

线性代数 (A) 课程讲义

王新茂

xinmao@ustc.edu.cn

东区管理科研楼 12-05 室

2021 年 9 月 6 日

前言

本书是作者在讲授数学专业线性代数课程时所编写的课程讲义。教学内容是向数学专业本科学生介绍关于线性代数的基本概念、基本理论和基本方法。教学目标是使得学生能够快速入门线性代数，扎实且全面地掌握基础知识，并能够熟练地运用线性代数知识解决相关的数学问题。为此，本书按照从具体到抽象、从特殊到一般、从简单到复杂的原则安排教学内容。

线性代数起源于线性方程组的求解问题。本书从线性方程组的消元解法开始，引入矩阵和初等变换的概念。向量可以看作是数的推广，矩阵又可以看作是向量的推广，自然有矩阵的加法、数乘运算。矩阵也可以看作是线性映射的表示，从而引入矩阵的乘法运算。为了定义乘法的逆运算以及判断矩阵是否可逆，我们对矩阵作初等方阵乘积分解，从而得到相抵标准形。行列式是矩阵理论的重要分支，与各种矩阵问题有密切的联系。对于整数矩阵和一元多项式矩阵，相抵的概念可推广到模相抵的概念。为了解决矩阵方幂的计算问题，相似的概念被引入。矩阵的相抵、相似、相合关系都可以看作是某种形式的矩阵乘积分解。为了研究相似标准形问题，多种概念和方法被引入其中。本书还详细介绍了正交(酉)方阵、对称(Hermite)方阵、正定方阵的良好性质及其在二次型问题中的应用。

在学生掌握了矩阵的基础知识之后，本书开始介绍线性代数的空间理论，主要包括一般域上的线性空间、线性变换和内积理论，空间的维数不作限制。线性空间理论是对向量和矩阵知识的抽象和扩充。有限维线性空间上的线性代数问题，可以转化为矩阵问题加以研究。面对无限维线性空间上的线性代数问题时，有限维线性空间中成立的结论在无限维线性空间中有可能不成立，矩阵方法无法替代空间理论。线性代数的矩阵理论和空间理论是两套不同的数学语言，分别具有代数属性和几何属性，它们是线性代数的一体两面，各有所长。综合掌握这两套数学语言，有利于从多个角度深入思考具体问题。

多项式(尤其是一元多项式)是代数学的基本工具，在线性代数的研究中有着重要的作用。关于多项式以及群、环、域的一些代数学基础知识应该在先修课程中学习，本书不再详细介绍这部分代数学知识，而把与线性代数课程密切相关的多项式内容列为附录，供读者参考。

学完了主要知识之后，学生能够对于线性代数学科有一定的了解，能够解决一些简单的常见类型的线性代数问题。想要熟练运用所学知识，学生还需要通过大量习题的磨练，来巩固知识的掌握。本书安排了各种难度的习题供不同层次的学生练习。习题只是学习的辅助手段，习题的目的是检验学习的效果。建议学生首先深入理解基本概念，掌握定理的证明和例题的求解思路，然后再做习题。遇到有难度的习题，要主动思考，百思不得其解后，再向他人寻求帮助。

目录

第一章 线性方程组	7
§1.1 消元解法	7
§1.2 矩阵表示	11
第二章 矩阵运算	15
§2.1 基本概念	15
§2.2 分块矩阵	24
§2.3 初等方阵	28
§2.4 可逆矩阵	31
第三章 行列式	37
§3.1 行列式的定义	37
§3.2 Binet-Cauchy 公式	42
§3.3 Laplace 展开	49
§3.4 行列式与几何	54
第四章 矩阵的相抵	57
§4.1 矩阵的秩与相抵	57
§4.2 相抵标准形的应用	63
§4.3 Smith 标准形	71
第五章 矩阵的相似	77
§5.1 相似的概念	77
§5.2 相似三角化	84
§5.3 Jordan 标准形	89
§5.4 最小多项式	94
§5.5 特征方阵	98
第六章 正交方阵	103
§6.1 正交方阵	103
§6.2 正交相似	109
§6.3 正交相抵	112
§6.4 酉方阵	116

第七章 二次型	121
§7.1 二次型的化简	121
§7.2 正定方阵	126
§7.3 一些例子	130
第八章 线性空间	133
§8.1 基本概念	133
§8.2 线性相关	137
§8.3 向量组的秩	141
§8.4 基与坐标	144
§8.5 交空间与和空间	147
§8.6 直和与补空间	150
§8.7 直积与商空间	153
第九章 线性变换	155
§9.1 基本概念	155
§9.2 线性映射的运算	159
§9.3 对偶空间	162
§9.4 核空间与像空间	165
§9.5 不变子空间	168
§9.6 根子空间	170
§9.7 循环子空间	174
第十章 内积空间	179
§10.1 基本概念	179
§10.2 标准正交基	184
§10.3 正交变换	189
§10.4 伴随变换	192
§10.5 复内积空间	197
§10.6 内积的推广	201
附录 A 多项式	205
§A.1 一元多项式	205
§A.2 多元多项式	215
参考文献	223

第一章 线性方程组

线性方程组是最简单也是最重要的一类代数方程组，在科学研究和生产实践等领域中都有着广泛的应用。线性方程组求解问题具有非常悠久的历史，中国古代数学名著《九章算术》（成书于约公元前 150 年）中就记载着线性方程组的消元解法。线性方程组求解问题的研究促进了线性空间、线性变换以及矩阵理论的建立和发展，构成了线性代数这门数学分支学科的中心内容。

本章主要介绍一般的 n 元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \cdots \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (1.1)$$

的求解方法，其中 $a_{11}, a_{12}, \cdots, a_{mn}, b_1, b_2, \cdots, b_m$ 是已知的数， x_1, x_2, \cdots, x_n 是待求解的变量。特别，当 $b_1 = b_2 = \cdots = b_m = 0$ 时，线性方程组 (1.1) 称为齐次线性方程组。

关于线性方程组，有下列几个基本问题。在本章以及后续的章节中，将从不同的角度来研究这些问题。

1. 解的存在性问题。线性方程组 (1.1) 是否有解？
2. 解的唯一性问题。线性方程组 (1.1) 是否有唯一解？
3. 解集的结构问题。线性方程组 (1.1) 的解集有何性质？

在本书中，数指的是某个数域中的元素。数域 \mathbb{F} 是一个定义了加、减、乘、除运算，并且满足特定运算律的非空集合。详见定义 8.2。若存在素数 p 使得 $\underbrace{x + \cdots + x}_{p \text{ 个}} = 0, \forall x \in \mathbb{F}$ ，则 p 称为 \mathbb{F} 的特征，记作 $\text{Char } \mathbb{F} = p$ 。否则，规定 $\text{Char } \mathbb{F} = 0$ 。例如，无限域 $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ 的特征为 0， q 元有限域 \mathbb{F}_q 的特征为 p ，其中 q 是素数 p 的方幂。

§1.1 消元解法

求解代数方程组的基本思想是“消元法”：由原方程不断产生新方程，并且不断减少每个新方程所包含未知数的个数，直到求出一个未知数；然后如法炮制求出其它未知数。在求解过程中有可能得到“增根”，因此还需要把所求得的“可能解”代入原方程组进行检验，这样才能得到方程组的真正解。对于非线性方程组，很难做到不产生“增根”。然而对于线性方程组，可以进行如下的同解变形，以避免“验根”所带来的额外运算。

例 1.1. 求解线性方程组

$$\begin{cases} x_2 - x_3 = -1 & \textcircled{1} \\ x_1 - 2x_2 = -4 & \textcircled{2} \\ 3x_1 - 2x_2 + x_4 = -7 & \textcircled{3} \\ 3x_1 + x_2 + x_3 + 3x_4 = 0 & \textcircled{4} \end{cases}$$

解答. $\textcircled{4} - \textcircled{3}$, 得

$$3x_2 + x_3 + 2x_4 = 7 \quad \textcircled{5}$$

方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{4}$ 的解一定满足方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{5}$. 反之, $\textcircled{4} = \textcircled{3} + \textcircled{5}$, 方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{5}$ 的解一定满足方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{4}$. 因此, 方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{4}$ 与方程组 $\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{5}$ 的解集相同.

同理, $\textcircled{3} - 3 \times \textcircled{2}$, 得同解方程组 $\textcircled{1}, \textcircled{2}, \textcircled{6}, \textcircled{5}$, 其中

$$4x_2 + x_4 = 5 \quad \textcircled{6}$$

$\textcircled{5} - 3 \times \textcircled{1}$, 得同解方程组 $\textcircled{1}, \textcircled{2}, \textcircled{6}, \textcircled{7}$, 其中

$$4x_3 + 2x_4 = 10 \quad \textcircled{7}$$

$\textcircled{6} - 4 \times \textcircled{1}$, 得同解方程组 $\textcircled{1}, \textcircled{2}, \textcircled{8}, \textcircled{7}$, 其中

$$4x_3 + x_4 = 9 \quad \textcircled{8}$$

$\textcircled{7} - \textcircled{8}$, 得同解方程组 $\textcircled{1}, \textcircled{2}, \textcircled{8}, \textcircled{9}$, 其中

$$x_4 = 1 \quad \textcircled{9}$$

把 x_4 代入 $\textcircled{8}$, 得到 $x_3 = 2$. 把 x_3 代入 $\textcircled{1}$, 得到 $x_2 = 1$. 把 x_2 代入 $\textcircled{2}$, 得到 $x_1 = -2$. □



思考: 把已解出的未知数代入某个方程求得其它未知数的过程是否有可能产生“增根”?

定理 1.1. 对于线性方程组的下列操作不改变线性方程组的解.

1. 交换某两个方程在线性方程组中的位置, 其它方程不变.
2. 把某个方程替换成它的非零常数倍, 其它方程不变.
3. 把某个方程替换成它与另一方程的常数倍之和, 其它方程不变.

其中方程 $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ 的常数 λ 倍定义为

$$(\lambda a_1)x_1 + (\lambda a_2)x_2 + \cdots + (\lambda a_n)x_n = (\lambda b).$$

两个方程 $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ 与 $a'_1x_1 + a'_2x_2 + \cdots + a'_nx_n = b'$ 之和定义为

$$(a_1 + a'_1)x_1 + (a_2 + a'_2)x_2 + \cdots + (a_n + a'_n)x_n = (b + b').$$

定义 1.1. 以上三类操作, 称为对于线性方程组的初等变换.

通过第 (3) 类初等变换, 我们可以消去两个方程所包含的公共未知数. 例如, 假设第一个方程含有 x_1 , 即 $a_{11} \neq 0$, 则第一个方程的 $-\frac{a_{i1}}{a_{11}}$ 倍与第 i 个方程之和不含 x_1 , 从而可以得到 $m-1$ 个不含 x_1 的方程; 待由这 $m-1$ 个方程求得 x_2, \dots, x_n 之后, 代回第一个方程, 解得 x_1 . 当 $a_{11} = 0$ 时, 某个方程含有 x_1 , 第 (1) 类初等变换可以把方程组化为 $a_{11} \neq 0$ 情形. 第 (2) 类初等变换则可以把 $a_{11} \neq 0$ 情形化为 $a_{11} = 1$ 情形.

例 1.1 中的同解变形消元过程可以表示为如下初等变换.

$$\begin{array}{c}
 \begin{array}{c} \textcircled{1} \\ \textcircled{2} \\ \textcircled{3} \\ \textcircled{4} \end{array} \xrightarrow{\text{换行}} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{3} \\ \textcircled{4} \end{array} \xrightarrow{\text{消去 } x_1} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{3} \\ \textcircled{5} = \textcircled{4} - \textcircled{3} \end{array} \xrightarrow{\text{消去 } x_1} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{6} = \textcircled{3} - 3 \times \textcircled{2} \\ \textcircled{5} \end{array} \\
 \\
 \xrightarrow{\text{消去 } x_2} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{6} \\ \textcircled{7} = \textcircled{5} - 3 \times \textcircled{1} \end{array} \xrightarrow{\text{消去 } x_2} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{8} = \textcircled{6} - 4 \times \textcircled{1} \\ \textcircled{7} \end{array} \xrightarrow{\text{消去 } x_3} \begin{array}{c} \textcircled{2} \\ \textcircled{1} \\ \textcircled{8} \\ \textcircled{9} = \textcircled{7} - \textcircled{8} \end{array}
 \end{array}$$

定理 1.2. 线性方程组 (1.1) 可以通过一系列初等变换, 化为如下阶梯形的线性方程组.

$$\left\{ \begin{array}{l} a'_{1p_1} x_{p_1} + \dots + a'_{1n} x_n = b'_1 \\ a'_{2p_2} x_{p_2} + \dots + a'_{2n} x_n = b'_2 \\ \dots \dots \dots \\ a'_{rp_r} x_{p_r} + \dots + a'_{rn} x_n = b'_r \\ 0 = b'_{r+1} \\ \vdots \\ 0 = b'_m \end{array} \right. \quad (1.2)$$

其中 $0 \leq r \leq \min(m, n)$, $1 \leq p_1 < p_2 < \dots < p_r \leq n$ 并且 $a'_{1p_1} a'_{2p_2} \dots a'_{rp_r} \neq 0$.

证明. 对 m 使用数学归纳法. 当 $m = 1$ 时, 结论显然成立. 当 $m \geq 2$ 时, 设 p 是最小的正整数使得存在 $a_{ip} \neq 0$. 不妨设 $i = 1$, 否则交换第 i 个方程与第 1 个方程的位置. 通过第 (3) 类初等变换, 利用第 1 个方程, 可以把第 $2, \dots, m$ 个方程中 x_1, \dots, x_p 的系数都变成 0. 对这 $m-1$ 个方程应用归纳假设, 可以通过一系列初等变换, 化为阶梯形的线性方程组. 添上第 1 个方程, 仍然是阶梯形的线性方程组. \square

定理 1.3.

1. 线性方程组 (1.1) 有解的充分必要条件是 $b'_{r+1} = \dots = b'_m = 0$.
2. 线性方程组 (1.1) 有唯一解的充分必要条件是 $r = n$ 且 $b'_{r+1} = \dots = b'_m = 0$.

证明. 把线性方程组 (1.1) 变成 (1.2) 的初等变换操作不改变方程组的解.

1. 若 (1.2) 有解, 则 $b'_{r+1} = \cdots = b'_m = 0$. 若 $b'_{r+1} = \cdots = b'_m = 0$, 则可以解出 $x_{p_1}, x_{p_2}, \cdots, x_{p_r}$ 为其它变元的一次函数.
2. 设 $b'_{r+1} = \cdots = b'_m = 0$. 当 $r = n$ 时, 由 (1.2) 的前 r 个方程, 可依次解出 $x_n, x_{n-1}, \cdots, x_1$, 解是唯一的. 当 $r < n$ 时, 设 $i \notin \{p_1, p_2, \cdots, p_r\}$, 则 x_i 可取任意值, (1.2) 的解不唯一. \square

习题

1. 证明定理 1.1.
2. 求解下列线性方程组.

$$(1) \begin{cases} 2x_1 + x_2 - 2x_4 = 1 \\ x_1 + x_2 - x_3 = 1 \\ 2x_2 - x_3 - x_4 = 1 \\ x_1 + 2x_2 + 2x_3 - 2x_4 = 1 \end{cases}$$

$$(3) \begin{cases} x_1 + x_2 - x_3 + 2x_5 = 1 \\ x_1 - x_3 - x_4 - x_5 = 1 \\ 2x_1 - 2x_2 + x_3 + x_4 + 2x_5 = 1 \\ 2x_1 + x_2 + x_3 - x_5 = 1 \end{cases}$$

$$(5) \begin{cases} x_1 + x_2 - x_3 + 2x_4 + x_5 = 1 \\ x_1 + 3x_2 + 4x_4 + 5x_5 = 1 \\ 2x_1 + x_2 - 4x_3 + 5x_4 = 1 \\ 2x_2 + x_3 + 2x_4 + 4x_5 = 1 \\ 2x_1 + 3x_2 + 3x_3 - x_4 + 4x_5 = 1 \end{cases}$$

$$(2) \begin{cases} x_1 - 4x_3 + 2x_4 = 0 \\ x_1 + 2x_2 - 2x_3 = 0 \\ x_1 + x_2 - x_3 + x_4 = 0 \\ x_1 + 3x_2 - 3x_3 - x_4 = 0 \end{cases}$$

$$(4) \begin{cases} x_1 + x_2 - 2x_3 + 2x_5 = 1 \\ x_2 - 2x_3 + x_4 + 2x_5 = 1 \\ x_1 - x_3 + x_4 + x_5 = 1 \\ x_1 + x_2 - x_3 - 2x_4 + x_5 = 1 \end{cases}$$

$$(6) \begin{cases} x_1 + x_2 + x_4 = -2 \\ 2x_1 + 2x_3 - x_4 = 1 \\ 2x_1 + x_2 - x_3 - x_4 = 0 \\ x_1 - 2x_2 - 2x_3 - 2x_4 = 4 \\ 2x_2 - 2x_3 + x_4 = -3 \end{cases}$$

3. 证明: 当 $m < n$ 且 $b_1 = b_2 = \cdots = b_m = 0$ 时, 齐次线性方程组 (1.1) 必有非零解.

§1.2 矩阵表示

在对线性方程组施行初等变换的时候，我们注意到每个方程的左端都是数与字母的乘积之和的形式，只需要对未知数的系数进行操作；并且方程的个数、未知数的个数、未知数的顺序均保持不变。因此，我们可以隐去未知数和运算符号，用一个 m 行 $n+1$ 列的矩形表格

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \quad (1.3)$$

来表示线性方程组 (1.1)，把对于线性方程组的操作转化为对于矩形表格的操作。《九章算术·方程》中就记载着这样的解法：用算筹把线性方程组的系数和常数项排成长方形阵，然后用加减消元法求解。秦九韶^[1]的《数书九章》（成书于 1247 年）中则用互乘相消法和代入法求解线性方程组。

例 1.2. ^[2]今有上禾三秉、中禾二秉、下禾一秉，实三十九斗；上禾二秉、中禾三秉、下禾一秉，实三十四斗；上禾一秉、中禾二秉、下禾三秉，实二十六斗。问上、中、下禾实一秉各几何？答曰：上禾一秉，九斗、四分斗之一；中禾一秉，四斗、四分斗之一；下禾一秉，二斗、四分斗之三。

方程术曰：置上禾三秉，中禾二秉，下禾一秉，实三十九斗，于右方。中、左行列如右方。以右行上禾遍乘中行而以直除。又乘其次，亦以直除。然以中行中禾不尽者遍乘左行而以直除。左方下禾不尽者，上为法，下为实。实即下禾之实。求中禾，以法乘中行下实，而除下禾之实。余如上禾乘数而一，即中禾之实。求上禾亦以法乘右行下实，而除下禾、中禾之实。余如上禾乘数而一，即上禾之实。实皆如法，各得一斗。

下面，我们来学习古人的“方程术”。用现代数学的语言来说，原问题等价于求解线性方程组

$$\begin{cases} 3x + 2y + z = 39 \\ 2x + 3y + z = 34 \\ x + 2y + 3z = 26 \end{cases}$$

首先，把线性方程组的系数按照古人的书写顺序（从上到下，从右到左，行是竖直方向）排成数表。数表逆时针旋转 90° ，即为线性方程组的增广矩阵。

	左行	中行	右行
上禾	1	2	3
中禾	2	3	2
下禾	3	1	1
实	26	34	39

然后，用（右行，上禾）位置的数 3 遍乘中行的每个元素，再利用右行对中行消元，得

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 6 & 3 \\ 2 & 9 & 2 \\ 3 & 3 & 1 \\ 26 & 102 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{pmatrix}.$$

^[1]秦九韶，约 1202–1261，南宋著名数学家，四川安岳人。

^[2]《九章算术》第 8 章《方程》第 1 题。

再对左行作同样的处理, 得

$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 0 & 3 \\ 6 & 5 & 2 \\ 9 & 1 & 1 \\ 78 & 24 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{pmatrix}.$$

然后, 用 (中行, 中禾) 位置的数 5 遍乘左行的每个元素, 再利用中行对左行消元, 得

$$\begin{pmatrix} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 3 \\ 20 & 5 & 2 \\ 40 & 1 & 1 \\ 195 & 24 & 39 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{pmatrix}.$$

(左行, 下禾) 位置的数 36 称为法, (左行, 实) 位置的数 99 称为下禾之实. 即 36 乘下禾实 99 斗.

下面计算中禾. 已知 5 乘中禾、1 乘下禾实 24 斗. 用 36 乘 (中行, 实) 位置处的数 24, 减去下禾之实, 再除以中禾乘数 5, 所得数 $\frac{36 \times 24 - 99}{5} = 153$ 为中禾之实. 即 36 乘中禾实 153 斗.

对于上禾, 已知 3 乘上禾、2 乘中禾、1 乘下禾实 39 斗. 用 36 乘 (右行, 实) 位置处的数 39, 减去下禾之实与 2 倍的中禾之实, 再除以上禾乘数 3, 所得数 $\frac{36 \times 39 - 99 - 2 \times 153}{3} = 333$ 为上禾之实. 即 36 乘上禾实 333 斗.

因此, 1 乘上、中、下禾之实的斗数分别是 $\frac{333}{36} = 9\frac{1}{4}$, $\frac{153}{36} = 4\frac{1}{4}$, $\frac{99}{36} = 2\frac{3}{4}$. \square

古人的做法几乎就是 Gauss 消元法, 先把系数矩阵初等变换成上三角形, 再依次回代求解未知数. 在求解过程中, 古人始终把运算限制在整数范围内, 只是在最后才以分数的形式给出答案. 究其原因, 古人使用算筹来表示数字, 追求计算过程的简洁和精确, 与其使用分数, 不如先通分, 而后始终在整数范围内运算. 事实上, 《九章算术》的第一章《方田》中就已经给出了分数的加、减、乘、除、通分、约分运算. 分母称为“法”, 分子称为“实”. 例 1.2 中的“实”, 有时指的是“农作物的收成”, 有时指的是“分数的分子”, 读者需仔细区分.

定义 1.2. (1.3) 式中的矩形表格 M 称为线性方程组 (1.1) 的增广矩阵. 删去 M 的最后一列得到的 m 行 n 列的矩形表格称为线性方程组 (1.1) 的系数矩阵. M 的第 i 行从左至右排成的数组, 称为 M 的第 i 个行向量. M 的第 j 列从上至下排成的数组, 称为 M 的第 j 个列向量.

 矩阵的元素并不仅限于数, 也可以是含有未知数和字母的代数表达式.

定义 1.3. 由数域 \mathbb{F} 中的 n 个数构成的数组 $x = (x_1, x_2, \dots, x_n)$ 称为数域 \mathbb{F} 上的 n 维数组向量, 简称 n 维向量. 数域 \mathbb{F} 上所有 n 维数组向量的集合记作 \mathbb{F}^n . 元素都是 0 的向量称为零向量, 记作 0 . 第 i 个元素是 1、其它元素都是 0 的向量称为第 i 个标准单位向量, 记作 e_i .

设 $y = (y_1, y_2, \dots, y_n) \in \mathbb{F}^n$, $\lambda \in \mathbb{F}$. 两个向量的加法运算定义为

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

数与向量的数乘运算定义为

$$\lambda x = (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$

显然, 向量的加法和数乘运算是数的加法和乘法运算的推广. 类似地, 可定义两个向量的减法运算

$$x - y = (x_1 - y_1, x_2 - y_2, \cdots, x_n - y_n),$$

和线性组合运算

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \cdots, \lambda x_n + \mu y_n),$$

并推广为多个向量 $\alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{F}^n$ 的线性组合运算

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_k \alpha_k$$

其中 $\mu, \lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{F}$.

以下是对于线性方程组 (1.1) 的两种常见看法.

1. 定义两个 n 维向量 $x = (x_1, x_2, \cdots, x_n)$ 和 $y = (y_1, y_2, \cdots, y_n)$ 的数量积运算为

$$x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

记 n 维向量 $\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in})$. 线性方程组 (1.1) 可以看作是:

• 求向量 $x = (x_1, x_2, \cdots, x_n)$ 使得 $\alpha_i \cdot x = b_i, \forall i = 1, 2, \cdots, m$.

2. 记 m 维向量 $\beta_j = (a_{1j}, a_{2j}, \cdots, a_{mj})$, $b = (b_1, b_2, \cdots, b_m)$. 线性方程组 (1.1) 还可以看作是:

• 求线性组合的系数 x_1, x_2, \cdots, x_n 使得 $x_1 \beta_1 + x_2 \beta_2 + \cdots + x_n \beta_n = b$.

定理 1.4. 数组向量的加法和数乘运算具有下列性质.

$$(A1) \quad (x + y) + z = x + (y + z)$$

$$(A2) \quad x + y = y + x$$

$$(A3) \quad x + \mathbf{0} = \mathbf{0} + x = x$$

$$(A4) \quad x + (-x) = (-x) + x = \mathbf{0}$$

$$(M1) \quad (\lambda \mu)x = \lambda(\mu x)$$

$$(M2) \quad 1x = x$$

$$(D1) \quad \lambda(x + y) = \lambda x + \lambda y$$

$$(D2) \quad (\lambda + \mu)x = \lambda x + \mu x$$

其中 $x, y, z \in \mathbb{F}^n$, $\lambda, \mu \in \mathbb{F}$.

线性方程组 (1.1) 与矩阵 M 之间存在一一对应关系. 方程 $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ 与行向量 $(a_{i1}, a_{i2}, \cdots, a_{in}, b_i)$ 一一对应, 方程的常数倍对应于行向量的数乘运算, 两个方程之和对应于两个行向量的加法运算. 对于线性方程组 (1.1) 作初等变换的消元过程, 可以表示为对矩阵 M 的行向量作初等变换, 把 M 化为阶梯形的矩阵

$$M' = \begin{pmatrix} a'_{1p_1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a'_{1n} & b'_1 \\ & & a'_{2p_2} & \cdots & \cdots & \cdots & \cdots & a'_{2n} & b'_2 \\ & & & & \cdots & \cdots & \cdots & \vdots & \vdots \\ & & & & & a'_{rp_r} & \cdots & a'_{rn} & b'_r \\ & & & & & & & & b'_{r+1} \\ & & & & & & & & \vdots \\ & & & & & & & & b'_m \end{pmatrix}$$

其中 M' 中空白处的元素都是 0.

例 1.3. 用矩阵方法求解例 1.1.

解答.

$$\begin{aligned}
 & \begin{pmatrix} 0 & 1 & -1 & 0 & -1 \\ 1 & -2 & 0 & 0 & -4 \\ 3 & -2 & 0 & 1 & -7 \\ 3 & 1 & 1 & 3 & 0 \end{pmatrix} \xrightarrow{\text{交换①,②}} \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 3 & -2 & 0 & 1 & -7 \\ 3 & 1 & 1 & 3 & 0 \end{pmatrix} \xrightarrow{\substack{\text{③}-=3\text{①} \\ \text{④}-=3\text{①}}} \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 4 & 0 & 1 & 5 \\ 0 & 7 & 1 & 3 & 12 \end{pmatrix} \xrightarrow{\substack{\text{③}-=4\text{②} \\ \text{④}-=7\text{②}}} \\
 & \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & 1 & 9 \\ 0 & 0 & 8 & 3 & 19 \end{pmatrix} \xrightarrow{\text{④}-=2\text{③}} \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & 1 & 9 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{③}-=④} \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{③}=\frac{1}{4}\text{③}} \\
 & \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{②}+=\text{③}} \begin{pmatrix} 1 & -2 & 0 & 0 & -4 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{①}+=2\text{②}} \begin{pmatrix} 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.
 \end{aligned}$$

上式最后一个矩阵的最后一列即为方程组的解, $(x_1, x_2, x_3, x_4) = (-2, 1, 2, 1)$. □

习题

1. 证明定理 1.4.
2. 用矩阵方法求解 §1.1 习题 2.
3. 设 $A(1,0)$, $B(0,2)$, $C(3,3)$. 求 $\triangle ABC$ 的外接圆方程.
4. 设 $A(1,0,0)$, $B(0,2,0)$, $C(0,0,3)$. 求 $\triangle ABC$ 的外接圆方程.
5. 求满足 $f(1) = f(2) = f(3) = 1$, $f(-1) = f(-2) = f(-3) = -1$ 的所有多项式 $f(x)$.
6. 求满足 $f(1) = f'(1) = f''(1) = 1$, $f(-1) = f'(-1) = f''(-1) = -1$ 的所有多项式 $f(x)$.
7. 当 λ 取何值时, 如下线性方程组有解? 有唯一解? 并求其所有解.

$$\begin{cases} \lambda x_1 + x_2 + x_4 = 1 \\ x_1 + \lambda x_2 + x_3 = 1 \\ x_2 + \lambda x_3 + x_4 = 1 \\ x_1 + x_3 + \lambda x_4 = 1 \end{cases}$$

8. 当 a, b 取何值时, 如下线性方程组有解? 有唯一解? 并求其所有解.

$$\begin{cases} ax_1 + x_2 + x_3 + x_4 = 0 \\ 2x_1 - x_2 + 2x_3 + 3x_4 = b \\ 2x_1 + x_2 - x_3 + 2x_4 = 1 \\ x_2 + x_3 + 2x_4 = 1 \\ x_2 + 2x_3 + 3x_4 = 2 \end{cases}$$

第二章 矩阵运算

与线性方程组一样, 矩阵的概念也具有悠久的历史, 如河图、洛书、幻方、拉丁方等. 矩阵作为现代数学的研究对象, 则是在行列式的研究发展起来之后. 许多与矩阵有关的性质已经在行列式的研究中被发现, 此时引入矩阵的概念变得水到渠成. 1801 年, Gauss^[1]把线性变换的所有系数当作一个整体使用, 并用一个字母表示. 1844 年, Eisenstein^[2]讨论了线性变换(矩阵)的乘积, 并强调了乘积的不可交换性. 1850 年, Sylvester^[3]首先使用了“矩阵”一词. 从 1858 年开始, Cayley^[4]发表了一系列关于矩阵的论文, 系统地阐述了矩阵的运算及其性质. 因此, Cayley 被公认是矩阵论的奠基人.

§2.1 基本概念

一个 m 行 n 列的表格

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

称为 $m \times n$ 矩阵, 简记 $A = (a_{ij})_{m \times n}$, 其中 a_{ij} 称为 A 的第 i 行 j 列元素, 或者 A 的 (i, j) 位置元素. 有序整数对 (m, n) 称为 A 的维数或阶数. 当 $m = n$ 时, A 称为 n 阶方阵. 当 $mn = 0$ 时, A 称为空矩阵.

设 A 的元素都属于某个集合 S . A 称为 S 上的矩阵. 例如, 当 $S = \mathbb{C}, \mathbb{R}, \mathbb{Q}, \mathbb{Z}$ 时, A 称为复数矩阵、实数矩阵、有理数矩阵、整数矩阵. S 上的所有 $m \times n$ 矩阵的集合记作 $S^{m \times n}$. 矩阵 $A \in S^{m \times n}$ 还可以看作是从集合 $\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ 到 S 的映射: $(i, j) \mapsto a_{ij}$.

a_{ii} 称为 A 的对角元素. A 的所有对角元素之和称为 A 的迹, 记作 $\text{tr}(A)$. 当 $i < j$ 时, a_{ij} 称为 A 的上三角元素. 当 $i > j$ 时, a_{ij} 称为 A 的下三角元素. 下三角元素都是 0 的矩阵称为上三角矩阵. 上三角元素都是 0 的矩阵称为下三角矩阵. 上三角矩阵和下三角矩阵统称为三角矩阵.

元素都是 0 的矩阵称为零矩阵, 记作 O . 通常也把空矩阵视为零矩阵. 非对角元素都是 0 的矩阵称为对角矩阵. 以 a_1, a_2, \dots, a_n 为对角元素的 n 阶对角方阵常记作 $\text{diag}(a_1, a_2, \dots, a_n)$. $\text{diag}(a, a, \dots, a)$ 称为纯量方阵. $\text{diag}(1, 1, \dots, 1)$ 称为单位方阵, 记作 I . (i, j) 位置元素为 1, 其它元素为 0 的矩阵称为基础矩阵, 记作 E_{ij} .

^[1]Johann Carl Friedrich Gauss, 1777–1855, 德国数学家, 近代数学奠基人之一.

^[2]Ferdinand Gotthold Max Eisenstein, 1823–1852, 德国数学家.

^[3]James Joseph Sylvester, 1814–1897, 英国数学家.

^[4]Arthur Cayley, 1821–1895, 英国数学家.

满足每个 $b_{ij} = a_{ji}$ 的矩阵 $B = (b_{ij})_{n \times m}$ 称为 $A = (a_{ij})_{m \times n}$ 的转置, 记作 $B = A^T$. 满足 $A^T = A$ 的矩阵 A 称为对称方阵. 满足 $A^T = -A$ 的矩阵 A 称为反对称方阵. 当 A 是复数矩阵时, 矩阵 $B = (\overline{a_{ij}})_{m \times n}$ 称为 A 的共轭, 记作 $B = \overline{A}$. 矩阵 $\overline{A^T}$ 称为 A 的共轭转置, 常记作 A^H . 若 $A^H = A$, 则 A 称为 Hermite^[5] 方阵. 若 $A^H = -A$, 则 A 称为反 Hermite 方阵.

矩阵 A 可以看作是 mn 个元素排成的 m 行 n 列的表格, 也可以看作是 m 个行向量排成一列, 又或者是 n 个列向量排成一行.

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n)$$

其中行向量 $\alpha_i = (a_{i1}, a_{i2}, \dots, a_{in})$, 列向量 $\beta_j = (a_{1j}, a_{2j}, \dots, a_{mj})^T$. 矩阵概念可以看作是向量概念的推广, 向量可以看作是矩阵的特殊形式. 行向量是 $1 \times n$ 的矩阵, 列向量是 $m \times 1$ 的矩阵. 向量的运算可以推广到矩阵的运算.

定义 2.1. 设 \mathbb{F} 是数域, $A = (a_{ij}) \in \mathbb{F}^{m \times n}$, $B = (b_{ij}) \in \mathbb{F}^{m \times n}$, $\lambda \in \mathbb{F}$, 定义

1. A 与 B 的加法运算 $A + B = (a_{ij} + b_{ij}) \in \mathbb{F}^{m \times n}$.
2. A 与 B 的减法运算 $A - B = (a_{ij} - b_{ij}) \in \mathbb{F}^{m \times n}$.
3. λ 与 A 的数乘运算 $\lambda A = (\lambda a_{ij}) \in \mathbb{F}^{m \times n}$.
4. $\lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k$ 称为矩阵的线性组合运算, 其中每个 $A_i \in \mathbb{F}^{m \times n}$, $\lambda_i \in \mathbb{F}$.

定理 2.1. 矩阵的加法和数乘运算具有与定理 1.4 类似的性质, 并且对于任意 $A, B \in \mathbb{F}^{m \times n}$, $\lambda \in \mathbb{F}$,

$$(A + B)^T = A^T + B^T, \quad (\lambda A)^T = \lambda A^T.$$

例 2.1. 设映射 $f = (f_1, f_2, \dots, f_m) : \mathbb{F}^n \rightarrow \mathbb{F}^m$, 其中每个 f_i 是 n 元齐次线性函数,

$$f_i(x_1, x_2, \dots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n, \quad i = 1, 2, \dots, m.$$

f 称为线性映射. $A = (a_{ij})_{m \times n}$ 与 f 一一对应. A 称为 f 的矩阵表示.

例 2.2. 设 A 和 B 分别是 $\mathbb{F}^n \rightarrow \mathbb{F}^m$ 的线性映射 $f = (f_1, f_2, \dots, f_m)$ 和 $g = (g_1, g_2, \dots, g_m)$ 的矩阵表示, $\lambda \in \mathbb{F}$, 则 $f + g = (f_1 + g_1, f_2 + g_2, \dots, f_m + g_m)$ 和 $\lambda f = (\lambda f_1, \lambda f_2, \dots, \lambda f_m)$ 也是 $\mathbb{F}^n \rightarrow \mathbb{F}^m$ 的线性映射, 并且 $A + B$ 是 $f + g$ 的矩阵表示, λA 是 λf 的矩阵表示.

例 2.3. 设 $f = (f_1, f_2, \dots, f_m)$ 是 $\mathbb{F}^n \rightarrow \mathbb{F}^m$ 的线性映射,

$$f_i(x_1, x_2, \dots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n, \quad i = 1, 2, \dots, m,$$

$g = (g_1, g_2, \dots, g_p)$ 是 $\mathbb{F}^p \rightarrow \mathbb{F}^n$ 的线性映射,

$$g_i(x_1, x_2, \dots, x_p) = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{ip}x_p, \quad i = 1, 2, \dots, p,$$

^[5]Charles Hermite, 1822–1901, 法国数学家.

则 $h = (h_1, h_2, \dots, h_m) = f \circ g$ 是 $\mathbb{F}^p \rightarrow \mathbb{F}^m$ 的线性映射, 其中

$$h_i(x_1, x_2, \dots, x_p) = \sum_{k=1}^n a_{ik} g_k(x_1, x_2, \dots, x_p) = \sum_{k=1}^n a_{ik} \sum_{j=1}^p b_{kj} x_j = \sum_{j=1}^p \left(\sum_{k=1}^n a_{ik} b_{kj} \right) x_j.$$

设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$, $C = (c_{ij})_{m \times p}$ 分别是 f, g, h 的矩阵表示, 则有

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}, \quad \forall i, j. \quad (2.1)$$

定义 2.2. 上式 (2.1) 定义的矩阵 C 称为 A 与 B 的乘积, 记作 $C = AB$.

设 A 的行向量分别为 $\alpha_1, \alpha_2, \dots, \alpha_m$, B 的列向量分别为 $\beta_1, \beta_2, \dots, \beta_p$, 则 $C = AB$ 也可以看作是由 $c_{ij} = \alpha_i \cdot \beta_j$ (n 维向量的数量积运算) 排成的矩阵.

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \quad B = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_p) \quad \Rightarrow \quad AB = \begin{pmatrix} \alpha_1 \cdot \beta_1 & \alpha_1 \cdot \beta_2 & \cdots & \alpha_1 \cdot \beta_p \\ \alpha_2 \cdot \beta_1 & \alpha_2 \cdot \beta_2 & \cdots & \alpha_2 \cdot \beta_p \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_m \cdot \beta_1 & \alpha_m \cdot \beta_2 & \cdots & \alpha_m \cdot \beta_p \end{pmatrix}.$$

例 2.4. 设 $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, 则有 $AB = B$, $BA = O$.

并非任意两个矩阵 A, B 都可以作乘法运算 AB . 只有当 A 的列数与 B 的行数相等时, 运算 AB 才有意义. AB 有意义也不能保证 BA 有意义. 矩阵的乘法运算不满足交换律. 为了避免歧义, 有时称 AB 为“ A 右乘上 B ”或“ B 左乘上 A ”.

例 2.5. 利用矩阵的乘法运算, 可以很简明地表示数学式.

- 线性方程组 (1.1) 可写成 $Ax = b$.
- 例 2.1 中的线性映射 f 可写成 $f(x) = Ax$.
- 例 2.2 表明 $Ax + Bx = (A + B)x$, $\lambda(Ax) = (\lambda A)x$.
- 例 2.3 表明 $A(Bx) = (AB)x$. 换言之, $f \circ g$ 的矩阵表示是 AB , 线性映射的复合与它们的矩阵表示的乘积一一对应.

例 2.6. 许多数学问题可以通过矩阵转化为线性代数问题. 在许多实际问题中, 矩阵也同样有着重要的作用.

- (几何变换) \mathbb{R}^2 中向量逆时针旋转 θ 的旋转变换可以表示为矩阵乘积的形式

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

\mathbb{R}^2 上的仿射变换 $(x, y) \mapsto (a_1x + b_1y + c_1, a_2x + b_2y + c_2)$ 可以表示为

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad \text{或} \quad \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

- (二次方程) \mathbb{R}^2 上二次曲线的一般方程

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0$$

可以表示为矩阵乘积形式

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

类似地, \mathbb{R}^3 上二次曲面的一般方程

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0$$

也可以表示为矩阵乘积形式

$$\begin{pmatrix} x & y & z & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{12} & a_{22} & a_{23} & b_2 \\ a_{13} & a_{23} & a_{33} & b_3 \\ b_1 & b_2 & b_3 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0.$$

- (数系的扩充) 设映射 $f: \mathbb{C} \rightarrow \mathbb{R}^{2 \times 2}$, $a + bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. 容易验证, f 是单射并且满足

$$f(z+w) = f(z) + f(w), \quad f(zw) = f(z)f(w), \quad \forall z, w \in \mathbb{C}.$$

因此, \mathbb{C} 可以看作是 $\mathbb{R}^{2 \times 2}$ 的子集, 复数的加、减、乘法运算对应矩阵的加、减、乘法运算.

类似地, 考虑四元数体^[6]

$$\mathbb{K} = \{a_0 + a_1i + a_2j + a_3k \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$$

其中 i, j, k 是 \mathbb{K} 的生成元, 满足

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

设映射

$$g: \mathbb{K} \rightarrow \mathbb{C}^{2 \times 2}, \quad a_0 + a_1i + a_2j + a_3k \mapsto \begin{pmatrix} a_0 + a_1i & a_2 + a_3i \\ -a_2 + a_3i & a_0 - a_1i \end{pmatrix}.$$

容易验证, g 是单射并且满足

$$g(a+b) = g(a) + g(b), \quad g(ab) = g(a)g(b), \quad \forall a, b \in \mathbb{K}. \quad (2.2)$$

因此, \mathbb{K} 可以看作是 $\mathbb{C}^{2 \times 2}$ 的子集, 四元数的加、减、乘法运算对应矩阵的加、减、乘法运算.

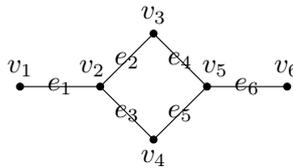
同理, 我们可以构造单射 $g: \mathbb{K} \rightarrow \mathbb{R}^{4 \times 4}$ 满足 (2.2), 把 \mathbb{K} 看作是 $\mathbb{R}^{4 \times 4}$ 的子集.

^[6]William Rowan Hamilton 于 1843 年发明了四元数的概念.

- (计算机图像) 在计算机系统中, 一幅分辨率为 $w \times h$ 的黑白图片通常可以用一个 $h \times w$ 的实数矩阵 $A = (a_{ij})$ 表示, 其中 w 和 h 分别是图片的宽度和高度, a_{ij} 是第 i 行 j 列的像素点的灰度. 一幅彩色图片则可以用三个矩阵 R, G, B 表示, 分别对应红、绿、蓝三色的分量. 图像处理的过程就是矩阵运算的过程.
- (投入产出模型) 假设共有 m 种产品和 n 种原料, 生产单位数量的产品 i 需要消耗原料 j 的数量为 a_{ij} , 则 $m \times n$ 矩阵 $A = (a_{ij})$ 称为**消耗系数矩阵**. 假设需要生产出数量分别为 x_1, x_2, \dots, x_m 的 m 种产品, 则需要投入的 n 种原料的数量 y_1, y_2, \dots, y_n 满足

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{即 } y = xA.$$

- (图的矩阵表示) 如下的图形 G 由 6 个点 v_1, v_2, \dots, v_6 和 6 条线段 e_1, e_2, \dots, e_6 构成.



通常记 $G = (V, E)$, $V = \{v_1, v_2, \dots, v_6\}$ 称为 G 的**顶点集**, $\{e_1, e_2, \dots, e_6\}$ 称为 G 的**边集**.

我们可以用一个 6 阶方阵 $A = (a_{ij})$ 来表示 G , 其中 $a_{ij} = \begin{cases} 1, & v_i \text{ 与 } v_j \text{ 有线段相连;} \\ 0, & v_i \text{ 与 } v_j \text{ 没有线段相连.} \end{cases}$

我们也可以用一个 6 阶方阵 $B = (b_{ij})$ 来表示 G , 其中 $b_{ij} = \begin{cases} 1, & v_j \text{ 是 } e_i \text{ 的端点;} \\ 0, & v_j \text{ 不是 } e_i \text{ 的端点.} \end{cases}$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

A 称为 G 的**邻接矩阵**, B 称为 G 的**关联矩阵**. 它们之间具有联系

$$B^T B = D + A, \quad D = \text{diag}(d_1, d_2, \dots, d_6)$$

其中每个 d_j 是 A 的第 j 行 (列) 元素之和, 也是 B 的第 j 列元素之和. D 称为 G 的**度矩阵**. 感兴趣的读者可参阅文献 [10] 组合矩阵论, 了解关于图和矩阵的更多知识.

- (有限射影平面) 设 $V = \{v_1, v_2, \dots, v_n\}$, V 中的元素称为**点**; $E = \{e_1, e_2, \dots, e_m\}$ 由 V 的某些子集构成, E 中的元素称为**线**. 当满足下列条件时, 集合对 $\pi = (V, E)$ 称为**有限射影平面**:

(i) 对于任意两线 $e, f \in E$, 存在唯一的点 $v \in V$, 使得 $v \in e$ 且 $v \in f$;

- (ii) 对于任意两点 $u, v \in V$, 存在唯一的线 $e \in E$, 使得 $u \in e$ 且 $v \in e$;
 (iii) 存在四点 $a, b, c, d \in V$, 使得 $\{a, b, c, d\}$ 与任意 $e \in E$ 至多有两个公共点.

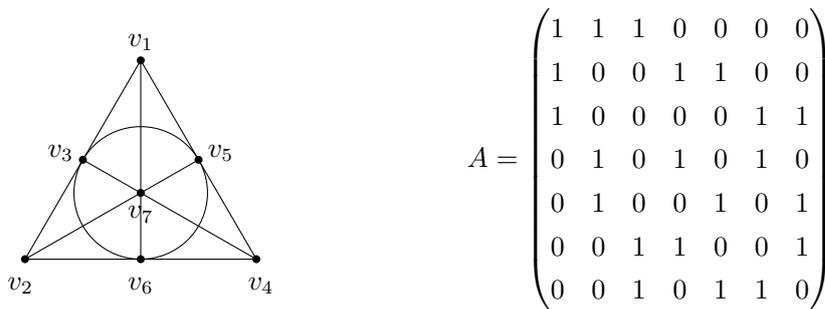
类似地, 可定义 π 的关联矩阵 $A = (a_{ij})_{m \times n}$, 其中 $a_{ij} = \begin{cases} 1, & v_j \in e_i \\ 0, & v_j \notin e_i \end{cases}$. 由条件 (i),(ii) 可得

$$AA^T = \begin{pmatrix} k_1 & 1 & \cdots & 1 \\ 1 & k_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & k_m \end{pmatrix}, \quad A^T A = \begin{pmatrix} d_1 & 1 & \cdots & 1 \\ 1 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & d_n \end{pmatrix}. \quad (2.3)$$

进一步由条件 (iii) 可得

$$k_1 = k_2 = \cdots = k_m = d_1 = d_2 = \cdots = d_n = \lambda, \quad m = n = \lambda^2 - \lambda + 1. \quad (2.4)$$

例如, 下图是最小的有限射影平面, 由 7 个点 v_1, v_2, \dots, v_7 和 7 条线 $\{v_1, v_2, v_3\}$, $\{v_1, v_4, v_5\}$, $\{v_1, v_6, v_7\}$, $\{v_2, v_4, v_6\}$, $\{v_2, v_5, v_7\}$, $\{v_3, v_4, v_7\}$, $\{v_3, v_5, v_6\}$ 构成.



定理 2.2. 矩阵的乘法运算具有下列性质.

1. 对于任意 $A \in \mathbb{F}^{m \times n}$ 和 $\beta_1, \beta_2, \dots, \beta_p \in \mathbb{F}^{n \times 1}$, 有

$$A \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \end{pmatrix} = \begin{pmatrix} A\beta_1 & A\beta_2 & \cdots & A\beta_p \end{pmatrix}.$$

2. 对于任意 $A \in \mathbb{F}^{m \times n}$ 和 $\lambda \in \mathbb{F}$, 有 $(\lambda I_m)A = \lambda A = A(\lambda I_n)$.

3. 对于任意 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$ 和 $C \in \mathbb{F}^{p \times q}$, 有 $(AB)C = A(BC)$.

4. 对于任意 $A, B \in \mathbb{F}^{m \times n}$ 和 $C, D \in \mathbb{F}^{n \times p}$, 有

$$(A + B)C = (AC) + (BC), \quad A(C + D) = (AC) + (AD).$$

5. 对于任意 $A \in \mathbb{F}^{m \times n}$ 和 $B \in \mathbb{F}^{n \times p}$, 有 $(AB)^T = B^T A^T$.

6. 对于任意 $A \in \mathbb{F}^{m \times n}$ 和 $B \in \mathbb{F}^{n \times m}$, 有 $\text{tr}(AB) = \text{tr}(BA)$.

由定理 2.2 可知, 矩阵的乘法运算满足结合律, 从而可以定义矩阵的方幂和多项式运算.

定义 2.3. 设 A 是 n 阶方阵, m 是正整数. m 个 A 的乘积称为 A 的 m 次方幂, 记作 A^m . 特别规定 $A^0 = I_n$ (包括 $A = O$ 情形). 对于 $f(x) = \sum_{k=0}^m c_k x^k \in \mathbb{F}[x]$, 定义

$$f(A) = \sum_{k=0}^m c_k A^k = c_0 I + c_1 A + \cdots + c_m A^m.$$

计算一般矩阵 A 的方幂 A^k 和多项式 $f(A)$ 的通常算法是: 首先把 A 分解为 $A = PBP^{-1}$ 的形式, 其中 B 是对角方阵或 Jordan 标准形, 然后计算 B^k 和 $f(B)$, 最后计算 $A^k = PB^kP^{-1}$ 和 $f(A) = Pf(B)P^{-1}$. 相关内容详见第五章. 对于某些特殊矩阵, 有可能存在更简便的方法.

例 2.7. 求线性递推数列 $x_{n+1} = ax_n + bx_{n-1}$ 的通项公式.

解答. $\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix} = \cdots = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} x_1 \\ x_0 \end{pmatrix} \Rightarrow x_n = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}.$ \square

例 2.8. 求连分数 $[a_0, a_1, \cdots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{\ddots + \frac{1}{a_{n-1} + \frac{1}{a_n}}}}$ 的展开分数.

解答. 设 $[a_k, a_{k+1}, \cdots, a_n] = \frac{p_k}{q_k}$, $0 \leq k \leq n$. 由 $\frac{p_{k-1}}{q_{k-1}} = a_{k-1} + \frac{q_k}{p_k} = \frac{a_{k-1}p_k + q_k}{p_k}$, 可设

$$\begin{pmatrix} p_{k-1} \\ q_{k-1} \end{pmatrix} = \begin{pmatrix} a_{k-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p_k \\ q_k \end{pmatrix}.$$

因此, $[a_0, a_1, \cdots, a_n] = \frac{p}{q}$, 其中 $\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$ \square

习题

1. 证明定理 2.1 和定理 2.2.

2. 设 $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{pmatrix}$. 计算 AB , BC , B^2 , ABC .

3. 设 $f(x) = a_0 + a_1x + \cdots + a_nx^n$. 求 $n+1$ 阶方阵 $B = (b_{ij})$ 和 n 阶方阵 $C = (c_{ij})$, 满足

$$f(x+y) = \sum_{i,j=1}^{n+1} b_{ij} x^{i-1} y^{j-1}, \quad \frac{f(x) - f(y)}{x - y} = \sum_{i,j=1}^n c_{ij} x^{i-1} y^{j-1}.$$

并说明 B 和 C 都是对称方阵.

4. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times n}$. 证明:

(1) 若 B 是对称方阵, 则 ABA^T 也是对称方阵.

(2) 若 B 是对角元素都是 0 的反对称方阵, 则 ABA^T 也是对角元素都是 0 的反对称方阵.

5. 对于 $H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 和 $H = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 分别求所有 2 阶实方阵 A 满足 $AHA^T = H$.

6. 设 $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. 分别举例 3 阶非对角的实方阵 X 满足如下条件.

(1) $X^2 = I$ (2) $X^3 = I$ (3) $X^2 = -A$ (4) $X^3 = -A$ (5) $X^2 = B$ (6) $X^3 = B$

7. 设 m 是正整数, 求 A^m , 其中 $A =$

(1) $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ (2) $\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$ (3) $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ (4) $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
 (5) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (6) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (7) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ (8) $\begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{pmatrix}$

8. 求与方阵 A 乘积可交换 (即满足 $AB = BA$) 的所有方阵 B , 其中 $A =$

(1) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ (2) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ (3) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ (4) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
 (5) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}$ (6) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ (7) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (8) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

9. (1) 求与任意 n 阶方阵都乘积可交换的方阵.

(2) 求与任意 n 阶对称方阵都乘积可交换的方阵.

(3) 求与任意 n 阶反对称方阵都乘积可交换的方阵.

10. 设 m, n 是正整数, $A, B \in \mathbb{F}^{n \times n}$, $f \in \mathbb{F}[x]$. 证明:

(1) $(A+B)^2 + (A-B)^2 = 2A^2 + 2B^2$.

(2) 若 $AB = BA$, 则 $(A+B)^m = \sum_{k=0}^m C_m^k A^k B^{m-k}$, 其中 C_m^k 是组合数.

(3) 若 $AB = BA$ 且 $B^m = O$, 则 $f(A+B) = f(A) + \sum_{k=1}^{m-1} \frac{1}{k!} f^{(k)}(A) B^k$.

(4) 若 A, B 都是对称方阵, 则 AB 是对称方阵 $\Leftrightarrow AB = BA$.

11. (1) 设 $A, B \in \mathbb{C}^{m \times n}$. 证明: $\text{tr}(AA^H) \text{tr}(BB^H) \geq \text{tr}(AB^H) \text{tr}(BA^H) \geq 0$.

(2) 设 $A \in \mathbb{C}^{m \times n}$ 满足 $\text{tr}(AA^H) = 0$. 证明: $A = O$.

(3) 设 $A \in \mathbb{C}^{n \times n}$ 满足 $\text{tr}(AA^H) = \text{tr}(A^2)$. 证明: $A = A^H$.

12. (1) 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$ 都是上三角矩阵. 证明: AB 也是上三角矩阵.

(2) 设 $A \in \mathbb{F}^{n \times n}$ 是上三角方阵, $f \in \mathbb{F}[x]$. 证明: $f(A)$ 也是上三角方阵.

13. 设 n 阶方阵 $A = (a_{ij}(x))$ 和 $B = (b_{ij}(x))$ 的元素都是可微函数. 记 $\frac{dA}{dx} = (a'_{ij}(x))$.

(1) 证明: $\frac{d(AB)}{dx} = \frac{dA}{dx}B + A\frac{dB}{dx}$.

(2) A 与 $\frac{dA}{dx}$ 是否一定乘积可交换? 证明你的结论.

14. 证明例 2.6 (有限射影平面) 中的 (2.3) 和 (2.4) 式.

15. 设 $Q_0 = 1$, $Q_1 = x$, $Q_n = \frac{Q_{n-1}^2 - 1}{Q_{n-2}}$, $\forall n \geq 2$.

(1)^[7]证明: Q_n 是以 x 为变元的整数系数多项式.

(2) 用矩阵方幂表示 Q_n 的通项公式.

16. 利用矩阵乘积求递推数列 $\{x_n\}_{n \in \mathbb{N}}$ 的通项公式.

(1) $x_n = a_1x_{n-1} + a_2x_{n-2} + \cdots + a_kx_{n-k}$, 其中 k 是给定的正整数, a_1, a_2, \cdots, a_k 是常数.

(2) $x_n = \frac{ax_{n-1} + b}{cx_{n-1} + d}$, 其中 a, b, c, d 是常数.

17. ^[8]设 n 是正整数, $a_1, a_2, \cdots, a_{n-1}$ 是任意实数, 定义递推数列 $u_0 = u_1 = v_0 = v_1 = 1$,

$$u_{k+1} = u_k + a_k u_{k-1}, \quad v_{k+1} = v_k + a_{n-k} v_{k-1}, \quad \forall k = 1, \cdots, n-1.$$

证明: $u_n = v_n$.

18. ^[9]设 a_0, a_1, a_2, \cdots 是整数, b_0, b_1, b_2, \cdots 是正整数, 满足 $a_0 = 0$, $a_1 = 1$,

$$a_{n+1} = \begin{cases} a_n b_n + a_{n-1}, & b_{n-1} = 1 \\ a_n b_n - a_{n-1}, & b_{n-1} > 1 \end{cases}, \quad n = 1, 2, \cdots$$

证明: $\max(a_n, a_{n+1}) \geq n$.

^[7]2017 年 Putnam 数学竞赛 A2.

^[8]2013 年 IMO 预选题 A1.

^[9]2017 年 IMO 预选题 A7.

§2.2 分块矩阵

定义 2.4. 设 $A = (a_{ij})_{m \times n}$, $I = (i_1, i_2, \dots, i_r)$ 是 $\{1, 2, \dots, m\}$ 中 r 个元素的排列, $J = (j_1, j_2, \dots, j_s)$ 是 $\{1, 2, \dots, n\}$ 中 s 个元素的排列. 由 A 的第 i_1, i_2, \dots, i_r 行和第 j_1, j_2, \dots, j_s 列元素排成的矩阵

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_s} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i_r j_1} & a_{i_r j_2} & \cdots & a_{i_r j_s} \end{pmatrix}$$

称为 A 的子矩阵, 记作 $A[I, J]$ 或 $A \begin{bmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_s \end{bmatrix}$. 特别, 当 $I = J$ 时, $A[I, I]$ 称为 A 的主子矩阵. $A \begin{bmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{bmatrix}$ 称为 A 的第 k 个顺序主子矩阵.

定义 2.5. 设 $m = m_1 + m_2 + \cdots + m_p$, $n = n_1 + n_2 + \cdots + n_q$, 把 $m \times n$ 矩阵 A 分块成

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}$$

其中每个 A_{ij} 是 $m_i \times n_j$ 的子矩阵. 这种分块方式 $(A_{ij})_{p \times q}$ 称为分块矩阵.

- 若 $A_{ij} = O$, $\forall i > j$, 则 $(A_{ij})_{p \times q}$ 称为准上三角矩阵.
- 若 $A_{ij} = O$, $\forall i < j$, 则 $(A_{ij})_{p \times q}$ 称为准下三角矩阵.
- 若 $A_{ij} = O$, $\forall i \neq j$, 则 $(A_{ij})_{p \times p}$ 称为准对角矩阵, 记作 $\text{diag}(A_{11}, A_{22}, \dots, A_{pp})$.



准上三角/准下三角/准对角矩阵的概念都是针对分块矩阵 $(A_{ij})_{p \times q}$ 而言, 不涉及原矩阵 A .

例 2.9. 考虑如下的网格图 G 的邻接矩阵 $A = (a_{ij})_{20 \times 20}$. 把 A 的 400 个元素都写出来显然很费事, 逐个指明 A 中所有 1 的位置也很麻烦, 而把 A 表示为分块矩阵的形式则相对容易一些.

$$A = \begin{pmatrix} P_5 & I_5 & O & O \\ I_5 & P_5 & I_5 & O \\ O & I_5 & P_5 & I_5 \\ O & O & I_5 & P_5 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

定理 2.3. 分块矩阵的运算具有下列性质.

1. 设 $A = (A_{ij})_{p \times q}$, 则 $A^T = (B_{ij})_{q \times p}$, 其中 $B_{ij} = A_{ji}^T$, $\forall i, j$.
2. 设 $A = (A_{ij})_{p \times q}$, λ 是数, 则 $\lambda A = (\lambda A_{ij})_{p \times q}$.
3. 设 $A = (A_{ij})_{p \times q}$, $B = (B_{ij})_{p \times q}$ 满足 A_{ij} 与 B_{ij} 的大小相同, 则 $A + B = (A_{ij} + B_{ij})_{p \times q}$.

4. 设 $A = (A_{ij})_{p \times q}$, $B = (B_{ij})_{q \times r}$ 满足 A_{ik} 的列数等于 B_{kj} 的行数, 则 $AB = (C_{ij})_{p \times r}$, 其中

$$C_{ij} = \sum_{k=1}^q A_{ik} B_{kj}, \quad \forall i, j.$$

例 2.10. 求与 $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ 乘积可交换的所有方阵 B .

解答. 设 $A = \begin{pmatrix} I & O \\ O & 2I \end{pmatrix}$, $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, 其中每个 B_{ij} 是 2×2 的方阵. 由 $AB = \begin{pmatrix} B_{11} & B_{12} \\ 2B_{21} & 2B_{22} \end{pmatrix}$, $BA = \begin{pmatrix} B_{11} & 2B_{12} \\ B_{21} & 2B_{22} \end{pmatrix}$ 可得 $B_{12} = B_{21} = O$. 容易验证, 形如 $\begin{pmatrix} B_{11} & O \\ O & B_{22} \end{pmatrix}$ 的方阵 B 都满足 $AB = BA$, 即为所求. \square

定义 2.6. 设 $A = (a_{ij})$ 是 $m \times n$ 矩阵, B 是 $p \times q$ 矩阵, 则 $mp \times nq$ 矩阵

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

称为 A 和 B 的张量积或 **Kronecker 积**, 记作 $C = A \otimes B$. 例如,

- 例 2.10 中的矩阵 $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \otimes I_2$.
- 例 2.9 中的矩阵 $A = I_4 \otimes P_5 + P_4 \otimes I_5$.

例 2.11. 设 2 阶方阵 $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$, $X = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix}$.

1. 矩阵方程 $AX = Y$ 可以表示为线性方程组 $Px = y$ 的形式,

$$\begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix},$$

其中系数矩阵 $P = A \otimes I$, x 由 X 的行向量拼接而成, y 由 Y 的行向量拼接而成.

2. 矩阵方程 $XB = Y$ 可以表示为线性方程组 $Qx = y$ 的形式,

$$\begin{pmatrix} b_1 & b_3 & 0 & 0 \\ b_2 & b_4 & 0 & 0 \\ 0 & 0 & b_1 & b_3 \\ 0 & 0 & b_2 & b_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix},$$

其中系数矩阵 $Q = I \otimes B^T$.

3. 矩阵方程 $AXB = Y$ 可以表示为线性方程组 $Mx = y$ 的形式,

$$\begin{pmatrix} a_1b_1 & a_1b_3 & a_2b_1 & a_2b_3 \\ a_1b_2 & a_1b_4 & a_2b_2 & a_2b_4 \\ a_3b_1 & a_3b_3 & a_4b_1 & a_4b_3 \\ a_3b_2 & a_3b_4 & a_4b_2 & a_4b_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix},$$

其中系数矩阵 $M = A \otimes B^T$.

4. 结合 1,2,3 可得 $M = PQ = QP$, 即 $A \otimes B^T = (A \otimes I)(I \otimes B^T) = (I \otimes B^T)(A \otimes I)$.

习题

1. 证明定理 2.3.

2. 设 A_i, X_i, Y_i, X, Y 都是 n 阶方阵, m 是正整数, 计算下列矩阵乘积或方幂.

$$\begin{aligned} (1) & \begin{pmatrix} X_1 & X_2 \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} & (2) & \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix} \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix} \\ (3) & \begin{pmatrix} I & X \\ O & I \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} I & -X \\ O & I \end{pmatrix} & (4) & \begin{pmatrix} I & X \\ O & I \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} I & O \\ Y & I \end{pmatrix} \\ (5) & \begin{pmatrix} O & X \\ Y & O \end{pmatrix}^m & (6) & \begin{pmatrix} X & O \\ Y & X \end{pmatrix}^m & (7) & \begin{pmatrix} O & X & O \\ O & O & X \\ X & O & O \end{pmatrix}^m & (8) & \begin{pmatrix} X & I & O \\ O & X & I \\ O & O & X \end{pmatrix}^m \end{aligned}$$

3. 设方阵 $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} O & A & O \\ O & O & A \\ A & O & O \end{pmatrix}$. 求与 B 乘积可交换的所有方阵.

4. 设 $m \times n$ 网格图 G 的顶点集 $V = \{(x, y) \mid x = 1, 2, \dots, m; y = 1, 2, \dots, n\}$, 两个顶点 (x_1, y_1) 与 (x_2, y_2) 相邻当且仅当 $|x_1 - x_2| + |y_1 - y_2| = 1$. 把 V 按照字典顺序排列. 用分块矩阵表示 G 的邻接矩阵.

5. 定义两个图 G_1, G_2 的直积 $G = G_1 \times G_2$ 如下. G 的顶点集 $V = \{(v_1, v_2) \mid v_i \text{ 是 } G_i \text{ 的顶点}\}$, $(u_1, u_2) \rightarrow (v_1, v_2)$ 是 G 的边当且仅当 $(u_1 = v_1 \text{ 且 } u_2 \rightarrow v_2 \text{ 是 } G_2 \text{ 的边})$ 或者 $(u_2 = v_2 \text{ 且 } u_1 \rightarrow v_1 \text{ 是 } G_1 \text{ 的边})$. 用 G_1, G_2 的邻接矩阵表示 G 的邻接矩阵.

6. 设 n 维超立方体图 Q_n 的顶点集 $V = \{0, 1\}^n$, 两个顶点 (x_1, x_2, \dots, x_n) 与 (y_1, y_2, \dots, y_n) 相邻当且仅当 $|x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n| = 1$. 把 V 按照字典顺序排列. 思考 Q_n 与 Q_{n-1} 的关系, 并用分块矩阵递推表示 Q_n 的邻接矩阵.

7. 设矩阵乘积 AC, BD 有意义. 证明: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

8. 设 A, A_1, A_2 是 $m \times n$ 矩阵, B, B_1, B_2 是 $p \times q$ 矩阵, C 是 $s \times t$ 矩阵. 证明:

$$\begin{aligned} (1) & (A_1 + A_2) \otimes B = (A_1 \otimes B) + (A_2 \otimes B) & (2) & A \otimes (B_1 + B_2) = (A \otimes B_1) + (A \otimes B_2) \\ (3) & (A \otimes B)^T = (A^T) \otimes (B^T) & (4) & (A \otimes B) \otimes C = A \otimes (B \otimes C) \end{aligned}$$

9. 设 A 是 m 阶方阵, B 是 n 阶方阵, X, Y 都是 $m \times n$ 矩阵, x, y 分别是 X, Y 的列向量依次拼接而成的 mn 维列向量. 把下列矩阵方程表示为 $y = Mx$ 的形式.

(1) $Y = AX$

(2) $Y = XB$

(3) $Y = AXB$

(4) $Y = AX - XB$

§2.3 初等方阵

在第一章中, 介绍了对于线性方程组, 或者说是对于矩阵的行向量进行操作的三种初等变换. 事实上, 它们都可以表示为矩阵乘积的形式. 例如, 设矩阵

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}, \quad \alpha_i \text{ 是行向量.}$$

- 交换 A 的第 1, 3 行, 得到

$$B_1 = \begin{pmatrix} \alpha_3 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = P_1 A.$$

- 把 A 的第 2 行乘上常数 λ , 得到

$$B_2 = \begin{pmatrix} \alpha_1 \\ \lambda\alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = P_2 A.$$

- 把 A 的第 1 行的 λ 倍加至第 3 行, 得到

$$B_3 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 + \lambda\alpha_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = P_3 A.$$

上述方阵 P_1, P_2, P_3 都可以看作是对单位方阵作同样的初等变换的结果.

定义 2.7. 对单位方阵的行向量作三类初等变换所得到的方阵称为初等方阵. 它们分别是

- $S_{ij} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}, \quad i \neq j.$
- $D_i(\lambda) = I_n + (\lambda - 1)E_{ii}, \quad \lambda \neq 0.$
- $T_{ij}(\lambda) = I + \lambda E_{ij}, \quad i \neq j.$

定理 2.4. 初等方阵具有下列性质.

- S_{ij} 是对称方阵. $S_{ij}^2 = I$. 当 $\{i, j\} \cap \{k, l\} = \emptyset$ 时, S_{ij} 与 S_{kl} 乘积可交换.
- $D_i(\lambda)$ 是对角方阵. $D_i(\lambda)D_i(\frac{1}{\lambda}) = I$. 任意 $D_i(\lambda)$ 与 $D_j(\mu)$ 乘积可交换.
- $T_{ij}(\lambda)$ 是三角方阵. $T_{ij}(\lambda)T_{ij}(-\lambda) = I$. 当 $i \neq l$ 且 $j \neq k$ 时, $T_{ij}(\lambda)$ 与 $T_{kl}(\mu)$ 乘积可交换.
- S_{ij} 可以表示为其它两类初等方阵的乘积. $S_{ij} = D_i(-1)T_{ij}(-1)T_{ji}(1)T_{ij}(-1).$

对单位方阵的列向量作初等变换得到的方阵也是初等方阵. 对矩阵的行向量作初等变换等价于矩阵左乘上初等方阵, 对矩阵的列向量作初等变换等价于矩阵右乘上初等方阵. 对于分块矩阵的行与列, 也可以作类似的“分块初等变换”.

$$\begin{aligned} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} C & D \\ A & B \end{pmatrix}, & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} O & I \\ I & O \end{pmatrix} &= \begin{pmatrix} B & A \\ D & C \end{pmatrix}, \\ \begin{pmatrix} P & O \\ O & Q \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} PA & PB \\ QC & QD \end{pmatrix}, & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} P & O \\ O & Q \end{pmatrix} &= \begin{pmatrix} AP & BQ \\ CP & DQ \end{pmatrix}, \\ \begin{pmatrix} I & O \\ X & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} A & B \\ XA+C & XB+D \end{pmatrix}, & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & O \\ X & I \end{pmatrix} &= \begin{pmatrix} A+BX & B \\ C+DX & D \end{pmatrix}. \end{aligned}$$

定理 2.5. 对于任意 $A \in \mathbb{F}^{m \times n}$, 有如下结论.

1. 存在非负整数 r 和一系列初等方阵 $P_1, \dots, P_s \in \mathbb{F}^{m \times m}$, $Q_1, \dots, Q_t \in \mathbb{F}^{n \times n}$, 使得

$$P_s \cdots P_1 A Q_1 \cdots Q_t = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

2. 存在非负整数 r 和一系列初等方阵 $\tilde{P}_1, \dots, \tilde{P}_s \in \mathbb{F}^{m \times m}$, $\tilde{Q}_1, \dots, \tilde{Q}_t \in \mathbb{F}^{n \times n}$, 使得

$$A = \tilde{P}_1 \cdots \tilde{P}_s \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{Q}_t \cdots \tilde{Q}_1.$$

证明. 当 $A = O$ 时, 设 $P_i = \tilde{P}_i = I_m$, $Q_i = \tilde{Q}_i = I_n$, $r = 0$, 结论显然成立. 下设 $A = (a_{ij}) \neq O$.

1. 对 m 使用数学归纳法. 当 $m = 1$ 时, 设 $a_{1j} \neq 0$, 则有

$$A \cdot D_j\left(\frac{1}{a_{1j}}\right) \cdot \prod_{k \neq j} T_{jk}(-a_{1k}) \cdot S_{1j} = (1 \ 0 \ \cdots \ 0).$$

当 $m \geq 2$ 时, 设 $a_{ij} \neq 0$, 则有

$$S_{1i} \cdot \prod_{k \neq i} T_{ki}(-a_{kj}) \cdot A \cdot D_j\left(\frac{1}{a_{1j}}\right) \cdot \prod_{k \neq j} T_{jk}(-a_{ik}) \cdot S_{1j} = \begin{pmatrix} 1 & \\ & B \end{pmatrix}.$$

对于矩阵 B 应用归纳假设, 存在非负整数 r 和一系列初等方阵 $P_1, \dots, P_s, Q_1, \dots, Q_t$, 使得

$$P_s \cdots P_1 B Q_1 \cdots Q_t = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}, \text{ 从而有}$$

$$\begin{pmatrix} 1 & \\ & P_s \end{pmatrix} \cdots \begin{pmatrix} 1 & \\ & P_1 \end{pmatrix} \begin{pmatrix} 1 & \\ & B \end{pmatrix} \begin{pmatrix} 1 & \\ & Q_1 \end{pmatrix} \cdots \begin{pmatrix} 1 & \\ & Q_t \end{pmatrix} = \begin{pmatrix} I_{r+1} & \\ & O \end{pmatrix},$$

其中所有 $\begin{pmatrix} 1 & \\ & P_i \end{pmatrix}, \begin{pmatrix} 1 & \\ & Q_i \end{pmatrix}$ 也都是初等方阵. 故定理结论成立.

2. 根据定理 2.4, 对于任意初等方阵 P_i, Q_i , 都存在初等方阵 \tilde{P}_i, \tilde{Q}_i , 使得 $\tilde{P}_i P_i = I_m, Q_i \tilde{Q}_i = I_n$.

因此, $\tilde{P}_1 \cdots \tilde{P}_s \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{Q}_t \cdots \tilde{Q}_1 = \tilde{P}_1 \cdots \tilde{P}_s P_s \cdots P_1 A Q_1 \cdots Q_t \tilde{Q}_t \cdots \tilde{Q}_1 = A. \quad \square$

习题

1. 证明定理 2.4.
2. 证明: 对于三类初等方阵 P , 均存在列向量 α, β 使得 $P = I + \alpha\beta^T$.
3. 若 n 维列向量 α, β 满足 $\alpha^T\beta = 0$, 则 n 阶方阵 $P = I + \alpha\beta^T$ 称为平延. 证明: 对于任意两个不共线的列向量 u, v , 存在平延 P 使得 $Pu = v$.
4. 设 n 阶方阵 A 的每行每列均恰有一个元素是 1、其它元素都是 0. A 称为置换方阵. 证明:
 - (1) $A^T A = AA^T = I_n$.
 - (2) 可经过至多 $n - 1$ 次交换两行 (列) 的操作, 把 A 变为单位方阵.
5. 设 A 是 $m \times n$ 矩阵, i_1, i_2, \dots, i_m 是 $1, 2, \dots, m$ 的排列, j_1, j_2, \dots, j_n 是 $1, 2, \dots, n$ 的排列. 证明: $P = I_m \begin{bmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{bmatrix}$ 和 $Q = I_n \begin{bmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{bmatrix}$ 都是置换方阵, 使得

$$PAQ = A \begin{bmatrix} i_1 & i_2 & \dots & i_m \\ j_1 & j_2 & \dots & j_n \end{bmatrix}.$$

特别, 当 $m = n$ 且 $(i_1, i_2, \dots, i_m) = (j_1, j_2, \dots, j_n)$ 时, $Q = P^T$.

6. 证明: 任意 $T_{ij}(\lambda)$ 都可以表示为

$$T_{pq}(a)T_{uv}(b)T_{pq}(-a)T_{uv}(-b)$$

的形式^[10], 其中 $i \neq j, p \neq q, u \neq v$.

7. 证明: 对于任意 $m \times n$ 矩阵 A , 存在一系列 m 阶初等方阵 P_1, P_2, \dots, P_s 和 n 阶置换方阵 Q , 使得 $P_s \cdots P_2 P_1 A Q$ 形如 $\begin{pmatrix} I_r & * \\ O & O \end{pmatrix}$, 其中 $0 \leq r \leq \min(m, n)$.
8. 对角元素都是 1 的上 (下) 三角矩阵称为单位上 (下) 三角矩阵. 证明: 对于任意 $m \times n$ 矩阵 A , 存在一系列 m 阶单位下三角的初等方阵 P_1, P_2, \dots, P_s 和 m 阶置换方阵 Q , 使得 $P_s \cdots P_2 P_1 Q A$ 是上三角矩阵.
9. 设 \mathbb{R}^3 中以 x, y, z 轴为转轴的三类旋转变换分别对应矩阵

$$P_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, P_2(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}, P_3(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- (1) 证明: $P_i(\alpha)P_i(\beta) = P_i(\alpha + \beta), \forall i, \alpha, \beta$.

- (2) 求 $\theta_1, \theta_2, \theta_3$ 使得 $P_1(\theta_1)P_2(\theta_2)P_3(\theta_3) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

- (3) 证明: 对于任意 3 阶实数方阵 A , 存在 $\theta_1, \theta_2, \theta_3$ 使得 $P_1(\theta_1)P_2(\theta_2)P_3(\theta_3)A$ 是上三角方阵.

^[10] 在一个乘法群中, 形如 $ghg^{-1}h^{-1}$ 的元素称为换位子.

§2.4 可逆矩阵

在 §2.1 中定义了矩阵的加法、减法、数乘、乘法，以及正整数次幂的运算，定理 2.2 也表明单位方阵是矩阵乘法的单位元，我们自然考虑是否能够定义矩阵的除法运算或负整数次幂运算。由于矩阵乘法运算不满足交换律，我们有如下定义。

定义 2.8. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times m}$. 若 $AB = I_m$ 是单位方阵，则称 A 是 B 的一个左逆， B 是 A 的一个右逆。若 A 既有右逆 B 又有左逆 C ，则有 $C = C(AB) = (CA)B = B$ ，此时 A 称为可逆矩阵， B 称为 A 的逆矩阵，记作 A^{-1} 。

显然，并非所有矩阵都是可逆矩阵。例如，零矩阵既没有左逆也没有右逆，不是可逆矩阵。那么，如何判断一个矩阵 A 是否可逆矩阵？定理 2.9 以线性方程组的形式给出了 A 是可逆矩阵的一些充分必要条件。在第三章和第四章中，我们还将通过行列式和秩的概念给出 A 是可逆矩阵的其它充分必要条件。

例 2.12. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{F}^{2 \times 2}$. 当 $\delta = ad - bc \neq 0$ 时， A 是可逆矩阵，

$$A^{-1} = \frac{1}{\delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

例 2.13. 设 $A = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \in \mathbb{F}^{3 \times 3}$. 当 $\delta = \mathbf{u} \times \mathbf{v} \cdot \mathbf{w} \neq 0$ 时， A 是可逆矩阵，

$$A^{-1} = \frac{1}{\delta} \begin{pmatrix} \mathbf{v} \times \mathbf{w} & \mathbf{w} \times \mathbf{u} & \mathbf{u} \times \mathbf{v} \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} v_2 w_3 - v_3 w_2 & w_2 u_3 - w_3 u_2 & u_2 v_3 - u_3 v_2 \\ v_3 w_1 - v_1 w_3 & w_3 u_1 - w_1 u_3 & u_3 v_1 - u_1 v_3 \\ v_1 w_2 - v_2 w_1 & w_1 u_2 - w_2 u_1 & u_1 v_2 - u_2 v_1 \end{pmatrix}.$$

在 $\mathbf{u} \times \mathbf{v} \cdot \mathbf{w}$ 式中，

$$(x_1, x_2, x_3) \times (y_1, y_2, y_3) = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1)$$

是 \mathbb{F}^3 上的向量积运算，

$$(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = x_1 y_1 + x_2 y_2 + x_3 y_3$$

是 \mathbb{F}^3 上的数量积运算。

例 2.14. 所有初等方阵都是可逆矩阵。

例 2.15. $m \times n$ 矩阵 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是可逆矩阵的充分必要条件是 $m = n = r$ 。

根据定义，可逆矩阵具有以下常用性质。

定理 2.6. 若 A 是可逆矩阵，则 A^{-1} 和 A^T 也是可逆矩阵，并且

$$(A^{-1})^{-1} = A, \quad (A^T)^{-1} = (A^{-1})^T.$$

- 类似地, 当 A_4 是可逆方阵时, 有矩阵乘积分解

$$A = \begin{pmatrix} I & A_2 A_4^{-1} \\ O & I \end{pmatrix} \begin{pmatrix} A_1 - A_2 A_4^{-1} A_3 & O \\ O & A_4 \end{pmatrix} \begin{pmatrix} I & O \\ A_4^{-1} A_3 & I \end{pmatrix}.$$

上式也称为 **Schur 公式**, $T = A_1 - A_2 A_4^{-1} A_3$ 称为 A_4 的 **Schur 补**. A 是可逆方阵当且仅当 T 是可逆方阵. 此时,

$$\begin{aligned} A^{-1} &= \begin{pmatrix} I & O \\ -A_4^{-1} A_3 & I \end{pmatrix} \begin{pmatrix} T^{-1} & O \\ O & A_4^{-1} \end{pmatrix} \begin{pmatrix} I & -A_2 A_4^{-1} \\ O & I \end{pmatrix} \\ &= \begin{pmatrix} T^{-1} & -T^{-1} A_2 A_4^{-1} \\ -A_4^{-1} A_3 T^{-1} & A_4^{-1} + A_4^{-1} A_3 T^{-1} A_2 A_4^{-1} \end{pmatrix}. \end{aligned}$$

- 当 A_1, A_4 和 A 都是可逆方阵时, 比较以上两个关于 A^{-1} 的表达式, 可得 S 和 T 之间的联系

$$S^{-1} = A_4^{-1} + A_4^{-1} A_3 T^{-1} A_2 A_4^{-1}, \quad T^{-1} = A_1^{-1} + A_1^{-1} A_2 S^{-1} A_3 A_1^{-1}.$$

例 2.19. 设 n 阶下三角方阵 $A = (C_i^j)$, 其中 $C_i^j = \frac{i!}{j!(i-j)!}$ 是组合数. 求 A^{-1} .

解答. 设 $M = (C_i^j) = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{1} & A \end{pmatrix}$, 其中 $0 \leq i, j \leq n$. 根据二项式定理,

$$\begin{aligned} \begin{pmatrix} 1 \\ 1+x \\ \vdots \\ (1+x)^n \end{pmatrix} &= \begin{pmatrix} C_0^0 & & & \\ C_1^0 & C_1^1 & & \\ \vdots & \vdots & \ddots & \\ C_n^0 & C_n^1 & \cdots & C_n^n \end{pmatrix} \begin{pmatrix} 1 \\ x \\ \vdots \\ x^n \end{pmatrix}, \\ \begin{pmatrix} 1 \\ x \\ \vdots \\ x^n \end{pmatrix} &= \begin{pmatrix} C_0^0 & & & \\ -C_1^0 & C_1^1 & & \\ \vdots & \vdots & \ddots & \\ (-1)^n C_n^0 & (-1)^{n-1} C_n^1 & \cdots & C_n^n \end{pmatrix} \begin{pmatrix} 1 \\ 1+x \\ \vdots \\ (1+x)^n \end{pmatrix}. \end{aligned}$$

因此, $M^{-1} = \begin{pmatrix} 1 & \mathbf{0} \\ * & A^{-1} \end{pmatrix} = ((-1)^{i-j} C_i^j)$, $A^{-1} = ((-1)^{i-j} C_i^j)$. □

习题

1. 证明定理 2.6, 2.7, 2.8.
2. 证明: \mathbb{F}^3 上的向量积运算具有下列性质, 其中 $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}^3$, $\lambda, \mu \in \mathbb{F}$.
 - (1) 双线性 $(\lambda \mathbf{a} + \mu \mathbf{b}) \times \mathbf{c} = \lambda(\mathbf{a} \times \mathbf{c}) + \mu(\mathbf{b} \times \mathbf{c})$, $\mathbf{a} \times (\lambda \mathbf{b} + \mu \mathbf{c}) = \lambda(\mathbf{a} \times \mathbf{b}) + \mu(\mathbf{a} \times \mathbf{c})$
 - (2) 反对称 $\mathbf{a} \times \mathbf{a} = \mathbf{0}$, $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
 - (3) 混合积 $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}$
 - (4) 双重向量积 $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a}$
 - (5) Jacobi 恒等式 $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} + (\mathbf{b} \times \mathbf{c}) \times \mathbf{a} + (\mathbf{c} \times \mathbf{a}) \times \mathbf{b} = \mathbf{0}$

3. 判断下列方阵 A 是否可逆方阵, 或者求 a, b, c, d 应满足的充分必要条件, 使得 A 是可逆方阵. 当 A 是可逆方阵时, 求 A^{-1} .

$$(1) \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$(2) \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$(3) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$(4) \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

$$(5) \begin{pmatrix} 0 & b & c & d \\ -b & 0 & b & c \\ -c & -b & 0 & b \\ -d & -c & -b & 0 \end{pmatrix}$$

$$(6) \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}$$

$$(7) \begin{pmatrix} a & a+1 & a+2 & a+3 \\ b & b+1 & b+2 & b+3 \\ c & c+1 & c+2 & c+3 \\ d & d+1 & d+2 & d+3 \end{pmatrix}$$

$$(8) \begin{pmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{pmatrix}$$

4. 设 $A = (a_{ij})$ 是 n 阶三角方阵. 证明: A 是可逆方阵当且仅当 $\prod_{i=1}^n a_{ii} \neq 0$.

5. 设 A 是 n 阶可逆方阵. 证明:

- (1) 若 A 是上/下三角方阵, 则 A^{-1} 也是上/下三角方阵.
- (2) 若 A 是 (反) 对称方阵, 则 A^{-1} 也是 (反) 对称方阵.
- (3) 若 A 是 (反) Hermite 方阵, 则 A^{-1} 也是 (反) Hermite 方阵.

6. 设 n 阶方阵 A 满足 $A^k = \lambda I$, k 是正整数, $\lambda \neq 1$. 证明: $I - A$ 是可逆方阵, 并且

$$(I - A)^{-1} = \frac{1}{1 - \lambda}(I + A + A^2 + \cdots + A^{k-1}).$$

7. 设 n 阶方阵 $A = (a_i b_j)$, $B = \lambda I - A$, $\mu = \sum_{i=1}^n a_i b_i$, $n \geq 2$. 证明:

$$(1) A^2 = \mu A, B^2 = (2\lambda - \mu)B - \lambda(\lambda - \mu)I.$$

$$(2) B \text{ 是可逆方阵当且仅当 } \lambda \notin \{0, \mu\}, \text{ 并且 } B^{-1} = \frac{1}{\lambda}I + \frac{1}{\lambda(\lambda - \mu)}A.$$

8. 设 A, B 分别是 $m \times n$ 和 $n \times m$ 的矩阵. 何时 $M = \begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 是可逆方阵? 并求 M^{-1} .

9. 设 A, B 是方阵. 证明: $A \otimes B$ 是可逆方阵当且仅当 A, B 都是可逆方阵, 并且

$$(A \otimes B)^{-1} = (A^{-1}) \otimes (B^{-1}).$$

10. (Sherman-Morrison-Woodbury 公式) 设 A 是 m 阶可逆方阵, B 是 $m \times n$ 矩阵, C 是 $n \times m$ 矩阵. 证明: $A + BC$ 是可逆方阵当且仅当 $I + CA^{-1}B$ 是可逆方阵, 并且

$$(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1}.$$

特别, $I_m - BC$ 是可逆方阵当且仅当 $I_n - CB$ 是可逆方阵, 并且

$$(I_m - BC)^{-1} = I_m + B(I_n - CB)^{-1}C.$$

11. 设 n 阶可逆方阵 $A = (a_{ij}(x))$ 的元素都是可微函数. 证明: $\frac{d(A^{-1})}{dx} = -A^{-1} \frac{dA}{dx} A^{-1}$.

12. 设 n 阶复数方阵 $A = (a_{ij})$ 满足 $|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \forall i$. A 称为严格行对角优的.

证明: 线性方程组 $Ax = \mathbf{0}$ 只有零解, 从而 A 是可逆方阵.

13. 设 n 阶整数方阵 $A = (a_{ij})$, 当 j 整除 i 时, $a_{ij} = 1$, 否则 $a_{ij} = 0$. 证明:

$$A^{-1} = \left(\mu\left(\frac{i}{j}\right) a_{ij} \right), \quad \text{其中 } \mu(x) = \begin{cases} (-1)^k, & x \text{ 是 } k \text{ 个不同素数的乘积;} \\ 0, & \text{否则.} \end{cases}$$

注: A 是 Möbius 变换 $f(x) \mapsto g(x) = \sum_{d|x} f(d)$ 的矩阵表示, $\mu(x)$ 称为 Möbius 函数.

14. 首先给出几个定义.

- 对于任意 $M \in \mathbb{F}^{n \times n}$, 若存在置换方阵 P 使得 $P^T M P = \begin{pmatrix} M_{11} & M_{12} \\ O & M_{22} \end{pmatrix}$, 其中 M_{11}, M_{22} 都是阶数 ≥ 1 的方阵, 则 M 称为可约的, 否则 M 称为不可约的.
- 若 $M \in \mathbb{R}^{m \times n}$ 的元素都 ≥ 0 , 则 M 称为非负矩阵.
- 对于非负矩阵 $M \in \mathbb{R}^{n \times n}$, 若存在正整数 k 使得 M^k 的元素都 > 0 , 则 M 称为本原的, 否则 M 称为非本原的.

设 $A \in \mathbb{R}^{n \times n}$ 是非负矩阵. 证明:

- (1) A 是不可约的 \Leftrightarrow 存在正整数 $k \leq n$ 使得 $I + A + \cdots + A^{k-1}$ 的元素都是正实数
 \Leftrightarrow 存在正实数 λ 使得 $(\lambda I - A)^{-1}$ 的元素都是正实数.
- (2) 存在置换方阵 P 使得 $P^T A P$ 是准上三角方阵, 其中每个准对角块都是不可约方阵.
- (3) 若 A 是本原的, 则 A 是不可约的.
- (4) 若 A 是非本原且不可约的, 则存在正整数 $m \geq 2$ 和置换方阵 P , 使得

$$P^T A P = \begin{pmatrix} O & A_1 & & & \\ & O & \ddots & & \\ & & & \ddots & \\ & & & & A_{m-1} \\ A_m & & & & O \end{pmatrix}$$

其中每个准对角块都是零方阵, 空白处元素都是 0.

第三章 行列式

行列式的概念来源于求解 n 个方程和 n 个未知数的线性方程组. 分别在 1683 和 1693 年, 关孝和^[1]与 Leibnitz^[2]独立地提出了行列式的概念, 用于求解线性方程组. 1750 年, Cramer^[3]发表了著名的 Cramer 法则. 1841 年, Cauchy^[4]首先提出了现代的行列式概念及其记号. 在 1750 至 1900 这段时期, 对行列式的研究成为线性代数的重要内容. 自 1858 年 Cayley 创立矩阵论以来, 矩阵逐渐发展为线性代数的主要组成部分和重要的数学工具, 而行列式则逐渐变成矩阵理论的一个小分支.

§3.1 行列式的定义

关于行列式的概念, 有许多形式不同但实质等价的定义. 大致有下列几种方式:

- 把行列式看作具有特定性质的函数.
- 直接给出 n 阶行列式的完全展开式.
- 通过 $n-1$ 阶行列式归纳定义 n 阶行列式.
- 定义行列式为平行多面体的有向体积.
- 通过 Grassmann^[5]代数定义行列式.

定义 3.1. 具有下列性质的 \mathbb{F}^n 上的 n 元函数 $\det(\alpha_1, \alpha_2, \dots, \alpha_n)$ 称为数域 \mathbb{F} 上的 n 阶行列式.

1. (多重线性) 行列式关于每个变量是线性的.

$$\det(\dots, \lambda\alpha_i + \mu\beta_i, \dots) = \lambda \det(\dots, \alpha_i, \dots) + \mu \det(\dots, \beta_i, \dots)$$

其中 $\lambda, \mu \in \mathbb{F}$, $\alpha_i, \beta_i \in \mathbb{F}^n$.

2. (反对称性) 若存在两个变量相等, 则行列式为 0. 从而, 交换两个变量的位置, 行列式反号.

$$\begin{aligned} \det(\dots, \alpha_i, \dots, \alpha_i, \dots) &= 0 \\ \det(\dots, \alpha_j, \dots, \alpha_i, \dots) &= -\det(\dots, \alpha_i, \dots, \alpha_j, \dots) \end{aligned}$$

^[1]Seki Takakazu, 约 1642–1708, 日本数学家.

^[2]Gottfried Wilhelm Leibniz, 1646–1716, 德国数学家、哲学家.

^[3]Gabriel Cramer, 1704–1752, 瑞士数学家.

^[4]Augustin-Louis Cauchy, 1789–1857, 法国数学家.

^[5]Hermann Günther Grassmann, 1809–1877, 德国数学家.

3. (规范性) 标准单位向量组的行列式等于 1.

$$\det(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$$

简而言之, 行列式 $\det(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是关于 n 个向量 $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}^n$ 的多元函数. 行列式 $\det(\alpha_1, \alpha_2, \dots, \alpha_n)$ 也可以看作是关于方阵 $A \in \mathbb{F}^{n \times n}$ 的函数, 记作 $\det(A)$ 或 $|A|$, 其中 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 A 的行向量. 设 $A = (a_{ij})_{n \times n}$, 则 $\det(A)$ 还可以看作是关于 n^2 个变元 $a_{ij} \in \mathbb{F}$ 的多元多项式.

定理 3.1. 由定义 3.1 可知, 行列式具有下列性质, 其中 $\lambda, \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$.

1. 若存在 $\alpha_i = \mathbf{0}$, 则 $\det(\alpha_1, \alpha_2, \dots, \alpha_n) = 0$.
2. 若存在 $\alpha_i = \lambda\alpha_j, i \neq j$, 则 $\det(\alpha_1, \alpha_2, \dots, \alpha_n) = 0$.
3. $\det(\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n) = \lambda_1\lambda_2 \cdots \lambda_n \det(\alpha_1, \alpha_2, \dots, \alpha_n)$.
4. $\det(\dots, \alpha_i, \dots, \alpha_j, \dots) = \det(\dots, \alpha_i + \lambda\alpha_j, \dots, \alpha_j, \dots), \forall i \neq j$.



定义 3.1 的合理性尚存在疑问. 如定义 3.1 所述的行列式函数是否存在? 是否唯一?

定义 3.2. 设 $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ 是 n 个两两不同的实数的一个排列. 满足 $i < j$ 且 $\sigma_i > \sigma_j$ 的有序实数对 (σ_i, σ_j) 称为 σ 中的一个**逆序**. σ 中逆序的个数称为 σ 的**逆序数**, 记作 $\tau(\sigma)$.

$1, 2, \dots, n$ 的所有排列的集合记作 S_n . 每个排列 σ 都可以看作是集合 $\{1, 2, \dots, n\}$ 上的一一映射. 在映射的复合运算下, S_n 构成群, 称为 n 次**对称群**.

例 3.1. 降序排列 $\sigma = (n, n-1, \dots, 2, 1)$ 的逆序数 $\tau(\sigma) = 1 + 2 + \dots + (n-1) = \frac{n(n-1)}{2}$.

定理 3.2. 交换排列 σ 中任意两项的位置, 所得排列记作 $\tilde{\sigma}$, 则 $\tau(\sigma)$ 与 $\tau(\tilde{\sigma})$ 具有不同的奇偶性.

证明. 设 $\tilde{\sigma}$ 由交换 σ 的第 i, j 项得到, $i < j$. 考虑 σ 和 $\tilde{\sigma}$ 中包含 σ_i 或 σ_j 的数对.

- 当 $k < i$ 或 $k > j$ 时, (σ_k, σ_i) 是 σ 中的逆序 $\Leftrightarrow (\sigma_k, \sigma_i)$ 是 $\tilde{\sigma}$ 中的逆序, (σ_j, σ_k) 是 σ 中的逆序 $\Leftrightarrow (\sigma_j, \sigma_k)$ 是 $\tilde{\sigma}$ 中的逆序.
- 当 $i < k < j$ 时, (σ_i, σ_k) 是 σ 中的逆序 $\Leftrightarrow (\sigma_k, \sigma_i)$ 不是 $\tilde{\sigma}$ 中的逆序, (σ_k, σ_j) 是 σ 中的逆序 $\Leftrightarrow (\sigma_j, \sigma_k)$ 不是 $\tilde{\sigma}$ 中的逆序.
- (σ_i, σ_j) 是 σ 中的逆序 $\Leftrightarrow (\sigma_j, \sigma_i)$ 不是 $\tilde{\sigma}$ 中的逆序.

综上, 共有 $2(j-i)-1$ 个数对的逆序性发生改变. 因此, $\tau(\tilde{\sigma}) - \tau(\sigma)$ 是奇数. \square

对于任意排列 $\sigma, \tilde{\sigma} \in S_n$, 如果可以经过 x 次交换把 σ 变成 $\tilde{\sigma}$, 则可以经过 x 次交换把 $\tilde{\sigma}$ 变成 σ . 如果还可以经过 y 次交换把 σ 变成 $\tilde{\sigma}$, 则可以经过 $x+y$ 次交换把 σ 变成 σ . 根据定理 3.2, $x+y$ 一定是偶数, 即 $(-1)^x = (-1)^y$.

例 3.2. 设 $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n) \in S_n$. $\det(\mathbf{e}_{\sigma_1}, \mathbf{e}_{\sigma_2}, \dots, \mathbf{e}_{\sigma_n}) = (-1)^{\tau(\sigma)}$.

解答. 依次取 $k = 1, 2, \dots, n-1$, 设 σ 中形如 $(*, k)$ 的逆序共有 m_k 个, 则可经过 m_k 次相邻元素的交换, 把 k 换到位置 k , 同时保持其它数的相对顺序不变. 合计经过 $\tau(\sigma) = \sum_{k=1}^{n-1} m_k$ 次交换把 σ 变成排列 $(1, 2, \dots, n)$. 因此, $\det(\mathbf{e}_{\sigma_1}, \mathbf{e}_{\sigma_2}, \dots, \mathbf{e}_{\sigma_n}) = (-1)^{\tau(\sigma)} \det(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = (-1)^{\tau(\sigma)}$. \square

下面, 我们推导行列式函数的代数表达式.

设 $\alpha_i = (a_{i1}, a_{i2}, \dots, a_{in})$, $i = 1, 2, \dots, n$. 根据定义 3.1,

$$\begin{aligned} \det(\alpha_1, \alpha_2, \dots, \alpha_n) &= \det\left(\sum_{j=1}^n a_{1j}\mathbf{e}_j, \sum_{j=1}^n a_{2j}\mathbf{e}_j, \dots, \sum_{j=1}^n a_{nj}\mathbf{e}_j\right) \\ &= \sum_{1 \leq j_1, j_2, \dots, j_n \leq n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_n}) \\ &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_n}) \\ &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} (-1)^{\tau(j_1, j_2, \dots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}. \end{aligned} \quad (3.1)$$

(3.1) 式称为 $\det(\alpha_1, \alpha_2, \dots, \alpha_n)$ 或 $\det(a_{ij})$ 的完全展开式. 许多教科书直接把它作为行列式的定义. 记 (3.1) 式定义的多元函数为 $\Delta(\alpha_1, \alpha_2, \dots, \alpha_n)$. 容易验证 $\Delta(\alpha_1, \alpha_2, \dots, \alpha_n)$ 满足定义 3.1 所述的多重线性和规范性. 由定理 3.2 可得 $\Delta(\alpha_1, \alpha_2, \dots, \alpha_n)$ 满足反对称性. 因此, 定义 3.1 中的行列式函数 $\det(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是存在且唯一的, 就是 $\Delta(\alpha_1, \alpha_2, \dots, \alpha_n)$.

例 3.3. 1 阶行列式 $\det(a) = a$, 2 阶行列式 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$, 3 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

例 3.4. 设 $A = (a_{ij})_{n \times n}$ 是上三角方阵, 则 $\det(A) = a_{11}a_{22} \cdots a_{nn}$.

证明. 在 $\det(A)$ 的完全展开式中, 当 $(j_1, j_2, \dots, j_n) \neq (1, 2, \dots, n)$ 时, $a_{1j_1}a_{2j_2} \cdots a_{nj_n} = 0$. \square

例 3.5. 设分块矩阵 $A = (A_{ij})_{p \times p}$ 是准上三角方阵, 并且每个 A_{ii} 是方阵, 则

$$\det(A) = \prod_{i=1}^p \det(A_{ii}).$$

证明. 设 $A = (a_{ij})_{n \times n}$. 当 $p = 2$ 时, 设 A_{11} 是 r 阶方阵, A_{22} 是 $n - r$ 阶方阵, 则

$$\begin{aligned} \det(A) &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} (-1)^{\tau(j_1, j_2, \dots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \\ &= \sum_{\substack{(j_1, \dots, j_r) \text{ 是 } (1, \dots, r) \text{ 的排列} \\ (j_{r+1}, \dots, j_n) \text{ 是 } (r+1, \dots, n) \text{ 的排列}}} (-1)^{\tau(j_1, \dots, j_r) + \tau(j_{r+1}, \dots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \\ &= \det(A_{11}) \det(A_{22}). \end{aligned}$$

当 $p \geq 3$ 时, 可先把 A 分块成 $\begin{pmatrix} A_{11} & * \\ O & B \end{pmatrix}$, 得 $\det(A) = \det(A_{11}) \det(B)$. 再对 B 继续分块. 由此可得 $\det(A) = \det(A_{11}) \det(A_{22}) \cdots \det(A_{pp})$. \square

行列式的完全展开式含有 $n!$ 个单项式. 当 n 较大时, 利用完全展开式求一般 n 阶行列式的计算量非常巨大, 不容易实现. 通常是利用行列式的定义及性质, 通过初等变换把方阵变成三角方阵, 从而求得原方阵的行列式.

例 3.6. 计算行列式
$$\begin{vmatrix} 0 & 2 & 1 & -1 \\ 2 & -2 & 1 & -1 \\ 1 & -1 & -2 & 2 \\ 1 & 1 & -1 & -2 \end{vmatrix}.$$

解答.
$$\begin{vmatrix} 0 & 2 & 1 & -1 \\ 2 & -2 & 1 & -1 \\ 1 & -1 & -2 & 2 \\ 1 & 1 & -1 & -2 \end{vmatrix} \xrightarrow{\text{交换①,③行}} - \begin{vmatrix} 1 & -1 & -2 & 2 \\ 2 & -2 & 1 & -1 \\ 0 & 2 & 1 & -1 \\ 1 & 1 & -1 & -2 \end{vmatrix} \xrightarrow{\substack{\text{行②}=-2\text{①} \\ \text{行④}=-\text{①}}} - \begin{vmatrix} 1 & -1 & -2 & 2 \\ 0 & 0 & 5 & -5 \\ 0 & 2 & 1 & -1 \\ 0 & 2 & 1 & -4 \end{vmatrix}$$

$$\xrightarrow{\text{交换②,③行}} \begin{vmatrix} 1 & -1 & -2 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 5 & -5 \\ 0 & 2 & 1 & -4 \end{vmatrix} \xrightarrow{\text{行④}=-\text{②}} \begin{vmatrix} 1 & -1 & -2 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & -3 \end{vmatrix} = 1 \cdot 2 \cdot 5 \cdot (-3) = -30. \quad \square$$

定理 3.3. 对于任意 n 阶方阵 $A = (a_{ij})$, $\det(A^T) = \det(A)$.

证明. 根据行列式的完全展开式 (3.1),

$$\det(A^T) = \sum_{(i_1, i_2, \dots, i_n) \in S_n} (-1)^{\tau(i_1, i_2, \dots, i_n)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

其中每个 $a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$ 可以重新排列成 $a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ 的形式. 即可以经过 $\tau(i_1, i_2, \dots, i_n)$ 次交换, 把 $\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ 1 & 2 & \cdots & n \end{pmatrix}$ 变成 $\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$. 同理, 可以经过 $\tau(j_1, j_2, \dots, j_n)$ 次交换, 把 $\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$ 变成 $\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ 1 & 2 & \cdots & n \end{pmatrix}$. 因此, $(-1)^{\tau(i_1, i_2, \dots, i_n)} = (-1)^{\tau(j_1, j_2, \dots, j_n)}$, 故 $\det(A^T) = \det(A)$. \square

定理 3.3 表明, $\det(A)$ 既可以看作是 A 的行向量的函数, 也可以看作是 A 的列向量的函数, 行与列的地位是平等的, 关于行向量的性质可以自然适用于列向量, 反之亦然. 在求行列式时, 既可以对行向量施行初等变换, 也可以同时对列向量施行初等变换.

例 3.7. 计算行列式
$$\begin{vmatrix} 2 & 1 & 1 & -1 \\ -1 & 0 & -1 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 0 & 2 & 0 \end{vmatrix}.$$

解答.
$$\begin{vmatrix} 2 & 1 & 1 & -1 \\ -1 & 0 & -1 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 0 & 2 & 0 \end{vmatrix} \xrightarrow{\substack{\text{交换①,④行} \\ \text{交换②,③列}}} \begin{vmatrix} 1 & 2 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & 2 & -2 & 1 \\ 2 & 1 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix} \cdot \begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = 1. \quad \square$$

习题

1. 证明完全展开式 (3.1) 定义的多元函数 $\Delta(\alpha_1, \alpha_2, \dots, \alpha_n)$ 满足定义 3.1 所述的三个性质.

2. 求下列排列的逆序数.

(1) (1, 7, 6, 10, 2, 5, 8, 4, 9, 3)

(2) (2, 10, 7, 9, 4, 6, 5, 3, 1, 8)

(3) (3, 5, 4, 9, 10, 1, 2, 8, 7, 6)

(4) (4, 2, 1, 3, 6, 9, 7, 10, 8, 5)

(5) (1, 3, 5, \dots , $2k-1$, 2, 4, 6, \dots , $2k$)

(6) (2, 4, 6, \dots , $2k$, 1, 3, 5, \dots , $2k-1$)

(7) (1, 3, 5, \dots , $2k-1$, $2k$, \dots , 6, 4, 2)

(8) ($2k$, \dots , 6, 4, 2, 1, 3, 5, \dots , $2k-1$)

3. 分别计算初等方阵 $S_{ij}, D_i(\lambda), T_{ij}(\lambda)$ 的行列式.

4. 设 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 是三角方阵, $f \in \mathbb{F}[x]$. 证明: $\det(f(A)) = f(a_{11})f(a_{22}) \cdots f(a_{nn})$.

5. 计算下列方阵的行列式.

$$(1) \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$(2) \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$(3) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$(4) \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$$

$$(5) \begin{pmatrix} 0 & a & b & c \\ a & 0 & d & e \\ b & d & 0 & f \\ c & e & f & 0 \end{pmatrix}$$

$$(6) \begin{pmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{pmatrix}$$

$$(7) \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

$$(8) \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix}$$

$$(9) \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}$$

$$(10) \begin{pmatrix} 1+a & 1+a^2 & 1+a^3 & 1+a^4 \\ 1+b & 1+b^2 & 1+b^3 & 1+b^4 \\ 1+c & 1+c^2 & 1+c^3 & 1+c^4 \\ 1+d & 1+d^2 & 1+d^3 & 1+d^4 \end{pmatrix}$$

$$(11) \begin{pmatrix} a & a+1 & a+2 & a+3 \\ b & b+1 & b+2 & b+3 \\ c & c+1 & c+2 & c+3 \\ d & d+1 & d+2 & d+3 \end{pmatrix}$$

$$(12) \begin{pmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{pmatrix}$$

$$(13) \begin{pmatrix} a^3 & (a+1)^3 & (a+2)^3 & (a+3)^3 \\ b^3 & (b+1)^3 & (b+2)^3 & (b+3)^3 \\ c^3 & (c+1)^3 & (c+2)^3 & (c+3)^3 \\ d^3 & (d+1)^3 & (d+2)^3 & (d+3)^3 \end{pmatrix}$$

$$(14) \begin{pmatrix} 0 & 1 & 2 & \cdots & n-1 \\ 1 & 0 & 1 & \ddots & \vdots \\ 2 & 1 & 0 & \ddots & 2 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ n-1 & \cdots & 2 & 1 & 0 \end{pmatrix}$$

$$(15) \begin{pmatrix} x & 1 & 2 & \cdots & n-1 \\ -1 & x & 1 & \ddots & \vdots \\ -2 & -1 & x & \ddots & 2 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 1-n & \cdots & -2 & -1 & x \end{pmatrix}$$

$$(16) \begin{pmatrix} O & \cdots & O & A_1 \\ \vdots & \ddots & A_2 & O \\ O & \ddots & \ddots & \vdots \\ A_p & O & \cdots & O \end{pmatrix}, \text{ 其中 } A_i = a_i I \text{ 是 } r_i \text{ 阶纯量方阵, } \forall i.$$

§3.2 Binet-Cauchy 公式

上节中给出了行列式的定义以及一些最基本的性质. 本节及下节将给出更多关于行列式和矩阵的重要性质, 以及运用这些性质求行列式和逆矩阵的方法. 基本思想方法是

1. 化繁为简, 把复杂矩阵分解为简单矩阵的乘积形式.
2. 分而治之, 把高阶行列式展开成低阶行列式.

定理 3.4. 对于任意 n 阶方阵 A, B , $\det(AB) = \det(A) \det(B)$.

证明. 设 $A = (a_{ij})$, B 的行向量分别为 $\beta_1, \beta_2, \dots, \beta_n$, 则 AB 的第 i 个行向量 $\gamma_i = \sum_{j=1}^n a_{ij} \beta_j$.

$$\begin{aligned} \det(AB) &= \det \left(\sum_{j=1}^n a_{1j} \beta_j, \sum_{j=1}^n a_{2j} \beta_j, \dots, \sum_{j=1}^n a_{nj} \beta_j \right) \\ &= \sum_{1 \leq j_1, j_2, \dots, j_n \leq n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det(\beta_{j_1}, \beta_{j_2}, \dots, \beta_{j_n}) \\ &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det(\beta_{j_1}, \beta_{j_2}, \dots, \beta_{j_n}) \\ &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} (-1)^{\tau(j_1, j_2, \dots, j_n)} \det(\beta_1, \beta_2, \dots, \beta_n) \\ &= \det(A) \det(B). \end{aligned} \quad \square$$

数域 \mathbb{F} 上行列式等于 1 的 n 阶方阵的全体, 在矩阵乘法运算下构成群, 称为特殊线性群, 记作 $SL(n, \mathbb{F})$.

定理 3.5. 方阵 A 是可逆方阵当且仅当 $\det(A) \neq 0$, 并且 $\det(A^{-1}) = \frac{1}{\det(A)}$.

证明. 根据定理 2.5, 存在一系列初等方阵 $P_1, \dots, P_s, Q_1, \dots, Q_t$, 使得

$$B = P_s \cdots P_1 A Q_1 \cdots Q_t = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

根据定理 2.7 和定理 2.14, A 可逆 $\Leftrightarrow B$ 可逆 $\Leftrightarrow r = n$. 根据定理 3.4, $\det(A) \neq 0 \Leftrightarrow \det(B) = \det(P_1) \cdots \det(P_s) \det(Q_1) \cdots \det(Q_t) \det(A) \neq 0 \Leftrightarrow r = n$. 因此, A 可逆 $\Leftrightarrow \det(A) \neq 0$.

由 $\det(A) \det(A^{-1}) = \det(AA^{-1}) = 1$ 可知 $\det(A^{-1}) = \frac{1}{\det(A)}$. □

定理 3.4 可以推广到 A, B 不是方阵的情形.

定理 3.6 (Binet^[6]-Cauchy 公式). 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵.

$$\det(AB) = \begin{cases} \sum_{1 \leq i_1 < i_2 < \cdots < i_m \leq n} \det \left(A \begin{bmatrix} 1 & 2 & \cdots & m \\ i_1 & i_2 & \cdots & i_m \end{bmatrix} \right) \det \left(B \begin{bmatrix} i_1 & i_2 & \cdots & i_m \\ 1 & 2 & \cdots & m \end{bmatrix} \right), & m \leq n; \\ 0, & m > n. \end{cases}$$

^[6]Jacques Philippe Marie Binet, 1786–1856, 法国数学家、物理学家.

证明. 当 $m > n$ 时, 设方阵 $\tilde{A} = \begin{pmatrix} A & O \end{pmatrix}$, $\tilde{B} = \begin{pmatrix} B \\ O \end{pmatrix}$, 得

$$\det(AB) = \det(\tilde{A}\tilde{B}) = \det(\tilde{A})\det(\tilde{B}) = 0.$$

当 $m \leq n$ 时, 设 $A = (a_{ij})$, B 的行向量分别为 $\beta_1, \beta_2, \dots, \beta_m$, 则

$$\begin{aligned} \det(AB) &= \det\left(\sum_{j=1}^n a_{1j}\beta_j, \sum_{j=1}^n a_{2j}\beta_j, \dots, \sum_{j=1}^n a_{mj}\beta_j\right) \\ &= \sum_{1 \leq j_1, j_2, \dots, j_m \leq n} a_{1j_1} a_{2j_2} \cdots a_{mj_m} \det(\beta_{j_1}, \beta_{j_2}, \dots, \beta_{j_m}) \\ &= \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_m \leq n \\ (j_1, j_2, \dots, j_m) \text{ 是 } (i_1, i_2, \dots, i_m) \text{ 的排列}}} a_{1j_1} a_{2j_2} \cdots a_{mj_m} (-1)^{\tau(j_1, j_2, \dots, j_m)} \det(\beta_{i_1}, \beta_{i_2}, \dots, \beta_{i_m}) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} \det\left(A \begin{bmatrix} 1 & 2 & \dots & m \\ i_1 & i_2 & \dots & i_m \end{bmatrix}\right) \det\left(B \begin{bmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{bmatrix}\right). \quad \square \end{aligned}$$

例 3.8. 方阵 $V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$ 称为 **Vandermonde**^[7] 方阵. 计算 $\det(V)$ 和 V^{-1} .

解答. 首先计算 $\det(V)$.

解法一. 对 V 作初等列变换, 从第 n 列开始, 每一列减去前一列的 x_1 倍, 得

$$\begin{aligned} \det(V) &= \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix} \\ &= (x_2 - x_1) \cdots (x_n - x_1) \begin{vmatrix} 1 & x_2 & \cdots & x_2^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} \end{vmatrix}. \end{aligned}$$

利用数学归纳法, $\det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

解法二. 视 $\det(V)$ 为 x_n 的多项式 $f(x_n)$. 由 $f(x_1) = \cdots = f(x_{n-1}) = 0$ 可得

$$f(x) = c(x - x_1) \cdots (x - x_{n-1}).$$

再由行列式的完全展开式, 可得 $f(x)$ 的最高次项系数 c 是 $n-1$ 阶 Vandermonde 方阵 $\begin{pmatrix} x_i^{j-1} \end{pmatrix}$ 的行列式. 同上, $\det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

^[7]Alexandre-Théophile Vandermonde, 1735–1796, 法国音乐家、数学家、化学家.

解法三. 设 $f_1(t) = 1$, $f_j(t) = \prod_{i=1}^{j-1} (t - x_i) = \sum_{i=1}^j a_{ij} t^{i-1}$, $2 \leq j \leq n$. 对等式

$$\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} f_1(x_1) & 0 & \cdots & 0 \\ f_1(x_2) & f_2(x_2) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{pmatrix}$$

两边分别取行列式, 得 $\det(V) = f_1(x_1)f_2(x_2)\cdots f_n(x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

下面计算 $W = (w_{ij}) = V^{-1}$. 根据定理 3.5, V 是可逆方阵当且仅当 $\det(V) \neq 0$, 即 x_1, \dots, x_n 两两不同. 设 $g_j(t) = \sum_{i=1}^n w_{ij} t^{i-1}$. 由 $VW = I$ 可得 $g_j(x_i) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. 因此,

$$g_j(t) = \prod_{k \neq j} \frac{t - x_k}{x_j - x_k}, \quad w_{ij} = \frac{(-1)^{n-i} \sigma_{n-i}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)}{\prod_{k \neq j} (x_j - x_k)},$$

其中 $\sigma_{n-i}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ 是 $n-1$ 个变元的 $n-i$ 次基本对称多项式. □

在例 3.8 中, Vandermonde 方阵 V 被分解为 $V = LU^{-1}$ 的形式, 其中 L 是下三角方阵, U 是单位上三角方阵. 记号 δ_{ij} 可以看作是关于 i, j 的二元函数, 通常称为 **Kronecker**^[8] δ 函数.

例 3.9. 方阵 $C = \begin{pmatrix} \frac{1}{s_1 - t_1} & \frac{1}{s_1 - t_2} & \cdots & \frac{1}{s_1 - t_n} \\ \frac{1}{s_2 - t_1} & \frac{1}{s_2 - t_2} & \cdots & \frac{1}{s_2 - t_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{s_n - t_1} & \frac{1}{s_n - t_2} & \cdots & \frac{1}{s_n - t_n} \end{pmatrix}$ 称为 **Cauchy 方阵**. 计算 $\det(C)$ 和 C^{-1} .

解答. 设 $p(x) = \prod_{i=1}^n (x - t_i)$, $q_j(x) = \frac{p(x)}{x - t_j} = \sum_{i=1}^n a_{ij} x^{i-1}$. 由矩阵乘积分解

$$C = \begin{pmatrix} \frac{1}{p(s_1)} & & & \\ & \frac{1}{p(s_2)} & & \\ & & \ddots & \\ & & & \frac{1}{p(s_n)} \end{pmatrix} \begin{pmatrix} 1 & s_1 & \cdots & s_1^{n-1} \\ 1 & s_2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & s_n & \cdots & s_n^{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

和

$$\begin{pmatrix} q_1(t_1) & & & \\ & q_2(t_2) & & \\ & & \ddots & \\ & & & q_n(t_n) \end{pmatrix} = \begin{pmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ 1 & t_2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

^[8]Leopold Kronecker, 1823–1891, 德国数学家.

可得

$$C = \begin{pmatrix} \frac{1}{p(s_1)} & & & \\ & \frac{1}{p(s_2)} & & \\ & & \ddots & \\ & & & \frac{1}{p(s_n)} \end{pmatrix} \begin{pmatrix} 1 & s_1 & \cdots & s_1^{n-1} \\ 1 & s_2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & s_n & \cdots & s_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ 1 & t_2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{pmatrix}^{-1} \begin{pmatrix} q_1(t_1) & & & \\ & q_2(t_2) & & \\ & & \ddots & \\ & & & q_n(t_n) \end{pmatrix},$$

$$C^{-1} = \begin{pmatrix} \frac{1}{q_1(t_1)} & & & \\ & \frac{1}{q_2(t_2)} & & \\ & & \ddots & \\ & & & \frac{1}{q_n(t_n)} \end{pmatrix} \begin{pmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ 1 & t_2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & s_1 & \cdots & s_1^{n-1} \\ 1 & s_2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & s_n & \cdots & s_n^{n-1} \end{pmatrix}^{-1} \begin{pmatrix} p(s_1) & & & \\ & p(s_2) & & \\ & & \ddots & \\ & & & p(s_n) \end{pmatrix}.$$

因此,

$$\det(C) = \prod_{i=1}^n \frac{q_i(t_i)}{p(s_i)} \prod_{1 \leq i < j \leq n} \frac{s_j - s_i}{t_j - t_i} = \prod_{1 \leq i < j \leq n} (s_j - s_i)(t_i - t_j) \prod_{i,j=1}^n (s_i - t_j)^{-1}. \quad \square$$

例 3.9 表明, 尽管形式不同, Cauchy 方阵与 Vandermonde 方阵之间存在紧密的联系, 而多项式恰是架设这种联系的桥梁.

例 3.10. 复数方阵 $C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ c_1 & \cdots & c_{n-1} & c_0 \end{pmatrix}$ 称为循环方阵. 计算 $\det(C)$ 和 C^{-1} .

解答. 注意到 $C = f(Z)$, 其中 $f(t) = \sum_{i=0}^{n-1} c_i t^i$, $Z = \begin{pmatrix} & & & I_{n-1} \\ 1 & & & \end{pmatrix}$ 满足 $Z^n = I$. 设 n 阶复数方阵 $\Omega = (\omega^{(i-1)(j-1)})$, 其中 $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ 是 n 次单位根. 容易验证 $\Omega \bar{\Omega} = nI$. 故 Ω 是可逆方阵, $\Omega^{-1} = \frac{1}{n} \bar{\Omega}$. 再由 $Z\Omega = \Omega \text{diag}(1, \omega, \dots, \omega^{n-1})$, 得

$$Z = \Omega \text{diag}(1, \omega, \dots, \omega^{n-1}) \Omega^{-1}, \quad C = f(Z) = \Omega \text{diag}(f(1), f(\omega), \dots, f(\omega^{n-1})) \Omega^{-1}.$$

因此, $\det(C) = \prod_{i=0}^{n-1} f(\omega^i)$, $C^{-1} = \Omega \text{diag}\left(\frac{1}{f(1)}, \frac{1}{f(\omega)}, \dots, \frac{1}{f(\omega^{n-1})}\right) \Omega^{-1} = \left(\frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega^{(i-j)k}}{f(\omega^k)}\right) \Omega^{-1}$. \square

在例 3.10 中, 循环方阵 C 被分解为 $\Omega D \Omega^{-1}$ 的形式 (称为相似对角化, 详见第五章), 其中 D 是对角方阵, Ω 是 Vandermonde 方阵, $\frac{1}{\sqrt{n}} \Omega$ 是酉方阵 (详见第六章).

在利用矩阵乘积分解 $A = P_1 P_2 \cdots P_k$ 计算某个方阵 A 的行列式 (或逆矩阵) 时, 通常需要这些 P_i 的行列式 (或逆矩阵) 容易求得. 初等方阵和“分块”初等方阵是常见的选择.

例 3.11. 当 $A_1 A_3 = A_3 A_1$ 时, $\det \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \det(A_1 A_4 - A_3 A_2)$.

证明. 当 $\det(A_1) \neq 0$ 时, 由 Schur 公式可得

$$\det \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \det(A_1) \det(A_4 - A_3 A_1^{-1} A_2) = \det(A_1 (A_4 - A_3 A_1^{-1} A_2)) = \det(A_1 A_4 - A_3 A_2).$$

当 $\det(A_1) = 0$ 时, 设 x 是变元, 则有 $(xI + A_1)A_3 = A_3(xI + A_1)$, 并且多项式 $\det(xI + A_1) \neq 0$. 根据前述结论, $\det \begin{pmatrix} xI + A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \det \left((xI + A_1)A_4 - A_3A_2 \right)$. 等式两边都是多项式, 令 $x = 0$, 得 $\det \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \det(A_1A_4 - A_3A_2)$. \square

例 3.12. 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵, x 是变元, 则有

$$\det(xI - AB) = x^{m-n} \det(xI - BA).$$

证明. 由

$$\begin{pmatrix} I & A \\ O & I \end{pmatrix} \begin{pmatrix} xI - AB & O \\ B & I \end{pmatrix} = \begin{pmatrix} xI & A \\ B & I \end{pmatrix} = \begin{pmatrix} I & O \\ x^{-1}B & I \end{pmatrix} \begin{pmatrix} xI & A \\ O & I - x^{-1}BA \end{pmatrix}$$

可得

$$\det(xI - AB) = x^m \det(I - x^{-1}BA) = x^{m-n} \det(xI - BA). \quad \square$$

习题

1. 设 $A \in \mathbb{C}^{m \times n}$. 利用 Binet-Cauchy 公式证明 $\det(AA^H) \geq 0$.

2. 设 A, B, C 分别是 $k \times m, m \times n, n \times k$ 矩阵, $k \leq \min(m, n)$. 证明:

$$\det(ABC) = \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_k \leq m \\ 1 \leq j_1 < j_2 < \dots < j_k \leq n}} \det \left(A \begin{bmatrix} 1 & 2 & \dots & k \\ i_1 & i_2 & \dots & i_k \end{bmatrix} \right) \det \left(B \begin{bmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{bmatrix} \right) \det \left(C \begin{bmatrix} j_1 & j_2 & \dots & j_k \\ 1 & 2 & \dots & k \end{bmatrix} \right).$$

3. 计算下列行列式.

$$(1) \begin{vmatrix} 1 & \cos \theta_1 & \dots & \cos(n-1)\theta_1 \\ 1 & \cos \theta_2 & \dots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \dots & \vdots \\ 1 & \cos \theta_n & \dots & \cos(n-1)\theta_n \end{vmatrix}$$

$$(3) \begin{vmatrix} \sin \theta_1 & \sin 2\theta_1 & \dots & \sin n\theta_1 \\ \sin \theta_2 & \sin 2\theta_2 & \dots & \sin n\theta_2 \\ \vdots & \vdots & \dots & \vdots \\ \sin \theta_n & \sin 2\theta_n & \dots & \sin n\theta_n \end{vmatrix}$$

$$(5) \begin{vmatrix} 1^2 & 2^2 & \dots & n^2 \\ 2^2 & 3^2 & \dots & (n+1)^2 \\ \vdots & \vdots & \dots & \vdots \\ n^2 & (n+1)^2 & \dots & (2n-1)^2 \end{vmatrix}$$

$$(7) \begin{vmatrix} c_0 & c_1 & \dots & c_{n-1} \\ -c_{n-1} & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ -c_1 & \dots & -c_{n-1} & c_0 \end{vmatrix}$$

$$(2) \begin{vmatrix} \cos \frac{1}{2}\theta_1 & \cos \frac{3}{2}\theta_1 & \dots & \cos \frac{2n-1}{2}\theta_1 \\ \cos \frac{1}{2}\theta_2 & \cos \frac{3}{2}\theta_2 & \dots & \cos \frac{2n-1}{2}\theta_2 \\ \vdots & \vdots & \dots & \vdots \\ \cos \frac{1}{2}\theta_n & \cos \frac{3}{2}\theta_n & \dots & \cos \frac{2n-1}{2}\theta_n \end{vmatrix}$$

$$(4) \begin{vmatrix} \sin \frac{1}{2}\theta_1 & \sin \frac{3}{2}\theta_1 & \dots & \sin \frac{2n-1}{2}\theta_1 \\ \sin \frac{1}{2}\theta_2 & \sin \frac{3}{2}\theta_2 & \dots & \sin \frac{2n-1}{2}\theta_2 \\ \vdots & \vdots & \dots & \vdots \\ \sin \frac{1}{2}\theta_n & \sin \frac{3}{2}\theta_n & \dots & \sin \frac{2n-1}{2}\theta_n \end{vmatrix}$$

$$(6) \begin{vmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & \vdots & \dots & \vdots \\ s_{n-1} & s_n & \dots & s_{2n-1} \end{vmatrix}, \quad s_k = \sum_{i=1}^n x_i^k.$$

$$(8) \begin{vmatrix} C_{p+1}^{q+1} & C_{p+1}^{q+2} & \dots & C_{p+1}^{q+n} \\ C_{p+2}^{q+1} & C_{p+2}^{q+2} & \dots & C_{p+2}^{q+n} \\ \vdots & \vdots & \dots & \vdots \\ C_{p+n}^{q+1} & C_{p+n}^{q+2} & \dots & C_{p+n}^{q+n} \end{vmatrix}, \quad C_{p+i}^{q+j} \text{ 是组合数.}$$

$$4. \text{ [9]证明: } \begin{vmatrix} 1+a^2-b^2-c^2 & 2(ab+c) & 2(ca-b) \\ 2(ab-c) & 1+b^2-c^2-a^2 & 2(bc+a) \\ 2(ca+b) & 2(bc-a) & 1+c^2-a^2-b^2 \end{vmatrix} = (1+a^2+b^2+c^2)^3.$$

5. 设 n 阶复数方阵 Ω 如例 3.10 所述. 证明: $\det(\Omega) = n^{\frac{n}{2}} e^{i\theta}$, 其中 $\theta = \frac{(n-1)(3n-2)}{4}\pi$.

6. 计算下列方阵的行列式和逆矩阵.

$$(1) \begin{pmatrix} 1+a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & 1+a_nb_n \end{pmatrix} \quad (2) \begin{pmatrix} 1+a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & 1+a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n+b_1 & a_n+b_2 & \cdots & 1+a_n+b_n \end{pmatrix}$$

7. 举例满足条件的方阵 A, B .

$$(1) \det(A-B) = 0 \text{ 且 } \det(A^2-B^2) = 1 \quad (2) \det(A-B) = 1 \text{ 且 } \det(A^2-B^2) = 0$$

8. 设 A, B 是 n 阶复数方阵. 证明:

$$(1) \det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det(A+B) \det(A-B) \quad (2) \det \begin{pmatrix} A & B \\ -B & A \end{pmatrix} = \det(A+iB) \det(A-iB)$$

9. 证明: 存在以 $\{a_{ij} \mid 1 \leq i < j \leq n\}$ 为变元的整数系数多项式 P , 使得

$$\det \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ -a_{1n} & \cdots & -a_{n-1,n} & 0 \end{pmatrix} = P^2.$$

关于 P 的表达式可参阅文献 [2] 《典型群》第七章 §2 定理 2.

10. 设 $U = (i^{j-1})$ 和 $V = (a_i^{j-1})$ 都是 n 阶整数方阵. 证明: $\det(U)$ 整除 $\det(V)$.

11. [10] 设 S_1, \dots, S_{2^n-1} 是 $\{1, 2, \dots, n\}$ 的所有非空子集, $A = (a_{ij})$ 是 $2^n - 1$ 阶整数方阵, 其中

$$a_{ij} = \begin{cases} 1, & S_i \cap S_j \neq \emptyset; \\ 0, & S_i \cap S_j = \emptyset. \end{cases}$$

证明: 当 $n \geq 2$ 时, $\det(A) = -1$.

12. 设 x_1, \dots, x_n 两两不同, $f_i(y) = \prod_{j \neq i} \frac{y-x_j}{x_i-x_j}$, $g_i(y) = \prod_{j \neq i} \frac{1-x_j y}{x_i-x_j}$, $i = 1, 2, \dots, n$. 证明:

$$(1) \begin{pmatrix} f_1(y) & f_2(y) & \cdots & f_n(y) \end{pmatrix} = \begin{pmatrix} 1 & y & \cdots & y^{n-1} \end{pmatrix} V^{-1}, \text{ 其中 } V = \begin{pmatrix} x_i^{j-1} \end{pmatrix} \text{ 同例 3.8.}$$

$$(2) \begin{pmatrix} g_1(y) & g_2(y) & \cdots & g_n(y) \end{pmatrix} = \begin{pmatrix} y^{n-1} & \cdots & y & 1 \end{pmatrix} V^{-1}.$$

[9] 1941 年 Putnam 数学竞赛 A7 第 (i) 问.

[10] 2018 年 Putnam 数学竞赛 A2.

§3.3 Laplace 展开

定理 3.7. 设 $A = (a_{ij})$ 是 n 阶方阵, B_j 是删除 A 的第 1 行和第 j 列后所得的 $n-1$ 阶方阵, 则有

$$\det(A) = \sum_{j=1}^n (-1)^{j-1} a_{1j} \det(B_j). \quad (3.2)$$

证明. 注意到 $\tau(j, j_2, \dots, j_n) = j-1 + \tau(j_2, \dots, j_n)$. 因此,

$$\det(A) = \sum_{(j_1, j_2, \dots, j_n) \in S_n} (-1)^{\tau(j, j_2, \dots, j_n)} a_{1j} a_{2j_2} \cdots a_{nj_n} = \sum_{j=1}^n (-1)^{j-1} a_{1j} \det(B_j). \quad \square$$

(3.2) 式称为 $\det(A)$ 的 Laplace^[14] 展开. 许多教科书首先定义 1 阶行列式 $\det(a) = a$, 然后通过 (3.2) 式归纳定义 $n \geq 2$ 阶行列式. 为了说明如此归纳定义的行列式与定义 3.1 是相同的, 我们只需证明它的完全展开式就是 (3.1) 式. 可使用数学归纳法证明, 与定理 3.7 的证明类似.

通过 Laplace 展开, 一个高阶行列式的计算被转化为多个低阶行列式的计算. 这有利于使用并行算法来计算大规模的行列式. 这种方法对于含有很多零元素的稀疏矩阵尤其有效.

例 3.13. 利用 Laplace 展开计算行列式 $\Delta = \begin{vmatrix} 1 & 2 & 0 & 0 \\ 3 & 1 & 2 & 0 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 3 & 1 \end{vmatrix}$.

解答. $\Delta = \begin{vmatrix} 1 & 2 & 0 \\ 3 & 1 & 2 \\ 0 & 3 & 1 \end{vmatrix} - 2 \begin{vmatrix} 3 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} - 2 \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} - 6 \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} + 4 \begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix} = -5 - 6 + 30 = 19. \quad \square$

例 3.14. 计算 n 阶行列式 $\Delta_n = \begin{vmatrix} x & 1 & & \\ 1 & x & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & x \end{vmatrix}$, 其中空白处的元素都是 0.

解答. 易得 $\Delta_1 = x, \Delta_2 = x^2 - 1$. 当 $n \geq 3$ 时, 由 Laplace 展开, 可得递推关系式 $\Delta_n = x\Delta_{n-1} - \Delta_{n-2}$. 由例 2.7, $\Delta_n = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. □

与定理 3.7 类似, 我们有以下常用结论.

定理 3.8. 设 A 是 n 阶方阵, B_{ij} 是删除 A 的第 i 行和第 j 列后所得的 $n-1$ 阶方阵.

- $\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(B_{ij}), \quad \forall i = 1, 2, \dots, n.$

- $\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(B_{ij}), \quad \forall j = 1, 2, \dots, n.$

^[14]Pierre-Simon Laplace, 1749–1827, 法国数学家、物理学家.

证明. 当 $i = 1$ 时, 结论 1 即为定理 3.7. 当 $i \geq 2$ 时, 可经 $i - 1$ 次相邻两行的交换, 把 A 的第 i 行换到第 1 行, 其它行的顺序不变, 所得矩阵记为 \tilde{A} . 由 $\det(\tilde{A}) = \sum_{j=1}^n (-1)^{j-1} a_{ij} \det(B_{ij})$ 可得 $\det(A) = (-1)^{i-1} \det(\tilde{A}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(B_{ij})$. 对 A^T 应用结论 1 即得结论 2. \square

例 3.15. 计算 n 阶行列式 $\Delta = \begin{vmatrix} x & & & c_0 \\ -1 & \ddots & & \vdots \\ & \ddots & x & c_{n-2} \\ & & -1 & x + c_{n-1} \end{vmatrix}$, 其中空白处的元素都是 0.

解答. 对行列式的第 n 列作 Laplace 展开, 得

$$\Delta = x^{n-1}(x + c_{n-1}) + \sum_{i=1}^{n-1} (-1)^{n+i} c_{i-1} x^{i-1} (-1)^{n-i} = x^n + \sum_{i=0}^{n-1} c_i x^i. \quad \square$$

定义 3.3. 设 $A = (a_{ij})_{n \times n}$, (i_1, i_2, \dots, i_n) 和 (j_1, j_2, \dots, j_n) 都是 $(1, 2, \dots, n)$ 的排列, $1 \leq k < n$.

- 子矩阵的行列式简称子式. 主子矩阵的行列式简称主子式. 顺序主子矩阵的行列式简称顺序主子式. $A \begin{bmatrix} i_{k+1} & i_{k+2} & \cdots & i_n \\ j_{k+1} & j_{k+2} & \cdots & j_n \end{bmatrix}$ 称为 $A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix}$ 的余子矩阵. $\det \left(A \begin{bmatrix} i_{k+1} & i_{k+2} & \cdots & i_n \\ j_{k+1} & j_{k+2} & \cdots & j_n \end{bmatrix} \right)$ 称为 $A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix}$ 的余子式.

$$(-1)^{\tau(i_1, i_2, \dots, i_n) + \tau(j_1, j_2, \dots, j_n)} \det \left(A \begin{bmatrix} i_{k+1} & i_{k+2} & \cdots & i_n \\ j_{k+1} & j_{k+2} & \cdots & j_n \end{bmatrix} \right) \quad (3.3)$$

称为 $A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix}$ 的代数余子式.

注: 代数余子式 (3.3) 仅依赖于排列 (i_1, i_2, \dots, i_k) 和 (j_1, j_2, \dots, j_k) , 与 $(i_{k+1}, i_{k+2}, \dots, i_n)$ 和 $(j_{k+1}, j_{k+2}, \dots, j_n)$ 的排序无关. 特别, 当这四个排列都是升序排列时,

$$(-1)^{\tau(i_1, i_2, \dots, i_n) + \tau(j_1, j_2, \dots, j_n)} = (-1)^{i_1 + \cdots + i_k + j_1 + \cdots + j_k}$$

- 由 a_{ij} 的代数余子式 A_{ij} 构成的 n 阶方阵

$$A^* = (A_{ij})^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

称为 A 的伴随方阵 (adjugate matrix) 或附属方阵 (adjunct matrix).

使用记号 A_{ij} , 定理 3.8 可表述为: (1) $\det(A) = \sum_{j=1}^n a_{ij} A_{ij}, \forall i$. (2) $\det(A) = \sum_{i=1}^n a_{ij} A_{ij}, \forall j$. 一般地, 设 $i \neq j$, 把 A 的第 j 行换成 $(a_{i1}, a_{i2}, \dots, a_{in})$ 后所得方阵记为 \tilde{A} , 则有

$$\sum_{k=1}^n a_{ik} A_{jk} = \det(\tilde{A}) = 0.$$

同理, $\sum_{k=1}^n a_{ki} A_{kj} = 0$. 由此可得

定理 3.9. $AA^* = A^*A = \det(A)I$, 其中 A 是方阵.

结合定理 3.5 和定理 3.9 可得, A 是可逆方阵当且仅当 $\det(A) \neq 0$, 并且

$$A^{-1} = \frac{1}{\det(A)} A^*. \quad (3.4)$$

例 3.16. 计算 n 阶方阵 $A = \begin{pmatrix} x & 1 & & \\ 1 & x & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & x \end{pmatrix}$ 的逆矩阵, 其中空白处的元素都是 0.

解答. $A^{-1} = \left(\frac{A_{ji}}{\Delta_n} \right)$, 其中 $\Delta_n = \det(A)$ 同例 3.14, $A_{ji} = A_{ij} = \begin{cases} (-1)^{i+j} \Delta_{i-1} \Delta_{n-j}, & i \leq j; \\ (-1)^{i+j} \Delta_{j-1} \Delta_{n-i}, & i \geq j. \end{cases}$ \square

定理 3.7 还可以推广为如下更一般的形式.

定理 3.10 (Laplace 展开定理). 设 $A = (a_{ij})$ 是 n 阶方阵, 正整数 $k < n$, $(i_1, i_2, \dots, i_k) \in S_n$ 满足 $i_1 < \dots < i_k$, $i_{k+1} < \dots < i_n$.

$$\det(A) = \sum_{1 \leq j_1 < \dots < j_k \leq n} (-1)^{i_1 + \dots + i_k + j_1 + \dots + j_k} \det \left(A \begin{bmatrix} i_1 & \dots & i_k \\ j_1 & \dots & j_k \end{bmatrix} \right) \det \left(A \begin{bmatrix} i_{k+1} & \dots & i_n \\ j_{k+1} & \dots & j_n \end{bmatrix} \right)$$

其中 (j_{k+1}, \dots, j_n) 是 $\{1, 2, \dots, n\} \setminus \{j_1, \dots, j_k\}$ 的升序排列.

证明. 首先考虑 $(i_1, i_2, \dots, i_k) = (1, 2, \dots, k)$ 的情形. 此时,

$$\det(A) = \sum_{\substack{1 \leq j_1 < j_2 < \dots < j_k \leq n \\ (p_1, p_2, \dots, p_k) \text{ 是 } (j_1, j_2, \dots, j_k) \text{ 的排列} \\ (p_{k+1}, p_{k+2}, \dots, p_n) \text{ 是 } (j_{k+1}, j_{k+2}, \dots, j_n) \text{ 的排列}}} (-1)^{\tau(p_1, p_2, \dots, p_n)} a_{1p_1} a_{2p_2} \dots a_{np_n},$$

其中

$$\begin{aligned} \tau(p_1, p_2, \dots, p_n) &= \tau(p_1, p_2, \dots, p_k) + \tau(p_{k+1}, p_{k+2}, \dots, p_n) + \tau(j_1, j_2, \dots, j_n), \\ \tau(j_1, j_2, \dots, j_n) &= (j_1 - 1) + (j_2 - 2) + \dots + (j_k - k). \end{aligned}$$

由此可得

$$\begin{aligned} \det(A) &= \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} (-1)^{\tau(j_1, j_2, \dots, j_n)} \det \left(A \begin{bmatrix} 1 & 2 & \dots & k \\ j_1 & j_2 & \dots & j_k \end{bmatrix} \right) \det \left(A \begin{bmatrix} k+1 & k+2 & \dots & n \\ j_{k+1} & j_{k+2} & \dots & j_n \end{bmatrix} \right) \\ &= \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} (-1)^{1+2+\dots+k+j_1+j_2+\dots+j_k} \det \left(A \begin{bmatrix} 1 & 2 & \dots & k \\ j_1 & j_2 & \dots & j_k \end{bmatrix} \right) \det \left(A \begin{bmatrix} k+1 & k+2 & \dots & n \\ j_{k+1} & j_{k+2} & \dots & j_n \end{bmatrix} \right). \end{aligned}$$

对于一般情形, 可经 $\tau(i_1, i_2, \dots, i_n) = (i_1 - 1) + (i_2 - 2) + \dots + (i_k - k)$ 次相邻两行的交换, 化为 $(i_1, i_2, \dots, i_k) = (1, 2, \dots, k)$ 情形. 结论仍然成立. \square

下面介绍 n 个方程和 n 个未知数的线性方程组的求解公式, 称为 **Cramer 法则**. Cramer 在 1750 年给出了这个公式. 实际上, Leibniz, Maclaurin 等人在此几十年以前就已经发现了类似结果, 只不过他们的记法不如 Cramer 的记法优越、流传广泛.

定理 3.11 (Cramer's Rule). 若 n 阶方阵 A 满足 $\Delta = \det(A) \neq 0$, 则线性方程组 $Ax = b$ 有唯一解 $x = (\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \dots, \frac{\Delta_n}{\Delta})^T$, 其中 Δ_j 是把 A 的第 j 列换成 b 所得方阵的行列式.

证明. 由 $\Delta \neq 0$ 知 A 可逆, $Ax = b$ 有唯一解 $x = A^{-1}b$. 设 $x = (x_i)$, $b = (b_i)$. 下面计算 x .

解法一. 由 $A^{-1} = \frac{1}{\Delta}A^*$, 得 $x_j = \frac{1}{\Delta} \sum_{i=1}^n A_{ij}b_i$. 再由 Laplace 展开, 得 $\Delta_j = \sum_{i=1}^n b_i A_{ij}$. 故 $x_j = \frac{\Delta_j}{\Delta}$.

解法二. 设 A 的列向量分别是 $\beta_1, \beta_2, \dots, \beta_n$. 把 $b = \sum_{i=1}^n x_i \beta_i$ 代入 $\Delta_j = \det(\dots, \beta_{j-1}, b, \beta_{j+1}, \dots)$, 得 $\Delta_j = x_j \Delta$. 故 $x_j = \frac{\Delta_j}{\Delta}$. \square

例 3.17. 用 Cramer 法则求解例 1.1.

解答. 经计算,

$$\Delta = \begin{vmatrix} 0 & 1 & -1 & 0 \\ 1 & -2 & 0 & 0 \\ 3 & -2 & 0 & 1 \\ 3 & 1 & 1 & 3 \end{vmatrix} = -4, \quad \Delta_1 = \begin{vmatrix} -1 & 1 & -1 & 0 \\ -4 & -2 & 0 & 0 \\ -7 & -2 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{vmatrix} = 8, \quad \Delta_2 = \begin{vmatrix} 0 & -1 & -1 & 0 \\ 1 & -4 & 0 & 0 \\ 3 & -7 & 0 & 1 \\ 3 & 0 & 1 & 3 \end{vmatrix} = -4,$$

$$\Delta_3 = \begin{vmatrix} 0 & 1 & -1 & 0 \\ 1 & -2 & -4 & 0 \\ 3 & -2 & -7 & 1 \\ 3 & 1 & 0 & 3 \end{vmatrix} = -8, \quad \Delta_4 = \begin{vmatrix} 0 & 1 & -1 & -1 \\ 1 & -2 & 0 & -4 \\ 3 & -2 & 0 & -7 \\ 3 & 1 & 1 & 0 \end{vmatrix} = -4.$$

故线性方程组有唯一解 $x = (-2, 1, 2, 1)^T$. \square

利用 Cramer 法则求解 n 阶线性方程组, 需要计算 $n+1$ 个 n 阶行列式. 除非这些行列式容易计算, 否则 Cramer 法则的计算量较大, 不是实用的求解方法.

习题

1. 利用 Laplace 展开定理以及 (3.4) 式计算下列 n 阶方阵 $A = (a_{ij})$ 的行列式和逆矩阵.

$$(1) \begin{pmatrix} x & 1 & & \\ -1 & x & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & x \end{pmatrix} \quad (2) \begin{pmatrix} x & 1 & \cdots & 1 \\ -1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ -1 & \cdots & -1 & x \end{pmatrix} \quad (3) \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & 2 & -1 \\ & & -1 & x \end{pmatrix}$$

$$(4) a_{ij} = \min(i, j) \quad (5) a_{ij} = \max(i, j) \quad (6) a_{ij} = \frac{1}{i+j-1}$$

2. 设 n 阶三对角行列式 $\Delta = \begin{vmatrix} a_1 & b_1 & & \\ c_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n \end{vmatrix}$, 其中 b_0, b_n, c_0, c_n 是任意数. 证明:

$$(1) \Delta = \begin{pmatrix} 1 & 0 \\ -c_0 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ -c_1 & 0 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ -c_2 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ -c_{n-2} & 0 \end{pmatrix} \begin{pmatrix} a_n & b_n \\ -c_{n-1} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

$$(2) \Delta = \begin{pmatrix} 1 & 0 \\ -c_n & 0 \end{pmatrix} \begin{pmatrix} a_n & b_{n-1} \\ -c_{n-1} & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} & b_{n-2} \\ -c_{n-2} & 0 \end{pmatrix} \cdots \begin{pmatrix} a_2 & b_1 \\ -c_2 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_0 \\ -c_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

3. (1) 设 n 阶方阵 $A = (a_{ij})$ 满足 $a_{i1} = 1, \forall i$. 证明: $\sum_{i,j=1}^n A_{ij} = \det(A)$.

(2) 设 n 阶方阵 $A = (a_{ij})$ 满足 $\sum_{j=1}^n a_{ij} = 1, \forall i$. 证明: $\sum_{i,j=1}^n A_{ij} = n \det(A)$.

4. (1) 设 n 阶方阵 $A = (a_{ij})$ 满足 $\sum_{j=1}^n a_{ij} = 0, \forall i$. 证明: A^* 的每列元素相同 (行向量相同).

(2) 设 n 阶方阵 $A = (a_{ij})$ 满足 $\sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji} = 0, \forall i$. 证明: $A^* = \lambda J$,

其中 J 是全一矩阵 (即 J 的元素都是 1), $\lambda = \det(A + \frac{1}{n^2}J)$.

(3) 证明 Cayley 公式 $(nI_n - J)^* = n^{n-2}J$.

5. 设 n 阶方阵 $A = (a_{ij}(x))$ 由可微函数构成. 证明: $\frac{d}{dx} \det(A) = \sum_{i,j=1}^n A_{ij} a'_{ij}(x)$.

6. 设 $A = (a_{ij})$ 和 $B = (x_i y_j)$ 都是 n 阶方阵. 证明:

(1) $\det(A + B) = \det(A) + \sum_{i,j=1}^n x_i y_j A_{ij}$.

$$(2) \begin{vmatrix} a_{11} & \cdots & a_{1n} & x_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & x_n \\ y_1 & \cdots & y_n & z \end{vmatrix} = z \det(A) - \sum_{i,j=1}^n x_i y_j A_{ij}.$$

7. 设 A, B 都是 n 阶方阵, λ 是数. 证明:

$$(AB)^* = B^* A^*, \quad (\lambda A)^* = \lambda^{n-1} A^*, \quad (A^T)^* = (A^*)^T, \quad (A^*)^* = (\det(A))^{n-2} A.$$

8. 设 A 是 m 阶方阵, B 是 n 阶方阵. 证明:

$$\det(A \otimes B) = (\det(A))^n (\det(B))^m, \quad (A \otimes B)^* = (\det(A))^{n-1} (\det(B))^{m-1} A^* \otimes B^*.$$

9. 设 $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}), D = (d_{ij})$ 都是 n 阶上三角方阵, 求 $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

10. 设 n 阶复数方阵 $A = \begin{pmatrix} B \\ C \end{pmatrix}$. 证明: $\det(AA^H) \leq \det(BB^H) \det(CC^H)$.

11. 设 n 阶方阵 $A = (a_{ij})$. 证明:

$$\det \left(A^* \begin{bmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{bmatrix} \right) = (-1)^{\tau(i_1, i_2, \dots, i_n) + \tau(j_1, j_2, \dots, j_n)} \det \left(A \begin{bmatrix} j_{r+1} & j_{r+2} & \cdots & j_n \\ i_{r+1} & i_{r+2} & \cdots & i_n \end{bmatrix} \right) (\det(A))^{r-1}$$

其中 (i_1, i_2, \dots, i_n) 和 (j_1, j_2, \dots, j_n) 是 $(1, 2, \dots, n)$ 的任意排列, $1 \leq r < n$.

12. 设 $\text{Char } \mathbb{F} = 2, A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 是对称方阵, $b = (a_{11}, a_{22}, \dots, a_{nn})^T$. 证明:

(1) 线性方程组 $Ax = b$ 一定有解.

(2) 若 $\det(A) \neq 0$ 且 $b = \mathbf{0}$, 则 A^{-1} 的对角元素都是 0.

(3) 若 $\det(A) \neq 0$ 且 $\prod_{i=1}^n a_{ii} \neq 0$, 则存在置换方阵 P , 使得 $P^T A P$ 的顺序主子式都 $\neq 0$.

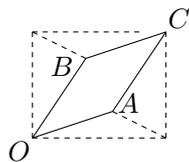
§3.4 行列式与几何

本章的前两节给出了行列式的定义及其性质. 行列式本质上是一个具有特殊性质的多元多项式. 对于任意 $A \in \mathbb{F}^{n \times n}$, 无论使用何种方法计算 $\det(A)$, 它不过是函数 $\det(X) : \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ 在点 $X = A$ 处的取值. 在本节中, 我们从几何的观点看行列式, 赋予其几何解释. 不妨设 $\mathbb{F} = \mathbb{R}$.

在 \mathbb{R}^1 中建立一维坐标系, 向量 \overrightarrow{OA} 的有向长度就是点 A 的坐标.



在 \mathbb{R}^2 中建立平面直角坐标系 Oxy . 设点 A, B 的坐标分别是 $(a_1, a_2), (b_1, b_2)$, $\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB}$,



则 $\overrightarrow{OA}, \overrightarrow{OB}$ 生成的有向平行四边形 $OACB$ 的有向面积

$$S = (a_1 + b_1)(a_2 + b_2) - (a_1 + b_1)a_2 - b_1(a_2 + b_2) = a_1b_2 - a_2b_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

在 \mathbb{R}^3 中建立空间直角坐标系 $Oxyz$. 设点 A, B, C 的坐标分别是 $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$, 则 $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ 生成的有向平行六面体的有向体积

$$V = \overrightarrow{OA} \times \overrightarrow{OB} \cdot \overrightarrow{OC} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} c_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} c_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} c_3,$$

恰是 $\det(\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC})$ 的 Laplace 展开, 并且 $\overrightarrow{OA} \times \overrightarrow{OB}$ 的三个坐标分量是 $\overrightarrow{OA}, \overrightarrow{OB}$ 生成的空间平行四边形分别在坐标平面 Oyz, Ozx, Oxy 上的投影的有向面积.

当 $n \geq 4$ 时, \mathbb{R}^n 中的几何体的长度、面积、体积和“高维体积”又应当如何定义并计算呢?

以 $\alpha \wedge \beta$ 表示 \mathbb{R}^n 中任意两个向量 α 和 β 生成的有向平行四边形, 运算 $\alpha \wedge \beta$ 应满足**二重线性**和**反对称性**, 即对于任意 $\alpha_i, \beta_i \in \mathbb{R}^n$ 和 $\lambda_i \in \mathbb{R}$ 都有

$$\begin{aligned} (\lambda_1 \alpha_1 + \lambda_2 \alpha_2) \wedge \beta &= \lambda_1 (\alpha_1 \wedge \beta) + \lambda_2 (\alpha_2 \wedge \beta), \\ \alpha \wedge (\lambda_1 \beta_1 + \lambda_2 \beta_2) &= \lambda_1 (\alpha \wedge \beta_1) + \lambda_2 (\alpha \wedge \beta_2), \\ \alpha \wedge \beta &= -(\beta \wedge \alpha). \end{aligned}$$

由此可得, 设 $\alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n)$, 则

$$\alpha \wedge \beta = \sum_{1 \leq i, j \leq n} a_i b_j \mathbf{e}_i \wedge \mathbf{e}_j = \sum_{1 \leq i < j \leq n} (a_i b_j - b_i a_j) \mathbf{e}_i \wedge \mathbf{e}_j.$$

也就是说, 任意有向平行四边形 $\alpha \wedge \beta$ 可表示为 $\frac{n(n-1)}{2}$ 个有向单位正方形 $\{\mathbf{e}_i \wedge \mathbf{e}_j \mid 1 \leq i < j \leq n\}$ 的线性组合, 各项系数 $a_i b_j - b_i a_j$ 恰是 $\alpha \wedge \beta$ 在 $\mathbf{e}_i \wedge \mathbf{e}_j$ 平面上的投影的有向面积.

同理, 以 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m$ 表示 \mathbb{R}^n 中任意 m 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 生成的 m 维有向平行多面体, 运算 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m$ 也应满足多重线性和反对称性. 由此可得,

$$\begin{aligned} \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m &= \sum_{1 \leq j_1, j_2, \dots, j_m \leq n} a_{1j_1} a_{2j_2} \cdots a_{mj_m} \mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_m} \\ &= \sum_{\substack{1 \leq i_1 < i_2 < \cdots < i_m \leq n \\ (j_1, j_2, \dots, j_m) \text{ 是 } (i_1, i_2, \dots, i_m) \text{ 的排列}}} a_{1j_1} a_{2j_2} \cdots a_{mj_m} \mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_m} \\ &= \sum_{\substack{1 \leq i_1 < i_2 < \cdots < i_m \leq n \\ (j_1, j_2, \dots, j_m) \text{ 是 } (i_1, i_2, \dots, i_m) \text{ 的排列}}} a_{1j_1} a_{2j_2} \cdots a_{mj_m} (-1)^{\tau(j_1, j_2, \dots, j_m)} \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \\ &= \sum_{1 \leq i_1 < i_2 < \cdots < i_m \leq n} \det \left(A \begin{bmatrix} 1 & 2 & \cdots & m \\ i_1 & i_2 & \cdots & i_m \end{bmatrix} \right) \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \end{aligned}$$

其中 $\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in})$ 是 $A = (a_{ij})_{m \times n}$ 的行向量, $1 \leq i \leq m$. 也就是说, 任意 m 维有向平行多面体 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m$ 可表示为 $\frac{n!}{m!(n-m)!}$ 个 m 维单位方体

$$\{\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \mid 1 \leq i_1 < i_2 < \cdots < i_m \leq n\}$$

的线性组合, 各项系数 $\det \left(A \begin{bmatrix} 1 & 2 & \cdots & m \\ i_1 & i_2 & \cdots & i_m \end{bmatrix} \right)$ 恰是 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m$ 在 $\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m}$ 空间中的投影的 m 维有向体积. 特别, 当 $m = n$ 时,

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n = \det(A) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_n.$$

$\det(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 就是 n 维有向体积 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n$ 与 $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_n$ 的比值.

更一般地, 我们可以在实数域上的线性空间

$$W = \left\{ \sum_{\{i_1, \dots, i_m\} \subset \{1, 2, \dots, n\}} c_{i_1, \dots, i_m} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_m} \mid c_{i_1, \dots, i_m} \in \mathbb{R} \right\}$$

上定义外积 (wedge) 运算 \wedge , 并且满足下列运算律 (其中 $X, Y, Z \in W$, $\lambda, \mu \in \mathbb{R}$)

1. (结合律) $(X \wedge Y) \wedge Z = X \wedge (Y \wedge Z)$.
2. (分配律) $(\lambda X + \mu Y) \wedge Z = \lambda(X \wedge Z) + \mu(Y \wedge Z)$, $X \wedge (\lambda Y + \mu Z) = \lambda(X \wedge Y) + \mu(X \wedge Z)$.
3. (反交换律) 当 j_1, j_2, \dots, j_m 中有重复元素时, $\mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_m} = \mathbf{0}$; 否则

$$\mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_m} = (-1)^{\tau(j_1, j_2, \dots, j_m)} \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m}$$

其中 (i_1, i_2, \dots, i_m) 是 (j_1, j_2, \dots, j_m) 的升序排列.

外积运算 \wedge 可看作是 \mathbb{R}^3 上的向量积运算 \times 的高维推广. 在外积运算下, 线性空间 W 构成外代数 (也称为 **Grassmann 代数**). 把有向平行多面体看作线性空间 W 中的向量, Binet-Cauchy 公式和 Laplace 展开定理可作如下理解:

- 设 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 A 的行向量, $\beta_1, \beta_2, \dots, \beta_m$ 是 B 的列向量, “ \cdot ” 是向量的数量积运算,

$$\text{则 } \det \begin{pmatrix} \alpha_1 \cdot \beta_1 & \alpha_1 \cdot \beta_2 & \cdots & \alpha_1 \cdot \beta_m \\ \alpha_2 \cdot \beta_1 & \alpha_2 \cdot \beta_2 & \cdots & \alpha_2 \cdot \beta_m \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_m \cdot \beta_1 & \alpha_m \cdot \beta_2 & \cdots & \alpha_m \cdot \beta_m \end{pmatrix} = (\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m) \cdot (\beta_1 \wedge \beta_2 \wedge \cdots \wedge \beta_m).$$

- 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 A 的行向量, 则 $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n = (\alpha_1 \wedge \cdots \wedge \alpha_k) \wedge (\alpha_{k+1} \wedge \cdots \wedge \alpha_n)$.

习题

1. 设 \mathbb{R}^2 上三点 A, B, C 的坐标分别是 (x_i, y_i) , $i = 1, 2, 3$. 证明: $\triangle ABC$ 的有向面积

$$S_{\triangle ABC} = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

2. ^[15] 设四边形 $ABCD$ 的顶点共圆, O 是所在平面上任意点. 证明:

$$OA^2 \cdot S_{\triangle BCD} - OB^2 \cdot S_{\triangle ACD} + OC^2 \cdot S_{\triangle ABD} - OD^2 \cdot S_{\triangle ABC} = 0.$$

3. 设 E, F, G, H 和 E', F', G', H' 分别是空间六面体 $ABCD-A'B'C'D'$ 的两个面 $ABCD$ 和 $A'B'C'D'$ 各边的中点. 证明: 六面体 $EFGH-E'F'G'H'$ 的体积是六面体 $ABCD-A'B'C'D'$ 的体积的一半.

4. 证明: \mathbb{R}^3 中的 4 个平面 $a_i x + b_i y + c_i z = 1$ ($i = 1, 2, 3, 4$) 围成的四面体的体积

$$V = \left| \frac{\Delta^3}{6\Delta_1\Delta_2\Delta_3\Delta_4} \right|,$$

$$\text{其中 } \Delta = \begin{vmatrix} a_1 & b_1 & c_1 & 1 \\ a_2 & b_2 & c_2 & 1 \\ a_3 & b_3 & c_3 & 1 \\ a_4 & b_4 & c_4 & 1 \end{vmatrix}, \Delta_i \text{ 是删除 } \Delta \text{ 的第 } i \text{ 行和第 4 列所得子式.}$$

5. 把上题结论推广至 \mathbb{R}^2 中的三角形面积^[16]和 \mathbb{R}^n 中的 n 维单纯形的有向体积, $n \geq 4$.

^[15] 单墀《解析几何的技巧》例 21.6, 中国科学技术大学出版社, 2015 年 6 月第 4 版.

^[16] 1940 年 Putnam 数学竞赛 A8.

第四章 矩阵的相抵

§4.1 矩阵的秩与相抵

根据定理 2.5, 对于任意 $m \times n$ 矩阵 A , 存在初等方阵 $P_1, \dots, P_s, Q_1, \dots, Q_t$ 和整数 r , 使得

$$P_s \cdots P_1 A Q_1 \cdots Q_t = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

由此可得 A 的一个矩阵乘积分解

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$$

其中 P, Q 分别是 m, n 阶的可逆方阵. 由定理 2.9 的证明可知, r 与线性方程组 $Ax = b$ 的解的存在性、唯一性都有着密切的关系.

 思考: r 是否唯一? r 还与 A 的哪些性质有关? 如何有效地求出 r ?

在解答上述问题之前, 首先引入矩阵相抵的概念.

定义 4.1. 设 $A, B \in \mathbb{F}^{m \times n}$. 若存在可逆方阵 $P \in \mathbb{F}^{m \times m}$ 和 $Q \in \mathbb{F}^{n \times n}$ 使得 $A = PBQ$, 则称 A 与 B 在 \mathbb{F} 上相抵.

容易验证, 矩阵的相抵关系具有下列性质, 构成一个等价关系^[1].

1. (自反性) A 与 A 相抵.
2. (对称性) 若 A 与 B 相抵, 则 B 与 A 相抵.
3. (传递性) 若 A 与 B 相抵, B 与 C 相抵, 则 A 与 C 相抵.

对于任意矩阵 $A \in \mathbb{F}^{m \times n}$, 与 A 相抵的矩阵 $B \in \mathbb{F}^{m \times n}$ 的全体构成一个等价类, 称为 A 所在的相抵等价类, 每个相抵等价类中可选出一个矩阵作为该等价类的代表元素, 称为相抵标准形.

定理 4.1. 当 $r \neq s$ 时, $m \times n$ 矩阵 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 与 $\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$ 不相抵.

证明. 反证法. 不妨设 $r > s$, 假设存在可逆矩阵 P, Q 使得 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = P \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} Q$. 把 P, Q 分

块成 $P = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}$, $Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix}$, 其中 P_1, Q_1 分别是 $r \times s, s \times r$ 矩阵, 则有 $P_1 Q_1 = I_r$. 根据定理 3.6, $\det(P_1 Q_1) = 0$, 矛盾. □

^[1]定义在某个集合 S 上的具有自反性、对称性、传递性的二元关系称为 S 上的一个等价关系. 根据等价关系, S 的元素被分为若干个等价类, 同一个等价类中的两个元素等价, 不同等价类中的两个元素不等价.

由定理 2.5 和定理 4.1 可知, 对于任意 A , 存在唯一的 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 与 A 相抵. 因此, 可以把 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 作为相抵标准形. 显然, 使得 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$ 的可逆矩阵 P, Q 有可能不唯一.

定义 4.2. 设 A 的相抵标准形为 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. r 称为 A 的秩, 记作 $\text{rank}(A)$. 显然, $r \leq \min(m, n)$. 若 $r = m$, 则称 A 是行满秩的. 若 $r = n$, 则称 A 是列满秩的. 行满秩与列满秩统称满秩.

求 $\text{rank}(A)$ 的一般方法是利用初等变换把 A 化为标准形或对角形.

例 4.1. 设 $A = \begin{pmatrix} 1 & 2 & 1 & -2 \\ 3 & 1 & 3 & -1 \\ -3 & 0 & -5 & 4 \\ 1 & 0 & 1 & 0 \end{pmatrix}$. 求 $\text{rank}(A)$.

解答. $\begin{pmatrix} 1 & 2 & 1 & -2 \\ 3 & 1 & 3 & -1 \\ -3 & 0 & -5 & 4 \\ 1 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & -2 \\ 0 & -5 & 0 & 5 \\ 0 & 6 & -2 & -2 \\ 0 & -2 & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & -2 \\ 0 & -5 & 0 & 5 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 5 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
 $\rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. 故 $\text{rank}(A) = 3$. \square

例 4.2. 设 n 阶方阵 $A = \begin{pmatrix} 1 & & & 1 \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{pmatrix}$, 其中空白处的元素都是 0. 求 $\text{rank}(A)$.

解答. $\begin{pmatrix} 1 & & & 1 \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{pmatrix} \xrightarrow[\text{每行减去前一}]{\text{从第 2 行起}} \begin{pmatrix} 1 & & & 1 \\ & \ddots & & \vdots \\ & & 1 & (-1)^{n-2} \\ & & & 1 + (-1)^{n-1} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 1 - (-1)^n \end{pmatrix}$.
 因此, 当 n 是奇数时, $\text{rank}(A) = n$; 当 n 是偶数时, $\text{rank}(A) = n - 1$. \square

关于矩阵的秩, 我们有以下常用结论.

定理 4.2. 设 $A \in \mathbb{F}^{n \times n}$. A 是可逆方阵当且仅当 $\text{rank}(A) = n$.

定理 4.3. $\text{rank}(A) = \text{rank}(A^T)$. 当 A 是复数矩阵时, $\text{rank}(A) = \text{rank}(A^H)$.

证明. 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 是可逆方阵. 由 $A^T = Q^T \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^T$ 可得 $\text{rank}(A^T) = r$.

当 A 是复数矩阵时, $A^H = Q^H \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^H$, 也有 $\text{rank}(A^H) = r$. \square

定理 4.4. $\text{rank} \begin{pmatrix} A & O \\ O & B \end{pmatrix} = \text{rank}(A) + \text{rank}(B)$.

证明. 设 $A = P_1 \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q_1$, $B = P_2 \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} Q_2$, 其中 P_1, P_2, Q_1, Q_2 是可逆方阵. 由

$$\begin{aligned} \begin{pmatrix} A & O \\ O & B \end{pmatrix} &= \begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix} \begin{pmatrix} I_r & O \\ O & I_s \end{pmatrix} \begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} \\ &= \begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix} \begin{pmatrix} I & O & O \\ O & I & O \\ O & O & I \end{pmatrix} \begin{pmatrix} I_r & I_s & O \\ O & O & O \end{pmatrix} \begin{pmatrix} I & O & O \\ O & I & O \\ O & O & I \end{pmatrix} \begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix}, \end{aligned}$$

可得 $\text{rank} \begin{pmatrix} A & O \\ O & B \end{pmatrix} = r + s$. □

例 4.3. 设 A 是 $m \times n$ 矩阵, $A \begin{bmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{bmatrix}$ 是 A 的阶数最大的可逆子矩阵. 不妨设

$$(i_1, i_2, \cdots, i_r) = (j_1, j_2, \cdots, j_r) = (1, 2, \cdots, r).$$

作矩阵分块 $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, 其中 A_1 是 r 阶可逆方阵, 则

$$A = \begin{pmatrix} I & O \\ A_3 A_1^{-1} & I \end{pmatrix} \begin{pmatrix} A_1 & O \\ O & S \end{pmatrix} \begin{pmatrix} I & A_1^{-1} A_2 \\ O & I \end{pmatrix}.$$

若 S 的 (p, q) 位置元素 $s_{pq} \neq 0$, 则

$$A \begin{bmatrix} 1 & 2 & \cdots & r & r+p \\ 1 & 2 & \cdots & r & r+q \end{bmatrix} = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{e}_p^T A_3 A_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & s_{pq} \end{pmatrix} \begin{pmatrix} I & A_1^{-1} A_2 \mathbf{e}_q \\ \mathbf{0} & 1 \end{pmatrix}$$

是 $r+1$ 阶可逆方阵, 与 r 的最大性矛盾. 故 $S = O$, 从而 $\text{rank}(A) = \text{rank}(A_1) = r$.

上例实际上给出了矩阵的秩的另一个常见的定义.

定义 4.3. 矩阵 A 的可逆子矩阵 (或非零子式) 的最大阶数称为 A 的秩.

在第八章 §8.3 中还将通过矩阵的行向量组或列向量组给出矩阵的秩的其它等价定义.

定理 4.5. $\text{rank} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \geq \text{rank}(A)$.

证明. 定义 4.3 的推论. □

定理 4.6. $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$.

证明. 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 则 $AB = P \begin{pmatrix} C \\ O \end{pmatrix}$, 其中 P, Q 是可逆方阵, C 是 QB 的前 r 行构成的子矩阵. 一方面, $\text{rank}(C) \leq C$ 的行数 $r = \text{rank}(A)$. 另一方面, $\text{rank}(C) \leq \text{rank}(QB) = \text{rank}(B)$.

综上, $\text{rank}(AB) = \text{rank} \begin{pmatrix} C \\ O \end{pmatrix} = \text{rank}(C) \leq \min(\text{rank}(A), \text{rank}(B))$. □

定理 4.7. $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} \geq \text{rank}(A) + \text{rank}(D)$.

证明. 设 A_1 是 A 的 $\text{rank}(A)$ 阶可逆子矩阵, D_1 是 D 的 $\text{rank}(D)$ 阶可逆子矩阵, 则 $\begin{pmatrix} A & B \\ O & D \end{pmatrix}$ 的 $\text{rank}(A) + \text{rank}(D)$ 阶子矩阵 $\begin{pmatrix} A_1 & B_1 \\ O & D_1 \end{pmatrix}$ 也可逆. 故 $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} \geq \text{rank}(A) + \text{rank}(D)$. \square

例 4.4. 设 $A = \begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix} \in \mathbb{F}^{n \times n}$. 求 $\text{rank}(A)$.

解答. $A = (a-1)I + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}$. 由例 3.12 可得 $\det(A) = (a-1)^{n-1}(a-1+n)$.

- 当 $a \neq 1$ 且 $a \neq 1-n$ 时, $\det(A) \neq 0$. 故 $\text{rank}(A) = n$.
- 当 $a = 1$ 时, 显然有 $\text{rank}(A) = 1$.
- 当 $a = 1-n$ 时, $\det(A[1 \cdots n-1]) = (a-1)^{n-2}(a-2+n) \neq 0$. 故 $\text{rank}(A) = n-1$. \square

习题

1. 计算下列矩阵的秩.

$$\begin{array}{lll}
 (1) \begin{pmatrix} 2 & 5 & 3 & 2 & 0 \\ 2 & 3 & 1 & 2 & 3 \\ 2 & 4 & 4 & 7 & 7 \\ 4 & 9 & 7 & 9 & 7 \end{pmatrix} & (2) \begin{pmatrix} 4 & 2 & 0 & 1 & 1 \\ 1 & 2 & 5 & 3 & 4 \\ 9 & 6 & 5 & 5 & 2 \\ 16 & 14 & 20 & 15 & 18 \end{pmatrix} & (3) \begin{pmatrix} 1 & -2 & -4 & 2 & 5 \\ 1 & -2 & -2 & 2 & 3 \\ 1 & -2 & -10 & 2 & 11 \\ -1 & 2 & -3 & -2 & 2 \end{pmatrix} \\
 (4) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{pmatrix} & (5) \begin{pmatrix} 1^2 & 2^2 & 3^2 & 4^2 & 5^2 \\ 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \\ 3^2 & 4^2 & 5^2 & 6^2 & 7^2 \\ 4^2 & 5^2 & 6^2 & 7^2 & 8^2 \end{pmatrix} & (6) \begin{pmatrix} 1^3 & 2^3 & 3^3 & 4^3 & 5^3 \\ 2^3 & 3^3 & 4^3 & 5^3 & 6^3 \\ 3^3 & 4^3 & 5^3 & 6^3 & 7^3 \\ 4^3 & 5^3 & 6^3 & 7^3 & 8^3 \end{pmatrix}
 \end{array}$$

2. 求 x, y 使得 $\text{rank}(A)$ 最小, 其中 $A =$

$$\begin{array}{lll}
 (1) \begin{pmatrix} x & -1 & -1 & -1 \\ -1 & 2 & -1 & -1 \\ -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & 2 \end{pmatrix} & (2) \begin{pmatrix} x & -1 & -1 & -1 \\ -1 & y & -1 & -1 \\ -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & 2 \end{pmatrix} & (3) \begin{pmatrix} x & y & -1 & -1 \\ y & x & -1 & -1 \\ -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & 2 \end{pmatrix}
 \end{array}$$

3. 证明: 任意秩为 r 的矩阵都可以表示为 r 个秩为 1 的矩阵之和的形式.

4. 证明: 任意非零矩阵 A 都可以表示为 PQ 的形式, 其中 P 是列满秩, Q 是行满秩.

5. 设 $A \in \mathbb{F}^{m \times n}$. 证明:

(1) A 有右逆 $\Leftrightarrow A$ 是行满秩 $\Leftrightarrow xA = \mathbf{0}$ 只有零解 \Leftrightarrow 存在 $B \in \mathbb{F}^{(n-m) \times n}$ 使得 $\begin{pmatrix} A \\ B \end{pmatrix}$ 可逆.

(2) A 有左逆 $\Leftrightarrow A$ 是列满秩 $\Leftrightarrow Ax = \mathbf{0}$ 只有零解 \Leftrightarrow 存在 $B \in \mathbb{F}^{m \times (m-n)}$ 使得 $\begin{pmatrix} A & B \end{pmatrix}$ 可逆.

6. 设 $A, C \in \mathbb{F}^{m \times n}$, $B, D \in \mathbb{F}^{n \times p}$. 证明:

(1) 若 A, B 都是行满秩的, 则 AB 也是行满秩的.

(2) 若 C, D 都是列满秩的, 则 CD 也是列满秩的.

(3) 若 AB 是行满秩的, 则 A 也是行满秩的.

(4) 若 CD 是列满秩的, 则 D 也是列满秩的.

(5) 若 A 与 C 相抵, B 与 D 相抵, 是否一定有 AB 与 CD 相抵? 证明你的结论.

7. 设 $A \in \mathbb{F}^{n \times n}$, A^* 是 A 的伴随方阵.

(1) 证明: $\text{rank}(A^*) = \begin{cases} n, & \text{rank}(A) = n; \\ 1, & \text{rank}(A) = n - 1; \\ 0, & \text{rank}(A) \leq n - 2. \end{cases}$

(2) 求所有 A 满足 $A^* = A^T$.

8. 证明下列关于矩阵的秩的等式和不等式.

(1) $\text{rank}(A + B) \leq \text{rank} \begin{pmatrix} A & B \end{pmatrix}$

(2) $\text{rank} \begin{pmatrix} A & B \end{pmatrix} \leq \text{rank}(A) + \text{rank}(B)$

(3) $\text{rank}(A \otimes B) = \text{rank}(A) \text{rank}(B)$

(4) $\text{rank} \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \text{rank}(A + B) + \text{rank}(A - B)$

(5) $\text{rank}(AC - BD) \leq \text{rank}(A - B) + \text{rank}(C - D)$, 其中 $A, B \in \mathbb{F}^{m \times n}$, $C, D \in \mathbb{F}^{n \times p}$.

(6) $-\frac{n}{2} \leq \text{rank}(AB) - \text{rank}(BA) \leq \frac{m}{2}$, 其中 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times m}$.

9. 设 n 阶复数方阵 $A = (a_{ij})$ 满足 $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$, $\forall i$. A 称为行对角优的. 参考第二章 §2.4 习题 12 和习题 14. 证明:

(1) 若 A 是不可约的, 则 $\text{rank}(A) \geq n - 1$.

(2) 若 $\text{rank}(A) = n - k$ 且 $k \geq 1$, 则存在置换方阵 P 使得 $P^T A P = \text{diag}(A_1, \dots, A_k)$, 其中每个 A_i 是 n_i 阶方阵, $\text{rank}(A_i) = n_i - 1$, 并且 A_i 的每行元素之和都是 0.

§4.2 相抵标准形的应用

本节介绍矩阵的秩和相抵标准形在线性方程组和一些矩阵问题中的应用. 基本思想是通过矩阵相抵, 把关于一般矩阵的问题转化为关于标准形矩阵的问题.

§4.2.1 线性方程组

考虑数域 \mathbb{F} 上的线性方程组 $Ax = b$, 其中 $A \in \mathbb{F}^{m \times n}$, $b \in \mathbb{F}^{m \times 1}$ 已知, $x \in \mathbb{F}^{n \times 1}$ 待求. 设

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$$

其中 $P \in \mathbb{F}^{m \times m}$, $Q \in \mathbb{F}^{n \times n}$ 都是可逆方阵. 作换元 $y = Qx$, 线性方程组 $Ax = b$ 同解变形为

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} y = P^{-1}b = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

线性方程组 $Ax = b$ 的增广矩阵

$$\left(A \quad b \right) = P \begin{pmatrix} I_r & O & \beta_1 \\ O & O & \beta_2 \end{pmatrix} \begin{pmatrix} Q & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}.$$

因此, $Ax = b$ 有解 $\Leftrightarrow \beta_2 = \mathbf{0} \Leftrightarrow \text{rank} \begin{pmatrix} A & b \end{pmatrix} = r$, 解都形如

$$x = Q^{-1}y = Q^{-1} \begin{pmatrix} \beta_1 \\ \mathbf{0} \end{pmatrix} + y_{r+1}Q^{-1}e_{r+1} + \cdots + y_nQ^{-1}e_n,$$

其中 $y_{r+1}, \dots, y_n \in \mathbb{F}$ 为自由参数. $\alpha_0 = Q^{-1} \begin{pmatrix} \beta_1 \\ \mathbf{0} \end{pmatrix}$ 称为 $Ax = b$ 的一个特解. α_0 是 Q^{-1} 的前 r 个列向量的线性组合. Q^{-1} 的后 $n-r$ 个列向量 $\alpha_i = Q^{-1}e_i$ 都是 $Ax = \mathbf{0}$ 的解, $i = r+1, r+2, \dots, n$. $Ax = \mathbf{0}$ 的任意解 x 也都是 $\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_n$ 的线性组合, 解集 $V = \{x \in \mathbb{F}^{n \times 1} \mid Ax = \mathbf{0}\}$ 称为线性方程组 $Ax = \mathbf{0}$ 的解空间, 向量组 $\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_n$ 称为 $Ax = \mathbf{0}$ 的一个基础解系, 也称为解空间 V 的一组基. 综合以上结论, 可得

定理 4.8. 设 $A \in \mathbb{F}^{m \times n}$, $b \in \mathbb{F}^{m \times 1}$.

1. $Ax = b$ 有解 $x \in \mathbb{F}^{n \times 1}$ 当且仅当 $\text{rank} \begin{pmatrix} A & b \end{pmatrix} = \text{rank}(A)$.
2. 设 $Ax = b$ 有特解 α_0 . $Ax = b$ 的解集为

$$\{\alpha_0 + \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k \mid \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}\}$$

其中 $\alpha_1, \alpha_2, \dots, \alpha_k$ 是 $Ax = \mathbf{0}$ 的基础解系, $k = n - \text{rank}(A)$.

3. 特别, 当 $\text{rank}(A) = n$ 时, $Ax = b$ 的解是唯一的.

例 4.5. 对于任意 $A \in \mathbb{C}^{m \times n}$, $\text{rank}(A^H A) = \text{rank}(A)$.

证明. 设 $x \in \mathbb{C}^{n \times 1}$. 若 $Ax = \mathbf{0}$, 显然有 $A^H Ax = \mathbf{0}$. 若 $A^H Ax = \mathbf{0}$, 由 $x^H A^H Ax = 0$ 可得 $Ax = \mathbf{0}$. 线性方程组 $Ax = \mathbf{0}$ 与 $A^H Ax = \mathbf{0}$ 的解集相同, 基础解系的大小也相同. 故 $\text{rank}(A^H A) = \text{rank}(A)$. \square

§4.2.2 广义逆矩阵

定义 4.4. 设 $m \times n$ 矩阵 $A = P \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 都是可逆方阵, 则 $n \times m$ 矩阵

$$B = Q^{-1} \begin{pmatrix} I & O \\ O & O \end{pmatrix} P^{-1}$$

称为 A 的一个广义逆矩阵.

在定义 4.4 中, P, Q 有可能是不唯一的, 故广义逆矩阵也有可能是不唯一的. 例如,

$$\begin{aligned} \begin{pmatrix} I & X \\ O & I \end{pmatrix} \begin{pmatrix} I & O \\ O & O \end{pmatrix} \begin{pmatrix} I & O \\ Y & I \end{pmatrix} &= \begin{pmatrix} I & O \\ O & O \end{pmatrix}, \\ \begin{pmatrix} I & O \\ Y & I \end{pmatrix}^{-1} \begin{pmatrix} I & O \\ O & O \end{pmatrix} \begin{pmatrix} I & X \\ O & I \end{pmatrix}^{-1} &= \begin{pmatrix} I & -X \\ -Y & YX \end{pmatrix}. \end{aligned}$$

由此可知, 对于任意 X, Y , 矩阵 $B = Q^{-1} \begin{pmatrix} I & -X \\ -Y & YX \end{pmatrix} P^{-1}$ 都是 $A = P \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q$ 的广义逆矩阵. A 有唯一的广义逆矩阵当且仅当 A 是可逆方阵.

定理 4.9. 矩阵 B 是 A 的一个广义逆矩阵的充分必要条件是 $ABA = A$ 并且 $BAB = B$.

证明. 由广义逆矩阵的定义可得必要性. 下证充分性. 设

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q, \quad B = Q^{-1} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} P^{-1}.$$

由 $ABA = A$, $BAB = B$ 可得

$$\begin{aligned} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} &= \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}, \\ \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} &= \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}. \end{aligned}$$

从而有 $B_1 = I_r$, $B_4 = B_3 B_2$. 设 $\tilde{P} = P \begin{pmatrix} I_r & -B_2 \\ O & I_{m-r} \end{pmatrix}$, $\tilde{Q} = \begin{pmatrix} I_r & O \\ -B_3 & I_{n-r} \end{pmatrix} Q$, 则有

$$A = \tilde{P} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{Q}, \quad B = \tilde{Q}^{-1} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{P}^{-1}.$$

故 B 是 A 的一个广义逆矩阵. □

下面给出广义逆矩阵在线性方程组问题中的一个应用.

定理 4.10. 设 $A \in \mathbb{F}^{m \times n}$, $b \in \mathbb{F}^{m \times 1}$, $B \in \mathbb{F}^{n \times m}$ 是 A 的一个广义逆矩阵.

1. $Ax = b$ 有解 $x \in \mathbb{F}^{n \times 1}$ 的充分必要条件是 Bb 是 $Ax = b$ 的特解.

2. 当 $Ax = b$ 有解时, $Ax = b$ 的解集为 $\{Bb + (I - BA)y \mid y \in \mathbb{F}^{n \times 1}\}$. 由此可得, $Ax = b$ 有唯一解的充分必要条件是 “ $ABb = b, BA = I$ ”.

证明.

1. 充分性显然. 下证必要性. $Ax = b \Rightarrow A(Bb) = A(BAx) = (ABA)x = Ax = b$.

2. $A(Bb + (I - BA)y) = ABb = b$. 另外, $Ax = b \Rightarrow Bb + (I - BA)x = Bb + (x - Bb) = x$. \square

第六章 §6.4 定义 6.13 还将介绍复数矩阵的 Moore^[3]-Penrose^[4]广义逆矩阵. 每个复数矩阵的 Moore-Penrose 广义逆矩阵是存在且唯一的.

§4.2.3 更多例子

例 4.6. $\text{rank} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \text{rank}(A)$ 的充分必要条件是存在矩阵 X, Y 使得

$$AX = B, \quad YA = C, \quad YAX = D.$$

证明. (充分性) 设 $B = AX, C = YA, D = YAX$. 由

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & O \\ Y & I \end{pmatrix} \begin{pmatrix} A & O \\ O & O \end{pmatrix} \begin{pmatrix} I & X \\ O & I \end{pmatrix},$$

得 $\text{rank} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \text{rank} \begin{pmatrix} A & O \\ O & O \end{pmatrix} = \text{rank}(A)$.

(必要性) 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 都是可逆方阵. 设

$$\begin{pmatrix} P^{-1} & O \\ O & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} Q^{-1} & O \\ O & I \end{pmatrix} = \begin{pmatrix} I_r & O & B_1 \\ O & O & B_2 \\ C_1 & C_2 & D \end{pmatrix},$$

则

$$\begin{pmatrix} I_r & O & O \\ O & I & O \\ -C_1 & O & I \end{pmatrix} \begin{pmatrix} I_r & O & B_1 \\ O & O & B_2 \\ C_1 & C_2 & D \end{pmatrix} \begin{pmatrix} I_r & O & -B_1 \\ O & I & O \\ O & O & I \end{pmatrix} = \begin{pmatrix} I_r & O & O \\ O & O & B_2 \\ O & C_2 & D - C_1 B_1 \end{pmatrix}$$

与 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 相抵. 由 $\text{rank} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = r$, 得 $B_2 = O, C_2 = O, D = C_1 B_1$. 从而,

$$B = P \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 \\ O \end{pmatrix} = A Q^{-1} \begin{pmatrix} B_1 \\ O \end{pmatrix} = AX,$$

$$C = (C_1 \ C_2) Q = (C_1 \ O) \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q = (C_1 \ O) P^{-1} A = YA,$$

$$YAX = (C_1 \ O) P^{-1} A Q^{-1} \begin{pmatrix} B_1 \\ O \end{pmatrix} = (C_1 \ O) \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 \\ O \end{pmatrix} = C_1 B_1 = D. \quad \square$$

^[3]Eliakim Hastings Moore, 1862–1932, 美国数学家.

^[4]Roger Penrose, 1931–, 英国数学家、物理学家.

例 4.7. $\text{rank}(AB) = \text{rank}(A)$ 的充分必要条件是存在矩阵 X 使得 $ABX = A$.

证明. (充分性) 设 $ABX = A$. 根据定理 4.6,

$$\text{rank}(ABX) \leq \text{rank}(AB) \leq \text{rank}(A) = \text{rank}(ABX).$$

故 $\text{rank}(AB) = \text{rank}(A)$.

(必要性) 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 都是可逆方阵. 由 $\text{rank}(AB) = r$ 可得 $AB = P \begin{pmatrix} C \\ O \end{pmatrix}$, 其中 C 是行满秩矩阵. 设矩阵 Y 是 C 的一个右逆, 则 $X = \begin{pmatrix} Y & O \end{pmatrix} Q$ 满足 $ABX = A$. \square

例 4.8. $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = \text{rank}(A) + \text{rank}(D)$ 的充分必要条件是存在矩阵 X, Y 使得 $AX + YD = B$.

证明. (充分性) 设 $AX + YD = B$, 则有

$$\begin{pmatrix} A & B \\ O & D \end{pmatrix} = \begin{pmatrix} I & Y \\ O & I \end{pmatrix} \begin{pmatrix} A & O \\ O & D \end{pmatrix} \begin{pmatrix} I & X \\ O & I \end{pmatrix}.$$

从而, $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = \text{rank} \begin{pmatrix} A & O \\ O & D \end{pmatrix} = \text{rank}(A) + \text{rank}(D)$.

(必要性) 设 $A = P_1 \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q_1$, $D = P_2 \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} Q_2$, 其中 P_1, Q_1, P_2, Q_2 都是可逆方阵. 设

$$\begin{pmatrix} P_1^{-1} & O \\ O & P_2^{-1} \end{pmatrix} \begin{pmatrix} A & B \\ O & D \end{pmatrix} \begin{pmatrix} Q_1^{-1} & O \\ O & Q_2^{-1} \end{pmatrix} = \begin{pmatrix} I_r & O & B_1 & B_2 \\ O & O & B_3 & B_4 \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix},$$

则

$$\begin{pmatrix} I_r & O & -B_1 & O \\ O & I & -B_3 & O \\ O & O & I_s & O \\ O & O & O & I \end{pmatrix} \begin{pmatrix} I_r & O & B_1 & B_2 \\ O & O & B_3 & B_4 \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix} \begin{pmatrix} I_r & O & O & -B_2 \\ O & I & O & O \\ O & O & I_s & O \\ O & O & O & I \end{pmatrix} = \begin{pmatrix} I_r & O & O & O \\ O & O & O & B_4 \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix}$$

与 $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix}$ 相抵. 由 $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = r + s$, 得 $B_4 = O$. 从而,

$$B = P_1 \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} Q_2 = A Q_1^{-1} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} Q_2 + P_1 \begin{pmatrix} O & O \\ B_3 & O \end{pmatrix} P_2^{-1} D. \quad \square$$

例 4.9. 对于任意 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$, $C \in \mathbb{F}^{p \times q}$, 有 **Frobenius**^[5]秩不等式

$$\text{rank}(AB) + \text{rank}(BC) \leq \text{rank}(ABC) + \text{rank}(B).$$

特别, 当 $B = I_n$ 时, 上式成为 **Sylvester** 秩不等式

$$\text{rank}(A) + \text{rank}(C) \leq \text{rank}(AC) + n.$$

^[5]Ferdinand Georg Frobenius, 1849–1917, 德国数学家. Karl Theodor Wilhelm Weierstrass 的学生.

证明. 注意到

$$\begin{pmatrix} I & -A \\ O & I \end{pmatrix} \begin{pmatrix} AB & O \\ B & BC \end{pmatrix} \begin{pmatrix} I & -C \\ O & I \end{pmatrix} = \begin{pmatrix} O & -ABC \\ B & O \end{pmatrix}.$$

根据定理 4.7,

$$\text{rank}(ABC) + \text{rank}(B) = \text{rank} \begin{pmatrix} AB & O \\ B & BC \end{pmatrix} \geq \text{rank}(AB) + \text{rank}(BC). \quad \square$$

根据例 4.8 结论, Frobenius 秩不等式取等号 $\Leftrightarrow \text{rank} \begin{pmatrix} AB & O \\ B & BC \end{pmatrix} = \text{rank}(AB) + \text{rank}(BC)$

\Leftrightarrow 存在矩阵 X, Y 使得 $XAB + BCY = B$.

例 4.10. 求所有 $A \in \mathbb{F}^{n \times n}$ 满足 $A^2 = A$.

解答. 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 都是可逆方阵. 设 $QP = \begin{pmatrix} R_1 & R_2 \\ R_3 & R_4 \end{pmatrix}$, 其中 R_1 是 r 阶方阵, 则有

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QPP^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1}, \quad A^2 = P \begin{pmatrix} R_1^2 & R_1R_2 \\ O & O \end{pmatrix} P^{-1}.$$

由 $A^2 = A$, 得 $R_1 \begin{pmatrix} R_1 & R_2 \end{pmatrix} = \begin{pmatrix} R_1 & R_2 \end{pmatrix}$. 再由 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 是行满秩, 得 $R_1 = I$. 从而,

$$A = P \begin{pmatrix} I_r & R_2 \\ O & O \end{pmatrix} P^{-1} = P \begin{pmatrix} I & -R_2 \\ O & I \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} I & R_2 \\ O & I \end{pmatrix} P^{-1} = \tilde{P} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{P}^{-1},$$

其中 $\tilde{P} = P \begin{pmatrix} I & -R_2 \\ O & I \end{pmatrix}$ 是可逆方阵. 容易看出, 对于任意可逆方阵 $P \in \mathbb{F}^{n \times n}$ 和非负整数 $r \leq n$,

$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^{-1}$ 满足 $A^2 = A$, 此为所求. □

例 4.11. 求所有 $A \in \mathbb{F}^{n \times n}$ 满足 $A^2 = O$.

解答. 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 都是可逆方阵. 设 $QP = \begin{pmatrix} R_1 & R_2 \\ R_3 & R_4 \end{pmatrix}$, 其中 R_1 是 r 阶方阵, 则有

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QPP^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1}, \quad A^2 = P \begin{pmatrix} R_1^2 & R_1R_2 \\ O & O \end{pmatrix} P^{-1}.$$

由 $A^2 = O$, 得 $R_1 \begin{pmatrix} R_1 & R_2 \end{pmatrix} = O$. 再由 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 是行满秩, 得 $R_1 = O$, R_2 是行满秩, $r \leq n - r$. 设 $R_2 = \begin{pmatrix} O & I_r \end{pmatrix} \tilde{Q}$, 其中 \tilde{Q} 是 $n - r$ 阶可逆方阵. 从而,

$$A = P \begin{pmatrix} O & R_2 \\ O & O \end{pmatrix} P^{-1} = P \begin{pmatrix} I & O \\ O & \tilde{Q}^{-1} \end{pmatrix} \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} \begin{pmatrix} I & O \\ O & \tilde{Q} \end{pmatrix} P^{-1} = \tilde{P} \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} \tilde{P}^{-1},$$

其中 $\tilde{P} = P \begin{pmatrix} I & -R_2 \\ O & I \end{pmatrix}$ 是可逆方阵. 容易看出, 对于任意可逆方阵 $P \in \mathbb{F}^{n \times n}$ 和非负整数 $r \leq \frac{n}{2}$,

$A = P \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} P^{-1}$ 满足 $A^2 = O$, 此为所求. □

例 4.12. 设 $H = \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix}$. 若 $A \in \mathbb{F}^{2n \times 2n}$ 满足 $A^T H A = H$, 则 A 称为辛方阵. 有下列结论.

1. 设 $P \in \mathbb{F}^{n \times n}$ 是可逆的, $S \in \mathbb{F}^{n \times n}$ 是对称的, 则 $\begin{pmatrix} P & O \\ O & P^{-T} \end{pmatrix}, \begin{pmatrix} I & S \\ O & I \end{pmatrix}, \begin{pmatrix} I & O \\ S & I \end{pmatrix}$ 是辛方阵.
2. 辛方阵是可逆方阵. 辛方阵的逆矩阵和转置矩阵是辛方阵. 辛方阵的乘积是辛方阵.
3. 任意辛方阵可以表示成结论 1 中的三类方阵的若干乘积.
4. 辛方阵的行列式都是 1.

证明.

1. 容易验证.

2. 设 A, B 都是辛方阵. 由 $B^T A^T H A B = B^T H B = H$, 得 AB 是辛方阵.

由 $\det(H) = 1$ 和 $\det(A^T H A) = (\det(A))^2 = 1$, 得 $\det(A) = 1$ 或 -1 . 故 A 是可逆方阵.

由 $A^T H A = H$, 得 $H = A^{-T} H A^{-1}$. 故 A^{-1} 是辛方阵.

再由 $H^{-1} = (A^{-T} H A^{-1})^{-1} = A H^{-1} A^T$ 和 $H^{-1} = -H$, 得 $H = A H A^T$. 故 A^T 是辛方阵.

3. 设辛方阵 $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, 其中每个 A_i 是 n 阶方阵. 由 $A^T H A = H = A H A^T$ 可得

$$\begin{aligned} A_1^T A_3 &= A_3^T A_1, & A_2^T A_4 &= A_4^T A_2, & A_1^T A_4 - A_3^T A_2 &= I, \\ A_1 A_2^T &= A_2 A_1^T, & A_3 A_4^T &= A_4 A_3^T, & A_1 A_4^T - A_2 A_3^T &= I. \end{aligned} \quad (4.1)$$

• 当 $\det(A_1) \neq 0$ 时,

$$A = \begin{pmatrix} I & O \\ A_3 A_1^{-1} & I \end{pmatrix} \begin{pmatrix} A_1 & \\ & A_4 - A_3 A_1^{-1} A_2 \end{pmatrix} \begin{pmatrix} I & A_1^{-1} A_2 \\ O & I \end{pmatrix}.$$

由 (4.1) 可知, $A_1^{-1} A_2$ 和 $A_3 A_1^{-1}$ 都是对称方阵, 并且

$$A_4 - A_3 A_1^{-1} A_2 = A_1^{-T} (I + A_3^T A_2) - A_3 A_1^{-1} A_2 = A_1^{-T}.$$

• 当 $\det(A_1) = 0$ 时, 首先证明: 存在对称方阵 S 使得 $A_1 + A_2 S$ 是可逆方阵. 设

$$A_1 = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q, \quad A_2 = P \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} Q^{-T}$$

其中 B_1 是 r 阶方阵, P 和 Q 是 n 阶可逆方阵. 由 $A_1 A_2^T = A_2 A_1^T$, 得 $B_3 = O$. 再

由 $\text{rank} \begin{pmatrix} A_1 & A_2 \end{pmatrix} = n$, 得 B_4 是可逆方阵. 设 $S = Q^T \begin{pmatrix} O & O \\ O & I_{n-r} \end{pmatrix} Q$, 则 $A_1 + A_2 S =$

$P \begin{pmatrix} I_r & B_2 \\ O & B_4 \end{pmatrix} Q$ 是可逆方阵. 根据前述情形, 辛方阵 $\tilde{A} = A \begin{pmatrix} I & O \\ S & I \end{pmatrix} = \begin{pmatrix} A_1 + A_2 S & A_2 \\ A_3 + A_4 S & A_4 \end{pmatrix}$

是三类方阵的乘积. 故 $A = \tilde{A} \begin{pmatrix} I & O \\ -S & I \end{pmatrix}$ 也是三类方阵的乘积.

4. 结论 1 中的三类方阵的行列式都是 1. 由结论 3 易知结论 4 成立. \square

在矩阵乘法运算下, 数域 \mathbb{F} 上 $2n$ 阶辛方阵的全体构成 $SL(2n, \mathbb{F})$ 的子群, 称为辛群, 记作 $Sp(2n, \mathbb{F})$. 例 4.12 中的三类方阵是 $Sp(2n, \mathbb{F})$ 的生成元.

习题

1. 对于参数 a , 讨论求解线性方程组 $Ax = b$, 其中

$$A = \begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ a \\ \vdots \\ a^{n-1} \end{pmatrix}.$$

2. 设 $A_1, A_2 \in \mathbb{F}^{m \times n}$, $b_1, b_2 \in \mathbb{F}^{m \times 1}$. 证明: 若线性方程组 $A_1x = b_1$ 与 $A_2x = b_2$ 在 $\mathbb{F}^{n \times 1}$ 中有相同的非空解集, 则存在可逆方阵 $P \in \mathbb{F}^{m \times m}$ 使得 $P \begin{pmatrix} A_1 & b_1 \end{pmatrix} = \begin{pmatrix} A_2 & b_2 \end{pmatrix}$.

3. 设矩阵 A 是行(列)满秩. 证明: B 是 A 的广义逆矩阵 $\Leftrightarrow B$ 是 A 的右(左)逆.

4. 设矩阵 $A = PQ$, 其中 P 是列满秩, Q 是行满秩. 证明:

(1) 若 P^\dagger, Q^\dagger 分别是 P, Q 的广义逆矩阵, 则 $A^\dagger = Q^\dagger P^\dagger$ 是 A 的广义逆矩阵.

(2) A 的每个广义逆矩阵都可以表示为 $Q^\dagger P^\dagger$ 的形式, 其中 P^\dagger, Q^\dagger 分别是 P, Q 的广义逆矩阵.

5. (1) 设 A 是对称方阵. 证明: 对于任意正整数 $k \leq \text{rank}(A)$, A 有 k 或 $k+1$ 阶可逆主子矩阵.

(2) 设 A 是对角元素都是 0 的反对称方阵. 证明: $r = \text{rank}(A)$ 是偶数, 并且对于任意正偶数 $k \leq r$, A 有 k 阶可逆主子矩阵.

6. 设 $\text{Char } \mathbb{F} = 2$, $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 是对称方阵, $a_{11} \neq 0$, $r = \text{rank}(A)$. 证明:

(1) 对于任意正整数 $k \leq r$, A 有 k 阶可逆主子矩阵.

(2) 存在置换方阵 P , 使得 PAP^T 的 $1, 2, \dots, r$ 阶顺序主子式都 $\neq 0$.

7. 设矩阵 $A, B \in \mathbb{F}^{m \times n}$ 满足 $\text{rank}(A+B) = \text{rank}(A) + \text{rank}(B)$. 证明: 存在可逆方阵 P, Q 使得

$$A = P \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q, \quad B = P \begin{pmatrix} O & O \\ O & I \end{pmatrix} Q.$$

8. 设 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times m}$ 满足 $AB = O, BA = O$. 证明: 存在可逆方阵 P, Q 使得

$$A = P \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q, \quad B = Q^{-1} \begin{pmatrix} O & O \\ O & I \end{pmatrix} P^{-1}.$$

9. 设方阵 $A \in \mathbb{F}^{n \times n}$, 整数 $k \geq 2$. 证明:

(1) 若 $A^2 = I$, 则存在可逆方阵 P 使得 $P^{-1}AP$ 形如 $\text{diag}(I, -I)$.

(2) 若 $A^3 = A$, 则存在可逆方阵 P 使得 $P^{-1}AP$ 形如 $\text{diag}(I, -I, O)$.

(3) 若 $\text{rank}(A^{k-1}) = \text{rank}(A^k)$, 则 $\text{rank}(A^k) = \text{rank}(A^{k+1})$.

(4) 若 $\text{rank}(A^k) = \text{rank}(A) = r$, 则存在可逆方阵 $P \in \mathbb{F}^{n \times n}, Q \in \mathbb{F}^{r \times r}$ 使得

$$A = P \begin{pmatrix} Q & O \\ O & O \end{pmatrix} P^{-1}.$$

10. 设方阵 A 满足 $A^m = O$, $A^{m-1} \neq O$, $m \geq 1$. A 称为**幂零方阵**, m 称为 A 的**幂零指数**. 证明:

(1) 存在可逆方阵 P 使得 $P^{-1}AP = (B_{ij})_{m \times m}$ 是准上三角方阵, 其中对角块 B_{ii} 都是零方阵.

(2) 存在可逆方阵 P 使得 $P^{-1}AP = \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-1} \\ & & & O \end{pmatrix}$, 其中对角块都是零方阵, 每个

B_i 形如 $\begin{pmatrix} I & O \end{pmatrix}$, 其余空白处元素都是 0.

(3) 存在可逆方阵 P 使得 $P^{-1}AP = \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-1} \\ & & & O \end{pmatrix}$, 其中对角块都是零方阵, 每个

B_i 形如 $\begin{pmatrix} I \\ O \end{pmatrix}$, 其余空白处元素都是 0.

(4) 存在可逆方阵 P 使得 $P^{-1}AP = \text{diag}(J_1, J_2, \dots, J_s)$, 每个 $J_i = \begin{pmatrix} I_{n_i-1} \\ 0 \end{pmatrix}$ 是 n_i 阶方阵, $m = n_1 \geq n_2 \geq \dots \geq n_s$. $\text{diag}(J_1, J_2, \dots, J_s)$ 称为**Jordan 方阵**, J_i 称为**Jordan 块**.

§4.3 Smith 标准形

在定义 4.1 中, A, B, P, Q 都是数域 \mathbb{F} 上的矩阵. 在某些问题中, 对矩阵 A, B, P, Q 有特殊的要求. 例如, 要求矩阵的元素都是整数, 矩阵的元素都是多项式等. 在这样的限制条件下, 什么样的矩阵可以作为相抵标准形?

定义 4.5. 设 P 是整数方阵. 若 P 是可逆方阵, 并且 P^{-1} 也是整数方阵, 则 P 称为 \mathbb{Z} 上的模方阵. 设 $A, B \in \mathbb{Z}^{m \times n}$. 若存在 \mathbb{Z} 上的模方阵 P, Q 使得 $A = PBQ$, 则称 A 与 B 在 \mathbb{Z} 上模相抵.

容易验证, 整数矩阵的模相抵关系构成一个等价关系, n 阶整数模方阵的全体, 在矩阵乘法运算下构成群.

定理 4.11. 设 P 是整数方阵. P 是 \mathbb{Z} 上的模方阵当且仅当 $\det(P) = \pm 1$.

证明. 若 P 是模方阵, 则 P^{-1} 是整数方阵, $\det(P)$ 和 $\det(P^{-1})$ 是乘积为 1 的整数, 故 $\det(P) = \pm 1$. 另一方面, 若 $\det(P) = \pm 1$, 则 P^* 和 $P^{-1} = \frac{1}{\det(P)}P^*$ 都是整数方阵. \square

例 4.13. 任意 2 阶整数模方阵可表示为一系列 S_{12} 和 $T_{12}(\pm 1)$ 的乘积.

证明. 设 G 是一系列 S_{12} 和 $T_{12}(\pm 1)$ 的乘积构成的集合. 先证明所有 2 阶初等模方阵都在 G 中.

- $T_{12}(k) = (T_{12}(1))^k \in G$, $T_{21}(k) = S_{12}T_{12}(k)S_{12} \in G$, $\forall k \in \mathbb{Z}$.
- $P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in G$.
- $D_1(-1) = PS_{12} \in G$, $D_2(-1) = S_{12}P \in G$.

对于任意 2 阶整数模方阵 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 总可以通过如下初等变换, 把 M 变成单位方阵, 从而 $M \in G$.

- 若 $c = 0$, 则 $ad = \pm 1$, $a = \pm 1$, $d = \pm 1$, $T_{12}(-ab)D_2(d)D_1(a)M = I$.
- 若 $0 < |a| \leq |c|$, 则可对 M 作初等变换 $T_{21}(k)M$, 化为 $|c| < |a|$ 情形.
- 若 $0 < |c| < |a|$ 或 $a = 0$, 则对 M 作初等变换 $S_{12}M$, 化为前两种情形. \square

定理 4.12. 对于任意整数矩阵 $A \in \mathbb{Z}^{m \times n}$, 存在 \mathbb{Z} 上的模方阵 P, Q 使得

$$PAQ = \text{diag}(d_1, d_2, \dots, d_r, O) \quad (4.2)$$

其中 d_1, d_2, \dots, d_r 是正整数, 并且每个 d_k 整除 d_{k+1} , $1 \leq k \leq r-1$.

证明. 对矩阵的维数 (m, n) 应用数学归纳法. 不妨设 $A \neq O$ 并且 $mn \geq 2$. 当 $A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ 时, 由辗转相除法, 存在模方阵 P 使得 $PA = \begin{pmatrix} d \\ 0 \end{pmatrix}$, 其中 $d = \gcd(a_1, a_2)$. 当 $A = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$ 时, 同样存在模方阵 Q 使得 $AQ = \begin{pmatrix} d & 0 \end{pmatrix}$. 当 $m \geq 2, n \geq 2$ 时, 设 $B = (b_{ij})$ 与 A 模相抵, 满足 $b_{11} > 0$, 并且使得 b_{11} 尽可能小. 我们断言 b_{11} 整除所有 b_{ij} . 分三种情形来反证.

- (1) 存在 b_{i1} 不能被 b_{11} 整除. 同上, 存在模方阵 P 使得 $PB = (c_{ij})$, 其中 $c_{11} = \gcd(b_{11}, b_{i1}) < b_{11}$, 与 b_{11} 的最小性矛盾.
- (2) 存在 b_{1j} 不能被 b_{11} 整除. 同上, 存在模方阵 Q 使得 $BQ = (c_{ij})$, 其中 $c_{11} = \gcd(b_{11}, b_{1j}) < b_{11}$, 与 b_{11} 的最小性矛盾.
- (3) 存在 $i \neq 1, j \neq 1$ 使得 b_{ij} 不能被 b_{11} 整除. 设 $BT_{1j}(-\frac{b_{1j}}{b_{11}})T_{j1}(1) = (c_{ij})$, 则 $c_{11} = b_{11}$, $c_{i1} = (1 - \frac{b_{1j}}{b_{11}})b_{i1} + b_{ij}$ 不能被 c_{11} 整除. 化为情形 (1), 矛盾. 断言证毕.

由于 b_{11} 整除所有 b_{ij} , 故存在初等模方阵 P, Q 使得 $PBQ = \begin{pmatrix} b_{11} & & \\ & b_{11}C & \end{pmatrix}$. 对整数方阵 C 应用归纳假设, 即可完成证明. \square



思考: 定理 4.12 中的 d_1, d_2, \dots, d_r 是否唯一?

显然 $r = \text{rank}(A)$. 设 g 是 A 的所有元素的最大公约数. 由 $d_1 = \mathbf{e}_1^T PAQ \mathbf{e}_1$ 知 g 整除 d_1 . 再由 $A = P^{-1} \text{diag}(d_1, \dots, d_r, O) Q^{-1}$ 知 d_1 整除 A 的所有元素. 故 $d_1 = g$. 受此启发, 我们引入如下定义.

定义 4.6. 设 A 是整数矩阵. A 的所有 k 阶子式的最大公约数 D_k 称为 A 的第 k 个行列式因子. 特别规定, $D_0 = 1$; 当 $k > \text{rank}(A)$ 时, $D_k = 0$.

定理 4.13. $D_k = d_1 d_2 \cdots d_k, 1 \leq k \leq r$, 其中 d_1, d_2, \dots, d_r 如定理 4.12 所述.

证明. 设 P, Q 如定理 4.12 所述, 则

$$\text{diag}(d_1, d_2, \dots, d_k) = P \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & k \end{bmatrix} A Q \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & k \end{bmatrix}.$$

根据 Binet-Cauchy 公式,

$$d_1 d_2 \cdots d_k = \sum_{\substack{1 \leq i_1 < i_2 < \cdots < i_k \leq m \\ 1 \leq j_1 < j_2 < \cdots < j_k \leq n}} \det \left(P \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & k \end{bmatrix} \right) \det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right) \det \left(Q \begin{bmatrix} j_1 & j_2 & \cdots & j_k \\ 1 & 2 & \cdots & k \end{bmatrix} \right)$$

可被 D_k 整除. 另一方面, 对于任意 $1 \leq i_1 < i_2 < \cdots < i_k \leq m, 1 \leq j_1 < j_2 < \cdots < j_k \leq n$,

$$A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} = P^{-1} \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & r \end{bmatrix} \text{diag}(d_1, d_2, \dots, d_r) Q^{-1} \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & k \end{bmatrix}.$$

再根据 Binet-Cauchy 公式,

$$\det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right) = \sum_{1 \leq t_1 < t_2 < \cdots < t_k \leq r} d_{t_1} d_{t_2} \cdots d_{t_k} \det \left(P^{-1} \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ t_1 & t_2 & \cdots & t_k \end{bmatrix} \right) \det \left(Q^{-1} \begin{bmatrix} t_1 & t_2 & \cdots & t_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right)$$

可被 $d_1 d_2 \cdots d_k$ 整除. 综上, $D_k = d_1 d_2 \cdots d_k$. \square

定义 4.7. 根据定理 4.13, $d_k = \frac{D_k}{D_{k-1}}$ 由 A 唯一确定, 称为 A 的第 k 个不变因子, $1 \leq k \leq r$. 定理 4.12 (4.2) 式中的 $m \times n$ 矩阵 $\text{diag}(d_1, d_2, \dots, d_r, O)$ 称为 A 的模相抵标准形或 Smith^[6]标准形.

利用整数方阵的模相抵, 我们可以像通常的线性方程组一样求解线性同余方程组.

^[6]Henry John Stephen Smith, 1826–1883, 爱尔兰数学家.

例 4.14. [7] 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？

解答. 问题可以转化为求如下线性方程组 $Ax = b$ 的正整数解.

$$\begin{pmatrix} 1 & -3 & 0 & 0 \\ 1 & 0 & -5 & 0 \\ 1 & 0 & 0 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$$

对线性方程组的增广矩阵作初等模变换，得

$$\begin{aligned} & \begin{pmatrix} 1 & -3 & 0 & 0 & 2 \\ 1 & 0 & -5 & 0 & 3 \\ 1 & 0 & 0 & -7 & 2 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & -3 & 0 & 0 & 2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 3 & 0 & -7 & 0 \end{pmatrix} \xrightarrow{\text{列变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 3 & 0 & -7 & 0 \end{pmatrix} \\ & \xrightarrow{\text{列变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 3 & 1 & 0 & 1 \\ 0 & 3 & 6 & -7 & 0 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 3 & 1 & 0 & 1 \\ 0 & -15 & 0 & -7 & -6 \end{pmatrix} \xrightarrow{\text{列变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & -7 & -6 \end{pmatrix} \\ & \xrightarrow{\text{列变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 & -6 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}, \end{aligned}$$

其中所有列变换的乘积对应模方阵

$$Q = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -15 & 6 & 105 \\ 0 & -5 & 2 & 35 \\ 0 & -3 & 1 & 21 \\ 0 & -2 & 0 & 15 \end{pmatrix}.$$

也就是说，经过一系列初等模变换，线性方程组 $Ax = b$ 同解变形为 $PAQy = Pb$,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 1 \end{pmatrix}.$$

解得

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & -15 & 6 & 105 \\ 0 & -5 & 2 & 35 \\ 0 & -3 & 1 & 21 \\ 0 & -2 & 0 & 15 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \\ 1 \\ y_4 \end{pmatrix} = y_4 \begin{pmatrix} 105 \\ 35 \\ 21 \\ 15 \end{pmatrix} - \begin{pmatrix} 82 \\ 28 \\ 17 \\ 12 \end{pmatrix}, \quad y_4 \in \mathbb{Z}.$$

因此，原问题所求的“物之数” $x_1 = 105k - 82$, k 是正整数. □

感兴趣的读者可参阅华罗庚《数论导引》第 13–14 章，了解关于整数矩阵及其应用的更多内容.

以上关于整数矩阵的模相抵的结果，可以略作修改，推广到一元多项式构成的矩阵，以及主理想整环上的矩阵.

[7] 《孙子算经》卷下第二十六题. 《孙子算经》是中国古代数学名著《算经十书》之一，成书于约公元 4 世纪.

定义 4.8. 设 $P \in \mathbb{F}[x]^{n \times n}$. 当且仅当 $0 \neq \det(P) \in \mathbb{F}$ 时, $P^{-1} \in \mathbb{F}[x]^{n \times n}$, P 称为 $\mathbb{F}[x]$ 上的模方阵. 设 $A, B \in \mathbb{F}[x]^{m \times n}$. 若存在 $\mathbb{F}[x]$ 上的模方阵 P, Q 使得 $A = PBQ$, 则称 A 与 B 在 $\mathbb{F}[x]$ 上模相抵.

容易验证, $\mathbb{F}[x]$ 上多项式矩阵的模相抵关系构成一个等价关系. 在矩阵乘法运算下, $\mathbb{F}[x]^{n \times n}$ 中模方阵的全体构成群.

定义 4.9. 设 $A \in \mathbb{F}[x]^{m \times n}$. A 的所有 k 阶子式的最大公因式 D_k (规定是首一多项式^[8]) 称为 A 的第 k 个行列式因子. 特别规定, $D_0 = 1$; 当 $k > \text{rank}(A)$ 时, $D_k = 0$.

定理 4.14. 对于任意多项式矩阵 $A \in \mathbb{F}[x]^{m \times n}$, 存在 $\mathbb{F}[x]$ 上的模方阵 P, Q 使得

$$PAQ = \text{diag}(d_1, d_2, \dots, d_r, O)$$

其中 $d_1, d_2, \dots, d_r \in \mathbb{F}[x]$ 是首一多项式, 并且每个 d_k 整除 d_{k+1} , $1 \leq k \leq r-1$. 特别, $d_k = \frac{D_k}{D_{k-1}}$ 由 A 唯一确定, $1 \leq k \leq r$.

定义 4.10. 设 $A \in \mathbb{F}[x]^{m \times n}$. 定理 4.14 中的 $m \times n$ 矩阵 $\text{diag}(d_1, d_2, \dots, d_r, O)$ 称为 A 的模相抵标准形或 Smith 标准形, d_k 称为 A 的第 k 个不变因子, $1 \leq k \leq r = \text{rank}(A)$.

例 4.15. 首一多项式 $f = a_0 + a_1x + \dots + a_{n-1}x^{n-1} + x^n \in \mathbb{F}[x]$ 的友方阵定义为

$$A = \begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{F}^{n \times n},$$

其中空白处的元素都是 0. A 的特征方阵定义为

$$xI - A = \begin{pmatrix} x & & & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & x & a_{n-2} \\ & & -1 & x + a_{n-1} \end{pmatrix} \in \mathbb{F}[x]^{n \times n}.$$

由 $D_{n-1}(xI - A) = 1$, $\det(xI - A) = f$, 得 $xI - A$ 的 Smith 标准形为 $\text{diag}(1, \dots, 1, f)$.

习题

1. 计算下列整数或多项式矩阵的 Smith 标准形.

$$\begin{array}{ll} (1) \begin{pmatrix} 12 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{pmatrix} & (2) \begin{pmatrix} 15 & 22 & 19 \\ 18 & 25 & 27 \\ 21 & 32 & 23 \end{pmatrix} & (3) \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix} & (4) \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 10 \end{pmatrix} \\ (5) \begin{pmatrix} x^2 + x & 0 & 0 \\ 0 & x^2 - 1 & 0 \\ 0 & 0 & x^2 + x \end{pmatrix} & (6) \begin{pmatrix} 2x & x + x^2 & x + x^3 \\ x + x^2 & 2x^2 & x^2 + x^3 \\ x + x^3 & x^2 + x^3 & 2x^3 \end{pmatrix} \\ (7) \begin{pmatrix} x & x^2 & x^3 & x^4 \\ 2x & 4x^2 & 8x^3 & 16x^4 \\ 3x & 9x^2 & 27x^3 & 81x^4 \end{pmatrix} & (8) \begin{pmatrix} x & x-1 & x-2 & x-3 \\ x^2 & x^2-1 & x^2-2 & x^2-3 \\ x^3 & x^3-1 & x^3-2 & x^3-3 \end{pmatrix} \end{array}$$

^[8]最高项系数是 1 的非零多项式称为首一多项式.

2. 设 $a_1, a_2, \dots, a_n \in \mathbb{N}^*$, $\text{diag}(d_1, d_2, \dots, d_n)$ 是 $\text{diag}(a_1, a_2, \dots, a_n)$ 的 Smith 标准形. 证明:
- (1) 若 a_1, a_2, \dots, a_n 两两互素, 则 $d_1 = \dots = d_{n-1} = 1$, $d_n = \prod_{i=1}^n a_i$.
 - (2) 设 $a_i = \prod_{j=1}^k p_j^{\alpha_{ij}}$, 其中 p_1, p_2, \dots, p_k 是两两不同的素数, $\alpha_{ij} \in \mathbb{N}$, 则 $d_i = \prod_{j=1}^k p_j^{\beta_{ij}}$, 其中 $\beta_{1j} \leq \beta_{2j} \leq \dots \leq \beta_{nj}$ 是 $\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{nj}$ 的升序排列, $\forall j$.
3. 设 $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^{m \times 1}$. 证明: 线性方程组 $Ax = b$ 有解 $x \in \mathbb{Z}^{n \times 1}$ 的充分必要条件是 $\begin{pmatrix} A & b \end{pmatrix}$ 与 $\begin{pmatrix} A & \mathbf{0} \end{pmatrix}$ 在 \mathbb{Z} 上模相抵.
4. 设 $S_{ij}, T_{ij}(\pm 1)$ 都是 n 阶初等方阵, $i \neq j$, $P = \begin{pmatrix} & & & I_{n-1} \\ & & & \\ & & & \\ 1 & & & \end{pmatrix}$. 证明:
- (1) 任意 n 阶整数模方阵可以表示为 $\{S_{ij}, T_{ij}(\pm 1) \mid i \neq j\}$ 中一系列元素的乘积.
 - (2) 任意 S_{ij} 可以表示为一系列 P, S_{12} 的乘积.
 - (3) 任意 $T_{ij}(1)$ 可以表示为一系列 $P, T_{12}(\pm 1)$ 的乘积.
 - (4) 任意 S_{ij} 可以表示为 $\{S_{12}, T_{kl}(\pm 1) \mid k \neq l\}$ 中一系列元素的乘积.
 - (5) 任意 $T_{ij}(1)$ 可以表示为 $\{S_{kl}, T_{12}(1) \mid k \neq l\}$ 中一系列元素的乘积.
5. 设 $T_{12}(\pm 1)$ 都是 n 阶初等方阵, $P = \begin{pmatrix} & & & I_{n-1} \\ & & & \\ & & & \\ (-1)^{n-1} & & & \end{pmatrix}$. 证明:
- (1) 行列式等于 1 的 n 阶整数模方阵都可以表示为 $\{T_{ij}(\pm 1) \mid i \neq j\}$ 中一系列元素的乘积.
 - (2) 任意 $T_{ij}(1)$ 可以表示为一系列 $P, T_{12}(\pm 1)$ 的乘积.
6. 仿照定理 4.12 和定理 4.13 的证明过程, 证明定理 4.14.
7. (1) 证明: 任意整数方阵 $A \in \mathbb{Z}^{n \times n}$ 与 A^T 模相抵.
 (2) 证明: 任意多项式方阵 $A \in \mathbb{F}[x]^{n \times n}$ 与 A^T 模相抵.
8. (1) 证明: $(a_1, a_2, \dots, a_n) \in \mathbb{Z}^n$ 是某个整数模方阵的行向量 $\Leftrightarrow \gcd(a_1, a_2, \dots, a_n) = 1$.
 (2) 证明: $(p_1, p_2, \dots, p_n) \in \mathbb{F}[x]^n$ 是某个多项式模方阵的行向量 $\Leftrightarrow \gcd(p_1, p_2, \dots, p_n) = 1$.
9. 设 $A, B \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 证明: 多项式方阵 $xA + B$ 的任意不变因子 $d_k \in \mathbb{R}[x]$.
10. 设 $A, B \in \mathbb{C}[x]^{n \times n}$ 满足 $A(z)$ 与 $B(z)$ 在 \mathbb{C} 上相抵, $\forall z \in \mathbb{C}$. A 与 B 是否一定在 $\mathbb{C}[x]$ 上模相抵? 证明你的结论.

第五章 矩阵的相似

§5.1 相似的概念

定义 5.1. 设 $A, B \in \mathbb{F}^{n \times n}$. 若存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $A = PBP^{-1}$, 则称 A 与 B 在 \mathbb{F} 上相似.

容易验证, 矩阵的相似关系构成一个等价关系.

例 5.1. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2 \times 2}$, $\lambda \in \mathbb{C}$. A 与下列方阵相似.

$$\begin{aligned} S_{12}AS_{12} &= \begin{pmatrix} d & c \\ b & a \end{pmatrix}, & T_{12}(\lambda)AT_{12}(-\lambda) &= \begin{pmatrix} a+c\lambda & b+(d-a)\lambda-c\lambda^2 \\ c & d-c\lambda \end{pmatrix}, \\ D_{12}(\lambda)AD_{12}(\lambda^{-1}) &= \begin{pmatrix} a & \lambda b \\ \lambda^{-1}c & d \end{pmatrix}, & T_{21}(\lambda)AT_{21}(-\lambda) &= \begin{pmatrix} a-b\lambda & b \\ c+(a-d)\lambda-b\lambda^2 & d+b\lambda \end{pmatrix}. \end{aligned}$$

特别,

- 当 A 是上(下)三角方阵时, $S_{12}AS_{12}$ 是下(上)三角方阵.
- 当 $b \neq 0$ 时, 存在 λ 使得 $T_{21}(\lambda)AT_{21}(-\lambda)$ 是上三角方阵.
- 当 $c \neq 0$ 时, 存在 λ 使得 $T_{12}(\lambda)AT_{12}(-\lambda)$ 是下三角方阵.

因此, 任意 2 阶复数方阵可以相似成上三角方阵, 也可以相似成下三角方阵.

定理 5.1. 设 A, A_i, B, B_i 都是 \mathbb{F} 上方阵. 有如下结论.

1. 若 A 与 B 相似, 则 $\text{tr}(A) = \text{tr}(B)$, $\det(A) = \det(B)$, $\text{rank}(A) = \text{rank}(B)$.
2. 若 A 与 B 相似, 则 A^T 与 B^T 相似.
3. 若 A 与 B 相似, 则 $f(A)$ 与 $f(B)$ 相似, $\forall f \in \mathbb{F}[x]$.
4. 若 A_i 与 B_i 相似, $i = 1, 2, \dots, k$, 则 $\text{diag}(A_1, A_2, \dots, A_k)$ 与 $\text{diag}(B_1, B_2, \dots, B_k)$ 相似.
5. 对于任意排列 $(i_1, i_2, \dots, i_k) \in S_k$, $\text{diag}(A_1, A_2, \dots, A_k)$ 与 $\text{diag}(A_{i_1}, A_{i_2}, \dots, A_{i_k})$ 相似.

由定理 5.1 可知, 方阵的迹、行列式、秩在相似关系下保持不变, 它们都称为相似不变量. 还有哪些矩阵性质是相似不变量?

我们希望在每个相似等价类中, 都能找到一个形式简单的矩阵作为相似标准形. 首先想到的是对角方阵. 是否每个方阵都能与某个对角方阵相似呢?

$$A = P \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^{-1} \Leftrightarrow A\alpha_i = \lambda_i\alpha_i \text{ 即 } (A - \lambda_i I)\alpha_i = \mathbf{0}$$

其中 $P = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)$. 为了更方便地研究矩阵的相似问题, 我们引入如下定义.

定义 5.2. 设 $A \in \mathbb{F}^{n \times n}$.

- n 次首一多项式 $\varphi_A(x) = \det(xI - A) \in \mathbb{F}[x]$ 称为 A 的特征多项式.
- 设 $\varphi_A(x)$ 在 \mathbb{F} 的某个扩域^[1] \mathbb{K} 上可分解为 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同. 每个 λ_i 称为 A 的一个特征值. n_i 称为 λ_i 的代数重数.
- 满足 $A\alpha = \lambda_i\alpha$ 的非零向量 $\alpha \in \mathbb{K}^{n \times 1}$ 称为 λ_i 对应的一个特征向量.
- 线性方程组 $(\lambda_i I - A)x = \mathbf{0}$ 的解集 $V_i = \{\alpha \in \mathbb{K}^{n \times 1} \mid A\alpha = \lambda_i\alpha\}$ 称为 λ_i 对应的特征子空间.
- 线性方程组 $(\lambda_i I - A)x = \mathbf{0}$ 的每个基础解系恰有 $m_i = n - \operatorname{rank}(\lambda_i I - A)$ 个向量. m_i 称为 V_i 的维数和 λ_i 的几何重数.

显然, 特征多项式、特征值、代数重数、几何重数都是相似不变量.

例 5.2.

- 对于任意 2 阶对称实数方阵 $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, $\varphi_A(x) = x^2 - (a+c)x + ac - b^2$, 两个特征值 $\lambda_{1,2} = \frac{a+c \pm \sqrt{(a-c)^2 + 4b^2}}{2}$ 都是实数.
- 对于任意 2 阶反对称实数方阵 $A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$, $\varphi_A(x) = x^2 + b^2$, 两个特征值 $\lambda_{1,2} = \pm bi$ 都是纯虚数或 0.
- 对于任意 n 阶三角方阵 $A = (a_{ij})$, $\varphi_A(x) = \prod_{i=1}^n (x - a_{ii})$, A 的对角元素即为所有特征值.

定理 5.2. 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶方阵 A 的所有特征值, 则有

$$\varphi_A(x) = x^n - \sigma_1 x^{n-1} + \dots + (-1)^{n-1} \sigma_{n-1} x + (-1)^n \sigma_n$$

其中

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_k} = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \det \left(A \begin{bmatrix} i_1 & i_2 & \dots & i_k \\ i_1 & i_2 & \dots & i_k \end{bmatrix} \right)$$

既是关于 $\lambda_1, \lambda_2, \dots, \lambda_n$ 的 k 次基本对称多项式, 也是 A 的所有 k 阶主子式之和. 特别,

$$\sigma_1 = \lambda_1 + \lambda_2 + \dots + \lambda_n = \operatorname{tr}(A)$$

$$\sigma_n = \lambda_1 \lambda_2 \dots \lambda_n = \det(A).$$

^[1]若在 \mathbb{K} 的加、减、乘、除运算下, 数域 \mathbb{K} 的子集 \mathbb{F} 构成数域, 则 \mathbb{F} 称为 \mathbb{K} 的子域, \mathbb{K} 称为 \mathbb{F} 的扩域.

证明. 一方面, 由 Vieta 定理可得

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k}.$$

另一方面, 设 $\alpha_1, \cdots, \alpha_n$ 是 A 的所有行向量, 则

$$\begin{aligned} \varphi_A(x) &= \det(x\mathbf{e}_1 - \alpha_1, x\mathbf{e}_2 - \alpha_2, \cdots, x\mathbf{e}_n - \alpha_n) \\ &= \sum_{k=0}^n \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} (-1)^{\tau(i_1, i_2, \cdots, i_n)} \det(-\alpha_{i_1}, \cdots, -\alpha_{i_k}, x\mathbf{e}_{i_{k+1}}, \cdots, x\mathbf{e}_{i_n}) \\ &= \sum_{k=0}^n \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} (-1)^k x^{n-k} \det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{bmatrix} \right), \end{aligned}$$

其中 $\{i_{k+1}, \cdots, i_n\}$ 是 $\{i_1, \cdots, i_k\}$ 在 $\{1, 2, \cdots, n\}$ 中的补集. 故

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{bmatrix} \right). \quad \square$$

定理 5.3. 设 $A \in \mathbb{F}^{n \times n}$ 有 n 个两两不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{F}$, 则 A 与 $\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 在 \mathbb{F} 上相似.

证明. 设 $P = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)$, 其中 $\alpha_i \in \mathbb{F}^{n \times 1}$ 是 λ_i 对应的特征向量, $i = 1, 2, \cdots, n$. 下面证明 P 是可逆方阵. 设 $x = (x_i) \in \mathbb{F}^{n \times 1}$ 是线性方程组 $Px = \mathbf{0}$ 的解, 即满足 $\sum_{i=1}^n x_i \alpha_i = \mathbf{0}$. 由

$$A^k \sum_{i=1}^n x_i \alpha_i = \sum_{i=1}^n \lambda_i^k x_i \alpha_i = \mathbf{0}, \quad \forall k \in \mathbb{N}$$

可得

$$\begin{pmatrix} x_1 \alpha_1 & x_2 \alpha_2 & \cdots & x_n \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{pmatrix} = O.$$

由于 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 两两不同, n 阶 Vandermonde 方阵 (λ_i^{j-1}) 是可逆方阵. 从而,

$$\begin{pmatrix} x_1 \alpha_1 & x_2 \alpha_2 & \cdots & x_n \alpha_n \end{pmatrix} = O \Rightarrow x = \mathbf{0}.$$

由线性方程组 $Px = \mathbf{0}$ 的解集结构, 知 P 是可逆方阵. 因此, $A = P \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) P^{-1}$ 与 $\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 在 \mathbb{F} 上相似. \square

例 5.3. 计算实数方阵 $A = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ 的所有特征值和所有特征向量, 并把 A 相似成对角方阵.

解答. $\varphi_A(x) = \begin{vmatrix} x & 0 & -1 \\ 1 & x-1 & -1 \\ -1 & 0 & x \end{vmatrix} = x^3 - x^2 - x + 1 = (x-1)^2(x+1)$. A 的所有特征值为 $1, 1, -1$.

求解 $(A - I)\alpha = \mathbf{0}$, 得 $\lambda_1 = 1$ 对应的特征向量 $\alpha_1 = (s, t, s)^T$, 其中 $(s, t) \neq (0, 0)$.

求解 $(A + I)\alpha = \mathbf{0}$, 得 $\lambda_2 = -1$ 对应的特征向量 $\alpha_2 = (s, s, -s)^T$, 其中 $s \neq 0$.

故 $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ 满足 $AP = P \operatorname{diag}(1, 1, -1)$. 由 $\det(P) = -2$ 得 A 与 $\operatorname{diag}(1, 1, -1)$ 相似. \square

定理 5.4. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值都属于 \mathbb{F} . A 可以在 \mathbb{F} 上相似成对角方阵的充分必要条件是每个特征值的几何重数等于代数重数.

证明. (充分性) 设 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同. 设 $\{\alpha_j \mid s_{i-1} + 1 \leq j \leq s_i\}$ 是线性方程组 $(\lambda_i I - A)x = \mathbf{0}$ 的一个基础解系, $s_0 = 0$, $s_i = n_1 + \dots + n_i$, $1 \leq i \leq k$. 易知 $P = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)$ 满足 $AP = P \operatorname{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k})$. 假设 $Px = \mathbf{0}$, 同定理 5.3 的证明可得 $\sum_{j=s_{i-1}+1}^{s_i} x_j \alpha_j = \mathbf{0}$, $\forall i$. 故 $x = \mathbf{0}$. 因此, P 是可逆方阵, A 与 $\operatorname{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k})$ 相似.

(必要性) 设 A 与 $B = \operatorname{diag}(b_1, b_2, \dots, b_n)$ 在 \mathbb{F} 上相似. 几何重数和代数重数都是相似不变量. 容易验证, B 的每个特征值 b_i 的几何重数等于代数重数. \square

定理 5.3 说明, 当 A 的特征值不重复时, A 可以相似成对角方阵. 定理 5.4 说明, 即使 A 的特征值有重数, 只要特征向量“足够多”, A 也可以相似成对角方阵.

例 5.4. 计算实数方阵 $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ 的所有特征值和所有特征向量, 并把 A 相似成对角方阵.

解答. $\varphi_A(x) = \begin{vmatrix} x-1 & -1 & -1 \\ -2 & x & -1 \\ 2 & -1 & x \end{vmatrix} = x^3 - x^2 - x + 1 = (x-1)^2(x+1)$. A 的所有特征值为 $1, 1, -1$.

求解 $(A - I)\alpha = \mathbf{0}$, 得 $\lambda_1 = 1$ 对应的特征向量 $\alpha_1 = (s, s, -s)^T$, 其中 $s \neq 0$.

求解 $(A + I)\alpha = \mathbf{0}$, 得 $\lambda_2 = -1$ 对应的特征向量 $\alpha_2 = (0, s, -s)^T$, 其中 $s \neq 0$.

λ_1 的代数重数 $n_1 = 2$, 几何重数 $m_1 = 1$. 根据定理 5.4, A 不可以相似成对角方阵. \square

例 5.5 (例 2.7 续). 设 $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$, 计算 A^n , 其中 $a, b \in \mathbb{C}$.

解答. $\varphi_A(x) = x^2 - ax - b$, 特征值 $\lambda_1 = \frac{a + \sqrt{a^2 + 4b}}{2}$, $\lambda_2 = \frac{a - \sqrt{a^2 + 4b}}{2}$.

- 当 $a^2 + 4b \neq 0$ 时, $\lambda_1 \neq \lambda_2$.

设 $\alpha_i = \begin{pmatrix} \lambda_i \\ 1 \end{pmatrix}$ 是 λ_i 对应的特征向量, 则 $P = (\alpha_1 \ \alpha_2) = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}$ 是可逆方阵, 满足 $AP = P \operatorname{diag}(\lambda_1, \lambda_2)$. 故 $A^n = P \operatorname{diag}(\lambda_1^n, \lambda_2^n) P^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1^{n+1} - \lambda_2^{n+1} & (\lambda_1^n - \lambda_2^n)b \\ \lambda_1^n - \lambda_2^n & (\lambda_1^{n-1} - \lambda_2^{n-1})b \end{pmatrix}$.

- 当 $a^2 + 4b = 0$ 时, $\lambda_1 = \lambda_2 = \frac{a}{2}$.

设 $\alpha_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\alpha_2 = (A - \lambda_1 I)\alpha_1 = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$, 则 $P = (\alpha_1 \ \alpha_2) = \begin{pmatrix} 1 & \lambda_1 \\ 0 & 1 \end{pmatrix}$ 是可逆方阵, 满足 $AP = P \begin{pmatrix} \lambda_1 & 0 \\ 1 & \lambda_1 \end{pmatrix}$. 故 $A^n = P \begin{pmatrix} \lambda_1 & 0 \\ 1 & \lambda_1 \end{pmatrix}^n P^{-1} = \begin{pmatrix} 1 & \lambda_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ n\lambda_1^{n-1} & \lambda_1^n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\lambda_1 & 1 \end{pmatrix} = \begin{pmatrix} (n+1)\lambda_1^n & -n\lambda_1^{n+1} \\ n\lambda_1^{n-1} & (1-n)\lambda_1^n \end{pmatrix}$. \square

习题

1. 证明定理 5.1.
2. 证明: 对于任意上三角方阵 A , 存在可逆方阵 P 使得 $P^{-1}AP$ 是下三角方阵.
3. 对下列复数方阵 A , 试求可逆方阵 P 使得 $P^{-1}AP$ 为对角方阵. 若 P 不存在, 请说明理由.

$$(1) \begin{pmatrix} 0 & 1 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \quad (2) \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 4 \\ 1 & -1 & 3 \end{pmatrix} \quad (3) \begin{pmatrix} 3 & -2 & -2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \quad (4) \begin{pmatrix} 0 & 2 & -1 \\ -1 & 3 & -1 \\ 2 & -2 & 3 \end{pmatrix}$$

$$(5) \begin{pmatrix} \mathbf{0} & I_{n-1} \\ a & \mathbf{0} \end{pmatrix}, a \neq 0 \quad (6) a_{ij} = ij, 1 \leq i, j \leq n \quad (7) a_{ij} = i + j - 1, 1 \leq i, j \leq n$$

4. 设 $A \in \mathbb{R}^{n \times n}$. 证明:
 - (1) 若 n 是奇数, 则 A 有实特征值.
 - (2) 若 A 无实特征值, 则 $\det(A) > 0$.
 - (3) 若 A 是对称方阵, 则 A 的特征值都是实数.
 - (4) 若 A 是反对称方阵, 则 A 的特征值的实部都是 0.
5. 设 $A \in \mathbb{F}^{n \times n}$, $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 A 的一些两两不同的特征值, α_i 是 λ_i 对应的特征向量. 证明: $n \times k$ 矩阵 $(\alpha_1 \ \alpha_2 \ \dots \ \alpha_k)$ 一定是列满秩的.

6. 设 $A \in \mathbb{F}^{n \times n}$, $\varphi_A(x) = \sum_{k=0}^n c_k x^k$. 证明:

$$(1) \text{ 当 } c_0 \neq 0 \text{ 时, 设 } B = A^{-1}, \text{ 则 } \varphi_B(x) = \sum_{k=0}^n \frac{c_{n-k}}{c_0} x^k.$$

$$(2) \text{ 当 } c_0 = 0 \text{ 时, 设 } B = A^*, \text{ 则 } \varphi_B(x) = x^n + (-1)^n c_1 x^{n-1}.$$

7. 设 $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, $\varphi_A(x) = \sum_{k=0}^n c_k x^k$, λ 是 A 的任意特征值. 证明:

$$(1) |\lambda| \leq 1 + \max_{0 \leq k \leq n-1} |c_k|.$$

$$(2) |\lambda| \leq 2 \max_{0 \leq k \leq n-1} |c_k|^{\frac{1}{n-k}}.$$

$$(3) (\text{Gershgorin}^{[2]} \text{圆盘定理}) \text{ 存在 } k \text{ 使得 } |\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}|.$$

^[2]Semyon Aranovich Gershgorin, 1901–1933, 苏联数学家.

8. 设 A, B 分别是首一多项式 $f, g \in \mathbb{F}[x]$ 的友方阵. 证明: A 与 B 相似 $\Leftrightarrow f = g$.
9. 设 $A, B \in \mathbb{F}^{n \times n}$ 满足 $AB = O$. 证明: $\varphi_A(x)\varphi_B(x) = x^n\varphi_{A+B}(x)$.
10. 设 $A, B \in \mathbb{F}^{n \times n}$ 满足 $\text{rank}(A) = \text{rank}(B) = 1$. 证明:
- (1) $\varphi_A(x) = x^n - \text{tr}(A)x^{n-1}$.
 - (2) $\text{tr}(A) \neq 0$ 是 A 可以相似成对角方阵的充分必要条件.
 - (3) $\text{tr}(A) = \text{tr}(B)$ 是 A 与 B 相似的充分必要条件.
11. 设 $A, B \in \mathbb{F}^{n \times n}$. 证明:
- (1) $\varphi_{AB} = \varphi_{BA}$.
 - (2) 若 $\text{rank}(ABA) = \text{rank}(A)$, 则 AB 与 BA 相似.
 - (3) 若 $\text{rank}(ABA) = \text{rank}(B)$, 则 AB 与 BA 相似.
 - (4) 举例: $\text{rank}(AB) = \text{rank}(BA)$, AB 与 BA 不相似.
12. 设 $H_0 = 0$. 求 2^n 阶整数方阵 $H_n = \begin{pmatrix} H_{n-1} & I \\ I & H_{n-1} \end{pmatrix}$ 的特征多项式和所有特征值, $n \geq 1$.
13. 设 n 阶实数方阵 $A = (a_{ij})$, $B = (b_{ij})$, 其中 $1 \leq i, j \leq n$ 并且

$$a_{ij} = \begin{cases} 1, & |i-j| = 1; \\ 0, & |i-j| \neq 1; \end{cases} \quad b_{ij} = \begin{cases} 1, & |i-j| = 1 \text{ 或 } i=j=1 \text{ 或 } i=j=n; \\ 2, & i=j \text{ 且 } 1 < i < n; \\ 0, & |i-j| \geq 2. \end{cases}$$

证明: (1) $\varphi_A(x) = \prod_{k=1}^n \left(x - 2 \cos \frac{k\pi}{n+1}\right)$. (2) $\varphi_B(x) = \prod_{k=1}^n \left(x - 2 - 2 \cos \frac{k\pi}{n}\right)$.

14. 设 $n+1$ 阶实数方阵 $A = (a_{ij})$, $B = (b_{ij})$, 其中 $0 \leq i, j \leq n$ 并且

$$a_{ij} = \begin{cases} n-2i, & j=i; \\ i, & j=i-1; \\ i-n, & j=i+1; \\ 0, & |i-j| \geq 2; \end{cases} \quad b_{ij} = \begin{cases} n^2 + (4i+2)n - 8i^2, & j=i; \\ 2i(2i-1), & j=i-1; \\ (2i-2n)(2i+1), & j=i+1; \\ 0, & |i-j| \geq 2. \end{cases}$$

证明: (1) $\varphi_A(x) = x^{n+1}$. (2) $\varphi_{tA+B}(x) = \prod_{k=-n}^n (x - (n-2k)^2)$, $\forall t \in \mathbb{R}$.

15. 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 满足 (i) $a_{ij} \leq 0$, $\forall i \neq j$; (ii) $\sum_{j=1}^n a_{ij} = 0$, $\forall i$. 证明:

- (1) $(1, \dots, 1)^T$ 是 A 的特征值 0 对应的特征向量.
- (2) 设 $\lambda \in \mathbb{C}$ 是 A 的任意特征值, 则 $\text{Re}(\lambda) \geq 0$.
- (3) A 可置换相似成 $\text{diag}(A_1, \dots, A_k)$ 的形式, 其中 A_i 是 n_i 阶方阵, $\text{rank}(A_i) = n_i - 1$, $\forall i$.
- (4) 若 $\text{rank}(A) = n - 1$, $x = (x_1, \dots, x_n) \in \mathbb{R}^{1 \times n}$ 满足 $xA = \mathbf{0}$, 则 $x_i x_j \geq 0$, $\forall i, j$.

16. 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 是非负方阵 (参考第二章 §2.4 习题 14), $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ 是 A 的所有特征值. 不妨设 $|\lambda_1| \geq \dots \geq |\lambda_n|$. 证明:
- (1) 存在 k 使得 $\lambda_k = |\lambda_1|$. 若 A 是不可约的, 则 k 是唯一的.
 - (2) 设 A 是不可约的, $\alpha = (x_1, \dots, x_n)^T \in \mathbb{R}^{n \times 1}$ 是 $|\lambda_1|$ 对应的特征向量, 则 $x_i x_j > 0, \forall i, j$.
 - (3) 若 $\sum_{j=1}^n a_{ij} = 1, \forall i$, 则 $|\lambda_1| = 1$. 这样的非负方阵 A 常称为**概率矩阵**或 **Markov 矩阵**.
 - (4) A 是本原的 $\Leftrightarrow \lambda_1 > |\lambda_2|$.

§5.2 相似三角化

定理 5.5. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$, 则存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $P^{-1}AP$ 是上三角方阵, 并且 $P^{-1}AP$ 的对角元素依次是 $\lambda_1, \lambda_2, \dots, \lambda_n$.

证明. 对 n 使用数学归纳法. 当 $n = 1$ 时, 结论显然成立. 当 $n \geq 2$ 时, 设 $\alpha_1 \in \mathbb{F}^{n \times 1}$ 是 λ_1 对应的一个特征向量, 则存在可逆方阵 $P_1 \in \mathbb{F}^{n \times n}$ 以 α_1 为第 1 列, 使得 $AP_1 = P_1 \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{pmatrix}$. 注意到 $\varphi_A(x) = (x - \lambda_1)\varphi_B(x)$, 根据归纳假设, 存在可逆方阵 P_2 使得 $P_2^{-1}BP_2$ 是上三角方阵, 并且 $P_2^{-1}BP_2$ 的对角元素依次是 $\lambda_2, \dots, \lambda_n$. 从而, $P = P_1 \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}$ 满足题设. \square

根据定理 5.5, 可以得到下列推论.

定理 5.6. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$, $f(x) \in \mathbb{F}[x]$.

1. $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$ 是 $f(A)$ 的所有特征值.
2. $f(A)$ 是可逆方阵 $\Leftrightarrow f(x)$ 与 $\varphi_A(x)$ 互素.
3. 当 A 是可逆方阵时, $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$ 是 A^{-1} 的所有特征值.
4. $\text{rank}(A - \lambda_i I)^k = n - n_i, \forall k \geq n_i$, 其中 n_i 是 A 的特征值 λ_i 的代数重数.
5. A 的特征值 λ_i 的几何重数不超过 λ_i 的代数重数.

定理 5.7 (Cayley-Hamilton^[3]定理). 对于任意方阵 A 都有 $\varphi_A(A) = O$.

证明. 不妨设 $A \in \mathbb{F}^{n \times n}$, $\varphi_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$, $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$. 根据定理 5.5, 存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得

$$B = P^{-1}AP = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{pmatrix}.$$

利用数学归纳法可以证明: $(B - \lambda_1 I)(B - \lambda_2 I) \cdots (B - \lambda_k I)$ 的前 k 列元素都是 0, $k = 1, 2, \dots, n$. 从而, $\varphi_A(A) = P\varphi_A(B)P^{-1} = O$. \square

定理 5.8. 设 $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$ 满足 $\varphi_A(x)$ 与 $\varphi_B(x)$ 互素. 证明: 对于任意 $C \in \mathbb{F}^{m \times n}$, 存在唯一的 $X \in \mathbb{F}^{m \times n}$ 使得 $AX - XB = C$. 从而, $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 与 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 相似.

证法一. 设 $B = (b_{ij})$, $C = (\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n)$, $X = (\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n)$. 矩阵方程 $AX - XB = C$ 可以表示为线性方程组的形式.

$$\begin{pmatrix} A - b_{11}I & -b_{21}I & \cdots & -b_{n1}I \\ -b_{12}I & A - b_{22}I & \cdots & -b_{n2}I \\ \vdots & \vdots & \ddots & \vdots \\ -b_{1n}I & -b_{2n}I & \cdots & A - b_{nn}I \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_n \end{pmatrix}. \quad (5.1)$$

^[3]William Rowan Hamilton, 1805-1865, 爱尔兰物理学家、数学家.

线性方程组 (5.1) 的系数矩阵 $M = I_n \otimes A - B^T \otimes I_m$. 设 $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ 是 A 的所有特征值, $\mu_1, \dots, \mu_n \in \mathbb{K}$ 是 B 的所有特征值, \mathbb{K} 是 \mathbb{F} 的某个扩域. 根据定理 5.5, 存在可逆方阵 $P \in \mathbb{K}^{m \times m}$ 和 $Q \in \mathbb{K}^{n \times n}$ 使得 $\tilde{A} = P^{-1}AP$ 和 $\tilde{B} = Q^{-1}B^TQ$ 都是上三角方阵. 注意到

$$(Q \otimes P)^{-1}M(Q \otimes P) = I_n \otimes \tilde{A} - \tilde{B} \otimes I_m$$

也是上三角方阵, 故

$$\det(M) = \det(I_n \otimes \tilde{A} - \tilde{B} \otimes I_m) = \prod_{i=1}^m \prod_{j=1}^n (\lambda_i - \mu_j) \neq 0.$$

因此, 线性方程组 (5.1) 有唯一解, 即矩阵方程 $AX - XB = C$ 有唯一解.

从而, $\begin{pmatrix} A & C \\ O & B \end{pmatrix} = \begin{pmatrix} I & -X \\ O & I \end{pmatrix} \begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} I & X \\ O & I \end{pmatrix}$ 与 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 相似. □

证法二. 由 (5.1) 式, $AX - XB = C$ 有唯一解 $\Leftrightarrow \det(M) \neq 0 \Leftrightarrow AX - XB = O$ 有唯一解 $X = O$. 根据定理 5.6 结论 2, $\varphi_B(A)$ 是可逆方阵. 再根据 Cayley-Hamilton 定理,

$$AX = XB \Rightarrow \varphi_B(A)X = X\varphi_B(B) = O \Rightarrow X = O.$$

从而 $AX - XB = C$ 有唯一解. 同证法一, $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 与 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 相似. □

定理 5.9. 设 $A \in \mathbb{F}^{n \times n}$, $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}$ 并且两两不同, 则存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $P^{-1}AP = \text{diag}(A_1, A_2, \dots, A_k)$, 其中每个 A_i 是 n_i 阶上三角方阵, 并且 A_i 的对角元素都是 λ_i .

证明. 根据定理 5.5, 存在可逆方阵 $P_1 \in \mathbb{F}^{n \times n}$ 使得

$$B = P_1^{-1}AP_1 = \begin{pmatrix} A_1 & * & \cdots & * \\ & A_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & A_k \end{pmatrix}, \quad \text{其中 } A_i = \begin{pmatrix} \lambda_i & * & \cdots & * \\ & \lambda_i & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_i \end{pmatrix} \in \mathbb{F}^{n_i \times n_i}.$$

根据定理 5.8, $B = \begin{pmatrix} A_1 & * \\ O & \tilde{A} \end{pmatrix}$ 与 $\begin{pmatrix} A_1 & O \\ O & \tilde{A} \end{pmatrix}$ 相似. 再对 \tilde{A} 分块, 反复运用定理 5.8, 可得 A 与 $\text{diag}(A_1, A_2, \dots, A_k)$ 相似. □

例 5.6. 设 $A = \begin{pmatrix} 0 & -5 & -6 & 2 & 2 \\ 1 & 4 & 4 & -2 & -2 \\ -1 & -5 & -5 & 2 & 2 \\ -2 & -5 & -5 & 3 & 2 \\ 3 & 1 & 0 & -2 & -1 \end{pmatrix}$. 求可逆方阵 P 使得 $P^{-1}AP$ 如定理 5.9 所示.

解答. 计算特征多项式 $\varphi_A(x) = (x-1)^3(x+1)^2$, 得 A 的特征值 $\lambda_1 = 1$ (3重), $\lambda_2 = -1$ (2重).

$$(A - \lambda_1 I)^3 = \begin{pmatrix} -8 & -28 & -28 & 8 & 8 \\ 8 & 20 & 20 & -8 & -8 \\ -8 & -28 & -28 & 8 & 8 \\ -8 & -28 & -28 & 8 & 8 \\ 8 & 12 & 12 & -8 & -8 \end{pmatrix}, \quad (A - \lambda_2 I)^2 = \begin{pmatrix} 4 & -8 & -12 & 4 & 4 \\ 0 & 8 & 8 & -4 & -4 \\ 0 & -8 & -8 & 4 & 4 \\ -4 & -8 & -8 & 8 & 4 \\ 8 & 0 & -4 & -4 & 0 \end{pmatrix}.$$

$(A - \lambda_1 I)^3 x = \mathbf{0}$ 的基础解系 $\alpha_1 = (0, -1, 1, 0, 0)^T$, $\alpha_2 = (1, 0, 0, 1, 0)^T$, $\alpha_3 = (1, 0, 0, 0, 1)^T$.

$(A - \lambda_2 I)^2 x = \mathbf{0}$ 的基础解系 $\alpha_4 = (2, -1, 2, 2, 0)^T$, $\alpha_5 = (0, 1, 0, 0, 2)^T$.

设 $P_1 = (\alpha_1 \cdots \alpha_5)$, 则

$$P_1^{-1} = \begin{pmatrix} -1 & -2 & -1 & 1 & 1 \\ -1 & -2 & -2 & 2 & 1 \\ 1 & 0 & 0 & -1 & 0 \\ \frac{1}{2} & 1 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad P_1^{-1} A P_1 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix},$$

其中 $A_1 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}$, $A_2 = \begin{pmatrix} -\frac{3}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$, 满足 $(A_1 - I)^2 = O$, $(A_2 + I)^2 = O$.

$(A_1 - I)x = \mathbf{0}$ 的通解 $x = (s+t, s, t)^T$. 取 β_1 使得 $\beta_2 = (A_1 - I)\beta_1 \neq \mathbf{0}$, 取 β_3 使得 $(A_1 - I)\beta_3 = \mathbf{0}$ 且 β_3 与 β_2 线性无关, 则 $Q_1 = (\beta_2 \beta_1 \beta_3)$ 满足

$$A_1 Q_1 = Q_1 \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = Q_1 B_1.$$

$(A_2 + I)x = \mathbf{0}$ 的通解 $x = (t, -t)^T$. 取 γ_1 使得 $\gamma_2 = (A_2 + I)\gamma_1 \neq \mathbf{0}$, 则 $Q_2 = (\gamma_2 \gamma_1)$ 满足

$$A_2 Q_2 = Q_2 \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = Q_2 B_2.$$

例如, 可取

$$\beta_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \quad \beta_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$\text{则 } P = P_1 \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 & -1 & 2 \\ 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 2 \\ 0 & 0 & 1 & -1 & 2 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix} \text{ 满足 } P^{-1} A P = \begin{pmatrix} B_1 & \\ & B_2 \end{pmatrix}. \quad \square$$

习题

1. 证明定理 5.6.

2. 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶方阵 $A = (a_{ij})$ 的所有特征值. 证明: $\sum_{i=1}^n \lambda_i^2 = \sum_{i,j=1}^n a_{ij}a_{ji}$.

3. 设正整数 $m \geq 2$, $\omega = \cos \frac{2\pi}{m} + i \sin \frac{2\pi}{m}$. 证明:

(1) 对于任意 $f \in \mathbb{C}[x]$, 存在 $g \in \mathbb{C}[x]$ 使得 $\prod_{k=1}^m f(\omega^k x) = g(x^m)$.

(2) 对于任意 $A \in \mathbb{C}^{n \times n}$, 设 $B = A^m$, 则 $\varphi_B(x) = (-1)^{(m-1)n} \prod_{k=1}^m \varphi_A(\omega^k x^{\frac{1}{m}})$.

4. 设可逆方阵 $P = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}$ 使得 $\begin{pmatrix} A & C_1 \\ O & B \end{pmatrix} = P \begin{pmatrix} A & C_2 \\ O & B \end{pmatrix} P^{-1}$, 其中 A 和 P_1 是 m 阶方阵, B 和 P_4 是 n 阶方阵, 并且 $\varphi_A(x)$ 与 $\varphi_B(x)$ 互素. 证明: $P_3 = O$.

5. 设 $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, $C \in \mathbb{F}^{m \times n}$, $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ 是 A 的所有特征值, $\mu_1, \dots, \mu_n \in \mathbb{F}$ 是 B 的所有特征值, $\varphi_B(x) = \sum_{k=0}^n b_k x^k$. 证明:

(1) 当 $\lambda_1 = \mu_1$ 时, 设 $\alpha \in \mathbb{F}^{m \times 1}$ 是 λ_1 对应的 A 的特征向量, $\beta \in \mathbb{F}^{n \times 1}$ 是 μ_1 对应的 B^T 的特征向量, 则 $X = \alpha\beta^T$ 是方程 $AX - XB = O$ 的一个非零解.

(2) 当且仅当 $\lambda_i \neq \mu_j, \forall i, j$ 时, 方程 $AX - XB = C$ 有唯一解 X , 并且

$$X = (\varphi_B(A))^{-1} \sum_{i,j \geq 0} b_{i+j+1} A^i C B^j.$$

(3) 当且仅当 $\lambda_i \mu_j \neq 1, \forall i, j$ 时, 方程 $X - AXB = C$ 有唯一解 X .

6. 设 $A \in \mathbb{F}^{n \times n}$ 满足 $A^k = O$, 其中 k 是某个正整数. 证明: $\varphi_A(x) = x^n$.

7. 设 $A \in \mathbb{F}^{n \times n}$ 满足 $A^k = A$, 其中 k 是某个正整数. 证明:

(1) 若 $k = 2$, 则 A 与 $\text{diag}(I_r, O)$ 相似, $r = \text{rank}(A)$.

(2) 若 $k = 3$, 则 A 与 $\text{diag}(I_p, -I_q, O)$ 相似, $p = n - \text{rank}(A - I)$, $q = n - \text{rank}(A + I)$.

(3) 若 $\mathbb{F} = \mathbb{C}$, 则 A 一定可以相似成对角方阵.

8. 设方阵 $H = (a_{ij}) \in \mathbb{F}^{n \times n}$ 满足 $a_{ij} = 0, \forall i \geq j + 2$. H 称为 **Hessenberg**^[4] 方阵. 证明:

(1) 对于任意正整数 k , $H^k = (b_{ij})$ 满足 $b_{ij} = \begin{cases} \prod_{t=j}^{j+k-1} a_{t+1,t}, & i = j + k; \\ 0, & i > j + k. \end{cases}$

(2) 任意 $A \in \mathbb{F}^{n \times n}$ 可以在 \mathbb{F} 上相似成 Hessenberg 方阵.

9. 设 $A \in \mathbb{F}^{n \times n}$ 不是纯量方阵. 证明: 若 $a_1, a_2, \dots, a_n \in \mathbb{F}$ 满足 $a_1 + a_2 + \dots + a_n = \text{tr}(A)$, 则存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $P^{-1}AP$ 的对角元素依次是 a_1, a_2, \dots, a_n .

^[4]Karl Adolf Hessenberg, 1904–1959, 德国数学家、工程师.

10. 设 I 是指标集合, $\{A_i \mid i \in I\}$ 是一组两两乘积可交换的复数方阵.
- (1) 证明: 存在复数向量 α 是所有 A_i 的公共特征向量.
 - (2) 证明: 存在可逆复数方阵 P 使得 $P^{-1}A_iP$ 都是上三角方阵.
 - (3) 对实数域 \mathbb{R} 和任意数域 \mathbb{F} 上的方阵, 推广上述结论.

§5.3 Jordan 标准形

定义 5.3.

• n 阶方阵 $J_n(a) = \begin{pmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{pmatrix}$ 称为 **Jordan**^[5]块, 其中空白处元素都是 0.

• 形如 $\text{diag}(J_{n_1}(a_1), J_{n_2}(a_2), \dots, J_{n_k}(a_k))$ 的准对角方阵称为 **Jordan 方阵**.

定理 5.10. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$, 则 A 可以在 \mathbb{F} 上相似成 Jordan 方阵.

证明. 根据定理 5.9, 只需证明 $\varphi_A(x) = (x - \lambda)^n$ 情形. 此时, 也只需证明 $\lambda = 0$ 情形.

对 n 使用数学归纳法. 当 $n = 1$ 时, 结论显然成立. 当 $n \geq 2$ 时, 根据归纳假设, $A \begin{bmatrix} 2 & & \\ & \ddots & \\ & & n \end{bmatrix}$ 可以相似成 Jordan 方阵 $\text{diag}(J_{n_1}(0), \dots, J_{n_k}(0))$. 故 A 可以相似成 $B = \begin{pmatrix} 0 & b_2 & \cdots & \cdots & \cdots & b_n \\ & J_{n_1}(0) & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & J_{n_k}(0) & \end{pmatrix}$,

其中 $n_1 \geq n_2 \geq \dots \geq n_k$. B 可以相似成

$$C = \prod_{j=3}^n T_{1,j-1}(-b_j) \cdot B \cdot \prod_{j=3}^n T_{1,j-1}(b_j) = \begin{pmatrix} 0 & c_1 0 \cdots 0 & c_2 0 \cdots 0 & \cdots & c_k 0 \cdots 0 \\ & J_{n_1}(0) & & & \\ & & J_{n_2}(0) & & \\ & & & \ddots & \\ & & & & J_{n_k}(0) \end{pmatrix}.$$

• 若 $c_1 = 0$, 则 C 可以置换相似成 $\begin{pmatrix} J_{n_1}(0) & & & \\ & 0 & c_2 0 \cdots 0 & \cdots & c_k 0 \cdots 0 \\ & & J_{n_2}(0) & & \\ & & & \ddots & \\ & & & & J_{n_k}(0) \end{pmatrix} = \begin{pmatrix} J_{n_1}(0) & O \\ O & C_1 \end{pmatrix}.$

根据归纳假设, C_1 可以相似成 Jordan 方阵, 故 C 也可以相似成 Jordan 方阵.

• 若 $c_1 \neq 0$, 则 C 可以相似成 $\begin{pmatrix} 0 & 1 0 \cdots 0 & \frac{c_2}{c_1} 0 \cdots 0 & \cdots & \frac{c_k}{c_1} 0 \cdots 0 \\ & J_{n_1}(0) & & & \\ & & J_{n_2}(0) & & \\ & & & \ddots & \\ & & & & J_{n_k}(0) \end{pmatrix} = \begin{pmatrix} J_{n_1+1}(0) & E \\ O & D \end{pmatrix}.$

故可设 $c_1 = 1$. 由 $X = \begin{pmatrix} 0 \cdots 0 & \cdots & 0 \cdots 0 \\ c_2 I_{n_2} & \cdots & c_k I_{n_k} \\ O & \cdots & O \end{pmatrix} \in \mathbb{F}^{(n_1+1) \times (n-n_1)}$ 满足 $J_{n_1+1}(0)X - XD = E$,

得 C 与 $\text{diag}(J_{n_1+1}(0), J_{n_2}(0), \dots, J_{n_k}(0))$ 相似. \square

^[5]Marie Ennemond Camille Jordan, 1838–1922, 法国数学家.



第四章 §4.2 习题 10 是定理 5.10 的另一个证明途径.

定理 5.11. 两个 Jordan 方阵相似的充分必要条件是它们有相同的 Jordan 块 (不考虑块的顺序).

证明. 充分性显然. 设 $A = \text{diag}(J_{m_1}(a_1), \dots, J_{m_p}(a_p))$ 与 $B = \text{diag}(J_{n_1}(b_1), \dots, J_{n_q}(b_q))$ 相似, 则对任意 $\lambda \in \mathbb{F}$ 和正整数 k , 有 $\text{rank}(A - \lambda I)^k = \text{rank}(B - \lambda I)^k$. 注意到

$$\begin{aligned} \text{rank}(A - \lambda I)^k &= \sum_{i=1}^p \text{rank}(J_{m_i}(a_i) - \lambda I)^k = n - \sum_{a_i=\lambda} \min(m_i, k), \\ \text{rank}(A - \lambda I)^{k-1} - \text{rank}(A - \lambda I)^k &= \sum_{a_i=\lambda} \min(m_i, k) - \min(m_i, k-1) = \#\{i \mid a_i = \lambda, m_i \geq k\}. \end{aligned}$$

故

$$\text{rank}(A - \lambda I)^{k-1} - 2\text{rank}(A - \lambda I)^k + \text{rank}(A - \lambda I)^{k+1} = \#\{i \mid a_i = \lambda, m_i = k\}$$

是 A 的 Jordan 块中 $J_k(\lambda)$ 的个数, 也是 B 的 Jordan 块中 $J_k(\lambda)$ 的个数. 因此, A, B 有相同的 Jordan 块. \square

设 $A \in \mathbb{F}^{n \times n}$ 的特征值都属于 \mathbb{F} . 由定理 5.10 和定理 5.11 可知, 存在唯一的 Jordan 方阵 J (不考虑 Jordan 块的顺序) 与 A 相似. J 称为 A 的 **Jordan 标准形**.

定理 5.11 的证明实际上给出了计算 A 的 Jordan 标准形 J 的一种方法, 算法流程如下.

1. 计算 A 的特征多项式, 完全分解成 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$ 的形式, 其中 $\lambda_1, \dots, \lambda_k$ 两两不同.
2. 计算 $r_{i,j} = \text{rank}(A - \lambda_i I)^j$, $1 \leq i \leq k$, $1 \leq j \leq s_i$, 直至 $r_{i,s_i} = n - n_i$.
3. 计算 $t_{i,j} = r_{i,j-1} - 2r_{i,j} + r_{i,j+1}$, $1 \leq i \leq k$, $1 \leq j \leq s_i$, 其中 $r_{i,0} = n$, $r_{i,s_i+1} = n - n_i$.
4. A 的 Jordan 标准形 J 由 $t_{i,j}$ 个 $J_j(\lambda_i)$ 构成, $1 \leq i \leq k$, $1 \leq j \leq s_i$.
5. 求解线性方程组 $AP = PJ$, 得可逆方阵 P 使得 $P^{-1}AP = J$.

例 5.7. 计算如下复数方阵 A 的 Jordan 标准形.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ -2 & 2 & 1 & 1 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 2 & 1 \\ 1 & -1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

解答. 首先计算 A 的特征多项式, 得 $\varphi_A(x) = (x-1)^7$. 然后计算 $r_k = \text{rank}(I - A)^k$, 得 $r_1 = 4$, $r_2 = 1$, $r_3 = 0$. 对序列 $(7, 4, 1, 0)$ 作两次“前项减后项”的运算, 得 $(3, 3, 1, 0)$ 和 $(0, 2, 1, 0)$. 故 A 的 Jordan 标准形 $J = \text{diag}(J_2(1), J_2(1), J_3(1))$. \square

对于一个 Jordan 块, 我们可以很容易地计算其矩阵方幂和多项式. 设 $a \in \mathbb{F}$, $f \in \mathbb{F}[x]$, 并且 $\text{Char } \mathbb{F} = 0$, 则

$$f(J_n(a)) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (J_n(0))^k = \begin{pmatrix} f(a) & f'(a) & \cdots & \frac{f^{(n-1)}(a)}{(n-1)!} \\ & f(a) & \ddots & \vdots \\ & & \ddots & f'(a) \\ & & & f(a) \end{pmatrix}.$$

从而, 对于任意方阵 A , 我们都可以利用 Jordan 标准形分解 $A = P \text{diag}(J_{n_1}(a_1), \dots, J_{n_k}(a_k)) P^{-1}$, 算出 A 的矩阵方幂和多项式 $f(A) = P \text{diag}(f(J_{n_1}(a_1)), \dots, f(J_{n_k}(a_k))) P^{-1}$.

定义 5.4. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值都属于 \mathbb{F} , $J = \text{diag}(J_{n_1}(a_1), \dots, J_{n_k}(a_k))$ 是 A 的 Jordan 标准形, 则每个 Jordan 块 $J_{n_i}(a_i)$ 的特征多项式 $(x - a_i)^{n_i} \in \mathbb{F}[x]$ 称为 A 的一个初等因子, A 的所有初等因子构成 A 的初等因子组.

易见, 初等因子组是相似不变量. 两个方阵相似当且仅当它们的初等因子组相同.

定理 5.12. 设 $\{(x - \lambda_i)^{m_{ij}} \mid 1 \leq i \leq k, 1 \leq j \leq s_i\}$ 是方阵 A 的初等因子组, 其中 $\lambda_1, \dots, \lambda_k$ 是 A 的所有两两不同的特征值, 则有如下结论.

1. A 的特征值 λ_i 的代数重数 $= \sum_{j=1}^{s_i} m_{ij}$, 几何重数 $= s_i$.
2. A 可以相似成对角方阵的充分必要条件是 $m_{ij} = 1, \forall i, j$.
3. $\varphi_A(x) = d_A(x)$ 的充分必要条件是 $s_i = 1, \forall i$, 其中 d_A 是 A 的最小多项式 (见 §5.4 定义 5.5).

例 5.8. 设方阵 A 的初等因子组为 $x^2, x^2, x+1, (x+1)^2, x-1, (x-1)^3$, 求 $\varphi_A(x)$ 和 $d_A(x)$.

解答. $\varphi_A(x)$ 是所有初等因子的乘积, $\varphi_A(x) = x^4(x+1)^3(x-1)^4$. 根据例 5.13 和定理 5.13 结论 3, $d_A(x)$ 是所有初等因子的最小公倍式, $d_A(x) = x^2(x+1)^2(x-1)^3$. \square

下面是 Jordan 标准形的一些应用例子.

例 5.9. 任意复数方阵 A 与 A^T 相似.

证明. 设 A 与 $\text{diag}(J_{n_1}(a_1), \dots, J_{n_k}(a_k))$ 相似, 则 A^T 与 $\text{diag}(J_{n_1}(a_1)^T, \dots, J_{n_k}(a_k)^T)$ 相似. 注意到 $J_{n_i}(a_i)^T = S J_{n_i}(a_i) S$ 与 $J_{n_i}(a_i)$ 相似, 其中 $S = \begin{pmatrix} & & 1 \\ & \ddots & \\ & & 1 \end{pmatrix}$ 是 n_i 阶方阵. 故 A 与 A^T 相似. \square

例 5.10. 对于任意可逆复数方阵 A 和正整数 m , 存在复数方阵 B 使得 $B^m = A$.

解答. 根据复数方阵的 Jordan 标准形, 只需考虑 $A = J_n(\lambda)$, $\lambda \neq 0$ 情形. 设 $\mu \in \mathbb{C}$ 满足 $\mu^k = \lambda$, 则 λ 是 $M = (J_n(\mu))^m$ 的 n 重特征值. 由 $\text{rank}(M - \lambda I) = n - 1$ 得 $J_n(\lambda)$ 是 M 的 Jordan 标准形. 故存在可逆复数方阵 P 使得 $P^{-1} M P = A$, $B = P^{-1} J_n(\mu) P$ 满足 $B^m = A$. \square

例 5.11. 对于任意复数方阵 A , $\lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{k!} A^k$ 存在, 记作 $e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$. 映射 $\exp: A \mapsto e^A$ 称为矩阵的指数函数. 矩阵的指数函数具有下列常用性质:

1. 当 $AB = BA$ 时, $e^{A+B} = e^A e^B$.

$$2. (e^A)^T = e^{(A^T)}, (e^A)^{-1} = e^{-A}.$$

$$3. \det(e^A) = e^{\operatorname{tr}(A)}.$$

证明. 对于任意 $X = (x_{ij}) \in \mathbb{C}^{n \times n}$, 记 $\|X\| = \max_{1 \leq i, j \leq n} |x_{ij}|$. 容易验证, $\|X^k\| \leq (n\|X\|)^k$. 从而,

$$\text{当 } m \rightarrow +\infty \text{ 时, } \left\| \sum_{k=m}^M \frac{1}{k!} A^k \right\| \leq \sum_{k=m}^M \frac{(n\|A\|)^k}{k!} \leq \frac{(n\|A\|)^m}{m!(1 - \frac{n\|A\|}{m})} \rightarrow 0.$$

根据 Cauchy 收敛准则, $\sum_{k=0}^{\infty} \frac{1}{k!} A^k$ 收敛. 下面证明矩阵的指数函数的性质.

$$1. \text{ 当 } AB = BA \text{ 时, } e^{A+B} = \sum_{k=0}^{\infty} \frac{1}{k!} (A+B)^k = \sum_{k=0}^{\infty} \sum_{i=0}^k \frac{C_k^i}{k!} A^i B^{k-i} = \sum_{i,j=0}^{\infty} \frac{1}{i!j!} A^i B^j = e^A e^B.$$

$$2. \text{ 由定义可得 } (e^A)^T = e^{(A^T)}. \text{ 由性质 1 可得 } e^A e^{-A} = I, \text{ 故 } (e^A)^{-1} = e^{-A}.$$

$$3. \text{ 设 } A = P \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ & & \lambda_n \end{pmatrix} P^{-1}, \text{ 则 } e^A = P \begin{pmatrix} e^{\lambda_1} & * & * \\ & \ddots & * \\ & & e^{\lambda_n} \end{pmatrix} P^{-1}.$$

$$\text{故 } \det(e^A) = \prod_{i=1}^n e^{\lambda_i} = e^{\operatorname{tr}(A)}. \quad \square$$

例 5.12. 求解常系数微分方程 $y''(x) = ay'(x) + by(x)$.

解答. 二阶微分方程 $y''(x) = ay'(x) + by(x)$ 可以表示为如下一阶微分方程组的形式,

$$Z'(x) = A \cdot Z(x) \quad (5.2)$$

其中 $Z(x) = \begin{pmatrix} y'(x) \\ y(x) \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$. 容易验证, $Z(x) = e^{xA} Z(0)$ 是 (5.2) 的解. 根据例 5.5,

$$\bullet \text{ 当 } a^2 + 4b \neq 0 \text{ 时, } \lambda_{1,2} = \frac{a}{2} \pm \sqrt{\frac{a^2}{4} + b},$$

$$A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}, e^{xA} = P \begin{pmatrix} e^{\lambda_1 x} & 0 \\ 0 & e^{\lambda_2 x} \end{pmatrix} P^{-1}, \text{ 得 } y(x) \text{ 形如 } c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$

$$\bullet \text{ 当 } a^2 + 4b = 0 \text{ 时, } \lambda = \frac{a}{2},$$

$$A = P \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} P^{-1}, e^{xA} = P \begin{pmatrix} e^{\lambda x} & x e^{\lambda x} \\ 0 & e^{\lambda x} \end{pmatrix} P^{-1}, \text{ 得 } y(x) \text{ 形如 } (c_1 + c_2 x) e^{\lambda x}. \quad \square$$

由以上几个例子可以看出, 通过 Jordan 标准形, 微积分中的许多结果可以应用于矩阵, 从而建立矩阵的微积分运算.

习题

1. 求下列复数方阵的 Jordan 标准形.

$$\begin{array}{l}
 (1) \begin{pmatrix} 1 & 2 & -2 & 2 & -4 \\ 0 & -1 & 2 & -2 & 4 \\ 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 (4) \begin{pmatrix} 0 & -1 & 1 & 3 & -4 \\ 1 & 2 & -3 & -5 & 4 \\ 0 & 0 & -1 & 0 & -2 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 (2) \begin{pmatrix} 1 & -2 & 1 & 0 & -2 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \\
 (5) \begin{pmatrix} -1 & 2 & 3 & -5 & 3 \\ 0 & 1 & -1 & -1 & 3 \\ 0 & 0 & 2 & -3 & 2 \\ 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -1 & 1 \end{pmatrix} \\
 (3) \begin{pmatrix} -1 & 2 & -1 & 3 & 3 \\ 0 & 1 & -2 & 2 & 3 \\ 0 & 0 & -1 & 2 & 4 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 (6) \begin{pmatrix} 1 & -2 & 2 & 0 & 4 \\ 0 & -1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & -2 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}
 \end{array}$$

2. 证明定理 5.12.

3. 设复数方阵 A 的特征值都是 1. 证明: 对于任意正整数 k , A 与 A^k 相似.

4. 设复数 A 是可逆方阵, 并且存在正整数 $i \neq j$ 使得 A^i 与 A^j 相似. 证明: 存在正整数 k 使得 A^k 的特征值都是 1.

5. 求与 A^2 相似的所有复数方阵 A .

6. 设复数方阵 $A = \text{diag}(J_{n_1}(0), J_{n_2}(0), \dots, J_{n_k}(0))$, 其中 $n_1 \geq n_2 \geq \dots \geq n_k$. 求与 A 乘积可交换的所有复数方阵 B .

7. 设 n 阶复数方阵 A, B 满足: 与 A 乘积可交换的复数方阵 C 都与 B 乘积可交换. 证明: 存在 $f \in \mathbb{C}[x]$ 使得 $B = f(A)$.

8. 证明: 任意复数方阵 A 可唯一地表示为 $A = B + C$ 的形式, 其中 B 可以相似成对角方阵, C 是幂零方阵, 并且 $BC = CB$. 这种表示形式称为 A 的 **Jordan 分解**.

9. (1) 设 $J_{n_1}(0), \dots, J_{n_k}(0)$ 是复数方阵 A 的所有幂零的 Jordan 块, m 是正整数. 证明: 下面两个叙述等价.

① 存在复数方阵 B 使得 $B^m = A$.

② 可以把 n_1, \dots, n_k 排成 m 行 $\lceil \frac{k}{m} \rceil$ 列 (补充一些 0), 使得每列中任意两数的差 ≤ 1 .

(2) 对于哪些实数方阵 A , 存在实数方阵 B 使得 $B^m = A$? 证明你的结论.

10. 设 A 是 n 阶复数方阵. 证明: $e^A = \lim_{k \rightarrow +\infty} (I + \frac{1}{k}A)^k$.

11. 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶复数方阵 A 的所有特征值. $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$ 称为 A 的谱半径.

(1) 证明: $\lim_{k \rightarrow +\infty} A^k = O$ 的充分必要条件是 $\rho(A) < 1$.

(2) 设 r 是幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 的收敛半径. 证明: 当 $\rho(A) < r$ 时, 矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 收敛.

12. (1) 求所有 $X \in \mathbb{R}^{n \times n}$ 使得 $e^X = I$.

(2) 对哪些实数方阵 $A \in \mathbb{R}^{n \times n}$, 存在 $X \in \mathbb{R}^{n \times n}$ 使得 $e^X = A$? 证明你的结论.

(3) 对哪些实数方阵 $A \in \mathbb{R}^{n \times n}$, 存在 $X \in \mathbb{C}^{n \times n}$ 使得 $e^X = A$? 证明你的结论.

§5.4 最小多项式

设 A 是 \mathbb{F} 上的方阵. 在 §5.3 中, 我们证明了, 当 A 的特征值都在 \mathbb{F} 中时, A 可以在 \mathbb{F} 上相似成 Jordan 标准形. 当 A 有特征值不在 \mathbb{F} 中时, 结论如何? 在 §5.4 和 §5.5 中, 我们将借助多项式来研究相似标准形问题.

定义 5.5. 设 $A \in \mathbb{F}^{n \times n}$. 若 $f \in \mathbb{F}[x]$ 满足 $f(A) = O$, 则 f 称为 A 在 \mathbb{F} 上的一个化零多项式.

若 $f, g \in \mathbb{F}[x]$ 都是 A 的化零多项式, 则根据 Bézout 定理, 存在 $u, v \in \mathbb{F}[x]$ 使得 $\gcd(f, g) = uf + vg$ 也是 A 的化零多项式.

A 在 \mathbb{F} 上的所有化零多项式的最大公因式, 是次数最小的首一化零多项式, 称为 A 在 \mathbb{F} 上的最小多项式, 记作 d_A .

 设 \mathbb{K} 是 \mathbb{F} 的任意扩域, $A \in \mathbb{F}^{n \times n}$ 也可以看作是 $A \in \mathbb{K}^{n \times n}$. d_A 由 A 唯一确定, 与 \mathbb{K} 无关.

定理 5.13. 关于 \mathbb{F} 上方阵的最小多项式, 有如下结论.

1. 若 A 与 B 相似, 则 $d_A = d_B$.
2. 对于任意 n 阶方阵 A , 有 d_A 整除 φ_A , φ_A 整除 $(d_A)^n$.
3. 设 $A = \text{diag}(A_1, A_2, \dots, A_k)$, 每个 A_i 是方阵, 则 $d_A = \text{lcm}(d_{A_1}, d_{A_2}, \dots, d_{A_k})$.^[6]

证明.

1. 设 $A = PBP^{-1}$, 则 $f(A) = Pf(B)P^{-1}$, $\forall f \in \mathbb{F}[x]$. $f(A) = O \Leftrightarrow f(B) = O$. 故 $d_A = d_B$.
2. 根据 Cayley-Hamilton 定理, d_A 整除 φ_A . 对于 A 的任意特征值 $\lambda_i \in \mathbb{K}$, 设 $\alpha \in \mathbb{K}^{n \times 1}$ 是 λ_i 对应的一个特征向量. 由 $A\alpha = \lambda_i\alpha$ 可得 $d_A(A)\alpha = d_A(\lambda_i)\alpha = \mathbf{0}$. 因此, $d_A(\lambda_i) = 0$, $x - \lambda_i$ 整除 $d_A(x)$. 从而有 $\varphi_A(x) = \prod_{i=1}^n (x - \lambda_i)$ 整除 $(d_A(x))^n$.
3. 设 $f \in \mathbb{F}[x]$. $f(A) = O \Leftrightarrow f(A_i) = O \Leftrightarrow d_{A_i}$ 整除 $f \Leftrightarrow \text{lcm}(d_{A_1}, d_{A_2}, \dots, d_{A_k})$ 整除 f . 因此, $d_A = \text{lcm}(d_{A_1}, d_{A_2}, \dots, d_{A_k})$. \square

例 5.13. 求 Jordan 块 $A = J_n(a)$ 的最小多项式.

解答. $\varphi_A(x) = (x-a)^n$. 由定理 5.13 结论 2, 得 $d_A(x)$ 形如 $(x-a)^m$. 当 $m < n$ 时, $(A-aI)^m \neq O$. 因此, $d_A(x) = (x-a)^n = \varphi_A(x)$. \square

例 5.14. 求友方阵 $A = \begin{pmatrix} & & & -a_0 \\ & & & -a_1 \\ & & \ddots & \vdots \\ & & & 1 & -a_{n-1} \\ 1 & & & & \end{pmatrix}$ 的最小多项式.

解答. 对于任意多项式 $f(x) = \sum_{i=0}^{n-1} c_i x^i$, 有 $f(A)\mathbf{e}_1 = \sum_{i=0}^{n-1} c_i A^i \mathbf{e}_1 = \sum_{i=0}^{n-1} c_i \mathbf{e}_{i+1} = (c_0 \ c_1 \ \cdots \ c_{n-1})^T$.

故 $\deg(d_A) \geq n$. 因此, $d_A(x) = \varphi_A(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$. \square

^[6] $\text{lcm}(\dots)$ 表示若干个多项式的最小公倍式, 或若干个整数的最小公倍数.

例 5.15. 求实数方阵 $A = \begin{pmatrix} 1 & 2 & \cdots & n \\ 2 & 3 & \cdots & n+1 \\ \vdots & \vdots & \cdots & \vdots \\ n & n+1 & \cdots & 2n-1 \end{pmatrix}$ 的最小多项式, $n \geq 2$.

解答. 由满秩分解

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ n & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & n-1 \end{pmatrix} = BC$$

可得

$$\varphi_A(x) = x^{n-2} \det(xI - CB) = x^{n-1} \left(x^2 - n^2x - \frac{n^4 - n^2}{12} \right).$$

A 的所有特征值为 $\lambda_1, \lambda_2, 0, \dots, 0$, 其中 $\lambda_1 \neq \lambda_2$ 且 $\lambda_1 \lambda_2 \neq 0$. 由定理 5.9 以及 $\text{rank}(A) = 2$, 得 A 与 $\text{diag}(\lambda_1, \lambda_2, O)$ 相似. 因此,

$$d_A(x) = x(x - \lambda_1)(x - \lambda_2) = x^3 - n^2x^2 - \frac{n^4 - n^2}{12}x. \quad \square$$

定理 5.14. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值都属于 \mathbb{F} , 则 A 可以在 \mathbb{F} 上相似成对角方阵的充分必要条件是 $d_A(x)$ 无重根.

证明. 设 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同. 根据定理 5.9, 存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $A = P \text{diag}(A_1, A_2, \dots, A_k) P^{-1}$, 其中 $\varphi_{A_i}(x) = (x - \lambda_i)^{n_i}$.

(充分性) 设 $d_A(x)$ 无重根. 根据定理 5.13, $d_A(x) = \prod_{i=1}^k (x - \lambda_i)$. 当 $i \neq j$ 时, $A_i - \lambda_j I$ 是可逆方阵. $d_A(A) = O \Rightarrow d_A(A_i) = O \Rightarrow A_i - \lambda_i I = O$. 因此, $P^{-1}AP$ 是对角方阵.

(必要性) 若 A 可以相似成对角方阵, 则 A 与 $B = \text{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k})$ 相似. 因此, $d_A(x) = d_B(x) = \prod_{i=1}^k (x - \lambda_i)$ 无重根. \square

例 5.16. 任意置换方阵可以在复数域上相似成对角方阵.

证明. 设 A 是 n 阶置换方阵. A^k ($k \in \mathbb{N}$) 都是置换方阵. n 阶置换方阵只有 $n!$ 个, 故存在正整数 m 使得 $A^m = I$. A 的化零多项式 $x^m - 1$ 无重根, 故 $d_A(x)$ 也无重根. 根据定理 5.14, A 可以在复数域上相似成对角方阵. \square

定义 5.6. 设 $A \in \mathbb{F}^{n \times n}$, $\alpha \in \mathbb{F}^{n \times 1}$ 是非零向量. 若 $f \in \mathbb{F}[x]$ 满足 $f(A)\alpha = \mathbf{0}$, 则 f 称为 A 在 \mathbb{F} 上关于 α 的一个化零多项式.

若 f 和 g 都是 A 关于 α 的化零多项式, 则 $\text{gcd}(f, g)$ 也是 A 关于 α 的化零多项式.

A 在 \mathbb{F} 上关于 α 的所有化零多项式的最大公因式, 是次数最小的关于 α 的首一化零多项式, 称为 A 在 \mathbb{F} 上关于 α 的最小多项式, 记作 $d_{A, \alpha}$.

显然, $d_{A, \alpha}$ 整除 d_A . 此外, $d_{A, \alpha}$ 由 A 和 α 唯一确定, 与它们所在的数域无关.

定理 5.15. 对于任意 $A \in \mathbb{F}^{n \times n}$, 存在 $\alpha \in \mathbb{F}^{n \times 1}$ 使得 $d_{A, \alpha} = d_A$.

证明. 设 $d_A = p_1^{n_1} \cdots p_k^{n_k}$, 其中 $p_1, \cdots, p_k \in \mathbb{F}[x]$ 是两两互素的首一不可约多项式, $\deg(p_i)$ 和 n_i 都是正整数, $\forall i$. 对于每个 $f_i = \frac{d_A}{p_i}$, 存在 $\alpha_i \in \mathbb{F}^{n \times 1}$ 使得 $f_i(A)\alpha_i \neq \mathbf{0}$. 设 $g_i = \frac{d_A}{p_i^{n_i}}$, $\beta_i = g_i(A)\alpha_i$. 由 $p_i^{n_i}(A)\beta_i = \mathbf{0}$ 和 $p_i^{n_i-1}(A)\beta_i \neq \mathbf{0}$, 得 $d_{A,\beta_i} = p_i^{n_i}$. 设 $\alpha = \beta_1 + \cdots + \beta_k$, $d_{A,\alpha} = p_1^{m_1} \cdots p_k^{m_k}$, 其中 $m_i \leq n_i, \forall i$. 若存在 $m_i \neq n_i$, 则 $d_{A,\alpha} \nmid f_i$, 得 $\mathbf{0} = f_i(A)\alpha = f_i(A)\beta_i$, 与 $d_{A,\beta_i} = p_i^{n_i} \nmid f_i$ 矛盾. 因此, $d_{A,\alpha} = d_A$. \square

定理 5.16. 若 $A \in \mathbb{F}^{n \times n}$ 满足 $d_A = \varphi_A$, 则 A 与 φ_A 的友方阵相似.

证明. 根据定理 5.15, 存在 $\alpha_1 \in \mathbb{F}^{n \times 1}$ 使得 $d_{A,\alpha_1} = d_A = \varphi_A = x^n + \sum_{i=0}^{n-1} c_i x^i$. 记 $\alpha_k = A^{k-1}\alpha_1$, $k \geq 2$. 由 $\varphi_A(A) = O$ 可得 $A\alpha_n = A^n\alpha_1 = -\sum_{i=0}^{n-1} c_i A^i \alpha_1 = -\sum_{i=1}^n c_{i-1} \alpha_i$. 设 $P = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)$, 则 $AP = PB$, 其中 $B = \begin{pmatrix} & & & -c_0 \\ & & & -c_1 \\ & & & \vdots \\ & & & 1 \ -c_{n-1} \\ 1 & & & \\ & \ddots & & \\ & & & \end{pmatrix}$ 是 $\varphi_A(x)$ 的友方阵. 下证 P 是可逆方阵. 若 $\sum_{i=1}^n t_i \alpha_i = \mathbf{0}$, 则 $f(x) = \sum_{i=1}^n t_i x^{i-1}$ 是 A 关于 α_1 的化零多项式. 由 d_{A,α_1} 整除 f 且 $\deg(f) < n$, 得 $f = 0$, 即 $Px = \mathbf{0}$ 只有零解. 故 P 是可逆方阵, 从而 A 与 B 相似. \square

由定理 5.16 的证明, 我们可以得到如下结论. 若 $A \in \mathbb{F}^{n \times n}$ 和 $\alpha \in \mathbb{F}^{n \times 1}$ 满足 $d_{A,\alpha} = \varphi_A$, 则 $\mathbb{F}^{n \times 1}$ 中的所有向量都可以唯一地表示成 $p(A)\alpha$ 的形式, 其中 $p \in \mathbb{F}[x]$ 且 $\deg(p) \leq n-1$.

例 5.17. 设 $A, B \in \mathbb{F}^{n \times n}$ 满足 $AB = BA$ 且 $d_A = \varphi_A$. 证明: 存在 $f \in \mathbb{F}[x]$ 使得 $B = f(A)$.

证明. 根据定理 5.15, 存在 $\alpha \in \mathbb{F}^{n \times 1}$ 满足 $d_{A,\alpha} = d_A = \varphi_A$. 设 $B\alpha = f(A)\alpha$, 其中 $f \in \mathbb{F}[x]$. 对于任意 $\beta \in \mathbb{F}^{n \times 1}$, 设 $\beta = p(A)\alpha$, 其中 $p \in \mathbb{F}[x]$, 则有

$$B\beta = Bp(A)\alpha = p(A)B\alpha = p(A)f(A)\alpha = f(A)p(A)\alpha = f(A)\beta.$$

因此, $B = f(A)$. \square

习题

1. 求下列方阵的特征多项式和最小多项式.

$$\begin{aligned} (1) & \begin{pmatrix} 3 & -2 & -1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} & (2) & \begin{pmatrix} -1 & 2 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix} & (3) & \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 2 & -2 & 1 \end{pmatrix} & (4) & \begin{pmatrix} -4 & 3 & -5 \\ -1 & 0 & -1 \\ 2 & -2 & 3 \end{pmatrix} \\ (5) & \begin{pmatrix} & I_{n-k} \\ I_k & \end{pmatrix} & (6) & \begin{pmatrix} & I_{n-k} \\ -I_k & \end{pmatrix} & (7) & \begin{pmatrix} & I_{n-k} \\ O_k & \end{pmatrix} & (8) & \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & & -1 & 0 \end{pmatrix} \end{aligned}$$

$$(9) \begin{pmatrix} a_1 + b_1 & a_1 + b_2 & \cdots & a_1 + b_n \\ a_2 + b_1 & a_2 + b_2 & \cdots & a_2 + b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n + b_1 & a_n + b_2 & \cdots & a_n + b_n \end{pmatrix} \quad (10) \begin{pmatrix} 1 + a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & 1 + a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \cdots & 1 + a_n b_n \end{pmatrix}$$

2. 设 $\text{Char } \mathbb{F} = 0$, $f \in \mathbb{F}[x]$ 在 \mathbb{F} 的扩域 \mathbb{K} 上可分解为 $f = \prod_{i=1}^n (x - \lambda_i)$. 证明: 在下列 4 个叙述中, ①是④的充分必要条件, ②是④的充分不必要条件, ③是④的既不充分也不必要条件.

① $\gcd(f, f') = 1$. ② f 在 $\mathbb{F}[x]$ 中不可约. ③ $\lambda_1, \dots, \lambda_n \notin \mathbb{F}$. ④ $\lambda_1, \dots, \lambda_n$ 两两不同.

3. 证明: 对于任意 n 阶置换方阵 A , 存在置换方阵 P 使得 $P^{-1}AP = \text{diag}(C_1, C_2, \dots, C_k)$, 其中每个 C_i 形如 $\begin{pmatrix} I_{n_i} \\ 1 \end{pmatrix}$. 由此可得, A 的复数特征值都是次数不超过 n 的单位根.

4. 设 $f_1, f_2, \dots, f_k \in \mathbb{F}[x]$ 是两两互素的首一多项式, A_i 是 f_i 的友方阵. 证明: $f_1 f_2 \cdots f_k$ 的友方阵与 $\text{diag}(A_1, A_2, \dots, A_k)$ 相似.

5. 设 $A \in \mathbb{F}^{n \times n}$, $\alpha \in \mathbb{F}^{n \times 1}$ 是非零向量. 证明: $\text{rank}(\alpha \ A \ \alpha \ \cdots \ A^{k-1} \alpha) = \min(k, \deg(d_{A, \alpha}))$.

6. (1) 设 $A \in \mathbb{F}^{n \times n}$, $\alpha, \beta \in \mathbb{F}^{n \times 1}$ 满足 $\gcd(d_{A, \alpha}, d_{A, \beta}) = 1$. 证明: $d_{A, \alpha + \beta} = d_{A, \alpha} d_{A, \beta}$.

(2) 试把上述结论推广到多个向量 $\alpha_1, \alpha_2, \dots, \alpha_k$.

7. 设 Hessenberg 方阵 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 满足 $\prod_{i=2}^n a_{i, i-1} \neq 0$. 证明: $d_A = \varphi_A$.

8. 设 $B \in \mathbb{F}^{n \times n}$ 满足 $d_B = \varphi_B$ 在 $\mathbb{F}[x]$ 中不可约, $A = \begin{pmatrix} B & I & & \\ & B & \ddots & \\ & & \ddots & I \\ & & & B \end{pmatrix}$. 证明: $d_A = \varphi_A$.

9. 设 $A \in \mathbb{F}^{n \times n}$ 满足 $d_A = \varphi_A$, $f \in \mathbb{F}[x]$, $g = \gcd(d_A, f)$. 证明: $\text{rank}(f(A)) = n - \deg(g)$.

10. 证明: 对于任意 $A \in \mathbb{F}^{n \times n}$, 存在 $f \in \mathbb{F}[x]$ 使得 A 的伴随方阵 $A^* = f(A)$.

§5.5 特征方阵

定义 5.7. 设 $A \in \mathbb{F}^{n \times n}$. 多项式方阵 $xI - A \in \mathbb{F}[x]^{n \times n}$ 称为 A 的特征方阵.

定理 5.17. 设 $A, B \in \mathbb{F}^{n \times n}$. A 与 B 在 \mathbb{F} 上相似的充分必要条件是 $xI - A$ 与 $xI - B$ 在 $\mathbb{F}[x]$ 上模相抵.

证明. (必要性) 设可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $A = PBP^{-1}$, 则 P 是 $\mathbb{F}[x]$ 上的模方阵, 使得 $xI - A = P(xI - B)P^{-1}$. 故 $xI - A$ 与 $xI - B$ 在 $\mathbb{F}[x]$ 上模相抵.

(充分性) 设 $xI - A = P(xI - B)Q$, 其中 P, Q 是 $\mathbb{F}[x]$ 上的模方阵. 设

$$P = \sum_{i=0}^m P_i x^i, \quad Q = \sum_{i=0}^m Q_i x^i, \quad R = Q^{-1} = \sum_{i=0}^m R_i x^i,$$

其中 $P_i, Q_i, R_i \in \mathbb{F}^{n \times n}$, m 是某个充分大的正整数. 由 $(xI - A)R = P(xI - B)$ 可得

$$R_{j-1} - AR_j = P_{j-1} - P_j B, \quad \forall j.$$

从而,

$$\sum_{j=0}^{m+1} (R_{j-1} - AR_j) B^j = \sum_{j=0}^{m+1} (P_{j-1} - P_j B) B^j = O.$$

故 $X = \sum_{j=0}^m R_j B^j$ 满足 $AX = XB$, 进而 $A^i X = X B^i, \forall i$. 另由 $\sum_{i,j=0}^m Q_i R_j x^{i+j} = I$ 可得

$$Q_0 R_0 = I, \quad \sum_{i=0}^k Q_i R_{k-i} = O, \quad \forall k \geq 1.$$

从而,

$$\sum_{i=0}^m Q_i A^i X = \sum_{i=0}^m Q_i X B^i = \sum_{i,j=0}^m Q_i R_j B^{i+j} = I.$$

即 X 是可逆方阵, $X^{-1} = \sum_{i=0}^m Q_i A^i$. 因此, $A = X B X^{-1}$, A 与 B 相似. \square

与定理 5.17 的证明类似, 我们有 Cayley-Hamilton 定理的如下证明.

例 5.18. 设 $A \in \mathbb{F}^{n \times n}$, $\varphi_A(x) = \sum_{i=0}^n c_i x^i$, $(xI - A)^* = \sum_{i=0}^{n-1} P_i x^i$. 由 $(xI - A)^*(xI - A) = \varphi_A(x)I$ 可得 $P_{i-1} - P_i A = c_i I, \forall i$. 从而, $\varphi_A(A) = \sum_{i=0}^n c_i A^i = \sum_{i=0}^n (P_{i-1} - P_i A) A^i = O$.

根据定理 5.17, 我们还可以证明如下结论.

例 5.19 (例 5.9 的推广). 对于任意 $A \in \mathbb{F}^{n \times n}$, $xI - A$ 与 $xI - A^T$ 有相同的 Smith 标准形. 因此, A 与 A^T 在 \mathbb{F} 上相似.

定理 5.18. 设 $A \in \mathbb{F}^{n \times n}$, $\varphi_A = \prod_{i=1}^k p_i^{n_i}$, 其中 $p_1, \dots, p_k \in \mathbb{F}[x]$ 是两两互素的首一不可约多项式.

1. 设 d_s, \dots, d_n 是 $xI - A$ 的所有 $\neq 1$ 的不变因子, 则 A 与 $\text{diag}(A_s, \dots, A_n)$ 在 \mathbb{F} 上相似, 其中每个 A_j 是 d_j 的友方阵.

2. d_n 是 A 的最小多项式.

3. 设 $d_j = \prod_{i=1}^k p_i^{m_{ij}}$, 则 A_j 与 $B_j = \text{diag}(B_{1j}, \dots, B_{kj})$ 在 \mathbb{F} 上相似, 其中每个 B_{ij} 是 $p_i^{m_{ij}}$ 的友方阵.

4. B_{ij} 与 $m_{ij} \times m_{ij}$ 分块矩阵 $M_{ij} = \begin{pmatrix} C_i & & & & \\ E & C_i & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & E & C_i \end{pmatrix}$ 在 \mathbb{F} 上相似, 其中 C_i 是 p_i 的友方阵, E 是右上角元素为 1、其它元素为 0 的基础方阵.

证明.

1. 注意到 $\varphi_A = \prod_{s \leq j \leq n} d_j$, $B = \text{diag}(A_s, \dots, A_n)$ 也是 n 阶方阵. 由例 4.15 可得 $xI - A_j$ 的不变因子为 $1, \dots, 1, d_j$. 故 $xI - A$ 与 $xI - B$ 的不变因子相同, $xI - A$ 与 $xI - B$ 模相抵. 由定理 5.17, 得 A 与 B 相似.

2. 根据定理 5.13、例 5.15 和结论 1, $d_A(x) = \text{lcm}(d_{A_s}, \dots, d_{A_n}) = \text{lcm}(d_s, \dots, d_n) = d_n$.

3. $xI - B_{ij}$ 的不变因子为 $1, \dots, 1, p_i^{m_{ij}}$, 故 $xI - B_j$ 与 $\text{diag}(1, \dots, 1, p_1^{m_{1j}}, \dots, p_k^{m_{kj}})$ 模相抵, 与 $\text{diag}(1, \dots, 1, d_j)$ 模相抵, 与 $xI - A_j$ 模相抵. 由定理 5.17, A_j 与 B_j 相似.

4. 设 M_{ij} 是 r 阶方阵. 注意到 $xI - M_{ij}$ 的左下角 $r-1$ 阶子式为 $(-1)^{r-1}$, 故 $xI - M_{ij}$ 与 $\text{diag}(1, \dots, 1, p_i^{m_{ij}})$ 模相抵, 与 $xI - C_{ij}$ 模相抵. 由定理 5.17, C_{ij} 与 M_{ij} 相似. \square

定义 5.8. 定理 5.18 中由友方阵构成的准对角方阵 $\text{diag}(A_s, \dots, A_n)$ 或 $\text{diag}(\dots, B_{ij}, \dots)$ 称为 A 的有理标准形或 Frobenius 标准形, 对应于 \mathbb{F}^n 的循环子空间分解, 详见第九章 §9.7. 每个 $p_i^{m_{ij}} \neq 1$ 称为 $xI - A$ 的一个初等因子, $xI - A$ 的所有初等因子构成 $xI - A$ 的初等因子组. 把 $xI - A$ 的初等因子组重新编号排列如下表所示 (每行每列都是一个因式分解).

φ_A	d_n	d_{n-1}	\dots	d_1
$p_1^{n_1}$	$p_1^{m_{11}}$	$p_1^{m_{12}}$	\dots	$p_1^{m_{1n}}$
$p_2^{n_2}$	$p_2^{m_{21}}$	$p_2^{m_{22}}$	\dots	$p_2^{m_{2n}}$
\vdots	\vdots	\vdots	\dots	\vdots
$p_k^{n_k}$	$p_k^{m_{k1}}$	$p_k^{m_{k2}}$	\dots	$p_k^{m_{kn}}$

设 $m_{i1} \geq \dots \geq m_{i,s_i} > m_{i,s_i+1} = \dots = m_{in} = 0$, 则 A 可相似成准对角方阵

$$M = \text{diag}(M_{11}, \dots, M_{1,s_1}, \dots, M_{k1}, \dots, M_{k,s_k}).$$

当 A 的所有特征值 $\lambda_j \in \mathbb{F}$ 时, $p_j = x - \lambda_j$, $M_{ij}^T = J_{m_{ij}}(\lambda_j)$ 是 Jordan 块, M 是 A 的 Jordan 标准形, $xI - A$ 的初等因子组与 A 的初等因子组也完全相同, 两个概念是一致的. 因此, M_{ij} 和 M 可以看作是 Jordan 块和 Jordan 方阵概念的推广. 定理 5.18 完全解决了任意数域 \mathbb{F} 上相似等价类的标准形问题.

例 5.20. 已知实数方阵 A 的特征方阵 $xI - A$ 与 $\text{diag}(I, (x^2 - x)^2, (x^2 + x)^3, (x^2 - 1)^4)$ 相抵, 求 A 的 Jordan 标准形.

解答. $xI - A$ 的初等因子组为 $x^2, (x-1)^2, x^3, (x+1)^3, (x+1)^4, (x-1)^4$. 故 A 的 Jordan 标准形为 $\text{diag}(J_2(0), J_3(0), J_3(-1), J_4(-1), J_2(1), J_4(1))$. \square

$xI - A$ 的每个初等因子 $p_j^{m_{ij}}$ 所对应的方阵 M_{ij} 可以有多种选取方式, 并不局限于定理 5.18 中的形式.

例 5.21. 任意 n 阶实数方阵 A 在 \mathbb{R} 中可以相似成 $\text{diag}(A_1, A_2, \dots, A_k)$ 的形式, 其中

$$A_i = \begin{pmatrix} a_i & 1 & & \\ & a_i & \ddots & \\ & & \ddots & 1 \\ & & & a_i \end{pmatrix} \quad \text{或者} \quad A_i = \begin{pmatrix} a_i & -b_i & & & \\ b_i & a_i & & & \\ & & a_i & -b_i & \\ & & b_i & a_i & \\ & & & & \ddots & \\ & & & & & \ddots & \\ & & & & & & a_i & -b_i \\ & & & & & & b_i & a_i \end{pmatrix}, \quad b_i \neq 0.$$

证明. $xI - A$ 在 $\mathbb{R}[x]$ 中可模相抵成 $\text{diag}(1, \dots, 1, p_1^{m_1}, p_2^{m_2}, \dots, p_k^{m_k})$ 的形式, 其中每个 p_i 不可约. 由于 $p_i(x)$ 的非实数根共轭成对出现, 故 $\deg(p_i) \leq 2$. 只需验证上述 A_i 的特征方阵与 $\text{diag}(1, \dots, 1, p_i^{m_i})$ 模相抵.

- 对于前一个方阵 $A_i = J_{m_i}(a_i)$, $\varphi_{A_i}(x) = p_i^{m_i}$, $p_i = x - a_i$. 由 $xI - A_i$ 的右上角 $m_i - 1$ 阶子式为 $(-1)^{m_i-1}$, 可得 $xI - A_i$ 与 $\text{diag}(1, \dots, 1, p_i^{m_i})$ 模相抵.
- 对于后一个 $2m_i$ 阶方阵 A_i , $\varphi_{A_i} = p_i^{m_i}$, $p_i(x) = (x - a_i)^2 + b_i^2$. 由 $xI - A_i$ 的右上角 $2m_i - 1$ 阶子式为 $(-1)^{m_i-1} b_i^{m_i} \neq 0$, 可得 $xI - A_i$ 与 $\text{diag}(1, \dots, 1, p_i^{m_i})$ 模相抵. \square

例 5.22. 求 n 阶循环方阵 $A = \begin{pmatrix} & & & I_{n-1} \\ & & & \\ & & & \\ 1 & & & \end{pmatrix}$ 在 \mathbb{R} 上的相似标准形.

解答. A 的所有特征值为 $\lambda_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$, $k = 1, 2, \dots, n$. 根据例 3.10 或者定理 5.14, A 与 $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 相似. 由于 $B_k = \begin{pmatrix} \cos \frac{2k\pi}{n} & -\sin \frac{2k\pi}{n} \\ \sin \frac{2k\pi}{n} & \cos \frac{2k\pi}{n} \end{pmatrix}$ 与 $\text{diag}(\lambda_k, \lambda_{n-k})$ 相似, 故

- 当 $n = 2m$ 为偶数时, A 与 $B = \text{diag}(1, -1, B_1, B_2, \dots, B_{m-1})$ 在 \mathbb{C} 上相似;
- 当 $n = 2m + 1$ 为奇数时, A 与 $B = \text{diag}(1, B_1, B_2, \dots, B_m)$ 在 \mathbb{C} 上相似.

$xI - A$ 与 $xI - B$ 在 $\mathbb{C}[x]$ 上模相抵, 从而在 $\mathbb{R}[x]$ 上也模相抵. 因此, A 与 B 在 \mathbb{R} 上相似. \square

习题

1. 举例说明: “ $\forall x \in \mathbb{F}, xI - A$ 与 $xI - B$ 相抵” 不是 “ A 与 B 相似” 的充分条件.
2. 设 $A \in \mathbb{F}^{n \times n}$, $xI - A$ 与 $\text{diag}(f_1, f_2, \dots, f_n)$ 在 $\mathbb{F}[x]$ 上模相抵. 证明:
 - (1) 存在 $P \in \mathbb{F}[x]^{n \times n}$ 使得 $(xI - A)P = \lambda I$, 其中 $\lambda = \text{lcm}(f_1, f_2, \dots, f_n)$.
 - (2) 仿照例 5.18, 证明 $\lambda(A) = O$.
 - (3) A 与 $\text{diag}(A_1, A_2, \dots, A_n)$ 相似, 其中 A_i 是首一多项式 f_i 的友方阵.
3. 设 $A, B \in \mathbb{F}^{n \times n}$, \mathbb{K} 是 \mathbb{F} 的扩域. 证明: 若 A 与 B 在 \mathbb{K} 上相似, 则 A 与 B 在 \mathbb{F} 上相似.

4. 证明: (1) 定理 5.18 中的 M_{ij} 与以下两个 $m_{ij} \times m_{ij}$ 分块矩阵都相似.

$$\begin{pmatrix} C_i & & & \\ I & C_i & & \\ & \ddots & \ddots & \\ & & I & C_i \end{pmatrix}, \begin{pmatrix} C_i & I & & \\ & C_i & \ddots & \\ & & \ddots & I \\ & & & C_i \end{pmatrix}.$$

(2) 例 5.21 中的 $\begin{pmatrix} \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} & E_{21} & & \\ & \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} & \ddots & \\ & & \ddots & E_{21} \\ & & & \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} \end{pmatrix}$ 与 $\begin{pmatrix} \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} & I & & \\ & \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} & \ddots & \\ & & \ddots & I \\ & & & \begin{matrix} a_i & -b_i \\ b_i & a_i \end{matrix} \end{pmatrix}$ 相似.

5. 求 n 阶实数方阵 $A = \begin{pmatrix} & I_{n-1} \\ -1 & \end{pmatrix}$ 在 \mathbb{R} 上的相似标准形.

6. 设实数方阵 A 满足 $d_A(x) = \varphi_A(x) = (x^2 - x^4)^6$. 分别求 A^k 的特征多项式、最小多项式和 Jordan 标准形, $k = 1, 2, \dots, 6$.

7. 设 A 是可逆实数方阵, m 是正整数. 参考例 5.10, 判断下列叙述是否正确, 并证明你的结论.

(1) 若 $\det(A) > 0$, 则存在实数方阵 B 使得 $B^2 = A$.

(2) 若 m 是奇数, 则存在实数方阵 B 使得 $B^m = A$.

8. 设 $A, B \in \mathbb{F}^{n \times n}$. 证明: 若 $\text{diag}(A, \dots, A)$ 与 $\text{diag}(B, \dots, B)$ 相似, 则 A 与 B 相似.

9. 设 $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, $xI - B$ 与 $\text{diag}(f_1, f_2, \dots, f_n)$ 在 $\mathbb{F}[x]$ 上模相抵. 证明: $I \otimes A - B \otimes I$ 与 $\text{diag}(f_1(A), f_2(A), \dots, f_n(A))$ 在 $\mathbb{F}[x]$ 上模相抵.

10. 设 $A \in \mathbb{F}^{n \times n}$, $xI - A$ 的不变因子和初等因子组如定义 5.8 所述. 证明:

(1) $\text{rank}(I \otimes A - A^T \otimes I) = n^2 - \dim\{X \in \mathbb{F}^{n \times n} \mid AX = XA\}$.

(2) $\text{rank}(I \otimes A - A^T \otimes I) = \sum_{i=1}^n \text{rank}(d_i(A)) = \sum_{i=1}^n (2i - n - 1) \deg(d_i)$.

(3) $\text{rank}(I \otimes A - A^T \otimes I) = \sum_{i=1}^k \sum_{j=1}^{s_i} \text{rank}(p_i^{m_{ij}}(A)) = n^2 - \sum_{i=1}^k \left(\sum_{j,l=1}^{s_i} \min(m_{ij}, m_{il}) \right) \deg(p_i)$.

11. 设 A, B 是可逆方阵, 并且对充分大的正整数 k , A^k 与 B^k 相似. 证明: A 与 B 相似.

12. (1) 设二元域 \mathbb{F}_2 上的 n^2 阶方阵 $A = \begin{pmatrix} B & I & \cdots & I \\ I & B & \ddots & \vdots \\ \vdots & \ddots & \ddots & I \\ I & \cdots & I & B \end{pmatrix}$, 其中 $B = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \cdots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$ 是 n 阶

全一方阵. 求 A 的行列式、秩、特征多项式和相似标准形.

(2) 设二元域 \mathbb{F}_2 上的 n^2 阶方阵 $A = \begin{pmatrix} B & I & & \\ I & B & \ddots & \\ & \ddots & \ddots & I \\ & & I & B \end{pmatrix}$, 其中 $B = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{pmatrix}$ 是 n 阶方阵. 研究 A 的行列式、秩、特征多项式和相似标准形.

第六章 正交方阵

正交方阵是一类具有良好性质的实数方阵，在数学、物理、电子、信息等科学技术和生产、生活各方面都有广泛应用。本章介绍与正交方阵相关的一些知识，为接下来学习内积空间理论打下基础，并提供具体实例和研究工具。

§6.1 正交方阵

定义 6.1. 设 $\alpha, \beta \in \mathbb{R}^{n \times 1}$. 实数 $\alpha^T \beta$ 称为 α 与 β 的数量积. $\sqrt{\alpha^T \alpha}$ 称为 α 的长度, 记作 $\|\alpha\|$. 若 $\alpha^T \beta = 0$, 则称 α 与 β 正交, 记作 $\alpha \perp \beta$.

满足 $PP^T = P^T P = I$ 的实数方阵 P 称为正交方阵. 正交方阵可以看作是 \mathbb{R}^n 中一组两两正交、长度为 1 的行/列向量.

例 6.1. 任意置换方阵都是正交方阵.

例 6.2. 任意 2 阶正交方阵形如 $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ 或 $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$.

定理 6.1. 设 P 是 n 阶正交方阵. 有如下结论.

1. $\det(P) = \pm 1$.
2. $P^{-1} = P^T$ 也是正交方阵.
3. 若 Q 是 n 阶正交方阵, 则 PQ 也是正交方阵.
4. 对于任意 $\alpha, \beta \in \mathbb{R}^{n \times 1}$, 有 $(P\alpha)^T(P\beta) = \alpha^T \beta$.
5. 若 P 是三角方阵, 则 $P = \text{diag}(\pm 1, \pm 1, \dots, \pm 1)$.
6. 若 A 是 n 阶对称 (反对称) 方阵, 则 $P^{-1}AP$ 也是对称 (反对称) 方阵.

在矩阵乘法运算下, n 阶正交方阵的全体构成 $GL(n, \mathbb{R})$ 的子群, 称为正交群, 记作 $O(n, \mathbb{R})$. 行列式等于 1 的 n 阶正交方阵的全体构成 $SL(n, \mathbb{R})$ 的子群, 称为特殊正交群, 记作 $SO(n, \mathbb{R})$.

定理 6.2. 设 P 是 n 阶正交方阵, 则 P 的任意特征值 $\lambda \in \mathbb{C}$ 满足 $|\lambda| = 1$. 特别, 当 $\lambda \in \mathbb{R}$ 时, $\lambda = \pm 1$. 当 $\lambda \notin \mathbb{R}$ 时, 设 $\alpha = u + iv$ 是 $\lambda = \cos \theta + i \sin \theta$ 对应的特征向量, 其中 $u, v \in \mathbb{R}^{n \times 1}$, 则有

$$u \perp v, \quad \|u\| = \|v\|, \quad P \begin{pmatrix} u & v \end{pmatrix} = \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

2. 对于任意非零向量 $\alpha \in \mathbb{R}^{n \times 1}$, 当 $\beta = (\|\alpha\|, \underbrace{0, \dots, 0}_{n-1})^T \neq \alpha$ 时, $v = \alpha - \beta$ 满足 $H_v \alpha = \beta$.

3. 由前述结论可得 QR 分解的存在性. 若 $A = Q_1 R_1 = Q_2 R_2$, 则 $M = Q_1^{-1} Q_2 = R_1 R_2^{-1}$ 既是正交方阵, 又是上三角方阵, 并且对角元素都是正数. 因此, $M = I$, $Q_1 = Q_2$, $R_1 = R_2$. \square

对于任意可逆方阵 A , 一般有两种方式产生 A 的 QR 分解. 一种是通过左乘正交方阵, 把 A 变成上三角方阵; 另一种是通过右乘上三角方阵, 把 A 变成正交方阵. 定理 6.3 采用的是第一种方式. 而 Gram^[3]-Schmidt^[4]标准正交化过程则采用了第二种方式.

定义 6.2. 设 $A = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)$ 是 n 阶可逆实数方阵. 定义

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\gamma_i^T \alpha_k) \gamma_i = \alpha_k - \sum_{i=1}^{k-1} \frac{\beta_i^T \alpha_k}{\beta_i^T \beta_i} \beta_i, \quad \gamma_k = \frac{1}{\|\beta_k\|} \beta_k, \quad k = 1, 2, \dots, n. \quad (6.1)$$

A 的可逆性保证了每个 $\beta_k \neq \mathbf{0}$. 容易验证, $\beta_1, \beta_2, \dots, \beta_n$ 两两正交, $Q = (\gamma_1 \ \gamma_2 \ \dots \ \gamma_n)$ 是正交方阵. 由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 生成 $\beta_1, \beta_2, \dots, \beta_n$ 的过程称为 **Gram-Schmidt 正交化过程**. 由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 生成 $\gamma_1, \gamma_2, \dots, \gamma_n$ 的过程称为 **Gram-Schmidt 标准正交化过程**.

由 (6.1) 式, 可得 QR 分解

$$(\alpha_1 \ \alpha_2 \ \dots \ \alpha_n) = (\gamma_1 \ \gamma_2 \ \dots \ \gamma_n) \begin{pmatrix} \|\beta_1\| & \gamma_1^T \alpha_2 & \dots & \gamma_1^T \alpha_n \\ & \|\beta_2\| & \ddots & \vdots \\ & & \ddots & \gamma_{n-1}^T \alpha_n \\ & & & \|\beta_n\| \end{pmatrix}.$$

Gram-Schmidt 标准正交化过程的几何意义如下. 设 V_k 是 $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}^{n \times 1}$ 生成的 k 维线性空间, 则 V_k 可以由一组两两正交、长度为 1 的向量 $\gamma_1, \gamma_2, \dots, \gamma_k$ 生成. $\gamma_1, \gamma_2, \dots, \gamma_k$ 称为**标准正交向量组**. α_k 可以分解为两个向量 $\alpha_k - \beta_k \in V_{k-1}$ 与 $\beta_k \perp V_{k-1}$ 之和. γ_k 是 β_k 的单位化.

例 6.4. 求方阵 $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix}$ 的 QR 分解.

解法一: 左乘 Givens 方阵.

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2\sqrt{2} & 2\sqrt{2} & \frac{3}{\sqrt{2}} \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 1 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} \begin{pmatrix} 2\sqrt{2} & 2\sqrt{2} & \frac{3}{\sqrt{2}} \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{3}{\sqrt{2}} \end{pmatrix} = R.$$

^[3]Jørgen Pedersen Gram, 1850–1916, 丹麦数学家.

^[4]Erhard Schmidt, 1876–1959, 德国数学家.

由此可得, $A = QR$, 其中

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix}. \quad \square$$

解法二: 左乘 Householder 方阵.

$$\begin{aligned} u &= \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, & H_u &= \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}, & H_u \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix} &= \begin{pmatrix} 3 & 3 & 3 \\ 0 & -1 & -2 \\ 0 & -1 & 1 \end{pmatrix} \\ v &= \begin{pmatrix} -1 - \sqrt{2} \\ -1 \end{pmatrix}, & H_v &= \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, & H_v \begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix} &= \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{2}} \end{pmatrix} \end{aligned}$$

由此可得, $A = QR$, 其中

$$Q = H_u \begin{pmatrix} 1 & & \\ & H_v & \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix}, \quad R = \begin{pmatrix} 3 & 3 & 3 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{3}{\sqrt{2}} \end{pmatrix}. \quad \square$$

解法三: Gram-Schmidt 标准正交化.

$$\begin{aligned} \alpha_1 &= (2, 2, 1), & \beta_1 &= \alpha_1, & \gamma_1 &= \frac{1}{3}\beta_1 = \left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}\right), \\ \alpha_2 &= (1, 3, 1), & \beta_2 &= \alpha_2 - 3\gamma_1 = (-1, 1, 0), & \gamma_2 &= \frac{1}{\sqrt{2}}\beta_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \\ \alpha_3 &= (1, 2, 3), & \beta_3 &= \alpha_3 - 3\gamma_1 - \frac{1}{\sqrt{2}}\gamma_2 = \left(-\frac{1}{2}, -\frac{1}{2}, 2\right), & \gamma_3 &= \frac{\sqrt{2}}{3}\beta_3 = \left(-\frac{1}{3\sqrt{2}}, -\frac{1}{3\sqrt{2}}, \frac{2\sqrt{2}}{3}\right). \end{aligned}$$

由此可得, $A = QR$, 其中

$$\begin{aligned} Q &= (\gamma_1 \ \gamma_2 \ \gamma_3) = \begin{pmatrix} \frac{2}{3} & -\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix}, \\ R &= \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{3}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{3}{\sqrt{2}} \end{pmatrix}. \quad \square \end{aligned}$$

 上例中的三种方法均可求得可逆实方阵的 QR 分解. 一般说来, “左乘正交方阵” 方法适合于计算机编程求解, 具有良好的数值稳定性; “Gram-Schmidt 标准正交化” 方法适合于小规模的手算. 针对具体问题时, 我们需要选择恰当的方法.

定理 6.4 (Hadamard^[5]不等式). 设 $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}^{n \times 1}$ 都是非零向量, 则

$$|\det(\alpha_1, \alpha_2, \dots, \alpha_n)| \leq \prod_{k=1}^n \|\alpha_k\|.$$

不等式取等号当且仅当 $\alpha_1, \alpha_2, \dots, \alpha_n$ 两两正交.

^[5]Jacques Salomon Hadamard, 1865–1963, 法国数学家.

证明. 设 $A = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)$. 当 $\det(A) = 0$ 时, 不等式显然成立, 并且取不到等号. 当 $\det(A) \neq 0$ 时, 设 $A = QR$, 其中 Q 是正交方阵, $R = (r_{ij}) = (\beta_1 \ \beta_2 \ \cdots \ \beta_n)$ 是上三角方阵, 则

$$|\det(A)| = \prod_{k=1}^n |r_{kk}| \leq \prod_{k=1}^n \|\beta_k\| = \prod_{k=1}^n \|Q\beta_k\| = \prod_{k=1}^n \|\alpha_k\|.$$

不等式取等号当且仅当 R 是对角方阵, 即 $\beta_1, \beta_2, \dots, \beta_n$ 两两正交. 根据定理 6.1 结论 3, $\beta_1, \beta_2, \dots, \beta_n$ 两两正交当且仅当 $\alpha_1, \alpha_2, \dots, \alpha_n$ 两两正交. \square

习题

- 证明定理 6.1.
- 设 $P \in \mathbb{R}^{n \times n}$ 满足 $\|P\alpha\| = \|\alpha\|, \forall \alpha \in \mathbb{R}^{n \times 1}$. 证明: P 是正交方阵.
- 设 $\alpha = (2, 1, 1)^T, \beta = (1, 1, 2)^T$. 求所有正交方阵 P 使得 $P\alpha = \beta$.
- 设 $\alpha, \beta \in \mathbb{R}^{3 \times 1}, \alpha \times \beta$ 是向量积运算, P 是 3 阶可逆实数方阵.
 - 证明: 若 $P \in SO(3, \mathbb{R})$, 则 $(P\alpha) \times (P\beta) = P(\alpha \times \beta), \forall \alpha, \beta$.
 - 设 $(P\alpha) \times (P\beta) = P(\alpha \times \beta), \forall \alpha, \beta$. P 是否必为正交方阵? 证明你的结论.
- (1) 证明: 任意 $P \in SO(3, \mathbb{R})$ 既可以表示为三个 Givens 方阵的乘积, 也可以表示为两个 Householder 方阵的乘积.
 - 对 $P \in SO(n, \mathbb{R})$, 推广上述结论并证明.
- 求 n 阶 Givens 方阵 $G_{ij}(\theta)$ 和 Householder 方阵 H_v 在 \mathbb{C} 上的 Jordan 标准形.
- (1) 设 H 是 Householder 方阵. 证明: $\begin{pmatrix} H & \\ & I \end{pmatrix}$ 和 $\begin{pmatrix} I & \\ & H \end{pmatrix}$ 也都是 Householder 方阵.
 - 设 n 阶对称实方阵 H 满足 $\det(H) = -1$ 且 $\text{rank}(H - I) = 1$. 证明: H 是 Householder 方阵.
- 对下列方阵 A 的列向量施行 Gram-Schmidt 标准正交化, 并给出 A 的 QR 分解.

(1) $\begin{pmatrix} 1 & 2 & 6 & -11 \\ 2 & -1 & 2 & 3 \\ 4 & -2 & 9 & 6 \\ 2 & 4 & 2 & 3 \end{pmatrix}$	(2) $\begin{pmatrix} -1 & 4 & 2 & 2 \\ -2 & 3 & 2 & -4 \\ 0 & 2 & -1 & 5 \\ -2 & 4 & 6 & 3 \end{pmatrix}$	(3) $\begin{pmatrix} 1 & 6 & -3 & -1 \\ 4 & 9 & -7 & -1 \\ 2 & 2 & -1 & -1 \\ 2 & 2 & 4 & 1 \end{pmatrix}$
----------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------
- (1) 设 A 是反对称实数方阵. 证明: $I + A$ 可逆, 并且 $B = (I + A)^{-1}(I - A)$ 是正交方阵.
 - 设 B 是正交方阵, 并且 $I + B$ 可逆. 证明: 存在反对称实数方阵 A 使得

$$B = (I + A)^{-1}(I - A).$$

- 证明: 正交方阵的任意子矩阵的特征值 λ 满足 $|\lambda| \leq 1$.

11. 设 P 是 n 阶可逆实数方阵, 并且对于任意 n 阶对称实数方阵 A , $P^{-1}AP$ 都是对称方阵. 证明: 存在实数 λ 使得 λP 是正交方阵.
12. 设实数方阵 A 的行向量两两正交, 列行向量也两两正交. 证明: 存在置换方阵 P_1, P_2 和正交方阵 Q_1, \dots, Q_k 和实数 $\lambda_1 > \dots > \lambda_k \geq 0$, 使得

$$A = P_1 \begin{pmatrix} \lambda_1 Q_1 & & \\ & \ddots & \\ & & \lambda_k Q_k \end{pmatrix} P_2.$$

§6.2 正交相似

定义 6.3. 设 $A, B \in \mathbb{R}^{n \times n}$. 若存在正交方阵 P 使得 $A = PBP^{-1}$, 则称 A 与 B 正交相似.

容易验证, 实数矩阵的正交相似关系构成一个等价关系.

对定理 5.5 及其证明略作修改, 可得

定理 6.5. 设 $A \in \mathbb{R}^{n \times n}$ 的所有特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$. 当 $1 \leq k \leq s$ 时, $\lambda_{2k} = \overline{\lambda_{2k-1}} \notin \mathbb{R}$. 当 $2s+1 \leq k \leq n$ 时, $\lambda_k \in \mathbb{R}$. A 可以正交相似成如下形式的准上三角方阵

$$\begin{pmatrix} A_1 & * & * & * & * & * \\ & \ddots & * & * & * & * \\ & & A_s & * & * & * \\ & & & \lambda_{2s+1} & * & * \\ & & & & \ddots & * \\ & & & & & \lambda_n \end{pmatrix}, \quad (6.2)$$

其中每个 A_k 是 2 阶方阵, A_k 的特征值为 λ_{2k-1} 和 λ_{2k} , $1 \leq k \leq s$.

证明. 对 n 使用数学归纳法. 当 $n=1$ 时, 结论显然成立. 下设 $n \geq 2$.

- 情形一: $s=0$. 设 α_1 是 λ_1 对应的一个特征向量, $\|\alpha_1\|=1$, 正交方阵 P_1 以 α_1 为第 1 列, 则有 $AP_1 = P_1 \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{pmatrix}$. B 的所有特征值为 $\lambda_2, \lambda_3, \dots, \lambda_n$. 根据归纳假设, 存在正交方阵 P_2 使得 $P_2^{-1}BP_2$ 是上三角方阵, 并且对角元素依次是 $\lambda_2, \lambda_3, \dots, \lambda_n$. 设 $P = P_1 \begin{pmatrix} 1 & \\ & P_2 \end{pmatrix}$, 则 P 是正交方阵, $P^{-1}AP$ 形如 (6.2).
- 情形二: $s \neq 0$. 设 $u+iv$ 是 $\lambda_1 = a_1 + b_1i$ 对应的一个特征向量, 其中 $u, v \in \mathbb{R}^{n \times 1}$, $a_1, b_1 \in \mathbb{R}$, 则有 $A \begin{pmatrix} u & v \end{pmatrix} = \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$. 由 $\lambda_1 \notin \mathbb{R}$ 知 u 与 v 不平行. 把 u, v 标准正交化为标准正交向量组 α_1, α_2 . 设 $A \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} A_1$, 则 A_1 与 $\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$ 相似, A_1 的特征值为 λ_1, λ_2 . 把 $\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix}$ 扩充为正交方阵 $P_1 = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$, 则 $AP_1 = P_1 \begin{pmatrix} A_1 & * \\ \mathbf{0} & B \end{pmatrix}$. B 的所有特征值为 $\lambda_3, \lambda_4, \dots, \lambda_n$. 根据归纳假设, 存在正交方阵 P_2 使得 $P_2^{-1}BP_2$ 是准上三角方阵. 设 $P = P_1 \begin{pmatrix} I_2 & \\ & P_2 \end{pmatrix}$, 则 P 是正交方阵, $P^{-1}AP$ 形如 (6.2). □

根据定理 6.5 可得下列常用推论.

定理 6.6. 任意正交方阵可以正交相似成如下形式的准对角方阵.

$$\text{diag} \left(\begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_s & \sin \theta_s \\ -\sin \theta_s & \cos \theta_s \end{pmatrix}, I, -I \right)$$

证明. 设 n 阶正交方阵 A 可以正交相似成 (6.2) 式中的方阵 B , 则 B 也是正交方阵. 由 $B^T = B^{-1}$ 既是准上三角方阵, 又是准下三角方阵, 知 $B = \text{diag}(A_1, \dots, A_s, \lambda_{2s+1}, \dots, \lambda_n)$. 每个 A_k 是 2 阶正交方阵, 特征值都不是实数, 故 $A_k = \begin{pmatrix} \cos \theta_k & \sin \theta_k \\ -\sin \theta_k & \cos \theta_k \end{pmatrix}$. 根据定理 6.2, $\lambda_{2s+1}, \dots, \lambda_n \in \{-1, 1\}$. 因此, 可设 $B = \text{diag}(A_1, \dots, A_s, I, -I)$. \square

定理 6.7. 任意对称实数方阵可以正交相似成对角方阵.

证明. 设 n 阶对称实数方阵 A 可以正交相似成 (6.2) 式中的方阵 B , 则 B 也是对称方阵. 由例 5.2 可知 $B = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. \square

定理 6.8. 任意反对称实数方阵可以正交相似成如下形式的准对角方阵.

$$\text{diag} \left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, O \right)$$

证明. 设 n 阶反对称实数方阵 A 可以正交相似成 (6.2) 式中的方阵 B , 则 B 也是反对称方阵, 从而 $B = \text{diag}(A_1, \dots, A_s, O)$. 每个 A_k 是 2 阶反对称方阵, 故 $A_k = \begin{pmatrix} 0 & b_k \\ -b_k & 0 \end{pmatrix}$. \square

正交方阵、对称方阵、反对称方阵都满足 $AA^T = A^T A$. 由此引入规范方阵的定义.

定义 6.4. 若实数方阵 A 满足 $AA^T = A^T A$, 则 A 称为规范方阵.

例 6.5. 设 $A, P \in \mathbb{R}^{n \times n}$, A 是规范方阵, P 是正交方阵. A^T 和 $P^{-1}AP$ 都是规范方阵. 当 A 可逆时, A^{-1} 也是规范方阵.

例 6.6. 任意 2 阶规范实数方阵 A 形如 $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ 或 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. 因此, A 是对称方阵或正交方阵的数乘.

解答. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. $AA^T = A^T A \Leftrightarrow \begin{cases} (b+c)(b-c) = 0 \\ (a-d)(b-c) = 0 \end{cases} \Leftrightarrow$ 下列两种情形之一成立.

(i) $b = c$. A 是对称方阵. (ii) $a = d, b = -c \neq 0$. $A = \sqrt{a^2 + b^2} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$. \square

定理 6.9. 任意规范实数方阵可正交相似成如下形式的准对角方阵.

$$\text{diag} \left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \dots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \dots, \lambda_n \right)$$

证明. 设 A 可以正交相似成 (6.2) 式中的方阵 B . 由例 6.5, 得 B 也是规范方阵. 设 $BB^T = (C_{ij})$, $B^T B = (D_{ij})$. 由 $\text{tr}(C_{ii}) = \text{tr}(D_{ii})$, 得 $B = \text{diag}(A_1, \dots, A_s, \lambda_{2s+1}, \dots, \lambda_n)$. 每个 A_k 是 2 阶规范方阵且特征值不是实数. 由例 6.6, $A_k = \begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}$, $b_k \neq 0$. \square

习题

1. (1) 证明: 任意 3 阶正交方阵 A 可以正交相似成 $\begin{pmatrix} \det(A) & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}$ 的形式.
- (2) 说明上述结论的几何含义.
2. 设 A 是实数方阵, $k \in \mathbb{N}^*$. 证明: A 可以相似成正交方阵 $\Leftrightarrow A^k$ 可以相似成正交方阵.
3. 设 A 是实数方阵. 证明: A 可以在 \mathbb{R} 上相似成规范方阵 $\Leftrightarrow A$ 可以在 \mathbb{C} 上相似成对角方阵.
4. 证明: 上三角的规范实数方阵一定是对角方阵.
5. 设 n 阶实数方阵 $A = (a_{ij})$ 满足 (i) $a_{ij} = 0, \forall |i - j| \geq 2$; (ii) $a_{i,i+1}a_{i+1,i} > 0, \forall i < n$. 证明:
 - (1) $d_A = \varphi_A$.
 - (2) A 可以在 \mathbb{R} 上相似成对称方阵, 进而可以相似成对角方阵.
 - (3) A 的 n 个特征值两两不同.
6. 设实数方阵 A 与 B 乘积可交换. 证明:
 - (1) 若 A 是规范方阵, 则 A 与 B^T 乘积可交换.
 - (2) 若 A 和 B 都是规范方阵, 则 $A + B$ 和 AB 也都是规范方阵.
7. 分别举例满足下列条件的 n 阶实数方阵 A, B .
 - (1) A 与 B 乘积可交换, A 与 B^T 乘积不交换.
 - (2) A 和 B 都是规范方阵, $A + B$ 和 AB 都不是规范方阵.
8. 设 A 是规范实数方阵. 证明: A 与 A^T 正交相似, 并且存在 $f \in \mathbb{R}[x]$ 使得 $A^T = f(A)$.
9. 设 I 是指标集合, $\{A_i \mid i \in I\}$ 是一组两两乘积可交换的规范方阵. 参考第五章 §5.2 习题 10. 证明: 存在正交方阵 P 使得每个 $P^{-1}A_iP$ 是定理 6.9 中的准对角方阵.
10. 设 A, B 都是 n 阶实数方阵. 证明:
 - (1) 若 A 是反对称方阵, 则 $B = e^A$ 是正交方阵, $\det(B) = 1$.
 - (2) 若 B 是正交方阵, $\det(B) = 1$, 则存在反对称方阵 A 使得 $B = e^A$.
 - (3) 若 $B = e^A$ 是正交方阵, A 是否一定是反对称方阵? 证明你的结论.

§6.3 正交相抵

定义 6.5. 设 $A, B \in \mathbb{R}^{m \times n}$. 若存在正交方阵 P, Q 使得 $A = PBQ$, 则称 A 与 B 正交相抵.

容易验证, 实数矩阵的正交相抵关系构成一个等价关系.

定理 6.10. 对于任意 $A \in \mathbb{R}^{m \times n}$, 存在正交方阵 P, Q 使得

$$A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q \quad (6.3)$$

其中 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, $r = \text{rank}(A)$.

证明. $A^T A$ 是对称方阵, 根据定理 6.7, 存在正交方阵 Q 使得

$$Q(A^T A)Q^{-1} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

设 $AQ^{-1} = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)$, 则 $\lambda_k = \alpha_k^T \alpha_k \geq 0, \forall k$. 当 $k \leq r$ 时, $\|\alpha_k\| = \sigma_k = \sqrt{\lambda_k} > 0$; 当 $k > r$ 时, $\alpha_k = \mathbf{0}$. 由 $\alpha_1, \dots, \alpha_r$ 两两正交, 可把 $(\frac{1}{\sigma_1}\alpha_1 \ \frac{1}{\sigma_2}\alpha_2 \ \dots \ \frac{1}{\sigma_r}\alpha_r)$ 扩充为 m 阶正交方阵 P , 即得 (6.3). \square

定义 6.6. (6.3) 式称为 A 的一个**奇异值分解**. $\begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix}$ 可以作为 A 所在的正交相抵等价类的代表元素, 称为 A 的**正交相抵标准形**. (6.3) 式中的正交方阵 P, Q 一般是不唯一的, 但 $\sigma_1, \sigma_2, \dots, \sigma_r$ 由 A 唯一确定. σ_k 称为 A 的第 k 个**奇异值**, 记作 $\sigma_k(A)$.

例 6.7. 设 $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, $r = \text{rank}(A)$, $x \in \mathbb{R}^{n \times 1}$. 有下列结论.

1. $\sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\sum_{k=1}^r \sigma_k^2(A)}$. 此值称为 A 的**Frobenius 范数**, 记作 $\|A\|_F$.
2. $\max_{\|x\|=1} \|Ax\| = \sigma_1(A)$. 此值称为 A 的**矩阵范数**, 记作 $\|A\|$.
3. 当 A 是列满秩时, $\min_{\|x\|=1} \|Ax\| = \sigma_n(A)$.

证明. 设 $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 是奇异值分解, $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 则有

$$\begin{aligned} \sum_{i,j} a_{ij}^2 &= \text{tr}(A^T A) = \text{tr} \left(Q^T \begin{pmatrix} \Sigma^2 & O \\ O & O \end{pmatrix} Q \right) = \text{tr}(\Sigma^2) = \sum_{k=1}^r \sigma_k^2, \\ \max_{\|x\|=1} \|Ax\| &= \max_{\|x\|=1} \sqrt{x^T A^T A x} = \max_{\|y\|=1} \sqrt{y^T \begin{pmatrix} \Sigma^2 & O \\ O & O \end{pmatrix} y} = \sigma_1, \\ \min_{\|x\|=1} \|Ax\| &= \min_{\|x\|=1} \sqrt{x^T A^T A x} = \min_{\|y\|=1} \sqrt{y^T \begin{pmatrix} \Sigma^2 & O \\ O & O \end{pmatrix} y} = \sigma_n. \end{aligned}$$

\square

例 6.8. 设线性映射 $f(x, y) = (ax + by, cx + dy)$, 其中 $ad - bc \neq 0$. 求圆 $x^2 + y^2 = 1$ 在 f 下的像 E 的标准方程, 以及 E 的对称轴所在直线的方程.

解答. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 有奇异值分解 $P \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} Q$, $P = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. 设点 (u, v) 在 E 上.

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} u \\ v \end{pmatrix} \Rightarrow (u \ v) A^{-T} A^{-1} \begin{pmatrix} u \\ v \end{pmatrix} = 1 \Rightarrow (u \ v) P \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix} P^T \begin{pmatrix} u \\ v \end{pmatrix} = 1.$$

E 的标准方程为 $\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = 1$, 故 E 是椭圆.

E 的长轴顶点 (u, v) 满足 $(u, v)P = (\pm\sigma_1, 0) \Rightarrow (u, v) = \pm\sigma_1(\cos \theta, \sin \theta)$.

E 的短轴顶点 (u, v) 满足 $(u, v)P = (0, \pm\sigma_2) \Rightarrow (u, v) = \pm\sigma_2(-\sin \theta, \cos \theta)$.

故 E 的长轴方程 $x \sin \theta - y \cos \theta = 0$, 短轴方程 $x \cos \theta + y \sin \theta = 0$. □

例 6.9. $m \times n$ 实系数线性方程组 $Ax = b$ 可能无解, 也可能有许多解. 在实际问题中, 经常把线性方程组问题转化为优化问题: 求 x 使得 $\|Ax - b\|$ 最小. 优化问题的解总是存在的, 称为原线性方程组问题的最小二乘解. 当最小二乘解 x 不唯一时, 长度最小的最小二乘解 x 是唯一的.

解答. 设 $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 是奇异值分解. 由

$$\|Ax - b\| = \|P^{-1}(Ax - b)\| = \left\| \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Qx - P^{-1}b \right\|,$$

可得 $Ax = b$ 的所有最小二乘解

$$x = Q^{-1} \begin{pmatrix} \Sigma^{-1}c \\ t \end{pmatrix}$$

其中 $c \in \mathbb{R}^{r \times 1}$ 由 $P^{-1}b$ 的前 r 个元素构成, $t \in \mathbb{R}^{(n-r) \times 1}$ 是自由参数. 在这些最小二乘解中, 唯有

$$x = Q^{-1} \begin{pmatrix} \Sigma^{-1}c \\ \mathbf{0} \end{pmatrix} = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1}b$$

使得 $\|x\|$ 最小. □

例 6.10. 设实数矩阵 A 有奇异值分解 $P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$, $X = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1}$, 则有

(1) X 是线性方程组 $AX = I$ 的最小二乘解, 并且使得 $\|X\|_F$ 最小.

(2) X 是矩阵方程组 $AXA = A$, $XAX = X$, $(AX)^T = AX$, $(XA)^T = XA$ 的唯一解.

证明. 结论 (1) 的证明同例 6.9. 下证结论 (2). 设 $X = Q^{-1} \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} P^{-1}$ 满足矩阵方程组. $AXA = A \Rightarrow X_1 = \Sigma^{-1}$. $(AX)^T = AX \Rightarrow X_2 = O$. $(XA)^T = XA \Rightarrow X_3 = O$. $XAX = X \Rightarrow X_4 = X_3 \Sigma^{-1} X_2 = O$. 因此, $X = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1}$. 容易验证, X 满足矩阵方程组. □

习题

1. 设实数矩阵 A 与 B 正交相抵. 证明: $\|A\|_F = \|B\|_F$.
2. 设实数矩阵 $A, B \in \mathbb{R}^{m \times n}$ 满足 $A^T A = B^T B$. 证明: 存在正交方阵 P 使得 $PA = B$.
3. 给定实数矩阵 A , 求所有实数矩阵 B 使得 $BB^T = AA^T$ 且 $B^T B = A^T A$.
4. 证明:
 - (1) 对于任意实数方阵 A , 存在规范的实数方阵 P_1, P_2 使得 $A = P_1 + P_2$.
 - (2) 对于任意实数方阵 A , 存在规范的实数方阵 P_1, P_2 使得 $A = P_1 P_2$.
 - (3) 对于任意正交方阵 A , 存在对称的正交方阵 P_1, P_2 使得 $A = P_1 P_2$.
5. 证明下列关于实数矩阵的范数 $\|\cdot\|$ 的结论.
 - (1) $\|A + B\| \leq \|A\| + \|B\|$, $\|AB\| \leq \|A\| \cdot \|B\|$.
 - (2) 设 B 是 A 的子矩阵, 则 $\|B\| \leq \|A\|$.
 - (3) $\|A\| \leq \|A\|_F \leq \sqrt{r} \|A\|$, 其中 $r = \text{rank}(A)$.
 - (4) 设 A 是方阵, 则 $\rho(A) \leq \|A\|$, 其中 $\rho(A)$ 是 A 的谱半径 (见第五章 §5.3 习题 10).
6. 设 $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^{n \times 1}$, $p, q \in [1, +\infty]$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$. 定义向量和矩阵的 p 范数

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad \|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p.$$

证明:

- (1) $\|A + B\|_p \leq \|A\|_p + \|B\|_p$, $\forall B \in \mathbb{R}^{m \times n}$. (2) $\|AB\|_p \leq \|A\|_p \cdot \|B\|_p$, $\forall B \in \mathbb{R}^{n \times k}$.
 - (3) $\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^m |a_{ij}| \right)$ (4) $\frac{\|A\|_1}{\sqrt{m}} \leq \|A\|_2 \leq \sqrt{n} \|A\|_1$
 - (5) $\|A\|_p = \|A^T\|_q$ (6) $\|A\|_2 \leq \sqrt{\|A\|_p \cdot \|A\|_q}$
7. 设线性映射 $f(x, y) = (ax + by, cx + dy)$, 其中 $ad - bc \neq 0$.
 - (1) 设 E 是圆 $x^2 + y^2 = 1$ 在 f 下的像. 求 E 的焦点的坐标.
 - (2) 设 H 是双曲线 $x^2 - y^2 = 1$ 在 f 下的像. 求 H 的标准方程、对称轴所在直线的方程、渐近线方程、顶点和焦点的坐标.
 - (3) 设 P 是抛物线 $y = x^2$ 在 f 下的像. 求 P 的标准方程、对称轴所在直线的方程、顶点和焦点的坐标.
 8. 给定 $A, B \in \mathbb{R}^{n \times n}$.
 - (1) 求实数 λ 使得 $\|\lambda A - B\|_F$ 最小.
 - (2) 求正交方阵 P 使得 $\|PA - B\|_F$ 最小.
 - (3) 求实数 λ 和正交方阵 P 使得 $\|\lambda PA - B\|_F$ 最小.

9. 设 A 是实数矩阵, $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 是奇异值分解, $X = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1}$. 证明:

(1) X 是唯一的, 与正交方阵 P, Q 的选取无关.

(2) 若 $A = BC$, B, C 分别是列满秩和行满秩矩阵, 则 $X = C^T(CC^T)^{-1}(B^TB)^{-1}B^T$.

(3) A 的任意广义逆矩阵 (见第四章 §4.4.2) 形如 $Y = Q^{-1} \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2\Sigma Y_1 \end{pmatrix} P^{-1}$. 从而,

$$\|Y\|_F \geq \|X\|_F, \quad \|Y\| \geq \|X\|.$$

10. 设 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ 是 $A \in \mathbb{R}^{n \times n}$ 的所有特征值. 证明: A 与 $\text{diag}(|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|)$ 正交相抵当且仅当 A 是规范方阵.

§6.4 酉方阵

本章前面几节中关于实数向量和实数方阵的结论, 可以略作修改, 推广到复数向量和复数矩阵. 我们仅叙述结论, 证明留作习题.

定义 6.7. 设 $\alpha, \beta \in \mathbb{C}^{n \times 1}$. 复数 $\alpha^H \beta$ 称为 α 与 β 的酉内积. 非负实数 $\sqrt{\alpha^H \alpha}$ 称为 α 的长度, 记作 $\|\alpha\|$. 若 $\alpha^H \beta = 0$, 则称 α 与 β 正交, 记作 $\alpha \perp \beta$.

满足 $PP^H = P^H P = I$ 的复数方阵 P 称为酉方阵. 酉方阵也可以看作是 \mathbb{C}^n 中一组两两正交、长度为 1 的 (行或列) 向量.

定理 6.11. 设 P 是 n 阶酉方阵. 有下列常用结论.

1. $|\det(P)| = 1$.
2. $P^T, \bar{P}, P^{-1} = P^H$ 都是酉方阵.
3. 若 Q 是 n 阶酉方阵, 则 PQ 也是酉方阵.
4. 对于任意 $\alpha, \beta \in \mathbb{C}^{n \times 1}$, 有 $(P\alpha)^H (P\beta) = \alpha^H \beta$.
5. 若 P 是三角方阵, 则 $P = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 其中每个 $|\lambda_i| = 1$.
6. 若 A 是 n 阶 Hermite (反 Hermite) 方阵, 则 $P^{-1}AP$ 也是 Hermite (反 Hermite) 方阵.

在矩阵乘法运算下, n 阶酉方阵的全体构成 $GL(n, \mathbb{C})$ 的子群, 称为酉群, 记作 $U(n, \mathbb{C})$. 行列式等于 1 的 n 阶酉方阵的全体构成 $SL(n, \mathbb{C})$ 的子群, 称为特殊酉群, 记作 $SU(n, \mathbb{C})$. $O(n, \mathbb{R})$ 是 $U(n, \mathbb{C})$ 的子群, $SO(n, \mathbb{R})$ 是 $SU(n, \mathbb{C})$ 的子群.

§6.4.1 酉方阵的构造

例 6.11. 任意 2 阶酉方阵形如 $\begin{pmatrix} z & w \\ -\mu\bar{w} & \mu\bar{z} \end{pmatrix}$, 其中 $z, w, \mu \in \mathbb{C}$ 满足 $|z|^2 + |w|^2 = 1$, $|\mu| = 1$.

例 6.12. 对于任意非零向量 $v \in \mathbb{C}^{n \times 1}$, Hermite 方阵 $H_v = I - \frac{2}{v^H v} v v^H$ 是酉方阵, $\det(H_v) = -1$.

定理 6.12 (QR 分解). 对于任意复数方阵 A , 存在酉方阵 P 使得 $R = PA$ 是上三角方阵, 并且 R 的对角元素都是非负实数. 特别, 当 A 是可逆方阵时, 存在唯一的酉方阵 Q 和上三角方阵 R , 使得 $A = QR$ 并且 R 的对角元素都是正数. 矩阵乘积分解 $A = QR$ 称为 A 的 QR 分解.

定义 6.8. 设 $A = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n) \in GL(n, \mathbb{C})$. 定义

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\gamma_i^H \alpha_k) \gamma_i = \alpha_k - \sum_{i=1}^{k-1} \frac{\beta_i^H \alpha_k}{\beta_i^H \beta_i} \beta_i, \quad \gamma_k = \frac{1}{\|\beta_k\|} \beta_k, \quad k = 1, 2, \dots, n.$$

容易验证, $\beta_1, \beta_2, \dots, \beta_n$ 两两正交, $Q = (\gamma_1 \ \gamma_2 \ \cdots \ \gamma_n)$ 是酉方阵. 上述由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 生成 $\beta_1, \beta_2, \dots, \beta_n$ 的过程称为 Gram-Schmidt 正交化过程. 由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 生成 $\gamma_1, \gamma_2, \dots, \gamma_n$ 的过程称为 Gram-Schmidt 标准正交化过程.

例 6.13. 对向量组 $\alpha_1 = (1, i, i)$, $\alpha_2 = (i, 1, i)$, $\alpha_3 = (i, i, 1)$ 作 Gram-Schmidt 标准正交化.

解答. $\beta_1 = \alpha_1 = (1, i, i)$, $\gamma_1 = \frac{1}{\sqrt{3}}\beta_1 = (\frac{1}{\sqrt{3}}, \frac{i}{\sqrt{3}}, \frac{i}{\sqrt{3}})$,
 $\beta_2 = \alpha_2 - \frac{1}{\sqrt{3}}\gamma_1 = (\frac{-1+3i}{3}, \frac{3-i}{3}, \frac{2i}{3})$, $\gamma_2 = \frac{3}{\sqrt{24}}\beta_2 = (\frac{-1+3i}{2\sqrt{6}}, \frac{3-i}{2\sqrt{6}}, \frac{2i}{2\sqrt{6}})$,
 $\beta_3 = \alpha_3 - \frac{1}{\sqrt{3}}\gamma_1 - \frac{1}{\sqrt{6}}\gamma_2 = (\frac{-1+3i}{4}, \frac{-1+3i}{4}, \frac{2-i}{2})$, $\gamma_3 = \frac{2}{\sqrt{10}}\beta_3 = (\frac{-1+3i}{2\sqrt{10}}, \frac{-1+3i}{2\sqrt{10}}, \frac{2-i}{\sqrt{10}})$. \square

定理 6.13 (Hadamard 不等式). 设 $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}^{n \times 1}$ 都是非零向量, 则

$$|\det(\alpha_1, \alpha_2, \dots, \alpha_n)| \leq \prod_{k=1}^n \|\alpha_k\|.$$

不等式取等号当且仅当 $\alpha_1, \alpha_2, \dots, \alpha_n$ 两两正交.

§6.4.2 酉相似

定义 6.9. 设 $A, B \in \mathbb{C}^{n \times n}$. 若存在酉方阵 P 使得 $A = PBP^{-1}$, 则称 A 与 B 酉相似.

容易验证, 复数矩阵的酉相似关系构成一个等价关系.

定理 6.14 (Schur 定理). 任意复数方阵可以酉相似成上三角方阵.

定义 6.10. 若复数方阵 A 满足 $AA^H = A^HA$, 则 A 称为规范方阵.

例如, 酉方阵、Hermite 方阵、反 Hermite 方阵都是规范方阵.

定理 6.15. 复方阵 A 是规范方阵的充分必要条件是 A 可以酉相似成对角方阵. 特别,

- 当 A 是酉方阵时, A 的任意特征值 λ 满足 $|\lambda| = 1$.
- 当 A 是 Hermite 方阵时, A 的任意特征值 $\lambda \in \mathbb{R}$.
- 当 A 是反 Hermite 方阵时, A 的任意特征值 λ 满足 $\operatorname{Re} \lambda = 0$.

定理 6.16 (Schur 不等式). 设 n 阶复数方阵 $A = (a_{ij})$ 的所有特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则

$$\sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i,j=1}^n |a_{ij}|^2.$$

不等式取等号当且仅当 A 是规范方阵.

§6.4.3 酉相抵

定义 6.11. 设 $A, B \in \mathbb{C}^{m \times n}$. 若存在酉方阵 P, Q 使得 $A = PBQ$, 则称 A 与 B 酉相抵.

容易验证, 复数矩阵的酉相抵关系构成一个等价关系.

定理 6.17. 对于任意 $A \in \mathbb{C}^{m \times n}$, 存在酉方阵 P, Q 使得

$$A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q \quad (6.4)$$

其中 $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, $r = \operatorname{rank}(A)$.

定义 6.12. (6.4) 式称为 A 的一个奇异值分解. $\begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix}$ 可以作为 A 所在的酉相抵等价类的代表元素, 称为 A 的酉相抵标准形. (6.4) 式中的酉方阵 P, Q 一般是不唯一的, 但 $\sigma_1, \sigma_2, \dots, \sigma_r$ 由 A 唯一确定. σ_k 称为 A 的第 k 个奇异值, 记作 $\sigma_k(A)$.

例 6.14. 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, $r = \text{rank}(A)$, $x \in \mathbb{C}^{n \times 1}$. 有下列结论.

1. $\sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\sum_{k=1}^r \sigma_k^2(A)}$. 此式称为 A 的 **Frobenius-范数**, 记作 $\|A\|_F$.
2. $\max_{\|x\|=1} \|Ax\| = \sigma_1(A)$. 此式称为 A 的**矩阵范数**, 记作 $\|A\|$ 或 $\|A\|_2$.
3. 当 A 列满秩时, $\min_{\|x\|=1} \|Ax\| = \sigma_n(A)$.

定义 6.13. 设 $m \times n$ 复数矩阵 A 有奇异值分解 (6.4), 则 $n \times m$ 复数矩阵

$$A^+ = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1} \quad (6.5)$$

称为 A 的 **Moore-Penrose 广义逆矩阵**.

定理 6.18. (6.5) 式中的 A^+ 是唯一的, 与酉方阵 P, Q 的选取无关.

证明. 设 $A = P_1 \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q_1 = P_2 \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q_2$. 由 $P_2^{-1}P_1 \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} = \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q_2 Q_1^{-1}$, 得 $P_2^{-1}P_1 = \begin{pmatrix} X_1 & X_2 \\ O & X_4 \end{pmatrix}$, $Q_2 Q_1^{-1} = \begin{pmatrix} \Sigma^{-1} X_1 \Sigma & O \\ Y_3 & Y_4 \end{pmatrix}$. 再由 $P_2^{-1}P_1$ 和 $Q_2 Q_1^{-1}$ 都是酉方阵, 得 $X_2 = O$, $Y_3 = O$. 因此, $Q_2 Q_1^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} = \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P_2^{-1}P_1$, 故 $Q_1^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P_1^{-1} = Q_2^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P_2^{-1}$. \square

习题

1. 证明例 6.11, 6.12, 6.14 和定理 6.11—6.17.

2. 证明:

(1) 任意 2 阶酉方阵形如 $\mu \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$, 其中 $z, w, \mu \in \mathbb{C}$ 满足 $|z|^2 + |w|^2 = 1$, $|\mu| = 1$.

(2) 任意 2 阶酉方阵形如 $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_1} \end{pmatrix} \begin{pmatrix} \cos \theta_2 & \sin \theta_2 \\ -\sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} e^{i\theta_3} & 0 \\ 0 & e^{i\theta_4} \end{pmatrix}$, 其中每个 $\theta_i \in \mathbb{R}$.

(3) 任意 2 阶规范方阵形如 $\lambda I + \mu P$, 其中 $\lambda, \mu \in \mathbb{C}$, P 是 Hermite 方阵.

(4) 任意 2 阶规范方阵形如 $\lambda I + \mu Q$, 其中 $\lambda, \mu \in \mathbb{C}$, Q 是酉方阵.

3. 设 $a, b, c, d \in \mathbb{C}$. 证明: $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ 与 $\begin{pmatrix} a & c \\ 0 & d \end{pmatrix}$ 酉相似当且仅当 $|b| = |c|$.

4. 计算 $A = \begin{pmatrix} 0 & i & 1 \\ 1 & 0 & i \\ i & 1 & 0 \end{pmatrix}$ 的 QR 分解、酉相似标准形和奇异值分解.

5. 求所有 n 阶酉方阵 A 使得 $\text{rank}(A - I) = 1$.

6. 设准上三角复数方阵 $A = (A_{ij})_{k \times k}$ 是规范方阵, 每个 A_{ii} 都是方阵. 证明: $A_{ij} = O, \forall i \neq j$.

7. 设映射 $\rho: \mathbb{C}^{n \times n} \rightarrow \mathbb{R}^{2n \times 2n}, A + Bi \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, 其中 $A, B \in \mathbb{R}^{n \times n}$.

证明: ρ 是一一映射, 并且对于任意 $X, Y \in \mathbb{C}^{n \times n}$, 有

$$(1) \rho(X + Y) = \rho(X) + \rho(Y), \rho(XY) = \rho(X)\rho(Y), \rho(X^H) = \rho(X)^T.$$

$$(2) X \text{ 是可逆方阵} \Leftrightarrow \rho(X) \text{ 是可逆方阵.}$$

$$(3) X \text{ 是酉方阵} \Leftrightarrow \rho(X) \text{ 是正交方阵.}$$

$$(4) X \text{ 是 Hermite 方阵} \Leftrightarrow \rho(X) \text{ 是对称方阵.}$$

$$(5) X \text{ 是反 Hermite 方阵} \Leftrightarrow \rho(X) \text{ 是反对称方阵.}$$

$$(6) X \text{ 是规范方阵} \Leftrightarrow \rho(X) \text{ 是规范方阵.}$$

8. 设 $A, B \in \mathbb{C}^{m \times n}$. 证明: $|\text{tr}(A^H B)| \leq \|A\|_F \|B\|_F$.

9. 设 $A \in \mathbb{C}^{m \times n}$. 证明: $X = A^+$ 是如下矩阵方程组的唯一解.

$$AXA = A, \quad XAX = X, \quad (AX)^H = AX, \quad (XA)^H = XA.$$

10. 设 $A \in \mathbb{C}^{m \times n}, b \in \mathbb{C}^{m \times 1}$. 使得 $\|Ax - b\|$ 最小的 $x \in \mathbb{C}^{n \times 1}$ 称为线性方程组 $Ax = b$ 的最小二乘解. 证明:

(1) $x = A^+b$ 是 $Ax = b$ 的一个最小二乘解.

(2) 在 $Ax = b$ 的所有最小二乘解中, $\|A^+b\|$ 最小.

11. 设 $\lambda \in \mathbb{C}$ 是 $A \in \mathbb{C}^{n \times n}$ 的任意特征值, $x_1 \leq \cdots \leq x_n$ 和 $y_1 \leq \cdots \leq y_n$ 分别是 Hermite 方阵 $X = \frac{1}{2}(A + A^H)$ 和 $Y = \frac{1}{2i}(A - A^H)$ 的所有特征值. 证明: $x_1 \leq \text{Re}(\lambda) \leq x_n, y_1 \leq \text{Im}(\lambda) \leq y_n$.

12. 证明: 对于任意酉方阵 $A \in U(n, \mathbb{C})$, 存在正交方阵 $P, Q \in O(n, \mathbb{R})$ 使得 PAQ 是对角阵.

13. 设 $A, B \in \mathbb{R}^{n \times n}$. 证明: 若 A 与 B 酉相似, 则 A 与 B 正交相似.

第七章 二次型

定义 7.1. 数域 \mathbb{F} 上的 n 元二次齐次函数

$$Q(x_1, x_2, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$$

称为 \mathbb{F} 上的 n 元二次型.

二次型 $Q(x_1, x_2, \dots, x_n)$ 可以表示为 $Q(x) = x^T A x$ 的形式, 其中 $x \in \mathbb{F}^{n \times 1}$, $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 是上三角方阵. 当 $\text{Char } \mathbb{F} \neq 2$ 时, $Q(x) = x^T S x$, 其中 $S = \frac{1}{2}(A + A^T)$ 是对称方阵.

在本章中, 我们始终假设 $\text{Char } \mathbb{F} \neq 2$, 二次型 $Q(x) = x^T A x$, 其中 A 是对称方阵, 称为 $Q(x)$ 的矩阵表示. 容易验证, 二次型的矩阵表示是存在且唯一的. 留作习题.

§7.1 二次型的化简

给定 \mathbb{F} 上的一个二次型 $Q(x) = x^T A x$, 我们希望通过可逆变量代换 $x = P y$ 或 $y = P^{-1} x$ 使

$$Q(x) = x^T A x = y^T (P^T A P) y = \tilde{Q}(y)$$

的形式简单. 也就是说, 求可逆方阵 P 使得 $P^T A P$ 的形式简单. 常用的方法是通过配方或者相合变换, 把二次型的矩阵表示化为对角形.

定义 7.2. 设 $A, B \in \mathbb{F}^{n \times n}$. 若存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $A = P^T B P$, 则称 A 与 B 在 \mathbb{F} 上相合.

容易验证, 矩阵的相合关系构成一个等价关系. 矩阵的对称性在相合关系下保持不变. 即若 A 与 B 相合, 则 A 是 (反) 对称方阵 $\Leftrightarrow B$ 是 (反) 对称方阵. 实数矩阵的正交相似关系是一种特殊的相合关系.

通过配方, 把 \mathbb{F} 上的二次型化为对角形的算法流程为:

1. 输入数域 \mathbb{F} 上的二次型 $Q(x_1, x_2, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$. 不妨设 $Q(x_1, x_2, \dots, x_n) \not\equiv 0$.
2. 若所有 $a_{ii} = 0$, 转至步骤 3. 否则, 存在 $a_{ii} \neq 0$. 不妨设 $a_{11} \neq 0$, 则

$$Q(x_1, x_2, \dots, x_n) = a_{11} \left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{2a_{11}} x_j \right)^2 + \tilde{Q}(x_2, \dots, x_n).$$

设 $y_1 = x_1 + \sum_{j=2}^n \frac{a_{1j}}{2a_{11}} x_j$. 把 $\tilde{Q}(x_2, \dots, x_n)$ 配方成 $d_2 y_2^2 + \dots + d_n y_n^2$ 的形式. 输出可逆变量代换 $y = Px$ 和

$$Q(x_1, x_2, \dots, x_n) = a_{11} y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2.$$

3. 所有 $a_{ii} = 0$, 但存在 $i \neq j$ 使得 $a_{ij} \neq 0$. 不妨设 $a_{12} \neq 0$, 则

$$Q(x_1, x_2, \dots, x_n) = a_{12} \left(x_1 + \sum_{j=3}^n \frac{a_{2j}}{a_{12}} x_j \right) \left(x_2 + \sum_{j=3}^n \frac{a_{1j}}{a_{12}} x_j \right) + \tilde{Q}(x_3, \dots, x_n).$$

设 $y_1 = x_1 + x_2 + \sum_{j=3}^n \frac{a_{1j} + a_{2j}}{a_{12}} x_j$, $y_2 = x_2 - x_1 + \sum_{j=3}^n \frac{a_{1j} - a_{2j}}{a_{12}} x_j$. 把 $\tilde{Q}(x_3, \dots, x_n)$ 配方成 $d_3 y_3^2 + \dots + d_n y_n^2$ 的形式. 输出可逆变量代换 $y = Px$ 和

$$Q(x_1, x_2, \dots, x_n) = \frac{a_{12}}{4} y_1^2 - \frac{a_{12}}{4} y_2^2 + d_3 y_3^2 + \dots + d_n y_n^2.$$

例 7.1. 通过配方, 把 \mathbb{R} 上的二次型 $Q(x) = x_1^2 + x_1 x_2 + x_1 x_3 + x_2^2 + x_2 x_3 + x_3^2$ 化为对角形.

解答.

$$\begin{aligned} Q(x) &= x_1^2 + (x_2 + x_3)x_1 + x_2^2 + x_2 x_3 + x_3^2 \\ &= (x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3)^2 + \frac{3}{4}x_2^2 + \frac{1}{2}x_2 x_3 + \frac{3}{4}x_3^2 \\ &= (x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3)^2 + \frac{3}{4}(x_2 + \frac{1}{3}x_3)^2 + \frac{2}{3}x_3^2 \\ &= y_1^2 + y_2^2 + y_3^2 \end{aligned}$$

其中 $y_1 = x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3$, $y_2 = \frac{\sqrt{3}}{2}(x_2 + \frac{1}{3}x_3)$, $y_3 = \frac{\sqrt{6}}{3}x_3$. □

例 7.2. 通过配方, 把 \mathbb{R} 上的二次型 $Q(x) = x_1 x_2 - 2x_1 x_3 + 3x_2 x_3$ 化为对角形.

解答.

$$\begin{aligned} Q(x) &= (x_1 + 3x_3)(x_2 - 2x_3) + 6x_3^2 \\ &= \frac{1}{4}(x_1 + x_2 + x_3)^2 - \frac{1}{4}(x_1 - x_2 + 5x_3)^2 + 6x_3^2 \\ &= y_1^2 + y_2^2 - y_3^2 \end{aligned}$$

其中 $y_1 = \frac{1}{2}(x_1 + x_2 + x_3)$, $y_2 = \sqrt{6}x_3$, $y_3 = \frac{1}{2}(x_1 - x_2 + 5x_3)$. □

通过相合变换, 把 \mathbb{F} 上的二次型化为对角形的算法流程为:

1. 输入数域 \mathbb{F} 上的二次型 $Q(x) = x^T A x$, 其中 $A = (a_{ij})$ 是对称方阵.

2. 对 A 作相合变换, 化为对角形. 具体过程如下. 不妨设 $A \neq O$.

(1) 若 $a_{11} \neq 0$, 则 $\prod_{i=2}^n T_{i1}(-\frac{a_{i1}}{a_{11}}) \cdot A \cdot \prod_{i=2}^n T_{1i}(-\frac{a_{1i}}{a_{11}}) = \begin{pmatrix} a_{11} & & \\ & B & \end{pmatrix}$. 再把 B 相合为对角形.

(2) 若 $a_{11} = 0$, 存在 $a_{ii} \neq 0$, 则 $\tilde{A} = S_{1i} A S_{i1}$ 的 $(1, 1)$ 元素为 a_{ii} , 化为情形 (1).

(3) 若所有 $a_{ii} = 0$, 则存在 $a_{ij} \neq 0$, $\tilde{A} = T_{ij}(1) A T_{ji}(1)$ 的 (i, i) 元素为 $2a_{ij}$, 化为情形 (2).

3. 设 $P^T A P = \text{diag}(d_1, d_2, \dots, d_n)$. 输出可逆变量代换 $x = Py$ 和

$$Q(x) = d_1 y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2.$$

例 7.3. 通过相合变换, 把 \mathbb{R} 上的二次型 $Q(x) = x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$ 化为对角形.

解答. $Q(x) = (x_1 \ x_2 \ x_3) \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. 对 $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$ 作相合变换,

$$\begin{pmatrix} 1 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ -\frac{1}{2} & 0 & 1 & 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & \frac{3}{4} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & 1 & 0 & 0 & \frac{2}{3} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{3}} & 0 & 0 & 1 & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{3}{\sqrt{6}} & 0 & 0 & 1 \end{pmatrix}.$$

因此, $Q(x) = y_1^2 + y_2^2 + y_3^2$, 其中 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{3}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. □

例 7.4. 通过相合变换, 把 \mathbb{R} 上的二次型 $Q(x) = x_1x_2 - 2x_1x_3 + 3x_2x_3$ 化为对角形.

解答. $Q(x) = (x_1 \ x_2 \ x_3) \begin{pmatrix} 0 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & \frac{3}{2} \\ -1 & \frac{3}{2} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. 对 $\begin{pmatrix} 0 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & \frac{3}{2} \\ -1 & \frac{3}{2} & 0 \end{pmatrix}$ 作相合变换,

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{2} & -1 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 0 & 1 & -1 & \frac{3}{2} & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{3}{2} & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{4} & \frac{5}{4} \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & \frac{5}{4} & -\frac{1}{4} \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{4} & 0 \\ -3 & 2 & 1 & 0 & 0 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 & 0 \\ -\frac{3}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ -\frac{3}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

因此, $Q(x) = y_1^2 + y_2^2 - y_3^2$, 其中 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{3}{\sqrt{6}} & -1 \\ 1 & \frac{2}{\sqrt{6}} & 1 \\ 0 & \frac{1}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. □

我们还可以通过正交相似变换, 把 \mathbb{R} 上的二次型化为对角形. 设对称实数方阵

$$A = P \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^T$$

其中 P 是正交方阵, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ 是 A 的所有特征值. 作正交变换 $x = Py$ 或 $y = P^T x$, 得

$$Q(x) = x^T A x = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2.$$

设 p, q 分别是 $\lambda_1, \lambda_2, \dots, \lambda_n$ 中的正数和负数的个数, 作变量代换

$$z = \operatorname{diag}(|\lambda_1|^{\frac{1}{2}}, \dots, |\lambda_p|^{\frac{1}{2}}, 1, \dots, 1, |\lambda_{n-q+1}|^{\frac{1}{2}}, \dots, |\lambda_n|^{\frac{1}{2}}) P^T x,$$

则二次型又可以化为

$$Q(x) = z_1^2 + \dots + z_p^2 - z_{n-q+1}^2 - \dots - z_n^2.$$

例 7.5. 求 \mathbb{R} 上的 n 元二次型 $Q(x) = (x_1 - x_n)^2 + \sum_{i=2}^n (x_i - x_{i-1})^2$ 的相合标准形.

解答. $Q(x) = x^T A x$, 其中 $A = \begin{pmatrix} 2 & -1 & & -1 \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ -1 & & -1 & 2 \end{pmatrix} = 2I - C - C^{-1}$, $C = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix}$.

由 $\varphi_C(x) = x^n - 1$, 得 C 的所有特征值为 $\{\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \mid k = 1, 2, \dots, n\}$, 故 A 的所有特征值为 $\{2 - 2\cos \frac{2k\pi}{n} \mid k = 1, 2, \dots, n\}$. 因此, A 与 $\begin{pmatrix} I_{n-1} & \\ & 0 \end{pmatrix}$ 相合, 存在可逆变量代换 $x = Py$ 使得 $Q(x) = y_1^2 + y_2^2 + \dots + y_{n-1}^2$. \square

由前面相合变换的算法流程, 可得

定理 7.1. \mathbb{F} 上任意对称方阵可以在 \mathbb{F} 上相合成对角方阵. 特别, 对称实数方阵可以在 \mathbb{R} 上相合成 $\text{diag}(I_p, -I_q, O)$ 的形式. 对称复数方阵可以在 \mathbb{C} 上相合成 $\text{diag}(I_r, O)$ 的形式.

判断任意数域 \mathbb{F} 上的两个方阵是否相合, 并不是一件平凡的事情, 这依赖于 \mathbb{F} 的性质. 可以证明^[1], 在数域 \mathbb{F} 上, 任意方阵 A 与 A^T 一定相合.

如果两个复数方阵相合, 则它们相抵. 因此, 可以把相抵标准形 $\text{diag}(I_r, O)$ 作为对称复数方阵在 \mathbb{C} 上的相合等价类的代表元素. 那么, 是否可以把 $\text{diag}(I_p, -I_q, O)$ 作为对称实数方阵在 \mathbb{R} 上的相合等价类的代表元素呢?

定理 7.2. 若实数方阵 $A = \text{diag}(I_p, -I_q, O)$ 与 $B = \text{diag}(I_r, -I_s, O)$ 在 \mathbb{R} 上相合, 则 $A = B$.

证明. 由 $\text{rank}(A) = \text{rank}(B)$ 可得 $p + q = r + s$. 设

$$\begin{pmatrix} I_p & O & O \\ O & -I_q & O \\ O & O & O \end{pmatrix} = \begin{pmatrix} P_{11}^T & P_{21}^T & P_{31}^T \\ P_{12}^T & P_{22}^T & P_{32}^T \\ P_{13}^T & P_{23}^T & P_{33}^T \end{pmatrix} \begin{pmatrix} I_r & O & O \\ O & -I_s & O \\ O & O & O \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}$$

其中 $P = (P_{ij})_{3 \times 3}$ 是可逆实数方阵, $P_{11} \in \mathbb{R}^{r \times p}$, $P_{22} \in \mathbb{R}^{s \times q}$. 假设 $A \neq B$, 不妨设 $p > r$, 则存在非零向量 $x \in \mathbb{R}^{p \times 1}$ 使得 $P_{11}x = \mathbf{0}$. 由 $I_p = P_{11}^T P_{11} - P_{21}^T P_{21}$, 得 $x^T x = -x^T P_{21}^T P_{21} x \leq 0$, 与 $x \neq \mathbf{0}$ 矛盾. \square

定义 7.3. 根据定理 7.2, $\text{diag}(I_p, -I_q, O)$ 可以作为对称实数方阵在 \mathbb{R} 上的相合等价类的代表元素, 称为对称实数方阵的**实相合标准形**. 整数对 p, q 是相合不变量, p 称为**正惯性指数**, q 称为**负惯性指数**.

利用线性空间和线性变换的概念, 正、负惯性指数有如下几何解释. 设 p, q 分别是 n 阶对称实数方阵 A 的正、负惯性指数, $Q_A(x) = x^T A x$. 线性空间 $V = \mathbb{R}^{n \times 1}$ 可以分解成 3 个两两不相交的子集, $V = V_A^- \cup V_A^0 \cup V_A^+$, 其中

$$V_A^- = \{x \in V \mid Q_A(x) < 0\}, \quad V_A^0 = \{x \in V \mid Q_A(x) = 0\}, \quad V_A^+ = \{x \in V \mid Q_A(x) > 0\}.$$

注意到

^[1]Dragomir Ž. Đoković and Khakim D. Ikramov, A square matrix is congruent to its transpose, Journal of Algebra **257** (2002), no.1, 97-105.

- V_A^0 包含一个 $n-p-q$ 维子空间.
- $V_A^+ \cup \{0\}$ 包含一个 p 维子空间, $V_A^- \cup V_A^0$ 包含一个 $n-p$ 维子空间.
- $V_A^- \cup \{0\}$ 包含一个 q 维子空间, $V_A^+ \cup V_A^0$ 包含一个 $n-q$ 维子空间.

因此,

- p 是集合 $V_A^+ \cup \{0\}$ 所包含线性空间的维数的最大值.
- q 是集合 $V_A^- \cup \{0\}$ 所包含线性空间的维数的最大值.

此外, 设 P 是 n 阶可逆实数方阵, $B = P^T A P$, 则 V 上的可逆线性变换 $x \mapsto P^{-1}x$ 分别把 V_A^-, V_A^0, V_A^+ 映射成 V_B^-, V_B^0, V_B^+ . 因此, A 和 B 的正(负)惯性指数相同.

习题

1. 设 $Q(x)$ 是数域 \mathbb{F} 上的二次型, $\text{Char } \mathbb{F} \neq 2$. 证明: 存在唯一的对称方阵 A 使得 $Q(x) = x^T A x$.

2. 证明: 在 $\text{Char } \mathbb{F} = 2$ 的数域 \mathbb{F} 上, 对称方阵 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 不能相合成对角方阵.

3. 把下列二次实系数多项式表示为 $(x-b)^T A(x-b) + c$ 的形式, 其中 b, c 是常量.

(1) $x_1^2 + x_2^2 + x_1 x_3 + x_2 x_3 + x_2 x_4 - 2x_1 + 3x_2 + x_4$

(2) $x_1^2 + x_1 x_2 - x_1 x_3 + x_1 x_4 + x_2^2 + x_3^2 - 3x_2 + x_3 - x_4$

(3) $x_1 x_3 - 2x_2 x_3 - x_2 x_4 - 2x_3 x_4 - x_1 + 3x_2 + 4x_3 + 3x_4$

(4) $x_1 x_2 - 2x_1 x_3 + 2x_2 x_3 - x_2 x_4 - 2x_3 x_4 - x_1 + 2x_2 - 2x_3 + x_4$

4. 通过配方或相合变换, 把下列实数域上的二次型化为标准形.

(1) $x_1^2 + 2x_1 x_2 - 2x_2^2 + 2x_2 x_3 + x_3^2 - 3x_3 x_4$

(2) $x_1^2 + 2x_1 x_3 - 4x_1 x_4 + x_2^2 + x_3^2 + 2x_3 x_4$

(3) $2x_1 x_2 + 2x_1 x_4 + 2x_2^2 - x_3^2 + 2x_3 x_4 + x_4^2$

(4) $x_1 x_3 - x_1 x_4 + 2x_2 x_3 - 2x_2 x_4 + 4x_3 x_4$

5. 求下列实数域上的二次型的标准形 (无需求变换矩阵).

(1) $\sum_{1 \leq i < j \leq n} (x_i - x_j)^2$ (2) $\sum_{1 \leq i < j \leq n} x_i x_j$ (3) $x_n x_1 + \sum_{i=1}^{n-1} x_i x_{i+1}$

6. (1) 证明: 任意 2 阶实数方阵 A 与 A^T 正交相似.

(2) 证明: 实数方阵 $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 与 A^T 既相似又相合, 但不正交相似.

7. (1) 设实数方阵 $A \neq -A^T$. 证明: 存在可逆实数方阵 P 使得 $P^T A P$ 是上三角方阵.

(2) 设实数方阵 $A = -A^T$. 证明: 存在可逆实数方阵 P 使得 $P^T A P = \begin{pmatrix} O & I_s & O \\ -I_s & O & O \\ O & O & O \end{pmatrix}$.

§7.2 正定方阵

定义 7.4. 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 若对于任意非零向量 $x \in \mathbb{C}^{n \times 1}$ 都有 $x^H Ax > 0$, 则 A 称为正定方阵. 类似地, 若对于任意 $x \in \mathbb{C}^{n \times 1}$ 都有 $x^H Ax \geq 0$, 则 A 称为半正定方阵.

方阵的正定性与二次型之间有着紧密的联系, 但是两者并不完全相同.

定义 7.5. 设 $A, B \in \mathbb{C}^{n \times n}$. 若存在可逆方阵 $P \in \mathbb{C}^{n \times n}$ 使得 $A = P^H B P$, 则称 A 与 B 在 \mathbb{C} 上共轭相合.

容易验证, 复数矩阵的共轭相合关系构成一个等价关系.

定理 7.3. 设 $P \in \mathbb{C}^{n \times n}$ 是可逆方阵. $A \in \mathbb{C}^{n \times n}$ 是 Hermite/反 Hermite/正定/半正定方阵当且仅当 PAP^H 是 Hermite/反 Hermite/正定/半正定方阵.

下面的定理给出了判断一个 Hermite 方阵是否正定方阵的一些充分必要条件.

定理 7.4. 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 下列叙述相互等价.

1. A 是正定的.
2. A 的特征值都是正实数.
3. 存在正定的 Hermite 方阵 P 使得 $A = P^2$.
4. 存在可逆的复数方阵 P 使得 $A = P^H P$.
5. 存在列满秩的复数矩阵 P 使得 $A = P^H P$.
6. A 的所有主子矩阵都是正定的.
7. A 的所有主子式都是正实数.
8. A 的顺序主子式都是正实数.

证明. 下面证明 $5 \Rightarrow 1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 6 \Rightarrow 7 \Rightarrow 8 \Rightarrow 4$. 从而叙述 1 ~ 8 相互等价.

由于 A 是 Hermite 方阵, 根据定理 6.15, 存在酉方阵 Q 使得 $A = Q \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) Q^H$, 其中每个 λ_i 都是实数.

$1 \Rightarrow 2$: 若某个 $\lambda_i \leq 0$, 则 $x = Qe_i$ 满足 $x^H Ax = \lambda_i \leq 0$, 与 A 的正定性矛盾.

$2 \Rightarrow 3$: $P = Q \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}) Q^H$ 满足要求.

$3 \Rightarrow 4$ 和 $4 \Rightarrow 5$ 显然.

$5 \Rightarrow 1$: 对于任意非零向量 $x \in \mathbb{C}^{n \times 1}$, $y = Px \neq 0$. 故 $x^H Ax = y^H y > 0$.

$5 \Rightarrow 6$ 和 $6 \Rightarrow 7$: 设 $P \in \mathbb{R}^{m \times n}$. A 的任意 k 阶主子矩阵均形如 $B^H B$, 其中 B 由 P 的某 k 列构成. 由 P 是列满秩的, 得 B 也是列满秩的. 由 $5 \Rightarrow 1$, $B^H B$ 是正定. 再由 $1 \Rightarrow 2$, $\det(B^H B) > 0$. $7 \Rightarrow 8$ 显然.

$8 \Rightarrow 4$: 设 Δ_k 是 A 的 k 阶顺序主子式. 根据 Schur 公式, 存在单位下三角方阵 L 使得

$$A = L \operatorname{diag} \left(\Delta_1, \frac{\Delta_2}{\Delta_1}, \dots, \frac{\Delta_n}{\Delta_{n-1}} \right) L^H = P^H P,$$

其中 $P = \operatorname{diag} \left(\sqrt{\Delta_1}, \sqrt{\frac{\Delta_2}{\Delta_1}}, \dots, \sqrt{\frac{\Delta_n}{\Delta_{n-1}}} \right) L^H$ 是可逆上三角方阵. 详细证明留作习题. □

判断 Hermite 方阵的半正定性, 也有与定理 7.4 类似的结果.

定理 7.5. 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 下列叙述相互等价.

1. A 是半正定的.
2. A 的特征值都是非负实数.
3. 存在半正定的 Hermite 方阵 P 使得 $A = P^2$.
4. 存在复数方阵 P 使得 $A = P^H P$.
5. 存在复数矩阵 P 使得 $A = P^H P$.
6. A 的所有主子矩阵都是半正定的.
7. A 的所有主子式都是非负实数.
8. 对于任意 $k = 1, 2, \dots, n$, A 的 k 阶主子式之和都是非负实数.

证明. 与定理 7.4 的上述证明过程类似, 可证 $5 \Rightarrow 1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 6 \Rightarrow 7 \Rightarrow 8$, 留作习题. 下证 $8 \Rightarrow 2$. 从而叙述 1 ~ 8 相互等价. 根据定理 5.2, $\varphi_A(x) = x^n + \sum_{k=1}^n (-1)^k \sigma_k x^{n-k}$, 其中 σ_k 是 A 的 k 阶主子式之和. 根据题设, $\sigma_k \geq 0, \forall k$. 由于 A 是 Hermite 方阵, A 的任意特征值 $\lambda \in \mathbb{R}$. 当 $x < 0$ 时, $(-1)^n \varphi_A(x) = |x|^n + \sum_{k=1}^n \sigma_k |x|^{n-k} > 0$. 故 $\lambda \geq 0$. \square



上述证明定理 7.5 ($8 \Rightarrow 2$) 的方法同样适用于证明定理 7.4 ($8 \Rightarrow 2$).

例 7.6. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是 Hermite 方阵, A 是正定的. 证明: AB 的任意特征值 $\lambda \in \mathbb{R}$. 特别, 若 B 是正定的, 则 $\lambda > 0$; 若 B 是半正定的, 则 $\lambda \geq 0$.

证明. 根据定理 7.4, 存在可逆方阵 $P \in \mathbb{C}^{n \times n}$ 使得 $A = P^H P$. 从而, $AB = P^H P B P^H P^{-H}$ 与 $C = P B P^H$ 相似, λ 是 Hermite 方阵 C 的特征值. 因此, $\lambda \in \mathbb{R}$. 当 B 是正定的时, C 是正定的, $\lambda > 0$. 当 B 是半正定的时, C 是半正定的, $\lambda \geq 0$. \square

例 7.7. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是半正定的 Hermite 方阵, 则存在可逆方阵 $P \in \mathbb{C}^{n \times n}$ 使得 $P^H A P$ 和 $P^H B P$ 都是对角方阵.

证明. 易知 $C = A + B$ 也是半正定的. 故存在可逆方阵 $Q \in \mathbb{C}^{n \times n}$ 使得 $Q^H C Q = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 设

$$Q^H A Q = \begin{pmatrix} A_1 & A_2 \\ A_2^H & A_3 \end{pmatrix}, \quad Q^H B Q = \begin{pmatrix} I - A_1 & -A_2 \\ -A_2^H & -A_3 \end{pmatrix}$$

其中 A_1 和 A_3 分别是 r 和 $n - r$ 阶 Hermite 方阵. 由 $Q^H A Q$ 和 $Q^H B Q$ 都是半正定的, 得 $\pm A_3$ 都是半正定的, 从而 $A_3 = O$. 若 $A_2 \neq O$, 则存在 $u \in \mathbb{C}^{r \times 1}$ 使得 $v = A_2^H u \neq 0$. 当 $\lambda \rightarrow -\infty$ 时,

$$\begin{pmatrix} u^H & \lambda v^H \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_2^H & O \end{pmatrix} \begin{pmatrix} u \\ \lambda v \end{pmatrix} = u^H A_1 u + 2\lambda u^H A_2 v < 0$$

与 $Q^H A Q$ 的半正定性矛盾. 故 $A_2 = O$. 设酉方阵 U 使得 $D = U^H A_1 U$ 是对角方阵, 则可逆方阵 $P = Q \begin{pmatrix} U & \\ & I \end{pmatrix}$ 使得 $P^H A P = \begin{pmatrix} D & \\ & O \end{pmatrix}$ 和 $P^H B P = \begin{pmatrix} I - D & \\ & O \end{pmatrix}$ 都是对角方阵. \square

习题

1. 完成定理 7.3 和定理 7.5 的证明.

2. 设 $A \in \mathbb{R}^{n \times n}$ 是正定的对称方阵. 分别求 $\int_{x^T A x \leq 1} dx_1 \cdots dx_n$ 和 $\int_{\mathbb{R}^n} e^{-x^T A x} dx_1 \cdots dx_n$, 并且表示成 A 的函数形式.

3. 设 $0 \leq \theta_1 \leq \theta_2 \leq \theta_3 \leq \pi$. 证明:

$$\begin{pmatrix} 1 & \cos \theta_1 & \cos \theta_2 \\ \cos \theta_1 & 1 & \cos \theta_3 \\ \cos \theta_2 & \cos \theta_3 & 1 \end{pmatrix} \text{ 是正定的} \Leftrightarrow \theta_3 < \theta_1 + \theta_2 < 2\pi - \theta_3.$$

4. 证明: 对于任意 Hermite 方阵 A , 存在可逆方阵 P 使得 $P^H A P = \text{diag}(I_p, -I_q, O)$, 其中 p, q 由 A 唯一确定.

5. 设 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^H & A_{22} \end{pmatrix}$ 是正定的 Hermite 方阵, $A^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^H & B_{22} \end{pmatrix}$, A_{11} 与 B_{11} 同阶. 证明:

(1) $A_{11}, A_{22}, A^{-1}, B_{11}, B_{22}$ 都是正定的.

(2) $\det(A) \leq \det(A_{11}) \det(A_{22})$, $\det(A_{11}) \det(B_{11}) \geq 1$, $\det(A_{22}) \det(B_{22}) \geq 1$.

(3) $A_{11} B_{11}$ 和 $A_{22} B_{22}$ 的任意特征值 λ 都是实数, 并且 $\lambda \geq 1$.

6. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是正定的 Hermite 方阵. 证明: 若 $A - B$ 是正定/半正定的, 则 $B^{-1} - A^{-1}$ 是正定/半正定的.

7. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是半正定的 Hermite 方阵, m 是正整数.

(1) 证明: 存在唯一的半正定的 Hermite 方阵 X 使得 $X^m = A$.

(2) 举例: $A - B$ 是正定的, $A^m - B^m$ 不是半正定的.

(3) 证明: 若 $A^2 - B^2$ 是正定/半正定的, 则 $A - B$ 是正定/半正定的.

(4) 给定 $m \geq 3$. 证明或否定: 若 $A^m - B^m$ 是正定/半正定的, 则 $A - B$ 是正定/半正定的.

8. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是 Hermite 方阵. 证明:

(1) 若 A 是正定的, 则存在上(下)三角可逆方阵 P 使得 $A = P^H P$.

(2) 若 A 是正定的, 则存在可逆方阵 P 使得 $P^H A P$ 和 $P^H B P$ 都是对角方阵.

(3) 若 A 是半正定的, 则 AB 的特征值都是实数.

(4) 若 A, B 都是半正定的, 则 AB 的特征值都是非负实数.

9. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是 Hermite 方阵. 回答下列问题, 并证明你的结论.

(1) 若 A 是半正定的, 是否一定存在上(下)三角方阵 P 使得 $A = P^H P$?

(2) 若 A 是半正定的, 是否一定存在可逆方阵 P 使得 $P^H A P$ 和 $P^H B P$ 都是对角方阵?

(3) 若 A, B 都是半正定的, $AB + BA$ 是否一定是半正定的?

(4) AB 的特征值是否一定是实数?

10. 设 $A = (a_{ij})$ 和 $B = (b_{ij})$ 都是 n 阶复数方阵, $C = (a_{ij}b_{ij})$, $D = A \otimes B$. 证明:
- (1) 若 A, B 都是 Hermite 方阵, 则 C, D 都是 Hermite 方阵.
 - (2) 若 A, B 都是正定的, 则 C, D 都是正定的.
 - (3) 若 A, B 都是半正定的, 则 C, D 都是半正定的.
11. 设 $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 证明:
- (1) 若 $\forall i, a_{ii} > \sum_{j \neq i} |a_{ij}|$, 则 A 是正定的.
 - (2) 若 $\forall i, a_{ii} \geq \sum_{j \neq i} |a_{ij}|$, 则 A 是半正定的.
12. 设 A 和 D 分别是图 G 的邻接矩阵和度矩阵 (见第二章 §2.1 例 2.6). $L = D - A$ 称为 G 的 **Laplace 矩阵** 或 **Kirchhoff 矩阵**. 证明: 存在整数矩阵 P 使得 $L = P^T P$, 从而 L 是半正定的.
13. 设 $A \in \mathbb{C}^{n \times n}$ 不是 Hermite 方阵, $B = \frac{1}{2}(A + A^H)$ 是正定的. 证明:
- (1) A 的任意特征值 λ 满足 $\operatorname{Re}(\lambda) > 0$.
 - (2) $|\det(A)| > \det(B)$. 特别, 当 $A \in \mathbb{R}^{n \times n}$ 时, $\det(A) > \det(B)$.
14. 设 $A, B, C \in \mathbb{C}^{n \times n}$, A, B 都是正定的 Hermite 方阵, $A = B + C + C^H$. 证明: $(B + C)^{-1}C$ 的任意特征值 λ 满足 $|\lambda| < 1$.

§7.3 一些例子

例 7.8. 对于任意复数方阵 A , 存在酉方阵 P 和半正定 Hermite 方阵 S 使得 $A = SP$, 其中 S 是唯一的. 此矩阵乘积分解称为 A 的极分解.

证明. 由奇异值分解易得极分解的存在性, $A = U\Sigma V = (U\Sigma U^H)(UV)$, 其中 U, V 是酉方阵, Σ 是对角阵. 下证 S 的唯一性. 设 $S_1 P_1$ 和 $S_2 P_2$ 都是 A 的极分解,

$$S_1 = Q_1 \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) Q_1^H, \quad S_2 = Q_2 \operatorname{diag}(\mu_1, \mu_2, \dots, \mu_n) Q_2^H,$$

其中 Q_1, Q_2 都是酉方阵, 则

$$\begin{aligned} AA^H = S_1^2 = S_2^2 &\Rightarrow Q_2^H Q_1 \operatorname{diag}(\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2) = \operatorname{diag}(\mu_1^2, \mu_2^2, \dots, \mu_n^2) Q_2^H Q_1 \\ &\Rightarrow Q_2^H Q_1 \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \operatorname{diag}(\mu_1, \mu_2, \dots, \mu_n) Q_2^H Q_1 \Rightarrow S_1 = S_2. \quad \square \end{aligned}$$

例 7.9. 设 $A \in \mathbb{R}^{n \times n}$ 是正定对称方阵, $b \in \mathbb{R}^{n \times 1}$. 求 n 元二次函数 $f(x) = x^T A x + 2b^T x$ 的最小值, 其中 $x \in \mathbb{R}^{n \times 1}$.

解答. 设 $A = P^T P$, $y = Px$, 则

$$f(x) = y^T y + 2(P^{-T}b)^T y = \|y + P^{-T}b\|^2 - \|P^{-T}b\|^2.$$

当且仅当 $y = -P^{-T}b$ 即 $x = -A^{-1}b$ 时, $f(x)$ 取得最小值 $-\|P^{-T}b\|^2 = -b^T A^{-1}b$. \square

例 7.10. 设 $A \in \mathbb{R}^{n \times n}$ 是正定对称方阵, $c \in \mathbb{R}^{n \times 1}$. 求 n 元线性函数 $f(x) = c^T x$ 在约束条件 $x^T A x \leq 1$ 下的最大值, 其中 $x \in \mathbb{R}^{n \times 1}$.

解答. 设 $A = P^T P$, $y = Px$, 则 $f(x) = (P^{-T}c)^T y$, 条件 $x^T A x \leq 1$ 化为 $\|y\| \leq 1$. 根据 Cauchy 不等式, 当且仅当 $y = \frac{1}{\|P^{-T}c\|} P^{-T}c$ 即 $x = \frac{1}{\sqrt{c^T A^{-1}c}} A^{-1}c$ 时, $f(x)$ 取得最大值 $\sqrt{c^T A^{-1}c}$. \square

定义 7.6. 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵, $x \in \mathbb{C}^{n \times 1}$ 是非零向量. 向量函数

$$R_A(x) = \frac{x^H A x}{x^H x}$$

称为 Rayleigh^[2]商.

例 7.11 (Rayleigh-Ritz^[3]定理). 设 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ 是 Hermite 方阵 A 的所有特征值, 则有

$$\lambda_1 = \max_{x \in \mathbb{C}^{n \times 1}} R_A(x), \quad \lambda_n = \min_{x \in \mathbb{C}^{n \times 1}} R_A(x).$$

证明. 设 $A = P \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^H$, 其中 P 是酉方阵, 再设 $y = (y_i) = P^H x$, 则

$$R_A(x) = \frac{x^H A x}{x^H x} = \frac{\sum_{i=1}^n \lambda_i |y_i|^2}{\sum_{i=1}^n |y_i|^2} \in [\lambda_n, \lambda_1].$$

当 $y = \mathbf{e}_1$ 时, $R_A(x)$ 取得最大值 λ_1 . 当 $y = \mathbf{e}_n$ 时, $R_A(x)$ 取得最小值 λ_n . \square

^[2]John William Strutt, 3rd Baron Rayleigh, 1842–1919, 英国物理学家. 获 1904 年 Nobel 物理学奖.

^[3]Walther Ritz, 1878–1909, 瑞士物理学家.

例 7.12. 设 $A \in \mathbb{C}^{m \times n}$, ε 是给定的正数, $\|\cdot\|$ 是例 6.14 中的矩阵范数.

$$\text{rank}_\varepsilon(A) = \min_{\|B-A\| \leq \varepsilon} \text{rank}(B)$$

称为 A 的 ε -近似秩. 证明: $\text{rank}_\varepsilon(A) = \max\{i \in \mathbb{N} \mid \sigma_i(A) > \varepsilon\}$.

证明. 设 $A = P \text{diag}(\sigma_1, \dots, \sigma_r, O)Q$ 是奇异值分解, $k = \max\{i \in \mathbb{N} \mid \sigma_i > \varepsilon\}$.

一方面, 设 $B = P \text{diag}(\sigma_1, \dots, \sigma_k, O)Q$. 由 $\|B-A\| = \sigma_{k+1} \leq \varepsilon$, 得 $\text{rank}_\varepsilon(A) \leq \text{rank}(B) = k$.

另一方面, 设 $B = P \begin{pmatrix} B_1 & B_2 \end{pmatrix} Q \in \mathbb{C}^{m \times n}$ 满足 $\|B-A\| \leq \varepsilon$, $B_1 \in \mathbb{C}^{m \times k}$. 假设 $\text{rank}(B) < k$,

则 $\text{rank}(B_1) < k$, 从而存在非零向量 $x \in \mathbb{C}^{k \times 1}$ 使得 $B_1 x = \mathbf{0}$. 设 $y = Q^{-1} \begin{pmatrix} x \\ \mathbf{0} \end{pmatrix}$. 由 $By = \mathbf{0}$ 可得

$$\|B-A\| \geq \frac{\|(B-A)y\|}{\|y\|} = \frac{\|Ay\|}{\|y\|} = \sqrt{\frac{|\sigma_1 x_1|^2 + \dots + |\sigma_k x_k|^2}{|x_1|^2 + \dots + |x_k|^2}} \geq \sigma_k > \varepsilon, \text{ 矛盾. 故 } \text{rank}(B) \geq k.$$

综上, $\text{rank}_\varepsilon(A) = k$. □

习题

1. 证明: 对于任意规范复数方阵 A , 存在酉方阵 P 和半正定 Hermite 方阵 S 使得 $A = PS = SP$, 其中 S 是唯一的.
2. 设 $A, B \in \mathbb{R}^{n \times n}$ 都是对称方阵, B 是正定的, $b \in \mathbb{R}^{n \times 1}$. 求函数 $f(x) = x^T A x + 2b^T x$ 在约束条件 $x^T B x \leq 1$ 下的取值范围, 其中 $x \in \mathbb{R}^{n \times 1}$.
3. 设实数 x_1, x_2, \dots, x_n 满足 $x_1^2 + \sum_{i=2}^n (x_{i-1} + x_i)^2 = 1$. 求每个 x_i 的取值范围.
4. 设实数 x_1, x_2, \dots, x_n 满足 $\sum_{i=1}^n x_i^2 + \frac{1}{n} \sum_{1 \leq i < j \leq n} x_i x_j = 1$. 求 $\sum_{i=1}^n x_i$ 的取值范围.
5. 设 $A \in \{-1, 1\}^{n \times n}$ 满足 $AA^T = nI$. 证明: 若 A 的某个 $p \times q$ 子矩阵的元素都是 1, 则 $pq \leq n$.
6. 设 $A \in \mathbb{C}^{m \times n}$, ε 是给定的正数, $\|\cdot\|_F$ 是矩阵的 Frobenius-范数. 证明:

$$\min_{\|X-A\|_F \leq \varepsilon} \text{rank}(X) = \max \left\{ k \in \mathbb{N} \mid \sum_{i \geq k+1} \sigma_i^2(A) \leq \varepsilon^2 \right\}.$$

7. (Courant-Fischer 定理) 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵, $\lambda_k(A)$ 是 A 的第 k 大特征值. 证明:

$$\lambda_k(A) = \max_{\dim V=k} \min_{x \in V} R_A(x), \quad \lambda_{n-k+1}(A) = \min_{\dim V=k} \max_{x \in V} R_A(x),$$

其中 V 是 \mathbb{C}^n 的任意 k 维子空间.

8. 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵, $B = A \begin{bmatrix} 1 & & \\ & \frac{1}{2} & \\ & & \ddots \\ & & & m \end{bmatrix}$, $\lambda_k(X)$ 是 X 的第 k 大特征值. 证明:

$$\lambda_{i+n-m}(A) \leq \lambda_i(B) \leq \lambda_i(A), \quad \forall i \leq m.$$

9. 设 $A, B \in \mathbb{C}^{n \times n}$ 都是 Hermite 方阵, $\lambda_k(X)$ 是 X 的第 k 大特征值. 证明:

(1) (Weyl^[4]不等式)

$$\lambda_{i+j-1}(A+B) \leq \lambda_i(A) + \lambda_j(B) \leq \lambda_{i+j-n}(A+B), \quad \forall i+j \geq n+1$$

$$(2) \sum_{k=1}^n |\lambda_k(A) - \lambda_k(B)|^2 \leq \|A - B\|_F^2.$$

$$(3) |\lambda_k(A) - \lambda_k(B)| \leq \|A - B\|, \quad \forall k.$$

10. ^[5] 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, 当 i, j 都整除 n 时, $a_{ij} = \frac{1}{\text{lcm}(i, j)}$, 否则 $a_{ij} = 0$. 证明:

$$x^T A x \geq \frac{\varphi(n)}{n} x_1^2, \quad \forall x = (x_1, \dots, x_n)^T \in \mathbb{R}^{n \times 1}$$

其中 $\varphi(n)$ 是 **Euler 函数**, 即不大于 n 且与 n 互素的正整数的个数.

11. 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, 其中 $a_{ij} = \text{gcd}(i, j)$. 证明:

(1) A 是正定方阵.

$$(2) A = P^T P, \text{ 其中 } P = (p_{ij}) \in \mathbb{R}^{m \times n}, m = \text{lcm}(1, 2, \dots, n), p_{ij} = \begin{cases} \frac{j}{\sqrt{m}}, & j | i; \\ 0, & j \nmid i. \end{cases}$$

$$(3) A = Q^T Q, \text{ 其中 } Q = (q_{ij}) \in \mathbb{R}^{n \times n}, q_{ij} = \begin{cases} \sqrt{\varphi(i)}, & i | j; \\ 0, & i \nmid j. \end{cases}$$

12. ^[6] 求实二次型 $Q = \sum_{1 \leq i < j \leq 2n} \min((j-i)^2, (2n-j+i)^2) x_i x_j$ 在条件 $\sum_{i=1}^{2n} |x_i| = 1$ 下的最大值.

13. 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, 其中 $a_{ij} = \frac{1}{\max(i, j)}$. 证明: A 是正定方阵, 并且 $\|A\| < 4$.

14. 设 λ_n 是 n 阶方阵 $A = (a_{ij})$ 的最小特征值, 其中 $a_{ij} = n - |i - j|$. 证明:

$$(1) \lambda_n > \frac{1}{2} \text{ 并且 } \lim_{n \rightarrow \infty} \lambda_n = \frac{1}{2}.$$

(2) 当 $n = 2m$ 时, λ_n 是 m 阶方阵 $B = (b_{ij})$ 的最小特征值, 其中 $b_{ij} = 2 \min(i, j) - 1$.

(3) 当 $n = 2m - 1$ 时, λ_n 是 DB 的最小特征值, 其中 B 同上, $D = \text{diag}(I_{m-1}, m)$.

(4) 当 $n \geq 3$ 时, $\lambda_n < \lambda_{n-1}$.

^[4]Hermann Klaus Hugo Weyl, 1885-1995, 德国数学家、物理学家.

^[5]2018年CGMO第3题.

^[6]2019年IMO中国国家集训队考试二第3题.

第八章 线性空间

§8.1 基本概念

如同“集合”和“元素”是数学的最基本概念一样，“线性空间”和“向量”是线性代数的最基本概念。粗略地说，具有加法和数乘运算的非空集合称为**线性空间**。也就是说，一个线性空间具有四个要素 $(V, \mathbb{F}, +, \cdot)$ ，其中 V 是非空集合， \mathbb{F} 是数域，加法运算 $+$ 是 $V \times V \rightarrow V$ 的映射，数乘运算 \cdot 是 $\mathbb{F} \times V \rightarrow V$ 的映射。但是，并非任意一个 $V \times V \rightarrow V$ 的映射都可以作为线性空间的“加法”，也并非任意一个 $\mathbb{F} \times V \rightarrow V$ 的映射都可以作为线性空间的“数乘”，它们必须满足一定的运算律，称为线性空间的公理。严格地说，

定义 8.1. 设 \mathbb{F} 是一个数域， \mathbb{F} 的元素称为**数**。设 V 是一个非空集合， V 的元素称为**向量**。在 V 上定义了加法运算 $\alpha + \beta \in V$ 和数乘运算 $\lambda \cdot \alpha \in V$ ，其中 $\alpha, \beta \in V$ ， $\lambda \in \mathbb{F}$ 。若这两种运算满足下列运算律

- | | |
|----------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| (A1) 加法结合律 | $\forall \alpha, \beta, \gamma \in V, (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ |
| (A2) 加法交换律 | $\forall \alpha, \beta \in V, \alpha + \beta = \beta + \alpha$ |
| (A3) 加法有单位元 | 存在零向量 $\mathbf{0} \in V$ 使得 $\forall \alpha \in V, \mathbf{0} + \alpha = \alpha$ |
| (A4) 加法有逆元 | $\forall \alpha \in V$ ，存在 $\beta \in V$ 使得 $\alpha + \beta = \mathbf{0}$ |
| (M1) 数乘结合律 | $\forall \alpha \in V$ 和 $\lambda, \mu \in \mathbb{F}, \lambda \cdot (\mu \cdot \alpha) = (\lambda\mu) \cdot \alpha$ |
| (M2) 数乘有单位元 | $\forall \alpha \in V, 1 \cdot \alpha = \alpha$ |
| (D1) 数乘对数的分配律 | $\forall \alpha \in V$ 和 $\lambda, \mu \in \mathbb{F}, (\lambda + \mu) \cdot \alpha = (\lambda \cdot \alpha) + (\mu \cdot \alpha)$ |
| (D2) 数乘对向量的分配律 | $\forall \alpha, \beta \in V$ 和 $\lambda \in \mathbb{F}, \lambda \cdot (\alpha + \beta) = (\lambda \cdot \alpha) + (\lambda \cdot \beta)$ |

则代数结构 $(V, \mathbb{F}, +, \cdot)$ 称为**线性空间**。通常称 V 在加法和数乘运算下构成数域 \mathbb{F} 上的线性空间。在不引起歧义的情况下，可以简称 V 是 \mathbb{F} 上的线性空间。

当 $\mathbb{F} = \mathbb{R}$ 时， V 称为**实线性空间**；当 $\mathbb{F} = \mathbb{C}$ 时， V 称为**复线性空间**。

只有一个元素的线性空间 $V = \{\mathbf{0}\}$ 称为**零空间**。

 每个线性空间由“非空集合 V 、数域 \mathbb{F} 、加法运算 $+$ 、数乘运算 \cdot ”这四个基本要素构成。对于任意两个线性空间，只要它们有一个要素不同，它们就是不同的线性空间。

例 8.1. 在通常的加法和乘法运算下， $(\mathbb{C}^2, \mathbb{C}, +, \cdot)$ 和 $(\mathbb{C}^2, \mathbb{R}, +, \cdot)$ 是不同的线性空间。

例 8.2. 在 \mathbb{C}^2 上定义加法运算 $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$ 和两种数乘运算

$$\lambda \cdot (a, b) = (\lambda a, \lambda b), \quad \lambda * (a, b) = (\bar{\lambda} a, \bar{\lambda} b),$$

则 $(\mathbb{C}^2, \mathbb{C}, +, \cdot)$ 和 $(\mathbb{C}^2, \mathbb{C}, +, *)$ 是不同的线性空间。

定理 8.1. 数域 \mathbb{F} 上的线性空间 V 具有下列性质.

1. V 的零向量 $\mathbf{0}$ 是唯一的.
2. 任意向量 $\alpha \in V$ 的加法逆元 β 是唯一的, 并且 $\beta = (-1) \cdot \alpha$. 常记作 $\beta = -\alpha$.
通过加法逆元, 可定义两个向量的减法运算 $\alpha_1 - \alpha_2 = \alpha_1 + (-\alpha_2)$.
3. 对于任意 $\lambda \in \mathbb{F}$, $\alpha \in V$, 都有 $\lambda \cdot \mathbf{0} = 0 \cdot \alpha = \mathbf{0}$.
4. 设 $\lambda \in \mathbb{F}$, $\alpha \in V$. 若 $\lambda\alpha = \mathbf{0}$, 则 $\lambda = 0$ 或 $\alpha = \mathbf{0}$.

定义 8.2. 设 \mathbb{F} 是定义了加法运算 $a + b$ 和乘法运算 $a * b$ 的非空集合. 若下列运算律成立, 则代数结构 $(\mathbb{F}, +, *)$ 称为**数域**, 或者简称 \mathbb{F} 是一个数域.

- | | |
|-------------|-----------------------------------------------------------------------------------|
| (A1) 加法结合律 | $\forall a, b, c \in \mathbb{F}, (a + b) + c = a + (b + c)$ |
| (A2) 加法交换律 | $\forall a, b \in \mathbb{F}, a + b = b + a$ |
| (A3) 加法有单位元 | 存在零元素 $0 \in \mathbb{F}$ 使得 $\forall a \in \mathbb{F}, 0 + a = a$ |
| (A4) 加法有逆元 | $\forall a \in \mathbb{F}$, 存在 $b \in \mathbb{F}$ 使得 $a + b = 0$ |
| (M1) 乘法结合律 | $\forall a, b, c \in \mathbb{F}, (a * b) * c = a * (b * c)$ |
| (M2) 乘法交换律 | $\forall a, b \in \mathbb{F}, a * b = b * a$. |
| (M3) 乘法有单位元 | 存在幺元素 $1 \in \mathbb{F}$ 使得 $\forall a \in \mathbb{F}, 1 * a = a$ |
| (M4) 乘法有逆元 | $\forall a \in \mathbb{F} \setminus \{0\}$, 存在 $b \in \mathbb{F}$ 使得 $a * b = 1$ |
| (D1) 加乘分配律 | $\forall a, b, c \in \mathbb{F}, (a + b) * c = (a * c) + (b * c)$ |

比较以上关于“线性空间”和“数域”的定义, 两者有许多相似之处. 我们可以认为“向量”是“数”的推广, 向量运算降低了对“乘法”的要求, 从而有更广泛的应用.

例 8.3. 容易验证, 下面是一些常见的线性空间的例子.

1. 在复数的加法和乘法运算下, 复数域 \mathbb{C} 是实数域 \mathbb{R} 上的线性空间, 也是有理数域 \mathbb{Q} 上的线性空间.
2. 在多项式的加法和数乘运算下, 多元多项式的全体 $\mathbb{F}[x_1, \dots, x_m]$ 是 \mathbb{F} 上的线性空间.
3. 在多项式的加法和数乘运算下, 次数小于 n 的一元多项式的全体 $\mathbb{F}_n[x]$ 是 \mathbb{F} 上的线性空间.
4. 在函数的加法和数乘运算下, 区间 $[a, b]$ 上的连续实函数的全体 $\mathcal{C}[a, b]$ 是 \mathbb{R} 上的线性空间.
5. 在级数的加法和数乘运算下, 实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体 $\mathbb{R}[[x]]$ 是 \mathbb{R} 上的线性空间.
6. 在数组向量的加法和数乘运算下, \mathbb{F} 上的 n 维数组向量的全体 \mathbb{F}^n 是 \mathbb{F} 上的线性空间.
7. 在数列的加法和数乘运算下, \mathbb{F} 上的无穷数列 $\{a_n\}_{n \in \mathbb{N}}$ 的全体是 \mathbb{F} 上的线性空间.
8. 给定 $A \in \mathbb{F}^{m \times n}$, 在数组向量的加法和数乘运算下, $\{x \in \mathbb{F}^{n \times 1} \mid Ax = \mathbf{0}\}$ 是 \mathbb{F} 上的线性空间, 称为线性方程组 $Ax = \mathbf{0}$ 在 \mathbb{F} 上的解空间.
9. 在矩阵的加法和数乘运算下, $m \times n$ 矩阵的全体 $\mathbb{F}^{m \times n}$ 是 \mathbb{F} 上的线性空间.
10. 在矩阵的加法和数乘运算下, \mathbb{F} 上的 n 元二次型 $Q(x)$ 的全体是 \mathbb{F} 上的线性空间.

定义 8.3. 设 U 是 V 的非空子集. 若 U 在 V 的加法和数乘运算下封闭, 即满足

$$\alpha + \beta \in U, \quad \lambda\alpha \in U, \quad \forall \alpha, \beta \in U, \lambda \in \mathbb{F}$$

则线性空间 $(U, \mathbb{F}, +, \cdot)$ 称为 $(V, \mathbb{F}, +, \cdot)$ 的**子空间**, 简称 U 是 V 的子空间. 特别, $\{\mathbf{0}\}$ 和 V 称为 V 的**平凡子空间**.

例 8.4. 下面是一些常见的子空间的例子.

1. 设 $A \in \mathbb{F}^{m \times n}$. $\{x \in \mathbb{F}^{n \times 1} \mid Ax = \mathbf{0}\}$ 是 $\mathbb{F}^{n \times 1}$ 的子空间.
2. 设 $A \in \mathbb{F}^{n \times n}$. $\{X \in \mathbb{F}^{n \times n} \mid AX = XA\}$ 是 $\mathbb{F}^{n \times n}$ 的子空间.
3. $\mathbb{F}_n[x]$ 是 $\mathbb{F}[x]$ 的子空间.
4. $\mathbb{R}[x]$ 是 $\mathbb{R}[[x]]$ 的子空间.

定义 8.4. 设 $(U, \mathbb{F}, \oplus, *)$ 和 $(V, \mathbb{F}, +, \cdot)$ 都是数域 \mathbb{F} 上的线性空间. 如果存在一一映射 $\rho: U \rightarrow V$ 满足

$$\rho(\alpha \oplus \beta) = \rho(\alpha) + \rho(\beta), \quad \rho(\lambda * \alpha) = \lambda \cdot \rho(\alpha), \quad \forall \alpha, \beta \in U, \lambda \in \mathbb{F}$$

则 ρ 称为从 $(U, \mathbb{F}, \oplus, *)$ 到 $(V, \mathbb{F}, +, \cdot)$ 的**同构映射**, 简称 \mathbb{F} 上的线性空间 U 与 V **同构**.

对于不同数域上的两个线性空间 $(V_1, \mathbb{F}_1, \oplus, *)$ 与 $(V_2, \mathbb{F}_2, +, \cdot)$, 可以类似地定义同构的概念. 如果存在一一映射 $\rho: V_1 \rightarrow V_2$ 和一一映射 $\sigma: \mathbb{F}_1 \rightarrow \mathbb{F}_2$ 满足

$$\begin{aligned} \rho(\alpha \oplus \beta) &= \rho(\alpha) + \rho(\beta), & \rho(\lambda * \alpha) &= \sigma(\lambda) \cdot \rho(\alpha), \\ \sigma(\lambda + \mu) &= \sigma(\lambda) + \sigma(\mu), & \sigma(\lambda\mu) &= \sigma(\lambda)\sigma(\mu), \end{aligned} \quad \forall \alpha, \beta \in V_1, \lambda, \mu \in \mathbb{F}_1$$

则 (ρ, σ) 称为从 $(V_1, \mathbb{F}_1, \oplus, *)$ 到 $(V_2, \mathbb{F}_2, +, \cdot)$ 的**同构映射**. 也就是说, 同构映射不仅需要保持向量之间的一一对应, 保持数之间的一一对应, 也需要保持运算之间的一一对应, 既保加法又保数乘.

例 8.5. 下面是一些线性空间之间的同构映射.

1. 作为 \mathbb{R} 上的线性空间, $z \mapsto (\operatorname{Re}(z), \operatorname{Im}(z))$ 是 $\mathbb{C} \rightarrow \mathbb{R}^2$ 的同构映射.
2. $A \mapsto Q(x) = x^T Ax$ 是“ \mathbb{F} 上的 n 阶对称方阵”到“ \mathbb{F} 上的 n 元二次型”的同构映射.
3. $\sum_{k=0}^{n-1} a_k x^k \mapsto (a_0, a_1, \dots, a_{n-1})$ 是 $\mathbb{F}_n[x] \rightarrow \mathbb{F}^n$ 的同构映射.
4. 设 $x_1, x_2, \dots, x_n \in \mathbb{F}$ 两两不同, 则 $f(x) \mapsto (f(x_1), f(x_2), \dots, f(x_n))$ 也是 $\mathbb{F}_n[x] \rightarrow \mathbb{F}^n$ 的同构映射.

习题

以下涉及到的加法和数乘运算都是通常意义下的加法和数乘运算.

1. 证明定理 8.1.

2. 判断下列集合是否 \mathbb{R} 上的线性空间, 并证明你的结论.

- (1) 满足 $f(1) = 0$ 的 \mathbb{R} 上连续实函数 $f(x)$ 的全体.
- (2) 满足 $f(0) = 1$ 的 \mathbb{R} 上连续实函数 $f(x)$ 的全体.
- (3) \mathbb{R} 上有界实函数 $f(x)$ 的全体.
- (4) 满足 $\max_{x \in \mathbb{R}} |f(x)| < 1$ 的 \mathbb{R} 上实函数 $f(x)$ 的全体.
- (5) 对所有 $|x| < 1$ 都收敛的实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体.
- (6) 收敛半径为 1 的实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体.

3. 作为 \mathbb{F} 上的线性空间, 下列 U 是否 V 的子空间? 请简要说明理由.

- (1) $U = \mathbb{F}^2, V = \mathbb{F}^3$.
- (2) $U = \mathbb{R}[x], V = \mathbb{R}[[x]], \mathbb{F} = \mathbb{R}$.
- (3) $U = \mathcal{C}[0, 1], V = \mathcal{C}(0, 1), \mathbb{F} = \mathbb{R}$.
- (4) $U = \{A \in V \mid A^H = A\}, V = \mathbb{C}^{n \times n}, \mathbb{F} = \mathbb{R}$.
- (5) $U = \{A \in V \mid A^H = A\}, V = \mathbb{C}^{n \times n}, \mathbb{F} = \mathbb{C}$.
- (6) $U = \{A \in V \mid AA^H = A^H A\}, V = \mathbb{C}^{n \times n}, \mathbb{F} = \mathbb{C}$.

4. 证明: \mathbb{F} 上的线性空间 $\mathbb{F}^{m \times n}$ 与 \mathbb{F}^{mn} 同构.

5. 证明: 当 $m \neq n$ 时, \mathbb{F} 上的线性空间 \mathbb{F}^m 与 \mathbb{F}^n 不同构.

6. 证明: $\mathbb{F}[x]$ 的子集 V_1, V_2 在多项式运算下都是 $\mathbb{F}[x]$ 的子空间, 并且都与 $\mathbb{F}[x]$ 同构, 其中

$$V_1 = \{p \in \mathbb{F}[x] \mid p(-x) = p(x)\}, \quad V_2 = \{p \in \mathbb{F}[x] \mid p'(1) = 0\}.$$

§8.2 线性相关

设 V 是 \mathbb{F} 上的线性空间. V 中最基本的运算是加法和数乘运算. 多个加法和数乘运算的混合运算称为线性组合运算. 严格地说,

定义 8.5. 设 $\alpha_1, \alpha_2, \dots, \alpha_k \in V$, $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}$. 向量

$$\beta = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k$$

称为向量组 $\alpha_1, \alpha_2, \dots, \alpha_k$ 的一个线性组合, 或者称 β 可以被 $\alpha_1, \alpha_2, \dots, \alpha_k$ 线性表出, $\lambda_1, \lambda_2, \dots, \lambda_k$ 称为线性组合的系数. $\alpha_1, \alpha_2, \dots, \alpha_k$ 的所有线性组合的集合

$$U = \{\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k \mid \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}\}$$

构成 V 的子空间, 称为 $\alpha_1, \alpha_2, \dots, \alpha_k$ 生成的子空间, 记作 $U = \text{Span}(\alpha_1, \alpha_2, \dots, \alpha_k)$.

若存在不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_k$ 使得 $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k = \mathbf{0}$, 则称 $\alpha_1, \alpha_2, \dots, \alpha_k$ 是线性相关的, 否则称 $\alpha_1, \alpha_2, \dots, \alpha_k$ 是线性无关的. 也就是说, “ $\alpha_1, \alpha_2, \dots, \alpha_k$ 是线性无关的” 当且仅当 “以 $\lambda_1, \lambda_2, \dots, \lambda_k$ 为变元的方程 $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k = \mathbf{0}$ 只有解 $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ ”.

线性组合/线性相关/线性无关的概念可以推广到 V 的任意子集 S . S 中任意有限多个向量的线性组合 β 称为 S 一个的线性组合, 或者称 β 可以被 S 线性表出. S 的所有线性组合的集合

$$U = \{\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k \mid \lambda_1, \dots, \lambda_k \in \mathbb{F}, \alpha_1, \dots, \alpha_k \in S, k \in \mathbb{N}^*\}$$

构成 V 的子空间, 称为 S 生成的子空间, 记作 $U = \text{Span}(S)$.

若 S 中存在有限多个向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是线性相关的, 则 S 称为线性相关的. 否则, 若 S 中的任意有限多个向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 都是线性无关的, 则 S 称为线性无关的. 特别, 规定零向量 $\mathbf{0}$ 是空集 \emptyset 的线性组合, $\text{Span}(\emptyset) = \{\mathbf{0}\}$, \emptyset 是线性无关的.

例 8.6. 仅有一个向量的向量组 $\{\alpha\}$ 是线性相关的当且仅当 $\alpha = \mathbf{0}$.

例 8.7. 线性空间 $\mathbb{R}[x]$ 中的向量组 $f_1 = x(x+1)$, $f_2 = x(x-1)$, $f_3 = x^2 - 1$ 是线性无关的.

证明. 设实数 $\lambda_1, \lambda_2, \lambda_3$ 使得 $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 = 0$ 是零函数. 在等式中分别令 $x = 1, -1, 0$, 可得 $\lambda_1 = \lambda_2 = \lambda_3 = 0$. 故 f_1, f_2, f_3 是线性无关的. \square

例 8.8. 对于任意 $A \in \mathbb{R}^{n \times n}$, 线性空间 $\mathbb{R}^{n \times n}$ 中的向量组 I, A, A^2, \dots, A^n 是线性相关的. 另外, $I, A, A^2, \dots, A^{n-1}$ 是线性无关的当且仅当 $d_A = \varphi_A$.

证明. 设 $\varphi_A(x) = \sum_{i=0}^n a_i x^i$, $a_n = 1$. 根据 Cayley-Hamilton 定理, $\sum_{i=0}^n a_i A^i = O$. 故 I, A, A^2, \dots, A^n 是线性相关的. 另外, $I, A, A^2, \dots, A^{n-1}$ 是线性相关的 \Leftrightarrow 存在非零多项式 $f \in \mathbb{R}[x]$ 满足 $f(A) = O$ 并且 $0 \leq \deg(f) \leq n-1$. 因此, $I, A, A^2, \dots, A^{n-1}$ 是线性无关的 $\Leftrightarrow d_A = \varphi_A$. \square

定理 8.2. 设 $S \subset T \subset V$. 由定义 8.5 可得下列结论.

1. 若 S 包含零向量或者两个成比例的向量, 则 S 是线性相关的.
2. 若 S 是线性相关的, 则 T 是线性相关的.

3. 若 T 是线性无关的, 则 S 是线性无关的.

定理 8.3. 设 $A = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n) \in \mathbb{F}^{m \times n}$, 则 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的 $\Leftrightarrow \text{rank}(A) = n$.

证明. 根据定义 8.5, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的 \Leftrightarrow 线性方程组 $Ax = \mathbf{0}$ 有唯一解 $x = \mathbf{0}$. 另由 A 的相抵标准形, $Ax = \mathbf{0}$ 有唯一解 $x = \mathbf{0} \Leftrightarrow \text{rank}(A) = n$. \square

定理 8.4. 设 $S \subset V$. S 是线性相关的 \Leftrightarrow 存在 $\alpha \in S$ 是 $S \setminus \{\alpha\}$ 的线性组合.

证明. (\Rightarrow) 根据定义 8.5, 存在有限多个不全是 0 的数 $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{F}$ 和两两不同的向量 $\alpha_1, \alpha_2, \cdots, \alpha_k \in S$, 使得 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = \mathbf{0}$. 不妨设 $\lambda_1 \neq 0$, 则 α_1 是 $\alpha_2, \cdots, \alpha_k$ 的线性组合. 再根据定义 8.5, α_1 是 $S \setminus \{\alpha_1\}$ 的线性组合.

(\Leftarrow) 设 $\alpha \in S$ 是 $S \setminus \{\alpha\}$ 的线性组合. 根据定义 8.5, 存在有限多个数 $\lambda_1, \cdots, \lambda_k \in \mathbb{F}$ 和两两不同的向量 $\alpha_1, \cdots, \alpha_k \in S \setminus \{\alpha\}$, 使得 $\alpha = \lambda_1\alpha_1 + \cdots + \lambda_k\alpha_k$. 因此, $\alpha, \alpha_1, \cdots, \alpha_k$ 是线性相关的. 再根据定义 8.5, S 是线性相关的. \square

定理 8.5. 设 β 是 $\alpha_1, \alpha_2, \cdots, \alpha_k \in V$ 的线性组合. $\alpha_1, \alpha_2, \cdots, \alpha_k$ 是线性无关的 \Leftrightarrow 存在唯一的 $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{F}$ 使得 $\beta = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k$.

证明. (\Rightarrow) 若 $\beta = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = \mu_1\alpha_1 + \mu_2\alpha_2 + \cdots + \mu_k\alpha_k$, 则

$$(\lambda_1 - \mu_1)\alpha_1 + (\lambda_2 - \mu_2)\alpha_2 + \cdots + (\lambda_k - \mu_k)\alpha_k = \mathbf{0} \quad \Rightarrow \quad \lambda_i = \mu_i, \quad \forall i.$$

(\Leftarrow) 若 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 是线性相关的, 则存在不全为 0 的数 $\mu_1, \mu_2, \cdots, \mu_k \in \mathbb{F}$ 使得 $\mu_1\alpha_1 + \mu_2\alpha_2 + \cdots + \mu_k\alpha_k = \mathbf{0}$. 从而有

$$\beta = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = (\lambda_1 + \mu_1)\alpha_1 + (\lambda_2 + \mu_2)\alpha_2 + \cdots + (\lambda_k + \mu_k)\alpha_k.$$

与题设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 的唯一性矛盾. \square

定义 8.6. 设 $S_1, S_2 \subset V$. 若 $\text{Span}(S_1) = \text{Span}(S_2)$, 则称 S_1 与 S_2 等价.

定理 8.6. 设 $S_1, S_2 \subset V$. 若 S_1 与 S_2 可以相互线性表出, 则 S_1 与 S_2 等价.

证明. 由 $S_1 \subset \text{Span}(S_2)$, 得 $\text{Span}(S_1) \subset \text{Span}(S_2)$. 由 $S_2 \subset \text{Span}(S_1)$, 得 $\text{Span}(S_2) \subset \text{Span}(S_1)$. 因此, $\text{Span}(S_1) = \text{Span}(S_2)$, 即 S_1 与 S_2 等价. \square

例 8.9. 在线性空间 $\mathcal{C}[0, 2\pi]$ 中, $S_1 = \{1, \cos x, \cos(2x), \cos(3x)\}$ 与 $S_2 = \{1, \cos x, (\cos x)^2, (\cos x)^3\}$ 等价.

证明. 一方面, $\cos(2x) = 2(\cos x)^2 - 1, \cos(3x) = 4(\cos x)^3 - 3\cos x$; 另一方面, $(\cos x)^2 = \frac{1}{2} + \frac{1}{2}\cos(2x), (\cos x)^3 = \frac{3}{4}\cos x + \frac{1}{4}\cos(3x)$. 根据定理 8.6, S_1 与 S_2 等价. \square

定理 8.7 (Steinitz^[1] Exchange Lemma). 设 V 中的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是线性无关的, 并且可以被向量组 $\beta_1, \beta_2, \cdots, \beta_n$ 线性表出, 则有 (1) $m \leq n$; (2) 可以用 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 替换 $\beta_1, \beta_2, \cdots, \beta_n$ 中的 m 个向量, 不妨设为 $\beta_1, \beta_2, \cdots, \beta_m$, 使得 $\alpha_1, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_n$ 与 $\beta_1, \beta_2, \cdots, \beta_n$ 等价.

^[1]Ernst Steinitz, 1871–1928, 德国数学家.

证明. (1) 设 $\alpha_i = \sum_{j=1}^n a_{ij}\beta_j$, $A = (a_{ij}) \in \mathbb{F}^{m \times n}$. 若 $m > n$, 则线性方程组 $xA = \mathbf{0}$ 有非零解 $x = (x_1, x_2, \dots, x_m) \neq \mathbf{0}$. 从而, $\sum_{i=1}^m x_i \alpha_i = \mathbf{0}$, 与 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是线性无关的矛盾. 因此, $m \leq n$.

(2) 由于 $\alpha_1 \neq \mathbf{0}$, 故存在 $a_{1j} \neq 0$. 不妨设 $a_{11} \neq 0$, 则 β_1 可以由 $\alpha_1, \beta_2, \dots, \beta_n$ 线性表出. 因此, $\alpha_1, \beta_2, \dots, \beta_n$ 与 $\beta_1, \beta_2, \dots, \beta_n$ 等价.

假设已知 $\alpha_1, \dots, \alpha_{k-1}, \beta_k, \dots, \beta_n$ 与 $\beta_1, \beta_2, \dots, \beta_n$ 等价. 下面证明可以用 α_k 替换 β_k, \dots, β_n 中的某个向量, 不妨设为 β_k , 使得 $\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n$ 与 $\beta_1, \beta_2, \dots, \beta_n$ 等价.

设 $\alpha_k = \lambda_1 \alpha_1 + \dots + \lambda_{k-1} \alpha_{k-1} + \lambda_k \beta_k + \dots + \lambda_n \beta_n$. 由于 $\alpha_1, \dots, \alpha_k$ 线性无关, 必存在 $j \geq k$ 使得 $\lambda_j \neq 0$. 不妨设 $\lambda_k \neq 0$, 则 β_k 可以由 $\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n$ 线性表出. 因此, $\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n$ 与 $\alpha_1, \dots, \alpha_{k-1}, \beta_k, \dots, \beta_n$ 等价, 进而与 $\beta_1, \beta_2, \dots, \beta_n$ 等价. \square

定理 8.7 的结论可以推广到 V 中任意多个向量.

定理 8.8. 设 $S_1, S_2 \subset V$, S_1 是线性无关的, $S_1 \subset \text{Span}(S_2)$, 则存在单射 $f: S_1 \rightarrow S_2$ 使得 $(S_2 \setminus f(S_1)) \cup S_1$ 与 S_2 等价. 由此可得, $|S_1| \leq |S_2|$. [2]

证明. 在集合 $X = \{A \mid A \subset S_1 \text{ 并且存在单射 } f: A \rightarrow S_2 \text{ 使得 } (S_2 \setminus f(A)) \cup A \text{ 与 } S_2 \text{ 等价}\}$ 上定义偏序 $A \prec B \Leftrightarrow A \subset B$. 根据 Zorn 引理, 存在 X 关于偏序 \prec 的一个极大元素 T . 下证 $T = S_1$, 从而结论成立.

假设存在 $\alpha \in S_1 \setminus T$. 由 $\text{Span}(S_2) = \text{Span}((S_2 \setminus f(T)) \cup T)$, 存在 $\lambda_1, \dots, \lambda_t, \mu_1, \dots, \mu_s \in \mathbb{F}$ 和 $\alpha_1, \dots, \alpha_t \in T$, $\beta_1, \dots, \beta_s \in S_2 \setminus f(T)$, 使得 $\alpha = \lambda_1 \alpha_1 + \dots + \lambda_t \alpha_t + \mu_1 \beta_1 + \dots + \mu_s \beta_s$. 再由 S_1 是线性无关的, $(\mu_1, \dots, \mu_s) \neq \mathbf{0}$. 不妨设 $\mu_1 \neq 0$, 则 β_1 可以由 $\alpha_1, \dots, \alpha_t, \alpha$ 线性表出. 设 $T' = T \cup \{\alpha\}$. 把 f 扩充为 $T' \rightarrow S_2$ 的单射, 使得 $f(\alpha) = \beta_1$. $T' \cup (S_2 \setminus f(T'))$ 与 $T \cup (S_2 \setminus f(T))$ 等价, 进而与 S_2 等价. 这与 T 的极大性矛盾. \square

 **Zorn 引理**[3]: 在任何一个非空的偏序集中, 若任意全序子集都有上界, 则这个偏序集必然存在一个极大元素.

良序定理: 对于任意集合都存在一种排序方式, 使得它的所有子集都有极小元素.

选择公理[4]: 从任意一族非空集合 $\{S_i \mid i \in I\}$ 中可选出一族元素 $x_i \in S_i, \forall i \in I$.

Zorn 引理、良序定理、选择公理三者等价.

习题

1. 证明定理 8.2.

2. 判断 \mathbb{R} 上的线性空间 $\mathbb{R}[x]$ 的子集 S 是否线性相关, 并证明你的结论.

$$(1) S = \{1, x-1, (x-1)^2, \dots, (x-1)^k, \dots\}$$

$$(2) S = \{1, x, x(x+1), \dots, x(x+1) \cdots (x+k-1), \dots\}$$

$$(3) S = \{x^i(1-x)^j \mid 0 \leq i < j \leq n\}, \text{ 其中 } n \text{ 是给定的正整数.}$$

$$(4) S = \{(x-a_0)^n, (x-a_1)^n, \dots, (x-a_n)^n\}, \text{ 其中 } a_0, a_1, \dots, a_n \in \mathbb{R} \text{ 两两不同.}$$

[2] $|X|$ 表示集合 X 的元素个数或势. 通常记 $\aleph_0 = |\mathbb{N}|$, $\aleph_1 = |\mathbb{R}|$. 关于集合论的知识, 可见参考文献 [9] *Set Theory*.

[3] 波兰数学家 Kazimierz Kuratowski 于 1922 年首先证明了此结论, 德国数学家 Max Zorn 于 1935 年独立地证明了此结论.

[4] 德国数学家 Ernst Friedrich Ferdinand Zermelo 于 1904 年最早提出了选择公理.

3. 判断 \mathbb{R} 上的线性空间 $\mathcal{C}[0, 2\pi]$ 的子集 S 是否线性相关, 并证明你的结论.
- (1) $S = \{\sin(nx) \mid n \in \mathbb{N}^*\}$ (2) $S = \{\cos(nx) \mid n \in \mathbb{N}\}$
 (3) $S = \{\sin(mx) \cos(nx) \mid m, n \in \mathbb{N}^*\}$ (4) $S = \{(\sin x)^m (\cos x)^n \mid m, n \in \mathbb{N}^*\}$
4. 证明: 坐标平面中的三点 $A(x_1, y_1)$, $B(x_2, y_2)$, $C(x_3, y_3)$ 共线的充分必要条件是线性空间 \mathbb{R}^3 中的向量组 $\alpha_1 = (x_1, y_1, 1)$, $\alpha_2 = (x_2, y_2, 1)$, $\alpha_3 = (x_3, y_3, 1)$ 线性相关.
5. 证明: 若 $k > n$, 则线性空间 $\mathbb{F}_n[x]$ 中的向量组 $f_1(x), f_2(x), \dots, f_k(x)$ 是线性相关的.
6. 证明: 若 $k > mn$, 则线性空间 $\mathbb{F}^{m \times n}$ 中的向量组 A_1, A_2, \dots, A_k 是线性相关的.
7. 设线性空间 V 中的向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 满足每个 α_k 都不是 $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$ 的线性组合, $k = 1, 2, \dots, n$. 证明: $\alpha_1, \alpha_2, \dots, \alpha_n$ 是线性无关的.
8. 设线性空间 V 的子集 $S_1 = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $S_2 = \{\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \dots, \alpha_{n-1} + \alpha_n\}$. 证明或否定:
- (1) 若 S_1 线性相关, 则 S_2 线性相关. (2) 若 S_1 线性无关, 则 S_2 线性无关.
9. 设线性空间 V 的子集 $S_1 = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $S_2 = \{\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \dots, \alpha_{n-1} + \alpha_n, \alpha_n + \alpha_1\}$. 证明或否定:
- (1) 若 S_1 线性相关, 则 S_2 线性相关. (2) 若 S_1 线性无关, 则 S_2 线性无关.
10. 设 V 为闭区间 $[0, 1]$ 上全体实函数构成的实向量空间. 证明: V 中向量组 f_1, \dots, f_n 线性无关当且仅当存在 $x_1, \dots, x_n \in [0, 1]$ 使得 n 阶实数方阵 $(f_i(x_j))$ 可逆.

§8.3 向量组的秩

定义 8.7. 设 V 是线性空间, $T \subset S \subset V$. 若 T 是线性无关的, 并且对于任意 $\alpha \in S \setminus T$, $T \cup \{\alpha\}$ 是线性相关的, 则 T 称为 S 的一个极大线性无关组.

 定义 8.7 并没有说明 S 的极大线性无关组是否存在. 当 S 是有限集合时, S 的极大线性无关组显然存在. 当 S 是无限集合时, 由 Zorn 引理可得 S 的极大线性无关组一定存在.

例 8.10. 求 \mathbb{R}^4 中的向量组 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的所有极大线性无关组, 其中

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} -1 \\ 3 \\ 3 \\ 3 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 1 \\ 1 \\ -4 \\ 0 \end{pmatrix}.$$

解答. 对矩阵 $A = (\alpha_1 \alpha_2 \alpha_3 \alpha_4)$ 作初等行变换, 化为阶梯形矩阵 $B = (\beta_1 \beta_2 \beta_3 \beta_4)$, 则 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的极大线性无关组与 $\beta_1, \beta_2, \beta_3, \beta_4$ 的极大线性无关组一一对应.

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & -1 & 3 & 1 \\ -2 & 2 & 3 & -4 \\ 1 & 1 & 3 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & -1 & 3 & 1 \\ 0 & 2 & 1 & -2 \\ 0 & 1 & 4 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 4 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = B$$

$\beta_1, \beta_2, \beta_3, \beta_4$ 的所有极大线性无关组为 $\{\beta_1, \beta_2, \beta_3\}$, $\{\beta_1, \beta_3, \beta_4\}$, $\{\beta_2, \beta_3, \beta_4\}$. 因此, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的所有极大线性无关组为 $\{\alpha_1, \alpha_2, \alpha_3\}$, $\{\alpha_1, \alpha_3, \alpha_4\}$, $\{\alpha_2, \alpha_3, \alpha_4\}$. \square

定理 8.9. 设 $T \subset S \subset V$. T 是 S 的极大线性无关组 $\Leftrightarrow T$ 是线性无关的并且 $S \subset \text{Span}(T)$.

证明. (\Rightarrow) 任取 $\alpha \in S \setminus T$. 由于 $T \cup \{\alpha\}$ 是线性相关的, 存在 T 的有限子集 $\{\beta_1, \dots, \beta_k\}$ 和不全是 0 的数 $\lambda, \mu_1, \dots, \mu_k \in \mathbb{F}$ 使得 $\lambda\alpha + \mu_1\beta_1 + \dots + \mu_k\beta_k = \mathbf{0}$. 再由 T 是线性无关的, 得 β_1, \dots, β_k 是线性无关的, 故 $\lambda \neq 0$, $\alpha = -\frac{1}{\lambda}(\mu_1\beta_1 + \dots + \mu_k\beta_k) \in \text{Span}(T)$. 因此, $S \subset \text{Span}(T)$.

(\Leftarrow) 根据定义 8.7, T 是 S 的极大线性无关组. \square

定理 8.10. 设 $T \subset S \subset V$. T 的任意一个极大线性无关组可以扩充为 S 的极大线性无关组.

证明. 设 T_1 是 T 的任意一个极大线性无关组, 则 T_1 是 S 的线性无关组. 在集合

$$X = \{A \mid T_1 \subset A \subset S \text{ 并且 } A \text{ 是线性无关的}\}$$

上定义偏序 $A \prec B \Leftrightarrow A \subset B$. 根据 Zorn 引理, 存在 X 关于偏序 \prec 的一个极大元素 S_1 , 则 S_1 是 S 的极大线性无关组. \square

定理 8.11. 设 $S \subset V$. S 的任意两个极大线性无关组的元素个数相同.

证明. 定理 8.8 和定理 8.9 的推论. \square

定义 8.8. 设 $S \subset V$. S 的任意一个极大线性无关组的元素个数称为 S 的秩, 记作 $\text{rank}(S)$.

例 8.11. 设 $A \in \mathbb{F}^{m \times n}$ 的行向量组为 $\alpha_1, \alpha_2, \dots, \alpha_m$, 列向量组为 $\beta_1, \beta_2, \dots, \beta_n$, 则有

$$\text{rank}(\alpha_1, \alpha_2, \dots, \alpha_m) = \text{rank}(\beta_1, \beta_2, \dots, \beta_n) = \text{rank}(A).$$

证明. 设 $r = \text{rank}(A)$ 并且 $A \begin{bmatrix} i_1 & i_2 & \dots & i_r \\ j_1 & j_2 & \dots & j_r \end{bmatrix}$ 可逆, 则线性方程组 $\mathbf{x}A \begin{bmatrix} i_1 & i_2 & \dots & i_r \\ 1 & 2 & \dots & n \end{bmatrix} = \mathbf{0}$ 只有解 $\mathbf{x} = \mathbf{0}$, 得 $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$ 线性无关. 再由定理 8.10, 得 $r \leq \text{rank}(S)$. 另一方面, 设 $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_k}$ 是 $\alpha_1, \alpha_2, \dots, \alpha_m$ 的一个极大线性无关组, 则线性方程组 $x_1\alpha_{i_1} + x_2\alpha_{i_2} + \dots + x_k\alpha_{i_k} = \mathbf{0}$ 只有解 $x_1 = x_2 = \dots = x_k = 0$. 故 $k = \text{rank}(A \begin{bmatrix} i_1 & i_2 & \dots & i_k \\ 1 & 2 & \dots & n \end{bmatrix}) \leq \text{rank}(A)$. 综上, $\text{rank}(\alpha_1, \alpha_2, \dots, \alpha_m) = \text{rank}(A)$.

同理, $\text{rank}(\beta_1, \beta_2, \dots, \beta_n) = \text{rank}(A^T) = \text{rank}(A)$. \square

例 8.11 利用线性空间的语言, 从几何的角度阐述了矩阵的秩的含义. 从而为一些关于矩阵的秩的结论赋予了几何的解释.

例 8.12. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{m \times p}$. 证明: $\text{rank} \begin{pmatrix} A & B \end{pmatrix} \leq \text{rank}(A) + \text{rank}(B)$.

证明. 设 T_1 是 A 的列向量组的极大线性无关组, T_2 是 B 的列向量组的极大线性无关组, 则 $T_1 \cup T_2$ 的极大线性无关组是 (A, B) 的列向量组的极大线性无关组. 因此,

$$\text{rank} \begin{pmatrix} A & B \end{pmatrix} \leq |T_1 \cup T_2| \leq |T_1| + |T_2| = \text{rank}(A) + \text{rank}(B). \quad \square$$

例 8.13. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$. 证明: $\text{rank}(AB) \leq \text{rank}(A)$.

证明. 设 T_1 是 A 的列向量组的极大线性无关组, T_2 是 AB 的列向量组的极大线性无关组. T_2 可以由 T_1 线性表出. 根据定理 8.7, $\text{rank}(AB) = |T_2| \leq |T_1| = \text{rank}(A)$. \square

习题

1. 求 \mathbb{R}^4 中下列向量组的所有极大线性无关组.

$$(1) \alpha_1 = (-1, -2, 2, 0), \alpha_2 = (-1, -1, 3, 1), \alpha_3 = (-1, 0, 0, -2), \alpha_4 = (-1, -2, 2, 0)$$

$$(2) \alpha_1 = (-1, 1, 1, 0), \alpha_2 = (1, 1, 1, 1), \alpha_3 = (0, -1, 0, -1), \alpha_4 = (0, 0, 1, -1), \alpha_5 = (1, -1, -1, 1)$$

2. 求 $\mathbb{R}^{2 \times 2}$ 中下列向量组的所有极大线性无关组.

$$(1) A_1 = \begin{pmatrix} 3 & -1 \\ -2 & 2 \end{pmatrix}, A_2 = \begin{pmatrix} 2 & -2 \\ -2 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}$$

$$(2) A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A_5 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

3. 求 $\mathbb{R}[x]$ 中下列向量组的所有极大线性无关组.

$$(1) p_1 = x(x-1), p_2 = (x-1)(x-2), p_3 = (x-2)(x-3), p_4 = (x-3)(x-4)$$

$$(2) p_1 = (x-1)^3, p_2 = (x-2)^3, p_3 = (x-3)^3, p_4 = (x-4)^3, p_5 = (x-5)^3$$

4. 设 $S_1, S_2 \subset V$. 证明:

$$(1) \text{若 } \text{Span}(S_1) = \text{Span}(S_2), \text{ 则 } \text{rank}(S_1) = \text{rank}(S_2).$$

$$(2) \text{若 } S_1 \subset S_2 \text{ 且 } \text{rank}(S_1) = \text{rank}(S_2) < \infty, \text{ 则 } \text{Span}(S_1) = \text{Span}(S_2).$$

$$(3) \text{举例: } S_1 \subset S_2 \text{ 且 } \text{rank}(S_1) = \text{rank}(S_2), \text{ 但 } \text{Span}(S_1) \neq \text{Span}(S_2).$$

5. 设 $\alpha_1, \alpha_2, \dots, \alpha_m \in V$, $n < m$. 证明:

$$\text{rank}(\alpha_1, \alpha_2, \dots, \alpha_m) - \text{rank}(\alpha_1, \alpha_2, \dots, \alpha_n) \leq m - n.$$

利用线性空间的语言证明下列矩阵秩不等式问题.

6. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$, $\text{rank}(AB) = \text{rank}(A)$. 证明:

(1) 存在 $X \in \mathbb{F}^{p \times n}$ 使得 $ABX = A$.

(2) 对于任意 $Y \in \mathbb{F}^{q \times m}$, 都有 $\text{rank}(YAB) = \text{rank}(YA)$.

7. 设 $A \in \mathbb{F}^{n \times n}$, $\text{rank}(A^k) = \text{rank}(A^{k-1})$. 证明: $\text{rank}(A^{k+1}) = \text{rank}(A^k)$.

8. 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{m \times p}$, $\text{rank}(A, B) = \text{rank}(A)$. 证明: 存在 $X \in \mathbb{F}^{n \times p}$ 使得 $AX = B$.

9. 设 $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbb{F}^{m \times n}$, $A \in \mathbb{F}^{p \times q}$. 证明: $\text{rank}(M) \geq \text{rank}(A)$, 等号成立的充分必要条件是存在 $X \in \mathbb{F}^{q \times (n-q)}$ 和 $Y \in \mathbb{F}^{(m-p) \times p}$ 使得 $B = AX$, $C = YA$, $D = YAX$.

10. 设 $A, B \in \mathbb{F}^{m \times n}$. 证明: $|\text{rank}(A) - \text{rank}(B)| \leq \text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$.

§8.4 基与坐标

定义 8.9. 线性空间 V 的任意一个极大线性无关组 S 称为 V 的一个基, S 的元素个数 $|S|$ 称为 V 的维数, 记作 $\dim V$.

例 8.14.

1. 空集 \emptyset 是零空间的基, 零空间的维数是 0.
2. 设 U 是 V 的子空间, 则 $\dim U \leq \dim V$.
3. 标准单位向量组 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 构成 \mathbb{F}^n 的标准基, $\dim \mathbb{F}^n = n$.
4. 基础矩阵组 $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ 构成 $\mathbb{F}^{m \times n}$ 的标准基, $\dim \mathbb{F}^{m \times n} = mn$.
5. 单项式 $\{x^k \mid k \in \mathbb{N}\}$ 是 $\mathbb{F}[x]$ 的基, $\dim \mathbb{F}[x] = \aleph_0$.
6. $U = \{f \in \mathbb{F}[x] \mid f(0) = 0\}$ 是 $\mathbb{F}[x]$ 的子空间, $\dim U = \aleph_0$.

例 8.15. 作为 \mathbb{C} 上的线性空间, \mathbb{C} 是 1 维的, 任意一个非零复数 a 都构成 \mathbb{C} 的基. 作为 \mathbb{R} 上的线性空间, \mathbb{C} 是 2 维的, 任意两个满足 $\frac{a}{b} \notin \mathbb{R}$ 的非零复数 a, b 都构成 \mathbb{C} 的基.

定义 8.10. 设 V 是 n 维线性空间. 取定 V 的基 $\alpha_1, \alpha_2, \dots, \alpha_n$. 对于任意 $\beta \in V$, 存在唯一的 $x = (x_1, x_2, \dots, x_n) \in \mathbb{F}^n$ 使得 $\beta = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n$. x 称为 β 在 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标. 定义坐标映射 $\rho: V \rightarrow \mathbb{F}^n, \beta \mapsto x$. 容易验证, ρ 是同构映射.

类似地, 当 V 是无穷维线性空间时, 设 $S = \{\alpha_i \mid i \in I\}$ 是 V 的基, I 是指标集合. 对于任意 $\beta \in V$, 存在唯一的数列 $X = (x_i)_{i \in I}$ 使得 $\beta = \sum_{i \in I} x_i \alpha_i$, 其中 $x_i \in \mathbb{F}$ 且只有有限多个 $x_i \neq 0$. 数列 X 称为 β 在 S 下的坐标.

定理 8.12. 设 U, V 都是 \mathbb{F} 上的线性空间, 则 U 与 V 同构 $\Leftrightarrow \dim U = \dim V$.

证明. (\Rightarrow) 设 $\rho: U \rightarrow V$ 是同构映射, $S = \{\alpha_i \mid i \in I\}$ 是 U 的基, $T = \{\rho(\alpha_i) \mid i \in I\}$. 一方面, 对任意 $v \in V$, 设 $\rho^{-1}(v) = \sum_{i \in I} x_i \alpha_i$, 其中 $x_i \in \mathbb{F}$ 且只有有限多个 $x_i \neq 0$, 则 $v = \sum_{i \in I} x_i \rho(\alpha_i) \in \text{Span}(T)$. 故 $\text{Span}(T) = V$. 另一方面, 若 $\sum_{i \in I} x_i \rho(\alpha_i) = \mathbf{0}$, 其中 $x_i \in \mathbb{F}$ 且只有有限多个 $x_i \neq 0$, 则 $\sum_{i \in I} x_i \alpha_i = \mathbf{0}$, 得 $x_i = 0, \forall i \in I$. 故 T 线性无关. 因此, T 是 V 的基, $\dim U = |S| = |T| = \dim V$.

(\Leftarrow) 设 S, T 分别是 U, V 的基, 由 $|S| = |T|$, 存在一一映射 $\rho: S \rightarrow T$. 设 $S = \{\alpha_i \mid i \in I\}$, $T = \{\beta_i \mid i \in I\}$. 把 ρ 扩充为 $U \rightarrow V$ 的映射 $\sum_{i \in I} x_i \alpha_i \mapsto \sum_{i \in I} x_i \beta_i$, 其中 $x_i \in \mathbb{F}$ 且只有有限多个 $x_i \neq 0$. 容易验证, ρ 是同构映射. \square

例 8.16. 设 \mathbb{R} 上的线性空间 V 由所有 2 阶对称实方阵构成. $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 是 V 的基. 任意 $\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in V$ 在 A_1, A_2, A_3 下的坐标是 $\left(\frac{a+c}{2}, b, \frac{a-c}{2}\right)$.

例 8.17. 设 $a \in \mathbb{R}, f(x) \in \mathbb{R}_n[x]$. 根据 Taylor 展开, $f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k$. 故 $f(x)$ 在 $\mathbb{R}_n[x]$ 的基 $\{(x-a)^k \mid k=0, 1, 2, \dots, n-1\}$ 下的坐标是 $\left(f(a), f'(a), \frac{f''(a)}{2!}, \dots, \frac{f^{(n-1)}(a)}{(n-1)!}\right)$.

一个线性空间的基通常是不唯一的, 一个向量在不同基下的坐标通常是不同的. 下面我们讨论线性空间 V 的不同基之间的联系, 以及同一个向量在不同基下的坐标之间的联系.

定义 8.11. 设 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 和 $T = \{\beta_1, \beta_2, \dots, \beta_n\}$ 都是 \mathbb{F} 上的 n 维线性空间 V 的基,

$$\beta_j = \sum_{i=1}^n p_{ij} \alpha_i, \quad j = 1, 2, \dots, n$$

或者形式上记作

$$\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \quad (8.1)$$

则 $P = (p_{ij}) \in \mathbb{F}^{n \times n}$ 称为从 S 到 T 的过渡矩阵. 也就是说, 过渡矩阵由新基向量在原基向量下的坐标构成. (8.1) 式称为从 S 到 T 的基变换公式.

定理 8.13. 设 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 和 $T = \{\beta_1, \beta_2, \dots, \beta_n\}$ 都是线性空间 V 的基.

1. 设 P 是从 S 到 T 的过渡矩阵, 则 P^{-1} 是从 T 到 S 的过渡矩阵.

2. 设向量 $\alpha \in V$ 在 S 和 T 下的坐标分别是列向量 x 和 y , 则有 $y = P^{-1}x$.

证明. 设 $Q = (q_{ij})$ 是从 T 到 S 的过渡矩阵, 即 $\alpha_j = \sum_{i=1}^n q_{ij} \beta_i, \forall j$.

1. 由 $\alpha_j = \sum_{i=1}^n q_{ij} \left(\sum_{k=1}^n p_{ki} \alpha_k \right) = \sum_{k=1}^n \left(\sum_{i=1}^n p_{ki} q_{ij} \right) \alpha_k$, 得 $PQ = I_n, Q = P^{-1}$.

2. 由 $\alpha = \sum_{j=1}^n x_j \alpha_j = \sum_{j=1}^n x_j \left(\sum_{i=1}^n q_{ij} \beta_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n q_{ij} x_j \right) \beta_i = \sum_{i=1}^n y_i \beta_i$, 得 $y = Qx = P^{-1}x$. \square

定义 8.12. 一个向量在不同基下的坐标之间的转换关系

$$x = Py \quad \text{和} \quad y = P^{-1}x \quad (8.2)$$

称为坐标变换公式.

习题

1. 求下列线性空间 V 的基和维数, 并求 V 中每个向量在这个基下的坐标.

(1) \mathbb{F} 上线性空间 $V = \{(a_1, a_2, \dots, a_n) \in \mathbb{F}^n \mid a_1 + a_2 + \dots + a_n = 0\}$

(2) \mathbb{F} 上线性空间 $V = \{A \in \mathbb{F}^{n \times n} \mid A \text{ 是下三角方阵}\}$

(3) \mathbb{F} 上线性空间 $V = \{A \in \mathbb{F}^{n \times n} \mid \text{tr}(A) = 0\}$

(4) \mathbb{F} 上线性空间 $V = \{f \in \mathbb{F}_n[x] \mid f'(1) = 0\}$

(5) \mathbb{C} 上线性空间 $V = \{A \in \mathbb{C}^{n \times n} \mid A^T = -A\}$

(6) \mathbb{R} 上线性空间 $V = \{A \in \mathbb{C}^{n \times n} \mid A^T = -A\}$

(7) \mathbb{C} 上线性空间 $V = \{f \in \mathbb{C}_n[x] \mid f(i) = 0\}$

(8) \mathbb{R} 上线性空间 $V = \{f \in \mathbb{C}_n[x] \mid f(i) = 0\}$

2. 设 $\{\alpha_i\}$ 和 $\{\beta_i\}$ 都是 \mathbb{R} 上的线性空间 V 的基, 求从 $\{\alpha_i\}$ 到 $\{\beta_i\}$ 的过渡矩阵.
- (1) $\alpha_1 = (1, -1, 1), \alpha_2 = (-1, 0, 1), \alpha_3 = (-1, 1, 0), \beta_1 = (0, 1, 2), \beta_2 = (2, 1, 0), \beta_3 = (1, 2, 1)$
- (2) $\alpha_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \beta_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \beta_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \beta_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
- (3) $\alpha_1 = x, \alpha_2 = x(x-1), \alpha_3 = x(x-1)(x-2), \beta_1 = x, \beta_2 = x^2, \beta_3 = x^3$
- (4) $\alpha_1 = \cos \frac{\pi}{3}, \alpha_2 = \cos(x + \frac{\pi}{3}), \alpha_3 = \sin(x + \frac{\pi}{3}), \beta_1 = \cos \frac{2\pi}{3}, \beta_2 = \cos(x + \frac{2\pi}{3}), \beta_3 = \sin(x + \frac{2\pi}{3})$
3. 设 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 n 维线性空间 V 的基, $P = (p_{ij}) \in \mathbb{F}^{n \times n}, \beta_j = \sum_{i=1}^n p_{ij} \alpha_i, j = 1, 2, \dots, n$.
证明: $\{\beta_1, \beta_2, \dots, \beta_n\}$ 是 V 的基 $\Leftrightarrow \det(P) \neq 0$.
4. 设 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}, \{\beta_1, \beta_2, \dots, \beta_n\}, \{\gamma_1, \gamma_2, \dots, \gamma_n\}$ 都是 n 维线性空间 V 的基, P 是从 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 到 $\{\beta_1, \beta_2, \dots, \beta_n\}$ 的过渡矩阵, Q 是从 $\{\beta_1, \beta_2, \dots, \beta_n\}$ 到 $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ 的过渡矩阵. 证明: PQ 是从 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 到 $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ 的过渡矩阵.
5. 在 \mathbb{R}^2 中以 $\alpha_1 = (\cos \theta, \sin \theta), \alpha_2 = (-\sin \theta, \cos \theta)$ 为基向量建立新的平面直角坐标系. 求平面曲线 $f(x, y) = 0$ 在新坐标系下的方程.
6. 在 \mathbb{R}^2 中建立新的平面直角坐标系, 化椭圆 $2x^2 + (x+y)^2 = 1$ 为标准方程.
7. 作为有理数域 \mathbb{Q} 上的线性空间, 实数域 \mathbb{R} 与复数域 \mathbb{C} 是否同构? 证明你的结论.
8. 作为 \mathbb{R} 上的线性空间, 实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体 $\mathbb{R}[[x]]$ 与实系数多项式的全体 $\mathbb{R}[x]$ 是否同构? 证明你的结论.
9. 设 $\mathbb{F}^{n \times n}$ 的子空间 V 满足 $\dim V \geq kn + 1$. 证明: 存在 $A \in V$ 使得 $\text{rank}(A) \geq k + 1$.
10. [5] 设 $\mathbb{F}^{n \times n}$ 的子空间 V 中任意两个方阵乘积可交换. 证明: $\dim V \leq \frac{n^2}{4} + 1$.

[5] Von Herrn J. Schur, Zur Theorie der vertauschbaren Matrizen, Journal für die Reine und Angewandte Mathematik, **130** (1905), 66-76.

§8.5 交空间与和空间

设 I 是指标集合, $\{V_i \mid i \in I\}$ 是线性空间 V 的一组子空间.

定理 8.14. $U = \bigcap_{i \in I} V_i$ 是 V 的子空间. U 称为 $\{V_i \mid i \in I\}$ 的交空间.

证明. 由于每个 V_i 都包含 $\mathbf{0}$, 故 $\mathbf{0} \in U$, U 非空. 对于任意 $\alpha, \beta \in U$, $\lambda \in \mathbb{F}$, 由 $\alpha, \beta \in V_i$ 可得 $\alpha + \beta, \lambda\alpha \in V_i$. 从而, $\alpha + \beta, \lambda\alpha \in U$. 因此, U 是 V 的子空间. \square

例 8.18. 设 V_i 是所有满足 $f^{(i)}(0) = 0$ 的实系数多项式 $f(x)$ 构成的 $\mathbb{R}[x]$ 的子空间, $i \in \mathbb{N}$. 对于任意正整数 n , $V_0 \cap V_1 \cap \cdots \cap V_n$ 与 $\mathbb{R}[x]$ 同构. 但是, $\bigcap_{i \in \mathbb{N}} V_i = \{\mathbf{0}\}$.

假设已知每个 V_i 的基, 如何计算 $\bigcap_{i \in I} V_i$ 的基呢? 我们先看一个例子.

例 8.19. 设 $V_1 = \text{Span}(\alpha_1, \alpha_2)$, $V_2 = \text{Span}(\beta_1, \beta_2)$ 都是 \mathbb{R}^3 的子空间, 求 $V_1 \cap V_2$ 的基, 其中

$$\alpha_1 = (1, 1, -1), \quad \alpha_2 = (2, -1, 2), \quad \beta_1 = (1, -3, 3), \quad \beta_2 = (0, -1, 3).$$

解答. $V_1 \cap V_2$ 中的向量 α 都可以表示为 $\alpha = x_1\alpha_1 + x_2\alpha_2 = y_1\beta_1 + y_2\beta_2$ 的形式, 其中 $x_1, x_2, y_1, y_2 \in \mathbb{R}$. 通过求解线性方程组

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 1 & -1 & 3 & 1 \\ -1 & 2 & -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ y_1 \end{pmatrix} = y_2 \begin{pmatrix} -\frac{11}{4} \\ 2 \\ \frac{5}{4} \end{pmatrix}$$

得 $V_1 \cap V_2$ 可由 $\alpha_3 = -11\alpha_1 + 8\alpha_2 = (5, -19, 27)$ 生成. 因此, $\{\alpha_3\}$ 是 $V_1 \cap V_2$ 的基. \square

下面给出计算 $V_1 \cap V_2$ 的一种方法. 设 V_1, V_2 是 \mathbb{F}^n 的任意两个子空间, V_1 可由 $\alpha_1, \cdots, \alpha_s$ 生成, V_2 可由 β_1, \cdots, β_t 生成, 则

$$V_1 \cap V_2 = \{x_1\alpha_1 + \cdots + x_s\alpha_s \mid x_1\alpha_1 + \cdots + x_s\alpha_s + x_{s+1}\beta_1 + \cdots + x_{s+t}\beta_t = \mathbf{0}\}.$$

对 $A = (\alpha_1 \cdots \alpha_s \beta_1 \cdots \beta_t)$ 作初等变换, 化为标准形 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 P, Q 是 \mathbb{F} 上的可逆方阵. 设 $Q = (q_1 \ q_2 \ \cdots \ q_{s+t})$, 则线性方程组 $Ax = \mathbf{0}$ 的通解

$$x = y_{r+1}q_{r+1} + y_{r+2}q_{r+2} + \cdots + y_{s+t}q_{s+t},$$

其中 $y_{r+1}, \cdots, y_{s+t} \in \mathbb{F}$. 从而,

$$\sum_{i=1}^s x_i\alpha_i = (\alpha_1 \cdots \alpha_s) \begin{pmatrix} I_s & O \end{pmatrix} x = (\alpha_1 \cdots \alpha_s) \begin{pmatrix} I_s & O \end{pmatrix} Q \begin{pmatrix} O \\ I_{s+t-r} \end{pmatrix} y = \sum_{j=r+1}^{s+t} y_j\gamma_j.$$

即 $V_1 \cap V_2 = \text{Span}(\gamma_{r+1}, \cdots, \gamma_{s+t})$, 其中

$$(\gamma_{r+1} \cdots \gamma_{s+t}) = (\alpha_1 \cdots \alpha_s) \tilde{Q}, \quad \tilde{Q} = \begin{pmatrix} I_s & O \end{pmatrix} Q \begin{pmatrix} O \\ I_{s+t-r} \end{pmatrix} = Q_{[r+1 \ \cdots \ s+t]}.$$

特别, 当 $\alpha_1, \dots, \alpha_s$ 和 β_1, \dots, β_t 分别是 V_1 和 V_2 的基时, 根据维数定理 8.15, $\dim V_1 \cap V_2 = s + t - r$, 从而 $\gamma_{r+1}, \dots, \gamma_{s+t}$ 是 $V_1 \cap V_2$ 的基.

一般说来, V 的子空间的并集 $S = \bigcup_{i \in I} V_i$ 不一定是 V 的子空间. 例如: $V_1 = \{(x, 0) \mid x \in \mathbb{F}\}$ 和 $V_2 = \{(0, y) \mid y \in \mathbb{F}\}$ 都是 \mathbb{F}^2 的子空间, 然而 $V_1 \cup V_2$ 在加法运算下不封闭, $V_1 \cup V_2$ 不是 \mathbb{F}^2 的子空间.

尽管 S 不一定是 V 的子空间, $\text{Span}(S)$ 总是 V 的子空间. $\text{Span}(S)$ 中的每个向量 α 都可以表示为 S 中有限多个向量的线性组合形式, 因此可以表示为 $\alpha = \sum_{i \in I} \alpha_i$ 的形式, 其中 $\alpha_i \in V_i$ 并且仅有限多个 $\alpha_i \neq 0$. 设 S_i 是 V_i 的基, 则 $\bigcup_{i \in I} S_i$ 的极大线性无关组是 $\sum_{i \in I} V_i$ 的基.

定义 8.13. $\text{Span}(\bigcup_{i \in I} V_i)$ 称为 $\{V_i \mid i \in I\}$ 的**和空间**, 记作 $\sum_{i \in I} V_i$.

例 8.20. 设 V_i 是 x^i 生成的 $\mathbb{R}[[x]]$ 的子空间, $i \in \mathbb{N}$, 则

$$\sum_{i \in \mathbb{N}} V_i = \text{Span}(1, x, x^2, \dots) = \mathbb{R}[x] \neq \mathbb{R}[[x]].$$

关于交空间与和空间的维数, 有以下常用结论.

定理 8.15 (维数定理). 设 V_1, V_2 是线性空间 V 的任意两个有限维子空间, 则有

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2.$$

证明. 设 $\alpha_1, \dots, \alpha_r$ 是 $V_1 \cap V_2$ 的基. 把它分别扩充为 V_1 的基 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s$ 和 V_2 的基 $\alpha_1, \dots, \alpha_r, \gamma_1, \dots, \gamma_t$. 下证 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s, \gamma_1, \dots, \gamma_t$ 线性无关. 若

$$a_1 \alpha_1 + \dots + a_r \alpha_r + b_1 \beta_1 + \dots + b_s \beta_s + c_1 \gamma_1 + \dots + c_t \gamma_t = 0,$$

其中 $a_1, \dots, a_r, b_1, \dots, b_s, c_1, \dots, c_t \in \mathbb{F}$, 则

$$a_1 \alpha_1 + \dots + a_r \alpha_r + b_1 \beta_1 + \dots + b_s \beta_s = -c_1 \gamma_1 - \dots - c_t \gamma_t \in V_1 \cap V_2.$$

从而, $b_1 = \dots = b_s = 0, c_1 = \dots = c_t = 0$, 得 $a_1 = \dots = a_r = 0$. 因此, $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s, \gamma_1, \dots, \gamma_t$ 是 $V_1 + V_2$ 的基. 由此可得, $\dim(V_1 + V_2) = r + s + t = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2$. \square

例 8.21. 设 \mathbb{F} 是无限域, V 是 \mathbb{F} 上的线性空间, V_1, V_2, \dots, V_k 都是 V 的真子空间. 证明:

$$V_1 \cup V_2 \cup \dots \cup V_k \neq V.$$

证明. 对 k 使用数学归纳法. 当 $k = 1$ 时, 结论显然成立. 下设 $k \geq 2$. 记 $W = V_1 \cup V_2 \cup \dots \cup V_{k-1}$.

- 当 $V_k \subset W$ 时, 根据归纳假设, $W \cup V_k = W \neq V$. 当 $W \subset V_k$ 时, $W \cup V_k = V_k \neq V$.
- 当 $V_k \not\subset W$ 且 $W \not\subset V_k$ 时, 存在 $\alpha \in V_k \setminus W, \beta \in W \setminus V_k$. 如果 $V_1 \cup V_2 \cup \dots \cup V_k = V$, 任取 $k+1$ 个两两不同的数 $\lambda_1, \lambda_2, \dots, \lambda_{k+1} \in \mathbb{F}$, 根据抽屉原理, 向量组 $\{\lambda_i \alpha + \beta \mid 1 \leq i \leq k+1\}$ 中必有两个向量同属于某个 V_j . 由此可得 $\alpha, \beta \in V_j$, 矛盾. \square

习题

1. 设 $V_1 = \text{Span}(\alpha_1, \alpha_2, \alpha_3)$, $V_2 = \text{Span}(\beta_1, \beta_2, \beta_3)$ 都是 \mathbb{R}^4 的子空间, 求 $V_1 \cap V_2$ 的基, 其中

(1) $\alpha_1 = (1, -2, 2, 0)$, $\alpha_2 = (2, 2, 1, 1)$, $\alpha_3 = (-1, 2, 0, 2)$,
 $\beta_1 = (2, -2, 1, 0)$, $\beta_2 = (1, 1, 2, 0)$, $\beta_3 = (-2, 2, -1, 2)$;

(2) $\alpha_1 = (2, -2, -1, 2)$, $\alpha_2 = (0, -1, -1, 2)$, $\alpha_3 = (-1, 0, -1, -1)$,
 $\beta_1 = (0, 1, 2, -2)$, $\beta_2 = (-2, -1, 2, 1)$, $\beta_3 = (-2, 2, -1, 0)$.
2. 设 $V_1 = \text{Span}(\alpha_1, \alpha_2, \alpha_3)$, $V_2 = \text{Span}(\beta_1, \beta_2, \beta_3)$ 都是 \mathbb{R}^5 的子空间, 求 $V_1 + V_2$ 的基, 其中

(1) $\alpha_1 = (1, 0, 2, -1, 0)$, $\alpha_2 = (0, 1, 0, 0, 2)$, $\alpha_3 = (-2, 1, -3, 2, 1)$,
 $\beta_1 = (0, 1, 1, 0, 1)$, $\beta_2 = (-1, 1, 0, 1, 0)$, $\beta_3 = (-1, 1, -2, 1, 2)$;

(2) $\alpha_1 = (2, 0, 1, 2, 1)$, $\alpha_2 = (2, -2, -3, 1, 2)$, $\alpha_3 = (2, -1, -1, 2, 2)$,
 $\beta_1 = (1, -1, 0, 1, 1)$, $\beta_2 = (2, 0, 1, 3, 2)$, $\beta_3 = (-1, 1, 3, 0, -1)$.
3. 分别求 $V_1 \cap V_2$ 和 $V_1 + V_2$ 的基, 其中 V_1, V_2 分别是

(1) $\{\cos^n x \mid n \in \mathbb{N}^*\}$ 和 $\{\sin^n x \mid n \in \mathbb{N}^*\}$ 生成的 $\mathcal{C}[0, 2\pi]$ 的子空间.

(2) $\{\cos(nx) \mid n \in \mathbb{N}^*\}$ 和 $\{\sin(nx) \mid n \in \mathbb{N}^*\}$ 生成的 $\mathcal{C}[0, 2\pi]$ 的子空间.
4. 设 $A = \begin{pmatrix} & I_{n-1} \\ 0 & \end{pmatrix}$, $V_1 = \{X \in \mathbb{R}^{n \times n} \mid AX = XA\}$, $V_2 = \{X \in \mathbb{R}^{n \times n} \mid A^T X = XA^T\}$.
 证明: V_1, V_2 都是 $\mathbb{R}^{n \times n}$ 的子空间. 并分别求 $V_1, V_2, V_1 \cap V_2, V_1 + V_2$ 的维数.
5. 设 V_1 是被 $x^2 + x$ 整除的实系数多项式的全体, V_2 是被 $x^2 - x$ 整除的实系数多项式的全体.
 证明: V_1, V_2 都是 $\mathbb{R}[x]$ 的子空间. 并分别求 $V_1 \cap V_2$ 和 $V_1 + V_2$ 的基.
6. 设 V_1, V_2 都是线性空间 V 的子空间. 证明: 若 $V_1 \cup V_2 = V_1 + V_2$, 则 $V_1 \subset V_2$ 或 $V_2 \subset V_1$.
7. 设 V_1, V_2, W 都是线性空间 V 的子空间. 证明:

(1) $(V_1 \cap W) + (V_2 \cap W)$ 是 $(V_1 + V_2) \cap W$ 的子空间.

(2) $(V_1 \cap W) + (V_2 \cap W) = (V_1 + V_2) \cap W$ 有可能不成立.

(3) $(V_1 \cap W) + (V_2 \cap W) = ((V_1 \cap W) + V_2) \cap W = (V_1 + (V_2 \cap W)) \cap W$.
8. 设 V_1, V_2, W 都是线性空间 V 的子空间. 证明:

(1) $(V_1 \cap V_2) + W$ 是 $(V_1 + W) \cap (V_2 + W)$ 的子空间.

(2) $(V_1 \cap V_2) + W = (V_1 + W) \cap (V_2 + W)$ 有可能不成立.

(3) $((V_1 + W) \cap V_2) + W = (V_1 \cap (V_2 + W)) + W = (V_1 + W) \cap (V_2 + W)$.
9. 设 V_1, V_2, V_3 都是线性空间 V 的有限维子空间. 证明或否定 $\dim(V_1 + V_2 + V_3) = \dim V_1 + \dim V_2 + \dim V_3 - \dim V_1 \cap V_2 - \dim V_1 \cap V_3 - \dim V_2 \cap V_3 + \dim V_1 \cap V_2 \cap V_3$.

§8.6 直和与补空间

设 I 是指标集合, $\{V_i | i \in I\}$ 是线性空间 V 的一组子空间.

定义 8.14. 若和空间 $W = \sum_{i \in I} V_i$ 中的每个向量 α 都可以唯一地表示为 $\alpha = \sum_{i \in I} \alpha_i$ 的形式, 其中每个 $\alpha_i \in V_i$, 并且只有有限多个 $\alpha_i \neq \mathbf{0}$, 则子空间运算 $\sum_{i \in I} V_i$ 称为直和, W 也称为 $\{V_i | i \in I\}$ 的直和, 记作 $W = \bigoplus_{i \in I} V_i$.

定理 8.16. 设 V_1, V_2 都是 V 的子空间. $V_1 + V_2$ 是直和的充分必要条件是 $V_1 \cap V_2 = \{\mathbf{0}\}$.

证明. (充分性) 设 $\alpha_1, \beta_1 \in V_1, \alpha_2, \beta_2 \in V_2$ 满足 $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$, 则 $\alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in V_1 \cap V_2$, 故 $\alpha_1 = \beta_1, \alpha_2 = \beta_2$. 因此, $V_1 + V_2 = V_1 \oplus V_2$.

(必要性) 若存在非零向量 $\alpha \in V_1 \cap V_2$, 则 $\mathbf{0} = \mathbf{0} + \mathbf{0} = \alpha + (-\alpha)$, 表示方式不唯一, 与 $V_1 + V_2$ 是直和矛盾. \square

例 8.22. 设 V 的子空间 U_1, U_2, W 满足 $U_1 \subset U_2$ 并且 $V = U_1 \oplus W = U_2 \oplus W$, 则 $U_1 = U_2$.

证明. 任取 $\alpha_2 \in U_2$. 由 $V = U_1 \oplus W$ 知, 存在 $\alpha_1 \in U_1$ 和 $\beta \in W$ 使得 $\alpha_2 = \alpha_1 + \beta$. 故 $\beta = \alpha_2 - \alpha_1 \in U_2 \cap W$. 根据定理 8.16, $\beta = \mathbf{0}$, 从而 $\alpha_2 = \alpha_1 \in U_1$. 因此, $U_2 = U_1$. \square

定理 8.17. 设 V_1, V_2, \dots, V_k 都是线性空间 V 的有限维子空间. $\sum_{i=1}^k V_i$ 是直和的充分必要条件是 $\dim \sum_{i=1}^k V_i = \sum_{i=1}^k \dim V_i$.

证明. 设 S_i 是 V_i 的基, $S = \bigcup_{i=1}^k S_i$, 则 $U = \sum_{i=1}^k V_i = \text{Span}(S)$, 并且 $|S| \leq \sum_{i=1}^k |S_i| = \sum_{i=1}^k \dim V_i$.

(充分性) 设 $\dim U = \sum_{i=1}^k \dim V_i$, 则 S 是 U 的基, U 中的每个向量 α 都可以唯一地表示为 $\alpha = \sum_{i=1}^k \alpha_i$ 的形式, 其中 $\alpha_i \in V_i$. 因此, $U = \bigoplus_{i=1}^k V_i$.

(必要性) 设 $U = \bigoplus_{i=1}^k V_i$, 则 S 线性无关. 因此, S 是 U 的基, $\dim U = |S| = \sum_{i=1}^k \dim V_i$. \square

定理 8.18. $\sum_{i \in I} V_i$ 是直和的充分必要条件是对于任意 $i \in I, V_i \cap \left(\sum_{j \neq i} V_j \right) = \{\mathbf{0}\}$.

证明. (充分性) 设 $\alpha_j, \beta_j \in V_j$ 满足 $\sum_{j \in I} \alpha_j = \sum_{j \in I} \beta_j$. 若存在 $\alpha_i \neq \beta_i$, 则有 $\alpha_i - \beta_i = \sum_{j \neq i} (\beta_j - \alpha_j) \in V_i \cap \left(\sum_{j \neq i} V_j \right) = \{\mathbf{0}\}$, 与 $\alpha_i \neq \beta_i$ 矛盾. 因此, $\sum_{j \in I} V_j$ 是直和.

(必要性) 对于任意 $i \in I$, 若存在非零向量 $\alpha \in V_i \cap \left(\sum_{j \neq i} V_j \right)$, 则存在 $\{j_1, \dots, j_k\} \subset I \setminus \{i\}$ 和 $\beta_1 \in V_{j_1}, \dots, \beta_k \in V_{j_k}$ 使得 $\alpha = \beta_1 + \dots + \beta_k$. 从而, $\mathbf{0} = \mathbf{0} + \mathbf{0} + \dots + \mathbf{0} = (-\alpha) + \beta_1 + \dots + \beta_k$, 表示方式不唯一, 与 $\sum_{i \in I} V_i$ 是直和矛盾. \square

定义 8.15. 设 V_1, V_2 都是 V 的子空间. 若 $V = V_1 \oplus V_2$, 则 V_1 称为 V_2 的一个补空间.

例 8.23. 求 \mathbb{R}^5 的子空间 $\text{Span}(\alpha_1, \alpha_2, \alpha_3)$ 的一个补空间, 其中

$$\alpha_1 = (1, 0, -2, 2, 2), \quad \alpha_2 = (1, -1, 2, 1, 1), \quad \alpha_3 = (2, 1, 2, 1, 0).$$

解答. 作初等行变换,

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 & 2 & 2 \\ 1 & -1 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -2 & 2 & 2 \\ 0 & -1 & 4 & -1 & -1 \\ 0 & 1 & 6 & -3 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -2 & 2 & 2 \\ 0 & 1 & -4 & 1 & 1 \\ 0 & 0 & 10 & -4 & -5 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix},$$

则 $\text{Span}(\alpha_1, \alpha_2, \alpha_3) = \text{Span}(\beta_1, \beta_2, \beta_3)$. 由 $\beta_1, \beta_2, \beta_3, \mathbf{e}_4, \mathbf{e}_5$ 是 \mathbb{R}^5 的基, 可得 $\text{Span}(\mathbf{e}_4, \mathbf{e}_5)$ 是 $\text{Span}(\alpha_1, \alpha_2, \alpha_3)$ 的一个补空间. \square

下面给出计算 \mathbb{F}^n 的子空间 $V_1 = \text{Span}(\alpha_1, \alpha_2, \dots, \alpha_m)$ 的一个补空间的通常方法.

设 $m \times n$ 矩阵 $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 都是 \mathbb{F} 上的可逆方阵, Q 的行向量分别为 $\beta_1, \beta_2, \dots, \beta_n$, 则向量组 $\alpha_1, \alpha_2, \dots, \alpha_m$ 与 $\beta_1, \beta_2, \dots, \beta_r$ 等价, $\beta_1, \beta_2, \dots, \beta_r$ 是 V_1 的一个基, $V_2 = \text{Span}(\beta_{r+1}, \beta_{r+2}, \dots, \beta_n)$ 是 V_1 的一个补空间, $\beta_{r+1}, \beta_{r+2}, \dots, \beta_n$ 是 V_2 的一个基.

尽管补空间通常是不唯一的, 我们有如下结论.

定理 8.19. 设 U 是 V 的子空间, V_1, V_2 都是 U 的补空间, 则 V_1 与 V_2 同构.

证明. 对于任意 $v \in V_1$, 由 $V = U \oplus V_2$ 知, 存在唯一的 $u \in U$ 和 $w \in V_2$ 使得 $v = u + w$. 故可以定义映射 $\rho: V_1 \rightarrow V_2$, $v \mapsto w$. 为了证明 ρ 是同构映射, 只需验证 ρ 是一一映射并且 ρ 既保加法又保数乘. 设 $v_1, v_2 \in V_1$, $\lambda \in \mathbb{F}$, $v_1 = u_1 + w_1$, $v_2 = u_2 + w_2$, 其中 $u_1, u_2 \in U$, $w_1, w_2 \in V_2$.

(1) 若 $\rho(v_1) = \rho(v_2)$, 即 $w_1 = w_2$, 则 $u_1 - u_2 = v_1 - v_2 \in U \cap V_1$, 得 $v_1 = v_2$. 故 ρ 是单射.

(2) 对于任意 $w \in V_2$, 由 $V = U \oplus V_1$ 知, 存在 $u \in U$ 和 $v \in V_1$ 使得 $w = u + v$, 从而有 $v = (-u) + w$, $w = \rho(v)$. 故 ρ 是满射.

(3) 由 $v_1 + v_2 = (u_1 + u_2) + (w_1 + w_2)$ 可得 $\rho(v_1 + v_2) = w_1 + w_2 = \rho(v_1) + \rho(v_2)$.

(4) 由 $\lambda v = (\lambda u) + (\lambda w)$ 可得 $\rho(\lambda v) = \lambda w = \lambda \rho(v)$.

综上, ρ 是同构映射. \square

习题

1. 设 $\mathbb{R}^{n \times n}$ 的子空间 V_1, V_2 分别由对称方阵和反对称方阵构成. 证明: $\mathbb{R}^{n \times n} = V_1 \oplus V_2$.

2. 设 $\mathbb{R}[x]$ 的子空间 $V_1 = \{f \in \mathbb{R}[x] \mid f(-x) = f(x)\}$, $V_2 = \{f \in \mathbb{R}[x] \mid f(-x) = -f(x)\}$.

证明: $\mathbb{R}[x] = V_1 \oplus V_2$.

3. 设 $A \in \mathbb{R}^{m \times n}$, \mathbb{R}^n 的子空间 V_1 由 A 的行向量生成, V_2 是方程组 $Ax = \mathbf{0}$ 在 \mathbb{R} 上的解空间.

(1) 证明: $\mathbb{R}^n = V_1 \oplus V_2$.

(2) 把实数域 \mathbb{R} 换成任意数域 \mathbb{F} , (1) 中结论是否仍然成立? 证明你的结论.

4. 证明: $\sum_{i=1}^k V_i$ 是直和的充分必要条件是 $\left(\sum_{j=1}^{i-1} V_j\right) \cap V_i = \{\mathbf{0}\}$, $\forall i$.

5. 设 I 是指标集合, $\{V_i \mid i \in I\}$ 是线性空间 V 的一组子空间. 证明: 下列 4 个叙述相互等价.
- ① $\sum_{i \in I} V_i$ 是直和.
 - ② 任取 I 的有限子集 J , $\sum_{j \in J} V_j$ 是直和.
 - ③ 任取 V_i 的基 S_i , 则 $\{S_i \mid i \in I\}$ 两两不相交, 且 $\bigcup_{i \in I} S_i$ 是 $\sum_{i \in I} V_i$ 的基.
 - ④ 任取非零向量 $\alpha_i \in V_i$, $\{\alpha_i \mid i \in I\}$ 是线性无关的.
6. 求实线性空间 $V = \{f \in \mathbb{R}[x] \mid f(1) = 0\}$ 的子空间 $U = \{f \in V \mid f(-1) = 0\}$ 的一个补空间.
7. 求实线性空间 $V = \{A \in \mathbb{C}^{n \times n} \mid \text{tr}(A) = 0\}$ 的子空间 $U = \{A \in V \mid A^T = A\}$ 的一个补空间.
8. 求复线性空间 $V = \{A \in \mathbb{C}^{n \times n} \mid \text{tr}(A) = 0\}$ 的子空间 $U = \{A \in V \mid A^T = A\}$ 的一个补空间.
9. 设 I 是指标集合, $A_i \in \mathbb{R}^{m_i \times n}$, \mathbb{R}^n 的子空间 U_i 由 A_i 的行向量生成, V_i 是方程组 $A_i x = \mathbf{0}$ 在 \mathbb{R} 上的解空间, $\forall i \in I$. $\bigcap_{i \in I} V_i$ 是否一定是 $\sum_{i \in I} U_i$ 的补空间? 证明你的结论.
10. 设 I 是指标集合, V_i 是线性空间 V 的子空间 U_i 的补空间, $\forall i \in I$. $\bigcap_{i \in I} V_i$ 是否一定是 $\sum_{i \in I} U_i$ 的补空间? 证明你的结论.

§8.7 直积与商空间

直积是一个容易与直和相混淆的概念. 商空间是一个容易与补空间相混淆的概念. 这两组概念之间既存在区别又有着联系.

定义 8.16. 设 I 是指标集合, $\{V_i \mid i \in I\}$ 是 \mathbb{F} 上的一组线性空间. 集合

$$V = \{(x_i)_{i \in I} \mid x_i \in V_i\}$$

称为集合族 $V_i (i \in I)$ 的直积或 Cartesian^[6]积, 记作 $V = \prod_{i \in I} V_i$. 在 V 上定义加法和数乘运算

$$(x_i) + (y_i) = (x_i + y_i), \quad \lambda \cdot (x_i) = (\lambda x_i), \quad \lambda \in \mathbb{F}.$$

容易验证, $(V, \mathbb{F}, +, \cdot)$ 构成线性空间, 称为线性空间族 $V_i (i \in I)$ 的直积或 Cartesian 积.

例 8.24. $\mathbb{F}^n = \prod_{i=1}^n \mathbb{F}$.

定理 8.20. 设 V_1, \dots, V_k 都是线性空间 V 的子空间, 且 $\sum_{i=1}^k V_i$ 是直和, 则 $\prod_{i=1}^k V_i$ 与 $\bigoplus_{i=1}^k V_i$ 同构.

证明. 设映射 $\rho: \prod_{i=1}^k V_i \rightarrow \bigoplus_{i=1}^k V_i, (\alpha_1, \dots, \alpha_k) \mapsto \alpha_1 + \dots + \alpha_k$. 由 $\sum_{i=1}^k V_i$ 是直和, 可得 ρ 是一一映射. 对于任意 $\alpha = (\alpha_1, \dots, \alpha_k), \beta = (\beta_1, \dots, \beta_k) \in \prod_{i=1}^k V_i$ 和 $\lambda \in \mathbb{F}$, 有

$$\rho(\alpha + \beta) = \sum_{i=1}^k (\alpha_i + \beta_i) = \rho(\alpha) + \rho(\beta), \quad \rho(\lambda\alpha) = \sum_{i=1}^k \lambda\alpha_i = \lambda\rho(\alpha).$$

因此, ρ 是同构映射. □

根据定理 8.20, 可得如下推论.

定理 8.21. 设 V_1, \dots, V_k 都是 \mathbb{F} 上的有限维线性空间, 则 $\dim \prod_{i=1}^k V_i = \sum_{i=1}^k \dim V_i$.

证明. 设 $\prod_{i=1}^k V_i$ 的子空间 $U_i = \{\mathbf{0}, \dots, \mathbf{0}, \alpha_i, \mathbf{0}, \dots, \mathbf{0} \mid \alpha_i \in V_i\}$, 则 U_i 与 V_i 同构且 $\prod_{i=1}^k V_i = \bigoplus_{i=1}^k U_i$. □

当指标集合 I 是无限集合时, $\prod_{i \in I} V_i$ 与 $\bigoplus_{i \in I} V_i$ 具有完全不同的结构.

例 8.25. 设 $V_i = \{ax^i \mid a \in \mathbb{R}\}, i \in \mathbb{N}$. 每个 V_i 与 \mathbb{R} 同构. $\prod_{i \in \mathbb{N}} V_i$ 与 $\mathbb{R}[[x]]$ 同构. $\bigoplus_{i \in \mathbb{N}} V_i = \mathbb{R}[x]$.

定义 8.17. 设 V 是 \mathbb{F} 上的线性空间, W 是 V 的子空间.

对于任意 $\alpha \in V$, 集合 $\{\alpha + w \mid w \in W\}$ 称为 α 所在的模 W 的同余类, 记作 $\alpha + W$ 或 $[\alpha]$. 在所有模 W 的同余类的集合 $U = \{[\alpha] \mid \alpha \in V\}$ 上定义加法和数乘运算

$$[\alpha] + [\beta] = [\alpha + \beta], \quad \lambda \cdot [\alpha] = [\lambda\alpha], \quad \lambda \in \mathbb{F}.$$

容易验证, ρ 的定义是合理的, 与 α, β 的选取无关. 代数结构 $(U, \mathbb{F}, +, \cdot)$ 构成线性空间, 称为 V 关于 W 的商空间, 记作 $U = V/W$.

^[6]René Descartes, 1596–1650, 法国哲学家、数学家.

定理 8.22. 设 W 是线性空间 V 的子空间, U 是 W 的一个补空间, 则 V/W 与 U 同构.

证明. 设映射 $\rho: U \rightarrow V/W, u \mapsto [u]$. 由于 $V = U \oplus W$, 对于任意 $\alpha \in V$, 存在唯一的 $u \in U$ 和 $w \in W$ 使得 $\alpha = u + w$. $[\alpha] = [u]$ 并且 $[\alpha] \cap U = \{u\}$. 故 ρ 是一一映射. 根据定义 8.17, 对于任意 $\alpha, \beta \in U$ 和 $\lambda \in \mathbb{F}$, 有 $\rho(\alpha + \beta) = \rho(\alpha) + \rho(\beta)$, $\rho(\lambda\alpha) = \lambda\rho(\alpha)$. 因此, ρ 是同构映射. \square

根据定理 8.22, 可得如下推论.

定理 8.23. 设 W 是有限维线性空间 V 的子空间, 则 $\dim(V/W) = \dim V - \dim W$.

例 8.26. 设 $A \in \mathbb{F}^{m \times n}$, U 是 A 的列向量生成的 \mathbb{F}^m 的子空间, W 是线性方程组 $Ax = \mathbf{0}$ 的解空间, 构造从 \mathbb{F}^n/W 到 U 的同构映射.

解答. 设映射 $\rho: \mathbb{F}^n/W \rightarrow U, [\alpha] \mapsto A\alpha$. 对于任意 $\alpha, \beta \in \mathbb{F}^n$, $[\alpha] = [\beta] \Leftrightarrow A\alpha = A\beta$. 故 ρ 的定义是合理的, 并且 ρ 是单射. 对于任意 $u \in U$, u 是 A 的列向量的线性组合, 线性方程组 $Ax = u$ 有解. 故 ρ 是满射. 容易验证, ρ 既保加法又保数乘. 因此, ρ 是同构映射. \square

习题

1. 设 V 是定义在区间 I 上的所有实值函数 $f: I \rightarrow \mathbb{R}$ 构成的集合. 证明: V 在函数的加法、数乘运算下构成 \mathbb{R} 上的线性空间, 并且 V 与 $\prod_{i \in I} \mathbb{R}$ 同构.
2. 设 U_i, W_i 都是 \mathbb{F} 上的线性空间 V_i 的子空间, $i = 1, 2$. 证明:
 - (1) $(U_1 + W_1) \times (U_2 + W_2)$ 是 $V_1 \times V_2$ 的子空间.
 - (2) 设 $U_1 + W_1$ 是直和. 是否一定有 $(U_1 \oplus W_1) \times V_2 = (U_1 \times V_2) \oplus (W_1 \times V_2)$? 证明你的结论.
3. 设 U, W 都是线性空间 V 的子空间. 证明: $(U + W)/W$ 与 $U/(U \cap W)$ 同构.
4. 设 V 是 \mathbb{R}^n 的 r 维子空间, $r < n$. 证明: 存在 $A \in \mathbb{R}^{(n-r) \times n}$ 使得 $V = \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\}$.
5. 设 $A \in \mathbb{F}^{m \times n}$, U 是 A 的行向量生成的 \mathbb{F}^m 的子空间, W 是线性方程组 $Ax = \mathbf{0}$ 的解空间, 构造从 \mathbb{F}^n/W 到 U 的同构映射.
6. 设 I 是指标集合, W_i 是 \mathbb{F} 上的线性空间 V_i 的子空间, $\forall i \in I$. 证明: $W = \prod_{i \in I} W_i$ 是 $V = \prod_{i \in I} V_i$ 的子空间, 并且 V/W 与 $\prod_{i \in I} (V_i/W_i)$ 同构.
7. 设 I 是指标集合, V_i 是线性空间 V 的子空间, W_i 是 V_i 的子空间, $\forall i \in I$. 记 $U = \bigoplus_{i \in I} V_i$, $W = \bigoplus_{i \in I} W_i$. U/W 与 $\prod_{i \in I} (V_i/W_i)$ 是否一定同构? 证明你的结论.

第九章 线性变换

§9.1 基本概念

定义 9.1. 设 $(U, \mathbb{F}, \oplus, *)$ 和 $(V, \mathbb{F}, +, \cdot)$ 都是 \mathbb{F} 上的线性空间. 若映射 $\mathcal{A}: U \rightarrow V$ 满足

$$\mathcal{A}(\alpha \oplus \beta) = \mathcal{A}(\alpha) + \mathcal{A}(\beta), \quad \mathcal{A}(\lambda * \alpha) = \lambda \cdot \mathcal{A}(\alpha), \quad \forall \alpha, \beta \in U, \lambda \in \mathbb{F}$$

则 \mathcal{A} 称为 $U \rightarrow V$ 的线性映射或线性算子. 在不引起歧义的情况下, 可以把 $\mathcal{A}(\alpha)$ 记作 $\mathcal{A}\alpha$, 表示线性算子 \mathcal{A} 作用于向量 α . $U \rightarrow V$ 的线性映射的全体记作 $L(U, V)$. 特别, 当 $U = V$ 时, 线性映射 $\mathcal{A}: V \rightarrow V$ 也称为 V 上的线性变换. V 上的线性变换的全体记作 $L(V)$.

例 9.1. 设 U, V 都是 \mathbb{F} 上的线性空间. 容易验证, 下列映射都是线性映射.

1. 零映射 $\mathcal{O}: U \rightarrow V, \alpha \mapsto \mathbf{0}$.
2. 恒等变换 $\mathcal{I}: V \rightarrow V, \alpha \mapsto \alpha$.
3. 转置映射 $\mathcal{T}: \mathbb{F}^{m \times n} \rightarrow \mathbb{F}^{n \times m}, X \mapsto X^T$.
4. 乘积映射 $\mathcal{M}: \mathbb{F}^{m \times n} \rightarrow \mathbb{F}^{p \times q}, X \mapsto PXQ$, 其中 $P \in \mathbb{F}^{p \times m}, Q \in \mathbb{F}^{n \times q}$ 是给定的常矩阵.
5. 微分变换 $\mathcal{D}: \mathbb{F}[x] \rightarrow \mathbb{F}[x], p(x) \mapsto p'(x)$.
6. 积分变换 $\mathcal{S}: \mathbb{F}[x] \rightarrow \mathbb{F}[x], p(x) \mapsto \int_0^x p(t)dt$, 其中 $\text{Char } \mathbb{F} = 0$.
7. 嵌入映射 $\sigma: U \rightarrow U \times V, \alpha \mapsto (\alpha, \mathbf{0})$.
8. 投影映射 $\pi: U \rightarrow U/V, \alpha \mapsto [\alpha]$, 其中 V 是 U 的子空间.

定理 9.1. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$. 有下列结论.

1. $\mathcal{A}(\mathbf{0}) = \mathbf{0}$.
2. $\mathcal{A}(\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_n \alpha_n) = \lambda_1 \mathcal{A}(\alpha_1) + \lambda_2 \mathcal{A}(\alpha_2) + \cdots + \lambda_n \mathcal{A}(\alpha_n), \forall \alpha_i \in U, \lambda_i \in \mathbb{F}$.
3. 若 $\alpha_1, \alpha_2, \cdots, \alpha_n \in U$ 线性相关, 则 $\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n \in V$ 线性相关.
4. 若 $\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n \in V$ 线性无关, 则 $\alpha_1, \alpha_2, \cdots, \alpha_n \in U$ 线性无关.
5. \mathcal{A} 是单射当且仅当对于任意非零向量 $\alpha \in U, \mathcal{A}\alpha \neq \mathbf{0}$.
6. \mathcal{A} 是满射当且仅当存在 $S \subset U$ 使得 $\mathcal{A}(S)$ 是 V 的基.

定理 9.2. 设 U, V 是 \mathbb{F} 上的线性空间, I 是指标集合, $\{\alpha_i \mid i \in I\}$ 是 U 的基, $\{\beta_i \mid i \in I\}$ 是 V 中任意一组向量, 则存在唯一的线性映射 $\mathcal{A} \in L(U, V)$ 使得 $\mathcal{A}\alpha_i = \beta_i, \forall i \in I$.

证明. (存在性) 构造映射 $\mathcal{A}: U \rightarrow V$ 如下. 设 $\mathcal{A}(\mathbf{0}) = \mathbf{0}$. 对于任意非零向量 $\alpha \in U$, 存在唯一的正整数 k 和 $i_1, i_2, \dots, i_k \in I$ 和 $x_1, x_2, \dots, x_k \in \mathbb{F} \setminus \{0\}$ 使得 $\alpha = x_1\alpha_{i_1} + x_2\alpha_{i_2} + \dots + x_k\alpha_{i_k}$. 设 $\mathcal{A}(\alpha) = x_1\beta_{i_1} + x_2\beta_{i_2} + \dots + x_k\beta_{i_k}$. 容易验证 \mathcal{A} 是线性映射.

(唯一性) 若线性映射 \mathcal{B} 也满足 $\mathcal{B}\alpha_i = \beta_i, \forall i$, 则 $\mathcal{B}\alpha = x_1\mathcal{B}\alpha_{i_1} + x_2\mathcal{B}\alpha_{i_2} + \dots + x_k\mathcal{B}\alpha_{i_k} = \mathcal{A}\alpha, \forall \alpha \in V$. 因此, $\mathcal{B} = \mathcal{A}$. \square

当 U, V 都是有限维线性空间时, 我们可以定义线性映射 $\mathcal{A} \in L(U, V)$ 的矩阵表示.

定义 9.2. 取定 U 的基 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 和 V 的基 $T = \{\beta_1, \beta_2, \dots, \beta_m\}$. 设

$$\mathcal{A}\alpha_j = \sum_{i=1}^m a_{ij}\beta_i, \quad \forall j$$

或者形式上记作

$$\begin{pmatrix} \mathcal{A}\alpha_1 & \mathcal{A}\alpha_2 & \cdots & \mathcal{A}\alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

则矩阵 $A = (a_{ij}) \in \mathbb{F}^{m \times n}$ 称为线性映射 \mathcal{A} 在 S, T 下的矩阵表示. 也就是说, \mathcal{A} 在 S, T 下的矩阵表示由 $\mathcal{A}(S)$ 在 T 下的坐标构成. \mathcal{A} 与 A 的关系如下图所示.

$$\begin{array}{ccc} \alpha \in U & \xrightarrow{\mathcal{A}} & \mathcal{A}\alpha \in V \\ \downarrow S & & \downarrow T \\ x \in \mathbb{F}^n & \xrightarrow{A} & Ax \in \mathbb{F}^m \end{array}$$

特别, 当 $U = V$ 且 $\beta_i = \alpha_i$ 时, A 也称为线性变换 \mathcal{A} 在 S 下的矩阵表示.

对于任意 $\alpha \in U$, 设 $\alpha = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n$, $\mathcal{A}\alpha = y_1\beta_1 + y_2\beta_2 + \dots + y_m\beta_m$. 由

$$\mathcal{A}\alpha = \sum_{j=1}^n x_j \mathcal{A}\alpha_j = \sum_{i=1}^m x_j \left(\sum_{j=1}^n a_{ij}\beta_i \right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}x_j \right) \beta_i$$

可得 $y_i = \sum_{j=1}^n a_{ij}x_j, \forall i$, 即 $y = Ax$. 这称为线性映射 \mathcal{A} 在 S, T 下的坐标表示.

例 9.2. 转置映射 $f: \mathbb{F}^{2 \times 3} \rightarrow \mathbb{F}^{3 \times 2}, X \mapsto X^T$ 在 $\mathbb{F}^{2 \times 3}$ 的基 $E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}$ 和 $\mathbb{F}^{3 \times 2}$ 的基 $E_{11}, E_{12}, E_{21}, E_{22}, E_{31}, E_{32}$ 下的矩阵表示为

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

例 9.3. 微分映射 $\mathcal{D}: \mathbb{F}_n[x] \rightarrow \mathbb{F}_n[x]$, $f(x) \mapsto f'(x)$ 在 $\mathbb{F}_n[x]$ 的基 $1, x, \dots, x^{n-1}$ 下的矩阵表示为

$$\begin{pmatrix} 0 & 1 & & & & \\ & 0 & 2 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & n-1 & \\ & & & & 0 & \end{pmatrix}.$$

例 9.4. 在例 2.11 中, $y = (A \otimes I)x$ 是 $\mathbb{F}^{2 \times 2}$ 上的线性变换 $X \mapsto AX$ 在 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的坐标表示, $y = (I \otimes B^T)x$ 是 $X \mapsto XB$ 的坐标表示, $y = (A \otimes B^T)x$ 是 $X \mapsto AXB$ 的坐标表示.

由于每个线性空间可以取不同的基, 一个线性映射在不同基下的矩阵表示一般是不同的, 那么它们之间有何联系呢?

定理 9.3. 设 U, V 都是 \mathbb{F} 上的有限维线性空间, S_1, S_2 是 U 的两组基, P 是从 S_1 到 S_2 的过渡矩阵, T_1, T_2 是 V 的两组基, Q 是从 T_1 到 T_2 的过渡矩阵, 线性映射 $\mathcal{A}: U \rightarrow V$ 在 S_i, T_i 下的矩阵表示为 A_i , 则有 $A_2 = Q^{-1}A_1P$. 因此, A_1 与 A_2 相抵. 特别, 当 $U = V$, $S_1 = T_1$, $S_2 = T_2$ 时, A_1 与 A_2 相似.

证明. P 的列向量是 S_2 在 S_1 下的坐标. 根据 \mathcal{A} 的坐标表示, A_1P 的列向量是 $\mathcal{A}(S_2)$ 在 T_1 下的坐标. 再根据坐标变换公式, $\mathcal{A}(S_2)$ 在 T_2 下的坐标 $A_2 = Q^{-1}A_1P$. \square

习题

1. 证明定理 9.1.

2. 判断下列映射 \mathcal{A} 是否 $V \rightarrow V$ 的线性映射, 并证明你的结论.

(1) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{C}^{n \times n}$, $X \mapsto X^T$.

(2) $\mathbb{F} = \mathbb{C}$, $V = \mathbb{C}^{n \times n}$, $X \mapsto X^T$.

(3) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{C}^{n \times n}$, $X \mapsto X^H$.

(4) $\mathbb{F} = \mathbb{C}$, $V = \mathbb{C}^{n \times n}$, $X \mapsto X^H$.

(5) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{C}^{2 \times 2}$, $X \mapsto X^*$ (伴随方阵).

(6) $\mathbb{F} = \mathbb{C}$, $V = \mathbb{C}^{2 \times 2}$, $X \mapsto X^+$ (Moore-Penrose 广义逆).

(7) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}[x]$, $p(x) \mapsto p(x^2)$.

(8) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}[x]$, $p(x) \mapsto (p(x))^2$.

(9) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}[x]$, $p(x) \mapsto \int_{1-x^2}^{1+x^2} p(t)dt$.

(10) $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}[x]$, $p(x) \mapsto \int_{-\infty}^{+\infty} e^{-(x+t)^2} p(t^2)dt$.

3. 视 \mathbb{C} 为 \mathbb{R} 上的线性空间. 映射 $\mathcal{A}: \mathbb{C} \rightarrow \mathbb{R}^{2 \times 2}$, $a + bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ 是否线性映射/单射/满射/可逆映射? 证明你的结论.

4. 设 U, V 是 \mathbb{F} 上的线性空间, I 是指标集合, $\{\alpha_i \mid i \in I\} \subset U$, $\{\beta_i \mid i \in I\} \subset V$. 证明: 下列两个叙述等价.
- ① 存在线性映射 $\mathcal{A}: U \rightarrow V$ 使得 $f(\alpha_i) = \beta_i, \forall i \in I$.
 - ② 若 \mathbb{F} 中仅有限多项非零的数列 $\{x_i \mid i \in I\}$ 满足 $\sum_{i \in I} x_i \alpha_i = \mathbf{0}$, 则 $\sum_{i \in I} x_i \beta_i = \mathbf{0}$.
5. 写出下列线性变换 $\mathcal{A} \in L(V)$ 在标准基或指定基下的矩阵表示.
- (1) $V = \mathbb{R}^2$, \mathcal{A} 把任意平面向量绕原点顺时针旋转 θ .
 - (2) $V = \mathbb{R}^3$, \mathcal{A} 把任意空间向量绕 $(1, 1, 1)$ 右手定则旋转 θ .
 - (3) $V = \mathbb{R}^3$, \mathcal{A} 把任意点垂直投影到平面 $x_1 + x_2 + x_3 = 0$ 上.
 - (4) $V = \mathbb{R}^3$, \mathcal{A} 把任意点映到它关于平面 $x_1 + x_2 + x_3 = 0$ 的对称点.
 - (5) $V = \mathbb{F}^{2 \times 2}$, $\mathcal{A}(X) = \text{tr}(X)A$, 其中 $A \in V$ 是常矩阵.
 - (6) $V = \mathbb{F}^{2 \times 2}$, $\mathcal{A}(X) = PXQ$, 其中 $P, Q \in V$ 是常矩阵.
 - (7) $V = \mathbb{F}^{2 \times 2}$, $\mathcal{A}(X) = PX^TQ$, 其中 $P, Q \in V$ 是常矩阵.
 - (8) $V = \mathbb{R}_4[x]$, 基 $1, x - a, (x - a)^2, (x - a)^3$, \mathcal{A} 是微分变换, 其中 $a \in \mathbb{R}$ 是常数.
 - (9) $V = \{(c_0 + c_1x + c_2x^2 + c_3x^3)e^x \mid c_0, c_1, c_2, c_3 \in \mathbb{R}\}$, \mathcal{A} 是微分变换.
 - (10) $V = \{a_1 \cos x + a_2 \cos(2x) + b_1 \sin x + b_2 \sin(2x) \mid a_1, a_2, b_1, b_2 \in \mathbb{R}\}$, \mathcal{A} 是微分变换.
6. 已知 \mathbb{R}^3 上的线性变换 $\mathcal{A}: \alpha_i \mapsto \beta_i, \forall i = 1, 2, 3$. 分别求 \mathcal{A} 在 $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}, \{\alpha_1, \alpha_2, \alpha_3\}, \{\beta_1, \beta_2, \beta_3\}$ 下的矩阵表示, 其中
- (1) $\alpha_1 = (0, 1, 0), \alpha_2 = (1, 0, 1), \alpha_3 = (1, 1, 0), \beta_1 = (0, 1, 1), \beta_2 = (0, 0, 1), \beta_3 = (1, 1, 1)$.
 - (2) $\alpha_1 = (0, 1, 1), \alpha_2 = (1, 0, 1), \alpha_3 = (1, 1, 0), \beta_1 = (2, 1, 1), \beta_2 = (1, 2, 1), \beta_3 = (1, 1, 2)$.
7. 分别求满足条件的所有线性变换 $\mathcal{A} \in L(\mathbb{F}^{n \times n})$.
- (1) $\mathcal{A}(X^T) = \mathcal{A}(X)$ (2) $\mathcal{A}(X^T) = (\mathcal{A}(X))^T$ (3) $\mathcal{A}(XY) = \mathcal{A}(YX)$
 - (4) $\mathcal{A}(XY) = \mathcal{A}(X)\mathcal{A}(Y)$ ^[1] (5) $\mathcal{A}(XY) = \mathcal{A}(Y)\mathcal{A}(X)$

^[1]参考第五届全国大学生数学竞赛数学类预赛第六题.

§9.2 线性映射的运算

定义 9.3. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in L(U, V)$, $\lambda \in \mathbb{F}$. 定义线性映射的加法运算 $\mathcal{A} + \mathcal{B}$ 和数乘运算 $\lambda\mathcal{A}$ 如下.

$$(\mathcal{A} + \mathcal{B})\alpha = \mathcal{A}\alpha + \mathcal{B}\alpha, \quad (\lambda\mathcal{A})\alpha = \lambda(\mathcal{A}\alpha), \quad \forall \alpha \in U.$$

容易验证, $\mathcal{A} + \mathcal{B}$ 和 $\lambda\mathcal{A}$ 都是 $U \rightarrow V$ 的线性映射. 从而, $L(U, V)$ 和 $L(V)$ 在线性映射的加法和数乘运算下均构成 \mathbb{F} 上的线性空间.

定理 9.4. 设 U, V 都是 \mathbb{F} 上的有限维线性空间, $\dim U = n$, $\dim V = m$, S, T 分别是 U, V 的基. 对于任意线性映射 $\mathcal{A} \in L(U, V)$, 设 A 是 \mathcal{A} 在 S, T 下的矩阵表示, 则映射

$$\rho: L(U, V) \rightarrow \mathbb{F}^{m \times n}, \quad \mathcal{A} \mapsto A$$

是线性映射, 也是同构映射. 从而, $\dim L(U, V) = mn$.

定义 9.4. 设 U, V, W 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$, $\mathcal{B} \in L(V, W)$, 则复合映射

$$\mathcal{B} \circ \mathcal{A}: U \rightarrow W, \quad \alpha \mapsto \mathcal{B}(\mathcal{A}(\alpha))$$

也是线性映射. $\mathcal{B} \circ \mathcal{A}$ 称为线性映射 \mathcal{B} 与 \mathcal{A} 的乘积. 在不引起歧义的情况下, 可以把 $\mathcal{B} \circ \mathcal{A}$ 记作 BA , 称为两个线性算子的乘积.

设 $\mathcal{A} \in L(V)$, k 是正整数. $\mathcal{A}^k = \underbrace{\mathcal{A} \circ \cdots \circ \mathcal{A}}_k$ 称为 \mathcal{A} 的 k 次方幂. 特别, 规定 $\mathcal{A}^0 = \mathcal{I}$ 是恒等变换. 对于任意 $p(x) = c_0 + c_1x + \cdots + c_kx^k \in \mathbb{F}[x]$, 线性变换 $c_0\mathcal{I} + c_1\mathcal{A} + \cdots + c_k\mathcal{A}^k \in L(V)$ 称为 \mathcal{A} 的多项式, 记作 $p(\mathcal{A})$.

容易验证, 对于任意 $\mathcal{A} \in L(V)$ 和 $p(x), q(x) \in \mathbb{F}[x]$, 有

$$(p + q)(\mathcal{A}) = p(\mathcal{A}) + q(\mathcal{A}), \quad (p \cdot q)(\mathcal{A}) = p(\mathcal{A})q(\mathcal{A}).$$

也就是说, 对于取定的线性变换 \mathcal{A} , 映射 $p(x) \mapsto p(\mathcal{A})$ 保持多项式的加法、乘法运算与线性变换的加法、乘法运算之间的对应关系, 这种映射称为从 $\mathbb{F}[x]$ 到 $\mathbb{F}[\mathcal{A}]$ 的同态映射.

定理 9.5. 设 S_U, S_V, S_W 分别是 \mathbb{F} 上的有限维线性空间 U, V, W 的基, A 是 $\mathcal{A} \in L(U, V)$ 在 S_U, S_V 下的矩阵表示, B 是 $\mathcal{B} \in L(V, W)$ 在 S_V, S_W 下的矩阵表示, 则 BA 是 $\mathcal{B} \circ \mathcal{A}$ 在 S_U, S_W 下的矩阵表示.

由定理 9.4 和定理 9.5 可知, 在取定有限维线性空间的基的情况下, 线性映射与其矩阵表示一一对应, 并且线性映射的乘积 (复合) 与矩阵的乘积一一对应.

$$\begin{array}{ccccc} \alpha \in U & \xrightarrow{\mathcal{A}} & \mathcal{A}\alpha \in V & \xrightarrow{\mathcal{B}} & \mathcal{B}(\mathcal{A}\alpha) \in W \\ S_U \downarrow & & \downarrow S_V & & \downarrow S_W \\ x & \xrightarrow{A} & y = Ax & \xrightarrow{B} & z = By = BAx \end{array}$$

定义 9.5. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$, $\mathcal{B} \in L(V, U)$. 若 $\mathcal{A}\mathcal{B} = \mathcal{I}_V$, 则称 \mathcal{A} 是 \mathcal{B} 的左逆, \mathcal{B} 是 \mathcal{A} 的右逆. 若 \mathcal{A} 既有左逆 \mathcal{B} 又有右逆 \mathcal{C} , 则有 $\mathcal{B} = \mathcal{B}(\mathcal{A}\mathcal{C}) = (\mathcal{B}\mathcal{A})\mathcal{C} = \mathcal{C}$, \mathcal{B} 称为 \mathcal{A} 的逆映射, 记作 $\mathcal{B} = \mathcal{A}^{-1}$, 并称 \mathcal{A} 是可逆的.

定理 9.6. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$. \mathcal{A} 有左逆/有右逆/可逆当且仅当 \mathcal{A} 是单射/满射/一一映射.

证明. 当 \mathcal{A} 有左逆 \mathcal{B} 时, $\mathcal{A}\alpha = \mathcal{A}\beta \Rightarrow \mathcal{B}(\mathcal{A}\alpha) = \mathcal{B}(\mathcal{A}\beta) \Rightarrow \alpha = \beta$. 故 \mathcal{A} 是单射.

当 \mathcal{A} 是单射时, 设 $\{\alpha_i \mid i \in I\}$ 是 U 的一组基, $\beta_i = \mathcal{A}\alpha_i$, 则 $\{\beta_i \mid i \in I\}$ 线性无关, 可以扩充为 V 的一组基 $\{\beta_j \mid j \in J\}$, 其中 $I \subset J$. 根据定理 9.2, 存在 $\mathcal{B} \in L(V, U)$ 使得 $\mathcal{B}\beta_j = \begin{cases} \alpha_j, & j \in I \\ 0, & j \notin I \end{cases}$.

\mathcal{B} 是 \mathcal{A} 的左逆.

当 \mathcal{A} 有右逆 \mathcal{B} 时, 对于任意 $\alpha \in V$, $\beta = \mathcal{B}\alpha$ 满足 $\mathcal{A}\beta = \alpha$. 故 \mathcal{A} 是满射.

当 \mathcal{A} 是满射时, 设 $\{\alpha_i \mid i \in I\}$ 是 V 的一组基, $\alpha_i = \mathcal{A}\beta_i$. 根据定理 9.2, 存在 $\mathcal{B} \in L(V, U)$ 使得 $\mathcal{B}\alpha_i = \beta_i, \forall i \in I$. \mathcal{B} 是 \mathcal{A} 的右逆.

综上, \mathcal{A} 可逆 $\Leftrightarrow \mathcal{A}$ 既有左逆又有右逆 $\Leftrightarrow \mathcal{A}$ 既是单射又是满射 $\Leftrightarrow \mathcal{A}$ 是一一映射. \square

定理 9.7. 设 U, V 都是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(U, V)$, S, T 分别是 U, V 的基, A 是 \mathcal{A} 在 S, T 下的矩阵表示. \mathcal{A} 有左逆/有右逆/可逆当且仅当 A 有左逆/有右逆/可逆.

证明. 设 B 是 $\mathcal{B} \in L(V, U)$ 在 T, S 下的矩阵表示. $\mathcal{B}\mathcal{A} = \mathcal{I}_U \Leftrightarrow BA = I. \mathcal{A}\mathcal{B} = \mathcal{I}_V \Leftrightarrow AB = I. \square$

习题

1. 证明定理 9.4 和定理 9.5.
2. 设 $U = \mathbb{F}^m, V = \mathbb{F}^n$. 求 $L(U, V)$ 的一个基 S , 以及任意 $\mathcal{A} \in L(U, V)$ 在 S 下的坐标.
3. 设 $\mathcal{A}, \mathcal{B} \in L(\mathbb{F}^n)$, $\text{Char } \mathbb{F} \nmid n$. 证明: $\mathcal{C} = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$ 一定不是恒等变换.
4. 设 $U = \mathbb{F}^{m \times n}, V = \mathbb{F}^{p \times q}, \mathcal{A} \in L(U, V): X \mapsto PXQ, P \in \mathbb{F}^{p \times m}, Q \in \mathbb{F}^{n \times q}$.
 - (1) 当 P, Q 满足何条件时, \mathcal{A} 有左逆? 并求 \mathcal{A} 的所有左逆.
 - (2) 当 P, Q 满足何条件时, \mathcal{A} 有右逆? 并求 \mathcal{A} 的所有右逆.
 - (3) 当 P, Q 满足何条件时, \mathcal{A} 可逆? 并求 \mathcal{A}^{-1} .
5. 设 $V = \mathbb{F}^{n \times n}, \mathcal{A} \in L(V): X \mapsto aX + bX^T, a, b \in \mathbb{F}$.
 - (1) 求 a, b 应满足的充分必要条件使得 \mathcal{A} 可逆.
 - (2) 求 \mathcal{A} 的最小多项式.
6. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$. 证明:
 - (1) 存在非零多项式 $p(x) \in \mathbb{F}[x]$ 使得 $p(\mathcal{A}) = \mathcal{O}$.
 - (2) \mathcal{A} 是单射 $\Leftrightarrow \mathcal{A}$ 是满射 $\Leftrightarrow \mathcal{A}$ 可逆.
 - (3) 当 V 是无限维时, (1) 和 (2) 是否仍然成立? 证明你的结论.
7. 设 $V = \mathbb{F}_n[x], \text{Char } \mathbb{F} = 0, \mathcal{A} \in L(V): f(x) \mapsto f(x+1)$.
 - (1) 证明: \mathcal{A} 是 V 上的可逆线性变换.
 - (2) 求 \mathcal{A} 在 $1, x, x^2, \dots, x^{n-1}$ 下的矩阵表示.
 - (3) 设 $\mathcal{D} \in L(V)$ 是微分变换, $p(x) = \sum_{k=0}^{n-1} \frac{x^k}{k!} \in \mathbb{F}[x]$. 证明: $\mathcal{A} = p(\mathcal{D})$.

8. 设 $V = \mathbb{F}[x]$, $\text{Char } \mathbb{F} = 0$, $\mathcal{D}, \mathcal{S} \in L(V)$ 分别是例 9.1 中的微分变换和积分变换.

(1) 求 $\mathcal{D} \circ \mathcal{S}$ 和 $\mathcal{S} \circ \mathcal{D}$, 并说明 \mathcal{D} 和 \mathcal{S} 是否有左逆/有右逆/可逆.

(2) 设 $\mathcal{M} \in L(V): f(x) \mapsto xf(x)$. 证明: $\mathcal{D} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D} = \mathcal{I}$ 是恒等变换.

(3) 设 $\mathcal{A} \in L(V)$ 满足 $\mathcal{A} \circ \mathcal{D} = \mathcal{D} \circ \mathcal{A}$. 证明: 存在数列 $\{a_k\} \subset \mathbb{F}$ 使得 $\mathcal{A}(f) = \sum_{k=0}^{\deg f} a_k \mathcal{D}^k(f)$.

(4) 设 $\mathcal{A} \in L(V)$ 满足 $\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A}$. 证明: 存在 $p \in \mathbb{F}[x]$ 使得 $\mathcal{A} = p(\mathcal{S})$.

§9.3 对偶空间

定义 9.6. 设 V 是 \mathbb{F} 上的线性空间. $V^* = L(V, \mathbb{F})$ 称为 V 的对偶空间. 任意 $f \in V^*$ 也称为 V 上的线性函数或线性泛函.

例 9.5. 设 V 是 \mathbb{F} 上的线性空间, $S = \{\alpha_i \mid i \in I\}$ 是 V 的基. 映射 $p_k : \sum_{i \in I} x_i \alpha_i \mapsto x_k$ 是 V 上的线性函数, $\forall k \in I$.

例 9.6. 设 $V = \mathcal{C}[0, 1]$, $a \in [0, 1]$, $g \in V$. 映射 $\tau : f \mapsto f(a)$ 和 $\sigma : f \mapsto \int_0^1 f(t)g(t)dt$ 都是 V 上的线性函数.

定理 9.8. 设 $\dim V = n < \infty$, 则 $\dim V^* = n$, 从而 V^* 与 V 同构.

证明. 任取 V 的基 $\alpha_1, \alpha_2, \dots, \alpha_n$, 设 $p_k \in V^* : \sum_{i=1}^n x_i \alpha_i \mapsto x_k$. 注意到

(1) 若 $\sum_{i=1}^n \lambda_i p_i$ 是零函数, 则 $\left(\sum_{i=1}^n \lambda_i p_i\right)(\alpha_k) = \lambda_k = 0, \forall k$. 故 p_1, p_2, \dots, p_n 线性无关.

(2) 对任意 $f \in V^*$ 和 $v = \sum_{i=1}^n x_i \alpha_i \in V$, $f(v) = \sum_{i=1}^n x_i f(\alpha_i) = \sum_{i=1}^n p_i(v) f(\alpha_i) = \left(\sum_{i=1}^n f(\alpha_i) p_i\right)(v)$.
故 $f = \sum_{i=1}^n f(\alpha_i) p_i$.

因此, p_1, p_2, \dots, p_n 是 V^* 的基, $\dim V^* = n$. 根据定理 8.12, V^* 与 V 同构. \square

定义 9.7. 设 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 n 维线性空间 V 的基. 定理 9.8 的证明中的 $\{p_1, p_2, \dots, p_n\}$ 称为 S 的对偶基. 通常记 $S^* = \{\alpha_1^*, \alpha_2^*, \dots, \alpha_n^*\}$, 其中 $\alpha_i^* = p_i$. 任意 $f \in V^*$ 在 S^* 下的坐标为 $(f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n))$.

例 9.7. 设 $V = \mathbb{F}^{n \times n}$, $S = \{E_{ij} \mid 1 \leq i, j \leq n\}$ 是 V 的基, S 的对偶基 $S^* = \{E_{ij}^* \mid 1 \leq i, j \leq n\}$, 其中 $E_{ij}^* : A \mapsto a_{ij}$. 迹函数 $\text{tr} \in V^*$ 可以表示成 $\text{tr} = \sum_{i=1}^n E_{ii}^*$.

例 9.8. 设 $V = \mathbb{F}_n[x]$, $S = \{(x-a)^{i-1} \mid 1 \leq i \leq n\}$ 是 V 的基, $a \in \mathbb{F}$. 由例 8.17 可知, S 的对偶基 $S^* = \{\alpha_i^* \mid 1 \leq i \leq n\}$, 其中 $\alpha_i^*(f) = \frac{f^{(i-1)}(a)}{(i-1)!}$.

例 9.9. 设 $V = \mathbb{R}[x]$. 对于任意 $a \in \mathbb{R}$, 定义 $\pi_a \in V^* : f(x) \mapsto f(a)$, 则 $S = \{\pi_a \mid a \in \mathbb{R}\}$ 是线性无关的. 由此可知, $\dim V$ 是可数的, $\dim V^*$ 是不可数的.

证明. 对于 S 的任意有限子集 $\{\pi_{a_1}, \pi_{a_2}, \dots, \pi_{a_n}\}$, 若 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$ 使得 $\lambda_1 \pi_{a_1} + \lambda_2 \pi_{a_2} + \dots + \lambda_n \pi_{a_n} = 0$, 则 $(\lambda_1 \pi_{a_1} + \lambda_2 \pi_{a_2} + \dots + \lambda_n \pi_{a_n}) \left(\prod_{\substack{1 \leq j \leq n \\ j \neq i}} (x - a_j) \right) = \lambda_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} (a_i - a_j) = 0$, 得 $\lambda_i = 0, \forall 1 \leq i \leq n$. 因此, S 是线性无关的. \square

设 V 是无限维线性空间, $S = \{\alpha_i \mid i \in I\}$ 是 V 的基, $p_k \in V^* : \sum_{i \in I} x_i \alpha_i \mapsto x_k$. 容易验证, 向量组 $S^* = \{p_i \mid i \in I\}$ 线性无关. 由下例可知 S^* 不是 V^* 的基.

例 9.10. 设 $f \in V^* : \sum_{i \in I} x_i \alpha_i \mapsto \sum_{i \in I} x_i$, 则 $f \notin \text{Span}(S^*)$.

证明. 假设 $f \in \text{Span}(S^*)$, 则 $f = \lambda_1 \alpha_{i_1}^* + \cdots + \lambda_k \alpha_{i_k}^*$, 其中 $i_1, \cdots, i_k \in I$, $\lambda_1, \cdots, \lambda_k \in \mathbb{F}$. 任取 $j \in I \setminus \{i_1, \cdots, i_k\}$, 则 $f(\alpha_j) = \lambda_1 \alpha_{i_1}^*(\alpha_j) + \cdots + \lambda_k \alpha_{i_k}^*(\alpha_j) = 0$, 与 $f(\alpha_j) = 1$ 矛盾. \square

一般说来, V^* 与 $\prod_{i \in I} \mathbb{F}$ 同构, 与 V 可能不同构. 容易验证, $\rho: V^* \rightarrow \prod_{i \in I} \mathbb{F}$, $f \mapsto (f(\alpha_i))_{i \in I}$ 是同构映射.

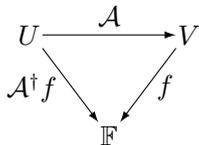
定义 9.8. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$. 线性映射

$$\mathcal{A}^\dagger \in L(V^*, U^*): f \mapsto f \circ \mathcal{A}$$

称为 \mathcal{A} 的伴随映射 (operator adjoint). 换句话说,

$$(\mathcal{A}^\dagger f)(x) = f(\mathcal{A}x), \quad \forall x \in U.$$

各映射之间的关系如下图所示.



定理 9.9. 设 U, V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(U, V)$, S, T 分别是 U, V 的基, S^*, T^* 分别是 S, T 的对偶基, A 是 \mathcal{A} 在 S, T 下的矩阵表示, 则 A^T 是 \mathcal{A}^\dagger 在 T^*, S^* 下的矩阵表示.

证明. \mathcal{A} 在 S, T 下的坐标表示 $\mathcal{A}(x) = Ax$. 任取 $f \in V^*$, 设 f 在 T^* 下的坐标是 c , 则 f 的坐标表示 $f(x) = c^T x$, $f \circ \mathcal{A}$ 的坐标表示 $(f \circ \mathcal{A})(x) = c^T (Ax) = (c^T A)x$, 得 $f \circ \mathcal{A}$ 在 S^* 下的坐标是 $A^T c$. 因此, A^T 是 \mathcal{A}^\dagger 在 T^*, S^* 下的矩阵表示. \square

习题

1. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 \mathbb{F}^n 的基, $\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*$ 是其对偶基,

$$\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in}), \quad \alpha_i^*(x_1, x_2, \cdots, x_n) = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{in}x_n.$$

记 n 阶方阵 $A = (a_{ij})$, $B = (b_{ij})$. 证明: $B = A^{-T}$.

2. 设 $V = \mathbb{R}_n[x]$, $\sigma \in V^*: f(x) \mapsto \int_0^1 f(x)dx$.
- (1) 求 V 的基 $S = \{x^i \mid i = 0, 1, \cdots, n-1\}$ 的对偶基 S^* , 并求 σ 在 S^* 下的坐标.
- (2) 求 V 的基 $T = \{(x-1)^i \mid i = 0, 1, \cdots, n-1\}$ 的对偶基 T^* , 并求 σ 在 T^* 下的坐标.
3. 证明: 在例 9.9 中, $\text{Span}(S) \neq V^*$.
4. 设 S, T 都是 \mathbb{F} 上的有限维线性空间 V 的基, P 是从 S 到 T 的过渡矩阵, S^*, T^* 分别是 S, T 的对偶基. 求从 S^* 到 T^* 的过渡矩阵.
5. 设 S 是 \mathbb{F} 上的有限维线性空间 V 的基, S^* 是 S 的对偶基.
- (1) 设 $\alpha \in V$, 定义 $f_\alpha: V^* \rightarrow \mathbb{F}$, $g \mapsto g(\alpha)$. 证明: $f_\alpha \in V^{**}$. 求 f_α 在 $(S^*, 1)$ 下的矩阵表示.
- (2) 定义自然映射 $\tau: \alpha \mapsto f_\alpha$. 证明: $\tau \in L(V, V^{**})$, 并且 τ 是同构映射.

6. 设 V 是 \mathbb{F} 上的线性空间, S 是 V 的任意非空子集.

$$\text{Ann}(S) = \{f \in V^* \mid f(x) = 0, \forall x \in S\}$$

称为 S 的零化子. 证明:

(1) $\text{Ann}(S)$ 是 V^* 的子空间, 并且 $\text{Ann}(S) = \text{Ann}(\text{Span}(S))$.

(2) 当 V 是有限维线性空间时, $\dim \text{Ann}(S) = \dim V - \text{rank}(S)$.

(3) 对于 V 的任意子空间 V_1, V_2 , 有

$$\text{Ann}(V_1 \cap V_2) = \text{Ann}(V_1) + \text{Ann}(V_2), \quad \text{Ann}(V_1 + V_2) = \text{Ann}(V_1) \cap \text{Ann}(V_2).$$

(4) 若 $V = V_1 \oplus V_2$, 则 $V^* = \text{Ann}(V_1) \oplus \text{Ann}(V_2)$.

§9.4 核空间与像空间

定义 9.9. 设 U, V 都是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(U, V)$. 容易验证,

- $\text{Ker } \mathcal{A} = \mathcal{A}^{-1}(\mathbf{0}) = \{\alpha \in U \mid \mathcal{A}\alpha = \mathbf{0}\}$ 是 U 的子空间.
- $\text{Im } \mathcal{A} = \mathcal{A}(U) = \{\mathcal{A}\alpha \mid \alpha \in U\}$ 是 V 的子空间.

$\text{Ker } \mathcal{A}$ 称为 \mathcal{A} 的核空间, $\dim(\text{Ker } \mathcal{A})$ 称为 \mathcal{A} 的零度, 记作 $\text{null}(\mathcal{A})$.

$\text{Im } \mathcal{A}$ 称为 \mathcal{A} 的像空间, $\dim(\text{Im } \mathcal{A})$ 称为 \mathcal{A} 的秩, 记作 $\text{rank}(\mathcal{A})$.

例 9.11. 设线性映射 $\mathcal{A}: \mathbb{F}^{n \times 1} \rightarrow \mathbb{F}^{m \times 1}$, $x \mapsto Ax$, 其中 $A \in \mathbb{F}^{m \times n}$. 有如下结论.

1. $\text{Ker } \mathcal{A}$ 是线性方程组 $Ax = \mathbf{0}$ 的解空间, $\text{null}(\mathcal{A}) = n - \text{rank}(A)$.
2. $\text{Im } \mathcal{A}$ 是 A 的列向量生成的 \mathbb{F}^m 的子空间, $\text{rank}(\mathcal{A}) = \text{rank}(A)$.

例 9.12. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $f, g \in \mathbb{F}[x]$ 且 $g \mid f$, 则有

$$\text{Ker } g(\mathcal{A}) \subset \text{Ker } f(\mathcal{A}), \quad \text{Im } g(\mathcal{A}) \supset \text{Im } f(\mathcal{A}).$$

例 9.13. 对于任意 $\mathcal{A} \in L(V)$, 有如下结论.

1. $\text{Ker } \mathcal{A}^k \subset \text{Ker } \mathcal{A}^{k+1}$, $\text{Im } \mathcal{A}^k \supset \text{Im } \mathcal{A}^{k+1}$, $\forall k \in \mathbb{N}$.
2. 若 $\text{Ker } \mathcal{A}^{k-1} = \text{Ker } \mathcal{A}^k$, 则 $\text{Ker } \mathcal{A}^k = \text{Ker } \mathcal{A}^{k+1}$.
3. 若 $\text{Im } \mathcal{A}^{k-1} = \text{Im } \mathcal{A}^k$, 则 $\text{Im } \mathcal{A}^k = \text{Im } \mathcal{A}^{k+1}$.

定理 9.10. 设 $\mathcal{A} \in L(U, V)$, 则 $U/\text{Ker } \mathcal{A}$ 与 $\text{Im } \mathcal{A}$ 同构.

证明. 首先, 定义映射 $\rho: U/\text{Ker } \mathcal{A} \rightarrow \text{Im } \mathcal{A}$, $[\alpha] \mapsto \mathcal{A}\alpha$. $[\alpha] = [\beta] \Leftrightarrow \alpha - \beta \in \text{Ker } \mathcal{A} \Leftrightarrow \mathcal{A}\alpha - \mathcal{A}\beta = \mathcal{A}(\alpha - \beta) = \mathbf{0}$. 故 $\rho([\alpha])$ 与 α 的选取无关, ρ 的定义是合理的.

其次, $\rho([\alpha]) = \rho([\beta]) \Leftrightarrow \mathcal{A}\alpha = \mathcal{A}\beta \Leftrightarrow \alpha - \beta \in \text{Ker } \mathcal{A} \Leftrightarrow [\alpha] = [\beta]$. 故 ρ 是单射. 对于任意 $\beta \in \text{Im } \mathcal{A}$, 存在 $\alpha \in V$ 使 $\mathcal{A}\alpha = \beta$, 即 $\rho([\alpha]) = \beta$. 故 ρ 是满射. 因此, ρ 是一一映射.

再次, 对于任意 $[\alpha], [\beta] \in U/\text{Ker } \mathcal{A}$ 和 $\lambda \in \mathbb{F}$, $\rho([\alpha] + [\beta]) = \rho([\alpha + \beta]) = \mathcal{A}(\alpha + \beta) = \mathcal{A}\alpha + \mathcal{A}\beta = \rho([\alpha]) + \rho([\beta])$, $\rho([\lambda\alpha]) = \rho([\lambda\alpha]) = \mathcal{A}(\lambda\alpha) = \lambda\mathcal{A}\alpha = \lambda\rho([\alpha])$.

综上, ρ 是同构映射. □

定理 9.11. 设 U, V 都是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(U, V)$. 证明: 存在 U 的基 S 和 V 的基 T , 使得 \mathcal{A} 在 S, T 下的矩阵表示为 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r = \text{rank}(\mathcal{A})$.

证明. 任取 $\text{Ker } \mathcal{A}$ 的基, 扩充为 U 的基 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, 其中 $\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_n$ 是 $\text{Ker } \mathcal{A}$ 的基. 设 $\beta_i = \mathcal{A}\alpha_i$, $1 \leq i \leq r$, 则 $\beta_1, \beta_2, \dots, \beta_r$ 线性无关; 否则, 存在不全为 0 的 $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{F}$ 使得 $\mathcal{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_r\alpha_r) = \lambda_1\beta_1 + \lambda_2\beta_2 + \dots + \lambda_r\beta_r = \mathbf{0}$, 从而 $\text{Span}(\alpha_1, \alpha_2, \dots, \alpha_r) \cap \text{Ker } \mathcal{A} \neq \{\mathbf{0}\}$, 矛盾. 故 $\beta_1, \beta_2, \dots, \beta_r$ 是 $\text{Im } \mathcal{A}$ 的基, $r = \text{rank}(\mathcal{A})$. 把 $\beta_1, \beta_2, \dots, \beta_r$ 扩充为 V 的基 $T = \{\beta_1, \beta_2, \dots, \beta_n\}$, 则 \mathcal{A} 在 S, T 下的矩阵表示为 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. □

定义 9.10. 设 U_1, V_1 分别是 \mathbb{F} 上的线性空间 U, V 的子空间, $\mathcal{A} \in L(U, V)$, $\mathcal{B} \in L(U_1, V_1)$. 若

$$\mathcal{B}\alpha = \mathcal{A}\alpha, \quad \forall \alpha \in U_1,$$

则 \mathcal{B} 称为 \mathcal{A} 在 (U_1, V_1) 上的限制映射.

例 9.14. 设 $\mathcal{A} \in L(U, V)$, $U = \text{Ker } \mathcal{A} \oplus W$, 则 \mathcal{A} 在 $(W, \text{Im } \mathcal{A})$ 上的限制映射 \mathcal{B} 是同构映射.

证明. 对于任意 $\alpha \in \text{Ker } \mathcal{B}$, $\alpha \in \text{Ker } \mathcal{A} \cap W = \{\mathbf{0}\}$. 故 \mathcal{B} 是单射. 对于任意 $\beta \in \text{Im } \mathcal{A}$, 存在 $\alpha \in V$ 使 $\beta = \mathcal{A}\alpha$. 设 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in W$, $\alpha_2 \in \text{Ker } \mathcal{A}$, 则 $\beta = \mathcal{B}\alpha_1$. 故 \mathcal{B} 是满射. 从而 \mathcal{B} 是一映射. 再由 \mathcal{B} 是线性映射, 得 \mathcal{B} 是同构映射. \square

例 9.15. 设 $\mathcal{A} \in L(V)$ 满足 $\mathcal{A}^2 = \mathcal{A}$, 则 \mathcal{A} 在 $\text{Im } \mathcal{A}$ 上的限制映射是恒等变换, 并且

$$V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A}.$$

证明. 对于任意 $\beta = \mathcal{A}\alpha \in \text{Im } \mathcal{A}$, $\mathcal{A}\beta = \mathcal{A}^2\alpha = \mathcal{A}\alpha = \beta$. 故 \mathcal{A} 在 $\text{Im } \mathcal{A}$ 上的限制映射是恒等变换. 对于任意 $\beta \in \text{Im } \mathcal{A} \cap \text{Ker } \mathcal{A}$, $\beta = \mathcal{A}\beta = \mathbf{0}$. 故 $\text{Im } \mathcal{A} \cap \text{Ker } \mathcal{A} = \{\mathbf{0}\}$. 对于任意 $\alpha \in V$, 设 $\gamma = \alpha - \mathcal{A}\alpha$. 由 $\mathcal{A}\gamma = \mathcal{A}\alpha - \mathcal{A}^2\alpha = \mathbf{0}$, 得 $\gamma \in \text{Ker } \mathcal{A}$, 从而 $\alpha = \mathcal{A}\alpha + \gamma \in \text{Im } \mathcal{A} + \text{Ker } \mathcal{A}$. 根据定理 8.16, $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A}$. \square

例 9.16 (Frobenius 秩不等式). 设 V 是有限维线性空间, $\mathcal{A}, \mathcal{B}, \mathcal{C} \in L(V)$, 则有

$$\text{rank}(\mathcal{A}\mathcal{B}) + \text{rank}(\mathcal{B}\mathcal{C}) \leq \text{rank}(\mathcal{A}\mathcal{B}\mathcal{C}) + \text{rank}(\mathcal{B}).$$

证明. 设 \mathcal{A}_1 是 \mathcal{A} 在 $\text{Im } \mathcal{B}$ 上的限制映射, \mathcal{A}_2 是 \mathcal{A} 在 $\text{Im } \mathcal{B}\mathcal{C}$ 上的限制映射, 得

$$\dim(\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}) = \dim(\text{Ker } \mathcal{A}_1) = \dim(\text{Im } \mathcal{B}) - \dim(\text{Im } \mathcal{A}_1) = \text{rank}(\mathcal{B}) - \text{rank}(\mathcal{A}\mathcal{B})$$

$$\dim(\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}\mathcal{C}) = \dim(\text{Ker } \mathcal{A}_2) = \dim(\text{Im } \mathcal{B}\mathcal{C}) - \dim(\text{Im } \mathcal{A}_2) = \text{rank}(\mathcal{B}\mathcal{C}) - \text{rank}(\mathcal{A}\mathcal{B}\mathcal{C}).$$

再由 $\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}\mathcal{C}$ 是 $\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}$ 的子空间, 得

$$\text{rank}(\mathcal{B}\mathcal{C}) - \text{rank}(\mathcal{A}\mathcal{B}\mathcal{C}) \leq \text{rank}(\mathcal{B}) - \text{rank}(\mathcal{A}\mathcal{B}). \quad \square$$

习题

1. 设 $V = \mathbb{F}^{2 \times 2}$, $\mathcal{A} \in L(V) : X \mapsto AX + XA$, 其中 $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. 分别求 $\text{Ker } \mathcal{A}$ 和 $\text{Im } \mathcal{A}$ 的一个基.
2. 设 $V = \mathbb{F}^{3 \times 3}$, $\mathcal{A} \in L(V) : X \mapsto X + X^T$. 分别求 $\text{Ker } \mathcal{A}$ 和 $\text{Im } \mathcal{A}$ 的一个基.
3. 设 $\mathcal{A} \in L(U, V)$, $W = \text{Ker } \mathcal{A}$, U_1, U_2 都是 U 的有限维子空间, $U_1 \subset U_2$. 证明:

$$\dim(U_2 \cap W) - \dim(U_1 \cap W) \leq \dim U_2 - \dim U_1$$

并求等号成立的充分必要条件.

4. 设 $\mathcal{A} \in L(U, V)$, $\mathcal{B} \in L(V, W)$. 证明:

(1) $\text{Ker } \mathcal{A} \subset \text{Ker } \mathcal{B}\mathcal{A}$, $\text{Im } \mathcal{B}\mathcal{A} \subset \text{Im } \mathcal{B}$.

(2) 若 $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}\mathcal{A}$, 则存在 $\mathcal{C} \in L(W, V)$ 使得 $\mathcal{A} = \mathcal{C}\mathcal{B}\mathcal{A}$.

(3) 若 $\text{Im } \mathcal{B}\mathcal{A} = \text{Im } \mathcal{B}$, 则存在 $\mathcal{C} \in L(V, U)$ 使得 $\mathcal{B} = \mathcal{B}\mathcal{A}\mathcal{C}$.

5. 设 $\mathcal{A} \in L(U, V)$. 证明:

(1) 存在 $\mathcal{B} \in L(V, U)$ 满足 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$. \mathcal{B} 称为 \mathcal{A} 的一个广义逆映射.

(2) \mathcal{B} 是唯一的 $\Leftrightarrow \mathcal{A}$ 是可逆映射.

6. 设 $\mathcal{A} \in L(U, V)$, $\mathcal{B} \in L(V, U)$ 满足 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$. 证明:

(1) $U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{B}$, $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{B}$.

(2) \mathcal{A} 在 $\text{Im } \mathcal{B}$ 上的限制映射与 \mathcal{B} 在 $\text{Im } \mathcal{A}$ 上的限制映射互为逆映射.

7. 设 $\mathcal{A}, \mathcal{B} \in L(U, V)$ 满足 $\text{Im}(\mathcal{A} + \mathcal{B}) = \text{Im } \mathcal{A} \oplus \text{Im } \mathcal{B}$. 证明: 存在 U 的子空间 U_1, U_2, U_3 , 使得 \mathcal{A} 在 $(U_1, \text{Im } \mathcal{A})$ 上的限制映射和 \mathcal{B} 在 $(U_2, \text{Im } \mathcal{B})$ 上的限制映射都是可逆映射, 并且

$$U = U_1 \oplus U_2 \oplus U_3, \quad \text{Ker } \mathcal{A} = U_2 \oplus U_3, \quad \text{Ker } \mathcal{B} = U_1 \oplus U_3.$$

8. 设 $\mathcal{A}, \mathcal{B} \in L(V)$ 满足 $\mathcal{A}^2 = \mathcal{A}$, $\mathcal{B}^2 = \mathcal{B}$. 证明:

(1) $\text{Im } \mathcal{A} = \text{Im } \mathcal{B} \Leftrightarrow \mathcal{A}\mathcal{B} = \mathcal{B}$ 且 $\mathcal{B}\mathcal{A} = \mathcal{A}$.

(2) $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B} \Leftrightarrow \mathcal{A}\mathcal{B} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A} = \mathcal{B}$.

(3) 设 V 是有限维. $\text{rank}(\mathcal{A}) = \text{rank}(\mathcal{B}) \Leftrightarrow$ 存在可逆映射 $\mathcal{C} \in L(V)$ 使得 $\mathcal{A}\mathcal{C} = \mathcal{C}\mathcal{B}$.

9. 设 $\mathcal{A} \in L(V)$. 证明:

(1) $V = \text{Im } \mathcal{A} + \text{Ker } \mathcal{A} \Leftrightarrow \mathcal{A}$ 在 $\text{Im } \mathcal{A}$ 上的限制映射是满射.

(2) $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A} \Leftrightarrow \mathcal{A}$ 在 $\text{Im } \mathcal{A}$ 上的限制映射是可逆映射.

10. 设 $\mathcal{A} \in L(V)$ 满足 $\mathcal{A}^m = \mathcal{O}$, 其中 m 是给定的正整数. \mathcal{A} 称为幂零变换. 证明: 存在 V 的子空间 U 使得 $V = \bigoplus_{i=1}^m \mathcal{A}^{i-1}(U)$.

§9.5 不变子空间

定义 9.11. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$. 若 V 的子空间 U 满足

$$\mathcal{A}(\alpha) \in U, \quad \forall \alpha \in U$$

则 U 称为 \mathcal{A} -不变子空间. 即, \mathcal{A} 在 U 上的限制映射存在当且仅当 U 是 \mathcal{A} -不变子空间.

例 9.17. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$. 容易验证, V 的下列子空间 U 都是 \mathcal{A} -不变子空间.

1. $U = \text{Ker } p(\mathcal{A})$, 其中 $p \in \mathbb{F}[x]$.
2. $U = \text{Im } p(\mathcal{A})$, 其中 $p \in \mathbb{F}[x]$.
3. $U = \bigcap_{i \in I} U_i$, 其中每个 U_i 都是 \mathcal{A} -不变子空间, I 是指标集合.
4. $U = \sum_{i \in I} U_i$, 其中每个 U_i 都是 \mathcal{A} -不变子空间, I 是指标集合.

例 9.18. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in L(V)$ 满足 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, 则 $\text{Ker } \mathcal{B}$ 和 $\text{Im } \mathcal{B}$ 都是 \mathcal{A} -不变子空间.

证明. 对于任意 $\alpha \in \text{Ker } \mathcal{B}$, 由 $\mathcal{B}(\mathcal{A}\alpha) = \mathcal{A}(\mathcal{B}\alpha) = \mathbf{0}$ 得 $\mathcal{A}\alpha \in \text{Ker } \mathcal{B}$. 对于任意 $\alpha \in \text{Im } \mathcal{B}$, 存在 $\beta \in V$ 使得 $\mathcal{A}\alpha = \mathcal{A}p(\mathcal{A})\beta = p(\mathcal{A})\mathcal{B}\beta \in \text{Im } p(\mathcal{A})$. \square

定理 9.12. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $f_1, f_2 \in \mathbb{F}[x]$, $g = \text{gcd}(f_1, f_2)$, $h = \text{lcm}(f_1, f_2)$, 则有

$$\begin{aligned} (1) \text{Ker } g(\mathcal{A}) &= \text{Ker } f_1(\mathcal{A}) \cap \text{Ker } f_2(\mathcal{A}) & (2) \text{Im } g(\mathcal{A}) &= \text{Im } f_1(\mathcal{A}) + \text{Im } f_2(\mathcal{A}) \\ (3) \text{Ker } h(\mathcal{A}) &= \text{Ker } f_1(\mathcal{A}) + \text{Ker } f_2(\mathcal{A}) & (4) \text{Im } h(\mathcal{A}) &= \text{Im } f_1(\mathcal{A}) \cap \text{Im } f_2(\mathcal{A}) \end{aligned}$$

特别,

$$(5) \text{ 当 } g = 1 \text{ 时, } \text{Ker } h(\mathcal{A}) = \text{Ker } f_1(\mathcal{A}) \oplus \text{Ker } f_2(\mathcal{A}).$$

$$(6) \text{ 当 } h(\mathcal{A}) = \mathcal{O} \text{ 时, } \text{Im } g(\mathcal{A}) = \text{Im } f_1(\mathcal{A}) \oplus \text{Im } f_2(\mathcal{A}).$$

证明. 注意到 $f_1 f_2 = gh$ 并且存在 $u_1, u_2 \in \mathbb{F}[x]$ 使得 $g = u_1 f_1 + u_2 f_2$.

$$(1) g \mid f_i \Rightarrow \text{Ker } g(\mathcal{A}) \subset \text{Ker } f_i(\mathcal{A}) \Rightarrow \text{Ker } g(\mathcal{A}) \subset \text{Ker } f_1(\mathcal{A}) \cap \text{Ker } f_2(\mathcal{A}).$$

$$\forall \alpha \in \text{Ker } f_1(\mathcal{A}) \cap \text{Ker } f_2(\mathcal{A}), \quad g(\mathcal{A})\alpha = u_1(\mathcal{A})f_1(\mathcal{A})\alpha + u_2(\mathcal{A})f_2(\mathcal{A})\alpha = \mathbf{0} \Rightarrow \alpha \in \text{Ker } g(\mathcal{A}).$$

$$(2) g \mid f_i \Rightarrow \text{Im } f_i(\mathcal{A}) \subset \text{Im } g(\mathcal{A}) \Rightarrow \text{Im } f_1(\mathcal{A}) + \text{Im } f_2(\mathcal{A}) \subset \text{Im } g(\mathcal{A}).$$

$$\forall \alpha \in \text{Im } g(\mathcal{A}), \quad \alpha = g(\mathcal{A})\beta = f_1(\mathcal{A})u_1(\mathcal{A})\beta + f_2(\mathcal{A})u_2(\mathcal{A})\beta \in \text{Im } f_1(\mathcal{A}) + \text{Im } f_2(\mathcal{A}).$$

$$(3) f_i \mid h \Rightarrow \text{Ker } f_i(\mathcal{A}) \subset \text{Ker } h(\mathcal{A}) \Rightarrow \text{Ker } f_1(\mathcal{A}) + \text{Ker } f_2(\mathcal{A}) \subset \text{Ker } h(\mathcal{A}).$$

$$\forall \alpha \in \text{Ker } h(\mathcal{A}), \quad \alpha = \frac{u_1 f_1}{g}(\mathcal{A})\alpha + \frac{u_2 f_2}{g}(\mathcal{A})\alpha. \quad \beta = \frac{u_1 f_1}{g}(\mathcal{A})\alpha \text{ 满足 } f_2(\mathcal{A})\beta = u_1(\mathcal{A})h(\mathcal{A})\alpha = \mathbf{0},$$

$$\gamma = \frac{u_2 f_2}{g}(\mathcal{A})\alpha \text{ 满足 } f_1(\mathcal{A})\gamma = u_2(\mathcal{A})h(\mathcal{A})\alpha = \mathbf{0}. \text{ 故 } \alpha = \beta + \gamma \in \text{Ker } f_2(\mathcal{A}) + \text{Ker } f_1(\mathcal{A}).$$

(4) $f_i | h \Rightarrow \text{Im } h(\mathcal{A}) \subset \text{Im } f_i(\mathcal{A}) \Rightarrow \text{Im } h(\mathcal{A}) \subset \text{Im } f_1(\mathcal{A}) \cap \text{Im } f_2(\mathcal{A})$.

$\forall \alpha \in \text{Im } f_1(\mathcal{A}) \cap \text{Im } f_2(\mathcal{A})$, 存在 $\alpha_1, \alpha_2 \in V$ 使得 $\alpha = f_1(\mathcal{A})\alpha_1 = f_2(\mathcal{A})\alpha_2$. 故 $\alpha = \frac{u_1 f_1}{g}(\mathcal{A})\alpha + \frac{u_2 f_2}{g}(\mathcal{A})\alpha = \frac{u_1 f_1 f_2}{g}(\mathcal{A})\alpha_2 + \frac{u_2 f_2 f_1}{g}(\mathcal{A})\alpha_1 = h(\mathcal{A})\left(u_1(\mathcal{A})\alpha_2 + u_2(\mathcal{A})\alpha_1\right) \in \text{Im } h(\mathcal{A})$.

(5) 由 (1) 得 $\text{Ker } f_1(\mathcal{A}) \cap \text{Ker } f_2(\mathcal{A}) = \{\mathbf{0}\}$, 再由 (3) 得 $\text{Ker } h(\mathcal{A}) = \text{Ker } f_1(\mathcal{A}) \oplus \text{Ker } f_2(\mathcal{A})$.

(6) 由 (4) 得 $\text{Im } f_1(\mathcal{A}) \cap \text{Im } f_2(\mathcal{A}) = \{\mathbf{0}\}$, 再由 (2) 得 $\text{Im } g(\mathcal{A}) = \text{Im } f_1(\mathcal{A}) \oplus \text{Im } f_2(\mathcal{A})$. \square

定理 9.13. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $f_1, \dots, f_k \in \mathbb{F}[x]$ 两两互素, $k \geq 2$, 则有

$$\text{Ker } \prod_{i=1}^k f_i(\mathcal{A}) = \bigoplus_{i=1}^k \text{Ker } f_i(\mathcal{A}).$$

证明. 当 $k=2$ 时, 由定理 9.12 知结论成立. 当 $k \geq 3$ 时, 由数学归纳法可得

$$\text{Ker } \prod_{i=1}^k f_i(\mathcal{A}) = \left(\text{Ker } \prod_{i=1}^{k-1} f_i(\mathcal{A}) \right) \oplus \text{Ker } f_k(\mathcal{A}) = \bigoplus_{i=1}^k \text{Ker } f_i(\mathcal{A}). \quad \square$$

习题

1. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, U 是 \mathcal{A} -不变子空间. 证明: $\mathcal{A}(U) = \{\mathcal{A}\alpha \mid \alpha \in U\}$ 和 $\mathcal{A}^{-1}(U) = \{\alpha \in V \mid \mathcal{A}\alpha \in U\}$ 都是 \mathcal{A} -不变子空间.
2. 设 $V = \mathbb{F}^2$, $\mathcal{A} \in L(V) : (x, y) \mapsto (x+y, x-y)$. 求所有 1 维 \mathcal{A} -不变子空间.
3. 设 $V = \mathbb{F}^{2 \times 2}$, $\mathcal{A} \in L(V) : X \mapsto AX + XA$, $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. 求所有 1 维 \mathcal{A} -不变子空间.
4. 设 $V = \mathbb{F}^{3 \times 3}$, $\mathcal{A} \in L(V) : X \mapsto X + X^T$. 求所有 k 维 \mathcal{A} -不变子空间, $k = 1, 2, 3$.
5. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$. 证明: 若 V 的任意子空间都是 \mathcal{A} -不变子空间, 则存在 $a \in \mathbb{F}$ 使得 $\mathcal{A} = a\mathcal{I}$.
6. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, W 是 \mathcal{A} -不变子空间, $\mathcal{B} \in L(V/W) : [\alpha] \mapsto [\mathcal{A}\alpha]$. 证明:
 - (1) 对于任意 \mathcal{A} -不变子空间 U , U/W 是 \mathcal{B} -不变子空间.
 - (2) 对于任意 \mathcal{B} -不变子空间 \tilde{U} , 存在 \mathcal{A} -不变子空间 U 使得 $U/W = \tilde{U}$.
7. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in L(V)$, U 是 V 的任意一个 \mathcal{A} -不变子空间. 思考下列问题, 并证明你的结论.
 - (1) 是否一定存在 $p \in \mathbb{F}[x]$ 使得 $U = \text{Ker } p(\mathcal{A})$?
 - (2) 是否一定存在 $p \in \mathbb{F}[x]$ 使得 $U = \text{Im } p(\mathcal{A})$?
 - (3) 是否一定存在 \mathcal{A} -不变子空间 W 使得 $V = U \oplus W$?
 - (4) 若 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, U 是否一定是 \mathcal{B} -不变子空间?
 - (5) 若 \mathcal{A} 可逆, $\mathcal{B} = \mathcal{A}^{-1}$, U 是否一定是 \mathcal{B} -不变子空间?

§9.6 根子空间

定义 9.12. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$. 若存在 $\lambda \in \mathbb{F}$ 和非零向量 $\alpha \in V$, 满足

$$\mathcal{A}\alpha = \lambda\alpha$$

则 λ 称为 \mathcal{A} 的一个特征值, α 称为 λ 对应的一个特征向量.

$$\text{Ker}(\lambda\mathcal{I} - \mathcal{A}) \quad \text{和} \quad \bigcup_{k=1}^{\infty} \text{Ker}(\lambda\mathcal{I} - \mathcal{A})^k$$

分别称为 λ 对应的特征子空间和根子空间.

\mathcal{A} 的特征向量 α 生成的子空间 $\text{Span}(\alpha)$ 是 1 维 \mathcal{A} -不变子空间, \mathcal{A} 的特征子空间和根子空间都是 \mathcal{A} -不变子空间.

当 V 是有限维线性空间时, 设 \mathcal{A} 在 V 的基 S 下的矩阵表示为 A , 则 \mathcal{A} 的特征值与 A 在 \mathbb{F} 中的特征值一一对应, \mathcal{A} 的特征向量与 A 在 \mathbb{F}^n 中的特征向量一一对应. 由于 $\varphi_{\mathcal{A}}(x)$ 与基 S 的选取无关, 可定义 $\varphi_{\mathcal{A}}(x)$ 为 \mathcal{A} 的特征多项式, 记作 $\varphi_{\mathcal{A}}(x)$. 若 $\varphi_{\mathcal{A}}(x)$ 在 \mathbb{F} 中没有根, 则 \mathcal{A} 没有特征值和特征向量.

当 V 是无限维线性空间时, \mathcal{A} 可能没有特征值, 也可能有无穷多个特征值, 例如本节的习题 2.

例 9.19. 设 $V = \mathbb{F}[x]$, $\mathcal{A} \in L(V) : f(x) \mapsto xf(x)$. 对于任意非零多项式 $f(x) \in V$, 不存在常数 $\lambda \in \mathbb{F}$ 使得 $xf(x) = \lambda f(x)$. 因此, \mathcal{A} 没有特征值.

例 9.20. 设 $V = \mathbb{C}^{2 \times 2}$ 是 \mathbb{R} 上的 8 维线性空间, $\mathcal{A} \in L(V) : X \mapsto X^H$. 把 V 分解成 \mathcal{A} 的特征子空间的直和.

解法一. 设 $\lambda \in \mathbb{R}$ 和 $X \in V$ 分别是 \mathcal{A} 的任意一个特征值和相应的特征向量. 由 $\mathcal{A}(X) = \lambda X$ 可得

$$X^H = \lambda X \quad \Rightarrow \quad X = \lambda X^H = \lambda^2 X \quad \Rightarrow \quad \lambda^2 = 1 \quad \Rightarrow \quad \lambda = \pm 1.$$

若 $\lambda = 1$, 则 $X^H = X$, X 是 Hermite 方阵; 若 $\lambda = -1$, 则 $X^H = -X$, X 是反 Hermite 方阵. 因此, 所有 2 阶 Hermite 方阵构成 $\lambda = 1$ 对应的特征子空间 W_1 , 所有 2 阶反 Hermite 方阵构成 $\lambda = -1$ 对应的特征子空间 W_2 . 由于任意 $X \in V$ 都可以唯一地表示为 $X = X_1 + X_2$ 的形式, 其中 $X_1 = \frac{1}{2}(X + X^H) \in W_1$, $X_2 = \frac{1}{2}(X - X^H) \in W_2$, 故 $V = W_1 \oplus W_2$. \square

解法二. \mathcal{A} 在 V 的基 $S = \{E_{11}, iE_{11}, E_{12}, iE_{12}, E_{21}, iE_{21}, E_{22}, iE_{22}\}$ 下的矩阵表示

$$A = \begin{pmatrix} B & O & O & O \\ O & O & B & O \\ O & B & O & O \\ O & O & O & B \end{pmatrix}, \quad \text{其中 } B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

由此可得, $\varphi_{\mathcal{A}}(x) = (x+1)^4(x-1)^4$, $\lambda_1 = -1$ 和 $\lambda_2 = 1$ 是 \mathcal{A} 的特征值. 因此, 任意 2 阶反 Hermite 的非零方阵是 λ_1 对应的特征向量, 任意 2 阶 Hermite 的非零方阵是 λ_2 对应的特征向量. $W_1 = \text{Ker}(\mathcal{A} + \mathcal{I})$ 和 $W_2 = \text{Ker}(\mathcal{A} - \mathcal{I})$ 分别是 λ_1 和 λ_2 对应的特征子空间 (也是根子空间). $\dim W_1 = \dim W_2 = 4$, $V = W_1 \oplus W_2$. \square

解法三. 注意到 $\mathcal{A}^2 = \mathcal{I}$, 故 $p(x) = x^2 - 1 = (x-1)(x+1)$ 满足 $p(\mathcal{A}) = \mathcal{O}$. 根据定理 9.13,

$$V = \text{Ker } p(\mathcal{A}) = \text{Ker}(\mathcal{A} - \mathcal{I}) \oplus \text{Ker}(\mathcal{A} + \mathcal{I})$$

其中 $W_1 = \text{Ker}(\mathcal{A} - \mathcal{I})$ 由 Hermite 方阵构成, $W_2 = \text{Ker}(\mathcal{A} + \mathcal{I})$ 由反 Hermite 方阵构成. \square

定理 9.14. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$. 若 $\varphi_{\mathcal{A}} = \prod_{i=1}^k f_i$, 其中 $f_1, \dots, f_k \in \mathbb{F}[x]$ 两两互素, 则有 $V = \bigoplus_{i=1}^k \text{Ker } f_i(\mathcal{A})$. 特别, 若 $\varphi_{\mathcal{A}} = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 其中 $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ 两两不同, 则有 $V = \bigoplus_{i=1}^k \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i}$. 这称为 V 的根子空间分解.

证明. Cayley-Hamilton 定理和定理 9.13 的推论. \square

设 V, \mathcal{A} 如定理 9.14 所述, $V_i = \text{Ker } f_i(\mathcal{A})$, $d_i = \dim V_i$. 任取 V_i 的基 $S_i = \{\alpha_{i,j} \mid 1 \leq j \leq d_i\}$, 则 $S = \{\alpha_{i,j} \mid 1 \leq i \leq k, 1 \leq j \leq d_i\}$ 构成 V 的基. \mathcal{A} 在 S 下的矩阵表示

$$A = \text{diag}(A_1, A_2, \dots, A_k),$$

其中每个 $A_i \in \mathbb{F}^{d_i \times d_i}$ 是 \mathcal{A} 在 V_i 上的限制映射在 S_i 下的矩阵表示, 满足 $f_i(A_i) = \mathcal{O}$, $\forall i$. 由 f_1, f_2, \dots, f_k 两两互素, 可得 $f_i(A_j)$ 是可逆方阵, 从而 f_i 与 φ_{A_j} 互素, $\forall i \neq j$. 再由 $\prod_{i=1}^k f_i = \varphi_{\mathcal{A}} = \prod_{i=1}^k \varphi_{A_i}$, 不妨设 f_1, f_2, \dots, f_k 都是首一多项式, 可得 $f_i = \varphi_{A_i}$, 从而 $d_i = \deg(f_i)$. 另外, 根据定理 9.13, 对于任意正整数 $m_i \geq n_i$ 都有 $V = \bigoplus_{i=1}^k \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{m_i}$. 因此

$$\text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i} = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i+1} = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i+2} = \dots$$

就是 λ_i 对应的根子空间.

当 V 是无限维线性空间时, 我们无法定义线性变换 \mathcal{A} 的特征多项式, 但是可以通过 \mathcal{A} 的最小多项式 (或 \mathcal{A} 关于某个向量 α 的最小多项式) 求 \mathcal{A} 的特征值和对应的特征向量.

定义 9.13. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$. 若 $f \in \mathbb{F}[x]$ 满足 $f(\mathcal{A}) = \mathcal{O}$, 则 f 称为 \mathcal{A} 的一个化零多项式. \mathcal{A} 的所有化零多项式的最大公因式 (规定是首一多项式) 称为 \mathcal{A} 的最小多项式, 记作 $d_{\mathcal{A}}(x)$.

不是所有线性变换都有最小多项式. 例如: $\mathbb{R}[x]$ 上的微分变换 \mathcal{D} 有特征值, 没有最小多项式; 积分变换 \mathcal{S} 既没有特征值, 也没有最小多项式.

当 $d_{\mathcal{A}}(x)$ 存在时, \mathcal{A} 的所有特征值就是 $d_{\mathcal{A}}(x)$ 在 \mathbb{F} 中的所有根. 留作习题.

定理 9.15. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, f 是 \mathcal{A} 的一个化零多项式. 若 $f = \prod_{i=1}^k f_i$, 其中 $f_1, \dots, f_k \in \mathbb{F}[x]$ 两两互素, 则 $V = \bigoplus_{i=1}^k \text{Ker } f_i(\mathcal{A})$.

证明. 定理 9.13 的推论. \square

例 9.21. 设 $f \in L^1(\mathbb{R})$, 即 f 是 \mathbb{R} 上的复数值连续函数并且 $\int_{-\infty}^{\infty} |f(x)|dx$ 存在, 则

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{ixt} dt$$

也是 \mathbb{R} 上的复数值连续函数. 根据逆变换公式

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) e^{-ixt} dt$$

容易验证, $V = \{f \in L^1(\mathbb{R}) \mid \hat{f} \in L^1(\mathbb{R})\}$ 构成 \mathbb{C} 上的线性空间, 映射 $\mathcal{F}: f \mapsto \hat{f}$ 是 V 上的线性变换. \mathcal{F} 称为 **Fourier**^[2]变换. 例如:

$f(x)$	$e^{-\frac{x^2}{2}}$	$x e^{-\frac{x^2}{2}}$	$(2x^2 - 1) e^{-\frac{x^2}{2}}$	$(2x^3 - 3x) e^{-\frac{x^2}{2}}$
$\hat{f}(x)$	$e^{-\frac{x^2}{2}}$	$i x e^{-\frac{x^2}{2}}$	$-(2x^2 - 1) e^{-\frac{x^2}{2}}$	$-i(2x^3 - 3x) e^{-\frac{x^2}{2}}$

由上表可知, $\lambda_1 = 1, \lambda_2 = i, \lambda_3 = -1, \lambda_4 = -i$ 是 \mathcal{F} 的特征值. 对于任意 $f \in V$, 我们有

$$f(x) \xrightarrow{\mathcal{F}} \hat{f}(x) \xrightarrow{\mathcal{F}} f(-x) \xrightarrow{\mathcal{F}} \hat{f}(-x) \xrightarrow{\mathcal{F}} f(x)$$

即 $\mathcal{F}^4 = \mathcal{I}$. 从而, \mathcal{F} 有最小多项式 $d_{\mathcal{F}}(x) = x^4 - 1$. 任意 $f \in V$ 还可以唯一地表示为

$$f = f_1 + f_2 + f_3 + f_4$$

的形式, 其中

$$\begin{aligned} f_1(x) &= \frac{1}{4}[f(x) + \hat{f}(x) + f(-x) + \hat{f}(-x)] \in \text{Ker}(\lambda_1 \mathcal{I} - \mathcal{A}), \\ f_2(x) &= \frac{1}{4}[f(x) - i\hat{f}(x) - f(-x) + i\hat{f}(-x)] \in \text{Ker}(\lambda_2 \mathcal{I} - \mathcal{A}), \\ f_3(x) &= \frac{1}{4}[f(x) - \hat{f}(x) + f(-x) - \hat{f}(-x)] \in \text{Ker}(\lambda_3 \mathcal{I} - \mathcal{A}), \\ f_4(x) &= \frac{1}{4}[f(x) + i\hat{f}(x) - f(-x) - i\hat{f}(-x)] \in \text{Ker}(\lambda_4 \mathcal{I} - \mathcal{A}). \end{aligned}$$

即 $V = \bigoplus_{i=1}^4 \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})$.

定义 9.14. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha \in V$ 是非零向量. 若 $f \in \mathbb{F}[x]$ 满足 $f(\mathcal{A})\alpha = \mathbf{0}$, 则 f 称为 \mathcal{A} 关于 α 的一个**化零多项式**. \mathcal{A} 关于 α 的所有化零多项式的最大公因式 (规定是首一多项式) 称为 \mathcal{A} 关于 α 的**最小多项式**, 记作 $d_{\mathcal{A},\alpha}(x)$.

$d_{\mathcal{A},\alpha}(x)$ 有可能不存在. 例如: $\mathbb{F}[x]$ 上的积分变换 \mathcal{S} 关于任意非零向量 α 都没有最小多项式. 当 $d_{\mathcal{A},\alpha}(x)$ 存在时, $d_{\mathcal{A},\alpha}(x)$ 在 \mathbb{F} 中的根一定是 \mathcal{A} 的特征值.

例 9.22. 设 $V = \mathbb{F}[x]$. 求 $\mathcal{A} \in L(V): f(x) \mapsto f(x+1)$ 关于任意非零向量 $\alpha \in V$ 的最小多项式.

解答. 设 $\alpha = \sum_{i=0}^k a_i x^i$, $a_0, \dots, a_k \in \mathbb{F}$ 且 $a_k \neq 0$, 则 $\mathcal{A}\alpha = \sum_{i=0}^k a_i (x+1)^i = \sum_{0 \leq j \leq i \leq k} a_i C_i^j x^j$, $\deg(\mathcal{A}\alpha - \alpha) = k-1$. 由 $(\mathcal{A} - \mathcal{I})^k \alpha = a_k k! \neq 0$, $(\mathcal{A} - \mathcal{I})^{k+1} \alpha = 0$, 得 $d_{\mathcal{A},\alpha}(x) = (x-1)^{k+1}$. \square

例 9.22 中的线性变换 \mathcal{A} 有唯一的特征值 $\lambda = 1$, 相应的特征子空间 $\text{Ker}(\lambda \mathcal{I} - \mathcal{A}) = \mathbb{F}$, 根子空间 $\bigcup_{k=1}^{\infty} \text{Ker}(\lambda \mathcal{I} - \mathcal{A})^k = V$, 最小多项式 $d_{\mathcal{A}}(x)$ 不存在.

^[2]Jean-Baptiste Joseph Fourier, 1768-1830, 法国数学家、物理学家.

习题

1. 设 $V = \mathbb{R}_4[x]$, $\mathcal{A} \in L(V) : f(x) \mapsto (1-x)^2 f(0) + 2x(1-x)f(1) + x^2 f(2)$. 求 \mathcal{A} 的所有特征值以及相应的特征子空间和根子空间.
2. 设 \mathbb{R} 上的线性空间 $V = \mathcal{C}^\infty(-\infty, \infty)$ 由所有无穷次可微的实函数构成. 求微分变换 $\mathcal{D} \in L(V)$ 的所有特征值以及相应的特征子空间和根子空间.
3. 设 V 是 \mathbb{F} 上的 n 维线性空间, $\mathcal{A} \in L(V)$. 证明: 下列 4 个叙述相互等价.
 - ① V 可以分解为 n 个 1 维 \mathcal{A} -不变子空间的直和.
 - ② V 可以分解为 \mathcal{A} 的特征子空间的直和.
 - ③ 存在 V 的基 S , 使得 \mathcal{A} 在 S 下的矩阵表示 A 是对角阵.
 - ④ \mathcal{A} 的最小多项式 $d_{\mathcal{A}} = \prod_{i=1}^k (x - \lambda_i)$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}$ 两两不同.
4. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 的最小多项式 $d_{\mathcal{A}}$ 存在, $\lambda \in \mathbb{F}$. 证明: λ 是 \mathcal{A} 的特征值当且仅当 $d_{\mathcal{A}}(\lambda) = 0$.
5. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, I 是指标集合, $\{\lambda_i \mid i \in I\}$ 是 \mathcal{A} 的一组两两不同的特征值, W_i 是 λ_i 对应的根子空间. 证明: $\sum_{i \in I} W_i$ 是直和.
6. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $d_{\mathcal{A}}(x) = \prod_{i=1}^k (x - \lambda_i)$, $\lambda_i \in \mathbb{F}$ 两两不同. 证明: 任意 $\alpha \in V$ 可以表示为 $\alpha = \sum_{i=1}^k \alpha_i$ 的形式, 其中 $\alpha_i = f_i(\mathcal{A})\alpha \in \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})$, $f_i(x) = \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j}$.
7. 设 V 是 \mathbb{C} 上的线性空间, $\mathcal{A} \in L(V)$ 满足 $\mathcal{A}^n = \mathcal{I}$. 证明: 任意 $\alpha \in V$ 可以表示为 $\alpha = \sum_{k=0}^{n-1} \alpha_k$ 的形式, 其中 $\alpha_k = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{-kj} \mathcal{A}^j \alpha \in \text{Ker}(\omega^k \mathcal{I} - \mathcal{A})$, $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$.
8. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $d_{\mathcal{A}}(x) = \prod_{i=1}^k (x - \lambda_i)^{m_i}$, $\lambda_i \in \mathbb{F}$ 两两不同. 证明:
 - (1) $W_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{m_i}$ 是 λ_i 对应的根子空间.
 - (2) $d_i(x) = (x - \lambda_i)^{m_i}$ 是 \mathcal{A} 在 W_i 上的限制映射的最小多项式.
 - (3) 当 W_i 是有限维时, $m_i \leq \dim W_i \leq m_i \dim V_i$, 其中 $V_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})$.
9. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha \in V$ 是非零向量, W 是 \mathcal{A} 的特征值 $\lambda \in \mathbb{F}$ 对应的根子空间. 证明: $\alpha \in W$ 当且仅当 $d_{\mathcal{A}, \alpha}(x)$ 存在并且形如 $(x - \lambda)^{m(\alpha)}$.
10. 设 $V = \mathbb{F}[x]$, $\mathcal{A} \in L(V) : f(x) \mapsto x f'(x)$.
 - (1) 求 \mathcal{A} 的所有特征值以及相应的特征子空间和根子空间.
 - (2) 求 \mathcal{A} 关于任意非零向量 $\alpha \in V$ 的最小多项式.

§9.7 循环子空间

定义 9.15. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha \in V$ 是非零向量.

$$U = \{f(\mathcal{A})\alpha \mid f(x) \in \mathbb{F}[x]\}$$

是包含 α 的最小 \mathcal{A} -不变子空间, 称为 α 生成的 \mathcal{A} -循环子空间 或 Krylov^[3]子空间, 常记作 $U = \mathbb{F}[\mathcal{A}]\alpha$. 特别, 若 $V = \mathbb{F}[\mathcal{A}]\alpha$, 则 \mathcal{A} 称为循环变换, α 称为 \mathcal{A} 的循环向量.

循环子空间与最小多项式这两个概念之间有着紧密的联系.

定理 9.16. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha \in V$ 是非零向量. $U = \mathbb{F}[\mathcal{A}]\alpha$ 是 V 的有限维子空间当且仅当 $d_{\mathcal{A},\alpha}(x)$ 存在, 并且 $\dim U = \deg(d_{\mathcal{A},\alpha})$.

证明. $\dim U = n \Leftrightarrow \{\alpha, \mathcal{A}\alpha, \dots, \mathcal{A}^{n-1}\alpha\}$ 是 U 的基 $\Leftrightarrow \deg(d_{\mathcal{A},\alpha}) = n$. \square

设 $n = \deg(d_{\mathcal{A},\alpha})$, \mathcal{B} 是 \mathcal{A} 在 $U = \mathbb{F}[\mathcal{A}]\alpha$ 上的限制映射, 则 \mathcal{B} 在 U 的基 $\alpha, \mathcal{A}\alpha, \dots, \mathcal{A}^{n-1}\alpha$ 下的矩阵表示是 $d_{\mathcal{A},\alpha}(x)$ 的友方阵.

定理 9.17. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha_1, \dots, \alpha_k \in V$ 都是非零向量, $d_{\mathcal{A},\alpha_1}, \dots, d_{\mathcal{A},\alpha_k}$ 存在并且两两互素, 则 $\alpha = \alpha_1 + \dots + \alpha_k$ 满足 (1) $d_{\mathcal{A},\alpha} = \prod_{i=1}^k d_{\mathcal{A},\alpha_i}$; (2) $\mathbb{F}[\mathcal{A}]\alpha = \bigoplus_{i=1}^k \mathbb{F}[\mathcal{A}]\alpha_i$.

证明. 只需证明 $k=2$ 情形. 记 $d_i = d_{\mathcal{A},\alpha_i}$, $f = d_{\mathcal{A},\alpha}$. 根据 Bézout 定理, 存在 $u, v \in \mathbb{F}[x]$ 使得 $u(x)d_1(x) + v(x)d_2(x) = 1$.

(1) 由 $\mathbf{0} = d_1(\mathcal{A})f(\mathcal{A})\alpha = d_1(\mathcal{A})f(\mathcal{A})\alpha_2$, 得 $d_2 \mid (d_1f)$, 故 $d_2 \mid f$. 同理, $d_1 \mid f$. 故 $(d_1d_2) \mid f$. 另由 $d_1(\mathcal{A})d_2(\mathcal{A})(\alpha_1 + \alpha_2) = \mathbf{0}$, 得 $f \mid (d_1d_2)$. 因此, $f = d_1d_2$.

(2) 易知 $\mathbb{F}[\mathcal{A}]\alpha \subset \mathbb{F}[\mathcal{A}]\alpha_1 + \mathbb{F}[\mathcal{A}]\alpha_2$. 另由

$$\begin{aligned}\alpha_1 &= \left(u(\mathcal{A})d_1(\mathcal{A}) + v(\mathcal{A})d_2(\mathcal{A})\right)\alpha_1 = v(\mathcal{A})d_2(\mathcal{A})\alpha_1 = v(\mathcal{A})d_2(\mathcal{A})\alpha \in \mathbb{F}[\mathcal{A}]\alpha \\ \alpha_2 &= \left(u(\mathcal{A})d_1(\mathcal{A}) + v(\mathcal{A})d_2(\mathcal{A})\right)\alpha_2 = u(\mathcal{A})d_1(\mathcal{A})\alpha_2 = u(\mathcal{A})d_1(\mathcal{A})\alpha \in \mathbb{F}[\mathcal{A}]\alpha\end{aligned}$$

得 $\mathbb{F}[\mathcal{A}]\alpha_1 + \mathbb{F}[\mathcal{A}]\alpha_2 \subset \mathbb{F}[\mathcal{A}]\alpha$. 对于任意 $\beta \in \mathbb{F}[\mathcal{A}]\alpha_1 \cap \mathbb{F}[\mathcal{A}]\alpha_2$, 由 $d_1(\mathcal{A})\beta = d_2(\mathcal{A})\beta = \mathbf{0}$, 得

$$\beta = u(\mathcal{A})d_1(\mathcal{A})\beta + v(\mathcal{A})d_2(\mathcal{A})\beta = \mathbf{0}.$$

故 $\mathbb{F}[\mathcal{A}]\alpha_1 \cap \mathbb{F}[\mathcal{A}]\alpha_2 = \{\mathbf{0}\}$, $\mathbb{F}[\mathcal{A}]\alpha_1 + \mathbb{F}[\mathcal{A}]\alpha_2$ 是直和. 因此, $\mathbb{F}[\mathcal{A}]\alpha = \mathbb{F}[\mathcal{A}]\alpha_1 \oplus \mathbb{F}[\mathcal{A}]\alpha_2$. \square

一般说来, 若干个 \mathcal{A} -循环子空间的和空间 (无论是否直和) 仍然是 \mathcal{A} -不变子空间, 但不一定是 \mathcal{A} -循环子空间. 另一方面, 当 $d_{\mathcal{A},\alpha}$ 可以分解时, 循环子空间 $\mathbb{F}[\mathcal{A}]\alpha$ 也可以相应地分解为若干个 \mathcal{A} -循环子空间的直和.

定理 9.18. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, $\alpha \in V$ 是非零向量, $d_{\mathcal{A},\alpha} = \prod_{i=1}^k f_i$, 其中 $f_1, f_2, \dots, f_k \in \mathbb{F}[x]$ 是两两互素的首一多项式, 则存在非零向量 $\alpha_1, \alpha_2, \dots, \alpha_k \in V$ 满足下列三个条件: (1) $\alpha = \alpha_1 + \dots + \alpha_k$; (2) $d_{\mathcal{A},\alpha_i} = f_i, \forall i$; (3) $\mathbb{F}[\mathcal{A}]\alpha = \bigoplus_{i=1}^k \mathbb{F}[\mathcal{A}]\alpha_i$.

^[3]Alexey Nikolaevich Krylov, 1863–1945, 俄国海军工程师、应用数学家.

证明. 只需证明 $k = 2$ 情形. 根据 Bézout 定理, 存在 $u, v \in \mathbb{F}[x]$ 使得 $u(x)f_1(x) + v(x)f_2(x) = 1$.

(1) 设 $\alpha_1 = v(\mathcal{A})f_2(\mathcal{A})\alpha$, $\alpha_2 = u(\mathcal{A})f_1(\mathcal{A})\alpha$, $d_i = d_{\mathcal{A},\alpha_i}$, $i = 1, 2$. 显然, $\alpha = \alpha_1 + \alpha_2$.

(2) 由 $f_i(\mathcal{A})\alpha_i = \mathbf{0}$ 得 $d_i \mid f_i$. 由 $d_1(\mathcal{A})d_2(\mathcal{A})(\alpha_1 + \alpha_2) = \mathbf{0}$ 得 $(f_1f_2) \mid (d_1d_2)$. 故 $d_i = f_i$.

(3) 根据 (1)、(2) 和定理 9.17, $\mathbb{F}[\mathcal{A}]\alpha = \mathbb{F}[\mathcal{A}]\alpha_1 \oplus \mathbb{F}[\mathcal{A}]\alpha_2$. □

定理 9.19. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 有最小多项式. 则存在非零向量 α 使得 $d_{\mathcal{A},\alpha} = d_{\mathcal{A}}$.

证明. 与定理 5.15 的证明类似. □

定理 9.20. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 有最小多项式. 对于满足 $d_{\mathcal{A},\alpha} = d_{\mathcal{A}}$ 的任意非零向量 α , 存在 \mathcal{A} -不变子空间 U 使得 $V = \mathbb{F}[\mathcal{A}]\alpha \oplus U$.

证明. 设 U 是满足 $U \cap \mathbb{F}[\mathcal{A}]\alpha = \{\mathbf{0}\}$ 的一个极大的 \mathcal{A} -不变子空间, 下证 $V = \mathbb{F}[\mathcal{A}]\alpha \oplus U$. 假设存在 $\beta \notin \mathbb{F}[\mathcal{A}]\alpha \oplus U$. 设 $f \in \mathbb{F}[x]$ 使得 $f(\mathcal{A})\beta \in \mathbb{F}[\mathcal{A}]\alpha \oplus U$ 且 $\deg(f)$ 最小, 则 $f \mid d_{\mathcal{A}}$.

设 $f(\mathcal{A})\beta = g(\mathcal{A})\alpha + u$, 其中 $g \in \mathbb{F}[x]$, $u \in U$. 设 $p = d_{\mathcal{A}}/f$, 由 $p(\mathcal{A})(g(\mathcal{A})\alpha + u) = \mathbf{0}$, $\mathbb{F}[\mathcal{A}]\alpha \cap U = \{\mathbf{0}\}$, 得 $d_{\mathcal{A}} \mid (pg)$, 故 $f \mid g$. 设 $q = g/f$, $\gamma = \beta - q(\mathcal{A})\alpha$, $W = U + \mathbb{F}[\mathcal{A}]\gamma$.

任取 $w = v + h(\mathcal{A})\gamma \in W \cap \mathbb{F}[\mathcal{A}]\alpha$, 其中 $v \in U$, $h \in \mathbb{F}[x]$. 由 $h(\mathcal{A})\beta = w + h(\mathcal{A})q(\mathcal{A})\alpha - v \in \mathbb{F}[\mathcal{A}]\alpha \oplus U$, 得 $f \mid h$. 再由 $f(\mathcal{A})\gamma = u \in U$, 得 $w \in U$. 故 $w \in U \cap \mathbb{F}[\mathcal{A}]\alpha = \{\mathbf{0}\}$.

\mathcal{A} -不变子空间 W 满足 $W \cap \mathbb{F}[\mathcal{A}]\alpha = \{\mathbf{0}\}$, 与 U 的极大性矛盾. □

定理 9.21. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 有最小多项式. 则存在一组向量 $\{\alpha_i \mid i \in I\}$ 使得

$$V = \bigoplus_{i \in I} \mathbb{F}[\mathcal{A}]\alpha_i$$

其中 I 是指标集合. 上式称为 V 的循环子空间分解.

证明. 考虑形如 $\bigoplus_{i \in I} \mathbb{F}[\mathcal{A}]\alpha_i$ 的 \mathcal{A} -不变子空间, 其中每个 α_i 满足 $d_{\mathcal{A},\alpha_i} = d_{\mathcal{A}}$. 设 U_0 是这些子空间在集合的包含关系下的一个极大元. 设 U_1 是满足 $U_1 \cap U_0 = \{\mathbf{0}\}$ 的一个极大的 \mathcal{A} -不变子空间. 下证 $V = U_0 \oplus U_1$.

假设存在 $\beta \notin U_0 \oplus U_1$. 设 $f \in \mathbb{F}[x]$ 使得 $f(\mathcal{A})\beta \in U_0 \oplus U_1$ 并且 $\deg(f)$ 最小, 则 $f \mid d_{\mathcal{A}}$. 设 $p = d_{\mathcal{A}}/f$, $f(\mathcal{A})\beta = u + \sum_{i \in I} g_i(\mathcal{A})\alpha_i$, 其中 $u \in U_1$, $g_i \in \mathbb{F}[x]$ 并且仅有有限多个 $g_i \neq 0$. 由 $p(\mathcal{A})(u + \sum_{i \in I} g_i(\mathcal{A})\alpha_i) = \mathbf{0}$, 得 $d_{\mathcal{A}} \mid pg_i$, 从而 $f \mid g_i$, $\forall i$. 设 $q_i = g_i/f$, $\gamma = \beta - \sum_{i \in I} q_i(\mathcal{A})\alpha_i$, $U_2 = U_1 + \mathbb{F}[\mathcal{A}]\gamma$. 任取 $w \in U_2 \cap U_0$. 设 $w = v + h(\mathcal{A})\gamma$, 其中 $v \in U_1$, $h \in \mathbb{F}[x]$. 由 $h(\mathcal{A})\beta = w + \sum_{i \in I} h(\mathcal{A})q_i(\mathcal{A})\alpha_i - v \in U_0 \oplus U_1$, 得 $f \mid h$. 再由 $f(\mathcal{A})\gamma = u \in U_1$, 得 $w \in U_1$. 故 $w \in U_1 \cap U_0$, $w = \mathbf{0}$. \mathcal{A} -不变子空间 U_2 满足 $U_2 \cap U_0 = \{\mathbf{0}\}$, 与 U_1 的极大性矛盾.

由 U_0 的极大性可知 $d_{\mathcal{A},u} \neq d_{\mathcal{A}}$, $\forall u \in U_1$. 对线性变换的最小多项式的次数应用数学归纳法, 根据归纳假设, U_1 是若干 \mathcal{A} -循环子空间的直和, 从而 V 也是若干 \mathcal{A} -循环子空间的直和. □

利用线性空间的根子空间分解和循环子空间分解, 可以很容易地得到复数方阵的 Jordan 标准形, 以及任意数域 \mathbb{F} 上方阵的相似标准形.

定理 9.22. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$, $\varphi_{\mathcal{A}} = \prod_{i=1}^k p_i^{n_i}$, 其中 $p_1, \dots, p_k \in \mathbb{F}[x]$ 是两两互素的首一不可约多项式, 则有

$$V = \text{Ker } p_1^{n_1}(\mathcal{A}) \oplus \dots \oplus \text{Ker } p_k^{n_k}(\mathcal{A}) \quad (\text{根子空间分解})$$

$$\text{Ker } p_i^{n_i}(\mathcal{A}) = \mathbb{F}[\mathcal{A}]\alpha_{i1} \oplus \dots \oplus \mathbb{F}[\mathcal{A}]\alpha_{it_i} \quad (\text{循环子空间分解})$$

其中 $d_{\mathcal{A}, \alpha_{ij}} = p_i^{m_{ij}}$, 并且 $m_{i1} \geq m_{i2} \geq \dots \geq m_{it_i}$ 由 \mathcal{A} 唯一确定. 具体说来, 恰有

$$\frac{1}{\deg(p_i)} \left(\text{rank}(p_i^{m-1}(\mathcal{A})) - 2\text{rank}(p_i^m(\mathcal{A})) + \text{rank}(p_i^{m+1}(\mathcal{A})) \right)$$

个 α_{ij} 满足 $d_{\mathcal{A}, \alpha_{ij}} = p_i^m$.

对于每个 $V_{ij} = \mathbb{F}[\mathcal{A}]\alpha_{ij}$, 设 \mathcal{B}_{ij} 是 \mathcal{A} 在 V_{ij} 上的限制映射, 则 \mathcal{B}_{ij} 在 V_{ij} 的基

$$\left\{ \underbrace{\alpha_{ij}, \mathcal{A}\alpha_{ij}, \dots, \mathcal{A}^{s-1}\alpha_{ij}}_{\text{第一组}}, \underbrace{p_i(\mathcal{A})\alpha_{ij}, p_i(\mathcal{A})\mathcal{A}\alpha_{ij}, \dots, p_i(\mathcal{A})\mathcal{A}^{s-1}\alpha_{ij}}_{\text{第二组}}, \dots, \right. \\ \left. \underbrace{p_i(\mathcal{A})^{m_{ij}-1}\alpha_{ij}, p_i(\mathcal{A})^{m_{ij}-1}\mathcal{A}\alpha_{ij}, \dots, p_i(\mathcal{A})^{m_{ij}-1}\mathcal{A}^{s-1}\alpha_{ij}}_{\text{第三组}} \right\}$$

下的矩阵表示

$$M_{ij} = \begin{pmatrix} C_i & & & & & \\ E & C_i & & & & \\ & & \ddots & \ddots & & \\ & & & & E & C_i \end{pmatrix}$$

其中 $s = \deg(p_i)$, C_i 是 p_i 的友方阵, E 是右上角元素为 1、其它元素为 0 的基础方阵. 特别, 当 $p_i(x) = x - \lambda_i$ 时, M_{ij}^T 是 Jordan 块方阵 $J_{m_{ij}}(\lambda_i)$.

定理 9.22 与定理 5.18 有异曲同工之效. 定理 9.22 依据的是线性空间 V 的不变子空间分解, 首先把 V 分解成若干根子空间的直和, 然后把每个根子空间分解为若干循环子空间的直和, 最后在每个循环子空间中适当地选取基, 得到相似标准形.

定理 5.18 依据的是特征方阵的模相抵, 首先把特征多项式分解成不变因子的乘积, 然后又分解成初等因子的乘积, 最后得到相似标准形. 这种做法对应于首先把 V 分解成若干循环子空间的直和, 然后把每个循环子空间分解为若干根子空间的直和.

这两种推导 Jordan 标准形的方法各有特点. 不变子空间分解方法对于线性空间的维数没有限制, 同样适用于无限维线性空间. 特征方阵方法对于方阵所在的数域没有限制, 可以根据需要随时调整, 便于得到各种形式的相似标准形.

习题

1. 设 $V = \mathbb{F}[x]$, $\text{Char } \mathbb{F} = 0$, $\mathcal{D}, \mathcal{S} \in L(V)$ 分别是微分变换和积分变换. 证明:
 - (1) 每个非平凡的 \mathcal{D} -不变子空间都是 \mathcal{D} -循环子空间.
 - (2) 每个非平凡的 \mathcal{S} -不变子空间都是 \mathcal{S} -循环子空间.
2. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$, U_1 是 \mathcal{A} -循环子空间, U_2 是 \mathcal{A} -不变子空间.
 - (1) 证明: $U_1 \cap U_2$ 是 \mathcal{A} -循环子空间.

- (2) 举例: $U_1 \oplus U_2$ 不是 \mathcal{A} -循环子空间.
- (3) 举例: 不存在 U_2 使得 $V = U_1 \oplus U_2$.
3. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$. 证明: 下列 4 个叙述相互等价.
- ① \mathcal{A} 是循环变换. ② V 的任意 \mathcal{A} -不变子空间都是 \mathcal{A} -循环子空间.
- ③ $d_{\mathcal{A}} = \varphi_{\mathcal{A}}$. ④ \mathcal{A} 在 V 的某个基下的矩阵表示是友方阵.
4. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$ 是循环变换. 证明: 下列叙述等价.
- ① 任意非零向量 $\alpha \in V$ 都是 \mathcal{A} 的循环向量. ② $\varphi_{\mathcal{A}}(x)$ 在 $\mathbb{F}[x]$ 中不可约.
5. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 是幂零变换. 对 $\deg(d_{\mathcal{A}})$ 应用数学归纳法证明: V 可以分解为若干个 \mathcal{A} -循环子空间的直和.
6. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in L(V)$ 满足 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$. 证明:
- (1) 若 \mathcal{A} 是循环变换, 则存在 $f(x) \in \mathbb{F}[x]$ 使得 $\mathcal{B} = f(\mathcal{A})$.
- (2) 若 U 是 \mathcal{A} -循环子空间, U 是否一定是 \mathcal{B} -不变子空间? 证明你的结论.

第十章 内积空间

一个线性空间 V 是定义了“加法”和“数乘”运算的非空集合. 数域 \mathbb{F} 上的两个线性空间同构当且仅当它们的维数相等. 这种看法是对 V 的结构的一种抽象. V 上的线性变换从侧面更深入地反映了 V 的子空间结构. 在实际应用中, V 可能具有更加丰富的结构, 如度量、测度、拓扑等等. 本章考虑定义了“内积”函数的线性空间, 称为**内积空间**. 利用内积的概念, 我们可以在内积空间上定义度量, 然后引入极限、收敛、开集、闭集、连续、稠密、完备等概念, 使 V 成为一个拓扑空间.

§10.1 基本概念

定义 10.1. 设 V 是 \mathbb{R} 上的线性空间. 若 V 上的二元函数 $\rho: V \times V \rightarrow \mathbb{R}$ 具有下列性质,

1. (对称性) $\rho(\alpha, \beta) = \rho(\beta, \alpha)$
2. (双线性) $\rho(\lambda\alpha + \mu\beta, \gamma) = \lambda\rho(\alpha, \gamma) + \mu\rho(\beta, \gamma)$, $\rho(\alpha, \lambda\beta + \mu\gamma) = \lambda\rho(\alpha, \beta) + \mu\rho(\alpha, \gamma)$
3. (正定性) 当 $\alpha \neq \mathbf{0}$ 时, $\rho(\alpha, \alpha) > 0$

其中 $\alpha, \beta, \gamma \in V$, $\lambda, \mu \in \mathbb{R}$, 则 ρ 称为 V 上的**内积**, 代数结构 (V, ρ) 称为**实内积空间**. 设 U 是 V 的子空间, 则 ρ 可以自然地限制在 U 上, 成为内积 $\tilde{\rho}: U \times U \rightarrow \mathbb{R}$. 实内积空间 $(U, \tilde{\rho})$ 称为实内积空间 (V, ρ) 的**子空间**.

对于 \mathbb{R} 上的一个线性空间 V , 可能存在多个函数 $\rho: V \times V \rightarrow \mathbb{R}$ 满足内积的定义. V 配上不同的内积 ρ 构成不同的实内积空间. 对于一个实内积空间 V , ρ 是取定的. 在不引起歧义的情况下, 可以略去内积名称 ρ , 把 $\rho(\alpha, \beta)$ 记作 (α, β) . 许多教科书也使用记号 $\langle \alpha, \beta \rangle$ 表示内积.

例 10.1. 设 $V = \mathbb{R}^n$. 二元函数 $\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i$ 是 V 上的内积, 称为**标准内积**. \mathbb{R}^n 连同标准内积称为 n 维 **Euclid 空间**.

例 10.2. 设 $V = \mathbb{R}^{m \times n}$. 二元函数 $\rho(X, Y) = \text{tr}(X^T Y)$ 是 V 上的内积.

例 10.3. 设 $V = \mathcal{C}[a, b]$. 二元函数 $\rho(f, g) = \int_a^b f(x)g(x)dx$ 是 V 上的内积.

定义 10.2. 设 (V, ρ) 是 n 维实内积空间, $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 V 的一个基. 对于任意 $\alpha, \beta \in V$, 设 α 和 β 在 S 下的坐标分别是 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 和 $\mathbf{y} = (y_1, y_2, \dots, y_n)$, 则有

$$\rho(\alpha, \beta) = \rho\left(\sum_{i=1}^n x_i \alpha_i, \sum_{j=1}^n y_j \beta_j\right) = \sum_{1 \leq i, j \leq n} x_i y_j \rho(\alpha_i, \alpha_j) = \mathbf{x}^T G \mathbf{y},$$

其中 \mathbf{x}, \mathbf{y} 被视为列向量. $\rho(\alpha, \beta) = \mathbf{x}^T G \mathbf{y}$ 称为 ρ 在 S 下的坐标表示. 方阵 $G = (\rho(\alpha_i, \alpha_j))$ 称为 ρ 在 S 下的度量矩阵或 Gram 方阵. 当 $G = I_n$ 时, S 称为实内积空间 (V, ρ) 的一个标准正交基.

例 10.4. 设 $V = \mathbb{R}^{2 \times 2}$, 内积 $\rho(X, Y) = \text{tr}(X^T Y)$ 在 V 的基 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的度量矩阵 G 是单位方阵 I_4 .

例 10.5. 设 V 是所有 2 阶对称实数方阵构成的 $\mathbb{R}^{2 \times 2}$ 的子空间, 内积 $\rho(X, Y) = \text{tr}(X^T Y)$ 在 V 的

基 $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ 下的度量矩阵 $G = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 3 \end{pmatrix}$.

例 10.6. 设 $V = \mathbb{R}_3[x]$, 内积 $\rho(f, g) = \int_0^1 f(x)g(x)dx$ 在 V 的基 $1, x, x^2$ 下的度量矩阵

$$G = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}.$$

例 10.7. 设 $V = \mathbb{R}_3[x]$, 内积 $\rho(f, g) = \int_0^\infty f(x)g(x)e^{-x} dx$ 在 V 的基 $1, x, x^2$ 下的度量矩阵

$$G = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3! \\ 2 & 3! & 4! \end{pmatrix}.$$

取定线性空间 V 的基 S , 由内积的坐标表示可知, V 上的内积 ρ 由其在 S 下的度量矩阵 G 唯一确定. 显然, 度量矩阵是对称实数方阵. 但是, 并非所有对称实数方阵都可以成为度量矩阵.

定理 10.1. 设 (V, ρ) 是 n 维实内积空间.

1. ρ 在 V 的任意基 S 下的度量矩阵 G 是正定对称方阵.
2. 设 G_i 是 ρ 在 V 的基 S_i 下的度量矩阵, P 是从 S_1 到 S_2 的过渡矩阵, 则 $G_2 = P^T G_1 P$.
3. 对于任意正定对称方阵 $G \in \mathbb{R}^{n \times n}$, 存在 V 的基 S 使得 ρ 在 S 下的度量矩阵是 G .

证明.

1. 由内积的对称性, 知 G 是对称方阵. 再由内积的正定性和坐标表示, 得 $\mathbf{x}^T G \mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0}$. 故 G 是正定对称方阵.
2. 设 $S_2 = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$. $P = (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n)$, 其中 \mathbf{x}_i 是 α_i 在 S_1 下的坐标. 由 ρ 在 S_1 下的坐标表示, 可得 $G_2 = (\rho(\alpha_i, \alpha_j)) = (\mathbf{x}_i^T G_1 \mathbf{x}_j) = P^T G_1 P$.
3. 任取 V 的基 S_1 , 设 G_1 是 ρ 在 S_1 下的度量矩阵. 由于 S_1 和 S 都是 n 阶正定对称方阵, 存在可逆方阵 P 使得 $G = P^T G_1 P$. 取 V 的基 S , 使得 P 是从 S_1 到 S 的过渡矩阵. 由结论 2, G 是 ρ 在 S 下的度量矩阵. \square

定理 10.1 表明, 任意 n 维实内积空间 (V, ρ) 都有标准正交基, ρ 在标准正交基下的坐标表示与 Euclid 空间 \mathbb{R}^n 上的标准内积完全相同. 存在一一映射 $\tau: V \rightarrow \mathbb{R}^n$ 把 V 的标准正交基映成 \mathbb{R}^n 的标准正交基. τ 既保持两个空间的线性运算之间的对应, 也保持内积运算之间的对应. V 与 \mathbb{R}^n 同构. 因此, 许多教科书把有限维实内积空间都称为 Euclid 空间.

定理 10.2. 设 (V, ρ) 是实内积空间, I 是指标集合, $\{\alpha_i \mid i \in I\}$ 是 V 的基. 若 $\beta \in V$ 满足对于任意 $i \in I$ 都有 $\rho(\alpha_i, \beta) = 0$, 则 $\beta = \mathbf{0}$.

证明. $\beta = x_1\alpha_{i_1} + x_2\alpha_{i_2} + \cdots + x_k\alpha_{i_k}$, 其中 $i_1, i_2, \cdots, i_k \in I$, $x_1, x_2, \cdots, x_k \in \mathbb{R}$. 故 $\rho(\beta, \beta) = x_1\rho(\alpha_{i_1}, \beta) + x_2\rho(\alpha_{i_2}, \beta) + \cdots + x_k\rho(\alpha_{i_k}, \beta) = 0$. 因此, $\beta = \mathbf{0}$. \square

定理 10.3 (Cauchy 不等式^[1]). 设 (V, ρ) 是实内积空间. 对于任意 $\alpha, \beta \in V$, 有

$$(\rho(\alpha, \beta))^2 \leq \rho(\alpha, \alpha)\rho(\beta, \beta). \quad (10.1)$$

等号成立当且仅当 α, β 线性相关.

证明. 当 $\alpha = \mathbf{0}$ 时, 结论显然成立. 下设 $\alpha \neq \mathbf{0}$. 取 $\gamma = \beta - \lambda\alpha \in V$, 其中 $\lambda = \frac{\rho(\alpha, \beta)}{\rho(\alpha, \alpha)} \in \mathbb{R}$. 由内积的对称性和双线性,

$$\rho(\gamma, \gamma) = \rho(\beta, \beta) - 2\lambda \cdot \rho(\alpha, \beta) + \lambda^2\rho(\alpha, \alpha) = \rho(\beta, \beta) - \frac{(\rho(\alpha, \beta))^2}{\rho(\alpha, \alpha)}.$$

由内积的正定性 $\rho(\gamma, \gamma) \geq 0$, 可得 $\rho(\alpha, \alpha)\rho(\beta, \beta) \geq (\rho(\alpha, \beta))^2$. 等号成立的充分必要条件是 $\gamma = \mathbf{0}$, 即 $\beta = \lambda\alpha$. \square

在任意实内积空间中, 可以定义一个向量的长度和两个向量之间的夹角, 使空间具有度量结构.

定义 10.3. 设 (V, ρ) 是实内积空间, $\alpha, \beta \in V$.

- $\|\alpha\| = \sqrt{\rho(\alpha, \alpha)}$ 称为 α 的长度. 当 $\|\alpha\| = 1$ 时, α 称为单位向量.
- $\arccos \frac{\rho(\alpha, \beta)}{\|\alpha\|\|\beta\|}$ 称为 α 和 β 的夹角. 当 $\rho(\alpha, \beta) = 0$ 时, 称 α 和 β 正交, 记作 $\alpha \perp \beta$.

使线性空间具有度量结构的方式还有许多. 常见的是引入范数和距离的概念.

定义 10.4. 设 V 是 \mathbb{R} 上的线性空间. 若函数 $p: V \rightarrow \mathbb{R}$ 具有下列性质,

1. (拟线性) $p(\lambda\alpha) = |\lambda| p(\alpha)$
2. (正定性) 当 $\alpha \neq \mathbf{0}$ 时, $p(\alpha) > 0$
3. (三角不等式) $p(\alpha + \beta) \leq p(\alpha) + p(\beta)$

其中 $\alpha, \beta \in V$, $\lambda \in \mathbb{R}$, 则 p 称为 V 上的范数, 代数结构 (V, p) 称为赋范线性空间. 特别, 实内积空间 (V, ρ) 上的函数 $p(\alpha) = \sqrt{\rho(\alpha, \alpha)}$ 构成 V 上的范数, 称为内积 ρ 的导出范数.

例 10.8. 设 $p > 0$ 或 $p = \infty$. 在 $V = \mathbb{R}^n$ 上定义函数

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}, \quad \|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|, \cdots, |x_n|).$$

当 $p \geq 1$ 时, $\|\cdot\|_p$ 满足定义 10.4, 是 V 上的范数, 称为 p -范数. 当 $p \geq 1$ 且 $p \neq 2$ 时, $\|\cdot\|_p$ 不是 V 上任何内积的导出范数.

^[1]Augustin-Louis Cauchy 在 1821 年首先发表了求和形式的 Cauchy 不等式, Viktor Bunyakovsky 在 1859 年证明了不等式的积分形式, 而 Hermann Amandus Schwarz 则在 1888 年独立地证明了不等式的积分形式. 因此, 也被称为 Cauchy-Bunyakovsky-Schwarz 不等式.

类似地, 在 $V = \mathcal{C}[a, b]$ 上定义函数

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{\frac{1}{p}}, \quad \|f\|_\infty = \max_{a \leq x \leq b} |f(x)|.$$

当 $p \geq 1$ 时, $\|\cdot\|_p$ 满足定义 10.4, 是 V 上的范数, 也称为 p -范数.

定义 10.5. 设 V 是一个非空集合. 若函数 $d: V \times V \rightarrow \mathbb{R}$ 具有下列性质,

1. (对称性) $d(x, y) = d(y, x)$
2. (正定性) $d(x, x) = 0$. 当 $x \neq y$ 时, $d(x, y) > 0$
3. (三角不等式) $d(x, z) \leq d(x, y) + d(y, z)$

其中 $x, y, z \in V$, 则 d 称为 V 上的距离, 代数结构 (V, d) 称为距离空间. 特别, 赋范线性空间 (V, p) 上的函数 $d(\alpha, \beta) = p(\alpha - \beta)$ 构成 V 上的距离, 称为范数 p 的导出距离.

例 10.9. $V = \mathbb{R}^n$ 上的二元函数 $d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$ 满足定义 10.5, 是 V 上的距离. 显然, d 不是 V 上任何范数的导出距离.

习题

1. 设 V 是 \mathbb{R} 上的线性空间. 判断下列函数 ρ 是否 V 上的内积, 并说明理由.

- (1) $V = \mathbb{R}^3$, $\rho(\mathbf{x}, \mathbf{y}) = x_1(y_2 + y_3) + x_2(y_1 + y_3) + x_3(y_1 + y_2)$.
- (2) $V = \mathbb{R}^3$, $\rho(\mathbf{x}, \mathbf{y}) = x_1(y_1 + y_2) + x_2(y_2 + y_3) + x_3(y_1 + y_3)$.
- (3) $V = \mathbb{R}^3$, $\rho(\mathbf{x}, \mathbf{y}) = x_1(2y_1 - y_2) + x_2(2y_2 - y_1 - y_3) + x_3(2y_3 - y_2)$.
- (4) $V = \mathbb{R}^{2 \times 2}$, $\rho(X, Y) = \text{tr}(XY)$.
- (5) $V = \mathbb{C}^{2 \times 2}$, $\rho(X, Y) = \text{tr}(X^H Y)$.
- (6) $V = \mathbb{C}^{2 \times 2}$, $\rho(X, Y) = \text{tr}(X^H Y) + \overline{\text{tr}(X^H Y)}$.

2. 设 (V, ρ) 是实内积空间. 计算 ρ 在 V 的基 S 下的度量矩阵.

- (1) $V = \{\mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0\}$, $\rho(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$, $S = \{\mathbf{e}_2 - \mathbf{e}_1, \mathbf{e}_3 - \mathbf{e}_1, \dots, \mathbf{e}_n - \mathbf{e}_1\}$.
- (2) $V = \{X \in \mathbb{R}^{n \times n} \mid \text{tr} X = 0\}$, $\rho(X, Y) = \text{tr}(X^T Y)$,
 $S = \{E_{ii} - E_{11} \mid 2 \leq i \leq n\} \cup \{E_{ij} \mid 1 \leq i \neq j \leq n\}$.
- (3) $V = \{f(x) = \sum_{k=0}^n a_k \cos(kx) + \sum_{k=1}^n b_k \sin(kx) \mid a_k, b_k \in \mathbb{R}\}$, $\rho(f, g) = \int_0^{2\pi} f(x)g(x)dx$,
 $S = \{1, \cos x, \cos(2x), \dots, \cos(nx), \sin x, \sin(2x), \dots, \sin(nx)\}$.
- (4) $V = \mathbb{R}_n[x]$, $\rho(f, g) = \int_0^1 f(x)g(x)|\ln x|dx$, $S = \{1, x, x^2, \dots, x^{n-1}\}$.

3. 证明: 下列方阵 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 是正定方阵, $1 \leq i, j \leq n$.

$$(1) a_{ij} = \frac{1}{i+j} \quad (2) a_{ij} = \frac{1}{(i+j)^2} \quad (3) a_{ij} = \frac{(i+j)!}{i!j!}$$

4. 设 $V = \mathbb{R}^{m \times n}$, $S \in \mathbb{R}^{m \times m}$. 证明: $\rho(X, Y) = \text{tr}(X^T S Y)$ 是 V 上的内积 $\Leftrightarrow S$ 是正定的.
5. 设 $w(x)$ 是定义在 $[0, 1]$ 上的连续函数, 并且 $w(x)$ 不恒为 0.
- (1) 证明: $\rho(f, g) = \int_0^1 f(x)g(x)w(x)dx$ 是 $V = \mathbb{R}[x]$ 上的内积 $\Leftrightarrow w(x) \geq 0, \forall x \in [0, 1]$.
- (2) 对于 $V = \mathbb{R}_n[x]$, 上述结论是否仍然成立? 证明你的结论.
6. 设 (V, ρ) 是实内积空间, $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\} \subset V$, $G = (\rho(\alpha_i, \alpha_j)) \in \mathbb{R}^{n \times n}$. 证明:
- (1) G 是半正定的.
- (2) G 是正定的 $\Leftrightarrow \det G \neq 0 \Leftrightarrow S$ 是线性无关的.
7. 设 V 是实内积空间. 证明: 对于任意 $\alpha, \beta \in V$, 有
- (1) (勾股定理) 若 $\alpha \perp \beta$, 则 $\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$.
- (2) (余弦定理) $\|\alpha - \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 - 2\|\alpha\|\|\beta\|\cos\theta$, 其中 θ 是 α 和 β 的夹角.
- (3) (三角不等式) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.
- (4) (菱形对角线性性质) 若 $\|\alpha\| = \|\beta\|$, 则 $(\alpha + \beta) \perp (\alpha - \beta)$.
- (5) (平行四边形法则) $\|\alpha + \beta\|^2 + \|\alpha - \beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2$.
- (6) (Apollonius 等式) $\|\alpha - \gamma\|^2 + \|\beta - \gamma\|^2 = \frac{1}{2}\|\alpha - \beta\|^2 + \frac{1}{2}\|\alpha + \beta - 2\gamma\|^2$.
- (7) (极化恒等式) $\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2 = 4(\alpha, \beta)$.
8. 设 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 n 维实内积空间 (V, ρ) 中的一组非零向量. 证明:
- (1) 若对于任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) < 0$, 则 $m \leq n + 1$.
- (2) 若对于任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) \leq 0$, 则 $m \leq 2n$.
9. 设 $V = \mathbb{R}^n$, $\|\cdot\|_p$ 是例 10.8 中的函数. 证明: 当 $0 < p < 1$ 时, $\|\cdot\|_p$ 不是 V 上的范数.

§10.2 标准正交基

定义 10.6. 设 S 是实内积空间 V 中的一个向量组.

- 若 S 中任意两个向量相互正交, 则 S 称为**正交向量组**.
- 若 S 是正交向量组且 S 中每个向量的长度都是 1, 则 S 称为**标准正交向量组**.
- 若 S 是 V 的基且是正交向量组, 则 S 称为 V 的一个**正交基**.
- 若 S 是 V 的基且是标准正交向量组, 则 S 称为 V 的一个**标准正交基**.

对于给定的实内积空间 V , 我们并不知道 V 是否有标准正交基. 自然地, 我们提出如下问题:

1. 什么样的实内积空间有标准正交基? 是否有简便的判定方法?
2. 当标准正交基存在时, 如何构造出标准正交基? 不同的标准正交基之间有何联系?
3. 当标准正交基不存在时, 能够找到标准正交基的替代? 如何描述实内积空间的度量结构?

借助度量矩阵, 定理 10.1 证明了有限维实内积空间的标准正交基的存在性. 事实上, 实内积空间 V 中任意一组线性无关的向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 都可以通过如下 **Gram-Schmidt 标准正交化过程**, 改造成为一个标准正交向量组 $\gamma_1, \gamma_2, \dots, \gamma_n$.

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\gamma_i, \alpha_k) \gamma_i, \quad \gamma_k = \frac{1}{\|\beta_k\|} \beta_k, \quad k = 1, 2, \dots, n. \quad (10.2)$$

上述过程可以看作是 (6.2) 和 (6.3) 的推广. 特别, 当 V 是可数维时, V 的任意一组基 $\{\alpha_n \mid n \in \mathbb{N}\}$ 也可以通过 Gram-Schmidt 标准正交化过程, 改造成为一个标准正交向量组 $\{\gamma_n \mid n \in \mathbb{N}\}$. 由每个 α_k 是 $\gamma_1, \dots, \gamma_k$ 的线性组合, 可知 $\{\gamma_n \mid n \in \mathbb{N}\}$ 是 V 的一个标准正交基. 因此, 当 $\dim V \leq \aleph_0$ 时, V 必有标准正交基.

例 10.10. 设 $V = \mathbb{R}_3[x]$, 内积 $(f, g) = \int_0^1 f(x)g(x)dx$. 求 V 的标准正交基.

解法一. 对基 $1, x, x^2$ 作 Gram-Schmidt 标准正交化.

$$\begin{aligned} \beta_1 &= 1, & \|\beta_1\| &= 1, & \gamma_1 &= 1, \\ \beta_2 &= x - \frac{1}{2}\gamma_1 = x - \frac{1}{2}, & \|\beta_2\| &= \frac{1}{\sqrt{12}}, & \gamma_2 &= \sqrt{12}(x - \frac{1}{2}), \\ \beta_3 &= x^2 - \frac{1}{3}\gamma_1 - \frac{\sqrt{3}}{6}\gamma_2 = x^2 - x + \frac{1}{6}, & \|\beta_3\| &= \frac{1}{\sqrt{180}}, & \gamma_3 &= \sqrt{180}(x^2 - x + \frac{1}{6}). \end{aligned}$$

$\gamma_1, \gamma_2, \gamma_3$ 是 V 的标准正交基. □

解法二. 对基 $1, x, x^2$ 下的度量矩阵作相合变换.

$$\begin{pmatrix} 1 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & \frac{1}{12} & \frac{1}{12} \\ -\frac{1}{3} & 0 & 1 & 0 & \frac{1}{12} & \frac{4}{45} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & \frac{1}{12} & 0 \\ \frac{1}{6} & -1 & 1 & 0 & 0 & \frac{1}{180} \end{pmatrix}.$$

由此可得, $e_1 = 1, e_2 = \sqrt{12}(x - \frac{1}{2}), e_3 = \sqrt{180}(x^2 - x + \frac{1}{6})$ 是 V 的标准正交基. □

解法三. 设 $f_i = \frac{d^i}{dx^i}[x^i(1-x)^i]$, 由分部积分可得 $\int_0^1 x^j f_i(x) dx = 0, \forall 0 \leq j < i \leq 2$.

$$\begin{aligned} f_0 &= 1, & f_1 &= (x-x^2)' = 1-2x, & f_2 &= (x^2-2x^3+x^4)'' = 2-12x+12x^2, \\ g_0 &= \frac{f_0}{\|f_0\|} = 1, & g_1 &= \frac{f_1}{\|f_1\|} = \sqrt{12}(x-\frac{1}{2}), & g_2 &= \frac{f_2}{\|f_2\|} = \sqrt{180}(x^2-x+\frac{1}{6}). \end{aligned}$$

g_0, g_1, g_2 是 V 的标准正交基. □

 以上三种方法均可以求得 $\mathbb{R}_3[x]$ 的一个标准正交基. 解法一和解法二都是通用方法, 解法三是特殊方法. 解法一的计算量较大, 解法二的计算量适中, 并且容易编程实现.

定理 10.4. 设 V 是有限维实内积空间, S_1, S_2 是 V 的两个标准正交基, 则从 S_1 到 S_2 的过渡矩阵 P 是正交方阵. 对于任意 $A \in L(V)$, 设 A 在 S_1, S_2 下的矩阵表示分别是 A_1, A_2 , 则 $A_2 = P^{-1}A_1P$, 即 A_1 与 A_2 正交相似.

证明. 定理 9.3 和定理 10.1 (2) 的推论. □

在线性空间中, “极大线性无关组”与“线性空间的基”是同义词. 在实内积空间 V 中, 标准正交基一定是(在集合的包含关系下)极大的标准正交向量组, 反之不一定成立. 当 V 是有限维时, 极大标准正交向量组一定是标准正交基. 当 V 是无限维时, 极大标准正交向量组一定存在, 标准正交基有可能不存在. 当 V 的标准正交基不存在时, 我们试图通过具有标准正交基的子空间来研究 V 的结构.

例 10.11. 设 V 是满足 $\sum_{n=0}^{\infty} a_n^2$ 收敛的实数数列 $\{a_n\}_{n \in \mathbb{N}}$ 的全体. 在数列的加法和数乘运算下, V 构成实线性空间. $\rho(\{a_n\}, \{b_n\}) = \sum_{n=0}^{\infty} a_n b_n$ 是 V 上内积. 实内积空间 (V, ρ) 通常称为 l^2 空间.

易知, 当 $0 < q < 1$ 时, 等比数列 $\alpha_q = \{q^n\}_{n \in \mathbb{N}} \in V$. 对于任意两两不同的 $q_1, \dots, q_k \in (0, 1)$, $\alpha_{q_1}, \dots, \alpha_{q_k}$ 是线性无关的. 因此, $\{\alpha_q \mid 0 < q < 1\}$ 是线性无关的, $\dim V$ 是不可数的.

设数列 e_n 的第 n 项是 1, 其它项是 0, $n \in \mathbb{N}$. 易知, $S = \{e_n \mid n \in \mathbb{N}\}$ 是 V 的一个标准正交向量组. 若 $\{a_n\}_{n \in \mathbb{N}} \in V$ 与每个 e_n 都正交, 则 $a_n = 0, \forall n$. 因此, S 是极大标准正交向量组.

假设 V 有标准正交基 $T = \{\alpha_i \mid i \in I\}$, 其中 I 是不可数的指标集合. 由于每个 e_n 都是 T 中有限多个向量的线性组合, 故存在可数子集 $J \subset I$ 使得 $S \subset \text{Span}(\{\alpha_j \mid j \in J\})$. 设 $i \in I \setminus J$. 由 $\alpha_i \perp \alpha_j, \forall j \in J$ 可得 $\alpha_i \perp e_n, \forall n \in \mathbb{N}$. 故 $\alpha_i = \mathbf{0}$. 矛盾. 因此, V 没有标准正交基.

定义 10.7. 设 V 是实内积空间, $S \subset V$. 容易验证,

$$S^\perp = \{\alpha \in V \mid \alpha \perp \beta, \forall \beta \in S\}$$

是 V 的子空间. S^\perp 称为 $\text{Span}(S)$ 的正交补空间.

容易证明, $\{\mathbf{0}\}^\perp = V, V^\perp = \{\mathbf{0}\}, S^\perp \cap \text{Span}(S) = \{\mathbf{0}\}$. 从而, $S^\perp + \text{Span}(S)$ 是直和.

例 10.12. 设 $V = \mathbb{R}^{n \times n}$, 内积 $(X, Y) = \text{tr}(X^T Y)$. V 的子空间 $U = \{X \in V \mid X^T = X\}$ 和 $W = \{X \in V \mid X^T = -X\}$ 互为正交补空间.

证明. $A = (a_{ij}) \in U^\perp \Leftrightarrow (A, E_{ij} + E_{ji}) = 0, \forall i, j \Leftrightarrow a_{ji} + a_{ij} = 0, \forall i, j \Leftrightarrow A^T = -A$.

$$B = (b_{ij}) \in W^\perp \Leftrightarrow (B, E_{ij} - E_{ji}) = 0, \forall i, j \Leftrightarrow b_{ji} - b_{ij} = 0, \forall i, j \Leftrightarrow B^T = B. \quad \square$$

定理 10.5. 设 S 是实内积空间 V 中的正交向量组. $S^\perp = \{\mathbf{0}\} \Leftrightarrow S$ 是极大正交向量组.

证明. (\Leftarrow) 若 $S^\perp \neq \{\mathbf{0}\}$, 则存在非零向量 $\alpha \in S^\perp$, $S \cup \{\alpha\}$ 也是正交向量组, 与 S 的极大性矛盾.

(\Rightarrow) 若存在正交向量组 $T \not\subseteq S$, 则 $T \setminus S \subset S^\perp$, 与 $S^\perp = \{\mathbf{0}\}$ 矛盾. \square

定理 10.6. 设 U 是实内积空间 V 的子空间.

1. 若 U 是有限维的, 则 $V = U \oplus U^\perp$.

2. 若 $V = U \oplus U^\perp$, 则 $(U^\perp)^\perp = U$.

证明.

1. 取 U 的标准正交基 e_1, e_2, \dots, e_n . 对于任意 $\alpha \in V$, $\beta = \sum_{i=1}^n (e_i, \alpha) e_i \in U$, $\alpha - \beta \in U^\perp$. 故 $V = U + U^\perp$. 再由 $U \cap U^\perp = \{\mathbf{0}\}$, 得 $V = U \oplus U^\perp$.

2. 一方面, 对于任意 $\alpha \in U$, 由 $\alpha \perp U^\perp$ 可得 $\alpha \in (U^\perp)^\perp$. 故 $U \subset (U^\perp)^\perp$. 另一方面, 对于任意 $\alpha \in (U^\perp)^\perp$, 设 $\alpha = u + w$, 其中 $u \in U$, $w \in U^\perp$. 由 $\alpha \perp w$ 且 $u \perp w$ 可得 $w = \mathbf{0}$, 从而 $\alpha \in U$. 故 $(U^\perp)^\perp \subset U$. 综上, $(U^\perp)^\perp = U$. \square

 在定理 10.6 中, 结论 1 的条件“ U 是有限维的”不可省略. 例如, 在例 10.11 中, $U = \text{Span}(S)$ 满足 $U^\perp = \{\mathbf{0}\}$, $V \neq U \oplus U^\perp$, $(U^\perp)^\perp \neq U$. 结论 2 的条件 $V = U \oplus U^\perp$ 也不是必要的.

例 10.13. 设 V 是实内积空间. 给定 $\alpha_1, \alpha_2, \dots, \alpha_n, \alpha \in V$, 求 $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ 使

$$f(\mathbf{x}) = \left\| \alpha - \sum_{i=1}^n x_i \alpha_i \right\|^2$$

取得最小值. 本问题是例 6.9 最小二乘问题的推广.

解答. 设 $U = \text{Span}(\alpha_1, \alpha_2, \dots, \alpha_n)$. 根据定理 10.6, α 可以唯一地表示为 $\beta + \gamma$ 的形式, 其中 $\beta \in U$, $\gamma \in U^\perp$. 从而, $f(\mathbf{x}) = \left\| \beta - \sum_{i=1}^n x_i \alpha_i \right\|^2 + \|\gamma\|^2$. 当且仅当 $\beta = \sum_{i=1}^n x_i \alpha_i$ 时, $f(\mathbf{x})$ 取得最小值 $\|\gamma\|^2$.

注意到 $\beta = \sum_{i=1}^n x_i \alpha_i$ 等价于 $\alpha - \sum_{i=1}^n x_i \alpha_i \in U^\perp$, 即 $(\alpha_i, \alpha - \sum_{i=1}^n x_i \alpha_i) = 0, \forall i$. 由此得线性方程组

$$\begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_n) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_n) \\ \vdots & \vdots & \cdots & \vdots \\ (\alpha_n, \alpha_1) & (\alpha_n, \alpha_2) & \cdots & (\alpha_n, \alpha_n) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} (\alpha_1, \alpha) \\ (\alpha_2, \alpha) \\ \vdots \\ (\alpha_n, \alpha) \end{pmatrix}.$$

上述线性方程组的任意解 \mathbf{x} 都是 $f(\mathbf{x})$ 的最小值点. \square

定理 10.7. 设 V 是实内积空间, I 是指标集合, $\{e_i \mid i \in I\}$ 是 V 的一个标准正交基.

1. (Parseval^[2]等式) 对于任意 $\alpha, \beta \in V$, $(\alpha, \beta) = \sum_{i \in I} (e_i, \alpha)(e_i, \beta)$.

2. (Bessel^[3]等式) 对于任意 $\alpha \in V$, $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$.

^[2]Marc-Antoine Parseval, 1755–1836, 法国数学家.

^[3]Friedrich Wilhelm Bessel, 1784–1846, 德国天文学家、数学家、物理学家.

证明. 对于任意 $\alpha, \beta \in V$, 存在有限子集 $J \subset I$ 使得 $\alpha = \sum_{j \in J} x_j e_j$, $\beta = \sum_{j \in J} y_j e_j$, 其中 $x_j, y_j \in \mathbb{R}$.

$$(e_i, \alpha) = \begin{cases} x_i, & i \in J; \\ 0, & i \notin J, \end{cases} \quad (e_j, \beta) = \begin{cases} y_j, & j \in J; \\ 0, & j \notin J. \end{cases}$$

因此, $(\alpha, \beta) = \sum_{j \in J} x_j y_j = \sum_{i \in I} (e_i, \alpha)(e_i, \beta)$. 特别, 当 $\alpha = \beta$ 时, $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$. \square

定理 10.8. 设 V 是实内积空间, I 是指标集合, $S = \{e_i \mid i \in I\}$ 是 V 中的标准正交向量组.

1. (Bessel 不等式) 对于任意 $\alpha \in V$, $\|\alpha\|^2 \geq \sum_{i \in I} (e_i, \alpha)^2$.
2. 对于任意 $\alpha \in V$ 和 $\beta \in \text{Span}(S)$, $\|\alpha - \beta\|^2 \geq \|\alpha\|^2 - \sum_{i \in I} (e_i, \alpha)^2$.
3. 对于任意 $\alpha \in V$, $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2 \Leftrightarrow$ 存在一列 $\beta_n \in \text{Span}(S)$ 满足 $\lim_{n \rightarrow \infty} \|\alpha - \beta_n\| = 0$.

证明.

1. 对于任意有限子集 $J \subset I$, 设 $\beta = \sum_{j \in J} (e_j, \alpha) e_j$. 由 $\alpha - \beta \perp \{e_j \mid j \in J\}$ 可得 $(\beta, \alpha - \beta) = 0$. 从而, $\|\alpha\|^2 = \|\beta\|^2 + \|\alpha - \beta\|^2 \geq \|\beta\|^2 = \sum_{j \in J} (e_j, \alpha)^2$. 因此, $\|\alpha\|^2 \geq \sum_{i \in I} (e_i, \alpha)^2$.
2. 设 $\beta = \sum_{j \in J} x_j e_j$, 其中 J 是 I 的有限子集. 记 $\gamma = \sum_{j \in J} (e_j, \alpha) e_j$. 由 $\alpha - \gamma \perp \{e_j \mid j \in J\}$ 可得 $\|\alpha - \beta\|^2 = \|\alpha - \gamma\|^2 + \|\beta - \gamma\|^2 \geq \|\alpha - \gamma\|^2 = \|\alpha\|^2 - \sum_{j \in J} (e_j, \alpha)^2 \geq \|\alpha\|^2 - \sum_{i \in I} (e_i, \alpha)^2$.
3. 一方面, 若 $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$, 则存在一列有限子集 $J_n \subset I$ 使得 $\lim_{n \rightarrow \infty} \sum_{i \in J_n} (e_i, \alpha)^2 = \|\alpha\|^2$. 故 $\beta_n = \sum_{j \in J_n} (e_j, \alpha) e_j$ 满足 $\lim_{n \rightarrow \infty} \|\alpha - \beta_n\| = 0$. 另一方面, 若 $\beta_n \in \text{Span}(S)$ 满足 $\lim_{n \rightarrow \infty} \|\alpha - \beta_n\| = 0$, 则 $J_n = \{i \in I \mid (e_i, \beta_n) \neq 0\}$ 使得 $\|\alpha\|^2 - \sum_{i \in J_n} (e_i, \alpha)^2 \leq \|\alpha - \beta_n\|^2$. 故 $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$. \square

 设 U 是实内积空间 V 的无限维子空间. 对于任意 $\alpha \in V$, 存在一系列 $\beta_n \in U$, $n \in \mathbb{N}$ 使得 $\|\beta_n - \alpha\|$ 单调递减. 但是, 这并不能说明存在 $\beta \in U$ 使得 $\|\beta - \alpha\|$ 取得最小值.

设 (V, d) 是距离空间, $\{\alpha_n\}_{n \in \mathbb{N}}$ 是 V 中的点列. 若存在 $\alpha \in V$ 使得 $\lim_{n \rightarrow \infty} d(\alpha_n, \alpha) = 0$, 则称 α 是 $\{\alpha_n\}_{n \in \mathbb{N}}$ 的极限, 或者称 $\{\alpha_n\}_{n \in \mathbb{N}}$ 收敛到 α , 记作 $\lim_{n \rightarrow \infty} \alpha_n = \alpha$. 由三角不等式可知, 若 $\{\alpha_n\}_{n \in \mathbb{N}}$ 在 V 中收敛, 则 $\{\alpha_n\}_{n \in \mathbb{N}}$ 满足如下的 **Cauchy 准则**:

对于任意 $\varepsilon > 0$, 存在 N 使得 $d(\alpha_m, \alpha_n) < \varepsilon$, $\forall m > n > N$.

若 V 中满足 Cauchy 准则的点列都在 V 中收敛, 则称 V 是完备的. 完备的赋范线性空间 (V, p) 称为 **Banach**^[4]空间, 完备的实内积空间 (V, ρ) 称为 **Hilbert**^[5]空间.

例如, 所有有限维实内积空间都是 Hilbert 空间. 在例 10.11 中, V 是 Hilbert 空间, 没有标准正交基; V 的子空间 $\text{Span}(S)$ 不是 Hilbert 空间, 有标准正交基.

^[4]Stefan Banach, 1892–1945, 波兰数学家, 泛函分析创始人之一.

^[5]David Hilbert, 1862–1943, 德国数学家. 1900 年 8 月提出的 23 个 Hilbert 问题对现代数学的研究和发展产生了深刻的影响.

习题

1. 设 V 分别是例 10.5 和例 10.7 中的实内积空间. 求 V 的一个标准正交基.
2. 分别求 $V = \mathbb{R}[x]$ 上的一个内积 ρ , 使得 S 构成 (V, ρ) 的正交基.
 - (1) $S = \{P_n(x) \mid n \in \mathbb{N}\}$, 其中 $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]$ 是 Legendre^[6] 多项式.
 - (2) $S = \{T_n(x) \mid n \in \mathbb{N}\}$, 其中 $T_n(\cos \theta) = \cos(n\theta)$ 是第一类 Chebyshev^[7] 多项式.
 - (3) $S = \{U_n(x) \mid n \in \mathbb{N}\}$, 其中 $U_n(\cos \theta) = \frac{\sin((n+1)\theta)}{\sin \theta}$ 是第二类 Chebyshev 多项式.
3. 设 S 是有限维实内积空间 V 的标准正交基. 证明: 向量组 T 是 V 的标准正交基 \Leftrightarrow 从 S 到 T 的过渡矩阵是正交方阵.
4. 设实内积空间 $V = \mathbb{R}^{2 \times 2}$, 内积 $(X, Y) = \text{tr}(X^T Y)$, $\mathcal{A} \in L(V) : X \mapsto AX$, 其中 $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. 分别求 $\text{Ker } \mathcal{A}$, $(\text{Ker } \mathcal{A})^\perp$, $\text{Im } \mathcal{A}$, $(\text{Im } \mathcal{A})^\perp$ 的一个标准正交基.
5. 设 $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{k \times n}$, $\alpha \in \mathbb{R}^{m \times 1}$, $x \in \mathbb{R}^{n \times 1}$. 求 $\|Ax - \alpha\|$ 在条件 $Bx = \mathbf{0}$ 下的最小值.
6. 证明: 在实内积空间中, 不含零向量的正交向量组一定是线性无关的.
7. 举例: 实内积空间 V 有标准正交基, 并且 $\dim V$ 是不可数的.
8. 举例: U 是实内积空间 V 的子空间, $U = (U^\perp)^\perp$, $V \neq U \oplus U^\perp$.
9. 设 U_1, U_2 是实内积空间 V 的子空间. 证明:
 - (1) $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$, $(U_1 \cap U_2)^\perp \supset U_1^\perp + U_2^\perp$.
 - (2) 若 $U_1 + U_2$ 是有限维, 则 $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.
 - (3) 举例: $(U_1 \cap U_2)^\perp \neq U_1^\perp + U_2^\perp$.
10. 设 V 是实内积空间, I 是指标集合, $\{e_i \mid i \in I\}$ 是 V 的极大标准正交向量组. 是否对于任意 $\alpha \in V$, 都有 $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$? 证明你的结论.

^[6]Adrien-Marie Legendre, 1752–1833, 法国数学家.

^[7]Pafnuty Lvovich Chebyshev, 1821–1894, 俄国数学家.

§10.3 正交变换

定义 10.8. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 可逆. 若 \mathcal{A} 保持任意两个向量的内积不变, 即

$$(\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta), \quad \forall \alpha, \beta \in V$$

则 \mathcal{A} 称为 V 上的正交变换.

许多教科书也把以下定理作为正交变换的一个等价的定义.

定理 10.9. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 可逆. 若 \mathcal{A} 保持每个向量的长度不变, 即

$$\|\mathcal{A}\alpha\| = \|\alpha\|, \quad \forall \alpha \in V$$

则 \mathcal{A} 是正交变换.

证明. 对于任意 $\alpha, \beta \in V$, 由余弦定理,

$$(\mathcal{A}\alpha, \mathcal{A}\beta) = \frac{1}{2} \left(\|\mathcal{A}\alpha + \mathcal{A}\beta\|^2 - \|\mathcal{A}\alpha\|^2 - \|\mathcal{A}\beta\|^2 \right) = \frac{1}{2} \left(\|\alpha + \beta\|^2 - \|\alpha\|^2 - \|\beta\|^2 \right) = (\alpha, \beta).$$

因此, \mathcal{A} 是正交变换. □

例 10.14. 设 V 是实内积空间. 任取非零向量 $v \in V$. v 称为 $\pi = \{v\}^\perp$ 的法向量.

- 恒等变换 $\mathcal{I}: \alpha \mapsto \alpha$ 显然是 V 上的正交变换.
- 线性变换 $\mathcal{A}: \alpha \mapsto 2\frac{(v, \alpha)}{(v, v)}v - \alpha$ 把每个点 α 映射到其关于 v 所在直线的对称点. \mathcal{A} 是 V 上的正交变换, 称为关于对称轴 v 的轴对称变换.
- 线性变换 $\mathcal{B} = -\mathcal{A}$ 把每个点 α 被映射到其关于 π 的对称点. \mathcal{B} 也是 V 上的正交变换, 称为关于法向量 v 的镜面反射变换.

例 10.15. 设 $V = \left\{ \sum_{i=-m}^n a_i x^i \mid a_i \in \mathbb{R}, m, n \in \mathbb{N} \right\}$, 内积 $\left(\sum_i a_i x^i, \sum_i b_i x^i \right) = \sum_i a_i b_i$.

V 上的线性变换 $\mathcal{A}: f(x) \mapsto xf(x)$ 是正交变换. $\mathbb{R}[x]$ 是 V 的 \mathcal{A} -不变子空间. \mathcal{A} 在 $\mathbb{R}[x]$ 上的限制映射 \mathcal{B} 不是满射. 故 \mathcal{B} 不是 $\mathbb{R}[x]$ 上的正交变换.

当实内积空间 V 是无限维时, V 的结构可能非常复杂, 验证一个线性变换是否正交变换, 或者构造一个满足要求的正交变换, 通常不是一件容易的事. 当 V 是有限维时, 由于存在标准正交基, V 的结构相对简单. 有下列常用结论.

定理 10.10. 设 V 是 n 维实内积空间, $\mathcal{A} \in L(V)$, $S = \{e_1, e_2, \dots, e_n\}$ 是 V 的一个标准正交基. 下列叙述相互等价.

1. \mathcal{A} 是正交变换.
2. $\mathcal{A}(S) = \{\mathcal{A}e_1, \mathcal{A}e_2, \dots, \mathcal{A}e_n\}$ 也是 V 的一个标准正交基.
3. \mathcal{A} 在 S 下的矩阵表示是正交方阵.

证明. $1 \Rightarrow 2$: 设 $x_1 \mathcal{A}e_1 + \cdots + x_n \mathcal{A}e_n = \mathcal{A}(x_1 e_1 + \cdots + x_n e_n) = \mathbf{0}$. 由 \mathcal{A} 是可逆变换, 得 $x_1 e_1 + \cdots + x_n e_n = \mathbf{0}$, 故 $x_1 = \cdots = x_n = 0$. 因此, $\mathcal{A}(S)$ 线性无关, $\mathcal{A}(S)$ 是 V 的基. 再由 $(\mathcal{A}e_i, \mathcal{A}e_j) = (e_i, e_j) = \delta_{ij}$, 得 $\mathcal{A}(S)$ 是 V 的标准正交基.

$2 \Rightarrow 1$: 设 $\alpha = x_1 e_1 + \cdots + x_n e_n \in \text{Ker } \mathcal{A}$. 由 $\mathcal{A}\alpha = x_1 \mathcal{A}e_1 + \cdots + x_n \mathcal{A}e_n = \mathbf{0}$, 得 $x_1 = \cdots = x_n = 0$, 即 $\alpha = \mathbf{0}$. 故 \mathcal{A} 是单射. 根据定理 9.10, \mathcal{A} 也是满射. 对于任意 $\alpha = \sum_{i=1}^n x_i e_i$, $\beta = \sum_{i=1}^n y_i e_i$, 有 $(\mathcal{A}\alpha, \mathcal{A}\beta) = \sum_{i,j=1}^n x_i y_j (\mathcal{A}e_i, \mathcal{A}e_j) = \sum_{i=1}^n x_i y_i = (\alpha, \beta)$. 因此, \mathcal{A} 是正交变换.

$2 \Leftrightarrow 3$: 设 P 是 \mathcal{A} 在 S 下的矩阵表示, 则 $P^T P = ((\mathcal{A}e_i, \mathcal{A}e_j))$. 因此, $\mathcal{A}(S)$ 是 V 的标准正交基 $\Leftrightarrow (\mathcal{A}e_i, \mathcal{A}e_j) = \delta_{ij} \Leftrightarrow P^T P = I \Leftrightarrow P$ 是正交方阵. \square

当 V 是无穷维时, 我们有与定理 10.10 类似的结论.

定理 10.11. 设 V 是实内积空间, $\mathcal{A} \in L(V)$, I 是指标集合, $S = \{e_i \mid i \in I\}$ 是 V 的一个标准正交基. \mathcal{A} 是正交变换 $\Leftrightarrow \mathcal{A}(S) = \{\mathcal{A}e_i \mid i \in I\}$ 也是 V 的一个标准正交基.

定理 10.12. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 是正交变换, U 是有限维的 \mathcal{A} -不变子空间, 则 U^\perp 也是 \mathcal{A} -不变子空间.

证明. 设 $\mathcal{B} \in L(U)$ 是 \mathcal{A} 在 U 上的限制映射. 若 $\alpha \in \text{Ker } \mathcal{B}$, 则由 $\|\alpha\| = \|\mathcal{A}\alpha\| = 0$ 可得 $\alpha = \mathbf{0}$. 故 \mathcal{B} 是单射. 再由 U 是有限维, 得 \mathcal{B} 可逆. 从而, 对于任意 $\alpha \in U$ 和 $\beta \in U^\perp$, 有

$$\mathcal{A}^{-1}\alpha = \mathcal{B}^{-1}\alpha \in U, \quad (\alpha, \mathcal{A}\beta) = (\mathcal{A}^{-1}\alpha, \beta) = 0.$$

因此, $\mathcal{A}(U^\perp) \subset U^\perp$. \square

 定理 10.12 中的条件“ U 是有限维的”不可省略.

例 10.16. 设 V, \mathcal{A} 同例 10.15, $U = \mathbb{R}[x]$ 是 \mathcal{A} -不变子空间, $U^\perp = \left\{ \sum_{i=1}^n a_i x^{-i} \mid a_i \in \mathbb{R}, n \in \mathbb{N} \right\}$ 不是 \mathcal{A} -不变子空间, 尽管有 $V = U \oplus U^\perp$.

如下定理可以看作是定理 10.12 的推广.

定理 10.13. 设 V 是实内积空间, $\mathcal{A} \in L(V)$, U 和 U^\perp 都是 \mathcal{A} -不变子空间, 并且 $V = U \oplus U^\perp$. \mathcal{A} 是正交变换 $\Leftrightarrow \mathcal{A}$ 在 U 和 U^\perp 上的限制映射都是正交变换.

证明. (\Rightarrow) 设 \mathcal{B} 是 \mathcal{A} 在 U 上的限制映射. 由 \mathcal{A} 是正交变换, 可得 \mathcal{B} 保持 U 中任意两个向量的内积不变, 并且 \mathcal{B} 是单射. 对于任意 $\alpha \in U$, 设 $\mathcal{A}^{-1}\alpha = u + v$ ($u \in U, v \in U^\perp$). 由 $\alpha = \mathcal{A}u + \mathcal{A}v$, 可得 $\alpha = \mathcal{A}u, \mathcal{A}v = \mathbf{0}$. 故 \mathcal{B} 是满射. 因此, \mathcal{B} 是正交变换. 同理, \mathcal{A} 在 U^\perp 上的限制映射也是正交变换.

(\Leftarrow) 对于任意 $\alpha = u + v \in V$ ($u \in U, v \in U^\perp$), 设 $u = \mathcal{A}u_1, v = \mathcal{A}v_1$ ($u_1 \in U, v_1 \in U^\perp$), 则 $\alpha = \mathcal{A}(u_1 + v_1)$. 故 \mathcal{A} 是满射. 若 $\mathcal{A}\alpha = \mathbf{0}$, 则 $\mathcal{A}u + \mathcal{A}v = \mathbf{0} \Rightarrow \mathcal{A}u = \mathcal{A}v = \mathbf{0} \Rightarrow u = v = \mathbf{0} \Rightarrow \alpha = \mathbf{0}$. 故 \mathcal{A} 是单射. 对于任意 $\beta = u' + v' \in V$ ($u' \in U, v' \in U^\perp$),

$$(\mathcal{A}\alpha, \mathcal{A}\beta) = (\mathcal{A}u + \mathcal{A}u', \mathcal{A}v + \mathcal{A}v') = (u, u') + (v, v') = (u + u', v + v') = (\alpha, \beta).$$

因此, \mathcal{A} 是正交变换. \square

例 10.17. 设 V 是实内积空间, $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \in V$ 满足 $(\alpha_i, \alpha_j) = (\beta_i, \beta_j), \forall i, j$, 则存在 V 上的正交变换 \mathcal{A} 使得 $\mathcal{A}\alpha_i = \beta_i, \forall i$.

证明. 任取 $U = \text{Span}(\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k)$ 的标准正交基 $S = \{e_1, \dots, e_r\}$. 设 $A, B \in \mathbb{R}^{r \times k}$ 的列向量分别由 $\alpha_1, \dots, \alpha_k$ 和 β_1, \dots, β_k 在 S 下的坐标构成. 由 $(\alpha_i, \alpha_j) = (\beta_i, \beta_j), \forall i, j$ 可得 $A^T A = B^T B$. 根据第六章 §6.3 习题 2, 存在正交方阵 P 使得 $B = PA$. 设 U 上的正交变换 \mathcal{B} 在 S 下的矩阵表示是 P , 则有 $\mathcal{B}\alpha_i = \beta_i, \forall i$. 下面, 我们把 \mathcal{B} 延拓为 V 上的正交变换. 根据定理 10.6, $V = U \oplus U^\perp$. 定义 $\mathcal{A} \in L(V): u + v \mapsto \mathcal{B}u + v, \forall u \in U, v \in U^\perp$. \mathcal{A} 在 U 和 U^\perp 上的限制映射 (分别是 \mathcal{B} 和恒等映射) 都是正交变换. 根据定理 10.13, \mathcal{A} 是正交变换. \square

例 10.17 的结论不可以推广到无穷多个向量. 例如, 在 $V = \mathbb{R}[x]$ 上定义内积

$$\left(\sum_i a_i x^i, \sum_i b_i x^i \right) = \sum_i a_i b_i.$$

把所有 $x^n (n \in \mathbb{N})$ 映成 x^{n+1} 的线性变换 $\mathcal{A} \in L(V)$ 不是满射. 故 \mathcal{A} 不是正交变换.

习题

1. 设 \mathcal{A}, \mathcal{B} 都是实内积空间 V 上的正交变换. 证明: $\mathcal{A}\mathcal{B}$ 和 \mathcal{A}^{-1} 也都是 V 上的正交变换.
2. 设 V 是实内积空间, $\mathcal{A}: V \rightarrow V$ 是任意映射.
 - (1) 证明: 若 $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta), \forall \alpha, \beta \in V$, 则 \mathcal{A} 是线性映射.
 - (2) 举例: 线性映射 \mathcal{A} 满足 $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta), \forall \alpha, \beta \in V$, 但 \mathcal{A} 不可逆.
 - (3) 举例: 可逆映射 \mathcal{A} 满足 $\|\mathcal{A}\alpha\| = \|\alpha\|, \forall \alpha \in V$, 但 \mathcal{A} 不是线性映射.
3. 完成定理 10.11 的证明.
4. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 可逆. 证明: 若 \mathcal{A} 保持任意两个向量的正交关系, 即

$$\mathcal{A}\alpha \perp \mathcal{A}\beta \Leftrightarrow \alpha \perp \beta, \quad \forall \alpha, \beta \in V$$

则存在 $\lambda > 0$ 使得 $\lambda\mathcal{A}$ 是正交变换.

5. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 是正交变换. 证明:
 - (1) 若 $\lambda \in \mathbb{R}$ 是 \mathcal{A} 的特征值, 则 $\lambda = 1$ 或 $\lambda = -1$.
 - (2) 若 $\alpha, \beta \in V$ 满足 $\mathcal{A}\alpha = \alpha$ 且 $\mathcal{A}\beta = -\beta$, 则 $\alpha \perp \beta$.
 - (3) $\text{Ker}(\mathcal{I} - \mathcal{A}) = \text{Ker}(\mathcal{I} - \mathcal{A})^2, \text{Ker}(\mathcal{I} + \mathcal{A}) = \text{Ker}(\mathcal{I} + \mathcal{A})^2$.
 - (4) 若 $d_{\mathcal{A}}(x) = \prod_{i=1}^k (x - \lambda_i)$, 其中 $\lambda_i \in \mathbb{C}$, 则 $|\lambda_i| = 1$ 并且 $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同.

§10.4 伴随变换

在第九章 §9.3 中介绍了一般线性空间 V 的对偶空间 V^* 的概念, 以及线性映射 $\mathcal{A} \in L(U, V)$ 的伴随映射 $\mathcal{A}^\dagger \in L(V^*, U^*)$ 的概念. 在本章中, 实数域上的线性空间 V 被赋予内积运算, 成为了实内积空间. 我们将重新审视对偶空间和伴随变换的性质.

定理 10.14 (Riesz^[8]表示定理). 设 V 是有限维实内积空间, $f \in V^*$, 则存在唯一的 $\alpha \in V$ 使得

$$f(x) = (\alpha, x), \quad \forall x \in V.$$

证明. 任取 V 的标准正交基 e_1, e_2, \dots, e_n , 设 $x = \sum_{i=1}^n x_i e_i$, $\alpha = \sum_{i=1}^n f(e_i) e_i \in V$, 则

$$f(x) = \sum_{i=1}^n x_i f(e_i) = \sum_{i=1}^n f(e_i) (e_i, x) = (\alpha, x).$$

若 $\beta \in V$ 也满足 $f(x) = (\beta, x)$, $\forall x \in V$, 则有 $(\beta - \alpha, x) = 0$, $\forall x \in V$. 取 $x = \beta - \alpha$, 得 $\beta = \alpha$. 因此, α 是唯一的. \square

定理 10.14 表明, 当 V 是有限维时, V 上的任意线性函数 $f(x)$ 都可以表示为 $f(x) = (\alpha, x)$ 的形式. 然而, 当 V 是无限维时, 存在 V 上的线性函数 $f(x)$ 无法表示为 (α, x) 的形式.

例 10.18. 在 $V = \mathbb{R}[x]$ 上定义内积 $(f, g) = \int_0^1 f(x)g(x)dx$. 映射 $g \mapsto g(0)$ 是 V 上的线性函数. 不存在 $f \in V$ 使得 $g(0) = (f, g)$, $\forall g \in V$.

证明. 假设存在 $f \in \mathbb{R}[x]$ 使得 $g(0) = (f, g)$ 恒成立. 取 $g(x) = xf(x)$, 由 $\int_0^1 xf^2(x)dx = (f, g) = 0$, 得 $f = 0$. 再取 $g(x) = 1$, 得 $(f, 1) = 1$, 与 $f = 0$ 矛盾. \square

我们自然会提出问题: 什么样的线性函数 $f(x)$ 可以表示为 $f(x) = (\alpha, x)$ 的形式? 此类问题是更深层次的线性代数和泛函分析学科的研究内容. 感兴趣者可参阅相关文献, 例如 [16] *Advanced Linear Algebra*. 显然, 对于任意 $\alpha \in V$, $f_\alpha(x) = (\alpha, x)$ 是 V 上的线性函数, $\tilde{V} = \{f_\alpha \mid \alpha \in V\}$ 是 V^* 的子空间. V 与 \tilde{V} 之间可以建立一一对应

$$\tau: V \rightarrow \tilde{V}, \quad \alpha \mapsto f_\alpha.$$

容易验证, τ 是同构映射. 根据定义 9.8, 任意线性变换 $\mathcal{A} \in L(V)$ 有伴随变换

$$\mathcal{A}^\dagger \in L(V^*), \quad f(x) \mapsto f(\mathcal{A}x).$$

若 \tilde{V} 是 \mathcal{A}^\dagger -不变子空间, 则 $\mathcal{A}^\dagger: f_\alpha \mapsto f_\beta$ 诱导出 V 上的线性变换 $\mathcal{A}^*: \alpha \mapsto \beta$, 如下图所示.

$$\begin{array}{ccc} \alpha \in V & \xrightarrow{\mathcal{A}^*} & \beta \in V \\ \tau \downarrow & & \uparrow \tau^{-1} \\ f_\alpha \in \tilde{V} & \xrightarrow{\mathcal{A}^\dagger} & f_\beta \in \tilde{V} \end{array}$$

由 $\mathcal{A}^* = \tau^{-1} \circ \mathcal{A}^\dagger \circ \tau$, 我们可以认为 \mathcal{A}^* 与 \mathcal{A}^\dagger 是“相似”的. \mathcal{A}^* 也常称为“伴随变换”.

^[8]Frigyes Frédéric Riesz, 1880–1956, 匈牙利数学家, 泛函分析创始人之一.

定义 10.9. 设 V 是实内积空间, $\mathcal{A} \in L(V)$. 若存在 $\mathcal{B} \in L(V)$, 使得

$$(\mathcal{B}\alpha, \beta) = (\alpha, \mathcal{A}\beta), \quad \forall \alpha, \beta \in V$$

则 \mathcal{B} 是唯一的, \mathcal{B} 称为 \mathcal{A} 的伴随变换, 记作 \mathcal{A}^* .

当 V 是有限维时, 根据定理 10.14, $\tilde{V} = V^*$. 因此, 任意 $\mathcal{A} \in L(V)$ 都有伴随变换 \mathcal{A}^* , 并且有如下结论.

定理 10.15. 设 V 是有限维实内积空间, $\mathcal{A} \in L(V)$, S 是 V 的一个标准正交基, A 是 \mathcal{A} 在 S 下的矩阵表示, 则 A^T 是 \mathcal{A}^* 在 S 下的矩阵表示.

证明. 任取 $\alpha, x \in V$, 设 $\beta = \mathcal{A}^*\alpha$, $\mathbf{a}, \mathbf{b}, \mathbf{x}$ 分别是 α, β, x 在 S 下的坐标. 由

$$(\alpha, \mathcal{A}x) = \mathbf{a}^T A \mathbf{x}, \quad (\beta, x) = \mathbf{b}^T \mathbf{x}$$

可得 $\mathbf{b} = A^T \mathbf{a}$, 此即为 \mathcal{A}^* 在 S 下的坐标表示. 故 A^T 是 \mathcal{A}^* 在 S 下的矩阵表示. \square

例 10.19. 设 $V = \mathbb{R}_3[x]$ 的内积 $(f, g) = \int_0^1 f(x)g(x)dx$. 求微分变换 $\mathcal{D} \in L(V)$ 的伴随变换 \mathcal{D}^* .

解法一. 在 V 的基 $S = \{1, x, x^2\}$ 下, 度量矩阵 $G = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$, \mathcal{D} 的矩阵表示 $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$,

\mathcal{D}^* 的矩阵表示 B 满足 $B^T G = GA$. 故 $B = G^{-1} A^T G = \begin{pmatrix} -6 & 2 & 3 \\ 12 & -24 & -26 \\ 0 & 30 & 30 \end{pmatrix}$. 即 $\mathcal{D}^*1 = 12x - 6$,

$\mathcal{D}^*x = 30x^2 - 24x + 2$, $\mathcal{D}^*(x^2) = 30x^2 - 26x + 3$. \square

解法二. 根据例 10.10, 取 V 的标准正交基 $\gamma_1 = 1$, $\gamma_2 = \sqrt{12}(x - \frac{1}{2})$, $\gamma_3 = \sqrt{180}(x^2 - x + \frac{1}{6})$, 则

$\mathcal{D}\gamma_1 = 0$, $\mathcal{D}\gamma_2 = \sqrt{12}\gamma_1$, $\mathcal{D}\gamma_3 = \sqrt{60}\gamma_2$. 在 $\{\gamma_1, \gamma_2, \gamma_3\}$ 下, \mathcal{D} 的矩阵表示 $A = \begin{pmatrix} 0 & \sqrt{12} & 0 \\ 0 & 0 & \sqrt{60} \\ 0 & 0 & 0 \end{pmatrix}$,

\mathcal{D}^* 的矩阵表示 $B = A^T$. 即 $\mathcal{D}^*\gamma_1 = \sqrt{12}\gamma_2$, $\mathcal{D}^*\gamma_2 = \sqrt{60}\gamma_3$, $\mathcal{D}^*\gamma_3 = 0$. \square

当 V 是无限维时, 不是所有线性变换 $\mathcal{A} \in L(V)$ 都有伴随变换 \mathcal{A}^* .

例 10.20. 设 $V = \mathbb{R}[x]$, $\mathcal{D} \in L(V)$ 是微分变换.

1. 在 V 上定义内积 $(\sum_{i=0}^n f_i x^i, \sum_{i=0}^n g_i x^i) = \sum_{i=0}^n f_i g_i$, 其中 $f_i, g_i \in \mathbb{R}$. 求 \mathcal{D} 的伴随变换.

2. 在 V 上定义内积 $(f, g) = \int_0^1 f(x)g(x)dx$. 证明 \mathcal{D} 没有伴随变换.

解答.

1. 设 $\mathcal{A} \in L(V)$, $\sum_{i=0}^n f_i x^i \mapsto \sum_{i=0}^n (i+1)f_i x^{i+1}$, 则

$$(\mathcal{A}x^i, x^j) = ((i+1)x^{i+1}, x^j) = j\delta_{i+1,j} = (x^i, jx^{j-1}) = (x^i, \mathcal{D}x^j), \quad \forall i, j \in \mathbb{N}.$$

故 \mathcal{A} 是 \mathcal{D} 的伴随变换.

2. 假设 \mathcal{D}^* 存在. 设 $f = \mathcal{D}^*1$, 则

$$\int_0^1 f(x)g(x)dx = (\mathcal{D}^*1, g) = (1, \mathcal{D}g) = \int_0^1 g'(x)dx = g(1) - g(0), \quad \forall g \in V.$$

取 $g = x(1-x)f(x)$, 得 $\int_0^1 x(1-x)f^2(x)dx = 0$, 故 $f = 0$. 再取 $g = x$, 得 $\int_0^1 f(x)xdx = 1$, 与 $f = 0$ 矛盾. \square

定理 10.16. 设 V 是实内积空间, $\mathcal{A}, \mathcal{B} \in L(V)$, $\lambda \in \mathbb{R}$. 若 \mathcal{A}^* 和 \mathcal{B}^* 都存在, 则 \mathcal{A}^* , $\lambda\mathcal{A}$, $\mathcal{A} + \mathcal{B}$, $\mathcal{A}\mathcal{B}$ 的伴随变换也都存在, 并且 $(\mathcal{A}^*)^* = \mathcal{A}$, $(\lambda\mathcal{A})^* = \lambda\mathcal{A}^*$, $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$, $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$.

证明. 设 α, β 是 V 中任意向量.

- $(\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}^*\beta)$. 故 \mathcal{A} 是 \mathcal{A}^* 的伴随变换.
- $(\lambda\mathcal{A}^*\alpha, \beta) = \lambda(\mathcal{A}^*\alpha, \beta) = \lambda(\alpha, \mathcal{A}\beta) = (\alpha, \lambda\mathcal{A}\beta)$. 故 $\lambda\mathcal{A}^*$ 是 $\lambda\mathcal{A}$ 的伴随变换.
- $((\mathcal{A}^* + \mathcal{B}^*)\alpha, \beta) = (\mathcal{A}^*\alpha + \mathcal{B}^*\alpha, \beta) = (\mathcal{A}^*\alpha, \beta) + (\mathcal{B}^*\alpha, \beta) = (\alpha, \mathcal{A}\beta) + (\alpha, \mathcal{B}\beta) = (\alpha, \mathcal{A}\beta + \mathcal{B}\beta) = (\alpha, (\mathcal{A} + \mathcal{B})\beta)$. 故 $\mathcal{A}^* + \mathcal{B}^*$ 是 $\mathcal{A} + \mathcal{B}$ 的伴随变换.
- $(\mathcal{B}^*\mathcal{A}^*\alpha, \beta) = (\mathcal{A}^*\alpha, \mathcal{B}\beta) = (\alpha, \mathcal{A}\mathcal{B}\beta)$. 故 $\mathcal{B}^*\mathcal{A}^*$ 是 $\mathcal{A}\mathcal{B}$ 的伴随变换. \square

定理 10.17. 设 V 是实内积空间, $\mathcal{A} \in L(V)$, \mathcal{A}^* 存在.

1. 若 U 是 \mathcal{A} -不变子空间, 则 U^\perp 是 \mathcal{A}^* -不变子空间.
2. $\text{Ker } \mathcal{A}^*\mathcal{A} = \text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp = (\text{Im } \mathcal{A}^*\mathcal{A})^\perp$.
当 $\text{Im } \mathcal{A}^*$ 是有限维时, $V = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*$ 且 $\text{Im } \mathcal{A}^* = \text{Im } \mathcal{A}^*\mathcal{A}$.
3. $\text{Ker } \mathcal{A}\mathcal{A}^* = \text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp = (\text{Im } \mathcal{A}\mathcal{A}^*)^\perp$.
当 $\text{Im } \mathcal{A}$ 是有限维时, $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A}^*$ 且 $\text{Im } \mathcal{A} = \text{Im } \mathcal{A}\mathcal{A}^*$.

证明.

1. 对于任意 $\alpha \in U^\perp$, 由 $(\mathcal{A}^*\alpha, \beta) = (\alpha, \mathcal{A}\beta) = 0, \forall \beta \in U$ 可得 $\mathcal{A}^*\alpha \in U^\perp$. 故 U^\perp 是 \mathcal{A}^* -不变子空间.
2. 易知 $\text{Ker } \mathcal{A}^*\mathcal{A} \supset \text{Ker } \mathcal{A}$. 对于任意 $\alpha \in \text{Ker } \mathcal{A}^*\mathcal{A}$, 由 $(\mathcal{A}\alpha, \mathcal{A}\alpha) = (\mathcal{A}^*\mathcal{A}\alpha, \alpha) = 0$ 可得 $\mathcal{A}\alpha = \mathbf{0}$, 即 $\alpha \in \text{Ker } \mathcal{A}$. 故 $\text{Ker } \mathcal{A}^*\mathcal{A} \subset \text{Ker } \mathcal{A}$. 因此, $\text{Ker } \mathcal{A}^*\mathcal{A} = \text{Ker } \mathcal{A}$.

对于任意 $\alpha \in V$, $\alpha \in \text{Ker } \mathcal{A} \Leftrightarrow “(\mathcal{A}\alpha, \beta) = 0, \forall \beta” \Leftrightarrow “(\alpha, \mathcal{A}^*\beta) = 0, \forall \beta” \Leftrightarrow \alpha \in (\text{Im } \mathcal{A}^*)^\perp$. 因此, $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$.

设 $\mathcal{B} = \mathcal{A}^*\mathcal{A}$, 则 $\mathcal{B}^* = \mathcal{B}$. 由 $\text{Ker } \mathcal{B} = (\text{Im } \mathcal{B}^*)^\perp$ 可得 $\text{Ker } \mathcal{A}^*\mathcal{A} = (\text{Im } \mathcal{A}^*\mathcal{A})^\perp$.

当 $\text{Im } \mathcal{A}^*$ 是有限维时, $\text{Im } \mathcal{A}^*\mathcal{A} \subset \text{Im } \mathcal{A}^*$ 也是有限维的. 根据定理 10.6,

$$V = \text{Im } \mathcal{A}^* \oplus (\text{Im } \mathcal{A}^*)^\perp = \text{Im } \mathcal{A}^*\mathcal{A} \oplus (\text{Im } \mathcal{A}^*\mathcal{A})^\perp.$$

再由 $(\text{Im } \mathcal{A}^*)^\perp = (\text{Im } \mathcal{A}^*\mathcal{A})^\perp$, 得 $\text{Im } \mathcal{A}^* = \text{Im } \mathcal{A}^*\mathcal{A}$.

3. 把结论 2 中的 \mathcal{A} 换成 \mathcal{A}^* , 由 $(\mathcal{A}^*)^* = \mathcal{A}$ 可得结论 3 成立. \square

定义 10.10. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 的伴随变换 \mathcal{A}^* 存在. 若 $\mathcal{A}^* = \mathcal{A}$, 则 \mathcal{A} 称为自伴变换. 若 $\mathcal{A}^* = -\mathcal{A}$, 则 \mathcal{A} 称为斜自伴变换. 若 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$, 则 \mathcal{A} 称为规范变换.

例 10.21. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 可逆. \mathcal{A} 是正交变换 $\Leftrightarrow \mathcal{A}^* = \mathcal{A}^{-1}$.

例 10.22. 设 \mathbb{R} 上的线性空间 V 由所有形如 $f(x) = \sum_{k=0}^n a_k \cos(kx) + \sum_{k=1}^n b_k \sin(kx)$ 的三角函数构成. 在 V 上定义内积 $(f, g) = \int_0^{2\pi} f(x)g(x)dx$, 使 V 成为实内积空间. 设 $\mathcal{D} \in L(V)$ 是微分变换, 则 \mathcal{D} 是斜自伴变换, \mathcal{D}^2 是自伴变换. 对于任意多项式 $p(x) \in \mathbb{R}[x]$, $p(\mathcal{D})$ 是规范变换.

证明. $\forall f, g \in V$, $(f, \mathcal{D}g) + (\mathcal{D}f, g) = \int_0^{2\pi} (f(x)g'(x) + f'(x)g(x))dx = f(2\pi)g(2\pi) - f(0)g(0) = 0$. 因此, $\mathcal{D}^* = -\mathcal{D}$, \mathcal{D} 是斜自伴变换. 由 $(\mathcal{D}^2)^* = (\mathcal{D}^*)^2 = \mathcal{D}^2$, 得 \mathcal{D}^2 是自伴变换. 由 $(p(\mathcal{D}))^* = p(\mathcal{D}^*) = p(-\mathcal{D})$ 与 $p(\mathcal{D})$ 乘积可交换, 得 $p(\mathcal{D})$ 是规范变换. \square

定理 10.18. 设 V 是实内积空间, $\mathcal{A} \in L(V)$. 下列叙述相互等价.

1. \mathcal{A} 是规范变换.
2. 对于任意 $\alpha \in V$, $\|\mathcal{A}\alpha\| = \|\mathcal{A}^*\alpha\|$.
3. 对于任意 $\alpha, \beta \in V$, $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\mathcal{A}^*\alpha, \mathcal{A}^*\beta)$.

证明. $1 \Rightarrow 2$: $\|\mathcal{A}\alpha\|^2 = (\mathcal{A}\alpha, \mathcal{A}\alpha) = (\mathcal{A}^*\mathcal{A}\alpha, \alpha) = (\mathcal{A}\mathcal{A}^*\alpha, \alpha) = (\mathcal{A}^*\alpha, \mathcal{A}^*\alpha) = \|\mathcal{A}^*\alpha\|^2$.

$2 \Rightarrow 3$: $(\mathcal{A}\alpha, \mathcal{A}\beta) = \frac{1}{4}(\|\mathcal{A}(\alpha + \beta)\|^2 - \|\mathcal{A}(\alpha - \beta)\|^2) = \frac{1}{4}(\|\mathcal{A}^*(\alpha + \beta)\|^2 - \|\mathcal{A}^*(\alpha - \beta)\|^2) = (\mathcal{A}^*\alpha, \mathcal{A}^*\beta)$.

$3 \Rightarrow 1$: 对于任意 $\alpha, \beta \in V$, $(\mathcal{A}^*\mathcal{A}\alpha, \beta) = (\mathcal{A}\alpha, \mathcal{A}\beta) = (\mathcal{A}^*\alpha, \mathcal{A}^*\beta) = (\mathcal{A}\mathcal{A}^*\alpha, \beta)$. 由 α, β 的任意性, 得 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$, \mathcal{A} 是规范变换. \square

定理 10.19. 设 V 是有限维实内积空间, $\mathcal{A} \in L(V)$.

1. 设 S 是 V 的标准正交基, A 是 \mathcal{A} 在 S 下的矩阵表示, 则 \mathcal{A} 是正交变换/自伴变换/斜自伴变换/规范变换当且仅当 A 是正交方阵/对称方阵/反对称方阵/规范方阵.
2. 设 S_1 和 S_2 都是 V 的标准正交基, A_i 是 \mathcal{A} 在 S_i 下的矩阵表示, P 是从 S_1 到 S_2 的过渡矩阵, 则 $A_2 = PA_1P^{-1}$, P 是正交方阵, 即 A_1 和 A_2 正交相似.

如下定理 10.20 是定理 10.12 的推广.

定理 10.20. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 是规范变换, U 是有限维的 \mathcal{A} -不变子空间, 则 U 是 \mathcal{A}^* -不变子空间, 从而 U^\perp 是 \mathcal{A} -不变子空间.

证明. 任取 U 的一个标准正交基 e_1, e_2, \dots, e_n . 根据定理 10.8, 对于任意 $i = 1, 2, \dots, n$,

$$\|\mathcal{A}e_i\|^2 = \sum_{j=1}^n (\mathcal{A}e_i, e_j)^2, \quad \|\mathcal{A}^*e_i\|^2 \geq \sum_{j=1}^n (\mathcal{A}^*e_i, e_j)^2 = \sum_{j=1}^n (e_i, \mathcal{A}e_j)^2.$$

由此可得, $\sum_{i=1}^n \|\mathcal{A}^*e_i\|^2 \geq \sum_{i=1}^n \|\mathcal{A}e_i\|^2$. 由于 \mathcal{A} 是规范变换, $\|\mathcal{A}^*e_i\| = \|\mathcal{A}e_i\|$. 故上述不等式均取等号, 得 $\mathcal{A}^*e_i \in U$. 因此, U 是 \mathcal{A}^* -不变子空间. 根据定理 10.17, U^\perp 是 \mathcal{A} -不变子空间. \square

当 V 是有限维实内积空间时, 设 $\mathcal{A} \in L(V)$ 是规范变换, $p(x)$ 是 $\varphi_{\mathcal{A}}(x)$ 的不可约因子, U 是任意非零向量 $\alpha \in \text{Ker } p(\mathcal{A})$ 生成的 \mathcal{A} -循环子空间, 则 $\dim U = \deg(p) \leq 2$. 根据定理 10.20 和定理 10.6, U^\perp 是 \mathcal{A} -不变子空间, $V = U \oplus U^\perp$. 考虑 \mathcal{A} 在 U^\perp 上的限制映射, 同理可以分解 U^\perp . 重复以上过程, 从而可以把 V 分解为若干个两两正交的 \mathcal{A} -循环子空间的直和, $V = \bigoplus_{i \in I} V_i$, 其中每个 V_i 的维数不超过 2. 由此可得定理 6.9 的几何解释.

习题

1. 设 V 是实内积空间, $\mathcal{A} \in L(V)$. 证明: 若 \mathcal{A}^* 存在, 则 \mathcal{A}^* 是唯一的.
2. 证明定理 10.19.
3. 设 $V = \mathbb{R}^n$, 内积 $(x, y) = x^T G y$, $\alpha, \beta \in V$, $\mathcal{A}(x) = x + (\alpha, x)\beta$. 求 \mathcal{A}^* .
4. 设 $V = \mathbb{R}^{m \times n}$, 内积 $(X, Y) = \text{tr}(X^T Y)$, $P \in \mathbb{R}^{m \times m}$, $Q \in \mathbb{R}^{n \times n}$, $\mathcal{A}(X) = P X Q$. 求 \mathcal{A}^* .
5. 设 $V = \mathbb{R}[x]$, 内积 $(f, g) = \int_{-1}^1 f(x)g(x)dx$, $\mathcal{A} \in L(V) : f(x) \mapsto x f(-x)$. 求 \mathcal{A}^* .
6. 设 $V = \mathbb{R}_3[x]$, 内积 $(f, g) = \int_{-1}^1 f(x)g(x)dx$, $\mathcal{D} \in L(V)$ 是微分变换. 求 \mathcal{D}^* .
7. 设 V 是实内积空间, $\mathcal{A}, \mathcal{B} \in L(V)$ 都是自伴变换. 证明: $\mathcal{A}\mathcal{B}$ 是自伴变换 $\Leftrightarrow \mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$.
8. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 可逆, \mathcal{A}^* 存在.
 - (1) 证明: 若 $(\mathcal{A}^{-1})^*$ 存在, 则 \mathcal{A}^* 可逆并且 $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$.
 - (2) 证明: 若 \mathcal{A}^* 可逆, 则 $(\mathcal{A}^{-1})^*$ 存在并且 $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$.
 - (3) 证明: \mathcal{A}^* 是单射并且 $(\text{Im } \mathcal{A}^*)^\perp = \{0\}$.
 - (4) 举例: \mathcal{A}^* 不是满射, 从而 \mathcal{A}^* 不可逆, $(\mathcal{A}^{-1})^*$ 不存在.
9. 设 V 是实内积空间, $\mathcal{A} \in L(V)$, \mathcal{A}^* 存在. 证明:
 - (1) \mathcal{A} 是斜自伴变换当且仅当对于任意 $\alpha \in V$ 有 $(\alpha, \mathcal{A}\alpha) = 0$.
 - (2) 若 V 是有限维, \mathcal{A} 是斜自伴变换, 则 $(\mathcal{I} + \mathcal{A})^{-1}(\mathcal{I} - \mathcal{A})$ 是正交变换.
 - (3) 若 \mathcal{A} 是斜自伴变换. $\mathcal{I} + \mathcal{A}$ 和 $\mathcal{I} - \mathcal{A}$ 是否一定是可逆变换? 证明你的结论.
10. 设 V 是实内积空间, $\mathcal{A} \in L(V)$ 的伴随变换 \mathcal{A}^* 和最小多项式 $d_{\mathcal{A}}$ 都存在. 证明:
 - (1) $d_{\mathcal{A}^*} = d_{\mathcal{A}}$.
 - (2) \mathcal{A} 是规范变换 $\Rightarrow d_{\mathcal{A}}$ 无重根.
 - (3) \mathcal{A} 是规范变换 \Leftrightarrow 存在 $f(x) \in \mathbb{R}[x]$ 使得 $\mathcal{A}^* = f(\mathcal{A})$.

§10.5 复内积空间

由于 \mathbb{C} 是 \mathbb{R} 的代数闭包, 我们试图在复数域上的线性空间 V 上定义内积, 赋予 V 几何结构. 注意到 \mathbb{C} 有一个自然的度量 $|z|^2 = z \cdot \bar{z}$. 因此, 我们保留内积的“正定性”, 而对“对称性”和“双线性”略作修改. 由此得到如下的复内积.

定义 10.11. 设 V 是 \mathbb{C} 上的线性空间. 若函数 $\rho: V \times V \rightarrow \mathbb{C}$ 具有下列性质,

1. (共轭对称性) $\rho(\alpha, \beta) = \overline{\rho(\beta, \alpha)}$
2. (共轭双线性) $\rho(\lambda\alpha + \mu\beta, \gamma) = \bar{\lambda}\rho(\alpha, \gamma) + \bar{\mu}\rho(\beta, \gamma)$, $\rho(\alpha, \lambda\beta + \mu\gamma) = \lambda\rho(\alpha, \beta) + \mu\rho(\alpha, \gamma)$
3. (正定性) 当 $\alpha \neq \mathbf{0}$ 时, $\rho(\alpha, \alpha) > 0$

其中 $\alpha, \beta, \gamma \in V$, $\lambda, \mu \in \mathbb{C}$, 则 ρ 称为 V 上的一个复内积, 代数结构 (V, ρ) 称为复内积空间. 设 U 是 V 的子空间, 则 ρ 可以自然地限制在 U 上, 成为复内积 $\tilde{\rho}: U \times U \rightarrow \mathbb{C}$. 复内积空间 $(U, \tilde{\rho})$ 称为 (V, ρ) 的子空间. 在不引起歧义的情况下, 可以略去复内积名称 ρ , 把 $\rho(\alpha, \beta)$ 记作 (α, β) .

通过复内积可以定义一个向量的长度 $\|\alpha\| = \sqrt{(\alpha, \alpha)}$ 和两个向量的正交关系 $\alpha \perp \beta \Leftrightarrow (\alpha, \beta) = 0$. 长度是 1 的向量称为单位向量. 复内积没有对称性. 两个向量的正交关系却是对称的, $\alpha \perp \beta \Leftrightarrow \beta \perp \alpha$.

设 $S \subset V$, $S^\perp = \{\alpha \in V \mid \alpha \perp \beta, \forall \beta \in S\}$ 是 V 的子空间, 称为 $\text{Span}(S)$ 的正交补空间. 对于任意向量组 S , 有 $\text{Span}(S) \cap S^\perp = \{\mathbf{0}\}$, 即 $\text{Span}(S) + S^\perp$ 是直和.

当 V 是有限维时, 设 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 V 的基, n 阶方阵 $G = (\rho(\alpha_i, \alpha_j))$ 称为复内积 ρ 在 S 下的度量矩阵或 Gram 方阵. 对于任意 $\alpha, \beta \in V$, 设 \mathbf{x}, \mathbf{y} 分别是 α, β 在 S 下的坐标, 则有 $\rho(\alpha, \beta) = \mathbf{x}^H G \mathbf{y}$, 称为复内积 ρ 在 S 下的坐标表示.

例 10.23. 设 $V = \mathbb{C}^n$. 二元函数 $\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \bar{x}_i y_i$ 是 V 上的复内积, 称为标准复内积. \mathbb{C}^n 连同标准复内积称为 n 维酉空间.

例 10.24. 设 $V = \mathbb{C}^{m \times n}$. 二元函数 $\rho(X, Y) = \text{tr}(X^H Y)$ 是 V 上的复内积.

例 10.25. 设 V 是 $[a, b]$ 上的复数值连续函数全体. 二元函数 $\rho(f, g) = \int_a^b \overline{f(x)} g(x) dx$ 是 V 上的复内积.

定理 10.21. 设 V 是复内积空间. 对于任意 $\alpha, \beta \in V$, 有

1. (Cauchy 不等式) $|(\alpha, \beta)| \leq \|\alpha\| \cdot \|\beta\|$. 等号成立当且仅当 α, β 线性相关.
2. (三角不等式) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$. 等号成立当且仅当存在 $\lambda \geq 0$ 使得 $\alpha = \lambda\beta$ 或 $\beta = \lambda\alpha$.
3. (勾股定理) 若 $\alpha \perp \beta$, 则 $\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$.

定理 10.22. 设 (V, ρ) 是 n 维复内积空间.

1. ρ 在 V 的任意基 S 下的度量矩阵 G 是正定 Hermite 方阵.
2. 设 G_i 是 ρ 在 V 的基 S_i 下的度量矩阵, P 是从 S_1 到 S_2 的过渡矩阵, 则 $G_2 = P^H G_1 P$.
3. 对于任意正定 Hermite 方阵 $G \in \mathbb{C}^{n \times n}$, 存在 V 的基 S , 使得 ρ 在 S 下的度量矩阵是 G .

上述定理表明, 任意 n 维复内积空间 V 有标准正交基. 存在一一映射 $\tau: V \rightarrow \mathbb{C}^n$ 把 V 的标准正交基映成 \mathbb{C}^n 的标准正交基. τ 既保持两个空间的线性运算之间的对应, 也保持内积运算之间的对应. V 与酉空间 \mathbb{C}^n 同构. 因此, 许多教科书把有限维复内积空间统称为酉空间.

对于无限维复内积空间 V , 标准正交基有可能不存在. 但是, 总是可以通过 **Gram-Schmidt 标准正交化过程**把有限或可数多个线性无关的向量改造成两两正交的单位向量组.

设 $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ 线性无关, 令

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\gamma_i, \alpha_k) \gamma_i, \quad \gamma_i = \frac{1}{\|\beta_k\|} \beta_k, \quad k = 1, 2, \dots, n, \dots \quad (10.3)$$

则 $\{\gamma_1, \gamma_2, \dots, \gamma_n, \dots\}$ 构成标准正交向量组. 注意: 与 (10.2) 式不同, (10.3) 式中的 (γ_i, α_k) 不可以写成 (α_k, γ_i) .

定理 10.23. 设 U 是复内积空间 V 的子空间.

1. 若 U 是有限维的, 则 $V = U \oplus U^\perp$.
2. 若 $V = U \oplus U^\perp$, 则 $(U^\perp)^\perp = U$.

定理 10.24. 设 V 是复内积空间, I 是指标集合, $S = \{e_i \mid i \in I\}$ 是 V 中的标准正交向量组.

1. (Parseval 等式) 若 S 是 V 的标准正交基, 则 $(\alpha, \beta) = \sum_{i \in I} (\alpha, e_i)(e_i, \beta)$, $\forall \alpha, \beta \in V$.
2. (Bessel 等式) 若 S 是 V 的标准正交基, 则 $\|\alpha\|^2 = \sum_{i \in I} |(e_i \alpha)|^2$, $\forall \alpha \in V$.
3. (Bessel 不等式) $\|\alpha\|^2 \geq \sum_{i \in I} |(e_i \alpha)|^2$, $\forall \alpha \in V$.
4. $\|\alpha\|^2 = \sum_{i \in I} |(e_i \alpha)|^2$ 当且仅当存在一系列 $\beta_n \in \text{Span}(S)$ 满足 $\lim_{n \rightarrow \infty} \|\alpha - \beta_n\| = 0$.

考虑 V 上的线性变换 $\mathcal{A} \in L(V)$. 当 V 是有限维时, \mathcal{A} 在 V 的某组基 S 下的矩阵表示 A 与 \mathcal{A} 一一对应. 通常取 S 为 V 的标准正交基. 当 V 是无限维时, 通常考虑 \mathcal{A} 在某个有限维的 \mathcal{A} -不变子空间 U 上的限制映射, 以及 \mathcal{A} 在 U^\perp 上的限制映射. 当 \mathcal{A} 具有良好的性质时, U^\perp 也是 \mathcal{A} -不变子空间.

定理 10.25. 设 V 是有限维复内积空间, S_1, S_2 都是 V 的标准正交基, P 是从 S_1 到 S_2 的过渡矩阵, 则 P 是酉方阵. 对于任意 $\mathcal{A} \in L(V)$, 设 \mathcal{A} 在 S_1, S_2 下的矩阵表示分别是 A_1, A_2 , 则 $A_2 = P^{-1}A_1P$, 即 A_1 与 A_2 酉相似.

定义 10.12. 设 V 是复内积空间, $\mathcal{A} \in L(V)$. 若存在 $\mathcal{B} \in L(V)$, 使得

$$(\mathcal{B}\alpha, \beta) = (\alpha, \mathcal{A}\beta), \quad \forall \alpha, \beta \in V,$$

则 \mathcal{B} 是唯一的, \mathcal{B} 称为 \mathcal{A} 的伴随变换, 记作 \mathcal{A}^* . 若 $\mathcal{A}^* = \mathcal{A}^{-1}$, 则 \mathcal{A} 称为酉变换; 若 $\mathcal{A}^* = \mathcal{A}$, 则 \mathcal{A} 称为自伴变换; 若 $\mathcal{A}^* = -\mathcal{A}$, 则 \mathcal{A} 称为斜自伴变换; 若 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$, 则 \mathcal{A} 称为规范变换.

酉变换、自伴变换、斜自伴变换都是规范变换的特例.

定理 10.26. 设 V 是有限维复内积空间, $\mathcal{A} \in L(V)$, S 是 V 的一个标准正交基.

1. 设 A 是 \mathcal{A} 在 S 下的矩阵表示, 则 A^H 是 \mathcal{A}^* 在 S 下的矩阵表示.
2. \mathcal{A} 是酉变换/自伴变换/斜自伴变换/规范变换当且仅当 A 是酉方阵/*Hermite* 方阵/反 *Hermite* 方阵/规范方阵.

定理 10.27. 设 V 是复内积空间, $\mathcal{A}, \mathcal{B} \in L(V)$, $\lambda \in \mathbb{C}$. 若 \mathcal{A}^* 和 \mathcal{B}^* 都存在, 则 \mathcal{A}^* , $\lambda\mathcal{A}$, $\mathcal{A} + \mathcal{B}$, $\mathcal{A}\mathcal{B}$ 的伴随变换也都存在, 并且 $(\mathcal{A}^*)^* = \mathcal{A}$, $(\lambda\mathcal{A})^* = \bar{\lambda}\mathcal{A}^*$, $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$, $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$.

定理 10.28. 设 V 是复内积空间, $\mathcal{A} \in L(V)$, 并且 \mathcal{A}^* 存在.

1. 若 U 是 \mathcal{A} -不变子空间, 则 U^\perp 是 \mathcal{A}^* -不变子空间.
2. $\text{Ker } \mathcal{A}^*\mathcal{A} = \text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp = (\text{Im } \mathcal{A}^*\mathcal{A})^\perp$.
当 $\text{Im } \mathcal{A}^*$ 是有限维时, $V = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*$, $\text{Im } \mathcal{A}^* = \text{Im } \mathcal{A}^*\mathcal{A}$.
3. $\text{Ker } \mathcal{A}\mathcal{A}^* = \text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp = (\text{Im } \mathcal{A}\mathcal{A}^*)^\perp$.
当 $\text{Im } \mathcal{A}$ 是有限维时, $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A}^*$, $\text{Im } \mathcal{A} = \text{Im } \mathcal{A}\mathcal{A}^*$.

定理 10.29. 设 V 是复内积空间, $\mathcal{A} \in L(V)$, I 是指标集合, $S = \{e_i \mid i \in I\}$ 是 V 的一个标准正交基. 下列叙述相互等价.

1. \mathcal{A} 是 V 上的酉变换.
2. $\mathcal{A}(S) = \{Ae_i \mid i \in I\}$ 也是 V 的一个标准正交基.
3. (V 是有限维情形) \mathcal{A} 在 S 下的矩阵表示是酉方阵.

定理 10.30. 设 V 是复内积空间, $\mathcal{A} \in L(V)$. 下列叙述相互等价.

1. \mathcal{A} 是规范变换.
2. 对于任意 $\alpha \in V$, $\|\mathcal{A}\alpha\| = \|\mathcal{A}^*\alpha\|$.
3. 对于任意 $\alpha, \beta \in V$, $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\mathcal{A}^*\alpha, \mathcal{A}^*\beta)$.

定理 10.31. 设 V 是复内积空间, $\mathcal{A} \in L(V)$ 是规范变换, U 是有限维 \mathcal{A} -不变子空间, 则 U 是 \mathcal{A}^* -不变子空间, U^\perp 是 \mathcal{A} -不变子空间. 由此可得, 当 $\dim V = n < \infty$ 时, V 可以分解为 n 个两两正交的 1 维 \mathcal{A} -不变子空间的直和.

习题

1. 对复内积空间 V , 推广 §10.1 习题 7 中的结论, 并给出证明.
2. 补充完成本节中所有定理的证明.
3. 对酉空间 \mathbb{C}^3 中向量组 $\alpha_1 = (i, 1, 0)$, $\alpha_2 = (0, i, 1)$, $\alpha_3 = (1, 0, i)$ 作 Gram-Schmidt 标准正交化.
4. 在 $\mathbb{C}_3[x]$ 上定义复内积 $(f, g) = \int_0^1 \overline{f(x)}g(x)dx$. 对基 $1, x, x^2$ 作 Gram-Schmidt 标准正交化.

5. 设 V 是复内积空间. 求下列线性变换 $\mathcal{A} \in L(V)$ 的伴随变换.

(1) $V = \mathbb{C}^3$, 复内积 $(x, y) = x^H G y$, $\mathcal{A}(X) = (\alpha, x)\beta$, $\alpha, \beta \in V$

(2) $V = \mathbb{C}^{3 \times 3}$, 复内积 $(X, Y) = \text{tr}(X^H Y)$, $\mathcal{A}(X) = P X Q$, $P, Q \in V$

(3) $V = \mathbb{C}_3[x]$, 复内积 $(f, g) = \int_{-1}^1 \overline{f(x)} g(x) dx$, $\mathcal{A}: f(x) \mapsto f(ix)$

6. 设复线性空间 $V = \mathbb{C}^{m \times n}$, $\rho(X, Y) = X^H S Y$, $\mathcal{A}(X) = P X Q$, 其中 S, P, Q 都是复数矩阵.

(1) 求 S 应满足的充分必要条件, 使得 ρ 是 V 上的复内积.

(2) 分别求矩阵对 (P, Q) 应满足的充分必要条件, 使得 \mathcal{A} 是 (V, ρ) 上的酉变换/自伴变换/斜自伴变换/规范变换.

§10.6 内积的推广

对于任意数域 \mathbb{F} 上的线性空间 V , 我们保留内积的“双线性”, 放松“对称性”, 用“非退化性”替代“正定性”, 把具有“良好性质”的双线性函数作为 V 上的内积概念的推广.

定义 10.13. 设 V 是 \mathbb{F} 上的线性空间, 映射 $\rho: V \times V \rightarrow \mathbb{F}$ 满足

$$\rho(\lambda\alpha + \mu\beta, \gamma) = \lambda\rho(\alpha, \gamma) + \mu\rho(\beta, \gamma), \quad \rho(\alpha, \lambda\beta + \mu\gamma) = \lambda\rho(\alpha, \beta) + \mu\rho(\alpha, \gamma),$$

其中 $\alpha, \beta, \gamma \in V$, $\lambda, \mu \in \mathbb{F}$, 则 ρ 称为 V 上的**双线性函数**.

若存在非零向量 $\alpha \in V$ 使得“ $\rho(\alpha, \beta) = 0, \forall \beta \in V$ ”或者“ $\rho(\beta, \alpha) = 0, \forall \beta \in V$ ”, 则 ρ 称为**退化的**. 否则 ρ 称为**非退化的**.

若对于任意 $\alpha, \beta \in V$ 有 $\rho(\alpha, \beta) = \rho(\beta, \alpha)$, 则 ρ 称为**对称的**. 若 $\rho(\alpha, \alpha) = 0, \forall \alpha \in V$, 则 ρ 称为**反对称的**. 容易证明, 若 ρ 是反对称的, 则有 $\rho(\alpha, \beta) = -\rho(\beta, \alpha), \forall \alpha, \beta \in V$.

若 $\alpha, \beta \in V$ 满足 $\rho(\alpha, \beta) = 0$, 则称 α 与 β **正交**, 记作 $\alpha \perp \beta$.

一般地, 由 $\alpha \perp \beta$ 并不能够推出 $\beta \perp \alpha$. 为了符合传统的几何直观, 我们希望正交关系是对称的, 即 $\alpha \perp \beta \Leftrightarrow \beta \perp \alpha$.

定理 10.32. 设 V 是 \mathbb{F} 上的线性空间, ρ 是 V 上的双线性函数. 若对于任意 $\alpha, \beta \in V$ 都有

$$\rho(\alpha, \beta) = 0 \Leftrightarrow \rho(\beta, \alpha) = 0$$

则“ ρ 是对称的”和“ ρ 是反对称的”两者之一成立.

证明. 假设 ρ 不是反对称的, 则存在 $v \in V$ 满足 $\rho(v, v) \neq 0$. 对于任意 $\alpha \in V$, $u = \alpha - \frac{\rho(\alpha, v)}{\rho(v, v)}v$ 满足 $\rho(u, v) = 0$, 故 $\rho(v, u) = 0$, 得 $\rho(\alpha, v) = \rho(v, \alpha)$. 对于任意 $\alpha, \beta \in V$, 存在 $\lambda, \mu \in \mathbb{F}$ 使得 $\rho(\alpha + \lambda v, \beta + \mu v) = 0$. 再由 $\rho(\beta + \mu v, \alpha + \lambda v) = 0$, 得 $\rho(\alpha, \beta) = \rho(\beta, \alpha)$. 从而 ρ 是对称的. \square

当 V 是有限维时, 设 $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 V 的一组基. $G = (\rho(\alpha_i, \alpha_j)) \in \mathbb{F}^{n \times n}$ 称为 ρ 在 S 下的**矩阵表示**. 对于任意 $\alpha, \beta \in V$, 设 \mathbf{x}, \mathbf{y} 分别是 α, β 在 S 下的坐标, 则有

$$\rho(\alpha, \beta) = \mathbf{x}^T G \mathbf{y}.$$

上式称为 ρ 在 S 下的**坐标表示**.

定理 10.33. 设 V 是 \mathbb{F} 上的有限维线性空间, S_1, S_2 都是 V 的基, P 是从 S_1 到 S_2 的过渡矩阵, V 上的双线性函数 ρ 在 S_1 和 S_2 下的矩阵表示分别是 G_1 和 G_2 , 则 $G_2 = P^T G_1 P$, 即 G_1 与 G_2 在 \mathbb{F} 上相合.

定理 10.34. 设 V 是 \mathbb{F} 上的有限维线性空间, ρ 是 V 上的双线性函数.

1. ρ 是非退化的 $\Leftrightarrow \rho$ 在 V 的任意一组基 S 下的矩阵表示是可逆方阵.

2. 若 ρ 是反对称的, 则存在 V 的基 S , 使得 ρ 在 S 下的矩阵表示形如
$$\begin{pmatrix} O & I_s \\ -I_s & O \\ & & O \end{pmatrix}.$$

3. 若 ρ 是对称的且不是反对称的, 则存在 V 的基 S , 使得 ρ 在 S 下的矩阵表示为对角方阵.

证明. 设 S 是 V 的任意一组基, $G = (g_{ij}) \in \mathbb{F}^{n \times n}$ 是 ρ 在 S 下的矩阵表示. 不妨设 $G \neq O$.

1. ρ 是非退化的 \Leftrightarrow 线性映射 $\mathbf{x} \mapsto \mathbf{x}^T G$ 和 $\mathbf{x} \mapsto G\mathbf{x}$ 都是 \mathbb{F}^n 上的一一映射 $\Leftrightarrow G$ 是可逆方阵.
2. ρ 是反对称的 $\Rightarrow G$ 是反对称方阵且对角元素都是 0. 根据定理 10.33, 只需证明 G 可相合成

$$\begin{pmatrix} O & I_s & \\ -I_s & O & \\ & & O \end{pmatrix}.$$

设 $g_{ij} \neq 0$, 则 G 可相合成 $D_1(\frac{1}{g_{ij}})S_{2j}S_{1i}GS_{i1}S_{j2}D_1(\frac{1}{g_{ij}}) = \begin{pmatrix} A & B \\ -B^T & C \end{pmatrix}$, 进而 G 可相合成 $\begin{pmatrix} A & \\ & \tilde{G} \end{pmatrix}$, 其中 $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\tilde{G} = C + B^T A^{-1} B$ 是反对称方阵且对角元素都是 0. 利用数

学归纳法容易证明, G 可相合成 $\text{diag}(\underbrace{A, \dots, A}_{s \uparrow}, O)$. 因此, G 可相合成 $\begin{pmatrix} O & I_s & \\ -I_s & O & \\ & & O \end{pmatrix}$.

3. ρ 是对称的 $\Rightarrow G$ 是对称方阵. 根据定理 10.33, 只需证明 G 可相合成对角方阵.

情形一: $\text{Char } \mathbb{F} \neq 2$.

- 若存在 $g_{ii} \neq 0$, 则 $S_{1i}GS_{i1}$ 的 $(1,1)$ 元素不是 0. 故不妨设 $g_{11} \neq 0$. $G = \begin{pmatrix} g_{11} & \alpha^T \\ \alpha & B \end{pmatrix}$ 可相合成 $\begin{pmatrix} g_{11} & \\ & \tilde{G} \end{pmatrix}$, 其中 $\tilde{G} = B - \frac{1}{g_{11}}\beta\beta^T$ 是对称方阵. 利用数学归纳法容易证明, G 可相合成对角方阵.
- 若所有 $g_{ii} = 0$, 则存在 $g_{ij} \neq 0$. $T_{ij}(1)GT_{ji}(1)$ 的 (i,i) 元素为 $2g_{ij}$, 化为前述情形.

情形二: $\text{Char } \mathbb{F} = 2$. 由 ρ 不是反对称的, 可知 G 的对角元素不都是 0. 设 $g_{ii} = a \neq 0$, 则 G 可相合成 $S_{1i}GS_{i1} = \begin{pmatrix} a & \alpha^T \\ \alpha & B \end{pmatrix}$, 进而 G 可相合成 $\begin{pmatrix} a & \\ & C \end{pmatrix}$, 其中 $C = B - \frac{1}{a}\alpha\alpha^T$ 是反对称方阵.

- 若 C 的对角元素不都是 0, 利用数学归纳法容易证明, G 可相合成对角方阵.
- 若 C 的对角元素都是 0, 则由结论 2, C 可相合成 $\text{diag}(A, \dots, A, O)$, 其中 $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

$$\text{由 } \begin{pmatrix} 1 & 0 & a \\ 1 & 1 & 0 \\ 1 & 1 & a \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ a & 0 & a \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}, G \text{ 可相合成 } \begin{pmatrix} aI & \\ & O \end{pmatrix}. \quad \square$$

定义 10.14. 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in L(V)$ 可逆.

- V 上的对称非退化双线性函数 ρ 称为内积, 代数结构 (V, ρ) 称为内积空间. 若 \mathcal{A} 满足

$$\rho(\mathcal{A}\alpha, \mathcal{A}\beta) = \rho(\alpha, \beta), \quad \forall \alpha, \beta \in V,$$

则 \mathcal{A} 称为正交变换. 在线性变换的乘积运算下, V 上的正交变换的全体构成群, 称为正交群.

- V 上的反对称非退化双线性函数称为辛内积, 代数结构 (V, ρ) 称为辛空间. 若 \mathcal{A} 满足

$$\rho(\mathcal{A}\alpha, \mathcal{A}\beta) = \rho(\alpha, \beta), \quad \forall \alpha, \beta \in V,$$

则 \mathcal{A} 称为辛变换. 在线性变换的乘积运算下, 辛空间 V 上的辛变换的全体构成群, 称为辛群.

例 10.26. 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in L(V)$, A 是 \mathcal{A} 在 V 的某组基 S 下的矩阵表示.

- 设 ρ 是 V 上的对称非退化双线性函数, G 是 ρ 在 S 下的矩阵表示. \mathcal{A} 是 (V, ρ) 上的正交变换当且仅当 $A^T G A = G$. 特别, 若 \mathcal{A} 是 (V, ρ) 上的正交变换, 则 $\det(A) = \pm 1$.
- 设 ρ 是 V 上的反对称非退化双线性函数, H 是 ρ 在 S 下的矩阵表示. \mathcal{A} 是 (V, ρ) 上的辛变换当且仅当 $A^T H A = H$. 特别, 若 \mathcal{A} 是 (V, ρ) 上的辛变换, 则 $\det(A) = 1$ (见例 4.12).

例 10.27. Minkowski 空间是 Minkowski^[9]于 1907 年为了适应狭义相对论的需要而提出来的时空模型. 一般的 n 维 Minkowski 空间是内积空间 $V_n = (\mathbb{R}^n, \rho)$, 其中 ρ 在标准基下的矩阵表示 $G = \text{diag}(1, \dots, 1, -1)$. 狭义相对论中采用的是 4 维 Minkowski 空间. 在 4 维时空世界中, 每个时空点有 4 个坐标分量 (x, y, z, t) , 其中 x, y, z 代表空间坐标, t 代表时间坐标.

V_4 上的正交变换称为 Lorentz^[10]变换. Lorentz 变换的矩阵表示 A 满足 $A^T G A = G$. 保持 t 分量的正负号不变的 Lorentz 变换称为正常 Lorentz 变换. 狭义相对论要求物理定律在正常 Lorentz 变换下是不变的.

习题

1. 设 V 是 \mathbb{F} 上的线性空间, $\text{Char } \mathbb{F} \neq 2$, ρ 是 V 上的双线性函数. 证明: 存在 V 上的对称双线性函数 f 和反对称双线性函数 g 使得 $\rho = f + g$, 并且这种表示方式是唯一的.
2. 设 V 是 \mathbb{F} 上的线性空间, ρ 是 V 上的双线性函数.
 - (1) 证明: 若 ρ 是反对称的, 则有 $\rho(\alpha, \beta) = -\rho(\beta, \alpha)$, $\forall \alpha, \beta \in V$.
 - (2) 举例: ρ 不是反对称的, 亦有 $\rho(\alpha, \beta) = -\rho(\beta, \alpha)$, $\forall \alpha, \beta \in V$.
3. 证明定理 10.33.
4. 设 ρ 是线性空间 V 上的双线性函数. 证明: 下列两个叙述等价.
 - ① 存在 $f, g \in V^*$ 使得 $\rho(\alpha, \beta) = f(\alpha)g(\beta)$.
 - ② $\forall \alpha_i, \beta_i \in V$, $\begin{vmatrix} \rho(\alpha_1, \beta_1) & \rho(\alpha_1, \beta_2) \\ \rho(\alpha_2, \beta_1) & \rho(\alpha_2, \beta_2) \end{vmatrix} = 0$.
5. 设 \mathbb{F} 是二元域, $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 $\mathbb{F}^{n \times 1}$ 中一组两两不同的向量. 证明:
 - (1) 若对于任意 i, j 都有 $\alpha_i^T \alpha_j = 0$, 则 $m \leq 2^{\lfloor n/2 \rfloor}$.
 - (2) 若对于任意 i, j 都有 $\alpha_i^T \alpha_j = 1$, 则 $m \leq 2^{\lfloor (n-1)/2 \rfloor}$.
6. 设 $V = (\mathbb{F}^{2n}, \rho)$ 是辛空间, 其中辛内积 ρ 在 \mathbb{F}^{2n} 的标准基下的矩阵表示 $G = \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix}$. 研究 V 上的辛变换、自伴变换、斜自伴变换、规范变换的性质.

^[9]Hermann Minkowski, 1864–1909, 德国数学家.

^[10]Hendrik Antoon Lorentz, 1853–1928, 荷兰物理学家、数学家. 获 1902 年 Nobel 物理学奖.

附录 A 多项式

多项式是代数学中最基本的研究对象, 也是最简单的初等函数. 多项式在数学的各个分支中都有着广泛的应用. 多项式与线性代数之间存在非常密切的联系. 许多关于多项式的概念、理论和方法都可以应用于矩阵多项式. 多项式方法是研究矩阵性质和线性变换的重要工具. 多项式空间 $\mathbb{F}[x]$ 也是无限维线性空间的一个典型例子, 对于多项式空间的研究有助于深入理解线性代数理论中的抽象概念. 在本章中, 我们主要介绍任意数域上的多项式理论的基本概念、基本理论、基本方法, 以及它们与线性代数学科的紧密联系.

§A.1 一元多项式

定义 A.1. 设 \mathbb{F} 是一个数域 (见第八章 §8.1 定义 8.2), \mathbb{F} 中的元素称为**常数**, x 是定义在 \mathbb{F} 上的未知数, 称为**变元**. 形如

$$f = c_0 + c_1x + \cdots + c_nx^n$$

的代数式 f 称为 \mathbb{F} 上以 x 为变元的一元多项式, 其中 $c_0, c_1, \dots, c_n \in \mathbb{F}$, $n \in \mathbb{N}$. 当 $c_n \neq 0$ 时, c_nx^n 称为 f 的首项, 记作 $\text{lt}(f)$; c_n 称为 f 的首项系数, 记作 $\text{lc}(f)$; n 称为 f 的次数, 记作 $\text{deg}(f)$. 约定 $x^0 = 1$, $\text{deg}(0) = -\infty$.

\mathbb{F} 上以 x 为变元的一元多项式的全体记作 $\mathbb{F}[x]$. 在 $\mathbb{F}[x]$ 上可以自然地定义加法和乘法运算

$$\sum_{i=0}^n a_i x^i + \sum_{i=0}^n b_i x^i = \sum_{i=0}^n (a_i + b_i) x^i, \quad \sum_{i=0}^m a_i x^i \cdot \sum_{j=0}^n b_j x^j = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j \right) x^k$$

其中 $a_i, b_j \in \mathbb{F}$, $m, n \in \mathbb{N}$. 容易验证, 在多项式的加法和乘法运算下, $\mathbb{F}[x]$ 满足定义 A.2 中的运算律, 构成一个整环. $\mathbb{F}[x]$ 称为 \mathbb{F} 上以 x 为变元的一元多项式环.

定义 A.2. 设 R 是定义了加法运算 $a + b$ 和乘法运算 $a * b$ 的非空集合. 若下列运算律成立,

- | | |
|-------------|----------------------------------------------------------|
| (A1) 加法结合律 | $\forall a, b, c \in R, (a + b) + c = a + (b + c)$ |
| (A2) 加法交换律 | $\forall a, b \in R, a + b = b + a$ |
| (A3) 加法有单位元 | 存在零元素 $0 \in R$ 使得 $\forall a \in R, 0 + a = a$ |
| (A4) 加法有逆元 | $\forall a \in R$, 存在 $b \in R$ 使得 $a + b = 0$ |
| (M1) 乘法结合律 | $\forall a, b, c \in R, (a * b) * c = a * (b * c)$ |
| (M2) 乘法交换律 | $\forall a, b \in R, a * b = b * a$ |
| (M3) 乘法有单位元 | 存在幺元素 $1 \in R$ 使得 $\forall a \in R, 1 * a = a$ |
| (M4) 零元素无因子 | 若 $a, b \in R$ 使得 $a * b = 0$, 则 $a = 0$ 或 $b = 0$ |
| (D1) 加乘分配律 | $\forall a, b, c \in R, (a + b) * c = (a * c) + (b * c)$ |

则代数结构 $(R, +, *)$ 称为**整环**. 在不引起歧义的情况下, 可以简称 R 是整环.

例如, 任意数域 \mathbb{F} 是整环, 整数集 \mathbb{Z} 是整环. 在实数的加法和乘法运算下, $\{a + \sqrt{2}b \mid a, b \in \mathbb{Z}\}$ 构成整环.

把定义 A.1 中的数域 \mathbb{F} 换成整环 R , 可得 R 上的一元多项式环 $R[x]$ 的概念. 容易验证, 在多项式的加法和乘法运算下, $R[x]$ 也构成整环. $R[x]$ 与 $\mathbb{F}[x]$ 具有许多类似的性质, 也有许多不同的性质, $\mathbb{F}[x]$ 的性质不能生搬硬套至 $R[x]$.

定理 A.1. 设 $f, g \in \mathbb{F}[x]$. 有如下常用结论.

1. $\deg(f + g) \leq \max(\deg(f), \deg(g))$.
2. 当 $\deg(f) \neq \deg(g)$ 时, $\deg(f + g) = \max(\deg(f), \deg(g))$.
3. $\deg(fg) = \deg(f) + \deg(g)$.
4. 若 $fg = 0$, 则 $f = 0$ 或 $g = 0$.
5. 当 $g \neq 0$ 时, $\deg(f(g)) = \deg(f) \deg(g)$.

下面考虑 $\mathbb{F}[x]$ 上的除法运算. 设 $f, g \in \mathbb{F}[x]$ 且 $g \neq 0$. 代数式 $\frac{f}{g}$ 称为**有理分式**或**有理函数**. \mathbb{F} 上以 x 为变元的有理函数的全体记作 $\mathbb{F}(x)$. 在 $\mathbb{F}(x)$ 上可以自然地定义加法和乘法运算

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} = \frac{f_1g_2 + f_2g_1}{g_1g_2}, \quad \frac{f_1}{g_1} \cdot \frac{f_2}{g_2} = \frac{f_1f_2}{g_1g_2}$$

及其逆运算减法和除法. 容易验证, 在有理函数的加法和乘法运算下, $\mathbb{F}(x)$ 满足定义 8.2 中的运算律, 构成一个数域. $\mathbb{F}(x)$ 称为 \mathbb{F} 上以 x 为变元的**有理函数域**.

定理 A.2. 设 $f, g \in \mathbb{F}[x]$ 且 $g \neq 0$, 则存在唯一的 $q, r \in \mathbb{F}[x]$ 使得

$$f = qg + r \quad \text{且} \quad \deg(r) < \deg(g). \quad (\text{A.1})$$

证明. (存在性) 对 $\deg(f)$ 归纳. 当 $\deg(f) < \deg(g)$ 时, $q = 0$ 和 $r = f$ 满足题设. 当 $\deg(f) \geq \deg(g)$ 时, $h = f - \frac{\text{lt}(f)}{\text{lt}(g)}g \in \mathbb{F}[x]$ 满足 $\deg(h) < \deg(f)$. 根据归纳假设, 存在 $p, r \in \mathbb{F}[x]$ 使得 $h = pg + r$ 且 $\deg(r) < \deg(g)$. 因此, $q = p + \frac{\text{lt}(f)}{\text{lt}(g)}$ 和 r 满足题设.

(唯一性) 若 q_1, r_1 和 q_2, r_2 都满足题设, 则由 $q_1g + r_1 = q_2g + r_2$, 得 $\deg((q_1 - q_2)g) = \deg(r_2 - r_1) < \deg(g)$. 根据定理 A.1 结论 3, $q_1 - q_2 = r_1 - r_2 = 0$. \square

定义 A.3. (A.1) 式称为 f/g 的**带余除法**, q 和 r 分别称为 f/g 的**商式**和**余式**, 记作

$$q = \text{quo}(f, g), \quad r = \text{rem}(f, g).$$

设 $f_1, f_2 \in \mathbb{F}[x]$. 若 $\text{rem}(f_1, g) = \text{rem}(f_2, g)$, 则称 f_1 与 f_2 模 g **同余**, 记作

$$f_1 \equiv f_2 \pmod{g}.$$

例 A.1. 设 $f \in \mathbb{F}[x]$, $g = (x - a)^n$, 其中 $\text{Char } \mathbb{F} = 0$, $a \in \mathbb{F}$, $n \geq 1$. 由 Taylor 展开公式, 可得

$$\text{rem}(f, g) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x - a)^k.$$

特别, $\text{rem}(f, x - a) = f(a)$.

例 A.2. 设 $f \in \mathbb{F}[x]$, $g = (x - a_1) \cdots (x - a_n)$, 其中 $a_1, \dots, a_n \in \mathbb{F}$ 两两不同, $n \geq 2$.

- 由 Lagrange 插值公式, 可得

$$\text{rem}(f, g) = \sum_{i=1}^n f(a_i) \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}.$$

- 由 Newton 插值公式, 可得

$$\text{rem}(f, g) = \sum_{k=1}^n f[a_1, \dots, a_k] (x - a_1) \cdots (x - a_{k-1})$$

其中 $f[a_1, \dots, a_k]$ 称为 $k-1$ 阶差商, 可由下式递推定义:

$$f[x_1] = f(x_1), \quad f[x_1, \dots, x_k] = \frac{f[x_2, \dots, x_k] - f[x_1, \dots, x_{k-1}]}{x_k - x_1}, \quad k \geq 2.$$

例 A.3. 设实系数多项式 $f = x + 2x^2 + 3x^3 + \cdots + nx^n$, $g = 1 + x + x^2$. 求 $\text{rem}(f, g)$.

解答. 设 $\text{rem}(f, g) = a + bx$, 其中 $a, b \in \mathbb{R}$. 把 $w = \frac{-1 + \sqrt{3}i}{2}$ 代入等式, 得

$$a + bw = f(w) = \frac{nw^{n+2} - (n+1)w^{n+1} + w}{(w-1)^2} = \begin{cases} k - kw, & n = 3k; \\ k + (2k+1)w, & n = 3k+1; \\ -2(k+1) - (k+1)w, & n = 3k+2. \end{cases}$$

$$\text{因此, } \text{rem}(f, g) = \begin{cases} \frac{n}{3}(1-x), & n \equiv 0 \pmod{3}; \\ \frac{n-1}{3} + \frac{2n+1}{3}x, & n \equiv 1 \pmod{3}; \\ -\frac{n+1}{3}(2+x), & n \equiv 2 \pmod{3}. \end{cases} \quad \square$$

接下来考虑 $\mathbb{F}[x]$ 上的整除问题.

定义 A.4. 设 $f, f_1, \dots, f_k, g \in \mathbb{F}[x]$.

- 若 $g \neq 0$ 并且 $\text{rem}(f, g) = 0$, 则称 f 是 g 的倍式, g 是 f 的因式, 或者称 f 被 g 整除, g 整除 f , 记作 $g \mid f$.
- 若 f_1, \dots, f_k 不全为 0, $g \neq 0$ 并且 g 是每个 f_i 的因式, 则 g 称为 f_1, \dots, f_k 的公因式. 次数最大的公因式称为最大公因式. 若 f_1, \dots, f_k 的所有公因式 $g \in \mathbb{F}$, 则称 f_1, \dots, f_k 互素.
- 若 f_1, \dots, f_k 全不为 0, 并且 g 是每个 f_i 的倍式, 则 g 称为 f_1, \dots, f_k 的公倍式. 次数最小的非零公倍式称为最小公倍式.

根据定义 A.4, 0 是任何非零多项式的倍式, 0 不是任何多项式的因式. 最大公因式和最小公倍式都有可能不唯一.

定理 A.3. 设 $f, g, h, u, v \in \mathbb{F}[x]$.

1. 若 $f \mid g$ 且 $g \mid h$, 则 $f \mid h$.

2. 若 $f \mid g$ 且 $f \mid h$, 则 $f \mid ug + vh$.
3. 若 $f \mid g$ 且 $g \mid f$, 则存在非零常数 $\lambda \in \mathbb{F}$ 使得 $f = \lambda g$.

定理 A.4. 设 $f_1, \dots, f_k, g, h \in \mathbb{F}[x]$.

1. 若 g 是 f_1, \dots, f_k 的最小公倍式, h 是 f_1, \dots, f_k 的公倍式, 则 $g \mid h$.
2. 若 g 是 f_1, \dots, f_k 的最大公因式, h 是 f_1, \dots, f_k 的公因式, 则 $h \mid g$.

证明.

1. 易知 $\text{rem}(h, g)$ 也是 f_1, \dots, f_k 的公倍式. 由 $\deg(g)$ 的最小性, 得 $\text{rem}(h, g) = 0$, 即 $g \mid h$.
2. 设 $p = \prod_{i=1}^k f_i$, 则 p/g 是 $p/f_1, \dots, p/f_k$ 的最小公倍式, p/h 是 $p/f_1, \dots, p/f_k$ 的公倍式. 由结论 1, 得 p/g 整除 p/h , 从而 $h \mid g$. \square

在 $\mathbb{F}[x]$ 中, 两个多项式的整除关系 $f \mid g$ 构成一个偏序. 定理 A.4 说明: 在 f_1, \dots, f_k 的所有公因式构成的集合中, 最大公因式不仅次数最大, 也是整除偏序下的最大元素; 在 f_1, \dots, f_k 的所有公倍式中, 最小公倍式不仅次数最小, 也是整除偏序下的最小元素. 许多教科书通过整除偏序来定义最大公因式和最小公倍式.

根据定理 A.3 结论 3 和定理 A.4, f_1, \dots, f_k 的任意两个最大公因式 (最小公倍式) 相互整除, 仅相差一个非零常数因子. 因此, f_1, \dots, f_k 的首一最大公因式和首一最小公倍式都是唯一的, 分别记作 $\text{gcd}(f_1, \dots, f_k)$ 和 $\text{lcm}(f_1, \dots, f_k)$.

定理 A.5. 设 $f, g \in \mathbb{F}[x]$ 都是首一多项式, 则 $\text{gcd}(f, g)\text{lcm}(f, g) = fg$.

证明. 设 $u = fg/\text{gcd}(f, g)$, $v = fg/\text{lcm}(f, g)$. 一方面, 由 $f \mid u$ 和 $g \mid u$, 得 $\text{lcm}(f, g) \mid u$, 故 $\text{gcd}(f, g)\text{lcm}(f, g) \mid fg$. 另一方面, 由 $v \mid f$ 和 $v \mid g$, 得 $v \mid \text{gcd}(f, g)$, 故 $fg \mid \text{gcd}(f, g)\text{lcm}(f, g)$. 因此, $\text{gcd}(f, g)\text{lcm}(f, g) = fg$. \square

定义 A.5. 设 $f, g \in \mathbb{F}[x]$ 且 $g \neq 0$,

$$r_0 = f, \quad r_1 = g, \quad r_{i+1} = \text{rem}(r_{i-1}, r_i), \quad i = 1, \dots, k \quad (\text{A.2})$$

直至 $r_{k+1} = 0$, 则多项式序列 $\{r_i\}_{i=0}^k$ 称为 f/g 的辗转相除序列.

定理 A.6. $\text{gcd}(f, g) = \frac{r_k}{\text{lc}(r_k)}$.

证明. (A.2) 式可以表示为矩阵乘积的形式

$$\begin{pmatrix} r_{i-1} \\ r_i \end{pmatrix} = \begin{pmatrix} q_i & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_i \\ r_{i+1} \end{pmatrix}, \quad \begin{pmatrix} r_i \\ r_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \begin{pmatrix} r_{i-1} \\ r_i \end{pmatrix}.$$

其中 $q_i = \text{quo}(r_{i-1}, r_i)$. 由此可得,

$$\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} q_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} q_k & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_k \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} r_k \\ 0 \end{pmatrix}, \quad (\text{A.3})$$

$$\begin{pmatrix} r_k \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_k \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}. \quad (\text{A.4})$$

一方面, 由 (A.3) 式可得 $f = a_{11}r_k$, $g = a_{21}r_k$, 故 $r_k \mid \text{gcd}(f, g)$. 另一方面, 由 (A.4) 式可得 $r_k = b_{11}f + b_{12}g$, 故 $\text{gcd}(f, g) \mid r_k$. 因此, $r_k = \text{lc}(r_k)\text{gcd}(f, g)$. \square

定理 A.6 的证明过程附带证明了

定理 A.7 (Bezout 定理). 设 $f, g \in \mathbb{F}[x]$ 不全为 0, 则存在 $u, v \in \mathbb{F}[x]$ 使得

$$uf + vg = \gcd(f, g). \quad (\text{A.5})$$



关于定理 A.6 证明的几点注记:

- 多项式方阵 $\begin{pmatrix} q_i & 1 \\ 1 & -q_i \end{pmatrix}$ 和 $\begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix}$ 可以看作是某种“初等方阵”, 称为初等模方阵.
- 辗转相除过程可以看作是对 2 维多项式向量实施某种“初等变换”, 称为初等模变换.
- $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ 和 $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ 互为逆矩阵, 称为 $\mathbb{F}[x]$ 上的模方阵. 见本书 §4.3.
- $\det(A) = \det(B) = (-1)^k$, $\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (-1)^k \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.
- (A.2) 式称为计算 $\gcd(f, g)$ 的辗转相除法或 **Euclid 算法**.
- (A.4) 式称为计算 Bezout 等式 (A.3) 的扩展 **Euclid 算法**.
- 当已知 $\gcd(f, g)$ 时, 也可以用待定系数法求解 u, v 使得 Bezout 等式成立. 设

$$f = \sum_{i=0}^m a_i x^i, \quad g = \sum_{i=0}^n b_i x^i, \quad \gcd(f, g) = \sum_{i=0}^k c_i x^i, \quad u = \sum_{i=0}^p u_i x^i, \quad v = \sum_{i=0}^q v_i x^i$$

其中 $p = n - k - 1$, $q = m - k - 1$, 则 Bezout 等式可以化成线性方程组

$$\begin{pmatrix} a_0 & & & b_0 & & & \\ \vdots & \ddots & & \vdots & \ddots & & \\ a_m & \vdots & a_0 & b_n & \vdots & b_0 & \\ & \ddots & \vdots & & \ddots & \vdots & \\ & & a_m & & & b_n & \end{pmatrix} \begin{pmatrix} u_0 \\ \vdots \\ u_p \\ v_0 \\ \vdots \\ v_q \end{pmatrix} = \begin{pmatrix} c_0 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

相关内容可见本书第三章 §3.2 习题 13 和第四章 §4.1 习题 10.

例 A.4. 设 $f = x^6 - x^3 - 2$, $g = x^4 - 1$, 求 u, v 使得 $uf + vg = \gcd(f, g)$.

解答. 计算 f/g 的辗转相除序列 $\{r_i\}$, 设 $r_i = u_i f + v_i g$, 则 $(u_{i+1}, v_{i+1}) = (u_{i-1}, v_{i-1}) - q_i(u_i, v_i)$, 其中 $q_i = \text{quo}(r_{i-1}, r_i)$. 具体过程如下表所示.

i	u_i	v_i	r_i	q_i
0	1	0	$x^6 - x^3 - 2$	
1	0	1	$x^4 - 1$	x^2
2	1	$-x^2$	$-x^3 + x^2 - 2$	$-x - 1$
3	$x + 1$	$-x^3 - x^2$	$x^2 - 2x - 3$	$-x - 1$
4	$x^2 + 2x + 2$	$-x^4 - 2x^3 - 2x^2 + x + 1$	$-5x - 5$	$-\frac{1}{5}(x - 3)$
5	$\frac{1}{5}(x^3 - x^2 + x - 1)$	$\frac{1}{5}(-x^5 + x^4 - x^3 + 2x^2 - 2)$	0	

因此, 所求 $u = -\frac{1}{5}(x^2 + 2x + 2)$, $v = \frac{1}{5}(x^4 + 2x^3 + 2x - x - 1)$, $\gcd(f, g) = x + 1$. \square

由 Bezout 定理可得以下常用结论.

定理 A.8. 设 $f, g \in \mathbb{F}[x]$ 全不为 0, 并且 $\gcd(f, g) = 1$.

1. 存在 $h \in \mathbb{F}[x]$ 使得 $gh \equiv 1 \pmod{f}$.
2. 对于任意 $h \in \mathbb{F}[x]$, $\gcd(f, gh) = \gcd(f, h)$.
3. 对于任意 $h \in \mathbb{F}[x]$, 若 $f \mid h$ 且 $g \mid h$, 则 $fg \mid h$.

证明. 设 $u, v \in \mathbb{F}[x]$ 使得 $uf + vg = 1$.

1. $h = v$ 满足题设.
2. 由 $\gcd(f, h)$ 是 f 和 gh 的公因式, 得 $\gcd(f, h) \mid \gcd(f, gh)$. 由 $h = ufh + vgh$ 是 $\gcd(f, gh)$ 的倍式, 得 $\gcd(f, gh) \mid \gcd(f, h)$. 因此, $\gcd(f, gh) = \gcd(f, h)$.
3. 设 $h = af = bg$, 则 $h = (uf + vg)h = (bu + av)fg$. 因此, $fg \mid h$. \square

定理 A.9 (中国剩余定理). 设 $p_1, \dots, p_k, r_1, \dots, r_k \in \mathbb{F}[x]$, 并且 p_1, \dots, p_k 两两互素, 则存在 $f \in \mathbb{F}[x]$ 使得 $f \equiv r_i \pmod{p_i}, \forall i$.

证明. $\forall i, q_i = \prod_{j \neq i} p_j$ 与 p_i 互素, 根据 Bezout 定理, 存在 $u_i, v_i \in \mathbb{F}[x]$ 使得 $u_i p_i + v_i q_i = 1$. 容易验证, $f = \sum_{i=1}^k v_i q_i r_i$ 满足题设. \square

接下来考虑 $\mathbb{F}[x]$ 中的因式分解问题, 即把一个多项式分解成若干个非平凡因式的乘积形式, 并且每个因式无法作进一步的分解.

定义 A.6. 设 $f \in \mathbb{F}[x]$ 且 $f \neq 0$. 若存在 $g, h \in \mathbb{F}[x] \setminus \mathbb{F}$ 使得 $f = gh$, 则称 f 在 $\mathbb{F}[x]$ 中是**可约的**; 否则称 f 在 $\mathbb{F}[x]$ 中是**不可约的**. 在不引起歧义的情况下, 可简称“ f 是可约的”或“ f 是不可约的”. 约定 0 是可约的.

多项式的可约性与其所在的多项式环有关. 例如, $f = x^2 + 1$ 在 $\mathbb{R}[x]$ 中是不可约的, 在 $\mathbb{C}[x]$ 中是可约的. 关于多项式的可约性, 有以下常用结论.

定理 A.10. 设 $f \in \mathbb{F}[x]$ 是不可约的.

1. 对于任意 $g \in \mathbb{F}[x]$, $\gcd(f, g) = 1$ 或者 $f \mid g$.
2. 对于任意 $g_1, \dots, g_k \in \mathbb{F}[x]$, 若 $f \mid g_1 \cdots g_k$, 则 f 整除某个 g_i .

证明.

1. 设 $d = \gcd(f, g)$, 则 $d \in \mathbb{F}$ 或 $f/d \in \mathbb{F}$. 前者, 得 $d = 1$. 后者, 得 $f = \text{lc}(f)d$ 整除 g .
2. 假设 f 不整除任意 g_i . 由结论 1 和 Bezout 定理, 存在 $u_i, v_i \in \mathbb{F}[x]$ 使得 $u_i f + v_i g_i = \gcd(f, g_i) = 1$. 从而 $\prod_{i=1}^k (u_i f + v_i g_i) = 1$ 可被 f 整除. 矛盾. \square

定理 A.11. 对于任意 $f \in \mathbb{F}[x] \setminus \mathbb{F}$, 存在不可约的首一多项式 $f_1, \dots, f_k \in \mathbb{F}[x] \setminus \mathbb{F}$ 使得 $f = \text{lc}(f) \prod_{i=1}^k f_i$. 不考虑因式的顺序, 这种分解方式是唯一的.

证明. 由定义可知分解方式的存在性. 下证唯一性. 假设 $f_1, \dots, f_k, g_1, \dots, g_t \in \mathbb{F}[x] \setminus \mathbb{F}$ 都是不可约的首一多项式, 并且满足 $\prod_{i=1}^k f_i = \prod_{i=1}^t g_i$. 根据定理 A.10 结论 2, f_1 整除某个 g_i . 不妨设 $i = 1$, 则 $f_1 = g_1$, $\prod_{i=2}^k f_i = \prod_{i=2}^t g_i$. 重复上述过程, 可得 $k = t$ 并且 f_1, \dots, f_k 是 g_1, \dots, g_t 的排列. \square

多项式的可约性与其根 (零点) 有非常紧密的联系.

例 A.5 (域的扩张). 设 $f \in \mathbb{F}[x]$ 是 n 次不可约多项式, $n \geq 2$. 易知不存在 $\lambda \in \mathbb{F}$ 使得 $f(\lambda) = 0$. 下面构造数域 $\mathbb{K} \supset \mathbb{F}$ 使得方程 $f(\lambda) = 0$ 有解 $\lambda \in \mathbb{K}$. 这样的 \mathbb{K} 称为 \mathbb{F} 的 n 次代数扩域.

设 t 是变元, $\mathbb{K} = \{g \in \mathbb{F}[t] \mid \deg(g) < n\}$. 在 \mathbb{K} 上定义加法和乘法运算

$$g_1 + g_2 = \text{rem}(g_1 + g_2, f(t)), \quad g_1 * g_2 = \text{rem}(g_1 g_2, f(t)).$$

根据定理 A.10 和定理 A.8 结论 1, 当 $g \in \mathbb{K}$ 且 $g \neq 0$ 时, $\gcd(f(t), g) = 1$, g 有乘法逆元 $h \in \mathbb{K}$ 使得 $g * h = 1$. 容易验证, $(\mathbb{K}, +, *)$ 满足定义 8.2 中的运算律, 构成数域.

\mathbb{K} 中的元素 $\lambda = t$ 满足 $f(\lambda) = 0$. 故 f 可以因式分解为 $f = (x - \lambda)g$, 其中

$$g = \frac{f(x) - f(t)}{x - t} = \sum_{i,j=0}^{n-1} g_{ij} t^i x^j \in \mathbb{K}[x].$$

定理 A.12 (代数基本定理). 任意 $n \geq 1$ 次多项式 $f \in \mathbb{C}[x]$ 有根 $z \in \mathbb{C}$.

尽管形式上是代数, 代数基本定理的证明本质上都是拓扑的, 离不开实函数和复函数的连续性概念. 有兴趣的读者可查阅相关文献, 此处就不给出证明了. 通常认为 Gauss 于 1799 年给出了代数基本定理的第一个严格的证明. 根据代数基本定理, 可得以下推论.

定理 A.13. 设 $f \in \mathbb{C}[x]$ 是不可约的, 则 $\deg(f) \leq 1$.

证明. 假设 $\deg(f) \geq 2$, 则存在 $z \in \mathbb{C}$ 使得 $f(z) = 0$. 故 $q = f/(x - z) \in \mathbb{C}[x]$, $\deg(q) \geq 1$, 与 f 的不可约性矛盾. \square

定理 A.14. 设 $f \in \mathbb{R}[x]$ 是不可约的, 则 $\deg(f) \leq 1$ 或者 “ $\deg(f) = 2$ 且 f 没有实数根”.

证明. 假设 $\deg(f) \geq 1$, 则存在 $z \in \mathbb{C}$ 使得 $f(z) = 0$.

(1) 若 $z \in \mathbb{R}$, 则 $x - z$ 整除 f . 由 f 的不可约性, 得 $f = c(x - z)$, c 是非零实数.

(2) 若 $z \notin \mathbb{R}$, 则由 $f(z) = 0$ 可得 $f(\bar{z}) = 0$. 设 $g = (x - z)(x - \bar{z}) \in \mathbb{R}[x]$. 由 $g \mid f$ 和 f 的不可约性, 得 $f = cg$, c 是非零实数. \square

对于有理数域上的多项式, 判断其是否可约通常是比较困难的, 仅有部分的结果. 一种常用方法是: 首先把 $f \in \mathbb{Q}[x]$ 表示成 g/λ 的形式, 其中 $\lambda \in \mathbb{Z}$, $g \in \mathbb{Z}[x]$; 于是 f 在 $\mathbb{Q}[x]$ 中的可约性等价于 g 在 $\mathbb{Z}[x]$ 中的可约性; 然后, 任取素数 p , 把 g 对应到 p 元域 \mathbb{Z}_p 上的多项式 $\tilde{g} = g \bmod p$; 若在模 p 运算下 \tilde{g} 是不可约的, 则 g 在 $\mathbb{Z}[x]$ 中是不可约的; 若在模 p 运算下 \tilde{g} 是可约的, 则 g 在 $\mathbb{Z}[x]$ 中的可约性还需进一步的检验.

定理 A.15. 设 $f \in \mathbb{Z}[x]$. 若 f 在 $\mathbb{Q}[x]$ 中是可约的, 则存在 $g, h \in \mathbb{Z}[x]$ 使得 $f = gh$ 并且 $\deg(g) \geq 1$, $\deg(h) \geq 1$.

证明. 不妨设 $\deg(f) \geq 2$. 由题设知, 存在 $f_1, f_2 \in \mathbb{Q}[x]$ 使得 $f = f_1 f_2$ 并且 $\deg(f_1) = k \geq 1$, $\deg(f_2) = n - k \geq 1$. 故存在 $\lambda \in \mathbb{Q}$ 和 $a_0, \dots, a_k, b_0, \dots, b_{n-k} \in \mathbb{Z}$ 使得 $f = \lambda \sum_{i=0}^k a_i x^i \sum_{i=0}^{n-k} b_i x^i$ 并且

$$\gcd(a_0, a_1, \dots, a_k) = \gcd(b_0, b_1, \dots, b_{n-k}) = 1. \quad (*)$$

记 $\sum_{i=0}^k a_i x^i \sum_{i=0}^{n-k} b_i x^i = \sum_{i=0}^n c_i x^i$, $d = \gcd(c_0, c_1, \dots, c_n)$. 假设 d 有素因子 p , i, j 是满足 $p \nmid a_i$, $p \nmid b_j$ 的最小非负整数, 则 $l = i + j$ 是满足 $p \nmid c_l$ 的最小非负整数. 由于 l 不存在, 只可能 $p \mid a_i, \forall i$ 或 $p \mid b_i, \forall i$, 与 (*) 矛盾. 故 $d = 1$, 得 $\lambda \in \mathbb{Z}$. 因此, f 在 $\mathbb{Z}[x]$ 中可约. \square

例 A.6 (Eisenstein 准则^[1]). 设 $f = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$. 若存在素数 p 满足

$$p \mid a_0, p \mid a_1, \dots, p \mid a_{n-1}, p \nmid a_n, p^2 \nmid a_0$$

则 f 在 $\mathbb{Q}[x]$ 中是不可约的.

证明. 假设 f 在 $\mathbb{Q}[x]$ 中是可约的. 根据定理 A.15, 存在 $g, h \in \mathbb{Z}[x]$ 使得 $f = gh$ 并且 $\deg(g) \geq 1$, $\deg(h) \geq 1$. 由 $f \equiv a_n x^n \pmod{p}$ 可得 $g \equiv \lambda x^k \pmod{p}$, $h \equiv \mu x^{n-k} \pmod{p}$, 其中 $1 \leq k \leq n-1$. 因此, $g(0) \equiv h(0) \equiv 0 \pmod{p}$, 与 $p^2 \nmid a_0 = g(0)h(0)$ 矛盾. \square

例 A.7. 设 p 是素数. $f = 1 + x + \dots + x^{p-1}$ 在 $\mathbb{Q}[x]$ 中是不可约的.

证明. 考虑 $g = f(x+1) = \sum_{k=1}^p C_p^k x^{k-1}$. 当 $1 \leq k \leq p-1$ 时, $p \mid C_p^k = \frac{p!}{k!(p-k)!}$. 此外, $p \nmid C_p^p = 1$, $p^2 \nmid C_p^1 = p$. 根据 Eisenstein 准则, g 在 $\mathbb{Q}[x]$ 中是不可约的. 因此, f 在 $\mathbb{Q}[x]$ 中是不可约的. \square

例 A.8. 设 n 是正整数, $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$.

- $1, \omega, \dots, \omega^{n-1}$ 是方程 $x^n = 1$ 的所有复数根, 称为 n 次单位根.
- 当 $\gcd(k, n) = 1$ 时, $z = \omega^k$ 满足 $1, z, z^2, \dots, z^{n-1}$ 两两不同, z 称为 n 次本原单位根.
- $\Phi_n(x) = \prod_{\substack{1 \leq k \leq n \\ \gcd(k, n) = 1}} (x - \omega^k)$ 以所有 n 次本原单位根为根, 称为 n 次分圆多项式. 例如,

$\Phi_1(x) = x - 1$	$\Phi_2(x) = x + 1$
$\Phi_3(x) = x^2 + x + 1$	$\Phi_4(x) = x^2 + 1$
$\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$	$\Phi_6(x) = x^2 - x + 1$
$\Phi_7(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$	$\Phi_8(x) = x^4 + 1$
$\Phi_9(x) = x^6 + x^3 + 1$	$\Phi_{10}(x) = x^4 - x^3 + x^2 - x + 1$

易知, 当 p 是素数时, $\Phi_p(x) = \frac{x^p - 1}{x - 1} = 1 + x + x^2 + \dots + x^{p-1}$. 关于分圆多项式, 还有下列常用结论: (1) $x^n - 1 = \prod_{d \mid n} \Phi_d(x)$. (2) $\Phi_n(x) \in \mathbb{Z}[x]$. (3) $\Phi_n(x)$ 在 $\mathbb{Q}[x]$ 中是不可约的. 证明留作习题.

^[1]1846 年, 德国数学家 Theodor Schönemann 在 Journal für die reine und angewandte Mathematik 上首先发表了此结果. 1850 年, 德国数学家 Gotthold Eisenstein 重新发现了这一结果, 也发表在 Journal für die reine und angewandte Mathematik 上.

一般地, 设 $f \in \mathbb{F}[x]$, 为了判断 f 在 $\mathbb{F}[x]$ 中是否可约, 我们可以首先在 \mathbb{F} 的某个充分大的扩域上把 $f(x)$ 完全分解为一次因式的乘积, 然后检验是否可以把某些因式相乘为 $\mathbb{F}[x]$ 中的多项式.

例 A.9. $f = x^{12} - 1$ 在 $\mathbb{C}[x]$ 中可以分解为

$$f = \prod_{k=0}^{11} (x - e^{\frac{k\pi i}{6}}) = (x-1)(x - e^{\frac{\pi i}{6}})(x - e^{\frac{2\pi i}{6}})(x-i)(x - e^{\frac{4\pi i}{6}})(x - e^{\frac{5\pi i}{6}}) \\ \cdot (x+1)(x - e^{\frac{7\pi i}{6}})(x - e^{\frac{8\pi i}{6}})(x+i)(x - e^{\frac{10\pi i}{6}})(x - e^{\frac{11\pi i}{6}}).$$

由此可得, f 在 $\mathbb{R}[x]$ 中的不可约因式分解

$$f = (x-1)(x+1)(x^2+1)(x^2-x+1)(x^2+x+1)(x^2-\sqrt{3}x+1)(x^2+\sqrt{3}x+1)$$

和 f 在 $\mathbb{Q}[x]$ 中的不可约因式分解

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

习题

1. 证明定理 A.1 和 A.3.

2. 设 $f, g \in \mathbb{F}[x]$, $f = \sum_{i=0}^m a_i x^i$, $g = \sum_{j=0}^n b_j x^j$, $fg = \sum_{k=0}^{m+n} c_k x^k$, $a_i, b_j, c_k \in \mathbb{F}$. 证明:

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_{m+n} \end{pmatrix} = \begin{pmatrix} a_0 & \cdots & a_m \end{pmatrix} \begin{pmatrix} b_0 & b_1 & \cdots & b_n & & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & b_0 & b_1 & \cdots & b_n \end{pmatrix}.$$

3. 对于 $\mathbb{F}[x]$ 上的矩阵 $A = (a_{ij})$, 定义 $\deg(A) = \max_{i,j} \deg(a_{ij})$. 证明:

$$\deg(A+B) \leq \max(\deg(A), \deg(B)), \quad \deg(AB) \leq \deg(A) + \deg(B).$$

4. 设正整数 $n \geq 2$. 证明: 当且仅当 n 是素数时, $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$ 在整数模 n 的加法和乘法运算下构成整环.

5. 设 n 是正整数, $f = x^{2n} - x^n + 1$, $g = x^2 - x + 1$. 求 $\text{rem}(f, g)$.

6. 求所有 $f \in \mathbb{R}[x]$ 使得 $f \equiv x-1 \pmod{x^2-x+1}$ 且 $f \equiv x+1 \pmod{x^2+x+1}$.

7. 给定 $n \in \mathbb{N}^*$. 求所有 $f \in \mathbb{R}[x]$ 使得 $(x-1)^n$ 整除 $f+1$ 且 $(x+1)^n$ 整除 $f-1$.

8. 求 $a, b \in \mathbb{R}$ 使得 f 可被 g 整除, 其中

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

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

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

9. 设 $f, g \in \mathbb{R}[x]$, 求 $\gcd(f, g)$, 其中

$$(1) f = x^5 + 2x^4 + x^2 - x - 1, \quad g = x^5 - x^4 + 2x^3 + x + 1.$$

$$(2) f = 2x^5 + 3x^4 - 5x^3 + 3x^2 - x - 2, \quad g = 2x^5 + x^4 - 2x - 1.$$

$$(3) f = x^5 - 3x^4 - 2x^3 + 4x^2 + 8x + 4, \quad g = x^5 - 4x^4 + 9x^2 + 2x - 2.$$

10. 求次数最小的 $u, v \in \mathbb{R}[x]$ 使得 $uf + vg = 1$, 其中

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

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

$$(3) f = (1 - x)^n, \quad g = x^n.$$

11. 设 $f_1, \dots, f_k \in \mathbb{F}[x]$ 不全为 0. 证明: 存在 $u_1, \dots, u_k \in \mathbb{F}[x]$ 使得

$$u_1 f_1 + \dots + u_k f_k = \gcd(f_1, \dots, f_k).$$

12. 设 $f, g \in \mathbb{F}[x]$ 都不是常数, a_{ij} 如 (A.3) 式所述. 证明:

$$\deg(a_{11}) = \deg(f) - \deg(r_k) > \deg(a_{12}),$$

$$\deg(a_{21}) = \deg(g) - \deg(r_k) > \deg(a_{22}).$$

13. 设 $a, b, c, d \in \mathbb{F}$ 两两不同, 并且 $a + b \neq c + d$. 证明:

$$f = (x - a)(x - b)(x - c)(x - d) \left(\frac{1}{x - a} + \frac{1}{x - b} - \frac{1}{x - c} - \frac{1}{x - d} \right) \in \mathbb{F}[x]$$

无重根.

14. 设 $f \in \mathbb{R}[x]$, $g = \gcd(f, f', \dots, f^{(k)})$, $k \in \mathbb{N}^*$. 证明: $g^{k+1} \mid f$.

15. 分别求 f 在 $\mathbb{C}[x]$, $\mathbb{R}[x]$, $\mathbb{Q}[x]$ 中的不可约因式分解, 其中

$$(1) f = x^8 + 1. \quad (2) f = x^8 + x^4 + 1. \quad (3) f = (x + 1)^8 + (x - 1)^8.$$

16. 设 $f \in \mathbb{Q}[x]$, $a, b \in \mathbb{Q}$. 证明: 若 $f(a + b\sqrt{5}) = 0$, 则 $f(a - b\sqrt{5}) = 0$.

17. 设 $f \in \mathbb{R}[x]$. 证明: 若对于任意 $a \in \mathbb{R}$ 都有 $f(a) \geq 0$, 则存在 $g \in \mathbb{C}[x]$ 使得 $f = g\bar{g}$, 其中

$$\bar{g} = \sum_{k=0}^n \bar{c}_k x^k, \quad \text{称为 } g = \sum_{k=0}^n c_k x^k \text{ 的共轭多项式.}$$

18. 设 $f \in \mathbb{Z}[x]$ 是首一多项式, $a \in \mathbb{Q}$. 证明: 若 $f(a) = 0$, 则 $a \in \mathbb{Z}$.

19. 设整数 a_1, \dots, a_n 两两不同. 证明: $f = 1 + \prod_{i=1}^n (x - a_i)$ 在 $\mathbb{Q}[x]$ 中是不可约的.

20. 设 $f = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$, $a_i \in \mathbb{Z}$. $\text{cont}(f) = \gcd(a_0, a_1, \dots, a_n)$ 称为 f 的容度. 若 $\text{cont}(f) = 1$, 则 f 称为本原^[2]的. 证明: $\text{cont}(f_1 f_2) = \text{cont}(f_1) \text{cont}(f_2)$, $\forall f_1, f_2 \in \mathbb{Z}[x]$.

21. 当正整数 n 是合数时, $f = 1 + x + \dots + x^{n-1}$ 在 $\mathbb{Q}[x]$ 中是否可约的? 证明你的结论.

22. 证明例 A.8 中的 3 个结论.

23. 设 R 是整环. 本节中关于 $\mathbb{F}[x]$ 的结论, 对于 $R[x]$ 是否成立? 证明你的结论或举出反例.

^[2]Gauss 提出了本原多项式的概念, 并证明了 Gauss 引理: 两个本原多项式的乘积仍是本原多项式.

§A.2 多元多项式

本节主要研究数域 \mathbb{F} 上的多元多项式.

定义 A.7. 设 x_1, \dots, x_n 是定义在 \mathbb{F} 上的两两不同的变元, $n \in \mathbb{N}^*$. 若干个系数与变元的乘积

$$p = cx_1^{k_1} \cdots x_n^{k_n}$$

称为 \mathbb{F} 上的**单项式**, 其中 $c \in \mathbb{F}$, $k_1, \dots, k_n \in \mathbb{N}$. 当 $c \neq 0$ 时, c 称为 p 的**系数**; 非负整数向量 (k_1, \dots, k_n) 称为 p 的**次数**. 次数相同的单项式称为**同类项**. 对于每个变元 x_i , 约定 $x_i^0 = 1$.

在所有非零单项式的集合上定义**字典序**

$$cx_1^{k_1} \cdots x_n^{k_n} < \tilde{c}x_1^{l_1} \cdots x_n^{l_n} \quad \text{或} \quad \tilde{c}x_1^{l_1} \cdots x_n^{l_n} > cx_1^{k_1} \cdots x_n^{k_n}$$

当且仅当

$$\text{存在 } 1 \leq r \leq n \text{ 使得 } k_r < l_r, k_{r+1} = l_{r+1}, \dots, k_n = l_n.$$

若干个单项式的和式称为**多项式**. 合并同类项之后, 把多项式 f 的所有单项式按照 $>$ 排序, 第一个单项式 p 称为 f 的**首项**, 记作 $\text{lt}(f)$; p 的系数和次数分别称为 f 的**首项系数**和**次数**, 记作 $\text{lc}(f)$ 和 $\text{deg}(f)$.

\mathbb{F} 上以 x_1, \dots, x_n 为变元的多项式的全体记作 $\mathbb{F}[x_1, \dots, x_n]$. 在 $\mathbb{F}[x_1, \dots, x_n]$ 上可以自然地定义多项式的加法、减法、乘法运算. 在多项式的加法和乘法运算下, $\mathbb{F}[x_1, \dots, x_n]$ 构成整环, 称为 \mathbb{F} 上以 x_1, \dots, x_n 为变元的**多元多项式环**.

\mathbb{F} 上形如 $\frac{f}{g}$ 的有理函数的全体记作 $\mathbb{F}(x_1, \dots, x_n)$, 其中 $f, g \in \mathbb{F}[x_1, \dots, x_n]$ 且 $g \neq 0$. 在有理函数的加法和乘法运算下, $\mathbb{F}(x_1, \dots, x_n)$ 构成数域, 称为 \mathbb{F} 上以 x_1, \dots, x_n 为变元的**有理函数域**. 注意到, $g \in \mathbb{F}[x_1, \dots, x_n]$ 有乘法逆元 $\frac{1}{g} \in \mathbb{F}(x_1, \dots, x_n)$ 当且仅当 $g \in \mathbb{F} \setminus \{0\}$.

$\mathbb{F}[x_1, \dots, x_n]$ 可以被看作是整环 $R = \mathbb{F}[x_1, \dots, x_{n-1}]$ 上的一元多项式环 $R[x_n]$, 也可以被看作是数域 $K = \mathbb{F}(x_1, \dots, x_{n-1})$ 上的一元多项式环 $K[x_n]$ 的子集. 从而, 上一节中关于一元多项式的许多概念和结论可以推广至多元多项式.

定理 A.16. 设 R 是整环, $f = \sum_{i=0}^s a_i x^i$, $g = \sum_{j=0}^t b_j x^j$, 其中 $a_i, b_j \in R$ 且 $a_s b_t \neq 0$, 则存在 $q, r \in R[x]$ 和非负整数 $k \leq \max(s-t+1, 0)$, 使得

$$b_t^k f = qg + r \quad \text{且} \quad \text{deg}(r) < t. \tag{A.6}$$

上式 (A.6) 称为 f/g 的**带余伪除**, r 称为 f/g 的**伪余式**.

证明. 对 s 归纳. 当 $s < t$ 时, $q = 0$, $r = f$, $k = 0$ 满足题设. 当 $s \geq t$ 时, $h = b_t f - (a_s x^{s-t})g \in R[x]$ 满足 $s' = \text{deg}(h) < s$. 根据归纳假设, 存在 $q', r \in R[x]$ 和非负整数 $k' \leq \max(s' - t + 1, 0)$ 使得 $b_t^{k'} h = q'g + r$ 且 $\text{deg}(r) < t$. 因此, $b_t^{k'+1} f = (q' + g_t^{k'} a_s x^{s-t})g + r$, 其中 $q = q' + b_t^{k'} a_s x^{s-t} \in R[x]$, $k = k' + 1 \leq \max(s-t+1, 1) = s-t+1$ 满足题设. \square

定义 A.8. 设 $f, f_1, \dots, f_k, g, h \in \mathbb{F}[x_1, \dots, x_n]$ (或者 $f, f_1, \dots, f_k, g, h \in R[x]$, 其中 R 是整环).

- 若 $g \neq 0$ 并且存在 h 使得 $f = gh$, 则称 f 是 g 的**倍式**, g 是 f 的**因式**, 或者称 f 被 g **整除**, g **整除** f , 记作 $g \mid f$.

- 若 f_1, \dots, f_k 不全为 0, $g \neq 0$ 并且 g 是每个 f_i 的因式, 则 g 称为 f_1, \dots, f_k 的公因式. 若 f_1, \dots, f_k 的所有公因式 $g \in \mathbb{F}$ (或者 g 有乘法逆元 $\frac{1}{g} \in R[x]$), 则称 f_1, \dots, f_k 互素.
- 若 f_1, \dots, f_k 全不为 0, 并且 g 是每个 f_i 的倍式, 则 g 称为 f_1, \dots, f_k 的公倍式.

定理 A.17. 设 R 是整环, $f_1, \dots, f_k, g, h \in R[x] \setminus \{0\}$.

1. 若 g, h 都是 f_1, \dots, f_k 的公倍式, 并且 $\deg(g)$ 最小, 则存在 $\lambda \in R \setminus \{0\}$ 使得 $g \mid \lambda h$.
2. 若 g, h 都是 f_1, \dots, f_k 的公因式, 并且 $\deg(g)$ 最大, 则存在 $\lambda \in R \setminus \{0\}$ 使得 $h \mid \lambda g$.
3. 若 g 是 f_1, \dots, f_k 的次数最大的公因式, 则存在 $u_1, \dots, u_k \in R[x]$ 和 $\lambda \in R \setminus \{0\}$ 使得

$$u_1 f_1 + \dots + u_k f_k = \lambda g.$$

证明.

1. 设 r 是 h/g 的伪余式, 则 r 也是 f_1, \dots, f_k 的公倍式. 由 $\deg(g)$ 的最小性, 得 $r = 0$. 故存在 $\lambda \in R \setminus \{0\}$ 使得 $g \mid \lambda h$.
2. 设 $p = \prod_{i=1}^k f_i$, 则 p/g 和 p/h 都是 $p/f_1, \dots, p/f_k$ 的公倍式. 由结论 1, 存在 $\lambda \in R \setminus \{0\}$ 使得 p/g 整除 $\lambda p/h$. 故 $h \mid \lambda g$. (注: 也可以由结论 3 直接得到结论 2.)
3. 设 h 是集合 $S = \{u_1 f_1 + \dots + u_k f_k \mid u_1, \dots, u_k \in R[x]\}$ 中次数最小的非零多项式. 易知 $g \mid h$. 设 r_i 是 f_i/h 的伪余式, 则 $r_i \in S$. 由 $\deg(h)$ 的最小性, 得 $r_i = 0$. 故存在 $\lambda_i \in R \setminus \{0\}$ 使得 $h \mid \lambda_i f_i$. 设 $\mu = \lambda_1 \cdots \lambda_k$, 则 μg 和 h 都是 $\mu f_1, \dots, \mu f_k$ 的公因式. 由 $\deg(\mu g) \geq \deg(h)$ 和 $g \mid h$, 得 $\lambda = h/g \in R \setminus \{0\}$. 因此, 存在 $u_1, \dots, u_k \in R[x]$ 使得 $u_1 f_1 + \dots + u_k f_k = h = \lambda g$. \square

定义 A.9. 设 $f \in \mathbb{F}[x_1, \dots, x_n]$. 若存在 $g, h \in \mathbb{F}[x_1, \dots, x_n] \setminus \mathbb{F}$ 使得 $f = gh$, 则称 f 在 $\mathbb{F}[x_1, \dots, x_n]$ 中是可约的; 否则称 f 在 $\mathbb{F}[x_1, \dots, x_n]$ 中是不可约的. 在不引起歧义的情况下, 可简称“ f 是可约的”或“ f 是不可约的”. 约定 0 是可约的.

定理 A.18. 设 $f \in \mathbb{F}[x_1, \dots, x_n]$ 是不可约的.

1. 对于任意 $g \in \mathbb{F}[x_1, \dots, x_n]$, f, g 互素或者 $f \mid g$.
2. 对于任意 $g_1, \dots, g_k \in \mathbb{F}[x_1, \dots, x_n]$, 若 $f \mid g_1 \cdots g_k$, 则 f 整除某个 g_i .

证明.

1. 若 f, g 不互素, 则存在公因式 $d \in \mathbb{F}[x_1, \dots, x_n] \setminus \mathbb{F}$. 由 f 是不可约的, 得 $\lambda = f/d \in \mathbb{F}$. 故 $f = \lambda d$ 整除 g .
2. 假设 f 不整除任意 g_i . 不妨设 f 关于 x_n 的次数 ≥ 1 . 记 $R = \mathbb{F}[x_1, \dots, x_{n-1}]$. 由结论 1 和定理 A.17, 存在 $u_i, v_i \in R[x_n]$ 使得 $u_i f + v_i g_i \in R \setminus \{0\}$. 从而 $\prod_{i=1}^k (u_i f + v_i g_i) \in R \setminus \{0\}$ 可被 f 整除. 矛盾. \square

定理 A.19 (唯一因式分解定理). 对于任意 $f \in \mathbb{F}[x_1, \dots, x_n] \setminus \mathbb{F}$, 存在不可约的 $f_1, \dots, f_k \in \mathbb{F}[x_1, \dots, x_n] \setminus \mathbb{F}$ 使得 $f = \prod_{i=1}^k f_i$. 不考虑因式的顺序以及非零常数因子, 这种分解方式是唯一的.

解答. 把 $f = x^2 + y^2 - 4$ 和 $g = 3x^2 + 7xy + 2y^2 - 8$ 看作是 关于 y 的一元多项式, 构造 Sylvester

方阵 $S = \begin{pmatrix} x^2 - 4 & 0 & 1 & 0 \\ 0 & x^2 - 4 & 0 & 1 \\ 3x^2 - 8 & 7x & 2 & 0 \\ 0 & 3x^2 - 8 & 7x & 2 \end{pmatrix}$, 算得 $\det(S) = 50x^4 - 196x^2$. 设 $f(x, y) = g(x, y) = 0$, 则

$\det(S) = uf + vg = 0$, 解得 $x \in \left\{ 0, \frac{7\sqrt{2}}{5}, -\frac{7\sqrt{2}}{5} \right\}$. 当 $x = 0$ 时, 解得 $y = \pm 2$. 当 $x = \frac{7\sqrt{2}}{5}$ 时, 解得 $y = -\frac{\sqrt{2}}{5}$. 当 $x = -\frac{7\sqrt{2}}{5}$ 时, 解得 $y = \frac{\sqrt{2}}{5}$. 综上, 所求交点共有 4 个. \square

例 A.12 (Vieta^[3]定理). 设 x_1, \dots, x_n, y 是两两不同的变元, 则有

$$\prod_{i=1}^n (y - x_i) = \sum_{k=0}^n (-1)^k \sigma_k y^{n-k} \quad (\text{A.7})$$

其中 $\sigma_0 = 1$, $\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} x_{i_1} \cdots x_{i_k}$ 称为 k 次基本对称多项式.

例 A.13. 考虑 Jacobi 矩阵

$$A = \begin{pmatrix} \frac{\partial \sigma_1}{\partial x_1} & \cdots & \frac{\partial \sigma_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial \sigma_n}{\partial x_1} & \cdots & \frac{\partial \sigma_n}{\partial x_n} \end{pmatrix}.$$

取定 j , (A.7) 式两端对 x_j 求偏导数, 得

$$-\prod_{i \neq j} (y - x_i) = \sum_{k=1}^n (-1)^k \frac{\partial \sigma_k}{\partial x_j} y^{n-k}.$$

再在上式中令 $y = x_k$, $1 \leq k \leq n$, 得

$$\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{pmatrix} = (-1)^{n-1} \begin{pmatrix} (-x_1)^{n-1} & \cdots & -x_1 & 1 \\ \vdots & \cdots & \vdots & \vdots \\ (-x_n)^{n-1} & \cdots & -x_n & 1 \end{pmatrix} A,$$

其中 $w_k = \prod_{i \neq k} (x_k - x_i)$. 由此可得, A 与 Vandermonde 矩阵之间具有如下联系.

$$A = \begin{pmatrix} & & (-1)^0 & & \\ & \ddots & & & \\ (-1)^{n-1} & & & & \end{pmatrix} \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}^{-1} \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{pmatrix},$$

$$A^{-1} = \begin{pmatrix} \frac{1}{w_1} & & \\ & \ddots & \\ & & \frac{1}{w_n} \end{pmatrix} \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} & & & (-1)^{n-1} \\ & \ddots & & \\ & & & (-1)^0 \end{pmatrix}.$$

^[3]François Viète, 1540–1603, 法国数学家.

例 A.14. 设 x_1, \dots, x_n 是两两不同的变元, 则 $s_k = \sum_{i=1}^n x_i^k$ 称为 k 次等幂和. 下面考虑等幂和与基本对称多项式之间的联系. 构造生成函数 $f(y) = \sum_{k=1}^{\infty} s_k y^k$, 则

$$f(y) = \sum_{i=1}^n \frac{x_i y}{1 - x_i y} = -y \frac{g'(y)}{g(y)} \quad \text{其中} \quad g(y) = \prod_{i=1}^n (1 - x_i y) = \sum_{k=0}^n (-1)^k \sigma_k y^k.$$

比较等式 $f(y)g(y) = -yg'(y)$ 两端的 y^k 系数, 得 **Newton 恒等式**

$$\sum_{i=0}^{k-1} (-1)^i \sigma_i s_{k-i} = (-1)^{k-1} k \sigma_k$$

其中 $\sigma_k = 0, \forall k > n$.

根据 Newton 恒等式, 我们可以把 σ_k 递推地表示为关于 s_1, \dots, s_k 的有理数系数多项式

$$\sigma_k = \frac{1}{k} \sum_{i=1}^k (-1)^{i-1} s_i \sigma_{k-i}, \quad k = 1, \dots, n.$$

或者, 把 s_k 递推地表示为关于 $\sigma_1, \dots, \sigma_k$ 的整数系数多项式

$$s_k = (-1)^{k-1} \sigma_k k + \sum_{i=1}^{k-1} (-1)^{i-1} \sigma_i s_{k-i}, \quad k \geq 1.$$

或者, 把 Newton 恒等式表示为矩阵乘积的形式

$$\begin{pmatrix} \sigma_0 & & & \\ \sigma_1 & \sigma_0 & & \\ \vdots & \ddots & \ddots & \\ \sigma_{k-1} & \cdots & \sigma_1 & \sigma_0 \end{pmatrix} \begin{pmatrix} s_1 \\ -s_2 \\ \vdots \\ (-1)^{k-1} s_k \end{pmatrix} = \begin{pmatrix} \sigma_1 \\ 2\sigma_2 \\ \vdots \\ k\sigma_k \end{pmatrix}.$$

又或者, 利用 Cramer 法则, 把 s_k 表示为如下 k 阶行列式的形式

$$s_k = \begin{vmatrix} \sigma_1 & \sigma_0 & & \\ 2\sigma_2 & \sigma_1 & \ddots & \\ \vdots & \vdots & \ddots & \sigma_0 \\ k\sigma_k & \sigma_{n-1} & \cdots & \sigma_1 \end{vmatrix}, \quad k \geq 1.$$

前例中的等幂和 s_k 与基本对称多项式 σ_k 都属于一类特殊的多元多项式——对称多项式.

定义 A.11. 设 $f \in \mathbb{F}[x_1, \dots, x_n]$. 若对于 $1, \dots, n$ 的任意排列 i_1, \dots, i_n , 都有

$$f(x_{i_1}, \dots, x_{i_n}) = f(x_1, \dots, x_n),$$

则 f 称为关于变元 x_1, \dots, x_n 的**对称多项式**.

例如, 对于任意 $f \in \mathbb{F}[x_1, \dots, x_n]$, $g = \sum_{(i_1, \dots, i_n) \in S_n} f(x_{i_1}, \dots, x_{i_n})$ 是对称多项式.

定理 A.20. 对于任意对称多项式 $f \in \mathbb{F}[x_1, \dots, x_n]$, 存在唯一的 $\varphi \in \mathbb{F}[\sigma_1, \dots, \sigma_n]$, 使得

$$f = \varphi(\sigma_1, \dots, \sigma_n),$$

其中 σ_k 是关于 x_1, \dots, x_n 的 k 次基本对称多项式.

证明. 对 $\text{lt}(f)$ 归纳. 当 $\text{lt}(f) \in \mathbb{F}$ 时, $f \in \mathbb{F}$, 结论显然成立. 设 $\text{lt}(f) = cx_1^{k_1} \cdots x_n^{k_n}$, $c \in \mathbb{F} \setminus \{0\}$. 由 f 是对称多项式, 得 $k_1 \leq \cdots \leq k_n$. 设 $\varphi_1 = cx_n^{k_1} x_{n-1}^{k_2 - k_1} \cdots x_1^{k_n - k_{n-1}}$, $g = \varphi_1(\sigma_1, \cdots, \sigma_n)$. 由

$$\text{lt}(g) = c(x_1 \cdots x_n)^{k_1} (x_2 \cdots x_n)^{k_2 - k_1} \cdots x_n^{k_n - k_{n-1}} = cx_1^{k_1} x_2^{k_2} \cdots x_{n-1}^{k_{n-1}} x_n^{k_n} = \text{lt}(f)$$

得 $\text{lt}(f - g) \prec \text{lt}(f)$. 根据归纳假设, 存在唯一的 $\varphi_2 \in \mathbb{F}[x_1, \cdots, x_n]$ 使得 $f - g = \varphi_2(\sigma_1, \cdots, \sigma_n)$. $\varphi = \varphi_1 + \varphi_2$ 满足题设. \square



把定理 A.20 中的数域 \mathbb{F} 换成整环 R , 结论仍然成立, 证明同上.

例 A.15. 设 $n \geq 2$. 把 $f = \sum_{1 \leq i < j \leq n} (x_i^2 x_j^3 + x_i^3 x_j^2)$ 表示成关于 $\sigma_1, \cdots, \sigma_n$ 的多项式.

解答. $\text{lt}(f) = x_{n-1}^2 x_n^3 = \text{lt}(\sigma_1 \sigma_2^2)$. 由 $g = f - \sigma_1 \sigma_2^2$ 是 5 次齐次^[4]对称多项式, 并且 $\text{lt}(g) \prec \text{lt}(f)$, 得 g 的每个单项式为下列 4 种形式之一

$$cx_{i_1}^3 x_{i_2} x_{i_3}, \quad cx_{i_1}^2 x_{i_2}^2 x_{i_3}, \quad cx_{i_1}^2 x_{i_2} x_{i_3} x_{i_4}, \quad cx_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5}$$

其中 c 是常数, i_1, i_2, i_3, i_4, i_5 两两不同. 故可设

$$f = \sigma_1 \sigma_2^2 + c_1 \sigma_1^2 \sigma_3 + c_2 \sigma_2 \sigma_3 + c_3 \sigma_1 \sigma_4 + c_4 \sigma_5$$

其中 c_1, c_2, c_3, c_4 是常数. 在上式中分别令 $(x_1, \cdots, x_n) = (1, 1, -1, 0, \cdots, 0)$ 或 $(1, 1, 1, 0, \cdots, 0)$ 或 $(1, 1, 1, 1, 0, \cdots, 0)$ 或 $(1, 1, 1, 1, 1, 0, \cdots, 0)$, 得

$$\begin{cases} 2 = 1 - c_1 + c_2 \\ 6 = 27 + 9c_1 + 3c_2 \\ 12 = 144 + 64c_1 + 24c_2 + 4c_3 \\ 20 = 500 + 250c_1 + 100c_2 + 25c_3 + c_4 \end{cases}$$

解得 $c_1 = -2$, $c_2 = -1$, $c_3 = 5$, $c_4 = -5$. 即 $f = \sigma_1 \sigma_2^2 - 2\sigma_1^2 \sigma_3 - \sigma_2 \sigma_3 + 5\sigma_1 \sigma_4 - 5\sigma_5$. \square

习题

1. 设 R 是整环, $f, g \in R[x]$. 证明: 定理 A.1 的结论仍然成立.
2. 设 $f, g \in \mathbb{F}[x_1, \cdots, x_n]$, $n \geq 2$. 证明:
 - (1) $\text{lt}(f + g) \prec \text{lt}(f)$ 或 $\text{lt}(f + g) \prec \text{lt}(g)$.
 - (2) 当 $\text{lt}(f) \prec \text{lt}(g)$ 时, $\text{lt}(f + g) = \text{lt}(g)$.
 - (3) $\text{lt}(fg) = \text{lt}(f) \text{lt}(g)$, $\deg(fg) = \deg(f) + \deg(g)$.
 - (4) 若 $fg = 0$, 则 $f = 0$ 或 $g = 0$.
3. 设 R 是整环, $r_0, r_1, \cdots, r_k \in R[x] \setminus \{0\}$ 满足 r_{i+1} 是 r_i/r_{i-1} 的伪余式, $i = 1, 2, \cdots, k$, 其中 $r_{k+1} = 0$. 证明: 存在 $\lambda \in R \setminus \{0\}$ 使得 $r_k = \lambda \gcd(r_0, r_1)$.

^[4]若 $g \in \mathbb{F}[x_1, \cdots, x_n]$ 满足 $g(\lambda x_1, \cdots, \lambda x_n) = \lambda^k g(x_1, \cdots, x_n)$, 则 g 称为 k 次齐次多项式.

4. 求复数 c_1, c_2, c_3, c_4 应满足的充分必要条件, 使得 $f = c_1x^2 + c_2xy + c_3y^2 + c_4$ 在 $\mathbb{C}[x, y]$ 中不可约.
5. 已知平面曲线的参数方程为 $x = t^3 + t^2, y = t^3 + t$. 用结式消元法求 $f \in \mathbb{R}[x, y]$ 使得 $f(x(t), y(t)) = 0, \forall t$.
6. 已知空间曲面的参数方程为 $x = s + t, y = s^2 - t^2, z = s^2 + t^2$. 用结式消元法求 $f \in \mathbb{R}[x, y, z]$ 使得 $f(x(s, t), y(s, t), z(s, t)) = 0, \forall s, t$.

7. 设 A 和 w_k 如例 A.13 所述. 证明:

$$(1) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} (-1)^{n-1} \sigma_{n-1} \\ \vdots \\ (-1)^0 \sigma_0 \end{pmatrix}.$$

$$(2) \det(A) = \prod_{1 \leq i < j \leq n} (x_i - x_j).$$

8. 设 $n \geq 3$, 把下列对称多项式 $f \in \mathbb{F}[x_1, \dots, x_n]$ 表示成关于 $\sigma_1, \dots, \sigma_n$ 的多项式.

$$(1) f = \sum_{1 \leq i < j < k \leq n} (x_i^2 x_j^2 x_k + x_i^2 x_j x_k^2 + x_i x_j^2 x_k^2).$$

$$(2) f = \prod_{1 \leq i < j \leq n} (x_j - x_i)^2.$$

9. (1) 设 $f \in \mathbb{F}[x_1, \dots, x_n]$ 满足交换任意两个变元后表达式不变, 即

$$f(\cdots, x_j, \cdots, x_i, \cdots) = f(\cdots, x_i, \cdots, x_j, \cdots), \quad \forall i < j.$$

证明: f 是对称多项式.

(2) 设 $\text{Char } \mathbb{F} \neq 2$, $f \in \mathbb{F}[x_1, \dots, x_n]$ 满足交换任意两个变元后表达式反号^[5], 即

$$f(\cdots, x_j, \cdots, x_i, \cdots) = -f(\cdots, x_i, \cdots, x_j, \cdots), \quad \forall i < j.$$

证明: 存在对称多项式 $g \in \mathbb{F}[x_1, \dots, x_n]$ 使得 $f = g \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

(3) 设 $\text{Char } \mathbb{F} = 0$, $f \in \mathbb{F}[x_1, \dots, x_n]$ 满足轮换任意三个不同变元后表达式不变, 即

$$f(\cdots, x_j, \cdots, x_k, \cdots, x_i, \cdots) = f(\cdots, x_i, \cdots, x_j, \cdots, x_k, \cdots), \quad \forall i < j < k.$$

证明: 存在 $\mathbb{F}[x_1, \dots, x_n]$ 中的对称多项式 g 和反对称多项式 h 使得 $f = g + h$.

10. 设 $\text{Char } \mathbb{F} = 0$. 求所有 $f \in \mathbb{F}[x_1, \dots, x_n]$ 满足 $f(x_2, \dots, x_n, x_1) = f(x_1, x_2, \dots, x_n)$.

^[5] f 称为关于变元 x_1, \dots, x_n 的反对称多项式.

参考文献

1. 数论导引, 华罗庚著, 科学出版社, 1957 年 7 月.
华罗庚文集: 数论卷 II, 华罗庚著, 贾朝华审校, 科学出版社, 2010 年 5 月.
2. 典型群, 华罗庚、万哲先著, 上海科学技术出版社, 1963 年 7 月.
华罗庚文集: 代数卷 I, 华罗庚、万哲先著, 万哲先审校, 科学出版社, 2010 年 5 月.
3. 中国数学史, 钱宝琮主编, 科学出版社, 1964 年 11 月.
4. 代数学引论, 许以超编著, 上海科学技术出版社, 1966 年 5 月.
5. Linear Algebra (Second Edition), Kenneth Hoffman and Ray Kunze, Prentice-Hall, 1971.
6. Linear Algebra (Fourth Edition), Werner Hildbert Greub, GTM 23, Springer-Verlag, 1974.
7. 高等数学引论 (余篇), 华罗庚著, 科学出版社, 1984 年 7 月.
高等数学引论 (第四册), 华罗庚著、王元校, 高等教育出版社, 2009 年 4 月.
8. 数书九章新释, (宋) 秦九韶原著, 王守义遗著, 李俨审校, 安徽科学技术出版社, 1992 年 10 月.
9. Set Theory (The Third Millennium Edition, revised and expanded), Thomas Jech, Springer, 2003.
10. 组合矩阵论 (第 2 版), 柳柏濂著, 科学出版社, 2005 年 1 月.
11. 线性代数, 李尚志编著, 高等教育出版社, 2006 年 5 月.
12. Handbook of Linear Algebra, edited by Leslie Hogben, Chapman & Hall / CRC, 2006.
13. 代数学引论第一卷: 基础代数 (第 2 版), А. И. Кострикин 著, 张英伯译, 高等教育出版社, 2006 年 12 月.
14. 代数学引论第二卷: 线性代数 (第 3 版), А. И. Кострикин 著, 牛凤文译, 高等教育出版社, 2008 年 1 月.
15. 线性代数与矩阵论 (第二版), 许以超编著, 高等教育出版社, 2008 年 6 月.
16. Advanced Linear Algebra (Third Edition), Steven Roman, GTM 135, Springer, 2008.
17. 线性代数 (第 2 版), 李炯生、查建国、王新茂编著, 中国科学技术大学出版社, 2010 年 1 月.

18. 多项式代数, 王东明、牟晨琪、李晓亮、杨静、金萌、黄艳丽编著, 高等教育出版社, 2011 年 5 月.
19. 典型流形与典型域, 陆启铿著, 科学出版社, 2011 年 8 月.
20. Matrix Computations (Fourth Edition), Gene H. Golub and Charles F. Van Loan, Johns Hopkins University Press, 2012.
21. Spectra of Graphs, Andries E. Brouwer and Willem H. Haemers, Springer, 2012.
22. 高等代数, 丘维声著, 科学出版社, 2013 年 3 月.
23. 古今数学思想, 第 1 至 3 册, Morris Kline 著, 张理京、石生明、邓东皋等译, 上海科学技术出版社, 2014 年 1 月.
24. 九章算术新校, 上、下册, 郭书春汇校, 中国科学技术大学出版社, 2014 年 1 月.
25. 高等代数 (第三版), 上、下册, 丘维声著, 高等教育出版社, 2015 年 3 月.
26. 线性代数与解析几何 (第 2 版), 陈发来、陈效群、李思敏、王新茂编著, 高等教育出版社, 2015 年 8 月.
27. 线性代数及其应用 (原书第 4 版), David C. Lay 著, 刘深泉、张万芹、陈玉珍、包东娥、陆博译, 机械工业出版社, 2017 年 3 月.
28. 图论及其应用 (第 4 版), 徐俊明编著, 中国科学技术大学出版社, 2019 年 3 月.
29. 线性代数高级教程 — 矩阵理论及应用, Stephan Ramon Garcia and Roger A. Horn 著, 张明尧译, 机械工业出版社, 2020 年 1 月.