# 第 1 章　矩阵理论基础

本章主要介绍矩阵理论的基本概念、基本运算，为理解这一数学工具打下坚实的基础. 通过本章的学习，读者将掌握矩阵的基本构造，了解它们如何通过简单的规则进行变换，以及这些变换如何影响我们对世界的理解和解释.

## 1.1　向量与矩阵

### 1.1.1　基本概念

**1. 引入**

在研究科学和工程问题时，其中的很多问题都可以通过数学建模转化成一个线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1-1}$$

在线性方程组 (1-1) 中，使用 $m$ 个方程描述 $n$ 个未知量之间的线性关系. 这种表示方式在研究问题时不够简练，为了简化问题，可以采用向量和矩阵. 下面介绍向量与矩阵的概念.

**2. 向量与矩阵的概念**

线性方程组(1-1)中所有变量的系数可按照相对位置排成 $m$ 行 $n$ 列的数集，按照如下格式写成：

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

称 $\boldsymbol{A}$ 为 $m \times n$ 矩阵，简记为 $\boldsymbol{A} = (a_{ij})_{m\times n}$，其中 $a_{ij}$ 表示矩阵 $\boldsymbol{A}$ 的第 $i$ 行、第 $j$ 列元素，简称第 $(i, j)$ 个元素. 当 $a_{ij}$ 取实数时，称 $\boldsymbol{A}$ 为**实矩阵**；当 $a_{ij}$ 取复数时，称 $\boldsymbol{A}$ 为**复矩阵**.

当 $m=n$ 时，矩阵 $\boldsymbol{A}$ 称为 **$\boldsymbol{n}$ 阶正方矩阵**，简称 **$\boldsymbol{n}$ 阶方阵**；若 $m<n$，则称矩阵 $\boldsymbol{A}$ 为**宽矩阵**；而当 $m>n$ 时，便称矩阵 $\boldsymbol{A}$ 为**高矩阵**.

特别地，当 $n=1$ 时，$\boldsymbol{A}$ 退化为如下形式：

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}. \tag{1-2}$$

称式(1-2) 的形式为 **$\boldsymbol{m}$ 维列向量**，简称 **$\boldsymbol{m}$ 维向量**，一般用小写字母表示，如 $\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{x}$，$\boldsymbol{y}$，$\cdots$. 若其元素 $a_i$ 取实数，即 $a_i \in \mathbf{R}$，则称其为 **$\boldsymbol{m}$ 维实（数）向量**，并记作 $\boldsymbol{a} \in \mathbf{R}^{m\times 1}$，或者简记为 $\boldsymbol{a} \in \mathbf{R}^m$. 类似地，若 $a_i \in \mathbf{C}$，则称其为 **$\boldsymbol{m}$ 维复向量**，并记作 $\boldsymbol{a} \in \mathbf{C}^{m\times 1}$.

类似地，称 $[a_1, a_2, \cdots, a_n]$ 为一个 **$\boldsymbol{n}$ 维行向量**，一般记作 $\boldsymbol{a}^{\mathrm{T}} \in \mathbf{R}^{1\times n}$ 或 $\boldsymbol{a}^{\mathrm{T}} \in \mathbf{C}^{1\times n}$.

有了以上概念，方程组(1-1)可以方便地记为矩阵形式：

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

一个 $n$ 阶方阵 $\boldsymbol{A}$ 的**主对角线**是指从左上角到右下角沿 $i=j$，$j=1,2,\cdots,n$ 连接的线段. 位于主对角线上的元素称为 $\boldsymbol{A}$ 的对角元素，它们是 $a_{11}, a_{22}, \cdots, a_{nn}$.

主对角线以外元素全部为零的 $n$ 阶方阵称为**对角矩阵**，记作

$$\boldsymbol{D} = \mathrm{diag}(d_{11}, d_{22}, \cdots, d_{nn}).$$

若 $n$ 阶对角矩阵主对角线元素全部等于 1，则称为**单位矩阵**，用符号 $\boldsymbol{I}_n$ 表示. 所有元素为零的 $m\times n$ 矩阵称为**零矩阵**，记为 $\boldsymbol{O}_{m\times n}$. 一个全部元素为零的向量称为**零向量**，记为 $\mathbf{0}$.

为了书写简洁，单位矩阵、零矩阵分别记为 $\boldsymbol{I}, \boldsymbol{O}$.

### 1.1.2 矩阵的基本运算

矩阵的基本运算包括矩阵的转置、共轭、共轭转置和求逆等.

#### 1. 矩阵的转置

若 $\boldsymbol{A} = (a_{ij})_{m\times n}$，称矩阵

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

为矩阵 $\boldsymbol{A}$ 的**转置矩阵**, 记为 $\boldsymbol{A}^{\mathrm{T}}$. 若 $\boldsymbol{A}$ 是一个复矩阵, 对其每个元素 $a_{ij}(i=1,2,\cdots,m;$ $j=1,2,\cdots,n)$ 取共轭后得到的矩阵称为 $\boldsymbol{A}$ 的**共轭矩阵**，记为 $\boldsymbol{A}^*$. 易见，$\boldsymbol{A}^*$ 仍然是一个 $m\times n$ 矩阵，即

$$\boldsymbol{A}^*=\begin{bmatrix} a_{11}^* & a_{12}^* & \cdots & a_{1n}^* \\ a_{21}^* & a_{22}^* & \cdots & a_{2n}^* \\ \vdots & \vdots & & \vdots \\ a_{m1}^* & a_{m2}^* & \cdots & a_{mn}^* \end{bmatrix}.$$

矩阵 $\boldsymbol{A}$ 求转置矩阵后再求共轭矩阵所得到的矩阵称为 $\boldsymbol{A}$ 的**共轭转置矩阵**，记为 $\boldsymbol{A}^{\mathrm{H}}$，易见，$\boldsymbol{A}^{\mathrm{H}}$ 是一个 $n\times m$ 矩阵，且

$$\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix} a_{11}^* & a_{21}^* & \cdots & a_{m1}^* \\ a_{12}^* & a_{22}^* & \cdots & a_{m2}^* \\ \vdots & \vdots & & \vdots \\ a_{1n}^* & a_{2n}^* & \cdots & a_{mn}^* \end{bmatrix}.$$

若 $\boldsymbol{A}$ 为实方阵，且满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为**对称矩阵**；若 $\boldsymbol{A}$ 为复方阵，且满足 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为 **Hermite 矩阵**（**或复共轭对称矩阵**）. 下面是矩阵的共轭、转置、共轭转置的性质，感兴趣的读者可以自己证明.

（1）矩阵的共轭、转置和共轭转置满足分配律：

$$(\boldsymbol{A}+\boldsymbol{B})^*=\boldsymbol{A}^*+\boldsymbol{B}^*,\quad (\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}},\quad (\boldsymbol{A}+\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}+\boldsymbol{B}^{\mathrm{H}}.$$

（2）矩阵乘积的转置、共轭转置满足关系式

$$(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}},\quad (\boldsymbol{A}\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}.$$

**2. 矩阵的求逆**

在很多工程问题中，需要将一个向量 $\boldsymbol{x}$ 进行变换，如用一个矩阵 $\boldsymbol{A}$ 乘以向量 $\boldsymbol{x}$，记

$$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}. \tag{1-3}$$

该类问题往往涉及从向量 $\boldsymbol{y}$ 到向量 $\boldsymbol{x}$ 的逆变换，这会用到下面的计算.

设 $\boldsymbol{A}$ 为方阵，若存在方阵 $\boldsymbol{B}$，使得

$$\boldsymbol{B}\boldsymbol{A}=\boldsymbol{I},$$

此时，式(1-3) 等号两边同时左乘 $\boldsymbol{B}$，得到

$$\boldsymbol{B}\boldsymbol{y}=\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{I}\boldsymbol{x}=\boldsymbol{x}.$$

一般地，若 $\boldsymbol{A}$ 为 $n$ 阶方阵，存在矩阵 $\boldsymbol{B}$，使得

$$\boldsymbol{AB}=\boldsymbol{I}\quad(\text{或}\boldsymbol{BA}=\boldsymbol{I})$$

成立，则称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的**逆矩阵**（或 $\boldsymbol{A}$ 也是 $\boldsymbol{B}$ 的逆矩阵），记为

$$\boldsymbol{B}=\boldsymbol{A}^{-1}\quad(\text{或}\boldsymbol{A}=\boldsymbol{B}^{-1}).$$

若 $\boldsymbol{A}$ 存在逆矩阵，则 $\boldsymbol{A}$ 为可逆矩阵（或非奇异矩阵）.

逆矩阵在数学和工程问题中有广泛的应用，对于理解和解决许多实际问题至关重要. 本章最后一节将会给出利用 Python 求解逆矩阵的方法.

**3. 正交矩阵与酉矩阵**

正交矩阵在数学、工程、计算机图形学和物理学等领域有着广泛应用，其独特的性质使其在简化运算、保持几何特性、实现变换等方面发挥重要作用，是众多领域不可或缺的数学工具. 下面给出正交矩阵的定义.

**定义 1.1.1** 若 $\boldsymbol{A}$ 为 $n$ 阶实方阵，且满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$（或 $\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{I}$），则称 $\boldsymbol{A}$ 为**正交矩阵**.

由正交矩阵的定义不难看出，若 $\boldsymbol{A}$ 为 $n$ 阶实方阵，以下五个命题等价：

（1）$\boldsymbol{A}$ 为正交矩阵；

（2）$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$；

（3）$\boldsymbol{A}^{\mathrm{T}}$ 为正交矩阵；

（4）$\boldsymbol{A}^{-1}$ 为正交矩阵；

（5）$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$.

将正交矩阵的定义推广到复数域，可以得到如下酉矩阵的定义.

**定义 1.1.2** 若 $\boldsymbol{A}$ 为 $n$ 阶复方阵，且满足 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{I}$(或 $\boldsymbol{AA}^{\mathrm{H}}=\boldsymbol{I}$)，则称复方阵 $\boldsymbol{A}$ 为**酉矩阵**.

类似地，若 $\boldsymbol{A}$ 为 $n$ 阶复方阵，则以下五个命题等价：

（1）$\boldsymbol{A}$ 为酉矩阵；

（2）$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}^{-1}$；

（3）$\boldsymbol{A}^{\mathrm{H}}$ 为酉矩阵；

（4）$\boldsymbol{A}^{-1}$ 为酉矩阵；

（5）$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{I}$.

正交矩阵和酉矩阵都是具有特殊性质的方阵，它们在数学和应用领域中都扮演着重要角色. 正交矩阵的元素是实数，具有保持向量的长度和夹角不变等性质，在几何变换、信号处理和控制系统等领域应用广泛. 酉矩阵则是正交矩阵在复数域的推广，同样可以保持向量的长度和夹角，在量子力学、复信号处理和复系统分析中具有重要应用.

## 1.2 矩阵的初等变换与初等矩阵

### 1.2.1 矩阵的初等变换

矩阵的初等变换是线性代数中的基本概念，它们是对矩阵行或列操作的集合，用于简化矩阵形式，从而降低所解决问题的复杂度. 这些变换包括行（列）交换、数乘行（列）以及行（列）加法等.

下面给出矩阵的初等行变换的几种具体形式.

（1）行交换

行交换是将矩阵的两行位置相互交换. 矩阵的第 $i$ 行 $r_i$ 和第 $j$ 行 $r_j$ 交换可以表示为

$$r_i \leftrightarrow r_j.$$

（2）数乘行

数乘行是将矩阵的某一行乘以一个非零常数. 矩阵的第 $k$ 行 $r_k$ 乘以常数 $\alpha$ 可以表示为

$$\alpha r_k.$$

（3）行加法

行加法是将矩阵的某一行乘以一个常数后加到另一行上. 矩阵的第 $j$ 行乘以常数 $\alpha$ 然后加到第 $i$ 行上可以表示为

$$r_i + \alpha r_j.$$

以上三种对矩阵行的变换称为矩阵的**初等行变换**，是研究线性方程组理论及求解优化问题的重要工具. 对矩阵的列进行类似于上面的三种操作，称为矩阵的初等列变换. 常用 $c_i \leftrightarrow c_j$ 表示矩阵的第 $i$ 列和第 $j$ 列互换，用 $\alpha c_k$ 表示矩阵的第 $k$ 列乘以常数 $\alpha$，用 $c_i + \alpha c_j$ 表示将矩阵的第 $j$ 列乘以常数 $\alpha$ 后加到第 $i$ 列上.

若矩阵 $\boldsymbol{A}$ 经过有限次初等变换变成 $\boldsymbol{B}$，则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价，记为 $\boldsymbol{A} \sim \boldsymbol{B}$.

### 1.2.2 初等矩阵

初等变换虽然很直观，但是由于形式的限制不便于深入研究，因此有必要将变换过程描述为更数学化的形式.

对单位矩阵进行一次初等行变换所得到的矩阵称为**初等矩阵**，三种初等行变换对应于三种初等矩阵.

（1）$r_i \leftrightarrow r_j$ 对应的初等矩阵记为 $\boldsymbol{E}(i,j)$，其中

$$
\boldsymbol{E}(i,j)=\begin{array}{c@{}l}
\left[\begin{array}{ccccccccccc}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 &&&& 1 &&& \\
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& 1 &&&& 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{array}\right]
&
\begin{array}{l}
\\ \\ \\
\longleftarrow\quad 第i行 \\
\\ \\ \\
\longleftarrow\quad 第j行 \\
\\ \\ \\
\end{array}
\end{array}.
$$

$\boldsymbol{E}(i,j)$ 是单位矩阵 $\boldsymbol{I}$ 第 $i$ 行与第 $j$ 行互换得到的. 容易验证，矩阵 $\boldsymbol{A}$ 的第 $i$ 行与第 $j$ 行互换的结果等于 $\boldsymbol{E}(i,j)$ 左乘 $\boldsymbol{A}$，矩阵 $\boldsymbol{A}$ 的第 $i$ 列与第 $j$ 列互换的结果等于 $\boldsymbol{E}(i,j)$ 右乘 $\boldsymbol{A}$.

（2）$\alpha r_k$ 对应的初等矩阵记为 $\boldsymbol{E}(k,\alpha)$，其中

$$
\boldsymbol{E}(k,\alpha)=\begin{array}{c@{}l}
\left[\begin{array}{ccccc}
1 &&&& \\
& \ddots &&& \\
&& \alpha && \\
&&& \ddots & \\
&&&& 1
\end{array}\right]
&
\begin{array}{l}
\\ \\
\longleftarrow\quad 第k行. \\
\\ \\
\end{array}
\end{array}
$$

$\boldsymbol{E}(k,\alpha)$ 是单位矩阵 $\boldsymbol{I}$ 第 $k$ 行乘以常数 $\alpha$ 得到的. 容易验证，矩阵 $\boldsymbol{A}$ 第 $k$ 行乘以常数 $\alpha$ 的结果等于 $\boldsymbol{E}(k,\alpha)$ 左乘 $\boldsymbol{A}$，$\boldsymbol{A}$ 第 $k$ 列乘以常数 $\alpha$ 的结果等于 $\boldsymbol{E}(k,\alpha)$ 右乘 $\boldsymbol{A}$.

（3）$r_i+\alpha r_j$ 对应的初等矩阵记为 $\boldsymbol{E}(i,j,\alpha)$，其中

$$
\boldsymbol{E}(i,j,\alpha)=\begin{array}{c@{}l}
\left[\begin{array}{ccccccc}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 && \alpha && \\
&&& \ddots &&& \\
&&&& 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{array}\right]
&
\begin{array}{l}
\\ \\
\longleftarrow\quad 第i行 \\
\\
\longleftarrow\quad 第j行 \\
\\ \\
\end{array}
\end{array}.
$$

$\boldsymbol{E}(i,j,\alpha)$ 是单位矩阵 $\boldsymbol{I}$ 第 $j$ 行乘以常数 $\alpha$ 后加到第 $i$ 行或者第 $i$ 列乘以常数 $\alpha$ 后加到第 $j$ 列得到的. 容易验证，矩阵 $\boldsymbol{A}$ 第 $j$ 行乘以常数 $\alpha$ 后加到第 $i$ 行的结果等

于 $\boldsymbol{E}(i,j,\alpha)$ 左乘 $\boldsymbol{A}$，$\boldsymbol{A}$ 第 $i$ 列乘以常数 $\alpha$ 后加到第 $j$ 列的结果等于 $\boldsymbol{E}(i,j,\alpha)$ 右乘 $\boldsymbol{A}$.

因此，初等矩阵是执行矩阵初等变换的矩阵. 通过这种形式上的变化，使得研究矩阵理论问题变得更加方便.

初等变换和初等矩阵是矩阵理论中的基础工具，它们在求解线性方程组、矩阵分解和矩阵求逆等操作中扮演着关键角色. 通过这些变换，可以在将矩阵保持某些重要特征的情况下，简化所研究问题的求解过程.

## 1.3 行阶梯形矩阵、行最简形矩阵

### 1.3.1 行阶梯形矩阵

设方程组

$$\begin{cases} x_1 - 2x_2 + x_3 = -2 \\ x_1 + x_2 - x_3 = 1 \\ 3x_1 - 5x_2 + x_3 = -13 \end{cases},$$

则该方程组的求解过程可以用其增广矩阵 $\boldsymbol{B}$ 的初等行变换来描述：

$$\boldsymbol{B} = \begin{bmatrix} 1 & -2 & 1 & -2 \\ 1 & 1 & -1 & 1 \\ 3 & -5 & 1 & -13 \end{bmatrix} \xrightarrow[r_3-3r_1]{r_2-r_1} \begin{bmatrix} 1 & -2 & 1 & -2 \\ 0 & 3 & -2 & 3 \\ 0 & 1 & -2 & -7 \end{bmatrix} \xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & -2 & 1 & -2 \\ 0 & 1 & -2 & -7 \\ 0 & 3 & -2 & 3 \end{bmatrix}$$

$$\xrightarrow{r_3-3\times r_2} \begin{bmatrix} \underline{1} & -2 & 1 & -2 \\ 0 & \underline{1} & -2 & -7 \\ 0 & 0 & \underline{4} & \underline{24} \end{bmatrix} \xrightarrow{r_3\times\frac{1}{4}} \begin{bmatrix} 1 & -2 & 1 & -2 \\ 0 & 1 & -2 & -7 \\ 0 & 0 & 1 & 6 \end{bmatrix} \tag{1-4}$$

$$\xrightarrow{r_2+2\times r_3} \begin{bmatrix} 1 & -2 & 1 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 6 \end{bmatrix} \xrightarrow[r_1-r_3]{r_1+2\times r_2} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 6 \end{bmatrix}. \tag{1-5}$$

式(1-4)及式 (1-5)中的矩阵具有如下特点：可画出一条阶梯线，线的下方元素全是 0；每个台阶只有一行，台阶数即是非零行的行数，阶梯线的竖线（每段竖线的长度为一行）后面的第一个元素为非零元，也就是非零行的第一个非零元素，这种矩阵称为**行阶梯形矩阵**.

### 1.3.2 行最简形矩阵

式(1-5)右侧的行阶梯形矩阵还具有如下特点：非零行的第一个非零元为 1，且其所在列的其余元素都是 0，称为**行最简形矩阵**.

将线性方程组的增广矩阵化为行最简形矩阵后，直接就能看出该方程组的解. 在解决一些优化问题时，也经常将矩阵变换为行最简形矩阵来简化计算.

可以证明：任何矩阵 $\boldsymbol{A}_{m\times n}$ 总可经有限次初等行变换化为行阶梯形矩阵和行最简形矩阵. 在通过将矩阵化成行最简形解线性方程组的讨论中，我们发现，方程组的增广矩阵经过初等行变换后，尽管形式上发生了很大的改变，但是却保留了关于方程组的解的一些本质的特性.

对行最简形矩阵再进行初等列变换，可变成形式更简单的矩阵，例如

$$\begin{bmatrix}1 & 0 & 0 & 2\\ 0 & 1 & 0 & 5\\ 0 & 0 & 1 & 6\end{bmatrix}\xrightarrow{c_4-2\times c_1-5\times c_2-6\times c_3}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}. \tag{1-6}$$

称式(1-6)中右侧的矩阵为**矩阵 $\boldsymbol{B}$ 的标准形**. 矩阵的标准形具有如下特点：左上角是一个单位矩阵，其余元素全为零.

任意矩阵 $\boldsymbol{A}$ 总可经有限次初等行变换和初等列变换化为标准形，标准形常简单地记为如下形式：

$$\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}_{m\times n}.$$

标准形由 $m$，$n$，$r$ 三个数完全确定，$r$ 是行阶梯形矩阵中非零行的行数. 显然，在一个矩阵的所有等价矩阵中，标准形是形式最简单的.

## 1.4 矩阵的行列式、特征值、迹和秩

在矩阵的工程应用中，经常希望能够使用一个数或者一个标量来反映一个矩阵的性能. 下面介绍评价矩阵性质的几个重要标量指标：矩阵的行列式、特征值、迹和秩.

### 1.4.1 矩阵的行列式

一个 $n$ 阶方阵 $\boldsymbol{A}$ 的全部元素 (各元素位置保持不变) 所对应的行列式称为矩阵 $\boldsymbol{A}$ 的**行列式**，记作 $\det(\boldsymbol{A})$ 或 $|\boldsymbol{A}|$.

**定义 1.4.1** 行列式不等于零的矩阵称为**非奇异矩阵**.

下面不加证明地给出矩阵行列式的一些性质：

（1）单位矩阵的行列式等于 1，即 $\det(\boldsymbol{I})=1$.

（2）任何一个方阵 $\boldsymbol{A}$ 和它的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 具有相同的行列式，即 $\det(\boldsymbol{A})=\det(\boldsymbol{A}^{\mathrm{T}})$，但 $\det(\boldsymbol{A}^{\mathrm{H}})=[\det(\boldsymbol{A}^{\mathrm{T}})]^*$.

（3）两个方阵乘积的行列式等于它们的行列式的乘积，即 $\det(\boldsymbol{AB})=\det(\boldsymbol{A})\det(\boldsymbol{B})$，$\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$.

（4）给定一个任意的常数 $\alpha$，则 $\det(\alpha\boldsymbol{A})=\alpha^n\det(\boldsymbol{A})$，其中 $\boldsymbol{A}$ 为 $n$ 阶方阵.

（5）若 $\boldsymbol{A}$ 是非奇异矩阵，则 $\det(\boldsymbol{A}^{-1})=\dfrac{1}{\det(\boldsymbol{A})}$.

方阵的行列式是方阵的一个基础指标，可以用于判断矩阵是否非奇异，主要刻画矩阵的奇异性.

### 1.4.2 矩阵的特征值与特征向量

在研究物理学、控制论及解析几何等很多问题时，经常会遇到这样一个问题：对于方阵 $\boldsymbol{A}$，能否找到数 $\lambda$ 和向量 $\boldsymbol{x}$，使得 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，即向量 $\boldsymbol{Ax}$ 与 $\boldsymbol{x}$ “平行”? 其对应的数学问题即为以下求方阵的特征值与特征向量的问题.

**定义 1.4.2** 设 $\boldsymbol{A}$ 为 $n$ 阶方阵，$\boldsymbol{x}$ 为非零列向量，若存在数 $\lambda$，使

$$\boldsymbol{Ax}=\lambda\boldsymbol{x}, \tag{1-7}$$

则称 $\lambda$ 为矩阵 $\boldsymbol{A}$ 的**特征值**，称 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 $\lambda$ 的**特征向量** (也称 $\boldsymbol{x}$ 为对应于特征值 $\lambda$ 的特征向量).

矩阵 $\boldsymbol{A}$ 的特征值常用符号 $\mathrm{eig}(\boldsymbol{A})$ 表示. 下面给出特征值的一些基本性质：

（1）矩阵乘积的特征值：$\mathrm{eig}(\boldsymbol{AB})=\mathrm{eig}(\boldsymbol{BA})$.

（2）逆矩阵的特征值：$\mathrm{eig}(\boldsymbol{A}^{-1})=\dfrac{1}{\mathrm{eig}(\boldsymbol{A})}$.

（3）令 $\boldsymbol{I}$ 为单位矩阵，$\alpha$ 为标量，则

$$\mathrm{eig}(\boldsymbol{I}+\alpha\boldsymbol{A})=1+\alpha\ \mathrm{eig}(\boldsymbol{A}),$$

$$\mathrm{eig}(\boldsymbol{A}-\alpha\boldsymbol{I})=\mathrm{eig}(\boldsymbol{A})-\alpha.$$

除了以上基本性质，特征值与特征向量还具有如下性质：

（4）若 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $n$ 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的 $n$ 个特征值，则 $\det(\boldsymbol{A})=\lambda_1\lambda_2\cdots\lambda_n$.

（5）设 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 $\lambda$ 的特征向量，由于 $\boldsymbol{A}(k\boldsymbol{x})=k(\boldsymbol{Ax})=k(\lambda\boldsymbol{x})=\lambda(k\boldsymbol{x})$，故 $k\boldsymbol{x}$ $(k\neq 0)$ 也为 $\boldsymbol{A}$ 的属于特征值 $\lambda$ 的特征向量.

（6）设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是方阵 $\boldsymbol{A}$ 的 $m$ 个特征值，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m$ 依次是与之对应的特征向量. 若 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 互不相等，则 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m$ 线性无关 (定义参考文献 [3]).

（7）设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是 $n$ 阶方阵 $\boldsymbol{A}$ 的互不相同的特征值，且 $\boldsymbol{p}_{i1},\boldsymbol{p}_{i2},\cdots,\boldsymbol{p}_{ir_i}$ 是 $\boldsymbol{A}$ 的属于特征值 $\lambda_i(i=1,2,\cdots,m)$ 的线性无关的特征向量，则向量组

$$\boldsymbol{p}_{11},\boldsymbol{p}_{12},\cdots,\boldsymbol{p}_{1r_1},\boldsymbol{p}_{21},\boldsymbol{p}_{22},\cdots,\boldsymbol{p}_{2r_2},\cdots,\boldsymbol{p}_{m1},\boldsymbol{p}_{m2},\cdots,\boldsymbol{p}_{mr_m}$$

也线性无关.

根据性质（4）容易得出结论：若矩阵 $\boldsymbol{A}$ 是非奇异矩阵，则其所有特征值都不等于 0.

特征值刻画了矩阵在特定方向上的“拉伸”或“压缩”程度，在用主成分分析进行数据降维中扮演着关键的角色.

### 1.4.3 矩阵的迹

**定义 1.4.3** $n$ 阶方阵 $\boldsymbol{A}$ 的对角元素之和称为 $\boldsymbol{A}$ 的**迹** (trace)，记作 $\mathrm{tr}(\boldsymbol{A})$，即

$$\mathrm{tr}(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}. \tag{1-8}$$

在通信工程中，一个具有 $n$ 个信号分量的信号向量 $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_n(t)]^{\mathrm{T}}$ 的自相关矩阵 $\boldsymbol{R} = E[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)]$ 的迹 $\mathrm{tr}(\boldsymbol{R}) = \mathrm{tr}(E[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)]) = E[|s_1(t)|^2] + E[|s_2(t)|^2] + \cdots + E[|s_n(t)|^2]$，表示 $n$ 个信号分量的能量之和，这里的 $E$ 指的是数学期望.

下面介绍矩阵的迹具有的一些基本性质[3].

（1）线性性质：$\mathrm{tr}(\boldsymbol{A} \pm \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}) \pm \mathrm{tr}(\boldsymbol{B})$，$\mathrm{tr}(\alpha\boldsymbol{A}) = \alpha\mathrm{tr}(\boldsymbol{A})$，$\alpha$ 为任意常数.

（2）矩阵 $\boldsymbol{A}$ 的转置、复数共轭和复共轭转置的迹分别为 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{tr}(\boldsymbol{A})$，$\mathrm{tr}(\boldsymbol{A}^*) = [\mathrm{tr}(\boldsymbol{A})]^*$ 和 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}) = [\mathrm{tr}(\boldsymbol{A})]^*$.

（3）若 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，$\boldsymbol{B} \in \mathbf{C}^{n\times m}$，则 $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$.

（4）若 $\boldsymbol{A}$ 是一个 $m \times n$ 矩阵，则 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = 0 \Longleftrightarrow \boldsymbol{A} = \boldsymbol{O}_{m\times n}$.

（5）矩阵 $\boldsymbol{A}$ 的迹等于该矩阵所有特征值之和，即 $\mathrm{tr}(\boldsymbol{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

（6）分块矩阵（定义参考文献 [3]）的迹满足

$$\mathrm{tr}\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \mathrm{tr}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{D}).$$

式中，$\boldsymbol{A} \in \mathbf{C}^{m\times m}$，$\boldsymbol{B} \in \mathbf{C}^{m\times n}$，$\boldsymbol{C} \in \mathbf{C}^{n\times m}$，$\boldsymbol{D} \in \mathbf{C}^{n\times n}$.

由性质（3）可以得出，矩阵 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})$，且容易计算得

$$\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2. \tag{1-9}$$

对于同阶方阵 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$，由性质（3）及矩阵乘法的结合律容易得出：

$$\mathrm{tr}(\boldsymbol{ABC}) = \mathrm{tr}(\boldsymbol{BCA}) = \mathrm{tr}(\boldsymbol{CAB}). \tag{1-10}$$

根据式(1-10)还易知，若矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为同阶方阵，且 $\boldsymbol{B}$ 非奇异，则

$$\mathrm{tr}(\boldsymbol{BAB}^{-1}) = \mathrm{tr}(\boldsymbol{B}^{-1}\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{ABB}^{-1}) = \mathrm{tr}(\boldsymbol{A}).$$

通过性质（5）可知，矩阵的迹所反映的矩阵的性能指标是所有特征值之和.

### 1.4.4 矩阵的秩

一组 $m$ 维向量 $\boldsymbol{x}_i \in \mathbf{C}^m (i = 1, 2, \cdots, n)$ 称为**线性无关**，若关于 $k_1, k_2, \cdots, k_n$ 的方程 $k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_n\boldsymbol{x}_n = \mathbf{0}$ 只有零解，即只有当 $k_1 = k_2 = \cdots = k_n = 0$ 时方程成立. 若存在一组不全部为零的系数 $k_1, k_2, \cdots, k_n$ 满足上述方程，则称向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ **线性相关**.