



线性代数习题课讲义

作者：陈文轩

组织：KFRC

时间：August 1, 2026

版本：2.0.1.2880



目录

第一章 矩阵之前	1
1.1 向量及其计算	1
1.1.1 加法与数乘的定义	1
1.1.2 向量的内积与夹角	1
1.1.3 $\mathbb{R}^3$ 向量的外积与混合积	2
1.2 数与数域	3
1.2.1 复数与单位根	3
1.2.2 数环、数域	5
第二章 求和符号, 组合恒等式与线性方程组	7
2.1 补充: 线性方程组	7
2.1.1 线性方程组的求解	7
2.1.2 唯一解, 无解, 无穷解问题	8
2.1.3 插值问题	8
2.2 求和符号练习	9
2.3 组合恒等式	10
第三章 行列式	12
3.1 行列式的定义 *	12
3.2 行列式计算练习	13
3.2.1 定义法/降阶法/化为上三角	13
3.2.2 加边	15
3.2.3 逐行/列加减	19
3.2.4 拆项与递推	23
3.2.5 乘法分解	25
3.3 函数与行列式 *	28
3.3.1 多项式与行列式	28
3.3.2 行列式的导数与积分	29
3.3.3 行列式与 Taylor 公式	30
3.4 综合练习 *	31
第四章 秩与相抵	34
4.1 秩与相抵的基本概念	34
4.2 秩不等式	35
4.3 初等变换与秩	36
4.4 补充: 利用矩阵结构的求逆技巧	38
第五章 线性空间的基与线性映射	39
5.0 选择公理 *	39
5.1 线性空间的基 *	40

5.1.1 Hamel 基	40
5.1.2 Schauder 基	40
5.2 相关练习	41
5.3 线性映射	44
5.3.1 线性映射与线性变换的定义	44
5.3.2 线性映射的矩阵描述	44
5.3.3 线性映射的性质 *	45
5.3.4 从线性变换观点看线性方程组	47
第六章 期中复习	49
6.1 复数与向量	49
6.2 线性方程组的 Gauss 消元法	50
6.3 行列式	50
6.4 矩阵	53
6.5 秩与相抵	56
6.6 线性空间	58
6.7 线性方程组解集的结构	61
第七章 线性变换与矩阵的相似	63
7.1 矩阵的相似	63
7.2 一些其他结论 *	65
7.2.1 Caylay-Hamilton 定理	65
7.2.2 Gershgorin 圆盘定理	66
7.2.3 同时对角化问题	66
第八章 Jordan 标准型 *	68
8.1 多项式矩阵的相抵	68
8.2 Jordan 标准型的概念与求法	73
8.3 Jordan 标准型的应用	76
第九章 矩阵分解	79
9.1 复习与总览	79
9.1.1 Gauss 消元与 LU 分解	79
9.1.2 与秩相关的分解	79
9.1.3 与特征值相关的分解	80
9.1.4 一般矩阵的分解	80
9.1.5 对称/ Hermitian 矩阵分解	81
9.2 奇异值分解	82
9.3 极分解	83
9.4 其他内容	84
第十章 期末复习	85
10.1 线性变换与矩阵的相似	85
10.2 正交变换与内积空间	86

10.3 二次型与对称矩阵的相合	89
----------------------------	----

第一章 矩阵之前

可能使用到的行列式计算： $\det \begin{bmatrix} a & b \\ x & y \end{bmatrix} = ay - bx$,

$$\det \begin{bmatrix} a & b & c \\ x & y & z \\ p & q & r \end{bmatrix} = a \det \begin{bmatrix} y & z \\ q & r \end{bmatrix} - b \det \begin{bmatrix} x & z \\ p & r \end{bmatrix} + c \det \begin{bmatrix} x & y \\ p & q \end{bmatrix} = a(yr - zq) - b(xr - pz) + c(xq - py)。$$

1.1 向量及其计算

既有大小又有方向的量称为向量，通常使用粗体字母表示向量，例如 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 等。向量可以用有序数组表示，例如 $\mathbf{a} = (a_1, a_2, \dots, a_n)$ ，其中 a_i 称为向量的第 i 个分量。

1.1.1 加法与数乘的定义

如果两个向量 $\mathbf{a}, \mathbf{b}$ 大小相等且方向相同，则称它们相等，记为 $\mathbf{a} = \mathbf{b}$ ，如果大小相等而方向相反，则称它们互为反向量，记为 $\mathbf{a} = -\mathbf{b}$ 。向量的长度称为模，记为 $|\mathbf{a}|$ 。模为 1 的向量称为单位向量，模为 0 的向量称为零向量，零向量没有确定的方向。

向量的加法通过以下方法定义对于 $\mathbf{a}, \mathbf{b}$ ，将 $\mathbf{a}$ 的终点与 $\mathbf{b}$ 的起点重合，连接 $\mathbf{a}$ 的起点与 $\mathbf{b}$ 的终点得到的向量称为 $\mathbf{a} + \mathbf{b}$ 。向量的减法定义为 $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$ 。向量的数乘定义为对于实数 λ ， $\lambda \mathbf{a}$ 是一个与 $\mathbf{a}$ 方向相同且模为 $|\lambda||\mathbf{a}|$ 的向量，当 $\lambda < 0$ 时， $\lambda \mathbf{a}$ 与 $\mathbf{a}$ 方向相反。显然 $\forall \mathbf{a} \neq \mathbf{0}, \frac{\mathbf{a}}{|\mathbf{a}|}$ 是与 $\mathbf{a}$ 同向的单位向量。

显然向量的加法与数乘满足以下运算律：

1. 交换律： $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$;
2. 结合律： $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$;
3. 零元： $\mathbf{a} + \mathbf{0} = \mathbf{a}$;
4. 负元： $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$;
5. 分配律： $\lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$;
6. 数乘结合律： $(\lambda \mu) \mathbf{a} = \lambda(\mu \mathbf{a})$;
7. 数乘分配律： $(\lambda + \mu) \mathbf{a} = \lambda \mathbf{a} + \mu \mathbf{a}$;
8. 单位元： $1 \mathbf{a} = \mathbf{a}$ 。

当使用有序数组 $\mathbf{a} = (a_1, a_2, \dots, a_n)$ 描述向量时，向量的加法与数乘可以通过分量来定义： $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$ ， $\lambda \mathbf{a} = (\lambda a_1, \lambda a_2, \dots, \lambda a_n)$ 。

1.1.2 向量的内积与夹角

定义 1.1

对于使用有序数组描述的向量 $\mathbf{a} = (a_1, a_2, \dots, a_n), \mathbf{b} = (b_1, b_2, \dots, b_n)$ ，定义它们的内积为

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i$$

对于任意向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 与任意实数 λ, μ ，内积满足以下性质：

命题 1.2

1. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$;
2. $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$;
3. $(\lambda \mathbf{a}) \cdot (\mu \mathbf{b}) = (\lambda \mu)(\mathbf{a} \cdot \mathbf{b})$;
4. $\mathbf{a} \cdot \mathbf{a} \geq 0$, 且当且仅当 $\mathbf{a} = \mathbf{0}$ 时等号成立。

定义 1.3

对于任意的向量 $\mathbf{a}$, 其模长可以通过内积定义为

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

。

对于任意非零向量 $\mathbf{a}, \mathbf{b}$, 定义它们的夹角为

$$\theta = \arccos \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$

其中 $0 \leq \theta \leq \pi$ 。

若 $\theta = 0$ 或 π , 则称 $\mathbf{a}, \mathbf{b}$ 共线 (平行); 若 $\theta = \frac{\pi}{2}$, 则称 $\mathbf{a}, \mathbf{b}$ 正交 (垂直)。

例题 1.1 计算 $\mathbf{a} = (1, 2, 3), \mathbf{b} = (4, 5, 6)$ 的内积与夹角。

解 $\mathbf{a} \cdot \mathbf{b} = 1 \times 4 + 2 \times 5 + 3 \times 6 = 32, |\mathbf{a}| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}, |\mathbf{b}| = \sqrt{4^2 + 5^2 + 6^2} = \sqrt{77}, \theta = \arccos \frac{32}{7\sqrt{22}}$ 。

例题 1.2 证明: $|a + b| \leq |a| + |b|$

证明 由内积的定义, 有 $|a + b|^2 = (a + b) \cdot (a + b) = a \cdot a + 2a \cdot b + b \cdot b = |a|^2 + 2|a||b| \cos \theta + |b|^2 \leq |a|^2 + 2|a||b| + |b|^2 = (|a| + |b|)^2$, 取算术平方根即得。

例题 1.3 证明 Cauchy 不等式: $\left(\sum_{i=1}^n a_i b_i\right)^2 \leq \left(\sum_{i=1}^n a_i^2\right) \left(\sum_{i=1}^n b_i^2\right)$

证明 记 $\mathbf{a} = (a_1, \dots, a_n), \mathbf{b} = (b_1, \dots, b_n)$, 则 $\left(\sum_{i=1}^n a_i b_i\right)^2 = (\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \cos^2 \theta \leq |\mathbf{a}|^2 |\mathbf{b}|^2 = \left(\sum_{i=1}^n a_i^2\right) \left(\sum_{i=1}^n b_i^2\right)$ 。

1.1.3 $\mathbb{R}^3$ 向量的外积与混合积

定义 1.4

对于 $\mathbf{a} = (a_1, a_2, a_3), \mathbf{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$, 记 $\mathbf{i} = (1, 0, 0), \mathbf{j} = (0, 1, 0), \mathbf{k} = (0, 0, 1)$, 则定义 $\mathbf{a}, \mathbf{b}$ 的外积为

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$$

显然 $\mathbf{a} \times \mathbf{b}$ 与 $\mathbf{a}, \mathbf{b}$ 都垂直, 且 $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \theta$, 其中 θ 为 $\mathbf{a}, \mathbf{b}$ 的夹角。

命题 1.5

对于任意 $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ 与任意实数 λ, μ , 外积满足以下性质:

1. $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$;
2. $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$;
3. $(\lambda \mathbf{a}) \times (\mu \mathbf{b}) = (\lambda \mu)(\mathbf{a} \times \mathbf{b})$;
4. $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}$;
5. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ 。

对于 $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$, 如下定义其混合积:

定义 1.6

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

混合积的几何意义为 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 所张立的平行六面体的有向体积。

混合积有以下性质:

命题 1.7

对于任意 $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ 与任意实数 λ, μ , 混合积满足以下性质:

1. $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b});$
2. $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -(\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}));$

例题 1.4 证明 Lagrange's Identity: $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$ 。

证明 $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = \mathbf{a} \cdot (\mathbf{b} \times (\mathbf{c} \times \mathbf{d})) = \mathbf{a} \cdot ((\mathbf{b} \cdot \mathbf{d})\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$

1.2 数与数域

1.2.1 复数与单位根

复数是形如 $a + bi$ 的数, 其中 $a, b \in \mathbb{R}, i^2 = -1$ 。复数的实部为 a , 虚部为 b 。复数的加法通过以下方式定义: $(a + bi) + (c + di) = (a + c) + (b + d)i$, 乘法与实数的乘法类似: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$, 除法定义为 $\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i$, 其中 $c + di \neq 0$ 。

在平面直角坐标系中使用向量 (a, b) 表示复数 $a + bi$, 则复数的加法与向量的加法相同。向量 (a, b) 的模长为 $\sqrt{a^2 + b^2}$, 称为复数的模。将 x 轴正半轴逆时针旋转至向 (a, b) 的方向所经过的角度称为复数的辐角, 记为 $\text{Arg}(a + bi)$ 。一个复数的辐角是不唯一的, 因为可以加上任意整数倍的 2π 。通常使用辐角在 $[0, 2\pi)$ 中的值作为复数的辐角, 称为辐角主值, 记作 $\arg z$ 。复数 $a + bi$ 的共轭复数定义为 $\overline{a + bi} = a - bi$ 。

复数的模长与辐角也可以唯一确定一个复数, 称为复数的三角形式: $r(\cos \theta + i \sin \theta)$, 其中 $r = |a + bi|, \theta = \text{Arg}(a + bi)$ 。复数三角形式的乘法与除法有简单的形式: $r_1(\cos \theta_1 + i \sin \theta_1)r_2(\cos \theta_2 + i \sin \theta_2) = r_1r_2[\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]$, $\frac{r_1(\cos \theta_1 + i \sin \theta_1)}{r_2(\cos \theta_2 + i \sin \theta_2)} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)]$ 。

定理 1.8 (Euler)

对于任意实数 θ , 有 $e^{i\theta} = \cos \theta + i \sin \theta$ 。特别地, 当 $\theta = \pi$ 时, 有 $e^{i\pi} + 1 = 0$ 。

利用 Euler 公式, 可以将复数的三角形式写为指数形式: $r(\cos \theta + i \sin \theta) = re^{i\theta}$ 。此时复数的乘法与除法可以写为: $r_1e^{i\theta_1}r_2e^{i\theta_2} = r_1r_2e^{i(\theta_1+\theta_2)}$, $\frac{r_1e^{i\theta_1}}{r_2e^{i\theta_2}} = \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)}$ 。

例题 1.5 计算 $\sum_{j=0}^n \cos j\theta$ 和 $\sum_{j=0}^n \sin j\theta$ 。

解 对 $\theta = 2k\pi$, 有 $\cos j\theta = 1$, $\sin j\theta = 0$, 故 $\sum_{j=0}^n \cos j\theta = n + 1$, $\sum_{j=1}^n \sin j\theta = 0$ 。

对 $\theta \neq 2k\pi$, 令 $T = \sum_{j=0}^n e^{ij\theta} = \sum_{j=0}^n \cos j\theta + i \sum_{j=0}^n \sin j\theta$, 有 $e^{i\theta}T = T + e^{i(n+1)\theta} - 1$, 故 $T = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}}$ 。

分别计算实部虚部, 有 $\sum_{j=0}^n \cos j\theta = \frac{1 - \cos \theta + \cos n\theta - \cos(n+1)\theta}{2(1 - \cos \theta)}$, $\sum_{j=0}^n \sin j\theta = \frac{-\sin(n+1)\theta + \sin n\theta + \sin \theta}{2(1 - \cos \theta)}$.

利用和差化积公式, 可以化简为 $\sum_{j=0}^n \cos j\theta = \frac{\sin \frac{(n+1)\theta}{2} \cos \frac{n\theta}{2}}{\sin \frac{\theta}{2}}$, $\sum_{j=0}^n \sin j\theta = \frac{\sin \frac{(n+1)\theta}{2} \sin \frac{n\theta}{2}}{\sin \frac{\theta}{2}}$.

定义 1.9

n 次单位根是指

$$\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, k = 0, 1, \dots, n-1$$

这 n 个复数, 记 $\omega_n = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$, 则他们可以写为 $\omega_n^0, \omega_n^1, \dots, \omega_n^{n-1}$ 。

之所以叫单位根, 是因为其满足

命题 1.10 (单位根的性质)

n 次单位根是方程 $x^n = 1$ 的全部根。

根据因式定理, 可以推出

$$x^n - 1 = \prod_{k=0}^{n-1} (x - \omega_n^k)$$

对比各项系数可以得到一些恒等式。事实上最常用的为:

命题 1.11 (单位根恒等式)

1.

$$\sum_{k=0}^{n-1} (\omega_n^k)^m = \begin{cases} n & n \mid m \\ 0 & n \nmid m \end{cases}$$

2.

$$\omega_n^a = \omega_n^{a \pm n}$$

3.

$$\overline{\omega_n^a} = \omega_n^{-a}$$

4.

$$\omega_{mn}^{ma} = \omega_n^a$$

证明 只证明 (1).

对 $n \mid m$, 有 $\omega_n^m = 1$, 故 $\sum_{k=0}^{n-1} (\omega_n^k)^m = \sum_{k=0}^{n-1} 1 = n$.

对 $n \nmid m$, 有 $\omega_n^m \neq 1$, 故 $\sum_{k=0}^{n-1} (\omega_n^k)^m = \sum_{k=0}^{n-1} \omega_n^{km} = \frac{1 - \omega_n^{nm}}{1 - \omega_n^m} = 0$.

第二个恒等式就是最重要的循环性质。出于循环, 它可以将一些东西按照模 n 的余数分类。

例题 1.6 计算 $\sum_{k=0}^{\lfloor \frac{n}{3} \rfloor} \binom{n}{3k}$, $\binom{n}{m}$ 为组合数 C_n^m , $\lfloor x \rfloor$ 为向下取整, 即不大于 x 的最大整数。

解 考虑三次单位根 $\omega_3 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, 由二项式定理有

$$(1+1)^n = \sum_{k=0}^n \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \binom{n}{4} + \cdots \quad (1.1)$$

$$(1+\omega_3)^n = \sum_{k=0}^n \binom{n}{k} \omega_3^k = \binom{n}{0} + \binom{n}{1} \omega_3 + \binom{n}{2} \omega_3^2 + \binom{n}{3} + \binom{n}{4} \omega_3 + \cdots \quad (1.2)$$

$$(1+\omega_3^2)^n = \sum_{k=0}^n \binom{n}{k} \omega_3^{2k} = \binom{n}{0} + \binom{n}{1} \omega_3^2 + \binom{n}{2} \omega_3 + \binom{n}{3} + \binom{n}{4} \omega_3^2 + \cdots \quad (1.3)$$

将 (1.1), (1.2) 和 (1.3) 相加, 注意到 $1 + \omega_3 + \omega_3^2 = 0$, 有

$$(1+1)^n + (1+\omega_3)^n + (1+\omega_3^2)^n = 3 \sum_{k=0}^{\lfloor \frac{n}{3} \rfloor} \binom{n}{3k}$$

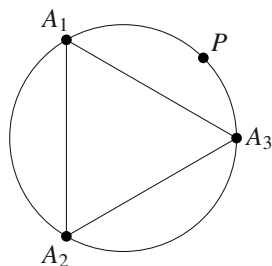
即有

$$\sum_{k=0}^{\lfloor \frac{n}{3} \rfloor} \binom{n}{3k} = \frac{1}{3} (2^n + (1+\omega_3)^n + (1+\omega_3^2)^n)$$

练习 1.1 计算 $\sum_{k=0}^{\lfloor \frac{n-1}{3} \rfloor} \binom{n}{3k+1}$.

例题 1.7 单位圆内接正 n 边形 $A_1 A_2 A_3 \dots A_n$ 中, P 为单位圆上一点, 证明: $\sum_{k=1}^n |PA_k|^2 = 2n$.

证明



放在复平面上, 设 $A_1 = \omega_n$, 则 $A_k = \omega_n^k, \forall k = 1, 2, \dots, n$. 设 $P = z$.

$$\Rightarrow \sum_{k=1}^n |PA_k|^2 = \sum_{k=1}^n |z - \omega_n^k|^2 = \sum_{k=1}^n (z - \omega_n^k)(\bar{z} - \overline{\omega_n^k}) = \sum_{k=1}^n (1 - \omega_n^k \bar{z} - z \overline{\omega_n^k} + 1) = \sum_{k=1}^n 2 - \bar{z} \sum_{k=1}^n \omega_n^k - z \sum_{k=1}^n \overline{\omega_n^k} = 2n$$

1.2.2 数环、数域

定义 1.12 (数环、数域)

若复数域 $\mathbb{C}$ 的某个包含 1 的子集 K 满足以下条件:

1. $a, b \in K \Rightarrow a + b \in K$;
2. $a, b \in K \Rightarrow a - b \in K$;
3. $a, b \in K \Rightarrow ab \in K$;

则称 K 为数环。

若还满足 $a \in K, a \neq 0 \Rightarrow a^{-1} \in K$, 则称 K 为数域。

不难验证:

命题 1.13 (数环、数域-例子)

有理数集 $\mathbb{Q}$ 、实数集 $\mathbb{R}$ 、复数集 $\mathbb{C}$ 都是数域。整数集 $\mathbb{Z}$ 是数环但不是数域。

 **笔记** 当然，还有形态更复杂的数环和数域，例如所有 $a + bi$ 在 $a, b \in \mathbb{Z}$ 时构成数环， $a, b \in \mathbb{Q}$ 构成数域。
对于究竟怎样的集合是数环/数域，有一个简单的结论：

命题 1.14 (最小的数环、数域)

任何数环包含 $\mathbb{Z}$ ，任何数域包含 $\mathbb{Q}$ 。

第二章 求和符号，组合恒等式与线性方程组

2.1 补充：线性方程组

2.1.1 线性方程组的求解

在求解线性方程组的时候，需要牢记根本方法只有一个：初等行变换，即使用 Gauss 消元法把矩阵化为阶梯型。但是，在此之后我们还可以进行进一步的操作，把矩阵化为更简单的形式（以齐次方程组为例，只考虑系数矩阵 $C = (c_{ij})_{m \times n}$ ）：先对每行乘某个数使得首个非零元素为 1，然后通过初等行变换使得这个 1 所在列的其他元素为 0：

$$\begin{aligned} & \begin{bmatrix} c_{11} & \cdots & c_{1j_2} & \cdots & c_{1j_r} & \cdots & c_{1n} \\ 0 & \cdots & c_{2j_2} & \cdots & c_{2j_r} & \cdots & c_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & c_{rj_r} & \cdots & c_{rn} \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \xrightarrow{\frac{1}{c_{kj_k}} r_k} \begin{bmatrix} 1 & \cdots & \tilde{c}_{1j_2} & \cdots & \tilde{c}_{1j_r} & \cdots & \tilde{c}_{1n} \\ 0 & \cdots & 1 & \cdots & \tilde{c}_{2j_r} & \cdots & \tilde{c}_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & \tilde{c}_{rn} \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \\ & \xrightarrow[-i=k-1, \dots, 1; k=r, \dots, 1]{-\tilde{c}_{ij_k} \rightarrow r_i} \begin{bmatrix} 1 & \tilde{c}_{12} & \cdots & 0 & \tilde{c}_{1,j_2+1} & \cdots & 0 & \tilde{c}_{1,j_r+1} & \cdots & \tilde{c}_{1n} \\ 0 & 0 & \cdots & 1 & \tilde{c}_{2,j_2+1} & \cdots & 0 & \tilde{c}_{2,j_r+1} & \cdots & \tilde{c}_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & \tilde{c}_{r,j_r+1} & \cdots & \tilde{c}_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \end{aligned}$$

可以发现这个矩阵具有以下好处：

- 省去了最后抄回原方程组代入的步骤；
- 这样操作得到的矩阵会是唯一的，避免了操作方式不同导致得到不同的答案；
- 解可以简便地写出来，甚至可以说已经解出来了：除了只有一个 1 的列（注意这个 1 要是它所在的那一行的第一个非零元），其余列可直接按序命为自由变量，直接写出方程的解，并且可以很清楚地看出自由变量个数为 $n - r$ 。

除此之外，正如解方程时可以把 x_i 和 x_j 在方程组中出现位置交换，用行变换法解方程时也是可以交换列的（也可以认为是从其他方向造阶梯），面对具体问题时不用拘泥于一般阶梯的固定形式，但是要记清楚交换后变量与交换前变量的对应关系。以交换 x_1, x_2 为例，这可以使用矩阵乘法解释：

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \iff \begin{bmatrix} a_{12} & a_{11} & \cdots & a_{1n} \\ a_{22} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

最后留给大家一个问题：一个非齐次线性方程组的解和它的系数矩阵对应的齐次线性方程组的解有什么关系？

2.1.2 唯一解, 无解, 无穷解问题

例题 2.1 已知线性方程组
$$\begin{cases} 3x_1 + 2x_2 - x_3 = 6 \\ x_1 + ax_2 + 2x_3 = 9 \\ 2x_1 - x_2 + 3x_3 = 3 \end{cases}$$
 有唯一解, 则 a 应满足什么条件?

解

$$\begin{bmatrix} 3 & 2 & -1 & 6 \\ 1 & a & 2 & 9 \\ 2 & -1 & 3 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 2 & -1 & 6 \\ 7 & a+4 & 0 & 21 \\ 11 & 5 & 0 & 21 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 2 & -1 & 6 \\ 0 & a+\frac{9}{11} & 0 & * \\ 11 & 5 & 0 & 21 \end{bmatrix}$$

因此有唯一解的条件是 $a \neq -\frac{9}{11}$ 。

例题 2.2 已知线性方程组
$$\begin{cases} \lambda x_1 + x_2 + x_3 = 1 \\ x_1 + \lambda x_2 + x_3 = \lambda \\ x_1 + x_2 + \lambda x_3 = \lambda^2 \end{cases}$$
 有无穷多解, 则 λ 应满足什么条件?

解

$$\begin{bmatrix} \lambda & 1 & 1 & 1 \\ 1 & \lambda & 1 & \lambda \\ 1 & 1 & \lambda & \lambda^2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1-\lambda & 1-\lambda^2 & 1-\lambda^3 \\ 0 & \lambda-1 & 1-\lambda & \lambda-\lambda^2 \\ 1 & 1 & \lambda & \lambda^2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 2-\lambda-\lambda^2 & 1+\lambda-\lambda^2-\lambda^3 \\ 0 & \lambda-1 & 1-\lambda & \lambda-\lambda^2 \\ 1 & 1 & \lambda & \lambda^2 \end{bmatrix}$$

因此有无穷多解的条件是
$$\begin{cases} 2-\lambda-\lambda^2=0 \\ 1+\lambda-\lambda^2-\lambda^3=0 \end{cases}, \text{ 即 } \lambda=1.$$

2.1.3 插值问题

例题 2.3(Hermite 插值) 求三次多项式 $f(x) = ax^3 + bx^2 + cx + d$ 满足 $f(0) = 2, f'(0) = 2, f(1) = 3, f'(1) = 1$ 。

解 易得 $f'(x) = 3ax^2 + 2bx + c$ 。

$$\begin{cases} d=2 \\ c=2 \\ a+b+c+d=3 \\ 3a+2b+c=1 \end{cases} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 1 & 1 & 1 & 1 & 3 \\ 3 & 2 & 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & -1 \\ 3 & 2 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 2 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 2 \end{bmatrix} \Rightarrow f(x) = x^3 - 2x^2 + 2x + 2$$

例题 2.4 求二次多项式 $f(x) = ax^2 + bx + c$ 满足 $f(0) = 2, f(1) = 1, f'\left(\frac{1}{2}\right) = 2$ 。

解 易得 $f'(x) = 2ax + b$ 。

$$\begin{cases} c=2 \\ a+b+c=1 \\ a+b=2 \end{cases} \rightarrow \begin{bmatrix} 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -3 \\ 1 & 1 & 0 & 2 \end{bmatrix} \quad \text{显然无解。}$$

例题 2.5 求三次多项式 $f(x) = ax^3 + bx^2 + cx + d$ 满足 $f(0) = 2, f(1) = 1, f'(\frac{1}{2}) = 2$ 。

解 易得 $f'(x) = 3ax^2 + 2bx + c$ 。

$$\begin{cases} d = 2 \\ a + b + c + d = 1 \\ \frac{3}{4}a + b + c = 2 \end{cases} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ \frac{3}{4} & 1 & 1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & -1 \\ \frac{3}{4} & 1 & 1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & -1 \\ -\frac{1}{4} & 0 & 0 & 0 & 3 \end{bmatrix}$$

因此解为 $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -12 \\ -t + 11 \\ t \\ 2 \end{bmatrix}$, 对应 $f(x) = -12x^3 + (-t + 11)x^2 + tx + 2, \forall t \in \mathbb{R}$ 。

2.2 求和符号练习

线性代数中求和符号是一个重要的工具, 有时候可以用来简化问题, 有时候可以用来构造问题。在中学时, 我们知道 $\sum_{i=1}^n a_i = a_1 + a_2 + \cdots + a_n$, 除此之外, 还有多重求和及集合求和。以下给出定义:

定义 2.1 (多重求和及集合求和)

(1)(多重求和) $\sum_{a_1, a_2, \dots, a_m=1}^n x_{a_1 a_2 \dots a_m} := \sum_{a_1, a_2, \dots, a_{m-1}=1}^n x_{a_1 a_2 \dots a_m} \sum_{a_m=1}^n = \cdots = \sum_{a_1=1}^n \sum_{a_2=1}^n \cdots \sum_{a_m=1}^n x_{a_1 a_2 \dots a_m}$

(2)(集合求和) $\sum_{e \in \Lambda} x_e$ 用于表示对集合 Λ 中所有元素 e 的 x_e 值求和。集合 Λ 是索引集合, 可以是有限的或无限的。索引集合包含所有我们要求和的索引 e 。对于集合 Λ 中的每个元素 e , x_e 表示与 e 相关联的数值。 $\sum$ 表示对所有 x_e 进行累加, 其中 e 遍历集合 Λ 的所有元素。例如 $\sum_{i=1}^n a_i = \sum_{e \in \{1, 2, \dots, n\}} a_e$ 。

集合 Λ 也可以被替换为某一条件, 例如 $\sum_{i=1}^n a_i = \sum_{1 \leq i \leq n} a_i, \sum_{1 \leq i < j \leq n} a_{ij} = \sum_{i=1}^n \sum_{j=i+1}^n a_{ij}$

下面是一些求和符号的练习。

例题 2.6 证明: $\sum_{i=1}^n \sum_{j=1}^m a_i b_j = \sum_{j=1}^m \sum_{i=1}^n a_i b_j = \left(\sum_{i=1}^n a_i \right) \left(\sum_{j=1}^m b_j \right)$

证明 显然矩阵 $\begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_m \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_m \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n b_m \end{bmatrix}$ 的元素按行求和等于按列求和, 故得证。

例题 2.7 计算 $\sum_{i=1}^n \sum_{j=1}^n (i+j)^2$

解

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n (i+j)^2 &= \sum_{i=1}^n \sum_{j=1}^n (i^2 + j^2 + 2ij) = n \sum_{i=1}^n i^2 + n \sum_{j=1}^n j^2 + 2 \sum_{i=1}^n \sum_{j=1}^n ij = 2n \cdot \frac{n(n+1)(2n+1)}{6} + 2 \left(\sum_{i=1}^n i \right) \left(\sum_{j=1}^n j \right) \\ &= \frac{n^2(n+1)(2n+1)}{3} + 2 \left(\frac{n(n+1)}{2} \right)^2 = \frac{n^2(n+1)(7n+5)}{6} \end{aligned}$$

例题 2.8(Chebyshev) 若 $a_1 \leq a_2 \leq \cdots \leq a_n, b_1 \leq b_2 \leq \cdots \leq b_n$, 则 $\left(\sum_{i=1}^n a_i \right) \left(\sum_{j=1}^n b_j \right) \leq n \sum_{k=1}^n a_k b_k$ 。

证明 只需证明 $n \sum_{k=1}^n a_k b_k - \left(\sum_{i=1}^n a_i \right) \left(\sum_{j=1}^n b_j \right) = \sum_{1 \leq j < i \leq n} (a_i - a_j)(b_i - b_j) \geq 0$ 。令 $S = \sum_{1 \leq j < i \leq n} (a_i - a_j)(b_i - b_j)$ 。

$$\begin{aligned} 2S &= \sum_{1 \leq j < i \leq n} (a_i - a_j)(b_i - b_j) + \sum_{1 \leq i < j \leq n} (a_j - a_i)(b_j - b_i) = \sum_{i,j=1}^n (a_j - a_i)(b_j - b_i) \\ &= \sum_{i,j=1}^n (a_j b_j + a_i b_i - a_i b_j - a_j b_i) = 2n \sum_{k=1}^n a_k b_k - 2 \sum_{i,j=1}^n a_i b_j = 2 \left(n \sum_{k=1}^n a_k b_k - \left(\sum_{i=1}^n a_i \right) \left(\sum_{j=1}^n b_j \right) \right) \end{aligned}$$

例题 2.9 记 $H_k = \sum_{j=1}^k \frac{1}{j}$, (1) 求 $\sum_{1 \leq j < k \leq n} \frac{1}{k-j}$ (用 H_k 表示), (2) 证明 $\sum_{j=1}^{n-1} H_j = nH_n - n$ 。

解

$$\begin{aligned} (1) \quad \sum_{1 \leq j < k \leq n} \frac{1}{k-j} &= \sum_{k=2}^n \sum_{j=1}^{k-1} \frac{1}{k-j} = \sum_{k=2}^n \sum_{j=1}^{k-1} \frac{1}{j} = \sum_{k=2}^n H_{k-1} = \sum_{k=1}^{n-1} H_k \\ (2) \quad \sum_{j=1}^{n-1} H_j &= \sum_{j=1}^{n-1} \sum_{l=1}^j \frac{1}{l} = \sum_{l=1}^{n-1} \sum_{j=l}^{n-1} \frac{1}{l} = \sum_{l=1}^{n-1} \frac{n-l}{l} = \sum_{l=1}^n \frac{n-l}{l} = \sum_{l=1}^n \left(\frac{n}{l} - 1 \right) = nH_n - n \end{aligned}$$

例题 2.10(Abel) 记 $B_0 = 0, B_n = \sum_{i=1}^n b_i$, 则 $S = \sum_{i=1}^n a_i b_i = B_n a_n - \sum_{i=1}^{n-1} B_i (a_{i+1} - a_i)$ 。

证明

$$\begin{aligned} \sum_{i=1}^n a_i b_i &= \sum_{i=1}^n a_i (B_i - B_{i-1}) = \sum_{i=1}^n a_i B_i - \sum_{i=1}^n a_i B_{i-1} = a_n B_n + \sum_{i=1}^{n-1} a_i B_i - \sum_{i=0}^{n-1} a_{i+1} B_i \\ &= a_n B_n + \sum_{i=1}^{n-1} (a_i - a_{i+1}) B_i - a_1 B_0 = B_n a_n - \sum_{i=1}^{n-1} B_i (a_{i+1} - a_i) \end{aligned}$$

2.3 组合恒等式

引理 2.2

1.

$$\binom{n}{k} = \binom{n}{n-k} \quad (0 \leq k \leq n)$$

2.

$$k \binom{n}{k} = n \binom{n-1}{k-1}$$

3.

$$\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}$$



定理 2.3 (二项式定理)

对 $\forall x \in \mathbb{R}$ 及 $n \in \mathbb{N}^*$, 有 $(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k$ 。



证明 归纳。利用引理 2.2(3), 有

$$(1+x)^n = (1+x) \cdot (1+x)^{n-1} = (1+x) \sum_{k=0}^{n-1} \binom{n-1}{k} x^k = \sum_{k=1}^n \left(\binom{n-1}{k-1} + \binom{n-1}{k} \right) x^k + 1 + x^n = \sum_{k=0}^n \binom{n}{k} x^k$$

推论 2.4

利用上述定理和引理, 可以得到:

1.

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

2.

$$\sum_{\substack{0 \leq k \leq n \\ k \text{ 为偶数}}} \binom{n}{k} = \sum_{\substack{0 \leq k \leq n \\ k \text{ 为奇数}}} \binom{n}{k} = 2^{n-1}$$

3.

$$\binom{m+n}{k} = \sum_{i=0}^k \binom{m}{i} \binom{n}{k-i}. \text{ 特别地, } \binom{2n}{n} = \sum_{i=0}^n \binom{n}{i}^2$$

4.

$$\sum_{k=0}^n k \binom{n}{k} = n 2^{n-1}$$

5.

$$\sum_{k=m}^n \binom{k}{m} = \binom{n+1}{m+1}$$

6.

$$\sum_{k=0}^n \binom{m+k}{k} = \binom{m+n+1}{n}$$



证明 (1)(2) 对 $(1+x)^n$ 使用二项式定理, 分别取 $x=1$ 和 $x=-1$ 即可。

(3) 注意到 $(1+x)^{m+n} = (1+x)^n \cdot (1+x)^m$, 分别使用二项式定理, 展开后比较 x^k 系数。

(4) $(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k$ 两侧对 x 求导, 并取 $x=1$ 即可。

(5)

$$\binom{k}{m} = \text{const} \left(\frac{(1+x)^k}{x^m} \right), \text{ 其中 } \text{const}(f(x)) \text{ 表示 } f(x) \text{ 的常数项}$$

$$\Rightarrow \sum_{k=m}^n \binom{k}{m} = \text{const} \left(\sum_{k=m}^n \frac{(1+x)^k}{x^m} \right) = \text{const} \left(\frac{1}{x^m} \frac{(1+x)^{n+1} - (1+x)^m}{x} \right) = \text{const} \left(\frac{(1+x)^{n+1}}{x^{m+1}} \right) = \binom{n+1}{m+1}.$$

(6) 利用引理 2.2(3) 归纳即可。

例题 2.11 设 $n \geq m$, 证明: $\sum_{k=0}^m \binom{m}{k} \binom{n+k}{m} = \sum_{k=0}^m \binom{m}{k} \binom{n}{k} 2^k$ 。

证明

$$\begin{aligned} \sum_{k=0}^m \binom{m}{k} \binom{n+k}{m} &= [x^m](1+x)^n \sum_{k=0}^m \binom{m}{k} (1+x)^k = [x^m](1+x)^n (2+x)^m \\ &= [x^m](1+x)^n \sum_{k=0}^m \binom{m}{k} 2^k x^{m-k} = \sum_{k=0}^m \binom{m}{k} \binom{n}{k} 2^k \end{aligned}$$

其中 $[x^m]f(x)$ 表示多项式 $f(x)$ 的 m 次项系数。

第三章 行列式

3.1 行列式的定义 *

行列式具有多种定义，以下只列出常见的三种，请读者自行验证它们的等价性。

逆序数定义又称为行列式的完全展开定义。具体表述为：

定义 3.1 (逆序数定义)

设 $A = (a_{ij})_{n \times n}$ 是 n 阶方阵，定义行列式

$$\det A = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} := \sum_{\sigma \in S_n} 2(\chi_{A_n}(\sigma)) - \frac{1}{2} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

其中 χ 是示性函数， S_n 是 n 阶置换群， A_n 是 n 阶交错群。

 **笔记** $2(\chi_{A_n}(\sigma)) - \frac{1}{2}$ 本质上是检验置换 σ 的奇偶性。对 $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$, $2(\chi_{A_n}(\sigma)) - \frac{1}{2} \equiv (-1)^{\tau(j_1, j_2, \cdots, j_n)}$ 。

归纳法定义又称为按行展开法定义。该定义直接给出了行列式的计算式（展开式）。

定义 3.2 (归纳法定义)

设 $A = (a_{ij})_{n \times n}$ 是 n 阶方阵，定义行列式 $\det A$ 是按下列法则确定的一个数：

1. 当 $n = 1$ 时， $\det a_{11} = a_{11}$ ；
2. 当 $n \geq 2$ 时， $\det A = \sum_{j=1}^n a_{1j} A_{1j}$ ，其中 A_{1j} 是 A 中元 a_{1j} 的代数余子式。

由这一定义可以得到：

定理 3.3 (Laplace 展开定理)

设 A 是 n 阶方阵， $A = (a_{ij})_{n \times n}$ ，对任意正整数 $r < n$ ，任意取定 r 个指标 $1 \leq i_1 < i_2 < \cdots < i_r \leq n$ ，则 $\det A$ 的值等于它的第 $i_1, i_2, \cdots, i_r$ 行（或列）元组成的所有 r 阶子式分别与它们的代数余子式的乘积之和。即

$$\begin{aligned} \det A &= \sum_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} A \begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} (-1)^{\sum_{k=1}^r (i_k + j_k)} A \begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ j_{r+1} & j_{r+2} & \cdots & j_n \end{pmatrix} \\ &= \sum_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} A \begin{pmatrix} j_1 & j_2 & \cdots & j_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} (-1)^{\sum_{k=1}^r (i_k + j_k)} A \begin{pmatrix} j_{r+1} & j_{r+2} & \cdots & j_n \\ i_{r+1} & i_{r+2} & \cdots & i_n \end{pmatrix} \end{aligned}$$

其中 $i_{r+1}, i_{r+2}, \cdots, i_n$ 是由 $1, 2, \cdots, n$ 去掉 $i_1, i_2, \cdots, i_r$ 后剩下的数按从小到大排列得到， $j_{r+1}, j_{r+2}, \cdots, j_n$ 是由 $1, 2, \cdots, n$ 去掉 $j_1, j_2, \cdots, j_r$ 后剩下的数按从小到大排列得到。

证明 此定理的证明已超出线性代数 B1 的范围，但是可以参考《线性代数（李炯生 查建国 王新茂）》教材的 2.3 节。

例题 3.1 设 $\det A$ 是 n 阶行列式，正整数 $r < n$ 。若 $\det A$ 的所有 r 阶子式都等于零，证明： $\det A = 0$ 。

证明 由 Laplace 展开定理， $\det A$ 的 r 阶子式等于零，有 $\det A$ 的展开式中求和的每一项都是 0，故 $\det A = 0$ 。
行列式的公理化定义将行列式归结为满足 3 条性质的线性映射，而不是直接给出行列式的具体形式。

定义 3.4 (公理化定义)

n 维向量空间 $\mathbb{F}^n$ 上的规范反对称 n 重线性函数称为数域 $\mathbb{F}$ 上 n 阶行列式函数, 简称 n 阶行列式。

换言之, 行列式函数 $\det: \mathbb{F}^n \times \mathbb{F}^n \times \cdots \times \mathbb{F}^n \rightarrow \mathbb{F}$ 是满足以下三条性质的函数:

1. 规范性: $\det(e_1, e_2, \cdots, e_n) = 1$, 其中 $e_1, e_2, \cdots, e_n$ 是 $\mathbb{F}^n$ 的一组标准正交基;
2. 反对称性: $\forall 1 \leq i < j \leq n, \det(\cdots, \alpha_i, \cdots, \alpha_j, \cdots) = -\det(\cdots, \alpha_j, \cdots, \alpha_i, \cdots)$;
3. 线性性: $\forall \lambda, \mu \in \mathbb{F}, \det(\lambda \alpha_1 + \mu \beta_1, \alpha_2, \cdots, \alpha_n) = \lambda \det(\alpha_1, \alpha_2, \cdots, \alpha_n) + \mu \det(\beta_1, \alpha_2, \cdots, \alpha_n)$ 。



3.2 行列式计算练习

矩阵/行列式打洞的神秘技巧并不是线性代数的核心, 但为了考试, 请大家务必掌握。

3.2.1 定义法/降阶法/化为上三角

例题 3.2 计算 $\det \begin{bmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & -4 & -2 \\ 2 & -1 & 4 & 4 \\ 2 & 3 & -3 & 2 \end{bmatrix}$ 。

解

$$\begin{aligned} \det \begin{bmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & -4 & -2 \\ 2 & -1 & 4 & 4 \\ 2 & 3 & -3 & 2 \end{bmatrix} &\xrightarrow[r_1 \rightarrow r_2]{\substack{-2r_1 \rightarrow r_3 \\ -2r_1 \rightarrow r_4}} \det \begin{bmatrix} 1 & 0 & 1 & -4 \\ 0 & -3 & -3 & -6 \\ 0 & -1 & 2 & 12 \\ 0 & 3 & -5 & 10 \end{bmatrix} \xrightarrow{r_2 \leftrightarrow r_3} -\det \begin{bmatrix} 1 & 0 & 1 & -4 \\ 0 & -1 & 2 & 12 \\ 0 & -3 & -3 & -6 \\ 0 & 3 & -5 & 10 \end{bmatrix} \\ &\xrightarrow[3r_2 \rightarrow r_4]{-3r_2 \rightarrow r_3} -\det \begin{bmatrix} 1 & 0 & 1 & -4 \\ 0 & -1 & 2 & 12 \\ 0 & 0 & -9 & -42 \\ 0 & 0 & 1 & 46 \end{bmatrix} = -1 \cdot (-1) \cdot \det \begin{bmatrix} -9 & -42 \\ 1 & 46 \end{bmatrix} = -372 \end{aligned}$$

例题 3.3 计算 $\det \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 0 & 5 & 4 & 0 & 0 \\ 9 & 8 & 7 & 6 & 0 \\ 13 & 12 & 11 & 0 & 10 \end{bmatrix}$ 。

解

$$\begin{aligned} \det \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 0 & 5 & 4 & 0 & 0 \\ 9 & 8 & 7 & 6 & 0 \\ 13 & 12 & 11 & 0 & 10 \end{bmatrix} &\xrightarrow{\text{第一行展开}} (-1)^{1+4} \det \begin{bmatrix} 0 & 0 & 3 & 0 \\ 0 & 5 & 4 & 0 \\ 9 & 8 & 7 & 0 \\ 13 & 12 & 11 & 10 \end{bmatrix} \xrightarrow{\text{第一行展开}} -(-1)^{1+3} \cdot 3 \det \begin{bmatrix} 0 & 5 & 0 \\ 9 & 8 & 0 \\ 13 & 12 & 10 \end{bmatrix} \\ &\xrightarrow{\text{第一行展开}} -(-1)^{1+2} \cdot 15 \det \begin{bmatrix} 9 & 0 \\ 13 & 10 \end{bmatrix} = 15 \cdot 90 = 1350 \end{aligned}$$

例题 3.4 计算 $\det \begin{bmatrix} x & y & & \\ & x & y & \\ & & x & \ddots \\ & & & \ddots & y \\ y & & & & x \end{bmatrix}_{n \times n}$ 。

解 $\det \begin{bmatrix} x & y & & \\ & x & y & \\ & & x & \ddots \\ & & & \ddots & y \\ y & & & & x \end{bmatrix} \xrightarrow{\text{第一列展开}} x \cdot \det \begin{bmatrix} x & y & & \\ & x & y & \\ & & x & \ddots \\ & & & \ddots & y \\ & & & & x \end{bmatrix} + (-1)^{n+1} y \cdot \det \begin{bmatrix} y & & & \\ x & y & & \\ & x & y & \\ & & x & \ddots \\ & & & x & y \end{bmatrix} = x^n + (-1)^{n+1} y^n$

 笔记 有一定规律 (元素大量重复) 矩阵化为上三角的重要结构: $(1, 1, \dots, 1)$

例题 3.5 计算 $\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix}$

解

$$\begin{aligned} \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix} &\xrightarrow{\substack{-ar_1 \rightarrow r_2 \\ -a^2 r_1 \rightarrow r_3 \\ -a^4 r_1 \rightarrow r_4}} \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & b-a & c-a & d-a \\ 0 & b^2-a^2 & c^2-a^2 & d^2-a^2 \\ 0 & b^4-a^4 & c^4-a^4 & d^4-a^4 \end{bmatrix} = \det \begin{bmatrix} b-a & c-a & d-a \\ b^2-a^2 & c^2-a^2 & d^2-a^2 \\ b^4-a^4 & c^4-a^4 & d^4-a^4 \end{bmatrix} \\ &= (b-a)(c-a)(d-a) \det \begin{bmatrix} 1 & 1 & 1 \\ a+b & a+c & a+d \\ (a^2+b^2)(a+b) & (a^2+c^2)(a+c) & (a^2+d^2)(a+d) \end{bmatrix} \\ &= (b-a)(c-a)(d-a) \det \begin{bmatrix} 1 & 1 & 1 \\ 0 & c-b & d-b \\ 0 & (a^2+c^2)(a+c) - (a^2+b^2)(a+b) & (a^2+d^2)(a+d) - (a^2+b^2)(a+b) \end{bmatrix} \\ &= (b-a)(c-a)(d-a) \det \begin{bmatrix} c-b & d-b \\ (c-b)(a^2+b^2+c^2+ab+bc+ac) & (d-b)(a^2+b^2+d^2+ab+bd+ad) \end{bmatrix} \\ &= (b-a)(c-a)(d-a)(c-b)(d-b)(d-c)(a+b+c+d) \end{aligned}$$

 笔记 如果没有 $(1, 1, \dots, 1)$, 一般先尝试逐行/逐列求和。

例题 3.6 计算 $\det \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$

解 $\det \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} 3 & 3 & 3 & 3 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} = 3 \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\substack{c_2 \leftrightarrow c_3 \\ c_1 \leftrightarrow c_4}} 3 \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = -3$

例题 3.7 计算 $\det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix}_{n \times n}$

解

$$\begin{aligned} \det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix} &= \det \begin{bmatrix} x+(n-1)a & x+(n-1)a & \cdots & x+(n-1)a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix} \\ &= (x+(n-1)a) \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & x-a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x-a \end{bmatrix} = (x+(n-1)a)(x-a)^{n-1} \end{aligned}$$

3.2.2 加边

首先回忆这个典型：

例题 3.8 计算 $\Delta = \det \begin{bmatrix} a & b & b & \cdots & b \\ c & d & & & \\ c & & d & & \\ \vdots & & & \ddots & \\ c & & & & d \end{bmatrix}_{n \times n}$

解 若 $d = 0$ ，显然有 $\Delta = 0$ 。

$$d \neq 0 \text{ 时, } \det \begin{bmatrix} a & b & b & \cdots & b \\ c & d & & & \\ c & & d & & \\ \vdots & & & \ddots & \\ c & & & & d \end{bmatrix} \xrightarrow[\forall k=2, \dots, n]{-\frac{c}{d}c_k \rightarrow c_1} \det \begin{bmatrix} a - \frac{bc}{d}(n-1) & b & b & \cdots & b \\ 0 & d & & & \\ 0 & & d & & \\ \vdots & & & \ddots & \\ 0 & & & & d \end{bmatrix} = d^{n-1} \left(a - \frac{bc}{d}(n-1) \right)$$

这启发了对前面一道例题的另外一种解法。

例题 3.9 计算 $\det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix}_{n \times n}$

$$\begin{aligned} \det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix} &= \det \begin{bmatrix} 1 & a & a & \cdots & a \\ 0 & x & a & \cdots & a \\ 0 & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a & a & \cdots & x \end{bmatrix} = \det \begin{bmatrix} 1 & a & a & \cdots & a \\ -1 & x-a & & & \\ -1 & & x-a & & \\ \vdots & & & \ddots & \\ -1 & & & & x-a \end{bmatrix} = (x-a)^n \left(1 + \frac{a}{x-a}n \right) \end{aligned}$$

显然，两种方法的结果是一致的。

 **笔记** 如果出现大量重复元素，但是找不到 $(1, 1, \dots, 1)$ 时，可以尝试加边。

先介绍一个结论：

命题 3.5

设准上三角形方阵 $A = (A_{ij})_{k \times k}$ 的每个对角块 A_{kk} 都是方阵, 则有

$$\det \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ O & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & A_{kk} \end{bmatrix} = \prod_{j=1}^k \det A_{jj}$$



证明 首先考虑 $k=2$ 的情况, 设

$$A = (a_{ij})_{n \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{bmatrix}$$

考虑 $\det A$ 的完全展开式中的非零项, 若 $\prod_{i=1}^n a_{ij_i} \neq 0$, 则必然有 $(j_{r+1}, j_{r+2}, \cdots, j_n)$ 是 $(r+1, r+2, \cdots, n)$ 的一个排列, 进而 $(j_1, j_2, \cdots, j_r)$ 是 $(1, 2, \cdots, r)$ 的一个排列, 故 $\tau(j_1, j_2, \cdots, j_n) = \tau(j_1, j_2, \cdots, j_r) + \tau(j_{r+1}, j_{r+2}, \cdots, j_n)$ 。因此,

$$\begin{aligned} \det A &= \sum_{\substack{1 \leq j_1, \cdots, j_r \leq r \\ r+1 \leq j_{r+1}, \cdots, j_n \leq n}} (-1)^{\tau(j_1, j_2, \cdots, j_r) + \tau(j_{r+1}, j_{r+2}, \cdots, j_n)} \prod_{i=1}^n a_{ij_i} \\ &= \left(\sum_{1 \leq j_1, \cdots, j_r \leq r} (-1)^{\tau(j_1, j_2, \cdots, j_r)} \prod_{i=1}^r a_{ij_i} \right) \left(\sum_{r+1 \leq j_{r+1}, \cdots, j_n \leq n} (-1)^{\tau(j_{r+1}, j_{r+2}, \cdots, j_n)} \prod_{i=r+1}^n a_{ij_i} \right) \\ &= \det A_{11} \det A_{22} \end{aligned}$$

当 $k \geq 3$ 时 A 可以分块成 $\begin{bmatrix} A_{11} & A_{12} \\ O & B \end{bmatrix}$, 从而 $\det A = \det A_{11} \det B$ 。同理 B 可以继续分块, 故有 $\det A = \prod_{j=1}^k \det A_{jj}$ 。

例题 3.10 计算 $\det \begin{bmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{bmatrix}$

解

$$\det \begin{bmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1 & 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{bmatrix} = \det \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & 1+x_n^2 & \cdots & x_n^n \end{bmatrix}$$

$$\begin{aligned}
&= \det \begin{bmatrix} -1 & -1 & -1 & \cdots & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & 1+x_n^2 & \cdots & x_n^n \end{bmatrix} + \det \begin{bmatrix} 2 & 0 & 0 & \cdots & 0 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & 1+x_n^2 & \cdots & x_n^n \end{bmatrix} \\
&\stackrel{\text{前者 } -c_i \rightarrow c_{i-1}}{=} \det \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & x_1-1 & x_1(x_1-1) & \cdots & x_1^{n-1}(x_1-1) \\ 1 & x_2-1 & x_2(x_2-1) & \cdots & x_2^{n-1}(x_2-1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n-1 & x_n(x_n-1) & \cdots & x_n^{n-1}(x_n-1) \end{bmatrix} + 2 \det \begin{bmatrix} x_1 & x_1^2 & \cdots & x_1^n \\ x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \\
&= - \prod_{i=1}^n (x_i - 1) \cdot \det \begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{bmatrix} + 2 \prod_{i=1}^n x_i \cdot \det \begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{bmatrix} \\
&= (2 \prod_{i=1}^n x_i - \prod_{i=1}^n (x_i - 1)) \prod_{1 \leq i < j \leq n} (x_j - x_i)
\end{aligned}$$

最后一步使用到 Vandermonde 行列式的性质，会在后续提及。

除了加 $(1, 1, \dots, 1)$ 之外，还有可能需要添加其他类型的边，这需要根据行列式的特殊性质来判断。

例题 3.11 计算 $\det$

$$\begin{bmatrix} 1 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & \cdots & 3 \\ 3 & 3 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & & \vdots \\ 3 & 3 & 3 & \cdots & n \end{bmatrix}$$

解

$$\begin{aligned}
\det \begin{bmatrix} 1 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & \cdots & 3 \\ 3 & 3 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & & \vdots \\ 3 & 3 & 3 & \cdots & n \end{bmatrix} &= \det \begin{bmatrix} 1 & 3 & 3 & 3 & \cdots & 3 \\ 0 & 1 & 3 & 3 & \cdots & 3 \\ 0 & 3 & 2 & 3 & \cdots & 3 \\ 0 & 3 & 3 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 3 & 3 & 3 & \cdots & n \end{bmatrix} = \det \begin{bmatrix} 1 & 3 & 3 & 3 & \cdots & 3 \\ -1 & -2 & 0 & 0 & \cdots & 0 \\ -1 & 0 & -1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & 0 & \cdots & n-3 \end{bmatrix} \\
&\stackrel{\text{第四行展开}}{=} (-1)(-1)^{1+4} \det \begin{bmatrix} 3 & 3 & 3 & 3 & \cdots & 3 & 3 \\ -2 & 0 & 0 & & & & \\ 0 & -1 & 0 & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & n-4 & \\ & & & & & & n-3 \end{bmatrix} \stackrel{\text{命题 3.5}}{=} \det \begin{bmatrix} 3 & 3 & 3 \\ -2 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} (n-3)! = 6(n-3)!
\end{aligned}$$

例题 3.12 证明: $\det \begin{bmatrix} a_{11}+x_1 & a_{12}+x_2 & \cdots & a_{1n}+x_n \\ a_{21}+x_1 & a_{22}+x_2 & \cdots & a_{2n}+x_n \\ \vdots & \vdots & & \vdots \\ a_{n1}+x_1 & a_{n2}+x_2 & \cdots & a_{nn}+x_n \end{bmatrix} = \det A + \sum_{j=1}^n x_j \sum_{k=1}^n A_{kj}.$

证明

$$\begin{aligned} \det \begin{bmatrix} a_{11}+x_1 & a_{12}+x_2 & \cdots & a_{1n}+x_n \\ a_{21}+x_1 & a_{22}+x_2 & \cdots & a_{2n}+x_n \\ \vdots & \vdots & & \vdots \\ a_{n1}+x_1 & a_{n2}+x_2 & \cdots & a_{nn}+x_n \end{bmatrix} &= \det \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 0 & a_{11}+x_1 & a_{12}+x_2 & \cdots & a_{1n}+x_n \\ 0 & a_{21}+x_1 & a_{22}+x_2 & \cdots & a_{2n}+x_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_{n1}+x_1 & a_{n2}+x_2 & \cdots & a_{nn}+x_n \end{bmatrix} \\ &= \det \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ -1 & a_{11} & a_{12} & \cdots & a_{1n} \\ -1 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \det \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 0 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \det \begin{bmatrix} 0 & x_1 & x_2 & \cdots & x_n \\ -1 & a_{11} & a_{12} & \cdots & a_{1n} \\ -1 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\ &\stackrel{\text{第一行展开}}{=} \det A + \sum_{j=1}^n x_j (-1)^{j+1+1} (-1) \det \begin{bmatrix} 1 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix} \\ &\stackrel{\text{第一列展开}}{=} \det A + \sum_{j=1}^n x_j (-1)^{j+3} \sum_{k=1}^n (-1)^{1+k} M_{kj} = \det A + \sum_{j=1}^n x_j \sum_{k=1}^n (-1)^{j+k+4} A_{kj} \\ &= \det A + \sum_{j=1}^n x_j \sum_{k=1}^n A_{kj} \end{aligned}$$

例题 3.13 计算 $\det \begin{bmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{bmatrix}$

解 $n=1, 2$ 时自行计算。 $n \geq 3$ 时

$$\begin{aligned} \det \begin{bmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{bmatrix} &= \det \begin{bmatrix} 1 & b_1 & b_2 & \cdots & b_n \\ 0 & a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ 0 & a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{bmatrix} = \det \begin{bmatrix} 1 & b_1 & b_2 & \cdots & b_n \\ 1 & a_1 & a_1 & \cdots & a_1 \\ 1 & a_2 & a_2 & \cdots & a_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n & \cdots & a_n \end{bmatrix} \\ &\stackrel{\text{第一列展开}}{=} \sum_{k=1}^n (-1)^k \det \begin{bmatrix} b_1 & \cdots & b_n \\ a_1 & \cdots & a_1 \\ \vdots & & \vdots \\ a_{k-1} & \cdots & a_{k-1} \\ a_{k+1} & \cdots & a_{k+1} \\ \vdots & & \vdots \\ a_n & \cdots & a_n \end{bmatrix} = 0 \end{aligned}$$

例题 3.14 计算 $\det \begin{bmatrix} 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & & \vdots \\ x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{bmatrix}$

解

$$\begin{aligned} \det \begin{bmatrix} 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & & \vdots \\ x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{bmatrix} &= \det \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 0 & 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ 0 & x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{bmatrix} = \det \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ -x_1 & 1 & & & \\ -x_2 & & 1 & & \\ \vdots & & & \ddots & \\ -x_n & & & & 1 \end{bmatrix} \\ &= \det \begin{bmatrix} 1+\sum_{i=1}^n x_i^2 & x_1 & x_2 & \cdots & x_n \\ 0 & 1 & & & \\ 0 & & 1 & & \\ \vdots & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} = 1 + \sum_{i=1}^n x_i^2 \end{aligned}$$

3.2.3 逐行/列加减

对相邻行与行/列与列之间关联度高 (如相差一定倍数) 时, 可以采用逐行/列加减的方法。该方法也适用于行之间差距不大的情形。

例题 3.15 计算 $\det \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & x & a_2 & \cdots & a_{n-1} & a_n \\ 1 & a_1 & x & \cdots & a_{n-1} & a_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & a_1 & a_2 & \cdots & x & a_n \\ 1 & a_1 & a_2 & \cdots & a_{n-1} & x \end{bmatrix}$

解 $\det \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & x & a_2 & \cdots & a_{n-1} & a_n \\ 1 & a_1 & x & \cdots & a_{n-1} & a_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & a_1 & a_2 & \cdots & x & a_n \\ 1 & a_1 & a_2 & \cdots & a_{n-1} & x \end{bmatrix} \xrightarrow[\forall k=2, \dots, n+1]{-r_1 \rightarrow r_k} \det \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 0 & x-a_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & x-a_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x-a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & x-a_n \end{bmatrix} = \prod_{k=1}^n (x-a_k)$

例题 3.16 计算 $\det \begin{bmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{bmatrix}$

解

$$\det \begin{bmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{bmatrix} = \det \begin{bmatrix} a^2 & 2a+1 & 2a+3 & 2a+5 \\ b^2 & 2b+1 & 2b+3 & 2b+5 \\ c^2 & 2c+1 & 2c+3 & 2c+5 \\ d^2 & 2d+1 & 2d+3 & 2d+5 \end{bmatrix} = \det \begin{bmatrix} a^2 & 2a+1 & 2 & 2 \\ b^2 & 2b+1 & 2 & 2 \\ c^2 & 2c+1 & 2 & 2 \\ d^2 & 2d+1 & 2 & 2 \end{bmatrix} = 0$$

例题 3.17 计算 $\det \begin{bmatrix} 1 & a & a^2 & \cdots & a^{n-1} \\ a^{n-1} & 1 & a & \cdots & a^{n-2} \\ a^{n-2} & a^{n-1} & 1 & \cdots & a^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a^2 & a^3 & \cdots & 1 \end{bmatrix}$

解 $\det \begin{bmatrix} 1 & a & a^2 & \cdots & a^{n-1} \\ a^{n-1} & 1 & a & \cdots & a^{n-2} \\ a^{n-2} & a^{n-1} & 1 & \cdots & a^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a^2 & a^3 & \cdots & 1 \end{bmatrix} \xrightarrow[\forall k=1, \dots, n-1]{-ar_{k+1} \rightarrow r_k} \det \begin{bmatrix} 1-a^n & 0 & 0 & \cdots & 0 \\ 0 & 1-a^n & 0 & \cdots & 0 \\ 0 & 0 & 1-a^n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a^2 & a^3 & \cdots & 1-a^n \end{bmatrix} = (1-a^n)^{n-1}$

 **笔记** 这题中的矩阵被称为轮换矩阵，我们会在后面讨论一般情形下的行列式。

例题 3.18 计算 $f(x) = \det \begin{bmatrix} x & -1 & & & \\ & x & -1 & & \\ & & x & -1 & \\ & & & \ddots & \ddots \\ & & & & x & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & x+a_{n-1} \end{bmatrix}$

解

$$\det \begin{bmatrix} x & -1 & & & \\ & x & -1 & & \\ & & x & -1 & \\ & & & \ddots & \ddots \\ & & & & x & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & x+a_{n-1} \end{bmatrix} \xrightarrow[\forall k=n-1, \dots, 1]{xc_{k+1} \rightarrow c_k} \begin{bmatrix} 0 & -1 & & & \\ & 0 & -1 & & \\ & & 0 & -1 & \\ & & & \ddots & \ddots \\ & & & & 0 & -1 \\ x^n + \sum_{k=0}^{n-1} a_k x^k & * & * & \cdots & * & x+a_{n-1} \end{bmatrix}$$

$$\xrightarrow{\text{第一列展开}} (x^n + \sum_{k=0}^{n-1} a_k x^k)(-1)^{n+1} \det(-I_{n-1}) = x^n + \sum_{k=0}^{n-1} a_k x^k$$

 **笔记** 实际上，这个例子并不平凡。题中所给出的矩阵与一个特殊的矩阵相关，这个矩阵是多项式 $f(x)$ 的友矩阵。以下给出定义：

定义 3.6 (友矩阵)

设 $f(x) = x^n + \sum_{k=0}^{n-1} a_k x^k$ 是 $\mathbb{F}$ 上的 n 次首一多项式，则称 $A = - \begin{bmatrix} 0 & -1 & & & \\ & 0 & -1 & & \\ & & 0 & -1 & \\ & & & \ddots & \ddots \\ & & & & 0 & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \end{bmatrix}$ 为 $f(x)$ 的友矩阵。此时显然有 $\det(xI_n - A) = f(x)$ ，即 A 的特征多项式为 $f(x)$ 。

 **笔记** 友矩阵可用于求数列通项，同时也与最小多项式，给定域上的相似标准型有关。

例题 3.19 计算 $\Delta = \det$

$$\begin{bmatrix} 2 & -1 & & & & \\ & 2 & -1 & & & \\ & & 2 & -1 & & \\ & & & \ddots & \ddots & \\ & & & & 2 & -1 \\ 2 & 2 & 2 & \cdots & 2 & 2 \end{bmatrix}$$

解 令 $f(x) = x^n + \sum_{k=0}^{n-2} 2x^k$, 则 $\Delta = f(2) = 2^{n+1} - 2$ 。

例题 3.20 计算 Vandermonde 行列式:

$$\Delta_n(a_1, a_2, \dots, a_n) = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$$

解

$$\begin{aligned} \Delta_n(a_1, a_2, \dots, a_n) &\xrightarrow[\forall k=n, \dots, 2]{-a_n c_{k-1} \rightarrow c_k} \det \begin{bmatrix} 1 & a_1 - a_n & \cdots & a_1^{n-2}(a_1 - a_n) \\ 1 & a_2 - a_n & \cdots & a_2^{n-2}(a_2 - a_n) \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} = \prod_{k=1}^{n-1} (a_k - a_n) \det \begin{bmatrix} 1 & 1 & a_1 & \cdots & a_1^{n-2} \\ 1 & 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \\ &\xrightarrow{\text{第一列展开}} (-1)^{n+1} \prod_{k=1}^{n-1} (a_k - a_n) \cdot \det \begin{bmatrix} 1 & a_1 & \cdots & a_1^{n-2} \\ 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & \cdots & a_{n-1}^{n-2} \end{bmatrix} = \prod_{k=1}^{n-1} (a_n - a_k) \Delta_{n-1}(a_1, a_2, \dots, a_{n-1}) \end{aligned}$$

又有 $\Delta_1 = 1, \Delta_2(a_1, a_2) = a_2 - a_1$, 故归纳有 $\Delta_n(a_1, a_2, \dots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$ 。

以下是一些 Vandermonde 行列式的应用。

例题 3.21 计算 $\det$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1+x_1 & 1+x_2 & 1+x_3 & \cdots & 1+x_{n+1} \\ x_1+x_1^2 & x_2+x_2^2 & x_3+x_3^2 & \cdots & x_{n+1}+x_{n+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1}+x_1^n & x_2^{n-1}+x_2^n & x_3^{n-1}+x_3^n & \cdots & x_{n+1}^{n-1}+x_{n+1}^n \end{bmatrix}$$

解

$$\begin{aligned} \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1+x_1 & 1+x_2 & 1+x_3 & \cdots & 1+x_{n+1} \\ x_1+x_1^2 & x_2+x_2^2 & x_3+x_3^2 & \cdots & x_{n+1}+x_{n+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1}+x_1^n & x_2^{n-1}+x_2^n & x_3^{n-1}+x_3^n & \cdots & x_{n+1}^{n-1}+x_{n+1}^n \end{bmatrix} &\xrightarrow[\forall k=1, \dots, n]{-r_k \rightarrow r_{k+1}} \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_{n+1} \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_{n+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^n & x_2^n & x_3^n & \cdots & x_{n+1}^n \end{bmatrix} \\ &= \prod_{1 \leq i < j \leq n+1} (x_j - x_i) \end{aligned}$$

例题 3.22 计算 $\det \begin{bmatrix} a_1^n & a_1^{n-1}b_1 & a_1^{n-2}b_1^2 & \cdots & a_1b_1^{n-1} & b_1^n \\ a_2^n & a_2^{n-1}b_2 & a_2^{n-2}b_2^2 & \cdots & a_2b_2^{n-1} & b_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n+1}^n & a_{n+1}^{n-1}b_{n+1} & a_{n+1}^{n-2}b_{n+1}^2 & \cdots & a_{n+1}b_{n+1}^{n-1} & b_{n+1}^n \end{bmatrix}$

解

$$\det \begin{bmatrix} a_1^n & a_1^{n-1}b_1 & a_1^{n-2}b_1^2 & \cdots & a_1b_1^{n-1} & b_1^n \\ a_2^n & a_2^{n-1}b_2 & a_2^{n-2}b_2^2 & \cdots & a_2b_2^{n-1} & b_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n+1}^n & a_{n+1}^{n-1}b_{n+1} & a_{n+1}^{n-2}b_{n+1}^2 & \cdots & a_{n+1}b_{n+1}^{n-1} & b_{n+1}^n \end{bmatrix} = \prod_{k=1}^{n+1} a_k^n \det \begin{bmatrix} 1 & \frac{b_1}{a_1} & \left(\frac{b_1}{a_1}\right)^2 & \cdots & \left(\frac{b_1}{a_1}\right)^n \\ 1 & \frac{b_2}{a_2} & \left(\frac{b_2}{a_2}\right)^2 & \cdots & \left(\frac{b_2}{a_2}\right)^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \frac{b_{n+1}}{a_{n+1}} & \left(\frac{b_{n+1}}{a_{n+1}}\right)^2 & \cdots & \left(\frac{b_{n+1}}{a_{n+1}}\right)^n \end{bmatrix}$$

$$= \prod_{k=1}^{n+1} a_k^n \cdot \prod_{1 \leq i < j \leq n+1} \left(\frac{b_j}{a_j} - \frac{b_i}{a_i} \right) = \prod_{1 \leq i < j \leq n+1} (a_i b_j - a_j b_i)$$

例题 3.23 (Lagrange 插值) 设 $a_1, \dots, a_n \in \mathbb{F}$ 各不相同, $b_1, \dots, b_n \in \mathbb{F}$, 证明: 存在唯一的 $n-1$ 次多项式 $f(x) \in \mathbb{F}[x]$, 使得 $f(a_i) = b_i, \forall i$.

证明 设 $f(x) = \sum_{k=0}^{n-1} c_k x^k$ 满足条件, 则

$$\begin{cases} c_0 + c_1 a_1 + \cdots + c_{n-1} a_1^{n-1} = b_1 \\ c_0 + c_1 a_2 + \cdots + c_{n-1} a_2^{n-1} = b_2 \\ \vdots \\ c_0 + c_1 a_n + \cdots + c_{n-1} a_n^{n-1} = b_n \end{cases} \Leftrightarrow \begin{bmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

由 Vandermonde 行列式的性质知系数矩阵可逆, 故解存在唯一。

例题 3.24 证明: $\begin{cases} \sum_{k=1}^n x_k = 0 \\ \sum_{k=1}^n x_k^2 = 0 \\ \vdots \\ \sum_{k=1}^n x_k^n = 0 \end{cases}$ 在 $\mathbb{C}$ 上只有零解。

证明 设 $(x_1, x_2, \dots, x_n)^\top$ 是一组非零解, $(y_1, y_2, \dots, y_k)$ 是 $(x_1, x_2, \dots, x_n)$ 中互异非零的值, $(l_1, l_2, \dots, l_k)$ 是对应出现次数, 则 $\sum_{j=1}^k l_j > 0 \Rightarrow l_j$ 不全为 0。此时原方程组即

$$\begin{cases} l_1 y_1 + l_2 y_2 + \cdots + l_k y_k = 0 \\ l_1 y_1^2 + l_2 y_2^2 + \cdots + l_k y_k^2 = 0 \\ \vdots \\ l_1 y_1^n + l_2 y_2^n + \cdots + l_k y_k^n = 0 \end{cases} \xrightarrow{\text{取前 } k \text{ 个}} \begin{bmatrix} y_1 & y_2 & \cdots & y_k \\ y_1^2 & y_2^2 & \cdots & y_k^2 \\ \vdots & \vdots & & \vdots \\ y_1^k & y_2^k & \cdots & y_k^k \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow l_1 = l_2 = \cdots = l_k = 0, \text{ 矛盾。}$$

例题 3.25 证明: 对 $a_1, a_2, \dots, a_n \in \mathbb{N}^*$, $\prod_{k=1}^{n-1} k!$ 整除 $\Delta_n = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$ 。

证明 令 $f_k(x) = x(x-1)\cdots(x-k+1)$, 则 $\frac{f_k(x)}{k!} = \binom{x}{k}$ 。在 $m < n$ 时视 $\binom{m}{n} = 0$ 。

$$\begin{aligned}
 \text{则 } \Delta_n &= \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} \xrightarrow{\text{适当的列变换}} \det \begin{bmatrix} 1 & f_1(a_1) & f_2(a_1) & \cdots & f_{n-1}(a_1) \\ 1 & f_1(a_2) & f_2(a_2) & \cdots & f_{n-1}(a_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & f_1(a_n) & f_2(a_n) & \cdots & f_{n-1}(a_n) \end{bmatrix} \\
 \Rightarrow \frac{\Delta_n}{\prod_{k=1}^{n-1} k!} &= \det \begin{bmatrix} 1 & \frac{f_1(a_1)}{1!} & \frac{f_2(a_1)}{2!} & \cdots & \frac{f_{n-1}(a_1)}{(n-1)!} \\ 1 & \frac{f_1(a_2)}{1!} & \frac{f_2(a_2)}{2!} & \cdots & \frac{f_{n-1}(a_2)}{(n-1)!} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \frac{f_1(a_n)}{1!} & \frac{f_2(a_n)}{2!} & \cdots & \frac{f_{n-1}(a_n)}{(n-1)!} \end{bmatrix} = \begin{bmatrix} 1 & \binom{a_1}{1} & \binom{a_1}{2} & \cdots & \binom{a_1}{n-1} \\ 1 & \binom{a_2}{1} & \binom{a_2}{2} & \cdots & \binom{a_2}{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \binom{a_n}{1} & \binom{a_n}{2} & \cdots & \binom{a_n}{n-1} \end{bmatrix} \in \mathbb{Z}, \text{ 得证。}
 \end{aligned}$$

3.2.4 拆项与递推

对于一些特殊的行列式，可以将其某一行/列拆成两项，从而简化计算或得到递推。

例题 3.26 计算 $\Delta = \det \begin{bmatrix} x+a & x+b & x+c \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{bmatrix}$

证明 记 $e = (1, 1, 1)^\top, \lambda = (x, y, z)^\top$, 则

$$\Delta = \det(\lambda + ae, \lambda + be, \lambda + ce) = \det(\lambda, \lambda + be, \lambda + ce) + \det(ae, \lambda + be, \lambda + ce) = \det(\lambda, be, ce) + \det(ae, \lambda, \lambda) = 0.$$

例题 3.27 证明: $\det \begin{bmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{bmatrix} = 2 \det \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix}$

证明

$$\begin{aligned}
 \det \begin{bmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{bmatrix} &= \det \begin{bmatrix} b & c+a & a+b \\ q & r+p & p+q \\ y & z+x & x+y \end{bmatrix} + \det \begin{bmatrix} c & c+a & a+b \\ r & r+p & p+q \\ z & z+x & x+y \end{bmatrix} \\
 &= \det \begin{bmatrix} b & c+a & a \\ q & r+p & p \\ y & z+x & x \end{bmatrix} + \det \begin{bmatrix} c & a & a+b \\ r & p & p+q \\ z & x & x+y \end{bmatrix} = \det \begin{bmatrix} b & c & a \\ q & r & p \\ y & z & x \end{bmatrix} + \det \begin{bmatrix} c & a & b \\ r & p & q \\ z & x & y \end{bmatrix} = 2 \det \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix}
 \end{aligned}$$

例题 3.28 计算 $\Delta_n = \det \begin{bmatrix} a & b & b & \cdots & b \\ c & a & b & \cdots & b \\ c & c & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ c & c & c & \cdots & a \end{bmatrix}_{n \times n}$

解

$$\Delta_n = \det \begin{bmatrix} c & b & b & \cdots & b \\ c & a & b & \cdots & b \\ c & c & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ c & c & c & \cdots & a \end{bmatrix}_{n \times n} + \det \begin{bmatrix} ca-c & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & c & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & c & c & \cdots & a \end{bmatrix}_{n \times n}$$

$$= \det \begin{bmatrix} c & b & b & \cdots & b \\ 0 & a-b & 0 & \cdots & 0 \\ 0 & c-b & a-b & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & c-b & c-b & \cdots & a-b \end{bmatrix}_{n \times n} + (a-c)\Delta_{n-1} = c(a-b)^{n-1} + (a-c)\Delta_{n-1}$$

对 $b = c$, 由于 $\Delta_1 = a$, 计算得 $\Delta_n = nb(a-b)^{n-1} + (a-b)^n$ 。

对 $b \neq c$, 同理有 $\Delta_n = (a-c)^{n-1} + (a-b)\Delta_{n-1}$, 联立并消去 Δ_{n-1} 得 $\Delta_n = \frac{c(a-b)^n - b(a-c)^n}{c-b}$ 。

 **练习 3.1** (1) 计算 $\Delta_n = \det$

$$\begin{bmatrix} a_1 & b & b & \cdots & b \\ c & a_2 & b & \cdots & b \\ c & c & a_3 & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ c & c & c & \cdots & a_n \end{bmatrix}$$

(2) 计算 $\Delta_n = \det$

$$\begin{bmatrix} d & b & b & \cdots & b & b \\ c & x & a & \cdots & a & a \\ c & a & x & \cdots & a & a \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ c & a & a & \cdots & x & a \\ c & a & a & \cdots & a & x \end{bmatrix}_{n \times n}$$

例题 3.29 计算 $\Delta_n = \det$

$$\begin{bmatrix} a+b & a & & & \\ & b & a+b & a & \\ & & b & a+b & \ddots \\ & & & \ddots & \ddots & a \\ & & & & b & a+b \end{bmatrix}_{n \times n}$$

解

$$\Delta_n \xrightarrow{\text{第一行展开}} (a+b)\Delta_{n-1} - a \det \begin{bmatrix} a+b & a & & \\ & b & a+b & \ddots \\ & & \ddots & \ddots & a \\ & & & b & a+b \end{bmatrix}_{(n-1) \times (n-1)} \xrightarrow{\text{第一列展开}} (a+b)\Delta_{n-1} - ab\Delta_{n-2}$$

$$\Rightarrow \Delta_n - b\Delta_{n-1} = a(\Delta_{n-1} - b\Delta_{n-2}). \text{ 又 } \Delta_1 = a+b, \Delta_2 = a^2 + b^2 + ab \Rightarrow \Delta_n = a^n + b\Delta_{n-1}.$$

若 $a = b$, 直接计算有 $\Delta_n = (n+1)a^n$ 。

若 $a \neq b$, 同理有 $\Delta_n = b^n + a\Delta_{n-1}$ 。联立消去 Δ_{n-1} 得 $\Delta_n = \frac{b^{n+1} - a^{n+1}}{b-a} = \sum_{k=0}^n a^k b^{n-k}$ 。

 **笔记** $a \neq b$ 时最终答案形式上不要求 $a \neq b$, 这是否意味着不必要进行分类讨论?

 **练习 3.2** 计算 $\Delta_n = \det$

$$\begin{bmatrix} a+b & ab & & & \\ 1 & a+b & ab & & \\ & 1 & a+b & \ddots & \\ & & \ddots & \ddots & ab \\ & & & 1 & a+b \end{bmatrix}_{n \times n}$$

事实上, 这类行列式被称为三对角行列式。我们有如下结论:

命题 3.7

$$\text{设 } \Delta_n = \det \begin{bmatrix} a & b & & \\ c & a & b & \\ & c & a & \ddots \\ & & \ddots & \ddots & b \\ & & & c & a \end{bmatrix}_{n \times n}, \quad x, y \in \mathbb{C} \text{ 是 } \lambda^2 - a\lambda + bc = 0 \text{ 的根, 则有 } \Delta_n = \sum_{k=0}^n x^k y^{n-k}.$$

证明 留作练习。

3.2.5 乘法分解

首先介绍几个引理:

引理 3.8

设 $A, B \in \mathbb{F}^{n \times n}$, 则 $\det(AB) = \det A \cdot \det B$, $\det A = \det A^\top$ 。

引理 3.9

- (1) 若 A 可逆, 则 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det(D - CA^{-1}B)$ 。
- (2) 若 D 可逆, 则 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det D \cdot \det(A - BD^{-1}C)$ 。

证明 自行打洞即可。

引理 3.10

- (1) 若 $A \in \mathbb{F}^{n \times n}, B \in \mathbb{F}^{m \times n}$, 则 $\det(A - B^\top B) = \det A \cdot \det(I_m - BA^{-1}B^\top)$
- (2) 若 $A \in \mathbb{F}^{n \times m}, B \in \mathbb{F}^{m \times n}$, 则 $\lambda^m \det(\lambda I_n - AB) = \lambda^n \det(\lambda I_m - BA)$ 。

证明 (1) $\det \begin{bmatrix} A & B^\top \\ B & I_m \end{bmatrix} \xrightarrow{-c_2 B \rightarrow c_1} \det \begin{bmatrix} A - B^\top B & B^\top \\ 0 & I_m \end{bmatrix} = \det(A - B^\top B)$

又由引理 3.9(1), 有 $\det \begin{bmatrix} A & B^\top \\ B & I_m \end{bmatrix} = \det A \cdot \det(I_m - BA^{-1}B^\top)$, 即有 $\det(A - B^\top B) = \det A \cdot \det(I_m - BA^{-1}B^\top)$ 。

$$(2) \det \begin{bmatrix} \lambda I_n & \lambda A \\ B & \lambda I_m \end{bmatrix} = \lambda^n \det \begin{bmatrix} I_n & A \\ B & \lambda I_m \end{bmatrix} \xrightarrow{\text{引理 3.9(1)}} \lambda^n \det(\lambda I_m - BA),$$

$$\det \begin{bmatrix} \lambda I_n & \lambda A \\ B & \lambda I_m \end{bmatrix} = \lambda^m \det \begin{bmatrix} \lambda I_n & A \\ B & I_m \end{bmatrix} \xrightarrow{\text{引理 3.9(2)}} \lambda^m \det(\lambda I_n - AB)$$

例题 3.30 设 $A = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}$, 计算 $\det A$ 。

解 A 行与行相互正交, 列与列相互正交, 因此可以考虑 $AA^\top$ 。 $\det(AA^\top) = \det((a^2+b^2+c^2+d^2)I_4) = (a^2+b^2+c^2+d^2)^4$,

又 $\det(AA^T) = \det A \cdot \det A^T = (\det A)^2$, 故 $\det A = \pm(a^2 + b^2 + c^2 + d^2)$ 。验证符号后可得 $\det A = (a^2 + b^2 + c^2 + d^2)^2$ 。

例题 3.31 计算 $\Delta_n = \det$

$$\begin{bmatrix} a_1 & & & & b_1 \\ & \ddots & & & \\ & & a_n & b_n & \\ & & c_n & d_n & \\ & \ddots & & & \ddots \\ c_1 & & & & d_1 \end{bmatrix}$$

解

$$\begin{aligned} \text{记 } A &= \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}, B = \begin{bmatrix} & & b_1 \\ & \ddots & \\ b_n & & \end{bmatrix}, C = \begin{bmatrix} & & c_n \\ & \ddots & \\ c_1 & & \end{bmatrix}, D = \begin{bmatrix} d_n & & \\ & \ddots & \\ & & d_1 \end{bmatrix}, \text{ 则若 } A \text{ 可逆,} \\ \Delta_n &= \det A \det(D - CA^{-1}B) = \prod_{k=1}^n a_k \cdot \det \left(\begin{bmatrix} d_n & & \\ & \ddots & \\ & & d_1 \end{bmatrix} - \begin{bmatrix} & & c_n \\ & \ddots & \\ c_1 & & \end{bmatrix} \begin{bmatrix} \frac{1}{a_1} & & \\ & \ddots & \\ & & \frac{1}{a_n} \end{bmatrix} \begin{bmatrix} & & b_1 \\ & \ddots & \\ b_n & & \end{bmatrix} \right) \\ &= \prod_{k=1}^n a_k \cdot \det \left(\begin{bmatrix} d_n & & \\ & \ddots & \\ & & d_1 \end{bmatrix} - \begin{bmatrix} \frac{b_n c_n}{a_n} & & \\ & \ddots & \\ \frac{b_1 c_1}{a_1} & & \end{bmatrix} \right) = \prod_{k=1}^n (a_k d_k - b_k c_k) \end{aligned}$$

若 A 不可逆, 则 $f(t) = \det(tI_n + A)$ 关于 t 的 n 次多项式, 由代数基本定理, $f(t)$ 在 $\mathbb{C}$ 有 n 个根, 设为 $x_1, \dots, x_n$ 。记 $\delta = \min_{\substack{1 \leq k \leq n \\ x_k \neq 0}} |x_k| > 0$, 则 $\forall t \in \dot{U}(0, \delta), f(t) \neq 0$ 。故在 $\dot{U}(0, \delta)$ 内取一列 $\{t_m\} \rightarrow 0$, 则对 $\forall m \in \mathbb{N}^*$, 有

$$\det \begin{bmatrix} A + tI_n & B \\ C & D \end{bmatrix} = \prod_{k=1}^n ((a_k + t)d_k - b_k c_k). \text{ 由行列式及伴随矩阵元素的连续性, 可令 } m \rightarrow \infty, \text{ 得到与上面一致的答案。}$$

 **笔记** 本题对不可逆的 A 的处理方法称为扰动法, 在比较多的题型中均有使用场景。

 **练习 3.3** 使用递推法完成例题 3.31。

例题 3.32* 计算 $\det$

$$\begin{bmatrix} 10 & 1 & 2 & 3 & 3 \\ 1 & 10 & 2 & 1 & 0 \\ 2 & 2 & 10 & 3 & x \\ 3 & 1 & 3 & 10 & x \\ 3 & 0 & x & x & 10 \end{bmatrix}.$$

解

$$\begin{aligned} \det \begin{bmatrix} 10 & 1 & 2 & 3 & 3 \\ 1 & 10 & 2 & 1 & 0 \\ 2 & 2 & 10 & 3 & x \\ 3 & 1 & 3 & 10 & x \\ 3 & 0 & x & x & 10 \end{bmatrix} &= \det \begin{bmatrix} 10 & 1 \\ 1 & 10 \end{bmatrix} \det \left(\begin{bmatrix} 10 & 3 & x \\ 3 & 10 & x \\ x & x & 10 \end{bmatrix} - \frac{1}{99} \begin{bmatrix} 2 & 2 \\ 3 & 1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 10 & -1 \\ -1 & 10 \end{bmatrix} \begin{bmatrix} 2 & 3 & 3 \\ 2 & 1 & 0 \end{bmatrix} \right) \\ &= 99 \det \left(\begin{bmatrix} 10 & 3 & x \\ 3 & 10 & x \\ x & x & 10 \end{bmatrix} - \begin{bmatrix} \frac{8}{11} & \frac{8}{11} & \frac{6}{11} \\ \frac{8}{11} & \frac{94}{99} & \frac{29}{33} \\ \frac{6}{11} & \frac{29}{33} & \frac{10}{11} \end{bmatrix} \right) = 99 \det \begin{bmatrix} \frac{102}{11} & \frac{25}{11} & x - \frac{6}{11} \\ \frac{25}{11} & \frac{896}{99} & x - \frac{29}{33} \\ x - \frac{6}{11} & x - \frac{29}{33} & \frac{100}{11} \end{bmatrix} = -1364x^2 + 1950x + 70122 \end{aligned}$$

例题 3.33 计算 $\det \begin{bmatrix} 0 & 2 & 3 & 4 & \cdots & n \\ 2 & 3 & 6 & 8 & \cdots & 2n \\ 3 & 6 & 8 & 12 & \cdots & 3n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ n & 2n & 3 & 4n & \cdots & n^2 - 1 \end{bmatrix}$

解

$$\begin{aligned} \det \begin{bmatrix} 0 & 2 & 3 & 4 & \cdots & n \\ 2 & 3 & 6 & 8 & \cdots & 2n \\ 3 & 6 & 8 & 12 & \cdots & 3n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ n & 2n & 3 & 4n & \cdots & n^2 - 1 \end{bmatrix} &= \det \left(\begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix} - I_n \right) = (-1)^n \det \left(I_n - \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix} \right) \\ &= (-1)^n \det \left(1 - \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \right) = (-1)^n \left(1 - \sum_{k=1}^n k^2 \right) = (-1)^n \left(1 - \frac{n(n+1)(2n+1)}{6} \right) \end{aligned}$$

例题 3.34(2021B2Mid)* $A = (a_{ij})_{n \times n}$, 其中 $a_{ij} = i + j + \delta_{ij}$, $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ 为 Kronecker 符号。求 $\det A$ 。

解

$$\begin{aligned} \det A &= \det \left(I_n + \begin{bmatrix} 2 & 3 & 4 & \cdots & n+1 \\ 3 & 4 & 5 & \cdots & n+2 \\ 4 & 5 & 6 & \cdots & n+2 \\ \vdots & \vdots & \vdots & & \vdots \\ n+1 & n+2 & n+3 & \cdots & 2n \end{bmatrix} \right) = \det \left(I_n + \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & n \end{bmatrix} \begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \right) \\ &= \det \left(I_2 + \begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & n \end{bmatrix} \right) = \det \begin{bmatrix} 1 + \sum_{k=1}^n k & \sum_{k=1}^n k^2 \\ \sum_{k=1}^n 1 & 1 + \sum_{k=1}^n k \end{bmatrix} = -\frac{1}{12}n^4 + \frac{13}{12}n^2 + n + 1 \end{aligned}$$

 笔记 这题还可以使用加边方法完成，但是需要加两次。读者可以自行尝试。

 练习 3.4* 计算 $\det \begin{bmatrix} 1+x_1+y_1 & x_1+y_2 & \cdots & x_1+y_n \\ x_2+y_1 & 1+x_2+y_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_{n-1}+y_n \\ x_n+y_1 & \cdots & x_n+y_{n-1} & x_n+y_n \end{bmatrix}$

例题 3.35* 计算 $\det \begin{bmatrix} (a_0+b_0)^n & (a_0+b_1)^n & \cdots & (a_0+b_n)^n \\ (a_1+b_0)^n & (a_1+b_1)^n & \cdots & (a_1+b_n)^n \\ \vdots & \vdots & & \vdots \\ (a_n+b_0)^n & (a_n+b_1)^n & \cdots & (a_n+b_n)^n \end{bmatrix}$

解

$$\det \begin{bmatrix} (a_0+b_0)^n & (a_0+b_1)^n & \cdots & (a_0+b_n)^n \\ (a_1+b_0)^n & (a_1+b_1)^n & \cdots & (a_1+b_n)^n \\ \vdots & \vdots & & \vdots \\ (a_n+b_0)^n & (a_n+b_1)^n & \cdots & (a_n+b_n)^n \end{bmatrix} = \det \begin{bmatrix} \binom{n}{0}a_0^n & \binom{n}{1}a_0^{n-1} & \cdots & \binom{n}{n} \\ \binom{n}{0}a_1^n & \binom{n}{1}a_1^{n-1} & \cdots & \binom{n}{n} \\ \vdots & \vdots & & \vdots \\ \binom{n}{0}a_n^n & \binom{n}{1}a_n^{n-1} & \cdots & \binom{n}{n} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ b_0 & b_1 & \cdots & b_n \\ \vdots & \vdots & & \vdots \\ b_0^n & b_1^n & \cdots & b_n^n \end{bmatrix}$$

$$= (-1)^{\frac{n(n+1)}{2}} \prod_{k=0}^n \binom{n}{k} \cdot \prod_{0 \leq i < j \leq n} (a_j - a_i)(b_j - b_i)$$

3.3 函数与行列式 *

3.3.1 多项式与行列式

首先回到之前的一个例题:

例题 3.36 计算 $\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix}$

解 考虑 $f(x) = \det \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ a & b & c & d & x \\ a^2 & b^2 & c^2 & d^2 & x^2 \\ a^3 & b^3 & c^3 & d^3 & x^3 \\ a^4 & b^4 & c^4 & d^4 & x^4 \end{bmatrix} \xrightarrow{\text{最后一列展开}} A_4x^4 + A_3x^3 + A_2x^2 + A_1x + A_0$, 所求即为 $-A_3$ 。

由代数学基本定理, $f(x)$ 所有零点为 a, b, c, d 。由韦达定理, $-\frac{A_3}{A_4} = a + b + c + d$ 。

$$\text{又 } A_4 = \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix} = (d-a)(d-b)(d-c)(c-b)(c-a)(b-a),$$

$$\text{故 } \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix} = -A_3 = (a+b+c+d)A_4 = (d-a)(d-b)(d-c)(c-b)(c-a)(b-a)(a+b+c+d).$$

重新考虑 Vandermonde 行列式, 可以得到:

例题 3.37 计算 $\Delta_n(a_1, a_2, \cdots, a_n) = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$ 。

解 考虑 $f(x) = \det \begin{bmatrix} 1 & x & x^2 & \cdots & x^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$ 为 $n-1$ 次多项式, 则 $a_2, a_3, \cdots, a_n$ 为 $f(x)$ 的 $n-1$ 个零点。

$$\Rightarrow f(x) = C \prod_{k=2}^n (x - a_k). \text{ 又 } f(0) = (-1)^{n-1} C \prod_{k=2}^n a_k, \quad f(a_1) = \Delta_n(a_1, a_2, \cdots, a_n),$$

$$f(0) = \det \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \det \begin{bmatrix} a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \prod_{k=2}^n a_k \cdot \Delta_{n-1}(a_2, a_3, \cdots, a_n).$$

$$\Rightarrow C = (-1)^{n-1} \Delta_{n-1}(a_2, a_3, \cdots, a_n), \quad \Delta_n(a_1, a_2, \cdots, a_n) = \Delta_{n-1}(a_2, a_3, \cdots, a_n) \prod_{k=2}^n (a_k - a_1).$$

$$\text{递推得 } \Delta_n(a_1, a_2, \cdots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

3.3.2 行列式的导数与积分

命题 3.11

设 $A(t) = (a_{ij}(t))_{n \times n}$, $a_{ij}(t)$ 连续可导, 则 $\frac{d}{dt} \det A(t) = \sum_{k=1}^n \det \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a'_{k1}(t) & \cdots & a'_{kn}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{bmatrix}.$

特别地, 若 $A(t)$ 可逆, 则 $\frac{d}{dt} \det A(t) = \det A(t) \cdot \text{tr} \left(A^{-1}(t) \cdot \frac{dA(t)}{dt} \right)$, 其中 $\frac{dA(t)}{dt}$ 表示对 $A(t)$ 元素逐个求导后得到的矩阵。



证明 $A(t) = \sum_{j_1 \cdots j_n} (-1)^{\tau(j_1, \cdots, j_n)} \prod_{k=1}^n a_{kj_k}(t)$

$$\begin{aligned} A'(t) &= \sum_{j_1 \cdots j_n} (-1)^{\tau(j_1, \cdots, j_n)} \sum_{l=1}^n \left(a'_{lj_l}(t) \prod_{\substack{k=1 \\ k \neq l}}^n a_{kj_k}(t) \right) = \sum_{l=1}^n \left(\sum_{j_1 \cdots j_n} (-1)^{\tau(j_1, \cdots, j_n)} a'_{lj_l}(t) \prod_{\substack{k=1 \\ k \neq l}}^n a_{kj_k}(t) \right) \\ &= \sum_{l=1}^n \det \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a'_{l1}(t) & \cdots & a'_{ln}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{bmatrix} \end{aligned}$$

例题 3.38 n 是偶数, 计算 $\det$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{bmatrix}$$

解

$$\det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{bmatrix} = n! \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{bmatrix} = (n-1)! \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ \frac{n^2}{2} & \frac{n^3}{3} & \cdots & \frac{n^{n+1}}{n+1} & \frac{n^{n+2}}{n+2} \end{bmatrix}$$

$$\text{令 } f(x) = (n-1)! \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ \frac{x^2}{2} & \frac{x^3}{3} & \cdots & \frac{x^{n+1}}{n+1} & \frac{x^{n+2}}{n+2} \end{bmatrix}, \text{ 则 } f(n) = \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{bmatrix}, f(0) = 0.$$

$$\begin{aligned} \text{由命题 3.11, } f'(x) &= (n-1)! \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ x & x^2 & \cdots & x^n & x^{n+1} \end{bmatrix} = (n-1)! x \cdot \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ 1 & x & \cdots & x^{n-1} & x^n \end{bmatrix} \\ &= (n-1)! x \cdot \prod_{i \leq j \leq n} (j-i) \cdot \prod_{k=1}^n (x-k) \triangleq C \cdot \prod_{k=0}^n (x-k), \text{ 其中 } C = (n-1)! \prod_{i \leq j \leq n} (j-i) \end{aligned}$$

$$\begin{aligned} \Rightarrow f(n) &= \int_0^n f'(x) dx + f(0) = C \int_0^n x(x-1) \cdots (x-n) dx \\ &\stackrel{y=n-x}{=} \frac{C}{2} \left(\int_0^n x(x-1) \cdots (x-n) dx - \int_n^0 (-y)(1-y) \cdots (n-y) dy \right) \\ &= \frac{C}{2} \left(\int_0^n x(x-1) \cdots (x-n) dx + (-1)^{n+3} \int_0^n y(y-1) \cdots (y-n) dy \right) \stackrel{2|n}{=} 0 \end{aligned}$$

$$\text{故 } \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{bmatrix} = 0.$$

3.3.3 行列式与 Taylor 公式

还是回到之前的例题:

$$\text{例题 3.39 计算 } D_n(x) = \det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix}$$

解 在 a 处进行 Taylor 展开, 其中 $(x-a)(a-a) < 0$:

$$D_n(x) = D_n(a) + D'_n(a)(x-a) + \frac{D''_n(a)}{2!}(x-a)^2 + \frac{D_n^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \frac{D_n^{(n)}(a)}{n!}(x-a)^n + \frac{D_n^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

$$D'_n(x) = \det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix} + \det \begin{bmatrix} x & a & \cdots & a \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{bmatrix} + \cdots + \det \begin{bmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = n D_{n-1}(x)$$

$$\Rightarrow D_n^{(k)}(x) = n(n-1) \cdots (n-k+1) D_{n-k}(x), \text{ 又 } D_2(a) = D_3(a) = \cdots = D_n(a) = 0, \text{ 有 } D'_n(a) = \cdots = D_n^{(n-2)}(a) = 0$$

$$\text{而 } D_n^{(n-1)}(x) = n(n-1) \cdots 2 \cdot D_1(x) = n!x, D_n^{(n)}(x) = n!D_0(x) = n!, \text{ 由 } D_n(x) \text{ 是 } n \text{ 次多项式, } D_n^{(n+1)}(x) = 0.$$

$$\Rightarrow D_n(x) = \frac{D_n^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \frac{D_n^{(n)}(a)}{n!}(x-a)^n = na(x-a)^{n-1} + (x-a)^n.$$

例题 3.40 设 $f(x)$ 在区间 I 上 n 阶可导, $\forall x \in I, |f(x)| < M_0, |f^{(n)}(x)| < M_n$. 证明: $\exists M_1, \dots, M_{n-1}$, 使得 $\forall x \in I, k = 1, \dots, n-1, |f^{(k)}(x)| < M_k$.

证明 取非 0 互异 $a_1, \dots, a_{n-1} \in I$, 在 x 处作 Taylor 展开: $f(x+a_i) = f(x) + \sum_{k=1}^n \frac{f^{(k)}(x)}{k!} a_i^k + \frac{f^{(n)}(\xi_i)}{n!} a_i^n$.

$$\Rightarrow \left| \sum_{k=1}^n \frac{f^{(k)}(x)}{k!} a_i^k \right| = \left| f(x+a_i) - f(x) - \frac{f^{(n)}(\xi_i)}{n!} a_i^n \right| \leq 2M_0 + \frac{A}{n!} M_n, \text{ 其中 } A = \max_{1 \leq i \leq n-1} |a_i|^n.$$

令 $A_i(x) = \sum_{k=1}^n \frac{f^{(k)}(x)}{k!} a_i^k$, 则 $|A_i(x)| \leq 2M_0 + \frac{A}{n!} M_n$, 且有

$$\begin{bmatrix} a_1 & \frac{a_1^2}{2!} & \cdots & \frac{a_1^{n-1}}{(n-1)!} \\ a_2 & \frac{a_2^2}{2!} & \cdots & \frac{a_2^{n-1}}{(n-1)!} \\ \vdots & \vdots & & \vdots \\ a_n & \frac{a_n^2}{2!} & \cdots & \frac{a_n^{n-1}}{(n-1)!} \end{bmatrix} \begin{bmatrix} f'(x) \\ f''(x) \\ \vdots \\ f^{(n-1)}(x) \end{bmatrix} = \begin{bmatrix} A_1(x) \\ A_2(x) \\ \vdots \\ A_n(x) \end{bmatrix}, \forall x \in I. \text{ 记 } D_n = \begin{bmatrix} a_1 & \frac{a_1^2}{2!} & \cdots & \frac{a_1^{n-1}}{(n-1)!} \\ a_2 & \frac{a_2^2}{2!} & \cdots & \frac{a_2^{n-1}}{(n-1)!} \\ \vdots & \vdots & & \vdots \\ a_n & \frac{a_n^2}{2!} & \cdots & \frac{a_n^{n-1}}{(n-1)!} \end{bmatrix},$$

$$\text{由 } \det D_n = \prod_{k=1}^{n-1} \frac{a_k}{k!} \cdot \prod_{1 \leq i < j \leq n-1} (a_j - a_i) \neq 0, \text{ 方程组有唯一解 } \begin{bmatrix} f'(x) \\ f''(x) \\ \vdots \\ f^{(n-1)}(x) \end{bmatrix} = D_n^{-1} \begin{bmatrix} A_1(x) \\ A_2(x) \\ \vdots \\ A_n(x) \end{bmatrix},$$

不妨设为 $f^{(k)}(x) = \sum_{j=1}^n \lambda_{jk} A_j(x)$, 则 $|f^{(k)}(x)| \leq \sum_{j=1}^n |\lambda_{jk}| |A_j(x)| \leq \left(2M_0 + \frac{A}{n!} M_n \right) \sum_{j=1}^n |\lambda_{jk}| \triangleq M_k$.

3.4 综合练习 *

例题 3.41(轮换矩阵的行列式) 设 $\text{circ}(a_1, \dots, a_n) = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \vdots & \vdots & & \vdots \\ a_2 & a_3 & \cdots & a_1 \end{bmatrix}$, 计算 $\det \text{circ}(a_1, \dots, a_n)$.

解 令 $f(x) = \sum_{k=0}^{n-1} a_{k+1} x^k$, 记 $\omega_j = e^{\frac{2j\pi i}{n}}$, 则 $\omega_j \cdot \omega_k = \omega_{j+k}$. 记 $\Omega = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \omega_1 & \omega_2 & \cdots & \omega_n \\ \vdots & \vdots & & \vdots \\ \omega_1^{n-1} & \omega_2^{n-1} & \cdots & \omega_n^{n-1} \end{bmatrix}$,

则 $\det \Omega = \prod_{i \leq j < k \leq n} (\omega_k - \omega_j) \neq 0$. 记 $\Lambda = (\lambda_{ij})_{n \times n} = \text{circ}(a_1, \dots, a_n) \Omega$, 则 $\det \Lambda = \det \text{circ}(a_1, \dots, a_n) \cdot \det \Omega$.

考虑 $\lambda_{11} = 1a_1 + \omega_1 a_2 + \cdots + \omega_1^{n-1} a_n = f(\omega_1)$, $\lambda_{21} = 1a_n + \omega_1 a_1 + \cdots + \omega_1^{n-1} a_{n-1} = \omega_1 f(\omega_1)$, 同理可得

$$\Lambda = \begin{bmatrix} f(\omega_1) & f(\omega_2) & \cdots & f(\omega_n) \\ \omega_1 f(\omega_1) & \omega_2 f(\omega_2) & \cdots & \omega_n f(\omega_n) \\ \vdots & \vdots & & \vdots \\ \omega_1^{n-1} f(\omega_1) & \omega_2^{n-1} f(\omega_2) & \cdots & \omega_n^{n-1} f(\omega_n) \end{bmatrix} \Rightarrow \det \Lambda = \prod_{l=1}^n f(\omega_l) \cdot \prod_{i \leq j < k \leq n} (\omega_k - \omega_j)$$

$$\Rightarrow \det \text{circ}(a_1, \dots, a_n) = \frac{\det \Lambda}{\det \Omega} = \prod_{k=1}^n f(\omega_k).$$

 **练习 3.5** 计算 $\text{circ}(a_1, \dots, a_n)^{-1}$

例题 3.42(Cauchy 行列式) 设复数列 $\{x_i\}_{i=1}^n, \{y_j\}_{j=1}^n$ 满足 $x_i \neq y_j, \forall i, j$, 计算

$$\det C_n(x, y) = \det \begin{bmatrix} \frac{1}{x_1 - y_1} & \frac{1}{x_1 - y_2} & \cdots & \frac{1}{x_1 - y_n} \\ \frac{1}{x_2 - y_1} & \frac{1}{x_2 - y_2} & \cdots & \frac{1}{x_2 - y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n - y_1} & \frac{1}{x_n - y_2} & \cdots & \frac{1}{x_n - y_n} \end{bmatrix}$$

解

$$\begin{aligned} & \det \begin{bmatrix} \frac{1}{x_1 - y_1} & \frac{1}{x_1 - y_2} & \cdots & \frac{1}{x_1 - y_n} \\ \frac{1}{x_2 - y_1} & \frac{1}{x_2 - y_2} & \cdots & \frac{1}{x_2 - y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n - y_1} & \frac{1}{x_n - y_2} & \cdots & \frac{1}{x_n - y_n} \end{bmatrix} \\ & \xrightarrow[\forall k=1, \dots, n-1]{-r_n \rightarrow r_k} \det \begin{bmatrix} \frac{x_n - x_1}{(x_1 - y_1)(x_n - y_1)} & \frac{x_n - x_1}{(x_1 - y_2)(x_n - y_2)} & \cdots & \frac{x_n - x_1}{(x_1 - y_n)(x_n - y_n)} \\ \frac{x_n - x_2}{(x_2 - y_1)(x_n - y_1)} & \frac{x_n - x_2}{(x_2 - y_2)(x_n - y_2)} & \cdots & \frac{x_n - x_2}{(x_2 - y_n)(x_n - y_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n - y_1} & \frac{1}{x_n - y_2} & \cdots & \frac{1}{x_n - y_n} \end{bmatrix} \\ & = \frac{1}{x_n - y_n} \prod_{j=1}^{n-1} \frac{x_n - x_j}{x_n - y_j} \cdot \det \begin{bmatrix} \frac{1}{x_1 - y_1} & \frac{1}{x_1 - y_2} & \cdots & \frac{1}{x_1 - y_n} \\ \frac{1}{x_2 - y_1} & \frac{1}{x_2 - y_2} & \cdots & \frac{1}{x_2 - y_n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \\ & \xrightarrow[\forall k=1, \dots, n-1]{-c_n \rightarrow c_k} \frac{1}{x_n - y_n} \prod_{j=1}^{n-1} \frac{x_n - x_j}{x_n - y_j} \cdot \begin{bmatrix} \frac{y_1 - y_n}{(x_1 - y_1)(x_1 - y_n)} & \frac{y_2 - y_n}{(x_1 - y_2)(x_1 - y_n)} & \cdots & \frac{1}{x_1 - y_n} \\ \frac{y_2 - y_n}{(x_2 - y_1)(x_2 - y_n)} & \frac{y_2 - y_n}{(x_2 - y_2)(x_2 - y_n)} & \cdots & \frac{1}{x_2 - y_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ & = \frac{1}{x_n - y_n} \prod_{j=1}^{n-1} \frac{x_n - x_j}{x_n - y_j} \cdot \begin{bmatrix} \frac{y_1 - y_n}{(x_1 - y_1)(x_1 - y_n)} & \frac{y_2 - y_n}{(x_1 - y_2)(x_1 - y_n)} & \cdots & \frac{y_{n-1} - y_n}{(x_1 - y_{n-1})(x_1 - y_n)} \\ \frac{y_2 - y_n}{(x_2 - y_1)(x_2 - y_n)} & \frac{y_2 - y_n}{(x_2 - y_2)(x_2 - y_n)} & \cdots & \frac{y_{n-1} - y_n}{(x_2 - y_{n-1})(x_2 - y_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{y_1 - y_n}{(x_{n-1} - y_1)(x_{n-1} - y_n)} & \frac{y_2 - y_n}{(x_{n-1} - y_2)(x_{n-1} - y_n)} & \cdots & \frac{y_{n-1} - y_n}{(x_{n-1} - y_{n-1})(x_{n-1} - y_n)} \end{bmatrix} \\ & = \frac{1}{x_n - y_n} \prod_{j=1}^{n-1} \frac{(x_n - x_j)(y_j - y_n)}{(x_n - y_j)(x_j - y_n)} \cdot \det \begin{bmatrix} \frac{1}{x_1 - y_1} & \frac{1}{x_1 - y_2} & \cdots & \frac{1}{x_1 - y_{n-1}} \\ \frac{1}{x_2 - y_1} & \frac{1}{x_2 - y_2} & \cdots & \frac{1}{x_2 - y_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_{n-1} - y_1} & \frac{1}{x_{n-1} - y_2} & \cdots & \frac{1}{x_{n-1} - y_{n-1}} \end{bmatrix} \\ & \xrightarrow{\text{递推}} \frac{\prod_{1 \leq i < j \leq n} (x_i - x_j)(y_j - y_i)}{\prod_{i,j=1}^n (x_i - y_j)} \end{aligned}$$

 **笔记** $x_n = n - 1, y_n = -n$ 时, $C_n(x, y)$ 称为 Hilbert 矩阵。Hilbert 矩阵正定且高度病态 (对任意一个元素的微小扰动都会导致行列式和逆的巨大变化)。

第四章 秩与相抵

4.1 秩与相抵的基本概念

在开始之前，先回顾一些秩的定义和性质。

定义 4.1

设 A 至少有一个 r 阶非零子式，且所有 $r+1$ 阶子式均为零，则称 r 为 A 的秩。当 $A = O$ 时， A 不含任何非零子式，定义其秩为 0。

还有一个等价的定义：

定义 4.2

对于 $A \in \mathbb{F}^{m \times n}$ ，存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$ 和 $Q \in \mathbb{F}^{n \times n}$ ，使得 $PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ 。此时，称 r 为 A 的秩。

显然 $m \times n$ 矩阵 A 的非零子式的最高阶数 $\text{rank}(A) \leq \min\{m, n\}$ 。如果 $m = n$ ，当 $\text{rank}(A)$ 达到最大值 n 时称 A 为满秩方阵。 A 是满秩方阵 $\Leftrightarrow \det A \neq 0 \Leftrightarrow A$ 可逆。当 $\text{rank} A = m < n$ 时，称 A 为行满秩矩阵，它的行向量组线性无关。当 $\text{rank} A = n < m$ 时，称 A 为列满秩矩阵，它的列向量组线性无关。以下给出相抵的定义：

定义 4.3

设 $A, B \in \mathbb{F}^{m \times n}$ ，若 A 可通过有限次初等变换变为 B ，则称 A 与 B 相抵。

定理 4.4

设 $A, B \in \mathbb{F}^{m \times n}$ ，则 A 与 B 相抵的充分必要条件是存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$ 和 $Q \in \mathbb{F}^{n \times n}$ ，使得 $B = PAQ$ 。

矩阵的相抵是一个等价关系，因此其满足以下性质：

命题 4.5

设 $A, B, C \in \mathbb{F}^{m \times n}$ ，则

1. 反身性： A 与 A 相抵；
2. 对称性：若 A 与 B 相抵，则 B 与 A 相抵；
3. 传递性：若 A 与 B 相抵，且 B 与 C 相抵，则 A 与 C 相抵。

因此，可以将矩阵集 $\mathbb{F}^{m \times n}$ 划分为若干个相抵等价类，每一个 $A \in \mathbb{F}^{m \times n}$ 属于其中唯一的一类，同一类的矩阵互相相抵，不同类的矩阵不相抵。从每一类中任选一个矩阵作为该类的代表元，则这些代表元构成了矩阵集 $\mathbb{F}^{m \times n}$ 的一个划分。以选取最简单的矩阵 $S_r = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ 为例，每个矩阵都可以通过初等变换化为某个 S_r ，因此选取这些 S_r 作为等价类的代表元，称为矩阵的相抵标准形。

定理 4.6

设 $A, B \in \mathbb{F}^{m \times n}$ ，则 A 与 B 相抵的充分必要条件是 $\text{rank}(A) = \text{rank}(B)$ 。

秩满足以下性质：

命题 4.7

1. $\text{rank}(A) + \text{rank}(B) = \text{rank} \begin{bmatrix} A & O \\ O & B \end{bmatrix}$.
2. $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$.
3. 若 P 为可逆矩阵, 则 $\text{rank}(PA) = \text{rank}(AP) = \text{rank}(A)$.
4. $\text{rank} \begin{bmatrix} A & O \\ O & B \end{bmatrix} \leq \text{rank} \begin{bmatrix} A & C \\ O & B \end{bmatrix}$
5. $\text{rank}(A+B) \leq \text{rank} \begin{bmatrix} A & B \end{bmatrix} \leq \text{rank}(A) + \text{rank}(B)$

4.2 秩不等式

定理 4.8 (Frobenius)

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

证明

$$\begin{aligned} \begin{bmatrix} ABC & O \\ O & B \end{bmatrix} &\rightarrow \begin{bmatrix} ABC & AB \\ O & B \end{bmatrix} \rightarrow \begin{bmatrix} O & AB \\ -BC & B \end{bmatrix} \\ \Rightarrow \text{rank}(AB) + \text{rank}(BC) &\leq \text{rank} \begin{bmatrix} O & AB \\ -BC & B \end{bmatrix} = \text{rank}(ABC) + \text{rank}(B) \end{aligned}$$

特别的, 取 $B = I_n$, 就可以得到:

推论 4.9 (Sylvester)

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

例题 4.1 (2014-2015 期中) 证明: 若 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$ 且 $AB = O$, 则 $\text{rank}(A) + \text{rank}(B) \leq n$.**证明** 由 Sylvester 不等式, $\text{rank}(A) + \text{rank}(B) \leq \text{rank}(AB) + n = n$ **例题 4.2 (2012-2013 期中)** 证明: 对于任意 n 阶方阵 A 均有 $\text{rank}(I_n - A) + \text{rank}(I_n + A) \geq n$, 等号成立当且仅当 $A^2 = I_n$ **证明** 首先有 $\text{rank}(I_n - A) + \text{rank}(I_n + A) \geq \text{rank}((I_n - A) + (I_n + A)) = n$. 考虑等号成立条件:

$$\begin{aligned} (\Leftrightarrow) & \text{由 Sylvester 不等式, } \text{rank}(I_n - A) + \text{rank}(I_n + A) \leq \text{rank}((I_n - A)(I_n + A)) + n = n. \\ (\Rightarrow) & \begin{bmatrix} I_n - A & O \\ O & I_n + A \end{bmatrix} \rightarrow \begin{bmatrix} I_n - A & I_n - A \\ O & I_n + A \end{bmatrix} \rightarrow \begin{bmatrix} I_n - A & 2I_n \\ & I_n + A \end{bmatrix} \rightarrow \begin{bmatrix} I_n - A & 2I_n \\ \frac{1}{2}(I_n - A^2) & O \end{bmatrix} \rightarrow \begin{bmatrix} O & 2I_n \\ \frac{1}{2}(I_n - A^2) & O \end{bmatrix} \\ & \text{此时由 } \text{rank} \begin{bmatrix} I_n - A & O \\ O & I_n + A \end{bmatrix} = n \text{ 即有 } I_n - A^2 = O, \text{ 即 } A^2 = I_n. \end{aligned}$$

例题 4.3 证明: 对于任意 n 阶方阵 A 均有 $\text{rank}(A) + \text{rank}(I_n - A) \geq n$, 等号成立当且仅当 $A^2 = A$ **证明** 首先有 $\text{rank}(A) + \text{rank}(I_n - A) \geq \text{rank}((I_n - A) + A) = n$. 考虑等号成立条件:

$$\begin{aligned} (\Leftrightarrow) & \text{由 Sylvester 不等式, } \text{rank}(A) + \text{rank}(I_n - A) \leq \text{rank}(A(I_n - A)) + n = n. \\ (\Rightarrow) & \begin{bmatrix} A & O \\ O & I_n - A \end{bmatrix} \rightarrow \begin{bmatrix} A & I_n - A \\ O & I_n - A \end{bmatrix} \rightarrow \begin{bmatrix} A & I_n \\ O & I_n - A \end{bmatrix} \rightarrow \begin{bmatrix} A & I_n \\ A - A^2 & O \end{bmatrix} \rightarrow \begin{bmatrix} O & I_n \\ A - A^2 & O \end{bmatrix} \end{aligned}$$

此时由 $\text{rank} \begin{bmatrix} O & I_n \\ A - A^2 & O \end{bmatrix} = n$ 即有 $A - A^2 = O$, 即 $A^2 = A$ 。

 **练习 4.1 (2013-2014 期中)** 证明: 对于任意 n 阶方阵 A 均有 $\text{rank}(A) + \text{rank}(2I_n - A) \geq n$, 等号成立当且仅当 $A^2 = 2A$

例题 4.4 (2019-2020 期中) 若 $A \in \mathbb{R}^{m \times n}$ 满足 $A^2 = O$, 证明: $\text{rank}(A) \leq \lfloor \frac{n}{2} \rfloor$. 并对每个 n 给出一个 n 阶方阵 A 使得 $A^2 = O$ 且 $\text{rank}(A) = \lfloor \frac{n}{2} \rfloor$.

证明 不等式是显然的, 只给出构造。记 $J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, 取 $A = \text{diag}(J, \dots, J)$ 或 $A = \text{diag}(J, \dots, J, 0)$ 即可。

例题 4.5 A 是 n 阶方阵, $\exists N \in \mathbb{N}^*$ 使得 $\text{rank}(A^N) = \text{rank}(A^{N+1})$, 证明: $\forall m > N, \text{rank}(A^m) = \text{rank}(A^N)$

证明 只证明 $\text{rank}(A^N) = \text{rank}(A^{N+1}) = \text{rank}(A^{N+2})$, 剩余递推即可。

一方面, 由 Frobenius 不等式, $\text{rank}(A^N A) + \text{rank}(A A^N) \leq \text{rank}(A A^N A) + \text{rank}(A^N) \Rightarrow \text{rank}(A^N) \leq \text{rank}(A^{N+2})$

另一方面, $\text{rank}(A^{N+2}) = \text{rank}(A^N A^2) \leq \text{rank}(A^N)$, 因此 $\text{rank}(A^N) = \text{rank}(A^{N+2})$ 。

4.3 初等变换与秩

例题 4.6 (2011-2012 期中) 若 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}$, 则 $I_n - BA, I_m - AB, \begin{bmatrix} I_m & A \\ B & I_n \end{bmatrix}$ 的秩有什么关系?

解 $\begin{bmatrix} I_m - AB & O \\ O & I_n \end{bmatrix} \leftarrow \begin{bmatrix} I_m - AB & O \\ B & I_n \end{bmatrix} \leftarrow \begin{bmatrix} I_m & A \\ B & I_n \end{bmatrix} \rightarrow \begin{bmatrix} I_m & O \\ B & I_n - BA \end{bmatrix} \rightarrow \begin{bmatrix} I_m & O \\ O & I_n - BA \end{bmatrix}$

故 $n + \text{rank}(I_m - AB) = m + \text{rank}(I_n - BA) = \text{rank} \begin{bmatrix} I_m & A \\ B & I_n \end{bmatrix}$

例题 4.7 求下列矩阵的秩:

$$(1) A = \begin{bmatrix} 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_m \\ b_1 & b_2 & \cdots & b_m & 1 \end{bmatrix} \quad (2) B = \begin{bmatrix} 1 - a_1 b_1 & -a_1 b_2 & \cdots & -a_1 b_n \\ -a_2 b_1 & 1 - a_2 b_2 & \cdots & -a_2 b_n \\ \vdots & \vdots & & \vdots \\ -a_n b_1 & -a_n b_2 & \cdots & 1 - a_n b_n \end{bmatrix}$$

$$(3) C = \begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ b & b & \cdots & a \end{bmatrix}_{n \times n}$$

解 (1) 记 $A = \begin{bmatrix} I_m & \alpha \\ \beta^T & 1 \end{bmatrix}$, 则 $\text{rank}(A) = m + \text{rank}(1 - \beta^T \alpha)$ 。若 $\beta^T \alpha = 1$, 则 $\text{rank}(A) = m + 1$, 否则 $\text{rank}(A) = m$ 。

$$(2) B = I_n - \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = I_n - \alpha \beta^T, \text{ 此时 } 1 + \text{rank}(B) = n + \text{rank}(1 - \beta^T \alpha)。$$

$$(3) C = (a-b)I_n + \begin{bmatrix} b \\ b \\ \vdots \\ b \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}. \text{ 若 } a=b=0, \text{ 则 } \text{rank}(C)=0; \text{ 若 } a=b \neq 0, \text{ 则 } \text{rank}(C)=1.$$

$$\text{对 } a \neq b, C = (a-b) \left(I_n - \begin{bmatrix} -\frac{b}{a-b} \\ -\frac{b}{a-b} \\ \vdots \\ -\frac{b}{a-b} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \right), \text{ 若 } b = -\frac{a}{n-1}, \text{ 则 } \text{rank}(C) = n-1, \text{ 否则 } \text{rank}(C) = n.$$

例题 4.8 (2014-2015 期中)(满秩分解) 证明: 若 $A \in \mathbb{F}^{m \times n}$, 则 $\text{rank}(A) = r$ 等价于存在列满秩矩阵 $B \in \mathbb{F}^{m \times r}$ 与行满秩矩阵 $C \in \mathbb{F}^{r \times n}$ 使得 $A = BC$ 。

证明 存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$ 和 $Q \in \mathbb{F}^{n \times n}$ 使得 $A = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q = P \begin{bmatrix} I_r \\ O \end{bmatrix} \begin{bmatrix} I_r & O \end{bmatrix} Q$ 。

取 $B = P \begin{bmatrix} I_r \\ O \end{bmatrix}$, $C = \begin{bmatrix} I_r & O \end{bmatrix} Q$ 即可。

例题 4.9 (2016-2017 期中) A 是 n 阶方阵, $\text{rank}(A) = 1, c = \text{tr}(A)$, 证明: $A^2 = cA$, 并求 $\det(I + A)$

解 对 A 作满秩分解 $A = \beta \alpha^T$, 其中 α, β 是列向量, 则 $A^2 = \beta \alpha^T \beta \alpha^T = \beta (\alpha^T \beta) \alpha^T = cA$ 。

其中 $c = \text{tr}(A) = \text{tr}(\beta \alpha^T) = \text{tr}(\alpha^T \beta) = \alpha^T \beta$. $\det(I + A) = \det(I + \beta \alpha^T) = \det(1 + \alpha^T \beta) = 1 + c$ 。

例题 4.10 (2015-2016 期中) 判断: 若 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}$, 则 $\text{rank}(AB) = \text{rank}(BA)$

证明 错误。 $A = \begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix}, B = \begin{bmatrix} 2 & 4 \\ -3 & -6 \end{bmatrix}$ 满足 $AB \neq O, BA = O$ 。

例题 4.11 设 A, B 为 n 阶方阵, $A^2 = A, B^2 = B, I - A - B$ 可逆, 证明 $\text{rank}(A) = \text{rank}(B)$

证明 由 Sylvester 不等式, $\text{rank}(A) = \text{rank}(I - A - B) - n + \text{rank}(A) \leq \text{rank}(A - A^2 - BA) = \text{rank}(BA)$

而 $\text{rank}(BA) \leq \text{rank}(B)$, 得到 $\text{rank}(A) \leq \text{rank}(B)$ 。同理 $\text{rank}(B) \leq \text{rank}(A)$, 故 $\text{rank}(A) = \text{rank}(B)$ 。

如果注意力足够好, 则有更简单的做法:

证明 注意到 $(I - A - B)B = -AB = A(I - A - B)$, 由左乘或右乘可逆矩阵不改变秩, 有 $\text{rank}(B) = \text{rank}(A)$

例题 4.12 设 $A \in \mathbb{R}^{n \times n}$, 证明: $A^2 = A$ 时 $\text{rank}(A) = \text{tr}(A)$ 。

证明 记 $r = \text{rank}(A)$, 则 $\exists P, Q \in GL_n(\mathbb{R}), s.t. A = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q$, 此时 $A^2 = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q = A$ 。

消去可逆矩阵, $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ 。对 QP 作分块 $\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$, 代入前式, 得到

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} = \begin{bmatrix} R_{11} & O \\ O & O \end{bmatrix} = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \Rightarrow R_{11} = I_r, \text{ 此时}$$

$$\text{tr}(A) = \text{tr} \left(P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q \right) = \text{tr} \left(\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q P \right) = \text{tr} \left(\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \begin{bmatrix} I_r & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right) = \text{tr} \begin{bmatrix} I_r & R_{12} \\ O & O \end{bmatrix} = r = \text{rank}(A)。$$

4.4 补充：利用矩阵结构的求逆技巧

定理 4.10 (Schur 补)

设 $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ ，其中 A 和 D 均为方阵，且 A 可逆，定义 Schur 补为 $S = D - CA^{-1}B$ ，则当 S 可逆时， M 也可逆，且其逆矩阵为

$$M^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{bmatrix}$$

对 D 可逆的情况亦然。



定理 4.11 (Sherman-Morrison 秩 1 修正)

设 $A \in \mathbb{R}^{n \times n}$ 可逆， $\alpha, \beta \in \mathbb{R}^{n \times 1}$ ，则 $A + \alpha\beta^T$ 可逆的充分必要条件是 $1 + \beta^T A^{-1}\alpha \neq 0$ ，且此时

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



定理 4.12 (Woodbury 秩 k 修正)

设 $A \in \mathbb{R}^{n \times n}$ 可逆， $U, V \in \mathbb{R}^{n \times k}$ ，则 $A + UV^T$ 可逆的充分必要条件是 $I_k + V^T A^{-1}U$ 可逆，且此时

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I_k + V^T A^{-1}U)^{-1}V^T A^{-1}$$

进一步地，若 $C \in \mathbb{R}^{k \times k}$ 可逆，则 $A + UCV^T$ 可逆的充分必要条件是 $C^{-1} + V^T A^{-1}U$ 可逆，且此时

$$(A + UCV^T)^{-1} = A^{-1} - A^{-1}U(C^{-1} + V^T A^{-1}U)^{-1}V^T A^{-1}$$



第五章 线性空间的基与线性映射

5.0 选择公理 *

选择公理是现代数学的一个基本公理，尽管其陈述看起来非常直观，但它却能引出一些非常强大且惊人的结论。

公理 5.1 (选择公理)

对于任意一族非空集合，总能构造出一个新的集合，其元素是从原先集合族的每一个集合中恰好各取一个元素组成的。也可表述如下：

设 I 是非空集合族 $C := \{A_i \mid i \in I\}$ 的指标集，且 $\forall i \in I, A_i \neq \emptyset$ ，则存在映射 $f: I \rightarrow \bigcup_{i \in I} A_i$ ，使得 $\forall i \in I, f(i) \in A_i$ 。



如果集合族 C 是有限的，那么选择公理是显然成立的，因为可以逐个选择每个集合中的元素，这个过程可以在有限步内完成。如果集合族 C 是无限的，但是但每个集合内部的元素有明确的规则可以依循，此时通常也不需要选择公理。例如，假设 C 是所有正整数集的非空子集构成的集族，可以定义一个选择函数 $f(A) = \min A$ 。这个规则是明确的，不需要任意选择。但是选择公理的威力体现在当集合族 C 是无限的，并且其中集合的元素没有可辨识的特征或明确的排序规则时。例如，想象一个无限集族，其中每个集合都包含一对无法区分的袜子。选择公理断言，即使没有任何标准（比如大小、颜色），也能够为每一双袜子都指定一只“左脚”的袜子。这个“指定”的过程是非构造性的，公理只保证了选择的可能性，但没有告知如何去选择。这种非构造性是选择公理争议的来源。它断言了某个对象（选择函数或由所选元素构成的集合）的存在，但没有给出具体构造它的方法。

以下是一些选择公理的等价表述：

引理 5.2 (Zorn 引理)

设 $(P, \leq)$ 是偏序集，且每个链（全序子集）都有上界，则 P 中存在极大元。



引理 5.3 (良序原理)

每个集合都可以良序化，即存在一个全序关系，使得每个非空子集都有最小元。



选择公理是现代数学的基石之一，许多习以为常的定理都依赖于它。以下是一些依赖选择公理的著名定理：

- 每个向量空间都存在一个 Hamel 基。
- Hilbert 空间中每个有界序列都有一个弱收敛子列。
- 任意多个紧拓扑空间的积空间仍然是紧的。
- 在赋范线性空间中，任何定义在子空间上的有界线性泛函都可以扩展到整个空间，且保持其范数不变。

但是，当承认选择公理时，也会引入一些反直觉的结果。

命题 5.4 (Banach-Tarski 悖论)

在三维欧氏空间中，存在一种将一个单位球体分割成有限个不重叠的部分，并通过平移和旋转将这些部分重新组合成两个与原球体大小相同的球体的方法。



关于 Banach-Tarski 悖论，可以观看 [BV1Ps411m7a8](#) 以获得更深入的理解。

5.1 线性空间的基 *

5.1.1 Hamel 基

课程中同学们学习的基的定义只在有限维线性空间上，而一般的不确定维度的线性空间上，一种较为简单的和课本涉及的有限维的基定义相容的定义被称为 Hamel 基。

定义 5.5 (Hamel 基)

设 V 是域 $\mathbb{F}$ 上的线性空间，称 V 中的一组向量 $\{\alpha_i \mid i \in S\}$ 是 V 的 Hamel 基，若 $\forall x \in V$ ，存在唯一的有限子集 $T \subseteq S$ 和 $\{c_i \in \mathbb{F} \mid i \in T\}$ ，使得 $x = \sum_{i \in T} c_i \alpha_i$ 。

需要注意的是，Hamel 基的定义中要求线性组合是有限的，这与某些空间的自然基不同。例如，考虑以下线性空间 $\mathbb{R}^{\mathbb{N}}$ ，即形如 $(x_1, x_2, \dots)$ 的所有实数序列构成的线性空间。该空间的自然基为 $\{e_i \mid i \in \mathbb{N}\}$ ，其中 e_i 是第 i 个位置为 1，其余位置为 0 的序列。然而，这个自然基并不是 Hamel 基，因为某些序列（例如 $a = (1, 1, 1, \dots)$ ）无法表示为有限个基向量的线性组合。事实上，在 Hamel 基的定义下， $a, e_1, e_2, \dots$ 是线性无关的，即便有看似成立的等式 $x = \sum_{i=1}^{\infty} e_i$ 。

事实上，一个空间是否存在 Hamel 基是一个深刻的问题。Zorn 引理保证了每个非零线性空间都存在 Hamel 基，但构造这样的基通常需要使用选择公理，且在实际应用中并不常见。

5.1.2 Schauder 基

在刚才的例子中，我们看到 Hamel 基的定义限制了线性组合的形式，导致某些自然基无法满足该定义。为了解决这个问题，我们引入了 Schauder 基的概念。但是，Schauder 基的定义需要在线性空间上引入拓扑结构，因此我们需要先了解一些基本的拓扑概念。

定义 5.6 (度量空间)

设 X 是一个非空集合， $d: X \times X \rightarrow \mathbb{R}$ 是一个函数，若对任意 $x, y, z \in X$ ，满足以下条件：

1. 非负性： $d(x, y) \geq 0$ ，且当且仅当 $x = y$ 时， $d(x, y) = 0$ ；
2. 对称性： $d(x, y) = d(y, x)$ ；
3. 三角不等式： $d(x, z) \leq d(x, y) + d(y, z)$ ；

则称 d 为 X 上的度量（距离）， (X, d) 称为度量空间。

注 距离概念是欧氏空间中两点间距离的抽象。例如，在 $\mathbb{R}^n$ 中 $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ 满足以上三个条件，即欧氏距离。

引入距离的目的是刻画收敛。

定义 5.7 (收敛与完备性)

设 (X, d) 是度量空间， $\{x_n\}$ 是 X 中的一个序列， $\{x_n\}$ 收敛到 x_0 是指 $\lim_{n \rightarrow \infty} d(x_n, x_0) = 0$ ，此时记作 $\lim_{n \rightarrow \infty} x_n = x_0$ 或 $x_n \rightarrow x_0$ 。

对于任意序列 $\{x_n\}$ ，若对任意 $\varepsilon > 0$ ，存在 $N \in \mathbb{N}$ ，使得当 $m, n > N$ 时，有 $d(x_n, x_m) < \varepsilon$ ，则称 $\{x_n\}$ 是 Cauchy 列或基本列。如果 (X, d) 中的任意 Cauchy 列均收敛于 X 中的某个元素，则称 (X, d) 是完备的。

回到线性空间，现在需要把在线性空间中建立度量，于是引出了范数与赋范线性空间的概念。

定义 5.8 (范数与赋范线性空间)

设 V 是域 $\mathbb{F}$ 上的线性空间, $\|\cdot\|: V \rightarrow \mathbb{R}$ 是一个函数, 若对任意 $x, y \in V$ 和 $\alpha \in \mathbb{F}$, 满足以下条件:

1. 非负性: $\|x\| \geq 0$, 且当且仅当 $x = 0$ 时, $\|x\| = 0$;
2. 齐次性: $\|\alpha x\| = |\alpha| \|x\|$;
3. 三角不等式: $\|x + y\| \leq \|x\| + \|y\|$;

则称 $\|\cdot\|$ 为 V 上的范数, $(V, \|\cdot\|)$ 称为赋范线性空间。



注 每个范数都可以诱导出一个度量, 定义为 $d(x, y) = \|x - y\|$ 。因此, 赋范线性空间 $(V, \|\cdot\|)$ 也是一个度量空间 (V, d) 。但是反之不然。例如 $\mathbb{R}$ 上的离散度量 $d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$ 不满足齐次性, 故无法由任何范数诱导得到。

度量空间的完备性在赋范线性空间中也有对应的概念, 即 Banach 空间。

定义 5.9 (Banach 空间)

设 $(V, \|\cdot\|)$ 是赋范线性空间, 若 (V, d) 是完备的, 则称 $(V, \|\cdot\|)$ 为 Banach 空间。



在了解了赋范线性空间与 Banach 空间后, 我们可以定义 Schauder 基。它将 Hamel 基的概念推广到了允许无限线性组合的情况。

定义 5.10 (Schauder 基)

设 $(V, \|\cdot\|)$ 是域 $\mathbb{F}$ 上的 Banach 空间。若 V 中一个可数序列 $\{e_n\}$ 满足 $\forall x \in V$, 存在唯一的数列 $\{a_n\} \subseteq \mathbb{F}$, 使得 $x = \sum_{n=1}^{\infty} a_n e_n \Leftrightarrow \lim_{N \rightarrow \infty} \left\| x - \sum_{n=1}^N a_n e_n \right\| = 0$, 则称 $\{e_n\}$ 为 V 的 Schauder 基。



例如, 考虑平方可和的数列空间 $l^2 = \left\{ (x_1, x_2, \dots) : x_i \in \mathbb{R}, \sum_{i=1}^{\infty} x_i^2 < \infty \right\}$, 配备范数 $\|(x_1, x_2, \dots)\|_{l^2} = \sqrt{\sum_{i=1}^{\infty} x_i^2}$ 。该空间是一个 Banach 空间, 其 Schauder 基为 $\{e_n \mid n \in \mathbb{N}\}$, 其中 e_n 是第 n 个位置为 1, 其余位置为 0 的序列。对于任意 $x = (x_1, x_2, \dots) \in l^2$, 都有 $x = \sum_{n=1}^{\infty} x_n e_n$, 且该表示唯一。此时, l^2 的 Schauder 基与其自然基相同。

在 Banach 空间上, Hamel 基组成集合的势也有特殊性质: 其若非有限则不可数。这与 Schauder 基的可数性(通常)形成对比, 表明在无限维空间中, Hamel 基通常比 Schauder 基“更大”。关于可数与不可数区别与性质, 可以观看 [BV12N411o7MU](#) 以获得更深入的理解。

5.2 相关练习

例题 5.1 判断下述命题是否正确, 并说明理由。

1. 若 $a_1, a_2, \dots, a_s (s \geq 2)$ 线性相关, 则其中每一个向量都可以表示成其它向量的线性组合。
2. 如果向量组的任何一个非它本身的子向量组都线性无关, 则该向量组也线性无关。
3. 若向量组线性无关, 则它的任何子向量组都线性无关。
4. F^n 的 $n+1$ 个向量组成的向量组必线性相关。
5. 设 $a_1, a_2, \dots, a_s$ 线性无关, 则 $a_1 + a_2, a_2 + a_3, \dots, a_s + a_1$ 必线性无关。
6. 设 $a_1, a_2, \dots, a_s$ 线性相关, 则 $a_1 + a_2, a_2 + a_3, \dots, a_s + a_1$ 必线性相关。
7. 设 $a_1, a_2, \dots, a_s \in F^n$ 线性无关, 则它们的加长向量组也必线性无关。
8. 设 $a_1, a_2, \dots, a_s \in F^n$ 线性相关, 则它们的加长向量组也必线性相关。

解

1. 若 $a_1, a_2, \dots, a_s (s \geq 2)$ 线性相关, 则其中每一个向量都可以表示成其它向量的线性组合。

否。考虑作为一维数组向量的 $0, 1, 1$ 不能被 0 表示为线性组合。

2. 如果向量组的任何一个非它本身的子向量组都线性无关, 则该向量组也线性无关。

否。考虑二维数组向量组 $(1, 0), (0, 1), (1, 1)$ 。

3. 若向量组线性无关, 则它的任何子向量组都线性无关。

是。不妨设原向量组集合为 V , 子向量组为 $V' \subset V$, 则

V 线性无关

$$\begin{aligned} &\Leftrightarrow \sum_{\alpha \in V} \lambda_{\alpha} \alpha = 0 \text{ 只有 } \lambda_{\alpha} = 0, \forall \alpha \in V \text{ 唯一解} \\ &\Rightarrow \sum_{\alpha \in V'} \lambda_{\alpha} \alpha = \sum_{\alpha \in V'} \lambda_{\alpha} \alpha + \sum_{\alpha \in V-V'} 0 \cdot \alpha = 0 \text{ 只有 } \lambda_{\alpha} = 0, \forall \alpha \in V' \text{ 唯一解} \\ &\Leftrightarrow V' \text{ 线性无关} \end{aligned}$$

4. F^n 的 $n+1$ 个向量组成的向量组必线性相关。

是。它们可作为列向量排成一个规格为 $n \times (n+1)$ 的矩阵, 这个矩阵作为系数矩阵对应的齐次线性方程组的有非零解, 该非零解表明它们可线性组合为零向量, 因而线性相关。

5. 设 $a_1, a_2, \dots, a_s$ 线性无关, 则 $a_1 + a_2, a_2 + a_3, \dots, a_s + a_1$ 必线性无关。

否。记矩阵

$$P = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \end{bmatrix} \Rightarrow P \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ a_2 + a_3 \\ \vdots \\ a_s + a_1 \end{bmatrix}$$

根据 s 为偶数时 $\det P = 0$, P 不可逆知该向量组构成的矩阵一定无 s 阶可逆子矩阵 (假如线性无关那么秩为 s , 列秩也为 s , 从而一定可选出 s 个列得到一个可逆的 s 阶子矩阵), 一定线性相关。

6. 设 $a_1, a_2, \dots, a_s$ 线性相关, 则 $a_1 + a_2, a_2 + a_3, \dots, a_s + a_1$ 必线性相关。

是。 s 为偶数时理由同 5, s 为奇数时 $\det(P) \neq 0$, 该二向量组等价, 因此秩都严格小于 s 。

7. 设 $a_1, a_2, \dots, a_s \in F^n$ 线性无关, 则它们的加长向量组也必线性无关。

是。利用定义验证时仅考虑它们加长前的分量便可得全部的线性组合系数为 0。

8. 设 $a_1, a_2, \dots, a_s \in F^n$ 线性相关, 则它们的加长向量组也必线性相关。

否。每一个线性无关的数组向量组都可添上零保持线性无关性, 而它们现在都可看成是由线性相关的零向量加长得到。

例题 5.2 设 $f_1(x) = e^{\lambda_1 x}, f_2(x) = e^{\lambda_2 x}, \dots, f_n(x) = e^{\lambda_n x}, \lambda_1, \dots, \lambda_n$ 两两不同。证明: $f_1(x), f_2(x), \dots, f_n(x)$ 线性无关。

证明 若 $f_1(x), f_2(x), \dots, f_n(x)$ 线性无关, 则存在 $a_1, \dots, a_n$ 不全为 0 使得 $g(x) = \sum a_i e^{\lambda_i x} = 0$ 。此时 $g^{(k)}(x) = \sum a_i \lambda_i^k e^{\lambda_i x} = 0$ 。取 $x = 0, k = 0, 1, \dots, n-1$ 并把方程组写成矩阵形式, 即有

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

系数矩阵行列式为 $\prod_{1 \leq i < j \leq n} (\lambda_j - \lambda_i) \neq 0$, 此时方程组只有零解, 故 $f_1(x), f_2(x), \dots, f_n(x)$ 线性无关。

例题 5.3 对 $A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n$ 是列向量, $\exists k \in \mathbb{N}$ 使得 $A^k x = 0, A^{k-1} x \neq 0$, 证明: $x, Ax, A^2 x, \dots, A^{k-1} x$ 线性无关。

证明 若 $x, Ax, A^2x, \dots, A^{k-1}x$ 线性相关, 则 $\exists m \in \mathbb{N}$ 使得 $A^m x, A^{m+1}x, \dots, A^{k-1}x$ 线性相关, $A^{m+1}x, A^{m+2}x, \dots, A^{k-1}x$ 线性无关。此时 $\exists \lambda_m, \lambda_{m+1}, \dots, \lambda_{k-1}$ 不全为 0, 使得 $\sum_{j=m}^{k-1} \lambda_j A^j x = 0$ 。

此时 $0 = A \sum_{j=m}^{k-1} \lambda_j A^j x = \sum_{j=m}^{k-1} \lambda_j A^{j+1} x = \sum_{j=m+1}^{k-1} \lambda_j A^j x$ 。由 $A^{m+1}x, A^{m+2}x, \dots, A^{k-1}x$ 线性无关,

有 $\lambda_{m+1} = \dots = \lambda_{k-1} = 0$, 代入回前式, 有 $\lambda_m = 0$, 与 $\lambda_m, \lambda_{m+1}, \dots, \lambda_{k-1}$ 不全为 0 矛盾。

故 $x, Ax, A^2x, \dots, A^{k-1}x$ 线性无关。

例题 5.4 给定 $A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, A_3 = \begin{bmatrix} -6 & 4 \\ 1 & 0 \end{bmatrix}, A_4 = \begin{bmatrix} 4 & -2 \\ 0 & 1 \end{bmatrix}, B_1 = E_{11}, B_2 = E_{12}, B_3 = E_{21}, B_4 = E_{22}$ 。

1. 证明: A_1, A_2, A_3, A_4 构成 $\mathbb{R}^{2 \times 2}$ 的一组基。

2. 求向量组 A_1, A_2, A_3, A_4 到 B_1, B_2, B_3, B_4 的过渡矩阵 T 。

3. 求 $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ 在基 A_1, A_2, A_3, A_4 下的坐标。

解

1. 自行验证即可。

2. 把矩阵 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 视为 $aE_{11} + bE_{12} + cE_{21} + dE_{22}$, 进而等价于坐标 $(a, b, c, d)^\top$, 则有 $I_4 = \begin{bmatrix} 1 & 0 & -6 & 4 \\ 0 & 1 & 4 & -2 \\ 1 & 2 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{bmatrix} T$,

$$\text{故 } T = \begin{bmatrix} -1 & 0 & 6 & 4 \\ 0 & -1 & 4 & -2 \\ 1 & 2 & -1 & 0 \\ 2 & 3 & 0 & -1 \end{bmatrix}.$$

3. 形式上有 $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 & B_3 & B_4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \end{bmatrix} T \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, 故坐标是 $T \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ 2 \\ 4 \end{bmatrix}$ 。

例题 5.5 设 $P \in \mathbb{R}^{n \times n}$ 是可逆方阵, $A = P^{-1} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)P$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 两两不同, 设 $V = \{B \in \mathbb{R}^{n \times n} \mid AB = BA\}$ 。

1. 证明: V 是 $\mathbb{R}^{n \times n}$ 的子空间。

2. 求 V 的维数与一组基。

解

1. 显然 $0 \in V$, 且 $\forall B_1, B_2 \in V, \forall \lambda, \mu \in \mathbb{R}, (\lambda B_1 + \mu B_2)A = \lambda B_1 A + \mu B_2 A = \lambda A B_1 + \mu A B_2 = A(\lambda B_1 + \mu B_2)$, 故 $\lambda B_1 + \mu B_2 \in V$, 因此 V 是 $\mathbb{R}^{n \times n}$ 的子空间。

2. $AB = BA \Leftrightarrow PAP^{-1}PBP^{-1} = PBP^{-1}PAP^{-1} \Leftrightarrow \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)PBP^{-1} = PBP^{-1}\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 。设 $PBP^{-1} = C = (c_{ij})_{n \times n}$, 则 $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)C = C\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \Rightarrow \forall i \neq j, (\lambda_i - \lambda_j)c_{ij} = 0$, 故 C 为对角矩阵, C 组成空间的基为 $\{E_{11}, E_{22}, \dots, E_{nn}\}$, 维数为 n 。因此 V 的基为 $\{P^{-1}E_{11}P, P^{-1}E_{22}P, \dots, P^{-1}E_{nn}P\}$, 维数为 n 。

5.3 线性映射

5.3.1 线性映射与线性变换的定义

定义 5.11 (线性映射与线性变换)

设 V, V' 为 $\mathbb{F}$ 上的两个线性空间, 若映射 $f: V \rightarrow V'$ 满足

1. $\forall a, b \in V$, 有 $f(a+b) = f(a) + f(b)$
2. $\forall a \in V, \lambda \in \mathbb{F}$, 有 $f(\lambda a) = \lambda f(a)$

则称 f 为 V 到 V' 的线性映射。若 $V' = V$, 则称 f 为线性变换。

容易验证 f 为 V 到 V' 的线性变换的一个等价定义是 $\forall a, b \in V, \forall \lambda, \mu \in \mathbb{F}$, $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$ 。

需要小心的是, 线性变换并不像初等行变换, 平移旋转等几何变换是可逆的, 虽然在一些语境下变换确实有可逆的意味。例如一个把全部元素都映为零的映射当然也符合线性变换的定义。

5.3.2 线性映射的矩阵描述

根据线性映射的线性性, 对于 $f: V \rightarrow V'$, 设 V 的基为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 对任意向量 $a \in V$, 它有唯一的线性组合表示 $a = \sum_{i=1}^n x_i \alpha_i$, 则 $f(a) = f(\sum_{i=1}^n x_i \alpha_i) = \sum_{i=1}^n x_i f(\alpha_i)$ 。因此只要知道一组基的像, 就可以计算出每个元素的像了, 因此这个线性映射也被完全地确定了下来。假设 $\beta_1, \beta_2, \dots, \beta_m$ 为 V' 的基, 并设 $f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n)$ 在该组基下的线性表示为

$$\begin{cases} f(\alpha_1) = a_{11}\beta_1 + a_{21}\beta_2 + \dots + a_{m1}\beta_m \\ f(\alpha_2) = a_{12}\beta_1 + a_{22}\beta_2 + \dots + a_{m2}\beta_m \\ \vdots \\ f(\alpha_n) = a_{1n}\beta_1 + a_{2n}\beta_2 + \dots + a_{mn}\beta_m \end{cases}$$

$$\text{写成矩阵形式即 } \begin{bmatrix} f(\alpha_1) & f(\alpha_2) & \dots & f(\alpha_n) \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}。$$

上面这个式子即是确定一个线性映射的三要素, 即原空间的一组基, 像所在空间的一组基, 和一个矩阵。这个矩阵的每一列是原空间的基的像在像所在空间的基下的坐标。把上面这个矩阵 A 称为 f 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\beta_1, \beta_2, \dots, \beta_m$ 下的矩阵, 则这个式子可以简单记为 $f(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_m)A$ 。现在可以用上面这两

组基和一个矩阵描述每个元素的像: $f(a) = \begin{bmatrix} f(\alpha_1) & f(\alpha_2) & \dots & f(\alpha_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_m \end{bmatrix} A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ 。这

表明, 线性映射把在 V 中以 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为基下坐标为 $(x_1, x_2, \dots, x_n)^T$ 的向量映射为在 V' 中以 $\beta_1, \beta_2, \dots, \beta_m$ 为基下坐标为 $A(x_1, x_2, \dots, x_n)^T$ 的向量。反过来看, 对于确定的两组基, 任意一个恰当规格的矩阵, 它都可以确定一个线性映射。因此可以将线性映射 f 与映射 $\mathcal{A}: \mathbb{F}^{n \times 1} \rightarrow \mathbb{F}^{m \times 1}, x \mapsto Ax$ 近乎等同看待, 并把这个映射称为 f 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\beta_1, \beta_2, \dots, \beta_m$ 下的坐标表示。对于线性变换, 就是上述讨论在两个空间和两组基完全相同的情形。

进一步讨论一个问题, 假如 V 有另一组基 $\alpha'_1, \alpha'_2, \dots, \alpha'_n$, V' 有另一组基 $\beta'_1, \beta'_2, \dots, \beta'_m$, 它们和原来基的过

渡矩阵 T, T' 满足 $(\alpha'_1, \alpha'_2, \dots, \alpha'_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)T$, $(\beta'_1, \beta'_2, \dots, \beta'_m) = (\beta_1, \beta_2, \dots, \beta_m)T'$, 由 f 的线性性可得出 $(f(\alpha'_1), f(\alpha'_2), \dots, f(\alpha'_n)) = (f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n))T$, 从而有 $f(\alpha'_1, \alpha'_2, \dots, \alpha'_n) = (\beta'_1, \beta'_2, \dots, \beta'_m)(T')^{-1}AT$ 。这说明不同基下的矩阵它们是彼此相抵的, 总可以找到恰当的可逆阵 T, T' 来得到两组基, 使得 f 在这两组基下的矩阵为 $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, $r = \text{rank } A$ 。在该矩阵下, f 的坐标表示是简单的取前 r 分量的投影。因此在坐标表示的意义下, 若原空间和像所在空间不同, 线性映射总能看成是投影。而线性变换即是上述 $T = T'$ 且两组基相同的情形, 它的矩阵为 $T^{-1}AT$, 则未必能得到类似投影的矩阵了。这也是后续矩阵相似部分的内容。

5.3.3 线性映射的性质 *

所有的性质都是由映射的性质和线性性导出的。

假设有线性空间 V_1, V_2, V_3 , 它们的基向量组分别为 $M_1 = (\alpha_1, \dots, \alpha_n), M_2 = (\beta_1, \dots, \beta_m), M_3 = (\gamma_1, \dots, \gamma_l)$, 对于线性映射 $f: V_1 \rightarrow V_2, g: V_2 \rightarrow V_3$, 假设它们在前面基下的矩阵为 A, B , 则根据定义, 对于 $g \circ f: V_1 \rightarrow V_3$, 它在 M_1 与 M_3 下的矩阵满足 $(g \circ f(\alpha_1), g \circ f(\alpha_2), \dots, g \circ f(\alpha_n)) = (\gamma_1, \gamma_2, \dots, \gamma_l)C$, 而根据 g 的线性性, 对于 $f(\alpha_i) = \sum_{j=1}^m a_{ji}\beta_j$, 有 $g \circ f(\alpha_i) = g\left(\sum_{j=1}^m a_{ji}\beta_j\right) = \sum_{j=1}^m a_{ji}g(\beta_j) = \sum_{j=1}^m a_{ji} \sum_{k=1}^l b_{kj}\gamma_k$, 写成矩阵形式, 即 $(g \circ f(\alpha_1), \dots, g \circ f(\alpha_n)) = (g(\beta_1), \dots, g(\beta_m))A = (\gamma_1, \dots, \gamma_l)BA$ 。这表明 $C = BA$, 因此在线性空间的基都给定的情形下, 对于可复合的线性映射, 它们的复合在这些基下的矩阵就是按照复合的次序进行矩阵乘法得到的矩阵。

对于同一个空间 V 上的线性变换的复合, 由于原空间和像所在空间完全相同, 所选取的基也完全相同, 是允许一个线性变换和它自己复合的, 复合得到的矩阵依旧为各个参与复合的映射按照复合的次序相乘, 可记作 $\mathcal{A}^P(\alpha_1, \alpha_2, \dots, \alpha_n) = (f^P(\alpha_1), f^P(\alpha_2), \dots, f^P(\alpha_n)) = (\alpha_1, \alpha_2, \dots, \alpha_n)A^P$ 。

以下是线性映射的像与核的定义:

定义 5.12 (像与核)

设 $f: V \rightarrow V'$ 是线性映射, 则称集合 $\text{Im } f = \{f(a) \mid a \in V\}$ 为 f 的像, 称集合 $\ker f = \{a \in V \mid f(a) = 0\}$ 为 f 的核。

以下是线性映射的像与核的性质:

定理 5.13

设 $f: V \rightarrow V'$ 是线性映射, 则:

- $\text{Im } f$ 是 V' 的子空间;
- 设 U' 是 V' 的子空间, 则 $f^{-1}(U') = \{a \in V \mid f(a) \in U'\}$ 是 V 的子空间; 特别的, $\ker f = f^{-1}(\{0\})$ 是 V 的子空间。

证明是平凡的, 依照线性映射的定义, 它们都对加法与数乘封闭, 当然构成原向量空间的子空间。关于零空间, 两个常用的定理如下:

定理 5.14

设 $f: V \rightarrow V'$ 是线性映射, 则 f 为单射的充分必要条件是 $\ker f = \{0\}$ 。

证明 $(\Rightarrow)$ 设 $a \in \ker f$, 则 $f(a) = 0 = f(0)$, 由 f 的单射性可得 $a = 0$, 故 $\ker f = \{0\}$ 。

$(\Leftarrow)$ 设 $a_1, a_2 \in V$ 且 $f(a_1) = f(a_2)$, 则 $f(a_1 - a_2) = f(a_1) - f(a_2) = 0$, 由 $\ker f = \{0\}$ 可得 $a_1 - a_2 = 0$, 即 $a_1 = a_2$, 故 f 为单射。

定理 5.15

设 $f: V \rightarrow V'$ 是线性映射, 且 V 为有限维线性空间, 则 $\dim V = \dim \ker f + \dim \operatorname{Im} f$ 。

特别的, 若 f 为单射, 则 $\dim V = \dim \operatorname{Im} f$; 若 f 为满射, 则 $\dim V' = \dim \operatorname{Im} f = \dim V - \dim \ker f$, 若 f 为双射, 则 $\dim V = \dim V'$ 。



证明 设 $(\alpha_1, \dots, \alpha_n)$ 为 V 基, 由 $(f(\alpha_1), \dots, f(\alpha_n))$ 可线性组合成 $f(V)$, 有 $\dim f(V) = \operatorname{rank}(f(\alpha_1), \dots, f(\alpha_n)) := l$ 。设 $(f(\alpha_1), \dots, f(\alpha_n))$ 的极大线性无关组为 $(f(\alpha_{i_1}), \dots, f(\alpha_{i_l}))$, 其余的向量为 $(f(\alpha_{i_{l+1}}), \dots, f(\alpha_{i_n}))$, 则根据 $(f(\alpha_{i_1}), \dots, f(\alpha_{i_l}))$ 为 $f(V)$ 的一组基, $(f(\alpha_{i_{l+1}}), \dots, f(\alpha_{i_n}))$ 可被其线性表示, 设为 $f(\alpha_{i_{l+k}}) = \sum_{j=1}^l c_{jk} f(\alpha_{i_j}) \Rightarrow f\left(\alpha_{i_{l+k}} - \sum_{j=1}^l c_{jk} \alpha_{i_j}\right) = 0$ 。而可知 $(\alpha_{i_1}, \dots, \alpha_{i_l}, \alpha_{i_{l+1}} - \sum_{j=1}^l c_{j1} \alpha_{i_j}, \alpha_{i_{l+2}} - \sum_{j=1}^l c_{j2} \alpha_{i_j}, \dots, \alpha_{i_n} - \sum_{j=1}^l c_{j(n-l)} \alpha_{i_j})$ 为 V 的一组基, 不妨把它改记为 $(\beta_1, \dots, \beta_l, \beta_{l+1}, \dots, \beta_n)$, 则前 l 个向量的像构成了 $f(V)$ 的基, 后 $n-l$ 个基向量构成了 $\ker(f)$ 里的线性无关组。且设 $a = \sum_{i=1}^n x_i \beta_i, a \in \ker f \Leftrightarrow f(a) = \sum_{i=1}^l x_i f(\beta_i) = 0 \Leftrightarrow x_i = 0, \forall i \leq l$, 这表明了 $(\beta_{l+1}, \dots, \beta_n)$ 构成了 $\ker(f)$ 的基, 从而原题等式成立。

以上是关于线性映射是单射的使用零空间的描述方法, 接下来会尝试用线性映射的矩阵来描述。以下先介绍一个关于映射的常用命题。

定理 5.16

设 $f: A \rightarrow B$ 为一个映射, 则

1. f 为单射当且仅当存在 $g: B \rightarrow A$, 使得 $g \circ f = \operatorname{id}_A$, 即在映射复合的意义下 f 有左逆;
2. f 为满射当且仅当存在 $g: B \rightarrow A$, 使得 $f \circ g = \operatorname{id}_B$, 即在映射复合的意义下 f 有右逆;
3. f 为双射当且仅当存在 f 在映射复合的意义下既有左逆也有右逆, 且此时左逆和右逆相等且唯一, 故 f 有可定义的逆映射。

**证明**

1. $(\Rightarrow)$ 设 f 为单射, 定义映射 $g: B \rightarrow A$ 如下: $g(b) = \begin{cases} a, & b \in f(A), f(a) = b \\ a_0, & b \notin f(A) \end{cases}$, 其中 $a_0 \in A$ 为任意选定的元素。则对任意 $a \in A$, 有 $g(f(a)) = a$, 即 $g \circ f = \operatorname{id}_A$ 。

的元素。则对任意 $a \in A$, 有 $g(f(a)) = a$, 即 $g \circ f = \operatorname{id}_A$ 。

$(\Leftarrow)$ 设存在映射 $g: B \rightarrow A$, 使得 $g \circ f = \operatorname{id}_A$ 。若 $f(a_1) = f(a_2)$, 则有 $a_1 = g(f(a_1)) = g(f(a_2)) = a_2$, 故 f 为单射。

2. $(\Rightarrow)$ 设 f 为满射, 则可利用选择公理, 从 B 中每个元素的原像中选取一个元素定义 g , 即 $g(b) = a_b \in \{a \mid f(a) = b\}$, 此时对任意 $b \in B$, 有 $f(g(b)) = b$, 即 $f \circ g = \operatorname{id}_B$ 。

$(\Leftarrow)$ 设存在映射 $g: B \rightarrow A$, 使得 $f \circ g = \operatorname{id}_B$ 。则由 $\operatorname{Im} f \circ g \subset \operatorname{Im} f$ 和 $\operatorname{Im} f \circ g = B$ 可知 $\operatorname{Im} f = B$, 故 f 为满射。

3. 由前两点可知, f 为双射当且仅当存在映射 $g: B \rightarrow A$, 使得 $g \circ f = \operatorname{id}_A$ 与 $g': B \rightarrow A$, 使得 $f \circ g' = \operatorname{id}_B$ 。又由映射复合的结合律有 $g = g \circ (f \circ g') = (g \circ f) \circ g' = g'$, 故左逆与右逆相等且唯一。

现在就可以将上述定理和现有知识进行类比了。

定理 5.17

设 $f: V \rightarrow V'$ 是线性映射, V, V' 有基 M_1, M_2 , 且 V 为有限维线性空间, 设 f 在基 M_1, M_2 下的矩阵为 A , 则:

1. f 为单射的充分必要条件是存在线性映射 $g: V' \rightarrow V$, 使得 $g \circ f = \operatorname{id}_V$, 即在映射复合的意义下 f 有左逆, 也即 A 有左逆矩阵;

2. f 为满射的充分必要条件是存在线性映射 $g: V' \rightarrow V$, 使得 $f \circ g = \text{id}_{V'}$, 即在映射复合的意义下 f 有右逆, 也即 A 有右逆矩阵;
3. f 为双射的充分必要条件是 f 有左逆也有右逆, 且此时左逆和右逆相等且唯一, 故 f 有可定义的逆映射, 也即 A 可逆。



证明 无论 V 是否和 V' 是否等同, 总可以改变两组基的选取使得 f 在它们下的矩阵左上角为 I_r , 其余全为零, 即 A 可和上述形式的矩阵相抵。在相抵意义下, 是否有左逆, 有右逆等为没有改变的性质。这样的矩阵共有四类分别为 $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_r & 0 \end{bmatrix}, I_r$, 它们的一些讨论如下:

1. $\text{rank}(A) < \min\{\dim V', \dim V\}$, 此时非单射也非满射, 无左逆也无右逆, A 亦同;
2. $\text{rank}(A) = \dim V \leq \dim V'$, 此时 f 为单射, 有左逆, A 亦同;
3. $\text{rank}(A) = \dim V' \leq \dim V$, 此时 f 为满射, 有右逆, A 亦同;
4. $\text{rank}(A) = \dim V = \dim V'$, 此时 f 为双射, 有左右逆, 且左右逆相等且唯一, A 亦同。

综上所述, 命题得证。

上述定理表明线性映射的单满射性仅仅取决于 $\text{rank}(A)$, 原空间维数, 像所在空间维数的大小关系。对于原空间和像所在空间维数相等的线性映射, 讨论会简单一些。设 f 为这样的映射, 有以下几点相互等价:

1. $\ker f = \{0\}$;
2. f 为单射;
3. f 在任意一组基下的矩阵列满秩;
4. f 在任意一组基下的矩阵行满秩;
5. f 为满射;
6. f 为双射。

5.3.4 从线性变换观点看线性方程组

先考虑一个练习

例题 5.6 设 W 是 $\mathbb{F}^n$ 的子空间, 证明存在 $A \in \mathbb{F}^{m \times n}$ 使得 $Ax = 0$ 的解空间为 W 。

解 设 W 的基为 $(\alpha_1, \alpha_2, \dots, \alpha_k)$, 则可将其扩充为 $\mathbb{F}^n$ 的一组基 $(\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_{n-k})$ 。设这些基组成矩阵为 B , 则取 $A = \text{diag}(O_k, I_{n-k})B^{-1}$, 则 $A\alpha_i = 0, A\beta_i = e_i$, 且对任意 $x \in \mathbb{F}^n$, 有 $x = \sum_{i=1}^k c_i \alpha_i + \sum_{j=1}^{n-k} d_j \beta_j$, 则 $Ax = \sum_{j=1}^{n-k} d_j e_j$, 因此 $Ax = 0 \Leftrightarrow d_j = 0, \forall j \Leftrightarrow x \in W$ 。故 $Ax = 0$ 的解空间为 W 。

考虑更一般的情形, 设 $A \in \mathbb{F}^{m \times n}$, 则线性方程组 $Ax = b$ 可看成对于某个线性映射的坐标表示 $\mathcal{A}: \mathbb{F}^{n \times 1} \rightarrow \mathbb{F}^{m \times 1}, x \mapsto Ax$, 寻找 $b \in \mathbb{F}^{m \times 1}$ 在 $\mathbb{F}^{n \times 1}$ 中的原像。此时可以将方程组的有解情况和线性映射的单满射情况联系起来:

1. $\forall b \in \mathbb{F}^{m \times 1}, Ax = b$ 至多有一个解当且仅当 A 为单射。
2. $\forall b \in \mathbb{F}^{m \times 1}, Ax = b$ 有解当且仅当 A 为满射。
3. $\forall b \in \mathbb{F}^{m \times 1}, Ax = b$ 有唯一解当且仅当 A 为双射。

因此方程组的解集合就是线性映射的原像集合, 可以从线性映射的角度考虑问题。对于解集非空的方程组, 设 $Ax = b$ 有解为 x_0 , 则对于 $\mathbb{F}^{n \times 1}$ 中任意一个元素 x , 有 $x \in \mathcal{A}^{-1}(b) \Leftrightarrow \mathcal{A}(x - x_0) = b - b = 0 \Leftrightarrow x - x_0 \in \ker \mathcal{A}$, 从而可得 $\mathcal{A}^{-1}(b) = \{x + x_0 \mid x \in \ker \mathcal{A}\} := x_0 + \ker \mathcal{A}$ 。对于齐次线性方程组, 它的解集就是 $\ker \mathcal{A}$, 这个集合是 $\mathbb{F}^{n \times 1}$ 的一个子空间, 以下是它的一些性质:

1. 方程的自由度, 即零空间的维度 $\dim \ker \mathcal{A} = n - \text{rank } A \geq n - \min\{m, n\}$;

2. $\ker A$ 的基就是方程的基础解系。

第六章 期中复习

6.1 复数与向量

以下是一些基本的概念与性质。

定义 6.1

对于使用有序数组描述的向量 $a = (a_1, a_2, \dots, a_n), b = (b_1, b_2, \dots, b_n)$, 其内积定义为 $a \cdot b = \sum_{i=1}^n a_i b_i$, a 的模长定义为 $\|a\| = \sqrt{a \cdot a}$, a 与 b 的夹角 θ 定义为 $\theta = \arccos \frac{a \cdot b}{\|a\| \|b\|}$, 其中 $0 \leq \theta \leq \pi$ 。若 $\theta = 0$ 或 π , 则称 a 与 b 平行; 若 $\theta = \frac{\pi}{2}$, 则称 a 与 b 正交。

命题 6.2

1. 内积满足正定性, 线性性与对称性;
2. $\|a + b\| \leq \|a\| + \|b\|$;
3. $|a \cdot b| \leq \|a\| \|b\|$ (柯西不等式)。

定义 6.3

对于 $\mathbb{R}^3$ 中的向量 $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3)$, 其外积定义为 $a \times b = \det \begin{bmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$, 若 $c = (c_1, c_2, c_3)$, 定义其混合积为 $(a, b, c) = a \cdot (b \times c) = (a \times b) \cdot c = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ 。

命题 6.4

1. 外积满足反对称性与分配律;
2. $\|a \times b\| = \|a\| \|b\| \sin \theta$, 其中 θ 为 a 与 b 的夹角;
3. 混合积 $(a, b, c) = (b, c, a) = (c, a, b) = -(b, a, c) = -(a, c, b) = -(c, b, a)$ 。
4. 拉格朗日恒等式: $(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c)$ 。

定义 6.5

复数 z 是形如 $z = a + bi$ 的数, 其中 $a, b \in \mathbb{R}$, i 为虚数单位, 满足 $i^2 = -1$ 。其中 a 称为实部, 记为 $\operatorname{Re}(z)$; b 称为虚部, 记为 $\operatorname{Im}(z)$ 。复数的共轭复数定义为 $\bar{z} = a - bi$ 。复数 $a + bi$ 的模长定义为 $|z| = \sqrt{a^2 + b^2}$, 三角形式定义为 $z = |z|(\cos \theta + i \sin \theta)$, 其中 θ 称为幅角, 记为 $\operatorname{Arg}(z)$ 。把在 $[0, 2\pi)$ 范围内的幅角称为幅角主值, 记为 $\arg(z)$ 。

定理 6.6 (Euler 公式)

对于任意实数 θ , 有 $e^{i\theta} = \cos \theta + i \sin \theta$ 。

定义 6.7

n 次单位根是指 $\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, k = 0, 1, \dots, n-1$ 这 n 个复数。记 $\omega_n = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$, 则 n 次单位根可表示为 $\omega_n^k, k = 0, 1, \dots, n-1$ 。 n 次单位根是方程 $z^n = 1$ 的全部解。

6.2 线性方程组的 Gauss 消元法

可以把线性方程组写成矩阵形式 $Ax = b$, 其中 A 称为系数矩阵, x 是未知数向量, b 称为常数向量, Gauss 消元法的步骤如下:

算法 6.8

1. 将线性方程组写成增广矩阵形式 $[A|b]$;
2. 通过初等行变换将增广矩阵化为行阶梯形矩阵 $[U|b']$;
3. 若 $[U|b']$ 存在形如 $[0 \cdots 0 | c]$ 的行 ($c \neq 0$), 则方程组无解, 否则将不作为某行首个非零元的未知数设为自由变量, 通过回代求出方程组的通解。

例题 6.1 求线性方程组的解:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 10 \\ 2x_1 - x_2 + 2x_3 - x_4 = 2 \\ 3x_1 + x_2 - x_3 + 2x_4 = 10 \\ 4x_1 + 2x_2 + 3x_4 = 20 \end{cases}$$

$$\text{解} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 10 \\ 2 & -1 & 2 & -1 & 2 \\ 3 & 1 & -1 & 2 & 10 \\ 4 & 2 & 0 & 3 & 20 \end{bmatrix} \xrightarrow[r_4 - 4r_1]{\begin{matrix} r_2 - 2r_1 \\ r_3 - 3r_1 \end{matrix}} \begin{bmatrix} 1 & 1 & 1 & 1 & 10 \\ 0 & -3 & 0 & -3 & -18 \\ 0 & -2 & -4 & -1 & -20 \\ 0 & -2 & -4 & -1 & -20 \end{bmatrix} \xrightarrow[r_2 \div (-3)]{\begin{matrix} r_4 - r_3 \\ r_3 - \frac{2}{3}r_2 \end{matrix}} \begin{bmatrix} 1 & 1 & 1 & 1 & 10 \\ 0 & 1 & 0 & 1 & 6 \\ 0 & 0 & -4 & 1 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

取 x_4 为自由变量, 则由第三行得 $x_3 = 2 + \frac{1}{4}x_4$, 由第二行得 $x_2 = 6 - x_4$, 由第一行得 $x_1 = 2 - \frac{1}{4}x_4$, 故通解

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -\frac{1}{4} \\ -1 \\ \frac{1}{4} \\ 1 \end{bmatrix}, \quad x_4 \in \mathbb{R}$$

6.3 行列式

以下是几种常见的行列式定义方式:

定义 6.9 (逆序数定义)

设 $A = (a_{ij})_{n \times n}$ 是一个 n 阶方阵, 则其行列式定义为 $\det A = \sum_{(j_1, \dots, j_n) \in S_n} (-1)^{\tau(j_1, \dots, j_n)} \prod_{i=1}^n a_{ij_i}$ 。其中 S_n 是 n 个元素的全排列集合, $\tau(j_1, \dots, j_n)$ 是排列 $(j_1, \dots, j_n)$ 的逆序数。

定义 6.10 (递归定义)

设 $A = (a_{ij})_{n \times n}$ 是一个 n 阶方阵, 则其行列式定义为 $\det A = \begin{cases} a_{11}, & n = 1 \\ \sum_{j=1}^n (-1)^{1+j} a_{1j} M_{1j}, & n > 1 \end{cases}$ 。其中 M_{1j} 是由 A 删除第 1 行与第 j 列后得到的 $(n-1)$ 阶子矩阵的行列式, 称为 a_{1j} 的余子式。 a_{ij} 的代数余子式定义为 $A_{ij} = (-1)^{i+j} M_{ij}$ 。

定义 6.11 (公理化定义)

n 维向量空间 $\mathbb{F}^n$ 上的规范反对称 n 重线性函数称为数域 $\mathbb{F}$ 上 n 阶行列式函数, 简称 n 阶行列式。

换言之, 行列式函数 $\det: \mathbb{F}^n \times \mathbb{F}^n \times \cdots \times \mathbb{F}^n \rightarrow \mathbb{F}$ 是唯一满足以下三条性质的函数:

1. 规范性: $\det(e_1, e_2, \dots, e_n) = 1$, 其中 $e_1, e_2, \dots, e_n$ 是 $\mathbb{F}^n$ 的一组标准正交基;
2. 反对称性: $\forall 1 \leq i < j \leq n, \det(\dots, \alpha_i, \dots, \alpha_j, \dots) = -\det(\dots, \alpha_j, \dots, \alpha_i, \dots)$;
3. 线性性: $\forall \lambda, \mu \in \mathbb{F}, \det(\lambda \alpha_1 + \mu \beta_1, \alpha_2, \dots, \alpha_n) = \lambda \det(\alpha_1, \alpha_2, \dots, \alpha_n) + \mu \det(\beta_1, \alpha_2, \dots, \alpha_n)$ 。

命题 6.12

设 $A = (a_{ij})_{n \times n}$, 则有以下性质:

1. $\det A = \det A^T$;
2. 若 A 的某一行(列)全为零, 则 $\det A = 0$;
3. 若 A 的两行(列)相等, 则 $\det A = 0$;
4. 若 A 的两行(列)互换位置, 则行列式变号;
5. 若 A 的某一行(列)乘以 λ , 则行列式也乘以 λ ;
6. 若 A 的某一行(列)加上另一行(列)的 λ 倍, 则行列式不变;
7. 若 A 为上三角矩阵或下三角矩阵, 则 $\det A = \prod_{i=1}^n a_{ii}$ 。

以下是一些常见的计算行列式的方法:

1. 直接使用定义计算(适用于不超过 4 阶的行列式);
2. 加边法(加的边根据行列式的规律而定, 最好可以消掉方阵中大多数元素);
3. 逐行/列加减(目标是得到(准)上/下三角矩阵);
4. 拆项与递推(适用于行列式元素有规律的情况);
5. 进行乘法分解;
6. 考虑使用多项式技巧。

定理 6.13 (Laplace 展开定理)

设 A 是 n 阶方阵, $A = (a_{ij})_{n \times n}$, 对任意正整数 $r < n$, 任意取定 r 个指标 $1 \leq i_1 < i_2 < \cdots < i_r \leq n$, 则 $\det A$ 的值等于它的第 $i_1, i_2, \dots, i_r$ 行(或列)元组成的所有 r 阶子式分别与它们的代数余子式的乘积之和。即

$$\begin{aligned} \det A &= \sum_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} A \begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} (-1)^{\sum_{k=1}^r (i_k + j_k)} A \begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ j_{r+1} & j_{r+2} & \cdots & j_n \end{pmatrix} \\ &= \sum_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} A \begin{pmatrix} j_1 & j_2 & \cdots & j_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} (-1)^{\sum_{k=1}^r (i_k + j_k)} A \begin{pmatrix} j_{r+1} & j_{r+2} & \cdots & j_n \\ i_{r+1} & i_{r+2} & \cdots & i_n \end{pmatrix} \end{aligned}$$

其中 $i_{r+1}, i_{r+2}, \dots, i_n$ 是由 $1, 2, \dots, n$ 去掉 $i_1, i_2, \dots, i_r$ 后剩下的数按从小到大排列得到, $j_{r+1}, j_{r+2}, \dots, j_n$ 是由 $1, 2, \dots, n$ 去掉 $j_1, j_2, \dots, j_r$ 后剩下的数按从小到大排列得到。

定理 6.14 (Binet-Cauchy 公式)

设 A 为 $m \times n$ 矩阵, B 为 $n \times m$ 矩阵, 则有

$$\det(AB) = \begin{cases} \det A \det B, & m = n \\ 0, & m > n \\ \sum_{1 \leq j_1 < j_2 < \cdots < j_m \leq n} \det A \begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} \det B \begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}, & m < n \end{cases}$$



命题 6.15 (乘法分解)

- 若 A 可逆, 则 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det(D - CA^{-1}B)$ 。
- 若 D 可逆, 则 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det D \cdot \det(A - BD^{-1}C)$ 。
- 若 $A \in \mathbb{F}^{n \times m}, B \in \mathbb{F}^{m \times n}$, 则 $\lambda^m \det(\lambda I_n - AB) = \lambda^n \det(\lambda I_m - BA)$ 。
- 若 $A \in \mathbb{F}^{n \times n}, B \in \mathbb{F}^{m \times n}$, 则 $\det(A - B^T B) = \det A \cdot \det(I_m - BA^{-1}B^T)$ 。



命题 6.16 (三对角线矩阵的行列式)

设 $\Delta_n = \det \begin{bmatrix} a & b & & \\ c & a & b & \\ & c & a & \ddots \\ & & \ddots & \ddots & b \\ & & & c & a \end{bmatrix}_{n \times n}$, $x, y \in \mathbb{C}$ 是 $\lambda^2 - a\lambda + bc = 0$ 的根, 则有 $\Delta_n = \sum_{k=0}^n x^k y^{n-k}$ 。



定理 6.17 (Cramer 法则)

对 n 阶矩阵 A , 若 $\det A \neq 0$, 则线性方程组 $Ax = b$ 有唯一解, 则该解的第 i 个分量为 $x_i = \frac{\det A_i}{\det A}$, 其中 A_i 是由 A 的第 i 列替换为常数向量 b 后得到的矩阵。



例题 6.2 判断以下命题是否正确。若正确, 请证明; 若不正确, 请给出反例。

1. 若线性方程组有唯一解, 则可用 Cramer 法则求解;
2. 若矩阵 A, B 满足 AB, BA 均有定义, 则 $\det(AB) = \det(BA)$;
3. 设 A, B 为 n 阶矩阵, 则 $\det \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \det(A+B) \cdot \det(A-B)$ 。

解

1. 不正确。如 $\begin{cases} x = 1 \\ 2x = 2 \end{cases}$ 有唯一解 $x = 1$, 但系数矩阵 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 非方阵, 不能使用 Cramer 法则。
2. 不正确。考虑 $A = \begin{bmatrix} 1 & 1 \end{bmatrix}, B = A^T$, 则 $AB = \begin{bmatrix} 2 \end{bmatrix}, BA = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, 故 $\det(AB) = 2, \det(BA) = 0$ 。
3. 正确。 $\det \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \det \begin{bmatrix} A+B & A+B \\ B & A \end{bmatrix} = \det \begin{bmatrix} A+B & O \\ B & A-B \end{bmatrix} = \det(A+B) \cdot \det(A-B)$ 。

例题 6.3 计算行列式 $\det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix}_{n \times n}$ 。

解 $\det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix} \xrightarrow[i=2, \dots, n]{-r_1 \rightarrow r_i} \det \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ -2 & 0 & \cdots & 0 & 0 \\ -2 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ -2 & -2 & \cdots & -2 & 0 \end{bmatrix} \xrightarrow{\text{最后一列展开}} (-1)^{n+1} (-2)^{n-1} = 2^{n-1}$

例题 6.4 $A \in \mathbb{F}^{n \times n}$, 证明: 存在非零矩阵 $B \in \mathbb{F}^{n \times p}$, 使得 $AB = O$ 当且仅当 $\det A = 0$ 。

证明 充分性: 若 $\det A = 0$, 则 A 不可逆, 故 $\exists x \neq 0$, 使得 $Ax = 0$ 。取 $B = (x, 0, \dots, 0)$ 即可。

必要性: 若存在非零矩阵 $B \in \mathbb{F}^{n \times p}$, 使得 $AB = O$, 则 B 的每一列都是 $Ax = 0$ 的解, 故 $Ax = 0$ 有非零解, 故 $\det A = 0$ 。

6.4 矩阵

定义 6.18 (矩阵及其运算)

矩阵是一个由数域 $\mathbb{F}$ 中元素按长方阵列排列成的二维数组。设 $A = (a_{ij})_{m \times n}$, 则称 A 为一个 m 行 n 列的矩阵, a_{ij} 称为矩阵 A 的第 i 行第 j 列元素, m 称为矩阵的行数, n 称为矩阵的列数。若 $m = n$, 则称 A 为方阵, 称其阶数为 n 。

设 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{m \times n}$, 他们的加法定义为 $A + B = (a_{ij} + b_{ij})_{m \times n}$; 对于 $\lambda \in \mathbb{F}$, 矩阵的数乘定义为 $\lambda A = (\lambda a_{ij})_{m \times n}$ 。

对于 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{p \times q}$, 当且仅当 $n = p$ 时, 矩阵的乘法有定义, 定义为 $AB = C = (c_{ij})_{m \times q}$, 其中 $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ 。

定义 6.19 (转置与迹)

设 $A = (a_{ij})_{m \times n}$, 则矩阵 A 的转置定义为 $A^T = (a_{ji})_{n \times m}$ 。若 A 为 n 阶方阵, 则称矩阵 A 的迹为 $\text{tr} A = \sum_{i=1}^n a_{ii}$ 。

命题 6.20

设 A, B, C 为适当阶数的矩阵, $\lambda, \mu \in \mathbb{F}$, 则有以下性质:

1. $(A^T)^T = A$;
2. $(A + B)^T = A^T + B^T$;
3. $(\lambda A)^T = \lambda A^T$;
4. $(AB)^T = B^T A^T$;
5. $\text{tr}(A + B) = \text{tr} A + \text{tr} B$;
6. $\text{tr}(\lambda A) = \lambda \text{tr} A$;
7. $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$ 。
8. $\text{tr}(AA^T) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$, 特别地, 对于实矩阵 A , $\text{tr}(AA^T) = 0$ 当且仅当 $A = O$ 。

定义 6.21 (逆矩阵)

设 $A \in \mathbb{R}^{n \times n}$, 若存在矩阵 $B \in \mathbb{R}^{n \times n}$, 使得 $AB = BA = I_n$, 则称 A 可逆, 称 B 为 A 的逆矩阵, 记为 A^{-1} 。不可逆的方阵称为奇异阵。矩阵 A 可逆的充分必要条件是 $\det A \neq 0$ 。

命题 6.22

设 A, B 为可逆矩阵, 则有以下性质:

1. $(A^{-1})^{-1} = A$;
2. $(AB)^{-1} = B^{-1}A^{-1}$;
3. $(A^T)^{-1} = (A^{-1})^T$ 。
4. 上/下三角矩阵可逆的充分必要条件是其对角线元素均不为零, 且其逆矩阵仍为上/下三角矩阵。
5. 对称矩阵/反对称矩阵的逆矩阵仍为对称矩阵/反对称矩阵。

定义 6.23 (伴随矩阵)

设 $A = (a_{ij})_{n \times n}$, 则矩阵 A 的伴随矩阵定义为 $A^* = (A_{ji})_{n \times n}$, 其中 A_{ji} 是 a_{ji} 的代数余子式。 $AA^* = \det A \cdot I_n$ 。

命题 6.24 (伴随矩阵性质)

设 A, B 为 n 阶矩阵, $\lambda \in \mathbb{R}$, 则有以下性质:

1. $\det A^* = (\det A)^{n-1}$;
2. $(A^*)^* = (\det A)^{n-2}A (n \geq 2)$;
3. $(A^T)^* = (A^*)^T$;
4. $(\lambda A)^* = \lambda^{n-1}A^*$;
5. $(AB)^* = B^*A^*$ 。

希望同学们可以尝试证明这些性质。

算法 6.25 (计算逆矩阵)

可以通过以下方法计算逆矩阵:

1. 若 $\det A \neq 0$, 则 $A^{-1} = \frac{1}{\det A} A^*$;
2. 将增广矩阵 $[A|I_n]$ 通过初等行变换化为 $[I_n|A^{-1}]$ 。注意一定不能使用列变换。
3. 使用 Sherman-Morrison 公式 $(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}$, 适用于 A 可逆且 $1 + v^T A^{-1}u \neq 0$ 的情况。

命题 6.26 (分块矩阵运算)

分块矩阵是将矩阵分割为多个子矩阵, 以便于处理较大的矩阵。设矩阵 A 可以分块表示为 $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, B

可以分块表示为 $\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$, 则在维度匹配时具有如下性质:

- 加法: $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}$;
- 乘法: $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$ 。

除了分块为 2×2 的矩阵外, 分块矩阵的加法与乘法也可以推广到更高维度的情况。

例题 6.5 判断以下命题是否正确。若正确, 请证明; 若不正确, 请给出反例。

1. 设 A, B, C 是三个实对称矩阵, 则 $\text{tr}(ABC) = \text{tr}(ACB)$;
2. 对 $A, B \in \mathbb{R}^{m \times n}$, $\text{tr}((A-B)(A^\top - B^\top)) = 0$, 则 $A = B$;

解

1. 正确。 $\text{tr}(ABC) = \text{tr}((ABC)^\top) = \text{tr}(C^\top B^\top A^\top) = \text{tr}(CBA) = \text{tr}(ACB)$ 。
2. 正确。 $\text{tr}((A-B)(A^\top - B^\top)) = \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - b_{ij})^2$, 故 $\text{tr}((A-B)(A^\top - B^\top)) = 0$ 当且仅当 $A = B$ 。

例题 6.6

1. 计算 $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{2025}$;
2. 计算 $\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix}_{n \times n}^{-1}$ 。

解

$$1. \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{2025} = \left(I + \begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & \\ 0 & & 0 \end{bmatrix} \right)^{2025} = I + 2025 \begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & \\ 0 & & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4050 \\ 0 & 1 & -2025 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.

$$\begin{aligned} & \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \ddots \\ -1 & -1 & -1 & \cdots & 1 & 1 \end{bmatrix} \xrightarrow[i=2, \dots, n]{-r_{i-1} \rightarrow r_i} \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 \\ -2 & 0 & \cdots & 0 & 0 & -1 & 1 \\ 0 & -2 & \cdots & 0 & 0 & -1 & 1 \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \ddots \\ 0 & 0 & \cdots & -2 & 0 & & -1 & 1 \end{bmatrix} \\ & \xrightarrow[r_i \div 2, i \neq n]{\text{交换行}} \begin{bmatrix} -1 & & & & -\frac{1}{2} & \frac{1}{2} \\ & -1 & & & -\frac{1}{2} & \frac{1}{2} \\ & & \ddots & & & \ddots & \ddots \\ & & & -1 & & -\frac{1}{2} & \frac{1}{2} \\ 1 & 1 & \cdots & 1 & 1 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ & \xrightarrow[i=1, \dots, n-1]{r_i \rightarrow r_n} \begin{bmatrix} -1 & & & & -\frac{1}{2} & \frac{1}{2} \\ & -1 & & & -\frac{1}{2} & \frac{1}{2} \\ & & \ddots & & & \ddots & \ddots \\ & & & -1 & & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \cdots & 0 & 1 & \frac{1}{2} & 0 & \cdots & 0 & \frac{1}{2} \end{bmatrix} \\ & \xrightarrow[i=1, \dots, n-1]{r_i \times (-1)} \begin{bmatrix} 1 & & & & \frac{1}{2} & -\frac{1}{2} \\ & 1 & & & \frac{1}{2} & -\frac{1}{2} \\ & & \ddots & & & \ddots & \ddots \\ & & & 1 & & \frac{1}{2} & -\frac{1}{2} \\ & & & & 1 & \frac{1}{2} & 0 & \cdots & 0 & \frac{1}{2} \end{bmatrix} \end{aligned}$$

$$\text{故} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & & & \\ & \frac{1}{2} & -\frac{1}{2} & & \\ & & \ddots & \ddots & \\ & & & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & \cdots & 0 & \frac{1}{2} \end{bmatrix}.$$

例题 6.7

1. 设 A, B 是 n 阶可逆方阵, $\det A = \lambda, \det B = \mu$, 计算 $\det \begin{bmatrix} O & A^* \\ 2B & O \end{bmatrix}$.
2. 设矩阵 $B = (b_{ij})_{3 \times 3}$ 满足 $B^* = B^T$, 若 b_{11}, b_{12}, b_{13} 为 3 个相等正数, 求 b_{11} .

解

1. $\det \begin{bmatrix} O & A^* \\ 2B & O \end{bmatrix} = (-1)^n \det \begin{bmatrix} 2B & \\ & A^* \end{bmatrix} = (-2)^n \det B \cdot \det A^* = (-2)^n \mu \lambda^{n-1}$.
2. $(\det B)^2 = \det B^* = \det B^T = \det B \Rightarrow \det B = 0$ 或 $\det B = 1$, $\det B \cdot \delta_{ij} = (BB^*)_{ij} = (BB^T)_{ij} = \sum_{k=1}^3 b_{ik}b_{jk}$, 取 $i = j = 1$, 即有 $0 < \sum_{k=1}^3 b_{1k}^2 = 3b_{11}^2 = \det B$, 故 $\det B = 1, b_{11}^2 = \frac{1}{3}, b_{11} = \frac{\sqrt{3}}{3}$.

6.5 秩与相抵

定义 6.27 (初等矩阵)

回忆一下三种初等变换: 交换两行(列)、某一行(列)乘以非零常数、某一行(列)加上另一行(列)的若干倍。通过对单位矩阵进行一次初等行变换(或列变换)得到的矩阵称为初等矩阵。初等矩阵分为三类:

1. 交换两行(列)的初等矩阵 S_{ij} ;
2. 某一行(列)乘以非零常数的初等矩阵 $D_i(\lambda)$;
3. 某一行(列)加上另一行(列)的若干倍的初等矩阵 $T_{ij}(\lambda)$ 。

对矩阵进行初等行变换(列变换)相当于将该矩阵左乘(右乘)对应的初等矩阵。

命题 6.28

初等矩阵的逆矩阵如下: $S_{ij}^{-1} = S_{ij}, D_i(\lambda)^{-1} = D_i\left(\frac{1}{\lambda}\right), T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$ 。

定理 6.29

对于任意矩阵 $A \in \mathbb{F}^{m \times n}$, 都存在初等矩阵 $P_1, P_2, \dots, P_s$ 和 $Q_1, Q_2, \dots, Q_t$, 使得 $P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, 其中 r 是非负整数且由矩阵 A 唯一确定, 称为矩阵 A 的秩, 记为 $\text{rank } A = r$ 。这等价于存在可

逆矩阵 $P \in \mathbb{F}^{m \times m}, Q \in \mathbb{F}^{n \times n}$, 使得 $PAQ = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ 。

定义 6.30 (矩阵的相抵)

设 $A, B \in \mathbb{F}^{m \times n}$, 若可以经过有限次初等行变换与初等列变换, 将矩阵 A 变为矩阵 B , 则称矩阵 A 与矩阵 B 相抵。这等价于存在可逆矩阵 $P \in \mathbb{F}^{m \times m}, Q \in \mathbb{F}^{n \times n}$, 使得 $B = PAQ$ 。

命题 6.31

相抵是等价关系, 即相抵满足以下性质:

- 自反性: $\forall A \in \mathbb{F}^{m \times n}$, A 与 A 相抵;
- 对称性: $\forall A, B \in \mathbb{F}^{m \times n}$, 若 A 与 B 相抵, 则 B 与 A 相抵;
- 传递性: $\forall A, B, C \in \mathbb{F}^{m \times n}$, 若 A 与 B 相抵, 且 B 与 C 相抵, 则 A 与 C 相抵。

定理 6.32

设 $A, B \in \mathbb{F}^{m \times n}$, 则 A 与 B 相抵的充分必要条件是 $\text{rank } A = \text{rank } B$ 。矩阵 $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ (其中 r 为非负整数) 是所有 $m \times n$ 矩阵中秩为 r 的矩阵的唯一的相抵标准形。若 $\text{rank } A = m$, 称矩阵 A 为行满秩矩阵; 若 $\text{rank } A = n$, 称矩阵 A 为列满秩矩阵。

初等变换不改变矩阵的秩, 即 $\text{rank}(PAQ) = \text{rank } A$, 其中 P, Q 分别为适当阶数的可逆矩阵。

定理 6.33

矩阵 A 的秩等于其最大非零子式的阶数。

命题 6.34

设 A, B, C 为适当阶数的矩阵, 则有以下性质:

1. $\text{rank } A + \text{rank } B = \text{rank} \begin{bmatrix} A & O \\ O & B \end{bmatrix}$;
2. $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$;
3. 若 P 为可逆矩阵, 则 $\text{rank}(PA) = \text{rank}(AP) = \text{rank}(A)$;
4. 矩阵 A 某个子块的秩不大于矩阵 A 的秩;
5. 对实矩阵 A , $\text{rank}(A) = \text{rank}(A^T A) = \text{rank}(AA^T)$;
6. $\text{rank} \begin{bmatrix} A & O \\ O & B \end{bmatrix} \leq \text{rank} \begin{bmatrix} A & C \\ O & B \end{bmatrix}$
7. $\text{rank}(A+B) \leq \text{rank} \begin{bmatrix} A & B \end{bmatrix} \leq \text{rank } A + \text{rank } B$
8. Sylvester 不等式: $\text{rank}(AC) + n \geq \text{rank } A + \text{rank } C$
9. Frobenius 不等式: $\text{rank}(AB) + \text{rank}(BC) \leq \text{rank}(ABC) + \text{rank } B$
10. $n + \text{rank}(I_m - AB) = \text{rank} \begin{bmatrix} I_m & A \\ B & I_n \end{bmatrix} = m + \text{rank}(I_n - BA)$

命题 6.35

$$\text{rank } A^* = \begin{cases} 0, & \text{rank } A < n-1 \\ 1, & \text{rank } A = n-1 \\ n, & \text{rank } A = n \end{cases}, \text{rank}(A^*)^* = \begin{cases} n, & \text{rank } A = n \\ 0, & \text{rank } A < n \end{cases}.$$

例 6.8 判断以下命题是否正确。若正确, 请证明; 若不正确, 请给出反例。

1. 设 $A_1, A_2, \dots, A_n \in \mathbb{F}^{n \times n}$, 且 $A_1 A_2 \dots A_n = O$, 则 $\sum_{i=1}^n \text{rank } A_i \leq n^2 - n$;
2. 设 $A \in \mathbb{F}^{n \times n}$, n 是偶数, 则 $A^2 - 2A + I = O$ 的充分必要条件是 $\text{rank}(A - I) = \frac{n}{2}$ 。

解

1. 正确。 $\sum_{i=1}^n \text{rank } A_i \leq n + \text{rank}(A_1 A_2) + \sum_{i=3}^n \text{rank } A_i \leq 2n + \text{rank}(A_1 A_2 A_3) + \sum_{i=4}^n \text{rank } A_i \leq \cdots \leq (n-1)n + \text{rank}(A_1 A_2 \cdots A_n) = (n-1)n$ 。

2. 错误。考虑 $A = I$ 即可。

 **笔记** 此题第二问与例4.2与例4.3类似，实际上有更强的结论：对 $A \in \mathbb{F}^{n \times n}$, $p_1, p_2, \dots, p_m \in \mathbb{F}[x]$ 是两两互素的多项式，且 $\prod_{k=1}^m p_k(A) = O$ ，则 $\mathbb{F}^n = \bigoplus_{k=1}^m \ker P_k(A)$

例题 6.9 对于 n 阶方阵 $D_n = \begin{bmatrix} 0 & 1 & & \\ -1 & 0 & 1 & \\ & -1 & 0 & \ddots \\ & & \ddots & \ddots & 1 \\ & & & -1 & 0 \end{bmatrix}$,

1. 计算 $\text{rank } D_n$;

2. 记 $B_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \ddots \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$, $C_n = \begin{bmatrix} 0 & & & \\ -1 & 0 & & \\ & -1 & 0 & \\ & & \ddots & \ddots \\ & & & 1 & 0 \end{bmatrix}$, 能否利用 $D_n = B_n + C_n$ 的二项式展开式计算出 D_n^k ? 请说明理由。

解

1. $\det D_n \xrightarrow{\text{第一行展开}} -\det \begin{bmatrix} & 1 & & \\ -1 & 0 & 1 & \\ & -1 & 0 & \ddots \\ & & \ddots & \ddots & 1 \\ & & & -1 & 0 \end{bmatrix} \xrightarrow{\text{第一列展开}} \det D_{n-2}$, 且 $\det D_1 = 0, \det D_2 = 1$,

故 $\det D_n = \begin{cases} 0, & n \text{ 为奇数} \\ 1, & n \text{ 为偶数} \end{cases}$, 从而 $\text{rank } D_n = \begin{cases} n-1, & n \text{ 为奇数} \\ n, & n \text{ 为偶数} \end{cases}$ 。

2. 不能。因为 $B_n C_n \neq C_n B_n$, 故 B_n, C_n 不可交换, 不能使用二项式展开式。

6.6 线性空间

定义 6.36 (线性空间及其子空间)

设 $\mathbb{F}$ 是数域, V 是一个非空集合, V 上定义了加法与数乘两种运算, 且对任意 $u, v, w \in V, \alpha, \beta \in \mathbb{F}$, 满足以下八条公理:

1. 交换律: $u + v = v + u$;
2. 结合律: $(u + v) + w = u + (v + w)$;
3. 存在零元: $\exists 0 \in V$, 使得 $\forall v \in V, 0 + v = v$;
4. 存在逆元: $\forall v \in V, \exists -v \in V$, 使得 $v + (-v) = 0$;
5. 数乘结合律: $\alpha(\beta v) = (\alpha\beta)v$;
6. 数乘分配律 1: $(\alpha + \beta)v = \alpha v + \beta v$;
7. 数乘分配律 2: $\alpha(u + v) = \alpha u + \alpha v$;

8. 数乘单位元: $1v = v$ 。

则称 $(V, +, \cdot)$ 为数域 $\mathbb{F}$ 上的线性空间 (向量空间), 其中的元素称为向量。若 $U \subset V$, 且 $\forall x, y \in U, \lambda, \mu \in \mathbb{F}, \lambda x + \mu y \in U$, 则称 U 为线性空间 V 的子空间。

定义 6.37 (线性组合)

设 V 是数域 $\mathbb{F}$ 上的线性空间, $v_1, v_2, \dots, v_m \in V, \alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{F}$, 则向量 $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m$ 称为向量 $v_1, v_2, \dots, v_m$ 的线性组合, 称 $\alpha_1, \alpha_2, \dots, \alpha_m$ 为该线性组合的系数。 $\text{span}(v_1, v_2, \dots, v_m) = \mathbb{F}\langle v_1, v_2, \dots, v_m \rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m | \alpha_i \in \mathbb{F}, i = 1, 2, \dots, m\}$ 称为由向量 $v_1, v_2, \dots, v_m$ 所生成的子空间。

定义 6.38 (向量组的等价)

设 V 是数域 $\mathbb{F}$ 上的线性空间, 向量组 $S_1 = \{v_1, v_2, \dots, v_m\}, S_2 = \{u_1, u_2, \dots, u_n\} \subset V$, 若 S_1 中的每个向量都能由 S_2 线性表示, 称为 S_1 能由 S_2 线性表示; 如果 S_1 能由 S_2 线性表示, 且 S_2 能由 S_1 线性表示, 则称向量组 S_1 与向量组 S_2 等价。

命题 6.39

向量组 S_1 能由向量组 S_2 线性表示的充分必要条件是 $\text{span}(S_1) \subset \text{span}(S_2)$ 。向量组 S_1 与向量组 S_2 等价的充分必要条件是 $\text{span}(S_1) = \text{span}(S_2)$ 。向量组的等价关系满足自反性、对称性与传递性, 是一种等价关系。

定义 6.40 (线性相关与线性无关)

设 V 是数域 $\mathbb{F}$ 上的线性空间, 向量组 $S = \{v_1, v_2, \dots, v_m\} \subset V$, 若存在不全为零的 $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{F}$, 使得 $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0$, 则称向量组 S 线性相关; 否则称向量组 S 线性无关。向量组 S 线性相关

当且仅当方程
$$\begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = 0$$
 有非零解。

命题 6.41

若对向量组 S, T 有 $S \subset T$, 则 T 线性无关时 S 线性无关, S 线性相关时 T 线性相关。

若向量组 T 中向量由 S 中向量加长得到, 则 S 线性无关时 T 线性无关, T 线性相关时 S 线性相关。

定义 6.42 (极大无关组)

设 S 是一个向量组, S_1 是 S 的一个子集, 且 S_1 线性无关, 若 S 中任一向量加入 S_1 后所得的新向量组均线性相关, 则称 S_1 为 S 的一个极大无关组。一个向量组的极大无关组不唯一, 但其向量个数相同。一个向量组与它的极大无关组等价, 一个向量组的任意两个极大无关组等价。

命题 6.43

1. 设线性无关的向量组 $\{a_1, a_2, \dots, a_s\}$ 可以由线性无关的向量组 $\{b_1, b_2, \dots, b_t\}$ 线性表示, 则 $s \leq t$;
2. 设两个线性无关的向量组 $\{a_1, a_2, \dots, a_s\}, \{b_1, b_2, \dots, b_t\}$ 等价, 则 $s = t$ 。

定义 6.44 (向量组的秩)

向量组 $a_1, a_2, \dots, a_m$ 的秩定义为其极大无关组中向量的个数, 记为 $\text{rank}(a_1, a_2, \dots, a_m)$ 。

命题 6.45

1. 向量组 $a_1, a_2, \dots, a_m$ 线性无关的充分必要条件是 $\text{rank}(a_1, a_2, \dots, a_m) = m$;
2. 向量组 $a_1, a_2, \dots, a_m$ 线性相关的充分必要条件是 $\text{rank}(a_1, a_2, \dots, a_m) < m$;
3. 向量组 $a_1, a_2, \dots, a_m$ 能由 $b_1, b_2, \dots, b_n$ 线性表示时 $\text{rank}(a_1, a_2, \dots, a_m) \leq \text{rank}(b_1, b_2, \dots, b_n)$;
4. 向量组 $a_1, a_2, \dots, a_m$ 与 $b_1, b_2, \dots, b_n$ 等价时 $\text{rank}(a_1, a_2, \dots, a_m) = \text{rank}(b_1, b_2, \dots, b_n)$;
5. b 能由 $a_1, a_2, \dots, a_m$ 线性表示的充分必要条件是 $\text{rank}(a_1, a_2, \dots, a_m) = \text{rank}(a_1, a_2, \dots, a_m, b)$ 。
6. 设 $A \in \mathbb{F}^{m \times n}$, 则 $\text{rank}(A)$ 等于其列向量组的秩, 也等于其行向量组的秩。
7. 可逆矩阵的列向量组与行向量组均线性无关。
8. 设 $\text{rank } A = r$, 则 A 不为 0 的 r 阶子式所在的行(列)向量组构成矩阵 A 的行(列)向量组的一个极大无关组。

定义 6.46 (基与维数)

设 V 是数域 $\mathbb{F}$ 上的线性空间, $S = \{v_1, v_2, \dots, v_m\} \subset V$, 若 S 线性无关, 且 $V = \text{span}(S)$, 则称 S 为线性空间 V 的一个基。有限维线性空间 V 的任意两个基均含有相同个数的向量, 称该个数为线性空间 V 的维数, 记为 $\dim V$ 。若 $\dim V = n$, 则称 V 为 n 维线性空间。若 V 有基 $\alpha_1, \alpha_2, \dots, \alpha_n$, 则任意向量 $v \in V$ 都可唯一表示为 $v = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n$, 其中 $x_1, x_2, \dots, x_n \in \mathbb{F}$, 称 v 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标为 $(x_1, x_2, \dots, x_n)^\top$ 。

命题 6.47 (坐标变换)

设线性空间 V 的两个基分别为 $\alpha_1, \alpha_2, \dots, \alpha_n$ 与 $\beta_1, \beta_2, \dots, \beta_n$, 则存在唯一的可逆矩阵 $P \in \mathbb{F}^{n \times n}$, 使得 $\begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} P$, 且该矩阵 P 称为从基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 到基 $\beta_1, \beta_2, \dots, \beta_n$ 的过渡矩阵。若向量 $v \in V$ 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标为 $(x_1, x_2, \dots, x_n)^\top$, 在基 $\beta_1, \beta_2, \dots, \beta_n$ 下的坐标为

$$(y_1, y_2, \dots, y_n)^\top, \text{ 则有 } \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = P^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}。$$

命题 6.48

设 V 是数域 $\mathbb{F}$ 上的有限维线性空间, 则有

1. 设 $U \subset V$ 是 r 维子空间, 则 U 中任意 $r+1$ 个向量线性相关, 任意 r 个线性无关向量构成 U 的一个基;
2. 设 $W \subset U \subset V$, 则 $\dim W \leq \dim U \leq \dim V$, 且 $\dim W = \dim U$ 当且仅当 $W = U$ 。
3. 设 $\dim V = n$, $a_1, a_2, \dots, a_s$ 是线性无关向量, 则存在 V 中线性无关向量 $a_{s+1}, a_{s+2}, \dots, a_n$, 使得 $a_1, a_2, \dots, a_s, a_{s+1}, a_{s+2}, \dots, a_n$ 构成 V 的一个基。

例题 6.10 判断以下命题是否正确。若正确, 请证明; 若不正确, 请给出反例。

1. $W = \{A \in \mathbb{R}^{n \times n} \mid \det A = 0\}$ 是 $\mathbb{R}^{n \times n}$ 的子空间。
2. 设向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关, $\beta = \sum_{i=1}^n \lambda_i \alpha_i$ 且 $\forall i, \lambda_i \neq 0$, 则 $\forall j, \alpha_1, \dots, \alpha_{j-1}, \beta, \alpha_{j+1}, \alpha_n$ 线性无关;
3. 若 $\text{span}(\alpha_1, \dots, \alpha_s) \subset \text{span}(\beta_1, \dots, \beta_t)$, 则 $s \leq t$;
4. 设 $A, B, C \in \mathbb{R}^{n \times n}, AB = C$, 且 B 可逆, 则 C 的行向量组与 A 的行向量组等价;
5. 设 $A, B, C \in \mathbb{R}^{n \times n}, AB = C$, 且 B 可逆, 则 C 的列向量组与 A 的列向量组等价。

解

1. 错误。反例: $n=2$ 时, $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, 则 $\det A = 0, \det B = 0$, 但 $\det(A+B) = 1 \neq 0$ 。
2. 正确。设 $\alpha_1, \dots, \alpha_{j-1}, \beta, \alpha_{j+1}, \dots, \alpha_n$ 线性相关, 则存在不全为零的 $\mu_1, \dots, \mu_n$, 使得 $\sum_{i \neq j} \mu_i \alpha_i + \mu_j \beta = 0$, 即 $\sum_{i \neq j} \mu_i \alpha_i + \mu_j \left(\sum_{i=1}^n \lambda_i \alpha_i \right) = 0$, 从而 $\sum_{i \neq j} (\mu_i + \lambda_i \mu_j) \alpha_i + \mu_j \lambda_j \alpha_j = 0$, 由于 $\alpha_1, \dots, \alpha_n$ 线性无关, 故 $\mu_j = 0, \forall i \neq j, \mu_i + \lambda_i \mu_j = 0 \Rightarrow \forall i, \mu_i = 0$, 与假设矛盾。
3. 错误。反例: $\alpha_1 = \alpha_2 = (1, 0)^\top, \beta_1 = (1, 0)^\top$ 。
4. 错误。如 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 。
5. 正确。由于 C 的每一列都是 A 的列的线性组合, 其系数由 B 的对应列给出, 故 $\text{span col } C \subset \text{span col } A$, 又由 B 可逆有 $A = CB^{-1}$, 同理可得 $\text{span col } A \subset \text{span col } C$, 从而 $\text{span col } A = \text{span col } C$, 即 C 的列向量组与 A 的列向量组等价。

例题 6.11 对于正整数 $m \geq n$, 在数组空间 $\mathbb{R}^n$ 中找出 m 个向量 $a_1, a_2, \dots, a_m$, 使得任意 n 个向量都线性无关。

解 任取 m 个互异实数 $x_1, \dots, x_m$, 令 $a_i = (1, x_i, x_i^2, \dots, x_i^{n-1})^\top, i = 1, 2, \dots, m$, 则任意 n 个向量 $a_{i_1}, a_{i_2}, \dots, a_{i_n}$ 组成的矩阵的行列式为 Vandermonde 行列式, 且不为零, 故任意 n 个向量线性无关。

例题 6.12 设 $A \in \mathbb{F}^{n \times n}, A^2 = I$, 设 $W_1 = \{x \in \mathbb{F}^n | Ax = x\}, W_2 = \{x \in \mathbb{F}^n | Ax = -x\}$, 证明: W_1, W_2 是 $\mathbb{F}^n$ 的子空间且 $\mathbb{F}^n = W_1 \oplus W_2$ 。

证明 由于 $\forall x, y \in W_1, \alpha, \beta \in \mathbb{F}, A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha x + \beta y$, 故 W_1 是 $\mathbb{F}^n$ 的子空间。同理可证 W_2 是 $\mathbb{F}^n$ 的子空间。

对任意 $x \in \mathbb{F}^n$, 令 $x_1 = \frac{x + Ax}{2}, x_2 = \frac{x - Ax}{2}$, 则 $Ax_1 = x_1, Ax_2 = -x_2$, 即 $x_1 \in W_1, x_2 \in W_2$, 且 $x = x_1 + x_2$, 故 $\mathbb{F}^n = W_1 + W_2$ 。又由于 $W_1 \cap W_2 = \{0\}$, 故 $\mathbb{F}^n = W_1 \oplus W_2$ 。

6.7 线性方程组解集的结构

定理 6.49 (解存在与唯一的条件)

对于线性方程组 $Ax = b$, 其中 $A \in \mathbb{R}^{m \times n}$, 其有解的充分必要条件是 $\text{rank } A = \text{rank}(A|b)$, 有唯一解的充分必要条件是 $\text{rank } A = \text{rank}(A|b) = n$ 。

定理 6.50 (齐次线性方程组解集的结构)

对于 $A \in \mathbb{F}^{m \times n}$, 线性方程组 $Ax = 0$ 的解集是 $\mathbb{F}^n$ 的一个子空间, 且其维数为 $n - \text{rank } A$ 。这个空间称为矩阵 A 的右零空间, 记为 $\mathcal{N}(A)$ 。解空间的任一基称为基础解系。

定理 6.51 (非齐次线性方程组解集的结构)

设线性方程组 $Ax = b$ 有解 x_0 , 则其解集为 $x_0 + \mathcal{N}(A) = \{x_0 + v | v \in \mathcal{N}(A)\}$ 。

例题 6.13 判断以下命题是否正确。若正确, 请证明; 若不正确, 请给出反例。

1. 若齐次方程组 $Ax = 0$ 有非零解, 则对非齐次方程组 $Ax = b$ 有无穷多解;
2. 设矩阵 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times p}$, 则矩阵方程 $AX = B$ 有解的充分必要条件是 $\text{rank } A = \text{rank}(A, B)$
3. 若 $x_1, \dots, x_s$ 是方程组 $Ax = b$ 的解, 则 $\sum_{i=1}^s \lambda_i x_i$ 是方程组的解时 $\sum_{i=1}^s \lambda_i = 1$ 。

解

1. 错误。反例: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 。

2. 正确。矩阵方程 $AX = B$ 有解的充分必要条件是对 B 的每一列 b_i , 方程组 $Ax = b_i$ 都有解, 即 $\text{rank } A = \text{rank}(A|b_i), i = 1, 2, \dots, p$, 等价于 $\text{rank } A = \text{rank}(A, B)$ 。
3. 错误。 $b = 0$ 时显然不成立。

例题 6.14 设某个 4 元线性方程组的系数矩阵为 A , 满足 $\text{rank } A = 3$, 已知 $\alpha_1, \alpha_2, \alpha_3$ 是其三个解, 满足 $\alpha_1 = (1, -2, -3, 4)^\top, 5\alpha_2 - 2\alpha_3 = (2, 0, 2, 0)^\top$ 。

1. 证明: 这个方程组是非齐次的;
2. 求这个方程组的通解。

解

1. 若方程组为齐次方程组, 则解空间 V 满足 $\dim V = 1$, 故其中任意两个向量均线性相关, 但 $\alpha_1 \in V, 5\alpha_2 - 2\alpha_3 \in V$, 且显然 $\alpha_1, 5\alpha_2 - 2\alpha_3$ 线性无关, 矛盾。
2. 设方程组为 $Ax = b$, 则 $A\alpha_1 = b, A(5\alpha_2 - 2\alpha_3) = 3b \Rightarrow A(5\alpha_2 - 2\alpha_3 - 3\alpha_1) = 0$, 故 $5\alpha_2 - 2\alpha_3 - 3\alpha_1$ 是齐次方程组 $Ax = 0$ 的一个非零解, 且 $\dim \mathcal{N}(A) = 1$, 故齐次方程组的通解为 $k(5\alpha_2 - 2\alpha_3 - 3\alpha_1), k \in \mathbb{R}$, 非齐次方程组的通解为 $\alpha_1 + k(5\alpha_2 - 2\alpha_3 - 3\alpha_1), k \in \mathbb{R}$ 。

第七章 线性变换与矩阵的相似

7.1 矩阵的相似

与相抵类似, $\forall A \in \mathbb{C}^{n \times n}$, A 的相抵标准形是 $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, 为了给出相似的标准型和最简单的代表元, 需要使用到以下特征值的知识。

定义 7.1 (矩阵的特征值和特征向量)

对 $A \in \mathbb{F}^{n \times n}$, 若 $\exists 0 \neq x \in \mathbb{F}^n, \lambda \in \mathbb{F}, s.t. Ax = \lambda x$, 则称 λ 是 A 的特征值, x 是属于特征值 λ 的对应特征向量。

笔记 由于矩阵可以看作在特定基下的线性变换, 此时 Ax 是向量 x 在线性变换 $\mathcal{A}$ 下的像, 此时 $Ax = \lambda x$ 是指线性变换 $\mathcal{A}$ 对 x 的作用是将 x 拉伸 λ 倍。

定义 7.2 (线性变换的特征值和特征向量)

对 $\mathcal{A} \in \mathcal{L}(V, V)$, 若 $\exists 0 \neq \alpha \in V, \lambda \in \mathbb{F}, s.t. \mathcal{A}\alpha = \lambda\alpha$, 则称 λ 是 $\mathcal{A}$ 的特征值, α 是属于特征值 λ 的对应特征向量。

定理 7.3 (线性变换的特征值与其在一组基下特征值的关系)

设 $\mathcal{A} \in \mathcal{L}(V, V), \dim V = n, \mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A, A \in \mathbb{F}^{n \times n}$, 则 $\mathcal{A}\alpha = \lambda\alpha \Leftrightarrow Ax = \lambda x$, 其中 x 为 α 在基 $\{\alpha_1, \dots, \alpha_n\}$ 下的坐标。

笔记 该定理给出了一般线性变换与矩阵的特征值与特征向量的关系, 从而以下只讨论矩阵的特征值。

算法 7.4 (特征值与特征向量的计算)

对 $A \in \mathbb{C}^{n \times n}$, 可以通过以下步骤计算其特征值与特征向量:

1. 计算特征多项式 $p_A(\lambda) = \det(\lambda I - A) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \sum_{i=1}^s \lambda_i = n$, 这是关于 λ 的一个 n 次多项式, 从而有 n 个根;
2. 令 $p_A(\lambda) = 0$, 得到特征值 $\lambda_1(n_1 \text{ 重}), \lambda_2(n_2 \text{ 重}) \cdots \lambda_s(n_s \text{ 重})$, 这些是 A 的特征值;
3. 对每一个 $\lambda_i (1 \leq i \leq s)$, 求出线性方程组 $(\lambda_i I - A)x = 0$ 的非零解, 即为 λ_i 的特征向量;

笔记

1. $(\lambda_i I - A)x = 0$ 的解空间 V_i 称为特征值 λ_i 对应的特征子空间。
2. λ_i 作为 $p_A(\lambda) = 0$ 的根的重数称为 λ_i 的代数重数; $\dim V_i := m_i$ 称为 λ_i 的几何重数
3. 可以证明 $m_i \leq n_i$, 即几何重数小于等于代数重数。

命题 7.5 (特征值的性质)

- 若 λ 是 A 的特征值, x 是对应特征向量, $f(\lambda)$ 是多项式, 则 $f(\lambda)$ 是 $f(A)$ 的特征值, x 是 $f(A)$ 的对应特征向量;
- 若 λ 是 A 的特征值, x 是对应特征向量, 则 λ 是 A^T 的特征值, 但 x 不一定是对应特征向量;
- 若 λ 是 A 的特征值, 且 $A^T A = I$, 则 $|\lambda| = 1$;
- 若 λ 是实矩阵 A 的特征值, 且 $A = A^T$, 则 λ 是实数, 且存在 P 满足 $P^T P = I$, 使得 $P^{-1}AP$ 是对角

矩阵;

- 若 λ 是实矩阵 A 的特征值, 且 $A = -A^T$, 则 λ 是纯虚数或零。

定理 7.6

特征值和特征向量是相似不变量, 即若 A, B 相似, 则 A, B 有相同的特征值。

定理 7.7

若 $\lambda_1, \dots, \lambda_n$ 是 A 的全部特征值, 则 $\operatorname{tr} A = \sum_{i=1}^n \lambda_i, \det A = \prod_{i=1}^n \lambda_i$ 。

现在考虑相似标准形的一种简单情形, 即可相似于对角矩阵。一般情形, 即相似于 Jordan 标准形会在后面提及。

从定义出发, A 可对角化 $\Leftrightarrow \exists T \in GL_n(\mathbb{C}), \Lambda$ 是对角矩阵, 使得 $A = T\Lambda T^{-1}$ 。设 $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n), T = (x_1, \dots, x_n)$, 则相似等价于 $AT = \Lambda T = (\lambda_1 x_1, \dots, \lambda_n x_n)$ 。故只需要找到一组线性无关的特征向量排成 T 即可。如此我们也找到了求出相似对角化的过渡矩阵的方法。

引理 7.8

对 $A \in \mathbb{F}^{n \times n}$, A 属于不同特征值的特征向量线性无关。

定理 7.9

$A \in \mathbb{F}^{n \times n}$ 可相似对角化当且仅当 A 有 n 个线性无关的特征向量。

因此可以得到一些可对角化的条件:

命题 7.10

- 若 n 阶方阵 A 有 n 个不同的特征值, 则 A 可对角化;
- 若方阵 A 的所有特征值的代数重数等于几何重数, 则 A 可对角化;

例题 7.1 记 $\mathbb{R}_n[x] = \left\{ f(x) \mid f(x) = \sum_{i=0}^n a_i x^i, a_i \in \mathbb{R} \right\}$, 定义线性变换 $\mathcal{D}: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x], \mathcal{D}f(x) = f'(x)$, 求一组基

使得 $\mathcal{D}$ 在该基下的矩阵为 $J_n(0) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$ 。

解 设 $\{f_0, f_1, \dots, f_{n-1}\}$ 是满足题意的基, 则 $\mathcal{D}f_0 = 0, \mathcal{D}f_1 = f_0, \mathcal{D}f_2 = f_1, \dots, \mathcal{D}f_{n-1} = f_{n-2}$, 从而其关系是从前往后的逐项积分关系, 且 f_0 是常数项非零的常数多项式, 取 $f_0 = c_0$, 则 $f_1 = c_0 x + c_1, f_2 = \frac{c_0}{2} x^2 + c_1 x + c_2, \dots, f_{n-1} = \frac{c_0}{(n-1)!} x^{n-1} + \frac{c_1}{(n-2)!} x^{n-2} + \cdots + c_{n-2} x + c_{n-1}$, 其中 $c_0, c_1, \dots, c_{n-1}$ 为任意常数, 故 $\{f_0, f_1, \dots, f_{n-1}\}$ 是满足题意的基。

例题 7.2 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times m}$, 证明: AB 与 BA 的非零特征值相同。

证明

$$\lambda^n \det(\lambda I_m - AB) = \det \begin{bmatrix} \lambda I_n & O \\ A & \lambda I_m - AB \end{bmatrix} = \det \left(\begin{bmatrix} \lambda I_n & B \\ A & \lambda I_m \end{bmatrix} \begin{bmatrix} \lambda I_n & -B \\ O & \lambda I_m \end{bmatrix} \right)$$

$$= \det \left(\begin{bmatrix} \lambda I_n & -B \\ O & \lambda I_m \end{bmatrix} \begin{bmatrix} \lambda I_n & B \\ A & \lambda I_m \end{bmatrix} \right) = \det \begin{bmatrix} \lambda I_n - BA & O \\ A & \lambda I_m \end{bmatrix} = \lambda^m \det(\lambda I_n - BA)$$

故 $\lambda^n p_{AB}(\lambda) = \lambda^m p_{BA}(\lambda)$, 即 AB 与 BA 的非零特征值相同

例题 7.3 $A = \begin{bmatrix} 4 & 2 & 2 \\ 0 & 4 & 0 \\ 0 & -2 & 2 \end{bmatrix}$, 求 A^n 。

解 $p_A(\lambda) = \det(\lambda I_n - A) = (\lambda - 2)(\lambda - 4)^2$, 故 A 有单重特征值 2 和二重特征值 4。

$(2I - A)x = 0$ 有非零解 $(1, 0, -1)^\top$, $(4I - A)x = 0$ 有线性无关非零解 $(1, 0, 0)^\top$ 和 $(0, 1, -1)^\top$,

$$\text{记 } T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix}, \Lambda = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \text{ 有 } A = T\Lambda T^{-1},$$

$$\text{故 } A^n = T\Lambda^n T^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 4^n & 0 \\ 0 & 0 & 4^n \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 4^n & 4^n - 2^n & 4^n - 2^n \\ 0 & 4^n & 0 \\ 0 & 2^n - 4^n & 2^n \end{bmatrix}.$$

例题 7.4 设 f 是 $\mathbb{R}^{n \times n}$ 上的线性函数, 且 $\forall A, B \in \mathbb{R}^{n \times n}, f(AB) = f(BA)$, 证明: $\exists \lambda \in \mathbb{R}, s.t. f(A) = \lambda \operatorname{tr} A, \forall A \in \mathbb{R}^{n \times n}$ 。

证明 不妨设 $n \geq 2$, 则 $\forall i \neq j, f(E_{ij}) = f(E_{ii}E_{ij}) = f(O) = 0, f(E_{ii}) = f(E_{i1}E_{1i}) = f(E_{1i}E_{i1}) = f(E_{11})$, 故 $f(A) = f(E_{11}) \operatorname{tr} A$ 。

例题 7.5 判断以下命题是否正确。

1. 若 $AB = O$, 则 $\det(xI - A) \det(xI - B) = x^n \det(xI - A - B)$;
2. AB 的特征值总能写成 A 的特征值和 B 的特征值的乘积;
3. 若 $\operatorname{rank} A = 1$, 则 $\det(xI - A) = x^n - \operatorname{tr} A \cdot x^{n-1}$ 。

解

1. 正确。 $\det(xI - A) \det(xI - B) = \det(x^2I - x(A + B) + AB) = \det(x^2I - x(A + B)) = x^n \det(xI - A - B)$;
2. 错误。反例: $A = E_{12}, B = E_{21}$, 则 $AB = E_{11}, BA = E_{22}$, AB 的特征值为 1, 0, 而 A, B 的特征值均为 0;
3. 正确。设 $A = uv^\top, u, v \in \mathbb{F}^n$, 则 $\det(xI - A) = \det(xI - uv^\top) = x^n \det\left(I - \frac{1}{x}uv^\top\right) = x^n \left(1 - \frac{1}{x}v^\top u\right) = x^n - (v^\top u)x^{n-1} = x^n - \operatorname{tr} A \cdot x^{n-1}$ 。

7.2 一些其他结论 *

7.2.1 Caylay-Hamilton 定理

定理 7.11 (Caylay-Hamilton 定理)

对 $A \in \mathbb{F}^{n \times n}$, $p_A(A) = 0$, 其中 $p_A(\lambda) = \det(\lambda I - A)$ 。

证明 设 $\lambda I - A$ 的伴随方阵是 B , 则 B 中元素都是次数至多为 $n-1$ 的多项式, 记 $b_{ij} = \sum_{k=0}^{n-1} b_{ij}^{(k)} \lambda^k$, 记 $B_k = (b_{ij}^{(k)})_{n \times n}$,

则 $B = \sum_{k=0}^{n-1} B_k \lambda^k$, 故 $p_A(\lambda)I_n = (\lambda I - A)B = \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \cdots + \lambda(B_0 - AB_1) - AB_0$,

设 $p_A(\lambda) = \lambda^n + \sum_{k=0}^{n-1} a_k \lambda^k$, 比较系数有 $B_{n-1} = I_n, B_{n-2} - AB_{n-1} = a_{n-1}I_n, \cdots, B_0 - AB_1 = a_1I_n, -AB_0 = a_0I_n$,

以上等式乘 $A^n, A^{n-1}, \cdots, A, I_n$ 后相加, 得到 $p_A(A) = A^n + \sum_{k=0}^{n-1} a_k A^k = 0$ 。

推论 7.12

对 $A \in \mathbb{F}^{n \times n}$, A^n 可以用 $A^{n-1}, \dots, A^2, A, I$ 线性表示, 即 $I, A, A^2, \dots$ 张成空间的维数至多为 n 。

推论 7.13

Caylay-Hamilton 定理也给出了计算逆矩阵的一种方法: 若 A 可逆, 则由 $p_A(A) = 0$ 可得 $A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \dots + a_2A + a_1I)$ 。

7.2.2 Gershgorin 圆盘定理

矩阵的特征值的估计是一个重要的问题, Gershgorin 定理给出了特征值的一个上界。为了方便叙述, 我们先给出以下定义:

定义 7.14

对方阵 $A = (a_{ij})_{n \times n} \in \mathbb{F}^{n \times n}$,

- 若 $\forall 1 \leq i \leq n, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$, 则称 A 是严格行对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$, 则称 A 是行对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| > \sum_{j \neq i} |a_{ji}|$, 则称 A 是严格列对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| \geq \sum_{j \neq i} |a_{ji}|$, 则称 A 是列对角占优方阵。

定理 7.15 (Levy-Desplanques)

若 A 是行或列严格对角占优矩阵, 则 $\det A \neq 0$ 。

证明 只证明行严格对角占优矩阵的情形, 对列严格对角占优矩阵, 考虑其转置。

对 $A = (a_{ij})_{n \times n}$, 若 $\det A = 0$, 则 $Ax = 0$ 有非零解 $x = (x_1, \dots, x_n)^T$, 设 $\max\{|x_1|, \dots, |x_n|\} = |x_k| > 0$, 则 $\sum_{j=1}^n a_{kj}x_j = 0$, 即 $a_{kk}x_k = -\sum_{j \neq k} a_{kj}x_j$, 两边取模, $|a_{kk}||x_k| = |\sum_{j \neq k} a_{kj}x_j| \leq \sum_{j \neq k} |a_{kj}||x_j| \leq |x_k| \sum_{j \neq k} |a_{kj}|$, $\Rightarrow |a_{kk}| \leq \sum_{j \neq k} |a_{kj}|$, 与 A 行严格对角占优矛盾, 故 $\det A \neq 0$ 。

定理 7.16 (Gershgorin 圆盘定理)

n 阶复方阵 $A = (a_{ij})_{n \times n}$ 的特征值落在复平面上 n 个圆盘 $|z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$ 的并集内, 其中 $1 \leq i \leq n$ 。

证明 设 λ 是 A 的特征值, 则 $\det(\lambda I - A) = 0$, 故由 Levy-Desplanques 定理有 $\lambda I - A$ 不是行严格对角占优矩阵。

故 $\exists 1 \leq i \leq n$, 使得 $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$, 即特征值落在圆盘内。

7.2.3 同时对角化问题**定理 7.17**

A, B 为 n 阶方阵, 且 A, B 均可对角化, 则 $\exists P \in GL_n(\mathbb{C})$ s.t. $P^{-1}AP, P^{-1}BP$ 均为对角矩阵当且仅当 $AB = BA$ 。

证明 ($\Rightarrow$)

由于对角矩阵乘法可交换, 故 $P^{-1}APP^{-1}BP = P^{-1}BPP^{-1}AP$, 即 $P^{-1}ABP = P^{-1}BAP$ 。由 P 可逆有 $AB = BA$ 。

($\Leftarrow$)

由 A 可对角化有 $\exists Q \in GL_n(\mathbb{C})$ s.t. $Q^{-1}AQ = \begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix}$, 其中 λ_i 互不相同, $\sum_{i=1}^s n_i = n$ 。

由 $AB = BA$ 有 $Q^{-1}AQQ^{-1}BQ = Q^{-1}BQQ^{-1}AQ$ 。对 $Q^{-1}BQ$ 分块为 $\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix}$

$$\begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix} \begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix}$$

计算得到对 $i \neq j$ 均有 $B_{ij} = 0$, 即 $Q^{-1}BQ = \begin{bmatrix} B_{11} & & \\ & B_{22} & \\ & & \ddots \\ & & & B_{ss} \end{bmatrix}$ 为准对角矩阵。

由 B 可对角化有 $Q^{-1}BQ$ 可对角化, 故 B_{ii} 可对角化, 即 $\exists N_i$ 可逆 s.t. $N_i^{-1}B_{ii}N_i = \Sigma_i$ 为对角矩阵。

$$\begin{bmatrix} N_1^{-1} & & \\ & N_2^{-1} & \\ & & \ddots \\ & & & N_s^{-1} \end{bmatrix} Q^{-1}AQ \begin{bmatrix} N_1 & & \\ & N_2 & \\ & & \ddots \\ & & & N_s \end{bmatrix} = \begin{bmatrix} \lambda_1 N_1^{-1}N_1 & & \\ & \lambda_2 N_2^{-1}N_2 & \\ & & \ddots \\ & & & \lambda_s N_s^{-1}N_s \end{bmatrix} \text{ 是对角矩阵,}$$

$$\text{而 } \begin{bmatrix} N_1^{-1} & & \\ & N_2^{-1} & \\ & & \ddots \\ & & & N_s^{-1} \end{bmatrix} Q^{-1}BQ \begin{bmatrix} N_1 & & \\ & N_2 & \\ & & \ddots \\ & & & N_s \end{bmatrix} = \begin{bmatrix} \Sigma_1 & & \\ & \Sigma_2 & \\ & & \ddots \\ & & & \Sigma_s \end{bmatrix} \text{ 也是对角矩阵。}$$

记 $N = \begin{bmatrix} N_1 & & \\ & N_2 & \\ & & \ddots \\ & & & N_s \end{bmatrix}$, 取 $P = QN$ 即为所求。

例题 7.6(2020~2021 期末) 对 A, B 为 n 阶实对称方阵, 且 $AB = BA$, 证明: $\exists P \in GL_n(\mathbb{C})$ s.t. $P^{-1}AP, P^{-1}BP$ 均为对角矩阵。

解 这是定理的直接推论。注意到实对称矩阵均可对角化即可。

第八章 Jordan 标准型 *

8.1 多项式矩阵的相抵

定义 8.1 (多项式矩阵)

设 $\mathbb{F}$ 是数域, λ 是未定元, 数域 $\mathbb{F}$ 上所有关于未定元 λ 的多项式集合记为 $\mathbb{F}[\lambda]$ 。取 $m \times n$ 个多项式 $a_{ij}(\lambda) \in \mathbb{F}[\lambda], 1 \leq i \leq m, 1 \leq j \leq n$, 则

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

称为数域 $\mathbb{F}$ 上的 $m \times n$ 多项式矩阵, 简称 λ 矩阵或多项式矩阵。数域 $\mathbb{F}$ 上所有 $m \times n$ 多项式矩阵的集合记为 $(\mathbb{F}[\lambda])^{m \times n}$ 。

多项式矩阵的加法, 纯量与多项式矩阵的乘法, 多项式矩阵的乘法以及多项式矩阵的行列式, 转置, 迹的定义和通常数域 $\mathbb{F}$ 上方阵相同, 但应注意多项式矩阵的行列式, 迹是关于未定元 λ 的多项式。

定义 8.2 (多项式矩阵的秩)

设 $A(\lambda) \in (\mathbb{F}[\lambda])^{m \times n}$, $A(\lambda)$ 的秩是 $A(\lambda)$ 中非零子式的最高阶数, 仍记为 $\text{rank } A(\lambda)$ 。若 n 阶方阵 $A(\lambda)$ 的秩等于 n , 则称 $A(\lambda)$ 是满秩的。显然 $A(\lambda)$ 满秩 $\Leftrightarrow \det A(\lambda) \neq 0$ 。

定义 8.3 (多项式矩阵的逆矩阵)

对于 n 阶方阵 $A(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 若存在 n 阶方阵 $B(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 使得 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n$, 则称 $A(\lambda)$ 是可逆的, $B(\lambda)$ 是 $A(\lambda)$ 的逆矩阵, 记为 $A(\lambda)^{-1}$ 。

定理 8.4

设 $A(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 则 $A(\lambda)$ 可逆 $\Leftrightarrow \det A(\lambda) \in \mathbb{F} \setminus \{0\}$ 。

证明 $\Rightarrow$: 设 $A(\lambda)$ 可逆, 则存在 $B(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 使 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n$ 。由 $\det A(\lambda) \det B(\lambda) = \det I_n = 1$, $\det A(\lambda), \det B(\lambda) \in \mathbb{F}[\lambda], \deg(1) = 0$, 有 $\deg(A(\lambda)) = \deg(B(\lambda)) = 0$, 故 $\det A(\lambda) \in \mathbb{F} \setminus \{0\}$ 。

$\Leftarrow$: 设 $\det A(\lambda) \in \mathbb{F} \setminus \{0\}$, 设 $a = \det A(\lambda)$, 考虑 $A^*(\lambda)$ 是 $A(\lambda)$ 的伴随矩阵, 由 $A^*(\lambda)$ 中元素是 $A(\lambda)$ 的 $n-1$ 阶子式, 故 $A^*(\lambda), \frac{1}{a}A^*(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 而 $A(\lambda) \left(\frac{1}{a}A^*(\lambda) \right) = I_n = \left(\frac{1}{a}A^*(\lambda) \right) A(\lambda)$, 故 $A(\lambda)$ 可逆,

$$A(\lambda)^{-1} = \frac{1}{a}A^*(\lambda)。$$

 **笔记** 若 $A(\lambda)$ 可逆, 则 $A(\lambda)$ 满秩, 但是反之不成立。这与通常方阵不同。

定义 8.5 (多项式矩阵的初等变换, 初等 λ 矩阵)

多项式矩阵的初等 λ 变换分为以下三类:

1. 对调矩阵的两行或两列;
2. 把某一行或某一列乘以非零多项式加到另一行或另一列;
3. 以非零常数乘矩阵的某一行或某一列。

依次称为对多项式矩阵的第一、第二和第三种初等变换。

记 $P_{ij} = I_n + E_{ij} + E_{ji} - E_{ii} - E_{jj}$, $T_{ij}(f(\lambda)) = I_n + f(\lambda)E_{ij}$, $D_i(a) = I_n + (a-1)E_{ii}$, 则 $S_{ij}, T_{ij}(f(\lambda)), D_i(a)$ 被称为第一、第二、第三类初等 λ 矩阵。



分别用第一、第二、第三种初等 λ 矩阵左乘 $A(\lambda)$, 相当于对 $A(\lambda)$ 做第一、第二、第三种初等行变换。分别用第一、第二、第三种初等 λ 矩阵右乘 $A(\lambda)$, 相当于对 $A(\lambda)$ 做第一、第二、第三种初等列变换。

初等 λ 变换不改变多项式矩阵的秩。

定理 8.6

设 $m \times n$ 多项式矩阵 $A(\lambda)$ 的秩为 r , 则 $A(\lambda)$ 可以经过有限次初等 λ 变换化为 $B(\lambda) = \begin{bmatrix} D(\lambda) & O \\ O & O \end{bmatrix}$ 的形式, 其中 $D(\lambda) = \text{diag}(d_1(\lambda), \dots, d_r(\lambda))$ 是 r 阶对角多项式方阵, $d_1(\lambda), \dots, d_r(\lambda)$ 是首一多项式且 $d_j(\lambda) | d_{j+1}(\lambda), \forall j = 1, \dots, r-1$ 。

等价的, 存在 $s, t \in \mathbb{N}$, m 阶初等 λ 方阵 $P_1(\lambda), \dots, P_s(\lambda)$ 和 n 阶初等 λ 方阵 $Q_1(\lambda), \dots, Q_t(\lambda)$ 使得

$$P_s(\lambda) \cdots P_1(\lambda) A(\lambda) Q_1(\lambda) \cdots Q_t(\lambda) = \begin{bmatrix} D(\lambda) & O \\ O & O \end{bmatrix}.$$



证明 对零矩阵 $A(\lambda)$, 定理显然成立。因此可设 $A = (a_{ij}(\lambda))_{n \times n} \neq 0$ 。于是存在某个元素 $a_{ij}(\lambda) \neq 0$ 。对调矩阵第 1 行与第 i 行, 第 1 列与第 j 列, 则有 $a_{11}(\lambda) \neq 0$ 。因此可不妨假设 $a_{11}(\lambda) \neq 0$ 。

首先证明: 多项式矩阵 $A(\lambda)$ 可以经有限次初等 λ 变换化为多项式矩阵 $C(\lambda) = (c_{ij}(\lambda))_{n \times n}$, 其中 $c_{11}(\lambda) \neq 0$, 且 $c_{11}(\lambda) | c_{ij}(\lambda), \forall 1 \leq i \leq m, 1 \leq j \leq n$ 。为此对次数 $\deg a_{11}(\lambda)$ 归纳。

当 $\deg a_{11}(\lambda) = 0$ 时, $a_{11}(\lambda)$ 是 $\mathbb{F}$ 中非零常数, 显然 $a_{11}(\lambda) | a_{ij}(\lambda), \forall 1 \leq i \leq m, 1 \leq j \leq n$ 。因此结论对 $\deg a_{11}(\lambda) = 0$ 成立。假设结论对 $\deg a_{11}(\lambda) \leq t-1$ 成立, 现在考虑 $\deg a_{11}(\lambda) = t$ 的情况。假设 $\exists a_{ij}(\lambda)$ 使得 $a_{11}(\lambda) \nmid a_{ij}(\lambda)$ 。

1. $i = 1$, 即 $a_{11}(\lambda) \nmid a_{1j}(\lambda)$ 。由带余除法, $\exists q(\lambda), r(\lambda) \in \mathbb{F}[\lambda]$, 使得 $a_{1j}(\lambda) = q(\lambda)a_{11}(\lambda) + r(\lambda), \deg r(\lambda) < t$ 。把第 1 列的 $-q(\lambda)$ 倍加到第 j 列, 再对调第 1 列与第 j 列, 这个矩阵 (1,1) 位置上的元素就是 $r(\lambda)$, 且 $\deg r(\lambda) < t$, 使用归纳假设, 有结论成立。
2. $j = 1$, 将上一情形的列改为行即可。
3. $a_{11}(\lambda)$ 整除每一个 a_{1j} 与 a_{i1} , 且 $a_{ij}(\lambda) \nmid a_{kl}(\lambda), k \neq 1, l \neq 1$ 。此时设 $a_{1j}(\lambda) = q_j(\lambda)a_{11}(\lambda), a_{i1}(\lambda) = p_i(\lambda)a_{11}(\lambda), p_i(\lambda), q_j(\lambda) \in \mathbb{F}[\lambda]$, 把第 1 行的 $-p_i(\lambda)$ 倍加到第 i 行, 再把第 1 列的 $-q_j(\lambda)$ 倍加到第 j 列, 得到多项式矩阵:

$$\begin{bmatrix} a_{11}(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & \tilde{a}_{22}(\lambda) & \tilde{a}_{23}(\lambda) & \cdots & \tilde{a}_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \tilde{a}_{m2}(\lambda) & \tilde{a}_{m3}(\lambda) & \cdots & \tilde{a}_{mn}(\lambda) \end{bmatrix}$$

此时如果 $a_{11}(\lambda)$ 整除每个 $\tilde{a}_{ij}(\lambda)$, 则结论成立。否则, 设 $a_{11}(\lambda) \nmid \tilde{a}_{kl}(\lambda)$, 把这个多项式矩阵的第 m 行加到第 1 行, 则化为情形 1, 所以结论成立。

由上述讨论可知, 多项式矩阵 $A(\lambda)$ 可以经有限次初等 λ 变换化为多项式矩阵 $C(\lambda) = (c_{ij}(\lambda))_{n \times n}$, 其中 $c_{11}(\lambda) \neq 0$, 且 $c_{11}(\lambda) | c_{ij}(\lambda), \forall 1 \leq i \leq m, 1 \leq j \leq n$ 。设 $c_{i1}(\lambda) = f_i(\lambda)c_{11}(\lambda), c_{1j}(\lambda) = g_j(\lambda)c_{11}(\lambda)$, 把第 1 行的

$-f_i(\lambda)$ 倍加到第 i 行, 再把第 1 列的 $-g_j(\lambda)$ 倍加到第 j 列, 得到多项式矩阵:

$$\begin{bmatrix} c_{11}(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & b_{22}(\lambda) & b_{23}(\lambda) & \cdots & b_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{m2}(\lambda) & b_{m3}(\lambda) & \cdots & b_{mn}(\lambda) \end{bmatrix} \quad (8.1)$$

此时仍有 $c_{11} \mid b_{ij}(\lambda)$ 。记 $A_1(\lambda) = (b_{kl}(\lambda))_{(m-1) \times (n-1)}, 2 \leq k \leq m, 2 \leq l \leq n$ 。对多项式矩阵 $A_1(\lambda)$ 重复上述过程, 则多项式矩阵 $A_1(\lambda)$ 可经有限次初等 λ 变换化为

$$\begin{bmatrix} c_{22}(\lambda) & 0 \\ 0 & A_2(\lambda) \end{bmatrix}$$

其中 $c_{22}(\lambda)$ 整除 $A_2(\lambda)$ 的每个元素。于是原来的矩阵可化为

$$\begin{bmatrix} c_{11}(\lambda) & 0 & 0 \\ 0 & c_{22}(\lambda) & 0 \\ 0 & 0 & A_2(\lambda) \end{bmatrix} \quad (8.2)$$

由于对 $A_1(\lambda)$ 施加初等 λ 不改变其元素被 $c_{11}(\lambda)$ 整除的性质, 故 $c_{11}(\lambda)$ 整除 $c_{22}(\lambda)$ 及 $A_2(\lambda)$ 的所有元素。且 $c_{22}(\lambda)$ 整除 $A_2(\lambda)$ 的所有元素。

重复上述过程, 多项式矩阵 $A(\lambda)$ 可经有限次初等 λ 变换化为以下形式:

$$\begin{bmatrix} \begin{bmatrix} c_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & c_{22}(\lambda) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_{kk}(\lambda) \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix}$$

其中非零多项式 $c_{11}(\lambda), c_{22}(\lambda), \dots, c_{kk}(\lambda)$ 满足 $c_{ii}(\lambda) \mid c_{i+1, i+1}(\lambda), \forall 1 \leq i \leq k-1$ 。由于初等 λ 变换不改变多项式矩阵的秩, 故 $k = \text{rank}(A(\lambda)) = r$ 。设 c_{ii} 的首项系数为 b_i , 则把以上矩阵的第 i 行乘 b_i^{-1} , 就可以把多项式矩阵 $A(\lambda)$ 化为 $B(\lambda)$, 因此定理得证。

定理 8.7

设 $A(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 则 $A(\lambda)$ 可逆 $\Leftrightarrow A(\lambda)$ 可表示为有限个初等 λ 矩阵的乘积。



证明 由定理 8.6 易得。

定义 8.8 (多项式矩阵的相抵)

设 $A(\lambda), B(\lambda) \in (\mathbb{F}[\lambda])^{m \times n}$, 若 $A(\lambda)$ 可经有限次初等变换化为 $B(\lambda)$, 即 $\exists P(\lambda) \in (\mathbb{F}[\lambda])^{m \times m}, Q(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$ 可逆, 使得 $P(\lambda)A(\lambda)Q(\lambda) = B(\lambda)$, 则称 $A(\lambda)$ 与 $B(\lambda)$ 相抵。



容易验证, 多项式矩阵的相抵关系是等价关系。于是, 根据多项式矩阵的相抵关系可以把数域 $\mathbb{F}$ 上所有 $m \times n$ 多项式矩阵分类: 彼此相抵的多项式矩阵属于同一类, 不相抵的多项式矩阵属于不同的类。和一般数域上的相抵一样, 两个基本问题是: 多项式矩阵在相抵下的标准型是什么? 多项式矩阵在相抵下的全系不变量是什么?

首先, 由 $A(\lambda) = I_2, B(\lambda) = \text{diag}(1, \lambda)$ 不相抵可以得出多项式矩阵的秩不足以构成多项式矩阵的全系不变量, 所以必须寻找多项式矩阵在相抵下的其他不变量。

定义 8.9

设 $A(\lambda) \in (\mathbb{F}[\lambda])^{m \times n}$, 多项式矩阵 $A(\lambda)$ 中所有 k 阶非零子式的最大公因子称为 $A(\lambda)$ 的 k 阶行列式因子, 记为 $D_k(\lambda)$ 。若 $A(\lambda)$ 的所有 k 阶子式均为 0, 则约定 $D_k(\lambda) = 0$ 。

容易验证, 若 $\text{rank } A(\lambda) = r$, 记 $\min\{m, n\} = s$, 则 $D_{r+1}(\lambda) = \cdots = D_s(\lambda) = 0$, 对于 $1 \leq k \leq r$, $D_k(\lambda)$ 是非零多项式, 且 $D_k(\lambda) | D_{k+1}(\lambda), \forall k = 1, \cdots, r-1$ 。

我们先回忆一下 Binet-Cauchy 公式。

引理 8.10 (Binet-Cauchy 公式)

对 $A \in \mathbb{F}^{p \times q}, B \in \mathbb{F}^{q \times p}$, 有

$$\det(AB) = \begin{cases} 0, & q < p \\ \det A \cdot \det B, & q = p \\ \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq q} A \begin{pmatrix} 1 & 2 & \cdots & p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} B \begin{pmatrix} j_1 & j_2 & \cdots & j_p \\ 1 & 2 & \cdots & p \end{pmatrix}, & q > p \end{cases}$$

证明 只考虑 $p \leq q$ 的情形。

$$\begin{aligned} \det(AB) &= \det \begin{bmatrix} \sum_{i_1=1}^q a_{1i_1} b_{i_1 1} & \cdots & \sum_{i_p=1}^q a_{1i_p} b_{i_p p} \\ \vdots & & \vdots \\ \sum_{i_1=1}^q a_{pi_1} b_{i_1 1} & \cdots & \sum_{i_p=1}^q a_{pi_p} b_{i_p p} \end{bmatrix} = \sum_{i_1=1}^q \cdots \sum_{i_p=1}^q \det \begin{bmatrix} a_{1i_1} b_{i_1 1} & \cdots & a_{1i_p} b_{i_p p} \\ \vdots & & \vdots \\ a_{pi_1} b_{i_1 1} & \cdots & a_{pi_p} b_{i_p p} \end{bmatrix} \\ &= \sum_{i_1=1}^q \cdots \sum_{i_p=1}^q \det \begin{bmatrix} a_{1i_1} & \cdots & a_{1i_p} \\ \vdots & & \vdots \\ a_{pi_1} & \cdots & a_{pi_p} \end{bmatrix} \prod_{j=1}^p b_{i_j j} \\ &\stackrel{(*)}{=} \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq q} \sum_{j_1, j_2, \cdots, j_p} (-1)^{\tau(j_1, j_2, \cdots, j_p)} \det \begin{bmatrix} a_{1i_1} & \cdots & a_{1i_p} \\ \vdots & & \vdots \\ a_{pi_1} & \cdots & a_{pi_p} \end{bmatrix} \prod_{k=1}^p b_{k j_k} \\ &= \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq q} A \begin{pmatrix} 1 & 2 & \cdots & p \\ i_1 & i_2 & \cdots & i_p \end{pmatrix} B \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ 1 & 2 & \cdots & p \end{pmatrix} \end{aligned}$$

其中 (*) 处取等是因为若存在某个 $i_k = i_l$, 则对应项中行列式部分为 0, 故可以把所有为 0 的项去除然后进行排序, 根据行列式的性质即得。

命题 8.11

对 $A \in \mathbb{F}^{p \times q}, B \in \mathbb{F}^{q \times p}$, 记 $C = AB$, 则对 C 的 r 阶子式有

$$C \begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \begin{cases} 0, & q < p \\ \sum_{1 \leq k_1 < k_2 < \cdots < k_r \leq q} A \begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}, & q \geq p \end{cases}$$

证明 对 $A_1 = \begin{bmatrix} a_{i_1 1} & a_{i_1 2} & \cdots & a_{i_1 q} \\ a_{i_2 1} & a_{i_2 2} & \cdots & a_{i_2 q} \\ \vdots & \vdots & & \vdots \\ a_{i_r 1} & a_{i_r 2} & \cdots & a_{i_r q} \end{bmatrix}, B_1 = \begin{bmatrix} b_{1 j_1} & b_{1 j_2} & \cdots & b_{1 j_r} \\ b_{2 j_1} & b_{2 j_2} & \cdots & b_{2 j_r} \\ \vdots & \vdots & & \vdots \\ b_{q j_1} & b_{q j_2} & \cdots & b_{q j_r} \end{bmatrix}$ 使用引理8.10即可。

定理 8.12

设 $A(\lambda), B(\lambda) \in (\mathbb{F}[\lambda])^{m \times n}$, 则 $A(\lambda), B(\lambda)$ 相抵 $\Leftrightarrow A(\lambda), B(\lambda)$ 的行列式因子相同。换言之, 多项式矩阵的行列式因子是其在相抵下的全系不变量。



证明 以下设 $A(\lambda), B(\lambda)$ 的 k 阶行列式因子分别为 $D_k(\lambda), \tilde{D}_k(\lambda)$ 。

(必要性) 设多项式矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 则存在可逆多项式矩阵 $P(\lambda) \in (\mathbb{F}[\lambda])^{m \times m}, Q(\lambda) \in (\mathbb{F}[\lambda])^{n \times n}$, 使得 $P(\lambda)A(\lambda)Q(\lambda) = B(\lambda)$ 。由命题8.11有:

$$\begin{aligned} B(\lambda) \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} &= \sum_{1 \leq l_1 < l_2 < \cdots < l_k \leq m} P(\lambda) \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ l_1 & l_2 & \cdots & l_k \end{pmatrix} (A(\lambda)Q(\lambda)) \begin{pmatrix} l_1 & l_2 & \cdots & l_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \\ &= \sum_{1 \leq l_1 < l_2 < \cdots < l_k \leq m} \sum_{1 \leq t_1 < t_2 < \cdots < t_k \leq n} P(\lambda) \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ l_1 & l_2 & \cdots & l_k \end{pmatrix} A(\lambda) \begin{pmatrix} l_1 & l_2 & \cdots & l_k \\ t_1 & t_2 & \cdots & t_k \end{pmatrix} Q(\lambda) \begin{pmatrix} t_1 & t_2 & \cdots & t_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \end{aligned}$$

若 $A(\lambda)$ 的 k 阶行列式因子 $D_k(\lambda) = 0$, 则 $A(\lambda)$ 的每个 k 阶子式为 0, 故 $B(\lambda)$ 的每个 k 阶子式为 0, 即 $B(\lambda)$ 的 k 阶行列式因子 $\tilde{D}_k(\lambda) = 0$ 。反之亦然。

若 $D_k(\lambda) \neq 0$, 则 $D_k(\lambda)$ 整除每个 $A(\lambda)$ 的 k 阶子式, 故 $D_k(\lambda)$ 整除每个 $B(\lambda)$ 的 k 阶子式, 即 $D_k(\lambda)$ 整除 $B(\lambda)$ 的每个 k 阶子式, 故 $\tilde{D}_k(\lambda) | D_k(\lambda)$ 。同理可证 $D_k(\lambda) | \tilde{D}_k(\lambda)$, 又作为最大公因子, $D_k(\lambda), \tilde{D}_k(\lambda)$ 均为首一多项式, 故 $D_k(\lambda) = \tilde{D}_k(\lambda)$ 。

这说明 $A(\lambda), B(\lambda)$ 的行列式因子相同。

(充分性) 若 $A(\lambda), B(\lambda)$ 的行列式因子相同, 并设 $\text{rank } A(\lambda) = r$, 则由定理8.6, $A(\lambda)$ 相抵于如下多项式矩阵:

$$C(\lambda) = \begin{bmatrix} D(\lambda) & O \\ O & O \end{bmatrix} = \begin{bmatrix} \text{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)) & O \\ O & O \end{bmatrix}$$

多项式矩阵 $C(\lambda)$ 的 k 阶行列式因子同样为 $D_k(\lambda)$, 易得 $D_1(\lambda) = d_1(\lambda), \cdots, D_r(\lambda) = \prod_{j=1}^r d_j(\lambda), D_r(\lambda) = \cdots =$

$$D_{\min\{m,n\}}(\lambda) = 0 \Rightarrow d_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} \text{ 由 } D_k(\lambda) \text{ 唯一确定 } (D_0(\lambda) = 1)。$$

由于 $A(\lambda), B(\lambda)$ 的行列式因子相同, 有 $B(\lambda), C(\lambda)$ 相抵, 由传递性, $A(\lambda), B(\lambda)$ 相抵。

定义 8.13 (不变因子)

设多项式矩阵 $A(\lambda) \in (\mathbb{F})^{m \times n}$ 的秩为 r , k 阶行列式因子为 $D_k(\lambda)$, 约定 $D_0(\lambda) = 1$, 定义 $d_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}$ 为 $A(\lambda)$ 的不变因子。

**定理 8.14**

设多项式矩阵 $A(\lambda) \in (\mathbb{F})^{m \times n}$ 的秩为 r , 阶不变因子为 $d_1(\lambda), \cdots, d_r(\lambda)$, 则 $A(\lambda)$ 相抵于 Smith 标准型

$$\begin{bmatrix} \text{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)) & O \\ O & O \end{bmatrix}, \text{ 且 } A(\lambda) \text{ 的不变因子是其在相抵下的全系不变量。}$$



证明 由定理8.12的证明易得。

定理8.12和定理8.14解决了多项式矩阵在相抵下的标准型和全系不变量的问题,下面将给出 $\mathbb{C}$ 上多项式矩阵在相抵下的另一种全系不变量。

定义 8.15 (初等因子, 初等因子组)

多项式矩阵 $A(\lambda) \in (\mathbb{F})^{m \times n}$ 的秩为 r , 阶不变因子为 $d_1(\lambda), \dots, d_r(\lambda)$, 由 $\mathbb{C}$ 是代数闭域, 故 $d_k(\lambda)$ 可分解为一次因式的乘积。由此可设

$$d_1(\lambda) = (\lambda - \lambda_1)^{e_{11}} (\lambda - \lambda_2)^{e_{12}} \cdots (\lambda - \lambda_t)^{e_{1t}}$$

$$d_2(\lambda) = (\lambda - \lambda_1)^{e_{21}} (\lambda - \lambda_2)^{e_{22}} \cdots (\lambda - \lambda_t)^{e_{2t}}$$

.....

$$d_r(\lambda) = (\lambda - \lambda_1)^{e_{r1}} (\lambda - \lambda_2)^{e_{r2}} \cdots (\lambda - \lambda_t)^{e_{rt}}$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_t$ 是两两不同的复数, $e_{ij} \in \mathbb{N}, i = 1, 2, \dots, r, j = 1, 2, \dots, t$, 由 $d_k(\lambda) | d_{k+1}(\lambda)$, 有 $0 \leq e_{1j} \leq e_{2j} \leq \cdots \leq e_{rj}, j = 1, 2, \dots, t$ 。当 $e_{kl} > 0$ 时, 因子 $(\lambda - \lambda_k)^{e_{kl}}$ 被称为矩阵 $A(\lambda)$ 属于 λ_k 的初等因子, $A(\lambda)$ 的初等因子的全体称为 $A(\lambda)$ 的初等因子组。

定理 8.16

复数域上的多项式矩阵的秩和初等因子组是其在相抵下的全系不变量。

证明 由定义可得不变因子可以唯一确定初等因子组和秩。

反之, 设秩为 r 的 $A(\lambda)$ 的初等因子组为 $\{(\lambda - \lambda_p)^{e_{pq}}\}_{1 \leq p \leq t}$, 将其下标按 p 分组并按 e_{pq} 递增排列, 并在最前面补充足够的 1 使得每组均有 r 个元素, 得到增广因子组 $\{(\lambda - \lambda_p)^{e_{pq}}\}_{1 \leq p \leq t, 1 \leq q \leq r}$, 由 $d_k(\lambda) | d_{k+1}(\lambda)$, 有 $d_k(\lambda) = \prod_{p=1}^t (\lambda - \lambda_p)^{e_{pk}}, 1 \leq k \leq r$ (思考原因)。如此初等因子组可以唯一确定不变因子。

由上, 初等因子组与秩、不变因子是等价的, 故秩和初等因子组是多项式矩阵在相抵下的全系不变量。

 **笔记** 该定理中秩的条件是不可以删去的。

8.2 Jordan 标准型的概念与求法

下面的定理给出方阵的相似与多项式矩阵相抵之间的重要联系。

定理 8.17

n 阶复方阵 A, B 相似 $\Leftrightarrow$ 多项式矩阵 $\lambda I_n - A, \lambda I_n - B$ 相抵。

证明

若方阵 A, B 相似, 则存在可逆方阵 P 使得 $B = P^{-1}AP$ 。此时, P 作为可逆多项式矩阵有 $\lambda I_n - B = P^{-1}(\lambda I_n - A)P$, 故 $\lambda I_n - A, \lambda I_n - B$ 相抵。

反之, 若 $\lambda I_n - A, \lambda I_n - B$ 相抵, 则存在可逆多项式矩阵 $P(\lambda), Q(\lambda)$ 使得 $P(\lambda)(\lambda I_n - A)Q(\lambda) = \lambda I_n - B$ 。设 $Q(\lambda) = \sum_{j=0}^k \lambda^j Q_j, Q(\lambda)^{-1} = \sum_{j=0}^m \lambda^j R_j$, 设 $W = Q(B) = \sum_{j=0}^k Q_j B^j$ 。由 $Q(\lambda)^{-1}Q(\lambda) = I_n$, 有 $\sum_{j=0}^m R_j Q(\lambda) \lambda^j = I_n \Rightarrow \sum_{j=0}^m R_j W B^j = I_n$ 。又 $P(\lambda)(\lambda I_n - A)Q(\lambda) = \lambda I_n - B$, 有 $P(\lambda)^{-1}(\lambda I_n - B) = (\lambda I_n - A)Q(\lambda) = Q(\lambda)\lambda - AQ(\lambda)$ 。用 B 代替 λ , 有 $Q(B)B = AQ(B)$, 即 $WB = AW$ 。由此得到 $WB^l = A^l W, \forall l \in \mathbb{N}$ 。于是有 $I_n = \sum_{j=0}^m R_j W B^j = \sum_{j=0}^m R_j A^j W = (\sum_{j=0}^m R_j A^j)W$ 。这表明 W 是可逆矩阵且 $B = W^{-1}AW$ 。

 **笔记** 对 n 阶复方阵 A , $\lambda I_n - A$ 被称为 A 的特征方阵。显然特征方阵是满秩的。

由定理8.16和定理8.17, 可以得到:

定理 8.18

n 阶复方阵 A 与 B 相似的充要条件是它们的特征方阵 $\lambda I_n - A, \lambda I_n - B$ 的初等因子组相同。也就是说, 复方阵的特征方阵的初等因子组是复方阵在相似下的全系不变量。

在介绍复方阵的 Jordan 标准型的概念之前, 还需要一个引理。

引理 8.19

准对角多项式矩阵 $A(\lambda) = \text{diag}(A_1(\lambda), A_2(\lambda))$ 的初等因子组由对角块 $A_1(\lambda), A_2(\lambda)$ 的初等因子组合并而成。

证明 设 $d_1(\lambda), \dots, d_r(\lambda)$ 和 $d_{r+1}(\lambda), \dots, d_{r+s}(\lambda)$ 分别是多项式矩阵 $A_1(\lambda)$ 和 $A_2(\lambda)$ 的不变因子, 则存在可逆多项式方阵 $P_1(\lambda), P_2(\lambda), Q_1(\lambda), Q_2(\lambda)$, 使得:

$$P_1(\lambda)A_1(\lambda)Q_1(\lambda) = \begin{bmatrix} \text{diag}(d_1(\lambda), \dots, d_r(\lambda)) & 0 \\ 0 & 0 \end{bmatrix}, P_2(\lambda)A_2(\lambda)Q_2(\lambda) = \begin{bmatrix} \text{diag}(d_{r+1}(\lambda), \dots, d_{r+s}(\lambda)) & 0 \\ 0 & 0 \end{bmatrix}$$

记 $P(\lambda) = \text{diag}(P_1(\lambda), P_2(\lambda)), Q(\lambda) = \text{diag}(Q_1(\lambda), Q_2(\lambda))$, 则 $P(\lambda), Q(\lambda)$ 可逆,

且 $P(\lambda)A(\lambda)Q(\lambda) = \text{diag}(D_1, 0, D_2, 0)$ 。其中 $D_1 = \text{diag}(d_1(\lambda), \dots, d_r(\lambda)), D_2 = \text{diag}(d_{r+1}(\lambda), \dots, d_{r+s}(\lambda))$ 。记 $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$, 则 $A(\lambda), B(\lambda)$ 相抵, 因此它们的初等因子组相同。

接下来求多项式矩阵 $B(\lambda)$ 的初等因子组。设

$$d_1(\lambda) = (\lambda - \lambda_1)^{e_{11}}(\lambda - \lambda_2)^{e_{12}} \cdots (\lambda - \lambda_k)^{e_{1k}}, \dots, d_r(\lambda) = (\lambda - \lambda_1)^{e_{r1}}(\lambda - \lambda_2)^{e_{r2}} \cdots (\lambda - \lambda_k)^{e_{rk}}$$

$$d_{r+1}(\lambda) = (\lambda - \lambda_1)^{e_{r+1,1}}(\lambda - \lambda_2)^{e_{r+1,2}} \cdots (\lambda - \lambda_k)^{e_{r+1,k}}, \dots, d_{r+s}(\lambda) = (\lambda - \lambda_1)^{e_{r+s,1}}(\lambda - \lambda_2)^{e_{r+s,2}} \cdots (\lambda - \lambda_k)^{e_{r+s,k}}$$

其中 $e_{ij} \in \mathbb{N}, \lambda_1, \dots, \lambda_k$ 互不相同。由于不变因子依次一个整除另一个, 有 $0 \leq e_{1j} \leq e_{2j} \leq \cdots \leq e_{rj}, 0 \leq e_{r+1,j} \leq e_{r+2,j} \leq \cdots \leq e_{r+s,j}, j = 1, 2, \dots, k$ 。

显然, 多项式矩阵 $A_1(\lambda)$ 的初等因子组为 $(\lambda - \lambda_j)^{e_{ij}} (1 \leq i \leq r)$, 而多项式矩阵 $A_2(\lambda)$ 的初等因子组为 $(\lambda - \lambda_j)^{e_{r+i,j}} (1 \leq i \leq s)$ 。

对于每个 j , 将指数 $e_{1j}, \dots, e_{rj}, e_{r+1,j}, \dots, e_{r+s,j}$ 重排为 $\tilde{e}_{1j}, \tilde{e}_{2j}, \dots, \tilde{e}_{r+s,j}$, 使得 $0 \leq \tilde{e}_{1j} \leq \tilde{e}_{2j} \leq \cdots \leq \tilde{e}_{r+s,j}$ 。容易看出, 多项式矩阵 $B(\lambda)$ 的 $r+s$ 阶行列式因子 $D_{r+s}(\lambda)$ 为:

$$D_{r+s}(\lambda) = d_1(\lambda) \cdots d_r(\lambda) d_{r+1}(\lambda) \cdots d_{r+s}(\lambda) = (\lambda - \lambda_1)^{\tilde{e}_{11} + \cdots + \tilde{e}_{r+s,1}} \cdots (\lambda - \lambda_k)^{\tilde{e}_{1k} + \cdots + \tilde{e}_{r+s,k}}$$

$$B(\lambda) \text{ 的所有 } r+s-1 \text{ 阶子式为 } \frac{D_{r+s}(\lambda)}{d_i(\lambda)}, i = 1, 2, \dots, r+s \text{ 且}$$

$$\frac{D_{r+s}(\lambda)}{d_i(\lambda)} = (\lambda - \lambda_1)^{\tilde{e}_{11} + \cdots + \tilde{e}_{r+s,1} - \tilde{e}_{i1}} \cdots (\lambda - \lambda_k)^{\tilde{e}_{1k} + \cdots + \tilde{e}_{r+s,k} - \tilde{e}_{ik}}$$

由于 $\tilde{e}_{r+s,j} \geq e_{ij}$, 因此 $\tilde{e}_{1j} + \cdots + \tilde{e}_{r+s,j} - \tilde{e}_{ij} \geq \tilde{e}_{1j} + \cdots + \tilde{e}_{r+s-1,j}$, 取等当且仅当 $\tilde{e}_{r+s,j} = e_{ij}$ 。

所以多项式 $\frac{D_{r+s}(\lambda)}{d_1(\lambda)}, \dots, \frac{D_{r+s}(\lambda)}{d_{r+s}(\lambda)}$ 的最大公因子, 即 $B(\lambda)$ 的 $r+s-1$ 阶行列式因子 $D_{r+s-1}(\lambda)$ 与 $B(\lambda)$ 的第 $r+s$ 个不变因子分别为:

$$D_{r+s-1}(\lambda) = (\lambda - \lambda_1)^{\tilde{e}_{11} + \cdots + \tilde{e}_{r+s-1,1}} \cdots (\lambda - \lambda_k)^{\tilde{e}_{1k} + \cdots + \tilde{e}_{r+s-1,k}}, \tilde{d}_{r+s}(\lambda) = (\lambda - \lambda_1)^{\tilde{e}_{r+s,1}} \cdots (\lambda - \lambda_k)^{\tilde{e}_{r+s,k}}$$

同理可得 $B(\lambda)$ 的第 l 个不变因子 $\tilde{d}_l(\lambda) = (\lambda - \lambda_1)^{\tilde{e}_{l1}} \cdots (\lambda - \lambda_k)^{\tilde{e}_{lk}}$

因此 $B(\lambda)$ 的初等因子组为 $\{(\lambda - \lambda_j)^{\tilde{e}_{lj}} \mid \tilde{e}_{lj} > 0, 1 \leq l \leq r+s, 1 \leq j \leq k\}$ 。

由于 $\tilde{e}_{1j}, \tilde{e}_{2j}, \dots, \tilde{e}_{r+s,j}$ 是 $e_{1j}, \dots, e_{rj}, e_{r+1,j}, \dots, e_{r+s,j}$ 的重排, 故多项式矩阵 $B(\lambda)$ 的初等因子组由 $A_1(\lambda), A_2(\lambda)$ 的初等因子组合并而成。

以下给出复方阵的 Jordan 标准型的概念。

定义 8.20 (Jordan 块, Jordan 标准型)

记 $N_m = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}_{m \times m}$, 则形如 $\lambda I_m + N_m$ 的矩阵被称为一个 m 阶 Jordan 块, 记作 $J_m(\lambda)$ 。Jordan 标准型是由对角线由若干个 Jordan 块组成的准对角矩阵。

定理 8.21

设 n 阶复方阵 A 的特征方阵 $\lambda I_n - A$ 的初等因子组为

$$\begin{aligned} &(\lambda - \lambda_1)^{m_{11}}, (\lambda - \lambda_1)^{m_{12}}, \dots, (\lambda - \lambda_1)^{m_{1k_1}}, \\ &(\lambda - \lambda_2)^{m_{21}}, (\lambda - \lambda_2)^{m_{22}}, \dots, (\lambda - \lambda_2)^{m_{2k_2}}, \\ &\dots\dots\dots \\ &(\lambda - \lambda_t)^{m_{t1}}, (\lambda - \lambda_t)^{m_{t2}}, \dots, (\lambda - \lambda_t)^{m_{tk_t}}, \end{aligned}$$

其中 $m_{j1} \geq m_{j2} \geq \dots \geq m_{jk_j} > 0, j = 1, 2, \dots, t$, 且 $\lambda_1, \lambda_2, \dots, \lambda_t$ 两两不同。则复方阵 A 相似于如下的 Jordan 标准型:

$$J = \text{diag}(J_{m_{11}}(\lambda_1), \dots, J_{m_{1k_1}}(\lambda_1), J_{m_{21}}(\lambda_2), \dots, J_{m_{tk_t}}(\lambda_t))$$

证明 对 $k \in \mathbb{N}, \lambda_0 \in \mathbb{C}$, 记 $M = \lambda I_k - J_k(\lambda_0) =$

$$\begin{bmatrix} \lambda - \lambda_0 & -1 & & \\ & \lambda - \lambda_0 & -1 & \\ & & \ddots & \ddots \\ & & & \ddots & -1 \\ & & & & \lambda - \lambda_0 \end{bmatrix},$$

由 $M \begin{pmatrix} 1 & 2 & \dots & k-1 \\ 2 & 3 & \dots & k \end{pmatrix} = (-1)^{k-1}$, 故行列式因子 $D_1(\lambda) = \dots = D_{k-1}(\lambda) = 1, D_k(\lambda) = (\lambda - \lambda_0)^k$, 对应不变

因子 $d_1(\lambda) = \dots = d_{k-1}(\lambda) = 1, d_k(\lambda) = (\lambda - \lambda_0)^k$, 因此 M 的初等因子组为 $\{(\lambda - \lambda_0)^k\}$ 。由引理 8.19, $\lambda I_n - J, \lambda I_n - A$ 的初等因子组相同, 故 A 与 J 相似。

最后总结求复方阵 A 的 Jordan 标准型的过程如下:

算法 8.22

1. 求出特征方阵 $\lambda I_n - A$ 的不变因子;
2. 把不变因子分解为一次因式的乘积, 得到初等因子组;
3. 写出每个初等因子的 Jordan 块;
4. 把所有 Jordan 块进行合并, 得到 Jordan 标准型。

同时也给出另外一种算法:

算法 8.23

1. 求出特征值 $\lambda_1, \lambda_2, \dots, \lambda_t$;
2. 记 $r_{ik} = \text{rank}(A - \lambda_i)^k$, 约定 $r_{i0} = n$;
3. 记 $d_{ik} = r_{i,k-1} - r_{ik}, \forall k \geq 1$;

4. 记 $\delta_{ik} = d_{ik} - d_{i,k+1}, \forall k \geq 1$;
5. 此时 $J_k(\lambda_i)$ 有 δ_{ik} 个。



证明 设 $\phi_A(\lambda) = \det(\lambda I_n - A) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_t)^{n_t}$, 其中 $\sum_{i=1}^t n_i = n$ 。记 $J = \text{diag}(J_1, \cdots, J_t)$,

其中 $J_i = \text{diag}(J_{m_{i1}}(\lambda_i), \cdots, J_{m_{ij_i}}(\lambda_i)), m_{i1} \geq \cdots \geq m_{ij_i}$ 且 $\sum_{k=1}^{j_i} m_{ik} = n_i$ 。

由于 $\forall p \neq q, \lambda_p \neq \lambda_q$, 且此时 $\forall 1 \leq k \leq j_p, J_{m_{pk}}(\lambda_p) - \lambda_q I_{m_{pk}}$ 是满秩上三角阵, 故 $\forall l > 0, \text{rank}(J_{m_{pk}}(\lambda_p) - \lambda_q I_{m_{pk}})^l = m_{pk}$ 。

此时 $r_{ik} = \text{rank}(A - \lambda_i I_n)^k = \text{rank}(J - \lambda_i I_n)^k = \text{rank} \text{diag}(J_1 - \lambda_i I_{n_1}, \cdots, J_t - \lambda_i I_{n_t}) = n - n_i + \sum_{k=1}^{j_i} \text{rank}(J_{m_{ik}}(0))^k$,

$d_{ik} = r_{i,k-1} - r_{ik} = \sum_{k=1}^{j_i} (\text{rank}(J_{m_{ik}}(0))^{k-1} - \text{rank}(J_{m_{ik}}(0))^k) = \langle \text{属于特征值 } \lambda_i \text{ 的阶数 } \geq k \text{ 的 Jordan 块个数} \rangle$,

$\delta_{ik} = d_{ik} - d_{i,k+1} = \langle \text{属于特征值 } \lambda_i \text{ 的阶数为 } k \text{ 的 Jordan 块个数} \rangle = \langle J_k(\lambda_i) \text{ 的个数} \rangle$

使得 $P^{-1}AP$ 为 Jordan 标准型的可逆矩阵 P 称为过渡矩阵。过渡矩阵的求法如下:

算法 8.24

1. 求出矩阵的特征值, 进而得到其 Jordan 标准形 $J = \text{diag}(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \dots, J_{n_k}(\lambda_k))$;
2. 对每个 Jordan 块 $J_{n_i}(\lambda_i)$ 构造 Jordan 链: 首先求解 $(\lambda_i I - A)x_{i1} = 0$ 得到第一个特征向量 x_{i1} , 然后依次求解 $(\lambda_i I - A)x_{ij} = x_{i,j-1}$ 得到后续的广义特征向量 x_{ij} , 直到得到 n_i 个向量;
3. 如果上式中某个方程组无解, 则说明需要调整第一个特征向量的选择, 重新选择后继续上述过程, 直到为每个 Jordan 块构造出完整的 Jordan 链;
4. 将所有 Jordan 链中的向量按顺序排列, 组成矩阵 P , 则 P 即为所求的过渡矩阵。

若多个 Jordan 块对应同一特征值, 则按照 Jordan 块的大小顺序降序排列依次构造 Jordan 链, 以一个矩阵 A 具有一个 $J_3(a)$ 与两个 $J_2(a)$ 为例, 记 $N = A - aI$, 以下是计算方法:

1. 取 $x_{13} \in \ker(N^3) \setminus \ker(N^2)$, 此时有一个 Jordan 链 $x_{13}, Nx_{13} := x_{12}, N^2x_{13} := x_{11}$, 也即求出了最大的 Jordan 块 $J_3(a)$ 对应的三个向量;
2. 取 $x_{22}, x_{32} \in \ker(N^2) \setminus \ker(N)$, 且满足 x_{22}, x_{32} 与 x_{11} 线性无关, 此时有两个 Jordan 链 $x_{22}, Nx_{22} := x_{21}$ 与 $x_{32}, Nx_{32} := x_{31}$, 也即求出了两个 $J_2(a)$ 对应的四个向量;
3. 以上七个向量线性无关, 可以与其他向量一起组成矩阵 P 。



以下是一种不依赖于 Jordan 链的过渡矩阵的求法:

算法 8.25

1. 对特征方阵 $\lambda I_n - A$ 施加初等 λ 变换, 把 $\lambda I_n - A$ 化为 $\lambda I_n - J$, 从而求出 n 阶可逆多项式矩阵 $P(\lambda), Q(\lambda)$ 使得 $P(\lambda)(\lambda I_n - A)Q(\lambda) = \lambda I_n - J$;
2. 求出 $R(\lambda) = Q(\lambda)^{-1}$;
3. 把 $R(\lambda)$ 写成矩阵的多项式 $R(\lambda) = \sum_{j=1}^m R_m \lambda^m$, 则 $P = R(A)$ 。



该方法正确性的证明留给读者。

8.3 Jordan 标准型的应用

例题 8.1 证明: A, A^T 相似。

证明 设 A 的 Jordan 标准型为 J , 即证 J, J^T 相似。设 $J = \text{diag}(J_{m_{11}}(\lambda_1), \cdots, J_{m_{kt}}(\lambda_t))$ 。

对 $\forall 1 \leq i \leq t, 1 \leq j \leq k_i$, 容易求得 $J_{m_{ij}}(\lambda_i), J_{m_{ij}}(\lambda_i)^\top$ 的初等因子组均为 $\{(\lambda - \lambda_i)^{m_{ij}}\}$, 故存在可逆方阵 P_{ij} 使得 $P_{ij}^{-1} J_{m_{ij}}(\lambda_i) P_{ij} = J_{m_{ij}}(\lambda_i)^\top$, 令 $P = \text{diag}(P_{11}, \dots, P_{tk_t})$, 即有 $P^{-1} J P = J^\top$, 即 $A, A^\top$ 相似。

 **笔记** 注意到 $S_{ij} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \end{bmatrix}$ 满足 $S^{-1} J_{m_{ij}}(\lambda_i) S_{ij} = J_{m_{ij}}(\lambda_i)^\top$, 设 $S = \text{diag}(S_{11}, \dots, S_{tk_t}), Q^{-1} A Q = J$, 则

$(Q S Q^\top)^{-1} A (Q S Q^\top) = A^\top$, 由 $Q S Q^\top$ 对称可逆知存在对称可逆矩阵 P 把 A 相似至 $A^\top$ 。

例题 8.2 证明: 任意一个 n 阶复方阵 A 都可以表示为两个 n 阶对称方阵的乘积, 而且可以指定其中一个是可逆的。

证明 不妨设要求第一个矩阵可逆。由上, 存在对称可逆矩阵 S 使得 $S^{-1} A^\top S = A \Leftrightarrow A^\top S = S A$,

记 $P_1 = S^{-1}$ 对称可逆, $P_2 = A^\top S$ 满足 $P_2^\top = S^\top A = S A = A^\top S = P_2$, 即 P_2 对称。此时 $A = P_1 P_2$ 即为所求。

例题 8.3 设 A 是 n 阶可逆复方阵, 证明: $\exists B \in \mathbb{C}^{n \times n}$ 使得 $A = B^2$ 。

证明 首先证明 $\forall a \neq 0, J_m(a^2)$ 与 $(J_m(a))^2$ 相似。首先有 $J_m(a) = a I_m + N_m, J_m(a^2) = a^2 I_m + N_m$,

$(J_m(a))^2 = a^2 I_m + 2a N_m + N_m^2$, 由 $\text{rank}(J_m(a^2) - a^2 I_m) = \text{rank}((J_m(a))^2 - a^2 I_m) = m - 1$ 即有其相似。

对 A 的某个 Jordan 块 $J_m(a)$, $\exists P_{ma}$ 可逆使得 $P_{ma}^{-1} (J_m(\sqrt{a}))^2 P_{ma} = J_m(a)$, 记 $B_{ma} = P_{ma}^{-1} J_m(\sqrt{a}) P_{ma}$,

设 $A = P^{-1} \text{diag}(J_{m_{11}}(\lambda_1), \dots, J_{m_{tk_t}}(\lambda_t)) P = P^{-1} \text{diag}(B_{m_{11}\lambda_1}^2, \dots, B_{m_{tk_t}\lambda_t}^2) P$, (A 可逆 $\Rightarrow 0$ 不是特征值)

令 $B = P^{-1} \text{diag}(B_{m_{11}\lambda_1}, \dots, B_{m_{tk_t}\lambda_t}) P$, 即有 $A = B^2$ 。

例题 8.4 $A, B \in \mathbb{C}^{3 \times 3}, B \neq O, AB = O$, 证明: $\exists D \neq O, s.t. AD = DA = O$ 。

证明 设 $P^{-1} A P = J$, 记 $\tilde{B} = P^{-1} B P \neq O$, 则 $J \tilde{B} = O \Rightarrow \text{rank}(J) \leq 2$ 。以下对 J 的可能形式分类。

1. $J = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{11} P$ 即可;
2. $J = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{33} P$ 即可;
3. $J = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{33} P$ 即可;
4. $J = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{13} P$ 即可;
5. $J = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{23} P$ 即可。
6. $J = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, 此时取 $D = P^{-1} E_{33} P$ 即可。

例题 8.5 证明: n 阶复矩阵 A 是幂零矩阵当且仅当 $\text{tr}(A^k) = 0, \forall 1 \leq k \leq n$ 。

证明 充分性: A 是幂零矩阵时, A 的特征值全为 0, 故 A 的 Jordan 标准形是 $\text{diag}(J_{n_1}(0), \dots, J_{n_t}(0))$, 则 $\text{tr}(A^k) = \text{tr}(J^k) = 0, \forall k \in \mathbb{N}$ 。

必要性: 设 A 的 Jordan 标准形为 $J = \text{diag}(J_{m_1}(\lambda_1), \dots, J_{m_t}(\lambda_t))$, 并记其对角元为 $\mu_1, \dots, \mu_n$, 则 $\text{tr}(A^k) =$

$\text{tr}(J^k) = \sum_{i=1}^n \mu_i^k = 0, \forall 1 \leq k \leq n$, 由例题3.24可知此方程组只有零解, 即 $\lambda_1 = \lambda_2 = \cdots = \lambda_t = 0$, 故 A 是幂零矩阵。

第九章 矩阵分解

9.1 复习与总览

截至目前，我们已经学习了线性代数中的多个重要主题，包括矩阵的基本运算、行列式、向量空间、线性变换、特征值与特征向量以及内积空间等。这些主题为我们理解和处理线性系统提供了坚实的基础。然而，在实际应用中，直接处理大型矩阵往往效率低下且计算复杂。因此，矩阵分解技术应运而生，成为简化矩阵运算和揭示矩阵结构的重要工具。矩阵分解通过将一个复杂的矩阵表示为多个更简单矩阵的乘积，极大地简化了矩阵的分析和计算过程。以下是按照章节顺序对矩阵分解的主要内容进行的总览。

9.1.1 Gauss 消元与 LU 分解

Gauss 消元法是求解线性方程组的基本方法，通过一系列初等行变换将矩阵化为上三角形，从而简化求解过程。LU 分解则是将矩阵分解为一个下三角矩阵 L 和一个上三角矩阵 U 的乘积，即 $A = LU$ 。这种分解不仅加速了线性方程组的求解，还在数值分析中具有重要应用。

然而，LU 分解并不总是适用，特别是当矩阵不可逆或存在零主元时。为了解决这些问题，引入了带行交换的 LU 分解（也称为 PLU 分解）。以下是其具体形式：

定理 9.1

设 $A \in \mathbb{F}^{n \times n}$ ，若 A 可逆，则存在置换矩阵 $P \in \mathbb{F}^{n \times n}$ ，下三角矩阵 $L \in \mathbb{F}^{n \times n}$ ，上三角矩阵 $U \in \mathbb{F}^{n \times n}$ ，使得 $PA = LU$ 。

9.1.2 与秩相关的分解

在求解矩阵的相抵标准形时，引入了矩阵的秩概念。秩表示矩阵中线性无关行或列的最大数量，是衡量矩阵信息量的重要指标。

引理 9.2

设 $A \in \mathbb{F}^{m \times n}$ ， $r = \text{rank}(A)$ ，则存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$ ， $Q \in \mathbb{F}^{n \times n}$ ，使得 $PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ 。

这实际上对应了将矩阵 A 分解为以下形式：

定理 9.3

设 $A \in \mathbb{F}^{m \times n}$ ， $r = \text{rank}(A)$ ，则存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$ ， $Q \in \mathbb{F}^{n \times n}$ ，使得 $A = P^{-1} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q^{-1}$ 。

在上式中把 $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ 拆分为 $\begin{bmatrix} I_r \\ 0 \end{bmatrix} \begin{bmatrix} I_r & 0 \end{bmatrix}$ ，分别与 P^{-1} ， Q^{-1} 相乘，可以得到矩阵 A 的一种低秩分解形式：

定理 9.4

设 $A \in \mathbb{F}^{m \times n}$ ， $r = \text{rank}(A)$ ，则存在满秩矩阵 $D \in \mathbb{F}^{m \times r}$ ， $F \in \mathbb{F}^{r \times n}$ ，使得 $A = DF$ 。

9.1.3 与特征值相关的分解

特征值是矩阵的重要性质。但是不是所有矩阵都能通过相似变换对角化。为了解决这一问题，引入了 Jordan 标准形，它将矩阵分解为 Jordan 块的形式，从而揭示了矩阵的结构。

定理 9.5

设 $A \in \mathbb{C}^{n \times n}$ 矩阵, $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 A 的全部不同特征值, 则存在可逆矩阵 $P \in \mathbb{F}^{n \times n}$, 使得 $A = PJP^{-1}$, 其中 $J = \text{diag}(J_{n_{11}}(\lambda_1), J_{n_{12}}(\lambda_1), \dots, J_{n_{l_1}}(\lambda_1), J_{n_{21}}(\lambda_2), \dots, J_{n_{k_l}}(\lambda_k))$, 其中 $J_{n_{ij}}(\lambda_i)$ 是 Jordan 块, 且 $\sum_{i,j} n_{ij} = n$ 。

当需要对过渡矩阵提出更高要求时, 可以使用 Schur 分解:

定理 9.6

设 $A \in \mathbb{C}^{n \times n}$, 存在酉矩阵 $U \in \mathbb{C}^{n \times n}$, 使得 $A = UTU^H$, 其中 T 为上三角矩阵, 且 T 的对角线元素为 A 的特征值。

证明 使用归纳完成证明。对 $n=1$ 显然成立, 假设对 $n-1$ 成立, 则考虑 n 阶矩阵 A 。

由代数学基本定理, A 至少有一个特征值 λ_1 , 设对应的单位特征向量为 u_1 。将 u_1 扩展为 $\mathbb{C}^n$ 的一组标准正交基 $\{u_1, \dots, u_n\}$, 构成酉矩阵 $U_1 = [u_1, \dots, u_n]$, 则 $U_1^H A U_1 = U_1^H [A u_1, A u_2, \dots, A u_n] = U_1^H [\lambda_1 u_1, A u_2, \dots, A u_n] = \begin{bmatrix} \lambda_1 & b^H \\ 0 & A_1 \end{bmatrix}$, 其中 $b \in \mathbb{C}^{n-1}, A_1 \in \mathbb{C}^{(n-1) \times (n-1)}$ 。由归纳假设, 存在酉矩阵 $U_2 \in \mathbb{C}^{(n-1) \times (n-1)}$, 使得 $U_2^H A_1 U_2 = T_2$,

其中 T_2 为上三角矩阵, 且 T_2 的对角线元素为 A_1 的特征值。令 $U = \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}$, 则 U 为酉矩阵, 进而 $V = U U_1$ 是酉矩阵, 且 $V^H A V = U^H U_1^H A U_1 U = \begin{bmatrix} 1 & 0 \\ 0 & U_2^H \end{bmatrix} \begin{bmatrix} \lambda_1 & b^H \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & b^H U_2 \\ 0 & T_2 \end{bmatrix}$, 显然该矩阵为上三角矩阵, 且其对角线元素为 A 的特征值。由此完成归纳, 定理得证。

9.1.4 一般矩阵的分解

实际应用中, 矩阵往往不是方阵, 此时便引入了一些对矩阵自身条件要求较弱的分解方法, 如基于 Gram-Schmidt 正交化的 QR 分解和奇异值分解。

定理 9.7

设 $A \in \mathbb{C}^{m \times n}$, 存在酉矩阵 $Q \in \mathbb{C}^{m \times m}$ 和上三角矩阵 $R \in \mathbb{C}^{m \times n}$, 使得 $A = QR$ 。

QR 分解的转置情形也成立:

定理 9.8

设 $A \in \mathbb{C}^{m \times n}$, 存在酉矩阵 $Q \in \mathbb{C}^{n \times n}$ 和下三角矩阵 $L \in \mathbb{C}^{m \times n}$, 使得 $A = LQ$ 。

除了 QR 分解, 奇异值分解 (SVD) 也是一种重要的矩阵分解方法。定义矩阵 A 的奇异值为矩阵 $A^H A$ 的非零特征值的平方根, 则奇异值数量等于矩阵 A 的秩, 且有分解:

定理 9.9

设 $A \in \mathbb{C}^{m \times n}$ 矩阵, $r = \text{rank}(A)$, 则存在酉矩阵 $U \in \mathbb{C}^{m \times m}, V \in \mathbb{C}^{n \times n}$ 和矩阵 $\Sigma \in \mathbb{C}^{m \times n}$, 使得 $A = U \Sigma V^H$, 其中 Σ 的前 r 个对角线元素为 A 的奇异值, 其他元素均为零。

与秩的分解类似, 把 Σ 拆分为 $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} D_r \begin{bmatrix} I_r \\ 0 \end{bmatrix}$, 左右两侧矩阵分别与 U, V^H 相乘, 可以得到矩阵 A 的另一种奇异值分解形式:

定理 9.10

设 $A \in \mathbb{C}^{m \times n}$, $r = \text{rank}(A)$, 则存在矩阵 U, V 作为某个酉矩阵的前 r 列和 A 的奇异值组成的对角矩阵 D_r , 使得 $A = UD_rV^H$.

为了把握矩阵的更多结构信息, 还引入了极分解:

定理 9.11

设 $A \in \mathbb{C}^{m \times n}$, $m \geq n$, 则存在矩阵 $U \in \mathbb{C}^{m \times n}$ 满足 $U^H U = I_n$ 和半正定矩阵 $P \in \mathbb{C}^{n \times n}$, 使得 $A = UP$ 。若 A 可逆, 则 P 正定且分解唯一。

同理有左极分解:

定理 9.12

设 $A \in \mathbb{C}^{m \times n}$, $m \leq n$, 则存在矩阵 $V \in \mathbb{C}^{m \times n}$ 满足 $VV^H = I_m$ 和半正定矩阵 $Q \in \mathbb{C}^{m \times m}$, 使得 $A = QV$ 。若 A 可逆, 则 Q 正定且分解唯一。

奇异值分解和极分解是本节课重点关注的矩阵分解方法。

9.1.5 对称/ Hermitian 矩阵分解

我们熟知对称矩阵可以正交相似对角化, 对于满足 $A^H = A$ 的矩阵 A 也有类似结论:

定理 9.13

设 $A \in \mathbb{C}^{n \times n}$ 满足 $A^H = A$, 则存在酉矩阵 $U \in \mathbb{C}^{n \times n}$ 和实数 $\lambda_1, \lambda_2, \dots, \lambda_n$, 使得 $A = U\Lambda U^H$, 其中 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 。

对于正定矩阵与半正定矩阵, 还有更强的 Cholesky 分解形式:

定理 9.14

设 $A \in \mathbb{C}^{n \times n}$ 为正定矩阵, 则存在唯一的上三角矩阵 $R \in \mathbb{C}^{n \times n}$, 使得 $A = R^H R$, 且 R 的对角线元素均为正实数。设 $A \in \mathbb{C}^{n \times n}$ 为半正定矩阵, 则存在上三角矩阵 $R \in \mathbb{C}^{n \times n}$, 使得 $A = R^H R$, 且 R 的对角线元素均为非负实数。

下三角情形同理成立:

定理 9.15

设 $A \in \mathbb{C}^{n \times n}$ 为正定矩阵, 则存在唯一的下三角矩阵 $L \in \mathbb{C}^{n \times n}$, 使得 $A = LL^H$, 且 L 的对角线元素均为正实数。设 $A \in \mathbb{C}^{n \times n}$ 为半正定矩阵, 则存在下三角矩阵 $L \in \mathbb{C}^{n \times n}$, 使得 $A = LL^H$, 且 L 的对角线元素均为非负实数。

Bunch 与 Kaufman 提出了一种推广的 Cholesky 分解方法, 适用于所有 Hermitian 矩阵:

定理 9.16

设 $A \in \mathbb{C}^{n \times n}$ 满足 $A^H = A$, 则存在置换矩阵 $P \in \mathbb{C}^{n \times n}$, 单位下三角矩阵 $L \in \mathbb{C}^{n \times n}$ 和准对角矩阵 $B \in \mathbb{C}^{n \times n}$, 使得 $PAP^H = LBL^H$, 其中 B 的对角块为 1×1 或 2×2 矩阵。



9.2 奇异值分解

先证明奇异值分解的存在性:

证明 对复矩阵 A , 考虑 $M = A^H A$ 作为半正定矩阵, 其特征值均为非负实数。设 M 的特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$, 对应的单位正交特征向量为 $v_1, v_2, \dots, v_n$, 则有 $Mv_i = \lambda_i v_i$ 。定义 $\sigma_i = \sqrt{\lambda_i}$, 则 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 为 A 的奇异值。构造矩阵 $V = [v_1, v_2, \dots, v_n]$, 则 V 为酉矩阵。

对于所有 $\sigma_i > 0$, 定义 $u_i = \frac{1}{\sigma_i} A v_i$, 则有 $u_i^H u_j = \frac{1}{\sigma_i \sigma_j} v_i^H A^H A v_j = \frac{\lambda_j}{\sigma_i \sigma_j} v_i^H v_j = \delta_{ij}$, 即 $u_1, u_2, \dots, u_r$ 是标准正交集, 将其扩充为标准正交基 $u_1, u_2, \dots, u_m$, 构造矩阵 $U = [u_1, u_2, \dots, u_m]$, 则 U 也是酉矩阵。以下验证 $AV = U\Sigma$:

对于 $1 \leq i \leq r$, 有 $Av_i = \sigma_i u_i$; 对于 $r+1 \leq i \leq n$, 有 $\|Av_i\|^2 = v_i^H A^H A v_i = v_i^H \lambda_i v_i = 0$, 即 $Av_i = 0$ 。综上所述, $AV = U\Sigma$, 其中 Σ 为 $m \times n$ 矩阵, 其前 r 个对角线元素为 $\sigma_1, \sigma_2, \dots, \sigma_r$, 其他元素均为零。两边右乘 V^H 即得 $A = U\Sigma V^H$ 。

例题 9.1 $A \in \mathbb{C}^{m \times n}$, $\sigma_1, \dots, \sigma_r$ 为 A 的奇异值, 证明: $\|A\|_F^2 := \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 = \sum_{i=1}^r \sigma_i^2$ 。

证明 $\sum_{i=1}^r \sigma_i^2$ 是 $A^H A$ 的非零特征值之和, 即 $A^H A$ 的特征值之和, 故 $\sum_{i=1}^r \sigma_i^2 = \text{tr}(A^H A) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2$ 。

例题 9.2 证明: $A \in \mathbb{C}^{n \times n}$ 的奇异值是其所有非零特征值当且仅当 A 是半正定 Hermitian 矩阵。

证明 充分性: 由 A 为对称半正定矩阵可知, 存在酉矩阵 Q , 使得 $A = Q\Lambda Q^H$, 其中可取 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$ 。则这也是 A 的奇异值分解形式, 故 A 的奇异值即为其所有非零特征值。必要性: 设 A 的奇异值分解为 $A = U\Sigma V^H$, 并记 $T = V^H U = (t_{ij})_{n \times n}$, 则 $\text{tr} A = \text{tr}(U\Sigma V^H) = \text{tr}(\Sigma T) = \sum_{i=1}^r \sigma_i t_{ii} \leq \sum_{i=1}^r \sigma_i$, 等号成立当且仅当 $t_{ii} = 1, i = 1, 2, \dots, r$ 。又由于 T 为酉矩阵, 故 $t_{ii} = 1$ 蕴含 $t_{ij} = 0, i \neq j, i < r, j < r$, 故 $T = \begin{bmatrix} I_r & 0 \\ 0 & T_1 \end{bmatrix}$, 其中 T_1 为酉矩阵。此时 $A = U\Sigma V^H = U \begin{bmatrix} I_r & 0 \\ 0 & T_1 \end{bmatrix} \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix} V^H = U \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix} V^H$, 即 A 可以酉相似对角化, 且其特征值为其奇异值和零, 是非负实数, 故 A 为对称半正定矩阵。

例题 9.3 证明: $A \in \mathbb{C}^{n \times n}$ 的奇异值是其所有非零特征值的模长当且仅当 A 是正规矩阵, 即 $AA^H = A^H A$ 。

证明 充分性: 先证明 A 是正规矩阵 $\iff A$ 可酉对角化。($\Rightarrow$) 设 A 可酉相似到上三角阵 T , 则 $A = UTU^H$, 其中 U 是酉矩阵, T 是上三角矩阵。由 $AA^H = A^H A$ 可得 $TT^H = T^H T := (t_{ij})_{n \times n}$ 。比较其对角线元素, 有 $\sum_{k=1}^n |t_{ik}|^2 = \sum_{k=1}^n |t_{ki}|^2, i = 1$ 可推导出 $t_{1j} = 0, j > 1$, 再由 $i = 2$ 可推导出 $t_{2j} = 0, j > 2$, 依此类推可得 T 为对角矩阵, 故 A 可酉对角化。($\Leftarrow$) 设 A 可酉对角化为 $A = U\Lambda U^H$, 则 $AA^H = U\Lambda\Lambda^H U^H = U\Lambda^H \Lambda U^H = A^H A$, 故 A 是正规矩阵。

由 A 可酉对角化可知, 存在酉矩阵 U , 使得 $A = U\Lambda U^H$, 其中 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 则 $A^H A = U\Lambda^H \Lambda U^H$, 故 $A^H A$ 的特征值为 $|\lambda_1|^2, |\lambda_2|^2, \dots, |\lambda_r|^2, 0, \dots, 0$, 其奇异值即为 $|\lambda_1|, |\lambda_2|, \dots, |\lambda_r|$, 即 A 的奇异值是其所有非零特征值的模长。

必要性: 先证明左乘或右乘酉矩阵不改变矩阵的 Frobenius 范数。设 $Q \in \mathbb{C}^{m \times m}$ 为酉矩阵, 则 $\|QA\|_F^2 = \text{tr}((QA)^H QA) = \text{tr}(A^H Q^H QA) = \text{tr}(A^H A) = \|A\|_F^2 = \text{tr}(AA^H) = \text{tr}(AQ Q^H A^H) = \text{tr}((AQ)(AQ)^H) = \|AQ\|_F^2$ 。

再证明 Schur 不等式: $\|A\|_F^2 \geq \sum_{i=1}^n |\lambda_i|^2$, 其中 λ_i 为 A 的特征值, 且取到等号当且仅当 A 是正规矩阵。设 A 可酉相似到上三角阵 T , 则 $A = UTU^H$, 其中 U 是酉矩阵, T 是上三角矩阵。此时 $\|A\|_F^2 = \|T\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n |t_{ij}|^2 \geq$

$\sum_{i=1}^n |t_{ii}|^2 = \sum_{i=1}^n |\lambda_i|^2$ 。等号成立条件为 $t_{ij} = 0, i \neq j$, 即 T 为对角矩阵, 故 A 是可酉对角化矩阵, 是正规矩阵。

此时由 $\|A\|_F^2 = \sum_{i=1}^r \sigma_i^2 = \sum_{i=1}^n |\lambda_i|^2$ 可知 A 是正规矩阵。

注 方阵的奇异值与其特征值有什么关系?

笔记 方阵的特征值是矩阵作为线性空间自映射时的固有伸缩因子, 是其固有的代数性质, 不需要空间中有“长度”与“角度”的概念, 也就是不需要度量/内积即可定义。而奇异值是矩阵作为双线性型 $B: V \times V \rightarrow \mathbb{F}: B(u, v) = u^T A v$ 的性质, 其定义依赖于距离与角度的概念, 因此需要空间中有内积结构才能定义。奇异值可以看作是矩阵在“长度”和“角度”意义下的固有伸缩因子。例如对 $A = \begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}$, 不论 t 取何值, 其特征值均为 1, 但其奇异值会随着 t 的增大而变化, 具体表现为一个会变大, 另一个变小, 反映了矩阵在“长度”和“角度”意义下的伸缩作用。可以说, 这两者没有直接联系。

在一定程度上, 转置/伴随也依赖度量结构。回忆线性映射的伴随的定义: 对定义在内积空间上的线性映射 $\mathcal{A} \in \mathcal{L}(V)$, 其伴随被定义为满足 $\langle \mathcal{A}x, y \rangle = \langle x, \mathcal{A}^*y \rangle$ 的映射 $\mathcal{A}^* \in \mathcal{L}(V)$ 。如果没有内积结构, 就无法定义伴随映射, 而矩阵的转置一定程度上依赖于伴随的概念。因此, 提及伴随也意味着度量的引入。

例题 9.4 设 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$, 证明: $\|AB\|_F^2 \leq \lambda_{\max}(BB^T) \|A\|_F^2$, 其中 $\lambda_{\max}(M)$ 表示方阵 M 的最大特征值。

证明 设 B 的奇异值分解为 $B = U \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0) V^T$, 则有 $BB^T = U \operatorname{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0) U^T$, 故 $(AB)(AB)^T = AU \operatorname{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0) U^T A^T$ 。记 $AU = (a_{ij})_{m \times n}$, 则 $\operatorname{tr}((AB)(AB)^T) = \sum_{i=1}^m \sum_{j=1}^n \sigma_j^2 a_{ij}^2 \leq \sigma_1^2 \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 = \lambda_{\max}(BB^T) \operatorname{tr}(AA^T)$ 。

例题 9.5 设 $A, B \in \mathbb{C}^{m \times n}$, 证明: $AA^H = BB^H$ 当且仅当存在酉矩阵 $V \in \mathbb{C}^{n \times n}$, 使得 $A = BV$ 。

证明 必要性: 由于 $AA^H = BB^H$, 有 A, B 的奇异值分解具有完全相同的奇异值与左奇异向量。设 $A = U \Sigma V_1^H, B = U \Sigma V_2^H$, 则 $A = B V_2 V_1^H$, 记 $V = V_2 V_1^H$ 即得所需结论。

充分性: 由 $A = BV$ 可得 $AA^H = B V V^H B^H = BB^H$ 。

练习 9.1 请给出 $A^H A = B^H B$ 时 A 与 B 之间关系的类似结论, 并证明。

9.3 极分解

同样的, 先证明极分解的存在性。

证明 对 $m \geq n$, 设 A 的奇异值分解为 $A = U \begin{bmatrix} D_r & O \\ O & O_{n-r} \\ O & O \end{bmatrix} V^H = U \begin{bmatrix} I_n \\ O \end{bmatrix} V^H V \begin{bmatrix} D_r & O \\ O & O_{n-r} \end{bmatrix} V^H$, 令 $\tilde{U} = U \begin{bmatrix} I_n \\ O \end{bmatrix} V^H, P = V \begin{bmatrix} D_r & O \\ O & O_{n-r} \end{bmatrix} V^H$, 则 P 正交相似于半正定矩阵 $\begin{bmatrix} D_r & O \\ O & O_{n-r} \end{bmatrix}$, 进而半正定, $\tilde{U}^H \tilde{U} = U^H \begin{bmatrix} I_n & O \end{bmatrix} V V^H \begin{bmatrix} I_n \\ O \end{bmatrix} U = I_n$ 。对于可逆的 A , 有 $m = n = r$, 对应 D_r 是可逆的对角矩阵, 故 P 正定。又由于 $A^H A = P^H U^H U P = P P^H$ 正定, 由正定矩阵平方根的唯一性可知 P 唯一。

对 $m \leq n$, 考虑其转置的极分解, 将其转置即得到左极分解。

例题 9.6 若 n 阶方阵 A 的极分解唯一, 证明: A 可逆。

证明 考虑证明其逆否命题。若不可逆矩阵 A 的极分解中 P 正定, 则 P 可逆, 进而 A 可逆, 矛盾, 故 P 可酉相似于 $\operatorname{diag}(s_1, \dots, s_{n-1}, 0)$, 记过渡矩阵为 Q , 令 $\tilde{U} = U Q^H \operatorname{diag}(I_{n-1}, -1) Q$, 则 $\tilde{U}$ 仍为酉矩阵, 且 $A = U P = \tilde{U} P$ 。因此, A 的极分解不唯一。

例题 9.7 证明: A 规范当且仅当 A 存在一组极分解 $A = UP$ 满足 $UP = PU$ 。

证明 必要性: 规范矩阵可以酉对角化, 设 $A = U \Lambda U^H, \Lambda = \operatorname{diag}(r_1 e^{i\theta_1}, \dots, r_n e^{i\theta_n}), R = \operatorname{diag}(r_1, \dots, r_n), \Theta =$

$\text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, 则 $\Lambda = R\Theta = \Theta R$, 且 R 为半正定矩阵, Θ 为酉矩阵。此时 $A = URU^H U\Theta U^H$, 记 $P = URU^H, U_1 = U\Theta U^H$, 则 $A = U_1 P = P U_1$ 符合要求。

充分性: 设 $A = UP = PU$, 则 $AA^H = UPP^H U^H = PUU^H P^H = P^2, A^H A = UP^2 U^H = PUU^H P = P^2$, 故 $AA^H = A^H A$, 即 A 规范。

9.4 其他内容

例题 9.8 设 $\mathbb{R}_n[x]$ 是所有次数不超过 n 的实系数多项式构成的线性空间, 内积为 $\langle f(x), g(x) \rangle = \int_{-1}^1 f(x)g(x) dx$ 。令 $P_0(x) = 1, P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n)$, 证明: $\{P_0(x), P_1(x), \dots, P_n(x)\}$ 是 $\mathbb{R}_n[x]$ 的正交基。称 $P_n(x)$ 为 Legendre 多项式。

证明 $P_n(x)$ 是一个 $2n$ 次多项式 $f_n(x) = (x-1)^n(x+1)^n$ 的 n 阶导数, 其次数为 n , 故 $\{P_0(x), P_1(x), \dots, P_n(x)\}$ 线性无关, 构成 $\mathbb{R}_n[x]$ 的基。显然 $\forall 0 \leq k < n, f_n^{(k)}(1) = f_n^{(k)}(-1) = 0$ 。以下证明 $P_n(x)$ 与任意次数小于 n 的多项式 $Q(x)$ 正交, 记 $C_n = \frac{1}{2^n n!}$:

$$\begin{aligned} C_n \langle P_n(x), Q(x) \rangle &= \int_{-1}^1 f_n^{(n)}(x) Q(x) dx = \int_{-1}^1 Q(x) df_n^{(n-1)}(x) = f_n^{(n-1)}(x) Q(x) \Big|_{-1}^1 - \int_{-1}^1 f_n^{(n-1)}(x) Q^{(1)}(x) dx \\ &= - \int_{-1}^1 f_n^{(n-1)}(x) Q^{(1)}(x) dx = - \int_{-1}^1 Q^{(1)}(x) df_n^{(n-2)}(x) = - f_n^{(n-1)}(x) Q(x) \Big|_{-1}^1 + \int_{-1}^1 f_n^{(n-2)}(x) Q^{(2)}(x) dx \\ &= \dots = (-1)^n \int_{-1}^1 f_n(x) Q^{(n)}(x) dx = 0 \end{aligned}$$

故 $\{P_0(x), P_1(x), \dots, P_n(x)\}$ 是 $\mathbb{R}_n[x]$ 的正交基。

例题 9.9 设 $\mathbb{R}_n[x]$ 是所有次数不超过 n 的实系数多项式构成的线性空间, 内积为 $\langle f(x), g(x) \rangle = \int_{-1}^1 \frac{f(x)g(x)}{\sqrt{1-x^2}} dx$ 。令 $T_0(x) = 1, T_n(x) = \cos(n \arccos x)$, 证明: $\{T_0(x), T_1(x), \dots, T_n(x)\}$ 是 $\mathbb{R}_n[x]$ 的正交基。称 $T_n(x)$ 为 Chebyshev 多项式。

证明 由 $\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(n\theta)\cos\theta$ 可知 $T_n(x)$ 满足递推关系 $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$, 又 $T_0(x) = 1, T_1(x) = x$ 有 $T_n(x)$ 的次数为 n , 故 $\{T_0(x), T_1(x), \dots, T_n(x)\}$ 线性无关, 构成 $\mathbb{R}_n[x]$ 的基。以下证明其正交性: $\forall m \neq n$,

$$\begin{aligned} \langle T_n(x), T_m(x) \rangle &= \int_{-1}^1 \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx \stackrel{x=\cos\theta}{dx=-\sin\theta d\theta} = \int_{\pi}^0 \frac{\cos(n\theta)\cos(m\theta)}{\sin\theta} (-\sin\theta) d\theta \\ &= \int_0^{\pi} \cos(n\theta)\cos(m\theta) d\theta = \frac{1}{2} \int_0^{\pi} [\cos((n-m)\theta) + \cos((n+m)\theta)] d\theta = 0 \end{aligned}$$

故 $\{T_0(x), T_1(x), \dots, T_n(x)\}$ 是 $\mathbb{R}_n[x]$ 的正交基。

第十章 期末复习

10.1 线性变换与矩阵的相似

首先回顾一些比较重要的算法和性质。

命题 10.1

$\mathcal{A}$ 是 n 维线性空间 V 上的线性映射, 在基 $\{\alpha_1, \dots, \alpha_n\}$ 下矩阵为 A , P 是可逆矩阵, $\beta_i = P\alpha_i$, 则 $\mathcal{A}$ 在基 $\{\beta_1, \dots, \beta_n\}$ 下矩阵是 $P^{-1}AP$ 。

算法 10.2 (特征值与特征向量的计算)

对 $A \in \mathbb{C}^{n \times n}$, 可以通过以下步骤计算其特征值与特征向量:

1. 计算特征多项式 $p_A(\lambda) = \det(\lambda I - A) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$, $\sum_{i=1}^s \lambda_i = n$, 这是关于 λ 的一个 n 次多项式, 从而有 n 个根;
2. 令 $p_A(\lambda) = 0$, 得到特征值 $\lambda_1(n_1 \text{ 重}), \lambda_2(n_2 \text{ 重}) \cdots \lambda_s(n_s \text{ 重})$, 这些是 A 的特征值;
3. 对每一个 $\lambda_i (1 \leq i \leq s)$, 求出线性方程组 $(\lambda_i I - A)x = 0$ 的非零解, 即为 λ_i 的特征向量;

命题 10.3 (特征值的性质)

- 若 λ 是 A 的特征值, x 是对应特征向量, $f(\lambda)$ 是多项式, 则 $f(\lambda)$ 是 $f(A)$ 的特征值, x 是 $f(A)$ 的对应特征向量;
- 若 $\lambda \neq 0$ 是 A 的特征值, x 是对应特征向量, 则 λ^{-1} 是 A^{-1} 的特征值, x 是对应的特征向量;
- 若 λ 是 A 的特征值, x 是对应特征向量, 则 λ 是 A^T 的特征值, 但 x 不一定是对应特征向量;
- 若 λ 是正交矩阵 A 的特征值, 则 $|\lambda| = 1$;
- 若 λ 是实矩阵 A 的特征值, 且 $A = A^T$, 则 λ 是实数, 且存在正交矩阵 P , 使得 $P^{-1}AP$ 是对角矩阵;
- 若 λ 是实矩阵 A 的特征值, 且 $A = -A^T$, 则 λ 是纯虚数或零。

定理 10.4

- 特征值是相似不变量, 即若 A, B 相似, 则 A, B 有相同的特征值;
- 若 $\lambda_1, \dots, \lambda_n$ 是 A 的全部特征值, 则 $\operatorname{tr} A = \sum_{i=1}^n \lambda_i, \det A = \prod_{i=1}^n \lambda_i$;
- 可对角化矩阵 A, B 乘法可交换当且仅当 A, B 可以同时相似对角化。

不是所有方阵都可以对角化, 这里给出一些可对角化的充分条件:

命题 10.5

- 若 n 阶方阵 A 有 n 个不同的特征值, 则 A 可对角化;
- 若方阵 A 的所有特征值的代数重数等于几何重数, 则 A 可对角化;
- 若 n 阶方阵 A 的所有特征值的几何重数等于 n , 则 A 可对角化;
- 若 A 的某个零化多项式没有重根, 则 A 可对角化。

例题 10.1 判断以下命题是否正确, 若正确, 给出证明; 若错误, 给出反例。

1. 若 $(A - I)^2 = O, A\mathbf{1} = \mathbf{1}$, 其中 $\mathbf{1} = (1, 1, \dots, 1)^T$, 则 $A = I$;
2. 若 A 每行元素之和为 2, 则 A^5 每行元素之和为 32。

解

1. 错误。反例: $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$;

2. 正确。A 每行元素之和为 2, 则 $A\mathbf{1} = 2\mathbf{1}$, 因此 2 是 A 的特征值, 故 $2^5 = 32$ 是 A^5 的特征值, 因此 $A^5\mathbf{1} = 32\mathbf{1}$, 即 A^5 每行元素之和为 32。

例题 10.2 证明: $N = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{n \times n}$ 与 $A = \begin{bmatrix} n & 0 & \cdots & 0 \\ n-1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$ 相似。

证明 显然 $N, N - nI$ 是奇异矩阵, 故 $0, n$ 是 N 的特征值, 且 $\dim \ker N = n - 1, \dim \ker(N - nI) = 1$, 故特征值 0 的几何重数是 $n - 1$, 特征值 n 的几何重数是 1, 因此 N 可对角化, 相似标准形是 nE_{11} 。

同理对 A , 0 是几何重数为 $n - 1$ 的特征值, n 是几何重数为 1 的特征值, 故 A 可对角化, 相似标准形也是 nE_{11} , 故 A, N 相似。

例题 10.3 假设 $e_1 = (1, 0, 0)^T, e_2 = (0, 1, 0)^T, e_3 = (0, 0, 1)^T, \beta_1 = (-1, 0, 1)^T, \beta_2 = (0, 1, 1)^T, \beta_3 = (2, -1, -2)^T; M_1 = \{e_1, e_2, e_3\}, M_2 = \{\beta_1, \beta_2, \beta_3\}$ 是 $\mathbb{R}^3$ 的两组基。设 σ 是 $\mathbb{R}^3$ 上的一个线性变换, 且 $\sigma(\beta_1) = (1, 0, -3)^T, \sigma(\beta_2) = (0, -1, -1)^T, \sigma(\beta_3) = (-5, -1, 0)^T$ 。分别求出 σ 在 M_1, M_2 下的矩阵 T_1, T_2 。

解 $\sigma(\beta_1) = (1, 0, -3)^T = -5\beta_1 - 2\beta_2 - 2\beta_3, \sigma(\beta_2) = (0, -1, -1)^T = -\beta_2, \sigma(\beta_3) = (-5, -1, 0)^T = -3\beta_1 - 5\beta_2 - 4\beta_3$ 。

$$\Rightarrow \sigma(\beta_1, \beta_2, \beta_3) = (\beta_1, \beta_2, \beta_3) \begin{bmatrix} -5 & 0 & -3 \\ -2 & -1 & -5 \\ -2 & 0 & -4 \end{bmatrix} \Rightarrow T_2 = \begin{bmatrix} -5 & 0 & -3 \\ -2 & -1 & -5 \\ -2 & 0 & -4 \end{bmatrix}。记 B = (\beta_1, \beta_2, \beta_3), 则$$

$$\sigma(e_1, e_2, e_3) = \sigma((\beta_1, \beta_2, \beta_3)B^{-1}) = (\sigma(\beta_1), \sigma(\beta_2), \sigma(\beta_3))B^{-1} = BT_2B^{-1} = (e_1, e_2, e_3)BT_2B^{-1}$$

$$\Rightarrow T_1 = BT_2B^{-1} = \begin{bmatrix} -4 & 3 & -3 \\ -2 & 1 & -2 \\ -4 & 6 & -7 \end{bmatrix}$$

例题 10.4 A 是 n 阶实对称矩阵, 且 $A^2 + 3A + 2I = 0$ 。

(1) 给出 A 的所有可能的互异特征值。

(2) 若 n 是奇数, 证明 A 的伴随矩阵 A^* 正定。

证明 (1) A 的特征值是方程 $\lambda^2 + 3\lambda + 2 = 0$ 的根, 即 $\lambda_1 = -1, \lambda_2 = -2$ 。

(2) 由 A 对称有 A^{-1} 对称, A^{-1} 的全部可能特征值为 $-1, -\frac{1}{2}$ 均为负数, 故 $A^{-1} \prec 0$ 。

又 n 是奇数, 有 $\det A < 0$, 因此 $A^* = \det A \cdot A^{-1} \succ 0$ 。

10.2 正交变换与内积空间

定义 10.6

$\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, X 是 $\mathbb{K}$ 上线性空间, 若函数 $\|\cdot\|: X \rightarrow \mathbb{R}$ 满足:

- 正定性: $\|x\| \geq 0, \forall x \in X$, 且 $\|x\| = 0$ 当且仅当 $x = 0$;
- 齐次性: $\|\alpha x\| = |\alpha| \cdot \|x\|, \forall x \in X, \alpha \in \mathbb{K}$;
- 三角不等式: $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in X$ 。

则称 $\|\cdot\|$ 是 X 上的一个范数, $(X, \|\cdot\|)$ 称为赋范空间。 $d(x, y) := \|x - y\|$ 是 X 上的一个度量, 称为典则度量。

定义 10.7

X 是 $\mathbb{K}$ 上线性空间, 若函数 $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{K}$ 满足:

- 对第一变元线性: $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle, \forall x, y, z \in X, \alpha, \beta \in \mathbb{K};$
- 对第二变元共轭线性: $\langle x, \alpha y + \beta z \rangle = \bar{\alpha} \langle x, y \rangle + \bar{\beta} \langle x, z \rangle, \forall x, y, z \in X, \alpha, \beta \in \mathbb{K};$
- 共轭对称: $\langle x, y \rangle = \overline{\langle y, x \rangle}, \forall x, y \in X;$
- 正定性: $\langle x, x \rangle \geq 0, \forall x \in X$, 且 $\langle x, x \rangle = 0$ 当且仅当 $x = 0$ 。

则称 $\langle \cdot, \cdot \rangle$ 是 X 上的一个内积, $(X, \langle \cdot, \cdot \rangle)$ 称为内积空间。 $\|x\| := \sqrt{\langle x, x \rangle}$ 是 X 上的一个范数, 称为由内积诱导的范数。

引理 10.8 (Cauchy-Schwarz)

X 是内积空间, 对任意 $x, y \in X$, 有 $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$, 等号成立当且仅当 x, y 线性相关。

定义 10.9

X 是内积空间, 若对 $x, y \in X$, 有 $\langle x, y \rangle = 0$, 则称 x, y 正交, 记作 $x \perp y$ 。

对 $M \subset X$, 若 $\forall y \in M, x \perp y$, 则称 x 正交于 M , 记作 $x \perp M$ 。

$M^\perp := \{x \in X | x \perp M\}$ 称为 M 的正交补。此时若 M 是 X 的子空间, 有 $X = M \oplus M^\perp$ 。

命题 10.10

- (极化恒等式 over $\mathbb{R}$) $(X, \langle \cdot, \cdot \rangle)$ 是实内积空间, 则 $\langle x, y \rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$
- (极化恒等式 over $\mathbb{C}$) $(X, \langle \cdot, \cdot \rangle)$ 是复内积空间, 则 $\langle x, y \rangle = \frac{1}{4}(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2)$
- (平行四边形法则) $(X, \|\cdot\|)$ 是赋范空间, 则 $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$
- (勾股定理) $(X, \langle \cdot, \cdot \rangle)$ 是内积空间, 则 $x \perp y \Leftrightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$

定义 10.11

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $\alpha_1, \dots, \alpha_n$ 是 X 的一组基, 则 $G = (\langle \alpha_i, \alpha_j \rangle)_{n \times n}$ 称为基 $\alpha_1, \dots, \alpha_n$ 的 Gram 矩阵, $G \succ 0$ 。对 $\alpha = (\alpha_1, \dots, \alpha_n)x, \beta = (\alpha_1, \dots, \alpha_n)y$, 有 $\langle \alpha, \beta \rangle = x^T G y$ 。

命题 10.12

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $\alpha_1, \dots, \alpha_n; \beta_1, \dots, \beta_n$ 分别是 X 的一组基, G_1 是基 $\alpha_1, \dots, \alpha_n$ 的 Gram 矩阵, G_2 是基 $\beta_1, \dots, \beta_n$ 的 Gram 矩阵, 若可逆矩阵 P 使得 $(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n)P$, 使得 $G_2 = P^T G_1 P$ 。

定义 10.13

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $e_1, \dots, e_n$ 是 X 的一组基, 若 $\langle e_i, e_j \rangle = \delta_{ij}$, 则称 $e_1, \dots, e_n$ 是 X 的标准正交基。

算法 10.14 (Gram-Schmidt 正交化)

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $\alpha_1, \dots, \alpha_n$ 是 X 的一组基, 对 $i = 1, 2, \dots, n$, 令

$$\begin{aligned}\beta_1 &= \alpha_1, u_1 = \frac{\beta_1}{\|\beta_1\|} \\ \beta_2 &= \alpha_2 - \langle \alpha_2, u_1 \rangle u_1, u_2 = \frac{\beta_2}{\|\beta_2\|} \\ \beta_3 &= \alpha_3 - \langle \alpha_3, u_1 \rangle u_1 - \langle \alpha_3, u_2 \rangle u_2, u_3 = \frac{\beta_3}{\|\beta_3\|}\end{aligned}$$

$$\vdots$$

$$\beta_n = \alpha_n - \sum_{j=1}^{n-1} \langle \alpha_n, u_j \rangle u_j, u_n = \frac{\beta_n}{\|\beta_n\|}$$

则 $u_1, \dots, u_n$ 是 X 的一组标准正交基。

定义 10.15

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $T: X \rightarrow X$ 是线性变换,

- 若对 $\forall x, y \in X$, 有 $\langle Tx, Ty \rangle = \langle x, y \rangle$, 则称 T 是 X 上的正交变换。
- 若对 $\forall x, y \in X$, 有 $\langle x, Ty \rangle = \langle Tx, y \rangle$, 则称 T 是 X 上的对称变换。

命题 10.16

$(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $T: X \rightarrow X$ 是线性变换, 则以下各条性质等价:

- T 是正交变换;
- T 保范数: $\|Tx\| = \|x\|, \forall x \in X$;
- T 把标准正交基映成标准正交基;
- T 在某一组基下的矩阵是正交矩阵。

命题 10.17

实对称阵具有以下性质:

- 实对称阵的特征值是实数;
- 实对称阵的特征向量对应于不同特征值的特征向量是正交的;
- 实对称阵可以正交对角化。

例题 10.5 判断以下命题是否正确, 若正确, 给出证明; 若错误, 给出反例。

1. 设 $V = \{a_0 + a_1x + a_2x^2 | a_0, a_1, a_2 \in \mathbb{R}\}$, 则 $\langle f(x), g(x) \rangle = f(0)g(0)$ 定义了 V 上的一个内积;
2. 设 $\{a_1, a_2, \dots, a_n\}$ 是内积空间 X 的一组基, x 与 $a_1, \dots, a_{n-1}$ 都正交, 则 x 与 a_n 线性相关。

解

1. 错误。不满足正定性。
2. 错误。反例: $X = \mathbb{R}^2, \langle x, y \rangle = x^\top y, a_1 = (1, 0)^\top, a_2 = (1, 1)^\top, x = (0, 1)^\top$ 。

例题 10.6 设 $V = f(x) = a_0 + a_1x + a_2x^2$ 为次数不超过 2 的实系数多项式构成的线性空间。

- (1) 证明: $\langle f(x), g(x) \rangle = f(0)g(0) + f(1)g(1) + f(2)g(2)$ 定义了 V 上的一个内积。
- (2) 应用 Schmidt 正交化方法将向量值 $1, x$ 改造成相对 (1) 中所定义内积的标准正交向量组。

解 (1) 只验证正定性。 $\langle f(x), f(x) \rangle = f(0)^2 + f(1)^2 + f(2)^2 \geq 0$, 取等当且仅当
$$\begin{cases} a_0 = 0 \\ a_1 + a_1 + a_2 = 0 \\ a_0 + 2a_1 + 4a_2 = 0 \end{cases} \Leftrightarrow f = 0.$$

$$(2) f_1 = 1, f_2 = x; u_1 = \frac{f_1}{\|f_1\|} = \frac{\sqrt{3}}{3}, \beta_2 = f_2 - \langle f_2, u_1 \rangle u_1 = x - 1, u_2 = \frac{\beta_2}{\|\beta_2\|} = \frac{\sqrt{2}}{2}(x - 1).$$

例题 10.7 $(X, \langle \cdot, \cdot \rangle)$ 是内积空间, $\alpha_1, \dots, \alpha_s \in X, A = (\langle \alpha_i, \alpha_j \rangle)_{s \times s}$. 证明: $\text{rank } A = \text{rank}(\alpha_1, \dots, \alpha_s)$

证明 设 A_i 是 A 的第 i 列, $\{\alpha_1, \dots, \alpha_r\}$ 是 $\{\alpha_1, \dots, \alpha_s\}$ 的极大线性无关组, 先证明 $A_1, \dots, A_r$ 线性无关。

若 $\sum_{i=1}^r \lambda_i A_i = 0$, 则 $\forall 1 \leq k \leq s, \sum_{i=1}^r \lambda_i \langle \alpha_k, \alpha_i \rangle = 0 \Rightarrow \forall 1 \leq k \leq s, \langle \alpha_k, \sum_{i=1}^r \lambda_i \alpha_i \rangle = 0$

$\Rightarrow \langle \sum_{k=1}^r \lambda_k \alpha_k, \sum_{i=1}^r \lambda_i \alpha_i \rangle = 0 \Rightarrow \sum_{i=1}^r \lambda_i \alpha_i = 0 \Rightarrow \lambda_i = 0, \forall 1 \leq i \leq r$. 因此 $A_1, \dots, A_r$ 确实线性无关。

对其他的 A_k , 设 $\alpha_k = \sum_{i=1}^r \mu_i \alpha_i$, 则 $A_k = \begin{bmatrix} \langle \alpha_1, \alpha_k \rangle \\ \langle \alpha_2, \alpha_k \rangle \\ \vdots \\ \langle \alpha_s, \alpha_k \rangle \end{bmatrix} = \begin{bmatrix} \langle \alpha_1, \sum_{i=1}^r \mu_i \alpha_i \rangle \\ \langle \alpha_2, \sum_{i=1}^r \mu_i \alpha_i \rangle \\ \vdots \\ \langle \alpha_s, \sum_{i=1}^r \mu_i \alpha_i \rangle \end{bmatrix} = \sum_{i=1}^r \mu_i \begin{bmatrix} \langle \alpha_1, \alpha_i \rangle \\ \langle \alpha_2, \alpha_i \rangle \\ \vdots \\ \langle \alpha_s, \alpha_i \rangle \end{bmatrix} = \sum_{i=1}^r \mu_i A_i,$

故 $\text{rank } A = r = \text{rank}(\alpha_1, \dots, \alpha_s)$ 。

 **笔记** 以下是一个典型错误解答。请思考错误原因: $A = \begin{bmatrix} \alpha_1^\top \\ \vdots \\ \alpha_s^\top \end{bmatrix} \begin{bmatrix} \alpha_1 & \cdots & \alpha_s \end{bmatrix} := P^\top P$, 由 $\text{rank}(P^\top P) = \text{rank } P$ 即得。

例题 10.8 已知三阶实对称矩阵 A 的特征值为 $1, 2, 3$, 并且有 $\alpha_1 = (-1, -1, 1)^\top, \alpha_2 = (1, -2, -1)^\top$ 分别为属于特征值 1 和 2 的特征向量。

(1) 求 A 的属于特征值 3 的特征向量的全体,

(2) 求矩阵 A 。

解 (1) A 实对称 $\Rightarrow A$ 属于不同特征值的特征向量相互正交, 故 $V_3 \cup \{0\} = (\text{span}(\alpha_1) \oplus \text{span}(\alpha_2))^\perp = t(1, 0, 1)^\top$ 。

$V_3 = t(1, 0, 1)^\top, t \neq 0$ 。

(2) 把 $\alpha_1, \alpha_2, \alpha_3 = (1, 0, 1)^\top$ 归一化, 得到 $e_1 = \left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)^\top, e_2 = \left(\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right)^\top, e_3 = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right)^\top$

$$P = (e_1, e_2, e_3) \text{ 满足 } P^{-1}AP = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix} \Rightarrow A = P \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix} P^\top = \begin{bmatrix} \frac{13}{6} & -\frac{1}{3} & \frac{5}{6} \\ -\frac{1}{3} & \frac{5}{3} & \frac{1}{3} \\ \frac{5}{6} & \frac{1}{3} & \frac{13}{6} \end{bmatrix}.$$

10.3 二次型与对称矩阵的相合

定义 10.18

$A, B \in \mathbb{R}^{n \times n}$ 是对称矩阵, 若存在可逆矩阵 P 使得 $A = P^\top B P$, 则称 A, B 相合。

相合是一种等价关系, 即满足自反性、对称性、传递性。

 **笔记** 每一个二次型都和一个对称矩阵一一对应, 由此可知矩阵的相合等价于对二次型换元。

类似于相抵/相似标准形, 我们想找到相合关系中等价类的代表元, 因此引入相合标准型的概念。

以下是几种求相合标准形的方法:

算法 10.19

1. 正交相似/正交相合: A 是 n 阶实对称矩阵, 可以正交相似对角化, 即存在 P 是正交矩阵, $P^\top A P$ 是对角矩阵, 此即为 A 的相合对角化;

2. 配方法(对二次型操作): 若 $Q(x)$ 有含平方项, 则直接配方即可, 否则利用 $ab = \frac{1}{4}((a+b)^2 - (a-b)^2)$ 构造出平方项;

3. 初等变换: 对 $\begin{bmatrix} A \\ I \end{bmatrix}$ 进行成对的行初等变换与列初等变换, 最终得到 $\begin{bmatrix} P^\top A P \\ P \end{bmatrix}$ 。

由于相合标准型中对角项系数并不唯一, 因此我们可以进一步求出唯一的相合规范型。

定理 10.20

A 是对称矩阵, 则存在可逆矩阵 P 使得 $P^T A P = \begin{bmatrix} I_r & & \\ & -I_s & \\ & & O \end{bmatrix}$, 其中 $r+s = \text{rank } A$, r 称为正惯性指数, s 称为负惯性指数, $r-s$ 称为符号差, r, s 由 A 唯一确定。

在学习相合标准型之后, 我们就可以对 $\mathbb{R}^3$ 中的曲面进行分类了, 具体结论如下:

命题 10.21

假设已经把二次曲面化为标准形式, 则有: (以下假设 $a, b, c > 0$)

- 若标准形式为 $ax^2 + by^2 + cz^2 = 1$, 则曲面为椭球面;
- 若标准形式为 $ax^2 + by^2 - cz^2 = 1$, 则曲面为单叶双曲面;
- 若标准形式为 $ax^2 - by^2 - cz^2 = 1$, 则曲面为双叶双曲面;
- 若标准形式为 $ax^2 + by^2 - cz = 0$, 则曲面为椭圆抛物面;
- 若标准形式为 $ax^2 - by^2 - cz = 0$, 则曲面为双曲抛物面;
- 若标准形式为 $ax^2 + by^2 - cz^2 = 0$, 则曲面为圆锥面;
- 若标准形式为 $ax^2 + by^2 = 1$, 则曲面为椭圆圆柱面;
- 若标准形式为 $ax^2 - by^2 = 1$, 则曲面为双曲圆柱面;
- 若标准形式为 $ax^2 - by = 0$, 则曲面为抛物圆柱面;

定义 10.22

对 $A \in \mathbb{R}^{n \times n}$, $A^T = A$,

- 若 $\forall 0 \neq x \in \mathbb{R}^n, x^T A x > 0$, 则称 A 是正定的, 记作 $A > 0$;
- 若 $\forall x \in \mathbb{R}^n, x^T A x \geq 0$, 则称 A 是半正定的, 记作 $A \geq 0$;
- 若 $\forall 0 \neq x \in \mathbb{R}^n, x^T A x < 0$, 则称 A 是负定的, 记作 $A < 0$;
- 若 $\forall x \in \mathbb{R}^n, x^T A x \leq 0$, 则称 A 是半负定的, 记作 $A \leq 0$;

命题 10.23

对 $A \in \mathbb{R}^{n \times n}$, $A^T = A$, 以下命题等价:

- A 是正定的;
- A 的正惯性指数是 n , 负惯性指数是 0;
- A 的特征值全是正的;
- 存在唯一的 $B > 0$ 使得 $A = B^2$;
- 存在可逆矩阵 P 使得 $A = P^T P$ (相合与单位阵);
- A 的顺序主子式全为正。

命题 10.24

对 $A \in \mathbb{R}^{n \times n}$, $A^T = A$, 以下命题等价:

- A 是半正定的;
- A 的正惯性指数是 $r = \text{rank } A$, 负惯性指数是 0;
- A 的特征值全是非负的;
- 存在 $B \geq 0$ 使得 $A = B^2$ 且 $\text{rank } A = \text{rank } B$;

- 存在 P 使得 $A = P^T P$ 且 $\text{rank } A = \text{rank } P$;
- A 的所有主子式全为非负。
- A 的所有 k 阶主子式之和非负。

例题 10.9 判断以下命题的正确性, 若正确, 给出证明; 若错误, 给出反例。

1. 设 $A = \begin{bmatrix} A_1 & B \\ B^T & A_2 \end{bmatrix}$, 其中 A_1, A_2 是方阵, 若 A 正定, 则 A_1, A_2 也正定;
2. 设 A 是 n 阶正定矩阵, $\alpha_1, \alpha_2, \dots, \alpha_s \in \mathbb{R}^n$, 且 $\alpha_i \neq 0, \alpha_i^T A \alpha_j = 0 (i \neq j)$, 则 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性无关。

解

1. 正确。 A_1 的顺序主子式是 A 的顺序主子式的一部分, 故 A_1 正定。同理 A_2 也正定。
2. 正确。若 $\sum_{i=1}^s \lambda_i \alpha_i = 0$, 则 $0 = \left(\sum_{i=1}^s \lambda_i \alpha_i \right)^T A \left(\sum_{i=1}^s \lambda_i \alpha_i \right) = \sum_{i=1}^s \lambda_i^2 \alpha_i^T A \alpha_i$, 由 A 正定知 $\alpha_i^T A \alpha_i > 0$, 故 $\lambda_i = 0, \forall 1 \leq i \leq s$ 。

例题 10.10 设实二次型 $Q(x) = x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 - 8x_1x_3 - 4x_2x_3$ 。

- (1) 利用正交变换把二次型化为标准型, 并写出正交变换矩阵。
- (2) 判断 $Q(x) = -1$ 在 $\mathbb{R}^3$ 中表示曲面的类型。

解 (1) 二次型对应矩阵为 $A = \begin{bmatrix} 1 & 2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{bmatrix}$,

特征值和特征向量为 $\lambda_1 = \lambda_2 = 5, \lambda_3 = -4; x_1 = a(-1, 0, 1)^T, x_2 = b(-1, 2, 0)^T, x_3 = (2, 1, 2)^T$ 。

对 x_1, x_2 作 G-S 正交化, 把 x_3 归一化, 得到 $e_1 = \left(-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right), e_2 = \left(-\frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}, \frac{\sqrt{2}}{6} \right), e_3 = \left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3} \right)$ 。

$$\Rightarrow P = \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & \frac{1}{3} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix} \text{ 使得 } P^T A P = \begin{bmatrix} 5 & & \\ & 5 & \\ & & -4 \end{bmatrix}。$$

因此经过正交变换 $x = Py$ 后二次型化为标准型 $\tilde{Q}(y) = 5y_1^2 + 5y_2^2 - 4y_3^2$ 。

- (2) 由 (1) 知曲面为双叶双曲面。

例题 10.11 A 是 n 阶实对称矩阵且 $\det A < 0$, 证明: $\exists x \in \mathbb{R}^n$ 使得 $x^T A x < 0$ 。

证明 对 A 作正交相似对角化使得 $P^T A P = \text{diag}(\lambda_1, \dots, \lambda_n)$, 由 $\det A < 0$ 有至少有一个 $\lambda_i < 0$ 。

取 $x = P e_i$ 即有 $x^T A x = \lambda_i < 0$ 。

例题 10.12 A, B 为 n 阶实对称矩阵, 证明以下结论:

(1) 若 A 正定, 则存在可逆矩阵 P 使得 $P^T A P = I_n, P^T B P$ 为对角矩阵, 且 $P^T B P$ 的对角元全是 $\det(\lambda_i A - B) = 0$ 的根。

* (2) 若 A, B 均半正定, 则存在可逆矩阵 H 使得 $H^T A H, H^T B H$ 为对角矩阵。

证明 (1) 由 A 正定有存在可逆矩阵 Q 使得 $Q^T A Q = I_n$, 且由 B 实对称有 $Q^T B Q$ 是实对称矩阵, 此时存在正交矩阵 S 使得 $S^T Q^T B Q S = \text{diag}(\lambda_1, \dots, \lambda_n)$, 且 $S^T Q^T A Q S = S^T I_n S = I_n$, 故 $P = QS$ 即为所求, 且 $\det(\lambda_i A - B) = \det(\lambda_i (Q^T)^{-1} Q^{-1} - B) = \det Q^T \det(\lambda_i (Q^T)^{-1} Q^{-1} - B) \det Q \cdot \frac{1}{\det(QQ^T)} = \det(\lambda_i I - Q^T B Q) \cdot \frac{1}{\det(QQ^T)} = 0$ 。

(2) 由 A, B 半正定有 $A + B$ 半正定, 故存在可逆矩阵 P 使得 $P^T (A + B) P = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ 。设 $P^T A P = \begin{bmatrix} A_1 & A_2 \\ A_2^T & A_4 \end{bmatrix}$,

则 $P^T B P = \begin{bmatrix} I_r - A_1 & -A_2 \\ -A_2^T & -A_4 \end{bmatrix}$, 由于 $P^T A P \succeq 0$, 有 $A_4 \succeq 0$; 同理 $-A_4 \succeq 0 \Rightarrow A_4 = O$ 。若 $A_2 \neq O$, 则存

在 $u \in \mathbb{R}^r, v \in \mathbb{R}^{n-r}, u^\top A_2 v < 0$, 此时 $\begin{bmatrix} u^\top & \lambda v^\top \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ A_2^\top & O \end{bmatrix} \begin{bmatrix} u \\ \lambda v \end{bmatrix} = u^\top A_1 u + 2\lambda u^\top A_2 v$, 取 λ 足够大, 即有此时

$\begin{bmatrix} u^\top & \lambda v^\top \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ A_2^\top & O \end{bmatrix} \begin{bmatrix} u \\ \lambda v \end{bmatrix} < 0$ 与 $P^\top A P \succeq 0$ 矛盾, 故 $A_2 = O$ 。此时 $P^\top A P = \begin{bmatrix} A_1 & O \\ O & O \end{bmatrix}$, $P^\top B P = \begin{bmatrix} I_r - A_1 & O \\ O & O \end{bmatrix}$ 。对于

对称矩阵 A_1 , 存在正交矩阵 Q 使得 $Q^\top A_1 Q$ 为对角矩阵, 而 $Q^\top (I_r - A_1) Q$ 也是对角矩阵, 因此取 $H = P \begin{bmatrix} Q & O \\ O & I_{n-r} \end{bmatrix}$

即可。