矩阵计算及其理论

潘建瑜

华东师范大学 数学科学学院

http://math.ecnu.edu.cn/~jypan/Teaching/MCT/index.html

第一讲　线性代数基础

1　矩阵基础知识

2　对称正定矩阵

3　向量范数与矩阵范数

4　Kronecker 积

1 矩阵基础知识

- 矩阵: 方阵, 长方阵, 基本运算, 转置, 共轭转置, 逆矩阵, 矩阵初等变换
- 行列式: 代数余子式, 伴随矩阵, 基本性质
- 线性相关与线性无关
- Gram-Schmidt 正交化过程

- 矩阵的秩

关于矩阵秩的相关性质

- $\mathrm{rank}(A) \le \min\{m, n\}, \quad A \in \mathbb{R}^{m\times n}$;
- $\mathrm{rank}(A^\intercal) = \mathrm{rank}(A)$;
- $\mathrm{rank}(A^\intercal A) = \mathrm{rank}(AA^\intercal) = \mathrm{rank}(A)$;
- $\mathrm{rank}(PA) = \mathrm{rank}(AQ) = \mathrm{rank}(A), \quad P, Q$ 非奇异;
- $\mathrm{rank}(A + B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$;
- $\mathrm{rank}(A) = \mathrm{rank}(B) \Longleftrightarrow A = PBQ, \quad P, Q$ 非奇异;
- $\mathrm{rank}(A) + \mathrm{rank}(B) - k \le \mathrm{rank}(AB) \le \min\{\mathrm{rank}(A),\ \mathrm{rank}(B)\}$.

- 两类等价关系

定义 (矩阵相似) 设 $A, B \in \mathbb{R}^{n\times n}$, 若存在非奇异矩阵 $X \in \mathbb{R}^{n\times n}$ 使得 $A = XBX^{-1}$, 则称 A 与 B **相似**.

定义 (矩阵合同) 设 $A, B \in \mathbb{R}^{n\times n}$ 对称, 若存在非奇异矩阵 $X \in \mathbb{R}^{n\times n}$ 使得 $A = XBX^{\intercal}$, 则称 A 与 B **合同**.

- 特征值与特征向量

定义 (特征多项式) 设 $A \in \mathbb{R}^{n\times n}$, 称 $p_A(\lambda) = \det(A - \lambda I)$ 为 A 的**特征多项式**, 其零点就是 A 的**特征值**.

✍ 特征多项式 $p_A(\lambda)$ 是 n 次多项式, 其零点就是 A 的特征值, 记为 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 即 $p_A(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$

✍ 本讲义中, 如果没有特别指出, 我们通常假定

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

定义 (特征值与特征向量) 设 $A \in \mathbb{R}^{n\times n}$. 若存在 $\lambda \in \mathbb{C}$ 和非零向量 $x, y \in \mathbb{C}^n$, 满足

$$Ax = \lambda x, \qquad y^* A = \lambda y^* \ (\text{或 } A^* y = \bar{\lambda} y),$$

则称 λ 为 A 的**特征值**, x 为 A 对应于 λ 的（右）**特征向量**, y 为 A 对应于 λ 的**左特征向量**, 并称 (λ, x) 为 A 的一个**特征对** (**eigenpair**).

根据多项式零点关于多项式系数的连续性, 我们可以得到下面的结论.

定理 矩阵的特征值关于矩阵元素是连续的. 即当矩阵的元素发生变化时, 其特征值的变化是连续的.

关于特征值的几点说明

- 只有当 A 是方阵时, 才具有特征值与特征向量;
- 实矩阵的特征值与特征向量也有可能是复的, 如果实矩阵存在复特征值, 则一定成对出现 (相互共轭);
- n 阶矩阵总是存在 n 个特征值 (其中可能有相等的);
- 特征值有代数重数和几何重数, 几何重数不超过代数重数;
- 相似变换不改变矩阵的特征值;
- 合同变换不改变矩阵的惯性指数.

设 $A \in \mathbb{R}^{n\times n}$, 则 $A^{\intercal}$ 和 A 的特征值和特征向量是什么关系?

A^{-1} 和 A 的特征值和特征向量是什么关系?

设 $p(t)$ 是多项式, 则 $p(A)$ 和 A 的特征值与特征向量是什么关系?

- 矩阵的迹

定义 (矩阵的迹) 设 $A \in \mathbb{R}^{n\times n}$, 称 $\mathrm{tr}(A) \triangleq a_{11} + a_{22} + \cdots + a_{nn}$ 为矩阵 A 的迹 (trace).

定理 设 $A \in \mathbb{R}^{n\times n}$, 则

$$\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad \det(A) = \lambda_1\lambda_2\cdots\lambda_n.$$

- **矩阵的谱与谱半径**

  所有特征值组成的集合 $\sigma(A)$ 称为 A 的**谱**, 称 $\rho(A)$ 为 A 的**谱半径**, 其中

$$\rho(A) \triangleq \max_{\lambda \in \sigma(A)} \{|\lambda|\}, \quad \sigma(A) \triangleq \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$

- 可对角化

  若存在非奇异矩阵 $X \in \mathbb{R}^{n\times n}$, 使得

$$\boxed{A = X\Lambda X^{-1}, \quad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)} \tag{1.1}$$

则称 A 是**可对角化**的. 称 (1.1) 为 A 的**特征值分解**或**谱分解**.

定理 设 $A \in \mathbb{R}^{n\times n}$.

(1) A 可对角化的充要条件是 A 具有 n 个线性无关的特征向量;

(2) A 可对角化的充要条件是 A 的每个特征值的几何重数等于代数重数;

(3) 若 A 的特征值互不相同, 则 A 可对角化.

例 若 $A \in \mathbb{R}^{n\times n}$ 对称, 则 A 可对角化.

- 特征值的简单估计

定义 (Gerschgorin 圆盘) 设 $A=[a_{ij}]\in\mathbb{R}^{n\times n}$, 定义集合

$$\mathcal{D}_i \triangleq \left\{ z\in\mathbb{C} \,:\, |z-a_{ii}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}| \right\}, \quad i=1,2,\ldots,n. \tag{1.2}$$

我们称 $\mathcal{D}_i$ 为 A 的 n 个 **Gerschgorin 圆盘**.

定理 (Gerschgorin 圆盘定理) 设 $A=[a_{ij}]\in\mathbb{R}^{n\times n}$, 则 A 的所有特征值都包含在 A 的 Gerschgorin 圆盘的并集中, 即

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

将 A 的非对角线元素换成 τa_{ij}, 其中 $0 \le \tau \le 1$, 并利用特征值关于矩阵元素的连续性, 我们就可以得到下面的结论.

定理 设 $A = [a_{ij}] \in \mathbb{R}^{n\times n}$, 如果 $\bigcup\limits_{i=1}^{n} \mathcal{D}_i$ 可分解成两个不相交的子集 S 和 T, 即

$$\bigcup_{i=1}^{n} \mathcal{D}_i = \mathsf{S} \bigcup \mathsf{T} \quad 且 \quad \mathsf{S} \bigcap \mathsf{T} = \emptyset.$$

假定 S 由 k 的圆盘组成, 而 T 由其它 $n-k$ 个圆盘组成, 则 S 中恰好包含 A 的 k 个特征值 (重特征值按重数计算), 而 T 中则包含 A 的其它 $n-k$ 个特征值.

- **矩阵的直和**

矩阵 $A_i \in \mathbb{R}^{n_i \times n_i}$ $(i = 1, 2, \ldots, k)$ 的**直和**定义为

$$A_1 \oplus A_2 \oplus \cdots \oplus A_k \triangleq \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix},$$ 即块对角矩阵

直和的简单性质

- $\det(A \oplus B) = \det(A) \cdot \det(B)$
- $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$
- $\alpha(A \oplus B) = (\alpha A) \oplus (\alpha B)$

- 一些特殊矩阵
  - 单位矩阵, 数量矩阵, 对角矩阵, 三对角矩阵, 上 (下) 三角矩阵, 上 (下) Hessenberg 矩阵, 带状矩阵
  - 置换 (排列) 矩阵, 对称矩阵 (Hermitian 矩阵), 反向单位矩阵, 反向对称矩阵, 斜对称矩阵
  - Vandermonde 矩阵, Toeplitz 矩阵, 循环矩阵, Hankel 矩阵
  - 正定矩阵, 半正定矩阵, 对角占优矩阵, 不可约矩阵
  - 正交矩阵 (酉矩阵), 对合矩阵 ($A^2 = I$), 幂等矩阵 ($A^2 = A$, 也称投影矩阵), 幂零矩阵 ($A^k = 0$)
  - 分块矩阵 (块对角, 块三角, ...)

2 对称正定矩阵

定义 设 $A \in \mathbb{R}^{n\times n}$.

- 若对所有向量 $x \in \mathbb{C}^n$ 有 $\mathrm{Re}(x^* A x) \geq 0$, 则称 A 是**半正定**的;
- 若对所有非零向量 $x \in \mathbb{C}^n$ 有 $\mathrm{Re}(x^* A x) > 0$, 则称 A 是**正定**的;
- 若 $A \in \mathbb{R}^{n\times n}$ 是对称的且半正定, 则称 A 为 **对称半正定**;
- 若 $A \in \mathbb{R}^{n\times n}$ 是对称的且正定, 则称 A 为 **对称正定**.

✍ 本讲义中, 正定和半正定矩阵不要求是对称的.

定理 设 $A \in \mathbb{R}^{n\times n}$, 则

- A 正定的充要条件是对任意非零向量 $x \in \mathbb{R}^n$ 有 $x^\intercal Ax > 0$;
- A 半正定的充要条件是对任意非零向量 $x \in \mathbb{R}^n$ 有 $x^\intercal Ax \geq 0$.

(留作练习)

定理 设 $A \in \mathbb{R}^{n\times n}$ 对称, 则

- A 正定的充要条件是其所有特征值都是正实数;
- A 半正定的充要条件是其所有特征值都是非负实数.

如果 A 是对称 (半) 正定矩阵, 则可以定义其平方根, 即存在唯一的对称 (半) 正定矩阵 B, 使得 $B^2 = A$. 事实上, 我们有下面更一般的性质.

定理 设 $A \in \mathbb{R}^{n\times n}$ 是一个对称半正定矩阵, k 是一个给定的正整数. 则存在一个唯一的对称半正定矩阵 $B \in \mathbb{C}^{n\times n}$ 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) $\mathrm{rank}(B) = \mathrm{rank}(A)$, 因此, 若 A 是正定的, 则 B 也正定.

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

3 向量范数与矩阵范数

3.1 向量范数

定义 (向量范数) 若函数 $f(x): \mathbb{R}^n \to \mathbb{R}$ 满足

(1) $f(x) \geq 0, \forall\, x \in \mathbb{R}^n$ 且等号当且仅当 $x = 0$ 时成立; (非负性, nonnegativity)

(2) $f(\alpha x) = |\alpha| \cdot f(x), \forall\, x \in \mathbb{R}^n, \alpha \in \mathbb{R}$ (正齐次性, homogeneity)

(3) $f(x+y) \leq f(x) + f(y), \forall x, y \in \mathbb{R}^n$; (三角不等式, triangular inequality)

则称 $f(x)$ 为 $\mathbb{R}^n$ 上的 **范数** (**norm**), 通常记作 $\|x\|$.

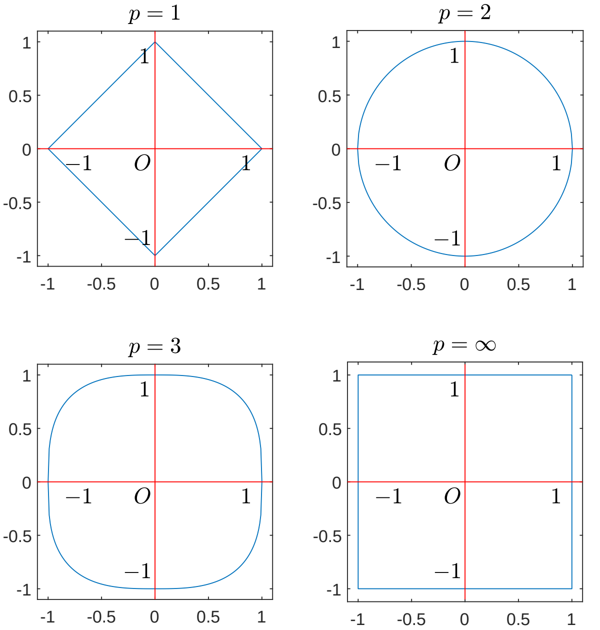
✍ 为简单起见, 我们这里只考虑实数域情形, 复数域情形完全类似.

例 $\mathbb{R}^n$ 上的常用范数: 设 $x=[x_1,x_2,\ldots,x_n]^{\mathsf{T}}\in\mathbb{R}^n$,

- 1-范数: $\|x\|_1=\sum_{i=1}^{n}|x_i|$;
- 2-范数: $\|x\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$;
- ∞-范数: $\|x\|_\infty=\max_{1\le i\le n}|x_i|$;
- p-范数: $\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}},\quad p\ge 1.$

✍ 1-范数, 2-范数, ∞-范数可以看作是 p-范数在 $p=1,2,\infty$ 时的特殊情形.

不同范数下的单位圆



向量范数的一些基本性质

定理 (向量范数的连续性) 设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的一个向量范数, 则 $f(x) \triangleq \|x\|$ 关于 x 的每个分量是连续的.

(证明留作课外练习, 利用范数的三角不等式)

定义 (范数的等价性) 设 $\|\cdot\|_\alpha$ 与 $\|\cdot\|_\beta$ 是 $\mathbb{R}^n$ 空间上的两个向量范数, 若存在正常数 c_1, c_2, 使得

$$c_1\|x\|_\alpha \leq \|x\|_\beta \leq c_2\|x\|_\alpha \quad \text{或} \quad c_1 \leq \frac{\|x\|_\beta}{\|x\|_\alpha} \leq c_2$$

对任意 $x \in \mathbb{R}^n$ 都成立, 则称 $\|\cdot\|_\alpha$ 与 $\|\cdot\|_\beta$ 是**等价**的.

定理 (向量范数的等价性) $\mathbb{R}^n$ 上的所有向量范数都是等价的, 特别地, 有

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

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

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

(板书, 以第一个为例, 其余留作练习)

3.2 矩阵范数

定义 (矩阵范数) 若函数 $f(X): \mathbb{R}^{n\times n} \to \mathbb{R}$ 满足

(1) $f(A) \geq 0, \forall\, A \in \mathbb{R}^{n\times n}$, 且等号当且仅当 $A = 0$ 时成立;

(2) $f(\alpha A) = |\alpha| \cdot f(A), \forall\, A \in \mathbb{R}^{n\times n}, \alpha \in \mathbb{R}$;

(3) $f(A+B) \leq f(A) + f(B), \forall\, A, B \in \mathbb{R}^{n\times n}$;

(4) $f(AB) \leq f(A) f(B), \forall\, A, B \in \mathbb{R}^{n\times n}$;

则称 $f(X)$ 为 $\mathbb{R}^{n\times n}$ 上的矩阵范数, 通常记作 $\|X\|$.

✍ 在矩阵范数的定义中, 条件 (4) 称为**相容性条件**.

✍ 类似地, 我们可以定义 $\mathbb{R}^{m\times n}$ 上的矩阵范数 (相容性条件会复杂一些).

由向量范数导出的算子范数

引理 (算子范数) 设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的向量范数, 则

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

是 $\mathbb{R}^{n\times n}$ 上的范数, 称为**算子范数**, 也称为**诱导范数**或**导出范数**.

(板书)

✍ 对于算子范数, 我们可以立即得到下面的性质: 设 $A \in \mathbb{R}^{n\times n}$, $x \in \mathbb{R}^n$, 则

$$\|Ax\| \leq \|A\| \cdot \|x\|. \tag{1.3}$$

如果没有特别指出哪类范数, 对向量和矩阵使用同样的范数时, 默认是指算子范数.

✍ 如果矩阵范数和向量范数满足 (1.3), 则称它们是**相容**的.

例 $\mathbb{R}^{n\times n}$ 上常见的矩阵范数:

- Frobenius 范数, 简称 F-范数

$$\|A\|_F \triangleq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}\ ;$$

- p-范数 (算子范数)

$$\|A\|_p \triangleq \sup_{x\neq 0}\frac{\|Ax\|_p}{\|x\|_p},\quad p=1,2,\infty.$$

✍ 可以证明: $\|A\|_F^2 = \mathrm{tr}(A^\intercal A)$, 留作练习.

定理 设 $A \in \mathbb{R}^{n\times n}$, 则

(1) 1-范数 (也称为 **列范数**): $\|A\|_1 = \max\limits_{1\le j\le n} \left(\sum_{i=1}^{n} |a_{ij}|\right)$;

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

(3) 2-范数 (也称为 **谱范数**): $\|A\|_2 = \sqrt{\rho(A^\intercal A)}$.

(板书, 以 ∞-范数和 2-范数为例, 1-范数留作练习)

例 设 $A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$, 计算 $\|A\|_1, \|A\|_2, \|A\|_\infty, \|A\|_F$.

(板书)

✍ 计算 2-范数时需要求谱半径, 因此通常比计算 1-范数和 ∞-范数更困难. 但在某些情况下可以用下面的范数等价性来估计一个矩阵的 2-范数.

定理 (矩阵范数的等价性) $\mathbb{R}^{n\times n}$ 上的所有范数都是等价的, 特别地, 有

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

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

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

(板书)

定理 (矩阵范数的连续性) 设 $\|\cdot\|$ 是 $\mathbb{R}^{n\times n}$ 上的一个矩阵范数, 则 $f(A) \triangleq \|A\|$ 关于 A 的每个分量是连续的.

(留作课外练习)

定理 设 $\|\cdot\|$ 是 $\mathbb{R}^{n\times n}$ 上的任一算子范数, 若 $\|B\|<1$, 则 $I\pm B$ 非奇异, 且

$$\left\|(I\pm B)^{-1}\right\| \le \frac{1}{1-\|B\|}.$$

(板书)

矩阵范数的更多性质

(1) 对任意算子范数 $\|\cdot\|$, 有 $\|A^k\| \leq \|A\|^k$;

(2) 对任意算子范数 $\|\cdot\|$, 有 $\|Ax\| \leq \|A\| \cdot \|x\|$, $\|AB\| \leq \|A\| \cdot \|B\|$;

(3) $\|Ax\|_2 \leq \|A\|_F \cdot \|x\|_2$, $\|AB\|_F \leq \|A\|_F \cdot \|B\|_F$;

(4) F-范数不是算子范数;

(5) $\|\cdot\|_2$ 和 $\|\cdot\|_F$ 是 **酉不变范数**, 即对任意正交矩阵 (或酉矩阵) U, V, 有

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

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

(6) $\|A^\intercal\|_2 = \|A\|_2, \quad \|A^\intercal\|_1 = \|A\|_\infty$.

(留作课外练习)

定理 (谱半径与范数的关系) 设 $A \in \mathbb{R}^{n\times n}$, 则对任意算子范数, 有 $\rho(A) \leq \|A\|$.

(板书)

✍ 事实上, 上述定理中的结论 (1) 对任意矩阵范数都成立.

4 Kronecker 积

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

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

- ✍ Kronecker 积也称为**直积** (direct product) 或**张量积** (tensor product).
- ✍ 任意两个矩阵都存在 Kronecker 积, 且 $A \otimes B$ 和 $B \otimes A$ 是同阶矩阵, 但通常 $A \otimes B \neq B \otimes A$.

定理 矩阵的 Kronecker 积有以下性质:

(1) $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B), \quad \forall\, \alpha \in \mathbb{R}$;

(2) $(A \otimes B)^{\intercal} = A^{\intercal} \otimes B^{\intercal}, \quad (A \otimes B)^* = A^* \otimes B^*$;

(3) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$;

(4) $(A + B) \otimes C = A \otimes C + B \otimes C$;

(5) $A \otimes (B + C) = A \otimes B + A \otimes C$;

(6) 混合积: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

(7) $$(A_1 \otimes A_2 \otimes \cdots \otimes A_k)(B_1 \otimes B_2 \otimes \cdots \otimes B_k) = (A_1B_1) \otimes (A_2B_2) \otimes \cdots \otimes (A_kB_k);$$

(8) $(A_1 \otimes B_1)(A_2 \otimes B_2) \cdots (A_k \otimes B_k) = (A_1A_2 \cdots A_k) \otimes (B_1B_2 \cdots B_k)$;

(9) $\operatorname{rank}(A \otimes B) = \operatorname{rank}(A) \operatorname{rank}(B)$;

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

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

(1) $\mathrm{tr}(A\otimes B)=\mathrm{tr}(A)\mathrm{tr}(B)$;

(2) $\det(A\otimes B)=\det(A)^n\det(B)^m$;

(3) $A\otimes I_n+I_m\otimes B$ 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 $(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$;

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

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

$A\otimes B$ 与 $B\otimes A$ 通常是不相等的, 但它们之间存在下面的关系式.

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

$$P^{\mathsf{T}}(A \otimes B)P = B \otimes A.$$

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

$$\mathrm{vec}(X)=[x_1^\intercal,x_2^\intercal,\ldots,x_n^\intercal]^\intercal,$$

则有

$$\mathrm{vec}(AX)=(I\otimes A)\mathrm{vec}(X),\quad \mathrm{vec}(XB)=(B^\intercal\otimes I)\mathrm{vec}(X),$$

以及

$$(A\otimes B)\mathrm{vec}(X)=\mathrm{vec}(BXA^\intercal).$$

Kronecker 积一个重要应用是可以将某些矩阵方程转化成一般的代数方程.

定理 矩阵方程

$$AX + XB = D$$

等价于代数方程

$$(I \otimes A + B^{\mathsf{T}} \otimes I)\text{vec}(X) = \text{vec}(D).$$