

矩 阵 论 讲 稿

作者：张凯院

使用教材：《矩阵论》

科学出版社 / 西北工业大学出版社

张凯院 徐 仲 等编

辅助教材：《矩阵论导教导学导考》

《矩阵论辅导讲案》

西北工业大学出版社

张凯院 徐 仲 编

课时分配：第一章 18 学时 第四章 8 学时

第二章 6 学时 第五章 8 学时

第三章 8 学时 第六章 8 学时

西北工业大学主页（<http://www.nwpu.edu.cn>）

翱翔学堂—流媒体课程—矩阵论（课堂教学录像）

精品课程—研究生高水平课程 2011—矩阵论（网络教学课件）

矩阵论

张凯院 徐仲 等编

科学出版社 (2013)

西北工业大学出版社 (2017)



矩阵论简明教程 (第3版)

徐仲 张凯院 陆全 等编

科学出版社 (2014)

矩阵论导教导学导考 (第3版)

张凯院 徐仲 编

西北工业大学出版社 (2014)



矩阵论辅导讲案

张凯院 徐仲 编

西北工业大学出版社 (2007)

矩阵论典型题解析及自测试题 (第2版)

张凯院 徐仲 陆全 编

西北工业大学出版社 (2003)



矩阵论同步学习辅导

张凯院 徐仲 编

西北工业大学出版社 (2002)

矩阵论网络教学课件 (音像版)

张凯院 主编

科学出版社 (2008)

矩阵论流媒体课程 (校园网)

西北工业大学主页 (<http://www.nwpu.edu.cn>)

翱翔学堂—流媒体课程—矩阵论

精品课程—研究生高水平课程 2011—矩阵论



第一章 线性空间与线性变换

§ 1.1 线性空间

一、集合与映射

1. 集合: 能够作为整体看待的一堆东西.

列举法: $S = \{a_1, a_2, a_3, \dots\}$

其它表示方法?

性质法: $S = \{a \mid a \text{ 所具有的性质}\}$

相等 ($S_1 = S_2$): 指下面二式同时成立

$$\forall a \in S_1 \Rightarrow a \in S_2, \text{ 即 } S_1 \subset S_2$$

$$\forall b \in S_2 \Rightarrow b \in S_1, \text{ 即 } S_2 \subset S_1$$

交: $S_1 \cap S_2 = \{a \mid a \in S_1 \text{ 且 } a \in S_2\}$

并: $S_1 \cup S_2 = \{a \mid a \in S_1 \text{ 或 } a \in S_2\}$

指定加法规则!

和: $S_1 + S_2 = \{a = a_1 + a_2 \mid a_1 \in S_1, a_2 \in S_2\}$

例 1 $S_1 = \{A = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \mid a_{ij} \in \mathbf{R}\}$

$S_2 = \{A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \mid a_{ij} \in \mathbf{R}\}, S_1 \neq S_2$

$S_1 \cap S_2 = \{A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \mid a_{11}, a_{22} \in \mathbf{R}\}$

$S_1 \cup S_2 = \{A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{12}a_{21} = 0, a_{ij} \in \mathbf{R}\}$

$S_1 + S_2 = \{A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{ij} \in \mathbf{R}\}$

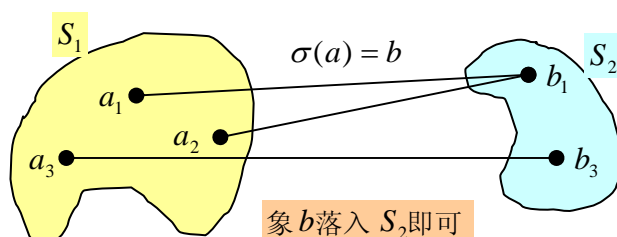
为什么?

2. 数域: 关于四则运算封闭的数的集合.

例如: 实数域 \mathbf{R} , 复数域 \mathbf{C} , 有理数域 \mathbf{Q} , 等等.

3. 映射: 设集合 S_1 与 S_2 , 若对任意的 $a \in S_1$, 按照法则 σ , 对应唯一的 $b \in S_2$, 记作 $\sigma(a) = b$. 称 σ 为由 S_1 到 S_2 的映射; 称 b 为 a 的象, a 为 b 的象源. (b 未必取遍集合 S_2)

变换: 当 $S_1 = S_2$ 时, 称映射 σ 为 S_1 上的变换.



例 2 $S = \{A = (a_{ij})_{n \times n} \mid a_{ij} \in \mathbb{R}\} \quad (n \geq 2).$

映射 $\sigma_1: \sigma_1(A) = \det A \quad (S \rightarrow \mathbb{R})$

变换 $\sigma_2: \sigma_2(A) = (\det A)I_n \quad (S \rightarrow S)$

二、线性空间及其性质

向量 \rightarrow 向量组 \rightarrow 向量集合 (引入线性运算、满足八条算律) \rightarrow 向量空间

矩阵 \rightarrow 矩阵组 \rightarrow 矩阵集合 (引入线性运算、满足八条算律) \rightarrow 矩阵空间

元素 \rightarrow 元素组 \rightarrow 元素集合 (引入线性运算、满足八条算律) \rightarrow 线性空间

1. 线性空间: 集合 V 非空, 给定数域 K , 若在 V 中

(I) 定义的加法运算封闭, 即

$\forall x, y \in V$, 对应唯一元素 $z \in V$, 记作 $z = x + y$, 且满足

(1) 交换律: $x + y = y + x$

(2) 结合律: $x + (y + z) = (x + y) + z \quad (\forall z \in V)$

(3) 有零元: $\exists \theta \in V$, 使得 $x + \theta = x \quad (\forall x \in V)$

(4) 有负元: $\forall x \in V, \exists (-x) \in V$, 使得 $x + (-x) = \theta$.

零元唯一吗?
负元唯一吗?

(II) 定义的数乘运算封闭, 即

$\forall x \in V, \forall k \in K$, 对应唯一元素 $z \in V$, 记作 $z = kx$, 且满足

(5) 数对元素分配律: $k(x + y) = kx + ky \quad (\forall y \in V)$

(6) 元素对数分配律: $(k + l)x = kx + lx \quad (\forall l \in K)$

(7) 数因子结合律: $k(lx) = (kl)x \quad (\forall l \in K)$

(8) 有单位数: 单位数 $1 \in K$, 使得 $1x = x$.

则称 V 为 K 上的线性空间.

例 3 $K = \mathbb{R}$ 时, \mathbb{R}^n —向量空间; $\mathbb{R}^{m \times n}$ —矩阵空间

$P_n[t]$ —多项式空间; $C[a, b]$ —函数空间

$K = \mathbb{C}$ 时, \mathbb{C}^n —复向量空间; $\mathbb{C}^{m \times n}$ —复矩阵空间

例 4 集合 $\mathbb{R}^+ = \{m \mid m \text{ 是正实数}\}$, 数域 $\mathbb{R} = \{k \mid k \text{ 是实数}\}$.

加法: $m, n \in \mathbb{R}^+, m \oplus n = mn$

数乘: $m \in \mathbb{R}^+, k \in \mathbb{R}, k \circ m = m^k$

验证 \mathbb{R}^+ 是 \mathbb{R} 上的线性空间.

任意指定运算方式
都能构成线空吗?

证 加法封闭, 且(1)~(2)成立.

(3) $m \oplus \theta = m \Rightarrow m\theta = m \Rightarrow \theta = 1$

(4) $m \oplus (-m) = \theta \Rightarrow m(-m) = 1 \Rightarrow (-m) = 1/m$

数乘封闭, (5)~(8)成立. 故 \mathbf{R}^+ 是 \mathbf{R} 上的线性空间.

例 5 集合 $\mathbf{R}^2 = \{\alpha = (\xi_1, \xi_2) \mid \xi_i \in \mathbf{R}\}$, 数域 \mathbf{R} . 设 $\beta = (\eta_1, \eta_2), k \in \mathbf{R}$.

运算方式 1 加法: $\alpha + \beta = (\xi_1 + \eta_1, \xi_2 + \eta_2)$

数乘: $k\alpha = (k\xi_1, k\xi_2)$

运算方式 2 加法: $\alpha \oplus \beta = (\xi_1 + \eta_1, \xi_2 + \eta_2 + \xi_1\eta_1)$

数乘: $k \circ \alpha = (k\xi_1, k\xi_2 + \frac{1}{2}k(k-1)\xi_1^2)$

可以验证 $\mathbf{R}^2(+\cdot)$ 与 $\mathbf{R}^2(\oplus \circ)$ 都是 \mathbf{R} 上的线性空间.

两个不同的
线性空间!

[注] 在 $\mathbf{R}^2(\oplus \circ)$ 中, $\theta = (0, 0)$, $-\alpha = (-\xi_1, -\xi_2 + \xi_1^2)$.

Th1 线性空间 V 中的零元素唯一, 负元素也唯一.

证 设 θ_1 与 θ_2 都是 V 的零元素, 则 $\theta_1 = \theta_1 + \theta_2 = \theta_2 + \theta_1 = \theta_2$

设 x_1 与 x_2 都是 x 的负元素, 则由 $x + x_1 = \theta$ 及 $x + x_2 = \theta$ 可得

$$\begin{aligned} x_1 &= x_1 + \theta = x_1 + (x + x_2) = (x_1 + x) + x_2 \\ &= (x + x_1) + x_2 = \theta + x_2 = x_2 + \theta = x_2 \end{aligned}$$

例 6 在线性空间 V 中, 下列结论成立.

$$0x = \theta: 1x + 0x = (1+0)x = 1x \Rightarrow 0x = \theta$$

$$k\theta = \theta: kx + k\theta = k(x + \theta) = kx \Rightarrow k\theta = \theta$$

$$(-1)x = (-x): (-1)x = (-1)x + [x + (-x)] = [(-1)x + 1x] + (-x) = (-x)$$

可以作为
算律应用

2. 减法运算: 线性空间 V 中, $x - y = x + (-y)$.

3. 线性组合: $x, x_i \in V$, 若存在 $c_i \in K$, 使 $x = c_1x_1 + \cdots + c_mx_m$, 则称 x 是 x_1, \cdots, x_m 的**线性组合**, 或者 x 可由 x_1, \cdots, x_m **线性表示**.

4. 线性相关: 若有 c_1, \cdots, c_m 不全为零, 使得 $c_1x_1 + \cdots + c_mx_m = \theta$, 则称元素组 x_1, \cdots, x_m **线性相关**.

5. 线性无关: 仅当 c_1, \cdots, c_m 全为零时, 才有 $c_1x_1 + \cdots + c_mx_m = \theta$, 则称元素组 x_1, \cdots, x_m **线性无关**.

[注] 在 $\mathbf{R}^2(\oplus \circ)$ 中, $\alpha_1 = (1, 1)$, $\alpha_2 = (2, 2)$ 线性无关;

$\alpha_1 = (1, 1)$, $\alpha_2 = (2, 3)$ 线性相关.

特殊运算
特殊结论

在例 4 的 \mathbf{R}^+ 中, 任意两个正数**线性相关**.

三、基与坐标

基不是唯一的!

1. 基与维数: 线性空间 V 中, 若元素组 x_1, \cdots, x_n 满足

(1) x_1, \cdots, x_n 线性无关; (2) $\forall x \in V$ 都可由 x_1, \cdots, x_n 线性表示.

称 x_1, \cdots, x_n 为 V 的一个基, n 为 V 的维数, 记作 $\dim V = n$, 或者 V^n .

[注]
$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} &= \begin{bmatrix} a_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ a_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & a_{22} \end{bmatrix} \\ &= a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22} = \sum_{i=1}^2 \sum_{j=1}^2 a_{ij}E_{ij} \end{aligned}$$

熟练掌握

例 7 矩阵空间 $\mathbf{R}^{m \times n}$ 中, 易见

(1) E_{ij} ($i=1,2,\dots,m; j=1,2,\dots,n$) 线性无关;

(2) $A = (a_{ij})_{m \times n} = \sum_{i=1}^m \sum_{j=1}^n a_{ij}E_{ij}.$

故 E_{ij} ($i=1,2,\dots,m; j=1,2,\dots,n$) 是 $\mathbf{R}^{m \times n}$ 的一个基, $\dim \mathbf{R}^{m \times n} = mn.$

2. 坐标: 给定线性空间 V^n 的基 x_1, \dots, x_n , 当 $x \in V^n$ 时, 有

$$x = \xi_1 x_1 + \dots + \xi_n x_n, \text{ 称 } \xi_1, \dots, \xi_n \text{ 为 } x \text{ 在给定基 } x_1, \dots, x_n \text{ 下的}$$

坐标, 记作列向量 $\alpha = (\xi_1, \dots, \xi_n)^T.$

坐标是列向量

Th2 线性空间 V^n 中, 元素在给定基下的坐标唯一.

证 设 V^n 的基为 x_1, \dots, x_n , 对于 $x \in V^n$, 若

$$x = \xi_1 x_1 + \dots + \xi_n x_n = \eta_1 x_1 + \dots + \eta_n x_n$$

$$\text{则有 } (\xi_1 - \eta_1)x_1 + \dots + (\xi_n - \eta_n)x_n = \theta$$

因为 x_1, \dots, x_n 线性无关, 所以 $\xi_i - \eta_i = 0$, 即 $\xi_i = \eta_i$ ($i=1,2,\dots,n$).

故 x 的坐标唯一.

[注] 设 $x, y \in V^n$ 在给定基 x_1, \dots, x_n 下的坐标为 α, β (列向量), 则

(1) $x + y$ 在该基下的坐标为 $\alpha + \beta.$

(2) kx 在该基下的坐标为 $k\alpha$ ($k \in K$).

抽象运算转化为向量的通常运算

(3) $x = \xi_1 x_1 + \dots + \xi_n x_n = (x_1, \dots, x_n)\alpha$ (矩阵乘法形式).

例 8 矩阵空间 $\mathbf{R}^{2 \times 2}$ 中, 设 $A = (a_{ij})_{2 \times 2}.$

(1) 取基 $E_{11}, E_{12}, E_{21}, E_{22}$, $A = a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22}$

$$\text{坐标为 } \alpha = (a_{11}, a_{12}, a_{21}, a_{22})^T$$

(2) 取基 $B_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, B_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$$A = a_{11}(B_1 - B_2) + a_{12}(B_2 - B_3) + a_{21}(B_3 - B_4) + a_{22}B_4$$

$$= a_{11}B_1 + (a_{12} - a_{11})B_2 + (a_{21} - a_{12})B_3 + (a_{22} - a_{21})B_4$$

$$\text{坐标为 } \beta = (a_{11}, a_{12} - a_{11}, a_{21} - a_{12}, a_{22} - a_{21})^T$$

[注] 一个元素在两个不同的基下的坐标可能相同, 也可能不同.

例如: $A = E_{22}$ 在上述两个基下的坐标都是 $(0, 0, 0, 1)^T$.

$A = E_{11}$ 在上述两个基下的坐标不同.

例 9 设线性空间 V^n 的基为 x_1, \dots, x_n , 元素 y, y_1, \dots, y_m 在该基下的坐标为

$\alpha, \alpha_1, \dots, \alpha_m$, 则有

(1) 元素 y 可由 y_1, \dots, y_m 线性表示 \Leftrightarrow 向量 α 可由 $\alpha_1, \dots, \alpha_m$ 线性表示;

(2) 元素组 y_1, \dots, y_m 线性相关 \Leftrightarrow 向量组 $\alpha_1, \dots, \alpha_m$ 线性相关;

(3) y_{i_1}, \dots, y_{i_r} 是元素组 y_1, \dots, y_m 的最大无关组 \Leftrightarrow

$\alpha_{i_1}, \dots, \alpha_{i_r}$ 是向量组 $\alpha_1, \dots, \alpha_m$ 的最大无关组.

证(2) 对于数组 k_1, \dots, k_m , 因为

$$k_1 y_1 + \dots + k_m y_m = (x_1, \dots, x_n)(k_1 \alpha_1 + \dots + k_m \alpha_m) = \theta$$

等价于 $k_1 \alpha_1 + \dots + k_m \alpha_m = 0$, 所以结论成立.

四、基变换与坐标变换

1. 基变换: 设线性空间 V^n 的基(I)为 x_1, \dots, x_n , 基(II)为 y_1, \dots, y_n , 则

$$\begin{cases} y_1 = c_{11}x_1 + c_{21}x_2 + \dots + c_{n1}x_n \\ y_2 = c_{12}x_1 + c_{22}x_2 + \dots + c_{n2}x_n \\ \dots\dots\dots \\ y_n = c_{1n}x_1 + c_{2n}x_2 + \dots + c_{nn}x_n \end{cases} \quad C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}$$

写成矩阵乘法形式为 $(y_1, \dots, y_n) = (x_1, \dots, x_n)C$

称上式为**基变换公式**, C 为由基(I)改变为基(II)的**过渡矩阵**.

[注] 过渡矩阵 C 一定可逆. 否则 C 的 n 个列向量线性相关, 从而 y_1, \dots, y_n

线性相关 (例 9). 矛盾! 由此可得 $(x_1, \dots, x_n) = (y_1, \dots, y_n)C^{-1}$

称 C^{-1} 为由基(II)改变为基(I)的过渡矩阵.

2. 坐标变换: 设 $x \in V^n$ 在两个基下的坐标分别为 α 和 β , 则有

$$x = \xi_1 x_1 + \dots + \xi_n x_n = (x_1, \dots, x_n)\alpha$$

$$x = \eta_1 y_1 + \dots + \eta_n y_n = (y_1, \dots, y_n)\beta = (x_1, \dots, x_n)C\beta$$

由定理 2 可得 $\alpha = C\beta$, 或者 $\beta = C^{-1}\alpha$, 称为**坐标变换公式**.

例 10 矩阵空间 $R^{2 \times 2}$ 中, 取基

$$(I) \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$(II) \quad B_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

- (1) 求由基(I)改变为基(II)的过渡矩阵;
 (2) 求由基(II)改变为基(I)的坐标变换公式.

解 采用中介法求过渡矩阵.

$$\text{基(0): } E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$(0) \rightarrow (I): (A_1, A_2, A_3, A_4) = (E_{11}, E_{12}, E_{21}, E_{22})C_1$$

$$(0) \rightarrow (II): (B_1, B_2, B_3, B_4) = (E_{11}, E_{12}, E_{21}, E_{22})C_2$$

$$C_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$(I) \rightarrow (II): (B_1, B_2, B_3, B_4) = (A_1, A_2, A_3, A_4)C_1^{-1}C_2$$

$$C = C_1^{-1}C_2 = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix} C_2 = \frac{1}{2} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = C \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2\eta_1 + \eta_2 + \eta_3 + \eta_4 \\ \eta_2 + \eta_3 + \eta_4 \\ 2\eta_1 + 2\eta_2 + \eta_3 \\ \eta_3 \end{bmatrix}$$

五、线性子空间

1. 定义: 线性空间 V 中, 若子集 V_1 非空, 且对 V 中的线性运算封闭, 即

$$(1) \quad \forall x, y \in V_1 \Rightarrow x + y \in V_1$$

$$(2) \quad \forall x \in V_1, \forall k \in K \Rightarrow kx \in V_1$$

称 V_1 为 V 的线性子空间, 简称为子空间.

[注] (1) 子空间 V_1 也是线性空间, 而且 $\dim V_1 \leq \dim V$.

(2) $\{\theta\}$ 是 V 的线性子空间, 规定 $\dim\{\theta\} = 0$.

(3) 子空间 V_1 的零元素就是 V 的零元素.

例 11 线性空间 V 中, 子集 V_1 是 V 的子空间 \Leftrightarrow

对 $\forall x, y \in V_1, \forall k, l \in K$, 有 $kx + ly \in V_1$.

等价命题

证 充分性. $k = l = 1$: $\forall x, y \in V_1 \Rightarrow x + y \in V_1$

$$l = 0: \quad \forall x \in V_1, \forall k \in K \Rightarrow kx = kx + 0y \in V_1$$

故 V_1 是 V 的子空间.

必要性. $\forall x \in V_1, \forall k \in K \Rightarrow kx \in V_1$ (数乘封闭)

$\forall y \in V_1, \forall l \in K \Rightarrow ly \in V_1$ (数乘封闭)

故 $kx + ly \in V_1$ (加法封闭)

例 12 在线性空间 V 中, 设 $x_i \in V (i=1, 2, \dots, m)$, 则

$$V_1 = \{x = k_1 x_1 + \dots + k_m x_m \mid k_i \in K\}$$

构造方法

是 V 的子空间, 称 V_1 为由 x_1, \dots, x_m 生成的子空间.

证 $\forall x \in V_1 \Rightarrow x = k_1 x_1 + \dots + k_m x_m, \forall y \in V_1 \Rightarrow y = l_1 x_1 + \dots + l_m x_m$

$\forall k, l \in K: kx + ly = (kk_1 + ll_1)x_1 + \dots + (kk_m + ll_m)x_m \in V_1$

根据例 11 知, V_1 是 V 的子空间.

[注] (1) 将 V_1 记作 $\text{span}\{x_1, \dots, x_m\}$ 或者 $L(x_1, \dots, x_m)$.

重要结论

(2) 元素组 x_1, \dots, x_m 的最大无关组是 $L(x_1, \dots, x_m)$ 的基.

(3) 若线性空间 V^n 的基为 x_1, \dots, x_n , 则 $V^n = L(x_1, \dots, x_n)$.

2. 矩阵的值域:

划分 $A = (a_{ij})_{m \times n} = (\beta_1, \dots, \beta_n) \quad (\beta_j \in C^m)$

称 $R(A) = L(\beta_1, \dots, \beta_n)$ 为矩阵 A 的值域 (列空间)

易见 $\dim R(A) = \text{rank} A$.

等价定义

例 13 矩阵 A 的值域 $R(A) = \{\beta = Ax \mid x \in C^n\}$.

证 $\forall \beta \in \text{左}, \text{有 } \beta = k_1 \beta_1 + \dots + k_n \beta_n = (\beta_1, \dots, \beta_n) \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix} = Ax \in \text{右}$

$\forall \beta \in \text{右}, \text{有 } \beta = Ax = (\beta_1, \dots, \beta_n) \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix} = k_1 \beta_1 + \dots + k_n \beta_n \in \text{左}$

3. 矩阵的零空间:

设 $A \in C^{m \times n}$, 称 $N(A) = \{x \mid Ax = 0, x \in C^n\}$ 为矩阵 A 的零空间.

易见 $\dim N(A) = n - \text{rank} A$.

Th3 线性空间 V^n 中, 设子空间 V_1 的基为 $x_1, \dots, x_m (m < n)$, 则存在

$x_{m+1}, \dots, x_n \in V^n$, 使得 $x_1, \dots, x_m, x_{m+1}, \dots, x_n$ 为 V^n 的基.

证 $m < n \Rightarrow \exists x_{m+1} \in V^n$ 不能由 x_1, \dots, x_m 线性表示

$\Rightarrow x_1, \dots, x_m, x_{m+1}$ 线性无关

子空间的基能够扩充为线性空间的基

若 $m+1 = n$, 则 x_1, \dots, x_m, x_{m+1} 是 V^n 的基;

否则, $m+1 < n \Rightarrow \exists x_{m+2} \in V^n$ 不能由 x_1, \dots, x_m, x_{m+1} 线性表示

$\Rightarrow x_1, \dots, x_m, x_{m+1}, x_{m+2}$ 线性无关

若 $m+2 = n$, 则 $x_1, \dots, x_m, x_{m+1}, x_{m+2}$ 是 V^n 的基;

否则, $m+2 < n \Rightarrow \dots$. 依此类推, 即得所证.

六、子空间的交与和

1. 子空间的交: $V_1 \cap V_2 = \{x \mid x \in V_1 \text{ 且 } x \in V_2\}$

Th4 设 V_1, V_2 是线性空间 V 的子空间, 则 $V_1 \cap V_2$ 是 V 的子空间.

证 $\theta \in V_1, \theta \in V_2 \Rightarrow \theta \in V_1 \cap V_2 \Rightarrow V_1 \cap V_2$ 非空

$$\forall x, y \in V_1 \cap V_2 \Rightarrow \begin{cases} x, y \in V_1 \Rightarrow x+y \in V_1 \\ x, y \in V_2 \Rightarrow x+y \in V_2 \end{cases} \Rightarrow x+y \in V_1 \cap V_2$$

$$\forall k \in K, \forall x \in V_1 \cap V_2 \Rightarrow \begin{cases} x \in V_1 \Rightarrow kx \in V_1 \\ x \in V_2 \Rightarrow kx \in V_2 \end{cases} \Rightarrow kx \in V_1 \cap V_2$$

所以 $V_1 \cap V_2$ 是 V 的子空间.

2. 子空间的和: $V_1 + V_2 = \{x = x_1 + x_2 \mid x_1 \in V_1, x_2 \in V_2\}$

Th5 设 V_1, V_2 是线性空间 V 的子空间, 则 $V_1 + V_2$ 是 V 的子空间.

证 $\theta \in V_1, \theta \in V_2 \Rightarrow \theta = \theta + \theta \in V_1 + V_2 \Rightarrow V_1 + V_2$ 非空

$$\begin{aligned} \forall x, y \in V_1 + V_2 &\Rightarrow \begin{cases} x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2 \\ y = y_1 + y_2, y_1 \in V_1, y_2 \in V_2 \end{cases} \\ &\Rightarrow x+y = (x_1+y_1) + (x_2+y_2), \quad x_1+y_1 \in V_1, x_2+y_2 \in V_2 \\ &\Rightarrow x+y \in V_1 + V_2 \end{aligned}$$

$$\begin{aligned} \forall k \in K, \forall x \in V_1 + V_2 &\Rightarrow x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2 \\ &\Rightarrow kx = kx_1 + kx_2, \quad kx_1 \in V_1, kx_2 \in V_2 \\ &\Rightarrow kx \in V_1 + V_2 \end{aligned}$$

所以 $V_1 + V_2$ 是 V 的子空间.

[注] $V_1 \cup V_2$ 不一定是 V 的子空间.

例如: 在 \mathbb{R}^2 中, $V_1 = L(e_1)$ 与 $V_2 = L(e_2)$ 的并集为

$$V_1 \cup V_2 = \{\alpha = (\xi_1, \xi_2) \mid \xi_1 \cdot \xi_2 = 0, \xi_i \in \mathbb{R}\}$$

易见 $e_1, e_2 \in V_1 \cup V_2$, 但 $e_1 + e_2 = (1, 1) \notin V_1 \cup V_2$, 故加法运算不封闭.

Th6 设 V_1, V_2 是线性空间 V 的有限维子空间, 则

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$$

维数公式

证 记 $\dim V_1 = n_1, \dim V_2 = n_2, \dim V_1 \cap V_2 = m$

欲证 $\dim(V_1 + V_2) = n_1 + n_2 - m$

(1) $m = n_1$: $(V_1 \cap V_2) \subset V_1 \Rightarrow V_1 \cap V_2 = V_1$

$$(V_1 \cap V_2) \subset V_2 \Rightarrow V_1 \subset V_2 \Rightarrow V_1 + V_2 = V_2$$

$$\dim(V_1 + V_2) = \dim V_2 = n_2 = n_1 + n_2 - m$$

(2) $m = n_2$: $(V_1 \cap V_2) \subset V_2 \Rightarrow V_1 \cap V_2 = V_2$

$$(V_1 \cap V_2) \subset V_1 \Rightarrow V_2 \subset V_1 \Rightarrow V_1 + V_2 = V_1$$

$$\dim(V_1 + V_2) = \dim V_1 = n_1 = n_1 + n_2 - m$$

(3) $m < n_1, m < n_2$: 设 $V_1 \cap V_2$ 的基为 x_1, \dots, x_m , 那么

扩充为 V_1 的基: $x_1, \dots, x_m, y_1, \dots, y_{n_1-m}$ (I)

扩充为 V_2 的基: $x_1, \dots, x_m, z_1, \dots, z_{n_2-m}$ (II)

考虑元素组: $x_1, \dots, x_m, y_1, \dots, y_{n_1-m}, z_1, \dots, z_{n_2-m}$ (III)

因为 $V_1 = L(I)$, $V_2 = L(II)$, 所以 $V_1 + V_2 = L(III)$ (自证).

下面证明元素组(III)线性无关:

设数组 $k_1, \dots, k_m, p_1, \dots, p_{n_1-m}, q_1, \dots, q_{n_2-m}$ 使得

$$k_1 x_1 + \dots + k_m x_m + p_1 y_1 + \dots + p_{n_1-m} y_{n_1-m} + q_1 z_1 + \dots + q_{n_2-m} z_{n_2-m} = \theta$$

$$\text{由 } x = \begin{cases} k_1 x_1 + \dots + k_m x_m + p_1 y_1 + \dots + p_{n_1-m} y_{n_1-m} \in V_1 \\ -(q_1 z_1 + \dots + q_{n_2-m} z_{n_2-m}) \in V_2 \end{cases} \quad (*)$$

得 $x \in V_1 \cap V_2 \Rightarrow x = l_1 x_1 + \dots + l_m x_m$

结合(*)中第二式得

$$l_1 x_1 + \dots + l_m x_m + q_1 z_1 + \dots + q_{n_2-m} z_{n_2-m} = \theta$$

$$(II) \text{ 线性无关} \Rightarrow l_1 = \dots = l_m = 0, q_1 = \dots = q_{n_2-m} = 0$$

结合(*)中第一式得

$$k_1 x_1 + \dots + k_m x_m + p_1 y_1 + \dots + p_{n_1-m} y_{n_1-m} = \theta$$

$$(I) \text{ 线性无关} \Rightarrow k_1 = \dots = k_m = 0, p_1 = \dots = p_{n_1-m} = 0$$

故元素组(III)线性无关, 从而是 $V_1 + V_2$ 的一个基.

因此 $\dim(V_1 + V_2) = n_1 + n_2 - m$.

3. 子空间的直和:

$$V_1 + V_2 = \{x = x_1 + x_2 \mid \text{唯一 } x_1 \in V_1, \text{唯一 } x_2 \in V_2\}$$

记作: $V_1 + V_2 = V_1 \oplus V_2$

例如:

(1) 在 \mathbb{R}^3 中, 取 $V_1 = L(e_1, e_2), V_2 = L(e_2, e_3)$ 时, 有 $V_1 + V_2 = \mathbb{R}^3$, 此时

$$x = (1, 2, 3) = (1, 2, 0) + (0, 0, 3), (1, 2, 0) \in V_1, (0, 0, 3) \in V_2$$

$$x = (1, 2, 3) = (1, 1, 0) + (0, 1, 3), (1, 1, 0) \in V_1, (0, 1, 3) \in V_2 \text{ 不唯一}$$

(2) 在 \mathbb{R}^3 中, 取 $V_1 = L(e_1, e_2), V_2 = L(e_3)$ 时, 有 $V_1 + V_2 = \mathbb{R}^3$, 此时

$$x = (1, 2, 3) = (1, 2, 0) + (0, 0, 3), \text{ 而 } (1, 2, 0) \in V_1, (0, 0, 3) \in V_2 \text{ 唯一}$$

Th7 设 V_1, V_2 是线性空间 V 的子空间, 则 $V_1 + V_2$ 是直和 $\Leftrightarrow V_1 \cap V_2 = \{\theta\}$.

证 充分性. 已知 $V_1 \cap V_2 = \{\theta\}$: 对于 $\forall z \in V_1 + V_2$, 若

$$\begin{cases} z = x_1 + x_2, x_1 \in V_1, x_2 \in V_2 \\ z = y_1 + y_2, y_1 \in V_1, y_2 \in V_2 \end{cases}$$

$$\text{则有 } (x_1 - y_1) + (x_2 - y_2) = \theta, x_1 - y_1 \in V_1, x_2 - y_2 \in V_2$$

$$\Rightarrow x_1 - y_1 = -(x_2 - y_2) \in V_1 \cap V_2$$

$$\Rightarrow x_1 - y_1 = \theta, x_2 - y_2 = \theta$$

$$\Rightarrow x_1 = y_1, x_2 = y_2$$

故 $z \in V_1 + V_2$ 的分解式唯一, 从而 $V_1 + V_2 = V_1 \oplus V_2$.

必要性. 若 $V_1 \cap V_2 \neq \{\theta\}$, 则有 $\theta \neq x \in V_1 \cap V_2$. 对于 $\theta \in V_1 + V_2$, 有

$$\theta = \theta + \theta, \theta \in V_1, \theta \in V_2$$

$$\theta = x + (-x), x \in V_1, (-x) \in V_2$$

即 $\theta \in V_1 + V_2$ 有两种不同的分解式. 这与 $V_1 + V_2$ 是直和矛盾.

故 $V_1 \cap V_2 = \{\theta\}$.

推论 1 $V_1 + V_2$ 是直和 $\Leftrightarrow \dim(V_1 + V_2) = \dim V_1 + \dim V_2$

推论 2 设 $V_1 + V_2$ 是直和, V_1 的基为 x_1, \dots, x_k , V_2 的基为 y_1, \dots, y_l ,

则 $V_1 + V_2$ 的基为 $x_1, \dots, x_k, y_1, \dots, y_l$.

证 因为 $V_1 + V_2 = L(x_1, \dots, x_k, y_1, \dots, y_l)$, 且

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 = k + l$$

所以 $x_1, \dots, x_k, y_1, \dots, y_l$ 线性无关, 故 $x_1, \dots, x_k, y_1, \dots, y_l$ 是 $V_1 + V_2$ 的基.

典型题解析

例 1 设多项式空间 $P_3[t]$ 的两个基为

$$(I) \quad f_1(t)=1, \quad f_2(t)=1+t, \quad f_3(t)=1+t+t^2, \quad f_4(t)=1+t+t^2+t^3$$

$$(II) \quad g_1(t)=1+t^2+t^3, \quad g_2(t)=t+t^2+t^3, \quad g_3(t)=1+t+t^2, \quad g_4(t)=1+t+t^3$$

(1) 求由基 (I) 改变为基 (II) 的过渡矩阵;

(2) 求 $P_3[t]$ 中在基 (I) 和基 (II) 下有相同坐标的全体多项式.

解 (1) 采用中介基法求过渡矩阵: 取 $P_3[t]$ 的简单基为 $1, t, t^2, t^3$, 写出由简单基改变为基 (I) 和基 (II) 过渡矩阵

$$C_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

由基 (I) 改变为基 (II) 的过渡矩阵为

$$C = C_1^{-1}C_2 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

(2) 设 $f(t) \in P_3[t]$ 在基 (I) 和基 (II) 下的坐标分别为

$$\alpha = (\xi_1, \xi_2, \xi_3, \xi_4)^T, \quad \beta = (\eta_1, \eta_2, \eta_3, \eta_4)^T$$

由坐标变换公式 $\alpha = C\beta$ 及题设 $\alpha = \beta$ 可得 $(I - C)\alpha = 0$

该齐次线性方程组的通解为 $\alpha = k(0, 0, 1, 0)^T \quad (\forall k \in \mathbb{R})$

在基 (I) 和基 (II) 下有相同坐标的全体多项式

$$\begin{aligned} f(t) &= (f_1(t), f_2(t), f_3(t), f_4(t))\alpha \\ &= k f_3(t) = k + kt + kt^2 \quad (\forall k \in \mathbb{R}) \end{aligned}$$

§ 1.2 线性变换及其矩阵

一、线性变换

1. 定义 线性空间 V , 数域 K , T 是 V 中的变换. 若对 $\forall x, y \in V, \forall k, l \in K$, 都有 $T(kx + ly) = k(Tx) + l(Ty)$, 称 T 是 V 中的线性变换.

性质 (1) $T(\theta) = T(0x + 0y) = 0(Tx) + 0(Ty) = \theta$

(2) $T(-x) = T((-1)x + 0y) = (-1)(Tx) + 0(Ty) = -(Tx)$

(3) $x_1, \dots, x_m \in V$ 线性相关 $\Rightarrow Tx_1, \dots, Tx_m$ 线性相关

(4) $x_1, \dots, x_m \in V$ 线性无关时, 不能推出 Tx_1, \dots, Tx_m 线性无关.

(5) T 是线性变换 $\Leftrightarrow T(x + y) = Tx + Ty, T(kx) = k(Tx)$

$(\forall x, y \in V, \forall k \in K)$

等价命题

例 1 矩阵空间 $R^{n \times n}$, 给定矩阵 $B_{n \times n}$, 则变换 $T(X) = BX + XB$ ($\forall X \in R^{n \times n}$) 是 $R^{n \times n}$ 的线性变换.

2. 线性变换的值域: $R(T) = \{y \mid y = Tx, x \in V\}$

3. 线性变换的核: $N(T) = \{x \mid Tx = \theta, x \in V\}$

Th8 设 T 是线性空间 V 的线性变换, 则 $R(T)$ 和 $N(T)$ 都是 V 的子空间.

证 (1) V 非空 $\Rightarrow R(T)$ 非空.

$$\forall y_1 \in R(T) \Rightarrow \exists x_1 \in V, \text{st } y_1 = Tx_1$$

$$\forall y_2 \in R(T) \Rightarrow \exists x_2 \in V, \text{st } y_2 = Tx_2$$

$$y_1 + y_2 = Tx_1 + Tx_2 = T(x_1 + x_2) \in R(T) \quad (\because x_1 + x_2 \in V)$$

$$ky_1 = k(Tx_1) = T(kx_1) \in R(T) \quad (\because \forall k \in K, kx_1 \in V)$$

故 $R(T)$ 是 V 的子空间.

(2) $\theta \in V, T\theta = \theta \Rightarrow \theta \in N(T)$, 即 $N(T)$ 非空.

$$\forall x, y \in N(T) \Rightarrow T(x + y) = Tx + Ty = \theta, \text{ 即 } x + y \in N(T).$$

$$\forall x \in N(T), \forall k \in K \Rightarrow T(kx) = k(Tx) = \theta, \text{ 即 } kx \in N(T).$$

故 $N(T)$ 是 V 的子空间.

[注] 定义: T 的秩 = $\dim R(T)$, T 的亏 = $\dim N(T)$

例 2 设线性空间 V^n 的基为 x_1, \dots, x_n , T 是 V^n 的线性变换, 则

$$R(T) = L(Tx_1, \dots, Tx_n), \dim R(T) + \dim N(T) = n$$

证 (1) 先证 $R(T) \subset L(Tx_1, \dots, Tx_n)$: $\forall y \in R(T) \Rightarrow \exists x \in V^n, \text{st } y = Tx$

$$x = c_1x_1 + \dots + c_nx_n \Rightarrow y = c_1(Tx_1) + \dots + c_n(Tx_n) \in L(Tx_1, \dots, Tx_n)$$

再证 $R(T) \supset L(Tx_1, \dots, Tx_n)$:

$$\forall y \in L(Tx_1, \dots, Tx_n) \Rightarrow \exists c_1, \dots, c_n, \text{st } y = c_1(Tx_1) + \dots + c_n(Tx_n)$$

$$x_i \in V^n \Rightarrow Tx_i \in R(T) \Rightarrow y = c_1(Tx_1) + \dots + c_n(Tx_n) \in R(T)$$

(2) 设 $\dim N(T) = m$, 且 $N(T)$ 的基为 y_1, \dots, y_m , 扩充为 V^n 的基:

$$y_1, \dots, y_m, y_{m+1}, \dots, y_n$$

特殊方法

则 $R(T) = L(Ty_1, \dots, Ty_m, Ty_{m+1}, \dots, Ty_n) = L(Ty_{m+1}, \dots, Ty_n)$

设数组 k_{m+1}, \dots, k_n 使得 $k_{m+1}(Ty_{m+1}) + \dots + k_n(Ty_n) = \theta$, 则

$$T(k_{m+1}y_{m+1} + \dots + k_n y_n) = \theta$$

因为 T 是线性变换, 所以 $k_{m+1}y_{m+1} + \dots + k_n y_n \in N(T)$, 故

$$k_{m+1}y_{m+1} + \dots + k_n y_n = l_1 y_1 + \dots + l_m y_m$$

即 $(-l_1)y_1 + \dots + (-l_m)y_m + k_{m+1}y_{m+1} + \dots + k_n y_n = \theta$

因为 $y_1, \dots, y_m, y_{m+1}, \dots, y_n$ 线性无关, 所以 $k_{m+1} = 0, \dots, k_n = 0$. 因此

Ty_{m+1}, \dots, Ty_n 线性无关, 从而 $\dim R(T) = n - m$, 即 $\dim R(T) + m = n$.

例 3 矩阵空间 $\mathbf{R}^{2 \times 2}$ 中, 给定矩阵 $B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, 线性变换 T 为

$$T(X) = BX - XB \quad (\forall X \in \mathbf{R}^{2 \times 2})$$

求 $R(T)$ 和 $N(T)$ 的基与维数.

解(1) 取 $\mathbf{R}^{2 \times 2}$ 的简单基 $E_{11}, E_{12}, E_{21}, E_{22}$, 计算

$$T(E_{11}) = BE_{11} - E_{11}B = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix}$$

$$T(E_{12}) = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix}, \quad T(E_{21}) = \begin{bmatrix} 2 & 0 \\ 2 & -2 \end{bmatrix}, \quad T(E_{22}) = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

基象组的坐标为

$$\gamma_1 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \quad \gamma_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \quad \gamma_3 = \begin{bmatrix} 2 \\ 0 \\ 2 \\ -2 \end{bmatrix}, \quad \gamma_4 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

该向量组的一个最大线性无关组为 γ_1, γ_3 .

故 $\dim R(T) = 2$, 且 $R(T)$ 的一个基为 $T(E_{11}), T(E_{21})$.

(2) 设 $X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \in N(T)$, 由 $T(X) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 可得

$$\begin{bmatrix} 2x_3 & -2x_1 - 2x_2 + 2x_4 \\ 2x_3 & -2x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 2x_3 = 0 \\ -2x_1 - 2x_2 + 2x_4 = 0 \\ 2x_3 = 0 \\ -2x_3 = 0 \end{cases} \quad \text{基础解系为} \quad \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ 线性无关}$$

故 $\dim N(T) = 2$, 且 $N(T)$ 的一个基为 A_1, A_2 .

隐含: 满足 $T(X) = O$ 的
线性无关矩阵最多两个!

4. 单位变换: 线性空间 V 中, 定义变换 T 为 $Tx = x \quad (\forall x \in V)$

则 T 是线性变换, 记作 T_e .

也称恒等变换

5. 零变换: 线性空间 V 中, 定义变换 T 为 $Tx = \theta \quad (\forall x \in V)$

则 T 是线性变换, 记作 T_0 .

6. 线性变换的运算: 线性空间 V , 数域 K , 线性变换 T_1 与 T_2 .

(1) 相等: 若 $T_1x = T_2x \quad (\forall x \in V)$, 称 $T_1 = T_2$.

(2) 加法: 定义变换 T 为 $Tx = T_1x + T_2x \quad (\forall x \in V)$

则 T 是线性变换, 记作 $T = T_1 + T_2$.

负变换: 定义变换 T 为 $Tx = -(T_1x) \quad (\forall x \in V)$

则 T 是线性变换, 记作 $T = -T_1$.

(3) 数乘: 给定 $k \in K$, 定义变换 T 为 $Tx = k(T_1x) \quad (\forall x \in V)$

则 T 是线性变换, 记作 $T = kT_1$.

[注] 集合 $\text{Hom}(V, V) \stackrel{\text{def}}{=} \{T \mid T \text{ 是数域 } K \text{ 上的线性空间 } V \text{ 的线性变换}\}$

按照线性运算(2)和(3)构成数域 K 上的线性空间, 称为 V 的同态.

(4) 乘法: 定义变换 T 为 $Tx = T_1(T_2x) \quad (\forall x \in V)$

则 T 是线性变换, 记作 $T = T_1T_2$.

对比复合
函数概念

7. 逆变换: 设 T 是线性空间 V 的线性变换, 若 V 的线性变换 S 满足

$$ST = TS = T_e \text{ 或者 } (ST)x = (TS)x = x \quad (\forall x \in V)$$

则称 T 为可逆变换, 且 S 为 T 的逆变换, 记作 $T^{-1} = S$.

8. 幂变换: 设 T 是线性空间 V 的线性变换, 则

$$T^m \stackrel{\text{def}}{=} T^{m-1}T \quad (m = 2, 3, \dots) \text{ 也是 } V \text{ 的线性变换.}$$

9. 多项式变换: 设 T 是线性空间 V 的线性变换, 多项式

$$f(t) = a_0 + a_1t + \dots + a_mt^m \quad (a_i \in K)$$

则 $f(T) = a_0T_e + a_1T + \dots + a_mT^m$ 也是 V 的线性变换.

二、线性变换的矩阵表示

1. 线性变换在给定基下的矩阵

设线性空间 V^n 的基为 x_1, \dots, x_n , T 是 V^n 的线性变换, 则 $Tx_i \in V^n$, 且有

$$\begin{cases} Tx_1 = a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n \\ Tx_2 = a_{12}x_1 + a_{22}x_2 + \dots + a_{n2}x_n \\ \dots\dots\dots \\ Tx_n = a_{1n}x_1 + a_{2n}x_2 + \dots + a_{nn}x_n \end{cases} \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

写成矩阵乘法形式 $T(x_1, \dots, x_n) \stackrel{\text{def}}{=} (Tx_1, \dots, Tx_n) = (x_1, \dots, x_n)A$

称 A 为线性变换 T 在基 x_1, \dots, x_n 下的矩阵.

[注] (1) 给定 V^n 的基 x_1, \dots, x_n 和线性变换 T 时, 矩阵 A 唯一.

(2) 给定 V^n 的基 x_1, \dots, x_n 和矩阵 A 时, 基象组 Tx_1, \dots, Tx_n 确定.

$\forall x \in V^n \Rightarrow x = c_1x_1 + \dots + c_nx_n$, 定义变换

$$Tx = c_1(Tx_1) + \dots + c_n(Tx_n)$$

则 T 是线性变换. 因此线性变换 T 与方阵 A 是一一对应关系.

(3) 单位变换在任何基下的矩阵为单位矩阵.

(4) 零变换在任何基下的矩阵为零矩阵.

例 4 在 $\mathbb{R}^{2 \times 2}$ 中, 给定矩阵 $B = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix}$, 线性变换 $T(X) = XB$ ($\forall X \in \mathbb{R}^{2 \times 2}$).

基(I): $E_{11}, E_{12}, E_{21}, E_{22}$

$$\text{基像组 } T(E_{11}) = E_{11}B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad T(E_{12}) = E_{12}B = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$$

$$T(E_{21}) = E_{21}B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad T(E_{22}) = E_{22}B = \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}$$

$$T \text{ 在基(I)下的矩阵为 } A_1 = \begin{bmatrix} 0 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{基(II): } B_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\text{基像组 } T(B_1) = B_1B = \begin{bmatrix} 4 & 1 \\ 4 & 1 \end{bmatrix} = 1B_1 + 3B_2 - 3B_3 + 3B_4$$

$$T(B_2) = B_2B = \begin{bmatrix} 4 & 1 \\ 0 & 1 \end{bmatrix} = 1B_1 - 1B_2 + 1B_3 + 3B_4$$

$$T(B_3) = B_3B = \begin{bmatrix} 4 & 1 \\ 0 & 0 \end{bmatrix} = 0B_1 + 0B_2 + 1B_3 + 3B_4$$

$$T(B_4) = B_4 B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0B_1 + 0B_2 + 1B_3 - 1B_4$$

$$T \text{ 在基(II)下的矩阵为 } A_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ -3 & 1 & 1 & 1 \\ 3 & 3 & 3 & -1 \end{bmatrix}, \text{ 易见 } A_1 \neq A_2.$$

[注] 混合法求 A_2 :

$$(B_1, B_2, B_3, B_4) = (E_{11}, E_{12}, E_{21}, E_{22})C$$

$$(T(B_1), \dots, T(B_4)) = (E_{11}, E_{12}, E_{21}, E_{22})A_0 = (B_1, \dots, B_4)C^{-1}A_0$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 4 & 4 & 4 & 0 \\ 1 & 1 & 1 & 1 \\ 4 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad A_2 = C^{-1}A_0$$

例 5 线性空间 V^n 中, 设线性变换 T 在基 x_1, \dots, x_n 下的矩阵为 A , 则

$$\dim R(T) = \text{rank} A, \quad \dim N(T) = n - \text{rank} A.$$

证 $\text{rank} A = m \Leftrightarrow A$ 的列向量组 β_1, \dots, β_n 中最大无关组含 m 个向量

\Leftrightarrow 元素组 Tx_1, \dots, Tx_n 中最大无关组含 m 个向量

$$\Leftrightarrow \dim R(T) = \dim L(Tx_1, \dots, Tx_n) = m$$

由例 2 知另一结论成立.

例 求 $N(T)$ 基的一般方法. 设线性变换 T 在基 x_1, \dots, x_n 下的矩阵为 A , 则

$$T(x) = \theta \Leftrightarrow T[(x_1, \dots, x_n)\alpha] = \theta \Leftrightarrow (x_1, \dots, x_n)A\alpha = \theta \Leftrightarrow A\alpha = 0$$

求 $A\alpha = 0$ 的基础解系: $\alpha_1, \dots, \alpha_{n-r}$ ($r = \text{rank} A$)

构造 $N(T)$ 的基: $y_1 = (x_1, \dots, x_n)\alpha_1, \dots, y_{n-r} = (x_1, \dots, x_n)\alpha_{n-r}$

2. 线性变换运算的矩阵表示 (将线性变换运算转化为矩阵运算)

Th9 线性空间 V^n 的基为 x_1, \dots, x_n , 线性变换 T_1 与 T_2 的矩阵为 A 与 B , 则

(1) $T_1 + T_2$ 在该基下的矩阵为 $A + B$; (2) kT_1 在该基下的矩阵为 kA ;

(3) $T_1 T_2$ 在该基下的矩阵为 AB ; (4) T_1^{-1} 在该基下的矩阵为 A^{-1} .

证 $T_1(x_1, \dots, x_n) = (x_1, \dots, x_n)A$, $T_2(x_1, \dots, x_n) = (x_1, \dots, x_n)B$

$$\begin{aligned} (1) \quad (T_1 + T_2)(x_1, \dots, x_n) &= ((T_1 + T_2)x_1, \dots, (T_1 + T_2)x_n) \\ &= (T_1 x_1 + T_2 x_1, \dots, T_1 x_n + T_2 x_n) = (T_1 x_1, \dots, T_1 x_n) + (T_2 x_1, \dots, T_2 x_n) \\ &= (x_1, \dots, x_n)A + (x_1, \dots, x_n)B = (x_1, \dots, x_n)(A + B) \end{aligned}$$

(2) 略.

(3) 先证: $\forall C = (c_{ij})_{n \times m}$, $T[(x_1, \dots, x_n)C] = [T(x_1, \dots, x_n)]C$

$$\text{左} = T\left[\sum c_{i1}x_i, \dots, \sum c_{im}x_i\right] = \left[\sum c_{i1}(Tx_i), \dots, \sum c_{im}(Tx_i)\right]$$

常用结论

$$= (Tx_1, \dots, Tx_n)C = \text{右}$$

$$\begin{aligned} \text{由此可得 } (T_1 T_2)(x_1, \dots, x_n) &= T_1[T_2(x_1, \dots, x_n)] = T_1[(x_1, \dots, x_n)B] \\ &= [T_1(x_1, \dots, x_n)]B = (x_1, \dots, x_n)AB \end{aligned}$$

$$(4) \text{ 记 } T_1^{-1} = T_2, \text{ 则 } T_1 T_2 = T_2 T_1 = T_e \stackrel{(3)}{\Rightarrow} AB = BA = I \Rightarrow B = A^{-1}.$$

3. 象与原象坐标间的关系

Th10 线性空间 V^n 的基为 x_1, \dots, x_n , 线性变换 T 在该基下的矩阵为 A ,

$x \in V^n$ 的坐标为 α , $T(x)$ 的坐标为 β , 则 $\beta = A\alpha$.

证 $x = \xi_1 x_1 + \dots + \xi_n x_n = (x_1, \dots, x_n)\alpha$

$$T(x) = T[(x_1, \dots, x_n)\alpha] = (x_1, \dots, x_n)A\alpha, \text{ 由定理 2 知 } \beta = A\alpha.$$

4. 线性变换在不同基下矩阵之间的关系

Th11 线性空间 V^n 的基(I): x_1, \dots, x_n , 基(II): y_1, \dots, y_n , 线性变换 T :

$$T(x_1, \dots, x_n) = (x_1, \dots, x_n)A, \quad T(y_1, \dots, y_n) = (y_1, \dots, y_n)B$$

由基(I)到基(II)的过渡矩阵为 C , 则 $B = C^{-1}AC$.

证 因为 $T(y_1, \dots, y_n) = T(x_1, \dots, x_n)C$

$$= (x_1, \dots, x_n)AC = (y_1, \dots, y_n)C^{-1}AC$$

所以 $B = C^{-1}AC$.

核心结论

三、线性变换的特征值与特征向量

1. 定义 线性空间 V , 线性变换 T , 若 $\lambda_0 \in K$ 及 $\theta \neq x \in V$ 满足 $T(x) = \lambda_0 x$, 则称 λ_0 为 T 的特征值, x 为 T 的对应于 λ_0 的特征向量 (元素).

2. 算法 设线性空间 V^n 的基为 x_1, \dots, x_n , 线性变换 T 的矩阵为 $A_{n \times n}$.

T 的特征值为 λ_0 , 对应的特征向量为 x .

x 的坐标为 α , $T(x)$ 的坐标为 $A\alpha$, $\lambda_0 x$ 的坐标为 $\lambda_0 \alpha$.

因为 $T(x) = \lambda_0 x \Leftrightarrow A\alpha = \lambda_0 \alpha$, 所以 T 的特征值与 A 的特征值相同;

T 的对应于 λ_0 的特征向量的坐标就是 A 的对应于 λ_0 的特征向量.

例 6 设 $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, 线性空间 $V = \{X = (x_{ij})_{2 \times 2} \mid x_{11} + x_{22} = 0, x_{ij} \in \mathbb{R}\}$,

线性变换为 $TX = B^T X - X^T B$ ($X \in V$), 求 T 的特征值与特征向量.

$$\text{解 } X \in V \Rightarrow X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{bmatrix} = \begin{bmatrix} x_{11} & 0 \\ 0 & -x_{11} \end{bmatrix} + \begin{bmatrix} 0 & x_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ x_{21} & 0 \end{bmatrix}$$

$$= x_{11} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + x_{12} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + x_{21} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

坐标:

$$\begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \end{bmatrix}$$

可得 V 的简单基为 $X_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $X_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

由公式求得 $TX_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $TX_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $TX_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

故 T 在简单基下的矩阵为 $A = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$

A 的特征值与线性无关的特征向量为

$$\lambda_1 = \lambda_2 = 0, \alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}; \quad \lambda_3 = 2, \alpha_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

T 的特征值与线性无关的特征向量为

$$\lambda_1 = \lambda_2 = 0, Y_1 = (X_1, X_2, X_3)\alpha_1 = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$Y_2 = (X_1, X_2, X_3)\alpha_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\lambda_3 = 2, Y_3 = (X_1, X_2, X_3)\alpha_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

α_i 不是 T 的特征向量

例 7 线性空间 V , 线性变换 T , $V_{\lambda_0} = \{x | Tx = \lambda_0 x, x \in V\}$ 是 V 的子空间.

证 $\theta \in V, T\theta = \lambda_0 \theta \Rightarrow \theta \in V_{\lambda_0}$, 即 V_{λ_0} 非空.

$$\forall x, y \in V_{\lambda_0} \Rightarrow T(x+y) = Tx + Ty = \lambda_0 x + \lambda_0 y = \lambda_0(x+y)$$

$$\Rightarrow x+y \in V_{\lambda_0}$$

$$\forall k \in K, \forall x \in V_{\lambda_0} \Rightarrow T(kx) = k(Tx) = k(\lambda_0 x) = \lambda_0(kx) \Rightarrow kx \in V_{\lambda_0}$$

故 V_{λ_0} 是 V 的子空间.

[注] 若 λ_0 是线性变换的特征值, 则称 V_{λ_0} 为 T 的特征子空间.

3. 矩阵的迹: $A = (a_{ij})_{n \times n}$, $\text{tr} A = \sum_{i=1}^n a_{ii}$.

Th12 $A_{m \times n}, B_{n \times m} \Rightarrow \text{tr}(AB) = \text{tr}(BA)$.

证 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times m}$, $AB = (u_{ij})_{m \times m}$, $BA = (v_{ij})_{n \times n}$:

$$u_{ii} = (a_{i1}, \dots, a_{in}) \begin{pmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{pmatrix} = \sum_{k=1}^n a_{ik} b_{ki}, \quad v_{kk} = (b_{k1}, \dots, b_{km}) \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} = \sum_{i=1}^m b_{ki} a_{ik}$$

$$\text{tr}(AB) = \sum_{i=1}^m u_{ii} = \sum_{i=1}^m \left[\sum_{k=1}^n a_{ik} b_{ki} \right] = \sum_{k=1}^n \left[\sum_{i=1}^m b_{ki} a_{ik} \right] = \sum_{k=1}^n v_{kk} = \text{tr}(BA)$$

Th13 若 A 相似于 B , 则 $\text{tr}A = \text{tr}B$.

证 由 $B = P^{-1}AP$ 可得 $\text{tr}B = \text{tr}(P^{-1}AP) = \text{tr}((AP)P^{-1}) = \text{tr}A$

[注] 因为相似矩阵有相同的特征值 (Th14 — 线性代数课程结论)

所以线性变换的特征值与线性空间中基的选取无关

4. 三角相似

Th17 $A_{n \times n}$ 相似于上三角矩阵.

证 归纳法. $n=1$ 时, $A = (a_{11})$ 是上三角矩阵 $\Rightarrow A$ 相似于上三角矩阵.

假设 $n=k-1$ 时定理成立, 下证 $n=k$ 时定理也成立.

$A_{k \times k}$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$, 对应 λ_1 的特征向量为 $x_1 \Rightarrow Ax_1 = \lambda_1 x_1$.

扩充 x_1 为 C^k 的基: x_1, x_2, \dots, x_k (列向量)

$P_1 = (x_1, x_2, \dots, x_k)$ 可逆, $AP_1 = (Ax_1, Ax_2, \dots, Ax_k)$

$Ax_j \in C^k \Rightarrow Ax_j = b_{1j}x_1 + b_{2j}x_2 + \dots + b_{kj}x_k \quad (j=2, \dots, k)$

$$AP_1 = (x_1, x_2, \dots, x_k) \begin{bmatrix} \lambda_1 & b_{12} & \dots & b_{1k} \\ 0 & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & & \vdots \\ 0 & b_{k2} & \dots & b_{kk} \end{bmatrix}, \quad P_1^{-1}AP_1 = \left[\begin{array}{c|ccc} \lambda_1 & b_{12} & \dots & b_{1k} \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array} \right]$$

A_1 的特征值为 $\lambda_2, \dots, \lambda_k$, 由假设知, 存在 $k-1$ 阶可逆矩阵 Q 使得

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & \dots & * \\ & \ddots & \vdots \\ & & \lambda_k \end{bmatrix}, \quad P_2 = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array} \right]$$

$$P \stackrel{\Delta}{=} P_1P_2 \Rightarrow P^{-1}AP = P_2^{-1}(P_1^{-1}AP_1)P_2 = \dots = \begin{bmatrix} \lambda_1 & * & \dots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_k \end{bmatrix}$$

由归纳法原理, 对任意 n , 定理成立.

5. Hamilton-Cayley 定理

Th18 设 $A_{n \times n}$, $\varphi(\lambda) = \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$, 则

$$\varphi(A) = A^n + a_1A^{n-1} + \dots + a_{n-1}A + a_nI = O_{n \times n}$$

证 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n \Rightarrow \varphi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$.

$$\text{由 Th17 知, 存在可逆矩阵 } P_{n \times n}, \text{ 使得 } P^{-1}AP = \begin{bmatrix} \lambda_1 & \dots & * \\ & \ddots & \vdots \\ & & \lambda_n \end{bmatrix}.$$

$$\varphi(P^{-1}AP) = (P^{-1}AP - \lambda_1 I)(P^{-1}AP - \lambda_2 I) \cdots (P^{-1}AP - \lambda_n I)$$

$$\begin{aligned}
&= \begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2 - \lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n - \lambda_1 \end{bmatrix} \begin{bmatrix} \lambda_1 - \lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n - \lambda_2 \end{bmatrix} \cdots \\
&= \begin{bmatrix} \lambda_1 - \lambda_n & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \lambda_{n-1} - \lambda_n & * \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & \vdots \\ \vdots & \vdots & & \ddots & * \\ 0 & 0 & & & * \end{bmatrix} \begin{bmatrix} \lambda_1 - \lambda_3 & * & * & \cdots & * \\ & \lambda_2 - \lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_n - \lambda_3 \end{bmatrix} \cdots \\
&= \begin{bmatrix} \lambda_1 - \lambda_n & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \lambda_{n-1} - \lambda_n & * \\ & & & 0 \end{bmatrix} = O
\end{aligned}$$

即 $P^{-1}\varphi(A)P = O \Rightarrow \varphi(A) = O$.

[注] (1) $|A| \neq 0 \Rightarrow a_n \neq 0, A^{-1} = -\frac{1}{a_n}(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-2}A + a_{n-1}I)$

(2) $A^n \in \text{span}\{A^{n-1}, \cdots, A, I\}$

矩阵多项式
表示逆矩阵

例 8 $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix}$, 计算 $A^{100} + 2A^{50}$.

解 $f(\lambda) = \lambda^{100} + 2\lambda^{50}$, $\varphi(\lambda) = \det(\lambda I - A) = (\lambda - 1)^2(\lambda - 2)$

φ 除 f : $f(\lambda) = \varphi(\lambda)g(\lambda) + (b_0 + b_1\lambda + b_2\lambda^2)$

$f'(\lambda) = [\varphi(\lambda)g(\lambda)]' + (b_1 + 2b_2\lambda)$

由 $f(1) = 3, f'(1) = 200, f(2) = 2^{100} + 2^{51}$ 可得

$$\left. \begin{aligned} b_0 + b_1 + b_2 &= 3 \\ b_1 + 2b_2 &= 200 \\ b_0 + 2b_1 + 4b_2 &= 2^{100} + 2^{51} \end{aligned} \right\} \Rightarrow \begin{cases} b_0 = 2^{100} + 2^{51} - 400 \\ b_1 = -2^{101} - 2^{52} + 606 \\ b_2 = 2^{100} + 2^{51} - 203 \end{cases}$$

$\varphi(A) = O \Rightarrow f(A) = b_0I + b_1A + b_2A^2$

6. 最小多项式

以 $A_{n \times n}$ 为根, 且次数最低的首 1 多项式, 记作 $m(\lambda)$.

$f(\lambda) = 1 \Rightarrow f(A) = I \neq O \Rightarrow \partial m(\lambda) \geq 1$

$$\varphi(\lambda) = \det(\lambda I - A), \text{ Th18} \Rightarrow \varphi(A) = O \Rightarrow \partial m(\lambda) \leq n$$

$$\text{例 9 } A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}, \quad \varphi(\lambda) = (\lambda - 2)^2(\lambda - 4)$$

$$f_1(\lambda) = \lambda + k (\forall k \in \mathbf{R}): f_1(A) = A + kI \neq O \Rightarrow \partial m(\lambda) > 1$$

$$f_2(\lambda) = (\lambda - 2)(\lambda - 4): f_2(A) = (A - 2I)(A - 4I) = O \Rightarrow m(\lambda) = f_2(\lambda)$$

Th19 (1) 多项式 $f(\lambda)$ 满足 $f(A) = O \Rightarrow m(\lambda) \mid f(\lambda)$;

(2) $m(\lambda)$ 唯一.

证 (1) 反证法.

$$m(\lambda) \nmid f(\lambda) \Rightarrow f(\lambda) = m(\lambda)g(\lambda) + r(\lambda), \quad r(\lambda) \neq 0 \text{ 且 } \partial r(\lambda) < \partial m(\lambda)$$

$$\Rightarrow f(A) = m(A)g(A) + r(A)$$

$$f(A) = O, m(A) = O \Rightarrow r(A) = O, \partial r(\lambda) < \partial m(\lambda)$$

$\Rightarrow m(\lambda)$ 不是 A 的最小多项式, 矛盾!

(2) 设 $m(\lambda)$ 与 $\tilde{m}(\lambda)$ 都是 A 的最小多项式, 则

$$\left. \begin{array}{l} \tilde{m}(A) = O \Rightarrow m(\lambda) \mid \tilde{m}(\lambda) \\ m(A) = O \Rightarrow \tilde{m}(\lambda) \mid m(\lambda) \end{array} \right\} \Rightarrow m(\lambda) = \tilde{m}(\lambda)$$

Th20 $m(\lambda)$ 与 $\varphi(\lambda)$ 的零点相同 (不计重数).

证 Th19 $\Rightarrow m(\lambda)$ 的零点是 $\varphi(\lambda)$ 的零点. 再设 λ_0 是 $\varphi(\lambda)$ 的零点, 则有

$$Ax = \lambda_0 x (x \neq 0) \Rightarrow m(A)x = m(\lambda_0)x$$

$$m(A) = O \Rightarrow m(\lambda_0)x = 0 \Rightarrow m(\lambda_0) = 0, \text{ 故 } \lambda_0 \text{ 也是 } m(\lambda) \text{ 的零点.}$$

[注] Th20 $\Rightarrow m(\lambda)$ 一定含 $\varphi(\lambda)$ 的全部单因式.

但 $m(\lambda)$ 不一定是 $\varphi(\lambda)$ 的全部单因式的乘积.

$$\text{例如: } A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \varphi(\lambda) = (\lambda - 1)^2, \quad m(\lambda) \neq (\lambda - 1).$$

7. 最小多项式求法

Th21 对 $A_{n \times n}$, 设 $\lambda I - A$ 的第 i 行第 j 列元素的余子式为 $M_{ij}(\lambda)$, 记

$$d(\lambda) = \max_{i,j} [M_{ij}(\lambda)], \text{ 则 } m(\lambda) = \frac{\det(\lambda I - A)}{d(\lambda)}.$$

$$\text{例 10 设 } A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}, \text{ 求 } m(\lambda).$$

$$\text{解 } \lambda I - A = \begin{bmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{bmatrix}, \quad \varphi(\lambda) = \det(\lambda I - A) = (\lambda - 2)^2(\lambda - 4)$$

$$\begin{aligned}
 M_{11} &= \lambda^2 - 5\lambda + 6, \quad M_{21} = 3(\lambda - 2), \quad M_{31} = 2(\lambda - 2) \\
 M_{12} &= \lambda - 2, \quad M_{22} = \lambda^2 - 3\lambda + 2, \quad M_{32} = 2(\lambda - 2) \\
 M_{13} &= -(\lambda - 2), \quad M_{23} = -3(\lambda - 2), \quad M_{33} = \lambda^2 - 8\lambda + 12 \\
 d(\lambda) &= \lambda - 2, \quad m(\lambda) = \frac{\varphi(\lambda)}{d(\lambda)} = (\lambda - 2)(\lambda - 4) = \lambda^2 - 6\lambda + 8
 \end{aligned}$$

例 11 $A_{n \times n}$ 相似于 $B_{n \times n} \Rightarrow m_A(\lambda) = m_B(\lambda)$.

证 $B = P^{-1}AP \Rightarrow A = PBP^{-1}$

取 $f(\lambda) = m_A(\lambda)$, 则 $f(A) = m_A(A) = O$, 从而有

$$f(B) = f(P^{-1}AP) = P^{-1}f(A)P = O$$

$$\stackrel{\text{Th19}}{\Rightarrow} m_B(\lambda) \mid f(\lambda), \text{ 即 } m_B(\lambda) \mid m_A(\lambda) \quad (1)$$

取 $g(\lambda) = m_B(\lambda)$, 则 $g(B) = m_B(B) = O$, 从而有

$$g(A) = g(PBP^{-1}) = P g(B) P^{-1} = O$$

$$\stackrel{\text{Th19}}{\Rightarrow} m_A(\lambda) \mid g(\lambda), \text{ 即 } m_A(\lambda) \mid m_B(\lambda) \quad (2)$$

①+②得: $m_A(\lambda) = m_B(\lambda)$.

四、对角矩阵

Th24 在线性空间 V^n 中, 线性变换 T 在某基下的矩阵为对角矩阵

$\Leftrightarrow T$ 有 n 个线性无关的特征向量 (元素).

证 必要性. 设 V^n 的基为 x_1, \dots, x_n , 且 $T(x_1, \dots, x_n) = (x_1, \dots, x_n)A$,

$$A = \text{diag}(\lambda_1, \dots, \lambda_n), \text{ 则有 } (Tx_1, \dots, Tx_n) = (\lambda_1 x_1, \dots, \lambda_n x_n)$$

$$\Rightarrow Tx_j = \lambda_j x_j \quad (j = 1, 2, \dots, n)$$

$\Rightarrow x_1, \dots, x_n$ 是 T 的 n 个线性无关的特征向量

充分性. 设 T 有 n 个线性无关的特征向量 y_1, \dots, y_n , 即

$$T y_j = \lambda_j y_j \quad (j = 1, 2, \dots, n)$$

取 y_1, \dots, y_n 为 V^n 的基, 则有

$$T(y_1, \dots, y_n) = (T y_1, \dots, T y_n) = (\lambda_1 y_1, \dots, \lambda_n y_n)$$

$$= (y_1, \dots, y_n) \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

线性变换在这个基下的矩阵是对角形

Th25 $A_{n \times n}$ 相似于对角矩阵 $\Leftrightarrow A$ 有 n 个线性无关的特征向量 (列向量).

证 A 相似于 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$

\Leftrightarrow 存在可逆矩阵 $P = (x_1, \dots, x_n)$, 使得 $P^{-1}AP = \Lambda$

$\Leftrightarrow A(x_1, \dots, x_n) = (x_1, \dots, x_n)\Lambda \Leftrightarrow Ax_j = \lambda_j x_j \quad (j = 1, 2, \dots, n)$

$\Leftrightarrow A$ 有 n 个线性无关的特征向量 x_1, \dots, x_n

Th26 $A_{n \times n}$ 有 n 个互异的特征值 $\Rightarrow A$ 相似于对角矩阵.

问题: 线性空间 V^n 的基 x_1, \dots, x_n , 线性变换 T 在该基下的矩阵 A 相似于 $A = \text{diag}(\lambda_1, \dots, \lambda_n)$, 确定 V^n 的新基 y_1, \dots, y_n , 使得 T 在新基下的矩阵为 A .

基变换公式

算法: 求 P 使 $P^{-1}AP = A$, 令 $(y_1, \dots, y_n) = (x_1, \dots, x_n)P$, 则有

$$\begin{aligned} T(y_1, \dots, y_n) &= T(x_1, \dots, x_n)P = (x_1, \dots, x_n)AP \\ &= (y_1, \dots, y_n)P^{-1}AP = (y_1, \dots, y_n)A \end{aligned}$$

例 12 在 $\mathbb{R}^{2 \times 2}$ 中, 给定 $B = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix}$, 线性变换 $T(X) = XB \quad (\forall X \in \mathbb{R}^{2 \times 2})$,

求 $\mathbb{R}^{2 \times 2}$ 的一个基, 使线性变换 T 在该基下的矩阵为对角矩阵.

解 取 $\mathbb{R}^{2 \times 2}$ 的简单基 $E_{11}, E_{12}, E_{21}, E_{22}$, 求得 T 在该基下的矩阵为

$$A = \begin{bmatrix} 0 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{见例 4})$$

求 P 使得 $P^{-1}AP = A$:

$$A = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & -2 & \\ & & & -2 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

由 $(B_1, B_2, B_3, B_4) = (E_{11}, E_{12}, E_{21}, E_{22})P$ 可得

$$B_1 = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}, B_4 = \begin{bmatrix} 0 & 0 \\ 2 & -1 \end{bmatrix}$$

故 T 在基 B_1, B_2, B_3, B_4 下的矩阵为 A .

典型题解析

例 3 设矩阵空间 $\mathbb{R}^{2 \times 2}$ 的子空间为

$$V = \{X = (x_{ij})_{2 \times 2} \mid x_{11} + x_{12} + x_{21} = 0, x_{ij} \in \mathbb{R}\}$$

V 中的线性变换为 $T(X) = X + X^T \quad (\forall X \in V)$, 求 V 的一个基, 使 T 在该基下的矩阵为对角矩阵.

解 (1) 先求 V 的简单基

$$\begin{aligned} X \in V &\Rightarrow X = \begin{bmatrix} -x_{12} - x_{21} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} -x_{12} & x_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -x_{21} & 0 \\ x_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & x_{22} \end{bmatrix} \\ &= x_{12} \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + x_{21} \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} + x_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

容易验证 $X_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$, $X_2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$, $X_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ 线性无关, 故

X_1, X_2, X_3 是 V 的简单基.

(2) 由公式计算

$$T(X_1) = X_1 + X_1^T = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} = 1X_1 + 1X_2 + 0X_3$$

$$T(X_2) = X_2 + X_2^T = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} = 1X_1 + 1X_2 + 0X_3$$

$$T(X_3) = X_3 + X_3^T = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} = 0X_1 + 0X_2 + 2X_3$$

$$T \text{ 在基 } X_1, X_2, X_3 \text{ 下的矩阵为 } A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\text{求 } P \text{ 使得 } P^{-1}AP = \Lambda: \quad A = \begin{bmatrix} 0 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

由 $(Y_1, Y_2, Y_3) = (X_1, X_2, X_3)P$ 求得

$$Y_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, Y_2 = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}, Y_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

故 T 在基 Y_1, Y_2, Y_3 下的矩阵为 Λ .

五、不变子空间

线性空间 V , 子空间 V_1 , 线性变换 T .

若对 $\forall x \in V_1$, 有 $Tx \in V_1$, 称 V_1 是 T 的不变子空间.

[注] V_1 是 T 的不变子空间时, 可将 T 看作 V_1 中的线性变换.

例 ① 子空间 $V_{\lambda_0} = \{x | Tx = \lambda_0 x, x \in V\}$ 是 T 的不变子空间.

$$\because \forall x \in V_{\lambda_0} \Rightarrow Tx = \lambda_0 x \in V_{\lambda_0}$$

② 子空间 $R(T)$ 是 T 的不变子空间.

$$\because \forall x \in R(T) \subset V \Rightarrow Tx \in R(T)$$

③ 子空间 $N(T)$ 是 T 的不变子空间.

$$\because \forall x \in N(T) \Rightarrow Tx = \theta \in N(T)$$

④ V_1 与 V_2 是 T 的不变子空间 $\Rightarrow V_1 \cap V_2, V_1 + V_2$ 亦是 T 的不变子空间.

$$1^\circ \quad \forall x \in V_1 \cap V_2 \Rightarrow \begin{cases} x \in V_1, Tx \in V_1 \\ x \in V_2, Tx \in V_2 \end{cases} \Rightarrow Tx \in V_1 \cap V_2$$

$$2^\circ \quad \forall x \in V_1 + V_2 \Rightarrow x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2$$

子空间相对于线性空间
不变子空间相对于线性变换

$$\Rightarrow Tx = Tx_1 + Tx_2, Tx_1 \in V_1, Tx_2 \in V_2$$

$$\Rightarrow Tx \in V_1 + V_2$$

Th27 线性空间 V^n , 线性变换 T , 设 V_1 与 V_2 是 T 的不变子空间, 且

$V^n = V_1 \oplus V_2$, T 在 V_1 的基 x_1, \dots, x_{n_1} 下的矩阵为 A_1 , T 在 V_2 的

基 y_1, \dots, y_{n_2} 下的矩阵为 A_2 , 则 T 在 V^n 的基 $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$

$$\text{下的矩阵为 } A = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}.$$

证 因为 $(Tx_1, \dots, Tx_{n_1}) = (x_1, \dots, x_{n_1})A_1$, $(Ty_1, \dots, Ty_{n_2}) = (y_1, \dots, y_{n_2})A_2$

$$\text{所以 } T(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = [(Tx_1, \dots, Tx_{n_1}), (Ty_1, \dots, Ty_{n_2})]$$

$$= [(x_1, \dots, x_{n_1})A_1, (y_1, \dots, y_{n_2})A_2]$$

$$= [(x_1, \dots, x_{n_1}), (y_1, \dots, y_{n_2})] \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix} = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2})A$$

[注] 若 T 在 V^n 的基 $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ 下的矩阵 $A = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}$, 则

$V_1 = L(x_1, \dots, x_{n_1})$ 与 $V_2 = L(y_1, \dots, y_{n_2})$ 都是 T 的不变子空间, 且

$$V^n = V_1 \oplus V_2.$$

试题八(M2002B)

设线性空间 V^3 的一个基为 x_1, x_2, x_3 , 线性变换 T 在该基下的矩阵为

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 3 & 1 \\ 1 & -1 & 1 \end{bmatrix}. \text{ 令 } y_1 = x_1 + x_2, y_2 = x_1 + x_2 + x_3, \text{ 证明: 子空间}$$

$W = L(y_1, y_2)$ 是 T 的不变子空间.

$$\text{证 } T(y_1) = T(x_1, x_2, x_3) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = (x_1, x_2, x_3)A \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = (x_1, x_2, x_3) \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = 2y_1 \in W$$

$$T(y_2) = T(x_1, x_2, x_3) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (x_1, x_2, x_3)A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (x_1, x_2, x_3) \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = 2y_1 + y_2 \in W$$

对任意的 $y \in W$, 存在常数 k_1, k_2 使得 $y = k_1 y_1 + k_2 y_2$, 且有

$$T(y) = k_1 \cdot T(y_1) + k_2 \cdot T(y_2) \in W$$

故 W 是 T 的不变子空间.

试题八(M2002A)

设 V 是数域 K 上的线性空间, T_1 和 T_2 是线性空间 V 中的两个线性变换, 且满足 $T_1 T_2 = T_2 T_1$, 证明:

请勿上传公共网站

关于 T_1 的
不变子空间

1. T_2 的值域 $R(T_2)$ 是 T_1 的不变子空间;
2. T_2 的核 $N(T_2)$ 也是 T_1 的不变子空间.

证 1. $\forall y \in R(T_2) \Rightarrow \exists x \in V, \text{st } y = T_2(x)$

$$T_1(y) = T_1[T_2(x)] = (T_1 T_2)(x) = (T_2 T_1)(x) = T_2[T_1(x)] \in R(T_2)$$

故 $R(T_2)$ 是 T_1 的不变子空间.

2. $\forall x \in N(T_2) \Rightarrow T_2(x) = \theta$

$$\text{因为 } T_2[T_1(x)] = (T_2 T_1)(x) = (T_1 T_2)(x) = T_1[T_2(x)] = T_1(\theta) = \theta$$

所以 $T_1(x) \in N(T_2)$. 故 $N(T_2)$ 是 T_1 的不变子空间.

六、Jordan 标准形

1. λ -矩阵: $A(\lambda) = (a_{ij}(\lambda))_{n \times n}$, $a_{ij}(\lambda)$ 是 λ 的多项式函数.

$A(\lambda)$ 的秩: $A(\lambda)$ 中不恒等于零的子式的最高阶数.

λ -矩阵的初等变换: 行变换 列变换

(1) 对调: $r_i \leftrightarrow r_j$ $c_i \leftrightarrow c_j$

(2) 数乘 ($k \neq 0$): kr_i kc_i

(3) 倍加 ($p(\lambda)$ 是多项式): $r_i + p(\lambda)r_j$ $c_i + p(\lambda)c_j$

2. 行列式因子: $D_k(\lambda) = \text{最大公因式} \{ A(\lambda) \text{ 的所有 } k \text{ 阶子式} \}$

$$\text{不变因子: } d_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} \quad (D_0(\lambda) = 1, k = 1, 2, \dots, n)$$

$$\text{满足 } d_k(\lambda) \mid d_{k+1}(\lambda) \quad (k = 1, 2, \dots, n-1)$$

初等因子: 次数大于零的 $d_k(\lambda)$ 的不可约因式

[注] 考虑 λ -矩阵 $\lambda I - A$ 可得 A 的最小多项式 $m(\lambda) = d_n(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}$

例 13 $A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$, 求 $\lambda I - A$ 的全体初等因子.

解 $\lambda I - A = \begin{bmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{bmatrix}$, $D_1(\lambda) = 1$ (用特殊方法求 $D_2(\lambda)$)

$$\text{因为 } \begin{vmatrix} 4 & \lambda-3 \\ -1 & 0 \end{vmatrix} = \lambda-3 \text{ 与 } \begin{vmatrix} 4 & 0 \\ -1 & \lambda-2 \end{vmatrix} = 4(\lambda-2) \text{ 互质}$$

$$\text{所以 } D_2(\lambda) = 1, D_3(\lambda) = \det(\lambda I - A) = (\lambda-2)(\lambda-1)^2$$

$$\text{不变因子为 } d_1(\lambda) = 1, d_2(\lambda) = 1, d_3(\lambda) = (\lambda-1)^2(\lambda-2)$$

$$\text{全体初等因子为 } (\lambda-1)^2, \lambda-2.$$

3. 初等变换法求初等因子

$$A(\lambda) \rightarrow \begin{bmatrix} f_1(\lambda) & & \\ & \ddots & \\ & & f_n(\lambda) \end{bmatrix} \quad (f_k(\lambda) \text{ 是首 1 多项式, 或者 } 0)$$

次数大于零的 $f_k(\lambda)$ 的不可约因式为 $A(\lambda)$ 的初等因子

[注] 当 $f_k(\lambda) \mid f_{k+1}(\lambda)$ 时, $d_k(\lambda) = f_k(\lambda)$.

例如: 在例 13 中

$$\begin{aligned} \lambda I - A &= \begin{bmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{bmatrix} \xrightarrow{c_1 \leftrightarrow c_2} \begin{bmatrix} -1 & \lambda+1 & 0 \\ \lambda-3 & 4 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix} \\ &\xrightarrow{r_2 + (\lambda-3)r_1} \begin{bmatrix} -1 & \lambda+1 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix} \xrightarrow{\begin{matrix} (-1)c_1 \\ c_2 - (\lambda+1)c_1 \end{matrix}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix} \\ &\xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & \lambda-2 \\ 0 & (\lambda-1)^2 & 0 \end{bmatrix} \xrightarrow{r_3 + (\lambda-1)^2 r_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & \lambda-2 \\ 0 & 0 & (\lambda-1)^2(\lambda-2) \end{bmatrix} \\ &\xrightarrow{\begin{matrix} (-1)c_2 \\ c_3 - (\lambda-2)c_2 \end{matrix}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)^2(\lambda-2) \end{bmatrix} \end{aligned}$$

于是 $f_1(\lambda) = 1, f_2(\lambda) = 1, f_3(\lambda) = (\lambda-1)^2(\lambda-2)$.

故 $\lambda I - A$ 的全体初等因子为 $(\lambda-1)^2, \lambda-2$.

[注] 设 $A = (a_{ij})_{n \times n}$, 称 $\lambda I - A$ 的行列式因子 (不变因子, 初等因子) 为 A 的行列式因子 (不变因子, 初等因子).

4. Jordan 标准形

设 $A = (a_{ij})_{n \times n}$ 的全体初等因子为

$$(\lambda - \lambda_1)^{m_1}, \dots, (\lambda - \lambda_i)^{m_i}, \dots, (\lambda - \lambda_s)^{m_s}$$

则有 $\varphi(\lambda) = \det(\lambda I - A) = D_n(\lambda) = D_{n-1}(\lambda)d_n(\lambda) = \dots$

$$= D_0(\lambda)d_1(\lambda) \cdots d_n(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_i)^{m_i} \cdots (\lambda - \lambda_s)^{m_s}$$

而且 $m_1 + \dots + m_i + \dots + m_s = n$

对于第 i 个初等因子 $(\lambda - \lambda_i)^{m_i}$ 构造 m_i 阶 Jordan 块矩阵 J_i , 以及准对角矩阵 J 如下:

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{m_i \times m_i}, \quad J = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_s \end{bmatrix}$$

称 J 为矩阵 A 的 **Jordan 标准形**.

这是上 Jordan 块!

Th29 设 A 的 Jordan 标准形为 J , 则存在可逆矩阵 P , 使得 $P^{-1}AP = J$.

例如: 在例 13 中, A 的 Jordan 标准形为 $J = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 2 \end{bmatrix}$.

一般结论 (举例): 设特征矩阵 $\lambda I - A_{9 \times 9}$ 的全体初等因子为

$$(\lambda - 1)^2, (\lambda - 1)^2, (\lambda - 1)$$

$$(\lambda - 2)^2, (\lambda - 2)$$

$$(\lambda - 3)$$

那么 $d_9(\lambda) = (\lambda - 1)^2(\lambda - 2)^2(\lambda - 3)$ (注意 $\text{rank}(\lambda I - A_{9 \times 9}) = 9$)

$$d_8(\lambda) = (\lambda - 1)^2(\lambda - 2)$$

$$d_7(\lambda) = (\lambda - 1)$$

$$d_6(\lambda) = \cdots = d_1(\lambda) = 1$$

由初等因子组
确定不变因子

5. 特征向量分析法求初等因子

设 $\varphi(\lambda) = \det(\lambda I - A)$ 的一个不可约因式为 $(\lambda - \lambda_0)^r$, 则

$(\lambda - \lambda_0)^r$ 是 A 的 k 个初等因子的乘积

$\Leftrightarrow (\lambda_0 I - A)x = 0$ 的基础解系含 k 个解向量 (证明略去)

\Leftrightarrow 对应特征值 λ_0 有 k 个线性无关的特征向量

$\Leftrightarrow k = n - \text{rank}(\lambda_0 I - A)$

例 14 求 $A = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 1 & 3 & 2 & \\ 2 & 3 & 1 & 1 \end{bmatrix}$ 的 Jordan 标准形.

解 $\varphi(\lambda) = \det(\lambda I - A) = (\lambda - 1)^3(\lambda - 2)$

由 $\text{rank}(1I - A) = 2$ 知, $(\lambda - 1)^3$ 是 A 的 $4 - 2 = 2$ 个初等因子的乘积, 即 $(\lambda - 1)^2$ 和 $(\lambda - 1)$ 的乘积, 故 A 的全体初等因子为 $(\lambda - 1)^2, \lambda - 1, \lambda - 2$.

A 的 Jordan 标准形为 $J = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 2 \end{bmatrix}$.

[注] 在例 14 中, 将 $a_{33} = 2, a_{43} = 1$ 改作 $a_{33} = 1, a_{43} = 0$ 时, 此法失效.

6. 相似变换矩阵的求法

仅适用于初等因子组中 $\lambda_i \neq \lambda_j (i \neq j)$ 的情形.

$$P = (P_1, \dots, P_s), P_i = (X_1^{(i)}, \dots, X_{m_i}^{(i)})$$

$$AP = PJ \Leftrightarrow AP_i = P_i J_i \quad (i = 1, 2, \dots, s)$$

特殊情形

$$\text{考虑: } A(X_1^{(i)}, \dots, X_{m_i}^{(i)}) = (X_1^{(i)}, \dots, X_{m_i}^{(i)}) \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

$$(AX_1^{(i)}, AX_2^{(i)}, \dots, AX_{m_i}^{(i)}) = (\lambda_i X_1^{(i)}, X_1^{(i)} + \lambda_i X_2^{(i)}, \dots, X_{m_i-1}^{(i)} + \lambda_i X_{m_i}^{(i)})$$

$$\begin{cases} (\lambda_i I - A)X_1^{(i)} = 0 & X_1^{(i)} \text{ 是 } (\lambda_i I - A)x = 0 \text{ 的非零解} \\ (\lambda_i I - A)X_2^{(i)} = -X_1^{(i)} & X_2^{(i)} \text{ 是 } (\lambda_i I - A)x = -X_1^{(i)} \text{ 的一个解} \\ \dots\dots\dots & \dots\dots\dots \\ (\lambda_i I - A)X_{m_i}^{(i)} = -X_{m_i-1}^{(i)} & X_{m_i}^{(i)} \text{ 是 } (\lambda_i I - A)x = -X_{m_i-1}^{(i)} \text{ 的一个解} \end{cases}$$

可以证明: $X_1^{(i)}, X_2^{(i)}, \dots, X_{m_i}^{(i)}$ 线性无关. (以 $m_i = 3$ 为例证明)

在例 13 中, $\lambda_1 = 1, m_1 = 2$, 求 $X_1^{(1)}, X_2^{(1)}$:

$$\lambda_1 I - A = \begin{bmatrix} 2 & -1 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & -1 \end{bmatrix}, \quad X_1^{(1)} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$[\lambda_1 I - A, -X_1^{(1)}] = \left[\begin{array}{ccc|c} 2 & -1 & 0 & -1 \\ 4 & -2 & 0 & -2 \\ -1 & 0 & -1 & 1 \end{array} \right], \quad X_2^{(1)} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$\lambda_2 = 2, m_2 = 1$, 求 $X_1^{(2)}$:

$$\lambda_2 I - A = \begin{bmatrix} 3 & -1 & 0 \\ 4 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad X_1^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad \text{故 } P = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$

$$\text{例 15 解线性微分方程组 } \begin{cases} \xi_1'(t) = -\xi_1 + \xi_2 \\ \xi_2'(t) = -4\xi_1 + 3\xi_2 \\ \xi_3'(t) = \xi_1 + 2\xi_3 \end{cases}.$$

$$\text{解 } x(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \end{bmatrix}, x'(t) = \begin{bmatrix} \xi_1'(t) \\ \xi_2'(t) \\ \xi_3'(t) \end{bmatrix}, A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}: x'(t) = Ax(t)$$

$$\text{已求得 } P = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \text{ 使得 } P^{-1}AP = J = \left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & & \\ & & & 2 \end{array} \right], \text{ 则有}$$

$$P^{-1}x'(t) = P^{-1}APP^{-1}x(t) \Rightarrow [P^{-1}x(t)]' = J[P^{-1}x(t)]$$

$$y(t) \stackrel{\Delta}{=} P^{-1}x(t) = \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \\ \eta_3(t) \end{bmatrix} \Rightarrow y'(t) = J y(t)$$

$$\begin{cases} \eta_1'(t) = \eta_1 + \eta_2 \\ \eta_2'(t) = \eta_2 \\ \eta_3'(t) = 2\eta_3 \end{cases} \Rightarrow \begin{cases} \eta_2(t) = c_2 e^t \\ \eta_1'(t) = \eta_1 + c_2 e^t \Rightarrow \eta_1(t) = c_1 e^t + c_2 t e^t \\ \eta_3(t) = c_3 e^{2t} \end{cases}$$

$$x(t) = P y(t) = \begin{bmatrix} \eta_1 \\ 2\eta_1 + \eta_2 \\ -\eta_1 - \eta_2 + \eta_3 \end{bmatrix} \Rightarrow \begin{cases} \xi_1(t) = c_1 e^t + c_2 t e^t \\ \xi_2(t) = 2c_1 e^t + c_2 (2t+1) e^t \\ \xi_3(t) = -c_1 e^t - c_2 (t+1) + c_3 e^{2t} \end{cases}$$

(c_1, c_2, c_3 为任意常数)

[注] $\eta_1(t) = e^t \{ c_1 + \int_0^t e^{-\tau} \eta_2(\tau) d\tau \}$

例 16 准对角矩阵 Jordan 标准形的计算方法.

$$A = \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}, \quad P_1^{-1} A_1 P_1 = J_{A_1}, \quad P_2^{-1} A_2 P_2 = J_{A_2}$$

$$P = \begin{bmatrix} P_1 & \\ & P_2 \end{bmatrix}, \quad P^{-1} A P = \cdots = \begin{bmatrix} J_{A_1} & \\ & J_{A_2} \end{bmatrix} = J_A$$

例 17 矩阵的下 Jordan 块相似于上 Jordan 块.

$$J_{\text{上}} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \quad J_{\text{下}} = \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{bmatrix}, \quad P = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \end{bmatrix}$$

$$P^{-1} J_{\text{下}} P = J_{\text{上}}$$

求线性变换在给定基下的矩阵——方法总结:

给定线性空间 V^n 的基 x_1, \dots, x_n , 设线性变换 T 在该基下的矩阵为 A .

一、直接法

- (1) 计算基象组 $T(x_1), \dots, T(x_n)$, 并求出 $T(x_j)$ 在基 x_1, \dots, x_n 下的坐标 (列向量) $\beta_j (j = 1, 2, \dots, n)$;
- (2) 写出 T 在给定基 x_1, \dots, x_n 下的矩阵 $A = (\beta_1, \dots, \beta_n)$.

二、中介法

- (1) 选取 V^n 的简单基, 记作 $\varepsilon_1, \dots, \varepsilon_n$;
(要求 V^n 中元素在该基下的坐标能够直接写出)
- (2) 写出由简单基改变为给定基的过渡矩阵 C (采用直接法);
- (3) 采用直接法求 T 在简单基下的矩阵:
计算基象组 $T(\varepsilon_1), \dots, T(\varepsilon_n)$, 并写出 $T(\varepsilon_j)$ 在简单基 $\varepsilon_1, \dots, \varepsilon_n$ 下的坐标 (列向量) $\beta_j (j = 1, 2, \dots, n)$ 及 T 在简单基下的矩阵 $A_0 = (\beta_1, \dots, \beta_n)$;
- (4) 计算 T 在给定基 x_1, \dots, x_n 下的矩阵 $A = C^{-1} A_0 C$.

三、混合法

- (1) 选取 V^n 的简单基, 记作 $\varepsilon_1, \dots, \varepsilon_n$;
- (2) 写出由简单基改变为给定基的过渡矩阵 C (采用直接法), 则有
$$(x_1, \dots, x_n) = (\varepsilon_1, \dots, \varepsilon_n)C$$
- (3) 计算基象组 $T(x_1), \dots, T(x_n)$, 并写出 $T(x_j)$ 在简单基 $\varepsilon_1, \dots, \varepsilon_n$ 下的坐标 (列向量) $\beta_j (j = 1, 2, \dots, n)$, 以及矩阵 $B = (\beta_1, \dots, \beta_n)$, 则有

$$\begin{aligned} T(x_1, \dots, x_n) &= (T(x_1), \dots, T(x_n)) \\ &= (\varepsilon_1, \dots, \varepsilon_n)B = (x_1, \dots, x_n)C^{-1}B \end{aligned}$$

- (4) 计算 T 在给定基 x_1, \dots, x_n 下的矩阵 $A = C^{-1} B$.

§ 1.3 欧氏空间与酉空间

一、欧氏空间

1. 内积: 线性空间 V , 数域 \mathbf{R} , 对 $\forall x, y \in V$, 定义实数 (x, y) , 且满足

- (1) 交换律 $(x, y) = (y, x)$
 - (2) 分配律 $(x, y + z) = (x, y) + (x, z), \forall z \in V$
 - (3) 齐次性 $(kx, y) = k(x, y), \forall k \in \mathbf{R}$
 - (4) 非负性 $(x, x) \geq 0, (x, x) = 0 \Leftrightarrow x = \theta$
- 称实数 (x, y) 为 x 与 y 的 **内积**.

例 ① 线性空间 \mathbf{R}^n 中: $x = (\xi_1, \dots, \xi_n), y = (\eta_1, \dots, \eta_n)$

内积 1: $(x, y) = \xi_1 \eta_1 + \dots + \xi_n \eta_n$

内积 2: $(x, y)_h = h_1 \xi_1 \eta_1 + \dots + h_n \xi_n \eta_n \quad (h_i > 0)$

多种对应规则

② 线性空间 $\mathbf{R}^{m \times n}$ 中: $A = (a_{ij})_{m \times n}, B = (b_{ij})_{m \times n}$

内积: $(A, B) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij} = \text{tr}(A^T B) \quad (= \text{tr}(AB^T) = \dots)$

③ 线性空间 $C[a, b]$ 中: $f(t), g(t)$ 是区间 $[a, b]$ 上的连续函数

内积: $(f(t), g(t)) = \int_a^b f(t)g(t)dt$

2. 欧氏空间: 定义了内积运算的实线性空间.

设欧氏空间 V^n 的基为 x_1, \dots, x_n , 对 $\forall x, y \in V^n$ 有

$$\left. \begin{array}{l} x = \xi_1 x_1 + \dots + \xi_n x_n \\ y = \eta_1 x_1 + \dots + \eta_n x_n \end{array} \right\} \Rightarrow (x, y) = \sum_{i,j=1}^n \xi_i \eta_j (x_i, x_j)$$

令 $a_{ij} = (x_i, x_j), A = (a_{ij})_{n \times n}$, 则有

$$(x, y) = \sum_{i,j=1}^n a_{ij} \xi_i \eta_j = (\xi_1, \dots, \xi_n) A \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}$$

矩阵乘法表示

则称 A 为基 x_1, \dots, x_n 的 **度量矩阵** (Gram Matrix) .

① A 对称: $(x_i, x_j) = (x_j, x_i) \Rightarrow a_{ij} = a_{ji}$

② A 正定: $\forall \xi_1, \dots, \xi_n \in \mathbf{R}$ 不全为零 $\Rightarrow x = \overset{\Delta}{=} \xi_1 x_1 + \dots + \xi_n x_n \neq \theta$

$$(x, x) > 0 \Rightarrow \text{二次型} (\xi_1, \dots, \xi_n) A \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} > 0, \text{ 即 } A \text{ 正定.}$$

③ V^n 中不同基的度量矩阵是合同的:

基(I): x_1, \dots, x_n ; 基(II): y_1, \dots, y_n

基(I)到基(II)的过渡矩阵为 $C: (y_1, \dots, y_n) = (x_1, \dots, x_n)C$

$$y_i = c_{1i}x_1 + \cdots + c_{ni}x_n, \quad y_j = c_{1j}x_1 + \cdots + c_{nj}x_n$$

$$b_{ij} = (y_i, y_j) = (c_{1i}, \cdots, c_{ni}) A \begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$$

$$(b_{i1}, \cdots, b_{in}) = (c_{1i}, \cdots, c_{ni}) AC$$

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & & \vdots \\ c_{1n} & \cdots & c_{nn} \end{bmatrix} AC = C^T AC$$

合同变换矩阵
是过渡矩阵!

④ “内积值”的计算与基的选取无关: $\forall x, y \in V^n$

$$\left. \begin{aligned} x &= (x_1, \cdots, x_n) \alpha_1 \\ x &= (y_1, \cdots, y_n) \alpha_2 = (x_1, \cdots, x_n) C \alpha_2 \end{aligned} \right\} \Rightarrow \alpha_1 = C \alpha_2$$

$$\left. \begin{aligned} y &= (x_1, \cdots, x_n) \beta_1 \\ y &= (y_1, \cdots, y_n) \beta_2 = (x_1, \cdots, x_n) C \beta_2 \end{aligned} \right\} \Rightarrow \beta_1 = C \beta_2$$

$$\text{基(I)下: } (x, y) = \alpha_1^T A \beta_1$$

$$\text{基(II)下: } (x, y) = \alpha_2^T B \beta_2 = \alpha_2^T C^T A C \beta_2 = \alpha_1^T A \beta_1$$

3. 元素的模: 欧氏空间 V , $\forall x \in V$, 称 $|x| = \sqrt{(x, x)}$ 为 x 的模 (长度).

$$\text{① } \forall k \in \mathbf{R}, |kx| = |k| |x|$$

$$\text{② } |(x, y)| \leq |x| |y|; \text{ 当 } x \neq \theta, y \neq \theta \text{ 时, 等号成立} \Leftrightarrow x, y \text{ 线性相关.}$$

$$\text{证 } \forall t \in \mathbf{R}, x - ty \in V \Rightarrow |x - ty|^2 = (x - ty, x - ty) \geq 0$$

$$(y, y)t^2 - 2(x, y)t + (x, x) \geq 0$$

$$\Delta \leq 0: 4(x, y)^2 - 4(y, y)(x, x) \leq 0 \Rightarrow |(x, y)| \leq |x| |y|$$

充分性. 已知 x, y 线性相关, 且 $x \neq \theta$, 所以 $y = kx$, 从而

$$|(x, y)| = |(x, kx)| = |k| |(x, x)| = |k| |x| |x| = |x| |y|$$

必要性. 已知 $|(x, y)| = |x| |y|$, 则

$$(x, y) \geq 0 \text{ 时, 取 } t = \frac{|x|}{|y|} \Rightarrow (x - ty, x - ty) = \cdots = 0 \Rightarrow x - ty = \theta$$

$$(x, y) < 0 \text{ 时, 取 } t = -\frac{|x|}{|y|} \Rightarrow (x - ty, x - ty) = \cdots = 0 \Rightarrow x - ty = \theta$$

故 x, y 线性相关.

$$\text{③ } |x + y| \leq |x| + |y|; \quad \text{④ } |x - y| \geq ||x| - |y||$$

单位元素: 模为 1 的元素.

单位化: $x \neq \theta$, $x_0 = \frac{1}{|x|} x$ (同方向), $x_0 = \frac{-1}{|x|} x$ (反方向).

4. 元素的夹角: 欧氏空间 V 中, 若 $x \neq \theta, y \neq \theta$, 称 $\phi = \arccos \frac{(x, y)}{|x||y|} \in [0, \pi]$ 为 x 与 y 之间的夹角.

二、正交性

欧氏空间 V 中, 若 $(x, y) = 0$, 称 x 与 y 正交, 记作 $x \perp y$;

若 x_1, \dots, x_m 满足 $(x_i, x_j) = 0 (i \neq j)$, 称 x_1, \dots, x_m 为正交元素组.

Th 31 欧氏空间 V 中, $x \perp y \Rightarrow |x+y|^2 = |x|^2 + |y|^2$.

勾股定理

证 $|x+y|^2 = (x+y, x+y) = (x, x) + 2(x, y) + (y, y) = |x|^2 + |y|^2$

Th 32 欧氏空间 V 中, x_1, \dots, x_m 两两正交且非零 $\Rightarrow x_1, \dots, x_m$ 线性无关.

[注] 欧氏空间 V^n 中, 两两正交的非零元素的个数不超过 n .

1. 正交基: 欧氏空间 V^n 中, 若 x_1, \dots, x_n 两两正交且非零, 称 x_1, \dots, x_n 为 V^n 的正交基; 若还有 $|x_i| = 1 (i = 1, 2, \dots, n)$, 称 x_1, \dots, x_n 为 V^n 的标准正交基.

Schmidt 正交化方法: 设欧氏空间 V^n 的一个基为 x_1, \dots, x_n , 构造元素组

y_1, \dots, y_n , 使满足 $y_i \perp y_j (i \neq j)$ 且 $y_i \neq \theta (i = 1, 2, \dots, n)$.

$$y_1 = x_1, \quad y_1 \neq \theta$$

$$y_2 = x_2 + k_{21}y_1, \quad y_2 \neq \theta$$

$$(y_2, y_1) = 0 \Rightarrow 0 = (x_2, y_1) + k_{21}(y_1, y_1) \Rightarrow k_{21} = -\frac{(x_2, y_1)}{(y_1, y_1)}$$

$$y_3 = x_3 + k_{32}y_2 + k_{31}y_1, \quad y_3 \neq \theta$$

$$(y_3, y_1) = 0 \Rightarrow k_{31} = -\frac{(x_3, y_1)}{(y_1, y_1)}$$

$$(y_3, y_2) = 0 \Rightarrow k_{32} = -\frac{(x_3, y_2)}{(y_2, y_2)}$$

.....

$$y_n = x_n + k_{n,n-1}y_{n-1} + \dots + k_{n1}y_1, \quad y_n \neq \theta$$

$$(y_n, y_j) = 0 \Rightarrow k_{nj} = -\frac{(x_n, y_j)}{(y_j, y_j)}, \quad j = 1, 2, \dots, n-1$$

因为 y_1, \dots, y_n 两两正交且非零, 所以线性无关, 从而是 V^n 的正交基.

Th 33 欧氏空间 $V^n (n \geq 1)$ 存在标准正交基.

证 对 V^n 的基 x_1, \dots, x_n 进行正交化, 可得正交基 y_1, \dots, y_n ;

再进行单位化, 可得标准正交基 z_1, \dots, z_n ($z_j = \frac{1}{|y_j|} y_j$).

例 1 标准正交基的特殊性质: 设欧氏空间 V^n 的标准正交基为 x_1, \dots, x_n ,

且 $x = \xi_1 x_1 + \dots + \xi_n x_n$, $y = \eta_1 x_1 + \dots + \eta_n x_n$, 则有

(1) $\xi_i = (x, x_i)$, $\eta_j = (y, x_j)$; (2) 基 x_1, \dots, x_n 的度量矩阵 $A = I$;

(3) $(x, y) = \xi_1 \eta_1 + \dots + \xi_n \eta_n$. (等号右端: 坐标向量的通常内积)

2. 子空间的正交性: 欧氏空间 V 的子空间 V_1 , 给定 $y \in V$, 若 $\forall x \in V_1$, 都有 $y \perp x$, 则称 y 正交于 V_1 , 记作 $y \perp V_1$.

例 2 欧氏空间 V^n , 子空间 V_1 的基为 x_1, x_2, \dots, x_m , 则

$$y \perp V_1 \Leftrightarrow y \perp x_j \quad (j=1, 2, \dots, m)$$

判别方法

证 必要性. (略)

充分性. $\forall x \in V_1 \Rightarrow x = k_1 x_1 + \dots + k_m x_m$

$$y \perp x_j \Rightarrow (y, x_j) = 0 \Rightarrow (y, x) = 0, \text{ 即 } y \perp V_1$$

注 1. 若 $V_1 = L(x_1, x_2, \dots, x_m)$, 则 $y \perp V_1 \Leftrightarrow y \perp x_j \quad (j=1, 2, \dots, m)$

注 2. 若对 $\forall x \in V_1, \forall y \in V_2$, 都有 $x \perp y$, 则称 $V_1 \perp V_2$.

注 3. 若 $V_1 = L(x_1, x_2, \dots, x_m)$, $V_2 = L(y_1, y_2, \dots, y_l)$, 则

$$V_1 \perp V_2 \Leftrightarrow x_i \perp y_j \quad (i=1, \dots, m; j=1, \dots, l).$$

例 3 欧氏空间 V , 子空间 V_1 , 则 $V_1^\perp = \{y \mid y \in V \text{ 且 } y \perp V_1\}$ 是 V 的子空间.

证 $\theta \in V_1^\perp \Rightarrow V_1^\perp$ 非空. $\forall y, z \in V_1^\perp, \forall x \in V_1, \forall k \in \mathbb{R}$, 有

$$(y+z, x) = (y, x) + (z, x) = 0 \Rightarrow (y+z) \perp V_1: (y+z) \in V_1^\perp$$

$$(ky, x) = k(y, x) = 0 \Rightarrow (ky) \perp V_1: (ky) \in V_1^\perp$$

故 V_1^\perp 是 V 的子空间 (称 V_1^\perp 为 V_1 的正交补).

正交补分解

Th 34 设欧氏空间 V^n , 子空间 V_1 , 则 $V^n = V_1 \oplus V_1^\perp$.

证 若 $V_1 = \{\theta\}$, 则 $V_1^\perp = V^n \Rightarrow V^n = \{\theta\} \oplus V^n = V_1 \oplus V_1^\perp$

若 $V_1 \neq \{\theta\}$, 记 $\dim V_1 \stackrel{\text{def}}{=} m \leq n$, 设 V_1 的标准正交基为 x_1, \dots, x_m .

① 先证 $V^n = V_1 + V_1^\perp$: 只需证明 $V^n \subset V_1 + V_1^\perp$ 即可.

任意 $x \in V^n$, 设 $x = y + z$ ($y \in V_1$), 则有

$$y \in V_1 \Rightarrow y = a_1 x_1 + \dots + a_m x_m$$

选取 $a_i = (x, x_i)$ (下式成立需要), 由 $z = x - y$ 可得

$$(z, x_i) = (x, x_i) - (y, x_i) = a_i - a_i = 0 \quad (i=1, 2, \dots, m)$$

故 $z \perp V_1$, 即 $z \in V_1^\perp$. 于是 $x = y + z \in V_1 + V_1^\perp$, 故 $V^n \subset V_1 + V_1^\perp$.

② 再证 $V_1 \cap V_1^\perp = \{\theta\}$: $\forall x \in V_1 \cap V_1^\perp \Rightarrow \begin{cases} x \in V_1 \\ x \in V_1^\perp \end{cases} \Rightarrow (x, x) = 0 \Rightarrow x = \theta$

①+②: 结论成立.

Th 35 设 $A = (a_{ij})_{m \times n} \in \mathbb{R}^{m \times n}$, 则

(1) $[R(A)]^\perp = N(A^T)$, 且 $\mathbb{R}^m = R(A) \oplus N(A^T)$;

(2) $[R(A^T)]^\perp = N(A)$, 且 $\mathbb{R}^n = R(A^T) \oplus N(A)$.

正交补分解的特例

证 划分 $A = (\beta_1, \dots, \beta_n)$ (列向量 $\beta_j \in \mathbb{R}^m$)

$$V_1 = R(A) = L(\beta_1, \dots, \beta_n) \subset \mathbb{R}^m$$

$$V_1^\perp = \{y \mid y \in \mathbb{R}^m \text{ 且 } y \perp (k_1\beta_1 + \dots + k_n\beta_n)\} \subset \mathbb{R}^m$$

$$= \{y \mid y \in \mathbb{R}^m \text{ 且 } y \perp \beta_j (j=1, 2, \dots, n)\}$$

$$= \{y \mid y \in \mathbb{R}^m \text{ 且 } \beta_j^T y = 0 (j=1, 2, \dots, n)\}$$

$$= \{y \mid y \in \mathbb{R}^m \text{ 且 } A^T y = 0\} = N(A^T)$$

由 Th34 可得 $\mathbb{R}^m = V_1 \oplus V_1^\perp = R(A) \oplus N(A^T)$.

对 A^T 应用上述结果可得 $[R(A^T)]^\perp = N(A)$

再由 Th34 可得 $\mathbb{R}^n = R(A^T) \oplus N(A)$.

三、正交变换与正交矩阵

正交变换: 欧氏空间 V 中, 线性变换 T 满足 $(Tx, Tx) = (x, x) (\forall x \in V)$.

注: $(Tx, Tx) = (x, x) \Leftrightarrow |Tx| = |x|$

模不变

Th36 欧氏空间 V , 线性变换 T .

T 是正交变换 $\Leftrightarrow \forall x, y \in V, (Tx, Ty) = (x, y)$.

证 充分性. 取 $y = x$, 则 $\forall x \in V, (Tx, Tx) = (x, x)$.

内积不变

必要性. T 是正交变换: $\forall x, y \in V \Rightarrow x - y \in V$

$$(T(x-y), T(x-y)) = (x-y, x-y) \Rightarrow \dots \Rightarrow (Tx, Ty) = (x, y)$$

注: 若 T 是正交变换, 则有 $\frac{(Tx, Ty)}{|Tx| \cdot |Ty|} = \frac{(x, y)}{|x| \cdot |y|}$.

夹角不变

Th 37 欧氏空间 V^n 的标准正交基为 x_1, \dots, x_n

线性变换 $T: T(x_1, \dots, x_n) = (x_1, \dots, x_n)A$

则 T 是正交变换 $\Leftrightarrow A$ 是正交矩阵.

判别方法

证 必要性. T 是正交变换: 划分 $A = (a_{ij})_{n \times n} = (\alpha_1, \dots, \alpha_n)$, 则

$$Tx_i = (x_1, \dots, x_n)\alpha_i, \quad Tx_j = (x_1, \dots, x_n)\alpha_j$$

$$(Tx_i, Tx_j) = (x_i, x_j) \Rightarrow (\alpha_i, \alpha_j) = \delta_{ij} \Rightarrow A^T A = I$$

充分性. A 是正交矩阵:

$$\forall x \in V \Rightarrow x = (x_1, \dots, x_n)\alpha, \quad Tx = (x_1, \dots, x_n)A\alpha$$

$$(Tx, Tx) = (A\alpha)^T (A\alpha) = \alpha^T A^T A \alpha = \alpha^T \alpha = (x, x)$$

[注] 欧氏空间 V^n 的标准正交基具有如下性质:

- (1) x_1, \dots, x_n 是标准正交基, T 是正交变换 $\Rightarrow Tx_1, \dots, Tx_n$ 是标准正交基
- (2) $\left. \begin{array}{l} x_1, \dots, x_n \text{ 和 } y_1, \dots, y_n \text{ 都是标准正交基} \\ (y_1, \dots, y_n) = (x_1, \dots, x_n)C \end{array} \right\} \Rightarrow C \text{ 是正交矩阵}$

四、对称变换与对称矩阵

对称变换: 欧氏空间 V 中, 线性变换 T 满足 $(Tx, y) = (x, Ty) \quad (\forall x, y \in V)$.

Th 38 欧氏空间 V^n 的标准正交基为 x_1, \dots, x_n

线性变换 $T: T(x_1, \dots, x_n) = (x_1, \dots, x_n)A$

则 T 是对称变换 $\Leftrightarrow A$ 是对称矩阵.

判别方法

证 设 $A = (a_{ij})_{n \times n}$, 则 $Tx_i = a_{1i}x_1 + \dots + a_{ni}x_n \Rightarrow (Tx_i, x_j) = a_{ji}$

$$Tx_j = a_{1j}x_1 + \dots + a_{nj}x_n \Rightarrow (x_i, Tx_j) = a_{ij}$$

必要性. T 是对称变换 $\Rightarrow (Tx_i, x_j) = (x_i, Tx_j) \Rightarrow a_{ji} = a_{ij}$, 即 $A^T = A$

充分性. A 是对称矩阵, 则有

$$\forall x \in V \Rightarrow x = (x_1, \dots, x_n)\alpha, \quad Tx = (x_1, \dots, x_n)A\alpha$$

$$\forall y \in V \Rightarrow y = (x_1, \dots, x_n)\beta, \quad Ty = (x_1, \dots, x_n)A\beta$$

$$(Tx, y) = (A\alpha)^T \cdot \beta = \alpha^T \cdot A\beta = (x, Ty)$$

Th 39 $A \in \mathbb{R}^{n \times n}$ 且 $A^T = A \Rightarrow \lambda_A \in \mathbb{R}$.

证 设 $Ax = \lambda x \quad (x \neq \theta)$, 则 $x^H Ax = \begin{cases} x^H(Ax) = \lambda(x^H x) \\ (Ax)^H x = \bar{\lambda}(x^H x) \end{cases}$.

故 $(\lambda - \bar{\lambda})(x^H x) = 0 \Rightarrow \lambda = \bar{\lambda} \quad (\because x^H x > 0)$, 即 $\lambda \in \mathbb{R}$.

[注] 因为 $(\lambda I - A)x = 0$ 是实系数齐次线性方程组, 所以可求得非零解向量 $x \in \mathbb{R}^n$. **约定:** 实对称矩阵的特征向量为实向量.

Th 40 设实对称矩阵 A 的特征值 $\lambda_1 \neq \lambda_2$, 对应的特征向量分别为 x_1 和 x_2 , 则 $(x_1, x_2) = 0$.

证 $\left. \begin{array}{l} Ax_1 = \lambda_1 x_1 \\ Ax_2 = \lambda_2 x_2 \end{array} \right\} \Rightarrow x_1^T Ax_2 = \begin{cases} x_1^T(Ax_2) = \lambda_2(x_1^T x_2) \\ (Ax_1)^T x_2 = \lambda_1(x_1^T x_2) \end{cases}$
 $\Rightarrow (\lambda_2 - \lambda_1)(x_1^T x_2) = 0 \Rightarrow x_1^T x_2 = 0$, 即 $(x_1, x_2) = 0$.

五、酉空间简介

1. 复内积 线性空间 V , 复数域 K , 对 $\forall x, y \in V$, 定义复数 (x, y) , 且满足

$$(1) (x, y) = \overline{(y, x)}; \quad (2) (x, y+z) = (x, y) + (x, z), \forall z \in V$$

$$(3) (kx, y) = k(x, y), \forall k \in K; \quad (4) (x, x) \geq 0, (x, x) = 0 \Leftrightarrow x = \theta$$

称复数 (x, y) 为 x 与 y 的复内积.

2. 酉空间 定义了复内积运算的复线性空间.

$$\textcircled{1} (x, ky) = \overline{(ky, x)} = \overline{k(y, x)} = \bar{k}(y, x) = \bar{k}(x, y)$$

② 基的度量矩阵为 Hermite 正定矩阵 (Hermite 矩阵 $A^H = A$)

$$\textcircled{3} (x, y) = (\xi_1, \dots, \xi_n) A \begin{bmatrix} \bar{\eta}_1 \\ \vdots \\ \bar{\eta}_n \end{bmatrix}$$

④ 酉变换 (复正交变换): $(Tx, Tx) = (x, x) \quad (\forall x \in V)$

T 是酉变换 $\Leftrightarrow T$ 在标准正交基下的矩阵 A 是酉矩阵, 即 $A^H A = I$.

⑤ Hermite-变换 (复对称变换): $(Tx, y) = (x, Ty) \quad (\forall x, y \in V)$

T 是 Hermite-变换 \Leftrightarrow

T 在标准正交基下的矩阵 A 是 Hermite 矩阵, 即 $A^H = A$.

Th41 (1) 设 $A_{n \times n}$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则存在酉矩阵 $P_{n \times n}$, 使得

$$P^H A P = \begin{bmatrix} \lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_n \end{bmatrix}$$

(2) 设 $A \in \mathbb{R}^{n \times n}$ 且 $\lambda_A \in \mathbb{R}$, 则存在正交矩阵 $Q_{n \times n}$, 使得 $Q^T A Q$ 为上三角矩阵.

证 在 Th17 的证明过程中, 将“扩充 x_1 为 \mathbb{C}^k 的基”修改为

“扩充 x_1 为 \mathbb{C}^k 的标准正交基”即可.

3. 正规矩阵: 指 $A_{n \times n}$ 满足 $A^H A = A A^H$.

例如 $\mathbb{C}^{n \times n}$ 中: ① $A^H = A \Rightarrow A$ 正规; ② $A^H A = I \Rightarrow A$ 正规.

$\mathbb{R}^{n \times n}$ 中: ① $A^T = A \Rightarrow A$ 正规; ② $A^T A = I \Rightarrow A$ 正规.

$$B = \begin{bmatrix} 5+4j & 1+6j \\ 1+6j & 5+4j \end{bmatrix}; \quad B^H B = \begin{bmatrix} 78 & 58 \\ 58 & 78 \end{bmatrix} = B B^H \Rightarrow B \text{ 是正规矩阵}$$

但是 $B^H \neq B, B^H B \neq I$.

Th42 (1) $A \in \mathbb{C}^{n \times n}, A$ 正规 $\Leftrightarrow \exists$ 酉矩阵 $P_{n \times n}$, 使得 $P^H A P = \Lambda$;

(2) $A \in \mathbb{R}^{n \times n}$ 且 $\lambda_A \in \mathbb{R}, A$ 正规 $\Leftrightarrow \exists$ 正交矩阵 $Q_{n \times n}$ 使得 $Q^T A Q = \Lambda$.

证 (1) 充分性. $A = P \Lambda P^H, A^H = P \bar{\Lambda} P^H$

$$A^H A = P \bar{\Lambda} \Lambda P^H = P \Lambda \bar{\Lambda} P^H = A A^H$$

必要性. $A^H A = A A^H$:

A 不是正规矩阵
 A 不能酉相似 Λ
 A 有可能相似 Λ

$$\text{th41(1)} \Rightarrow \exists \text{ 酉矩阵 } P \text{ 使得 } P^H A P = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix} \stackrel{\Delta}{=} B$$

$$B^H B = P^H A^H A P = P^H A A^H P = B B^H$$

$$\begin{bmatrix} \bar{b}_{11} & & & \\ \bar{b}_{12} & \bar{b}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{b}_{1n} & \bar{b}_{2n} & \cdots & \bar{b}_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix} \begin{bmatrix} \bar{b}_{11} & & & \\ \bar{b}_{12} & \bar{b}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{b}_{1n} & \bar{b}_{2n} & \cdots & \bar{b}_{nn} \end{bmatrix}$$

比较第一行第一列元素可得

$$|b_{11}|^2 = |b_{11}|^2 + |b_{12}|^2 + \cdots + |b_{1n}|^2 \Rightarrow b_{12} = 0, b_{13} = 0, \cdots, b_{1n} = 0$$

一般地, 有

$$\left. \begin{array}{l} i=1: b_{12}=0, b_{13}=0, \cdots, b_{1n}=0 \\ i=2: \quad b_{23}=0, \cdots, b_{2n}=0 \\ \cdots \cdots \\ i=n-1: \quad b_{n-1,n}=0 \end{array} \right\} \Rightarrow B = \begin{bmatrix} b_{11} & & & \\ & b_{22} & & \\ & & \ddots & \\ & & & b_{nn} \end{bmatrix}$$

(2) 利用 th41(2) 可得.

例如, 对于 $B = \begin{bmatrix} 5+4j & 1+6j \\ 1+6j & 5+4j \end{bmatrix}$, 可求得酉矩阵 $P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, 使得

$$P^H B P = \begin{bmatrix} 6+10j & \\ & 4-2j \end{bmatrix}$$

推论 1 $A_{n \times n}$ 实对称 $\Rightarrow \exists$ 正交矩阵 $Q_{n \times n}$, 使得 $Q^T A Q = \Lambda$.

推论 2 欧氏空间 V^n , 对称变换 $T \Rightarrow$

存在标准正交基 y_1, \cdots, y_n , 使得 $T(y_1, \cdots, y_n) = (y_1, \cdots, y_n) \Lambda$.

证 设 V^n 的标准正交基为 x_1, \cdots, x_n , T 在该基下的矩阵为 A , 则

A 是实对称矩阵 $\Rightarrow \exists$ 正交矩阵 $Q_{n \times n}$ 使得 $Q^T A Q = \Lambda$

构造 V^n 的标准正交基: $(y_1, \cdots, y_n) = (x_1, \cdots, x_n) Q$ (为什么?)

$$\begin{aligned} \text{则 } T(y_1, \cdots, y_n) &= T(x_1, \cdots, x_n) Q = (x_1, \cdots, x_n) A Q \\ &= (y_1, \cdots, y_n) Q^{-1} A Q = (y_1, \cdots, y_n) \Lambda \end{aligned}$$

推论 3 A 实对称 $\Rightarrow A$ 有 n 个线性无关的特征向量 (\mathbb{R}^n 中).

推论 4 T 是欧氏空间 V^n 的对称变换 \Rightarrow

T 有 n 个线性无关的特征向量 (V^n 中)

谱分解: $A_{n \times n}$ 是 Hermite 矩阵 $\xrightarrow{\text{Th42}} \exists$ 酉矩阵 $P_{n \times n}$, 使得 $A = P \Lambda P^H$

划分 $P = (p_1, \cdots, p_n)$, 则有

$$\begin{aligned} A &= (p_1, \cdots, p_n) \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^H = (\lambda_1 p_1, \cdots, \lambda_n p_n) \begin{bmatrix} p_1^H \\ \vdots \\ p_n^H \end{bmatrix} \\ &= \lambda_1 (p_1 p_1^H) + \cdots + \lambda_n (p_n p_n^H) \end{aligned}$$

注: 矩阵组 $B_1 = p_1 p_1^H, \cdots, B_n = p_n p_n^H$ 线性无关, 且 $\text{rank } B_j = 1$.

例 4 给定欧氏空间 V^n 的标准正交基 x_1, x_2, \dots, x_n , 设 T 是 V^n 的正交变换, $W = L(x_1, x_2, \dots, x_r)$ 是 T 的不变子空间, 证明: V^n 的子空间 $W^\perp = \{y \mid y \in V^n, y \perp W\}$ 也是 T 的不变子空间.

证 因为 T 是 V^n 的正交变换, 而 x_1, x_2, \dots, x_n 是 V^n 的**标准正交基**

所以 $T(x_1), T(x_2), \dots, T(x_n)$ 也是 V^n 的**标准正交基** (定理 37 之证明)

$$W = L(x_1, x_2, \dots, x_r) \Rightarrow W^\perp = L(x_{r+1}, \dots, x_n)$$

理由: $\dim W = r \Rightarrow \dim W^\perp = n - r$, $x_1, \dots, x_r \perp x_{r+1}, \dots, x_n$ (线性无关)

W 是 T 的不变子空间 $\Rightarrow T(x_1), \dots, T(x_r)$ 是 W 的标准正交基

$$T(x_{r+1}), \dots, T(x_n) \perp W \Rightarrow T(x_{r+1}), \dots, T(x_n) \in W^\perp$$

理由: $T(x_{r+1}), \dots, T(x_n) \perp T(x_1), \dots, T(x_r)$ (W 的标准正交基)

$$\forall x \in W^\perp \Rightarrow x = k_{r+1}x_{r+1} + \dots + k_n x_n$$

$$\Rightarrow T(x) = k_{r+1}T(x_{r+1}) + \dots + k_n T(x_n) \in W^\perp$$

故 W^\perp 是 T 的不变子空间.

证 2 因为 T 是正交变换, 所以 T 是可逆变换, 可将 T 看作 W 中的可逆变换.

$$\forall \alpha \in W \Rightarrow \text{存在 } \beta \in W \text{ 使得 } T(\beta) = \alpha \quad (\text{可逆变换的性质})$$

$$\forall \gamma \in W^\perp \Rightarrow (T(\gamma), \alpha) = (T(\gamma), T(\beta)) = (\gamma, \beta) = 0 \Rightarrow T(\gamma) \in W^\perp$$

故 W^\perp 是 T 的不变子空间.

例 5 设线性空间 V 中的线性变换 T 满足 $T^2 = T$, $R(T)$ 表示 T 的值域, $N(T)$ 表示 T 的核, T_e 表示 V 中的单位变换, 证明 $N(T) = R(T_e - T)$.

证 先证 $N(T) \subset R(T_e - T)$: $\forall \alpha \in N(T)$, 有

$$T(\alpha) = \theta \Rightarrow \alpha = T_e(\alpha) = T_e(\alpha) - T(\alpha) = (T_e - T)(\alpha) \in R(T_e - T)$$

再证 $R(T_e - T) \subset N(T)$: $\forall \alpha \in R(T_e - T)$, 存在 $\beta \in V$, 使得

$$\alpha = (T_e - T)(\beta) = \beta - T(\beta)$$

$$T(\alpha) = T(\beta) - T[T(\beta)] = T(\beta) - T^2(\beta) = T(\beta) - T(\beta) = \theta$$

故 $\alpha \in N(T)$.

典型题解析

例 6 (M2007A 七) 设 3 维欧氏空间 V 中元素 α_0 在 V 的标准正交基 ξ_1, ξ_2, ξ_3 下的坐标为 $(1, -1, 0)^T$, 定义 V 中的变换如下

$$T(\alpha) = \alpha + (\alpha, \alpha_0)\alpha_0 \quad (\forall \alpha \in V)$$

其中 (α, α_0) 表示 α 与 α_0 的内积.

1. 证明 T 是线性变换;
2. 证明 T 是对称变换;
3. 求 V 的一组标准正交基 η_1, η_2, η_3 , 使 T 在该基下的矩阵为对角矩阵.

解 T 是线性变换 (略)

$$\text{计算基像组} \quad \begin{cases} T(\xi_1) = \xi_1 + 1\alpha_0 = 2\xi_1 - \xi_2 \\ T(\xi_2) = \xi_2 - 1\alpha_0 = -\xi_1 + 2\xi_2 \\ T(\xi_3) = \xi_3 + 0\alpha_0 = \xi_3 \end{cases}$$

$$\text{设 } T(\xi_1, \xi_2, \xi_3) = (\xi_1, \xi_2, \xi_3)A, \text{ 则 } A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

因为 ξ_1, ξ_2, ξ_3 是标准正交基, A 是对称矩阵, 所以 T 是对称变换.

求正交矩阵 Q , 使得 $Q^{-1}AQ = \Lambda$:

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 3 \end{bmatrix}$$

由 $(\eta_1, \eta_2, \eta_3) = (\xi_1, \xi_2, \xi_3)Q$ 求得标准正交基

$$\eta_1 = \frac{1}{\sqrt{2}}\xi_1 + \frac{1}{\sqrt{2}}\xi_2, \quad \eta_2 = \xi_3, \quad \eta_3 = -\frac{1}{\sqrt{2}}\xi_1 + \frac{1}{\sqrt{2}}\xi_2$$

且有 $T(\eta_1, \eta_2, \eta_3) = (\eta_1, \eta_2, \eta_3)\Lambda$.

M2009A

八、(7 分) 设线性空间 V^4 的一个基为 x_1, x_2, x_3, x_4 , 线性变换 T 在该基下的

矩阵为 $\begin{bmatrix} A & B \\ O & C \end{bmatrix}$, 其中 A, B, C 都是 2 阶方阵, O 是 2 阶零矩阵, 证明:

1. 子空间 $V_1 = \text{span}\{x_1, x_2\}$ 是 T 的不变子空间;
2. 若 $B \neq O$, 则子空间 $V_2 = \text{span}\{x_3, x_4\}$ 不是 T 的不变子空间.

证 1. 记 $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, 由 $T(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3, x_4) \begin{bmatrix} A & B \\ O & C \end{bmatrix}$ 可得

$$T(x_1) = a_{11}x_1 + a_{21}x_2 \in V_1, \quad T(x_2) = a_{12}x_1 + a_{22}x_2 \in V_1$$

任意 $x \in V_1$, 存在 k_1, k_2 使得 $x = k_1x_1 + k_2x_2$, 于是有

$$T(x) = k_1 \cdot T(x_1) + k_2 \cdot T(x_2) \in V_1$$

故 $V_1 = \text{span}\{x_1, x_2\}$ 是 T 的不变子空间.

2. 记 $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$, 不妨设 $b_{11} \neq 0$, 因为

$$x_3 \in V_2, \quad T(x_3) = b_{11}x_1 + b_{21}x_2 + c_{11}x_3 + c_{21}x_4 \notin V_2 \quad (\text{反证法})$$

所以 $V_2 = \text{span}\{x_3, x_4\}$ 不是 T 的不变子空间.

M2008B

八、(8 分) 设 $A \in C^{n \times n}$ 满足 $A^2 = -A$, 证明 $C^n = R(A) \oplus N(A)$, 即 C^n 可分解为子空间 $R(A)$ 与 $N(A)$ 的直和.

证 (1) 先证 $C^n = R(A) + N(A)$: 对任意 $x \in C^n$, 有 $x = -Ax + (I + A)x$

因为 $y = -Ax = A(-x) \in R(A)$, $z = (I + A)x \in N(A) \quad (\because Az = 0)$

所以 $x = y + z \in R(A) + N(A)$, 故 $C^n \subset R(A) + N(A)$

由此可得 $C^n = R(A) + N(A)$

(2) 再证 $R(A) \cap N(A) = \{0\}$: 任意 $\beta \in R(A) \cap N(A)$, 有

$$\beta \in R(A) \Rightarrow \text{存在 } \alpha \in C^n, \text{ 使得 } \beta = A\alpha; \quad \beta \in N(A) \Rightarrow A\beta = 0$$

$$\text{因为 } \beta = A\alpha = -A^2\alpha = -A(A\alpha) = -A\beta = 0$$

$$\text{所以 } R(A) \cap N(A) = \{0\}$$

综述(1)和(2)可得 $C^n = R(A) \oplus N(A)$.

证法 2 (1) 先证 $R(A) \cap N(A) = \{0\}$: (同证法 1)

(2) 再证 $C^n = R(A) + N(A)$: 利用维数公式可得

$$\dim[R(A) + N(A)] = \dim R(A) + \dim N(A) - \dim[R(A) \cap N(A)]$$

$$= \text{rank}(A) + [n - \text{rank}(A)] - 0 = n$$

因为 $R(A) + N(A) \subset C^n$ (定理 5), 所以 $C^n = R(A) + N(A)$.

综述(1)和(2)可得 $C^n = R(A) \oplus N(A)$.

M2005A

八、(9 分) 设线性空间 V^n 的两个基分别为 (I) x_1, x_2, \dots, x_n 和 (II) y_1, y_2, \dots, y_n ,

元素 $x \in V^n$ 在基 (I) 下的坐标为 α , 在基 (II) 下的坐标为 β , 证明子集

$W = \{x \mid x \in V^n, \alpha = 2\beta\}$ 是 V^n 的线性子空间.

证 $\theta \in V^n$ 在基 (I) 下的坐标为 0 (零向量), 在基 (II) 下的坐标为 0

$$0 = 2 \cdot 0 \Rightarrow \theta \in W \Rightarrow W \text{ 非空}$$

设 $x \in W$ 在基 (I) 和基 (II) 下的坐标分别为 α_1, β_1 , 则 $\alpha_1 = 2\beta_1$

$y \in W$ 在基 (I) 和基 (II) 下的坐标分别为 α_2, β_2 , 则 $\alpha_2 = 2\beta_2$

于是 $x + y$ 在基 (I) 和基 (II) 下的坐标分别为 $\alpha_1 + \alpha_2, \beta_1 + \beta_2$, 且有

$$\alpha_1 + \alpha_2 = 2\beta_1 + 2\beta_2 = 2(\beta_1 + \beta_2) \Rightarrow x + y \in W$$

而 kx 在基 (I) 和基 (II) 下的坐标分别为 $k\alpha_1, k\beta_1$, 且有

$$k\alpha_1 = k(2\beta_1) = 2(k\beta_1) \Rightarrow kx \in W$$

故 W 是 V^n 的子空间.

M2006A

七、(15 分) 已知矩阵空间 $R^{2 \times 2}$ 的一个基为

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{线性变换 } T \text{ 满足: } T(A_1) = B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, T(A_2) = B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$T(A_3) = B_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, T(A_4) = B_4 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

1. 求 T 在基 A_1, A_2, A_3, A_4 下的矩阵;
2. 判断 T 能否在 $R^{2 \times 2}$ 的某个基下的矩阵为对角矩阵; (要求写出判断依据)
3. 求 $N(T)$ 的一个基.

解 ① $(A_1, \dots, A_4) = (E_{11}, \dots, E_{22})C_1, (B_1, \dots, B_4) = (E_{11}, \dots, E_{22})C_2$

$$C_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & 1 \end{bmatrix}, C_1^{-1}C_2 = \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & -2 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

$$(B_1, \dots, B_4) = (A_1, \dots, A_4)C_1^{-1}C_2$$

$$T(A_1, \dots, A_4) = (B_1, \dots, B_4) = (A_1, \dots, A_4)C_1^{-1}C_2, \quad A = C_1^{-1}C_2$$

$$\textcircled{2} \quad |\lambda I - A| = \lambda^2(\lambda - 1)^2$$

$r(0I - A) = 2 \Rightarrow \lambda_1 = \lambda_2 = 0$ 有 2 个线性无关的特征向量 (略计算过程)

$r(1I - A) = 3 \Rightarrow \lambda_3 = \lambda_4 = 1$ 只有 1 个线性无关的特征向量 (略计算过程)

故 A 不可对角化 $\Rightarrow T$ 不能在某个基下的矩阵为对角矩阵

$$\textcircled{3} \quad N(T) = \{X \mid T(X) = O\} = \{X \mid X = (A_1, \dots, A_4)\alpha, A\alpha = 0\}$$

$A\alpha = 0$ 的基础解系为 $\alpha_1 = (-1, -1, 1, 0)^T, \alpha_2 = (-1, 1, 0, 1)^T$

$$N(T) \text{ 的基为 } X_1 = (A_1, \dots, A_4)\alpha_1 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \quad X_2 = (A_1, \dots, A_4)\alpha_2 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

M2004B

七、(15 分) 在向量空间 \mathbb{R}^4 中, 对于向量 $x = (\xi_1, \xi_2, \xi_3, \xi_4)$, 定义线性变换为

$$T(x) = (\xi_2 + \xi_4, \xi_1 + \xi_3, \xi_1 + \xi_3, \xi_2 + \xi_4)$$

1. 求 T 的值域 $R(T)$ 的一个基;

2. 将 T 看作 $R(T)$ 中的线性变换, 求 $R(T)$ 的一个基, 使 T 在该基下的矩阵为对角矩阵. (注意: $R(T)$ 是 T 的不变子空间)

解 1. 取 \mathbb{R}^4 的基为 e_1, e_2, e_3, e_4 , 计算基象组

$$T(e_1) = (0, 1, 1, 0), \quad T(e_2) = (1, 0, 0, 1)$$

$$T(e_3) = (0, 1, 1, 0), \quad T(e_4) = (1, 0, 0, 1)$$

$$\text{因为 } R(T) = L(T(e_1), T(e_2), T(e_3), T(e_4))$$

所以 $R(T)$ 的基为基象组的最大无关组: $\alpha_1 = (0, 1, 1, 0), \alpha_2 = (1, 0, 0, 1)$

$$2. \quad T(\alpha_1) = (1, 1, 1, 1) = \alpha_1 + \alpha_2, \quad T(\alpha_2) = (1, 1, 1, 1) = \alpha_1 + \alpha_2$$

$$T(\alpha_1, \alpha_2) = (\alpha_1, \alpha_2)A, \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{求 } P \text{ 使得 } P^{-1}AP = \Lambda: \quad \Lambda = \begin{bmatrix} 0 & \\ & 2 \end{bmatrix}, \quad P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

令 $(\beta_1, \beta_2) = (\alpha_1, \alpha_2)P$, 求得另一基:

$$\beta_1 = -\alpha_1 + \alpha_2 = (1, -1, -1, 1), \quad \beta_2 = \alpha_1 + \alpha_2 = (1, 1, 1, 1)$$

于是可得 $T(\beta_1, \beta_2) = (\beta_1, \beta_2)\Lambda$.

第二章 范数理论及其应用

§ 2.1 向量范数

一、 \mathbb{C}^n 中向量序列的收敛性

设 $x^{(k)} = (\xi_1^{(k)}, \dots, \xi_n^{(k)})$, 若 $\lim_{k \rightarrow \infty} \xi_i^{(k)} = \xi_i$ ($i = 1, 2, \dots, n$), 称 $\{x^{(k)}\}$ 收敛于

$x = (\xi_1, \dots, \xi_n)$, 记作 $\lim_{k \rightarrow \infty} x^{(k)} = x$ 或 $x^{(k)} \rightarrow x$ ($k \rightarrow \infty$).

[注] 判断一个向量序列收敛等价于判断 n 个数列同时收敛.

用模刻画: \mathbb{C}^n 中向量 x 的模为 $|x| = \left(|\xi_1|^2 + \dots + |\xi_n|^2\right)^{\frac{1}{2}}$

$$\lim_{k \rightarrow \infty} x^{(k)} = x \Leftrightarrow \lim_{k \rightarrow \infty} \xi_i^{(k)} = \xi_i \quad (i = 1, 2, \dots, n) \Leftrightarrow \lim_{k \rightarrow \infty} |\xi_i^{(k)} - \xi_i|^2 = 0 \quad (i = 1, 2, \dots, n)$$

$$\Leftrightarrow \lim_{k \rightarrow \infty} \left(|\xi_1^{(k)} - \xi_1|^2 + \dots + |\xi_n^{(k)} - \xi_n|^2\right)^{\frac{1}{2}} = 0 \Leftrightarrow \lim_{k \rightarrow \infty} |x^{(k)} - x| = 0$$

二、线性空间 V 的向量范数

线性空间 V , 数域 K , $\forall x \in V$, 定义实数 $\|x\|$, 且满足

$$(1) \|x\| \geq 0; \|x\| = 0 \Leftrightarrow x = \theta$$

$$(2) \|kx\| = |k| \|x\|, \forall k \in K$$

$$(3) \|x + y\| \leq \|x\| + \|y\|, \forall y \in V$$

实数的对应
规则不唯一

称 $\|x\|$ 为向量 x 的范数.

例 1 欧氏空间 V 中, $\|x\| = |x| = \sqrt{(x, x)}$ 是一种向量范数.

例 2 线性空间 \mathbb{C}^n 中, $\|x\|_p = \left(\sum_{i=1}^n |\xi_i|^p\right)^{\frac{1}{p}}$ ($1 \leq p < \infty$) 是向量范数.

证 ① 略. ② 略. ③ 设 $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$, 则

$$p=1: \|x + y\|_1 = \sum |\xi_i + \eta_i| \leq \sum (|\xi_i| + |\eta_i|) = \|x\|_1 + \|y\|_1;$$

$p > 1$: $x + y = \theta$ 时, 结论成立; $x + y \neq \theta$ 时, 应用 Hölder 不等式

$$\sum |a_i b_i| \leq \left(\sum |a_i|^p\right)^{\frac{1}{p}} \left(\sum |b_i|^q\right)^{\frac{1}{q}} \quad \left(p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1\right)$$

可得 (利用 $(p-1)q = p$)

$$\begin{aligned} (\|x + y\|_p)^p &= \sum |\xi_i + \eta_i|^p = \sum (|\xi_i + \eta_i| |\xi_i + \eta_i|^{p-1}) \\ &\leq \sum |\xi_i| |\xi_i + \eta_i|^{p-1} + \sum |\eta_i| |\xi_i + \eta_i|^{p-1} \\ &\leq \left(\sum |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum (|\xi_i + \eta_i|^{p-1})^q\right)^{\frac{1}{q}} + \left(\sum |\eta_i|^p\right)^{\frac{1}{p}} \left(\sum (|\xi_i + \eta_i|^{p-1})^q\right)^{\frac{1}{q}} \\ &= (\|x\|_p + \|y\|_p) (\|x + y\|_p)^{p-1} \end{aligned}$$

(利用 $\frac{1}{q} = \frac{1}{p} \cdot (p-1)$)

$$\text{故 } \|x+y\|_p \leq \|x\|_p + \|y\|_p$$

因此, 当 $1 \leq p < +\infty$ 时, $\|x\|_p$ 是向量范数.

例 证明极限式 $\lim_{p \rightarrow +\infty} \|x\|_p = \max_i |\xi_i|$.

证 当 $x = \theta$ 时成立. 当 $x \neq \theta$ 时, 设 $|\xi_{i_0}| = \max_i |\xi_i|$, 则有

$$\|x\|_p = |\xi_{i_0}| \left(\sum \left| \frac{\xi_i}{\xi_{i_0}} \right|^p \right)^{\frac{1}{p}}, \quad 1 \leq \left(\sum \left| \frac{\xi_i}{\xi_{i_0}} \right|^p \right)^{\frac{1}{p}} \leq n^{\frac{1}{p}} \rightarrow 1 \quad (p \rightarrow +\infty)$$

$$\lim_{p \rightarrow +\infty} \|x\|_p = |\xi_{i_0}| = \max_i |\xi_i|$$

例 证明 $\|x\|_\infty = \max_i |\xi_i|$ 是 C^n 中的向量范数.

证 ① 略. ② 略. ③ 设 $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$, 则

$$\|x+y\|_\infty = \max_i |\xi_i + \eta_i| \leq \max_i [|\xi_i| + |\eta_i|] \leq \max_i |\xi_i| + \max_i |\eta_i| = \|x\|_\infty + \|y\|_\infty$$

故 $\|x\|_\infty = \max_i |\xi_i|$ 是向量范数.

常用向量范数: **1-范数** $\|x\|_1 = \sum |\xi_i|$; **2-范数** $\|x\|_2 = \left(\sum |\xi_i|^2 \right)^{\frac{1}{2}}$

$$\infty\text{-范数 } \|x\|_\infty = \max_i |\xi_i|$$

例 3 线性空间 V^n 中, 给定基 x_1, \dots, x_n , 对任意 $x \in V^n$, 有

$$x = \xi_1 x_1 + \dots + \xi_n x_n \xleftrightarrow{1-1} \alpha = (\xi_1, \dots, \xi_n)^T \in C^n$$

定义 $\|x\|_p = \|\alpha\|_p$, 则 $\|x\|_p$ 是 V^n 中元素 x 的范数 (自证), 称为 p -范数.

例 4 设 $A \in \mathbf{R}^{n \times n}$ 对称正定, 线性空间 \mathbf{R}^n 中, $\|\alpha\|_A = (\alpha^T A \alpha)^{\frac{1}{2}}$ 是向量范数.

证 ① 略. ② 略.

$$\text{③ } A \text{ 实对称} \Rightarrow \exists \text{ 正交矩阵 } Q \text{ 使得 } Q^T A Q = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$A \text{ 正定} \Rightarrow \lambda_i > 0$$

$$\Rightarrow A = Q \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} Q^T \stackrel{\Delta}{=} B^T B$$

$$\text{因为 } \|\alpha\|_A = (\alpha^T B^T B \alpha)^{\frac{1}{2}} = [(B\alpha)^T (B\alpha)]^{\frac{1}{2}} = \|B\alpha\|_2 \quad (\mathbf{R}^n \text{ 中})$$

$$\text{所以 } \|\alpha + \beta\|_A = \|B(\alpha + \beta)\|_2 = \|(B\alpha) + (B\beta)\|_2$$

$$\leq \|B\alpha\|_2 + \|B\beta\|_2 = \|\alpha\|_A + \|\beta\|_A$$

三、范数等价

对于线性空间 V^n 的向量范数 $\|x\|_\alpha$ 与 $\|x\|_\beta$, 若有正常数 c_1 和 c_2 , 使得

$$c_1 \|x\|_\beta \leq \|x\|_\alpha \leq c_2 \|x\|_\beta \quad (\forall x \in V^n)$$

成立, 称 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 等价.

(1) 自反性: $1 \cdot \|x\|_\alpha \leq \|x\|_\alpha \leq 1 \cdot \|x\|_\alpha, \forall x \in V^n$

(2) 对称性: $\frac{1}{c_2} \|x\|_\alpha \leq \|x\|_\beta \leq \frac{1}{c_1} \|x\|_\alpha, \forall x \in V^n$

(3) 传递性: $\left. \begin{array}{l} c_1 \|x\|_\beta \leq \|x\|_\alpha \leq c_2 \|x\|_\beta \\ c'_1 \|x\|_\gamma \leq \|x\|_\beta \leq c'_2 \|x\|_\gamma \end{array} \right\} \Rightarrow c_1'' \|x\|_\gamma \leq \|x\|_\alpha \leq c_2'' \|x\|_\gamma, \forall x \in V^n$

例 5 向量空间 C^n 中, 对任意向量 $x = (\xi_1, \dots, \xi_n)^T$, 有

$$(1) \|x\|_1 = \sum |\xi_i| \leq n \cdot \max_i |\xi_i| = n \|x\|_\infty$$

$$\|x\|_1 \geq \max_i |\xi_i| = 1 \cdot \|x\|_\infty, \text{ 故 } 1 \cdot \|x\|_\infty \leq \|x\|_1 \leq n \cdot \|x\|_\infty.$$

$$(2) \|x\|_2 = \left(\sum |\xi_i|^2 \right)^{\frac{1}{2}} \leq \left(n \cdot \max_i |\xi_i|^2 \right)^{\frac{1}{2}} = \sqrt{n} \cdot \|x\|_\infty$$

$$\|x\|_2 \geq \left(\max_i |\xi_i|^2 \right)^{\frac{1}{2}} = 1 \cdot \|x\|_\infty, \text{ 故 } 1 \cdot \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \cdot \|x\|_\infty.$$

$$(3) \frac{1}{\sqrt{n}} \cdot \|x\|_2 \leq \|x\|_1 \leq n \cdot \|x\|_2$$

Th1 线性空间 V^n 中, 任意两种向量范数等价.

证 只需证明任一向量范数 $\|x\|_\alpha$ 与向量范数 $\|x\|_2$ 等价即可.

给定 V^n 的基 x_1, \dots, x_n , 对任意 $x \in V^n$, 有

$$x = \xi_1 x_1 + \dots + \xi_n x_n, \quad \|x\|_2 = \left(\sum_{i=1}^n |\xi_i|^2 \right)^{\frac{1}{2}} \quad (\text{见例 3})$$

C^n 的子集 $S = \{(\xi_1, \dots, \xi_n) \mid |\xi_1|^2 + \dots + |\xi_n|^2 = 1\}$ 是闭区域, 实值函数

$$f(\xi_1, \dots, \xi_n) \stackrel{\Delta}{=} \|x\|_\alpha = \|\xi_1 x_1 + \dots + \xi_n x_n\|_\alpha$$

在 S 上连续, 故 f 在 S 上取得最值 $\min f = c_1, \max f = c_2$.

$$\forall (\xi_1, \dots, \xi_n) \in S \Rightarrow x = \xi_1 x_1 + \dots + \xi_n x_n \neq \theta$$

$$\Rightarrow f(\xi_1, \dots, \xi_n) = \|x\|_\alpha > 0 \Rightarrow c_1 > 0$$

对 $\forall x \in V^n$, 当 $x \neq \theta$ 时, 由 $x = \xi_1 x_1 + \dots + \xi_n x_n$ 导出

$$y = \frac{x}{\|x\|_2} = \sum_{i=1}^n \frac{\xi_i}{\|x\|_2} x_i = \sum_{i=1}^n \eta_i x_i \quad (\eta_i = \frac{\xi_i}{\|x\|_2})$$

$$|\eta_1|^2 + \dots + |\eta_n|^2 = 1 \Rightarrow (\eta_1, \dots, \eta_n) \in S$$

$$\begin{aligned} \text{故 } 0 < c_1 \leq f(\eta_1, \dots, \eta_n) \leq c_2 &\Rightarrow 0 < c_1 \leq \|y\|_\alpha \leq c_2 \\ &\Rightarrow c_1 \|x\|_2 \leq \|x\|_\alpha \leq c_2 \|x\|_2 \end{aligned}$$

当 $x = \theta$ 时, 上式显然成立.

Th2 向量空间 C^n 中, $\lim_{k \rightarrow \infty} x^{(k)} = x \Leftrightarrow \forall \|x\|, \lim_{k \rightarrow \infty} \|x^{(k)} - x\| = 0$.

证 只需对 $\|x\| = \|x\|_1$ 证明即可.

在 V^n 中也成立

$$\begin{aligned} x^{(k)} \rightarrow x &\Leftrightarrow \xi_i^{(k)} \rightarrow \xi_i \quad (i = 1, 2, \dots, n) \Leftrightarrow |\xi_i^{(k)} - \xi_i| \rightarrow 0 \quad (i = 1, 2, \dots, n) \\ &\Leftrightarrow \sum_{i=1}^n |\xi_i^{(k)} - \xi_i| \rightarrow 0 \Leftrightarrow \|x^{(k)} - x\|_1 \rightarrow 0 \end{aligned}$$

§ 2.2 矩阵范数

一、矩阵范数

集合 $C^{m \times n}$ 中, $\forall A \in C^{m \times n}$, 定义实数 $\|A\|$, 且满足

- (1) $\|A\| \geq 0$; $\|A\| = 0 \Leftrightarrow A = O_{m \times n}$
- (2) $\|kA\| = |k| \cdot \|A\|$, $\forall k \in C$
- (3) $\|A + B\| \leq \|A\| + \|B\|$, $\forall B \in C^{m \times n}$
- (4) AB 有意义: $\|AB\| \leq \|A\| \cdot \|B\|$, $\forall B \in C^{n \times l}$

实数的对应
规则不唯一

称 $\|A\|$ 为矩阵 A 的范数.

在三个矩阵集合中
实数对应规则相同

二、矩阵范数与向量范数相容

设 $C^{m \times n}$ 的矩阵范数 $\|A\|_M$, C^m 与 C^n 中的同类向量范数 $\|x\|_V$, 若

$$\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V \quad (\forall A \in C^{m \times n}, \forall x \in C^n)$$

称矩阵范数 $\|A\|_M$ 与向量范数 $\|x\|_V$ 相容.

预备: 设 $a_i \geq 0$, $b_i \geq 0$, $\alpha = (a_1, \dots, a_n)$, $\beta = (b_1, \dots, b_n)$

$$\text{则有 } (\alpha, \beta) \leq \sqrt{(\alpha, \alpha)} \cdot \sqrt{(\beta, \beta)}$$

$$\text{或者 } \sum_{i=1}^n a_i b_i \leq \sqrt{\sum_{i=1}^n a_i^2} \cdot \sqrt{\sum_{i=1}^n b_i^2}$$

两组非负实数
对应相乘求和 \leq

例 6 $A = (a_{ij})_{m \times n}$, $x = (\xi_1, \xi_2, \dots, \xi_n)^T$.

(1) $\|A\|_{m_1} = \sum_{i,j} |a_{ij}|$ 是矩阵范数, 且与 $\|x\|_1$ 相容.

(2) $\|A\|_{m_2} = \left(\sum_{i,j} |a_{ij}|^2 \right)^{\frac{1}{2}}$ 是矩阵范数, 且与 $\|x\|_2$ 相容.

(3) $\|A\|_{m_\infty} = n \cdot \max_{i,j} |a_{ij}|$ 是矩阵范数, 且与 $\|x\|_\infty$ 相容.

证 (1) $1^0 \sim 3^0$ 成立. 先验证 $\|Ax\|_1 \leq \|A\|_{m_1} \cdot \|x\|_1$:

$$\begin{aligned}
\|Ax\|_1 &= \sum_{i=1}^m |a_{i1}\xi_1 + \cdots + a_{in}\xi_n| \leq \sum_{i=1}^m (|a_{i1}||\xi_1| + \cdots + |a_{in}||\xi_n|) \\
&\leq \sum_{i=1}^m [(|a_{i1}| + \cdots + |a_{in}|) \cdot (|\xi_1| + \cdots + |\xi_n|)] \\
&= \left[\sum_{i=1}^m (|a_{i1}| + \cdots + |a_{in}|) \right] \cdot (|\xi_1| + \cdots + |\xi_n|) = \|A\|_{m_1} \cdot \|x\|_1
\end{aligned}$$

4⁰ 划分 $B_{n \times l} = (b_1, \cdots, b_l)$, 则 $AB = (Ab_1, \cdots, Ab_l)$, 且有

$$\begin{aligned}
\|AB\|_{m_1} &= \|Ab_1\|_1 + \cdots + \|Ab_l\|_1 \leq \|A\|_{m_1} \|b_1\|_1 + \cdots + \|A\|_{m_1} \|b_l\|_1 \\
&= \|A\|_{m_1} \cdot (\|b_1\|_1 + \cdots + \|b_l\|_1) = \|A\|_{m_1} \cdot \|B\|_{m_1}
\end{aligned}$$

(2) 1⁰ 成立. 2⁰ 成立.

3⁰ 设 $B_{m \times n}$, 划分 $A = (a_1, \cdots, a_n)$, $B = (b_1, \cdots, b_n)$, 则有

$$\begin{aligned}
\|A+B\|_{m_2}^2 &= \|a_1+b_1\|_2^2 + \cdots + \|a_n+b_n\|_2^2 \\
&\leq (\|a_1\|_2 + \|b_1\|_2)^2 + \cdots + (\|a_n\|_2 + \|b_n\|_2)^2 \\
&\leq \|A\|_{m_2}^2 + 2(\|a_1\|_2 \|b_1\|_2 + \cdots + \|a_n\|_2 \|b_n\|_2) + \|B\|_{m_2}^2 \\
&\leq \|A\|_{m_2}^2 + 2\left(\sum \|a_i\|_2^2\right)^{\frac{1}{2}} \cdot \left(\sum \|b_i\|_2^2\right)^{\frac{1}{2}} + \|B\|_{m_2}^2 = (\|A\|_{m_2} + \|B\|_{m_2})^2
\end{aligned}$$

4⁰ 设 $B_{n \times l}$, $AB = \left(\sum_{k=1}^n a_{ik} b_{kj} \right)_{m \times l}$, 则有

$$\begin{aligned}
\|AB\|_{m_2}^2 &= \sum_{i,j} \left| \sum_k a_{ik} b_{kj} \right|^2 \leq \sum_{i,j} \left(\sum_k |a_{ik}| \cdot |b_{kj}| \right)^2 \\
&\leq \sum_{i,j} \left[\left(\sum_k |a_{ik}|^2 \right) \cdot \left(\sum_k |b_{kj}|^2 \right) \right] = \sum_i \left\{ \left(\sum_k |a_{ik}|^2 \right) \cdot \sum_j \left(\sum_k |b_{kj}|^2 \right) \right\} \\
&= \left(\sum_{i,k} |a_{ik}|^2 \right) \cdot \left(\sum_{k,j} |b_{kj}|^2 \right) = \|A\|_{m_2}^2 \cdot \|B\|_{m_2}^2
\end{aligned}$$

特别的, 取 $B = x \in C^{n \times 1}$, 则有

$$\|Ax\|_2 = \|AB\|_{m_2} \leq \|A\|_{m_2} \cdot \|B\|_{m_2} = \|A\|_{m_2} \cdot \|x\|_2$$

(3) 1⁰ ~ 3⁰ 成立. 4⁰ 设 $B = (b_{ij})_{n \times l}$, 则 $AB = \left(\sum_{k=1}^n a_{ik} b_{kj} \right)_{m \times l}$, 且有

$$\begin{aligned}
\|AB\|_{m_\infty} &= l \cdot \max_{i,j} \left| \sum_{k=1}^n a_{ik} b_{kj} \right| \leq l \cdot \max_{i,j} \left(\sum_{k=1}^n |a_{ik}| |b_{kj}| \right) \\
&\leq \left(n \cdot \max_{i,j} |a_{ij}| \right) \cdot \left(l \cdot \max_{i,j} |b_{ij}| \right) = \|A\|_{m_\infty} \cdot \|B\|_{m_\infty} \\
\|Ax\|_\infty &= \max_i |a_{i1}\xi_1 + \cdots + a_{in}\xi_n| \leq \max_i (|a_{i1}||\xi_1| + \cdots + |a_{in}||\xi_n|)
\end{aligned}$$

$$\leq \max_i (|a_{i1}| + \cdots + |a_{in}|) \cdot \max_i |\xi_i| \leq \left(n \cdot \max_i |a_{ij}| \right) \cdot \max_i |\xi_i| = \|A\|_{m_\infty} \cdot \|x\|_\infty$$

[注] ① $\|A\|_{m_2} = [\text{tr}(A^H A)]^{\frac{1}{2}} = [\text{tr}(A A^H)]^{\frac{1}{2}}$, 记作 $\|A\|_F$.

② $C^{m \times n}$ 中的矩阵范数等价: 对任意的两种矩阵范数 $\|A\|_\alpha$ 与 $\|A\|_\beta$,

存在 $0 < c_1 \leq c_2$, 使得 $c_1 \|A\|_\beta \leq \|A\|_\alpha \leq c_2 \|A\|_\beta$ ($\forall A_{m \times n}$).

③ $C^{m \times n}$ 中, $\lim_{k \rightarrow \infty} A^{(k)} = A \Leftrightarrow \forall \|A\|, \lim_{k \rightarrow \infty} \|A^{(k)} - A\| = 0$.

Th3 对于 $A_{m \times n}$ 及酉矩阵 $P_{m \times m}$ 和 $Q_{n \times n}$, 有

$$\|PA\|_F = \|A\|_F, \quad \|AQ\|_F = \|A\|_F$$

证 $\|PA\|_F^2 = \text{tr}((PA)^H(PA)) = \text{tr}(A^H P^H P A) = \text{tr}(A^H A) = \|A\|_F^2$

$$\|AQ\|_F^2 = \text{tr}(AQ(AQ)^H) = \text{tr}(AQQ^H A^H) = \text{tr}(AA^H) = \|A\|_F^2$$

引理 对 $C^{m \times n}$ 的矩阵范数 $\|A\|$, 存在向量范数 $\|x\|_V$, 使得 $\|Ax\|_V \leq \|A\| \cdot \|x\|_V$.

例 7 对 $C^{n \times n}$ 的矩阵范数 $\|A\|_M$, 任取 $0 \neq y \in C^n$ (列向量), 则

(1) $\|x\|_V = \|xy^T\|_M$ 是 C^n 的向量范数; (2) $\|A\|_M$ 与 $\|x\|_V$ 相容.

证 (1) 非负性. $x = 0$ 时, $\|x\|_V = \|0y^T\|_M = \|O\|_M = 0$;

$$x \neq 0 \text{ 时, } xy^T \neq O, \quad \|x\|_V = \|xy^T\|_M > 0.$$

齐次性. $\|kx\|_V = \|(kx)y^T\|_M = \|k(xy^T)\|_M = |k| \|xy^T\|_M = |k| \|x\|_V$ ($\forall k \in C$)

三角不等式. 任意 $x_1, x_2 \in C^n$, 有

$$\begin{aligned} \|x_1 + x_2\|_V &= \|(x_1 + x_2)y^T\|_M = \|x_1 y^T + x_2 y^T\|_M \\ &\leq \|x_1 y^T\|_M + \|x_2 y^T\|_M = \|x_1\|_V + \|x_2\|_V \end{aligned}$$

(2) $\|Ax\|_V = \|(Ax)y^T\|_M = \|A(xy^T)\|_M \leq \|A\|_M \cdot \|xy^T\|_M = \|A\|_M \cdot \|x\|_V$.

练习 给定非零列向量 $y \in C^m$, 对 $C^{m \times n}$ 的矩阵范数 $\|A\|_F$, 定义 $\|x\|_V = \|yx^T\|_F$

(列向量 $x \in C^n$), 则 $\|x\|_V$ 是 C^n 的向量范数, 且有 $\|Ax\|_V \leq \|A\|_F \cdot \|x\|_V$.

三、从属范数

定理 4 对 C^m 与 C^n 中的“同类向量范数” $\|x\|_V$, 定义实数

$$\|A\| = \max_{\|x\|_V=1} \|Ax\|_V \quad (\forall A_{m \times n}, x \in C^n)$$

则 $\|A\|$ 是 $C^{m \times n}$ 中矩阵 A 的范数, 且 $\|A\|$ 与 $\|x\|_V$ 相容.

[注] 等价定义 $\max_{\|x\|_V=1} \|Ax\|_V = \max_{x \neq \theta} \frac{\|Ax\|_V}{\|x\|_V}$ ($= \max_{x \neq \theta} \left\| A \cdot \frac{x}{\|x\|_V} \right\|_V = \max_{\|y\|_V=1} \|Ay\|_V$)

证 $1^0 \quad A = O : \|A\| = \max_{\|x\|_V=1} \|Ax\|_V = \max_{\|x\|_V=1} \|0\|_V = 0$

$A \neq O : \exists x_0$ 满足 $\|x_0\|_V = 1, \text{ st } Ax_0 \neq \theta \Rightarrow \|A\| \geq \|Ax_0\|_V > 0$

2^0 略

3^0 对 $A+B : \exists x_1$ 满足 $\|x_1\|_V = 1, \text{ st } \max_{\|x\|_V=1} \|(A+B)x\|_V = \|(A+B)x_1\|_V.$

$$\|A+B\| = \|Ax_1 + Bx_1\|_V \leq \|Ax_1\|_V + \|Bx_1\|_V \leq \|A\| + \|B\|$$

4^0 先证 $\|Ay\|_V \leq \|A\| \cdot \|y\|_V \quad (y \in C^n)$

$y = \theta$: 显然成立.

$y \neq \theta$: $y_0 = \frac{y}{\|y\|_V}$ 满足 $\|y_0\|_V = 1 \Rightarrow \|Ay_0\|_V \leq \max_{\|x\|_V=1} \|Ax\|_V = \|A\|$

$$\text{故 } \|Ay\|_V = \|A(\|y\|_V y_0)\|_V = \|Ay_0\|_V \cdot \|y\|_V \leq \|A\| \cdot \|y\|_V$$

对 AB , $\exists x_2$ 满足 $\|x_2\|_V = 1$, 使得 $\max_{\|x\|_V=1} \|(AB)x\|_V = \|(AB)x_2\|_V$, 于是

$$\|AB\| = \|(AB)x_2\|_V = \|A(Bx_2)\|_V \leq \|A\| \cdot \|Bx_2\|_V \leq \|A\| \cdot \|B\|$$

故定理成立.

[注] ① 一般的矩阵范数: $I = I \cdot I \Rightarrow \|I\| \leq \|I\| \cdot \|I\| \Rightarrow \|I\| \geq 1.$

例如 $\|I\|_{m_1} = n, \quad \|I\|_F = \sqrt{n}.$

② 矩阵的从属范数: $\|I\| = \max_{\|x\|_V=1} \|Ix\|_V = 1.$

③ 常用从属范数: $\|x\|_V \quad \|x\|_1 \quad \|x\|_2 \quad \|x\|_\infty$

$$\|A\|_M \quad \|A\|_1 \quad \|A\|_2 \quad \|A\|_\infty$$

Th5 设 $A = (a_{ij})_{m \times n}$, 则

(1) 列和范数 $\|A\|_1 = \max_j \left\{ \sum_{i=1}^m |a_{ij}| \right\}$

如何记忆?

(2) 谱范数 $\|A\|_2 = \sqrt{\lambda_1}, \quad \lambda_1 = \max \{ \lambda(A^H A) \}$

(3) 行和范数 $\|A\|_\infty = \max_i \left\{ \sum_{j=1}^n |a_{ij}| \right\}$

证 (1) 记 $t = \max_j \left(\sum_{i=1}^m |a_{ij}| \right)$. 左 \leq 右: 若 $x \in C^n$ 满足 $\|x\|_1 = 1$, 则

$$\begin{aligned} \|Ax\|_1 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} \xi_j \right| \leq \sum_i \sum_j |a_{ij}| \cdot |\xi_j| = \sum_j \left[|\xi_j| \cdot \left(\sum_i |a_{ij}| \right) \right] \\ &\leq \sum_j \left[|\xi_j| t \right] = t \cdot \|x\|_1 = \text{右} \end{aligned}$$

故 $\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 \leq \text{右}.$

左 \geq 右: 选取 k 使得 $t = \max_j \left(\sum_{i=1}^m |a_{ij}| \right) = \sum_{i=1}^m |a_{ik}|$, 令

$e_k = (0, \dots, 0, 1, 0, \dots, 0)^T$, 则 $\|e_k\|_1 = 1$, 且有

$$\|A\|_1 \geq \|Ae_k\|_1 = \left\| \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} \right\|_1 = \sum_{i=1}^m |a_{ik}| = \text{右}$$

(2) $A^H A$ 是 Hermite 矩阵 $\Rightarrow \lambda(A^H A) \in \mathbb{R}$ 且 $\lambda(A^H A) \geq 0$

$A^H A$ 的特征值: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$

$A^H A$ 的特征向量: x_1, x_2, \dots, x_n 两两正交且满足 $\|x_i\|_2 = 1$ 标准正交基

左 \leq 右: 若 $x \in \mathbb{C}^n$ 满足 $\|x\|_2 = 1$, 则 $x = \xi_1 x_1 + \dots + \xi_n x_n$, 且有

$$(x, x) = x^H x = \|x\|_2^2 = 1 \Rightarrow |\xi_1|^2 + \dots + |\xi_n|^2 = 1 \quad \text{复内积}$$

$$(A^H A)x = \lambda_1 \xi_1 x_1 + \dots + \lambda_n \xi_n x_n$$

$$\|Ax\|_2^2 = (Ax)^H (Ax) = x^H (A^H A)x = (x, A^H Ax)$$

$$= \lambda_1 |\xi_1|^2 + \dots + \lambda_n |\xi_n|^2 \leq \lambda_1$$

$$\text{故 } \|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 \leq \sqrt{\lambda_1}$$

$$\text{左} \geq \text{右: } \|Ax_1\|_2^2 = (Ax_1)^H (Ax_1) = x_1^H (A^H A)x_1 = \lambda_1 x_1^H x_1 = \lambda_1$$

$$\text{故 } \|A\|_2 \geq \|Ax_1\|_2 = \sqrt{\lambda_1}$$

(3) 记 $t = \max_i \left(\sum_{j=1}^n |a_{ij}| \right)$. 左 \leq 右: 若 $x \in \mathbb{C}^n$ 满足 $\|x\|_\infty = 1$, 则

$$\|Ax\|_\infty = \max_i \left| \sum_{j=1}^n a_{ij} \xi_j \right| \leq \max_i \left(\sum_{j=1}^n |a_{ij}| \cdot |\xi_j| \right)$$

$$\leq \left(\max_i \sum_{j=1}^n |a_{ij}| \right) \cdot \max_j |\xi_j| = t \|x\|_\infty = t$$

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty \leq \text{右}$$

左 \geq 右: 选取 k 使得 $t = \max_i \left(\sum_{j=1}^n |a_{ij}| \right) = \sum_{j=1}^n |a_{kj}|$, 令

$$y_0 = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}, \quad \eta_j = \begin{cases} 1 & (a_{kj} = 0) \\ \frac{|a_{kj}|}{a_{kj}} & (a_{kj} \neq 0) \end{cases} \Rightarrow |\eta_j| = 1$$

$$\text{则 } \|y_0\|_\infty = 1, Ay_0 = \begin{bmatrix} \sum_{j=1}^n a_{1j}\eta_j \\ \dots \\ \sum_{j=1}^n |a_{kj}| \\ \dots \\ \sum_{j=1}^n a_{mj}\eta_j \end{bmatrix} \Rightarrow \|Ay_0\|_\infty = t$$

由此可得 $\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty \geq \|Ay_0\|_\infty = t = \text{右}$

[注] (1) 若 P, Q 是酉矩阵, 则 $\|PAQ\|_2 = \|A\|_2$;

(2) 若 $A^H = A$, 则 $\|A\|_2 = \max |\lambda_A|$;

(3) 若 $A^H A = AA^H$, 则 $\|A\|_2 = \max |\lambda_A|$.

证(3) $A = PAP^H \Rightarrow \lambda_1(A^H A) = \lambda_1(P\bar{A}P^H P A P^H) = \lambda_1(\bar{A}A) = |\lambda_1(A)|^2$

§ 2.3 范数的应用

Th6 $\|A_{n \times n}\| < 1 \Rightarrow (I - A)$ 可逆, 且 $\|(I - A)^{-1}\| \leq \frac{\|I\|}{1 - \|A\|}$.

证 选取向量范数 $\|x\|_v$, 使得 $\|A\|$ 与 $\|x\|_v$ 相容 (例 7).

若 $\det(I - A) = 0$, 则 $(I - A)x = 0$ 有非零解 x_0 , 由 $(I - A)x_0 = 0$ 可得

$$x_0 = Ax_0 \Rightarrow \|x_0\|_v = \|Ax_0\|_v \leq \|A\| \cdot \|x_0\|_v < \|x_0\|_v$$

产生矛盾, 故 $(I - A)$ 可逆.

$$\begin{aligned} I &= (I - A) + A \Rightarrow (I - A)^{-1} = I + (I - A)^{-1}A \\ \Rightarrow \|(I - A)^{-1}\| &\leq \|I\| + \|(I - A)^{-1}\| \cdot \|A\| \Rightarrow \|(I - A)^{-1}\| \leq \frac{\|I\|}{1 - \|A\|} \end{aligned}$$

Th7 $\|A_{n \times n}\| < 1 \Rightarrow \|I - (I - A)^{-1}\| \leq \frac{\|A\|}{1 - \|A\|}$.

证 恒等式: $(I - A) - I = -A$

右乘 $(I - A)^{-1}$: $I - (I - A)^{-1} = -A(I - A)^{-1}$

左乘 A : $A - A(I - A)^{-1} = -A^2(I - A)^{-1}$

$$\Rightarrow A(I - A)^{-1} = A + A \cdot A(I - A)^{-1}$$

$$\Rightarrow \|A(I - A)^{-1}\| \leq \|A\| + \|A\| \cdot \|A(I - A)^{-1}\|$$

$$\Rightarrow \|A(I - A)^{-1}\| \leq \frac{\|A\|}{1 - \|A\|}$$

$$\Rightarrow \|I - (I - A)^{-1}\| = \|-A(I - A)^{-1}\| \leq \frac{\|A\|}{1 - \|A\|}$$

Th8 设 $A_{n \times n}$ 可逆, $B_{n \times n}$, 且满足 $\|A^{-1}B\| < 1$, 则

(1) $A+B$ 可逆;

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

(2) $F \stackrel{\Delta}{=} I - (I+A^{-1}B)^{-1} : \|F\| \leq \frac{\|A^{-1}B\|}{1-\|A^{-1}B\|};$

(3) $\frac{\|A^{-1} - (A+B)^{-1}\|}{\|A^{-1}\|} \leq \frac{\|A^{-1}B\|}{1-\|A^{-1}B\|}.$

$$A^{-1} - (A+B)^{-1} = [I - (I+A^{-1}B)^{-1}]A^{-1}$$

证 利用定理 6 和定理 7 可得.

谱半径: $\rho(A_{n \times n}) = \max_i |\lambda_i(A)|$

Th9 对 $\forall A_{n \times n}, \forall \|\cdot\|_M$, 有 $\rho(A) \leq \|A\|_M$.

证 对矩阵范数 $\|\cdot\|_M$, 存在向量范数 $\|\cdot\|_V$, 使得 $\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V$.

设 $Ax_i = \lambda_i x_i$ ($x_i \neq \theta$), 则有

$$|\lambda_i| \cdot \|x_i\|_V = \|\lambda_i x_i\|_V = \|Ax_i\|_V \leq \|A\|_M \cdot \|x_i\|_V$$

$$|\lambda_i| \leq \|A\|_M \Rightarrow \rho(A) \leq \|A\|_M$$

Th10 给定 $A_{n \times n}$, 对 $\forall \varepsilon > 0, \exists$ 矩阵范数 $\|\cdot\|_M$, st $\|A\|_M \leq \rho(A) + \varepsilon$.

证 根据矩阵的 Jordan 标准形理论: 对于矩阵 A , 存在可逆矩阵 $P_{n \times n}$, 使得 $P^{-1}AP = J$. 记

$$A = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}, \quad \tilde{I} = \begin{bmatrix} 0 & \delta_1 & & \\ & 0 & \delta_2 & \\ & & \ddots & \ddots \\ & & & 0 & \delta_{n-1} \\ & & & & 0 \end{bmatrix}$$

则 $J = A + \tilde{I}$ ($\delta_i = 1$ 或 0), 于是有

$$D = \begin{bmatrix} 1 & & & \\ & \varepsilon & & \\ & & \ddots & \\ & & & \varepsilon^{n-1} \end{bmatrix}, \quad (PD)^{-1}A(PD) = D^{-1}JD = A + \varepsilon \tilde{I}$$

$S \stackrel{\Delta}{=} PD$ 可逆: $\|S^{-1}AS\|_1 = \|A + \varepsilon \tilde{I}\|_1 \leq \rho(A) + \varepsilon$

可证 $\|B\|_M = \|S^{-1}BS\|_1$ ($\forall B \in C^{n \times n}$) 是 $C^{n \times n}$ 中的矩阵范数, 于是有

$$\|A\|_M = \|S^{-1}AS\|_1 \leq \rho(A) + \varepsilon.$$

[注] 因为 $\|\cdot\|_M$ 与给定的矩阵 A 有关, 所以当 $B_{n \times n} \neq A$ 时, 针对 A 构造的矩阵范数 $\|\cdot\|_M$, 不等式 $\|B\|_M \leq \rho(B) + \varepsilon$ 不一定成立!

讨论: ① $\|A\|_{M_1}, \|A\|_{M_2}$ 可与同一种 $\|x\|_V$ 相容?

② $\|A\|_M$ 可与不同的 $\|x\|_{V_1}, \|x\|_{V_2}$ 相容?

③ $\forall \|A\|_M$ 与 $\forall \|x\|_V$ 不一定相容?

分析: ① $\|A\|_{m_1}, \|A\|_1$ 与 $\|x\|_1$ 相容.

② $\|A\|_{m_1}$ 与 $\|x\|_p$ ($p \geq 1$) 相容. (例如 $p=1, p=2$ 及 $\|x\|_\infty$)

$$x = (\xi_1, \dots, \xi_n)^T, \quad E_{ij}x = (0, \dots, 0, \xi_j, 0, \dots, 0)^T \Rightarrow \|E_{ij}x\|_p \leq \|x\|_p$$

$$Ax = \sum_{i,j} a_{ij} E_{ij}x, \quad \|Ax\|_p \leq \sum_{i,j} |a_{ij}| \cdot \|E_{ij}x\|_p \leq \left(\sum_{i,j} |a_{ij}| \right) \cdot \|x\|_p = \|A\|_{m_1} \cdot \|x\|_p$$

③ $\|A\|_1 = \max_j \left(\sum_{i=1}^m |a_{ij}| \right)$ 与 $\|x\|_\infty = \max_i |\xi_i|$ 不相容.

$$n > 1: A_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m \times n}, \quad x_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n \times 1}, \quad A_0 x_0 = \begin{bmatrix} n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\|A_0\|_1 = 1, \quad \|x_0\|_\infty = 1, \quad \|A_0 x_0\|_\infty = n$$

$$\|A_0 x_0\|_\infty = n > 1 = \|A_0\|_1 \cdot \|x_0\|_\infty$$

构造方法

(1) 由向量范数构造新的向量范数 (已知 C^m 中的向量范数, 构造 C^n 中的向量范数): $S_{m \times n}$ 列满秩, $\|x\| = \|Sx\|_V$ 是 C^n 中的向量范数.

(2) 由矩阵范数构造向量范数 (已知 $C^{m \times n}$ 中的矩阵范数, 构造 C^m 中的向量范数): 非零列向量 $y_0 \in C^n$, $\|x\| = \|x y_0^T\|_M$ 是 C^m 中的向量范数.

$$\text{例如: } y_0 = e_i \text{ 时, } x y_0^T = \begin{bmatrix} 0 & \cdots & 0 & \xi_1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \xi_m & 0 & \cdots & 0 \end{bmatrix}$$

$$\|A\|_M = \|A\|_{m_1} \Rightarrow \|x\| = \|x\|_1; \quad \|A\|_M = \|A\|_2 \Rightarrow \|x\| = \|x\|_2$$

$$\|A\|_M = \|A\|_{m_2} \Rightarrow \|x\| = \|x\|_2; \quad \|A\|_M = \|A\|_\infty \Rightarrow \|x\| = \|x\|_\infty$$

(3) 由向量范数构造矩阵范数:

$$\|A\| = \max_{\|x\|_V=1} \|Ax\|_V \text{ 是 } C^{m \times n} \text{ 中的矩阵范数.}$$

(4) 由矩阵范数构造新的矩阵范数:

$$S_{n \times n} \text{ 可逆, } \|A\| = \|S^{-1}AS\|_M \text{ 是 } C^{n \times n} \text{ 中的矩阵范数.}$$

证(4) (1) $A = O$: $\|A\| = \|S^{-1}OS\|_M = \|O\|_M = 0$

$A \neq O$: 若 $S^{-1}AS = O$, 则 $A = O$, 矛盾! 故 $S^{-1}AS \neq O$.

从而 $\|A\| = \|S^{-1}AS\|_M > 0$

$$(2) \quad \|kA\| = \|S^{-1}(kA)S\|_M = \|k(S^{-1}AS)\|_M = |k| \|S^{-1}AS\|_M = |k| \|A\|$$

$$(3) \quad \|A+B\| = \|S^{-1}(A+B)S\|_M = \|S^{-1}AS + S^{-1}BS\|_M \\ \leq \|S^{-1}AS\|_M + \|S^{-1}BS\|_M = \|A\| + \|B\|$$

$$(4) \quad \|AB\| = \|S^{-1}(AB)S\|_M = \|(S^{-1}AS)(S^{-1}BS)\|_M \\ \leq \|S^{-1}AS\|_M \|S^{-1}BS\|_M = \|A\| \|B\|$$

例 8 (M2004A) 在向量空间 C^3 中, 对于向量 $x = (\xi_1, \xi_2, \xi_3)$, 定义实数

$\|x\| = \sqrt{|\xi_1|^2 + 2|\xi_2|^2 + 3|\xi_3|^2}$, 证明: $\|x\|$ 是 C^3 中的向量范数.

证 (1)-(2) 略! (3)

$$\text{方法 1.} \quad \|x+y\|^2 = |\xi_1 + \eta_1|^2 + 2|\xi_2 + \eta_2|^2 + 3|\xi_3 + \eta_3|^2 \\ \leq (|\xi_1| + |\eta_1|)^2 + 2(|\xi_2| + |\eta_2|)^2 + 3(|\xi_3| + |\eta_3|)^2 \\ = [|\xi_1|^2 + 2|\xi_2|^2 + 3|\xi_3|^2] + [|\eta_1|^2 + 2|\eta_2|^2 + 3|\eta_3|^2] + \\ 2[|\xi_1||\eta_1| + \sqrt{2}|\xi_2| \cdot \sqrt{2}|\eta_2| + \sqrt{3}|\xi_3| \cdot \sqrt{3}|\eta_3|] \\ \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$$

$$\text{方法 2.} \quad B = \begin{bmatrix} 1 & & \\ & \sqrt{2} & \\ & & \sqrt{3} \end{bmatrix}: \|x\| = \sqrt{|\xi_1|^2 + 2|\xi_2|^2 + 3|\xi_3|^2} = \|Bx^T\|_2$$

$$\|x+y\| = \|B(x+y)^T\|_2 = \|Bx^T + By^T\|_2 \leq \|Bx^T\|_2 + \|By^T\|_2 = \|x\| + \|y\|$$

例 9 (M2009A) 已知 C^n 上的向量范数 $\|\bullet\|_a$, 向量 $z = (z_1, \dots, z_n)^T \in C^n$,

对于向量 $x = (x_1, \dots, x_n)^T \in C^n$, 定义实数

$$\|x\|_b = \max \left\{ \frac{|x^H z|}{\|z\|_a} \mid 0 \neq z \in C^n \right\}$$

其中 $|x^H z|$ 表示复数 $x^H z$ 的模, 证明:

1. $\|x\|_b$ 是 C^n 上的向量范数;
2. 若取 $\|x\|_a = \|x\|_1$ (向量的 1-范数), 则 $\|x\|_b = \|x\|_\infty$ (向量的 ∞ -范数).

证 1. 任意非零向量 $z = (z_1, \dots, z_n)^T \in C^n$.

$$\textcircled{1} \quad x = 0: \|x\|_b = \max \frac{|\mathbf{0}^H z|}{\|z\|_a} = \max \frac{0}{\|z\|_a} = 0$$

$$x \neq 0: \|x\|_b = \max \frac{|x^H z|}{\|z\|_a} \geq \frac{|x^H x|}{\|x\|_a} > 0$$

② 略. ③ 设 $y \in C^n$, 则有

$$\|x + y\|_b = \max \frac{|(x + y)^H z|}{\|z\|_a} \leq \max \frac{|x^H z|}{\|z\|_a} + \max \frac{|y^H z|}{\|z\|_a} = \|x\|_b + \|y\|_b$$

故 $\|x\|_b$ 是 C^n 上的向量范数.

2. 因为 $|x^H z| = \left| \sum_{i=1}^n \bar{x}_i z_i \right| \leq \sum_{i=1}^n |x_i| |z_i| \leq (\max_i |x_i|) \sum_{i=1}^n |z_i| = \|x\|_\infty \|z\|_1$, 所以

$$\|x\|_b = \max \left\{ \frac{|x^H z|}{\|z\|_1} \mid 0 \neq z \in C^n \right\} \leq \|x\|_\infty$$

设 $|x_k| = \max_i |x_i| = \|x\|_\infty$, 则有 $|x^H e_k| = |x_k| = \|x\|_\infty \|e_k\|_1$, 由此可得

$$\|x\|_b = \max \left\{ \frac{|x^H z|}{\|z\|_1} \mid 0 \neq z \in C^n \right\} \geq \frac{|x^H e_k|}{\|e_k\|_1} = \|x\|_\infty$$

故 $\|x\|_b = \|x\|_\infty$.

例 10 设 A 是 n 阶实对称正定矩阵, $f(t)$ 是 m 次实系数多项式, 则对任意 $x \in R^n$, 有

$$\|f(A)x\|_A \leq [\max_{1 \leq i \leq n} |f(\lambda_i)|] \|x\|_A$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值, $\|x\|_A = \sqrt{x^T A x}$ 是 R^n 中的向量范数.

证 设 A 的对应于特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 的特征向量为 y_1, y_2, \dots, y_n , 且标准正交, 则有

$$x = \sum_{i=1}^n c_i y_i, \quad f(A)x = \sum_{i=1}^n c_i f(\lambda_i) y_i$$

$$\|x\|_A^2 = x^T A x = \left[\sum_{i=1}^n c_i y_i \right]^T \left[\sum_{i=1}^n c_i \lambda_i y_i \right] = \sum_{i=1}^n c_i^2 \lambda_i$$

$$\|f(A)x\|_A^2 = [f(A)x]^T A [f(A)x] = \left[\sum_{i=1}^n c_i f(\lambda_i) y_i \right]^T \left[\sum_{i=1}^n c_i f(\lambda_i) \lambda_i y_i \right]$$

$$= \sum_{i=1}^n c_i^2 \lambda_i f^2(\lambda_i) \leq [\max_{1 \leq i \leq n} f^2(\lambda_i)] \sum_{i=1}^n c_i^2 \lambda_i = [\max_{1 \leq i \leq n} |f(\lambda_i)|]^2 \|x\|_A^2$$

上式两端开平方即得所证.

第三章 矩阵分析及其应用

引言: 一元多项式 $f(t) = c_0 + c_1 t + \cdots + c_m t^m$

矩阵多项式 $f(A) = c_0 I + c_1 A + \cdots + c_m A^m \quad (\forall A_{n \times n})$

易见, $f(A)$ 是以矩阵为自变量且取值为矩阵的一类函数.

本章研究一般的以矩阵为自变量且取值为矩阵的函数——**矩阵函数**

§ 3.1 矩阵序列 (无穷多个有序的同阶矩阵)

一、敛散性: 将矩阵序列 $A^{(k)} = (a_{ij}^{(k)})_{m \times n} \quad (k=1, 2, 3, \dots)$ 记作 $\{A^{(k)}\}$.

若 $\lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij} \quad (\forall i, j)$, 称 $\{A^{(k)}\}$ 收敛于 $A = (a_{ij})_{m \times n}$, 记作

$$\lim_{k \rightarrow \infty} A^{(k)} = A \quad \text{或者} \quad A^{(k)} \rightarrow A \quad (k \rightarrow \infty)$$

若数列 $\{a_{ij}^{(k)}\}$ 之一发散, 称 $\{A^{(k)}\}$ 发散.

类似数列

性质 1 若 $A^{(k)} \rightarrow A_{m \times n}, B^{(k)} \rightarrow B_{m \times n}$, 则

$$A^{(k)} + B^{(k)} \rightarrow A + B, \quad \lambda A^{(k)} \rightarrow \lambda A \quad (\forall \lambda \in \mathbb{C})$$

性质 2 若 $A^{(k)} \rightarrow A_{m \times n}, B^{(k)} \rightarrow B_{n \times l}$, 则 $A^{(k)} B^{(k)} \rightarrow AB$.

性质 3 若 $A^{(k)}$ 与 A 是可逆矩阵, 且 $A^{(k)} \rightarrow A$, 则 $(A^{(k)})^{-1} \rightarrow A^{-1}$.

Th1 (1) $A^{(k)} \rightarrow O_{m \times n} \Leftrightarrow \forall \|\cdot\|, \|A^{(k)}\| \rightarrow 0$;

(2) $A^{(k)} \rightarrow A \Leftrightarrow \forall \|\cdot\|, \|A^{(k)} - A\| \rightarrow 0$.

证 (1) 考虑矩阵范数 $\|\cdot\|_{m_1}$:

$$\begin{aligned} A^{(k)} \rightarrow O_{m \times n} &\Leftrightarrow a_{ij}^{(k)} \rightarrow 0 \quad (\text{all } i, j) \Leftrightarrow |a_{ij}^{(k)}| \rightarrow 0 \quad (\text{all } i, j) \\ &\Leftrightarrow \sum_i \sum_j |a_{ij}^{(k)}| \rightarrow 0 \Leftrightarrow \|A^{(k)}\|_{m_1} \rightarrow 0 \quad (k \rightarrow \infty) \end{aligned}$$

$$\begin{aligned} (2) \quad A^{(k)} \rightarrow A &\Leftrightarrow a_{ij}^{(k)} \rightarrow a_{ij} \quad (\text{all } i, j) \Leftrightarrow (a_{ij}^{(k)} - a_{ij}) \rightarrow 0 \\ &\Leftrightarrow (A^{(k)} - A) \rightarrow O_{m \times n} \Leftrightarrow \forall \|\cdot\|, \|A^{(k)} - A\| \rightarrow 0 \end{aligned}$$

二、收敛矩阵: 若 $A_{n \times n}$ 满足 $A^k \rightarrow O_{n \times n}$, 称 A 为收敛矩阵.

Th2 A 为收敛矩阵 $\Leftrightarrow \rho(A) < 1$.

证 充分性. 已知 $\rho(A) < 1$, 对 $\varepsilon = \frac{1}{2}[1 - \rho(A)] > 0$, 存在矩阵范数 $\|\cdot\|_M$,

$$\text{使得} \quad \|A\|_M \leq \rho(A) + \varepsilon = \frac{1}{2}[1 + \rho(A)] < 1$$

于是有 $\|A^k\|_M \leq \|A\|_M^k \rightarrow 0$, 故由 **定理 1** 可得 $A^k \rightarrow O$.

必要性. 已知 $A^k \rightarrow O$, 设 $Ax = \lambda x \quad (x \neq 0)$, 则有

$$\lambda^k x = A^k x \rightarrow 0 \Rightarrow \lambda^k \rightarrow 0 \Rightarrow |\lambda| < 1$$

故 $\rho(A) < 1$.

Th3 若矩阵范数 $\|\bullet\|_M$ 使 $\|A\|_M < 1$, 则 $A^k \rightarrow O$.

证 $\rho(A) \leq \|A\|_M < 1 \Rightarrow A^k \rightarrow O$.

例如: $A = \begin{bmatrix} 0.1 & 0.3 \\ 0.7 & 0.6 \end{bmatrix}$, $\|A\|_1 = 0.9 < 1 \Rightarrow A^k \rightarrow O$

§ 3.2 矩阵级数 (无穷多个有序的同阶矩阵之和)

$$A^{(k)} = (a_{ij}^{(k)})_{m \times n} : A^{(0)} + A^{(1)} + \cdots + A^{(k)} + \cdots = \sum_{k=0}^{\infty} A^{(k)}$$

部分和 $S^{(N)} = \sum_{k=0}^N A^{(k)}$ 构成矩阵序列 $\{S^{(N)}\}$.

一、敛散性 若 $\lim_{N \rightarrow \infty} S^{(N)} = S$, 称 $\sum A^{(k)}$ 收敛于 S , 记作 $\sum A^{(k)} = S$;

若 $\{S^{(N)}\}$ 发散, 称 $\sum A^{(k)}$ 发散.

性质 1 $\sum A^{(k)} = S \Leftrightarrow \sum a_{ij}^{(k)} = s_{ij} \text{ (all } i, j)$

证 左端 $\Leftrightarrow \lim_{N \rightarrow \infty} S^{(N)} = S \Leftrightarrow \lim_{N \rightarrow \infty} s_{ij}^{(N)} = s_{ij} \text{ (all } i, j)$

$$\Leftrightarrow \lim_{N \rightarrow \infty} \sum_{k=0}^N a_{ij}^{(k)} = s_{ij} \text{ (all } i, j) \Leftrightarrow \text{右端}$$

性质 2 若 $\sum |a_{ij}^{(k)}|$ 收敛 (all i, j), 称 $\sum A^{(k)}$ **绝对收敛**.

(1) $\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum A^{(k)}$ 收敛;

(2) 若 $\sum A^{(k)}$ 绝对收敛于 S , 对 $\sum A^{(k)}$ 任意**重组重排**得 $\sum B^{(k)}$, 则 $\sum B^{(k)}$ 绝对收敛于 S .

性质 3 $\sum A^{(k)}$ 绝对收敛 $\Leftrightarrow \forall \|\bullet\|, \sum \|A^{(k)}\|$ 收敛.

证 只对矩阵范数 $\|\bullet\|_{m_1}$ 证明即可.

必要性. $\sum A^{(k)}$ 绝对收敛时, 正项级数 $\sum |a_{ij}^{(k)}|$ 收敛, 那么正项数列

$$p_{ij}^{(N)} = \sum_{k=0}^N |a_{ij}^{(k)}| \text{ 收敛, 于是 } \exists M_{ij} > 0, \text{ 使得 } p_{ij}^{(N)} \leq M_{ij} \text{ (} \forall N \text{),}$$

$$\text{从而有 } p_{ij}^{(N)} \leq M = \max_{i,j} M_{ij} \text{ (} \forall i, j, N \text{)}$$

$$\sum_{k=0}^N \|A^{(k)}\|_{m_1} = \sum_{k=0}^N \left(\sum_{i,j} |a_{ij}^{(k)}| \right) = \sum_{i,j} \left(\sum_{k=0}^N |a_{ij}^{(k)}| \right) \leq (mn)M$$

故 $\sum \|A^{(k)}\|_{m_1}$ 收敛. (正项级数收敛 \Leftrightarrow 部分和有上界)

充分性. $\sum \|A^{(k)}\|_{m_1}$ 收敛时, 因为 $|a_{ij}^{(k)}| \leq \sum_{i,j} |a_{ij}^{(k)}| = \|A^{(k)}\|_{m_1}$, 所以

$$\sum |a_{ij}^{(k)}| \text{ 收敛 (all } i, j) \Rightarrow \sum A^{(k)} \text{ 绝对收敛}$$

性质 4 $\sum A^{(k)} \overset{\text{收}}{=} S \Rightarrow \sum (PA^{(k)}Q) \overset{\text{收}}{=} PSQ$

$\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum (PA^{(k)}Q)$ 绝对收敛

证 (1) $S^{(N)} = \sum_{k=0}^N A^{(k)} \rightarrow S \Rightarrow \sum_{k=0}^N PA^{(k)}Q = PS^{(N)}Q \rightarrow PSQ$
 $\Rightarrow \sum PA^{(k)}Q = PSQ$

(2) 指定矩阵范数 $\|\bullet\|$, 由性质 3 知 $\sum \|A^{(k)}\|$ 收敛.

因为 $\|PA^{(k)}Q\| \leq \|P\| \|A^{(k)}\| \|Q\| = M \|A^{(k)}\| \quad \left(M \stackrel{\Delta}{=} \|P\| \|Q\| \right)$

所以 $\sum_{k=0}^N \|PA^{(k)}Q\| \leq \sum_{k=0}^N (M \|A^{(k)}\|) = M \cdot \sum_{k=0}^N \|A^{(k)}\|$ 有界

故 $\sum \|PA^{(k)}Q\|$ 收敛 $\Rightarrow \sum (PA^{(k)}Q)$ 绝对收敛

性质 5 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛于 $S_{m \times n}$, $\sum_{k=0}^{\infty} B^{(k)}$ 绝对收敛于 $T_{n \times l}$, 则 **Cauchy 乘积**

$$A^{(0)}B^{(0)} + [A^{(0)}B^{(1)} + A^{(1)}B^{(0)}] + [A^{(0)}B^{(2)} + A^{(1)}B^{(1)} + A^{(2)}B^{(0)}] + \dots \\ + [A^{(0)}B^{(k)} + A^{(1)}B^{(k-1)} + \dots + A^{(k)}B^{(0)}] + \dots$$

绝对收敛于 ST , 记作 $\sum A^{(k)} \cdot \sum B^{(k)} = ST$.

二、Neumann-级数: $I + A + A^2 + \dots = \sum_{k=0}^{\infty} A^k \quad (A_{n \times n}, A^0 = I)$

Th4 对 $A_{n \times n}$, $\sum A^k$ 收敛 $\Leftrightarrow A^k \rightarrow O$; $\sum A^k$ 收敛时, 其和为 $(I - A)^{-1}$.

证 必要性. 当 $\sum A^k$ 收敛时, 根据性质 1 可得

$$\sum (A^k)_{ij} \text{ 收敛 (all } i, j) \Rightarrow (A^k)_{ij} \rightarrow 0, \text{ 即 } A^k \rightarrow O.$$

充分性. 当 $A^k \rightarrow O$ 时, 由定理 2 知 $\rho(A) < 1$, 故 $I - A$ 可逆, 且有

$$(I + A + A^2 + \dots + A^N)(I - A) = I - A^{N+1} \\ (I + A + A^2 + \dots + A^N) = (I - A)^{-1} - A^{N+1}(I - A)^{-1} \\ \rightarrow (I - A)^{-1} \quad (N \rightarrow \infty)$$

$$\text{即 } \sum A^k = (I - A)^{-1}$$

Th5 $A_{n \times n}, \|A\| < 1 \Rightarrow \left\| (I - A)^{-1} - \sum_{k=0}^N A^k \right\| \leq \frac{\|A\|^{N+1}}{1 - \|A\|} \quad (N = 0, 1, 2, \dots)$

证 $\|A\| < 1 \Rightarrow \rho(A) < 1 \Rightarrow I - A$ 可逆

$$(I + A + A^2 + \dots + A^N)(I - A) = I - A^{N+1}$$

右乘 $(I - A)^{-1}$, 移项可得: $(I - A)^{-1} - \sum_{k=0}^N A^k = A^{N+1}(I - A)^{-1}$

$$\text{恒等式 } A^{N+1} = A^{N+1}(I - A)^{-1}(I - A)$$

$$= A^{N+1}(I-A)^{-1} - A^{N+1}(I-A)^{-1}A$$

$$\text{即 } A^{N+1}(I-A)^{-1} = A^{N+1} + A^{N+1}(I-A)^{-1} \cdot A$$

$$\|A^{N+1}(I-A)^{-1}\| \leq \|A^{N+1}\| + \|A^{N+1}(I-A)^{-1}\| \cdot \|A\|$$

$$\text{故 } \|A^{N+1}(I-A)^{-1}\| \leq \frac{\|A^{N+1}\|}{1-\|A\|} \Rightarrow \left\| (I-A)^{-1} - \sum_{k=0}^N A^k \right\| \leq \frac{\|A\|^{N+1}}{1-\|A\|}$$

三、矩阵幂级数

对于函数 $f(z) = \sum_{k=0}^{\infty} c_k z^k$ ($|z| < r$) 与方阵 $A_{n \times n}$, 构造 **矩阵幂级数** $\sum_{k=0}^{\infty} c_k A^k$.

Th6 (1) $\rho(A) < r \Rightarrow \sum c_k A^k$ 绝对收敛;

(2) $\rho(A) > r \Rightarrow \sum c_k A^k$ 发散.

证(1) 对 $A_{n \times n}$, 取 $\varepsilon = \frac{1}{2}[r - \rho(A)] > 0$, 存在矩阵范数 $\|\bullet\|_M$, 使得

$$\|A\|_M \leq \rho(A) + \varepsilon = \frac{1}{2}[r + \rho(A)] < r$$

当 $|z| < r$ 时, $\sum c_k z^k$ 绝对收敛, 即 $\sum |c_k| |z|^k$ 收敛. 因为 $\|A\|_M < r$, 所以

$\sum |c_k| \|A\|_M^k$ 收敛. 于是有

$$\|c_k A^k\|_M \leq |c_k| \|A\|_M^k \Rightarrow \sum \|c_k A^k\|_M \text{ 收敛} \Rightarrow \sum c_k A^k \text{ 绝对收敛}$$

(2) 设 A 的特征值为 $\lambda_1, \dots, \lambda_n$, 则存在可逆矩阵 P , 使得

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix} = S, \quad S^k = \begin{bmatrix} \lambda_1^k & \tilde{*} & \cdots & \tilde{*} \\ & \lambda_2^k & \ddots & \vdots \\ & & \ddots & \tilde{*} \\ & & & \lambda_n^k \end{bmatrix}$$

$\sum c_k S^k$ 的对角元素为 $\sum c_k \lambda_i^k$ ($i=1, 2, \dots, n$)

$\rho(A) > r \Rightarrow \exists |\lambda_i| > r \Rightarrow \sum c_k \lambda_i^k \text{ 发散} \Rightarrow \sum c_k S^k \text{ 发散} \Rightarrow \sum c_k A^k \text{ 发散}$

(否则: $\sum c_k A^k \text{ 收敛} \Rightarrow \sum (P^{-1}(c_k A^k)P) \text{ 收敛} \Rightarrow \sum c_k S^k \text{ 收敛}$)

例 讨论矩阵幂级数 $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^k$ 的收敛性.

根据定义

分析 矩阵 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ 的谱半径 $\rho(A) = 1$, 级数 $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} z^k$ 的收敛半径 $r = 1$.

$$A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^k = \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

因为 $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}$ 收敛, $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ 收敛, 所以 $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} A^k$ 收敛.

§ 3.3 矩阵函数

一、矩阵函数 设一元函数 $f(z)$ 在 $z=0$ 点的幂级数为

$$f(z) = c_0 + c_1 z + \cdots + c_k z^k + \cdots \quad (|z| < r, r > 0)$$

构造矩阵函数 $f(A) = c_0 I + c_1 A + \cdots + c_k A^k + \cdots \quad (\rho(A) < r)$

例 1 $e^z = 1 + \frac{1}{1!}z + \cdots + \frac{1}{k!}z^k + \cdots \quad (r = +\infty)$

$$e^A = I + \frac{1}{1!}A + \cdots + \frac{1}{k!}A^k + \cdots \quad (\forall A_{n \times n})$$

矩阵指数函数

$$\sin z = z - \frac{1}{3!}z^3 + \cdots + (-1)^k \frac{1}{(2k+1)!}z^{2k+1} + \cdots \quad (r = +\infty)$$

$$\sin A = A - \frac{1}{3!}A^3 + \cdots + (-1)^k \frac{1}{(2k+1)!}A^{2k+1} + \cdots \quad (\forall A_{n \times n})$$

例 2 $f(z) = \frac{1}{1-z} = \sum_{k=0}^{\infty} z^k \quad (|z| < 1), \quad f(A) = \sum_{k=0}^{\infty} A^k \quad (\rho(A) < 1).$

例 3 $\forall A_{n \times n}, \quad e^{jA} = \cos A + j \sin A \quad (j = \sqrt{-1})$

$$\cos A = \frac{1}{2}(e^{jA} + e^{-jA}), \quad \sin A = \frac{1}{2j}(e^{jA} - e^{-jA})$$

证 在第一式中, 视“ jA ”为整体, 并按“奇偶次幂”组项可得

$$\begin{aligned} e^{jA} &= \left[I + \frac{1}{2!}(jA)^2 + \frac{1}{4!}(jA)^4 + \cdots \right] + \left[\frac{1}{1!}(jA) + \frac{1}{3!}(jA)^3 + \cdots \right] \\ &= \cos A + j \sin A \end{aligned}$$

例 4 $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$, 求 e^A, e^B 及 e^{A+B} .

解 $A^2 = A: \quad e^A = I + \left(\frac{1}{1!} + \frac{1}{2!} + \cdots \right) A = I + (e-1)A = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}$

$$B^2 = B: \quad e^B = I + \left(\frac{1}{1!} + \frac{1}{2!} + \cdots \right) B = I + (e-1)B = \begin{bmatrix} e & 1-e \\ 0 & 1 \end{bmatrix}$$

$$A+B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 2I \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (A+B)^k = 2^k \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$e^{A+B} = I + \left(\frac{2^1}{1!} + \frac{2^2}{2!} + \cdots \right) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = I + (e^2-1) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$$

[注] $e^A e^B = \begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix}, \quad e^B e^A = \begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$

Th7 $A_{n \times n}, B_{n \times n}, AB = BA \Rightarrow e^A e^B = e^B e^A = e^{A+B}$

证 $e^A e^B = \left[I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots \right] \left[I + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \cdots \right]$

$$\begin{aligned}
 &= I + (A+B) + \frac{1}{2!}(A^2 + 2AB + B^2) + \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots \\
 &= I + (A+B) + \frac{1}{2!}(A+B)^2 + \frac{1}{3!}(A+B)^3 + \cdots = e^{A+B}
 \end{aligned}$$

同理 $e^B e^A = e^{B+A} = e^{A+B}$

e^A 一定可逆

[注] ① $e^A e^{-A} = e^O = I \Rightarrow (e^A)^{-1} = e^{-A} \quad (\forall A_{n \times n})$

② $(e^A)^m = e^{mA} \quad (m=2,3,\cdots)$

例 5 $A_{n \times n}, B_{n \times n}, AB = BA:$

$$\cos(A+B) = \cos A \cdot \cos B - \sin A \cdot \sin B \quad (1)$$

$$\sin(A+B) = \sin A \cdot \cos B + \cos A \cdot \sin B \quad (2)$$

证 式(1)中, 根据欧拉公式 (例 3) 可得

$$\begin{aligned}
 \text{右端} &= \frac{1}{2}[e^{jA} + e^{-jA}] \cdot \frac{1}{2}[e^{jB} + e^{-jB}] - \frac{1}{2j}[e^{jA} - e^{-jA}] \cdot \frac{1}{2j}[e^{jB} - e^{-jB}] \\
 &= \frac{1}{4}[e^{j(A+B)} + \cdots + \cdots + e^{-j(A+B)}] + \frac{1}{4}[e^{j(A+B)} - \cdots - \cdots + e^{-j(A+B)}] \\
 &= \frac{1}{2}[e^{j(A+B)} + e^{-j(A+B)}] = \text{左端}
 \end{aligned}$$

二、矩阵函数值的计算方法

$$f(A) = \sum_{k=0}^{\infty} c_k A^k \quad (\rho(A) < r)$$

1. 待定系数法 设以 $A_{n \times n}$ 为根的首 1 多项式为 (一般取 $\varphi(\lambda)$ 或者 $m(\lambda)$)

$$\psi(\lambda) = \lambda^m + c_1 \lambda^{m-1} + \cdots + c_{m-1} \lambda + c_m \quad (1 \leq m \leq n)$$

且满足 $\psi(\lambda) \mid \varphi(\lambda)$, 分解因式

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_s)^{m_s} \quad (\lambda_i \neq \lambda_j, \sum m_i = m)$$

因为 λ_i 是 A 的特征值, 而 $|\lambda_i| \leq \rho(A) < r$, 故 $f(\lambda_i) = \sum c_k \lambda_i^k$ 绝对收敛.

设 $f(z) = \sum c_k z^k = \psi(z) \cdot g(z) + r(z)$

$$r(z) = b_0 + b_1 z + \cdots + b_{m-1} z^{m-1}$$

系数待定

由 $\psi(\lambda_i) = 0, \psi'(\lambda_i) = 0, \cdots, \psi^{(m_i-1)}(\lambda_i) = 0$ 可得

$$r(\lambda_i) = f(\lambda_i), r'(\lambda_i) = f'(\lambda_i), \cdots, r^{(m_i-1)}(\lambda_i) = f^{(m_i-1)}(\lambda_i) \quad (i=1,2,\cdots,s)$$

解此方程组 得出 $b_0, b_1, \cdots, b_{m-1}$. 因为 $\psi(A) = O$, 所以

$$f(A) = \sum c_k A^k = \psi(A) \cdot g(A) + r(A) = r(A)$$

即 $f(A) = b_0 I + b_1 A + \cdots + b_{m-1} A^{m-1}$

例 6 $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{bmatrix}$, 求 $e^A, e^{tA} (t \in \mathbb{R})$.

解 $\varphi(\lambda) = |\lambda I - A| = (\lambda - 2)^3$, $A - 2I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$, $(A - 2I)^2 = O$

取 $\psi(\lambda) = m(\lambda) = (\lambda - 2)^2$

(1) $f(\lambda) = e^\lambda = \psi(\lambda) \cdot g(\lambda) + (a + b\lambda)$, $f'(\lambda) = e^\lambda = [\psi(\lambda) \cdot g(\lambda)]' + b$

$$\begin{cases} f(2) = e^2 : a + 2b = e^2 \\ f'(2) = e^2 : b = e^2 \end{cases} \Rightarrow \begin{cases} a = -e^2 \\ b = e^2 \end{cases}, \quad e^A = e^2(A - I) = e^2 \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

(2) $f(\lambda) = e^{t\lambda} = \psi(\lambda) \cdot g(\lambda) + (a + b\lambda)$, $f'(\lambda) = te^{t\lambda} = [\psi(\lambda) \cdot g(\lambda)]' + b$

$$\begin{cases} f(2) = e^{2t} : a + 2b = e^{2t} \\ f'(2) = te^{2t} : b = te^{2t} \end{cases} \Rightarrow \begin{cases} a = (1 - 2t)e^{2t} \\ b = te^{2t} \end{cases}$$

$$e^{tA} = e^{2t}[(1 - 2t)I + tA] = e^{2t} \begin{bmatrix} 1 & 0 & 0 \\ t & 1 - t & t \\ t & -t & 1 + t \end{bmatrix}$$

2. 数项级数求和法 设以 $A_{n \times n}$ 为根的首 1 多项式 $\psi(\lambda)$ 如前所述, 则有

$$\begin{aligned} \psi(A) = O &\Rightarrow \begin{cases} A^m = k_0 I + k_1 A + \cdots + k_{m-1} A^{m-1} \\ A^{m+1} = k_0^{(1)} I + k_1^{(1)} A + \cdots + k_{m-1}^{(1)} A^{m-1} \\ \dots\dots\dots \\ A^{m+l} = k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)} A^{m-1} \end{cases} \\ f(A) = \sum c_k A^k &= (c_0 I + c_1 A + \cdots + c_{m-1} A^{m-1}) + \\ &\quad c_m (k_0 I + k_1 A + \cdots + k_{m-1} A^{m-1}) + \cdots + \\ &\quad c_{m+l} (k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)} A^{m-1}) + \cdots \\ &= \left[c_0 + \sum_{l=0}^{\infty} c_{m+l} k_0^{(l)} \right] I + \left[c_1 + \sum_{l=0}^{\infty} c_{m+l} k_1^{(l)} \right] A + \cdots + \left[c_{m-1} + \sum_{l=0}^{\infty} c_{m+l} k_{m-1}^{(l)} \right] A^{m-1} \end{aligned}$$

例 7 $A = \begin{bmatrix} \pi & 0 & 0 & 0 \\ & -\pi & 0 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$, 求 $\sin A$.

解 $\varphi(\lambda) = |\lambda I - A| = \lambda^4 - \pi^2 \lambda^2$, 取 $\psi(\lambda) = \varphi(\lambda)$:

$$\psi(A) = O \Rightarrow A^4 = \pi^2 A^2, A^5 = \pi^2 A^3, A^7 = \pi^4 A^3, \dots$$

$$\sin A = A - \frac{1}{3!} A^3 + \frac{1}{5!} A^5 - \frac{1}{7!} A^7 + \dots$$

$$\begin{aligned}
&= A + \left[-\frac{1}{3!} + \frac{\pi^2}{5!} - \frac{\pi^4}{7!} + \cdots \right] A^3 = A + \frac{1}{\pi^3} \left[-\frac{\pi^3}{3!} + \frac{\pi^5}{5!} - \frac{\pi^7}{7!} + \cdots \right] A^3 \\
&= A + \frac{1}{\pi^3} (\sin \pi - \pi) A^3 = A - \frac{1}{\pi^2} A^3 \\
&= \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}; \quad A^3 = \begin{bmatrix} \pi^3 & & & \\ & -\pi^3 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}
\end{aligned}$$

3. 对角形法 考虑函数 $f(z) = \sum_{k=0}^{\infty} c_k z^k$ ($|z| < r$), 设

$$\begin{aligned}
P^{-1}AP &= \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \Lambda, \text{ 则 } A^k = P\Lambda^k P^{-1}, \text{ 且有} \\
\sum_{k=0}^N c_k A^k &= P \cdot \sum_{k=0}^N c_k \Lambda^k \cdot P^{-1} = P \cdot \begin{bmatrix} \sum_{k=0}^N c_k \lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^N c_k \lambda_n^k \end{bmatrix} \cdot P^{-1} \\
f(A) &= \sum_{k=0}^{\infty} c_k A^k = P \cdot \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} \cdot P^{-1}
\end{aligned}$$

例 8 $P^{-1}AP = \Lambda$: $e^A = P \cdot \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \cdot P^{-1}$

$$e^{tA} = P \cdot \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \cdot P^{-1}$$

$$\sin A = P \cdot \text{diag}(\sin \lambda_1, \dots, \sin \lambda_n) \cdot P^{-1}$$

4. Jordan 标准形法 设

$$P^{-1}AP = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}, \quad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{m_i \times m_i} = \lambda_i I + I^{(1)}$$

易证: $I^{(k)} I^{(1)} = I^{(1)} I^{(k)} = I^{(k+1)}, \quad I^{(m_i)} = O$

$$k \leq m_i - 1: J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{k-1} \lambda_i I^{(k-1)} + I^{(k)}$$

$$k \geq m_i: J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{m_i-1} \lambda_i^{k-(m_i-1)} I^{(m_i-1)}$$

$$f(J_i) = \sum c_k J_i^k = f(\lambda_i) I + \frac{f'(\lambda_i)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}(\lambda_i)}{(m_i-1)!} I^{(m_i-1)}$$

$$f(A) = \sum c_k A^k = P \cdot \sum c_k J^k \cdot P^{-1}$$

$$= P \cdot \begin{bmatrix} \sum c_k J_1^k & & \\ & \ddots & \\ & & \sum c_k J_s^k \end{bmatrix} \cdot P^{-1} = P \cdot \begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{bmatrix} \cdot P^{-1}$$

三、矩阵函数的拓宽定义

展开式 $f(z) = \sum c_k z^k$ ($|z| < r, r > 0$) 要求:

$$\textcircled{1} f^{(k)}(0) \text{ 存在 } (k=0,1,2,\dots); \quad \textcircled{2} \lim_{k \rightarrow \infty} \frac{f^{(k+1)}(\xi)}{(k+1)!} z^{k+1} = 0 \quad (|z| < r)$$

对于一元函数 $f(z) = \frac{1}{z}$ 等, 还不能定义矩阵函数. 基于矩阵函数值的

Jordan 标准形算法, 拓宽定义如下: 设 $f(z) \in C^{m_i-1}[\delta(\lambda_i)]$, 令

$$f(J_i) = f(\lambda_i)I + \frac{f'(\lambda_i)}{1!}I^{(1)} + \dots + \frac{f^{(m_i-1)}(\lambda_i)}{(m_i-1)!}I^{(m_i-1)}$$

$$f(A) = P \cdot \text{diag}(f(J_1), \dots, f(J_s)) \cdot P^{-1}$$

称 $f(A)$ 为对应于 $f(z)$ 的矩阵函数.

[注] ① 拓宽定义不要求 $f(z)$ 能展为 z 的幂级数, 但要求 $f(z)$ 在 A 的特征值 λ_i (重数为 m_i) 处有 m_i-1 阶导数, 后者较前者弱!

② 当 $f(z)$ 能够展为“ z ”的幂级数时, 矩阵函数 $f(A)$ 的拓宽定义与级数定义是一致的.

例 9 $A = \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}, \quad f(z) = \frac{1}{z}, \quad \text{求 } f(A).$

矩阵 A 可逆时
不需要 $f^{(k)}(0)$
只需要 $f^{(k)}(\lambda_i)$

解 $f(z) = \frac{1}{z}, \quad f'(z) = -z^{-2}, \quad f''(z) = 2z^{-3}, \quad f'''(z) = -6z^{-4}$

$$f(A) = f(J) = f(2) \cdot I + f'(2) \cdot I^{(1)} + \frac{f''(2)}{2!} \cdot I^{(2)} + \frac{f'''(2)}{3!} \cdot I^{(3)}$$

$$= \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & \frac{1}{8} & -\frac{1}{16} \\ & \frac{1}{2} & -\frac{1}{4} & \frac{1}{8} \\ & & \frac{1}{2} & -\frac{1}{4} \\ & & & \frac{1}{2} \end{bmatrix}$$

例 10 $A = \begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 2 \end{bmatrix}, \quad f(z) = \sqrt{z}, \quad \text{求 } f(A).$

解 $f(z) = \sqrt{z}$, $f'(z) = \frac{1}{2\sqrt{z}}$

$$J_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}: f(J_1) = f(1) \cdot I + f'(1) \cdot I^{(1)} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}$$

$$J_2 = (2): f(J_2) = f(2) \cdot I = (\sqrt{2})$$

$$f(A) = f(J) = \begin{bmatrix} f(J_1) & \\ & f(J_2) \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 0 \\ & 1 & 0 \\ & & \sqrt{2} \end{bmatrix}$$

四、矩阵函数的性质

由级数定义或者拓宽定义给出的矩阵函数具有下列性质:

$$(1) f(z) = f_1(z) + f_2(z) \Rightarrow f(A) = f_1(A) + f_2(A)$$

$$f^{(l)}(\lambda_i) = f_1^{(l)}(\lambda_i) + f_2^{(l)}(\lambda_i) \Rightarrow f(J_i) = f_1(J_i) + f_2(J_i)$$

$$\begin{aligned} f(A) &= P \cdot \left[\begin{pmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{pmatrix} + \begin{pmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{pmatrix} \right] \cdot P^{-1} \\ &= f_1(A) + f_2(A) \end{aligned}$$

$$(2) f(z) = f_1(z) \cdot f_2(z) \Rightarrow f(A) = f_1(A) \cdot f_2(A) = f_2(A) \cdot f_1(A)$$

$$\begin{aligned} f_1(J_i) \cdot f_2(J_i) &= \left[f_1 \cdot I + \frac{f_1'}{1!} \cdot I^{(1)} + \frac{f_1''}{2!} \cdot I^{(2)} + \dots + \frac{f_1^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right] \cdot \\ &\quad \left[f_2 \cdot I + \frac{f_2'}{1!} \cdot I^{(1)} + \frac{f_2''}{2!} \cdot I^{(2)} + \dots + \frac{f_2^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right] \\ &= (f_1 f_2) \cdot I + \frac{1}{1!} (f_1' f_2 + f_1 f_2') \cdot I^{(1)} + \frac{1}{2!} (f_1'' f_2 + f_1 f_2'' + 2 f_1' f_2') \cdot I^{(2)} + \dots \\ &= (f_1 f_2) \cdot I + \frac{1}{1!} (f_1 f_2)' \cdot I^{(1)} + \frac{1}{2!} (f_1 f_2)'' \cdot I^{(2)} + \dots + \frac{(f_1 f_2)^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \\ &= f(J_i) \\ f(A) &= P \begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{bmatrix} P^{-1} \cdot P \begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{bmatrix} P^{-1} \\ &= f_1(A) \cdot f_2(A) \end{aligned}$$

§ 3.4 矩阵的微分与积分

一、函数矩阵的导数

设 $A(t) = (a_{ij}(t))_{m \times n}$, 若 $a_{ij}(t)$ 可导, 称 $A(t)$ 可导, 记作

$$\frac{d}{dt} A(t) = (a'_{ij}(t))_{m \times n} \quad \text{或者} \quad A'(t) = (a'_{ij}(t))_{m \times n}$$

Th8 设 $A(t), B(t)$ 可导, 则有

- (1) $A_{m \times n}, B_{m \times n}$, $\frac{d}{dt}[A(t) + B(t)] = A'(t) + B'(t)$
- (2) $A_{m \times n}, f(t)$ 可导, $\frac{d}{dt}[f(t)A(t)] = f'(t)A(t) + f(t)A'(t)$
- (3) $A_{m \times n}, B_{n \times l}$, $\frac{d}{dt}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t)$

证 (3) 左 = $\frac{d}{dt} \left(\sum_k a_{ik}(t)b_{kj}(t) \right)_{m \times l} = \left(\sum_k (a'_{ik}(t)b_{kj}(t) + a_{ik}(t)b'_{kj}(t)) \right)_{m \times l}$

$$= \left(\sum_k a'_{ik}(t)b_{kj}(t) \right)_{m \times l} + \left(\sum_k a_{ik}(t)b'_{kj}(t) \right)_{m \times l} = \text{右}$$

Th9 设 $A_{n \times n}$ 为常数矩阵, 则有

- (1) $\frac{d}{dt} e^{tA} = A \cdot e^{tA} = e^{tA} \cdot A$
- (2) $\frac{d}{dt} \cos(tA) = -A \cdot \sin(tA) = -\sin(tA) \cdot A$
- (3) $\frac{d}{dt} \sin(tA) = A \cdot \cos(tA) = \cos(tA) \cdot A$

证(1) $e^{tA} = I + \frac{1}{1!}(tA) + \frac{1}{2!}(tA)^2 + \cdots + \frac{1}{k!}(tA)^k + \cdots$ 绝对收敛

$$(e^{tA})_{ij} = \delta_{ij} + \frac{t}{1!}(A)_{ij} + \frac{t^2}{2!}(A^2)_{ij} + \cdots + \frac{t^k}{k!}(A^k)_{ij} + \cdots$$
 绝对收敛

$$\frac{d}{dt}(e^{tA})_{ij} = 0 + (A)_{ij} + \frac{t}{1!}(A^2)_{ij} + \cdots + \frac{t^{k-1}}{(k-1)!}(A^k)_{ij} + \cdots$$
 绝对收敛

$$\frac{d}{dt} e^{tA} = A + \frac{t}{1!}A^2 + \cdots + \frac{t^{k-1}}{(k-1)!}A^k + \cdots$$
 绝对收敛

$$= \begin{cases} A \cdot \left[I + \frac{1}{1!}(tA) + \cdots + \frac{1}{(k-1)!}(tA)^{k-1} + \cdots \right] = A \cdot e^{tA} \\ \left[I + \frac{1}{1!}(tA) + \cdots + \frac{1}{(k-1)!}(tA)^{k-1} + \cdots \right] \cdot A = e^{tA} \cdot A \end{cases}$$

二、函数矩阵的积分

设 $A(t) = (a_{ij}(t))_{m \times n}$, 若 $a_{ij}(t)$ 在区间 $[t_0, t]$ 上可积, 定义 $A(t)$ 的积分为

$$\int_{t_0}^t A(\tau) d\tau = \left(\int_{t_0}^t a_{ij}(\tau) d\tau \right)_{m \times n} ; \quad \int A(t) dt = \left(\int a_{ij}(t) dt \right)_{m \times n}$$

$$(1) \int_{t_0}^t [A(\tau) + B(\tau)] d\tau = \int_{t_0}^t A(\tau) d\tau + \int_{t_0}^t B(\tau) d\tau$$

$$(2) A \text{ 为常数矩阵: } \int_{t_0}^t [A \cdot B(\tau)] d\tau = A \cdot \left[\int_{t_0}^t B(\tau) d\tau \right]$$

$$B \text{ 为常数矩阵: } \int_{t_0}^t A(\tau) \cdot B d\tau = \left[\int_{t_0}^t A(\tau) d\tau \right] \cdot B$$

$$(3) \text{ 设 } a_{ij}(t) \in C[t_0, t_1], \text{ 且 } a \in [t_0, t_1], \text{ 则 } \frac{d}{dt} \int_a^t A(\tau) d\tau = A(t) .$$

$$(4) \text{ 设 } a'_{ij}(t) \in C[t_0, t_1], \text{ 则 } \int_{t_0}^{t_1} A'(t) dt = A(t_1) - A(t_0) .$$

$$\text{左} = \left(\int_{t_0}^{t_1} a'_{ij}(t) dt \right)_{m \times n} = (a_{ij}(t_1) - a_{ij}(t_0))_{m \times n} = (a_{ij}(t_1))_{m \times n} - (a_{ij}(t_0))_{m \times n} = \text{右}$$

三、函数对矩阵的导数

设 $X = (\xi_{ij})_{m \times n}$, 多元函数 $f(X) = f(\xi_{11}, \xi_{12}, \dots, \xi_{1n}, \dots, \xi_{mn})$, 定义

$$\frac{df}{dX} \triangleq \left(\frac{\partial f}{\partial \xi_{ij}} \right)_{m \times n} = \begin{bmatrix} \frac{\partial f}{\partial \xi_{11}} & \dots & \frac{\partial f}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \xi_{m1}} & \dots & \frac{\partial f}{\partial \xi_{mn}} \end{bmatrix}$$

例 1 $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$: $f(x) = f(\xi_1, \dots, \xi_n)$, $\frac{df}{dx} = \begin{bmatrix} \frac{\partial f}{\partial \xi_1} \\ \vdots \\ \frac{\partial f}{\partial \xi_n} \end{bmatrix}$

例 2 $A = (a_{ij})_{n \times n}$, $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$: $f(x) = x^T A x$, 求 $\frac{df}{dx}$.

解 $f(x) = \xi_1 \cdot \sum_{j=1}^n a_{1j} \xi_j + \dots + \xi_k \cdot \sum_{j=1}^n a_{kj} \xi_j + \dots + \xi_n \cdot \sum_{j=1}^n a_{nj} \xi_j$

$$\frac{\partial f}{\partial \xi_k} = \xi_1 \cdot a_{1k} + \dots + \xi_{k-1} \cdot a_{k-1,k} + \left(\sum_{j=1}^n a_{kj} \xi_j + \xi_k \cdot a_{kk} \right) +$$

$$\xi_{k+1} \cdot a_{k+1,k} + \dots + \xi_n \cdot a_{nk} = \sum_{j=1}^n a_{kj} \xi_j + \sum_{i=1}^n a_{ik} \xi_i$$

$$\frac{df}{dx} = \begin{bmatrix} \sum_{j=1}^n a_{1j} \xi_j + \sum_{i=1}^n a_{i1} \xi_i \\ \vdots \\ \sum_{j=1}^n a_{nj} \xi_j + \sum_{i=1}^n a_{in} \xi_i \end{bmatrix} = (A + A^T)x \quad \text{“ } A^T = A \Rightarrow \frac{df}{dx} = 2Ax \text{ ”}$$

例 3 $X = (\xi_{ij})_{n \times n}$: $f(X) = [\text{tr}(X)]^2$, 求 $\frac{df}{dX} \Big|_{X=I_n}$.

解 $f(X) = (\xi_{11} + \xi_{22} + \cdots + \xi_{nn})^2$

$$\frac{df}{dX} = 2(\xi_{11} + \xi_{22} + \cdots + \xi_{nn})I_n, \quad \frac{df}{dX} \Big|_{X=I_n} = 2nI_n.$$

例 4 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, 若 $x \in \mathbb{R}^n$ 使得 $\|Ax - b\|_2 = \min$, 则 $A^T Ax = A^T b$.

证 $f(x) \stackrel{\Delta}{=} \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b) = x^T A^T Ax - 2b^T Ax + b^T b$

$$g(x) \stackrel{\Delta}{=} b^T Ax = b_1 \cdot \sum_{j=1}^n a_{1j} \xi_j + \cdots + b_m \cdot \sum_{j=1}^n a_{mj} \xi_j$$

$$\frac{dg}{dx} = \begin{bmatrix} \frac{\partial g}{\partial \xi_1} \\ \vdots \\ \frac{\partial g}{\partial \xi_n} \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + \cdots + b_m a_{m1} \\ \vdots \\ b_1 a_{1n} + \cdots + b_m a_{mn} \end{bmatrix} = A^T b$$

$$\frac{df}{dx} = 2A^T Ax - 2A^T b = 0 \Rightarrow A^T Ax = A^T b$$

[注] $r(A^T A) = r(A) \Rightarrow r(A^T A \mid A^T b) = r(A^T A) \Rightarrow A^T Ax = A^T b$ 有解

四、函数矩阵对矩阵的导数

设 $X = (\xi_{ij})_{m \times n}$, 多元函数 $f_{kl}(X) \stackrel{\Delta}{=} f_{kl}(\xi_{11}, \xi_{12}, \cdots, \xi_{1n}, \cdots, \xi_{mn})$, 定义

$$F = \begin{bmatrix} f_{11} & \cdots & f_{1s} \\ \vdots & & \vdots \\ f_{r1} & \cdots & f_{rs} \end{bmatrix}, \quad \frac{\partial F}{\partial \xi_{ij}} \stackrel{\Delta}{=} \begin{bmatrix} \frac{\partial f_{11}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial \xi_{ij}} \\ \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial \xi_{ij}} \end{bmatrix}, \quad \frac{dF}{dX} \stackrel{\Delta}{=} \begin{bmatrix} \frac{\partial F}{\partial \xi_{11}} & \cdots & \frac{\partial F}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial \xi_{m1}} & \cdots & \frac{\partial F}{\partial \xi_{mn}} \end{bmatrix}_{m \times n \text{ 块}}$$

$$\text{例 5} \quad x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \quad F = (f_1(x), \cdots, f_l(x)): \quad \frac{dF}{dx} = \begin{bmatrix} \frac{\partial f_1}{\partial \xi_1} & \cdots & \frac{\partial f_1}{\partial \xi_n} \\ \vdots & & \vdots \\ \frac{\partial f_l}{\partial \xi_1} & \cdots & \frac{\partial f_l}{\partial \xi_n} \end{bmatrix}$$

例 6 $A = (a_{ij})_{n \times n}$

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \quad Ax = \begin{bmatrix} \sum_{j=1}^n a_{1j} \xi_j \\ \vdots \\ \sum_{j=1}^n a_{nj} \xi_j \end{bmatrix}, \quad \frac{d(Ax)}{dx^T} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = A$$

§ 3.5 应用 (一阶线性常系数微分方程组的矩阵函数解法)

$$\begin{cases} \xi_1'(t) = a_{11}\xi_1(t) + a_{12}\xi_2(t) + \cdots + a_{1n}\xi_n(t) + b_1(t) \\ \xi_2'(t) = a_{21}\xi_1(t) + a_{22}\xi_2(t) + \cdots + a_{2n}\xi_n(t) + b_2(t) \\ \cdots \cdots \\ \xi_n'(t) = a_{n1}\xi_1(t) + a_{n2}\xi_2(t) + \cdots + a_{nn}\xi_n(t) + b_n(t) \end{cases}$$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad x(t) = \begin{bmatrix} \xi_1(t) \\ \vdots \\ \xi_n(t) \end{bmatrix}, \quad b(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

齐次微分方程组(1): $x'(t) = A \cdot x(t)$

非齐次微分方程组(2): $x'(t) = A \cdot x(t) + b(t) \quad (b(t) \neq 0)$

一、齐次方程组 $x'(t) = A \cdot x(t)$ 的解法

Th10 齐次方程组 $x'(t) = A \cdot x(t)$ 满足 $x(t_0) = x_0$ 的解存在且唯一.

证 存在性. 设 $x(t) = e^{(t-t_0)A} x_0$, 则

$$x'(t) = A e^{(t-t_0)A} x_0 = A \cdot x(t), \quad x(t_0) = e^0 x_0 = x_0$$

唯一性. 设 $x(t)$ 满足 $x'(t) = A \cdot x(t)$, $x(t_0) = x_0$, 则有

$$x'(t) - A x(t) = 0 \Rightarrow e^{-tA} x'(t) + e^{-tA} (-A) x(t) = 0$$

$$[e^{-tA} x(t)]' = 0 \Rightarrow e^{-tA} x(t) = c \Rightarrow x(t) = e^{tA} c \quad (\text{通解})$$

因为 $x(t_0) = x_0$, 所以 $x_0 = e^{t_0 A} c \Rightarrow c = e^{-t_0 A} x_0$

$$\text{故 } x(t) = e^{tA} e^{-t_0 A} x_0 = e^{(t-t_0)A} x_0 \quad (\text{特解})$$

[注] 齐次方程组 $x'(t) = A \cdot x(t)$ 满足 $x(0) = x_0$ 的特解为 $x(t) = e^{tA} x_0$.

例 1 设 $A = \begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 2 \end{bmatrix}$, 求 $x'(t) = A \cdot x(t)$ 的通解.

解 $e^{tA} = \begin{bmatrix} e^t & te^t & 0 \\ & e^t & 0 \\ & & e^{2t} \end{bmatrix}, \quad x(t) = e^{tA} \cdot c = \begin{bmatrix} c_1 e^t + c_2 t e^t \\ c_2 e^t \\ c_3 e^{2t} \end{bmatrix}$

例 2 矩阵函数 e^{tA} 的列向量 $x_1(t), \dots, x_n(t)$ 构成齐次方程组 $x'(t) = A \cdot x(t)$ 的基础解系.

证 e^{tA} 可逆 $\Rightarrow x_1(t), \dots, x_n(t)$ 线性无关

取 $c = e_j \Rightarrow x_j(t) = e^{tA} c$ 是 $x'(t) = A \cdot x(t)$ 的一个解

通解 $x(t) = e^{tA} \cdot c = c_1 \cdot x_1(t) + \cdots + c_n \cdot x_n(t)$

二、非齐次方程组 $x'(t) = A \cdot x(t) + b(t)$ 的解法

方程(1): $x'(t) = A \cdot x(t)$

方程(2): $x'(t) = A \cdot x(t) + b(t)$

$$\left. \begin{array}{l} \tilde{x}(t) \text{ 是 (2) 的特解} \\ x(t) \text{ 是 (2) 的通解} \end{array} \right\} \Rightarrow \begin{cases} \tilde{x}'(t) = A \cdot \tilde{x}(t) + b(t) \\ x'(t) = A \cdot x(t) + b(t) \end{cases}$$

$$\Rightarrow [x(t) - \tilde{x}(t)]' = A[x(t) - \tilde{x}(t)] \Rightarrow x(t) - \tilde{x}(t) \text{ 是 (1) 的解}$$

例2

$$\Rightarrow x(t) - \tilde{x}(t) = c_1 \cdot x_1(t) + \cdots + c_n \cdot x_n(t) \Rightarrow x(t) = e^{tA} \cdot c + \tilde{x}(t)$$

采用常向量变异法求 $\tilde{x}(t)$. 设 $\tilde{x}(t) = e^{tA} c(t)$ 满足(2), 则有

$$Ae^{tA} \cdot c(t) + e^{tA} \cdot c'(t) = A \cdot e^{tA} c(t) + b(t)$$

$$c'(t) = e^{-tA} b(t) \Rightarrow c(t) = \int_{t_0}^t e^{-\tau A} b(\tau) d\tau \quad (\text{原函数之一})$$

$$\text{故 (2) 的通解为 } x(t) = e^{tA} \cdot \left[c + \int_{t_0}^t e^{-\tau A} b(\tau) d\tau \right]$$

$$\text{特解为 } x(t) \Big|_{x(t_0)=x_0} = e^{tA} \cdot \left[e^{-t_0 A} x_0 + \int_{t_0}^t e^{-\tau A} b(\tau) d\tau \right]$$

$$[\text{注}] \text{ 当 } t_0 = 0 \text{ 时, 特解为 } x(t) \Big|_{x(0)=x_0} = e^{tA} \cdot \left[x_0 + \int_0^t e^{-\tau A} b(\tau) d\tau \right]$$

$$\text{例 3 设 } A = \begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 2 \end{bmatrix}, b(t) = \begin{bmatrix} 1 \\ 0 \\ e^{2t} \end{bmatrix}, x(0) = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

求 $x'(t) = A \cdot x(t) + b(t)$ 满足初始条件 $x(0)$ 的特解.

$$\text{解 例 1 求得 } e^{tA} = \begin{bmatrix} e^t & te^t & 0 \\ & e^t & 0 \\ & & e^{2t} \end{bmatrix}, \text{ 计算}$$

$$e^{-\tau A} \cdot b(\tau) = \begin{bmatrix} e^{-\tau} \\ 0 \\ 1 \end{bmatrix}, \int_0^t e^{-\tau A} b(\tau) d\tau = \begin{bmatrix} 1 - e^{-t} \\ 0 \\ t \end{bmatrix}$$

$$x(t) = e^{tA} \cdot \left[\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 - e^{-t} \\ 0 \\ t \end{pmatrix} \right] = \begin{bmatrix} te^t - 1 \\ e^t \\ te^{2t} \end{bmatrix}$$

$$\text{例 4 设 } A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix}, b(t) = \begin{bmatrix} e^{9t} \\ e^{9t} \\ 0 \end{bmatrix}, x(0) = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

求 $x'(t) = A \cdot x(t) + b(t)$ 满足初始条件 $x(0)$ 的特解.

解 用 3 种方法求 e^{tA} .

(1) 待定法: $|\lambda I - A| = \lambda(\lambda - 9)^2$, 因为 $A(A - 9I) = O$, 所以 A 的最小多项式为 $m(\lambda) = \lambda(\lambda - 9)$. 设 $f(\lambda) = e^{t\lambda} = m(\lambda)g(\lambda) + (a + b\lambda)$, 则

$$\begin{cases} f(0) = 1 = a \\ f(9) = e^{9t} = a + 9b \end{cases} \quad \begin{cases} a = 1 \\ b = (e^{9t} - 1)/9 \end{cases}$$

$$\begin{aligned} \text{于是 } e^{tA} &= aI + bA = \frac{1}{9}(9I - A) + \frac{e^{9t}}{9}A \\ &= \frac{1}{9} \begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix} + \frac{e^{9t}}{9} \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix} \\ &= \frac{1}{9} \begin{bmatrix} 4+5y & -4+4y & -2+2y \\ -4+4y & 4+5y & 2-2y \\ -2+2y & 2-2y & 1+8y \end{bmatrix} \quad (y = e^{9t}) \end{aligned}$$

(2) 对角形法: A 是实对称矩阵, 存在可逆矩阵 P 使得 $P^{-1}AP = \Lambda$, 其中

$$A = \begin{bmatrix} 0 & & \\ & 9 & \\ & & 9 \end{bmatrix}, \quad P = \begin{bmatrix} -2 & 1 & 1 \\ 2 & 1 & -1 \\ 1 & 0 & 4 \end{bmatrix}, \quad P^{-1} = \frac{1}{18} \begin{bmatrix} -4 & 4 & 2 \\ 9 & 9 & 0 \\ 1 & -1 & 4 \end{bmatrix}$$

$$\begin{aligned} \text{故 } e^{tA} &= P \begin{bmatrix} 1 & & \\ & e^{9t} & \\ & & e^{9t} \end{bmatrix} P^{-1} = P \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix} P^{-1} + e^{9t} P \begin{bmatrix} 0 & & \\ & 1 & \\ & & 1 \end{bmatrix} P^{-1} \\ &= \frac{1}{9} \begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix} + \frac{e^{9t}}{9} \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix} \end{aligned}$$

(3) 级数求和法: $|\lambda I - A| = \lambda(\lambda - 9)^2$, 因为 $A(A - 9I) = O$, 所以 A 的最小多项式为 $m(\lambda) = \lambda(\lambda - 9)$. 由 $m(A) = O$ 知 $A^2 = 9A$, 于是可得

$$\begin{aligned} e^{tA} &= I + \frac{(tA)}{1!} + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \cdots = I + \left(\frac{t}{1!} + \frac{9t^2}{2!} + \frac{9^2 t^3}{3!} + \cdots \right) A \\ &= I + \frac{1}{9} \left(\frac{9t}{1!} + \frac{(9t)^2}{2!} + \frac{(9t)^3}{3!} + \cdots \right) A \\ &= I + \frac{1}{9} (e^{9t} - 1) A = \frac{1}{9} (9I - A) + \frac{e^{9t}}{9} A \\ &= \frac{1}{9} \begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix} + \frac{e^{9t}}{9} \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix} \\ e^{-\tau A} b(\tau) &= \left\{ \frac{1}{9} \begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix} + \frac{e^{-9\tau}}{9} \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix} \right\} \begin{bmatrix} e^{9\tau} \\ e^{9\tau} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ x(t) &= e^{tA} \left\{ x(0) + \int_0^t e^{-\tau A} b(\tau) d\tau \right\} \end{aligned}$$

$$= \left\{ \frac{1}{9} \begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix} + \frac{e^{9t}}{9} \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix} \right\} \cdot \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} t \\ t \\ 0 \end{bmatrix} \right\} = e^{9t} \begin{bmatrix} 1+t \\ t \\ 2 \end{bmatrix}$$

例 5(2002B) 设 $A = \begin{bmatrix} -1 & 2 & 0 \\ -2 & 3 & 1 \\ 2 & -2 & -2 \end{bmatrix}$, $b(t) = \begin{bmatrix} e^{-t} \\ 0 \\ 2e^{-t} \end{bmatrix}$, $x(0) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

1. 求 e^{tA} ;

2. 用矩阵函数方法求 $\frac{d}{dt}x(t) = Ax(t) + b(t)$ 满足初始条件 $x(0)$ 的解.

解 $\varphi(\lambda) = |\lambda I - A| = \lambda(\lambda - 1)(\lambda + 1)$

(1) **对角化法:** 求 P 使得 $P^{-1}AP = \Lambda$

$$A = \begin{bmatrix} 0 & & \\ & 1 & \\ & & -1 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 2 & -2 & -1 \\ -2 & 3 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

$$\begin{aligned} e^{tA} &= P \begin{bmatrix} 1 & & \\ & e^t & \\ & & e^{-t} \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix} P^{-1} + e^t P \begin{bmatrix} 0 & & \\ & 1 & \\ & & 0 \end{bmatrix} P^{-1} + e^{-t} P \begin{bmatrix} 0 & & \\ & 0 & \\ & & 1 \end{bmatrix} P^{-1} \\ &= \begin{bmatrix} 4 & -4 & -2 \\ 2 & -2 & -1 \\ 2 & -2 & -1 \end{bmatrix} + e^t \begin{bmatrix} -2 & 3 & 1 \\ -2 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix} + e^{-t} \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 0 \\ -2 & 2 & 2 \end{bmatrix} \end{aligned}$$

(2) **待定法:** 设 $f(\lambda) = e^{t\lambda} = \varphi(\lambda)g(\lambda) + (a + b\lambda + c\lambda^2)$, 则

$$\begin{cases} f(0) = 1 = a \\ f(1) = e^t = a + b + c \\ f(-1) = e^{-t} = a - b + c \end{cases} \Rightarrow \begin{cases} a = 1 \\ b = (e^t - e^{-t})/2 \\ c = (e^t + e^{-t} - 2)/2 \end{cases}$$

$$\text{于是 } e^{tA} = aI + bA + cA^2 = (I - A^2) + \frac{e^t}{2}(A^2 + A) + \frac{e^{-t}}{2}(A^2 - A)$$

$$= \begin{bmatrix} 4 & -4 & -2 \\ 2 & -2 & -1 \\ 2 & -2 & -1 \end{bmatrix} + e^t \begin{bmatrix} -2 & 3 & 1 \\ -2 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix} + e^{-t} \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 0 \\ -2 & 2 & 2 \end{bmatrix}$$

$$e^{-\tau A} b(\tau) = \left\{ (I - A^2) + \frac{e^{-\tau}}{2}(A^2 + A) + \frac{e^{\tau}}{2}(A^2 - A) \right\} \cdot e^{-\tau} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

$$x(t) = e^{tA} \left\{ x(0) + \int_0^t e^{-\tau A} b(\tau) d\tau \right\}$$

$$\begin{aligned}
 &= \left\{ (I - A^2) + \frac{e^t}{2}(A^2 + A) + \frac{e^{-t}}{2}(A^2 - A) \right\} \cdot \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\} \\
 &= e^t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t e^{-t} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} e^t + t e^{-t} \\ e^t \\ 2 t e^{-t} \end{bmatrix}
 \end{aligned}$$

例 6 (2004A)

设 $\sin At = \begin{bmatrix} \sin 2t + \sin t & 2 \sin 2t + \sin t & 3 \sin 2t + \sin t \\ \sin 2t + 2 \sin t & 2 \sin 2t + 2 \sin t & 3 \sin 2t + 2 \sin t \\ \sin 2t + 3 \sin t & 2 \sin 2t + 3 \sin t & 3 \sin 2t + 3 \sin t \end{bmatrix}$, 求 A .

解 $\frac{d \sin At}{dt} \Big|_{t=0} = (A \cos At) \Big|_{t=0} \Rightarrow A = \begin{bmatrix} 3 & 5 & 7 \\ 4 & 6 & 8 \\ 5 & 7 & 9 \end{bmatrix}$

例 7 (2004B) 设 A 为可逆矩阵, 求 $\int_0^1 \sin At \, dt$.

解 原式 $= -A^{-1} \int_0^1 d(\cos At) = A^{-1}(I - \cos A)$

第四章 矩阵分解

§ 4.1 三角分解

目的: 将 $A_{n \times n}$ 分解为下三角矩阵与上三角矩阵的乘积.

一、分解原理: 以 $n=4$ 为例

$$\textcircled{1} \quad \Delta_1(A) = a_{11}: \quad a_{11} \neq 0 \Rightarrow c_{i1} = \frac{a_{i1}}{a_{11}} \quad (i=2,3,4)$$

$$L_1 = \begin{bmatrix} 1 & & & \\ c_{21} & 1 & & \\ c_{31} & 0 & 1 & \\ c_{41} & 0 & 0 & 1 \end{bmatrix}, \quad L_1^{-1} = \begin{bmatrix} 1 & & & \\ -c_{21} & 1 & & \\ -c_{31} & 0 & 1 & \\ -c_{41} & 0 & 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

$$L_1^{-1}A = \left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{array} \right] \triangleq A^{(1)}$$

$$\textcircled{2} \quad \Delta_2(A) = \Delta_2(A^{(1)}) = a_{11}a_{22}^{(1)}: \quad a_{22}^{(1)} \neq 0 \Rightarrow c_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} \quad (i=3,4)$$

$$L_2 = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & c_{32} & 1 & \\ 0 & c_{42} & 0 & 1 \end{bmatrix}, \quad L_2^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & -c_{32} & 1 & \\ 0 & -c_{42} & 0 & 1 \end{bmatrix}$$

$$L_2^{-1}A^{(1)} = \left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ \hline & & a_{33}^{(2)} & a_{34}^{(2)} \\ & & a_{43}^{(2)} & a_{44}^{(2)} \end{array} \right] \triangleq A^{(2)}$$

$$\textcircled{3} \quad \Delta_3(A) = \Delta_3(A^{(2)}) = a_{11}a_{22}^{(1)}a_{33}^{(2)}: \quad a_{33}^{(2)} \neq 0 \Rightarrow c_{43} = \frac{a_{43}^{(2)}}{a_{33}^{(2)}}$$

$$L_3 = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & c_{43} & 1 \end{bmatrix}, \quad L_3^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & -c_{43} & 1 \end{bmatrix}$$

$$L_3^{-1}A^{(2)} = \left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & & a_{33}^{(2)} & a_{34}^{(2)} \\ & & & a_{44}^{(3)} \end{array} \right] \triangleq A^{(3)}$$

$$\text{即 } L_3^{-1}L_2^{-1}L_1^{-1}A = A^{(3)} \Rightarrow A = L_1L_2L_3A^{(3)}$$

$$\text{令 } L \stackrel{\Delta}{=} L_1 L_2 L_3 = \begin{bmatrix} 1 & & & \\ c_{21} & 1 & & \\ c_{31} & c_{32} & 1 & \\ c_{41} & c_{42} & c_{43} & 1 \end{bmatrix}, \text{ 则 } A = LA^{(3)}.$$

$$\text{分解 } A^{(3)} = \begin{bmatrix} a_{11} & & & \\ & a_{22}^{(1)} & & \\ & & a_{33}^{(2)} & \\ & & & a_{44}^{(3)} \end{bmatrix} \begin{bmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{bmatrix} = DU, \text{ 则 } A = LDU.$$

Th1 $A_{n \times n}, \Delta_k(A) \neq 0 (k=1, 2, \dots, n-1) \Rightarrow A = LDU$ 存在且唯一.

二、紧凑格式算法: $A = LDU \stackrel{\Delta}{=} \tilde{L}U$ (Crout 分解)

$$\tilde{L} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, \quad U = \begin{bmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$(i, 1)\text{元: } a_{i1} = l_{i1} \cdot 1 \Rightarrow l_{i1} = a_{i1} \quad (i=1, \dots, n)$$

$$(1, j)\text{元: } a_{1j} = l_{11} \cdot u_{1j} \Rightarrow u_{1j} = \frac{a_{1j}}{l_{11}} \quad (j=2, \dots, n)$$

$$(i, k)\text{元: } a_{ik} = l_{i1} \cdot u_{1k} + \cdots + l_{i, k-1} \cdot u_{k-1, k} + l_{ik} \cdot 1 \quad (i \geq k) \\ \Rightarrow l_{ik} = a_{ik} - (l_{i1} \cdot u_{1k} + \cdots + l_{i, k-1} \cdot u_{k-1, k})$$

$$(k, j)\text{元: } a_{kj} = l_{k1} \cdot u_{1j} + \cdots + l_{k, k-1} \cdot u_{k-1, j} + l_{kk} \cdot u_{kj} \quad (j > k)$$

$$\Rightarrow u_{kj} = \frac{1}{l_{kk}} [a_{kj} - (l_{k1} \cdot u_{1j} + \cdots + l_{k, k-1} \cdot u_{k-1, j})]$$

计算框图:

l_{11}	u_{12}	u_{13}	u_{14}	\cdots	第1框
l_{21}	l_{22}	u_{23}	u_{24}	\cdots	第2框
l_{31}	l_{32}	l_{33}	u_{34}	\cdots	第3框
l_{41}	l_{42}	l_{43}	l_{44}	\cdots	第4框
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots

例 1 $A = \begin{bmatrix} 5 & 2 & -4 & 0 \\ 2 & 1 & -2 & 1 \\ -4 & -2 & 5 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$, 计算框图:

5	2/5	-4/5	0	
2	1/5	-2	5	
-4	-2/5	1	2	
0	1	2	-7	

$$\tilde{L} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 1/5 & 0 & 0 \\ -4 & -2/5 & 1 & 0 \\ 0 & 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 2/5 & 1 & & \\ -4/5 & -2 & 1 & \\ 0 & 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 & & & \\ & 1/5 & & \\ & & 1 & \\ & & & -7 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 2/5 & -4/5 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad A = \tilde{L}U = LDU$$

§ 4.2 QR 分解

目的: 将 $A_{n \times n}$ 分解为正交矩阵与上三角矩阵之积.

约定: 本节涉及的矩阵为实矩阵, 向量为实向量, 数为实数.

一、Givens 矩阵

$$T_{ij}(c, s) = \begin{bmatrix} I & & & \\ & c & & s \\ & & I & \\ & -s & & c \\ & & & I \end{bmatrix} \quad \begin{matrix} (i) \\ (j) \end{matrix} \quad (c^2 + s^2 = 1)$$

$$1. \quad T_{ij}^T T_{ij} = I, \quad [T_{ij}(c, s)]^{-1} = [T_{ij}(c, s)]^T = T_{ij}(c, -s), \quad \det T_{ij} = 1.$$

$$2. \quad x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \quad T_{ij}x = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \Rightarrow \begin{cases} \eta_i = c\xi_i + s\xi_j \\ \eta_j = -s\xi_i + c\xi_j \\ \eta_k = \xi_k \quad (k \neq i, j) \end{cases}$$

$$\text{若 } \xi_i^2 + \xi_j^2 \neq 0, \text{ 取 } c = \frac{\xi_i}{(\xi_i^2 + \xi_j^2)^{\frac{1}{2}}}, \quad s = \frac{\xi_j}{(\xi_i^2 + \xi_j^2)^{\frac{1}{2}}}$$

$$\text{则 } \eta_i = (\xi_i^2 + \xi_j^2)^{\frac{1}{2}} > 0, \quad \eta_j = 0.$$

Th3 $x \neq 0 \Rightarrow \exists$ 有限个 G-矩阵之积 T , 使得 $Tx = |x|e_1$.

证 ① $\xi_1 \neq 0$ (以 $n=4$ 为例)

$$\text{构造 } T_{12}(c, s), \quad c = \frac{\xi_1}{(\xi_1^2 + \xi_2^2)^{\frac{1}{2}}}, \quad s = \frac{\xi_2}{(\xi_1^2 + \xi_2^2)^{\frac{1}{2}}}: \quad T_{12}x = \begin{bmatrix} (\xi_1^2 + \xi_2^2)^{\frac{1}{2}} \\ 0 \\ \xi_3 \\ \xi_4 \end{bmatrix} \stackrel{\Delta}{=} x^{(2)}$$

$$\text{构造 } T_{13}(c, s), \quad c = \frac{(\xi_1^2 + \xi_2^2)^{\frac{1}{2}}}{(\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}}}, \quad s = \frac{\xi_3}{(\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}}}: \quad$$

$$T_{13}x^{(2)} = \begin{bmatrix} (\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}} \\ 0 \\ 0 \\ \xi_4 \end{bmatrix} \stackrel{\Delta}{=} x^{(3)}$$

构造 $T_{14}(c, s)$, $c = \frac{(\xi_1^2 + \xi_2^2 + \xi_3^2)^{\frac{1}{2}}}{(\xi_1^2 + \dots + \xi_4^2)^{\frac{1}{2}}}$, $s = \frac{\xi_4}{(\xi_1^2 + \dots + \xi_4^2)^{\frac{1}{2}}}$: $T_{14}x^{(3)} = |x|e_1$

于是可得 $T_{14}T_{13}T_{12}x = |x|e_1$

② $\xi_1 = \dots = \xi_{k-1} = 0, \xi_k \neq 0$ ($1 < k \leq n$): $|x| = (\xi_k^2 + \dots + \xi_n^2)^{\frac{1}{2}}$
由 T_{1k} 开始即可.

推论 \mathbf{R}^n 中, $\forall x \neq 0, \forall$ 单位向量 $z \Rightarrow \exists$ 有限个 G-矩阵之积 T , st. $Tx = |x|z$.

证 $T^{(1)}x = |x|e_1, T^{(2)}z = |z|e_1 \Rightarrow [T^{(2)}]^{-1}T^{(1)}x = |x|z$
 $T = [T^{(2)}]^{-1}T^{(1)} = [T_{1n}^{(2)} \dots T_{12}^{(2)}]^{-1}T^{(1)} = [(T_{12}^{(2)})^T \dots (T_{1n}^{(2)})^T][T_{1n}^{(1)} \dots T_{12}^{(1)}]$

例 1 $x = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$, 求 G-矩阵之积 T 使得 $Tx = |x|e_1$.

解 $T_{12}(c, s)$ 中, $c = \frac{3}{5}, s = \frac{4}{5}$. $T_{12}x = \begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix}$

$T_{13}(c, s)$ 中, $c = \frac{1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$. $T_{13}(T_{12}x) = \begin{bmatrix} 5\sqrt{2} \\ 0 \\ 0 \end{bmatrix} = |x|e_1$

$$T = T_{13}T_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \frac{1}{5} \begin{bmatrix} 3 & 4 & 0 \\ -4 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \frac{1}{5\sqrt{2}} \begin{bmatrix} 3 & 4 & 5 \\ -4\sqrt{2} & 3\sqrt{2} & 0 \\ -3 & -4 & 5 \end{bmatrix}$$

$$Tx = 5\sqrt{2}e_1$$

二、Householder 矩阵

$H_u = I_n - 2uu^T$ ($u \in \mathbf{R}^n$ 是单位列向量)

(1) $H^T = H$ 对称矩阵 (2) $H^T H = I$ 正交矩阵

(3) $H^2 = I$ 对合矩阵 (4) $H^{-1} = H$ 自逆矩阵 (5) **$\det H = -1$**

验证(5):

$$\begin{bmatrix} I & 0 \\ -u^T & 1 \end{bmatrix} \begin{bmatrix} I & 2u \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I - 2uu^T & 0 \\ u^T & 1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ -u^T & 1 \end{bmatrix} \begin{bmatrix} I & 2u \\ u^T & 1 \end{bmatrix} = \begin{bmatrix} I & 2u \\ 0^T & -1 \end{bmatrix}$$

$$\begin{vmatrix} I - 2uu^T & 0 \\ u^T & 1 \end{vmatrix} = \begin{vmatrix} I & 2u \\ 0^T & -1 \end{vmatrix} \Rightarrow \det(I - 2uu^T) = -1$$

Th4 \mathbf{R}^n 中 ($n > 1$), $\forall x \neq 0, \forall$ 单位向量 $z \Rightarrow \exists H_u$, st $H_u x = |x|z$.

证 ① $x = |x|z$: $n > 1$ 时, 可取单位向量 u 使得 $u \perp x$, 于是

$$H_u = I - 2uu^T : H_u x = Ix - 2uu^T x = x - |x|z$$

② $x \neq |x|z$: 取 $u = \frac{x - |x|z}{|x - |x|z|}$, 有

$$\begin{aligned} H_u x &= \left[I - 2 \frac{(x - |x|z)(x - |x|z)^T}{|x - |x|z|^2} \right] x = x - \frac{2(x - |x|z, x)}{|x - |x|z|^2} (x - |x|z) \\ &= x - 1 \cdot (x - |x|z) = |x|z \end{aligned}$$

例 2 $x = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, 求 H-矩阵 H 使得 $Hx = |x|e_1$.

解 $|x| = 3, x - |x|e_1 = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}, u = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$$H = I - \frac{2}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix} : Hx = 3e_1$$

三、G-矩阵与 H-矩阵的关系

Th5 G-矩阵 $T_{ij}(c, s) \Rightarrow \exists$ H-矩阵 H_u 与 H_v , st $T_{ij} = H_u H_v$.

证 $c^2 + s^2 = 1 \Rightarrow$ 取 $\theta = \arctg \frac{s}{c}$, 则 $\cos \theta = c, \sin \theta = s$.

$$T_{ij}(c, s) = \begin{bmatrix} I & & & \\ & \cos \theta & \sin \theta & \\ & & I & \\ & -\sin \theta & \cos \theta & \\ & & & I \end{bmatrix} \quad \begin{matrix} (i) \\ (j) \end{matrix}$$

$$v = \left(0 \quad \cdots \quad 0 \quad \sin \frac{\theta}{4} \quad 0 \quad \cdots \quad 0 \quad \cos \frac{\theta}{4} \quad 0 \quad \cdots \quad 0 \right)^T$$

$$H_v = \begin{bmatrix} I & & & \\ & 1 & & \\ & & I & \\ & & & 1 \\ & & & & I \end{bmatrix} - 2 \begin{bmatrix} O & & & \\ & \sin^2 \frac{\theta}{4} & & \sin \frac{\theta}{4} \cos \frac{\theta}{4} \\ & & O & \\ & \cos \frac{\theta}{4} \sin \frac{\theta}{4} & & \cos^2 \frac{\theta}{4} \\ & & & & O \end{bmatrix}$$

$$= \begin{bmatrix} I & & & \\ & \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} & \\ & -\sin \frac{\theta}{2} & -\cos \frac{\theta}{2} & \\ & & & I \end{bmatrix}$$

$$u = \left(0 \quad \cdots \quad 0 \quad \sin \frac{3\theta}{4} \quad 0 \quad \cdots \quad 0 \quad \cos \frac{3\theta}{4} \quad 0 \quad \cdots \quad 0 \right)^T$$

$$H_u = \begin{bmatrix} I & & & \\ & \cos \frac{3\theta}{2} & -\sin \frac{3\theta}{2} & \\ & -\sin \frac{3\theta}{2} & -\cos \frac{3\theta}{2} & \\ & & & I \end{bmatrix}, \quad T_{ij}(c, s) = H_u H_v$$

[注] H-矩阵不能由若干个 G-矩阵的乘积表示.

因为 $\det H = -1$, 而 $\det T_{ij} = 1$.

例 3 G-矩阵 $T_{12}(0,1) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ 中, $c=0, s=1 \Rightarrow \theta = \frac{\pi}{2}$

$$H_u = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}, H_v = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \Rightarrow H_u H_v = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

四、QR 分解

1. Schmidt 正交化方法

Th6 $A_{n \times n}$ 可逆 $\Rightarrow \exists$ 正交矩阵 Q , 可逆上三角矩阵 R , 使得 $A = QR$.

证 $A = (a_1, a_2, \dots, a_n)$ 可逆 $\Rightarrow a_1, a_2, \dots, a_n$ 线性无关, 正交化可得:

$$\begin{cases} b_1 = a_1 \\ b_2 = a_2 - k_{21}b_1 \\ \dots\dots \\ b_n = a_n - k_{n,n-1}b_{n-1} - \dots - k_{n1}b_1 \end{cases} \quad \begin{cases} a_1 = b_1 \\ a_2 = k_{21}b_1 + b_2 \\ \dots\dots \\ a_n = k_{n1}b_1 + \dots + k_{n,n-1}b_{n-1} + b_n \end{cases}$$

$$(a_1, a_2, \dots, a_n) = (b_1, b_2, \dots, b_n)K = (|b_1|q_1, |b_2|q_2, \dots, |b_n|q_n)K$$

$$= (q_1, q_2, \dots, q_n) \begin{bmatrix} |b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n| \end{bmatrix} \begin{bmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$\text{令 } Q = (q_1, q_2, \dots, q_n), R = \begin{bmatrix} |b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n| \end{bmatrix} \begin{bmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

则 $A = QR$, 其中 $q_i = \frac{b_i}{|b_i|} \quad (i=1,2,\cdots,n)$.

例 4 求 $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 的 QR 分解.

解 $b_1 = a_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $b_2 = a_2 - 1 \cdot b_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $b_3 = a_3 - \frac{1}{3}b_2 - \frac{7}{6}b_1 = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{6} & & \\ & \sqrt{3} & \\ & & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & \frac{7}{6} \\ & 1 & \frac{1}{3} \\ & & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \sqrt{6} & \frac{7}{\sqrt{6}} \\ & \sqrt{3} & \frac{1}{\sqrt{3}} \\ & & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Th7 $A_{m \times n}$ 列满秩 $\Rightarrow \exists Q_{m \times n}$ 满足 $Q^T Q = I$, 可逆上三角矩阵 $R_{n \times n}$, 使 $A = QR$.

(证明过程同 Th6)

2. G-变换方法

Th8 $A_{n \times n}$ 可逆 $\Rightarrow \exists$ 有限个 G-矩阵之积 T , 使得 TA 为可逆上三角矩阵.

(记 $TA = R$, 则有 $A = T^{-1}R = T^T R = QR$)

证 以 $n=4$ 为例

(1) $|A| \neq 0$: $\beta^{(0)} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \neq 0$, 存在有限个 G-矩阵之积 T_0 , 使得

$$T_0 \beta^{(0)} = \begin{bmatrix} |\beta^{(0)}| \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad a_{11}^{(1)} = |\beta^{(0)}| > 0, \quad T_0 A = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ 0 & & & \\ 0 & & A^{(1)} & \\ 0 & & & \end{bmatrix}$$

(2) $A^{(1)} = \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix}$, $|A^{(1)}| \neq 0$: $\beta^{(1)} = \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ a_{42}^{(1)} \end{bmatrix} \neq 0$

存在有限个 G-矩阵之积 T_1 , 使得

$$T_1 \beta^{(1)} = \begin{bmatrix} |\beta^{(1)}| \\ 0 \\ 0 \end{bmatrix}, \quad a_{22}^{(2)} = |\beta^{(1)}| > 0, \quad T_1 A^{(1)} = \begin{bmatrix} a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

$$(3) \quad A^{(2)} = \begin{bmatrix} a_{33}^{(2)} & a_{34}^{(2)} \\ a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}, \quad |A^{(2)}| \neq 0: \quad \beta^{(2)} = \begin{bmatrix} a_{33}^{(2)} \\ a_{43}^{(2)} \end{bmatrix} \neq 0$$

存在 G-矩阵 T_2 , 使得

$$T_2 \beta^{(2)} = \begin{bmatrix} |\beta^{(2)}| \\ 0 \end{bmatrix}, \quad a_{33}^{(3)} = |\beta^{(2)}| > 0, \quad T_2 A^{(2)} = \begin{bmatrix} a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & a_{44}^{(3)} \end{bmatrix}$$

令 $T = \begin{bmatrix} I_2 & \\ & T_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} \cdot T_0$, 则 T 为有限个 G-矩阵之积, 且有

$$TA = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ & & a_{33}^{(3)} & a_{34}^{(3)} \\ & & & a_{44}^{(3)} \end{bmatrix} \triangleq R$$

[注] $\det T = 1 \Rightarrow \det A = a_{11}^{(1)} a_{22}^{(2)} \cdots a_{n-1, n-1}^{(n-1)} a_{nn}^{(n-1)}$, 故 $a_{nn}^{(n-1)}$ 与 $\det A$ 同符号.

在 Th8 中, 当 $A_{n \times n}$ 不可逆时, 仍可得 $TA = R$, 但 R 是不可逆矩阵.

例 5 用 G-变换求 $A = \begin{bmatrix} 3 & 5 & 5 \\ 0 & 3 & 4 \\ 4 & 0 & 5 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}: T_{13}(c, s) \text{ 中 } c = \frac{3}{5}, s = \frac{4}{5}. \quad T_{13} \beta^{(0)} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$

$$T_0 \triangleq T_{13} = \begin{bmatrix} 3/5 & 0 & 4/5 \\ 0 & 1 & 0 \\ -4/5 & 0 & 3/5 \end{bmatrix}, \quad T_0 A = \begin{bmatrix} 5 & 3 & 7 \\ 0 & 3 & 4 \\ 0 & -4 & -1 \end{bmatrix}$$

$$(2) \quad A^{(1)} = \begin{bmatrix} 3 & 4 \\ -4 & -1 \end{bmatrix}, \beta^{(1)} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}, T_{12}(c, s) \text{ 中 } c = \frac{3}{5}, s = -\frac{4}{5}. \quad T_{12} \beta^{(1)} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$T_1 \triangleq T_{12} = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}, \quad T_1 A^{(1)} = \begin{bmatrix} 5 & 16/5 \\ 0 & 13/5 \end{bmatrix}$$

$$\text{令 } T = \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} T_0 = \frac{1}{5} \begin{bmatrix} 5 & & \\ & 3 & -4 \\ & 4 & 3 \end{bmatrix} \cdot \frac{1}{5} \begin{bmatrix} 3 & 0 & 4 \\ 0 & 5 & 0 \\ -4 & 0 & 3 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 15 & 0 & 20 \\ 16 & 15 & -12 \\ -12 & 20 & 9 \end{bmatrix}$$

$$\text{则 } Q = T^{-1} = T^T, \quad R = \begin{bmatrix} 5 & 3 & 7 \\ & 5 & 16/5 \\ & & 13/5 \end{bmatrix}, \quad A = QR$$

例 6 用 Givens 变换求 $A = \begin{bmatrix} 0 & 0 & 1 & 3 \\ 0 & 4 & 3 & 2 \\ 1 & 1 & 2 & 4 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ 的 QR 分解.

解(1) $\beta^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, 构造 $T_{13}(c, s)$, $c = 0, s = 1$, 则

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} \alpha_3 \\ \alpha_2 \\ -\alpha_1 \\ \alpha_4 \end{bmatrix}$$

$$T_{13}\beta^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, T_0 = T_{13} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, T_0 A = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 4 & 3 & 2 \\ 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

(2) $A^{(2)} = \begin{bmatrix} -1 & -3 \\ 1 & 2 \end{bmatrix}$, $\beta^{(2)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, 构造 $T_{12}(c, s)$, $c = \frac{-1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$, 则

$$T_{12}\beta^{(2)} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}, T_2 = T_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, T_2 A^{(2)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 5 \\ 0 & 1 \end{bmatrix}$$

$$\text{令 } T = \begin{bmatrix} I_2 & O \\ O & T_2 \end{bmatrix} T_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} & 0 & & \\ 0 & \sqrt{2} & & \\ & & -1 & 1 \\ & & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & \sqrt{2} & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix} \quad [\beta_1 \ \beta_2 \ \beta_3 \ \beta_4] \cdot \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = [-\beta_3 \ \beta_2 \ \beta_1 \ \beta_4]$$

$$\text{则 } Q = T^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & \sqrt{2} & 0 & 0 \\ \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, R = \begin{bmatrix} 1 & 1 & 2 & 4 \\ & 4 & 3 & 2 \\ & & \sqrt{2} & 5/\sqrt{2} \\ & & & 1/\sqrt{2} \end{bmatrix}, A = QR$$

3. H-变换方法:

Th10 $A_{n \times n}$ 可逆 $\Rightarrow \exists$ 有限个 H-矩阵之积 S , 使得 SA 为可逆上三角矩阵.

(记 $SA = R$, 则有 $A = S^{-1}R = S^T R = QR$)

证 以 $n = 4$ 为例

$$(1) |A| \neq 0: \beta^{(0)} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \neq 0 \Rightarrow \exists \text{ H-矩阵 } H_0, \text{ 使得 } H_0 \beta^{(0)} = \begin{bmatrix} |\beta^{(0)}| \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$a_{11}^{(1)} = |\beta^{(0)}| > 0$$

$$H_0 A = \left[\begin{array}{c|ccc} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ \hline 0 & & & \\ 0 & & A^{(1)} & \\ 0 & & & \end{array} \right], \quad A^{(1)} = \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix}$$

$$(2) \quad |A^{(1)}| \neq 0: \quad \beta^{(1)} = \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ a_{42}^{(1)} \end{bmatrix} \neq 0 \Rightarrow \exists \text{H-矩阵 } H_1, \text{ 使得 } H_1 \beta^{(1)} = \begin{bmatrix} |\beta^{(1)}| \\ 0 \\ 0 \end{bmatrix}.$$

$$a_{22}^{(2)} = |\beta^{(1)}| > 0, \quad H_1 A^{(1)} = \left[\begin{array}{c|cc} a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ \hline 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{array} \right], \quad A^{(2)} = \begin{bmatrix} a_{33}^{(2)} & a_{34}^{(2)} \\ a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

$$(3) \quad |A^{(2)}| \neq 0: \quad \beta^{(2)} = \begin{bmatrix} a_{33}^{(2)} \\ a_{43}^{(2)} \end{bmatrix} \neq 0 \Rightarrow \exists \text{H-矩阵 } H_2, \text{ 使得 } H_2 \beta^{(2)} = \begin{bmatrix} |\beta^{(2)}| \\ 0 \end{bmatrix}.$$

$$a_{33}^{(3)} = |\beta^{(2)}| > 0, \quad H_2 A^{(2)} = \begin{bmatrix} a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & a_{44}^{(3)} \end{bmatrix}$$

令 $S = \begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix} \cdot H_0$, 则 S 为有限个 H-矩阵之积, 且有

$$SA = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ & & a_{33}^{(3)} & a_{34}^{(3)} \\ & & & a_{44}^{(3)} \end{bmatrix} \triangleq R$$

[注 1] 设 $H_l = I_{n-l} - 2u_{n-l}u_{n-l}^T$ ($u_{n-l}^T u_{n-l} = 1$), 则

$$\begin{aligned} \begin{bmatrix} I_l & \\ & H_l \end{bmatrix} &= \begin{bmatrix} I_l & \\ & I_{n-l} \end{bmatrix} - 2 \begin{bmatrix} O & O \\ O & u_{n-l}u_{n-l}^T \end{bmatrix} = I_n - 2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_{n-l} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & u_{n-l}^T \end{bmatrix} \\ &= I_n - 2u_n u_n^T \quad (u_n^T u_n = u_{n-l}^T u_{n-l} = 1) \end{aligned}$$

故 $\begin{bmatrix} I_l & \\ & H_l \end{bmatrix}$ 是 H-矩阵.

[注 2] 在 Th10 中, 当 $A_{n \times n}$ 不可逆时, 仍可得 $SA = R$, 但 R 是不可逆矩阵.

[注 3] $\begin{bmatrix} H & \\ & I \end{bmatrix}$ 和 $\begin{bmatrix} I & \\ & H \end{bmatrix}$ 也是 Householder 矩阵.

例 7 用 H-变换求 $A = \begin{bmatrix} 3 & 14 & 9 \\ 6 & 43 & 3 \\ 6 & 22 & 15 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}$, $\beta^{(0)} - |\beta^{(0)}|e_1 = \begin{bmatrix} -6 \\ 6 \\ 6 \end{bmatrix} = 6 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $u = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$$H_0 = I - 2uu^T = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}, \quad H_0 A = \begin{bmatrix} 9 & 48 & 15 \\ 0 & 9 & -3 \\ 0 & -12 & 9 \end{bmatrix}$$

(2) $A^{(1)} = \begin{bmatrix} 9 & -3 \\ -12 & 9 \end{bmatrix}$, $\beta^{(1)} = \begin{bmatrix} 9 \\ -12 \end{bmatrix}$

$$\beta^{(1)} - |\beta^{(1)}|e_1 = \begin{bmatrix} -6 \\ -12 \end{bmatrix} = (-6) \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad u = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$H_1 = I - 2uu^T = \frac{1}{5} \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}, \quad H_1 A^{(1)} = \begin{bmatrix} 15 & -9 \\ 0 & -3 \end{bmatrix}$$

令 $S = \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix} H_0$

则 $Q = S^{-1} = S^T = H_0 \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 5 & -2 & -14 \\ 10 & 11 & 2 \\ 10 & -10 & 5 \end{bmatrix}$

$$R = \begin{bmatrix} 9 & 48 & 15 \\ & 15 & -9 \\ & & -3 \end{bmatrix}, \quad A = QR$$

例 2002(B) 用 Householder 变换求 $A = \begin{bmatrix} 1 & 2 & 5 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 2 & -1 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\beta^{(0)} - |\beta^{(0)}|e_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $u = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$H_0 = I - 2uu^T = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \quad H_0 A = \begin{bmatrix} 2 & 4 & 4 & 4 \\ 0 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \end{bmatrix}$$

$$(2) \quad A^{(2)} = \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}, \quad \beta^{(2)} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \quad \beta^{(2)} - |\beta^{(2)}| e_1 = 4 \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad u = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$H_2 = I - 2uu^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad H_2 A^{(2)} = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\text{令 } S = \begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix} H_0, \quad Q = S^T = H_0 \begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\text{而 } R = \left[\begin{array}{cc|cc} 2 & 4 & 4 & 4 \\ & 0 & 2 & -4 \\ \hline & & 4 & 0 \\ & & & 0 \end{array} \right], \quad \text{则 } A = QR.$$

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{bmatrix}$$

[注] 因为 $H_0 A$ 的 3 行 2 列和 4 行 2 列元素都是 0, 所以不必进行第二步.

例 (2002B) 设 T 是 n 阶 Givens 矩阵, H_u 是 n 阶 Householder 矩阵, 则 $TH_u T^{-1}$ 也是 Householder 矩阵. (☒)

分析: $TH_u T^{-1} = T(I - 2uu^T)T^{-1} = I - 2(Tu)(Tu)^T$ (注意 $|Tu| = 1$)

五、化方阵与 Hessenberg 矩阵相似

$$\text{上 Hessenberg 矩阵: } F_{\perp} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & \ddots & \ddots & \ddots & \vdots \\ & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n,n-1} & a_{nn} \end{bmatrix}$$

Th11 设 $A \in \mathbb{R}^{n \times n}$, 则存在有限个 G-矩阵之积 Q , 使得 $Q A Q^T = F_{\perp}$.

证 以 $n=4$ 为例

$$(1) \text{ 对 } A: \text{ 若 } \beta^{(0)} = \begin{bmatrix} a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \neq 0, \text{ 则存在有限个 G-矩阵之积 } T_0, \text{ 使得}$$

$$T_0 \beta^{(0)} = |\beta^{(0)}| e_1 = a_{21}^{(1)} e_1 \quad (e_1 \in \mathbb{R}^3)$$

$$\begin{bmatrix} 1 & \\ & T_0 \end{bmatrix} A \begin{bmatrix} 1 \\ & T_0 \end{bmatrix}^T = \left[\begin{array}{c|ccc} a_{11} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ a_{21}^{(1)} & \hline 0 & & A^{(1)} & \\ 0 & & & \end{array} \right]$$

若 $\beta^{(0)} = 0$, 转入(2).

(2) 对 $A^{(1)}$: 若 $\beta^{(1)} = \begin{bmatrix} a_{32}^{(1)} \\ a_{42}^{(1)} \end{bmatrix} \neq 0$, 则存在一个 G-矩阵 T_1 , 使得

$$T_1 \beta^{(1)} = |\beta^{(1)}| e_1 \stackrel{\Delta}{=} a_{32}^{(2)} e_1 \quad (e_1 \in \mathbb{R}^2)$$

$$\begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} A^{(1)} \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix}^T = \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(2)} & a_{24}^{(2)} \\ a_{32}^{(2)} & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

若 $\beta^{(1)} = 0$, 结束.

令 $Q = \begin{bmatrix} I_2 & \\ & T_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & T_0 \end{bmatrix}$, 则有

$$\begin{aligned} QAQ^T &= \begin{bmatrix} I_2 & \\ & T_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & T_0 \end{bmatrix} A \begin{bmatrix} 1 & \\ & T_0^T \end{bmatrix} \begin{bmatrix} I_2 & \\ & T_0^T \end{bmatrix} \\ &= \begin{bmatrix} I_2 & \\ & T_1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix} \begin{bmatrix} I_2 & \\ & T_0^T \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12}^{(1)} & * & * \\ a_{21}^{(1)} & a_{22}^{(1)} & * & * \\ 0 & a_{32}^{(2)} & * & * \\ 0 & 0 & * & * \end{bmatrix} \end{aligned}$$

Th12 设 $A \in \mathbb{R}^{n \times n}$, 则存在有限个 H-矩阵之积 Q , 使得 $QAQ^T = F_{\perp}$.

证 类似于 Th11 的证明.

推论 $A_{n \times n}$ 实对称 $\Rightarrow \exists$ 有限个 H-矩阵 (G-矩阵) 之积 Q , 使得

$$QAQ^T = \text{“实对称三对角矩阵”}$$

证 由 Th12 知, 存在 $Q = H_{u_{n-2}} \cdots H_{u_1}$, 使得 $QAQ^T = F_{\perp}$. 于是有

$$A^T = A \Rightarrow QAQ^T = (F_{\perp})^T$$

即 $F_{\perp} = (F_{\perp})^T$, 故 F_{\perp} 是 “实对称三对角矩阵”.

例 8 用 H-变换化 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 正交相似于 “三对角矩阵”.

解 $\beta^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$: $\beta^{(0)} - |\beta^{(0)}| e_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $u = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$H_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & \\ & H_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$QA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad QAQ^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

§ 4.3 满秩分解

目的: 对 $A \in C_r^{m \times n}$ ($r \geq 1$), 求 $F \in C_r^{m \times r}$ 及 $G \in C_r^{r \times n}$ 使得 $A = FG$.

一、分解原理

$$\text{rank} A = r \Rightarrow A \xrightarrow{\text{行}} \text{梯形} B = \begin{bmatrix} G \\ O \end{bmatrix}: G \in C_r^{r \times n}$$

\Rightarrow 存在有限个初等矩阵之积 $P_{m \times m}$ 使得 $PA = B$

$$\Rightarrow A = P^{-1}B = \begin{bmatrix} F_{m \times r} & S_{m \times (m-r)} \end{bmatrix} \begin{bmatrix} G \\ O \end{bmatrix} = FG: F \in C_r^{m \times r}$$

例 1 $A = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 2 & 2 & -2 & -1 \end{bmatrix}$, 求 $A = FG$.

解(1) $(A | I) = \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & & \\ 1 & 2 & -1 & 1 & & 1 & \\ 2 & 2 & -2 & -1 & & & 1 \end{array} \right] \rightarrow \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]$

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}, F = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -2 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 2 & 0 & 3 \end{bmatrix}: \text{满秩分解为 } A = FG.$$

解(2) $(A | I) \rightarrow \dots \rightarrow \left[\begin{array}{cccc|ccc} 1 & 0 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 3/2 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]$

$$A = P^{-1}B = \begin{bmatrix} F & S \end{bmatrix} \begin{bmatrix} I_2 & B_{12} \\ O & O \end{bmatrix} = \begin{bmatrix} F & FB_{12} \end{bmatrix}$$

不需计算 P^{-1}

故 $F = \text{"A 的前 2 列"} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \\ 2 & 2 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 0 & 3/2 \end{bmatrix}.$

[注] ① 满秩分解不唯一.

② 解(2)中 " F 的列向量组" 是 " A 的列向量组的部分组".

二、Hermite 标准形方法

定义: 若 $B \in C_r^{m \times n}$ ($r \geq 1$) 满足

- (1) B 的后 $m-r$ 行元素均为零 (隐含: 前 r 个行向量线性无关);
- (2) B 中有 r 列, 设为 c_1, \dots, c_r 列, 构成 I_m 的前 r 个列.

则称 B 为拟 Hermite 标准形.

[注] 矩阵的 Hermite 标准形就是矩阵的行最简形 (初等变换意义下).

使用初等行变换可将任何非零矩阵化为拟 Hermite 标准形.

Th14 设 $A \in \mathbb{C}_r^{m \times n}$ 的拟 Hermite 标准形为 B , $F = "A \text{ 的 } c_1, \dots, c_r \text{ 列}"$,
 $G = "B \text{ 的前 } r \text{ 行}"$, 则 A 的满秩分解为 $A = FG$.

证 $A \xrightarrow{\text{行}} B$ (拟 Hermite 标准形) $\Rightarrow PA = B \stackrel{\text{def}}{=} \begin{bmatrix} G \\ O \end{bmatrix}$

分解原理: $A = P^{-1}B = [F \mid S] \begin{bmatrix} G \\ O \end{bmatrix} = FG$

划分单位矩阵 $I_n = (e_1, e_2, \dots, e_n)$, 构造矩阵 $P_1 = (e_{c_1}, \dots, e_{c_r})$, 则有

$$GP_1 = "G \text{ 的 } c_1, \dots, c_r \text{ 列}" = I_r$$

$$AP_1 = (FG)P_1 = F(GP_1) = F$$

也就是 $F = "A \text{ 的 } c_1, \dots, c_r \text{ 列}"$.

提取 A 的列

例 2 $A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 2 & 0 & 2 & 2 \\ 4 & 2 & 4 & 2 \end{bmatrix}$, 求 $A = FG$.

解 $A \xrightarrow{\text{行}} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 \end{bmatrix} \xrightarrow{\text{行}} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B$

(1) $c_1 = 1, c_2 = 2$:

$$F = "A \text{ 的第 } 1, 2 \text{ 列}" = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 2 & 0 \\ 4 & 2 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

(2) $A \xrightarrow{\text{行}} B \xrightarrow{\text{行}} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = C$: 可取 $c_1 = 1, c_2 = 4$

$$F = "A \text{ 的第 } 1, 4 \text{ 列}" = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \\ 4 & 2 \end{bmatrix}, G = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

例 3 $A_1 \in \mathbb{C}_{r_1}^{m \times n}, A_2 \in \mathbb{C}_{r_2}^{m \times n} \Rightarrow \text{rank}(A_1 + A_2) \leq r_1 + r_2$.

证 $r_1 \cdot r_2 = 0$: 成立.

$$r_1 \cdot r_2 \neq 0: A_1 = F_1 G_1, A_2 = F_2 G_2 \quad (F_1 \in \mathbb{C}_{r_1}^{m \times r_1}, F_2 \in \mathbb{C}_{r_2}^{m \times r_2})$$

$$A_1 + A_2 = [F_1 \mid F_2] \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \Rightarrow \text{秩}(A_1 + A_2) \leq \text{秩}(F_1 \mid F_2) \leq r_1 + r_2$$

§ 4.4 奇异值分解

一、预备知识

(1) $\forall A_{m \times n}$, $A^H A$ 是 Hermite (半) 正定矩阵.

$$\forall x \neq 0, x^H A^H A x = (Ax)^H (Ax) = |Ax|^2 \geq 0$$

一组结论
特殊证法

(2) 齐次方程组 $Ax = 0$ 与 $A^H Ax = 0$ 同解.

若 $Ax = 0$, 则 $A^H Ax = 0$;

反之 $A^H Ax = 0 \Rightarrow |Ax|^2 = (Ax)^H (Ax) = x^H (A^H Ax) = 0 \Rightarrow Ax = 0$

(3) $\text{rank } A = \text{rank}(A^H A)$

由 $Ax = 0$ 与 $A^H Ax = 0$ 同解可得 $n - r_A = n - r_{A^H A} \Rightarrow r_A = r_{A^H A}$

(4) $A = O_{m \times n} \Leftrightarrow A^H A = O_{n \times n}$

必要性. 左乘 A^H 即得; 充分性. $r_A = r_{A^H A} = 0 \Rightarrow A = O$.

二、正交对角分解

Th15 $A_{n \times n}$ 可逆 $\Rightarrow \exists$ 酉矩阵 $U_{n \times n}$ 及 $V_{n \times n}$, 使得

$$U^H A V = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \triangleq D \quad (\sigma_i > 0)$$

证 $A^H A$ 是 Hermite 正定矩阵, \exists 酉矩阵 $V_{n \times n}$, 使得

$$V^H (A^H A) V = \text{diag}(\lambda_1, \dots, \lambda_n) \triangleq \Lambda \quad (\lambda_i > 0)$$

改写为 $D^{-1} V^H A^H \cdot A V D^{-1} = I \quad (\sigma_i = \sqrt{\lambda_i})$

令 $U = A V D^{-1}$, 则有 $U^H U = I$, 从而 U 是酉矩阵.

由此可得 $U^H A V = U^H U D = D$.

[注] 称 $A = U D V^H$ 为 A 的正交对角分解.

三、奇异值分解

$A \in C_r^{m \times n} (r \geq 1) \Rightarrow A^H A \in C_r^{n \times n}$ 半正定 半正定: 特征值大于等于0

$A^H A$ 的特征值: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$

A 的奇异值: $\sigma_i = \sqrt{\lambda_i} \quad (i = 1, 2, \dots, n)$

特点: (1) A 的奇异值个数等于 A 的列数;

(2) A 的非零奇异值个数 = $\text{rank } A$.

Th16 $A \in C_r^{m \times n} (r \geq 1)$, $\Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$, $D = \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix}_{m \times n} \Rightarrow$

存在酉矩阵 $U_{m \times m}$ 及 $V_{n \times n}$, 使得 $U^H A V = D$.

证 对于 Hermite 半正定矩阵 $A^H A$, 存在酉矩阵 $V_{n \times n}$, 使得

$$V^H (A^H A) V = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \Sigma^2 & O \\ O & O \end{bmatrix}_{n \times n}$$

划分 $V = [V_1 \mid V_2]$: V_1 是 V 的前 r 列, V_2 是 V 的后 $n-r$ 列, 则有

$$A^H A V = V \begin{bmatrix} \Sigma^2 & O \\ O & O \end{bmatrix}_{n \times n}, \text{ 即 } A^H A [V_1 \mid V_2] = [V_1 \mid V_2] \begin{bmatrix} \Sigma^2 & O \\ O & O \end{bmatrix}_{n \times n}$$

也就是 $[A^H A V_1 \mid A^H A V_2] = [V_1 \Sigma^2 \mid O]$:

$$(1) A^H A V_2 = O: V_2^H A^H A V_2 = O \Rightarrow (A V_2)^H (A V_2) = O \Rightarrow A V_2 = O$$

$$(2) A^H A V_1 = V_1 \Sigma^2: V_1^H A^H A V_1 = \Sigma^2 \Rightarrow (A V_1 \Sigma^{-1})^H (A V_1 \Sigma^{-1}) = I_r$$

令 $U_1 \stackrel{\Delta}{=} A V_1 \Sigma^{-1}$, 则有 $U_1^H U_1 = I_r$

设 U_1 的列为 u_1, \dots, u_r , 扩充为 C^m 的**标准正交基** $u_1, \dots, u_r, u_{r+1}, \dots, u_m$, 课本定理 1.3

则 $U_2 = (u_{r+1}, \dots, u_m)$ 满足 $U_2^H U_1 = O_{(m-r) \times r}$. 记 $U = [U_1 \mid U_2]$, 则有

$$\begin{aligned} U^H A V &= U^H [A V_1 \mid A V_2] = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} [U_1 \Sigma \mid O] \\ &= \begin{bmatrix} U_1^H U_1 \Sigma & O \\ U_2^H U_1 \Sigma & O \end{bmatrix} = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix}_{m \times n} \end{aligned}$$

[注] 称 $A = U D V^H$ 为 A 的奇异值分解.

(1) U 与 V 不唯一;

性质 (2) 见课本 P163 页

(2) U 的列向量为 $A A^H$ 的特征向量, V 的列向量为 $A^H A$ 的特征向量;

(3) 称 U 的列向量为 A 的**左奇异向量**, V 的列向量为 A 的**右奇异向量**.

例 1 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, 求 $A = U D V^T$.

解 $A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = B, \quad |\lambda I - B| = \lambda(\lambda - 1)(\lambda - 3)$

$$\lambda_1 = 3: 3I - B = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \quad \xi_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\lambda_2 = 1: 1I - B = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \quad \xi_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\lambda_3 = 0: 0I - B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{bmatrix}, \xi_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$r_A = 2, \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}, V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{bmatrix}, V_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{bmatrix}$$

$$U_1 = AV_1\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix}, \text{ 取 } U_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \text{ 则 } U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U^T AV = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = D, \underline{A = UDV^T}$$

Th17 在 $A \in \mathbb{C}^{m \times n}$ ($r > 0$) 的奇异值分解 $A = U \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} V^H$ 中, 划分

$U = (u_1, u_2, \dots, u_m), V = (v_1, v_2, \dots, v_n)$, 则有

- (1) $N(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_n\}$;
- (2) $R(A) = \text{span}\{u_1, u_2, \dots, u_r\}$;
- (3) $A = \sigma_1 u_1 v_1^H + \sigma_2 u_2 v_2^H + \dots + \sigma_r u_r v_r^H$.

比较谱分解

证 $A = [U_1 \mid U_2] \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} = U_1 \Sigma V_1^H$

容易验证: $U_1 \Sigma V_1^H x = 0 \Leftrightarrow V_1^H x = 0$

$$\begin{aligned} (1) \quad N(A) &= \{x \mid Ax = 0\} = \{x \mid U_1 \Sigma V_1^H x = 0\} \\ &= \{x \mid V_1^H x = 0\} = N(V_1) = R^\perp(V_1) \\ &= R(V_2) = \text{span}\{v_{r+1}, \dots, v_n\} \end{aligned}$$

(Th1-35) 正交补:
P65

(因为 $R(V_1) \perp R(V_2)$, $R(V_1) \oplus R(V_2) = \mathbb{C}^n$, 所以 $R^\perp(V_1) = R(V_2)$)

$$(2) \quad \text{任意 } y \in R(A) \Rightarrow y = Ax = (U_1 \Sigma V_1^H)x = U_1 (\Sigma V_1^H x) \in R(U_1)$$

故 $R(A) \subset R(U_1)$

$$\text{任意 } y \in R(U_1) \Rightarrow y = U_1 z = (AV_1 \Sigma^{-1})z = A(V_1 \Sigma^{-1} z) \in R(A)$$

故 $R(U_1) \subset R(A)$

由此可得 $R(A) = R(U_1) = \text{span}\{u_1, \dots, u_r\}$

$$(3) \quad A = (u_1, \dots, u_r) \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^H \\ \vdots \\ v_r^H \end{bmatrix} = \sigma_1 u_1 v_1^H + \dots + \sigma_r u_r v_r^H$$

四、正交相抵

$A_{m \times n}, B_{m \times n}$, 若有酉矩阵 $U_{m \times m}$ 及 $V_{n \times n}$ 使 $U^H A V = B$, 称 A 与 B 正交相抵.

性质: (1) A 与 A 正交相抵;

(2) A 与 B 正交相抵 $\Rightarrow B$ 与 A 正交相抵;

(3) A 与 B 正交相抵, B 与 C 正交相抵 $\Rightarrow A$ 与 C 正交相抵.

Th18 A 与 B 正交相抵 $\Rightarrow \sigma_A = \sigma_B$.

证 $B = U^H A V \Rightarrow B^H B = \dots = V^{-1} (A^H A) V$

$$\Rightarrow \lambda_{B^H B} = \lambda_{A^H A} \geq 0 \Rightarrow \sigma_B = \sigma_A$$

例 2 $A^H = A \Rightarrow \sigma_A = |\lambda_A| \quad (\because \lambda_{A^H A} = \lambda_{A^2} = (\lambda_A)^2)$

$$A^H = -A \Rightarrow \sigma_A = |\lambda_A| \quad (\because \lambda_{A^H A} = \lambda_{(jA)^2} = (j\lambda_A)^2)$$

($A^H = -A \Rightarrow \lambda_A$ 为 0 或纯虚数, $j\lambda_A$ 为实数)

矩阵分解的应用: 设方程组 $A_{m \times n} x = b$ 有解, 则有

(1) $m = n: A = LU \Rightarrow Ly = b, Ux = y$

(2) $m = n: A = QR \Rightarrow Rx = Q^T b$

(3) $A = UDV^H \Rightarrow Dy = U^H b \stackrel{\text{def}}{=} c, V^H x = y$

$$\text{其中 } D = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix}_{m \times n}, \quad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

$$Dy = c \Leftrightarrow \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}; \quad \text{隐含} \begin{bmatrix} c_{r+1} \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (\text{判断可解性})$$

$$\text{通解为} \quad \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} c_1/\sigma_1 \\ \vdots \\ c_r/\sigma_r \\ k_1 \\ \vdots \\ k_{n-r} \end{bmatrix} \quad (k_1, \dots, k_{n-r} \text{ 是任意常数})$$

$$x = Vy = \left(\frac{c_1}{\sigma_1} v_1 + \dots + \frac{c_r}{\sigma_r} v_r \right) + (k_1 v_{r+1} + \dots + k_{n-r} v_n)$$

[注] $k_1 v_{r+1} + \dots + k_{n-r} v_n$ 是 $A_{m \times n} x = 0$ 的通解 (Th17(1))

$$\text{因为 } A\left(\frac{c_1}{\sigma_1}v_1 + \cdots + \frac{c_r}{\sigma_r}v_r\right) = AV_1\Sigma^{-1}\begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} = U_1\begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} = [U_1 \mid U_2]c = b$$

所以 $\frac{c_1}{\sigma_1}v_1 + \cdots + \frac{c_r}{\sigma_r}v_r$ 是 $A_{m \times n}x = b$ 的一个特解

矩阵论辅导讲案-例 4.7

设欧氏空间 V^n 的一个基为 x_1, x_2, \dots, x_n , 它的度量矩阵为 A , 证明: 存在实对称正定矩阵 C , 使得由

$$(y_1, y_2, \dots, y_n) = (x_1, x_2, \dots, x_n)C$$

确定的基 y_1, y_2, \dots, y_n 为 V^n 的标准正交基.

分析 只要能够确定实对称正定矩阵 C , 使得 $C^T AC = I$ 即可.

证 因为 A 对称正定, 所以存在正交矩阵 Q , 使得

$$Q^T A Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

其中 $\lambda_i (i = 1, 2, \dots, n)$ 是 A 的特征值, 它们都是正数. 记

$$D = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$$

则有

$$A = QD^2Q^T = (QDQ^T)(QDQ^T)$$

令

$$C = (QDQ^T)^{-1} = QD^{-1}Q^T$$

则 C 为对称正定矩阵, 且有 $C^T AC = I$. 此时, 基 y_1, y_2, \dots, y_n 的度量矩阵为 $C^T AC = I$, 从而是 V^n 的一个标准正交基.

注 1 本例给出由已知基及其度量矩阵构造标准正交基的矩阵分解方法.

注 2 如果不要要求 C 为对称正定矩阵, 而只要求 C 为可逆矩阵, 那么由

$$A = QD^2Q^T = (QD)(DQ^T) = (DQ^T)^T(DQ^T)$$

知, 取 $C = (DQ^T)^{-1} = QD^{-1}$ 时, 有

$$A = (C^{-1})^T C^{-1}, \quad C^T AC = C^T \cdot (C^{-1})^T C^{-1} \cdot C = I$$

此时, 基 y_1, y_2, \dots, y_n 是 V^n 的一个标准正交基.

注 3 因为 A 对称正定, 所以可对 A 进行三角分解 (Cholesky 分解), 即

$$A = LL^T \quad (L \text{ 为可逆下三角矩阵})$$

取 $C = (L^T)^{-1}$ 时, 有

$$A = (C^{-1})^T C^{-1}, \quad C^T AC = L^{-1} \cdot LL^T \cdot (L^T)^{-1} = I$$

此时, 基 y_1, y_2, \dots, y_n 是 V^n 的一个标准正交基.

加例 设 $A_{m \times n}$ 的奇异值分解为 $A = U \begin{pmatrix} \Sigma_r & O \\ O & O \end{pmatrix} V^H$, 求 $\begin{pmatrix} A \\ A \end{pmatrix}$ 的奇异值分解.

解 在 A 的奇异值分解中有下列关系:

$$V^H(A^H A)V = \begin{pmatrix} \Sigma^2 & O \\ O & O \end{pmatrix}, \quad V_1^H(A^H A)V_1 = \Sigma^2, \quad U_1 = AV_1\Sigma^{-1}$$

考虑 $B = \begin{pmatrix} A \\ A \end{pmatrix}$, 由 $B^H B = 2A^H A$ 可得

$$V^H(B^H B)V = \begin{pmatrix} 2\Sigma^2 & O \\ O & O \end{pmatrix}, \quad V_1^H(B^H B)V_1 = (\sqrt{2}\Sigma)^2$$

构造矩阵 $\tilde{U}_1 = BV_1(\sqrt{2}\Sigma)^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}}U_1 \\ \frac{1}{\sqrt{2}}U_1 \end{pmatrix}$, $\tilde{U}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}}U_1 & U_2 & O \\ \frac{-1}{\sqrt{2}}U_1 & O & U_2 \end{pmatrix}$

则 $\tilde{U} = (\tilde{U}_1 \mid \tilde{U}_2)$ 是 $2m$ 阶酉矩阵, 且有

$$\tilde{U}^H B V = \begin{pmatrix} \frac{1}{\sqrt{2}}U_1^H & \frac{1}{\sqrt{2}}U_1^H \\ \frac{1}{\sqrt{2}}U_1^H & \frac{-1}{\sqrt{2}}U_1^H \\ U_2^H & O \\ O & U_2^H \end{pmatrix} \begin{pmatrix} A \\ A \end{pmatrix} (V_1 \mid V_2) = \begin{pmatrix} \sqrt{2}U_1^H A V_1 & \sqrt{2}U_1^H A V_2 \\ O & O \\ U_2^H A V_1 & U_2^H A V_2 \\ U_2^H A V_1 & U_2^H A V_2 \end{pmatrix}$$

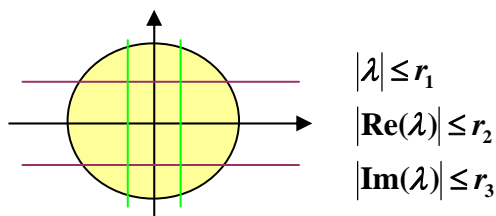
注意到 $AV_1 = U_1\Sigma$, $AV_2 = O$, $U_2^H U_1 = O$ 可得 $\tilde{U}^H B V = \begin{pmatrix} \sqrt{2}\Sigma & O \\ O & O \end{pmatrix}$.

于是 B 的奇异值分解为 $B = \tilde{U} \begin{pmatrix} \sqrt{2}\Sigma & O \\ O & O \end{pmatrix} V^H$.

第五章 特征值的估计

§ 5.1 特征值估计

一、特征值的上界



$$\text{Th1 } A \in \mathbf{R}^{n \times n}, M = \frac{1}{2} \cdot \max_{i,j} \{ |a_{ij} - a_{ji}| \} \Rightarrow |\text{Im}(\lambda_A)| \leq M \left[\frac{1}{2} n(n-1) \right]^{\frac{1}{2}}$$

证明略去.

$$\leq M \frac{n}{\sqrt{2}} \quad (\text{比较 Th2})$$

例 1 $A_{n \times n}$ 是实对称矩阵 $\Rightarrow \text{Im}(\lambda_A) = 0$, 即 $\lambda_A \in \mathbf{R}$.

$$\text{引理 1 } \forall B = (b_{ij})_{n \times n}, \forall y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \text{ 满足 } \|y\|_2 = 1 \Rightarrow |y^H B y| \leq \|B\|_{m_\infty}.$$

$$\begin{aligned} \text{证 } |y^H B y| &= \left| \sum_{i,j} b_{ij} \bar{\eta}_i \eta_j \right| \leq \max_{i,j} |b_{ij}| \cdot \sum_{i,j} |\eta_i| |\eta_j| \\ &\leq \max_{i,j} |b_{ij}| \cdot \frac{1}{2} \sum_{i,j} (|\eta_i|^2 + |\eta_j|^2) = \max_{i,j} |b_{ij}| \cdot \frac{1}{2} (n + n) = \|B\|_{m_\infty} \end{aligned}$$

$$\text{Th2 } A_{n \times n}, |\lambda_A| \leq \|A\|_{m_\infty}, |\text{Re}(\lambda_A)| \leq \frac{1}{2} \|A + A^H\|_{m_\infty}, |\text{Im}(\lambda_A)| \leq \frac{1}{2} \|A - A^H\|_{m_\infty}.$$

证 设 A 的特征值为 λ , 特征向量 x 满足 $\|x\|_2 = 1$, 则

$$Ax = \lambda x \Rightarrow \lambda = x^H Ax, \quad \bar{\lambda} = x^H A^H x$$

$$(1) |\lambda| = |x^H Ax| \leq \|A\|_{m_\infty}$$

$$(2) \lambda + \bar{\lambda} = x^H (A + A^H) x \Rightarrow |\text{Re}(\lambda)| = \frac{1}{2} |\lambda + \bar{\lambda}| \leq \frac{1}{2} \|A + A^H\|_{m_\infty}$$

$$(3) \lambda - \bar{\lambda} = x^H (A - A^H) x \Rightarrow |\text{Im}(\lambda)| = \frac{1}{2} |\lambda - \bar{\lambda}| \leq \frac{1}{2} \|A - A^H\|_{m_\infty}$$

例 2 $A^H = A \Rightarrow \text{Im}(\lambda_A) = 0$, 即 $\lambda_A \in \mathbf{R}$.

$A^H = -A \Rightarrow \text{Re}(\lambda_A) = 0$, 即 λ_A 是零或纯虚数.

$$\text{Th4 } A_{n \times n} = (a_1, a_2, \dots, a_n) \Rightarrow |\lambda_1| |\lambda_2| \cdots |\lambda_n| \leq \|a_1\|_2 \|a_2\|_2 \cdots \|a_n\|_2.$$

证 $\det A = 0$ 时, 结论成立.

$\det A \neq 0 \Rightarrow a_1, a_2, \dots, a_n$ 线性无关. 正交化可得

$$\begin{cases} b_1 = a_1 \\ b_2 = a_2 - k_{21} b_1 \\ \dots \\ b_n = a_n - k_{n,n-1} b_{n-1} - \dots - k_{n1} b_1 \end{cases}, \quad \begin{cases} a_1 = b_1 \\ a_2 = k_{21} b_1 + b_2 \\ \dots \\ a_n = k_{n1} b_1 + \dots + k_{n,n-1} b_{n-1} + b_n \end{cases}$$

$$(a_1, a_2, \dots, a_n) = (b_1, b_2, \dots, b_n) \begin{bmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$B \stackrel{\Delta}{=} (b_1, b_2, \dots, b_n): \det A = \det B, \quad b_i \perp b_j \quad (i \neq j)$$

$$\begin{aligned} \|a_i\|_2^2 &= \|(k_{i1}b_1 + \cdots + k_{i,i-1}b_{i-1}) + b_i\|_2^2 \\ &= \|k_{i1}b_1 + \cdots + k_{i,i-1}b_{i-1}\|_2^2 + \|b_i\|_2^2 \geq \|b_i\|_2^2 \end{aligned}$$

$$|\det B|^2 = \overline{\det B} \cdot \det B = \det \bar{B} \cdot \det B = \det B^H \cdot \det B = \det(B^H B)$$

$$= \det \begin{bmatrix} \|b_1\|_2^2 & & \\ & \ddots & \\ & & \|b_n\|_2^2 \end{bmatrix} = (\|b_1\|_2 \cdots \|b_n\|_2)^2$$

$$|\lambda_1| \cdots |\lambda_n| = |\det A| = |\det B| = \|b_1\|_2 \cdots \|b_n\|_2 \leq \|a_1\|_2 \cdots \|a_n\|_2$$

比较:

$$\rho(A) \leq \|A\|_F$$

Th5 设 $A_{n \times n}$ 的特征值为 $\lambda_1, \dots, \lambda_n$, 则 $|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|A\|_F^2$.

证 对 A , 存在酉矩阵 $U_{n \times n}$, 使得 $U^H A U = \begin{bmatrix} \lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_n \end{bmatrix} \stackrel{\Delta}{=} T$

$$\text{由此可得 } |\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|T\|_F^2 = \|A\|_F^2 \quad (\text{Th2.3})$$

例 3 设 $A_{n \times n}$ 的特征值为 $\lambda_1, \dots, \lambda_n$, 则

$$A^H A = A A^H \Leftrightarrow |\lambda_1|^2 + \cdots + |\lambda_n|^2 = \|A\|_F^2$$

证 必要性. $A^H A = A A^H$: 存在酉矩阵 P , 使得

$$P^H A P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \stackrel{\Delta}{=} A, \quad |\lambda_1|^2 + \cdots + |\lambda_n|^2 = \|A\|_F^2 = \|A\|_F^2$$

充分性. 对 A , 存在酉矩阵 U , 使得 $U^H A U = \begin{bmatrix} \lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_n \end{bmatrix} \stackrel{\Delta}{=} T$

$$\|T\|_F^2 = \|A\|_F^2 = |\lambda_1|^2 + \cdots + |\lambda_n|^2 \Rightarrow t_{ij} = 0 \quad (i < j)$$

$$T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Rightarrow T^H T = T T^H \Rightarrow A^H A = \cdots = A A^H$$

二、特征值的包含域

1. 盖尔圆: 设 $A = (a_{ij})_{n \times n}$, 记 $R_i = \sum_{j=1, j \neq i}^n |a_{ij}|$ ($i = 1, 2, \dots, n$), 称

$G_i = \{z \mid |z - a_{ii}| \leq R_i\}$ 为 A 的第 i 个盖尔圆 (Gerschgorin 圆).

Th6 矩阵 $A_{n \times n}$ 的特征值 $\lambda_A \in \bigcup_{i=1}^n G_i$.

证 设 A 的特征值为 λ , 对应的特征向量为 $x = (\xi_1, \xi_2, \dots, \xi_n)^T$,

选取 i_0 使得 $|\xi_{i_0}| = \max_i |\xi_i|$, 则 $\xi_{i_0} \neq 0$, 且有

$$Ax = \lambda x \Rightarrow \sum_{j=1}^n a_{i_0 j} \xi_j = \lambda \xi_{i_0} \quad (i_0 - \text{分量}) \Rightarrow (\lambda - a_{i_0 i_0}) \xi_{i_0} = \sum_{j \neq i_0} a_{i_0 j} \xi_j$$

$$\Rightarrow |\lambda - a_{i_0 i_0}| = \left| \sum_{j \neq i_0} a_{i_0 j} \frac{\xi_j}{\xi_{i_0}} \right| \leq \sum_{j \neq i_0} |a_{i_0 j}| \left| \frac{\xi_j}{\xi_{i_0}} \right| \leq R_{i_0}$$

$$\Rightarrow \lambda \in G_{i_0} \subset \bigcup_{i=1}^n G_i$$

例 4 $A = \begin{bmatrix} 10 & 1 & 2 & 3 \\ 5 & 30 & 1 & 2 \\ 10 & 3 & -10 & 5 \\ 2 & -3 & -1 & -40 \end{bmatrix}$, 估计 λ_A .

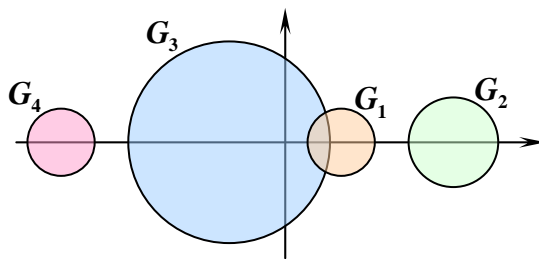
$$G_1: |z - 10| \leq 6$$

$$G_2: |z - 30| \leq 8$$

$$G_3: |z + 10| \leq 18$$

$$G_4: |z + 40| \leq 6$$

$$\lambda_A \in G_1 \cup G_2 \cup G_3 \cup G_4$$



2. 连通部分: (1) 孤立的 G-圆, 或

(2) 相交的 G-圆构成的最大连通区域.

Th7 设 S 为 $A_{n \times n}$ 的 G-圆的一个连通部分, 则

S 由 k 个 G-圆构成 $\Leftrightarrow S$ 中恰好有 A 的 k 个特征值.

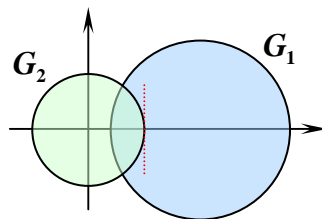
其中 G-圆重叠时重复计数, 特征值相同是亦重复计数.

例 5 $A = \begin{bmatrix} 10 & -8 \\ 5 & 0 \end{bmatrix}$, 估计 λ_A .

$$G_1: |z - 10| \leq 8; \quad G_2: |z| \leq 5$$

$$S = G_1 \cup G_2 \text{ 是一个连通部分} \stackrel{\text{Th7}}{\Rightarrow} \lambda_{1,2} \in S.$$

$$\text{计算得 } \lambda_{1,2} = 5 \pm j\sqrt{15}, \quad |\lambda_{1,2}| = \sqrt{40} > 5 \Rightarrow \lambda_{1,2} \notin G_2.$$



3. 特征值的分离问题

找 n 个孤立的 G-圆, 使其中各包含 $A_{n \times n}$ 的一个特征值.

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \quad (d_i > 0)$$

$$B \stackrel{\Delta}{=} DAD^{-1} = \left(a_{ij} \frac{d_i}{d_j} \right)_{n \times n} = \begin{pmatrix} a_{11} & a_{12} \frac{d_1}{d_2} & \cdots & a_{1n} \frac{d_1}{d_n} \\ a_{21} \frac{d_2}{d_1} & a_{22} & \cdots & a_{2n} \frac{d_2}{d_n} \\ \vdots & \vdots & & \vdots \\ a_{n1} \frac{d_n}{d_1} & a_{n2} \frac{d_n}{d_2} & \cdots & a_{nn} \end{pmatrix}$$

$$B \text{ 的 G-圆: } G'_i = \{z \mid |z - a_{ii}| \leq r_i\}, \quad r_i \stackrel{\Delta}{=} \sum_{j \neq i} |a_{ij}| \frac{d_i}{d_j}$$

$$A \text{ 相似于 } B \Rightarrow \lambda_A \in \bigcup_{i=1}^n G'_i \quad (\text{Th6})$$

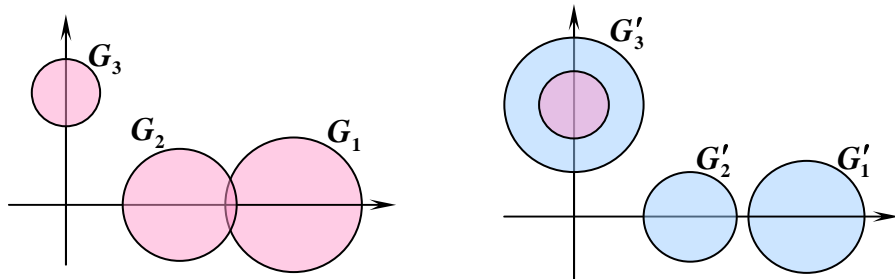
[注] 当 G'_i 是 B 的孤立 G-圆时, G'_i 中恰有 A 的一个特征值 (Th7).

例 6 $A = \begin{bmatrix} 20 & 5 & 0.8 \\ 4 & 10 & 1 \\ 1 & 2 & 10j \end{bmatrix}$, 分离 λ_A .

解 $G_1: |z - 20| \leq 5.8 \quad G_2: |z - 10| \leq 5 \quad G_3: |z - 10j| \leq 3$

对 A , 取 $D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}$, 则 $B = DAD^{-1} = \begin{bmatrix} 20 & 5 & 0.4 \\ 4 & 10 & 0.5 \\ 2 & 4 & 10j \end{bmatrix}$.

$$G'_1: |z - 20| \leq 5.4 \quad G'_2: |z - 10| \leq 4.5 \quad G'_3: |z - 10j| \leq 6$$



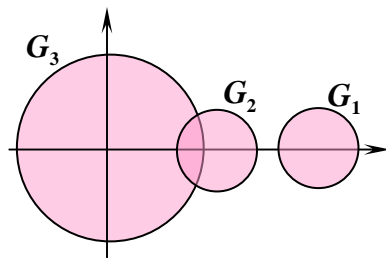
结论: G'_1, G'_2 及 G_3 中各有 A 的一个特征值.

[注] 在例 6 中, G_3 孤立, G'_3 孤立, 故可取较小者;

A^T 的 G-圆称为 A 的列 G-圆, A^T 与 A 的特征值相同.

例 7 $A = \begin{bmatrix} 20 & 3 & 1 \\ 2 & 10 & 2 \\ 8 & 1 & 0 \end{bmatrix}$, 分离 λ_A .

解 (1) $G_1: |z-20| \leq 4$
 $G_2: |z-10| \leq 4$
 $G_3: |z| \leq 9$
 G_1 孤立, G_2 与 G_3 相交.

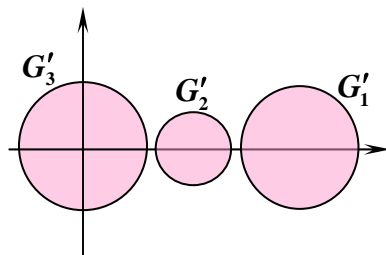


(2) 对 A , 取 $D = \begin{bmatrix} 1.5 & & \\ & 1 & \\ & & 1 \end{bmatrix}$, 则 $B = DAD^{-1} = \begin{bmatrix} 20 & 4.5 & 1.5 \\ 4/3 & 10 & 2 \\ 16/3 & 1 & 0 \end{bmatrix}$.

$$G'_1: |z-20| \leq 6$$

$$G'_2: |z-10| \leq \frac{10}{3}$$

$$G'_3: |z| \leq \frac{19}{3}$$



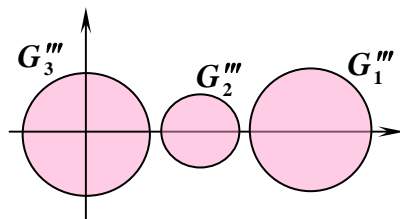
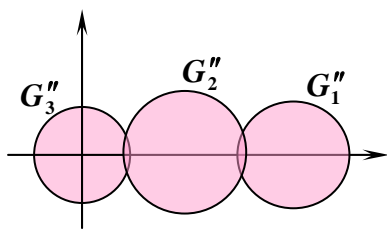
G'_1, G'_2, G'_3 都是孤立的盖尔圆, 其中各有 A 的一个特征值.

结论: 结合(1)可得 G_1, G_2 及 G_3 中各有 A 的一个特征值.

(3) 对 A , 取 $D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0.5 \end{bmatrix}$, 则 $B = DAD^{-1} = \begin{bmatrix} 20 & 3 & 2 \\ 2 & 10 & 4 \\ 4 & 0.5 & 0 \end{bmatrix}$.

$$G''_1: |z-20| \leq 5 \quad G''_2: |z-10| \leq 6 \quad G''_3: |z| \leq 4.5$$

G''_1 与 G''_2 相交, G''_2 与 G''_3 相交.



考虑 B^T . $G'''_1: |z-20| \leq 6 \quad G'''_2: |z-10| \leq 3.5 \quad G'''_3: |z| \leq 6$

G'''_1, G'''_2, G'''_3 都是孤立的盖尔圆, 其中各有 A 的一个特征值.

结论: 结合(1)可得 G_1, G_2 及 G_3 中各有 A 的一个特征值.

例 8 $A_{n \times n}$ 严格对角占优 $\Rightarrow \det A \neq 0$.

证 只证“行优”情形: $|a_{ii}| > \sum_{j \neq i} |a_{ij}| = R_i \quad (i=1, 2, \dots, n)$

$$\forall \lambda_A, \exists i_0, \text{st. } \lambda_A \in G_{i_0} \Rightarrow |\lambda_A - a_{i_0 i_0}| \leq R_{i_0}$$

若 $\lambda_A = 0$, 则 $|a_{i_0 i_0}| \leq R_{i_0}$ 与 A “行优”矛盾. 故 $\lambda_A \neq 0 \Rightarrow \det A \neq 0$.

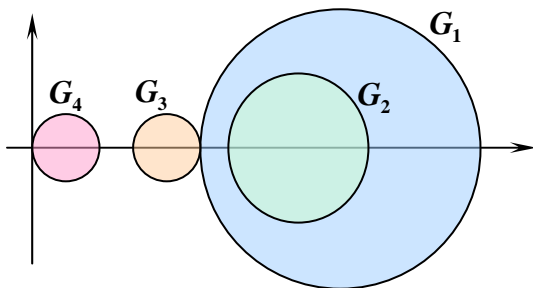
加例 应用盖尔圆定理说明 $A = \begin{bmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ 至少有两个实特征值.

解 $G_1: |z-9| \leq 4$

$G_2: |z-8| \leq 2$

$G_3: |z-4| \leq 1$

$G_4: |z-1| \leq 1$



两个连通部分关于实轴对称:

$S_1 = G_4$ 中只有 A 的一个特征值, 该特征值为实数;

$S_2 = G_1 \cup G_2 \cup G_3$ 中有 A 的三个特征值, 其中至少有一个特征值为实数.

[注] 实矩阵的复特征值一定成对共轭出现.

4. 矩阵 $A = (a_{ij})_{n \times n}$ ($n > 1$) 的 **Cassini 卵形**

$$\Omega_{ij} = \left\{ z \mid |z - a_{ii}| \cdot |z - a_{jj}| \leq R_i \cdot R_j \right\} \quad (i \neq j)$$

Th12 矩阵 $A_{n \times n}$ ($n > 1$) 的特征值 $\lambda_A \in \bigcup_{i \neq j} \Omega_{ij}$.

证 设 A 的特征值为 λ , 对应的特征向量为 $x = (\xi_1, \xi_2, \dots, \xi_n)^T$.

选取 $i_0 \neq j_0$ 满足 $|\xi_{i_0}| \geq |\xi_{j_0}| \geq |\xi_k|$ ($k \neq i_0, j_0$), 下证 $\lambda \in \Omega_{i_0 j_0}$.

(1) $\xi_{j_0} = 0$: 此时 $\xi_k = 0$ ($k \neq i_0, j_0$), $\xi_{i_0} \neq 0$ ($\because x \neq 0$)

$$\begin{aligned} Ax = \lambda x &\Rightarrow \lambda \xi_{i_0} = \sum_{k=1}^n a_{i_0 k} \xi_k = a_{i_0 i_0} \xi_{i_0} \Rightarrow \lambda = a_{i_0 i_0} \\ &\Rightarrow |\lambda - a_{i_0 i_0}| \cdot |\lambda - a_{j_0 j_0}| = 0 \leq R_{i_0} R_{j_0} \end{aligned}$$

(2) $\xi_{j_0} \neq 0$: 此时 $|\xi_{i_0}| \geq |\xi_{j_0}| > 0$

$$Ax = \lambda x \Rightarrow \sum_{k=1}^n a_{lk} \xi_k = \lambda \xi_l \Rightarrow (\lambda - a_{ll}) \xi_l = \sum_{k \neq l} a_{lk} \xi_k \quad (l = 1, 2, \dots, n)$$

$$l = i_0: |\lambda - a_{i_0 i_0}| \cdot |\xi_{i_0}| \leq \sum_{k \neq i_0} |a_{i_0 k}| \cdot |\xi_k| \leq |\xi_{j_0}| \cdot R_{i_0}$$

$$l = j_0: |\lambda - a_{j_0 j_0}| \cdot |\xi_{j_0}| \leq \sum_{k \neq j_0} |a_{j_0 k}| \cdot |\xi_k| \leq |\xi_{i_0}| \cdot R_{j_0}$$

$$\text{故 } |\lambda - a_{i_0 i_0}| \cdot |\lambda - a_{j_0 j_0}| \leq R_{i_0} R_{j_0}$$

因此 $\lambda \in \Omega_{i_0 j_0} \subset \bigcup_{i \neq j} \Omega_{ij}$.

广义行优

例 9 $A_{n \times n}$ ($n > 1$) 满足 $|a_{ii}| \cdot |a_{jj}| > R_i R_j$ ($i \neq j$) $\Rightarrow \det A \neq 0$.

证 因为 $\forall i \neq j, |0 - a_{ii}| \cdot |0 - a_{jj}| = |a_{ii}| \cdot |a_{jj}| > R_i R_j$

所以 $0 \notin \Omega_{ij} \Rightarrow 0$ 不是 A 的特征值 $\Rightarrow \det A \neq 0$

例 10 $A = \begin{bmatrix} 20 & 11 & 10 \\ -8 & 30 & 20 \\ 13 & 11 & 30 \end{bmatrix}, R_1 = 21, R_2 = 28, R_3 = 24$

$$|a_{11}| \cdot |a_{22}| = 600 > 588 = R_1 R_2, |a_{11}| \cdot |a_{33}| = 600 > 504 = R_1 R_3$$

$$|a_{22}| \cdot |a_{33}| = 900 > 672 = R_2 R_3. \text{ 由例 9 可得 } \det A \neq 0.$$

例 11 $\bigcup_{i \neq j} \Omega_{ij} \subset \bigcup_{k=1}^n G_k$

证 $\forall z \in \text{左}, \exists i \neq j, \text{st. } z \in \Omega_{ij} \Rightarrow |z - a_{ii}| \cdot |z - a_{jj}| \leq R_i R_j$

$\Rightarrow |z - a_{ii}| > R_i$ 与 $|z - a_{jj}| > R_j$ 至少之一不成立

$\Rightarrow |z - a_{ii}| \leq R_i$ 与 $|z - a_{jj}| \leq R_j$ 至少之一成立

$\Rightarrow z \in G_i$ 或者 $z \in G_j$, 即 $z \in \text{右}$.

§ 5.2 广义特征值问题

$A_{n \times n}$ 实对称, $B_{n \times n}$ 实对称正定, 求常数 λ 及非零向量 x 使得 $Ax = \lambda Bx$.

直接解法: $Ax = \lambda Bx \Leftrightarrow (\lambda B - A)x = 0$

由 $|\lambda B - A| = 0$ 求广义特征值 λ_i

由 $(\lambda_i B - A)x = 0$ 求广义特征向量 x

一般不限制
 A, B 实对称

一、等价形式

(1) $Ax = \lambda Bx \Leftrightarrow (B^{-1}A)x = \lambda x$: $B^{-1}A$ 不一定对称

(2) B 对称正定 $\Rightarrow B = GG^T$: G 可逆

$$Ax = \lambda Bx \Leftrightarrow G^{-1}A(G^{-1})^T G^T x = \lambda G^T x \Leftrightarrow Sy = \lambda y$$

其中 $S \stackrel{\Delta}{=} G^{-1}A(G^{-1})^T$ 对称, $y \stackrel{\Delta}{=} G^T x$.

二、正交性

1. 按 B -正交: 设 B 为实对称正定矩阵, 列向量 $x_1, \dots, x_m \in \mathbb{R}^n$.

若 $x_i^T B x_j = 0$ ($i \neq j$), 称 x_1, \dots, x_m 按 B -正交.

性质 非零向量 x_1, \dots, x_m 按 B -正交 $\Rightarrow x_1, \dots, x_m$ 线性无关.

证 设 $k_1 x_1 + \dots + k_i x_i + \dots + k_m x_m = 0$ (注意 B 是对称正定矩阵)

左乘 $x_i^T B$: $k_i (x_i^T B x_i) = 0 \Rightarrow k_i = 0$, 故 x_1, \dots, x_m 线性无关.

2. 按 B -标准正交: 设 x_1, \dots, x_m 按 B -正交, 若 $\|x_i\|_B = 1$ ($i = 1, \dots, m$),

称 x_1, \dots, x_m 按 B -标准正交. ($\|x\|_B = \sqrt{x^T B x}$)

三、广义特征向量的正交性

S 实对称 \Rightarrow 特征值: $\lambda_1, \dots, \lambda_n \in \mathbf{R}$

特征向量: $y_1, \dots, y_n \in \mathbf{R}^n$ 标准正交

$$\begin{aligned} x_i = (G^{-1})^T y_i &\Rightarrow x_i^T B x_j = y_i^T G^{-1} \cdot G G^T \cdot (G^{-1})^T y_j = y_i^T y_j = \delta_{ij} \\ &\Rightarrow x_1, \dots, x_n \text{ 按 } B\text{-标准正交} \Rightarrow x_1, \dots, x_n \text{ 线性无关} \end{aligned}$$

§ 5.3 对称矩阵特征值的极值

一、常义 Rayleigh 商:

问题: 若 x 是实对称矩阵 A 的特征向量, 则有 $Ax = \lambda x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$.

下面讨论多元函数 $R(x) = \frac{x^T Ax}{x^T x}$ ($x \neq 0$) 的极值问题.

(1) $R(x)$ 连续;

(2) \forall 实数 $k \neq 0$, $R(kx) = R(x)$;

(3) $\forall x_0 \neq 0$, 当 $0 \neq x \in L(x_0)$ 时, $R(x) = R(x_0)$;

$$\because x \in L(x_0) \Rightarrow x = kx_0 \Rightarrow R(x) = R(kx_0) = R(x_0)$$

(4) $\min_{x \neq 0} R(x)$ 与 $\max_{x \neq 0} R(x)$ 存在, 且在 $S = \{x \mid x \in \mathbf{R}^n, \|x\|_2 = 1\}$ 上达到.

证 因为子集 S 是闭集, 且 $R(x)$ 在 S 上连续, 所以

$$\min_{x \in S} R(x) = m_1, \max_{x \in S} R(x) = m_2$$

$$\forall 0 \neq y \in \mathbf{R}^n, x = \frac{1}{\|y\|_2} y \in S \Rightarrow m_1 \leq R(x) \leq m_2$$

$$\text{由(2): } R(x) = R(y) \Rightarrow m_1 \leq R(y) \leq m_2$$

约定: 实对称矩阵 $A_{n \times n}$ 的特征值排序为 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

Th16 $A_{n \times n}$ 实对称 $\Rightarrow \min_{x \neq 0} R(x) = \lambda_1, \max_{x \neq 0} R(x) = \lambda_n$.

证 设 A 的特征值为 $\lambda_1, \dots, \lambda_n$, 特征向量 p_1, \dots, p_n 标准正交.

$$\forall x \neq 0 \Rightarrow x = c_1 p_1 + \dots + c_n p_n \quad (c_1^2 + \dots + c_n^2 \neq 0)$$

$$Ax = c_1 \lambda_1 p_1 + \dots + c_n \lambda_n p_n, \quad x^T Ax = c_1^2 \lambda_1 + \dots + c_n^2 \lambda_n$$

$$x^T x = c_1^2 + \dots + c_n^2: \quad k_i = \frac{c_i^2}{c_1^2 + \dots + c_n^2} \Rightarrow k_1 + \dots + k_n = 1$$

$$\left. \begin{aligned} R(x) &= k_1 \lambda_1 + \dots + k_n \lambda_n \Rightarrow \lambda_1 \leq R(x) \leq \lambda_n \\ R(p_1) &= \lambda_1, \quad R(p_n) = \lambda_n \end{aligned} \right\} \Rightarrow \begin{cases} \min_{x \neq 0} R(x) = \lambda_1 \\ \max_{x \neq 0} R(x) = \lambda_n \end{cases}$$

[注] ① 在闭集 S 上, p_1 是 $R(x)$ 的极小点, p_n 是 $R(x)$ 的极大点.

② p_1, \dots, p_n 构成 \mathbf{R}^n 的标准正交基:

$$\mathbf{R}^n = L(p_1, p_n) + L(p_2, \dots, p_{n-1}), \quad L(p_2, \dots, p_{n-1}) = [L(p_1, p_n)]^\perp$$

$$[L(p_1, p_n)]^\perp = \{x \mid x \perp p_1, x \perp p_n\} = \left\{x \mid \begin{pmatrix} p_1^T \\ p_n^T \end{pmatrix} x = 0\right\}$$

Th16 的进一步讨论: 设 $0 \neq x \in L(p_2, \dots, p_{n-1})$, 则有

$$x = c_2 p_2 + \dots + c_{n-1} p_{n-1} \quad (c_2^2 + \dots + c_{n-1}^2 \neq 0)$$

$$x^T A x = c_2^2 \lambda_2 + \dots + c_{n-1}^2 \lambda_{n-1}$$

$$x^T x = c_2^2 + \dots + c_{n-1}^2 : \quad k_i = \frac{c_i^2}{c_2^2 + \dots + c_{n-1}^2} \Rightarrow k_2 + \dots + k_{n-1} = 1$$

$$\left. \begin{aligned} R(x) &= k_2 \lambda_2 + \dots + k_{n-1} \lambda_{n-1} \Rightarrow \lambda_2 \leq R(x) \leq \lambda_{n-1} \\ R(p_2) &= \lambda_2, \quad R(p_{n-1}) = \lambda_{n-1} \end{aligned} \right\} \Rightarrow \begin{cases} \min_{x \neq 0} R(x) = \lambda_2 \\ \max_{x \neq 0} R(x) = \lambda_{n-1} \end{cases}$$

Th17 设 $A_{n \times n}$ 实对称, 特征向量 p_1, \dots, p_n 标准正交, 对于 $1 \leq r \leq s \leq n$,

当 $x \in L(p_r, p_{r+1}, \dots, p_s)$ 时, 有 $\min_{x \neq 0} R(x) = \lambda_r, \max_{x \neq 0} R(x) = \lambda_s$.

Th18 设 $A_{n \times n}$ 实对称, $\forall V_k \subset \mathbf{R}^n$ 且 $\dim V_k = k$, 则 A 的第 k 个特征值

$$\text{为 } \lambda_k = \min_{V_k} [\max \{R(x) \mid 0 \neq x \in V_k\}].$$

证 A 的特征向量: p_1, \dots, p_n 标准正交

(1) 左 \leq 右:

$$W_k = L(p_k, \dots, p_n) \subset \mathbf{R}^n \Rightarrow \dim W_k = n - k + 1$$

$$V_k \subset \mathbf{R}^n \Rightarrow V_k + W_k \subset \mathbf{R}^n \quad (\text{注意 } \dim V_k = k)$$

$$\begin{aligned} \text{因为 } n &\geq \dim(V_k + W_k) = \dim V_k + \dim W_k - \dim(V_k \cap W_k) \\ &= n + 1 - \dim(V_k \cap W_k) \end{aligned}$$

所以 $\dim(V_k \cap W_k) \geq 1 \Rightarrow \exists x_0 \in V_k \cap W_k$ 且 $x_0 \neq 0$. 由 $x_0 \in W_k$ 可得

$$x_0 = c_k p_k + \dots + c_n p_n \quad (c_k^2 + \dots + c_n^2 \neq 0)$$

$$\text{故 } R(x_0) = \frac{x_0^T A x_0}{x_0^T x_0} = \frac{c_k^2 \lambda_k + \dots + c_n^2 \lambda_n}{c_k^2 + \dots + c_n^2} \geq \lambda_k$$

$$\max \{R(x) \mid x \in V_k\} \geq R(x_0) \geq \lambda_k \Rightarrow \text{右} \geq \text{左}$$

(2) 右 \leq 左: 令 $V_k^0 = L(p_1, \dots, p_k)$, 对 $\forall x \in V_k^0 (x \neq 0)$, 有

$$x = l_1 p_1 + \dots + l_k p_k \quad (l_1^2 + \dots + l_k^2 \neq 0)$$

$$R(x) = \frac{x^T A x}{x^T x} = \frac{l_1^2 \lambda_1 + \dots + l_k^2 \lambda_k}{l_1^2 + \dots + l_k^2} \leq \lambda_k$$

$$\max \{R(x) \mid x \in V_k^0\} \leq \lambda_k \Rightarrow \text{右} \leq \text{左}$$

极大值集合
中的最小值

二、广义 Rayleigh 商:

问题: 若 x 是 $Ax = \lambda Bx$ 的广义特征向量, 则有 $\lambda = \frac{x^T Ax}{x^T Bx}$.

下面讨论多元函数 $R_B(x) = \frac{x^T Ax}{x^T Bx}$ ($x \neq 0$) 的极值问题.

- (1) $R_B(x)$ 连续;
- (2) \forall 实数 $k \neq 0, R_B(kx) = R_B(x)$;
- (3) $\forall x_0 \neq 0$, 当 $0 \neq x \in L(x_0)$ 时, $R_B(x) = R_B(x_0)$;
- (4) $\min_{x \neq 0} R_B(x)$ 与 $\max_{x \neq 0} R_B(x)$ 存在, 且在 $S_B = \{x \mid x \in \mathbb{R}^n, \|x\|_B = 1\}$ 上达到.

Th22 $x_0 \neq 0$ 是 $R_B(x)$ 的驻点 $\Leftrightarrow Ax_0 = \lambda Bx_0$.

证 $(x^T Bx) \cdot R_B(x) = (x^T Ax)$, $2Bx \cdot R_B(x) + (x^T Bx) \cdot \frac{dR_B(x)}{dx} = 2Ax$

$$\frac{dR_B(x)}{dx} = \frac{2}{x^T Bx} \cdot [Ax - R_B(x) \cdot Bx]$$

必要性. $x_0 \neq 0$ 使 $\left. \frac{dR_B}{dx} \right|_{x=x_0} = 0$: $Ax_0 = R_B(x_0) \cdot Bx_0$ $\left(\lambda \stackrel{\Delta}{=} R_B(x_0) \right)$

充分性. $x_0 \neq 0$ 使 $Ax_0 = \lambda Bx_0$: $\lambda = R_B(x_0)$ 且有 $\left. \frac{dR_B}{dx} \right|_{x=x_0} = 0$

约定: $Ax = \lambda Bx$ 的特征值排序为 $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$

广义特征向量 p_1, p_2, \cdots, p_n “按 B -标准正交”

Th23 $\forall V_k \subset \mathbb{R}^n$ 且 $\dim V_k = k \Rightarrow \lambda_k = \min_{V_k} [\max \{R_B(x) \mid 0 \neq x \in V_k\}]$

证 类似于 Th18.

Th23' $\forall V_k \subset \mathbb{R}^n$ 且 $\dim V_k = k \Rightarrow \lambda_{n-k+1} = \max_{V_k} [\min \{R_B(x) \mid 0 \neq x \in V_k\}]$

证 $(-A)x = (-\lambda)Bx$ 的特征值为 $-\lambda_n \leq -\lambda_{n-1} \leq \cdots \leq -\lambda_1$, 由 Th23 可得

$$\begin{aligned} -\lambda_{n-k+1} &= \min_{V_k} [\max \{-R_B(x) \mid 0 \neq x \in V_k\}] & \frac{x^T (-A)x}{x^T Bx} &= -R_B(x) \\ &= \min_{V_k} [(-1) \cdot \min \{R_B(x) \mid 0 \neq x \in V_k\}] \\ &= (-1) \max_{V_k} [\min \{R_B(x) \mid 0 \neq x \in V_k\}] \end{aligned}$$

极小值集合
中的最大值

特殊情形:

$$\text{Th23} \Rightarrow \lambda_n = \max \{R_B(x) \mid 0 \neq x \in \mathbb{R}^n\} = \max \{x^T Ax \mid x \in \mathbb{R}^n \text{ 且 } \|x\|_B = 1\}$$

$$\text{Th23'} \Rightarrow \lambda_1 = \min \{R_B(x) \mid 0 \neq x \in \mathbb{R}^n\} = \min \{x^T Ax \mid x \in \mathbb{R}^n \text{ 且 } \|x\|_B = 1\}$$

三、矩阵奇异值的极性

$$A \in \mathbb{R}_r^{m \times n}: \sigma(A) = [\lambda(A^T A)]^{\frac{1}{2}}$$

$$A^T A \text{ 的特征值: } 0 = \lambda_1 = \cdots = \lambda_{n-r} < \lambda_{n-r+1} \leq \cdots \leq \lambda_n$$

A 的奇异值: $0 = \sigma_1 = \cdots = \sigma_{n-r} < \sigma_{n-r+1} \leq \cdots \leq \sigma_n$

Th24 $\forall V_k \subset \mathbb{R}^n$ 且 $\dim V_k = k$, 则有

$$\sigma_k = \min_{V_k} \left\{ \max_{0 \neq x \in V_k} \frac{\|Ax\|_2}{\|x\|_2} \right\}, \quad \sigma_{n-k+1} = \max_{V_k} \left\{ \min_{0 \neq x \in V_k} \frac{\|Ax\|_2}{\|x\|_2} \right\}$$

证 对 $A^T A$ 应用 Th23 (取 $B = I$) 可得

$$\sigma_k = \lambda_k^{\frac{1}{2}} = \left[\min_{V_k} \left\{ \max_{0 \neq x \in V_k} \frac{x^T (A^T A) x}{x^T x} \right\} \right]^{\frac{1}{2}} = \left[\min_{V_k} \left\{ \max_{0 \neq x \in V_k} \left(\frac{\|Ax\|_2}{\|x\|_2} \right)^2 \right\} \right]^{\frac{1}{2}} = \text{右}$$

应用 Th23' 可得

$$\sigma_{n-k+1} = \lambda_{n-k+1}^{\frac{1}{2}} = \left[\max_{V_k} \left\{ \min_{0 \neq x \in V_k} \frac{x^T (A^T A) x}{x^T x} \right\} \right]^{\frac{1}{2}} = \left[\max_{V_k} \left\{ \min_{0 \neq x \in V_k} \left(\frac{\|Ax\|_2}{\|x\|_2} \right)^2 \right\} \right]^{\frac{1}{2}} = \text{右}$$

§ 5.4 矩阵的直积及应用

$$A = (a_{ij})_{m \times n}, \quad B = (b_{ij})_{p \times q}, \quad A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

例如 $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

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

易见 $A \otimes B \neq B \otimes A$

一、基本性质

(1) $k(A \otimes B) = (kA) \otimes B = A \otimes (kB)$ (k 为常数)

(2) A_1 与 A_2 同阶: $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$

$$B \otimes (A_1 + A_2) = B \otimes A_1 + B \otimes A_2$$

(3) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$

证 $[A \otimes B]_{ij\text{块}} = a_{ij}B = \begin{bmatrix} a_{ij}b_{11} & \cdots & a_{ij}b_{1q} \\ \vdots & & \vdots \\ a_{ij}b_{p1} & \cdots & a_{ij}b_{pq} \end{bmatrix}$

$$\begin{aligned}
 [A \otimes B]_{ij\text{块}} \otimes C &= \begin{bmatrix} a_{ij}b_{11}C & \cdots & a_{ij}b_{1q}C \\ \vdots & & \vdots \\ a_{ij}b_{p1}C & \cdots & a_{ij}b_{pq}C \end{bmatrix} = a_{ij}(B \otimes C) \\
 (A \otimes B) \otimes C &= \begin{bmatrix} [A \otimes B]_{11\text{块}} \otimes C & \cdots & [A \otimes B]_{1n\text{块}} \otimes C \\ \vdots & & \vdots \\ [A \otimes B]_{m1\text{块}} \otimes C & \cdots & [A \otimes B]_{mn\text{块}} \otimes C \end{bmatrix} \\
 &= \begin{bmatrix} a_{11}(B \otimes C) & \cdots & a_{1n}(B \otimes C) \\ \vdots & & \vdots \\ a_{m1}(B \otimes C) & \cdots & a_{mn}(B \otimes C) \end{bmatrix} = A \otimes (B \otimes C)
 \end{aligned}$$

(4) 设 $A_1 = (a_{ij}^{(1)})_{m \times n}$, $A_2 = (a_{ij}^{(2)})_{n \times l}$, $B_1 = (b_{ij}^{(1)})_{p \times q}$, $B_2 = (b_{ij}^{(2)})_{q \times r}$, 则

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$$

证 $A_1 \otimes B_1 = \begin{bmatrix} \cdots & \cdots & \cdots \\ a_{i1}^{(1)}B_1 & a_{i2}^{(1)}B_1 & \cdots & a_{in}^{(1)}B_1 \\ \cdots & \cdots & \cdots \end{bmatrix}$, $A_2 \otimes B_2 = \begin{bmatrix} \cdots & a_{1j}^{(2)}B_2 & \cdots \\ \cdots & a_{2j}^{(2)}B_2 & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & a_{nj}^{(2)}B_2 & \cdots \end{bmatrix}$

$$[\text{左}]_{ij\text{块}} = \sum_{k=1}^n (a_{ik}^{(1)}B_1)(a_{kj}^{(2)}B_2) = \sum_{k=1}^n a_{ik}^{(1)}a_{kj}^{(2)} \cdot (B_1 B_2) = (A_1 A_2)_{ij} (B_1 B_2) = [\text{右}]_{ij\text{块}}$$

(5) $A_{m \times m}$ 与 $B_{n \times n}$ 都可逆 $\Rightarrow (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$

验证: $(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1}) \otimes (BB^{-1}) = I_m \otimes I_n = I_{(mn)}$

(6) $A_{m \times m}$ 与 $B_{n \times n}$ 都是上(下)三角矩阵 $\Rightarrow A \otimes B$ 也是上(下)三角矩阵

(7) $(A \otimes B)^T = A^T \otimes B^T$, $(A \otimes B)^H = A^H \otimes B^H$

例如: 左端 $= \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix}^H = \begin{bmatrix} \bar{a}_{11}B^H & \bar{a}_{21}B^H \\ \bar{a}_{12}B^H & \bar{a}_{22}B^H \end{bmatrix} = \text{右端}$

(8) $A_{m \times m}$ 与 $B_{n \times n}$ 都是(酉)正交矩阵 $\Rightarrow A \otimes B$ 也是(酉)正交矩阵

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

证 记 $r_1 = \text{rank}A$, $r_2 = \text{rank}B$

$$\text{存在满秩矩阵 } P_1, Q_1, \text{ 使得 } P_1 A Q_1 = \begin{bmatrix} I_{r_1} & O \\ O & O \end{bmatrix} \triangleq A_1$$

$$\text{存在满秩矩阵 } P_2, Q_2, \text{ 使得 } P_2 B Q_2 = \begin{bmatrix} I_{r_2} & O \\ O & O \end{bmatrix} \triangleq B_1$$

$$(P_1 \otimes P_2)(A \otimes B)(Q_1 \otimes Q_2) = \cdots = A_1 \otimes B_1$$

因为 $P_1 \otimes P_2$ 与 $Q_1 \otimes Q_2$ 都可逆, 所以

$$\text{rank}(A \otimes B) = \text{rank}(A_1 \otimes B_1) = "A_1 \otimes B_1 \text{ 中 } 1 \text{ 的个数}" = r_1 r_2$$

直积矩阵的特征值: 二元多项式函数 $f(x, y) = \sum_{i=0}^{l_1} \sum_{j=0}^{l_2} c_{ij} x^i y^j$

对于矩阵 $A_{m \times m}$ 和 $B_{n \times n}$, 定义矩阵函数 $f(A, B) = \sum_{i=0}^{l_1} \sum_{j=0}^{l_2} c_{ij} A^i \otimes B^j$

Th27 设 $A_{m \times m}$ 的特征值为 $\lambda_1, \dots, \lambda_m$, $B_{n \times n}$ 的特征值为 μ_1, \dots, μ_n , 则 $f(A, B)$ 的特征值为 $f(\lambda_i, \mu_j)$ ($i = 1, \dots, m, j = 1, \dots, n$).

证 对于 $A_{m \times m}$ 和 $B_{n \times n}$, 存在可逆矩阵 $P_{m \times m}$ 与 $Q_{n \times n}$, 使得

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_m \end{bmatrix} \triangleq T_1, \quad Q^{-1}BQ = \begin{bmatrix} \mu_1 & \cdots & * \\ & \ddots & \vdots \\ & & \mu_n \end{bmatrix} \triangleq T_2$$

因为 $P \otimes Q$ 可逆, 所以

$$(P \otimes Q)^{-1}(A^i \otimes B^j)(P \otimes Q) = \cdots = T_1^i \otimes T_2^j \text{ 是上三角矩阵}$$

$$(P \otimes Q)^{-1} \cdot f(A, B) \cdot (P \otimes Q) = \cdots = f(T_1, T_2) \text{ 也是上三角矩阵}$$

又 $T_1^i \otimes T_2^j = \begin{bmatrix} \lambda_1^i T_2^j & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_m^i T_2^j \end{bmatrix}, \quad \lambda_k^i T_2^j = \begin{bmatrix} \lambda_k^i \mu_1^j & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_k^i \mu_n^j \end{bmatrix}$

故 $f(T_1, T_2)$ 的主对角线元素为 $f(\lambda_k, \mu_s)$ ($k = 1, \dots, m, s = 1, \dots, n$).

由此可得: $f(A, B)$ 的特征值为 $f(\lambda_i, \mu_j)$ ($i = 1, \dots, m, j = 1, \dots, n$).

推论 1 $A_{m \times m}, B_{n \times n}$, $A \otimes B$ 的特征值为 $\lambda_i \mu_j$ ($i = 1, \dots, m, j = 1, \dots, n$).

取 $f(x, y) = x^1 y^1$, 则 $f(A, B) = A \otimes B$.

推论 2 $A_{m \times m}, B_{n \times n}$: $|A \otimes B| = |A|^n |B|^m$.

$$\text{左} = \prod_{i=1}^m \prod_{j=1}^n \lambda_i \mu_j = \prod_{i=1}^m \left(\lambda_i^n \cdot \prod_{j=1}^n \mu_j \right) = \prod_{i=1}^m \lambda_i^n \cdot \left(\prod_{j=1}^n \mu_j \right)^m = \text{右}$$

二、线性矩阵方程的可解性

$$A = (a_{ij})_{m \times p}, B_{q \times n}, X_{p \times q} = \begin{bmatrix} x_1^T \\ \vdots \\ x_p^T \end{bmatrix}, \overline{\text{vec}}(X) \triangleq \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} = \bar{X} \quad (\text{后者为简写记号})$$

$$AXB = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} x_1^T \\ \vdots \\ x_p^T \end{bmatrix} B = \begin{bmatrix} (a_{11}x_1^T + \cdots + a_{1p}x_p^T)B \\ \vdots \\ (a_{m1}x_1^T + \cdots + a_{mp}x_p^T)B \end{bmatrix}$$

$$\overrightarrow{AXB} = \begin{bmatrix} B^T(a_{11}x_1 + \cdots + a_{1p}x_p) \\ \vdots \\ B^T(a_{m1}x_1 + \cdots + a_{mp}x_p) \end{bmatrix} = \begin{bmatrix} a_{11}B^T & \cdots & a_{1p}B^T \\ \vdots & & \vdots \\ a_{m1}B^T & \cdots & a_{mp}B^T \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$$

$$= (A \otimes B^T) \vec{X}$$

性质: $\overline{\text{vec}(X+Y)} = \overline{\text{vec}(X)} + \overline{\text{vec}(Y)}$, $\overline{\text{vec}(kX)} = k \cdot \overline{\text{vec}(X)}$ (k 为常数)

复习: 划分 $A = (a_1, a_2, \dots, a_n)$, 则有

$$(1) R(A) = \text{span}\{a_1, a_2, \dots, a_n\};$$

$$(2) Ax = b \text{ 有解} \Leftrightarrow (a_1, \dots, a_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b \text{ 有解} \Leftrightarrow b \in R(A).$$

Th28 $A_i \in \mathbb{C}^{m \times p}$, $B_i \in \mathbb{C}^{q \times n}$, $F \in \mathbb{C}^{m \times n}$:

$$\sum_{i=1}^k A_i X B_i = F \text{ 有解 } X_{p \times q} \Leftrightarrow \vec{F} \in R\left(\sum_{i=1}^k A_i \otimes B_i^T\right)$$

证 左 $\Leftrightarrow \sum A_i X B_i = \vec{F}$ 有解 $X \Leftrightarrow \sum (A_i \otimes B_i^T) \vec{X} = \vec{F}$ 有解 \vec{X}
 $\Leftrightarrow \left(\sum A_i \otimes B_i^T\right) \vec{X} = \vec{F}$ 有解 $\vec{X} \Leftrightarrow \vec{F} \in R\left(\sum A_i \otimes B_i^T\right)$

Th29 $A_{m \times m}, B_{n \times n}, F_{m \times n}$, 矩阵方程 $AX + XB = F$ 有唯一解 $X_{m \times n} \Leftrightarrow$

$$\lambda_i(A) + \mu_j(B) \neq 0 \quad (i=1, \dots, m, j=1, \dots, n)$$

证 左 $\Leftrightarrow AXI_n + I_m XB = F$ 有唯一解 $X_{m \times n}$

$$\Leftrightarrow (A \otimes I_n + I_m \otimes B^T) \vec{X} = \vec{F} \text{ 有唯一解 } \vec{X}$$

$$\Leftrightarrow \det(A \otimes I_n + I_m \otimes B^T) \neq 0$$

令 $f(x, y) = x^1 y^0 + x^0 y^1$, 并注意 $\mu(B^T) = \mu(B)$, 由定理 27 知

$$f(A, B^T) = A \otimes I_n + I_m \otimes B^T$$

的特征值为 $f(\lambda_i, \mu_j) = \lambda_i + \mu_j$, 因此可得

$$\det(A \otimes I_n + I_m \otimes B^T) \neq 0 \Leftrightarrow \lambda_i + \mu_j \neq 0 \quad (i=1, \dots, m, j=1, \dots, n)$$

推论 $A_{m \times m}, B_{n \times n}, AX + XB = O_{m \times n}$ 有非零解 $X \Leftrightarrow \exists i_0, j_0, \text{st. } \lambda_{i_0} + \mu_{j_0} = 0$.

例 1 $A_{m \times m}, B_{n \times n}, C_{m \times n}$: $\lambda(A) \neq \mu(B) \Rightarrow \begin{bmatrix} A & C \\ O & B \end{bmatrix}$ 相似于 $\begin{bmatrix} A & O \\ O & B \end{bmatrix}$.

证 设 $P = \begin{bmatrix} I_m & X \\ O & I_n \end{bmatrix}$ ($X_{m \times n}$ 待定), 要求

$$\begin{aligned} P^{-1} \begin{bmatrix} A & C \\ O & B \end{bmatrix} P = \begin{bmatrix} A & O \\ O & B \end{bmatrix} &\Leftrightarrow \begin{bmatrix} A & C \\ O & B \end{bmatrix} \begin{bmatrix} I_m & X \\ O & I_n \end{bmatrix} = \begin{bmatrix} I_m & X \\ O & I_n \end{bmatrix} \begin{bmatrix} A & O \\ O & B \end{bmatrix} \\ &\Leftrightarrow \begin{bmatrix} A & AX+C \\ O & B \end{bmatrix} = \begin{bmatrix} A & XB \\ O & B \end{bmatrix} \\ &\Leftrightarrow AX+C = XB \Leftrightarrow (-A)X + XB = C \end{aligned}$$

因为 $\lambda(A) \neq \mu(B) \Rightarrow \lambda_i(-A) + \mu_j(B) \neq 0$

所以 $(-A)X + XB = C$ 有唯一解 $X_{m \times n}^*$

取 $P = \begin{bmatrix} I_m & X^* \\ O & I_n \end{bmatrix}$ 时, 有 $P^{-1} \begin{bmatrix} A & C \\ O & B \end{bmatrix} P = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$.

[注] $\begin{bmatrix} A & C \\ O & B \end{bmatrix}$ 相似于 $\begin{bmatrix} A & O \\ O & B \end{bmatrix}$ 的充要条件是 $\bar{C} \in R((-A) \otimes I_n + I_m \otimes B^T)$.

Th30 $A_{m \times m}$ 的特征值为 $\lambda_1, \dots, \lambda_m$, $B_{n \times n}$ 的特征值为 μ_1, \dots, μ_n .

(1) $\sum_{k=0}^l A^k XB^k = F$ 有唯一解 $\Leftrightarrow 1 + (\lambda_i \mu_j) + \dots + (\lambda_i \mu_j)^l \neq 0 \quad (\forall i, j)$;

(2) $\sum_{k=0}^l A^k XB^k = O$ 有非零解 $\Leftrightarrow \exists i_0, j_0, \text{st. } 1 + (\lambda_{i_0} \mu_{j_0}) + \dots + (\lambda_{i_0} \mu_{j_0})^l = 0$.

证 将 $\sum_{k=0}^l A^k XB^k = F$ 按行拉直可得 $\left(\sum_{k=0}^l A^k \otimes (B^T)^k \right) \bar{X} = \bar{F}$

$$f(x, y) = \sum_k x^k y^k \Rightarrow f(A, B^T) = \sum_{k=0}^l A^k \otimes (B^T)^k$$

$$f(A, B^T) \text{ 的特征值为 } f(\lambda_i, \mu_j) = \sum_{k=0}^l \lambda_i^k \mu_j^k = \sum_{k=0}^l (\lambda_i \mu_j)^k$$

故 $\sum_{k=0}^l A^k XB^k = F$ 有唯一解 $\Leftrightarrow f(\lambda_i, \mu_j) \neq 0$

$$\sum_{k=0}^l A^k XB^k = O \text{ 有非零解 } \Leftrightarrow \exists i_0, j_0, \text{st. } f(\lambda_{i_0}, \mu_{j_0}) = 0$$

三、线性矩阵方程的矩阵函数解法

引理 3 设 $A_{m \times m}, B_{n \times n}, F_{m \times n}$, 若 $\operatorname{Re}(\lambda_A) < 0, \operatorname{Re}(\mu_B) < 0$, 则广义积分

$$\int_0^{+\infty} e^{At} F e^{Bt} dt \text{ 存在.}$$

Th31 设 $A_{m \times m}, B_{n \times n}, F_{m \times n}, \lambda_A + \mu_B \neq 0$, 且 $\int_0^{+\infty} e^{At} F e^{Bt} dt$ 存在, 则

$$AX + XB = F \text{ 的唯一解 } X = -\int_0^{+\infty} e^{At} F e^{Bt} dt.$$

证 $Y(t) = e^{At} F e^{Bt} : Y(t)|_{t=0} = F, \int_0^{+\infty} Y(t) dt \text{ 存在 } \Rightarrow \lim_{t \rightarrow +\infty} Y(t) = O$

$$\frac{dY}{dt} = A \cdot Y(t) + Y(t) \cdot B \quad (1)$$

$$\text{积分 } (1): Y(t)|_0^{+\infty} = A \cdot \int_0^{+\infty} Y(t) dt + \int_0^{+\infty} Y(t) dt \cdot B$$

$$-F = A(-X) + (-X)B \Rightarrow AX + XB = F$$

推论 1 设 $A_{m \times m}, B_{n \times n}, F_{m \times n}$, $\operatorname{Re}(\lambda_A) < 0, \operatorname{Re}(\mu_B) < 0$, 则 $AX + XB = F$ 存在

$$\text{唯一解 } X = -\int_0^{+\infty} e^{At} F e^{Bt} dt.$$

推论 2 设 $A_{m \times m}, F_{m \times m}, \operatorname{Re}(\lambda_A) < 0$, 则 $A^H X + X A = -F$ 存在唯一解

$$X = \int_0^{+\infty} e^{A^H t} F e^{A t} dt; \text{ 若 } F \text{ 是正定矩阵, 则 } X \text{ 也是正定矩阵.}$$

$$(\lambda(A^H) = \overline{\lambda(A)} \Rightarrow \operatorname{Re}(\lambda(A^H)) < 0)$$

证 ② $0 \neq x \in \mathbb{C}^n: e^{A t} \text{ 可逆} \Rightarrow e^{A t} x \neq 0$

$$\begin{aligned} (e^{A t} x)^H F (e^{A t} x) &> 0 \Rightarrow x^H \cdot e^{A^H t} F e^{A t} \cdot x > 0 \\ &\Rightarrow x^H X x = \int_0^{+\infty} x^H (e^{A^H t} F e^{A t}) x dt > 0 \end{aligned}$$

四、线性矩阵微分方程的矩阵函数解法

性质 1 设 $A \in \mathbb{C}^{n \times n}$, 则 $e^{I \otimes A} = I \otimes e^A, e^{A \otimes I} = e^A \otimes I$.

$$\begin{aligned} \text{证 } e^{I \otimes A} &= I \otimes I + \frac{1}{1!} (I \otimes A) + \frac{1}{2!} (I \otimes A)^2 + \cdots \\ &= I \otimes I + \frac{1}{1!} (I \otimes A) + \frac{1}{2!} (I \otimes A^2) + \cdots \\ &= I \otimes \left(I + \frac{1}{1!} A + \frac{1}{2!} A^2 + \cdots \right) = I \otimes e^A \end{aligned}$$

性质 2 设 $A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}$, 则 $e^{(A \otimes I_n + I_m \otimes B)} = e^A \otimes e^B$.

$$\begin{aligned} \text{证 因为 } (A \otimes I_n)(I_m \otimes B) &= (A I_m) \otimes (I_n B) = A \otimes B \\ &= (I_m A) \otimes (B I_n) = (I_m \otimes B)(A \otimes I_n) \end{aligned}$$

$$\text{所以 } e^{(A \otimes I_n + I_m \otimes B)} = e^{A \otimes I_n} e^{I_m \otimes B} = (e^A \otimes I_n)(I_m \otimes e^B) = e^A \otimes e^B$$

Th32 设 $A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}, X(t) \in \mathbb{C}^{m \times n}$, 则矩阵微分方程初值问题

$$\frac{d}{dt} X(t) = A \cdot X(t) + X(t) \cdot B, \quad X(0) = X_0$$

的解为 $X(t) = e^{A t} X_0 e^{B t}$.

证 将矩阵微分方程按行拉直, 并利用第三章的结论可得

$$\begin{aligned} \frac{d}{dt} \overrightarrow{X(t)} &= (A \otimes I_n + I_m \otimes B^T) \overrightarrow{X(t)}, \quad \overrightarrow{X(0)} = \overrightarrow{X_0} \\ \overrightarrow{X(t)} &= e^{(A \otimes I_n + I_m \otimes B^T)t} \overrightarrow{X_0} = (e^{A t} \otimes e^{B^T t}) \overrightarrow{X_0} \\ X(t) &= e^{A t} X_0 (e^{B^T t})^T = e^{A t} X_0 e^{B t} \end{aligned}$$

典型题分析:

加例 设 $A \in \mathbb{C}^{n \times n}$ 的 n 个线性无关的特征向量为 x_1, x_2, \dots, x_n , 证明: $A \otimes A$ 有 n^2 个线性无关的特征向量, 并将它们构造出来.

证 设 $Ax_i = \lambda_i x_i$ ($i = 1, 2, \dots, n$), $P = [x_1 | x_2 | \dots | x_n]$, 则有

$$\begin{aligned} P^{-1}AP &= \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \stackrel{\Delta}{=} \Lambda \\ (P \otimes P)^{-1}(A \otimes A)(P \otimes P) &= (P^{-1} \otimes P^{-1})(A \otimes A)(P \otimes P) \\ &= (P^{-1}AP) \otimes (P^{-1}AP) = \Lambda \otimes \Lambda \end{aligned}$$

因为 $\Lambda \otimes \Lambda$ 是对角矩阵, 所以 $P \otimes P$ 的 n^2 个列向量是 $A \otimes A$ 的线性无关的特征向量. 根据矩阵直积的定义, 有

$$\begin{aligned} x_i \otimes P &= x_i \otimes [x_1 | \dots | x_n] = [x_i \otimes x_1 | \dots | x_i \otimes x_n] \\ P \otimes P &= [x_1 | \dots | x_n] \otimes P = [x_1 \otimes P | \dots | x_n \otimes P] \\ &= [x_1 \otimes x_1 | \dots | x_1 \otimes x_n | x_2 \otimes x_1 | \dots | x_n \otimes x_n] \end{aligned}$$

故 $x_i \otimes x_j$ ($i, j = 1, 2, \dots, n$) 是 $A \otimes A$ 的 n^2 个线性无关的特征向量.

解释:
$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes P = \begin{bmatrix} 1P \\ 2P \end{bmatrix} = \begin{bmatrix} 1x_1 & \dots & 1x_n \\ 2x_1 & \dots & 2x_n \end{bmatrix} = \left[\begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes x_1 \mid \dots \mid \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes x_n \right]$$

M2011A 一(2)

设 $A = \begin{bmatrix} & & 1 \\ & 2 & \\ & \ddots & \\ n & & \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$, 则 $\|A \otimes x\|_2 = (\sqrt{n^3})$.

分析: $\|A \otimes x\|_2^2 = \rho((A \otimes x)^H (A \otimes x)) = \rho((A^H A) \otimes n) = \rho(n(A^H A)) = n^3$

M2011B 一(3)

设 $A \in \mathbb{C}^{m \times m}$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, $B \in \mathbb{C}^{n \times n}$ 的特征值为 $\mu_1, \mu_2, \dots, \mu_n$, 则矩阵方程 $AX - XB = X$ 有非零解的充要条件是 $(\lambda_{i_0} - \mu_{j_0} = 1)$.

分析: 将 $AX - XB = X$ 按行拉直可得

$$(A \otimes I - I \otimes B^T - I \otimes I) \vec{X} = \vec{0} \quad (*)$$

令 $f(x, y) = x - y - 1$, 则 $f(A, B^T) = A \otimes I - I \otimes B^T - I \otimes I$ 的特征值为

$$f(\lambda_i, \mu_j) = \lambda_i - \mu_j - 1 \quad (i = 1, \dots, m; j = 1, \dots, n)$$

于是齐次方程组 (*) 有非零解的充要条件是

$$\det[f(A, B^T)] = 0 \Rightarrow \exists i_0, j_0, \text{ st } f(\lambda_{i_0}, \mu_{j_0}) = 0, \text{ 即 } \lambda_{i_0} - \mu_{j_0} = 1$$

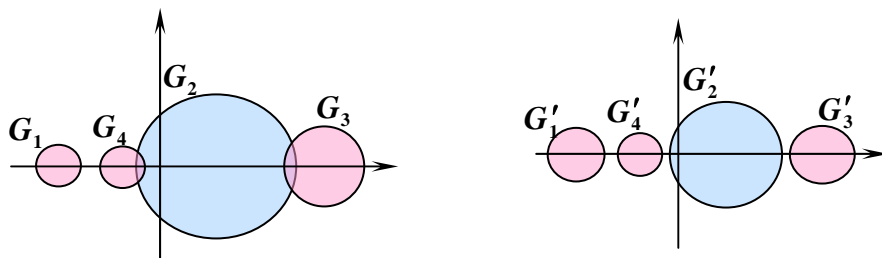
M2010(A)五 用 Gerschgorin 定理隔离矩阵 $A = \begin{bmatrix} -5 & 1 & 0 & 0 \\ 1 & 3 & 1.3 & -2 \\ 1 & 0 & 9 & 1.4 \\ 0 & 1 & 0 & -2 \end{bmatrix}$ 的特征值,

并根据实矩阵特征值的性质改进所得结果.

解 A 的 4 个盖尔圆为

$$G_1: |z+5| \leq 1; \quad G_2: |z-3| \leq 4.3; \quad G_3: |z-9| \leq 2.4; \quad G_4: |z+2| \leq 1$$

易见 G_1 孤立, 而 G_2, G_3, G_4 相交.



方案一 (放大 G_1):

$$D = \begin{bmatrix} 5/3 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \quad B = DAD^{-1} = \begin{bmatrix} -5 & 5/3 & 0 & 0 \\ 0.6 & 3 & 1.3 & -2 \\ 0.6 & 0 & 9 & 1.4 \\ 0 & 1 & 0 & -2 \end{bmatrix}$$

B 的 4 个孤立盖尔圆为

$$G'_1: |z+5| \leq \frac{5}{3}; \quad G'_2: |z-3| \leq 3.9; \quad G'_3: |z-9| \leq 2; \quad G'_4: |z+2| \leq 1$$

其中各含 B 的一个特征值. G_1, G'_2, G'_3, G'_4 中各含 A 的一个特征值.

注: 可取 $d_1 = 1.6 \sim 1.9$

因为实矩阵的复特征值一定成对共轭出现, 所以孤立盖尔圆中的特征值是实数 (在实轴上). 因此, A 的特征值所在的 4 个闭区间为

$$[-6, -4], \quad [-3, -1], \quad [-0.9, 6.9], \quad [7, 11]$$

方案二 (缩小 G_2): $D = \text{diag}(1, d_2, 1, 1)$, 可取 $d_2 = 0.7 \sim \frac{5}{6} = 0.8\bar{3}$

第六章 广义逆矩阵

§ 6.1 广义逆矩阵的概念与性质

1. 定义: 对 $A_{m \times n}$, 若有 $X_{n \times m}$ 满足 Penrose 方程

$$(1) AXA = A$$

$$(2) XAX = X$$

注意序号!

$$(3) (AX)^H = AX$$

$$(4) (XA)^H = XA$$

称 X 为 A 的 M-P 逆, 记作 A^+ . (Moore 1920, Penrose 1955)

例如 $A_{n \times n}$ 可逆, $X = A^{-1}$ 满足 P-方程: $A^+ = A^{-1}$

$A = O_{m \times n}$, $X = O_{n \times m}$ 满足 P-方程: $O_{m \times n}^+ = O_{n \times m}$

$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, X = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \text{ 满足 P-方程: } A^+ = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

例 1 $F \in C_r^{m \times r} (r \geq 1) \Rightarrow F^+ = (F^H F)^{-1} F^H$, 且 $F^+ F = I_r$;

$G \in C_r^{r \times n} (r \geq 1) \Rightarrow G^+ = G^H (GG^H)^{-1}$, 且 $GG^+ = I_r$.

证 验证第一式: 令 $X = (F^H F)^{-1} F^H$, 则有

$$FXF = F \cdot (F^H F)^{-1} F^H F = F, \quad XFX = (F^H F)^{-1} F^H F \cdot X = X$$

$$(FX)^H = X^H F^H = F (F^H F)^{-1} F^H = FX, \quad (XF)^H = I_r^H = I_r = XF$$

Th1 任意 $A_{m \times n}$, A^+ 存在且唯一.

证 存在性. $A = O_{m \times n} \Rightarrow A^+ = O_{n \times m}$

$$A \neq O \Rightarrow \text{rank} A \geq 1: A = FG, F \in C_r^{m \times r}, G \in C_r^{r \times n}$$

令 $X = G^+ F^+$, 则有

$$AXA = FG \cdot G^+ F^+ \cdot FG = FG = A$$

$$XAX = G^+ F^+ \cdot FG \cdot G^+ F^+ = G^+ F^+ = X$$

$$(AX)^H = (FG \cdot G^+ F^+)^H = (FF^+)^H = FF^+ = F \cdot GG^+ \cdot F^+ = AX$$

$$(XA)^H = (G^+ F^+ \cdot FG)^H = (G^+ G)^H = G^+ G = G^+ \cdot F^+ F \cdot G = XA$$

故 $A^+ = G^+ F^+ = G^H (F^H A G^H)^{-1} F^H$. (注意: 后者计算复杂)

唯一性. 对 $A_{m \times n}$, 若 $X_{n \times m}$ 与 $Y_{n \times m}$ 都满足 P-方程, 则

$$X = XAX = X \cdot AYA \cdot X = X \cdot (AY)^H \cdot (AX)^H$$

$$= X \cdot (AXAY)^H = X \cdot (AY)^H = XAY = X \cdot AYA \cdot Y$$

$$= (XA)^H \cdot (YA)^H \cdot Y = (YAXA)^H \cdot Y = (YA)^H \cdot Y = YAY = Y$$

Th2 设 $A \in C_r^{m \times n}$ 的奇异值分解为 $A = U \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix}_{m \times n} V^H$, 则

$$A^+ = V \begin{bmatrix} \Sigma_r^{-1} & O \\ O & O \end{bmatrix}_{n \times m} U^H. \quad (\text{直接验证})$$

广义逆矩阵的分类: 对 $A_{m \times n}$, 若 $X_{n \times m}$ 满足 **P-方程**

(i): 称 X 为 A 的 $\{i\}$ -逆, 记作 $A^{(i)}$. 全体记作 $A\{i\}$.

(i),(j): 称 X 为 A 的 $\{i,j\}$ -逆, 记作 $A^{(i,j)}$. 全体记作 $A\{i,j\}$.

(i),(j),(k): 称 X 为 A 的 $\{i,j,k\}$ -逆, 记作 $A^{(i,j,k)}$. 全体记作 $A\{i,j,k\}$.

(1)~(4): 则 X 为 A^+ .

合计: 15 类

常用广义逆矩阵: $A\{1\}, A\{1,2\}, A\{1,3\}, A\{1,4\}$ 及 A^+ .

2. 求 $A^{(1)}$ 与 $A^{(1,2)}$ 的初等变换方法

置换矩阵: 划分单位矩阵 $I_n = (e_1, e_2, \dots, e_n)$, 称 $P = (e_{j_1}, e_{j_2}, \dots, e_{j_n})$ 为置换矩阵, 其中 $j_1 j_2 \dots j_n$ 是 $1, 2, \dots, n$ 的一个排列.

重要性质: 划分 $A_{m \times n} = (a_1, a_2, \dots, a_n)$, 那么 $AP = (a_{j_1}, a_{j_2}, \dots, a_{j_n})$.

$A \in C_r^{m \times n}$, $A \xrightarrow{\text{行}} B \Rightarrow$ 存在可逆矩阵 $Q_{m \times m}$, 使得 $QA = B$

其中 B 为 **拟 Hermite 标准形**, 它的后 $m-r$ 行元素全为零.

$B \xrightarrow{\text{列对换}} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m \times n} = C \Rightarrow$ 存在置换矩阵 $P_{n \times n}$, 使得 $BP = C$

于是 $QAP = C \Rightarrow A = Q^{-1}CP^{-1}$.

Th3 已知 A, P, Q 如上所述, 对任意 $L_{(n-r) \times (m-r)}$, 有

$$X = P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}_{n \times m} Q \in A\{1\}, \quad X_0 = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}_{n \times m} Q \in A\{1,2\}$$

$$\begin{aligned} \text{证} \quad AXA &= Q^{-1}CP^{-1} \cdot P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}_{n \times m} Q \cdot Q^{-1}CP^{-1} \\ &= Q^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m \times n} \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}_{n \times m} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m \times n} P^{-1} \\ &= Q^{-1} \begin{bmatrix} I_r & KL \\ O & O \end{bmatrix}_{m \times m} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m \times n} P^{-1} = Q^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m \times n} P^{-1} = A \end{aligned}$$

故 $X \in A\{1\}$; 显然 $AX_0A = A$, 且有

$$X_0AX_0 = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q \cdot Q^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1} \cdot P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} Q = \dots = X_0$$

故 $X_0 \in A\{1,2\}$.

例 2 $A = \begin{bmatrix} 2 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 2 & 1 & 1 & 4 \end{bmatrix}$, 求 $A^{(1)}, A^{(1,2)}$ 及 A^+ .

解 $(A | I) \xrightarrow{\text{行}} \left[\begin{array}{cccc|ccc} 2 & 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{array} \right] : c_1 = 2, c_2 = 3$

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, P = (e_2, e_3, e_1, e_4) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$$A^{(1)} = P \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & b & 0 \end{array} \right] Q = \begin{bmatrix} -a & -a & a \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b & -b & b \end{bmatrix}, A^{(1,2)} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$A = FG : F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, F^T F = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, F^+ = \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} 2 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}, GG^T = \begin{bmatrix} 9 & 4 \\ 4 & 5 \end{bmatrix}, G^+ = \frac{1}{29} \begin{bmatrix} 10 & -8 \\ 5 & -4 \\ -4 & 9 \\ 2 & 10 \end{bmatrix}$$

$$A^+ = G^+ F^+ = \frac{1}{87} \begin{bmatrix} 28 & -26 & 2 \\ 14 & -13 & 1 \\ -17 & 22 & 5 \\ -6 & 18 & 12 \end{bmatrix}$$

例 3 设 $A_{m \times n} \neq O$, 且 A^+ 已知, 记 $B = \begin{pmatrix} A \\ A \end{pmatrix}$, 求 B^+ .

解 $\text{rank} A = r \geq 1 \Rightarrow A = FG : F \in C_r^{m \times r}, G \in C_r^{r \times n}$

$$B = \begin{pmatrix} FG \\ FG \end{pmatrix} = \begin{pmatrix} F \\ F \end{pmatrix} G : \begin{pmatrix} F \\ F \end{pmatrix} \in C_r^{2m \times r}, G \in C_r^{r \times n}$$

$$\begin{aligned} B^+ &= G^+ \begin{pmatrix} F \\ F \end{pmatrix}^+ = G^+ \cdot \left[\begin{pmatrix} F^H & F^H \end{pmatrix} \begin{pmatrix} F \\ F \end{pmatrix} \right]^{-1} \begin{pmatrix} F^H & F^H \end{pmatrix} \\ &= G^+ \cdot \frac{1}{2} (F^H F)^{-1} \cdot \begin{pmatrix} F^H & F^H \end{pmatrix} = \frac{1}{2} (G^+ F^+ | G^+ F^+) = \frac{1}{2} (A^+ | A^+) \end{aligned}$$

其它方法
见 Th10(7)

3. 广义逆矩阵的性质

Th4 $A_{m \times n}$, $A^{(1)}$ 唯一 $\Leftrightarrow A$ 为可逆方阵, 且 $A^{(1)} = A^{-1}$.

证 设 $X_{n \times m} \in A\{1\}$:

只有可逆方阵
的 $\{1\}$ -逆唯一

(1) 划分 $X = (x_1, \dots, x_m)$, $\forall x \in N(A)$, 令 $Y = (x_1 + x, x_2, \dots, x_m)$, 则

$$AYA = A \cdot [X + (x, \theta, \dots, \theta)] \cdot A = AXA = A \Rightarrow Y \in A\{1\}$$

$$(2) \text{ 划分 } X = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \forall \beta \in N(A^H), \text{ 令 } Z = \begin{bmatrix} \alpha_1 + \beta^H \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}, \text{ 则}$$

$$\begin{aligned} AZA &= [A^H Z^H A^H]^H = [A^H \cdot (X^H + (\beta, \theta, \dots, \theta)) \cdot A^H]^H \\ &= [A^H X^H A^H]^H = AXA = A \Rightarrow Z \in A\{1\} \end{aligned}$$

必要性. 已知 $A^{(1)}$ 唯一, 由(1)和(2)可得 $N(A) = \{\theta\}, N(A^H) = \{\theta\}$

故 $n - r_A = 0, m - r_{A^H} = 0 \Rightarrow m = r_{A^H} = r_A = n \Rightarrow A$ 可逆

$$AA^{(1)}A = A \Rightarrow A^{(1)} = A^{-1}$$

充分性. 已知 A 为可逆方阵, 则有

$$AA^{(1)}A = A \Rightarrow A^{(1)} = A^{-1} \text{ 唯一}$$

右乘矩阵-
值域变小
左乘矩阵-
零空间变大

引理 1 $A_{m \times n}, B_{n \times p}, R(AB) \subset R(A), N(B) \subset N(AB).$

证 $\forall y \in R(AB) \Rightarrow y = ABx \ (x \in C^p) \Rightarrow y = A(Bx) \in R(A)$

故 $R(AB) \subset R(A)$

$$\forall x \in N(B) \Rightarrow Bx = 0 \Rightarrow ABx = 0 \Rightarrow x \in N(AB)$$

故 $N(B) \subset N(AB)$

引理 2 $A_{m \times n}, B_{n \times p}, R(AB) = R(A) \Rightarrow \exists C_{p \times n}, \text{ st. } A = ABC.$

证 划分 $A = (a_1, \dots, a_n)$, 则有

$$a_i \in R(A) = R(AB) \Rightarrow \exists x_i \in C^p, \text{ st. } a_i = ABx_i$$

$$(a_1, \dots, a_n) = AB(x_1, \dots, x_n) \Rightarrow A = ABC: C_{p \times n} = (x_1, \dots, x_n)$$

Th5 $A_{m \times n}, B_{n \times p}, \lambda \in C, \lambda^+ = \begin{cases} 1/\lambda & (\lambda \neq 0) \\ 0 & (\lambda = 0) \end{cases}.$

$$(1) [A^{(1)}]^H \in A^H\{1\}: AA^{(1)}A = A \Rightarrow A^H(A^{(1)})^H A^H = A^H$$

$$(2) \lambda^+ A^{(1)} \in (\lambda A)\{1\}: (\lambda A)(\lambda^+ A^{(1)})(\lambda A) = (\lambda \lambda^+ \lambda)(AA^{(1)}A) = \lambda A$$

$$(3) S_{m \times m} \text{ 和 } T_{n \times n} \text{ 都可逆} \Rightarrow T^{-1}A^{(1)}S^{-1} \in (SAT)\{1\}$$

$$(4) r_A \leq r_{A^{(1)}}: r_A = r_{AA^{(1)}A} \leq r_{A^{(1)}}$$

$$(5) AA^{(1)} \text{ 与 } A^{(1)}A \text{ 都是幂等矩阵, 且 } r_{AA^{(1)}} = r_A = r_{A^{(1)}A}.$$

$$\text{易见 } r_A = r_{AA^{(1)}A} \leq \begin{cases} r_{AA^{(1)}} \\ r_{A^{(1)}A} \end{cases} \leq r_A$$

$$(6) R(AA^{(1)}) = R(A): R(A) = R(AA^{(1)}A) \stackrel{\text{引理}}{\subset} R(AA^{(1)}) \stackrel{\text{引理}}{\subset} R(A)$$

$$N(A^{(1)}A) = N(A): N(A) \subset N(A^{(1)}A) \subset N(AA^{(1)}A) = N(A)$$

$$(7) \textcircled{1} r_A = n \Leftrightarrow A^{(1)}A = I_n \quad \text{“} A \text{ 列满秩”}$$

② $r_A = m \Leftrightarrow AA^{(1)} = I_m$ “A 行满秩”

证 ① 充分性. (5) $\Rightarrow r_A = r_{A^{(1)}A} = n$

必要性. (5) $\Rightarrow r_{A^{(1)}A} = r_A = n \Rightarrow A^{(1)}A$ 可逆

$$(A^{(1)}A)^2 = A^{(1)}A \Rightarrow A^{(1)}A = I_n$$

(8) ① $r_{AB} = r_A \Leftrightarrow (AB)(AB)^{(1)}A = A$

② $r_{AB} = r_B \Leftrightarrow B(AB)^{(1)}(AB) = B$

证 ① 充分性. $r_A \leq r_{AB} \leq r_A \Rightarrow r_{AB} = r_A$

必要性. 已知 $r_{AB} = r_A$, 则

引理 1 $\Rightarrow R(AB) \subset R(A)$, 故 $R(AB) = R(A)$

引理 2 $\Rightarrow \exists C$, st. $A = ABC$

$$\Rightarrow (AB)(AB)^{(1)}A = (AB)(AB)^{(1)} \cdot ABC = (AB)C = A$$

② 必要性. $r_{B^H A^H} = r_{(AB)^H} = r_{AB} = r_B = r_{B^H}$

引理 1 $\Rightarrow R(B^H A^H) \subset R(B^H) \Rightarrow R(B^H A^H) = R(B^H)$

引理 2 $\Rightarrow B^H = B^H A^H \cdot C \Rightarrow B = C^H AB$

$$\Rightarrow B(AB)^{(1)}(AB) = C^H AB \cdot (AB)^{(1)}(AB) = C^H AB = B$$

Th6 $A_{m \times n}, Y \in A\{1\}, Z \in A\{1\} \Rightarrow X \stackrel{\Delta}{=} YAZ \in A\{1,2\}$.

证 $AXA = A \cdot YAZ \cdot A = AY(AZA) = AYA = A$

$$XAX = YAZ \cdot A \cdot YAZ = Y(AZA)YAZ = Y \cdot AYA \cdot Z = YAZ = X$$

推论 $A_{m \times n}, Y \in A\{1\} \Rightarrow X \stackrel{\Delta}{=} YAY \in A\{1,2\}$.

Th7 设 $X \in A\{1\}$, 则 $r_X = r_A \Leftrightarrow X \in A\{1,2\}$.

证 必要性. 已知 $X \in A\{1\}$ 且 $r_X = r_A$, 则

$$\left. \begin{array}{l} \text{Th5(5)} \Rightarrow r_{XA} = r_A = r_X \\ \text{引理 1} \Rightarrow R(XA) \subset R(X) \end{array} \right\} \Rightarrow R(XA) = R(X)$$

引理 2 $\Rightarrow \exists Y_{n \times m}$, st. $X = XAY$

$$\Rightarrow XAX = XA \cdot XAY = X \cdot AXA \cdot Y = XAY = X$$

故 $X \in A\{1,2\}$.

充分性. 已知 $X \in A\{1,2\}$, 则

$$\left. \begin{array}{l} AXA = A \Rightarrow r_A \leq r_X \\ XAX = X \Rightarrow r_X \leq r_A \end{array} \right\} \Rightarrow r_X = r_A$$

[注] $A\{1,2\} = \{X \mid X \in A\{1\} \text{ 且 } r(X) = r(A)\}$

引理 3 $r_{A^H A} = r_A = r_{A^H} = r_{AA^H}$

Th8 $Y \stackrel{\Delta}{=} (A^H A)^{(1)} A^H \in A\{1,2,3\}, Z \stackrel{\Delta}{=} A^H (AA^H)^{(1)} \in A\{1,2,4\}$.

构造方法
判断方法

使用 $R(AA^H) = R(A)$

证 验证第一式 (科学版教材)

$$\left. \begin{array}{l} \text{引理 1: } R(A^H A) \subset R(A^H) \\ \text{引理 3: } r_{A^H A} = r_{A^H} \end{array} \right\} \Rightarrow R(A^H A) = R(A^H)$$

$$\text{引理 2: } \exists B_{n \times m}, \text{ st. } A^H = A^H A B, \text{ 即 } A = B^H A^H A.$$

$$A Y A = B^H A^H A \cdot (A^H A)^{(1)} A^H \cdot A = B^H \cdot A^H A = A$$

$$\begin{aligned} Y A Y &= (A^H A)^{(1)} A^H \cdot A \cdot (A^H A)^{(1)} A^H \\ &= (A^H A)^{(1)} \cdot (A^H A) (A^H A)^{(1)} \cdot A^H A B \\ &= (A^H A)^{(1)} \cdot A^H A \cdot B = (A^H A)^{(1)} A^H = Y \end{aligned}$$

$$A Y = A \cdot (A^H A)^{(1)} A^H = B^H A^H A \cdot (A^H A)^{(1)} \cdot A^H A B = B^H (A^H A) B$$

由此可得 $(A Y)^H = A Y$. 因此 $Y \in A\{1, 2, 3\}$.

Th9 $A^+ = A^{(1,4)} A A^{(1,3)}$.

证 $X = A^{(1,4)} A A^{(1,3)}$: Th6 $\Rightarrow X$ 满足 Penrose 方程(1)和(2).

$$A X = A A^{(1,4)} A A^{(1,3)} = A A^{(1,3)} \Rightarrow (A X)^H = (A A^{(1,3)})^H = A A^{(1,3)} = A X$$

$$X A = A^{(1,4)} A A^{(1,3)} A = A^{(1,4)} A \Rightarrow (X A)^H = (A^{(1,4)} A)^H = A^{(1,4)} A = X A$$

$$\text{故 } A^+ = X = A^{(1,4)} A A^{(1,3)}.$$

Th10 (1) $r_{A^+} = r_A$ (2) $(A^+)^+ = A$ (对 A^+ 取 $X = A$ 验证)

(3) $(A^T)^+ = (A^+)^T$ 对 A^T 取 $X = (A^+)^T$ 验证

$(A^H)^+ = (A^+)^H$ 对 A^H 取 $X = (A^+)^H$ 验证

(4) $(A^H A)^+ = A^+ (A^H)^+$ 对 $(A^H A)$ 取 $X = A^+ (A^H)^+$ 验证

$(A A^H)^+ = (A^H)^+ A^+$ 对 $(A A^H)$ 取 $X = (A^H)^+ A^+$ 验证

(5) $A^+ = (A^H A)^+ A^H = A^H (A A^H)^+$

$$A^+ = A^+ A A^+ = \begin{cases} A^+ (A A^+)^H = A^+ (A^+)^H A^H \stackrel{4(1)}{=} (A^H A)^+ A^H \\ \left[(A^+ A)^H A^+ = A^H (A^+)^H A^+ \stackrel{4(2)}{=} A^H (A A^H)^+ \right] \end{cases}$$

(6) $R(A^+) = R(A^H), N(A^+) = N(A^H).$

$$\left. \begin{array}{l} (5) \Rightarrow R(A^+) = R(A^H (A A^H)^+) \subset R(A^H) \\ (1) \Rightarrow r_{A^+} = r_A = r_{A^H} \end{array} \right\} \Rightarrow R(A^+) = R(A^H)$$

$$\left. \begin{array}{l} (5) \Rightarrow N(A^+) = N((A^H A)^+ A^H) \supset N(A^H) \\ (1) \Rightarrow r_{A^+} = r_{A^H} \Rightarrow m - r_{A^+} = m - r_{A^H} \end{array} \right\} \Rightarrow N(A^+) = N(A^H)$$

(7) $(A \otimes B)^+ = A^+ \otimes B^+$ (习题——直接验证)

§ 6.2 投影矩阵与 Moor 逆

一、投影变换: 向量空间 C^n 中, 子空间 L 与 M 满足 $C^n = L \oplus M$, 对 $\forall x \in C^n$, 分解式 $x = y + z$, $y \in L, z \in M$ 唯一. 称变换 $T_{L,M}(x) = y$ 为沿着 M 到 L 的投影变换, 称 y 为 x 沿着 M 到 L 的投影.

性质(1) $T_{L,M}$ 是线性变换.

称 y 为 x 在 L 中的分量

证 $\forall x_1 \in C^n, x_1 = y_1 + z_1, y_1 \in L, z_1 \in M$ 唯一 $\Rightarrow T_{L,M}(x_1) = y_1$

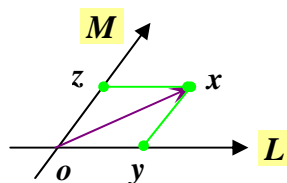
$\forall x_2 \in C^n, x_2 = y_2 + z_2, y_2 \in L, z_2 \in M$ 唯一 $\Rightarrow T_{L,M}(x_2) = y_2$

因为 $x_1 + x_2 = (y_1 + y_2) + (z_1 + z_2), y_1 + y_2 \in L, z_1 + z_2 \in M$ 唯一

所以 $T_{L,M}(x_1 + x_2) = y_1 + y_2 = T_{L,M}(x_1) + T_{L,M}(x_2)$

又 $kx_1 = (ky_1) + (kz_1), ky_1 \in L, kz_1 \in M$ 唯一

所以 $T_{L,M}(kx_1) = ky_1 = k \cdot T_{L,M}(x_1)$



性质(2) $R(T_{L,M}) = L, N(T_{L,M}) = M$.

证 (1) $\forall x \in L \subset C^n: x = x + \theta, x \in L, \theta \in M$ 唯一

故 $x = T_{L,M}(x) \in R(T_{L,M})$, 即 $L \subset R(T_{L,M})$;

$\forall x \in C^n$, 由 $T_{L,M}(x) \in L$ 可得 $R(T_{L,M}) \subset L$. 故第一式成立.

(2) $\forall x \in M \subset C^n: x = \theta + x, \theta \in L, x \in M$ 唯一

故 $T_{L,M}(x) = \theta \Rightarrow x \in N(T_{L,M})$, 即 $M \subset N(T_{L,M})$;

$\forall x \in N(T_{L,M}) \subset C^n: x = y + z, y \in L, z \in M$ 唯一

因为 $\theta = T_{L,M}(x) = y$, 所以 $x = \theta + z = z \in M$

从而 $N(T_{L,M}) \subset M$. 故第二式成立.

性质(3) $\forall x \in L \Rightarrow T_{L,M}(x) = x, \forall x \in M \Rightarrow T_{L,M}(x) = \theta$.

证 $\forall x \in L \subset C^n: x = x + \theta, x \in L, \theta \in M$ 唯一 $\Rightarrow T_{L,M}(x) = x$

$\forall x \in M \subset C^n: x = \theta + x, \theta \in L, x \in M$ 唯一 $\Rightarrow T_{L,M}(x) = \theta$

[注] $R(T_{L,M})$ 和 $N(T_{L,M})$ 都是 $T_{L,M}$ 的不变子空间 (Th1-27 之前)

$L = R(T_{L,M})$ 是 $T_{L,M}$ 的不变子空间 $\Rightarrow T_{L,M}$ 是 L 中的单位变换

$M = N(T_{L,M})$ 是 $T_{L,M}$ 的不变子空间 $\Rightarrow T_{L,M}$ 是 M 中的零变换

二、投影矩阵

取线性空间 C^n 的基为 e_1, \dots, e_n 时, 元素 x 与它的坐标 “形式一致”.

称 $T_{L,M}$ 在该基下的矩阵为投影矩阵, 记作 $P_{L,M}$.

性质(4) $T_{L,M}(x) = y \Leftrightarrow P_{L,M}x = y$

$x \in L \Rightarrow T_{L,M}(x) = x \Rightarrow P_{L,M}x = x$

$$x \in M \Rightarrow T_{L,M}(x) = \theta \Rightarrow P_{L,M}x = \theta$$

预备: $R(A) = \{y \mid y = Ax, x \in C^n\}$, $N(A) = \{x \mid Ax = \theta, x \in C^n\}$

引理 4 $A_{n \times n}, A^2 = A \Rightarrow N(A) = R(I - A)$.

证 $A^2 = A \Rightarrow A(I - A) = O$. 先证 $R(I - A) \subset N(A)$:

$$\forall x \in R(I - A) \Rightarrow \exists u \in C^n, \text{st. } x = (I - A)u$$

$$Ax = A(I - A)u = \theta \Rightarrow x \in N(A)$$

再证 $N(A) \subset R(I - A)$: $\forall \alpha \in N(A) \Rightarrow A\alpha = \theta$ (无条件成立)

$$\alpha = \alpha - A\alpha = (I - A)\alpha \in R(I - A)$$

故 $N(A) = R(I - A)$.

Th11 方阵 $P = P_{L,M} \Leftrightarrow P^2 = P$.

证 必要性. $C^n = L \oplus M$

$$\forall x \in C^n, x = y + z, y \in L, z \in M \text{ 唯一} \Rightarrow P_{L,M}x = y$$

$$P_{L,M}^2 x = P_{L,M}(P_{L,M}x) = P_{L,M}y \stackrel{\text{性质 (4)}}{=} y = P_{L,M}x$$

$$P^2 x = Px \Rightarrow P^2(e_1, \dots, e_n) = P(e_1, \dots, e_n) \Rightarrow P^2 = P$$

充分性. $\forall x \in C^n \Rightarrow x = Px + (I - P)x$

$$\text{令 } y = Px \in R(P), z = (I - P)x \in R(I - P) \stackrel{\text{引 4}}{=} N(P)$$

则 $C^n = R(P) + N(P)$. 下证 $R(P) \cap N(P) = \{\theta\}$:

$$\text{对 } \forall \beta \in R(P) \cap N(P), \beta \in R(P) \Rightarrow \exists u \in C^n, \text{st. } \beta = Pu$$

$$\beta \in N(P) \Rightarrow P\beta = \theta$$

$$\text{故 } \beta = Pu = P^2u = P \cdot Pu = P\beta = \theta.$$

于是可得 $C^n = R(P) \oplus N(P)$, 从而有

$$x = y + z, y = Px \in R(P), z = (I - P)x \in N(P) \text{ 唯一}$$

因为投影变换 $T_{R(P), N(P)}$ 满足

$$T_{R(P), N(P)}(x) = y \Rightarrow P_{R(P), N(P)}x = y \quad (\forall x \in C^n)$$

$$\text{所以 } Px = P_{R(P), N(P)}x \quad (\forall x \in C^n) \Rightarrow P = P_{R(P), N(P)}$$

三、投影矩阵的确定方法

$\dim L = r$, L 的基为 x_1, \dots, x_r : $X = (x_1, \dots, x_r)$

$\dim M = n - r$, M 的基为 y_1, \dots, y_{n-r} : $Y = (y_1, \dots, y_{n-r})$

$$\left. \begin{array}{l} P_{L,M}x_i = x_i \Rightarrow P_{L,M}X = X \\ P_{L,M}y_j = \theta \Rightarrow P_{L,M}Y = O \end{array} \right\} \Rightarrow P_{L,M}(X \mid Y) = (X \mid O)$$

$$\Rightarrow P_{L,M} = (X \mid O) \cdot (X \mid Y)^{-1}$$

例 4 \mathbf{R}^2 中: $\alpha_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, L = L(\alpha_1), M = L(\alpha_2)$, 求 $P_{L,M}$.

解 $P_{L,M} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ ($\alpha_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 时 $P_{L,M} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$)

注: $P_{L,M}$ 与 L 和 M 的基的选择无关.

证 L 的基 x_1, \dots, x_r ; 另一基 $\tilde{x}_1, \dots, \tilde{x}_r$:

$$X = (x_1, \dots, x_r), \tilde{X} = (\tilde{x}_1, \dots, \tilde{x}_r) \Rightarrow \tilde{X} = XC_{r \times r}$$

M 的基 y_1, \dots, y_{n-r} ; 另一基 $\tilde{y}_1, \dots, \tilde{y}_{n-r}$:

$$Y = (y_1, \dots, y_{n-r}), \tilde{Y} = (\tilde{y}_1, \dots, \tilde{y}_{n-r}) \Rightarrow \tilde{Y} = YD_{(n-r) \times (n-r)}$$

$$\begin{aligned} (\tilde{X} \mid O) \cdot (\tilde{X} \mid \tilde{Y})^{-1} &= (XC \mid O) \cdot \left[(X \mid Y) \begin{pmatrix} C & O \\ O & D \end{pmatrix} \right]^{-1} \\ &= (XC \mid O) \cdot \begin{bmatrix} C^{-1} & O \\ O & D^{-1} \end{bmatrix} \cdot (X \mid Y)^{-1} = (X \mid O) \cdot (X \mid Y)^{-1} \end{aligned}$$

四、正交投影变换

酉空间 C^n 中, 子空间 L 给定, 取 $M = L^\perp$, 则 $C^n = L \oplus M$.

正交投影变换 $T_L = T_{L,M}$; 正交投影矩阵 $P_L = P_{L,M}$.

Th12 方阵 $P = P_L \Leftrightarrow P^2 = P, P^H = P$.

证 必要性. 已知 $P = P_L$: **Th11** $\Rightarrow P^2 = P$

$$\begin{aligned} &\left. \begin{aligned} \forall x_1 \in C^n \Rightarrow x_1 = y_1 + z_1, y_1 = Px_1 \in L, z_1 \in L^\perp \\ \forall x_2 \in C^n \Rightarrow x_2 = y_2 + z_2, y_2 = Px_2 \in L, z_2 \in L^\perp \end{aligned} \right\} \\ &x_1^H Px_2 = (y_1 + z_1)^H y_2 = y_1^H y_2 + z_1^H y_2 = y_1^H y_2 \\ &= y_1^H y_2 + y_1^H z_2 = y_1^H (y_2 + z_2) = (Px_1)^H x_2 = x_1^H P^H x_2 \\ &x_1^H (P^H - P)x_2 = 0 \Rightarrow (P^H - P)_{ij} = e_i^H (P^H - P)e_j = 0 \end{aligned}$$

故 $P^H = P$.

充分性. 已知 $P^2 = P$: **Th11** $\Rightarrow P = P_{R(P), N(P)}$

$$P^H = P: N(P) = N(P^H) = R^\perp(P) \quad (\text{Th1-35})$$

故 $P = P_{R(P)}$.

五、正交投影矩阵的确定方法

$$\left. \begin{aligned} L \text{ 的基 } x_1, \dots, x_r: X = (x_1, \dots, x_r) \\ L^\perp \text{ 的基 } y_1, \dots, y_{n-r}: Y = (y_1, \dots, y_{n-r}) \end{aligned} \right\} \Rightarrow \begin{cases} X^H Y = O \\ Y^H X = O \end{cases}$$

已求得 $P_L = P_{L, L^\perp} = (X \mid O) \cdot (X \mid Y)^{-1}$

$$\text{因为 } (X \mid Y)^H \cdot (X \mid Y) = \begin{pmatrix} X^H \\ Y^H \end{pmatrix} \cdot (X \mid Y) = \begin{bmatrix} X^H X & O \\ O & Y^H Y \end{bmatrix}$$

$$\text{所以 } (X \mid Y)^{-1} = \begin{bmatrix} (X^H X)^{-1} & O \\ O & (Y^H Y)^{-1} \end{bmatrix} \cdot \begin{bmatrix} X^H \\ Y^H \end{bmatrix} = \begin{bmatrix} (X^H X)^{-1} X^H \\ (Y^H Y)^{-1} Y^H \end{bmatrix}$$

$$\text{于是 } P_L = (X \mid O) \cdot \begin{bmatrix} (X^H X)^{-1} X^H \\ (Y^H Y)^{-1} Y^H \end{bmatrix} = X \cdot (X^H X)^{-1} \cdot X^H = XX^+$$

$$\text{例 5 向量空间 } \mathbf{R}^3 \text{ 中: } \alpha = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \beta = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, L = L(\alpha, \beta), \text{ 求 } P_L.$$

$$\text{解 } X = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}, X^T X = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, (X^T X)^{-1} = \frac{1}{6} \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$$

$$P_L = X \cdot (X^T X)^{-1} X^T = \frac{1}{6} \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & 1 \\ -2 & 1 & 5 \end{bmatrix}$$

[注] 正交投影矩阵 P_L 与子空间 L 的基的选择无关.

➤ M-P 逆的等价定义

Moore 逆: 对 $A_{m \times n}$, 若有 $X_{n \times m}$ 满足 $AX = P_{R(A)}$ 和 $XA = P_{R(X)}$, 称 X 为 A 的 Moore 逆. (P_L 表示子空间 L 上的正交投影矩阵)

Th13 M-逆与 P-逆等价.

证 (1) 设 X 是 A 的 M-逆:

$$AXA = P_{R(A)} \cdot (a_1, \dots, a_n) = (a_1, \dots, a_n) = A$$

$$XAX = P_{R(X)} \cdot (x_1, \dots, x_m) = (x_1, \dots, x_m) = X$$

$$(AX)^H = P_{R(A)}^H = P_{R(A)} = AX, (XA)^H = P_{R(X)}^H = P_{R(X)} = XA$$

故 X 是 A 的 P-逆.

(2) 设 X 是 A 的 P-逆:

$$\left. \begin{array}{l} (AX)^2 = AXAX = AX \\ (AX)^H = AX \end{array} \right\} \xRightarrow{\text{Th12}} AX = P_{R(AX)} \xRightarrow{\text{Th5}} P_{R(A)}$$

(将 X 看作 $A^{(1)}$, $R(AA^{(1)}) = R(A)$)

$$\left. \begin{array}{l} (XA)^2 = XAXA = XA \\ (XA)^H = XA \end{array} \right\} \Rightarrow XA = P_{R(XA)} = P_{R(X)}$$

(将 A 看作 $X^{(1)}$, $R(XX^{(1)}) = R(X)$)

故 X 是 A 的 M-逆.

§ 6.3 广义逆矩阵的计算方法 (略)

§ 6.4 广义逆矩阵与线性方程组的求解

考虑 $A_{m \times n}, x \in C^n, b \in C^m$, 线性方程组 $Ax = b$:

有解 (相容) 时, 求极小范数解 x_0 满足 $\|x_0\|_2 = \min_{Ax=b} \|x\|_2$;

无解 (不相容) 时, 求极小范数最小二乘解 x_0 满足 $\|x_0\|_2 = \min_{\|Ax-b\|_2=\min} \|x\|_2$.

一、矩阵的 $\{1\}$ -逆

Th26 $A_{m \times n}, B_{p \times q}, C_{m \times q}$.

(1) $AXB = C$ 有解 $\Leftrightarrow AA^{(1)}CB^{(1)}B = C$;

(2) $AXB = C$ 有解 \Rightarrow 通解 $X = A^{(1)}CB^{(1)} + Y - A^{(1)}AYBB^{(1)}$ ($\forall Y_{n \times p}$)

证 (1) 充分性. 取 $X = A^{(1)}CB^{(1)}$ 即可.

必要性. $AXB = C$: $A = AA^{(1)}A, B = BB^{(1)}B$

$$AA^{(1)}AXB BB^{(1)}B = C \Rightarrow AA^{(1)}CB^{(1)}B = C$$

(2) $A \cdot (A^{(1)}CB^{(1)} + Y - A^{(1)}AYBB^{(1)}) \cdot B = \dots = C$

设 $AXB = C$ 的一个解为 X , 则

$$\begin{aligned} X &= X + A^{(1)}CB^{(1)} - A^{(1)}CB^{(1)} \\ &= A^{(1)}CB^{(1)} + X - A^{(1)}AXB BB^{(1)} \quad (\text{取 } Y = X_{n \times p}) \end{aligned}$$

[注] $AXB = C$ 的特解为 $A^{(1)}CB^{(1)}$

$AXB = O$ 的通解为 $Y - A^{(1)}AYBB^{(1)}$ ($\forall Y_{n \times p}$)

例 1 $A_{m \times n}, AXA = A$ 的通解为 $X = A^{(1)}AA^{(1)} + Y - A^{(1)}AYA A^{(1)}$ ($\forall Y_{n \times m}$).

令 $Y - A^{(1)}AA^{(1)} = Z$

$$\begin{aligned} \text{则 } X &= A^{(1)}AA^{(1)} + (A^{(1)} + Z) - (A^{(1)}AA^{(1)}AA^{(1)} + A^{(1)}AZAA^{(1)}) \\ &= A^{(1)} + Z - A^{(1)}AZAA^{(1)} \quad (\forall Z_{n \times m}) \end{aligned}$$

故 $A\{1\} = \{A^{(1)} + Z - A^{(1)}AZAA^{(1)} \mid Z \in C^{n \times m}\}$.

Th27 $A_{m \times n}, b \in C^m$.

(B 可逆时 $B^{(1)} = B^{-1}$ 唯一)

(1) $Ax = b$ 有解 $\Leftrightarrow AA^{(1)}b = b$;

(2) $Ax = b$ 有解 \Rightarrow 通解 $x = A^{(1)}b + (I - A^{(1)}A)y$ ($\forall y \in C^n$).

[注] $Ax = b$ 的特解为 $A^{(1)}b$, $Ax = 0$ 的通解为 $(I - A^{(1)}A)y$ ($\forall y \in C^n$).

例 2 $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, b = \begin{bmatrix} l \\ k \\ k \end{bmatrix}$. (**Hermite 标准形中取 $c_1 = 1, c_2 = 3$**)

(1) 求 $A^{(1)}$; (2) l 与 k 取何值时, $Ax = b$ 有解, 并求通解.

解 置换矩阵 $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A^{(1)} = P \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

$$AA^{(1)}b = b \Rightarrow \begin{bmatrix} l \\ k \\ 0 \end{bmatrix} = \begin{bmatrix} l \\ k \\ k \end{bmatrix} \Rightarrow k = 0, l \text{ 任意}$$

$$\text{通解 } x = \begin{bmatrix} l \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} = \begin{bmatrix} l \\ 0 \\ 0 \end{bmatrix} + \eta_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad (\forall \eta_2 \in \mathbb{C})$$

Th28 给定 $A_{m \times n}, X_{n \times m}$. 若对任意 $b \in R(A)$, $x = Xb$ 都是 $Ax = b$ 的解, 则 $X \in A\{1\}$.

证 划分 $A = (a_1, \dots, a_n)$, 则有

$$\begin{aligned} a_i \in R(A) &\Rightarrow A(Xa_i) = a_i \quad (i = 1, 2, \dots, n) \\ &\Rightarrow AX(a_1, \dots, a_n) = (a_1, \dots, a_n) \Rightarrow AXA = A \Rightarrow X \in A\{1\} \end{aligned}$$

二、矩阵的 $\{1,4\}$ -逆

引理 8 $A_{m \times n}$, 给定 $A^{(1,4)}$, 则 $A\{1,4\} = S_4 \stackrel{\text{def}}{=} \{X \mid XA = A^{(1,4)}A, X \in \mathbb{C}^{n \times m}\}$.

证 (1) 先证 $S_4 \subset A\{1,4\}$: $\forall X \in S_4$

$$\left. \begin{aligned} AXA &= A \cdot A^{(1,4)}A = A \\ (XA)^H &= (A^{(1,4)}A)^H = A^{(1,4)}A = XA \end{aligned} \right\} \Rightarrow X \in A\{1,4\}$$

(2) 再证 $A\{1,4\} \subset S_4$: $\forall X \in A\{1,4\}$

$$\begin{aligned} A^{(1,4)}A &= A^{(1,4)} \cdot AXA = (A^{(1,4)}A)^H (XA)^H \\ &= (XAA^{(1,4)}A)^H = (XA)^H = XA. \quad \text{故 } X \in S_4. \end{aligned}$$

Th29 $A_{m \times n}$, 给定 $A^{(1,4)}$, 则 $A\{1,4\} = \{A^{(1,4)} + Z(I - AA^{(1,4)}) \mid Z \in \mathbb{C}^{n \times m}\}$.

证 显然, $XA = A^{(1,4)}A$ 有解 $X = A^{(1,4)}$.

$$\text{通解 } X = (A^{(1,4)}A)A^{(1,4)} + Y - YAA^{(1,4)} \quad (\forall Y_{n \times m})$$

$$\text{令 } Y - A^{(1,4)} = Z$$

$$\begin{aligned} \text{则 } X &= A^{(1,4)}AA^{(1,4)} + (A^{(1,4)} + Z) - (A^{(1,4)}AA^{(1,4)} + ZAA^{(1,4)}) \\ &= A^{(1,4)} + Z(I - AA^{(1,4)}) \quad (\forall Z_{n \times m}) \end{aligned}$$

由引理 8 即得所证.

引理 9 若方程组 $Ax = b$ 有解, 则极小范数解 $x_0 \in R(A^H)$, 且 x_0 唯一, 在 $R(A^H)$ 中只有 $Ax = b$ 的一个解.

证 (1) 先证 $x_0 \in R(A^H)$: 反证法. 若 $x_0 \notin R(A^H)$, 由 **Th1-35** 知

$$\mathbb{C}^n = R(A^H) \oplus N(A), \quad R(A^H) \perp N(A)$$

$$x_0 = y_0 + z_0, \quad y_0 \in R(A^H), \quad z_0 \in N(A), \quad \text{且 } z_0 \neq 0$$

$$\|x_0\|_2^2 = \|y_0\|_2^2 + \|z_0\|_2^2 > \|y_0\|_2^2$$

$$Ax_0 = b, Az_0 = 0 \Rightarrow Ay_0 = Ay_0 + Az_0 = Ax_0 = b$$

故 x_0 不是 $Ax = b$ 的极小范数解, 矛盾!

(2) 再证唯一性: 设 $y_0 \in R(A^H)$, 且 $Ay_0 = b$, 则

$$x_0 - y_0 \in R(A^H) = N^\perp(A)$$

$$A(x_0 - y_0) = b - b = 0 \Rightarrow x_0 - y_0 \in N(A)$$

故 $x_0 - y_0 = 0 \Rightarrow x_0 = y_0$, 即 $R(A^H)$ 中只有 $Ax = b$ 的一个解.

Th30 (1) $Ax = b$ 有解 $\Rightarrow x_0 = A^{(1,4)}b$ 是唯一极小范数解;

(2) 给定 $A_{m \times n}$, $X_{n \times m}$, 若对任意的 $b \in R(A)$, $x = Xb$ 都是 $Ax = b$ 的极小范数解, 则 $X \in A\{1,4\}$.

证(1) **Th27**: $Ax = b$ 有解 $\Rightarrow AA^{(1,4)}b = b$, 故 $x_0 = A^{(1,4)}b$ 是 $Ax = b$ 的解.
下证 $x_0 \in R(A^H)$:

$$Ax = b \text{ 有解} \Rightarrow b \in R(A) \Rightarrow \exists u \in C^n, \text{ st. } b = Au$$

$$x_0 = A^{(1,4)}b = A^{(1,4)}Au = (A^{(1,4)}A)^H u = A^H [A^{(1,4)}]^H u \in R(A^H)$$

由引理 9 的唯一性证明知, x_0 是 $Ax = b$ 的唯一极小范数解.

(2) 划分 $A = (a_1, \dots, a_n)$, 则有

$a_i \in R(A) \Rightarrow Xa_i$ 是 $Ax = a_i$ 的极小范数解 (已知)

$$\stackrel{(1)}{\Rightarrow} Xa_i = A^{(1,4)}a_i \quad (i = 1, \dots, n)$$

$$\Rightarrow X(a_1, \dots, a_n) = A^{(1,4)}(a_1, \dots, a_n)$$

$$\Rightarrow XA = A^{(1,4)}A \Rightarrow X \in S_4 = A\{1,4\} \quad (\text{引理 8})$$

三、矩阵的 $\{1,3\}$ -逆

引理 10 $A_{m \times n}$, 给定 $A^{(1,3)}$, 则 $A\{1,3\} = S_3 \stackrel{\text{def}}{=} \{X \mid AX = AA^{(1,3)}, X \in C^{n \times m}\}$.

证 类似于引理 8 的证明.

Th31 $A_{m \times n}$, 给定 $A^{(1,3)}$, 则 $A\{1,3\} = \{A^{(1,3)} + (I - A^{(1,3)}A)Z \mid Z \in C^{n \times m}\}$.

证 利用 **Th26**, 类似于 **Th29** 的证明.

Th32 $Ax = b$ 无解 $\Rightarrow \tilde{x}_0 = A^{(1,3)}b$ 是最小二乘解, 即

$$\|A\tilde{x}_0 - b\|_2 = \min \|Ax - b\|_2$$

证 $P \stackrel{\Delta}{=} AA^{(1,3)} \Rightarrow Pb = A \cdot A^{(1,3)}b \in R(A)$

$$b = Pb + (I - P)b \Rightarrow Ax - b = (Ax - Pb) + [-(I - P)b]$$

$$\text{因为 } (Ax)^H \cdot (I - P)b = x^H A^H (I - AA^{(1,3)})b$$

$$= x^H (A^H - A^H (AA^{(1,3)})^H)b = x^H (A^H - (AA^{(1,3)}A)^H)b = 0$$

$$(Pb)^H \cdot (I - P)b = b^H \cdot AA^{(1,3)}(I - AA^{(1,3)})b = 0$$

所以 $(Ax - Pb) \perp (I - P)b \Rightarrow \|Ax - b\|_2^2 = \|Ax - Pb\|_2^2 + \|(I - P)b\|_2^2$

故 $\|Ax - b\|_2 = \min \Leftrightarrow \|Ax - Pb\|_2 = \min \Leftrightarrow Ax = Pb \quad (\because Pb \in R(A))$

又 $A\tilde{x}_0 = AA^{(1,3)}b = Pb$, 即 \tilde{x}_0 是 $Ax = Pb$ 的解.

所以 \tilde{x}_0 是 $\|Ax - b\|_2 = \min$ 的解.

例 3 (1) $Ax = b$ 有解: $Ax_0 = b \Leftrightarrow A^H Ax_0 = A^H b$;

(2) $Ax = b$ 无解: $\|Ax_0 - b\|_2 = \min \|Ax - b\|_2 \Leftrightarrow A^H Ax_0 = A^H b$.

证 (1) $Ax_0 = b \Rightarrow A^H Ax_0 = A^H b$;

$$\begin{aligned} A^H Ax_0 = A^H b &\Rightarrow A^H (Ax_0 - b) = 0 \\ &\Rightarrow Ax_0 - b \in N(A^H) = R^\perp(A) \end{aligned} \quad (\text{Th1-35})$$

$$Ax = b \text{ 有解} \Rightarrow b \in R(A) \Rightarrow Ax_0 - b \in R(A)$$

$$\text{故 } Ax_0 - b = 0 \Rightarrow Ax_0 = b.$$

(2) **必要性.** Th32: x_0 是最小二乘解 $\Rightarrow Ax_0 = Pb = AA^{(1,3)}b$

$$\text{于是 } A^H Ax_0 = A^H AA^{(1,3)}b = A^H (AA^{(1,3)})^H b = (AA^{(1,3)}A)^H b = A^H b$$

充分性. 若 $A^H Ax_0 = A^H b$, 则有

$$\begin{aligned} A^H (Ax_0 - AA^{(1,3)}b) &= A^H Ax_0 - A^H (AA^{(1,3)})^H b \\ &= A^H b - (AA^{(1,3)}A)^H b = 0 \end{aligned}$$

$$\text{于是 } Ax_0 - AA^{(1,3)}b \in N(A^H) = R^\perp(A) \quad (\text{Th1-35})$$

$$\text{又 } Ax_0 - AA^{(1,3)}b = A(x_0 - A^{(1,3)}b) \in R(A)$$

$$\text{所以 } Ax_0 - AA^{(1,3)}b = 0 \Rightarrow Ax_0 = AA^{(1,3)}b = Pb$$

$$\stackrel{\text{Th32}}{\Rightarrow} \|Ax_0 - b\|_2 = \min \|Ax - b\|_2$$

Th33 $Ax = b$ 无解 $\Rightarrow x_0 = A^+b$ 是极小范数最小二乘解, 且唯一.

证 Th32: $\|Ax - b\|_2 = \min \Leftrightarrow Ax = AA^{(1,3)}b \quad (\text{有解 } x = A^{(1,3)}b)$

Th30: $Ax = AA^{(1,3)}b$ 有解 \Rightarrow

$$\text{极小范数解 } x_0 = A^{(1,4)} \cdot AA^{(1,3)}b \stackrel{\text{Th9}}{=} A^+b \text{ 唯一}$$

结论: (1) $Ax = b$ 有解 $\stackrel{\text{Th27}}{\Leftrightarrow} AA^+b = b$.

(2) $Ax = b$ 有解 $\stackrel{\text{Th27}}{\Rightarrow}$ 通解 $x = A^+b + (I - A^+A)y \quad (\forall y \in C^n)$.

(3) $Ax = b$ 有解 $\stackrel{\text{Th30}}{\Rightarrow} x_0 = A^+b$ 是极小范数解, 且唯一.

(4) $Ax = b$ 无解 $\stackrel{\text{Th33}}{\Rightarrow} x_0 = A^+b$ 是极小范数最小二乘解, 且唯一.

例 10 已知 $A = \begin{bmatrix} 1 & 1 & 0 & 2 & 2 \\ 2 & 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

1. 求 A 的满秩分解;
2. 求 A^+ ;
3. 用广义逆矩阵方法判断线性方程组 $Ax = b$ 是否有解;
4. 求线性方程组 $Ax = b$ 的极小范数解或极小范数最小二乘解 x_0 .

(要求指出所求的是哪种解).

解 1. $A \xrightarrow{\text{行}} \begin{bmatrix} 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$; $c_1 = 1, c_2 = 3$

$$F = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix}; \quad A = FG$$

$$\begin{aligned} 2. \quad F^+ &= (F^T F)^{-1} F^T = \begin{bmatrix} 6 & 7 \\ 7 & 10 \end{bmatrix}^{-1} F^T \\ &= \frac{1}{11} \begin{bmatrix} 10 & -7 \\ -7 & 6 \end{bmatrix} F^T = \frac{1}{11} \begin{bmatrix} 10 & -1 & 3 \\ -7 & 4 & -1 \end{bmatrix} \end{aligned}$$

$$G^+ = G^T (GG^T)^{-1} = G^T \begin{bmatrix} 10 & -4 \\ -4 & 3 \end{bmatrix}^{-1} = G^T \cdot \frac{1}{14} \begin{bmatrix} 3 & 4 \\ 4 & 10 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 3 & 4 \\ 3 & 4 \\ 4 & 10 \\ 2 & -2 \\ 2 & -2 \end{bmatrix}$$

$$A^+ = G^+ F^+ = \frac{1}{154} \begin{bmatrix} 2 & 13 & 5 \\ 2 & 13 & 5 \\ -30 & 36 & 2 \\ 34 & -10 & 8 \\ 34 & -10 & 8 \end{bmatrix}$$

$$3 \sim 4. \quad x_0 = A^+ b = \frac{1}{14} [3 \quad 3 \quad 4 \quad 2 \quad 2]^T$$

$AA^+b = Ax_0 = b \Rightarrow Ax = b$ 有解, 故 x_0 是 $Ax = b$ 的极小范数解.

例 11 设非零矩阵 $A_{m \times n}$ 的 M-P 逆为 A^+ , $B_{n \times n}$ 为酉矩阵, 则 $[A \mid AB]^+ = \underline{\hspace{1cm}}$.

分析: 设 A 的满秩分解为 $A = FG$, 则 $A^+ = G^+ F^+$.

满秩分解 $[A \mid AB] = F[G \mid GB] = F\tilde{G}$ (\tilde{G} 为行满秩矩阵)

$$\tilde{G}^+ = \begin{bmatrix} G^H \\ B^H G^H \end{bmatrix} \left([G \mid GB] \begin{bmatrix} G^H \\ B^H G^H \end{bmatrix} \right)^{-1} = \begin{bmatrix} G^H \\ B^H G^H \end{bmatrix} (GG^H + GBB^H G^H)^{-1}$$

$$= \begin{bmatrix} G^H \\ B^H G^H \end{bmatrix} (2GG^H)^{-1} = \begin{bmatrix} G^H \\ B^H G^H \end{bmatrix} \cdot \frac{1}{2} (GG^H)^{-1} = \frac{1}{2} \begin{bmatrix} G^+ \\ B^H G^+ \end{bmatrix}$$

$$[A \mid AB]^+ = \tilde{G}^+ F^+ = \frac{1}{2} \begin{bmatrix} G^+ \\ B^H G^+ \end{bmatrix} F^+ = \frac{1}{2} \begin{bmatrix} G^+ F^+ \\ B^{-1} G^+ F^+ \end{bmatrix} = \frac{1}{2} \begin{bmatrix} A^+ \\ B^+ A^+ \end{bmatrix}$$

例 12 利用 $(A \otimes B)^+ = A^+ \otimes B^+$ 求分块矩阵的 M-P 逆.

$$[A \mid 2A]^+ = \{(1 \mid 2) \otimes A\}^+ = (1 \mid 2)^+ \otimes A^+ = \frac{1}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes A^+ = \frac{1}{5} \begin{bmatrix} A^+ \\ 2A^+ \end{bmatrix}$$

四、矩阵方程 $AXB = D$ 的极小范数解与极小范数最小二乘解

定理 26: 矩阵方程 $AXB = D$ 有解 $\Leftrightarrow AA^+DB^+B = D$;

矩阵 A 的 F-范数等于向量 $\overline{\text{vec}}(A)$ 的 2-范数. 将矩阵方程 $AXB = D$ 按行拉直可得线性代数方程组

$$(A \otimes B^T) \overline{\text{vec}}(X) = \overline{\text{vec}}(D)$$

因为 $\overline{\text{vec}}(X_0) = (A \otimes B^T)^+ \overline{\text{vec}}(D) = (A^+ \otimes (B^+)^T) \overline{\text{vec}}(D)$

所以 $X_0 = A^+DB^+$.

定理 30: $AXB = D$ 有解, $X_0 = A^+DB^+$ 是唯一极小 F-范数解;

定理 33: $AXB = D$ 无解, $X_0 = A^+DB^+$ 是唯一极小 F-范数最小二乘解.