第一讲 科学计算引论

线性代数基础

目录

https://math.ecnu.edu.cn/~jypan/Teaching/SC

1-1 线性代数基础

1.2 线性代数基础

1-1-1 线性空间基本概念

- 数域, 如: $\mathbb{R}$, $\mathbb{C}$
- 线性空间, 如: $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$, $C[a,b]$, $C^p[a,b]$
- 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)
- 像空间 (列空间, 值域) $\mathrm{Ran}(A)$, 零空间 (核空间) $\mathrm{Ker}(A)$, 行空间 $\mathrm{Ran}(A^{\top})$, 左零空间 $\mathrm{Ker}(A^{\top})$

张成的线性空间

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

$$\operatorname{span}\{x_1, x_2, \ldots, x_k\} \triangleq \left\{ \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \ : \ \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F} \right\},$$

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

$\operatorname{span}(A) \rightarrow$ 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

1-1-2 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

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

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 $\lambda_1, \lambda_2, \ldots, \lambda_n$;
- 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.

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:

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

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 $\mathrm{tr}(A) \triangleq a_{11} + a_{22} + \cdots + a_{nn}$ 为矩阵 A 的**迹** (trace), 有

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

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

$$X^{-1}AX = \Lambda, \tag{1.1}$$

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**.

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 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 都是实数, 且存在正交矩阵 $Q \in \mathbb{R}^{n\times n}$ (即 $Q^{\mathsf{T}}Q = QQ^{\mathsf{T}} = I$), 使得

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

其中 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 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 是复的酉矩阵, 但 Λ 仍然是实对角矩阵.

1-1-3 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 $\mathbb{S}$ 为数域 $\mathbb{R}$ 上的线性空间, 若对任意 $x \in \mathbb{S}$, 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) $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\forall\ \alpha \in \mathbb{F}$; (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 $\mathbb{S}$ 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 $\mathbb{S}$ 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: $\|f\|_1=\displaystyle\int_a^b |f(x)|\,\mathrm{d}x$;
- 2-范数: $\|f\|_2=\left(\displaystyle\int_a^b |f(x)|^2\,\mathrm{d}x\right)^{\frac{1}{2}}$;
- p-范数: $\|f\|_p=\left(\displaystyle\int_a^b |f(x)|^p\,\mathrm{d}x\right)^{\frac{1}{p}}$; ($p$ 为正整数)
- ∞-范数: $\|f\|_\infty=\max\limits_{a\leq x\leq b}|f(x)|$.

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

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

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

$$
\begin{aligned}
\|x\|_2 &\le \|x\|_1 \le \sqrt{n}\ \|x\|_2, \\
\|x\|_\infty &\le \|x\|_1 \le n\ \|x\|_\infty, \\
\|x\|_\infty &\le \|x\|_2 \le \sqrt{n}\ \|x\|_\infty.
\end{aligned}
$$

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 $f(X): \mathbb{R}^{n\times n} \to \mathbb{R}$ 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 $\|X\|$.

✍ 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**.

✍ 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

$$\|Ax\| \leq \|A\| \cdot \|x\|. \tag{1.2}$$

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**): $\|A\|_1 = \max\limits_{1\le j\le n}\left(\sum\limits_{i=1}^{n}|a_{ij}|\right)$;

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

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

(4) F-范数: $\|A\|_F = \sqrt{\mathrm{tr}(A^\intercal A)}$.

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) $\|A^{\mathsf{T}}\|_2 = \|A\|_2, \quad \|A^{\mathsf{T}}\|_1 = \|A\|_\infty$;

(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

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

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

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, $6, \sqrt{15+\sqrt{221}}, 7, \sqrt{30}$)

✍ 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

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

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

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

更多不等式

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

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 $\|A\|_\varepsilon < 1$;

✍ 事实上, 上述定理中的结论 (1) 对任意矩阵范数都成立.

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 $\|A\|_\varepsilon < 1$;

✍ 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 $\|A\|_2^2 \leq \|A\|_1 \cdot \|A\|_\infty$.

1-1-4 内积与内积空间

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

(1) $(x,y)=\overline{(y,x)},\quad \forall\ x,y\in\mathbb{S}$;

(2) $(\alpha x,y)=\alpha(x,y),\quad \forall\ \alpha\in\mathbb{R},\ x,y\in\mathbb{S}$;

(3) $(x+y,z)=(x,z)+(y,z),\quad \forall\ x,y,z\in\mathbb{S}$;

(4) $(x,x)\geq 0$, 等号当且仅当 $x=0$ 时成立;

则称 $(\cdot,\cdot)$ 为 $\mathbb{S}$ 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**.

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 $x=[x_1,x_2,\dots,x_n]^{\mathsf T}\in\mathbb{R}^n$, $y=[y_1,y_2,\dots,y_n]^{\mathsf T}\in\mathbb{R}^n$, 定义

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

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 $(x,y)\triangleq y^*x=\sum_{i=1}^{n}x_i\bar{y}_i$.

✍ 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

$$(f,g) \triangleq \int_a^b f(x)g(x)\,\mathrm{d}x,$$

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

$$\left|(x,y)\right|^2 \le (x,x)\,(y,y).$$

证明概要. 对任意 α 有

$$0 \le (x-\alpha y, x-\alpha y) = (x,x) - \alpha(y,x) - \bar{\alpha}(x,y) + |\alpha|^2 (y,y).$$

取 $\alpha = \dfrac{(x,y)}{(y,y)}$, 代入即可得结论成立.

该结论对 $\mathbb{C}^n$ 上的内积也成立.

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.

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.

谢谢

THANK YOU