# 第一章 距离和范数（1）

## 常见矩阵

### 1. 实对称矩阵

- **定义**：一个实矩阵（元素均为实数）的转置等于其自身，即 $A^T = A$。实对称矩阵的特征值都是实数，且特征向量相互正交。
- **例子**：

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

验证：$A^T = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = A$。

### 2. Hermit矩阵

- **定义**：一个复矩阵的共轭转置等于其自身，即 $A^H = A$，其中 $A^H$ 表示共轭转置（先取共轭再转置）。Hermit矩阵的特征值都是实数，且特征向量相互正交。
- **例子**：

$$B = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$$

验证：共轭转置 $B^H = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}^T = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix} = B$。

### 3. 正交矩阵

- **定义**：一个实矩阵的转置等于其逆，即 $Q^TQ = I$ 或 $Q^T = Q^{-1}$。正交矩阵的行和列都是正交单位向量，保持向量长度和夹角不变。
- **例子**：

$$Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

验证：$Q^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，且 $Q^TQ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$.

## 4. 酉矩阵

- **定义**：一个复矩阵的共轭转置等于其逆，即 $U^HU = I$ 或 $U^H = U^{-1}$。酉矩阵是正交矩阵的复推广，保持内积不变。
- **例子**：

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix}$$

验证：共轭转置 $U^H = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}$，且
$U^HU = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$

## 5. 正规矩阵

- **定义**：一个矩阵与它的共轭转置可交换，即 $AA^H = A^HA$。正规矩阵包括实对称矩阵、Hermit矩阵、正交矩阵、酉矩阵等。正规矩阵可以被酉对角化。
- **例子**：

$$C = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

验证：共轭转置 $C^H = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$（因为实矩阵，共轭转置即转置），计算
$CC^H = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$，且
$C^HC = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$，所以 $CC^H = C^HC.$

---

# 矩阵范数

## 1. 矩阵的 1 范数（列和范数）

- **定义**：矩阵的 1 范数是所有列向量的 1 范数的最大值（即 **所有列元素绝对值之和最大值**）。数学表示为：

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$$

它反映了矩阵列方向上的最大扩展程度。

- **例子**：对于 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$：
    - 第一列绝对值和：$|1| + |3| = 4$
    - 第二列绝对值和：$|2| + |4| = 6$
    - 因此，$\|A\|_1 = \max(4, 6) = 6.$

## 2. 矩阵的无穷范数（行和范数）

- **定义**：矩阵的无穷范数是所有行向量的 1 范数的最大值（即 **所有行元素绝对值之和最大值**）。数学表示为：

  $$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

  它反映了矩阵行方向上的最大扩展程度。

- **例子**：对于 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$：
  - 第一行绝对值和：$|1| + |2| = 3$
  - 第二行绝对值和：$|3| + |4| = 7$
  - 因此，$\|A\|_\infty = \max(3, 7) = 7.$

## 3. 矩阵的 2 范数（谱范数）

- **定义**：矩阵的 2 范数是矩阵的最大奇异值，即矩阵 $A^T A$ **的最大特征值的平方根**。数学表示为：

  $$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

  它反映了矩阵在向量变换中的最大拉伸倍数。

- **例子**：对于 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$：
  - 计算 $A^T A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}.$
  - 求特征值：解特征方程 $\det(A^T A - \lambda I) = 0$，即 $\det \begin{pmatrix} 10 - \lambda & 14 \\ 14 & 20 - \lambda \end{pmatrix} = (10 - \lambda)(20 - \lambda) - 196 = \lambda^2 - 30\lambda + 4 = 0.$
  - 特征值为 $\lambda = 15 \pm \sqrt{221}$，最大特征值 $\lambda_{\max} = 15 + \sqrt{221}.$
  - 因此，$\|A\|_2 = \sqrt{15 + \sqrt{221}} \approx 5.465$（数值近似）。

## 4. 矩阵的 F 范数（Frobenius 范数）

- **定义**：矩阵的 F 范数是 **所有元素绝对值的平方和的平方根**。数学表示为：

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

  它类似于向量的欧几里得范数的扩展，常用于衡量矩阵的整体大小。

- **例子**：对于 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$：
  - 计算元素平方和：$1^2 + 2^2 + 3^2 + 4^2 = 1 + 4 + 9 + 16 = 30.$
  - 因此，$\|A\|_F = \sqrt{30} \approx 5.477.$

# 向量范数

假设有一个 n 维列向量 **x**：

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

## 1. 向量的 1 范数 ($L^1$ Norm) - 曼哈顿范数

- **定义**：所有向量元素绝对值的和。

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

- **几何意义**：在二维空间中，它表示从原点到点 ($x_1$, $x_2$) 的曼哈顿距离（只能沿坐标轴方向行走的总距离）。
- **例子**：对于向量 $\mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$：

$$\|\mathbf{x}\|_1 = |1| + |-2| + |3| = 1 + 2 + 3 = 6$$

## 2. 向量的 2 范数 ($L^2$ Norm) - 欧几里得范数

- **定义**：向量元素平方和的平方根，是最常见的向量长度概念。

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

- **几何意义**：在二维/三维空间中，这就是我们熟悉的欧几里得距离。
- **例子**：对于向量 $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$：

$$\|\mathbf{x}\|_2 = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$

## 3. 向量的无穷范数 ($L^\infty$ Norm) - 最大范数

- **定义**：所有向量元素绝对值的最大值。

$$\|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |x_i| = \max(|x_1|, |x_2|, \ldots, |x_n|)$$

- **几何意义**：在二维空间中，它表示从原点到点 ($x_1$, $x_2$) 的切比雪夫距离。
- **例子**：对于向量 $\mathbf{x} = \begin{pmatrix} 1 \\ -5 \\ 3 \end{pmatrix}$：

$$\|\mathbf{x}\|_\infty = \max(|1|, |-5|, |3|) = \max(1, 5, 3) = 5$$

---

# 特征值计算

## 核心概念：特征值与特征向量

对于一个 **n x n** 的方阵 **A**，如果存在一个非零向量 **v** 和一个标量 **λ**，使得：
**A v = λ v**
那么：

- **λ** 被称为矩阵 **A** 的一个 **特征值**。
- **v** 被称为对应于 **λ** 的 **特征向量**。

---

## 示例 1：

计算矩阵 $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ 的特征值。

1. **写出特征方程**：$\det \begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix} = 0$
2. **计算行列式**：$(4-\lambda)(3-\lambda) - (1)(2) = 0$
3. **展开整理**：
   $(12 - 4\lambda - 3\lambda + \lambda^2) - 2 = 0$
   $\lambda^2 - 7\lambda + 10 = 0$
4. **解二次方程**：
   $(\lambda - 2)(\lambda - 5) = 0$
   $\lambda_1 = 2, \quad \lambda_2 = 5$

所以，矩阵 **A** 的特征值是 **2** 和 **5**。

## 示例 2：

计算矩阵 $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ 的特征值。

1. **写出特征方程**：
   $$\det \begin{pmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{pmatrix} = 0$$
2. **计算行列式（使用余子式展开第一行）**：
   $$\begin{aligned} & (2-\lambda) \cdot \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 1 \\ 0 & 2-\lambda \end{pmatrix} + 0 \\ = & (2-\lambda)\left[(2-\lambda)(2-\lambda) - (1)(1)\right] - 1\left[(1)(2-\lambda) - (1)(0)\right] \\ = & (2-\lambda)\left[(2-\lambda)^2 - 1\right] - (2-\lambda) \end{aligned}$$
3. **展开并整理**：
   提取公因式 $(2-\lambda)$：
   $(2-\lambda)\left[(2-\lambda)^2 - 1 - 1\right] = (2-\lambda)\left[(2-\lambda)^2 - 2\right]$
   展开括号：
   $(2-\lambda)(\lambda^2 - 4\lambda + 4 - 2) = (2-\lambda)(\lambda^2 - 4\lambda + 2)$
   最终得到特征多项式：
   $-\lambda^3 + 6\lambda^2 - 10\lambda + 4 = 0$

两边乘以 -1：
$\lambda^3 - 6\lambda^2 + 10\lambda - 4 = 0$

4. **解三次方程**：
   - **试根**：常数项是 -4，因子有 ±1, ±2, ±4。
     尝试 λ = 1: $1 - 6 + 10 - 4 = 1$ (不为0)
     尝试 λ = 2: $8 - 24 + 20 - 4 = 0$ (成立！)
     所以 $\lambda_1 = 2$ 是一个特征值。
   - **多项式除法**：将 $(\lambda^3 - 6\lambda^2 + 10\lambda - 4)$ 除以 $(\lambda - 2)$。
     $(\lambda^3 - 6\lambda^2 + 10\lambda - 4) \div (\lambda - 2) = \lambda^2 - 4\lambda + 2$
   - **解二次方程**：解 $\lambda^2 - 4\lambda + 2 = 0$
     使用求根公式：
     $\lambda = \frac{4 \pm \sqrt{16-8}}{2} = \frac{4 \pm \sqrt{8}}{2} = \frac{4 \pm 2\sqrt{2}}{2} = 2 \pm \sqrt{2}$

所以，矩阵 **A** 的三个特征值是：
$\lambda_1 = 2, \quad \lambda_2 = 2 + \sqrt{2}, \quad \lambda_3 = 2 - \sqrt{2}$

---

# 谱半径

## 定义

> 对于一个 $n \times n$ 的方阵 $A$，其所有特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$ (这些特征值构成了矩阵的"谱") 的绝对值的最大值，称为矩阵 $A$ 的谱半径。

记作 $\rho(A)$ (读作 "rho A")：
$\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$

**简单来说，谱半径就是矩阵所有特征值中，绝对值最大的那个。**

## 示例

2x2 矩阵例子：$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$

求得它的特征值为 $\lambda_1 = 2$ 和 $\lambda_2 = 5$。

那么，它的谱半径计算如下：

- $|\lambda_1| = |2| = 2$
- $|\lambda_2| = |5| = 5$

取最大值：
$\rho(A) = \max(2, 5) = 5$

所以，矩阵 $A$ 的谱半径是 **5**。

## 性质

- **与矩阵范数的关系**
  对于任意矩阵范数 $\|\cdot\|$，谱半径永远小于或等于该矩阵的范数：
  $\rho(A) \leq \|A\|$

- **对于特殊矩阵**：如果矩阵 $A$ 是**对称矩阵**（或埃尔米特矩阵），那么它的谱半径就等于它的**2-范数**（即最大的奇异值），即 $\rho(A) = \|A\|_2$。
- **矩阵的收敛性**：矩阵 $A$ 为收敛矩阵的充要条件为 $A$ 的谱半径 $\rho(A) < 1$。

---

# 算子范数

算子范数的核心思想是：**用一个已知的向量范数 $\|\cdot\|$ 来“诱导”或“定义”一个矩阵范数。**它将矩阵 $A$ 视为一个线性算子，并衡量这个算子对向量长度的最大“拉伸”倍数。

## 数学定义

对于一个 $m \times n$ 的矩阵 $A$，由其向量范数 $\|\cdot\|_p$ 所诱导的算子范数定义为：

$$\|A\|_p = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p = 1} \|A\mathbf{x}\|_p$$

**几何解释**：

- 分子 $\|A\mathbf{x}\|_p$：表示向量 $\mathbf{x}$ 被矩阵 $A$ 变换后的新长度。
- 分母 $\|\mathbf{x}\|_p$：表示向量 $\mathbf{x}$ 的原始长度。
- 这个比值衡量了变换的“放大率”。
- 算子范数就是在所有非零向量中，找到这个“放大率”的最大值。等价地，就是在所有长度为1的向量（单位球面）上，找到被变换后的向量的最大长度。

## 重要性质

1. **相容性（次乘性）**：

   $$\|A\mathbf{x}\|_p \leq \|A\|_p \cdot \|\mathbf{x}\|_p$$

   这个性质是算子范数定义的直接结果，也是其核心价值所在。
2. **次乘性（对于矩阵乘法）**：

   $$\|AB\|_p \leq \|A\|_p \cdot \|B\|_p$$

   这个性质在分析算法误差和系统稳定性时至关重要。
3. **定义的一致性**：对于单位矩阵 $I$，有 $\|I\|_p = 1$。

---

# 例题

## 例题 1

> 已知 $\|A\|_1 = 4$，求 $\max_{\|\mathbf{x}\|_1 = 4} \|A\mathbf{x}\|_1$

解答：

$$\max_{\|\mathbf{x}\|_1 = 4} \|A\mathbf{x}\|_1 = \max_{\|\frac{\mathbf{x}}{4}\|_1 = 1} \|A\mathbf{x}\|_1 = 4 \cdot \max_{\|\frac{\mathbf{x}}{4}\|_1 = 1} \|A \cdot \frac{\mathbf{x}}{4}\|_1 = 4 \cdot \|A\|_1 = 16$$

---

## 例题 2

> 设 $V$ 为酉矩阵，$A=\begin{bmatrix}-1 & i & 0\\ -i & 0 & -i\\ 0 & i & -1\end{bmatrix}$，求 $\max_{\|V\mathbf{x}\|_2=3}\|A\mathbf{x}\|_2$

解答：

由于 $A=A^H$，故 $A$ 为 Hermite 矩阵，属于正规矩阵。因此，有 $\|A\|_2=\rho(A)$，即矩阵 $A$ 的 2 范数为 $A$ 的谱半径。谱半径即 $A$ 的特征值中绝对值最大的特征值的绝对值。

求 $A$ 的特征值，解特征方程 $|A-\lambda I|=0$ 得到特征值分别为 $\lambda_1=-2,\lambda_2=-1,\lambda_3=1$。

因此，$\|A\|_2=\rho(A)=2$。

设 $V\mathbf{x}=\mathbf{y}$，由于 $V$ 为酉矩阵，可得 $\mathbf{x}=V^H\mathbf{y}$。代入原式：

$$\max_{\|\mathbf{y}\|_2=3}\|AV^H\mathbf{y}\|_2=3\cdot\max_{\|\frac{\mathbf{y}}{3}\|_2=1}\|AV^H\frac{\mathbf{y}}{3}\|_2=3\cdot\|AV^H\|_2=3\cdot\|A\|_2=6$$

---