第二章 线性代数基础

魏华祎，易年余，陆讯

湘潭大学数学与计算科学学院

September 27, 2019

计算数学的主要任务

- 数域, 如: Q, R, C
- 线性空间, 如: $R^n, C^n, R^{m\times n}$
- 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) Ran(A), 零空间 (核) Ker(A)
- 张成子空间:$\mathrm{span}\{x_1, x_2, ..., x_k\}, \mathrm{span}(A) = \mathrm{Ran}(A)$

直和

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

$$x = x_1 + x_2, x_1 \in S1, x2 \in S2,$$

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 $A \in \mathbb{C}^{m\times n}$, 则

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

内积空间

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

正交与正交补:

- 正交: 向量正交, 子空间正交
- 正交补空间

定义

(向量范数) 若函数 $f: C^n \to R$ 满足
(1) $f(x) \geq 0, x \in C^n$, 等号当且仅当 $x = 0$ 时成立;
(2) $f(\alpha x) = |\alpha| f(x), x \in C^n, \in C$;
(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:$||x||_1 = |x_1| + |x_2| + ... + |x_n|$
- 2-范数:$||x||_2 = \sqrt{|x_1|^2 + |x_2|^2 + ... + |x_n|^2}$
- ∞-范数:$||x||_\infty = \max_{1\leq i\leq n} |x_i|$
- p-范数:$||x||_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, \quad 1 \leq p < \infty$

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

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

定理

$\mathbb{C}^n$ 空间上的所有向量范数都是等价的, 特别地, 有

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

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

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

定理

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

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

推论

设 $(.,.)$ 是 $\mathbb{C}^n$ 上的内积，则 $||x|| \triangleq \sqrt{(x,x)}$ 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 $f(x) \triangleq ||x||$ 是 $\mathbb{C}^n$ 上的连续函数。

定义

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

相容的矩阵范数：$f(AB) \leq f(A)f(B), \forall A, B \in \mathbb{C}^{n\times n}$。
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 是 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):$\|A\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$

(2) ∞-范数 (行范数): $\|A\|_\infty = \max_{1\le i\le n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$

(3) 2-范数 (谱范数):$\|A\|_2 = \sqrt{\rho\left(A^\top A\right)}$

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

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

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

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

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

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

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

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

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

定理

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

$$\lim_{k\to\infty} \left\| x^{(k)} - x \right\| = 0$$

设点列 $k_{k=1}^{inf}$ 收敛, 且 $\lim_{k=\infty}\varepsilon_k = 0$. 若存在一个有界常数 $0 < c <$, 使得
$\lim_{k\to\infty}\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} = c$
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.

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, ..., 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

$$
\begin{aligned}
\lambda_{\min}(H) \le \mathrm{Re}(\lambda(A)) \le \lambda_{\max}(H) \\
\lambda_{\min}(iS) \le \mathrm{Im}(\lambda(A)) \le \lambda_{\max}(iS)
\end{aligned}
$$

其中 Re(.) 和 Im(.) 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.

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

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

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 $\sigma(A)\subset\bigcup_{i=1}^{n}\mathcal{D}_i$

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 $P : S \to S$. 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.

投影变换与投影矩阵

引理

设 $P \in \mathbb{R}^{nn}$ 是一个投影矩阵, 则

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

反之, 若 (1.3) 成立, 则 P 是沿 Ker(P) 到 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 $\mathbb{R}^n = S_1 S_2$, 则存在唯一的投影矩阵 P, 使得

$$\mathrm{Ran}(P) = \mathcal{S}_1, \quad \mathrm{Ker}(P) = \mathcal{S}_2$$

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 $S_1 \oplus \mathcal{S}_2^{\perp} = \mathbb{R}^n$, 则存在唯一的投影矩阵 P, 使得

$$\mathrm{Ran}(P) = \mathcal{S}_1, \quad \mathrm{Ker}(P) = \mathcal{S}_2^{\perp}$$

此时, 我们称 P 是 S1 上与 S2 正交的投影矩阵, 且有

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

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

正交投影

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

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

, 其中 x_1 称为 x 在 S_1 上的正交投影.

- 若 P 是沿 S_1 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 $P \in \mathbb{R}^{nn}$ 是正交投影矩阵的充要条件 $P = P$.

推论

设 P 是子空间 S1 上的正交投影变换. 令 $v_1, v_2, ..., v_m$ 是 S_1 的一组标准正交基, 则

$$P = VV^\top$$

其中 $V = [v_1, v_2, ..., v_m]$.

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

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

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

$$\min_{x \in \mathcal{S}_1} \|x - z\|_2$$

的唯一解为

$$x_* = P_{\mathcal{S}_1} z$$

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 $A \in \mathbb{R}^{nn}$ 对称正定, 向量 $x \in S_1 R_n$. 则 x 是最佳逼近问题

$$\min_{x \in \mathcal{S}_1} \|x - z\|_A$$

的解的充要条件是

$$A(x_* - z) \perp \mathcal{S}_1$$

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

不变子空间

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

$$A\mathcal{S} \triangleq \{Ax : x \in \mathcal{S}\}$$

定义

若 ASS, 则称 S 为 A 的一个不变子空间.

定理

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

$$\operatorname{span}\{x_1, x_2, \ldots, x_m\}$$

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

$$AX = XB,$$

此时, B 的特征值都是 A 的特征值.

推论

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

矩阵标准型

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.

定理

设 $A \in \mathbb{C}^{nn}$ 有 p 个不同特征值，则存在非奇异矩阵 $X \in \mathbb{C}^{nn}$，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 $_i$ 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 $_i$ 为 $_i$ 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.

Schur 标准型

定理

设 $A \in \mathbb{C}^{n\times n}$, 则存在一个酉矩阵 $U \in \mathbb{C}^{n\times n}$ 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 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 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上述定理中的 R 是对角矩阵;
- A 是 Hermite 矩阵当且仅当上述定理中的 R 是实对角矩阵.

实 Schur 标准型

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

$$Q^\top AQ = T$$

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.

对称正定矩阵

设 $A \in \mathbb{C}^{n\times n}$.

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

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

A 是 Hermite 半正定 $\iff$ A Hermite 且半正定

A 是 Hermite 正定 $\iff$ A Hermite 且正定

正定和半正定矩阵不要求是对称或 Hermite 的

对称正定矩阵

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 $H = \frac{1}{2}(A + A)$ 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 $x^\top Ax > 0$(或 $x^\top Ax \geq 0$).

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下面的关系

定理

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

$$(x,y)=yAx.$$

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

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

是 $\mathbb{C}^n$ 上的一个内积.

上述性质在实数域中也成立.

对角占优矩阵

定义

设 $A \in \mathbb{C}^{n\times n}$, 若

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

对所有 $i = 1, 2, ..., n$ 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 $i = 1, 2, ..., n$ 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 $PAP^\top$ 为块上三角, 即

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

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

定理

设 $A \in \mathbb{R}^{n\times n}$, 指标集 $\mathbb{Z}_n = \{1, 2, \ldots, \ldots, n\}$. 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

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

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

可约与不可约

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: $a_{ij} = 0 \quad \text{for} \quad i-j > 1$,

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

- 下 Hessenberg 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):

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

Hankel 矩阵:

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

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 $A \oplus B$ 和 $B \oplus A$ 是同阶矩阵, 但通常 $A \otimes B \neq B \otimes A$

基本性质

(1) $(A) \oplus B = A \oplus (B) = (A \oplus B), \ \in \mathbb{C}$

(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) $(A_1 \oplus A_2 \oplus \ \oplus A_k)(B_1 \oplus B_2 \oplus \ \oplus B_k) = (A_1B_1) \oplus (A_2B_2) \oplus \ \oplus (A_kB_k)$

(8) $(A_1 \oplus B_1)(A_2 \oplus B_2)(A_k \oplus B_k) = (A_1A_2A_k) \oplus (B_1B_2B_k)$

(9) $\mathrm{rank}(A \oplus B) = \mathrm{rank}(A)\mathrm{rank}(B)$

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 $(\lambda\mu, x \oplus y)$ 是 $A \oplus B$ 的一个特征对. 由此可知,$B \oplus A$ 与 $A \oplus B$ 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) $\mathrm{tr}(A \oplus B) = \mathrm{tr}(A)\mathrm{tr}(B)$;
(2) $\det(A \oplus B) = \det(A)^n \det(B)^m$;
(3) $A \oplus I_n + I_m \oplus B$ 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 $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$;

推论

设 $A = Q_1\Lambda_1Q_1^{-1}, B = Q_2\Lambda_2Q_2^{-1}$, 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

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

定理

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

$$\mathrm{vec}(X) = \left[x_1^\top, x_2^\top, \ldots, x_N^\top\right]^\top$$

则有

$$\mathrm{vec}(AX) = (I \oplus A)\mathrm{vec}(X), \mathrm{vec}(XB) = (B^\top \oplus I)\mathrm{vec}(X),$$

以及

$$(A \oplus B)\mathrm{vec}(X) = \mathrm{vec}(BXA^\top)$$