=f(x+a)$. 证明: $f(x)$有重因式当且仅当$g(x)$有重因式.
A4. 设$(x-1)^2\mid Ax^4+Bx+1$, 求A,B.
A5. 设$f(x)=x^4+5x^3+ax^2+bx+c\in\mathbb{Q}[x]$, 如果$-2$是$f(x)$的3重根, 求$a,b,c$.
A6. 设$p(x)\in\mathbb{F}[x]$是不可约多项式, 证明: $p(x)$在复数域中无重根.
B1. (1) 求多项式$x^3+ax+b\in\mathbb{F}[x]$有重因式的条件;
(2) 求多项式$x^3+a_2x^2+a_1x+a_0\in\mathbb{F}[x]$有重因式的条件.
B2. 证明: 如果a是$f'''(x)$的一个k重根, 则a是

$$g(x)=\frac{x-a}{2}[f'(x)+f'(a)]-f(x)+f(a)$$

的一个$(k+3)$重根.
B3. 设$n>m>0$, 证明: $x^n+ax^{n-m}+b$不能有不为0的重数大于2的根.
B4. 设$f(x)\in\mathbb{F}[x]$是n次多项式, 其中$n\geqslant 1$. 证明: $f'(x)\mid f(x)$的充分必要条件是, $f(x)$和一个一次因式的n次幂相伴.
B5. 设n是正整数, 求$2n-1$次多项式$f(x)$, 使得$f(x)+1$被$(x-1)^n$整除, $f(x)-1$被$(x+1)^n$整除.

§ 1.10 复系数多项式的因式分解

设$f(x)\in\mathbb{F}[x]$, 且$\deg(f(x))\geqslant 1$, 则$f(x)$有标准分解

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x).$$

前面我们对$f(x)$的不可约因式$p_i(x)$的重数r_i进行了讨论, 现在我们具体看如何求出分解式(或者分解式长成什么样子). 为此, 当然希望知道$\mathbb{F}[x]$中哪些多项式不可约. 但是多项式的不可约性与数域有关, 为此需要对不同的数域分别讨论. 本节考虑复数域, 后面两节依次讨论实数域和有理数域.

本节中取$\mathbb{F}=\mathbb{C}$, 考虑复系数多项式的因式分解.

1.10.1 复数的指数(复习)

设$i=\sqrt{-1}$是虚数单位, 即是$x^2+1=0$的一个根, 则复数有形式$z=a+bi$, 其中$a,b\in\mathbb{R}$. 称a为z的实部, 记为$\Re(z)$, 而b称为z的虚部, 记为$\Im(z)$. 可以定义$z=a+bi$的共轭复数

$$\overline{z}=a-bi,$$

则

$$a=\Re(z)=\frac{z+\overline{z}}{2},\quad b=\Im(z)=\frac{z-\overline{z}}{2i},$$

且

$$z \in \mathbb{R} \Longleftrightarrow z = \overline{z}.$$

容易证明: 对任意$z, w \in \mathbb{C}$, 成立

$$\overline{z+w} = \overline{z} + \overline{w}, \quad \overline{zw} = \overline{z}\,\overline{w}.$$

复数$z = a + bi$的模长$|z|$定义为

$$|z| := \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}.$$

我们有

$$z\overline{z} = a^2 + b^2 = |z|^2.$$

在几何上, $z = a + bi$和平面$\mathbb{R}^2$上的点(a, b)一一对应, 其中(a, b)和原点的距离就是$|z|$, 而(a, b)和正实轴的夹角φ是幅角(见图1.1). 于是可得

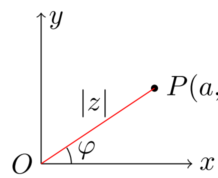


图 1.1: 复数$z = a + bi$的几何表示

$$a = |z|\cos\varphi, \quad b = |z|\sin\varphi.$$

进而就得到复数的三角表示

$$z = a + bi = |z|(\cos\varphi + i\sin\varphi).$$

设$z_1 = |z_1|(\cos\varphi_1 + i\sin\varphi_1)$和$z_2 = |z_2|(\cos\varphi_2 + i\sin\varphi_2)$是两个复数, 则有

$$\begin{aligned} z_1 z_2 &= |z_1|(\cos\varphi_1 + i\sin\varphi_1) \cdot |z_2|(\cos\varphi_2 + i\sin\varphi_2) \\ &= |z_1||z_2|[(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2)] \\ &= |z_1||z_2|(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)). \end{aligned}$$

于是对于复数$z = |z|(\cos\varphi + i\sin\varphi)$和$n \in \mathbb{N}$, 有

$$z^n = |z|^n(\cos n\varphi + i\sin n\varphi).$$

设$z^{\frac{1}{n}} = w = |w|(\cos\theta + i\sin\theta)$, 则$z = w^n$得到

$$|z|(\cos\varphi + i\sin\varphi) = |w|^n(\cos n\theta + i\sin n\theta).$$

进而有

$$|w|^n = |z|, \quad n\theta = \varphi + 2k\pi, \quad (k \in \mathbb{Z}),$$

即

$$|w| = |z|^{\frac{1}{n}}, \quad \theta = \frac{\varphi + 2k\pi}{n}, \quad (k \in \mathbb{Z}).$$

考虑到三角函数的周期性, 得到

$$z^{\frac{1}{n}} = |z|^{\frac{1}{n}} \left(\cos \frac{\varphi + 2k\pi}{n} + i \sin \frac{\varphi + 2k\pi}{n} \right), \quad (k = 0, 1, \dots, n-1).$$

也可以用指数形式来计算复数的指数. 对于$\varphi \in \mathbb{R}$, 我们有

$$\begin{aligned} e^{i\varphi} &= \sum_{n=0}^{\infty} \frac{i^n}{n!} \varphi^n = \sum_{n=0}^{\infty} (-1)^n \frac{\varphi^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{\varphi^{2n+1}}{(2n+1)!} \\ &= \cos \varphi + i \sin \varphi. \end{aligned}$$

于是得到复数的指数形式

$$z = |z|(\cos \varphi + i \sin \varphi) = |z| e^{i\varphi}.$$

比如令$z = -1$, 则有著名公式$e^{i\pi} = -1$, 或者

$$e^{i\pi} + 1 = 0.$$

这个公式包含了数学中的五个常数: $0, 1, \pi, e$和i.

使用指数形式, 可以用指数函数的性质计算复数的乘法, 例如当$z_1 = |z_1| e^{i\varphi_1}$, $z_2 = |z_2| e^{i\varphi_2}$时, 有

$$z_1 z_2 = |z_1| e^{i\varphi_1} \cdot |z_2| e^{i\varphi_2} = |z_1||z_2| e^{i(\varphi_1 + \varphi_2)}.$$

于是对于$z = |z| e^{i\varphi}$和$n \in \mathbb{N}$, 我们有

$$z^n = |z|^n e^{in\varphi}$$

和

$$z^{\frac{1}{n}} = |z|^{\frac{1}{n}} e^{i \frac{\varphi + 2k\pi}{n}}, \quad k = 0, 1, \dots, n-1.$$

我们看两个例子.

例1.22. 在$\mathbb{C}$上解方程$x^n = 1$.

解. 方程的解为$1^{\frac{1}{n}}$. 由于

$$1^{\frac{1}{n}} = e^{\frac{2k\pi}{n} i}, \quad (k = 0, 1, \dots, n-1),$$

所以如果记

$$\zeta_n = e^{\frac{2\pi}{n} i} = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n},$$

则方程的所有解为$1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1}$. 称$\zeta_n$为一个$n$次本原单位根. □

例1.23. 计算$i^{\frac{1}{2}}$和$(-i)^{\frac{1}{2}}$.

解. 我们有

$$i^{\frac{1}{2}} = (e^{\frac{\pi}{2} i})^{\frac{1}{2}} = e^{i \frac{\frac{\pi}{2} + 2k\pi}{2}} \overset{k=0,1}{=\!=\!=} \pm e^{\frac{\pi}{4} i} = \pm \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i \right)$$

和

$$(-i)^{\frac{1}{2}} = (e^{\frac{3\pi}{2} i})^{\frac{1}{2}} = e^{i \frac{\frac{3\pi}{2} + 2k\pi}{2}} \overset{k=0,1}{=\!=\!=} \pm e^{\frac{3\pi}{4} i} = \pm \left(-\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i \right).$$

解毕. □

1.10.2 复系数多项式的因式分解

我们先确定$\mathbb{C}[x]$中的不可约多项式. 显然, 一次多项式都是不可约的. 下面说明, 这就是$\mathbb{C}[x]$中所有的不可约多项式. 这需要一个基本结论[21]: 代数基本定理.

定理1.18 (代数基本定理)**.** 设$f(x) \in \mathbb{C}[x]$且$\deg(f(x)) \geqslant 1$, 则$f(x)$在$\mathbb{C}$中至少有一个根.

现在设$p(x) \in \mathbb{C}[x]$不可约, 则$\deg(p(x)) \geqslant 1$. 由代数基本定理, 存在$c \in \mathbb{C}$, 使得$p(c) = 0$. 利用因式定理有

$$x - c \mid p(x).$$

由于$p(x)$不可约, 所以$p(x) \sim x - c$. 则$p(x)$是一次多项式. 于是我们就得到复系数多项式因式分解的基本定理.

定理1.19. (1) 设$p(x) \in \mathbb{C}[x]$, 则

$$p(x)\text{ 不可约} \iff p(x)\text{是一次的;}$$

(2) (复系数多项式的因式分解) 任意次数$\geqslant 1$的复系数多项式在$\mathbb{C}$中可以唯一分解成一次因式的乘积;

(3) 任意n次复系数多项式有n个复根(重根计算重数).

证明. (1) 已证.
(2) 由(1)和因式分解存在唯一定理.
(3) 由(2). □

由上面的结果, 对任意$f(x) \in \mathbb{C}[x]$, $\deg(f(x)) \geqslant 1$, 其标准分解式有形式

$$f(x) = c(x - c_1)^{r_1}(x - c_2)^{r_2} \cdots (x - c_s)^{r_s},$$

其中$c \in \mathbb{C}^\times$, $c_1, c_2, \ldots, c_s \in \mathbb{C}$是两两不同的复数, $r_1, r_2, \ldots, r_s \in \mathbb{N}$. 则$c_i$是$f(x)$的$r_i$重根, 于是求$f(x)$在$\mathbb{C}[x]$中的因式分解, 等价于求$f(x)$的所有复根(包括重数).

例1.24. 求$f(x) = x^8 - 1$在$\mathbb{C}[x]$中的标准分解式.

解. 我们有

$$\begin{aligned} f(x) =&(x^4+1)(x^4-1) = (x^2+i)(x^2-i)(x^2+1)(x^2-1) \\ =&\left(x + \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right)\left(x - \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right)\left(x + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right)\left(x - \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right) \\ &\times (x+i)(x-i)(x+1)(x-1), \end{aligned}$$

其中我们利用了例1.23的计算结果. 当然也可以直接算$1^{\frac{1}{8}}$得到$x^8 = 1$的所有根. □

最后, 我们介绍根与系数关系的**Vièta公式**, 并做一些历史注记.

[21]从实数到复数, 我们失去了一些东西, 比如与实数不同, 并不是任意两个复数可以比较大小; 但是我们又得到了一些东西, 任意代数方程一定有根, 特别地, $x^2 + 1 = 0$在复数范围可以求解.

设$\mathbb{F}$是数域, 多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{F}[x],\quad (n\in\mathbb{N},a_n\neq 0).$$

将$f(x)$看成复系数多项式, 设它的n个复根为$c_1,c_2,\ldots,c_n$. 则$f(x)$在$\mathbb{C}$上的因式分解为

$$f(x)=a_n(x-c_1)(x-c_2)\cdots(x-c_n).$$

将右边展开并比较系数, 就得到Vièta公式

$$\begin{cases} c_1+c_2+\cdots+c_n=-\frac{a_{n-1}}{a_n},\\ \sum\limits_{1\leqslant i<j\leqslant n}c_ic_j=\frac{a_{n-2}}{a_n},\\ \qquad\vdots\\ \sum\limits_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}c_{i_1}c_{i_2}\cdots c_{i_k}=(-1)^k\frac{a_{n-k}}{a_n},\\ \qquad\vdots\\ c_0c_1\cdots c_n=(-1)^n\frac{a_0}{a_n}. \end{cases}$$

对于n次多项式$f(x)$, 我们知道$f(x)$有n个复根. 那么如何求出这些根, 有没有求根公式? 这些问题是古典代数学的基本问题之一.

- 当$n=1$时, 容易求解.
- 当$n=2$时, 有求根公式: $ax^2+bx+c=0$的根为
$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$
- 当$n=3,4$时也是根式可解的[22](即根可以由系数经过有限步加, 减, 乘, 除及开方运算表示).
- 当$n\geqslant 5$时, 是否也根式可解? 22岁的Abel证明了此时没有求根公式, 即不是根式可解的; 而Galois则在不超过21岁时给出了一个具体的多项式根式可解的充要条件, 更重要的是他提出了群和Galois扩张等新的概念, 这可以看成是近世代数的开端. Galios理论是数学中最优美的理论之一.

习题1.10

A1. 求下面多项式在$\mathbb{C}[x]$中的标准分解式:

(1) x^4-2;　　(2) x^n-1;　　(3) x^n+1.

A2. 设$p(x)\in\mathbb{F}[x]$是不可约多项式, 并且存在一个$a\neq 0$使得$p(a)=0$, $p\left(\frac{1}{a}\right)=0$. 证明: 对于任意$b\neq 0$, 如果$p(b)=0$, 则$p\left(\frac{1}{b}\right)=0$.

A3. 设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$的$n$个复根为$x_1,x_2,\ldots,x_n$, 且$x_i\neq 0$, $i=1,2,\ldots,n$. 求以$\frac{1}{x_1}$, $\frac{1}{x_2}$, $\ldots$, $\frac{1}{x_n}$为根的n次多项式(系数用$a_0,a_1,\ldots,a_n$表示).

[22]参看本章附录.

A4. (1) 设整数$m,n,l\geqslant 0$, 证明: 在$\mathbb{Q}[x]$中有

$$x^2+x+1 \mid x^{3m}+x^{3n+1}+x^{3l+2};$$

(2) 设整数$m_1,m_2,m_3,m_4,m_5\geqslant 0$, 证明: 在$\mathbb{Q}[x]$中有

$$x^4+x^3+x^2+x+1 \mid x^{5m_1}+x^{5m_2+1}+x^{5m_3+2}+x^{5m_4+3}+x^{5m_5+4};$$

(3) 设$f(x)=\sum\limits_{i=1}^{n}x^{a_i}$, 其中$a_1,a_2,\ldots,a_n$都为非负整数. 求$f(x)$被$x^2+x+1$整除的充分和必要条件.

A5. 设$1,\omega_1,\ldots,\omega_{n-1}$是$x^n-1$的全部不同的$n$个复根, 证明:

$$(1-\omega_1)(1-\omega_2)\cdots(1-\omega_{n-1})=n.$$

A6. 证明: $\cos\frac{\pi}{7}-\cos\frac{2\pi}{7}+\cos\frac{3\pi}{7}=\frac{1}{2}$.

B1. 求下面多项式在$\mathbb{C}[x]$中的标准分解式[23]:

(1) $x^{2n}+x^n+1$;

(2) $(x+\cos\theta+i\sin\theta)^n+(x+\cos\theta-i\sin\theta)^n$;

(3) $(x-1)^n+(x+1)^n$;

(4) $x^n-\binom{2n}{2}x^{n-1}+\binom{2n}{4}x^{n-2}+\cdots+(-1)^n\binom{2n}{2n}$;

(5) $x^{2n}+\binom{2n}{2}x^{2n-2}(x^2-1)+\binom{2n}{4}x^{2n-4}(x^2-1)^2+\cdots+(x^2-1)^n$;

(6) $x^{2n+1}+\binom{2n+1}{2}x^{2n-1}(x^2-1)+\binom{2n+1}{4}x^{2n-3}(x^2-1)^2+\cdots+\binom{2n+1}{2n}x(x^2-1)^n$.

B2. 如果复数$\zeta\in\mathbb{C}$满足$\zeta^n=1$, 而对任意$1\leqslant l<n$, 有$\zeta^l\neq 1$, 则称ζ 是一个n次本原单位根. 例如, $\zeta_n=e^{\frac{2\pi i}{n}}$就是一个$n$次本原单位根. 设$\zeta$是一个$n$次本原单位根, $m,k\in\mathbb{N}$, 证明:

(1) $\zeta^m=1 \iff n\mid m$;

(2) ζ^k是$\frac{n}{(n,k)}$次本原单位根;

(3) ζ^k是n次本原单位根的充分必要条件是$(n,k)=1$.

B3. 判断15次单位根各是多少次本原单位根, 并且这些单位根各是$x^{15}-1$在有理数域上的分解式

$$x^{15}-1=(x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)(x^8-x^7+x^5-x^4+x^3-x+1)$$

中哪个因式的根?

B4. 设$n\in\mathbb{N}$, 令$\Phi_n(x)$是恰以所有的n次本原单位根为单根的首一多项式, 即(参看上面习题)

$$\Phi_n(x)=\prod_{\substack{1\leqslant k\leqslant n\\(k,n)=1}}\left(x-\zeta_n^k\right),\qquad (\zeta_n=e^{\frac{2\pi i}{n}}).$$

称$\Phi_n(x)$为n次分圆多项式. 证明:

(1) $\Phi_n(x)$是整系数多项式(即所有的系数都是整数);

(2) $x^n-1=\prod\limits_{1\leqslant d\mid n}\Phi_d(x)$.

[23]这里, $\binom{n}{k}:=\frac{n!}{k!(n-k)!}$是二项式系数, 我们有下面的二项展开公式

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^ky^{n-k}.$$

§ 1.11 实系数多项式的因式分解

本节考虑实系数多项式的因式分解, 与复系数多项式一样, 我们先要考察不可约实系数多项式. 一次多项式当然不可约, 下面考察二次实系数多项式的不可约性. 设$f(x) = ax^2 + bx + c$是二次实系数多项式, 如果$f(x)$可约, 则$f(x)$有一次因式; 反之当然成立. 于是

$$f(x)\text{可约} \Longleftrightarrow f(x)\text{有一次因式} \Longleftrightarrow f(x)\text{有实根}.$$

而$f(x)$的两个(复)根为

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

所以

$$f(x)\text{可约} \Longleftrightarrow b^2 - 4ac \geqslant 0.$$

这样, 我们就得到$f(x)$不可约的充分必要条件是$f(x)$的判别式$\Delta = b^2 - 4ac < 0$. 于是, 与复系数多项式不同, 除了一次多项式不可约外, 还存在二次的实系数不可约多项式, 例如, 多项式$f(x) = x^2 + 1$. 那么, 还有其它的实系数不可约多项式吗?

看三次实系数多项式$f(x) = ax^3 + bx^2 + cx + d$, 其中$a \neq 0$. 不妨设$a > 0$, 则

$$\lim_{x \to +\infty} f(x) = +\infty, \quad \lim_{x \to -\infty} f(x) = -\infty.$$

所以由连续性, 存在$c \in \mathbb{R}$, 使得$f(c) = 0$. 因此$f(x)$有一次因式, 进而可约. 我们就证明了所有的三次实系数多项式都可约.

类似可以证明, 奇数次的实系数多项式一定有一个实根, 进而当次数$\geqslant 3$时都可约. 那么次数$\geqslant 4$的偶数次实系数多项式呢?

从二次多项式$f(x)$的求根公式可以看出, 如果$f(x)$不可约, 则$f(x)$的两个根互为共轭复数. 一般地, 我们有下面的观察. 设

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x],$$

如果$c \in \mathbb{C}$是$f(x)$的一个复根, 即$f(c) = 0$, 则

$$a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0 = 0.$$

取共轭得

$$\overline{a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0} = \overline{0} = 0,$$

即

$$a_n \overline{c}^n + a_{n-1} \overline{c}^{n-1} + \cdots + a_1 \overline{c} + a_0 = 0.$$

所以$f(\overline{c}) = 0$, 即$\overline{c}$也是$f(x)$的根.

引理1.20. 设$f(x) \in \mathbb{R}[x]$, $c \in \mathbb{C}$满足$f(c) = 0$, 则$f(\overline{c}) = 0$.

引理1.20表明实系数多项式的虚部不为零的复根共轭成对出现, 而且可以证明其重数也相同.

例1.25. (1) 设$f(x)\in\mathbb{R}[x]$, $c\in\mathbb{C}$为$f(x)$的k重根, 证明: $\overline{c}$也为$f(x)$的k重根;

(2) 求$f(x)\in\mathbb{R}[x]$, 使其以$2,i,1+i$为根, 其中$2,i$为二重根, 且$f(x)$的次数尽可能低.

☞ **解.** (1) 因为$c\in\mathbb{C}$为$f(x)$的k重根, 所以

$$f(c)=f'(c)=\cdots=f^{(k-1)}(c)=0,\quad f^{(k)}(c)\neq 0.$$

对实系数多项式$f(x),f'(x),\ldots,f^{(k-1)}(x)$用引理1.20, 得到

$$f(\overline{c})=f'(\overline{c})=\cdots=f^{(k-1)}(\overline{c})=0.$$

如果$f^{(k)}(\overline{c})=0$, 则对$f^{(k)}(x)\in\mathbb{R}[x]$用引理1.20, 得到

$$f^{(k)}(c)=f^{(k)}(\overline{\overline{c}})=0,$$

矛盾. 所以$f^{(k)}(\overline{c})\neq 0$, 进而$\overline{c}$为$f(x)$的$k$重根.
(2) $f(x)$还有根$-i,1-i$, 且$-i$是二重根. 于是次数最低的$f(x)$是

$$\begin{aligned}f(x)=&(x-2)^2[(x+i)(x-i)]^2(x-1-i)(x-1+i)\\=&(x-2)^2(x^2+1)^2(x^2-2x+2).\end{aligned}$$

解毕. □

下面用引理1.20来考察实系数不可约多项式. 假设$p(x)\in\mathbb{R}[x]$是首一不可约的, 且$\deg(p(x))>1$. 由代数基本定理, 存在$c\in\mathbb{C}$, 使得$p(c)=0$. 如果$c\in\mathbb{R}$, 则由因式定理, $p(x)$有一次因式$x-c$, 这矛盾于$p(x)$的不可约性. 所以$c\notin\mathbb{R}$, 进而$c\neq\overline{c}$. 由引理1.20, $\overline{c}$也是$p(x)$的复根, 于是在$\mathbb{C}$上有分解

$$p(x)=(x-c)(x-\overline{c})g(x),\quad(\exists g(x)\in\mathbb{C}[x]).$$

但是

$$(x-c)(x-\overline{c})=x^2-(c+\overline{c})x+c\overline{c}\in\mathbb{R}[x],$$

所以也有$g(x)\in\mathbb{R}[x]$, 即有$\mathbb{R}$上的分解式

$$p(x)=(x^2-(c+\overline{c})x+c\overline{c})g(x).$$

因为$p(x)$不可约, 所以$g(x)=1$, 于是

$$p(x)=x^2+px+q,\quad p=-(c+\overline{c}),\ q=c\overline{c}\in\mathbb{R}.$$

而且

$$p^2-4q=(c+\overline{c})^2-4c\overline{c}=(c-\overline{c})^2<0.$$

即$p(x)$是判别式小于零的二次多项式.

于是我们就找到了所有的不可约实系数多项式, 进而得到实系数多项式的因式分解定理.

定理1.21. (1) 设$0\neq p(x)\in\mathbb{R}[x]$是首一多项式, 则

$$p(x)\text{不可约}\iff\begin{array}{l}p(x)=x-c\ (\exists c\in\mathbb{R})\quad\text{或者}\\p(x)=x^2+px+q\ (\exists p,q\in\mathbb{R},p^2-4q<0);\end{array}$$

(2) (实系数多项式的因式分解) 任意次数$\geqslant 1$的实系数多项式在$\mathbb{R}$上可以唯一地分解成一次因式和二次不可约因式的乘积.

由上面的定理, 对于$f(x)\in\mathbb{R}[x]$, $\deg(f(x))\geqslant 1$, 其标准分解式有形式

$$f(x)=c(x-c_1)^{l_1}\cdots(x-c_s)^{l_s}(x^2+p_1x+q_1)^{k_1}\cdots(x^2+p_rx+q_r)^{k_r},$$

其中$c_1,\ldots,c_s\in\mathbb{R}$互不相等, $(p_1,q_1),\ldots,(p_r,q_r)$为互不相等的实数对, 满足

$$p_i^2-4q_i<0,\quad i=1,\ldots,r,$$

且$l_1,\ldots,l_s,k_1,\ldots,k_r\in\mathbb{N}$.

习题1.11

A1. 求下面多项式在实数域上的标准分解式
(1) x^4+1;　(2) x^6+27;
(3) $x^4+4x^3+4x^2+1$;　(4) x^4-ax^2+1, $-2<a<2$.
A2. 证明: 奇数次实系数多项式一定有实根.
A3. 设实系数多项式$f(x)=x^3+a_2x^2+a_1x+a_0$的3个复根都是实数, 证明: $a_2^2\geqslant 3a_1$.
B1. 求下面多项式在实数域上的标准分解式:
(1) $x^{2n}-2x^n+2$;　(2) $x^{2n}+x^n+1$;
(3) x^n-1;　(4) x^n+1;
(5) x^n-a^n, $(a\in\mathbb{R}^\times)$;　(6) x^n+a^n, $(a\in\mathbb{R}^\times)$.
B2. 设$m\in\mathbb{N}$, 证明:
(1) $\prod\limits_{k=1}^{m-1}\sin\frac{k\pi}{2m}=\frac{\sqrt{m}}{2^{m-1}}$;　(2) $\prod\limits_{k=1}^{m-1}\cos\frac{k\pi}{2m}=\frac{\sqrt{m}}{2^{m-1}}$;
(3) $\prod\limits_{k=1}^{m}\sin\frac{(2k-1)\pi}{2(2m+1)}=\frac{1}{2^m}$;　(4) $\prod\limits_{k=1}^{m}\cos\frac{k\pi}{2m+1}=\frac{1}{2^m}$;
(5) $\prod\limits_{k=1}^{m}\sin\frac{k\pi}{2m+1}=\frac{\sqrt{2m+1}}{2^m}$;　(6) $\prod\limits_{k=1}^{m}\cos\frac{(2k-1)\pi}{2(2m+1)}=\frac{\sqrt{2m+1}}{2^m}$;
(7) $\prod\limits_{k=1}^{m}\sin\frac{(2k-1)\pi}{4m}=\frac{\sqrt{2}}{2^m}$;　(8) $\prod\limits_{k=1}^{m}\cos\frac{(2k-1)\pi}{4m}=\frac{\sqrt{2}}{2^m}$.
B3. 证明: 实系数多项式$f(x)$对所有实数x恒取正值的充分必要条件是, 存在复系数多项式$\varphi(x)$, $\varphi(x)$没有实数根, 使得$f(x)=|\varphi(x)|^2$.
B4. 设$f(x)\in\mathbb{R}[x]$, 如果对于任意$a\in\mathbb{R}$都有$f(a)\geqslant 0$, 证明: 存在$\varphi(x),\psi(x)\in\mathbb{R}[x]$, 使得$f(x)=\varphi^2(x)+\psi^2(x)$.

§ 1.12　有理系数多项式的因式分解

从前两节可以看出, 复系数以及实系数不可约多项式都比较简单, 复系数和实系数多项式因式分解理论上都有漂亮的结果. 那么有理系数多项式呢? 有理数有其容易处理的地方, 比如它们和整数紧密相关, 可以用上整数的结论和办法; 但是从函数的观点看, 有理数其实很难处理, 比如取极限不封闭. 前面复系数和实系数多项式的因式分解理论的基石是代数基本定理, 在这里却没有这样类似的

结论. 于是可以预见, 有理系数多项式的因式分解的情况将复杂很多. 事实上, 我们没有统一的办法判定一个有理系数多项式是否可约; 于是更不能写出有理系数多项式因式分解的好的定理.

所以, 我们只能介绍人们得到的关于有理系数多项式因式分解的一些结果. 我们主要做三件事. 首先, 我们给出有理系数(事实上为整系数)多项式的有理根的求法, 这个可用来判断二次和三次有理系数多项式是否可约; 接着, 我们证明一个整系数多项式在$\mathbb{Q}$上的不可约性等价于其在$\mathbb{Z}$上的不可约性; 最后给出判别整系数多项式不可约的一种常用的判别法: Eisenstein判别法. 从后面的讨论可以看到, 这里有浓重的数论色彩. 事实上, 有理数(整数)是数论关心的研究对象.

为此, 先引入一个记号. 我们记$\mathbb{Z}[x]$为所有整系数多项式做成的集合, 即

$$\mathbb{Z}[x] = \{a_n x^n + \cdots + a_1 x + a_0 \mid n \in \mathbb{Z}_{\geqslant 0}, a_n, \ldots, a_1, a_0 \in \mathbb{Z}\}.$$

在$\mathbb{Z}[x]$中可以类似地定义加法和乘法, 在这两个运算下, $\mathbb{Z}[x]$成为有单位元的交换环. 有包含关系$\mathbb{Z}[x] \subset \mathbb{Q}[x]$. 还要注意的是, 在$\mathbb{Z}[x]$中一般不能进行带余除法. 但是当$g(x) \in \mathbb{Z}[x]$是首一多项式时, 容易得到, 对任意$f(x) \in \mathbb{Z}[x]$, 存在唯一一对多项式$q(x), r(x) \in \mathbb{Z}[x]$, 满足

$$f(x) = q(x)g(x) + r(x), \qquad \deg(r(x)) < \deg(g(x)).$$

1.12.1 有理系数多项式的有理根

本小节考虑有理系数多项式的有理根, 这是一个数论问题. 首先, 我们可以将系数整数化. 设$0 \neq f(x) \in \mathbb{Q}[x]$, 我们要求$f(x)$的有理根. 存在非零整数$a$, 使得$af(x) \in \mathbb{Z}[x]$. 对于任意的$c \in \mathbb{Q}$, 显然有

$$f(c) = 0 \Longleftrightarrow af(c) = 0.$$

于是只需考虑整系数多项式的有理根, 我们有下面的结果.

命题1.22. *设$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$, $n \in \mathbb{N}$, $a_n \neq 0$, 设$c = \frac{r}{s} \in \mathbb{Q}$为$f(x)$的根, 其中$r, s \in \mathbb{Z}$且$(r, s) = 1$, 则有*

$$s \mid a_n, \qquad r \mid a_0.$$

特别地, 若$a_n = 1$, 则$\alpha \in \mathbb{Z}$且$\alpha \mid a_0$.

☞ **证明.** 由于$f(c) = 0$, 所以

$$a_n \frac{r^n}{s^n} + a_{n-1}\frac{r^{n-1}}{s^{n-1}} + \cdots + a_1 \frac{r}{s} + a_0 = 0.$$

两边乘s^n, 得到

$$a_n r^n + a_{n-1} r^{n-1} s + \cdots + a_1 r s^{n-1} + a_0 s^n = 0.$$

于是有

$$s \mid a_n r^n, \qquad r \mid a_0 s^n.$$

而$(r, s) = 1$, 则有$s \mid a_n$和$r \mid a_0$[24]. □

例1.26. *求方程$2x^4 - 2x^3 + x - 3 = 0$的全部有理解.*

[24]这里用了性质: 如果整数r和s互素, 而$r \mid sa$, 其中$a \in \mathbb{Z}$, 则$r \mid a$.

☞ **解.** 记$f(x)=2x^4-2x^3+x-3$, 则$a_n=2$, 因子有±1, ±2; 而$a_0=-3$, 因子有±1, ±3. 于是有理根只可能为±1, ±3, $\pm\frac{1}{2}$, $\pm\frac{3}{2}$. 直接计算可得

$$\begin{array}{llll} f(1)=-2, & f(3)=108, & f\left(\frac{1}{2}\right)=-\frac{21}{8}, & f\left(\frac{3}{2}\right)=\frac{15}{8}, \\ f(-1)=0, & f(-3)=210, & f\left(-\frac{1}{2}\right)=-\frac{25}{8}, & f\left(-\frac{3}{2}\right)=\frac{99}{8}, \end{array}$$

所以该方程的有理解有$x=-1$. □

例1.27. **(1)** 设$f(x)\in\mathbb{Q}[x]$, 且$\deg(f(x))=2$或3, 证明:

$$f(x)\text{在}\mathbb{Q}\text{上不可约}\iff f(x)\text{无有理根};$$

(2) 证明: $f(x)=x^3-5x+1$在$\mathbb{Q}$上不可约.

☞ **证明.** (1) 由于$\deg(f(x))=2$或3, 所以$f(x)$在$\mathbb{Q}$上可约的充分必要条件是, $f(x)$在$\mathbb{Q}[x]$中有一次因式, 而这又等价于$f(x)$有有理根.

(2) 多项式$f(x)$的有理根只可能为±1, 而

$$f(1)=-3,\quad f(-1)=5,$$

所以$f(x)$没有有理根. 则由(1)知$f(x)$在$\mathbb{Q}$上不可约. □

注意, 当$\deg(f(x))\geqslant4$时, 如果$f(x)$在$\mathbb{Q}$上不可约, 则$f(x)$无有理根; 但是如果$f(x)$没有有理根, 那么只得到$f(x)$无一次因式, 并不能得到$f(x)$在$\mathbb{Q}$上不可约, 因为$f(x)$可以有高次因式. 例如, $f(x)=(x^2-2)^2$ 是可约的, 但是它没有有理根.

1.12.2 $\mathbb{Q}[x]$中可约性与$\mathbb{Z}[x]$中可约性的等价性

本小节考虑有理系数多项式的不可约性. 任取$f(x)\in\mathbb{Q}[x]$, $\deg(f(x))\geqslant1$, 则存在$0\neq a\in\mathbb{Z}$, 使得$af(x)\in\mathbb{Z}[x]$. 容易知道

$$f(x)\text{在}\mathbb{Q}[x]\text{中可约}\iff af(x)\text{在}\mathbb{Q}[x]\text{中可约}.$$

于是只需考虑整系数多项式在$\mathbb{Q}[x]$中的不可约性.

类似于数域上的多项式, 我们可以定义整系数多项式在$\mathbb{Z}[x]$中的不可约性.

定义1.15. 设$f(x)\in\mathbb{Z}[x]$, $\deg(f(x))\geqslant1$. 如果$f(x)$不能写成两个次数更小的整系数多项式的乘积, 则称$f(x)$(在$\mathbb{Z}[x]$中)不可约. 否则称其可约.

任取$f(x)\in\mathbb{Z}[x]$, 由定义得

$$f(x)\text{在}\mathbb{Z}[x]\text{中可约}\Longrightarrow f(x)\text{在}\mathbb{Q}[x]\text{中可约}.$$

本小节的主要结果是上面结果的逆命题也成立.

定理1.23. 设$f(x)\in\mathbb{Z}[x]$, $\deg(f(x))\geqslant1$, 则

$$f(x)\text{在}\mathbb{Z}[x]\text{中可约}\iff f(x)\text{在}\mathbb{Q}[x]\text{中可约}.$$

定理1.23表明, 一个整系数多项式在$\mathbb{Q}[x]$中不可约当且仅当它在$\mathbb{Z}[x]$中不可约. 为了证明定理1.23中的充分性, 我们需要一些准备, 下面的处理极富数论色彩.

定义1.16. 设$f(x)=a_nx^n+\cdots+a_1x+a_0\in\mathbb{Z}[x]$, $n\geqslant 1$, $a_n\neq 0$, 称$f(x)$为本原多项式, 如果系数$a_0,a_1,\ldots,a_n$互素(即: $a_0,a_1,\ldots,a_n$的公因子只有± 1).

下面的结果表明, 有理系数多项式和本原多项式密切相关.

引理1.24. 设$f(x)\in\mathbb{Q}[x]$, $\deg(f(x))\geqslant 1$.

(1) 存在$r\in\mathbb{Q}$, 存在本原多项式$g(x)\in\mathbb{Z}[x]$, 使得$f(x)=rg(x)$. 如果$f(x)\in\mathbb{Z}[x]$, 则可取$r\in\mathbb{Z}$;

(2) 如果$f(x)=r_1g_1(x)=r_2g_2(x)$, 其中$r_1,r_2\in\mathbb{Q}$, $g_1(x),g_2(x)\in\mathbb{Z}[x]$是本原多项式, 则

$$r_1=r_2,\ g_1(x)=g_2(x)\quad \text{或者}\quad r_1=-r_2,\ g_1(x)=-g_2(x).$$

☞ **证明.** (1) 记$f(x)=a_nx^n+\cdots+a_1x+a_0$, $a_n\neq 0$, $a_i\in\mathbb{Q}$. 设

$$a_i=\frac{b_i}{c_i},\quad b_i,c_i\in\mathbb{Z},c_i\neq 0,$$

取$c_0,c_1,\ldots,c_n$的一个非零公倍数c(例如, $c=c_0c_1\cdots c_n$), 则

$$cf(x)=(ca_n)x^n+\cdots+(ca_1)x+(ca_0)\in\mathbb{Z}[x].$$

令b为$ca_0,ca_1,\ldots,ca_n$的最大公因子, 则$ca_i/b\in\mathbb{Z}$, 且$ca_0/b,ca_1/b,\ldots,ca_n/b$互素[25], 于是

$$\frac{c}{b}f(x)=\frac{ca_n}{b}x^n+\cdots+\frac{ca_1}{b}x+\frac{ca_0}{b}\in\mathbb{Z}[x]$$

是本原多项式, 将其记为$g(x)$. 取$r=\frac{b}{c}\in\mathbb{Q}$, 则$f(x)=rg(x)$.

(2) 记

$$r_1=\frac{p_1}{q_1},\ r_2=\frac{p_2}{q_2},\quad p_1,p_2,q_1,q_2\in\mathbb{Z},\ q_1,q_2\neq 0,$$

则有

$$p_1q_2g_1(x)=p_2q_1g_2(x).$$

比较两边系数的最大公因子[26], 得到

$$p_1q_2=\pm p_2q_1\Longrightarrow r_1=\pm r_2.$$

进而有$g_1(x)=\pm g_2(x)$. □

关于本原多项式, 有下面著名的Gauss引理.

引理1.25 (Gauss引理)**.** 两个本原多项式的乘积仍本原.

---

[25]这里使用了性质: 如果$d=(a_1,a_2,\ldots,a_n)$, 则

$$\left(\frac{a_1}{d},\frac{a_2}{d},\ldots,\frac{a_n}{d}\right)=1.$$

[26]这里使用了性质: 设$d\in\mathbb{N}$, 则

$$d(a_1,a_2,\ldots,a_n)=(da_1,da_2,\ldots,da_n).$$

☞ **证明.** 设

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{j=0}^{m}b_jx^j\in\mathbb{Z}[x]$$

是本原多项式, 记

$$h(x)=f(x)g(x)=\sum_{k=0}^{m+n}c_kx^k,$$

则

$$c_k=\sum_{\substack{i,j\geqslant 0\\ i+j=k}}a_ib_j.$$

如果$h(x)$不是本原多项式, 则存在素数p, 使得

$$p\mid c_k,\quad k=0,1,\ldots,m+n.$$

由于$f(x)$本原, 所以存在i $(0\leqslant i\leqslant n)$, 使得

$$p\mid a_0,\ldots,p\mid a_{i-1},p\nmid a_i.$$

类似地, $g(x)$本原导出存在j $(0\leqslant j\leqslant m)$, 使得

$$p\mid b_0,\ldots,p\mid b_{j-1},p\nmid b_j.$$

考查c_{i+j}, 由于

$$c_{i+j}=a_0b_{i+j}+\cdots+a_{i-1}b_{j+1}+a_ib_j+a_{i+1}b_{j-1}+\cdots+a_{i+j}b_0,$$

所以$p\mid a_ib_j$. 这得到$p\mid a_i$或者$p\mid b_j$[27], 矛盾. 所以$h(x)$是本原多项式. □

下面给出定理1.23的证明.

☞ **定理1.23的证明.** 只需证明充分性. 假设$f(x)$在$\mathbb{Q}[x]$中可约, 则存在$g(x),h(x)\in\mathbb{Q}[x]$，使得

$$f(x)=g(x)h(x),\quad \deg(g(x)),\deg(h(x))<\deg(f(x)).$$

由引理1.24, 存在$a\in\mathbb{Z}$, 存在$r,s\in\mathbb{Q}$, 存在$f_1(x),g_1(x),h_1(x)\in\mathbb{Z}[x]$是本原多项式, 使得

$$f(x)=af_1(x),\quad g(x)=rg_1(x),\quad h(x)=sh_1(x).$$

于是有

$$af_1(x)=rsg_1(x)h_1(x).$$

由Gauss引理, $g_1(x)h_1(x)$是本原多项式. 再用引理1.24, 就得到

$$rs=\pm a,\quad g_1(x)h_1(x)=\pm f_1(x).$$

特别有$rs\in\mathbb{Z}$. 所以

$$f(x)=(rsg_1(x))h_1(x)$$

是$f(x)$在$\mathbb{Z}[x]$中的一个非平凡分解, 即$f(x)$在$\mathbb{Z}[x]$中可约. □

类似可以证明, 如果$f(x),g(x)\in\mathbb{Z}[x]$, 其中$g(x)$本原, 且

$$f(x)=g(x)h(x),\quad h(x)\in\mathbb{Q}[x],$$

则有$h(x)\in\mathbb{Z}[x]$.

[27] 如果p是素数, $p\mid ab$, 则$p\mid a$或者$p\mid b$.

1.12.3 Eisenstein判别法

本节最后, 我们给出判别整系数多项式不可约性常用的一种判别法: **Eisenstein判别法**.

定理1.26 (Eisenstein判别法)**.** *设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{Z}[x]$, 其中$n\geqslant 1$. 如果存在素数$p$, 满足*

(i) $p\nmid a_n$,

(ii) $p\mid a_i$, $i=0,1,\ldots,n-1$,

(iii) $p^2\nmid a_0$,

则$f(x)$在$\mathbb{Q}[x]$中不可约.

☞ **证明.** 如果$f(x)$在$\mathbb{Q}$上可约, 由定理1.23, 存在$g(x),h(x)\in\mathbb{Z}[x]$, 使得

$$f(x)=g(x)h(x),\quad \deg(g(x)),\deg(h(x))<n.$$

记

$$g(x)=b_lx^l+\cdots+b_0,\quad b_l\neq 0,l<n,\qquad h(x)=c_mx^m+\cdots+c_0,\quad c_m\neq 0,m<n,$$

则$n=l+m$且$a_n=b_lc_m$, $a_0=b_0c_0$. 由于$p\mid b_0c_0$, 所以$p\mid b_0$或者$p\mid c_0$. 下面不妨设$p\mid b_0$, 而$p^2\nmid a_0$, 所以$p\nmid c_0$. 由于$p\nmid a_n$, 所以$p\nmid b_l$. 于是存在k, $1\leqslant k\leqslant l$, 使得

$$p\mid b_0,p\mid b_1,\ldots,p\mid b_{k-1},\quad 而\quad p\nmid b_k.$$

因为

$$a_k=b_kc_0+b_{k-1}c_1+\cdots+b_1c_{k-1}+b_0c_k,\quad p\mid a_k,$$

所以$p\mid b_kc_0$. 这得到$p\mid b_k$或者$p\mid c_0$, 矛盾. 所以$f(x)$不可约. □

下面是一些例子.

例1.28. *设$f(x)=x^n+2$, $n\in\mathbb{N}$, 取$p=2$, 由Eisenstein判别法, 可知$f(x)$在$\mathbb{Q}$上不可约. 于是$\mathbb{Q}[x]$中有任意正次数的不可约多项式.*

例1.29. *证明: $f(x)=x^6+x^3+1$在$\mathbb{Q}$上不可约.*

☞ **证明.** 本题不能直接用Eisenstein判别法, 但是有如下处理. 令$x=y+1$, 有

$$\begin{aligned}g(y)=&f(y+1)=(y+1)^6+(y+1)^3+1\\=&y^6+6y^5+15y^4+21y^3+18y^2+9y+3.\end{aligned}$$

取素数$p=3$, 由Eisenstein判别法, 可知$g(y)$在$\mathbb{Q}$上不可约. 如果$f(x)$在$\mathbb{Q}$上可约, 则

$$f(x)=f_1(x)f_2(x),\quad \exists f_1(x),f_2(x)\in\mathbb{Q}[x],\ \deg(f_1(x)),\deg(f_2(x))<6.$$

这得到

$$g(y)=f(y+1)=f_1(y+1)f_2(y+1),$$

于是$g(y)$也可约, 矛盾. 所以$f(x)$在$\mathbb{Q}$上不可约. □

例1.30. 设p是一个素数, 证明: p次分圆多项式

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$$

在$\mathbb{Q}$上不可约.

☞ **证明.** 记$f(y) = \Phi_p(y+1)$. 由于

$$\Phi_p(x) = \frac{x^p - 1}{x - 1},$$

所以得到

$$f(y) = \frac{(y+1)^p - 1}{y} = \sum_{i=1}^{p} \binom{p}{i} y^{i-1} = y^{p-1} + py^{p-2} + \binom{p}{p-2} y^{p-3} + \cdots + \binom{p}{2} y + p.$$

对于$1 \leqslant i \leqslant p-1$, 有

$$\binom{p}{i} = \frac{p(p-1)\cdots(p-i+1)}{i!} = p\frac{(p-1)\cdots(p-i+1)}{i!} \in \mathbb{Z}.$$

这得到

$$p \mid \binom{p}{i}, \quad (1 \leqslant i \leqslant p-1).$$

而$p \nmid 1$且$p^2 \nmid p$, 于是由Eisenstein判别法, $f(y)$在$\mathbb{Q}$上不可约. 进而得$\Phi_p(x)$在$\mathbb{Q}$上不可约. □

例1.31 (Schur定理的特殊情形)**.** 证明: 当n是素数时, 多项式

$$f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

在有理数域上不可约.

☞ **证明.** 对整系数多项式$n!f(x)$用Eisenstein判别法即可(取素数n). □

最后给出一个不好用Eisenstein判别法判别不可约性的一个例子.

例1.32. 判别多项式$f(x) = x^4 + 3x + 1$在$\mathbb{Q}$上的可约性.

☞ **解.** 由于$f(x)$的有理根只有± 1, 且

$$f(1) = 5, \quad f(-1) = -1,$$

所以$f(x)$没有一次因式. 如果$f(x)$在$\mathbb{Q}$上可约, 则有分解

$$f(x) = (a_2x^2 + a_1x + a_0)(b_2x^2 + b_1x + b_0), \quad (\exists a_2, a_1, a_0, b_2, b_1, b_0 \in \mathbb{Z}, a_2, b_2 \neq 0).$$

比较上式的首项系数得$a_2b_2 = 1$. 于是$a_2 = b_2 = \pm 1$. 不妨设$a_2 = b_2 = 1$. 于是

$$\begin{aligned} f(x) =& (x^2 + a_1x + a_0)(x^2 + b_1x + b_0) \\ =& x^4 + (a_1 + b_1)x^3 + (a_0 + a_1b_1 + b_0)x^2 + (a_0b_1 + a_1b_0)x + a_0b_0, \end{aligned}$$

这得到

$$\begin{cases} a_1 + b_1 = 0, \\ a_0 + a_1b_1 + b_0 = 0, \\ a_0b_1 + a_1b_0 = 3, \\ a_0b_0 = 1. \end{cases}$$

由最后一式得$a_0 = b_0 = \pm 1$, 结合第三式得到$a_1 + b_1 = \pm 3$, 这矛盾于第一式. 所以$f(x)$在$\mathbb{Q}$上不可约. □

习题1.12

A1. 求下列多项式的有理根.

(1) $x^3-6x^2+15x-14$; (2) $4x^4-7x^2-5x-1$;

(3) $x^5+x^4-6x^3-14x^2-11x-3$.

A2. 下列多项式在有理数域上是否可约?

(1) x^3+x^2-3x+2; (2) $x^4-4x^3+12x^2-6x-2$;

(3) x^4+1; (4) $4x^4-8x^3+12x^2+2$;

(5) x^4-2x^3+2x-3; (6) x^5+5x^3+1;

(7) x^p+px+1, p为奇素数; (8) x^p+px^r+1, p为奇素数, $0\leqslant r\leqslant p$;

(9) $x^4+4kx+1$, k为整数; (10) x^4-5x+1.

A3. 设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{Z}[x]$, 其中$n\geqslant 1$. 如果存在素数$p$, 满足

(i) $p\nmid a_0$, (ii) $p\mid a_i$, $i=1,\ldots,n$, (iii) $p^2\nmid a_n$,

证明: $f(x)$在$\mathbb{Q}[x]$中不可约.

A4. 设$p_1,p_2,\ldots,p_s$是两两互不相同的素数, $n>1$是整数, 证明: $\sqrt[n]{p_1p_2\cdots p_s}$是无理数.

A5. 设$m,n\in\mathbb{N}$且$m<n$, $f(x)\in\mathbb{Q}[x]$是m次多项式, $p_1,p_2,\ldots,p_s$是两两不等的素数. 证明: $\sqrt[n]{p_1p_2\cdots p_s}$不是$f(x)$ 的实根.

A6. 证明: $\cos 20°$是无理数.

A7. (1) 设$f(x),g(x)\in\mathbb{Z}[x]$, 在$\mathbb{Q}[x]$中有$g(x)\mid f(x)$, 记$g(x)$除$f(x)$的商为$q(x)\in\mathbb{Q}[x]$. 证明: 如果$g(x)$是本原多项式, 那么$q(x)\in\mathbb{Z}[x]$;

(2) 设$f(x)$是整系数多项式, 既约分数$\frac{q}{p}$是$f(x)$的根, 证明: 存在整系数多项式$g(x)$, 使得$f(x)=(px-q)g(x)$.

B1. 设$f(x)$是一个整系数多项式, 证明: 如果$f(0)$和$f(1)$都是奇数, 则$f(x)$不能有整数根.

B2. 设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是整系数多项式, 证明: 如果$f(0)$, $f(1)$和$f(-1)$都不能被3整除, 则$f(x)$没有整数根.

B3. 设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是整系数多项式, 证明: 如果$a_n$和$a_0$都是奇数, 而且$f(1)$与$f(-1)$中至少有一个是奇数, 则$f(x)$没有有理根.

B4. 设$p(x)\in\mathbb{Q}[x]$是不可约多项式, $\deg(p(x))>1$是奇数, α_1和α_2是$p(x)$两个不同的复根. 证明: $\alpha_1+\alpha_2$不是有理数.

B5. 设$a_1,a_2,\ldots,a_n$是两两不等的整数. 证明:

(1) $f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)-1$在$\mathbb{Q}$上不可约;

(2) $g(x)=(x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2+1$在$\mathbb{Q}$上不可约.

B6. 设$a_1,a_2,\ldots,a_n$是两两不等的整数, $f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)+1$.

(1) 证明: 当n是奇数或者是$\geqslant 6$的偶数时, $f(x)$在$\mathbb{Q}$上不可约;

(2) 当$n=2$或4时, $f(x)$是否在$\mathbb{Q}$上不可约?

B7. 设$f(x)=x^4+px^2+q\in\mathbb{Q}[x]$, 证明: $f(x)$在$\mathbb{Q}$上不可约的充分必要条件是, p^2-4q不是有理数的平方, 或者q不是有理数的平方, 且$\pm 2\sqrt{q}-p$ 不是有理数的平方.

B8. (1) 在$\mathbb{Q}[x]$中将$x^8+x^7+x^6+x^5+x^4+x^3+x^2+x+1$因式分解;

(2) 设$n>1$是整数, $f(x)=\sum\limits_{i=0}^{n-1}x^i$. 证明: 如果$n$不是素数, 则$f(x)$在$\mathbb{Q}$上可约.

B9. 设$f(x)=a_nx^n+\cdots+a_1x+a_0\in\mathbb{Z}[x]$, $n\geqslant 1$, $k\leqslant n$是正整数. 如果存在素数p, 满足

(i) $p\mid a_i$, $i=0,1,\ldots,k-1$,　(ii) $p\nmid a_k$, $p\nmid a_n$,　(iii) $p^2\nmid a_0$,

证明: $f(x)$有一个次数大于或等于k的在$\mathbb{Q}$上不可约的整系数因式.

B10. 设$f(x)=a_{2n+1}x^{2n+1}+\cdots+a_1x+a_0\in\mathbb{Z}[x]$. 证明: 如果存在素数$p$, 满足

(i) $p\nmid a_{2n+1}$,　(ii) $p^2\mid a_i$, $i=0,1,\ldots,n$,　(iii) $p\mid a_j$, $j=n+1,n+2,\ldots,2n$,　(iv) $p^3\nmid a_0$,

那么, $f(x)$在$\mathbb{Q}$上不可约.

B11. 设$n\in\mathbb{N}$, 证明: n次分圆多项式$\Phi_n(x)$在$\mathbb{Q}$上不可约, 从而

$$x^n-1=\prod_{1\leqslant d\mid n}\Phi_d(x)$$

是x^n-1在$\mathbb{Q}$上的标准分解式.

C1. 设$n\in\mathbb{N}$, 证明: 多项式$f(x)=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}$在$\mathbb{Q}[x]$中不可约[28].

§ 1.13　多元多项式

前面我们介绍了一元多项式的一些理论, 而世界是多元的, 所以很自然地要考虑多个不定元的情形. 本节将不定元的个数从一个推广到任意有限个, 介绍多元多项式的一些基本概念.

设$n\in\mathbb{N}$, $x_1,x_2,\ldots,x_n$是n个无关的不定元.

定义1.17. (1) 称

$$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},\quad a\in\mathbb{F},i_1,i_2,\ldots,i_n\in\mathbb{Z}_{\geqslant 0}$$

为数域$\mathbb{F}$上的一个n元单项式. 称a为这个单项式的系数, 而$i_1+\cdots+i_n$为这个单项式的次数;

(2) 如果两个单项式

$$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},\quad bx_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$$

中相同不定元的幂相等, 即

$$i_1=j_1,i_2=j_2,\ldots,i_n=j_n,$$

则称它们为同类项.

将有限个单项式的形式和称为多项式, 具体地

定义1.18. 称数域$\mathbb{F}$上有限个两两不是同类项的n元单项式的形式和

$$f(x_1,x_2,\ldots,x_n)=\sum_{i_1,i_2,\ldots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

为$\mathbb{F}$上的一个n元多项式.

---

[28]一般的, I. Schur在1929年证明了[13]: 对$c_i\in\mathbb{Z}$, 多项式

$$1+c_1x+c_2\frac{x^2}{2!}+\cdots+c_{n-1}\frac{x^{n-1}}{(n-1)!}\pm\frac{x^n}{n!}$$

在$\mathbb{Q}$上不可约.

一元多项式的许多概念可以推广到多元多项式. 设

$$f(x_1,x_2,\ldots,x_n)=\sum_{i_1,i_2,\ldots,i_n} a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

是一个n元多项式, 有时为了简单起见, 我们把$f(x_1,x_2,\ldots,x_n)$简记为f.

- 称$a_{i_1i_2\cdots i_n}$为项$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$的系数;
- 称系数不为零的各项次数的最大值为$f(x_1,x_2,\ldots,x_n)$的次数, 记为$\deg f(x_1,x_2,\ldots,x_n)$. 也简记为$\deg f$;
- 如果每一项的系数都为零, 则称$f(x_1,x_2,\ldots,x_n)$为零多项式, 记为0. 约定$\deg 0=-\infty$.

数域$\mathbb{F}$上的所有n元多项式组成的集合记为$\mathbb{F}[x_1,x_2,\ldots,x_n]$. 我们可以按照一元多项式一样在这个集合中定义运算. 首先, 对于两个多项式$f(x_1,x_2,\ldots,x_n)$和$g(x_1,x_2,\ldots,x_n)$, 如果它们的所有同类项的系数对应相当, 则称它们相等, 记为$f(x_1,x_2,\ldots,x_n)=g(x_1,x_2,\ldots,x_n)$. 其次, 可以按照我们熟悉的方式定义加法, 减法和乘法: 设

$$\begin{aligned}f(x_1,x_2,\ldots,x_n)&=\sum_{i_1,i_2,\ldots,i_n} a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\in\mathbb{F}[x_1,x_2,\ldots,x_n],\\ g(x_1,x_2,\ldots,x_n)&=\sum_{j_1,j_2,\ldots,j_n} b_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}\in\mathbb{F}[x_1,x_2,\ldots,x_n],\end{aligned}$$

(1) 将$f(x_1,x_2,\ldots,x_n)$和$g(x_1,x_2,\ldots,x_n)$的同类项的系数分别相加(减), 得到f和g的和(差), 记为

$$f(x_1,x_2,\ldots,x_n)\pm g(x_1,x_2,\ldots,x_n);$$

(2) 将$f(x_1,x_2,\ldots,x_n)$的每一项和$g(x_1,x_2,\ldots,x_n)$的每一项如下相乘

$$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\cdot b_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}=a_{i_1i_2\cdots i_n}b_{j_1j_2\cdots j_n}x_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n},$$

再把所有这样相乘得到的项相加并合并同类项, 得到的多项式称为f和g的积, 记为

$$f(x_1,x_2,\ldots,x_n)\cdot g(x_1,x_2,\ldots,x_n)=f(x_1,x_2,\ldots,x_n)g(x_1,x_2,\ldots,x_n).$$

容易验证, 在上面的运算下, $\mathbb{F}[x_1,x_2,\ldots,x_n]$满足类似于$\mathbb{F}[x]$满足的运算律(A1-A4), (M1-M3)和(D), 成为有单位元1的交换环. 称$\mathbb{F}[x_1,x_2,\ldots,x_n]$为$\mathbb{F}$上的$n$元多项式环.

我们知道, 在$\mathbb{F}[x]$中, 两个非零多项式的积仍然非零(即$\mathbb{F}[x]$无零因子, 是整环), 而且两个多项式的和(差)以及积的次数满足一些关系式, 这些都可以推广到多元多项式环.

引理1.27. *设$f(x_1,x_2,\ldots,x_n),g(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$.*

(1) *如果f和g都非零, 则$f(x_1,x_2,\ldots,x_n)g(x_1,x_2,\ldots,x_n)\neq 0$;*

(2) *成立*

$$\begin{aligned}&\deg(f(x_1,x_2,\ldots,x_n)\pm g(x_1,x_2,\ldots,x_n))\leqslant\max\{\deg f,\deg g\},\\ &\deg(f(x_1,x_2,\ldots,x_n)g(x_1,x_2,\ldots,x_n))=\deg f+\deg g.\end{aligned}$$

由定义, 容易证明引理1.27(2)的第一个式子成立, 为了证明其它的结论, 我们需要在单项式之间引入序, 定义首项的概念; 还需要介绍齐次多项式的概念.

定义1.19. **(1)** 对于任意两个由非负整数组成的n元有序数组$(i_1, i_2, \ldots, i_n)$和$(j_1, j_2, \ldots, j_n)$, 如果存在$s(1 \leqslant s \leqslant n)$, 满足

$$i_1 = j_1, i_2 = j_2, \ldots, i_{s-1} = j_{s-1}, i_s > j_s,$$

则称$(i_1, i_2, \ldots, i_n)$大于$(j_1, j_2, \ldots, j_n)$, 记为$(i_1, i_2, \ldots, i_n) > (j_1, j_2, \ldots, j_n)$. 称这种序为字典序;

(2) 如果$(i_1, i_2, \ldots, i_n) > (j_1, j_2, \ldots, j_n)$, 则称单项式$a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$比$b_{j_1 j_2 \cdots j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}$大;

(3) 称非零多项式的所有非零项中最大者为该多项式的首项, 首项的系数称为该多项式的首项系数.

例如, 设$f(x_1, x_2, x_3) = 4x_1x_2x_3 + x_3^3 + x_1^2 + x_1x_3 + x_2$是一个3元3次多项式, 将它的各项按从大到小的次序排列为

$$f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2x_3 + x_1x_3 + x_2 + x_3^3.$$

请注意, 首项不一定具有最大的次数.

☞ **引理1.27(1)的证明.** 我们只要证明: 两个非零多项式的首项的乘积是这两个多项式的积的首项. 事实上, 设$f(x_1, x_2, \ldots, x_n)$的首项是$ax_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}$ $(a \neq 0)$, 而$g(x_1, x_2, \ldots, x_n)$的首项是$bx_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}$ $(b \neq 0)$. 对满足

$$(p_1, p_2, \ldots, p_n) > (i_1, i_2, \ldots, i_n), \quad (q_1, q_2, \ldots, q_n) > (j_1, j_2, \ldots, j_n)$$

的非负整数n元数组$(i_1, i_2, \ldots, i_n)$和$(j_1, j_2, \ldots, j_n)$, 容易知道下面成立

$$\begin{aligned}
(p_1 + q_1, p_2 + q_2, \ldots, p_n + q_n) &> (p_1 + j_1, p_2 + j_2, \ldots, p_n + j_n), \\
(p_1 + q_1, p_2 + q_2, \ldots, p_n + q_n) &> (i_1 + q_1, i_2 + q_2, \ldots, i_n + q_n), \\
(p_1 + q_1, p_2 + q_2, \ldots, p_n + q_n) &> (i_1 + q_1, i_2 + q_2, \ldots, i_n + q_n) \\
&> (i_1 + j_1, i_2 + j_2, \ldots, i_n + j_n).
\end{aligned}$$

于是$abx_1^{p_1+q_1} x_2^{p_2+q_2} \cdots x_n^{p_n+q_n}$比$f(x_1, x_2, \ldots, x_n)g(x_1, x_2, \ldots, x_n)$中其它项都大, 是多项式积的首项. □

多元多项式除了按照字典序排列各项外, 还可以按照各项的次数来排列.

定义1.20. 设$f(x_1, x_2, \ldots, x_n)$是$\mathbb{F}$上的一个非零多项式.

(1) 如果所有非零项的次数都等于m, 则称$f(x_1, x_2, \ldots, x_n)$是一个m次齐次多项式. 我们约定零多项式可以看成任意次数的齐次多项式;

(2) 多项式$f(x_1, x_2, \ldots, x_n)$可以唯一写为

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{m} f_i(x_1, x_2, \ldots, x_n),$$

其中$f_i(x_1, x_2, \ldots, x_n)$是$i$次齐次多项式, 称它为$f(x_1, x_2, \ldots, x_n)$的$i$次齐次成分.

容易证明, 两个齐次多项式的积仍是齐次多项式, 且次数等于这两个多项式次数之和.

☞ **引理1.27(2)的证明.** 我们只要证明第二个等式, 且可以不妨设这两个多项式都是非零多项式. 设$\deg f = m$而$\deg g = s$, 将这两个多项式写为

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{m} f_i(x_1, x_2, \ldots, x_n), \quad g(x_1, x_2, \ldots, x_n) = \sum_{j=0}^{s} g_j(x_1, x_2, \ldots, x_n),$$

其中$f_i(x_1, x_2, \ldots, x_n)$是$i$次齐次多项式, 而$g_j(x_1, x_2, \ldots, x_n)$是$j$次齐次多项式. 这得到

$$\begin{aligned} & f(x_1, x_2, \ldots, x_n)g(x_1, x_2, \ldots, x_n) = f_m(x_1, x_2, \ldots, x_n)g_s(x_1, x_2, \ldots, x_n) \\ & \quad + [f_m(x_1, x_2, \ldots, x_n)g_{s-1}(x_1, x_2, \ldots, x_n) + f_{m-1}(x_1, x_2, \ldots, x_n)g_s(x_1, x_2, \ldots, x_n)] \\ & \quad + \cdots + f_0(x_1, x_2, \ldots, x_n)g_0(x_1, x_2, \ldots, x_n). \end{aligned}$$

由于$f_m(x_1, x_2, \ldots, x_n), g_s(x_1, x_2, \ldots, x_n) \neq 0$, 所以它们的积是非零的$m+s$次齐次多项式. 因此

$$\deg(f(x_1, x_2, \ldots, x_n)g(x_1, x_2, \ldots, x_n)) = m + s,$$

证毕. □

与$\mathbb{F}[x]$不同, 对于$\mathbb{F}[x_1, x_2, \ldots, x_n]$中的两个非零多项式, 一般不能作带余除法. 但是与$\mathbb{F}[x]$类似, 在$\mathbb{F}[x_1, x_2, \ldots, x_n]$中也可定义整除, 最大公因式, 不可约多项式等概念, 并且可证$\mathbb{F}[x_1, x_2, \ldots, x_n]$中的非常数多项式可以唯一分解为不可约多项式的乘积. 我们把这些留作练习.

处理多元多项式的一种常用的办法是把它看成某个不定元的一元多项式, 这时的系数是多项式. 例如, 任意n元多项式$f(x_1, x_2, \ldots, x_n)$可唯一写成

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) = & a_m(x_2, \ldots, x_n)x_1^m + a_{m-1}(x_2, \ldots, x_n)x_1^{m-1} \\ & + \cdots + a_1(x_2, \ldots, x_n)x_1 + a_0(x_2, \ldots, x_n), \end{aligned}$$

其中$a_i(x_2, \ldots, x_n) \in \mathbb{F}[x_2, \ldots, x_n]$. 用集合的记号写即为

$$\mathbb{F}[x_1, x_2, \ldots, x_n] = (\mathbb{F}[x_2, \ldots, x_n])[x_1].$$

将多元多项式看成一元多项式后, 就可以利用一元多项式的某些结论来解决问题, 例如可以重新如下证明$\mathbb{F}[x_1, \ldots, x_n]$没有零因子.

☞ **引理1.27(1)的另证.** 对n归纳. 当$n = 1$时已经证明. 设$n > 1$且$n-1$时成立. 看成x_1的多项式, 有

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) = & a_m(x_2, \ldots, x_n)x_1^m + a_{m-1}(x_2, \ldots, x_n)x_1^{m-1} \\ & + \cdots + a_1(x_2, \ldots, x_n)x_1 + a_0(x_2, \ldots, x_n), \\ g(x_1, x_2, \ldots, x_n) = & b_l(x_2, \ldots, x_n)x_1^l + a_{l-1}(x_2, \ldots, x_n)x_1^{l-1} \\ & + \cdots + b_1(x_2, \ldots, x_n)x_1 + b_0(x_2, \ldots, x_n), \end{aligned}$$

其中$m, l \geqslant 0$, $a_i(x_2, \ldots, x_n), b_j(x_2, \ldots, x_n) \in \mathbb{F}[x_2, \ldots, x_n]$, 且

$$a_m(x_2, \ldots, x_n) \neq 0, \quad b_l(x_2, \ldots, x_n) \neq 0.$$

于是

$$fg = a_m b_l x_1^{m+l} + (x_1\text{的低次项}),$$

而由归纳假设, $a_m b_l \neq 0$, 所以$fg \neq 0$. □

要注意的是, 将多元多项式看成某个不定元的一元多项式时, 这时的系数是多项式, 不一定可以作除法. 如果除式是首一多项式, 则可以做带余除法.

命题1.28. 设$D = \mathbb{Z}$或者$\mathbb{F}[x_1, \ldots, x_n]$, $f(x) = a_m x^m + \cdots + a_1 x + a_0 \in D[x]$, $a \in D$. 则存在唯一的商$q(x) \in D[x]$和余式$r = f(a) \in D$, 使得

$$f(x) = q(x)(x-a) + r.$$

特别有

$$x - a \mid f(x) \Longleftrightarrow f(a) = 0.$$

☞ **证明.** 类似于带余除法, 余式定理和因式定理的证明. 请注意, 虽然对$D[x]$中的任意两个非零多项式一般不能作带余除法, 但这里$x-a$的首项系数是一. □

例1.33. 分别在有理数域和复数域上分解因式$f(x,y,z) = x^3 + y^3 + z^3 - 3xyz$.

☞ **解.** 将$f(x,y,z)$看成x的多项式

$$g(x) = f(x,y,z) = x^3 - 3yzx + y^3 + z^3.$$

由于

$$g(-y-z) = (-y-z)^3 - 3yz(-y-z) + y^3 + z^3 = 0,$$

所以

$$x + y + z \mid g(x).$$

作带余除法

$$\begin{array}{r|l}
 & x^2 - (y+z)x + (y^2 - yz + z^2) \\ \hline
x+y+z & x^3 - 3yzx + y^3 + z^3 \\
 & \underline{x^3 + (y+z)x^2} \\
 & -(y+z)x^2 - 3yzx + y^3 + z^3 \\
 & \underline{-(y+z)x^2 - (y+z)^2 x} \\
 & (y^2 - yz + z^2)x + y^3 + z^3 \\
 & \underline{(y^2 - yz + z^2)x + y^3 + z^3} \\
 & 0
\end{array}$$

可得

$$f(x,y,z) = (x+y+z)(x^2 - (y+z)x + y^2 - yz + z^2).$$

再用配方法分解二次因式

$$\begin{aligned}
& x^2 - (y+z)x + y^2 - yz + z^2 \\
= & \left[x - \frac{1}{2}(y+z)\right]^2 - \frac{1}{4}(y+z)^2 + y^2 - yz + z^2 \\
= & \left[x - \frac{1}{2}(y+z)\right]^2 + \frac{3}{4}(y-z)^2 \\
= & \left[x - \frac{1}{2}(y+z) + \frac{\sqrt{3}i}{2}(y-z)\right]\left[x - \frac{1}{2}(y+z) - \frac{\sqrt{3}i}{2}(y-z)\right]
\end{aligned}$$

$$=(x+\omega y+\omega^2 z)(x+\omega^2 y+\omega z),$$

其中$\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i=e^{\frac{2\pi i}{3}}$. 于是

$$f(x,y,z)=(x+y+z)(x+\omega y+\omega^2 z)(x+\omega^2 y+\omega z)$$

就是$f(x,y,z)$在复数域上的分解式.

由于$x+\omega y+\omega^2 z$和$x+\omega^2 y+\omega z$都不是有理系数的, 所以

$$x^2-(y+z)x+y^2-yz+z^2=x^2+y^2+z^2-xy-yz-zx$$

在有理数域上不可约. 这得到

$$f(x,y,z)=(x+y+z)(x^2+y^2+z^2-xy-yz-zx)$$

是$f(x,y,z)$在有理数域上的分解式. □

最后我们说明将多元多项式看成多元函数其身份不变.

例1.34. *设$f(x_1,x_2,\dots,x_n),g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 证明: f和g作为多项式相等的充分必要条件是, 对任意$a_1,a_2,\dots,a_n\in\mathbb{F}$, 有*

$$f(a_1,a_2,\dots,a_n)=g(a_1,a_2,\dots,a_n).$$

☞ **证明.** 只要证明充分性. 设f和g作为多元函数相等. 令

$$h(x_1,x_2,\dots,x_n)=f(x_1,x_2,\dots,x_n)-g(x_1,x_2,\dots,x_n),$$

如果h不是零多项式, 则由本节习题, 存在$a_1,a_2,\dots,a_n\in\mathbb{F}$, 使得$h(a_1,a_2,\dots,a_n)\neq 0$. 这得到

$$f(a_1,a_2,\dots,a_n)\neq g(a_1,a_2,\dots,a_n),$$

矛盾. 于是h是零多项式, f和g作为多项式相等. □

习题1.13

A1. 将下面三元多项式的各项按照字典序从大到小的次序重新排列, 并写出次数最高的齐次成分:

(1) $f(x_1,x_2,x_3)=4x_1x_2^5x_3^2+5x_1^2x_2x_3-x_1^3x_3^4+x_1^3x_2+x_1x_2^4x_3^3$;

(2) $f(x_1,x_2,x_3)=x_1^2x_2^3+x_2^2x_3^3+x_1^3x_3^2+x_1^4+x_1^2x_2^4+x_2^2x_3^4$.

A2. 设$0\neq f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, $m\in\mathbb{N}$, 证明: $f(x_1,x_2,\dots,x_n)$是m次齐次多项式的充分必要条件是, 对任意$t\in\mathbb{F}$, 成立

$$f(tx_1,tx_2,\dots,tx_n)=t^m f(x_1,x_2,\dots,x_n).$$

A3. 设$f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$是非零多项式, 证明: f是非零函数, 即存在$b_1,b_2,\dots,b_n\in\mathbb{F}$, 使得$f(b_1,b_2,\dots,b_n)\neq 0$.

A4. 设$f(x_1,x_2,\dots,x_n),g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 且$g(x_1,x_2,\dots,x_n)$是非零多项式. 如果对任意满足$g(a_1,a_2,\dots,a_n)\neq 0$的$a_1,a_2,\dots,a_n\in\mathbb{F}$, 都成立$f(a_1,a_2,\dots,a_n)=0$, 证明: $f(x_1,x_2,\dots,x_n)$是零多项式.

A5. 设$f(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 如果存在$g(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 使得

$$f(x_1,x_2,\ldots,x_n)g(x_1,x_2,\ldots,x_n)=1,$$

则称$f(x_1,x_2,\ldots,x_n)$是可逆元. 多项式环$\mathbb{F}[x_1,x_2,\ldots,x_n]$中所有可逆元的全体做成的集合记为$\mathbb{F}[x_1,x_2,\ldots,x_n]^\times$, 证明: $\mathbb{F}[x_1,x_2,\ldots,x_n]^\times=\mathbb{F}^\times$.

A6. 设$f,g\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 如果存在$h\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 使得$f=gh$, 则称$g$整除$f$, 记为$g\mid f$, 否则称$g$不整除$f$, 记为$g\nmid f$. 当$g\mid f$时, 称$g$是$f$的一个因式, 而称$f$是$g$的一个倍式. 如果$f\mid g$且$g\mid f$, 则称$f$和$g$相伴, 记为$f\sim g$. 证明: $f\sim g$当且仅当存在$c\in\mathbb{F}^\times$, 使得$f=cg$.

A7. 设$p\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, $\deg p>0$. 如果p可以写为两个次数更低的多项式的乘积, 则称p是$\mathbb{F}$上的可约多项式, 否则称p是$\mathbb{F}$上的不可约多项式. 证明: p是$\mathbb{F}$上的不可约多项式的充分必要条件是, p的因式只有$\mathbb{F}$中的非零数, 以及它的相伴元.

A8. 分别在实数域和复数域上考虑下面多项式的不可约性

(1) $f(x,y)=x^2+y^2-1$;

(2) $f(x,y)=x^2-y$;

(3) $f(x,y)=x^3+y$;

(4) $f(x,y)=x^2-2xy+y^2+y$;

(5) $f(x,y,z)=x^2+y^2+z^2$;

(6) $f(x,y,z)=x^2+y^2+z^2+xy+yz+zx$.

A9. 在复数域上分解因式: $f(x,y,z)=-x^3-y^3-z^3+x^2(y+z)+y^2(x+z)+z^2(x+y)-2xyz$.

B1. 设非零多项式$f,g\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 如果$f$和$g$的最大公因式是$\mathbb{F}$中的非零数, 则称$f$和$g$互素. 设$n>1$, 是否成立命题: 对非零多项式$f,g\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, f和g互素的充分必要条件是, 存在$u,v\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 使得$uf+vg=1$?

B2. (唯一分解定理) 设多项式$f(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 且$\deg f\geqslant 1$, 证明: f可以分解为$\mathbb{F}[x_1,x_2,\ldots,x_n]$中有限个不可约多项式的乘积; 并且, 如果不计常数因子以及乘积中因式的次序, 这些不可约多项式由f唯一确定[29].

B3. 设$f,g\in\mathbb{F}[x_1,x_2,\ldots,x_n]$不全为零, 如果$d\in\mathbb{F}[x_1,x_2,\ldots,x_n]$满足: d是f和g的公因式, 且对f和g的任意公因式h, 有$d\mid h$, 则称d是f和g的一个最大公因式. 证明: f和g的最大公因式存在.

§ 1.14 对称多项式

在多元多项式中, 经常遇到的是所谓的对称多项式. 通俗地讲, 对称多项式就是该多项式中不定元的地位等同, 可如下严格定义.

定义1.21. 设$f(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$, 如果对自然数$1,2,\ldots,n$的任意排列$i_1$, i_2, $\ldots$, i_n, 都有

$$f(x_{i_1},x_{i_2},\ldots,x_{i_n})=f(x_1,x_2,\ldots,x_n),$$

[29]这是抽象代数中的Gauss定理的特例, Gauss定理断言: 如果D是唯一分解整环, 则多项式环$D[x]$也是唯一分解整环. 证明的一个方法是对n归纳. 假设$D=\mathbb{F}[x_1,\ldots,x_{n-1}]$中唯一分解成立, 则可以定义最大公因式, 以及本元多项式等概念, 并证明类似Gauss引理成立. 类似于从$\mathbb{Z}$到$\mathbb{Q}$, 可以从D得到它的分式域$K=\mathbb{F}(x_1,\ldots,x_{n-1})=\{f/g\mid f,g\in D,g\neq 0\}$, 然后证明$K[x_n]$中带余除法成立, 从而$K[x_n]$中唯一分解成立. 进而类似$\mathbb{Z}[x]$和$\mathbb{Q}[x]$中多项式可约性的等价性的处理办法, 可以证明$D[x_n]$中唯一分解成立.

则称$f(x_1,x_2,\dots,x_n)$是一个n元对称多项式.

可以证明[30], 多项式$f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$是对称多项式的充分和必要条件是, 对任意$1\leqslant i<j\leqslant n$, 成立

$$f(x_1,\dots,x_i,\dots,x_j,\dots,x_n)=f(x_1,\dots,x_j,\dots,x_i,\dots,x_n).$$

由定义, $\mathbb{F}$中的数是对称多项式. 且容易看出

$$\begin{aligned}
\sigma_1&=x_1+x_2+\cdots+x_n=\sum_{i=1}^{n}x_i,\\
\sigma_2&=x_1x_2+x_1x_3+\cdots+x_1x_n+\cdots+x_{n-1}x_n=\sum_{1\leqslant i_1<i_2\leqslant n}x_{i_1}x_{i_2},\\
&\ \ \vdots\\
\sigma_k&=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}x_{i_1}x_{i_2}\cdots x_{i_k},\\
&\ \ \vdots\\
\sigma_n&=x_1x_2\cdots x_n
\end{aligned}$$

都是n元对称多项式. 称它们为n元基本对称多项式, 也称为n元初等对称多项式. 关于多项式

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=(x-c_1)(x-c_2)\cdots(x-c_n)$$

的根与系数关系的Vièta公式即为

$$a_k=(-1)^{n-k}\sigma_{n-k}(c_1,c_2,\dots,c_n),\quad (k=0,1,\dots,n-1).$$

容易验证, 两个n元对称多项式的和, 差, 积仍是n元对称多项式[31]. 而且, 如果f_1, f_2, $\dots$, f_m是m个n元对称多项式, 则对任意$g\in\mathbb{F}[y_1,y_2,\dots,y_m]$, 用$f_i$替换$y_i$得到的多项式$g(f_1,f_2,\dots,f_m)$仍是一个关于不定元$x_1,x_2,\dots,x_n$的对称多项式. 特别地, 取$m=n$, 将$y_i$用$\sigma_i$替换后得到的多项式$g(\sigma_1,\sigma_2,\dots,\sigma_n)$是对称多项式. 下面证明每个$n$元对称多项式都有这种表达式.

定理1.29 (对称多项式基本定理)**.** *设$f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$是$n$元对称多项式, 则$f$可唯一表示为初等对称多项式$\sigma_1,\sigma_2,\dots,\sigma_n$的多项式, 即存在唯一的$g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 使得*

$$f(x_1,x_2,\dots,x_n)=g(\sigma_1,\sigma_2,\dots,\sigma_n).$$

☞ **证明.** 先证明存在性. 设$f\neq 0$, f的首项是$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$, $a\neq 0$.

断言. $i_1\geqslant i_2\geqslant\cdots\geqslant i_n$.

否则, 存在k, 使得$i_k<i_{k+1}$. 这时, 有

$$(i_1,\dots,i_{k-1},i_{k+1},i_k,i_{k+2},\dots,i_n)>(i_1,\dots,i_{k-1},i_k,i_{k+1},i_{k+2},\dots,i_n).$$

[30]参看本节习题.

[31]即所有n元对称多项式组成的集合是$\mathbb{F}[x_1,x_2,\dots,x_n]$的一个子环.

由于f是对称多项式, 所以f中含有项

$$ax_1^{i_1}\cdots x_{k-1}^{i_{k-1}}x_{k+1}^{i_k}x_k^{i_{k+1}}x_{k+2}^{i_{k+2}}\cdots x_n^{i_n},$$

它比f的首项大, 矛盾.

下面寻找一组非负整数$d_1,d_2,\ldots,d_n$, 使得

$$a\sigma_1^{d_1}\sigma_2^{d_2}\cdots\sigma_n^{d_n}$$

的首项等于f的首项$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$.

由于

$$a\sigma_1^{d_1}\sigma_2^{d_2}\cdots\sigma_n^{d_n}=a(x_1+\cdots)^{d_1}(x_1x_2+\cdots)^{d_2}\cdots(x_1x_2\cdots x_n)^{d_n},$$

所以它的首项是

$$ax_1^{d_1+d_2+\cdots+d_n}x_2^{d_2+\cdots+d_n}\cdots x_n^{d_n},$$

所以我们要取

$$d_1=i_1-i_2,\quad d_2=i_2-i_3,\quad d_{n-1}=i_{n-1}-i_n,\quad d_n=i_n.$$

于是, 令

$$\varphi_1(x_1,x_2,\ldots,x_n)=a\sigma_1^{i_1-i_2}\sigma_2^{i_2-i_3}\cdots\sigma_{n-1}^{i_{n-1}-i_n}\sigma_n^{i_n},$$

则对称多项式

$$f_1(x_1,x_2,\ldots,x_n)=f(x_1,x_2,\ldots,x_n)-\varphi_1(x_1,x_2,\ldots,x_n)$$

的首项小于f的首项.

当$f_1\neq 0$时, 重复这一过程, 则存在$\sigma_1,\sigma_2,\ldots,\sigma_n$的多项式$\varphi_2$, 使得$\varphi_2$和$f_1$有相同的首项, 而

$$f_2=f_1-\varphi_2$$

的首项更小. 如此, 就得到$\sigma_1,\sigma_2,\ldots,\sigma_n$的多项式$\varphi_1,\varphi_2,\ldots$, 使得

$$f,f_1=f-\varphi_1,f_2=f_1-\varphi_2,\cdots,f_j=f_{j-1}-\varphi_j,\cdots$$

的每一个的首项都比前一个的首项小. 但是比f的首项小的单项式(不计系数)只有有限个, 于是有限步后必有$f_t=0$. 于是

$$f=\varphi_1+\varphi_2+\cdots+\varphi_t$$

是$\sigma_1,\sigma_2,\ldots,\sigma_n$的多项式.

再证明唯一性. 设有n元多项式g,h, 使得

$$f(x_1,x_2,\ldots,x_n)=g(\sigma_1,\sigma_2,\ldots,\sigma_n)=h(\sigma_1,\sigma_2,\ldots,\sigma_n).$$

定义多项式

$$\varphi(x_1,x_2,\ldots,x_n)=g(x_1,x_2,\ldots,x_n)-h(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n].$$

如果$\varphi\neq 0$, 则由上一节习题, 存在$b_1,b_2,\ldots,b_n\in\mathbb{F}$, 使得

$$\varphi(b_1,b_2,\ldots,b_n)\neq 0.$$

令多项式

$$\Phi(x)=x^n-b_1x^{n-1}+\cdots+(-1)^kb_kx^{n-k}+\cdots+(-1)^nb_n\in\mathbb{F}[x]$$

的n个复根为$c_1,c_2,\ldots,c_n$, 则由Vièta公式

$$b_k=\sigma_k(c_1,c_2,\ldots,c_n),\quad k=1,2,\ldots,n.$$

于是有

$$
\begin{aligned}
&\varphi(b_1,b_2,\ldots,b_n)=g(b_1,b_2,\ldots,b_n)-h(b_1,b_2,\ldots,b_n)\\
=&g(\sigma_1(c_1,\ldots,c_n),\ldots,\sigma_n(c_1,\ldots,c_n))-h(\sigma_1(c_1,\ldots,c_n),\ldots,\sigma_n(c_1,\ldots,c_n))\\
=&f(c_1,c_2,\ldots,c_n)-f(c_1,c_2,\ldots,c_n)=0,
\end{aligned}
$$

矛盾. 于是$\varphi=0$, 进而$g=h$. □

上面定理1.29的存在性的证明事实上给出了将对称多项式表示成初等对称多项式的多项式的算法, 我们总结如下.

算法1.1. *求多项式g, 使得对称多项式f满足$f(x_1,x_2,\ldots,x_n)=g(\sigma_1,\sigma_2,\ldots,\sigma_n)$的步骤:*

(1) *写出f的首项$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$, 计算*

$$f_1(x_1,x_2,\ldots,x_n)=f(x_1,x_2,\ldots,x_n)-a\sigma_1^{i_1-i_2}\sigma_2^{i_2-i_3}\cdots\sigma_{n-1}^{i_{n-1}-i_n}\sigma_n^{i_n};$$

(2) *当$f_1\neq 0$, 对f_1用(1), 得到f_2. 继续这一过程, 直到得到$f_t=0$, 则可求出g.*

例1.35. *设$f(x_1,x_2,x_3)=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2\in\mathbb{F}[x_1,x_2,x_3]$, 用初等对称多项式表示$f$.*

☞ **解.** f的首项为$x_1^2x_2^2$, 我们有

$$
\begin{aligned}
&f(x_1,x_2,x_3)-\sigma_1^{2-2}\sigma_2^2=f(x_1,x_2,x_3)-\sigma_2^2\\
=&x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2-(x_1x_2+x_1x_3+x_2x_3)^2\\
=&-2(x_1^2x_2x_3+x_1x_2^2x_3+x_1x_2x_3^2)\\
=&-2x_1x_2x_3(x_1+x_2+x_3)=-2\sigma_1\sigma_3,
\end{aligned}
$$

于是, $f=\sigma_2^2-2\sigma_1\sigma_3$. □

设f是m次对称多项式, 则有

$$f=f_m+f_{m-1}+\cdots+f_0,$$

其中, f_i是f的i次齐次成分. 容易证明, f_i也是对称多项式[32]. 于是, 如果可以把f的每个齐次成分表示为初等对称多项式的多项式, 则f也可表示.

设f是齐次对称多项式, 在定理1.29的存在性证明中

$$f=\varphi_1+\varphi_2+\cdots+\varphi_t,$$

其中φ_i都有形式

$$\varphi_i(x_1,x_2,\ldots,x_n)=a_i\sigma_1^{l_{i1}-l_{i2}}\sigma_2^{l_{i2}-l_{i3}}\cdots\sigma_{n-1}^{l_{i,n-1}-l_{in}}\sigma_n^{l_{in}},$$

这里

- $a_1x_1^{l_{11}}x_2^{l_{12}}\cdots x_n^{l_{1n}}$是$f$的首项,

[32]参看本节习题.

- 对$i=2,3,\ldots,t$, 有

  (a) $l_{i1}+l_{i2}+\cdots+l_{in}=l_{11}+l_{12}+\cdots+l_{1n}$,

  (b) $l_{i1}\geqslant l_{i2}\geqslant\cdots\geqslant l_{in}$,

  (c) $(l_{i1},l_{i2},\ldots,l_{in})<(l_{11},l_{12},\ldots,l_{1n})$.

于是也可以用下面的待定系数法用初等对称多项式表示齐次对称多项式.

算法1.2. *用待定系数法将m次齐次对称多项式f表示为初等对称多项式的步骤:*

(1) 写出f的首项$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$, 求集合

$$M=\left\{(l_1,l_2,\ldots,l_n)\;\middle|\;\begin{array}{c}l_j\geqslant 0,\sum\limits_{j=1}^{n}l_j=m,\\ l_1\geqslant l_2\geqslant\cdots\geqslant l_n,(l_1,l_2,\ldots,l_n)<(i_1,i_2,\ldots,i_n)\end{array}\right\};$$

(2) 假设

$$f=a\sigma_1^{i_1-i_2}\sigma_2^{i_2-i_3}\cdots\sigma_{n-1}^{i_{n-1}-i_n}\sigma_n^{i_n}+\sum_{(l_1,l_2,\ldots,l_n)\in M}A_{l_1l_2\cdots l_n}\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\cdots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n};$$

(3) 取$x_1,x_2,\ldots,x_n$的一些特殊值, 得到待定常数$A_{j_1j_2\cdots j_n}$的方程, 求出待定常数. 通常取$x_1=\cdots=x_k=1$, $x_{k+1}=\cdots=x_n=0$, 此时

$$\sigma_j(\underbrace{1,\ldots,1}_{k},0,\ldots,0)=\begin{cases}\binom{k}{j} & 1\leqslant j\leqslant k,\\ 0 & k+1\leqslant j\leqslant n.\end{cases}$$

例1.36 (例1.35). 设$f(x_1,x_2,x_3)=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2\in\mathbb{F}[x_1,x_2,x_3]$, 用初等对称多项式表示$f$.

☞ **解.** f是4次齐次多项式, f的首项为$x_1^2x_2^2$, 对应$(2,2,0)$. 于是

$$M=\{(2,1,1)\},$$

可设

$$f=\sigma_1^{2-2}\sigma_2^2+A\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3=\sigma_2^2+A\sigma_1\sigma_3.$$

令$x_1=x_2=x_3=1$, 则

$$\sigma_1=\binom{3}{1}=3,\quad \sigma_2=\binom{3}{2}=3,\quad \sigma_3=\binom{3}{3}=1,$$

而

$$f(1,1,1)=3.$$

于是有

$$3=3^2+A\cdot 3\cdot 1.$$

这得到$A=-2$, 进而$f=\sigma_2^2-2\sigma_1\sigma_3$. □

我们再看一个例子.

例1.37. 将对称多项式$f(x_1,x_2,\dots,x_n)=\sum\limits_{1\leqslant i<j\leqslant n}x_i^2x_j^2$用初等对称多项式表示, 这里$n\geqslant 4$.

解. f是4次齐次多项式, 首项是$x_1^2x_2^2$, 对应$(2,2,\underbrace{0,\dots,0}_{n-2})$. 于是

$$M=\{(2,1,1,\underbrace{0,\dots,0}_{n-3}),(1,1,1,1,\underbrace{0,\dots,0}_{n-4})\}.$$

所以可设

$$f=\sigma_2^2+A\sigma_1\sigma_3+B\sigma_4.$$

取$x_1=x_2=x_3=1$, $x_4=\dots=x_n=0$, 得到

$$3=3^2+A\cdot 3\cdot 1;$$

取$x_1=x_2=x_3=x_4=0$, $x_5=\dots=x_n=0$, 得到

$$6=6^2+A\cdot 4\cdot 4+B\cdot 1.$$

于是解得$A=-2$, $B=2$. 进而$f=\sigma_2^2-2\sigma_1\sigma_3+2\sigma_4$. □

除了初等对称多项式, 还有一组(无穷个)重要的对称多项式, 即**等幂和**:

$$s_k=s_k(x_1,x_2,\dots,x_n):=x_1^k+x_2^k+\dots+x_n^k,\quad (k=0,1,2,\dots)$$

由定理1.29, 等幂和可以用初等对称多项式表示, 下面的**Newton公式**给出了具体的表示方法.

定理1.30 (Newton公式). 在$\mathbb{F}[x_1,x_2,\dots,x_n]$中, 当$1\leqslant k\leqslant n$时, 有

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\dots+(-1)^{k-1}\sigma_{k-1}s_1+(-1)^kk\sigma_k=0;$$

当$k\geqslant n$时, 有

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\dots+(-1)^{n-1}\sigma_{n-1}s_{k-n+1}+(-1)^n\sigma_ns_{k-n}=0.$$

证明. 在$\mathbb{F}[x_1,\dots,x_n,x]$中定义多项式

$$\begin{aligned}f(x)=&\prod_{i=1}^n(x-x_i)\\=&x^n-\sigma_1(x_1,\dots,x_n)x^{n-1}+\dots+(-1)^{n-1}\sigma_{n-1}(x_1,\dots,x_n)x+(-1)^n\sigma_n(x_1,\dots,x_n)\\=&\sum_{i=0}^n(-1)^i\sigma_ix^{n-i},\quad(\sigma_0:=1)\end{aligned}$$

和

$$\begin{aligned}g(x)=&\sum_{i=0}^ks_ix^{k-i}=\sum_{i=0}^k(x_1^i+x_2^i+\dots+x_n^i)x^{k-i}\\=&\sum_{j=1}^n\sum_{i=0}^kx_j^ix^{k-i}=\sum_{j=1}^n(x^k+x_jx^{k-1}+\dots+x_j^{k-1}x+x_j^k)\\=&\sum_{j=1}^n\frac{x^{k+1}-x_j^{k+1}}{x-x_j}.\end{aligned}$$

令$h(x)=f(x)g(x)$. 一方面

$$\begin{aligned}h(x)&=\left(\sum_{i=0}^{n}(-1)^i\sigma_i x^{n-i}\right)\left(\sum_{i=0}^{k}s_i x^{k-i}\right)\\&=\sum_{l=0}^{n+k}\left(\sum_{\substack{i+j=l\\0\leqslant i\leqslant n,0\leqslant j\leqslant k}}(-1)^i\sigma_i s_j\right)x^{n+k-l},\end{aligned}$$

这得到$h(x)$中x^n的系数为

$$\sum_{\substack{i+j=k\\0\leqslant i\leqslant n,0\leqslant j\leqslant k}}(-1)^i\sigma_i s_j=\sum_{i=0}^{\min\{k,n\}}(-1)^i\sigma_i s_{k-i}=\begin{cases}\sum\limits_{i=0}^{k}(-1)^i\sigma_i s_{k-i} & 1\leqslant k\leqslant n,\\ \sum\limits_{i=0}^{n}(-1)^i\sigma_i s_{k-i} & k\geqslant n.\end{cases}$$

另一方面, 由上面的计算, 得到

$$\begin{aligned}h(x)=&\sum_{j=1}^{n}(x^{k+1}-x_j^{k+1})\prod_{\substack{i=1\\i\neq j}}^{n}(x-x_i)\\=&x^{k+1}f'(x)-\sum_{j=1}^{n}x_j^{k+1}\prod_{\substack{i=1\\i\neq j}}^{n}(x-x_i).\end{aligned}$$

上面等式右边的第二项x的次数都小于n, 所以x^n只可能出现在右边的第一项. 注意到

$$f'(x)=\sum_{i=0}^{n-1}(-1)^i(n-i)\sigma_i x^{n-i-1},$$

所以又可得x^n在$h(x)$中的系数为

$$\begin{cases}(-1)^k(n-k)\sigma_k & 1\leqslant k\leqslant n-1,\\ 0 & k\geqslant n.\end{cases}$$

由此就可得Newton公式. □

利用Newton公式和数学归纳法, 容易证明, 等幂和可以用初等对称多项式表示; 反之, 类似可以证明, 初等对称多项式$\sigma_1,\sigma_2,\ldots,\sigma_n$可以用等幂和$s_1,s_2,\ldots,s_n$表示. 例如, 设$n\geqslant 3$, 则由Newton公式, 我们得到

$$\begin{aligned}&s_1-\sigma_1=0,\\&s_2-\sigma_1 s_1+2\sigma_2=0,\\&s_3-\sigma_1 s_2+\sigma_2 s_1-3\sigma_3=0.\end{aligned}$$

由此, 容易得到

$$s_1=\sigma_1,\quad s_2=\sigma_1^2-2\sigma_2,\quad s_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3$$

和

$$\sigma_1=s_1,\quad \sigma_2=\frac{1}{2}(s_1^2-s_2),\quad \sigma_3=\frac{1}{6}(s_1^3-3s_1s_2+2s_3).$$

事实上, 可以用行列式给出一般的公式, 我们把这留作练习. 由此可得

推论1.31. 设$f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$是$n$元对称多项式, 则$f$可以唯一表示为等幂和$s_1,s_2,\dots,s_n$的多项式, 即存在唯一的多项式$g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 使得

$$f(x_1,x_2,\dots,x_n)=g(s_1,s_2,\dots,s_n).$$

☞ **证明.** 唯一性可类似于定理1.29的唯一性的证明. □

习题1.14

A1. 记$1,2,\dots,n$的所有排列做成的集合为S_n. 设$i_1,i_2,\dots,i_n\in S_n$, 可以将该排列看成映射

$$\sigma:\{1,2,\dots,n\}\to\{1,2,\dots,n\},j\mapsto\sigma(j)=i_j,j=1,2,\dots,n.$$

于是利用映射的复合, 可以定义两个排列的乘积. 即对于$\sigma,\tau\in S_n$, 有$\sigma\tau\in S_n$, 使得对任意$1\leqslant j\leqslant n$, 有$(\sigma\tau)(j)=\sigma(\tau(j))$.

(1) 设$\sigma=1,3,2,4\in S_4$, $\tau=2,4,3,1\in S_4$, 计算排列$\sigma\tau$;

(2) 如果$\sigma\in S_n$满足: 存在$1\leqslant i<j\leqslant n$, 使得

$$\sigma(i)=j,\quad\sigma(j)=i,\quad\sigma(k)=k,\quad 1\leqslant\forall k\neq i,j\leqslant n,$$

则称σ是一个对换. 证明: S_n中的任意一个排列都可以写为有限个对换的乘积(自然排列$1,2,\dots,n$看成是零个对换的乘积);

(3) 证明: 多项式$f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$是对称多项式的充分和必要条件是, 对任意$1\leqslant i<j\leqslant n$, 成立

$$f(x_1,\dots,x_i,\dots,x_j,\dots,x_n)=f(x_1,\dots,x_j,\dots,x_i,\dots,x_n).$$

A2. 在$\mathbb{F}[x_1,x_2,x_3]$含有项$x_1^3x_3$的对称多项式中, 写出项数最少的那个对称多项式.

A3. 设$f(x_1,x_2,\dots,x_n)$是m次齐次多项式, 记$f=f_0+f_1+\cdots+f_m$, 其中f_i是f的i次齐次部分, $i=0,1,\dots,m$. 证明: 对$i=0,1,\dots,m$, $f_i(x_1,x_2,\dots,x_n)$都是对称多项式.

A4. 在$\mathbb{F}[x_1,x_2,x_3]$中, 用初等对称多项式表示下列对称多项式:

(1) $x_1^3x_2^2+x_1^3x_3^2+x_1^2x_2^3+x_1^2x_3^3+x_2^3x_3^2+x_2^2x_3^3$;

(2) $x_1^4+x_2^4+x_3^4$;

(3) $(x_1^2+x_2^2)(x_1^2+x_3^2)(x_2^2+x_3^2)$;

(4) $(x_1x_2+x_3^2)(x_2x_3+x_1^2)(x_1x_3+x_2^2)$;

(5) $(2x_1x_2+x_3^2)(2x_2x_3+x_1^2)(2x_1x_3+x_2^2)$;

(6) $(2x_1-x_2-x_3)(2x_2-x_1-x_3)(2x_3-x_1-x_2)$.

A5. 在$\mathbb{F}[x_1,x_2,\dots,x_n]$中, 用初等对称多项式表示下列对称多项式:

(1) $\sum\limits_{1\leqslant i<j<k\leqslant n}(x_i^2x_j^2x_k+x_i^2x_jx_k^2+x_ix_j^2x_k^2)$, $(n\geqslant 5)$;

(2) $\sum\limits_{1\leqslant i<j\leqslant n}(x_i-x_j)^2$;

(3) $(-x_1+x_2+\cdots+x_n)(x_1-x_2+x_3+\cdots+x_n)\cdots(x_1+\cdots+x_{n-1}-x_n)$.

A6. 设x_1,x_2,x_3是方程$x^3+px+q=0$的3个复根, 将$D=(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$用$p,q$表示.

A7. 设c_1,c_2,c_3是$x^3+a_2x^2+a_1x+a_0$的三个复根, 计算

$$(c_1^2+c_1c_2+c_2^2)(c_2^2+c_2c_3+c_3^2)(c_1^2+c_1c_3+c_3^2).$$

A8. 解方程组$\begin{cases}x+y+z+w=10,\\ x^2+y^2+z^2+w^2=30,\\ x^3+y^3+z^3+w^3=100,\\ xyzw=24.\end{cases}$

A9. 已知复数x,y,z满足$x+y+z=3$, $x^2+y^2+z^2=5$, $x^3+y^3+z^3=7$, 求$x^4+y^4+z^4$.

A10. 设$k\in\mathbb{N}$, 把下列n元对称多项式用n元等幂和的多项式表示:

(1) $\sum\limits_{1\leqslant i<j\leqslant n} x_i^k x_j^k$;

(2) $\sum\limits_{1\leqslant i<j\leqslant n} (x_i+x_j)^k$;

(3) $\sum\limits_{1\leqslant i<j\leqslant n} (x_i-x_j)^{2k}$.

B1. 设$f(x)=x^3+a_2x^2+a_1x+a_0\in\mathbb{F}[x]$. 证明:

(1) $f(x)$的三个复根成等差数列的充分必要条件是, $2a_2^3-9a_1a_2+27a_0=0$;

(2) $f(x)$的三个复根成等比数列的充分必要条件是, $a_2^3a_0-a_1^3=0$.

B2. 设$1\leqslant k\leqslant n$, 证明: 在$\mathbb{F}[x_1,x_2,\ldots,x_n]$中有等式

$$s_k=\begin{vmatrix}\sigma_1 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0\\ 2\sigma_2 & \sigma_1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0\\ 3\sigma_3 & \sigma_2 & \sigma_1 & 1 & 0 & 0 & \cdots & 0 & 0\\ 4\sigma_4 & \sigma_3 & \sigma_2 & \sigma_1 & 1 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ (k-1)\sigma_{k-1} & \sigma_{k-2} & \sigma_{k-3} & \sigma_{k-4} & \sigma_{k-5} & \sigma_{k-6} & \cdots & \sigma_1 & 1\\ k\sigma_k & \sigma_{k-1} & \sigma_{k-2} & \sigma_{k-3} & \sigma_{k-4} & \sigma_{k-5} & \cdots & \sigma_2 & \sigma_1\end{vmatrix}$$

和

$$\sigma_k=\frac{1}{k!}\begin{vmatrix}s_1 & 1 & 0 & 0 & \cdots & 0 & 0\\ s_2 & s_1 & 2 & 0 & \cdots & 0 & 0\\ s_3 & s_2 & s_1 & 3 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ s_{k-1} & s_{k-2} & s_{k-3} & s_{k-4} & \cdots & s_1 & k-1\\ s_k & s_{k-1} & s_{k-2} & s_{k-3} & \cdots & s_2 & s_1\end{vmatrix}.$$

B3. 设整数$n\geqslant 2$, 求数域$\mathbb{F}$上的n次首一多项式$f(x)$, 使得它的n个复根的k次幂之和等于0, 其中

(1) $1\leqslant k\leqslant n-1$;

(2) $2\leqslant k\leqslant n$.

B4. 设三次方程$x^3+ax^2+bx+c=0$的三个根分别是某个三角形的三个角的正弦, 证明:

$$a(4ab-a^3-8c)=4c^2.$$

B5. 求$\sigma_k(x_1,x_2,\ldots,x_n)$关于$x_i$的偏导数$\frac{\partial\sigma_k}{\partial x_i}$, 并计算

$$\sum_{i=1}^{n}\frac{\partial\sigma_k}{\partial x_i}.$$

B6. 求多项式

$$f(x)=x^n+(a+b)x^{n-1}+(a^2+b^2)x^{n-2}+\cdots+(a^{n-1}+b^{n-1})x+(a^n+b^n)$$

的根的等幂和$s_1,s_2,\ldots,s_n$.

§ 1.15 附录: 代数基本定理的几个证明

代数基本定理(定理1.18)的第一个严格证明由Gauss于1799年给出. 这里我们给出几个证明.

1.15.1 使用微积分证明

这个证明是Gauss在1816年给出的. 我们只要证明: 如果$f(z)\in\mathbb{C}[z]$没有复数根, 那么$f(z)$一定是一个非零常数.

可以不妨设$f(z)$是首一的:

$$f(z)=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0,\quad (a_j\in\mathbb{C}).$$

采用极坐标$z=re^{i\theta}$, 则$f(z)=P(r,\theta)+iQ(r,\theta)$, 其中

$$\begin{aligned}P(r,\theta)=&r^n\cos n\theta+(\Re(a_{n-1})\cos(n-1)\theta-\Im(a_{n-1})\sin(n-1)\theta)r^{n-1}\\&+\cdots+(\Re(a_1)\cos\theta-\Im(a_1)\sin\theta)r+\Re(a_0),\\Q(r,\theta)=&r^n\sin n\theta+(\Re(a_{n-1})\sin(n-1)\theta+\Im(a_{n-1})\cos(n-1)\theta)r^{n-1}\\&+\cdots+(\Re(a_1)\sin\theta+\Im(a_1)\cos\theta)r+\Im(a_0).\end{aligned}$$

令$U=\arctan\left(\frac{Q}{P}\right)$, 则

$$\frac{\partial U}{\partial r}=\frac{PQ_r-QP_r}{P^2+Q^2},\qquad \frac{\partial U}{\partial\theta}=\frac{PQ_\theta-QP_\theta}{P^2+Q^2}.$$

请注意这里U的定义可能会有问题: 存在(r,θ), 使得$P(r,\theta)=0$. 但是由于$f(z)$无复根, 所以P,Q不同时为零, 即$\frac{\partial U}{\partial r}$和$\frac{\partial U}{\partial\theta}$对任意$(r,\theta)$都有定义, 而我们只需要使用这两个函数, 并不需要使用U.

直接计算可得

$$\frac{\partial}{\partial\theta}\left(\frac{\partial U}{\partial r}\right)=\frac{\partial}{\partial r}\left(\frac{\partial U}{\partial\theta}\right),$$

于是对任意$R>0$, 有

$$\int_0^R\left(\int_0^{2\pi}\frac{\partial}{\partial\theta}\frac{\partial U}{\partial r}d\theta\right)dr=\int_0^R\left(\int_0^{2\pi}\frac{\partial}{\partial r}\frac{\partial U}{\partial\theta}d\theta\right)dr=\int_0^{2\pi}\left(\int_0^R\frac{\partial}{\partial r}\frac{\partial U}{\partial\theta}dr\right)d\theta.$$

下面分别计算上面连等式中的第一和第三个累次积分.

由于P和Q都是关于θ的周期为2π的周期函数, 所以P_r和Q_r也是, 进而$\frac{\partial U}{\partial r}$也是$\theta$的周期为$2\pi$的周期函数. 这得到

$$\int_0^{2\pi} \frac{\partial}{\partial \theta}\frac{\partial U}{\partial r} d\theta = \left.\frac{\partial U}{\partial r}\right|_{\theta=0}^{\theta=2\pi} = 0.$$

另一方面, 我们有

$$\int_0^{R} \frac{\partial}{\partial r}\frac{\partial U}{\partial \theta} dr = \left.\frac{\partial U}{\partial \theta}\right|_{r=0}^{r=R}.$$

由P和Q的表达式, 可得

$$\begin{aligned}
P_\theta =& - nr^n \sin n\theta - (n-1)(\Re(a_{n-1})\sin(n-1)\theta + \Im(a_{n-1})\cos(n-1)\theta)r^{n-1} \\
&+ \cdots - (\Re(a_1)\sin\theta + \Im(a_1)\cos\theta)r, \\
Q_\theta =& \, nr^n \cos n\theta + (n-1)(\Re(a_{n-1})\cos(n-1)\theta - \Im(a_{n-1})\sin(n-1)\theta)r^{n-1} \\
&+ \cdots + (\Re(a_1)\cos\theta - \Im(a_1)\sin\theta)r.
\end{aligned}$$

特别有

$$\left.\frac{\partial P}{\partial \theta}\right|_{r=0} = \left.\frac{\partial Q}{\partial \theta}\right|_{r=0} = 0.$$

这得到

$$\int_0^{R} \frac{\partial}{\partial r}\frac{\partial U}{\partial \theta} dr = \left.\frac{\partial U}{\partial \theta}\right|_{r=R}.$$

所以有

$$\int_0^{2\pi} \left.\frac{\partial U}{\partial \theta}\right|_{r=R} d\theta = 0,$$

进而得到

$$0 = \lim_{R\to\infty}\int_0^{2\pi} \left.\frac{\partial U}{\partial \theta}\right|_{r=R} d\theta = \int_0^{2\pi} \lim_{R\to\infty}\left.\frac{\partial U}{\partial \theta}\right|_{r=R} d\theta.$$

由于

$$PQ_\theta - QP_\theta = nr^{2n} + r\text{的低次项}, \quad P^2 + Q^2 = r^{2n} + r\text{的低次项},$$

且$\sin x$和$\cos x$是有界函数, 我们得到

$$\lim_{R\to\infty}\left.\frac{\partial U}{\partial \theta}\right|_{r=R} = n.$$

进而就有$2\pi n = 0$, $n = 0$. □

1.15.2 使用矩阵证明

这里介绍H. Derksen在2003年给出的一个证明([2]). 我们证明

定理1.32. 设$n \geqslant 1$, 则任意n阶复矩阵有特征值. 等价地, 任意n维复线性空间上的线性变换有特征值.

代数基本定理和定理1.32是等价的. 事实上, 这由于矩阵

$$A=\begin{pmatrix}0&0&0&\cdots&0&-a_0\\1&0&0&\cdots&0&-a_1\\0&1&0&\cdots&0&-a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&-a_{n-2}\\0&0&0&\cdots&1&-a_{n-1}\end{pmatrix}$$

的特征多项式

$$\varphi_A(z)=|zE-A|=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0.$$

引理1.33. *设整数$m>1$. 如果对任意维数不是m倍数的$\mathbb{F}$上的线性空间V, V上的每个线性变换都有特征向量, 则对任意维数不是m倍数的$\mathbb{F}$上的线性空间V, V上的任意两个可交换的线性变换有公共的特征向量.*

☞ **证明.** 维数不是m倍数的$\mathbb{F}$上的线性空间可以按照维数分类:

$$\mathcal{V}_d=\{V: d\text{维线性空间}\},\quad m\nmid d.$$

当$d=1$时, 对任意$V\in\mathcal{V}_1$和V上的线性变换$\mathscr{A}$, V上的任意非零向量都是$\mathscr{A}$的特征向量. 由此可知结论对$\mathcal{V}_1$中的线性空间成立.

下设$d>1$, $m\nmid d$, 且结论对$\mathcal{V}_{d'}$ $(d'<d, m\nmid d')$ 中的线性空间成立. 任取$V\in\mathcal{V}_d$和V上的两个可交换的线性变换$\mathscr{A}$和$\mathscr{B}$. 由条件, $\mathscr{A}$有特征值$\lambda_0\in\mathbb{F}$. 定义V的子空间

$$U=\operatorname{Im}(\mathscr{A}-\lambda_0\mathscr{E}),\qquad W=\operatorname{Ker}(\mathscr{A}-\lambda_0\mathscr{E}).$$

则可得$\dim W\geqslant 1$. 显然, U和W是$\mathscr{A}$的不变子空间. 因为$\mathscr{A}$和$\mathscr{B}$可交换, 所以U和W也是$\mathscr{B}$的不变子空间.

如果U和W都是V的真子空间, 则由$\dim V=\dim U+\dim W$和$m\nmid d$, 可知$1\leqslant\dim U,\dim W<d$, 且

$$m\nmid\dim U\quad\text{或者}\quad m\nmid\dim W.$$

当$m\nmid\dim U$时, 由假设可得U上可交换的线性变换$\mathscr{A}|_U$和$\mathscr{B}|_U$有公共的特征向量, 进而就得到$\mathscr{A}$和$\mathscr{B}$有公共的特征向量$(\in U)$. 类似的, 如果$m\nmid\dim W$, 则可得$\mathscr{A}$和$\mathscr{B}$有公共的特征向量$(\in W)$.

如果U或者W不是V的真子空间, 则只能$W=V$. 由假设, $\mathscr{B}$有特征向量$\alpha\in V=W$. 而W中的任意非零向量都是$\mathscr{A}$的特征向量, 所以α是$\mathscr{A}$和$\mathscr{B}$的公共特征向量. □

引理1.34. *奇数次实系数多项式至少有一个实根. 等价地, 奇数维实线性空间上的线性变换一定有实特征值.*

☞ **证明.** 设

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{R}[x],$$

其中n是奇数. 令$a=|a_0|+\cdots+|a_{n-1}|+1$, 则

$$f(a)>0,\quad f(-a)<0.$$

所以由介值定理, 存在$\lambda_0\in(-a,a)$, 使得$f(\lambda_0)=0$. □

推论1.35. 设V是奇数维实线性空间, 则V上任意两个可交换的线性变换有公共的特征向量.

☞ **证明.** 由引理1.34, 取$\mathbb{F}=\mathbb{R}$和$m=2$满足引理1.33的假设. □

引理1.36. 奇数阶复矩阵必有特征值.

☞ **证明.** 任取n阶复矩阵A, 其中n是奇数. 考虑所有n阶Hermitian矩阵全体

$$\mathcal{H}_n=\{M\in\mathbb{C}^{n\times n}\mid M^*=M\},\quad M^*=\overline{M}^T.$$

容易证明, 在矩阵的加法和实数数乘下, $\mathcal{H}_n$成为$\mathbb{R}$上的n^2维线性空间, 是奇数维实线性空间.

注意到对任意$C\in\mathbb{C}^{n\times n}$, 有

$$C=\frac{C+C^*}{2}+i\frac{C-C^*}{2i},$$

其中

$$\frac{C+C^*}{2},\frac{C-C^*}{2i}\in\mathcal{H}_n.$$

于是对任意$B\in\mathcal{H}_n$, 有

$$\begin{aligned}AB=&\frac{AB+(AB)^*}{2}+i\frac{AB-(AB)^*}{2i}=\frac{AB+B^*A^*}{2}+i\frac{AB-B^*A^*}{2i}\\=&\frac{AB+BA^*}{2}+i\frac{AB-BA^*}{2i}.\end{aligned}$$

这得到映射

$$\mathscr{B},\mathscr{C}:\mathcal{H}_n\longrightarrow\mathcal{H}_n,$$

使得

$$\mathscr{B}(B)=\frac{AB+BA^*}{2},\quad\mathscr{C}(B)=\frac{AB-BA^*}{2i},\quad(\forall B\in\mathcal{H}_n).$$

可以验证$\mathscr{B}$和$\mathscr{C}$是可交换的$\mathbb{R}$线性变换, 于是由推论1.35, 存在$\lambda_1,\lambda_2\in\mathbb{R}$, 存在$O\neq B\in\mathcal{H}_n$, 使得

$$\mathscr{B}(B)=\lambda_1B,\qquad\mathscr{C}(B)=\lambda_2B.$$

于是

$$AB=\mathscr{B}(B)+\mathscr{C}(B)=(\lambda_1+i\lambda_2)B,$$

即$\lambda_1+i\lambda_2$是A的特征值, 而B的任意非零列是A的特征向量. □

我们最后来证明定理1.32.

☞ **定理1.32的证明.** 设$n=2^km$, 其中$k\geqslant 0$, 而m是奇数. 我们对k归纳证明. 由引理1.36可知$k=0$成立.

下设$k\geqslant 1$, 且任意满足$2^k\nmid d$的d阶复矩阵都有特征值. 于是由引理1.33, 对任意d维复线性空间V, 其中$2^k\nmid d$, V上任意两个可交换的线性变换有公共的特征向量.

现在设$n=2^km$, 其中m是奇数, 而$A\in\mathbb{C}^{n\times n}$, 我们证明$A$有特征值. 考虑$n$阶复对称阵全体所成的集合

$$\mathcal{S}_n=\{M\in\mathbb{C}^{n\times n}\mid A^T=A\}.$$

容易知道, $\mathcal{S}_n$是$\frac{n(n+1)}{2}$维复线性空间. 因为

$$\frac{n(n+1)}{2}=2^{k-1}m(n+1),$$

而$m(n+1)$是奇数, 所以$\mathcal{S}_n$上的任意两个可交换的线性变换有公共的特征向量.

考虑$\mathcal{S}_n$上的两个可交换的线性变换$\mathscr{B}$和$\mathscr{C}$, 其中

$$\mathscr{B}(B)=AB+BA^T,\quad\mathscr{C}(B)=ABA^T,\quad\forall B\in\mathcal{S}_n.$$

所以存在$\lambda, \mu \in \mathbb{C}$, 存在$O \neq B \in \mathcal{S}_n$, 使得

$$\mathscr{B}(B) = \lambda B, \quad \mathscr{C}(B) = \mu B,$$

即

$$AB + BA^T = \lambda B, \quad ABA^T = \mu B.$$

这得到

$$A^2B + \mu B = \lambda AB,$$

即

$$(A^2 - \lambda A + \mu E)B = O.$$

考虑二次多项式$z^2 - \lambda z + \mu$. 由引理1.37, 存在$\alpha, \beta \in \mathbb{C}$, 使得

$$z^2 - \lambda z + \mu = (z - \alpha)(z - \beta).$$

所以有

$$(A - \alpha E)(A - \beta E)B = O.$$

当$(A - \beta E)B = O$时, B的任意非零列是A的特征向量, β是B的特征值; 当$(A - \beta E)B \neq O$时, $(A - \beta E)B$的任意非零列是A的特征向量, α是A的特征值. □

引理1.37. *任意复数都有平方根.*

证明. 设$a + bi \in \mathbb{C}$, 其中$a, b \in \mathbb{R}$. 直接验证可得

$$\left(\sqrt{\frac{\sqrt{a^2+b^2}+a}{2}} \pm i\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}}\right)^2 = a + bi.$$

证毕. □

1.15.3 使用复变函数证明

我们需要复变函数中的**Liouvill定理**.

定理1.38 (Liouvill定理)**.** *在复平面上解析的有界函数一定是常值函数.*

证明. 参看[5]. □

下面开始证明代数基本定理. 设

$$f(x) = a_nx^n + \cdots + a_x + a_0 \in \mathbb{C}[x]$$

没有复根, 其中$a_n \neq 0$且$n \geqslant 1$, 我们要得出矛盾. 考查复变函数

$$F(z) = \frac{1}{f(z)},$$

它是$\mathbb{C}$上的解析函数. 我们证明$F(z)$是$\mathbb{C}$上的有界函数.

记$M=\max\{|a_{n-1}|,\dots,|a_1|,|a_0|\}$. 我们有

$$\begin{aligned}|f(z)|\geqslant&|a_nz^n|-|a_{n-1}z^{n-1}+\cdots+a_1z+a_0|\\ \geqslant&|a_n||z|^n-(|a_{n-1}||z|^{n-1}+\cdots+|a_1||z|+|a_0|)\\ \geqslant&|a_n||z|^n-M(|z|^{n-1}+\cdots+|a|+1)\\ =&|a_n||z|^n-M\frac{|z|^n-1}{|z|-1}\\ \geqslant&|a_n||z|^n-M\frac{|z|^n}{|z|-1}\quad(\text{当}|z|>1\text{时})\\ =&|z|^n\left(|a_n|-\frac{M}{|z|-1}\right),\end{aligned}$$

这得到

$$\lim_{z\to\infty}f(z)=\infty,$$

进而有

$$\lim_{z\to\infty}F(z)=0.$$

所以存在$R>0$, 存在$M_1>0$, 使得$|z|>R$时,

$$|F(z)|\leqslant M_1.$$

而$F(z)$在$|z|\leqslant R$上连续, 所以有界, 即存在$M_2>0$, 使得当$|z|\leqslant R$时有

$$|F(z)|\leqslant M_2.$$

最后就得到对任意$z\in\mathbb{C}$, 成立

$$|F(z)|\leqslant\max\{M_1,M_2\}.$$

于是$F(z)$是$\mathbb{C}$上的有界函数. 于是由Liouvill定理, $F(z)$是常值函数: 存在$b\in\mathbb{C}$, 使得

$$F(z)=b,\quad\forall z\in\mathbb{C}.$$

显然$b\neq0$, 所以

$$f(z)=\frac{1}{b},\quad\forall z\in\mathbb{C},$$

这矛盾于$n\geqslant1$. □

§ 1.16 附录: 一元三次和一元四次多项式的求根公式

1.16.1 Cardan公式

设$f(x)\in\mathbb{C}[x]$是三次多项式, 我们来求它的根. 可以不妨设$f(x)$是首一的, 即设

$$f(x)=x^3+ax^2+bx+c.$$

配方, 消去平方项

$$f(x)=\left(x+\frac{1}{3}a\right)^3+\left(b-\frac{1}{3}a^2\right)\left(x+\frac{1}{3}a\right)+c-\frac{1}{27}a^3-\frac{1}{3}a\left(b-\frac{1}{3}a^2\right).$$

令$y=x+\frac{1}{3}a$, 则

$$f(x)=y^3+py+q,\quad p=b-\frac{1}{3}a^2,\quad q=c+\frac{2}{27}a^3-\frac{1}{3}ab.$$

于是我们可以不妨设

$$f(x)=x^3+px+q.$$

如果$p=0$, 则该多项式的根为

$$x_1=\sqrt[3]{-q},\quad x_2=\sqrt[3]{-q}\omega,\quad x_3=\sqrt[3]{-q}\omega^2,$$

其中

$$\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i.$$

如果$q=0$, 则多项式的根为

$$x_1=0,\quad x_2=\sqrt{-p},\quad x_3=-\sqrt{-p}.$$

下面假设$p\neq 0$且$q\neq 0$. 引入新的变量, 设$x=u+v$. 则有

$$x^3=u^3+3u^2v+3uv^2+v^3=u^3+v^3+3uvx,$$

即

$$x^3-3uvx-(u^3-v^3)=0.$$

于是, 如果u和v满足

$$\begin{cases}uv=-\frac{1}{3}p,\\ u^3+v^3=-q,\end{cases}\tag{1.1}$$

则$x=u+v$就是$f(x)$的一个根. 如果u和v满足(1.1), 则满足

$$\begin{cases}u^3v^3=-\frac{1}{27}p^3,\\ u^3+v^3=-q,\end{cases}$$

即u^3,v^3是二次多项式

$$y^2+qy-\frac{1}{27}p^3$$

的两个根. 于是

$$u^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}},\quad v^3=-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}.$$

由于u, v要满足(1.1), 所以$f(x)$的三个根为

$$\begin{cases} x_1 = \sqrt[3]{-\frac{q}{2}+\sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2}-\sqrt{\Delta}}, \\ x_2 = \omega\sqrt[3]{-\frac{q}{2}+\sqrt{\Delta}} + \omega^2\sqrt[3]{-\frac{q}{2}-\sqrt{\Delta}}, \\ x_3 = \omega^2\sqrt[3]{-\frac{q}{2}+\sqrt{\Delta}} + \omega\sqrt[3]{-\frac{q}{2}-\sqrt{\Delta}}, \end{cases}$$

其中,

$$\Delta = \frac{q^2}{4} + \frac{p^3}{27}.$$

通常称上面的求根公式为**Cardan公式**.

1.16.2 Ferrari解法

考察四次首一多项式

$$f(x) = x^4 + ax^3 + bx^2 + cx + d$$

的根. 类似地, 令$y = x + \frac{1}{4}a$, 可以消去三次项, 所以可以不妨设

$$f(x) = x^4 + ax^2 + bx + d.$$

引入新的变量u, 配方得

$$\begin{aligned} f(x) =& x^4 + ux^2 + \frac{u^2}{4} - \left[(u-a)x^2 - bx + \frac{u^2}{4} - c\right] \\ =& \left(x^2 + \frac{u}{2}\right)^2 - \left[(u-a)x^2 - bx + \frac{u^2}{4} - c\right]. \end{aligned}$$

取u使得$u \neq a$且$(u-a)x^2 - bx + \frac{u^2}{4} - c$是完全平方, 这等价于$u$满足

$$b^2 - 4(u-a)\left(\frac{u^2}{4} - c\right) = 0.$$

上面是关于u的三次方程, 任取一个根u, 则

$$\begin{aligned} f(x) &= \left(x^2 + \frac{u}{2}\right)^2 - \left(\sqrt{u-a}x - \frac{b}{2\sqrt{u-a}}\right)^2 \\ &= \left(x^2 + \sqrt{u-a}x + \frac{u}{2} - \frac{b}{2\sqrt{u-a}}\right)\left(x^2 - \sqrt{u-a}x + \frac{u}{2} + \frac{b}{2\sqrt{u-a}}\right). \end{aligned}$$

由此可以求出$f(x)$的四个根. 称这种解法为**Ferrari解法**.

§ 1.17 附录: Schur定理的一个证明

本节证明下面的结论([13]).

定理1.39 (Schur定理)**.** 设$n \in \mathbb{N}$, $c_1, \ldots, c_{n-1} \in \mathbb{Z}$, 则多项式

$$1 + c_1 x + c_2 \frac{x^2}{2!} + \cdots + c_{n-1} \frac{x^{n-1}}{(n-1)!} \pm \frac{x^n}{n!}$$

在$\mathbb{Q}$上不可约. 特别地, 多项式

$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

在$\mathbb{Q}$上不可约.

☞ **证明.** 设

$$f(x) = \sum_{i=0}^{n} \frac{n!}{i!} c_i x^i = \pm x^n + n c_{n-1} x^{n-1} + \cdots + n! c_1 x + n! \in \mathbb{Z}[x], \quad (c_0 = 1, c_n = \pm 1).$$

则只要证$f(x)$在$\mathbb{Q}$上不可约. 我们假设$f(x)$在$\mathbb{Q}$上可约, 则$f(x)$有不可约因式

$$p(x) = x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x], \quad (m \leqslant n/2).$$

断言一. 设p是素数, 且$p \mid \frac{n!}{(n-m)!} = n(n-1)\cdots(n-m+1)$, 则$p \mid a_0$.

事实上, 当$i \leqslant n-m$时, 有

$$\frac{n!}{i!} = \frac{n!}{(n-m)!} \frac{(n-m)!}{i!}.$$

于是对这些i, 有$p \mid \frac{n!}{i!}$, 这得到

$$f(x) \equiv \sum_{i=n-m+1}^{n} \frac{n!}{i!} c_i x^i \pmod{p}.$$

记$\overline{f}(x)$为$f(x)$的系数模p后在$\mathbb{F}_p[x]$中的像, 则有

$$x^{n-m+1} \mid \overline{f}(x).$$

再设$f(x) = p(x)q(x)$, 其中$q(x) \in \mathbb{Z}[x]$为首项系数为± 1的$n-m$次多项式. 模p就得到

$$x^{n-m+1} \mid \overline{p}(x)\overline{q}(x).$$

而$\overline{q}(x)$的次数为$n-m$, 所以由$\mathbb{F}_p[x]$中的唯一分解可知$x \mid \overline{p}[x]$. 这就得到$p \mid a_0$.

断言二. 设p是素数, 且$p \mid a_0$, 则$p \leqslant m$.

这个需要使用代数数论([4]). 设α是$p(x)$的一个复根, 代数数域$K = \mathbb{Q}(\alpha)$, 而K的代数整数环记为$\mathcal{O}_K$. 于是$[K:\mathbb{Q}] = m$, $\alpha \in \mathcal{O}_K$, 且范数

$$N_{K/\mathbb{Q}}(\alpha) = \pm a_0 \equiv 0 \pmod{p}.$$

设有素理想分解

$$(p) = \mathfrak{p}^e \mathfrak{b}, \quad \mathfrak{p}: \text{素理想}, \mathfrak{p} \nmid \mathfrak{b}, 1 \leqslant e = e(\mathfrak{p}|p) \leqslant m.$$

由于$p \mid N((\alpha))$, 所以有素理想分解

$$(\alpha) = \mathfrak{p}^d \mathfrak{a}, \quad \mathfrak{p} \nmid \mathfrak{a}, d \geqslant 1.$$

由于$f(\alpha) = 0$, 所以得到

$$-n! = \sum_{i=1}^{n} \frac{n!}{i!} c_i \alpha^i.$$

下面来计算左右两边的$\mathfrak{p}$进赋值. 记

$$s_r = \mathrm{ord}_p(r!) = \sum_{j \geqslant 1} \left[\frac{r}{p^j}\right] < \sum_{j \geqslant 1} \frac{r}{p^j} = \frac{r}{p-1}.$$

则有

$$\mathrm{ord}_{\mathfrak{p}}(r!) = e\,\mathrm{ord}_p(r!) = es_r.$$

于是有

$$\mathrm{ord}_{\mathfrak{p}}(n!) = es_n,$$

进而存在i $(1 \leqslant i \leqslant n)$, 使得$c_i \neq 0$, 且

$$\mathrm{ord}_{\mathfrak{p}}\left(\frac{n!}{i!}c_i\alpha^i\right) \leqslant es_n.$$

由于

$$\mathrm{ord}_{\mathfrak{p}}\left(\frac{n!}{i!}c_i\alpha^i\right) = es_n - es_i + \mathrm{ord}_{\mathfrak{p}}(c_i) + id \geqslant es_n - es_i + id,$$

所以有

$$es_n - es_i + id \leqslant es_n \Longrightarrow id \leqslant es_i < e\frac{i}{p-1} \Longrightarrow (p-1)d < e \leqslant m.$$

所以$p \leqslant m$.

由断言一和断言二可得: $n-m+1, n-m+2, \ldots, n-1, n$这连续$m$个整数的任意素因子一定$\leqslant m$, 这矛盾于引理1.40. □

引理1.40. *设$k, l \in \mathbb{N}$, 且$k \leqslant l$, 则*

$$l+1, l+2, \ldots, l+k$$

中至少有一个整数有$> k$的素因子. 特别地, 对任意$k \in \mathbb{N}$, 存在素数p, 使得$k < p \leqslant 2k$.

☞ **证明.** 参看[13]或者[3]. □

§1.18 附录: 判别式和结式

1.18.1 判别式

我们在中学学过二次三项式的判别式. 设$f(x) = ax^2 + bx + c \in \mathbb{R}[x]$, 其中$a \neq 0$, 则$f$的判别式定义为

$$\Delta = b^2 - 4ac.$$

可以用判别式判别$f(x)$根的情况:

$$\Delta > 0 \Longleftrightarrow \text{有两个不相等的实根},$$
$$\Delta = 0 \Longleftrightarrow \text{有两个相等的实根},$$
$$\Delta < 0 \Longleftrightarrow \text{没有实根}.$$

特别地, 可以用Δ判别f是否有重根. 我们可以如下推导, 设α_1, α_2是f的两个复根, 则

$$f\text{有重根} \Longleftrightarrow \alpha_1 = \alpha_2 \Longleftrightarrow (\alpha_1 - \alpha_2)^2 = 0.$$

由根与系数的关系得到

$$\alpha_1 + \alpha_2 = -\frac{b}{a}, \quad \alpha_1\alpha_2 = \frac{c}{a},$$

所以得到

$$(\alpha_1-\alpha_2)^2=(\alpha_1+\alpha_2)^2-4\alpha_1\alpha_2=\frac{b^2}{a^2}-4\frac{c}{a}=\frac{\Delta}{a^2},$$

即有

$$\Delta=a^2(\alpha_1-\alpha_2)^2.$$

于是就有

$$f\text{有重根}\iff\Delta=0.$$

可以将上面的讨论推广到任意次数的多项式.

定义1.22. 设$f(x)=a_nx^n+\cdots+a_1x+a_0\in\mathbb{F}[x]$, 其中$a_n\neq 0$, $n\geqslant 1$. 设f的n个复根为α_1, α_2, $\ldots$, α_n, 定义f的判别式为

$$D(f)=a_n^{2n-2}\prod_{1\leqslant j<i\leqslant n}(\alpha_i-\alpha_j)^2.$$

由定义可得, 当$n=1$时$D(f)=1$. 而且一般有

$$f(x)\text{在复数域中有重根}\iff D(f)=0.$$

下面说明可以用f的系数来计算$D(f)$. 由于$\prod\limits_{1\leqslant j<i\leqslant n}(\alpha_i-\alpha_j)^2$是关于$\alpha_1,\alpha_2,\ldots,\alpha_n$的对称多项式, 所以存在多项式$g(x_1,x_2,\ldots,x_n)$, 使得

$$\prod_{1\leqslant j<i\leqslant n}(\alpha_i-\alpha_j)^2=g(\sigma_1(\alpha_1,\ldots,\alpha_n),\ldots,\sigma_n(\alpha_1,\ldots,\alpha_n)).$$

而由Vièta公式, 有

$$\begin{aligned}\sigma_1(\alpha_1,\ldots,\alpha_n)&=-\frac{a_{n-1}}{a_n},\\&\vdots\\\sigma_k(\alpha_1,\ldots,\alpha_n)&=(-1)^k\frac{a_{n-k}}{a_n},\\&\vdots\\\sigma_n(\alpha_1,\ldots,\alpha_n)&=(-1)^n\frac{a_0}{a_n}.\end{aligned}$$

于是就得到

$$D(f)=a_n^{2n-2}g\left(-\frac{a_{n-1}}{a_n},\frac{a_{n-2}}{a_n},\ldots,(-1)^n\frac{a_0}{a_n}\right)\in\mathbb{F}.$$

还可以利用Vandermonde行列式的公式来计算$D(f)$. 我们有

$$D(f)=a_n^{2n-2}\begin{vmatrix}1&1&\cdots&1\\\alpha_1&\alpha_2&\cdots&\alpha_n\\\alpha_1^2&\alpha_2^2&\cdots&\alpha_n^2\\\vdots&\vdots&&\vdots\\\alpha_1^{n-1}&\alpha_2^{n-1}&\cdots&\alpha_n^{n-1}\end{vmatrix}\begin{vmatrix}1&\alpha_1&\alpha_1^2&\cdots&\alpha_1^{n-1}\\1&\alpha_2&\alpha_2^2&\cdots&\alpha_2^{n-1}\\\vdots&\vdots&\vdots&&\vdots\\1&\alpha_n&\alpha_n^2&\cdots&\alpha_n^{n-1}\end{vmatrix}$$

$$=a_n^{2n-2}\begin{vmatrix} n & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ \vdots & \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{vmatrix},$$

其中

$$s_i = s_i(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

为$\alpha_1, \alpha_2, \ldots, \alpha_n$的等幂和, 它可以由初等对称多项式表示.

例1.38. *设$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{F}[x]$是三次多项式, 求$f(x)$的判别式$D(f)$. 并且当$f(x) \in \mathbb{R}[x]$时, 利用$D(f)$判别$f(x)$根的情况.*

☞ **解.** 设$f(x)$的三个复根为$\alpha_1, \alpha_2, \alpha_3$, 则

$$\sigma_1 = -\frac{a_2}{a_3}, \quad \sigma_2 = \frac{a_1}{a_3}, \quad \sigma_3 = -\frac{a_0}{a_3}.$$

而$(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2$的首项为$\alpha_1^4\alpha_2^2$, 对应到$(4,2,0)$. 需要考虑的比$(4,2,0)$小的数组有

$$(4,1,1), (3,3,0), (3,2,1), (2,2,2).$$

于是可设

$$(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2 = \sigma_1^2\sigma_2^2 + A\sigma_1^3\sigma_3 + B\sigma_2^3 + C\sigma_1\sigma_2\sigma_3 + D\sigma_3^2.$$

取$\alpha_1 = \alpha_2 = 1, \alpha_3 = 0$, 可得

$$0 = 4 + B \Longrightarrow B = -4.$$

取$\alpha_1 = \alpha_2 = -1$, $\alpha_3 = 2$, 可得

$$0 = -27B + 4D \Longrightarrow D = -27.$$

取$\alpha_1 = \alpha_2 = -2$, $\alpha_3 = 1$, 可得

$$0 = -27 \cdot 4A + 16D \Longrightarrow A = -4.$$

最后取$\alpha_1 = \alpha_2 = \alpha_3$, 可得

$$0 = 3^4 + 3^3A + 3^3B + 3^2C + D \Longrightarrow C = 18.$$

于是

$$(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2 = \sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 - 4\sigma_2^3 + 18\sigma_1\sigma_2\sigma_3 - 27\sigma_3^2,$$

进而

$$\begin{aligned} D(f) =& a_3^4\left(\frac{a_1^2a_2^2}{a_3^4} - 4\frac{a_0a_2^3}{a_3^4} - 4\frac{a_1^3}{a_3^3} + 18\frac{a_0a_1a_2}{a_3^3} - 27\frac{a_0^2}{a_3^2}\right) \\ =& a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 + 18a_0a_1a_2a_3 - 27a_0^2a_3^2. \end{aligned}$$

下设$f(x)$是三次实系数多项式. 首先, $f(x)$必有实根, 设为α_1, 再设另外两个根为α_2, α_3.

当$D(f) = 0$时, $f(x)$有重根. 如果$\alpha_1 = \alpha_2$, 则由于实系数多项式的虚根共轭成对出现, 所以有$\alpha_3 \in \mathbb{R}$. 类似地, 如果$\alpha_1 = \alpha_3$, 则有$\alpha_2 \in \mathbb{R}$; 如果有$\alpha_2 = \alpha_3$, 则有$\alpha_2, \alpha_3 \in \mathbb{R}$. 所以, 当$D(f) = 0$时, $f(x)$有三个实根且有重根.

当$D(f) \neq 0$时, $f(x)$无重根. 当$\alpha_2, \alpha_3 \in \mathbb{R}$时, 有

$$D(f) = a_3^4(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2 > 0;$$

当α_2和α_3是共轭虚数时, 设$\alpha_2 = a + bi$, $\alpha_3 = a - bi$, $a, b \in \mathbb{R}$, $b \neq 0$, 则

$$D(f) = a_3^4[(\alpha_1 - a)^2 + b^2]^2(2bi)^2 = -4a_3^4b^2[(\alpha_1 - a)^2 + b^2]^2 < 0.$$

因此, 我们得到$f(x)$的根的情况: 当$D(f) \geqslant 0$时, $f(x)$有三个实根(重根计算重数), 其中当$D(f) = 0$时有重根, 当$D(f) > 0$ 时无重根; 当$D(f) < 0$时, $f(x)$有一个实根和一对共轭虚根. □

例1.39. 求$f(x) = x^n + a \in \mathbb{F}[x]$的判别式.

解. 我们用$D(f)$的行列式公式来求. 首先, 由Vièta公式, 可得

$$\sigma_1 = \cdots = \sigma_{n-1} = 0, \quad \sigma_n = (-1)^n a.$$

于是当$1 \leqslant k < n$时, 有[33]

$$s_k = \begin{vmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{vmatrix} = 0;$$

当$k = n$时, 有

$$s_n = \begin{vmatrix} & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & 0 & 1 \\ n(-1)^n a & & & & & 0 \end{vmatrix} = (-1)^{n-1}n(-1)^n a = -na;$$

当$n < k < 2n$时, 由Newton公式有

$$s_k = (-1)^n \sigma_n s_{k-n} = 0.$$

所以

$$D(f) = \begin{vmatrix} n & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -na \\ 0 & 0 & 0 & \cdots & -na & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & -na & 0 & \cdots & 0 & 0 \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} n^n a^{n-1}.$$

特别地, 当$a \neq 0$时, $f(x)$在复数域上没有重根. □

1.18.2 结式

设$f(x) \in \mathbb{F}[x]$, 则$f(x)$在复数域上有重根的充分必要条件是$D(f) = 0$. 而$f(x)$在复数域上有重根当且仅当$f(x)$在复数域上有重因式, 这又等价于$(f(x), f'(x)) \neq 1$. 再利用$(f(x), f'(x)) \neq 1$当且仅当$f(x)$和$f'(x)$在复数域上有公共根, 我们得到

$$D(f) = 0 \Longleftrightarrow f(x)\text{和}f'(x)\text{在复数域上有公共根}.$$

[33]参看习题1.14.

下面研究两个多项式何时在复数域上有公共根.

设有$\mathbb{F}[x]$中的两个非常数多项式

$$\begin{aligned}f(x) =& a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0, \quad a_n \neq 0, n > 0,\\ g(x) =& b_mx^m + b_{m-1}x^{m-1} + \cdots + b_1x + b_0, \quad b_m \neq 0, m > 0.\end{aligned}$$

假设$f(x)$和$g(x)$在复数域上有公共根, 则存在$d(x) \in \mathbb{F}[x]$, $\deg(d(x)) > 0$, 使得

$$d(x) \mid f(x), \quad d(x) \mid g(x).$$

于是存在$f_1(x), g_1(x) \in \mathbb{F}[x]$, 使得

$$f(x) = d(x)f_1(x), \quad g(x) = d(x)g_1(x).$$

由此得

$$\deg(f_1(x)) < n, \quad \deg(g_1(x)) < m.$$

所以可以假设

$$\begin{aligned}f_1(x) &= c_{n-1}x^{n-1} + \cdots + c_1x + c_0,\\ g_1(x) &= d_{m-1}x^{m-1} + \cdots + d_1x + d_0.\end{aligned}$$

由于$f_1(x) \neq 0$且$g_1(x) \neq 0$, 所以

$$(c_{n-1}, \ldots, c_1, c_0) \neq 0, \quad (d_{m-1}, \ldots, d_1, d_0) \neq 0.$$

而

$$f(x)g_1(x) = d(x)f_1(x)g_1(x) = g(x)f_1(x),$$

比较系数得到

$$\left\{\begin{aligned} a_nd_{m-1} &= b_mc_{n-1},\\ a_{n-1}d_{m-1} + a_nd_{m-2} &= b_{m-1}c_{n-1} + b_mc_{n-2},\\ \vdots \quad &= \quad \vdots\\ a_0d_1 + a_1d_0 &= b_0c_1 + b_1c_0,\\ a_0d_0 &= b_0c_0.\end{aligned}\right.$$

于是系数矩阵为

$$A = \begin{pmatrix} a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 & & & \\ & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 & & & \\ & b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & & b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m\text{行} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n\text{行} \end{array}$$

的齐次线性方程组$XA=0$有非零解

$$(d_{m-1},\dots,d_1,d_0,-c_{n-1},\dots,-c_1,-c_0).$$

这得到系数行列式为零, 即$|A|=0$. 我们引入下面的定义.

定义1.23. $\mathbb{F}[x]$中的两个非常数多项式

$$\begin{aligned} f(x)=&a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0, \quad a_n\neq 0, n>0,\\ g(x)=&b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0, \quad b_m\neq 0, m>0 \end{aligned}$$

的**结式**(也称为**Sylvester行列式**)定义为

$$\operatorname{Res}(f,g):=\left|\begin{array}{ccccccccc} a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 & & \\ & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 & & \\ & b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & b_m & b_{m-1} & \cdots & \cdots & \cdots & \cdots & b_0 \end{array}\right| \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m\text{行} \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} n\text{行} \end{array}$$

所以由上面的讨论, 如果$f(x)$和$g(x)$在复数域上有公共根, 则$\operatorname{Res}(f,g)=0$. 事实上, 逆命题也成立.

定理1.41. 设有$\mathbb{F}[x]$中的多项式

$$\begin{aligned} f(x)=&a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0, \quad a_n\neq 0, n>0,\\ g(x)=&b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0, \quad b_m\neq 0, m>0, \end{aligned}$$

则$f(x)$和$g(x)$有公共复根的充分必要条件是$\operatorname{Res}(f,g)=0$.

☞ **证明.** 我们只要证明充分性. 假设$\operatorname{Res}(f,g)=0$, 则上面的齐次线性方程组$XA=0$有非零解. 设

$$(d_{m-1},\dots,d_1,d_0,-c_{n-1},\dots,-c_1,-c_0)$$

是一个非零解. 定义

$$\begin{aligned} f_1(x)&=c_{n-1}x^{n-1}+\cdots+c_1x+c_0,\\ g_1(x)&=d_{m-1}x^{m-1}+\cdots+d_1x+d_0, \end{aligned}$$

则可得

$$f(x)g_1(x)=g(x)f_1(x).$$

如果$(f(x),g(x))=1$, 则由上式可得$f(x)\mid f_1(x)$. 但是

$$\deg(f_1(x))<n, \quad \deg(f(x))=n,$$

所以只能$f_1(x)=0$. 这得到$f(x)g_1(x)=0$. 而$f(x)\neq 0$, 所以$g_1(x)=0$. 这得到

$$(d_{m-1},\ldots,d_1,d_0,-c_{n-1},\ldots,-c_1,-c_0)=0,$$

矛盾. 所以$(f(x),g(x))\neq 1$, 即$f(x)$和$g(x)$有公共复根. □

由于两个多项式有公共复根的充分必要条件是它们不互素, 所以

推论1.42. 设$f(x),g(x)\in\mathbb{F}[x]$, 则

$$(f(x),g(x))=1\Longleftrightarrow \mathrm{Res}(f,g)\neq 0.$$

下面是一个例子.

例1.40. 设$f(x)=x^3-x+2$, $g(x)=x^4+x-1$, 判断$f(x)$和$g(x)$有无公共复根.

☞ **解.** 由于

$$\mathrm{Res}(f,g)=\begin{vmatrix}1&0&-1&2&&&\\&1&0&-1&2&&\\&&1&0&-1&2&\\&&&1&0&-1&2\\1&0&0&1&-1&&\\&1&0&0&1&-1&\\&&1&0&0&1&-1\end{vmatrix}=11,$$

所以没有公共复根. □

可以用结式来求两个二元多项式$f(x,y),g(x,y)\in\mathbb{F}[x,y]$在$\mathbb{C}^2$中的公共零点. 我们将这两个二元多项式看成以$\mathbb{F}[y]$为系数的$x$的多项式

$$f(x,y),g(x,y)\in\mathbb{F}[y][x],$$

且设$f(x,y)$和$g(x,y)$都含有x, 则可以定义关于x的结式$\mathrm{Res}_x(f,g)\in\mathbb{F}[y]$. 类似地, 当二元多项式$f(x,y)$和$g(x,y)$都含有$y$时, 可以定义关于$y$的结式$\mathrm{Res}_y(f,g)\in\mathbb{F}[x]$. 请读者证明可以按照如下步骤求两个二元多项式在$\mathbb{C}^2$上的公共零点.

算法1.3. 求$f(x,y),g(x,y)\in\mathbb{F}[x,y]$在$\mathbb{C}^2$中的公共零点的步骤:

(1) 求$\mathrm{Res}_x(f,g)$(求$\mathrm{Res}_y(f,g)$);

(2) 求$\mathrm{Res}_x(f,g)$的所有复根(求$\mathrm{Res}_y(f,g)$的所有复根);

(3) 对$\mathrm{Res}_x(f,g)$的每个复根y_0, 求$f(x,y_0)$和$g(x,y_0)$的所有公共根

(对$\mathrm{Res}_y(f,g)$的每个复根x_0, 求$f(x_0,y)$和$g(x_0,y)$的所有公共根);

(4) 写出在$\mathbb{C}^2$中的公共复根.

我们看一个例子.

例1.41. 求方程组

$$\begin{cases} x^2 - xy + y^2 - 3 = 0, \\ x^2 + xy + 2y^2 + y - 1 = 0 \end{cases}$$

的有理根.

☞ **解.** 等价于求$f(x,y) = x^2 - xy + y^2 - 3$和$g(x,y) = x^2 + xy + 2y^2 + y - 1$的公共有理根. 由于

$$f(x,y) = x^2 - yx + (y^2 - 3), \quad g(x,y) = x^2 + yx + (2y^2 + y - 1),$$

所以

$$\begin{aligned}
\mathrm{Res}_x(f,g) &= \begin{vmatrix} 1 & -y & y^2-3 & 0 \\ 0 & 1 & -y & y^2-3 \\ 1 & y & 2y^2+y-1 & 0 \\ 0 & 1 & y & 2y^2+y-1 \end{vmatrix} \\
&\xlongequal[r_4-r_2]{r_3-r_1} \begin{vmatrix} 1 & -y & y^2-3 & 0 \\ 0 & 1 & -y & y^2-3 \\ 0 & 2y & y^2+y+2 & 0 \\ 0 & 0 & 2y & y^2+y+2 \end{vmatrix} = \begin{vmatrix} 1 & -y & y^2-3 \\ 2y & y^2+y+2 & 0 \\ 0 & 2y & y^2+y+2 \end{vmatrix} \\
&\xlongequal{c_2+yc_1} \begin{vmatrix} 1 & 0 & y^2-3 \\ 2y & 3y^2+y+2 & 0 \\ 0 & 2y & y^2+y+2 \end{vmatrix} \\
&= (3y^2+y+2)(y^2+y+2) + 4y^2(y^2-3) \\
&= 7y^4 + 4y^3 - 3y^2 + 4y + 4.
\end{aligned}$$

验证可知-1是$\mathrm{Res}_x(f,g)$的一个根. 而

$$\mathrm{Res}_x(f,g) = (y+1)(7y^3 - 3y^2 + 4),$$

且$7y^3 - 3y^2 + 4$无有理根, 所以-1是$\mathrm{Res}_x(f,g)$的唯一的有理根.

将$y = -1$代入原方程组, 得到

$$\begin{cases} x^2 + x - 2 = 0, \\ x^2 - x = 0. \end{cases}$$

它的解为$x = 1$. 所以原方程组的有理数解为$(x,y) = (1,-1)$. □

设$f(x), g(x) \in \mathbb{F}[x]$分别是$n$和$m$次多项式, 其中$n, m > 0$. 上面我们用$\mathrm{Res}(f,g)$是否为零来判断$f(x)$和$g(x)$是否有公共的复根. 我们设$f(x)$的所有复根为$\alpha_1, \alpha_2, \ldots, \alpha_n$, 而$g(x)$的所有复根为$\beta_1, \beta_2, \ldots, \beta_m$, 则

$$f(x)\text{和}g(x)\text{有公共的复根} \Longleftrightarrow \prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j) = 0.$$

于是, 我们有理由相信$\mathrm{Res}(f,g)$应该和$\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{m}(\alpha_i - \beta_j)$有关系. 这正是下面定理所陈述的结论.

定理1.43. 设有$\mathbb{F}[x]$中的多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad a_n \neq 0, n > 0,$$

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, \quad b_m \neq 0, m > 0,$$

其中$f(x)$的n个复根为$\alpha_1, \alpha_2, \ldots, \alpha_n$, 而$g(x)$的$m$个复根为$\beta_1, \beta_2, \ldots, \beta_m$, 则

$$\operatorname{Res}(f,g) = a_n^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{nm} b_m^n \prod_{j=1}^{m} f(\beta_j) = a_n^m b_m^n \prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j).$$

☞ **证明.** 我们有

$$\begin{aligned}
f(x) &= a_n(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)\\
&= a_n(x^n - \sigma_1(\alpha_1,\ldots,\alpha_n)x^{n-1} + \cdots + (-1)^{n-1}\sigma_{n-1}(\alpha_1,\ldots,\alpha_n)x + (-1)^n\sigma_n(\alpha_1,\ldots,\alpha_n)),\\
g(x) &= b_m(x-\beta_1)(x-\beta_2)\cdots(x-\beta_m)\\
&= b_m(x^m - \sigma_1(\beta_1,\ldots,\beta_m)x^{m-1} + \cdots + (-1)^{m-1}\sigma_{m-1}(\beta_1,\ldots,\beta_m)x + (-1)^m\sigma_m(\beta_1,\ldots,\beta_m)).
\end{aligned}$$

所以$\operatorname{Res}(f,g)$等于

$$\begin{vmatrix}
a_n & -a_n\sigma_1(\alpha_1,\ldots,\alpha_n) & \cdots & a_n(-1)^n\sigma_n(\alpha_1,\ldots,\alpha_n) & & \\
 & \cdots & \cdots & \cdots & \cdots & \\
 & & a_n & -a_n\sigma_1(\alpha_1,\ldots,\alpha_n) & \cdots & a_n(-1)^n\sigma_n(\alpha_1,\ldots,\alpha_n)\\
b_m & -b_m\sigma_1(\beta_1,\ldots,\beta_m) & \cdots & b_m(-1)^m\sigma_m(\beta_1,\ldots,\beta_m) & & \\
 & \cdots & \cdots & \cdots & \cdots & \\
 & & b_m & -b_m\sigma_1(\beta_1,\ldots,\beta_m) & \cdots & b_m(-1)^m\sigma_m(\beta_1,\ldots,\beta_m)
\end{vmatrix}.$$

我们将$\alpha_1,\ldots,\alpha_n$和$\beta_1,\ldots,\beta_m$看成不定元, 则可以认为

$$\operatorname{Res}(f,g) \in \mathbb{F}[\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_m].$$

下面先确定这个多项式的次数. 将该行列式的第一行乘α_1, 第二行乘α_1^2, …, 第m行乘α_1^m; 第$m+1$行乘α_1, 第$m+2$行乘α_1^2, …, 第$m+n$行乘α_1^n, 得到行列式D. 则D的第i列是i次齐次多项式, 进而D是

$$1 + 2 + \cdots + (m+n) = \frac{1}{2}(m+n)(m+n+1)$$

次齐次多项式. 而

$$D = \alpha_1^{(1+\cdots+m)+(1+\cdots+n)}\operatorname{Res}(f,g) = \alpha_1^{\frac{m(m+1)+n(n+1)}{2}}\operatorname{Res}(f,g),$$

所以$\operatorname{Res}(f,g)$是

$$\frac{1}{2}[(m+n)(m+n+1) - m(m+1) - n(n+1)] = mn$$

次齐次多项式.

设$1 \leqslant i \leqslant n$. 将行列式$\operatorname{Res}(f,g)$的第一列乘$\alpha_i^{n+m-1}$, 第二列乘$\alpha_i^{n+m-2}$, …, 第$n+m-1$列乘$\alpha_i$, 然后都加到第$n+m$列上, 得到的新的行列式的最后一列是

$$\begin{pmatrix} a_n\alpha_i^{m-1}f(\alpha_i)\\ a_n\alpha_i^{m-2}f(\alpha_i)\\ \vdots\\ a_nf(\alpha_i)\\ b_m\alpha_i^{n-1}g(\alpha_i)\\ b_m\alpha_i^{n-2}g(\alpha_i)\\ \vdots\\ b_mg(\alpha_i)\end{pmatrix} = \begin{pmatrix} 0\\ 0\\ \vdots\\ 0\\ b_m\alpha_i^{n-1}g(\alpha_i)\\ b_m\alpha_i^{n-2}g(\alpha_i)\\ \vdots\\ b_mg(\alpha_i)\end{pmatrix}.$$

所以有

$$g(\alpha_i) \mid \operatorname{Res}(f,g).$$

由于当$i \neq j$时, $g(\alpha_i)$和$g(\alpha_j)$没有相同的一次因式, 所以由唯一分解定理可得

$$\prod_{i=1}^{n} g(\alpha_i) \mid \operatorname{Res}(f,g).$$

而$\prod\limits_{i=1}^{n} g(\alpha_i)$的次数为$mn$, 所以

$$\operatorname{Res}(f,g) = c\prod_{i=1}^{n} g(\alpha_i), \quad (\exists c \in \mathbb{F}).$$

取$\alpha_1 = \cdots = \alpha_n = 0$, $\beta_1 = \cdots = \beta_m = 1$, 则$\operatorname{Res}(f,g)$成为下三角行列式, 对角元为

$$\underbrace{a_n, \ldots, a_n}_{m}, \underbrace{(-1)^m b_m, \ldots, (-1)^m b_m}_{n},$$

所以此时

$$\operatorname{Res}(f,g) = (-1)^{nm} a_n^m b_m^n.$$

而此时

$$g(\alpha_i) = b_m(-1)^m.$$

所以

$$(-1)^{nm} a_n^m b_m^n = c b_m^n (-1)^{nm} \Longrightarrow c = a_n^m.$$

因此

$$\operatorname{Res}(f,g) = a_n^m \prod_{i=1}^{n} g(\alpha_i).$$

由此可得另两个等号成立. □

利用定理1.43中结式的表达式, 可以证明结式满足的一些性质.

例1.42. 设$f(x), f_1(x), f_2(x), g(x), g_1(x), g_2(x) \in \mathbb{F}[x]$都是正次数多项式, 证明:

(1) $\operatorname{Res}(f,g) = (-1)^{mn} \operatorname{Res}(g,f)$, $(n = \deg(f(x)), m = \deg(g(x)))$;

(2) $\operatorname{Res}(f_1 f_2, g) = \operatorname{Res}(f_1, g) \operatorname{Res}(f_2, g)$;

(3) $\operatorname{Res}(f, g_1 g_2) = \operatorname{Res}(f, g_1) \operatorname{Res}(f, g_2)$.

☞ **证明.** (1) 由定理1.43中结式的表达式

$$\operatorname{Res}(f,g) = a_n^m b_m^n \prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j)$$

得到.

(2) 设$\deg(f_i(x)) = n_i$, $\deg(g(x)) = m$, $g(x)$的首项系数为b_m, m个复根为$\beta_1, \ldots, \beta_m$, 则由定理1.43, 得到

$$\begin{aligned}\operatorname{Res}(f_1 f_2, g) =& (-1)^{(n_1+n_2)m} b_m^{n_1+n_2} \prod_{j=1}^{m} (f_1 f_2)(\beta_j) = (-1)^{(n_1+n_2)m} b_m^{n_1+n_2} \prod_{j=1}^{m} f_1(\beta_j) f_2(\beta_j) \\ =& \left[(-1)^{n_1 m} b_m^{n_1} \prod_{j=1}^{m} f_1(\beta_j)\right]\left[(-1)^{n_2 m} b_m^{n_2} \prod_{j=1}^{m} f_2(\beta_j)\right]\end{aligned}$$

$$= \mathrm{Res}(f_1, g)\,\mathrm{Res}(f_2, g).$$

(3) 类似于(2)证明. 或者利用(1)和(2). □

前面提到, 为了研究多项式$f(x)$的判别式, 我们需要考虑$f(x)$和$f'(x)$在复数域上有无公共根. 而两个多项式有无公共根可以用它们的结式判别, 于是$f(x)$的判别式应该与结式$\mathrm{Res}(f, f')$有关系, 这正是下面结论所揭示的.

定理1.44. 设$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{F}[x]$, $a_n \neq 0$, $n > 1$, 则

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_n^{-1}\,\mathrm{Res}(f, f').$$

☞ **证明.** 设$f(x)$的n个复根为$\alpha_1, \ldots, \alpha_n$, 则

$$f(x) = a_n(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n).$$

于是得到

$$f'(x) = a_n \sum_{i=1}^{n} \prod_{\substack{j=1 \\ j\neq i}} (x-\alpha_j).$$

这得到

$$f'(\alpha_i) = a_n \prod_{\substack{j=1 \\ j\neq i}} (\alpha_i-\alpha_j).$$

由定理1.43, 我们有

$$\begin{aligned}
\mathrm{Res}(f, f') =& a_n^{n-1} \prod_{i=1}^{n} f'(\alpha_i) = a_n^{2n-1} \prod_{i=1}^{n} \prod_{\substack{j=1 \\ j\neq i}}^{n} (\alpha_i-\alpha_j) \\
=& a_n^{2n-1} \prod_{1\leqslant i\neq j\leqslant n} (\alpha_i-\alpha_j) = a_n^{2n-1} \prod_{1\leqslant j<i\leqslant n} (\alpha_i-\alpha_j)(\alpha_j-\alpha_i) \\
=& (-1)^{\frac{n(n-1)}{2}} a_n^{2n-1} \prod_{1\leqslant j<i\leqslant n} (\alpha_i-\alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} a_n D(f).
\end{aligned}$$

由此得结论成立. □

我们看几个例子.

例1.43. 设$f(x), g(x) \in \mathbb{F}[x]$是正次数多项式, 证明:

$$D(fg) = D(f)D(g)\left[\mathrm{Res}(f, g)\right]^2.$$

☞ **证明.** 设$\deg(f(x)) = n$, $\deg(g(x)) = m$, 且

$$\begin{aligned}
f(x) &= a_n(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n), \quad (a_n \neq 0), \\
g(x) &= b_m(x-\beta_1)(x-\beta_2)\cdots(x-\beta_m), \quad (b_m \neq 0),
\end{aligned}$$

则

$$f(x)g(x) = a_n b_m (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)(x-\beta_1)(x-\beta_2)\cdots(x-\beta_m).$$

所以有

$$D(fg) = (a_n b_m)^{2(n+m)-2} \prod_{1\leqslant j<i\leqslant n} (\alpha_i-\alpha_j)^2 \prod_{1\leqslant j<i\leqslant m} (\beta_i-\beta_j)^2 \prod_{i=1}^{n}\prod_{j=1}^{m} (\alpha_i-\beta_j)^2$$

$$=D(f)D(g)\left[\mathrm{Res}(f,g)\right]^2,$$

证毕. □

例1.44. 求多项式$f(x)$的判别式, 其中

(1) $f(x)=x^n+ax+b$;

(2) $f(x)=x^{n-1}+x^{n-2}+\cdots+x+1,\ n>1$.

☞ **解.** (1) 我们有

$$f'(x)=nx^{n-1}+a,$$

于是

$$\mathrm{Res}(f,f')=\begin{vmatrix}1&0&\cdots&0&a&b&0&\cdots&0&0\\0&1&\cdots&0&0&a&b&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&0&0&0&\cdots&a&b\\n&0&\cdots&0&a&0&0&\cdots&0&0\\0&n&\cdots&0&0&a&0&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&0&n&0&0&\cdots&0&a\end{vmatrix}$$

$$\xlongequal{r_n-nr_1}\begin{vmatrix}1&0&\cdots&0&a&b&0&\cdots&0&0\\0&1&\cdots&0&0&a&b&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&0&0&0&\cdots&a&b\\0&0&\cdots&0&(1-n)a&-nb&0&\cdots&0&0\\0&n&\cdots&0&0&a&0&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&0&n&0&0&\cdots&0&a\end{vmatrix}$$

$$=\begin{vmatrix}1&\cdots&0&0&a&b&\cdots&0&0\\\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&\cdots&1&0&0&0&\cdots&a&b\\0&\cdots&0&(1-n)a&-nb&0&\cdots&0&0\\n&\cdots&0&0&a&0&\cdots&0&0\\\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&\cdots&0&n&0&0&\cdots&0&a\end{vmatrix}.$$

继续这一过程: r_n-nr_1, 再按照第一列展开, 一共$n-1$次后, 得到

$$\mathrm{Res}(f,f')=\begin{vmatrix}(1-n)a&-nb&0&0&\cdots&0&0\\0&(1-n)a&-nb&0&\cdots&0&0\\0&0&(1-n)a&-nb&\cdots&0&0\\\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&0&\cdots&(1-n)a&-nb\\n&0&0&0&\cdots&0&a\end{vmatrix}$$

$$
\begin{aligned}
&= n(-1)^{n+1}(-nb)^{n-1} + a[(1-n)a]^{n-1} \\
&= (-1)^{n-1}(n-1)^{n-1}a^n + n^n b^{n-1}.
\end{aligned}
$$

所以

$$D(f) = (-1)^{\frac{n(n-1)}{2}} \operatorname{Res}(f, f') = (-1)^{\frac{n(n-1)}{2}} \left[(-1)^{n-1}(n-1)^{n-1}a^n + n^n b^{n-1}\right].$$

(2) 设$g(x) = x - 1$, 则

$$f(x)g(x) = x^n - 1.$$

由(1), 可得

$$D(fg) = (-1)^{\frac{n(n-1)}{2}} n^n (-1)^{n-1} = (-1)^{\frac{(n-1)(n-2)}{2}} n^n.$$

由$D(g) = 1$, 以及

$$\operatorname{Res}(f,g) = \begin{vmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & & \ddots & \ddots & \\ & & & & 1 & -1 \end{vmatrix} \xlongequal[i=2,\ldots,n]{c_1 + c_i} \begin{vmatrix} n & 1 & 1 & 1 & \cdots & 1 \\ 0 & -1 & & & & \\ & 1 & -1 & & & \\ & & & \ddots & \ddots & \\ & & & & 1 & -1 \end{vmatrix} = (-1)^{n-1} n,$$

结合例1.43, 就得到

$$D(f) = (-1)^{\frac{(n-1)(n-2)}{2}} n^{n-2}.$$

解毕. □

结式还可以化平面曲线的参数方程为直角坐标方程.

例1.45. 求曲线

$$x = \frac{-t^2 + 2t}{t^2 + 1}, \quad y = \frac{2t^2 + 2t}{t^2 + 1}$$

的直角坐标方程.

☞ **解.** 任取曲线上的一点(x, y), 则存在$t_0 \in \mathbb{R}$, 使得

$$x = \frac{-t_0^2 + 2t_0}{t_0^2 + 1}, \quad y = \frac{2t_0^2 + 2t_0}{t_0^2 + 1},$$

即

$$(x+1)t_0^2 - 2t_0 + x = 0, \quad (y-2)t_0^2 - 2t_0 + y = 0.$$

令

$$f(t) = (x+1)t^2 - 2t + x, \quad g(t) = (y-2)t^2 - 2t + y \in \mathbb{R}[t],$$

则$f(t)$和$g(t)$有公共根t_0, 这得到$\operatorname{Res}(f, g) = 0$.

反之, 如果(x, y)满足$\operatorname{Res}(f, g) = 0$, 则$f(t)$和$g(t)$不互素. 而$f(t)$和$g(t)$不相伴, 且次数至多为2, 所以$f(t)$和$g(t)$有公共根$t_1 \in \mathbb{R}$. 因此$(x, y)$在该曲线上. 因此, 该曲线的直角坐标方程为$\operatorname{Res}(f, g) = 0$. 下面具体计算.

当$x \neq -1$且$y \neq 2$时, 我们有

$$\operatorname{Res}(f,g) = \begin{vmatrix} x+1 & -2 & x & 0 \\ 0 & x+1 & -2 & x \\ y-2 & -2 & y & 0 \\ 0 & y-2 & -2 & y \end{vmatrix} = 8x^2 - 4xy + 5y^2 + 12x - 12y;$$

当$x \neq -1$而$y=2$时, 我们有

$$\mathrm{Res}(f,g)=\begin{vmatrix} x+1 & -2 & x \\ -2 & 2 & 0 \\ 0 & -2 & 2 \end{vmatrix}=4(2x-1);$$

当$x=-1$而$y \neq 2$时, 我们有

$$\mathrm{Res}(f,g)=\begin{vmatrix} -2 & -1 & 0 \\ 0 & -2 & -1 \\ y-2 & -2 & y \end{vmatrix}=5y+2;$$

当$x=-1$而$y=2$时, 我们有

$$\mathrm{Res}(f,g)=\begin{vmatrix} -2 & -1 \\ -2 & 2 \end{vmatrix}=-6.$$

由于$(-1,2),(1/2,2),(-1,-2/5)$都满足$8x^2-4xy+5y^2+12x-12y=0$, 所以曲线的直角坐标方程为

$$8x^2-4xy+5y^2+12x-12y=0,$$

并且$(x,y) \neq (-1,2)$. □

§ 1.19 附录: 群, 环, 域和主理想整环

这里简要介绍群, 环, 域和主理想整环的定义, 并给出一些例子. 数域是域, 而域上的一元多项式环是主理想整环.

1.19.1 群

群是最简单的代数结构, 它是Galosi在研究多项式的根式可解性时引入的, 这标志着代数学从研究线性方程组等转向研究代数结构.

定义1.24. *设G是非空集合, G上有一个二元运算: 对任意的$a,b \in G$, 存在唯一的$ab=a\cdot b \in G$与之对应, 且满足下面公理*

(G1) *(结合律) 对任意$a,b,c \in G$, 有$(ab)c=a(bc)$;*

(G2) *(存在单位元) 存在$e \in G$, 使得对任意$a \in G$, 有$ae=ea=a$;*

(G3) *(存在逆元) 对任意$a \in G$, 存在$b \in G$, 使得$ab=ba=e$,*

*则称G为一个**群**. 如果还满足*

(G4) *(交换律) 对任意$a,b \in G$, 有$ab=ba$,*

*则称G是**交换群**, 也称为**Abel群**.*

容易证明, 公理(G2)中的e唯一, 公理(G3)中, 满足$ab=ba=e$的b唯一, 称其为a的逆, 记为a^{-1}.

上面定义中群的运算写成了乘法, 有时也将群的运算写成加法, 这时候通常称群中的单位元为零元.

我们看一些例子.

例1.46. **(1)** $\mathbb{Z}$(或$\mathbb{Q}$, 或$\mathbb{R}$, 或$\mathbb{C}$)在加法下成为无限$Abel$群;

(2) $\mathbb{Q}$(或$\mathbb{R}$, 或$\mathbb{C}$)的非零元全体在乘法下成为无限$Abel$群;

(3) $\mathbb{N}$在加法下不成为群, 比如没有零元;

(3) 非零整数全体在乘法下不成为群, 比如整数2没有乘法逆.

例1.47. 设S是一个非空集合, 集合S到S的双射全体做成的集合记为$\mathrm{Aut}(S)$, 则$\mathrm{Aut}(S)$在映射的复合下成为群, 其单位元为S的恒等映射. 称$\mathrm{Aut}(S)$为S的置换群, 该群中的元素称为S上的置换. 如果$S=\{1,2,\dots,n\}$, 则将$\mathrm{Aut}(S)$记为S_n, 并称为n阶对称群.

例1.48. 设m是一个固定的正整数. 对整数a,b, 如果$m\mid(a-b)$, 则称a和b模m同余, 并记为$a\equiv b \pmod m$. 可以证明模m同余关系是$\mathbb{Z}$上的一个等价关系, 于是这得到整数集的模m分类. 具体的, 对$a\in\mathbb{Z}$, 记

$$\bar{a}=\{b\in\mathbb{Z}\mid b\equiv a \pmod m\},$$

称其为a所在的模m同余类. 所有模m同余类所成的集合记为$\mathbb{Z}/m\mathbb{Z}$, 则在加法

$$\bar{a}+\bar{b}:=\overline{a+b},\quad a,b\in\mathbb{Z}$$

下, $\mathbb{Z}/m\mathbb{Z}$成为$Abel$群.

为了和后面的线性空间比较, 我们这里都定义一下子结构和同态的概念.

保持代数结构的运算的映射就称为同态, 具体地, 对群有定义

定义1.25. 设G和H是群, 如果映射$\varphi: G\to H$满足, 对任意$a,b\in G$, 有

$$\varphi(ab)=\varphi(a)\varphi(b),$$

则称φ为一个群同态. 此时, 有$\varphi(e_G)=e_H$, 且对任意$a\in G$, 有$\varphi(a^{-1})=\varphi(a)^{-1}$.

子结构就是对运算封闭的非空子集. 具体的

定义1.26. 设G是群, H是G的非空子集, 如果

(i) 对任意$a,b\in H$, 有$ab\in H$;

(ii) 对任意$a\in H$, 有$a^{-1}\in H$,

则称H为G的一个子群. 此时, H在G的运算下也成为群.

1.19.2 环

群只有一个运算, 而环则有两个运算.

定义1.27. 设非空集合R有两个二元运算, 一个记为加法+, 一个记为乘法, 满足

(R1) $(R, +)$是$Abel$群;

(R2) (结合律) 对任意$a, b, c \in R$, 有$(ab)c = a(bc)$;

(R3) (分配律) 对任意a, b, c, 有

$$a(b+c) = ab + ac, \qquad (a+b)c = ac + bc,$$

则称R是一个环. 如果环R还满足

(R4) (交换律) 对任意$a, b \in R$, 有$ab = ba$,

则称R是交换环. 如果环R还满足

(R5) 存在$1_R \in R$, 使得对任意$a \in R$, 有

$$1_R a = a 1_R = a,$$

则称R是含幺环.

设R是环, R中有些元比较特殊. 设$0 \neq a, b \in R$, 如果$ab = 0$, 则称a是R的一个左零因子, 而b是R的一个右零因子. 即是左零因子, 又是右零因子的元称为R的零因子.

定义1.28. 设R是含幺交换环, 且满足

(i) $1_R \neq 0$;

(ii) R无零因子,

则称R是整环.

下面看一些环的例子.

例1.49. $\mathbb{Z}$(或$\mathbb{Q}$, 或$\mathbb{R}$, 或$\mathbb{C}$)在数的加法和乘法下成为整环. 称$\mathbb{Z}$为整数环. 设$\mathbb{F}$是数域, 则$\mathbb{F}[x]$在多项式的加法和乘法下成为整环, 称其为一元多项式环.

例1.50. 设m是固定的正整数, 则$\mathbb{Z}/m\mathbb{Z}$在模m同余类的加法和乘法:

$$\bar{a} + \bar{b} = \overline{a+b}, \quad \bar{a}\bar{b} = \overline{ab}, \quad a, b \in \mathbb{Z}$$

下成为含幺交换环. 称$\mathbb{Z}/m\mathbb{Z}$为模m同余类环. 可以证明, $\mathbb{Z}/m\mathbb{Z}$是整环的充分必要条件是, m是素数.

例1.51. 设A是一个$Abel$群, 令$\mathrm{End}(A)$为A到A的群同态全体做成的集合, 则$\mathrm{End}(A)$在加法和乘法:

$$(f+g)(a)=f(a)+g(a),\quad (fg)(a)=f(g(a)),\quad \forall f,g\in \mathrm{End}(A), \forall a\in A$$

下成为含幺环, 称之为A的自同态环.

本小节最后定义环同态和子环.

定义1.29. 设R,S是环, 如果映射$f:R\to S$满足: 对任意$a,b\in R$, 有

$$f(a+b)=f(a)+f(b),\qquad f(ab)=f(a)f(b),$$

则称f是从R到S的一个环同态.

定义1.30. 设R是环, R的非空子集S满足: 对任意$a,b\in S$, 有

$$a-b\in S,\quad ab\in S,$$

则称S是R的一个子环.

1.19.3 域

如果R是含幺环, $a\in R$满足: 存在$b\in R$, 使得$ab=ba=1_R$, 则称a是R的一个可逆元. 此时, b唯一, 称为a的逆, 并记为a^{-1}. 环R的所有可逆元做成的集合记为$R^\times$, 可以证明$R^\times$在R的乘法下成为群, 称为R的单位群.

例1.52. 我们有

$$\mathbb{Z}^\times=\{1,-1\},\quad \mathbb{Q}^\times=\mathbb{Q}-\{0\},\quad \mathbb{R}^\times=\mathbb{R}-\{0\},\quad \mathbb{C}^\times=\mathbb{C}-\{0\},$$

和

$$(\mathbb{F}[x])^\times=\mathbb{F}^\times=\mathbb{F}-\{0\},\quad (\mathbb{Z}/m\mathbb{Z})^\times=\{\bar{a}\mid a\in\mathbb{Z},(a,m)=1\}.$$

如果环的乘法群最大, 则成为体.

定义1.31. 设D是含幺环, 如果$1_D\neq 0$, 且D的任意非零元都是可逆元, 则称D是体(也称为除环). 交换体称为域.

例1.53. $\mathbb{Q}$, $\mathbb{R}$和$\mathbb{C}$都是域, 分别称为有理数域, 实数域和复数域. 设p为素数, 则$\mathbb{Z}/p\mathbb{Z}$为域, 称为模p剩余类域.

1.19.4 主理想整环

设R是环, I是R的非空子集, 满足

(i) 对任意$a, b \in I$, 有$a - b \in I$;

(ii) 对任意$a \in I$和$r \in R$, 有$ar \in I$和$ra \in I$,

则称I是R的一个理想.

上面的条件(i)相当于说I对R中的加法成为Abel群.

设R是含幺交换环, S是R的非空子集, 则集合

$$\langle S \rangle := \{r_1a_1 + \cdots + r_sa_s \mid s \geqslant 1, r_1, \ldots, r_s \in R, a_1, \ldots, a_s \in S\}$$

是R的理想, 称为由S生成的理想. 设$a \in R$, 称

$$\langle a \rangle = Ra = \{ra \mid r \in R\}$$

为由a生成的主理想.

定义1.32. *每个理想都是主理想的整环称为主理想整环.*

例1.54. 证明: *整数环是主理想整环.*

☞ **证明.** 设I是$\mathbb{Z}$的非零理想, 则I中含正整数, 令m是I中的最小正整数, 我们证明: $I = \langle m \rangle$. 事实上, 由$m \in I$, 知$\langle m \rangle \subset I$. 反之, 设$a \in I$, 由带余除法, 存在$q, r \in \mathbb{Z}$, 使得

$$a = mq + r, \quad 0 \leqslant r < m.$$

由$a, m \in I$, 得$r \in I$. 再由m的最小性得$r = 0$, 即$a = mq \in \langle m \rangle$. 所以$I = \langle m \rangle$是主理想. □

类似有

例1.55. 设$\mathbb{F}$是域, 证明: *一元多项式环$\mathbb{F}[x]$是主理想整环.*

☞ **证明.** 首先, $\mathbb{F}[x]$是整环. 其次, 任取$\mathbb{F}[x]$的非零理想I, 我们证明I是主理想. 设$m(x)$是I中次数最低的首一多项式, 它是唯一的. 我们有$\langle m(x) \rangle \subset I$. 反之, 任取$f(x) \in I$, 由带余除法, 存在$q(x), r(x) \in \mathbb{F}[x]$, 使得

$$f(x) = q(x)m(x) + r(x), \quad \deg(r(x)) < \deg(m(x)).$$

由$f(x), m(x) \in I$, 得$r(x) \in I$. 再由$m(x)$的次数最低性, 得$r(x) = 0$. 所以$f(x) = q(x)m(x) \in \langle m(x) \rangle$. 于是$I = \langle m(x) \rangle$是主理想. □

例1.56. 设$\mathbb{F}$是域, $f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$, 证明:

$$\langle f_1(x), f_2(x), \ldots, f_s(x) \rangle = \langle (f_1(x), f_2(x), \ldots, f_s(x)) \rangle.$$

☞ **证明.** 记$d(x)=(f_1(x),f_2(x),\dots,f_s(x))$, 则存在$u_1(x),u_2(x),\dots,u_s(x)\in\mathbb{F}[x]$, 使得

$$d(x)=u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x)\in\langle f_1(x),f_2(x),\dots,f_s(x)\rangle.$$

于是$\langle d(x)\rangle\subset\langle f_1(x),f_2(x),\dots,f_s(x)\rangle$. 反之, 对任意$r_1(x),r_2(x),\dots,r_s(x)\in\mathbb{F}[x]$, 有

$$\begin{aligned}&r_1(x)f_1(x)+r_2(x)f_2(x)+\cdots+r_s(x)f_s(x)\\=&\left(r_1(x)\frac{f_1(x)}{d(x)}+r_2(x)\frac{f_2(x)}{d(x)}+\cdots+r_s(x)\frac{f_s(x)}{d(x)}\right)d(x)\in\langle d(x)\rangle,\end{aligned}$$

即有$\langle f_1(x),f_2(x),\dots,f_s(x)\rangle\subset\langle d(x)\rangle$. □

整数环和一元多项式环中的唯一分解定理可以推广到主理想整环上(事实上, 可以推广到所谓的唯一分解整环), 为此需要将整数环和一元多项式环中的整除理论一般化. 设R是整环, $a,b\in R$. 如果存在$c\in R$, 使得$a=bc$, 则称b**整除**a, 记为$b\mid a$. 如果存在$u\in R^\times$, 使得$a=ub$, 则称a和b**相伴**. 容易证明, a和b相伴当且仅当a和b相互整除.

定义1.33. *设$p\in R$非零非单位.*

(1) *如果$p=ab$, 其中$a,b\in R$, 可导出$a\in R^\times$或者$b\in R^\times$, 则称p是**不可约元**;*

(2) *如果$p\mid ab$, 其中$a,b\in R$, 可导出$p\mid a$或者$p\mid b$, 则称p是**素元**.*

整数环中的不可约元和素元一致, 就是素数及其相反数; 一元多项式环中的不可约元和素元也一致, 就是不可约多项式. 这个对任意的主理想整环都成立.

引理1.45. *设R是整环. 则*

(1) *R中的素元一定是不可约元;*

(2) *如果R是主理想整环, 则R中的不可约元是素元.*

☞ **证明.** (1) 设p是素元. 令$p=ab$, 其中$a,b\in R$, 则$p\mid ab$. 于是$p\mid a$或者$p\mid b$. 不妨设$p\mid a$, 于是存在$c\in R$, 使得$a=pc$. 代入得到$p=pcb$, 进而由$p\neq0$得$cb=1$. 所以$b\in R^\times$, 得到p是不可约元.

(2) 设p是R中的不可约元, 先证明$\langle p\rangle$是R的**极大理想**, 即$\langle p\rangle\neq R$, 且不存在R的理想I, 使得$\langle p\rangle\subsetneqq I\subsetneqq R$.

事实上, 由于$p\notin R^\times$, 所以$\langle p\rangle\neq R$. 设$I=\langle a\rangle$是R的理想, 且$\langle p\rangle\subset\langle a\rangle\subset R$. 因为$p\in\langle p\rangle$, 所以$p\in\langle a\rangle$, 于是存在$b\in R$, 使得$p=ab$. 而$p$不可约, 所以$a\in R^\times$或者$b\in R^\times$. 如果$a\in R^\times$, 则$\langle a\rangle=R$; 如果$b\in R^\times$, 则$a\in\langle p\rangle$, 进而$\langle a\rangle=\langle p\rangle$. 所以$\langle p\rangle$是极大理想.

下设$p\mid ab$, 其中$a,b\in R$. 如果$p\nmid a$, 我们证明$p\mid b$. 此时, 有$\langle p\rangle\subsetneqq\langle p,a\rangle$. 由于$\langle p\rangle$是极大理想, 所以$\langle p,a\rangle=R$. 因此存在$x,y\in R$, 使得

$$xp+ya=1.$$

这得到$b=bxp+yab$. 由$p\mid ab$, 就得$p\mid b$. □

还需要证明主理想整环是Noether环.

定义1.34. *设R是环.*

(1) *如果R的理想序列$I_1,I_2,\dots$满足*

$$I_1\subset I_2\subset\cdots\subset I_i\subset I_{i+1}\subset\cdots,$$

*则称$\{I_i\}$为一个**理想升链**;*

(2) 称R的理想升链$\{I_i\}$是**稳定的**, 如果存在$k \in \mathbb{N}$, 使得

$$I_k = I_{k+1} = I_{k+2} = \cdots;$$

(3) 如果R的任意理想升链都是稳定的, 则称R是**Noether环**.

命题1.46. 主理想整环是$Noether$环.

证明. 设R是主理想整环, $\{I_i\}$是R的一个理想升链. 令$I = \bigcup\limits_{i \geqslant 1} I_i$, 我们证明$I$是$R$的理想.

事实上, 对任意$a, b \in I$和$r \in R$, 存在$i, j \geqslant 1$, 使得$a \in I_i$, $b \in I_j$. 不妨设$i \leqslant j$, 则

$$a, b \in I_j.$$

于是

$$a - b \in I_j \subset I.$$

又

$$ra \in I_i \subset I,$$

所以I是理想.

因为R是主理想整环, 所以存在$c \in R$, 使得$I = \langle c \rangle$. 存在$k \geqslant 1$, 使得$c \in I_k$. 于是

$$I \subset I_k \subset I_{k+1} \subset \cdots \subset I.$$

这得到

$$I = I_k = I_{k+1} = \cdots,$$

即理想升链$\{I_i\}$稳定. □

下面可以证明主理想整环的唯一分解定理.

定理1.47. 设R是主理想整环, $a \in R$非零非单位, 则a可写为R中有限个不可约元的乘积

$$a = p_1 p_2 \cdots p_s.$$

而且在不计因子的次序下, 这种分解唯一. 即如果还有不可约分解$a = q_1 q_2 \cdots q_t$, 则$s = t$, 且必要时重新下标, 有$p_i \sim q_i$, $i = 1, 2, \ldots, s$.

证明. 先证明分解的存在性. 如果a是不可约元, 则$a = a$为所求分解; 否则$a = a_1 a_2$, 其中a_1, a_2都不是单位. 如果a_1, a_2都是不可约元, 则得到不可约分解; 否则可以继续写为非单位元的乘积. 于是继续这一过程, 我们得到如下形式的图: 该图中每个顶点都是R中的非单位元, 每个顶点是下方和它相连的两个顶点(如果有的话)的乘积, 位于每个

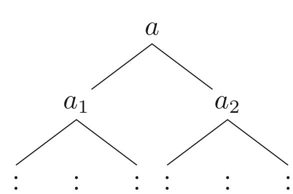


图 1.2: a的分解树

枝条(从a出发向下的链)最下端的顶点(叶子)都是不可约元. 如果这个图中有无穷长的枝条, 设为$a, b_1, b_2, \ldots$, 则得到理想的严格升链

$$\langle a\rangle \subsetneqq \langle b_1\rangle \subsetneqq \langle b_2\rangle \subsetneqq \cdots,$$

这矛盾于R是Noether环. 所以图1.2中的每个枝条都是有限的, 即上面的过程有限步后停止. 假设所有的叶子为$p_1, p_2, \ldots, p_s$, 则

$$a = p_1p_2\cdots p_s$$

为不可约元的乘积.

再证分解的唯一性. 设有不可约分解$a = p_1p_2\cdots p_s = q_1q_2\cdots q_t$. 则$p_1 \mid q_1q_2\cdots q_t$. 而$p_1$为素元, 所以存在$q_j$, 使得$p_1 \mid q_j$. 重新下标, 可设$j = 1$, 即$p_1 \mid q_1$. 于是存在$x \in R$, 使得$q_1 = xp_1$. 但是$q_1$不可约, p_1不是单位, 所以$x \in R^{\times}$, 即$p_1 \sim q_1$, 且有

$$p_2\cdots p_s = xq_1\cdots q_t.$$

于是对s归纳可以得到唯一性的结论. □

将上面的定理用到主理想整环$\mathbb{Z}$和$\mathbb{F}[x]$, 就可以分别得到算术基本定理和一元多项式环的因式分解存在唯一定理. 注意整数环$\mathbb{Z}$中的不可约元就是$\pm p$, 其中p是素数, 而一元多项式环$\mathbb{F}[x]$中的不可约元就是不可约多项式. 读者从这里应该可以体会到抽象化的威力.

§ 1.20 补充题

A1. 求包含$\sqrt[3]{2}$的最小的数域.

A2. 设$f(x), g(x) \in \mathbb{F}[x]$, $(f(x), g(x)) = d(x)$, 证明: 对任意正整数n, 成立

$$(f^n(x), f^{n-1}(x)g(x), \ldots, f(x)g^{n-1}(x), g^n(x)) = d^n(x).$$

A3. 设多项式$f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$全不为零, $m(x)$是$f_1(x), f_2(x), \ldots, f_s(x)$的公倍式, 证明: $m(x)$是$f_1(x), f_2(x), \ldots, f_s(x)$的最小公倍式的充分和必要条件是,

$$\left(\frac{m(x)}{f_1(x)}, \frac{m(x)}{f_2(x)}, \ldots, \frac{m(x)}{f_s(x)}\right) = 1.$$

A4. 证明: $y = \sin x$在实数域内不能表示为x的多项式.

A5. 设$\mathbb{F}$是数域, $f(x) \in \mathbb{F}[x]$, $0 \neq a \in \mathbb{F}$. 证明: 如果$f(x + a) = f(x)$, 则$f(x)$为常数多项式.

A6. 设$f(x) \in \mathbb{F}[x]$, 对任意$a, b \in \mathbb{F}$, 有$f(a + b) = f(a) + f(b)$, 证明: 存在$\lambda \in \mathbb{F}$, 使得$f(x) = \lambda x$.

A7. 设n为正整数, $f(x) = x^{2n+1} - (2n + 1)x^{n+1} + (2n + 1)x^n - 1$, 证明: 1是$f(x)$的三重根.

A8. 设多项式$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, k是正整数, 证明: $(x - 1)^{k+1} \mid f(x)$的充分必要条件是,

$$\begin{cases} a_0 + a_1 + a_2 + \cdots + a_n = 0, \\ a_1 + 2a_2 + \cdots + na_n = 0, \\ a_1 + 2^2a_2 + \cdots + n^2a_n = 0, \\ \qquad\vdots \\ a_1 + 2^ka_2 + \cdots + n^ka_n = 0. \end{cases}$$

A9. 设$f_1(x), f_2(x), \ldots, f_n(x) \in \mathbb{F}[x]$, 记

$$g(x) = f_1(x^{n+1}) + xf_2(x^{n+1}) + \cdots + x^{n-1}f_n(x^{n+1}).$$

证明:

$$1 + x + \cdots + x^n \mid g(x) \Longleftrightarrow x - 1 \mid f_i(x), i = 1, 2, \ldots, n.$$

(提示: 可以学过行列式后再来做这个习题.)

A10. 设m, n是正整数, 证明:

$$1 + x + \cdots + x^m \mid 1 + x^n + x^{2n} + \cdots + x^{mn} \Longleftrightarrow (m+1, n) = 1.$$

A11. 设$f(x) \in \mathbb{C}[x]$的次数小于n, 其中n是正整数, 设$\zeta_n = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$, 证明:

$$f(0) = \frac{1}{n}\sum_{j=0}^{n-1} f(\zeta_n^j).$$

A12. 设$p(x), f(x) \in \mathbb{F}[x]$, $p(x)$不可约, 且$p(x)$和$f(x)$有公共复根, 证明: $p(x)$的所有复根都是$f(x)$的根.

A13. 设n是奇数, 证明: $(x+y)(y+z)(z+x) \mid (x+y+z)^n - x^n - y^n - z^n$.

A14. 设$c \neq 0$, 求三次方程$x^3 + ax^2 + bx + c = 0$的三个根的倒数的平方和.

A15. 设三次方程$x^3 + ax^2 + bx + c = 0$的三个根是x_1, x_2, x_3, 求一个三次方程, 使其三个根为

(1) x_1^2, x_2^2, x_3^2;

(2) x_1x_2, x_1x_3, x_2x_3.

B1. 设多项式$f(x)$被$x-1$, $x-2$, $x-3$除后, 余式分别为4, 8, 16, 求$f(x)$被$(x-1)(x-2)(x-3)$除后的余式.

B2. 设多项式除以x^2+1, x^2+2的余式分别为$2x+3$, $x+4$, 求$f(x)$除以$(x^2+1)(x^2+2)$的余式.

B3. 求以$\sqrt{2} + \sqrt[3]{3}$为根的次数最小的非零首一有理系数多项式$f(x)$, 并求$f(x)$的所有复根.

B4. 设$f(x) \in \mathbb{Q}[x]$, $a, b, c, d \in \mathbb{Q}$, 其中$c, d > 0$. 已知$\sqrt{c}$, $\sqrt{d}$, $\sqrt{cd}$都是无理数, 证明:

(1) 如果$a + b\sqrt{d}$是$f(x)$的根, 则$a - b\sqrt{d}$也是$f(x)$的根;

(2) 如果$a\sqrt{c} + b\sqrt{d}$是$f(x)$的根, 则$a\sqrt{c} - b\sqrt{d}$, $-a\sqrt{c} + b\sqrt{d}$, $-a\sqrt{c} - d\sqrt{d}$也都是$f(x)$的根.

B5. 设$f(x) \in \mathbb{Q}[x]$, $d \in \mathbb{Q}$且$\sqrt[3]{d}$是无理数, 证明: 如果$\sqrt[3]{d}$是$f(x)$的根, 则$\sqrt[3]{d}\omega$和$\sqrt[3]{d}\omega^2$也是$f(x)$的根, 其中$\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$.

B6. 设$f(x)$是整系数多项式, a_1, a_2, a_3, a_4是互不相同的整数, p是素数. 如果对$i = 1, 2, 3, 4$, 都有$f(a_i) = p$, 证明: $f(x)$没有整数根.

B7. 证明: 实系数多项式$f(x) = x^3 + ax^2 + bx + c$的根的实部全是负数的充分必要条件是

$$a > 0,\ c > 0,\ ab > c.$$

B8. 设$A \in \mathbb{C}^{4\times 4}$满足: $\mathrm{Tr}(A^i) = i$, $i = 1, 2, 3, 4$. 求行列式$\det(A)$.

B9. 求解方程组

$$\begin{cases} x_1 + x_2 + \cdots + x_n = n, \\ x_1^2 + x_2^2 + \cdots + x_n^2 = n, \\ \qquad\qquad \vdots \\ x_1^n + x_2^n + \cdots + x_n^n = n. \end{cases}$$

B10. 求多项式

$$\begin{aligned} f(x) = & x^n + (a+b)x^{n-1} + (a^2+ab+b^2)x^{n-2} + \cdots \\ & + (a^{n-1} + a^{n-2}b + \cdots + a^{n-2}b + b^{n-1})x + (a^n + a^{n-1}b + \cdots + a^{n-1}b + b^n) \end{aligned}$$

的根的等幂和$s_1, s_2, \ldots, s_n$.

第二章　线性方程组与矩阵

现在我们正式开始线性代数之旅. 本章从大家熟悉的线性方程组出发, 从理论的角度阐述用Gauss消元法解线性方程组的过程; 并自然引入矩阵, 用矩阵的初等变换来实现Gauss消元法; 以求解线性方程组这一具体例子来阐明线性代数的代数工具矩阵及其初等变换的重要性.

本章中$\mathbb{F}$表示任意一个数域.

§ 2.1　线性方程组的初等变换

线性方程组是一次方程(或者: 次数不超过1的方程)所组成的方程组. 例如, 在前言的"鸡兔同笼"例子中, 我们得到的方程组

$$\begin{cases} x+y=35, \\ 2x+4y=94 \end{cases}$$

就是线性方程组. 这里x和y是未知数, 每个方程都是一次方程. 我们称这是一个二元线性方程组, 因为它有两个未知数. 我们在中学学过它们的解法, 例如, 由第一个方程得到

$$x=35-y,$$

代入第二个方程得到

$$2(35-y)+4y=94,$$

即$2y=24$. 求得$y=12$, 再得到

$$x=35-12=23.$$

容易看到, 上面的解法可以求解任意(有两个方程的)二元线性方程组. 我们当然希望可以类似地求解有三个未知数的线性方程组, 有四个未知数的线性方程组, 一般地, 有若干个未知数的线性方程组.

一般的, 取定$n\in\mathbb{N}$. 称有n个未知数的线性方程组为n元线性方程组. 通常用$x_1, x_2, \ldots, x_n$来表示这n个未知数, 于是一个有m个方程的n元线性方程组可以表示为

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases}$$

其中$a_{ij}, b_i \in \mathbb{F}$是常数. 我们称$a_{ij}$是系数, b_i是常数项. 注意系数a_{ij}的双下标的意义: 第一个下标i对应第i个方程, 第二个下标j对应第j个未知数x_j. 我们还要强调的是, 线性方程组解的情况与未知数所用的符号无关, 而仅取决于系数与常数项.

如果一个线性方程组有解, 我们就称它是相容的线性方程组. 显然, 线性方程组的基本问题是如何求出解, 我们要讲述的就是线性方程组的求解故事. 可以将线性方程组的求解具体细化为下面的几个小问题:

(a) 线性方程组何时相容?

(b) 有解时, 有几个解?

(c) 给出具体的求解算法.

对于问题(a), 我们要找到一个有解规则, 每个线性方程组不必求解, 只要去对照这个规则, 满足规则则有解. 对于问题(b), 我们希望有个解数的"公式", 给定一个线性方程组, 不必求解, 代入这个公式就可以求出解数. 对于问题(c), 我们希望有一个计算机看得懂的求解算法, 给定一个线性方程组, 输入它的系数和常数项后, 按照算法可以很快地得到解. 咋一看, 这些问题我们可能不知道如何下手. 给定一个二元甚至三元线性方程组, 我们知道如何具体地求出解. 而现在我们考察的是n元线性方程组, 这里n是任意的正整数. 下面我们试图解决这些问题, 其实, 在中学数学中, 我们已经学过回答这些问题的工具和想法.

2.1.1 Gauss消元法

解决问题的一个常用想法是先看简单的情形. 这里最简单的情形是: $m = n = 1$, 此时方程组就是一个一元线性方程: $ax = b$. 它的解是非常明确的:

(i) 当$a \neq 0$时有唯一解$x = a^{-1}b$;

(ii) 当$a = 0$且$b = 0$时, $\mathbb{F}$中任意数都是解;

(iii) 当$a = 0$而$b \neq 0$时, 无解.

再看$m = n = 2$的情形, 此时方程组为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

从几何上看, 如果a_{11}, a_{12}不同时为0, 且a_{21}, a_{22}不同时为0, 则解这个方程组对应到平面上求直线$a_{11}x_1 + a_{12}x_2 = b_1$和直线$a_{21}x_1 + a_{22}x_2 = b_2$的交点问题. 于是我们得到

(i) 如果两直线相交, 那么有唯一解;

(ii) 如果两直线重合, 那么任意的数对都是解;

(iii) 如果两直线平行, 那么没有解.

在这两种简单情形下, 解都是三种情况. 于是我们有理由猜想一般的线性方程组的解的个数有三种可能性: 一个解, 无穷多解, 以及无解. 后面我们将证明这个猜想成立. 另一方面, 我们看到, 未知数越少越容易求解. 所以一个自然的想法是, 是否可以把一般的线性方程组转化为未知数很少(最好一个)的“简单”线性方程来求解? 从很多未知数变到很少未知数, 就是消元, 而这正是中学求解线性方程组的方法.

消元通常有两种过程: 一种是逐步消去未知数, 另一种是一次消去其余未知数. 我们称前者为**Gauss消元法**, 后者为加减消元法. 解线性方程组常用方法是Gauss消元法, 于是线性方程组的一种解法是

$$\boxed{\text{线性方程组}} \xrightarrow{\text{Gauss消元法}} \boxed{\text{简单线性方程组}}$$

下面是一个具体例子.

例2.1. *在平面直角坐标系中, 作抛物线$y = ax^2 + bx + c$, 使其经过点$(2, 10)$, $(1, 1)$, $(-3, 5)$, 求抛物线方程.*

☞ **解.** 由条件得到线性方程组

$$(\clubsuit)\begin{cases} 4a + 2b + c = 10, \\ a + b + c = 1, \\ 9a - 3b + c = 5. \end{cases}$$

下面用Gauss消元法解这个线性方程组:

$$(\clubsuit) \xrightarrow{(1)\leftrightarrow(2)} \begin{cases} a + b + c = 1, \\ 4a + 2b + c = 10, \\ 9a - 3b + c = 5 \end{cases} \xrightarrow[(3)-9\times(1)]{(2)-4\times(1)} \begin{cases} a + b + c = 1, \\ \quad -2b - 3c = 6, \\ \quad -12b - 8c = -4 \end{cases}$$

$$\xrightarrow{(3)-6\times(2)} \begin{cases} a + b + c = 1, \\ \quad -2b - 3c = 6, \\ \qquad\quad 10c = -40 \end{cases} \xrightarrow{\frac{1}{10}\times(3)} \begin{cases} a + b + c = 1, \\ \quad -2b - 3c = 6, \\ \qquad\quad c = -4 \end{cases}$$

$$\xrightarrow{(2)+3\times(3)} \begin{cases} a + b + c = 1, \\ \quad -2b \quad = -6, \\ \qquad c = -4 \end{cases} \xrightarrow{-\frac{1}{2}\times(2)} \begin{cases} a + b + c = 1, \\ \quad b \quad = 3, \\ \qquad c = -4 \end{cases}$$

$$\xrightarrow[(1)-(3)]{(1)-(2)} \begin{cases} a = 2, \\ b = 3, \\ c = -4. \end{cases}$$

这里, 我们把方程组看成一个整体. 对于每个方程组, 用(1)表示它的第一个方程, 其余类似. 记号$(1) \leftrightarrow (2)$表示交换方程组中第一个和第二个方程的顺序, 记号$(2) - 4 \times (1)$表示第二个方程的两边分别对应加上第一个方程两边的-4倍, 记号$\frac{1}{10} \times (3)$表示第三个方程的两边同时乘$\frac{1}{10}$, 其余类似.

于是所求的抛物线方程为$y = 2x^2 + 3x - 4$. □

2.1.2 初等变换

在上面的Gauss消元法中, 我们将方程组变到另一个方程组, 共用到三种变换.

定义2.1. *称下面三类线性方程组的变换为初等变换:*

(i) *交换某两个方程的位置;*

(ii) *用非零常数乘某个方程;*

(iii) *用数乘某个方程再加到另一个方程上.*

类似于上面, 我们用$(i) \leftrightarrow (j)$表示交换方程组中第(i)和第(j)两个方程的位置; 用$a \times (i)$表示用非零数a乘第(i)个方程; 用$(j) + a \times (i)$表示第(i)个方程乘数a后再加到第(j)个方程上.

一个自然的问题是: 在Gauss消元法将方程组变换的过程中, 方程组的解是否会改变? 为此定义一个自然的概念.

定义2.2. *设(A)和(B)是两个n元线性方程组.*

(1) *如果(A)和(B)有完全相同的解, 则称(A)和(B)同解.*

(2) *设变换$\mathcal{T}$将(A)变成(B), 如果(A)和(B)同解, 则称$\mathcal{T}$是一个同解变换.*

容易证明线性方程组的三个初等变换都是同解变换.

引理2.1. *线性方程组的初等变换是同解变换.*

所以Gauss消元法解线性方程组的过程可以描述如下:

$$\boxed{\text{线性方程组}} \xrightarrow{\text{初等变换}} \boxed{\text{简单线性方程组}}$$

在上面的求解过程中, 线性方程组中未知数所用符号并不是本质的, 本质的是系数和常数项. 为了抓住本质的东西, 在下节中, 我们引入矩阵的概念, 用矩阵来记录线性方程组的系数和常数项.

习题2.1

A1. 用Gauss消元法解线性方程组$\begin{cases} x+y+z=6, \\ 2x+y-3z=-5, \\ 3x-y+z=4. \end{cases}$

A2. 证明: 线性方程组的初等变换是同解变换.

§ 2.2 矩阵和线性方程组

将线性方程组$\begin{cases} 4a+2b+c=10, \\ a+b+c=1, \\ 9a-3b+c=5 \end{cases}$的未知数和等号去掉, 我们得到下面的数表

$$\begin{matrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 1 \\ 9 & -3 & 1 & 5 \end{matrix}$$

注意为了能够恢复原来的线性方程组, 我们将这些数排得比较整齐: 每行对应一个方程, 前面三列是系数, 最后一列是常数项, 第一列是第一个未知数的系数, 等等. 这样整齐的一个数表就是矩阵. 为了告诉别人这个数表是一个整体, 我们用圆括号[1]把它括起来, 得到

$$\begin{pmatrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 1 \\ 9 & -3 & 1 & 5 \end{pmatrix}.$$

一般地, 矩形(长方形)的数表称为矩阵. 于是矩阵的特点是外部框形, 内部成行列. 一个m行n列的$\mathbb{F}$系数矩阵($m\times n$矩阵)可以表示为

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

其中$a_{ij}\in\mathbb{F}$, 称其为(i,j)位置的元素, i表示所在的行, 而j表示所在的列. 常用A,B,C等大写的英文字母表示矩阵, 上面的矩阵可以表示为

$$A=A_{m\times n}=(a_{ij})=(a_{ij})_{m\times n}.$$

行数和列数称为这个矩阵的型号, 两个矩阵如果型号相同, 则称它们同型. 设$A=(a_{ij})$和$B=(b_{ij})$是两个矩阵, 如果A和B同型, 且对任意的i和j有$a_{ij}=b_{ij}$, 则称A和B相等, 记为$A=B$.

数域$\mathbb{F}$上所有$m\times n$矩阵全体记为$\mathbb{F}^{m\times n}$, 即[2]

$$\mathbb{F}^{m\times n}=\{(a_{ij})_{m\times n}\mid a_{ij}\in\mathbb{F},\forall i,j\}.$$

矩阵是本讲义的主角之一, 所以我们花费一些笔墨来介绍一些特殊的矩阵.

[1] 也有用方括号的.

[2] 也用记号$M_{m\times n}(\mathbb{F})$表示$\mathbb{F}^{m\times n}$, 用记号$M_n(\mathbb{F})$表示$\mathbb{F}^{n\times n}$.

首先, $\mathbb{F}$中的数可以看成是1×1阵, 所以矩阵是数的推广. 其次, 称$m\times 1$ 的矩阵为m维列向量, 其一般形式为

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \quad (a_i \in \mathbb{F});$$

称$1\times n$的矩阵为n维行向量, 例如,

$$(a_1, a_2, \ldots, a_n), \quad (a_i \in \mathbb{F})$$

就是一般的n维行向量. 所有m维列向量做成的集合为$\mathbb{F}^{m\times 1}$, 我们将这个简记为$\mathbb{F}^m$. 再次, 我们称$n\times n$的矩阵为n阶方阵. 下面是一个一般的n阶方阵

$$A = A_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

这个方阵有两个对角线(正方形的两个对角线), 其中一个对角线上的所有元素的行和列下标相等(即a_{ii}所在的对角线), 称其为主对角, 简称为对角; 另一对角线上元素的行标和列标之和为$n+1$, 称之为副对角. 最后, 我们称所有元素为0的矩阵为零矩阵, 记为O或者0, 即

$$0 = O = O_{m\times n} = (0)_{m\times n}$$

是$m\times n$的零矩阵. 根据定义, 零矩阵很多, 例如

$$O_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq O_{2\times 1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

下面再介绍一些特殊的方阵. 如果一个方阵的主对角下方的元素都等于零, 则称其为上三角阵, 于是一般形式是

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}, \quad (\text{空白位置为}0).$$

由定义, $A=(a_{ij})_{n\times n}$是上三角阵的充分必要条件是, 对任意$1\leqslant j<i\leqslant n$, 有$a_{ij}=0$. 类似的, 如果一个方阵的主对角上方的元素都等于零, 则称其为下三角阵. 如果一个方阵同时是上三角阵和下三角阵, 则称它是对角阵, 即对角元外的元素都是零的方阵是对角阵. 对角阵常用下面的记号

$$\mathrm{diag}(a_1, a_2, \ldots, a_n) := \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}, \quad (a_1, a_2, \ldots, a_n \in \mathbb{F}).$$

所有对角元都相等的对角阵称为纯量阵, 例如

$$\operatorname{diag}(\underbrace{a,\ldots,a}_{n}), \quad (a \in \mathbb{F})$$

是a对应的n阶纯量阵. 数1对应的纯量阵就是单位阵, 记为[3]E, 即n阶单位阵为

$$E = E_n := \operatorname{diag}(\underbrace{1,\ldots,1}_{n}) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}_{n\times n}.$$

下面回到线性方程组. 设有n元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

我们用矩阵来记录这个方程组本质的东西. 令

$$A = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{F}^{m\times n},$$

称其为系数矩阵. 记

$$\beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{F}^m,$$

它是m维列向量, 称其为常数项列向量. 最后记

$$\widetilde{A} = (A, \beta) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \in \mathbb{F}^{m\times(n+1)},$$

称其为增广矩阵.

增广矩阵记录了这个线性方程组的所有信息: 每一行对应一个方程, 最后一列对应常数项, 前面的列对应相应未知数的系数. 于是, 将每个线性方程组对应到它的增广矩阵, 我们得到双射

$$\{m\text{个方程的}n\text{元线性方程组}\} \xleftrightarrow[\text{onto}]{1-1} \mathbb{F}^{m\times(n+1)}.$$

在这个对应下, Gauss消元法解线性方程组的过程可以描述如下:

[3]也记为I.

$$\begin{tikzcd}
\boxed{\text{线性方程组}} \arrow[r, "\text{初等变换}"] \arrow[d, leftrightarrow] & \boxed{\text{简单线性方程组}} \arrow[d, leftrightarrow] \\
\boxed{\text{矩阵}} \arrow[r, "?"] & \boxed{\text{简单矩阵}}
\end{tikzcd}$$

下面我们的任务之一是将线性方程组的初等变换翻译到矩阵世界, 对应的变换我们称之为矩阵的初等(行)变换. 另一个任务是回答什么是简单矩阵, 这些都留待下一节解决.

习题2.2

A1. 写出线性方程组 $\begin{cases} 2x_2+3x_3-x_4=1, \\ -x_1+2x_3+11x_4=0, \\ 3x_1-2x_2-4x_3+x_4=-1 \end{cases}$ 的增广矩阵 $\widetilde{A}$.

A2. 写出以矩阵 $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 2 & -5 & 0 \\ 3 & -2 & 4 & -2 \end{pmatrix}$ 为增广矩阵的线性方程组.

§ 2.3 矩阵的初等变换

线性方程组的方程对应到其增广矩阵的行, 对方程进行操作, 相当于对增广矩阵的行进行相应操作. 于是, 对应到线性方程组的初等变换, 我们有下面的定义.

定义2.3. *称下面三类变换为矩阵的*行初等变换*:*

(i) *调行变换: $r_i \leftrightarrow r_j$(交换第 i 行和第 j 行);*

(ii) *行数乘变换: $a \times r_i$, $a \neq 0$(用 a 乘第 i 行所有元素);*

(iii) *行消去变换: $r_j + a \times r_i$(用 a 乘第 i 行所有元素, 再加到第 j 行的对应位置元素上).*

为了后面的应用, 我们还给出如下定义.

定义2.4. *称下面三类变换为矩阵的*列初等变换*:*

(i) *调列变换: $c_i \leftrightarrow c_j$;*

(ii) *列数乘变换: $a \times c_i$, $a \neq 0$;*

(iii) *列消去变换: $c_j + a \times c_i$.*

*行初等变换和列初等变换统称为*初等变换.

由于通常需要许多次行初等变换才能将一个矩阵化为“简单”矩阵, 所以我们引入下面概念.

定义2.5. *设 $A, B \in \mathbb{F}^{m\times n}$.*

(1) *如果 A 可经过有限次行初等变换化为 B, 则称 A 与 B* 行等价*, 记为 $A \overset{r}{\sim} B$;*

(2) 如果A可经过有限次列初等变换化为B, 则称A与B列等价, 记为$A \overset{c}{\sim} B$;

(3) 如果A可经过有限次初等变换化为B, 则称A与B等价[4], 记为$A \sim B$.

线性方程组的初等变换为同解变换对应到下面的对称性.

引理2.2. 矩阵的(行, 列)等价满足[5](以等价为例):

(i) (自反性) $\forall A \in \mathbb{F}^{m\times n}$, 有$A \sim A$;

(ii) (对称性) $\forall A, B \in \mathbb{F}^{m\times n}$, 若$A \sim B$, 则$B \sim A$;

(iii) (传递性) $\forall A, B, C \in \mathbb{F}^{m\times n}$, 若$A \sim B$, $B \sim C$, 则$A \sim C$.

☞ **证明.** 我们以行等价为例. (i)和(iii)显然成立. 注意到

$$A \overset{r_i \leftrightarrow r_j}{\rightsquigarrow} B \Longrightarrow B \overset{r_i \leftrightarrow r_j}{\rightsquigarrow} A,$$

$$A \overset{a\times r_i}{\rightsquigarrow} B \Longrightarrow B \overset{\frac{1}{a}\times r_i}{\rightsquigarrow} A, \quad (a \neq 0),$$

$$A \overset{r_j + a r_i}{\rightsquigarrow} B \Longrightarrow B \overset{r_j - a r_i}{\rightsquigarrow} A,$$

则可知(ii)成立. □

下面我们回到矩阵化简. 我们已经知道了化简的工具: (行)初等变换, 那么什么样的矩阵是“简单”的呢? 目前看来, 简单是为了可以直接写出对应的线性方程组的解. 还是看个例子.

例2.2. 例2.1中用$Gauss$消元法解线性方程组(♣)的过程, 用对应的增广矩阵表示为:

$$\widetilde{A} = \begin{pmatrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 1 \\ 9 & -3 & 1 & 5 \end{pmatrix} \overset{r_1 \leftrightarrow r_2}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 10 \\ 9 & -3 & 1 & 5 \end{pmatrix}$$

$$\underset{r_3 - 9r_1}{\overset{r_2 - 4r_1}{\rightsquigarrow}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -3 & 6 \\ 0 & -12 & -8 & -4 \end{pmatrix} \overset{r_3 - 6r_2}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -3 & 6 \\ 0 & 0 & 10 & -40 \end{pmatrix}$$

$$\overset{\frac{1}{10}r_3}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -3 & 6 \\ 0 & 0 & 1 & -4 \end{pmatrix} \overset{r_2 + 3r_3}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -6 \\ 0 & 0 & 1 & -4 \end{pmatrix}$$

$$\overset{-\frac{1}{2}r_2}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \end{pmatrix} \underset{r_1 - r_3}{\overset{r_1 - r_2}{\rightsquigarrow}} \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \end{pmatrix}.$$

观察上面例子中用行初等变换化简矩阵的过程, 我们主要做了两件事: 一个是利用行消去变换不断的产生零[6], 以至于最后得到的矩阵除最后一列外, 每一行至多只有一个数非零, 这些非零的数位

[4]行等价也称为行相抵, 列等价也称为列相抵, 等价也称为相抵.
[5]也就是说矩阵的(行, 列)等价是$\mathbb{F}^{m\times n}$的等价关系.
[6]用初等变换化简矩阵也称为矩阵打洞.

于不同的列[7], 而且从上往下看, 这些非零数不断右移[8]; 另一个是把这些非零的数变成[9]1. 这些非零的数我们称之为首元, 即矩阵中非零行的从左往右看的第一个非零元是首元. 称首元所在的列为首元列.

例2.3. 下面两个矩阵中加灰色的数都是首元, 首元列都是第1, 2, 3列.

$$\begin{pmatrix} 0 & \colorbox{gray}{3} & -1 & 0 & 2 \\ \colorbox{gray}{-2} & 0 & 3 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \colorbox{gray}{-1} & 5 & 1 \end{pmatrix}, \quad \begin{pmatrix} \colorbox{gray}{1} & 0 & 0 & 2 \\ 0 & \colorbox{gray}{1} & 0 & 3 \\ 0 & 0 & \colorbox{gray}{1} & -4 \end{pmatrix}.$$

结合上面的讨论, 我们给出关于何为简单矩阵的定义.

定义2.6. **(1)** 称一个矩阵是阶梯形阵, 如果它满足

(i) 全零行下方无非零行,

(ii) 各非零行的首元在上一行首元的右边;

(2) 称一个矩阵是简化阶梯形阵(或者行等价标准形阵[10]), 如果它满足

(i) 是阶梯形阵,

(ii) 首元都是1,

(iii) 各首元列只有一个非零元.

请读者仔细体会为什么我们觉得如此定义的(简化)阶梯形阵为简单矩阵. 显然, 在阶梯形阵中

$$\text{非零行数} = \text{首元数} = \text{首元列数}.$$

且由定义, 在阶梯形阵中可以画一个台阶, 其一般形式是

$$\begin{pmatrix} 0 & \cdots & 0 & \Delta & & & & & & \\ 0 & \cdots & 0 & & \Delta & & & * & & \\ 0 & \cdots & 0 & & & \Delta & & & & \\ \vdots & & \vdots & & & & \ddots & & & \\ 0 & \cdots & 0 & & & & & \Delta & & \\ 0 & \cdots & 0 & & & & & & \Delta & \\ \vdots & & \vdots & & & & & & & \\ 0 & \cdots & 0 & & & & & & & \end{pmatrix}, \quad \Delta \neq 0: \text{首元},$$

[7]如此就把未知数分离开了.
[8]如此未知数越来越少, 实现了消元.
[9]如此就可以直接写出解.
[10]行等价标准形也称为行相抵标准形.

而简化阶梯形阵的一般形式是

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & 0 & * & 0 & & & 0 & & 0 & \\ 0 & \cdots & 0 & & & 1 & * & 0 & & & 0 & & 0 & \\ 0 & \cdots & 0 & & & & & 1 & & & 0 & * & 0 & * \\ \vdots & & \vdots & & & & & & & \ddots & \vdots & & \vdots & \\ 0 & \cdots & 0 & & & & & & & & 1 & & 0 & \\ 0 & \cdots & 0 & & & & & & & & & & 1 & \\ \vdots & & \vdots & & & & & & & & & & & \\ 0 & \cdots & 0 & & & & & & & & & & & \end{pmatrix}.$$

例2.4. **(1)** $\begin{pmatrix} 0 & \boxed{3} & -1 & 0 & 2 \\ \boxed{-2} & 0 & 3 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{-1} & 5 & 1 \end{pmatrix}$: 非阶梯形阵.

(2) $\begin{pmatrix} 0 & \boxed{-1} & 2 & 3 & 1 \\ 0 & & \boxed{3} & 4 & 2 \\ 0 & & & & \boxed{2} \\ 0 & & & & \end{pmatrix}$, $\begin{pmatrix} \boxed{1} & 1 & 1 & 0 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & \boxed{1} & -4 \end{pmatrix}$: 阶梯形阵, 但非简化阶梯形阵.

(3) $\begin{pmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & \boxed{1} & -4 \end{pmatrix}$: 简化阶梯形阵.

下面我们进入本节的主题: 用行初等变换将一个矩阵化为阶梯形阵和简化阶梯形阵.

定理2.3. 设$A \in \mathbb{F}^{m\times n}$, 则

(1) $A \overset{r}{\sim}$ 阶梯形阵;

(2) $A \overset{r}{\sim}$ 简化阶梯形阵.

☞ **证明.** 这个证明就是给出化简的算法.

(1) 算法一: $A \overset{r}{\sim}$ 阶梯形阵.

第一步 若$A = O$, 则结束; 否则, 可设

$$A = \begin{pmatrix} 0 & \cdots & 0 & a_{1j} & \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & a_{mj} & \end{pmatrix}, \quad \text{第 } j \text{ 列是非零列}.$$

于是必要时总可用调行变换, 使得A行等价于

$$A_1 = \begin{pmatrix} 0 & \cdots & 0 & \Delta & \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & * & \end{pmatrix}, \quad \Delta \neq 0.$$

再用行消去变换$r_i - \frac{*}{\Delta}r_1$, 可得$A$行等价于

$$A_2 = \begin{pmatrix} 0 & \cdots & 0 & \Delta & & * & \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & B & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}.$$

第二步 对B重复第一步(由于B的左边都是零, 所以不论对B所在的行进行何种行初等变换, B的左边还是零. 于是只要考虑B.).

由于A的列数n是有限数, 所以有限步后必有$B=0$, 即将A行等价于阶梯形阵.

(2) 算法二: $A \overset{r}{\sim}$ 简化阶梯形阵.

只需

$$A \overset{\text{算法一}}{\sim} \text{阶梯形阵} \overset{\text{行数乘}}{\sim} \text{首元为1} \overset{\text{行消去}}{\sim} \text{首元列其余元为0}$$

即可. □

与A行等价的简化阶梯形阵是唯一的(为什么?); 与A行等价的阶梯形阵不唯一, 但是非零行数是由A唯一确定的.

还要注意的是, 在算法一中, 我们只用到了调行变换和行消去变换.

例2.5. 求与矩阵$A = \begin{pmatrix} 0 & 0 & 0 & 2 & 2 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & 1 & 2 & 4 & 3 \end{pmatrix}$行等价的简化阶梯形阵.

解. 我们进行如下的行初等变换

$$A \overset{r_1 \leftrightarrow r_3}{\sim} \begin{pmatrix} 0 & \boxed{1} & 2 & 4 & 3 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix} \overset{r_2 - 2r_1}{\sim} \begin{pmatrix} 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

$$\underset{\frac{1}{2}r_3}{\overset{-\frac{1}{5}r_2}{\sim}} \begin{pmatrix} 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \underset{r_3 - r_2}{\overset{r_1 - 4r_2}{\sim}} \begin{pmatrix} 0 & \boxed{1} & 2 & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

最后一个矩阵就是所求的简化阶梯形阵. □

习题2.3

A1. 用行初等变换, 将下面矩阵化为简化阶梯形阵.

(1) $\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & 1 & 3 \\ 5 & -2 & 3 & 2 \end{pmatrix}$; (2) $\begin{pmatrix} 0 & 2 & 1 & 1 & -1 \\ 0 & 3 & -1 & 4 & 0 \\ 0 & 4 & 2 & 2 & 1 \\ 0 & -2 & 3 & -5 & 1 \end{pmatrix}$.

A2. 设$A, B \in \mathbb{F}^{m\times n}$, A是阶梯形矩阵, B是简化阶梯形矩阵. 证明:

(1) 如果$m > n$, 那么A一定有全零行;

(2) 如果$m = n$, 那么A一定是上三角阵;

(3) 如果$m = n$且B无全零行, 那么一定有$B = E_n$.

§ 2.4 线性方程组的求解

2.4.1 线性方程组的求解

现在我们可以写出用Gauss消元法解线性方程组的过程:

$$
\begin{tikzcd}
\boxed{\text{线性方程组}} \arrow[r, "\text{初等变换}"] \arrow[d] & \boxed{\text{简单线性方程组}} \\
\boxed{\text{增广矩阵}} \arrow[r, "\text{行初等变换}"] & \boxed{\text{简化阶梯形阵}} \arrow[u]
\end{tikzcd}
$$

即先写出线性方程组的增广矩阵$\widetilde{A}$, 然后用行初等变换将$\widetilde{A}$化为简化阶梯形阵$\widetilde{R}$, 最后从$\widetilde{R}$对应的同解方程组写出原方程组的解.

我们看一些例子.

例2.6. 解线性方程组$\begin{cases} 2x_1 - x_2 + 3x_3 = 1, \\ 4x_1 - 2x_2 + 5x_3 = 4, \\ 2x_1 - x_2 + 4x_3 = 0. \end{cases}$

☞ **解.** 对增广矩阵进行行初等变换

$$\widetilde{A} = \begin{pmatrix} 2 & -1 & 3 & 1 \\ 4 & -2 & 5 & 4 \\ 2 & -1 & 4 & 0 \end{pmatrix} \xrightarrow[r_3-r_1]{r_2-2r_1} \begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & -1 \end{pmatrix} \xrightarrow{r_3+r_2} \begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

得到同解方程组

$$\begin{cases} 2x_1 - x_2 + 3x_3 = 1, \\ -x_3 = 2, \\ 0 = 1. \end{cases}$$

最后一个方程是矛盾方程, 于是原方程组无解. □

例2.7. 解线性方程组$\begin{cases} 2x_1 - x_2 + 3x_3 = 1, \\ 4x_1 + 2x_2 + 5x_3 = 4, \\ 2x_1 + 2x_3 = 6. \end{cases}$

☞ **解.** 对增广矩阵进行行初等变换

$$\widetilde{A} = \begin{pmatrix} 2 & -1 & 3 & 1 \\ 4 & 2 & 5 & 4 \\ 2 & 0 & 2 & 6 \end{pmatrix} \xrightarrow[r_3-r_1]{r_2-2r_1} \begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 4 & -1 & 2 \\ 0 & 1 & -1 & 5 \end{pmatrix} \xrightarrow{r_2 \leftrightarrow r_3} \begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 1 & -1 & 5 \\ 0 & 4 & -1 & 2 \end{pmatrix}$$

$$\xrightarrow[r_3-4r_2]{r_1+r_2} \begin{pmatrix} 2 & 0 & 2 & 6 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 3 & -18 \end{pmatrix} \xrightarrow[\frac{1}{3}r_3]{\frac{1}{2}r_1} \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 1 & -6 \end{pmatrix} \xrightarrow[r_2+r_3]{r_1-r_3} \begin{pmatrix} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -6 \end{pmatrix},$$

得到同解方程组

$$\begin{cases} x_1 = 9, \\ x_2 = -1, \\ x_3 = -6. \end{cases}$$

这就是原方程组的解. □

例2.8. 解线性方程组 $\begin{cases} -x_2 - 2x_3 + 11x_4 - 3x_5 = 15, \\ -2x_1 + x_2 + 2x_3 - 4x_4 + 3x_5 = -3, \\ 6x_1 - x_2 - 2x_3 - 10x_4 - 3x_5 = -21, \\ -5x_1 + 2x_2 + 4x_3 - 5x_4 + 6x_5 = -1. \end{cases}$

☞ **解.** 对增广矩阵进行行初等变换

$$\widetilde{A} = \begin{pmatrix} 0 & -1 & -2 & 11 & -3 & 15 \\ -2 & 1 & 2 & -4 & 3 & -3 \\ 6 & -1 & -2 & -10 & -3 & -21 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix} \xrightarrow[r_3+3r_2]{r_4+r_3} \begin{pmatrix} 0 & -1 & -2 & 11 & -3 & 15 \\ -2 & 1 & 2 & -4 & 3 & -3 \\ 0 & 2 & 4 & -22 & 6 & -30 \\ 1 & 1 & 2 & -15 & 3 & -22 \end{pmatrix}$$

$$\xrightarrow[\substack{r_1 \leftrightarrow r_4 \\ \frac{1}{2}r_3}]{r_2+2r_4} \begin{pmatrix} 1 & 1 & 2 & -15 & 3 & -22 \\ 0 & 3 & 6 & -34 & 9 & -47 \\ 0 & 1 & 2 & -11 & 3 & -15 \\ 0 & -1 & -2 & 11 & -3 & 15 \end{pmatrix} \xrightarrow[\substack{r_2-2r_3 \\ r_4+r_3}]{r_1-r_3} \begin{pmatrix} 1 & 0 & 0 & -4 & 0 & -7 \\ 0 & 0 & 0 & -1 & 0 & -2 \\ 0 & 1 & 2 & -11 & 3 & -15 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[r_2 \leftrightarrow r_3]{-r_2} \begin{pmatrix} 1 & 0 & 0 & -4 & 0 & -7 \\ 0 & 1 & 2 & -11 & 3 & -15 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[r_2+11r_3]{r_1+4r_3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 3 & 7 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

得到同解方程组

$$\begin{cases} x_1 = 1, \\ x_2 + 2x_3 + 3x_5 = 7, \\ x_4 = 2, \end{cases} \quad \text{即} \quad \begin{cases} x_1 = 1, \\ x_2 = -2x_3 - 3x_5 + 7, \\ x_4 = 2. \end{cases}$$

令$x_3 = c_1$, $x_5 = c_2$, 其中c_1, c_2是任意常数, 则有

$$\begin{cases} x_1 = 1, \\ x_2 = -2c_1 - 3c_2 + 7, \\ x_3 = c_1, \\ x_4 = 2, \\ x_5 = c_2. \end{cases}$$

称这个是原方程组的通解. □

上面例子中得到的同解方程组有5个未知数, 有3个方程, 这三个方程都是“真正”的方程(即没有一个方程是多余的), 于是有$5-3=2$个自由度. 这里我们把x_3和x_5看成自由的, 令它们为独立的自由

常数. 注意到这两个未知数对应到非首元列. 我们称非首元列对应的未知数为自由未知数, 首元列对应的未知数为非自由未知数.

于是, Gauss消元法解线性方程组的过程就是一个去伪存真的过程, 我们通过化简将方程组中的多余方程去掉, 最后剩下就是真正的方程. 而要求简化阶梯形一方面是为了每一非零行对应的方程只有一个非自由未知数(即非自由未知数可分离), 另一方面将自由未知数移项后就可以表示非自由未知数. 下面是n元线性方程组的求解步骤:

算法2.1. *n元线性方程组的求解步骤:*

(1) 写出增广矩阵$\widetilde{A}$;

(2) 将$\widetilde{A}$行等价于阶梯形阵$\widetilde{A_1}$. 如果$\widetilde{A_1}$的最后一列是首元列, 则无解; 否则将$\widetilde{A_1}$行等价于简化阶梯形阵$\widetilde{R}$;

(3) 设$\widetilde{R}$有r个首元,

- **(i)** 当$r = n$时, 直接写出唯一解;
- **(ii)** 当$r < n$时, 有$n - r$个自由度,
  - **(a)** 将$n - r$个自由未知数取为独立常数$c_1, c_2, \ldots, c_{n-r}$;
  - **(b)** 将前r行对应方程中的自由未知数移项后, r个非自由未知数可用c_1, c_2, $\ldots$, c_{n-r}表示.

由上面的求解步骤, 可得线性方程组的解情况判定定理[11].

定理2.4. 设n元线性方程组的增广矩阵为$\widetilde{A}$, 而$\widetilde{R}$是与$\widetilde{A}$行等价的(简化)阶梯形阵, 则

(1) 方程组有解$\Longleftrightarrow$ $\widetilde{R}$的最后一列不是首元列;

(2) 当方程组有解时,

- **(i)** 有唯一解$\Longleftrightarrow$ $\widetilde{R}$恰有n个首元(即: $\widetilde{R}$的前n列都是首元列);
- **(ii)** 有无穷多解$\Longleftrightarrow$ $\widetilde{R}$的首元数$< n$.

于是, 线性方程组解的情况只有三种可能: 无解, 唯一解, 无穷多解. 即我们在前面通过简单例子得到的线性方程组解情况的猜想是正确的.

例2.9. 当a为何值时, 线性方程组

$$\begin{cases} 3x_1 + x_2 - x_3 = 2, \\ x_1 - 5x_2 + 2x_3 = -1, \\ 2x_1 + 6x_2 - 3x_3 = a \end{cases}$$

有解? 当有解时, 求出它的通解.

[11] 严格来说, 对这个判定定理我们并不是特别满意, 因为它基本等同于解方程组.

☞ **解.** 对增广矩阵进行行初等变换

$$\widetilde{A}=\begin{pmatrix}3&1&-1&2\\1&-5&2&-1\\2&6&-3&a\end{pmatrix}\underset{\substack{r_2-3r_1\\r_3-2r_1}}{\overset{r_1\leftrightarrow r_2}{\sim}}\begin{pmatrix}1&-5&2&-1\\0&16&-7&5\\0&16&-7&a+2\end{pmatrix}\overset{r_3-r_2}{\sim}\begin{pmatrix}1&-5&2&-1\\0&16&-7&5\\0&0&0&a-3\end{pmatrix},$$

可得当且仅当$a=3$时原方程组有解. 此时, 有

$$\widetilde{A}\overset{r}{\sim}\begin{pmatrix}1&-5&2&-1\\0&1&-\frac{7}{16}&\frac{5}{16}\\0&0&0&0\end{pmatrix}\overset{r_1+5r_2}{\sim}\begin{pmatrix}1&0&-\frac{3}{16}&\frac{9}{16}\\0&1&-\frac{7}{16}&\frac{5}{16}\\0&0&0&0\end{pmatrix}.$$

于是, 原方程组的通解是

$$\begin{cases}x_1=\frac{3}{16}c+\frac{9}{16},\\x_2=\frac{7}{16}c+\frac{5}{16},\\x_3=c,\end{cases}$$

其中, c是任意常数. □

2.4.2 齐次线性方程组的求解

有一类线性方程组一定有解. 如果线性方程组中所有的常数项都为零, 则称之为齐次线性方程组. 有非零的常数项的线性方程组则称之为非齐次线性方程组. 一个n元齐次线性方程组有形式

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\\qquad\qquad\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0.\end{cases}$$

它一定有一个解: $x_1=0$, $x_2=0$, $\ldots$, $x_n=0$, 称这个解为零解; 其余的解称为非零解.

齐次线性方程组的增广矩阵的最后一列是全零列, 不论进行何种行初等变换后仍是全零列, 所以我们只需考虑系数矩阵即可. 将上面讨论特殊化到齐次线性方程组, 我们得到:

算法2.2. *n元齐次线性方程组的求解步骤:*

(1) *写出系数矩阵A;*

(2) *将A行等价于简化阶梯形阵R, 设R有r个首元.*

- **(i)** *当$r=n$时, 只有零解;*
- **(ii)** *当$r<n$时, 有$n-r$个自由度,*
  - **(a)** *将$n-r$个自由未知数取为独立常数$c_1, c_2, \ldots, c_{n-r}$;*
  - **(b)** *将前r行对应方程中的自由未知数移项后, r个非自由未知数可用c_1, c_2, $\ldots$, c_{n-r}表示.*

还可以得到定理2.4的如下推论.

推论2.5. 设n元齐次线性方程组的系数矩阵是A, 而R是与A行等价的(简化)阶梯形阵, 则

(1) 只有零解$\Longleftrightarrow$ R有n个首元(即: R的所有列为首元列);

(2) 有非零解$\Longleftrightarrow$ R的首元数$< n$.

特别的, 当

$$\text{“方程个数”} < \text{“未知数个数”} \quad (\text{即: 当 } A \in \mathbb{F}^{m\times n} \text{ 时, 有 } m < n)$$

时, 方程组一定有非零解.

下面是一个求解齐次线性方程组的例子.

例2.10. 解线性方程组$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 0, \\ 2x_1 + x_2 - 2x_3 - 2x_4 = 0, \\ x_1 - x_2 - 4x_3 - 3x_4 = 0. \end{cases}$

☞ **解.** 对系数矩阵进行行初等变换

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & -2 & -2 \\ 1 & -1 & -4 & -3 \end{pmatrix} \underset{r_3 - r_1}{\overset{r_2 - 2r_1}{\sim}} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & -3 & -6 & -4 \end{pmatrix}$$

$$\underset{-\frac{1}{3}r_2}{\overset{r_3 - r_2}{\sim}} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 2 & \frac{4}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{r_1 - 2r_2}{\sim} \begin{pmatrix} 1 & 0 & -2 & -\frac{5}{3} \\ 0 & 1 & 2 & \frac{4}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

得到同解方程组

$$\begin{cases} x_1 - 2x_3 - \frac{5}{3}x_4 = 0, \\ x_2 + 2x_3 + \frac{4}{3}x_4 = 0, \end{cases} \qquad \text{即} \qquad \begin{cases} x_1 = 2x_3 + \frac{5}{3}x_4, \\ x_2 = -2x_3 - \frac{4}{3}x_4. \end{cases}$$

令$x_3 = c_1$, $x_4 = c_2$, 则有

$$\begin{cases} x_1 = 2c_1 + \frac{5}{3}c_2, \\ x_2 = -2c_1 - \frac{4}{3}c_2, \\ x_3 = c_1, \\ x_4 = 2, \end{cases} \qquad c_1, c_2: \text{任意常数.}$$

这是原方程组的通解. □

至此, 我们用Gauss消元法给出了本章一开始提出的关于线性方程组的求解的几个问题的一种回答. 对于问题(a)和问题(b), 我们并不是特别满意这个回答, 希望后面可以更好地回答这两个问题. 消元法除了Gauss消元法外, 还有加减消元法, 后面我们将用加减消元法来求解一类特殊的线性方程组.

习题2.4

A1. 解下面的线性方程组.

(1) $\begin{cases} x_1-2x_2+3x_3-4x_4=4, \\ x_2-x_3+x_4=-3, \\ x_1+3x_2+x_4=1, \\ -7x_2+3x_3+x_4=-3; \end{cases}$ (2) $\begin{cases} 2x_1+x_2-x_3+x_4=1, \\ 3x_1-2x_2+2x_3-3x_4=2, \\ 5x_1+x_2-x_3+2x_4=-1, \\ 2x_1-x_2+x_3-3x_4=4; \end{cases}$

(3) $\begin{cases} x_1+x_2+2x_3+2x_4=1, \\ 2x_2+x_3+5x_4=-1, \\ 2x_1+3x_3-x_4=3, \\ x_1+x_2+4x_4=-1; \end{cases}$ (4) $\begin{cases} x_1+x_2+2x_3-x_4=0, \\ x_1+x_2+x_3+x_4=0, \\ 2x_1+2x_2-x_3+2x_4=0. \end{cases}$

A2. 是否存在抛物线$y=ax^2+bx+c$, 使其经过点$(1,2)$, $(-1,3)$和$(-2,5)$? 如果没有, 说明理由; 如果有, 求出所有满足条件的抛物线.

A3. 某工厂在一次投料的生产过程中能同时获得4种产品, 但是对每种产品的单位成本难以确定, 于是通过几次测试来求解. 现通过4次测试所得的总成本如下表:

批次	产品/kg				总成本/千元
	A	B	C	D	
第1批生产	200	100	100	50	2900
第2批生产	500	250	200	100	7050
第3批生产	100	40	40	20	1360
第4批生产	400	180	160	60	5500

试求每种产品的单位成本(即每千克的成本).

A4. 当a为何值时, 线性方程组$\begin{cases} x_1-4x_2+2x_3=-1, \\ -x_1+11x_2-x_3=3, \\ 3x_1-5x_2+7x_3=a \end{cases}$ 有解? 当有解时, 求出它的通解.

A5. 讨论线性方程组$\begin{cases} x_1+x_2+x_3=3, \\ x_1+2x_2-ax_3=9, \\ 2x_1-x_2+3x_3=6 \end{cases}$ 解的情况.

B1. 解线性方程组

(1) $\begin{cases} (1+a_1)x_1+x_2+x_3+\cdots+x_n=b_1, \\ x_1+(1+a_2)x_2+x_3+\cdots+x_n=b_2, \\ \qquad\vdots \\ x_1+x_2+x_3+\cdots+(1+a_n)x_n=b_n, \end{cases}$ $(a_1,a_2,\ldots,a_n\neq 0,\ \frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_n}\neq -1)$;

(2) $\begin{cases} x_1+2x_2+3x_3+\cdots+(n-1)x_{n-1}+nx_n=b_1, \\ nx_1+x_2+2x_3+\cdots+(n-2)x_{n-1}+(n-1)x_n=b_2, \\ \qquad\qquad\vdots \\ 2x_1+3x_2+4x_3+\cdots+nx_{n-1}+x_n=b_n. \end{cases}$

B2. 当a取何值时, 线性方程组$\begin{cases} ax_1+x_2+x_3=1, \\ x_1+ax_2+x_3=a, \\ x_1+x_2+ax_3=a^2 \end{cases}$ 无解, 有唯一解和无穷多解? 并在有无穷多解时求出其通解.

B3. 某食品厂收到了2000kg食品的订单, 要求这种食品含脂肪5%, 碳水化合物12%, 蛋白质15%. 该厂准备用5种原料配制这种食品, 其中每一种原料含脂肪, 碳水化合物, 蛋白质的百分比和每千克的成本(元)如下表所示:

	A_1	A_2	A_3	A_4	A_5
脂肪	8	6	3	2	4
碳水化合物	5	25	10	15	5
蛋白质	15	5	20	10	10
每千克成本	4.4	2	2.4	2.8	3.2

(1) 用上述5种原料能不能配制出2000kg这种食品? 如果能, 配料方法唯一吗? 写出所有可能的配料方法;

(2) 对于(1)中的每种配料方法, 写出所花费成本的表达式, 并且求成本最低的配料方式(有的原料可以不用);

(3) 用A_1, A_2, A_3, A_4这4种原料能配制2000kg这种食品吗? 如果能, 配料方式唯一吗? 求出这时所花费的成本;

(4) 用A_2, A_3, A_4, A_5这4种原料能配制2000kg这种食品吗?

(5) 用A_3, A_4, A_5这三种原料呢?

第三章　矩阵的运算

在上一章中, 利用矩阵的初等变换, 我们讨论了线性方程组求解的Gauss消元法. 矩阵是线性代数的主角之一, 作为代数工具, 我们在本章中要定义矩阵的四则运算. 矩阵的加法和减法的定义比较自然, 为了和后面的几何(线性空间之间的线性映射)相对应, 我们给出了看上去不是非常自然的矩阵乘法定义, 最后我们用了很大的篇幅讲矩阵的“除法”. 可以看到, 定义除法本质上需要定义乘法可逆的矩阵, 而矩阵的初等变换在讨论矩阵可逆性时起到了重大作用. 我们还介绍了矩阵的分块运算, 这是处理矩阵问题的重要工具和方法.

本章中$\mathbb{F}$表示任意一个数域.

§ 3.1　矩阵的运算

矩阵是数的推广. 数有四则运算, 自然希望可以将这些运算推广到矩阵上. 本节我们定义矩阵的加法, 减法和乘法, 以及一种特殊的乘法: 数乘, 而将“除法”的定义留到后面. 读者应该将矩阵的运算和数的运算相比较, 理解相似的地方, 注意不同的地方. 本节还将定义矩阵的转置和复矩阵的共轭.

3.1.1　加法

定义3.1 (矩阵的加法)**.** *设A和B是两个同型矩阵, 将所有对应位置元素相加得到的同型矩阵称为A和B的和, 记为$A+B$. 即*

$$(a_{ij})_{m\times n}+(b_{ij})_{m\times n}:=(a_{ij}+b_{ij})_{m\times n}.$$

例3.1.

$$\begin{pmatrix}1&2\\3&4\end{pmatrix}+\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}1+a&2+b\\3+c&4+d\end{pmatrix}.$$

由于矩阵加法相当于同时做许多数的加法, 所以容易知道: 设A,B,C是同型矩阵, 则

(i) (交换律) $A+B=B+A$;

(ii) (结合律) $(A+B)+C=A+(B+C)$;

(iii) (零元存在) $O_{m\times n}+A_{m\times n}=A=A+O_{m\times n}$;

(iv) (负元存在) 设$A=(a_{ij})_{m\times n}$, 记$-A:=(-a_{ij})_{m\times n}$, 称$-A$为$A$的负矩阵. 我们有

$$A+(-A)=O_{m\times n}.$$

于是, 矩阵集合$\mathbb{F}^{m\times n}$在矩阵的加法下成为交换群.

3.1.2 减法

数的减法可以通过加法来定义: 减去一个数等于加上这个数的相反数. 类似的, 有下面定义.

定义3.2 (矩阵的减法)**.** 设$A=(a_{ij})_{m\times n}$和$B=(b_{ij})_{m\times n}$是同型矩阵, 定义

$$A-B:=A+(-B)=(a_{ij}-b_{ij})_{m\times n}.$$

即$A-B$是对应元素相减得到的同型矩阵.

容易知道, 对于矩阵等式, 移项法则和消去律等成立. 例如, 如果$A+B=C+D$, 则有$A+B-C=D$.

3.1.3 数乘

定义3.3 (矩阵的数乘)**.** 设$a\in\mathbb{F}$, 而$A=(a_{ij})_{m\times n}\in\mathbb{F}^{m\times n}$, 定义$a$与$A$的数量乘积(数乘)为

$$aA=Aa:=(aa_{ij})_{m\times n},$$

即aA是将A的所有元素遍乘a得到的同型矩阵.

例3.2.

$$7\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}7a & 7b\\ 7c & 7d\end{pmatrix}.$$

容易证明: 设a,b是数, A和B是同型矩阵, 则

(i) $(ab)A=a(bA)=b(aA)$;

(ii) $(a+b)A=aA+bA$;

(iii) $a(A+B)=aA+aB$;

(iv) $1\times A=A$, $(-1)\times A=-A$;

(v) $0\times A_{m\times n}=O_{m\times n}$, $a\times O_{m\times n}=O_{m\times n}$, 且

$$aA=O\Longleftrightarrow a=0 \text{ 或者 } A=O;$$

(vi) $\operatorname{diag}(\underbrace{a,\dots,a}_{n})=aE_n$.

3.1.4 乘法

下面定义矩阵的乘法. 类比于矩阵的加法, 一个比较"自然"的想法是: 设$A=(a_{ij})_{m\times n}$和$B=(b_{ij})_{m\times n}$是两个同型矩阵, 将它们的对应位置元素相乘得到的同型矩阵记为$A\circ B$, 即

$$A\circ B:=(a_{ij}b_{ij})_{m\times n}=\begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n}\\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn}\end{pmatrix}.$$

称$A\circ B$为A和B的**Hadamard乘积**. 但是Hadamard乘积不是我们通常说的矩阵乘法, 通常用的矩阵乘法定义如下[1].

定义3.4 (矩阵的乘法). 设$A=(a_{ij})_{m\times s}\in\mathbb{F}^{m\times s}$, $B=(b_{ij})_{s\times n}\in\mathbb{F}^{s\times n}$满足$A$的列数和$B$的行数相等(都为$s$), 则定义$A$和$B$的**乘积**$AB$为$m\times n$矩阵$C=(c_{ij})_{m\times n}$, 其中

$$c_{ij}:=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}=\sum_{k=1}^{s}a_{ik}b_{kj},\quad (1\leqslant i\leqslant m, 1\leqslant j\leqslant n),$$

即c_{ij}等于A的第i行和B的第j列的对应位置元素乘积之和.

由定义, 我们有:

$$\boxed{A\text{的列数}}=\boxed{B\text{的行数}}\Longrightarrow AB\text{存在};$$

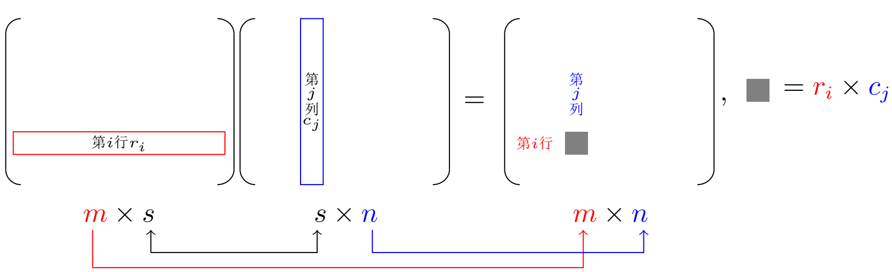
且此时有

$$\boxed{AB\text{的行数}}=\boxed{A\text{的行数}},\quad \boxed{AB\text{的列数}}=\boxed{B\ \text{的列数}}.$$

又由于

$$(a_{i1},a_{i2},\ldots,a_{is})\begin{pmatrix} b_{1j}\\ b_{2j}\\ \vdots\\ b_{sj}\end{pmatrix}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj},$$

所以AB的(i,j)位置元素等于A的第i行对应的行向量和B的第j列对应的列向量的乘积. 于是矩阵乘法可图示如下:



[1]这个定义才是自然的, 因为由此可得代数(矩阵)和几何(线性映射)的对应.

我们看一些例子.

例3.3. **(1)** 设$A=\begin{pmatrix}1&3&2\\4&-1&0\end{pmatrix}$, $B=\begin{pmatrix}1&1\\1&2\\0&1\end{pmatrix}$, 则

$$AB=\begin{pmatrix}4&9\\3&2\end{pmatrix},\quad BA=\begin{pmatrix}5&2&2\\9&1&2\\4&-1&0\end{pmatrix};$$

(2) 同阶对角阵相乘为同阶对角阵, 只要把相应的对角元相乘, 即

$$\begin{pmatrix}a_1&&&\\&a_2&&\\&&\ddots&\\&&&a_n\end{pmatrix}\begin{pmatrix}b_1&&&\\&b_2&&\\&&\ddots&\\&&&b_n\end{pmatrix}=\begin{pmatrix}a_1b_1&&&\\&a_2b_2&&\\&&\ddots&\\&&&a_nb_n\end{pmatrix}.$$

矩阵的乘法满足下面一些运算律(这里假定写出的矩阵乘积都有意义):

(i) (结合律) $(AB)C=A(BC)$;

(ii) (分配律) $A(B+C)=AB+AC$, $(B+C)D=BD+CD$;

(iii) $a(AB)=(aA)B=A(aB)=:aAB,\quad (a\in\mathbb{F})$;

(iv) (单位元存在) $E_mA_{m\times n}=A=AE_n$;

(v) $aA_{m\times n}=(aE_m)A=A(aE_n),\quad (a\in\mathbb{F})$;

(vi) $A_{m\times n}O_{n\times s}=O_{m\times s}$, $O_{s\times m}A_{m\times n}=O_{s\times n}$.

这里验证一下乘法结合律[2]. 设$A=(a_{ij})_{m\times s}$, $B=(b_{ij})_{s\times t}$, $C=(c_{ij})_{t\times n}$, 则$(AB)C$和$A(BC)$都是$m\times n$矩阵. 而对于$1\leqslant i\leqslant m$和$1\leqslant j\leqslant n$, 有

$$\begin{aligned}&(AB)C\text{ 的 }(i,j)\text{ 元}\\=&\sum_{k=1}^{t}(AB\text{ 的 }(i,k)\text{ 元 })c_{kj}=\sum_{k=1}^{t}\left(\sum_{l=1}^{s}a_{il}b_{lk}\right)c_{kj}\\=&\sum_{k=1}^{t}\sum_{l=1}^{s}a_{il}b_{lk}c_{kj}\end{aligned}$$

和

$$A(BC)\text{ 的 }(i,j)\text{ 元}$$

[2]在对应的几何世界这是显然成立的: 映射的复合是满足结合律的.

$$\begin{aligned}&=\sum_{l=1}^{s}a_{il}(BC\text{ 的 }(l,j)\text{ 元 })=\sum_{l=1}^{s}a_{il}\left(\sum_{k=1}^{t}b_{lk}c_{kj}\right)\\&=\sum_{l=1}^{s}\sum_{k=1}^{t}a_{il}b_{lk}c_{kj}=\sum_{k=1}^{t}\sum_{l=1}^{s}a_{il}b_{lk}c_{kj}.\end{aligned}$$

于是$(AB)C=A(BC)$.

由于矩阵的乘法满足结合律, 所以可以定义多个矩阵相乘: 按照任意次序相乘. 具体地, 设$A_1=(a_{ij}^{(1)}), A_2=(a_{ij}^{(2)}),\dots,A_r=(a_{ij}^{(r)})$为$r$个矩阵, 其中$A_1\in\mathbb{F}^{m\times s_1}$, $A_2\in\mathbb{F}^{s_1\times s_2}$, $\dots$, $A_{r-1}\in\mathbb{F}^{s_{r-2}\times s_{r-1}}$, $A_r\in\mathbb{F}^{s_{r-1}\times n}$, 则有$m\times n$矩阵$A_1A_2\cdots A_r$, 其$(i,j)$元为

$$\sum_{k_1=1}^{s_1}\sum_{k_2=1}^{s_2}\cdots\sum_{k_{r-1}=1}^{s_{r-1}}a_{i,k_1}^{(1)}a_{k_1,k_2}^{(2)}\cdots a_{k_{r-2},k_{r-1}}^{(r-1)}a_{k_{r-1},j}^{(r)}.$$

特别的, 可以定义方阵的幂. 即对于$A\in\mathbb{F}^{n\times n}$, 定义

$$\begin{cases}A^0:=E_n,\\A^k:=A^{k-1}A=\underbrace{AA\cdots A}_{k}, & (k\in\mathbb{N}).\end{cases}$$

下面显然成立

$$A^kA^l=A^{k+l},\quad (A^k)^l=A^{kl},\quad (\forall k,l\in\mathbb{Z}_{\geqslant 0}).$$

进而可以定义方阵$A\in\mathbb{F}^{n\times n}$的矩阵多项式, 即对于$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0\in\mathbb{F}[x]$, 我们定义[3]

$$f(A):=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E_n\in\mathbb{F}^{n\times n}.$$

容易证明: 如果有$\mathbb{F}[x]$中等式

$$f(x)+g(x)=h_1(x),\quad f(x)g(x)=h_2(x),$$

那么成立

$$f(A)+g(A)=h_1(A),\quad f(A)g(A)=h_2(A).$$

于是将$\mathbb{F}[x]$中的代数等式中的x用A替换, 等式仍然成立.

我们看两个例子.

例3.4. 设E_{ij}是(i,j)位置元素为1, 其余位置元素为0的n阶方阵, 证明:

$$E_{ij}E_{kl}=\delta_{jk}E_{il},$$

其中

$$\delta_{jk}=\begin{cases}1, & \text{如果}j=k,\\0, & \text{如果}j\neq k\end{cases}$$

是**Kronecker符号**.

[3] 多项式中不定元用方阵代入!

☞ **证明.** 设$E_{ij}E_{kl}=(a_{pq})$, 则容易知道当$p\neq i$或者$q\neq l$时有$a_{pq}=0$, 而

$$a_{il}=1\times(E_{kl}\text{的}(j,l)\text{元})=\begin{cases}1, & \text{如果}j=k,\\ 0, & \text{如果}j\neq k.\end{cases}$$

于是$E_{ij}E_{kl}=\delta_{jk}E_{il}$. □

方阵E_{ij}有特殊的重要性, 称其为基本矩阵. 例如, 设$A=(a_{ij})\in\mathbb{F}^{n\times n}$, 则有

$$A=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}E_{ij}.$$

请读者自己计算$E_{ij}B$和CE_{ij}.

例3.5. 设$N=\begin{pmatrix}0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1\\ & & & & 0\end{pmatrix}$是$n$阶方阵, 其中$n\geqslant 3$, 计算$N^2$.

☞ **解.** 可以按照矩阵的乘法直接计算, 也可以如下计算: 由于

$$N=E_{12}+E_{23}+\cdots+E_{n-1,n}=\sum_{i=1}^{n-1}E_{i,i+1},$$

所以

$$\begin{aligned}N^2&=\left(\sum_{i=1}^{n-1}E_{i,i+1}\right)\left(\sum_{j=1}^{n-1}E_{j,j+1}\right)=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}E_{i,i+1}E_{j,j+1}\\&=\sum_{i=1}^{n-2}E_{i,i+2}=E_{13}+E_{24}+\cdots+E_{n-2,n}.\end{aligned}$$

我们在例3.22中给出了另一种算法. □

- **矩阵乘法无交换律和消去律**

下面我们看矩阵乘法与数的乘法不一样的性质. 先看几个例子.

例3.6. **(1)** (向量相乘) $(a_1,a_2,\ldots,a_n)\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}=a_1b_1+a_2b_2+\cdots+a_nb_n$,

$$\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}(c_1,c_2,\ldots,c_m)=\begin{pmatrix}b_1c_1 & b_1c_2 & \cdots & b_1c_m\\ b_2c_1 & b_2c_2 & \cdots & b_2c_m\\ \vdots & \vdots & & \vdots\\ b_nc_1 & b_nc_2 & \cdots & b_nc_m\end{pmatrix}=(b_ic_j)_{n\times m};$$

(2) 设$A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $B=\begin{pmatrix}1\\0\end{pmatrix}$, 则$AB=\begin{pmatrix}1\\0\end{pmatrix}$, 而$BA$不存在;

(3) 设$A=\begin{pmatrix}1&0\\0&0\end{pmatrix}=E_{11}$, $B=\begin{pmatrix}0&1\\0&0\end{pmatrix}=E_{12}$, 则

$$AB=E_{11}E_{12}=E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad BA=E_{12}E_{11}=O=\begin{pmatrix}0&0\\0&0\end{pmatrix}.$$

由上面例子可以看出, 矩阵乘法的交换律不成立, 即通常$AB\neq BA$. 于是, 许多代数公式不能直接推广到矩阵上. 例如, 通常

$$(AB)^2\neq A^2B^2,\quad (A+B)^2\neq A^2+2AB+B^2.$$

上面的例子还告诉我们, 矩阵乘法有零因子: 即存在$A\neq O$和$B\neq O$, 但是$AB=O$. 于是, 从矩阵等式$AB=O$, 通常不能得到$A=O$或者$B=O$. 进而在矩阵的运算中, (乘法)消去律不一定成立. 即由$AB=AC$和$A\neq O$, 通常不能得到$B=C$.

- **可交换矩阵**

设A,B是两个矩阵, 如果AB和BA都存在, 且$AB=BA$, 则称A和B可交换. 如果两个矩阵可交换, 那么它们一定是同阶方阵.

设A,B是同阶方阵, 定义

$$[A,B]:=AB-BA,$$

称其为A和B的李括号(也称为换位元素). 于是, A和B可交换, 当且仅当$[A,B]=O$.

例3.7. *下面方阵可交换:*

(1) A_n和aE_n $(a\in\mathbb{F})$, 特别的, A_n和E_n, A_n和O_n;

(2) 同阶对角阵;

(3) 对方阵A, A^k和A^l, 进而, 关于A的两个矩阵多项式$f(A)$和$g(A)$.

如果矩阵A和B可交换, 那么中学学过的许多乘法公式仍然成立. 例如, 我们有[4]

$$\begin{aligned}&(AB)^k=A^kB^k;\\&(A+B)(A-B)=A^2-B^2;\\&(A+B)^2=A^2+2AB+B^2;\\&(A+B)^k=\binom{k}{0}A^k+\binom{k}{1}A^{k-1}B+\binom{k}{2}A^{k-2}B^2\\&\qquad+\cdots+\binom{k}{k-1}AB^{k-1}+\binom{k}{k}B^k=\sum_{i=0}^{k}\binom{k}{i}A^{k-i}B^i.\end{aligned}$$

[4]最后一个公式称为二项展开公式, 其中

$$\binom{k}{i}=C_k^i:=\frac{k!}{i!(k-i)!}$$

是二项式系数.

例3.8. 设$A=\begin{pmatrix}1 & a\\ 0 & 1\end{pmatrix}$, 求$A^k$, 其中$k\geqslant 0$.

☞ **解.** 我们有

$$A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+a\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=E_2+aN,\quad N=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

而$N^2=O$, 所以对任意的$i\geqslant 2$有$N^i=O$. 进而当$k\geqslant 1$时, 就有

$$\begin{aligned}A^k=&(E_2+aN)^k=\binom{k}{0}E_2^k+\binom{k}{1}E_2^{k-1}aN+O\\=&E_2+kaN=\begin{pmatrix}1 & ka\\ 0 & 1\end{pmatrix}.\end{aligned}$$

这个结果对$k=0$也成立, 于是$A^k=\begin{pmatrix}1 & ka\\ 0 & 1\end{pmatrix}$. □

- **线性方程组的矩阵形式**

本小节最后, 我们用矩阵乘法来表示线性方程组. 设有线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \qquad\qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m.\end{cases}\tag{3.1}$$

则有

$$\begin{aligned}(3.1)\iff&\begin{pmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\ \vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n\end{pmatrix}=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}\\ \iff&\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}\iff AX=\beta,\end{aligned}$$

其中$A=(a_{ij})\in\mathbb{F}^{m\times n}$是系数矩阵, $X=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$是未知数列向量, 而$\beta=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}$是常数项列向量. 于是线性方程组可以表示为$AX=\beta$, 注意到从形式上看这类似于一元线性方程$ax=b$, 一些线性方程组的性质可以用一元线性方程来理解. 这种表示方法的另一个好处是, 可以用矩阵来研究线性方程组解的结构; 反之, 也可以用线性方程组解的结构来讨论矩阵的性质.

3.1.5 转置

定义3.5 (矩阵的转置)**.** 设$A=(a_{ij})=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\in\mathbb{F}^{m\times n}$, 则$A$的转置矩阵$A^T$定义为[5]

$$A^T:=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}\in\mathbb{F}^{n\times m},$$

即A^T为$n\times m$阵, 且(i,j)元为a_{ji}.

矩阵转置本质上是行和列互换. 于是, 如果要处理关于列的问题, 取转置后就变成关于行的问题.

下面性质成立(假设所涉及的矩阵运算都可进行).

(i) $(A^T)^T=A$;

(ii) $(A+B)^T=A^T+B^T$;

(iii) $(aA)^T=aA^T, \quad (a\in\mathbb{F})$;

(iv) (穿脱原理) $(AB)^T=B^TA^T$, 进而$(A_1A_2\cdots A_k)^T=A_k^T\cdots A_2^TA_1^T$.

我们只验证穿脱原理. 设$A=(a_{ij})_{m\times s}$, $B=(b_{ij})_{s\times n}$, 则$(AB)^T$和B^TA^T都是$n\times m$矩阵. 对于$1\leqslant i\leqslant n$和$1\leqslant j\leqslant m$, 有

$$(AB)^T \text{ 的 } (i,j) \text{ 元 } = AB \text{ 的 } (j,i) \text{ 元 } = \sum_{k=1}^{s} a_{jk}b_{ki}$$

和

$$B^TA^T \text{ 的 } (i,j) \text{ 元 } = \sum_{k=1}^{s}(B^T \text{ 的 } (i,k) \text{ 元})(A^T \text{ 的 } (k,j) \text{ 元}) = \sum_{k=1}^{s} b_{ki}a_{jk}.$$

于是$(AB)^T=B^TA^T$.

例3.9. 设$A=\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, 则有

$$AB=\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}=\begin{pmatrix} 2 \\ 0 \end{pmatrix}\Longrightarrow (AB)^T=(2,0).$$

而

$$B^TA^T=(1,1)\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}=(2,0),$$

所以$(AB)^T=B^TA^T$. 但是A^TB^T不存在.

[5] 也有的教材将矩阵A的转置记为A'.

- **对称阵与反对称阵**

通常, 一个矩阵转置后得到不同的矩阵. 我们对转置后不变的矩阵和转置后成为其负阵的矩阵特别感兴趣.

定义3.6. 如果矩阵A满足$A^T = A$, 则称A是对称阵; 如果满足$A^T = -A$, 则称A是反对称阵.

设$A = (a_{ij})$, 则可得

$$A\text{ 是对称阵} \iff A\text{ 是方阵, 且 } a_{ij} = a_{ji}, \quad \forall i, j,$$
$$A\text{ 是反对称阵} \iff A\text{ 是方阵, 且 } a_{ij} = -a_{ji}, \quad \forall i, j.$$

特别有, 当n阶方阵$A = (a_{ij})$是反对称阵时成立

$$a_{11} = a_{22} = \cdots = a_{nn} = 0.$$

由此可以看出, 这里的对称(反对称)是相对于主对角而言.

例3.10. **(1)** 对角阵是对称阵;

(2) 设$A \in \mathbb{F}^{m\times n}$, 则$AA^T$和$A^TA$都是对称阵.

☞ **证明.** (2) 由于

$$(AA^T)^T = (A^T)^TA^T = AA^T,$$

所以AA^T是对称阵. □

例3.11. 设$X = (x_1, \ldots, x_n)^T$, $H = E_n - 2XX^T$. 证明:

(1) H是对称阵;

(2) 如果$X \neq 0$, 则: $HH^T = E_n \iff X^TX = 1$.

☞ **证明.** (1) 由于

$$\begin{aligned} H^T =&(E_n - 2XX^T)^T = E_n^T - 2(XX^T)^T \\ =&E_n - 2(X^T)^TX^T = E_n - 2XX^T = H, \end{aligned}$$

所以H是对称阵.
(2) 记$a = X^TX$, 则

$$\begin{aligned} HH^T =&H^2 = (E_n - 2XX^T)^2 = E_n - 4XX^T + 4XX^TXX^T \\ =&E_n - 4XX^T + 4X(X^TX)X^T = E_n - 4XX^T + 4aXX^T \\ =&E_n + 4(a-1)XX^T. \end{aligned}$$

于是

$$HH^T = E_n \iff (a-1)XX^T = O.$$

但是$X \neq 0$, 所以$XX^T \neq O$. 于是上面等价于$a = 1$, 即$X^TX = 1$. □

3.1.6 共轭

所有的元素都是实数的矩阵称为实矩阵, 所有元素都是复数的矩阵称为复矩阵. 复数有共轭复数, 类似地, 将矩阵看成复矩阵时, 可以定义它的共轭.

定义3.7 (复矩阵的共轭)**.** 设$A=(a_{ij})\in\mathbb{C}^{m\times n}$, A的共轭矩阵$\overline{A}$定义为

$$\overline{A}:=(\overline{a_{ij}})_{m\times n}=\begin{pmatrix}\overline{a_{11}} & \overline{a_{12}} & \cdots & \overline{a_{1n}}\\ \overline{a_{21}} & \overline{a_{22}} & \cdots & \overline{a_{2n}}\\ \vdots & \vdots & & \vdots\\ \overline{a_{m1}} & \overline{a_{m2}} & \cdots & \overline{a_{mn}}\end{pmatrix}\in\mathbb{C}^{m\times n}.$$

容易证明下面性质成立(假设所涉及的矩阵运算都可进行).

(i) $\overline{\overline{A}}=A$,

(ii) $\overline{A+B}=\overline{A}+\overline{B}$;

(iii) $\overline{aA}=\overline{a}\,\overline{A}, \qquad (a\in\mathbb{C})$;

(iv) $\overline{AB}=\overline{A}\,\overline{B}$;

(v) $\overline{A^T}=\overline{A}^T$.

习题3.1

A1. 设

$$A=\begin{pmatrix}a & b & c\\ c & b & a\\ 1 & 1 & 1\end{pmatrix},\quad B=\begin{pmatrix}1 & a & c\\ 1 & b & b\\ 1 & c & a\end{pmatrix},$$

计算AB和$[A,B]$.

A2. 计算:

(1) $\begin{pmatrix}2 & 1 & 1\\ 3 & 1 & 0\\ 0 & 1 & 2\end{pmatrix}^2$; (2) $(2,3,-1)\begin{pmatrix}1\\ -1\\ 1\end{pmatrix}$;

(3) $\begin{pmatrix}1\\ -1\\ 1\end{pmatrix}(2,3,-1)$; (4) $(x,y,1)\begin{pmatrix}a_{11} & a_{12} & b_1\\ a_{12} & a_{22} & b_2\\ b_1 & b_2 & c\end{pmatrix}\begin{pmatrix}x\\ y\\ 1\end{pmatrix}$;

(5) $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; (6) $(1,1,1)\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$;

(7) $\begin{pmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{pmatrix}\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$; (8) $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\begin{pmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{pmatrix}$;

(9) $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}^n$; (10) $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n$.

A3. 设$A=\begin{pmatrix} 1 & 2 & 1 & 2 \\ -1 & -2 & -1 & -2 \\ 1 & 2 & 1 & 2 \\ -1 & -2 & -1 & -2 \end{pmatrix}$, 计算$A^n$ $(n \geqslant 1)$.

A4. 计算方阵A的矩阵多项式$f(A)$, 其中

(1) $A=\begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix}$, $f(x)=x^2-x-1$;

(2) $A=\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$, $f(x)=x^2-5x+3$.

A5. 举例说明下列命题是错误的:

(1) 如果$A^2=O$, 则$A=O$;

(2) 如果$A^2=A$, 则$A=O$或者$A=E$;

(3) 如果$AX=AY$且$A \neq O$, 则$X=Y$.

A6. 求出与方阵A可交换的所有矩阵, 其中

(1) $A=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{pmatrix}$; (2) $A=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.

A7. 设n阶对角阵A的对角元两两不等, 证明: n阶方阵B与A可交换的充分必要条件是B为对角阵.

A8. 证明: (1) 和所有n阶方阵都可交换的方阵一定是纯量阵;

(2) 和所有n阶可逆方阵都可交换的方阵一定是纯量阵.

A9. 证明: (1) 如果A为实对称阵, 且$A^2=O$, 则$A=O$;

(2) 如果A和B都是n阶对称阵, 则AB是对称阵的充分必要条件是A和B可交换;

(3) 如果A和B都是n阶反对称阵, 则AB是反对称阵的充分必要条件是$AB=-BA$;

(4) 如果A和B都是n阶反对称阵, 则AB是对称阵的充分必要条件是A和B可交换;

(5) 如果A和B都是n阶对称阵, 则$[A,B]$是反对称阵;

(6) 如果A和B都是n阶反对称阵, 则$[A,B]$也是反对称阵;

(7) 如果A和B中有一个是n阶对称阵, 另一个是n阶反对称阵, 则$[A,B]$是对称阵;

(8) 任意n阶方阵可以唯一地写为一个对称阵和一个反对称阵之和.

A10. 设A为n阶方阵. 如果A满足$A^2 = A$, 则称A是幂等阵; 如果A满足$A^2 = E_n$, 则称A是对合阵.

(1) 求出所有2阶实幂等阵;

(2) 求出所有2阶实对合阵;

(3) 设n阶方阵A和B满足$2A = B + E_n$, 证明: A是幂等阵当且仅当B是对合阵.

B1. 计算:

(1) $\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}^n$; (2) $\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}^n$;

(3) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}^n$; (4) $\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}_{n\times n}^n$.

B2. 设n阶方阵A, B的元素都是非负实数, 证明: 如果AB有全零行, 则A或者B有全零行.

§ 3.2 可逆矩阵的定义和性质

本节的目标是定义矩阵的"除法", 那么应该如何自然的定义呢? 我们回忆数的除法的定义. 如果数$a \neq 0$, 则a有乘法逆(倒数)a^{-1}, 此时定义

$$b \div a = \frac{b}{a} = a^{-1} \times b = b \times a^{-1}.$$

即我们是通过乘法逆用乘法来定义除法的. 类比于此, 矩阵的除法应该如下:

如果矩阵A满足$\boxed{?}$, 则A有乘法逆A^{-1}, 此时定义

$$B \div A \overset{?}{=} \begin{cases} A^{-1} \times B, & \text{左除} \\ B \times A^{-1}. & \text{右除} \end{cases}$$

但是通常$A^{-1}B \neq BA^{-1}$, 所以我们要区分左除还是右除. 于是我们不讲矩阵的除法, 而直接用乘法逆去左乘或者右乘一个矩阵, 即矩阵的"除法"事实上应该为$A^{-1}B$和BA^{-1}两种.

由上面的分析, 我们需要知道何种矩阵有乘法逆, 以及A^{-1}是什么. 那么, 如何定义矩阵可逆? 回到数上, 对于$a \in \mathbb{F}$, 我们知道

$$a\text{ 可逆} \iff a \neq 0 \iff \exists b \in \mathbb{F}, \text{使得 } ab = ba = 1.$$

此时, $b = a^{-1}$. 类比到矩阵上, 对于矩阵A, 我们有下面两种可能的选择

(i) A可逆$\overset{?}{\iff} A \neq O$;

(ii) A可逆$\overset{?}{\iff}$ 存在矩阵B, 使得$AB = BA = E$.

如果用第(i)种定义, 那么$A=(1,0)$是可逆的. 取$B=\begin{pmatrix}0\\1\end{pmatrix}$, 则

$$AB=(1,0)\begin{pmatrix}0\\1\end{pmatrix}=0.$$

由于A可逆, 所以上面等式两边可以“左除”A, 得到$B=A^{-1}AB=A^{-1}0=0$. 这得到矛盾, 所以定义(i)不是一个好的选择. 而做除法时我们需要可逆阵的逆, 所以采用定义(ii)可能时比较好的选择[6]. 注意到此时蕴含A和B可交换, 所以A和B是同阶方阵.

定义3.8. 设$A\in\mathbb{F}^{n\times n}$是方阵, 如果存在矩阵$B\in\mathbb{F}^{n\times n}$, 使得

$$AB=BA=E_n,$$

则称矩阵A可逆(或: 非奇异). 否则, 称A不可逆(或: 奇异).

如果还有矩阵C, 使得$AC=CA=E$, 则

$$B=BE=B(AC)=(BA)C=EC=C.$$

所以满足$AB=BA=E$的矩阵B是唯一的, 称为A的逆阵, 记为A^{-1}.

例3.12. 设矩阵A可逆, 证明:

(1) $AB=C\implies B=A^{-1}C$;

(2) $DA=C\implies D=CA^{-1}$.

☞ **证明.** 以(1)为例, (2)类似. 我们有

$$B=EB=(A^{-1}A)B=A^{-1}(AB)=A^{-1}C,$$

即得到(1). □

从可逆矩阵的定义看出, 要判别一个矩阵是否可逆并不是一件轻松的事. 对于n阶方阵A, 我们要在无穷多的矩阵中找到那个唯一的B, 使得$AB=BA=E$. 如果用待定系数法, 将B的n^2个元素看成未知数, 这相当于问一个有n^2个未知数的线性方程组是否相容. 当n比较大时, 该线性方程组并不容易求解. 在解决这一困难前, 我们需要有一些例子给我们一些感性的认识. 特别地, 我们当然希望零矩阵和单位阵像数的0和1一样分别不可逆和可逆, 下面的例子给出了更多.

例3.13. **(1)** 设方阵A有全零行(列), 则A不可逆. 特别的, O_n不可逆;

(2) 设$A=\operatorname{diag}(a_1,a_2,\ldots,a_n)$是对角阵, 则

$$A\text{可逆}\iff a_1,a_2,\ldots,a_n\neq 0.$$

此时,

$$A^{-1}=\operatorname{diag}(a_1^{-1},a_2^{-1},\ldots,a_n^{-1}).$$

特别的, E_n可逆, 且$E_n^{-1}=E_n$.

---

[6]对于任意有单位元1的环R, R中乘法可逆元也是如此定义的: $a\in R$可逆$\iff$ 存在$b\in R$, 使得$ab=ba=1$.

证明. (1) 设$A=(a_{kl})$的第i行是全零行, 即

$$a_{i1}=a_{i2}=\cdots=a_{in}=0.$$

则对于任意的n阶方阵$B=(b_{kl})$, 有AB的(i,j)元素为

$$\sum_{k=1}^{n} a_{ik}b_{kj}=\sum_{k=1}^{n} 0\cdot b_{kj}=0,$$

即AB的第i行也是全零行. 于是$AB\neq E$. 即A不可逆.

类似可证, 如果A有全零列, 则A也不可逆.

(2) 如果存在某个$a_i=0$, 则A的第i行为全零行, 由(1)得A不可逆. 如果任意的$a_i\neq 0$, 则由于

$$\operatorname{diag}(a_1,a_2,\ldots,a_n)\operatorname{diag}(a_1^{-1},a_2^{-1},\ldots,a_n^{-1})=E$$

和

$$\operatorname{diag}(a_1^{-1},a_2^{-1},\ldots,a_n^{-1})\operatorname{diag}(a_1,a_2,\ldots,a_n)=E,$$

得A可逆, 且$A^{-1}=\operatorname{diag}(a_1^{-1},a_2^{-1},\ldots,a_n^{-1})$. □

对方阵A, 前面定义了A的幂. 如果A可逆, 则可如下定义A的负指数

$$A^{-k}:=(A^{-1})^k,\quad \forall k\in\mathbb{N}.$$

可以证明下面成立

$$A^kA^l=A^{k+l},\quad (A^k)^l=A^{kl},\quad (\forall k,l\in\mathbb{Z}).$$

于是, 对于可逆矩阵$A\in\mathbb{F}^{n\times n}$, 除了$A$的矩阵多项式, 比如$A^2+2A+E$, 还有类似于$A^{-2}+2E+A$的方阵存在. 令$f(x)=x^{-2}+2+x$, 称其为一个Laurent多项式, 则

$$A^{-2}+2E+A=f(A).$$

一般地, 定义**Laurent多项式**集合为

$$\begin{aligned}\mathbb{F}[x,x^{-1}]:=\{&a_{-k}x^{-k}+a_{-k+1}x^{-k+1}+\cdots+a_0+a_1x+\cdots+a_mx^m\mid k,m\geqslant 0,\\ &a_{-k},a_{-k+1},\ldots,a_0,a_1,\ldots,a_m\in\mathbb{F}\}\supset\mathbb{F}[x],\end{aligned}$$

则对任意Laurent多项式$f(x)=a_{-k}x^{-k}+\cdots+a_0+a_1x+\cdots+a_mx^m\in\mathbb{F}[x,x^{-1}]$, 可以定义

$$f(A):=a_{-k}A^{-k}+\cdots+a_0E+a_1A+\cdots+a_mA^m\in\mathbb{F}^{n\times n}.$$

类似于多项式类, 可定义Laurent多项式的加法, 减法和乘法, 这里自然定义

$$x^n\cdot x^m=x^{n+m},\quad \forall n,m\in\mathbb{Z}.$$

例如, 我们有

$$\begin{aligned}&(-2x^{-2}+x^{-1}+2+6x)+(-3x^{-1}+5+2x)=-2x^{-2}-2x^{-1}+7+8x,\\ &(2x^{-1}+x)(3x^{-1}+x^2)=2x^{-1}\cdot 3x^{-1}+2x^{-1}\cdot x^2+x\cdot 3x^{-1}+x\cdot x^2\end{aligned}$$

$$= 6x^{-2} + 2x + 3 + x^3.$$

于是, 可以证明, 对任意$f(x), g(x) \in \mathbb{F}[x, x^{-1}]$, 如果

$$f(x) + g(x) = h_1(x), \quad f(x)g(x) = h_2(x),$$

则有

$$f(A) + g(A) = h_1(A), \quad f(A)g(A) = h_2(A).$$

进而, 将$\mathbb{F}[x, x^{-1}]$中的代数等式中的x用可逆矩阵A替换, 等式仍然成立.

根据定义, 类比于可逆的数的性质, 我们可以猜到并证明下面的性质.

命题3.1. *设$A, B \in \mathbb{F}^{n\times n}$是可逆阵, $a \in \mathbb{F}^{\times}$, 则$A^{-1}, aA, AB$和$A^T$都可逆, 且*

$$\begin{aligned}
&(A^{-1})^{-1} = A,\\
&(aA)^{-1} = a^{-1}A^{-1},\\
&(AB)^{-1} = B^{-1}A^{-1}, \quad (\text{穿脱原理})\\
&(A^T)^{-1} = (A^{-1})^T.
\end{aligned}$$

进而, 如果$A_1, A_2, \ldots, A_s \in \mathbb{F}^{n\times n}$都可逆, 则$A_1A_2\cdots A_s$可逆, 且

$$(A_1A_2\cdots A_s)^{-1} = A_s^{-1}\cdots A_2^{-1}A_1^{-1}.$$

☞ **证明.** 我们以可逆方阵的乘积为例, 其它类似. 事实上, 由于

$$\begin{aligned}
(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E,\\
(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E,
\end{aligned}$$

所以AB可逆, 且$(AB)^{-1} = B^{-1}A^{-1}$. □

我们看一些例子.

例3.14. *设$A, B \in \mathbb{F}^{n\times n}$是可逆阵, 如果$A + B$可逆, 证明: $A^{-1} + B^{-1}$可逆. 并求其逆.*

☞ **解.** 由于

$$A^{-1} + B^{-1} = A^{-1}(E + AB^{-1}) = A^{-1}(B + A)B^{-1},$$

且A^{-1}, B^{-1}和$A + B$都可逆, 所以$A^{-1} + B^{-1}$可逆, 且

$$\begin{aligned}
(A^{-1} + B^{-1})^{-1} =& (A^{-1}(A + B)B^{-1})^{-1}\\
=& (B^{-1})^{-1}(A + B)^{-1}(A^{-1})^{-1} = B(A + B)^{-1}A.
\end{aligned}$$

类似于上面, 也可以求出$(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B$. □

例3.15. *设$A \in \mathbb{F}^{n\times n}$, $A \neq O$, 且$A^2 = aA$, 其中$a \in \mathbb{F}$. 问a取何值时, 方阵$A + E$可逆? 在$A + E$可逆时, 求它的逆.*

☞ **解.** 由综合除法可得

$$x^2 - ax = (x+1)(x-a-1) + (a+1),$$

于是

$$A^2 - aA = (A+E)(A-(a+1)E) + (a+1)E,$$

即

$$(A+E)(A-(a+1)E) = -(a+1)E.$$

当$a+1 \neq 0$时, 有

$$(A+E)\left(E - \frac{1}{a+1}A\right) = E.$$

而上面等式左边的两个因子可以交换, 所以有

$$\left(E - \frac{1}{a+1}A\right)(A+E) = E.$$

于是$A+E$可逆, 且

$$(A+E)^{-1} = E - \frac{1}{a+1}A.$$

当$a+1=0$时, 有$(A+E)A=O$. 如果$A+E$可逆, 则$A=O$, 矛盾. 所以$A+E$不可逆.

综上, 当且仅当$a \neq -1$时$A+E$可逆. 且可逆时$(A+E)^{-1} = E - \frac{1}{a+1}A$. □

例3.16. 设$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$, *证明: A可逆, 并求A^{-1}.*

☞ **解.** 容易得到$A^2 = 4E$, 所以A可逆, 且

$$A^{-1} = \frac{1}{4}A = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

解毕. □

上面例子中的A有很大的特殊性. 下面遗留的问题是, 给定一个一般的n阶方阵A, 如何判别它是否可逆? 当可逆时如何求A^{-1}? 如前分析, 从定义来看, 我们找B, 使得$AB=BA=E$, 相当于求解一个有n^2个未知数的线性方程组. 当n比较大时, 这是一个困难的任务. 那么如何解决这个问题呢? 前面解线性方程组的分析问题的方法可以给我们一些有益的启示, 我们可以试着用初等变换化一般为简单来考虑方阵的可逆性. 为此需要考虑初等变换是否改变方阵的可逆性, 以及简单的矩阵何时可逆这两个问题. 在此之前, 作为工具, 也作为重要的方法, 我们先介绍矩阵的分块以及按照矩阵运算定义如何进行分块运算.

习题3.2

A1. 设A是n阶可逆阵. 证明:

(1) 如果A是对称阵, 则A^{-1}也是对称阵;

(2) 如果A是反对称阵, 则A^{-1}也是反对称阵.

A2. 设A, B是n阶方阵, 满足A, B和$AB-E$都可逆, 证明:

(1) $A-B^{-1}$可逆, 并求其逆阵;

(2) $(A-B^{-1})^{-1}-A^{-1}$也可逆, 并求其逆阵.

A3. 设n阶方阵A适合

$$a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mE_n=0,$$

其中$m\geqslant 1$, $a_0a_m\neq 0$. 证明: 方阵A可逆, 并求A^{-1}.

A4. 设A为n阶方阵, 如果存在正整数k, 使得$A^k=O$, 则称A是幂零阵. 此时, 使得$A^k=O$成立的最小正整数称为方阵A的幂零指数. 设A为幂零矩阵, 且幂零指数是k, 证明: $E-A$可逆, 并求$(E-A)^{-1}$.

A5. 设A是n阶方阵.

(1) 如果A满足$A^3+3A^2+3A=O$, 证明: $A+2E$可逆, 并求$(A+2E)^{-1}$;

(2) 如果A满足$A^3+E=O$, 证明: A^2+E可逆, 并求$(A^2+E)^{-1}$.

B1. 设A为$n\times m$矩阵, B为$m\times n$矩阵, 如果E_n-AB可逆, 证明: E_m-BA也可逆, 并求$(E_m-BA)^{-1}$.

§ 3.3 矩阵的分块

我们知道, 通常越低阶的矩阵越容易处理. 在处理有关矩阵的问题时, 有时将某些部分看成整体, 将矩阵看成低阶矩阵来处理会更加简单和方便, 这就有矩阵的分块. 设$A=(a_{ij})$是任意的$m\times n$矩阵, 设想用一些水平线和竖直线把A的元素分割成若干个长方形小块, 每一小块对应的矩阵用矩阵记号表示, 就得到A的一个分块. 具体的, 假设在A的第$m_1+\cdots+m_i$行与第$m_1+\cdots+m_i+1$行之间有一条水平线, 其中$i=1,\ldots,s-1$; 而在A的第$n_1+\cdots+n_j$列与第$n_1+\cdots+n_j+1$列之间有一条竖直线, 其中$j=1,\ldots,r-1$, 则将A分成$s\times r$个块

$$A=(a_{ij})_{m\times n}=\begin{array}{cc}\left(\begin{array}{c:c:c:c} A_{11} & A_{12} & \cdots & A_{1r} \\ \hdashline A_{21} & A_{22} & \cdots & A_{2r} \\ \hdashline \vdots & \vdots & & \vdots \\ \hdashline A_{s1} & A_{s2} & \cdots & A_{sr} \end{array}\right) & \begin{array}{c} m_1 \\ m_2 \\ \vdots \\ m_s \end{array} \\ \begin{array}{cccc} n_1 & n_2 & \cdots & n_r \end{array} & \end{array}$$

其中A_{ij}是$m_i\times n_j$阵, 称为A的子阵. 此时, A可简记为$A=(A_{ij})_{s\times r}$. 我们有$m_1+\cdots+m_s=m$, $n_1+\cdots+n_r=n$. 称$(m_1,m_2,\ldots,m_s)$为分块矩阵A的行型, 称$(n_1,n_2,\ldots,n_r)$为列型.

例3.17. (1) $\left(\begin{array}{c:cc} a_{11} & a_{12} & a_{13} \\ \hdashline a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right)=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, *行型为$(1,2)$, 列型为$(1,2)$;*

(2) $\left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ \hdashline a_{21} & a_{22} & a_{23} \\ \hdashline a_{31} & a_{32} & a_{33} \end{array}\right)=\begin{pmatrix} A_{11} \\ A_{21} \\ A_{31} \end{pmatrix}$, *行型为$(1,1,1)$, 列型为空.*

类比于三角阵, 我们特别喜欢下列形式的分块阵. 如果矩阵A分块后有如下形式

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ & A_{22} & \cdots & A_{2r} \\ & & \ddots & \vdots \\ & & & A_{rr} \end{pmatrix}, \quad \text{(空白的块为零子矩阵)},$$

则称A为分块上三角阵(或准上三角阵). 类似的可以定义分块下三角阵(或准下三角阵)为

$$\begin{pmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \vdots & \vdots & \ddots & \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{pmatrix}.$$

最后, 如果分块矩阵A即是分块上三角阵, 又是分块下三角阵, 则称A是分块对角阵(或准对角阵). 此时, 记

$$A=\operatorname{diag}(A_{11},A_{22},\ldots,A_{rr})=\begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{rr} \end{pmatrix}.$$

3.3.1 分块运算

将矩阵分块后, 按照矩阵运算的定义, 我们可以得到下面关于分块矩阵的运算规则. 由于证明都很简单, 这里省略.

命题3.2 (分块加法). 设$A,B\in\mathbb{F}^{m\times n}$按照同型分块$A=(A_{ij})_{s\times r}$, $B=(B_{ij})_{s\times r}$, 则有

$$A+B=(A_{ij}+B_{ij})=\begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1r}+B_{1r} \\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2r}+B_{2r} \\ \vdots & \vdots & & \vdots \\ A_{s1}+B_{s1} & A_{s2}+B_{s2} & \cdots & A_{sr}+B_{sr} \end{pmatrix}.$$

即$A+B$也是同型的分块阵, 且只需将A和B对应位置上的小块相加即可.

命题3.3 (分块数乘). 设$A\in\mathbb{F}^{m\times n}$有分块$A=(A_{ij})_{s\times r}$, 设$a\in\mathbb{F}$, 则有

$$aA=(aA_{ij})_{s\times r}=\begin{pmatrix} aA_{11} & aA_{12} & \cdots & aA_{1r} \\ aA_{21} & aA_{22} & \cdots & aA_{2r} \\ \vdots & \vdots & & \vdots \\ aA_{s1} & aA_{s2} & \cdots & aA_{sr} \end{pmatrix}.$$

命题3.4 (分块乘法)**.** 设$A \in \mathbb{F}^{m\times l}$和$B \in \mathbb{F}^{l\times n}$有分块:

$$A = (A_{ij})_{s\times t}, \quad B = (B_{ij})_{t\times r},$$

其中A分块的列型等于B分块的行型, 则AB也是分块矩阵$(C_{ij})_{s\times r}$, 其中

$$\text{分块的行型} = A\text{分块的行型}, \quad \text{分块的列型} = B\text{分块的列型},$$

且

$$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{it}B_{tj}, \quad (1 \leqslant i \leqslant s, 1 \leqslant j \leqslant r).$$

命题3.5 (分块转置)**.** 设$A \in \mathbb{F}^{m\times n}$有分块$A = (A_{ij})_{s\times r}$, 则有

$$A^T = (A_{ji}^T)_{r\times s} = \begin{pmatrix} A_{11}^T & A_{21}^T & \cdots & A_{s1}^T \\ A_{12}^T & A_{22}^T & \cdots & A_{s2}^T \\ \vdots & \vdots & & \vdots \\ A_{1r}^T & A_{2r}^T & \cdots & A_{sr}^T \end{pmatrix}.$$

命题3.6 (分块共轭)**.** 设$A \in \mathbb{C}^{m\times n}$有分块$A = (A_{ij})_{s\times r}$, 则有

$$\overline{A} = (\overline{A_{ij}})_{s\times r} = \begin{pmatrix} \overline{A_{11}} & \overline{A_{12}} & \cdots & \overline{A_{1s}} \\ \overline{A_{21}} & \overline{A_{22}} & \cdots & \overline{A_{2s}} \\ \vdots & \vdots & & \vdots \\ \overline{A_{r1}} & \overline{A_{r2}} & \cdots & \overline{A_{rs}} \end{pmatrix}.$$

从上面可以看出, 进行分块运算时, 我们需要正确分块, 然后把每一小块看成数, 按照矩阵的运算规则进行运算即可. 需要注意的是, 两个小块相乘时不能交换位置, 以及分块转置时, 每个小块也要转置.

例3.18. 设$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -3 & 1 & 0 \\ -4 & -5 & 0 & 1 \end{pmatrix}$, 求$AB$.

☞ **解.** 将A和B分块

$$A = \begin{pmatrix} E_2 & O \\ A_1 & E_2 \end{pmatrix}, \quad B = \begin{pmatrix} E_2 & O \\ -A_1 & E_2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix},$$

则有

$$\begin{aligned} AB &= \begin{pmatrix} E_2 & O \\ A_1 & E_2 \end{pmatrix} \begin{pmatrix} E_2 & O \\ -A_1 & E_2 \end{pmatrix} = \begin{pmatrix} E_2E_2 - OA_1 & E_2O + OE_2 \\ A_1E_2 - E_2A_1 & A_1O + E_2E_2 \end{pmatrix} \\ &= \begin{pmatrix} E_2 & O \\ O & E_2 \end{pmatrix} = E_4. \end{aligned}$$

当然也可以直接计算. □

例3.19. 设$A=\mathrm{diag}(A_1,A_2,\dots,A_r)$为分块对角阵, 其中$A_i$都是方阵, 证明:

$$A\text{ 可逆 }\Longleftrightarrow A_1,A_2,\dots,A_r\text{ 都可逆}.$$

当A可逆时求A^{-1}.

☞ **解.** 设A_i是n_i阶方阵, 对于n阶方阵B, 做分块$B=(B_{ij})$, 其中B_{ij}是$n_i\times n_j$阵. 由于

$$AB=\begin{pmatrix}A_1B_{11}&A_1B_{12}&\cdots&A_1B_{1r}\\A_2B_{21}&A_2B_{22}&\cdots&A_2B_{2r}\\\vdots&\vdots&&\vdots\\A_rB_{r1}&A_rB_{r2}&\cdots&A_rB_{rr}\end{pmatrix},$$

$$BA=\begin{pmatrix}B_{11}A_1&B_{12}A_2&\cdots&B_{1r}A_r\\B_{21}A_1&B_{22}A_2&\cdots&B_{2r}A_r\\\vdots&\vdots&&\vdots\\B_{r1}A_1&B_{r2}A_2&\cdots&B_{rr}A_r\end{pmatrix},$$

所以有

$$AB=BA=E\Longleftrightarrow\begin{cases}A_iB_{ii}=B_{ii}A_i=E_{n_i}, & 1\leqslant i\leqslant r,\\A_iB_{ij}=O,\quad B_{ij}A_j=O, & 1\leqslant i\neq j\leqslant r.\end{cases}$$

这又等价于

$$\begin{cases}A_i\text{可逆, 且}A_i^{-1}=B_{ii}, & 1\leqslant i\leqslant r,\\B_{ij}=O, & 1\leqslant i\neq j\leqslant r.\end{cases}$$

于是

$$A\text{可逆}\Longleftrightarrow A_1,A_2,\dots,A_r\text{都可逆}.$$

且当A可逆时, 有

$$A^{-1}=\mathrm{diag}(A_1^{-1},A_2^{-1},\dots,A_r^{-1}).$$

解毕[7]. □

3.3.2 按行(列)分块

我们常常需要讲矩阵的某行和某列, 于是下面两种特殊的矩阵分块特别重要. 设$A=(a_{ij})_{m\times n}\in\mathbb{F}^{m\times n}$, 则有分块

$$A=\left(\begin{array}{cccc}a_{11}&a_{12}&\cdots&a_{1n}\\\hdashline a_{21}&a_{22}&\cdots&a_{2n}\\\hdashline\vdots&\vdots&&\vdots\\\hdashline a_{m1}&a_{m2}&\cdots&a_{mn}\end{array}\right)=\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_m\end{pmatrix}$$

和

$$A=\left(\begin{array}{c:c:c:c}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{array}\right)=(\beta_1,\beta_2,\dots,\beta_n).$$

[7] 这个例子是例3.13的推广. 后面我们将用初等变换给出一个更简单的证明.

称前者为按行分块, 后者为按列分块. 例如, 设$B \in \mathbb{F}^{n\times s}$和$C \in \mathbb{F}^{t\times m}$, 则有常用的分块乘法

$$\begin{aligned} AB &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} B = \begin{pmatrix} \alpha_1 B \\ \alpha_2 B \\ \vdots \\ \alpha_m B \end{pmatrix}, \\ CA &= C(\beta_1, \beta_2, \ldots, \beta_n) = (C\beta_1, C\beta_2, \ldots, C\beta_n). \end{aligned}$$

将上面的A取为单位阵, 则可取出矩阵的行和列. 事实上, 设单位矩阵E_n按列分块为

$$E_n = (e_1, e_2, \ldots, e_n), \quad e_i = e_i^{(n)} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{第 } i \text{ 行},$$

则E_n的按行分块为

$$E_n = \begin{pmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_n^T \end{pmatrix}.$$

我们可以如下取出矩阵的任意行, 列和元素.

例3.20. 设$A = (a_{ij})_{m\times n} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = (\beta_1, \beta_2, \ldots, \beta_n)$, 证明:

(1) $\alpha_i = {e_i^{(m)}}^T A$;

(2) $\beta_j = Ae_j^{(n)}$;

(3) $a_{ij} = {e_i^{(m)}}^T Ae_j^{(n)}$.

☞ **证明.** 我们证明(2). 事实上, 我们有

$$(\beta_1, \beta_2, \ldots, \beta_n) = A = AE_n = A(e_1, e_2, \ldots, e_n) = (Ae_1, Ae_2, \ldots, Ae_n).$$

于是就有$\beta_j = Ae_j$.

类似可以证(1), 而利用(1)和(2)可以得到(3). □

我们再看两个例子.

例3.21. 设$A\in\mathbb{F}^{m\times n}$, 满足对任意的$X\in\mathbb{F}^n$有$AX=0$, 证明: $A=O$.

☞ **证明.** 设有列分块$A=(\beta_1,\beta_2,\dots,\beta_n)$. 对于$i=1,2,\dots,n$, 取$X=e_i$, 则有

$$\beta_i=Ae_i=0.$$

所以$A=O$. □

例3.22. 设矩阵$A=\begin{pmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&0&1\\&&&&0\end{pmatrix}_{n\times n}$, 计算$A^k$, 其中$k\geqslant 0$.

☞ **证明.** 按列分块, $A=(0,e_1,\dots,e_{n-1})$. 于是

$$A^2=A(0,e_1,\dots,e_{n-1})=(0,Ae_1,\dots,Ae_{n-1})=(0,0,e_1,\dots,e_{n-2}).$$

类似的, 假设已证$A^k=(\underbrace{0,\dots,0}_{k},e_1,\dots,e_{n-k})$, 其中$1\leqslant k\leqslant n-1$, 则

$$\begin{aligned}A^{k+1}=&A(\underbrace{0,\dots,0}_{k},e_1,\dots,e_{n-k})=(\underbrace{0,\dots,0}_{k},Ae_1,\dots,Ae_{n-k})\\=&(\underbrace{0,\dots,0}_{k+1},e_1,\dots,e_{n-k-1}).\end{aligned}$$

特别可得$A^n=O$. 所以当$0\leqslant k\leqslant n-1$时, 有

$$A^k=(\underbrace{0,\dots,0}_{k},e_1,\dots,e_{n-k})=\begin{pmatrix}0&\cdots&0&1&&\\&\ddots&&\ddots&\ddots&\\&&\ddots&&\ddots&1\\&&&\ddots&&0\\&&&&\ddots&\vdots\\&&&&&0\end{pmatrix}\left.\begin{matrix}\\\\\\\\\\\\\end{matrix}\right\}k;$$

而当$k\geqslant n$时, 有$A^n=O$. □

习题3.3

A1. 设A和B是n阶方阵, 计算

$$\begin{pmatrix}O&E_n\\E_n&O\end{pmatrix}\begin{pmatrix}A&O\\O&B\end{pmatrix}\begin{pmatrix}O&E_n\\E_n&O\end{pmatrix}.$$

A2. 设方阵A和B可交换, 求$\begin{pmatrix}A&B\\O&A\end{pmatrix}^n$.

A3. 证明: 用对角阵左(右)乘矩阵A, 相当于用该对角阵的对角元分别去乘A的相应的行(列). 将该结果推广到分块阵.

A4. 证明: (1) 上三角方阵的乘积仍是上三角的, 且对角元为相应对角元的乘积, 即如果A_1, A_2, $\ldots$, A_s都是n阶上三角阵, 其中A_j的(i,i)元为$a_{ii}^{(j)}$, $i=1,2,\ldots,n$, 则$A=A_1A_2\cdots A_s$也是上三角阵, 且A的(i,i)元为$a_{ii}^{(1)}a_{ii}^{(2)}\cdots a_{ii}^{(s)}$, $i=1,2,\ldots,n$. 将该结果推广到分块阵;

(2) 如果$A_1,A_2,\cdots,A_n$是n个n阶上三角方阵, 且方阵$A_1,A_2,\ldots,A_n$的对角元都为零(称对角元都为零的上三角阵为严格上三角阵), 则$A_1A_2\cdots A_n=O$. 将该结果推广到分块阵.

A5. (1) 设n阶方阵$N=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$, 证明: N是幂零阵. 并求N的幂零指数;

(2) 设A是n阶上三角阵, 证明: A是幂零阵的充分必要条件是A的对角元全为零, 并且此时A的幂零指数不超过n.

A6. 证明: 对角块都是方阵的分块上三角方阵为幂零的必要且充分条件是, 它的每一个对角块都是幂零的.

A7. 求出所有2阶实幂零阵.

A8. 称方阵

$$C=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{n\times n} = \begin{pmatrix} 0 & E_{n-1} \\ 1 & 0 \end{pmatrix}$$

为n阶循环移位矩阵. 证明:

(1) 用C左乘矩阵, 相当于将这个矩阵的行向上移一行, 而第一行移到最后一行; 用C右乘矩阵, 相当于将这个矩阵的列向右移一列, 而最后一列移到第一列;

(2) $C^k=\begin{cases} \begin{pmatrix} 0 & E_{n-k} \\ E_k & 0 \end{pmatrix}, & k=1,2,\ldots,n-1, \\ E_n, & k=n; \end{cases}$

(3) 设J是所有元素为1的n阶方阵, 则

$$E+C+C^2+\cdots+C^{n-1}=J;$$

(4) 设

$$A=\begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_n & a_1 & \cdots & a_{n-2} & a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_3 & a_4 & \cdots & a_1 & a_2 \\ a_2 & a_3 & \cdots & a_n & a_1 \end{pmatrix},$$

称A是循环矩阵, 则

$$A=a_1E+a_2C+a_3C^2+\cdots+a_nC^{n-1};$$

(5) 两个n阶循环阵的乘积仍是循环阵;

(6) $AB=B\operatorname{diag}(f(1),f(\zeta),f(\zeta^2),\ldots,f(\zeta^{n-1}))$, 其中$\zeta$是$n$次本原单位根,

$$A=\begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_n & a_1 & \cdots & a_{n-2} & a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_3 & a_4 & \cdots & a_1 & a_2 \\ a_2 & a_3 & \cdots & a_n & a_1 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta & \zeta^2 & \cdots & \zeta^{n-1} \\ 1 & \zeta^2 & \zeta^4 & \cdots & \zeta^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \zeta^{n-1} & \zeta^{2(n-1)} & \cdots & \zeta^{(n-1)^2} \end{pmatrix},$$

而$f(x)=a_1+a_2x+\cdots+a_{n-1}x^{n-2}+a_nx^{n-1}$.

A9. 设方阵A为分块对角阵

$$A=\operatorname{diag}(\lambda_1E_{n_1},\lambda_2E_{n_2},\ldots,\lambda_tE_{n_t}),$$

其中$\lambda_1,\lambda_2,\ldots,\lambda_t$两两互不相等. 证明: n阶方阵B与A可交换的必要且充分条件是B也是分块对角方阵, 且有如下的形式: $B=\operatorname{diag}(B_1,B_2,\cdots,B_t)$, 其中, B_i是n_i阶方阵, $1\leqslant i\leqslant t$.

B1. 设有非零矩阵$A\in\mathbb{F}^{m\times n}$, 证明: 存在$m$维非零列向量$\alpha$和$n$维非零列向量$\beta$, 使得$\alpha^TA\beta\neq 0$.

B2. 设$A=(a_{ij})\in\mathbb{F}^{m\times n}$, $B=(b_{ij})\in\mathbb{F}^{s\times t}$, 定义$A$和$B$的张量积(**Kronecker积**)为

$$A\otimes B=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}\in\mathbb{F}^{ms\times nt}.$$

证明(假设所写的运算都可进行):

(1) $(A\otimes B)^T=A^T\otimes B^T$;

(2) $A\otimes(B+C)=A\otimes B+A\otimes C$;

(3) $(B+C)\otimes A=B\otimes A+C\otimes A$;

(4) $A\otimes(aB)=(aA)\otimes B=a(A\otimes B)$, $a\in\mathbb{F}$;

(5) $E_m\otimes E_n=E_{mn}$;

(6) $(A\otimes B)\otimes C=A\otimes(B\otimes C)$;

(7) $(AC)\otimes(BD)=(A\otimes B)(C\otimes D)$;

(8) 如果A和B都是可逆矩阵, 则$A\otimes B$也可逆, 且$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$.

§ 3.4 初等阵与初等变换

下面我们尝试用初等变换来研究方阵的可逆性. 矩阵可逆性是从矩阵的乘法导出的概念, 所以为了用初等变换来研究矩阵的可逆逆, 我们需要用矩阵的乘法来实现初等变换, 而这只需要用矩阵乘法实现三类初等变换. 具体地, 对于初等变换$r_i \leftrightarrow r_j$, 我们希望找矩阵P, 使得

$$A \overset{r_i\leftrightarrow r_j}{\rightsquigarrow} B = PA(\text{或者}AP), \quad \forall A.$$

如何找P呢? 假设这样的P存在, 上面取A为单位阵, 则

$$E \overset{r_i\leftrightarrow r_j}{\rightsquigarrow} B = PE(\text{或者}EP) = P,$$

即P一定为对单位阵进行该类初等变换得到的矩阵. 我们称这样得到的方阵为初等阵, 下面我们先定义它们.

3.4.1 初等阵

对n阶单位阵E_n施行一次初等变换得到的矩阵称为n阶**初等阵**. 对应到三类初等变换, 我们一共有三类初等阵:

第I类

$$E_n \underset{(c_i\leftrightarrow c_j)}{\overset{r_i\leftrightarrow r_j}{\rightsquigarrow}} P_{ij} = \begin{array}{cc} & \begin{array}{ccccccccccc} & & & i\text{列} & & & & j\text{列} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i\text{行} \\ \\ \\ \\ j\text{行} \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & \vdots & & & & \vdots & & & \\ \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & 1 & & & \\ & & & \vdots & 1 & & & \vdots & & & \\ & & & \vdots & & \ddots & & \vdots & & & \\ & & & \vdots & & & 1 & \vdots & & & \\ \cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} \end{array}, \quad (i<j);$$

第II类

$$E_n \underset{(a\times c_i)}{\overset{a\times r_i}{\rightsquigarrow}} P_i(a) = \begin{array}{cc} & \begin{array}{ccccccc} & & & i\text{列} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i\text{行} \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \vdots & & & \\ \cdots & \cdots & \cdots & a & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}, \quad (a \neq 0);$$

第III类

$$E_n \underset{(c_i+ac_j)}{\overset{r_j+ar_i}{\sim}} P(i(a),j) = \begin{array}{cc} & \begin{matrix} & & & & i\text{列} & & j\text{列} & & & \end{matrix} \\ \begin{matrix} \\ \\ \\ i\text{行} \\ \\ j\text{行} \\ \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & \vdots & & \vdots & & & \\ & & & \cdots & 1 & & \vdots & & & \\ & & & & \vdots & \ddots & \vdots & & & \\ & & & \cdots & a & \cdots & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix} \end{array}. \quad (i \neq j).$$

用列分块, 这三类初等阵也可以写成

$$\begin{aligned} P_{ij} &= (e_1, \ldots, e_{i-1}, e_j, e_{i+1}, \ldots, e_{j-1}, e_i, e_{j+1}, \ldots, e_n), \quad (i<j), \\ P_i(a) &= (e_1, \ldots, e_{i-1}, ae_i, e_{i+1}, \ldots, e_n), \\ P(i(a),j) &= (e_1, \ldots, e_{i-1}, e_i + ae_j, e_{i+1}, \ldots, e_n), \quad (i \neq j). \end{aligned}$$

还可以用另一种方式表示初等阵. 设

$$E_{ij} = \begin{array}{cc} \begin{matrix} i\text{行} \\ \\ \\ \\ \end{matrix} & \begin{pmatrix} & \cdots & 1 & \\ & & \vdots & \\ & & & \\ & & & \end{pmatrix} \\ & \begin{matrix} & & j\text{列} & \end{matrix} \end{array} \in \mathbb{F}^{n \times n}$$

是基本矩阵, 则有

$$\begin{aligned} P_{ij} &= E_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}, \quad (i<j), \\ P_i(a) &= E_n + (a-1)E_{ii}, \\ P(i(a),j) &= E_n + aE_{ji}, \quad (i \neq j). \end{aligned}$$

从定义可以看出, P_{ij}和$P_i(a)$是对称阵, 即

$$P_{ij}^T = P_{ij}, \quad P_i(a)^T = P_i(a).$$

而

$$P(i(a),j)^T = (E_n + aE_{ji})^T = E_n + aE_{ij} = P(j(a),i).$$

例3.23. 设$n=3$, 则

$$P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P(1(a),2) = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3.4.2 初等变换和初等阵

下面我们看是否初等阵满足我们的需要, 当然还是先看一些例子. 设$A=(a_{ij})\in\mathbb{F}^{3\times 3}$, 我们有下面的观察:

(1) $A \overset{r_1\leftrightarrow r_2}{\rightsquigarrow} \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = P_{12}A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$;

(2) $A \overset{a\times r_2}{\rightsquigarrow} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ aa_{21} & aa_{22} & aa_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = P_2(a)A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$;

(3) $A \overset{r_2+ar_1}{\rightsquigarrow} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}+aa_{11} & a_{22}+aa_{12} & a_{23}+aa_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$= P(1(a),2)A = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

也就是对A施行这三个行初等变换, 都相当于用相应的初等方阵左乘A. 类似可以得到, 对A施行三类列初等变换, 都相当于用相应的初等方阵右乘A. 于是我们有理由猜想这个是一般规律, 经检验确实得到下面漂亮的结果.

定理3.7. 设$A\in\mathbb{F}^{m\times n}$, 则

(1) *对A施行一次行初等变换, 相当于用相应的m阶初等阵左乘A;*

(2) *对A施行一次列初等变换, 相当于用相应的n阶初等阵右乘A.*

☞ **证明.** (2) 设有列分块$A=(\beta_1,\beta_2,\dots,\beta_n)$. 注意到$\beta_j=Ae_j$, 于是
(i) 当$i<j$时有

$$\begin{aligned} AP_{ij} &= A(e_1,\dots,e_{i-1},e_j,e_{i+1},\dots,e_{j-1},e_i,e_{j+1},\dots,e_n) \\ &= (Ae_1,\dots,Ae_{i-1},Ae_j,Ae_{i+1},\dots,Ae_{j-1},Ae_i,Ae_{j+1},\dots,Ae_n) \\ &= (\beta_1,\dots,\beta_{i-1},\beta_j,\beta_{i+1},\dots,\beta_{j-1},\beta_i,\beta_{j+1},\dots,\beta_n), \end{aligned}$$

最后一个矩阵也可以对A施行$c_i\leftrightarrow c_j$得到;
(ii) 当$a\in\mathbb{F}$时有

$$\begin{aligned} AP_i(a) &= A(e_1,\dots,e_{i-1},ae_i,e_{i+1},\dots,e_n) \\ &= (Ae_1,\dots,Ae_{i-1},A(ae_i),Ae_{i+1},\dots,Ae_n) \\ &= (\beta_1,\dots,\beta_{i-1},aAe_i,\beta_{i+1},\dots,\beta_n) \\ &= (\beta_1,\dots,\beta_{i-1},a\beta_i,\beta_{i+1},\dots,\beta_n), \end{aligned}$$

最后一个矩阵也可以对A施行$a\times c_i$得到;

(iii) 当$i\neq j$时有

$$\begin{aligned}AP(i(a),j)=&A(e_1,\ldots,e_{i-1},e_i+ae_j,e_{i+1},\ldots,e_n)\\=&(Ae_1,\ldots,Ae_{i-1},A(e_i+ae_j),Ae_{i+1},\ldots,Ae_n)\\=&(\beta_1,\ldots,\beta_{i-1},Ae_i+aAe_j,\beta_{i+1},\ldots,\beta_n)\\=&(\beta_1,\ldots,\beta_{i-1},\beta_i+a\beta_j,\beta_{i+1},\ldots,\beta_n),\end{aligned}$$

最后一个矩阵也可以对A施行c_i+ac_j得到.

(1) 可以类似于(2)证明, 也可以通过取转置, 利用(2)的结论证明. 为了体会一下转置在处理矩阵问题中的作用, 我们给出第二种证明方法. 设A有行分块

$$A=\begin{pmatrix}\alpha_1^T\\\alpha_2^T\\\vdots\\\alpha_m^T\end{pmatrix},$$

则有$A^T=(\alpha_1,\alpha_2,\ldots,\alpha_m)$.

(i) 对于$i<j$和m阶初等阵P_{ij}, 我们有

$$A^TP_{ij}\xlongequal{(2)}(\alpha_1,\ldots,\alpha_{i-1},\alpha_j,\alpha_{i+1},\ldots,\alpha_{j-1},\alpha_i,\alpha_{j+1},\ldots,\alpha_m).$$

上面等式两边同时取转置, 就得到

$$P_{ij}A=\begin{pmatrix}\alpha_1^T\\\vdots\\\alpha_{i-1}^T\\\alpha_j^T\\\alpha_{i+1}^T\\\vdots\\\alpha_{j-1}^T\\\alpha_i^T\\\alpha_{j+1}^T\\\vdots\\\alpha_m^T\end{pmatrix},$$

而右边的矩阵可以通过对A施行$r_i\leftrightarrow r_j$得到.

(ii) 对于m阶初等阵$P_i(a)$, 我们有

$$A^TP_i(a)\xlongequal{(2)}(\alpha_1,\ldots,\alpha_{i-1},a\alpha_i,\alpha_{i+1},\ldots,\alpha_m).$$

上面等式两边同时取转置, 就得到

$$P_i(a)A=\begin{pmatrix}\alpha_1^T\\\vdots\\\alpha_{i-1}^T\\a\alpha_i^T\\\alpha_{i+1}^T\\\vdots\\\alpha_m^T\end{pmatrix},$$

而右边的矩阵可以通过对A施行$a \times r_i$得到.
(iii) 对于$i \neq j$和m阶初等阵$P(j(a), i)$, 我们有

$$A^T P(j(a), i) \xlongequal{(2)} (\alpha_1, \ldots, \alpha_{j-1}, \alpha_j + a\alpha_i, \alpha_{j+1}, \ldots, \alpha_m).$$

上面等式两边同时取转置, 就得到

$$P(i(a), j)A = \begin{pmatrix} \alpha_1^T \\ \vdots \\ \alpha_{j-1}^T \\ \alpha_j^T + a\alpha_i^T \\ \alpha_{j+1}^T \\ \vdots \\ \alpha_m^T \end{pmatrix},$$

而右边的矩阵可以通过对A施行$r_j + ar_i$得到. □

例3.24. 求3阶方阵Q, 使它左乘$A_{3\times 4}$相当于对A连续施行下面两个行初等变换:

(1) 用$-b$乘A的第1行加到第3行上;

(2) 再调换第2, 3行.

解. 行初等变换(1)即$r_3 - br_1$, 对应的初等阵是$P(1(-b), 3)$; 变换(2)是$r_2 \leftrightarrow r_3$, 对应初等阵P_{23}. 于是

$$Q = P_{23}P(1(-b), 3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -b & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

也可以如下求Q. 由于$Q = P_{23}P(1(-b), 3) = P_{23}P(1(-b), 3)E_3$, 所以

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{r_3 - br_1}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b & 0 & 1 \end{pmatrix} \overset{r_2 \leftrightarrow r_3}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -b & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = Q.$$

这个方法的好处是不必记忆初等阵的具体形式, 也不必进行矩阵乘法. □

同上例一样, 一般地, 对于$A, B \in \mathbb{F}^{m\times n}$, 如果$A \overset{r}{\sim} B$, 要找方阵$Q$, 使得$QA = B$, 那么由

$$E_m \underset{\text{变换}}{\overset{\text{同样行初等}}{\sim}} Q,$$

就可以得到Q.

习题3.4

A1. 设A是$3 \times n$矩阵, 求矩阵P, 使得P左乘A相当于对A依次施行如下的行初等变换:
(1) 交换第2, 3行;
(2) 第2行的-2倍加到第1行;
(3) 第3行乘-3.

A2. 证明: 调行变换可以通过一些行数乘变换和行消去变换实现, 进而初等矩阵都可以表示成形如$E + a_{ij}E_{ij}$这样矩阵的乘积.

§ 3.5　可逆矩阵求逆

3.5.1　用行初等变换判别可逆性

下面我们用初等变换来研究矩阵可逆. 用行初等变换可以把矩阵化为简化阶梯形阵, 如果每一步的初等变换都不改变矩阵的可逆性, 那么我们只要看简化阶梯形是否可逆就可以了. 这个思考过程具体如下图:

$$\boxed{\text{方阵}} \xrightarrow{\text{行初等变换}} \boxed{\text{简化阶梯形阵}}$$

$$\uparrow \qquad\qquad \uparrow$$

是否改变可逆性?　何时可逆?

首先说明(行)初等变换不改变矩阵的可逆性, 由于初等变换相当于用初等阵做乘法, 所以需要考虑初等阵的可逆性.

引理3.8. (1) *初等阵都可逆, 且逆为同类型初等阵.*

(2) *设对方阵A施行一次行初等变换得到B, 则*

$$A\text{ 可逆} \Longleftrightarrow B\text{ 可逆}.$$

证明. (1) 我们有

$$P_{ij}P_{ij} = E_n \Longrightarrow P_{ij}^{-1} = P_{ij},$$
$$P_i(a)P_i\left(\frac{1}{a}\right) = P_i\left(\frac{1}{a}\right)P_i(a) = E_n \Longrightarrow P_i(a)^{-1} = P_i\left(\frac{1}{a}\right), \quad (a \neq 0),$$
$$P(i(a),j)P(i(-a),j) = P(i(-a),j)P(i(a),j) = E_n \Longrightarrow P(i(a),j)^{-1} = P(i(-a),j).$$

事实上, 这等同于(线性方程组的)初等变换的可逆性.
(2) 由条件, 存在初等阵P, 使得$PA = B$. 由(1), P可逆, 于是有$A = P^{-1}B$. 而可逆矩阵的乘积仍可逆, 所以得证. □

其次, 对于简化阶梯形阵的可逆性, 我们有下面的结果.

引理3.9. *设$R \in \mathbb{F}^{n\times n}$是简化阶梯形阵, 则下面命题等价*

(i) *R可逆;*

(ii) *R没有全零行;*

(iii) *$R = E_n$.*

证明. "(i) $\Longrightarrow$ (ii)" 如果R有全零行, 则R不可逆.
"(ii) $\Longrightarrow$ (iii)" 如果R无全零行, 则R有n个首元. 设第i个首元列为第j_i列, 则

$$1 \leqslant j_1 < j_2 < \cdots < j_n \leqslant n.$$

于是有

$$j_i = i, \quad i = 1, 2, \ldots, n,$$

即R的n个对角元都是首元. 进而$R = E_n$.
"(iii) $\Longrightarrow$ (i)" 显然. □

最后, 我们就得到用初等变换判别方阵的可逆性的定理[8].

定理3.10. 设$A \in \mathbb{F}^{n\times n}$, R是与A行等价的阶梯形阵, 则下面命题等价:

(i) A可逆;

(ii) R没有全零行;

(iii) $A \overset{r}{\sim} E_n$;

(iv) A是一些n阶初等阵的乘积.

证明. "(i) $\Longrightarrow$ (ii)" 由引理3.8, A可逆当且仅当R可逆.
"(ii) $\Longrightarrow$ (iii)" 如果R没有全零行, 则与A行等价的简化阶梯形阵R_1也没有全零行. 由引理3.9, 一定有$R_1 = E_n$.
"(iii) $\Longrightarrow$ (iv)" 如果$A \overset{r}{\sim} E_n$, 则$E_n \overset{r}{\sim} A$. 所以由定理3.7, 存在初等阵$P_1, P_2, \ldots, P_s$, 使得

$$A = P_s \cdots P_2 P_1 E_n = P_s \cdots P_2 P_1.$$

"(iv) $\Longrightarrow$ (i)" 初等阵都可逆, 而可逆阵的乘积仍可逆. □

定理3.10中的判别法则(iv)与左, 右乘无关, 于是我们也可得到

$$A\text{可逆} \Longleftrightarrow A \overset{c}{\sim} E_n \Longleftrightarrow A \sim E_n.$$

事实上, 如果A可逆, 则A是一些初等阵的乘积

$$A = P_1 \cdots P_s = E_n P_1 \cdots P_s.$$

于是对E_n依次施行$P_1, \ldots, P_s$所对应的列初等变换, 可以得到A. 即有$E_n \overset{c}{\sim} A$, 也就是$A \overset{c}{\sim} E_n$. 反之, 如果$A \overset{c}{\sim} E_n$, 则$E_n \overset{c}{\sim} A$. 所以存在初等阵$P_1, \ldots, P_s$, 使得

$$E_n P_1 \cdots P_s = A,$$

即得到A是一些初等阵的乘积.

例3.25. 判别矩阵$A = \begin{pmatrix} 2 & -3 & 1 \\ 2 & 3 & -2 \\ 4 & 0 & -1 \end{pmatrix}$的可逆性.

解. 由于

$$A \underset{r_3-2r_1}{\overset{r_2-r_1}{\sim}} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 6 & -3 \\ 0 & 6 & -3 \end{pmatrix} \overset{r_3-r_2}{\sim} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 6 & -3 \\ 0 & 0 & 0 \end{pmatrix},$$

所以A不可逆. □

[8]这里的处理方法和用Gauss消元法解线性方程组是类似的: 我们先考虑简单的情形—(简化)阶梯形阵, 再考虑(行)初等变换是否会保持我们研究的对象(解, 可逆性). 我们称这种方法为"标准形"方法, 后面考虑矩阵秩时还要用到.

- **应用：矩阵可逆定义的简化**

我们知道同阶可逆阵之积仍可逆, 那么如果其中一个因子不可逆呢? 类比到数, 数域中不可逆的数就是零, 而零与任意数之积仍为零, 不可逆. 对于矩阵呢? 不可逆阵本质上可以认为是最后一行为全零行的阶梯形阵, 将它左乘同阶方阵, 所得矩阵的最后一行当然也为全零行, 不可逆. 所以有理由猜想不可逆阵与任意同阶方阵之积仍不可逆, 这正是下面要证明的.

推论3.11. 设$A, B \in \mathbb{F}^{n\times n}$.

(1) *如果A不可逆, 则AB和BA都不可逆;*

(2) *AB可逆$\Longleftrightarrow$ A和B都可逆.*

☞ **证明.** (1) 假设AB可逆. 设R是与A行等价的阶梯形阵, 由于A不可逆, 则R的最后一行是全零行. 由定理3.7, 存在n阶初等阵$P_1, P_2, \ldots, P_s$, 使得

$$P_s \ldots P_2 P_1 A = R.$$

所以

$$RB = (P_s \ldots P_2 P_1 A)B = P_s \ldots P_2 P_1 (AB).$$

于是RB是一些初等阵和可逆阵AB之积, 必可逆. 而R的最后一行为全零行, 可得RB的最后一行也为全零行, 所以RB不可逆. 矛盾, 于是AB不可逆.

又由于A^T不可逆, 所以由已证得$A^T B^T$不可逆, 进而$(A^T B^T)^T$不可逆, 即BA不可逆.

(2) 必要性由(1), 充分性已知. □

在矩阵可逆定义中, 我们要求$AB = BA = E$, 即等式$AB = E$和$BA = E$需要同时成立. 但是如果$AB = E$或者$BA = E$, 则利用上面的推论就得到A可逆. 所以要保证A的可逆性, 只需要其中某一个等式成立即可[9].

推论3.12. 设$A \in \mathbb{F}^{n\times n}$, *下列命题等价*

(i) *A可逆;*

(ii) *存在$B \in \mathbb{F}^{n\times n}$, 使得$AB = E_n$;*

(iii) *存在$B \in \mathbb{F}^{n\times n}$, 使得$BA = E_n$.*

当(ii)或者(iii)成立时, 有$A^{-1} = B$.

☞ **证明.** 我们已经证明了(i), (ii)和(iii)的等价性. 假设(ii)成立, 则

$$A^{-1} = A^{-1}E = A^{-1}(AB) = (A^{-1}A)B = EB = B.$$

类似可得(iii)成立时有$B = A^{-1}$. □

- **应用：初等变换基本定理**

由可逆阵与初等阵的关系, 以及初等变换和初等阵的关系, 我们可以得到下面的初等变换基本定理.

[9]利用可逆的行列式判别法则, 可以给出这一事实的更简单的证明.

定理3.13. 设$A, B \in \mathbb{F}^{m\times n}$, 则

(1) $A \overset{r}{\sim} B \iff$ 存在m阶可逆阵P, 使得$PA = B$;

(2) $A \overset{c}{\sim} B \iff$ 存在n阶可逆阵Q, 使得$AQ = B$;

(3) $A \sim B \iff$ 存在m阶可逆阵P和n阶可逆阵Q, 使得$PAQ = B$.

证明. 我们证明(1), 可以类似证明(2)和(3). 如果$A \overset{r}{\sim} B$, 则存在初等阵$P_1, P_2, \ldots, P_s$, 使得$P_s \cdots P_2 P_1 A = B$. 令$P = P_s \cdots P_2 P_1$, 则P可逆.

反之, 如果$B = PA$, 其中P可逆. 则P为初等阵之积: $P = P_s \cdots P_2 P_1$, 即有$P_s \cdots P_2 P_1 A = B$. 而用初等阵左乘相当于做行初等变换, 所以$A \overset{r}{\sim} B$. □

推论3.14. 设$A \in \mathbb{F}^{m\times n}$, 则

(1) 存在m阶可逆阵P, 使得PA为简化阶梯形阵;

(2) 存在整数r, $0 \leqslant r \leqslant \min\{m, n\}$, 使得$A \sim \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$. 进而存在$m$阶可逆阵$P$和$n$阶可逆阵$Q$, 使得

$$PAQ = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}.$$

定义3.9. 称推论3.14中的矩阵$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$为矩阵$A$的初等变换标准形[10].

要注意的是, 初等变换标准形也可能为(E_r, O), $\begin{pmatrix} E_r \\ O \end{pmatrix}$或者$E_r$. 这里的$r$由$A$唯一确定[11].

推论3.14的证明. (2) 设R是与A行等价的简化阶梯形阵, 且R有r个非零行. 用调列变换可以依次将第i个首元列换到第i列, $i = 1, 2, \ldots, r$, 即

$$R \overset{c}{\sim} \begin{pmatrix} E_r & * \\ O & O \end{pmatrix}.$$

再用列消去变换将前r行的后$n-r$列的非零元化为零即可. □

例3.26. 求矩阵$A = \begin{pmatrix} 0 & 0 & 0 & 2 & 2 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & 1 & 2 & 4 & 3 \end{pmatrix}$的初等变换标准形$F$.

解. 由例2.5, 我们有

$$A \overset{r}{\sim} \begin{pmatrix} 0 & \boxed{1} & 2 & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = R.$$

[10] 初等变换标准形也称为相抵标准形.

[11] 事实上, r是A的秩. 这里的三种特殊情形对应到A行满秩, 列满秩和可逆.

而

$$R \overset{c}{\sim} \begin{pmatrix} \boxed{1} & 0 & 2 & 0 & -1 \\ 0 & \boxed{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{c}{\sim} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

所以

$$F = \begin{pmatrix} E_2 & O \\ O & O \end{pmatrix}_{3\times 5}.$$

解毕. □

例3.27. *设$A \in \mathbb{F}^{n\times n}$不可逆, 证明: 存在非零矩阵$B, C \in \mathbb{F}^{n\times n}$, 使得$BA = O$, $AC = O$.*

☞ **证明.** 我们用标准形解法[12]. 如果用行等价标准形, 可如下证明. 设R是与A行等价的简化阶梯形阵, 则R的最后一行为全零行. 于是

$$\begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}_{n\times n} R = O.$$

又存在n阶可逆阵P, 使得$PA = R$. 取

$$B = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} P \neq O,$$

则

$$BA = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} PA = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} R = O.$$

又A不可逆蕴含A^T不可逆, 由上面的证明, 存在非零n阶方阵B_1, 使得$B_1A^T = O$. 取转置得$AB_1^T = O$, 所以取$C = B_1^T \neq O$即可.

如果用初等变换标准形, 可如下证明. 存在n阶可逆阵P和Q, 使得

$$PAQ = F = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{n\times n}.$$

如果$r = n$, 则$F = E_n$, 进而$A \sim E_n$, 得A可逆, 矛盾. 所以$r < n$. 注意到

$$\begin{pmatrix} O & O \\ O & E_{n-r} \end{pmatrix}\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} = O, \quad \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}\begin{pmatrix} O & O \\ O & E_{n-r} \end{pmatrix} = O,$$

所以取

$$B = \begin{pmatrix} O & O \\ O & E_{n-r} \end{pmatrix} P \neq O, \quad C = Q\begin{pmatrix} O & O \\ O & E_{n-r} \end{pmatrix} \neq O$$

[12]也可用定理3.15得到C的存在性.

时有

$$\begin{aligned}BA=&\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}PA=\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}(PAQ)Q^{-1}\\ =&\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q^{-1}=OQ^{-1}=O,\\ AC=&AQ\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}=P^{-1}(PAQ)\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}\\ =&P^{-1}\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}\begin{pmatrix}O & O\\ O & E_{n-r}\end{pmatrix}=P^{-1}O=O.\end{aligned}$$

证毕. □

- **应用: 线性方程组解的判定**

设$A\in\mathbb{F}^{m\times n}$, 考虑$n$元齐次线性方程组$AX=0$. 由推论2.5, 当$m<n$时(即方程个数比未知数个数少时)该齐次方程组必有非零解. 那么当$m=n$, 即A是方阵时呢? 答案是与A是否可逆有关.

定理3.15. *设$A\in\mathbb{F}^{n\times n}$, 则线性方程组$AX=\beta$有唯一解的充分必要条件是$A$可逆. 进而, 齐次线性方程组$AX=0$有非零解的充分必要条件是$A$不可逆.*

☞ **证明.** 如果A可逆, 则$AX=\beta$导出$X=A^{-1}\beta$, 有唯一解. 反之, 如果$AX=\beta$有唯一解, 设R是与A行等价的简化阶梯形阵, 则R的所有列为首元列. 因此R无全零行, 这得到$R=E_n$, 进而A可逆. □

3.5.2 用行初等变换求逆

不但可以用初等变换判别方阵是否可逆, 还可以用行初等变换求可逆矩阵的逆. 设$A\in\mathbb{F}^{n\times n}$, 则当$A\overset{r}{\sim}E_n$时$A$可逆. 此时, 存在初等阵$P_1,P_2,\ldots,P_s$, 使得

$$P_s\cdots P_2P_1A=E_n.$$

于是$A^{-1}=P_s\cdots P_2P_1$. 所以需要求$P_s\cdots P_2P_1$, 这可以对E_n进行同样的行初等变换得到.

命题3.16. *设$A\in\mathbb{F}^{n\times n}$可逆, $B\in\mathbb{F}^{n\times m}$, 如果*

$$(A,E_n)\overset{r}{\sim}(E_n,P),\quad (A,B)\overset{r}{\sim}(E_n,X),$$

那么$P=A^{-1}$, 而$X=A^{-1}B$.

☞ **证明.** 只需证明$X=A^{-1}B$. 存在n阶可逆阵P_1, 使得

$$P_1(A,B)=(E_n,X).$$

即

$$(P_1A,P_1B)=(E_n,X).$$

于是

$$P_1A=E_n,\quad P_1B=X.$$

从第一个等式得到$P_1 = A^{-1}$, 代入第二个等式就得$A^{-1}B = X$. □

上面命题给出了计算可逆矩阵A的逆A^{-1}和用A左除矩阵B得到$A^{-1}B$的方法. 那么如何计算用可逆阵右除? 即如何计算BA^{-1}. 一种方法是利用列的初等变换, 我们更推荐另一种方法: 取转置, 将列转化为行. 具体的, 利用

$$(BA^{-1})^T = (A^{-1})^T B^T = (A^T)^{-1} B^T$$

即可.

我们看一些例子.

例3.28. *问矩阵*$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 4 \\ 2 & -1 & 0 \end{pmatrix}$*是否可逆? 当*$A$*可逆时求*$A^{-1}$.

☞ **解.** 我们有

$$\begin{aligned}(A, E_3) = \begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 4 & 0 & 1 & 0 \\ 2 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} &\overset{r_1 \leftrightarrow r_2}{\widetilde{\qquad}} \begin{pmatrix} 1 & 1 & 4 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ &\overset{r_3 - 2r_1}{\widetilde{\qquad}} \begin{pmatrix} 1 & 1 & 4 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -8 & 0 & -2 & 1 \end{pmatrix} \underset{r_3 + 3r_2}{\overset{r_1 - r_2}{\widetilde{\qquad}}} \begin{pmatrix} 1 & 0 & 2 & -1 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -2 & 1 \end{pmatrix} \\ &\underset{r_2 + r_3}{\overset{r_1 + r_3}{\widetilde{\qquad}}} \begin{pmatrix} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 4 & -2 & 1 \\ 0 & 0 & -2 & 3 & -2 & 1 \end{pmatrix} \overset{-\frac{1}{2}r_3}{\widetilde{\qquad}} \begin{pmatrix} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 4 & -2 & 1 \\ 0 & 0 & 1 & -\frac{3}{2} & 1 & -\frac{1}{2} \end{pmatrix},\end{aligned}$$

于是A可逆, 且$A^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 4 & -2 & 1 \\ -\frac{3}{2} & 1 & -\frac{1}{2} \end{pmatrix}$. □

例3.29. *设*$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$*. 问: *$a, b, c, d$*满足何种条件时*$A$*可逆? 并在*$A$*可逆时求*$A$*的逆.*

☞ **解.** 当$a \neq 0$, 我们有

$$(A, E_2) = \begin{pmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{pmatrix} \overset{a \times r_2}{\widetilde{\qquad}} \begin{pmatrix} a & b & 1 & 0 \\ ac & ad & 0 & a \end{pmatrix} \overset{r_2 - cr_1}{\widetilde{\qquad}} \begin{pmatrix} a & b & 1 & 0 \\ 0 & ad - bc & -c & a \end{pmatrix}.$$

所以

$$A \text{可逆} \Longleftrightarrow ad - bc \neq 0.$$

此时有

$$(A, E_2) \overset{r}{\sim} \begin{pmatrix} a & 0 & \frac{ad}{ad-bc} & -\frac{ab}{ad-bc} \\ 0 & ad - bc & -c & a \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix},$$

即得

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

当$a=0$时, 有

$$A=\begin{pmatrix}0&b\\c&d\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}c&d\\0&b\end{pmatrix}.$$

所以

$$A\text{可逆}\iff b\neq 0, c\neq 0.$$

此时有

$$(A,E_2)=\begin{pmatrix}0&b&1&0\\c&d&0&1\end{pmatrix}\underset{\sim}{\overset{r_1\leftrightarrow r_2}{}}\begin{pmatrix}c&d&0&1\\0&b&1&0\end{pmatrix}\underset{\sim}{\overset{r_1-\frac{d}{b}r_2}{}}\begin{pmatrix}c&0&-\frac{d}{b}&1\\0&b&1&0\end{pmatrix}$$
$$\overset{r}{\sim}\begin{pmatrix}1&0&-\frac{d}{bc}&\frac{1}{c}\\0&1&\frac{1}{b}&0\end{pmatrix},$$

即

$$A^{-1}=-\frac{1}{bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}.$$

综上, A可逆的充分必要条件是$ad-bc\neq 0$, 且当A可逆时有

$$A^{-1}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}.$$

此即二阶方阵的求逆公式[13]. □

例3.30. 解矩阵方程[14]

$$\begin{pmatrix}2&5\\1&3\end{pmatrix}X=\begin{pmatrix}4&-6&1\\2&1&-1\end{pmatrix}.$$

☞ **解.** 由于

$$\begin{pmatrix}2&5&4&-6&1\\1&3&2&1&-1\end{pmatrix}\underset{\sim}{\overset{r_1\leftrightarrow r_2}{}}\begin{pmatrix}1&3&2&1&-1\\2&5&4&-6&1\end{pmatrix}$$
$$\underset{\sim}{\overset{r_2-2r_1}{}}\begin{pmatrix}1&3&2&1&-1\\0&-1&0&-8&3\end{pmatrix}\underset{-r_2}{\overset{r_1+3r_2}{\sim}}\begin{pmatrix}1&0&2&-23&8\\0&1&0&8&-3\end{pmatrix},$$

所以

$$X=\begin{pmatrix}2&5\\1&3\end{pmatrix}^{-1}\begin{pmatrix}4&-6&1\\2&1&-1\end{pmatrix}=\begin{pmatrix}2&-23&8\\0&8&-3\end{pmatrix}.$$

也可以用上例所得到的二阶方阵的求逆公式. 因为$2\times 3-5\times 1=1\neq 0$, 所以$\begin{pmatrix}2&5\\1&3\end{pmatrix}$可逆, 且

$$\begin{pmatrix}2&5\\1&3\end{pmatrix}^{-1}=\begin{pmatrix}3&-5\\-1&2\end{pmatrix}.$$

最后得到

$$X=\begin{pmatrix}2&5\\1&3\end{pmatrix}^{-1}\begin{pmatrix}4&-6&1\\2&1&-1\end{pmatrix}=\begin{pmatrix}3&-5\\-1&2\end{pmatrix}\begin{pmatrix}4&-6&1\\2&1&-1\end{pmatrix}=\begin{pmatrix}2&-23&8\\0&8&-3\end{pmatrix}.$$

解毕. □

[13]后面我们将使用伴随矩阵表示的求逆公式重新证明这个结论.
[14]含有未知矩阵的矩阵等式称为矩阵方程.

例3.31. 解矩阵方程

$$X\begin{pmatrix}1&1&-1\\0&2&2\\1&-1&0\end{pmatrix}=\begin{pmatrix}1&-1&1\\1&1&0\\2&1&1\end{pmatrix}.$$

☞ **解.** 方程等价于

$$\begin{pmatrix}1&0&1\\1&2&-1\\-1&2&0\end{pmatrix}X^T=\begin{pmatrix}1&1&2\\-1&1&1\\1&0&1\end{pmatrix}.$$

由于

$$\begin{pmatrix}1&0&1&1&1&2\\1&2&-1&-1&1&1\\-1&2&0&1&0&1\end{pmatrix}\underset{r_3+r_1}{\overset{r_2-r_1}{\sim}}\begin{pmatrix}1&0&1&1&1&2\\0&2&-2&-2&0&-1\\0&2&1&2&1&3\end{pmatrix}$$

$$\overset{r_3-r_2}{\sim}\begin{pmatrix}1&0&1&1&1&2\\0&2&-2&-2&0&-1\\0&0&3&4&1&4\end{pmatrix}\underset{\frac{1}{3}r_3}{\overset{\frac{1}{2}r_2}{\sim}}\begin{pmatrix}1&0&1&1&1&2\\0&1&-1&-1&0&-\frac{1}{2}\\0&0&1&\frac{4}{3}&\frac{1}{3}&\frac{4}{3}\end{pmatrix}$$

$$\underset{r_2+r_3}{\overset{r_1-r_3}{\sim}}\begin{pmatrix}1&0&0&-\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\\0&1&0&\frac{1}{3}&\frac{1}{3}&\frac{5}{6}\\0&0&1&\frac{4}{3}&\frac{1}{3}&\frac{4}{3}\end{pmatrix},$$

所以

$$X^T=\begin{pmatrix}1&0&1\\1&2&-1\\-1&2&0\end{pmatrix}^{-1}\begin{pmatrix}1&1&2\\-1&1&1\\1&0&1\end{pmatrix}=\begin{pmatrix}-\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\\\frac{1}{3}&\frac{1}{3}&\frac{5}{6}\\\frac{4}{3}&\frac{1}{3}&\frac{4}{3}\end{pmatrix}.$$

这得到

$$X=\begin{pmatrix}-\frac{1}{3}&\frac{1}{3}&\frac{4}{3}\\\frac{2}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{2}{3}&\frac{5}{6}&\frac{4}{3}\end{pmatrix},$$

解毕. □

例3.32. 设$A=\begin{pmatrix}1&1&-1\\-1&1&1\\1&-1&1\end{pmatrix}$, 且有矩阵等式$4X=2AX+E$, 求$X$.

☞ **解.** 原等式化为$(4E-2A)X=E$, 即$(2E-A)X=\frac{1}{2}E$. 由于

$$(2E-A,E)=\begin{pmatrix}1&-1&1&1&0&0\\1&1&-1&0&1&0\\-1&1&1&0&0&1\end{pmatrix}\underset{r_3+r_1}{\overset{r_2-r_1}{\sim}}\begin{pmatrix}1&-1&1&1&0&0\\0&2&-2&-1&1&0\\0&0&2&1&0&1\end{pmatrix}$$

$$\underset{\frac{1}{2}r_3}{\overset{\frac{1}{2}r_2}{\sim}}\begin{pmatrix}1&-1&1&1&0&0\\0&1&-1&-\frac{1}{2}&\frac{1}{2}&0\\0&0&1&\frac{1}{2}&0&\frac{1}{2}\end{pmatrix}\underset{r_2+r_3}{\overset{r_1+r_2}{\sim}}\begin{pmatrix}1&0&0&\frac{1}{2}&\frac{1}{2}&0\\0&1&0&0&\frac{1}{2}&\frac{1}{2}\\0&0&1&\frac{1}{2}&0&\frac{1}{2}\end{pmatrix},$$

所以$2E-A$可逆, 且

$$X=\frac{1}{2}(2E-A)^{-1}=\frac{1}{4}\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}.$$

解毕. □

习题3.5

A1. 下列方阵是否可逆? 可逆时求它的逆.

(1) $\begin{pmatrix}2&3\\4&5\end{pmatrix}$; (2) $\begin{pmatrix}1&1&-1\\2&1&0\\1&-1&0\end{pmatrix}$;

(3) $\begin{pmatrix}3&-1&2\\1&0&1\\-1&1&0\end{pmatrix}$; (4) $\begin{pmatrix}1&1&\cdots&1\\&1&\cdots&1\\&&\ddots&\vdots\\&&&1\end{pmatrix}_{n\times n}$.

A2. 求矩阵$A=\begin{pmatrix}1&2&3&3\\2&3&-1&-2\\0&1&3&4\\-2&-1&1&4\end{pmatrix}$的行初等变换标准形和初等变换标准形.

A3. 求矩阵X, 使得

(1) $\begin{pmatrix}4&3\\6&5\end{pmatrix}X=\begin{pmatrix}1&2&3\\6&5&4\end{pmatrix}$;

(2) $X\begin{pmatrix}-3&1&-1\\1&-2&-1\\-1&1&0\end{pmatrix}=\begin{pmatrix}1&1&1\\2&0&-1\\-1&1&2\end{pmatrix}$.

A4. 设矩阵A,B,C满足等式$2A^{-1}C+B=C$, 其中$A=\begin{pmatrix}1&1&-3\\0&1&1\\1&2&1\end{pmatrix}$, $B=\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}$, 求矩阵$C$.

B1. 求下列矩阵的逆矩阵:

(1) $\begin{pmatrix}2&&&\\1&2&&\\&\ddots&\ddots&\\&&1&2\end{pmatrix}_{n\times n}$; (2) $\begin{pmatrix}0&1&\cdots&1&1\\1&0&\cdots&1&1\\\vdots&\vdots&\ddots&\vdots&\vdots\\1&1&\cdots&0&1\\1&1&\cdots&1&0\end{pmatrix}_{n\times n}$, $(n\geqslant 2)$;

(3) $\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}_{n\times n}$, $(n \geqslant 2)$;

(4) $\begin{pmatrix} 1 & b & b^2 & \cdots & b^{n-2} & b^{n-1} \\ 0 & 1 & b & \cdots & b^{n-3} & b^{n-2} \\ 0 & 0 & 1 & \cdots & b^{n-4} & b^{n-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$; (5) $\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ & 1 & 2 & \cdots & n-1 \\ & & \ddots & & \vdots \\ & & & 1 & 2 \\ & & & & 1 \end{pmatrix}$;

(6) $\begin{pmatrix} 1 & a & a & \cdots & a \\ a & 1 & a & \cdots & a \\ a & a & 1 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & 1 \end{pmatrix}_{n\times n}$, $(a \neq 1, a \neq \frac{1}{1-n})$;

(7) $\begin{pmatrix} 1+a_1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a_n \end{pmatrix}$, $(a_1a_2\cdots a_n \neq 0$, 且 $\frac{1}{a_1}+\cdots+\frac{1}{a_n} \neq -1)$;

(8) $\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ n & 1 & \cdots & n-2 & n-1 \\ \vdots & \vdots & & \vdots & \vdots \\ 3 & 4 & \cdots & 1 & 2 \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$;

(9) $\begin{pmatrix} a & a+1 & \cdots & a+(n-2) & a+(n-1) \\ a+(n-1) & a & \cdots & a+(n-3) & a+(n-2) \\ \vdots & \vdots & & \vdots & \vdots \\ a+2 & a+3 & \cdots & a & a+1 \\ a+1 & a+2 & \cdots & a+(n-1) & a \end{pmatrix}$, $(a \neq \frac{1-n}{2})$;

(10) $$\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}_{n\times n}, (n\geqslant 2);$$

(11) $$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta & \zeta^2 & \cdots & \zeta^{n-1} \\ 1 & \zeta^2 & \zeta^4 & \cdots & \zeta^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \zeta^{n-1} & \zeta^{2(n-1)} & \cdots & \zeta^{(n-1)^2} \end{pmatrix}, \zeta = e^{\frac{2\pi i}{n}}.$$

B2. 设A为n阶方阵, 证明: 存在n阶方阵B和C, 使得B可逆, C满足$C^2=C$, 且$A=BC$.

§ 3.6 分块矩阵求逆

矩阵分块是一种重要的方法和手段, 前面介绍了分块矩阵如何进行加法和乘法, 那么如何求分块矩阵的逆呢? 上节我们用矩阵的初等变换来求逆, 类比于此, 我们首先要定义分块矩阵的初等变换, 然后以其为工具来求分块矩阵的逆.

3.6.1 分块矩阵的初等变换

定义3.10. 称下面三类变换为分块矩阵的行(列)初等变换:

(i) 调行(列)变换: $r_i \leftrightarrow r_j(c_i \leftrightarrow c_j)$(交换分块矩阵第$i$行(列)和第$j$行(列)对应位置的子阵);

(ii) 行(列)块乘变换: $P\times r_i(c_i\times P)$(用可逆矩阵P左(右)乘第i行的所有子阵);

(iii) 行(列)消去变换: $r_j+M\times r_i(c_j+c_i\times M)$(用矩阵$M$左(右)乘第$i$行(列)所有子阵, 再加到第$j$行(列)的对应位置子阵上).

分块矩阵的行和列初等变换统称为分块矩阵的初等变换.

如果分块阵A经过有限次分块矩阵的行(列)初等变换变为B, 我们也记为$A\overset{r}{\sim}B$ ($A\overset{c}{\sim}B$). 类似于初等矩阵, 单位阵进行行型和列型一样的分块后, 经过一次分块矩阵的初等变换得到的分块矩阵称为初等分块阵. 对应到三类分块矩阵的初等变换, 我们一共有三类初等分块阵: 设单位阵有如下分块

$$E=\mathrm{diag}(E_{k_1},E_{k_2},\ldots,E_{k_t}),$$

则

第I类 对应到$r_i \leftrightarrow r_j$和$c_i \leftrightarrow c_j$ $(i<j)$

$$P_{ij}=\begin{array}{cc}
 & \begin{array}{ccccccccccc} &&& i\text{列} &&&& j\text{列} &&& \end{array}\\
\begin{array}{c} \\ \\ \\ i\text{行} \\ \\ \\ \\ j\text{行} \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
E_{k_1} &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& E_{k_{i-1}} & \vdots &&&& \vdots &&& \\
\cdots & \cdots & \cdots & O & \cdots & \cdots & \cdots & E_{k_j} &&& \\
&&& \vdots & E_{k_{i+1}} &&& \vdots &&& \\
&&& \vdots && \ddots && \vdots &&& \\
&&& \vdots &&& E_{k_{j-1}} & \vdots &&& \\
\cdots & \cdots & \cdots & E_{k_i} & \cdots & \cdots & \cdots & O &&& \\
&&&&&&&& E_{k_{j+1}} && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& E_{k_t}
\end{pmatrix}
\end{array};$$

第II类 对应到$P\times r_i$和$c_i\times P$ (P可逆)

$$P_i(P)=\begin{array}{cc}
 & \begin{array}{ccccccc} &&& i\text{列} &&& \end{array}\\
\begin{array}{c} \\ \\ \\ i\text{行} \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
E_{k_1} &&&&&& \\
& \ddots &&&&& \\
&& E_{k_{i-1}} & \vdots &&& \\
\cdots & \cdots & \cdots & P &&& \\
&&&& E_{k_{i+1}} && \\
&&&&& \ddots & \\
&&&&&& E_{k_t}
\end{pmatrix}
\end{array};$$

第III类 对应到$r_j+M\times r_i$和$c_i+c_j\times M$ $(i\neq j)$

$$P(i(M),j)=\begin{array}{cc}
 & \begin{array}{ccccccccc} &&& i\text{列} && j\text{列} &&& \end{array}\\
\begin{array}{c} \\ \\ \\ i\ \text{行} \\ \\ j\ \text{行} \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
E_{k_1} &&&&&&&& \\
& \ddots &&&&&&& \\
&& E_{k_{i-1}} & \vdots && \vdots &&& \\
\cdots & \cdots && E_{k_i} && \vdots &&& \\
&&& \vdots & \ddots & \vdots &&& \\
\cdots & \cdots && M & \cdots & E_{k_j} &&& \\
&&&&&& E_{k_{j+1}} && \\
&&&&&&& \ddots & \\
&&&&&&&& E_{k_t}
\end{pmatrix}
\end{array}.$$

类似于初等阵和初等变换的关系, 进行分块阵的初等变换相当于用对应的初等分块阵左或者右乘该矩阵. 我们先看一个例子, 请读者体会何谓对应的初等分块阵.

例3.33. 设$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{pmatrix}=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}$, 其中$A_{11}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\end{pmatrix}$, 于是分块阵$A$的行型为$(2,1)$, 列型为$(3,1)$. 当考虑对$A$进行分块阵的行初等变换时, 要考虑$E_3$及对应到$A$的行型的分块

$$E_3=\begin{pmatrix}E_2&O\\O&E_1\end{pmatrix}.$$

我们有

$$A\underset{\frown}{\overset{r_1\leftrightarrow r_2}{}}\begin{pmatrix}A_{21}&A_{22}\\A_{11}&A_{12}\end{pmatrix}=\begin{pmatrix}O&E_1\\E_2&O\end{pmatrix}\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}.$$

而当考虑对A进行分块阵的列初等变换时, 要考虑E_4及对应到A的列型的分块

$$E_4=\begin{pmatrix}E_3&O\\O&E_1\end{pmatrix}.$$

我们有

$$A\underset{\frown}{\overset{c_1\leftrightarrow c_2}{}}\begin{pmatrix}A_{12}&A_{11}\\A_{22}&A_{21}\end{pmatrix}=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}\begin{pmatrix}O&E_3\\E_1&O\end{pmatrix}.$$

一般地, 我们有下面的结果.

命题3.17. (1) 对分块矩阵A施行一次分块矩阵的行(列)初等变换相当于用对应的初等分块阵左(右)乘A;

(2) 初等分块阵都可逆, 且逆也为同类型的初等分块阵.

☞ **证明.** (1) 类似于定理3.7的证明.
(2) 类似于引理3.8 (1)的证明. □

下面的例子是我们常常用到的.

例3.34. 设$S=\begin{pmatrix}A&B\\C&D\end{pmatrix}$, 其中$A$为可逆方阵. 则我们有如下的分块矩阵的初等变换

$$S\underset{\frown}{\overset{r_2-CA^{-1}r_1}{}}\begin{pmatrix}A&B\\O&D-CA^{-1}B\end{pmatrix},\tag{3.2}$$

$$S\underset{\frown}{\overset{c_2-c_1A^{-1}B}{}}\begin{pmatrix}A&O\\C&D-CA^{-1}B\end{pmatrix},\tag{3.3}$$

$$S\underset{\frown}{\overset{r_2-CA^{-1}r_1}{}}\begin{pmatrix}A&B\\O&D-CA^{-1}B\end{pmatrix}\underset{\frown}{\overset{c_2-c_1A^{-1}B}{}}\begin{pmatrix}A&O\\O&D-CA^{-1}B\end{pmatrix}.\tag{3.4}$$

如果对应到初等分块阵的左(右)乘, 则(3.2)对应的矩阵等式为

$$\begin{pmatrix}E&O\\-CA^{-1}&E\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A&B\\O&D-CA^{-1}B\end{pmatrix};$$

而(3.3)对应到
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} E & -A^{-1}B \\ O & E \end{pmatrix} = \begin{pmatrix} A & O \\ C & D-CA^{-1}B \end{pmatrix};$$
最后, (3.4)对应到
$$\begin{pmatrix} E & O \\ -CA^{-1} & E \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} E & -A^{-1}B \\ O & E \end{pmatrix} = \begin{pmatrix} A & O \\ O & D-CA^{-1}B \end{pmatrix}.$$
上面三个矩阵等式均称之为**Schur公式**.

3.6.2 分块矩阵求逆

类似于可以用行初等变换求矩阵的逆, 我们也可以用分块矩阵的行初等变换求分块矩阵的逆.

命题3.18. 设$A \in \mathbb{F}^{n\times n}$是分块矩阵, 如果矩阵$(A, E_n)$可经分块矩阵的行初等变换变为$(E_n, P)$, 则$A$可逆, 且$A^{-1} = P$.

☞ **证明.** 类似于命题3.16的证明. □

例3.35. 设$D = \begin{pmatrix} A & C \\ O & B \end{pmatrix}$, 其中$A$和$B$是可逆方阵. 证明: D可逆, 并求D^{-1}.

☞ **解.** 设A为k阶方阵, B为r阶方阵, 我们有分块矩阵的行初等变换
$$(D, E) = \begin{pmatrix} A_k & C & E_k & O \\ O & B_r & O & E_r \end{pmatrix} \underset{}{\overset{r_1 - CB^{-1}r_2}{\sim}} \begin{pmatrix} A_k & O & E_k & -CB^{-1} \\ O & B_r & O & E_r \end{pmatrix}$$
$$\underset{B^{-1}r_2}{\overset{A^{-1}r_1}{\sim}} \begin{pmatrix} E_k & O & A^{-1} & -A^{-1}CB^{-1} \\ O & E_r & O & B^{-1} \end{pmatrix}.$$
于是D可逆, 且
$$D^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ O & B^{-1} \end{pmatrix}.$$
解毕. □

特别地, 如果$C = O$, 则得到分块对角可逆阵的逆
$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix}.$$
这给出了例3.19的初等变换证明.

例3.36. 设$A = \begin{pmatrix} O & A_1 \\ A_2 & O \end{pmatrix}$, 其中$A_1$和$A_2$是方阵. 求$A$可逆的充分必要条件; 当$A$可逆时, 求$A^{-1}$.

☞ **解.** 设A_1是n阶方阵, A_2是m阶方阵. 如果A可逆, 则存在$m+n$阶方阵$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$, 其中$X \in \mathbb{F}^{m\times n}$, $Y \in \mathbb{F}^{m\times m}$, $Z \in \mathbb{F}^{n\times n}$, $W \in \mathbb{F}^{n\times m}$, 使得
$$\begin{pmatrix} O & A_1 \\ A_2 & O \end{pmatrix}\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = E.$$

这得到

$$A_1Z = E_n, \quad A_1W = O, \quad A_2X = O, \quad A_2Y = E_m.$$

于是A_1和A_2可逆, 且$Z = A_1^{-1}$, $Y = A_2^{-1}$, $W = O$, $X = O$. 反之, 容易得到

$$\begin{pmatrix} O & A_1 \\ A_2 & O \end{pmatrix}\begin{pmatrix} O & A_2^{-1} \\ A_1^{-1} & O \end{pmatrix} = E.$$

所以A可逆的充分必要条件是A_1和A_2都可逆, 且当A可逆时, 有$A^{-1} = \begin{pmatrix} O & A_2^{-1} \\ A_1^{-1} & O \end{pmatrix}$.

也可以用初等变换求A的逆:

$$(A, E) = \begin{pmatrix} O & A_1 & E & O \\ A_2 & O & O & E \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} A_2 & O & O & E \\ O & A_1 & E & O \end{pmatrix} \xrightarrow[A_1^{-1}r_2]{A_2^{-1}r_1} \begin{pmatrix} E & O & O & A_2^{-1} \\ O & E & A_1^{-1} & O \end{pmatrix},$$

这得到$A^{-1} = \begin{pmatrix} O & A_2^{-1} \\ A_1^{-1} & O \end{pmatrix}$. □

习题3.6

A1. 求下面矩阵的逆矩阵:

(1) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 1 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 0 & a_1 & 0 & \cdots & 0 \\ 0 & 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} \\ a_n & 0 & 0 & \cdots & 0 \end{pmatrix}$, $(a_1a_2\cdots a_n \neq 0)$.

A2. 设A, B和C是可逆方阵, 证明: 方阵$X = \begin{pmatrix} O & O & A \\ O & B & O \\ C & O & O \end{pmatrix}$也可逆, 并求$X^{-1}$.

A3. 设$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, 其中$A$和$D$是方阵, 证明:

(1) 当A可逆时, M可逆当且仅当$D - CA^{-1}B$可逆;

(2) 当D可逆时, M可逆当且仅当$A - BD^{-1}C$可逆.

A4. 设$A = (a_{ij})$是n阶上三角阵, 证明: A可逆的充分必要条件是A的n个对角元$a_{11}, a_{22}, \ldots, a_{nn}$都不等于零; 而且当$A$可逆时, A^{-1}也是上三角阵, 且有形式

$$A^{-1} = \begin{pmatrix} a_{11}^{-1} & * & \cdots & * \\ & a_{22}^{-1} & \cdots & * \\ & & \ddots & \vdots \\ & & & a_{nn}^{-1} \end{pmatrix}.$$

更一般的, 设$A = (A_{ij})_{t\times t}$是n阶分块上三角阵, 其中对角块$A_{11}, A_{22}, \ldots, A_{tt}$都是方阵. 证明: A可逆的充分必要条件是A的t个对角块$A_{11}, A_{22}, \ldots, A_{tt}$都可逆; 而且当$A$可逆时, A^{-1}也是分块上三角阵,

且有形式

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & * & \cdots & * \\ & A_{22}^{-1} & \cdots & * \\ & & \ddots & \vdots \\ & & & A_{tt}^{-1} \end{pmatrix}.$$

B1. 设$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, 其中$A, B, C, D$都是$n$阶方阵, 且两两可交换. 证明: X可逆当且仅当$AD - BC$可逆[15].

B2. 设n阶方阵A满足$A^2 = A$, 证明: 存在n阶可逆方阵P和非负整数r, 使得$P^{-1}AP = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$.

§ 3.7 补充题

A1. 设$E_{ij} \in \mathbb{F}^{n\times n}$, $e_i \in \mathbb{F}^n$, $A \in \mathbb{F}^{n\times n}$, 证明:

(1) $e_i^T e_j = \delta_{ij}$;

(2) $E_{ij} = e_i e_j^T$;

(3) $E_{ij}E_{kl} = \delta_{jk}E_{il}$;

(4) $E_{ij}A$将A的第j行变为第i行, 将其余元素全变为零;

(5) AE_{ij}将A的第i列变为第j列, 将其余元素全变为零.

A2. 设$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{F}^{2\times 2}$. 证明:

(1) 总成立

$$A^2 - (a+d)A + (ad - bc)E_2 = O;$$

(2) 如果存在正整数$n > 2$, 使得$A^n = O$, 则$A^2 = O$.

A3. 设A是n阶方阵, 证明: A是反对称阵的充分必要条件是, 对任意的n维向量X, 有$X^TAX = 0$.

A4. 证明: 不存在n阶不可逆阵A, 使得$A^2 + A + E_n = O$.

A5. 设A是n阶实反对称阵, 证明: $E_n - A$可逆.

A6. 设A是n阶可逆阵, 且A的每一行元素之和都等于常数c, 证明: $c \neq 0$, 且A^{-1}的每一行元素之和都等于c^{-1}.

A7. 设n阶方阵A和B满足$A + B = AB$, 证明: $E_n - A$可逆, 且$AB = BA$.

A8. (Sherman-Morrison公式) 设A是n阶可逆阵, α, β是n维列向量, 且$1 + \beta^T A^{-1}\alpha \neq 0$. 证明: $A + \alpha\beta^T$可逆, 且

$$(A + \alpha\beta^T)^{-1} = A^{-1} - \frac{1}{1 + \beta^T A^{-1}\alpha} A^{-1}\alpha\beta^T A^{-1}.$$

---

[15] 可参看习题4.5的B1.

第四章 行列式

解线性方程组的消元法除了Gauss消元法外, 还有加减消元法: 一次消去所有其他的未知数. 为了表示某些特殊的线性方程组用加减消元导出的解的公式, 我们需要行列式的记号和概念. 另一方面, 在上一章中我们用是否和单位阵(行)等价来判别方阵是否可逆. 这个和"数a可逆当且仅当$a \neq 0$"的判别法看上去相去甚远, 那有没有某个数不等于零可以得到方阵可逆呢? 我们发现方阵的行列式恰好是这个数.

本章介绍了行列式的定义和基本性质, 以及行列式的计算方法, 特别地, 我们介绍了在初等变换下矩阵的行列式如何改变. 利用行列式这一工具, 我们用加减消元法给出了一类特殊的线性方程组的解的公式, 即得到所谓的Cramer法则; 并给出了矩阵可逆的行列式判别法, 以及用行列式表示的可逆阵的求逆公式.

本章中$\mathbb{F}$表示任意一个数域.

§ 4.1 线性方程组与二, 三阶行列式

在上章中我们讲了矩阵的运算这一支线故事, 现在回到我们的主线故事—线性方程组的求解之旅. 解线性方程组的基本方法是消元法, 前面我们介绍了Gauss消元法, 现在我们讨论用加减消元法来解未知数个数和方程个数相等的线性方程组.

4.1.1 二阶行列式

设有二元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, & (1) \\ a_{21}x_1 + a_{22}x_2 = b_2. & (2) \end{cases}$$

所谓加减消元, 指的是方程组中的每个方程乘以一个适当的数后相加得到的方程只剩下一个未知数. 例如, 对于上面的线性方程组, 我们有

$$a_{22} \times (1) - a_{12} \times (2) \Longrightarrow (a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12},$$
$$a_{21} \times (1) - a_{11} \times (2) \Longrightarrow (a_{12}a_{21} - a_{11}a_{22})x_2 = b_1a_{21} - b_2a_{11}.$$

于是当$a_{11}a_{22} - a_{12}a_{21} \neq 0$时, 有

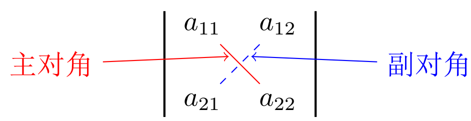
$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

为了更好地表示和记忆上面得到的求解公式, 我们引入二阶行列式的记号.

定义4.1. 对于二阶方阵$(a_{ij}) \in \mathbb{F}^{2\times 2}$, 记

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

我们常常用对角线法则记忆二阶行列式的公式: 主对角上两个元素相乘取正号, 副对角上两个元素相乘取负号.



采用二阶行列式的记号, 上面得到的求解公式可以写为

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

即x_i的分母是系数矩阵对应的二阶行列式, 而分子恰为将系数矩阵的第i列用常数项列向量替换后得到的方阵对应的二阶行列式.

4.1.2 三阶行列式

现在考虑三元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, & (1) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, & (2) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. & (3) \end{cases}$$

我们想用加减消元法来解这个方程组. 设$c_1, c_2, c_3 \in \mathbb{F}$, 则$c_1 \times (1) + c_2 \times (2) + c_3 \times (3)$得到方程

$$\begin{aligned} (c_1a_{11} + c_2a_{21} + c_3a_{31})x_1 &+ (c_1a_{12} + c_2a_{22} + c_3a_{32})x_2 \\ &+ (c_1a_{13} + c_2a_{23} + c_3a_{33})x_3 = c_1b_1 + c_2b_2 + c_3b_3. \end{aligned}$$

如果我们想消去x_2和x_3, 则c_1, c_2和c_3需要满足

$$\begin{cases} c_1a_{12} + c_2a_{22} + c_3a_{32} = 0, \\ c_1a_{13} + c_2a_{23} + c_3a_{33} = 0. \end{cases}$$

借助解析几何的记号, 设$\vec{c} = (c_1, c_2, c_3)$, $\vec{a_2} = (a_{12}, a_{22}, a_{32})$和$\vec{a_3} = (a_{13}, a_{23}, a_{33})$, 则上面的条件即

$$\vec{c} \perp \vec{a_2}, \quad \vec{c} \perp \vec{a_3}.$$

所以我们只需取$\overrightarrow{c}$为$\overrightarrow{a_2}$和$\overrightarrow{a_3}$的外积

$$\overrightarrow{c} = \overrightarrow{a_2} \times \overrightarrow{a_3} = (a_{22}a_{33} - a_{23}a_{32}, a_{13}a_{32} - a_{12}a_{33}, a_{12}a_{23} - a_{13}a_{22}).$$

用这样的c_1, c_2, c_3加减消元后得到

$$\begin{aligned}&(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{21}a_{13}a_{32} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22})x_1\\ =&b_1a_{22}a_{33} - b_1a_{23}a_{32} + b_2a_{13}a_{32} - b_2a_{12}a_{33} + b_3a_{12}a_{23} - b_3a_{13}a_{22},\end{aligned}$$

即

$$\begin{aligned}&(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_1\\ =&b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - b_1a_{23}a_{32} - a_{12}b_2a_{33} - a_{13}a_{22}b_3.\end{aligned}$$

类似于次, 我们可以消去x_1和x_3得到

$$\begin{aligned}&(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_2\\ =&a_{11}b_2a_{33} + b_1a_{23}a_{31} + a_{13}a_{21}b_3 - a_{11}a_{23}b_3 - b_1a_{21}a_{33} - a_{13}b_2a_{31};\end{aligned}$$

消去x_1和x_2得到

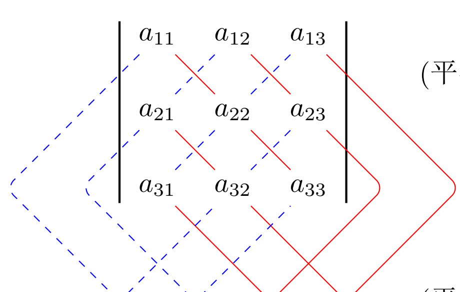
$$\begin{aligned}&(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_3\\ =&a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}.\end{aligned}$$

于是我们引入三阶行列式的记号.

定义4.2. 设有三阶方阵$(a_{ij}) \in \mathbb{F}^{3\times 3}$, 记

$$\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \in \mathbb{F}.$$

三阶行列式也有对角线法则:



(平行于)主对角的三个元相乘, 取正号

(平行于)副对角的三个元相乘, 取负号

采用三阶行列式的记号, 对上面的三元线性方程组, 我们得到: 当

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

时有

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

即x_i的分母是系数矩阵对应的三阶行列式, 而分子恰为将系数矩阵的第i列用常数项列向量替换后得到的方阵对应的三阶行列式.

4.1.3 n阶行列式的定义

继续上面的讨论, 我们需要考虑用加减消元法来解由n个方程组成的n元线性方程组. 我们当然希望可以引入n阶行列式的记号, 得到类似的求解公式:

$$x_i = \frac{n\text{阶行列式}}{n\text{阶行列式}}, \quad (i = 1, 2, \ldots, n),$$

其中分母为系数矩阵所对应的n阶行列式, 而分子为将系数矩阵的第i列用常数项列向量替换后得到的n阶方阵所对应的n阶行列式. 在$n = 3$时我们借助了解析几何中的内积和外积的概念, 因此为了得到上面的求解公式, 进而定义n阶行列式, 我们需要发展高维空间的解析几何, 定义这里的内积和外积. 这看上去是一个艰巨的任务.

也可以倒过来想问题, 我们先从二, 三阶行列式的定义中归纳出n阶行列式的(一个)定义, 然后证明这个定义是"好的", 即证明由n个方程组成的n元线性方程组在一定条件下确实有上面的求解公式. 本讲义采用这种方法来定义n阶行列式和用加减消元法解线性方程组, 因为这个途径看上去相对简单.

下面我们仔细分析一下二, 三阶行列式的定义. 对于二阶方阵$(a_{ij})_{2\times 2} \in \mathbb{F}^{2\times 2}$, 对应$\mathbb{F}$中一个数

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \in \mathbb{F};$$

对于三阶方阵$(a_{ij})_{3\times 3} \in \mathbb{F}^{3\times 3}$, 对应$\mathbb{F}$中一个数

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

$$- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \in \mathbb{F}.$$

可以看到, 右边是一些项的代数和(所谓代数和指的是有正有负), 每一项不看符号的话是几个a_{ij}的乘积(二阶行列式是2个, 三阶行列式是3个). 分析每一项中a_{ij}的双下标, 可以得到表4.1. 我们看到, 所有项的行标都是一样的(二阶行列式是$1,2$, 三阶行列式是$1,2,3$); 而列标则各不相同. 二阶行列式两项的列标分别是$1,2$和$2,1$, 他们是$1,2$的所有排列. 类似的, 三阶行列式六项的列标恰是$1,2,3$的所有排列.

表 4.1: 二, 三阶行列式各项的行标和列标

二阶行列式

项	行标	列标
$a_{11}a_{22}$	$1,2$	$1,2$
$a_{12}a_{21}$	$1,2$	$2,1$

三阶行列式

项	行标	列标
$a_{11}a_{22}a_{33}$	$1,2,3$	$1,2,3$
$a_{12}a_{23}a_{31}$	$1,2,3$	$2,3,1$
$a_{13}a_{21}a_{32}$	$1,2,3$	$3,1,2$
$a_{11}a_{23}a_{32}$	$1,2,3$	$1,3,2$
$a_{12}a_{21}a_{33}$	$1,2,3$	$2,1,3$
$a_{13}a_{22}a_{31}$	$1,2,3$	$3,2,1$

由此, 我们有理由给出下面的设想. 对于n阶方阵$A = (a_{ij})_{n\times n}$, 它对应$\mathbb{F}$中的一个数

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} := \sum_{\substack{p_1,\ldots,p_n\in\{1,2,\ldots,n\} \\ \text{互不相同}}} \pm a_{1p_1}a_{2p_2}\cdots a_{np_n} \in \mathbb{F}.$$

上面的求和是对$1,2,\ldots,n$的所有排列进行的, 这是个组合的概念. 那么每一项前面的符号如何选择呢? 从组合数学知道, 每个排列都有一个符号; 分析二, 三阶行列式每一项列标的排列的符号, 可以发现恰好与这项在行列式中的符号一致. 这提示我们, 应该把每一项前面的符号取成列标对应排列的符号. 这样, 我们就得到了一个可能的n阶行列式的定义, 在本章最后我们将说明这个定义是"好的".

于是, 为了给出这种n阶行列式的定义, 我们需要介绍排列和排列的符号这两个组合概念. 这并不是一项困难的任务, 我们将在下节给出.

习题4.1

A1. 求下面行列式的值

(1) $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$; (2) $\begin{vmatrix} 2 & 0 & -1 \\ 1 & -4 & -1 \\ -1 & 8 & 3 \end{vmatrix}$; (3) $\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$.

§ 4.2　排列及其符号

4.2.1　排列的符号

数$1, 2, \ldots, n$的按照任意次序排成的有序数组称为一个n元排列, 常记为

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix},$$

简记为$p_1, p_2, \ldots, p_n$. 所有n元排列做成的集合记为S_n. 容易知道$|S_n| = n!$.

例4.1. **(1)** $S_1 = \{1\}$, $S_2 = \{1, 2;\ 2, 1\}$,

$S_3 = \{1, 2, 3;\ 1, 3, 2;\ 2, 1, 3;\ 2, 3, 1;\ 3, 1, 2;\ 3, 2, 1\}$;

(2) $1, 2, \ldots, n \in S_n$, 称这个排列为自然排列;

(3) $3, 2, 1, 3 \notin S_4$.

排列的符号可通过逆序数来定义的.

定义4.3. 设$p_1, p_2, \ldots, p_n \in S_n$.

(1) 如果存在$1 \leqslant i < j \leqslant n$, 使得$p_i > p_j$, 则称$(p_i, p_j)$是该排列的一个逆序对;

(2) 逆序对的总数称为这个排列的逆序数, 记为$\tau(p_1, p_2, \ldots, p_n)$.

例4.2. 设$3, 2, 1, 4 \in S_4$, 则有逆序对$(3, 2), (3, 1), (2, 1)$, 所以$\tau(3, 2, 1, 4) = 3$.

可以用下面的"前大"或"后小"法计算逆序数. 设$p_1, p_2, \ldots, p_n \in S_n$, 令

$$\tau(p_i) := (p_1, \ldots, p_{i-1}\text{ 中比 } p_i \text{ 大的个数}),$$

则有

$$\tau(p_1, p_2, \ldots, p_n) = \sum_{i=2}^{n} \tau(p_i).$$

这就是前大法. 类似的, 令

$$\widetilde{\tau}(p_i) := (p_{i+1}, \ldots, p_n\text{ 中比 } p_i \text{ 小的个数}),$$

则有

$$\tau(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n-1} \widetilde{\tau}(p_i).$$

这就是后小法.

例4.3. 设$4, 5, 1, 3, 6, 2 \in S_6$, 则

$$\tau(4, 5, 1, 3, 6, 2) \xlongequal{\text{前大}} 0 + 2 + 2 + 0 + 4 = 8,$$

$$\tau(4, 5, 1, 3, 6, 2) \xlongequal{\text{后小}} 3 + 3 + 0 + 1 + 1 = 8.$$

下面定义排列的符号.

定义4.4. 设$p_1, p_2, \ldots, p_n \in S_n$.

(1) 记

$$\delta(p_1, p_2, \ldots, p_n) := (-1)^{\tau(p_1, p_2, \ldots, p_n)} \in \{\pm 1\},$$

称其为排列的符号;

(2) 如果$\delta(p_1, p_2, \ldots, p_n) = 1 (\Longleftrightarrow \tau(p_1, p_2, \ldots, p_n)$是偶数), 则称排列$p_1, p_2, \ldots, p_n$是偶排列; 如果$\delta(p_1, p_2, \ldots, p_n) = -1 (\Longleftrightarrow \tau(p_1, p_2, \ldots, p_n)$是奇数), 则称排列$p_1, p_2, \ldots, p_n$是奇排列.

例4.4. (1) 自然排列是偶排列;

(2) 由于$\delta(3, 2, 1, 4) = (-1)^3 = -1$, 所以排列$3, 2, 1, 4$是奇排列;

(3) 由于$\delta(4, 5, 1, 3, 6, 2) = (-1)^8 = 1$, 所以排列$4, 5, 1, 3, 6, 2$是偶排列.

4.2.2 对换与排列符号

为了后面证明行列式性质的需要, 我们引入排列的对换, 并考虑对换对符号的影响.

定义4.5. 设$1 \leqslant i < j \leqslant n$, 定义映射

$$\sigma_{ij} : S_n \longrightarrow S_n,$$

使得对任意$p_1, p_2, \ldots, p_n \in S_n$有

$$\sigma_{ij}(p_1, p_2, \ldots, p_n) = p_1, \ldots, p_{i-1}, p_j, p_{i+1}, \ldots, p_{j-1}, p_i, p_{j+1}, \ldots, p_n.$$

称σ_{ij}为对换. 当$j = i + 1$时, 称$\sigma_{i,i+1}$为相邻对换.

例4.5. $\sigma_{35}(3, 4, 1, 5, 2) = 3, 4, 2, 5, 1.$

我们看到, 排列$3, 4, 1, 5, 2$有逆序对$(3, 1), (3, 2), (4, 1), (4, 2), (5, 2)$, 而排列$3, 4, 2, 5, 1$除了这些逆序对外, 还多了$(2, 1)$这个逆序对. 于是对换$\sigma_{35}$改变排列$3, 4, 1, 5, 2$的符号, 下面证明这个是一般规律.

引理4.1. 对换改变排列的符号和奇偶性.

☞ **证明.** 先证明相邻对换的情形. 设有相邻对换

$$a_1, \ldots, a_k, p, l, b_1, \ldots, b_s \xmapsto{\sigma_{k+1,k+2}} a_1, \ldots, a_k, p, l, b_1, \ldots, b_s,$$

我们用前大法来比较两边的逆序数. 左边排列前大法中的τ记为τ_1, 右边排列前大法中的τ记为τ_2, 则有

$$\begin{aligned}
&\tau_1(a_i) = \tau_2(a_i), \quad i = 1, 2, \ldots, k,\\
&\tau_1(b_j) = \tau_2(b_j), \quad j = 1, 2, \ldots, s,\\
&\tau_1(p) = \tau_2(p), \quad \tau_1(l) = \tau_2(l) + 1, \quad 若\ p > l,\\
&\tau_1(p) = \tau_2(p) - 1, \quad \tau_1(l) = \tau_2(l), \quad 若\ p < l.
\end{aligned}$$

所以得到逆序数的关系

$$\tau(a_1,\dots,a_k,p,l,b_1,\dots,b_s)=\begin{cases}\tau(a_1,\dots,a_k,l,p,b_1,\dots,b_s)+1, & 若\,p>l,\\ \tau(a_1,\dots,a_k,l,p,b_1,\dots,b_s)-1, & 若\,p<l,\end{cases}$$

即逆序数的奇偶性改变, 于是排列的符号和奇偶性也改变(也可直接讨论逆序对).

下面再考虑一般的对换. 设有对换

$$a_1,\dots,a_i,t_1,\dots,t_k,a_j,\dots,a_n \xmapsto{\sigma_{ij}} a_1,\dots,a_{i-1},a_j,t_1,\dots,t_k,a_i,a_{j+1},\dots,b_s,$$

我们说明可以用相邻对换实现这个对换. 事实上, 下面的一系列相邻对换就实现了σ_{ij}:

$$\begin{array}{ccc}
a_1,\dots,a_i,t_1,\dots,t_k,a_j,\dots,a_n & \xmapsto{\sigma_{ij}} & a_1,\dots,a_j,t_1,\dots,t_k,a_i,\dots,b_s\\
\downarrow & & \uparrow\\
a_1,\dots,a_{i-1},t_1,a_i,\dots,t_k,a_j,\dots,a_n & & \vdots\\
\downarrow & & \uparrow\\
a_1,\dots,a_{i-1},t_1,t_2,a_i,\dots,t_k,a_j,\dots,a_n & & \vdots\\
\downarrow & & \uparrow\\
\vdots & & a_1,\dots,a_{i-1},t_1,\dots,t_{k-2},a_j,t_{k-1},t_k,a_i,\dots,a_n\\
\downarrow & & \uparrow\\
\vdots & & a_1,\dots,a_{i-1},t_1,\dots,t_{k-1},a_j,t_k,a_i,\dots,a_n\\
\downarrow & & \uparrow\\
a_1,\dots,a_{i-1},t_1,\dots,t_k,a_i,a_j,\dots,a_n & \mapsto & a_1,\dots,a_{i-1},t_1,\dots,t_k,a_j,a_i,\dots,a_n
\end{array}$$

一共有$2k+1$次相邻对换, 所以排列符号和奇偶性改变. □

这个引理有个显然的推论, 它给出了排列的符号的另一种刻画.

推论4.2. 设排列$p_1,p_2,\dots,p_n\in S_n$经M次对换后变为$q_1,q_2,\dots,q_n$, 则有

$$\delta(p_1,p_2,\dots,p_n)=(-1)^M\delta(q_1,q_2,\dots,q_n).$$

特别的, 设$p_1,p_2,\dots,p_n\in S_n$经M次对换后变为自然排列, 则

$$\delta(p_1,p_2,\dots,p_n)=(-1)^M.$$

最后我们证明奇偶排列各半.

例4.6. 设$n\geqslant 2$, 证明: S_n中奇, 偶排列各一半.

☞ **证明.** 记S_n中奇排列全体所成的集合为O_n, 偶排列全体所成的集合为E_n. 由引理4.1, 我们有映射

$$\sigma_{12}:O_n\longrightarrow E_n,\quad \sigma_{12}:E_n\longrightarrow O_n.$$

而合成$\sigma_{12}\circ\sigma_{12}$为恒等映射, 所以集合$O_n$和$E_n$同构. 进而$|O_n|=|E_n|$. □

习题4.2

A1. 确定下面排列的逆序数和符号:
(1) $2,1,7,9,8,6,3,5,4$; (2) $n,n-1,\ldots,2,1$;
(3) $1,n,n-1,\ldots,3,2$; (4) $2,4,\ldots,2n-2,2n,1,3,\ldots,2n-3,2n-1$;
(5) $1,3,\ldots,2n-3,2n-1,2,4,\ldots,2n-2,2n$.
A2. 选择i和j, 使得9元排列
(1) $1,2,7,4,i,5,6,j,9$成为偶排列;
(2) $1,i,2,5,j,4,8,9,7$成为奇排列.
A3. 证明: (1) 排列$p_1,p_2,\ldots,p_n\in S_n$可通过$\tau(p_1,p_2,\ldots,p_n)$次相邻对换变为自然排列, 其中$\tau(p_1,p_2,\ldots,p_n)$是排列$p_1,p_2,\ldots,p_n$的逆序数;
(2) 任意的n元排列都可以经过至多$n-1$次对换变为自然排列;
(3) 存在n元排列, 使得它不能经过小于$n-1$次对换变为自然排列.

§ 4.3 行列式的定义

现在我们给出我们心目中的行列式的定义.

定义4.6. 设$A=(a_{ij})_{n\times n}\in\mathbb{F}^{n\times n}$, 方阵$A$的行列式(或$n$阶行列式)定义为

$$\begin{aligned} D =D_n = \det(A) = |A| = |a_{ij}| = |a_{ij}|_{n\times n} = &\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ := \sum_{p_1,p_2,\ldots,p_n\in S_n} &\delta(p_1,p_2,\ldots,p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n}\in\mathbb{F}. \end{aligned}$$

所以n阶行列式是一个数, 它是$n!$项的代数和, 每项为n个数的乘积, 它们取自不同的行和列, 其中行标为自然排列, 而列标的排列的符号就是该项在和式中的符号.

4.3.1 低阶行列式

我们的定义当然要符合§4.1中二阶和三阶行列式的公式. 下面看$n=1,2,3$时, 按照上面的定义得到的行列式的公式.

由定义, 一阶行列式为$|a|=a$, 这要区别于数的模长. 又由表4.2, 我们有二阶行列式

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

和三阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}$$

表 4.2: S_2和S_3中各排列的符号

S_2	逆序数	符号
$1,2$	0	1
$2,1$	1	-1

S_3	逆序数	符号
$1,2,3$	0	1
$2,3,1$	2	1
$3,1,2$	2	1
$1,3,2$	1	-1
$2,1,3$	1	-1
$3,2,1$	3	-1

$$- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

这确实和§4.1中给出的二, 三阶行列式的公式是一致的.

二, 三阶行列式可以用对角线法则来计算, 但是四阶及更高阶行列式则无对角线法则. 比如, 如果四阶行列式有对角线法则, 则它是8项的和; 但是由定义, 四阶行列式是$4! = 24$项的和.

4.3.2 转置方阵的行列式

上面行列式的定义中每一项的行标为自然排列, 列标排列的符号为该项的符号. 这里强调了行, 而列是被动的, 这有点像现在的阅读习惯是一行一行地读. 但正如有左撇子一样, 并不是所有的人都更喜欢行, 也有人更喜欢列, 比如由于文字是写在狭长的竹片上, 古人是一列一列地阅读的. 下面替更喜欢列的人考察一下行列式的定义中是否可以先强调列.

交换每一项中因子的次序, 二阶和三阶行列式的公式可以改写为

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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{31}a_{12}a_{23} + a_{21}a_{32}a_{13}$$

$$- a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}.$$

此时, 每一项中列标为自然排列, 而行标排列的符号就是该项在和式中的符号. 于是我们有理由相信这是一般规律.

定理4.3. 设$A = (a_{ij})_{n\times n} \in \mathbb{F}^{n\times n}$, 则[1]

$$\det(A) = \sum_{p_1,p_2,\ldots,p_n \in S_n} \delta(p_1,\ldots,p_n) a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}.$$

进而有$\det(A^T) = \det(A)$.

[1] 每一项n个数的列标为自然排列, 而行标的排列符号就是该项在和式中的符号.

☞ **证明.** (1) 设$p_1, p_2, \ldots, p_n \in S_n$, $q_1, q_2, \ldots, q_n \in S_n$, 对任意的$1 \leqslant i < j \leqslant n$, 由引理4.1我们有

$$\begin{aligned}\delta(q_1, q_2, \ldots, q_n) &= -\delta(q_1, \ldots, q_{i-1}, q_j, q_{i+1}, \ldots, q_{j-1}, q_i, q_{j+1}, \ldots, q_n),\\ \delta(p_1, p_2, \ldots, p_n) &= -\delta(p_1, \ldots, p_{i-1}, p_j, p_{i+1}, \ldots, p_{j-1}, p_i, p_{j+1}, \ldots, p_n).\end{aligned}$$

所以有

$$\begin{aligned}&\delta(q_1, q_2, \ldots, q_n)\delta(p_1, p_2, \ldots, p_n)a_{q_1p_1}a_{q_2p_2}\cdots a_{q_np_n}\\ =&\delta(q_1, \ldots, q_j, \ldots, q_i, \ldots, q_n)\delta(p_1, \ldots, p_j, \ldots, p_i, \ldots, p_n)a_{q_1p_1}a_{q_2p_2}\cdots a_{q_np_n}\\ =&\delta(q_1, \ldots, q_j, \ldots, q_i, \ldots, q_n)\delta(p_1, \ldots, p_j, \ldots, p_i, \ldots, p_n)a_{q_1p_1}\cdots a_{q_jp_j}\cdots a_{q_ip_i}\cdots a_{q_np_n}.\end{aligned}$$

所以将行标排列与列标排列同时进行一样的对换, 上面连等式的最左边不改变.

注意到

$$\delta(p_1, p_2, \ldots, p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n} = \delta(1, 2, \ldots, n)\delta(p_1, p_2, \ldots, p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n},$$

列标排列$p_1, p_2, \ldots, p_n$可经一系列对换变成自然排列$1, 2, \ldots, n$, 设此时行标排列$1, 2, \ldots, n$用同样的对换变成$q_1, q_2, \ldots, q_n$, 即

$$p_1, p_2, \ldots, p_n \xrightarrow{\text{一系列对换}} 1, 2, \ldots, n, \qquad 1, 2, \ldots, n \xrightarrow{\text{同样对换}} q_1, q_2, \ldots, q_n,$$

则有

$$\begin{aligned}\delta(p_1, p_2, \ldots, p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n} =&\delta(q_1, q_2, \ldots, q_n)\delta(1, 2, \ldots, n)a_{q_11}a_{q_22}\cdots a_{q_nn}\\ =& \delta(q_1, q_2, \ldots, q_n)a_{q_11}a_{q_22}\cdots a_{q_nn}.\end{aligned}$$

又由于若$p_i = j$, 则$q_j = i$. 所以$q_1, q_2, \ldots, q_n$和$p_1, p_2, \ldots, p_n$相互唯一确定. 这得到

$$\begin{aligned}\det(A) &= \sum_{q_1, q_2, \ldots, q_n \in S_n} \delta(q_1, q_2, \ldots, q_n)a_{q_11}a_{q_22}\cdots a_{q_nn}\\ &= \sum_{p_1, p_2, \ldots, p_n \in S_n} \delta(p_1, p_2, \ldots, p_n)a_{p_11}a_{p_22}\cdots a_{p_nn}.\end{aligned}$$

(2) 设$A^T = (b_{ij})_{n\times n}$, 则$b_{ij} = a_{ji}$. 于是

$$\begin{aligned}\det(A^T) &\xlongequal{\text{定义}} \sum_{p_1, p_2, \ldots, p_n \in S_n} \delta(p_1, p_2, \ldots, p_n)b_{1p_1}b_{2p_2}\cdots b_{np_n}\\ &= \sum_{p_1, p_2, \ldots, p_n \in S_n} \delta(p_1, p_2, \ldots, p_n)a_{p_11}a_{p_22}\cdots a_{p_nn}\\ &\xlongequal{(1)} \det(A).\end{aligned}$$

证毕. □

定理4.3说明行列式中行和列的地位等同, 行列式关于行的性质必对应到相应的关于列的性质. 而且由上面的证明, 我们还有: 对任意固定的$i_1, i_2, \ldots, i_n \in S_n$成立

$$\begin{aligned}\det(A) &= \sum_{j_1, j_2, \ldots, j_n \in S_n} \delta(i_1, i_2, \ldots, i_n)\delta(j_1, j_2, \ldots, j_n)a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}\\ &= \sum_{j_1, j_2, \ldots, j_n \in S_n} \delta(i_1, i_2, \ldots, i_n)\delta(j_1, j_2, \ldots, j_n)a_{j_1i_1}a_{j_2i_2}\cdots a_{j_ni_n}.\end{aligned}$$

4.3.3 三角行列式

我们已经知道直到三阶的行列式的公式. 对于四阶行列式, 一共有24项, 要把这些项都正确地写出来, 并记住公式(四阶行列式没有对角线法则!)是一个困难的任务. 所以去记忆四阶及其高阶行列式的公式不是一个明智的决定. 另一方面, 由于零和任意数的乘积还是零, 所以如果一项中有某个元素为零, 则这一项必为零, 可以不考虑. 于是当方阵的零很多时, 常常可以由定义得出它的行列式的值. 零很多的方阵, 我们前面介绍过上三角和下三角方阵, 对它们我们确实可以由定义得出及其方便记忆的行列式公式.

例4.7. *证明: 上, 下三角方阵的行列式分别为*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}, \quad \begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

特别有, 对角方阵的行列式为

$$\begin{vmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

☞ **证明.** 取转置, 下三角方阵的公式可由上三角方阵的公式得到, 所以只需证明上三角方阵的行列式公式. 行列式定义的和式中的一般项为

$$\delta(p_1, p_2, \ldots, p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n}.$$

但是上三角方阵的最后一行除了a_{nn}外其余元素一定为零, 所以我们只需考虑这样的项

$$\delta(p_1, \ldots, p_{n-1}, n)a_{1p_1}\cdots a_{n-1,p_{n-1}}a_{nn}.$$

又第$n-1$行除了最后两个位置外的元素一定为零, 且$p_{n-1} \neq n$, 所以只需考虑$p_{n-1} = n-1$的项, 即考虑

$$\delta(p_1, \ldots, p_{n-2}, n-1, n)a_{1p_1}\cdots a_{n-2,p_{n-2}}a_{n-1,n-1}a_{nn}.$$

继续这个讨论, 我们得到上三角方阵的行列式的定义和式中除了

$$\delta(1, 2, \ldots, n)a_{11}a_{22}\cdots a_{nn}$$

外的其余项一定都是零. 由此得到上三角方阵的行列式公式. □

4.3.4 行列式的计算

我们定义了行列式, 当然希望任意给定一个n阶方阵, 可以求出它的行列式的值, 至少希望可以用计算机程序很快得到. 但是根据定义, n阶行列式是$n!$项的代数和, 所以如果按照定义计算, 我们(计算机)至少需要计算$n!$次. 而微积分中关于取$n!$的近似值的Stirling公式告诉我们, $n!$是非常高阶

的无穷大. 我们以$n = 50$为例说明, 因为计算50阶的行列式在工程上并不是一个万年不遇的问题. 我们有

$$\begin{aligned}50! =& 30414\quad 0932017133\quad 7804361260\quad 8166064768\quad 8443776415\\ & 6896051200\quad 0000000000 \approx 3\times 10^{64}.\end{aligned}$$

一年有

$$365\times 24\times 3600 = 31536000 \approx 3\times 10^{7}\quad 秒.$$

而亿亿$= 10^8\times 10^8 = 10^{16}$. 于是用每秒运算亿亿次的计算机, 计算50!个运算, 大约需要

$$\frac{3\times 10^{64}}{3\times 10^{7}\times 10^{16}} = 10^{41}\quad 年.$$

可见我们不可能等待这么长的时间去按定义计算一个50阶的行列式.

那么如何计算一般的n阶行列式呢? 希望读者不要由于定义了一个似乎很难计算的概念而沮丧, 让我们想办法吧. 前面章节中解线性方程组的Gauss消元法和判别方阵可逆性的思考过程可以给我们一些提示. 我们都是先考虑简单的情形, 然后将一般的情形转化为简单情形, 并思考在转化过程中所研究性质如何改变. 让我们也用先简单再一般的这个数学思考方法来想如何计算行列式的值. 我们对行列式并不是一无所知, 我们至少知道低阶和三角方阵的行列式公式. 所以是否可以如下计算行列式?

$$\begin{tikzcd} & \boxed{\text{三角方阵行列式}} \\ \boxed{n\text{阶行列式}} \arrow[ur, "?"] \arrow[dr, "?"'] & \\ & \boxed{\text{二阶行列式}} \end{tikzcd}$$

我们先想想可以如何实现这两个想法. 首先, 从方阵到上三角阵我们是有办法的—行初等变换. 事实上, 如果一个方阵是阶梯形阵, 那么它一定是上三角的[2]. 于是, 要把方阵的行列式转化为上三角方阵的行列式来计算, 我们需要考虑初等变换下方阵的行列式如何改变. 我们称这种计算行列式的方法为化三角. 其次, 要把一个n阶行列式转化为二阶行列式计算, 我们可以如下实现.

$$\begin{tikzcd} \boxed{n\text{阶}} \arrow[rr, dashed] \arrow[d] & & \boxed{2\text{阶}} \\ \boxed{n-1\text{阶}} \arrow[r] & \cdots \arrow[r] & \boxed{3\text{阶}} \arrow[u] \end{tikzcd}$$

于是我们需要考虑n阶行列式和$n-1$阶行列式的关系, 后面的行列式按照行(列)展开公式就给出了这种关系. 我们称这种计算行列式的方法为降阶法.

习题4.3

A1. 在6阶行列式中$a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$和$a_{32}a_{43}a_{14}a_{51}a_{66}a_{25}$这两项应带什么符号?

A2. 确定正整数i和j的值, 使得7阶行列式含有项

(1) $-a_{62}a_{i5}a_{33}a_{j4}a_{46}a_{21}a_{77}$;

[2] 参看习题2.3.

(2) $a_{1i}a_{24}a_{31}a_{47}a_{55}a_{63}a_{7j}$.

A3. 确定多项式

$$f(x)=\begin{vmatrix}3x & x & 1 & 2\\ 1 & x & -1 & 1\\ 3 & 2 & x & 1\\ 1 & 1 & 1 & x\end{vmatrix}$$

中x^4和x^3的系数.

A4. 证明:

$$\begin{vmatrix}* & * & \cdots & * & a_1\\ * & * & \cdots & a_2 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ * & a_{n-1} & \cdots & 0 & 0\\ a_n & 0 & \cdots & 0 & 0\end{vmatrix}=(-1)^{\frac{1}{2}n(n-1)}a_1a_2\cdots a_n.$$

A5. 设$a_{ij}(t)$都是t的可导函数, 证明:

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\begin{vmatrix}a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t)\\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t)\\ \vdots & \vdots & & \vdots\\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t)\end{vmatrix}$$
$$=\sum_{j=1}^{n}\begin{vmatrix}a_{11}(t) & \cdots & a_{1,j-1}(t) & \frac{\mathrm{d}}{\mathrm{d}\,t}a_{1j}(t) & a_{1,j+1}(t) & \cdots & a_{1n}(t)\\ a_{21}(t) & \cdots & a_{2,j-1}(t) & \frac{\mathrm{d}}{\mathrm{d}\,t}a_{2j}(t) & a_{2,j+1}(t) & \cdots & a_{2n}(t)\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n1}(t) & \cdots & a_{n,j-1}(t) & \frac{\mathrm{d}}{\mathrm{d}\,t}a_{nj}(t) & a_{n,j+1}(t) & \cdots & a_{nn}(t)\end{vmatrix}.$$

B1. 称映射$f:\underbrace{\mathbb{F}^n\times\cdots\times\mathbb{F}^n}_{k}\longrightarrow\mathbb{F}$为$\mathbb{F}^n$上的一个$k$元函数. 设$f$是$\mathbb{F}^n$上的一个$k$元函数, 如果对每个$i$, $1\leqslant i\leqslant k$, 都有

$$\begin{aligned}&f(\alpha_1,\ldots,\alpha_{i-1},a\alpha+b\beta,\alpha_{i+1},\ldots,\alpha_k)=af(\alpha_1,\ldots,\alpha_{i-1},\alpha,\alpha_{i+1},\ldots,\alpha_k)\\&+bf(\alpha_1,\ldots,\alpha_{i-1},\beta,\alpha_{i+1},\ldots,\alpha_k),\qquad(\forall a,b\in\mathbb{F},\forall\alpha_l,\alpha,\beta,\in\mathbb{F}^n),\end{aligned}$$

则称f是k重线性函数; 如果对任意i,j, $1\leqslant i<j\leqslant k$, 都有

$$f(\alpha_1,\ldots,\alpha_{i-1},\alpha,\alpha_{i+1},\ldots,\alpha_{j-1},\alpha,\alpha_{j+1},\ldots,\alpha_k)=0,\quad(\forall\alpha_l,\alpha\in\mathbb{F}^n),$$

则称f是反对称的. 设f是$\mathbb{F}^n$上的一个n元函数, 如果有

$$f(e_1,e_2,\ldots,e_n)=1,$$

则称f是规范的. 证明: 如果f是$\mathbb{F}^n$上的一个规范, 反对称, n重线性函数, 则f必是行列式函数, 即对任意$\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{F}^n$, 有

$$f(\alpha_1,\alpha_2,\ldots,\alpha_n)=\det(\alpha_1,\alpha_2,\ldots,\alpha_n).$$

B2. 设方阵$A = (a_{ij}) \in \mathbb{F}^{n\times n}$, 称

$$\mathrm{Per}(A) = \sum_{p_1,p_2,\ldots,p_n\in S_n} a_{1p_1}a_{2p_2}\cdots a_{np_n}$$

为方阵A所对应的n阶积和式(permanent).

(1) 如果$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, 证明: $\mathrm{Per}(A) = a_{11}a_{22} + a_{12}a_{21}$;

(2) 对任意$A \in \mathbb{F}^{n\times n}$, 证明: $\mathrm{Per}(A^T) = \mathrm{Per}(A)$;

(3) 如果f是$\mathbb{F}^n$上的一个n元函数, 使得对任意$\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}^n$, 有

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = \mathrm{Per}(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

证明: f是$\mathbb{F}^n$上的一个规范, 对称, n重线性函数. 这里, 对$\mathbb{F}^n$上的k元函数g, 如果对任意i, j, $1 \leqslant i < j \leqslant k$, 都有

$$\begin{aligned} &g(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \ldots, \alpha_k) \\ =&g(\alpha_1, \ldots, \alpha_{i-1}, \alpha_j, \alpha_{i+1}, \ldots, \alpha_{j-1}, \alpha_i, \alpha_{j+1}, \ldots, \alpha_k), \qquad (\forall \alpha_l \in \mathbb{F}^n), \end{aligned}$$

则称g是对称的;

(4) 如果f是$\mathbb{F}^n$上的一个规范, 对称, n重线性函数, 是否一定有$f = \mathrm{Per}$? 如果成立的话给出证明; 如果不成立的话给出反例, 并指出再加上什么条件可以得到$f = \mathrm{Per}$.

§ 4.4 初等变换和行列式

正如前面所预告的, 为了实现行列式的化三角算法, 我们需要研究在初等变换下方阵的行列式如何改变. 有了想法, 要得出一般规律就不困难了. 比如可以看看二阶和三阶行列式, 然后归纳出一般规律, 或者利用行列式的几何意义猜出结论. 我们下面省略了这些发现过程, 读者应该在阅读下文前自己先研究下初等变换下行列式的变化规律.

4.4.1 初等变换和行列式

- **第I类初等变换**

关于调行(列)变换, 我们有

定理4.4. 交换方阵的某两行(列), 其行列式变号, 即

$$\begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_n \end{vmatrix} \xlongequal{r_i \leftrightarrow r_j} - \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{vmatrix}.$$

证明. 设$A = (a_{kl})_{n\times n}$, 而$A \overset{r_i \leftrightarrow r_j}{\rightsquigarrow} B = (b_{kl})_{n\times n}$, 则

$$b_{kl} = \begin{cases} a_{kl}, & k \neq i, j, \\ a_{jl}, & k = i, \\ a_{il}, & k = j. \end{cases}$$

所以有

$$\begin{aligned} \det(B) &= \sum_{p_1,\ldots,p_n \in S_n} \delta(p_1,\ldots,p_n) b_{1p_1}\cdots b_{ip_i}\cdots b_{jp_j}\cdots b_{np_n} \\ &= \sum_{p_1,\ldots,p_n \in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots a_{jp_i}\cdots a_{ip_j}\cdots a_{np_n} \\ &= \sum_{p_1,\ldots,p_n \in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots a_{ip_j}\cdots a_{jp_i}\cdots a_{np_n} \\ &= -\sum_{p_1,\ldots,p_n \in S_n} \delta(p_1,\ldots,p_j,\ldots,p_i,\ldots,p_n) a_{1p_1}\cdots a_{ip_j}\cdots a_{jp_i}\cdots a_{np_n} \\ &= -\sum_{q_1,\ldots,q_n \in S_n} \delta(q_1,\ldots,q_n) a_{1q_1}\cdots a_{nq_n} = -\det(A). \end{aligned}$$

取转置就得到关于列的结论. □

- **第II类初等变换**

对于行(列)数乘变换, 我们有

定理4.5. 方阵的某行(列)乘数$a \in \mathbb{F}$, 其行列式也乘a, 即

$$\begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_{i-1} \\ a \times \alpha_i \\ \alpha_{i+1} \\ \vdots \\ \alpha_n \end{vmatrix} = a \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_{i-1} \\ \alpha_i \\ \alpha_{i+1} \\ \vdots \\ \alpha_n \end{vmatrix}.$$

证明. 设$A=(a_{kl})_{n\times n}$, 而$A \overset{a\times r_i}{\rightsquigarrow} B=(b_{kl})_{n\times n}$, 则

$$b_{kl}=\begin{cases} a_{kl}, & k\neq i, \\ aa_{il}, & k=i. \end{cases}$$

所以

$$\begin{aligned}\det(B)&=\sum_{p_1,\ldots,p_n\in S_n}\delta(p_1,\ldots,p_n)b_{1p_1}\cdots b_{ip_i}\cdots b_{np_n}\\ &=\sum_{p_1,\ldots,p_n\in S_n}\delta(p_1,\ldots,p_n)a_{1p_1}\cdots(aa_{ip_i})\cdots a_{np_n}\\ &=a\sum_{p_1,\ldots,p_n\in S_n}\delta(p_1,\ldots,p_n)a_{1p_1}\cdots a_{ip_i}\cdots a_{np_n}=a\det(A).\end{aligned}$$

证毕. □

定理4.5也可以看成行列式可以按照行(列)提取公因子, 但是一次只能提取一个行(列)的公因子. 请比较

$$a\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{vmatrix}=\begin{vmatrix} aa_{11} & aa_{12} \\ a_{21} & a_{22}\end{vmatrix}=\begin{vmatrix} aa_{11} & a_{12} \\ aa_{21} & a_{22}\end{vmatrix},\quad a\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}=\begin{pmatrix} aa_{11} & aa_{12} \\ aa_{21} & aa_{22}\end{pmatrix}.$$

例4.8 (数乘的行列式). 设$A\in\mathbb{F}^{n\times n}$, $a\in\mathbb{F}$, 则$|aA|=a^n|A|$.

证明. 每行都提取公因子a. □

定理4.4和定理4.5有下面的推论.

推论4.6. **(1)** 有全零行(列)的方阵的行列式的值为0;

(2) 有两行(列)相同的方阵的行列式的值为0;

(3) 有两行(列)成比例的方阵的行列式的值为0.

证明. (1) 定理4.5中取$a=0$.
(2) 设矩阵A的第i和j行相同, 则

$$\det(A)\xlongequal{r_i\leftrightarrow r_j}-\det(A)\Longrightarrow\det(A)=0.$$

(3) 提出比例系数, 再用(2). □

如果行列式的值为零, 不能得到该行列式必有两行或两列成比例, 例如

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4\end{vmatrix}=0.$$

- **第III类初等变换**

下面考虑行(列)消去变换对方阵行列式的影响, 为此需要下面的拆行(列)公式.

命题4.7 (拆行(列)公式). 设$A, A_1, A_2 \in \mathbb{F}^{n\times n}$, A的第i行(列)为A_1的第i行(列)和A_2的第i行(列)之和, 而对于任意$j \neq i$, 这三个方阵有相同的第j行(列), 则有$|A| = |A_1| + |A_2|$, 即

$$\begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_{i-1} \\ \alpha_i' + \alpha_i'' \\ \alpha_{i+1} \\ \vdots \\ \alpha_n \end{vmatrix} = \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_{i-1} \\ \alpha_i' \\ \alpha_{i+1} \\ \vdots \\ \alpha_n \end{vmatrix} + \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_{i-1} \\ \alpha_i'' \\ \alpha_{i+1} \\ \vdots \\ \alpha_n \end{vmatrix}.$$

☞ **证明.** 设$A = (a_{kl})_{n\times n}$, $A_1 = (b_{kl})_{n\times n}$, $A_2 = (c_{kl})_{n\times n}$, 则

$$\begin{cases} a_{il} = b_{il} + c_{il}, & l = 1, 2, \ldots, n, \\ a_{kl} = b_{kl} = c_{kl}, & \forall k \neq i, \forall l. \end{cases}$$

所以有

$$\begin{aligned} |A| &= \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots a_{ip_i}\cdots a_{np_n} \\ &= \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots (b_{ip_i} + c_{ip_i})\cdots a_{np_n} \\ &= \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots b_{ip_i}\cdots a_{np_n} \\ &\quad + \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) a_{1p_1}\cdots c_{ip_i}\cdots a_{np_n} \\ &= \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) b_{1p_1}\cdots b_{np_n} + \sum_{p_1,\ldots,p_n\in S_n} \delta(p_1,\ldots,p_n) c_{1p_1}\cdots c_{np_n} \\ &= |A_1| + |A_2|. \end{aligned}$$

证毕. □

注意, 使用拆行(列)公式时, 每次只拆一行(列), 而保持其它行(列)不变. 请比较

$$\begin{vmatrix} a+x & b+y \\ c+z & d+w \end{vmatrix} = \begin{vmatrix} a & b \\ c+z & d+w \end{vmatrix} + \begin{vmatrix} x & y \\ c+z & d+w \end{vmatrix}$$
$$= \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b \\ z & w \end{vmatrix} + \begin{vmatrix} x & y \\ c & d \end{vmatrix} + \begin{vmatrix} x & y \\ z & w \end{vmatrix}$$

和

$$\begin{pmatrix} a+x & b+y \\ c+z & d+w \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} x & y \\ z & w \end{pmatrix}.$$

于是对于同阶方阵A和B, 通常$|A+B| \neq |A| + |B|$.

下面可以证明行列消去变换下方阵的行列式不改变.

定理4.8. 方阵的某行(列)乘数a再加到另一行(列), 其行列式不改变.

证明. 设

$$A=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}\overset{r_i+ar_j}{\leadsto} B=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_i+a\alpha_j\\ \vdots\\ \alpha_j\\ \vdots\\ \alpha_n\end{pmatrix},$$

于是得到

$$|B|\xlongequal{拆r_i}\begin{vmatrix}\alpha_1\\ \vdots\\ \alpha_i\\ \vdots\\ \alpha_j\\ \vdots\\ \alpha_n\end{vmatrix}+\begin{vmatrix}\alpha_1\\ \vdots\\ a\alpha_j\\ \vdots\\ \alpha_j\\ \vdots\\ \alpha_n\end{vmatrix}=|A|+0=|A|,$$

其中上面等式中第三个行列式为零是由于它的第i行和第j行成比例. □

我们看一个例子.

例4.9. 设$A,B\in\mathbb{F}^{3\times3}$, 其中$A=(\alpha_1,\alpha_2,\alpha_3)$, $B=(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1)$, $\alpha_1,\alpha_2,\alpha_3$为3维列向量. 已知$|A|=2$, 求$|B|$.

解. 我们有

$$|B|\xlongequal{拆c_1}|\alpha_1,\alpha_2+\alpha_3,\alpha_3+\alpha_1|+|\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1|.$$

由于

$$|\alpha_1,\alpha_2+\alpha_3,\alpha_3+\alpha_1|\xlongequal{c_3-c_1}|\alpha_1,\alpha_2+\alpha_3,\alpha_3|\xlongequal{c_2-c_3}|\alpha_1,\alpha_2,\alpha_3|=|A|$$

和

$$\begin{aligned}&|\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1|\xlongequal{c_2-c_1}|\alpha_2,\alpha_3,\alpha_3+\alpha_1|\xlongequal{c_3-c_2}|\alpha_2,\alpha_3,\alpha_1|\\ \xlongequal{c_2\leftrightarrow c_3}&-|\alpha_2,\alpha_1,\alpha_3|\xlongequal{c_1\leftrightarrow c_2}|\alpha_1,\alpha_2,\alpha_3|=|A|,\end{aligned}$$

所以有

$$|B|=2|A|=4.$$

也可以如下计算

$$\begin{aligned}|B|&\xlongequal{c_1-c_2+c_3}|2\alpha_1,\alpha_2+\alpha_3,\alpha_3+\alpha_1|=2|\alpha_1,\alpha_2+\alpha_3,\alpha_3+\alpha_1|\\ &\xlongequal{c_3-c_1}2|\alpha_1,\alpha_2+\alpha_3,\alpha_3|\xlongequal{c_2-c_3}2|\alpha_1,\alpha_2,\alpha_3|=2|A|=4,\end{aligned}$$

解毕. □

4.4.2 行列式的三角化

设$A \in \mathbb{F}^{n\times n}$, 则定理2.3证明中的算法一将

$$A \overset{r}{\sim} \text{阶梯形阵}R.$$

而且算法一中只用到了调行变换和行消去变换, 于是如果假设其中共用了M次调行变换, 则有

$$|A| = (-1)^M |R|.$$

注意到阶梯形的方阵R必是上三角阵.

命题4.9. 设$A \in \mathbb{F}^{n\times n}$, 则

(1) 利用调行变换和行消去变换, 可以把A化为上(下)三角阵R;

(2) 利用调列变换和列消去变换, 可以把A化为上(下)三角阵R.

进而, (1)或者(2)中如果共用了M次调行(列)变换, 则有$|A| = (-1)^M |R|$.

☞ **证明.** 类似于定理2.3证明中的算法一. □

我们看一个具体的例子.

例4.10. 计算行列式$D = \begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}$.

☞ **解.** 我们有

$$D \xlongequal{c_1 \leftrightarrow c_2} -\begin{vmatrix} 1 & 3 & -1 & 2 \\ 1 & -5 & 3 & -4 \\ 0 & 2 & 1 & -1 \\ -5 & 1 & 3 & -3 \end{vmatrix} \xlongequal[r_4+5r_1]{r_2-r_1} -\begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & -8 & 4 & -6 \\ 0 & 2 & 1 & -1 \\ 0 & 16 & -2 & 7 \end{vmatrix}$$

$$\xlongequal{r_2 \leftrightarrow r_3} \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & -8 & 4 & -6 \\ 0 & 16 & -2 & 7 \end{vmatrix} \xlongequal[r_4-8r_2]{r_3+4r_2} \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 8 & -10 \\ 0 & 0 & -10 & 15 \end{vmatrix}$$

$$= 10 \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 4 & -5 \\ 0 & 0 & -2 & 3 \end{vmatrix} \xlongequal{r_3 \leftrightarrow r_4} -10 \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 4 & -5 \end{vmatrix}$$

$$\xlongequal{r_4+2r_3} -10 \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 40.$$

在上面的计算过程中, 为了简化用行消去变换打洞时的计算量(不出现分数等), 我们并不完全按照定理2.3证明中的算法一进行上三角化. □

4.4.3 分块三角阵的行列式

上(下)三角阵的行列式公式是否可以推广到分块上(下)三角阵呢? 下面我们考虑这个问题. 设

$$M=\begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

其中$A\in\mathbb{F}^{p\times p}$, $C\in\mathbb{F}^{q\times q}$, 我们要计算$M$的行列式. 用行初等变换将$M$上三角化, 可以分别将$A$和$C$上三角化. 具体地, 设

$$A \underset{\text{行消去变换}}{\overset{\text{调行变换}}{\sim}} \begin{pmatrix} a_{11} & \cdots & * \\ & \ddots & \vdots \\ & & a_{pp} \end{pmatrix}, \quad N_1=\text{调行变换次数},$$

$$C \underset{\text{行消去变换}}{\overset{\text{调行变换}}{\sim}} \begin{pmatrix} c_{11} & \cdots & * \\ & \ddots & \vdots \\ & & c_{qq} \end{pmatrix}, \quad N_2=\text{调行变换次数}.$$

在M中, 分别对A和C所在的行做上面一样的行初等变换, 得到

$$M=\begin{pmatrix} A & B \\ O & C \end{pmatrix} \underset{\text{行消去变换}}{\overset{\text{调行变换}}{\sim}} \begin{pmatrix} a_{11} & \cdots & * & & & \\ & \ddots & \vdots & & * & \\ & & a_{pp} & & & \\ & & & c_{11} & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & c_{qq} \end{pmatrix},$$

则所用的调行变换的次数为N_1+N_2. 所以有

$$\begin{aligned} |M| =&(-1)^{N_1+N_2}\begin{vmatrix} a_{11} & \cdots & * & & & \\ & \ddots & \vdots & & * & \\ & & a_{pp} & & & \\ & & & c_{11} & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & c_{qq} \end{vmatrix} \\ =&(-1)^{N_1+N_2}a_{11}\cdots a_{pp}c_{11}\cdots c_{qq} \\ =&(-1)^{N_1}a_{11}\cdots a_{pp}\cdot(-1)^{N_2}c_{11}\cdots c_{qq}=|A||C|. \end{aligned}$$

于是我们得到对角块是方阵的分块上(下)三角阵的行列式公式.

命题4.10. 设A,C是方阵, B,F是矩阵, 则有

$$\begin{vmatrix} A & B \\ O & C \end{vmatrix}=|A||C|, \qquad \begin{vmatrix} A & O \\ F & C \end{vmatrix}=|A||C|.$$

进而, 如果$A_1, A_2, \ldots, A_s$是方阵, 则有

$$\begin{vmatrix} A_1 & * & \cdots & * \\ & A_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & A_s \end{vmatrix} = |A_1||A_2|\cdots|A_s|, \quad \begin{vmatrix} A_1 & & & \\ * & A_2 & & \\ \vdots & \vdots & \ddots & \\ * & * & \cdots & A_s \end{vmatrix} = |A_1||A_2|\cdots|A_s|,$$

和

$$|\operatorname{diag}(A_1, A_2, \ldots, A_s)| = |A_1||A_2|\cdots|A_s|.$$

☞ **证明.** 取转置, 可从分块上三角阵的行列式公式得到分块下三角阵的行列式公式. 而有s个对角块的分块上三角阵的行列式公式可以对s归纳得到. □

4.4.4 方阵乘积的行列式

设$A, B \in \mathbb{F}^{n\times n}$, $a \in \mathbb{F}$, 我们知道下面结果成立

$$\begin{aligned} &|A^T| = |A|, \\ &|aA| = a^n|A|, \\ &\text{通常 } |A+B| \neq |A| + |B|. \end{aligned}$$

那么, $|AB|$和$|A|, |B|$有什么关系吗? 我们来研究这个问题.

设$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$和$B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$是二阶方阵, 则

$$AB = \begin{pmatrix} ax+bz & ay+bw \\ cx+dz & cy+dw \end{pmatrix}.$$

于是

$$\begin{aligned} |AB| =&(ax+bz)(cy+dw) - (ay+bw)(cx+dz) \\ =&(acxy+adxw+bcyz+bdzw) - (acxy+adyz+bcxw+bdzw) \\ =&ad(xw-yz) + bc(yz-xw) \\ =&(ad-bc)(xw-yz) = |A||B|. \end{aligned}$$

再设$A = \begin{pmatrix} a_1 & * & \cdots & * \\ & a_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & a_n \end{pmatrix}$和$B = \begin{pmatrix} b_1 & * & \cdots & * \\ & b_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & b_n \end{pmatrix}$都是上三角阵, 则

$$AB = \begin{pmatrix} a_1b_1 & * & \cdots & * \\ & a_2b_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & a_nb_n \end{pmatrix}.$$

于是

$$|AB| = (a_1b_1)(a_2b_2)\cdots(a_nb_n) = (a_1a_2\cdots a_n)(b_1b_2\cdots b_n) = |A||B|.$$

在这两种特殊情形下都成立

$$|AB| = |A||B|$$

这一神奇[3]的等式, 这是一般规律.

定理4.11. 设A和B是同阶方阵, 则有

$$|AB| = |A||B|.$$

进而, 设$A_1, A_2, \ldots, A_s$是同阶方阵, 则有

$$|A_1A_2\cdots A_s| = |A_1||A_2|\cdots|A_s|.$$

☞ **证明.** 只需证第一个等式, 第二个等式可以对s归纳证明. 设$A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$. 定义分块矩阵

$$C = \begin{pmatrix} A & O \\ -E_n & B \end{pmatrix},$$

则由命题4.10, $|C| = |A||B|$. 另一方面, 对C进行如下初等变换

$$C = \begin{pmatrix} a_{11} & \cdots & a_{1n} & & & \\ \vdots & & \vdots & & & \\ a_{n1} & \cdots & a_{nn} & & & \\ -1 & & & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ & & -1 & b_{n1} & \cdots & b_{nn} \end{pmatrix} \underbrace{\scriptstyle c_{n+1}+b_{11}c_1+\cdots+b_{n1}c_n}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & \sum\limits_{k=1}^n a_{1k}b_{k1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & \sum\limits_{k=1}^n a_{nk}b_{k1} & 0 & \cdots & 0 \\ -1 & & & 0 & b_{12} & \cdots & b_{1n} \\ & \ddots & & \vdots & \vdots & & \vdots \\ & & -1 & 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix} \underbrace{\scriptstyle c_{n+j}+b_{1j}c_1+\cdots+b_{nj}c_n}_{j=2,\ldots,n}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & \\ \vdots & & \vdots & AB \\ a_{n1} & \cdots & a_{nn} & \\ -1 & & & \\ & \ddots & & \\ & & -1 & \end{pmatrix} \underbrace{\scriptstyle r_i\leftrightarrow r_{n+i}}_{i=1,\ldots,n} \begin{pmatrix} -1 & & & \\ & \ddots & & \\ & & -1 & \\ a_{11} & \cdots & a_{1n} & \\ \vdots & & \vdots & AB \\ a_{n1} & \cdots & a_{nn} & \end{pmatrix}.$$

由于上面共用了n个调行变换, 所以

$$|C| = (-1)^n \begin{vmatrix} -E & O \\ A & AB \end{vmatrix} = (-1)^n|-E||AB| = |AB|.$$

[3]从行列式的几何意义—对应的线性变换下体积的伸缩比, 可以得到这一等式成立的一个理由.

最后就得到[4] $|AB| = |A||B|$. □

例4.11. 设$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $n \in \mathbb{N}$, 求$|A^n|$.

☞ **解.** 我们有

$$|A^n| = |A|^n = (-2)^n.$$

解毕. □

例4.12 (例4.9)**.** 设$A, B \in \mathbb{F}^{3\times 3}$, 其中$A = (\alpha_1, \alpha_2, \alpha_3)$, $B = (\alpha_1+\alpha_2, \alpha_2+\alpha_3, \alpha_3+\alpha_1)$, $\alpha_1, \alpha_2, \alpha_3$为3维列向量. 已知$|A| = 2$, 求$|B|$.

☞ **解.** 这里给出另一种解法. 由定义有

$$B = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = A \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

而(也可用对角线法则)

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} \xlongequal{r_2 - r_1} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} \xlongequal{r_3 - r_2} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{vmatrix} = 2,$$

所以$|B| = 2|A| = 4$. □

4.4.5 可逆阵的行列式

我们考察可逆阵的行列式. 设$A \in \mathbb{F}^{n\times n}$为可逆阵, 则有$AA^{-1} = E_n$. 取行列式得

$$|A||A^{-1}| = |AA^{-1}| = |E| = 1.$$

所以$|A| \neq 0$, 且$|A^{-1}| = |A|^{-1}$. 而且可以证明当$|A| \neq 0$时, 必有A可逆.

定理4.12. 设$A \in \mathbb{F}^{n\times n}$, 则

$$A\text{可逆} \iff |A| \neq 0.$$

且当A可逆时, $|A^{-1}| = |A|^{-1}$.

☞ **证明.** 必要性已证, 下证充分性. 设A不可逆, 而R是与A行等价的简化阶梯形阵, 则R有全零行, 所以$|R| = 0$. 又存在可逆阵P, 使得$A = PR$. 于是

$$|A| = |PR| = |P||R| = 0.$$

证毕. □

上面定理就是命题"$a \in \mathbb{F}$可逆$\iff a \neq 0$" 的矩阵对应, 即可用行列式是否为零来判别方阵是否可逆.

结合定理3.15, 可以得到如下系数矩阵为方阵时线性方程组有唯一解的行列式判别法.

[4] 在这个证明中, 我们对同一个矩阵C按照不同的方法去计算行列式, 从而得到一个关于行列式的等式. 这是证明恒等式常用的一种方法: 对一个对象去算同一个量, 按照不同的看法分别得到等式的两边. 那么, 这里如何想到(找到)这个矩阵C? 一个可能的回答是分块上三角阵的行列式公式的右边是方阵行列式的乘积, 和这里要证明的等式的右边相同.

推论4.13. 设有线性方程组$AX=\beta$, 其中A是方阵. 则

(1) 有唯一解$\Longleftrightarrow |A|\neq 0$;

(2) 特别的, $AX=0$有非零解$\Longleftrightarrow |A|=0$.

最后看一个例子.

例4.13. 参数a取何值时, 线性方程组$\begin{cases}(5-a)x+2y+2z=0,\\ 2x+(6-a)y=0,\\ 2x+(4-a)z=0\end{cases}$有非零解?

☞ **解.** 系数矩阵的行列式

$$\begin{aligned}\begin{vmatrix}5-a & 2 & 2\\ 2 & 6-a & 0\\ 2 & 0 & 4-a\end{vmatrix} &=(5-a)(6-a)(4-a)-4(4-a)-4(6-a)\\ &=(5-a)(2-a)(8-a),\end{aligned}$$

所以当且仅当$a=2,5,8$时方程组有非零解. □

习题4.4

A1. 计算下面的行列式

(1) $\begin{vmatrix}-ab & ac & ae\\ bd & -cd & de\\ bf & cf & -ef\end{vmatrix}$; (2) $\begin{vmatrix}2 & 0 & 1 & -1\\ 1 & 2 & 0 & 1\\ 0 & -1 & 3 & -2\\ -1 & 3 & 1 & 2\end{vmatrix}$.

A2. 证明: $\begin{vmatrix}ax+by & ay+bz & az+bx\\ ay+bz & az+bx & ax+by\\ az+bx & ax+by & ay+bz\end{vmatrix}=(a^3+b^3)\begin{vmatrix}x & y & z\\ y & z & x\\ z & x & y\end{vmatrix}$.

A3. 计算下面行列式:

(1) $\begin{vmatrix}a & b & c & d\\ a & a+b & a+b+c & a+b+c+d\\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d\\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d\end{vmatrix}$;

(2) $|A|$, 其中$A=(a_{ij})$是n阶方阵, 满足

$$\begin{cases}a_{i1}=a, & i=1,2,\dots,n,\\ a_{ij}=a_{i,j-1}+a_{i-1,j}, & i,j=2,\dots,n.\end{cases}$$

A4. 证明: 奇数阶反对称阵的行列式为零.

A5. 设$n\geqslant 2$, 利用行列式证明S_n中奇排列和偶排列各半.

A6. 设α,β,γ是3维列向量, 矩阵$A=(\alpha,\beta,\gamma)$, $B=(\alpha+2\beta,\beta+2\gamma,\gamma+2\alpha)$. 设$|A|=2$, 求$|B|$.

A7. 设$\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$是4维列向量, 矩阵$A = (\alpha_1, \alpha_2, 2\alpha_3, \beta_1)$, $B = (\alpha_3, \alpha_2, \alpha_1, \beta_2)$. 如果行列式$|A| = -2$, $|B| = -4$, 求行列式$|A+B|$.

A8. 设3阶方阵A, B满足$A^2B - A - B = E$, 其中$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ -2 & 0 & 1 \end{pmatrix}$, 求$|B|$.

A9. 用行列式判断方阵$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & -1 \\ 1 & -2 & 1 \end{pmatrix}$的可逆性.

A10. 设A是n阶复方阵, 证明: 存在无穷个复数λ, 使得$\lambda E_n + A$可逆.

A11. 参数a取何值时, 线性方程组$\begin{cases} x_2 + x_3 = ax_1, \\ x_1 + x_3 = ax_2, \\ x_1 + x_2 = ax_3 \end{cases}$ 有非零解?

A12. 设a, b, c和d是不全为零的实数. 证明: 线性方程组

$$\begin{cases} ax + by + cz + dt = 0, \\ bx - ay + dz - ct = 0, \\ cx - dy - az + bt = 0, \\ dx + cy - bz - at = 0 \end{cases}$$

具有唯一解, 其中x, y, z和t是未知数.

§ 4.5 分块矩阵的初等变换和行列式

由于我们常常需要对分块矩阵进行分块矩阵的初等变换, 所以在给出更多的行列式计算的例子之前, 我们先讨论一下在分块矩阵的初等变换下, 分块矩阵的行列式如何改变. 同样的, 在阅读下文前, 读者应该自己先研究一下这个问题; 我们省略得到结论的发现之旅.

- **第I类分块矩阵的初等变换**

定理4.14. *设方阵$A = (A_{ij})_{s\times r}$是分块矩阵, 其中A_{ij}是$m_i \times n_j$子阵.*

(1) *设$1 \leqslant i < j \leqslant s$, 记$l = m_{i+1} + m_{i+2} + \cdots + m_{j-1}$, 如果*

$$A = \begin{array}{c} \begin{pmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{i1} & \cdots & A_{ir} \\ \vdots & & \vdots \\ A_{j1} & \cdots & A_{jr} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{sr} \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \}m_i \\ \}l \\ \}m_j \\ \\ \\ \end{array} \xrightarrow[\sim]{r_i \leftrightarrow r_j} \begin{pmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{j1} & \cdots & A_{jr} \\ \vdots & & \vdots \\ A_{i1} & \cdots & A_{ir} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{sr} \end{pmatrix} = B,$$

则有$|B| = (-1)^{m_i m_j + m_i l + m_j l}|A|$;

(2) 设$1 \leqslant i < j \leqslant r$, 记$l = n_{i+1} + n_{i+2} + \cdots + n_{j-1}$, 如果

$$A \overset{c_i \leftrightarrow c_j}{\rightsquigarrow} B,$$

则$|B| = (-1)^{n_i n_j + n_i l + n_j l}|A|$.

☞ **证明.** 只需证明(1). 由条件$B = P_{ij}A$, 其中

$$P_{ij} = \begin{pmatrix} E & & & & \\ & O & O & E_{m_j} & \\ & O & E_l & O & \\ & E_{m_i} & O & O & \\ & & & & E \end{pmatrix}.$$

所以有$|B| = |P_{ij}||A|$, 而

$$|P_{ij}| = \begin{vmatrix} O & O & E_{m_j} \\ O & E_l & O \\ E_{m_i} & O & O \end{vmatrix}.$$

将第m_j行依次与第$m_j+1, m_j+2, \ldots, m_i+m_j+l$行交换, 再将第$m_j-1$行依次与第$m_j, m_j+1, \ldots, m_i+m_j+l-1$行交换, 依次类推, 最后将第1行依次与第$2, 3, \ldots, m_i+l+1$ 行交换, 得到

$$P_{ij} = (-1)^{m_j(m_i+l)} \begin{vmatrix} O & E_l & O \\ E_{m_i} & O & O \\ O & O & E_{m_j} \end{vmatrix} = (-1)^{m_j(m_i+l)} \begin{vmatrix} O & E_l \\ E_{m_i} & O \end{vmatrix}.$$

又将第l行依次与第$l+1, l+2, l+m_i$行交换, 将第$l-1$行依次与第$l, l+1, \ldots, l+m_i-1$行交换, 依次类推, 最后将第1行依次与第$2, 3, \ldots, m_i+1$行交换, 得到

$$|P_{ij}| = (-1)^{m_j(m_i+l)}(-1)^{m_i l} \begin{vmatrix} E_{m_i} & O \\ O & E_l \end{vmatrix} = (-1)^{m_i m_j + m_i l + m_j l}.$$

所以$|B| = (-1)^{m_i m_j + m_i l + m_j l}|A|$. □

- **第II类分块矩阵的初等变换**

定理4.15. 设方阵A是分块矩阵, 如果P是方阵, 而

$$A \overset{P \times r_i}{\rightsquigarrow} B, \quad (A \overset{c_i \times P}{\rightsquigarrow} B)$$

则有$|B| = |P||A|$.

☞ **证明.** 由条件, $B = P_i(P)A$, 其中

$$P_i(P) = \begin{pmatrix} E & & \\ & P & \\ & & E \end{pmatrix}.$$

所以$|B| = |P_i(P)||A| = |P||A|$. □

- **第III类分块矩阵的初等变换**

定理4.16. 设方阵A是分块矩阵, 如果

$$A \overset{r_j+M\times r_i}{\rightsquigarrow} B, \quad (A \overset{c_i+c_j\times M}{\rightsquigarrow} B)$$

则$|B|=|A|$.

☞ **证明.** 由条件, $B=P(i(M),j)A$, 其中

$$P(i(M),j)=\begin{pmatrix} E & & & & \\ & E & & & \\ & \vdots & \ddots & & \\ & M & \cdots & E & \\ & & & & E \end{pmatrix}.$$

所以$|B|=|P(i(M),j)||A|=|A|$. □

最后是一个例子.

例4.14. 记$S=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, 其中$A,D$是方阵, 且$A$可逆. 由于

$$S \overset{r_2-CA^{-1}r_1}{\rightsquigarrow} \begin{pmatrix} A & B \\ O & D-CA^{-1}B \end{pmatrix},$$

所以$|S|=|A||D-CA^{-1}B|$.

习题4.5

A1. 设A和B都是n阶方阵, 证明: $\begin{vmatrix} A & B \\ B & A \end{vmatrix}=|A+B||A-B|$.

A2. 设$A,B\in\mathbb{C}^{n\times n}$, 证明: $\begin{vmatrix} A & -B \\ B & A \end{vmatrix}=|A+iB||A-iB|$. 进而, 如果$AB=BA$, 则$\begin{vmatrix} A & -B \\ B & A \end{vmatrix}=|A^2+B^2|$.

A3. 设$A,B,C,D\in\mathbb{F}^{n\times n}$, 其中$A$可逆且$AC=CA$, 证明: $\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|$.

B1. 设$A,B,C,D\in\mathbb{F}^{n\times n}$, 其中$AC=CA$, 证明: $\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|$.

B2. (Burnside) 设$A=(a_{ij})$是n阶反对称阵, 将a_{ij}看成未定元, 证明: 当n是偶数时, $|A|$是完全平方.

§ 4.6 行列式计算之例一

本节给出一些行列式计算的例子, 并给出一些特殊的计算技巧. 虽然与故事情节无关, 但也是一些风景.

例4.15. 计算$2n$阶行列式$D_{2n}=\begin{vmatrix} a &&&&& b\\ &\ddots&&&\iddots&\\ &&a&b&&\\ &&c&d&&\\ &\iddots&&&\ddots&\\ c&&&&&d\end{vmatrix}$.

☞ **解.** 用分块矩阵的调行和调列变换, 有

$$D_{2n}=\left|\begin{array}{cccccccc} a&&&&&&&b\\ \hdashline &a&&&&&b&\\ &&\ddots&&&\iddots&&\\ &&&a&b&&&\\ &&&c&d&&&\\ &&\iddots&&&\ddots&&\\ &c&&&&&d&\\ \hdashline c&&&&&&&d \end{array}\right| \xlongequal{r_2\leftrightarrow r_3}(-1)^{2n-2}\left|\begin{array}{c:cccccc:c} a&&&&&&&b\\ c&&&&&&&d\\ &a&&&&&b&\\ &&\ddots&&&\iddots&&\\ &&&a&b&&&\\ &&&c&d&&&\\ &&\iddots&&&\ddots&&\\ &c&&&&&d& \end{array}\right|$$

$$\xlongequal{c_2\leftrightarrow c_3}(-1)^{2n-2}(-1)^{2n-2}\begin{vmatrix} a&b&&&&&&\\ c&d&&&&&&\\ &&a&&&&&b\\ &&&\ddots&&&\iddots&\\ &&&&a&b&&\\ &&&&c&d&&\\ &&&\iddots&&&\ddots&\\ &&c&&&&&d \end{vmatrix}.$$

所以得到递推公式[5]

$$D_{2n}=\begin{vmatrix} a&b\\ c&d\end{vmatrix}D_{2(n-1)}=(ad-bc)D_{2(n-1)}.$$

不断利用该递推公式以及$D_2=ad-bc$, 就得

$$D_{2n}=(ad-bc)^2D_{2(n-2)}=\cdots=(ad-bc)^{n-1}D_2=(ad-bc)^n.$$

解毕. □

例4.16. 证明爪形行列式公式

$$\begin{vmatrix} a_1&x_2&\cdots&x_n\\ y_2&a_2&&\\ \vdots&&\ddots&\\ y_n&&&a_n\end{vmatrix}=\left(a_1-\sum_{i=2}^n\frac{x_iy_i}{a_i}\right)a_2a_3\cdots a_n.$$

上面公式的右边事实上无分式.

[5]也可以对第1和$2n$行Laplace展开得到递推公式.

☞ **证明.** 记爪形行列式为D_n. 当$a_2,\dots,a_n$都不为零时, 有

$$D_n \xrightarrow[i=2,\dots,n]{c_1-\frac{y_i}{a_i}c_i} \begin{vmatrix} a_1-\sum\limits_{i=2}^n \frac{x_iy_i}{a_i} & x_2 & \cdots & x_n \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{vmatrix} = \left(a_1-\sum_{i=2}^n \frac{x_iy_i}{a_i}\right)a_2a_3\cdots a_n.$$

如果存在$2\leqslant i\leqslant n$, 使得$a_i=0$, 将第i行依次与第$i-1,i-2,\dots,1$行交换, 得

$$\begin{aligned} D_n =& (-1)^{i-1}\begin{vmatrix} y_i & & & & & & & \\ a_1 & x_2 & \cdots & x_{i-1} & x_i & x_{i+1} & \cdots & x_n \\ y_2 & a_2 & & & & & & \\ \vdots & & \ddots & & & & & \\ y_{i-1} & & & a_{i-1} & & & & \\ y_{i+1} & & & & & a_{i+1} & & \\ \vdots & & & & & & \ddots & \\ y_n & & & & & & & a_n \end{vmatrix} \\ =& (-1)^{i-1}y_i\begin{vmatrix} x_2 & \cdots & x_{i-1} & x_i & x_{i+1} & \cdots & x_n \\ a_2 & & & & & & \\ & \ddots & & & & & \\ & & a_{i-1} & & & & \\ & & & & a_{i+1} & & \\ & & & & & \ddots & \\ & & & & & & a_n \end{vmatrix}. \end{aligned}$$

再将第$i-1$列依次与第$i-2,\dots,1$列交换, 得

$$\begin{aligned} D_n =& (-1)^{i-1}y_i(-1)^{i-2}\begin{vmatrix} x_i & x_2 & \cdots & x_{i-1} & x_{i+1} & \cdots & x_n \\ & a_2 & & & & & \\ & & \ddots & & & & \\ & & & a_{i-1} & & & \\ & & & & a_{i+1} & & \\ & & & & & \ddots & \\ & & & & & & a_n \end{vmatrix} \\ =& -x_iy_ia_2\cdots a_{i-1}a_{i+1}\cdots a_n. \end{aligned}$$

这与一般公式中令$a_i=0$得到的结果一致. 所以爪形行列式的公式成立.

上面的过程也可以用矩阵分块实现. 设D_n所对应的方阵为A. 当$a_2,\dots,a_n$都不为零时, A有分块

$$A=\begin{pmatrix} a_1 & X \\ Y & A_1 \end{pmatrix},\quad A_1=\mathrm{diag}(a_2,\dots,a_n), X=(x_2,\dots,x_n), Y=(y_2,\dots,y_n)^T.$$

由初等变换

$$A \overset{c_1-c_2A^{-1}Y}{\sim} \begin{pmatrix} a_1-XA_1^{-1}Y & X \\ O & A_1 \end{pmatrix},$$

得

$$D_n=(a_1-XA_1^{-1}Y)|A_1|=\left(a_1-\sum_{i=2}^{n}\frac{x_iy_i}{a_i}\right)a_2a_3\cdots a_n.$$

如果存在$2\leqslant i\leqslant n$, 使得$a_i=0$, 进行分块矩阵的换行和换列变换

$$\begin{aligned}
|D_n|=&\left|\begin{array}{cccccccc}
a_1 & x_2 & \cdots & x_{i-1} & x_i & x_{i+1} & \cdots & x_n\\
y_2 & a_2 & & & & & & \\
\vdots & & \ddots & & & & & \\
y_{i-1} & & & a_{i-1} & & & & \\
\hdashline
y_i & & & & & & & \\
\hdashline
y_{i+1} & & & & & a_{i+1} & & \\
\vdots & & & & & & \ddots & \\
y_n & & & & & & & a_n
\end{array}\right|\\
\xlongequal{r_1\leftrightarrow r_2}&(-1)^{i-1}\left|\begin{array}{cccccccc}
y_i & & & & & & & \\
a_1 & x_2 & \cdots & x_{i-1} & x_i & x_{i+1} & \cdots & x_n\\
y_2 & a_2 & & & & & & \\
\vdots & & \ddots & & & & & \\
y_{i-1} & & & a_{i-1} & & & & \\
y_{i+1} & & & & & a_{i+1} & & \\
\vdots & & & & & & \ddots & \\
y_n & & & & & & & a_n
\end{array}\right|\\
=&(-1)^{i-1}y_i\left|\begin{array}{ccc:c:ccc}
x_2 & \cdots & x_{i-1} & x_i & x_{i+1} & \cdots & x_n\\
a_2 & & & & & & \\
 & \ddots & & & & & \\
 & & a_{i-1} & & & & \\
 & & & & a_{i+1} & & \\
 & & & & & \ddots & \\
 & & & & & & a_n
\end{array}\right|\\
\xlongequal{c_1\leftrightarrow c_2}&(-1)^{i-1}y_i(-1)^{i-2}\left|\begin{array}{ccccccc}
x_i & x_2 & \cdots & x_{i-1} & x_{i+1} & \cdots & x_n\\
 & a_2 & & & & & \\
 & & \ddots & & & & \\
 & & & a_{i-1} & & & \\
 & & & & a_{i+1} & & \\
 & & & & & \ddots & \\
 & & & & & & a_n
\end{array}\right|\\
=&-x_iy_ia_2\cdots a_{i-1}a_{i+1}\cdots a_n.
\end{aligned}$$

这与一般公式中令$a_i=0$得到的结果一致. □

例4.17. 计算n阶行列式$D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & x \end{vmatrix}$.

☞ **解.** 这个行列式有许多算法, 这里给出几种(还可以利用方阵的特征值来计算, 参看第七章例7.47).

可以如下计算

$$D_n \xlongequal[i=2,\ldots,n]{r_1+r_i} \begin{vmatrix} x+(n-1)a & x+(n-1)a & \cdots & x+(n-1)a \\ a & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & x \end{vmatrix}$$

$$\xlongequal[i=2,\ldots,n]{c_i-c_1} \begin{vmatrix} x+(n-1)a & & & \\ a & x-a & & \\ \vdots & & \ddots & \\ a & & & x-a \end{vmatrix} = [x+(n-1)a](x-a)^{n-1}.$$

上面第一步的变换称为行累加法(所有其它行加到一行), 第二步的变换称为主列消法(用某一列去消其它列).

也可以如下计算

$$D_n \xlongequal[i=2,\ldots,n]{r_i-r_1} \begin{vmatrix} x & a & \cdots & a \\ a-x & x-a & & \\ \vdots & & \ddots & \\ a-x & & & x-a \end{vmatrix}$$

$$\xlongequal[i=2,\ldots,n]{c_1+c_i} \begin{vmatrix} x+(n-1)a & a & \cdots & a \\ & x-a & & \\ & & \ddots & \\ & & & x-a \end{vmatrix} = [x+(n-1)a](x-a)^{n-1}.$$

上面第一步用了主行消法得到爪形行列式, 再用了列累加法.

还可以如下计算

$$D_n \xlongequal{\text{拆}r_1} \begin{vmatrix} x-a & 0 & \cdots & 0 \\ a & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & x \end{vmatrix} + \begin{vmatrix} a & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & x \end{vmatrix}.$$

右边第一个行列式为$(x-a)D_{n-1}$, 第二个行列式用主行消法$(r_i - r_1, i = 2, \ldots, n)$为

$$\begin{vmatrix} a & a & \cdots & a \\ & x-a & & \\ & & \ddots & \\ & & & x-a \end{vmatrix} = a(x-a)^{n-1}.$$

于是我们得到递推公式

$$D_n = (x-a)D_{n-1} + a(x-a)^{n-1}.$$

反复利用该公式, 得

$$\begin{aligned}
D_n =&(x-a)^2 D_{n-2} + (x-a)a(x-a)^{n-2} + a(x-a)^{n-1}\\
=&(x-a)^2 D_{n-2} + 2a(x-a)^{n-1}\\
=&\cdots\\
=&(x-a)^{n-1}D_1 + (n-1)a(x-a)^{n-1}\\
=&x(x-a)^{n-1} + (n-1)a(x-a)^{n-1},
\end{aligned}$$

其中用了事实$D_1 = x$. □

我们看上面例子的两个变形.

例4.18. 计算n阶行列式

$$(1)\ D_n = \begin{vmatrix} x_1 & a_2 & \cdots & a_n \\ a_1 & x_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & x_n \end{vmatrix}; \qquad (2)\ D_n = \begin{vmatrix} x & a & \cdots & a \\ b & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & x \end{vmatrix}.$$

☞ **解.** (1) 用第1行主行消去

$$D_n \xlongequal[i=2,\dots,n]{r_i - r_1} \begin{vmatrix} x_1 & a_2 & \cdots & a_n \\ a_1 - x_1 & x_2 - a_2 & & \\ \vdots & & \ddots & \\ a_1 - x_1 & & & x_n - a_n \end{vmatrix},$$

再用爪形行列式公式得

$$\begin{aligned}
D_n &= \left(x_1 - \sum_{i=2}^n \frac{a_i(a_1 - x_1)}{x_i - a_i}\right)(x_2 - a_2)\cdots(x_n - a_n)\\
&= \left(1 + \sum_{i=1}^n \frac{a_i}{x_i - a_i}\right)\prod_{i=1}^n (x_i - a_i).
\end{aligned}$$

(2) 拆第1行, 得

$$D_n = \begin{vmatrix} x-b & 0 & \cdots & 0 \\ b & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & x \end{vmatrix} + \begin{vmatrix} b & a & \cdots & a \\ b & x & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & x \end{vmatrix}.$$

上右边第一个行列式为$(x-b)D_{n-1}$, 第二个行列式$r_i - r_1\ (i = 2, \dots, n)$后为

$$\begin{vmatrix} b & a & \cdots & a \\ 0 & x-a & & \\ \vdots & \vdots & \ddots & \\ 0 & b-a & \cdots & x-a \end{vmatrix} = b \begin{vmatrix} x-a & & \\ \vdots & \ddots & \\ b-a & \cdots & x-a \end{vmatrix} = b(x-a)^{n-1}.$$

于是得到递推公式

$$D_n = (x-b)D_{n-1} + b(x-a)^{n-1}.$$

将上式中的a和b互换, 注意到此时D_n和D_{n-1}变成对应的转置矩阵的行列式而不改变, 得

$$D_n = (x-a)D_{n-1} + a(x-b)^{n-1}.$$

利用两个递推公式消去D_{n-1}, 得

$$(a-b)D_n = a(x-b)^n - b(x-a)^n.$$

所以当$a \neq b$时有

$$D_n = \frac{a(x-b)^n - b(x-a)^n}{a-b}.$$

当$a=b$时由前面例子, $D_n = [x+(n-1)a](x-a)^{n-1}$. 但是, 利用微积分中求极限的Hôspital法则我们有

$$\begin{aligned}&\lim_{b\to a}\frac{a(x-b)^n - b(x-a)^n}{a-b} = \lim_{b\to a}\frac{-na(x-b)^{n-1}-(x-a)^n}{-1}\\ =&na(x-a)^{n-1}+(x-a)^n = (na+x-a)(x-a)^{n-1} = [x+(n-1)a](x-a)^{n-1}.\end{aligned}$$

所以总有

$$D_n = \frac{a(x-b)^n - b(x-a)^n}{a-b}.$$

解毕. □

例4.19. 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$, 证明: $|E_m - AB| = |E_n - BA|$.

☞ **证明.** 令$C = \begin{pmatrix} E_m & A \\ B & E_n \end{pmatrix}$, 则有分块矩阵的初等变换

$$C \overset{r_2 - Br_1}{\sim} \begin{pmatrix} E_m & A \\ O & E_n - BA \end{pmatrix}, \qquad C \overset{c_1 - c_2 B}{\sim} \begin{pmatrix} E_m - AB & A \\ O & E_n \end{pmatrix},$$

得

$$\begin{vmatrix} E_m & A \\ O & E_n - BA \end{vmatrix} = \begin{vmatrix} E_m - AB & A \\ O & E_n \end{vmatrix},$$

即$|E_m - AB| = |E_n - BA|$. □

将A用$-A$代替, 也成立

$$|E_m + AB| = |E_n + BA|.$$

最后举些应用上例结果的例子.

例4.20. 计算n阶行列式

(1) $$D_n = \begin{vmatrix} 1+a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & 1+a_nb_n \end{vmatrix};$$

(2) $$D_n = \begin{vmatrix} a_1b_1 & 1+a_1b_2 & \cdots & 1+a_1b_n \\ 1+a_2b_1 & a_2b_2 & \cdots & 1+a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+a_nb_1 & 1+a_nb_2 & \cdots & a_nb_n \end{vmatrix}.$$

☞ **解.** (1) 设D_n为方阵A对应的行列式, 由于

$$A = E_n + \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (b_1, \ldots, b_n),$$

所以

$$D_n = 1 + (b_1, \ldots, b_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 1 + \sum_{i=1}^{n} a_i b_i.$$

也可以拆第1行得到D_n的递推公式.

(2) 设所给行列式对应矩阵为A, 则

$$\begin{aligned} A &= \begin{pmatrix} 1+a_1b_1 & 1+a_1b_2 & \cdots & 1+a_1b_n \\ 1+a_2b_1 & 1+a_2b_2 & \cdots & 1+a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+a_nb_1 & 1+a_nb_2 & \cdots & 1+a_nb_n \end{pmatrix} - E_n \\ &= \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} - E_n. \end{aligned}$$

于是有

$$\begin{aligned} D_n =& (-1)^n \left| E_n - \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \right| \\ =& (-1)^n \left| E_2 - \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \right| \\ =& (-1)^n \begin{vmatrix} 1-n & -\sum\limits_{i=1}^{n} a_i \\ -\sum\limits_{i=1}^{n} b_i & 1-\sum\limits_{i=1}^{n} a_ib_i \end{vmatrix} \\ =& (-1)^n \left[(1-n)\left(1-\sum_{i=1}^{n} a_ib_i\right) - \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right) \right]. \end{aligned}$$

解毕. □

习题4.6

A1. 计算下列行列式

(1) $\begin{vmatrix} 1 & 2 & \cdots & n-1 & n \\ 1 & -1 & \cdots & 0 & 0 \\ 0 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & n-1 & 1-n \end{vmatrix}$;

(2) $\begin{vmatrix} a & a+h & a+2h & \cdots & a+(n-2)h & a+(n-1)h \\ -a & a & & & & \\ & -a & a & & & \\ & & \ddots & \ddots & & \\ & & & -a & a & \\ & & & & -a & a \end{vmatrix}$;

(3) $\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix}$; (4) $\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix}$;

(5) $\begin{vmatrix} x_1 & a & \cdots & a \\ b & x_2 & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & x_n \end{vmatrix}$.

A2. 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$, λ是未定元. 证明:

$$\lambda^n|\lambda E_m - AB| = \lambda^m|\lambda E_n - BA|.$$

A3. 设A为n阶可逆矩阵, α为n维列向量, 证明:

$$\det(A - \alpha\alpha^T) = (1 - \alpha^T A^{-1}\alpha)\det(A).$$

A4. 设有n阶方阵$A = (a_{ij})$, 其中

$$a_{ij} = \begin{cases} x + a_i, & i = j, \\ a_j, & i \neq j. \end{cases}$$

求行列式$|A|$.

B1. 设$b_{ij} = (a_{i1} + a_{i2} + \cdots + a_{in}) - a_{ij}$, $1 \leqslant i, j \leqslant n$. 证明:

$$\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix} = (-1)^{n-1}(n-1)\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

如果$b_{ij} = (a_{i1} + a_{i2} + \cdots + a_{in}) - ca_{ij}$, 其中$1 \leqslant i, j \leqslant n$, $c \in \mathbb{F}$, 结论又怎样?

B2. 计算$n+1$阶行列式:

(1)
$$\begin{vmatrix} x & 1 & & & & & \\ -n & x-2 & 2 & & & & \\ & -(n-1) & x-4 & 3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -3 & x-(2n-4) & n-1 & \\ & & & & -2 & x-(2n-2) & n \\ & & & & & -1 & x-2n \end{vmatrix};$$

(2)
$$\begin{vmatrix} x+n & -n & & & & & \\ 1 & x+n-2 & -n+1 & & & & \\ & 2 & x+n-4 & -n+2 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & n-2 & x-n+4 & -2 & \\ & & & & n-1 & x-n+2 & -1 \\ & & & & & n & x-n \end{vmatrix}.$$

C1. 设$A\in\mathbb{R}^{2n\times 2n}$满足[6]

$$A\begin{pmatrix} O & E_n \\ -E_n & O \end{pmatrix}A^T=\begin{pmatrix} O & E_n \\ -E_n & O \end{pmatrix}.$$

证明: $|A|=1$.

C2. 设$A\in\mathbb{C}^{2n\times 2n}$满足

$$A\begin{pmatrix} O & E_n \\ -E_n & O \end{pmatrix}A^T=\begin{pmatrix} O & E_n \\ -E_n & O \end{pmatrix}.$$

证明: $|A|=1$.

§ 4.7 行列式的按行(列)展开

4.7.1 行列式的按行(列)展开

现在我们介绍行列式计算的降阶法. 按照我们的想法, 我们希望将n阶行列式用$n-1$阶行列式表示. 为了得到一些有益的提示, 先看三阶行列式如何用二阶行列式表示. 事实上, 我们有

$$\begin{aligned} D_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31} \\ &= a_{11}(a_{22}a_{33}-a_{23}a_{32})+a_{12}(a_{23}a_{31}-a_{21}a_{33})+a_{13}(a_{21}a_{32}-a_{22}a_{31}) \end{aligned}$$

[6]称满足这个条件的矩阵为辛方阵.

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix},$$

确实将三阶行列式用二阶行列式表示. 上面推导中, 我们将三阶行列式的六项按照一正一负分成三组, 于是等于第一行的三个元素与某些二阶行列式乘积的代数和. 那么, 这些元素乘怎样的二阶行列式? 可以看到, 都是把该元素所在行和列划去后, 剩下的二阶方阵的行列式. 最后还需确定代数和中的符号. 经过仔细分析后, 我们给出下面的定义.

定义4.7. 设$A = (a_{ij})_{n\times n} \in \mathbb{F}^{n\times n}$, 其中$n \geqslant 2$. 任取$1 \leqslant i, j \leqslant n$.

(1) 将A中第i行和第j列元素划去后, 剩下的$n-1$阶方阵的行列式称为(i,j)位置(或: a_{ij})的余子式, 记为M_{ij}. 即

$$M_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i1} & \cdots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

(2) (i,j)位置(或: a_{ij})的代数余子式A_{ij}定义为

$$A_{ij} := (-1)^{i+j} M_{ij}.$$

这里, 我们强调(i,j)位置的余子式和代数余子式, 是由于它们与A的第i行和第j列的具体元素没有关系(它们被划去了!), 而只与i,j有关系.

利用代数余子式的记号, 上面用二阶行列式表示三阶行列式的等式可以改为

$$D_3 = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13},$$

即D_3等于第一行的所有元素与对应位置的代数余子式乘积之和. 感兴趣的读者可以验证: 如果对三阶行列式的六项采用不同的一正一负的分组方式(共有6种分组方法), 则可知三阶行列式等于某一行(列)的三个元素与对应位置的代数余子式乘积之和. 于是我们有理由相信下面的定理成立.

定理4.17 (按行(列)展开定理)**.** 设$A = (a_{ij})_{n\times n} \in \mathbb{F}^{n\times n}$是$n$阶方阵, 其中$n \geqslant 2$, 则$|A|$等于某行(列)的所有元素与对应位置的代数余子式乘积之和. 即

(1) $|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}, \qquad (1 \leqslant i \leqslant n)$; (按第$i$行展开)

(2) $|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}, \qquad (1 \leqslant j \leqslant n)$. (按第$j$列展开)

下面的任务是证明这个定理成立. 我们先处理简单的情形.

引理4.18. 如果A的第i行除a_{ij}外其余元素都是零, 则$|A| = a_{ij}A_{ij}$.

☞ **证明.** 可以记

$$A = \begin{pmatrix} B_1 & * & B_2 \\ O & a_{ij} & O \\ B_3 & * & B_4 \end{pmatrix},$$

则

$$M_{ij} = \begin{vmatrix} B_1 & B_2 \\ B_3 & B_4 \end{vmatrix}.$$

由分块阵的初等变换

$$A \overset{r_1 \leftrightarrow r_2}{\rightsquigarrow} \begin{pmatrix} O & a_{ij} & O \\ B_1 & * & B_2 \\ B_3 & * & B_4 \end{pmatrix} \overset{c_1 \leftrightarrow c_2}{\rightsquigarrow} \begin{pmatrix} a_{ij} & O & O \\ * & B_1 & B_2 \\ * & B_3 & B_4 \end{pmatrix},$$

得

$$\begin{aligned} |A| =& (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} a_{ij} & O & O \\ * & B_1 & B_2 \\ * & B_3 & B_4 \end{vmatrix} = (-1)^{i+j} a_{ij} \begin{vmatrix} B_1 & B_2 \\ B_3 & B_4 \end{vmatrix} \\ =& (-1)^{i+j} a_{ij} M_{ij} = a_{ij} A_{ij}. \end{aligned}$$

证毕. □

☞ **定理4.17的证明.** (1) 拆第i行

$$\begin{aligned} |A| &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i1} + 0 + \cdots + 0 & 0 + a_{i2} + 0 + \cdots + 0 & \cdots & 0 + \cdots + 0 + a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i1} & 0 & \cdots & 0 \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & a_{i2} & \cdots & 0 \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \end{aligned}$$

$$+\cdots+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & 0 & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

上面右边的n个行列式都满足引理4.18的条件, 再注意到这n个行列式的第i行任意位置的代数余子式都与A的对应位置的代数余子式一致, 应用引理4.18就得到

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}.$$

(2) 利用(1)和方阵转置后行列式不变的事实. □

如果某行元素乘的不是该行对应的代数余子式, 而是另一行对应的代数余子式, 结果如何呢?

推论4.19. *设$A = (a_{ij})_{n\times n} \in \mathbb{F}^{n\times n}$是$n$阶方阵, 其中$n \geqslant 2$, 则对于$1 \leqslant i, j \leqslant n$有*

(1) $a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \delta_{ij}|A| = \begin{cases} |A|, & \text{如果} i = j, \\ 0, & \text{如果} i \neq j; \end{cases}$

(2) $a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = \delta_{ij}|A| = \begin{cases} |A|, & \text{如果} i = j, \\ 0, & \text{如果} i \neq j. \end{cases}$

☞ **证明.** 只需证明(1)及$i \neq j$的情形. 不妨设$i < j$, 我们有

$$0 = \begin{array}{c} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \end{array}\begin{array}{l} \\ \\ i\text{行} \\ \\ j\text{行} \\ \\ \\ \end{array} \xlongequal{\text{按第}j\text{行展开}} a_{i1}A_{j1} + \cdots + a_{in}A_{jn}.$$

这里读者要想清楚为什么上面连等式中构造的行列式的第j行位置的代数余子式与A的相应的代数余子式一致. □

按行(列)展开可以降阶, 为了不增加计算量, 应该选取零多的行和列展开; 当零不够多时, 用初等变换造出零(打洞降阶). 下面用这种想法重新计算例4.10.

例4.21 (例4.10)**.** 计算行列式$D = \begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}$.

☞ **解.** 我们有

$$D \xlongequal[c_4+c_3]{c_1-2c_3} \begin{vmatrix} 5 & 1 & -1 & 1 \\ -11 & 1 & 3 & -1 \\ 0 & 0 & 1 & 0 \\ -5 & -5 & 3 & 0 \end{vmatrix} = (-1)^{3+3} \begin{vmatrix} 5 & 1 & 1 \\ -11 & 1 & -1 \\ -5 & -5 & 0 \end{vmatrix}$$

$$\xlongequal{r_2+r_1} \begin{vmatrix} 5 & 1 & 1 \\ -6 & 2 & 0 \\ -5 & -5 & 0 \end{vmatrix} = (-1)^{1+3} \begin{vmatrix} -6 & 2 \\ -5 & -5 \end{vmatrix} = 40.$$

解毕. □

我们将在下一节给出更多的行列式计算的例子.

4.7.2 伴随矩阵和求逆公式

设$A = (a_{ij}) \in \mathbb{F}^{n\times n}$, 其中$n \geqslant 2$. 设$A_{ij}$是$A$的$(i,j)$位置的代数余子式, 则推论4.19断言下面$2n^2$个等式成立

$$\begin{aligned} a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} &= \delta_{ij}|A|, \quad (1 \leqslant i,j \leqslant n), \\ a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} &= \delta_{ij}|A|, \quad (1 \leqslant i,j \leqslant n). \end{aligned}$$

用矩阵来写恰是

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} |A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A| \end{pmatrix},$$

$$\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} |A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A| \end{pmatrix}.$$

我们引入如下的定义

定义4.8. 设$A \in \mathbb{F}^{n\times n}$, 其中$n \geqslant 2$, 设$A$的$(i,j)$位置的代数余子式是$A_{ij}$. 矩阵$A$的伴随矩阵定义为[7]

$$\operatorname{adj}(A) := \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} = (A_{ji})_{n\times n} \in \mathbb{F}^{n\times n}.$$

当$n = 1$时, 定义$\operatorname{adj}(A) = E_1$.

应用伴随矩阵的记号, 推论4.19可以改写成下面的定理.

定理4.20. 设$A \in \mathbb{F}^{n\times n}$, 则
$$A\,\mathrm{adj}(A) = \mathrm{adj}(A)A = |A|E_n.$$

☞ **证明.** 只需看$n=1$的情形. 此时$A=(a)$, $\mathrm{adj}(A)=E_1$, 进而$|A|=a$, 于是可知成立. □

于是我们得到下面的求逆公式.

定理4.21 (求逆公式)**.** 设$A \in \mathbb{F}^{n\times n}$是可逆矩阵, 则
$$A^{-1} = \frac{1}{|A|}\,\mathrm{adj}(A).$$

☞ **证明.** 由于A可逆, 所以$|A| \neq 0$. 于是定理4.20中伴随矩阵满足的等式可以改写为
$$A\left(\frac{1}{|A|}\,\mathrm{adj}(A)\right) = E_n.$$
因此$A^{-1} = \frac{1}{|A|}\,\mathrm{adj}(A)$. □

例4.22 (例3.29)**.** 设$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. 问: a,b,c,d满足何种条件时A可逆? 并在A可逆时求A的逆.

☞ **解.** 由于$|A| = ad-bc$, 所以
$$A\text{可逆} \iff ad-bc \neq 0.$$
而伴随矩阵
$$\mathrm{adj}(A) = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$
所以当$ad-bc\neq 0$时, 有求逆公式
$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
解毕. □

虽然用求逆公式很容易就求出了二阶可逆方阵的逆, 但是通常不用该公式具体计算高阶矩阵的逆, 因为这需要计算很多行列式. 我们还常常反过来用这个公式: 当A可逆时, 有$\mathrm{adj}(A) = |A|A^{-1}$.

习题4.7

A1. 计算下列行列式

(1) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & -5 \\ -1 & 3 & 1 & 3 \\ 2 & -4 & -1 & -3 \end{vmatrix}$; (2) $\begin{vmatrix} a & 1 & & \\ -1 & b & 1 & \\ & -1 & c & 1 \\ & & -1 & d \end{vmatrix}$; (3) $\begin{vmatrix} 1 & a & a^2 & a^3 \\ a & 1 & a & a^2 \\ a^2 & a & 1 & a \\ a^3 & a^2 & a & 1 \end{vmatrix}$.

A2. 计算下列n阶行列式

[7] 矩阵A的伴随矩阵也常常记为A^*. 由于后面我们将用A^*表示复矩阵A的转置共轭矩阵, 而伴随矩阵的英文是adjugate matrix, 所以本讲义采用了$\mathrm{adj}(A)$ 这个记号.

(1) $\begin{vmatrix} x & y & 0 & \cdots & 0 & 0 \\ 0 & x & y & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & y \\ y & 0 & 0 & \cdots & 0 & x \end{vmatrix}$, $(n \geqslant 2)$;

(2) $\begin{vmatrix} x & & & & & a_0 \\ -1 & x & & & & a_1 \\ & -1 & x & & & a_2 \\ & & \ddots & \ddots & & \vdots \\ & & & -1 & x & a_{n-2} \\ & & & & -1 & x+a_{n-1} \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ a & 1 & 2 & \cdots & n-2 & n-1 \\ a & a & 1 & \cdots & n-3 & n-2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a & a & a & \cdots & 1 & 2 \\ a & a & a & \cdots & a & 1 \end{vmatrix}$.

A3. 求下列方阵的逆阵

(1) $\begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$;　　(2) $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 2 \end{pmatrix}$.

A4. (1) 设$A = (a_{ij})$是n阶方阵, 其中$n \geqslant 2$, A_{ij}是行列式$|A|$中元素a_{ij}的代数余子式, $b_1, \ldots, b_n \in \mathbb{F}$. 证明: 对任意$1 \leqslant i \leqslant n$, 有

$$b_1 A_{i1} + b_2 A_{i2} + \cdots + b_n A_{in} = |B|, \quad b_1 A_{1i} + b_2 A_{2i} + \cdots + b_n A_{ni} = |C|,$$

其中B是将A的第i行用$(b_1, b_2, \ldots, b_n)$替换后得到的n阶方阵, 而C是将A的第i列用$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$替换后得到的$n$阶方阵;

(2) 设$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 10 & 2 & 1 \\ 3 & 1 & 4 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix}$, 求$|A|$中第二列元素的代数余子式之和.

A5. 设n阶方阵$A = (a_{ij})$的元素a_{ij}都是变量x的可微函数, $1 \leqslant i, j \leqslant n$. 证明:

$$\frac{\mathrm{d}(\det A)}{\mathrm{d}x} = \sum_{1 \leqslant i,j \leqslant n} \frac{\mathrm{d}a_{ij}}{\mathrm{d}x} A_{ij},$$

其中A_{ij}是行列式$|A|$中元素a_{ij}的代数余子式.

A6. 设A为3阶方阵, 且$|A|=\frac{1}{2}$, 求$|(2A)^{-1}-5\operatorname{adj}(A)|$.

A7. 设A是n阶方阵, 其中$n\geqslant 2$, 证明: $|\operatorname{adj}(A)|=|A|^{n-1}$.

A8. 设A和B是n阶方阵, 其中$n\geqslant 2$, 证明:

(1) $\operatorname{adj}(aA)=a^{n-1}\operatorname{adj}(A)$, 其中$a$是数;

(2) $\operatorname{adj}(AB)=\operatorname{adj}(B)\operatorname{adj}(A)$;

(3) $\operatorname{adj}(A^T)=\operatorname{adj}(A)^T$;

(4) 如果A可逆, 则$\operatorname{adj}(A)$也可逆, 且$\operatorname{adj}(A)^{-1}=\operatorname{adj}(A^{-1})$.

A9. 设A是n阶方阵, $n\geqslant 2$. 证明:

(1) 如果A是对称阵, 则$\operatorname{adj}(A)$也是对称阵;

(2) 如果A是反对称阵, 则当n为偶数时, $\operatorname{adj}(A)$是反对称阵, 而当n是奇数时, $\operatorname{adj}(A)$是对称阵.

B1. 给定n阶方阵$A=(a_{ij})$, 令

$$A(t)=\begin{pmatrix} a_{11}+t & a_{12}+t & \cdots & a_{1n}+t \\ a_{21}+t & a_{22}+t & \cdots & a_{2n}+t \\ \vdots & \vdots & & \vdots \\ a_{n1}+t & a_{n2}+t & \cdots & a_{nn}+t \end{pmatrix},$$

其中t是参数. 设A_{ij}是行列式$|A|$中元素a_{ij}的代数余子式, 证明:

(1)

$$\det(A(t))=\det(A)+t\sum_{i,j=1}^{n}A_{ij};$$

(2)

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{21}-a_{11} & a_{22}-a_{12} & \cdots & a_{2n}-a_{1n} \\ a_{31}-a_{11} & a_{32}-a_{12} & \cdots & a_{3n}-a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1}-a_{11} & a_{n2}-a_{12} & \cdots & a_{nn}-a_{1n} \end{vmatrix}=\sum_{1\leqslant i,j\leqslant n}A_{ij}.$$

B2. (1) 设n阶方阵$A=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ & 1 & 1 & \cdots & 1 \\ & & 1 & \cdots & 1 \\ & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix}$, 求$|A|$的所有代数余子式之和;

(2) 一般地, 如果n阶行列式D的某一行(或列)的所有元素都是1, 则D的所有元素的代数余子式之和和D有什么关系? 证明你的结论.

B3. 用$x_1, x_2, \ldots, x_{n-1}, 1$替换$n$阶行列式

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & 1 \\ a_{21} & \cdots & a_{2,n-1} & 1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 1 \end{vmatrix}$$

的第i行, 得到的行列式记为D_i, $i = 1, 2, \ldots, n$. 证明:

$$D = D_1 + D_2 + \cdots + D_n.$$

B4. 给定n阶方阵$A = (a_{ij})$. 证明:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & z \end{vmatrix} = z \det A - \sum_{1 \leqslant i,j \leqslant n} A_{ij} x_i y_j,$$

其中A_{ij}是行列式$\det A$的元素a_{ij}的代数余子式, $1 \leqslant i, j \leqslant n$.

B5. 设$A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^{m\times m}$, $C = \begin{pmatrix} A & O \\ O & B \end{pmatrix}$, 证明:

$$\operatorname{adj}(C) = \begin{pmatrix} |B| \operatorname{adj}(A) & O \\ O & |A| \operatorname{adj}(B) \end{pmatrix},$$

其中一阶方阵的伴随矩阵约定为E_1.

B6. 设$A \in \mathbb{F}^{n\times n}$, $\alpha, \beta \in \mathbb{F}^n$, 证明: $|A + \alpha\beta^T| = |A| + \beta^T \operatorname{adj}(A)\alpha$, 其中一阶方阵的伴随矩阵约定为$E_1$.

B7. 设$A \in \mathbb{F}^{n\times n}$为$n$阶反对称阵, 其中$n$是偶数, 任取$\alpha \in \mathbb{F}^n$和$a \in \mathbb{F}$, 证明:

$$|A| = |A + a\alpha\alpha^T|.$$

§ 4.8 行列式计算之例二

本节再举一些行列式计算的例子, 这是另一些风景.

例4.23. 计算n阶行列式$D_n = \begin{vmatrix} x & a & a & \cdots & a & a \\ b & a & c & \cdots & c & c \\ b & c & a & \cdots & c & c \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & c & c & \cdots & a & c \\ b & c & c & \cdots & c & a \end{vmatrix}$.

☞ **解.** 用主行消法和列累加法

$$
D_n \xlongequal[i=2,\dots,n-1]{r_i-r_n} \begin{vmatrix} x & a & a & \cdots & a & a \\ & a-c & & & & c-a \\ & & a-c & & & c-a \\ & & & \ddots & & \vdots \\ & & & & a-c & c-a \\ b & c & c & \cdots & c & a \end{vmatrix}
$$

$$
\xlongequal[i=2,\dots,n-1]{c_n+c_i} \begin{vmatrix} x & a & a & \cdots & a & (n-1)a \\ & a-c & & & & \\ & & a-c & & & \\ & & & \ddots & & \\ & & & & a-c & \\ b & c & c & \cdots & c & a+(n-2)c \end{vmatrix}.
$$

再按照第一列展开, 得

$$
\begin{aligned}
D_n =& x \begin{vmatrix} a-c & & & & \\ & a-c & & & \\ & & \ddots & & \\ & & & a-c & \\ c & c & \cdots & c & a+(n-2)c \end{vmatrix} \\
& + (-1)^{n+1} b \begin{vmatrix} a & a & \cdots & a & (n-1)a \\ a-c & & & & \\ & a-c & & & \\ & & \ddots & & \\ & & & a-c & \end{vmatrix} \\
=& x(a-c)^{n-2}[a+(n-2)c] \\
& + (-1)^{n+1}(-1)^{1+(n-1)} b(n-1)a \begin{vmatrix} a-c & & & \\ & a-c & & \\ & & \ddots & \\ & & & a-c \end{vmatrix} \\
=& x(a-c)^{n-2}[a+(n-2)c] - (n-1)ab(a-c)^{n-2} \\
=& (a-c)^{n-2}[ax+(n-2)cx-(n-1)ab].
\end{aligned}
$$

解毕. □

我们举一些三对角行列式的例子.

例4.24. 计算n阶行列式$D_n = \begin{vmatrix} \cos\theta & 1 & & & & \\ 1 & 2\cos\theta & 1 & & & \\ & 1 & 2\cos\theta & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 2\cos\theta & 1 \\ & & & & 1 & 2\cos\theta \end{vmatrix}$.

☞ **解.** 三对角行列式的计算方法通常是按照第1行(列)或者第n行(列)展开, 得到递推公式, 然后归纳得到. 这里, 当$n \geqslant 3$时, 按照第n行展开得(为什么不按照第1行展开?)

$$\begin{aligned} D_n =& 2\cos\theta D_{n-1} - \begin{vmatrix} \cos\theta & 1 & & & & \\ 1 & 2\cos\theta & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & 2\cos\theta & 1 & \\ & & & 1 & 2\cos\theta & \\ & & & & 1 & 1 \end{vmatrix} \\ =& 2\cos\theta D_{n-1} - D_{n-2}. \end{aligned}$$

现在$D_1 = \cos\theta$,

$$D_2 = \begin{vmatrix} \cos\theta & 1 \\ 1 & 2\cos\theta \end{vmatrix} = 2\cos^2\theta - 1 = \cos 2\theta,$$

而按照递推公式

$$D_3 = 2\cos\theta D_2 - D_1 = 2\cos\theta\cos 2\theta - \cos\theta = \cos 3\theta + \cos\theta - \cos\theta = \cos 3\theta,$$

这里, 我们用了三角公式

$$2\cos\alpha\cos\beta = \cos(\alpha+\beta) + \cos(\alpha-\beta).$$

于是我们猜想: $D_n = \cos n\theta$.

下面用数学归纳法证明这个猜想. 当$n = 1, 2$时猜想成立, 下设$n \geqslant 3$, 且对任意$k < n$有$D_k = \cos k\theta$. 利用递推公式和归纳假设, 就得

$$\begin{aligned} D_n =& 2\cos\theta D_{n-1} - D_{n-2} = 2\cos\theta\cos(n-1)\theta - \cos(n-2)\theta \\ =& \cos n\theta + \cos(n-2)\theta - \cos(n-2)\theta = \cos n\theta. \end{aligned}$$

所以有$D_n = \cos n\theta$. □

考虑例4.24的如下变形.

例4.25. 计算n阶行列式$D_n = \begin{vmatrix} 2\cos\theta & 1 & & & \\ 1 & 2\cos\theta & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2\cos\theta & 1 \\ & & & 1 & 2\cos\theta \end{vmatrix}$.

☞ **解.** 类似于例4.24, 按照第n行(或者第1行)展开得递推公式

$$D_n = 2\cos\theta D_{n-1} - D_{n-2}, \quad (n \geqslant 3).$$

现在$D_1 = 2\cos\theta = \frac{\sin 2\theta}{\sin\theta}$,

$$\begin{aligned} D_2 =& 4\cos^2\theta - 1 = 2\cos 2\theta + 1 = \frac{2\sin\theta\cos 2\theta + \sin\theta}{\sin\theta} \\ =& \frac{\sin 3\theta - \sin\theta + \sin\theta}{\sin\theta} = \frac{\sin 3\theta}{\sin\theta}, \end{aligned}$$

其中用了三角公式

$$2\sin\alpha\cos\beta = \sin(\alpha+\beta) + \sin(\alpha-\beta).$$

又由递推公式

$$D_3 = 2\cos\theta D_2 - D_1 = \frac{2\sin 3\theta\cos\theta - \sin 2\theta}{\sin\theta} = \frac{\sin 4\theta}{\sin\theta}.$$

所以我们猜想: $D_n = \frac{\sin(n+1)\theta}{\sin\theta}$.

假设$n \geqslant 3$, 且$n-1$和$n-2$时猜想成立, 则有

$$\begin{aligned} D_n =& 2\cos\theta D_{n-1} - D_{n-2} = \frac{2\sin n\theta\cos\theta - \sin(n-1)\theta}{\sin\theta} \\ =& \frac{\sin(n+1)\theta + \sin(n-1)\theta - \sin(n-1)\theta}{\sin\theta} = \frac{\sin(n+1)\theta}{\sin\theta}, \end{aligned}$$

即猜想成立. □

上面两个例子最后本质上归结为如下的数学问题:

设数列$\{D_n\}$满足递推公式

$$D_n = aD_{n-1} + bD_{n-2}, \qquad (n \geqslant 3),$$

并已知D_1和D_2, 求D_n的通项公式.

可以如下求解这个问题. 考虑递推公式对应的特征方程$x^2 - ax - b = 0$, 设它的两个根为$\alpha, \beta \in \mathbb{C}$, 则$\alpha + \beta = a$, $\alpha\beta = -b$. 所以递推公式为

$$D_n = (\alpha+\beta)D_{n-1} - \alpha\beta D_{n-2} \Longrightarrow D_n - \alpha D_{n-1} = \beta(D_{n-1} - \alpha D_{n-2}).$$

由等比数列的通项公式得

$$D_n - \alpha D_{n-1} = \beta^{n-2}(D_2 - \alpha D_1).$$

类似得

$$D_n - \beta D_{n-1} = \alpha^{n-2}(D_2 - \beta D_1).$$

消去D_{n-1}, 得

$$(\alpha-\beta)D_n = \alpha^{n-1}(D_2 - \beta D_1) - \beta^{n-1}(D_2 - \alpha D_1).$$

所以当$\alpha \neq \beta$时有

$$D_n = \frac{\alpha^{n-1}(D_2 - \beta D_1) - \beta^{n-1}(D_2 - \alpha D_1)}{\alpha - \beta}.$$

当$\alpha = \beta$时, 由连续性

$$D_n = \lim_{\beta\to\alpha} \frac{\alpha^{n-1}(D_2 - \beta D_1) - \beta^{n-1}(D_2 - \alpha D_1)}{\alpha - \beta}$$

$$
\begin{aligned}
&=\lim_{\beta\to\alpha}\frac{\alpha^{n-1}(-D_1)-(n-1)\beta^{n-2}(D_2-\alpha D_1)}{-1}\\
&=(n-1)\alpha^{n-2}D_2-(n-2)\alpha^{n-1}D_1
\end{aligned}
$$

所以得到通项公式

$$
D_n=\begin{cases}\frac{\alpha^{n-1}(D_2-\beta D_1)-\beta^{n-1}(D_2-\alpha D_1)}{\alpha-\beta}, & \alpha\neq\beta,\\ (n-1)\alpha^{n-2}D_2-(n-2)\alpha^{n-1}D_1, & \alpha=\beta.\end{cases}
$$

回到例4.24和例4.25, 通项公式是$D_n=2\cos\theta D_{n-1}-D_{n-2}$, 特征方程是$x^2-2\cos\theta x+1=0$, 根为

$$
x=\frac{2\cos\theta\pm\sqrt{4\cos^2\theta-4}}{2}=\cos\theta\pm i\sin\theta=e^{\pm i\theta}.
$$

代入通项公式就得到

$$
D_n=\frac{e^{i(n-1)\theta}(D_2-e^{-i\theta}D_1)-e^{-i(n-1)\theta}(D_2-e^{i\theta}D_1)}{2i\sin\theta}.
$$

由于$D_1,D_2\in\mathbb{R}$, 所以上面的分子为

$$
\begin{aligned}
&(\cos(n-1)\theta+i\sin(n-1)\theta)(D_2-D_1\cos\theta+iD_1\sin\theta)\\
&-(\cos(n-1)\theta-i\sin(n-1)\theta)(D_2-D_1\cos\theta-iD_1\sin\theta)\\
=&2i[D_1\sin\theta\cos(n-1)\theta+(D_2-D_1\cos\theta)\sin(n-1)\theta].
\end{aligned}
$$

进而有

$$
D_n=\frac{D_1\sin\theta\cos(n-1)\theta+(D_2-D_1\cos\theta)\sin(n-1)\theta}{\sin\theta}.
$$

对于例4.24, $D_1=\cos\theta$, $D_2=\cos 2\theta$, 所以有

$$
\begin{aligned}
D_n=&\frac{\cos\theta\sin\theta\cos(n-1)\theta+(\cos 2\theta-\cos^2\theta)\sin(n-1)\theta}{\sin\theta}\\
=&\cos\theta\cos(n-1)-\sin\theta\sin(n-1)\theta=\cos n\theta.
\end{aligned}
$$

对于例4.25, $D_1=2\cos\theta$, $D_2=4\cos^2\theta-1$, 所以有

$$
\begin{aligned}
D_n=&\frac{2\cos\theta\sin\theta\cos(n-1)\theta+(4\cos^2\theta-1-2\cos^2\theta)\sin(n-1)\theta}{\sin\theta}\\
=&\frac{\sin 2\theta\cos(n-1)\theta+\cos 2\theta\sin(n-1)\theta}{\sin\theta}=\frac{\sin(n+1)\theta}{\sin\theta}.
\end{aligned}
$$

考虑例4.25的一般化.

例4.26. 计算n阶三对角行列式$D_n=\begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix}$.

☞ **解.** 按照第一行展开, 得

$$D_n = aD_{n-1} - b\begin{vmatrix} c & b & & & & \\ & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{vmatrix} = aD_{n-1} - bcD_{n-2}.$$

特征方程为$x^2 - ax + bc = 0$. 设α和β是两个根, 则

$$\alpha + \beta = a, \quad \alpha\beta = bc.$$

当$\alpha \neq \beta$, 即$a^2 \neq 4bc$时有

$$D_n = \frac{\alpha^{n-1}(D_2 - \beta D_1) - \beta^{n-1}(D_2 - \alpha D_1)}{\alpha - \beta}.$$

而$D_1 = a = \alpha + \beta$, $D_2 = a^2 - bc = (\alpha+\beta)^2 - \alpha\beta$, 所以

$$\begin{aligned} D_2 - \beta D_1 &= (\alpha+\beta)^2 - \alpha\beta - \beta(\alpha+\beta) = \alpha^2, \\ D_2 - \alpha D_1 &= (\alpha+\beta)^2 - \alpha\beta - \alpha(\alpha+\beta) = \beta^2. \end{aligned}$$

这得到

$$D_n = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}.$$

当$a^2 = 4bc$时, $\alpha = \beta = \frac{a}{2}$. 于是

$$\begin{aligned} D_n =&(n-1)\alpha^{n-2}D_2 - (n-2)\alpha^{n-1}D_1 \\ =&(n-1)\left(\frac{a}{2}\right)^{n-2}(a^2 - bc) - (n-2)\left(\frac{a}{2}\right)^{n-1} a \\ =&(n-1)\left(\frac{a}{2}\right)^{n-2}\frac{3}{4}a^2 - (n-2)\left(\frac{a}{2}\right)^{n-1} a \\ =&[3(n-1) - 2(n-2)]\left(\frac{a}{2}\right)^{n} = (n+1)\left(\frac{a}{2}\right)^{n}. \end{aligned}$$

综上有

$$D_n = \begin{cases} \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}, & a^2 \neq 4bc, \\ (n+1)\left(\frac{a}{2}\right)^n, & a^2 = 4bc. \end{cases}$$

其中α, β为$x^2 - ax + bc = 0$的两个根. □

下面我们介绍一个重要的行列式: **Vandermonde行列式**.

例4.27 (Vandermonde行列式)**.** 设$n \geqslant 2$, 证明:

$$V_n(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i).$$

☞ **证明.** 自下而上, 后行减去前行的x_1倍, 得

$$V_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & \cdots & x_n - x_1 \\ 0 & x_2^2 - x_1x_2 & \cdots & x_n^2 - x_1x_2 \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-1} - x_1x_2^{n-2} & \cdots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix},$$

再按第一列展开, 并提取每列的公因子, 得

$$\begin{aligned} V_n = & \begin{vmatrix} x_2 - x_1 & \cdots & x_n - x_1 \\ x_2^2 - x_1x_2 & \cdots & x_n^2 - x_1x_2 \\ \vdots & & \vdots \\ x_2^{n-1} - x_1x_2^{n-2} & \cdots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix} \\ = & (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1)\begin{vmatrix} 1 & \cdots & 1 \\ x_2 & \cdots & x_n \\ \vdots & & \vdots \\ x_2^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}. \end{aligned}$$

于是得到递推公式

$$V_n = \prod_{j=2}^{n}(x_j - x_1)V_{n-1}(x_2, \ldots, x_n), \quad (n \geqslant 3).$$

下面用数学归纳法证明Vandermonde行列式的公式. 当$n = 2$时,

$$V_2 = \begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix} = x_2 - x_1,$$

于是此时公式成立. 下设$n \geqslant 3$且V_{n-1}的公式成立, 则

$$\begin{aligned} V_n &= \prod_{j=2}^{n}(x_j - x_1)V_{n-1}(x_2, \ldots, x_n) = \prod_{j=2}^{n}(x_j - x_1)\prod_{2\leqslant i<j\leqslant n}(x_j - x_i) \\ &= \prod_{1\leqslant i<j\leqslant n}(x_j - x_i). \end{aligned}$$

证毕.

如果将$x_1, x_2, \ldots, x_n$看成不定元, 也可以利用$\mathbb{F}[x_1, x_2, \ldots, x_n]$中因式分解唯一性定理证明Vandermonde行列式公式成立. □

从Vandermonde行列式的公式可以看出

$$V_n(x_1, x_2, \ldots, x_n) \neq 0 \Longleftrightarrow x_1, x_2, \ldots, x_n\text{两两不等}.$$

下面考虑Vandermonde行列式的一个变形.

例4.28. 计算n阶行列式$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ x_1^n & x_2^n & \cdots & x_n^n \end{vmatrix}.$

☞ **解.** 这里介绍计算行列式的一种特殊方法: 加边(将行列式的阶数升高). 定义关于x的多项式

$$f(x)=\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_n & x \\ x_1^2 & x_2^2 & \cdots & x_n^2 & x^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} & x^{n-2} \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} & x^{n-1} \\ x_1^n & x_2^n & \cdots & x_n^n & x^n \end{vmatrix}_{(n+1)\times(n+1)}.$$

首先, 按照第$n+1$列展开, 可得

$$f(x)=a_0+a_1x+\cdots+a_{n-2}x^{n-2}+a_{n-1}x^{n-1}+a_nx^n,$$

其中各项的系数均可表示出, 特别有

$$a_{n-1}=(-1)^{n+n+1}D_n=-D_n.$$

其次, 由Vandermonde行列式的公式, 可得

$$f(x)=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)[(x-x_1)(x-x_2)\cdots(x-x_n)].$$

而

$$(x-x_1)(x-x_2)\cdots(x-x_n)=x^n-(x_1+x_2+\cdots+x_n)x^{n-1}+\text{低次项},$$

所以$f(x)$的x^{n-1}的系数是

$$a_{n-1}=-(x_1+x_2+\cdots+x_n)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

最后就得

$$D_n=(x_1+x_2+\cdots+x_n)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

解毕. □

习题4.8

B1. 计算下列n阶行列式

(1) $\begin{vmatrix} 7 & 5 & & & \\ 2 & 7 & 5 & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & 7 & 5 \\ & & & 2 & 7 \end{vmatrix}$;

(2) $\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-3} & x_2^{n-3} & \cdots & x_n^{n-3} \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \\ x_1^n & x_2^n & \cdots & x_n^n \end{vmatrix}$;

(3) $\begin{vmatrix} a_1 & -a_2 & & & & \\ & a_2 & -a_3 & & & \\ & & a_3 & -a_4 & & \\ & & & \ddots & \ddots & \\ & & & & a_{n-1} & -a_n \\ 1 & 1 & 1 & \cdots & 1 & 1+a_n \end{vmatrix}$;

(4) $\begin{vmatrix} 1 & 2 & \cdots & n \\ 2 & 3 & \cdots & 1 \\ 3 & 4 & \cdots & 2 \\ \vdots & \vdots & & \vdots \\ n & 1 & \cdots & n-1 \end{vmatrix}$;

(5) $\begin{vmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix}$; (6) $\begin{vmatrix} \dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2} & \cdots & \dfrac{1}{a_1+b_n} \\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2} & \cdots & \dfrac{1}{a_2+b_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{1}{a_n+b_1} & \dfrac{1}{a_n+b_2} & \cdots & \dfrac{1}{a_n+b_n} \end{vmatrix}$;

(7) $\begin{vmatrix} 1 & \cos\theta_1 & \cos 2\theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos\theta_2 & \cos 2\theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos\theta_n & \cos 2\theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}$;

(8) $\begin{vmatrix} \sin n\theta_1 & \sin(n-1)\theta_1 & \cdots & \sin\theta_1 \\ \sin n\theta_2 & \sin(n-1)\theta_2 & \cdots & \sin\theta_2 \\ \vdots & \vdots & & \vdots \\ \sin n\theta_n & \sin(n-1)\theta_n & \cdots & \sin\theta_n \end{vmatrix}$.

B2. 设$a_1, a_2, \ldots, a_n$是正整数. 证明: n阶行列式

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

能被$1^{n-1}2^{n-2}\cdots(n-2)^2(n-1)$整除.

B3. (Levy-Desplanques) 设n阶方阵$A=(a_{ij})$的元素都是复数, 满足

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad i=1,2,\ldots,n,$$

则方阵A称为主角占优矩阵. 证明: 主角占优矩阵的行列式不为零.

B4. (Minkowski) 设n阶方阵$A=(a_{ij})$的元素都是实数, 满足

$$a_{ii} > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad i=1,2,\ldots,n,$$

证明: $\det(A) > 0$.

§ 4.9 Laplace展开

我们知道行列式可以按照一行(列)展开, 那么是否可以按照几行(列)展开呢? 下面我们来分析这个问题. 在此之前, 我们重温一下行列式的定义.

行列式的定义是从排列开始的. 设$I=\{k_1<k_2<\cdots<k_n\}$为n元有序集, 则$k_1, k_2, \ldots, k_n$按照任意次序排成的有序组$p_1, p_2, \ldots, p_n$称为$k_1, k_2, \ldots, k_n$(或者I)的一个排列, 记为$\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$.

称$\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$为自然排列. 我们前面讨论的$n$元排列就是取$I = \{1 < 2 < \cdots < n\}$的特殊情形; 类似于$n$元排列, 可以定义$I$的排列的逆序数和符号. 如果$i < j$而$p_i > p_j$, 则称$(p_i, p_j)$是排列$\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$的一个逆序对; 逆序对的总数称为逆序数; 而排列的符号则定义为

$$\delta \begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} := (-1)^{\text{逆序数}}.$$

类似于n元排列, 可以定义I的所有排列上的对换映射, 并且可以证明(引理4.1和推论4.2):

(i) 对换改变排列的符号;

(ii) 设排列$\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$经过$M$次对换变为自然排列$\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$, 则有

$$\delta \begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} = (-1)^M.$$

设$A \in \mathbb{F}^{n\times n}$是$n$阶方阵, 习惯上我们用$i, j \in \{1 < 2 < \cdots < n\}$来下标$A$中的元素, 但有时也会用其它的$n$元有序集来下标$A$中的元素. 设$I = \{k_1 < k_2 < \cdots < k_n\}$和$J = \{l_1 < l_2 < \cdots < l_n\}$是任意两个$n$元有序集, 而

$$A = (a_{ij})_{i\in I, j\in J} = \begin{pmatrix} a_{k_1l_1} & a_{k_1l_2} & \cdots & a_{k_1l_n} \\ a_{k_2l_1} & a_{k_2l_2} & \cdots & a_{k_2l_n} \\ \vdots & \vdots & & \vdots \\ a_{k_nl_1} & a_{k_nl_2} & \cdots & a_{k_nl_n} \end{pmatrix},$$

则由行列式的定义和定理4.3, 有

$$\begin{aligned} |A| &= \sum_{\begin{pmatrix} l_1 & l_2 & \cdots & l_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}} \delta \begin{pmatrix} l_1 & l_2 & \cdots & l_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} a_{k_1p_1}a_{k_2p_2}\cdots a_{k_np_n} \\ &= \sum_{\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}} \delta \begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} a_{p_1l_1}a_{p_2l_2}\cdots a_{p_nl_n}. \end{aligned}$$

例4.29. 设$A = (a_{ij})_{n\times n} \in \mathbb{F}^{n\times n}$, 且$1 \leqslant j \leqslant n$, 记

$$A_j = \begin{pmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3,j-1} & a_{3,j+1} & \cdots & a_{3n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{F}^{(n-1)\times(n-1)},$$

则这里A_j用了有序集$I=\{2<3<\cdots<n\}$和

$$J=\{1,2,\ldots,n\}-\{j\}=:\{k_2<k_3<\cdots<k_n\}$$

来下标. 所以A_j的行列式, 即a_{1j}的余子式为

$$M_{1j}=|A_j|=\sum_{\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}}\delta\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}a_{2p_2}a_{3p_3}\cdots a_{np_n}.$$

下面回到行列式的按行(列)展开. 设$A=(a_{ij})\in\mathbb{F}^{n\times n}$, 其中$n\geqslant 2$. 首先, 我们重新认识一下按照第一行展开的公式

$$|A|=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{11}.$$

这不就是把$a_{11},a_{12},\ldots,a_{1n}$看成变量, 算出它们的系数吗? 也就是说, 这事实上是个合并同类项的过程. 下面我们用这个观点来重新得到这个公式, 以期得到一些有益的提示.

回到定义

$$|A|=\sum_{p_1,p_2,\ldots,p_n\in S_n}\delta(p_1,p_2,\ldots,p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n},$$

所以有

$$|A|=a_{11}\widetilde{A_{11}}+a_{12}\widetilde{A_{12}}+\cdots+a_{1n}\widetilde{A_{1n}},$$

其中对$1\leqslant j\leqslant n$, a_{1j}的系数为

$$\begin{aligned}\widetilde{A_{1j}}&=\sum_{\substack{p_1,p_2,\ldots,p_n\in S_n\\p_1=j}}\delta(j,p_2,\ldots,p_n)a_{2p_2}\cdots a_{np_n}\\&=\sum_{\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}}\delta(j,p_2,\ldots,p_n)a_{2p_2}\cdots a_{np_n},\end{aligned}$$

其中

$$\{k_2<k_3<\cdots<k_n\}=\{1,2,\ldots,n\}-\{j\}.$$

设$k_2,k_3,\ldots,k_n$的某一排列$\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}$经过$M$次对换后变为自然排列$\begin{pmatrix}k_2&k_3&\cdots&k_n\\k_2&k_3&\cdots&k_n\end{pmatrix}$, 则有

$$\delta\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}=(-1)^M.$$

此时, 经过这M次对换后, n元排列$j,p_2,p_3,\ldots,p_n$变为$j,k_2,k_3,\ldots,k_n$. 再将j依次与后一位交换, $j-1$次相邻对换后该排列变为$1,2,\ldots,n$. 即有如下过程

$$j,p_2,p_3,\ldots,p_n\xrightarrow{M\text{次对换}}j,k_2,k_3,\ldots,k_n\xrightarrow{j-1\text{次相邻对换}}1,2,\ldots,n.$$

所以

$$\delta(j,p_2,p_3,\ldots,p_n)=(-1)^{M+j-1}=(-1)^{j+1}\delta\begin{pmatrix}k_2&k_3&\cdots&k_n\\p_2&p_3&\cdots&p_n\end{pmatrix}.$$

所以我们有

$$\begin{aligned}\widetilde{A_{1j}} =&(-1)^{j+1}\sum_{\begin{pmatrix}k_2 & k_3 & \cdots & k_n\\ p_2 & p_3 & \cdots & p_n\end{pmatrix}}\delta\begin{pmatrix}k_2 & k_3 & \cdots & k_n\\ p_2 & p_3 & \cdots & p_n\end{pmatrix}a_{2p_2}\cdots a_{np_n}\\ =&(-1)^{j+1}M_{1j}=A_{1j}.\end{aligned}$$

这就重新证明了行列式按照第一行展开的公式.

下面将对第一行的讨论推广到前r行$(1\leqslant r<n)$. 合并同类项, 可得

$$\begin{aligned}|A| &= \sum_{\substack{1\leqslant p_1,p_2,\ldots,p_r\leqslant n\\ p_1,p_2,\ldots,p_r:\text{互不相同}}} a_{1p_1}a_{2p_2}\cdots a_{rp_r}c_{p_1,p_2,\ldots,p_r}\\ &= \sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant n}\ \sum_{\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ p_1 & p_2 & \cdots & p_r\end{pmatrix}} a_{1p_1}a_{2p_2}\cdots a_{rp_r}c_{p_1,p_2,\ldots,p_r},\end{aligned}$$

其中系数

$$c_{p_1,p_2,\ldots,p_r} = \sum_{\begin{pmatrix}k_{r+1} & k_{r+2} & \cdots & k_n\\ p_{r+1} & p_{r+2} & \cdots & p_n\end{pmatrix}}\delta(p_1,\ldots,p_n)a_{r+1,p_{r+1}}\cdots a_{np_n},$$

这里

$$\{k_{r+1}<\cdots<k_n\}=\{1,2,\ldots,n\}-\{k_1<k_2<\cdots<k_r\}.$$

我们来计算n元排列$p_1,p_2,\ldots,p_n$的符号. 设排列$\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ p_1 & p_2 & \cdots & p_r\end{pmatrix}$经过$M$次对换变为自然排列$\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ k_1 & k_2 & \cdots & k_r\end{pmatrix}$, 而$\begin{pmatrix}k_{r+1} & k_{r+2} & \cdots & k_n\\ p_{r+1} & p_{r+2} & \cdots & p_n\end{pmatrix}$经过$N$次对换变为$\begin{pmatrix}k_{r+1} & k_{r+2} & \cdots & k_n\\ k_{r+1} & k_{r+2} & \cdots & k_n\end{pmatrix}$, 则有

$$\delta\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ p_1 & p_2 & \cdots & p_r\end{pmatrix}=(-1)^M,\quad \delta\begin{pmatrix}k_{r+1} & k_{r+2} & \cdots & k_n\\ p_{r+1} & p_{r+2} & \cdots & p_n\end{pmatrix}=(-1)^N.$$

于是, 可如下分三步用对换将n元排列$p_1,p_2,\ldots,p_n$变成自然排列:

$$\begin{aligned}&p_1,\ldots,p_r,p_{r+1},\ldots,p_n\xrightarrow{M\text{次对换}}k_1,\ldots,k_r,p_{r+1},\ldots,p_n\xrightarrow{N\text{次对换}}k_1,\ldots,k_r,k_{r+1},\ldots,k_n\\ &\xrightarrow{(\star)}1,2,\ldots,n,\end{aligned}$$

其中$(\star)$是将k_r依次向右经过k_r-r次相邻对换, k_{r-1}依次向右经过$k_{r-1}-(r-1)$次相邻对换, 依次类推, 直到k_1依次向右经过k_1-1次相邻对换. 于是有

$$\begin{aligned}\delta(p_1,\ldots,p_n) =&(-1)^{M+N+(k_1-1)+(k_2-2)+\cdots+(k_r-r)}\\ =&(-1)^{1+2+\cdots+r+k_1+k_2+\cdots+k_r}\delta\begin{pmatrix}k_1 & \cdots & k_r\\ p_1 & \cdots & p_r\end{pmatrix}\delta\begin{pmatrix}k_{r+1} & \cdots & k_n\\ p_{r+1} & \cdots & p_n\end{pmatrix}.\end{aligned}$$

这得到

$$
\begin{aligned}
c_{p_1,p_2,\dots,p_r} =&(-1)^{k_1+\cdots+k_r+1+\cdots+r}\delta\begin{pmatrix}k_1 & \cdots & k_r\\ p_1 & \cdots & p_r\end{pmatrix}\\
&\times\sum_{\begin{pmatrix}k_{r+1} & \cdots & k_n\\ p_{r+1} & \cdots & p_n\end{pmatrix}}\delta\begin{pmatrix}k_{r+1} & \cdots & k_n\\ p_{r+1} & \cdots & p_n\end{pmatrix}a_{r+1,p_{r+1}}\cdots a_{np_n}.
\end{aligned}
$$

注意到上面右边的和其实是一个行列式, 为了表示这个行列式, 我们引入子式的概念.

定义4.9. 设$A=(a_{ij})_{m\times n}\in\mathbb{F}^{m\times n}$, 对于满足$1\leqslant r\leqslant\min\{m,n\}$的整数$r$, 如果$1\leqslant i_1<i_2<\cdots<i_r\leqslant m$, $1\leqslant j_1<j_2<\cdots<j_r\leqslant n$, 将$A$的第$i_1,i_2,\dots,i_r$行和第$j_1,j_2,\dots,j_r$列的交叉位置的元素保持其相对位置组成的$r$阶方阵的行列式称为$A$的一个$r$阶子式, 记为$A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}$, 即

$$
A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}=\begin{vmatrix}a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r}\\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_r}\\ \vdots & \vdots & & \vdots\\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r}\end{vmatrix}_{r\times r}.
$$

上面的r阶子式用定义写就是

$$
A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}=\sum_{\begin{pmatrix}j_1 & j_2 & \cdots & j_r\\ p_1 & p_2 & \cdots & p_r\end{pmatrix}}\delta\begin{pmatrix}j_1 & j_2 & \cdots & j_r\\ p_1 & p_2 & \cdots & p_r\end{pmatrix}a_{i_1p_1}a_{i_2p_2}\cdots a_{i_rp_r}.
$$

利用子式的记号, 我们有

$$
c_{p_1,p_2,\dots,p_r} =(-1)^{1+\cdots+r+k_1+\cdots+k_r}\delta\begin{pmatrix}k_1 & \cdots & k_r\\ p_1 & \cdots & p_r\end{pmatrix}A\begin{pmatrix}r+1 & \cdots & n\\ k_{r+1} & \cdots & k_n\end{pmatrix}.
$$

于是

$$
\begin{aligned}
|A|=&\sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant n}(-1)^{1+\cdots+r+k_1+\cdots+k_r}A\begin{pmatrix}r+1 & \cdots & n\\ k_{r+1} & \cdots & k_n\end{pmatrix}\\
&\times\sum_{\begin{pmatrix}k_1 & \cdots & k_r\\ p_1 & \cdots & p_r\end{pmatrix}}\delta\begin{pmatrix}k_1 & \cdots & k_r\\ p_1 & \cdots & p_r\end{pmatrix}a_{1p_1}a_{2p_2}\cdots a_{rp_r}\\
=&\sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant n}A\begin{pmatrix}1 & \dots & r\\ k_1 & \dots & k_r\end{pmatrix}(-1)^{1+\cdots+r+k_1+\cdots+k_r}A\begin{pmatrix}r+1 & \cdots & n\\ k_{r+1} & \cdots & k_n\end{pmatrix}.
\end{aligned}
$$

这就是$|A|$按照前r行展开的公式.

一般的, 对于任意固定的$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$, 我们来推导类似的结论. 记

$$\{1, 2, \ldots, n\} - \{i_1 < i_2 < \cdots < i_r\} = \{i_{r+1} < i_{r+2} < \cdots < i_n\}.$$

设有行分块$A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$. 将$A$的第$i_1$行依次与前面的$i_1 - 1$行交换, 换到第1行; 再将第$i_2$行依次与前面的$i_2 - 2$行交换, 换到第2行; 依次类推, 最后将第$i_r$行与前面的$i_r - r$行相交换, 换到第$r$行得到矩阵$B$, 于是

$$B = \begin{pmatrix} \alpha_{i_1} \\ \vdots \\ \alpha_{i_r} \\ \alpha_{i_{r+1}} \\ \vdots \\ \alpha_{i_n} \end{pmatrix}.$$

所以

$$|A| = (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)}|B| = (-1)^{1+2+\cdots+r+i_1+i_2+\cdots+i_r}|B|.$$

再对$|B|$按照前r行展开, 就得

$$\begin{aligned} |A| =&(-1)^{1+2+\cdots+r+i_1+i_2+\cdots+i_r} \sum_{1 \leqslant k_1 < \cdots < k_r \leqslant n} B\begin{pmatrix} 1 & 2 & \cdots & r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} \\ & \times (-1)^{1+2+\cdots+r+k_1+k_2+\cdots+k_r} B\begin{pmatrix} r+1 & r+2 & \cdots & n \\ k_{r+1} & k_{r+2} & \cdots & k_n \end{pmatrix} \\ =& \sum_{1 \leqslant k_1 < \cdots < k_r \leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} \\ & \times (-1)^{i_1+i_2+\cdots+i_r+k_1+k_2+\cdots+k_r} A\begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ k_{r+1} & k_{r+2} & \cdots & k_n \end{pmatrix}. \end{aligned}$$

定义4.10. 设$A \in \mathbb{F}^{n\times n}$, $1 \leqslant r < n$, 而$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$, 记$D = A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$为$r$阶子式.

(1) 将A中第$i_1, i_2, \ldots, i_r$行和第$j_1, j_2, \ldots, j_r$列删去后的$n-r$阶方阵的行列式称为D的余子式, 记为$M\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$, 即

$$M\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = A\begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ j_{r+1} & j_{r+2} & \cdots & j_n \end{pmatrix};$$

(2) D的代数余子式定义为

$$\widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = (-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r} M\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}.$$

最后我们来陈述我们发现的展开定理.

定理4.22 (Laplace展开定理)**.** 设$A \in \mathbb{F}^{n\times n}$, $1 \leqslant r < n$, $1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$, 则

(1) (按第$i_1, \ldots, i_r$行展开) $|A|$等于第$i_1, \ldots, i_r$行所有的r阶子式与对应的代数余子式乘积之和, 即

$$|A| = \sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} \widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix};$$

(2) (按第$i_1, \ldots, i_r$列展开) $|A|$等于第$i_1, \ldots, i_r$列所有的r阶子式与对应的代数余子式乘积之和, 即

$$|A| = \sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant n} A\begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} \widehat{A}\begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}.$$

☞ **证明.** (2) 对$|A^T|$用(1). □

如果不在意Laplace展开定理的发现过程, 也可以如下证明该展开定理.

☞ **Laplace展开定理的另证.** 这里只证明定理中的(1). 记$A = (a_{ij})$. 将需证的等式两边按照行列式的定义进行完全展开, 则左边是$n!$项之和, 每项有形式$\pm a_{1p_1}a_{2p_2}\cdots a_{np_n}$, 其中$p_1, p_2, \ldots, p_n \in S_n$; 而右边为

$$\binom{n}{r} r!(n-r)! = n!$$

项之和, 每项有形式

$$\pm a_{i_1p_1}\cdots a_{i_rp_r}\cdot a_{i_{r+1}p_{r+1}}\cdots a_{i_np_n},$$

其中$p_1, \ldots, p_r$为$i_1, \ldots, i_r$的排列, $p_{r+1}, \ldots, p_n$为$i_{r+1}, \ldots, i_n$的排列, 于是这样一项也有形式$\pm a_{1p_1}a_{2p_2}\cdots a_{np_n}$, 其中$p_1, p_2, \ldots, p_n \in S_n$. 再注意到将$a_{ij}$看成不定元时, 左边的各项的列标排列互不相同, 右边的各项的列标排列也互不相同, 所以我们只要证明右边的每项均为左边的一项, 即右边的每项都有形式

$$\delta(p_1, p_2, \ldots, p_n)a_{1p_1}a_{2p_2}\cdots a_{np_n},$$

其中$p_1, p_2, \ldots, p_n \in S_n$.

设$1 \leqslant k_1 < k_2 < \cdots < k_r \leqslant n$, 则

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix}(-1)^{i_1+i_2+\cdots+i_r+k_1+k_2+\cdots+k_r} A\begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ k_{r+1} & k_{r+2} & \cdots & k_n \end{pmatrix}$$

展开后的标准的一项有形式

$$\delta\begin{pmatrix} k_1 & \cdots & k_r \\ p_1 & \cdots & p_r \end{pmatrix} a_{i_1p_1}\cdots a_{i_rp_r}(-1)^{i_1+\cdots+i_r+k_1+\cdots+k_r}\delta\begin{pmatrix} k_{r+1} & \cdots & k_n \\ p_{r+1} & \cdots & p_n \end{pmatrix} a_{i_{r+1}p_{r+1}}\cdots a_{i_np_n}. \tag{4.1}$$

可用$(i_1-1)+(i_2-2)+\cdots+(i_r-r)$次对换将排列$i_1, \ldots, i_r, i_{r+1}, \ldots, i_n$变成标准排列$1, 2, \ldots, n$; 设在同样的对换下, 排列$p_1, \ldots, p_r, p_{r+1}, \ldots, p_n$变成了排列$j_1, j_2, \ldots, j_n$, 则

$$a_{i_1p_1}\cdots a_{i_rp_r}\cdot a_{i_{r+1}p_{r+1}}\cdots a_{i_np_n} = a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

我们按照下面的对换过程来计算排列$j_1, j_2, \ldots, j_n$的符号:

$$j_1, j_2, \ldots, j_n \xrightarrow{(\bigstar)} p_1, \ldots, p_r, p_{r+1}, \ldots, p_n \xrightarrow{M\text{次对换}} k_1, \ldots, k_r, p_{r+1}, \ldots, p_n$$
$$\xrightarrow{N\text{次对换}} k_1, \ldots, k_r, k_{r+1}, \ldots, k_n \xrightarrow{(\star)} 1, 2, \ldots, n,$$

其中$\begin{pmatrix} k_1 & \cdots & k_r \\ p_1 & \cdots & p_r \end{pmatrix}$经过$M$次对换变为自然排列$\begin{pmatrix} k_1 & \cdots & k_r \\ k_1 & \cdots & k_r \end{pmatrix}$, 排列$\begin{pmatrix} k_{r+1} & \cdots & k_n \\ p_{r+1} & \cdots & p_n \end{pmatrix}$经过$N$次对换变为自然排列$\begin{pmatrix} k_{r+1} & \cdots & k_n \\ k_{r+1} & \cdots & k_n \end{pmatrix}$, $(\bigstar)$所用的对换次数为$(i_1 - 1) + (i_2 - 2) + \cdots + (i_r - r)$, 而$(\star)$所用的对换次数为$(k_1 - 1) + (k_2 - 2) + \cdots + (k_r - r)$. 这得到

$$\begin{aligned}
&\delta(j_1, j_2, \ldots, j_n) \\
=&(-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)}\delta\begin{pmatrix} k_1 & \cdots & k_r \\ p_1 & \cdots & p_r \end{pmatrix}\delta\begin{pmatrix} k_{r+1} & \cdots & k_n \\ p_{r+1} & \cdots & p_n \end{pmatrix} \\
&\quad \cdot (-1)^{(k_1-1)+(k_2-2)+\cdots+(k_r-r)} \\
=&(-1)^{i_1+\cdots+i_r+k_1+\cdots+k_r}\delta\begin{pmatrix} k_1 & \cdots & k_r \\ p_1 & \cdots & p_r \end{pmatrix}\delta\begin{pmatrix} k_{r+1} & \cdots & k_n \\ p_{r+1} & \cdots & p_n \end{pmatrix}.
\end{aligned}$$

最后可得, 需证等式右边的一项(4.1)有形式

$$\delta(j_1, j_2, \ldots, j_n)a_{1j_1}a_{2j_2}\cdots a_{nj_n},$$

证毕. □

我们用Laplace展开定理再算一次例4.10, 当然我们一般不会这么做.

例4.30 (例4.10). *按照第一和第二行Laplace展开, 计算行列式*$D = \begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}$.

☞ **解.** 第一和第二行共有$\binom{4}{2} = 6$个二阶子式:

$$A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{vmatrix} 3 & 1 \\ -5 & 1 \end{vmatrix} = 8, \quad A\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{vmatrix} 3 & -1 \\ -5 & 3 \end{vmatrix} = 4, \quad A\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} = \begin{vmatrix} 3 & 2 \\ -5 & -4 \end{vmatrix} = -2,$$
$$A\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 3 \end{vmatrix} = 4, \quad A\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & -4 \end{vmatrix} = -6, \quad A\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{vmatrix} -1 & 2 \\ 3 & -4 \end{vmatrix} = -2.$$

它们对应的代数余子式分别为

$$\widehat{A}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (-1)^{1+2+1+2}\begin{vmatrix} 1 & -1 \\ 3 & -3 \end{vmatrix} = 0, \quad \widehat{A}\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = (-1)^{1+2+1+3}\begin{vmatrix} 0 & -1 \\ -5 & -3 \end{vmatrix} = 5,$$
$$\widehat{A}\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} = (-1)^{1+2+1+4}\begin{vmatrix} 0 & -5 \\ 1 & 3 \end{vmatrix} = 5, \quad \widehat{A}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = (-1)^{1+2+2+3}\begin{vmatrix} 2 & -1 \\ 1 & -3 \end{vmatrix} = -5,$$
$$\widehat{A}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = (-1)^{1+2+2+4}\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = -5, \quad \widehat{A}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = (-1)^{1+2+3+4}\begin{vmatrix} 2 & 0 \\ 1 & -5 \end{vmatrix} = -10.$$

于是由Laplace展开定理, 得

$$D = 8\times 0 + 4\times 5 + (-2)\times 5 + 4\times(-5) + (-6)\times(-5) + (-2)\times(-10) = 40,$$

得到和例4.10一样的答案. □

通常当某些行(列)有很多零时, 按这些行(列)Laplace展开可以简化行列式的计算.

例4.31. 设$A\in\mathbb{F}^{n\times n}$, $B\in\mathbb{F}^{m\times m}$, 证明:

$$\begin{vmatrix} A & O \\ * & B \end{vmatrix} = \begin{vmatrix} A & * \\ O & B \end{vmatrix} = |A||B|, \quad \begin{vmatrix} O & A \\ B & * \end{vmatrix} = \begin{vmatrix} * & A \\ B & O \end{vmatrix} = (-1)^{mn}|A||B|.$$

☞ **证明.** 矩阵$\begin{pmatrix} A & O \\ * & B \end{pmatrix}$的前$n$行的$n$阶子式如果不取前$n$列则必有全零列, 所以此时为零, 因此按照前$n$行展开得

$$\begin{vmatrix} A & O \\ * & B \end{vmatrix} = |A|(-1)^{1+\cdots+n+1+\cdots+n}|B| = |A||B|.$$

类似的, 矩阵$\begin{pmatrix} O & A \\ B & * \end{pmatrix}$的前$n$行的$n$阶子式如果不取后$n$列则必为零, 因此按照前$n$行展开得

$$\begin{vmatrix} O & A \\ B & * \end{vmatrix} = |A|(-1)^{1+2+\cdots+n+(m+1)+(m+2)+\cdots+(m+n)}|B| = (-1)^{mn}|A||B|.$$

剩余等式取转置即可. □

习题4.9

A1. 利用Laplace展开定理计算下列行列式:

(1) $\begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{vmatrix}$; (2) $\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ x_1 & x_2 & 0 & 0 & 0 & x_3 \\ a_1 & b_1 & 1 & 1 & 1 & c_1 \\ a_2 & b_2 & x_1 & x_2 & x_3 & c_2 \\ a_3 & b_3 & x_1^2 & x_2^2 & x_3^2 & c_3 \\ x_1^2 & x_2^2 & 0 & 0 & 0 & x_3^2 \end{vmatrix}$; (4) $\begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & a \\ x_1 & c & b & \cdots & b & y_1 \\ x_2 & b & c & \cdots & b & y_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_n & b & b & \cdots & c & y_n \\ a & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix}$.

A2. 设A,B,C和D依次是由下表

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

中删去第1, 2, 3和4列而得到的三阶行列式. 证明:

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 & 0 & 0 \\ a_2 & b_2 & c_2 & d_2 & 0 & 0 \\ a_3 & b_3 & c_3 & d_3 & 0 & 0 \\ 0 & 0 & a_1 & b_1 & c_1 & d_1 \\ 0 & 0 & a_2 & b_2 & c_2 & d_2 \\ 0 & 0 & a_3 & b_3 & c_3 & d_3 \end{vmatrix} = AD - BC.$$

A3. 利用行列式证明恒等式:

$$(ab' - a'b)(cd' - c'd) - (ac' - a'c)(bd' - b'd) + (ad' - a'd)(bc' - b'c) = 0.$$

B1. 设方阵$A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^{m\times m}$, 证明: $\det(A \otimes B) = \det(A)^m \det(B)^n$.

§ 4.10 Binet-Cauchy公式

我们知道, 当A和B为同阶方阵时, 有$|AB| = |A||B|$. 如果$A \in \mathbb{F}^{n\times m}$, $B \in \mathbb{F}^{m\times n}$, 而$m$与$n$不必相等, 这时$AB \in \mathbb{F}^{n\times n}$ 也是方阵, 所以有$|AB|$. 那么这时$|AB|$该如何计算? 下面我们来研究这个问题.

类似于定理4.11, 我们考查分块矩阵

$$C = \begin{pmatrix} A_{n\times m} & O_{n\times n} \\ -E_m & B_{m\times n} \end{pmatrix}_{(m+n)\times(m+n)}.$$

一方面, 由分块矩阵的初等变换

$$C \overset{c_2-c_1B}{\rightsquigarrow} \begin{pmatrix} A & AB \\ -E_m & O \end{pmatrix} \overset{r_1\leftrightarrow r_2}{\rightsquigarrow} \begin{pmatrix} -E_m & O \\ A & AB \end{pmatrix},$$

有

$$|C| = (-1)^{mn} \begin{vmatrix} -E_m & O \\ A & AB \end{vmatrix} = (-1)^{mn}|-E_m||AB| = (-1)^{mn+m}|AB|.$$

这得到

$$|AB| = (-1)^{mn+m}|C|.$$

另一方面, 不同于定理4.11的证明, 虽然C是分块下三角阵, 但是对角块A和B不一定是方阵, 所以此时不能用分块下三角阵的行列式公式. 类似于例4.31, 对$|C|$的前n行Laplace展开, 得

$$|C| = \sum_{1\leqslant k_1<k_2<\ldots<k_n\leqslant m+n} C\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} \widehat{C}\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}.$$

设$\{1, 2, \ldots, m+n\} - \{k_1, \ldots, k_n\} = \{k_{n+1} < \cdots < k_{m+n}\}$, 则

$$\widehat{C}\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} = (-1)^{1+2+\cdots+n+k_1+k_2+\cdots+k_n} C\begin{pmatrix} n+1 & n+2 & \cdots & m+n \\ k_{n+1} & k_{n+2} & \cdots & k_{m+n} \end{pmatrix}.$$

下面根据m和n的大小关系分别讨论.

情形1. $n > m$

此时, 任意子式$C\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}$中一定有全零列, 即必为0. 所以有$|C| = 0$, 进而得$|AB| = 0$.

情形2. $m = n$

此时, C为对角块是方阵的分块下三角阵, 所以$|C| = |A||B|$. 进而

$$|AB| = (-1)^{n^2+n}|A||B| = (-1)^{n(n+1)}|A||B| = |A||B|.$$

情形3. $n < m$

此时, 去掉有全零列从而为零的子式$C\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}$, 我们有

$$\begin{aligned}|C| = &\sum_{1\leqslant k_1<k_2<\ldots<k_n\leqslant m} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}\\ &\times(-1)^{1+2+\cdots+n+k_1+k_2+\cdots+k_n} C\begin{pmatrix}n+1 & n+2 & \cdots & m+n\\ k_{n+1} & k_{n+2} & \cdots & k_{m+n}\end{pmatrix},\end{aligned}$$

且必有

$$k_{m+1} = m+1,\ k_{m+2} = m+2,\ \ldots,\ k_{m+n} = m+n.$$

再将每个子式

$$C\begin{pmatrix}n+1 & n+2 & \cdots & m+n\\ k_{n+1} & k_{n+2} & \cdots & k_{m+n}\end{pmatrix} = \begin{array}{c|cccccc|} k_{n+1} & -1 & & & & & \\ k_{n+2} & & -1 & & & & B\\ \vdots & & & \ddots & & & \\ \vdots & & & & \ddots & & \\ k_m & & & & & -1 & \end{array}$$

按照前$m-n$列展开, 得

$$\begin{aligned}&C\begin{pmatrix}n+1 & n+2 & \cdots & m+n\\ k_{n+1} & k_{n+2} & \cdots & k_{m+n}\end{pmatrix}\\ =&(-1)^{k_{n+1}+k_{n+2}+\cdots+k_m+1+2+\cdots+(m-n)}|-E_{m-n}|B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}\\ =&(-1)^{k_{n+1}+k_{n+2}+\cdots+k_m+1+2+\cdots+(m-n)+(m-n)}B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}.\end{aligned}$$

最后得到

$$|C| = \sum_{1\leqslant k_1<k_2<\ldots<k_n\leqslant m} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}(-1)^{1+2+\cdots+n+k_1+k_2+\cdots+k_n}$$
$$\times(-1)^{k_{n+1}+k_{n+2}+\cdots+k_m+1+2+\cdots+(m-n)+(m-n)}B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}.$$

由于

$$\begin{aligned}
&1+\cdots+n+k_1+\cdots+k_n+k_{n+1}+\cdots+k_m+1+\cdots+(m-n)+(m-n)\\
=&(1+\cdots+n)+(1+\cdots+m)+(1+\cdots+(m-n))+(m-n)\\
=&2(1+\cdots+n)+[(n+1)+\cdots+(n+(m-n))]\\
&+(1+\cdots+(m-n))+(m-n)\\
=&2(1+\cdots+n)+2(1+\cdots+(m-n))+n(m-n)+(m-n)\\
=&2(1+\cdots+n)+2(1+\cdots+(m-n))+(n+1)(m-n),
\end{aligned}$$

所以

$$|C| =(-1)^{(n+1)(m-n)}\sum_{1\leqslant k_1<k_2<\ldots<k_n\leqslant m} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}$$
$$\times B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}.$$

而

$$mn+m+(n+1)(m-n)=2m(n+1)-n(n+1)$$

是偶数, 所以

$$|AB| = \sum_{1\leqslant k_1<k_2<\ldots<k_n\leqslant m} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}.$$

于是我们得到了下面的定理.

定理4.23 (Binet-Cauchy公式)**.** 设$A\in\mathbb{F}^{n\times m}$, $B\in\mathbb{F}^{m\times n}$, 则

(1) 当$n>m$时, 有$|AB|=0$;

(2) 当$m=n$时, 有$|AB|=|A||B|$;

(3) 当$n<m$时, 有

$$|AB| = \sum_{1\leqslant k_1<k_2<\cdots<k_n\leqslant m} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix}.$$

也可以如下证明$n>m$的情形. 当$n>m$时, 有

$$(A, O_{n\times(n-m)})_{n\times n}\begin{pmatrix} B \\ O_{(n-m)\times n}\end{pmatrix}_{n\times n} = AB+O=AB.$$

于是

$$|AB| = |A, O|\begin{vmatrix} B \\ O\end{vmatrix} = 0.$$

不用Laplace展开定理, 也可如下利用拆列公式来证明Binet-Cauchy公式.

☞ **Binet-Cauchy公式的另证.** 设$A=(a_{ij})$和$B=(b_{ij})$, 则

$$AB = \begin{pmatrix} \sum\limits_{j_1=1}^m a_{1j_1}b_{j_11} & \sum\limits_{j_2=1}^m a_{1j_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^m a_{1j_n}b_{j_nn} \\ \sum\limits_{j_1=1}^m a_{2j_1}b_{j_11} & \sum\limits_{j_2=1}^m a_{2j_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^m a_{2j_n}b_{j_nn} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{j_1=1}^m a_{nj_1}b_{j_11} & \sum\limits_{j_2=1}^m a_{nj_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^m a_{nj_n}b_{j_nn} \end{pmatrix}.$$

拆列, 并提取每列的公因式, 得到

$$|AB| = \sum_{1\leqslant j_1, j_2, \ldots, j_n\leqslant m} b_{j_11}b_{j_22}\cdots b_{j_nn}\begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n}\end{vmatrix}.$$

如果$n>m$, 则上式右边的求和项中的行列式必有两列相同, 从而行列式的值为零, 进而$|AB|=0$. 下设$n\leqslant m$, 去掉一定为零的行列式, 有

$$|AB| = \sum_{\substack{1\leqslant j_1, j_2, \ldots, j_n\leqslant m \\ \text{互不相同}}} b_{j_11}b_{j_22}\cdots b_{j_nn}\begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n}\end{vmatrix}.$$

设$1\leqslant j_1, j_2, \ldots, j_n\leqslant m$且互不相同, 将$j_1, j_2, \ldots, j_n$重排后记为

$$1\leqslant k_1<k_2<\cdots<k_n\leqslant m.$$

将上面$|AB|$的表达式的求和中的行列式调列, 可得

$$\begin{aligned}\begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n}\end{vmatrix} =& \delta\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ j_1 & j_2 & \cdots & j_n\end{pmatrix}\begin{vmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ a_{2k_1} & a_{2k_2} & \cdots & a_{2k_n} \\ \vdots & \vdots & & \vdots \\ a_{nk_1} & a_{nk_2} & \cdots & a_{nk_n}\end{vmatrix} \\ =& \delta\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ j_1 & j_2 & \cdots & j_n\end{pmatrix} A\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n\end{pmatrix}.\end{aligned}$$

所以得到

$$|AB| = \sum_{1\leqslant k_1<k_2<\cdots<k_n\leqslant m}\ \sum_{\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ j_1 & j_2 & \cdots & j_n\end{pmatrix}} \delta\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ j_1 & j_2 & \cdots & j_n\end{pmatrix} b_{j_11}b_{j_22}\cdots b_{j_nn}$$

$$\cdot A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}$$

$$=\sum_{1\leqslant k_1<k_2<\cdots<k_n\leqslant m} B\begin{pmatrix}k_1 & k_2 & \cdots & k_n\\ 1 & 2 & \cdots & n\end{pmatrix} A\begin{pmatrix}1 & 2 & \cdots & n\\ k_1 & k_2 & \cdots & k_n\end{pmatrix}.$$

特别地, 当$n=m$时, 上面右边的和式只有一项, 即有$|AB|=|A||B|$. □

Binet-Cauchy公式可以用来计算矩阵乘积的子式.

推论4.24. 设$A\in\mathbb{F}^{n\times m}$, $B\in\mathbb{F}^{m\times l}$, $1\leqslant r\leqslant\min\{m,n,l\}$, $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$, $1\leqslant j_1<j_2<\cdots<j_r\leqslant l$, 则$r$阶子式

$$(AB)\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}$$

$$=\sum_{1\leqslant k_1<k_2<\cdots<k_r\leqslant m} A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ k_1 & k_2 & \cdots & k_r\end{pmatrix} B\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}.$$

☞ **证明.** 将A行分块, B列分块

$$A=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix},\qquad B=(\beta_1,\beta_2,\ldots,\beta_l),$$

则AB的(i,j)元是$\alpha_i\beta_j$. 于是令

$$A_1=\begin{pmatrix}\alpha_{i_1}\\ \alpha_{i_2}\\ \vdots\\ \alpha_{i_r}\end{pmatrix},\quad B_1=(\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_r}),$$

则有

$$(AB)\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}=|A_1B_1|.$$

对$|A_1B_1|$用Binet-Cauchy公式即得结论. □

如果$r=1$, 就得到两矩阵相乘的公式. 我们看几个例子.

例4.32. 设$n\geqslant 2$, $a_1,\ldots,a_n,b_1,\ldots,b_n\in\mathbb{C}$. 利用行列式证明*Lagrange*恒等式:

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right)-\left(\sum_{i=1}^n a_ib_i\right)^2=\sum_{1\leqslant i<j\leqslant n}(a_ib_j-a_jb_i)^2.$$

进而证明*Cauchy-Schwarz*不等式: 当$a_1,\ldots,a_n,b_1,\ldots,b_n\in\mathbb{R}$时, 有

$$(a_1b_1+a_2b_2+\cdots+a_nb_n)^2\leqslant(a_1^2+a_2^2+\cdots+a_n^2)(b_1^2+b_2^2+\cdots+b_n^2),$$

其中当且仅当$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\cdots=\frac{a_n}{b_n}$时等号成立.

☞ **证明.** 令$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}_{2\times n}$, 则

$$AA^T = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_ib_i \\ \sum\limits_{i=1}^n a_ib_i & \sum\limits_{i=1}^n b_i^2 \end{pmatrix}.$$

因此有

$$|AA^T| = \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_ib_i\right)^2.$$

另一方面, 由Binet-Cauchy公式

$$\begin{aligned} |AA^T| &= \sum_{1\leqslant i<j\leqslant n} A\begin{pmatrix} 1 & 2 \\ i & j \end{pmatrix} A^T\begin{pmatrix} i & j \\ 1 & 2 \end{pmatrix} \\ &= \sum_{1\leqslant i<j\leqslant n} \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}^2 = \sum_{1\leqslant i<j\leqslant n} (a_ib_j - a_jb_i)^2. \end{aligned}$$

于是Lagrange恒等式成立. 进而容易证明Cauchy-Schwarz不等式成立. □

例4.33. 设矩阵

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix},$$

称A为循环方阵. 求A的行列式.

☞ **解.** 设$\omega = e^{\frac{2\pi i}{n}}$是$n$次本原单位根. 令多项式

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1},$$

矩阵[8]

$$B = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix},$$

我们来计算AB. 由于矩阵AB的(i,j)元为(其中约定$a_n = a_0$)

$$(a_{n-i+1}, \ldots, a_{n-1}, a_1, \ldots, a_{n-i}) \begin{pmatrix} 1 \\ \omega^{j-1} \\ \vdots \\ \omega^{(n-1)(j-1)} \end{pmatrix}$$

[8]矩阵B如何来的? 一个可能的来源是: A是矩阵$C = \begin{pmatrix} 0 & E_{n-1} \\ 1 & 0 \end{pmatrix}$的矩阵多项式, 有相同的特征向量; 考虑$C$的对角化.

$$\begin{aligned}
&=a_{n-i+1}+a_{n-i+2}\omega^{j-1}+\cdots+a_{n-1}\omega^{(i-2)(j-1)}+a_0\omega^{(i-1)(j-1)}+\cdots+a_{n-i}\omega^{(n-1)(j-1)}\\
&=\omega^{(i-1)(j-1)}(a_{n-i+1}\omega^{-(i-1)(j-1)}+\cdots+a_{n-1}\omega^{(i-2)(j-1)-(i-1)(j-1)}\\
&\qquad+a_0+a_1\omega^{j-1}+\cdots+a_{n-i}\omega^{(n-i)(j-1)})\\
&=\omega^{(i-1)(j-1)}(a_{n-i+1}\omega^{n(j-1)-(i-1)(j-1)}+\cdots+a_{n-1}\omega^{n(j-1)-(j-1)}\\
&\qquad+a_0+a_1\omega^{j-1}+\cdots+a_{n-i}\omega^{(n-i)(j-1)})\\
&=\omega^{(i-1)(j-1)}f(\omega^{j-1}),
\end{aligned}$$

所以

$$AB=B\,\mathrm{diag}(f(1),f(\omega),f(\omega^2),\cdots,f(\omega^{n-1})).$$

两边取行列式就得到

$$|A||B|=|AB|=|B|\prod_{i=0}^{n-1}f(\omega^i).$$

由于Vandermonde行列式

$$|B|=\prod_{0\leqslant i<j\leqslant n-1}(\omega^j-\omega^i)\neq 0,$$

所以

$$|A|=\prod_{i=0}^{n-1}f(\omega^i).$$

解毕. □

习题4.10

A1. 计算下列行列式:

(1) $\begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} & 1\\ s_1 & s_2 & s_3 & \cdots & s_n & x\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} & x^{n-1}\\ s_n & s_{n+1} & s_{n+2} & \cdots & s_{2n-1} & x^n \end{vmatrix}$, 其中等幂和$s_k=x_1^k+x_2^k+\cdots+x_n^k$, $k=1,2,\ldots$;

(2) $\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n\\ 1+x_2y_1 & 1+x_2y_2 & \cdots & 1+x_2y_n\\ \vdots & \vdots & & \vdots\\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n \end{vmatrix}$;

(3) $\begin{vmatrix} \cos(\theta_1-\varphi_1) & \cos(\theta_1-\varphi_2) & \cdots & \cos(\theta_1-\varphi_n)\\ \cos(\theta_2-\varphi_1) & \cos(\theta_2-\varphi_2) & \cdots & \cos(\theta_2-\varphi_n)\\ \vdots & \vdots & & \vdots\\ \cos(\theta_n-\varphi_1) & \cos(\theta_n-\varphi_2) & \cdots & \cos(\theta_n-\varphi_n) \end{vmatrix}$;

(4) $\begin{vmatrix} (a_0+b_0)^n & (a_0+b_1)^n & \cdots & (a_0+b_n)^n\\ (a_1+b_0)^n & (a_1+b_1)^n & \cdots & (a_1+b_n)^n\\ \vdots & \vdots & & \vdots\\ (a_n+b_0)^n & (a_n+b_1)^n & \cdots & (a_n+b_n)^n \end{vmatrix}$.

A2. 设行列式$D = \begin{vmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0,n-1} \\ a_{10} & a_{11} & a_{12} & \cdots & a_{1,n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n-1,0} & a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix}$, 记$\varphi_i(x) = a_{0i} + a_{1i}x + \cdots + a_{n-1,i}x^{n-1}$, 其中$i = 0, 1, \ldots, n-1$, 求行列式

$$\begin{vmatrix} \varphi_0(x_1) & \varphi_0(x_2) & \cdots & \varphi_0(x_n) \\ \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_n) \\ \vdots & \vdots & & \vdots \\ \varphi_{n-1}(x_1) & \varphi_{n-1}(x_2) & \cdots & \varphi_{n-1}(x_n) \end{vmatrix}.$$

A3. 当$j_1 = i_1$, $j_2 = i_2$, $\ldots$, $j_k = i_k$时, 矩阵A的子式$A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$称为$A$的一个$k$阶**主子式**, 其中$1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$. 设$A \in \mathbb{C}^{m\times n}$, 证明: 矩阵$A\overline{A}^T$的每一个主子式都是非负实数.

A4. 适合$AA^T = E_n = A^TA$的n阶实方阵A称为**正交阵**. 证明:

(1) 正交方阵的行列式等于± 1;

(2) 位于正交方阵的k个行上的所有k阶子式的平方和等于1, $k = 1, 2, \cdots, n$.

A5. 适合$A\overline{A}^T = E_n = \overline{A}^TA$的$n$阶复方阵$A$称为**酉方阵**. 证明:

(1) 酉方阵的行列式的模为1;

(2) 位于酉方阵的k个行上的所有k阶子式的模的平方和为1.

B1. 计算下列行列式:

(1) $\begin{vmatrix} 1^2 & 2^2 & 3^2 & \cdots & n^2 \\ n^2 & 1^2 & 2^2 & \cdots & (n-1)^2 \\ \vdots & \vdots & \vdots & & \vdots \\ 2^2 & 3^2 & 4^2 & \cdots & 1^2 \end{vmatrix}$; (2) $\begin{vmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ -a_n & a_1 & a_2 & \cdots & a_{n-1} \\ -a_{n-1} & -a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_2 & -a_3 & -a_4 & \cdots & a_1 \end{vmatrix}$.

B2. 对正整数n, 定义**Euler函数**$\varphi(n)$为$1, 2, \ldots, n$中与n互素的整数的个数. 可以证明Euler函数满足等式

$$\sum_{d|n} \varphi(d) = n, \quad (\forall n \in \mathbb{N}),$$

其中上式左边表示对n的所有正因子d求和. 设矩阵$A = (a_{ij})_{n\times n}$, 其中a_{ij}等于整数i和j的最大公因子(i, j). 证明: $\det A = \varphi(1)\varphi(2)\cdots\varphi(n)$.

§ 4.11　Cramer法则

本章最后说明这里的行列式定义是“好的”, 即证明在一定条件下, 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, & (1) \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, & (2) \\ \qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n & (n) \end{cases}$$

有求解公式

$$x_j = \frac{D_j}{D}, \quad j = 1, 2, \ldots, n,$$

其中D为系数矩阵$(a_{ij})_{n\times n}$的行列式, 而D_j为将系数矩阵的第j列用常数项列向量$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$替换后得到的$n$阶方阵的行列式.

我们可以如下计算D_j:

$$\begin{aligned} D_j &= \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{11}x_1 + \cdots + a_{1n}x_n & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & a_{21}x_1 + \cdots + a_{2n}x_n & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n1}x_1 + \cdots + a_{nn}x_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} \\ &\xlongequal[i=1,\ldots,j-1,j+1,\ldots,n]{c_j - x_i c_i} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j}x_j & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & a_{2j}x_j & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj}x_j & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} \\ &= x_j \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = x_j D. \end{aligned}$$

于是当$D \neq 0$时, 如果上面的线性方程组有解, 则一定有

$$x_j = \frac{D_j}{D}, \quad j = 1, 2, \ldots, n.$$

这就说明了我们给出的行列式的定义是好的.

上面给出的推导显得有些神秘, 说好的加减消元法呢? 所以我们还是用加减消元法来解一下上面的线性方程组. 要实现加减消元, 对于每个$j(1 \leqslant j \leqslant n)$, 我们要找$n$个数$c_{1j}, c_{2j}, \ldots, c_{nj}$, 使得

$$c_{1j} \times (1) + c_{2j} \times (2) + \cdots + c_{nj} \times (n) \Longrightarrow \text{只余未知数 } x_j.$$

如同§4.1中分析一样, 这将得到a_{ij}和c_{ij}之间的某种"正交"关系.

为了从几何上更好地理解加减消元法, 我们先给出上面线性方程组的向量形式. 记

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \cdots, \quad \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}, \quad \beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

则上面的线性方程组等价于

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta.$$

称这种形式为线性方程组的向量形式.

接着我们推广解析几何中的内积概念. 设$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \delta = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$, 定义它们的内积为

$$\alpha \cdot \delta := a_1d_1 + a_2d_2 + \cdots + a_nd_n = \alpha^T\delta.$$

因此, 用列向量δ内积上面的线性方程组, 我们得到

$$x_1(\alpha_1 \cdot \delta) + x_2(\alpha_2 \cdot \delta) + \cdots + x_n(\alpha_n \cdot \delta) = \beta \cdot \delta.$$

所以, 加减消元等价于对每个$j(1 \leqslant j \leqslant n)$, 找$n$维列向量$\delta_j$, 使得

$$\alpha_k \cdot \delta_j = 0, \quad (\forall k \neq j).$$

可以理解为δ_j需要与$\alpha_k(\forall k \neq j)$都正交, 所以几何上可以取$\delta_j$为这$n-1$个向量的"外积".

设$A = (a_{ij})_{n\times n}$是上面线性方程组的系数矩阵, A_{ij}是(i,j)位置的代数余子式. 推论4.19告诉我们

$$a_{1k}A_{1j} + a_{2k}A_{2j} + \cdots + a_{nk}A_{nj} = \delta_{kj}|A|,$$

用内积记号就是

$$\alpha_k \cdot \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{nj} \end{pmatrix} = \delta_{kj}|A| = \begin{cases} |A|, & k = j, \\ 0, & k \neq j. \end{cases}$$

所以可取$\delta_j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{nj} \end{pmatrix}$. 用$\delta_j$加减消元后, 我们得到

$$x_j|A| = \beta \cdot \delta_j.$$

注意到上面右边也是n阶行列式

$$\begin{aligned}\beta \cdot \delta_j =& b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj} \\ =& \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.\end{aligned}$$

定理4.25 (Cramer法则). 设有n元线性方程组$AX = \beta$, 其中系数矩阵$A = (a_{ij}) \in \mathbb{F}^{n\times n}$, 常数项列向量$\beta = (b_1, b_2, \cdots, b_n)^T \in \mathbb{F}^n$, 未知数列向量$X = (x_1, x_2, \ldots, x_n)^T$. 如果系数矩阵的行列式满足$|A| \neq 0$, 则方程组有唯一解

$$x_j = \frac{D_j}{|A|}, \quad j = 1, 2, \ldots, n,$$

其中

$$D_j = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

☞ **证明.** 上面的分析事实上证明了解的唯一性, 可以利用代数余子式的性质证明解的存在性(即所给的解确实满足方程), 这里我们给出Cramer法则更简单的一个证明. 不妨设$n > 1$. 因为$|A| \neq 0$, 所以A可逆. 于是

$$AX = \beta \Longleftrightarrow X = A^{-1}\beta = \frac{1}{|A|}\,\mathrm{adj}(A)\beta.$$

即得方程组有唯一解

$$x_j = \frac{1}{|A|}(b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj}) = \frac{D_j}{|A|}, \quad j = 1, 2, \ldots, n,$$

其中A_{ij}是A的(i, j)位置的代数余子式. □

Cramer法则的好处是给出了一定条件下线性方程组的具体求解公式, 这在许多问题的理论分析中很有用处. 但是我们这里要求方程组的系数矩阵为方阵, 而且系数矩阵的行列式不为零. 我们一般不用该法则具体解线性方程组, 这一方面是由于行列式计算量大, 另一方面是由于我们已经有解线性方程组的好办法—对增广矩阵的行初等变换.

习题4.11

A1. 用Cramer法则解线性方程组$\begin{cases} x_1+x_2+x_3=1, \\ x_1-x_2+2x_3=2, \\ 2x_1+x_2+3x_3=-1. \end{cases}$

A2. 利用矩阵和行列式证明Lagrange插值定理: 设$a_1, a_2, \ldots, a_n$是$\mathbb{F}$中n个两两不同的数, b_1, b_2, $\ldots$, b_n是$\mathbb{F}$中任意n个数, 则存在唯一的$f(x)\in\mathbb{F}[x]$, 满足$\deg(f(x))\leqslant n-1$, 且

$$f(a_i)=b_i,\quad i=1,2,\ldots,n.$$

并求出这个多项式.

§ 4.12 附录: 行列式的几何意义

我们先看一下二阶行列式的几何意义. 我们来求图4.1(a)中平行四边形$OABC$的面积S. 假设A点的坐标为(a_{11}, a_{21}), C点的坐标为(a_{12}, a_{22}), 则B点的坐标为$(a_{11}+a_{12}, a_{21}+a_{22})$.

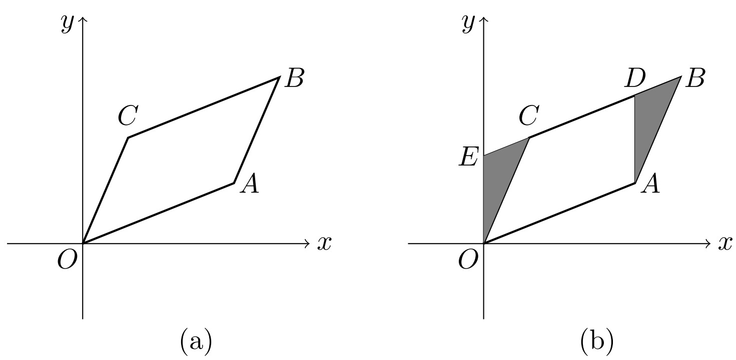


图 4.1: 求平行四边形的面积S

可以按照图4.1(b)来求S. 将BC延长与y轴交于点E; 过点A做y轴的平行线, 设与BC交于点D. 则两个阴影三角形的面积相等, 进而S等于平行四边形$OADE$的面积. 这个平行四边形底面OE的高为a_{11}, 于是只需求出线段OE的长度, 这是一个简单的几何问题. 过点B和C的直线方程是

$$y-a_{22}=\frac{a_{21}}{a_{11}}(x-a_{12}),$$

令$x=0$, 得到

$$y=a_{22}-\frac{a_{12}a_{21}}{a_{11}}=\frac{a_{11}a_{22}-a_{12}a_{21}}{a_{11}},$$

即得点E的纵坐标为$\frac{a_{11}a_{22}-a_{12}a_{21}}{a_{11}}$. 最后有

$$S=a_{11}\cdot\frac{a_{11}a_{22}-a_{12}a_{21}}{a_{11}}=a_{11}a_{22}-a_{12}a_{21}=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

所以二阶行列式表示平行四边形的有向面积. 这里“有向”指的是当OA和OC的相对位置改变时, 二阶行列式的值等于$-S$.

类似的, 三阶行列式表示平行六面体的有向体积. 具体的, 设图4.2中的以OA, OB和OC为邻棱的平行六面体的体积是V, 又设点A坐标为(a_{11}, a_{21}, a_{31}), 点B坐标为(a_{12}, a_{22}, a_{32}), 点C坐标为(a_{13}, a_{23}, a_{33}), 则有

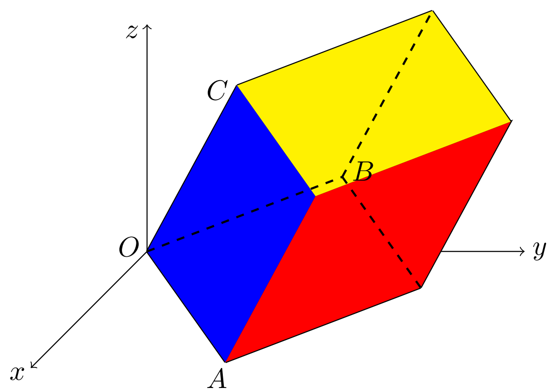


图 4.2: 平行六面体

$$V = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

这里的"有向"指的是当OA, OB, OC成左手系时, 三阶行列式的值等于$-V$. 感兴趣的读者可以试着给出这一事实的证明.

下面我们再给出行列式的另外一种几何意义[9].

设矩阵$A \in \mathbb{R}^{2\times 2}$, 则有$\mathbb{R}^2$上的线性变换$\mathscr{A}$, 使得对任意$X \in \mathbb{R}^2$, 有$\mathscr{A}(X) = AX$. 如果记$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, 则

$$\mathscr{A}(e_1) = \begin{pmatrix} a_1 \\ a_3 \end{pmatrix}, \quad \mathscr{A}(e_2) = \begin{pmatrix} a_2 \\ a_4 \end{pmatrix},$$

其中$e_1 = (1,0)^T$, $e_2 = (0,1)^T$. 于是$\mathbb{R}^2$中的单位正方形S被$\mathscr{A}$映到平行四边形P, 参看图4.3.

考虑面积, 可知在线性变换$\mathscr{A}$下, 单位正方形S的面积变成了平行四边形P的面积, 即$|A|$. 更一般地, 如果S是$\mathbb{R}^2$中一个可求面积的有界区域, 令$P = \mathscr{A}(S)$, 我们想求P和S的面积的关系. 设$X = (x_1, x_2)^T$, 而$Y = (y_1, y_2)^T = AX$, 则

$$y_1 = a_1x_1 + a_2x_2, \quad y_2 = a_3x_1 + a_4x_2.$$

于是(可参看[6])

$$dy_1 = a_1dx_1 + a_2dx_2, \quad dy_2 = a_3dx_1 + a_4dx_2.$$

这得到

$$dy_1 \wedge dy_2 = (a_1dx_1 + a_2dx_2) \wedge (a_3dx_1 + a_4dx_2) = a_1a_4dx_1 \wedge dx_2 + a_2a_3dx_2 \wedge dx_1$$

[9] 请学了线性变换后再看这部分内容.

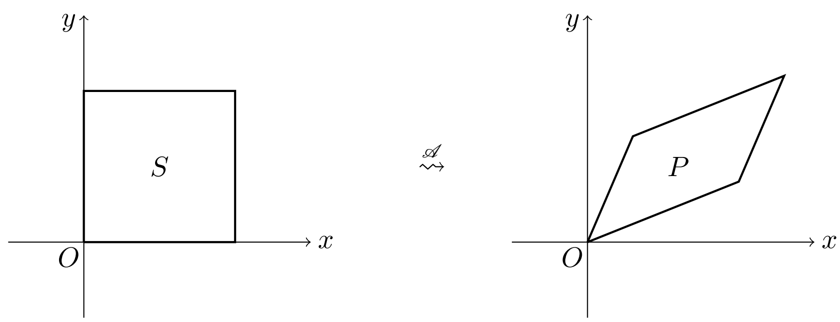


图 4.3: 线性变换$\mathscr{A}$下单位正方形S变为平行四边形P

$$=(a_1a_4-a_2a_3)dx_1\wedge dx_2=|A|dx_1\wedge dx_2,$$

这里应用了性质

$$dx_1\wedge dx_1=0,\quad dx_2\wedge dx_2=0,\quad dx_1\wedge dx_2=-dx_2\wedge dx_1.$$

进而有

$$(P\text{的面积})=\int_P dy_1\wedge dy_2=\int_S|A|dx_1\wedge dx_2=|A|\cdot(S\text{的面积}).$$

于是可以看到, $|A|$是线性变换$\mathscr{A}$下面积的伸缩比.

更一般的, 设矩阵$A\in\mathbb{R}^{n\times n}$, 则$A$导出线性变换$\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^n; X\mapsto AX$. 于是$A$的行列式表示线性变换$\mathscr{A}$下"体积"的伸缩比. 即如果$V$是$\mathbb{R}^n$中的有界区域, 则

$$(\mathscr{A}(V)\text{的体积})=|A|\cdot(V\text{的体积}).$$

§ 4.13 补充题

A1. 设有n阶方阵$A=(a_{ij})_{n\times n}$和

$$B=\begin{pmatrix} a_{11}b_1^2 & a_{12}b_1b_2 & \cdots & a_{1n}b_1b_n \\ a_{21}b_2b_1 & a_{22}b_2^2 & \cdots & a_{2n}b_2b_n \\ \vdots & \vdots & & \vdots \\ a_{n1}b_nb_1 & a_{n2}b_nb_2 & \cdots & a_{nn}b_n^2 \end{pmatrix},$$

其中$b_1,b_2,\dots,b_n$为常数. 如果$\det(A)=c$, 求$\det(B)$.

A2. 求下列行列式的值:

(1) $\begin{vmatrix} x & y & z & w \\ y & x & w & z \\ z & y & x & w \\ w & z & y & x \end{vmatrix}$; (2) $\begin{vmatrix} (a+b)^2 & c^2 & c^2 \\ a^2 & (b+c)^2 & a^2 \\ b^2 & b^2 & (c+a)^2 \end{vmatrix}$.

A3. 求下列行列式的值:

(1) $\begin{vmatrix} 1 & a_1 & a_2 & \cdots & a_n \\ 1 & a_1+b_1 & a_2 & \cdots & a_n \\ 1 & a_1 & a_2+b_2 & \cdots & a_n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_1 & a_2 & \cdots & a_n+b_n \end{vmatrix}$; (2) $\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & \binom{2}{1} & \cdots & \binom{n}{1} \\ 1 & \binom{3}{2} & \cdots & \binom{n+1}{2} \\ \vdots & \vdots & & \vdots \\ 1 & \binom{n}{n-1} & \cdots & \binom{2n-2}{n-1} \end{vmatrix}$;

(3) $\begin{vmatrix} a_1^{n-1} & a_1^{n-2}b_1 & \cdots & a_1b_1^{n-2} & b_1^{n-1} \\ a_2^{n-1} & a_2^{n-2}b_2 & \cdots & a_2b_2^{n-2} & b_2^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_n^{n-1} & a_n^{n-2}b_n & \cdots & a_nb_n^{n-2} & b_n^{n-1} \end{vmatrix}$;

(4) $\begin{vmatrix} 1-a_1 & a_2 & & & & \\ -1 & 1-a_2 & a_3 & & & \\ & -1 & 1-a_3 & a_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 1-a_{n-1} & a_n \\ & & & & -1 & 1-a_n \end{vmatrix}$;

(5) $\begin{vmatrix} a_0+a_1 & a_1 & & & & \\ a_1 & a_1+a_2 & a_2 & & & \\ & a_2 & a_2+a_3 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-2} & a_{n-2}+a_{n-1} & a_{n-1} \\ & & & & a_{n-1} & a_{n-1}+a_n \end{vmatrix}$;

(6) $\begin{vmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{vmatrix}$.

A4. 计算行列式

$$\begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{vmatrix},$$

其中等幂和$s_k = x_1^k + x_2^k + \cdots + x_n^k$, $k = 1, 2, \ldots$.

A5. 设矩阵$A, B \in \mathbb{F}^{n\times n}$, 整数$r$满足$1 \leqslant r < n$, 证明: AB和BA的所有r阶主子式之和相等.

A6. 设$f_k(x)=x^k+a_{k1}x^{k-1}+a_{k2}x^{k-2}+\cdots+a_{kk}$, 求下列行列式的值:

$$\begin{vmatrix} 1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1) \\ 1 & f_1(x_2) & f_2(x_2) & \cdots & f_{n-1}(x_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n) \end{vmatrix}.$$

A7. 设A是奇数阶矩阵, $|A|>0$, 且$AA^T=E_n$, 证明: E_n-A不可逆.

A8. 设同阶方阵A和B满足$A^2=B^2=E$且$|A|+|B|=0$, 证明: $A+B$不可逆.

A9. 设A是非零实矩阵且$\mathrm{adj}(A)=A^T$, 证明: A可逆.

A10. 已知$\mathrm{adj}(A)=\begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -2 \\ -1 & 2 & 1 \end{pmatrix}$, 求矩阵$A$.

A11. 证明: (1) 平面上互不相同的3点$(x_1,y_1),(x_2,y_2),(x_3,y_3)$共线的充分必要条件是

$$\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}=0;$$

(2) 平面上不共线的4点$(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)$在同一个圆上的充分必要条件是

$$\begin{vmatrix} 1 & x_1 & y_1 & x_1^2+y_1^2 \\ 1 & x_2 & y_2 & x_2^2+y_2^2 \\ 1 & x_3 & y_3 & x_3^2+y_3^2 \\ 1 & x_4 & y_4 & x_4^2+y_4^2 \end{vmatrix}=0.$$

A12. 证明: (1) 空间中互不相同的3点(x_i,y_i,z_i), $i=1,2,3$共线的充分必要条件是

$$\begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \end{vmatrix}=\begin{vmatrix} 1 & x_1 & z_1 \\ 1 & x_2 & z_2 \\ 1 & x_3 & z_3 \end{vmatrix}=\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}=0;$$

(2) 空间中不共线的4点(x_i,y_i,z_i), $i=1,2,3,4$共面的充分必要条件是

$$\begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}=0;$$

(3) 空间中不共面的5点(x_i,y_i,z_i), $i=1,2,3,4,5$在同一个球面上的充分必要条件是

$$\begin{vmatrix} 1 & x_1 & y_1 & z_1 & x_1^2+y_1^2+z_1^2 \\ 1 & x_2 & y_2 & z_2 & x_2^2+y_2^2+z_2^2 \\ 1 & x_3 & y_3 & z_3 & x_3^2+y_3^2+z_3^2 \\ 1 & x_4 & y_4 & z_4 & x_4^2+y_4^2+z_4^2 \\ 1 & x_5 & y_5 & z_5 & x_5^2+y_5^2+z_5^2 \end{vmatrix}=0;$$

(4) 两空间直线

$$L_1:\begin{cases}A_1x+B_1y+C_1z+D_1=0,\\A_2x+B_2y+C_2z+D_2=0,\end{cases}\qquad L_2:\begin{cases}A_3x+B_3y+C_3z+D_3=0,\\A_4x+B_4y+C_4z+D_4=0\end{cases}$$

共面的充分必要条件是

$$\begin{vmatrix}A_1&B_1&C_1&D_1\\A_2&B_2&C_2&D_2\\A_3&B_3&C_3&D_3\\A_4&B_4&C_4&D_4\end{vmatrix}=0.$$

B1. 设函数$f:\mathbb{F}^{n\times n}\to\mathbb{F}$满足:

(i) $f(E_n)=1$;

(ii) 对任意$A\in\mathbb{F}^{n\times n}$, 调换$A$的某两列得到$B$, 则$f(B)=-f(A)$;

(iii) 对任意$A\in\mathbb{F}^{n\times n}$, 将$A$的某一列乘以数$c\in\mathbb{F}$得到$B$, 则$f(B)=cf(A)$;

(iv) 如果$A,B,C\in\mathbb{F}^{n\times n}$满足$A$的第$i$列为$B$和$C$的第$i$列之和, 而$B$和$C$的其它列都和$A$的相应列相同, 这里$1\leqslant i\leqslant n$, 则$f(A)=f(B)+f(C)$.

证明: 对任意$A\in\mathbb{F}^{n\times n}$, 有$f(A)=\det(A)$.

B2. 设A,B,C,D都是n阶方阵, 令$M=\begin{pmatrix}A&B&C&D\\B&A&D&C\\C&D&A&B\\D&C&B&A\end{pmatrix}$, 证明:

$$|M|=|A+B+C+D||A+B-C-D||A-B+C-D||A-B-C+D|.$$

B3. 设A是n阶可逆阵, 证明: 经有限步第三类初等变换, 可将A化为$\mathrm{diag}(1,\ldots,1,|A|)$.

B4. 设$n>2$, n阶方阵A的所有元素为1或者-1, 证明: $|\det(A)|\leqslant(n-1)!(n-1)$.

B5. 设多项式$f_1(x),f_2(x),\ldots,f_n(x)\in\mathbb{F}[x]$的次数都不超过$n-2$, 其中$n\geqslant 2$, 证明: 对任意$a_1,a_2,\ldots,a_n\in\mathbb{F}$, 有

$$\begin{vmatrix}f_1(a_1)&f_1(a_2)&\cdots&f_1(a_n)\\f_2(a_1)&f_2(a_2)&\cdots&f_2(a_n)\\\vdots&\vdots&&\vdots\\f_n(a_1)&f_n(a_2)&\cdots&f_n(a_n)\end{vmatrix}=0.$$

B6. 求下列行列式的值:

(1) $\begin{vmatrix}(x-a_1)^2&a_2^2&\cdots&a_n^2\\a_1^2&(x-a_2)^2&\cdots&a_n^2\\\vdots&\vdots&&\vdots\\a_1^2&a_2^2&\cdots&(x-a_n)^2\end{vmatrix}$;

(2) $\begin{vmatrix} \cos\theta & \cos 2\theta & \cos 3\theta & \cdots & \cos n\theta \\ \cos n\theta & \cos\theta & \cos 2\theta & \cdots & \cos(n-1)\theta \\ \cos(n-1)\theta & \cos n\theta & \cos\theta & \cdots & \cos(n-2)\theta \\ \vdots & \vdots & \vdots & & \vdots \\ \cos 2\theta & \cos 3\theta & \cos 4\theta & \cdots & \cos\theta \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ 1 & x_2^2 & x_2^3 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n^2 & x_n^3 & \cdots & x_n^n \end{vmatrix}$.

B7. 设$A, B \in \mathbb{R}^{m\times n}$, 证明:

$$|AA^T||BB^T| \geqslant |AB^T|^2.$$

B8. 设$A, B \in \mathbb{R}^{n\times n}$, 满足$A^2+B^2=AB$, $AB-BA$可逆, 证明: $3 \mid n$.

B9. 对n阶方阵A, 约定$|A|$的零阶子式只有1, 其代数余子式为$|A|$, 再约定$|A|$的代数余子式为1. 证明: 对n阶方阵A和B, 它们的和$A+B$的行列式等于$|A|$的所有子式和$|B|$中相同位置的子式的代数余子式乘积之和, 即

$$|A+B| = \sum_{k=0}^{n} \sum_{\substack{1\leqslant i_1<i_2<\cdots<i_k\leqslant n \\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n}} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \widehat{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}.$$

第五章　向量组与矩阵的秩

本章的主题是数数: 数线性方程组中真正方程的个数, 数线性方程组解的个数, 数矩阵行(列)的真正个数. 而这些都可以统一为数向量组中向量的个数. 于是我们开始进入线性代数的几何理论. 从数线性方程组中的真正方程个数出发, 我们引入了向量组的线性表示, 线性相关, 极大无关组以及秩等概念; 并由此定义了矩阵的秩, 研究了线性方程组的解的个数与真正方程个数的关系. 我们还研究了最简单的线性空间: 向量空间.

本章中$\mathbb{F}$表示任意一个数域.

§5.1　向量与向量组

我们的故事—线性方程组的求解之旅, 已经快要结束了, 但是还有若干情节需要一定的笔墨. 一个是一个线性方程组有几个方程? 好像这不该成为问题, 数一下不就得出方程的个数了吗? 看下面的线性方程组

$$\begin{cases} x+y=1, \\ x-2y=2, \\ 2x-y=3. \end{cases}$$

这里, 第三个方程等于前两个方程之和, 所以它是多余的, 可以删去. 于是上面的方程组与

$$\begin{cases} x+y=1, \\ x-2y=2 \end{cases}$$

是同解的. 新的方程组中不能再删去任意一个方程, 否则就不同解了. 那么从原来的方程组中我们应该数出的方程个数是3还是2呢? 我想读者应该会同意数出的方程个数是2这个结论吧, 毕竟这可以看成真正方程的个数. 回想Gauss消元法, 其在化简的过程中就把多余的方程(全零行对应的方程)干掉了, 最后由剩下的方程(非零行对应的方程)得出解, 所以剩下的方程才是本质的.

另一个问题是线性方程组解的个数. 这个问题我们不是有答案了吗? 解的个数是零个, 1个或者无穷个. 但是在我们的直观中, 无穷们可能并不完全一样. 比如下面两个三元线性方程组

$$(A)\ x=0, \qquad (B)\ \begin{cases} x=0, \\ y=0, \end{cases}$$

从几何上看, 方程组(A)的解是yz坐标平面, 而方程组(B)的解则是该坐标平面上的直线z轴. 所以直观上看, 方程组(A)应该比(B)更多解. 那么, 如何理解这种多少呢? 即如何在这两个无穷中合理数出可以比较的两个数?

所以我们需要数两种数, 一个是在有限个方程中合理地数出一个数, 另一个是在无限个解中合理地数出个有限数. 如果可以统一处理它们就好了! 方程本质上对应到行向量, 比如方程组$\begin{cases} x+y=1, \\ x-2y=2, \\ 2x-y=3 \end{cases}$ 对应到三个行向量$(1,1,1),(1,-2,2),(2,-1,3)$, 我们相当于要数这三个行向量的数目. 方程组的解也对应到向量, 比如三元线性方程组$\begin{cases} x=0, \\ y=0 \end{cases}$ 的解可以写为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}, \quad c \in \mathbb{F},$$

我们需要数这无限个三维列向量的数目. 因此, 我们的任务是数若干个(可以有限个, 也可以无限个)行(或列)向量的个数. 在分析如何合理地数前, 我们先给向量以几何意义, 把它们看成我们生活的三维空间中的向量的推广. 这样读者可以用三维空间中的向量来想我们的数数问题, 从而更好地理解我们的想法.

- **空间向量**

力, 位移等许多物理量都是向量, 空间中的向量指的是有大小和方向的量. 可以用一个有向线段来表示向量, 其中线段的长度表示向量的大小, 线段的方向表示向量的方向, 如图5.1(a)表示一个向量$\vec{a}$.

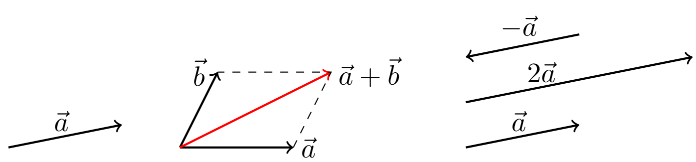


图 5.1: 向量和向量的运算

我们知道, 力可以进行累加, 这对应到向量的加法. 向量的加法按照平行四边形法则进行, 如图5.1(b). 向量还可以进行数乘, 如果用0乘$\vec{a}$, 则得到零向量(大小为零的向量); 如果用正数a乘$\vec{a}$, 则向量的方向不变, 长度变为原长度的a倍; 如果用负数$-a$乘$\vec{a}$, 则向量的方向反向, 长度变为原长度的a倍, 见图5.1(c). 我们还可以考虑向量间的夹角, 这就有向量的内积的概念; 还可以定义向量的外积, 等等.

- **高维向量**

虽然我们生活在三维空间, 但是我们也可以考虑4维, 5维等高维空间的物理现象, 考虑高维空间的向量. 那么如何自然的定义高维空间的向量呢? 由于我们的脑中并无高维空间的图像, 所以很难想象何为高维空间中有大小和方向的量. 但我们可以换一个角度来看问题. 解析几何告诉我们, 在三维空间建立了直角坐标系后, 将向量的起点移到坐标原点, 那么向量就与终点的坐标一一对应, 如图5.2.

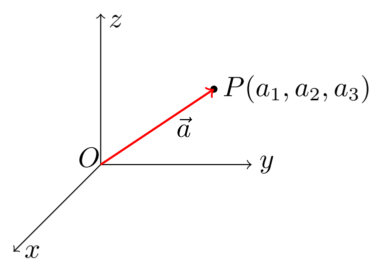


图 5.2: 向量与三元有序数组的对应

于是, 三维空间中的向量集合就和所有的三元有序数组所组成的集合

$$\mathbb{R}^3 = \{(a_1, a_2, a_3) \mid a_1, a_2, a_3 \in \mathbb{R}\}$$

一一对应, 而后者是一个代数对象. 向量的运算对应到代数运算, 例如如果向量$\vec{a}$对应到(a_1, a_2, a_3), 向量$\vec{b}$对应到(b_1, b_2, b_3), 那么$\vec{a}+\vec{b}$就对应到$(a_1+b_1, a_2+b_2, a_3+b_3)$; 数乘向量$a\vec{a}$对应到$(aa_1, aa_2, aa_3)$; 向量的内积$\vec{a}\cdot\vec{b}$等于$a_1b_1 + a_2b_2 + a_3b_3$; 等等. 于是要考虑几何上的三维空间的向量集合, 我们在代数上只要考虑集合$\mathbb{R}^3$及其上的运算

$$\begin{aligned}
&(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3),\\
&a(a_1, a_2, a_3) = (aa_1, aa_2, aa_3),\\
&(a_1, a_2, a_3) \cdot (b_1, b_2, b_3) = a_1b_1 + a_2b_2 + a_3b_3,\\
&\qquad\qquad\vdots
\end{aligned}$$

可以看到, 向量的这种代数化很容易进行高维推广. 空间向量(即三维向量)代数上就是三元有序数组, 那么4维向量不就是4元有序数组吗? 一般的, 我们有下面的定义.

定义5.1. 称数域$\mathbb{F}$中的n元有序数组为数域$\mathbb{F}$上的一个n维向量. 数域$\mathbb{F}$上所有的n维向量所组成的集合记为$\mathbb{F}^n$.

如果将n维向量写成行的形式, 则称为行向量, 例如$(a_1, a_2, \ldots, a_n)$; 如果写成列的形式, 则称为列向量, 例如$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. 这里行向量和列向量只是向量的不同表示形式, 并无差别. 后面我们将向量看成矩阵, 这时作为矩阵的话, 行向量和列向量是不同的, 但由于互为转置, 所以也没有本质的不同.

本讲义中习惯使用列向量, 所以记

$$\mathbb{F}^n=\left\{\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\mid a_1,a_2,\dots,a_n\in\mathbb{F}\right\},$$

称其为n维(列)向量空间. 将向量看成矩阵, 就有$\mathbb{F}^n=\mathbb{F}^{n\times 1}$.

这里, 我们只考虑向量的加法和数乘这两种运算(后面的章节将考虑向量的内积):

$$\text{加法}:\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}+\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=\begin{pmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_n+b_n\end{pmatrix};$$

$$\text{数乘}:a\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}=\begin{pmatrix}aa_1\\aa_2\\\vdots\\aa_n\end{pmatrix},\quad(a\in\mathbb{F}).$$

注意到这事实上就是矩阵的加法和数乘, 于是我们可以自由地应用矩阵运算的性质, 例如$\mathbb{F}^n$中的零元是$0=\begin{pmatrix}0\\\vdots\\0\end{pmatrix}$, 而对于$\alpha=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\in\mathbb{F}^n$, 其负元是$-\alpha=\begin{pmatrix}-a_1\\-a_2\\\vdots\\-a_n\end{pmatrix}$.

称一组同类(类别指行向量还是列向量)和同维的向量为向量组. 为了应用, 我们这里并未要求向量组中的向量两两不同, 所以α,α也是一个向量组. 而且向量组与表示时向量的顺序无关, 例如, α,β和β,α 是同一个向量组.

- **有限向量组与矩阵**

当向量组所含向量个数有限时, 我们称其为有限向量组. 我们有下面的对应

$$\boxed{\text{有限向量组}}\longleftrightarrow\boxed{\text{矩阵}}.$$

这是我们最早遇到的几何和代数的对应. 例如, 列向量组$\beta_1,\beta_2,\dots,\beta_m\in\mathbb{F}^n$对应到矩阵[1]

$$(\beta_1,\beta_2,\dots,\beta_m)\in\mathbb{F}^{n\times m},$$

[1]将有限向量组对应到矩阵, 就可以用矩阵这个代数工具来解决某些关于向量的几何问题. 这个对应并不唯一, 例如列向量组α,β对应到矩阵(α,β), 也对应到矩阵(β,α), 我们可以根据需要选取.

而行向量组$\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{F}^m$则对应到矩阵

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}\in\mathbb{F}^{n\times m}.$$

反之, 如果给定矩阵$A\in\mathbb{F}^{n\times m}$, 设

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}=(\beta_1,\beta_2,\ldots,\beta_m),$$

则我们得到m维行向量组$\alpha_1,\alpha_2,\ldots,\alpha_n$和$n$维列向量组$\beta_1,\beta_2,\ldots,\beta_m$, 前者称为$A$的行向量组, 后者称为$A$的列向量组[2].

- **线性方程组与向量组**

线性方程组也对应到向量组. 具体的, 设有线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \qquad\qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m,\end{cases}$$

记第i个方程对应的向量为

$$\alpha_i=(a_{i1},a_{i2},\ldots,a_{in},b_i),\quad i=1,2,\ldots,m,$$

则得到行向量组$\alpha_1,\alpha_2,\ldots,\alpha_m$, 它就是该线性方程组的增广矩阵的行向量组. 某两方程相加, 对应到向量做加法; 用数乘方程, 则对应到向量的数乘.

另一方面, 线性方程组有矩阵形式$AX=\beta$, 其中

$$A=(a_{ij})_{m\times n},\quad X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \beta=\begin{pmatrix}b_1\\ \vdots\\ b_m\end{pmatrix}.$$

如果$\begin{cases}x_1=a_1,\\ x_2=a_2,\\ \quad\vdots\\ x_n=a_n\end{cases}$是一个解, 则称$\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}\in\mathbb{F}^n$为一个**解向量**. 这相当于在$\mathbb{F}^n$中求解矩阵方程$AX=\beta$. 所有解向量所组成的集合是一个向量组, 称为解集合. 后面我们通常将线性方程组的解写成解向量的形式.

[2]考虑矩阵的行(列)向量组这一几何对象, 有时可以解决一些关于矩阵的问题.

• 线性方程组方程的个数和解的多少

下面回到本节开头提的数数问题. 数线性方程组所含方程的个数时, 我们考察了线性方程组

$$\begin{cases} x+y=1, \\ x-2y=2, \\ 2x-y=3. \end{cases}$$

这里第三个方程是前两个方程之和, 用对应的行向量表示就是

$$(2,-1,3)=(1,1,1)+(1,-2,2),$$

称第三个方程是前两个方程的线性组合. 于是数方程组中所含方程的个数要进行打假, 如果某个方程是其余方程的线性组合, 那么这个方程就是多余的, 要去掉; 而且要将打假进行到底, 即最后剩下的方程组中所有的方程都不是多余的, 一个都不能少, 这就得到极大无关组. 所以方程组中所含方程的真正个数, 就是它的极大无关组所含向量的个数. 我们要数出这个数.

数线性方程组的解的个数时, 我们考察了两个三元线性方程组

$$(A)\ x=0, \qquad (B)\ \begin{cases} x=0, \\ y=0, \end{cases}$$

则(A)的解集合为

$$S_1=\{(0,b,c)^T \mid b,c\in\mathbb{F}\}=\{b(0,1,0)^T+c(0,0,1)^T \mid b,c\in\mathbb{F}\},$$

而(B)的解集合为

$$S_2=\{(0,0,c)^T \mid c\in\mathbb{F}\}=\{c(0,0,1)^T \mid c\in\mathbb{F}\}.$$

所以S_1是由$(0,1,0)^T$和$(0,0,1)^T$通过加法和数乘生成, 而这两个向量不能互相生成. 此时, S_1中每个向量都是$(0,1,0)^T$和$(0,0,1)^T$的线性组合, 而$(0,1,0)^T,(0,0,1)^T$ 则是生成S_1的基础向量, 称为基(当然也是极大无关组). 类似的, S_2中的向量都是$(0,0,1)^T$的线性组合, $(0,0,1)^T$是基. 于是解集合的基所含向量个数可以用来衡量解集合的大小, 称其为维数. 我们要数出这个维数.

因此, 本章中我们要定义向量组的极大无关组, 它是向量组中的若干向量, 这些向量可以产生其余向量, 而且它们不能再减少(即任意一个不能由其余向量产生). 这里产生向量的机制就是加法和数乘, 我们称为线性组合, 所以我们要定义线性组合的概念. 极大无关组中任意一个向量不能由其余向量产生, 我们称极大无关组中向量没有关系, 所以我们要定义线性无关的概念. 有了这些必要的准备, 我们就可以定义极大无关组的概念, 并合理地数出向量组所含向量的个数. 最后, 我们将这些用到矩阵的行(列)向量组以及线性方程组对应的行向量组和解集合上, 从而得出数数问题的一个合理的解答.

在后面的章节中, 我们将考虑更一般的空间和其中的向量. 和解析几何一样, 为了得到代数对应, 我们的主要任务是建立这个空间的(直角)坐标系(即(标准正交)基), 然后该空间的每个向量就对应到它在这个坐标系中的坐标. 从这个观点看, 线性代数就是高维解析几何!

§5.2 向量组的线性组合

如前预告, 我们要定义向量组的线性组合.

定义5.2. 设S是$\mathbb{F}^n$中的列向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$, 而向量$\beta \in \mathbb{F}^n$, 如果存在$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 使得

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s,$$

则称β是向量组S的线性组合, 也称β可由S线性表示. 称$c_1, c_2, \ldots, c_s$为组合系数.

设$\alpha_1, \alpha_2, \ldots, \alpha_s$是$\mathbb{F}^n$中的任意向量组, 对于$0 \in \mathbb{F}^n$有

$$0 = 0 \cdot \alpha_1 + 0 \cdot \alpha_2 + \cdots + 0 \cdot \alpha_s,$$

即0是该向量组的线性组合. 称这个组合为平凡的组合. 又因为

$$\alpha_i = 0 \cdot \alpha_1 + \cdots + 0 \cdot \alpha_{i-1} + 1 \cdot \alpha_i + 0 \cdot \alpha_{i+1} + \cdots + 0 \cdot \alpha_s,$$

所以$\alpha_i\ (1 \leqslant i \leqslant s)$也是该向量组的线性组合.

例5.1. 设$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\alpha = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \in \mathbb{F}^3$, 则

$$\alpha = 3e_1 + 4e_2 + 5e_3,$$

即α是e_1, e_2, e_3的线性组合.

线性组合

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s$$

用矩阵写就是

$$\beta = (\alpha_1, \alpha_2, \ldots, \alpha_s) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_s \end{pmatrix}.$$

于是我们得到

命题5.1. 设$\alpha_1, \alpha_2, \ldots, \alpha_s, \beta \in \mathbb{F}^n$, 记$A = (\alpha_1, \alpha_2, \ldots, \alpha_s) \in \mathbb{F}^{n \times s}$, 则$\beta$是$\alpha_1, \alpha_2, \ldots, \alpha_s$的线性组合当且仅当线性方程组$AX = \beta$有解, 且任意解都是组合系数.

例5.2. 设$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, 问$\beta = \begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix}$是否可由$\alpha_1, \alpha_2$线性表示?

☞ **解.** 等价于线性方程组$(\alpha_1,\alpha_2)X=\beta$是否有解? 由于

$$(\alpha_1,\alpha_2,\beta)=\begin{pmatrix}1&4&9\\2&5&12\\3&6&15\end{pmatrix}\underset{r_3-3r_1}{\overset{r_2-2r_1}{\sim}}\begin{pmatrix}1&4&9\\0&-3&-6\\0&-6&-12\end{pmatrix}$$
$$\underset{-\frac{1}{3}r_2}{\overset{r_3-2r_2}{\sim}}\begin{pmatrix}1&4&9\\0&1&2\\0&0&0\end{pmatrix}\overset{r_1-4r_2}{\sim}\begin{pmatrix}1&0&1\\0&1&2\\0&0&0\end{pmatrix},$$

所以方程组有唯一解$X=\begin{pmatrix}1\\2\end{pmatrix}$, 则$\beta=\alpha_2+2\alpha_2$可由$\alpha_1,\alpha_2$线性表示. □

命题5.1本质上给出了线性方程组有解的几何解释. 线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\\qquad\qquad\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m,\end{cases}$$

有矩阵形式$AX=\beta$, 其中系数矩阵$A=(a_{ij})_{m\times n}$, 未知数向量$X=(x_1,x_2,\ldots,x_n)^T$, 常数项列向量$\beta=(b_1,b_2,\ldots,b_m)^T$. 记$A$的列分块为$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)$, 其中

$$\alpha_j=\begin{pmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{pmatrix}.$$

则线性方程组$AX=\beta$有向量形式

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=\beta.$$

由此看出, 线性方程组$AX=\beta$有解当且仅当β是$\alpha_1,\alpha_2,\ldots,\alpha_n$的线性组合. 这就是命题5.1. 例如, 线性方程组$\begin{cases}x+2y=1,\\2x+4y=3\end{cases}$ 的向量形式是

$$x\begin{pmatrix}1\\2\end{pmatrix}+y\begin{pmatrix}2\\4\end{pmatrix}=\begin{pmatrix}1\\3\end{pmatrix}.$$

记$\alpha_1=(1,2)^T$, $\alpha_2=(2,4)^T$, $\beta=(1,3)^T$, 从几何上看, α_1和α_2所有的线性组合是它们所在的直线, 而β不在这条直线上(见图5.3), 所以这个线性方程组无解.

习题5.2

A1. 将向量β写成向量组$\alpha_1,\alpha_2,\alpha_3,\alpha_4$的线性组合, 其中

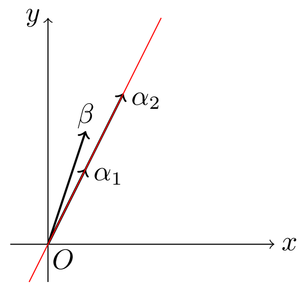


图 5.3: 线性方程组的几何解释

(1) $\beta=(0,0,0,1)^T$, $\alpha_1=(1,1,0,1)^T$, $\alpha_2=(2,1,3,1)^T$, $\alpha_3=(1,1,0,0)^T$, $\alpha_4=(0,1,-1,-1)^T$;
(2) $\beta=(1,2,1,1)$, $\alpha_1=(1,1,1,1)$, $\alpha_2=(1,1,-1,-1)$, $\alpha_3=(1,-1,1,-1)$, $\alpha_4=(1,-1,-1,1)$.

A2. 把线性方程组$\begin{cases}3x_1+5x_2-x_3=2,\\ x_1+7x_2-x_3=0,\\ 4x_1-6x_2=1\end{cases}$ 改写成向量形式.

A3. 如果一个线性方程组的向量形式是$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta$, 其中$\alpha_1=(1,2,6)^T$, $\alpha_2=(3,-1,7)^T$, $\alpha_3=(-1,2,5)^T$, $\beta=(1,0,3)^T$, 试写出该线性方程组.

§ 5.3 向量组的线性相关性

5.3.1 线性相关的定义

要数线性方程组中方程的真正个数, 我们要将可以用其它方程线性表示的方程去掉, 当方程组中所有的方程都不能由其它的方程线性表示时, 方程组中的方程之间没有关系, 都是真正需要的方程. 所以我们数方程的个数时, 要将原来有关系的线性方程组中多余的方程去掉, 直到得到没有关系的同解的线性方程组. 将有关系和没关系严格化, 就是下面向量组的线性相关和线性无关的概念.

定义5.3. 设$s\geqslant 2$, $\alpha_1,\alpha_2,\ldots,\alpha_s\in\mathbb{F}^n$, 如果存在$\alpha_i(1\leqslant i\leqslant s)$, 使得$\alpha_i$可由$\alpha_1$, $\ldots$, α_{i-1}, α_{i+1}, $\ldots$, α_s线性表示, 则称向量组$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关. 否则, 如果$\alpha_1,\alpha_2,\ldots,\alpha_s$中任意一个向量都不能由其余向量线性表示, 则称向量组$\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关.

设$\alpha,\beta,\gamma\in\mathbb{F}^n$, 则

$$\alpha,\beta\text{ 线性相关}\iff \alpha=a\beta(\exists a\in\mathbb{F}),\text{ 或者 }\beta=b\alpha(\exists b\in\mathbb{F}),$$

几何上即α和β共线; 而

$$\begin{aligned}\alpha,\beta,\gamma\text{ 线性相关}\iff& \alpha=a_1\beta+b_1\gamma(\exists a_1,b_1\in\mathbb{F}),\text{ 或者 }\beta=a_2\alpha+b_2\gamma(\exists a_2,b_2\in\mathbb{F}),\\ &\text{或者 }\gamma=a_3\alpha+b_3\beta(\exists a_3,b_3\in\mathbb{F}),\end{aligned}$$

几何上即α,β,γ共面. 所以线性相关是几何上共线和共面概念的推广.

我们看到, 要说明线性相关, 我们需要强调某个向量. 例如, 考察向量组α, β, 其中$\beta = 2\alpha$. 可以利用$\beta = 2\alpha$得出这个向量组线性相关, 也可以利用$\alpha = \frac{1}{2}\beta$得到同样的结论. 强调某个向量会对另一个向量不公平, 不公平的原因是一个在左边另一个在右边. 消除这种不公平性, 可以将上面的关系改写为

$$2\alpha - \beta = 0,$$

此时两个向量都在左边, 它们非平凡地产生了零向量, 这就是一种关系. 一般地, 为了公平和方便, 我们更常用这种等价定义(不再强调某个向量).

命题5.2. 设$\alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}^n$, 其中$s \geqslant 2$. 则

(1) 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关的充分必要条件是存在$c_1, c_2, \ldots, c_s \in \mathbb{F}$不全为零, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0,$$

即0为$\alpha_1, \alpha_2, \ldots, \alpha_s$的非平凡线性组合;

(2) 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性无关的充分必要条件是对任意$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 如果

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0,$$

那么必有$c_1 = c_2 = \cdots = c_s = 0$, 即0只能为$\alpha_1, \alpha_2, \ldots, \alpha_s$的平凡线性组合.

☞ **证明.** (1) 设$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关, 则存在$\alpha_i$ $(1 \leqslant i \leqslant s)$, 使得

$$\alpha_i = b_1\alpha_1 + \cdots + b_{i-1}\alpha_{i-1} + b_{i+1}\alpha_{i+1} + \cdots + b_s\alpha_s, \quad (\exists b_j \in \mathbb{F}).$$

所以

$$b_1\alpha_1 + \cdots + b_{i-1}\alpha_{i-1} + (-1)\alpha_i + b_{i+1}\alpha_{i+1} + \cdots + b_s\alpha_s = 0.$$

反之, 设存在不全为零的$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0.$$

设$c_i \neq 0$ $(1 \leqslant i \leqslant s)$, 则

$$\alpha_i = \frac{-c_1}{c_i}\alpha_1 + \cdots + \frac{-c_{i-1}}{c_i}\alpha_{i-1} + \frac{-c_{i+1}}{c_i}\alpha_{i+1} + \cdots + \frac{-c_s}{c_i}\alpha_s.$$

所以$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关.
(2) 是(1)的逆否命题. □

这个等价定义可以用来定义含一个向量的向量组的线性相关性. 设$\alpha \in \mathbb{F}^n$, 则

$$\alpha\text{ 线性相关} \Longleftrightarrow \alpha = 0.$$

由于

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0 \Longleftrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_s)\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_s \end{pmatrix} = 0,$$

所以又得到可以用齐次线性方程组有无非零解来判别线性相关性, 这也给出了齐次线性方程组有无非零解的一种几何解释.

命题5.3. 设$\alpha_1,\alpha_2,\dots,\alpha_s\in\mathbb{F}^n$, $A=(\alpha_1,\alpha_2,\dots,\alpha_s)\in\mathbb{F}^{n\times s}$, 则

(1) $\alpha_1,\alpha_2,\dots,\alpha_s$线性相关$\Longleftrightarrow$ 齐次线性方程组$AX=0$有非零解;

(2) $\alpha_1,\alpha_2,\dots,\alpha_s$线性无关$\Longleftrightarrow$ 齐次线性方程组$AX=0$只有零解.

结合推论2.5和推论4.13, 我们得到

推论5.4. 设$\alpha_1,\alpha_2,\dots,\alpha_s\in\mathbb{F}^n$, $A=(\alpha_1,\alpha_2,\dots,\alpha_s)\in\mathbb{F}^{n\times s}$, 则

(1) 当$s>n$(即: 向量个数>向量维数)时, $\alpha_1,\alpha_2,\dots,\alpha_s$线性相关.

(2) 当$s=n$时, 有

$$\alpha_1,\alpha_2,\dots,\alpha_s\text{ 线性相关}\Longleftrightarrow A\text{ 不可逆}\Longleftrightarrow |A|=0,$$
$$\alpha_1,\alpha_2,\dots,\alpha_s\text{ 线性无关}\Longleftrightarrow A\text{ 可逆}\Longleftrightarrow |A|\neq 0.$$

上面推论中的(1)说当向量比较多时应该有关系, 这也符合我们的直观感觉. 而(2)说明当向量组中向量的个数恰为向量的维数时, 可以用对应的方阵的行列式是否为零来判别有无关系, 这看上去挺神奇的.

5.3.2 例子

我们看一些例子.

例5.3. 设$\alpha_1=\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\alpha_2=\begin{pmatrix}2\\-1\\1\end{pmatrix}$, $\alpha_3=\begin{pmatrix}3\\0\\1\end{pmatrix}$, $\alpha_4=\begin{pmatrix}1\\2\\-1\end{pmatrix}$, 问它们是否线性相关?

☞ **解.** 由于向量个数为4, 大于向量的维数3, 所以线性相关(事实上, $\alpha_3=\alpha_1+\alpha_2$). □

例5.4. 设$\alpha_1=\begin{pmatrix}1\\2\\-1\\3\end{pmatrix}$, $\alpha_2=\begin{pmatrix}0\\4\\-1\\3\end{pmatrix}$, $\alpha_3=\begin{pmatrix}0\\0\\5\\4\end{pmatrix}$, $\alpha_4=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$, 问它们是否线性相关?

☞ **解.** 由于行列式

$$|\alpha_1,\alpha_2,\alpha_3,\alpha_4|=\begin{vmatrix}1&0&0&0\\2&4&0&0\\-1&-1&5&0\\3&3&4&1\end{vmatrix}=20\neq 0,$$

所以线性无关. □

例5.5. 设$\alpha_1=\begin{pmatrix}1\\2\\-1\\3\end{pmatrix}$, $\alpha_2=\begin{pmatrix}-1\\2\\-1\\4\end{pmatrix}$, $\alpha_3=\begin{pmatrix}0\\4\\3\\1\end{pmatrix}$, 问它们是否线性相关?

解. 由于

$$(\alpha_1,\alpha_2,\alpha_3)=\begin{pmatrix}1&-1&0\\2&2&4\\-1&-1&3\\3&4&1\end{pmatrix}\underset{r_4-3r_1}{\underbrace{\overset{r_2-2r_1}{r_3+r_1}}}\begin{pmatrix}1&-1&0\\0&4&4\\0&-2&3\\0&7&1\end{pmatrix}\underbrace{\tfrac{1}{4}r_2}\begin{pmatrix}1&-1&0\\0&1&1\\0&-2&3\\0&7&1\end{pmatrix}$$

$$\underset{r_4-7r_2}{\underbrace{r_3+2r_2}}\begin{pmatrix}1&-1&0\\0&1&1\\0&0&5\\0&0&-6\end{pmatrix}\underbrace{r_4+\tfrac{6}{5}r_3}\begin{pmatrix}1&-1&0\\0&1&1\\0&0&5\\0&0&0\end{pmatrix},$$

所以齐次线性方程组$(\alpha_1,\alpha_2,\alpha_3)X=0$只有零解, 线性无关. □

例5.6. 设$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, $e_2=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}$, …, $e_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}\in\mathbb{F}^n$, 证明: $e_1,e_2,\ldots,e_n$线性无关.

证明. 由于$(e_1,e_2,\ldots,e_n)=E_n$可逆, 所以$e_1,e_2,\ldots,e_n$线性无关. □

称$e_1,e_2,\ldots,e_n$为单位坐标向量. 当$n=3$时, 可以看成通常的三维空间中直角坐标系的三个坐标方向向量, 它们当然不共面, 没有关系.

例5.7. 设$\alpha_1,\alpha_2,\alpha_3\in\mathbb{F}^n$, 证明:

(1) $\alpha_1-\alpha_2$, $\alpha_2-\alpha_3$, $\alpha_3-\alpha_1$线性相关;

(2) $\alpha_1,\alpha_2,\alpha_3$线性无关$\Longleftrightarrow$ $\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$, $\alpha_3+\alpha_1$线性无关.

证明. 我们给出几何(定义)证明, 后面会给出代数证明.

(1) 由于

$$\alpha_3-\alpha_1=-(\alpha_1-\alpha_2)-(\alpha_2-\alpha_3),$$

所以线性相关.

(2) 设$\alpha_1,\alpha_2,\alpha_3$线性无关. 如果

$$c_1(\alpha_1+\alpha_2)+c_2(\alpha_2+\alpha_3)+c_3(\alpha_3+\alpha_1)=0,\quad (c_1,c_2,c_3\in\mathbb{F}),$$

则有

$$(c_1+c_3)\alpha_1+(c_1+c_2)\alpha_2+(c_2+c_3)\alpha_3=0.$$

因为$\alpha_1,\alpha_2,\alpha_3$线性无关, 所以得到

$$\begin{cases}c_1+c_3=0,\\c_1+c_2=0,\\c_2+c_3=0.\end{cases}$$

三个方程相加得到$c_1+c_2+c_3=0$, 进而可得$c_1=c_2=c_3=0$ (也可以算出系数矩阵的行列式为2, 得只有零解). 于是$\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$, $\alpha_3+\alpha_1$线性无关.

反之, 设$\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$, $\alpha_3+\alpha_1$线性无关. 如果

$$b_1\alpha_1+b_2\alpha_2+b_3\alpha_3=0,\quad (b_1,b_2,b_3\in\mathbb{F}),$$

则[3]

$$\begin{aligned}&(b_1+b_2-b_3)(\alpha_1+\alpha_2)+(-b_1+b_2+b_3)(\alpha_2+\alpha_3)+(b_1-b_2+b_3)(\alpha_3+\alpha_1)\\=&2b_1\alpha_1+2b_2\alpha_2+2b_3\alpha_3=0.\end{aligned}$$

所以有

$$\begin{cases}b_1+b_2-b_3=0,\\-b_1+b_2+b_3=0,\\b_1-b_2+b_3=0.\end{cases}$$

三个方程相加得到$b_1+b_2+b_3=0$, 进而可得$b_1=b_2=b_3=0$ (也可以算出系数矩阵的行列式为4, 得只有零解). 于是α_1, α_2, α_3线性无关. □

5.3.3 性质

下面我们介绍后面需要用到的一些性质, 从逻辑上看应该是在考虑我们想做的问题时发现要研究这些性质是否成立才去研究它们的, 但为了不让我们的陈述显得杂乱, 所以我们先讲述这些性质, 请读者相信它们是必要的.

首先, 如果原来一个向量组已经有关系了, 再加一些向量上去, 原来的关系当然还在. 也就是下面结论应该成立.

命题5.5. **(1)** *设$\alpha_1,\alpha_2,\ldots,\alpha_s\in\mathbb{F}^n$线性相关, 则对任意$\alpha_{s+1}$, $\ldots$, $\alpha_r\in\mathbb{F}^n$, 有α_1, $\ldots$, α_s, α_{s+1}, $\ldots$, α_r也线性相关;*

(2) *设$\alpha_1,\alpha_2,\ldots,\alpha_s\in\mathbb{F}^n$线性无关, 则$\alpha_1,\alpha_2,\ldots,\alpha_s$的任一子向量组也线性无关.*

☞ **证明.** (1) 由于$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关, 所以存在不全为零的$c_1,c_2,\ldots,c_s\in\mathbb{F}$, 使得

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s=0.$$

[3]如何想到这个等式? 事实上, 由

$$(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix},$$

求上式右边矩阵的逆后可得

$$(\alpha_1,\alpha_2,\alpha_3)=\frac{1}{2}(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1)\begin{pmatrix}1&1&-1\\-1&1&1\\1&-1&1\end{pmatrix}.$$

再代入

$$(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}=0$$

即可.

于是得到

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s+0\cdot\alpha_{s+1}+\cdots+0\cdot\alpha_r=0,$$

即得$\alpha_1,\ldots,\alpha_s,\alpha_{s+1},\ldots,\alpha_r$线性相关. □

由于0线性相关, 所以含有零向量的向量组必线性相关; 由于α,α线性相关, 所以含有重复向量的向量组也线性相关. 于是, 如果一个向量组线性无关, 则该向量组必无零向量, 且该向量组的向量必两两不同.

其次, 线性无关和线性组合还有下面的关系.

命题5.6. 设$\alpha_1,\alpha_2,\ldots,\alpha_s\in\mathbb{F}^n$.

(1) 向量组$\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关的充分必要条件是, 对任意$\beta\in\mathbb{F}^n$, 若$\beta$可由$\alpha_1,\alpha_2,\ldots,\alpha_s$线性表示, 则表示方式唯一;

(2) 设$\beta\in\mathbb{F}^n$, 如果$\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关, 而$\alpha_1,\alpha_2,\ldots,\alpha_s,\beta$线性相关, 则$\beta$可由$\alpha_1,\alpha_2,\ldots,\alpha_s$唯一地线性表示.

☞ **证明.** (1) 设$\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关. 如果

$$\beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_s\alpha_s=c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s,\quad(b_i,c_j\in\mathbb{F}),$$

则有

$$(b_1-c_1)\alpha_1+(b_2-c_2)\alpha_2+\cdots+(b_s-c_s)\alpha_s=0.$$

由于$\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关, 所以

$$b_1-c_1=b_2-c_2=\cdots=b_s-c_s=0,$$

即得$b_1=c_1,\ b_2=c_2,\ \ldots,\ b_s=c_s$.

反之, 设对任意$\beta\in\mathbb{F}^n$, 若β可由$\alpha_1,\alpha_2,\ldots,\alpha_s$线性表示, 则表示方式唯一. 如果$s=1$, 而$\alpha_1$线性相关, 则$\alpha_1=0$. 于是

$$\alpha_1=\alpha_1=0\times\alpha_1.$$

则α_1由α_1线性表示的方式不唯一, 矛盾. 如果$s\geqslant 2$, 设$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关, 则存在$i$, 使得

$$\alpha_i=c_1\alpha_1+\cdots+c_{i-1}\alpha_{i-1}+c_{i+1}\alpha_{i+1}+\cdots+c_s\alpha_s,\quad\exists c_1,\ldots,c_s\in\mathbb{F}.$$

而$\alpha_i=\alpha_i$, 所以$\alpha_1=\alpha_2=\cdots=\alpha_s=\alpha_i$. 此时,

$$1\times\alpha_1=2\times\alpha_1+(-1)\times\alpha_2,$$

线性表示的方式不唯一, 矛盾.

(2) 由于$\alpha_1,\alpha_2,\ldots,\alpha_s,\beta$线性相关, 所以存在不全为零的$a_1,a_2,\ldots,a_s,a\in\mathbb{F}$, 使得

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_s\alpha_s+a\beta=0.$$

若$a=0$, 则$a_1,a_2,\ldots,a_s$不全为零, 且

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_s\alpha_s=0.$$

这得到$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关, 矛盾. 所以$a\neq 0$. 进而

$$\beta=\frac{-a_1}{a}\alpha_1+\frac{-a_2}{a}\alpha_2+\cdots+\frac{-a_s}{a}\alpha_s,$$

即β可由$\alpha_1,\alpha_2,\ldots,\alpha_s$线性表示. 表示的唯一性由(1). □

最后, 如果将一个向量组的每个向量都取出同样的几个分量, 得到维数更小的一个向量组, 是否改变线性无关性? 用线性方程组来看, 这相当于删去几个未知数. 如果原来方程组中有某个方程多余, 则在删去了几个未知数的新的方程组中对应的方程也多余. 即如果原来的方程组线性相关, 则新的方程组也线性相关.

命题5.7. 设

$$\alpha_1=\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\alpha_2=\begin{pmatrix}a_{12}\\ \vdots\\ a_{n2}\end{pmatrix},\ldots,\alpha_s=\begin{pmatrix}a_{1s}\\ \vdots\\ a_{ns}\end{pmatrix}\in\mathbb{F}^n$$

是n维列向量组, 对正整数$m<n$和$1\leqslant i_1<i_2<\cdots<i_m\leqslant n$, 定义$m$维列向量组

$$\beta_1=\begin{pmatrix}a_{i_11}\\ a_{i_21}\\ \vdots\\ a_{i_m1}\end{pmatrix},\beta_2=\begin{pmatrix}a_{i_12}\\ a_{i_22}\\ \vdots\\ a_{i_m2}\end{pmatrix},\ldots,\beta_s=\begin{pmatrix}a_{i_1s}\\ a_{i_2s}\\ \vdots\\ a_{i_ms}\end{pmatrix}\in\mathbb{F}^m.$$

称$\beta_1,\beta_2,\ldots,\beta_s$为向量组$\alpha_1,\alpha_2,\ldots,\alpha_s$的$m$维缩短向量组. 则

(1) $\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关$\Longrightarrow$ $\beta_1,\beta_2,\ldots,\beta_s$线性相关;

(2) $\beta_1,\beta_2,\ldots,\beta_s$线性无关$\Longrightarrow$ $\alpha_1,\alpha_2,\ldots,\alpha_s$线性无关.

☞ **证明.** (1) 简单起见, 不妨设$i_1=1$, $i_2=2$, ..., $i_m=m$. 记$\gamma_j=\begin{pmatrix}a_{m+1,j}\\ \vdots\\ a_{nj}\end{pmatrix}$, 则$\alpha_j=\begin{pmatrix}\beta_j\\ \gamma_j\end{pmatrix}$, $j=1,2,\ldots,s$. 由于$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关, 所以存在不全为零的$c_1,c_2,\ldots,c_s\in\mathbb{F}$, 使得

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s=0.$$

则有

$$c_1\begin{pmatrix}\beta_1\\ \gamma_1\end{pmatrix}+c_2\begin{pmatrix}\beta_2\\ \gamma_2\end{pmatrix}+\cdots+c_s\begin{pmatrix}\beta_s\\ \gamma_s\end{pmatrix}=0,$$

这得到

$$c_1\beta_1+c_2\beta_2+\cdots+c_s\beta_s=0.$$

于是$\beta_1,\beta_2,\ldots,\beta_s$线性相关. □

习题5.3

A1. 举例说明下列各命题是错误的.

(1) 如果向量组$\alpha_1,\alpha_2,\ldots,\alpha_s$线性相关, 则$\alpha_1$可由$\alpha_2,\ldots,\alpha_s$线性表示;

(2) 如果α_1,α_2线性无关, α_2,α_3线性无关, α_3,α_1线性无关, 则$\alpha_1,\alpha_2,\alpha_3$也线性无关;

(3) 如果有不全为零的数$c_1, c_2, \ldots, c_s$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s + c_1\beta_1 + c_2\beta_2 + \cdots + c_s\beta_s = 0$$

成立, 则向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关, 向量组$\beta_1, \beta_2, \ldots, \beta_s$也线性相关;

(4) 如果只有当$c_1, c_2, \ldots, c_s$全为零时, 等式

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s + c_1\beta_1 + c_2\beta_2 + \cdots + c_s\beta_s = 0$$

才能成立, 则向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性无关, 向量组$\beta_1, \beta_2, \ldots, \beta_s$也线性无关.

A2. 判别下列向量组是否线性相关:

(1) $\alpha_1 = (1, 2, 3)^T$, $\alpha_2 = (2, 3, 4)^T$;

(2) $\alpha_1 = (1, 2, 3)^T$, $\alpha_2 = (2, -1, 1)^T$, $\alpha_3 = (-1, 3, 2)^T$;

(3) $\alpha_1 = (1, 2, 3, 4)^T$, $\alpha_2 = (2, 0, 1, 2)^T$, $\alpha_3 = (3, -2, -3, -1)^T$.

A3. 设$\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}^n$, 证明:

(1) 向量组$\alpha_1 - \alpha_2$, $\alpha_2 + \alpha_3$, $\alpha_3 + \alpha_1$线性相关;

(2) 如果$\alpha_1, \alpha_2, \alpha_3$线性无关, 则向量组$\alpha_1 + 2\alpha_2$, $\alpha_2 + 2\alpha_3$, $\alpha_3 + 2\alpha_1$也线性无关.

A4. 设向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性无关, 其中$s \geqslant 2$, 证明:

(1) 向量组$\alpha_1 - \alpha_2$, $\alpha_2 - \alpha_3$, $\ldots$, $\alpha_{s-1} - \alpha_s$也线性无关;

(2) 向量组$\alpha_1 - \alpha_2$, $\alpha_1 - \alpha_3$, $\ldots$, $\alpha_1 - \alpha_s$也线性无关.

A5. 设向量β可由向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性表示, 但是不能由该向量组的任意一个个数少于$s$的子向量组线性表示, 证明: 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性无关.

A6. 设A是n阶方阵, $\alpha_1, \alpha_2, \ldots, \alpha_s$是$n$维列向量组.

(1) 证明: 如果$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关, 则$A\alpha_1, A\alpha_2, \ldots, A\alpha_s$也线性相关;

(2) 证明: 如果A可逆, 则(1)的逆命题也成立;

(3) 如果A不可逆, (1)的逆命题是否还成立? 成立时给出证明, 不成立时给出反例;

(4) 如果A不可逆, 是否$A\alpha_1, A\alpha_2, \ldots, A\alpha_s$一定线性相关? 肯定时给出证明, 否定时给出反例.

B1. 设$s \geqslant 2$, 证明: 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$(其中$\alpha_1 \neq 0$)线性相关的充分必要条件是存在$i(1 < i \leqslant s)$, 使得$\alpha_i$可由$\alpha_1, \alpha_2, \ldots, \alpha_{i-1}$线性表示.

B2. 设$a_1, a_2, \ldots, a_s \in \mathbb{F}$是两两不同的数, 取$n \geqslant s$, 令$\alpha_i = (1, a_i, a_i^2, \ldots, a_i^{n-1})^T \in \mathbb{F}^n$, $i = 1, 2, \ldots, s$, 证明: 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性无关.

B3. 设$\mathbb{F}^n$中的向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$线性相关, 其中$s \geqslant 2$, 设$\beta \in \mathbb{F}^n$, 证明: 存在不全为零的数$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 使得向量组$\alpha_1 + c_1\beta$, $\alpha_2 + c_2\beta$, $\ldots$, $\alpha_s + c_s\beta$线性相关.

§ 5.4 向量组间的等价

5.4.1 定义

在删去线性方程组中假的方程时, 不同的人可能删去的方程会不一样. 例如, 考察线性方程组

$$\begin{cases} x+y+z=1, & (1) \\ 2x+y-2z=3, & (2) \\ 3x+2y-z=4. & (3) \end{cases}$$

如果注意到方程(3)是方程(1)和(2)的和, 则可以将(3)去掉, 得到同解方程组

$$(A) \begin{cases} x+y+z=1, & (1) \\ 2x+y-2z=3. & (2) \end{cases}$$

也可以注意到方程(2)是(3)和(1)之差, 所以可以去掉(2), 得到

$$(B) \begin{cases} x+y+z=1, & (1) \\ 3x+2y-z=4. & (3) \end{cases}$$

还可以注意到方程(1)是(3)和(2)之差, 所以可以去掉(1), 得到

$$(C) \begin{cases} 2x+y-2z=3, & (2) \\ 3x+2y-z=4. & (3) \end{cases}$$

这里的方程组$(A),(B),(C)$都可以代表原方程组, 原来方程组的每个方程都可以分别由它们产生, 即是它们的线性组合. 特别地, $(A),(B),(C)$中任两个方程组中的方程可以相互线性表示, 我们称它们是等价的.

一般地, 将给定向量组的多余向量删去时, 不同的人最后删剩下的子向量组可能不同, 但这些不同组之间的向量可以相互线性表示. 我们把这些严格化, 引入下面的定义.

定义5.4. 设有$\mathbb{F}^n$中的向量组

$$S:\alpha_1,\alpha_2,\dots,\alpha_s, \quad T:\beta_1,\beta_2,\dots,\beta_t.$$

(1) 如果任意的$\beta_j(1\leqslant j\leqslant t)$都可由$S$线性表示, 则称$T$可由$S$线性表示, 记为$T\leftarrow S$, 也记为$S\to T$.

(2) 如果S和T可以相互线性表示, 即$S\leftarrow T$且$T\leftarrow S$, 则称S与T等价, 记为$S\leftrightarrow T$.

和前面一样, 我们看矩阵表示. 我们有

$$T\leftarrow S \iff \begin{cases} \beta_1=c_{11}\alpha_1+c_{21}\alpha_2+\cdots+c_{s1}\alpha_s, \\ \beta_2=c_{12}\alpha_1+c_{22}\alpha_2+\cdots+c_{s2}\alpha_s, \\ \qquad\vdots \\ \beta_t=c_{1t}\alpha_1+c_{2t}\alpha_2+\cdots+c_{st}\alpha_s, \end{cases} \qquad (\exists c_{ij}\in\mathbb{F})$$

$$\iff (\beta_1,\beta_2,\dots,\beta_t)=(\alpha_1,\alpha_2,\dots,\alpha_s)\begin{pmatrix}c_{11} & c_{12} & \cdots & c_{1t}\\ c_{21} & c_{22} & \cdots & c_{2t}\\ \vdots & \vdots & & \vdots\\ c_{s1} & c_{s2} & \cdots & c_{st}\end{pmatrix},\quad (\exists c_{ij}\in\mathbb{F}).$$

这得到下面的等价刻画, 也给出矩阵方程$AX=B$有解的一种几何解释.

命题5.8. *设有$\mathbb{F}^n$中的向量组$S:\alpha_1,\alpha_2,\dots,\alpha_s$和$T:\beta_1,\beta_2,\dots,\beta_t$, 记矩阵$A=(\alpha_1,\alpha_2,\dots,\alpha_s)\in\mathbb{F}^{n\times s}$, $B=(\beta_1,\beta_2,\dots,\beta_t)\in\mathbb{F}^{n\times t}$, 则下面命题等价*

(i) *$T\leftarrow S$;*

(ii) *存在矩阵$C\in\mathbb{F}^{s\times t}$, 使得$B=AC$;*

(iii) *矩阵方程$AX=B$有解.*

有下面重要的例子.

例5.8. *设$A\in\mathbb{F}^{n\times s}$,$C\in\mathbb{F}^{s\times t}$, 记$B=AC$, 则有*

(1) *(B 的列向量组) $\leftarrow$ (A 的列向量组);*

(2) *(B 的行向量组) $\leftarrow$ (C 的行向量组).*

☞ **证明.** 用命题5.8可得(1), 对$B^T=C^TA^T$用(1)可得(2). □

• **矩阵等价与向量组等价**

我们学过两种等价, 一种是矩阵的(行, 列)等价, 另一种是向量组的等价. 为什么用一样的名字? 它们有什么关系? 具体的, 每个矩阵都带有两个向量组: 它的行向量组和列向量组, 那么两个矩阵等价与其对应的向量组等价之间有什么关系呢? 有了思考的方向, 那么离结论就不远了.

命题5.9. *设$A,B\in\mathbb{F}^{m\times n}$, 则*

(1) *$A\overset{c}{\sim}B \iff A$和$B$的列向量组等价;*

(2) *$A\overset{r}{\sim}B \iff A$和$B$的行向量组等价;*

(3) *如果$A\sim B$, 则A和B的(行)列向量组不一定等价.*

☞ **证明.** (1) 设$A\overset{c}{\sim}B$, 则存在n阶可逆阵Q, 使得$AQ=B$. 由此得

$$(B\text{ 的列向量组})\leftarrow(A\text{ 的列向量组}).$$

又有$BQ^{-1}=A$, 所以

$$(A\text{ 的列向量组})\leftarrow(B\text{ 的列向量组}).$$

就得到A和B的列向量组等价.

我们将在下节中证明充分性.

(2) 对A^T和B^T用(1).

(3) 例如, 取$A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$, 则由

$$A\overset{r_1\leftrightarrow r_2}{\rightsquigarrow}\begin{pmatrix}0&0\\1&0\end{pmatrix}\overset{c_1\leftrightarrow c_2}{\rightsquigarrow}B$$

得$A\sim B$. 而A的列向量组$\begin{pmatrix}1\\0\end{pmatrix}$, $\begin{pmatrix}0\\0\end{pmatrix}$与$B$的列向量组$\begin{pmatrix}0\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\end{pmatrix}$不等价, A的行向量组$(1,0)$, $(0,0)$与B的行向量组$(0,0)$, $(0,1)$也不等价. □

我们先证明在一定条件下命题5.9(1)的充分性成立.

例5.9. 设$A,B\in\mathbb{F}^{m\times n}$, 如果它们的列向量组都线性无关, 则

$$A\text{和}B\text{列向量组等价}\Longrightarrow A\overset{c}{\sim}B.$$

☞ **证明.** 由条件, 存在n阶方阵P和Q, 使得

$$A=BP,\quad B=AQ.$$

这得到

$$B=(BP)Q=B(PQ),$$

即

$$B(PQ-E_n)=O.$$

但是B的列向量组线性无关导出线性方程组$BX=0$只有零解, 这得到$PQ=E_n$. 所以P是可逆方阵, 于是$A\overset{c}{\sim}B$. □

5.4.2 性质

同样地, 我们给出一些后面要用到的性质, 它们在我们的数数问题中是必要的.

首先, 向量组之间的表示关系有自然的传递性.

命题5.10. **(1)** (表示的传递性) 设有$\mathbb{F}^n$中的有限向量组S_1,S_2和S_3, 则

$$S_1\leftarrow S_2, S_2\leftarrow S_3\Longrightarrow S_1\leftarrow S_3;$$

(2) 设S,T,R为$\mathbb{F}^n$中任意的有限向量组, 则有[4]

(i) (自反性) $S\leftrightarrow S$;

(ii) (对称性) $S\leftrightarrow T\Longrightarrow T\leftrightarrow S$;

(iii) (传递性) $S\leftrightarrow T$, $T\leftrightarrow R\Longrightarrow S\leftrightarrow R$.

☞ **证明.** (1) 设$S_1:\alpha_1,\alpha_2,\dots,\alpha_r$, $S_2:\beta_1,\beta_2,\dots,\beta_s$, $S_3:\gamma_1,\gamma_2,\dots,\gamma_t$, 记矩阵

$$A=(\alpha_1,\alpha_2,\dots,\alpha_r)\in\mathbb{F}^{n\times r}, B=(\beta_1,\beta_2,\dots,\beta_s)\in\mathbb{F}^{n\times s}, C=(\gamma_1,\gamma_2,\dots,\gamma_t)\in\mathbb{F}^{n\times t}.$$

由于$S_1\leftarrow S_2$且$S_2\leftarrow S_3$, 所以存在$C_1\in\mathbb{F}^{s\times r}$和$C_2\in\mathbb{F}^{t\times s}$, 使得

$$A=BC_1,\quad B=CC_2.$$

[4]即向量组之间的等价为$\mathbb{F}^n$中的有限向量组所成之集的等价关系.

于是

$$A = (CC_2)C_1 = C(C_2C_1),$$

这得到$S_1 \leftarrow S_3$.

(2) 自反性和对称性由定义, 而传递性从(1)得到. □

其次, 我们知道向量个数比较多的向量组是线性相关的. 于是, 如果一个向量组产生的新的向量组的向量个数很多时, 很有可能这个新的向量组是线性相关的. 这个确实是一般规律.

定理5.11. 设有$\mathbb{F}^n$中的向量组$S: \alpha_1, \alpha_2, \ldots, \alpha_s$和$T: \beta_1, \beta_2, \ldots, \beta_t$, 则

(1) $T \leftarrow S$且$t > s \Longrightarrow T$线性相关;

(2) $T \leftarrow S$且T线性无关$\Longrightarrow t \leqslant s$;

(3) $S \leftrightarrow T$且S与T均线性无关$\Longrightarrow s = t$.

☞ **证明.** (1) 记矩阵$A = (\alpha_1, \alpha_2, \ldots, \alpha_s) \in \mathbb{F}^{n\times s}$, $B = (\beta_1, \beta_2, \ldots, \beta_t) \in \mathbb{F}^{n\times t}$. 因为$T \leftarrow S$, 所以存在矩阵$C \in \mathbb{F}^{s\times t}$, 使得$B = AC$. 而$t > s$, 所以齐次线性方程组$CX = 0$有非零解. 这得到$ACX = 0$有非零解. 于是齐次线性方程组$BX = 0$有非零解, 所以$T$线性相关.

(2) 由(1).

(3) 由(2). □

习题5.4

A1. 设$s \geqslant 2$, $\alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}^n$, 定义$\beta_1 = \alpha_2 + \alpha_3 + \cdots + \alpha_s$, $\beta_2 = \alpha_1 + \alpha_3 + \cdots + \alpha_s$, ..., $\beta_s = \alpha_1 + \alpha_2 + \cdots + \alpha_{s-1}$, 证明: 向量组$\beta_1, \beta_2, \ldots, \beta_s$与向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$等价.

A2. 设$\alpha_1, \alpha_2, \ldots, \alpha_s$是$\mathbb{F}^n$中的向量组(其中$s \leqslant n$), 如果向量组$\beta_1, \beta_2, \ldots, \beta_s$线性无关且可以由$\alpha_1, \alpha_2, \ldots, \alpha_s$线性表示, 证明: 向量组$\alpha_1, \alpha_2, \ldots, \alpha_s$和向量组$\beta_1, \beta_2, \ldots, \beta_s$等价.

B1. (**Steinitz替换定理**) 设向量组$\alpha_1, \alpha_2, \ldots, \alpha_r$是线性无关的, 且可由向量组$\beta_1$, β_2, ..., β_s线性表示(于是, 有$r \leqslant s$), 证明: 可用$\alpha_1, \alpha_2, \ldots, \alpha_r$替换$\beta_1, \beta_2, \ldots, \beta_s$中的$r$个向量, 不妨设替换了$\beta_1, \beta_2, \ldots, \beta_r$, 使得向量组$\alpha_1, \ldots, \alpha_r, \beta_{r+1}, \ldots, \beta_s$和向量组$\beta_1, \beta_2, \ldots, \beta_s$等价.

§ 5.5 向量组的秩

我们要删去线性方程组中多余的方程, 即对线性方程组进行打假, 最后剩下的方程组中的方程没有多余的, 即它是线性无关的; 另一方面, 最后留下的方程要能代表原方程组, 即原方程组的每个方程都可以由剩下的方程组线性表示. 我们称进行完全打假后剩下的方程所组成的方程组为原方程组的一个极大无关组. 考察线性方程组, 本质上只要考察它的任意一个极大无关组就好了. 不同的人进行打假后得到的极大无关组可能不同, 但是每个极大无关组所含方程的个数是相同的(可由定理5.11(3)导出). 这个相同的数称为原方程组的秩, 这就是原方程组真正方程的个数, 数线性方程组中真正方程的个数就是要得到它的秩.

下面我们将讨论一般化.

5.5.1 极大无关组

引理5.12. 设S为$\mathbb{F}^n$中的向量组(所含向量个数可能无穷), $\alpha_1, \alpha_2, \ldots, \alpha_r \in S$满足

(i) $\alpha_1, \alpha_2, \ldots, \alpha_r$线性无关.

则下面等价

(ii) 任意$\alpha \in S$, α可由$\alpha_1, \alpha_2, \ldots, \alpha_r$线性表示.

(ii)′ S中任意$r+1$个向量线性相关.

(ii)″ 任意$\alpha \in S$, $\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha$线性相关.

定义5.5. 设S为$\mathbb{F}^n$中的向量组(所含向量个数可能无穷), α_1, α_2, …, $\alpha_r \in S$, 如果向量组α_1, α_2, …, α_r满足(i)和(ii)($\Longleftrightarrow$ 满足(i)和(ii)′ $\Longleftrightarrow$ 满足(i)和(ii)″), 则称α_1, α_2, …, α_r是S的一个极大无关组(极大线性无关组).

于是, 极大无关组是可表示原向量组[5]的线性无关子组, 可以代表原向量组; 也是所含向量个数最多的线性无关子组; 还是不能再线性无关扩充的无关子组. 从后两个刻画可以看出用"极大"这个定语的原因.

☞ **引理5.12的证明.** "(ii) $\Longrightarrow$ (ii)′" 设(ii)成立, 任取$\beta_1, \beta_2, \ldots, \beta_{r+1} \in S$, 则由(ii)得对$i = 1, 2, \ldots, r+1$, β_i可由$\alpha_1, \alpha_2, \ldots, \alpha_r$线性表示. 因此

$$\beta_1, \beta_2, \ldots, \beta_{r+1} \leftarrow \alpha_1, \alpha_2, \ldots, \alpha_r.$$

由定理5.11(1)就得$\beta_1, \beta_2, \ldots, \beta_{r+1}$线性相关.

"(ii)′ $\Longrightarrow$ (ii)″" 显然.

"(ii)″ $\Longrightarrow$ (ii)" 由命题5.6(2). □

我们看几个例子.

例5.10. 设$\alpha_1 = \begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\alpha_3 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, 证明: α_2, α_3; α_1, α_3和α_1, α_2都是向量组$\alpha_1, \alpha_2, \alpha_3$的极大无关组.

☞ **证明.** 由于

$$|\alpha_1, \alpha_2, \alpha_3| = \begin{vmatrix} 9 & 1 & 4 \\ 12 & 2 & 5 \\ 15 & 3 & 6 \end{vmatrix} \xlongequal[r_3 - 3r_1]{r_2 - 2r_1} \begin{vmatrix} 9 & 1 & 4 \\ -6 & 0 & -3 \\ -12 & 0 & -6 \end{vmatrix} = 0,$$

所以$\alpha_1, \alpha_2, \alpha_3$线性相关(或者由: $\alpha_1 = \alpha_2 + 2\alpha_3$). 而$\alpha_2, \alpha_3$不成比例, 所以$\alpha_2, \alpha_3$线性无关. 这得到$\alpha_2, \alpha_3$是$\alpha_1$, α_2, α_3的极大无关组.

类似可证, α_1, α_3和α_1, α_2均是向量组$\alpha_1, \alpha_2, \alpha_3$的极大无关组. □

例5.11. 证明: $e_1, e_2, \ldots, e_n$是$\mathbb{F}^n$的极大无关组.

[5]即和原向量组等价. 这里可以自然地将线性表示的概念推广到含有无穷个向量的向量组, 参看下一章.

证明. 我们已经证明过$e_1,e_2,\dots,e_n$线性无关, 而任意$n+1$个n维向量必线性相关, 所以$e_1,e_2,\dots,e_n$是$\mathbb{F}^n$的极大无关组. □

由例5.10, 极大无关组不唯一, 但是每个极大无关组所含向量的个数是唯一的.

命题5.13. *设S为$\mathbb{F}^n$中的向量组, $\alpha_1,\alpha_2,\dots,\alpha_r$和$\beta_1,\beta_2,\dots,\beta_s$都是$S$的极大无关组, 则$r=s$.*

证明. 由于$\alpha_1,\alpha_2,\dots,\alpha_r$是极大无关组, 所以$\beta_1,\beta_2,\dots,\beta_s$可由$\alpha_1,\alpha_2,\dots,\alpha_r$线性表示. 类似有, $\alpha_1,\alpha_2,\dots,\alpha_r$可由$\beta_1,\beta_2,\dots,\beta_s$线性表示. 即得

$$\alpha_1,\alpha_2,\dots,\alpha_r \leftrightarrow \beta_1,\beta_2,\dots,\beta_s.$$

而$\alpha_1,\alpha_2,\dots,\alpha_r$和$\beta_1,\beta_2,\dots,\beta_s$都线性无关, 所以定理5.11(3)推出$r=s$. □

最后, 对于$\mathbb{F}^n$中的任意向量组S, S是否一定有极大无关组? 这个问题必须回答, 否则我们可能在讨论一个不一定存在的对象. 我们可以用进行线性无关扩充来讨论这个问题. 如果一个线性无关子组不能再进行线性无关扩充, 则就是极大无关组; 否则可以线性无关扩充, 此时对该更大的线性无关子组继续线性无关扩充, 直到不能再扩充为止. 而这个过程有限步后必将停止, 因为线性无关的向量不能太多! 于是, 我们得到下面的存在性定理, 特别有用的是线性无关扩充定理.

命题5.14. *设S为$\mathbb{F}^n$中的任意向量组.*

(1) *(线性无关扩充定理) 如果$\alpha_1,\alpha_2,\dots,\alpha_s\in S$是线性无关子组, 则存在$\alpha_{s+1}$, $\dots$, $\alpha_r\in S$ $(r\geqslant s)$, 使得$\alpha_1,\dots,\alpha_s,\alpha_{s+1},\dots,\alpha_r$是$S$ 的极大无关组;*

(2) *如果S中有非零向量, 则S有极大无关组.*

证明. (1) 如果对任意的$\alpha\in S$, 有$\alpha_1,\dots,\alpha_s,\alpha$线性相关, 则$\alpha_1,\dots,\alpha_s$为$S$的极大无关组. 否则, 存在$\alpha_{s+1}\in S$, 使得$\alpha_1,\dots,\alpha_s,\alpha_{s+1}$线性无关. 如果对任意的$\alpha\in S$, 有$\alpha_1,\dots,\alpha_s,\alpha_{s+1},\alpha$线性相关, 则$\alpha_1,\dots,\alpha_{s+1}$为$S$的极大无关组. 否则, 存在$\alpha_{s+2}\in S$, 使得$\alpha_1,\dots,\alpha_{s+1},\alpha_{s+2}$线性无关. 继续下去, 该过程必有限步后停止($\mathbb{F}^n$中任意$n+1$个向量必相关), 即可得到$S$的极大无关组$\alpha_1,\dots,\alpha_s,\alpha_{s+1},\dots,\alpha_r$.

(2) 取$\alpha\in S$, $\alpha\neq 0$. 则向量组α线性无关, 将它进行线性无关扩充, 就可得S的极大无关组. □

5.5.2 向量组的秩

下面我们就可以定义向量组所含向量的真正个数.

定义5.6. *设S是$\mathbb{F}^n$中的向量组, 定义S的秩(rank)$R(S)$为*[6]

$$R(S):=\begin{cases}0, & S\text{ 只含零向量},\\ S\text{ 的任意极大无关组所含向量的个数}, & S\text{ 中有非零向量}.\end{cases}$$

例如, 设$\alpha_1=\begin{pmatrix}9\\12\\15\end{pmatrix}$, $\alpha_2=\begin{pmatrix}1\\2\\3\end{pmatrix}$, $\alpha_3=\begin{pmatrix}4\\5\\6\end{pmatrix}$, 则由例5.10, $R(\alpha_1,\alpha_2,\alpha_3)=2$.

我们看一些例子.

[6] 也有的书上记向量组S的秩为$r(S)$, 或者$\mathrm{rank}(S)$, 或者$\mathrm{Rank}(S)$.

例5.12. (1) $R(\mathbb{F}^n)=n$, 且对于$\mathbb{F}^n$中任意向量组S, 有$R(S)\leqslant n$;

(2) 设T是S的子向量组, 则$R(T)\leqslant R(S)$.

极大无关组是可表示原向量组的线性无关子组. 如果已知一个向量组的秩为r, 那么这个向量组中的r个向量成为极大无关组只需要满足一半条件即可, 即或者线性无关, 或者可表示原向量组.

例5.13. 设S是$\mathbb{F}^n$中的向量组, 且$R(S)=r$. 设$\alpha_1,\alpha_2,\ldots,\alpha_r\in S$, 证明: 下面等价

(i) $\alpha_1,\alpha_2,\ldots,\alpha_r$是$S$的极大无关组;

(ii) $\alpha_1,\alpha_2,\ldots,\alpha_r$线性无关;

(iii) 对任意$\alpha\in S$, α可由$\alpha_1,\alpha_2,\ldots,\alpha_r$线性表示.

☞ **证明.** "(i) $\Longrightarrow$ (ii)" 已知.
"(ii) $\Longrightarrow$ (iii)" 如果存在$\alpha\in S$, 使得α不能由$\alpha_1,\alpha_2,\ldots,\alpha_r$线性表示, 则由$\alpha_1,\alpha_2,\ldots,\alpha_r$线性无关可得$\alpha_1$, α_2, $\ldots$, α_r, α线性无关. 这得到$R(S)\geqslant r+1$, 矛盾.
"(iii) $\Longrightarrow$ (i)" 易知, $\alpha_1,\alpha_2,\ldots,\alpha_r$含有非零向量. 任取$\alpha_1,\alpha_2,\ldots,\alpha_r$的一个极大无关组$T$. 由于任意的$\alpha\in S$, α可由$\alpha_1,\alpha_2,\ldots,\alpha_r$线性表示, 且$\alpha_1,\alpha_2,\ldots,\alpha_r$可由$T$线性表示, 所以$\alpha$可由$T$线性表示. 而$T$线性无关, 所以$T$就是$S$的极大无关组. 这得到

$$R(\alpha_1,\alpha_2,\ldots,\alpha_r)=R(S)=r.$$

所以$\alpha_1,\alpha_2,\ldots,\alpha_r$线性无关是$S$的极大无关组. □

下面例子可视为命题5.13的推广.

例5.14. 设S和T为$\mathbb{F}^n$中等价的有限向量组, 证明: $R(S)=R(T)$.

☞ **证明.** 可以不妨设S和T均含非零向量. 取S的极大无关组$\alpha_1,\alpha_2,\ldots,\alpha_r$, 取$T$的极大无关组$\beta_1,\beta_2,\ldots,\beta_s$, 则有等价

$$S\leftrightarrow\alpha_1,\alpha_2,\ldots,\alpha_r,\qquad T\leftrightarrow\beta_1,\beta_2,\ldots,\beta_s.$$

而$S\leftrightarrow T$, 所以得到等价

$$\alpha_1,\alpha_2,\ldots,\alpha_r\leftrightarrow\beta_1,\beta_2,\ldots,\beta_s.$$

因为$\alpha_1,\alpha_2,\ldots,\alpha_r$和$\beta_1,\beta_2,\ldots,\beta_s$都线性无关, 所以有$r=s$, 即$R(S)=R(T)$. □

5.5.3 线性表示, 线性相关与向量组的秩

极大无关组在某种程度上可以代表(替代)原向量组, 我们用这个观点重新认识一下线性表示. 设有向量组S, 我们用它的极大无关组$\alpha_1,\alpha_2,\ldots,\alpha_r$替代它. 任意给定向量$\beta$, 如果$\beta$可由$S$线性表示, 则$\alpha_1,\ldots,\alpha_r,\beta$线性相关, 此时有$R(S,\beta)=r$; 如果$\beta$不能由$S$线性表示, 则$\alpha_1,\ldots,\alpha_r,\beta$线性无关, 于是把$\beta$加入到$S$后的向量组有极大无关组$\alpha_1,\ldots,\alpha_r,\beta$, 即得到$R(S,\beta)=r+1$. 所以$\beta$是否可以由$S$线性表示可以用$S$加入$\beta$后的向量组的秩是否还是$r$来判别, 数数就可以判别线性表示! 类似地, 向量组是否可以由另一个向量组线性表示, 向量组是否线性无关都可以用数数来判别. 所以, 线性表示, 线性相关除了线性方程组解的情况判别法外, 还有下面的秩的判别法.

命题5.15. 设$\alpha_1,\dots,\alpha_s,\beta_1,\dots,\beta_t,\beta\in\mathbb{F}^n$, 记如下向量组

$$\begin{aligned}&S:\alpha_1,\dots,\alpha_s, && T:\beta_1,\dots,\beta_t,\\ &S,\beta:\alpha_1,\dots,\alpha_s,\beta, && S,T:\alpha_1,\dots,\alpha_s,\beta_1,\dots,\beta_t.\end{aligned}$$

则

(1) β可由S线性表示$\Longleftrightarrow R(S)=R(S,\beta)$, 且

  (i) β可由S唯一地线性表示$\Longleftrightarrow R(S)=R(S,\beta)=s$;

  (ii) β可由S不唯一地线性表示$\Longleftrightarrow R(S)=R(S,\beta)<s$;

(2) $T\leftarrow S\Longleftrightarrow R(S)=R(S,T)$;

(3) $T\leftarrow S\Longrightarrow R(T)\leqslant R(S)$;

(4) $S\leftrightarrow T\Longleftrightarrow R(S)=R(T)=R(S,T)$;

(5) S线性相关$\Longleftrightarrow R(S)<s$;

(6) S线性无关$\Longleftrightarrow R(S)=s$.

☞ **证明.** (1) 由(2)和(5), (6).
(2) 设$T\leftarrow S$. 记$R(S)=r$. 如果$r=0$, 则$\alpha_1=\alpha_2=\cdots=\alpha_s=0$, 进而$\beta_1=\beta_2=\beta_t=0$. 所以此时$R(S,T)=0=R(S)$. 下设$r>0$, 必要时重新下标, 可以不妨设$\alpha_1,\dots,\alpha_r$是$S$的极大无关组. 下面证明$\alpha_1,\dots,\alpha_r$也是$S,T$的极大无关组. 事实上, 由假设和条件

$$S\leftarrow\alpha_1,\dots,\alpha_r,\quad T\leftarrow S,$$

得到

$$T\leftarrow\alpha_1,\dots,\alpha_r\Longrightarrow S,T\leftarrow\alpha_1,\dots,\alpha_r.$$

又$\alpha_1,\dots,\alpha_r$线性无关, 所以$\alpha_1,\dots,\alpha_r$是S,T的极大无关组. 因此$R(S,T)=r=R(S)$.

反之, 设$R(S)=R(S,T)=r$. 当$r=0$时, 有$\alpha_1=\cdots=\alpha_s=\beta_1=\cdots=\beta_t=0$. 这得到$T\leftarrow S$. 如果$r>0$, 不妨设$\alpha_1,\dots,\alpha_r$是$S$的极大无关组, 则$\alpha_1,\dots,\alpha_r$也是$S,T$的极大无关组. 于是

$$T\leftarrow\alpha_1,\dots,\alpha_r.$$

又显然

$$\alpha_1,\dots,\alpha_r\leftarrow S,$$

所以就有$T\leftarrow S$.
(3) 由(2)得$R(S)=R(S,T)$. 而$R(T)\leqslant R(S,T)$, 所以$R(T)\leqslant R(S)$.
(4) 由(2)得

$$S\leftrightarrow T\Longleftrightarrow R(S)=R(S,T),R(T)=R(T,S).$$

但是$R(S,T)=R(T,S)$, 所以上面等价于$R(S)=R(T)=R(S,T)$.
(5) 由(6).
(6) 如果S线性无关, 则S为S的极大无关组, 所以$R(S)=s$.

反之, 显然任意$\alpha\in S$, α可由S线性表示. 又$R(S)=s$, 则S是S的极大无关组. 特别地, S线性无关. □

现在我们可以给出命题5.9(1)充分性的证明.

命题5.9(1)充分性的证明. 我们要证: 对任意$A, B \in \mathbb{F}^{m\times n}$, 如果$A$和$B$的列向量组等价, 则$A \overset{c}{\sim} B$.

由于A和B的列向量组等价, 所以A和B的列向量组的秩相等, 记为r. 当$r=0$时, 有$A=O=B$. 此时显然$A \overset{c}{\sim} B$. 下设$r>0$, 设$\alpha_1,\alpha_2,\ldots,\alpha_r$是$A$的列向量组的极大无关组, $\beta_1,\beta_2,\ldots,\beta_r$是$B$的列向量组的极大无关组. 于是

$$\alpha_1,\alpha_2,\ldots,\alpha_r \leftrightarrow A\text{ 的列向量组},\quad \beta_1,\beta_2,\ldots,\beta_r \leftrightarrow B\text{ 的列向量组}.$$

又A和B的列向量组等价, 所以

$$\alpha_1,\alpha_2,\ldots,\alpha_r \leftrightarrow \beta_1,\beta_2,\ldots,\beta_r.$$

再利用$\alpha_1,\alpha_2,\ldots,\alpha_r$和$\beta_1,\beta_2,\ldots,\beta_r$都线性无关, 例5.9推出

$$(\alpha_1,\alpha_2,\ldots,\alpha_r) \overset{c}{\sim} (\beta_1,\beta_2,\ldots,\beta_r).$$

进而有

$$(\alpha_1,\alpha_2,\ldots,\alpha_r,\underbrace{0,\ldots,0}_{n-r}) \overset{c}{\sim} (\beta_1,\beta_2,\ldots,\beta_r,\underbrace{0,\ldots,0}_{n-r}).$$

设A的其它列向量是$\alpha_{r+1},\ldots,\alpha_n$, B的其它列向量是$\beta_{r+1},\ldots,\beta_n$, 用调列变换可得

$$A \overset{c}{\sim} (\alpha_1,\ldots,\alpha_r,\alpha_{r+1},\ldots,\alpha_n),\quad B \overset{c}{\sim} (\beta_1,\ldots,\beta_r,\beta_{r+1},\ldots,\beta_n).$$

又

$$\alpha_{r+1},\ldots,\alpha_n \leftarrow \alpha_1,\ldots,\alpha_r,\quad \beta_{r+1},\ldots,\beta_n \leftarrow \beta_1,\ldots,\beta_r,$$

所以可用列消去变换得

$$A \overset{c}{\sim} (\alpha_1,\ldots,\alpha_r,\underbrace{0,\ldots,0}_{n-r}),\quad B \overset{c}{\sim} (\beta_1,\ldots,\beta_r,\underbrace{0,\ldots,0}_{n-r}).$$

所以由列等价的对称性和传递性得$A \overset{c}{\sim} B$. □

习题5.5

A1. 求下列向量组的极大无关组与秩:

(1) $\alpha_1=(1,1,1)^T$, $\alpha_2=(1,-2,0)^T$, $\alpha_3=(-2,3,1)^T$;

(2) $\alpha_1=(1,3,-1)^T$, $\alpha_2=(2,1,1)^T$, $\alpha_3=(0,5,-3)^T$.

A2. 已知两个有限向量组有相同的秩, 且其中一个可以由另一个线性表示, 证明: 这两个向量组等价.

A3. 设有向量$\alpha_1=\begin{pmatrix}1\\2-a\\3-2a\end{pmatrix}$, $\alpha_2=\begin{pmatrix}1\\2-a\\2-a\end{pmatrix}$, $\alpha_3=\begin{pmatrix}2-a\\1\\1\end{pmatrix}$, $\beta=\begin{pmatrix}1\\1\\a\end{pmatrix}$, 对$a$的不同取值, 讨论$\beta$由$\alpha_1,\alpha_2,\alpha_3$线性表示的情况. 当$\beta$可由$\alpha_1,\alpha_2,\alpha_3$不唯一线性表示时, 求出具体的表示式.

§5.6 矩阵的秩—定义与计算

前面我们已经介绍了如何数向量组所含向量的真正个数, 从本节开始我们来数一些特殊的向量组的向量个数. 我们的目的是数线性方程组中所含方程的个数和数线性方程组解的个数. 前者可以认为是数线性方程组的增广矩阵的行数, 后者在一定程度上可以认为是数由解向量构成的一类特殊

的空间的向量个数. 于是一般地, 我们来数矩阵所含行(列)的个数, 以及数由向量构成的一类特殊的空间的向量个数, 然后再应用到线性方程组上.

本节的任务是数矩阵中真正行(列)的数目.

5.6.1 秩的定义

定义5.7. 设$A \in \mathbb{F}^{m\times n}$.

(1) A的行向量组的秩称为A为行秩, 记为$R_r(A)$;

(2) A的列向量组的秩称为A为列秩, 记为$R_c(A)$.

由定义, $R_r(A) \leqslant m$, $R_c(A) \leqslant n$. 如果$R_r(A) = m$, 即A的行向量组线性无关, 则称A行满秩; 如果$R_c(A) = n$, 即A的列向量组线性无关, 则称A列满秩.

同一个矩阵有两个秩, 这看上去比较奇怪. 它们会不会相等呢? 我们看二阶方阵$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. 如果$R_r(A) = 0$, 则$A$的行向量组只有零向量, 即有$A = O$. 所以$A$的列向量组也只有零向量, $R_c(A) = 0$. 如果$R_r(A) = 1$, 则可以不妨设

$$(a, b) \neq 0, \quad (c, d) = t(a, b), \quad \exists t \in \mathbb{F}.$$

如果$a \neq 0$, 则$(a, c)^T \neq 0$, 且

$$\frac{b}{a}\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} b \\ \frac{bc}{a} \end{pmatrix} = \begin{pmatrix} b \\ tb \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}.$$

于是$R_c(A) = 1$. 类似地, 当$b \neq 0$时, 也可得$R_c(A) = 1$. 最后, 如果$R_r(A) = 2$, 则A的行向量组线性无关, 这得到A^T的列向量组线性无关. 于是$|A^T| \neq 0$, 进而$|A| \neq 0$. 所以A的列向量组线性无关, $R_c(A)$=2. 因此, 总有$R_r(A) = R_c(A)$, 即行秩和列秩相等. 于是我们不要怀疑数学的美, 大胆地写下下面的结论.

定理5.16. 对任意$A \in \mathbb{F}^{m\times n}$, 有$R_r(A) = R_c(A)$.

既然行秩和列秩一致, 我们就可以统称为秩.

定义5.8. 设$A \in \mathbb{F}^{m\times n}$, A的秩$R(A)$定义为[7]

$$R(A) := R_r(A) = R_c(A).$$

下面我们想办法证明定理5.16. 我们先证明下面的结论, 注意我们如何行列互换.

引理5.17. 对$A \in \mathbb{F}^{m\times n}$, 有

(1) $R_r(A) = R_c(A^T)$, $R_c(A) = R_r(A^T)$;

(2) $R_r(A) \leqslant R_c(A)$.

---

[7]矩阵A的秩也用记号rank(A)或$r(A)$表示.

☞ **证明.** (1) 这由于A的行(列)向量组就是A^T的列(行)向量组.

(2) 设$R_c(A) = r$. 如果$r = 0$, 则$A = O$, 于是$R_r(A) = 0 \leqslant R_c(A)$. 下设$r > 0$, 于是存在$A$的$m \times r$子阵$A_1$, 使得$A_1$的$r$个列是$A$的列向量组的极大无关组. 因为$A$的列向量组可以由$A_1$的列向量组线性表示, 所以

$$A = A_1 C, \quad \exists C \in \mathbb{F}^{r \times n}.$$

由此又得A的行向量组可由C的行向量组线性表示, 所以

$$R_r(A) \leqslant R_r(C) \leqslant r = R_c(A).$$

证毕. □

☞ **定理5.16的证明.** 对A和A^T用引理5.17(2), 得

$$R_r(A) \leqslant R_c(A), \qquad R_r(A^T) \leqslant R_c(A^T).$$

再利用引理5.17(1)可知第二个不等式即$R_c(A) \leqslant R_r(A)$. 于是$R_r(A) = R_c(A)$. □

推论5.18. 设$A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times s}$, 则

(1) $0 \leqslant R(A) \leqslant \min\{m, n\}$, 且有

$$R(A) = 0 \Longleftrightarrow A = O;$$

(2) $R(A^T) = R(A)$;

(3) $R(AB) \leqslant \min\{R(A), R(B)\}$.

☞ **证明.** (3) 由于AB的列向量组可由A的列向量组线性表示, 所以$R(AB) \leqslant R(A)$. 又

$$R(AB) = R((AB)^T) = R(B^T A^T) \leqslant R(B^T) = R(B).$$

证毕. □

5.6.2 秩的等价定义

前面给出了矩阵的秩的几何定义: 行(列)向量组的秩. 向量组的秩就是极大无关组所含向量的个数, 而线性无关性又与方阵的行列式有关系. 设α_1, ..., α_n是$\mathbb{F}^n$中的向量组, 记矩阵$A = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^{n \times n}$. 则

$$\alpha_1, \ldots, \alpha_n\text{线性无关} \Longleftrightarrow |A| \neq 0.$$

于是线性无关对应到行列式非零, 下面证明一般情况下线性无关的向量组对应到矩阵的非零子式.

引理5.19. 设$\alpha_1, \alpha_2, \ldots, \alpha_r$是$\mathbb{F}^n$中的向量组, 其中$r \leqslant n$. 记矩阵$A = (\alpha_1, \alpha_2, \ldots, \alpha_r) \in \mathbb{F}^{n \times r}$, 则

$$\alpha_1, \alpha_2, \ldots, \alpha_r\text{线性无关} \Longleftrightarrow A\text{有}r\text{阶非零子式}.$$

☞ **证明.** 设$\alpha_1, \alpha_2, \ldots, \alpha_r$线性无关, 则$R(A) = r$. 于是$A$的行向量组的秩也是$r$, 这得到存在$A$的$r$个行为$A$的行向量组的极大无关组. 这$r$个行按照相对次序排成的子阵记为$A_1$, 则$A_1$是$r$阶方阵, 且$R(A_1) = r$. 因此$A_1$的列向量组线性无关, 这得到$|A_1| \neq 0$, 即$A$有$r$阶非零子式$|A_1|$.

反之, 设A有r阶非零子式$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix}$. 这个子式对应的$r$阶方阵记为$A_2$, 则$A_2$的列向量组线性无关. 而$A_2$的列向量组是$A$的列向量组的缩短向量组, 于是$A$的列向量组也线性无关, 即$\alpha_1, \alpha_2, \ldots, \alpha_r$线性无关. □

矩阵的秩等于它的行(列)向量组的极大无关组所含向量的个数, 极大无关组应该对应到矩阵的最大阶的非零子式. 于是下面秩的等价刻画成立也就不奇怪了.

定理5.20. *设$A \in \mathbb{F}^{m\times n}$, 则$R(A)$等于$A$的非零子式的最大阶数.*

☞ **证明.** 设$R(A) = r$, A的非零子式的最大阶数是s. 如果$r = 0$, 则$A = O$, 于是$s = 0 = r$.

下设$r > 0$. 由定义, A有r个列向量线性无关, 引理5.19推出A有r阶非零子式, 所以$s \geqslant r > 0$. 又A有s阶非零子式, 引理5.19推出A有s 个列向量线性无关, 所以$r \geqslant s$. 最后得$r = s$. □

由于n阶方阵的n阶子式只有一个: A的行列式$|A|$, 所以定理5.20有下面的推论, 这给出了矩阵可逆的更多的刻画.

推论5.21. *设$A \in \mathbb{F}^{n\times n}$, 则下面命题等价*

(i) *A可逆;*

(ii) *$|A| \neq 0$;*

(iii) *$R(A) = n$;* (此时称A满秩)

(iv) *A的行向量组线性无关;*

(v) *A的列向量组线性无关.*

例5.15. *设$A \in \mathbb{F}^{m\times n}$, 证明:*

(1) *如果A有r阶非零子式, 则$R(A) \geqslant r$;*

(2) *如果A的所有s阶子式为零, 则$R(A) < s$.*

☞ **证明.** (1) 由秩的等价定义立刻得到.

(2) 任取k $(s < k \leqslant \min\{m, n\})$, 考察$A$的一个$k$阶子式. 将这个$k$阶子式按照前$s$行Laplace展开, 其值为

$$\sum (\pm 1)(s\text{阶子式})(k - s\text{阶子式}) = 0.$$

所以A的非零子式的最大阶数一定小于s, 即$R(A) < s$. □

由秩的代数定义, 还容易得到:

$$R(aA) = R(A), \quad (\forall A \in \mathbb{F}^{m\times n}, 0 \neq \forall a \in \mathbb{F}).$$

5.6.3 初等变换与矩阵的秩

下面我们的首要任务是给出秩的有效的计算方法, 我们先看矩阵的秩. 矩阵的秩在几何上可以看成其行(列)向量组的秩, 在代数上是其非零子式的最大阶数. 于是按照定义来算, 可以看下面的例子.

例5.16. 设$A=\begin{pmatrix}2&4&3&3\\1&2&1&1\\1&2&3&3\end{pmatrix}$, 求$R(A)$.

☞ **解.** 我们记$\alpha_1=\begin{pmatrix}2\\1\\1\end{pmatrix}$, $\alpha_2=\begin{pmatrix}4\\2\\2\end{pmatrix}$, $\alpha_3=\begin{pmatrix}3\\1\\3\end{pmatrix}$, 则$A=(\alpha_1,\alpha_2,\alpha_3,\alpha_3)$. 于是$R(A)$等于向量组$\alpha_1,\alpha_2,\alpha_3$的秩. 而$\alpha_2=2\alpha_1$, 所以$\alpha_1,\alpha_2,\alpha_3$线性相关. 又由于$\alpha_1,\alpha_3$不成比例, 所以$\alpha_1,\alpha_3$线性无关. 于是$\alpha_1,\alpha_3$是$\alpha_1$, α_2, α_3 的极大无关组. 所以$R(A)=2$.

也可以计算A的非零子式的最大阶数. 先看最高阶子式: 3阶子式, 共有4个. 由于

$$A\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=\begin{vmatrix}2&4&3\\1&2&1\\1&2&3\end{vmatrix}=0,\quad A\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}=\begin{vmatrix}2&4&3\\1&2&1\\1&2&3\end{vmatrix}=0,$$

$$A\begin{pmatrix}1&2&3\\1&3&4\end{pmatrix}=\begin{vmatrix}2&3&3\\1&1&1\\1&3&3\end{vmatrix}=0,\quad A\begin{pmatrix}1&2&3\\2&3&4\end{pmatrix}=\begin{vmatrix}4&3&3\\2&1&1\\2&3&3\end{vmatrix}=0,$$

所以$R(A)<3$. 而二阶子式

$$A\begin{pmatrix}1&2\\1&3\end{pmatrix}=\begin{vmatrix}2&3\\1&1\end{vmatrix}=-1\neq 0,$$

所以$R(A)=2$. □

上面例子的第一种解法有很大的随机性, 当矩阵改变时就比较难求列向量组的极大无关组(事实上我们后面要介绍求法); 第二种解法虽然原则上都可以算出秩, 但是当矩阵的行和列都很大时, 需要算很多高阶行列式, 计算量会非常大. 于是我们必须寻找其它方法来计算矩阵的秩. 按照我们解决问题(解线性方程组, 判别可逆性, 计算行列式等等)的经验, 我们可以尝试考虑初等变换

$$\boxed{\text{矩阵}}\overset{\text{行初等变换}}{\sim}\boxed{\text{(简化)阶梯形阵}}$$

于是我们要考虑

(i) 在初等变换下矩阵的秩如何改变?

(ii) (简化)阶梯形阵的秩是否可以快速给出?

这两个问题. 我们先考虑问题(i). 我们知道等价的向量组的秩相等, 而矩阵等价基本等同于向量组等价, 于是可以得到初等变换下矩阵的秩不改变. 具体地, 设$A,B\in\mathbb{F}^{m\times n}$. 如果$A\overset{r}{\sim}B$, 则$A$和$B$的行向量组等价. 于是$R(A)=R_r(A)=R_r(B)=R(B)$. 类似的, 如果$A\overset{c}{\sim}B$, 则$A$和$B$的列向量组等价. 于是$R(A)=R_c(A)=R_c(B)=R(B)$. 特别的, 对矩阵施行一次初等变换不改变秩. 于是由$A\sim B$可得$R(A)=R(B)$. 进而我们得到下面的定理.

定理5.22. 设$A,B\in\mathbb{F}^{m\times n}$, 则

(1) $A\sim B \iff R(A)=R(B)$;

(2) 如果A的初等变换标准形是$\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}_{m\times n}$, 那么$r=R(A)$.

☞ **证明.** (1)的必要性已证, 下面先证(2), 再证(1)的充分性.
(2) 由条件, $A\sim\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}=F$. 利用(1)的必要性就得$R(A)=R(F)$. 但是$F$有$r$阶子式

$$F\begin{pmatrix}1 & 2 & \cdots & r\\ 1 & 2 & \cdots & r\end{pmatrix}=|E_r|=1\neq 0,$$

且F的任意$r+1$阶子式必有全零行, 为零. 这得到$R(F)=r$, 进而$r=R(A)$.
(1) "$\Longleftarrow$" 设$R(A)=R(B)=r$, 则由(2)得

$$A\sim\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}_{m\times n},\quad B\sim\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}_{m\times n}.$$

再利用矩阵等价的对称性和传递性(引理2.2), 就得$A\sim B$. □

注意由$R(A)=R(B)$不能得到A和B等价, 这由于不同型的矩阵的秩可以相等, 而矩阵等价必须同型. 例如$A=\begin{pmatrix}1\\1\end{pmatrix}$和$B=\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$的秩都是1, 但它们不等价.

由于矩阵等价相当于可逆阵做乘法, 所以可得乘以可逆阵不改变矩阵的秩, 进而可知在分块矩阵的初等变换下, 矩阵的秩也不改变.

推论5.23. 设$A,B\in\mathbb{F}^{m\times n}$, 则

(1) 如果P是m阶可逆方阵, Q是n阶可逆方阵, 那么$R(PAQ)=R(A)$;

(2) 如果A可用分块阵的初等变换变为B, 那么$R(A)=R(B)$.

☞ **证明.** (1) 由于$PAQ\sim A$, 再用定理5.22.
(2) 存在分块初等阵$P_1,\ldots,P_s$和$Q_1,\ldots,Q_r$, 使得

$$P_s\cdots P_1AQ_1\cdots Q_r=B.$$

再利用分块初等阵都可逆和(1). □

现在我们已经非常清楚矩阵等价的标准形以及两个同型的矩阵等价的充分必要条件:

(i) 设$A\in\mathbb{F}^{m\times n}$, 则$A\sim\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}_{m\times n}$, 其中$r=R(A)$;

(ii) 设$A,B\in\mathbb{F}^{m\times n}$, 则: $A\sim B\iff R(A)=R(B)$.

这事实上解决了矩阵等价分类的两个基本问题. 在我们看来, 矩阵的分类是认识该课程的一个重要观点, 所以我们下面花费些笔墨来阐述这个观点.

- **等价关系和集合的分类**

这里简单介绍一下等价关系和集合的分类. 设X是一个集合, $\sim$是X上的一个二元关系. 如果$\sim$满足

(i) (自反性) 对任意$a \in X$, 有$a \sim a$;

(ii) (对称性) 对任意$a, b \in X$, 如果$a \sim b$, 那么$b \sim a$;

(iii) (传递性) 对任意$a, b, c \in X$, 如果$a \sim b$且$b \sim c$, 那么$a \sim c$,

则称$\sim$是X上的一个等价关系.

例5.17. **(1)** *整除关系不是$\mathbb{Z}$上的等价关系(不满足对称性);*

(2) *设$m \in \mathbb{N}$, 对任意$a, b \in \mathbb{Z}$, 如果$m \mid a - b$, 则称a和b模m同余, 记为$a \equiv b \pmod{m}$. 容易证明, 模m同余关系是$\mathbb{Z}$上的等价关系;*

(3) *矩阵的(行, 列)等价关系是$\mathbb{F}^{m \times n}$上的等价关系(引理2.2).*

设$\sim$是集合X上的等价关系, 任意的$a \in X$, 记

$$[a] := \{x \in X \mid x \sim a\} \subset X.$$

由自反性, $a \in [a]$. 称$[a]$为a所在的等价类, 而a称为$[a]$的一个代表元, $[a]$中的任意元素都是它的代表元, 所以通常代表元不唯一. 由对称和传递性, 等价类中的元素两两等价. 任取$a, b \in X$, 如果$[a] \cap [b] \neq \varnothing$, 则存在$x \in X$, 使得$x \sim a$且$x \sim b$. 于是$a \sim b$, 进而容易得到$[a] = [b]$. 于是两个等价类或者相等, 或者交为空集; 不同的等价类之间的元素必不等价. 又X中的每个元素一定属于某个等价类, 所以X等于其所有等价类的无交并

$$X = \bigsqcup_a [a].$$

这就给出了集合X的一个分类.

反之, 如果集合X有一个分类, 即有无交并

$$X = \bigsqcup_{i \in I} X_i,$$

其中I是下标集, $X_i \subset X$, 则可如下定义X上的等价关系$\sim$: 对任意$a, b \in X$, 如果存在$i \in I$, 使得$a, b \in X_i$, 则定义$a \sim b$. 这个等价关系给出的X的等价类分类恰好是开始给出的分类.

将所有等价类做成的集合记为$X/\sim$, 即

$$X/\sim := \{[a] \mid a \in X\}.$$

称$X/\sim$为集合X关于等价关系$\sim$的商集. 由定义, 商集中的元素是X的子集. 可以看到, 通常商集比原来的集合简单.

例5.18. 我们看$\mathbb{Z}$在模2同余关系下的分类. 这时有两个等价类:

$$[0]=\{n\in\mathbb{Z}\mid n\equiv 0 \pmod 2\}=\{\text{偶数}\},$$
$$[1]=\{n\in\mathbb{Z}\mid n\equiv 1 \pmod 2\}=\{\text{奇数}\}.$$

所以模2同余关系恰好给出了整数的奇, 偶分类. 此时商集是

$$\{[0],[1]\},$$

它是二元集, 通常记为$\mathbb{Z}/2\mathbb{Z}$.

- **矩阵的等价分类**

分类问题是数学的一个基本问题, 线性代数从代数的观点来看就是研究矩阵在各种等价关系下的分类问题. 设$\equiv$是矩阵集合$\mathbb{F}^{m\times n}$上的一个等价关系, 这就得到$\mathbb{F}^{m\times n}$关于关系$\equiv$的等价类分类. 我们要解决两个基本问题:

(i) 全系不变量问题: 如果一个量为等价类中每个矩阵所共有, 则称这种量为矩阵在关系$\equiv$下的不变量. 如果一组$\equiv$下的不变量足以区分不同的等价类, 而且当这组不变量缺少某一个就不能区分不同的等价类, 那么这组不变量就称为矩阵在关系$\equiv$下的全系不变量. 判定两个矩阵是否属于同一个等价类的问题, 就是寻求$\equiv$下的全系不变量;

(ii) 标准形问题: 在每个等价类中选取代表元, 使它具有最简单的形式, 而且从它容易读出全系不变量.

由引理2.2, 矩阵等价是矩阵集合$\mathbb{F}^{m\times n}$上的等价关系, 于是这得到$\mathbb{F}^{m\times n}$的等价分类. 定理5.22解决了矩阵在矩阵等价下的分类问题, 等价标准形可以取成初等变换标准形, 而秩是矩阵在矩阵等价下的全系不变量. 见图5.4. 我们有商集合

$$\mathbb{F}^{m\times n}/\sim\ =\left\{\{\text{秩}r\text{的矩阵}\in\mathbb{F}^{m\times n}\}\mid 0\leqslant r\leqslant\min\{m,n\}\right\}.$$

$\mathbb{F}^{m\times n}$	秩0	秩1	秩2	$\cdots$	秩l	$l=\min\{m,n\}$
标准形	O	$\begin{pmatrix}E_1 & O\\ O & O\end{pmatrix}$	$\begin{pmatrix}E_2 & O\\ O & O\end{pmatrix}$	$\cdots$	$\begin{pmatrix}E_l & O\\ O & O\end{pmatrix}$	

图 5.4: 矩阵集合$\mathbb{F}^{m\times n}$在矩阵等价下的分类

在处理许多矩阵问题时, 常常可以先考虑矩阵在某种分类下的标准形(这通常比较简单), 然后再考虑一般情形. 这种解决矩阵问题的方法称为标准形方法.

例5.19. 设$A \in \mathbb{F}^{m\times n}$, 则

$$R(A)=1 \Longleftrightarrow 0 \neq \exists \alpha \in \mathbb{F}^m, 0 \neq \exists \beta \in \mathbb{F}^n,\ \text{使得}\ A=\alpha\beta^T.$$

☞ **证明.** 设$A=\alpha\beta^T$, 其中$\alpha \in \mathbb{F}^m$, $\beta \in \mathbb{F}^n$都非零. 于是

$$R(A) \leqslant R(\alpha)=1.$$

而$A \neq O$, 所以$R(A) \geqslant 1$. 这得到$R(A)=1$.

反之, 设$R(A)=1$, 则存在m阶可逆阵P和n阶可逆阵Q, 使得

$$A=P\begin{pmatrix} E_1 & O \\ O & O \end{pmatrix}_{m\times n} Q.$$

而有分解(先考虑等价标准形!)

$$\begin{pmatrix} E_1 & O \\ O & O \end{pmatrix}_{m\times n} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{m\times 1} (1,0,\ldots,0)_{1\times n},$$

所以记

$$\alpha = P\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{m\times 1} \in \mathbb{F}^m, \quad \beta^T=(1,0,\ldots,0)_{1\times n}Q(\Longrightarrow \beta \in \mathbb{F}^n),$$

则α,β非零, 且$A=\alpha\beta^T$. □

5.6.4 秩的计算

定理5.22告诉我们, 初等变换不改变矩阵的秩. 于是应该可以利用初等变换来求矩阵的秩, 进一步可以利用初等变换来求有限向量组的秩. 有时候我们需要求向量组的一个极大无关组, 所以我们需要考虑初等变换是否改变极大无关组. 分析发现, 确实行初等变换不改变矩阵的列向量组的线性关系, 进而不改变极大无关组.

引理5.24. 设$A,B \in \mathbb{F}^{m\times n}$, 且$A \overset{r}{\sim} B$, 则$A$和$B$的列向量组间有相同的线性关系. 即若$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)$, $B=(\beta_1,\beta_2,\ldots,\beta_n)$, 则对于$c_1,c_2,\ldots,c_n \in \mathbb{F}$, 有

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n=0 \Longleftrightarrow c_1\beta_1+c_2\beta_2+\cdots+c_n\beta_n=0.$$

☞ **证明.** 由于$A \overset{r}{\sim} B$, 所以存在m阶可逆阵P, 使得$PA=B$. 这得到

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n=0 \Longleftrightarrow A\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}=0$$

$$\Longleftrightarrow PA\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}=0 \Longleftrightarrow B\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}=0$$

$$\Longleftrightarrow c_1\beta_1+c_2\beta_2+\cdots+c_n\beta_n=0.$$

证毕. □

最后剩下的任务就是看如何求(简化)阶梯形阵的秩和其列向量组的一个极大无关组, 这个并不困难.

例5.20. 设$B=(\beta_1,\beta_2,\ldots,\beta_n)\in\mathbb{F}^{m\times n}$是阶梯形矩阵, 其首元列是第$j_1<j_2<\cdots<j_r$列. 证明:

(1) $R(B)=r$;

(2) $\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_r}$是$B$的列向量组的一个极大无关组;

(3) 如果B还是简化阶梯形阵, 且$\beta_j=(b_{1j},b_{2j},\ldots,b_{rj},0,\ldots,0)^T$, 则

$$\beta_j=b_{1j}\beta_{j_1}+b_{2j}\beta_{j_2}+\cdots+b_{rj}\beta_{j_r},\quad (j=1,2,\ldots,n).$$

☞ **证明.** 记$B=(b_{ij})$, 则B的r个首元从左到右为$b_{1j_1},b_{2j_2},\ldots,b_{rj_r}$. 于是$B$有$r$阶子式

$$B\begin{pmatrix}1&2&\cdots&r\\j_1&j_2&\cdots&j_r\end{pmatrix}=\begin{vmatrix}b_{1j_1}&*&\cdots&*\\&b_{2j_2}&\cdots&*\\&&\ddots&\vdots\\&&&b_{rj_r}\end{vmatrix}=b_{1j_1}b_{2j_2}\cdots b_{rj_r}\neq0.$$

而B的任意阶数大于r的子式一定有全零行, 进而为零. 这得到$R(B)=r$, 且$\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_r}$线性无关, 是$B$的列向量组的极大无关组.

当B是简化阶梯形阵时, 有

$$\beta_{j_i}=(0,\ldots,0,\underset{i}{1},0,\ldots,0)^T=e_i,\quad i=1,2,\ldots,r.$$

于是有

$$\beta_j=b_{1j}\beta_{j_1}+b_{2j}\beta_{j_2}+\cdots+b_{rj}\beta_{j_r}.$$

证毕. □

利用例5.20和引理5.24, 我们就得到一般矩阵的秩, 其列向量的一个极大无关组, 以及用该极大无关组表示其余列向量的行初等变换的方法.

推论5.25. 设$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathbb{F}^{m\times n}$, 而$B=(\beta_1,\beta_2,\ldots,\beta_n)$是与$A$行等价的阶梯形阵, 首元列为第$j_1<j_2<\cdots<j_r$列, 则

(1) $R(A)=r$;

(2) $\alpha_{j_1},\ \alpha_{j_2},\ \ldots,\ \alpha_{j_r}$为$A$的列向量组的一个极大无关组;

(3) 如果B为简化阶梯形阵, 且$\beta_j = (b_{1j}, b_{2j}, \dots, b_{rj}, 0, \dots, 0)^T$, 那么

$$\alpha_j = b_{1j}\alpha_{j_1} + b_{2j}\alpha_{j_2} + \cdots + b_{rj}\alpha_{j_r}, \quad (j = 1, 2, \dots, n).$$

总结一下秩的计算. 首先, 对于矩阵的秩

算法5.1. 矩阵A的秩的计算方法:

(1) $A \sim B$: 阶梯形阵;

(2) $R(A)$等于B中非零行的数目.

例5.21 (例5.16). 设$A = \begin{pmatrix} 2 & 4 & 3 & 3 \\ 1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 3 \end{pmatrix}$, 求$R(A)$.

☞ **解.** 进行行初等变换

$$A \overset{r_1 \leftrightarrow r_2}{\sim} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 3 & 3 \\ 1 & 2 & 3 & 3 \end{pmatrix} \underset{r_3 - r_1}{\overset{r_2 - 2r_1}{\sim}} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{pmatrix} \overset{r_3 - 2r_2}{\sim} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

于是$R(A) = 2$. □

其次, 对于有限向量组的秩的计算和极大无关组的寻找

算法5.2. 有限(列)向量组的秩和极大无关组的计算步骤: 设有$\mathbb{F}^n$中的向量组$\alpha_1, \alpha_2, \dots, \alpha_m$.

(1) 作矩阵$A = (\alpha_1, \alpha_2, \dots, \alpha_m) \in \mathbb{F}^{n \times m}$;

(2) $A \overset{r}{\sim} B$: 阶梯形阵;

(3) 秩等于B中非零行的数目;

(4) 如果B的首元列为第$j_1 < \cdots < j_r$列, 则$\alpha_{j_1}, \dots, \alpha_{j_r}$为极大无关组;

(5) 若要将其余向量用$\alpha_{j_1}, \dots, \alpha_{j_r}$线性表示,

  (i) $B \overset{r}{\sim} R$: 简化阶梯形阵;

  (ii) 设$R = (\beta_1, \beta_2, \dots, \beta_m)$, 而$\beta_j = (b_{1j}, b_{2j}, \dots, b_{rj}, 0, \dots, 0)^T$, 则

$$\alpha_j = b_{1j}\alpha_{j_1} + b_{2j}\alpha_{j_2} + \cdots + b_{rj}\alpha_{j_r}.$$

如果问题中给出的是行向量组怎么办? 当然是取转置转化为列向量组处理了.

例5.22. 求向量组$\alpha_1 = (1, 0, 2, 1)$, $\alpha_2 = (1, 2, 0, 1)$, $\alpha_3 = (2, 1, 3, 0)$, $\alpha_4 = (2, 5, -1, 4)$, $\alpha_5 = (1, -1, 3, -1)$的秩和一个极大无关组, 并用这个极大无关组表示其余向量.

☞ **解.** 由于

$$(\alpha_1^T,\alpha_2^T,\alpha_3^T,\alpha_4^T,\alpha_5^T)=\begin{pmatrix}1&1&2&2&1\\0&2&1&5&-1\\2&0&3&-1&3\\1&1&0&4&-1\end{pmatrix}\underset{r_4-r_1}{\underbrace{r_3-2r_1}}\begin{pmatrix}1&1&2&2&1\\0&2&1&5&-1\\0&-2&-1&-5&1\\0&0&-2&2&-2\end{pmatrix}$$

$$\underset{r_3\leftrightarrow r_4}{\underbrace{\substack{-\frac{1}{2}r_4\\ r_3+r_2}}}\begin{pmatrix}1&1&2&2&1\\0&2&1&5&-1\\0&0&1&-1&1\\0&0&0&0&0\end{pmatrix}\underset{r_2-r_3}{\underbrace{r_1-2r_3}}\begin{pmatrix}1&1&0&4&-1\\0&2&0&6&-2\\0&0&1&-1&1\\0&0&0&0&0\end{pmatrix}$$

$$\underset{r_1-r_2}{\underbrace{\tfrac{1}{2}r_2}}\begin{pmatrix}1&0&0&1&0\\0&1&0&3&-1\\0&0&1&-1&1\\0&0&0&0&0\end{pmatrix},$$

所以向量组的秩是3, 有极大无关组$\alpha_1,\alpha_2,\alpha_3$, 且

$$\alpha_4=\alpha_1+3\alpha_2-\alpha_3,\qquad \alpha_5=-\alpha_2+\alpha_3.$$

解毕. □

习题5.6

A1. 计算下列矩阵的秩:

(1) $\begin{pmatrix}2&-1&3&-2&4\\4&-2&5&1&7\\2&-1&1&8&2\end{pmatrix}$; (2) $\begin{pmatrix}1&-1&2&1&0\\2&-2&4&-2&0\\3&0&6&-1&1\\0&0&0&0&1\end{pmatrix}$;

(3) $A=(a_{ij})\in\mathbb{F}^{n\times n}$是对称阵, 并且当$1\leqslant i<j\leqslant n$时, $a_{ij}=j$, 而当$1\leqslant i\leqslant n$时, $a_{ii}=i$;

(4) $A=(a_{ij})\in\mathbb{F}^{n\times n}$是反对称阵, 并且当$1\leqslant i<j\leqslant n$时, $a_{ij}=i$.

A2. 设矩阵$A=\begin{pmatrix}3&1&1&4\\a&4&10&1\\1&7&17&3\\2&2&4&3\end{pmatrix}$, 求使得$A$的秩最小的$a$值.

A3. 求下列向量组的秩和一个极大无关组, 并把向量组中其余向量用极大无关组线性表示.

(1) $\alpha_1=(6,4,1,-1,2)^T$, $\alpha_2=(1,0,2,3,-4)^T$, $\alpha_3=(1,4,-9,-6,22)^T$, $\alpha_4=(7,1,0,-1,3)^T$;

(2) $\alpha_1=(1,-1,2,4)$, $\alpha_2=(2,-1,1,0)$, $\alpha_3=(-1,0,1,4)$, $\alpha_4=(1,-2,-1,1)$, $\alpha_5=(1,2,5,2)$.

A4. 在秩是r的矩阵中有没有可能有等于0的$r-1$阶子式? 有没有可能有等于0的r阶子式? 有没有可能有不等于0的$r+1$阶子式? 如果可能有的话, 请举个具体的例子; 如果一定没有的话, 请给出证明.

A5. 证明: 任意一个秩为r的矩阵可以表为r个秩为1的矩阵之和.

§ 5.7 矩阵的秩—性质与应用

在结束秩的讨论, 将前面的讨论用到线性方程组的两个数数问题前, 我们还有些意犹未尽. 让我

们再说说矩阵的秩吧.

5.7.1 矩阵的满秩分解

把上节的讨论矩阵化, 会得到什么呢? 设矩阵$A\in\mathbb{F}^{m\times n}$的秩$R(A)=r$, A有列分块$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)$, 其中$\alpha_j\in\mathbb{F}^m$. 如果$B=(\beta_1,\beta_2,\ldots,\beta_n)$是与$A$行等价的简化阶梯形阵, 则$B$有$r$个非零行, 所以可设$\beta_j=\begin{pmatrix}b_{1j}\\ \vdots\\ b_{rj}\\ 0\\ \vdots\\ 0\end{pmatrix}$. 设$B$的首元列为第$j_1<j_2<\cdots<j_r$列, 则对$j=1,2,\ldots,n$有

$$\alpha_j=b_{1j}\alpha_{j_1}+b_{2j}\alpha_{j_2}+\cdots+b_{rj}\alpha_{j_r}=(\alpha_{j_1},\alpha_{j_2},\ldots,\alpha_{j_r})\begin{pmatrix}b_{1j}\\ b_{2j}\\ \vdots\\ b_{rj}\end{pmatrix}.$$

这得到

$$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_{j_1},\alpha_{j_2},\ldots,\alpha_{j_r})\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\ b_{21}&b_{22}&\cdots&b_{2n}\\ \vdots&\vdots&&\vdots\\ b_{r1}&b_{r2}&\cdots&b_{rn}\end{pmatrix}.$$

注意到上式最右边的两个矩阵的秩都是r, 所以A就写成了一个列满秩阵和一个行满秩阵的乘积. 这个结论就是所谓的满秩分解, 它是"秩为1的矩阵A为非零列向量和非零行向量的乘积(例5.19)"这一结论的一般化.

命题5.26. *设$A\in\mathbb{F}^{m\times n}$, $R(A)=r$, 则存在$H\in\mathbb{F}^{m\times r}$, $L\in\mathbb{F}^{r\times n}$, 使得$H$列满秩, L行满秩(即$R(H)=R(L)=r$), 且$A=HL$. 称这个为A的满秩分解.*

☞ **证明.** 也可以如同例5.19一样, 利用初等变换标准形证明. 存在m阶可逆阵P和n阶可逆阵Q, 使得

$$A=P\begin{pmatrix}E_r&O\\ O&O\end{pmatrix}_{m\times n}Q.$$

而

$$\begin{pmatrix}E_r&O\\ O&O\end{pmatrix}_{m\times n}=\begin{pmatrix}E_r\\ O\end{pmatrix}_{m\times r}(E_r,O)_{r\times n},$$

所以令$H=P\begin{pmatrix}E_r\\ O\end{pmatrix}_{m\times r}\in\mathbb{F}^{m\times r}$, $L=(E_r,O)_{r\times n}Q\in\mathbb{F}^{r\times n}$, 则有$A=HL$, 且

$$R(H)=R\begin{pmatrix}E_r\\ O\end{pmatrix}=r,\quad R(L)=R(E_r,O)=r.$$

证毕. □

很遗憾的是, 矩阵的满秩分解不唯一. 例如, 如果$A=HL$为A的一个满秩分解, 设$R(A)=r$, 则任取r阶可逆阵P, 有$A=(HP)(P^{-1}L)$也是A的满秩分解. 但我们总可以如下求出一个满秩分解

算法5.3. 矩阵的满秩分解的求法步骤: 设$A=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathbb{F}^{m\times n}$.

(1) $A\overset{r}{\sim}B$: 简化阶梯形阵;

(2) 设B的首元列为第$j_1<j_2<\cdots<j_r$列, 令$H=(\alpha_{j_1},\alpha_{j_2},\ldots,\alpha_{j_r})$;

(3) 令L为B的前r行所组成的矩阵(即去掉全零行后的矩阵), 则得到A的满秩分解$A=HL$.

例5.23. 设$A=\begin{pmatrix}1&1&2&2&1\\0&2&1&5&-1\\2&0&3&-1&3\\1&1&0&4&-1\end{pmatrix}$, 求$A$的一个满秩分解.

☞ **证明.** 由例5.22的计算, 得

$$A\overset{r}{\sim}\begin{pmatrix}1&0&0&1&0\\0&1&0&3&-1\\0&0&1&-1&1\\0&0&0&0&0\end{pmatrix}.$$

令$H=\begin{pmatrix}1&1&2\\0&2&1\\2&0&3\\1&1&0\end{pmatrix}$, $L=\begin{pmatrix}1&0&0&1&0\\0&1&0&3&-1\\0&0&1&-1&1\end{pmatrix}$, 则$A=HL$是$A$的一个满秩分解. □

如果矩阵A的秩比较小, 利用A的满秩分解, 有时可以简化问题.

例5.24. 计算A^n, 其中$A=\begin{pmatrix}2&-1&0\\2&-3&-4\\-1&2&3\end{pmatrix}$.

☞ **解.** 进行行初等变换

$$A\overset{r_1\leftrightarrow r_3}{\sim}\begin{pmatrix}-1&2&3\\2&-3&-4\\2&-1&0\end{pmatrix}\underset{r_2+2r_1}{\overset{r_3-r_2}{\sim}}\begin{pmatrix}-1&2&3\\0&1&2\\0&2&4\end{pmatrix}\underset{r_3-2r_2}{\overset{r_1-2r_2}{\sim}}\begin{pmatrix}-1&0&-1\\0&1&2\\0&0&0\end{pmatrix}\overset{-r_1}{\sim}\begin{pmatrix}1&0&1\\0&1&2\\0&0&0\end{pmatrix},$$

于是A有满秩分解$A=HL$, 其中

$$H=\begin{pmatrix}2&-1\\2&-3\\-1&2\end{pmatrix},\qquad L=\begin{pmatrix}1&0&1\\0&1&2\end{pmatrix}.$$

我们有

$$A^n=(HL)^n=H(LH)^{n-1}L.$$

由于

$$LH = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

所以

$$(LH)^{n-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{n-1} = \begin{pmatrix} 1 & n-1 \\ 0 & 1 \end{pmatrix}.$$

这得到

$$A^n = \begin{pmatrix} 2 & -1 \\ 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & n-1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2n-3 & 4n-4 \\ 2 & 2n-5 & 4n-8 \\ -1 & -n+3 & -2n+5 \end{pmatrix},$$

解毕. □

为什么我们要讲满秩分解呢? 当然是因为满秩阵有好的性质, 容易处理. 事实上, 满秩阵很像可逆阵, 只是满秩阵只有行(或列)满秩, 所以类似的性质可能只成立一半. 例如, 对于n阶方阵A, 有

$$A\text{可逆} \iff A\text{的行向量组线性无关} \iff A\text{的列向量组线性无关};$$

而当A为任意矩阵时, 有

$$A\text{行满秩} \iff A\text{的行向量组线性无关}, \quad A\text{列满秩} \iff A\text{的列向量组线性无关}.$$

可逆阵还成立如下性质: 设A是n阶方阵, 则

$$\begin{aligned} A\text{ 可逆} \iff & \quad \exists B \in \mathbb{F}^{n\times n}, \text{使得 } AB = E_n \\ \iff & \quad \exists B \in \mathbb{F}^{n\times n}, \text{使得 } BA = E_n \\ \iff & \quad A \overset{r}{\sim} E_n \\ \iff & \quad A \overset{c}{\sim} E_n. \end{aligned}$$

且当A可逆时, 有

$$\begin{aligned} & AB_1 = AB_2 \Longrightarrow B_1 = B_2, \\ & C_1A = C_2A \Longrightarrow C_1 = C_2, \\ & R(AB) = R(B), \quad R(CA) = R(C). \end{aligned}$$

作为满秩方阵的推广, 列满秩矩阵成立类似的性质(转置后读者应该可以写出行满秩阵的类似结论).

例5.25. 设$A \in \mathbb{F}^{m\times n}$, 则下面命题等价

(i) A列满秩(即$R(A) = n$);

(ii) 存在$B \in \mathbb{F}^{n\times m}$, 使得$BA = E_n$;

(iii) $A \overset{r}{\sim} \begin{pmatrix} E_n \\ O \end{pmatrix}$.

进而, 如果A列满秩, 则

(1) (左消去) 对任意$B,C\in\mathbb{F}^{n\times s}$, 若$AB=AC$, 那么$B=C$;

(2) (左乘不改变秩) 对任意$B\in\mathbb{F}^{n\times s}$, 有$R(AB)=R(B)$.

☞ **证明.** "(i) $\Longrightarrow$ (ii),(iii)" 存在m阶可逆阵P和n阶可逆阵Q, 使得$A=P\begin{pmatrix}E_n\\O\end{pmatrix}Q$. 取

$$B=Q^{-1}(E_n,O)_{n\times m}P^{-1}\in\mathbb{F}^{n\times m},$$

则有

$$BA=Q^{-1}(E_n,O)\begin{pmatrix}E_n\\O\end{pmatrix}Q=Q^{-1}E_nQ=E_n.$$

又由于

$$\begin{pmatrix}E_n\\O\end{pmatrix}Q=\begin{pmatrix}Q\\O\end{pmatrix}=\begin{pmatrix}Q&O\\O&E_{m-n}\end{pmatrix}\begin{pmatrix}E_n\\O\end{pmatrix},$$

所以取$P_1=P\begin{pmatrix}Q&O\\O&E_{m-n}\end{pmatrix}\in\mathbb{F}^{m\times m}$, 则$P_1$可逆, 且$A=P_1\begin{pmatrix}E_n\\O\end{pmatrix}$, 得$A\overset{r}{\sim}\begin{pmatrix}E_n\\O\end{pmatrix}$.

"(ii) $\Longrightarrow$ (i)" 由于$A\in\mathbb{F}^{m\times n}$, 所以$R(A)\leqslant n$. 另一方面, 如果存在矩阵$B$, 使得$BA=E_n$, 则

$$n=R(E_n)=R(BA)\leqslant R(A).$$

所以$R(A)=n$.

"(iii) $\Longrightarrow$ (i)" 由于$A\overset{r}{\sim}\begin{pmatrix}E_n\\O\end{pmatrix}$, 所以$R(A)=R\begin{pmatrix}E_n\\O\end{pmatrix}=n$.

"(ii) $\Longrightarrow$ (1)" 由(ii), 存在矩阵B_1, 使得$B_1A=E_n$. 于是

$$B=E_nB=(B_1A)B=B_1(AB)=B_1(AC)=(B_1A)C=E_nC=C.$$

"(ii) $\Longrightarrow$ (2)" 由(ii), 存在矩阵B_1, 使得$B_1A=E_n$. 于是

$$R(B)=R(E_nB)=R(B_1(AB))\leqslant R(AB).$$

又$R(AB)\leqslant R(B)$, 所以$R(B)=R(AB)$. □

由于列满秩矩阵左乘不改变秩, 所以如果$A=HL\in\mathbb{F}^{m\times n}$, 其中$H\in\mathbb{F}^{m\times r}$, $L\in\mathbb{F}^{r\times n}$, 且$R(H)=R(L)=r$, 则

$$R(A)=R(L)=r.$$

即满秩分解的逆命题也成立.

下面我们看应用矩阵的满秩分解的一个例子, 为此我们引入一个重要概念. 设$A=(a_{ij})\in\mathbb{F}^{n\times n}$, 定义$A$的迹$\mathrm{Tr}(A)$为

$$\mathrm{Tr}(A):=a_{11}+a_{22}+\cdots+a_{nn}\in\mathbb{F}.$$

于是我们得到迹函数

$$\mathrm{Tr}:\mathbb{F}^{n\times n}\longrightarrow\mathbb{F};A\mapsto\mathrm{Tr}(A).$$

矩阵的迹有性质: 对任意$A,B\in\mathbb{F}^{n\times n}$和$a,b\in\mathbb{F}$, 有

$$\mathrm{Tr}(aA+bB)=a\,\mathrm{Tr}(A)+b\,\mathrm{Tr}(B),\quad \mathrm{Tr}(AB)=\mathrm{Tr}(BA).$$

事实上, 设$A=(a_{ij})_{n\times n}$, $B=(b_{ij})_{n\times n}$, 则$aA+bB=(aa_{ij}+bb_{ij})$. 这得到

$$\operatorname{Tr}(aA+bB)=\sum_{i=1}^{n}(aa_{ii}+bb_{ii})=a\sum_{i=1}^{n}a_{ii}+b\sum_{i=1}^{n}b_{ii}=a\operatorname{Tr}(A)+b\operatorname{Tr}(B).$$

设$AB=(c_{ij})_{n\times n}$, $BA=(d_{ij})_{n\times n}$, 则有

$$c_{ii}=\sum_{k=1}^{n}a_{ik}b_{ki},\quad d_{ii}=\sum_{k=1}^{n}b_{ik}a_{ki}.$$

这得到

$$\begin{aligned}\operatorname{Tr}(AB)&=\sum_{i=1}^{n}c_{ii}=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ki},\\ \operatorname{Tr}(BA)&=\sum_{i=1}^{n}d_{ii}=\sum_{i=1}^{n}\sum_{k=1}^{n}b_{ik}a_{ki}=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ki},\end{aligned}$$

即有$\operatorname{Tr}(AB)=\operatorname{Tr}(BA)$.

类似地, 可以证明: 对于$A\in\mathbb{F}^{m\times n}$和$B\in\mathbb{F}^{n\times m}$(这里$m$和$n$不一定相等), 有$\operatorname{Tr}(AB)=\operatorname{Tr}(BA)$.

例5.26. *设$A\in\mathbb{F}^{n\times n}$是幂等阵, 即有$A^2=A$, 证明: $R(A)=\operatorname{Tr}(A)$.*

☞ **证明.** 设$R(A)=r$, 则A有满秩分解$A=HL$, 其中$H\in\mathbb{F}^{n\times r}$, $L\in\mathbb{F}^{r\times n}$, 且$R(H)=R(L)=r$. 由于$A^2=A$, 所以$HLHL=HL$. 由于$H$列满秩, 所以可左消去; 由于$L$行满秩, 所以可以右消去. 这得到$LH=E_r$, 进而

$$\operatorname{Tr}(A)=\operatorname{Tr}(HL)=\operatorname{Tr}(LH)=\operatorname{Tr}(E_r)=r=R(A).$$

上面使用矩阵的满秩分解的证明有一定的技巧性, 我们通常使用初等变换标准形来证明这个结论. 设$R(A)=r$, 则存在n阶可逆阵P和Q, 使得$A=P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q$. 由于$A^2=A$, 所以

$$P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}QP\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q=P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q,$$

即

$$\begin{pmatrix}E_r&O\\O&O\end{pmatrix}QP\begin{pmatrix}E_r&O\\O&O\end{pmatrix}=\begin{pmatrix}E_r&O\\O&O\end{pmatrix}.$$

做分块$QP=\begin{pmatrix}R_1&R_2\\R_3&R_4\end{pmatrix}$, 其中$R_1\in\mathbb{F}^{r\times r}$, 则

$$\begin{pmatrix}E_r&O\\O&O\end{pmatrix}=\begin{pmatrix}E_r&O\\O&O\end{pmatrix}\begin{pmatrix}R_1&R_2\\R_3&R_4\end{pmatrix}\begin{pmatrix}E_r&O\\O&O\end{pmatrix}=\begin{pmatrix}R_1&O\\O&O\end{pmatrix}.$$

这得到$R_1=E_r$, 进而

$$A=P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}(QP)P^{-1}=P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}\begin{pmatrix}E_r&R_2\\R_3&R_4\end{pmatrix}P^{-1}=P\begin{pmatrix}E_r&R_2\\O&O\end{pmatrix}P^{-1}.$$

所以有

$$\operatorname{Tr}(A)=\operatorname{Tr}\left(P\begin{pmatrix}E_r&R_2\\O&O\end{pmatrix}P^{-1}\right)=\operatorname{Tr}\left(P^{-1}P\begin{pmatrix}E_r&R_2\\O&O\end{pmatrix}\right)=\operatorname{Tr}\begin{pmatrix}E_r&R_2\\O&O\end{pmatrix}=r=R(A).$$

事实上, 上面证明了A和$\begin{pmatrix} E_r & R_2 \\ O & O \end{pmatrix}$相似[8]. 再利用

$$\begin{pmatrix} E_r & -R_2 \\ O & E_{n-r} \end{pmatrix}\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}\begin{pmatrix} E_r & R_2 \\ O & E_{n-r} \end{pmatrix}=\begin{pmatrix} E_r & R_2 \\ O & O \end{pmatrix},$$

可知幂等阵A与$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$相似[9], 其中$r=R(A)$. □

5.7.2 秩的一些关系式

我们已经知道下面关于秩的关系式成立:

- $0\leqslant R(A_{m\times n})\leqslant \min\{m,n\}$;
- $R(A^T)=R(A)$;
- $R(cA)=R(A)$, $(\forall c\neq 0)$;
- 如果B是A的子阵, 则$R(B)\leqslant R(A)$;
- 当P和Q是可逆方阵时, $R(PAQ)=R(A)$;
- $R(AB)\leqslant \min\{R(A),R(B)\}$.

下面证明更多的关于秩的关系式, 有些关系式显得很神奇和漂亮, 是前人努力的结果, 请读者欣赏其中的美. 我们还要说明的是, 由于矩阵的秩有几何和代数的不同定义(后面还有更几何的刻画—对应的线性映射的像空间的维数), 所以这些关系式通常有许多不同的证明, 请读者尝试给出自己的证明.

命题5.27. (1) 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{m\times s}$, 则$\max\{R(A),R(B)\}\leqslant R(A,B)\leqslant R(A)+R(B)$.

(2) 设$A,B\in\mathbb{F}^{m\times n}$, 则$R(A+B)\leqslant R(A)+R(B)$.

☞ **证明.** (1) 只需要证第二个不等号成立. 设$R(A)=r$, $R(B)=s$. 如果$r=0$, 则$A=O$. 于是

$$R(A,B)=R(B)=R(A)+R(B).$$

类似的, 当$s=0$时也有$R(A,B)=R(A)+R(B)$.

下设$r>0$, $s>0$. 取A的列向量组的极大无关组$\alpha_1,\dots,\alpha_r$, B的列向量组的极大无关组$\beta_1,\dots,\beta_s$, 则(A,B)的列向量组与$\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s$等价. 这得到

$$R(A,B)=R(\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s)\leqslant r+s=R(A)+R(B).$$

(2) 注意到分块矩阵的列变换

$$(A+B,B)\xrightarrow{c_1-c_2}(A,B),$$

[8]矩阵相似的概念参看第七章.

[9]在第七章我们将给出这一结论的另一个证明.

所以

$$R(A+B) \leqslant R(A+B,B) = R(A,B) \overset{(1)}{\leqslant} R(A)+R(B).$$

证毕. □

下面给出两矩阵乘积的秩的一个下界.

命题5.28 (Sylvester秩不等式). 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times s}$, 则

$$R(AB) \geqslant R(A)+R(B)-n.$$

特别地, 如果$AB=O$, 则有

$$R(A)+R(B) \leqslant n.$$

为了证明Sylvester秩不等式, 我们证明下面更一般的结论.

命题5.29 (Frobenius秩不等式). 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times s}$, $C \in \mathbb{F}^{s\times t}$, 则

$$R(ABC)+R(B) \geqslant R(AB)+R(BC).$$

☞ **Frobenius秩不等式$\Longrightarrow$ Sylvester秩不等式.** 由Frobenius秩不等式, 有

$$R(AB)+n = R(AE_nB)+R(E_n) \geqslant R(AE_n)+R(E_nB) = R(A)+R(B).$$

得证. □

下面的任务是证明Frobenius秩不等式. 我们采用分块矩阵的初等变换的技巧, 为此, 先考虑分块对角和分块(上)下三角阵的秩与对角块秩的和的关系.

引理5.30. 设A,B,C是矩阵, 则

(1) $R\begin{pmatrix} A & O \\ O & B \end{pmatrix} = R(A)+R(B)$;

(2) $R\begin{pmatrix} A & O \\ C & B \end{pmatrix} \geqslant R(A)+R(B)$.

☞ **证明.** 设$R(A)=r$, $R(B)=s$, 则存在可逆方阵P_1,Q_1,P_2,Q_2, 使得

$$P_1AQ_1 = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}, \qquad P_2BQ_2 = \begin{pmatrix} E_s & O \\ O & O \end{pmatrix}.$$

令$P=\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix}$, $Q=\begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix}$, 则$P$和$Q$是可逆方阵.

(1) 由于

$$P\begin{pmatrix} A & O \\ O & B \end{pmatrix}Q = \begin{pmatrix} P_1AQ_1 & O \\ O & P_2BQ_2 \end{pmatrix} = \begin{pmatrix} \begin{matrix} E_r & O \\ O & O \end{matrix} & O \\ O & \begin{matrix} E_s & O \\ O & O \end{matrix} \end{pmatrix} \sim \begin{pmatrix} E_r & & \\ & E_s & \\ & & O \end{pmatrix},$$

所以有

$$R\begin{pmatrix} A & O \\ O & B \end{pmatrix} = r + s = R(A) + R(B).$$

(2) 类似的, 由于

$$P\begin{pmatrix} A & O \\ C & B \end{pmatrix}Q = \begin{pmatrix} P_1AQ_1 & O \\ P_2CQ_1 & P_2BQ_2 \end{pmatrix} = \begin{pmatrix} E_r & O & O & O \\ O & O & O & O \\ C_{11} & C_{12} & E_s & O \\ C_{21} & C_{22} & O & O \end{pmatrix}$$

$$\sim \begin{pmatrix} E_r & O & O & O \\ O & O & O & O \\ O & O & E_s & O \\ O & C_{22} & O & O \end{pmatrix} \sim \begin{pmatrix} E_r & O & O & O \\ O & E_s & O & O \\ O & O & C_{22} & O \\ O & O & O & O \end{pmatrix},$$

所以$R\begin{pmatrix} A & O \\ C & B \end{pmatrix} \geqslant r + s = R(A) + R(B)$.

当然, 上面证明过程也可以都改写为分块阵的初等变换. □

引理5.30(2)有可能取等号, 例如有下面的例子.

例5.27. (1) 设A是可逆方阵, 则$R\begin{pmatrix} A & O \\ C & B \end{pmatrix} = R(A) + R(B)$.

(2) 设B是可逆方阵, 则$R\begin{pmatrix} A & O \\ C & B \end{pmatrix} = R(A) + R(B)$.

☞ **证明.** (1) 当A可逆时有

$$\begin{pmatrix} A & O \\ C & B \end{pmatrix} \underbrace{\quad}_{}^{r_2 - CA^{-1}r_1} \begin{pmatrix} A & O \\ O & B \end{pmatrix}.$$

再由引理5.30(1)得结论. □

为了体会一下用引理5.30证明Frobenius秩不等式的矩阵技巧, 我们先用它来重新证明命题5.27.

例5.28. 用引理5.30重新证明命题5.27.

☞ **证明.** (1) 由于

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix} \overset{r_1 + r_2}{\sim} \begin{pmatrix} A & B \\ O & B \end{pmatrix},$$

所以

$$R(A, B) \leqslant R\begin{pmatrix} A & B \\ O & B \end{pmatrix} = R\begin{pmatrix} A & O \\ O & B \end{pmatrix} \xlongequal{\text{引理}5.30} R(A) + R(B).$$

(2) 由于

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix} \overset{r_2 + r_1}{\sim} \begin{pmatrix} A & O \\ A & B \end{pmatrix} \overset{c_1 + c_2}{\sim} \begin{pmatrix} A & O \\ A + B & B \end{pmatrix},$$

所以

$$R(A + B) \leqslant \begin{pmatrix} A & O \\ A + B & B \end{pmatrix} = R\begin{pmatrix} A & O \\ O & B \end{pmatrix} \xlongequal{\text{引理}5.30} R(A) + R(B).$$

证毕. □

☞ **命题5.29(Frobenius秩不等式)的证明.** 我们有下面的分块矩阵的初等变换

$$\begin{pmatrix} AB & O \\ B & BC \end{pmatrix} \overset{r_1-Ar_2}{\rightsquigarrow} \begin{pmatrix} O & -ABC \\ B & BC \end{pmatrix}$$
$$\overset{c_2-c_1C}{\rightsquigarrow} \begin{pmatrix} O & -ABC \\ B & O \end{pmatrix} \overset{c_1\leftrightarrow c_2}{\rightsquigarrow} \begin{pmatrix} -ABC & O \\ O & B \end{pmatrix},$$

于是有

$$\begin{aligned} R(AB)+R(BC) &\overset{\text{引理}5.30}{\leqslant} R\begin{pmatrix} AB & O \\ B & BC \end{pmatrix} = R\begin{pmatrix} -ABC & O \\ O & B \end{pmatrix} \\ &\xlongequal{\text{引理}5.30} R(-ABC)+R(B) = R(ABC)+R(B), \end{aligned}$$

得证. □

Sylvester秩不等式有许多的证明方法. 除了可以利用Frobenius秩不等式证明外[10], 还可以用初等变换标准形如下证明.

☞ **命题5.28(Sylvester秩不等式)的另证.** 设$R(A)=r$, 则存在m阶可逆阵P和n阶可逆阵Q, 使得$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$. 做分块$QB=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$, 其中$B_1\in\mathbb{F}^{r\times r}$, 则

$$AB = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}QB = P\begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}.$$

所以

$$\begin{aligned} R(B) =& R(QB) = R\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \leqslant R(B_1,B_2)+R(B_3,B_4) \\ =& R\left(P\begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}\right)+R(B_3,B_4) = R(AB)+R(B_3,B_4) \\ \leqslant& R(AB)+(n-r), \end{aligned}$$

整理得到Sylvester秩不等式. □

我们后面还会给出Sylvester秩不等式的其它一些证明.

5.7.3 线性表示, 线性相关和矩阵的秩

本节最后, 我们将命题5.15用矩阵的秩重新表述一下.

定理5.31. 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{m\times l}$, $\beta\in\mathbb{F}^m$, 记S为A的列向量组, T为B的列向量组, 则

---

[10] 当然, 也可用上面给出的证明Frobenius秩不等式的方法, 比如利用如下的初等变换

$$\begin{pmatrix} AB & O \\ O & E_n \end{pmatrix} \overset{r_1+Ar_2}{\rightsquigarrow} \begin{pmatrix} AB & A \\ O & E_n \end{pmatrix} \overset{c_1-c_2B}{\rightsquigarrow} \begin{pmatrix} O & A \\ -B & E_n \end{pmatrix} \overset{c_1\leftrightarrow c_2}{\rightsquigarrow} \begin{pmatrix} A & O \\ E_n & -B \end{pmatrix}$$

来证明.

(1) β可由S线性表示$\Longleftrightarrow AX=\beta$有解$\Longleftrightarrow R(A)=R(A,\beta)$. 特别有

- **(i)** β不能由S线性表示$\Longleftrightarrow AX=\beta$无解$\Longleftrightarrow R(A)<R(A,\beta)$;
- **(ii)** β可由S唯一线性表示$\Longleftrightarrow AX=\beta$有唯一解$\Longleftrightarrow R(A)=R(A,\beta)=n$;
- **(iii)** β可由S不唯一地线性表示$\Longleftrightarrow AX=\beta$有无穷多解$\Longleftrightarrow R(A)=R(A,\beta)<n$;

(2) $T\leftarrow S \Longleftrightarrow AX=B$有解$\Longleftrightarrow R(A)=R(A,B)$;

(3) $S\leftrightarrow T \Longleftrightarrow AX=B$和$BY=A$都有解$\Longleftrightarrow R(A)=R(B)=R(A,B)$;

(4) S线性相关$\Longleftrightarrow AX=0$有非零解$\Longleftrightarrow R(A)<n$;

(5) S线性无关$\Longleftrightarrow AX=0$只有零解$\Longleftrightarrow R(A)=n$.

于是我们可以用矩阵的秩来判别向量组的线性相关性, 来判别线性方程组(矩阵方程)解的情况, 这种判别法是我们所喜欢的. 我们看几个例子.

例5.29. 设有向量$\alpha_1=\begin{pmatrix}1\\1-\lambda\\5-3\lambda\end{pmatrix}$, $\alpha_2=\begin{pmatrix}1\\1-\lambda\\1-\lambda\end{pmatrix}$, $\alpha_3=\begin{pmatrix}1-\lambda\\1\\1\end{pmatrix}$, $\beta=\begin{pmatrix}1\\1\\\lambda\end{pmatrix}$, 其中$\lambda$是参数. 试讨论向量$\beta$由向量组$\alpha_1,\alpha_2,\alpha_3$线性表示的情况. 当$\beta$可由向量组$\alpha_1,\alpha_2,\alpha_3$不唯一线性表示时, 求出一般的表达式.

☞ **解.** 设$A=(\alpha_1,\alpha_2,\alpha_3)$, 则问题等价于讨论线性方程组$AX=\beta$的解的情况. 对增广矩阵进行行初等变换

$$(A,\beta)=\begin{pmatrix}1&1&1-\lambda&1\\1-\lambda&1-\lambda&1&1\\5-3\lambda&1-\lambda&1&\lambda\end{pmatrix}\overset{r_2-(1-\lambda)r_1}{\underset{r_3-(5-3\lambda)r_1}{\sim}}\begin{pmatrix}1&1&1-\lambda&1\\0&0&\lambda(2-\lambda)&\lambda\\0&2(\lambda-2)&(2-3\lambda)(\lambda-2)&4\lambda-5\end{pmatrix}$$
$$\overset{r_2\leftrightarrow r_3}{\sim}\begin{pmatrix}1&1&1-\lambda&1\\0&2(\lambda-2)&(2-3\lambda)(\lambda-2)&4\lambda-5\\0&0&\lambda(2-\lambda)&\lambda\end{pmatrix}.$$

于是当$\lambda\neq0$且$\lambda\neq2$时有$R(A)=R(A,\beta)=3$, 得到线性方程组$AX=\beta$有唯一解. 所以当$\lambda\neq0$且$\lambda\neq2$时, β可由向量组$\alpha_1,\alpha_2,\alpha_3$唯一地线性表示.

当$\lambda=2$时, 有$R(A)=1$而$R(A,\beta)=2$, 得到线性方程组$AX=\beta$无解. 所以当$\lambda=2$时, β不能由向量组α_1, α_2, α_3线性表示.

当$\lambda=0$时, 有$R(A)=R(A,\beta)=2$, 得到线性方程组$AX=\beta$有无穷多解. 所以当$\lambda=0$时, β可由向量组α_1, α_2, α_3不唯一地线性表示. 此时, 由

$$(A,\beta)\overset{r}{\sim}\begin{pmatrix}1&1&1&1\\0&-4&-4&-5\\0&0&0&0\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&0&0&-\frac14\\0&1&1&\frac54\\0&0&0&0\end{pmatrix},$$

得到$AX=\beta$的通解为

$$X=\begin{pmatrix}-\frac14\\\frac54-c\\c\end{pmatrix},$$

其中c为任意常数. 于是
$$\beta=-\frac{1}{4}\alpha_1+\left(\frac{5}{4}-c\right)\alpha_2+c\alpha_3,$$
其中c为任意常数.

由于矩阵A是方阵, 也可先利用行列式如下讨论. 由于
$$|A|\xlongequal{r_1-r_2}\begin{vmatrix}\lambda&\lambda&-\lambda\\1-\lambda&1-\lambda&1\\5-3\lambda&1-\lambda&1\end{vmatrix}\xlongequal[c_3+c_1]{c_2-c_1}\begin{vmatrix}\lambda&0&0\\1-\lambda&0&2-\lambda\\5-3\lambda&2\lambda-4&6-3\lambda\end{vmatrix}=-2\lambda(2-\lambda)^2,$$
所以当$|A|\neq 0$, 即$\lambda\neq 0$且$\lambda\neq 2$时$AX=\beta$有唯一解. 于是当$\lambda\neq 0$且$\lambda\neq 2$时, β可由向量组$\alpha_1,\alpha_2,\alpha_3$唯一地线性表示.

再用行初等变换分别讨论$\lambda=2$和$\lambda=0$时系数矩阵和增广矩阵的秩的关系, 这里省略. □

例5.30 (例5.7). 设$\alpha_1,\alpha_2,\alpha_3\in\mathbb{F}^n$, 证明:

(1) $\alpha_1-\alpha_2$, $\alpha_2-\alpha_3$, $\alpha_3-\alpha_1$线性相关;

(2) $\alpha_1,\alpha_2,\alpha_3$线性无关$\Longleftrightarrow$ $\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$, $\alpha_3+\alpha_1$线性无关.

☞ **证明.** 前面给出了该例的几何证明, 这里给出代数证明.

(1) 由于
$$(\alpha_1-\alpha_2,\alpha_2-\alpha_3,\alpha_3-\alpha_1)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{pmatrix},$$
而
$$C=\begin{pmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{pmatrix}\xrightarrow{r_2+r_1}\begin{pmatrix}1&0&-1\\0&1&-1\\0&-1&1\end{pmatrix}\xrightarrow{r_3+r_2}\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&0\end{pmatrix},$$
所以$R(\alpha_1-\alpha_2,\alpha_2-\alpha_3,\alpha_3-\alpha_1)\leqslant R(C)=2$(也可计算$C$的行列式得到$R(C)<3$), 于是$\alpha_1-\alpha_2$, $\alpha_2-\alpha_3$, $\alpha_3-\alpha_1$线性相关.

(2) 由于
$$(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix},$$
而
$$P=\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix}\xrightarrow{r_2-r_1}\begin{pmatrix}1&0&1\\0&1&-1\\0&1&1\end{pmatrix}\xrightarrow{r_3-r_2}\begin{pmatrix}1&0&1\\0&1&-1\\0&0&2\end{pmatrix},$$
所以P可逆(也可计算P的行列式得到P可逆), 进而
$$R(\alpha_1,\alpha_2,\alpha_3)=R(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1).$$
由此得结论. □

例5.31. 设$\alpha_1=\begin{pmatrix}0\\1\\1\end{pmatrix}$, $\alpha_2=\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\beta_1=\begin{pmatrix}-1\\0\\1\end{pmatrix}$, $\beta_2=\begin{pmatrix}1\\2\\1\end{pmatrix}$, $\beta_3=\begin{pmatrix}3\\2\\-1\end{pmatrix}$, 证明: α_1,α_2和β_1, β_2, β_3等价.

☞ **证明.** 由于

$$(\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3)=\begin{pmatrix}0&1&-1&1&3\\1&1&0&2&2\\1&0&1&1&-1\end{pmatrix}\underset{r_2-r_1}{\overset{r_1\leftrightarrow r_3}{\sim}}\begin{pmatrix}1&0&1&1&-1\\0&1&-1&1&3\\0&1&-1&1&3\end{pmatrix}$$

$$\overset{r_3-r_2}{\sim}\begin{pmatrix}1&0&1&1&-1\\0&1&-1&1&3\\0&0&0&0&0\end{pmatrix},$$

所以$R(\alpha_1,\alpha_2)=R(\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3)=2$. 而

$$(\beta_1,\beta_2,\beta_3)\overset{r}{\sim}\begin{pmatrix}1&1&-1\\-1&1&3\\0&0&0\end{pmatrix}\overset{r_2+r_1}{\sim}\begin{pmatrix}1&1&-1\\0&2&2\\0&0&0\end{pmatrix},$$

所以$R(\beta_1,\beta_2,\beta_3)=2$. 于是$\alpha_1,\alpha_2$和$\beta_1,\beta_2,\beta_3$等价. □

例5.32 (例5.25中的(i) $\Longleftrightarrow$ (ii))**.** 设$A\in\mathbb{F}^{m\times n}$, 证明: $R(A)=n$当且仅当存在$B\in\mathbb{F}^{n\times m}$, 使得$BA=E_n$.

☞ **证明.** 由于$R(A^T)=R(A)$, 所以只要证: $R(A^T)=n\Longleftrightarrow$存在矩阵$C$, 使得$A^TC=E_n$. 考虑矩阵方程$A^TX=E_n$. 由于

$$R(A^T,E_n)\geqslant R(E_n)=n,$$

且(A^T,E_n)为$n\times(n+m)$矩阵推出$R(A^T,E_n)\leqslant n$. 所以得$R(A^T,E_n)=n$. 于是$A^TX=E_n$有解的充分必要条件是$R(A^T)=R(A^T,E_n)=n$. □

习题5.7

A1. 设矩阵$A=\begin{pmatrix}1&0&-1\\2&1&2\\0&1&4\end{pmatrix}$.

(1) 求A的一个满秩分解$A=HL$;

(2) 求矩阵H_1和L_1, 使得$H_1H=LL_1=E_2$;

(3) 求矩阵X, 使得$AXA=A$.

A2. 是否存在n阶方阵A和B, 使得

(1) $[A,B]=E_n$? (2) $[A,B]$可逆?

说明理由.

A3. 设有线性方程组$\begin{cases}(2-a)x_1+2x_2-2x_3=1,\\2x_1+(5-a)x_2-4x_3=2,\\-2x_1-4x_2+(5-a)x_3=-a-1,\end{cases}$ 问a取何值时, 该方程组有唯一解, 无解或有无穷多解? 并在有无穷多解时求其通解.

A4. 设矩阵$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{n\times m}$, 证明:

(1) $R(A)=R(AB,A)$;

(2) $R(AB)=R(A)$的充分必要条件是, 存在$C\in\mathbb{F}^{m\times n}$, 使得$A=ABC$.

A5. 设A是n阶方阵($n\geqslant 2$), 证明:

(1) $R(\mathrm{adj}(A))=\begin{cases} n, & \text{如果 } R(A)=n,\\ 1, & \text{如果 } R(A)=n-1,\\ 0, & \text{如果 } R(A)<n-1;\end{cases}$

(2) $\mathrm{adj}(\mathrm{adj}(A))=\begin{cases} A, & \text{如果 } n=2,\\ (\det A)^{n-2}A, & \text{如果 } n>2.\end{cases}$

A6. 设A是n阶方阵, 证明:

(1) 如果$A^2=E$, 则$R(A+E)+R(A-E)=n$;

(2) 如果$A^2=A$, 则$R(A)+R(A-E)=n$.

A7. 设n阶方阵A有分块$A=\begin{pmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}$, 其中$A_{11}$是$r$阶可逆矩阵, 证明: $R(A)=r$的充分必要条件是$A_{22}=A_{21}A_{11}^{-1}A_{12}$.

A8. 设A是n阶方阵, $k\in\mathbb{N}$满足$R(A^k)=R(A^{k+1})$, 证明:

$$R(A^k)=R(A^{k+1})=R(A^{k+2})=\cdots.$$

A9. 设k是正整数, $A_1,A_2,\ldots,A_k$是n阶方阵, 证明[11]:

$$R(A_1)+R(A_2)+\cdots+R(A_k)\leqslant R(A_1A_2\cdots A_k)+(k-1)n.$$

特别地, 当$A_1A_2\cdots A_k=O$时, 有$R(A_1)+R(A_2)+\cdots+R(A_k)\leqslant(k-1)n$.

B1. 设A和B为n阶方阵, 证明:

$$R(AB-E_n)\leqslant R(A-E_n)+R(B-E_n).$$

B2. 设k是正整数, $B_1,B_2,\ldots,B_k$都是n阶幂等阵. 记$A=B_1B_2\cdots B_k$, 证明:

$$R(E_n-A)\leqslant k(n-R(A)).$$

B3. 设A为n阶方阵, 证明:

(1) $R(A^2-E_n)=R(A+E_n)+R(A-E_n)-n$, 进而

$$A^2=E_n\Longleftrightarrow R(A+E_n)+R(A-E_n)=n;$$

(2) $R(A^2-A)=R(A)+R(A-E_n)-n$, 进而

$$A^2=A\Longleftrightarrow R(A)+R(A-E_n)=n.$$

B4. 设$A_1,A_2,\ldots,A_k$都是n阶方阵, 其中$k\geqslant 2$, 且满足$A_1+A_2+\cdots+A_k=E_n$. 证明: A_1, A_2, $\ldots$, A_k都是幂等阵的充分且必要条件是, $R(A_1)+R(A_2)+\cdots+R(A_k)=n$.

[11] 习题8.1中有一般化.

B5. 设A是n阶实对称阵, 证明: 如果$R(A) = r$, 则A至少有一个r阶主子式非零, 同时所有的r阶非零主子式都同号.
B6. 设A是n阶实反对称阵, 证明: $R(A)$一定是偶数, 并且如果$R(A) = r$, 则A至少有一个r阶主子式非零, 同时所有的r阶非零主子式都同号.
B7. 设A和B是n阶方阵, 满足$AB = BA = O$, 证明:
(1) 如果$R(A^2) = R(A)$, 那么$R(A + B) = R(A) + R(B)$;
(2) 存在正整数k, 使得$R(A^k + B^k) = R(A^k) + R(B^k)$.

§ 5.8 向量空间

本节数由向量构成的一类特殊的空间的向量个数, 这类空间我们称为向量空间. 顾名思义, 向量空间就是向量生活的空间, 向量在该空间中可以自由地进行一切活动. 而向量可以进行的活动就是加法和数乘, 可以自由地进行加法和数乘就是指进行加法和数乘后产生的向量仍在这个空间中, 不会出现奇怪的东西.

定义5.9. *设$V \subset \mathbb{F}^n$是非空子集, 如果V对向量的加法和数乘封闭, 即*

(i) *$\forall \alpha, \beta \in V$, 有$\alpha + \beta \in V$;*

(ii) *$\forall c \in \mathbb{F}$, $\forall \alpha \in V$, 有$c\alpha \in V$,*

*则称V是*向量空间.

由定义, 我们可以自由地在向量空间中进行向量的加法和数乘, 进而进行线性组合. 而且容易知道, 对向量的加法和数乘封闭等价于对向量的线性组合封闭: $\forall c_1, \ldots, c_s \in \mathbb{F}$, $\forall \alpha_1, \ldots, \alpha_s \in V$, 有

$$c_1\alpha_1 + \cdots + c_s\alpha_s \in V.$$

设V是向量空间, 任取$\alpha \in V$, 则$0 = 0 \cdot \alpha \in V$, 即向量空间中一定有零向量. 又$-\alpha = (-1) \cdot \alpha \in V$, 进而可知$V$对向量的减法封闭. 如果$V \neq \{0\}$, 任取非零的$\alpha \in V$, 则

$$\alpha, 2\alpha, 3\alpha, \ldots \in V$$

为无穷个互不相同的向量, 特别地V是无限集.

例5.33. **(1)** *$\{0\}$是向量空间, 称为*零向量空间*; $\mathbb{F}^n$也是向量空间, 称为*n维(列)向量空间*;*

(2) *$V = \{(0, x_2, \ldots, x_n)^T \mid x_2, \ldots, x_n \in \mathbb{F}\}$是向量空间;*

(3) *$W = \{(1, x_2, \ldots, x_n)^T \mid x_2, \ldots, x_n \in \mathbb{F}\}$不是向量空间(比如$0 \notin W$).*

例5.34. *设$\mathbb{F} = \mathbb{R}$, $n = 3$, 则$\mathbb{R}^3$中过原点的直线和过原点的平面都是向量空间. 除了这些和$\{0\}$以及$\mathbb{R}^3$外, $\mathbb{R}^3$中是否还有其它的向量空间?*

非零的向量空间都含有无限个向量, 但其所含向量的真正个数(秩)确是有限数. 由于其重要性, 我们对向量空间的秩和极大无关组给以新的名称: 维数和基. 而称向量在极大无关组下的表示系数为坐标. 具体地, 我们给出下面的定义.

定义5.10. 设V是向量空间.

(1) V的维数定义为

$$\dim V = \dim_{\mathbb{F}} V := R(V);$$

(2) 当$V \neq \{0\}$时, V的极大无关组称为V的基;

(3) 设$\dim V = r > 0$, $\alpha_1, \alpha_2, \ldots, \alpha_r$是$V$的一组(有序)基, 则任意$\alpha \in V$, α可唯一表示为

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r, \quad (c_1, c_2, \ldots, c_r \in \mathbb{F}).$$

称$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \end{pmatrix} \in \mathbb{F}^r$为$\alpha$在基$\alpha_1, \alpha_2, \ldots, \alpha_r$下的坐标.

于是, 这可以理解为在向量空间中建立坐标系, 即我们在做高维的解析几何. 基就是向量空间的坐标系, 基中每个向量代表一个坐标轴; 维数就是坐标系中坐标轴的个数; 而某个向量的坐标则是它在这个坐标系中的坐标. 设α_1, α_2, $\ldots$, α_r是向量空间V的一组基, 则有

$$V = \{c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r \mid c_1, c_2, \ldots, c_r \in \mathbb{F}\},$$

且对应

$$V \longrightarrow \mathbb{F}^r; \alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r \mapsto \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \end{pmatrix}$$

是双射. 于是V可等同于r元有序数组.

例5.35. **(1)** $\dim \mathbb{F}^n = n$, 有自然基: $e_1, e_2, \ldots, e_n$;

(2) 设$V = \{(0, x_2, \ldots, x_n)^T \mid x_2, \ldots, x_n \in \mathbb{F}\}$, 则$V$有基$e_2, e_3, \ldots, e_n$, 所以$\dim V = n - 1$.

例5.36. 在$\mathbb{R}^3$中,

(1) 设L为过原点的直线, 取定L上的一个非零向量α, 则L上任意向量可表示为$c\alpha$, 其中$c \in \mathbb{R}$. 于是α是L的基, $\dim L = 1$, 即直线是一维的;

(2) 设P为过原点的平面, 取定P上不共线的两个向量β, γ, 则P上任意向量可以表示为$a\beta + b\gamma$, 其中$a, b \in \mathbb{R}$. 于是β, γ是P的基, $\dim P = 2$, 即平面是二维的.

习题5.8

A1. 问下面的向量集合V是否是向量空间? 请给出理由. 如果V是向量空间, 求V的维数和V的一组基.

(1) $V=\{(x_1,x_2,\dots,x_n)^T \mid x_1,x_2,\dots,x_n\in\mathbb{F}\text{满足}\,x_1+x_2+\cdots+x_n=0\}$;

(2) $V=\{(x_1,x_2,\dots,x_n)^T \mid x_1,x_2,\dots,x_n\in\mathbb{F}\text{满足}\,x_1+x_2+\cdots+x_n=1\}$.

A2. 证明$\alpha_1=(1,-1,0)^T$, $\alpha_2=(2,1,3)^T$, $\alpha_3=(3,1,2)^T$是$\mathbb{F}^3$的一组基, 并求向量$\alpha=(5,0,7)^T$在这组基下的坐标.

§ 5.9 线性方程组解的结构

最后我们来数线性方程组的解集合所含向量的真正个数. 设有线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \qquad\qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m.\end{cases}$$

它的矩阵形式为$AX=\beta$, 其中

$$A=(a_{ij})_{m\times n},\quad X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \beta=\begin{pmatrix}b_1\\ \vdots\\ b_m\end{pmatrix}.$$

该线性方程组的解集合记为S, 即

$$S=\{\eta\in\mathbb{F}^n \mid A\eta=\beta\}.$$

本节的主要目的是研究S的结构, 数向量组S的真正数目, 即求S的秩. 我们分齐次和非齐次两种情形讨论.

5.9.1 齐次线性方程组解的结构

当$b_1=b_2=\cdots=b_m=0$时得到齐次线性方程组, 此时的矩阵形式为$AX=0$. 我们知道, 齐次线性方程组一定有零解$X=0$, 即解集合过原点. 于是, 从几何上看, 当$\mathbb{F}=\mathbb{R}$时, 三元线性方程组的解集合或者是原点, 或者是过原点的直线, 或者是过原点的平面, 或者是整个$\mathbb{R}^3$. 总之, 解集合为$\mathbb{R}^3$中的向量空间. 那么, 是否任意的齐次线性方程组的解集合都是向量空间呢? 我们来考察齐次线性方程组$AX=0$的解集合S. 如果$\alpha,\beta\in S$, 则$A\alpha=0$, $A\beta=0$. 这得到

$$A(\alpha+\beta)=A\alpha+A\beta=0+0=0,$$

即$\alpha+\beta\in S$. 又对$c\in\mathbb{F}$, 有

$$A(c\alpha)=cA\alpha=c0=0,$$

即$c\alpha\in S$. 所以S确实是向量空间.

命题5.32. 齐次线性方程组$AX=0$的解集合S关于向量的加法和数乘封闭, 即

(i) $\forall\alpha,\beta\in S\Longrightarrow\alpha+\beta\in S$;

(ii) $\forall\alpha\in S,\forall c\in\mathbb{F}\Longrightarrow c\alpha\in S$.

进而, S为向量空间, 称为$AX=0$的解空间.

定义5.11. 当$S\neq\{0\}$(即$AX=0$有非零解)时, S的基称为$AX=0$的基础解系.

于是, 要比较齐次线性方程组解的多少, 我们只要比较解空间维数的大小就好了. 那么解空间的维数如何计算呢? 例如, 3元齐次线性方程组

$$\begin{cases}x=0,\\ y=0\end{cases}$$

的解空间为

$$S=\{(0,0,z)^T\in\mathbb{F}^3\mid z\in\mathbb{F}\}.$$

于是$\dim S=1$. 我们看到原来向量$(x,y,z)^T\in\mathbb{F}^3$中的x,y,z可以任取, 有3个自由度; 当这个向量在S中时, 要满足两个方程, 这相当于给了两个约束条件, 于是解的自由度就成为$3-2=1$, 也即解空间维数为1. 用这种朴素的想法, n元齐次线性方程组$AX=0$的解空间S的维数应该等于原来的自由度n减去方程组给出的约束条件个数, 而约束条件的个数就是方程组中真正方程的个数, 即方程组的秩, 也就是$R(A)$. 最后我们发现, n元齐次线性方程组$AX=0$的解空间S的维数应该等于$n-R(A)$, 即解的自由度等于未知数个数与真正方程个数之差.

定理5.33. 设S是n元齐次线性方程组$AX=0$的解空间, 则

$$\dim S=n-R(A).$$

☞ **证明.** 如果$R(A)=n$, 则$AX=0$只有零解, 即$S=\{0\}$. 所以$\dim S=0=n-R(A)$.

下设$R(A)=r<n$. 设R是与A行等价的简化阶梯形阵, 则R有r个非零行. 必要时重新下标未知量, 可以不妨设

$$R=\begin{pmatrix}1&0&\cdots&0&c_{1,r+1}&\cdots&c_{1n}\\ 0&1&\cdots&0&c_{2,r+1}&\cdots&c_{2n}\\ \vdots&\vdots&\ddots&\vdots&\vdots&&\vdots\\ 0&0&\cdots&1&c_{r,r+1}&\cdots&c_{rn}\\ 0&0&\cdots&0&0&\cdots&0\\ \vdots&\vdots&&\vdots&\vdots&&\vdots\\ 0&0&\cdots&0&0&\cdots&0\end{pmatrix}.$$

这得到同解方程组

$$(\clubsuit)\quad\begin{cases}x_1=-c_{1,r+1}x_{r+1}-c_{1,r+2}x_{r+2}-\cdots-c_{1n}x_n,\\ x_2=-c_{2,r+1}x_{r+1}-c_{2,r+2}x_{r+2}-\cdots-c_{2n}x_n,\\ \qquad\vdots\\ x_r=-c_{r,r+1}x_{r+1}-c_{r,r+2}x_{r+2}-\cdots-c_{rn}x_n.\end{cases}$$

取定$\mathbb{F}^{n-r}$的一组基$\beta_1, \beta_2, \ldots, \beta_{n-r}$. 令自由未知数取向量$\begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \beta_1$, 代入方程组(♣)得非自由未知数向量$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \alpha_1 \in \mathbb{F}^r$, 进而得$\eta_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \in S$. 类似的, 对$i = 1, 2, \ldots, n-r$, 令$\begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \beta_i$, 代入方程组(♣)得$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \alpha_i \in \mathbb{F}^r$, 进而得$\eta_i = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \in S$. 下证: $\eta_1, \eta_2, \ldots, \eta_{n-r}$是$AX = 0$的一个基础解系.

首先, 由于$\beta_1, \beta_2, \ldots, \beta_{n-r}$线性无关, 所以$\eta_1, \eta_2, \ldots, \eta_{n-r}$线性无关. 其次, 设$\eta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in S$, 其中$\alpha \in \mathbb{F}^r$, $\beta \in \mathbb{F}^{n-r}$, 则存在$a_1, a_2, \ldots, a_{n-r} \in \mathbb{F}$, 使得

$$\beta = a_1\beta_1 + a_2\beta_2 + \cdots + a_{n-r}\beta_{n-r}.$$

于是

$$\eta - (a_1\eta_1 + a_2\eta_2 + \cdots + a_{n-r}\eta_{n-r}) = \begin{pmatrix} \alpha' \\ 0 \end{pmatrix} \in S,$$

其中$\alpha' = \alpha - (a_1\alpha_1 + a_2\alpha_2 + \cdots + a_{n-r}\alpha_{n-r})$. 代入方程组(♣)就得$\alpha' = 0$, 所以

$$\eta = a_1\eta_1 + a_2\eta_2 + \cdots + a_{n-r}\eta_{n-r}.$$

这就证明了$\eta_1, \eta_2, \ldots, \eta_{n-r}$是$S$的基, 是$AX = 0$的一个基础解系. 于是

$$\dim S = n - r = n - R(A).$$

通常取$\beta_1, \beta_2, \ldots, \beta_{n-r}$为$\mathbb{F}^{n-r}$的自然基$e_1, e_2, \ldots, e_{n-r}$. 此时, 对$i = 1, 2, \ldots, n-r$, 令$\begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = e_i$, 得$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} -c_{1,r+i} \\ -c_{2,r+i} \\ \vdots \\ -c_{r,r+i} \end{pmatrix} \in \mathbb{F}^r$. □

上面的证明过程事实上给出了求齐次线性方程组$AX = 0$的基础解系的方法.

算法5.4. *n元齐次线性方程组$AX = 0$的基础解系的求解步骤:*

(1) $A \overset{r}{\sim} R$: *简化阶梯形阵;*

(2) *写出同解方程组;*

(3) 设R的首元列为第$j_1 < j_2 < \cdots < j_r$列, 非首元列为第$j_{r+1} < \cdots < j_n$列, 令自由未知数向量$\begin{pmatrix} x_{j_{r+1}} \\ \vdots \\ x_{j_n} \end{pmatrix}$分别取$\mathbb{F}^{n-r}$的某一组基中的$n-r$个向量(通常取自然基$e_1, e_2, \ldots, e_{n-r}$), 代入同解方程组得到非自由未知数$x_{j_1}, \ldots, x_{j_r}$, 从而得到解$\eta_1, \eta_2, \ldots, \eta_{n-r}$, 这就是$AX=0$ 的基础解系.

如果$\eta_1, \eta_2, \ldots, \eta_{n-r}$是$AX=0$的一个基础解系, 则

$$S = \{c_1\eta_1 + c_2\eta_2 + \cdots + c_{n-r}\eta_{n-r} \mid c_1, c_2, \ldots, c_{n-r} \in \mathbb{F}\}.$$

于是$AX=0$的通解是

$$X = c_1\eta_1 + c_2\eta_2 + \cdots + c_{n-r}\eta_{n-r}, \quad (c_1, c_2, \ldots, c_{n-r} \in \mathbb{F}).$$

由于基础解系不唯一, 所以通解形式也不唯一. 反之, 如果n元线性方程组$AX=0$的通解是

$$X = c_1\eta_1 + c_2\eta_2 + \cdots + c_{n-r}\eta_{n-r}, \quad (c_1, c_2, \ldots, c_{n-r} \in \mathbb{F}).$$

其中$R(A)=r$, 则$\eta_1, \eta_2, \ldots, \eta_{n-r}$是$AX=0$的一个基础解系. 另一方面, n元线性方程组$AX=0$的任意$n-R(A)$个线性无关的解是$AX=0$的一个基础解系.

例5.37. 求解方程组$\begin{cases} x_1+x_2+x_3+x_4=0, \\ x_1+x_2+2x_3-x_4=0, \\ x_1+x_2-x_3+5x_4=0, \\ 3x_1+3x_2+2x_3+5x_4=0. \end{cases}$

☞ **解.** 由于系数矩阵

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & -1 \\ 1 & 1 & -1 & 5 \\ 3 & 3 & 2 & 5 \end{pmatrix} \underbrace{\overset{\substack{r_2-r_1 \\ r_3-r_1}}{}}_{r_4-3r_1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & -1 & 2 \end{pmatrix} \underbrace{\overset{\substack{r_1-r_2 \\ r_3+2r_2}}{}}_{r_4+r_2} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

得同解方程组$\begin{cases} x_1+x_2+3x_4=0, \\ x_3-2x_4=0, \end{cases}$ 其中x_2, x_4是自由未知数.

一方面, 可以令$x_2=c_1$, $x_4=c_2$得通解

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -c_1-3c_2 \\ c_1 \\ 2c_2 \\ c_2 \end{pmatrix} = c_1\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} -3 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \quad (c_1, c_2 \in \mathbb{F}).$$

另一方面, 也可以分别令$\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, 得到$\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix}$. 于是有基础解系$\eta_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$,

$\eta_2 = \begin{pmatrix} -3 \\ 0 \\ 2 \\ 1 \end{pmatrix}$, 进而通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = c_1\eta_1 + c_2\eta_2, \quad (c_1, c_2 \in \mathbb{F}).$$

解毕. □

可以利用定理5.33来证明关于矩阵秩的一些关系式, 例如我们可以重新证明Sylvester秩不等式(命题5.28).

例5.38 (Sylvester秩不等式)**.** 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times s}$, 证明:

$$R(AB) \geqslant R(A) + R(B) - n.$$

证明. 记$AX = 0$的解空间为$V_A \subset \mathbb{F}^n$, $BX = 0$的解空间为$V_B \subset \mathbb{F}^s$, $ABX = 0$的解空间为$V_{AB} \subset \mathbb{F}^s$, 则有$V_B \subset V_{AB}$. 再定义

$$W = \{BX \mid X \in V_{AB}\} \subset \mathbb{F}^n,$$

则W是向量空间, 且$W \subset V_A$. 特别有$\dim W \leqslant \dim V_A$.

取V_B的基$\alpha_1, \dots, \alpha_r$, 其中$r = \dim V_B$, 将其扩充为$V_{AB}$的基$\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_t$, 其中$t = \dim V_{AB}$. 我们证明$B\alpha_{r+1}, \dots, B\alpha_t$是$W$的一组基.

事实上, 设

$$c_{r+1}B\alpha_{r+1} + \cdots + c_tB\alpha_t = 0, \quad c_{r+1}, \dots, c_t \in \mathbb{F},$$

则

$$B(c_{r+1}\alpha_{r+1} + \cdots + c_t\alpha_t) = 0.$$

这得到$c_{r+1}\alpha_{r+1} + \cdots + c_t\alpha_t \in V_B$, 进而

$$c_{r+1}\alpha_{r+1} + \cdots + c_t\alpha_t = c_1\alpha_1 + \cdots + c_r\alpha_r, \quad \exists c_1, \dots, c_r \in \mathbb{F}.$$

利用$\alpha_1, \alpha_2, \dots, \alpha_t$线性无关, 就得到$c_1 = c_2 = \cdots = c_t = 0$. 特别有$B\alpha_{r+1}, \dots, B\alpha_t$线性无关.

再任取$BX \in W$, 其中$X \in V_{AB}$. 于是

$$X = x_1\alpha_1 + \cdots + x_r\alpha_r + x_{r+1}\alpha_{r+1} + \cdots + x_t\alpha_t, \quad \exists x_i \in \mathbb{F},$$

进而

$$BX = x_1B\alpha_1 + \cdots + x_rB\alpha_r + x_{r+1}B\alpha_{r+1} + \cdots + x_tB\alpha_t = x_{r+1}B\alpha_{r+1} + \cdots + x_tB\alpha_t.$$

这就证明了$B\alpha_{r+1}, \dots, B\alpha_t$是$W$的一组基.

所以我们有

$$\dim V_{AB} = \dim V_B + \dim W \leqslant \dim V_B + \dim V_A.$$

代入解空间维数公式

$$\dim V_A = n - R(A), \quad \dim V_B = s - R(B), \quad \dim V_{AB} = s - R(AB),$$

就得到所需证明的Sylvester秩不等式. □

5.9.2 非齐次线性方程组解的结构

当线性方程组的常数项$b_1, b_2, \ldots, b_m$不全为零时, 我们得到非齐次线性方程组, 其矩阵形式是$AX = \beta$, 其中$\beta \neq 0$. 我们假设这个线性方程组有解, 其解集合记为S. 由于$0 \notin S$, 所以S不是向量空间. 那么如何研究S呢?

设$\mathbb{F} = \mathbb{R}$, 考查二元线性方程组

$$x - y = -1.$$

几何上它是平面上的一条直线, 并且不过原点. 我们把常数项改成零, 得到齐次线性方程组

$$x - y = 0.$$

几何上, 上面的齐次线性方程组是过原点的直线, 而非齐次线性方程组对应的直线可以由齐次的直线平移得到, 参看图5.5.

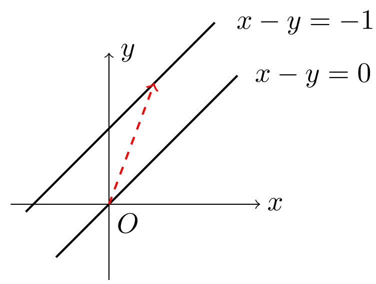


图 5.5: 非齐次线性方程组的解集和导出组的解空间的关系

那么这是一般现象吗? 对非齐次线性方程组$AX = \beta$, 称齐次线性方程组$AX = 0$为该线性方程组的导出组, 它的解空间记为S'. 我们看S和S'的关系. 设$\gamma_1, \gamma_2 \in S$, 则$A\gamma_1 = \beta$, $A\gamma_2 = \beta$. 于是

$$A(\gamma_1 - \gamma_2) = A\gamma_1 - A\gamma_2 = \beta - \beta = 0,$$

即有$\gamma_1 - \gamma_2 \in S'$, 非齐次的两个解之差为齐次的解. 又设$\gamma \in S$, $\eta \in S'$, 则$A\gamma = \beta$, $A\eta = 0$. 于是

$$A(\gamma + \eta) = A\gamma + A\eta = \beta + 0 = \beta,$$

于是$\gamma + \eta \in S$, 非齐次和齐次的解之和为非齐次的解.

命题5.34. **(1)** $\forall \gamma_1, \gamma_2 \in S \Longrightarrow \gamma_1 - \gamma_2 \in S'$;

(2) $\forall \gamma \in S, \forall \eta \in S' \Longrightarrow \gamma + \eta \in S$.

由此立刻可得

定理5.35. 设S为n元线性方程组$AX = \beta$的解集合且$S \neq \varnothing$, 令S'为导出组$AX = 0$的解空间.

(1) 任取$\gamma_0 \in S$(称γ_0为特解), 则

$$S = \gamma_0 + S' = \{\gamma_0 + \eta \mid \eta \in S'\};$$

(2) 若导出组$AX=0$的基础解系为$\eta_1,\eta_2,\dots,\eta_{n-r}$, γ_0为$AX=\beta$的特解, 则$AX=\beta$的通解是

$$X=\gamma_0+c_1\eta_1+c_2\eta_2+\cdots+c_{n-r}\eta_{n-r},\quad (c_1,c_2,\dots,c_{n-r}\in\mathbb{F}).$$

在定理5.35中, 我们并未要求$\beta\neq 0$, 它对齐次和非齐次线性方程组都成立. 当$\beta\neq 0$时, 定理5.35表明非齐次线性方程组的解集合是它的导出组的平移, 正如图5.5所显示.

☞ **定理5.35的证明.** (1) 任意的$\gamma\in S$, 有$\gamma-\gamma_0\in S'$, 即存在$\eta\in S'$, 使得$\gamma-\gamma_0=\eta$. 这得到

$$\gamma=\gamma_0+\eta\in\gamma_0+S'.$$

反之, 任意的$\gamma_0+\eta\in\gamma_0+S'$, 有$\gamma_0+\eta\in S$. 所以$S=\gamma_0+S'$.

(2) 由(1). □

可以用(非齐次)线性方程组解的结构证明关于矩阵秩的一些关系式, 例如我们可以如下再次证明Sylvester秩不等式.

例5.39 (Sylvester秩不等式)**.** 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{n\times s}$, 证明:

$$R(AB)\geqslant R(A)+R(B)-n.$$

☞ **证明.** 设$R(A)=r$, 取$AX=0$的基础解系$\eta_1,\eta_2,\dots,\eta_{n-r}$. 记$B=(\beta_1,\beta_2,\dots,\beta_s)$, 其中$\beta_i\in\mathbb{F}^n$, 则

$$AB=(A\beta_1,A\beta_2,\dots,A\beta_s).$$

设AB的列向量组的一个极大无关组为$A\beta_{i_1},A\beta_{i_2},\dots,A\beta_{i_t}$, 这里$t=R(AB)$.

取定整数i, 使得$1\leqslant i\leqslant s$, 考察线性方程组$AX=A\beta_i$. 由于存在$a_1,a_2,\dots,a_t\in\mathbb{F}$, 使得

$$A\beta_i=a_1A\beta_{i_1}+a_2A\beta_{i_2}+\cdots+a_tA\beta_{i_t}=A(a_1\beta_{i_1}+a_2\beta_{i_2}+\cdots+a_t\beta_{i_t}),$$

所以$a_1\beta_{i_1}+a_2\beta_{i_2}+\cdots+a_t\beta_{i_t}$为$AX=A\beta_i$的一个特解, 进而$AX=A\beta_i$的通解为

$$X=a_1\beta_{i_1}+a_2\beta_{i_2}+\cdots+a_t\beta_{i_t}+c_1\eta_1+c_2\eta_2+\cdots+c_{n-r}\eta_{n-r},\quad c_1,c_2,\dots,c_{n-r}\in\mathbb{F}.$$

而β_i为$AX=A\beta_i$的一个解, 所以存在$c_1,c_2,\dots,c_{n-r}\in\mathbb{F}$, 使得

$$\beta_i=a_1\beta_{i_1}+a_2\beta_{i_2}+\cdots+a_t\beta_{i_t}+c_1\eta_1+c_2\eta_2+\cdots+c_{n-r}\eta_{n-r}.$$

这就证明了β_i可由$\beta_{i_1},\dots,\beta_{i_t},\eta_1,\dots,\eta_{n-r}$线性表示.

所以, B的列向量组可由$\beta_{i_1},\dots,\beta_{i_t},\eta_1,\dots,\eta_{n-r}$线性表示, 进而得

$$R(B)\leqslant R(\beta_{i_1},\dots,\beta_{i_t},\eta_1,\dots,\eta_{n-r})\leqslant t+n-r,$$

此即Sylvester秩不等式. □

作为向量组, 非齐次线性方程组的解集合S的秩是多少呢? 当$AX=0$只有零解时, 如果$AX=\beta$有解, 则有唯一解, 于是$R(S)=1=\dim S'+1$. 这是一般规律, 即有下面结论.

命题5.36. 设S是n元非齐次线性方程组$AX=\beta$的解集合, 则当$S\neq\varnothing$时,

(1) $R(S)=n-R(A)+1$;

(2) 对任意线性无关的$\gamma_1, \gamma_2, \dots, \gamma_{n-r+1} \in S$, 其中$r = R(A)$, 有$AX = \beta$的通解是

$$X = c_1\gamma_1 + c_2\gamma_2 + \cdots + c_{n-r+1}\gamma_{n-r+1},$$

其中$c_1, c_2, \dots, c_{n-r+1} \in \mathbb{F}$, 满足$c_1 + c_2 + \cdots + c_{n-r+1} = 1$.

☞ **证明.** (1) 记$R(A) = r$, 不妨设$r < n$. 取$AX = 0$的基础解系$\eta_1, \eta_2, \dots, \eta_{n-r}$和$AX = \beta$的特解$\gamma_0$. 于是$\gamma_0 + \eta_i \in S$. 下面证明: γ_0, $\gamma_0 + \eta_1$, $\gamma_0 + \eta_2$, ..., $\gamma_0 + \eta_{n-r}$是S的极大无关组.

事实上, 如果

$$c\gamma_0 + c_1(\gamma_0 + \eta_1) + c_2(\gamma_0 + \eta_2) + \cdots + c_{n-r}(\gamma_0 + \eta_{n-r}) = 0,$$

则

$$(c + c_1 + c_2 + \cdots + c_{n-r})\gamma_0 + c_1\eta_1 + c_2\eta_2 + \cdots + c_{n-r}\eta_{n-r} = 0.$$

两边同时用A左乘, 利用$A\gamma_0 = \beta$, $A\eta_i = 0$得

$$(c + c_1 + c_2 + \cdots + c_{n-r})\beta = 0.$$

而$\beta \neq 0$, 所以有

$$c + c_1 + c_2 + \cdots + c_{n-r} = 0.$$

这得到

$$c_1\eta_1 + c_2\eta_2 + \cdots + c_{n-r}\eta_{n-r} = 0.$$

因为$\eta_1, \eta_2, \dots, \eta_{n-r}$线性无关, 所以有$c_1 = c_2 = \cdots = c_{n-r} = 0$. 进而$c = 0$, 即有$\gamma_0$, $\gamma_0+\eta_1$, $\gamma_0+\eta_2$, ..., $\gamma_0+\eta_{n-r}$线性无关.

另一方面, 任意$\gamma \in S$, 有$a_1, a_2, \dots, a_{n-r} \in \mathbb{F}$, 使得

$$\gamma = \gamma_0 + a_1\eta_1 + a_2\eta_2 + \cdots + a_{n-r}\eta_{n-r}.$$

这得到

$$\gamma = (1 - a_1 - a_2 - \cdots - a_{n-r})\gamma_0 + a_1(\gamma_0 + \eta_1) + a_2(\gamma_0 + \eta_2) + \cdots + a_{n-r}(\gamma_0 + \eta_{n-r}).$$

由此得γ_0, $\gamma_0 + \eta_1$, $\gamma_0 + \eta_2$, ..., $\gamma_0 + \eta_{n-r}$是S的极大无关组, 进而$R(S) = n - r + 1$.

(2) 由(1), $\gamma_1, \gamma_2, \dots, \gamma_{n-r+1}$是$S$的极大无关组, 所以任意$\gamma \in S$有

$$\gamma = c_1\gamma_1 + c_2\gamma_2 + \cdots + c_{n-r+1}\gamma_{n-r+1}, \quad (c_1, c_2, \dots, c_{n-r+1} \in \mathbb{F}).$$

于是

$$\beta = A\gamma = c_1A\gamma_1 + c_2A\gamma_2 + \cdots + c_{n-r+1}A\gamma_{n-r+1} = (c_1 + c_2 + \cdots + c_{n-r+1})\beta,$$

这得到$c_1 + c_2 + \cdots + c_{n-r+1} = 1$.

反之, 如果$c_1 + c_2 + \cdots + c_{n-r+1} = 1$, 则

$$\begin{aligned} & A(c_1\gamma_1 + c_2\gamma_2 + \cdots + c_{n-r+1}\gamma_{n-r+1}) \\ =& c_1A\gamma_1 + c_2A\gamma_2 + \cdots + c_{n-r+1}A\gamma_{n-r+1} \\ =& (c_1 + c_2 + \cdots + c_{n-r+1})\beta = \beta. \end{aligned}$$

证毕. □

上面关于线性方程组解的结构的讨论可以归结于下表.

n元线性方程组	$AX=0$	$AX=\beta\ (\beta\neq 0)$
S是否向量空间	是	否
S或通解	基础解系	特解+导出组的基础解系
$R(S)$	$n-R(A)$	$n-R(A)+1$

例5.40. 解方程组$\begin{cases} x_1+x_2-3x_3-x_4=1, \\ 3x_1-x_2-3x_3+4x_4=4, \\ x_1+5x_2-9x_3-8x_4=0. \end{cases}$

解. 由于增广矩阵

$$\widetilde{A}=\begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 3 & -1 & -3 & 4 & 4 \\ 1 & 5 & -9 & -8 & 0 \end{pmatrix} \underset{r_3-r_1}{\overset{r_2-3r_1}{\sim}} \begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 0 & -4 & 6 & 7 & 1 \\ 0 & 4 & -6 & -7 & -1 \end{pmatrix}$$

$$\underset{-\frac{1}{4}r_2}{\overset{r_3+r_2}{\sim}} \begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{7}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{r_1-r_2}{\sim} \begin{pmatrix} 1 & 0 & -\frac{3}{2} & \frac{3}{4} & \frac{5}{4} \\ 0 & 1 & -\frac{3}{2} & -\frac{7}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

得同解方程组$\begin{cases} x_1-\frac{3}{2}x_3+\frac{3}{4}x_4=\frac{5}{4}, \\ x_2-\frac{3}{2}x_3-\frac{7}{4}x_4=-\frac{1}{4}. \end{cases}$ 于是x_3, x_4是自由未知数. 令$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, 得$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} \frac{5}{4} \\ -\frac{1}{4} \end{pmatrix}$, 即有特解$\gamma_0=\begin{pmatrix} \frac{5}{4} \\ -\frac{1}{4} \\ 0 \\ 0 \end{pmatrix}$.

又导出组的同解方程组是$\begin{cases} x_1-\frac{3}{2}x_3+\frac{3}{4}x_4=0, \\ x_2-\frac{3}{2}x_3-\frac{7}{4}x_4=0. \end{cases}$

令$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}$, 得$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 3 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 7 \end{pmatrix}$. 于是得到基础解系$\eta_1=\begin{pmatrix} 3 \\ 3 \\ 2 \\ 0 \end{pmatrix}$, $\eta_2=\begin{pmatrix} -3 \\ 7 \\ 0 \\ 4 \end{pmatrix}$. 最后, 原方程组的通解是

$$X=\gamma_0+c_1\eta_1+c_2\eta_2, \quad (c_1, c_2\in\mathbb{F}).$$

解毕. □

例5.41. 设$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta\in\mathbb{F}^n$满足: $\alpha_2, \alpha_3, \alpha_4$线性无关, $\alpha_1=2\alpha_2-\alpha_3$, $\beta=\alpha_1+\alpha_2+\alpha_3+\alpha_4$, 令$A=(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 求线性方程组$AX=\beta$的通解.

解. 因为$\alpha_2, \alpha_3, \alpha_4$线性无关, $\alpha_1=2\alpha_2-\alpha_3$, 所以$\alpha_2, \alpha_3, \alpha_4$是$\alpha_1, \alpha_2, \alpha_3, \alpha_4$的极大无关组. 这得到$R(A)=3$, 导出组$AX=0$的基础解系含一个向量.

由于$\alpha_1=2\alpha_2-\alpha_3$, 所以

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}=0.$$

这得到$\eta = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}$是导出组的基础解系.

由于$\beta = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4$, 所以

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \beta.$$

即$\gamma_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$是特解.

最后得$AX = \beta$的通解是

$$X = \gamma_0 + k\eta = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + c \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \qquad (c \in \mathbb{F}).$$

解毕. □

习题5.9

A1. 求下列齐次线性方程组的一个基础解系:

(1) $\begin{cases} x_1 - 8x_2 + 10x_3 + 2x_4 = 0, \\ 2x_1 + 4x_2 + 5x_3 - x_4 = 0, \\ 3x_1 + 8x_2 + 6x_3 - 2x_4 = 0; \end{cases}$ (2) $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0, \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 0, \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 0, \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 0. \end{cases}$

A2. 解下列线性方程组:

(1) $\begin{cases} x_1 + x_2 = 5, \\ 2x_1 + x_2 + x_3 + 2x_4 = 1, \\ 5x_1 + 3x_2 + 2x_3 + 2x_4 = 3; \end{cases}$ (2) $\begin{cases} x_1 - 5x_2 + 2x_3 - 3x_4 = 11, \\ 5x_1 + 3x_2 + 6x_3 - x_4 = -1, \\ 2x_1 + 4x_2 + 2x_3 + x_4 = -6. \end{cases}$

A3. 设A是$m \times n$矩阵, α_i是A的第i个行向量, $\beta = (b_1, b_2, \ldots, b_n)$. 如果齐次线性方程组$AX = 0$的解全是方程$b_1x_1 + b_2x_2 + \cdots + b_nx_n = 0$的解, 证明: β是$\alpha_1, \alpha_2, \ldots, \alpha_m$的线性组合.

A4. 设四元非齐次线性方程组的系数矩阵的秩为3, 已知$\gamma_1, \gamma_2, \gamma_3$是它的三个解向量, 且

$$\gamma_1 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad \gamma_2 + \gamma_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix},$$

求该方程组的通解.

A5. 设$\alpha_1,\alpha_2,\alpha_3,\alpha_4$是非齐次线性方程组$AX=\beta$的4个解, 且

$$(\alpha_1+2\alpha_2,\alpha_2-2\alpha_3,3\alpha_3+2\alpha_4)=\begin{pmatrix}1&1&0\\1&0&1\\0&1&1\end{pmatrix},$$

求该线性方程组$AX=\beta$的通解.

A6. 设n阶方阵A的(i,j)位置的代数余子式为A_{ij}, 证明: 如果A不可逆且$A_{11}\neq 0$, 则齐次线性方程组$AX=0$的所有解有形式

$$c\begin{pmatrix}A_{11}\\A_{12}\\\vdots\\A_{1n}\end{pmatrix}.$$

A7. 设n阶方阵A满足$R(A)=R(A^2)$, 证明: 对任意正整数k, 有$R(A)=R(A^k)$.

A8. 设A是复矩阵, 证明: $R(AA^*)=R(A^*A)=R(A)$, 其中$A^*=\overline{A}^T$是A的共轭转置. 特别的, 当A是实矩阵时, 成立$R(AA^T)=R(A^TA)=R(A)$.

B1. 设$A\in\mathbb{C}^{m\times n}$, 证明: 对任意$\beta\in\mathbb{C}^m$, 线性方程组$A^*AX=A^*\beta$一定有解. 特别的, 当$A\in\mathbb{R}^{m\times n}$时, 对任意$\beta\in\mathbb{R}^m$, 线性方程组$A^TAX=A^T\beta$一定有解.

B2. 是否存在非齐次线性方程组$AX=\beta$, 使得$\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$是这个方程组的解, 且这个方程组的解都可以由向量组$\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$线性表示? 如果有, 请给出一个这样的方程组; 如果没有, 请说明理由. 其中

(1) $\alpha_1=\begin{pmatrix}1\\1\\1\\0\end{pmatrix},\alpha_2=\begin{pmatrix}1\\2\\0\\1\end{pmatrix},\alpha_3=\begin{pmatrix}1\\3\\-1\\2\end{pmatrix},\alpha_4=\begin{pmatrix}1\\1\\-1\\2\end{pmatrix},\alpha_5=\begin{pmatrix}6\\8\\-2\\8\end{pmatrix};$

(2) $\alpha_1=\begin{pmatrix}1\\1\\1\\0\end{pmatrix},\alpha_2=\begin{pmatrix}1\\1\\0\\1\end{pmatrix},\alpha_3=\begin{pmatrix}1\\1\\-1\\2\end{pmatrix},\alpha_4=\begin{pmatrix}1\\0\\1\\1\end{pmatrix},\alpha_5=\begin{pmatrix}1\\2\\2\\-2\end{pmatrix}.$

B3. 设A为n阶方阵, 其中$n\geqslant 2$. 如果A的每行之和都为0, 每列之和也都为零, 证明: A的伴随矩阵$\mathrm{adj}(A)$的所有位置的元素都相同.

§ 5.10 附录: 线性方程组和初等变换标准形

前面我们利用了Gauss消元法和几何讨论给出了线性方程组解的相容性定理和解的结构定理, 这里我们利用矩阵的初等变换标准形重新讨论一下. 这是一种比较狠的方法, 我们相当于解系数矩阵很简单的矩阵方程.

首先考虑齐次线性方程组$AX=0$, 其中$A\in\mathbb{F}^{m\times n}$, 且$R(A)=r$. 如果$r=n$, 即$A$列满秩. 前面

利用初等变换标准形我们已经知道列满秩阵可以左消去, 所以

$$AX=0 \Longleftrightarrow X=0,$$

即$AX=0$只有零解.

如果$r<n$, 存在m阶可逆阵P和n阶可逆阵Q, 使得$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$. 于是

$$AX=0 \Longleftrightarrow \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}QX=0.$$

记$QX=Y=\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, 其中$Y_1\in\mathbb{F}^r$, 则

$$AX=0 \Longleftrightarrow \begin{pmatrix} Y_1 \\ 0 \end{pmatrix}=0 \Longleftrightarrow Y_1=0.$$

所以$AX=0$的解为

$$X=Q^{-1}\begin{pmatrix} 0 \\ Y_2 \end{pmatrix}, \quad Y_2\in\mathbb{F}^{n-r}.$$

于是我们重新得到

定理5.37 (齐次线性方程组解的结构定理)**.** *设$A\in\mathbb{F}^{m\times n}$的秩$R(A)=r$, 则*

(1) *当$r=n$时, 齐次线性方程组$AX=0$只有零解;*

(2) *当$r<n$时, $AX=0$有非零解, 且其通解依赖于$n-r$个独立参数. 具体地, 设$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$, 其中$P$和$Q$分别是取定的$m$阶和$n$阶可逆阵, 则$AX=0$的通解是*

$$X=c_{r+1}Q^{-1}e_{r+1}+c_{r+2}Q^{-1}e_{r+2}+\cdots+c_nQ^{-1}e_n,$$

其中$c_{r+1},c_{r+2},\ldots,c_n\in\mathbb{F}$是任意常数.

下面再讨论非齐次线性方程组$AX=\beta$, 其中$A\in\mathbb{F}^{m\times n}$, $\beta\in\mathbb{F}^m$, 且$R(A)=r$. 取定m阶可逆阵P和n阶可逆阵Q, 使得$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$, 于是

$$AX=\beta \Longleftrightarrow \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}QX=P^{-1}\beta.$$

记$QX=\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$和$P^{-1}\beta=\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$, 其中$Y_1,Z_1\in\mathbb{F}^r$, 则

$$AX=\beta \Longleftrightarrow \begin{pmatrix} Y_1 \\ 0 \end{pmatrix}=\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \Longleftrightarrow Y_1=Z_1, Z_2=0.$$

这得到$AX=\beta$有解的充分和必要条件是$Z_2=0$; 且当$Z_2=0$时$AX=\beta$的解为

$$\begin{aligned}X=&Q^{-1}\begin{pmatrix}Z_1\\Y_2\end{pmatrix}=Q^{-1}\begin{pmatrix}Z_1\\0\end{pmatrix}+Q^{-1}\begin{pmatrix}0\\Y_2\end{pmatrix}\\=&Q^{-1}\begin{pmatrix}E_r&O\\O&O\end{pmatrix}_{n\times m}P^{-1}\beta+Q^{-1}\begin{pmatrix}0\\Y_2\end{pmatrix},\quad Y_2\in\mathbb{F}^{n-r}.\end{aligned}$$

下面来算增广矩阵的秩. 我们有

$$\begin{aligned}R(A,\beta)=&R\left(P\left[\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q,P^{-1}\beta\right]\right)\\=&R\left(\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q,\begin{pmatrix}Z_1\\Z_2\end{pmatrix}\right)\\=&R\left[\begin{pmatrix}E_r&O&Z_1\\O&O&Z_2\end{pmatrix}\begin{pmatrix}Q&O\\O&1\end{pmatrix}\right]\\=&R\begin{pmatrix}E_r&O&Z_1\\O&O&Z_2\end{pmatrix},\end{aligned}$$

所以

$$R(A,\beta)=r=R(A)\Longleftrightarrow Z_2=0.$$

结合上面的讨论, 就得到$AX=\beta$有解的充分必要条件是$R(A)=R(A,\beta)$. 于是我们重新得到

定理5.38 (非齐次线性方程组的相容性定理)**.** 非齐次线性方程组$AX=\beta$有解的充分和必要条件是它的系数矩阵和增广矩阵的秩相等.

定理5.39 (非齐次线性方程组解的结构定理)**.** 设非齐次线性方程组$AX=\beta$有解, 且$R(A)=r$. 则

(1) 当$r=n$时, 方程组的解唯一;

(2) 当$r<n$时, 方程组的通解依赖于$n-r$个独立参数, 且它的通解由其一个特解和对应的齐次线性方程组$AX=0$的通解构成. 具体地, 设$A=P\begin{pmatrix}E_r&O\\O&O\end{pmatrix}Q$, 其中$P$和$Q$分别是取定的$m$阶和$n$阶可逆阵, 则方程组$AX=\beta$的通解为

$$X=Q^{-1}\begin{pmatrix}E_r&O\\O&O\end{pmatrix}_{n\times m}P^{-1}\beta+c_{r+1}Q^{-1}e_{r+1}+c_{r+2}Q^{-1}e_{r+2}+\cdots+c_nQ^{-1}e_n,$$

其中$c_{r+1},c_{r+2},\ldots,c_n\in\mathbb{F}$是任意常数.

我们最后用这种方法再看几个例子.

例5.42. 求解矩阵方程$A^TX=X^TA$.

☞ **解.** 设$R(A)=r$, 取定m阶可逆阵P和n阶可逆阵Q, 使得$A=P\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q$, 则

$$\begin{aligned}A^TX=X^TA &\Longleftrightarrow Q^T\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}P^TX=X^TP\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q\\ &\Longleftrightarrow \begin{pmatrix}E_r & O\\ O & O\end{pmatrix}P^TXQ^{-1}=(P^TXQ^{-1})^T\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}.\end{aligned}$$

设$P^TXQ^{-1}=\begin{pmatrix}Y_1 & Y_2\\ Y_3 & Y_4\end{pmatrix}$, 其中$Y_1\in\mathbb{F}^{r\times r}$, 则

$$A^TX=X^TA\Longleftrightarrow\begin{pmatrix}Y_1 & Y_2\\ O & O\end{pmatrix}=\begin{pmatrix}Y_1^T & O\\ Y_2^T & O\end{pmatrix}\Longleftrightarrow Y_1^T=Y_1,Y_2=O.$$

这得到所求解的矩阵方程的解为

$$X=(P^T)^{-1}\begin{pmatrix}Y_1 & O\\ Y_3 & Y_4\end{pmatrix}Q,$$

其中$Y_1\in\mathbb{F}^{r\times r}$, $Y_3\in\mathbb{F}^{(m-r)\times r}$, $Y_4\in\mathbb{F}^{(m-r)\times(n-r)}$, 且$Y_1$是对称阵. □

例5.43 (Roth, 1952)**.** *设矩阵$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{p\times q}$, $C\in\mathbb{F}^{m\times q}$, 证明: 关于未知矩阵$X$和$Y$的矩阵方程$AX-YB=C$有解的充分必要条件是矩阵$\begin{pmatrix}A & O\\ O & B\end{pmatrix}$和矩阵$\begin{pmatrix}A & C\\ O & B\end{pmatrix}$等价.*

☞ **证明.** 假设存在矩阵X和Y, 使得$AX-YB=C$, 则

$$\begin{pmatrix}A & C\\ O & B\end{pmatrix}=\begin{pmatrix}A & AX-YB\\ O & B\end{pmatrix}\overset{r_1+Yr_2}{\sim}\begin{pmatrix}A & AX\\ O & B\end{pmatrix}\overset{c_2-c_1X}{\sim}\begin{pmatrix}A & O\\ O & B\end{pmatrix}.$$

反之, 假设$\begin{pmatrix}A & O\\ O & B\end{pmatrix}$和$\begin{pmatrix}A & C\\ O & B\end{pmatrix}$等价. 设$R(A)=r$, $R(B)=s$, 则存在可逆阵$P_1\in\mathbb{F}^{m\times m}$, $Q_1\in\mathbb{F}^{n\times n}$, $P_2\in\mathbb{F}^{p\times p}$, $Q_2\in\mathbb{F}^{q\times q}$, 使得

$$P_1AQ_1=\begin{pmatrix}E_r & O\\ O & O\end{pmatrix},\quad P_2BQ_2=\begin{pmatrix}E_s & O\\ O & O\end{pmatrix}.$$

这得到

$$\begin{pmatrix}P_1 & O\\ O & P_2\end{pmatrix}\begin{pmatrix}A & O\\ O & B\end{pmatrix}\begin{pmatrix}Q_1 & O\\ O & Q_2\end{pmatrix}=\begin{pmatrix}E_r & O & O & O\\ O & O & O & O\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix}$$

和

$$\begin{pmatrix}P_1 & O\\ O & P_2\end{pmatrix}\begin{pmatrix}A & C\\ O & B\end{pmatrix}\begin{pmatrix}Q_1 & O\\ O & Q_2\end{pmatrix}=\begin{pmatrix}E_r & O & C_{11} & C_{12}\\ O & O & C_{21} & C_{22}\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix},$$

其中$P_1CQ_2=\begin{pmatrix} C_{11} & C_{12}\\ C_{21} & C_{22}\end{pmatrix}\in\mathbb{F}^{m\times q}$, 而$C_{11}\in\mathbb{F}^{r\times s}$. 由于

$$\begin{pmatrix} E_r & O & C_{11} & C_{12}\\ O & O & C_{21} & C_{22}\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix}\sim\begin{pmatrix} E_r & O & O & O\\ O & O & O & C_{22}\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix},$$

所以

$$\begin{pmatrix} E_r & O & O & O\\ O & O & O & O\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix}\sim\begin{pmatrix} E_r & O & O & O\\ O & O & O & C_{22}\\ O & O & E_s & O\\ O & O & O & O\end{pmatrix}.$$

考虑秩得到$C_{22}=O$. 于是

$$\begin{aligned} C=&P_1^{-1}\begin{pmatrix} C_{11} & C_{12}\\ C_{21} & O\end{pmatrix}Q_2^{-1}=P_1^{-1}\begin{pmatrix} O & C_{12}\\ O & O\end{pmatrix}Q_2^{-1}+P_1^{-1}\begin{pmatrix} C_{11} & O\\ C_{21} & O\end{pmatrix}Q_2^{-1}\\ =&P_1^{-1}\begin{pmatrix} E_r & O\\ O & O\end{pmatrix}Q_1^{-1}Q_1\begin{pmatrix} O & C_{12}\\ O & O\end{pmatrix}_{n\times q}Q_2^{-1}+P_1^{-1}\begin{pmatrix} C_{11} & O\\ C_{21} & O\end{pmatrix}_{m\times p}P_2P_2^{-1}\begin{pmatrix} E_s & O\\ O & O\end{pmatrix}Q_2^{-1}\\ =&AQ_1\begin{pmatrix} O & C_{12}\\ O & O\end{pmatrix}_{n\times q}Q_2^{-1}+P_1^{-1}\begin{pmatrix} C_{11} & O\\ C_{21} & O\end{pmatrix}_{m\times p}P_2B.\end{aligned}$$

这得到解$X=Q_1\begin{pmatrix} O & C_{12}\\ O & O\end{pmatrix}_{n\times q}Q_2^{-1}$, $Y=-P_1^{-1}\begin{pmatrix} C_{11} & O\\ C_{21} & O\end{pmatrix}_{m\times p}P_2$. □

§ 5.11 附录: 矩阵的广义逆

设A是方阵, 则当A可逆时, 线性方程组$AX=\beta$有唯一解, 而且解为$X=A^{-1}\beta$. 下面利用矩阵的初等变换标准形, 将上面的结果推广到任意矩阵.

5.11.1 广义逆

如果方阵A可逆, 则有$AA^{-1}=E$, 进而

$$AA^{-1}A=A,$$

即A^{-1}为矩阵方程$AXA=A$的(唯一)解. 一般的, 对任意$A\in\mathbb{F}^{m\times n}$, 我们考察矩阵方程$AXA=A$.

定理5.40. *设$A\in\mathbb{F}^{m\times n}$, 则矩阵方程$AXA=A$恒有解. 具体地, 设$R(A)=r$, 且*

$$A=P\begin{pmatrix} E_r & O\\ O & O\end{pmatrix}Q,$$

其中P和Q分别是取定的m阶和n阶可逆阵, 则矩阵方程$AXA=A$的通解是

$$X=Q^{-1}\begin{pmatrix} E_r & B\\ C & D\end{pmatrix}P^{-1},$$

其中$B\in\mathbb{F}^{r\times(m-r)}$, $C\in\mathbb{F}^{(n-r)\times r}$, $D\in\mathbb{F}^{(n-r)\times(m-r)}$是任意矩阵.

定义5.12. 设$A\in\mathbb{F}^{m\times n}$, 矩阵方程$AXA=A$的每一个解都称为矩阵$A$的广义逆, 记为$A^-$.

由定义, 对于$A\in\mathbb{F}^{m\times n}$, 有

$$A^-\in\{X\in\mathbb{F}^{n\times m}\mid AXA=A\}.$$

当A为可逆方阵时, $A^-=A^{-1}$, 唯一. 而一般的, A^-不唯一. 事实上, 由于

$$A^-=Q^{-1}\begin{pmatrix}E_r & B\\ C & D\end{pmatrix}P^{-1},$$

所以

$$R(A^-)=R\begin{pmatrix}E_r & B\\ C & D\end{pmatrix}\geqslant R(E_r)=r,$$

所以

$$r\leqslant R(A^-)\leqslant\min\{m,n\}.$$

反之, 设整数k满足$r\leqslant k\leqslant\min\{m,n\}$, 取$D\in\mathbb{F}^{(n-r)\times(m-r)}$, 使得$R(D)=k-r$, 则

$$A^-=Q^{-1}\begin{pmatrix}E_r & O\\ O & D\end{pmatrix}P^{-1}$$

的秩$R(A^-)=k$.

☞ **定理5.40的证明.** 任取$B\in\mathbb{F}^{r\times(m-r)}$, $C\in\mathbb{F}^{(n-r)\times r}$, $D\in\mathbb{F}^{(n-r)\times(m-r)}$, 则

$$\begin{aligned}&A\cdot Q^{-1}\begin{pmatrix}E_r & B\\ C & D\end{pmatrix}P^{-1}\cdot A=P\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}\begin{pmatrix}E_r & B\\ C & D\end{pmatrix}\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q\\ =&P\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q=A,\end{aligned}$$

即矩阵$Q^{-1}\begin{pmatrix}E_r & B\\ C & D\end{pmatrix}P^{-1}$是矩阵方程$AXA=A$的解.

反之, 设X是$AXA=A$的一个解, 则

$$P\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}QXP\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q=P\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}Q,$$

进而有

$$\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}QXP\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}=\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}.$$

将QXP分块

$$QXP=\begin{pmatrix}F & B\\ C & D\end{pmatrix},\quad F\in\mathbb{F}^{r\times r},$$

则有

$$\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}=\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}\begin{pmatrix}F & B\\ C & D\end{pmatrix}\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}=\begin{pmatrix}F & O\\ O & O\end{pmatrix},$$

进而得$F = E_r$. 于是

$$X = Q^{-1}\begin{pmatrix} E_r & B \\ C & D \end{pmatrix} P^{-1},$$

证毕. □

设$A \in \mathbb{F}^{m\times n}$, 于是$AA^-A = A$. 如果$A$列满秩, 则可左削去, 即得$A^-A = E_n$. 反之, 如果$B \in \mathbb{F}^{n\times m}$, 使得$BA = E_n$, 则有$ABA = A$, 即$B$为$A$的广义逆. 于是列满秩矩阵$A$的广义逆也就是矩阵方程$XA = E_n$的解. 此时, 矩阵方程$YA = B$一定有解, 比如$Y = BA^-$就是一个解. 读者可以类似写出行满秩阵的相应结论.

下面利用矩阵的广义逆重新叙述线性方程组解的相容性定理和结构定理.

定理5.41 (线性方程组解的相容性定理). *线性方程组$AX = \beta$有解的充分和必要条件是, $\beta = AA^-\beta$.*

☞ **证明.** 设α是解, 则

$$\beta = A\alpha = AA^-A\alpha = AA^-\beta.$$

反之, 设$\beta = AA^-\beta$, 则$A^-\beta$是一个解. □

定理5.42 (线性方程组解的结构定理). *设线性方程组$AX = \beta$有解, 则它的通解是*

$$X = A^-\beta + (E_n - A^-A)Z,$$

其中A^-是A的某个取定的广义逆, $Z \in \mathbb{F}^n$任意. 进而, 当$\beta \neq 0$时, $AX = \beta$的通解是

$$X = A^-\beta,$$

其中A^-是A的任意一个广义逆.

☞ **证明.** 由于线性方程组$AX = \beta$有解, 所以$\beta = AA^-\beta$. 任取$Z \in \mathbb{F}^n$, 有

$$A[A^-\beta + (E_n - A^-A)Z] = AA^-\beta + (A - AA^-A)Z = \beta.$$

反之, 设α是$AX = \beta$的一个解. 任取A的广义逆A^-. 取$Z = \alpha \in \mathbb{F}^n$, 则

$$A^-\beta + (E_n - A^-A)Z = A^-\beta + \alpha - A^-A\alpha = A^-\beta + \alpha - A^-\beta = \alpha.$$

下证: 当$\beta \neq 0$时, 可取一个A的广义逆A^-, 使得$\alpha = A^-\beta$.

事实上, 设$R(A) = r$, 且$A = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$Q, 其中$P, Q$为可逆方阵, 则由$A\alpha = \beta$得

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q\alpha = P^{-1}\beta.$$

分块

$$Q\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad P^{-1}\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

其中$\alpha_1, \beta_1 \in \mathbb{F}^r$, 则有

$$\begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

这得到

$$\alpha_1 = \beta_1, \qquad \beta_2 = 0.$$

由于$\beta \neq 0$, 所以$P^{-1}\beta \neq 0$, 进而$\beta_1 \neq 0$. 记

$$\alpha_1 = \beta_1 = (b_1, \dots, b_r)^T,$$

其中$b_i \neq 0$.

任取A的广义逆A^-, 则存在B, C, D, 使得

$$A^- = Q^{-1}\begin{pmatrix} E_r & B \\ C & D \end{pmatrix} P^{-1}.$$

于是

$$QA^-\beta = \begin{pmatrix} E_r & B \\ C & D \end{pmatrix} P^{-1}\beta = \begin{pmatrix} E_r & B \\ C & D \end{pmatrix}\begin{pmatrix} \beta_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ C\beta_1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ C\beta_1 \end{pmatrix}.$$

取矩阵$C \in \mathbb{F}^{(n-r)\times r}$, 使其第$i$列为$\frac{1}{b_i}\alpha_2$, 其它列为零, 则

$$C\beta_1 = \alpha_2.$$

此时, 有$QA^-\beta = Q\alpha$, 即$A^-\beta = \alpha$. 所以取A的广义逆

$$A^- = Q^{-1}\begin{pmatrix} E_r & O \\ C & O \end{pmatrix} P^{-1},$$

其中, C的第i列为$\frac{1}{b_i}\alpha_2$, 其它列为零, 则有$\alpha = A^-\beta$. □

5.11.2 Moore-Penrose广义逆

下面介绍一种特殊的广义逆, 它是唯一的. 设$A \in \mathbb{C}^{m\times n}$, 考虑矩阵方程组

$$\begin{cases} AXA = A, & (P_1) \\ XAX = X, & (P_2) \\ \overline{(AX)^T} = AX, & (P_3) \\ \overline{(XA)^T} = XA. & (P_4) \end{cases}$$

称该方程组为**Penrose方程组**.

我们先证明: 如果Penrose方程组有解, 那么解唯一. 事实上, 设X_1和X_2是两个解, 则

$$\begin{aligned} X_1 &\xlongequal{(P_2)} X_1AX_1 \xlongequal{(P_1)} X_1AX_2AX_1 \xlongequal{(P_3)} X_1\overline{(AX_2)^T}\,\overline{(AX_1)^T} \\ &= X_1\overline{(AX_1AX_2)^T} \xlongequal{(P_1)} X_1\overline{(AX_2)^T} \xlongequal{(P_3)} X_1AX_2 \\ &\xlongequal{(P_1)} X_1AX_2AX_2 \xlongequal{(P_4)} \overline{(X_1A)^T}\,\overline{(X_2A)^T}X_2 \\ &= \overline{(X_2AX_1A)^T}X_2 \xlongequal{(P_1)} \overline{(X_2A)^T}X_2 \xlongequal{(P_4)} X_2AX_2 \xlongequal{(P_2)} X_2. \end{aligned}$$

下面再找Penrose方程组的解. 设$R(A) = r$, 且

$$A = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q,$$

其中P和Q分别是m阶和n阶可逆阵. 由定理5.40, 方程(P_1)的解是

$$X = Q^{-1}\begin{pmatrix} E_r & B \\ C & D \end{pmatrix} P^{-1},$$

其中$B \in \mathbb{C}^{r\times(m-r)}$, $C \in \mathbb{C}^{(n-r)\times r}$, $D \in \mathbb{C}^{(n-r)\times(m-r)}$. 代入$(P_2)$, 得

$$\begin{pmatrix} E_r & B \\ C & D \end{pmatrix}\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}\begin{pmatrix} E_r & B \\ C & D \end{pmatrix} = \begin{pmatrix} E_r & B \\ C & D \end{pmatrix},$$

即

$$\begin{pmatrix} E_r & B \\ C & CB \end{pmatrix} = \begin{pmatrix} E_r & B \\ C & D \end{pmatrix}.$$

于是$D = CB$, 进而

$$X = Q^{-1}\begin{pmatrix} E_r & B \\ C & CB \end{pmatrix} P^{-1}.$$

由此得

$$AX = P\begin{pmatrix} E_r & B \\ O & O \end{pmatrix} P^{-1}, \qquad XA = Q^{-1}\begin{pmatrix} E_r & O \\ C & O \end{pmatrix} Q.$$

代入(P_3)和(P_4), 得到

$$\begin{pmatrix} E_r & O \\ \overline{B^T} & O \end{pmatrix}\overline{P^T}P = \overline{P^T}P\begin{pmatrix} E_r & B \\ O & O \end{pmatrix},$$
$$\begin{pmatrix} E_r & O \\ C & O \end{pmatrix} Q\overline{Q^T} = Q\overline{Q^T}\begin{pmatrix} E_r & \overline{C^T} \\ O & O \end{pmatrix}.$$

做分块

$$\overline{P^T}P = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}, \quad Q\overline{Q^T} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},$$

其中$R_{11}, S_{11} \in \mathbb{C}^{r\times r}$, 则可得

$$R_{12} = R_{11}B, \quad S_{21} = CS_{11}.$$

由Binet-Cauchy公式, 可得

$$\begin{aligned}
\det(R_{11}) =& (\overline{P^T}P)\begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix} \\
=& \sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant m} \overline{P^T}\begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix} \\
=& \sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant m} \left| P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix}\right|^2 \geqslant 0.
\end{aligned}$$

如果对任意$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant m$, 有$P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix} = 0$, 则$\det(P)$按照前$r$列展开, 可得$\det(P) = 0$, 矛盾于$P$可逆. 所以$\det(R_{11}) > 0$, 即$R_{11}$可逆(也可以从$P$可逆得出$\overline{P^T}P$是正定的Hermite阵, 从而其$r$阶主子式大于零). 类似可证$S_{11}$可逆. 于是

$$B = R_{11}^{-1}R_{12}, \quad C = S_{21}S_{11}^{-1}.$$

最后得到

$$\begin{aligned} X =& Q^{-1}\begin{pmatrix} E_r & B \\ C & CB \end{pmatrix}P^{-1} = Q^{-1}\begin{pmatrix} E_r \\ C \end{pmatrix}(E_r, B)P^{-1} \\ =& Q^{-1}\begin{pmatrix} E_r \\ S_{21}S_{11}^{-1} \end{pmatrix}(E_r, R_{11}^{-1}R_{12})P^{-1} \\ =& Q^{-1}\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} S_{11}^{-1} \\ O \end{pmatrix}(R_{11}^{-1}, O)\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}P^{-1} \\ =& \overline{Q^T}\begin{pmatrix} E_r \\ O \end{pmatrix}S_{11}^{-1}R_{11}^{-1}(E_r, O)\overline{P^T}. \end{aligned}$$

令

$$H = P\begin{pmatrix} E_r \\ O \end{pmatrix}, \quad L = (E_r, O)Q,$$

则$A = HL$为满秩分解, 且有

$$R_{11} = \overline{H^T}H, \quad S_{11} = L\overline{L^T}.$$

于是得到

$$X = \overline{L^T}\left(L\overline{L^T}\right)^{-1}\left(\overline{H^T}H\right)^{-1}\overline{H^T}.$$

定理5.43. 设$A \in \mathbb{C}^{m\times n}$, 则*Penrose*方程组有唯一解. 具体地, 设$A = HL$是A的一个满秩分解, 则*Penrose*方程组的唯一解为

$$X = \overline{L^T}\left(L\overline{L^T}\right)^{-1}\left(\overline{H^T}H\right)^{-1}\overline{H^T}.$$

☞ **证明.** 唯一性已证, 下面只要验证所给为解即可. 对于(P_1), 有

$$AXA = HL\left[\overline{L^T}\left(L\overline{L^T}\right)^{-1}\left(\overline{H^T}H\right)^{-1}\overline{H^T}\right]HL = HL = A;$$

对于(P_2), 有

$$XAX = \left[\overline{L^T}\left(L\overline{L^T}\right)^{-1}\left(\overline{H^T}H\right)^{-1}\overline{H^T}\right]HL\left[\overline{L^T}\left(L\overline{L^T}\right)^{-1}\left(\overline{H^T}H\right)^{-1}\overline{H^T}\right] = X;$$

对于(P_3), 有

$$\overline{(AX)^T} = \overline{\left(H\left(\overline{H^T}H\right)^{-1}\overline{H^T}\right)^T} = H\left(\overline{H^T}H\right)^{-1}\overline{H^T} = AX;$$

对于(P_4), 有

$$\overline{(XA)^T} = \overline{\left(\overline{L^T}\left(L\overline{L^T}\right)^{-1}L\right)^T} = \overline{L^T}\left(L\overline{L^T}\right)^{-1}L = XA.$$

证毕. □

定义5.13. 设$A \in \mathbb{C}^{m\times n}$, 称$A$对应的*Penrose*方程组的唯一解为矩阵$A$的**Moore-Penrose广义逆**, 记为A^+.

例5.44. **(1)** 如果A为可逆方阵, 则$A^+ = A^- = A^{-1}$. 特别地, 对$a \in \mathbb{C}^\times$, 有$a^+ = \frac{1}{a}$;

(2) 零矩阵的*Moore-Penrose*广义逆为零矩阵. 具体有$O_{m\times n}^+ = O_{n\times m}$.

例5.45. 设$A \in \mathbb{C}^{m\times n}$, $a \in \mathbb{C}$, 证明:

$$(A^+)^+ = A, \quad \left(\overline{A^T}\right)^+ = \overline{(A^+)^T}, \quad (aA)^+ = a^+A^+.$$

证明. 由Penrose方程组中A和X的对称性, 可得$X^+ = A$, 即$(A^+)^+ = A$. 将Penrose方程组的每个方程同时取转置共轭, 可得$\overline{X^T} = \left(\overline{A^T}\right)^+$, 即有

$$\overline{(A^+)^T} = \left(\overline{A^T}\right)^+.$$

最后利用数和矩阵可交换, 容易得到最后一个等式. □

由上面例子可以看出, Moore-Penrose广义逆与可逆阵的逆有一些类似的性质, 但是对逆矩阵的穿脱原理对Moore-Penrose广义逆并不成立.

例5.46. 设

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,1) \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,1),$$

则

$$A^+ = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1,0) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad B^+ = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(0,1) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

由于$AB = A$, 所以

$$(AB)^+ = A^+ = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

而

$$A^+B^+ = \frac{1}{4}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad B^+A^+ = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

所以

$$(AB)^+ \neq A^+B^+, \quad (AB)^+ \neq B^+A^+.$$

习题5.11

A1. 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{m\times p}$, 证明: 矩阵方程$AX=B$有解的充分必要条件是$B=AA^-B$, 而且在有解时, 通解为

$$X=A^-B+(E_n-A^-A)W,$$

其中$W\in\mathbb{F}^{n\times p}$任意.

A2. 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{p\times q}$, $C\in\mathbb{F}^{m\times q}$, 证明: 矩阵方程$AXB=C$有解的充分必要条件是,

$$(E_m-AA^-)C=C(E_q-B^-B)=O,$$

而且在有解时, 通解为

$$X=A^-CB^-+(E_n-A^-A)Y+Z(E_p-BB^-)+(E_n-A^-A)W(E_p-BB^-),$$

其中$Y,Z,W\in\mathbb{F}^{n\times p}$任意.

A3. 设$A\in\mathbb{F}^{m\times p}$, $B\in\mathbb{F}^{q\times n}$, $C\in\mathbb{F}^{m\times n}$, 证明: 关于未知矩阵$X$和$Y$的矩阵方程$AX-YB=C$有解的充分必要条件是

$$(E_m-AA^-)C(E_n-B^-B)=O,$$

而且在有解时, 通解为

$$\begin{aligned}X=&A^-C+A^-ZB+(E_p-A^-A)W,\\ Y=&-(E_m-AA^-)CB^-+Z-(E_m-AA^-)ZBB^-,\end{aligned}$$

其中$W\in\mathbb{F}^{p\times n}$, $Z\in\mathbb{F}^{m\times q}$任意.

A4. 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{p\times q}$, $C\in\mathbb{F}^{m\times q}$, 证明: 存在$A$的广义逆$A^-$和$B$的广义逆$B^-$, 使得

$$R\begin{pmatrix}A & C\\ O & B\end{pmatrix}=R(A)+R(B)+R[(E_m-AA^-)C(E_q-B^-B)].$$

§ 5.12 补充题

A1. 设A是n阶方阵, $R(A)=r$. 证明: A是幂等阵的充分必要条件是, 存在秩等于r的$n\times r$阶矩阵H和秩等于r的$r\times n$阶矩阵L, 使得$A=HL$, $LH=E_r$.

A2. 设$A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{m\times k}$, B有列分块$B=(\beta_1,\beta_2,\ldots,\beta_k)$, 证明: 如果对$1\leqslant j\leqslant k$, 都有$R(A)=R(A,\beta_j)$, 则$R(A)=R(A,B)$.

A3. 设n元非齐次线性方程组$AX=\beta$有解, 取定整数k, $1\leqslant k\leqslant n$, 将A的第k列划去后得到的矩阵记为B, 证明: 如果$AX=\beta$的每个解向量的第k个分量都等于零, 那么$R(B)<R(A)$.

A4. 证明: 线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \qquad\qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m\end{cases}$$

有解的充分必要条件是线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{m1}x_m = 0, \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{m2}x_m = 0, \\ \qquad\qquad \vdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{mn}x_m = 0, \\ b_1x_1 + b_2x_2 + \cdots + b_mx_m = 1 \end{cases}$$

无解.

A5. 设矩阵$A \in \mathbb{F}^{m\times n}$, 其中$m < n$. 若齐次线性方程组$AX = 0$的基础解系为$\eta_i = (b_{i1}, b_{i2}, \ldots, b_{in})^T$, $i = 1, 2, \ldots, n-m$, 求齐次线性方程组

$$\begin{cases} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1n}y_n = 0, \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2n}y_n = 0, \\ \qquad\qquad \vdots \\ b_{n-m,1}y_1 + b_{n-m,2}y_2 + \cdots + b_{n-m,n}y_n = 0 \end{cases}$$

的基础解系.

A6. 设两个非齐次线性方程组的通解分别为

$$\gamma + c_1\eta_1 + c_2\eta_2, \quad \delta + a_1\xi_1 + a_2\xi_2,$$

其中

$$\gamma = \begin{pmatrix} 5 \\ -3 \\ 0 \\ 0 \end{pmatrix}, \eta_1 = \begin{pmatrix} -6 \\ 5 \\ 1 \\ 0 \end{pmatrix}, \eta_2 = \begin{pmatrix} -5 \\ 4 \\ 0 \\ 1 \end{pmatrix}; \delta = \begin{pmatrix} -11 \\ 3 \\ 0 \\ 0 \end{pmatrix}, \xi_1 = \begin{pmatrix} 8 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \xi_2 = \begin{pmatrix} 10 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

求这两个方程组的公共解.

A7. 设非齐次线性方程组$\begin{cases} 7x_1 - 6x_2 + 3x_3 = a, \\ 8x_1 - 9x_2 + bx_4 = 7 \end{cases}$ 和通解为

$$(1, 1, 0, 0)^T + c_1(1, 0, -1, 0)^T + c_2(2, 3, 0, 1)^T$$

的非齐次线性方程组有无穷多组公共解, 求a, b的值, 并求出公共解.

A8. 证明: 实平面上3条不同的直线

$$\begin{aligned} L_1: &\quad ax + by + c = 0, \\ L_2: &\quad bx + cy + a = 0, \\ L_3: &\quad cx + ay + b = 0 \end{aligned}$$

相交于一点的充分必要条件是$a + b + c = 0$.

A9. 设$V \subset \mathbb{F}^n$是向量空间, 证明: 存在矩阵A, 使得V是n元齐次线性方程组$AX = 0$的解空间.

A10. 设A和B是n阶方阵, 齐次线性方程组$AX = 0$和$BX = 0$同解, 且每个方程组的基础解系含有m个向量, 证明: $R(A - B) \leqslant n - m$.

A11. 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times k}$, 证明: 齐次线性方程组$ABX = 0$和$BX = 0$同解的充分必要条件是$R(AB) = R(B)$.

A12. 设$A = (a_{ij})$和$B = (b_{ij})$是n阶方阵, 且$b_{ij} = (-1)^{i+j}a_{ij}$, $i, j = 1, 2, \ldots, n$, 证明: $R(B) = R(A)$.

B1. 设有$n + 1$个人及供他们阅读的n本不同的书, 假设每人至少读了一本书, 证明: 在这$n + 1$个人中一定存在甲乙两组, 使得甲组读过书的种类与乙组读过书的种类相同.

B2. 设矩阵$A \in \mathbb{F}^{m\times n}$的行分块和列分块分别为

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = (\beta_1, \beta_2, \ldots, \beta_n),$$

且$\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_r}$和$\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_r}$分别是$A$的行向量组和列向量组的极大无关组, 其中$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant m$, $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$, 证明: A的r阶子式$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} \neq 0$.

B3. 设有$\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in}) \in \mathbb{C}^n$, 其中$i = 1, 2, \ldots, t$, 而$t \leqslant n$. 假设对$i = 1, 2, \ldots, t$, 成立

$$2|a_{ii}| > \sum_{k=1}^{n} |a_{ik}|,$$

证明: 向量组$\alpha_1, \alpha_2, \ldots, \alpha_t$线性无关.

B4. 设矩阵$A, B \in \mathbb{F}^{m\times n}$, 证明: $R(A \circ B) \leqslant R(A)R(B)$.

B5. 设矩阵$A, B \in \mathbb{F}^{m\times n}$, 证明: $R(A, B) + R\begin{pmatrix} A \\ B \end{pmatrix} \leqslant R(A) + R(B) + R(A + B)$.

B6. 设$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times k}$, 且$R(AB) = R(B)$. 证明: 对任意$C \in \mathbb{F}^{k\times l}$, 有$R(ABC) = R(BC)$.

B7. 证明: 线性方程组$AX = \beta$有解的充分必要条件是, 齐次线性方程组$A^TY = 0$的任一解α均满足$\alpha^T\beta = 0$.

B8. 设矩阵方程$AX = B$有解, 其中$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{m\times l}$. 证明:

(1) 存在$AX = B$的解X_0, 使得$R(X_0) = R(B)$. 称X_0为$AX = B$的秩最小的解;

(2) 若X_0为$AX = B$的秩最小的解, 则存在$M \in \mathbb{F}^{n\times m}$, 使得$X_0 = MB$.

B9. 设矩阵$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{p\times q}$, $C \in \mathbb{F}^{m\times q}$, 证明: 矩阵方程$AXB = C$有解的充分必要条件是, 矩阵方程$AY = C$和$ZB = C$都有解.

第六章　线性空间

从几何上看, 线性代数主要研究线性空间及其上的线性变换, 本章介绍线性空间的方方面面. 线性空间是几何空间的抽象, 是一种最简单的代数结构, 其基本模型是上一章介绍的向量空间, 所以我们应该用具体的几何模型(例如我们生活的三维空间)来理解抽象的概念. 所谓代数结构就是带有一些具有好的性质的运算的非空集合, 我们抽象地研究这些代数结构的性质, 然后将研究得到的结果用到具体的例子上. 考查一个对象时, 一方面可以研究它的内部结构和表象, 另一方面还可以研究这个对象和其它对象的关系. 我们通常也按照这两个方面来研究一个代数结构, 前者比如子结构, 商结构, 是否可以用最单纯的结构分解等; 后者考虑关系, 在数学上就是研究具有好的性质的映射. 对于线性空间这一代数结构, 我们在本章先考虑其内部结构, 比如子空间, 商空间和直和分解等; 在下章我们考虑线性空间之间的关系, 即考虑线性映射.

另一个看待线性代数几何理论的观点是我们在做解析几何. 所以我们首先要在空间中建立坐标系, 将空间中每个元素用对应于坐标系的坐标表示; 然后利用建立的坐标系来研究空间之间好的变换. 本章的一个主要任务就是建立坐标系, 每个坐标轴可以用其上的一个非零向量代表, 于是建立坐标系等价于找到没有关系的一组元素, 使其可以表示空间中的任意一个元素, 在线性代数里称其为基(我们在向量空间中已经这么做了). 当有了基后, 就可以用代数工具来研究几何空间了.

本章中$\mathbb{F}$表示任意一个数域.

§ 6.1　线性空间

前面我们讲了线性方程组的求解之旅这个故事, 故事结束了, 课程不就应该结束了吗? 怎么还有? 我们想说的是, 前面的故事仅仅是一个前传, 正式的故事才刚刚开始. 希望读者可以这么想, 前传都这么精彩, 那么正式的故事还能差吗? 从而对我们后面的故事充满期待.

我们在上一章研究了向量空间, 它是对向量的加法和数乘封闭的$\mathbb{F}^n$的非空子集. 除了向量外, 还有许多对象有加法和数乘. 比如矩阵, 多项式, 函数等等. 这些对象上的加法和数乘都满足一些共同的性质, 比如加法都是交换和结合的, 加法都有零元和负元等. 将这些进行抽象, 忘记这些对象具体的身份(向量, 矩阵, 多项式, 函数), 而抓住本质的地方: 有两个运算, 满足好的性质, 就得到线性空间的概念. 我们抽象地研究线性空间的性质, 然后再用到各个具体的例子上, 一方面使得应用范围非常之广, 不拘泥于向量, 矩阵, 多项式或者函数, 另一方面, 忘记对象的身份常常可以发现一些本质的东西. 需要提醒读者的是, 在学习或者思考抽象概念时, 不能只靠记忆, 而应该用具体的例子来理解抽象概念为何如此以及结论成立的本质原因. 毕竟艺术源于生活, 抽象源于实例, 好的抽象不是无源之

水, 而一定是自然和必要的.

下面是线性空间的抽象定义.

定义6.1. 设V是非空集合, 如果

(i) 在V中有二元运算, 称为加法: 对任意$\alpha,\beta\in V$, 存在唯一的$\gamma\in V$与之对应, 称为α与β的和, 记为$\gamma=\alpha+\beta$. 加法满足下面公理:

(A1) (加法结合律) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$, $\quad(\forall\alpha,\beta,\gamma\in V)$;

(A2) (加法交换律) $\alpha+\beta=\beta+\alpha$, $\quad(\forall\alpha,\beta\in V)$;

(A3) (零元存在性) $\exists 0\in V$, 使得: $\forall\alpha\in V$有$\alpha+0=\alpha$. 称0为V的零元素;

(A4) (负元存在性) $\forall\alpha\in V$, $\exists\beta\in V$, 使得$\alpha+\beta=0$. 称β为α的负元;

(ii) 在$\mathbb{F}$和V之间有运算, 称为数乘: 对任意$c\in\mathbb{F}$, 任意$\alpha\in V$, 存在唯一的$\delta\in V$与之对应, 称为c与α的数乘, 记为$\delta=c\alpha=c\cdot\alpha$. 数乘满足下面公理

(M1) (二次数乘律) $a(b\alpha)=(ab)\alpha$, $\quad(\forall a,b\in\mathbb{F},\forall\alpha\in V)$;

(M2) (单位数乘律) $1\cdot\alpha=\alpha$, $\quad(\forall\alpha\in V)$;

(iii) 加法和数乘还满足公理

(D1) (数对元的分配律) $c(\alpha+\beta)=c\alpha+c\beta$, $\quad(\forall c\in\mathbb{F},\forall\alpha,\beta\in V)$;

(D2) (元对数的分配律) $(a+b)\alpha=a\alpha+b\alpha$, $\quad(\forall a,b\in\mathbb{F},\forall\alpha\in V)$,

则称V是数域$\mathbb{F}$上的线性空间. 有时也称V为数域$\mathbb{F}$上的向量空间, 线性空间V中的任意元素α称为向量. 当$\mathbb{F}=\mathbb{R}$时, 称V是实线性空间; 当$\mathbb{F}=\mathbb{C}$时, 称V是复线性空间.

设V是$\mathbb{F}$上的线性空间. 由定义可知V中的零元素是唯一的. 事实上, 如果V有两个零元素0_1和0_2, 则有

$$0_1=0_1+0_2=0_2+0_1=0_2.$$

类似的, 对任意$\alpha\in V$, α的负元素也是唯一的, 我们将这唯一的负元素记为$-\alpha$. 事实上, 如果$\beta,\gamma\in V$都是α的负元素, 则$\alpha+\beta=0$, $\alpha+\gamma=0$. 于是有

$$\beta=\beta+0=\beta+(\alpha+\gamma)=(\beta+\alpha)+\gamma=(\alpha+\beta)+\gamma=0+\gamma=\gamma+0=\gamma.$$

利用元素的负元素, 可以定义V中的减法为: $\forall\alpha,\beta\in V$, 有

$$\alpha-\beta:=\alpha+(-\beta).$$

下面我们看一些线性空间的例子.

例6.1. **(1)** 数域$\mathbb{F}$关于$\mathbb{F}$的加法和乘法成为$\mathbb{F}$上的线性空间;

(2) 由(1), $\mathbb{C}$关于复数的加法和乘法成为复线性空间. 类似的, $\mathbb{C}$关于复数的加法和实数与复数的乘法成为实线性空间;

(3) 如果$\mathbb{K}$是数域, 且$\mathbb{F}\subset\mathbb{K}$, 则类似于(2), $\mathbb{K}$成为$\mathbb{F}$上的线性空间.

例6.2. **(1)** n维(列)向量空间

$$\mathbb{F}^n=\{(a_1,a_2,\ldots,a_n)^T\mid a_1,a_2,\ldots,a_n\in\mathbb{F}\}$$

关于向量的加法和向量的数乘成为$\mathbb{F}$上的线性空间. 其中, 零元为零向量$0=(\underbrace{0,\ldots,0}_{n})^T$, 而任意$\alpha=(a_1,a_2,\ldots,a_n)^T\in\mathbb{F}^n$的负元$-\alpha=(-a_1,-a_2,\ldots,-a_n)$;

(2) 类似的, $\mathbb{F}$上所有$m\times n$矩阵所成集合

$$\mathbb{F}^{m\times n}=\{(a_{ij})_{m\times n}\mid a_{ij}\in\mathbb{F},\forall i,j\}$$

关于矩阵的加法和矩阵的数乘成为$\mathbb{F}$上的线性空间. 其中, 零元为$m\times n$的零矩阵$O=(0)_{m\times n}$, 而矩阵$A=(a_{ij})_{m\times n}\in\mathbb{F}^{m\times n}$的负元$-A=(-a_{ij})_{m\times n}$.

例6.3. 数域$\mathbb{F}$上的所有数列做成的集合记为

$$\mathbb{F}^\infty:=\{(a_1,a_2,\ldots)\mid a_j\in\mathbb{F},j=1,2,\ldots,\},$$

定义如下的加法和数乘

$$\begin{aligned}&(a_1,a_2,\ldots)+(b_1,b_2,\ldots):=(a_1+b_1,a_2+b_2,\ldots),\\&c(a_1,a_2,\ldots):=(ca_1,ca_2,\ldots),\end{aligned}$$

则$\mathbb{F}^\infty$在这两个运算下成为$\mathbb{F}$上的线性空间, 其中零元是零数列$0=(0,0,\ldots)$.

例6.4. 数域$\mathbb{F}$上所有的一元多项式做成的集合$\mathbb{F}[x]$关于多项式的加法和多项式的数乘成为$\mathbb{F}$上的线性空间, 其中零元是零多项式, 而对$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{F}[x]$, $f(x)$的负元是$-f(x)=-a_nx^n-a_{n-1}x^{n-1}-\cdots-a_1x-a_0$.

例6.5. 设V是闭区间$[a,b]$上所有的实值函数组成的集合, 则V关于函数的加法和实数对函数的乘法成为实线性空间.

例6.6. 设$V=\mathbb{R}_{>0}$是所有正实数所成的集合.

(1) V关于实数的加法和实数乘法不成为实线性空间, 比如数乘不封闭;

(2) 定义V中如下的加法和数乘: $\forall a,b\in V$, $\forall\lambda\in\mathbb{R}$

$$a\oplus b:=ab,\qquad \lambda\circ a:=a^\lambda,$$

则V在这两个运算下成为实线性空间. 其中零元为1, 而对于$a\in V$, a的负元是a^{-1}.

从上面的例子可以看出, 许多大家遇到的数学对象都可以看成线性空间. 抽象地研究线性空间, 一定程度上相当于同时研究这些数学对象. 由此可预见抽象的威力和魅力, 请读者往后欣赏.

设V是数域$\mathbb{F}$上的线性空间, 由于V中每个元素都有负元, 所以加法消去律成立. 例如, 如果$\alpha, \beta, \gamma \in V$满足$\alpha+\beta=\alpha+\gamma$, 则有$\beta=\gamma$. 这由于

$$\begin{aligned}\beta &= 0+\beta=(-\alpha+\alpha)+\beta=-\alpha+(\alpha+\beta)\\ &=-\alpha+(\alpha+\gamma)=(-\alpha+\alpha)+\gamma=0+\gamma=\gamma.\end{aligned}$$

进而对于V中向量的等式, 移项法则成立. 例如, 如果$\alpha+\beta=\gamma$, 则有$\alpha+\beta=\gamma-\beta+\beta$, 进而$\alpha=\gamma-\beta$, 将$\beta$移到右边, 取其负元. 在$V$中还成立下面简单的性质.

性质1 $0\alpha=0, \quad (\forall \alpha \in V)$.

性质2 $c0=0, \quad (\forall c \in \mathbb{F})$.

性质3 设$c \in \mathbb{F}$, $\alpha \in V$, 则: $c\alpha=0 \Longrightarrow c=0$或者$\alpha=0$.

性质4 $(-1)\alpha=-\alpha, \quad (\forall \alpha \in V)$.

证明. 由于

$$0\alpha=(0+0)\alpha=0\alpha+0\alpha,$$

所以由消去律得$0\alpha=0$, 性质1成立.

对于性质2, 由于

$$c0=c(0+0)=c0+c0,$$

得$c0=0$.

设$c\alpha=0$, 而$c \neq 0$. 则有

$$c^{-1}(c\alpha)=c^{-1}0=0 \Longrightarrow (c^{-1}c)\alpha=0 \Longrightarrow 1\alpha=0 \Longrightarrow \alpha=0,$$

即性质3成立.

最后, 由于

$$\alpha+(-1)\alpha=(1-1)\alpha=0\alpha=0,$$

得$-\alpha=(-1)\alpha$. □

习题6.1

A1. 以下的集合V关于所规定的运算是否成为数域$\mathbb{F}$上的线性空间?

(1) 设$A \in \mathbb{F}^{n\times n}$, V为A的矩阵多项式全体所组成的集合, 关于矩阵的加法和数乘;

(2) V为$\mathbb{F}$上可逆的n阶方阵全体所组成的集合, 关于矩阵的加法和数乘;

(3) 设$A \in \mathbb{F}^{n\times n}$, V为与A可交换的矩阵全体所组成的集合, 关于矩阵的加法和数乘;

(4) 取$\mathbb{F}=\mathbb{R}$, 取V为所有平面向量所组成的集合, 关于向量的加法和如下定义的数乘

$$c\alpha=0, \quad \forall c \in \mathbb{R}, \forall \alpha \in V;$$

(5) 取$\mathbb{F}=\mathbb{R}$, 取V为所有平面向量所组成的集合, 关于向量的加法和如下定义的数乘

$$c\alpha=\alpha,\quad \forall c\in\mathbb{R},\forall\alpha\in V;$$

(6) 取$\mathbb{F}=\mathbb{R}$, 取V为不平行于已知向量α的所有平面向量所组成的集合, 关于向量的加法和数乘;

(7) $V=\mathbb{F}^2$, 定义加法和数乘为

$$(x_1,x_2)^T+(y_1,y_2)^T=(x_1+y_1,x_2+y_2+x_1y_1)^T,$$
$$a(x_1,x_2)^T=\left(ax_1,ax_2+\frac{a(a-1)}{2}x_1^2\right)^T;$$

(8) 取$\mathbb{F}=\mathbb{R}$, V是满足$f(x^2)=f(x)^2$的实值函数$f(x)$全体做成的集合, 关于函数的加法和数乘;

(9) 取$\mathbb{F}=\mathbb{R}$, V是满足$f(-1)=0$的实值函数$f(x)$全体做成的集合, 关于函数的加法和数乘.

§ 6.2 子空间

在建立线性空间的坐标系前, 我们给出更多的例子. 可以看到, 线性空间的某些子集关于该空间的运算仍然成为线性空间, 这就是子空间的概念.

6.2.1 子空间的定义和例

何谓线性空间, 朴素地讲, 就是一些向量生活的空间, 它们自由地生活, 即它们可以自由地做空间中的两个运算. 如果这个空间的一部分也可以自由地做两个运算, 做运算得到的向量仍在这部分内, 不需要求助这部分外面的世界就可以认识它们, 则当然可以认为这个部分也成为一个线性空间, 这就是子空间. 严格地, 设V是$\mathbb{F}$上的线性空间.

定义6.2. *设W是V的非空子集, 如果W对V中的加法和数乘运算封闭, 即*

(i) *$\forall\alpha,\beta\in W$, 有$\alpha+\beta\in W$;*

(ii) *$\forall\alpha\in W$, $\forall c\in\mathbb{F}$, 有$c\alpha\in W$,*

则称W是V的一个子空间.

确实子空间是线性空间.

引理6.1. *设W是V的非空子集, 则*

(1) *W是V的子空间$\Longleftrightarrow\forall\alpha,\beta\in W$, $\forall a,b\in\mathbb{F}$, 有$a\alpha+b\beta\in W$;*

(2) *如果W是V的子空间, 则W在V的向量加法和数乘下成为$\mathbb{F}$上的线性空间.*

☞ **证明.** (1) 设W是V的子空间, 则由(ii)得$a\alpha,b\beta\in W$; 再由(i)得$a\alpha+b\beta\in W$.
反之, 如果对任意$\alpha,\beta\in W$和$a,b\in\mathbb{F}$, 有$a\alpha+b\beta\in W$, 则有

$$\alpha+\beta=1\alpha+1\beta\in W,\quad c\alpha=c\alpha+0\alpha\in W.$$

即W是子空间.
(2) 由子空间定义, W中有封闭的加法和数乘. 于是只要验证线性空间定义的八条公理成立. 因为在V满足(A1), (A2), (M1), (M2)和(D1), (D2), 所以在W中这些也成立, 最后需要证明W中(A3)和(A4)成立. 事实上, 设$\alpha \in W$, 则

$$-\alpha = (-1)\alpha \in W,$$

进而

$$0 = \alpha + (-\alpha) \in W,$$

于是(A3)和(A4)成立. □

任意线性空间V都有子空间: 零子空间$\{0\}$和V, 称它们为V的平凡子空间, 其余的子空间称为非平凡子空间; 称不等于V的V的子空间为V的真子空间.

下面举一些子空间的例子, 由此可以看到更多有意思的线性空间.

例6.7. 实线性空间$\mathbb{C}$有子空间$\mathbb{R}$.

例6.8. 设$V = \mathbb{F}^n$, 则V的子空间就是我们前面定义的向量空间. 例如, 对矩阵$A \in \mathbb{F}^{m\times n}$, 齐次线性方程组$AX = 0$的解空间就是$\mathbb{F}^n$的子空间.

例6.9. 设$V = \mathbb{F}^{n\times n}$, 则$V$中所有$n$阶对角阵所成的子集是$V$的子空间, 所有$n$阶上(下)三角阵所成的子集也是$V$的子空间.

例6.10. 设$V = \mathbb{F}^{n\times n}$, 定义子集$\mathfrak{sl}(n, \mathbb{F}) = \{A \in V \mid \operatorname{Tr}(A) = 0\}$[1]. 证明: W是V的子空间.

☞ **证明.** 由于$\operatorname{Tr}(O) = 0$, 所以$O \in \mathfrak{sl}(n, \mathbb{F})$, $\mathfrak{sl}(n, \mathbb{F}) \neq \varnothing$. 任取$A, B \in \mathfrak{sl}(n, \mathbb{F})$, 则$\operatorname{Tr}(A) = \operatorname{Tr}(B) = 0$. 于是对任意$a, b \in \mathbb{F}$, 有

$$\operatorname{Tr}(aA + bB) = a\operatorname{Tr}(A) + b\operatorname{Tr}(B) = a0 + b0 = 0.$$

这得到$aA + bB \in \mathfrak{sl}(n, \mathbb{F})$. 所以$\mathfrak{sl}(n, \mathbb{F})$是$V$的子空间. □

例6.11. 设$V = \mathbb{F}[x]$, 对于$n \in \mathbb{N}$, 记$\mathbb{F}_n[x]$为$\mathbb{F}[x]$中所有次数不超过n的多项式所成的集合, 即

$$\mathbb{F}_n[x] := \{f(x) \in \mathbb{F}[x] \mid \deg(f(x)) \leqslant n\},$$

则$\mathbb{F}_n[x]$是$\mathbb{F}[x]$的子空间.

上例中, 如果取

$$W = \{f(x) \in \mathbb{F}[x] \mid \deg(f(x)) = n\},$$

则W不是$\mathbb{F}[x]$的子空间. 这比如可从$0 \notin W$得出.

例6.12. 设V是闭区间$[a, b]$上所有实值函数做成的线性空间, 则

$$C[a, b] := \{f(x) \in V \mid f(x) \text{ 连续}\}$$

和

$$C^1[a, b] := \{f(x) \in V \mid f(x) \text{ 可导}\}$$

都是V的子空间.

[1] $\mathfrak{sl}(n, \mathbb{F})$不但是线性空间, 还是单李代数.

例6.13. (1) 设$V=\mathbb{R}^2$是$Euclid$平面, 则$\mathbb{R}^2$的非平凡子空间恰为所有过原点的直线; 不过原点的直线不是V的子空间;

(2) 设$V=\mathbb{R}^3$是$Euclid$三维空间, 则$\mathbb{R}^3$的非平凡子空间恰为所有过原点的直线和平面; 不过原点的直线和平面不是V的子空间.

☞ **证明.** (1) 容易验证过原点的直线是V的子空间. 反之, 如果W是V的非平凡子空间, 则存在$0\neq\alpha\in W$. 令W'为α所在的过原点的直线, 则$W'\subset W$. 若存在$\beta\in W$, 而$\beta\notin W'$, 则$\beta\neq 0$, 且对任意$0\neq\gamma\in\mathbb{R}^2$, 从$\gamma$的终点分别做$W'$和$\beta$所在直线的平行线, 可知(请画图)

$$\gamma=a\alpha+b\beta\in W,\qquad(\exists a,b\in\mathbb{R}).$$

于是就得到$W=V$, 矛盾. 所以$W=W'$是过原点的直线[2].

(2) 容易验证过原点的直线和平面是V的子空间. 反之, 如果W是V的非平凡子空间, 则存在$0\neq\alpha\in W$. 如果W等于α所在的过原点的直线, 则证明结束. 否则, 存在$\beta\in W$, 而β不在α所在的过原点的直线上. 如果W等于α和β所在的过原点的平面, 则证明结束. 否则存在这个平面外的W的向量γ. 此时容易证明$\mathbb{R}^3=W$, 矛盾. □

6.2.2 子空间的运算

子空间作为集合, 当然可以做交和并运算, 我们来研究在这些运算下是否还是子空间, 由此可以得到更多的子空间.

- 子空间的交

设V是线性空间, $W_i\ (i\in I)$是V的子空间, 其中I是指标集(可能为无限集), 则有这些子空间的交

$$\bigcap_{i\in I}W_i=\{\alpha\in V\mid\alpha\in W_i,\forall i\in I\}.$$

当I为有限集, 比如$I=\{1,2,\ldots,m\}$时, 记

$$\bigcap_{i\in I}W_i=\bigcap_{i=1}^{m}W_i=W_1\cap W_2\cap\cdots\cap W_m.$$

子空间的交是否还是子空间呢? 用几何例子想一下, 考察Eucild平面$\mathbb{R}^2$的子空间的交. 此时, 非平凡的子空间是过原点的直线, 若干个不同的过原点的直线的交是原点, 是子空间. 类似地, 容易得到Eucild三维空间$\mathbb{R}^3$的任意子空间的交仍是子空间. 所以我们猜想子空间的交仍是子空间. 事实上, 设$W_i\ (i\in I)$是V的子空间, 记$W=\bigcap\limits_{i\in I}W_i$.

由于$\forall i\in I$, $0\in W_i$, 所以$0\in W$, 即$W\neq\varnothing$.

又设$\alpha,\beta\in W$, $a,b\in\mathbb{F}$, 则对任意$i\in I$, 有$\alpha,\beta\in W_i$. 由于W_i是子空间, 所以

$$a\alpha+b\beta\in W_i,\quad\forall i\in I.$$

于是有$a\alpha+b\beta\in W$. 所以W是V的子空间.

命题6.2. 任意多个子空间的交仍是子空间.

[2]也可以用维数证明.

例如, 设$W_1 \subset \mathbb{F}^n$为齐次线性方程组$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$的解空间, $W_2 \subset \mathbb{F}^n$是齐次线性方程组$\begin{cases} b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0, \\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0, \\ \qquad\vdots \\ b_{l1}x_1+b_{l2}x_2+\cdots+b_{ln}x_n=0 \end{cases}$的解空间, 则$W_1 \cap W_2$是齐次线性方程组$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ \qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0, \\ b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0, \\ \qquad\vdots \\ b_{l1}x_1+b_{l2}x_2+\cdots+b_{ln}x_n=0 \end{cases}$的解空间.

- **子空间的并**

设W_1, W_2是线性空间V的子空间, 则有集合并

$$W_1 \cup W_2 = \{\alpha \in V \mid \alpha \in W_1 \text{ 或 } \alpha \in W_2\}.$$

那么, $W_1 \cup W_2$是否还是V的子空间? 还是用几何例子想一下, 取$V = \mathbb{R}^2$, W_1为x轴, W_2为y轴(如图6.1). 由于$W_1 \cup W_2$不是过原点的直线, 所以不是$\mathbb{R}^2$的子空间(也可如下证明: $(1,0)^T, (0,1)^T \in W_1 \cup W_2$, 但是$(1,0)^T + (0,1)^T = (1,1)^T \notin W_1 \cup W_2$). 于是, 子空间的并通常不再是子空间.

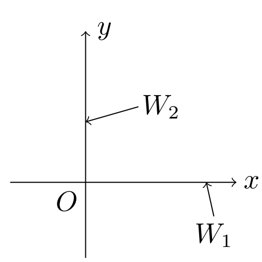


图 6.1: x轴和y轴的并不是子空间

- **子空间的和**

上面$V = \mathbb{R}^2$的子空间W_1和W_2的并$W_1 \cup W_2$不是子空间, 是由于存在$\alpha \in W_1$和$\beta \in W_2$, 使得$\alpha + \beta \notin W_1 \cup W_2$. 为了得到子空间, 我们应该把这些不认识的向量也加进去, 于是需要考虑子空间的和.

定义6.3. 设$W_1, W_2, \ldots, W_m$是线性空间V的子空间, 定义子空间$W_1, W_2, \ldots, W_m$的**和**为

$$W_1 + W_2 + \cdots + W_m := \{\alpha_1 + \alpha_2 + \cdots + \alpha_m \mid \alpha_1 \in W_1, \alpha_2 \in W_2, \ldots, \alpha_m \in W_m\}.$$

记$W = W_1 + W_2 + \cdots + W_m$, 由于

$$0 = 0 + 0 + \cdots + 0 \in W,$$

所以$W \neq \varnothing$. 又设$\alpha, \beta \in W$, $a, b \in \mathbb{F}$, 则

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m,\ \beta = \beta_1 + \beta_2 + \cdots + \beta_m,\ (\exists \alpha_i, \beta_i \in W_i, i = 1, 2, \ldots, m).$$

所以

$$a\alpha + b\beta = (a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2) + \cdots + (a\alpha_m + b\beta_m).$$

因为W_i是子空间, $\alpha_i, \beta_i \in W_i$, 所以$a\alpha_i + b\beta_i \in W_i$. 因此$a\alpha + b\beta \in W$, 即得$W$是子空间. 这得到子空间的和还是子空间.

命题6.3. 设$W_1, W_2, \ldots, W_m$是线性空间V的子空间, 则$W_1 + W_2 + \cdots + W_m$是V的子空间, 且是V中包含$W_1, W_2, \ldots, W_m$的最小子空间.

☞ **证明.** 记$W = W_1 + W_2 + \cdots + W_m$, 则只要证$W$的最小性. 设$U$是$V$的子空间, 且$W_1, W_2, \ldots, W_m \subset U$. 任取$\alpha \in W$, 则

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m, \quad (\exists \alpha_i \in W_i, i = 1, 2, \ldots, m).$$

因为$W_i \subset U$, 所以$\alpha_i \in U$. 而U是子空间, 所以

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m \in U.$$

这就得到$W \subset U$. □

最后看一个例子.

例6.14. 设$V = \mathbb{R}^3$, 取W_1为过原点的直线, W_2为过原点的平面, 且设直线W_1不在平面W_2上(见图6.2), 则有

$$W_1 \cap W_2 = \{0\}, \quad W_1 + W_2 = \mathbb{R}^3.$$

☞ **证明.** 由条件可得$W_1 \cap W_2 = \{0\}$. 任取$\alpha \in \mathbb{R}^3$, 如图6.2, 过α的终点作W_1的平行线得$\alpha_2 \in W_2$; 过α的终点作与α_2平行且与W_1相交的直线, 得$\alpha_1 \in W_1$. 由平行四边形法则, 得$\alpha = \alpha_1 + \alpha_2$. 于是$\mathbb{R}^3 = W_1 + W_2$. □

习题6.2

A1. 下列$\mathbb{F}^n$的子集W是否是$\mathbb{F}^n$的子空间? 请给出理由.

(1) $W = \{(a_1, a_2, \ldots, a_n)^T \in \mathbb{F}^n \mid a_1 + a_2 + \cdots + a_n = 0\}$;

(2) $W = \{(a_1, a_2, \ldots, a_n)^T \in \mathbb{F}^n \mid a_1 + a_2 + \cdots + a_n = 1\}$;

(3) $W = \{(a_1, a_2, \ldots, a_n)^T \in \mathbb{F}^n \mid a_1, a_2, \ldots, a_n$不同时大于零且不同时小于零$\}$, 这里$n \geqslant 2$;

(4) $W = \{(a_1, a_2, \ldots, a_n)^T \in \mathbb{F}^n \mid$ 存在某个i, 使得$a_i > 0\}$.

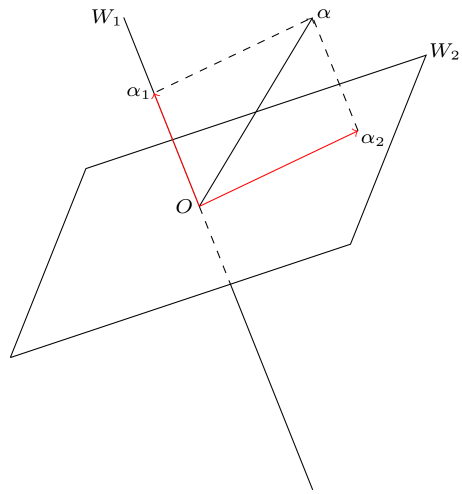


图 6.2: 直线和平面的交和和

A2. 设L_1, L_2和L_3是$\mathbb{R}^3$中过原点的三条直线, 它们都是$\mathbb{R}^3$的子空间. 问L_1+L_2和$L_1+L_2+L_3$可能构成$\mathbb{R}^3$的哪些类型的子空间?

A3. 设U, W_1, W_2是线性空间V的子空间, 证明[3]:

(1) $U\cap(W_1+W_2)\supset U\cap W_1+U\cap W_2$;

(2) $U\cap(W_1+U\cap W_2)=U\cap W_1+U\cap W_2$;

(3) $(U+W_1)\cap(U+W_2)=U+(U+W_1)\cap W_2$.

举例说明(1)中的等号不一定成立.

A4. 设W_1和W_2是线性空间V的子空间, 证明: 下列命题等价

(i) $W_1\cup W_2$是V的子空间;

(ii) $W_1\cup W_2=W_1+W_2$;

(iii) $W_1\subset W_2$或$W_2\subset W_1$.

B1. 证明: 数域$\mathbb{F}$上的线性空间$V(V\neq\{0\})$不能被它的有限个真子空间所覆盖, 即设W_1, W_2, $\ldots$, W_k是V的真子空间, 则存在$\alpha\in V$, 使得$\alpha\notin W_1\cup W_2\cup\cdots\cup W_k$.

§ 6.3　线性相关性

现在我们的脑子里已经有了许多具体的线性空间的例子, 线性空间已经不那么抽象了. 因此我们可以开始建立线性空间的坐标系. 要建立坐标系, 首先各坐标轴应该没有关系, 且个数不能再少, 于是就需要线性无关和极大无关组的概念; 其次, 我们的坐标系当然需要可以表示出空间中的每一个向量, 于是需要线性表示的概念. 本节先做这些建立坐标系的准备, 而把建立坐标系的任务留待下节. 这些概念是上一章中特殊的线性空间—向量空间中的相应概念的推广, 所以读者不该觉得奇怪和不自然.

[3]如果没有括号, 则集合运算次序是: 先算交和并, 再算加法.

6.3.1 线性相关性

设V是$\mathbb{F}$上的线性空间. 下面定义V中线性无关的概念. 和向量空间$\mathbb{F}^n$不同的是, V中可能会有无穷多个向量没有关系.

定义6.4. 设S是V的非空子集(可能为无限集).

(1) 如果存在$\alpha_1, \alpha_2, \ldots, \alpha_s \in S(\exists s \in \mathbb{N})$, 存在不全为零的$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0,$$

则称S线性相关;

(2) 如果S非线性相关, 即对任意$s \in \mathbb{N}$, 任意$\alpha_1, \alpha_2, \ldots, \alpha_s \in S$, 若

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s = 0, \quad (c_1, c_2, \ldots, c_s \in \mathbb{F}),$$

那么有$c_1 = c_2 = \cdots = c_s = 0$, 则称$S$线性无关.

如果$S = \{\alpha_1, \alpha_2, \ldots, \alpha_m\} \subset V$是有限集, 则容易证明:

(i) S线性相关$\Longleftrightarrow$ 存在不全为零的$c_1, c_2, \ldots, c_m \in \mathbb{F}$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m = 0.$$

(ii) S线性无关$\Longleftrightarrow$ 如果有$c_1, c_2, \ldots, c_m \in \mathbb{F}$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m = 0,$$

那么$c_1 = c_2 = \cdots = c_m = 0$.

我们看一个线性无关的无限集.

例6.15. 设$V = \mathbb{F}[x]$, $S = \{x^i \mid i = 0, 1, 2 \ldots\} \subset V$. 证明: S线性无关.

☞ **证明.** 任意$s \in \mathbb{N}$, 任意$x^{i_1}, x^{i_2}, \ldots, x^{i_s} \in S$, 其中$0 \leqslant i_1 < i_2 < \cdots < i_s$, 如果

$$a_1x^{i_1} + a_2x^{i_2} + \cdots + a_sx^{i_s} = 0, \quad (a_1, a_2, \ldots, a_s \in \mathbb{F}),$$

则由多项式相等的定义得$a_1 = a_2 = \cdots = a_s = 0$. 所以$S$线性无关. □

由定义容易证明下面性质成立:

性质1 若$0 \in S \subset V$, 则S线性相关.

性质2 若$\varnothing \neq S_1 \subset S_2 \subset V$, 则

$$S_1\text{ 线性相关} \Longrightarrow S_2\text{ 线性相关},$$
$$S_2\text{ 线性无关} \Longrightarrow S_1\text{ 线性无关}.$$

6.3.2 线性表示

设V是$\mathbb{F}$上的线性空间, 我们来定义线性表示的概念. 首先是一个向量由某个向量集线性表示的概念.

定义6.5. *设S是V的非空子集, $\alpha \in V$. 如果存在$\alpha_1, \alpha_2, \ldots, \alpha_s \in S(\exists s \in \mathbb{N})$, 存在$c_1, c_2, \ldots, c_s \in \mathbb{F}$, 使得*

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s,$$

则称α是S的线性组合, 或称α可由S线性表示.

如果$S = \{\alpha_1, \alpha_2, \ldots, \alpha_m\} \subset V$是有限集, 容易证明: $\alpha \in V$是S的线性组合当且仅当存在$c_1, c_2, \ldots, c_m \in \mathbb{F}$, 使得

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m.$$

然后是两个向量集之间线性表示的概念.

定义6.6. *设S_1, S_2是V的非空子集.*

(1) *如果任意$\alpha \in S_1$, α可由S_2线性表示, 则称S_1可由S_2线性表示, 记为$S_1 \leftarrow S_2$;*

(2) *如果$S_1 \leftarrow S_2$且$S_2 \leftarrow S_1$, 则称S_1和S_2等价, 记为$S_1 \leftrightarrow S_2$.*

容易证明, V的非空子集之间的$\leftrightarrow$关系为等价关系(我们把传递性留作练习), 即满足

(i) (自反性) $\varnothing \neq \forall S \subset V$, 有$S \leftrightarrow S$;

(ii) (对称性) $\varnothing \neq \forall S_1, S_2 \subset V$, 若$S_1 \leftrightarrow S_2$, 则$S_2 \leftrightarrow S_1$;

(iii) (传递性) $\varnothing \neq \forall S_1, S_2, S_3 \subset V$, 若$S_1 \leftrightarrow S_2$且$S_2 \leftrightarrow S_3$, 则$S_1 \leftrightarrow S_3$.

类似于向量空间, 可以证明下面性质(请读者在脑袋中想一下这些性质的证明):

性质1 如果S是V的至少含两个元素的子集, 则

$$S\text{ 线性相关} \Longleftrightarrow \exists \alpha \in S, \text{使得}\alpha\text{ 可由}S - \{\alpha\}\text{ 线性表示.}$$

性质2 如果S是V的线性无关子集, $\alpha \in V$可由S线性表示, 则α由S线性表示的方式唯一.

性质3 设S是V的线性无关子集, $\alpha \in V$. 如果$S \cup \{\alpha\}$线性相关, 则α可由S(唯一)线性表示.

性质4 设S_1, S_2是V的非空有限子集, 则

- **(1)** $S_1 \leftarrow S_2$且[4]$|S_1| > |S_2| \Longrightarrow S_1$线性相关;
- **(2)** S_1线性无关, $S_1 \leftarrow S_2 \Longrightarrow |S_1| \leqslant |S_2|$;
- **(3)** S_1和S_2都线性无关, $S_1 \leftrightarrow S_2 \Longrightarrow |S_1| = |S_2|$.

[4] 当B_1是有限集时, $|B_1|$表示集合B_1所含元素的个数; 当B_1是无限集时, $|B_1|$表示集合B_1的势. 注意无限集的势不一定相同, 例如, $\mathbb{Z}$的势是可数无穷, 而$\mathbb{R}$的势是不可数无穷, 我们有$|\mathbb{Z}| < |\mathbb{R}|$.

6.3.3 向量集生成的子空间

设V是$\mathbb{F}$上的线性空间, 我们希望V中的坐标系可以线性表示V中的每一个向量, 所以对于V的非空子集S, 我们要考虑S的所有线性组合所成的集合. 我们将这个集合记为[5]$\mathrm{Span}(S)$, 有时也记为$\mathrm{Span}_{\mathbb{F}}(S)$. 于是

$$\mathrm{Span}(S)=\{c_1\alpha_1+\cdots+c_s\alpha_s \mid s\geqslant 1,\alpha_1,\ldots,\alpha_s\in S,c_1,\ldots,c_s\in\mathbb{F}\}.$$

称$\mathrm{Span}(S)$为由S生成的子空间, S称为生成元集. 当$S=\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$时, 有

$$\mathrm{Span}(S)=\{c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m \mid c_1,c_2,\ldots,c_m\in\mathbb{F}\}.$$

如果$S=\{\alpha\}$, 有时也记$\mathrm{Span}(S)=\mathbb{F}\alpha$, 于是成立

$$\mathrm{Span}(\{\alpha_1,\alpha_2,\ldots,\alpha_m\})=\mathbb{F}\alpha_1+\mathbb{F}\alpha_2+\cdots+\mathbb{F}\alpha_m.$$

我们还约定, 对于$S=\varnothing$, 有

$$\mathrm{Span}(S)=\{0\}.$$

直观地想, $\mathrm{Span}(S)$就是S生活的世界, 而且这是最小的世界.

命题6.4. *设S是V的子集, 则$\mathrm{Span}(S)$是V的子空间, 且为V中包含S的最小子空间.*

证明. 可以不妨设$S\neq\varnothing$. 此时, 容易证明$\mathrm{Span}(S)$是V的子空间, 这里省略. 任取V的子空间U, 使得$S\subset U$, 我们要证明$\mathrm{Span}(S)\subset U$.

事实上, 任意的$\alpha\in\mathrm{Span}(S)$有

$$\alpha=c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s,\quad (\exists\alpha_1,\alpha_2,\ldots,\alpha_s\in S,\exists c_1,c_2,\ldots,c_s\in\mathbb{F}).$$

因为$S\subset U$, 所以$\alpha_1,\alpha_2,\ldots,\alpha_s\in U$. 而$U$是子空间, 所以

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s\in U,$$

即$\alpha\in U$. 于是$\mathrm{Span}(S)\subset U$. □

最后, 我们说明线性表示也可以用向量组生成的子空间来描述.

命题6.5. *设S,S_1,S_2是V的非空子集, $\alpha\in V$, 则*

(1) *α可由S线性表示$\iff \alpha\in\mathrm{Span}(S)$;*

(2) *$S_1\leftarrow S_2 \iff \mathrm{Span}(S_1)\subset\mathrm{Span}(S_2)$;*

(3) *$S_1\leftrightarrow S_2 \iff \mathrm{Span}(S_1)=\mathrm{Span}(S_2)$.*

证明. (1)就是定义, (3)由(2)导出, 下证(2). 如果$S_1\leftarrow S_2$, 则$\forall\alpha\in S_1$, α可由S_2表示, 于是$\alpha\in\mathrm{Span}(S_2)$. 这就得到$S_1\subset\mathrm{Span}(S_2)$. 由于$\mathrm{Span}(S_2)$是子空间, $\mathrm{Span}(S_1)$是包含S_1的最小子空间(命题6.4), 所以$\mathrm{Span}(S_1)\subset\mathrm{Span}(S_2)$. 反之, 如果$\mathrm{Span}(S_1)\subset\mathrm{Span}(S_2)$, 则$\forall\alpha\in S_1\subset\mathrm{Span}(S_1)$, 有$\alpha\in\mathrm{Span}(S_2)$, 即$\alpha$可由$S_2$线性表示. 于是$S_1\leftarrow S_2$. □

---

[5] 也记为$\langle S\rangle$.

习题6.3

A1. 设S_1, S_2, S_3为线性空间V的非空子集, 如果S_1可由S_2线性表示, S_2可由S_3线性表示, 证明: S_1可由S_3线性表示.

A2. 设S_1和S_2是线性空间V的两个非空子集, 证明:

(1) 如果$S_1 \subset S_2$, 则$\mathrm{Span}(S_1) \subset \mathrm{Span}(S_2)$;

(2) $\mathrm{Span}(S_1 \cup S_2) = \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$;

(3) $\mathrm{Span}(S_1 \cap S_2) \subset \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$. 举例说明等号不一定成立.

A3. 设V是所有实值函数构成的实数域$\mathbb{R}$上的线性空间, 判断V中的下列向量组的线性相关性, 并证明你的结论.

(1) $1, \sin^2 x, \cos^2 x$; (2) xe^x, e^{2x}; (3) $\sin x, e^x$.

A4. 设$V = \mathbb{F}[x]$, $f_1(x), f_2(x), f_3(x) \in V$满足: 这三个多项式互素, 但是其中任意两个都不互素. 证明: $f_1(x), f_2(x), f_3(x)$线性无关.

A5. 取集合V为实数域$\mathbb{R}$, 数域为有理数域$\mathbb{Q}$. 集合V的向量加法规定为实数的加法, 纯量与向量的乘法规定为有理数与实数的乘法, 则V成为有理数域$\mathbb{Q}$上的线性空间. 证明: 在线性空间V中, 实数1与α线性无关的必要且充分条件是, α为无理数.

B1. 设V是所有连续实值函数构成的实数域$\mathbb{R}$上的线性空间, 证明:

(1) 向量$\sin x$, $\sin 2x$, $\ldots$, $\sin nx$, $\ldots$线性无关;

(2) 向量1, $\cos x$, $\cos 2x$, $\ldots$, $\cos nx$, $\ldots$线性无关;

(3) 向量1, $\sin x$, $\cos x$, $\sin 2x$, $\cos 2x$, $\ldots$, $\sin nx$, $\cos nx$, $\ldots$线性无关.

§ 6.4 基与维数

现在我们来建立任意线性空间的坐标系.

6.4.1 基与维数

设V是$\mathbb{F}$上的线性空间, V的坐标系即基的定义是自然的: 坐标轴向量应该没有关系, 可以表示空间中每个向量.

定义6.7. *设B是V的非空子集, 满足*

(i) *B线性无关;*

(ii) $\mathrm{Span}(B) = V$ *($\Longleftrightarrow B \leftrightarrow V \Longleftrightarrow \forall \alpha \in V$, α可由B线性表示),*

则称B是V的一个(一组)基.

由定义, V的非空子集B是V的基的充分必要条件是: $\forall \alpha \in V$, α可由B唯一地线性表示. 我们将其证明留作练习.

我们要解决的第一个基本问题是: 线性空间是否都有基? 当然, 当$V = \{0\}$时, V没有基; 那么$V \neq \{0\}$时呢? 结论自然应该是存在基, 否则我们就不会在这里啰嗦了. 让我们回忆一下在向量空

间我们是如何证明极大无关组的存在性的. 假设已经有一个线性无关的子向量组, 我们把它进行线性无关扩充, 直到不能再线性无关扩充为止; 由于$\mathbb{F}^n$中任意$n+1$个向量一定有关系, 所以线性无关扩充的过程有限步后一定停止, 就得到一个极大无关组. 类似地, 对于线性空间V, 假设已有一个线性无关的子集, 我们对它进行线性无关扩充; 但这时和向量空间不一样, 我们会遇到有无穷多个向量的线性无关子集(例6.15), 所以这个线性无关扩充的过程可能会一直进行下去, 停不下来! 怎么办呢?

为了克服停不下来的困难, 数学家们给出了办法. 下面我们介绍集合论中的**Zorn引理**, 它与选择公理等价.

设S是一个集合, 如果S上有二元关系$\preccurlyeq$, 满足性质

(i) (自反性) $\forall a \in S$, 有$a \preccurlyeq a$;

(ii) (反对称性) $\forall a, b \in S$, $a \preccurlyeq b$且$b \preccurlyeq a \Longrightarrow a = b$;

(iii) (传递性) $\forall a, b, c \in S$, $a \preccurlyeq b$且$b \preccurlyeq c \Longrightarrow a \preccurlyeq c$,

则称$(S, \preccurlyeq)$是偏序集, 简称S是偏序集; 称关系$\preccurlyeq$为S上的偏序关系. 例如, 集合的包含关系$\subset$是偏序关系.

设S是一个偏序集. 元素$a, b \in S$称为是可比较的, 如果$a \preccurlyeq b$或者$b \preccurlyeq a$. 设$a \in S$, T是S的非空子集, 我们定义

(1) 如果$a \in T$, 且对每个与a可比较的$b \in T$, 必有$b \preccurlyeq a$, 则称a为T中的极大元;

(2) 如果对任意$b \in T$有$b \preccurlyeq a$, 则称a是T的上界;

(3) 如果T中任意两个元素都可比较, 则称T是S中的一个链.

下面可以陈述Zorn引理.

定理6.6 (Zorn引理). *设S是偏序集, 如果S的每个链都有上界, 则S有极大元.*

利用Zorn引理就可以证明非零线性空间基的存在性, 事实上我们证明更重要的基扩充定理.

定理6.7 (基扩充定理). *设B是线性空间V的线性无关子集, 则存在V的基B_1, 使得$B \subset B_1$. 特别的, 当V非零时, V一定有基.*

☞ **证明.** 记S为V中所有包含B的线性无关子集所成的集合, 即

$$S = \{B' \subset V \mid B'\text{线性无关且}B \subset B'\}.$$

由于$B \in S$, 所以$S \neq \varnothing$. 集合S关于集合的包含关系成为偏序集, 任取S的一个链T, 令

$$B_2 = \bigcup_{X \in T} X \subset V.$$

由于$\forall X \in T$, 有$B \subset X$, 所以$B \subset B_2$. 如果B_2线性相关, 则存在$\alpha_1, \alpha_2, \ldots, \alpha_m \in B_2$, 存在不全为零的$c_1$, c_2, $\ldots$, $c_m \in \mathbb{F}$, 使得

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m = 0.$$

由于$\alpha_i \in B_2$, 所以存在$X_i \in T$, 使得$\alpha_i \in X_i$, $i = 1, 2, \ldots, m$. 而T是S中的链, 所以必要时重新下标可以不妨设

$$X_1 \subset X_2 \subset \cdots \subset X_m.$$

这得到$\alpha_1, \alpha_2, \ldots, \alpha_m \in X_m$, 所以$X_m$线性相关, 矛盾于$X_m \in S$. 于是$B_2$线性无关, 即有$B_2 \in S$. 因此$B_2$是$T$的上界. 于是我们证明了$S$的每个链都有上界, 由Zorn引理, S有极大元B_1.

因为$B_1 \in S$, 所以B_1线性无关. 如果存在$\alpha \in V$, 使得α不能用B_1线性表示, 则$B_1 \cup \{\alpha\}$线性无关. 于是

$$B \subset B_1 \subsetneqq B_1 \cup \{\alpha\} \in S,$$

这矛盾于B_1的极大性. 所以V中任意向量可由B_1线性表示, 即B_1是包含B的V的基.

当$V \neq \{0\}$时, 任取$0 \neq \alpha \in V$, 则$\{\alpha\}$是V的线性无关子集, 可扩充成V的基, 即V有基. □

和向量空间一样, 非零线性空间的基不唯一, 但是基所含向量“个数”是唯一的.

引理6.8. 设B_1, B_2是向量空间V的任意两组基, 则$|B_1| = |B_2|$.

☞ **证明.** 首先, 我们证明如下事实: 如果V有有限基B, 则对于V的任意线性无关子集S, 有$|S| \leqslant |B|$.

事实上, 记$|B| = n$. 如果存在V的线性无关子集S, 使得$|S| > |B| = n$, 则S有子集S_1, 使得$|S_1| = n + 1$. 因为S线性无关, 所以S_1也线性无关. 而B是基, 所以$S_1 \leftarrow B$. 这得到

$$n + 1 = |S_1| \leqslant |B| = n,$$

矛盾.

下面回到引理的证明. 如果B_1或者B_2是有限集, 则由上面的事实, B_1和B_2都是有限集. 因为B_1和B_2都是基, 所以$B_1 \leftrightarrow B_2$. 这得到$|B_1| = |B_2|$.

如果B_1和B_2都是无限集, 则也可证明它们的势相等(参见本章附录). □

于是可以定义线性空间的维数.

定义6.8. 设V是$\mathbb{F}$上的线性空间, 则V的维数定义为

$$\dim V = \dim_{\mathbb{F}} V := \begin{cases} 0, & V = \{0\}, \\ |B|, & V\text{有基}B. \end{cases}$$

当$\dim V < \infty$时, 称V是有限维线性空间; 否则, 称V是无限维线性空间, 这时也简单记$\dim V = \infty$.

如果无特别说明, 我们都约定$\dim V > 0$, 即V有基. 和向量空间一样, 如果$\dim V = n$, 则有

(1) V中任意$n + 1$个向量线性相关;

(2) 对任意$\alpha_1, \alpha_2, \ldots, \alpha_n \in V$, 下面等价

- **(i)** $\alpha_1, \alpha_2, \ldots, \alpha_n$是$V$的基[6];
- **(ii)** $\alpha_1, \alpha_2, \ldots, \alpha_n$线性无关;
- **(iii)** $\forall \alpha \in V$, α可由$\alpha_1, \alpha_2, \ldots, \alpha_n$线性表示.

最后, 我们定义坐标的概念.

[6]写基时常常省略集合记号的大括号.

定义6.9. 设$\alpha_1,\alpha_2,\ldots,\alpha_n$是$n$维线性空间$V$的基, $\alpha\in V$, 则α可唯一写为

$$\alpha=c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n,\quad (c_1,c_2,\ldots,c_n\in\mathbb{F}).$$

称$\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}\in\mathbb{F}^n$为$\alpha$在$V$的(有序)基$\alpha_1,\alpha_2,\ldots,\alpha_n$下的坐标.

设$\dim V=n$, $\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的基, 则有双射

$$\varphi:V\longrightarrow\mathbb{F}^n;\alpha=c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n\mapsto\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}.$$

而且容易证明, 对任意$\alpha,\beta\in V$和$c\in\mathbb{F}$, 有

$$\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta),\quad \varphi(c\alpha)=c\varphi(\alpha),$$

即向量和的坐标等于坐标的和, 而向量数乘的坐标等于坐标的数乘. 于是取了基后, n维线性空间V可以等同于$\mathbb{F}^n$, 即我们在做解系几何.

6.4.2 例

我们看一些具体的线性空间的维数以及常取的基.

例6.16. 复数域$\mathbb{C}$作为复线性空间有基1, 维数$\dim_{\mathbb{C}}\mathbb{C}=1$; 而作为实线性空间有基$1,i$, 维数$\dim_{\mathbb{R}}\mathbb{C}=2$.

例6.17. n维列向量空间$\mathbb{F}^n$有自然基$e_1,e_2,\ldots,e_n$, 维数$\dim\mathbb{F}^n=n$.

例6.18. 设E_{ij}是(i,j)位置为1, 其余位置为0的$m\times n$矩阵, 证明: 矩阵空间$\mathbb{F}^{m\times n}$有自然基

$$\{E_{ij}\mid 1\leqslant i\leqslant m,1\leqslant j\leqslant n\},$$

进而维数$\dim\mathbb{F}^{m\times n}=mn$.

☞ **证明.** 如果数$a_{ij}\in\mathbb{F}$ $(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$使得

$$\sum_{i,j}a_{ij}E_{ij}=O,$$

则

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

于是$a_{ij}=0$, $(\forall i,j)$. 这得到$\{E_{ij}\mid 1\leqslant i\leqslant m, 1\leqslant j\leqslant n\}$线性无关. 这组向量生成$\mathbb{F}^{m\times n}$, 是由于对任意$A=(a_{ij})\in\mathbb{F}^{m\times n}$有

$$A=\sum_{i,j}a_{ij}E_{ij}.$$

所以$\{E_{ij}\mid 1\leqslant i\leqslant m, 1\leqslant j\leqslant n\}$是$\mathbb{F}^{m\times n}$的基. □

例6.19. 证明*:*

(1) 多项式空间$\mathbb{F}[x]$有自然基

$$\{x^i\mid i=0,1,2,\ldots\},$$

于是维数$\dim\mathbb{F}[x]=\infty$;

(2) $\mathbb{F}[x]$的子空间$\mathbb{F}_n[x]=\{f(x)\in\mathbb{F}[x]\mid \deg(f(x))\leqslant n\}$有自然基

$$1,x,x^2,\ldots,x^n,$$

于是维数$\dim\mathbb{F}_n[x]=n+1$. 任意$f(x)=b_0+b_1x+\cdots+b_nx^n\in\mathbb{F}_n[x]$在自然基下的坐标是$(b_0,b_1,\ldots,b_n)^T$. 注意到

$$b_i=\frac{f^{(i)}(0)}{i!}.$$

☞ **证明.** (1) 已证$B=\{x^i\mid i=0,1,2,\ldots\}$线性无关, 而任意$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in\mathbb{F}[x]$, 有$f(x)\in\mathrm{Span}(B)$. 于是$B$是$\mathbb{F}[x]$的基. □

例6.20. 设$V=\mathbb{F}_n[x]$. 取两两不相等的数$a_0,a_1,\ldots,a_n\in\mathbb{F}$, 对$i=0,1,\ldots,n$定义

$$f_i(x)=\frac{(x-a_0)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_0)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)}=\prod_{\substack{j=0\\ j\neq i}}^{n}\frac{x-a_j}{a_i-a_j},$$

称之为**Lagrange多项式**. 证明*:* $f_0(x),f_1(x),\ldots,f_n(x)$是$\mathbb{F}_n[x]$的基, 且任意$f(x)\in\mathbb{F}_n[x]$在这组基下的坐标是

$$(f(a_0),f(a_1),\ldots,f(a_n))^T.$$

☞ **证明.** 首先注意到

$$f_i(a_j)=\delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

如果

$$c_0f_0(x)+c_1f_1(x)+\cdots+c_nf_n(x)=0,\quad (c_0,c_1,\ldots,c_n\in\mathbb{F}),$$

则令$x=a_i$, 就得$c_i=0$. 于是$f_0(x),f_1(x),\ldots,f_n(x)$线性无关. 又$\dim\mathbb{F}_n[x]=n+1$, 所以$f_0(x)$, $f_1(x)$, …, $f_n(x)$就是$\mathbb{F}_n[x]$的基.

设

$$f(x)=c_0f_0(x)+c_1f_1(x)+\cdots+c_nf_n(x),$$

令$x=a_i$, 得$f(a_i)=c_i$. 所以$f(x)$在这组基下的坐标是$(f(a_0),f(a_1),\ldots,f(a_n))^T$. □

所谓的插值问题是: 任给互不相同的$a_0, a_1, \dots, a_n \in \mathbb{F}$, 任给$b_0, b_1, \dots, b_n \in \mathbb{F}$, 找$f(x) \in \mathbb{F}_n[x]$, 使得

$$f(a_i) = b_i, \quad (i = 0, 1, 2 \dots, n).$$

容易知道, 这样的$f(x) \in \mathbb{F}_n[x]$是唯一的, 由例6.20得

$$f(x) = \sum_{i=0}^{n} b_i f_i(x),$$

称这个公式为**Lagrange插值公式**.

例6.21. 设$\alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $\alpha_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $\alpha_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$, 证明: α_1, α_2, α_3, α_4是$\mathbb{R}^4$的基, 并求$\xi = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$在这组基下的坐标.

☞ **解.** 由于

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \xi) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix} \underbrace{\overset{r_2 - r_1}{\scriptstyle r_3 - r_1}}_{r_4 - r_1} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & 1 \\ 0 & -2 & 0 & -2 & 0 \\ 0 & -2 & -2 & 0 & 0 \end{pmatrix}$$

$$\underbrace{\overset{-\frac{1}{2}r_2}{\scriptstyle -\frac{1}{2}r_3}}_{-\frac{1}{2}r_4} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \underbrace{\overset{r_1 - r_3}{\scriptstyle r_4 - r_3}}_{r_2 \leftrightarrow r_3} \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \underbrace{r_1 - r_3}_{r_4 - r_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & -1 & \frac{3}{2} \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & -2 & \frac{1}{2} \end{pmatrix} \underbrace{-\frac{1}{2}r_4} \begin{pmatrix} 1 & 0 & 0 & -1 & \frac{3}{2} \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{pmatrix} \underbrace{\overset{r_1 + r_4}{\scriptstyle r_2 - r_4}}_{r_3 - r_4} \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{5}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{pmatrix},$$

所以$R(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = 4$且

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)^{-1}\xi = \begin{pmatrix} \frac{5}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{pmatrix}.$$

于是$\alpha_1, \alpha_2, \alpha_3, \alpha_4$线性无关, 是$\mathbb{R}^4$的基, 且$\xi$在这组基下的坐标是$\begin{pmatrix} \frac{5}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{pmatrix}$.

注意到这四个向量的特殊形式, 也可如下解. 记$A=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$, 则有$A^2=4E$, 于是$A$可逆, 且$A^{-1}=\frac{1}{4}A$. 这得到$\alpha_1,\alpha_2,\alpha_3,\alpha_4$线性无关, 是$\mathbb{R}^4$的基. 且$\xi$在这组基下的坐标是

$$A^{-1}\xi=\frac{1}{4}A\xi=\frac{1}{4}\begin{pmatrix}5\\1\\-1\\-1\end{pmatrix}.$$

解毕. □

6.4.3 维数公式

设V是$\mathbb{F}$上的线性空间, 本节最后, 我们讨论V的子空间的维数计算. 我们学过构造子空间的几种方法, 一个是对子空间进行交和和运算得到新的子空间, 另一个是非空子集可以生成子空间. 对前者, 我们看子空间的交和和的维数和原子空间维数的关系; 对后者, 我们看非空子集生成的子空间的维数如何计算.

我们先考虑后面一个问题. 设S是线性空间V的非空子集, 如何计算$\dim\operatorname{Span}(S)$, 并找$\operatorname{Span}(S)$的一个自然基? 维数本质上是数向量的个数, 数向量的个数时, 多余的向量要删去. 于是本质上只要数S中有几个真正的向量, 即S的极大无关组所含向量的个数. 当然, 我们还没有定义非空向量集的极大无关组的概念, 这里定义一下, 这是向量空间中相应概念的自然推广.

定义6.10. *设S是V的非空子集.*

(1) *如果S的非空子集S_1满足*

  (i) *S_1线性无关;*

  (ii) *$S_1\leftrightarrow S(\Longleftrightarrow\forall\alpha\in S$, α可由S_1线性表示),*

  *则称S_1是S的一个*极大无关组.

(2) *S的*秩*定义为*

$$R(S)=\begin{cases}0, & S=\{0\},\\ |S_1|, & S\text{ 有极大无关组 }S_1.\end{cases}$$

类似于向量空间, 利用Zorn引理可以证明当$S\neq\{0\}$时一定有极大无关组; 且如果S有两个极大无关组S_1和S_2, 则必有$|S_1|=|S_2|$. 当极大无关组含无穷多个向量时, 也简单记$R(S)=\infty$.

极大无关组是S中真正的向量, 于是下面关于线性空间$\operatorname{Span}(S)$的结论是自然的.

命题6.9. *设$S\neq\{0\}$是V的非空子集, 则S的极大无关组为$\operatorname{Span}(S)$的基, 且*

$$\dim\operatorname{Span}(S)=R(S).$$

☞ **证明.** 任取S的极大无关组S_1, 则$S_1\leftrightarrow S$. 于是$\operatorname{Span}(S)=\operatorname{Span}(S_1)$. 因为$S_1$线性无关, 所以$S_1$是$\operatorname{Span}(S)$的基. □

我们再考虑前一个问题. 设W_1和W_2都是V的子空间, 那么

$$\dim(W_1 \cap W_2), \quad \dim(W_1 + W_2), \quad \dim W_1, \quad \dim W_2$$

有什么关系? 我们还是考察一个具体的几何例子. 设$V = \mathbb{R}^3$, 取W_1为Oxy坐标平面, W_2为Oz轴, W_3为Oy 轴(参看图6.3). 由于

$$W_1 \cap W_2 = \{0\}, \quad W_1 + W_2 = \mathbb{R}^3, \quad W_1 \cap W_3 = W_3, \quad W_1 + W_3 = W_1,$$

所以

$$\dim(W_1 \cap W_2) = 0, \quad \dim(W_1 + W_2) = 3, \quad \dim(W_1 \cap W_3) = 1, \quad \dim(W_1 + W_3) = 2.$$

可见, 子空间的交和和的维数和它们的位置关系有关, 好像没有统一的公式. 但上面例子中交和和的维数之和都等于3, 是原来两个子空间的维数之和.

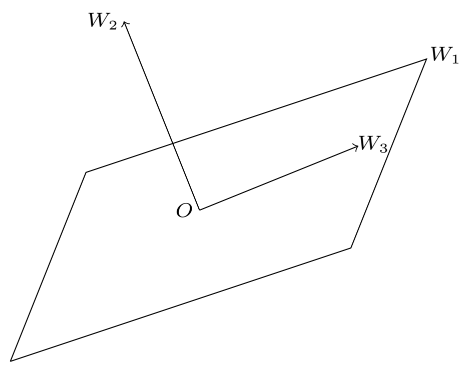


图 6.3: 子空间的交和和的维数

我们知道, 如果A和B都是有限集, 那么有下面的公式(参看图6.4)

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

如果我们用基来替代每个子空间, 则上面例子得到的结论有点像这个公式. 这个公式对任意有限集成立, 于是我们有理由相信成立下面的维数公式, 它是该公式在线性空间上的类比.

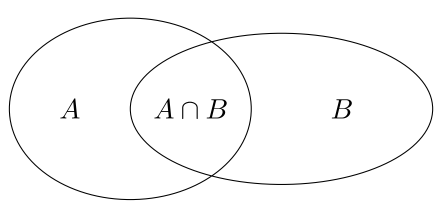


图 6.4: 集合的交和并

定理6.10 (维数公式)**.** 设W_1和W_2是线性空间V的有限维子空间, 则

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

证明. 请注意这里的基扩充的证明方法. 设$\dim W_1\cap W_2=t$, 取$W_1\cap W_2$的基$\alpha_1,\alpha_2,\ldots,\alpha_t$. 由于$W_1\cap W_2\subset W_1$, 所以可将$W_1\cap W_2$的基$\alpha_1,\alpha_2,\ldots,\alpha_t$扩充成$W_1$的基

$$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_r.$$

于是, $\dim W_1=t+r$. 类似地, 可得W_2的基

$$\alpha_1,\ldots,\alpha_t,\gamma_1,\ldots,\gamma_s.$$

于是$\dim W_2=t+s$. 下证: $\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_r,\gamma_1,\ldots,\gamma_s$是$W_1+W_2$的基.

首先, 这些向量都属于W_1+W_2. 其次, 如果

$$a_1\alpha_1+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_r\beta_r+c_1\gamma_1+\cdots+c_s\gamma_s=0,$$

其中$a_1,\ldots,a_t,b_1,\ldots,b_r,c_1,\ldots,c_s\in\mathbb{F}$, 则有

$$a_1\alpha_1+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_r\beta_r=-(c_1\gamma_1+\cdots+c_s\gamma_s)\in W_1\cap W_2.$$

于是

$$-(c_1\gamma_1+\cdots+c_s\gamma_s)=d_1\alpha_1+\cdots+d_t\alpha_t,\quad(\exists d_1,\ldots,d_t\in\mathbb{F}),$$

即

$$c_1\gamma_1+\cdots+c_s\gamma_s+d_1\alpha_1+\cdots+d_t\alpha_t=0.$$

因为$\alpha_1,\ldots,\alpha_t,\gamma_1,\ldots,\gamma_s$线性无关, 所以$c_1=\cdots=c_s=0$. 进而有

$$a_1\alpha_1+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_r\beta_r=0.$$

因为$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_r$线性无关, 所以$a_1=\cdots=a_t=b_1=\cdots=b_r=0$. 所以$\alpha_1,\ldots,\alpha_t$, $\beta_1,\ldots,\beta_r$, $\gamma_1,\ldots,\gamma_s$线性无关. 最后, 任意$\alpha\in W_1+W_2$, 有$\alpha=\beta+\gamma$, 其中$\beta\in W_1$, $\gamma\in W_2$. 而

$$\begin{aligned}\beta&=a_1\alpha_1+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_r\beta_r,\quad(\exists a_1,\ldots,a_t,b_1,\ldots,b_r\in\mathbb{F}),\\ \gamma&=c_1\alpha_1+\cdots+c_t\alpha_t+d_1\gamma_1+\cdots+d_s\gamma_s,\quad(\exists c_1,\ldots,c_t,d_1,\ldots,d_s\in\mathbb{F}),\end{aligned}$$

所以

$$\alpha=(a_1+c_1)\alpha_1+\cdots+(a_t+c_t)\alpha_t+b_1\beta_1+\cdots+b_r\beta_r+d_1\gamma_1+\cdots+d_s\gamma_s.$$

这得到W_1+W_2可由$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_r,\gamma_1,\ldots,\gamma_s$线性表示. 所以这组向量是$W_1+W_2$的基.

于是

$$\begin{aligned}\dim(W_1+W_2)=&t+r+s=(t+r)+(t+s)-t\\ =&\dim W_1+\dim W_2-\dim(W_1\cap W_2),\end{aligned}$$

证毕. □

定理6.10有下面显然的推论.

推论6.11. *设W_1和W_2是线性空间V的有限维子空间, 则*

(1) $\dim(W_1+W_2)\leqslant\dim W_1+\dim W_2$.

进而, 若V是有限维的, 则

(2) $\dim(W_1\cap W_2)\geqslant\dim W_1+\dim W_2-\dim V$;

(3) 若$\dim W_1+\dim W_2>\dim V$，则$\{0\}\subsetneqq W_1\cap W_2$.

我们最后看一个具体的例子.

例6.22. 设$W_1=\mathrm{Span}(\{\alpha_1,\alpha_2\})$，$W_2=\mathrm{Span}(\{\beta_1,\beta_2,\beta_3\})$，其中$\alpha_1=\begin{pmatrix}1\\2\\1\\0\end{pmatrix}$，$\alpha_2=\begin{pmatrix}-1\\1\\1\\1\end{pmatrix}$，$\beta_1=\begin{pmatrix}2\\1\\0\\-1\end{pmatrix}$，$\beta_2=\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}$，$\beta_3=\begin{pmatrix}4\\-1\\0\\1\end{pmatrix}$. 求$W_1+W_2$和$W_1\cap W_2$的基及其维数.

☞ **解.** 我们有

$$W_1+W_2=\mathrm{Span}(\{\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3\}),$$

而

$$(\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3)=\begin{pmatrix}1&-1&2&1&4\\2&1&1&-1&-1\\1&1&0&0&0\\0&1&-1&1&1\end{pmatrix}\underset{r_3-r_1}{\overset{r_2-2r_1}{\sim}}\begin{pmatrix}1&-1&2&1&4\\0&3&-3&-3&-9\\0&2&-2&-1&-4\\0&1&-1&1&1\end{pmatrix}$$

$$\underset{\substack{r_3-2r_4\\r_2\leftrightarrow r_4}}{\overset{\substack{r_1+r_4\\r_2-3r_4}}{\sim}}\begin{pmatrix}1&0&1&2&5\\0&1&-1&1&1\\0&0&0&-3&-6\\0&0&0&-6&-12\end{pmatrix}\underset{-\frac{1}{3}r_3}{\overset{r_4-2r_3}{\sim}}\begin{pmatrix}1&0&1&2&5\\0&1&-1&1&1\\0&0&0&1&2\\0&0&0&0&0\end{pmatrix}$$

$$\underset{r_2-r_3}{\overset{r_1-2r_3}{\sim}}\begin{pmatrix}1&0&1&0&1\\0&1&-1&0&-1\\0&0&0&1&2\\0&0&0&0&0\end{pmatrix},$$

所以$\alpha_1,\alpha_2,\beta_2$为$\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3$的极大无关组, 且

$$\beta_1=\alpha_1-\alpha_2,\quad \beta_3=\alpha_1-2\alpha_2+2\beta_2.$$

于是W_1+W_2有基$\alpha_1,\alpha_2,\beta_2$, 且$\dim(W_1+W_2)=3$.

由于

$$(\alpha_1,\alpha_2)\overset{r}{\sim}\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix},$$

所以α_1,α_2是W_1的基, $\dim W_1=2$. 由于

$$(\beta_1,\beta_2,\beta_3)\overset{r}{\sim}\begin{pmatrix}1&0&1\\-1&0&-1\\0&1&2\\0&0&0\end{pmatrix}\underset{r_2\leftrightarrow r_3}{\overset{r_2+r_1}{\sim}}\begin{pmatrix}1&0&1\\0&1&2\\0&0&0\\0&0&0\end{pmatrix},$$

所以β_1,β_2是W_2的基, $\dim W_2=2$. 维数公式给出

$$\dim(W_1\cap W_2)=\dim W_1+\dim W_2-\dim(W_1+W_2)=2+2-3=1.$$

因为$\beta_1=\alpha_1-\alpha_2\in W_1\cap W_2$, 所以$W_1\cap W_2$有基$\beta_1$. □

习题6.4

A1. 求下列线性空间V的维数和一组基:

(1) V为数域$\mathbb{F}$上所有n阶对称(反对称, 上三角)矩阵构成的$\mathbb{F}$上的线性空间;

(2) V为矩阵A的全体实系数多项式组成的实线性空间, 其中

$$A=\begin{pmatrix}1&0&0\\0&\omega&0\\0&0&\omega^2\end{pmatrix},\quad \omega=e^{\frac{2\pi i}{3}}=-\frac{1}{2}+\frac{\sqrt{3}}{2}i.$$

A2. 证明: 线性空间V为无穷维线性空间, 其中

(1) $V=\mathbb{F}^\infty$;

(2) V为闭区间$[0,1]$上所有连续实值函数做成的实线性空间.

A3. 在实线性空间$\mathbb{R}^4$中, 求向量ξ在基$\alpha_1,\alpha_2,\alpha_3,\alpha_4$下的坐标, 其中

(1) $\alpha_1=(1,1,1,1)^T$, $\alpha_2=(1,1,1,0)^T$, $\alpha_3=(1,1,0,0)^T$, $\alpha_4=(1,0,0,0)^T$, $\xi=(1,2,1,1)^T$;

(2) $\alpha_1=(1,1,0,1)^T$, $\alpha_2=(2,1,2,1)^T$, $\alpha_3=(1,1,0,0)^T$, $\alpha_4=(0,1,-1,-1)^T$, $\xi=(1,2,3,4)^T$.

A4. 求一个次数不超过2的实系数多项式$\varphi(x)$, 使得$\varphi(1)=8$, $\varphi(2)=5$, $\varphi(3)=-4$.

A5. 设W是实线性空间$\mathbb{R}^4$中由向量$\alpha_1,\alpha_2,\alpha_3,\alpha_4$所生成的子空间, 其中

$$\alpha_1=(2,1,3,1)^T,\ \alpha_2=(1,2,0,1)^T,\ \alpha_3=(-1,1,-3,0)^T,\ \alpha_4=(1,1,1,1)^T,$$

求W的维数与一组基.

A6. 设W_1和W_2是$\mathbb{F}^9$的子空间, 且$\dim W_1=\dim W_2=5$, 证明: $W_1\cap W_2\neq\{0\}$.

A7. 设$\alpha_1=(1,2,-1,-2)^T$, $\alpha_2=(3,1,1,1)^T$, $\alpha_3=(-1,0,1,-1)^T$, $\beta_1=(2,5,-6,-5)^T$, $\beta_2=(-1,2,-7,3)^T$, 而$W_1=\mathrm{Span}(\{\alpha_1,\alpha_2,\alpha_3\})$, $W_2=\mathrm{Span}(\{\beta_1,\beta_2\})$, 求$W_1+W_2$和$W_1\cap W_2$的维数与一组基.

A8. 设V是数域$\mathbb{F}$上的线性空间, B是V的非空子集, 证明: B是V的基的充分和必要条件是, 对任意$\alpha\in V$, α可由B唯一地线性表示.

A9. 设W是有限维线性空间V的子空间, 证明: $\dim W\leqslant\dim V$, 且$\dim W=\dim V$的充分必要条件是$W=V$.

A10. 设多项式$f_0(x),f_1(x),\dots,f_n(x)\in\mathbb{F}_n[x]$满足$f_j(2)=0$, 其中$j=0,1,\dots,n$, 证明: $f_0(x)$, $f_1(x)$, $\dots$, $f_n(x)$线性相关.

A11. 类比于三个有限集的并的所含元素个数公式, 猜想对于有限维线性空间V的三个子空间W_1,W_2,W_3, 成立

$$\dim(W_1+W_2+W_3)=\dim W_1+\dim W_2+\dim W_3-\dim(W_1\cap W_2)-\dim(W_1\cap W_3)$$

$$-\dim(W_2 \cap W_3) + \dim(W_1 \cap W_2 \cap W_3).$$

上面公式是否成立? 成立时给出证明, 不成立时举出反例.

A12. 证明: 数域$\mathbb{F}$上的无限维线性空间V一定有无限维真子空间.

A13. 设$\mathbb{F}$和$\mathbb{K}$都是数域, 且$\mathbb{F} \subset \mathbb{K}$, 将$\mathbb{K}$作为$\mathbb{F}$上线性空间的维数记为$[\mathbb{K} : \mathbb{F}]$. 设$\mathbb{E}$也是数域, 且$\mathbb{K} \subset \mathbb{E}$. 证明: 如果$[\mathbb{E} : \mathbb{K}] < \infty$且$[\mathbb{K} : \mathbb{F}] < \infty$, 那么$[\mathbb{E} : \mathbb{F}] = [\mathbb{E} : \mathbb{K}][\mathbb{K} : \mathbb{F}]$.

B1. 设W是$\mathbb{F}^{n\times n}$中所有形如$[A, B]$, $A, B \in \mathbb{F}^{n\times n}$的矩阵所生成的子空间. 证明:

(1) $W = \{A \in \mathbb{F}^{n\times n} \mid \mathrm{Tr}(A) = 0\}$;

(2) $\dim W = n^2 - 1$.

B2. 证明: 实数集$\mathbb{R}$作为有理数域$\mathbb{Q}$上的线性空间是无限维的. (**提示**: 实数集$\mathbb{R}$的势是不可数无穷或者利用超越数.)

B3. 设S是非空集合, $\mathbb{F}$是数域, 定义集合

$$F(S) = \{f : S \longrightarrow \mathbb{F} \mid \text{存在}S\text{的有限子集}T_f\text{使得对任意}x \in S - T_f, \text{有}f(x) = 0\}.$$

证明:

(1) 集合$F(S)$在函数的加法和数乘下成为$\mathbb{F}$上的线性空间, 称其为由集合S生成的自由线性空间;

(2) 对任意$s \in S$, 定义函数$\delta_s : S \longrightarrow \mathbb{F}$, 使得对任意$x \in S$, 有

$$\delta_s(x) = \begin{cases} 1, & \text{如果 } x = s, \\ 0, & \text{如果 } x \neq s, \end{cases}$$

则$\delta_s \in F(S)$, 且$\{\delta_s \mid s \in S\}$是$F(S)$的一组基. (于是, 由$S$生成的自由线性空间也可以看成是以$S$为基的线性空间, 即将$S$中的元素看成没有关系的形式记号, 且

$$F(S) = \left\{ \sum_{s\in S} c_s s \mid c_s \in \mathbb{F}, \text{只有有限个}c_s \neq 0 \right\},$$

而加法和数乘为

$$\sum_{s\in S} a_s s + \sum_{s\in S} b_s s = \sum_{s\in S} (a_s + b_s)s, \quad a\sum_{s\in S} c_s s = \sum_{s\in S} (ac_s)s.$$

应用时通常用这种观点.)

§ 6.5 基变换与坐标变换

现在对每个非零线性空间, 我们都有了坐标系. 对一个问题, 如果用解系几何, 我们要取一个坐标系, 将问题翻译到这个坐标系上. 坐标系通常不唯一(想想三维Euclid空间), 同一个问题在不同的坐标系中的翻译应该不一样, 有些坐标系中显得复杂, 有些坐标系中显得简单. 对一个问题, 我们当然希望找这个问题翻译形式最简单的坐标系. 这其实就是我们的故事所要讲述的, 对我们想考虑的问题找其合适的坐标系. 如果运气不好, 取了我们不满意的坐标系, 那么我们要换坐标系. 在这之前, 当然希望知道不同的坐标系之间的关系, 以及同一个向量在不同的坐标系中的坐标的关系. 本节来回答这些问题, 为此我们先引入形式矩阵记号.

设V是$\mathbb{F}$上的线性空间, $\alpha_1,\alpha_2,\ldots,\alpha_m\in V$, 对$c_1,c_2,\ldots,c_m\in\mathbb{F}$, 我们常常形式上记

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m=:(\alpha_1,\alpha_2,\ldots,\alpha_m)\begin{pmatrix}c_1\\c_2\\\vdots\\c_m\end{pmatrix},$$

即将$(\alpha_1,\alpha_2,\ldots,\alpha_m)$看成行向量, 右边理解成矩阵的乘法. 类似的, 对于矩阵$A=(a_{ij})_{m\times s}$, 记

$$(\alpha_1,\alpha_2,\ldots,\alpha_m)A=\left(\sum_{i=1}^m a_{i1}\alpha_i,\sum_{i=1}^m a_{i2}\alpha_i,\ldots,\sum_{i=1}^m a_{is}\alpha_i\right),$$

即上面看成两个行向量相等, 而左边理解成矩阵乘法. 还可以定义V中同维数的形式行向量的加法以及形式行向量的数乘:

$$\begin{aligned}&(\alpha_1,\alpha_2,\ldots,\alpha_m)+(\beta_1,\beta_2,\ldots,\beta_m):=(\alpha_1+\beta_1,\alpha_2+\beta_2,\ldots,\alpha_m+\beta_m),\\&c(\alpha_1,\alpha_2,\ldots,\alpha_m):=(c\alpha_1,c\alpha_2,\ldots,c\alpha_m),\quad(c\in\mathbb{F}).\end{aligned}$$

容易证明, 下面成立(假设运算都可进行):

$$\begin{aligned}&(\alpha_1,\alpha_2,\ldots,\alpha_m)(A+B)=(\alpha_1,\alpha_2,\ldots,\alpha_m)A+(\alpha_1,\alpha_2,\ldots,\alpha_m)B,\\&((\alpha_1,\alpha_2,\ldots,\alpha_m)+(\beta_1,\beta_2,\ldots,\beta_m))A=(\alpha_1,\alpha_2,\ldots,\alpha_m)A+(\beta_1,\beta_2,\ldots,\beta_m)A,\\&(c(\alpha_1,\alpha_2,\ldots,\alpha_m))A=(\alpha_1,\alpha_2,\ldots,\alpha_m)(cA)=c\left((\alpha_1,\alpha_2,\ldots,\alpha_m)A\right),\\&((\alpha_1,\alpha_2,\ldots,\alpha_m)A)C=(\alpha_1,\alpha_2,\ldots,\alpha_m)(AC).\end{aligned}$$

于是采用矩阵记号后, 就可以像矩阵一样自由进行运算了.

作为应用, 线性无关的矩阵结论为如下常用的结果.

引理6.12. *设$\alpha_1,\alpha_2,\ldots,\alpha_m\in V$线性无关, $A,B\in\mathbb{F}^{m\times s}$, 满足*

$$(\alpha_1,\alpha_2,\ldots,\alpha_m)A=(\alpha_1,\alpha_2,\ldots,\alpha_m)B,$$

则$A=B$.

☞ **证明.** 记$A=(a_{ij})$, $B=(b_{ij})$, 由上面的记号说明, 所给的等式即

$$\sum_{i=1}^m a_{ij}\alpha_i=\sum_{i=1}^m b_{ij}\alpha_i,\quad(j=1,2,\ldots,s).$$

因为$\alpha_1,\alpha_2,\ldots,\alpha_m$线性无关, 所以有

$$a_{ij}=b_{ij},\quad(1\leqslant i\leqslant m,1\leqslant j\leqslant s),$$

即$A=B$. □

下面回到我们的主题: 研究V的不同基之间的关系. 设$\dim V=n>0$, 而

$$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}\text{ 和 }\{\beta_1,\beta_2,\ldots,\beta_n\}$$

是V的两组基. 因为前者是基, 所以存在$a_{1j}, a_{2j}, \dots, a_{nj} \in \mathbb{F}$, 使得

$$\beta_j = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n, \quad (j = 1, 2, \dots, n).$$

将上面n个等式用矩阵记号表示就是

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{6.1}$$

称矩阵$P = (a_{ij}) \in \mathbb{F}^{n\times n}$为从基$\alpha_1, \alpha_2, \dots, \alpha_n$到基$\beta_1, \beta_2, \dots, \beta_n$的过渡矩阵. 注意到$P$的第$j$列就是$\beta_j$在基$\alpha_1, \alpha_2, \dots, \alpha_n$下的坐标. 式(6.1)给出了$V$的两组基$\alpha_1, \alpha_2, \dots, \alpha_n$和$\beta_1, \beta_2, \dots, \beta_n$之间的关系, 称式(6.1)为从基$\alpha_1, \alpha_2, \dots, \alpha_n$到基$\beta_1, \beta_2, \dots, \beta_n$的基变换公式.

上面考虑过渡矩阵和基变换公式是从$\alpha_1, \alpha_2, \dots, \alpha_n$是基得到的, 但是$\beta_1, \beta_2, \dots, \beta_n$也是基啊, 于是可设

$$(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_n)Q, \quad Q \in \mathbb{F}^{n\times n},$$

即Q是从基$\beta_1, \beta_2, \dots, \beta_n$到基$\alpha_1, \alpha_2, \dots, \alpha_n$的过渡矩阵. 则有

$$(\beta_1, \beta_2, \dots, \beta_n) = (\beta_1, \beta_2, \dots, \beta_n)QP,$$

由引理6.12得$QP = E_n$, 于是P可逆, 且$P^{-1} = Q$.

所以下面关于过渡矩阵的性质成立.

命题6.13. *设$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$和$\{\beta_1, \beta_2, \dots, \beta_n\}$是$V$的两组基, $P \in \mathbb{F}^{n\times n}$是从基$\alpha_1, \alpha_2, \dots, \alpha_n$到基$\beta_1, \beta_2, \dots, \beta_n$的过渡矩阵, 则$P$可逆, 且*

$$P^{-1} = \text{从基 } \beta_1, \beta_2, \dots, \beta_n \text{ 到基 } \alpha_1, \alpha_2, \dots, \alpha_n \text{ 的过渡矩阵}.$$

事实上, 容易证明: 对于线性空间V的一组基$\alpha_1, \alpha_2, \dots, \alpha_n$, 和矩阵$P \in \mathbb{F}^{n\times n}$, 由

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P$$

得到的向量组$\beta_1, \beta_2, \dots, \beta_n$是$V$的基的充分必要条件是矩阵$P$可逆. 我们将其证明留作练习.

下面考虑坐标之间的关系. 设V有两组基$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$和$\{\beta_1, \beta_2, \dots, \beta_n\}$, 基变换公式是

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P.$$

任取$\alpha \in V$, 设

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

$$\alpha = y_1\beta_1 + y_2\beta_2 + \cdots + y_n\beta_n = (\beta_1, \beta_2, \ldots, \beta_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

于是有

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)P\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

就得到

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

这就是坐标变换公式.

所以, 不同的基以及同一向量的不同坐标之间的关系都差一个可逆矩阵, 从一个坐标系到另一个坐标系, 要用一个可逆矩阵(通过乘法)过渡.

我们最后看几个例子.

例6.23. 设$V = \mathbb{R}^4$, $\alpha_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ -2 \end{pmatrix}$, $\alpha_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\alpha_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $\beta_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\beta_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$, $\beta_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $\beta_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, 求从基$\alpha_1, \alpha_2, \alpha_3, \alpha_4$到基$\beta_1, \beta_2, \beta_3, \beta_4$的过渡矩阵.

☞ **解.** 设所求过渡矩阵为P, 则

$$(\beta_1, \beta_2, \beta_3, \beta_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)P,$$

于是

$$P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)^{-1}(\beta_1, \beta_2, \beta_3, \beta_4).$$

由于

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4) = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 & -1 & -1 & 1 \\ 1 & 2 & 1 & 0 & 1 & -1 & 1 & -1 \\ 1 & -2 & 0 & 1 & 1 & 1 & -1 & -1 \end{pmatrix} \underbrace{\overset{r_2+r_1}{\overset{r_3-r_1}{}}}_{r_4-r_1}$$

$$
\begin{pmatrix}1&0&1&1&1&1&1&1\\0&-1&1&2&2&0&0&2\\0&2&0&-1&0&-2&0&-2\\0&-2&-1&0&0&0&-2&-2\end{pmatrix}\xrightarrow[r_3+2r_2]{r_4+r_3}\begin{pmatrix}1&0&1&1&1&1&1&1\\0&-1&1&2&2&0&0&2\\0&0&2&3&4&-2&0&2\\0&0&-1&-1&0&-2&-2&-4\end{pmatrix}
$$

$$
\xrightarrow[\substack{r_3+2r_4\\ r_3\leftrightarrow r_4}]{\substack{r_1+r_4\\ r_2+r_4}}\begin{pmatrix}1&0&0&0&1&-1&-1&-3\\0&-1&0&1&2&-2&-2&-2\\0&0&-1&-1&0&-2&-2&-4\\0&0&0&1&4&-6&-4&-6\end{pmatrix}\xrightarrow[r_3+r_4]{r_2-r_4}
$$

$$
\begin{pmatrix}1&0&0&0&1&-1&-1&-3\\0&-1&0&0&-2&4&2&4\\0&0&-1&0&4&-8&-6&-10\\0&0&0&1&4&-6&-4&-6\end{pmatrix}\xrightarrow[-r_3]{-r_2}\begin{pmatrix}1&0&0&0&1&-1&-1&-3\\0&1&0&0&2&-4&-2&-4\\0&0&1&0&-4&8&6&10\\0&0&0&1&4&-6&-4&-6\end{pmatrix},
$$

所以$P=\begin{pmatrix}1&-1&-1&-3\\2&-4&-2&-4\\-4&8&6&10\\4&-6&-4&-6\end{pmatrix}$. □

例6.24. 设$V=\mathbb{F}_2[x]$, 求从自然基$1,x,x^2$到基$1,x-1,(x-1)^2$的过渡矩阵, 并求$f(x)=1+2x+3x^2$在基$1,x-1,(x-1)^2$下的坐标.

☞ **解.** 由于

$$
1=1,\quad x-1=-1+x,\quad (x-1)^2=1-2x+x^2,
$$

所以所求的过渡矩阵为

$$
P=\begin{pmatrix}1&-1&1\\0&1&-2\\0&0&1\end{pmatrix}.
$$

因为

$$
\begin{pmatrix}1&-1&1&1\\0&1&-2&2\\0&0&1&3\end{pmatrix}\xrightarrow{r_1+r_2}\begin{pmatrix}1&0&-1&3\\0&1&-2&2\\0&0&1&3\end{pmatrix}\xrightarrow[r_2+2r_3]{r_1+r_3}\begin{pmatrix}1&0&0&6\\0&1&0&8\\0&0&1&3\end{pmatrix},
$$

所以$f(x)$在基$1,x-1,(x-1)^2$下的坐标是$\begin{pmatrix}6\\8\\3\end{pmatrix}$.

也可利用Taylor展开

$$
f(x)=f(1)+f'(1)(x-1)+\frac{f''(1)}{2}(x-1)^2
$$

求坐标. 我们有

$$
f(1)=6,\quad f'(x)=2+6x,\ f'(1)=8,\quad f''(x)=6,
$$

于是所求坐标是$(6,8,3)^T$. □

例6.25. 设$V=\mathbb{F}_2[x]$, 求从基$1+x,1-x,1+x^2$到基$1-x^2,x+x^2,x-x^2$的过渡矩阵.

☞ **解.** 我们用自然基过渡, 由于

$$(1+x,1-x,1+x^2)=(1,x,x^2)\begin{pmatrix}1&1&1\\1&-1&0\\0&0&1\end{pmatrix},$$

所以得到

$$\begin{aligned}(1-x^2,x+x^2,x-x^2)=&(1,x,x^2)\begin{pmatrix}1&0&0\\0&1&1\\-1&1&-1\end{pmatrix}\\=&(1+x,1-x,1+x^2)\begin{pmatrix}1&1&1\\1&-1&0\\0&0&1\end{pmatrix}^{-1}\begin{pmatrix}1&0&0\\0&1&1\\-1&1&-1\end{pmatrix}.\end{aligned}$$

由于

$$\begin{aligned}&\begin{pmatrix}1&1&1&1&0&0\\1&-1&0&0&1&1\\0&0&1&-1&1&-1\end{pmatrix}\xrightarrow{r_2-r_1}\begin{pmatrix}1&1&1&1&0&0\\0&-2&-1&-1&1&1\\0&0&1&-1&1&-1\end{pmatrix}\xrightarrow[r_2+r_3]{r_1-r_3}\\&\begin{pmatrix}1&1&0&2&-1&1\\0&-2&0&-2&2&0\\0&0&1&-1&1&-1\end{pmatrix}\xrightarrow{-\frac{1}{2}r_2}\begin{pmatrix}1&1&0&2&-1&1\\0&1&0&1&-1&0\\0&0&1&-1&1&-1\end{pmatrix}\xrightarrow{r_1-r_2}\\&\begin{pmatrix}1&0&0&1&0&1\\0&1&0&1&-1&0\\0&0&1&-1&1&-1\end{pmatrix},\end{aligned}$$

所以所求过渡矩阵为$\begin{pmatrix}1&0&1\\1&-1&0\\-1&1&-1\end{pmatrix}$. □

上例本质上就是求$\mathbb{F}^3$的从基$\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix}$到基$\begin{pmatrix}1\\0\\-1\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}0\\1\\-1\end{pmatrix}$的过渡矩阵. 参看下节.

习题6.5

A1. 设V是n维线性空间, 有三组基$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$, $\{\beta_1,\beta_2,\dots,\beta_n\}$和$\{\gamma_1,\gamma_2,\dots,\gamma_n\}$, 设从基$\alpha_1$, α_2, $\dots$, α_n到基$\beta_1,\beta_2,\dots,\beta_n$的过渡矩阵是$P$, 从基$\beta_1,\beta_2,\dots,\beta_n$到基$\gamma_1,\gamma_2,\dots,\gamma_n$的过渡矩阵是$Q$, 求从基$\alpha_1,\alpha_2,\dots,\alpha_n$到基$\gamma_1$, γ_2, $\dots$, γ_n的过渡矩阵.

A2. 求四维实向量空间$\mathbb{R}^4$中, 从基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$到基$\{\beta_1,\beta_2,\beta_3,\beta_4\}$的过渡矩阵, 其中

$$\begin{aligned}\alpha_1&=(1,2,-1,0)^T, &\beta_1&=(2,1,0,1)^T,\\\alpha_2&=(1,-1,1,1)^T, &\beta_2&=(0,1,2,2)^T,\\\alpha_3&=(-1,2,1,1)^T, &\beta_3&=(-2,1,1,2)^T,\\\alpha_4&=(-1,-1,0,1)^T, &\beta_4&=(1,3,1,2)^T.\end{aligned}$$

并求向量$\alpha = (1,0,0,0)^T$在基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$下的坐标.

A3. 在四维实线性空间$\mathbb{R}^4$的自然基$\{e_1,e_2,e_3,e_4\}$下, 超球面方程为$x_1^2+x_2^2+x_3^2+x_4^2 = 4$. 设$\alpha_1 = (1,1,1,1)^T$, $\alpha_2 = (1,1,-1,-1)^T$, $\alpha_3 = (1,-1,1,-1)^T$, $\alpha_4 = (1,-1,-1,1)^T$, 求该超球面在基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$下的方程.

A4. 设$\alpha_1,\alpha_2,\ldots,\alpha_n$是线性空间$V$的一组基, 矩阵$P\in\mathbb{F}^{n\times n}$, 由

$$(\beta_1,\beta_2,\ldots,\beta_n) = (\alpha_1,\alpha_2,\ldots,\alpha_n)P$$

得到向量$\beta_1,\beta_2,\ldots,\beta_n\in V$. 证明: $\beta_1,\beta_2,\ldots,\beta_n$是$V$的基的充分必要条件是矩阵$P$可逆.

A5. 设$a_1,a_2,\ldots,a_n\in\mathbb{F}$互不相同, 对$i=1,2,\ldots,n$定义

$$f_i(x) = (x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)\in\mathbb{F}_{n-1}[x].$$

(1) 证明: $\{f_1(x),\ f_2(x),\ \ldots,\ f_n(x)\}$是$\mathbb{F}_{n-1}[x]$的一组基;

(2) 设$\mathbb{F}=\mathbb{C}$, 取$a_1,a_2,\ldots,a_n$为全体n次单位根, 求$\mathbb{F}_{n-1}[x]$中从基$\{1,\ x, x^2,\ \ldots,\ x^{n-1}\}$到基$\{f_1(x),$ $f_2(x),\ \ldots,\ f_n(x)\}$的过渡矩阵.

§6.6 同构

对于非零线性空间, 我们已经有了坐标系, 而且知道了不同坐标系之间的变换关系, 下面考虑空间的相等, 这个可以认为是线性空间之间最简单的关系. 完全相同的两个线性空间当然相等; 有些线性空间的表现形式可能不一样, 但是它们中的向量一一对应, 加法和数乘也一一对应, 作为线性空间应该可以看成是一样的, 也要认为相等(线性空间的公理化定义时, 我们忘掉了对象的具体表现形式, 而只关注对象所成的集合, 以及加法和数乘两个运算). 所以, 在这里我们认为平行的两个空间是一样的.

下面我们要严格定义何谓向量一一对应, 何谓运算一一对应.

6.6.1 双射与可逆映射

为了说明两个向量空间的向量一一对应, 我们回忆一下集合论中的双射和可逆映射的概念和关系.

定义6.11. *设$\varphi: A\longrightarrow B$是集合间的映射.*

(1) *如果对任意$a_1\neq a_2\in A$, 有$\varphi(a_1)\neq\varphi(a_2)$, 则称$\varphi$是单射;*

(2) *如果对任意$b\in B$, 存在$a\in A$, 使得$\varphi(a)=b$, 则称φ是满射;*

(3) *如果φ既是满射又是单射, 则称φ是双射.*

由定义可得, $\varphi: A\longrightarrow B$是单射当且仅当: $\forall a_1,a_2\in A$, 若$\varphi(a_1)=\varphi(a_2)$, 则$a_1=a_2$. 满射还有另一种描述方式. 设$S$是$A$的(非空)子集, 定义$S$在$\varphi$下的像集为

$$\varphi(S) := \{\varphi(a)\mid a\in S\}\subset B.$$

特别地, 称

$$\varphi(A) = \{\varphi(a) \mid a \in A\} \subset B$$

为映射φ的像. 于是

$$\varphi\text{为满射} \iff \varphi(A) = B.$$

可以定义φ在S上的**限制映射**

$$\varphi|_S : S \longrightarrow \varphi(S) \subset B; a \mapsto \varphi|_S(a) := \varphi(a).$$

下面定义可逆映射.

定义6.12. **(1)** 对集合A, 记映射$\mathrm{id}_A : A \longrightarrow A; a \mapsto a$, 称$\mathrm{id}_A$为$A$的**恒等映射**;

(2) 如果$\varphi : A \longrightarrow B$和$\psi : B \longrightarrow C$是两个映射, 定义它们的**复合**为

$$\psi \circ \varphi : A \longrightarrow C; a \mapsto (\psi \circ \varphi)(a) := \psi(\varphi(a));$$

(3) 设$\varphi : A \longrightarrow B$是映射, 如果存在映射$\psi : B \longrightarrow A$, 使得$\psi \circ \varphi = \mathrm{id}_A$, $\varphi \circ \psi = \mathrm{id}_B$, 则称$\varphi$是**可逆映射**. 此时, ψ是唯一的, 称为φ的**逆映射**, 记为φ^{-1}.

可以将可逆映射的概念类比于可逆矩阵的定义. 下面证明一个映射可逆的充分必要条件是它是双射.

引理6.14. 设$\varphi : A \longrightarrow B$是映射, 则

$$\varphi\text{是双射} \iff \varphi\text{是可逆映射}.$$

证明. 设φ是双射. 对任意$b \in B$, 由于φ是满射, 所以存在$a \in A$, 使得$\varphi(a) = b$. 因为φ是单射, 所以a唯一. 于是我们得到映射

$$\psi : B \longrightarrow A; b \mapsto a, \quad (a\text{是满足}\varphi(a) = b\text{的唯一元} \in A).$$

于是对任意$a \in A$, 有

$$(\psi \circ \varphi)(a) = \psi(\varphi(a)) = a,$$

这得到$\psi \circ \varphi = \mathrm{id}_A$. 类似的, 对任意$b \in B$, 有

$$(\varphi \circ \psi)(b) = \varphi(\psi(b)) = b,$$

即有$\varphi \circ \psi = \mathrm{id}_B$. 所以$\varphi$是可逆映射.

反之, 设φ是可逆映射. 如果$a_1, a_2 \in A$, 满足$\varphi(a_1) = \varphi(a_2)$. 两边用$\varphi^{-1}$作用得

$$\varphi^{-1}(\varphi(a_1)) = \varphi^{-1}(\varphi(a_2)) \Longrightarrow a_1 = a_2,$$

于是φ是单射. 又任意$b \in B$, 有$\varphi^{-1}(b) \in A$且

$$\varphi(\varphi^{-1}(b)) = b,$$

于是φ是满射. 因此φ是双射. $\square$

6.6.2 线性映射

为了说明加法和数乘的对应, 我们引入重要的线性映射的概念, 这是下章研究的重点.

定义6.13. 设V和W是$\mathbb{F}$上的线性空间, 如果映射$\varphi : V \longrightarrow W$满足

(i) (保加法) $\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta), \quad \forall \alpha, \beta \in V$;

(ii) (保数乘) $\varphi(c\alpha)=c\varphi(\alpha), \quad \forall \alpha \in V, \forall c \in \mathbb{F}$,

则称φ是V到W的线性映射. 称线性映射$\varphi : V \longrightarrow V$为$V$上的线性变换, 称线性映射$\varphi : V \longrightarrow \mathbb{F}$为$V$上的线性函数.

例6.26. 设$\dim V=n$, $\alpha_1, \alpha_2, \ldots, \alpha_n$是$V$的基, 则映射

$$\varphi : V \longrightarrow \mathbb{F}^n; \alpha=c_1\alpha_1+\cdots+c_n\alpha_n \mapsto \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

既是双射, 也是线性映射.

线性映射满足下面的性质. 设$\varphi : V \longrightarrow W$是线性映射, 则

(i) (保零元) $\varphi(0)=0$;

(ii) (保负元) $\varphi(-\alpha)=-\varphi(\alpha), \quad \forall \alpha \in V$;

(iii) (保线性组合) $\forall \alpha_1, \alpha_2, \ldots, \alpha_s \in V, \forall c_1, c_2, \ldots, c_s \in \mathbb{F}$, 有

$$\varphi(c_1\alpha_1+c_2\alpha_2+\cdots+c_s\alpha_s)=c_1\varphi(\alpha_1)+c_2\varphi(\alpha_2)+\cdots+c_s\varphi(\alpha_s);$$

(iv) (保线性相关性) $\varnothing \neq \forall S \subset V$线性相关$\Longrightarrow \varphi(S) \subset W$线性相关;

(v) (保子空间) 对V的任意子空间U, 有$\varphi(U)$是W的子空间, 且限制映射$\varphi|_U : U \longrightarrow \varphi(U) \subset W$也是线性映射.

☞ **证明.** (i) 由于

$$\varphi(0)=\varphi(0+0)=\varphi(0)+\varphi(0),$$

所以$\varphi(0)=0$.

(ii) 我们有

$$\varphi(-\alpha)=\varphi((-1)\alpha)=(-1)\varphi(\alpha)=-\varphi(\alpha).$$

(iii) 对s归纳. 当$s=1$时, 由φ保数乘知成立. 设$s>1$, 且结论对$s-1$成立, 则有

$$\varphi(c_1\alpha_1+\cdots+c_{s-1}\alpha_{s-1})=c_1\varphi(\alpha_1)++\cdots+c_{s-1}\varphi(\alpha_{s-1}).$$

于是

$$\varphi(c_1\alpha_1+\cdots+c_{s-1}\alpha_{s-1}+c_s\alpha_s)=\varphi(c_1\alpha_1+\cdots+c_{s-1}\alpha_{s-1})+\varphi(c_s\alpha_s)$$

$$=c_1\varphi(\alpha_1)+\cdots+c_{s-1}\varphi(\alpha_{s-1})+c_s\varphi(\alpha_s).$$

(iv) 由(iii)和(i).
(v) 由于$0=\varphi(0)\in\varphi(U)$, 所以$\varphi(U)\neq\varnothing$. 对任意$\varphi(\alpha),\varphi(\beta)\in\varphi(U)$, 其中$\alpha,\beta\in U$, 有

$$\varphi(\alpha)+\varphi(\beta)=\varphi(\alpha+\beta)\in\varphi(U).$$

又若$a\in\mathbb{F}$, 则

$$a\varphi(\alpha)=\varphi(a\alpha)\in\varphi(U).$$

于是$\varphi(U)$是W的子空间. 进而由φ线性, 可得$\varphi|_U$线性. □

6.6.3 同构与可逆线性映射

现在可以定义线性空间之间的同构映射.

定义6.14. 设V和W是线性空间, $\varphi: V\longrightarrow W$是映射.

(1) 如果φ满足

(i) φ为线性映射;

(ii) φ为双射,

则称φ是同构映射, 简称为同构. 此时称V和W同构, 记为$V\cong W$;

(2) 如果φ满足

(i) φ为线性映射;

(ii)′ φ为可逆映射,

则称φ是可逆线性映射.

引理6.14表明

$$\varphi\text{是同构}\Longleftrightarrow\varphi\text{是可逆线性映射}.$$

我们看几个例子.

例6.27. 设V是n维线性空间, $\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的基, 则映射

$$\varphi: V\longrightarrow\mathbb{F}^n;\alpha=c_1\alpha_1+\cdots+c_n\alpha_n\mapsto\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$$

是同构映射, 进而$V\cong\mathbb{F}^n$.

例6.28. 设线性空间

$$V_1=\{A\in\mathbb{F}^{n\times n}\mid A\text{是下三角阵}\},\quad V_2=\{A\in\mathbb{F}^{n\times n}\mid A^T=A\},$$

证明: $V_1\cong V_2$.

☞ **证明.** 定义映射$\varphi: V_1 \longrightarrow V_2$, 使得$\forall A \in V_1$, 有

$$\varphi(A) = \frac{1}{2}(A + A^T) \in V_2.$$

对任意$A, B \in V_1$和$a \in \mathbb{F}$, 我们有

$$\varphi(A+B) = \frac{1}{2}[(A+B) + (A+B)^T] = \frac{1}{2}(A + A^T) + \frac{1}{2}(B + B^T) = \varphi(A) + \varphi(B),$$
$$\varphi(aA) = \frac{1}{2}(aA + (aA)^T) = \frac{1}{2}a(A + A^T) = a\varphi(A),$$

于是φ是线性映射.

注意到, 如果$A = (a_{ij}) \in V_1$, 记$\varphi(A) = (b_{ij})$, 则

$$b_{ij} = \begin{cases} a_{ii}, & i = j, \\ \frac{1}{2}(a_{ij} + a_{ji}), & i \neq j. \end{cases}$$

于是映射$\psi: V_2 \longrightarrow V_1$, 使得$\forall B = (b_{ij}) \in V_2$有$\psi(B) = (a_{ij})$, 其中

$$a_{ij} = \begin{cases} 0, & 1 \leqslant i < j \leqslant n, \\ b_{ii}, & 1 \leqslant i = j \leqslant n, \\ 2b_{ij}, & 1 \leqslant j < i \leqslant n, \end{cases}$$

满足$\psi \circ \varphi = \mathrm{id}_{V_1}$和$\varphi \circ \psi = \mathrm{id}_{V_2}$. 即$\varphi$是双射.

因此, φ是同构, $V_1 \cong V_2$.

事实上, 可以利用$\dim V_1 = \dim V_2$和定理6.16直接得到同构. □

下面的性质成立. 设$\varphi: V \longrightarrow W$是同构映射, 则

(i) $\forall \alpha \in V$, 有

$$\varphi(\alpha) = 0 \Longleftrightarrow \alpha = 0;$$

(ii) 设$\alpha_1, \ldots, \alpha_s \in V$, 数$c_1, \ldots, c_s \in \mathbb{F}$, 则

$$c_1\alpha_1 + \cdots + c_s\alpha_s = 0 \Longleftrightarrow c_1\varphi(\alpha_1) + \cdots + c_s\varphi(\alpha_s) = 0.$$

即$\alpha_1, \ldots, \alpha_s$和$\varphi(\alpha_1), \ldots, \varphi(\alpha_s)$有相同的线性关系;

(iii) 对任意$\varnothing \neq S \subset V$, 有

$$S \subset V\text{线性相关} \Longleftrightarrow \varphi(S) \subset W\text{线性相关},$$

和

$$S\text{是}V\text{的基} \Longleftrightarrow \varphi(S)\text{是}W\text{的基};$$

(iv) 对V的任意子空间U, 限制映射

$$\varphi|_U : U \to \varphi(U)$$

也是同构.

于是我们可以将同构的线性空间等同, 认为它们本质上就是同一个线性空间.

6.6.4 (有限维)线性空间在同构下的分类

我们定义了线性空间之间的同构关系, 它是等价关系, 即满足

(i) (自反性) $V \cong V$;

(ii) (对称性) $V \cong W \Longrightarrow W \cong V$;

(iii) (传递性) $V \cong W$且$W \cong U \Longrightarrow V \cong U$.

上面的对称性和传递性是下面引理的直接推论.

引理6.15. 设$\varphi : V \longrightarrow W$和$\psi : W \longrightarrow U$是线性空间的同构映射, 则

(1) $\varphi^{-1} : W \longrightarrow V$是同构映射;

(2) $\psi \circ \varphi : V \longrightarrow U$是同构映射.

证明. 容易证明所需证明为同构的映射是双射[7], 下面我们验证它们是线性映射.
(1) 对任意$\alpha, \beta \in W$和$c \in \mathbb{F}$, 我们有

$$\begin{aligned}\varphi(\varphi^{-1}(\alpha) + \varphi^{-1}(\beta)) &= \varphi(\varphi^{-1}(\alpha)) + \varphi(\varphi^{-1}(\beta)) = \alpha + \beta \\ &\Longrightarrow \varphi^{-1}(\alpha + \beta) = \varphi^{-1}(\alpha) + \varphi^{-1}(\beta), \\ \varphi(c\varphi^{-1}(\alpha)) &= c\varphi(\varphi^{-1}(\alpha)) = c\alpha \Longrightarrow \varphi^{-1}(c\alpha) = c\varphi^{-1}(\alpha),\end{aligned}$$

于是φ^{-1}是线性映射.
(2) 对任意$\alpha, \beta \in V$和$c \in \mathbb{F}$, 我们有

$$\begin{aligned}(\psi \circ \varphi)(\alpha + \beta) &= \psi(\varphi(\alpha + \beta)) = \psi(\varphi(\alpha) + \varphi(\beta)) = \psi(\varphi(\alpha)) + \psi(\varphi(\beta)) \\ &= (\psi \circ \varphi)(\alpha) + (\psi \circ \varphi)(\beta), \\ (\psi \circ \varphi)(c\alpha)) &= \psi(\varphi(c\alpha)) = \psi(c\varphi(\alpha)) = c\psi(\varphi(\alpha)) = c(\psi \circ \varphi)(\alpha),\end{aligned}$$

于是$\psi \circ \varphi$是线性映射. □

由于线性空间的同构是等价关系, 它给出了线性空间在同构下的分类. 于是我们希望解决关于这个分类的基本问题: (i) 同构的全系不变量? (ii) 同构标准形? 我们对有限维线性空间来回答这两个问题. 对于第(ii)个问题, 当$\dim V = n$时, 有$V \cong \mathbb{F}^n$. 下面的定理回答了第(i)个问题.

定理6.16. 设V和W是数域$\mathbb{F}$上的有限维线性空间, 则

$$V \cong W \Longleftrightarrow \dim V = \dim W.$$

证明. 设$\varphi : V \longrightarrow W$是同构映射, 任取$V$的基$\alpha_1, \alpha_2, \ldots, \alpha_n$, 则$\varphi(\alpha_1), \varphi(\alpha_2), \ldots, \varphi(\alpha_n) \in W$是$W$的基. 这得到$\dim V = \dim W$.

反之, 设$\dim V = \dim W = n$, 则$V \cong \mathbb{F}^n$, $W \cong \mathbb{F}^n$. 于是$V \cong W$. □

于是我们就完全解决了有限维线性空间在同构下的分类问题: 维数是有限维线性空间同构的全系不变量, 列向量空间$\mathbb{F}^n$是同构标准形. 见图6.5.

最后看几个例子.

[7] 参看本节习题.

有限维线性空间	$\dim=0$	$\dim=1$	$\cdots$	$\dim=n$	$\cdots$
标准形	$\{0\}$	$\mathbb{F}^1$	$\cdots$	$\mathbb{F}^n$	$\cdots$

图 6.5: 有限维线性空间在同构下的分类

例6.29. 设$f_1(x)=x^3-x^2-x+1$, $f_2(x)=x^3+x^2-x+1$, $f_3(x)=-x^3-x^2+x+1$, $g_1(x)=2x^3-2x+2$, $g_2(x)=x^3-3x^2-x+3$, $g_3(x)=2x^3+2x^2-2x$, 令$V=\mathrm{Span}(\{f_1(x),f_2(x),f_3(x)\})\subset\mathbb{F}_3[x]$, 证明: $g_1(x),g_2(x),g_3(x)$是V的一组基.

☞ **证明.** 取$\mathbb{F}_3[x]$的自然基$1,x,x^2,x^3$, 则有同构映射$\varphi:\mathbb{F}_3[x]\longrightarrow\mathbb{F}^4$. 我们有

$$\varphi(f_1(x))=\alpha_1=\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix},\ \varphi(f_2(x))=\alpha_2=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix},\ \varphi(f_3(x))=\alpha_3=\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix},$$

$$\varphi(g_1(x))=\beta_1=\begin{pmatrix}2\\-2\\0\\2\end{pmatrix},\ \varphi(g_2(x))=\beta_2=\begin{pmatrix}3\\-1\\-3\\1\end{pmatrix},\ \varphi(g_3(x))=\beta_3=\begin{pmatrix}0\\-2\\2\\2\end{pmatrix},$$

于是问题转化为证明: β_1,β_2,β_3是$W=\mathrm{Span}(\{\alpha_1,\alpha_2,\alpha_3\})$的一组基.

由于

$$(\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3)=\begin{pmatrix}1&1&1&2&3&0\\-1&-1&1&-2&-1&-2\\-1&1&-1&0&-3&2\\1&1&-1&2&1&2\end{pmatrix}\xrightarrow[r_4-r_1]{\substack{r_2+r_1\\r_3+r_1}}$$

$$\begin{pmatrix}1&1&1&2&3&0\\0&0&2&0&2&-2\\0&2&0&2&0&2\\0&0&-2&0&-2&2\end{pmatrix}\xrightarrow[\substack{\frac{1}{2}r_3\\r_2\leftrightarrow r_3}]{\substack{r_4+r_2\\\frac{1}{2}r_2}}\begin{pmatrix}1&1&1&2&3&0\\0&1&0&1&0&1\\0&0&1&0&1&-1\\0&0&0&0&0&0\end{pmatrix}$$

$$\xrightarrow[r_1-r_3]{r_1-r_2}\begin{pmatrix}1&0&0&1&2&0\\0&1&0&1&0&1\\0&0&1&0&1&-1\\0&0&0&0&0&0\end{pmatrix},$$

所以$\alpha_1,\alpha_2,\alpha_3$线性无关, 且$\beta_1,\beta_2,\beta_3$可由$\alpha_1,\alpha_2,\alpha_3$线性表示. 于是$\dim W=3$, 且$\beta_1,\ \beta_2,\ \beta_3\in W$. 又

$$(\beta_1,\beta_2,\beta_3)\overset{r}{\sim}\begin{pmatrix}1&2&0\\1&0&1\\0&1&-1\\0&0&0\end{pmatrix}\xrightarrow{r_2-r_1}\begin{pmatrix}1&2&0\\0&-2&1\\0&1&-1\\0&0&0\end{pmatrix}\xrightarrow[r_2\leftrightarrow r_3]{r_2+2r_3}\begin{pmatrix}1&2&0\\0&1&-1\\0&0&-1\\0&0&0\end{pmatrix},$$

所以β_1,β_2,β_3线性无关. 于是β_1,β_2,β_3是W的基. □

例6.30. 设$V=\mathbb{R}^{2\times 2}$, W是V的子空间, 它有两组基

$$A_1=\begin{pmatrix}0&-2\\2&2\end{pmatrix},\quad A_2=\begin{pmatrix}1&-1\\1&1\end{pmatrix},\quad A_3=\begin{pmatrix}3&-1\\-3&1\end{pmatrix}$$

和

$$B_1=\begin{pmatrix}2&-2\\0&2\end{pmatrix},\quad B_2=\begin{pmatrix}1&1\\-1&-1\end{pmatrix},\quad B_3=\begin{pmatrix}1&-1\\-1&1\end{pmatrix},$$

求从基A_1,A_2,A_3到基B_1,B_2,B_3的过渡矩阵.

☞ **解.** 取$\mathbb{R}^{2\times 2}$的自然基$E_{11},E_{12},E_{21},E_{22}$, 则有同构映射$\varphi:\mathbb{R}^{2\times 2}\longrightarrow\mathbb{R}^4$. 由于

$$\varphi(A_1)=\alpha_1=\begin{pmatrix}0\\-2\\2\\2\end{pmatrix},\quad \varphi(A_2)=\alpha_2=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix},\quad \varphi(A_3)=\alpha_3=\begin{pmatrix}3\\-1\\-3\\1\end{pmatrix},$$

$$\varphi(B_1)=\beta_1=\begin{pmatrix}2\\-2\\0\\2\end{pmatrix},\quad \varphi(B_2)=\beta_2=\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix},\quad \varphi(B_3)=\beta_3=\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix},$$

问题转化为求$\mathbb{R}^4$的子空间$\mathrm{Span}(\{\alpha_1,\alpha_2,\alpha_3\})(\cong W)$中从基$\alpha_1,\alpha_2,\alpha_3$到基$\beta_1,\beta_2,\beta_3$的过渡矩阵. 我们有

$$(\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3)=\begin{pmatrix}0&1&3&2&1&1\\-2&-1&-1&-2&1&-1\\2&1&-3&0&-1&-1\\2&1&1&2&-1&1\end{pmatrix}\underbrace{\overset{\substack{r_3+r_2\\r_4+r_2}}{\frown}}_{r_1\leftrightarrow r_2}$$

$$\begin{pmatrix}-2&-1&-1&-2&1&-1\\0&1&3&2&1&1\\0&0&-4&-2&0&-2\\0&0&0&0&0&0\end{pmatrix}\underbrace{\overset{r_1+r_2}{\frown}}_{-\frac14 r_3}\begin{pmatrix}-2&0&2&0&2&0\\0&1&3&2&1&1\\0&0&1&\frac12&0&\frac12\\0&0&0&0&0&0\end{pmatrix}\underbrace{\overset{-\frac12 r_1}{\frown}}_{r_2-3r_3}$$

$$\begin{pmatrix}1&0&-1&0&-1&0\\0&1&0&\frac12&1&-\frac12\\0&0&1&\frac12&0&\frac12\\0&0&0&0&0&0\end{pmatrix}\overset{r_1+r_3}{\frown}\begin{pmatrix}1&0&0&\frac12&-1&\frac12\\0&1&0&\frac12&1&-\frac12\\0&0&1&\frac12&0&\frac12\\0&0&0&0&0&0\end{pmatrix},$$

于是

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}\frac12&-1&\frac12\\\frac12&1&-\frac12\\\frac12&0&\frac12\end{pmatrix},$$

即所求的过渡矩阵是$\begin{pmatrix}\frac12&-1&\frac12\\\frac12&1&-\frac12\\\frac12&0&\frac12\end{pmatrix}$. □

习题6.6

A1. 设$\varphi: A \to B$是非空集合之间的映射. 证明: 下面命题等价
(i) φ是单射;
(ii) (存在左逆) 存在映射$\psi: B \to A$, 使得$\psi \circ \varphi = \mathrm{id}_A$;
(iii) (可左消去) 对任意非空集合C和映射$\psi_1, \psi_2: C \to A$, 如果$\varphi \circ \psi_1 = \varphi \circ \psi_2$, 则$\psi_1 = \psi_2$.
A2. 设$\varphi: A \to B$是非空集合之间的映射. 证明: 下面命题等价
(i) φ是满射;
(ii) (存在右逆) 存在映射$\psi: B \to A$, 使得$\varphi \circ \psi = \mathrm{id}_B$;
(iii) (可右消去) 对任意非空集合C和映射$\psi_1, \psi_2: B \to C$, 如果$\psi_1 \circ \varphi = \psi_2 \circ \varphi$, 则$\psi_1 = \psi_2$.
A3. 设$\varphi: A \to B$和$\psi: B \to C$是非空集合之间的可逆映射, 证明:
(1) $\varphi^{-1}: B \to A$也可逆, 且$(\varphi^{-1})^{-1} = \varphi$;
(2) $\psi \circ \varphi: A \to C$也可逆, 且$(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}$.
A4. 设有多项式

$$\begin{aligned} f_1(x) &= x^3 + x^2 + x + 1, & g_1(x) &= 2x^3 + 2x^2 - 3, \\ f_2(x) &= x^3 + x + 2, & g_2(x) &= 2x^3 + x^2 + 2x + 3, \\ f_3(x) &= 2x^3 + x^2 - 2, & g_3(x) &= x^3 + x^2 - 3x - 9, \end{aligned}$$

令$V = \mathrm{Span}(\{f_1(x), f_2(x), f_3(x)\}) \subset \mathbb{F}_3[x]$, 证明: $f_1(x), f_2(x), f_3(x)$和$g_1(x)$, $g_2(x)$, $g_3(x)$都是V的基, 并求从基$f_1(x), f_2(x), f_3(x)$ 到基$g_1(x), g_2(x), g_3(x)$的过渡矩阵.
A5. 设$\mathbb{R}_{>0}$是所有正实数组成的集合, 它在如下的加法和数乘下成为实线性空间:

$$a \oplus b = ab, \quad \lambda \circ a = a^{\lambda}, \quad \forall a, b \in \mathbb{R}_{>0},\ \forall \lambda \in \mathbb{R}.$$

证明: 实数域作为实线性空间与实线性空间$\mathbb{R}_{>0}$同构, 并给出一个具体的同构映射.
A6. 设$V = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$, 则$V$关于矩阵的加法和数乘成为实线性空间, 将$\mathbb{C}$看成实线性空间, 定义映射$\varphi: \mathbb{C} \longrightarrow V$, 使得对任意$\alpha = a + bi \in \mathbb{C}$, $a, b \in \mathbb{R}$, 有

$$\varphi(\alpha) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

证明: φ是线性空间之间的同构, 且对任意$\alpha, \beta \in \mathbb{C}$, 有

$$\varphi(\alpha\beta) = \varphi(\alpha)\varphi(\beta).$$

A7. 设$\alpha_1, \alpha_2, \ldots, \alpha_n$是$n$维线性空间$V$的一组基, $A \in \mathbb{F}^{n \times m}$, 由

$$(\beta_1, \beta_2, \ldots, \beta_m) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

得到向量$\beta_1, \beta_2, \ldots, \beta_m \in V$, 令$W = \mathrm{Span}(\{\beta_1, \beta_2, \ldots, \beta_m\})$, 证明: $\dim W = R(A)$.

§ 6.7 直和

数学中常常将一个对象分解成更简单的对象, 利用更简单的对象来研究一般对象. 例如, 将一个大于1的整数分解成素数的乘积, 就可以先研究素数; 将一个次数大于0的一元多项式分解为不可约多项式的乘积, 就可以先研究不可约多项式. 将这个想法用到线性空间上, 我们想用更简单的线性空间来分解一个线性空间, 于是自然就是用子空间来分解. 那么这里的分解是子空间的何种运算呢? 我们只学了子空间的交和和运算, 由于做交的话空间不会变大, 所以比较好的选择是和, 即我们希望把线性空间分解为一些子空间的和.

例如, 设V是线性空间, 它分解为子空间$W_1, W_2, \ldots, W_r$的和

$$V = W_1 + W_2 + \cdots + W_r.$$

我们可以先研究$W_1, W_2, \ldots, W_r$, 然后再回到V. 回到V时自然要考虑这些W_i之间公共的部分, 它们会对最后V的性质有影响. 如果这些子空间的公共部分很小, 比如就是零, 那么就不需要考虑公共部分, 直接将相互独立的W_1, W_2, $\ldots$, W_r的性质合起来就得到V的性质. 于是, 在应用上, 我们应该考虑这种分解: 将V分解为相互独立的子空间的和. 我们称这种和为直和, 下面就按照这种想法来定义直和.

设V是数域$\mathbb{F}$上的线性空间.

6.7.1 两个子空间的直和

如果$W_1, W_2 \subset V$为子空间, 则对任意$\alpha \in W_1 + W_2$, 有

$$\alpha = \alpha_1 + \alpha_2, \quad \exists \alpha_1 \in W_1, \exists \alpha_2 \in W_2.$$

一般情况下α_1和α_2不唯一; 如果唯一, 我们称是直和.

定义6.15. *设$W_1, W_2 \subset V$是子空间, 如果对任意$\alpha \in W_1 + W_2$, α可唯一写为*

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in W_1, \alpha_2 \in W_2,$$

则称和$W_1 + W_2$为直和, 记为$W_1 \oplus W_2$.

关于直和, 我们有下面的等价刻画, 其中(iii)可以认为说W_1和W_2相互独立.

定理6.17. *设$W_1, W_2 \subset V$是子空间, 则下面的命题等价*

(i) *$W_1 + W_2$是直和;*

(ii) *$0 \in W_1 + W_2$的分解式唯一: $0 = 0 + 0$;*

(iii) *$W_1 \cap W_2 = \{0\}$;*

(iv) *任取W_i的基$B_i (i = 1, 2)$, 有$B_1 \cap B_2 = \varnothing$, 且$B_1 \cup B_2$为$W_1 + W_2$的基;*

(v) *存在W_i的基$B_i (i = 1, 2)$, 使得$B_1 \cap B_2 = \varnothing$, 且$B_1 \cup B_2$为$W_1 + W_2$的基.*

进而, 当W_1和W_2是有限维子空间时, 上面还等价于

(vi) $\dim(W_1+W_2)=\dim W_1+\dim W_2$.

☞ **证明.** “(i) $\Longrightarrow$ (ii)” 显然.

“(ii) $\Longrightarrow$ (iii)” 假设(ii)成立, 任取$\alpha\in W_1\cap W_2$, 则$\alpha\in W_1$且$\alpha\in W_2$. 因为W_2是子空间, 所以$-\alpha\in W_2$. 由于

$$0=\alpha+(-\alpha)$$

为$0\in W_1+W_2$的分解, 所以(ii)推出$\alpha=0$. 即有$W_1\cap W_2=\{0\}$.

“(iii) $\Longrightarrow$ (iv)” 假设(iii)成立. 任取W_i的基B_i, $i=1,2$. 首先, 如果$\exists\alpha\in B_1\cap B_2$, 则$0\neq\alpha\in W_1\cap W_2$, 矛盾于(iii), 所以$B_1\cap B_2=\varnothing$. 其次, 如果存在$\alpha_1,\dots,\alpha_r\in B_1$和$\beta_1,\dots,\beta_s\in B_2$, 使得

$$a_1\alpha_1+\cdots+a_r\alpha_r+b_1\beta_1+\cdots+b_s\beta_s=0,\quad (a_1,\dots,a_r,b_1,\dots,b_s\in\mathbb{F}),$$

则

$$a_1\alpha_1+\cdots+a_r\alpha_r=-(b_1\beta_1+\cdots+b_s\beta_s)\in W_1\cap W_2.$$

于是有

$$a_1\alpha_1+\cdots+a_r\alpha_r=0,\quad b_1\beta_1+\cdots+b_s\beta_s=0.$$

因为B_1和B_2都线性无关, 所以$a_1=\cdots=a_r=0$, $b_1=\cdots=b_s=0$, 即$B_1\cup B_2$线性无关. 最后, 我们有

$$W_1+W_2=\mathrm{Span}(B_1)+\mathrm{Span}(B_2)=\mathrm{Span}(B_1\cup B_2).$$

所以$B_1\cup B_2$是W_1+W_2的基.

“(iv) $\Longrightarrow$ (v)” 显然.

“(v) $\Longrightarrow$ (iii)” 假设(v)成立. 对任意$\alpha\in W_1\cap W_2$, 则存在$\alpha_1,\alpha_2,\dots,\alpha_r\in B_1$和$\beta_1,\beta_2,\dots,\beta_s\in B_2$, 使得

$$\begin{aligned}\alpha=&a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r \qquad (\exists a_1,a_2,\dots,a_r\in\mathbb{F})\\ =&b_1\beta_1+b_2\beta_2+\cdots+b_s\beta_s, \qquad (\exists b_1,b_2,\dots,b_s\in\mathbb{F}).\end{aligned}$$

于是

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r-b_1\beta_1-b_2\beta_2-\cdots-b_s\beta_s=0.$$

由条件$\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s$线性无关, 所以$a_1=\cdots=a_r=0$. 这得到$\alpha=0$, 即$W_1\cap W_2=\{0\}$.

“(iii) $\Longrightarrow$ (i)” 假设(iii)成立, 任取$\alpha\in W_1+W_2$, 设

$$\alpha=\alpha_1+\alpha_2=\beta_1+\beta_2,\quad (\alpha_1,\beta_1\in W_1,\alpha_2,\beta_2\in W_2),$$

则有

$$\alpha_1-\beta_1=\beta_2-\alpha_2\in W_1\cap W_2=\{0\}.$$

于是$\alpha_1=\beta_1$且$\alpha_2=\beta_2$, 即W_1+W_2是直和.

下设W_1和W_2均为有限维子空间.

“(iv) $\Longrightarrow$ (vi)” 显然.

“(vi) $\Longrightarrow$ (iii)” 假设(v)成立, 由维数公式

$$\dim(W_1\cap W_2)=\dim W_1+\dim W_2-\dim(W_1+W_2)=0.$$

所以$W_1\cap W_2=\{0\}$. □

我们看一个几何例子.

例6.31. 设$V=\mathbb{R}^3$.

(1) 如图6.6(a), 取W_1为xOy坐标平面, W_2为Oz轴, 则

$$\mathbb{R}^3=W_1+W_2=W_1\oplus W_2;$$

(2) 如图6.6(b), 取W_1为xOy坐标平面, W_3为xOz坐标平面, 则

$$\mathbb{R}^3=W_1+W_3$$

不是直和. 事实上, $W_1\cap W_3$为Ox轴.

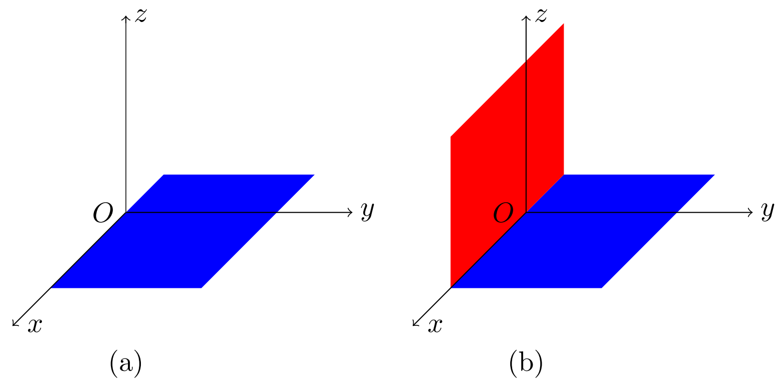


图 6.6: 子空间的和和直和

下面这个例子表明研究方阵时, 某种程度上只需要研究对称和反对称阵即可.

例6.32. 设$V=\mathbb{F}^{n\times n}$, 定义

$$W_1=\{A\in\mathbb{F}^{n\times n}\mid A^T=A\},\quad W_2=\{A\in\mathbb{F}^{n\times n}\mid A^T=-A\},$$

证明: W_1和W_2是V的子空间, 且$V=W_1\oplus W_2$.

证明. 由于$O\in W_1$, 所以$W_1\neq\varnothing$. 任取$A,B\in W_1$和$a,b\in\mathbb{F}$, 我们有

$$(aA+bB)^T=aA^T+bB^T=aA+bB\Longrightarrow aA+bB\in W_1.$$

于是W_1是V的子空间. 类似可证W_2是V的子空间.

如果$A\in W_1\cap W_2$, 则$A\in W_1$且$A\in W_2$, 这得到

$$A=A^T=-A\Longrightarrow A=O\Longrightarrow W_1\cap W_2=\{O\},$$

即W_1+W_2是直和.

最后, 任取$A\in V$, 我们有

$$A=\frac{A+A^T}{2}+\frac{A-A^T}{2}.$$

因为

$$\left(\frac{A+A^T}{2}\right)^T=\frac{A+A^T}{2}\Longrightarrow\frac{A+A^T}{2}\in W_1,$$

$$\left(\frac{A-A^T}{2}\right)^T = \frac{A^T-A}{2} \Longrightarrow \frac{A-A^T}{2} \in W_2,$$

所以$A \in W_1 + W_2$. 于是$V = W_1 + W_2$, 进而$V = W_1 \oplus W_2$. □

整数$n > 1$进行素分解时, 我们先找n的一个素因子p, 进行分解$n = pm$, 然后继续对m进行分解即可. 类似地, 给定线性空间V, 我们先找一个子空间(后面章节会类似于整数的素因子找好的子空间)W, 是否存在子空间U, 使得$V = W \oplus U$? 在整数世界问题很简单, 做除法$m = n/p$即可; 在线性空间的世界呢[8]? 可以用基来如下得到: 任取W的基B_1, 将其扩充为V的基B. 令$B_2 = B - B_1$和$U = \mathrm{Span}(B_2)$, 则U是V的子空间, 且

$$V = \mathrm{Span}(B) = \mathrm{Span}(B_1 \cup B_2) = \mathrm{Span}(B_1) + \mathrm{Span}(B_2) = W + U.$$

进而由定理6.17的(v)知, $W + U = W \oplus U$. 由U的寻找过程可以看出, 通常U不唯一.

推论6.18. 设W是线性空间V的子空间, 则存在V的子空间U, 使得$V = W \oplus U$. 称U为W在V中的直和补(通常不唯一).

我们看一个几何例子.

例6.33. 取$V = \mathbb{R}^2$, W为Ox轴, 设L是过原点的直线, 则当L不等于W时, L是W的直和补, 参看图6.7.

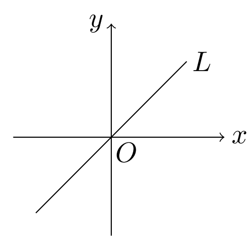


图 6.7: 子空间的直和补

我们最后做一个注记. 设$W_1, W_2, U \subset V$是子空间, 其中$W_1 + W_2 = W_1 \oplus W_2$, 则有$(W_1 \cap U) + (W_2 \cap U)$是直和, 且

$$(W_1 \cap U) \oplus (W_2 \cap U) \subset (W_1 \oplus W_2) \cap U. \tag{6.2}$$

事实上, 由

$$(W_1 \cap U) \cap (W_2 \cap U) \subset W_1 \cap W_2 = \{0\},$$

可知式(6.2)左边为直和; 由$W_i \cap U \subset (W_1 \oplus W_2) \cap U$, $i = 1, 2$, 得式(6.2)成立. 要注意的是, 式(6.2)的等号一般不成立. 例如, 取$V = \mathbb{R}^2$, W_1为Ox轴, W_2为Oy轴, U为过原点的直线L, 但是L不是坐标轴(参看图6.7), 于是$(W_1 \oplus W_2) \cap U = U$, 但是$W_1 \cap U = \{0\}$, $W_2 \cap U = \{0\}$, 得$(W_1 \cap U) \oplus (W_2 \cap U) = \{0\}$.

[8]从下节可以看出本质上我们也是将V除以W得到U.

6.7.2 几个子空间的直和

现在我们考虑几个子空间的直和.

定义6.16. 设$W_1, W_2, \dots, W_r \subset V$是子空间$(r \geqslant 2)$, 如果对任意$\alpha \in W_1 + W_2 + \cdots + W_r$, α可唯一写为

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_r, \quad (\alpha_i \in W_i, i = 1, 2, \dots, r),$$

则称和$W_1 + W_2 + \cdots + W_r$是直和, 记为$W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

类似于定理6.17, 我们有

定理6.19. 设$W_1, W_2, \dots, W_r \subset V$是子空间$(r \geqslant 2)$, 则下面命题等价

(i) $W_1 + W_2 + \cdots + W_r$是直和;

(ii) $0 \in W_1 + W_2 + \cdots + W_r$的分解式唯一: $0 = 0 + 0 + \cdots + 0$;

(iii) 将$W_1, W_2, \dots, W_r$任意分成两组: $W_{i_1}, \dots, W_{i_s}$和$W_{j_1}, \dots, W_{j_t}$ $(s, t > 0, s + t = r)$, 有

$$(W_{i_1} + \cdots + W_{i_s}) \cap (W_{j_1} + \cdots + W_{j_t}) = \{0\};$$

(iii)$'$ 对$i = 1, 2, \dots, r$, 有

$$W_i \cap \sum_{\substack{j=1\\ j\neq i}}^{r} W_j = \{0\};$$

(iv) 任取W_i的基B_i, $i = 1, 2, \dots, r$, 有$i \neq j$时$B_i \cap B_j = \varnothing$, 且$\bigcup\limits_{i=1}^{r} B_i$为$W_1 + W_2 + \cdots + W_r$的基;

(v) 存在W_i的基B_i, $i = 1, 2, \dots, r$, 使得$i \neq j$时$B_i \cap B_j = \varnothing$, 且$\bigcup\limits_{i=1}^{r} B_i$为$W_1 + W_2 + \cdots + W_r$的基.

进而, 若$W_1, W_2, \dots, W_r$均为有限维子空间, 则上面还等价于

(vi) $\dim(W_1 + W_2 + \cdots + W_r) = \dim W_1 + \dim W_2 + \cdots + \dim W_r$.

与两个子空间的直和不同, 如果只满足任意两个子空间W_i和W_j的交为零, 则不能得到$W_1 + W_2 + \cdots + W_r$是直和. 例如, 取$V = \mathbb{R}^2$, W_1为Ox轴, W_2为Oy轴, W_3 为直线$y = x$(参看图6.7), 则这三条直线都只交于原点, 但是$W_1 + W_2 + W_3$不是直和.

☞ **定理6.19的证明.** "(i) $\Longrightarrow$ (ii)" 显然.
"(ii) $\Longrightarrow$ (iii)" 假设(ii)成立, 任取$\alpha \in (W_{i_1} + \cdots + W_{i_s}) \cap (W_{j_1} + \cdots + W_{j_t})$, 有

$$\alpha = \alpha_{i_1} + \cdots + \alpha_{i_s} = \alpha_{j_1} + \cdots + \alpha_{j_t}, \quad (\exists \alpha_{i_k} \in W_{i_k}, \alpha_{j_l} \in W_{j_l}).$$

于是

$$\alpha_{i_1} + \cdots + \alpha_{i_s} - \alpha_{j_1} - \cdots - \alpha_{j_t} = 0.$$

由(ii), 得$\alpha_{i_1} = \cdots = \alpha_{i_s} = 0$, 所以$\alpha = 0$, 即(iii)成立.
"(iii) $\Longrightarrow$ (iii)$'$" 显然.

"(iii)$'$ $\Longrightarrow$ (iv)" 假设(iii)$'$成立. 如果存在$i \neq j$, 使得$\exists \alpha \in B_i \cap B_j$, 则有

$$0 \neq \alpha \in W_i \cap \sum_{k \neq i} W_k = \{0\},$$

矛盾. 所以对任意$i \neq j$, 有$B_i \cap B_j = \varnothing$.

如果存在$\alpha_{i1}, \ldots, \alpha_{it_i} \in B_i$, 使得

$$\sum_{i=1}^{r} (a_{i1}\alpha_{i1} + \cdots + a_{it_i}\alpha_{it_i}) = 0, \quad (a_{ij} \in \mathbb{F}),$$

则对固定i, 由于

$$a_{i1}\alpha_{i1} + \cdots + a_{it_i}\alpha_{it_i} = -\sum_{\substack{j=1\\ j \neq i}}^{r} (a_{j1}\alpha_{j1} + \cdots + a_{jt_j}\alpha_{jt_j}),$$

所以有

$$a_{i1}\alpha_{i1} + \cdots + a_{it_i}\alpha_{it_i} \in W_i \cap \sum_{j \neq i} W_j = \{0\}.$$

因此

$$a_{i1}\alpha_{i1} + \cdots + a_{it_i}\alpha_{it_i} = 0,$$

再利用B_i是基就得$a_{i1} = \cdots = a_{it_i} = 0$. 于是$B_1 \cup B_2 \cup \cdots \cup B_r$线性无关.

最后, 归纳可证

$$W_1 + W_2 + \cdots + W_r = \mathrm{Span}(B_1 \cup B_2 \cup \cdots \cup B_r).$$

于是$B_1 \cup B_2 \cup \cdots \cup B_r$是$W_1 + W_2 + \cdots + W_r$的基.

"(iv) $\Longrightarrow$ (v)" 显然.

"(v) $\Longrightarrow$ (i)" 设(v)成立, 即存在W_i的基B_i, 使得对任意$i \neq j$, 有$B_i \cap B_j = \varnothing$, 且

$$B = \bigcup_{i=1}^{r} B_i$$

是$W_1 + W_2 + \cdots + W_r$的基. 任取$\alpha \in W_1 + W_2 + \cdots + W_r$, 如果

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_r = \beta_1 + \beta_2 + \cdots + \beta_r, \quad (\alpha_i, \beta_i \in W_i),$$

则

$$(\alpha_1 - \beta_1) + (\alpha_2 - \beta_2) + \cdots + (\alpha_r - \beta_r) = 0.$$

因为$\alpha_i - \beta_i \in W_i$, 所以存在$\alpha_{i1}, \ldots, \alpha_{it_i} \in B_i$和$b_{ij} \in \mathbb{F}$, 使得

$$\alpha_i - \beta_i = \sum_{j=1}^{t_i} b_{ij}\alpha_{ij}.$$

这得到

$$\sum_{i=1}^{r} \sum_{j=1}^{t_i} b_{ij}\alpha_{ij} = 0,$$

而B是基, 所以$b_{ij} = 0$, $\forall i, j$. 于是$\alpha_i = \beta_i$, $\forall i$, 即$W_1 + W_2 + \cdots + W_r$是直和.

下设$W_1, W_2, \ldots, W_r$均是有限维的.

"(iv) $\Longrightarrow$ (vi)" 显然.

"(vi) $\Longrightarrow$ (iv)" 设(v)成立. 任取W_i的基B_i, 其中$i = 1, 2, \ldots, r$, 于是有

$$W_1 + W_2 + \cdots + W_r = \mathrm{Span}(B_1 \cup B_2 \cup \cdots \cup B_r).$$

这得到

$$\dim(W_1 + W_2 + \cdots + W_r) \leqslant |B_1 \cup B_2 \cup \cdots \cup B_r|.$$

进而利用(v), 就有

$$\dim W_1 + \dim W_2 + \cdots + \dim W_r \leqslant |B_1 \cup B_2 \cup \cdots \cup B_r|,$$

也就是

$$|B_1| + |B_2| + \cdots + |B_r| \leqslant |B_1 \cup B_2 \cup \cdots \cup B_r|.$$

于是, 对任意$i \neq j$, 有$B_i \cap B_j = \varnothing$, 且$B_1 \cup B_2 \cup \cdots \cup B_r$是$W_1 + W_2 + \cdots + W_r$的基. □

习题6.7

A1. 设$\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{F}^4$, $W_1 = \mathrm{Span}(\{\alpha_1, \alpha_2\})$和$W_2 = \mathrm{Span}(\{\alpha_3, \alpha_4\})$是$\mathbb{F}^4$的子空间, 试判断$\mathbb{F}^4 = W_1 \oplus W_2$是否成立, 并给出理由. 其中
(1) $\alpha_1 = (0,1,0,1)^T$, $\alpha_2 = (0,0,1,0)^T$, $\alpha_3 = (1,0,1,0)^T$, $\alpha_4 = (1,1,0,0)^T$;
(2) $\alpha_1 = (1,0,0,1)^T$, $\alpha_2 = (0,1,1,0)^T$, $\alpha_3 = (1,0,1,0)^T$, $\alpha_4 = (0,1,0,1)^T$.
A2. 设W_1和W_2分别是齐次线性方程组$x_1 + x_2 + \cdots + x_n = 0$和$x_1 = x_2 = \cdots = x_n$在$\mathbb{R}^n$中的解空间, 证明: $\mathbb{R}^n = W_1 \oplus W_2$.
A3. 设V是$\mathbb{R}$所有实值函数组成的实线性空间, W_1是全体偶函数组成的子集, W_2是全体奇函数组成的子集, 证明: W_1和W_2都是V的子空间, 且$V = W_1 \oplus W_2$.
A4. 设V是数域$\mathbb{F}$上的线性空间, W_1和W_2是V的子空间, 且$V = W_1 \oplus W_2$. 又设U_1和U_2是W_2的子空间, 且有$W_2 = U_1 \oplus U_2$. 证明: U_1和U_2也是V的子空间, 且有$V = W_1 \oplus U_1 \oplus U_2$.
A5. 设U和V是数域$\mathbb{F}$上的线性空间, 它们的**Cartesian积**(**笛卡尔积**)

$$U \times V := \{(\alpha, \beta) \mid \alpha \in U, \beta \in V\}$$

在如下的加法和数乘下成为$\mathbb{F}$上的线性空间:

$$\begin{aligned}
&(\alpha_1, \beta_1) + (\alpha_2, \beta_2) := (\alpha_1 + \alpha_2, \beta_1 + \beta_2), \quad (\alpha_1, \alpha_2 \in U, \beta_1, \beta_2 \in V),\\
&c(\alpha, \beta) := (c\alpha, c\beta), \quad (c \in \mathbb{F}, \alpha \in U, \beta \in V).
\end{aligned}$$

定义子集

$$\begin{aligned}
\widetilde{U} &:= \{(\alpha, 0) \in U \times V \mid \alpha \in U\} \subset U \times V,\\
\widetilde{V} &:= \{(0, \beta) \in U \times V \mid \beta \in V\} \subset U \times V,
\end{aligned}$$

和映射

$$\begin{aligned}
\phi_1 &: U \longrightarrow \widetilde{U}; \alpha \mapsto (\alpha, 0),\\
\phi_2 &: V \longrightarrow \widetilde{V}; \beta \mapsto (0, \beta),
\end{aligned}$$

证明:
(1) $\widetilde{U}$和$\widetilde{V}$是$U \times V$的子空间, 并且$U \times V = \widetilde{U} \oplus \widetilde{V}$;
(2) ϕ_1和ϕ_2是同构[9].

[9]于是在同构意义下, $U \times V$是U和V的直和, 可以记为$U \times V = U \oplus V$, 称这种直和为外直和. 而本节中线性空间为其若干子空间的直和也称为内直和.

§ 6.8 商空间

设V是$\mathbb{F}$上的线性空间, W是V的子空间. 如果我们对W不感兴趣, 考察问题时想将W去掉, 如何实现呢? 一种办法是取集合的差$V-W$, 此时由于$0\notin V-W$, 所以$V-W$不是线性空间, 似乎这不是一个好的选择. 数学家们想出了另一种办法: 用V"去除"W, 得到一个比V更简单的线性空间. 本节介绍这个空间的构造和基本性质.

定义6.17. 设$\alpha,\beta\in V$, 如果$\alpha-\beta\in W$, 则称α与β模W同余, 记为

$$\alpha\equiv\beta\pmod W.$$

容易证明模W同余是V中的等价关系, 即满足

(i) (自反性) $\forall\alpha\in V$, 有$\alpha\equiv\alpha\pmod W$;

(ii) (对称性) $\forall\alpha,\beta\in V$, $\alpha\equiv\beta\pmod W\Longrightarrow\beta\equiv\alpha\pmod W$;

(iii) (传递性) $\forall\alpha,\beta,\gamma\in V$, $\alpha\equiv\beta\pmod W$且$\beta\equiv\gamma\pmod W\Longrightarrow\alpha\equiv\gamma\pmod W$.

于是模W同余给出了V的一种分类. 对任意$\alpha\in V$, 记α所在的模W同余类为$\overline{\alpha}$, 即

$$\overline{\alpha}=\{\beta\in V\mid\beta\equiv\alpha\pmod W\}\subset V.$$

于是对$\alpha,\beta\in V$, 有

$$\overline{\alpha}=\overline{\beta}\Longleftrightarrow\alpha\equiv\beta\pmod W.$$

模W同余类还有下面的刻画.

引理6.20. 设$\alpha\in V$, 则

$$\overline{\alpha}=\alpha+W:=\{\alpha+\beta\mid\beta\in W\}.$$

☞ **证明.** 任取$\beta\in W$, 则

$$(\alpha+\beta)-\alpha=\beta\in W\Longrightarrow\alpha+\beta\equiv\alpha\pmod W\Longrightarrow\alpha+\beta\in\overline{\alpha}.$$

于是$\alpha+W\subset\overline{\alpha}$.

反之, 任取$\gamma\in\overline{\alpha}$, 则$\gamma\equiv\alpha\pmod W$, 即有$\gamma-\alpha\in W$, 所以

$$\gamma=\alpha+(\gamma-\alpha)\in\alpha+W.$$

这得到$\overline{\alpha}\subset\alpha+W$. □

我们将V对模W同余这一等价关系的商集合记为V/W, 即V/W是所有模W的同余类所成的集合, 也即

$$V/W=\{\overline{\alpha}\mid\alpha\in V\}.$$

请注意V/W中的元素都是V的子集. 由于$\overline{0}\in V/W$, 所以$V/W\neq\varnothing$. 定义V/W中的加法和数乘为

加法 $\overline{\alpha}+\overline{\beta}:=\overline{\alpha+\beta}$;

数乘 $c\overline{\alpha} := \overline{c\alpha}$.

需要说明上面定义与同余类中代表元的选取无关. 设$\overline{\alpha_1} = \overline{\alpha}$, $\overline{\beta_1} = \overline{\beta}$, 则

$$\alpha_1 \equiv \alpha \pmod{W}, \quad \beta_1 \equiv \beta \pmod{W}.$$

这得到$\alpha_1 - \alpha, \beta_1 - \beta \in W$. 但是$W$是子空间, 所以

$$\begin{aligned}&(\alpha_1 + \beta_1) - (\alpha + \beta) = (\alpha_1 - \alpha) + (\beta_1 - \beta) \in W,\\&c\alpha_1 - c\alpha = c(\alpha_1 - \alpha) \in W.\end{aligned}$$

这得到

$$\overline{\alpha_1 + \beta_1} = \overline{\alpha + \beta}, \qquad \overline{c\alpha_1} = \overline{c\alpha},$$

即与同余类中代表元的选取无关.

可以验证集合V/W在上面定义的加法和数乘下成为$\mathbb{F}$上的线性空间, 称其为V关于W的商空间. 商空间V/W的零元是$0 = \overline{0} = W$, 而$\overline{\alpha} \in V/W$的负元是$-\overline{\alpha} = \overline{-\alpha}$.

我们看一个具体的几何例子.

例6.34. *设$V = \mathbb{R}^2$, W为Ox轴, 即*

$$W = \{(x, 0)^T \mid x \in \mathbb{R}\}.$$

设$\alpha = (x_0, y_0)^T \in \mathbb{R}^2$, 有

$$\overline{\alpha} = \alpha + W = \{(x_0 + x, y_0)^T \mid x \in \mathbb{R}\}.$$

于是$\overline{\alpha}$是y坐标为y_0的平行于Ox轴的直线, 而商空间V/W就是所有平行于Ox轴的直线做成的集合. 参看图6.8.

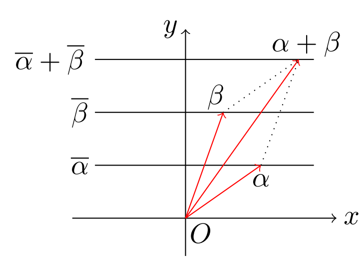


图 6.8: 商空间

下面通过所谓的典范同构来计算商空间的维数和找商空间的基, 并且可以看出直和补本质上是做除法得到的. 设W是V的子空间, 而U是W在V中的直和补, 于是$V = W \oplus U$. 定义映射

$$\phi : U \longrightarrow V/W; \alpha \mapsto \overline{\alpha},$$

我们来证明ϕ是同构, 称ϕ是典范同构.

首先, 对任意$\overline{\alpha} \in V/W$, 其中$\alpha \in V$, 我们有

$$\alpha = \beta + \gamma, \quad \beta \in W, \gamma \in U.$$

于是$\alpha - \gamma = \beta \in W$, 即$\overline{\alpha} = \overline{\gamma}$. 因此$\phi(\gamma) = \overline{\gamma} = \overline{\alpha}$, 即$\phi$是满射. 其次, 如果$\alpha_1, \alpha_2 \in U$满足$\phi(\alpha_1) = \phi(\alpha_2)$, 则$\overline{\alpha_1} = \overline{\alpha_2}$. 这得到$\alpha_1 - \alpha_2 \in W$, 进而有

$$\alpha_1 - \alpha_2 \in W \cap U = \{0\},$$

推出$\alpha_1 = \alpha_2$, 即ϕ是单射. 最后, 任意$\alpha_1, \alpha_2 \in U$和$c \in \mathbb{F}$, 有

$$\begin{aligned}
&\phi(\alpha_1 + \alpha_2) = \overline{\alpha_1 + \alpha_2} = \overline{\alpha_1} + \overline{\alpha_2} = \phi(\alpha_1) + \phi(\alpha_2), \\
&\phi(c\alpha_1) = \overline{c\alpha_1} = c\overline{\alpha_1} = c\phi(\alpha_1),
\end{aligned}$$

即ϕ是线性映射. 这就证明了ϕ是同构.

定理6.21. 设W是线性空间V的子空间.

(1) 设U是W在V中的任意直和补, 则典范映射

$$\phi : U \longrightarrow V/W; \alpha \mapsto \overline{\alpha}$$

是同构, $U \cong V/W$;

(2) 任取W的基B_1, 扩充为V的基B, 令$B_2 = B - B_1$, 则对任意$\alpha \neq \beta \in B_2$有$\overline{\alpha} \neq \overline{\beta}$, 且

$$\overline{B}_2 = \{\overline{\alpha} \mid \alpha \in B_2\}$$

是V/W的基;

(3) 设$\dim V < \infty$, 则

$$\dim(V/W) = \dim V - \dim W.$$

☞ **证明.** (1) 已证.
(2) 令$U = \mathrm{Span}(B_2)$, 则U是W在V中的直和补. 同构

$$\phi : U \longrightarrow V/W; \alpha \mapsto \overline{\alpha}$$

将U的基B_2映为V/W的基, 即$\overline{B}_2$是V/W的基.
(3) 任取W在V中的直和补U, 则有$V = W \oplus U$, $U \cong V/W$. 这得到

$$\dim V = \dim W + \dim U, \quad \dim U = \dim V/W,$$

进而得结论. □

习题6.8

A1. 设$V=\mathbb{F}[x]$, 取V的子空间

$$\begin{aligned} W_1&=\mathbb{F}_n[x]=\{f(x)\in\mathbb{F}[x]\mid \deg(f(x))\leqslant n\},\\ W_2&=\{f(x)\in\mathbb{F}[x]\mid f(-x)=f(x)\}.\end{aligned}$$

证明: 商空间V/W_1和V/W_2都是无限维的.
A2. 设矩阵$A\in\mathbb{F}^{m\times n}$, 向量$\beta\in\mathbb{F}^m$, 满足$AX=\beta$有解. 记齐次线性方程组$AX=0$的解空间为$W$, 任取$AX=\beta$的一个特解$\gamma$, 证明: 线性方程组$AX=\beta$的解集合恰为$\mathbb{F}^n$中向量$\gamma$所在的模$W$同余类.
B1. 设V和W是数域$\mathbb{F}$上的线性空间, $F(V\times W)$是由Cartesian积$V\times W$生成的自由线性空间, 令N是由下面所有的元素生成的$F(V\times W)$ 的子空间:

(i) $(\alpha_1+\alpha_2,\beta)-(\alpha_1,\beta)-(\alpha_2,\beta)$, $\forall\alpha_1,\alpha_2\in V$, $\forall\beta\in W$;

(ii) $(c\alpha,\beta)-c(\alpha,\beta)$, $\forall\alpha\in V$, $\forall\beta\in W$, $\forall c\in\mathbb{F}$;

(iii) $(\alpha,\beta_1+\beta_2)-(\alpha,\beta_1)-(\alpha,\beta_2)$, $\forall\alpha\in V$, $\forall\beta_1,\beta_2\in W$;

(iv) $(\alpha,c\beta)-c(\alpha,\beta)$, $\forall\alpha\in V$, $\forall\beta\in W$, $\forall c\in\mathbb{F}$.

记商空间$F(V\times W)/N$为$V\otimes W$, 称其为线性空间V和W的张量积.
(1) 对$\alpha\in V$和$\beta\in W$, 记$\overline{(\alpha,\beta)}\in V\otimes W$为$\alpha\otimes\beta$, 是否$V\otimes W$中的每个元素一定有形式$\alpha\otimes\beta$? 说明理由;
(2) 定义映射

$$\varphi: V\times W\longrightarrow V\otimes W; (\alpha,\beta)\mapsto\varphi(\alpha,\beta)=\alpha\otimes\beta,$$

其中这里只把$V\times W$看成集合. 证明: φ是双线性映射, 即对任意$\alpha,\alpha_1,\alpha_2\in V$, 任意$\beta,\beta_1,\beta_2\in W$和任意$c\in\mathbb{F}$, 有

$$\begin{aligned}\varphi(\alpha_1+\alpha_2,\beta)&=\varphi(\alpha_1,\beta)+\varphi(\alpha_2,\beta),\\ \varphi(\alpha,\beta_1+\beta_2)&=\varphi(\alpha,\beta_1)+\varphi(\alpha,\beta_2),\\ \varphi(c\alpha,\beta)&=c\varphi(\alpha,\beta)=\varphi(\alpha,c\beta);\end{aligned}$$

(3) 证明: 如果B_1是V的基, B_2是W的基, 则$\{\alpha\otimes\beta\mid\alpha\in B_1,\beta\in B_2\}$是$V\otimes W$的基. 进而, 如果$\dim V<\infty$且$\dim W<\infty$, 则$\dim V\otimes W=\dim V\dim W$;
(4) 证明: $\mathbb{F}^m\otimes\mathbb{F}^n\cong\mathbb{F}^{mn}$.
B2. 设V和W是数域$\mathbb{F}$上的线性空间, $V\otimes W$是V和W的张量积. 证明: 对任意线性空间U和任意双线性映射$\psi: V\times W\longrightarrow U$, 存在唯一的线性映射$\widetilde{\psi}: V\otimes W\longrightarrow U$, 使得$\psi=\widetilde{\psi}\circ\varphi$, 即有交换图

$$\begin{tikzcd} V\times W \arrow[r, "\varphi"] \arrow[dr, "\psi"'] & V\otimes W \arrow[d, dashed, "\widetilde{\psi}"] \\ & U \end{tikzcd}$$

其中$\varphi: V\times W\longrightarrow V\otimes W$是双线性映射

$$\varphi(\alpha,\beta)=\alpha\otimes\beta,\quad(\forall\alpha\in V,\forall\beta\in W).$$

称该性质为张量积的泛性质.

B3. 设V, W, U是数域$\mathbb{F}$上的线性空间, 证明:

(1) $V \otimes W \cong W \otimes V$;

(2) $(V \otimes W) \otimes U \cong V \otimes (W \otimes U)$.

§ 6.9 附录: 线性空间的无限基的势

设V是数域$\mathbb{F}$上的线性空间, B_1和B_2是V的两个无限基, 我们要证明集合的势$|B_1| = |B_2|$. 由对称性, 只需证明$|B_1| \leqslant |B_2|$. 这等价于构造一个单射$f : B_1 \longrightarrow B_2$. 当然, 这不一定可以做到. 这个证明取自[8].

由于B_2是基, 所以对任意$\alpha \in B_1$, 有唯一的分解

$$\alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n, \quad (\beta_i \in B_2, a_i \neq 0, i = 1, 2, \ldots, n).$$

记$K(B_2)$为B_2的所有有限子集所成的集合, 于是有映射

$$g : B_1 \longrightarrow K(B_2); \alpha \mapsto \{\beta_1, \beta_2, \ldots, \beta_n\}.$$

对于集合$K(B_2)$, 我们有下面的一般结论.

引理6.22. *设X是无限集, $K(X)$是X的所有有限子集所成之集, 则$|X| = |K(X)|$.*

证明. 定义映射$\varphi : X \longrightarrow K(X); x \mapsto \{x\}$, 则$\varphi$是单射, 所以$|X| \leqslant |K(X)|$. 又有映射

$$\psi : K(X) \longrightarrow \bigcup_{n=1}^{\infty} X^n; \{x_1, x_2, \ldots, x_n\} \mapsto (x_1, x_2, \ldots, x_n),$$

则ψ是单射, 于是

$$|K(X)| \leqslant \left|\bigcup_{n=1}^{\infty} X^n\right| = |\mathbb{N} \times X| = |X|.$$

因此$|X| = |K(X)|$. □

由上面的引理, 如果g是单射, 那么我们就证明了结论. 但是g不一定为单射, 所以我们需要更仔细地研究g. 我们有下面的断言:

(1) $\mathrm{Im}(g)$为无限集;

(2) $\forall T \in \mathrm{Im}(g)$, $g^{-1}(T)$为有限集;

(3) 有无交并$B_1 = \bigsqcup\limits_{T \in \mathrm{Im}(g)} g^{-1}(T)$.

证明. (1) 如果$\mathrm{Im}(g)$是有限集, 则存在B_2的有限子集B_3, 使得B_1可由B_3线性表示. 而B_2线性无关推出B_3线性无关; $V = \mathrm{Span}(B_1)$推出$V = \mathrm{Span}(B_3)$. 于是B_3是V的基, 这矛盾于V有无限基.

(2) 由于B_1是基, 所以T可由B_1线性表示. 但T是有限集, 所以存在$S \subset B_1$, 使得S是有限集, 且$T \subset \mathrm{Span}(S)$.

任取$\alpha \in g^{-1}(T)$, 有$\alpha \in \mathrm{Span}(T)$. 因此$\alpha \in \mathrm{Span}(S)$. 但是$\alpha \in B_1$, $S \subset B_1$, 所以有$\alpha \in S$(否则, $S \cup \{\alpha\}$线性相关, 这矛盾于B_1线性无关). 因此有$g^{-1}(T) \subset S$是有限集.

(3) 只需证明并是无交并. 设$T_1, T_2 \in \mathrm{Im}(g)$, 且$T_1 \neq T_2$. 如果存在$\alpha \in g^{-1}(T_1) \cap g^{-1}(T_2)$, 则

$$\begin{aligned}\alpha =& a_1\alpha_1 + \cdots + a_n\alpha_n, \quad T_1 = \{\alpha_1, \ldots, \alpha_n\}, a_1, \ldots, a_n \neq 0,\\ =& b_1\beta_1 + \cdots + b_m\beta_m, \quad T_2 = \{\beta_1, \ldots, \beta_m\}, b_1, \ldots, b_m \neq 0.\end{aligned}$$

而$T_1 \neq T_2$, 所以这得到$T_1 \cup T_2$线性相关. 这矛盾于B_2线性无关. □

对任意$T \in \mathrm{Im}(g)$, 记$g^{-1}(T) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, 并且固定这个排序. 于是有映射

$$h : B_1 = \bigsqcup_{T \in \mathrm{Im}(g)} g^{-1}(T) \longrightarrow \mathrm{Im}(g) \times \mathbb{N}; \alpha \in g^{-1}(T) \mapsto (T, k), \quad (\alpha_k = \alpha).$$

易知映射h是单射. 所以得到

$$|B_1| \leqslant |\mathrm{Im}(g) \times \mathbb{N}| \xlongequal{(1)} |\mathrm{Im}(g)| \leqslant |K(B_2)| = |B_2|.$$

于是我们就得到$|B_1| = |B_2|$.

§ 6.10 附录: Fibonacci数列和幻方

本节介绍线性空间的两个简单应用.

6.10.1 Fibonacci数列

Fibonacci数列的定义者, 是意大利数学家列Leonardo Fibonacci. 该数列指的是这样一个数列:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, \ldots$$

这个数列从第3项开始, 每一项都等于前两项之和. 于是记F_n为该数列的第n项, 则有

$$F_1 = F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad (n \geqslant 3).$$

下面的任务是求出该数列的通项公式.

一方面可以按照第四章中利用递归公式的特征方程计算, 另一方面也可以利用线性空间来计算. 设$n \geqslant 3$固定, 定义$\mathbb{C}^n$的一个子空间

$$V = \{(a_1, a_2, \ldots, a_n)^T \in \mathbb{C}^n \mid a_i = a_{i-1} + a_{i-2}, i = 3, \ldots, n\}.$$

于是有

$$(F_1, F_2, \ldots, F_n)^T \in V.$$

下面我们具体求出V的维数和一组基. 由于对任意的$(a_1, \ldots, a_n)^T \in V$, 有$a_3, \ldots, a_n$可由$a_1, a_2$确定, 所以可以想象$V$中向量的自由度是2, 即$V$的维数是2. 事实上, 定义映射

$$\varphi : \mathbb{C}^2 \to V; \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}, a_i = a_{i-1} + a_{i-2}, (\forall i \geqslant 3),$$

则容易验证φ是线性空间之间的同构. 这得到

$$\dim V = \dim \mathbb{C}^2 = 2,$$

且对$\mathbb{C}^2$的任意基α_1, α_2, 有$\varphi(\alpha_1), \varphi(\alpha_2)$是$V$的基. 进而如果

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = c_1\alpha_1 + c_2\alpha_2,$$

则用φ作用得到

$$\begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix} = c_1\varphi(\alpha_1) + c_2\varphi(\alpha_2).$$

于是就可以得到F_n的通项公式. 所以我们现在的目标是

(i) 找$\mathbb{C}^2$的基α_1, α_2, 使得$\varphi(\alpha_1), \varphi(\alpha_2)$容易确定;

(ii) 求c_1, c_2.

当α_1, α_2确定后, 容易得到c_1, c_2. 于是我们的主要任务的找合适的α_1, α_2. 我们先看V中可能有哪些数列. 最简单的数列当然是等差和等比数列了. 设

$$\alpha = \begin{pmatrix} a \\ a+d \\ a+2d \\ \vdots \\ a+(n-1)d \end{pmatrix}$$

是等差数列, 则$\alpha \in V$的充分和必要条件是

$$a + (i-1)d = a + (i-2)d + a + (i-3)d, \quad i = 3, \ldots, n.$$

而后者又等价于

$$a + (i-4)d = 0, \quad i = 3, \ldots, n.$$

这得到$\alpha \in V$当且仅当, 当$n = 3$时$a = d$, 而当$n \geqslant 4$时, $a = d = 0$. 所以当$n \geqslant 4$时, V中只有平凡的等差数列.

再看等比数列. 设

$$\alpha = \begin{pmatrix} a \\ aq \\ \vdots \\ aq^{n-1} \end{pmatrix}, \quad (a, q \neq 0)$$

是非平凡的等比数列, 则$\alpha \in V$的充分和必要条件是

$$aq^{i-1} = aq^{i-2} + aq^{i-3}, \quad i = 3, \ldots, n.$$

而后者又等价于

$$q^2 = q + 1,$$

即$q = \frac{1\pm\sqrt{5}}{2}$. 令

$$q_1 = \frac{1+\sqrt{5}}{2}, \quad q_2 = \frac{1-\sqrt{5}}{2},$$

则$\mathbb{C}^2$有基

$$\alpha_1 = \begin{pmatrix} 1 \\ q_1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 1 \\ q_2 \end{pmatrix},$$

而

$$\varphi(\alpha_1) = \begin{pmatrix} 1 \\ q_1 \\ \vdots \\ q_1^{n-1} \end{pmatrix}, \quad \varphi(\alpha_2) = \begin{pmatrix} 1 \\ q_2 \\ \vdots \\ q_2^{n-1} \end{pmatrix}$$

是V的基.

设

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = c_1\alpha_1 + c_2\alpha_2,$$

则可得

$$c_1 = \frac{q_1}{q_1 - q_2}, \quad c_2 = \frac{-q_2}{q_1 - q_2}.$$

于是

$$F_n = c_1 q_1^{n-1} + c_2 q_2^{n-1} = \frac{q_1^n - q_2^n}{q_1 - q_2} = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}.$$

在下一章中, 我们会给出另一种计算F_n的通项公式的线性代数的办法.

6.10.2 幻方

将数字安排在正方形格子中, 使每行, 每列和每条对角线上的数字和都相等, 所得到者称之为幻方. 具体的, 在$n \times n$的方格中填入正整数$1, 2, \ldots, n^2$, 使得每行, 每列和每条对角线上的n个数字之和都相等, 就得到一个n阶幻方. 例如, 下面就是一个3阶幻方:

6	1	8
7	5	3
2	9	4

下面以$n = 3$为例, 给出构造n阶幻方的一种方法.

可以将幻方看成方阵, 于是我们要找3阶方阵, 使得它的9个元素恰好为$1, 2, \ldots, 9$, 且这个方阵的三个行, 三个列, 两条对角线上的3个元之和都相等(等于15). 定义

$$V = \{A \in \mathbb{R}^{3\times 3} \mid A\text{的每行, 每列, 每条对角线上三个数之和相等}\} \subset \mathbb{R}^{3\times 3},$$

则三阶幻方是V中特殊的矩阵. 容易验证V是$\mathbb{R}^{3\times 3}$的子空间, 且零矩阵和

$$H=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$$

都属于V. 如果A是一个三阶幻方, 则$A-H\in V$, 且$A-H$的9个元为$0,1,\ldots,8$. 而这9个数字可由

$$a=3q+r,\quad (0\leqslant q,r\leqslant 2)$$

得到, 所以我们考虑每行, 每列的三个元都为$0,1,2$的V中的矩阵. 这样的矩阵有

$$A_1=\begin{pmatrix}1&2&0\\0&1&2\\2&0&1\end{pmatrix},\quad A_2=\begin{pmatrix}1&0&2\\2&1&0\\0&2&1\end{pmatrix}$$

和

$$B_1=\begin{pmatrix}0&2&1\\2&1&0\\1&0&2\end{pmatrix},\quad B_2=\begin{pmatrix}2&0&1\\0&1&2\\1&2&0\end{pmatrix}.$$

计算

$$3A_i+B_j+H\quad (i,j=1,2)$$

和

$$3B_j+A_i+H\quad (i,j=1,2),$$

就得到如下8个三阶幻方

4	9	2
3	5	7
8	1	6

6	7	2
1	5	9
8	3	4

4	3	8
9	5	1
2	7	6

6	1	8
7	5	3
2	9	4

2	9	4
7	5	3
6	1	8

2	7	6
9	5	1
4	3	8

8	3	4
1	5	9
6	7	2

8	1	6
3	5	7
4	9	2

事实上, 从上面的某一个三阶幻方出发, 通过绕中心逆(顺)时针旋转$\frac{\pi}{2},\pi,\frac{3\pi}{2}$, 以及相应于垂直(水平)中位线做反射, 可以得到其它的7个三阶幻方.

§ 6.11 附录: 模

6.11.1 模

线性空间是最简单的模.

定义6.18. 设R是环, 则左R-模是一个加法Abel群M, 且有运算$R \times M \to M; (r,a) \mapsto ra$, 满足

(M1) 对任意$r \in R$和$a, b \in M$, 有$r(a+b) = ra + rb$;

(M2) 对任意$r, s \in R$和$a \in M$, 有$(r+s)a = ra + sa$;

(M3) 对任意$r, s \in R$和$a \in M$, 有$r(sa) = (rs)a$.

如果R是含幺环, 且还满足

(M4) 对任意$a \in M$, 有$1_R a = a$,

则称M是含幺左R-模. 类似地, 可以定义右R-模, 含幺右R-模等概念.

当R是交换环时, 对于左R-模M, 定义

$$ar = ra, \quad a \in M, r \in R$$

可以自然得到M的一个右R-模结构, 所以交换环R上的左(右)R-模就简称为R-模.

例6.35. **(1)** $\mathbb{Z}$-模就是加法Abel群;

(2) 设$\mathbb{F}$是域, 则$\mathbb{F}$上的线性空间就是含幺$\mathbb{F}$-模.

例6.36. 设S是环, R是S的子环, 则S自然成为R-模: $rs (r \in R, s \in S)$为S中的乘法. 特别地, R为R-模.

可以将线性空间的许多概念平行推广到模上, 有一些类似的结果, 也有不一样的地方.

6.11.2 子模和商模

从现在开始, 都假设R是含幺交换环.

定义6.19. 设M是R-模. 如果M的非空子集N满足: 对任意$a, b \in N$和$r \in R$, 有

$$a + b \in N, \quad ra \in N,$$

则称N是M的一个子R-模.

可以证明, N是子R-模当且仅当对任意$r, s \in R$和任意$a, b \in N$, 有

$$ra + sb \in N.$$

设$\mathbb{F}$是域, 则子$\mathbb{F}$-模就是线性空间的子空间. 而R-模R的子模就是R的理想.

我们可以定义子模的运算. 设M是R-模, 则M的任意多个子模的交还是M的子模. 又若N_1, N_2, $\ldots$, N_n都是M的子模, 则它们的和

$$N_1 + N_2 + \cdots + N_n := \{a_1 + a_2 + \cdots + a_n \mid a_i \in N_i, i = 1, 2, \ldots, n\}$$

也是M的子模. 如果任意$a \in N_1 + N_2 + \cdots + N_n$, a写为$N_1, N_2, \ldots, N_n$中元素的和的形式唯一, 即如果

$$a = a_1 + a_2 + \cdots + a_n = b_1 + b_2 + \cdots + b_n, \quad a_i, b_i \in N_i,$$

那么

$$a_i = b_i, \quad i = 1, 2, \ldots, n,$$

则称和$N_1 + N_2 + \cdots + N_n$是直和, 记为$N_1 \oplus N_2 \oplus \cdots \oplus N_n$. 可以证明, 和$N_1 + N_2 + \cdots + N_n$是直和的充分必要条件是, 对任意$i = 1, 2, \ldots, n$, 有

$$N_i \cap \left(\sum_{j \neq i} N_j \right) = \{0\}.$$

设S是R-模M的非空子集, 令

$$\langle S \rangle = \mathrm{Span}(S) = \{r_1 a_1 + \cdots + r_n a_n \mid n \in \mathbb{N}, r_1, \ldots, r_n \in R, a_1, \ldots, a_n \in S\},$$

则Span(S)是M的子模, 称为由S生成的子模. 由一个元素生成的子模, 即

$$\langle a \rangle = Ra = \{ra \mid r \in R\},$$

称为由a生成循环子模.

定义6.20. *设M是R-模, 如果存在M的有限非空子集S, 使得$M = \mathrm{Span}(S)$, 则称M是*有限生成的.

设N是R-模M的子模, 称

$$a + N = \{a + x \mid x \in N\}, \quad a \in M$$

为N在M中的一个陪集. 记M/N为N在M中所有陪集做成的集合, 则M/N在运算

$$(a + N) + (b + N) = (a + b) + N, \quad r(a + N) = ra + N$$

下成为R-模, 称为M关于N的商模.

6.11.3 模同态

定义6.21. *设M和N是R-模, 如果映射$\varphi : M \to N$满足: 对任意$a, b \in M$和$r \in R$, 有*

$$\varphi(a + b) = \varphi(a) + \varphi(b), \qquad \varphi(ra) = r\varphi(a),$$

则称φ是一个R-模同态. *称单射的同态为*单同态, *满射的同态为*满同态, *双射的同态为*同构. *如果φ是同构, 则称M和N同构, 记为$M \cong N$.*

如果$\mathbb{F}$是域, 则$\mathbb{F}$-模同态就是线性空间之间的线性映射.

设$\varphi : M \to N$是R-模同态, 则φ的核和像分别定义为

$$\mathrm{Ker}(\varphi) = \{a \in M \mid \varphi(a) = 0\}$$

核
$$\mathrm{Im}(\varphi) = \{\varphi(a) \mid a \in M\},$$
它们分别是M和N的子模.

设N是R-模M的子模, 则有典范同态
$$\varphi : M \to M/N; a \mapsto a + N.$$
该同态是满同态, 且$\mathrm{Ker}(\varphi) = N$.

命题6.23 (第一同构定理)**.** *设$\varphi : M \to N$是R-模同态, 则有R-模同构: $M/\mathrm{Ker}(\varphi) \cong \mathrm{Im}(\varphi)$.*

☞ **证明.** 定义映射$\widetilde{\varphi} : M/\mathrm{Ker}(\varphi) \to \mathrm{Im}(\varphi)$, 使得对任意$a + \mathrm{Ker}(\varphi)$, 有
$$\widetilde{\varphi}(a + \mathrm{Ker}(\varphi)) = \varphi(a) \in \mathrm{Im}(\varphi).$$

首先, 该映射是良定的. 事实上, 设$a+\mathrm{Ker}(\varphi) = b+\mathrm{Ker}(\varphi)$, 则$a-b \in \mathrm{Ker}(\varphi)$. 于是$\varphi(a-b) = 0$, 即$\varphi(a) = \varphi(b)$. 其次, 容易验证$\widetilde{\varphi}$是$R$-模同态. 再次, 如果$a + \mathrm{Ker}(\varphi) \in \mathrm{Ker}(\widetilde{\varphi})$, 则$\varphi(a) = \widetilde{\varphi}(a + \mathrm{Ker}(\varphi)) = 0$. 于是$a \in \mathrm{Ker}(\varphi)$, 进而$a + \mathrm{Ker}(\varphi) = 0$, 即$\widetilde{\varphi}$是单射. 最后, 对任意$\varphi(a) \in \mathrm{Im}(\varphi)$, 有$\widetilde{\varphi}(a + \mathrm{Ker}(\varphi)) = \varphi(a)$, 即$\widetilde{\varphi}$是满射. □

6.11.4 自由模

下面将线性空间基的概念推广到R-模. 设M是R-模, S是M的非空子集. 如果存在$n \geqslant 1$, 存在$a_1, a_2, \ldots, a_n \in S$, 存在不全为零的$r_1, r_2, \ldots, r_n \in R$, 使得
$$r_1 a_1 + r_2 a_2 + \cdots + r_n a_n = 0,$$
则称S是R-线性相关的. 否则称S是R-线性无关的.

定义6.22. *设M是R-模, B是M的非空子集, 如果B线性无关, 且生成M, 则称B是M的R-基. 具有R-基的R-模M称为自由R-模, 如果B是M的基, 称M在B上自由.*

对任意$n \in \mathbb{N}$, 定义
$$R^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in R\}.$$
则R^n在分量加法和分量乘法下成为R-模. 对$i = 1, 2, \ldots, n$, 记
$$e_i = (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i}),$$
则$e_1, e_2, \ldots, e_n$是R^n的一组R-基, R^n是自由R-模.

类似于线性空间的任意两个基有相同的势, 自由模的任意两个基也有相同的势.

命题6.24. *设M是自由R-模, 则M的任意两个基有相同的势. 将M的基的势称为自由模M的秩, 记为$\mathrm{rank}(M)$.*

为了证明这个结论, 我们需要一些环论的结果. 设I是环R的理想, 则I是R-模R的子模, 而商模R/I可以成为环, 其中乘法为

$$(a+I)(b+I)=ab+I,\quad a,b\in R.$$

称R/I为R关于理想I的商环.

引理6.25. *设R是含幺交换环, 则*

(1) *R有极大理想;*

(2) *R的理想I是极大理想的充分必要条件是R/I是域.*

☞ **证明.** (1) 定义S为R的所有不等于R的理想做成的集合. 由于$\{0\}\in S$, 所以S不是空集. 任取S中的一个链(关于集合的包含关系)$\{I_i\}_{i\in\Lambda}$, 令

$$I=\bigcup_{i\in\Lambda}I_i,$$

则可证I是R的理想. 而$1\notin I$, 所以$I\in S$. 于是S的任意链有上界, 由Zorn引理, S有极大元I_m, 易知I_m是R的极大理想.

(2) 设I是R的极大理想, 任取$0\neq a+I\in R/I$, 则$a\notin I$. 于是$I\subsetneqq\langle I,a\rangle$, 这得到$\langle I,a\rangle=R$, 特别有$r\in R$和$b\in I$, 使得

$$1=ra+b.$$

于是

$$(a+I)(r+I)=ar+I=(1-b)+I=1+I,$$

即$a+I$可逆. 所以R/I是域.

反之, 设R/I是域. 理想J满足$I\subset J\subset R$. 如果$J\neq I$, 则存在$a\in J$, 而$a\notin I$. 于是R/I中的元$a+I\neq 0$, 进而存在$b\in R$, 使得

$$(a+I)(b+I)=1+I.$$

这得到

$$ab-1\in I\subset J.$$

由$a\in J$, 得$1\in J$. 所以$J=R$, 进而I是极大理想. □

下面可以证明命题6.24.

☞ **命题6.24的证明.** 任取R的极大理想I, 则$\mathbb{F}=R/I$是域. 令

$$IM=\{r_1a_1+r_2a_2+\cdots+r_na_n\mid n\geqslant 1,r_i\in I,a_i\in M,i=1,2,\ldots,n\},$$

则IM是M的子模. 令$\overline{M}=M/IM$.

定义$\mathbb{F}$对$\overline{M}$的数乘如下:

$$(r+I)(a+IM)=ra+IM,\quad r\in R,a\in M.$$

容易验证这个运算是良定的, 且$\overline{M}$在其加法和该数乘下成为$\mathbb{F}$上的线性空间.

任取M的R-基B, 令

$$\overline{B}=\{b+IM\mid b\in B\}\subset\overline{M}.$$

我们证明:

(i) B和$\overline{B}$有相同的势;

(ii) $\overline{B}$是$\mathbb{F}$上的线性空间$\overline{M}$的一组基.

由此得

$$|B| = \dim_{\mathbb{F}} \overline{M},$$

与B的选取无关.

先证(i). 有映射$f : B \to \overline{B}; b \mapsto b + IM$, 则$f$是满射. 设$b_1, b_2 \in B$, $b_1 \neq b_2$, 而$f(b_1) = f(b_2)$, 则$b_1 + IM = b_2 + IM$, 即$b_1 - b_2 \in IM$. 于是

$$b_1 - b_2 = r_1 a_1 + r_2 a_2 + \cdots + r_s a_s, \quad r_1, r_2, \ldots, r_s \in I, a_1, a_2, \ldots, a_s \in M.$$

由于B是M的R-基, 所以a_i可以写成B的R-组合. 设

$$a_i = l_i b_1 + \sum_{b \in B, b \neq b_1} l_b b, \quad l_i, l_b \in R,$$

则有

$$b_1 - b_2 = \left(\sum_{i=1}^{s} r_i l_i \right) b_1 + \text{其他项}.$$

所以得到

$$1 = \sum_{i=1}^{s} r_i l_i \in I,$$

进而$I = R$, 矛盾于I是极大理想. 于是f是单射, 进而为双射, (i)得证.

再证(ii). 由于$M = \operatorname{Span}_R(B)$, 所以$\overline{M} = \operatorname{Span}_{\mathbb{F}}(\overline{B})$. 如果有

$$\sum_{i=1}^{s} (r_i + I)(b_i + IM) = 0, \quad r_i \in R, b_i \in B,$$

则

$$\sum_{i=1}^{s} (r_i b_i + IM) = 0,$$

即

$$\sum_{i=1}^{s} r_i b_i \in IM.$$

于是

$$\sum_{i=1}^{s} r_i b_i = \sum_{j=1}^{n} l_j a_j, \quad \exists l_j \in I, a_j \in M.$$

再将a_j写成B的R-线性组合, 比较b_i的系数, 可得$r_i \in I$. 于是

$$r_i + I = 0, \quad i = 1, 2, \ldots, s,$$

即$\overline{B}$在$\mathbb{F}$上线性无关. □

两个(有限维)线性空间同构, 当且仅当它们有相同的维数. 在个结论对自由模也成立.

定理6.26. *两个自由R-模同构当且仅当它们有相同的秩. 特别地, 如果M为秩n的自由R-模, 则$M \cong R^n$.*

☞ **证明.** 设M和N是自由R-模. 如果有同构$\varphi : M \to N$, 任取M的基B, 则$\varphi(B)$是N的基. 于是

$$\operatorname{rank}(M) = |B| = |\varphi(B)| = \operatorname{rank}(N).$$

反之, 设$\operatorname{rank}(M) = \operatorname{rank}(N)$. 任取$M$的基$B_1$和$N$的基$B_2$, 则$|B_1| = |B_2|$. 于是有双射$\varphi : B_1 \to B_2$. 将$\varphi$线性扩充到$M$上, 就得到$R$-同构$\varphi : M \to N$. □

6.11.5 Noether模

最后讨论何时有限生成R-模的子模也是有限生成的. 首先给出Noether模的等价定义.

定理6.27. *设M是R-模. 则下面命题等价*

(i) *M的子模组成的任意非空集合有极大元(集合包含关系下);*

(ii) *M满足*子模的升链条件*, 即任给M的子模升链*

$$M_1 \subset M_2 \subset M_3 \subset \cdots \subset M,$$

存在$k \in \mathbb{N}$, 使得$M_k = M_{k+1} = M_{k+2} = \cdots$;

(iii) *M的每个子模都是有限生成的.*

*如果M满足上面的等价命题, 则称M为***Noether模**.

如果R-模R是Noether模, 即R的每个理想是有限生成的, 则称R是**Noether环**. 例如, 主理想整环是Noether环, 特别的域是Noether环.

有下面重要的

定理6.28. *如果R是Noether环, 则任意有限生成R-模是Noether模.*

于是Noether环上的有限生成模的子模也是有限生成的. 如果R是主理想整环, 则有更好的结论([7, 定理4.2.1]): 主理想整环R上的自由模M的子模N也是自由的, 且$\operatorname{rank}(N) \leqslant \operatorname{rank}(M)$.

著名的**Hilbert基定理**给出了构造新的Noether环的一种办法.

定理6.29 (Hilbert基定理)**.** *设R是Noether环, 则多项式环$R[x]$也是Noether环.*

我们省略Hilbert基定理的证明, 感兴趣的读者可以参看交换代数的教材. 这里给出定理6.27和定理6.28的证明.

☞ **定理6.27的证明.** 设(i)成立. 任取M的子模升链

$$M_1 \subset M_2 \subset M_3 \subset \cdots \subset M,$$

令

$$S = \{M_i \mid i = 1, 2, \ldots\}.$$

由(i), S有极大元M_k, $k \in \mathbb{N}$. 于是

$$M_i \subset M_k, \quad \forall i,$$

进而有

$$M_k = M_{k+1} = M_{k+2} = \cdots .$$

(ii)成立.

假设(ii)成立. 任取M的子模N. 任取$a_1 \in N$, 如果$N_1 = \langle a_1 \rangle = N$, 则$N$有限生成; 否则, 存在$a_2 \in N$, 而$a_2 \notin N_1$. 如果$N_2 = \langle a_1, a_2 \rangle = N$, 则$N$有限生成; 否则, 存在$a_3 \in N$, 而$a_3 \notin N_2$. 继续这一讨论, 如果这一过程有限次后停止, 则N是有限生成的. 否则就有N的子模链

$$N_1 \subsetneqq N_2 \subsetneqq N_3 \subsetneqq \cdots ,$$

这也是M的子模链, 矛盾于(ii). 所以N是有限生成的, (iii)成立.

再取M的子模组成的非空集合S. 由(ii), S中的任意链都有上界. 于是由Zorn引理, S有极大元, (i)成立.

最后假设(iii)成立. 任取M的子模升链

$$M_1 \subset M_2 \subset M_3 \subset \cdots \subset M,$$

令

$$N = \bigcup_{i \geqslant 1} M_i.$$

则N是M的子模, 由(iii)它是有限生成的. 设$N = \langle a_1, \ldots, a_n \rangle$, 其中$a_i \in N$. 存在$k_i \in \mathbb{N}$, 使得$a_i \in M_{k_i}$. 令$k = \max\{k_1, \ldots, k_n\}$, 则

$$a_i \in M_k, \quad i = 1, 2, \ldots, n.$$

于是

$$N \subset M_k \subset M_{k+1} \subset \cdots \subset N,$$

得到

$$M_k = M_{k+1} = M_{k+2} = \cdots .$$

这就证明了(ii)成立. □

☞ **定理6.28的证明.** 设$M = \mathrm{Span}(\{a_1, \ldots, a_n\})$是有限生成$R$-模. 有$R$-模满同态

$$\varphi : R^n \to M; (r_1, r_2, \ldots, r_n) \mapsto r_1 a_1 + r_2 a_2 + \cdots + r_n a_n.$$

设N是M的R-子模, 则$\varphi^{-1}(N)$是R^n的子模, 且$N = \varphi(\varphi^{-1}(N))$. 设$\varphi^{-1}(N)$是有限生成的, 即

$$\varphi^{-1}(N) = \langle \alpha_1, \alpha_2, \ldots, \alpha_k \rangle, \quad \exists \alpha_1, \alpha_2, \ldots, \alpha_k \in \varphi^{-1}(N).$$

对任意$a \in N$, 存在$\alpha \in \varphi^{-1}(N)$, 使得$a = \varphi(\alpha)$. 设

$$\alpha = r_1 \alpha_1 + r_2 \alpha_2 + \cdots + r_k \alpha_k, \quad \exists r_1, r_2, \ldots, r_k \in R,$$

则

$$a = \varphi(\alpha) = r_1 \varphi(\alpha_1) + r_2 \varphi(\alpha_2) + \cdots + r_k \varphi(\alpha_k).$$

因此有$N = \mathrm{Span}(\{\varphi(\alpha_1), \varphi(\alpha_2), \ldots, \varphi(\alpha_k)\})$, 即$N$是有限生成的.

于是只要证明R^n的每个子模都是有限生成的. 对n归纳. 当$n = 1$时, 由R是Noether环得结论. 下设$n > 1$, 任取R^n的子模N, 定义

$$N_1 = \{\alpha = (r_1, \ldots, r_{n-1}, 0) \in N \mid r_1, \ldots, r_{n-1} \in R\}$$

和

$$N_2 = \{(0, \ldots, 0, r_n) \mid \exists r_1, \ldots, r_{n-1} \in R, \text{ 使得 } (r_1, \ldots, r_{n-1}, r_n) \in N\}.$$

容易知道, N_1同构于R^{n-1}的一个子模, 而N_2同构于R的一个子模(将零坐标去掉). 于是由归纳假设, N_1和N_2都是有限生成的. 设

$$N_1 = \langle \alpha_1, \ldots, \alpha_k \}, \quad N_2 = \langle \beta_1, \ldots, \beta_l \rangle, \quad \beta_j = (0, \ldots, 0, b_j),$$

则存在

$$\gamma_j = (r_{j1}, \ldots, r_{j,n-1}, b_j) \in N, \quad j = 1, 2, \ldots, l.$$

下面证明: $N = \langle \alpha_1, \ldots, \alpha_k, \gamma_1, \ldots, \gamma_l \rangle$.

事实上, 任取$\alpha = (r_1, \ldots, r_n) \in N$, 则$(0, \ldots, 0, r_n) \in N_2$. 于是

$$(0, \ldots, 0, r_n) = \sum_{j=1}^{l} a_j \beta_j, \quad \exists a_j \in R.$$

比较最后一个分量得到

$$r_n = \sum_{j=1}^{l} a_j b_j.$$

于是$\alpha - \sum_{j=1}^{l} a_j\gamma_j \in N$的最后一个分量为零, 即$\alpha - \sum_{j=1}^{l} a_j\gamma_j \in N_1$. □

§ 6.12 补充题

A1. 设$V = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Q}\}$, 证明: 在通常的实数加法和乘法下, V成为有理数域上的线性空间. 并求V的维数.

A2. 取$\mathbb{R}^n$的通常拓扑, 证明: $\mathbb{R}^n$的任意子空间为$\mathbb{R}^n$中的闭子集.

A3. 在数域$\mathbb{F}$上n维向量空间$\mathbb{F}^n$中, 求向量$\alpha = (a_1, a_2, \cdots, a_n)^T$在基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$下的坐标, 其中对$j = 1, 2, \ldots, n$, 有

$$\alpha_j = (\underbrace{1, \ldots, 1}_{j\text{个}}, 0, \ldots, 0)^T.$$

A4. 设$W_1, W_2, \ldots, W_r$是线性空间V的子空间, 其中$r \geqslant 2$, 证明: 和$W_1 + W_2 + \cdots + W_r$是直和的充分必要条件是, 对任意的$2 \leqslant i \leqslant r$, 有

$$W_i \cap (W_1 + W_2 + \cdots + W_{i-1}) = \{0\}.$$

A5. 设I是指标集, W_i $(i \in I)$是线性空间V的子空间, 定义V的子集

$$\sum_{i\in I} W_i = \{\alpha_1 + \alpha_2 + \cdots + \alpha_s \mid s \geqslant 1, \alpha_j \in \cup_{i\in I} W_i, j = 1, 2, \ldots, s\},$$

证明: $\sum\limits_{i\in I} W_i$为V中包含所有W_i $(i \in I)$的最小的子空间. 称其为W_i $(i \in I)$的和.

A6. 设I是指标集, W_i $(i \in I)$是线性空间V的子空间, 如果对任意$i \in I$, 有

$$W_i \cap \left(\sum_{j\in I, j\neq i} W_j \right) = \{0\},$$

则称和$\sum\limits_{i\in I} W_i$为直和, 记为$\bigoplus\limits_{i\in I} W_i$. 证明: 下面命题等价

(i) 和$\sum\limits_{i\in I} W_i$是直和;

(ii) 0不能写为不同的W_i中的非零元之和;

(iii) $\sum\limits_{i\in I} W_i$中的非零元写为W_i $(i \in I)$的非零元之和的方式唯一.

B1. 设m是正整数, $R = \mathbb{Z}/m\mathbb{Z}$是模$m$的剩余类环, V是R上的所有复值函数做成的集合, 即

$$V = \{f : R \longrightarrow \mathbb{C} \mid f\text{是函数}\}.$$

(1) 证明: V在函数的加法和数乘运算下成为复线性空间;

(2) 令$\zeta=\zeta_m=e^{\frac{2\pi}{m}i}$是$m$次本原单位根, 对任意$a\in R$, 定义$\varphi_a,\delta_a\in V$, 使得对任意$x\in R$, 有

$$\varphi_a(x)=\zeta^{ax},\qquad \delta_a(x)=\begin{cases}1, & \text{如果}x=a,\\ 0, & \text{否则}.\end{cases}$$

证明:

(i) 对任意$a\in R$, 成立

$$\sum_{x\in R}\varphi_{-a}(x)\varphi_x=m\delta_a;$$

(ii) 对任意$f\in V$, 令$\widehat{f}=\frac{1}{m}\sum\limits_{x\in R}f(x)\varphi_{-x}$, 称其为$f$的有限Fourier变换, 则

$$f=\sum_{a\in R}\widehat{f}(a)\varphi_a,$$

称上式为f的Fourier展开;

(iii) V有基$\{\varphi_a\mid a\in R\}$, 进而$\dim V=m$.

B2. 求$\max\{\dim V\mid V$是$\mathbb{R}^{n\times n}$的子空间, 且对任意$X,Y\in V$, 有$\operatorname{Tr}(XY)=0\}$.

B3. 设$\alpha_1,\alpha_2,\ldots,\alpha_n$和$\beta_1,\beta_2,\ldots,\beta_n$是$n$维线性空间$V$的两组基, $1\leqslant m<n$, 证明: 存在$1,2,\ldots,n$的排列$i_1,i_2,\ldots,i_n$, 使得$\alpha_1,\ldots,\alpha_m,\beta_{i_{m+1}},\ldots,\beta_{i_n}$和$\beta_{i_1}$, …, β_{i_m}, α_{m+1}, …, α_n仍是V的两组基. (提示: 如果n阶方阵A可逆, 通过换列, 可使A的左上角的m阶方阵和右下角的$n-m$阶方阵都可逆.)

B4. 设$W_1,W_2,\ldots,W_k$都是线性空间V的真子空间, $V\neq\{0\}$, $k\in\mathbb{N}$, 证明: 存在V的基B, 使得$B\cap(W_1\cup W_2\cup\cdots\cup W_k)=\varnothing$.

B5. 设V是数域$\mathbb{F}$上的有限维线性空间, $W_1,W_2,\ldots,W_k$是V的维数相同的子空间, 证明: W_1, W_2, …, W_k有公共的直和补, 即存在V的子空间U, 使得对$i=1,2,\ldots,k$都有$V=W_i\oplus U$.

B6. 设$W_1,W_2,\ldots,W_k$是n维线性空间V的子空间, 且$m<n$. 证明: 如果对任意i $(1\leqslant i\leqslant k)$, 成立$\dim W_i\leqslant m$, 则存在$V$的子空间$U$, 使得$\dim U=n-m$, 且对任意$1\leqslant i\leqslant k$, 都有$U\cap W_i=\{0\}$.

B7. 设V是有限维实线性空间, 对任意$\alpha,\beta\in V$, V中连接α和β的线段定义为

$$L(\alpha,\beta)=\{t\alpha+(1-t)\beta\mid 0\leqslant t\leqslant 1\}.$$

设W是V的真子空间, 对任意$\alpha,\beta\in V-W$, 如果$L(\alpha,\beta)\cap W=\varnothing$, 则记$\alpha\equiv\beta$.

(1) 证明: 如果$\dim W=\dim V-1$, 则$\equiv$为集合$V-W$中的等价关系, 且商集合$(V-W)/\equiv$为二元集;

(2) 如果不要求$\dim W=\dim V-1$, 那么$\equiv$是否还是集合$V-W$中的等价关系? 说明理由.

参考文献

[1] 常庚哲, 史济怀, 数学分析教程, 第二册, 江苏教育出版社, 1999.

[2] H. Derksen, The fundamental theorem of algebra and linear algebra, *Amer. Math. Monthly* **110** (2003), 620-623.

[3] P. Erdős, A theorem of Sylvester and Schur, *J. London Math. Soc.* **9** (1934), 282-288.

[4] 冯克勤, 代数数论, 科学出版社, 2000.

[5] 龚升, 简明复分析, 北京大学出版社, 1996.

[6] 龚升, 话说微积分, 中国科技大学出版社,1998.

[7] 龚升, 线性代数五讲, 科学出版社, 2005.

[8] T. Hungerford, Algebra, 世界图书出版公司, 2006.

[9] C. R. Johnson, Positive definite matrices, *Amer. Math. Monthly* **77** (1970), 259-264.

[10] 李尚志, 线性代数, 高等教育出版社, 2007.

[11] 李炯生, 查建国, 王新茂, 线性代数, 中国科学技术大学出版社, 2005.

[12] 丘维声, 高等代数, 清华大学出版社, 2013.

[13] I. Schur, Einige Sätze über Primzahlen mit Anwendungen auf Irreduzibilitätsfragen I, *Sitzungsberichte Preuss. Akad. Wiss. Phys.-Math. Klasse* **14** (1929), 125-136. Also in *Gesammelte Abhandlungen*, Band III, 140-151.

[14] 同济大学数学系, 高等代数与解析几何, 高等教育出版社, 2015.

[15] 姚慕生, 吴泉水, 谢启鸿, 高等代数, 复旦大学出版社, 2014.

索引

R-模同态, 377
i次齐次成分, 57
k元函数, 181
k重线性函数, 181
n次本原单位根, 43
n阶行列式, 175
n维(列)向量空间, 249, 295
n元单项式, 54
n元多项式, 55
n元多项式环, 55
n元排列, 172
(在$\mathbb{Z}[x]$中)不可约, 48

Abel群, 91

Bezout等式, 17
Binet-Cauchy公式, 231

Cardan公式, 76

Eisenstein判别法, 51
Euler函数, 235

Ferrari解法, 76
Frobenius秩不等式, 288

Gauss消元法, 103
Gauss引理, 49

Hadamard乘积, 122
Hilbert基定理, 381

Kronecker符号, 125
Kronecker积, 144

Lagrange插值定理, 32, 239
Lagrange插值公式, 339
Lagrange多项式, 338
Laurent多项式, 134
Liouvill定理, 73

Moore-Penrose广义逆, 317

Newton公式, 65
Noether环, 97, 381
Noether模, 381

Penrose方程组, 314

Schur公式, 164
Steinitz替换定理, 265
Sylvester行列式, 83
Sylvester秩不等式, 288

Vandermonde行列式, 216
Vièta公式, 41

Zorn 引理, 335

按列分块, 141
按行分块, 141
伴随矩阵, 208

倍式, 11, 60
本原单位根, 40
本原多项式, 49
标准分解式, 26

不变量, 277
不可逆, 133

不可约多项式, 24, 60
不可约因式, 26
不可约元, 96
不整除, 60
插值问题, 339
差, 55

常数项, 3
常数项列向量, 107
初等变换, 104, 108
初等变换标准形, 153
初等对称多项式, 61
初等分块阵, 161
初等阵, 145
除环, 94

纯量阵, 107
次数, 3, 54, 55
带余除法, 8
代表元, 276
代数基本定理, 41
代数数, 28
代数余子式, 204, 225

单根, 30
单射, 351
单同态, 377
单位群, 94
单位阵, 107
单因式, 33
导出组, 302
导数, 34
等价, 109, 262, 332
等价关系, 276
等价类, 276
等幂和, 65
典范同构, 368

对称, 181
对称多项式, 61
对称群, 92
对称阵, 129
对合阵, 132
对换, 173
对角, 106
对角阵, 106
多项式环, 6
二阶行列式, 168
二阶行列式的对角线法则, 168
二项式系数, 126

反对称, 181
反对称阵, 129
方阵, 106
方阵的幂, 124
非零解, 116
非平凡子空间, 326
非奇异, 133
非齐次线性方程组, 116
非自由未知数, 115
分块对角阵, 138
分块矩阵的初等变换, 161
分块矩阵的列初等变换, 161
分块矩阵的行初等变换, 161
分块上三角阵, 138
分块下三角阵, 138
分圆多项式, 43
符号, 220

副对角, 106
复合, 352
复数域, 1, 94
复线性空间, 322
负矩阵, 121
负元素, 322
根, 30

公因式, 15
广义逆, 312

规范, 181
过渡矩阵, 347

含幺环, 93
含幺右R-模, 376
和, 55
恒等映射, 352

互素, 21, 23, 60
化三角, 179
环, 93
环同态, 94
换位元素, 126
幻方, 374
基, 296, 334
基本对称多项式, 61
基本矩阵, 125
基变换公式, 347
基础解系, 298
积, 55
积和式, 181

迹, 285
迹函数, 285
极大理想, 96
极大无关组, 266, 340
极大元, 335
极小多项式, 28
加边, 218
加法, 322
加减消元法, 103
简化阶梯形阵, 110

降阶法, 179
交换环, 93
交换群, 91
阶梯形阵, 110
结式, 83
解空间, 298
解向量, 250

矩阵, 105
矩阵的分块, 137
矩阵的共轭, 130
矩阵的加法, 120
矩阵的减法, 121
矩阵的转置, 128
矩阵多项式, 124
矩阵方程, 157

可比较的, 335
可交换, 126
可逆, 133
可逆线性映射, 354
可逆映射, 352
可逆元, 94
可约多项式, 24, 60

理想, 95
理想升链, 97
李括号, 126

链, 335
列初等变换, 108
列等价, 109
列满秩, 271
列向量, 106, 248
列秩, 271
零点, 30
零多项式, 3, 55
零解, 116
零矩阵, 106
零向量空间, 295
零因子, 93
零元素, 322

满射, 351
满同态, 377
满秩, 273
满秩分解, 282

幂等阵, 132
幂零阵, 137
幂零指数, 137

模m同余类环, 93
模p剩余类域, 94
模W同余, 367
内直和, 366
逆序对, 172, 220
逆序数, 172, 220
逆映射, 352
逆阵, 133

偶排列, 173
排列, 219
排列的符号, 173
判别式, 79

偏序关系, 335
偏序集, 335
平凡的组合, 252
平凡子空间, 326
奇排列, 173
奇异, 133
齐次多项式, 57
齐次线性方程组, 116

全系不变量, 277
群, 91
群同态, 92
三对角行列式, 212
三阶行列式, 169
三阶行列式的对角线法则, 169

商, 8
商环, 379
商集, 276
商空间, 368
上界, 335
上三角阵, 106

生成元集, 333

实数域, 1, 94
实线性空间, 322
首项, 3, 56
首项系数, 3, 56
首一多项式, 3
首元, 110
首元列, 110
数乘, 322
数域, 1

双射, 351
素元, 96

体, 94

通解, 114
同构, 354, 377
同构映射, 354
同解, 104
同解变换, 104
同类项, 54
外直和, 366

唯一分解定理, 60
维数, 296, 336
稳定的, 97
无限维线性空间, 336

系数, 3, 54, 55
系数矩阵, 107
下三角阵, 106
限制映射, 352
线性变换, 353
线性表示, 252, 332
线性方程组, 101
线性方程组的向量形式, 237
线性函数, 353
线性空间, 322
线性无关, 254, 331, 378

线性相关, 254, 331, 378
线性映射, 353
线性组合, 252, 332
相伴, 13, 60, 96
相等, 55
相邻对换, 173
相容的线性方程组, 102
像, 351
向量, 248, 322
向量空间, 295, 322
向量形式, 253
向量组, 249

行初等变换, 108
行等价, 109
行等价标准形阵, 110
行列式, 175
行满秩, 271
行向量, 106, 248
行秩, 271

循环方阵, 233
循环矩阵, 144
循环移位矩阵, 143
循环子模, 377
严格上三角阵, 143

一元多项式, 3
一元多项式环, 93
因式, 11, 60
因式定理, 30

由S生成的子空间, 333
酉方阵, 235
有理数域, 1, 94
有限生成, 377
有限维线性空间, 336
右R-模, 376
右零因子, 93
余式, 8
余式定理, 30
余子式, 204, 225
域, 1, 94

增广矩阵, 107

辗转相除法, 17
张量积, 144, 370
张量积的泛性质, 371
真子空间, 326
整除, 11, 60, 96
整环, 93
整数环, 6, 93
正交阵, 235

直和, 360, 364
直和补, 363
置换, 92
置换群, 92
秩, 267, 271, 340
中国剩余定理, 23
重根, 30
重数, 30
重因式, 33
主对角, 106
主角占优矩阵, 219
主理想, 95
主理想整环, 95
主子式, 235

爪形行列式, 196
准对角阵, 138
准上三角阵, 138
准下三角阵, 138
子R-模, 376
子环, 94
子空间, 325
子空间的和, 329
子模的升链条件, 381
子群, 92

子式, 223
子阵, 137
自然排列, 172, 220
自同态环, 94
自由模, 378
自由未知数, 115
自由线性空间, 345
字典序, 56
综合除法, 10
组合系数, 252
最大公因式, 15, 20, 60
最小公倍式, 20
左R-模, 376
左零因子, 93
坐标, 296, 337
坐标变换公式, 348