矩阵理论笔记

陈军成

2021 年 12 月 9 日

目录

1 线性代数基础

1.1 线性空间与子空间

定义 1.1 (线性空间) 设 V 是一个非空集合，P 是一个数域. 在集合 V 中定义加法运算：$\forall \alpha, \beta \in V$,$\exists$ 唯一的 $v = \alpha + \beta \in V$; 在集合 V 中定义数乘运算：$\forall \alpha \in V, \forall k \in P$,$\exists$ 唯一的 $\delta = k\alpha \in V$. 如果加法和数乘满足下述法则：

(1) $\alpha + \beta = \beta + \alpha$
(2) $(\alpha + \beta) + v = \alpha + (\beta + v)$
(3) $\exists\, 0 \in V, \forall\, \alpha \in V$, 有 $\alpha + 0 = \alpha$
(4) $\forall \alpha \in V, \exists \beta \in V, s.t.\ \alpha + \beta = 0$
(5) $1\alpha = \alpha$
(6) $k(l\alpha) = kl(\alpha)$
(7) $(k + l)\alpha = k\alpha + l\alpha$
(8) $k(\alpha + \beta) = k\alpha + k\beta$

则 V 称为数域 P 上的一个线性子空间.

注 线性空间同时满足加法封闭和数乘封闭.

定义 1.2 (线性空间的基和维数) 在 V 中有 n 个线性无关的向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 而 V 中任意 $n+1$ 个向量都线性相关，则称 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组基，n 就是线性空间的维数. 常记为 $\dim V = n$.

定义 1.3 (线性子空间) 如果数域 P 上的线性空间 V 的一非空子集 W 对于 V 的两种运算也构成线性空间，则称 W 是 V 的线性子空间（简称子空间）.

注 平凡空间就是都是零向量的空间，一个线性空间里面只要有一个非零向量，就是非平凡空间.

1.2 空间分解与维数定理

定义 1.4 (和空间) 设 V_1 和 V_2 是线性空间 V 的子空间，则 V_1 与 V_2 的和空间记为

$$V_1 + V_2 = \{\alpha_1 + \alpha_2 | \forall \alpha_1 \in V_1, \forall \alpha_2 \in V_2\}$$

注 $V_1 + V_2$ 是包含 V_1 和 V_2 的最小子空间.

定义 1.5 (交空间) 设 V_1 和 V_2 是线性空间 V 的子空间，则 V_1 与 V_2 的交空间记为

$$V_1 \cap V_2 = \{\alpha | \alpha \in V_1 \text{且} \alpha \in V_2\}$$

注 $V_1 \cap V_2$ 为包含于 V_1 与 V_2 的最大子空间.

定义 1.6 (生成子空间) $\alpha_1, \alpha_2, \cdots, \alpha_n \in V$, 则生成子空间 W 记为

$$\begin{aligned} W &= \{\beta | \beta = \sum_{i=1}^{r} k_i \alpha_i\} = \text{span}\{\alpha_1, \alpha_2, \cdots, \alpha_n\} \\ &= L(\alpha_1, \alpha_2, \cdots, \alpha_n) \end{aligned}$$

定理 1.1 设 V_1 和 V_2 是线性空间 V 的子空间，则

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

注 子空间和的维数一般比子空间的维数之和小.

定义 1.7 设 V_1 和 V_2 是线性空间 V 的子空间，若对 $\forall \alpha \in V_1 + V_2$，有 $\alpha = \alpha_1 + \alpha_2 (\alpha_1 \in V_1, \alpha_2 \in V_2)$ 且是唯一的，这个和 $V_1 + V_2$ 就称为直和，记为 $V_1 \oplus V_2$.

定理 1.2 设 V_1 和 V_2 是线性空间 V 的子空间，则下列命题等价

(1) $V_1 + V_2$ 是直和
(2) 零向量表示法唯一
(3) $V_1 \cap V_2 = \{0\}$
(4) $\dim(V_1 \cap V_2) = 0 \Leftrightarrow \dim(V_1) + \dim(V_1) = \dim V_1 + V_2)$

定义 1.8 设 $V_1, V_2, \cdots, V_s$ 线性空间 V 的子空间，如果和 $V_1 + V_2 + \cdots + V_s$ 中的每个向量 α 的分解式

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s, \alpha_i \in V_i (i = 1, 2, \ldots, s)$$

是唯一的，这个和 $V_1, V_2, \cdots, V_s$ 就称为直和，记为 $V_1 \oplus V_2 \oplus \cdots \oplus V_s$.

定理 1.3 设 $V_1, V_2, \cdots, V_s$ 线性空间 V 的子空间，则下列命题等价

(1) $W = V_1 + V_2 + \cdots + V_s$ 是直和
(2) 零向量表示法唯一
(3) $V_i \cap \sum\limits_{(j \neq i)} V_j = \{0\}$
(4) $\dim(W) = \sum \dim(V_i)$

1.3 特征值与特征向量

1.3.1 特征值和特征向量的概念

定义 1.9 $A \in C^{n \times n}$，如果存在 $\lambda \in C$ 和非零向量 $x \in C^n$, 使得

$$Ax = \lambda x$$

则 λ 叫 A 的特征值，x 叫 A 的属于特征值 λ 的特征向量.

注 (1) 矩阵的普：A 的所有特征值全体，叫做 A 的普，记为 $\lambda(A)$.

(2) 特征多项式：

$$|\lambda E-A|=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_i)^{n_i}\cdots(\lambda-\lambda_r)^{n_r}\left(\text{其中}\sum_{i=1}^{r}n_i=n_0\right).$$

(3) 代数重数：n_i 叫做 λ_i 的代数重数.(特征值重数.)

(4) 几何重数：$W=\{x|(\lambda_iE-A)x=0\},m_i=\dim(W)=n-rank(\lambda_iE-A)$ 叫做 λ_i 的几何重数.(特征值 λ_i 对应的线性无关的特征向量的个数，至多有 n_i 个.)

(5) $m_i<n_i$

(6) 不同特征值对应的特征向量线性无关.

定义 1.10 设 $A\in C^{n\times n}$，如果存在可逆矩阵 $P\in C^{n\times n}$, 使得

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

则称矩阵 A 叫做可对角化矩阵.

注 n 阶矩阵 A 可对角化 $\Leftrightarrow$ n 阶矩阵 A 有 n 个线性无关的特征向量.

定理 1.4 设 $A\in C^{n\times n}$，则存在可逆矩阵 $P\in C^{n\times n}$, 使得

$$P^{-1}AP=J=diag\left(J(\lambda_1),J(\lambda_2),\cdots,J(\lambda_r)\right)$$

则称 J 为 A 的 Jordan 标准型.

注 (1) Jordan 块的个数 r 是线性无关特征向量的个数，则 $r=n(m_i=n_i)\Leftrightarrow$ 矩阵 A 可对角化

(2) 由于同一特征值可以对应多个线性无关特征向量，所以 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 不一定不相同.

(3) 对应于同一特征值的 Jordan 块的个数是该特征值的几何重数, 它是相应的特征子空间的维数.

(4) 对应于同一特征值的所有 Jordan 块的阶数之和是该特征值的代数重数.

定理 1.5 设 $A\in C^{n\times n}$，则下列命题等价：

(1) 矩阵 A 可对角化

(2) C^n 存在由 A 的特征值向量构成的一组基底.

(3) A 的 Jordan 标准型中的 Jordan 块都是一阶的.

(4) $m_i=n_i(i=1,2,\cdots,r)$

1.3.2 特征值和特征向量的几何性质

(1) 变换

$$V\xrightarrow{T}V$$

$$\forall\alpha\in V,\alpha\xrightarrow{T}\alpha'\in V$$

(线性空间 V 中的任一元素 α, 都有 V 中唯一确定的元素 α' 与之对应) 则称 T 为 V 的变换.

(2) 线性变换

T 为 V 的变换且满足

$$\begin{cases} \forall \alpha, \beta \in V, T(\alpha+\beta) = T(\alpha)+T(\beta) \\ \forall k \in P, T(k\alpha) = kT(\alpha) \end{cases}$$

则称 T 为 V 的线性变换.

(3) 线性变换的特征值

定义 1.11 设 T 是线性空间 $V_n(C)$ 的一个线性变换，如果存在 $\lambda \in C$ 和非零向量 $\xi \in V_n(C)$，使得 $T\xi = \lambda\xi$, 则 λ 叫 T 的特征值，ξ 叫 T 的属于特征值 λ 的特征向量.

(4) 线性变换与矩阵

$$T(\epsilon_1, \epsilon_2, \cdots, \epsilon_n) = (T\epsilon_1, T\epsilon_2, \cdots, T\epsilon_n) = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)A$$

A 是线性变换 T 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的矩阵.

(5) 线性变换与矩阵特征值关系

$$\alpha = \sum_{i=1}^{n} x_i\epsilon_i, T\alpha = \lambda\alpha$$

$$\Rightarrow T\alpha = (T\epsilon_1, T\epsilon_2, \cdots, T\epsilon_n)x = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)Ax, \lambda\alpha = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)\lambda x$$

$$\Rightarrow Ax = \lambda x$$

(6) 线性变换在不同基下矩阵之间的关系

定理 1.6

$$\left.\begin{array}{l} T \xrightarrow{(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)} A, T \xrightarrow{(\epsilon_1', \epsilon_2', \cdots, \epsilon_n')} B \\ (\epsilon_1', \epsilon_2', \cdots, \epsilon_n') = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)C \end{array}\right. \Rightarrow B = C^{-1}AC$$

1.3.3 广义特征值问题

定义 1.12 设 $A, B \in C^{n\times n}$，如果存在 $\lambda \in C$ 和非零向量 $x \in C^n$, 使得

$$Ax = \lambda Bx$$

则 λ 叫 A 与 B 确定的广义的特征值，x 称为与 λ 的广义特征向量.

注 • $B=E \Rightarrow Ax=\lambda x(x\neq 0)$

- B 可逆，$B^{-1}Ax=\lambda x$
- 当 A,B 都是 Hermite 矩阵，即 $A=A^H, B=B^H$ 且 B 正定时，有

  $Qy=\lambda y\left(\text{其中}Q=(P^{-1})^H AP^{-1}, y=Px, B=P^HP\right)$

  $(B=B^H$ 且 B 正定 $\xrightarrow{\text{存在可逆矩阵}P} B=P^HP\ )$

  由于广义特征值是实数，则 $y_1,y_2,\cdots,y_n$ 构成标准正交基，当

$$\delta_{ij}=y_i^H y_j=(Px_i)^H(Px_j)=x_i^HP^HPx_j=x_i^HBx_j=\begin{cases}1 & ,i=j\\ 0 & ,i\neq j\end{cases}$$

  ，称 $x_1,x_2,\cdots,x_n$ 为 B 的共轭向量系.

定理 1.7 设 $n\times n$ 矩阵 $A=A^H, B=B^H$, 且 B 正定, 则 B 共轭向量系 $x_1,x_2,\cdots,x_n$ 具有以下性质:

- $x_i\neq 0(i=1,2,\cdots,n)$;
- $x_1,x_2,\cdots,x_n$ 线性无关;
- λ_i 与 x_i 满足方程 $Ax_i=\lambda_i Bx_i$;
- 若令 $X=(x_1,x_2,\cdots,x_n)$, $X^HBX=E, X^HAX=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$

1.4 欧氏空间和酉空间

1.4.1 内积空间

$$\begin{cases}V(P)\cdots\cdots P\text{上的线性空间}\\ (\bullet,\bullet)\cdots\cdots\ \text{内积 (特殊的二元函数)}\end{cases}\cdots\cdots V\text{ 为 }P\text{ 上的内积空间.}$$

内积: $V\ \rightarrow\ P$ 上的二元函数 (或 $V\times V\ \rightarrow\ P$) : $f(0,0)\ =\ (0,0)$ 满足: $\forall\alpha,\beta\in V, f(\alpha,\beta)=k\in P$

- (双线性): $\begin{cases}f\left(x,a_1y_1+a_2y_2\right)=a_1f\left(x,y_1\right)+a_2f\left(x,y_2\right)\\ f\left(a_1x_1+a_2x_2,y\right)=\bar{a}_1f\left(x_1,y\right)+\bar{a}_2f\left(x_2,y\right)\end{cases}, a_i\in P$
- (正定性及对称性): $\begin{cases}f(x,y)=\overline{f(y,x)} & \text{,对称性}\\ f(x,x)\geq 0\text{ 且}f(x,x)=0\Leftrightarrow x=0 & \text{,正定性}\end{cases}$

注 $\begin{cases}P=R: \text{内积空间称为欧氏空间}\\ P=C: \text{内积空间称为酉空间}\end{cases}.$

(1) 欧氏空间

定义 1.13 在线性空间 $V_n(R)$ 上, $\forall\alpha,\beta,\gamma\in V$, 若映射 (α,β) 满足

(a) (正定性): $(\alpha,\alpha)\geq 0; (\alpha,\alpha)=0\Leftrightarrow\alpha=0$,

(b) (齐次性): $(k\alpha,\beta)=k(\alpha,\beta)$
(c) (交换律): $(\alpha,\beta)=(\beta,\alpha)$
(d) (分配律): $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$
则映射 (α,β) 是 $V_n(R)$ 上的内积, 定义了内积的 V 为 n 维欧几里得空间, 简称欧氏空间.

(2) 酉空间

定义 1.14 在线性空间 $V_n(R)$ 上, $\forall\alpha,\beta,\gamma\in V$, 若映射 (α,β) 满足
(a) (正定性) $(\alpha,\alpha)\geq 0;(\alpha,\alpha)=0\Leftrightarrow\alpha=0$,
(b) (齐次性) $(k\alpha,\beta)=\bar{k}(\alpha,\beta)$
(c) (交换律): $(\alpha,\beta)=\overline{(\beta,\alpha)}$
(d) (分配律): $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$
则映射 (α,β) 是 $V_n(R)$ 上的内积, 定义了内积的 V 为 n 维酉空间, 简称欧氏空间.

1.4.2 欧氏 (酉) 空间的度量

1.4.3 内积的应用

1.4.4 补充: 初等矩阵

2 向量与矩阵的的范数

2.1 向量范数

2.1.1 向量范数的定义

定义 2.1 设映射 $\|\cdot\|:C^n\to R$ 满足:

(a) 正定性: $\|x\|\geq 0$, 当且仅当 $x=0$ 时, $\|x\|=0$;
(b) 齐次性: $\|\lambda x\|=|\lambda|\cdot\|x\|,\lambda\in C,x\in C^n$;
(c) 三角不等式: $\|x+y\|\leq\|x\|+\|y\|,\forall x,y\in C^n$.

则称映射 $\|\cdot\|$ 为 C^n 上向量 x 的范数.

向量范数的性质:

(a) $\|0\|=0$;
(b) $x\neq 0$ 时, $\left\|\frac{1}{\|x\|}x\right\|=1$;
(c) 对 $\forall x\in C^n$, 有 $\|-x\|=\|x\|$

(d) 对 $\forall x, y \in C^n$，有 $|\|x\| - \|y\|| \le \|x - y\|$

2.1.2 向量范数的一般形式

定理 2.1 设 $x = (x_1, x_2, \cdots, x_n)^T \in C^n$ 则 C^n 上的 Hölder 范数(p 范数)表述为

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \qquad 1 \le p < \infty$$

证明. young 不等式 $\Rightarrow$ Hölder 不等式 $\Rightarrow$ 定理2.1.

注 特别地，

(1) 向量 1 范数 ($p = 1$)：$\|x\|_1 = \sum_{i=1}^{n} |x_i|$

(2) 向量 2 范数 ($p = 2$)：$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$

(3) 向量 ∞ 范数 ($p \to \infty$)：$\|x\|_\infty = \max\limits_i |x_i|$

2.1.3 向量范数诱导向量范数

定理 2.2 设 $\|\cdot\|$ 是 C^m 上的范数, $A \in C_n^{m\times n}$, 则 $\|A\cdot\|$ 是 C^n 上的范数

证明. $\|x\|_\beta \overset{def}{=\!=} \|y\|_\alpha = \|Ax\|_\alpha$

- (正定性)$x \ne 0 \Rightarrow Ax \ne 0 \Rightarrow \|A > x\|0$
- (齐次性)$\|A(\lambda x)\| = \|\lambda Ax\| = |\lambda|\|Ax\|$
- (三角不等式)$\|A(x+y)\| = \|Ax + Ay\| \le \|Ax\| + \|Ay\|$

注 已知 $A^H = A, A$正定, $\|x\|$ 是向量范数，则 $\|x\|_A \overset{A=P^HP}{=\!=\!=} \sqrt{x^HAx} = \sqrt{(Px)^HPx} = \|Px\|_2$ 是 C^n 上的向量范数.

定理 2.3 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为线性空间 $V_n(P)$ 的一组基, $x = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\, \tilde{x}, \tilde{x} = (x_1, \cdots, x_n)^T$.

$$\|x\| \overset{def}{=\!=} \|\tilde{x}\| \text{(范数的稳定性)}$$

2.1.4 向量范数的等价性

定义 2.2 设在 $V_n(P)$ 上定义了 $\|x\|_a, \|x\|_b$ 两种向量范数，若存在常数 $C_1 > 0, C_2 > 0$, 使得

$$C_1\|x\|_a \le \|x\|_b \le C_2\|x\|_d \qquad \forall x \in V_n(P)$$

则称 $\|x\|_a$ 与 $\|x\|_b$ 等价.

定理 2.4 $V_n(P)$ 上的任意两个向量范数与等价.

2.1.5 向量范数的应用

收敛性

定理 2.5 设 $\|\cdot\|$ 是 C^n 上的任一向量范数, 则 (收敛归一化)

$$\lim_{k\to\infty} x^{(k)} = a \Leftrightarrow \lim_{k\to\infty} |x_i^{(k)} - a_i| = 0, 1 \le i \le n \Leftrightarrow \lim_{k\to\infty} \max_{1\le i\le n} \{|x_i^{(k)} - a_i|\} = 0 \Leftrightarrow \lim_{k\to\infty} \|x^{(k)} - a\| = 0$$

2.2 矩阵范数

2.2.1 矩阵范数的定义

定义 2.3 设 $A \in P^{m\times n}$, 若映射 $\|\cdot\| : P^{m\times n} \to R$ 满足:

(a) 正定性: $\|A\| \ge 0$, 当且仅当 $A = 0$ 时, $\|A\| = 0$;
(b) 齐次性: $\|\lambda A\| = |\lambda|\|A\|, \forall \lambda \in P, \forall A \in P^{m\times n}$;
(c) 三角不等式: $\|A + B\| \le \|A\| + \|B\|, \forall A, B \in P^{m\times n}$.

则称映射 $\|\cdot\|$ 为 $P^{m\times n}$ 上的矩阵范数.

矩阵范数的性质:

(a) $\|0\| = 0$;
(b) $A \ne 0$ 时, $\left\|\frac{1}{\|A\|}x\right\| = 1$;
(c) 对 $\forall A \in P$, 有 $\|-A\| = \|A\|$
(d) 对 $\forall A, B \in P^{m\times n}$, 有 $|\|A\| - \|B\|| \le \|A - B\|$

2.2.2 矩阵范数的一般形式

(1) 矩阵 m_1 范数 ($p = 1$): $\|A\|_{m_1} = \sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|$

(2) 矩阵 m_2 范数 ($p = 2$): $\|A\|_{m_2} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|^2\right)^{1/2}$

(3) 矩阵 m_∞ 范数 ($p \to \infty$): $\|A\|_{m_\infty} = \max_{i,j}\{|a_{ij}|\}, 1 \le i \le m, 1 \le j \le n$

2.2.3 矩阵范数的相容性

定义 2.4 设 $\|\cdot\|_a : P^{m\times l} \to R, \|\cdot\|_b : P^{l\times n} \to R, \|\cdot\|_c : P^{m\times n} \to R$ 是矩阵范数，如果

$$\|AB\|_c \le \|A\|_a \cdot \|B\|_b$$

则称矩阵范数 $\|\cdot\|_a, \|\cdot\|_b$ 和 $\|\cdot\|_c$ 相容. 特别地，如果

$$\|AB\| \le \|A\| \cdot \|B\|$$

则称 $\|\cdot\|$ 是自相容矩阵范数，简称 $\|\cdot\|$ 是相容的矩阵范数.

注 $\|\cdot\|_{m_1}$ 和 $\|\cdot\|_{m_2}$ 均是自相容的矩阵范数，但是 $\|\cdot\|_{m_\infty}$ 不是自相容的矩阵范数.

定理 2.6 ($\|\cdot\|_{m_2}$ 的性质) 设 $A \in C^{n\times n}$

(1) 若 $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 则

$$\|A\|_F^2 = \|A\|_{m_2}^2 = \sum_{j=1}^{n}\sum_{i=1}^{n} |a_{ij}|^2 = \sum_{j=1}^{n} \|\alpha_j\|_2^2 = \sum_{j=1}^{n} \alpha_j^H \alpha_j = \mathrm{tr}(A^H A) = \sum_{j=1}^{n} \lambda_i(A^H A)$$

(2) $A \in P^{m\times n}, U \in C^{m\times m}, U^H U = E, V \in C^{n\times n}, V^H V = E$, 则

$$\|A\|_{m_2} = \|U^H A V\|_{m_2} = \|U A V^H\|_{m_2}$$

$$\|A\|_{m_2} = \|UA\|_{m_2} = \|AV\|_{m_2} = \|UAV\|_{m_2}$$

2.3 算子范数

2.3.1 矩阵及向量范数的相容性

定义 2.5 设 $\|\cdot\|_a$ 是 P^n 上的向量范数，$\|\cdot\|_m$ 是 $P^{m\times n}$ 上的矩阵范数，且

$$\|Ax\|_a \le \|A\|_m \|x\|_a, A \in P^{m\times n}, x \in P^n$$

则称 $\|\cdot\|_m$ 为与向量范数 $\|\cdot\|_a$ 相容的矩阵范数.

注 $\|A\|_{m_1}$ 是与向量范数 $\|x\|_1$ 相容的矩阵范数, $\|A\|_{m_2}$ 是与向量范数 $\|x\|_2$ 相容的矩阵范数, 但 $\|A\|_{m_\infty}$ 不与向量范数 $\|x\|_\infty$ 相容

2.3.2 算子范数

定理 2.7 设 $\|x\|_a$ 是 P^n 上的向量范数，则

$$\|A\|_a = \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} (= \max_{\|u\|_a=1} \|Au\|_a)$$

是与向量范数 $\|x\|_a$ 相容的矩阵范数. 称此矩阵范数为从属于向量范数 $\|x\|_a$ 的算子范数

注 (1) 算子范数是自相容矩阵范数 ($\|AB\|_a \le \|A\|_a \cdot \|B\|_a$);

(2) 算子范数与向量范数 $\|x\|_a$ 相容 ($\|Ax\|_a \le \|A\|_a \cdot \|x\|_a$);

(3) $a = 1, 2, \infty$ 分别对应算子 $1, 2, \infty$ 范数;

(4) 性质:

(a) 是与向量范数 $\|x\|_a$ 相容的矩阵范数中最小的

(b) $\|\cdot\|_a$ 是算子范数 $\Rightarrow \|E\|_a = 1$, 故 $\|E\|_a \ne 1 \Rightarrow \|\cdot\|_a$ 不是算子范数

定理 2.8 设 $\|\cdot\|_m$ 是相容的矩阵范数, 则存向量范数 $\|x\|$ 使得

$$\|Ax\| \le \|A\|_m \|x\|$$

注 相容矩阵范数必存在与其相容的向量范数(即相容矩阵范数诱导的向量范数也相容).

定理 2.9 如果 $\|\cdot\|_m : C^{n\times n} \to R$ 是相容的矩阵范数, 则对 $\forall A \in C^{n\times n}$ 有

$$|\lambda_i| \le \|A\|_m (\lambda_i \text{是 } A \text{ 的特征值})$$

2.3.3 范数的谱估计

已经 $A \in C^{n\times n}$, 谱:$\lambda(A) = \{\lambda | Ax = \lambda x, \forall x \ne 0\}$ 谱半径:$r(A) = \max\limits_i |\lambda_i|, \lambda_i \in \lambda(A)$

由定理2.9知: $r(A) \le \|A\|_m$, 但存在 $\epsilon > 0$, 使得 $\|A\|_m \le r(A) + \epsilon$

2.3.4 算子范数的计算

(1) 极大列和范数 (算子 1 范数):$\|A\|_1 = \max\limits_j \left(\sum\limits_{i=1}^n |a_{ij}|\right)$

(2) 极大行和范数 (算子 ∞ 范数):$\|A\|_\infty = \max\limits_i \left(\sum\limits_{j=1}^n |a_{ij}|\right)$

(3) 谱范数 (算子 2 范数):$\|A\|_2 = \sqrt{r(A^H A)}$

证明. $\|A\|_2 = \max\limits_{\|u\|_2=1} \|Au\|_2$

谱范数不便计算, 但有好的性质 ($A \in C^{n\times n}$):

(a) $\|A\|_2 = \|A^H\|_2 = \|A^T\|_2 = \|\bar{A}\|_2$

(b) $\|AA^H\|_2 = \|AA^H\|_2 = \|A\|_2^2$

(c) 任意 n 阶酉矩阵 U 和 V 都有 (酉不变特性)

$$\|UA\|_2 = \|AV\|_2 = \|UAV\|_2 = \|A\|_2$$

(d) $\|A\|_2 = \max\limits_{\|x\|_2=\|y\|_2=1} |y^H Ax|$(柯西不等式证明)

(e) $\|A\|_2^2 \le \|A\|_1 \|A\|_\infty$

2.4 范数的应用（不太理解）

2.4.1 病态分析

若系数矩阵 A 或常数项 b 的微小变化，引起 $Ax=b$ 解的巨大变化，则称方程组为病态方程组，其系数矩阵 A 就叫做对于解方程组（或求逆）的病态矩阵. 反之，方程组就称为良态方程组，A 称为良态矩阵.

若矩阵 A 的微小变化引起特征值的巨大变化，则称矩阵 A 对求特征值来说是病态矩阵.

"病态" 是矩阵本身的特性，与所用的计算工具与计算方法无关. 但工具条件越好，病态表现相对不明显.

病态分析工具 (条件数)：$K_p(A)=\|A_p\|\|A_p^{-1}\|$(刻画矩阵的稳定性)

2.4.2 矩阵逆的摄动

定理 2.10 $A\in C^{n\times n}, \|A\|_a$ 是从属于向量范数 $\|x\|_a$ 的算子范数, 如果 $\|A\|_a<1$, 则 $E-A$ 可逆, 且

$$\left\|(E-A)^{-1}\right\|_a \leq (1-\|A\|_a)^{-1}.$$

定理 2.11 A 可逆, δA 为扰动矩阵, $\left\|A^{-1}\delta A\right\|_a<1$, 则

(1) $A+\delta A$ 可逆;
(2) $\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|_a}{\left\|A^{-1}\right\|_a} \leq \frac{\left\|A^{-1}\delta A\right\|_a}{1-\left\|A^{-1}\delta A\right\|_a}$.
(3) 若 $\left\|A^{-1}\right\|_a\|\delta A\|_a<1$, 则

$$\frac{\|A^{-1}-(A+\delta A)^{-1}\|_a}{\|A^{-1}\|_a} \leq \frac{k(A)\frac{\|\delta A\|_a}{\|A\|_a}}{1-k(A)\frac{\|\delta A\|_a}{\|A\|_a}},$$

(其中, $k(\boldsymbol{A})=\left\|\boldsymbol{A}^{-1}\right\|_a\|\boldsymbol{A}\|_a$)(条件数越大，误差上界越大)

2.4.3 线性方程组的摄动

定理 2.12 在方程组 $Ax=b$ 中, A 固定且可逆, 令 $b\neq 0$ 且有小的摄动 δb, 则解方程组

$$A\left(x_0+\delta x\right)=b+\delta b, (Ax_0=b)$$

得

$$\frac{\|(x_0+\delta x)-x_0\|}{\|x_0\|}=\frac{\|\delta x\|}{\|x_0\|} \leq K(A)\frac{\|\delta b\|}{\|b\|}$$

定理 2.13 在方程组 $Ax=b$ 中, b 固定且 $b\neq 0$ 可逆矩阵 A 有小的摄动 δA, 且 $\|A^{-1}\|\cdot\|\delta A\|<1$,

$$(A+\delta A)(x_0+\delta x)=b,(Ax_0=b)$$

得

$$\frac{\|(x_0+\delta x)-x_0\|}{\|x_0\|}=\frac{\|\delta x\|}{\|x_0\|}\leq\frac{K(A)\frac{\|\delta A\|}{\|A\|}}{1-K(A)\frac{\|\delta A\|}{\|A\|}}$$

定理 2.14 在方程组 $Ax=b$ 中, $b\neq 0$ 有小的摄动 δb, 可逆矩阵 A 有小的摄动 δA, 且 $\|A^{-1}\|\|\delta A\|<1$

$$(A+\delta A)(x_0+\delta x)=b+\delta b,(Ax_0=b)$$

得

$$\frac{\|(x_0+\delta x)-x_0\|}{\|x_0\|}=\frac{\|\delta x\|}{\|x_0\|}\leq\frac{K(A)}{r(A)}\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

其中, $r(A)=1-K(A)\frac{\|\delta A\|}{\|A\|}>0$

3 矩阵分解

3.1 矩阵的三角分解

3.1.1 基本知识

(1) (单位) 正线上 (下) 三角复 (实) 矩阵

(2) 上三角矩阵 R 的逆 R^{-1} 也是上三角矩阵, 且对角元是 R 对角元的倒数

(3) 两个上三角矩阵 R_1,R_2 的乘积也是上三角矩阵, 且对角元是 R_1 与 R_2 对角元之积;

(4) 酉矩阵 U 的逆 U^{-1} 也是酉矩阵

(5) 两个酉矩阵之积 U_1U_2 也是酉矩阵

3.1.2 n 阶矩阵的 QR 分解和上下三角分解

定理 3.1 (1) $A\in C_n^{n\times n}\Rightarrow A=U_1R=LU_2$(唯一表示), 其中 $L(R)$ 是正线下 (上) 复矩阵, U_1,U_2 为酉矩阵 $(U_1^HU_1=E,U_1\in C_n^{n\times n})$.

(2) $A\in R_n^{n\times n}\Rightarrow A=Q_1R=LQ_2$(唯一表示), 其中 $L(R)$ 是正线下 (上) 三角复矩阵, Q_1,Q_2 为正交矩阵.

(3) $A\in C_n^{n\times n},A^H=A$ 且正定 $\Rightarrow A=P^HP$(表示法不唯一, 且P可逆)

$$=R^HR(\text{表示法唯一, 且}R\text{是正线上三角复矩阵})$$

(4) $A\in R_n^{n\times n},A$ 实对称且正定 $\Rightarrow A=P^HP$(表示法不唯一, 且P可逆)

$$=R^HR(\text{表示法唯一, 且}R\text{是正线上三角实矩阵})$$

注 3 阶方阵 QR 分解：

$A=(\alpha_1,\alpha_2,\alpha_3)$ 可逆，则施密特正交化得

$$\beta_1=\alpha_1 \qquad \xi_1=\frac{\beta_1}{\|\beta_1\|}$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \qquad \xi_2=\frac{\beta_2}{\|\beta_2\|}$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2 \qquad \xi_2=\frac{\beta_3}{\|\beta_3\|}$$

$$Q=\begin{pmatrix}\xi_1 & \xi_2 & \xi_3\end{pmatrix} \quad R=\begin{pmatrix}k_{11} & k_{21} & k_{23}\\ 0 & k_{22} & k_{32}\\ 0 & 0 & k_{33}\end{pmatrix} \quad k_{ij}=\begin{cases}\|\beta_i\| & j=i\\ (\alpha_i,\xi_j) & j\neq i\end{cases}$$

定理 3.2 设 $A\in C_n^{n\times n}, L(\tilde{L})$ 为 (单位) 下三角复矩阵，$R(\tilde{R})$ 为 (单位) 上三角复矩阵，D 为对角矩阵，则下列命题等价：

(1) A 的顺序主子式均不为 0
(2) $A=L\tilde{R}$(唯一分解)，$l_{ii}\neq 0(\forall i)$
(3) $A=\tilde{L}R$(唯一分解)，$r_{ii}\neq 0(\forall i)$
(4) $A=\tilde{L}D\tilde{R}$(唯一分解)，$d_{ii}\neq 0(\forall i)$

3.1.3 任意矩阵的 QR 分解

(1) $A\in C_m^{m\times n}(m<n,$ 行满秩$)\Rightarrow A=\begin{pmatrix}L & O\end{pmatrix}V=\begin{pmatrix}L & O\end{pmatrix}\begin{pmatrix}V_1\\ V_2\end{pmatrix}$

其中，L 是正线下三角复矩阵，$V\in C^{n\times n}, V_1\in C_m^{m\times n}$

(a) A 可以唯一分解为

$$A=LU$$

其中，L 是 m 阶正线下三角矩阵，$U\in C_m^{m\times n}$

(2) $A\in C_n^{m\times n}(n<m,$ 列满秩$)\Rightarrow A=U\begin{pmatrix}R\\ O\end{pmatrix}=\begin{pmatrix}U_1 & U_2\end{pmatrix}\begin{pmatrix}R\\ O\end{pmatrix}$

其中，R 是正线上三角复矩阵，$U\in C^{m\times m}, U_1\in C_n^{m\times n}$

(a) A 可以唯一分解为

$$A=UR$$

其中，R 是 n 阶正线上三角矩阵，$U\in C_n^{m\times n}$

(3) $A\in C_r^{m\times n}$, 则存在酉矩阵 $U\in U^{m\times m}, V\in U^{n\times n}$ 及 r 阶正线下三角矩阵 L, 使得

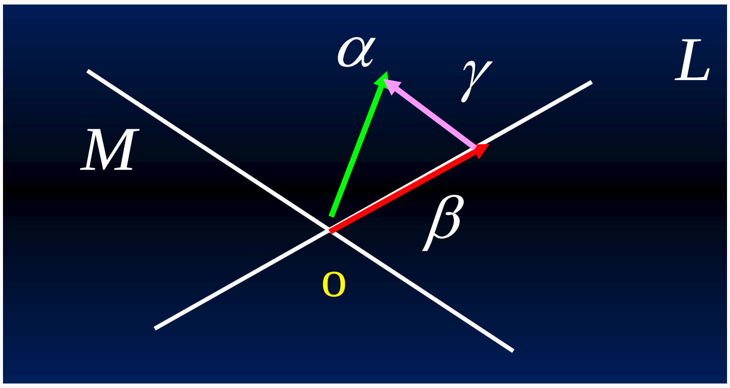
$$A=U\begin{bmatrix}L & 0\\ 0 & 0\end{bmatrix}V$$

3.2 投影

3.2.1 投影的概念

$L \oplus M = V_n(C)$ 直和 $\Leftrightarrow \forall \alpha \in V_n(C), \alpha = \beta + \gamma, \beta \in L, \gamma \in M$ 唯一.

$$P_{LM}\alpha = \beta, P_{LM}\beta = \beta \Rightarrow P_{LM}\alpha = \beta = P_{LM}\beta = P_{LM}P_{LM}\alpha \Rightarrow P_{LM} = P_{LM}^2$$



$$\forall \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} \in V, \alpha \in V, k \in C \Rightarrow \begin{cases} P_{LM}\left(\alpha_1 + \alpha_2\right) = P_{LM}\alpha_1 + P_{LM}\alpha_2 \\ P_{LM}(k\alpha) = k \bullet P_{LM}\alpha \end{cases}$$

3.2.2 投影变换与矩阵

$$\begin{aligned} &\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A = P_{LM}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) = P_{LM}^2\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) \\ &= P_{LM}\left(P_{LM}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)\right) = P_{LM}\left(\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A\right) \\ &= \left(P_{LM}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)\right) A = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A \bullet A = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A^2 \\ &\Rightarrow A = A^2 (\text{幂等矩阵}) \end{aligned}$$

3.2.3 正交投影与矩阵

$V_n(C), \alpha = \beta + \gamma, \beta \in L, \gamma \in M$ 且 $\beta \perp \gamma (\Leftrightarrow (\beta, \gamma) = 0)$, 则 β 为 α 的正交投影.

正交投影矩阵满足: $\Rightarrow A = A^2$(幂等矩阵) 且 $A = A^H$

3.2.4 幂等矩阵的性质

$A \in C^{m\times n}$, 则

$$\text{A 的值域：} R(A) = \{y|y = Ax, \forall x \in C^n\}$$

$$\text{A 的核：} N(A) = \{x|Ax = 0, \forall x \in C^n\}$$

$$A \in C^{m\times n} \Rightarrow \begin{cases} \dim R(A) + \dim N(A^H) = m \\ \dim R(A^H) + \dim N(A) = n \\ C^m = R(A) \oplus N(A^H) \\ C^n = R(A^H) \oplus N(A) \end{cases}$$

基本结论：

$$A \in C^{n\times n}, A = A^2 \Rightarrow \begin{cases} A^H, (E-A)\text{幂等矩阵} \\ A\text{的特征值非 0 即 1，且可对角化} \\ \operatorname{rank}(A) = \operatorname{tr}(A) \\ A(E-A) = (E-A)A = 0 \\ A\alpha = \alpha \Leftrightarrow \alpha \in R(A) \\ N(A) = R(E-A), R(A) = N(E-A) \end{cases}$$

3.3 矩阵的谱分解

3.3.1 单纯矩阵的谱分解

定义 3.1 若矩阵 A 的每个特征值的代数重数等于几何重数，则矩阵 A 是单纯矩阵.

定理 3.3 $A \in C^{n\times n}$ 是单纯矩阵 $\Rightarrow A = \sum\limits_{i=1}^{n} \lambda_i A_i (\Leftrightarrow A = \sum\limits_{i=1}^{k} \lambda_i A_i)$ 其中：

$$\begin{cases} \lambda_i(i = 1, 2, \cdots, n)\text{是 } A \text{ 的特征值} \\ (\lambda_i(i = 1, 2, \cdots, k)\text{是 } A \text{ 的 } k \text{ 个不同特征值}) \\ A_i\text{满足} \begin{cases} \text{幂等性：} A_i^2 = A_i(A_i = x_i y_i^T, x_i\text{是 } A_i \text{ 的特征向量}, y_i\text{是 } A_i^T \text{ 的特征向量}) \\ \text{分离性：} A_i A_j = 0 (j \neq i) \\ \text{可加性：} \sum\limits_{i=1}^{n} A_i = E_n (\sum\limits_{i=1}^{k} A_i = E_n) \end{cases} \end{cases}$$

证明. 由于 A 为单纯矩阵，则 A 有 n 个线性无光的特征向量

设 $Ax_i = \lambda_i x_i (i = 1, 2, \cdots, n)$, 其中 x_i 为特征值 λ_i 对应的特征向量. 令 $P = (x_1, x_2, \cdots, x_n), P^{-1} = \begin{pmatrix} y_1^T \\ y_2^T \\ \vdots \\ y_n^T \end{pmatrix}$, 则 $y_1^T, y_2^T, \cdots, y_n^T$ 也线性无关.

$$A = P\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)P^{-1} = \sum_{i=0}^{n} \lambda_i x_i y_i^T = \sum_{i=0}^{n} \lambda_i A_i, A_i = x_i y_i^T$$

$$P^{-1}P = E \Rightarrow y_i^T x_i = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} \Rightarrow A_i A_i = x_i y_i^T x_i y_i^T = A_i \Rightarrow A_i \text{ 为幂等矩阵}$$

注 $A^n = \sum\limits_{i=0}^{n} \lambda_i^n x_i y_i^T = \sum\limits_{i=0}^{n} \lambda_i^n A_i$

3.3.2 正规矩阵的谱分解

定义 3.2 $A \in C^{n\times n}$ 满足 $AA^H = A^H A$ 则称 A 为正规矩阵(正规矩阵不一定是 Hermite 矩阵).

几种正规矩阵:

(1) $A^H = A \Rightarrow A$ 正规
(2) $A^H = -A \Rightarrow A$ 正规
(3) $A^H = \mathrm{diag}(a_1, a_2, \cdots, a_n) \Rightarrow A$ 正规
(4) $U \in C^{n\times n}, U^H U = UU^H = E \Rightarrow U$ 正规

A 为正规矩阵，则:

(1) 存在 U 矩阵，使得 $U^H AU$ 和 $U^H A^H U$ 均为对角矩阵 (定理3.5)
(2) A 为单纯矩阵
(3) 若 $Ax = \lambda_i x (x \neq 0)$, 则 $A^H x = \bar{\lambda}_i x$
(4) A 不同特征值对应的特征向量必然正交.

证明. 设不相同特征值 λ, μ 对应得特征向量分别为 x, y，则 $Ax = \lambda x, Ay = \mu y$

$$\bar{\lambda}(x, y) = (\lambda x, y) = (Ax, y) = (x, A^H y) = (x, \bar{\mu} y) = \bar{\mu}(x, y)$$
$$\Rightarrow (\bar{\lambda} - \bar{\mu})(x, y) = 0, (\bar{\lambda} - \bar{\mu}) \neq 0 \Rightarrow (x, y) = 0$$

引理 3.4 (1) 设 A 为正规矩阵，$A = UBU^H (U \in C^{n\times n}, U^H U = E)$，则 B 为正规矩阵 (酉相似保正规).

(2) (Schur) 设 $A \in C^{n\times n}$，则存在酉矩阵 U, 使得

$$A = URU^H$$

其中，R 是一个上三角矩阵，且主对角元素为 A 的特征值

(3) A 为三角矩阵，则 A 是正规矩阵的充要条件是 A 是对角矩阵.

定理 3.5 (1) A 是正规矩阵 $\Leftrightarrow A = U\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)U^H$, 其中 $U \in C^{n\times n}, U^H U = E, \lambda_i (i = 1, 2, \cdots, n)$是 A 的特征值

(2) $A \in C^{n\times n}$ 是正规矩阵 $\Leftrightarrow A = \sum\limits_{i=1}^{k} \lambda_i A_i$. 其中：

$$\begin{cases} \lambda_i (i = 1, 2, \cdots, k)\text{是 } A \text{ 的 } k \text{ 个不同特征值} \\ A_i\text{满足} \begin{cases} \text{幂等性：} A_i^2 = A_i (A_i = x_i y_i^T, x_i\text{是 } A_i \text{ 的特征向量}, y_i\text{是 } A_i^T \text{ 的特征向量}) \\ \text{分离性：} A_i A_j = 0 (j \neq i) \\ \text{可加性：} \sum\limits_{i=1}^{k} A_i = E_n \\ A_i^H = A_i (i = 1, 2, \cdots, k) \end{cases} \end{cases}$$

证明. (a) 必要性：A 为正规矩阵，则 $A = U\mathrm{diag}(\lambda_1 E_{r_1}, \lambda_2 E_{r_2}, \cdots, \lambda_k E_{r_k})U^H$,$U$ 列分为

$$U = \begin{pmatrix} U_1 & U_2 & \cdots & U_k \end{pmatrix} \quad U^H = \begin{pmatrix} U_1^H \\ U_2^H \\ \vdots \\ U_k^H \end{pmatrix}$$

则 $A = \sum\limits_{i=1}^{k} \lambda_i U_i U_i^H = \sum\limits_{i=1}^{k} \lambda_i A_i$

$UU^H = U^H U = E \Rightarrow U_i^H U_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} \Rightarrow$ 以上结果.

(b) 充分性

$A^H A = \sum\limits_{i=1}^{k} \bar{\lambda}_i A_i^H \sum\limits_{j=1}^{k} \lambda_j A_i = \sum\limits_{i=1}^{k} |\lambda_i|^2 A_i^H A^i = \sum\limits_{i=1}^{k} |\lambda_i|^2 A_i$

$AA^H = \sum\limits_{j=1}^{k} \lambda_j A_i \sum\limits_{i=1}^{k} \bar{\lambda}_i A_i^H = \sum\limits_{i=1}^{k} |\lambda_i|^2 A_i A_i^H = \sum\limits_{i=1}^{k} |\lambda_i|^2 A_i$

故 $A^H A = AA^H \Rightarrow A$ 为 Hermite 矩阵.

3.4 矩阵最大秩分解

$$\begin{cases} A \in C_m^{m\times n} \Rightarrow AA^H \in C_m^{m\times m} \\ A \in C_n^{m\times n} \Rightarrow AA^H \in C_n^{n\times n} \end{cases}$$

定理 3.6 (最大秩分解) 设 $A \in C_r^{m\times n}$, 则存在矩阵 $B \in C_r^{m\times r}, D \in C_r^{r\times n}$, 使得

$$A = BD$$

最大秩分解步骤:

(1) 进行行初等变换，化为行标准型 $\tilde{A}$ (首非零元为 1，首非零元所在列其他元素为 0)
(2) 根据 $\tilde{A}$ 首非零元所在列的序号，在 A 中选取对应列，构成 $B_{n\times r}$
(3) 选取 $\tilde{A}$ 的前 r 行构成 $D_{r\times n}$

定理 3.7 设 $A \in C_r^{m\times n}$, 且 $A = B_1D_1 = B_2D_2$ 均为 A 的最大秩分解，则

(1) 存在 r 阶可逆矩阵 $Q(Q$ 可以取 $D_2D_1^H(D_1D_1^H)^{-1})$, 使得

$$B_1 = B_2Q \qquad D_1 = Q^{-1}D_2$$

(2)

$$\begin{aligned} &D_1(D_1D_1^H)^{-1}(B^HB_1)^{-1}B_1^H \\ =&D_2(D_2D_2^H)^{-1}(B^HB_2)^{-1}B_2^H \end{aligned} \tag{3-1}$$

注 最大秩矩阵分解形式不唯一，但由最大秩分解所作出的形式 (公式 (3-1)) 不变.

3.5 矩阵的奇异值分解

定理 3.8 设 $A \in C_r^{m\times n}$, 则有

(1) $\operatorname{rank}(A) = \operatorname{rank}(A^HA) = \operatorname{rank}(AA^H)$
(2) A^HA, AA^H 特征值均为非负实数
(3) A^HA, AA^H 的非零特征值相同

定义 3.3 (1) 设 $A \in C_r^{m\times n}$, A^HA 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$$

则称 $\sigma_i = \sqrt{\lambda_i}(i = 1, 2, \cdots, r)$ 为 A 的正奇异值，简称奇异值.

注 算子范数 $\|A\|_2 = \sqrt{r(A^HA)} = \sqrt{\lambda_1} = \sigma_1 = \max\limits_i \sigma_i$, 则 $\|AB\|_2 \leq \|A\|_2 \cdot \|B\|_2 \Rightarrow \max\limits_i \sigma_i(AB) \leq \max\limits_i \sigma_i(A) \cdot \max\limits_i \sigma_i(B)$

(2) 设 $A, B \in C^{m\times n}$, 如果存在酉矩阵 $U \in C^{m\times m}$ 和 $V \in C^{n\times n}$, 使得

$$A = UBV$$

则 A 与 B 酉等价.

定理 3.9 (1) 若 A 与 B 酉等价，则 A 与 B 有相同的正奇异值.

(酉相似，保酉等价 $\Rightarrow$ 保奇异值)

(2) 设 $A \in C_r^{m\times n}, \sigma_1, \sigma_2, \cdots, \sigma_r$ 是 A 的 r 个正奇异值，则存在酉矩阵 $U \in C^{m\times m}$ 和 $V \in C^{n\times n}$, 使得

$$A = U\begin{pmatrix} L(R) & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} V = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} V$$

其中，$D = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r), L(R)$ 是正线下 (上) 三角矩阵.

注 $\|A\|_{m_2} = \left\|\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}\right\|_{m_2} = \left(\sum\limits_{i=1}^{r} \sigma_i^2\right)^{1/2}$, 则

$\|AB\|_{m_2} \leq \|A\|_{m_2} \cdot \|B\|_{m_2} \Rightarrow \left(\sum\limits_{i=1}^{r} \sigma_i^2(AB)\right)^{1/2} \leq \left(\sum\limits_{i=1}^{r} \sigma_i^2(A)\right)^{1/2} \cdot \left(\sum\limits_{i=1}^{r} \sigma_i^2(B)\right)^{1/2}$

求解矩阵 A 奇异值步骤:

(1) 求解 D

(2) 构造酉矩阵 $U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} = E_m$, 其中 $U_1 \in C_r^{m\times r}$

(3) 构造酉矩阵 $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$, 其中 $V_1 = D^{-1}U_1^H A$, 进而 $V_2 V_1^H = 0 \Rightarrow V_2 \Rightarrow V$

4 特征值的估计

4.1 特征值的估计

4.1.1 特征值的上界估计

定理 4.1 (Schur 不等式) 设 $A \in C^{n\times n}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \sum_{i=1}^{n}\sum_{i=1}^{n} |a_{ij}|^2 = \|A\|_F^2$$

且等号成立当且仅当 A 为正规矩阵.

证明.

$$\begin{aligned}
A \in C^{n\times n} \Rightarrow A = URU^H (\text{引理3.4}) \Rightarrow \sum_{i=1}^{n} |\lambda_i|^2 &= \sum_{i=1}^{n} |r_{ii}|^2 \leq \sum_{i=1}^{n} |r_{ii}|^2 + \sum_{i=1}^{n} |r_{ij}|^2 \\
&= \|R\|_F^2 = \mathrm{tr}(R^H R) \\
&\big(X = UYU^H \Rightarrow X^H X = U(Y^H Y)U^H \Rightarrow \mathrm{tr}(X^H X) = \mathrm{tr}(X^H X)\big) \\
&= \mathrm{tr}(A^H A) = \|A\|_F^2
\end{aligned}$$

等号成立 $\Leftrightarrow \sum_{i=1}^{n}|r_{ij}|^2=0 \Leftrightarrow R$ 为对角矩阵 $\Leftrightarrow A$ 为正规矩阵 (定理3.5).

注 (1) X 与 Y 酉相似, 则

$\Rightarrow X$ 与 Y 迹相同

$\Rightarrow X^HX$ 与 Y^HY 酉相似 $\Rightarrow X$ 与 YX^HX 与 Y^HY 迹相同 $\Rightarrow \|X\|_F^2=\|Y\|_F^2$

(2) $|\lambda_i|^2 \leq \sum_{i=1}^{n}|\lambda_i|^2 \leq \|A\|_F^2 \Rightarrow \lambda_i| \leq \|A\|_F$

(3) $|\lambda_i|^2 \leq \sum_{i=1}^{n}|\lambda_i|^2 \leq \sum_{i=1}^{n}\sum_{i=1}^{n}|a_{ij}|^2 \leq n^2\max\{|a_{ij}|^2\} \Rightarrow |\lambda_i| \leq n\max\{|a_{ij}|\}$

符号约定:

(1) $B=\frac{1}{2}(A+A^H)=(b_{ij})_{n\times n}$ $C=\frac{1}{2}(A-A^H)=(c_{ij})_{n\times n}$($B^H=B, C^H=-C$)($B$ 为 Hermite 矩阵, 特征值为实数; C 为反 Hermite 矩阵, 特征值为纯虚数)

(2) A,B,C 特征值分别为 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\},\{\mu_1,\mu_2,\cdots,\mu_n\},\{i\gamma_1,i\gamma_2,\cdots,i\gamma_n\}$, 且满足 $|\lambda_1|\geq|\lambda_2|\cdots|\lambda_n|, \mu_1\geq\mu_2\cdots\mu_n, \gamma_1\geq\gamma_2\cdots\gamma_n$

定理 4.2 (Hirsh) 设 $A\in C^{n\times n}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则

(1) $|\lambda_i| \leq n\max\{|a_{ij}|\}$

(2) $|\mathrm{Re}\lambda_i| \leq n\max\{|b_{ij}|\}$

(3) $|\mathrm{Im}\lambda_i| \leq n\max\{|c_{ij}|\}$

特征值虚部模的估计可以更精确化.

定理 4.3 (Bendixsion) 设 $A\in R^{n\times n}$,A 的任一特征值 λ_i 满足:

$$|\mathrm{Im}\lambda_i| \leq \sqrt{\frac{n(n-1)}{2}}\max\{|c_{ij}|\}(c_{ii}=0)$$

4.1.2 特征值上下界的估计

定理 4.4 (1) 设 $A\in C^{n\times n}, B, C, \lambda_i, \mu_i, \gamma_i$ 同上, 则有

$$\mu_n \leq \mathrm{Re}\lambda_i \leq \mu_1$$

$$\gamma_n \leq \mathrm{Im}\lambda_i \leq \gamma_1$$

(2) (Browne) 设 $A\in C^{n\times n}$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 奇异值 $\sigma_1\geq\sigma_2\cdots\geq\sigma_n$, 则有

$$\sigma_n \leq |\lambda_i| \leq \sigma_1$$

酉矩阵任一特征值的模为 1

(3) (Hadamard 不等式) 设 $A\in C^{n\times n}$, 则

$$\prod_{i=1}^{n}|\lambda_i(A)|=|detA| \leq \left[\prod_{j=1}^{n}\left(\sum_{i=1}^{n}|a_{ij}|^2\right)\right]^{1/2}$$

4.2 圆盘定理

利用矩阵元素研究特征值分布.

4.2.1 盖尔圆盘基本定理

定义 4.1 设 $A=(a_{ij})\in C^{n\times n}$

$$\text{行盖尔圆} \Leftrightarrow S_i=\{z\in C:|z-a_{ii}|\le R_i=\sum_{j\ne i}|a_{ij}|\}$$

$$\text{列盖尔圆} \Leftrightarrow G_j=\{z\in C:|z-a_{jj}|\le C_i=\sum_{i\ne j}|a_{ij}|\}$$

定理 4.5 (1) (圆盘定理 1) 设 $A\in C^{n\times n}$，则 A 的任一特征值满足下列任一一个式子.

(a) $\lambda\in S=\bigcup\limits_{i=1}^{n}S_i$

(b) $\lambda\in G=\bigcup\limits_{j=1}^{n}G_j$

(c) $\lambda\in T=\left(\bigcup\limits_{i=1}^{n}S_i\right)\bigcap\left(\bigcup\limits_{j=1}^{n}G_j\right)$

证明. (a)

设 $A\in C^{n\times n}, Ax=\lambda x\left(x=(x_1,x_2,\cdots,x_n)^T\ne 0\right)\Rightarrow\sum\limits_{j=1}^{n}a_{ij}x_j=\lambda x_i$,

令 $|x_k|=\max\{|x_1|,|x_2|,\cdots,|x_n|\}>0$，则 $\sum\limits_{j=1}^{n}a_{kj}x_j=\lambda x_k\Rightarrow x_k(\lambda-a_{kk})=\sum\limits_{j=1,j\ne k}^{n}a_{kj}x_j\Rightarrow|x_k||\lambda-a_{kk}|\le\sum\limits_{j=1,j\ne k}^{n}|a_{kj}||x_j|\le|x_k|\sum\limits_{j=1,j\ne k}^{n}|a_{kj}|$

$\Rightarrow|\lambda-a_{kk}|\le\sum\limits_{j=1,j\ne k}^{n}|a_{kj}|=R_k\Rightarrow\lambda\in S_k\subset S=\bigcup\limits_{i=1}^{n}S_i$

(2) (圆盘定理 2) 设 $A\in C^{n\times n}$，A 的 n 个盖尔圆中有 k 个盖尔圆的并形成一个连通区域，且它与余下的 $n-k$ 个盖尔圆都不相交，则在该区域中恰好有 A 的 k 个特征值.

推论 (a) $A\in C^{n\times n}$,A 的 n 个盖尔圆两两互不相交 $\Rightarrow A$ 相似于对角矩阵 ($\Rightarrow A$ 有 n 个不同的特征值 $\Rightarrow A$ 为单纯矩阵)

(b) $A\in R^{n\times n}$,A 的 n 个盖尔圆两两互不相交 $\Rightarrow A$ 有 n 个不同的实特征值

4.2.2 特征值精细化与分离

分离某些盖尔圆 (圆心不变).

设 $A \in C^{n\times n}, B = D^{-1}AD$ A 与 B 相似, D 可逆, 且 $D = \operatorname{diag}(p_1, p_2, \cdots, p_n)(p_i > 0)$, 则有

$$R_i(B) = \frac{1}{p_i}\sum_{j=1,j\neq i}^{n}|a_{ij}|p_j, \quad Q_i = \{z \in C : |z - a_{ii}| \le R_i(B)\}$$

$$C_j(B) = p_j\sum_{i=1,i\neq j}^{n}\frac{|a_{ij}|}{p_i}, \quad P_j = \{z \in C : |z - a_{jj}| \le C_j(B)\}$$

注 若要分离第 i,j 个盖尔圆，则取 $p_i = p_j = 1 > p_l(l \neq i, j, p_l$可以取 0.1)

若要分离盖尔圆圆心，则 D 不能取对角矩阵.

定理 4.6 设 $A \in C^{n\times n}$，则 A 的任一特征值满足.

$$\lambda \in \left(\bigcup_{i=1}^{n} Q_i\right)\bigcap\left(\bigcup_{j=1}^{n} P_j\right)$$

定义 4.2 设 $A = (a_{ij}) \in C^{n\times n}$

$$\text{行 (严格) 对角占优} \Leftrightarrow |a_{ii}| \ge (>) R_i = \sum_{j\neq i}|a_{ij}|$$

$$\text{列 (严格) 对角占优} \Leftrightarrow |a_{jj}| \ge (>) C_i = \sum_{i\neq j}|a_{ij}|$$

定理 4.7 设 $A = (a_{ij}) \in C^{n\times n}$ 行 (列) 严格对角占优，则

(1) A 可逆，且 $\lambda \in \bigcup_{i=1}^{n}\tilde{S}_i \quad (\tilde{S}_i = \{z \in C : |z - a_{ii}| \le |a_{ii}|\})$

(2) 若 A 的所有对角元都为正数，则 A 的特征值位于右半平面 (对应一般矩阵特征值有正实部，Hermite 矩阵特征值为正数)

4.2.3 盖尔圆盘定理的推广

定理 4.8 (1) (Ostrowski) 设 $A = (a_{ij}) \in C^{n\times n}$，$\alpha \in [0,1]$ 为给定的数，则 A 的所有特征值位于 n 个圆盘的并集

$$\lambda \in \bigcup_{i=1}^{n}\{z \in C : |z - a_{ii}| \le R_i^{\alpha} + C_i^{1-\alpha}\}$$

(2) 设 $A = (a_{ij}) \in C^{n\times n}$，$\alpha in[0,1]$ 为给定的数，则 A 的所有特征值位于如下并集中

$$\lambda \in \bigcup_{i=1}^{n}\{z \in C : |z - a_{ii}| \le \alpha R_i + (1-\alpha)C_i\}$$

推论 设 $A=(a_{ij})\in C^{n\times n}$, 如果存在 $\alpha in[0,1]$ 使得

$$|a_{ii}|>\alpha R_i+(1-\alpha)C_i, i=1,2,\cdots,n$$

则 A 非奇异.

(3) 设 $A=(a_{ij})\in C^{n\times n}$, 则 A 的特征值位于 $\dfrac{n(n-1)}{2}$ 个 $Cassini$ 卵形域 O_{ij} 的并集中, 即

$$\lambda\in\bigcup_{j\neq i}^{n}O_{ij}=\bigcup_{j\neq i}^{n}\{z\in C:|z-a_{ii}||z-a_{jj}|\leq R_iR_j, i\neq j\}$$

4.3 Hermite 矩阵特征值的变分特征

4.3.1 Rayleigh 商的定义

定义 4.3 设 $A\in C^{n\times n}$ 为 Hermite 矩阵, $x\in C^n$ 则 A 的Rayleigh 商 表示为

$$R(x)=\frac{x^HAx}{x^Hx}\quad x\neq 0$$

注 (1) 另一种形式设 $u=\dfrac{x}{\|x\|_2}\Rightarrow x=ku(k=\|x\|_2), \|u\|_2=1, u^Hu=1$

$\Rightarrow R(x)=R(ku)=u^HAu$

(2) $R(kx)=R(x)=k^0R(x)\Rightarrow$Rayleigh 商是 0 次齐次函数 ($f(lx)=l^mf(x)$, 称 $f(x)$ 为 m 次齐次函数)

(3) $R_{A-kE}(x)=\dfrac{x^H(A-kE)x}{x^Hx}=R_A(x)-k$ 平移不变性.

(4) $x^H\left(Ax-R(x)x\right)=x^HAx-R(x)x^Hx=0\Rightarrow x\perp\left(Ax-R(x)x\right)$

4.3.2 Rayleigh 商性质

定理 4.9 (Rayleigh-Ritz) 设 $A\in C^{n\times n}$ 为 Hermite 矩阵, $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$, 则

(1) $\lambda_nx^Hx\leq x^HAx\leq\lambda_1x^Hx(\forall x\in C^n)\Rightarrow\lambda_n\leq R(x)\leq\lambda_1(x\neq 0)$

(2) $\lambda_{max}=\lambda_1=\max\limits_{x\neq 0}R(x)=\max\limits_{u^Hu=1}u^HAu$

(3) $\lambda_{mim}=\lambda_n=\min\limits_{x\neq 0}R(x)=\min\limits_{u^Hu=1}u^HAu$

证明. 内容...

定理 4.10 设 $A\in C^{n\times n}$ 为 Hermite 矩阵, $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$,$u_1,u_2,\cdots,u_n$ 为对应的标准正交特征向量, 令 $W=\text{span}\{u_s,\cdots,u_t\}(1\leq s\leq t\leq n)$(定理4.9在子空间中同样适用), 则

$$\lambda_t=\min_{0\neq x\in W}R(x)\quad\lambda_s=\max_{0\neq x\in W}R(x)(\lambda_s\leq R(x)\leq\lambda_t(0\neq x\in W))$$

定理 4.11 (Courant -Fischer) 设 $A \in C^{n\times n}$ 为 Hermite 矩阵，$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_i \geq \cdots \geq \lambda_{n-1} \geq \lambda_n$ 则

$$\lambda_i = \max_{W(\dim(W)=i)} \min_{x\in W, x\neq 0} R(x) = \max_{W(\dim(W)=i)} \min_{u\in W, \|u\|_2=1} u^H Au$$

$$\lambda_i = \max_{W(\dim(W)=n-i+1)} \min_{x\in W, x\neq 0} R(x) = \max_{W(\dim(W)=n-i+1)} \min_{u\in W, \|u\|_2=1} u^H Au$$

定理 4.12 设 $A, B \in C^{n\times n}$ 为 Hermite 矩阵，则 $\forall k = 0, 1, \cdots, n$ 有

$$\lambda_k(A) + \lambda_n(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_1(B)$$

5 矩阵分析

5.1 矩阵序列与矩阵级数

5.1.1 矩阵序列及收敛性

设 $\{A^{(k)}\}, k = 1, 2, \cdots$，是 $C^{m\times n}$ 矩阵序列，$A^{(k)} = (a_{ij})_{m\times n}$，$\|\cdot\|$ 是 $C^{m\times n}$ 上任一矩阵范数，矩阵序列 $\{A^{(k)}\}$ 收敛于 A, 有

$$\begin{aligned}\lim_{k\to\infty} A^{(k)} = A &\Leftrightarrow \lim_{k\to\infty} a_{ij}^{(k)} = a_{ij} \quad (1 \leq i \leq m, 1 < j < n) \\ &\Leftrightarrow \lim_{k\to\infty} |a_{ij}^{(k)} - a_{ij}| = 0 \quad (1 \leq i \leq m, 1 < j < n) \\ &\Leftrightarrow \lim_{k\to\infty} \|A^{(k)} - A\| = 0\end{aligned}$$

矩阵序列满足的性质:

定理 5.1 设 $\lim_{k\to\infty} A^{(k)} = A$, $\lim_{k\to\infty} B^{(k)} = B, \alpha, \beta \in C$, 则

(1) $\lim_{k\to\infty} (\alpha A^{(k)} + \beta B^{(k)}) = \alpha A + \beta B$

(2) $\lim_{k\to\infty} (A^{(k)} B^{(k)}) = AB$

(3) 当 $A^{(k)}$ 和 A 都可逆时，$\lim_{k\to\infty} (A^{(k)})^{-1} = A^{-1}$

方阵的幂构成的矩阵序列 (重要)

定理 5.2 设 $A \in C^{n\times n}$, 则 $\lim_{k\to\infty} A^k = 0$($A$ 为收敛矩阵) $\Leftrightarrow r(A) < 1 \Leftrightarrow$ 存在相容矩阵范数$\|A\|_m < 1$

5.1.2 矩阵级数及收敛性

设 $\{A^{(k)}\}, k = 1, 2, \cdots$, 是 $C^{m\times n}$ 矩阵序列, 则称 $\sum\limits_{k=1}^{\infty} A^{(k)}$ 为矩阵级数, 称 $S^N = \sum\limits_{k=1}^{N} A^{(k)}$ 为矩阵级数的部分和.

$\sum\limits_{k=1}^{\infty} A^{(k)}$ 收敛 ($\lim\limits_{N\to\infty} S^N = S$) $\Leftrightarrow \lim\limits_{N\to\infty} s_{ij}^N = \lim\limits_{N\to\infty} \sum\limits_{k=1}^{N} a_{ij}^{(k)} = s_{ij} (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$

$\Leftrightarrow \sum\limits_{k=1}^{\infty} a_{ij}^{(k)} (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$收敛 $\Rightarrow \lim\limits_{k\to\infty} a_{ij}^{(k)} = 0 \Rightarrow \lim\limits_{k\to\infty} A^{(k)} = 0$

此外，有:

$\sum\limits_{k=1}^{\infty} A^{(k)}$绝对收敛 $\Leftrightarrow \sum\limits_{k=1}^{\infty} |a_{ij}^{(k)}|$收敛 $\Rightarrow \sum\limits_{k=1}^{\infty} a_{ij}^{(k)}$收敛 $\Rightarrow \sum\limits_{k=1}^{\infty} A^{(k)}$收敛

在 $C^{n\times n}$ 中，$\sum\limits_{k=1}^{\infty} A^{(k)}$绝对收敛 $\Leftrightarrow \sum\limits_{k=1}^{\infty} \|A^{(k)}\|$ 收敛 (任一矩阵范数)

5.1.3 矩阵幂级数及收敛性 (方阵)

$$\sum_{k=1}^{\infty} A^{(k)} \xlongequal{A^{(k)}=c_{k-1}A^{k-1}} \sum_{k=0}^{\infty} c_k A^k (\text{矩阵幂级数}) \xlongequal{c_k=1} \sum_{k=0}^{\infty} A^k (\text{Neummam 级数})$$

定理 5.3 方阵 A 的 Neummam 级数 $\sum\limits_{k=0}^{\infty} A^k$ 收敛 $\Leftrightarrow$ A 为收敛矩阵($r(A) < 1$), 且收敛时, 其和为 $(E - A)^{-1}$(参考定理5.2)

更具一般性

定理 5.4 设幂级数 $f(z) = \sum\limits_{k=0}^{\infty} c_k z^k$ 的收敛半径为 $r \Rightarrow$ 方阵 A 的幂级数 $\sum\limits_{k=0}^{\infty} c_k A^k$ 绝对收敛 ($r(A) < r$)|发散 ($r(A) > r$)

5.2 矩阵函数

矩阵 A 的函数:

$$f(A) = \sum_{k=0}^{\infty} c_k A^k (r(A) < r, \text{即方阵 } A \text{ 的幂级数绝对收敛}) \tag{5-1}$$

性质:

$$AB = BA \Rightarrow \begin{cases} e^A e^B = e^B e^A = e^{A+B} \\ \cos(A + B) = \cos A \cos B - \sin A \sin B \\ \sin(A + B) = \sin A \cos B + \cos A \sin B \end{cases}$$

5.2.1 常见的矩阵函数

$$e^A = \sum_{k=0}^{\infty} A^k \quad \forall A \in C^{n\times n} \qquad (E-A)^{-1} = \sum_{k=0}^{\infty} A^k \quad r(A) < 1$$

$$\sin A = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} A^{2k+1} \quad \forall A \in C^{n\times n} \qquad \cos A = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} A^{2k} \quad \forall A \in C^{n\times n}$$

$$\ln(E+A) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} A^{k+1} \quad r(A) < 1$$

5.2.2 矩阵函数值的计算

(1) 单纯矩阵函数值的计算 (利用相似对角化)

根据式 (5-1),

$$f(At) = P\mathrm{diag}(f(\lambda_1 t), f(\lambda_2 t), \cdots, f(\lambda_n t))P^{-1}$$

(2) 一般矩阵函数值的计算 (利用 Jordan 标准型法)

$$f(A) = P\mathrm{diag}(f(J_1), f(J_2), \cdots, f(J_n))P^{-1}$$

任一矩阵相似于 Jordan 标准型，即 $P^{-1}AP = J = \mathrm{diag}(J_1, J_2, \cdots, J_n)$($P$ 不要求计算)

$f(J_i)$ 满足：

$$f(J_i) = \begin{bmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & f(\lambda_i) & \cdots & \frac{1}{(m_i-2)!}f^{(m_i-2)}(\lambda_i) \\ & & \ddots & \vdots \\ & & & f(\lambda_i) \end{bmatrix}$$

5.2.3 矩阵函数的几种特殊情形

(1) $A^2 = A$

$A^2 = A(\lambda = 0\ or\ 1) \Rightarrow A^k = A(k \geq 1)$

$\Rightarrow f(A) = \sum\limits_{k=0}^{\infty} c^k A^k = \sum\limits_{k=0}^{\infty} c_k A = c_0 E + \sum\limits_{k=1}^{\infty} c_k A = c_0 E + \sum\limits_{k=1}^{\infty} (c_k \cdot 1) A$

$= c_0 E + (f(1) - c_0) A$

注 $e^A = E + (e-1)A \quad \sin A = (\sin 1)A$

(2) $A^2 = E$(对和矩阵)

$$A^2 = E \Rightarrow \begin{cases} A^{2k} = E \\ A^{2k+1} = A \end{cases} \Rightarrow f(A) = \sum_{k=0}^{\infty} c_k A^k = \sum_{k=0}^{\infty} c_{2k} E + \sum_{k=0}^{\infty} c_{2k+1} A$$

注 $\sin A = (\sin 1)A \quad \cos A = \cos 1$

6 广义逆矩阵

矩阵的单边逆、广义逆、自反广义逆均不唯一，M-P 广义逆是唯一的

(1) $A \in C^{m\times n}, G \in C^{n\times m} \quad \begin{cases} GA = E_n \Rightarrow G = A_L^{-1} \\ AG = E_m \Rightarrow G = A_R^{-1} \end{cases}$

(2) $A \in C^{m\times n}, G \in C^{n\times m} \quad AGA = A \Rightarrow G = A^-$

(3) $A \in C^{m\times n}, G \in C^{n\times m} \quad AGA = A \quad GAG = G \Rightarrow G = A_r^-$

(4) $A\{1\} = \{G | AGA = A, \forall G \in C^{n\times m}\}$

(5) $A\{1,2\} = \{G | AGA = A, GAG = G, \forall G \in C^{n\times m}\}$

(6) $A \in C_n^{m\times n} \Leftrightarrow A_L^{-1}$ 存在 $\Leftrightarrow N(A) = \{0\} \Leftrightarrow A^- A = E_n$

(a) $P\begin{pmatrix} A & E_m \end{pmatrix} \xlongequal{\text{初等行变换}} \begin{pmatrix} E_n & G \\ * & * \end{pmatrix} \Rightarrow G = A_L^{-1}$

(b) $Ax = b$ 有解 $\Leftrightarrow (E_m - AA_L^{-1})b = 0$ 且 $x = A_L^{-1}b = (A^H A)^{-1} A^H b$(唯一解)

(7) $A \in C_m^{m\times n} \Leftrightarrow A_R^{-1}$ 存在 $\Leftrightarrow R(A) = C^m \Leftrightarrow AA^- = E_m$

(a) $\begin{pmatrix} A \\ E_n \end{pmatrix} Q \xlongequal{\text{初等列变换}} \begin{pmatrix} E_m & * \\ G & * \end{pmatrix} \Rightarrow G = A_R^{-1}$

(b) $Ax = b$ 对 $\forall b \in C^m$ 都有解，若 $b \neq 0$, 则 $x = A_R^{-1}b = A^H(AA^H)^{-1}b$

(8) A^- 矩阵的性质：设 $A \in Cm \times n, \lambda \in C$

(a) $(A^T)^- = (A^-)^T, (A^H)^- = (A^-)^H$

(b) AA^-, A^-A 都是幂等矩阵，且

$$\operatorname{rank}(A) = \operatorname{rank}(AA^-) = \operatorname{rank}(A^-A) \leq \operatorname{rank}(A^-)$$

(c) $\lambda^- A^-$ 为 λA 的广义逆矩阵，其中 $\lambda^- = \begin{cases} 0 & \lambda = 0 \\ \lambda^{-1} & \lambda \neq 0 \end{cases}$

(d) $A \in C^{m\times n}, S \in C_m^{m\times m}, T \in C_n^{n\times n}$ 且 $B = SAT$, 则 $B^- = T^{-1}A^-S^{-1}$

(e) $R(AA^-) = R(A), N(A^-A) = N(A)$

(9) $A \in C^{m\times n}, P \in C_m^{m\times m}, Q \in C_n^{n\times n} \Rightarrow Q(PAQ)^-P \in A\{1\}$