

第一讲 预备知识

线性代数基础



目录

1.1 数值分析引论

1.2 线性代数基础

1.3 数值计算中的误差

<https://math.ecnu.edu.cn/~jypan/Teaching/NA>

1-2 | 线性代数基础

1.2 线性代数基础

1.2.1 线性空间基本概念

1.2.2 矩阵特征值与谱半径

1.2.3 对称正定矩阵

1.2.4 向量范数与矩阵范数

1.2.5 内积与内积空间

1.2.6 矩阵标准型

1-2-1 | 线性空间基本概念

- 数域, 如: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$
- 线性空间, 如: $\mathbb{R}^n, \mathbb{C}^n, \mathbb{R}^{m \times n}, \mathbb{C}^{m \times n}, \mathbb{H}_n, C[a, b], C^p[a, b]$
- 线性相关与线性无关, 线性组合, 向量组的秩
- 线性空间的基, 维数
- 线性子空间 (简称子空间), 零空间 (核) $\text{Ker}(A)$

张成的线性空间

例 设 $\mathbb{S}$ 是数域 $\mathbb{F}$ 上的一个线性空间, $x_1, x_2, \dots, x_k \in \mathbb{S}$, 记

$$\text{span}\{x_1, x_2, \dots, x_k\} \triangleq \{ \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k : \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{F} \},$$

即由 $x_1, x_2, \dots, x_k$ 的所有线性组合构成的集合, 则 $\text{span}\{x_1, x_2, \dots, x_k\}$ 是 $\mathbb{S}$ 的一个线性子空间, 称为由 $x_1, x_2, \dots, x_k$ **张成的线性空间**.



$\text{span}(A) \rightarrow$ 由 A 的列向量所张成的线性空间, 即 A 的像空间/值域/列空间

1-2-2 | 矩阵特征值与谱半径

定义 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$), 若存在 $\lambda \in \mathbb{C}$ 和非零向量 $x, y \in \mathbb{C}^n$, 使得

$$Ax = \lambda x, \quad y^* A = \lambda y^*,$$

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

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;
- n 阶矩阵总是存在 n 个特征值 (其中可能有相等的), 通常记为 $\lambda_1, \lambda_2, \dots, \lambda_n$;
- 特征值也可以通过特征多项式的零点来定义;
- 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.

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;
- n 阶矩阵总是存在 n 个特征值 (其中可能有相等的), 通常记为 $\lambda_1, \lambda_2, \dots, \lambda_n$;
- 特征值也可以通过特征多项式的零点来定义;
- 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.

 **思考:** 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$), 则

- (1) A^* 和 A 的特征值和特征向量是什么关系?
- (2) A^{-1} 和 A , 它们的特征值和特征向量是什么关系?
- (3) 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? 有理函数呢?

谱半径

定义 (谱半径) 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$), 则称

$$\rho(A) \triangleq \max_{\lambda \in \sigma(A)} \{|\lambda|\}$$

为 A 的谱半径, 其中 $\sigma(A)$ 为 A 谱, 即所有特征值组成的集合:

$$\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}.$$

特征值与行列式和迹的关系

根据高次多项式的韦达定理, 我们可以得到下面的结论.

定理 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$), 称 $\text{tr}(A) \triangleq a_{11} + a_{22} + \cdots + a_{nn}$ 为矩阵 A 的迹 (trace), 有

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

特征值分解 (谱分解)

定义 (特征值分解) 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$). 若存在非奇异矩阵 $X \in \mathbb{C}^{n \times n}$, 使得

$$X^{-1}AX = \Lambda, \quad (1.1)$$

其中 $\Lambda \in \mathbb{C}^{n \times n}$ 是对角矩阵, 则称 A 是可对角化的, 矩阵 Λ 的对角线元素即为 A 的特征值, 分解 (1.1) 称为矩阵 A 的特征值分解或谱分解.

特征值分解 (谱分解)

定义 (特征值分解) 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$). 若存在非奇异矩阵 $X \in \mathbb{C}^{n \times n}$, 使得

$$X^{-1}AX = \Lambda, \quad (1.1)$$

其中 $\Lambda \in \mathbb{C}^{n \times n}$ 是对角矩阵, 则称 A 是**可对角化**的, 矩阵 Λ 的对角线元素即为 A 的特征值, 分解 (1.1) 称为矩阵 A 的**特征值分解**或**谱分解**.

定理 设 $A \in \mathbb{R}^{n \times n}$ (或 $\mathbb{C}^{n \times n}$), 则

- (1) A 可对角化当且仅当 A 有 n 个线性无关的特征向量;
 - (2) A 可对角化当且仅当 A 的所有特征值的代数重数与几何重数都相等.
- 特别地, 若 A 有 n 个互不相等的特征值, 则 A 可对角化.

对称矩阵的正交对角化

如果 A 是对称的, 则 A 的所有特征值都是实的, 而且可以正交对角化.

定理 设 $A \in \mathbb{R}^{n \times n}$ 对称, 则 A 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都是实数, 且存在正交矩阵 $Q \in \mathbb{R}^{n \times n}$, 使得

$$Q^T A Q = \Lambda \quad \text{或} \quad A = Q \Lambda Q^T,$$

其中 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 是实对角矩阵, Q 的第 i 列为 A 的对应于 λ_i 的特征向量.



该结论对复矩阵也成立, 此时 Q 是复的西矩阵, 但 Λ 仍然是实对角矩阵.

1-2-3 | 对称正定矩阵

定义 设 $A \in \mathbb{C}^{n \times n}$.

- 若对所有向量 $x \in \mathbb{C}^n$ 有 $\operatorname{Re}(x^*Ax) \geq 0$, 则称 A 是**半正定**的;
- 若对所有非零向量 $x \in \mathbb{C}^n$ 有 $\operatorname{Re}(x^*Ax) > 0$, 则称 A 是**正定**的;
- 若 A 是 Hermite 的且半正定, 则称 A 为 **Hermite 半正定**;
- 若 A 是 Hermite 的且正定, 则称 A 为 **Hermite 正定**;
- 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**.

 若对所有向量 $x \in \mathbb{C}^n$ 有 $x^*Ax \in \mathbb{R}$, 则 $A^* = A$.

正定矩阵的充要条件

定理 设 $A \in \mathbb{C}^{n \times n}$, 则 A 正定 (或半正定) 的充要条件是 $H = \frac{1}{2}(A + A^*)$ 正定 (或半正定). (留作练习)

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 $x^T A x \geq 0$). (留作练习)

定理 (Hermite 矩阵正定的充分条件) $A \in \mathbb{C}^{n \times n}$ 是 Hermite 正定矩阵的充要条件是 A 是 Hermite 的, 且所有特征值都大于零或者所有顺序主子式都大于零. (留作课外自习)

对称正定矩阵的更多性质

设 $A \in \mathbb{R}^{n \times n}$ 是对称正定矩阵, 则

- (1) A 非奇异, 且 A^{-1} 也对称正定;
- (2) A 的所有特征值都是正实数;
- (3) A 的所有顺序主子阵都对称正定的;
- (4) A 的所有顺序主子式都大于零.

 对于 Hermite 正定矩阵, 也有类似的结论.

平方根

如果 A 是 Hermite (半) 正定矩阵, 则可以定义其平方根, 即存在唯一的 Hermite (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 $A \in \mathbb{C}^{n \times n}$ 是一个 Hermite 半正定矩阵, k 是一个给定的正整数. 则存在一个唯一的 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) $\text{rank}(B) = \text{rank}(A)$, 因此, 若 A 是正定的, 则 B 也正定;
- (2) 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, 记为 $A^{\frac{1}{2}}$.

(留作课外自习, 证明可参见相关资料)

1-2-4 | 向量范数与矩阵范数

一般线性空间上的范数

定义 (范数与赋范线性空间) 设 S 为数域 F (F 可以是 $\mathbb{R}$ 或 $\mathbb{C}$) 上的线性空间, 若对任意 $x \in S$, 存在唯一实数与之对应, 记为 $\|x\|$, 满足:

- (1) $\|x\| \geq 0$, 等号当且仅当 $x = 0$ 时成立; (正定性)
- (2) $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\forall \alpha \in F$; (正齐次性)
- (3) $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in S$; (三角不等式)

则称 $\|\cdot\|$ 为线性空间 S 上的**范数**, 定义了范数的线性空间称为**赋范线性空间**.



范数是从 S 到 $\mathbb{R}_+ \cup \{0\}$ 的**一元函数**.

举例

例 $C[a, b]$ 上的常用范数: 设 $f(x) \in C[a, b]$

➤ 1-范数: $\|f\|_1 = \int_a^b |f(x)| \, dx;$

➤ 2-范数: $\|f\|_2 = \left(\int_a^b |f(x)|^2 \, dx \right)^{\frac{1}{2}};$

➤ p -范数: $\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}};$ (p 为正整数)

➤ ∞ -范数: $\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|.$

向量范数

 定义在向量空间 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 上的范数就是**向量范数**.

例 $\mathbb{C}^n$ 和 $\mathbb{R}^n$ 中常见的向量范数:

➤ 1-范数:
$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

➤ 2-范数:
$$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$$

➤ ∞ -范数:
$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

➤ p -范数:
$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty$$

范数的连续性

为简单起见, 我们下面只考虑实数域情形, 复数域情形完全类似.

定理 设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的向量范数, 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)

范数的等价性

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

$$c_1\|x\|_\alpha \leq \|x\|_\beta \leq c_2\|x\|_\alpha$$

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, 只证明等价性, 三个不等式留作练习)



事实上, 任意有限维赋范线性空间上的所有范数都是等价的

矩阵范数

矩阵范数除了满足一般范数的要求外,通常还要求满足相容性条件.

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

 条件 (4) 称为**相容性条件**. 有的教材中, 只需满足 (1), (2), (3), 并称满足 (4) 的矩阵范数为**相容矩阵范数**. 如果没有特别说明, 本讲义中矩阵范数就是指相容矩阵范数.

算子范数

一类常用的矩阵范数就是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

$$\|Ax\| \leq \|A\| \cdot \|x\|. \quad (1.2)$$

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.

 如果存在矩阵范数和向量范数满足 (1.2), 则称它们是**相容**的.

常用矩阵范数

例 $\mathbb{R}^{n \times n}$ 上常见的矩阵范数:

➤ p -范数 (算子范数)

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

➤ Frobenius 范数, 简称 F -范数

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

算子范数的计算公式

定理 可以证明:

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

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

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

(4) f -范数: $\|A\|_F = \sqrt{\text{tr}(A^\top A)} .$

(板书板书, 以 ∞ -范数和 2-范数为例, 1-范数和 F -范数留作练习)

矩阵范数的一些基本性质

- (1) 对任意范数 $\|\cdot\|$, 有 $\|I\| \geq 1$, 对任意的算子范数 $\|\cdot\|$, 有 $\|I\| = 1$;
- (2) 对任意范数 $\|\cdot\|$, 有 $\|A^k\| \leq \|A\|^k$;
- (3) $\|A^T\|_2 = \|A\|_2$, $\|A^T\|_1 = \|A\|_\infty$;
- (4) 若 A 是正规矩阵, 则 $\|A\|_2 = \rho(A)$, 特别地, 若 A 是对称矩阵, 则 $\|A\|_2 = \rho(A)$;
- (5) F -范数不是算子范数;
- (6) $\|\cdot\|_2$ 和 $\|\cdot\|_F$ 是酉不变范数, 即对任意正交矩阵 (或酉矩阵) U, V , 有

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

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

(留作课外自习)

举例

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范数等价性来估计一个矩阵的 2-范数.

矩阵范数的等价性

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, 等价性证明略, 只证明第一个不等式, 利用矩阵特征值与迹的关系式)

矩阵范数的等价性

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

$$\begin{aligned}\frac{1}{\sqrt{n}} \|A\|_1 &\leq \|A\|_2 \leq \sqrt{n} \|A\|_1, \\ \frac{1}{\sqrt{n}} \|A\|_\infty &\leq \|A\|_2 \leq \sqrt{n} \|A\|_\infty, \\ \frac{1}{n} \|A\|_\infty &\leq \|A\|_1 \leq n \|A\|_\infty.\end{aligned}$$

(板书, 等价性证明略, 只证明第一个不等式, 利用矩阵特征值与迹的关系式)

更多不等式

$$\frac{1}{\sqrt{n}} \|A\|_F \leq \|A\|_1 \leq \sqrt{n} \|A\|_F; \quad \|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2.$$

(留作练习)

矩阵范数的连续性

定理 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.
(留作课外自习)

矩阵范数的更多性质

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

$$\|(I \pm B)^{-1}\| \leq \frac{1}{1 - \|B\|}.$$

(板书)

谱半径与范数之间的关系

定理 (谱半径与范数的关系) 设 $A \in \mathbb{C}^{n \times n}$, 则

- (1) 对任意算子范数, 有 $\rho(A) \leq \|A\|$;
 - (2) 反之, 对任意 $\varepsilon > 0$, 都存在一个算子范数 $\|\cdot\|_\varepsilon$, 使得 $\|A\|_\varepsilon \leq \rho(A) + \varepsilon$, 其中范数 $\|\cdot\|_\varepsilon$ 依赖于 A 和 ε . 所以, 若 $\rho(A) < 1$, 则存在算子范数 $\|\cdot\|_\varepsilon$, 使得 $\|A\|_\varepsilon < 1$;
- (板书, 只证明结论 (1), 结论 (2) 可查询相关资料)

 事实上, 上述定理中的结论 (1) 对任意矩阵范数都成立.

推论 设 $A \in \mathbb{R}^{n \times n}$, 则 $\|A\|_2^2 \leq \|A\|_1 \cdot \|A\|_\infty$, 且

$$\max_{1 \leq i, j \leq n} \{|a_{ij}|\} \leq \|A\|_2 \leq n \max_{1 \leq i, j \leq n} \{|a_{ij}|\}.$$

(留作练习)

推论 设 $A \in \mathbb{R}^{n \times n}$, 则对任意矩阵范数 $\|\cdot\|$, 有

$$\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}}. \quad (1.3)$$

(留作课外自习, 证明见讲义)

 等式 (1.3) 有时也用来定义矩阵的谱半径.

1-2-5 | 内积与内积空间

定义 (内积与内积空间) 设 S 是数域 F ($\mathbb{R}$ 或 $\mathbb{C}$) 上的一个线性空间, 定义一个从 $S \times S$ 到 F 的代数运算, 记为 “ $(\cdot, \cdot)$ ”, 即对任意 $x, y \in S$, 都存在唯一的 $f \in F$, 使得 $f = (x, y)$. 如果该运算满足

- (1) $(x, y) = \overline{(y, x)}, \quad \forall x, y \in S;$
- (2) $(\alpha x, y) = \alpha(x, y), \quad \forall \alpha \in F, x, y \in S;$
- (3) $(x + y, z) = (x, z) + (y, z), \quad \forall x, y, z \in S;$
- (4) $(x, x) \geq 0$, 等号当且仅当 $x = 0$ 时成立;

则称 $(\cdot, \cdot)$ 为 S 上的一个**内积**, 有时也称**标量积**. 定义了内积的线性空间称为**内积空间**.

 定义在 $\mathbb{R}$ 上的内积空间称为**欧氏空间**, 定义在 $\mathbb{C}$ 上的内积空间称为**酉空间**.

 内积可以看作是从线性空间 S 到数域 F 的**二元函数**.

内积基本性质

- $(u, \beta v) = \bar{\beta}(u, v), \forall \beta \in \mathbb{F}$
- $(u, v + w) = (u, v) + (u, w)$
- $(u_1 + u_2 + \cdots + u_n, v) = (u_1, v) + (u_2, v) + \cdots + (u_n, v)$

举例: $\mathbb{R}^n/\mathbb{C}^n$ 上的内积

例 考虑线性空间 $\mathbb{R}^n$, 对任意 $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$, $y = [y_1, y_2, \dots, y_n]^T \in \mathbb{R}^n$, 定义

$$(x, y) \triangleq y^T x = \sum_{i=1}^n x_i y_i,$$

则 (x, y) 是 $\mathbb{R}^n$ 上的内积, 因此 $\mathbb{R}^n$ 是一个欧氏空间. 这种方式定义的内积称为**欧氏内积**.

 如果是线性空间 $\mathbb{C}^n$, 则定义内积 $(x, y) \triangleq y^* x = \sum_{i=1}^n x_i \bar{y}_i$. 此时 $\mathbb{C}^n$ 是一个酉空间.

 需要指出的是, 同一个线性空间中可以定义多个内积.

 以上定义的内积是 $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 上的标准内积, 若不特别声明, 我们一般采用标准内积.

举例: $C[a, b]$ 上的内积

例 考虑实数域 $\mathbb{R}$ 上的线性空间 $C[a, b]$, 对任意 $f(x), g(x) \in C[a, b]$, 定义

$$(f, g) \triangleq \int_a^b f(x)g(x) \, dx,$$

则 (f, g) 是 $C[a, b]$ 上的内积, 此时 $C[a, b]$ 是一个欧氏空间.

正交

有了内积以后, 我们就可以定义正交.

定义 设 $\mathbb{S}$ 是内积空间, $u, v \in \mathbb{S}$, 若 $(u, v) = 0$, 则称 u 与 v 是**正交**的.



正交可以看作是垂直概念的推广.

Cauchy-Schwartz 不等式

定理 设 $(\cdot, \cdot)$ 是 $\mathbb{R}^n$ 上的内积, 则对任意 $x, y \in \mathbb{R}^n$, 有

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

(板书)



该结论对 $\mathbb{C}^n$ 也成立.



思考: 等号成立的条件是什么?

Gram 矩阵

定理 设 $\mathbb{S}$ 是内积空间, $u_1, u_2, \dots, u_n \in \mathbb{S}$, 则 $u_1, u_2, \dots, u_n$ 线性无关的充要条件是矩阵 G 非奇异, 其中

$$G = \begin{bmatrix} (u_1, u_1) & (u_2, u_1) & \cdots & (u_n, u_1) \\ (u_1, u_2) & (u_2, u_2) & \cdots & (u_n, u_2) \\ \vdots & \vdots & & \vdots \\ (u_1, u_n) & (u_2, u_n) & \cdots & (u_n, u_n) \end{bmatrix}.$$

这个矩阵就称为 $u_1, u_2, \dots, u_n$ 的 **Gram 矩阵**.

(留作课外自习, 证明见讲义)

内积与范数

设 $\mathbb{S}$ 是内积空间, 对任意 $u \in \mathbb{S}$, 定义

$$\|u\| \triangleq (u, u)^{\frac{1}{2}},$$

则可以验证, $\|u\|$ 是 $\mathbb{S}$ 上的范数. 这就是由内积导出的范数. 即任意一个内积都可以定义一个范数. (留作练习)

例 $\mathbb{R}^n$ 上由标准内积导出的范数为

$$\|x\| = (x, x)^{\frac{1}{2}} = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}.$$

这就是 2-范数.

带权内积

设 $\omega_1, \omega_2, \dots, \omega_n$ 为任意给定的正实数, 对任意 $x, y \in \mathbb{R}^n$, 定义

$$(x, y)_\omega \triangleq \sum_{i=1}^n \omega_i x_i y_i .$$

可以验证, $(x, y)_\omega$ 是 $\mathbb{R}^n$ 上的内积, 我们称其为**带权内积**, 其中 ω_i 称为**权系数**.



同样, 可以在 $\mathbb{C}^n$ 中定义带权内积.

权函数

为了在 $C[a, b]$ 上定义带权内积, 我们首先定义权函数.

定义 设 $\omega(x)$ 是 $[a, b]$ 上的非负函数, 如果 $\omega(x)$ 满足

- (1) 对任意非负整数 k , $\int_a^b x^k \omega(x) dx$ 存在且为有限值;
- (2) 对 $[a, b]$ 上的任意非负连续函数 $g(x)$, 如果 $\int_a^b g(x) \omega(x) dx = 0$, 则 $g(x) \equiv 0$;
则称 $\omega(x)$ 是 $[a, b]$ 上的一个**权函数**.

 $[a, b]$ 可以是有限区间, 也可以是无限区间, 即 a 可以是 $-\infty$, b 可以是 ∞ ;

 权函数与定义区间相关.

举例: 常见的权函数

例 常见的权函数有

➤ $\omega(x) = 1$ 是 $[-1, 1]$ 上的权函数;

➤ $\omega(x) = \frac{1}{\sqrt{1-x^2}}$ 是 $[-1, 1]$ 上的权函数;

➤ $\omega(x) = e^{-x}$ 是 $[0, \infty)$ 上的权函数;

➤ $\omega(x) = e^{-x^2}$ 是 $(-\infty, \infty)$ 上的权函数.

举例: 带权内积

例 设 $\omega(x)$ 是 $[a, b]$ 上的权函数, 对任意 $f(x), g(x) \in C[a, b]$, 定义

$$(f, g)_\omega = \int_a^b \omega(x) f(x) g(x) \, dx,$$

则 $(f, g)_\omega$ 是 $C[a, b]$ 上的带权内积. 对应的带权内积导出范数为

$$\|f\|_\omega = (f, f)_\omega^{\frac{1}{2}} = \left(\int_a^b \omega(x) f^2(x) \, dx \right)^{\frac{1}{2}}.$$

特别地, 当 $\omega(x) \equiv 1$ 时, 内积导出范数为

$$\|f\| = (f, f)^{\frac{1}{2}} = \left(\int_a^b f^2(x) \, dx \right)^{\frac{1}{2}},$$

即为 $f(x)$ 的 2-范数.

1-2-6 | 矩阵标准型

计算矩阵特征值的一个基本思想

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.

两个非常有用的特殊矩阵是 Jordan 标准型和 Schur 标准型 / Schur 分解

Jordan 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.

另一种描述: 设 $A \in \mathbb{C}^{n \times n}$, 则存在非奇异矩阵 $X \in \mathbb{C}^{n \times n}$ 使得

$$X^{-1}AX = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_\ell \end{bmatrix},$$

其中

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \\ & & & \lambda_k \end{bmatrix}_{n_k \times n_k}, \quad n_k \geq 1, k = 1, 2, \dots, \ell, \text{ 且 } \sum_{k=1}^{\ell} n_k = n.$$

矩阵 J_k 就称为 **Jordan 块**, $\lambda_1, \lambda_2, \dots, \lambda_\ell$ 是 A 的特征值 (不要求互异).

另一种描述: 设 $A \in \mathbb{C}^{n \times n}$, 则存在非奇异矩阵 $X \in \mathbb{C}^{n \times n}$ 使得

$$X^{-1}AX = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_\ell \end{bmatrix},$$

其中

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \\ & & & \lambda_k \end{bmatrix}_{n_k \times n_k}, \quad n_k \geq 1, k = 1, 2, \dots, \ell, \text{ 且 } \sum_{k=1}^{\ell} n_k = n.$$

矩阵 J_k 就称为 **Jordan 块**, $\lambda_1, \lambda_2, \dots, \lambda_\ell$ 是 A 的特征值 (不要求互异).

 Jordan 标准型在理论研究中非常有用, 但数值计算比较困难.

 由 Jordan 标准型可知, 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 / Schur 分解

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

$$U^*AU = \begin{bmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ 0 & \lambda_2 & \cdots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \triangleq R \quad \text{或} \quad A = URU^*, \quad (1.4)$$

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

(板书)

关于 Schur 标准型的几点说明

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 (1.4) 中的 R 是对角矩阵;
- A 是 Hermite 矩阵当且仅当 (1.4) 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

$$Q^T A Q = T,$$

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.

(留作课外自习)

谢谢
THANK YOU

