

矩阵论基础知识

矩阵论是研究矩阵及其运算、性质与应用的数学分支，是线性代数的核心，在工程、物理、计算机科学、统计学等领域有广泛应用。

本文按 **由浅入深、循序渐进** 的方式组织：先建立矩阵的基本概念与运算，再引入行列式、秩、逆等方阵工具，然后讨论结构特殊的矩阵与复数情形，最后进入特征值、矩阵分解、范数等进阶主题，并以应用示例收尾。

导读：学习路径

阶段	章节	核心内容	目标
入门	§ 1–§ 2	定义、记号、特殊矩阵、基本运算	会读会写矩阵，掌握加减乘与转置
基础工具	§ 3–§ 5	行列式、秩、逆矩阵	判断可逆性，理解方程组解的结构
结构性质	§ 6–§ 7	对称/正交/正定矩阵；复数与共轭转置	识别常见矩阵类型，进入复数域
进阶理论	§ 8–§ 10	特征值、矩阵分解、范数与条件数	理解对角化、SVD 与数值稳定性
综合应用	§ 11	方程组、PCA、图论、机器学习等	将理论与实际问题对应

建议：每读完一节，先动手算一遍例题，再进入下一节。逆矩阵（§ 5）依赖行列式（§ 3）与秩（§ 4）；特征值（§ 8）依赖行列式；范数（§ 10）依赖 SVD 中的奇异值概念（§ 9）。

§ 1 矩阵的定义与基础记号

1.1 什么是矩阵

设 m, n 为正整数，由 $m \times n$ 个数排成 m 行 n 列的数表，称为 $m \times n$ **矩阵**：

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- a_{ij} 称为矩阵 A 的 **(i, j) 元**（第 i 行第 j 列的元素）
- $m \times n$ 称为矩阵的 **阶 或 维数**
- 行数与列数相等的矩阵称为 **方阵**（ n 阶方阵）

例：

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{是 } 2 \times 3 \text{ 矩阵, } a_{12} = 2, a_{21} = 4$$

$$B = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} \quad \text{是 } 2 \text{ 阶方阵}$$

1.2 基础记号

初学阶段常用记号如下；更进阶的符号在首次出现时会再说明，完整一览见 附录：记号一览。

记号	含义
a_{ij} 或 $(A)_{ij}$	第 i 行、第 j 列元素
$m \times n$	矩阵的阶（ m 行 n 列）
$\mathbb{R}^{m \times n}$ 、 $\mathbb{C}^{m \times n}$	所有 $m \times n$ 实（复）矩阵的集合
I_n 或 I	n 阶单位矩阵
O 或 $\mathbf{0}$	零矩阵
A^T	转置矩阵
$\mathbf{x}$ 、 $\mathbf{v}$	列向量（粗体小写字母）
A^k	A 的 k 次幂（ k 次自乘， $k \in \mathbb{Z}_{\geq 0}$ ）

§ 2 特殊矩阵与基本运算

先认识几种最简单的矩阵，再学习如何对它们做运算——这是后续一切讨论的基础。

2.1 常见特殊矩阵

名称	定义
零矩阵 O	所有元素均为 0
单位矩阵 I_n	主对角线为 1，其余为 0 的 n 阶方阵
对角矩阵	非主对角线元素全为 0，常记为 $\text{diag}(d_1, \dots, d_n)$

例：

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} \quad (\text{对角矩阵})$$

对任意同阶方阵 A ，有 $AI = IA = A$ ；对零矩阵，有 $A + O = A$ 。

2.2 加法与数乘

同阶矩阵可相加，对应元素相加；数 λ 与矩阵相乘，每个元素乘以 λ 。

例：

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}, \quad 3 \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ 0 & 6 \end{pmatrix}$$

2.3 矩阵乘法

设 A 为 $m \times n$ 矩阵， B 为 $n \times p$ 矩阵，则乘积 $C = AB$ 为 $m \times p$ 矩阵：

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

注意：矩阵乘法一般 **不满足交换律**，即 $AB \neq BA$ （即使两者都有定义）。

例：

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow AB = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \quad BA = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \Rightarrow AB \neq BA$$

行向量左乘列向量得到标量（内积）：

$$(1 \quad 2 \quad 3) \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$$

2.4 转置

$(A^T)_{ij} = a_{ji}$ ，即行列互换。重要性质： $(AB)^T = B^T A^T$ 。

例：

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

验证 $(AB)^T = B^T A^T$ ：用上式 A, B ， $AB = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ ，故 $(AB)^T = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$ ；而

$$B^T A^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}。$$

小结 (§2)：已掌握矩阵的"语法"——如何构造、相加、相乘、转置。下一节引入方阵的标量不变量 **行列式**。

§3 行列式

行列式是 **方阵** 对应的一个标量，记为 $\det(A)$ 、 $|A|$ 或（部分文献中） $\|A\|$ （勿与后文的矩阵范数混淆）。

3.1 基本性质

1. 交换两行，行列式变号
2. 某行乘以 k ，行列式乘以 k
3. 某行加上另一行的倍数，行列式不变
4. $\det(AB) = \det(A) \det(B)$
5. $\det(A^T) = \det(A)$

3.2 低阶计算公式

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

3 阶行列式可用 **Sarrus 法则** 或 **按行（列）展开** 计算。记

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Sarrus 法则（仅适用于 3 阶）：将矩阵的前两列再抄写到右侧，形成"扩展"的 3×5 数表。沿三条主对角线（左上 $\rightarrow$ 右下）各取三数相乘后相加，再减去沿三条副对角线（右上 $\rightarrow$ 左下）各取三数相乘之和：

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32})$$

此法便于手算记忆，但不推广到 4 阶及以上。

按行（列）展开（Laplace 展开）：删去第 i 行第 j 列后得到的 2 阶子矩阵的行列式，称为 (i, j) **余子式** M_{ij} ；**代数余子式** 定义为 $A_{ij} = (-1)^{i+j} M_{ij}$ 。沿任意一行或一列展开：

$$\det(A) = \sum_{j=1}^3 a_{ij} A_{ij} \quad (\text{按第 } i \text{ 行}), \quad \det(A) = \sum_{i=1}^3 a_{ij} A_{ij} \quad (\text{按第 } j \text{ 列})$$

展开时各项符号依 $(-1)^{i+j}$ 交替（棋盘格规律）。此法对任意阶方阵均有效，高阶行列式通常按此方法（或结合初等行变换）计算。

3.3 例

2 阶：

$$\det \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} = 2 \cdot 3 - 1 \cdot 5 = 1$$

3 阶（按第一行展开）：

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix} = 1 \cdot \det \begin{pmatrix} 1 & 4 \\ 6 & 0 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 0 & 4 \\ 5 & 0 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 0 & 1 \\ 5 & 6 \end{pmatrix} = -24 + 40 - 15 = 1$$

性质验证：交换前两行， $\det \begin{pmatrix} 5 & 3 \\ 2 & 1 \end{pmatrix} = 5 - 6 = -1$ ，符号相反。

小结（§3）：行列式 $\det(A) \neq 0$ 是方阵可逆的第一个判据。下一节从线性无关的角度引入秩，并与方程组解联系起来。

§4 秩与线性方程组

4.1 秩的定义

矩阵 A 的 **秩** $\text{rank}(A)$ 是其列向量组（或行向量组）的极大线性无关组所含向量的个数。

- $\text{rank}(A) \leq \min(m, n)$
- 可逆 n 阶方阵的秩等于 n

相关子空间记号：

记号	含义
$\text{ker}(A)$ 、 $\text{null}(A)$	核（零空间）：满足 $A\mathbf{x} = \mathbf{0}$ 的所有 $\mathbf{x}$
$\text{im}(A)$ 、 $\text{col}(A)$	像（列空间）
$\text{row}(A)$	行空间

秩的求法

方法	适用情形	要点
初等行变换	任意 $m \times n$ 矩阵（最常用）	化为行阶梯形，主元个数 = $\text{rank}(A)$
行列式	n 阶方阵	$\det(A) \neq 0 \Leftrightarrow \text{rank}(A) = n$
观察线性相关性	低阶、结构简单	若某行（列）为其余行（列）的线性组合，则秩小于阶数

方法一：初等行变换（推荐）

对 A 施以初等行变换（交换两行、某行乘以非零常数、某行加上另一行的倍数），化为 **行阶梯形** 或 **行最简形**（RREF）。**非零行的个数**即为秩；等价地，**主元列**的个数也为秩。

例（与 § 4.3 相同矩阵）：

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix}$$

第 1 步： $R_2 \leftarrow R_2 - 2R_1$, $R_3 \leftarrow R_3 - R_1$ ：

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & -1 & -2 \end{pmatrix}$$

第 2 步： $R_2 \leftrightarrow R_3$, $R_2 \leftarrow -R_2$ ：

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

两行非零，主元在第 1、2 列，故 $\text{rank}(A) = 2$ 。继续化为行最简形： $R_1 \leftarrow R_1 - 2R_2$ ，得

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \text{ 可读出 } c_3 = -c_1 + 2c_2.$$

方法二：行列式（仅方阵）

n 阶方阵 A 满秩 $\Leftrightarrow \det(A) \neq 0$ 。若 $\det(A) = 0$ ，则 $\text{rank}(A) < n$ ，需进一步化简求精确秩。

方法三：观察线性相关性

若可直接看出某行（列）是其余行（列）的倍数或线性组合，可快速判断秩。如 § 4.3 中第 2 行 = $2 \times$ 第 1 行，立知 $\text{rank}(A) \leq 2$ ；再结合第 3 行与第 1 行不成比例，得 $\text{rank}(A) = 2$ 。

注： $\text{rank}(A) = \text{rank}(A^T)$ （行秩 = 列秩）；初等 **列** 变换不改变秩。

4.2 秩与方程组

对 $m \times n$ 矩阵 A ，线性方程组 $A\mathbf{x} = \mathbf{b}$ 有解，当且仅当

$$\text{rank}(A) = \text{rank}([A \mid \mathbf{b}])$$

其中 $[A \mid \mathbf{b}]$ 为 **增广矩阵**（将列向量 $\mathbf{b}$ 拼在 A 右侧）。

4.3 例

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix}$$

第 2 行 = $2 \times$ 第 1 行，故列向量中线性无关的最多 2 个， $\text{rank}(A) = 2$ 。

解方程组 $A\mathbf{x} = \mathbf{b}$ ， $\mathbf{b} = (6, 12, 3)^T$ ：

$$[A \mid \mathbf{b}] = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & 4 & 6 & 12 \\ 1 & 1 & 1 & 3 \end{array} \right) \Rightarrow \text{rank}(A) = \text{rank}([A \mid \mathbf{b}]) = 2$$

方程有解（无穷多解）。若 $\mathbf{b} = (6, 12, 4)^T$ ，则 $\text{rank}([A \mid \mathbf{b}]) = 3 > 2$ ，无解。

小结（§ 4）：秩刻画了矩阵"有效维度"。结合行列式，可以完整讨论方阵的 **逆**。

§ 5 逆矩阵

若 n 阶方阵 A 满足 $AA^{-1} = A^{-1}A = I$, 则 A 可逆, A^{-1} 称为 A 的逆矩阵。

5.1 可逆的充要条件

以下等价 (A 为 n 阶方阵):

- $\det(A) \neq 0$
- $\text{rank}(A) = n$
- A 的列 (行) 向量线性无关
- 对任意 $\mathbf{b}$, $A\mathbf{x} = \mathbf{b}$ 有唯一解

5.2 2 阶求逆公式

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det(A) = ad - bc \neq 0 \Rightarrow A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

5.3 例

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}, \quad \det(A) = 2 \cdot 3 - 1 \cdot 5 = 1 \neq 0$$

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$$

验证: $AA^{-1} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = I_2$ 。

小结 (§ 5): 逆矩阵是"乘法意义下的倒数"。下面讨论具有特定 **结构** 的矩阵——它们在线性变换与二次型中反复出现。

§ 6 结构特殊的矩阵 (实数情形)

在掌握基本运算与逆之后, 可以识别并运用几类重要的 **实矩阵**。

名称	定义	几何 / 代数意义
对称矩阵	$A^T = A$	二次型 $\mathbf{x}^T A \mathbf{x}$ 的系数矩阵

名称	定义	几何 / 代数意义
反对称矩阵	$A^T = -A$	对角元必为 0
正交矩阵	$A^T A = I$	列向量两两正交且为单位长；保长度、保角度
正定矩阵	对称，且 $\mathbf{x}^T A \mathbf{x} > 0$ ($\mathbf{x} \neq \mathbf{0}$)	二次型恒正；Cholesky 分解的前提

迹 $\text{tr}(A)$ 定义为对角元之和；对称矩阵的迹等于其特征值之和（见 § 8）。

6.1 例

对称矩阵： $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ ，因为 $A^T = A$ ， $\text{tr}(A) = 1 + 3 = 4$ 。

反对称矩阵： $B = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$ ，因为 $B^T = -B$ 。

正交矩阵（旋转矩阵）：

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \theta = 90^\circ \Rightarrow R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad R^T R = I$$

正定矩阵： $P = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ，对 $\mathbf{x} = (1, -1)^T$ 有 $\mathbf{x}^T P \mathbf{x} = 2 > 0$ 。

小结（§ 6）：实数域上，转置 A^T 足以描述对称与正交。当元素为 **复数** 时，需要推广到共轭转置——见下一节。

§ 7 复数矩阵

在实数范围内，矩阵元素都是实数，转置 A^T 就足够了。当元素为 **复数** 时，需要引入 **共轭** 的概念。

7.1 复数的共轭

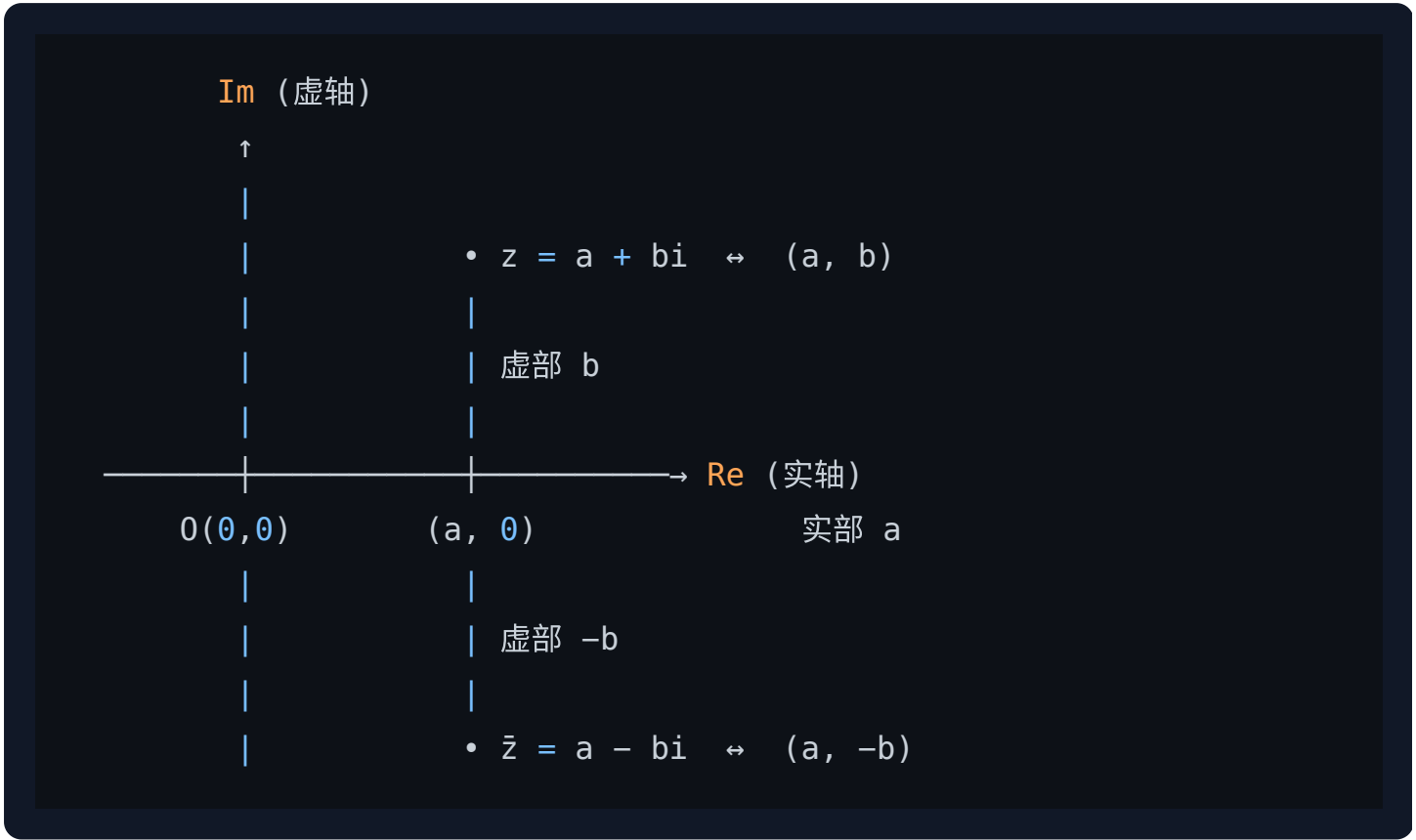
设 $z = a + bi$ ($a, b \in \mathbb{R}$, $i^2 = -1$)，其 **共轭复数**（简称 **共轭**）为：

$$\bar{z} = z^* = a - bi$$

几何上，**复平面** 用坐标 (a, b) 表示所有复数 $z = a + bi$ (a 为实部， b 为虚部)。横轴为 **实轴** (Re)，竖轴为 **虚轴** (Im)。共轭运算将 z 映为 $\bar{z}$ ，在坐标上即

$$z = a + bi \longleftrightarrow (a, b), \quad \bar{z} = a - bi \longleftrightarrow (a, -b)$$

实部 a 不变，虚部 b 变号——相当于在复平面上关于 **实轴**（横轴）做镜像对称。取原点 $O = (0, 0)$ 为实轴与虚轴的交点，则 z 、垂足 $(a, 0)$ 、 $\bar{z}$ 落在同一条竖直线 $x = a$ 上：



图中 $\bullet$ 表示复平面上的点； O 为坐标原点， $(a, 0)$ 是 z 在实轴上的垂足。 z 与 $\bar{z}$ 关于实轴对称，虚部互为相反数。

性质	公式
共轭的和	$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
共轭的积	$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$
模长平方	$z \bar{z} = \ z\ ^2 = a^2 + b^2$
实数不变	$z \in \mathbb{R} \Leftrightarrow \bar{z} = z$
纯虚数	$z = bi \Leftrightarrow \bar{z} = -z$

例：

$$z = 3 + 4i \Rightarrow \bar{z} = 3 - 4i, \quad z \bar{z} = 9 + 16 = 25 = \|z\|^2$$

$$\overline{(1+i)(2-i)} = (1-i)(2+i) = 3+i$$

复数的运算

设 $z_1 = a_1 + b_1i$, $z_2 = a_2 + b_2i$ ($a_k, b_k \in \mathbb{R}$)，常用 **实部** $\text{Re}(z) = a$ 、**虚部** $\text{Im}(z) = b$ （注意 $\text{Im}(z)$ 取实数值 b ，而非 bi ）。

运算	公式
加法	$z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i$
减法	$z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)i$
乘法	$z_1 z_2 = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i$
除法	$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{z_2 \bar{z}_2} = \frac{z_1 \bar{z}_2}{\ z_2\ ^2} \quad (z_2 \neq 0)$
模（长度）	$\ z\ = \sqrt{a^2 + b^2} = \sqrt{z \bar{z}}$
倒数	$\frac{1}{z} = \frac{\bar{z}}{\ z\ ^2} \quad (z \neq 0)$

加减法按实部、虚部分别相加减；乘法展开时利用 $i^2 = -1$ ；除法将分子分母同乘分母的共轭 $\bar{z}_2$ ，使分母变为实数 $\|z_2\|^2$ （**有理化**）。

例：

$$(1 + 2i) + (3 - i) = 4 + i$$

$$(1 + i)(2 - i) = 2 - i + 2i - i^2 = 3 + i$$

$$\frac{1 + i}{1 - i} = \frac{(1 + i)(1 + i)}{(1 - i)(1 + i)} = \frac{2i}{2} = i$$

$$\|3 + 4i\| = \sqrt{9 + 16} = 5$$

7.2 矩阵共轭与共轭转置

- **矩阵共轭** $\bar{A}$ ：对 A 的 **每个元素** 分别取共轭，即 $(\bar{A})_{ij} = \overline{a_{ij}}$ 。
- **共轭转置** $A^* = A^H = (\bar{A})^T$ ：先转置，再对每个元素取共轭（等价于先取共轭再转置）。

例：

$$A = \begin{pmatrix} 1+i & 2 \\ 0 & 3-i \end{pmatrix} \Rightarrow \bar{A} = \begin{pmatrix} 1-i & 2 \\ 0 & 3+i \end{pmatrix}, \quad A^* = \begin{pmatrix} 1-i & 0 \\ 2 & 3+i \end{pmatrix}$$

7.3 复数域上的特殊矩阵

实数情形	复数情形
对称矩阵: $A^T = A$	Hermite 矩阵: $A^* = A$
反对称矩阵: $A^T = -A$	反 Hermite 矩阵: $A^* = -A$
正交矩阵: $A^T A = I$	酉矩阵 (Unitary): $A^* A = I$

Hermite 矩阵的对角元必为 **实数** (因为 $a_{ii} = \overline{a_{ii}}$), 非对角元关于主对角线 **共轭对称**: $a_{ij} = \overline{a_{ji}}$ 。

例: $H = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$

- 对角元 2, 3 为实数 ✓
- $H_{12} = 1+i, H_{21} = 1-i = \overline{1+i}$ ✓
- 故 $H^* = H$, H 是 Hermite 矩阵

酉矩阵: $U = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, U^* = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}$, 故 $U^* U = I$ 。

小结 (§ 7): 共轭转置 A^* 是转置 A^T 在复数域上的自然推广; Hermite 矩阵、酉矩阵、量子力学中的可观测量都建立在其上。

§ 8 特征值与特征向量

对 n 阶方阵 A , 若存在非零向量 $\mathbf{v}$ 和标量 λ , 使得

$$A\mathbf{v} = \lambda\mathbf{v}$$

则 λ 称为 A 的 **特征值**, $\mathbf{v}$ 称为对应于 λ 的 **特征向量**。

直观理解: 特征向量在 A 作用下 **只被拉伸 (或压缩、反向)**, 方向不变。

8.1 特征方程

特征值是 **特征方程** 的根:

$$\det(A - \lambda I) = 0$$

该多项式称为 A 的 **特征多项式**, 次数为 n 。

8.2 重要结论

- $\text{tr}(A) = \sum_{i=1}^n \lambda_i$ (特征值之和等于迹)
- $\det(A) = \prod_{i=1}^n \lambda_i$ (特征值之积等于行列式)
- 相似矩阵** $A \sim B$ (即 $B = P^{-1}AP$ 对某可逆 P) 有相同的特征值

8.3 例

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

特征方程：

$$\det(A - \lambda I) = (4 - \lambda)(3 - \lambda) - 2 = \lambda^2 - 7\lambda + 10 = 0 \Rightarrow \lambda_1 = 5, \lambda_2 = 2$$

- $\text{tr}(A) = 7 = 5 + 2 \checkmark$
- $\det(A) = 12 - 2 = 10 = 5 \times 2 \checkmark$

对 $\lambda_1 = 5$ ： $(A - 5I)\mathbf{v} = \mathbf{0}$ 得 $\mathbf{v}_1 = (1, 1)^T$ 。

对 $\lambda_2 = 2$ ： $(A - 2I)\mathbf{v} = \mathbf{0}$ 得 $\mathbf{v}_2 = (1, -2)^T$ 。

验证： $A\mathbf{v}_1 = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5\mathbf{v}_1$ 。

小结 (§ 8)： 特征值把矩阵作用"分解"为沿各特征方向的缩放。当矩阵可对角化时，可进一步做 **特征值分解** (见 § 9)。

§ 9 矩阵分解

矩阵分解将复杂矩阵拆成结构简单的因子之积，便于理论分析与数值计算。由易到难：

分解	形式	适用条件 / 主要用途
LU 分解	$A = LU$ (L 下三角, U 上三角)	方阵；解线性方程组
Cholesky 分解	$A = LL^T$	对称正定矩阵
QR 分解	$A = QR$ (Q 正交, R 上三角)	任意 $m \times n$ 矩阵；最小二乘
满秩分解	$A = FG$ (F 列满秩, G 行满秩)	任意 $m \times n$ 矩阵；秩、列/行空间
特征值分解	$A = PDP^{-1}$ (D 对角)	A 可对角化时
奇异值分解 (SVD)	$A = U\Sigma V^T$	任意矩阵；降维、图像压缩

SVD 中 Σ 对角元 $\sigma_i \geq 0$ 称为 **奇异值**。

9.1 分步算例

以下用具体矩阵，按步骤手算各分解。符号与上表一致。

LU 分解

例： $A = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ ，解 $A\mathbf{x} = \mathbf{b}$ ， $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 。

第 1 步 ($k = 1$)：消去第 2 行第 1 列元素。乘子 $l_{21} = a_{21}/a_{11} = 4/2 = 2$ ；第 2 行减去 2 倍第 1 行：

$$\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \xrightarrow{\text{第 2 行} - 2 \times \text{第 1 行}} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

第 2 步 (读出 L 、 U)：

$$L = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

验证： $LU = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} = A$ 。

第 3 步 (解方程组) ——前代 $L\mathbf{y} = \mathbf{b}$ ：

$$\begin{cases} y_1 = 1 \\ 2y_1 + y_2 = 1 \end{cases} \Rightarrow y_1 = 1, y_2 = -1$$

回代 $U\mathbf{x} = \mathbf{y}$ ：

$$\begin{cases} 2x_1 + x_2 = 1 \\ x_2 = -1 \end{cases} \Rightarrow x_2 = -1, x_1 = 1$$

解为 $\mathbf{x} = (1, -1)^T$ 。

Cholesky 分解

例： $A = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$ (对称正定)。

设 $L = \begin{pmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{pmatrix}$ ，由 $A = LL^T$ 逐元素递推：

第 1 列：

$$l_{11} = \sqrt{a_{11}} = \sqrt{4} = 2, \quad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 1$$

第 2 列:

$$l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{2 - 1} = 1$$

故

$$L = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}, \quad LL^T = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} = A$$

QR 分解 (Gram-Schmidt)

例: $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, 列向量 $\mathbf{a}_1 = (1, 1, 0)^T$, $\mathbf{a}_2 = (1, 0, 1)^T$ 。

第 1 列:

$$r_{11} = \|\mathbf{a}_1\| = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad \mathbf{q}_1 = \frac{\mathbf{a}_1}{r_{11}} = \frac{1}{\sqrt{2}}(1, 1, 0)^T$$

第 2 列——先减去在 $\mathbf{q}_1$ 上的投影:

$$r_{12} = \mathbf{q}_1^T \mathbf{a}_2 = \frac{1}{\sqrt{2}}(1, 1, 0) \cdot (1, 0, 1) = \frac{1}{\sqrt{2}}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - r_{12}\mathbf{q}_1 = (1, 0, 1)^T - \frac{1}{2}(1, 1, 0)^T = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)^T$$

归一化:

$$r_{22} = \|\mathbf{u}_2\| = \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \sqrt{\frac{3}{2}}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{6}}(1, -1, 2)^T$$

结果:

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{pmatrix}, \quad R = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{pmatrix}$$

验证: $QR = A$ ($Q^T Q = I$, R 上三角)。

满秩分解

设 A 为 $m \times n$ 矩阵, $\text{rank}(A) = r$ 。满秩分解将 A 写成

$$A = FG$$

其中 F 为 $m \times r$ **列满秩** 矩阵 ($\text{rank}(F) = r$)， G 为 $r \times n$ **行满秩** 矩阵 ($\text{rank}(G) = r$)。 F 的列张成 **列空间** $\text{col}(A)$ ， G 的行张成 **行空间** $\text{row}(A)$ 。

例： $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix}$ (与 § 4 相同, 3×3 , $\text{rank}(A) = 2$)。

第 1 步 (行最简形) —— 对 A 做初等行变换：

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{R_2 - 2R_1, R_3 - R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & -1 & -2 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3, R_2 \leftarrow -R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 - 2R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = R$$

主元在第 1、2 列，故 $r = 2$ 。

第 2 步 (取 F) —— 在 **原矩阵** A 中取主元列 (第 1、2 列)：

$$F = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 1 \end{pmatrix}$$

第 3 步 (取 G) —— 在行最简形 R 中取前 r 个非零行 (主元行)：

$$G = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

第 4 步 (验证)：

$$FG = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix} = A$$

读出行最简形的列关系 (c_i 为 A 的第 i 列)： $c_3 = -c_1 + 2c_2$ ，即第 3 列由主元列线性表出，与 G 的第三列 $(-1, 2)^T$ 的系数一致。

注： 满秩分解不唯一 (主元列选取、行变换路径不同均可给出不同的 F, G)。对 $m \times n$ 矩阵， F 的阶为 $m \times r$ ， G 的阶为 $r \times n$ 。

特征值分解

例： $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ (与 § 8 相同)。

第 1 步 (特征值)：

$$\det(A - \lambda I) = (4 - \lambda)(3 - \lambda) - 2 = \lambda^2 - 7\lambda + 10 = 0 \Rightarrow \lambda_1 = 5, \lambda_2 = 2$$

$$D = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}$$

第 2 步 ($\lambda_1 = 5$ 的特征向量) —— 解 $(A - 5I)\mathbf{v} = \mathbf{0}$:

$$\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v}_1 = (1, 1)^T$$

第 3 步 ($\lambda_2 = 2$ 的特征向量) —— 解 $(A - 2I)\mathbf{v} = \mathbf{0}$:

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v}_2 = (1, -2)^T$$

第 4 步 (组装 P):

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \quad A = PDP^{-1}$$

验证: $AP = \begin{pmatrix} 5 & 2 \\ 5 & -4 \end{pmatrix} = PD$ 。

奇异值分解 (SVD)

例: $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ (非对角, 便于展示完整流程)。

第 1 步 (计算 $A^T A$):

$$A^T A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$$

第 2 步 ($A^T A$ 的特征值与 V):

$$\det(A^T A - \lambda I) = (5 - \lambda)^2 - 16 = \lambda^2 - 10\lambda + 9 = 0 \Rightarrow \lambda_1 = 9, \lambda_2 = 1$$

奇异值 $\sigma_1 = \sqrt{9} = 3$, $\sigma_2 = \sqrt{1} = 1$ 。

- $\lambda_1 = 9$: $(A^T A - 9I)\mathbf{v} = \mathbf{0}$ 得 $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1, 1)^T$
- $\lambda_2 = 1$: $(A^T A - I)\mathbf{v} = \mathbf{0}$ 得 $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1, -1)^T$

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

第 3 步 (求 U) —— 对 $\sigma_i > 0$, $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$:

$$\mathbf{u}_1 = \frac{1}{3} A\mathbf{v}_1 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(1, 1)^T$$

$$\mathbf{u}_2 = A\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(-1, 1)^T$$

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

第4步 (验证): $U\Sigma V^T = A$, 且 $U^T U = V^T V = I$ 。

注: 若 A 本身已是对角阵 (如 $\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$), 可直接取 $\sigma_i = |a_{ii}|$, $U = V = I$ (负对角元对应 V 中符号调整), 无需完整计算 $A^T A$ 。

小结 (§9): 分解是连接理论与算法的桥梁。SVD 的奇异值也用于定义矩阵 **范数** 与 **条件数**。

§ 10 范数与条件数

10.1 矩阵范数

Frobenius 范数 (最直观, 对应元素平方和):

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

谱范数 (2-范数): $\|A\|_2 = \sigma_{\max}(A)$, 即最大奇异值。

10.2 条件数

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|$$

条件数衡量 $A\mathbf{x} = \mathbf{b}$ 对数据扰动的敏感程度: κ 越大, 数值求解越 **病态** (不稳定)。

10.3 例

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Frobenius 范数:

$$\|A\|_F = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30} \approx 5.48$$

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}, \quad \|A^{-1}\|_F = \sqrt{4 + 1 + 2.25 + 0.25} = \sqrt{7.5} \approx 2.74。$$

条件数 (Frobenius 范数): $\kappa_F(A) \approx 5.48 \times 2.74 \approx 15$, 属中等敏感。

对比 Hilbert 矩阵 $H_2 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.333 \end{pmatrix}$, $\kappa(H_2) \approx 27$, 病态程度更高。

§ 11 典型应用

将前述工具与实际问题对应：

领域	用到的核心概念	示例
线性方程组	逆矩阵、LU 分解、秩	$A\mathbf{x} = \mathbf{b}$ 的求解与解的结构
线性变换	矩阵乘法、正交矩阵	旋转、投影等几何变换
主成分分析 (PCA)	对称矩阵、特征值分解	协方差矩阵的特征向量即主成分方向
图论	邻接矩阵、对角矩阵	拉普拉斯矩阵 $L = D - A$
量子力学	Hermite 矩阵、特征值	可观测量（厄米特算符）的本征值
机器学习	伪逆、矩阵乘法	线性回归 $\beta = (X^T X)^{-1} X^T \mathbf{y}$

11.1 例

线性方程组： $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, 解为 $\mathbf{x} = (1, -1)^T$ 。

线性变换：旋转矩阵 $R_{90^\circ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ 将 $(1, 0)^T$ 映为 $(0, 1)^T$ 。

PCA：数据协方差矩阵 $\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 的特征值为 3, 1, 最大特征值对应方向 $(1, 1)^T$ 为第一主成分。

图论：三角形图的邻接矩阵 $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, 度矩阵 $D = \text{diag}(2, 2, 2)$, 拉普拉斯矩阵 $L = D - A$ 。

机器学习：线性回归 $\hat{\mathbf{y}} = X\beta$, 正规方程 $\beta = (X^T X)^{-1} X^T \mathbf{y}$ 即矩阵求逆的应用。

附录：记号一览

本文及矩阵论中常用的记号汇总如下，按主题分类，便于查阅。

A. 矩阵、元素与维数

记号	含义
$A, B, C, \dots$	矩阵（大写字母）
a_{ij} 或 $(A)_{ij}$	矩阵 A 第 i 行、第 j 列的元素
c_{ij}	乘积矩阵 $C = AB$ 的第 (i, j) 元
$m \times n$	矩阵的阶（ m 行 n 列）
n 阶方阵	行数与列数均为 n 的方阵
$\mathbb{R}^{m \times n}$	所有 $m \times n$ 实矩阵的集合
$\mathbb{C}^{m \times n}$	所有 $m \times n$ 复矩阵的集合

B. 向量

记号	含义
$\mathbf{x}, \mathbf{v}, \mathbf{b}, \mathbf{y}$	列向量（粗体小写字母）
$\mathbf{x}^T$	行向量（列向量的转置）
$\mathbf{0}$	零向量
$(x_1, x_2, \dots, x_n)^T$	n 维列向量的分量表示

C. 特殊矩阵

记号	含义
I_n 或 I	n 阶单位矩阵（主对角线为 1，其余为 0）
O 或 $\mathbf{0}$	零矩阵（所有元素为 0）
$\text{diag}(d_1, \dots, d_n)$	以 $d_1, \dots, d_n$ 为主对角元的对角矩阵
Λ	对角矩阵的常用记法之一

D. 基本运算

记号	含义
$A + B$	矩阵加法（同阶矩阵，对应元素相加）

记号	含义
λA 或 kA	数乘（每个元素乘以标量 λ 或 k ）
AB	矩阵乘法（ A 为 $m \times n$ ， B 为 $n \times p$ ）
A^T	转置（行列互换）
A^k	A 的 k 次幂（ k 次自乘， $k \in \mathbb{Z}_{\geq 0}$ ）
A^{-1}	逆矩阵（若存在；满足 $AA^{-1} = A^{-1}A = I$ ）
$A^\dagger$ 或 A^+	Moore–Penrose 伪逆（广义逆）

E. 行列式、秩与迹

记号	含义
$\det(A)$ 、 $\ A\ $ 或 $\$$	A
$\text{rank}(A)$	秩（列向量组或行向量组的极大线性无关组大小）
$\text{tr}(A)$	迹（主对角元之和）

F. 子空间与线性方程组

记号	含义
$\ker(A)$ 、 $\text{null}(A)$	核（零空间）：满足 $A\mathbf{x} = \mathbf{0}$ 的所有 $\mathbf{x}$
$\text{im}(A)$ 、 $\text{col}(A)$	像（列空间）
$\text{row}(A)$	行空间
$A\mathbf{x} = \mathbf{b}$	线性方程组
$[A \mid \mathbf{b}]$	增广矩阵（将列向量 $\mathbf{b}$ 拼在 A 右侧）

G. 结构特殊的矩阵

记号	含义
$A^T = A$	对称矩阵
$A^T = -A$	反对称矩阵

记号	含义
$A^T A = I$	正交矩阵
$A^* = A$	Hermite 矩阵（厄米特矩阵）
$A^* = -A$	反 Hermite 矩阵
$A^* A = I$	酉矩阵（Unitary）
$\mathbf{x}^T A \mathbf{x} > 0 \ (\mathbf{x} \neq \mathbf{0})$	正定矩阵（ A 对称时）

H. 复数与共轭

记号	含义
i	虚数单位, $i^2 = -1$
$z = a + bi$	复数 (a 为实部, b 为虚部)
$\bar{z}$ 或 z^*	复数 z 的共轭
$\ z\ $	复数 z 的模 (长度)
$\bar{A}$	矩阵共轭 (逐元素取共轭)
A^* 或 A^H	共轭转置 ($A^* = (\bar{A})^T$)

I. 特征值与相似

记号	含义
λ, λ_i	特征值
$\mathbf{v}, \mathbf{v}_i$	特征向量
$A\mathbf{v} = \lambda\mathbf{v}$	特征值-特征向量关系
$\det(A - \lambda I) = 0$	特征方程
$A \sim B$	A 与 B 相似 (存在可逆 P 使 $B = P^{-1}AP$)
$A = PDP^{-1}$	特征值分解 (D 对角, P 可逆)

J. 矩阵分解

记号	含义
$A = LU$	LU 分解 (L 下三角, U 上三角)
$A = LL^T$	Cholesky 分解 (L 下三角; A 对称正定)
$A = QR$	QR 分解 (Q 正交, R 上三角)
$A = FG$	满秩分解 (F 为 $m \times r$ 列满秩, G 为 $r \times n$ 行满秩, $r = \text{rank}(A)$)
$A = U\Sigma V^T$	奇异值分解 (SVD)
σ_i 、 $\sigma_{\max}$	奇异值 (SVD 中对角元, $\sigma_i \geq 0$)
$L, U, Q, R, F, G, P, D, \Sigma$	各分解中的因子矩阵 (见上)

K. 范数与条件数

记号	含义
$\ A\ _F$	Frobenius 范数: $\sqrt{\sum_{i,j} a_{ij}^2}$
$\ A\ _2$	谱范数 (2-范数), 等于最大奇异值 $\sigma_{\max}(A)$
$\kappa(A)$	条件数: $\ A\ \cdot \ A^{-1}\ $
$\kappa_F(A)$	以 Frobenius 范数定义的条件数

L. 应用与其他常用符号

记号	含义
R 、 R_θ	旋转矩阵 (正交矩阵之一)
θ	旋转角
Σ	协方差矩阵 (如 PCA 中)
D	度矩阵 (图论中, 对角元为顶点度数)
$L = D - A$	拉普拉斯矩阵 (A 为邻接矩阵)
X	设计矩阵 / 数据矩阵 (统计、机器学习中)
β	回归系数向量

记号	含义
$\hat{\mathbf{y}}$	预测值向量
$\sum_{k=1}^n$	求和（如矩阵乘法中的指标求和）
$\Leftrightarrow, \Rightarrow$	等价、推出（逻辑关系）
$\min(m, n)$	m 与 n 中的较小者
$\mathbb{R}, \mathbb{C}$	实数集、复数集
$\mathbb{Z}_{\geq 0}$	非负整数集

参考资料

- 张贤达, 《矩阵分析与应用》
- Roger A. Horn & Charles R. Johnson, *Matrix Analysis*
- Gilbert Strang, *Introduction to Linear Algebra*
- 丘维声, 《高等代数》