

数值线性代数——

矩阵计算

潘建瑜

1 线性空间与内积空间

2 向量范数与矩阵范数

3 矩阵与投影

4 矩阵标准型

5 几类特殊矩阵

6 Kronecker 积

计算数学的主要任务

- 数域, 如: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$
- 线性空间, 如: $\mathbb{R}^n, \mathbb{C}^n, \mathbb{R}^{mn}$
- 线性相关与线性无关, 秩, 基, 维数
- 线性子空间
- 像空间 (列空间, 值域) $\text{Ran}(A)$, 零空间 (核) $\text{Ker}(A)$
- 张成子空间: $\text{span}x_1, x_2, \dots, x_k, \text{span}(A) = \text{Ran}(A)$

设 S_1, S_2 是子空间, 若 $S_1 + S_2$ 中的任一元素都可唯一表示成

$$x = x_1 + x_2, x_1 \in S_1, x_2 \in S_2,$$

则称 $S_1 + S_2$ 为直和, 记为 $S_1 \oplus S_2$.

定理

设 S_1 是 S 的子空间, 则存在另一个子空间 S_2 , 使得

$$S = S_1 \oplus S_2.$$

例： 设 $A \in \mathbb{C}^{mn}$, 则

$$\mathbb{C}^n = \text{Ker}(A) \oplus \text{Ran}(A^*), \quad \mathbb{C}^m = \text{Ker}(A^*) \oplus \text{Ran}(A)$$

- 内积, 内积空间, 欧氏空间, 酉空间
- 常见内积空间:
 - $C^n : (x, y) = yx$
 - $R^n : (x, y) = y^T x$
 - $R^{mn} : (A, B) = \text{tr}(B^T A)$

正交与正交补：

- 正交: 向量正交, 子空间正交
- 正交补空间

线性代数基础

- 1 线性空间与内积空间
- 2 向量范数与矩阵范数
- 3 矩阵与投影
- 4 矩阵标准型
- 5 几类特殊矩阵
- 6 Kronecker 积

向量范数与矩阵范数

定义

(向量范数) 若函数 $f: C^n \rightarrow R$ 满足

(1) $f(x) \geq 0, x \in C^n$, 等号当且仅当 $x = 0$ 时成立;

(2) $f(x) = \|f(x)\|, x \in C^n, C$;

(3) $f(x+y) \leq f(x) + f(y), x, y \in C^n$;

则称 $f(x)$ 为 C^n 上的范数, 通常记作 $\| \cdot \|$

相类似地, 我们可以定义实数空间 R^n 上的向量范数。

常见的向量范数

- 1-范数: $\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$
- 2-范数: $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$
- ∞ -范数: $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$
- p-范数: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}, \quad 1 \leq p < \infty$

常见的向量范数

定义

(范数等价性) \mathbb{C}^n 上的向量范数 $\|\cdot\|_\alpha$ 与 $\|\cdot\|_\beta$ 等价: 存在正常数 c_1, c_2 , 使得

$$c_1 \|x\|_\alpha \leq \|x\|_\beta \leq c_2 \|x\|_\alpha, \quad \forall x \in \mathbb{C}^n$$

定理

\mathbb{C}^n 空间上的所有向量范数都是等价的, 特别地, 有

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty$$

常见的向量范数

定理

(Cauchy-Schwartz 不等式) 设 (\cdot, \cdot) 是 \mathbb{C}^n 上的内积, 则对任意 $x, y \in \mathbb{C}^n$, 有

$$|(\mathbf{x}, \mathbf{y})|^2 \leq (\mathbf{x}, \mathbf{x}) \cdot (\mathbf{y}, \mathbf{y})$$

推论

设 (\cdot, \cdot) 是 \mathbb{C}^n 上的内积, 则 $\|\mathbf{x}\| \triangleq \sqrt{(\mathbf{x}, \mathbf{x})}$ 是 \mathbb{C}^n 上的一个向量范数

定理

设 $\|\cdot\|$ 是 \mathbb{C}^n 上的一个向量范数, 则 $f(\mathbf{x}) \triangleq \|\mathbf{x}\|$ 是 \mathbb{C}^n 上的连续函数。

定义

(矩阵范数) 若函数 $f: \mathbb{C}^{nn} \rightarrow \mathbb{R}$ 满足 (1) $f(A) \geq 0, A \in \mathbb{C}^{nn}$, 等号当且仅当 $A = 0$ 时成立; (2) $f(A) = \|f(A)\|, A \in \mathbb{C}^{nn}, \mathbb{C}$; (3) $f(A+B) \leq f(A) + f(B), A, B \in \mathbb{C}^{nn}$; 则称 $f(x)$ 为 \mathbb{C}^{nn} 上的范数, 通常记作 $\|\cdot\|$ 。

相容的矩阵范数: $f(AB) \leq f(A)f(B), \forall A, B \in \mathbb{C}^{nn}$ 。

若未明确指出, 讲义所涉及矩阵范数都指相容矩阵范数

引理

设 $\|\cdot\|$ 是 \mathbb{C}^n 上的向量范数, 则

$$\|A\| \triangleq \sup_{x \in \mathbb{C}^n, x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

是 $\mathbb{C}^{n \times n}$ 上的范数, 称为算子范数, 或诱导范数, 导出范数。

☒ 算子范数都是相容的, 且

$$\|Ax\| \leq \|A\| \cdot \|x\|, \quad A \in \mathbb{C}^{n \times n}, x \in \mathbb{C}^n$$

☒ 类似地, 我们可以定义 $\mathbb{C}^{m \times n}, \mathbb{R}^{n \times n}, \mathbb{R}^{m \times n}$ 上的矩阵范数。

引理

可以证明:

- (1) 1-范数 (列范数): $\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right)$
- (2) ∞ -范数 (行范数): $\|A\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right)$
- (3) 2-范数 (谱范数): $\|A\|_2 = \sqrt{\rho(A^T A)}$

另一个常用范数 F-范数 $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}$

定理

(矩阵范数的等价性) \mathbb{R}^{nn} 空间上的所有范数都是等价的, 特别地, 有

$$\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_1 \leq \sqrt{n}\|A\|_2,$$

$$\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_\infty \leq \sqrt{n}\|A\|_2,$$

$$\frac{1}{n}\|A\|_\infty \leq \|A\|_1 \leq n\|A\|_\infty,$$

$$\frac{1}{\sqrt{n}}\|A\|_1 \leq \|A\|_F \leq \sqrt{n}\|A\|_2,$$

矩阵范数的一些性质

- 对任意的算子范数 $\|\cdot\|$, 有 $\|I\| = 1$
- 对任意的相容范数 $\|\cdot\|$, 有 $\|I\| \leq 1$
- F-范数是相容的, 但不是算子范数
- $\|\cdot\|_2$ 和 $\|\cdot\|_F$ 酉不变范数
- $\|A^T\|_2 = \|A\|_2, \|A^T\|_1 = \|A\|_\infty$
- 若 A 是正规矩阵, 则 $\|A\|_2 = \rho(A)$

向量序列的收敛

设 $x^{(k)}_{k=1}^{\inf}$ 是 \mathbb{C}^n 中的一个向量序列, 如果存在 $x \in \mathbb{C}^n$, 使得

$$\lim_{k \rightarrow \infty} x_i^{(k)} = x_i, \quad i = 1, 2, \dots, n$$

则称 $x^{(k)}$ (按分量) 收敛到 x , 记为 $\lim_{k \rightarrow \infty} x^{(k)} = x$

定理

(矩阵范数的等价性) 设 $\|\cdot\|$ 是 \mathbb{C}^n 上的任意一个向量范数, 则 $\lim_{k \rightarrow \infty} x^{(k)} = x$ 的充要条件是

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\| = 0$$

收敛速度

设点列 $k_{k=1}^{\text{inf}}$ 收敛, 且 $\lim_{k \rightarrow \infty} \varepsilon_k = 0$. 若存在一个有界常数 $0 < c < 1$, 使得

$$\lim_{k \rightarrow \infty} \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} = c$$

则称点列 k 是 p 次 (渐进) 收敛的. 若 $1 < p < 2$ 或 $p = 1$ 且 $c = 0$, 则称点列是超线性收敛的.

类似地, 我们可以给出矩阵序列的收敛性和判别方法.

- 1 线性空间与内积空间
- 2 向量范数与矩阵范数
- 3 矩阵与投影**
- 4 矩阵标准型
- 5 几类特殊矩阵
- 6 Kronecker 积

特征值与特征向量

- 特征多项式, 特征值, 特征向量, 左特征向量, 特征对
- n 阶矩阵 A 的谱: $(A)1, 2, \dots, n$
- 代数重数和几何重数, 特征空间
- 最小多项式
- 可对角化, 特征值分解
- 可对角化的充要条件
- 特征值估计: Bendixson 定理, 圆盘定理

Bendixson 定理

设 $A \in \mathbb{C}^{n \times n}$, 令 $H = \frac{1}{2} (A + A^*)$, $S = \frac{1}{2} (A - A^*)$. 则有

$$\begin{aligned}\lambda_{\min}(H) &\leq \operatorname{Re}(\lambda(A)) \leq \lambda_{\max}(H) \\ \lambda_{\min}(iS) &\leq \operatorname{Im}(\lambda(A)) \leq \lambda_{\max}(iS)\end{aligned}$$

其中 $\operatorname{Re}(\cdot)$ 和 $\operatorname{Im}(\cdot)$ 分别表示实部和虚部。

☒ 一个矩阵的特征值的实部的取值范围由其 Hermite 部分确定, 而虚部则由其 Skew-Hermite 部分确定.

Gerschgorin 圆盘定理

设 $A = [a_{ij}] \in \mathbb{C}^{n \times n}$, 定义集合

$$\mathcal{D}_i \triangleq \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| \right\}, \quad i = 1, 2, \dots, n$$

这就是 A 的 n 个 Gerschgorin 圆盘。

定理

(Gerschgorin 圆盘定理) 设 $A = [a_{ij}] \in \mathbb{C}^{n \times n}$. 则 A 的所有特征值都包含在 A 的 Gerschgorin 圆盘的并集中, 即 $\sigma(A) \subset \bigcup_{i=1}^n \mathcal{D}_i$

投影变换与投影矩阵

设 $S = S_1 \oplus S_2$, 则 S 中的任意向量 x 都可唯一表示为 $x = x_1 + x_2, x_1 \in S_1, x_2 \in S_2$. 我们称 x_1 为 x 沿 S_2 到 S_1 上的投影, 记为 $x|_{S_1}$. 设线性变换 $P: S \rightarrow S$. 如果对任意 $x \in S$, 都有 $Px = x|_{S_1}$, 则称 P 是从 S 沿 S_2 到 S_1 上的投影变换 (或投影算子), 对应的变换矩阵称为投影矩阵.

引理

设 $P \in \mathbb{R}^{n \times n}$ 是一个投影矩阵, 则

$$\mathbb{R}^n = \text{Ran}(P) \oplus \text{Ker}(P) \quad (1)$$

反之, 若 (1.3) 成立, 则 P 是沿 $\text{Ker}(P)$ 到 $\text{Ran}(P)$ 上的投影

投影矩阵由其像空间和零空间唯一确定.

引理

若 S_1 和 S_2 是 \mathbb{R}^n 的两个子空间, 且 $\mathbb{R}^n = S_1 \oplus S_2$, 则存在唯一的投影矩阵 P , 使得

$$\text{Ran}(P) = S_1, \quad \text{Ker}(P) = S_2$$

投影矩阵的判别

定理

矩阵 $P \in \mathbb{R}^{n \times n}$ 是投影矩阵的充要条件是 $P^2 = P$

投影算子的矩阵表示

设 S_1 和 S_2 是 \mathbb{R}^n 的两个 m 维子空间. 如果 $S_1 \oplus S_2^\perp = \mathbb{R}^n$, 则存在唯一的投影矩阵 P , 使得

$$\text{Ran}(P) = S_1, \quad \text{Ker}(P) = S_2^\perp$$

此时, 我们称 P 是 S_1 上与 S_2 正交的投影矩阵, 且有

$$P = V(W^\top V)^{-1}W^\top$$

其中 $V = [v_1, v_2, \dots, v_m]$ 和 $W = [w_1, w_2, \dots, w_m]$ 的列向量组分别构成 S_1 和 S_2 的一组基.

设 S_1 是内积空间 S 的一个子空间, $x \in S$, 则 x 可唯一分解成

$$x = x_1 + x_2, x_1 \in S_1, x_2 \in S_1^\perp$$

, 其中 x_1 称为 x 在 S_1 上的正交投影.

- 若 P 是沿 S_1^\perp 到 S_1 上的投影变换, 则称 P 为 S_1 上的正交投影变换 (对应的矩阵为正交投影矩阵), 记为 P_{S_1}

- 如果 P 不是正交投影变换, 则称其为斜投影变换

定理

投影矩阵 $P \in \mathbb{R}^{n \times n}$ 是正交投影矩阵的充要条件 $P = P^2$.

推论

设 P 是子空间 S_1 上的正交投影变换. 令 v_1, v_2, \dots, v_m 是 S_1 的一组标准正交基, 则

$$P = VV^T$$

其中 $V = [v_1, v_2, \dots, v_m]$.

性质

设 $P \in \mathbb{R}^{nn}$ 是一个正交投影矩阵, 则

$$\|P\|_2 = 1$$

且对 $x \in \mathbb{R}^n$, 有

$$\|x\|_2^2 = \|Px\|_2^2 + \|(I - P)x\|_2^2$$

正交投影矩阵的一个重要应用

定理

设 S_1 是 R_n 的一个子空间, $z \in R_n$ 是一个向量. 则最佳逼近问题

$$\min_{x \in S_1} \|x - z\|_2$$

的唯一解为

$$x_* = P_{S_1} z$$

即 S_1 中距离 z 最近 (2-范数意义下) 的向量是 z 在 S_1 上的正交投影.

正交投影矩阵的一个重要应用

推论

设矩阵 $A \in \mathbb{R}^{n \times n}$ 对称正定, 向量 $x \in S_1 \subset \mathbb{R}^n$. 则 x 是最佳逼近问题

$$\min_{x \in S_1} \|x - z\|_A$$

的解的充要条件是

$$A(x_* - z) \perp S_1$$

这里 $\|x - z\|_A \triangleq \left\| A^{\frac{1}{2}}(x - z) \right\|_2$

不变子空间

设 $A \in \mathbb{R}^{n \times n}$, S 是 \mathbb{R}^n 的一个子空间, 记

$$AS \triangleq \{Ax : x \in S\}$$

定义

若 $AS \subseteq S$, 则称 S 为 A 的一个不变子空间.

定理

设 x_1, x_2, \dots, x_m 是 A 的一组线性无关特征向量, 则

$$\text{span} \{x_1, x_2, \dots, x_m\}$$

是 A 的一个 m 维不变子空间.

不变子空间的一个重要性质

定理

设 $A \in \mathbb{R}^{nn}$, $X \in \mathbb{R}^{nk}$ 且 $\text{rank}(X) = k$. 则 $\text{span}(X)$ 是 A 的不变子空间的充要条件是存在 $B \in \mathbb{R}^{kk}$ 使得

$$AX = XB,$$

此时, B 的特征值都是 A 的特征值.

推论

设 $A \in \mathbb{R}^{nn}$, $X \in \mathbb{R}^{nk}$ 且 $\text{rank}(X) = k$. 若存在一个矩阵 $B \in \mathbb{R}^{kk}$ 使得 $AX = XB$, 则 $(, v)$ 是 B 的一个特征对当且仅当 $(, Xv)$ 是 A 的一个特征对.

- 1 线性空间与内积空间
- 2 向量范数与矩阵范数
- 3 矩阵与投影
- 4 矩阵标准型**
- 5 几类特殊矩阵
- 6 Kronecker 积

计算矩阵特征值的一个基本思想是通过相似变换, 将其转化成一个形式尽可能简单的矩阵, 使得其特征值更易于计算. 其中两个非常有用的特殊矩阵是 Jordan 标准型和 Schur 标准型.

矩阵标准型

定理

设 $A \in \mathbb{C}^{n \times n}$ 有 p 个不同特征值, 则存在非奇异矩阵 $X \in \mathbb{C}^{n \times n}$, 使得

$$X^{-1}AX = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix} \triangleq J$$

其中 J_i 的维数等于 λ_i 的代数重数, 且具有下面的结构

$$J_i = \begin{bmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{i, \nu_i} \end{bmatrix} \quad J_{ik} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix}$$

这里 ν_i 为 λ_i 的几何重数, J_{ik} 称为 Jordan 块, 每个 Jordan 块对应一个特征向量

☒ Jordan 标准型在理论研究中非常有用, 但数值计算比较困难, 目前还没有找到十分稳定的数值算法.

推论

所有可对角化矩阵组成的集合在所有矩阵组成的集合中是稠密的.

定理

设 $A \in \mathbb{C}^{nn}$, 则存在一个酉矩阵 $U \in \mathbb{C}^{nn}$ 使得

$$U^*AU = \begin{bmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ 0 & \lambda_2 & \cdots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \triangleq R \text{ 或 } A = URU^*$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值 (排序任意).

关于 Schur 标准型的几点说明:

- Schur 标准型可以说是酉相似变化下的最简形式
- U 和 R 不唯一, R 的对角线元素可按任意顺序排列
- A 是正规矩阵当且仅当定理 (3.15) 中的 R 是对角矩阵;
- A 是 Hermite 矩阵当且仅当定理 (3.15) 中的 R 是实对角矩阵.

定理

设 $A \in \mathbb{R}^{n \times n}$, 则存在正交矩阵 $Q \in \mathbb{R}^{n \times n}$, 使得

$$Q^T A Q = T$$

其中 $T \in \mathbb{R}^{n \times n}$ 是拟上三角矩阵, 即 T 是块上三角的, 且对角块为 1×1 或 2×2 的块矩阵. 若对角块是 1×1 的, 则其就是 A 的一个特征值, 若对角块是 2×2 的, 则其特征值是 A 的一对共轭复特征值.

- 1 线性空间与内积空间
- 2 向量范数与矩阵范数
- 3 矩阵与投影
- 4 矩阵标准型
- 5 几类特殊矩阵**
- 6 Kronecker 积

对称正定矩阵

设 $A \in \mathbb{C}^{nn}$.

A 是半正定 $\iff \operatorname{Re}(x^*Ax) \geq 0, \forall x \in \mathbb{C}^n$

A 是正定 $\iff \operatorname{Re}(x^*Ax) > 0, \forall x \in \mathbb{C}^n, x \neq 0$

A 是 Hermite 半正定 $\iff A$ Hermite 且半正定

A 是 Hermite 正定 $\iff A$ Hermite 且正定

\boxtimes 正定和半正定矩阵不要求是对称或 Hermite 的

定理

设 $A \in \mathbb{C}^{n \times n}$. 则 A 正定 (半正定) 的充要条件是矩阵 $H = \frac{1}{2}(A + A^H)$ 正定 (半正定).

定理

设 $A \in \mathbb{R}^{n \times n}$. 则 A 正定 (或半正定) 的充要条件是对任意非零向量 $x \in \mathbb{R}^n$ 有 $x^T A x > 0$ (或 $x^T A x \geq 0$).

定理

设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 半正定, k 是正整数. 则存在唯一的 Hermite 半正定矩阵 $B \in \mathbb{C}^{n \times n}$ 使得

$$B^k = A.$$

同时, 我们还有下面的性质: (1) $BA = AB$, 且存在一个多项式 $p(t)$ 使得 $B = p(A)$; (2) $\text{rank}(B) = \text{rank}(A)$, 因此, 若 A 是正定的, 则 B 也正定; (3) 如果 A 是实矩阵的, 则 B 也是实矩阵.

特别地, 当 $k = 2$ 时, 称 B 为 A 的平方根, 通常记为 $A^{\frac{1}{2}}$.

Hermite 正定矩阵与内积之间有这样的关系

定理

设 $(,)$ 是 \mathbb{C}^n 上的一个内积, 则存在一个 Hermite 正定矩阵 $A \in \mathbb{C}^{n \times n}$ 使得

$$(x, y) = y^* A x.$$

反之, 若 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 正定矩阵, 则

$$f(x, y) \triangleq y^* A x$$

是 \mathbb{C}^n 上的一个内积.

☑ 上述性质在实数域中也成立.

定义

设 $A \in \mathbb{C}^{nn}$, 若

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$$

对所有 $i = 1, 2, \dots, n$ 都成立, 且至少有一个不等式严格成立, 则称 A 为弱行对角占优. 若对所有 $i = 1, 2, \dots, n$ 不等式都严格成立, 则称 A 是严格行对角占优. 通常简称为弱对角占优和严格对角占优.

☒ 类似地, 可以定义弱列对角占优和严格列对角占优.

可约与不可约

设 $A \in \mathbb{R}^{n \times n}$, 若存在置换矩阵 P , 使得 PAP^T 为块上三角, 即

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

其中 $A_{11} \in \mathbb{R}^{k \times k}$ ($1 \leq k < n$), 则称 A 为可约, 否则不可约.

定理

设 $A \in \mathbb{R}^{n \times n}$, 指标集 $\mathbb{Z}_n = \{1, 2, \dots, n\}$. 则 A 可约的充要条件是存在非空指标集 $J \subset \mathbb{Z}_n$ 且 $J \neq \mathbb{Z}_n$, 使得

$$a_{ij} = 0, \quad i \in J \text{ 且 } j \in \mathbb{Z}_n \setminus J$$

这里 $\mathbb{Z}_n \setminus J$ 表示 J 在 \mathbb{Z}_n 中的补集.

定理

若 $A \in \mathbb{C}^{n \times n}$ 严格对角占优, 则 A 非奇异

定理

若 $A \in \mathbb{C}^{n \times n}$ 不可约对角占优, 则 A 非奇异

其他常见特殊矩阵

- 带状矩阵: $a_{ij} \neq 0$ only if $-b_u \leq i - j \leq b_l$, 其中 b_u 和 b_l 为非负整数, 分别称为下带宽和上带宽, $b_u + b_l + 1$ 称为 A 的带宽
- 上 Hessenberg 矩阵: $a_{ij} = 0$ for $ij > 1$,

$$\begin{bmatrix} * & * & * & \dots & * \\ * & * & * & \dots & * \\ & * & * & \dots & * \\ & & \ddots & \ddots & \vdots \\ & & & * & * \end{bmatrix}$$

- 下 Hessenberg 矩阵

其他常见特殊矩阵

- Toeplitz 矩阵

$$T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-n+1} \\ t_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{bmatrix}$$

- 循环矩阵 (circulant):

$$C = \begin{bmatrix} c_0 & c_{n-1} & c_{n-2} & \cdots & c_1 \\ c_1 & c_0 & c_{n-1} & \cdots & c_2 \\ c_2 & c_1 & c_0 & \cdots & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_0 \end{bmatrix}$$

其他常见特殊矩阵

Hankel 矩阵:

$$H = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-2} & h_{n-1} \\ h_1 & \ddots & \ddots & \ddots & h_n \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{n-2} & \cdots & \cdots & \cdots & h_{2n-2} \\ h_{n-1} & h_n & \cdots & h_{2n-2} & h_{2n-1} \end{bmatrix}$$

- 1 线性空间与内积空间
- 2 向量范数与矩阵范数
- 3 矩阵与投影
- 4 矩阵标准型
- 5 几类特殊矩阵
- 6 Kronecker 积**

定义

设 $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{p \times q}$, 则 A 与 B 的 Kronecker 积定义为

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp \times nq}$$

Kronecker 积也称为直积, 或张量积.

† 任意两个矩阵都存在 Kronecker 积, 且 AB 和 BA 是同阶矩阵, 但通常 $A \otimes B \neq B \otimes A$

基本性质

- (1) $(A) \oplus B = A \oplus (B) = (A \oplus B), \mathbb{C}$
- (2) $(A \oplus B) = A \oplus B, (A \oplus B) = A \oplus B$
- (3) $(A \oplus B) \oplus C = A \oplus (B \oplus C)$
- (4) $(A + B) \oplus C = A \oplus C + B \oplus C$
- (5) $A \oplus (B + C) = A \oplus B + A \oplus C$
- (6) 混合积: $(A \oplus B)(C \oplus D) = (AC) \oplus (BD)$
- (7) $(A_1 \oplus A_2 \oplus \dots \oplus A_k)(B_1 \oplus B_2 \oplus \dots \oplus B_k) = (A_1 B_1) \oplus (A_2 B_2) \oplus \dots \oplus (A_k B_k)$
- (8) $(A_1 \oplus B_1)(A_2 \oplus B_2)(A_k \oplus B_k) = (A_1 A_2 A_k) \oplus (B_1 B_2 B_k)$
- (9) $\text{rank}(A \oplus B) = \text{rank}(A) + \text{rank}(B)$

定理

设 $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, 并设 (λ, x) 和 (μ, y) 分别是 A 和 B 的一个特征对, 则 $(\lambda\mu, xy)$ 是 AB 的一个特征对. 由此可知, BA 与 AB 具有相同的特征值.

定理

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

- (1) $\text{tr}(AB) = \text{tr}(A)\text{tr}(B)$;
- (2) $\det(AB) = \det(A)^n \det(B)^m$;
- (3) $A I_n + I_m B$ 的特征值为 $\lambda_i + \mu_j$, 其中 λ_i 和 μ_j 分别为 A 和 B 的特征值;
- (4) 若 A 和 B 都非奇异, 则 $(AB)^{-1} = A^{-1}B^{-1}$;

推论

设 $A = Q_1 \Lambda_1 Q_1^{-1}$, $B = Q_2 \Lambda_2 Q_2^{-1}$, 则

$$A \otimes B = (Q_1 \otimes Q_2) (\Lambda_1 \otimes \Lambda_2) (Q_1 \otimes Q_2)^{-1}$$

定理

设 $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, 则存在 $m + n$ 阶置换矩阵 P 使得

$$P^T (A \otimes B) P = B \otimes A$$

定理

设矩阵 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$, 记 $\text{vec}(X)$ 为 X 按列拉成的 mn 维列向量, 即

$$\text{vec}(X) = [x_1^\top, x_2^\top, \dots, x_n^\top]^\top$$

则有

$$\text{vec}(AX) = (IA)\text{vec}(X), \text{vec}(XB) = (B^\top I)\text{vec}(X),$$

以及

$$(AB)\text{vec}(X) = \text{vec}(BXA^\top)$$