



线性代数与解析几何

速通版复习资料

作者：2025 级 AI 学组

1. 主要编写：AI2501 范晨昊

2. 协助编写：AI2502 薛华林

3. 监管与审核：AI2502 王远棹、AI2502 孟祥杭、AI2501 张家畅

时间：2025/09-2026/01

版本：Final ver.



如有错误欢迎指正 QQ.3577646143 谢谢喵
本资料中大部分内容来源于授课幻灯片和课本

目录

第一章 行列式	2	3.1.4 空间直角坐标系	36
1.1 行列式的定义与性质	2	3.1.5 空间向量的坐标表示	37
1.1.1 2阶行列式	2	3.1.6 空间向量坐标的性质	38
1.1.2 3阶行列式和 n 阶行列式	3	3.1.7 正交射影	38
1.1.3 余子式和代数余子式	4	3.2 数量积、向量积和混合积	39
1.1.4 行列式的性质	5	3.2.1 数量积	39
1.2 行列式的计算	7	3.2.2 向量积	39
1.2.1 利用行列式的性质计算	7	3.2.3 混合积	40
1.2.2 Vandermonde 行列式	8	3.3 平面和空间直线	42
1.2.3 块行列式	9	3.3.1 平面的方程	42
1.3 Cramer 法则 (1)	11	3.3.2 平面与平面的位置关系	42
第一章 练习	12	3.3.3 直线的方程	43
第二章 矩阵	14	3.3.4 直线和直线的位置关系	44
2.1 矩阵及其运算	14	3.3.5 直线与平面的位置关系	45
2.1.1 矩阵的概念	14	3.3.6 点到平面的距离	45
2.1.2 矩阵的代数运算	15	3.3.7 点到直线的距离	46
2.1.3 矩阵的转置	17	3.3.8 异面直线的距离	46
2.1.4 方阵的行列式	18	第三章 练习	48
2.2 逆矩阵	19	第四章 n 维向量与线性方程组	49
2.2.1 逆矩阵的定义	19	4.1 消元法	49
2.2.2 逆矩阵的性质	21	4.1.1 线性方程组的基本概念	49
2.2.3 Cramer 法则 (2) 和矩阵方程	21	4.1.2 消元法	49
2.3 分块矩阵	23	4.1.3 线性方程组的解的判定	51
2.3.1 分块矩阵的定义	23	4.1.4 数域	53
2.3.2 分块矩阵的运算	24	4.2 向量组的线性相关性	54
2.4 初等变换与初等矩阵	27	4.2.1 向量组及其运算	54
2.4.1 矩阵初等变换的概念	27	4.2.2 线性表示和等价向量组	54
2.4.2 初等矩阵和初等变换的关系	28	4.2.3 线性相关和线性无关	56
2.4.3 阶梯型矩阵	29	4.2.4 线性相关性的基本性质	58
2.4.4 利用初等变换求逆矩阵	30	4.3 向量组的秩	59
2.5 矩阵的秩	31	4.3.1 向量组的极大无关组	59
2.5.1 秩的概念	31	4.3.2 向量组的秩	60
2.5.2 计算矩阵秩的一般方法	31	4.4 线性方程组解的结构	62
2.5.3 矩阵的标准型和满秩分解	32	4.4.1 齐次线性方程组的解	62
第二章 练习	34	4.4.2 非齐次线性方程组	64
第三章 几何向量及其应用	35	第四章 练习	65
3.1 向量及其线性运算	35	第五章 线性空间与欧式空间	66
3.1.1 向量的概念	35	5.1 线性空间	66
3.1.2 向量的线性运算	35	5.1.1 线性空间的定义	66
3.1.3 向量共线和共面	36	5.1.2 线性子空间	67

5.1.3 线性空间的基、维数和坐标表示	68	7.2.3 实对称矩阵的对角化讨论	94
5.1.4 基变换与坐标变换	69	第七章 练习	99
5.1.5 * 线性空间的同构	70	第八章 二次曲面和二次型	100
5.1.6 * 子空间的交与和	71	8.1 常见的二次曲面	100
5.2 欧式空间	74	8.1.1 曲面	100
5.2.1 内积	74	8.1.2 曲线	101
5.2.2 范数、夹角和距离	75	8.1.3 柱面、锥面和旋转面	102
5.2.3 标准正交基	76	8.1.4 五种常见的二次曲面	104
5.2.4 Gram-Schmidt 正交化	77	8.1.5 曲面的投影	108
5.2.5 正交矩阵	78	8.2 二次型	110
第五章 练习	80	8.2.1 二次型的表示	110
第六章 * 线性变换	81	8.2.2 二次型的标准型	111
6.1 线性变换的定义和性质	81	8.2.3 合同变换、规范型与 Sylvester 惯性定理	113
6.2 线性变换的核与值域	82	8.2.4 正定二次型和正定矩阵	114
6.3 线性变换的运算	84	8.2.5 二次曲面的标准方程	117
6.4 线性变换的矩阵表示	85	第八章 练习	119
第六章 练习	87	附录 A 章节练习参考答案	120
第七章 特征值和相似矩阵	88	A.1 第一章答案	120
7.1 特征值和特征向量	88	A.2 第二章答案	124
7.1.1 特征值和特征向量的定义	88	A.3 第三章答案	126
7.1.2 特征值的代数重数和几何重数	88	A.4 第四章答案	128
7.1.3 矩阵特征值的求法	89	A.5 第五章答案	129
7.1.4 特征值的性质	90	A.6 第六章答案	131
7.2 相似矩阵	92	A.7 第七章答案	133
7.2.1 相似矩阵的概念	92	A.8 第八章答案	135
7.2.2 相似对角化	92		

前言

本资料由 XJTU-2025AI 学组编写，自从 2025 年 9 月起启动项目并于 2025 年底完成编写、审核、校对等工作。《线性代数与解析几何》是大一上学期和高数紧密相连的数学课程，其中的部分概念主打一个“不直观”。本项目聚焦于魏志强老师的授课幻灯片和《线性代数与解析几何》的授课点整理而成，具体而详细的总结了其中的定义、定理、命题并附赠了知识点的相关练习题和章节末尾的练习题。

本资料按照《线性代数与解析几何》的结构整理了**线性代数基础知识和空间解析几何**的相关内容，包括行列式、矩阵、线性方程组、线性空间、二次型等重要内容。对于首次接触的新概念，资料整理了相关定义、定理的补充说明和证明过程。对于中学已经接触的解析几何部分，资料以练习题和图片等形式展现相关内容，并展现线性代数和解析几何之间的联系。

最后，感谢 Elegant-Book 项目对本资料编写工作的支持；感谢范晨昊同学的资料编写；感谢薛华林同学对章节习题的补充；感谢王远棹、孟祥杭同学的认真审核、校对；感谢张家畅同学的项目监管。在期中版本，我们的资料收到了许多同学的反馈建议。我们认真听取，并代表整个项目组忠心感谢每一位做出贡献的同学。鉴于水平有限，仍可能出现若干不足和错误之处，欢迎指出并提出改进建议！

第一章 行列式

1.1 行列式的定义与性质

一个含有 m 个方程, n 个未知数的线性方程组可以表示为1.1:

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

我们通常采用消元法去解线性方程组, 但当方程数量和未知数数量足够大时, 消元法带来的计算量也随之增加, 为此我们引入“行列式与矩阵”的概念(详见 1-4 章节内容), 这些概念可以帮助我们快速计算和解决此类问题。这一章节里, 我们主要考虑 $m = n$ 时的情况。

1.1.1 2 阶行列式

当 $m = n = 2$ 时, 我们有方程组1.2:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad (1.2)$$

小型线性方程组可以快速通过消元得到方程的解:

$$\begin{cases} x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \\ x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}} \end{cases}$$

观察到两解的分母相同, 分母和分子的形式也相似, 由此给出 2 阶行列式的定义:

定义 1.1 (2 阶行列式)

2 阶行列式是由 4 个数字 $a_{ij}(i, j = 1, 2)$ 排成 2(横) 行、2(竖) 列的算式, 表示为:

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (1.3)$$

其中, i 和 j 分别表示第 i 行和第 j 列。2 阶行列式满足对角线法则的计算, 即用主对角线之乘积减去副对角线之乘积。

根据该定义, 可以改写方程1.2的解为:

$$\begin{cases} x_1 = \frac{D_1}{D} \\ x_2 = \frac{D_2}{D} \end{cases}$$

其中, 我们规定 $D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $D_j(j = 1, 2)$ 表示分别把 b_1, b_2 替换 D 中的 a_{1j}, a_{2j} 得到的新行列式。这一讨论结果将在后续 Cramer 法则小节详细介绍。

 **练习 1.1** 利用上述结论求解线性方程组 1.4:

$$\begin{cases} 2x_1 - x_2 = 1 \\ 3x_1 + 2x_2 = 0 \end{cases} \quad (1.4)$$

解 只需代入上述式子有 $D = 7, D_1 = 2, D_2 = -3$, 故有
$$\begin{cases} x_1 = \frac{D_1}{D} = \frac{2}{7} \\ x_2 = \frac{D_2}{D} = -\frac{3}{7} \end{cases}$$

1.1.2 3 阶行列式和 n 阶行列式

当 $m = n = 3$ 时, 我们有方程组 1.5:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad (1.5)$$

通过消元, 我们得到较为复杂表示的解。为此引入 3 阶行列式的定义:

定义 1.2 (3 阶行列式)

3 阶行列式是由 9 个数字 $a_{ij}(i, j = 1, 2, 3)$ 排成 3(横) 行、3(竖) 列的算式, 表示为:

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \quad (1.6)$$

3 阶行列式同样满足对角线法则的计算, 即用 3 条主对角线之乘积和减去 3 条副对角线之乘积和。

为了方便观看, 用图 1.1 表示对角线法则的计算。注意, 该法则仅对 2, 3 阶行列式生效。

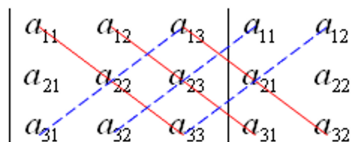


图 1.1: 3 阶行列式的对角线法则

借助定义 1.2, 我们可以把方程 1.5 的解以 $n = 2$ 的类似形式表示出, 即 $x_j = \frac{D_j}{D} (j = 1, 2, 3)$.

按照递归思路, 我们可以给出 n 阶行列式的定义:

定义 1.3 (n 阶行列式)

n 阶行列式是由 n^2 个数字 $a_{ij}(i, j = 1, 2, \dots, n)$ 排成 n (横) 行、 n (竖) 列的算式, 表示为:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (1.7)$$

可简记成 $\det(a_{ij})_{n \times n}$ 或者 $\det(a_{ij})$ 。特别地, 当 $n = 1$ 时, 规定 1 阶行列式 $|a_{11}| = a_{11}$

现在我们关心如何计算 n 阶行列式: 由于对角线法则对 $n > 3$ 的情况不再适用, 需要寻求新的方法总结出行列式计算的一般性规律, 我们在下一节给出讨论结果。

1.1.3 余子式和代数余子式

以 3 阶行列式为例，针对式 1.6 的右侧展开式，我们把第一行的元素 a_{1j} 提取为公因数，有：

$$\begin{aligned} D &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}(-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

注意到，每一项 a_{1j} 后面分别有着 $(-1)^{1+j}$ 和由 D 的部分元素按照原先相对顺序排成的“较小”的行列式。进一步把这种形式的式子归纳为定义 1.4：

定义 1.4 (余子式和代数余子式)

在 n 阶行列式 $D = \det(a_{ij})$ 中，称删去 a_{ij} 所在的第 i 行元素和第 j 列元素后由剩余元素按照它们原来的相对次序所形成的 $n-1$ 阶行列式为 a_{ij} 的余子式，记为 M_{ij} 。称：

$$A_{ij} = (-1)^{i+j} M_{ij} \quad (1.8)$$

为 a_{ij} 的代数余子式。

由代数余子式的定义可以直接把 3 阶行列式的计算改写为 1.9：

$$D = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = \sum_{j=1}^3 a_{1j}A_{1j} = \sum_{j=1}^3 a_{1j}(-1)^{1+j}M_{1j} \quad (1.9)$$

上述式子称为 3 阶行列式按照第 1 行展开的计算式，也有了下面的命题：

命题 1.1 (行列式按第 1 行展开)

n 阶行列式等于它的第 1 行各元素分别与其对应的代数余子式的乘积之和：

$$D = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} = \sum_{j=1}^n a_{1j}A_{1j} = \sum_{j=1}^n a_{1j}(-1)^{1+j}M_{1j} \quad (1.10)$$

 **练习 1.2** 分别用对角线法则和命题 1.1 的方法求行列式 $D_1 = \begin{vmatrix} 3 & -1 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & 3 \end{vmatrix}$ 和 $D_2 = \begin{vmatrix} 5 & -1 & 2 \\ 1 & 1 & 1 \\ 3 & 0 & 3 \end{vmatrix}$ 的值：

解 对角线法则：

$$D = 3 \times 1 \times 3 + (-1) \times 1 \times 2 + 0 \times 0 \times 2 - 2 \times 1 \times 2 - (-1) \times 0 \times 3 - 1 \times 0 \times 3 = 3$$

按第 1 行展开：

$$D_2 = 5 \times (-1)^{1+1} \times \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} + (-1) \times (-1)^{1+2} \times \begin{vmatrix} 1 & 1 \\ 3 & 3 \end{vmatrix} + 2 \times (-1)^{1+3} \times \begin{vmatrix} 1 & 1 \\ 3 & 0 \end{vmatrix} = 9$$

定理 1.1 (下三角行列式)

主对角线上方元素全为零的行列式，也即非零元素只出现在主对角线及下方的行列式，称为下三角形行列式(当 $i < j$ 时 $a_{ij} = 0$)。下三角行列式满足：

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdots a_{nn} = \prod_{i=1}^n a_{ii} \quad (1.11)$$

证明 采用数学归纳法, 当 $n=2$ 时 $D_2 = \begin{vmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}$ 成立;

假设 $n=k-1$ 时成立, 则 $n=k$ 时将 D_k 按照第 1 行展开有:

$$D_k = a_{11}(-1)^{1+1}M_{11} = a_{11} \begin{vmatrix} a_{22} & 0 & \cdots & 0 \\ a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{k2} & a_{k3} & \cdots & a_{kk} \end{vmatrix} = a_{11}a_{22} \cdots a_{kk} = \prod_{i=1}^k a_{ii}$$

由数学归纳法证毕。

同理, 上三角行列式 (当 $i > j$ 时 $a_{ij} = 0$) 的值也等于主对角元素之乘积。此类三角行列式是我们计算行列式的重要方法之一, 在后续小节中我们将学会如何利用行列式的性质对给定的行列式进行变换, 使之变成三角行列式的形式求值。

1.1.4 行列式的性质

行列式的行与列是否等价是一个重要的命题, 我们由定义 1.5 给出相关解释。

定义 1.5 (行列式的转置)

把行列式 $D = \det(a_{ij})$ 的行依次换成列 (或者说把列换成行) 所得到的行列式称为 D 的转置行列式, 记为 $D^T = \det(a_{ji})$:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad D^T = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}$$

并且有 $D = D^T$ 成立。

定义 1.4 表明行列式对行成立的性质, 对列也成立, 反之亦然。“行”与“列”的等价性使得我们后续的叙述更加具有一般性。在接下来的讨论中, 我们仅从“行”或“列”的角度去说明。

命题 1.2 (行列式交换行 (列))

互换行列式两行 (列) 的位置, 行列式的值反号。当行列式有两行 (列) 相同时, 行列式为 0。

 **练习 1.3** 证明副对角线行列式的计算公式:

$$\begin{vmatrix} 0 & \cdots & 0 & a_{1n} \\ 0 & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = (-1)^{n(n-1)/2} \prod_{i=1}^n a_{i,n-i+1} \quad (1.12)$$

证明 只需依次从第 n 列起向左交换次序, 一共交换了 $(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$ 次, 由命题 1.2 和定理 1.1 得证。

事实上, 由于行列式交换行列仅改变符号的性质, 可以进一步给出行列式按某一行展开的式子:

命题 1.3 (行列式按某行 (列) 展开)

n 阶行列式等于它的第 i_0 行 (第 j_0 列) 各元素分别与其对应的代数余子式的乘积之和:

$$D = \sum_{j=1}^n a_{i_0 j} (-1)^{i_0+j} M_{i_0 j} = \sum_{i=1}^n a_{i j_0} (-1)^{i+j_0} M_{i j_0} \quad (1.13)$$

命题1.3给了我们计算行列式的一种良好方法——如果行列式的某一行(列)中存在较多的0元素,那么按照此行(列)展开得到的计算式就会少很多项,称为“降阶法”。行列式按行展开时,如果用某一行的元素分别乘另一行元素对应的代数余子式,那么有:

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = \sum_{j=1}^n a_{ij}A_{kj} = \begin{cases} D, & \text{当 } k=i \text{ 时} \\ 0, & \text{当 } k \neq i \text{ 时} \end{cases}$$

我们给予 $k \neq i$ 情况的说明: $\sum_{j=1}^n a_{ij}A_{kj}$ 实际上等价于把行列式的 a_{kj} 元素变为 a_{ij} 元素,此时第 i 行和第 k 行相同,由命题1.2知行列式的值为0。

下面给出几个行列式的基本性质,不证。

命题 1.4 (行列式中提取公因数)

若行列式某行的各元素有公因子 k , 则可将 k 提到行列式符号外边来 (或者说, 用一个数 k 去乘行列式, 就等于用 k 去乘行列式某行的每个元素), 即:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (1.14)$$

命题 1.5 (行列式某一行的提取)

若行列式某行的每个元素都是两个数的和, 则可将此行列式写成如下两个行列式的和, 即:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \cdots & a_{in} + b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (1.15)$$

命题 1.6 (行列式某一行倍数加到另一行上)

行列式某行加上另一行的 k 倍, 行列式的值不变, 即:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1} + ka_{i1} & \cdots & a_{jn} + ka_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad (1.16)$$

 **练习 1.4** 证明: 若行列式 D 中有两行元素成比例, 则 $D = 0$ 。

证明 不妨假设 $a_{kj} = pa_{ij} (k \neq i)$, 由命题1.4提取出公因数 p , 此时两行相同, 由命题1.2得 $D = 0$ 。

1.2 行列式的计算

本小节将在行列式的基础上介绍几种常用的计算行列式的方法。

1.2.1 利用行列式的性质计算

 **练习 1.5** 求解方程 $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & x \\ 4 & 9 & x^2 \end{vmatrix} = 0$

解 按照对角线法则展开，有：

$$\begin{aligned} D &= 3x^2 + 4x + 18 - 9x - 2x^2 - 12 \\ &= x^2 - 5x + 6 \end{aligned}$$

解方程有 $x = 2$ 或 $x = 3$ 。

 **练习 1.6** 求行列式 $D = \begin{vmatrix} 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 2 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ n-1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & n \end{vmatrix}$

解 按照第 n 行展开，有：

$$D_n = n(-1)^{n+n} M_{nn} = n \begin{vmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 2 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ n-1 & \cdots & 0 & 0 \end{vmatrix} = (-1)^{(n-1)(n-2)/2} n!$$

通过一些变换把行列式化为三角形行列式是常见的计算方法，几种变换分别可以简记为表 1.1。这里以行 (row) 为例子，当进行列 (column) 变换操作时通常把 r 改写为 c 。利用这些记号，给出练习 1.7 和 1.8:

记号	含义
r_{ij}	交换行 i 和行 j
$r_i(k)$	用 k 数乘行 i
$r_i \div k$	从行 i 提取公因数 k
$r_{ij}(k)$	把第 i 行的 k 倍加到行 j 上

表 1.1: 行列式的变换

 **练习 1.7** 求行列式 $D = \begin{vmatrix} 1 & -1 & 1 & x-1 \\ 1 & -1 & x+1 & -1 \\ 1 & x-1 & 1 & -1 \\ x+1 & -1 & 1 & -1 \end{vmatrix}$

解 观察到每一列都有许多重复的元素，同时副对角线上存在未知数，因此考虑利用行列变换将行列式变为副对角线行列式：

$$D \xrightarrow[r=2,3,4]{r_{1j}(-1)} \begin{vmatrix} 1 & -1 & 1 & x-1 \\ 0 & 0 & x & -x \\ 0 & x & 0 & -x \\ x & 0 & 0 & -x \end{vmatrix} \xrightarrow[j=1,2,3]{c_{j4}(1)} \begin{vmatrix} 1 & -1 & 1 & x \\ 0 & 0 & x & 0 \\ 0 & x & 0 & 0 \\ x & 0 & 0 & 0 \end{vmatrix} = (-1)^{4(4-1)/2} x^4 = x^4$$

 **练习 1.8** 求行列式 $D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}$

解 该行列式的特点是每一行和每一列的元素和相同，故考虑行列变换使得某一行(列)可以提出公因数，再将其转化为三角行列式求解。从第 2 列起分别加到第 1 列上去，有：

$$D_n = \begin{vmatrix} x+(n-1)a & a & \cdots & a \\ x+(n-1)a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ x+(n-1)a & a & \cdots & x \end{vmatrix} = [x+(n-1)a] \begin{vmatrix} 1 & a & \cdots & a \\ 1 & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ 1 & a & \cdots & x \end{vmatrix} \xrightarrow[i=2, \dots, n]{r_{1i}(-1)} [x+(n-1)a] \begin{vmatrix} 1 & a & \cdots & a \\ 0 & x-a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x-a \end{vmatrix} = [x+(n-1)a](x-a)^{n-1}$$

1.2.2 Vandermonde 行列式

定理 1.2 (Vandermonde 行列式)

设有数 $a_1, a_2, \dots, a_n$ 组成的行列式 1.17 称为 n 阶 Vandermonde 行列式：

$$D_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq j < i \leq n} (x_i - x_j) \quad (1.17)$$

证明 采用数学归纳法：

i. 当 $n=2$ 时，有 $D_2 = \begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix} = x_2 - x_1$ 成立；

ii. 假设 $n=k$ 时成立，则当 $n=k+1$ 时，依次从第 $k+1$ 行起减去上一行的 x_1 倍，直到操作到第 2 行，那么有：

$$D_{k+1} = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_{k+1} \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_{k+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^k & x_2^k & x_3^k & \cdots & x_{k+1}^k \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_{k+1} - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_{k+1}(x_{k+1} - x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{k-1}(x_2 - x_1) & x_3^{k-1}(x_3 - x_1) & \cdots & x_{k+1}^{k-1}(x_{k+1} - x_1) \end{vmatrix}$$

按第 1 列展开，对 M_{11} 的每一列提取出公因数 $(x_i - x_1)$ ，其中 $i=2, 3, \dots, k+1$ ，有：

$$D_{k+1} = \prod_{i=2}^{k+1} (x_i - x_1) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_2 & x_3 & x_4 & \cdots & x_{k+1} \\ x_2^2 & x_3^2 & x_4^2 & \cdots & x_{k+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_2^{k-1} & x_3^{k-1} & x_4^{k-1} & \cdots & x_{k+1}^{k-1} \end{vmatrix} = \prod_{i=2}^{k+1} (x_i - x_1) \prod_{2 \leq j < i \leq k+1} (x_i - x_j) = \prod_{1 \leq j < i \leq k+1} (x_i - x_j)$$

证毕。

 **练习 1.9** 利用定理 1.2 求行列式 $D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2^2 & \cdots & 2^n \\ 3 & 3^2 & \cdots & 3^n \\ \vdots & \vdots & \ddots & \vdots \\ n & n^2 & \cdots & n^n \end{vmatrix}$

解 观察到第 i 行有公因数 i ，提取出来后得到一个 Vandermonde 转置行列式，即：

$$D_n = n! \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^{n-1} \\ 1 & 3 & 3^2 & \cdots & 3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \cdots & n^{n-1} \end{vmatrix} = n! \prod_{1 \leq j < i \leq n} (i - j) = \prod_{k=1}^n k!$$

1.2.3 块行列式

对于另一类比较特殊的行列式，命题 1.7 给出了这种“块行列式”的计算方法：

命题 1.7 (块行列式的计算)

形如 1.18 的行列式记为块状行列式，且有：

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1n} & b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} & b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} \quad (1.18)$$

证明 不妨假设由 a_{ij} 和 b_{ij} 组成的行列式记为 D_1 和 D_2 ，则需要证明 $D = D_1 D_2$ 即可。因为一个行列式总能化成下三角行列式，因此把 D_1 和 D_2 化成下三角形行列式有：

$$D_1 = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a'_{11} & & \\ \vdots & \ddots & \\ a'_{n1} & \cdots & a'_{nn} \end{vmatrix} = a'_{11} \cdots a'_{nn}$$

$$D_2 = \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} b'_{11} & & \\ \vdots & \ddots & \\ b'_{m1} & \cdots & b'_{mm} \end{vmatrix} = b'_{11} \cdots b'_{mm}$$

因此，在 D 中分别对 a_{ij} 所在行列和 b_{ij} 所在行列分别做变换，虽然会改变 c_{ij} ，但对 a_{ij} 和 b_{ij} 的行列变换不会互相影响，即：

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1n} & b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} & b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} a'_{11} & & & & & \\ \vdots & \ddots & & & & \\ a'_{n1} & \cdots & a'_{nn} & & & \\ c'_{11} & \cdots & c'_{1n} & b'_{11} & \cdots & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \\ c_{m1} & \cdots & c_{mn} & b_{m1} & \cdots & b'_{mm} \end{vmatrix} = a'_{11} \cdots a'_{nn} b'_{11} \cdots b'_{mm} = D_1 D_2$$

原命题得证。

块行列式的计算对于第 2 章节分块矩阵的矩阵具有重要意义，下一页给出块行列式变种的练习证明 1.10：

 **练习 1.10** 证明另一种块行列式的计算满足:

$$D = \begin{vmatrix} 0 & \cdots & 0 & a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n1} & \cdots & a_{nn} \\ b_{11} & \cdots & b_{1m} & c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} & c_{m1} & \cdots & c_{mn} \end{vmatrix} = (-1)^{mn} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} \quad (1.19)$$

证明 只需将第 $m+k$ 列 ($k=1,2,\dots,n$) 依次相邻列交换到第 k 列, 就转化为了式 1.18。在这一过程中一共交换列 mn 次, 即证。

 **练习 1.11** 称除了主(次)对角线及与其相邻的斜线, 再加上第一行或第一列外, 其他元素均为零的行列式为 Hessenberg 型行列式。例如计算:

$$D_n = \begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & n-2 & 2-n & 0 \\ 0 & 0 & 0 & \cdots & n-1 & 1-n \end{vmatrix}$$

解 观察到主对角线上元素和同行相邻元素为相反数, 因此考虑累加法, 即 $c_{j1}(1)(j=2,3,\dots,n)$:

$$D_n = \begin{vmatrix} \frac{n(n+1)}{2} & 2 & 3 & \cdots & n-1 & n \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & n-2 & 2-n & 0 \\ 0 & 0 & 0 & \cdots & n-1 & 1-n \end{vmatrix}$$

按照第一列展开, 有:

$$\begin{aligned} D_n &= \frac{n(n+1)}{2} \begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ 2 & -2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & n-2 & 2-n & 0 \\ 0 & \cdots & 0 & n-1 & 1-n \end{vmatrix} \\ &= (-1)^{n-1} \frac{(n+1)!}{2} \end{aligned}$$

1.3 Cramer 法则 (1)

定义 1.6 (齐次线性方程组与非齐次线性方程组)

设有 n 个未知量和 m 个方程的线性方程组:

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

若常数项 $b_1, b_2, \cdots, b_m$ 不全为零, 则该方程组是一个非齐次线性方程组; 如果所有的常数项都为 0, 则该方程组是一个齐次线性方程组。

我们来讨论线性方程组的解的计算。对于一个齐次线性方程组 1.20 来说, $x_1 = x_2 = \cdots = x_n = 0$ 一定是它的解, 称它为 n 元齐次线性方程组的零解。若有一组不全为零的数是齐次线性方程组的解, 则称其为齐次线性方程组的非零解。

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \quad (1.20)$$

对于一个更具有一般性的非齐次线性方程组, 有定理 1.3:

定理 1.3 (Cramer 法则)

对于有 n 个未知量和 n 个方程的线性方程组而言, 如果系数行列式不为 0, 即:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

那么方程组有唯一解, 且有 $x_j = \frac{D_j}{D}$ 。其中 D_j 是把系数行列式 D 中第 j 列的元素依次用方程组右端的常数项替代后得到的 n 阶行列式。

对于该法则, 我们将在逆矩阵小节给予证明。现给出重要推论:

命题 1.8 (齐次线性方程组的零解判定)

对于 n 个未知量 n 个方程的齐次线性方程组 1.20, 如果它的系数行列式 $D \neq 0$, 则齐次线性方程组只有零解; 反之, 如果方程组有非零解, 则 $D = 0$ 。

 **练习 1.12** λ 取何值时, 齐次方程组 $\begin{cases} \lambda x_1 + x_2 + x_3 = 0 \\ x_1 + \lambda x_2 + x_3 = 0 \\ x_1 + x_2 + \lambda x_3 = 0 \end{cases}$ 有非零解?

解 根据命题 1.8, 有:

$$D = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda+2 & 1 & 1 \\ \lambda+2 & \lambda & 1 \\ \lambda+2 & 1 & \lambda \end{vmatrix} = (\lambda+2) \begin{vmatrix} 1 & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} = (\lambda+2)(\lambda-1)^2 = 0$$

解得 $\lambda = -2$ 或 $\lambda = 1$ 。

第一章 练习

(1) 计算行列式 $D_n = \det(a_{ij})$, 其中 $a_{ij} = |i - j|$, $i, j = 1, 2, \dots, n$

(2) 证明:

$$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-2} & a_2^{n-2} & \cdots & a_n^{n-2} \\ a_1^n & a_2^n & \cdots & a_n^n \end{vmatrix} = \left(\sum_{i=1}^n a_i \right) \left(\prod_{1 \leq j < i \leq n} (a_i - a_j) \right)$$

(3) 设 n 阶行列式 $A = 3$, 将 D 的每一列减去其余各列得到的新行列式为 B , 求 B .

(4) 设 x_1, x_2, x_3 是方程 $x^3 + px + q = 0$ 的三个根, 求行列式

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ x_3 & x_1 & x_2 \\ x_2 & x_3 & x_1 \end{vmatrix}$$

(5) 证明 Pascal 行列式的两种变形满足:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 & C_n^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 & C_{n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} & C_{2n-3}^{n-2} \\ 1 & C_n^{n-1} & C_{n+1}^{n-1} & \cdots & C_{2n-3}^{n-1} & C_{2n-2}^{n-1} \end{vmatrix} = 1$$

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 & C_n^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 & C_{n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} & C_{2n-3}^{n-2} \\ 1 & C_n^{n-1} & C_{n+1}^{n-1} & \cdots & C_{2n-3}^{n-1} & C_{2n-2}^{n-1} - 1 \end{vmatrix} = 0$$

提示: Pascal 行列式是数字排布类似于杨辉三角的行列式, 考虑杨辉三角的数字分布规律。

(6) (Cramer 法则的应用) 已知 $\log_{18} 9 = a$, $18^b = 5$, 用 a 和 b 表示 $\log_{36} 45$

(7) 计算行列式

$$D_{n+1} = \begin{vmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & a_1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & a_2 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 & a_n \end{vmatrix}$$

(8) 计算行列式

$$D_n = \begin{vmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & \cdots & a_1 b_n \\ a_1 b_2 & a_2 b_2 & a_2 b_3 & \cdots & a_2 b_n \\ a_1 b_3 & a_2 b_3 & a_3 b_3 & \cdots & a_3 b_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 b_n & a_2 b_n & a_3 b_n & \cdots & a_n b_n \end{vmatrix}$$

(9) 求多项式 $p(x) = \begin{vmatrix} 1 & x & -1 & x^2 + 1 \\ x & 1 & x^2 + 1 & -1 \\ -1 & x^2 + 1 & 1 & x \\ x^2 + 1 & -1 & x & 1 \end{vmatrix}$ 的常数项。

(10) 计算行列式

$$D_n = \begin{vmatrix} \alpha + \beta & \alpha\beta & 0 & \cdots & 0 & 0 \\ 1 & \alpha + \beta & \alpha\beta & \cdots & 0 & 0 \\ 0 & 1 & \alpha + \beta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha + \beta & \alpha\beta \\ 0 & 0 & 0 & \cdots & 1 & \alpha + \beta \end{vmatrix}$$

(11) 计算行列式

$$D_n = \begin{vmatrix} a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \\ a_1^{n-2}b_1 & a_2^{n-2}b_2 & a_3^{n-2}b_3 & \cdots & a_n^{n-2}b_n \\ \vdots & \vdots & \vdots & & \vdots \\ a_1b_1^{n-2} & a_2b_2^{n-2} & a_3b_3^{n-2} & \cdots & a_nb_n^{n-2} \\ b_1^{n-1} & b_2^{n-1} & b_3^{n-1} & \cdots & b_n^{n-1} \end{vmatrix}$$

(12) 设 n 阶行列式 D 的 (i, j) 元素表示为 $a_i + b_j (i, j = 1, 2, \dots, n)$, 求 D 。

第二章 矩阵

2.1 矩阵及其运算

2.1.1 矩阵的概念

在第一章中，我们介绍了行列式。在本章节中我们将介绍高等代数中一种全新的代数形式，具体见定义2.1：

定义 2.1 (矩阵的定义)

由 $m \times n$ 个数 $a_{ij} (i = 1, \dots, m; j = 1, \dots, n)$ 排成的 m 行、 n 列的矩形数表

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

称为一个 $m \times n$ 矩阵. 其中 a_{ij} 称为该矩阵的第 i 行第 j 列元素，简称为该矩阵的 (i, j) 元素。矩阵一般记为大写字母 A 或 $A_{m \times n}$ 或 $(a_{ij})_{m \times n}$ 或 (a_{ij}) ；

当 $m = n$ 时，我们称 $A_{n \times n}$ 为 n 阶方阵或 n 阶矩阵，也记为 A_n ；

元素是实(复)数的矩阵称为实(复)矩阵。



矩阵和行列式的不同之处在于，其可以表示行列数不同的 $m \times n$ 按照一定顺序排列而成的数组，而行列式只是一个同行列数的数表按照一定规则计算得到的数字。在讨论矩阵的运算之前，我们先介绍几种重要且特殊的矩阵：

1. 零矩阵：所有元素都为零的矩阵，记为 $O_{m \times n}$ 或 O

$$O = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

2. 单位矩阵：主对角线元素都是 1，其他元素都是 0 的 n 阶方阵，记为 I_n 或 I

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

3. 行矩阵 (n 维行向量) 与列矩阵 (m 维列向量)：通常用小写希腊字母表示

$$\alpha = [a_1, a_2, \dots, a_n] \quad \beta = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

4. 上(下)三角矩阵：和上(下)三角行列式一样的元素排布规则的方阵。

5. 对角矩阵 (方阵): 主对角元素之外都为 0 的 n 阶方阵记为对角矩阵 (方阵), 记为 $D = \text{diag}(d_1, d_2, \dots, d_n)$

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

2.1.2 矩阵的代数运算

矩阵作为代数的一种形式, 应当满足代数运算。但不同矩阵之间能否进行代数运算存在一定的判定条件:

定义 2.2 (同型矩阵和相等矩阵)

同型矩阵: 设矩阵 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$, 这里 A 与 B 的行数相等, A 与 B 的列数也相等, 把这样的矩阵 A 与 B 称为同型矩阵。

矩阵相等: 如果两个同型矩阵 A 与 B 的对应元素都相等, 即 $a_{ij} = b_{ij} (i = 1, \dots, m; j = 1, \dots, n)$, 则称 A 与 B 相等, 记为 $A = B$ 。

矩阵的行列数是否满足特定条件是能否进行代数运算的必要条件。下定义矩阵的加减法:

定义 2.3 (矩阵的加减法)

设 $A = (a_{ij})_{m \times n}$ 和 $B = (b_{ij})_{m \times n}$ 是两个同型的矩阵, 规定 A 与 B 的和是由 A 与 B 的对应元素相加所得到的 $m \times n$ 矩阵记为 $A + B$, 即:

$$A + B = (a_{ij} + b_{ij})_{m \times n}$$

规定矩阵 $B = (b_{ij})_{m \times n}$ 的负矩阵为 $-B = (-b_{ij})_{m \times n}$, 那么有矩阵的减法定义为:

$$A - B = A + (-B) = (a_{ij} - b_{ij})_{m \times n}$$

不难验证, 矩阵的加减法满足命题 2.1 的运算规律。

命题 2.1 (矩阵加减法的运算规律)

假设 A, B 都为 $m \times n$ 矩阵, 则:

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + O = A$
4. $A + (-A) = O$

由此看出, 同型矩阵的加减运算满足交换律、结合律。下面给出矩阵和数字相乘的定义:

定义 2.4 (矩阵的数乘)

设矩阵 $A = (a_{ij})_{m \times n}$, k 为数规定 k 与 A 的乘积是用 k 去乘 A 的每个元素所得到的 $m \times n$ 矩阵, 记为 kA (或 Ak), 即 $kA = (ka_{ij})_{m \times n}$

命题 2.2 (矩阵数乘的运算规律)

假设 A, B 都为 $m \times n$ 矩阵, k, l 为任意常数, 则:

1. $1A = A, (-1)A = -A, 0A = O$
2. $k(lA) = (kl)A$
3. $(k + l)A = kA + lA$
4. $k(A + B) = kA + kB$

矩阵的加法与数乘运算合起来, 统称为矩阵的线性运算。

 **练习 2.1** 设矩阵 $A = \begin{bmatrix} 1 & 7 & 5 \\ 3 & -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 4 & -2 \end{bmatrix}$, 且 $3B + 2X = 2A$, 求 X 。

解 由题目可知 $X = \frac{2A - 3B}{2} = A - \frac{3}{2}B$, 即:

$$\begin{aligned} X &= \begin{bmatrix} 1 & 7 & 5 \\ 3 & -1 & 2 \end{bmatrix} - \frac{3}{2} \begin{bmatrix} 2 & 4 & -2 \\ 0 & 4 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 7 & 5 \\ 3 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 6 & -3 \\ 0 & 6 & -3 \end{bmatrix} \\ &= \begin{bmatrix} -2 & 1 & 8 \\ 3 & -7 & 5 \end{bmatrix} \end{aligned}$$

矩阵的线性运算要求矩阵同型, 但对于不同型的矩阵之间也能进行代数运算, 即矩阵的乘法:

定义 2.5 (矩阵的乘法)

设矩阵 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$ 规定 A 与 B 的乘积为矩阵 $C = (c_{ij})_{m \times n}$, 记为 $AB = C$, 其中:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj} = \sum_{k=1}^s a_{ik}b_{kj}$$

这里 $i = 1, 2, \dots, m; j = 1, 2, \dots, n$, 即 AB 的 (i, j) 元素为 A 的第 i 行各元素分别与 B 的第 j 列对应元素的乘积之和。注意, AB 被叫做 A 左乘 B 或者 B 右乘 A 。



注 必须指出, 左边矩阵列数等于右边矩阵行数, 两个矩阵才能相乘; 乘积矩阵的行数、列数分别等于左边矩阵的行数, 右边矩阵的列数。当左矩阵的列数不等于右矩阵的行数时, 矩阵的乘法无意义。借助练习 2.2 可以得知重要结论: $AB = BA$ 不一定总成立。

 **练习 2.2** 设矩阵 $A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, $B = [b_1, b_2, \dots, b_n]$, 分别求 AB 和 BA 。

解 由定义 2.5 知两式都有意义, 则:

$$\begin{aligned} AB &= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix} \\ BA &= \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n \end{aligned}$$

同样的, 我们可以给出矩阵乘法满足的运算规律:

命题 2.3 (矩阵乘法的运算规律)

假设矩阵 A, B, C 参与的运算都有意义, I 为单位矩阵, k 为任意常数, 则:

1. $I_m A_{m \times n} = A_{m \times n} I_n = A_{m \times n}$
2. $O_m A_{m \times n} = A_{m \times n} O_n = O_{m \times n}$

3. $(AB)C = A(BC)$
4. $(kA)B = A(kB) = k(AB)$
5. $A(B+C) = AB+AC$
6. $(A+B)C = AC+BC$

在介绍了矩阵的概念和其代数运算后，我们再次观察下列线性方程组：

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases}$$

利用矩阵，上式可以表示为2.1：

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \iff y = Ax \quad (2.1)$$

实际上，矩阵 A 可以看做是 $x_1, x_2, \cdots, x_n$ 经过线性计算得到 $y_1, y_2, \cdots, y_m$ 的桥梁，规范化地称这个过程为一个线性变换；当上式表示一个线性方程组时， $y = Ax$ 的形式也被称为线性方程组的矩阵形式，而 A 记为该线性方程组的系数矩阵。

最后，对于一个方阵而言，不难证明他可以乘以自身，即幂的形式：

定义 2.6 (方阵的幂)

对于方阵 A 而言，规定 A 的幂：

$$A^0 = I, \quad A^k = A \cdots A (k \text{ 个 } A \text{ 相乘})$$

对于任意非负整数 k, l ，有：

$$A^k A^l = A^{k+l}, \quad (A^k)^l = A^{kl}$$

注 必须指出，当 $A \neq B$ 时，一些代数中常用的幂展开式结论如 $(a+b)(a-b) = a^2 - b^2$ 等不一定成立。因为矩阵的左乘和右乘不等价，因此不一定能消去两矩阵的乘积项。

2.1.3 矩阵的转置

类比行列式的转置，我们给出矩阵转置的定义：

定义 2.7 (矩阵的转置)

把 $m \times n$ 矩阵 $A = (a_{ij})_{m \times n}$ 的行依次换成列（列依次换成行）所得到的 $n \times m$ 矩阵，称为 A 的转置矩阵，记为 A^T ，即：

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

矩阵的转置满足一下运算法则：

命题 2.4 (矩阵转置的运算规律)

假设矩阵 A, B 参与的运算都有意义, k 为任意常数, 则:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(kA)^T = kA^T$
4. $(AB)^T = B^T A^T$

 **练习 2.3** 证明命题2.4的性质 (4)。

证明 由矩阵乘法的定义有:

$$(AB)_{ij}^T = \sum_{k=1}^s a_{jk} b_{ki}$$

$$(B^T A^T)_{ij} = \sum_{k=1}^s (B^T)_{ik} (A^T)_{kj} = \sum_{k=1}^s a_{jk} b_{ki}$$

这样就说明了 $(AB)^T$ 和 $B^T A^T$ 同型且每一个元素都相同, 即 $(AB)^T = B^T A^T$ 证毕。

定理 2.1

任何一个方阵 A 可以表示成一个对称矩阵 P 与一个反对称矩阵 Q 之和。

其中, 对称矩阵 P 和反对称矩阵 Q 是分别满足 $P^T = P$ 和 $Q^T = -Q$ 的矩阵。

证明 $\because A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$, 记 $P = \frac{1}{2}(A + A^T)$, $Q = \frac{1}{2}(A - A^T)$ 。不难证明 $P^T = P$, $Q^T = -Q$ 即证。

2.1.4 方阵的行列式

本小节的最后, 我们讨论行列式和矩阵的关系。一个方阵因为行列数相同, 其元素按照原来顺序一定可以组成并对应唯一的行列式, 这种特殊的对应关系表示一个数组和一个数之间的联系, 因此有定义2.8:

定义 2.8 (方阵的行列式)

对于一个 n 阶方阵 $A = (a_{ij})_{n \times n}$, 定义其行列式为:

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

借助矩阵的运算法则和行列式的性质, 也可以给出命题2.5:

命题 2.5 (方阵行列式的运算规律)

设有 n 阶方阵 A, B , k 为任意常数, m 为非负整数, 则:

1. $\det(A^T) = \det(A)$
2. $\det(kA) = k^n \det(A)$
3. $\det(AB) = \det(A) \det(B)$
4. $\det(A^m) = [\det(A)]^m$

2.2 逆矩阵

2.2.1 逆矩阵的定义

上一节中介绍了矩阵的加法、减法、数乘和乘积运算，那么能不能对矩阵定义“除法运算”呢？然而在人们对矩阵的研究中，他们发现代数中的除法运算是不能照搬到矩阵上的。回顾代数内容中“倒数的概念”：一个非零数 a 的倒数定义为 $\frac{1}{a}$ ，这是因为他们的乘积等于 1。如果运用到矩阵学中，有重要定义 2.9：

定义 2.9 (逆矩阵)

设 A 为 n 阶方阵，如果存在 n 阶方阵 B ，使得：

$$AB = BA = I$$

则称方阵 A 是可逆的 (非奇异矩阵)，并称方阵 B 为方阵 A 的逆矩阵或逆阵，记为 $B = A^{-1}$ 。不存在逆矩阵的方阵也称为奇异矩阵。

需要指出，并不是所有的方阵都存在逆矩阵。那么，在什么条件下方阵可逆呢？有定理 2.2：

定理 2.2 (方阵可逆的充要条件)

对于 $n(n \geq 2)$ 阶方阵 A 可逆的充要条件是 $\det(A) \neq 0$ ，且当 A 可逆时，有：

$$A^{-1} = \frac{1}{\det(A)} A^* \quad (2.2)$$

其中， A^* 被定义为 A 的伴随矩阵，由 A 的 (j, i) 元素的代数余子式 A_{ji} 构成，记 $A^* = \text{adj}(A)$ ，且：

$$A^* = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

证明

- 先证必要性：设 A 可逆，则存在同型方阵 B 使得 $AB = I$ ，由方阵的行列式性质有：

$$\det(AB) = \det(A) \det(B) = \det(I) = 1$$

因此一定有 $\det(A) \neq 0$ ；

- 再证充分性：因为 $\det(A) \neq 0$ ，给定同型方阵 $B = \frac{1}{\det(A)} A^*$ ，那么有：

$$AB = A \left(\frac{1}{\det(A)} A^* \right) = \frac{1}{\det(A)} AA^*$$

$$\text{由小节 1.1 推论，可以计算 } AA^* = A^*A = \begin{bmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det(A) \end{bmatrix} = \det(A)I, \text{ 故 } AB = BA = I \text{ 证毕。}$$

注 要验证方阵 B 是方阵 A 的逆矩阵，只需验证 $AB = I$ 或 $BA = I$ 中的一个就可以了，这可以由定义证明。

由上面的证明过程，有命题 2.6：

命题 2.6

设 A 为 $n(n \geq 2)$ 阶方阵，则成立 $AA^* = A^*A = \det(A)I$ 。

 **练习 2.4** 判断 $A = \begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$ 是否可逆, 如果可逆, 求出其逆矩阵。

解 由于 $\det(A) = -10 \neq 0$, 故 A 可逆。

$$A^* = \begin{bmatrix} -15 & 11 & 12 \\ 10 & -10 & -10 \\ 5 & -7 & -4 \end{bmatrix}$$

根据 $A^{-1} = \frac{1}{\det(A)} A^*$, 计算得

$$A^{-1} = \frac{1}{-10} \begin{bmatrix} -15 & 11 & 12 \\ 10 & -10 & -10 \\ 5 & -7 & -4 \end{bmatrix}$$

命题 2.7

如果 $n(n \geq 2)$ 阶方阵 A 的行列式 $\det(A) \neq 0$, 则:

$$\det(A^*) = [\det(A)]^{n-1} \quad (2.3)$$

证明 在公式 $AA^* = \det(A)I$ 两端取行列式, 得 $\det(A) \det(A^*) = [\det(A)]^n$, 因为 $\det(A) \neq 0$, 即证 $\det(A^*) = [\det(A)]^{n-1}$ 。

对于一些特殊的矩阵, 比如对角矩阵等, 计算其逆矩阵有一定的规律, 具体有命题 2.8:

命题 2.8 (对角矩阵的逆矩阵)

对于对角矩阵 A 和 B 而言, 如果其对角线元素都不为零, 即:

$$A = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & \ddots \\ & & & a_n \end{pmatrix}, B = \begin{pmatrix} & & & b_1 \\ & & & \\ & & b_2 & \\ & & & \ddots \\ b_n & & & \end{pmatrix}$$

那么, A^{-1} 和 B^{-1} 可以表示为:

$$A^{-1} = \begin{pmatrix} a_1^{-1} & & \\ & a_2^{-1} & \\ & & \ddots \\ & & & a_n^{-1} \end{pmatrix}, B^{-1} = \begin{pmatrix} & & & b_1^{-1} \\ & & & \\ & & b_2^{-1} & \\ & & & \ddots \\ b_1^{-1} & & & \end{pmatrix}$$

对于上面的命题, 由于对角线矩阵的行列式一定包含对角线元素乘积这一项, 只要每一个元素非零, 那么矩阵一定可逆, 剩下的只需要代入验证就行了。一般的, 如果要从已知条件式中证明某矩阵可逆, 只需要通过适当的变换, 使得找到一个与其相乘等于单位矩阵 I 的同型矩阵即证, 可参考练习 2.5:

 **练习 2.5** 设 n 阶方阵 A 满足 $A^2 + A - 4I = O$, 证明 $A - I$ 可逆, 并求 $(A - I)^{-1}$ 。

解 因为:

$$O = A^2 + A - 4I = A^2 - A + 2A - 2I - 2I$$

移项化简得:

$$(A + 2I)(A - I) = 2I$$

因此有:

$$\frac{1}{2}(A + 2I)(A - I) = I$$

这说明 $A - I$ 可逆, 且 $(A - I)^{-1} = \frac{1}{2}(A + 2I)$ 。

2.2.2 逆矩阵的性质

命题 2.9 (逆矩阵的运算规律)

设有 n 阶可逆方阵 A, B , k 为任意非零常数, 则:

1. $(A^{-1})^{-1} = A$
2. $(A^T)^{-1} = (A^{-1})^T$
3. $(kA)^{-1} = \frac{1}{k}A^{-1}$
4. $(AB)^{-1} = B^{-1}A^{-1}$
5. $\det(A^{-1}) = \frac{1}{\det(A)}$

上述性质都不难证明, 先给出练习 2.6 以巩固:

 **练习 2.6** 设 A 为 3 阶方阵, $\det(A) = \frac{1}{2}$, A^* 为 A 的伴随矩阵, 求行列式 $\det[(3A)^{-1} - 2A^*]$ 的值。

解 由于 $(3A)^{-1} = \frac{1}{3}A^{-1}$, $2A^* = 2\det(A)A^{-1} = A^{-1}$, 所以:

$$\begin{aligned}\text{原式} &= \det\left(\frac{1}{3}A^{-1} - A^{-1}\right) = \det\left(-\frac{2}{3}A^{-1}\right) = \left(-\frac{2}{3}\right)^3 \det(A^{-1}) \\ &= \frac{-8}{27} \frac{1}{\det(A)} = -\frac{16}{27}\end{aligned}$$

关于逆矩阵性质的进一步探讨, 详见本章练习参考题。

2.2.3 Cramer 法则 (2) 和矩阵方程

小节 1.3 中的定理 1.3 即 Cramer 法则讨论了线性方程组的解的计算方法, 现在我们给予证明:

证明 假设方程组表示为 $Ax = b$ 的形式, 由于 $\det(A) \neq 0$, 因此等式两边左乘 A^{-1} 有:

$$x = A^{-1}b = \frac{1}{\det(A)}A^*b = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (2.4)$$

根据矩阵乘法, 有:

$$x_j = \frac{1}{\det(A)}(b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj})$$

而对于将 A 的第 j 列替换为 b 得到的方阵 A_j 而言, 有:

$$\det(A_j) = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj}$$

因此得到 $x_j = \frac{\det(A_j)}{\det(A)}$, 证毕。

上述证明告诉我们, 除了使用 Cramer 法则直接计算, 还可以通过求 A 的逆矩阵 A^{-1} 和 b 的矩阵乘积得解。这种以矩阵为未知数的方程也被称为**矩阵方程**。对于矩阵方程, 当系数矩阵是方阵时, 需要先判断是否可逆, 例如对于一般性矩阵方程:

$$AXB = F \quad (2.5)$$

当 A 和 B 均可逆时, 也就有解 $X = A^{-1}FB^{-1}$ 。此时, 只要掌握求解逆矩阵的方法就能快速求解此类矩阵方程。

 **练习 2.7** 设矩阵 X 满足矩阵方程 $AX = 2X + B$, 其中:

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 6 \\ 1 & 1 \\ 2 & -3 \end{bmatrix}$$

求矩阵 X 。

解 由 $AX = 2X + B$, 得 $AX - 2X = B$, 即 $(A - 2I)X = B$ 。经计算知 $\det(A - 2I) \neq 0$, 从而 $X = (A - 2I)^{-1}B$, 又因为:

$$(A - 2I)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

所以有:

$$X = (A - 2I)^{-1}B = \begin{bmatrix} \frac{3}{2} & 3 \\ -4 & 1 \\ 3 & -2 \end{bmatrix}$$

最后, 我们介绍矩阵多项式的概念和一个性质:

命题 2.10 (矩阵多项式)

设 m 次多项式 $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$, A 为方阵, 记 $f(A) = a_0I + a_1A + a_2A^2 + \cdots + a_mA^m$, 称 $f(A)$ 为方阵 A 的多项式。设 k 为正整数, 若存在方阵 P 及 B , 使 $A = PBP^{-1}$, 则 $A^k = PB^kP^{-1}$, 且 $f(A) = Pf(B)P^{-1}$ 。

证明 设 k 为正整数, 若存在可逆矩阵 P 及方阵 B , 使 $A = PBP^{-1}$, 则

$$A^k = (PBP^{-1})^k = PBP^{-1}PBP^{-1} \cdots PBP^{-1} = PB^kP^{-1}.$$

于是

$$\begin{aligned} f(A) &= a_0I + a_1A + a_2A^2 + \cdots + a_mA^m \\ &= a_0PIP^{-1} + a_1PBP^{-1} + a_2PB^2P^{-1} + \cdots + a_mPB^mP^{-1} \\ &= P(a_0I + a_1B + a_2B^2 + \cdots + a_mB^m)P^{-1} \\ &= Pf(B)P^{-1}. \end{aligned}$$

结论得证。

命题2.10给我们计算方阵多项式的一个简单思路, 当我们能找到 P 和对角矩阵 B , 使得 $A = PBP^{-1}$ 时, 由于对角矩阵的性质 (命题2.8) 可以快速计算出 $f(B)$, 因此就很简单的计算出了 $f(A)$ 。

2.3 分块矩阵

2.3.1 分块矩阵的定义

本节介绍在处理行数、列数较高的矩阵时的常用技巧——矩阵分块法。通过矩阵的分块，可将大矩阵的运算转化成小矩阵的运算，从而简化运算。

定义 2.10 (子矩阵)

对于行数、列数较高的矩阵，考虑它的若干行和若干列相交处的元素按原来的相对次序所构成的矩阵，称它为原矩阵的子矩阵。

特别地，对于 n 阶方阵，称删去最后 $n-k$ 行和 $n-k$ 列得到的子矩阵成为其 k 阶顺序主子矩阵或前主子矩阵。

举例而言，对于矩阵 $A = \begin{bmatrix} 1 & -1 & 8 & 2 & 3 \\ 0 & 7 & 9 & 5 & 2 \\ 3 & 4 & 5 & 6 & 7 \end{bmatrix}$ ，可以找到他的一部分子矩阵：

$$\begin{bmatrix} 1 & 8 \end{bmatrix} \quad \begin{bmatrix} -1 & 8 & 3 \\ 4 & 5 & 7 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 0 & 5 \\ 3 & 6 \end{bmatrix}$$

定义 2.11 (分块矩阵)

对于一些行列数较大的矩阵而言，通过将其分为若干个小型子矩阵的表示方法称为分块矩阵，即：

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{bmatrix}$$

其中， A_{ij} 为 $m_i \times n_j$ 矩阵 ($i = 1, 2, \dots, s; j = 1, 2, \dots, t$)。

分块矩阵可以直观地表示一些矩阵，以 $A = \begin{bmatrix} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 & -2 \\ 0 & 0 & 1 & -1 & 6 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$ 为例，常见的分块方式可以表示为：

- 按分块子矩阵分，可以把 A 表示为 $A = \begin{bmatrix} I_3 & A_{12} \\ O_{2 \times 3} & A_{22} \end{bmatrix}$ ；
- 按列分块，可以把 A 表示为 $A = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 \end{bmatrix}$ ；
- 按行分块，可以把 A 表示为 $A = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix}$ ；

注 分块矩阵的核心是把子矩阵看做是矩阵的元素，从而简化表示和计算。

对于主对角线上各子块都为方阵，其他部分子块为零的方阵，称之为分块对角矩阵或准对角方阵。

2.3.2 分块矩阵的运算

分块矩阵的运算具有和普通矩阵相似的代数运算规则:

命题 2.11 (分块矩阵的代数运算)

对同型矩阵 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{m \times n}$ 作同样的划分, 得分块矩阵

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & \ddots & \vdots \\ A_{s1} & \cdots & A_{st} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & \cdots & B_{1t} \\ \vdots & \ddots & \vdots \\ B_{s1} & \cdots & B_{st} \end{bmatrix}$$

1. 分块矩阵的加法, 按照矩阵的加法的定义, 有:

$$A + B = \begin{bmatrix} A_{11} + B_{11} & \cdots & A_{1t} + B_{1t} \\ \vdots & \ddots & \vdots \\ A_{s1} + B_{s1} & \cdots & A_{st} + B_{st} \end{bmatrix}$$

2. 分块矩阵的数乘:

$$kA = \begin{bmatrix} kA_{11} & \cdots & kA_{1t} \\ \vdots & \ddots & \vdots \\ kA_{s1} & \cdots & kA_{st} \end{bmatrix} \quad (k \text{ 为常数})$$

3. 分块矩阵的转置等于每一个子矩阵的转置矩阵按转置顺序排布:

$$A^T = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{bmatrix}^T = \begin{bmatrix} A_{11}^T & A_{21}^T & \cdots & A_{s1}^T \\ A_{12}^T & A_{22}^T & \cdots & A_{s2}^T \\ \vdots & \vdots & \ddots & \vdots \\ A_{1t}^T & A_{2t}^T & \cdots & A_{st}^T \end{bmatrix}$$

4. 分块矩阵的乘法, 需要重新对 A 和 B 分块, 即设矩阵 $A = (a_{ij})_{m \times l}, B = (b_{ij})_{l \times n}$ 分块成:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{sr} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rt} \end{bmatrix}$$

其中 $A_{i1}, A_{i2}, \dots, A_{ir}$ 的列数分别等于 $B_{1j}, B_{2j}, \dots, B_{rj}$ 的行数 ($i = 1, 2, \dots, s; j = 1, 2, \dots, t$), 则有:

$$AB = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1t} \\ C_{21} & C_{22} & \cdots & C_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ C_{s1} & C_{s2} & \cdots & C_{st} \end{bmatrix}$$

其中 $C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ir}B_{rj} = \sum_{k=1}^r A_{ik}B_{kj} (i = 1, 2, \dots, s; j = 1, 2, \dots, t)$

 **练习 2.8** 设 $\alpha_j (j = 1, 2, 3)$ 均为 3 维列向量, 方阵 $A = [\alpha_1, \alpha_2, \alpha_3]$, $B = [\alpha_1 + 2\alpha_2, 2\alpha_2 + 3\alpha_3, 3\alpha_3 + \alpha_1]$, 已知 $\det(A) = a$, 求 $\det(B)$.

解 由矩阵的乘法可得:

$$\alpha_1 + 2\alpha_2 = [\alpha_1, \alpha_2, \alpha_3] \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = A \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$2\alpha_2 + 3\alpha_3 = [\alpha_1, \alpha_2, \alpha_3] \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} = A \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$$

$$1\alpha_1 + 3\alpha_3 = [\alpha_1, \alpha_2, \alpha_3] \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$

$$\text{则 } B = \left[A \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad A \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \quad A \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \right] \Rightarrow B = AP = A \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 3 & 3 \end{bmatrix}. \text{ 所以有 } \det(B) = \det(A) \det(P) = 12a.$$

 **笔记** 特别地，设有两个分块对角矩阵：

$$A = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_m \end{bmatrix} \quad B = \begin{bmatrix} B_1 & & \\ & B_2 & \\ & & \ddots \\ & & & B_m \end{bmatrix}$$

其中 A_i, B_i 为同阶方阵， $i = 1, 2, \dots, m$ ，则有：

$$AB = \begin{bmatrix} A_1 B_1 & & \\ & A_2 B_2 & \\ & & \ddots \\ & & & A_m B_m \end{bmatrix} \quad A^n = \begin{bmatrix} A_1^n & & \\ & A_2^n & \\ & & \ddots \\ & & & A_m^n \end{bmatrix}$$

对于分块矩阵的逆矩阵求法，只需要按照逆矩阵的定义就可以直接求解，见练习2.9：

 **练习 2.9** 设 $A_{n \times n}, B_{m \times m}$ 均为方阵， C 为任意的 $n \times m$ 矩阵，问分块上三角形矩阵 $D = \begin{bmatrix} A & C \\ O & B \end{bmatrix}$ 何时可逆？当可逆时，求其逆矩阵。

解 由于 $\det(D) = \det(A) \det(B)$ ，所以矩阵 D 可逆 $\Leftrightarrow A, B$ 均可逆。

设矩阵 D 的逆为 $D^{-1} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ ，其中 X_{11} 是 n 阶方阵， X_{22} 是 m 阶方阵

根据逆矩阵的定义，有 $\begin{bmatrix} A & C \\ O & B \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} I_n & O \\ O & I_m \end{bmatrix}$

即得到：

$$AX_{11} + CX_{21} = I_n, AX_{12} + CX_{22} = O, BX_{21} = O, BX_{22} = I_m$$

$$\text{从而解得 } D^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}CB^{-1} \\ O & B^{-1} \end{bmatrix}$$

从练习2.9中看出，当 $C = O$ 时， $X_{12} = O$ ，从而 $D^{-1} = \begin{bmatrix} A^{-1} & O \\ O & B^{-1} \end{bmatrix}$ ，这暗示了分块对角矩阵的逆矩阵求法，

并类比命题2.8有：

命题 2.12 (分块对角矩阵的逆矩阵)

对于分块对角矩阵 A 和 B 而言：

$$A = \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_n \end{pmatrix}, B = \begin{pmatrix} & & & B_1 \\ & & & \\ & & B_2 & \\ & \ddots & & \\ B_n & & & \end{pmatrix}$$

那么, A^{-1} 和 B^{-1} 可以表示为：

$$A^{-1} = \begin{pmatrix} A_1^{-1} & & \\ & A_2^{-1} & \\ & & \ddots \\ & & & A_n^{-1} \end{pmatrix}, B^{-1} = \begin{pmatrix} & & & B_n^{-1} \\ & & & \\ & & B_1^{-1} & \\ & \ddots & & \\ B_2^{-1} & & & \end{pmatrix}$$

 **练习 2.10** 设 $A = \begin{bmatrix} 0 & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1} \\ a_n & 0 & 0 & \dots & 0 \end{bmatrix}$, 其中 $a_i \neq 0$, 求 A^{-1} 。

解 改写 $A = \begin{bmatrix} O & A_1 \\ A_2 & O \end{bmatrix}$, $A^{-1} = \begin{bmatrix} O & A_2^{-1} \\ A_1^{-1} & O \end{bmatrix}$, 其中 $A_2^{-1} = a_n^{-1}$,

$$\therefore A^{-1} = \begin{bmatrix} a_1^{-1} & & & \\ & \ddots & & \\ & & a_{n-1}^{-1} & \\ & & & \end{bmatrix}, \therefore A^{-1} = \begin{bmatrix} 0 & \dots & 0 & a_n^{-1} \\ a_1^{-1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & a_{n-1}^{-1} & 0 \end{bmatrix}.$$

2.4 初等变换与初等矩阵

2.4.1 矩阵初等变换的概念

本小节具有较强理论性，将给出数个矩阵学重要定义和命题。

定义 2.12 (初等行 (列) 变换)

对矩阵施行的下列 3 种变换称为初等行变换：

1. 交换第 i 行与第 j 行的位置，记为 r_{ij} ；
2. 用非零数 k 乘以矩阵的第 i 行，记为 $r_i(k)$ ；
3. 把矩阵的第 i 行的 k 倍加到第 j 行上去，记为 $r_{ij}(k)$ 。

类似的，对矩阵施行的下列 3 种变换称为初等列变换：

1. 交换第 i 列与第 j 列的位置，记为 c_{ij} ；
2. 用非零数 k 乘以矩阵的第 i 列，记为 $c_i(k)$ ；
3. 把矩阵的第 i 列的 k 倍加到第 j 列上去，记为 $c_{ij}(k)$ 。

矩阵的初等行变换和初等列变换统称为初等变换。

为了描述原矩阵经过初等变换带来的新矩阵之间的关系，有定义 2.13：

定义 2.13 (初等行 (列) 变换)

如果矩阵 A 经过有限次初等行 (列) 变换变成矩阵 B ，则称矩阵 A 与 B 行 (列) 等价。矩阵行等价和矩阵列等价统称为矩阵等价，记作 $A \cong B$ 。

矩阵等价关系具有下列性质：

1. 自反性： $A \cong A$ ；
2. 对称性： $A \cong B \Leftrightarrow B \cong A$ ；
3. 传递性： $A \cong B, B \cong C \Rightarrow A \cong C$ 。

进一步的，给出三种重要的矩阵：

定义 2.14 (初等矩阵)

对单位矩阵只作 1 次初等变换所得到的矩阵，称为初等矩阵。3 种初等变换对应 3 种初等矩阵：

1. 互换单位矩阵的第 i 行 (列) 和第 j 行 (列) 得到的初等矩阵：

$$P_{ij} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1_{ij} & \\ & & \vdots & \ddots & \vdots & \\ & & 1_{ji} & \cdots & 0 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad (2.6)$$

2. 用非零数 k 乘单位矩阵的第 i 行 (列) 得到的初等矩阵:

$$P_i(k) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & k_{ii} & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \quad (2.7)$$

3. 把单位矩阵的第 i 行的 k 倍加到第 j 行上去 (或把单位矩阵的第 j 列的 k 倍加到第 i 列上去) 得到的初等矩阵:

$$P_{ij}(k) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ & & k_{ji} & \cdots & 1 \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix} \quad (2.8)$$

不难看出三种初等矩阵都可逆, 且三种初等矩阵各自对应一种初等变换, 其逆矩阵也为初等矩阵 (并且是同类型的), 初等矩阵对应的初等变换与它的逆矩阵对应的初等变换互为逆变换。

2.4.2 初等矩阵和初等变换的关系

定理 2.3 (初等矩阵和初等变换的关系)

对矩阵 A 施行一次初等行变换, 相当于对 A 左乘一个相应的初等矩阵; 对矩阵 A 施行一次初等列变换, 相当于对 A 右乘一个相应的初等矩阵。

用矩阵乘法表示初等行变换:

$$\begin{aligned} A &\xrightarrow{r_{ij}} B, \text{ 则 } B = P_{ij}A \\ A &\xrightarrow{r_i(k)} B, \text{ 则 } B = P_i(k)A \\ A &\xrightarrow{r_{ij}(k)} B, \text{ 则 } B = P_{ij}(k)A \end{aligned}$$

用矩阵乘法表示初等列变换:

$$\begin{aligned} A &\xrightarrow{c_{ij}} B, \text{ 则 } B = AP_{ij} \\ A &\xrightarrow{c_i(k)} B, \text{ 则 } B = AP_i(k) \\ A &\xrightarrow{c_{ji}(k)} B, \text{ 则 } B = AP_{ij}(k) \end{aligned}$$

由定义 2.13 可知, 如果矩阵 A 和 B 是行等价的, 那么就相当于做了有限次初等行变换, 即存在初等矩阵 $P_1, P_2, \dots, P_s$ 使得:

$$P_s \cdots P_2 P_1 A = PA = B$$

其中 $P = P_s \cdots P_2 P_1$ 可逆; 类似的也可以表示列等价矩阵的形式。由此有定理 2.4:

定理 2.4 (等价矩阵的关系)

矩阵 A 和 B 等价的充要条件是存在初等矩阵 $P_1, P_2, \dots, P_s, Q_1, Q_2, \dots, Q_t$ 使得:

$$P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t = B \quad (2.9)$$

若令 $P = P_s \cdots P_2 P_1, Q = Q_1 Q_2 \cdots Q_t$, 则有:

$$PAQ = B \quad (2.10)$$

其中 P 和 Q 为可逆矩阵。因此, 当矩阵 A 与 B 等价时, 必存在可逆矩阵 P 和 Q 使 $PAQ = B$, 其中 P, Q 可以表示成若干初等矩阵之积。

**2.4.3 阶梯型矩阵****定义 2.15 (阶梯型矩阵和行最简形)**

称满足下列两个条件的矩阵为阶梯形矩阵:

1. 若有零行 (元素全为零的行), 则零行都位于底部;
2. 各非零行的首非零元素都位于前一行首非零元的右边。

特别地, 称满足下列条件的矩阵为简化行阶梯形矩阵 (行最简形):

1. 阶梯形矩阵;
2. 各非零行的首非零元均为 1;
3. 首非零元所在列其它元素均为 0。



事实上, 已经能够证明对于任一非零矩阵 $A = (a_{ij})_{m \times n}$ 都可通过有限次初等行变换把它化成阶梯形矩阵, 进一步可化为简化行阶梯型矩阵 (行最简形)。我们给出练习 2.11:

练习 2.11 用初等行变换将矩阵 $A = \begin{bmatrix} 0 & 4 & -12 & 2 \\ 3 & -1 & -6 & -2 \\ -1 & -1 & 6 & 2 \end{bmatrix}$ 化成阶梯型矩阵和行最简形。

解

$$\begin{aligned} A &= \begin{bmatrix} 0 & 4 & -12 & 2 \\ 3 & -1 & -6 & -2 \\ -1 & -1 & 6 & 2 \end{bmatrix} \xrightarrow{r_{13}} \begin{bmatrix} -1 & -1 & 6 & 2 \\ 3 & -1 & -6 & -2 \\ 0 & 4 & -12 & 2 \end{bmatrix} \xrightarrow{r_{12}(3)} \begin{bmatrix} -1 & -1 & 6 & 2 \\ 0 & -4 & 12 & 4 \\ 0 & 4 & -12 & 2 \end{bmatrix} \xrightarrow{r_{23}(1)} \begin{bmatrix} -1 & -1 & 6 & 2 \\ 0 & -4 & 12 & 4 \\ 0 & 0 & 0 & 6 \end{bmatrix} \\ &\xrightarrow{r_2(-1/4)} \begin{bmatrix} 1 & 1 & -6 & -2 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_{21}(-1)} \begin{bmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_{31}(1), r_{32}(1)} \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

如果对上述 A 再次施展初等列变换, 那么可以把 A 变换为形如 $A = \begin{bmatrix} I & O \\ O & O \end{bmatrix}$ 的形式, 称为 A 在等价意义下的标准形, 详见后一小节。

命题 2.13 (等价矩阵的等价描述)

设 A 为 n 阶方阵, 则下列条件相互等价:

1. A 是可逆矩阵;
2. A 可经有限次初等行变换化成同阶单位矩阵 (即 A 行等价于同阶单位阵);
3. A 可表示成若干个初等矩阵之积。



证明 采用循环证法

1 → 2: 由定义知一个可逆矩阵经过有限次初等行变换化成行最简形就是单位矩阵;

2 → 3: 存在初等矩阵 $P_1, P_2, \dots, P_m$ 使 $P_m \cdots P_2 P_1 A = I$ 。故 $A = (P_m \cdots P_2 P_1)^{-1} = P_1^{-1} P_2^{-1} \cdots P_m^{-1}$

3 → 1: 设 $A = Q_1 Q_2 \cdots Q_m$, 其中 $Q_i (i = 1, 2, \dots, m)$ 为初等矩阵。有 $\det(Q_i) \neq 0 (i = 1, 2, \dots, m)$ 。故 $\det(A) \neq 0$ 即 A 可逆。由此循环论证就表述了命题的正确性。

2.4.4 利用初等变换求逆矩阵

利用本小节中对初等矩阵和初等变换的讨论, 可以给出一个重要命题:

命题 2.14 (逆矩阵的初等变换求法)

设 A 是可逆矩阵, 则一定有:

$$\begin{aligned} \begin{bmatrix} A & I \end{bmatrix} &\xrightarrow{\text{若干次初等行变换}} \begin{bmatrix} I & A^{-1} \end{bmatrix} \\ \begin{bmatrix} A \\ I \end{bmatrix} &\xrightarrow{\text{若干次初等列变换}} \begin{bmatrix} I \\ A^{-1} \end{bmatrix} \end{aligned}$$

这个命题显然成立, 因为对矩阵进行初等变换操作等价于左右乘初等矩阵, 再结合逆矩阵的定义即证。

 **练习 2.12** 利用命题 2.14 求可逆矩阵 $A = \begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ 的逆矩阵。

解 简便起见省略行变换过程, 但可以得到:

$$\begin{bmatrix} 2 & 2 & 3 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{若干次初等行变换}} \begin{bmatrix} 1 & 0 & 0 & 1 & -4 & -3 \\ 0 & 1 & 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{bmatrix}$$

因此解得 $A^{-1} = \begin{bmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{bmatrix}$

2.5 矩阵的秩

2.5.1 秩的概念

定义 2.16 (k 阶子式)

在 $m \times n$ 矩阵 A 中, 任取 k 行与 k 列 ($k \leq m, k \leq n$), 位于这些行列交叉处的 k^2 个元素, 不改变它们在 A 中所处的位置次序而得的 k 阶行列式, 称为矩阵 A 的一个 k 阶子式。

注 需要区分 k 阶子式和子矩阵的概念, 前者是一个行列式, 而后者是一个矩阵。

为了表述一个矩阵在初等变换中更一般性的性质, 引入秩的概念:

定义 2.17 (矩阵的秩)

如果 $A = O$, 则称 A 的秩为零; 如果 $A \neq O$, 则称 A 中非零子式的最高阶数为 A 的秩, 记为 $r(A)$ 。

设 A 为 $m \times n$ 矩阵:

- 若 $r(A) = m$, 则称 A 为行满秩矩阵;
- 若 $r(A) = n$, 则称 A 为列满秩矩阵;

设 A 为 $n \times n$ 矩阵:

- 若 $r(A) = n$, 则称 A 为满秩方阵 (可逆矩阵、非奇异方阵);
- 若 $r(A) < n$, 则称 A 为降秩方阵 (不可逆方阵、奇异方阵)。

不难证明矩阵的秩满足下面的性质:

1. 若 $m \times n$ 矩阵 $A \neq O$, 则 $1 \leq r(A) \leq \min\{m, n\}$;
2. $r(A) = r(A^T)$;
3. $r(A) = r \Leftrightarrow A$ 中至少存在一个 r 阶非零子式, 而且 A 中所有的 $r+1$ 阶子式 (如果存在的话) 全为零。

更多关于秩的性质详见第四章的讨论。

练习 2.13 设 4 阶方阵 $A = \begin{bmatrix} 1 & a & a & a \\ a & 1 & a & a \\ a & a & 1 & a \\ a & a & a & 1 \end{bmatrix}$ 的秩为 3, 试求常数 a 的值。

解 由条件知 A 为降秩方阵, 所以有 $\det(A) = 0$ 。计算可得:

$$\det(A) = \begin{vmatrix} 1 & a & a & a \\ a & 1 & a & a \\ a & a & 1 & a \\ a & a & a & 1 \end{vmatrix} = (1+3a)(1-a)^3$$

故得 $a = -\frac{1}{3}$ 或 $a = 1$ 。若 $a = 1$, 显然有 $r(A) = 1$ 不合题意; 而当 $a = -\frac{1}{3}$ 时, A 的左上角的 3 阶子式等于 $\frac{16}{27}$, 故 $a = -\frac{1}{3}$ 。

2.5.2 计算矩阵秩的一般方法

现在讨论计算矩阵的秩的一般方法. 利用定义计算矩阵的秩, 需要计算一些子式, 当子式的阶数较高时, 计算量一般太大, 显然是不方便的。但我们知道阶梯形矩阵的秩由“数”出其非零行的个数就可得到, 而任一矩阵又可由初等变换化成阶梯形, 那么, 是否可通过化矩阵为阶梯形矩阵来求矩阵的秩呢? 这就涉及一个问题: 矩阵的初等变换是否保持矩阵的秩不改变呢? 下面的定理回答了这一问题。

定理 2.5 (等价矩阵的秩不变性)

设矩阵 A 经若干次初等变换变成了矩阵 B , 则 $r(A) = r(B)$, 即等价矩阵有相同的秩。

证明 我们以只做行变换为例, 列变换的情况同理。

先证明若 A 经 1 次初等行变换变成 B , 则成立 $r(A) \leq r(B)$:

设 $r(A) = k$, 则 A 中必存在非零的 k 阶子式 D_k 。

- 当 $A \xrightarrow{r_{ij}} B$ 时或 $A \xrightarrow{r_i(\lambda)} B(\lambda \neq 0)$ 时, 在 B 中总可找到与 D_k 对应的 k 阶子式 M_k , 且有 $M_k = D_k$ 或 $M_k = -D_k$ 或 $M_k = \lambda D_k$, 因而 $M_k \neq 0$, 从而有 $r(B) \geq k = r(A)$;
- 当 $A \xrightarrow{r_{ij}(\lambda)} B$ 时, 由于作变换 r_{ij} 时结论成立, 所以只需考虑当 $A \xrightarrow{r_{21}(\lambda)} B$ 时的特殊情形, 此时只有以下 3 种可能: (1) D_k 不含 A 的第 1 行; (2) D_k 含 A 的第 1 行和第 2 行; (3) D_k 含 A 的第 1 行但不含 A 的第 2 行。现在取 B 中与 D_k 同序号的行和列, 构成 B 的 k 阶子式 M_k , 在 (1)、(2) 两种情形, 有 $M_k = D_k \neq 0$, 所以有 $r(B) \geq k = r(A)$; 在情形 (3), 记 M_k 为

$$M_k = \begin{vmatrix} r_1 + \lambda r_2 \\ r_p \\ \vdots \\ r_q \end{vmatrix} = \begin{vmatrix} r_1 \\ r_p \\ \vdots \\ r_q \end{vmatrix} + \lambda \begin{vmatrix} r_2 \\ r_p \\ \vdots \\ r_q \end{vmatrix} = D_k + \lambda N_k$$

其中 r_i 是分量取自 A 的第 i 行的 k 维行向量, N_k 也是 B 的一个 k 阶子式, 由于 $M_k - \lambda N_k = D_k \neq 0$, 故 M_k 和 N_k 不同时为零, 即 B 中存在 k 阶非零子式, 因而也有 $r(B) \geq k = r(A)$ 。

以上证明了: 如果 A 经 1 次初等行变换变成 B , 则有 $r(A) \leq r(B)$ 。由于初等行变换是可逆的, B 也可经 1 次初等行变换变成 A , 所以同样有 $r(B) \leq r(A)$ 。

因此有 $r(A) = r(B)$, 而既然作 1 次初等行变换, 矩阵的秩不改变, 当然作若干次初等行变换, 矩阵的秩仍然不会改变, 证毕。

定理 2.5 也可以表述为命题 2.15:

命题 2.15

设 A 为 $m \times n$ 矩阵, P 为 m 阶满秩方阵, Q 为 n 阶满秩方阵, 则有

$$r(PA) = r(A) \quad r(AQ) = r(A) \quad r(PAQ) = r(A)$$

即用满秩方阵 (可逆矩阵) 乘矩阵后, 矩阵的秩不改变。

该命题显然, 因为小节 2.4 中已经表明, 对一个矩阵施加行 (列) 变换等价于左 (右) 乘一个可逆矩阵。

定理 2.5 或者命题 2.15 表明了求矩阵的秩的一般性方法——将矩阵经过初等变换成为阶梯型矩阵, 其非零行的个数就是矩阵的秩。

2.5.3 矩阵的标准型和满秩分解

在小节 2.4 中我们提到了标准型这一概念, 现在给予详细定义:

定义 2.18 (秩标准型)

设 $r(A_{m \times n}) = r > 0$, 则必存在 m 阶可逆方阵 P 和 n 阶可逆方阵 Q , 使:

$$PAQ = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}_{m \times n} \quad (2.11)$$

其中 I_r 为 r 阶单位矩阵, 并称式 2.11 右端的矩阵为矩阵 A 的秩标准形。

特别地, 如果 $A_{m \times n}$ 的秩为 m , 则 A 的秩标准形为:

$$\begin{bmatrix} I_m & O_{m(n-m)} \end{bmatrix}$$

如果 $A_{m \times n}$ 的秩为 n , 则 A 的秩标准形为:

$$\begin{bmatrix} I_n \\ O_{(m-n)n} \end{bmatrix}$$

如果 A 为 n 阶满秩方阵, 则 A 的秩标准形显然为 n 阶单位矩阵 I_n 。

命题 2.16 (满秩分解)

设 $r(A_{m \times n}) = r$ 。则存在列满秩矩阵 $G_{m \times r}$ 和行满秩矩阵 $H_{r \times n}$, 使 $A = GH$, 其中 $r(G) = r(H) = r$ 。这被称为矩阵 A 的满秩分解。



证明 由命题 2.15 知存在可逆矩阵 P 和 Q , 使:

$$PAQ = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}_{m \times n} = \begin{bmatrix} I_r \\ O \end{bmatrix}_{m \times r} \begin{bmatrix} I_r & O \end{bmatrix}_{r \times n}$$

用 P^{-1} 左乘上式两端, 用 Q^{-1} 右乘上式两端, 得:

$$A = P^{-1} \begin{bmatrix} I_r \\ O \end{bmatrix} \begin{bmatrix} I_r & O \end{bmatrix} Q^{-1} = GH$$

其中 $G = P^{-1} \begin{bmatrix} I_r \\ O \end{bmatrix}$ 是 $m \times r$ 矩阵, 且显然有 $r(G) = r$; $H = \begin{bmatrix} I_r & O \end{bmatrix} Q^{-1}$ 是 $r \times n$ 矩阵, 且显然有 $r(H) = r$ 证毕。

不难证明该命题的逆命题也成立。

第二章 练习

(1) 设有 n 阶可逆方阵 A, B , k 为任意非零常数, 证明下列关于伴随矩阵的运算性质:

(a). A^* 可逆且 $(A^*)^{-1} = (A^{-1})^* = \frac{1}{\det(A)} A$

(b). $(kA)^* = k^{n-1} A^*$

(c). $(A^*)^* = [\det(A)]^{n-2} A$

(d). $(AB)^* = B^* A^*$

(2) 设 A 和 B 都是 n 阶方阵, 并且对于任意的 n 维列向量 x 总有 $Ax = Bx$, 证明 $A = B$.

(3) 设 n 维向量 $\alpha = \begin{bmatrix} a & 0 & \cdots & 0 & a \end{bmatrix}^T$ ($a < 0$). 已知矩阵 $A = I - \alpha\alpha^T$ 的逆矩阵为 $B = I + \frac{1}{a}\alpha\alpha^T$, 求常数 a .

(4) 设 3 阶矩阵 A 满足 $A^* = A^T$, 且 $a_{11} = a_{12} = a_{13} = a > 0$, 求 a 的值.

(5) 设 $A_{n \times n}$ 可逆, α, β 均为 n 维列向量, 且 $1 + \beta^T A^{-1} \alpha \neq 0$. 证明 $A + \alpha\beta^T$ 可逆, 且有:

$$(A + \alpha\beta^T)^{-1} = A^{-1} - \frac{A^{-1}\alpha\beta^T A^{-1}}{1 + \beta^T A^{-1}\alpha}$$

(6) 设 A, B 为 n 阶可逆方阵, A^*, B^* 分别为 A, B 的伴随矩阵, 分块阵 $C = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$, 求 C^*

(7) 设 4 阶矩阵 B 满足 $\left(\left(\frac{1}{2}A\right)^*\right)^{-1} BA^{-1} = 2AB + 12I$, 其中 $A = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -1 & 0 \end{bmatrix}$, 求 B

(8) 设有 n ($n > 1$) 维向量 $\alpha = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$, I 为 n 阶单位阵, 求 n 阶矩阵 $A = I - \frac{1}{n}\alpha\alpha^T$ 的秩.

(9) 设有 n 阶方阵 A 和 n 维列向量 β , p, q 为常数, 若 $\det(A) = a$, $\begin{vmatrix} A & \beta \\ \beta^T & p \end{vmatrix} = 0$, 求 $\begin{vmatrix} A & \beta \\ \beta^T & q \end{vmatrix}$

(10) 设 $A, B, A+B$ 均为 n 阶可逆矩阵, 证明:

(a). $A^{-1} + B^{-1}$ 可逆, 且 $(A^{-1} + B^{-1})^{-1} = A(A+B)^{-1}B$;

(b). $A(A+B)^{-1}B = B(A+B)^{-1}A$

(11) 设 A 为 $m \times n$ 阶矩阵, B 为 $n \times m$ 阶矩阵, 证明 $|I_m + AB| = |I_n + BA|$, 其中 I_k 为 k 阶单位阵.

第三章 几何向量及其应用

3.1 向量及其线性运算

本小节和 3.2 会将中学所学的向量知识给予一般性的叙述，不会给出过多的定义、定理、命题。

3.1.1 向量的概念

在现实世界中有一些量，例如长度、面积、质量和温度等，只要给定了测量单位，就可以用实数来表示它们，称这类量为**数量**或**标量**。还有一些量例如位移、力和速度等，它们不仅有大小，还有方向，称这类既有大小又有方向的量为**向量**或**矢量**。

在几何上往往用带箭头的线段——有向线段来表示向量，有向线段的长度表示向量的大小，其指向表示向量的方向。起点为 A，终点为 B 的向量记为 $\vec{AB}$ 。为了方便，也用小黑体字母如 $a, \alpha, x, \dots$ 来表示向量（手写时可写为 $\vec{a}, \vec{\alpha}, \vec{x}, \dots$ ）。（本讲义中大部分非黑体字母也可以表示向量）

向量的大小称为向量的长度（或**模**、**范数**），向量 $\vec{AB}$ 和 $\vec{a}$ 的长度分别记为 $\|\vec{AB}\|$ 和 $\|\vec{a}\|$ 。长度为 1 的向量称为**单位向量**，与 $\vec{a}$ 同方向的单位向量通常记为 a° 。长度为零的向量称为**零向量**，记为 $\vec{0}$ 。零向量的起点与终点重合，因而代表一个点，它的方向是不确定的，也可以说它的方向是任意的，可以根据需要来选取它的方向。

向量 a 和 b 相等就是说两个向量方向相同，大小相同； a 的负向量记为 $-a$ ，其和 a 大小相同，方向相反；向量 a 和 b 平行，就是说两个向量方向相同或相反，记为 $a \parallel b$ ；向量 a 和 b 垂直或正交，就是说两个向量方向成 90 度角，记为 $a \perp b$ 。

3.1.2 向量的线性运算

定义 3.1 (向量加法的三角形法则)

对于向量 $\vec{a}, \vec{b}$ ，从一点 O 作向量 $\vec{OA} = \vec{a}$ ，再从点 A 作向量 $\vec{AB} = \vec{b}$ ，则称向量 $\vec{OB}$ 为向量 $\vec{a}$ 与 $\vec{b}$ 的和，记为 $\vec{a} + \vec{b}$ ，即 $\vec{a} + \vec{b} = \vec{OB}$ 或 $\vec{OA} + \vec{AB} = \vec{OB}$ 。

向量的加法可以由图 3.1 表示：

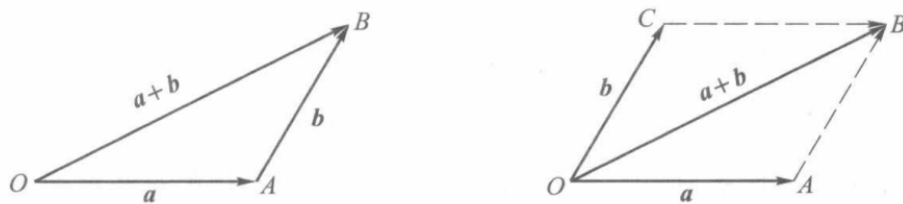


图 3.1: 向量的加法

向量的加法满足交换律，结合律；一个向量和零向量的和为其本身；一个向量和其负向量的和为**零向量**。

定义 3.2 (向量的数乘)

设 λ 是一个数，向量 a 与数 λ 的乘积 λa 规定为：

1. $\lambda > 0$ 时， λa 与 a 同向， $\|\lambda \vec{a}\| = \lambda \|\vec{a}\|$ ；
2. $\lambda = 0$ 时， $\lambda a = \vec{0}$ ；
3. $\lambda < 0$ 时， λa 与 a 反向， $\|\lambda a\| = |\lambda| \|a\|$ 。

向量的数乘满足交换律，结合律。

一个向量可以表示为其同方向的单位向量的数乘，该方向的单位向量也就表示为3.1

$$a^\circ = \frac{a}{\|a\|} \quad (3.1)$$

两个向量之间的关系还满足重要的三角不等式3.2

$$\|a + b\| \leq \|a\| + \|b\| \quad (3.2)$$

上述向量的加法和数乘统称为向量的线性运算。

3.1.3 向量共线和共面

定理 3.1 (向量共线定理)

两个向量 a 与 b 共线的充要条件是存在不全为零的常数 k_1 和 k_2 ，使得：

$$k_1 a + k_2 b = 0$$

向量的共面性就是说若干个向量在同一个平面上，有定理3.2：

定理 3.2 (向量共面定理)

三个向量 a, b, c 共面的充要条件是存在不全为零的常数 k_1, k_2 和 k_3 ，使得：

$$k_1 a + k_2 b + k_3 c = 0$$

给出一种重要定理3.3：

定理 3.3 (空间向量的线性表示)

设 e_1, e_2, e_3 是空间中不共面的 3 个向量，则空间中任一向量 a 都可由 e_1, e_2, e_3 惟一地表示为：

$$a = x e_1 + y e_2 + z e_3 \quad (3.3)$$

其中 x, y, z 为常数。

3.1.4 空间直角坐标系

空间直角坐标系是平面直角坐标系的推广。在空间取定一点 O ，过点 O 作互相垂直的有次序的 3 条数轴 Ox, Oy, Oz ，它们都以点 O 为原点，且有相同的长度单位，这样就构成了一个空间直角坐标系。坐标系有右手系和左手系之分。下图 (a) 和图 (b) 所示坐标系分别称为右手坐标系和左手坐标系，按照习惯通常取右手系。

通常，把与 Ox, Oy, Oz 轴的正向相同的单位向量分别记为 i, j, k ，并把这个直角坐标系记为 $\{O; i, j, k\}$ ，点 O 称为坐标原点， i, j, k 称为这个坐标系的基向量 (或基本单位向量，坐标向量)。轴 Ox, Oy, Oz 分别简称为 x 轴， y 轴和 z 轴，统称为坐标轴。3 个坐标轴中每两个都决定 1 个平面，由 x 轴和 y 轴所决定的平面称为 Oxy 面，由 y 轴和 z 轴所决定的平面称为 Oyz 面，由 x 轴和 z 轴所决定的平面称为 Ozx 面，统称为坐标面。

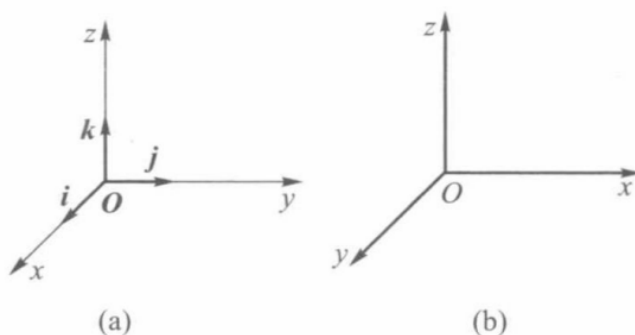


图 3.2: 空间直角坐标系

空间直角坐标系的三个坐标面将坐标系分为了八个部分，成为卦限，具体见图3.3

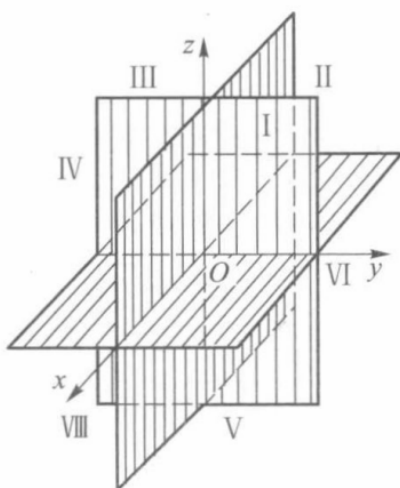


图 3.3: 空间直角坐标系的卦限

3.1.5 空间向量的坐标表示

根据定理3.3，空间直角坐标系中取三个正交基向量，则原点 O 到任意一点 P 的向量 $\vec{OP}$ 可以唯一表示为：

$$\vec{OP} = xi + yj + zk$$

因此，规定该点 P 在空间直角坐标系中的坐标为 (x, y, z) 。

起点为 $A(x_1, y_1, z_1)$ 且终点为 $B(x_2, y_2, z_2)$ 的向量 $\vec{AB} = \vec{OB} - \vec{OA} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ 。根据勾股定理不难计算任意一空间向量 $\vec{AB}$ 的长度 (模)，以上述 $\vec{AB} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ 为例，有：

$$\|\vec{AB}\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (3.4)$$

为了利用向量的坐标来刻画向量的方向，我们引入两个向量的夹角概念。对于两个非零向量 a, b ，任取空间一个定点 M ，作向量 $\vec{MA} = a$ ， $\vec{MB} = b$ ，则规定不超过 π 的 $\angle AMB$ 为 a 与 b 的夹角，即 $0 \leq \angle AMB \leq \pi$ 。如果 a 与 b 中有一个为零向量，则规定 a 与 b 的夹角可以是 0 到 π 之间的任意值。如果 a 与 b 的夹角为 $\frac{\pi}{2}$ ，就是 a 与 b 垂直。 a 与 b 的夹角可简记为 (a, b) 。

为进一步标定一个向量在空间直角坐标系中的方向，引入方向余弦的概念：

定义 3.3 (方向余弦)

在直角坐标系 $O; i, j, k$ 中，一个向量 $a = xi + yj + zk$ 分别与 3 个基向量 i, j, k 的夹角 (即分别与 x 轴、 y 轴、 z 轴正向之间的夹角) α, β, γ 称为向量 a 的方向角，方向角的余弦 $\cos \alpha, \cos \beta, \cos \gamma$ 称为向量 a 的方向余弦：

$$\begin{cases} \cos \alpha = \frac{x}{\|a\|} = \frac{x}{\sqrt{x^2 + y^2 + z^2}} \\ \cos \beta = \frac{y}{\|a\|} = \frac{y}{\sqrt{x^2 + y^2 + z^2}} \\ \cos \gamma = \frac{z}{\|a\|} = \frac{z}{\sqrt{x^2 + y^2 + z^2}} \end{cases}$$

三个方向余弦的平方和显然为 1；向量 a 方向的单位向量可以表示为：

$$a^\circ = (\cos \alpha, \cos \beta, \cos \gamma) \quad (3.5)$$

3.1.6 空间向量坐标的性质

设向量 $a = x_1i + y_1j + z_1k$, $b = x_2i + y_2j + z_2k$, 则由向量线性运算的运算规律易得:

$$a \pm b = (x_1 \pm x_2)i + (y_1 \pm y_2)j + (z_1 \pm z_2)k \quad (3.6)$$

$$\lambda a = (\lambda x_1)i + (\lambda y_1)j + (\lambda z_1)k \quad (3.7)$$

结合向量的共线共面定理, 有定理:

定理 3.4 (向量共线共面定理的坐标表示)

设向量 $a = x_1i + y_1j + z_1k$, $b = x_2i + y_2j + z_2k$, 则它们共线的充要条件是它们的对应坐标分量成比例;
 设还有向量 $c = x_3i + y_3j + z_3k$, 则它们共面的充要条件是它们坐标组成的行列式值为 0, 即:

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = 0$$



介绍定比分点性质: 设空间中两个不重合的已知点 $A(x_1, y_1, z_1)$ 和 $B(x_2, y_2, z_2)$, $C(x, y, z)$ 是直线 AB 上的一动点且 $\vec{AC} = \lambda \vec{CB} (\lambda \neq -1)$, 则 C 的坐标可以表示为:

$$(x, y, z) = \left(\frac{x_1 + \lambda x_2}{1 + \lambda}, \frac{y_1 + \lambda y_2}{1 + \lambda}, \frac{z_1 + \lambda z_2}{1 + \lambda} \right)$$

3.1.7 正交射影

定义 3.4 (正交射影向量和正交射影)

设向量 a 与 b 的夹角为 θ , 定义向量

$$\text{Proj}_a b = \|b\| \cos \theta a^\circ$$

为向量 b 在向量 a 上的正交射影向量, 简称为射影向量; 定义数值

$$(b)_a = \|b\| \cos \theta$$

为 b 在 a 上的正交射影 (或代数射影), 简称为射影。



笔记 一个空间点的坐标 (x, y, z) 中的每个分量都相当于其在对应基向量方向的正交射影。

令 $v = b - \text{Proj}_a b$, 则 $b = \text{Proj}_a b + v$ 表示为 b 的一个正交分解。射影满足数乘提取和射影分配律。

3.2 数量积、向量积和混合积

3.2.1 数量积

中学已经学过两个向量的点乘。设两个夹角为 θ 的向量 a 和 b ，有：

$$a \cdot b = \|a\| \|b\| \cos \theta \quad (3.8)$$

根据 3.1 节射影的定义，可以改写式 3.8 为：

$$a \cdot b = \|a\| (b)_a = \|b\| (a)_b \quad (3.9)$$

一般的，点乘将两个向量的运算转化为了一个数，因此也被称为数量积。其满足以下性质：

命题 3.1 (数量积的运算性质)

1. $a \cdot b = b \cdot a$
2. $(a+b) \cdot c = a \cdot c + b \cdot c$
3. $(ka) \cdot b = k(a \cdot b)$
4. $a \cdot a \geq 0$ 且 $a \cdot a = 0$ 时当且仅当 $a = \vec{0}$

设两个向量用坐标表示，则数量积的坐标形式如下：

$$(a_x, a_y, a_z) \cdot (b_x, b_y, b_z) = a_x b_x + a_y b_y + a_z b_z \quad (3.10)$$

借助式 3.8, 3.9, 3.10 可以利用坐标求以下量：

1. 求向量的模，设向量 $a = (a_x, a_y, a_z)$ ，则：

$$\|a\| = \sqrt{\|a\|^2} = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

2. 求向量的夹角，设向量 $a = (a_x, a_y, a_z)$ 和向量 $b = (b_x, b_y, b_z)$ ，夹角为 θ ，则：

$$\cos \theta = \frac{a \cdot b}{\|a\| \|b\|} = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2} \sqrt{b_x^2 + b_y^2 + b_z^2}} = a^\circ b^\circ$$

当两个向量垂直时，它们的数量积为 0；

3. 求射影，当 $a \neq 0$ 时，可得 b 在 a 上的射影为：

$$(b)_a = \frac{a \cdot b}{\|a\|} \quad \text{或} \quad b \cdot a^\circ$$

从而可得 b 在 a 上的射影向量为：

$$\text{Proj}_a b = \frac{a \cdot b}{\|a\|} a^\circ \quad \text{或} \quad (b \cdot a^\circ) a^\circ$$

3.2.2 向量积

两个向量的乘积还可以是一个新的向量，这种乘法叫做向量积：

定义 3.5 (向量积)

规定两个向量 a 与 b 的向量积是一个向量，记为 $a \times b$ ， $a \times b$ 的方向与 a, b 都垂直，且使 $a, b, a \times b$ 符合右手法则 (如图 3.4 所示)， $a \times b$ 的模等于以 a, b 为邻边的平行四边形的面积，即 $\|a \times b\| = \|a\| \|b\| \sin(a, b)$

向量积这个概念看上去陌生，但中学阶段的电磁物理等大量用到向量积。向量积也被称为外积或叉积，其中的 $\times$ 和点乘的 $\cdot$ 不能省去。其满足以下性质：

命题 3.2 (向量积的运算性质)

1. $a \times b = -b \times a$
2. $(a + b) \times c = a \times c + b \times c$
3. $(ka) \times b = k(a \times b)$

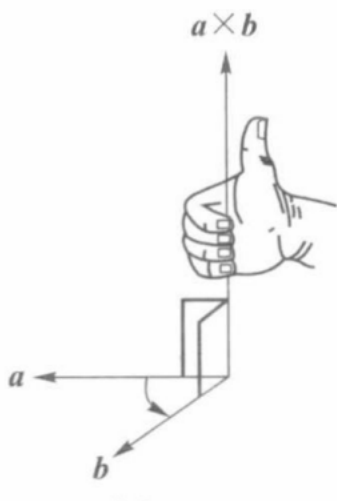


图 3.4: 右手法则

经过推导不难得到向量积的坐标形式:

$$a \times b = \begin{vmatrix} i & j & k \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} \quad (3.11)$$

根据向量积的定义, 当两向量共线时, $\sin \theta = 0$, 则有:

定理 3.5 (向量共线的向量积判定)

向量 a, b 共线当且仅当 $a \times b = \vec{0}$

3.2.3 混合积**定义 3.6 (混合积)**

规定 3 个向量 a, b, c 的混合积是一个数, 它等于 a 与 b 先作向量积, 然后再用 $a \times b$ 与 c 作数量积所得到的数, 记成 $[a \ b \ c]$, 即 $[a \ b \ c] = (a \times b) \cdot c$.

混合积有明显的几何意义。当 a, b, c 构成右手系时, 向量 c 与 $a \times b$ 的夹角为锐角, 此时, 考察以 a, b, c 为相邻棱的平行六面体 Ω (图), 由向量积的定义知 $\|a \times b\|$ 为 Ω 的底面面积, 由于 $a \times b$ 与 Ω 的底面垂直, 且 $a \times b$ 与 c 的夹角为锐角, 故 c 在 $a \times b$ 上的射影 $(c)_{a \times b}$ (此时为正值) 是 Ω 的高 h , 故该平行六面体的体积为

$$V = \|a \times b\| (c)_{a \times b}$$

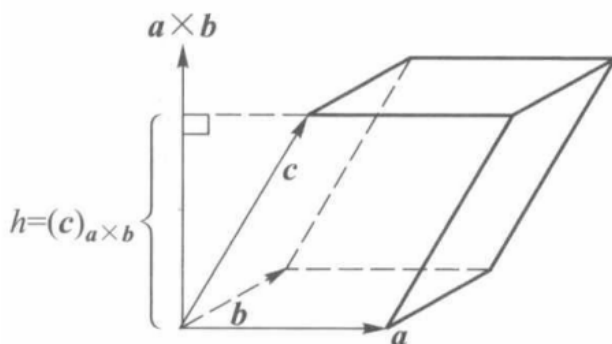


图 3.5: 混合积的几何意义

混合积满足以下运算法则:

命题 3.3 (混合积的运算性质)

1. $(a \times b) \cdot c = (b \times c) \cdot a = (c \times a) \cdot b$, 即 $[a \ b \ c] = [b \ c \ a] = [c \ a \ b]$;
2. 互换混合积中任意两个向量的位置, 则混合积变号。例如:

$$[b \ a \ c] = -[a \ b \ c]$$

根据向量积和数量积的坐标表示, 可以推导出混合积的坐标表示:

$$[a \ b \ c] = (a \times b) \cdot c = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \quad (3.12)$$

结合定理3.4, 有:

定理 3.6 (向量共面的混合积判定)

向量 a, b, c 共面当且仅当混合积 $[a \ b \ c] = 0$ 。

练习 3.1 求以 $A(1, 1, 1), B(2, 0, 1), C(0, 0, 1), D(1, 3, 2)$ 为顶点的四面体的体积 V 。

解 V 等于以 $\vec{AB} = (1, -1, 0), \vec{AC} = (-1, -1, 0), \vec{AD} = (0, 2, 1)$ 为相邻棱的平行六面体体积 Ω 的 $\frac{1}{6}$ 倍, 而由混合积的几何意义知 Ω 等于混合积 $[\vec{AB} \ \vec{AC} \ \vec{AD}]$ 的绝对值。由于:

$$[\vec{AB} \ \vec{AC} \ \vec{AD}] = \begin{vmatrix} 1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 2 & 1 \end{vmatrix} = -2$$

$$\text{故 } \Omega = |-2| = 2, V = \frac{1}{6}\Omega = \frac{1}{3}。$$

3.3 平面和空间直线

3.3.1 平面的方程

对于给定的平面 π , 假设有非零向量 $n = (a, b, c)$ 和平面垂直, 即对于平面上任意向量 s , 都有:

$$n \cdot s = 0$$

那么称 n 为平面 π 的一个法向量。

定理 3.7 (平面的点法式方程)

设平面 π 过点 $P(x_0, y_0, z_0)$, 且法向量 $n = (A, B, C)$, 那么平面方程可以表示为:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \quad (3.13)$$

把方程展开, 可以得到一个三元一次方程 $Ax + By + Cz + D = 0$, 这称为平面的一般式方程。几种特殊情况:

1. $D = 0$, 平面通过坐标原点;
2. $A = 0$, $\begin{cases} D = 0, & \text{平面过} x \text{轴;} \\ D \neq 0, & \text{平面平行于} x \text{轴;} \end{cases}$ 类似地可讨论 $B = 0, C = 0$ 情形;
3. $A = B = 0$, 平面平行于 xOy 坐标面 (即垂直于 z 轴), 类似可讨论 $A = C = 0, B = C = 0$ 情形。

定理 3.8 (平面的截距式方程)

设平面 π 在 x, y, z 三轴上分别有截距 $OP = a, OQ = b, OR = c$ (其中 a, b, c 均为非零常数), 则平面 π 的截距式方程可以表示为:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \quad (3.14)$$

证明 设平面为 $Ax + By + Cz + D = 0$, 由已知, 平面过点 $(a, 0, 0), (0, b, 0), (0, 0, c)$

$$\text{有 } \begin{cases} aA + D = 0, \\ bB + D = 0, \\ cC + D = 0 \end{cases} \Rightarrow A = -\frac{D}{a}, B = -\frac{D}{b}, C = -\frac{D}{c}, \text{ 代入一般式方程即证。}$$

定理 3.9 (平面的参数方程)

设平面 π 过点 $P_0(x_0, y_0, z_0)$, 且已知 π 上两个不共线的向量 $a = (L_1, M_1, N_1), b = (L_2, M_2, N_2)$ 设平面上任一点为 $P(x, y, z), r = OP, r_0 = OP_0$ 。则存在唯一的一组实数 s, t , 使得: $P_0P = sa + tb$ 或 $r = r_0 + sa + tb$ 。由此有平面的参数方程形式:

$$\begin{cases} x = x_0 + sL_1 + tL_2 \\ y = y_0 + sM_1 + tM_2 \\ z = z_0 + sN_1 + tN_2 \end{cases} \quad (3.15)$$

3.3.2 平面与平面的位置关系

设有两个平面:

$$\pi_1: a_1x + b_1y + c_1z + d_1 = 0 \quad \pi_2: a_2x + b_2y + c_2z + d_2 = 0$$

把这两个平面的法向量 $n_1 = (a_1, b_1, c_1)$ 与 $n_2 = (a_2, b_2, c_2)$ 的夹角 θ 称为这两个平面的夹角 (通常取锐角), 有:

$$\cos \theta = \frac{|n_1 \cdot n_2|}{\|n_1\| \|n_2\|} = \frac{|a_1a_2 + b_1b_2 + c_1c_2|}{\sqrt{a_1^2 + b_1^2 + c_1^2} \sqrt{a_2^2 + b_2^2 + c_2^2}}$$

根据向量的空间关系, 有:

命题 3.4 (平面之间的位置关系判断)

1. π_1 与 π_2 相交 $\Leftrightarrow n_1$ 与 n_2 不平行 $\Leftrightarrow a_1 : b_1 : c_1 \neq a_2 : b_2 : c_2$;
2. π_1 与 π_2 平行而不重合 $\Leftrightarrow \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} \neq \frac{d_1}{d_2}$;
3. π_1 与 π_2 重合 $\Leftrightarrow \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2}$

练习 3.2 求通过点 $P_1(-1, 0, 2)$, $P_2(1, 1, 1)$, 且与平面 $x + y + z + 1 = 0$ 垂直的平面的方程。

解 设所求平面的法向量为 n , 则因 $\vec{P_1P_2}$ 在平面内, 故有 $n \perp \vec{P_1P_2}$. 又由平面垂直的条件, 有 $n \perp (1, 1, 1)$, 于是可取:

$$n = \vec{P_1P_2} \times (1, 1, 1) = \begin{vmatrix} i & j & k \\ 2 & 1 & -1 \\ 1 & 1 & 1 \end{vmatrix} = (2, -3, 1)$$

所以, 所求平面的方程为 $2(x+1) - 3y + (z-2) = 0$ 或表示为 $2x - 3y + z = 0$ 。

3.3.3 直线的方程

定理 3.10 (直线的对称式方程)

设直线上一点 $M_0(x_0, y_0, z_0)$, 向量 $s = (l, m, n)$ 。若 $M(x, y, z)$ 为直线 L 上任一点, 都有 $\vec{M_0M} \parallel s$, 记 s 为直线 L 的方向向量, 且有直线的对称式方程:

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n} \quad (3.16)$$

注 l, m, n 为 0 时代表方向向量 s 在该方向的分量为 0, 这只是一特殊记法。

令 3.16 等于 t , 那么可以得到直线的参数式方程:

$$\begin{cases} x = x_0 + lt, \\ y = y_0 + mt, \\ z = z_0 + nt. \end{cases} \quad (-\infty < t < +\infty) \quad (3.17)$$

如果两平面:

$$\pi_1 : A_1x + B_1y + C_1z + D_1 = 0 \quad \pi_2 : A_2x + B_2y + C_2z + D_2 = 0$$

不平行, 则其交线是一条直线:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases} \quad (3.18)$$

这称为空间直线的一般式方程, 其方向向量为:

$$s = n_1 \times n_2 = (A_1, B_1, C_1) \times (A_2, B_2, C_2)$$

练习 3.3 化直线的一般式方程 $\begin{cases} 2x + 5y - 3z + 4 = 0 \\ -x - 3y + z - 1 = 0 \end{cases}$ 为对称式和参数式。

解 显然方程组中 x 与 y 的系数构成 2 阶非零行列式, 因此, 可通过令 z 取任意常数, 来求方程组的一个解。例如令 $z = 0$, 则有:

$$\begin{cases} 2x + 5y = -4 \\ -x - 3y = 1 \end{cases}$$

解之得 $x = -7, y = 2$, 于是 $P_0(-7, 2, 0)$ 便是直线上的一个点。

由于两平面法向量的向量积:

$$n_1 \times n_2 = \begin{vmatrix} i & j & k \\ 2 & 5 & -3 \\ -1 & -3 & 1 \end{vmatrix} = (-4, 1, -1) \parallel (4, -1, 1)$$

与直线的方向向量共线, 故可取直线的方向向量为 $s = (4, -1, 1)$ 。因此, 直线的对称式和参数式方程分别是:

$$\frac{x+7}{4} = \frac{y-2}{-1} = z \quad \text{与} \quad \begin{cases} x = -7 + 4t \\ y = 2 - t \\ z = t \end{cases}$$

3.3.4 直线和直线的位置关系

设有两条直线:

$$L_1: \frac{x-x_1}{l_1} = \frac{y-y_1}{m_1} = \frac{z-z_1}{n_1} \quad L_2: \frac{x-x_2}{l_2} = \frac{y-y_2}{m_2} = \frac{z-z_2}{n_2}.$$

把这两条直线的方向向量 s_1 与 s_2 的夹角 θ 称为这两条直线的夹角 (一般取锐角), 于是:

$$\cos \theta = \frac{|s_1 \cdot s_2|}{\|s_1\| \|s_2\|} = \frac{|l_1 l_2 + m_1 m_2 + n_1 n_2|}{\sqrt{l_1^2 + m_1^2 + n_1^2} \sqrt{l_2^2 + m_2^2 + n_2^2}}$$

从几何上看, 两条直线 L_1 与 L_2 的位置关系只有 4 种可能, 即异面、相交于一点、平行而不重合或重合。注意点 $P_1(x_1, y_1, z_1)$ 在 L_1 上, 点 $P_2(x_2, y_2, z_2)$ 在 L_2 上, 则可将上述 4 种情形的充要条件总结如下:

命题 3.5 (直线之间的位置关系判断)

1. L_1 与 L_2 异面 $\Leftrightarrow$ 3 个向量 $\vec{P_1 P_2}, s_1, s_2$ 不共面;
2. L_1 与 L_2 相交于一点 $\Leftrightarrow$ 3 个向量 $\vec{P_1 P_2}, s_1, s_2$ 共面, 且 $s_1 \nparallel s_2$;
3. L_1 与 L_2 平行而不重合 $\Leftrightarrow s_1 \parallel s_2 \nparallel \vec{P_1 P_2}$;
4. L_1 与 L_2 重合 $\Leftrightarrow s_1 \parallel s_2 \parallel \vec{P_1 P_2}$.

特别地, L_1 与 L_2 共面 $\Leftrightarrow$ 3 个向量 $\vec{P_1 P_2}, s_1, s_2$ 共面 $\Leftrightarrow \begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0$

 **练习 3.4** 设有两条直线:

$$L_1: \frac{x-2}{-1} = \frac{y+2}{2} = \frac{z+1}{3} \quad L_2: \frac{x-5}{2} = \frac{y-2}{1} = \frac{z-4}{1},$$

试讨论 L_1 与 L_2 的位置关系。若相交, 求出它们的交点; 若共面, 求出它们所在平面的方程。

解 L_1 与 L_2 的方向向量显然不平行, 现在来看它们是否相交。将 L_1 与 L_2 都写成参数方程:

$$L_1: x = 2 - t, y = -2 + 2t, z = -1 + 3t \quad L_2: x = 5 + 2r, y = 2 + r, z = 4 + r$$

若它们相交, 则在交点处应有:

$$2 - t = 5 + 2r, \quad -2 + 2t = 2 + r, \quad -1 + 3t = 4 + r$$

解得 $t = 1, r = -2$, 于是得 L_1 与 L_2 相交于一点 $(1, 0, 2)$, 故 L_1 与 L_2 共面。现在来求由 L_1 与 L_2 所确定的平面 π 的方程, 由于:

$$s_1 \times s_2 = \begin{vmatrix} i & j & k \\ -1 & 2 & 3 \\ 2 & 1 & 1 \end{vmatrix} = (-1, 7, -5)$$

故可取平面 π 的法向量为 $n = (1, -7, 5)$, 于是得平面 π 的方程为:

$$x - 2 - 7(y + 2) + 5(z + 1) = 0 \quad \text{或表示为} \quad x - 7y + 5z - 11 = 0$$

3.3.5 直线与平面的位置关系

直线和平面存在三种关系：相交、平行、重合 (直线在平面上)。一条直线 L 和平面 π 相交时，它们的夹角 φ 定义为直线 L 在平面 π 上投影直线和 L 的夹角 (一般取锐角)。当 $\varphi = \frac{\pi}{2}$ 时，直线和平面垂直。

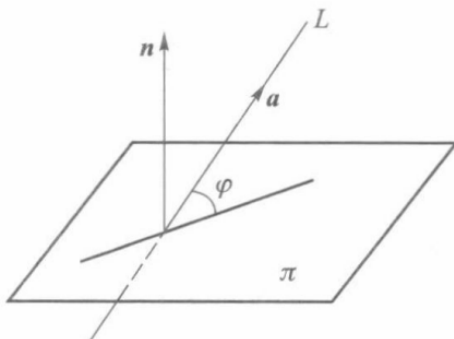


图 3.6: 直线和平面的夹角

设有平面 $\pi: Ax + By + Cz + D = 0$ 和直线 $L: \frac{x-x_0}{k} = \frac{y-y_0}{l} = \frac{z-z_0}{m}$ 。由于直线 L 的方向向量 $a = (k, l, m)$ 与平面 π 的法向量 n 的夹角为 $\frac{\pi}{2} - \varphi$ 或 $\frac{\pi}{2} + \varphi$ ，故夹角 φ 计算为：

$$\sin \varphi = \left| \cos\left(\frac{\pi}{2} \pm \varphi\right) \right| = |\cos(a, n)| = \frac{|Ak + Bl + Cm|}{\sqrt{A^2 + B^2 + C^2} \sqrt{k^2 + l^2 + m^2}}$$

命题 3.6 (直线和平面的位置关系判断)

1. 当 $Ak + Bl + Cm \neq 0$ 时, L 与 π 交于一点;
2. 当 $Ak + Bl + Cm = 0$ 且 $Ax_0 + By_0 + Cz_0 + D \neq 0$ 时, L 与 π 平行 (但 L 不在 π 上);
3. 当 $Ak + Bl + Cm = 0$ 且 $Ax_0 + By_0 + Cz_0 + D = 0$ 时, L 在平面 π 上。

3.3.6 点到平面的距离

给定平面 $\pi: Ax + By + Cz + D = 0$, 及 π 以外一点 $P_1(x_1, y_1, z_1)$, 从 P_1 作平面 π 的垂线, 设垂足为 P'_1 , 所谓 P_1 到 π 的距离 d , 是指 P_1 与 P'_1 两点之间的距离。在 π 上任取一点 $P_0(x_0, y_0, z_0)$, 记 π 的法向量 $n = (A, B, C)$, 则 d 显然等于向量 $\vec{P_0P_1}$ 在 n 上的射影的绝对值, 于是由计算射影的公式可得:

$$d = |\vec{P_0P_1} \cdot n^\circ| = \frac{|A(x_1 - x_0) + B(y_1 - y_0) + C(z_1 - z_0)|}{\sqrt{A^2 + B^2 + C^2}} \quad (3.19)$$

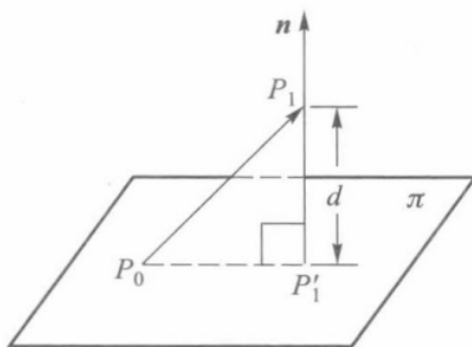


图 3.7: 点到平面的距离

3.3.7 点到直线的距离

给定直线 $L: \frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$, 及直线 L 外一点 $P_1(x_1, y_1, z_1)$, 设 P'_1 是 P_1 在直线 L 上的垂足点, 所谓点 P_1 到直线 L 的距离 d , 是指 P_1 与 P'_1 两点之间的距离。为求 d , 在 L 上取一点 $P_0(x_0, y_0, z_0)$, 记 L 的方向向量为 $a = (l, m, n)$, 则有:

$$d = \|\vec{P_0P_1}\| \cos \theta = \|\vec{P_0P_1}\| \sin(\angle(\vec{P_0P_1}, a))$$

于是由两个向量的向量积的模的定义, 得点到直线的距离公式:

$$d = \|\vec{P_0P_1} \times a^\circ\| = \frac{\|\vec{P_0P_1} \times a\|}{\|a\|} \quad (3.20)$$

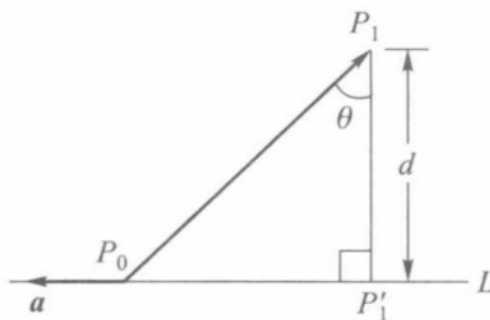


图 3.8: 点到直线的距离

练习 3.5 求点 $P_1(1, 0, -1)$ 到直线 $L: \frac{x-1}{2} = \frac{y+2}{0} = \frac{z-4}{1}$ 的距离 d 。

解 利用公式 3.20 来求 d 。在 L 上取一点 $P_0(1, -2, 4)$, L 的方向向量为 $a = (2, 0, 1)$, 于是有:

$$\vec{P_0P_1} \times a = (0, 2, -5) \times (2, 0, 1) = \begin{vmatrix} i & j & k \\ 0 & 2 & -5 \\ 2 & 0 & 1 \end{vmatrix} = (2, -10, -4)$$

从而有:

$$\|\vec{P_0P_1} \times a\| = \sqrt{2^2 + 10^2 + (-4)^2} = 2\sqrt{30}$$

$$\text{又 } \|a\| = \sqrt{5}, \text{ 所以 } d = \frac{2\sqrt{30}}{\sqrt{5}} = 2\sqrt{6}.$$

3.3.8 异面直线的距离

设异面直线 L_1, L_2 的方向向量分别为 a_1, a_2 。 P_1, P_2 分别为 L_1, L_2 上的点。所谓异面直线 L_1 与 L_2 的距离, 是指两直线上的点之间的最短距离, 因而是指它们的公垂线段的长。

我们来看 L_1 与 L_2 的公垂线, 过 L_2 可作惟一一个平面 π_2 与 $a_1 \times a_2$ 垂直, 因此向量 $a_1 \times a_2$ 可作为平面 π_2 的法向量。过 L_1 可作惟一一个平面 π_1 与 $a_1 \times a_2$ 平行。设 π_1 与 L_2 交于点 B , 过点 B 作平面 π_2 的垂线 (它必在平面 π_1 内), 设它交 L_1 于点 A , 则线段 BA 就是所求公垂线段。

公垂线段的长 $\|\vec{AB}\|$ 正是向量 $\vec{P_1P_2}$ 在 $a_1 \times a_2$ 上的射影的绝对值, 于是由计算射影的公式, 即得两条异面直线的距离公式:

$$d = \left| \vec{P_1P_2} \cdot \frac{a_1 \times a_2}{\|a_1 \times a_2\|} \right| = \frac{|\begin{vmatrix} a_1 & a_2 & \vec{P_1P_2} \end{vmatrix}|}{\|a_1 \times a_2\|} \quad (3.21)$$

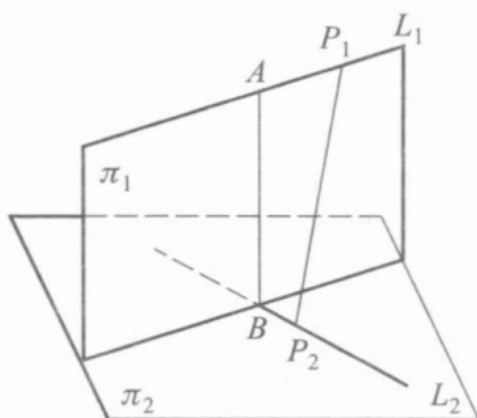


图 3.9: 异面直线的距离

练习 3.6 求下列两条异面直线 L_1 与 L_2 的距离 d :

$$L_1: \frac{x-5}{-4} = \frac{y-1}{1} = \frac{z-2}{1}, \quad L_2: \frac{x}{2} = \frac{y}{2} = \frac{z-8}{-3}$$

解 利用公式 3.21 来求 d 。现在有:

$$a_1 = (-4, 1, 1), \quad a_2 = (2, 2, -3), \quad P_1(5, 1, 2), \quad P_2(0, 0, 8)$$

由于:

$$[a_1 \quad a_2 \quad \vec{P_1P_2}] = \begin{vmatrix} -4 & 1 & 1 \\ 2 & 2 & -3 \\ -5 & -1 & 6 \end{vmatrix} = -25$$

$$a_1 \times a_2 = \begin{vmatrix} i & j & k \\ -4 & 1 & 1 \\ 2 & 2 & -3 \end{vmatrix} = (-5, -10, -10) = -5(1, 2, 2)$$

$$\|a_1 \times a_2\| = 5\|(1, 2, 2)\| = 15$$

代入公式 3.21 即得:

$$d = \frac{25}{15} = \frac{5}{3}$$

看到这，恭喜你已完成期中内容复习～



第三章 练习

- (1) 求过点 $M(-1, 0, 4)$ 且平行于平面 $\pi: 3x - 4y + z - 10 = 0$, 又与直线 $L: x + 1 = y - 3 = \frac{z}{2}$ 相交的直线的方程。
- (2) 已知点 $A(1, 0, 0)$ 及 $B(0, 2, 1)$, 试在 z 轴上求一点 C , 使 $\triangle ABC$ 面积最小。
- (3) 设直线 L 在平面 $\pi: 2x + 3y + 4z = 9$ 上且过点 $(1, 1, 1)$, 若 L 与 xOy 平面有最大交角, 求直线 l 的方程。
- (4) 已知 $\|a\| = \sqrt{3}, \|b\| = 1, (a, b) = \frac{\pi}{6}$, 求 $(a + b, a - b)$ 。
- (5) 求过点 $M_1(2, -3, -1)$ 与 x 轴的夹角为 $\frac{\pi}{3}$, 且与直线 $\frac{x-7}{1} = \frac{y-3}{2} = \frac{z-5}{2}$ 相交的直线。
- (6) 设一平面垂直于平面 $z = 0$, 并通过从点 $(1, -1, 1)$ 到直线 $\begin{cases} y - z + 1 = 0 \\ x = 0 \end{cases}$ 的垂线, 求此平面的方程。

第四章 n 维向量与线性方程组

4.1 消元法

4.1.1 线性方程组的基本概念

本章来讨论更加一般性的线性方程组, 考虑一个 n 元 m 个方程的线性方程组:

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

根据矩阵知识, 方程组可以表示为 $Ax = b$, 因此一个线性方程组的有效信息可以表示为一个矩阵:

$$\overline{A} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right] = [A | b]$$

此方程记为线性方程组 $Ax = b$ 的增广矩阵。线性方程组的一些基本概念由定义4.1给出:

定义 4.1 (线性方程组的一些基本概念)

- 方程组的解: $x_1 = c_1, x_2 = c_2, \cdots, x_n = c_n$ 使得每个方程两端都相等;
- 解向量: n 维列向量 $x_0 = (c_1, c_2, \dots, c_n)^T$;
- 解集合: 方程组的全部解所组成的集合; 解方程组就是要找出它的全部解, 即找出它的解集合;
- 一般解 (通解): 当方程组的解不唯一时, 称其全部解的表达式为方程组的一般解或通解;
- 同解 (等价): 如果两个线性方程组有相同的解集合, 称它们是同解的。

4.1.2 消元法

求解线性方程组通常采用消元法, 中学我们学习过消元步骤就是用某个式子的倍数加到另一个式子上, 以消去某一个变量。对于一个增广矩阵形式的线性方程组, 我们可以利用矩阵的初等变换来解决。下面的练习给出了一个消元的示例:

 **练习 4.1** 求解线性方程组:

$$\begin{cases} x_1 + x_2 + 2x_3 = 9 \\ 2x_1 + 2x_2 - 3x_3 = 11 \\ 3x_1 + 6x_2 - 6x_3 = 0 \\ x_1 + 3x_2 - 5x_3 = -8 \end{cases}$$

解 写出方程组的增广矩阵:

$$\overline{A} = \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 2 & -3 & 11 \\ 3 & 6 & -6 & 0 \\ 1 & 3 & -5 & -8 \end{array} \right]$$

利用第一行消去下属行 x_1 有：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 0 & -7 & -7 \\ 0 & 3 & -12 & -27 \\ 0 & 2 & -7 & -17 \end{array} \right]$$

把每一行的系数提取公因数，以简便运算：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -4 & -9 \\ 0 & 2 & -7 & -17 \end{array} \right]$$

交换第二行和第三行：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -4 & -9 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & -7 & -17 \end{array} \right]$$

利用第二行消去下属行 x_2 有：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -4 & -9 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

利用第三行消去下属行 x_3 有：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -4 & -9 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

此时，利用第三行消去上属行 x_3 ，再利用第二行消去上属行 x_2 ，得到左侧为行最简形的矩阵：

$$\bar{A} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 12 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

这个矩阵变换的过程实际上就对应着初始线性方程组经过相同操作得到的最终消元结果：

$$\begin{cases} x_1 & = 12 \\ & x_2 & = -5 \\ & & x_3 & = 1 \end{cases}$$

 **笔记** 从上述例题看出，消元法的思路就是将 $\bar{A}$ 通过初等变换变成行最简形。

有些时候，消元过程中会出现 $0x_k = 0$ 的情况，此时 x_k 可以取任意值；而同时其他变量 x_i 可能还需要 x_k 去表示。定义可以任意取值的未知量为**自由未知量**，其他的未知量称为**约束未知量**。因此，方程组的通解是由约束未知量和自由未知量共同表达的。举例：下面两个解是同一个解的不同表述形式：

$$\begin{cases} x_1 = 1 - 4c \\ x_2 = -5 + c \\ x_3 = c \end{cases} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ 0 \end{bmatrix} + c \begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}$$

 **练习 4.2** 求解线性方程组：

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_2 + 2x_3 = 1 \\ x_1 + 3x_2 + 5x_3 = 3 \end{cases}$$

解 对方程组的增广矩阵施行初等行变换：

$$\bar{A} = \left[A \mid b \right] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 3 & 5 & 3 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right],$$

与上式右端矩阵对应的方程组为：

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_2 + 2x_3 = 1, \\ 0 = 1 \end{cases}$$

该方程显然无解。

线性方程组并不是都有解，练习4.2展现的方程组就没有解。这种方程组也称为矛盾方程组或不相容方程组。

4.1.3 线性方程组的解的判定

对于一般的线性方程组 $Ax = b$ ，定理4.1阐述了其解的判定条件。

定理 4.1 (线性方程组的解的判定定理)

对于 n 元线性方程组 $Ax = b$ ，有：

1. 若 $r(A) = r$ ，而 $r(\bar{A}) = r + 1$ ，则方程组无解；
2. 若 $r(A) = r(\bar{A}) = n$ ，则方程组有惟一解；
3. 若 $r(A) = r(\bar{A}) = r < n$ ，则方程组有无穷多解。

证明 对于 n 元线性方程组：

$$Ax = b$$

设其系数矩阵 A 的秩为 r 。为以下讨论方便计，不妨设由 A 的前 r 列所构成的子矩阵中有一个 r 阶子式不等于零（此 r 阶非零子式对应的未知量为 $x_1, x_2, \dots, x_r$ ），则通过初等行变换，可将方程组 $Ax = b$ 的增广矩阵 $\bar{A}$ 化为简化行阶梯形矩阵：

$$\bar{A} = [A \mid b] \rightarrow \left[\begin{array}{cccc|ccc|c} 1 & 0 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} & d_1 \\ 0 & 1 & 0 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array} \right] \quad (4.1)$$

去掉式中那些 $0 = 0$ 的恒等式之后，方程组的最后一个方程是 $0 = d_{r+1}$ ，其中 d_{r+1} 可能为零（此时有 $r(A) = r(\bar{A}) = r$ ）；也可能不为零（此时有 $r(A) = r, r(\bar{A}) = r + 1$ ）。下面来说明，方程组是否有解取决于 d_{r+1} 是否为零，即 $r(A)$ 与 $r(\bar{A})$ 是否相等。

- 当 $d_{r+1} \neq 0$ 时，方程组显然无解；
- 当 $d_{r+1} = 0$ 时，分两种情形：

1. 若 $r = n$, 则有:

$$\begin{cases} x_1 = d_1, \\ x_2 = d_2, \\ \dots\dots\dots \\ x_n = d_n. \end{cases} \quad (4.2)$$

这表明方程组 $Ax = b$ 有惟一解 $x_1 = d_1, x_2 = d_2, \dots, x_n = d_n$

2. 若 $r < n$, 则由 4.1 式移项可得

$$\begin{cases} x_1 = d_1 - c_{1,r+1}x_{r+1} - \dots - c_{1n}x_n \\ x_2 = d_2 - c_{2,r+1}x_{r+1} - \dots - c_{2n}x_n \\ \dots \\ x_r = d_r - c_{r,r+1}x_{r+1} - \dots - c_{rn}x_n \end{cases} \quad (4.3)$$

这表明任给 $x_{r+1}, \dots, x_n$ 的一组值, 就可以惟一地解出 $x_1, \dots, x_r$ 的一组值, 从而解出方程组的一个解。因此, $x_1, \dots, x_r$ 可作为约束未知量, $x_{r+1}, \dots, x_n$ 可作为自由未知量, 而 4.3 式就是方程组 $Ax = b$ 的自由未知量表示的通解, 其中 $x_{r+1}, \dots, x_n$ 可任意取值。因此, 当 $r(A) = r(\bar{A}) = r < n$ 时, n 元线性方程组 (4.1.8) 有无穷多解, 且通解中有 $n - r$ 个自由未知量, 再令 $n - r$ 个自由未知量分别为 $n - r$ 个任意常数, 则参数式通解中含 $n - r$ 个任意常数。

 **练习 4.3** 若 $\bar{A}$ 经过初等行变换后得到 $\begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & a & -1 \\ 0 & 0 & (a-1)(a+3) & (a-1)^2 \end{bmatrix}$ 试判定方程组 $Ax = b$ 解的情况, 并在有解时, 求其全部解。

解 显然未知量个数 $n = 3$ 。

1. 当 $a \neq 1$ 且 $a \neq -3$ 时, $r(\bar{A}) = r(A) = 3$, 故方程组有唯一解。即将阶梯形矩阵进一步化为简化行阶梯形:

$$\bar{A} \rightarrow \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & a & -1 \\ 0 & 0 & 1 & \frac{a-1}{a+3} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 4 - \frac{3(a-1)}{a+3} \\ 0 & 1 & 0 & -1 - \frac{a+3}{a(a-1)} \\ 0 & 0 & 1 & \frac{a-1}{a+3} \end{bmatrix}$$

得唯一解为:

$$\begin{cases} x_1 = \frac{a+15}{a+3} \\ x_2 = \frac{-3-a^2}{a+3} \\ x_3 = \frac{a-1}{a+3} \end{cases}$$

2. 当 $a = 1$ 时, $r(A) = r(A) = 2 < 3$, 方程组有无穷多解, 此时:

$$\bar{A} \rightarrow \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

得到通解:

$$\begin{cases} x_1 = 4 - 3x_3 \\ x_2 = -1 - x_3 \end{cases} \quad \text{或表示为} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 - 3c \\ -1 - c \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}$$

3. 当 $a = -3$ 时, $r(A) = 2r(A) = 3$, 方程组无解。

对于 n 元齐次线性方程组 $Ax = 0$, 总有 $r(A) = r(\bar{A})$, 即方程组必有解, 因此有定理:

定理 4.2 (齐次线性方程组的解的判定)

- $Ax = 0$ 有唯一解 $\Leftrightarrow Ax = 0$ 只有零解 $\Leftrightarrow r(A) = r(\bar{A}) = n$
- $Ax = 0$ 有无穷多解 $\Leftrightarrow Ax = 0$ 有非零解 $\Leftrightarrow r(A) = r(\bar{A}) < n$

**4.1.4 数域****定义 4.2 (数域)**

如果数的集合 F 包含数 0 和 1, 并且 F 中任何两个数的和、差、积和商 (除数不为零) 都仍是 F 中的数, 即 F 对数的加法、减法、乘法和除法 (除数不为零) 这 4 种运算都是封闭的, 则称数集 F 是一个数域。



由定义易见, 全体有理数组成的集合、全体实数组成的集合和全体复数组成的集合都是数域, 我们分别称其为有理数域、实数域和复数域, 并分别用 $\mathbb{Q}, \mathbb{R}$ 和 $\mathbb{C}$ 来代表这 3 个数域, 这是最重要和最常用的 3 个数域。由于任何数域都包含数 0 和 1, 所以任何数域都包含有理数域。

根据定义, 也容易知道全体整数之集 $\mathbb{Z}$, 全体正实数之集 $\mathbb{R}^+$ 和全体无理数之集 $\hat{\mathbb{Q}}$ 都不是数域。事实上, 由于 $\mathbb{Z}$ 对除法不封闭, $\mathbb{R}^+$ 对减法不封闭, $\hat{\mathbb{Q}}$ 对乘法不封闭, 所以这 3 个数集都不是数域。

4.2 向量组的线性相关性

4.2.1 向量组及其运算

定义 4.3 (向量)

由数域 F 中的 n 个数 $a_1, a_2, \dots, a_n$ 组成的有序数组

$$\alpha = [a_1 \ a_2 \ \cdots \ a_n] \quad \text{行向量}$$

或者:

$$\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = [a_1 \ a_2 \ \cdots \ a_n]^T \quad \text{列向量}$$

称为一个 n 维向量, 其中 a_i 称为第 i 个分量 (坐标)。



注 向量是一种特殊的矩阵, 因此其线性运算和向量乘法规律可类比 2.1 矩阵的运算性质 (不再赘述)。向量可以用括号包裹起来 (以后的使用将不再区分方括号和括号)

基于定义 4.2, 不难证明一些向量可以组成一个数域, 即有:

定义 4.4 (向量空间 F^n)

数域 F 上 n 维向量的全体, 连同向量加法及数乘向量的运算, 称其为数域 F 上的一个 n 维向量空间, 记为 F^n 。特别地, 当 $F = \mathbb{R}$ 时的向量空间记为 $\mathbb{R}^n$, 当 $F = \mathbb{C}$ 时的向量空间记为 $\mathbb{C}^n$, 这是两个最重要且最常用的向量空间, 并分别称它们为实 n 维向量空间和复 n 维向量空间。



由向量的线性运算和向量相等的定义, 可将线性方程组 $Ax = b$ 写成:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{或} \quad x_1 \alpha_1 + x_2 \alpha_2 + \cdots + x_n \alpha_n = b$$

4.2.2 线性表示和等价向量组

定义 4.5 (线性表示)

设 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是一组 n 维向量, $k_1, k_2, \dots, k_s$ 是一组常数, 则称向量:

$$k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_s \alpha_s$$

为 $\alpha_1, \alpha_2, \dots, \alpha_s$ 的一个线性组合, 称常数 $k_1, k_2, \dots, k_s$ 为该线性组合的系数。又如果向量 β 可以表示为:

$$\beta = k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_s \alpha_s \quad (4.4)$$

则称向量 β 可由向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性表示或线性表出。



关于定义 4.5, 有注意事项:

- 若 $\alpha = k\beta$, 则称向量 α 与 β 成比例;
- 零向量 O 是任一向量组的线性组合;
- 向量组中每一向量都可由该向量组线性表示;
- 任一 n 维向量 $\alpha = (a_1 \ a_2 \ \cdots \ a_n)^T$ 都是基本单位向量组

$$\varepsilon_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$$

$$\varepsilon_2 = \begin{bmatrix} 0 & 1 & \cdots & 0 \end{bmatrix}^T$$

$$\vdots$$

$$\varepsilon_n = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}^T$$

的一个线性组合, 因为 $\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n$ 。

 **练习 4.4** 设有一组向量:

$$\alpha_1 = (1, 4, 0, 2)^T \quad \alpha_2 = (2, 7, 1, 3)^T$$

$$\alpha_3 = (0, 1, -1, a)^T \quad \beta = (3, 10, b, 4)^T$$

问 a, b 取何值时, β 可由向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性表示? 并在可以线性表示时, 求出此表示式。

 **解** 设有一组数 x_1, x_2, x_3 使得:

$$x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = \beta$$

对上面这个非齐次线性方程组的增广矩阵施行初等行变换

$$\bar{A} = \left[A \mid \beta \right] = \left[\alpha_1 \quad \alpha_2 \quad \alpha_3 \mid \beta \right] = \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 4 & 7 & 1 & 10 \\ 0 & 1 & -1 & b \\ 2 & 3 & a & 4 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & a-1 & 0 \\ 0 & 0 & 0 & b-2 \end{array} \right] = B$$

由阶梯形矩阵 B 可见:

1. 当 $b \neq 2$ 时, 方程组无解, 此时 β 不能由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示;
2. 当 $b = 2$ 且 $a \neq 1$ 时, 有 $r(A) = r(\bar{A}) = 3$ (未知量个数), 故此时方程组有惟一解。为求解, 将 B 再化成简化行阶梯形矩阵:

$$B \xrightarrow{r_3(\frac{1}{a-1})} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

由此得方程组的惟一解为 $x_1 = -1, x_2 = 2, x_3 = 0$, 故此时 β 可由 $\alpha_1, \alpha_2, \alpha_3$ 唯一地线性表示为 $\beta = -\alpha_1 + 2\alpha_2 + 0\alpha_3$;

3. 当 $b = 2$ 且 $a = 1$ 时, 有 $r(A) = r(\bar{A}) = 2 < 3$, 故此时方程组有解且有无穷多解。为求解, 将 B 再化成简化行阶梯形矩阵:

$$B \xrightarrow{r_{21}(-2)} \left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

由此得方程组的参数形式的通解为:

$$x_1 = -1 - 2c \quad x_2 = 2 + c \quad x_3 = c$$

因此, β 可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示为 $\beta = (-1 - 2c)\alpha_1 + (2 + c)\alpha_2 + c\alpha_3$ (其中 c 为任意常数)。

定义 4.6 (等价向量组)

设有两个向量组

$$(I) \alpha_1, \alpha_2, \dots, \alpha_s \quad (II) \beta_1, \beta_2, \dots, \beta_r$$

如果 (I) 中每个向量都可由 (II) 线性表示, 则称 (I) 可由 (II) 线性表示; 如果 (I) 与 (II) 可以互相线性表示, 则称向量组 (I) 与向量组 (II) 等价。

向量组的等价关系满足自反性、对称性和传递性。

练习 4.5 设有两个向量组

$$(I) : \alpha_1 = (1, 1, 0, 0)^T, \quad \alpha_2 = (1, 0, 1, 1)^T, \quad \alpha_3 = (1, 3, -2, -2)^T$$

$$(II) : \beta_1 = (2, -1, 3, 3)^T, \quad \beta_2 = (0, 1, -1, -1)^T$$

证明向量组 (I) 与 (II) 等价。

解 考虑下列矩阵的初等行变换:

$$\left[\begin{array}{ccc|cc} \alpha_1 & \alpha_2 & \alpha_3 & \beta_1 & \beta_2 \end{array} \right] = \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 0 \\ 1 & 0 & 3 & -1 & 1 \\ 0 & 1 & -2 & 3 & -1 \\ 0 & 1 & -2 & 3 & -1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 0 \\ 0 & -1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 3 & -1 \\ 0 & 1 & -2 & 3 & -1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 0 \\ 0 & -1 & 2 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

由阶梯形矩阵可见:

$$r \left[\begin{array}{ccc} \alpha_1 & \alpha_2 & \alpha_3 \end{array} \right] = r \left[\begin{array}{ccc|c} \alpha_1 & \alpha_2 & \alpha_3 & \beta_j \end{array} \right] = 2, \quad j = 1, 2$$

因此由线性方程组有解判定定理知方程组:

$$x_1 \alpha_1 + x_2 \alpha_2 + x_3 \alpha_3 = \beta_j, \quad j = 1, 2$$

有解, 即向量 β_1, β_2 均可由向量组 (I) 线性表示。由上面的阶梯形矩阵还可看出:

$$r \left[\begin{array}{cc} \beta_1 & \beta_2 \end{array} \right] = r \left[\begin{array}{cc|c} \beta_1 & \beta_2 & \alpha_j \end{array} \right] = 2 \quad j = 1, 2, 3$$

故向量 $\alpha_1, \alpha_2, \alpha_3$ 均可由向量组 (II) 线性表示, (I) 与 (II) 等价。

4.2.3 线性相关和线性无关**定义 4.7 (线性相关和线性无关)**

设 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是一组 n 维向量, 如果存在一组不全为零的常数 $k_1, k_2, \dots, k_s$, 使得:

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s = 0 \quad (4.5)$$

称向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是线性相关的。

一个向量组不是线性相关的, 则称它是线性无关的, 也就是说如果式 4.5 仅在 $k_1 = k_2 = \dots = k_s = 0$ 时才成立 (即若 $k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s = 0$, 则必有 $k_1 = k_2 = \dots = k_s = 0$), 称向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性无关。

定理 4.3 (定义 4.7 的等价叙述)

向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关 (线性无关)

$\Leftrightarrow$ 齐次线性方程组 $Ax = 0$ 有非零解 (只有零解)

$\Leftrightarrow$ 矩阵 $A = [\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_s]$ 的秩小于 s (等于 s)

也可以表述为: 矩阵 A 的列向量组线性相关 (线性无关)

$\Leftrightarrow$ 齐次线性方程组 $Ax = 0$ 有非零解 (只有零解)

$\Leftrightarrow A$ 的秩小于其列数 (A 为列满秩矩阵)

笔记

- n 个 n 维向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性相关 (线性无关) $\Leftrightarrow \det[\alpha_1 \ \alpha_2 \ \dots \ \alpha_n] = 0 (\neq 0)$;
- 若 $s > n$, 则 s 个 n 维向量 $\alpha_1, \alpha_2, \dots, \alpha_s$ 必线性相关。特别地, $n+1$ 个 n 维向量必线性相关。

练习 4.6 设向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, 又知向量:

$$\beta_1 = \alpha_1 + 2\alpha_2 \quad \beta_2 = 2\alpha_1 + 3\alpha_2 + 4\alpha_3 \quad \beta_3 = 5\alpha_1 + 6\alpha_2 + 7\alpha_3$$

试判别向量组 $\beta_1, \beta_2, \beta_3$ 的线性相关性。

解 设有一组数 x_1, x_2, x_3 , 使得:

$$x_1\beta_1 + x_2\beta_2 + x_3\beta_3 = 0$$

将已知的线性表示式代入上式, 得:

$$x_1(\alpha_1 + 2\alpha_2) + x_2(2\alpha_1 + 3\alpha_2 + 4\alpha_3) + x_3(5\alpha_1 + 6\alpha_2 + 7\alpha_3) = 0$$

即有:

$$(x_1 + 2x_2 + 5x_3)\alpha_1 + (2x_1 + 3x_2 + 6x_3)\alpha_2 + (4x_2 + 7x_3)\alpha_3 = 0$$

由于 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, 根据定义有:

$$\begin{cases} x_1 + 2x_2 + 5x_3 = 0 \\ 2x_1 + 3x_2 + 6x_3 = 0 \\ 4x_2 + 7x_3 = 0 \end{cases}$$

上面这个齐次线性方程组的系数矩阵 $P = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 3 & 6 \\ 0 & 4 & 7 \end{bmatrix}$ 为列满秩矩阵, 所以方程组只有零解, 即使得 x_1, x_2, x_3 必

全都为零, 故由定义知 $\beta_1, \beta_2, \beta_3$ 线性无关。

笔记 证明向量组的线性相关性时常常采用待定系数法, 结合线性方程组、矩阵等知识按照定义求解。

线性相关的判定也可以由定理 4.4 等价表述:

定理 4.4

向量组 $\alpha_1, \dots, \alpha_s$ ($s \geq 2$) 线性相关的充分必要条件是该组中至少存在 1 个向量可由该组中其余 $s-1$ 个向量线性表示。

证明 先证明必要性: 设向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关, 即存在不全为零的常数 $k_1, k_2, \dots, k_s$, 使得:

$$k_1\alpha_1 + k_2\alpha_2 + \dots + k_s\alpha_s = 0$$

由于 $k_1, k_2, \dots, k_s$ 不全为零, 故其中至少有 1 个不为零, 不妨设 $k_s \neq 0$, 则由上式可解得:

$$\alpha_s = -\frac{k_1}{k_s}\alpha_1 - \dots - \frac{k_{s-1}}{k_s}\alpha_{s-1}$$

这表明 α_s 可由该向量组中其余 $s-1$ 个向量线性表示, 证毕。

再证明充分性: 不妨设 α_s 可由该向量组中其余 $s-1$ 个向量线性表示为:

$$\alpha_s = \lambda_1\alpha_1 + \dots + \lambda_{s-1}\alpha_{s-1}$$

移项, 得:

$$\lambda_1\alpha_1 + \dots + \lambda_{s-1}\alpha_{s-1} - \alpha_s = 0$$

由于上式左端线性组合的系数

$$\lambda_1, \dots, \lambda_{s-1}, -1$$

不全为零, 所以 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关, 证毕。

4.2.4 线性相关性的基本性质

定理 4.5 (线性无关填补定理)

若向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性无关, 而向量组 $\alpha_1, \alpha_2, \dots, \alpha_s, \beta$ 线性相关, 则 β 可由向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性表示, 且表示法惟一。

证明 利用定义来证, 由条件知存在不全为零的常数 $k_1, \dots, k_{s+1}$, 使得:

$$\sum_{i=1}^s k_i \alpha_i + k_{s+1} \beta = 0$$

且必有 $k_{s+1} \neq 0$ (否则, $k_{s+1} = 0$, 则上述事实成为: 存在不全为零的常数 $k_1, \dots, k_s$, 使得 $\sum_{i=1}^s k_i \alpha_i = 0$, 这与已知 $\alpha_1, \dots, \alpha_s$ 线性无关矛盾), 于是解得:

$$\beta = \sum_{i=1}^s \left(-\frac{k_i}{k_{s+1}} \right) \alpha_i$$

即 β 可由 $\alpha_1, \dots, \alpha_s$ 线性表示。再证表示法惟一, 若有:

$$\beta = \sum_{i=1}^s \lambda_i \alpha_i$$

又有:

$$\beta = \sum_{i=1}^s \mu_i \alpha_i$$

两式相减, 得:

$$\sum_{i=1}^s (\lambda_i - \mu_i) \alpha_i = 0$$

因 $\alpha_1, \dots, \alpha_s$ 线性无关, 得 $\lambda_i - \mu_i = 0$, 即 $\lambda_i = \mu_i$ ($i = 1, \dots, s$), 所以 β 由 $\alpha_1, \dots, \alpha_s$ 线性表示的表示式惟一。

定理 4.6 (线性相关性的子集关系)

如果向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 有一个部分组 (非空子集合) 线性相关, 则向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关。逆命题同样成立。

证明 不妨设该向量组的部分组 $\alpha_1, \alpha_2, \dots, \alpha_r$ 线性相关, 则存在不全为零的常数 $k_1, k_2, \dots, k_r$, 使得:

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_r \alpha_r = 0$$

于是, 存在不全为零的 s 个常数 $k_1, k_2, \dots, k_r, 0, \dots, 0$, 使得:

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_r \alpha_r + 0 \alpha_{r+1} + \dots + 0 \alpha_s = 0$$

所以, 向量组 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关。

笔记 定理 4.6 可简单表述为: 部分组线性相关, 则整体组线性相关; 整体组线性无关, 则其任何部分组线性无关。

命题 4.1 (延伸 (缩短) 组的线性相关性)

设 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性无关, 则在每个向量中相同位置添加分量得到的新向量组称为延伸组, 且延伸组线性无关; 相反的, 若 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性相关, 则在每个向量中相同位置去掉分量得到的新向量组称为缩短组, 且缩短组线性相关。

4.3 向量组的秩

4.3.1 向量组的极大无关组

在一个向量组或向量集合中, 如何选取它的一个“最大”的线性无关的子集合有着重要意义。满足这样描述的一组向量可以由定义4.8表述:

定义 4.8 (向量组的极大无关组)

如果向量组 U 有一个部分组 $\alpha_1, \alpha_2, \dots, \alpha_r$ 满足:

1. $\alpha_1, \alpha_2, \dots, \alpha_r$ 线性无关;
2. U 中任意向量 α 都可由向量组 $\alpha_1, \alpha_2, \dots, \alpha_r$ 线性表示;

则称 $\alpha_1, \alpha_2, \dots, \alpha_r$ 为向量组 U 的一个极大线性无关组, 简称为极大无关组 (或最大无关组)。

显然, 一个向量组与他的一个极大无关组是等价的。下面用练习题介绍求极大无关组的一般方法:

练习 4.7 求向量组 (I): $\alpha_1 = (1, -2, 0, 3)^T, \alpha_2 = (2, -5, -3, 6)^T, \alpha_3 = (0, 1, 3, 0)^T, \alpha_4 = (2, -1, 4, -7)^T, \alpha_5 = (5, -8, 1, 2)^T$ 的一个极大无关组, 并用极大无关组线性表示该组中其他向量。

解 以向量组 (I) 为矩阵 A 的列向量组来构造矩阵 A , 并用初等行变换将 A 化成阶梯形矩阵

$$A = [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5] = \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = B$$

阶梯形矩阵 B 的首非零元所在的列为第 1、2、4 列, 下面我们说明: A 的相应的列向量——即 $\alpha_1, \alpha_2, \alpha_4$ 便可作为向量组 (I) 的一个极大无关组。这是因为, 首先, 由前面已作的初等行变换, 知矩阵

$$[\alpha_1 \quad \alpha_2 \quad \alpha_4] \xrightarrow{\text{初等行变换}} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{bmatrix}$$

的秩为 3, 所以 $\alpha_1, \alpha_2, \alpha_4$ 线性无关。其次, 考虑线性方程组:

$$x_1 \alpha_1 + x_2 \alpha_2 + x_3 \alpha_4 = \alpha_3$$

仍由前面已作的初等行变换, 有

$$[\alpha_1 \quad \alpha_2 \quad \alpha_4 \quad \alpha_3] \xrightarrow{\text{初等行变换}} \begin{bmatrix} 1 & 2 & 2 & 0 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

从而有 $r[\alpha_1 \quad \alpha_2 \quad \alpha_4] = r[\alpha_1 \quad \alpha_2 \quad \alpha_4 \mid \alpha_3] = 3$, 故由方程组有解判定定理知 α_3 可由 $\alpha_1, \alpha_2, \alpha_4$ 惟一地线性表示, 同理知 α_5 可由 $\alpha_1, \alpha_2, \alpha_4$ 惟一地线性表示。以上两方面说明 $\alpha_1, \alpha_2, \alpha_4$ 可作为向量组 (I) 的极大无关组。

为了用极大无关组 $\alpha_1, \alpha_2, \alpha_4$ 线性表示 α_3, α_5 , 再把矩阵 B 化成简化行阶梯形矩阵:

$$B \rightarrow \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \stackrel{\text{记为}}{=} C$$

由此即得:

$$\alpha_3 = 2\alpha_1 - \alpha_2 \quad \alpha_5 = \alpha_1 + \alpha_2 + \alpha_4$$

这是因为

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_4 & \alpha_3 \end{bmatrix} \xrightarrow{\text{初等行变换}} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

所以非齐次线性方程组:

$$x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_4 = \alpha_3$$

有惟一解 $x_1 = 2, x_2 = -1, x_3 = 0$, 从而有 $\alpha_3 = 2\alpha_1 - \alpha_2$ 。类似可说明 $\alpha_5 = \alpha_1 + \alpha_2 + \alpha_4$ 的理由。

定理 4.7

设向量组 (I): $\alpha_1, \alpha_2, \dots, \alpha_s$ 可由向量组 (II): $\beta_1, \beta_2, \dots, \beta_r$ 线性表示, 则:

- 当 $s > r$ 时, 向量组 (I) 线性相关;
- 当向量组 (I) 线性无关时, 必有 $s \leq r$



4.3.2 向量组的秩

定义 4.9 (向量组的秩)

如果向量组 U 仅含零向量, 规定 U 的秩为零; 否则称 U 的极大无关组所含向量的个数为向量组 U 的秩, U 的秩记为 $r(U)$ 。



注 通常, 称一个矩阵 A 的列 (行) 向量组的秩为 A 的列 (行) 秩。对任何矩阵 A , 有 $r(A) = A$ 的列秩 = A 的行秩。可以证明, 矩阵的初等行变换不改变矩阵列向量之间的线性关系。

练习 4.8 已知两个向量组

$$\begin{aligned} (I) : \alpha_1 &= (1, 2, -3)^T & \alpha_2 &= (3, 0, 1)^T & \alpha_3 &= (9, 6, -7)^T \\ (II) : \beta_1 &= (0, 1, -1)^T & \beta_2 &= (a, 2, 1)^T & \beta_3 &= (b, 1, 0)^T \end{aligned}$$

如果向量组 (II) 与向量组 (I) 有相同的秩, 且 β_3 可由 (I) 线性表示, 试求常数 a, b 的值。

解 先求出 (I) 的秩: 显然 α_1, α_2 线性无关, 又由计算可得 $\alpha_3 = 3\alpha_1 + 2\alpha_2$, 故 α_1, α_2 为 (I) 的极大无关组, 从而有 $r(I) = 2$ 由条件 $r(II) = r(I) = 2$, 及 (II) 含 3 个向量, 知 (II) 线性相关, 得行列式:

$$\det[\beta_1 \ \beta_2 \ \beta_3] = \begin{vmatrix} 0 & a & b \\ 1 & 2 & 1 \\ -1 & 1 & 0 \end{vmatrix} = 0$$

由此解得 $a = 3b$ 。再由 β_3 可由 (I) 线性表示, 知 β_3 可由 () 的极大无关组 α_1, α_2 线性表示, 故向量组 $\alpha_1, \alpha_2, \beta_3$ 线性相关, 得:

$$\det[\alpha_1 \ \alpha_2 \ \beta_3] = \begin{vmatrix} 1 & 3 & b \\ 2 & 0 & 1 \\ -3 & 1 & 0 \end{vmatrix} = 0$$

由此解得 $b = 5$, 所以 $a = 15, b = 5$

关于向量组的秩, 有一些基本性质:

命题 4.2

如果向量组 (I) 和 (II) 都是线性无关组, 且 (I) 与 (II) 等价, 则 (I) 与 (II) 所含向量的个数必相同。



命题 4.3

设 $r(U) = r$, 则 U 中任何 r 个线性无关的向量都可作为 U 的极大无关组。

定理 4.8

若向量组 (I) 可由向量组 (II) 线性表示, 则 $r(I) \leq r(II)$

证明 设 (I) 和 (II) 的极大无关组分别为 (I') 和 (II'), 由于向量组的极大无关组与向量组本身等价, 故由定理的条件知 (I') 可以由 (II') 线性表示, (I') 又是线性无关的, 于是由定理 4.7 知 (I') 所含向量的个数必不大于 (II') 所含向量的个数, 此即 $r(I) \leq r(II)$ 证毕。

根据定理 4.8, 当两向量组等价时, 他们的秩相同。

定理 4.9 (矩阵秩的重要定理)

- 对于矩阵 $A_{m \times n}, B_{m \times n}$, 有 $r(A+B) \leq r(A) + r(B)$;
- 对于矩阵 $A_{m \times n}, B_{n \times p}$, 有 $r(AB) \leq \min \{r(A), r(B)\}$

证明

- 设 A, B 按列分块分别为:

$$A = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \quad B = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]$$

则:

$$A+B = [\alpha_1 + \beta_1 \quad \alpha_2 + \beta_2 \quad \cdots \quad \alpha_n + \beta_n]$$

设 A 的列向量组的极大无关组为:

$$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r} \quad (\text{故 } r(A) = r)$$

设 B 的列向量组的极大无关组为:

$$\beta_{j_1}, \beta_{j_2}, \cdots, \beta_{j_s} \quad (\text{故 } r(B) = s)$$

显然 $A+B$ 的每个列向量均可由向量组

$$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \beta_{j_1}, \beta_{j_2}, \cdots, \beta_{j_s}$$

线性表示, 故可知 $A+B$ 的列向量组的秩 $r(A+B)$ 不会大于上述向量组的秩, 而该向量组的秩当然不会大于它所含向量的个数 $r+s$, 于是得

$$r(A+B) \leq r+s = r(A) + r(B)$$

- 只需要证 $r(AB) \leq r(A)$, 且 $r(AB) \leq r(B)$ 。先证 $r(AB) \leq r(A)$: 设矩阵 A 按列分块为 $A = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]$, 设矩阵 $B = (b_{ij})_{n \times p}$, 则由矩阵乘法知 AB 的第 j 列为:

$$A \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = b_{1j}\alpha_1 + b_{2j}\alpha_2 + \cdots + b_{nj}\alpha_n \quad (j = 1, 2, \cdots, p)$$

上式表明 AB 的列向量组可由 A 的列向量组线性表示, 于是由定理 4.8 知 AB 的列秩不大于 A 的列秩, 即 $r(AB) \leq r(A)$ 。

此式表明矩阵积的秩不大于其左因子矩阵之秩, 利用此结果及矩阵转置其秩不变, 得:

$$r(AB) = r(AB)^T = r(B^T A^T) \leq r(B^T) = r(B)$$

4.4 线性方程组解的结构

本节将完整阐述线性方程组的解的组成及性质。

4.4.1 齐次线性方程组的解

对于齐次线性方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \quad (4.6)$$

若 $x_1 = \xi_{11}, x_2 = \xi_{21}, \cdots, x_n = \xi_{n1}$ 使得方程 $Ax = 0$ 成立，则 $x = \xi = (\xi_{11} \ \xi_{21} \ \cdots \ \xi_{n1})^T$ 称为方程组 4.6 的解向量，也就是 $Ax = 0$ 的解。齐次线性方程组的解显然满足性质：

- 若 $x_1 = \xi_1, x_2 = \xi_2$ 为 $Ax = 0$ 的解，则 $x = \xi_1 + \xi_2$ 也是 $Ax = 0$ 的解；
- 若 $x_1 = \xi_1$ 为 $Ax = 0$ 的解， k 为实数，则 $x = k\xi_1$ 也是 $Ax = 0$ 的解。

因此，一个线性方程组的所有解可以构成一个向量空间（详见第 5 章）。称此向量空间为齐次线性方程组 $Ax = 0$ 的解空间（解集合）。

$$S = \{x \in F^n | Ax = 0\}$$

定义 4.10 (基础解系)

如果齐次方程组 $Ax = 0$ 的一组解 $\xi_1, \xi_2, \cdots, \xi_s$ 满足：

1. $\xi_1, \xi_2, \cdots, \xi_s$ 线性无关；
2. 方程组 $Ax = 0$ 的任一解都可由 $\xi_1, \xi_2, \cdots, \xi_s$ 线性表示

则称 $\xi_1, \xi_2, \cdots, \xi_s$ 为齐次线性方程组的一个基础解系。

 **笔记** 可以说明：齐次方程组的基础解系就是齐次方程组解集合的极大无关组，且和基础解系等价的线性无关向量组也是基础解系。

如果 $\xi_1, \xi_2, \cdots, \xi_s$ 为齐次线性方程组 $Ax = 0$ 的基础解系，则方程组 $Ax = 0$ 的通解可表示为：

$$x = k_1\xi_1 + k_2\xi_2 + \cdots + k_s\xi_s \quad (4.7)$$

其中 $k_1, k_2, \cdots, k_s$ 为任意实数。

定理 4.10 (基础解系和秩的关系)

设 A 为 $m \times n$ 矩阵， $r(A) = r < n$ ，则 n 元齐次线性方程组 $Ax = 0$ 必存在基础解系，且基础解系含 $n - r$ 个线性无关的向量。

求解齐次线性方程组的基础解系和通解一般是必考内容，见练习：

 **练习 4.9** 求齐次线性方程组的基础解系和通解：

$$\begin{cases} x_1 + 2x_2 + 4x_3 - 3x_4 = 0, \\ 3x_1 + 5x_2 + 6x_3 - 4x_4 = 0, \\ 4x_1 + 5x_2 - 2x_3 + 3x_4 = 0, \\ 3x_1 + 8x_2 + 24x_3 - 19x_4 = 0. \end{cases}$$

解 对方程组的系数矩阵 A 作初等行变换

$$A = \begin{bmatrix} 1 & 2 & 4 & -3 \\ 3 & 5 & 6 & -4 \\ 4 & 5 & -2 & 3 \\ 3 & 8 & 24 & -19 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -8 & 7 \\ 0 & 1 & 6 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

由阶梯形矩阵得方程组的由自由未知量表示的通解为:

$$\begin{cases} x_1 = 8x_3 - 7x_4, \\ x_2 = -6x_3 + 5x_4 \end{cases} \quad (x_3, x_4 \text{ 为自由未知量})$$

令 $x_3 = 1, x_4 = 0$, 得解向量 $\xi_1 = (8, -6, 1, 0)^T$; 令 $x_3 = 0, x_4 = 1$, 得解向量 $\xi_2 = (-7, 5, 0, 1)^T$, ξ_1, ξ_2 就是方程组的基础解系, 所以, 方程组的结构式通解为:

$$x = c_1 \xi_1 + c_2 \xi_2 \quad (c_1, c_2 \text{ 为任意常数})$$

或令自由未知量 $x_3 = c_1, x_4 = c_2$, 并将通解写成向量形式, 便得结构式通解:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8c_1 - 7c_2 \\ -6c_1 + 5c_2 \\ c_1 \\ c_2 \end{bmatrix} = c_1 \begin{bmatrix} 8 \\ -6 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -7 \\ 5 \\ 0 \\ 1 \end{bmatrix} \quad (c_1, c_2 \text{ 为任意常数})$$

定理 4.11 (基础解系的不唯一性)

设 A 为 $m \times n$ 矩阵, $r(A) = r < n$, 则齐次线性方程组 $Ax = 0$ 的任何 $n - r$ 个线性无关的解向量都可作为 $Ax = 0$ 的基础解系。

定理 4.11 表明, 基础解系的选取是不唯一的, 但是基础解系含有的向量个数是唯一的, 这由解空间的维度定义 (详见第 5 章)。

命题 4.4

设矩阵 $A_{m \times n}, B_{n \times p}$ 满足 $AB = O$, 则:

$$r(A) + r(B) \leq n \quad (4.8)$$

证明 如果 $B = O$, 显然成立; 如果 $B \neq O$, 则 B 至少有 1 列非零. 将 B 按列分块为 $B = [\beta_1 \ \beta_2 \ \cdots \ \beta_p]$, 则由:

$$O = AB = A[\beta_1 \ \beta_2 \ \cdots \ \beta_p] = [A\beta_1 \ A\beta_2 \ \cdots \ A\beta_p]$$

得到:

$$A\beta_j = 0 \quad (j = 1, 2, \cdots, p)$$

这表明矩阵 B 的每个列向量都是齐次线性方程组 $Ax = 0$ 的解向量, B 的列向量组中有 $r(B)$ 个线性无关的向量. 因此, 方程组 $Ax = 0$ 的解集合中至少含 $r(B)$ 个线性无关的解向量, 故方程组 $Ax = 0$ 的基础解系至少含 $r(B)$ 个向量, 而 $Ax = 0$ 的基础解系所含向量个数为 $n - r(A)$, 于是得:

$$n - r(A) \geq r(B)$$

即得:

$$r(A) + r(B) \leq n$$

定理 4.12 (同解等秩定理)

设齐次线性方程组 $Ax = 0$ 与 $Bx = 0$ 同解, 则 $r(A) = r(B)$

4.4.2 非齐次线性方程组

现在讨论非齐次线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (4.9)$$

通常称方程组 $Ax = 0$ 为与方程组 $Ax = b$ 对应的齐次线性方程组 (或称 $Ax = 0$ 为 $Ax = b$ 的导出组)。非齐次线性方程组和其导出组的解满足性质:

- 如果 η_1, η_2 都是非齐次线性方程组 $Ax = b$ 的解, 则 $\eta_1 - \eta_2$ 是对应齐次线性方程组 $Ax = 0$ 的解。
- 如果 η 是 $Ax = b$ 的一个解, ξ 是 $Ax = 0$ 的一个解, 则 $\eta + \xi$ 是 $Ax = b$ 的一个解。

定理 4.13 (非齐次线性方程组的解的结构)

设 η^* 为非齐次线性方程组 $Ax = b$ 的一个特解 (不含任意常数的一个确定解), 则 $Ax = b$ 的通解可表示为:

$$x = \eta^* + \sum_{i=1}^n c_i \xi_i \quad (4.10)$$

其中 $\xi_i (i = 1, 2, \dots, n-r)$ 为对应齐次线性方程组 $Ax = 0$ 的基础解系。

 练习 4.10 λ 取何值时, 线性方程组:

$$\begin{cases} (2-\lambda)x_1 + 2x_2 - 2x_3 = 1, \\ 2x_1 + (5-\lambda)x_2 - 4x_3 = 2, \\ -2x_1 - 4x_2 + (5-\lambda)x_3 = -\lambda - 1 \end{cases}$$

有惟一解、无解、有无穷多解? 并在有无穷多解时, 求其结构解。

解 由于含有参数 λ , 所以用初等行变换将 $\bar{A}$ 化成阶梯形矩阵的运算不是很方便。注意系数矩阵为方阵, 因此也可以先考虑系数行列式 $\det(A)$, 有:

$$\det(A) = \begin{vmatrix} 2-\lambda & 2 & -2 \\ 2 & 5-\lambda & -4 \\ -2 & -4 & 5-\lambda \end{vmatrix} = (\lambda-1)^2(10-\lambda)$$

1. 当 $\det(A) \neq 0$, 即 $\lambda \neq 1$ 且 $\lambda \neq 10$ 时, 由 Cramer 法则知方程组有惟一解;
2. 当 $\lambda = 10$ 时, 对增广矩阵作初等行变换:

$$\bar{A} = \left[\begin{array}{ccc|c} -8 & 2 & -2 & 1 \\ 2 & -5 & -4 & 2 \\ -2 & -4 & -5 & -11 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} -2 & -4 & -5 & -11 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

由上面的阶梯形矩阵可见 $r(A) = 2, r(\bar{A}) = 3$, 故由有解判定定理 4.1 知当 $\lambda = 10$ 时方程组无解;

3. 当 $\lambda = 1$ 时, 对增广矩阵作初等行变换:

$$\bar{A} = \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 2 & 4 & -4 & 2 \\ -2 & -4 & 4 & -2 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

可见 $r(A) = r(\bar{A}) = 1 < 3$, 所以当 $\lambda = 1$ 时方程组有无穷多解。可得通解:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 - 2c_1 + 2c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

其中 c_1, c_2 为任意常数。

第四章 练习

(1) 解下列线性方程组:

(a).

$$\begin{cases} x_1 - 5x_2 + 2x_3 - 3x_4 = 11 \\ -3x_1 + x_2 - 4x_3 + 2x_4 = -5 \\ -x_1 - 9x_2 - 4x_4 = 17 \\ 5x_1 + 3x_2 + 6x_3 - x_4 = -1 \end{cases}$$

(b).

$$x_1 - 4x_2 + 2x_3 - 3x_4 + 6x_5 = 4$$

(c).

$$\begin{cases} 2x_1 - 3x_2 + x_3 - 5x_4 = 1 \\ -5x_1 - 10x_2 - 2x_3 + x_4 = -21 \\ x_1 + 4x_2 + 3x_3 + 2x_4 = 1 \\ 2x_1 - 4x_2 + 9x_3 - 3x_4 = -16 \end{cases}$$

(2) 设有矩阵 $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ 满足:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad (i = 1, 2, \cdots, n)$$

证明 A 的列向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的秩等于 n 。

(3) 下述齐次线性方程组何时非零解? 何时只有零解?

$$\begin{cases} x_1 + 2x_2 - 11x_3 = 0 \\ 2x_1 - 5x_2 + 3x_3 = 0 \\ 5x_1 + x_2 + ax_3 = 0 \\ 6x_1 + 3x_2 + bx_3 = 0 \end{cases}$$

(4) 设向量组 $\{\alpha_1, \cdots, \alpha_s\}$ 的每个向量都可以由这个向量组的一个部分组 $\{\alpha_{i_1}, \cdots, \alpha_{i_r}\}$ 唯一线性表出。证明:

$\{\alpha_{i_1}, \cdots, \alpha_{i_r}\}$ 是 $\{\alpha_1, \cdots, \alpha_s\}$ 的一个极大线性无关组, 且 $r\{\alpha_1, \cdots, \alpha_s\} = r$

(5) 设向量组 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 线性无关。令:

$$\beta_1 = \alpha_1 - \alpha_2, \beta_2 = \alpha_2 - \alpha_3, \beta_3 = \alpha_3 - \alpha_4, \beta_4 = \alpha_4 - \alpha_1.$$

求 $\beta_1, \beta_2, \beta_3, \beta_4$ 的一个极大线性无关组。

(6) 设 $n-1$ 个方程的 n 元齐次线性方程组的系数矩阵为 B , 把 B 划去第 j 列得到的 $n-1$ 阶子式记作 D_j , 令:

$$\eta = \begin{bmatrix} D_1 \\ -D_2 \\ \vdots \\ (-1)^{n-1} D_n \end{bmatrix}$$

证明: η 是这个齐次线性方程组的一个解; 且若 $\eta \neq 0$, 则 η 是这个方程组的一个基础解系。

第五章 线性空间与欧式空间

本章来讨论近代数学以来有着重要地位的研究对象的一种数学抽象。

注 本章抽象概念较多，请读者打起精神。



5.1 线性空间

5.1.1 线性空间的定义

我们在第1章到第4章中讨论的数、矩阵、向量都满足线性运算性质，为了将这种线性运算的性质总结成某些集合的一般性质，我们引入线性空间的概念：

定义 5.1 (线性空间)

设 V 是一个以 $\alpha, \beta, \gamma, \dots$ 为元素的非空集合， F 是一个数域。在其中定义两种运算，一种叫加法： $\forall \alpha, \beta \in V, \alpha + \beta \in V$ ；另一种叫数量乘法： $\forall k \in F, \alpha \in V, k\alpha \in V$ ，并且满足下面八条运算法则：

1. 加法交换律： $\alpha + \beta = \beta + \alpha$ ；
2. 加法结合律： $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ ；
3. V 中存在零元素： $\exists \alpha_0 \in V, \forall \alpha \in V, \alpha + \alpha_0 = \alpha$ ，记 $\alpha_0 = \mathbf{0}$ ；
4. 负元素存在： $\forall \alpha \in V, \exists \beta \in V$ ，使 $\alpha + \beta = \mathbf{0}$ ，记 $\beta = -\alpha$ ；
5. 数乘结合律： $(kl)\alpha = k(l\alpha)$ ；
6. 存在 $1 \in F$ ： $1 \cdot \alpha = \alpha$ ；
7. 数乘分配律： $(k+l)\alpha = k\alpha + l\alpha$ ；
8. 元素分配律： $k(\alpha + \beta) = k\alpha + k\beta$ ，

则称 V 为数域 F 上的线性空间。 V 中元素称为向量。 F 为实(复)数域时，称 V 是实(复)线性空间。

线性空间也被成为向量空间，因此也称线性空间中的元素为向量。此处提及的“向量”和第4章中的 n 维向量不同，是针对所有可以被称为线性空间的集合的元素的一般称呼。命题5.1给出了几个常用的线性空间：

命题 5.1 (常见的线性空间)

- $F^n, \mathbb{R}^n, \mathbb{C}^n$ 分别是 $F, \mathbb{R}, \mathbb{C}$ 上的线性空间；
- 元素属于数域 F 的 $m \times n$ 矩阵的全体，按照矩阵的加法和数与矩阵的乘法，构成数域 F 上一个线性空间，记为 $F^{m \times n}$ ；
- 设 $A \in F^{m \times n}$ ，则齐次线性方程组 $Ax = 0$ 的解向量都是 F^n 中的向量。记 $Ax = 0$ 的解集合为 S ，不难验证 S 是一个线性空间。称 S 为方程组 $Ax = 0$ 的解空间；
- 定义在闭区间 $[a, b]$ 上的1元连续函数的全体，按照函数的加法和实数与函数的乘法，构成一个实线

性空间, 记为 $C[a, b]$;

- 在区间 I 上具有 m 阶连续导数的 1 元函数全体, 按照函数的加法和实数与函数的乘法, 构成一个实线性空间, 记为 $C^{(m)}(I)$;
- 记 $F[x]$ 是系数取自数域 F 的 1 元多项式全体, 即:

$$F[x] = \{a_0 + a_1x + \cdots + a_nx^n + \cdots \mid a_i \in F, i = 0, 1, \cdots\}$$

则 $F[x]$ 按照通常的多项式加法和数与多项式的乘法, 构成数域 F 上的一个线性空间。如果只考虑其中次数不超过 n 的多项式全体, 记为 $F[x]_n$, 即:

$$F[x]_n = \{a_0 + a_1x + \cdots + a_nx^n \mid a_i \in F, i = 0, 1, \cdots, n\}$$

则 $F[x]_n$ 也构成数域 F 上的一个线性空间。

利用线性空间的线性运算封闭性和 8 条线性运算性质, 可以验证一些集合不是线性空间:

 **练习 5.1** 对于全体 2 维实向量所组成的集合 W 中, 加法运算定义为通常的向量加法, 数乘运算定义为:

$$k \circ (a, b)^T = (ka, 0)^T$$

证明 W 不是线性空间。

证明 W 的数乘运算不满足定义 5.1 的运算规律 (6), 例如:

$$1 \circ (1, 2)^T = (1, 0)^T \neq (1, 2)^T$$

因此 W 不是线性空间。

线性空间满足一些基本性质:

定理 5.1 (线性空间的基本性质)

1. 线性空间的零元素是惟一的;
2. 线性空间的任一元素的负元素是惟一的;
3. $0\alpha = \mathbf{0}; (-1)\alpha = -\alpha; k\mathbf{0} = \mathbf{0}$;
4. 如果 $k\alpha = \mathbf{0}$, 则 $k = 0$ 或 $\alpha = \mathbf{0}$

5.1.2 线性子空间

定义 5.2 (线性子空间)

设 W 是线性空间 V 的一个非空子集, 如果 W 按照 V 中所定义的线性运算也构成一个线性空间, 则称 W 为 V 的一个子空间, 也称为线性子空间。

一般的, 称 V 本身和只有零元素的 $\{\mathbf{0}\}$ 子空间为 V 的平凡子空间。

根据定义 5.1 和 5.2, 线性空间 V 的非空子集 W 一定满足条件 (1, 2, 5, 6, 7, 8), 因此只需要判断其是否满足条件 (3, 4) 就可以断言 W 是不是 V 的子空间。结合线性空间的基本性质 (定理 5.1 的 (3)) 有:

定理 5.2 (线性子空间的判定定理)

设 W 是线性空间 V 的非空子集, 则:

$$W \text{ 是 } V \text{ 的子空间} \Leftrightarrow W \text{ 对于 } V \text{ 中的线性运算封闭}$$

 **练习 5.2** 设 $\alpha_1, \cdots, \alpha_n$ 是 F 上的线性空间 V 的向量组, 其非空子集 W 满足:

$$W = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \mid k_i \in F, i = 1, 2, \cdots, n\}$$

验证 W 是 V 的线性子空间。

证明 取 W 中的任意两个向量:

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \quad k_i \in F, i = 1, 2, \cdots, n$$

$$\beta = l_1\alpha_1 + l_2\alpha_2 + \cdots + l_n\alpha_n \quad l_i \in F, i = 1, 2, \cdots, n$$

则有:

$$\begin{aligned} \alpha + \beta &= k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n + l_1\alpha_1 + l_2\alpha_2 + \cdots + l_n\alpha_n \\ &= (k_1 + l_1)\alpha_1 + (k_2 + l_2)\alpha_2 + \cdots + (k_n + l_n)\alpha_n \in W \end{aligned}$$

则 W 对 V 的加法运算封闭, 同理可证明 W 对 V 的数乘运算封闭, 因此 W 是 V 的一个子空间。

一般记为 $W = \text{span}\{\alpha_1, \cdots, \alpha_n\}$ 或 $W = L\{\alpha_1, \cdots, \alpha_n\}$, 称 W 是向量组 $\alpha_1, \cdots, \alpha_n$ 张成的子空间。

5.1.3 线性空间的基、维数和坐标表示

通过第 4 章的讨论, 我们知道 F^n 中存在 n 个线性无关的向量 (例如 n 维基本单位向量组 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$), 而且对于 F^n 中任意 n 个线性无关的向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, F^n 中任一向量 α 都可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 唯一地线性表示。这些概念显示了向量空间 F^n 的基本结构和主要特征, 而将这些概念推广到一般的线性空间, 我们引入线性空间的基、维数及向量的坐标等基本概念。

定义 5.3 (线性空间的基、维数和坐标表示)

如果在线性空间 V 中存在一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 满足:

1. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关;
2. $\forall \alpha \in V, \alpha$ 可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表示:

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \quad (x_i \in F, i = 1, 2, \cdots, n) \quad (5.1)$$

则称 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 V 的一个基; 称基中所含向量的个数 n 为 V 的维数, 记为 $\dim(V) = n$, 并且称 V 为 n 维线性空间; 称式 5.1 中的 n 个有序数 $x_1, x_2, \cdots, x_n$ 为向量 α 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标, 记为 $(x_1, x_2, \cdots, x_n)^T$ 。如果在 V 中可以找到任意多个线性无关的向量, 则称 V 是无限维线性空间, 否则称 V 为有限维线性空间。

 **笔记** 根据第 4 章的讨论, 有几点注意事项:

- 线性空间的基不是唯一的, 但基中所含向量的个数是惟一确定的, 因此维数的定义是有意义的;
- 对于 n 维线性空间 V , V 中任意 n 个线性无关的向量都可作为 V 的基;
- V 中向量 α 用基向量线性表示的式 5.1 是惟一的。这表明在给定基下, 向量 α 的坐标是惟一确定的。

一般地, 令 E_{ij} 是 (i, j) 元素为 1, 其他元素全为零的 $m \times n$ 矩阵, 则下列 $m \times n$ 个矩阵:

$$E_{ij} \quad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$$

构成 $F^{m \times n}$ 的一个基, 从而可知 $F^{m \times n}$ 是 $m \times n$ 维线性空间。

 **练习 5.3** 设线性空间 $V = \{A \in F^{3 \times 3} \mid \text{tr}(A) = 0\}$, 其中 $\text{tr}(A)$ 为 A 的主对角元素之和 (方阵 A 的迹, 详见第 7 章)。求 V 的一组基和维数。

解 任取 $A \in V$, 设:

$$A = \begin{bmatrix} x_1 & a_{12} & a_{13} \\ a_{21} & x_2 & a_{23} \\ a_{31} & a_{32} & x_3 \end{bmatrix}$$

利用矩阵的加法性质有:

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix} + \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix} = B + C$$

B 可由 $E_{ij}(i=1,2,3, i \neq j)$ 线性表示; 而 C 的对角元满足:

$$x_1 + x_2 + x_3 = 0$$

其基础解系为:

$$\xi_1 = (-1, 1, 0)^T, \xi_2 = (-1, 0, 1)^T$$

故 C 可由 $\text{diag}(-1, 1, 0), \text{diag}(-1, 0, 1)$ 线性表示。

从而 V 的一组基为 $\text{diag}(-1, 1, 0), \text{diag}(-1, 0, 1)$ 和 $E_{ij}(i, j=1, 2, 3, i \neq j)$, 且 $\dim(V) = 8$

5.1.4 基变换与坐标变换

上面一部分指出, 一个线性空间的基是不唯一的。线性代数研究的一大问题便是考虑不同基下向量的坐标变换规则, 先给出过渡矩阵的定义:

定义 5.4 (过渡矩阵)

设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\beta_1, \beta_2, \dots, \beta_n$ 是 n 维线性空间 V 的两个基, 第 2 个基可由第 1 个基线性表示为:

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \dots + a_{n1}\alpha_n \\ \beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \dots + a_{n2}\alpha_n \\ \vdots \\ \beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \dots + a_{nn}\alpha_n \end{cases}$$

a_{ij} 为常数 ($i, j=1, 2, \dots, n$), 称矩阵 $A = (a_{ij})_{n \times n}$ 为由基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 到基 $\beta_1, \beta_2, \dots, \beta_n$ 的过渡矩阵。

利用矩阵乘法, 可将上式形式地写成:

$$[\beta_1 \ \beta_2 \ \dots \ \beta_n] = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (5.2)$$

注 注意这里的 α_i, β_i 都是一般线性空间中的向量, 因此 $[\alpha_1 \ \alpha_2 \ \dots \ \alpha_n]$ 及 $[\beta_1 \ \beta_2 \ \dots \ \beta_n]$ 未必是矩阵。

命题 5.2

矩阵 $A = (a_{ij})_{n \times n}$ 为由基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 到基 $\beta_1, \beta_2, \dots, \beta_n$ 的过渡矩阵; 则矩阵 A^{-1} 为由基 $\beta_1, \beta_2, \dots, \beta_n$ 到基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 的过渡矩阵。

过渡矩阵不仅仅表述了两个不同的基之间的变换, 更为坐标变换提供了方便, 有定理 5.3

定理 5.3 (坐标变换公式)

设 n 维线性空间 V 有两个基:

$$(I) : \alpha_1, \alpha_2, \dots, \alpha_n; \quad (II) : \beta_1, \beta_2, \dots, \beta_n$$

且由基 (I) 到基 (II) 的过渡矩阵为 A , 设 V 中向量 α 在基 (I) 的坐标为 $x = (x_1, x_2, \dots, x_n)^T$, α 在基 (II) 下的坐标为 $y = (y_1, y_2, \dots, y_n)^T$, 则有坐标变换公式:

$$x = Ay \quad \text{或} \quad y = A^{-1}x \quad (5.3)$$

练习 5.4 已知 $\mathbb{R}^3$ 有两个基:

$$(I) : \alpha_1 = (1, 1, 1)^T, \alpha_2 = (1, 0, -1)^T, \alpha_3 = (1, 0, 1)^T$$

$$(II) : \beta_1 = (1, 2, 1)^T, \beta_2 = (2, 3, 4)^T, \beta_3 = (3, 4, 3)^T$$

求向量 $\alpha = \alpha_1 + 2\alpha_2 - \alpha_3$ 在基 (II) 下的坐标。

解 解 (1) 设由基 (I) 到基 (II) 的过渡矩阵为 A , 则有:

$$[\beta_1 \ \beta_2 \ \beta_3] = [\alpha_1 \ \alpha_2 \ \alpha_3] A$$

两端左乘 $[\alpha_1 \ \alpha_2 \ \alpha_3]^{-1}$, 得:

$$A = [\alpha_1 \ \alpha_2 \ \alpha_3]^{-1} [\beta_1 \ \beta_2 \ \beta_3] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

已知向量 α 在基 (I) 下的坐标为 $x = (1, 2, -1)^T$ 。设 α 在基 (II) 下的坐标为 $y = (y_1, y_2, y_3)^T$, 则由式 5.3 得:

$$y = A^{-1}x = \begin{bmatrix} 2 & 3 & 4 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -3 \\ -4 \\ 5 \end{bmatrix}$$

5.1.5 * 线性空间的同构

在线性空间 V 中取定一个基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 之后, V 中每个向量在这个基下都有确定的坐标。向量的坐标是 F^n 中的向量, 反过来, 任给 F^n 中一个向量 $x = (x_1, x_2, \dots, x_n)^T$, 则由此可惟一确定 V 中一个向量 $\alpha = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n$, 其中 α 以 x 为其坐标。这样, 就在 V 的向量与它的坐标之间建立了一一对应的关系。 V 的向量与它的坐标之间的对应, 实际上就是 V 到 F^n 的一个映射, 根据前面的说明, 这个映射是 V 到 F^n 的一一映射。

根据这些讨论, 对 V 的性质讨论就可以转化为对数域 F^n 的讨论, 而我们引入同构的概念:

定义 5.5 (同构映射)

设 V_1 和 V_2 是数域 F 上两个线性空间, f 为 V_1 到 V_2 的一个映射, 如果

1. f 是 V_1 到 V_2 的双射;
2. $\forall \alpha, \beta \in V_1$, 恒有 $f(\alpha + \beta) = f(\alpha) + f(\beta)$;
3. $\forall \alpha \in V_1, k \in F$, 恒有 $f(k\alpha) = kf(\alpha)$,

则称映射 f 为同构映射。如果两个线性空间 V_1 与 V_2 之间可以建立一个同构映射, 那么就称 V_1 与 V_2 同构。

对于同构映射, 有性质:

定理 5.4 (线性空间同构的性质)

设 f 是线性空间 V_1 到 V_2 的同构映射, 则:

1. $f(\mathbf{0}_1) = \mathbf{0}_2$, 其中 $\mathbf{0}_1$ 和 $\mathbf{0}_2$ 分别是 V_1 和 V_2 的零向量;
2. $\forall \alpha \in V_1$, 有 $f(-\alpha) = -f(\alpha)$;
3. $\forall \alpha_i \in V_1, k_i \in F (i = 1, 2, \dots, m)$, 有 $f(\sum_{i=1}^m k_i \alpha_i) = \sum_{i=1}^m k_i f(\alpha_i)$;
4. V_1 中的向量组 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关 $\Leftrightarrow$ 它们的像 $f(\alpha_1), f(\alpha_2), \dots, f(\alpha_m)$ 线性相关。

同构作为线性空间之间的关系, 还具有下列简单性质:

- 自反性: V_1 与 V_1 同构;
- 对称性: 若 V_1 与 V_2 同构, 则 V_2 与 V_1 同构;
- 传递性: 若 V_1 与 V_2 同构, V_2 与 V_3 同构, 则 V_1 与 V_3 同构。

定理 5.5 (有限维线性空间同构的充要条件)

数域 F 上两个有限维线性空间同构的充要条件是它们的维数相同。

证明 如果 V_1 与 V_2 同构, 即存在 V_1 到 V_2 的同构映射 f , 因为维数就是空间中线性无关向量的最大个数, 而定理 5.4 的 (4) 表明, V_1 中的向量组线性相关 (线性无关), 当且仅当这组向量在 f 下的像构成的向量组 (作为 V_2 中的

向量组) 线性相关 (线性无关), 所以可知 $\dim(V_1) = \dim(V_2)$ 。反过来, 如果 $\dim(V_1) = \dim(V_2) = n$, 则 V_1 与 F^n 同构, V_2 也与 F^n 同构, 由同构的对称性和传递性, 即知 V_1 与 V_2 同构。

定理 5.5 表明了维数这一性质对于有限维线性空间的本质描述。

定理 5.6 (扩充定理)

设 $\alpha_1, \alpha_2, \dots, \alpha_r$ 是 n 维线性空间 V 中一个线性无关向量组, 且 $r < n$, 则由 $\alpha_1, \alpha_2, \dots, \alpha_r$ 扩充必能得到 V 的基, 也就是说, 必能找到 V 中的向量 $\alpha_{r+1}, \dots, \alpha_n$, 使 $\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n$ 成为 V 的基。

证明 若 $r < n$, 用反证法必存在一向量 $\alpha_{r+1} \in V$, 使得 $\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}$ 线性无关。

若对任意的非零向量 $\beta \in V$, 且 $\alpha_1, \alpha_2, \dots, \alpha_r, \beta$ 线性相关。则存在不全为 0 的数 $k_1, k_2, \dots, k_r, k$, 使得 $k_1\alpha_1 + k_2\alpha_2 + \dots + k_r\alpha_r + k\beta = 0$ 。显然 $k \neq 0$, 所以有:

$$\beta = -\frac{k_1}{k}\alpha_1 - \frac{k_2}{k}\alpha_2 - \dots - \frac{k_r}{k}\alpha_r$$

由此得到, V 任意向量均能被 $\alpha_1, \alpha_2, \dots, \alpha_r$ 线性表示, 所以 $\alpha_1, \alpha_2, \dots, \alpha_r$ 是 V^n 的基, 这与 $r < n$ 矛盾。故若 $r < n$, 必存在一向量 $\alpha_{r+1} \in V^n$, 使得 $\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}$ 线性无关。若 $r+1 = n$, 则证明完毕; 若 $r+1 < n$, 则重复上面过程即可。

5.1.6 * 子空间的交与和

本部分来讨论线性子空间的交与和, 先给出定义:

定义 5.6 (子空间的交与和)

如果 V_1, V_2 是线性空间 V 的两个子空间, 则定义它们的交:

$$V_1 \cap V_2 = \{\alpha | \alpha \in V_1 \text{ 且 } \alpha \in V_2\}$$

定义它们的和:

$$V_1 + V_2 = \{\alpha + \beta | \alpha \in V_1, \beta \in V_2\}$$

子空间的交与和的具有封闭性, 也就是说它们的交与和仍然是子空间, 具体有:

定理 5.7

如果 V_1, V_2 是线性空间 V 的两个子空间, 则 $V_1 \cap V_2$ 与 $V_1 + V_2$ 也是 V 的子空间。

证明 只需要借助定理 5.2 即证, 读者可自行具体给出证明。

定理 5.8 (维数公式)

设如果 V_1, V_2 是线性空间 V 的两个子空间, 则:

$$\dim(V_1) + \dim(V_2) = \dim(V_1 \cap V_2) + \dim(V_1 + V_2) \quad (5.4)$$

证明 证明设 $\dim(V_1 \cap V_2) = r$, $\dim(V_1) = m$, $\dim(V_2) = k$, 又设 $\alpha_1, \dots, \alpha_r$ 是 $V_1 \cap V_2$ 的基。由定理 5.6 知, 可将 $\alpha_1, \dots, \alpha_r$ 扩充成 V_1 的基:

$$\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}$$

同理可将 $\alpha_1, \dots, \alpha_r$ 扩充成 V_2 的基:

$$\alpha_1, \dots, \alpha_r, \xi_1, \dots, \xi_{k-r}$$

则有:

$$V_1 + V_2 = \text{span}\{\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}, \xi_1, \dots, \xi_{k-r}\}$$

下证 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}, \xi_1, \dots, \xi_{k-r}$ 线性无关。设：

$$x_1\alpha_1 + \dots + x_r\alpha_r + y_1\beta_1 + \dots + y_{m-r}\beta_{m-r} + z_1\xi_1 + \dots + z_{k-r}\xi_{k-r} = 0$$

则有：

$$x_1\alpha_1 + \dots + x_r\alpha_r + y_1\beta_1 + \dots + y_{m-r}\beta_{m-r} = -z_1\xi_1 - \dots - z_{k-r}\xi_{k-r} \in V_1 \cap V_2$$

所以 $-z_1\xi_1 - \dots - z_{k-r}\xi_{k-r}$ 能被 $V_1 \cap V_2$ 的基线性表示，即有：

$$-z_1\xi_1 - \dots - z_{k-r}\xi_{k-r} = l_1\alpha_1 + \dots + l_r\alpha_r$$

由此得到 $l_1\alpha_1 + \dots + l_r\alpha_r + z_1\xi_1 + \dots + z_{k-r}\xi_{k-r} = 0$ ，又因为 $\alpha_1, \dots, \alpha_r, \xi_1, \dots, \xi_{k-r}$ 线性无关，所以系数全为 0，即： $l_1 = \dots = l_r = z_1 = \dots = z_{k-r} = 0$ 。故有：

$$x_1\alpha_1 + \dots + x_r\alpha_r + y_1\beta_1 + \dots + y_{m-r}\beta_{m-r} = 0$$

又因为 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}$ 线性无关，所以有：

$$x_1 = \dots = x_r = y_1 = \dots = y_{m-r} = 0$$

由此得到 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}, \xi_1, \dots, \xi_{k-r}$ 线性无关，即 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{m-r}, \xi_1, \dots, \xi_{k-r}$ 是 $V_1 + V_2$ 的基。所以有： $\dim(V_1 + V_2) = r + (m-r) + (k-r) = m + k - r = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$ 证毕。

练习 5.5 设 $\alpha_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}$, $\alpha_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\beta_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, $\beta_2 = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 7 \end{bmatrix}$ ，令 $W_1 = \text{span}\{\alpha_1, \alpha_2\}$, $W_2 = \text{span}\{\beta_1, \beta_2\}$ 。试求

$W_1 + W_2$ 与 $W_1 \cap W_2$ 的维数与基。

解 因为 $W_1 + W_2 = \text{span}\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ，所以只要求出向量组 $\alpha_1, \alpha_2, \beta_1, \beta_2$ 的极大线性无关组即可。

$$[\alpha_1, \alpha_2, \beta_1, \beta_2] = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & -1 & -1 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

所以 $\dim(W_1 + W_2) = 3$ ， $\alpha_1, \alpha_2, \beta_1$ 是 $W_1 + W_2$ 的基。对任意的 $\xi \in W_1 \cap W_2$ ，则 ξ 既可写成向量组 α_1, α_2 的线性表示，也可写成向量组 β_1, β_2 的线性表示。设 $\xi = x_1\alpha_1 + x_2\alpha_2 = y_1\beta_1 + y_2\beta_2$ 由此得到方程组：

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = y_1 \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix} + y_2 \begin{bmatrix} 1 \\ -1 \\ 3 \\ 7 \end{bmatrix}$$

解上齐次线性方程组，得到基础解系：

$$\begin{bmatrix} x_1 & x_2 & y_1 & y_2 \end{bmatrix}^T = \begin{bmatrix} -1 & 4 & -3 & 1 \end{bmatrix}^T$$

将上基础解系代入 $\xi = x_1\alpha_1 + x_2\alpha_2 = y_1\beta_1 + y_2\beta_2$ ，即可得到 $W_1 \cap W_2$ 的一个基 $\begin{bmatrix} -5 & 2 & 3 & 7 \end{bmatrix}^T$ 。所以 $\dim(W_1 \cap W_2) = 1$

子空间的和的一个重要特殊情形便是和元素的表示法唯一，对于这种情况给出定义：

定义 5.7 (子空间的直和)

设 V_1, V_2 都是线性空间 V 的子空间，如果和 $V_1 + V_2$ 中每个向量 α 的表示式：

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in V_1, \alpha_2 \in V_2)$$

是惟一的，则称这个和为 V_1 与 V_2 的直和，并用 $V_1 \oplus V_2$ 表示它。

定理 5.9 (直和的充要条件) $V_1 + V_2$ 为直和 $\Leftrightarrow V_1 \cap V_2 = \{0\}$ 

证明 先证明必要性：设 $V_1 + V_2$ 为直和，即表示式惟一。任取 $\alpha \in V_1 \cap V_2$ ，则 $\alpha \in V_1$ ，且 $\alpha \in V_2$ ，因为：

$$\alpha + 0 = 0 + \alpha \in V_1 + V_2$$

由表示式的惟一性，得 $\alpha = 0$ ，所以 $V_1 \cap V_2 = \{0\}$

再证明充分性：设 $V_1 \cap V_2 = \{0\}$ ，任取 $V_1 + V_2$ 中的向量：

$$\alpha = \alpha_1 + \beta_1, \quad \text{其中 } \alpha_1 \in V_1, \beta_1 \in V_2$$

如果还有：

$$\alpha = \alpha_2 + \beta_2, \quad \text{其中 } \alpha_2 \in V_1, \beta_2 \in V_2$$

两式相减，得：

$$\alpha_1 - \alpha_2 = \beta_2 - \beta_1 \in V_1 \cap V_2 = \{0\}$$

因此有：

$$\alpha_1 = \alpha_2, \quad \beta_1 = \beta_2$$

可见 α 的表示式是惟一的，故 $V_1 + V_2$ 为直和。

结合维数公式，可得推论： $V_1 + V_2$ 为直和 $\Leftrightarrow \dim(V_1) + \dim(V_2) = \dim(V_1 + V_2)$

练习 5.6 设 $A \in F^{n \times n}$ 满足 $A^2 = 2A$ ，且 V_1, V_2 分别是方程 $Ax = 0$ 与 $(2I - A)x = 0$ 的解空间，证明：

$$F^n = V_1 \oplus V_2$$

证明 需要证明两个部分： $F^n = V_1 + V_2$ 且 $V_1 \cap V_2 = \{0\}$ 。

先证明 F^n 是 V_1 和 V_2 的和。对于任意向量 $x \in F^n$ 可以分解为：

$$x = (x - \frac{1}{2}Ax) + \frac{1}{2}Ax$$

验证第一部分属于 V_1 ，即：

$$A(x - \frac{1}{2}Ax) = Ax - \frac{1}{2}A^2x = Ax - \frac{1}{2} \cdot 2Ax = Ax - Ax = 0$$

所以 $x - \frac{1}{2}Ax \in V_1$ ；再验证第二部分属于 V_2 ，即：

$$(2I - A)(\frac{1}{2}Ax) = Ax - \frac{1}{2}A^2x = Ax - \frac{1}{2} \cdot 2Ax = Ax - Ax = 0$$

所以 $\frac{1}{2}Ax \in V_2$ 。因此任意向量都能表示为 V_1 和 V_2 中向量的和。

再来证明 $V_1 \cap V_2 = \{0\}$ 。设 $x \in V_1 \cap V_2$ ，则：

$$Ax = 0$$

$$(2I - A)x = 0$$

由第二个等式得到 $2x = Ax$ ，但 $Ax = 0$ ，所以 $x = 0$ 。因此 $V_1 \cap V_2 = \{0\}$ 证毕。

5.2 欧式空间

5.2.1 内积

中学时期，我们接触了欧式空间的一些基本概念。本小节为了给出欧式空间的定义，先引入内积的概念：

定义 5.8 (内积和 Euclid 空间)

设 V 是一个实线性空间，如果对于 V 中任意两个向量 α 和 β 都指定了一个实数与之对应，这个实数记为 $\langle \alpha, \beta \rangle$ ，并且满足以下条件 (称之为内积公理)：

1. 对称性： $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$
2. 加性： $\langle \alpha + \beta, \gamma \rangle = \langle \alpha, \gamma \rangle + \langle \beta, \gamma \rangle$
3. 齐性： $k\langle \alpha, \beta \rangle = \langle k\alpha, \beta \rangle$
4. 非负性： $\langle \alpha, \alpha \rangle \geq 0$ ，而且 $\langle \alpha, \alpha \rangle = 0 \Leftrightarrow \alpha = \mathbf{0}$

其中， α, β 和 γ 是 V 中任意向量， k 是任意实数，则称实数 $\langle \alpha, \beta \rangle$ 为 α 与 β 的内积。称定义了内积的实线性空间 V 为实内积空间或欧几里得 (Euclid) 空间，简称为欧氏空间。

利用定义中的四条内积公理，我们可以快速判断一个线性空间是否是欧式空间：

 **练习 5.7** 验证在实数域上所有 n 阶方阵构成的线性空间 $F^{n \times n}$ 中，定义两个方阵 A 和 B 的内积：

$$\langle A, B \rangle = \text{tr}(AB^T)$$

那么该线性空间是一个欧式空间。

证明 根据内积公理，可以验证：

1. 对称性： $\langle A, B \rangle = \text{tr}(AB^T) = \text{tr}((B^T A)^T) = \text{tr}(BA^T) = \langle B, A \rangle$
2. 加性： $\langle A + B, C \rangle = \text{tr}((A + B)C^T) = \text{tr}(AC^T + BC^T) = \text{tr}(AC^T) + \text{tr}(BC^T) = \langle A, C \rangle + \langle B, C \rangle$
3. 齐性： $\langle kA, B \rangle = \text{tr}((kA)B^T) = k\text{tr}(AB^T) = k\langle A, B \rangle$
4. 非负性： $\langle A, A \rangle = \text{tr}(AA^T)$ ，因为 AA^T 的主对角线元素都是非负的，所以 $\text{tr}(AA^T) \geq 0$ ，当且仅当 $A = O$ 时， $\text{tr}(AA^T) = 0$

因此，这就构成了一个欧式空间。

事实上，对于同一个线性空间，定义不同有效的内积，就可以得到不同的欧式空间。因此，讲某个欧氏空间 V ，一定是指实线性空间 V 连同定义在 V 上的某个内积这个整体而言的。特别地，以后我们说欧氏空间 $\mathbb{R}^n, \mathbb{R}^{n \times n}$ 和 $C[a, b]$ 时，它们的内积分别如下定义：

定义 5.9 ($\mathbb{R}^n, \mathbb{R}^{n \times n}$ 和 $C[a, b]$ 的标准内积)

对于线性空间 $\mathbb{R}^n, \mathbb{R}^{n \times n}$ 和 $C[a, b]$ ，下面三个内积式分别是其定义的标准内积。

$$\langle \alpha, \beta \rangle = a_1 b_1 + \cdots + a_n b_n = \alpha^T \beta \quad \alpha, \beta \in \mathbb{R}^n \quad (5.5)$$

$$\langle A, B \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij} \quad A, B \in \mathbb{R}^{n \times n} \quad (5.6)$$

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx \quad f, g \in C[a, b] \quad (5.7)$$

定理 5.10 (Cauchy-Schwarz 不等式)

设 α 和 β 是欧氏空间 V 中两个任意向量，则有不等式：

$$|\langle \alpha, \beta \rangle| \leq \sqrt{\langle \alpha, \alpha \rangle} \sqrt{\langle \beta, \beta \rangle} \quad (5.8)$$

其中等号成立的充要条件是 α 与 β 线性相关。

证明 如果 α 与 β 线性无关 (此时必有 $\alpha \neq 0$)，则对任意实数 t , $t\alpha + \beta \neq 0$ ，由内积的非负性便得：

$$\langle t\alpha + \beta, t\alpha + \beta \rangle > 0$$

即有：

$$t^2 \langle \alpha, \alpha \rangle + 2t \langle \alpha, \beta \rangle + \langle \beta, \beta \rangle > 0$$

上式左端是 t 的二次三项式，它恒正，由中学代数的知识知道它的判别式必小于零，有：

$$\langle \alpha, \beta \rangle^2 - \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle < 0$$

因此：

$$|\langle \alpha, \beta \rangle| < \sqrt{\langle \alpha, \alpha \rangle} \sqrt{\langle \beta, \beta \rangle}$$

如果 α 与 β 线性相关，则 α 与 β 成比例，不妨设 $\alpha = k\beta$ ($k \in \mathbb{R}$)，于是有

$$\begin{aligned} |\langle \alpha, \beta \rangle| &= |\langle k\beta, \beta \rangle| = |k \langle \beta, \beta \rangle| = \sqrt{k^2 \langle \beta, \beta \rangle} \sqrt{\langle \beta, \beta \rangle} \\ &= \sqrt{\langle k\beta, k\beta \rangle} \sqrt{\langle \beta, \beta \rangle} = \sqrt{\langle \alpha, \alpha \rangle} \sqrt{\langle \beta, \beta \rangle} \end{aligned}$$

对于定义5.9中描述的 $\mathbb{R}^n$, $\mathbb{R}^{n \times n}$ 和 $C[a, b]$ 的标准内积，可以有 Cauchy-Schwarz 不等式的一些常用特殊形式：

$$|a_1 b_1 + \cdots + a_n b_n| \leq \sqrt{a_1^2 + \cdots + a_n^2} \sqrt{b_1^2 + \cdots + b_n^2} \quad (5.9)$$

$$|w_1 a_1 b_1 + \cdots + w_n a_n b_n| \leq \sqrt{w_1 a_1^2 + \cdots + w_n a_n^2} \sqrt{w_1 b_1^2 + \cdots + w_n b_n^2} \quad (5.10)$$

$$\left| \int_a^b f(x)g(x)dx \right| \leq \left(\int_a^b f^2(x)dx \right)^{\frac{1}{2}} \left(\int_a^b g^2(x)dx \right)^{\frac{1}{2}} \quad (5.11)$$

5.2.2 范数、夹角和距离

在三维几何空间中，向量 $\alpha = (a_1, a_2, a_3)^T$ 的长度为

$$\|\alpha\| = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{\langle \alpha, \alpha \rangle}$$

现在将几何空间中向量长度的概念推广到一般的欧氏空间中：

定义 5.10 (向量的范数)

在欧氏空间 V 中，称非负实数 $\sqrt{\langle \alpha, \alpha \rangle}$ 为向量 α 的范数 (或长度)，记为 $\|\alpha\|$ ，即：

$$\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle} \quad (5.12)$$

按范数的定义，Cauchy-Schwarz 不等式现在可写成：

$$|\langle \alpha, \beta \rangle| \leq \|\alpha\| \|\beta\|$$

向量的范数是由内积定义的，所以对于不同的内积，范数可以不同。但是对于任何内积，由式5.12定义的范数都具有与几何空间中向量长度同样的基本性质，设 α 和 β 是欧氏空间 V 中任意向量， k 是任意实数，那么有：

1. $\|\alpha\| \geq 0$ ，而且 $\|\alpha\| = 0$ 当且仅当 $\alpha = 0$ ；
2. $\|k\alpha\| = |k| \|\alpha\|$ ；
3. $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$

范数为 1 的向量称为**单位向量**。若向量 $\alpha \neq 0$ ，则知向量 $\frac{1}{\|\alpha\|} \alpha$ 就是一个单位向量。由非零向量 α 得到单位向量 $\frac{1}{\|\alpha\|} \alpha$ ，称为把向量 α **单位化**。后面我们还会介绍把一组基转化为单位正交基的方法。

在几何空间中，向量 α 与 β 的内积定义为：

$$\langle \alpha, \beta \rangle = \|\alpha\| \|\beta\| \cos \varphi$$

其中 φ 为 α 与 β 的夹角, 因此可用内积来表示非零向量 α 与 β 的夹角 φ 为:

$$\varphi = \arccos \frac{\langle \alpha, \beta \rangle}{\|\alpha\| \|\beta\|} \quad (0 \leq \varphi \leq \pi)$$

在一般的欧氏空间 V 中, 知对任何非零向量 α 和 β 都有:

$$\left| \frac{\langle \alpha, \beta \rangle}{\|\alpha\| \|\beta\|} \right| \leq 1$$

这就使得我们可以将几何空间中非零向量夹角的概念推广到 V 中:

定义 5.11 (向量的夹角和正交)

在欧氏空间中, 规定两个非零向量 α 与 β 的夹角 φ 为

$$\varphi = \arccos \frac{\langle \alpha, \beta \rangle}{\|\alpha\| \|\beta\|} \quad (0 \leq \varphi \leq \pi) \quad (5.13)$$

如果 $\langle \alpha, \beta \rangle = 0$, 则称 α 与 β 正交或相互垂直, 记为 $\alpha \perp \beta$ 。

在内积公理的齐性中取数 $k = 0$, 便有 $\langle 0, \beta \rangle = 0$, 所以零向量与任何向量都正交。当两个非零向量 α 与 β 正交时, 由定义知 $\varphi = \frac{\pi}{2}$ 。可见在一般的欧氏空间中, 两个非零向量正交与几何空间中两个非零向量正交的概念是一致的。当 $\alpha \perp \beta$ 时, 有:

$$\|\alpha + \beta\|^2 = \langle \alpha + \beta, \alpha + \beta \rangle = \|\alpha\|^2 + \|\beta\|^2$$

这就是勾股定理。不难把勾股定理推广到多个向量的情形, 而类似的给出距离的定义:

定义 5.12 (距离)

对于欧氏空间 V 中两个向量 α 和 β , 称范数 $\|\alpha - \beta\|$ 为 α 与 β 的距离, 记为 $d(\alpha, \beta)$, 或者:

$$d(\alpha, \beta) = \|\alpha - \beta\|$$

由范数的基本性质知距离满足对称性, 非负性和三角不等式。

5.2.3 标准正交基

定义 5.13 (正交向量组和标准正交向量组)

对于欧氏空间 V 中的一个向量组, 如果其中不含零向量, 且其中的向量两两正交, 则称它为一个正交向量组。如果一个正交向量组中的每个向量都是单位向量, 则称它为一个正交单位向量组 (或标准正交向量组, 或正交规范向量组)。

 **笔记** 一组向量是正交向量组需要满足: 1. 不含零向量, 2. 两两正交。在此基础上, 所有向量都是单位向量时, 则变为正交单位向量组。

例如, 在欧氏空间 $\mathbb{R}^3$ 中, 向量组:

$$\alpha_1 = (1, 1, 1)^T, \quad \alpha_2 = (1, 0, -1)^T, \quad \alpha_3 = (1, -2, 1)^T$$

是一个正交向量组. 而向量组:

$$\begin{aligned} \beta_1 &= \frac{1}{\|\alpha_1\|} \alpha_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)^T, \\ \beta_2 &= \frac{1}{\|\alpha_2\|} \alpha_2 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right)^T, \\ \beta_3 &= \frac{1}{\|\alpha_3\|} \alpha_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right)^T \end{aligned}$$

是一个正交单位向量组。这说明一组正交向量组可以通过单位化转化为正交单位向量组

正交和线性无关是描述向量组的两个易混概念，而下面的定理阐述了它们的关系：

定理 5.11

正交向量组线性无关。

证明 设 $\alpha_1, \dots, \alpha_m$ 是一个正交向量组，则：

$$\langle \alpha_i, \alpha_j \rangle = \begin{cases} \|\alpha_i\|^2 > 0, & \text{当 } j = i \\ 0, & \text{当 } j \neq i \end{cases}$$

设有一组数 $k_1, \dots, k_m$ ，使：

$$k_1\alpha_1 + \dots + k_m\alpha_m = 0$$

用 α_i 与上式两边作内积，由于当 $j \neq i$ 时有 $\langle \alpha_i, \alpha_j \rangle = 0$ ，故得

$$k_i \langle \alpha_i, \alpha_i \rangle = \langle \alpha_i, 0 \rangle = 0$$

由 $\langle \alpha_i, \alpha_i \rangle > 0$ ，得 $k_i = 0$ ($i = 1, \dots, m$)，所以 $\alpha_1, \dots, \alpha_m$ 线性无关。

在 n 维欧氏空间 V 中，由 n 个向量组成的正交向量组称为 V 的正交基，由 n 个向量组成的正交单位向量组称为 V 的**标准正交基** (或规范正交基)。标准正交基为讨论欧氏空间的问题带来许多方便。下面的定理表明，在标准正交基下向量的坐标、内积、范数及距离的计算都变得非常容易：

定理 5.12 (标准正交基的性质)

设 $\alpha_1, \dots, \alpha_n$ 是 n 维欧氏空间 V 的一个标准正交基， α 和 β 是 V 中任意的向量，设：

$$\alpha = x_1\alpha_1 + \dots + x_n\alpha_n \quad \beta = y_1\alpha_1 + \dots + y_n\alpha_n$$

则有：

1. $x_i = \langle \alpha, \alpha_i \rangle$ ($i = 1, \dots, n$)，换言之有：

$$\alpha = \langle \alpha, \alpha_1 \rangle \alpha_1 + \dots + \langle \alpha, \alpha_n \rangle \alpha_n$$

2. $\langle \alpha, \beta \rangle = x_1y_1 + \dots + x_ny_n$

3. $\|\alpha\| = \sqrt{x_1^2 + \dots + x_n^2}$

4. $d(\alpha, \beta) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$

由上述定理还可以说明任何一个 n 维欧氏空间 V 与 n 维欧氏空间 $\mathbb{R}^n$ 是同构关系。

练习 5.8 在欧氏空间 $\mathbb{R}^3$ 中，求向量 $\alpha = (2, 3, 1)^T$ 在下述标准正交基的坐标。

$$\alpha_1 = (0, 1, 0)^T \quad \alpha_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)^T \quad \alpha_3 = \left(\frac{3}{5}, 0, \frac{4}{5}\right)^T$$

解 因为 $\langle \alpha, \alpha_1 \rangle = 3$ ， $\langle \alpha, \alpha_2 \rangle = -1$ ， $\langle \alpha, \alpha_3 \rangle = 2$ ，故由定理 5.12 知 α 在这个基下的坐标为 $(3, -1, 2)^T$ 。

5.2.4 Gram-Schmidt 正交化

前面已经强调标准正交基对研究欧式空间的重要作用，本节介绍由一个基转化为一个标准正交基的方法：

定理 5.13 (Gram-Schmidt 正交化过程)

设欧氏空间 V 有一组线性无关向量组 $\alpha_1, \alpha_2, \dots, \alpha_m$ ，则令：

$$\begin{aligned} \beta_1 &= \alpha_1 \\ \beta_2 &= \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 \\ \beta_3 &= \alpha_3 - \frac{\langle \alpha_3, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 - \frac{\langle \alpha_3, \beta_2 \rangle}{\langle \beta_2, \beta_2 \rangle} \beta_2 \end{aligned}$$

$$\beta_m = \alpha_m - \sum_{k=1}^{m-1} \frac{\langle \alpha_m, \beta_k \rangle}{\langle \beta_k, \beta_k \rangle} \beta_k \quad (5.14)$$

就可以得到 V 的一组正交向量组 $\beta_1, \beta_2, \dots, \beta_m$, 再对该向量组分别单位化, 即:

$$e_j = \frac{1}{\|\beta_j\|} \beta_j$$

就可以得到 V 的一组标准正交基 $e_1, e_2, \dots, e_m$ 。该过程称为 Gram-Schmidt 正交化过程。



笔记 Gram-Schmidt 正交化过程对于求欧式空间的标准正交基很有用, 读者需牢记。



练习 5.9 设 $\mathbb{R}^4$ 的子空间 W 由向量组:

$$\alpha_1 = (1, 2, 2, -1)^T \quad \alpha_2 = (1, 1, -5, 3)^T \quad \alpha_3 = (3, 2, 8, -7)^T$$

生成, 试利用 Gram-Schmidt 方法求 W 的一个标准正交基。

解 不难验证向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, 因此它就是 W 的基。将其先正交化, 即令

$$\begin{aligned} \beta_1 &= \alpha_1 = (1, 2, 2, -1)^T \\ \beta_2 &= \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 \\ &= (1, 1, -5, 3)^T - \frac{-10}{10} (1, 2, 2, -1)^T \\ &= (2, 3, -3, 2)^T \\ \beta_3 &= \alpha_3 - \frac{\langle \alpha_3, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 - \frac{\langle \alpha_3, \beta_2 \rangle}{\langle \beta_2, \beta_2 \rangle} \beta_2 \\ &= (3, 2, 8, -7)^T - \frac{30}{10} (1, 2, 2, -1)^T - \frac{-26}{26} (2, 3, -3, 2)^T \\ &= (2, -1, -1, -2)^T \end{aligned}$$

再对其单位化, 即令 $e_i = \frac{1}{\|\beta_i\|} \beta_i (i = 1, 2, 3)$, 便得到 W 的标准正交基:

$$e_1 = \frac{1}{\sqrt{10}} (1, 2, 2, -1)^T \quad e_2 = \frac{1}{\sqrt{26}} (2, 3, -3, 2)^T \quad e_3 = \frac{1}{\sqrt{10}} (2, -1, -1, -2)^T$$

5.2.5 正交矩阵

现在来介绍一种非常有用的矩阵:

定义 5.14 (正交矩阵)

若实方阵 A 满足 $A^T A = A A^T = I$ (等价于 $A^{-1} = A^T$), 则称 A 为正交矩阵。

设 A, B 为同阶正交矩阵, 不难证明一些有用的性质:

1. $\det(A) = \pm 1$, 即正交矩阵的行列式为 1 或 -1;
2. A^T, A^{-1} 及 A 的伴随矩阵 A^* 均为正交矩阵;
3. AB 为正交矩阵。

定理 5.14 (正交矩阵的判定)

实方阵 A 为正交矩阵的充要条件是 A 的列(行)向量组为标准正交向量组。

证明 设实方阵 A 按列分块为 $A = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]$, 则 $A^T A = I$ 就是说明:

$$\begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T \end{bmatrix} [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

也可以表示为:

$$\alpha_i^T \alpha_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (i, j = 1, 2, \cdots, n)$$

可见, A 为正交矩阵当且仅当 A 的列向量组为标准正交向量组。再由 A^T 为正交矩阵, 知这个结论对 A 的行向量组也成立, 因此证毕。

定义 5.15 (正交变换)

设 P 为 n 阶正交矩阵, 对于 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$, 称 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性变换 $T: T(x) = Px$ 为 $\mathbb{R}^n$ 上的正交变换。设 P 为 n 阶正交矩阵, x_1, x_2 是 $\mathbb{R}^n$ 中任意向量, 则有:

- $\langle Px_1, Px_2 \rangle = \langle x_1, x_2 \rangle$, 即正交变换保持向量的内积不变;
- $\|Px_1\| = \|x_1\|$, 即正交变换保持向量的范数不变。

笔记 根据定义 5.15 的第二条性质, 可知几何空间中的正交变换是一种保持距离不变的变换, 称等距变换。线性变换的具体概念将在下一章中介绍。

5.2.4 部分表明用 Gram-Schmidt 正交化方法, 可将任何一个线性无关向量组化为标准正交向量组。这样, 任意一个具有线性无关列向量组的矩阵 A , 都可以化成一个其列向量组为标准正交向量组的矩阵 Q , 我们来看矩阵 A 与 Q 的关系。以练习 5.9 所给向量 $\alpha_1, \alpha_2, \alpha_3$ 为例, 设 $A = [\alpha_1 \ \alpha_2 \ \alpha_3]$ 。现在考虑用 e_1, e_2, e_3 来线性表示原来的向量 $\alpha_1, \alpha_2, \alpha_3$, 将 $\beta_i = \|\beta_i\|e_i (i = 1, 2, 3)$ 代入练习 5.9 中正交化过程的各式, 并移项可得:

$$A = [\alpha_1 \ \alpha_2 \ \alpha_3] = [e_1 \ e_2 \ e_3] \begin{bmatrix} \sqrt{10} & -\sqrt{10} & 3\sqrt{10} \\ 0 & \sqrt{26} & -\sqrt{26} \\ 0 & 0 & \sqrt{10} \end{bmatrix}$$

用 Q 与 R 依次表示上式右端的两个矩阵, 则有 $A = QR$, 显然 Q 是一个列向量组为标准正交向量组的矩阵, R 是主对角线元素均非零的上三角形矩阵, 即 R 是一个可逆的上三角形矩阵。一般地, 有:

定理 5.15 (QR 分解)

若 $m \times n$ 矩阵 A 的列向量组线性无关, 则 A 能被分解成两个矩阵的乘积: $A = QR$, 其中 $m \times n$ 矩阵 Q 的列向量组是标准正交向量组, R 是可逆的 n 阶上三角形矩阵。如果 A 是 n 阶可逆方阵, 则 A 的上述分解式中的矩阵 Q 成为 n 阶正交矩阵。分解式 $A = QR$ 称为矩阵 A 的 QR 分解。

求一个矩阵 A 的 QR 分解, 可以按照以下步骤进行, 这里不做强制学习要求:

1. 明确 A 的 QR 分解形式为:

$$A_{m \times n} = Q_{m \times n} R_{n \times n}$$

2. 将 A 的列向量化为一组标准正交向量组, 从而求出 Q ;
3. 根据 $A = QR$, 两边同时左乘 Q^T 有 $R = Q^T A$, 从而求出 R 。

第五章 练习

- (1) 证明在实数域上的线性空间 $\mathbb{R}^n$ 中, 对任意自然数 n , 有 $1, \cos x, \cos 2x, \dots, \cos nx$ 线性无关。
- (2) 设 A 是 $\mathbb{R}$ 上的 n 阶矩阵, 证明: 如果 A 有 n 个不同的特征值, 那么 $\dim C(A) = n$ 。
- (3) 设 V_1, V_2 都是域 F 上线性空间 V 的有限维子空间, 且 $V_1 \subseteq V_2$ 。证明:

$$\dim V_1 \leq \dim V_2; \text{ 且如果 } \dim V_1 = \dim V_2, \text{ 那么 } V_1 = V_2$$

- (4) 设 A 是 $\mathbb{R}^3$ 上的一个线性变换:

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_3 - x_2 \\ x_2 - x_3 \end{bmatrix}$$

求 A 的标准基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵。

- (5) 已知 $\mathbb{R}^3$ 上的线性变换 A 在标准基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵是:

$$A = \begin{bmatrix} 15 & -11 & 5 \\ 20 & -15 & 8 \\ 8 & -7 & 6 \end{bmatrix}$$

求 A 在基 $\eta_1 = (2, 3, 1)^T, \eta_2 = (3, 4, 1)^T, \eta_3 = (1, 2, 2)^T$ 下的矩阵 B 。

- (6) 证明: 如果域 F 上两个 n 阶矩阵 A 与 B 相似, 那么它们可以看成是域 F 上 n 维线性空间 V 上同一个线性变换 T 在 V 的不同基下的矩阵。
- (7) 把 $\mathbb{C}$ 分别看作 $\mathbb{R}$ 和 $\mathbb{C}$ 上的线性空间。令 $T(z) = \bar{z}, \forall z \in \mathbb{C}$, T 是不是 $\mathbb{C}$ 上的线性变换?

第六章 * 线性变换

6.1 线性变换的定义和性质

定义 6.1 (线性变换)

设 $T: V \rightarrow W$ 是从线性空间 V 到线性空间 W 的一个映射, 而且 $\forall \alpha, \beta \in V, k \in F$, 恒成立:

1. $T(\alpha + \beta) = T(\alpha) + T(\beta)$
2. $T(k\alpha) = k(T(\alpha))$

则称 T 是 V 到 W 的一个线性变换 (或线性映射)。 V 到 W 的线性变换全体记为 $L(V, W)$, V 到 V 自身的线性变换全体记为 $L(V)$ 。如果 $T \in L(V)$, 则也称 T 是 V 上的线性算子。



注 同构映射一定是线性变换, 而线性变换不一定是同构映射 (无法保证双射条件)。

下面来介绍几种常见的线性变换, 读者可借助例子加深印象。

1. 定义 $I: V \rightarrow V$

$$I(\alpha) = \alpha \quad \forall \alpha \in V$$

则 I 是 V 上的线性算子, 称 I 是 V 上的恒等变换。

2. 设 k 为数域 F 中的一个数, 定义 $T: V \rightarrow V$

$$T(\alpha) = k\alpha \quad \forall \alpha \in V$$

则 T 是 V 上的线性算子, 称 T 为数乘变换。

3. 定义 $T: V \rightarrow W$

$$T(\alpha) = \mathbf{0} \quad \forall \alpha \in V$$

则有 $T \in L(V, W)$, 称 T 为零变换, 记为 O 。

4. 设 $e_1, e_2, \dots, e_n$ 是 n 维线性空间 V 的一个基。对于 $\alpha \in V$, 设 $\alpha = k_1 e_1 + k_2 e_2 + \dots + k_n e_n$, 定义 $T: V \rightarrow F^n$

$$T(\alpha) = (k_1, k_2, \dots, k_n)^T \quad \forall \alpha \in V$$

则有 $T \in L(V, F^n)$, 称 T 为坐标变换。

5. 设 V 是区间 $[a, b]$ 上具有 1 阶连续导数的 1 元函数全体所组成的线性空间, W 是 $[a, b]$ 上 1 元连续函数全体所组成的线性空间, 定义 $D: V \rightarrow W$

$$D(f(x)) = f'(x) \quad \forall f(x) \in V$$

则有 $T \in L(V, W)$, 称 T 为微商变换。

线性变换满足以下基本性质:

定理 6.1 (线性变换的性质)

设 $T \in L(V, W)$, 则:

1. $T(\mathbf{0}) = \mathbf{0}$;
2. $T(-\alpha) = -T(\alpha)$;
3. $T(k_1 \alpha_1 + \dots + k_m \alpha_m) = k_1 T(\alpha_1) + \dots + k_m T(\alpha_m)$;
4. T 把 V 中线性相关向量组映射为 W 中线性相关向量组。



根据上述性质 (4), 线性变换的本质是保持线性运算, 因此线性变换完全由其对作用空间的基的作用确定。设 $T, S \in L(V, W)$, 如果 $\forall \alpha \in V$, 有 $T(\alpha) = S(\alpha)$, 则称线性变换 T 与 S 相等, 记为 $T = S$ 。由上面的说明可知:

$$T = S \Leftrightarrow T(e_i) = S(e_i) \quad i = 1, \dots, n$$

6.2 线性变换的核与值域

定义 6.2 (核与值域)

设 $T \in L(V, W)$, 则称 V 为 T 的定义域。

称 V 的子集 $\{\alpha | \alpha \in V, T(\alpha) = \mathbf{0}\}$ 为 T 的核或零空间, 记为 $\ker(T)$ 或 $T^{-1}\{\mathbf{0}\}$;

称 W 的子集 $\{T(\alpha) | \alpha \in V\}$ 为 T 的值域或像空间, 记为 $R(T)$ 或 $\text{Im}(T)$ 。

核与值域可以直观表示为如下映射图:

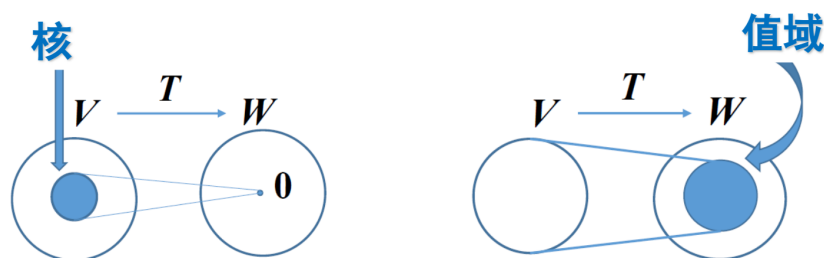


图 6.1: 核与值域直观映射图

定理 6.2

设 $T \in L(V, W)$, 则:

- $\ker(T)$ 是 V 的子空间;
- $R(T)$ 是 W 的子空间。

证明 先证明前半部分, 只要证明 $\ker(T)$ 非空且对 V 的线性运算封闭。因为 $T(\mathbf{0}) = \mathbf{0}$, 故 $\ker(T)$ 非空。 $\forall \alpha, \beta \in \ker(T)$, $k \in F$, 因为:

$$T(\alpha + \beta) = T(\alpha) + T(\beta) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$T(k\alpha) = kT(\alpha) = k\mathbf{0} = \mathbf{0}$$

故 $\alpha + \beta \in \ker(T)$, $k\alpha \in \ker(T)$, 所以 $\ker(T)$ 是 V 的子空间。

对后半部分, $R(T)$ 显然非空。 $\forall T(\alpha), T(\beta) \in R(T)$ (其中 $\alpha, \beta \in V$), $k \in F$, 因为:

$$T(\alpha) + T(\beta) = T(\alpha + \beta) \in R(T)$$

$$kT(\alpha) = T(k\alpha) \in R(T)$$

故 $R(T)$ 对 W 的线性运算封闭。所以 $R(T)$ 是 W 的子空间。

通常, 记线性映射 T 的核的维数为**零度 (零化度)**, 值域的维数为**秩**, 分别表示为:

$$\text{nullity}(T) = \dim(\ker(T)) \quad \text{rank}(T) = \dim(R(T)) \quad (6.1)$$

线性映射的零度和秩满足一重要定理:

定理 6.3 (秩-零化度定理)

对于线性映射 $T: V \rightarrow W$, 设其定义域的维数为 n , 即 $\dim(V) = n$, 则有:

$$\text{nullity}(T) + \text{rank}(T) = n \quad (6.2)$$

证明 设向量空间 V 的维数为 n , 且 $\dim \ker(T) = p, p \leq n$ 。设 $\ker(T)$ 的一组基底为 $\{u_1, \dots, u_p\}$, 并将其扩充为

V 的一组基底 $\{u_1, \dots, u_p, w_1, \dots, w_r\}$, 其中 $n = p + r$. 因此我们只需要证明:

$$\text{rank}(T) = r$$

空间 V 中任一向量 α 都可以表示为:

$$\alpha = a_1 u_1 + \dots + a_p u_p + b_1 w_1 + \dots + b_r w_r$$

使用线性变换 T 作用于 α , 有:

$$\begin{aligned} T(\alpha) &= T(a_1 u_1 + \dots + a_p u_p + b_1 w_1 + \dots + b_r w_r) \\ &= a_1 T(u_1) + \dots + a_p T(u_p) + b_1 T(w_1) + \dots + b_r T(w_r) \\ &= b_1 T(w_1) + \dots + b_r T(w_r) \end{aligned}$$

因为 α 是任意取的, 所以值域 $R(T)$ 可由 $T(w_1), \dots, T(w_r)$ 扩充得到. 接下来, 我们证明它们是线性无关的, 即 $\{T(w_1), \dots, T(w_r)\}$ 构成 $R(T)$ 的一组基底. 考虑:

$$c_1 T(w_1) + \dots + c_r T(w_r) = 0$$

上式可写成

$$T(c_1 w_1 + \dots + c_r w_r) = 0$$

所以有 $c_1 w_1 + \dots + c_r w_r \in \ker(T)$, 其可表示为 $\{u_1, \dots, u_p\}$ 的线性组合

$$c_1 w_1 + \dots + c_r w_r = d_1 u_1 + \dots + d_p u_p$$

又因为 $\{u_1, \dots, u_p, w_1, \dots, w_r\}$ 线性无关, 所以上式中的系数全部为零, 证得 $\{T(w_1), \dots, T(w_r)\}$ 线性无关, 即有 $\text{rank}(T) = r$, 原命题证毕。

 **练习 6.1** 基于秩-零化度定理证明命题6.1:

命题 6.1

设 $T \in L(V, W)$, 若 $\dim(V) > \dim(W)$, 则存在非零向量 $x_0 \in V$ 使得 $T(x_0) = \mathbf{0}$

证明 根据定义和秩-零化度定理有:

$$\dim(\ker(T)) = \dim(V) - \dim(R(T)) \geq \dim(V) - \dim(W) > 0$$

因此 T 的零化度不为零, 则必存在非零向量 $x_0 \in V$ 使得 $T(x_0) = \mathbf{0}$ 证毕。

读者不难举出几个例子验证秩-零化度定理的正确性。下面介绍一些关于线性映射的等价叙述条件。

命题 6.2

设 $T \in L(V, W)$, 则下列条件相互等价:

1. T 是单射;
2. $\ker(T) = \{\mathbf{0}\}$;
3. T 将 V 中线性无关组映射为 W 中线性无关组。

如果 $\dim(V) = n$, 即 V 是一个有限维空间, 则条件 (4) $\text{rank}(T) = n$ 和上述条件分别等价。

命题6.2表明了一个线性映射是单射的多个充要条件。

6.3 线性变换的运算

类似于映射的运算和复合，我们可以引出线性变换的运算。

定义 6.3 (线性映射的加法和数乘)

设 $T_1, T_2, T \in L(V, W)$, $k \in F$, 定义 T_1 与 T_2 的加法为:

$$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha) \quad \forall \alpha \in V$$

称 V 到 W 的映射 $T_1 + T_2$ 为 T_1 与 T_2 的和。定义数 k 与 T 的数量乘法 (简称数乘) 为:

$$(kT)(\alpha) = kT(\alpha) \quad \forall \alpha \in V$$

称 V 到 W 的映射 kT 为 k 与 T 的数量乘积。

定义 6.4 (映射的乘积)

设 $T_2: U \rightarrow V, T_1: V \rightarrow W$ 是两个映射, 定义 $T_1T_2: U \rightarrow W$ 为:

$$T_1T_2(\alpha) = T_1(T_2(\alpha)), \quad \forall \alpha \in U$$

称 U 到 W 的映射 T_1T_2 为 T_1 与 T_2 的乘积 (或复合)。

由定义可知映射的乘积满足结合律, 即:

$$(T_1T_2)T_3 = T_1(T_2T_3)$$

定理 6.4 (线性映射的乘积)

若 T_1, T_2 是线性变换, 则由定义中规定的映射乘积 T_1T_2 是线性变换。

如果 $T \in L(V, W)$, I_V 与 I_W 分别是 V 与 W 上的恒等变换, 则由线性变换乘积的定义可知:

$$TI_V = T, \quad I_W T = T$$

定义 6.5 ((线性) 映射的可逆运算)

设有映射 $T: V \rightarrow W$, 若存在映射 $S: W \rightarrow V$, 使得:

$$TS = I_W \quad ST = I_V$$

则称 T 为可逆映射, 并称 S 为 T 的逆映射, 且记 S 为 T^{-1} , 即有 $TT^{-1} = I_W, T^{-1}T = I_V$ 。如果线性变换 T 是可逆映射, 则称 T 为可逆线性变换。

类似方阵的幂运算, 可以定义 V 到 V 自身的映射 T 的幂为:

$$T^0 = I_V \quad T^1 = T \quad T^{k+1} = T^k T \quad (k = 1, 2, \dots)$$

由映射乘法满足结合律, 有:

$$T^{m+n} = T^m T^n, \quad (T^m)^n = T^{mn} \quad (m, n \text{ 为非负整数})$$

当 T 可逆时, 定义 T 的负整数幂为:

$$T^{-n} = (T^{-1})^n \quad (n \text{ 为正整数})$$

定理 6.5 (线性映射可逆的性质)

设 $T \in L(V, W)$ 是可逆线性变换, 则 T^{-1} 是线性变换。且 T 是可逆的 $\Leftrightarrow \ker(T) = \{0\}$ 且 $R(T) = W$ 。

6.4 线性变换的矩阵表示

设 V 和 W 的维数分别是 n 和 m ，它们的基分别是 B 和 B' ，其中：

$$B = \{\alpha_1, \dots, \alpha_n\} \quad B' = \{\beta_1, \dots, \beta_m\}$$

对于 $T \in L(V, W)$ ，因为基向量的像 $T(\alpha_j)$ 是 W 中的向量，所以 $T(\alpha_j)$ 可由 W 的基 B' 惟一线性表出 ($j = 1, \dots, n$)，即存在常数 a_{ij} ($i = 1, \dots, m; j = 1, \dots, n$)，使得：

$$\begin{cases} T(\alpha_1) = a_{11}\beta_1 + a_{21}\beta_2 + \dots + a_{m1}\beta_m, \\ T(\alpha_2) = a_{12}\beta_1 + a_{22}\beta_2 + \dots + a_{m2}\beta_m, \\ \vdots \\ T(\alpha_n) = a_{1n}\beta_1 + a_{2n}\beta_2 + \dots + a_{mn}\beta_m \end{cases}$$

称矩阵 $A = (a_{ij})_{m \times n}$ 为线性变换 T 在给定基 B, B' 下的矩阵，简称为线性变换 T 的矩阵或矩阵表示。特别地，当 $W = V$ 时， $m = n$ ，取 $B' = B$ ，则对于 $T \in L(V)$ ， A 为 n 阶方阵。称方阵 A 为线性算子 T 在基 B 下的矩阵，简称为线性算子 T 的矩阵。

令任意向量 α 在基 B 下的坐标向量和 $T(\alpha)$ 在基 B' 下的坐标向量为：

$$x = (x_1, x_2, \dots, x_n)^T \in F^n \quad y = (y_1, y_2, \dots, y_m)^T \in F^m$$

则由线性映射的性质可知：

$$\begin{aligned} T(\alpha) &= x_1 T(\alpha_1) + x_2 T(\alpha_2) + \dots + x_n T(\alpha_n) \\ &= [T(\alpha_1) \quad T(\alpha_2) \quad \dots \quad T(\alpha_n)]x \\ &= [\beta_1 \quad \beta_2 \quad \dots \quad \beta_m]Ax \end{aligned} \quad (6.3)$$

根据上式，有定理：

定理 6.6 (线性变换的矩阵表示)

如图 6.2 所示，线性变换 $T \in L(V, W)$ 的矩阵 A 可以用于计算 $T(\alpha)$ ：首先求出 V 中向量 α 在基 B 下的坐标 x ，再计算 Ax ，得到 $T(\alpha)$ 在基 B' 下的坐标 $y = Ax$ ，再由式 6.3 就可求出 $T(\alpha)$ 。

$$\begin{array}{ccc} \alpha & \xrightarrow{T} & T(\alpha) \\ \updownarrow & & \updownarrow \\ x & \xrightarrow{A} & y (=Ax) \end{array} \quad \begin{array}{l} \alpha \in V, T(\alpha) \in W, \\ x \in F^n, y \in F^m, \end{array}$$

图 6.2: 线性变换的坐标变换求法

根据前面的讨论，在线性空间 V 和 W 中取定基之后，对于任一线性变换 $T \in L(V, W)$ ，由 T 就可惟一地确定一个矩阵 (T 的矩阵) $A_{m \times n}$ 。反过来，对于任何一个矩阵 $A_{m \times n}$ ，可以根据式 6.3 定义一个映射 $T: V \rightarrow W$ ，且易知 T 是线性变换，它在基 B, B' 下的矩阵为 A 。这样，就在线性空间 $L(V, W)$ 与线性空间 $F^{m \times n}$ 之间建立了一一对应的关系。下述定理表明，它们各自的代数运算之间也保持这种一一对应的关系：

定理 6.7 (线性变换的矩阵运算)

由式 6.3 所表示的 $L(V, W)$ 与 $F^{m \times n}$ 之间的对应关系满足以下性质：

1. 线性变换的和对应于矩阵的和；
2. 线性变换的数量乘积对应于矩阵的数量乘积；
3. 线性变换的乘积对应于矩阵的乘积。

下面的定理描述了线性变换和其表示矩阵的重要关系:

定理 6.8 (线性变换和表示矩阵的关系)

设 $T \in L(V, W)$, T 在一个基下的矩阵为 A , 则:

- $R(T)$ 与矩阵 A 的列空间同构, 且 T 的秩等于 A 的秩;
- $\ker(T)$ 与齐次线性方程组 $Ax = 0$ 的解空间同构, 且 T 的零度等于 $n - r(A)$ 。

同构的相关性质见定理5.4。该定理可以通过求线性方程组 $Ax = 0$ 的元素从而求解其对应线性变换的元素。

 **练习 6.2** 设 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 是线性空间 V 的一个基。 T 是 V 上的线性算子, T 在该基下的矩阵为:

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{bmatrix}$$

求 T 的值域与核。

解 由定理6.8知 T 的值域与 A 的列空间同构。设 A 按列分块为 $A = [x_1 \ x_2 \ x_3 \ x_4]$, 计算可知 A 的前两列 x_1, x_2 是 A 的列空间的基, 故 $R(T)$ 中对应的向量:

$$T(\alpha_1) = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4]x_1 = \alpha_1 - \alpha_2 + \alpha_3 + 2\alpha_4$$

$$T(\alpha_2) = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4]x_2 = 2\alpha_2 + 2\alpha_3 - 2\alpha_4$$

是 $R(T)$ 的基, 即:

$$R(T) = \text{span}\{\alpha_1 - \alpha_2 + \alpha_3 + 2\alpha_4, 2\alpha_2 + 2\alpha_3 - 2\alpha_4\}$$

由于 T 的核与齐次线性方程组 $Ax = 0$ 的解空间同构, 计算可得 $Ax = 0$ 的基础解系为:

$$\xi_1 = (-4, -3, 2, 0)^T \quad \xi_2 = (-1, -2, 0, 1)^T$$

故有:

$$\ker(T) = \text{span}\{-4\alpha_1 - 3\alpha_2 + 2\alpha_3, -\alpha_1 - 2\alpha_2 + \alpha_4\}$$

定理 6.9 (线性变换可逆的充要条件)

设 $T \in L(V, W)$, 且 $\dim(V) = \dim(W) = n$, 且 B 和 B' 分别是 V, W 的基, 且 T 在 B, B' 下的矩阵为 A 。则 T 是可逆线性变换的充要条件为 A 是一个可逆矩阵。且 T^{-1} 在 B', B 下的矩阵为 A^{-1}

最后, 我们来说明线性变换在不同的基下, 表示矩阵总是相似的, 即:

定理 6.10 (线性变换的矩阵相似不变性)

设 $T \in L(V)$, $B = \{\alpha_1, \dots, \alpha_n\}$ 和 $B' = \{\beta_1, \dots, \beta_n\}$ 是 V 的两个不同基 T 在 B, B' 下的矩阵分别为 A 和 D , 且知 B 到 B' 的过渡矩阵为 C , 则 $C^{-1}AC = D$, 即线性算子 T 在不同基下的矩阵是相似的。

有关相似的概念详见第7章节。

第六章 练习

- (1) 设 $T \in L(\mathbb{R}^3)$, 定义为 $T(x) = Ax, \forall x \in \mathbb{R}^3$, 其中 $A = \begin{bmatrix} 1 & -1 & 3 \\ 5 & 6 & -4 \\ 7 & 4 & 2 \end{bmatrix}$ 。证明: 几何上, T 的核与值域分别对应

过原点的直线和平面, 并求其方程。

- (2) 设 V_1, V_2, V_3 都是有限维线性空间。 $T_2 \in L(V_1, V_2)$, $T_1 \in L(V_2, V_3)$ 。证明:

$$r(T_1 T_2) \leq \min\{r(T_1), r(T_2)\}$$

- (3) 设 V 上的线性算子 T 满足 $T^2 = T$, 证明:

$$V = \ker(T) \oplus R(T)$$

- (4) 设 $T \in L(\mathbb{R}^3)$, $T(\alpha_1) = (-5, 0, 3)^T$, $T(\alpha_2) = (0, -1, 6)^T$, $T(\alpha_3) = (-5, -1, 9)^T$ 。其中 $\alpha_1 = (-1, 0, 2)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (3, -1, 0)^T$ 。求在基 $\varepsilon_1 = (1, 0, 0)^T$, $\varepsilon_2 = (0, 1, 0)^T$, $\varepsilon_3 = (0, 0, 1)^T$ 下的矩阵。

- (5) 设 T 是 n 维线性空间 V 上的线性变换, $\xi \in V$, 如果 $T^{n-1}\xi \neq 0, T^n\xi = 0$ 。证明: $\xi, T\xi, \dots, T^{n-1}\xi$ 是 V 的一组基, 并求 T 在这组基下的矩阵。

- (6) 设 $a_0, a_1, a_2, \dots, a_n$ 是数域 F 中 $n+1$ 个不同的数。对每个 $0 \leq i \leq n$, 取:

$$f_i = \prod_{0 \leq j \leq n, j \neq i} (x - a_j) = (x - a_0) \cdots (x - a_{i-1})(x - a_{i+1}) \cdots (x - a_n)$$

证明: $f_0, f_1, f_2, \dots, f_n$ 是 $F[x]_n$ 的一组基, 并求基 $f_0, f_1, f_2, \dots, f_n$ 到基 $1, x, x^2, \dots, x^n$ 的过渡矩阵。

(提示: 考虑线性变换 $\sigma: F[x]_n \rightarrow F^{n+1}$)

- (7) 设 T 是 F^n 上线性空间 V 的算子, 且满足:

$$T^2 = T$$

而 A 为算子 T 在 V 的某组基下的矩阵。证明或求解:

(a). $T + I$ 是 V 上的可逆线性变换, 其中 I 是恒等变换。

(b). $r(A) = \text{tr}(A)$

(c). 求 $\det(2I - A)$, 其中 I 为和 A 同阶的单位矩阵。

第七章 特征值和相似矩阵

注 本章为自学内容，有需要的读者可参考本资料。

7.1 特征值和特征向量

7.1.1 特征值和特征向量的定义

定义 7.1 (特征值和特征向量)

设 $A = (a_{ij})_{n \times n}$ 是一个 n 阶矩阵，如果有一个复数 λ 及一个 n 维非零列向量 $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$ ，使得：

$$Ax = \lambda x \quad (7.1)$$

或表述为：

$$(\lambda I - A)x = 0 \quad (7.2)$$

则称 λ 为矩阵 A 的一个特征值，称非零列向量 x 为 A 的对应于（或属于）特征值 λ 的特征向量。

注 特征向量是非零向量，且属于同一特征值 λ_0 的特征向量不唯一。

下面，根据定义来看特征值与特征向量的计算问题。7.2式可看成是一个齐次线性方程组，其系数矩阵是 n 阶矩阵 $\lambda I - A$ ，由于特征向量 x 是它的非零解，故由 $n \times n$ 齐次线性方程组有非零解的充要条件，得：

$$\det(\lambda I - A) = 0 \quad (7.3)$$

或表述为：

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0 \quad (7.4)$$

定义 7.2 (特征方程、特征多项式和特征空间)

称关于 λ 的一元 n 次代数方程 7.3 或 7.4 为矩阵 A 的特征方程；称一元 n 次多项式 $f(\lambda) = \det(\lambda I - A)$ 为 A 的特征多项式；若 λ 为 A 的特征值，则称齐次线性方程组 7.2 的解空间为 λ 的特征空间，并记为 V_λ 。

7.1.2 特征值的代数重数和几何重数

前面的讨论表明， A 的特征值就是 A 的特征方程的根。特征方程 7.3 是 1 元 n 次代数方程，由代数学基本定理，它在复数范围内有 n 个根（重根按重数计算，以下均如此），所以 n 阶矩阵 A 一定有 n 个特征值。因此我们给出特征值的代数重数的定义：

定义 7.3 (特征值的代数重数)

设 A 为 n 阶方阵， A 的特征多项式为

$$\det(\lambda I - A) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s}$$

其中 $m_i (i = 1, 2, \dots, s)$ 均为正整数，且满足：

$$\sum_{i=1}^s m_i = n$$

不难得知 $\lambda_1, \lambda_2, \dots, \lambda_s$ 为 A 的不同特征值，且称 m_i 为 λ_i 的代数重数。

此外, 根据特征空间的定义, 我们还可以给出特征值的几何重数的定义:

定义 7.4 (特征值的几何重数)

把与 λ_i 对应的线性无关的特征向量的个数, 即其特征空间 V_{λ_i} 的维数称为 λ_i 的几何重数, 记为 α_i , 且有:

$$\alpha_i = n - r(\lambda_i I - A)$$

有关一个特征值的代数重数和几何重数, 有重要性质:

定理 7.1 (代数重数和几何重数的关系)

一个特征值的代数重数总是不小于其几何重数。

证明 设 n 阶矩阵 A 的特征值 λ_i 的代数重数为 m_i , 几何重数为 α_i , 其特征空间 V_{λ_i} 的维数为 α_i , 那么就有一组基 $B = \{\beta_1, \beta_2, \dots, \beta_{\alpha_i}\}$ 。根据基的扩充原理可以把这组基扩充为 $B' = \{\beta_1, \beta_2, \dots, \beta_{\alpha_i}, \dots, \beta_n\}$, 且 $A\beta_{\alpha_i+1}, \dots, A\beta_n$ 可分别由 B 的向量线性表示为:

$$A\beta_{\alpha_i+k} = \sum_{j=1}^n c_{jk}\beta_j \quad (c \text{ 为常数}, k = 1, 2, \dots, n - \alpha_i)$$

令矩阵 $P = [\beta_1, \beta_2, \dots, \beta_{\alpha_i}, \dots, \beta_n]$, 那么有:

$$\begin{aligned} AP &= A[\beta_1, \beta_2, \dots, \beta_{\alpha_i}, \dots, \beta_n] \\ &= [A\beta_1, A\beta_2, \dots, A\beta_{\alpha_i}, \dots, A\beta_n] \\ &= [\lambda_i\beta_1, \lambda_i\beta_2, \dots, \lambda_i\beta_{\alpha_i}, A\beta_{\alpha_i+1}, \dots, A\beta_n] \\ &= [\beta_1, \beta_2, \dots, \beta_{\alpha_i}, \dots, \beta_n] \begin{bmatrix} \lambda_i & & & c_{11} & \cdots & c_{1,n-\alpha_i} \\ & \lambda_i & & c_{21} & \cdots & c_{2,n-\alpha_i} \\ & & \ddots & \vdots & & \vdots \\ & & & \lambda_i & c_{\alpha_i,1} & \cdots & c_{\alpha_i,n-\alpha_i} \\ & & & & c_{\alpha_i+1,1} & \cdots & c_{\alpha_i+1,n-\alpha_i} \\ & & & & \vdots & & \vdots \\ & & & & c_{n,1} & \cdots & c_{n,n-\alpha_i} \end{bmatrix} \\ &= P \begin{bmatrix} \lambda_i I_{\alpha_i} & C_1 \\ O & C_2 \end{bmatrix} \\ &= PB \end{aligned}$$

因此有 $B = P^{-1}AP$, 即 A, B 相似。在 7.2 节中会说明 A 和 B 有相同的特征值, 而:

$$\det(\lambda I_{\alpha_i} - B) = (\lambda - \lambda_i)^{\alpha_i} \det(\lambda I_{n-\alpha_i} - C_2)$$

上述式子表明 B 的特征值 λ_i 的代数重数至少为 α_i , 则 $m_i \geq \alpha_i$, 再根据 A, B 相似即证。

7.1.3 矩阵特征值的求法

求 n 阶矩阵 A 的特征值与特征向量的一般步骤是:

- 第 1 步: 求出特征方程 $\det(\lambda I - A) = 0$ 的全部根 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则 $\lambda_1, \lambda_2, \dots, \lambda_n$, 就是 A 的全部特征值。
- 第 2 步: 对于 A 的特征值 λ_i , 求出齐次线性方程组 $(\lambda I - A)x = 0$ 的基础解系

$$\xi_{i1}, \xi_{i2}, \dots, \xi_{i\alpha_i}$$

则 A 的属于特征值 λ_i 的全部特征向量为:

$$x = c_1\xi_{i1} + c_2\xi_{i2} + \cdots + c_{\alpha_i}\xi_{i\alpha_i}$$

其中 $c_1, c_2, \dots, c_{\alpha_i}$ 是不全为零的任意常数。

 **练习 7.1** 求下列矩阵的特征值和特征向量

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$$

解 由特征方程:

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda - 1 & 3 & -3 \\ -3 & \lambda + 5 & -3 \\ -6 & 6 & \lambda - 4 \end{vmatrix} = \begin{vmatrix} \lambda + 2 & 3 & -3 \\ \lambda + 2 & \lambda + 5 & -3 \\ 0 & 6 & \lambda - 4 \end{vmatrix} \\ &= \begin{vmatrix} \lambda + 2 & 3 & -3 \\ 0 & \lambda + 2 & 0 \\ 0 & 6 & \lambda - 4 \end{vmatrix} = (\lambda + 2)^2(\lambda - 4) = 0 \end{aligned}$$

得 A 的全部特征值为: 2 重特征值 $\lambda_1 = \lambda_2 = -2$, 且单特征值 $\lambda_3 = 4$ 。

对于特征值 $\lambda_1 = \lambda_2 = -2$, 求解齐次线性方程组 $(-2I - A)x = 0$, 由:

$$-2I - A = \begin{bmatrix} -3 & 3 & -3 \\ -3 & 3 & -3 \\ -6 & 6 & -6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

得通解 $x_1 = x_2 - x_3$ (x_2, x_3 任意), 从而得基础解系:

$$\xi_1 = (1, 1, 0)^T \quad \xi_2 = (-1, 0, 1)^T$$

ξ_1, ξ_2 就是对应于特征值 -2 的线性无关特征向量, 而对应于特征值 -2 的全部特征向量也就是 $x = k_1\xi_1 + k_2\xi_2$ (其中 k_1, k_2 为不全为零的任意常数)。对于特征值 $\lambda_3 = 4$, 求解齐次线性方程 $(4I - A)x = 0$, 由:

$$4I - A = \begin{bmatrix} 3 & 3 & -3 \\ -3 & 9 & -3 \\ -6 & 6 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

得通解 $x_1 = \frac{1}{2}x_3, x_2 = \frac{1}{2}x_3$ (x_3 任意), 从而得基础解系:

$$\xi_3 = (1, 1, 2)^T$$

ξ_3 就是对应于特征值 4 的线性无关特征向量, 而对应于特征值 4 的全部特征向量就是 $x = k\xi_3$ (k 为任意非零常数)。这是因为特征向量不能是零向量。

7.1.4 特征值的性质

定理 7.2

设 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 的全部特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则:

$$\lambda_1 \lambda_2 \cdots \lambda_n = \det(A) \quad (7.5)$$

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn} \quad (7.6)$$

其中, 记 7.6 的右式为矩阵 A 的迹 (trace), 记为 $\text{tr}(A)$ 。



证明 由条件知 A 的特征方程:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0$$

的全部根为 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。注意特征多项式 $\det(\lambda I - A)$ 是最高次项系数为 1 的 n 次多项式, 且由因式分解定理知:

$$\begin{aligned} \det(\lambda I - A) &= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \\ &= \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \end{aligned}$$

另一方面, 利用数学归纳法可以证明: 如将 $\det(\lambda I - A)$ 的展开式按 λ 的方次降幂排列, 则有:

$$\det(\lambda I - A) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n \det(A)$$

比较上面两个等式右端的常数项和 λ^{n-1} 项, 即证。

定理 7.3

设 λ 为矩阵 A 的一个特征值且 x 为对应的一个特征向量, 则:

1. 对任何正整数 m , λ^m 为矩阵 A^m 的一个特征值且 x 为对应的一个特征向量;
2. 对任何多项式 $f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$, $f(\lambda)$ 为矩阵 $f(A) = a_m A^m + a_{m-1} A^{m-1} + \cdots + a_1 A + a_0 I$ 的一个特征值且 x 为对应的一个特征向量;
3. 设矩阵 A 可逆则 $\lambda \neq 0$, 且 $\frac{1}{\lambda}$ 为 A^{-1} 的一个特征值, $\frac{\det(A)}{\lambda}$ 为 A 的伴随矩阵 A^* 的一个特征值。

练习 7.2 3 阶矩阵 A 的行列式为 6, 且有两个特征值 $\lambda_1 = 1, \lambda_2 = 2$, 求 $A^* + A^{-1} - I$ 的特征值。

解 根据定理 7.2 可知 A 有剩下一个特征值 $\lambda_3 = 3$, 且对于 $A^* + A^{-1} - I$ 有特征值 $f(\lambda_i)$ 满足:

$$f(\lambda_i) = \frac{\det(A)}{\lambda_i} + \frac{1}{\lambda_i} - 1$$

分别代入可有特征值 $\mu_1 = 6, \mu_2 = \frac{5}{2}, \mu_3 = \frac{4}{3}$ 。

对于属于不相同的特征值的特征向量, 有重要定理:

定理 7.4 (相异特征值的特征向量的线性无关性)

设 $\lambda_1, \lambda_2, \dots, \lambda_k$ 为矩阵 A 的互不相同的特征值, x_i 为 A 的属于特征值 λ_i 的特征向量 ($i = 1, 2, \dots, k$), 则 $x_1, x_2, \dots, x_k$ 线性无关, 即属于互不相同特征值的特征向量线性无关。

进一步推广: 设 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是矩阵 A 的互不相同的特征值 $x_{i1}, x_{i2}, \dots, x_{i\alpha_i}$ 为属于 λ_i 的一组线性无关特征向量 ($i = 1, 2, \dots, k$), 则向量组:

$$x_{11}, x_{12}, \dots, x_{1\alpha_1}; x_{21}, x_{22}, \dots, x_{2\alpha_2}; \cdots; x_{k1}, x_{k2}, \dots, x_{k\alpha_k}$$

是线性无关的。

这一定理对下一小节相似矩阵有重要铺垫。

7.2 相似矩阵

7.2.1 相似矩阵的概念

定义 7.5 (相似矩阵)

设 A, B 都是 n 阶矩阵, 如果存在一个 n 阶可逆矩阵 P , 使得:

$$P^{-1}AP = B$$

则称 A 相似于 B , 或 A 与 B 相似, 记作 $A \sim B$ 。并称由 A 到 $B = P^{-1}AP$ 的变换为一个相似变换。如果 A 与一个对角矩阵相似, 则称 A 可相似对角化, 简称为 A 可对角化。



笔记 矩阵相似是矩阵之间的一种关系, 且容易证明满足自反性、对称性、传递性。

矩阵“相似”从字面上表明具有此关系的矩阵在一些特征上有相近之处, 具体而言有:

定理 7.5 (相似矩阵的性质)

设 n 阶矩阵 A 与 B 相似, 则:

1. $\det(A) = \det(B)$;
2. $r(A) = r(B)$, 特别当 A 与 B 都可逆时, A^{-1} 与 B^{-1} 也相似;
3. A 与 B 有相同的特征多项式, 从而有相同的特征值。

练习 7.3 已知矩阵 A 与 B 相似, 其中:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 4 & -2 \\ -3 & -3 & a \end{bmatrix} \quad B = \begin{bmatrix} 2 & & \\ & 2 & \\ & & b \end{bmatrix}$$

求 a, b 的值。

解 B 的特征值为 $2, 2, b$, 根据 A 与 B 相似和相似矩阵的性质直接得:

$$\begin{cases} 2 + 2 + b = 1 + 4 + a \\ 2 \times 2 \times b = \det(A) = 6(a - 1) \end{cases}$$

由此解得 $a = 5, b = 6$ 。

7.2.2 相似对角化

从上面的例题看出: 对角矩阵的全部特征值是其主对角线上的全部元素。因此, 若矩阵 A 相似于对角矩阵 $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 则由定理 7.5 的 (3) 知 A 的全部特征值就是 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

然而, 并非任何矩阵都可对角化, 因此研究矩阵对角化有两个基本问题:

- 矩阵可对角化的条件;
- 若矩阵 A 可对角化, 如何求可逆矩阵 P , 使 $P^{-1}AP$ 成对角矩阵, 这被称为相似对角化。

因此, 我们来说明一个矩阵可对角化的充要条件:

定理 7.6 (充要条件 I)

n 阶矩阵 A 可对角化的充要条件是 A 有 n 个线性无关的特征向量。

证明 必要性：设有可逆矩阵 P ，使得：

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = D$$

设 P 按列分块为：

$$P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

由于 P 可逆，则 $x_1, x_2, \dots, x_n$ 线性无关，且得 $AP = PD$ ，即：

$$A \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

亦即：

$$\begin{bmatrix} Ax_1 & Ax_2 & \cdots & Ax_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix}$$

或表述为：

$$Ax_i = \lambda_i x_i \quad i = 1, 2, \dots, n$$

因 $x_i \neq 0$ ，故 λ_i 为 A 的特征值， x_i 为对应于 λ_i 的特征向量 ($i = 1, 2, \dots, n$)。所以 A 有 n 个线性无关的特征向量 $x_1, x_2, \dots, x_n$ 。将以上证明倒推上去就是充分性的证明。

联立上一小节的定理 7.1 可以得到另一个充要条件：

定理 7.7 (充要条件 II)

n 阶矩阵 A 可对角化的充要条件是 A 的每个特征值的几何重数都等于它的代数重数，亦即每个特征值 λ_i 满足 $m_i = \alpha_i$ 。



证明 设 $A_{n \times n}$ 的互不相同的全部特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$ ，其代数重数分别为 $m_1, m_2, \dots, m_k$ ，其中 $m_1 + m_2 + \dots + m_k = n$ ；其几何重数分别为 $\alpha_1, \alpha_2, \dots, \alpha_k$ 。我们来证明：

$$A \text{ 可对角化} \Leftrightarrow m_i = \alpha_i (i = 1, 2, \dots, k)$$

设对应于特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的线性无关特征向量分别为：

$$x_{11}, x_{12}, \dots, x_{1\alpha_1}; x_{21}, x_{22}, \dots, x_{2\alpha_2}; \dots; x_{k1}, x_{k2}, \dots, x_{k\alpha_k}$$

可知上面的 $\alpha_1 + \alpha_2 + \dots + \alpha_k$ 个特征向量是线性无关的，它们又显然可以线性表示 A 的每个特征向量，因此上面的向量组就是 A 的全体特征向量的一个极大无关组。这就是说， A 的线性无关的特征向量有且只有 $\alpha_1 + \alpha_2 + \dots + \alpha_k$ 个。于是由充要条件 I 知：

$$A \text{ 可对角化} \Leftrightarrow \alpha_1 + \alpha_2 + \dots + \alpha_k = n$$

$$\Leftrightarrow \alpha_1 + \alpha_2 + \dots + \alpha_k = m_1 + m_2 + \dots + m_k$$

根据代数重数和几何重数的关系知 $\alpha_i \leq m_i (i = 1, 2, \dots, k)$ ，所以 A 可对角化的充要条件是 $\alpha_i = m_i (i = 1, 2, \dots, k)$ 。因此充要条件 II 证毕。

根据以上讨论，可得到判别 n 阶矩阵 A 可否对角化，以及在可对角化时如何求对角矩阵 D 和可逆矩阵 P 使得 $P^{-1}AP = D$ 的一般步骤：

首先求出 A 的全部特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。如果 A 的特征值都是单特征值，或者虽然有重特征值，但每个特征值的几何重数都等于它的代数重数，则 A 可对角化，否则不能对角化。在 A 可对角化时，求出对应于每个特征值的线性无关特征向量，从而得到 A 的 n 个线性无关特征向量 $\xi_1, \xi_2, \dots, \xi_n$ ，其中 ξ_i 是对应于特征值 λ_i 的特

征向量 ($i = 1, 2, \dots, n$)。以 $\xi_1, \xi_2, \dots, \xi_n$ 为列向量构成矩阵:

$$P = [\xi_1 \quad \xi_2 \quad \cdots \quad \xi_n]$$

则 P 可逆, 且有:

$$P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

 **笔记** 上述过程简记为: (1) 求特征值; (2) 判断是否可对角化; (3) 求特征空间的基; (4) 构造 P 和对角矩阵 D 。

 **练习 7.4** 下列矩阵能否对角化? 若能对角化, 求可逆矩阵 P , 使 $P^{-1}AP$ 成对角矩阵:

$$A_1 = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} \quad A_2 = \begin{bmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -7 \end{bmatrix}$$

解 A_1 就是练习 7.1 中的矩阵, 它有 3 个线性无关的特征向量:

$$\xi_1 = (1, 1, 0)^T, \xi_2 = (-1, 0, 1)^T, \xi_3 = (1, 1, 2)^T$$

且 ξ_1, ξ_2, ξ_3 分别属于特征值 $-2, -2, 4$ 。由充要条件 I 和相异特征值对应线性无关特征向量知 A_1 可对角化。令:

$$P = [\xi_1 \quad \xi_2 \quad \xi_3] = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

则有:

$$P^{-1}A_1P = \text{diag}(-2, -2, 4)$$

再求 A_2 的特征值:

$$\begin{aligned} \det(\lambda I - A_2) &= \begin{vmatrix} \lambda - 4 & -2 & 5 \\ -6 & \lambda - 4 & 9 \\ -5 & -3 & \lambda + 7 \end{vmatrix} = \begin{vmatrix} \lambda - 1 & -2 & 5 \\ \lambda - 1 & \lambda - 4 & 9 \\ \lambda - 1 & -3 & \lambda + 7 \end{vmatrix} \\ &= \lambda^2(\lambda - 1) = 0 \end{aligned}$$

知 A_2 有 2 重特征值 $\lambda_1 = \lambda_2 = 0$ 和单特征值 $\lambda_3 = 1$, 我们来确定特征值 $\lambda_1 = \lambda_2 = 0$ 的几何重数, 也就是齐次线性方程组 $(0I - A_2)x = 0$ 的基础解系所含向量个数, 由于该方程组的系数矩阵:

$$0I - A_2 = -A_2 = \begin{bmatrix} -4 & -2 & 5 \\ -6 & -4 & 9 \\ -5 & -3 & 7 \end{bmatrix}$$

的秩为 2, 故该方程组的基础解系含 $3-2=1$ 个向量, 即 2 重特征值 $\lambda_1 = \lambda_2 = 0$ 的几何重数为 1。所以, 由充要条件 II 知 A_2 不能对角化。

7.2.3 实对称矩阵的对角化讨论

定义 7.6 (实对称矩阵)

设 $A = (a_{ij})$ 是一个复矩阵, 称 $\bar{A} = (\bar{a}_{ij})$ 为 A 的共轭矩阵, 其中 $\bar{a}_{ij}$ 是复数 a_{ij} 的共轭数。容易验证矩阵的共轭运算满足:

$$\overline{A+B} = \bar{A} + \bar{B}, \quad \overline{\lambda A} = \lambda \bar{A}, \quad \overline{AB} = \bar{B}\bar{A}$$

其中 A, B 是同阶复矩阵, λ 是复数。当 $\bar{A} = A$ 时, A 就是实矩阵。如果实矩阵 A 为对称矩阵, 则称 A 为实对称矩阵。

实对称矩阵有以下重要性质:

定理 7.8 (实对称矩阵的性质)

设 A 是实对称矩阵, 那么有:

1. A 的特征值都是实数;
2. 设 λ_1, λ_2 是 A 的相异特征值, x_1 和 x_2 分别是与 λ_1, λ_2 对应的特征向量, 则 x_1 与 x_2 正交。
3. 必存在 n 阶正交矩阵 P , 使得 $P^{-1}AP = P^TAP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 为对角矩阵。

**证明**

1. 设复数 λ 为实对称矩阵 A 的特征值, 非零复向量 $x = (x_1, \dots, x_n)^T$ 为对应的特征向量, 即有 $Ax = \lambda x$. 对上式两端取共轭, 并注意 $\overline{A} = A$, 得 $\overline{Ax} = \overline{\lambda} \overline{x}$, 再对两端取转置, 并注意 $A^T = A$, 得:

$$\overline{x}^T A = \overline{\lambda} \overline{x}^T$$

用 x 右乘上式两端, 得:

$$\overline{x}^T Ax = \overline{\lambda} \overline{x}^T x$$

因 $Ax = \lambda x$, 从而有:

$$\lambda \overline{x}^T x = \overline{\lambda} \overline{x}^T x$$

移项, 得:

$$(\lambda - \overline{\lambda}) \overline{x}^T x = 0$$

由于

$$\overline{x}^T x = (\overline{x_1}, \dots, \overline{x_n}) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = |x_1|^2 + \dots + |x_n|^2 > 0$$

所以 $\lambda - \overline{\lambda} = 0$, 故 $\lambda = \overline{\lambda}$, 即 λ 为实数。

2. 不妨设:

$$x_1 = (a_1, \dots, a_n)^T \quad x_2 = (b_1, \dots, b_n)^T$$

我们要证明:

$$\langle x_1, x_2 \rangle = \sum_{i=1}^n a_i b_i = x_1^T x_2 = 0$$

由已知的:

$$Ax_1 = \lambda_1 x_1$$

两端取转置, 并注意 $A^T = A$, 得:

$$x_1^T A = \lambda_1 x_1^T$$

用 x_2 右乘上式两端, 得:

$$x_1^T Ax_2 = \lambda_1 x_1^T x_2$$

因 $Ax_2 = \lambda_2 x_2$, 得:

$$\lambda_2 x_1^T x_2 = \lambda_1 x_1^T x_2$$

亦即:

$$(\lambda_2 - \lambda_1) x_1^T x_2 = 0$$

由于 $\lambda_2 \neq \lambda_1$, 从而得 $x_1^T x_2 = 0$, 即正交。

3. 对矩阵 A 的阶数 n 用数学归纳法来证。当 $n = 1$ 时, 结论显然成立, 假设当阶数为 $n - 1$ 时结论成立, 下面来看阶数为 n 时的情形。设 λ_i 为 n 阶实对称矩阵 A 的一个特征值且 α_i 为对应的一个单位特征向量, 则

将 α_1 扩充可得到 $\mathbb{R}^n$ 的标准正交基 $\alpha_1, \alpha_2, \dots, \alpha_n$ (例如可取齐次线性方程组 $\alpha_1^T x = 0$ 的解空间的标准正交基作为 $\alpha_2, \dots, \alpha_n$), 令矩阵:

$$P_1 = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]$$

则 P_1 为正交矩阵, 且因 $A\alpha_1 = \lambda_1\alpha_1$, 所以有:

$$\begin{aligned} AP_1 &= [A\alpha_1 \quad A\alpha_2 \quad \cdots \quad A\alpha_n] = [\lambda_1\alpha_1 \quad A\alpha_2 \quad \cdots \quad A\alpha_n] \\ &= [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{bmatrix} \\ &= P_1 \begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix} \end{aligned}$$

其中, β^T 为 $n-1$ 维行向量, A_1 为 $n-1$ 阶方阵, 用 P_1^T 右乘两端, 并注意 $P_1 P_1^T = I$, 得:

$$A = P_1 \begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix} P_1^T$$

转置, 得:

$$A^T = P_1 \begin{bmatrix} \lambda_1 & 0 \\ \beta & A_1^T \end{bmatrix} P_1^T$$

由于 A 为对称矩阵, 即 $A = A^T$, 得:

$$\begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ \beta & A_1^T \end{bmatrix}$$

从而有 $\beta = 0, A_1 = A_1^T$, 即有 $A = P_1 \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} P_1^T$ 或 $P_1^T A P_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix}$ 。 A_1 是 $n-1$ 阶实对称矩阵, 于是由归纳假设, 存在 $n-1$ 阶正交矩阵 P_2 , 使得:

$$P_2^{-1} A_1 P_2 = P_2^T A_1 P_2 = \text{diag}(\lambda_2, \dots, \lambda_n)$$

令矩阵:

$$P = P_1 \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix}$$

则 P 为两个正交矩阵之积, 故 P 为 n 阶正交矩阵, 且使:

$$\begin{aligned} P^T A P &= \begin{bmatrix} 1 & 0 \\ 0 & P_2^T \end{bmatrix} P_1^T A P_1 \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & P_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & P_2^T A_1 P_2 \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \end{aligned}$$

所以, 结论对任何正整数 n 都成立。

对于实对称矩阵 A , 现在给出求一个正交矩阵 P , 使得 $P^{-1}AP = P^TAP$ 成为对角矩阵的一般计算步骤:

- 第 1 步: 求出 A 的全部特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$;
- 第 2 步: 对 A 的每个特征值 λ_i , 求出方程组

$$(\lambda_i I - A)x = 0$$

的一个基础解系 (即特征空间 V_{λ_i} 的一个基);

- 第3步: 对每个特征值 λ_i , 利用 Gram-Schmidt 正交化方法将 V_{λ_i} 的基中的向量先正交化, 再单位化 (如果 λ_i 为单特征值或 V_{λ_i} 的基中的向量已是正交向量组, 则只需单位化), 从而得到 V_{λ_i} 的标准正交基。再将所有特征空间的标准正交基合在一起, 便得到 A 的 n 个标准正交的特征向量 $e_1, e_2, \dots, e_n$, 这里利用了定理7.8的第(2)条;
- 第4步: 令矩阵 $P = [e_1 \ e_2 \ \cdots \ e_n]$, 则 P 为正交矩阵, 且有:

$$P^{-1}AP = P^TAP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

其中 P 的第 j 列 e_j 是属于特征值 λ_j 的特征向量 ($j = 1, 2, \dots, n$)。

以上过程, 也称为 n 阶实对称矩阵 A 的正交相似对角化过程。可以看出这个计算的关键在于求出 A 的 n 个标准正交的特征向量。下面给例题说明:

 **练习 7.5** 设实对称矩阵:

$$A = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix}$$

求一个正交矩阵 P , 使 $P^{-1}AP$ 成对角矩阵。

解 由 A 的特征方程:

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda - 1 & -1 & 2 \\ -1 & \lambda + 2 & -1 \\ 2 & -1 & \lambda - 1 \end{vmatrix} = \begin{vmatrix} \lambda - 1 & -1 & 2 \\ -1 & \lambda + 2 & -1 \\ 2 & -1 & \lambda - 1 \end{vmatrix} \\ &= \begin{vmatrix} \lambda & -1 & 2 \\ 0 & \lambda + 3 & -3 \\ 0 & 0 & \lambda - 3 \end{vmatrix} = \lambda(\lambda + 3)(\lambda - 3) = 0 \end{aligned}$$

得 A 的特征值为 $\lambda_1 = 3, \lambda_2 = -3, \lambda_3 = 0$ 。对于 $\lambda_1 = 3$, 解齐次线性方程组 $(3I - A)x = 0$, 由:

$$3I - A = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 5 & -1 \\ 2 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

得通解: $x_1 = -x_3, x_2 = 0$ (x_3 任意)。令 $x_3 = 1$, 得其基础解系 $\xi_1 = (-1, 0, 1)^T$, 将 ξ_1 单位化, 得属于 $\lambda_1 = 3$ 的单位特征向量 $e_1 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T$ 。

同理可求得属于 $\lambda_2 = -3, \lambda_3 = 0$ 的单位特征向量分别可取为:

$$e_2 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)^T, \quad e_3 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T$$

实对称矩阵的性质保证了 e_1, e_2, e_3 一定是两两正交的, 故 e_1, e_2, e_3 就是 A 的 3 个标准正交的特征向量, 于是得所求的正交矩阵 P 可取为:

$$P = [e_1 \ e_2 \ e_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

且有 $P^{-1}AP = P^TAP = \text{diag}(3, -3, 0)$ 。

 **练习 7.6** 设 3 阶实对称矩阵 A 的特征值为 $\lambda_1 = -1, \lambda_2 = \lambda_3 = 1$, 且 $\xi_1 = (0, 1, 1)^T$ 为对应于特征值 λ_1 的特征向量, 求这个实对称矩阵 A 。

解 实对称矩阵 A 必相似于对角矩阵, 即必存在可逆矩阵 P 和对角矩阵 D , 使 $P^{-1}AP = D$ 。因此, 若这样的矩阵 P 和对角矩阵 D 能求出, 则可由此解出 A 来。

由题设条件知 A 必相似于对角矩阵 $D = \text{diag}(-1, 1, 1)$ 。为求出上述的可逆矩阵 P , 需要求出 A 的属于特征

值 $\lambda_2 = \lambda_3 = 1$ 的特征向量。设属于 $\lambda_2 = \lambda_3 = 1$ 的特征向量为 $x = (x_1, x_2, x_3)^T$ ，则知 x 与 ξ_1 正交，即：

$$x_2 + x_3 = 0$$

解上面这个齐次方程，得其基础解系：

$$\xi_2 = (1, 0, 0)^T, \quad \xi_3 = (0, 1, -1)^T$$

ξ_2, ξ_3 就是属于 $\lambda_2 = \lambda_3 = 1$ 的线性无关特征向量。由特征值的性质，知 ξ_1, ξ_2, ξ_3 是 3 阶矩阵 A 的 3 个线性无关特征向量。令：

$$P = [\xi_1 \quad \xi_2 \quad \xi_3] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

则 P 可逆，且使 $P^{-1}AP = \text{diag}(-1, 1, 1)$ ，由此解得：

$$\begin{aligned} A &= P \text{diag}(-1, 1, 1) P^{-1} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \end{aligned}$$

第七章 练习

- (1) 证明：幂等矩阵一定有特征值，并且它的特征值是 1 或者 0。
 (2) 设 A 是数域 F 上的 n 阶矩阵， $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征多项式的全部根。令矩阵：

$$G = \begin{bmatrix} A & A^m \\ A^m & A \end{bmatrix}$$

其中 m 是正整数。求 G 的特征多项式的全部根。

- (3) 求 n 阶矩阵：

$$C = [\varepsilon_n, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_{n-1}]$$

的全部复特征值和复特征向量。

- (4) 证明：如果 α 与 β 是 n 级矩阵 A 的属于不同特征值的特征向量，那么 $\alpha + \beta$ 不是 A 的特征向量。
 (5) 对于下述实对称矩阵 A 求正交矩阵 T 使得 $T^{-1}AT$ 为对角矩阵：

$$A_1 = \begin{bmatrix} 0 & -2 & 2 \\ -2 & -3 & 4 \\ 2 & 4 & -3 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 2 & 4 \\ 2 & -2 & 2 \\ 4 & 2 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 4 & 1 & 0 & -1 \\ 1 & 4 & -1 & 0 \\ 0 & -1 & 4 & 1 \\ -1 & 0 & 1 & 4 \end{bmatrix}$$

- (6) 证明：正交矩阵 A 如果有两个不同的特征值，那么 A 的属于不同特征值的特征向量是正交的。

第八章 二次曲面和二次型

8.1 常见的二次曲面

8.1.1 曲面

像在平面解析几何中把曲线看作是动点的轨迹一样,把曲面也看成是一个动点或一条动曲线(直线)按一定条件或规律运动而形成的轨迹。这样,面上的任一点 P 必须满足一定的条件或规律,也只有面上的点才满足这个条件或规律。如果点 P 的坐标为 (x, y, z) ,那么从这个条件或规律就能导出一个含有变量 x, y, z 的方程

$$F(x, y, z) = 0 \quad (8.1)$$

如果当且仅当 P 为面上的点时,点 P 的坐标才满足上述方程,则称之为曲面的方程,而把这个曲面称为方程的图形。而方程 8.1 称为曲面的直角坐标方程。举例而言,球心为 (x_0, y_0, z_0) 且半径为 R 的球面的直角坐标方程表示为:

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2 \quad (8.2)$$

一般地,所谓曲面 Σ 的参数方程,是指由 $\mathbb{R}^2$ 的某区域 D 到 $\mathbb{R}^3$ 的一个映射,即 3 个 2 元函数:

$$\begin{cases} x = x(u, v) \\ y = y(u, v) \\ z = z(u, v) \end{cases} \quad (u, v) \in D \quad (8.3)$$

其中 D 是 u, v 平面的某个区域。具体地说,如果对于曲面 Σ 上的每一点 $P(x, y, z)$,其坐标 (x, y, z) 必是区域 D 内某点 (u, v) 在映射 8.3 下的像;反过来,对于区域 D 内每一点 (u, v) ,它在映射 8.3 下的像点必是曲面 Σ 上的一点,则称 8.3 式是曲面 Σ 的参数方程。举例而言,球心为 (x_0, y_0, z_0) 且半径为 R 的球面的参数方程表示为:

$$\begin{cases} x = x_0 + R \sin \theta \cos \varphi \\ y = y_0 + R \sin \theta \sin \varphi \\ z = z_0 + R \cos \theta \end{cases} \quad (0 \leq \theta \leq \pi, 0 \leq \varphi < 2\pi) \quad (8.4)$$

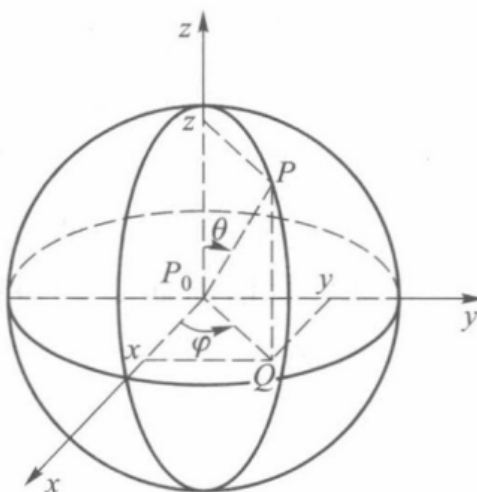


图 8.1: 球心为 (x_0, y_0, z_0) 且半径为 R 的球面

8.1.2 曲线

正像空间直线可看作是两个平面的交线一样, 空间曲线也可看作是两个曲面的交线。设 $F_1(x, y, z) = 0$ 和 $F_2(x, y, z) = 0$ 是两个曲面的方程, 它们的交线是空间曲线 C , 于是 C 上任意一点 $P(x, y, z)$ 必同时在这两个曲面上, 因而点 P 的坐标同时满足这两个方程。反之, 坐标同时满足这两个方程的点 P 必同时在两个曲面上, 因而 P 是这两个曲面交线 C 上的点。把两个曲面方程联立在一起的方程组:

$$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0 \end{cases} \quad (8.5)$$

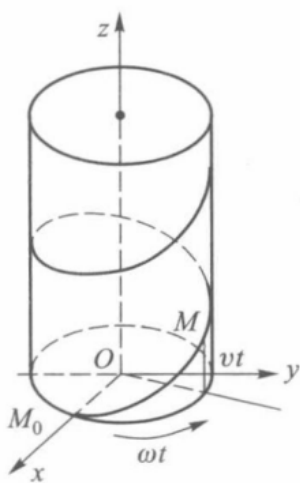
称为空间曲线 C 的方程。对于形如 8.5 式的 3 元方程组, 一般情况下它的图形是一条空间曲线, 它就是这两个方程所表示的曲面的交线。

上式是以点的直角坐标形式给出的, 所以也称它为曲线的直角坐标方程或一般式方程。同空间直线一样, 空间曲线也可用含有一个参数的参数方程来表示。

定义 8.1 (螺旋线)

设动点一方面在一平面内做匀速圆周运动, 角速率为 ω , 圆的半径为 a ; 另一方面, 该平面又沿着与平面垂直的方向作匀速直线运动, 速率为 v 。动点运动的轨迹曲线 C 称为螺旋线。当螺旋线绕 z 轴运动且初始位置位于 $M(a, 0, 0)$ 时, 记参数方程为:

$$\begin{cases} x = a \cos \theta \\ y = a \sin \theta \\ z = \frac{v}{\omega} \theta \end{cases} \quad (0 \leq \theta < +\infty) \quad (8.6)$$

图 8.2: 绕 z 轴的螺旋线

笔记 关于曲面 $F(x, y, z) = 0$ 的图形, 一般来说, 平面解析几何中的描点法已不适用。而且, 要判断方程所表示的曲面的形状, 通常都不是很容易的事情。应用专门软件, 计算机可以绘制出一个方程所表示曲面的形状。对于比较简单的图形, 我们可以采用所谓平行截面法或称截痕法来推知其形状。这个方法是用平面去截制曲面, 从而在截面上得到一条截线, 再令平面平行移动, 截线就随着移动, 从而描出一个曲面。于是, 我们就可利用平面解析几何的知识, 通过分析截线形状的变化, 推知曲面的大致形状。

练习 8.1 讨论方程 $x^2 + y^2 = 2z$ 的图形。

解 显然 $z \geq 0$, 故图形在 Oxy 坐标面的上方。我们用平行于 Oxy 坐标面的平面 $z = c (c \geq 0)$ 去截曲面, 其截线的方程不难得到是:

$$\begin{cases} x^2 + y^2 = 2c \\ z = c \end{cases}$$

这是一个中心在 $(0, 0, c)$, 半径为 $\sqrt{2c}$ 的圆。当 c 从 0 逐渐变大至无限时, 圆的半径也就相应地从 0 逐渐变大至无限, 而圆的中心也随着从原点 O 沿 z 轴无限上移, 因此方程的图形不是封闭的, 而是向 Oxy 坐标面的上方无限伸展, 如下图所示。

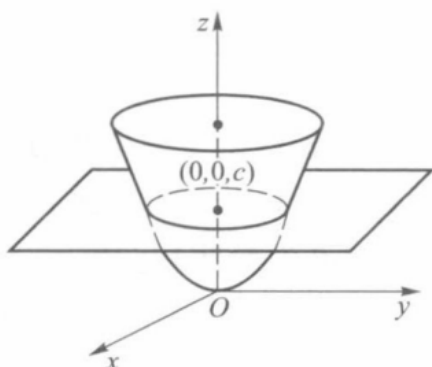


图 8.3: 练习 8.1 的图形

8.1.3 柱面、锥面和旋转面

定义 8.2 (柱面)

设 Γ 是空间一条定曲线, C 是空间一条定直线, 动直线 L 始终平行于直线 C 并沿着曲线 Γ 移动而形成的轨迹叫做柱面, Γ 称为柱面的准线, 动直线 L 称为柱面的母线。

注 一般来说, 空间中的一个 2 元方程代表一个母线与坐标轴平行的柱面, 而且, 方程中缺少某个变量, 则该柱面的母线就与该变量对应的那个轴平行。读者不可误将柱面的方程看作平面曲线的方程。

例如, 方程 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, $y^2 = 2px$ 都是母线平行于 z 轴的柱面 (见下图)。它们的准线分别是 Oxy 面的椭圆 $\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ z = 0 \end{cases}$ 双曲线 $\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ z = 0 \end{cases}$ 抛物线 $\begin{cases} y^2 = 2px \\ z = 0 \end{cases}$ 。这三个曲面分别称为椭圆柱面、双曲柱面和抛物柱面, 是几种常见的二次柱面。特别地, $x^2 + y^2 = 1$ 是大家常用的圆柱面。

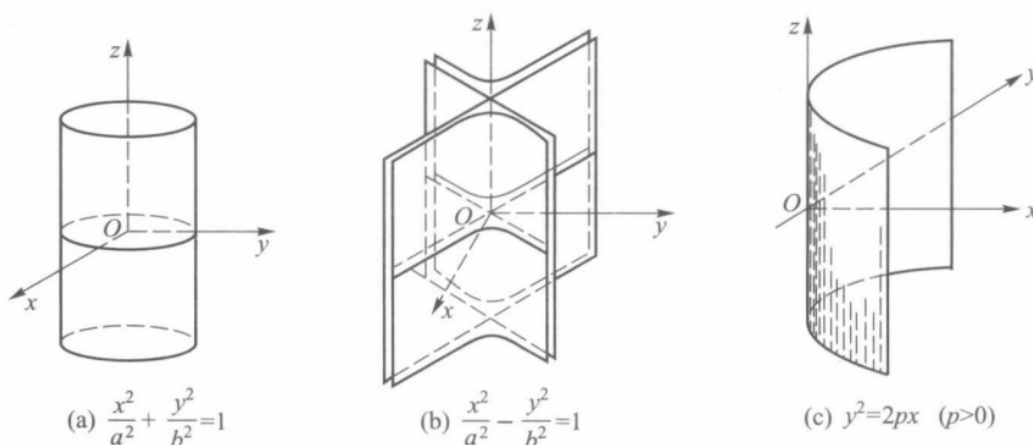


图 8.4: 几种常见的柱面

下面的练习说明了柱面方程的求解方法:

练习 8.2 建立母线平行于直线 $C: x = y = z$, 且准线为:

$$\Gamma: \begin{cases} x^2 + y^2 + z^2 = a^2, \\ x + y + z = 0 \end{cases}$$

的柱面方程。

解 设 $M(x, y, z)$ 是所求柱面上的任一点, 则由母线的方向向量为 $s = (1, 1, 1)^T$, 可得过点 M 的母线的参数方程是:

$$X = x + t \quad Y = y + t \quad Z = z + t$$

这条母线必与 Γ 相交, 故它们的交点的坐标 (X, Y, Z) 必满足 Γ 的方程, 即有:

$$\begin{cases} (x+t)^2 + (y+t)^2 + (z+t)^2 = a^2, \\ (x+t) + (y+t) + (z+t) = 0. \end{cases}$$

消去参数 t , 得:

$$2x^2 + 2y^2 + 2z^2 - 2xy - 2yz - 2zx = 3a^2$$

这就是所求柱面的方程。

现在来看柱面的参数方程. 设柱面的母线与向量 $(l, m, n)^T$ 平行, 准线 Γ 由下参数方程给定:

$$x = f_1(u) \quad y = f_2(u) \quad z = f_3(u) \quad \alpha \leq u \leq \beta$$

那么, 柱面的每一条母线都是过 Γ 上某一点 $(f_1(u), f_2(u), f_3(u))$, 以 $(l, m, n)^T$ 为方向向量的直线, 参数方程为:

$$\begin{cases} x = f_1(u) + lv, \\ y = f_2(u) + mv, \\ z = f_3(u) + nv \end{cases} \quad (\alpha \leq u \leq \beta, -\infty < v < +\infty) \quad (8.7)$$

由于点 $P(x, y, z)$ 在柱面上当且仅当点 P 在柱面的某条母线上, 故上式就是柱面的参数方程 (以 u, v 为参数)。

定义 8.3 (锥面)

设动直线 L 沿定曲线 Γ 移动, 移动时 L 始终通过定点 M_0 , 这样由动直线 L 移动所形成的曲面称为锥面。动直线 L 称为锥面的母线, 定点 M_0 称为锥面的顶点, 定曲线 Γ 称为锥面的准线。



注 易知锥面由它的顶点、准线唯一确定; 但反之, 同一锥面的准线不是惟一的。

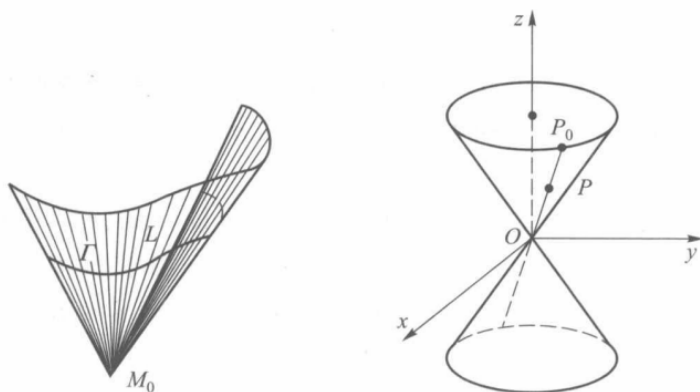


图 8.5: 锥面的图形

我们先来建立顶点为原点 $O(0, 0, 0)$, 准线为平面 $z = z_0$ 内的曲线 $\Gamma: \begin{cases} f(x, y) = 0, \\ z = z_0 \end{cases}$ 的锥面的方程. 任取锥面上一点 $P(x, y, z)$ (图 7.1.8), 则点 P 必在某条母线上, 设该母线与准线 Γ 交于点 $P_0(x_0, y_0, z_0)$, 则该母线的方程为:

$$\frac{x}{x_0} = \frac{y}{y_0} = \frac{z}{z_0}$$

由此得:

$$x_0 = \frac{z_0 x}{z}, \quad y_0 = \frac{z_0 y}{z}$$

因为点 $P_0(x_0, y_0, z_0)$ 在 Γ 上, 故有 $f(x_0, y_0) = 0$, 即

$$f\left(\frac{z_0 x}{z}, \frac{z_0 y}{z}\right) = 0 \quad (8.8)$$

显然, 当且仅当点 P 在锥面上时, 点 P 的坐标才满足 8.8 式, 因此 8.8 式就是所求锥面的方程.

现在来建立锥面的参数方程. 设锥面以 $M_0(x_0, y_0, z_0)$ 为顶点, 以 $\Gamma: x = f_1(u), y = f_2(u), z = f_3(u) (\alpha \leq u \leq \beta)$ 为准线. 由于锥面上任一点 (x, y, z) 都在联结 M_0 与 Γ 上某一点 $(f_1(u), f_2(u), f_3(u))$ 的母线上, 这条母线的方向向量为 $(f_1(u) - x_0, f_2(u) - y_0, f_3(u) - z_0)^T$, 故由直线的参数方程可得锥面的参数方程 (以 u, v 为参数) 为:

$$\begin{cases} x = x_0 + (f_1(u) - x_0)v \\ y = y_0 + (f_2(u) - y_0)v \\ z = z_0 + (f_3(u) - z_0)v \end{cases} \quad (\alpha \leq u \leq \beta, -\infty < v < +\infty) \quad (8.9)$$

 **练习 8.3** 求以原点 $O(0, 0, 0)$ 为顶点, 平面 $z = c$ 内的椭圆 $\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \\ z = c \end{cases}$ 为准线的锥面的方程.

解 由 8.8 式得所求锥面的方程 (其中 $z_0 = c$) 为:

$$\frac{1}{a^2} \left(\frac{cx}{z}\right)^2 + \frac{1}{b^2} \left(\frac{cy}{z}\right)^2 = 1$$

或记为:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

因为它的准线是一个椭圆, 所以称为椭圆锥面. 当 $a = b$ 时, 它就是大家熟悉的圆锥面:

$$x^2 + y^2 - k^2 z^2 = 0$$

其中常数 $k = \frac{a}{c}$. 由于准线的参数方程是:

$$x = a \cos u, \quad y = b \sin u, \quad z = c \quad (0 \leq u < 2\pi)$$

故锥面的参数方程为:

$$\begin{cases} x = av \cos u \\ y = bv \sin u \\ z = cv \end{cases} \quad (0 \leq u < 2\pi, -\infty < v < +\infty)$$

最后, 来看一个平面曲线绕轴旋转形成的**旋转面**的方程. 要求出曲线 $L: f(y, z) = 0, x = 0$ 绕 z 轴旋转所形成的旋转面的方程, 只需在方程 $f(y, z) = 0$ 中把 y 换成 $\pm\sqrt{x^2 + y^2}$ 即可. 按这个法则, L 绕 y 轴旋转所形成的旋转面方程就是 $f(y, \pm\sqrt{x^2 + z^2}) = 0$. Oxy 面及 Ozx 面内的曲线绕同一平面内的坐标轴旋转, 所得到的旋转面的方程可用类似方法写出, 读者可自行把它们写出来.

 **练习 8.4** 求曲线 $\begin{cases} z = \sqrt{y-1} \\ x = 0 \end{cases} \quad (1 \leq y \leq 3)$ 绕 y 轴旋转所得旋转面的方程.

解 因为是绕 y 轴旋转, 所以在方程 $z = \sqrt{y-1}$ 中保留 y 不变, 将 z 换成 $\sqrt{x^2 + z^2}$, 便得所求旋转面方程为 $y = 1 + x^2 + z^2 (1 \leq y \leq 3)$, 称之为旋转抛物面.

8.1.4 五种常见的二次曲面

1. 椭球面

它的标准方程是:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a > 0, b > 0, c > 0) \quad (8.10)$$

首先, 由这方程可知:

$$|x| \leq a \quad |y| \leq b \quad |z| \leq c$$

因此, 这曲面完全限制在由 6 个平面所围成的长方体中:

$$x = \pm a \quad y = \pm b \quad z = \pm c$$

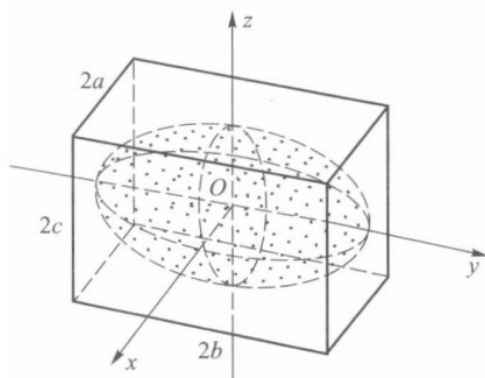


图 8.6: 椭球面

椭球面在三个坐标面上的截面显然是椭圆, 再看它在一族平行平面 $z = h$ 上的截线:

$$\begin{cases} \frac{x^2}{a^2 \left(1 - \frac{h^2}{c^2}\right)} + \frac{y^2}{b^2 \left(1 - \frac{h^2}{c^2}\right)} = 1 \\ z = h \end{cases}$$

按照椭球面的对称性和图形的范围, 只要看 $0 \leq h < c$ 的情形就行了, 这时截线总是一个椭圆, 它的两轴长分别是 $2a\sqrt{1 - \frac{h^2}{c^2}}$ 和 $2b\sqrt{1 - \frac{h^2}{c^2}}$ 。由此可见, 当 h 在 $[0, c)$ 中逐渐变大时, 椭圆由大变小, 最后缩成一点, 从这些截线的变化, 曲面的形状便大致清楚了。

椭球面的对称中心、对称轴和对称平面分别叫做它的**中心**、**主轴**和**主平面**。椭球面与 3 个坐标轴有 3 对交点, 叫做它的**顶点**。如果 $a > b > c$, 则把 a, b, c 分别称为椭球面的**半长轴**、**半中轴**和**半短轴**。当 a, b, c 中有两个相等时, 曲面为旋转椭球面; 当 $a = b = c$ 时, 即为球面。

2. 单叶双曲面

它的标准方程是:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (8.11)$$

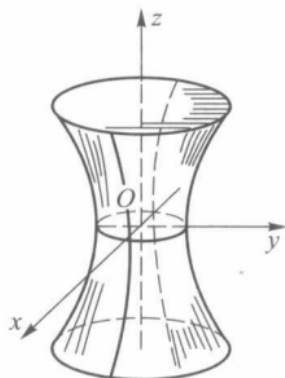


图 8.7: 单叶双曲面

这个曲面与 Ozx , Oyz 坐标面的截线都是双曲线:

$$\begin{cases} \frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 \\ y = 0 \end{cases} \quad \begin{cases} \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \\ x = 0 \end{cases}$$

这两条双曲线有共同的虚轴和虚轴长。而用平行于 Oxy 平面的平面 $z = h$ 去截它, 总有截线, 且截线总是一个椭圆:

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2} \\ z = h \end{cases}$$

当 $|h|$ 从 0 增大到 $+\infty$ 时, 椭圆的两个半轴的长度:

$$a\sqrt{1 + \frac{h^2}{c^2}}, \quad b\sqrt{1 + \frac{h^2}{c^2}}$$

也随着分别从 a, b 单调地增大到 $+\infty$, 所以单叶双曲面可以在 z 轴正、负方向无限地伸展开去。当 $a = b$ 时, 曲面成为旋转单叶双曲面。

3. 双叶双曲面

它的标准方程是:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

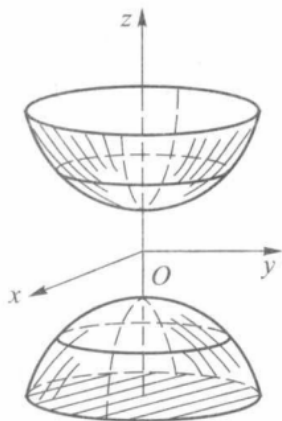


图 8.8: 双叶双曲面

这个曲面与 Ozx 和 Oyz 平面的截线都是双曲线:

$$\begin{cases} \frac{x^2}{a^2} - \frac{z^2}{c^2} = -1 \\ y = 0 \end{cases} \quad \begin{cases} \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1 \\ x = 0 \end{cases}$$

这两条双曲线有共同的实轴和实轴长。根据方程得 $|z| \geq c$, 所以曲面分成两叶 $z \geq c$ 和 $z \leq -c$ 。所以, 曲面与 Oxy 平面不相交。用平面 $z = h (|h| > c)$ 去截曲面, 截线总是椭圆 (当 $h = \pm c$ 时, 分别缩为两个点 $(0, 0, c), (0, 0, -c)$)。

当 $a = b$ 时, 曲面成为旋转双叶双曲面。



笔记 椭球面、单叶双曲面和双叶双曲面都有惟一的对称中心, 所以都叫做有心二次曲面。

4. 椭圆抛物面

它的标准方程是:

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z, \quad \text{其中常数 } p > 0, q > 0$$

显然它对称于 Ozx 面和 Oyz 面, 因而也对称于 z 轴, 但是它没有对称中心。它与对称轴的交点 (即原点)

叫做它的顶点。这个曲面与 Ozx, Oyz 平面的截线都是抛物线：

$$\begin{cases} x^2 = 2pz \\ y = 0 \end{cases} \quad \begin{cases} y^2 = 2qz \\ x = 0 \end{cases}$$

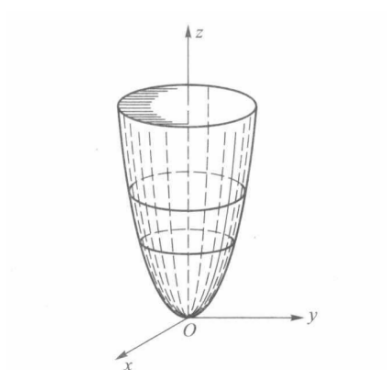


图 8.9: 椭圆抛物面

这两条抛物线有共同的顶点和轴，开口朝着同一个方向。用平行于 Oxy 平面的平面 $z = h (h > 0)$ 去截它，截线总是一个椭圆：

$$\begin{cases} \frac{x^2}{2ph} + \frac{y^2}{2qh} = 1 \\ z = h \end{cases}$$

且当 h 增大时，椭圆两个半轴的长度也随着增大，所以椭圆抛物面完全在 Oxy 平面上方，而且可以向 z 轴的正向无限地伸展开去。当 $p = q$ 时，曲面成为旋转抛物面。如果 $p < 0, q < 0$ ，则它是开口朝向 z 轴负向的椭圆抛物面。

5. 双曲抛物线

它的标准方程是：

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z, \text{ 其中常数 } p > 0, q > 0$$

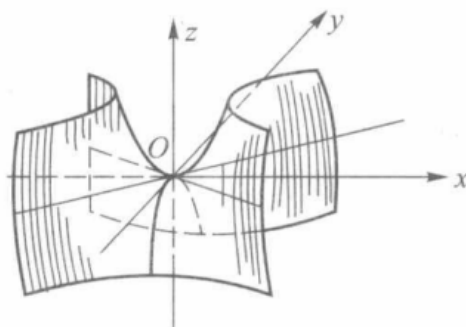


图 8.10: 双曲抛物面

显然它对称于 Ozx 面和 Oyz 面，因而也对称于 z 轴，但它没有对称中心。这个曲面与 Oxy 平面的截线是一对在原点相交的直线：

$$\begin{cases} y = \pm \sqrt{\frac{q}{p}}x, \\ z = 0. \end{cases}$$

曲面与 Ozx, Oyz 平面的截线都是抛物线：

$$\begin{cases} x^2 = 2pz \\ y = 0 \end{cases} \quad \begin{cases} y^2 = -2qz \\ x = 0 \end{cases}$$

前者开口向上, 对称轴是 z 轴的正半轴; 后者开口向下, 对称轴是 z 轴的负半轴。曲面与平面 $z = h (h \neq 0)$ 相交成双曲线:

$$\begin{cases} \frac{x^2}{2ph} - \frac{y^2}{2qh} = 1 \\ z = h \end{cases}$$

当 $h > 0$ 时, 双曲线的实轴在 Ozx 平面内且与 x 轴平行; 虚轴在 Oyz 平面内且与 y 轴平行。当 $h < 0$ 时, 双曲线的实轴和虚轴分别在 Oyz 和 Ozx 平面内。此外, 平面 $x = x_1$ 和 $y = y_1$ 与曲面的截线都是抛物线:

$$\begin{cases} y^2 = -2q \left(z - \frac{x_1^2}{2p} \right) \\ x = x_1 \end{cases} \quad \begin{cases} x^2 = 2p \left(z + \frac{y_1^2}{2q} \right) \\ y = y_1 \end{cases}$$

曲面在原点附近两边高中间低, 像个马鞍形, 所以双曲抛物面也称为**马鞍面**。



笔记 椭圆抛物面与双曲抛物面都没有对称中心, 所以都叫做无心二次曲面。

8.1.5 曲面的投影

设 Γ 是一条曲线, π 是一个平面, Γ 上每一点在 π 上都有一个垂足, 由这些垂足所形成的曲线称为 Γ 在 π 上的**投影**。经过 Γ 的每一点, 都有 π 的一条垂线, 这些垂线构成一个柱面, 称为从 Γ 到 π 的**投影柱面**。显然, Γ 在 π 上的投影就是该投影柱面与 π 的交线。

设已知曲线 Γ 的一般式方程是:

$$\begin{cases} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \end{cases}$$

现在我们来求 Γ 在 Oxy 坐标平面上的投影的方程。显然只要求出从 Γ 到 Oxy 平面的投影柱面的方程就行了。这个柱面的母线平行于 z 轴, 所以它的方程不包含 z 。因此, 问题是如何从 Γ 的方程得到一个不包含 z 的方程, 如果代表 Γ 的方程组中有一个这样的方程, 那么它就是 Γ 到 Oxy 平面的投影柱面方程。一般地, 如果从 Γ 的方程中消去 z , 得到方程:

$$\varphi(x, y) = 0$$

它表示一个母线平行于 z 轴的柱面。因为这个方程是从 Γ 的方程得到的, 所以 Γ 上点的坐标一定满足这个方程, 也就是说, 该柱面经过 Γ 。因此, 这个柱面就是 Γ 到 Oxy 平面的投影柱面, 于是 Γ 在 Oxy 平面的投影曲线的方程就是:

$$\begin{cases} \varphi(x, y) = 0 \\ z = 0 \end{cases}$$

同样地, 从 Γ 的方程中分别消去 x 或 y , 就可以得到从 Γ 到 Oyz 或 Ozx 平面的投影曲线的方程。下面给与例题加以说明。

练习 8.5 【近两年期末原题】求曲线 $C: \begin{cases} z = \sqrt{4 - x^2 - y^2} \\ x^2 + y^2 = 2y \end{cases}$ 在各坐标平面的投影曲线的方程。

解 显然 C 是上半球面 $z = \sqrt{4 - x^2 - y^2}$ 与圆柱面 $x^2 + y^2 = 2y$ 的交线, 如图 8.11:

由于 C 缠绕在圆柱面 $x^2 + y^2 = 2y$ 上, 故由 C 向 Oxy 坐标面的投影柱面就是 $x^2 + y^2 = 2y$, 于是得 C 在 Oxy 面的投影曲线方程为:

$$\begin{cases} x^2 + y^2 = 2y \\ z = 0 \end{cases} \quad (Oxy \text{ 面上的圆 } x^2 + y^2 = 2y)$$

从 C 的方程中消去 x , 得 $z = \sqrt{4-2y}$, 于是得 C 在 Oyz 面的投影曲线方程为:

$$\begin{cases} z = \sqrt{4-2y} \\ x = 0 \end{cases} \quad (0 \leq y \leq 2; Oyz \text{ 面上的抛物线 } z = \sqrt{4-2y} \text{ 的一段})$$

从 C 的方程中消去 y , 得 $x^2 + \frac{1}{4}z^4 - z^2 = 0$, 于是得 C 在 Ozx 面的投影曲线方程为:

$$\begin{cases} x^2 + \frac{1}{4}z^4 - z^2 = 0 \\ y = 0 \end{cases} \quad (|x| \leq 1, 0 \leq z \leq 2)$$

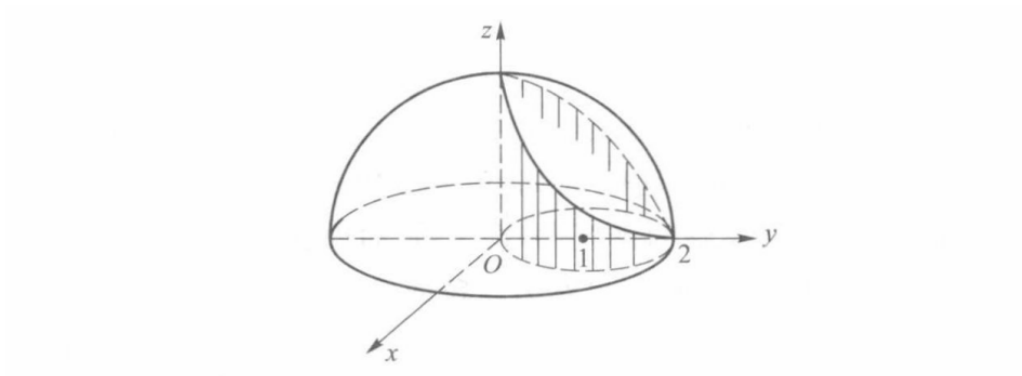


图 8.11: 立体示意图

8.2 二次型

8.2.1 二次型的表示

定义 8.4 (二次型)

称关于 n 个变量 $x_1, x_2, \dots, x_n$ 的二次齐次多项式函数:

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \dots + 2a_{1n}x_1x_n + \\ & a_{22}x_2^2 + 2a_{23}x_2x_3 + \dots + 2a_{2n}x_2x_n + \\ & \dots + a_{nn}x_n^2 \end{aligned} \quad (8.12)$$

为一个 n 元二次型, 如果其中的系数 $a_{ij} (i \leq j)$ 都是实(复)数, $x = (x_1, x_2, \dots, x_n)^T$ 是 $\mathbb{R}^n (\mathbb{C}^n)$ 中的向量, 则称它为一个 n 元实(复)二次型。根据所学内容, 下简称实二次型为二次型。

令 $a_{ji} = a_{ij}, i < j$, 由于 $x_i x_j = x_j x_i$, 所以可以写成:

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (8.13)$$

令 $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, 于是二次型可以用矩阵的乘积表示为:

$$\begin{aligned} x^T A x &= (x_1, x_2, \dots, x_n) \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= (x_1, x_2, \dots, x_n) \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{bmatrix} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \end{aligned}$$

故 $f(x_1, x_2, \dots, x_n) = x^T A x$ 。称 $A = (a_{ij})_{n \times n}$ 是二次型的矩阵, 当 $i \neq j$ 时, $a_{ji} = a_{ij}$ 正是它的 $x_i x_j$ 项的系数的一半, 而 a_{ii} 是 x_i^2 项的系数, 因此二次型和它的矩阵是相互唯一确定的。

定义 8.5 (二次型的矩阵表示)

n 元二次型的矩阵表示形式为:

$$f(x_1, x_2, \dots, x_n) = f(x) = x^T A x \quad (8.14)$$

其中 $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n, A = (a_{ij})_{n \times n}$ 为实对称矩阵。称实对称矩阵 A 为二次型 $f(x)$ 的矩阵, 并把矩阵 A 的秩称为二次型 $f(x)$ 的秩。

注 二次型对应的矩阵一定是对称矩阵, 但二次型 $f(x) = x^T A x$ 中 A 不一定是对称矩阵。当 A 不对称时, 令:

$$B = \frac{1}{2}(A + A^T) \quad (8.15)$$

此时二次型 $f(x)$ 对应的对称矩阵就是 B , 也就有 $f(x) = x^T B x$ 。

8.2.2 二次型的标准型

对于 n 元二次型 $f(x_1, x_2, \dots, x_n)$, 若令:

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n \\ \vdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases}$$

记 $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, 则上式写为:

$$x = Cy \quad (8.16)$$

其中8.16式称为线性变换。若线性变换的系数矩阵 C 可逆, 即 $|C| \neq 0$, 则称为可逆线性变换。现给出 $f(x) = x^T Ax$, 令 $x = Cy$, 则:

$$f(x) = (Cy)^T A(Cy) = y^T (C^T A C) y$$

记 $B = C^T A C$, 则:

$$f(x) = y^T B y = g(y)$$

这样, 通过可逆线性变换, 就可以对二次型进行变换。

定义 8.6 (二次型的标准型)

若二次型中只有平方项, 没有交叉项 (交叉项系数全为零), 形如:

$$f(x) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2 = x^T D x \quad (8.17)$$

称这样的二次型称为标准形。且对角矩阵 $D = \text{diag}(d_1, d_2, \dots, d_n)$ 。

二次型的标准型对于很多问题有重要意义, 因此下面来讨论如何求一个二次型的标准型。根据可逆变换的定义, 我们只需要找一个可逆矩阵 C 对应变换 $x = Cy$, 且 $C^T A C$ 为一个对角矩阵。下面来讨论如何实现这一目的:

1. 正交变换法:

在第7章我们已经知道, 对任何实对称矩阵 A , 总存在正交矩阵 P , 使得 $P^{-1}AP = P^T A P$ 成为对角矩阵。把这个结论用于二次型就有如下定理:

定理 8.1 (二次型的可逆正交变换)

对于二次型 $f(x) = x^T A x$ (其中 A 为 n 阶实对称矩阵), 总存在正交变换 $x = Py$ (P 为正交矩阵), 使得用它可将 f 化成标准形:

$$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为 A 的全部特征值。

利用实对称矩阵的正交相似对角化的方法, 就得到了用正交变换化二次型为标准形的方法 (戳习题7.5跳转到第7章相关内容处)。

 **练习 8.6** 求一个正交变换, 把下列二次型化为标准型:

$$f(x_1, x_2, x_3) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$$

解 f 的矩阵为:

$$A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$$

求一个正交变换 $x = Py$ 将 f 化为标准形, 其实质是求一个正交矩阵 P , 使得 $P^{-1}AP = P^TAP$ 成为对角矩阵。根据第 7 章的内容不难求得正交矩阵:

$$P = \begin{bmatrix} 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \end{bmatrix}$$

使得:

$$P^{-1}AP = P^TAP = \text{diag}(1, 1, 10)$$

因此, 通过正交变换 $x = Py$, 即:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = P \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

就可将 f 化为标准形 $f = y_1^2 + y_2^2 + 10y_3^2$ 的形式。

2. 配方法:

根据完全平方公式等把原二次型转化为标准型, 见下面例题。

 **练习 8.7** 用配方法将下述二次型化成标准形, 并写出相应的可逆线性变换。

$$f(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 - x_2x_3 + x_3^2$$

解 由于 f 中含变量 x_1 的平方项和交叉乘积项, 故把含 x_1 的项归并起来, 对 x_1 配方可得:

$$f = (x_1 + x_2)^2 - x_2^2 - x_2x_3 + x_3^2$$

上式右端除第 1 项外, 已不再含 x_1 , 继续对 x_2 配方可得:

$$f = (x_1 + x_2)^2 - (x_2 + \frac{1}{2}x_3)^2 + \frac{5}{4}x_3^2$$

作可逆线性变换 $y = C^{-1}x$:

$$\begin{cases} y_1 = x_1 + x_2 \\ y_2 = x_2 + \frac{1}{2}x_3 \\ y_3 = x_3 \end{cases}$$

若表示为 $x = Cy$ 的形式, 则:

$$\begin{cases} x_1 = y_1 - y_2 + \frac{1}{2}y_3 \\ x_2 = y_2 - \frac{1}{2}y_3 \\ x_3 = y_3 \end{cases} \quad (8.18)$$

就把 f 化成了标准形 $f = y_1^2 - y_2^2 + \frac{5}{4}y_3^2$, 而式 8.18 就是相应的可逆线性变换。



笔记 对于配方法, 一般从 x_1 开始一次性将含 x_1 的项配方掉, 再抓 x_2 并把含 x_2 的项配方掉, 以此类推。而如果没有平方项(如下面的例题), 就需要对其中含乘积的某两项作“一加一减”的变换, 这样就根据平方差公式创造了平方项, 因而得以继续配方。而最后的总变换就是多个变换的复合, 也就是对应变换矩阵的乘积, 而矩阵相乘的顺序遵从复合顺序。

 **练习 8.8** 用配方法化下面的二次型为标准形, 并写出相应的可逆线性变换的矩阵。

$$f(x_1, x_2, x_3) = x_1x_2 - x_2x_3$$

解 f 中不含平方项, 不好直接配方, 但 f 中含乘积项 x_1x_2 , 为了使 f 出现平方项 (然后就可利用上例的方法配方), 故先作可逆线性变换:

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases} \quad (8.19)$$

就将 f 化成为:

$$f = y_1^2 - y_2^2 - y_1y_3 + y_2y_3,$$

配方, 得:

$$f = (y_1 - \frac{1}{2}y_3)^2 - (y_2 - \frac{1}{2}y_3)^2$$

于是再作可逆线性变换:

$$\begin{cases} z_1 = y_1 - \frac{1}{2}y_3 \\ z_2 = y_2 - \frac{1}{2}y_3 \\ z_3 = y_3 \end{cases} \quad \begin{cases} y_1 = z_1 + \frac{1}{2}z_3 \\ y_2 = z_2 + \frac{1}{2}z_3 \\ y_3 = z_3 \end{cases}$$

就将 f 化成了标准形:

$$f = z_1^2 - z_2^2$$

把前面两个线性变换分别记为 $x = C_1y$ 和 $y = C_2z$, 它们的复合变换 $x = C_1y = C_1(C_2z) = (C_1C_2)z$ 就是化 f 为标准形的可逆线性变换, 其矩阵为:

$$C = C_1C_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

8.2.3 合同变换、规范型与 Sylvester 惯性定理

上一章中介绍了矩阵相似的概念, 这里给出一种新的矩阵之间的关系:

定义 8.7 (合同矩阵)

设 A, B 为 n 阶矩阵, 如果存在 n 阶可逆矩阵 C 使得:

$$C^TAC = B$$

则称矩阵 A 与 B 合同, 记为 $A \simeq B$, 并且称由 A 到 $B = C^TAC$ 的变换为合同变换。

合同也是同阶方阵之间的一种关系, 它显然具有下列简单性质:

1. 自反性: $A \simeq A$;
2. 对称性: 若 $A \simeq B$, 则 $B \simeq A$;
3. 传递性: 若 $A \simeq B$ 且 $B \simeq D$, 则 $A \simeq D$ 。
4. 等秩性: 若 $A \simeq B$, 则 $r(A) = r(B)$ 。

下面介绍一种特殊的标准型, 并在此基础上引入一重要概念:

定理 8.2 (规范型与 Sylvester 惯性定理)

对于实二次型 $f = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{nn}x_n^2$, 通过线性变换可以把其化简成唯一的规范型, 也就是系数只包含 ± 1 和 0 的标准型 $g = y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2$ 。

一个二次型的正项数称为正惯性系数 p 、负项数称为负惯性系数 $r - p$ 。正负惯性系数和系数为 0 的项数是被二次型唯一确定的, 不依赖可逆线性变换而改变。 r 为系数矩阵的秩。



Sylvester 惯性定理表明, 一个二次型的正惯性系数和负惯性系数是这个二次型的固有属性, 不随可逆线性变换而改变。进一步的有下面的推论:

命题 8.1 (特征值和惯性系数的关系)

矩阵的正(负)特征值个数等于其对应二次型的正(负)惯性系数。



要把一个二次型化为规范型, 一般只能采用配方法先化为标准型, 再做可逆线性变换把系数变成 ± 1 。

练习 8.9 设二次型:

$$f = x_1^2 + ax_2^2 + x_3^2 + 2x_1x_2 - 2x_2x_3 - 2ax_1x_3$$

的正负惯性指数都是 1, 求 f 的规范型及常数 a 。

解 二次型的矩阵为:

$$A = \begin{bmatrix} 1 & 1 & -a \\ 1 & a & -1 \\ -a & -1 & 1 \end{bmatrix}$$

因为 f 的正负惯性指数都是 1, 所以 f 的秩为 2, 1 即有 $r(A) = 2$:

$$A = \begin{bmatrix} 1 & 1 & -a \\ 1 & a & -1 \\ -a & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -a \\ 0 & a-1 & a-1 \\ 0 & 0 & (a-1)(a+2) \end{bmatrix}$$

当 $a = -2$ 时, $r(A) = 2$ 。因此对二次型配方:

$$\begin{aligned} f &= x_1^2 - 2x_2^2 + x_3^2 + 2x_1x_2 - 2x_2x_3 + 4x_1x_3 \\ &= (x_1 + x_2 + 2x_3)^2 - 3(x_2 + x_3)^2 \\ &= y_1^2 - 3y_2^2 \\ &= y_1^2 - (\sqrt{3}y_2)^2 \\ &= z_1^2 - z_2^2 \end{aligned}$$

f 的规范型即为 $z_1^2 - z_2^2$ 。事实上, 根据条件可以直接写出这个规范型。

8.2.4 正定二次型和正定矩阵**定义 8.8 (正定二次型和正定矩阵)**

设 $f(x) = x^T Ax$ 是一个 n 元二次型, 如果对任意非零向量 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$, 都有 $x^T Ax > 0$ (即对任意一组不全为零的实数 $x_1, x_2, \cdots, x_n$, 都有 $f(x_1, x_2, \cdots, x_n) > 0$), 则称 f 为正定二次型, 并称实对称矩阵 A 为正定矩阵。



注 正定二次型的正性必须要求向量 x 非零; 矩阵的正定性讨论前提条件必须是实对称矩阵。

下面介绍关于正定的重要性质:

定理 8.3 (正定的不变性)

二次型经可逆线性变换, 其正定性不变。



证明 设二次型 $f(x) = x^T A x$ 经可逆线性变换 $x = Cy$ 化成了二次型 $f(y) = y^T C^T A C y$ 。若 $x^T A x$ 正定, 即 $\forall x \in \mathbb{R}^n, x \neq 0$, 恒有 $x^T A x > 0$ 。于是, $\forall y \in \mathbb{R}^n, y \neq 0$, 有 $Cy \neq 0$, 得 $y^T C^T A C y = (Cy)^T A (Cy) > 0$, 因此二次型 $y^T C^T A C y$ 正定; 反过来也是对的。故 $x^T A x$ 与 $y^T C^T A C y$ 有相同的正定性。

从上述证明过程还能看出: 合同矩阵具有相同的正定性, 或都正定, 或都不正定。

定理 8.4 (正定的充要条件 I)

n 阶实对称矩阵 A 为正定矩阵的充要条件是 A 的所有特征值都大于零, 或者正惯性系数 $p = n$ 。



证明 设 A 为 n 阶实对称矩阵, 二次型 $f(x) = x^T A x$ 经正交变换化成的标准形为:

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

其中 $\lambda_1, \dots, \lambda_n$ 为 A 的全部特征值。要证明本定理, 根据可逆线性变换不改变正定性, 只要证明 f 的标准形正定的充要条件是 $\lambda_i > 0 (i = 1, \dots, n)$ 即可。

若 $\lambda_i > 0 (i = 1, \dots, n)$, 则对于 $y_1, \dots, y_n$ 不全为零, 不妨设 $y_1 \neq 0$, 有:

$$f \geq \lambda_1 y_1^2 > 0$$

由定义即知 f 是正定的。反之, 设 f 是正定的, 则必有 $\lambda_i > 0 (i = 1, \dots, n)$, 否则假设有某个 $\lambda_i \leq 0$, 则对于 $y_1 = 0, \dots, y_i = 1, y_{i+1} = 0, \dots, y_n = 0$, 有:

$$f = \lambda_i y_i^2 = \lambda_i \leq 0$$

这与 f 正定矛盾, 所以 $\lambda_i > 0 (i = 1, \dots, n)$ 。根据特征值和惯性系数的关系, 这也就等价于正惯性系数 $p = n$ 。

定理 8.5 (正定的充要条件 II)

实对称矩阵 A 为正定矩阵的充要条件是存在可逆矩阵 M , 使得:

$$A = M^T M \quad (8.20)$$

即表述为 A 与同阶单位矩阵 I 合同。



证明 对于必要性: 设 A 为正定矩阵, 则 A 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 均大于零, 且存在正交矩阵 P , 使得:

$$\begin{aligned} A &= P \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^T \\ &= P \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}) \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}) P^T \end{aligned}$$

在上式中令矩阵:

$$M = \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}) P^T$$

则 M 可逆, 且使 $A = M^T M$ 。

再证充分性: 如果存在可逆矩阵 M , 使得 $A = M^T M$, 则 $A^T = A$, 且 $\forall x \in \mathbb{R}^n, x \neq 0$, 有 $Mx \neq 0$ 。于是 $\forall x \in \mathbb{R}^n, x \neq 0$, 有:

$$x^T A x = x^T M^T M x = (Mx)^T (Mx) = \|Mx\|^2 > 0$$

由定义即知 A 为正定矩阵。

在式 8.20 两端取行列式, 有推论:

命题 8.2

如果 A 为正定矩阵, 则 $\det(A) > 0$ 。



在介绍判断正定的充要条件 III 之前, 先介绍顺序主子式的定义:

定义 8.9 (顺序主子式)

对于 n 阶矩阵 $A = (a_{ij})$, 称它的左上角的 r 阶主子方阵的行列式:

$$\Delta_r = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix} \quad (8.21)$$

为 A 的 r 阶顺序主子式 ($r = 1, 2, \dots, n$)。

利用顺序主子式的定义, 有:

定理 8.6 (正定的充要条件 III(Sylvester 判断))

实对称矩阵 A 为正定矩阵的充要条件是 A 的各阶顺序主子式都大于零。

 **练习 8.10** 请证明定理 8.6 的必要性。

证明 记 A_k 为矩阵 A 左上角的 k 阶矩阵, $k = 1, 2, \dots, n-1$, 由 $f = x^T A x$ 对任意 $x \neq 0$ 成立, 取 $x = (x_1, x_2, \dots, x_k, 0, \dots, 0)$, 并记 $x' = (x_1, x_2, \dots, x_k)$, 则:

$$x^T A x = x'^T A_k x' > 0$$

可以知道 A_k 也是正定矩阵。由于正定矩阵的行列式大于 0, 自然有 $|A_k| = \Delta_k > 0$ 。从而, 我们有正定矩阵 A 的顺序主子式全部大于 0。

 **练习 8.11** 试确定实数 t 的取值范围, 使得下面的二次型正定:

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + 5x_3^2 + 2tx_1x_2 - 2x_1x_3 + 4x_2x_3$$

解 f 的矩阵为:

$$A = \begin{bmatrix} 1 & t & -1 \\ t & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

A 的顺序主子式分别为:

$$\Delta_1 = 1 \quad \Delta_2 = \begin{vmatrix} 1 & t \\ t & 1 \end{vmatrix} = 1 - t^2 \quad \Delta_3 = \det(A) = -t(5t + 4)$$

已有 $\Delta_1 > 0$, 再令 $\Delta_2 > 0$, $\Delta_3 > 0$, 得 $-\frac{4}{5} < t < 0$ 。于是知当且仅当 t 满足 $-\frac{4}{5} < t < 0$ 时, f 是正定的。最后, 除了正定 (Positive Definite) 之外, 一个实对称矩阵或二次型还有其他的定性:

定义 8.10 (负定、半定和不定)

一个 n 阶实对称矩阵 A 和二次型 $x^T A x$ 称为:

- 半正定的, 如果对任意 $x \in \mathbb{R}^n, x \neq 0$, 都有 $x^T A x \geq 0$, 且存在 $x_0 \neq 0$ 使得:

$$x_0^T A x_0 = 0$$

- 负定的, 如果对任意 $x \in \mathbb{R}^n, x \neq 0$ 都有 $x^T A x < 0$;
- 半负定的, 如果对任意 $x \in \mathbb{R}^n, x \neq 0$ 都有 $x^T A x \leq 0$, 且存在 $x_0 \neq 0$ 使得:

$$x_0^T A x_0 = 0$$

- 不定的, 如果 $x^T A x$ 既能取到正值又能取到负值。

 **练习 8.12** 证明下面的二次型是半正定的:

$$n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \quad (8.22)$$

证明 首先要证明这个二次型非负, 且有非零向量使上式为 0。考虑 Cauchy-Schwarz 不等式:

$$n \sum_{i=1}^n x_i^2 = \left(\sum_{i=1}^n 1^2 \right) \left(\sum_{i=1}^n x_i^2 \right) \geq \left(\sum_{i=1}^n 1 \cdot x_i \right)^2$$

因此原二次型非负, 考虑等号成立时取 $x_i = k$, 其中 k 是个非零常数, 即这个二次型是半正定的。

8.2.5 二次曲面的标准方程

定理 8.7 (旋转变换)

设空间直角坐标系 $Oxyz: \{i, j, k\}$ 经过坐标旋转变换得到 $Ox'y'z': \{e_1, e_2, e_3\}$, 则:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (8.23)$$

其中 A 是行列式为 1 的正交矩阵。称变换 8.23 代表空间的一个旋转变换。

现在我们利用坐标系的旋转与平移来化简二次曲面方程。二次曲面的一般方程为:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0 \quad (8.24)$$

令实对称矩阵 $A = (a_{ij})_{3 \times 3} (a_{ij} = a_{ji}, i, j = 1, 2, 3)$, $x = (x, y, z)^T$, $B = (b_1, b_2, b_3)^T$, 则上式可写为:

$$x^T A x + 2B^T x + c = 0$$

可用适当的旋转变换 (正交变换) $x = Py$ 将二次型 $x^T A x$ 化为标准形 $y^T D y$ (其中 $P^T A P = D$ 为对角矩阵), 从而由此变换可将方程化为:

$$y^T D y + 2B^T y + c = 0$$

或展开为:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2b'_1 x' + 2b'_2 y' + 2b'_3 z' + c = 0$$

其中:

$$D = P^T A P = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \quad B' = B^T P = \begin{bmatrix} b'_1 \\ b'_2 \\ b'_3 \end{bmatrix} \quad y = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

这就说明了, 用坐标系的旋转可以消去 2 次方程中变量的交叉乘积项。接着作平移变换或者二次旋转变换就可以把一般的二次曲面方程化为带有如下特征的标准方程形式:

定义 8.11 (二次曲面的标准方程)

满足下面条件的二次曲面方程就叫做二次曲面的标准方程:

1. 没有交叉乘积项;
2. 如果有某个变量的平方项, 就没有它的 1 次项;
3. 如果有 1 次项, 就没有常数项;
4. 至多有 1 个 1 次项。

练习 8.13 将二次曲面方程:

$$2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3 + 4x_1 + 4x_2 = 2$$

化成标准方程, 并指出曲面的名称。

解 曲面方程可写成:

$$x^T A x + 2B^T x = 2$$

其中

$$A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad B^T = (2, 2, 0)$$

可求出行列式为 1 的正交矩阵:

$$P = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{bmatrix}$$

使得 $P^T A P = \text{diag}(-2, 4, 1) = D$ 。因此, 作正交变换 $x = Py$, 则方程化为:

$$y^T D y + 2B^T P y = 2$$

即有:

$$-2y_1^2 + 4y_2^2 + y_3^2 + 4y_1 + 4y_3 = 2$$

配方, 得:

$$-2(y_1 - 1)^2 + 4y_2^2 + (y_3 + 2)^2 = 4$$

将原点移到 $(1, 0, -2)$, 得曲面的标准方程为:

$$-\frac{x'^2}{2} + y'^2 + \frac{z'^2}{4} = 1$$

这是一个单叶双曲面。

 **练习 8.14** 将二次曲面方程 $z = xy$ 化成标准方程, 并指出曲面的名称。

解 二次型 $f(x, y, z) = xy$ 的矩阵为:

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

可求出行列式为 1 的正交矩阵:

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

使得:

$$P^T A P = \text{diag}\left(\frac{1}{2}, -\frac{1}{2}, 0\right)$$

因此, 经正交变换:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

便可将二次曲面方程 $z = xy$ 化成标准方程:

$$z' = \frac{1}{2}x'^2 - \frac{1}{2}y'^2$$

所以, 曲面为双曲抛物面。

第八章 练习

- (1) 下列实二次型中, 哪些是等价的?

$$f_1(x_1, x_2, x_3) = x_1^2 - x_2x_3$$

$$f_2(y_1, y_2, y_3) = y_1y_2 - y_3^2$$

$$f_3(z_1, z_2, z_3) = z_1z_2 + z_3^2$$

- (2) 求顶点为原点, 准线为下列椭圆的锥面方程。

$$\Gamma: \begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ z = c \end{cases}$$

- (3) 求抛物面 $y^2 + z^2 = x$ 与平面 $x + 2y - z = 0$ 的交线在各坐标平面上投影曲线的方程。

- (4) 判定下列二次型的正定性:

$$\sum_{i=1}^n x_i^2 + \sum_{i=1}^{n-1} x_i x_{i+1}$$

- (5) 设 A 为正定矩阵, 证明: 存在正定矩阵 S 使得 $A = S^2$

- (6) 设 A 是正定矩阵, c_i 为非零实常数 ($i = 1, \dots, n$), 令 $b_{ij} = a_{ij}c_ic_j$ 证明: $B = (b_{ij})_{m \times n}$ 是正定矩阵

- (7) 已知 $f(x_1, x_2, x_3) = x^T A x$ 经正交变换 $x = Qy$ 化成的标准形为 $y_1^2 + y_2^2$, 且 Q 的第 3 列为 $e_3 = \frac{1}{\sqrt{2}}(1, 0, 1)^T$ 。

求矩阵 A , 并证明 $A + I$ 为正定矩阵。

附录 A 章节练习参考题答案

A.1 第一章答案

$$(1) \text{ 写出 } D_n = \begin{vmatrix} 0 & 1 & 2 & \cdots & n-2 & n-1 \\ 1 & 0 & 1 & \cdots & n-3 & n-2 \\ 2 & 1 & 0 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-2 & n-3 & n-4 & \cdots & 0 & 1 \\ n-1 & n-2 & n-3 & \cdots & 1 & 0 \end{vmatrix};$$

此时，依次把第 $n-1$ 列乘 -1 加到第 n 列，再第 $n-2$ 列乘 -1 加到第 $n-1$ 列 ... 第 2 列乘 -1 加到第 3 列，第 1 列乘 -1 加到第 2 列可得：

$$D_n = \begin{vmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & -1 & 1 & \cdots & 1 & 1 \\ 2 & -1 & -1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-2 & -1 & -1 & \cdots & -1 & 1 \\ n-1 & -1 & -1 & \cdots & -1 & -1 \end{vmatrix} \xrightarrow[r_{ni}(1)]{i=1, \dots, n-1} \begin{vmatrix} n-1 & 0 & 0 & \cdots & 0 & 0 \\ n & -2 & 0 & \cdots & 0 & 0 \\ n+1 & -2 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 2n-3 & -2 & -2 & \cdots & -2 & 0 \\ n-1 & -1 & -1 & \cdots & -1 & -1 \end{vmatrix} = (n-1)(-2)^{n-2}(-1)$$

(2) 引入辅助 Vandermonde 行列式：

$$V_{n+1} = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & \cdots & a_n & x \\ a_1^2 & a_2^2 & \cdots & a_n^2 & x^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^{n-2} & a_2^{n-2} & \cdots & a_n^{n-2} & x^{n-2} \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} & x^{n-1} \\ a_1^n & a_2^n & \cdots & a_n^n & x^n \end{vmatrix} = \prod_{1 \leq j < i \leq n} (a_i - a_j) \prod_{k=1}^n (x - a_k) \quad (\text{A.1})$$

不难看出 D_n 是 V_{n+1} 的 $(n, n+1)$ 元素的余子式，现考虑对 V_{n+1} 的第 $n+1$ 列展开，则 V_{n+1} 可以看做是关于 x 的 n 次多项式，而 x^k 项的系数正是第 $(k+1, n+1)$ 元素的代数余子式。

现要求式子 A.1 中的 x^{n-1} 项的系数，因为：

$$\prod_{k=1}^n (x - a_k) = x^n - \left(\sum_{k=1}^n a_k \right) x^{n-1} + \cdots + (-1)^n \prod_{k=1}^n a_k$$

所以有：

$$(-1)^{n+(n+1)} D_n = - \left(\sum_{k=1}^n a_k \right) \left(\prod_{1 \leq j < i \leq n} (a_i - a_j) \right)$$

消去负号，故即证：

$$D_n = \left(\sum_{i=1}^n a_i \right) \left(\prod_{1 \leq j < i \leq n} (a_i - a_j) \right)$$

(3) 由题目，把 A 和 B 看做矩阵，其第 j 列元素组成的列向量分别为 α_j 和 β_j ，则：

$$\beta_j = \alpha_j - \sum_{k \neq j} \alpha_k$$

按照矩阵的计算法则，有：

$$B = AC = A \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ -1 & 1 & -1 & \cdots & -1 \\ -1 & -1 & 1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix}$$

所以有 $\det(B) = \det(AC) = \det(A) \det(C) = 3 \det(C)$ ，利用爪型行列式的计算方法不难有 $\det(C) = (2-n)2^{n-1}$ 。
故 $\det(B) = 3(2-n)2^{n-1}$

(4) 先原行列式展开：

$$D = x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3 = -p(x_1 + x_2 + x_3) - 3q - 3x_1x_2x_3 \quad (\text{A.2})$$

对于一般的三次方程 $ax^3 + bx^2 + cx + d = 0 (a \neq 0)$ 而言，其可以写成 $a(x-x_1)(x-x_2)(x-x_3) = 0$ 的形式，将该式子展开有：

$$ax^3 - a(x_1 + x_2 + x_3)x^2 + a(x_1x_2 + x_1x_3 + x_2x_3)x - ax_1x_2x_3 = 0$$

与原方程系数对比，有 $x_1 + x_2 + x_3 = -\frac{b}{a}$ 且 $x_1x_2x_3 = -\frac{d}{a}$ ，代入条件有 $x_1 + x_2 + x_3 = 0$ 且 $x_1x_2x_3 = -q$ 。再代入式子 A.2 得到 $D = 0$ 。

(5) 考虑 Pascal 行列式：

$$D_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 & C_n^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 & C_{n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} & C_{2n-3}^{n-2} \\ 1 & C_n^{n-1} & C_{n+1}^{n-1} & \cdots & C_{2n-3}^{n-1} & C_{2n-2}^{n-1} \end{vmatrix}$$

我们使用数学归纳法证明 $D_n = 1$ 。当 $n = 1$ 时， $D_1 = |1| = 1$ ，等式成立。

当 $n = 2$ 时， $D_2 = \begin{vmatrix} 1 & 1 \\ 1 & C_2^1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \cdot 2 - 1 \cdot 1 = 1$ 等式成立。

假设对于 $n = k$ ， $D_k = 1$ 成立。对于 $n = k+1$ ，我们通过行列式的性质进行变换。从最后一行起依次减去上一行，利用组合数恒等式 $C_{n+1}^m - C_n^m = C_n^{m-1}$ ，得到：

$$D_{k+1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & C_1^1 & C_2^1 & \cdots & C_{k-1}^1 & C_k^1 \\ 0 & C_2^2 & C_3^2 & \cdots & C_k^2 & C_{k+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & C_{k-1}^{k-1} & C_k^{k-1} & \cdots & C_{2k-3}^{k-1} & C_{2k-2}^{k-1} \\ 0 & C_k^k & C_{k+1}^k & \cdots & C_{2k-2}^k & C_{2k-1}^k \end{bmatrix}$$

按照第一列展开，并从最右侧列依次减去左一列，不难得到 $D_{k+1} = D_k$ ，根据归纳假设， $D_k = 1$ ，所以 $D_{k+1} = 1$ 。因此，对于所有正整数 n 都有 $D_n = 1$ 。

考虑修改后的 Pascal 行列式:

$$E_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 & C_n^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 & C_{n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} & C_{2n-3}^{n-2} \\ 1 & C_n^{n-1} & C_{n+1}^{n-1} & \cdots & C_{2n-3}^{n-1} & C_{2n-2}^{n-1} - 1 \end{vmatrix}$$

我们可以将 E_n 表示为 D_n 减去一个行列式, 其中只有最后一行最后一列的元素不同。具体地, 设 F_n 是与 D_n 相同的行列式, 但最后一行最后一列的元素为 -1 , 其他元素与 D_n 相同。那么 $E_n = D_n + F_n$ 。

根据行列式的线性性质, F_n 可以表示为:

$$F_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 & C_n^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 & C_{n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} & C_{2n-3}^{n-2} \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{vmatrix}$$

按最后一行展开, 得到:

$$F_n = (-1)^{n+n} \cdot (-1) \cdot \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & C_2^1 & C_3^1 & \cdots & C_{n-1}^1 \\ 1 & C_3^2 & C_4^2 & \cdots & C_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & C_{n-1}^{n-2} & C_n^{n-2} & \cdots & C_{2n-4}^{n-2} \end{vmatrix} = -D_{n-1}$$

因此, $E_n = D_n + F_n = D_n - D_{n-1}$ 。根据我们已经证明的第一个等式, $D_n = 1$ 和 $D_{n-1} = 1$, 所以:

$$E_n = 1 - 1 = 0$$

因此, 第二个等式得证。

(6) 设 $\log_{36} 45 = c$, 则可得

$$\begin{cases} a \ln 2 + 2(a-1) \ln 3 = 0 \\ b \ln 2 + 2b \ln 3 - \ln 5 = 0 \\ 2c \ln 2 + 2(c-1) \ln 3 - \ln 5 = 0 \end{cases}$$

将其看为是有非零解为 $(\ln 2, \ln 3, \ln 5)$ 的齐次线性方程组, 于是由 Cramer 法则得:

$$\begin{vmatrix} a & 2(a-1) & 0 \\ b & 2b & -1 \\ 2c & 2(c-1) & -1 \end{vmatrix} = 0$$

展开得 $ac + a + b - 2c = 0$, 所以

$$\log_{36} 45 = c = \frac{a+b}{2-a}$$

(7) 按最后一列展开, 在 $A_{1,n+1}$ 中从最后一行向上逐次互换移到第 1 行的位置得:

$$D_{n+1} = (-1)^{(n+1)+1+(n-1)} D_n + a_1 a_2 a_3 \cdots a_n$$

利用这个递推关系得:

$$\begin{aligned} D_{n+1} &= a_1 a_2 \cdots a_{n-1} a_n - a_1 a_2 \cdots a_{n-1} + a_1 a_2 \cdots a_{n-2} - \cdots + (-1)^{n-1} a_1 + (-1)^n \\ &= \sum_{k=1}^n \left((-1)^{n-k} \prod_{p=1}^k a_p \right) + (-1)^n \end{aligned}$$

(8) 首先从第 1 列提取公因子 a_1 , 然后从第 n 行提取公因子 b_n , 得

$$D_n = a_1 b_n \begin{vmatrix} b_1 & a_1 b_2 & a_1 b_3 & \cdots & a_1 b_n \\ b_2 & a_2 b_2 & a_2 b_3 & \cdots & a_2 b_n \\ b_3 & a_2 b_3 & a_3 b_3 & \cdots & a_3 b_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_2 & a_3 & \cdots & a_n \end{vmatrix}$$

用最后一行消去第一列的 b_i , 再按照第一列展开有:

$$\begin{aligned} D_n &= a_1 b_n \begin{vmatrix} 0 & a_1 b_2 - a_2 b_1 & a_1 b_3 - a_3 b_1 & \cdots & a_1 b_n - a_n b_1 \\ 0 & 0 & a_2 b_3 - a_3 b_2 & \cdots & a_2 b_n - a_n b_2 \\ 0 & 0 & 0 & \cdots & a_3 b_n - a_n b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_2 & a_3 & \cdots & a_n \end{vmatrix} \\ &= a_1 b_n (-1)^{n+1} (a_1 b_2 - a_2 b_1) (a_2 b_3 - a_3 b_2) \cdots (a_{n-1} b_n - a_n b_{n-1}) \\ &= (-1)^{n+1} a_1 b_n \prod_{i=1}^{n-1} (a_i b_{i+1} - a_{i+1} b_i) \end{aligned}$$

(9) 常数项即 $p(0)$ 的值, 代入有 $p(0) = 3$ 。

(10) 按第一行展开行列式 D_n , 得递推关系:

$$D_n = (\alpha + \beta) D_{n-1} - \alpha \beta D_{n-2}$$

其中初始条件为:

$$D_1 = \alpha + \beta \quad D_2 = (\alpha + \beta)^2 - \alpha \beta = \alpha^2 + \alpha \beta + \beta^2$$

对于该二阶线性递推关系, 采用特征根法求得满足 $D_n - x_1 D_{n-1} = x_2 (D_{n-1} - x_1 D_{n-2})$ 的 $x_1 = \alpha$ 和 $x_2 = \beta$ 。当 $\alpha \neq \beta$ 时, 有 $D_n = C_1 \alpha^n + C_2 \beta^n$, 代入初始值有 $C_1 = \frac{\alpha}{\alpha - \beta}$, $C_2 = -\frac{\beta}{\alpha - \beta}$ 因此:

$$D_n = \frac{\alpha}{\alpha - \beta} \cdot \alpha^n - \frac{\beta}{\alpha - \beta} \cdot \beta^n$$

上述式子不适用 $\alpha = \beta$ 的情况, 一般的有:

$$D_n = \alpha^n + \alpha^{n-1} \beta + \alpha^{n-2} \beta^2 + \cdots + \alpha \beta^{n-1} + \beta^n = \sum_{k=0}^n \alpha^{n-k} \beta^k$$

(11) 本题实际上是 Vandermonde 行列式的一个推广形式。我们可以通过列变换将其转化为标准的 Vandermonde 行列式。从第 j 列 ($j = 1, 2, \dots, n$) 中提取 a_j^{n-1} 。提取后, 行列式变为:

$$D_n = \left(\prod_{j=1}^n a_j^{n-1} \right) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \frac{b_1}{a_1} & \frac{b_2}{a_2} & \frac{b_3}{a_3} & \cdots & \frac{b_n}{a_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \left(\frac{b_1}{a_1} \right)^{n-2} & \left(\frac{b_2}{a_2} \right)^{n-2} & \left(\frac{b_3}{a_3} \right)^{n-2} & \cdots & \left(\frac{b_n}{a_n} \right)^{n-2} \\ \left(\frac{b_1}{a_1} \right)^{n-1} & \left(\frac{b_2}{a_2} \right)^{n-1} & \left(\frac{b_3}{a_3} \right)^{n-1} & \cdots & \left(\frac{b_n}{a_n} \right)^{n-1} \end{vmatrix}$$

这变成了一个 Vandermonde 行列式, 得到:

$$D_n = \left(\prod_{j=1}^n a_j^{n-1} \right) \prod_{1 \leq i < j \leq n} \left(\frac{b_j}{a_j} - \frac{b_i}{a_i} \right)$$

为了让表达式更简洁, 我们可以对后面的乘积项进行通分:

$$\frac{b_j}{a_j} - \frac{b_i}{a_i} = \frac{a_i b_j - a_j b_i}{a_i a_j}$$

将其代入原式:

$$D_n = \left(\prod_{j=1}^n a_j^{n-1} \right) \prod_{1 \leq i < j \leq n} \left(\frac{a_i b_j - a_j b_i}{a_i a_j} \right) = \left(\prod_{j=1}^n a_j^{n-1} \right) \frac{\prod_{1 \leq i < j \leq n} (a_i b_j - a_j b_i)}{\prod_{1 \leq i < j \leq n} (a_i a_j)}$$

来计算分母中 a_j 的总幂次。对于固定的 a_j , 当 i 从 1 到 $j-1$ 时它在分母 $\prod_{1 \leq i < j \leq n} (a_i a_j)$ 中出现了 $(j-1)$ 次; 当 j 从 $i+1$ 到 n 的过程中, $\prod_{1 \leq i < j \leq n} (a_i a_j)$ 中出现了 $(n-j)$ 次。所以 a_j 的总幂次是 $(j-1) + (n-j) = n-1$ 次。因此, 分母部分就是 $\prod_{j=1}^n a_j^{n-1}$ 。这正好与第一步提取的公因子 $\prod_{j=1}^n a_j^{n-1}$ 完全抵消。所以:

$$D_n = \prod_{1 \leq i < j \leq n} (a_i b_j - a_j b_i)$$

(12) 写出 D 的形式:

$$D = \begin{vmatrix} a_1 + b_1 & a_1 + b_2 & \cdots & a_1 + b_n \\ a_2 + b_1 & a_2 + b_2 & \cdots & a_2 + b_n \\ \vdots & \vdots & & \vdots \\ a_n + b_1 & a_n + b_2 & \cdots & a_n + b_n \end{vmatrix}$$

将第一列分别减到剩下的列上, 有:

$$D = \begin{vmatrix} a_1 + b_1 & b_2 - b_1 & \cdots & b_n - b_1 \\ a_2 + b_1 & b_2 - b_1 & \cdots & b_n - b_1 \\ \vdots & \vdots & & \vdots \\ a_n + b_1 & b_2 - b_1 & \cdots & b_n - b_1 \end{vmatrix}$$

显然, 当 $n > 2$ 时, 行列式的值为 0; 当 $n = 2$ 时, 为 $(a_1 - a_2)(b_2 - b_1)$; 当 $n = 1$ 时, 为 $a_1 + b_1$ 。

A.2 第二章答案

(1) (a). 由伴随矩阵的定义: $AA^* = A^*A = \det(A)I$ 。若 A 可逆, 则 $\det(A) \neq 0$, 于是:

$$A^* = \det(A)A^{-1}$$

从而:

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

最后一个等式是因为 $(A^{-1})^* = \det(A^{-1})(A^{-1})^{-1} = \frac{1}{\det(A)} A$ 。

(b). 由定义, $(kA)^*$ 的元素是 kA 的代数余子式, 而 kA 的每个代数余子式是 A 对应代数余子式的 k^{n-1} 倍, 因为每行提取因子 k 后, 余子式是 $n-1$ 阶行列式, 提出 k^{n-1} 。所以 $(kA)^* = k^{n-1}A^*$ 。

(c). 由 $AA^* = \det(A)I$, 若 A 可逆, 则 A^* 可逆, 且:

$$(A^*)^* = \det(A^*)(A^*)^{-1}$$

而 $\det(A^*) = [\det(A)]^{n-1}$, 且 $(A^*)^{-1} = \frac{1}{\det(A)} A$, 代入得:

$$(A^*)^* = [\det(A)]^{n-1} \cdot \frac{1}{\det(A)} A = [\det(A)]^{n-2} A$$

(d). 当 A, B 可逆时, 有:

$$(AB)^* = \det(AB)(AB)^{-1} = \det(A) \det(B) B^{-1} A^{-1}$$

而:

$$B^* A^* = \det(B) B^{-1} \cdot \det(A) A^{-1} = \det(A) \det(B) B^{-1} A^{-1}$$

所以 $(AB)^* = B^* A^*$ 。

(2) 记 n 阶单位矩阵的第 j 列列向量为 ε_j , 则一个方阵乘以 ε_j 得到该方阵的第 j 列列向量, 故取 $x = \varepsilon_j (j = 1, 2, \dots, n)$ 立刻得到 $A = B$ 。

(3) 由逆矩阵的定义有:

$$\begin{aligned} I &= AB = (I - \alpha\alpha^T)(I + \frac{1}{a}\alpha\alpha^T) \\ &= I + \frac{1}{a}\alpha\alpha^T - \alpha\alpha^T - \frac{1}{a}\alpha\alpha^T\alpha\alpha^T \\ &= I + \frac{1}{a}\alpha\alpha^T - \alpha\alpha^T - \frac{1}{a}\alpha(\alpha^T\alpha)\alpha^T \\ &= I + (\frac{1}{a} - 1 - 2a)\alpha\alpha^T \end{aligned}$$

因此有:

$$I + (\frac{1}{a} - 1 - 2a)\alpha\alpha^T = O$$

所以只能有 $\frac{1}{a} - 1 - 2a = 0$, 结合 $a < 0$ 解得 $a = -1$ 。

(4) 由 $A^* = A^T$ 可知 $A_{ij} = a_{ij}$, 也就不难得到 $A_{11} = a_{11} = a$, $A_{12} = a_{12} = a$, $A_{13} = a_{13} = a$ 。

因为:

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = 3a^2$$

而 $\det(A)I = AA^* = AA^T$, 即:

$$\begin{bmatrix} \det(A) & & \\ & \det(A) & \\ & & \det(A) \end{bmatrix} = AA^T$$

两边取行列式有 $[\det(A)]^3 = [\det(A)]^2$, 从而可知 $\det(A) = 1$ 或者 $\det(A) = 0$ 。由 $a > 0$, 知 $\det(A) = 3a^2 = 1$, 故 $a = \frac{\sqrt{3}}{3}$ 。

(5) 直接利用逆矩阵的定义:

$$\begin{aligned} &(A + \alpha\beta^T) \left(A^{-1} - \frac{A^{-1}\alpha\beta^TA^{-1}}{1 + \beta^TA^{-1}\alpha} \right) \\ &= I - \frac{\alpha\beta^TA^{-1}}{1 + \beta^TA^{-1}\alpha} + \alpha\beta^TA^{-1} - \frac{\alpha\beta^TA^{-1}\alpha\beta^TA^{-1}}{1 + \beta^TA^{-1}\alpha} \\ &= I + \alpha\beta^TA^{-1} - \frac{\alpha\beta^TA^{-1}}{1 + \beta^TA^{-1}\alpha} - (\beta^TA^{-1}\alpha) \frac{\alpha\beta^TA^{-1}}{1 + \beta^TA^{-1}\alpha} \\ &= I + \alpha\beta^TA^{-1} - \alpha\beta^TA^{-1} \\ &= I \end{aligned}$$

其中, 提取出来 $\beta^TA^{-1}\alpha$ 的依据是该因子是一个确定的数而非矩阵。

(6)

$$C^* = |C|C^{-1} = |A||B| \begin{bmatrix} A^{-1} & O \\ O & B^{-1} \end{bmatrix} = \begin{bmatrix} |A||B|A^{-1} & O \\ O & |A||B|B^{-1} \end{bmatrix} = \begin{bmatrix} |B|A^* & O \\ O & |A|B^* \end{bmatrix}$$

(7) 算有 $|A| = 2$, 则:

$$\left(\frac{1}{2}A\right)^* = \left|\frac{1}{2}A\right| \left(\frac{1}{2}A\right)^{-1} = \left(\frac{1}{2}\right)^4 |A| 2A^{-1} = \frac{1}{4}A^{-1}$$

而 $\left(\frac{1}{4}A^{-1}\right)^{-1} = 4A$, 故题干式同左乘 A^{-1} 并右乘 A 有:

$$A^{-1}4ABA^{-1}A = A^{-1}2ABA + A^{-1}12IA \implies 4B = 2BA + 12I$$

所以有:

$$B = 6(2I - A)^{-1} = 6 \begin{bmatrix} 1 & -2 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -4 & 0 & 0 \\ -2 & -2 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

$$(8) \alpha^T \alpha = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \text{ 设非零常数 } k, \text{ 有 } B = I - \frac{1}{k} \alpha^T \alpha = \begin{bmatrix} 1 - \frac{1}{k} & \cdots & -\frac{1}{k} \\ \vdots & \ddots & \vdots \\ -\frac{1}{k} & \cdots & 1 - \frac{1}{k} \end{bmatrix} \text{ 可以计算 } \det(B) = 1 - \frac{1}{k} n$$

因此 $|A| = 0$, 且 A 左上角的 $n-1$ 阶子式 $D_{n-1} = 1 - \frac{1}{n}(n-1) \neq 0$, 故 $r(A) = n-1$ 。

(9) 将 $\begin{vmatrix} A & \beta \\ \beta^T & p \end{vmatrix}$ 按最后一列展开, 注意到展开式中含 p 项的系数为 $\det(A)$, 而 $\begin{vmatrix} A & \beta \\ \beta^T & q \end{vmatrix}$ 和前式相比只有 $p \det(A)$ 和 $q \det(A)$ 一项之差, 移项有原式值为 $a(q-p)$ 。

(10) 要证明一个矩阵 M 可逆, 并且其逆矩阵是 N , 我们只需要验证 $MN = I$ (或 $NM = I$) 即可:

(a).

$$\begin{aligned} MN &= [(A^{-1} + B^{-1})][A(A+B)^{-1}B] \\ &= (A^{-1}A)(A+B)^{-1}B + B^{-1}A(A+B)^{-1}B \\ &= (A+B)^{-1}B + B^{-1}A(A+B)^{-1}B \\ &= (A+B)^{-1}B + \{B^{-1}[(A+B) - B]\}(A+B)^{-1}B \\ &= (A+B)^{-1}B + [B^{-1}(A+B) - I](A+B)^{-1}B \\ &= (A+B)^{-1}B + B^{-1}(A+B)(A+B)^{-1}B - I(A+B)^{-1}B \\ &= (A+B)^{-1}B + I - (A+B)^{-1}B \\ &= I \end{aligned}$$

第一部分得证。

(b). 从第一部分知道 $(A^{-1} + B^{-1})^{-1} = A(A+B)^{-1}B$, 交换左式子 A^{-1} 和 B^{-1} 的位置就有:

$$A(A+B)^{-1}B = (A^{-1} + B^{-1})^{-1} = (B^{-1} + A^{-1})^{-1} = B(B+A)^{-1}A = B(A+B)^{-1}A$$

第二部分得证。

(11)

$$\begin{aligned} \begin{bmatrix} I & -B \\ A & I \end{bmatrix} \begin{bmatrix} I & B \\ O & I \end{bmatrix} &= \begin{bmatrix} I & O \\ A & I+AB \end{bmatrix} \\ \begin{bmatrix} I & B \\ O & I \end{bmatrix} \begin{bmatrix} I & -B \\ A & I \end{bmatrix} &= \begin{bmatrix} I+BA & O \\ A & I \end{bmatrix} \end{aligned}$$

对以上两式两端分别取行列式得:

$$\begin{vmatrix} I & O \\ A & I+AB \end{vmatrix} = |I+AB|, \quad \begin{vmatrix} I+BA & O \\ A & I \end{vmatrix} = |I+BA|$$

故 $|I+AB| = |I+BA|$, 结论得证。

A.3 第三章答案

(1) 先找出过 M 点且平行于平面 π 的平面 π_1 , 则所求直线在平面 π_1 上, 如果再求出直线 L 与平面 π_1 的交点 M_1 , 则连接 M 和 M_1 的直线与直线 L 相交且平行于平面 π 。

过点 $M(-1, 0, 4)$ 且平行于平面 $\pi: 3x - 4y + z - 10 = 0$ 的平面方程为:

$$\pi_1: 3(x+1) - 4y + (z-4) = 0$$

直线 $L: x+1 = y-3 = \frac{z}{2}$ 与平面 $\pi_1: 3x - 4y + z - 1 = 0$ 的交点为 $M_1(15, 19, 32)$, 则所求直线为:

$$\frac{x+1}{16} = \frac{y}{19} = \frac{z-4}{28}$$

(2) 所求点位于 z 轴, 设其坐标为 $C(0, 0, z)$, 由向量的几何意义知 $S_{\triangle ABC} = \frac{1}{2} |\vec{AB} \times \vec{AC}|$,

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} i & j & k \\ -1 & 2 & 1 \\ -1 & 0 & z \end{vmatrix} = (2z, z-1, 2)$$

故 $S_{\triangle ABC} = \frac{1}{2} \sqrt{(2z)^2 + (z-1)^2 + 2^2} = \frac{1}{2} \sqrt{5z^2 - 2z + 5}$, 当 $z = \frac{1}{5}$, 即取 z 轴上一点 $C\left(0, 0, \frac{1}{5}\right)$, 可使得 $\triangle ABC$ 面积最小。

(3) 设直线 L 与 xOy 平面的交点为 $(x, y, 0)$, 由于直线 L 在平面 π 上, 故 $2x+3y=9$, 又由于 l 与 xOy 平面有最大交角, 则 $s = (x-1, y-1, -1)$ 与 $k = (0, 0, 1)$ 的夹角最小, 即求 $\max: |\cos(s, k)| = \frac{1}{\sqrt{(x-1)^2 + (y-1)^2 + 1}}$ 。

读者除了将 $2x+3y=9$ 代入求最值外, 还可以利用拉格朗日乘数法, 设:

$$f(x, y) = (x-1)^2 + (y-1)^2 + 1 + \lambda(2x+3y-9)$$

由 $f_x = 2(x-1) + 2\lambda = 0$, $f_y = 2(y-1) + 3\lambda = 0$ 及 $2x+3y=9$ 解得 $x = \frac{21}{13}$, $y = \frac{25}{13}$ 。所以直线方向为 $s = \left(\frac{8}{13}, \frac{12}{13}, -1\right) = \frac{1}{13}(8, 12, -13)$, 所求直线为:

$$\frac{x-1}{8} = \frac{y-1}{12} = \frac{z-1}{-13}$$

(4) 因为 $\cos(a+b, a-b) = \frac{(a+b) \cdot (a-b)}{\|a+b\| \|a-b\|}$, 其中:

$$a \cdot b = \|a\| \|b\| \cos \frac{\pi}{6} = \frac{3}{2} \quad (a+b) \cdot (a-b) = \|a\|^2 - \|b\|^2 = 2$$

$$\|a+b\| = \sqrt{(a+b) \cdot (a+b)} = \sqrt{\|a\|^2 + \|b\|^2 + 2a \cdot b} = \sqrt{7}$$

$$\|a-b\| = \sqrt{(a-b) \cdot (a-b)} = \sqrt{\|a\|^2 + \|b\|^2 - 2a \cdot b} = 1$$

所以:

$$\cos(a+b, a-b) = \frac{2}{\sqrt{7}} \quad (a+b, a-b) = \arccos \frac{2}{\sqrt{7}}$$

(5) 设所求直线的方向为 $s = (m, n, p)$, 因直线与 x 轴夹角为 $\frac{\pi}{3}$, 则:

$$\frac{s \cdot i}{\|s\| \|i\|} = \cos \frac{\pi}{3} = \frac{1}{2}, \text{ 即 } \frac{|m|}{\sqrt{m^2 + n^2 + p^2}} = \frac{1}{2}$$

已有直线过 $M_2(7, 3, 5)$, 方向 $s_1 = (1, 2, 2)$, 因所求直线与已知直线相交, 故向量 $s, s_1, \vec{M_1 M_2}$ 共面, 则:

$$\begin{vmatrix} m & n & p \\ 1 & 2 & 2 \\ 5 & 6 & 6 \end{vmatrix} = 0, \text{ 即 } n - p = 0$$

所以可解得 $m = \pm \frac{\sqrt{6}}{3} p$, $n = p$, 即得 $s = p \left(\pm \frac{\sqrt{6}}{3}, 1, 1\right)$, 所求直线为:

$$\frac{x-2}{\pm \frac{\sqrt{6}}{3}} = \frac{y+3}{1} = \frac{z+1}{1} \text{ 或表示为 } \frac{x-2}{\sqrt{6}} = \frac{y+3}{\pm 3} = \frac{z+1}{\pm 3}$$

A.4 第四章答案

(1) 表示法可能不唯一:

(a).

$$x = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix} + k_1 \begin{bmatrix} -9 \\ 1 \\ 7 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

(b).

$$x = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + k_1 \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + k_3 \begin{bmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + k_4 \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

(c).

$$x = \begin{bmatrix} 3 \\ 1 \\ -2 \\ 0 \end{bmatrix} + k \begin{bmatrix} 5 \\ -2 \\ -1 \\ 3 \end{bmatrix}$$

(2) 只要证 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关。假如 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性相关, 则有一组不全为 0 的数 $k_1, k_2, \dots, k_n$, 使得:

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n = 0$$

不妨设:

$$|k_l| = \max \{|k_1|, |k_2|, \dots, |k_n|\}$$

在上式中考虑第 l 个分量的等式:

$$k_1 \alpha_{l1} + k_2 \alpha_{l2} + \dots + k_l \alpha_{ll} + \dots + k_n \alpha_{ln} = 0$$

也就得:

$$\alpha_{ll} = -\frac{k_1}{k_l} \alpha_{l1} - \dots - \frac{k_{l-1}}{k_l} \alpha_{l,l-1} - \frac{k_{l+1}}{k_l} \alpha_{l,l+1} - \dots - \frac{k_n}{k_l} \alpha_{ln}$$

或者:

$$= -\sum_{\substack{j=1 \\ j \neq l}}^n \frac{k_j}{k_l} \alpha_{lj}$$

因此:

$$|\alpha_{ll}| \leq \sum_{\substack{j=1 \\ j \neq l}}^n \frac{|k_j|}{|k_l|} |\alpha_{lj}| \leq \sum_{\substack{j=1 \\ j \neq l}}^n |\alpha_{lj}|$$

这与已知条件矛盾。因此 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关。从而它的秩等于 n 。(3) 对系数矩阵 A 作初等行变换:

$$\begin{bmatrix} 1 & 2 & -11 \\ 2 & -5 & 3 \\ 5 & 1 & a \\ 6 & 3 & b \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -11 \\ 0 & -9 & 25 \\ 0 & -9 & 55+a \\ 0 & -9 & 66+b \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -11 \\ 0 & -9 & 25 \\ 0 & 0 & 30+a \\ 0 & 0 & 41+b \end{bmatrix}$$

当 $a = -30$ 且 $b = -41$ 时, $r(A) = 2 < 3$, 此时齐次线性方程组有非零解; 当 $a \neq -30$ 或 $b \neq -41$ 时, $\text{rank}(A) = 3$, 此时齐次线性方程组只有零解。

- (4) 在 $\alpha_1, \dots, \alpha_i$ 中任取一个向量 α_j 。由于 α_j 可以由 $\alpha_{i_1}, \dots, \alpha_{i_r}$ 唯一地线性表出, 因此 $\alpha_{i_1}, \dots, \alpha_{i_r}$ 线性无关。在其余向量中任取一个 α_i 添进去, 由于 α_i 可以由 $\alpha_{i_1}, \dots, \alpha_{i_r}$ 线性表出, 因此 $\alpha_{i_1}, \dots, \alpha_{i_r}, \alpha_i$ 线性相关。从而 $\alpha_{i_1}, \dots, \alpha_{i_r}$ 是 $\alpha_1, \dots, \alpha_i$ 的一个极大线性无关组。
- (5) $\beta_1, \beta_2, \beta_3$ 是 $\beta_1, \beta_2, \beta_3, \beta_4$ 的一个极大线性无关组。
- (6) 在所给的齐次线性方程组的下面添上一个方程:

$$0x_1 + 0x_2 + \dots + 0x_n = 0$$

得到 n 个方程的 n 元齐次线性方程组, 其系数矩阵 A 为:

$$A = \begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}$$

A 的 (n, j) 元的代数余子式为:

$$A_{nj} = (-1)^{n+j} D_j \quad (j = 1, 2, \dots, n)$$

原齐次线性方程组的第 i 个方程 ($i = 1, 2, \dots, n-1$) 为:

$$b_{i1}x_1 + b_{i2}x_2 + \dots + b_{in}x_n = 0$$

对 $|A|$ 用行列式按一行展开定理, 得:

$$b_{i1}A_{n1} + b_{i2}A_{n2} + \dots + b_{in}A_{nn} = 0 \quad (i = 1, 2, \dots, n-1)$$

由此得出 $\eta = (D_1, -D_2, \dots, (-1)^{n-1}D_n)^T$ 是原齐次线性方程组的一个解。如果 $\eta \neq 0$, 那么 B 有一个 $n-1$ 阶子式不为 0。从而 $r(B) \geq n-1$ 。又由于 B 只有 $n-1$ 行, 因此 $r(B) = n-1$ 。由于 $\eta = (D_1, -D_2, \dots, (-1)^{n-1}D_n)^T$ 是原齐次线性方程组的一个非零解, 因此 η 是原齐次线性方程组的一个基础解系。

A.5 第五章答案

- (1) 采用数学归纳法。 $n=0$ 时, 1 线性无关; 假设 $n-1$ 时命题为真, 来看 n 的情形。设:

$$k_0 1 + k_1 \cos x + k_2 \cos 2x + \dots + k_n \cos nx = 0$$

两边求 1 阶导数和 2 阶导数得:

$$-k_1 \sin x - 2k_2 \sin 2x - \dots - nk_n \sin nx = 0$$

$$-k_1 \cos x - 4k_2 \cos 2x - \dots - n^2 k_n \cos nx = 0$$

第一个式子两边乘 n^2 , 与第三个式子相加得:

$$n^2 k_0 1 + (n^2 - 1)k_1 \cos x + (n^2 - 4)k_2 \cos 2x + \dots + [n^2 - (n-1)^2]k_{n-1} \cos(n-1)x = 0$$

根据归纳假设得:

$$n^2 k_0 = 0, (n^2 - 1)k_1 = 0, (n^2 - 4)k_2 = 0, \dots, [n^2 - (n-1)^2]k_{n-1} = 0$$

从而 $k_0 = k_1 = k_2 = \dots = k_{n-1} = 0$, 即得 $k_n \cos nx = 0$ 。由于 $\cos nx$ 不是零函数, 因此 $k_n = 0$, 从而 $1, \cos x, \cos 2x, \dots, \cos(n-1)x, \cos nx$ 线性无关。

- (2) 设 A 有 n 个不同的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则 A 可对角化。从而存在 n 阶可逆矩阵 P , 使得 $P^{-1}AP = D$, 其中 $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 。于是对于 n 阶矩阵 X , 有:

$$\begin{aligned} XA = AX &\Leftrightarrow (P^{-1}XP)(P^{-1}AP) = (P^{-1}AP)(P^{-1}XP) \\ &\Leftrightarrow (P^{-1}XP)D = D(P^{-1}XP) \\ &\Leftrightarrow P^{-1}XP = \text{diag}(b_1, b_2, \dots, b_n) \\ &\Leftrightarrow X = P \text{diag}(b_1, b_2, \dots, b_n)P^{-1} \\ &\Leftrightarrow X = b_1(PE_{11}P^{-1}) + b_2(PE_{22}P^{-1}) + \dots + b_n(PE_{nn}P^{-1}) \end{aligned}$$

由于 $E_{11}, E_{22}, \dots, E_{nn}$ 线性无关, 因此容易证明 $PE_{11}P^{-1}, PE_{22}P^{-1}, \dots, PE_{nn}P^{-1}$ 也线性无关, 从而

$$PE_{11}P^{-1}, PE_{22}P^{-1}, \dots, PE_{nn}P^{-1}$$

是 A 的列空间 $C(A)$ 的一个基, 于是 $\dim C(A) = n$ 。

- (3) 取 V_1 的一组基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 。因为 $V_1 \subseteq V_2$, 所以 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V_2 中的线性无关向量组。根据线性空间基的性质: 有限维线性空间中, 任意线性无关组所含向量个数不超过空间的维数, 因此 $n \leq m$, 即 $\dim V_1 \leq \dim V_2$ 。

以下证明若 $\dim V_1 = \dim V_2$, 则 $V_1 = V_2$ 。设 $\dim V_1 = \dim V_2 = k$, 取 V_1 的基 $\alpha_1, \alpha_2, \dots, \alpha_k$ 。由 $V_1 \subseteq V_2$, 知 $\alpha_1, \alpha_2, \dots, \alpha_k$ 是 V_2 中含 k 个向量的线性无关组。根据有限维线性空间基的判定: k 维线性空间中, 含 k 个向量的线性无关组是该空间的基, 因此 $\alpha_1, \alpha_2, \dots, \alpha_k$ 也是 V_2 的基。

两个线性空间有相同的基, 则它们的元素可由同一组基线性表示, 故 $V_1 = V_2$ 。

(4)

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

- (5) 可知从 $[\varepsilon_i]$ 到 $[\eta_i]$ 的过渡矩阵为:

$$[\eta_1, \eta_2, \eta_3] = [\varepsilon_1, \varepsilon_2, \varepsilon_3] \begin{bmatrix} 2 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

求过渡矩阵 S 的逆矩阵为:

$$S^{-1} = \begin{bmatrix} -6 & 5 & -2 \\ 4 & -3 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

从而:

$$B = S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

- (6) 设 A 与 B 相似, 则存在 n 阶可逆矩阵 S 使得 $B = S^{-1}AS$ 。取 $\mathbb{R}$ 上一个 n 维线性空间 V , V 中取一个基 $\alpha_1, \alpha_2, \dots, \alpha_n$, 定义 V 上的一个线性变换 T 满足:

$$T(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A$$

则令:

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)S$$

由于 S 可逆, 因此 $\beta_1, \beta_2, \dots, \beta_n$ 是 V 的一个基, 且 T 在基 $\beta_1, \beta_2, \dots, \beta_n$ 下的矩阵为 $S^{-1}AS = B$ 。

- (7) 当 $\mathbb{C}$ 看作 $\mathbb{R}$ 上的线性空间时, 对任意 $z_1, z_2 \in \mathbb{C}$, $k \in \mathbb{R}$, 有:

$$T(z_1 + z_2) = \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} = T(z_1) + T(z_2)$$

$$T(kz_1) = \overline{kz_1} = k\overline{z_1} = kT(z_1)$$

因此 T 是 $\mathbb{R}$ 上线性空间 $\mathbb{C}$ 上的一个线性变换。当 $\mathbb{C}$ 看作自身上的线性空间时, 对于 $z \in \mathbb{C}$, 且 $z \neq 0$, 有:

$$A(iz) = \overline{iz} = \overline{i}\overline{z} = -i\overline{z} \neq i\overline{z} = iA(z)$$

因此 A 不是线性空间 $\mathbb{C}$ 上的一个线性变换。

A.6 第六章答案

(1) 线性变换 T 的值域就是 A 的列空间, 经计算得矩阵 A 的列空间的基为:

$$\alpha_1 = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix} \quad \alpha_2 = \begin{bmatrix} -1 \\ 6 \\ 4 \end{bmatrix}$$

所以 $R(T)$ 中的每一个向量 $\alpha = (x, y, z)^T$ 均可由 $\alpha_1 = (1, 5, 7)^T, \alpha_2 = (-1, 6, 4)^T$ 线性表示。即有 $\alpha = (x, y, z)^T, \alpha_1 = (1, 5, 7)^T, \alpha_2 = (-1, 6, 4)^T$ 是共面的, 故 $\det[\alpha \quad \alpha_1 \quad \alpha_2] = 0$ 。得平面方程:

$$2x + y - z = 0$$

很明显这是过原点的。又线性变换 T 的核就是 $Ax = 0$ 的解空间, 经计算得方程组 $Ax = 0$ 的解空间的基为:

$$\xi_1 = (-14, 19, 11)^T$$

所以 $\ker(T)$ 中的每一个向量 $\xi = (x, y, z)^T$ 均与 $\xi_1 = (-14, 19, 11)^T$ 共线。即 $\xi = k\xi_1$, 得过原点的直线方程:

$$\frac{x}{-14} = \frac{y}{19} = \frac{z}{11}$$

(2) 设 $T_2 \in L(V_1, V_2)$ 对应的矩阵为 A , $T_1 \in L(V_2, V_3)$ 对应的矩阵为 B 。因为 $r(T_2) = r(A)$, $r(T_1) = r(B)$ 。而有限维线性变换的乘积对应于矩阵的乘积。所以有 $T_1 T_2$ 对应的矩阵为 AB , 且 $r(T_1 T_2) = r(AB)$ 。故有结论:

$$r(T_1 T_2) = r(AB) \leq \min\{r(A), r(B)\} = \min\{r(T_1), r(T_2)\}$$

(3) 要证明原命题, 要分别证明 V 是核与值域的和, 且交集为 $\{0\}$:

易有: $\forall \alpha \in V, \alpha = (\alpha - T(\alpha)) + T(\alpha)$, 且:

$$T(\alpha - T(\alpha)) = T(\alpha) - T^2(\alpha) = T(\alpha) - T(\alpha) = 0$$

那么有 $\alpha - T(\alpha) \in \ker(T)$ 且 $T(\alpha) \in R(T)$, 所以有 $V = \ker(T) + R(T)$ 。下证二者交集为 $\{0\}$:

$\forall \beta \in R(T)$, 一定有 $\alpha \in V$, 使得 $T(\alpha) = \beta$ 。又因为 $T(\beta) = T(T(\alpha)) = T^2(\alpha) = T(\alpha) = \beta$ 。若 $\beta \neq 0$, $T(\beta) \neq 0$, 所以 $\beta \notin \ker(T)$, 由此有:

$$\ker(T) \cap R(T) = \{0\}$$

故得: $V = \ker(T) \oplus R(T)$

(4) 根据线性相关性判断不难得到 $\alpha_1, \alpha_2, \alpha_3$ 是 $\mathbb{R}^3$ 的一个基 B 。线性算子 T 在基 B 下的矩阵为 A :

$$T[\alpha_1, \alpha_2, \alpha_3] = [T(\alpha_1), T(\alpha_2), T(\alpha_3)] = [\alpha_1, \alpha_2, \alpha_3]A$$

解得:

$$A = \begin{bmatrix} 2 & 3 & 5 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}$$

从基 B 到基 $B' = \{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ 的过渡矩阵为 C , 即 $[\varepsilon_1, \varepsilon_2, \varepsilon_3] = [\alpha_1, \alpha_2, \alpha_3]C$, 联立 $T[\alpha_1, \alpha_2, \alpha_3] = [T(\alpha_1), T(\alpha_2), T(\alpha_3)] = [\alpha_1, \alpha_2, \alpha_3]A$ 可求出 C 。设线性算子 T 基 B' 下的矩阵为 D , 解得:

$$D = C^{-1}AC = \frac{1}{7} \begin{bmatrix} -5 & 20 & -20 \\ -4 & -5 & -2 \\ 27 & 18 & 24 \end{bmatrix}$$

(5) 因 $\xi, T\xi, \dots, T^{n-1}\xi$ 是 V 中的 n 个向量, 故只需要证明它们是线性无关的即可。因 $T^n\xi = 0$, 所以 $T^{n+k}\xi = 0, (k \geq 1)$ 。令 $l_1\xi + l_2T\xi + \dots + l_nT^{n-1}\xi = 0, l_1, \dots, l_n \in F$, 以 T^{n-1} 作用于上式两边, 可得 $l_1T^{n-1}\xi = 0$, 而已知 $T^{n-1}\xi \neq 0$, 故 $l_1 = 0$ 。

同理可得, $l_i = 0, (i = 2, \dots, n)$, 故 $\xi, T\xi, \dots, T^{n-1}\xi$ 线性无关; 又因 $\dim(V) = n$, 故 $\xi, T\xi, \dots, T^{n-1}\xi$ 是 V 的一组基。

为求得 T 在上述基下的矩阵, 对上述的基继续施加变换 T , 有:

$$\begin{aligned} T[\xi \quad T\xi \quad \cdots \quad T^{n-2}\xi \quad T^{n-1}\xi] &= [T\xi \quad T^2\xi \quad \cdots \quad T^{n-1}\xi \quad T^n\xi] \\ &= [T\xi \quad T^2\xi \quad \cdots \quad T^{n-1}\xi \quad 0] \\ &= [\xi \quad T\xi \quad \cdots \quad T^{n-2}\xi \quad T^{n-1}\xi]A \end{aligned}$$

$$\text{故 } T \text{ 在 } \xi, T\xi, \dots, T^{n-1}\xi \text{ 下的矩阵 } A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

(6) 考虑到线性映射 $\sigma: F[x]_n \rightarrow F^{n+1}$, 具体有:

$$f \in F[x]_n \rightarrow \begin{bmatrix} f(a_0) \\ f(a_1) \\ \vdots \\ f(a_n) \end{bmatrix} \in F^{n+1}$$

容易验证 σ 是 F 上的 n 维线性空间 $F[x]_n$ 到 F^{n+1} 的线性映射. 对于每个 $0 \leq i \leq n$, $j \neq i$, 有 $f_i(a_j) = 0$ 成立, 而 $f_i(a_i) \neq 0$, 因此:

$$\sigma(f_i) = [0, \dots, 0, f_i(a_i), \dots, 0]^T = f_i(a_i)e_i \text{ (其中 } e_i \text{ 是第 } i \text{ 个数组自然基向量)}$$

所以, $n+1$ 个像组成的向量组:

$$\{\sigma(f_0), \sigma(f_1), \dots, \sigma(f_n)\} = \{f_0(a_0)e_1, f_1(a_1)e_2, \dots, f_n(a_n)e_{n+1}\}$$

线性无关, 这使得 $n+1$ 个像组成的向量组 $f_0, f_1, \dots, f_n$ 线性无关, 因而是 $F[x]_n$ 的一组基. 又已知从 $f_0, f_1, \dots, f_n$ 到 $1, x, x^2, \dots, x^n$ 的过渡矩阵满足条件:

$$[1, x, x^2, \dots, x^n] = [f_0, f_1, f_2, \dots, f_n]P$$

两边同时用线性变换 σ 作用, 得:

$$[\sigma(1), \sigma(x), \sigma(x^2), \dots, \sigma(x^n)] = [\sigma(f_0), \sigma(f_1), \sigma(f_2), \dots, \sigma(f_n)]P$$

即:

$$\begin{bmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^n \end{bmatrix} = \begin{bmatrix} f_0(a_0) & & & & \\ & f_1(a_1) & & & \\ & & \ddots & & \\ & & & & f_n(a_n) \end{bmatrix} P$$

所以有:

$$\begin{aligned} P &= \begin{bmatrix} f_1(a_1) & & & & \\ & f_2(a_2) & & & \\ & & \ddots & & \\ & & & & f_n(a_n) \end{bmatrix}^{-1} \begin{bmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_0^{-1} & \lambda_0^{-1}a_0 & \lambda_0^{-1}a_0^2 & \cdots & \lambda_0^{-1}a_0^n \\ \lambda_1^{-1} & \lambda_1^{-1}a_1 & \lambda_1^{-1}a_1^2 & \cdots & \lambda_1^{-1}a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_n^{-1} & \lambda_n^{-1}a_n & \lambda_n^{-1}a_n^2 & \cdots & \lambda_n^{-1}a_n^n \end{bmatrix} \end{aligned}$$

其中 $\lambda_i = f_i(a_i) \neq 0, \forall 0 \leq i \leq n$.

(7) (a). 只需要证明 $\ker(T+I) = \{\mathbf{0}\}$ 即可。假设 $\alpha \in \ker(T+I)$, 那么有:

$$(T+I)(\alpha) = T(\alpha) + \alpha$$

即得到 $\alpha = -T(\alpha)$, 两边再次应用算子 T 有:

$$T(\alpha) = -T(T(\alpha)) = -(T^2)(\alpha) = -T(\alpha)$$

则有 $T(\alpha) = \mathbf{0} = -\alpha$, 则一定有 $\alpha = \{\mathbf{0}\}$ 即证。

(b). 事实上, 根据算子的性质可知:

$$A^2 = A$$

易知矩阵 A 有且仅有特征值 0 或 1, 设 $r(A) = R$, 由第 7 章可知 A 相似于矩阵 $P = \begin{bmatrix} I_R & O \\ O & O \end{bmatrix}$, 故:

$$r(A) = \text{tr}(P) = R$$

(c). 易知 $2I - A$ 相似于矩阵 $P = \begin{bmatrix} I_R & O \\ O & 2I_{n-R} \end{bmatrix}$, 立刻有:

$$\det(2I - A) = 2^{n-R}$$

A.7 第七章答案

(1) 设 A 是数域 F 上的 n 阶幂等矩阵。如果 λ_0 是 A 的特征值, 那么有 $\alpha \in F^n$ 且 $\alpha \neq 0$, 使得 $A\alpha = \lambda_0\alpha$ 。两边左乘 A , 得 $A^2\alpha = A\lambda_0\alpha = \lambda_0^2\alpha$ 。由于 $A^2 = A$, 因此 $A\alpha = \lambda_0^2\alpha$ 。于是 $\lambda_0\alpha = \lambda_0^2\alpha$ 。从而 $(\lambda_0 - \lambda_0^2)\alpha = 0$ 。由于 $\alpha \neq 0$, 因此 $\lambda_0 - \lambda_0^2 = 0$ 。因此推出 $\lambda_0 = 0$ 或 $\lambda_0 = 1$ 。

(2) 设 μ 为 G 的特征值, 特征方程为:

$$|\mu I_{2n} - G| = 0$$

即有:

$$\begin{vmatrix} \mu I_n - A & -A^m \\ -A^m & \mu I_n - A \end{vmatrix} = 0$$

根据分块行列式公式:

$$\begin{vmatrix} P & Q \\ Q & P \end{vmatrix} = |P+Q| \cdot |P-Q|$$

代入 $P = \mu I_n - A, Q = -A^m$ 得:

$$|(\mu I_n - A) - A^m| \cdot |(\mu I_n - A) + A^m| = 0$$

化简为:

$$|\mu I_n - (A + A^m)| \cdot |\mu I_n - (A - A^m)| = 0$$

由矩阵特征值的运算性质, 若 A 的特征值为 λ_i ($i = 1, 2, \dots, n$), 则 $A + A^m$ 的特征值为 $\lambda_i + \lambda_i^m$; $A - A^m$ 的特征值为 $\lambda_i - \lambda_i^m$ 。因此 $|\mu I_n - (A + A^m)| = 0$ 的根为 $\lambda_i + \lambda_i^m$, $|\mu I_n - (A - A^m)| = 0$ 的根为 $\lambda_i - \lambda_i^m$ ($i = 1, 2, \dots, n$)

(3) C 的特征多项式 $|\lambda I - C|$ 为:

$$\begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ -1 & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix} = \lambda^n - 1$$

于是矩阵 C 的全部特征值是 $1, \xi, \dots, \xi^{n-1}$, 其中 $\xi = e^{\frac{2\pi i}{n}}$ 。对于非负整数 $m (0 \leq m < n)$, 有:

$$C \begin{bmatrix} 1 \\ \xi^m \\ \xi^{2m} \\ \vdots \\ \xi^{(n-1)m} \end{bmatrix} = \begin{bmatrix} \xi^m \\ \xi^{2m} \\ \vdots \\ \xi^{(n-1)m} \\ 1 \end{bmatrix} = \xi^m \begin{bmatrix} 1 \\ \xi^m \\ \xi^{2m} \\ \vdots \\ \xi^{(n-1)m} \end{bmatrix}$$

因此 C 的属于特征值 ξ^m 的所有特征向量组成的集合是:

$$\{k(1, \xi^m, \xi^{2m}, \dots, \xi^{(n-1)m})^T \mid k \in \mathbb{C} \text{ 且 } k \neq 0\}$$

- (4) 设 α, β 分别是 A 的属于 λ_1, λ_2 的特征向量, 且 $\lambda_1 \neq \lambda_2$ 。如果 $\alpha + \beta$ 是 A 的特征向量, 那么它必属于 A 的某个特征值 λ_3 。于是 $A(\alpha + \beta) = \lambda_3(\alpha + \beta)$ 。又有:

$$A(\alpha + \beta) = A\alpha + A\beta = \lambda_1\alpha + \lambda_2\beta$$

从而:

$$\lambda_1\alpha + \lambda_2\beta = \lambda_3\alpha + \lambda_3\beta$$

即:

$$(\lambda_1 - \lambda_3)\alpha + (\lambda_2 - \lambda_3)\beta = 0$$

由于 A 的属于不同特征值的特征向量线性无关, 因此 α, β 线性无关。从而推出, $\lambda_1 = \lambda_3 = \lambda_2$ 矛盾。因此 $\alpha + \beta$ 不是 A 的特征向量。

(5)

$$T_1 = \begin{bmatrix} \frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ -\frac{1}{5}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ 0 & \frac{1}{3}\sqrt{5} & -\frac{2}{3} \end{bmatrix} \quad T_1^{-1}A_1T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -8 \end{bmatrix}$$

$$T_2 = \begin{bmatrix} \frac{1}{5}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ -\frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ 0 & -\frac{1}{3}\sqrt{5} & \frac{2}{3} \end{bmatrix} \quad T_2^{-1}A_2T_2 = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

$$T_3 = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \quad T_3^{-1}A_3T_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$T_4 = \begin{bmatrix} \frac{1}{2}\sqrt{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2}\sqrt{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\sqrt{2} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \quad T_4^{-1}A_4T_4 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

- (6) 设 n 阶正交矩阵 A 有两个不同的特征值 λ_1, λ_2 , 则 $|\lambda_i| = 1, i = 1, 2$. 由于 λ_i 是实数, 因此 $\lambda_i = \pm 1$. 不妨设 $\lambda_1 = 1, \lambda_2 = -1$. 设 α_i 是 A 的属于 λ_i 的一个特征向量, $i = 1, 2$, 则:

$$\langle \alpha_1, \alpha_2 \rangle = \lambda_1 \langle \alpha_1, \alpha_2 \rangle = \langle \lambda_1 \alpha_1, \alpha_2 \rangle = \langle A \alpha_1, \alpha_2 \rangle = (A \alpha_1)^T \alpha_2 = \alpha_1^T A^T \alpha_2$$

利用正交矩阵的性质有:

$$\alpha_1^T A^T \alpha_2 = \alpha_1^T A^{-1} \alpha_2 = \alpha_1^T \lambda_2^{-1} \alpha_2 = \alpha_1^T (-1) \alpha_2 = (-1) \langle \alpha_1, \alpha_2 \rangle$$

从而有 $\langle \alpha_1, \alpha_2 \rangle = 0$ 即证。

A.8 第八章答案

- (1) 设可逆线性变换:

$$\begin{cases} x_1 = y_1 \\ x_2 = y_2 + y_3 \\ x_3 = y_2 - y_3 \end{cases}$$

则有:

$$f_1(x_1, x_2, x_3) = y_1^2 - y_2^2 + y_3^2$$

再令:

$$\begin{cases} y_1 = w_1 + w_2 \\ y_2 = w_1 - w_2 \\ y_3 = w_3 \end{cases}$$

则有:

$$f_2(y_1, y_2, y_3) = w_1^2 - w_2^2 - w_3^2.$$

再令:

$$\begin{cases} z_1 = u_1 + u_2 \\ z_2 = u_1 - u_2 \\ z_3 = u_3 \end{cases}$$

则有:

$$f_3(z_1, z_2, z_3) = u_1^2 - u_2^2 + u_3^2$$

由此看出, $f_1(x_1, x_2, x_3)$ 与 $f_3(z_1, z_2, z_3)$ 的秩都是等于 3, 且正惯性指数都等于 2, 因此它们等价. 由于 $f_2(y_1, y_2, y_3)$ 的正惯性指数为 1, 因此 $f_2(y_1, y_2, y_3)$ 与 $f_1(x_1, x_2, x_3)$ 不等价, 与 $f_3(z_1, z_2, z_3)$ 也不等价.

(2) 设 $P(x, y, z)$ 是锥面上的点, 过 P 点的母线与准线交于 $P_0(x_0, y_0, z_0)$ 。则有:

$$\frac{x_0}{x} = \frac{y_0}{y} = \frac{c}{z} x_0 = \frac{cx}{z}, y_0 = \frac{cy}{z}$$

由 x_0, y_0, c 在准线上, 代入:

$$\frac{1}{a^2} \left(\frac{cx}{z} \right)^2 + \frac{1}{b^2} \left(\frac{cy}{z} \right)^2 = 1 \frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

当 $a = b$ 时为圆锥: $z^2 = \left(\frac{c}{a} \right)^2 (x^2 + y^2)$ 。

(3) 交线方程为 $\begin{cases} y^2 + z^2 = x \\ x + 2y - z = 0 \end{cases}$, 分别单独消去 x, y, z 有投影:

$$\begin{cases} y^2 + z^2 + 2y - z = 0 \\ x = 0 \end{cases} \quad (\text{消去 } x)$$

$$\begin{cases} x^2 + 5z^2 - 2xz - 4x = 0 \\ y = 0 \end{cases} \quad (\text{消去 } y)$$

$$\begin{cases} x^2 + 5y^2 + 4xy - x = 0 \\ z = 0 \end{cases} \quad (\text{消去 } z)$$

(4) 该二次型对应的矩阵为:

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & 0 & 0 \\ \frac{1}{2} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \frac{1}{2} \\ 0 & 0 & \cdots & \frac{1}{2} & 1 \end{bmatrix}$$

则各阶顺序主子式为:

$$\Delta_k = \left(\frac{1}{2} \right)^k \begin{vmatrix} 2 & 1 & \cdots & 0 & 0 \\ 1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & \cdots & 1 & 2 \end{vmatrix}$$

设:

$$D_k = \begin{vmatrix} 2 & 1 & \cdots & 0 & 0 \\ 1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & \cdots & 1 & 2 \end{vmatrix}$$

将 D_k 从第一行展开有 $D_k = 2D_{k-1} - D_{k-2}$, 也就有 $D_k - D_{k-1} = D_{k-1} - D_{k-2}$, $\{D_k\}$ 为等差数列。而 $D_1 = 2, D_2 = 3$, 可得 $D_k = k + 1$, 因此各阶顺序主子式为 $\Delta_k = \left(\frac{1}{2} \right)^k D_k = \left(\frac{1}{2} \right)^k (k + 1) > 0$ 正定。

(5) 设 A 正定, 则存在正交矩阵 Q , 使得 $Q^T A Q = \text{diag}(\lambda_1, \dots, \lambda_n)$, 且 $\lambda_1, \dots, \lambda_n$ 都大于零:

$$A = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} Q^T = Q \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} Q^T Q \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} Q^T = S^2$$

其中 S 与 $\begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix}$ 合同, 所以也是正定阵。

(6) 设 $C = \text{diag}(c_1, c_2, \dots, c_n)$, 则有 $B = C^{-1} A C$ 正定。

(7) f 对应的矩阵 A 的特征值为 $1, 1, 0$, 且 0 对应的特征向量为 e_3 。而 1 对应的特征向量与 e_3 正交。求解 $e_3^T x = 0$ 得 $\xi_1 = (1, 0, -1)^T$, $\xi_2 = (0, 1, 0)^T$, 单位化得 $e_1 = \frac{1}{\sqrt{2}}\xi_1$, $e_2 = \xi_2$ 。

令 $Q = (e_1, e_2, e_3)$, 则 $x = Qy$ 是正交变换。且

$$Q^T A Q = D = \text{diag}(1, 1, 0) \Rightarrow A = Q D Q^T = e_1 e_1^T + e_2 e_2^T = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

所以 $A + I = Q(D + I)Q^T$, 即 $A + I$ 与 $D + I = \text{diag}(2, 2, 1)$ 合同, 所以 $A + I$ 正定。