第一章 矩阵代数

- §1.1 定义
- §1.2 矩阵的运算
- §1.3 行列式
- §1.4 矩阵的逆
- §1.5 矩阵的秩
- §1.6 特征值、特征向量和矩阵的迹
- §1.7 正定矩阵和非负定矩阵
- §1.8 特征值的极值问题

§1.1 定义

$p\times q$矩阵：
$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{pmatrix}$$

p维列向量：
$$\boldsymbol{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$

q维行向量：$\boldsymbol{b}'=(b_1,b_2,\cdots,b_q)$

向量$\boldsymbol{a}$的长度：$\|\boldsymbol{a}\|=\sqrt{\boldsymbol{a}'\boldsymbol{a}}=\sqrt{a_1^2+a_2^2+\cdots+a_p^2}$

单位向量：$\|\boldsymbol{a}\|=1$

- 若$\boldsymbol{A}$的所有元素全为零，则称$\boldsymbol{A}$为零矩阵，记作$\boldsymbol{A}=\boldsymbol{0}_{pq}$或$\boldsymbol{A}=\boldsymbol{0}$。
- 若$p=q$，则称$\boldsymbol{A}$为p阶方阵，$a_{11},a_{22},\cdots,a_{pp}$称为它的对角线元素，其他元素$a_{ij}(i\neq j)$称为非对角线元素。
- 若方阵$\boldsymbol{A}$的对角线下方的元素全为零，则称$\boldsymbol{A}$为上三角矩阵。显然，$a_{ij}=0,i>j$。
- 若方阵$\boldsymbol{A}$的对角线上方的元素全为零，则称$\boldsymbol{A}$为下三角矩阵。显然，$a_{ij}=0,i<j$。
- 若方阵$\boldsymbol{A}$的所有非对角线元素均为零，则称$\boldsymbol{A}$为对角矩阵，简记为$\boldsymbol{A}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{pp})$。
- 若p阶对角矩阵$\boldsymbol{A}$的所有p个对角线元素均为1，则称$\boldsymbol{A}$为p阶单位矩阵，记作$\boldsymbol{A}=\boldsymbol{I}_p$或$\boldsymbol{A}=\boldsymbol{I}$。

- 若将矩阵$\boldsymbol{A}$的行与列互换，则得到的矩阵称为$\boldsymbol{A}$的转置，记作$\boldsymbol{A}'$，即

$$\boldsymbol{A}' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{p1} \\ a_{12} & a_{22} & \cdots & a_{p2} \\ \vdots & \vdots & & \vdots \\ a_{1q} & a_{2q} & \cdots & a_{pq} \end{pmatrix}$$

- 若方阵$\boldsymbol{A}$满足$\boldsymbol{A}'=\boldsymbol{A}$，则称$\boldsymbol{A}$为对称矩阵。显然，$a_{ij}=a_{ji}$。

§1.2 矩阵的运算

- 若$\boldsymbol{A}=(a_{ij})$：$p\times q$，$\boldsymbol{B}=(b_{ij})$：$p\times q$,则$\boldsymbol{A}$与$\boldsymbol{B}$的和定义为

$$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij}):\ p\times q$$

- 若c为一常数，则它与$\boldsymbol{A}$的积定义为

$$c\boldsymbol{A}=(ca_{ij}):\ p\times q$$

- 若$\boldsymbol{A}=(a_{ij})$：$p\times q$，$\boldsymbol{B}=(b_{ij})$：$q\times r$,则$\boldsymbol{A}$与$\boldsymbol{B}$的积定义为

$$\boldsymbol{AB}=\left(\sum_{k=1}^{q}a_{ik}b_{kj}\right):\ p\times r$$

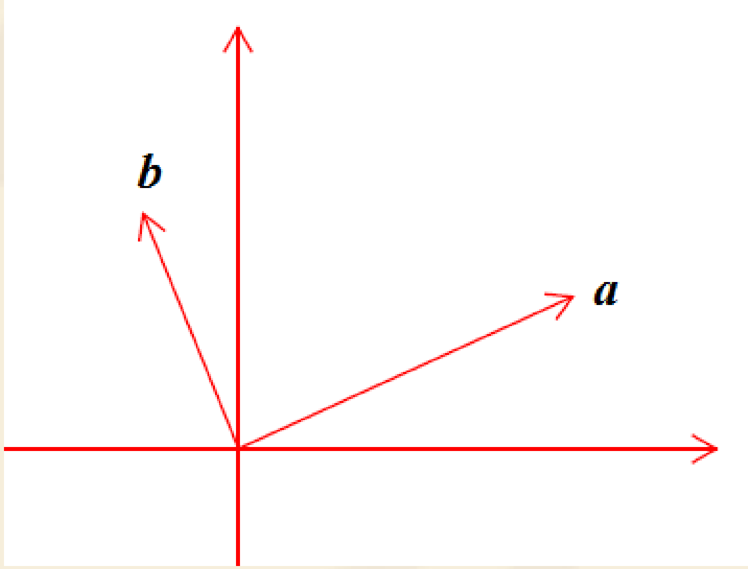
运算规律

- (1)$(\boldsymbol{A}+\boldsymbol{B})'=\boldsymbol{A}'+\boldsymbol{B}'$。
- (2)$(\boldsymbol{AB})'=\boldsymbol{B}'\boldsymbol{A}'$。
- (3)$\boldsymbol{A}(\boldsymbol{B}_1+\boldsymbol{B}_2)=\boldsymbol{AB}_1+\boldsymbol{AB}_2$。
- (4) $\boldsymbol{A}\left(\sum_{i=1}^{k}\boldsymbol{B}_i\right)=\sum_{i=1}^{k}\boldsymbol{AB}_i$ 。
- (5)$c(\boldsymbol{A}+\boldsymbol{B})=c\boldsymbol{A}+c\boldsymbol{B}$。

- 若两个p维向量$\boldsymbol{a}$和$\boldsymbol{b}$满足

$$\boldsymbol{a}'\boldsymbol{b}=a_1b_1+a_2b_2+\cdots+a_pb_p=0$$

则称$\boldsymbol{a}$和$\boldsymbol{b}$正交。



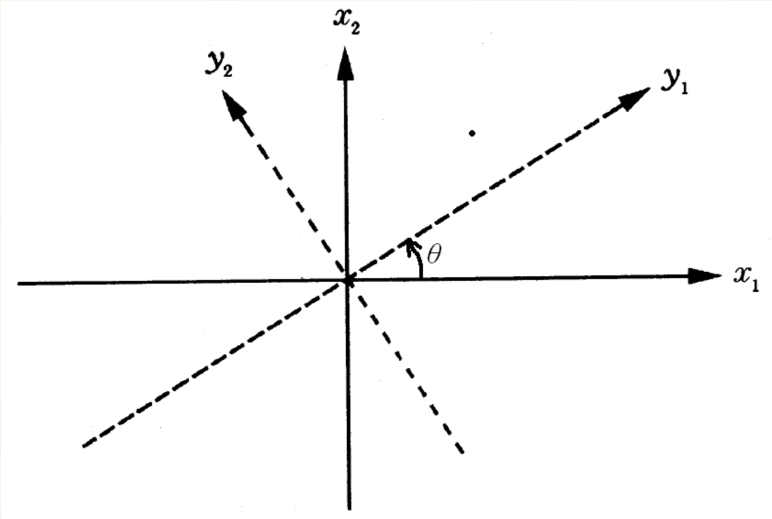
- 若方阵$\boldsymbol{A}$满足$\boldsymbol{AA}'=\boldsymbol{I}$，则称$\boldsymbol{A}$为正交矩阵。
- 正交矩阵的三个等价定义：

$$\boldsymbol{AA}'=\boldsymbol{I} \Leftrightarrow \boldsymbol{A}'=\boldsymbol{A}^{-1} \Leftrightarrow \boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$$

- 若方阵$\boldsymbol{A}$满足$\boldsymbol{A}^2=\boldsymbol{A}$，则称$\boldsymbol{A}$为幂等矩阵。
- 对称的幂等矩阵称为投影矩阵。

正交矩阵$\boldsymbol{A}$的几何意义

- 当p=2时，



$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\boldsymbol{A}\boldsymbol{x}$$

- 当p=3时，坐标系（刚性）旋转后新旧坐标的变换可表达为

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\boldsymbol{A}\boldsymbol{x}$$

其中的变换矩阵也一定为正交矩阵。

- 正交阵$\boldsymbol{A}$的行列式非1即−1。若$|\boldsymbol{A}|=1$，则正交变换$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$意味着对原$p$维坐标系作一刚性旋转（或称正交旋转），$\boldsymbol{y}$的各分量正是该点在新坐标系下的坐标；若$|\boldsymbol{A}|=-1$，则包含了一个镜面反射的坐标轴。
- 由于

$$\boldsymbol{y}'\boldsymbol{y}=(\boldsymbol{A}\boldsymbol{x})'(\boldsymbol{A}\boldsymbol{x})=\boldsymbol{x}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{x}$$

故在新、旧坐标系下，该点到原点的距离保持不变。

矩阵的分块

- 设$\boldsymbol{A}=(a_{ij})$：$p\times q$,将它分成四块，表示成

$$\boldsymbol{A}=\begin{array}{cc} \left(\begin{array}{cc} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{array}\right) & \begin{array}{l} k \\ p-k \end{array} \\ \begin{array}{cc} l & q-l \end{array} & \end{array}$$

- 若$\boldsymbol{A}$和$\boldsymbol{B}$有相同的分块，则

$$\boldsymbol{A}+\boldsymbol{B}=\left(\begin{array}{cc} \boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12} \\ \boldsymbol{A}_{21}+\boldsymbol{B}_{21} & \boldsymbol{A}_{22}+\boldsymbol{B}_{22} \end{array}\right)$$

❖ 若$\boldsymbol{C}$为$q \times r$矩阵，分成

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12}\\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22}\end{pmatrix}$$

则有

$$\boldsymbol{AC}=\begin{array}{c}\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}\\ \begin{matrix} l & \quad q-l\end{matrix}\end{array}\begin{array}{c}\begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12}\\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22}\end{pmatrix}\begin{matrix} l\\ q-l\end{matrix}\\ \end{array}$$

$$=\begin{pmatrix}\boldsymbol{A}_{11}\boldsymbol{C}_{11}+\boldsymbol{A}_{12}\boldsymbol{C}_{21} & \boldsymbol{A}_{11}\boldsymbol{C}_{12}+\boldsymbol{A}_{12}\boldsymbol{C}_{22}\\ \boldsymbol{A}_{21}\boldsymbol{C}_{11}+\boldsymbol{A}_{22}\boldsymbol{C}_{21} & \boldsymbol{A}_{21}\boldsymbol{C}_{12}+\boldsymbol{A}_{22}\boldsymbol{C}_{22}\end{pmatrix}$$

- 例1.2.2 证明正交矩阵**A**：$p\times p$的p个列向量和p个行向量都是一组正交单位向量。

- 证明 记

$$\boldsymbol{A}=\left(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p\right)=\begin{pmatrix}\boldsymbol{a}'_{(1)}\\ \boldsymbol{a}'_{(2)}\\ \vdots\\ \boldsymbol{a}'_{(p)}\end{pmatrix}$$

由$\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$，得

$$\begin{pmatrix}\boldsymbol{a}'_1\\ \boldsymbol{a}'_2\\ \vdots\\ \boldsymbol{a}'_p\end{pmatrix}\left(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p\right)=\boldsymbol{I}$$

于是

$$\begin{pmatrix} \boldsymbol{a}_1'\boldsymbol{a}_1 & \boldsymbol{a}_1'\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1'\boldsymbol{a}_p \\ \boldsymbol{a}_2'\boldsymbol{a}_1 & \boldsymbol{a}_2'\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_2'\boldsymbol{a}_p \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_p'\boldsymbol{a}_1 & \boldsymbol{a}_p'\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_p'\boldsymbol{a}_p \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

故有

$$\boldsymbol{a}_i'\boldsymbol{a}_j = \begin{cases} 1, & 若 i = j \\ 0, & 若 1 \le i \ne j \le p \end{cases}$$

即$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$为一组正交单位向量。同理，由$\boldsymbol{AA}'=\boldsymbol{I}$可证$\boldsymbol{a}_{(1)},\boldsymbol{a}_{(2)},\cdots,\boldsymbol{a}_{(p)}$也是一组正交单位向量。

§1.3 行列式

- p阶方阵$\boldsymbol{A}=(a_{ij})$的行列式定义为

$$\left|\boldsymbol{A}\right| = \sum_{j_1 j_2 \cdots j_p} (-1)^{\tau\left(j_1 j_2 \cdots j_p\right)} a_{1j_1} a_{2j_2} \cdots a_{pj_p}$$

这里 $\sum_{j_1 j_2 \cdots j_p}$ 表示对1,2,…,p的所有排列求和，$\tau(j_1 j_2 \cdots j_p)$是排列$j_1,j_2,\cdots,j_p$中逆序的总数，称它为这个排列的逆序数，一个逆序是指在一个排列中一对数的前后位置与大小顺序相反，即前面的数大于后面的数。例如，

$\tau(3142)=1+\tau(1342)=3+\tau(1234)=3$。

- 注：该定义只需了解即可，无需记住。应用中3阶及以上的行列式一般都用软件计算，除非特殊结构的。

行列式的一些基本性质

- (1)若$\boldsymbol{A}$的某行(或列)为零，则$|\boldsymbol{A}|=0$。
- (2)$|\boldsymbol{A}'|=|\boldsymbol{A}|$。
- (3)若将$\boldsymbol{A}$的某一行(或列)乘以常数c，则所得矩阵的行列式为$c|\boldsymbol{A}|$。
- (4)若$\boldsymbol{A}$是一个p阶方阵，c为一常数，则$|c\boldsymbol{A}|=c^p|\boldsymbol{A}|$。
- (5)若互换$\boldsymbol{A}$的任意两行(或列)，则行列式符号改变。
- (6)若$\boldsymbol{A}$的某两行(或列)相同，则行列式为零。
- (7)若将$\boldsymbol{A}$的某一行(或列)的倍数加到另一行(或列)，则所得行列式不变。
- (8)若$\boldsymbol{A}$的某一行(或列)是其他一些行(或列)的线性组合，则行列式为零。

- (9)若$\boldsymbol{A}$为上三角矩阵或下三角矩阵或对角矩阵，则$|\boldsymbol{A}| = \prod_{i=1}^{p} a_{ii}$
- (10)若$\boldsymbol{A}$和$\boldsymbol{B}$均为p阶方阵，则$|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$。
- (11)$|\boldsymbol{AA}'|\geq 0$。
- (12)若$\boldsymbol{A}$与$\boldsymbol{B}$都是方阵，则

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{vmatrix} = \begin{vmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{C} & \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{B}|$$

- (13)若$\boldsymbol{A}$：$p\times q$，$\boldsymbol{B}$：$q\times p$，则

$$|\boldsymbol{I}_p+\boldsymbol{AB}|=|\boldsymbol{I}_q+\boldsymbol{BA}|$$

- 证明　因为

$$\begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{A} \\ \boldsymbol{0} & \boldsymbol{I}_q \end{pmatrix}\begin{pmatrix} \boldsymbol{I}_p & -\boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}=\begin{pmatrix} \boldsymbol{I}_p+\boldsymbol{AB} & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{0} \\ -\boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}\begin{pmatrix} \boldsymbol{I}_p & -\boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}=\begin{pmatrix} \boldsymbol{I}_p & -\boldsymbol{A} \\ \boldsymbol{0} & \boldsymbol{I}_q+\boldsymbol{AB} \end{pmatrix}$$

上述两个等式两边各取行列式，故得

$$|\boldsymbol{I}_p+\boldsymbol{AB}|=|\boldsymbol{I}_q+\boldsymbol{BA}|$$

- 例1.3.3 设$\boldsymbol{x},\boldsymbol{y}$为两个$p$维向量，则

$$|\boldsymbol{I}_p+\boldsymbol{xy}'|=1+\boldsymbol{y}'\boldsymbol{x}$$

代数余子式

- 设$\boldsymbol{A}$为p阶方阵，将其元素a_{ij}所在的第i行与第j列划去之后所得$(p-1)$阶矩阵的行列式，称为元素a_{ij}的余子式，记为M_{ij}。$A_{ij}=(-1)^{i+j}M_{ij}$称为元素a_{ij}的代数余子式。有以下公式成立

$$|\boldsymbol{A}| = \sum_{j=1}^{p} a_{ij} A_{ij} = \sum_{i=1}^{p} a_{ij} A_{ij}$$

§1.4 矩阵的逆

- 若方阵$\boldsymbol{A}$满足$|\boldsymbol{A}|\neq 0$，则称$\boldsymbol{A}$为非退化（或非奇异）方阵；若$|\boldsymbol{A}|=0$，则称$\boldsymbol{A}$为退化（或奇异）方阵。
- 设$\boldsymbol{A}=(a_{ij})$是一非退化方阵，若方阵$\boldsymbol{C}$满足$\boldsymbol{AC}=\boldsymbol{I}$，则称$\boldsymbol{C}$为$\boldsymbol{A}$的逆矩阵，记为$\boldsymbol{C}=\boldsymbol{A}^{-1}$，$\boldsymbol{A}^{-1}$必是一个非退化矩阵。令

$$\boldsymbol{B}'=(A_{ij})/|\boldsymbol{A}|$$

其中A_{ij}是a_{ij}的代数余子式，则容易验证$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$。由于$\boldsymbol{C}=\boldsymbol{BAC}=\boldsymbol{B}$，因此$\boldsymbol{A}^{-1}$是唯一的，且$(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$。

- 注：应用中3阶及以上的逆矩阵一般都用软件计算，除非特殊结构的。

逆矩阵的基本性质

- (1)$\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$。
- (2)$(\boldsymbol{A}')^{-1}=(\boldsymbol{A}^{-1})'$。
- (3)若$\boldsymbol{A}$和$\boldsymbol{C}$均为p阶非退化方阵，则

$$(\boldsymbol{A}\boldsymbol{C})^{-1}=\boldsymbol{C}^{-1}\boldsymbol{A}^{-1}$$

- (4)$|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$。
- (5)若$\boldsymbol{A}$是正交矩阵，则$\boldsymbol{A}^{-1}=\boldsymbol{A}'$。
- (6)若$\boldsymbol{A}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{pp})$非退化(即$a_{ii}\neq 0$，$i=1,2,\cdots,p$)，则

$$\boldsymbol{A}^{-1}=\mathrm{diag}\left(a_{11}^{-1},a_{22}^{-1},\cdots,a_{pp}^{-1}\right)$$

- (7)若$\boldsymbol{A}$和$\boldsymbol{B}$为非退化方阵，则

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}^{-1}=\begin{pmatrix} \boldsymbol{A}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}^{-1} \end{pmatrix}$$

§1.5 矩阵的秩

- 一组同维向量$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$，若存在不全为零的常数$c_1,c_2,\cdots,c_n$，使得

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n=\boldsymbol{0}$$

  则称该组向量线性相关。若向量$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$不线性相关，就称为线性无关。
- 矩阵$\boldsymbol{A}$的线性无关行向量的最大数目称为行秩，其线性无关列向量的最大数目称为列秩。矩阵的行秩和列秩必相等，故统一将其称为$\boldsymbol{A}$的秩，记作rank($\boldsymbol{A}$)。

矩阵秩的基本性质

- (1)rank($\boldsymbol{A}$)=0，当且仅当$\boldsymbol{A}$=$\mathbf{0}$。
- (2)若$\boldsymbol{A}$为$p\times q$矩阵, 且$\boldsymbol{A}\neq\mathbf{0}$，则$1\le \text{rank}(\boldsymbol{A})\le\min\{p,q\}$[若rank($\boldsymbol{A}$)=$p$,则称$\boldsymbol{A}$为行满秩的；若rank($\boldsymbol{A}$)=$q$,则称$\boldsymbol{A}$为列满秩的]。
- (3)rank($\boldsymbol{A}$)=rank($\boldsymbol{A}'$)。
- (4) $\text{rank}\begin{pmatrix}\boldsymbol{A} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{B}\end{pmatrix}=\text{rank}\begin{pmatrix}\mathbf{0} & \boldsymbol{A}\\ \boldsymbol{B} & \mathbf{0}\end{pmatrix}=\text{rank}(\boldsymbol{A})+\text{rank}(\boldsymbol{B})$。
- (5)$\text{rank}(\boldsymbol{AB})\le\min\{\text{rank}(\boldsymbol{A}),\text{rank}(\boldsymbol{B})\}$。
- (6)若$\boldsymbol{A}$和$\boldsymbol{C}$为非退化方阵，则

$$\text{rank}(\boldsymbol{ABC})=\text{rank}(\boldsymbol{B})$$

- (7)p阶方阵$\boldsymbol{A}$是非退化的，当且仅当rank($\boldsymbol{A}$)=p(称作$\boldsymbol{A}$满秩)。
- (8)rank($\boldsymbol{AA}'$)=rank($\boldsymbol{A}'\boldsymbol{A}$)=rank($\boldsymbol{A}$)。

§1.6 特征值、特征向量和矩阵的迹

- 一、特征值和特征向量
- 二、矩阵的迹

一、特征值和特征向量

- 设$\boldsymbol{A}$是p阶方阵，若对于一个数λ，存在$\boldsymbol{x}\neq\boldsymbol{0}$，使得$\boldsymbol{Ax}=\lambda\boldsymbol{x}$，则称$\lambda$为$\boldsymbol{A}$的一个特征值或特征根，而称$\boldsymbol{x}$为$\boldsymbol{A}$的属于$\lambda$的一个特征向量。
- $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，$\boldsymbol{x}\neq\boldsymbol{0}$，故

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$$

$|\boldsymbol{A}-\lambda\boldsymbol{I}|$是$\lambda$的$p$次多项式，称为特征多项式。上式有$p$个根（可能有重根），记作$\lambda_1,\lambda_2,\cdots,\lambda_p$，可以为复数。反过来，若$\lambda_i$是上式的一个根，则存在$\boldsymbol{x}_i\neq\boldsymbol{0}$，使得

$$(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{x}_i=\boldsymbol{0}$$

- 今后，一般情况下取$\boldsymbol{x}_i$为单位向量，即满足$\|\boldsymbol{x}_i\|=1$。

特征值和特征向量的基本性质

- (1)$\boldsymbol{A}$和$\boldsymbol{A}'$有相同的特征值。
- (2)若$\boldsymbol{A}$和$\boldsymbol{B}$分别是$p\times q$和$q\times p$矩阵，则$\boldsymbol{AB}$和$\boldsymbol{BA}$有相同的非零特征值。
- 证明　因为

$$\begin{pmatrix} \boldsymbol{I}_p & -\boldsymbol{A} \\ \boldsymbol{0} & \lambda\boldsymbol{I}_q \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_p & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}=\begin{pmatrix} \lambda\boldsymbol{I}_p-\boldsymbol{AB} & \boldsymbol{0} \\ \lambda\boldsymbol{B} & \lambda\boldsymbol{I}_q \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{0} \\ -\boldsymbol{B} & \lambda\boldsymbol{I}_q \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_p & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_q \end{pmatrix}=\begin{pmatrix} \lambda\boldsymbol{I}_p & \boldsymbol{A} \\ \boldsymbol{0} & \lambda\boldsymbol{I}_q-\boldsymbol{BA} \end{pmatrix}$$

所以

$$\begin{vmatrix} \lambda\boldsymbol{I}_p-\boldsymbol{AB} & \boldsymbol{0} \\ \lambda\boldsymbol{B} & \lambda\boldsymbol{I}_q \end{vmatrix}=\begin{vmatrix} \lambda\boldsymbol{I}_p & \boldsymbol{A} \\ \boldsymbol{0} & \lambda\boldsymbol{I}_q-\boldsymbol{BA} \end{vmatrix}$$

$$\lambda^q\left|\lambda\boldsymbol{I}_p-\boldsymbol{AB}\right|=\lambda^p\left|\lambda\boldsymbol{I}_q-\boldsymbol{BA}\right|$$

由复系数多项式的因式分解定理可得出，两个关于λ的方程$|\lambda \boldsymbol{I}_p-\boldsymbol{AB}|=0$和$|\lambda \boldsymbol{I}_q-\boldsymbol{BA}|=0$有着完全相同的非零根(若有重根，则它们的重数也相同)，故而$\boldsymbol{AB}$和$\boldsymbol{BA}$有相同的非零特征值。

- 例1.6.2　设$\boldsymbol{A}$和$\boldsymbol{B}$为两个$p\times p$矩阵，则$\boldsymbol{AB}$和$\boldsymbol{BA}$有完全相同的特征值。
- 例1.6.3　设$\boldsymbol{a}=(2,-4,1)'$, $\boldsymbol{b}=(3,5,-1)'$，试求$\boldsymbol{ab}'$的特征值。
- 解　由于

$$\boldsymbol{b}'\boldsymbol{a}=(3,5,-1)\begin{pmatrix}2\\-4\\1\end{pmatrix}=-15$$

因此，$\boldsymbol{ab}'$有一个非零特征值-15，而另两个特征值为零。

- (3)若$\boldsymbol{A}$为实对称矩阵，则$\boldsymbol{A}$的特征值全为实数，p个特征值按大小依次表示为$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_p$。若$\lambda_i\neq\lambda_j$，则相应的特征向量$\mathbf{x}_i$和$\mathbf{x}_j$必正交，即$\mathbf{x}_i'\mathbf{x}_j=0$。
- (4)若$\boldsymbol{A}=\text{diag}(a_{11},a_{22},\cdots,a_{pp})$，则$a_{11},a_{22},\cdots,a_{pp}$为$\boldsymbol{A}$的$p$个特征值，相应的特征向量分别为$\boldsymbol{e}_1=(1,0,\cdots,0)'$，$\boldsymbol{e}_2=(0,1,0,\cdots,0)'$，$\cdots$，$\boldsymbol{e}_p=(0,\cdots,0,1)'$。
- (5)$|\boldsymbol{A}|=\prod_{i=1}^{p}\lambda_i$，即$\boldsymbol{A}$的行列式等于其特征值的乘积。可见，$\boldsymbol{A}$为非退化矩阵，当且仅当$\boldsymbol{A}$的特征值均不为零；$\boldsymbol{A}$为退化矩阵，当且仅当$\boldsymbol{A}$至少有一个特征值为零。
- 例1.6.4　设方阵$\boldsymbol{A}$：$p\times p$的p个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_p$，试证：

  (i)若$\boldsymbol{A}$可逆，相应于$\lambda_1,\lambda_2,\cdots,\lambda_p$的特征向量分别为$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_p$，则$\boldsymbol{A}^{-1}$的$p$个特征值为，相应的特征向量仍可为$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_p$；

  (ii)若$\boldsymbol{A}$为幂等矩阵，则$\boldsymbol{A}$的特征值为0或1；

  (iii)若$\boldsymbol{A}$为正交矩阵，则$\boldsymbol{A}$的特征值为1或−1。

❖ (6)若$\boldsymbol{A}$为p阶对称矩阵，则存在正交矩阵$\boldsymbol{T}$及对角矩阵$\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_p)$，使得

$$\boldsymbol{A}=\boldsymbol{T\Lambda T'}$$

得

$$\boldsymbol{AT}=\boldsymbol{T\Lambda}$$

记$\boldsymbol{T}=(\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_p)$，于是

$$(\boldsymbol{At}_1,\boldsymbol{At}_2,\cdots,\boldsymbol{At}_p)=(\lambda_1\boldsymbol{t}_1,\lambda_2\boldsymbol{t}_2,\cdots,\lambda_p\boldsymbol{t}_p)$$

$$\boldsymbol{At}_i=\lambda_i\boldsymbol{t}_i,\quad i=1,2,\cdots,p$$

这表明$\lambda_1,\lambda_2,\cdots,\lambda_p$是$\boldsymbol{A}$的$p$个特征值，而$\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_p$为相应的一组正交单位特征向量。

谱分解

$$\boldsymbol{A}=\boldsymbol{T\Lambda T}'=(\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_p)\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p\end{pmatrix}\begin{pmatrix}\boldsymbol{t}_1'\\ \boldsymbol{t}_2'\\ \vdots\\ \boldsymbol{t}_p'\end{pmatrix}=\sum_{i=1}^{p}\lambda_i\boldsymbol{t}_i\boldsymbol{t}_i'$$

奇异值分解

- (7)设$\boldsymbol{A}$：$p\times q$，rank($\boldsymbol{A}$)=k，则存在$\boldsymbol{U}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k)$：$p\times k$，$\boldsymbol{V}=(\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k)$：$q\times k$，$\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_k)$，使得

$$\boldsymbol{A}=\boldsymbol{U\Lambda V}'=\sum_{i=1}^{k}\lambda_i\boldsymbol{u}_i\boldsymbol{v}_i'$$

其中，$\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k$是一组$p$维正交单位向量，$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k$是一组$q$维正交单位向量，$\lambda_i>0$, $i=1,2,\cdots,k$。λ_i称为$\boldsymbol{A}$的奇异值。

- $$\boldsymbol{AA}'=\boldsymbol{U\Lambda}^2\boldsymbol{U}'，\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{V\Lambda}^2\boldsymbol{V}'$$

$$\boldsymbol{AA}'\boldsymbol{U}=\boldsymbol{U\Lambda}^2，\boldsymbol{A}'\boldsymbol{AV}=\boldsymbol{V\Lambda}^2$$

$$\boldsymbol{AA}'\boldsymbol{u}_i=\lambda_i^2\boldsymbol{u}_i，\quad i=1,2,\cdots,k$$

$$\boldsymbol{A}'\boldsymbol{A}\boldsymbol{v}_i=\lambda_i^2\boldsymbol{v}_i，\quad i=1,2,\cdots,k$$

二、矩阵的迹

- 设$\boldsymbol{A}$为p阶方阵，则$\boldsymbol{A}$的迹定义为

$$\text{tr}(\boldsymbol{A})=a_{11}+a_{22}+\cdots+a_{pp}$$

- 方阵的迹具有下述基本性质：

- (1)$\text{tr}(\boldsymbol{AB})=\text{tr}(\boldsymbol{BA})$。特别地，$\text{tr}(\boldsymbol{ab}')=\boldsymbol{b}'\boldsymbol{a}$。
- (2)$\text{tr}(\boldsymbol{A})=\text{tr}(\boldsymbol{A}')$。
- (3)$\text{tr}(\boldsymbol{A}+\boldsymbol{B})=\text{tr}(\boldsymbol{A})+\text{tr}(\boldsymbol{B})$。
- (4) $\text{tr}\left(\sum_{i=1}^{k}\boldsymbol{A}_i\right)=\sum_{i=1}^{k}\text{tr}(\boldsymbol{A}_i)$。

- (5)设$\boldsymbol{A}=(a_{ij})$为$p\times q$矩阵，则

$$\operatorname{tr}\left(\boldsymbol{A}'\boldsymbol{A}\right)=\operatorname{tr}\left(\boldsymbol{A}\boldsymbol{A}'\right)=\sum_{i=1}^{p}\sum_{j=1}^{q}a_{ij}^{2}$$

- (6)设$\lambda_1,\lambda_2,\cdots,\lambda_p$为方阵$\boldsymbol{A}$的特征值，则

$$\operatorname{tr}(\boldsymbol{A})=\lambda_1+\lambda_2+\cdots+\lambda_p$$

证明

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_p)$$

比较等式两边λ^{p-1}项的系数即得。

- (7)若$\boldsymbol{A}$为投影矩阵，则

$$\operatorname{tr}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A})$$

§1.7 正定矩阵和非负定矩阵

- 设$\boldsymbol{A}$是p阶对称矩阵，$\boldsymbol{x}$是一p维向量，则$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$称为$\boldsymbol{A}$的二次型。
- 若对一切$\boldsymbol{x}\neq\boldsymbol{0}$，有$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}>0$，则称$\boldsymbol{A}$为正定矩阵，记作$\boldsymbol{A}>0$。
- 若对一切$\boldsymbol{x}$，有$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}\geq 0$，则称$\boldsymbol{A}$为非负定矩阵，记作$\boldsymbol{A}\geq 0$。
- 对非负定矩阵$\boldsymbol{A}$和$\boldsymbol{B}$，$\boldsymbol{A}>\boldsymbol{B}$表示$\boldsymbol{A}-\boldsymbol{B}>0$；$\boldsymbol{A}\geq\boldsymbol{B}$表示$\boldsymbol{A}-\boldsymbol{B}\geq 0$。

基本性质

- (1)设$\boldsymbol{A}'=\boldsymbol{A}$，则$\boldsymbol{A}>0$(或$\geq 0$) $\Leftrightarrow \lambda_i>0$(或$\geq 0$)，$i=1,2,\cdots,p$。
- (2)设$\boldsymbol{A}\geq 0$，则$\boldsymbol{A}$的秩等于$\boldsymbol{A}$的正特征值个数。
- (3)若$\boldsymbol{A}>0$，则$\boldsymbol{A}^{-1}>0$。
- (4)设$\boldsymbol{A}\geq 0$，则$\boldsymbol{A}>0$，当且仅当$|\boldsymbol{A}|\neq 0$。
- (5)若$\boldsymbol{A}>0$(或≥ 0)，则$|\boldsymbol{A}|>0$(或≥ 0)。
- (6)$\boldsymbol{BB}'\geq 0$，对一切矩阵$\boldsymbol{B}$成立。
- (7)若$\boldsymbol{A}>0$(或≥ 0)，则存在$\boldsymbol{A}^{1/2}>0$ (或≥ 0)，使得$\boldsymbol{A}=\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2}$，$\boldsymbol{A}^{1/2}$称为$\boldsymbol{A}$的平方根矩阵。
- 注：当$p=1$时，$\boldsymbol{A}=a$是一个正数(或非负数),可有$a=a^{1/2}a^{1/2}$,而$a^{1/2}$也是一个正数(或非负数)。
- (8)设$\boldsymbol{A}\geq 0$是p阶秩为r的矩阵，则存在一个秩为r(即列满秩)的$p\times r$矩阵$\boldsymbol{B}$，使得$\boldsymbol{A}=\boldsymbol{BB}'$。

§1.8 特征值的极值问题

- (1)柯西-许瓦兹 (Cauchy−Schwarz) 不等式 设$\boldsymbol{x}$和$\boldsymbol{y}$是两个p维向量，则

$$(\boldsymbol{x}'\boldsymbol{y})^2 \le (\boldsymbol{x}'\boldsymbol{x})(\boldsymbol{y}'\boldsymbol{y})$$

等号成立当且仅当$\boldsymbol{y}=c\boldsymbol{x}$（或$\boldsymbol{x}=c\boldsymbol{y}$），这里$c$为一常数。

- (2)推广的柯西-许瓦兹不等式 设$\boldsymbol{B}>0$，则

$$(\boldsymbol{x}'\boldsymbol{y})^2 \le (\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x})(\boldsymbol{y}'\boldsymbol{B}^{-1}\boldsymbol{y})$$

等号成立当且仅当$\boldsymbol{x}=c\boldsymbol{B}^{-1}\boldsymbol{y}$（或$\boldsymbol{y}=c\boldsymbol{B}\boldsymbol{x}$），这里$c$为一常数。

❖ (3)设$\boldsymbol{A}$是p阶对称矩阵，其特征值依次是$\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_p$，相应的一组正交特征向量是$\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_p$，则

(i)

$$\max_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}\left(=\max_{\|\boldsymbol{x}\|=1}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}\right)=\lambda_1$$ （当$\boldsymbol{x}=\boldsymbol{t}_1$时达到）

$$\min_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}\left(=\min_{\|\boldsymbol{x}\|=1}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}\right)=\lambda_p$$ （当$\boldsymbol{x}=\boldsymbol{t}_p$时达到）

(ii)

$$\max_{\substack{\boldsymbol{x}'\boldsymbol{t}_k=0\\ k=1,\cdots,i-1\\ \boldsymbol{x}\neq\boldsymbol{0}}}\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}\left(=\max_{\substack{\boldsymbol{x}'\boldsymbol{t}_k=0\\ k=1,\cdots,i-1\\ \|\boldsymbol{x}\|=1}}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}\right)=\lambda_i$$ （当$\boldsymbol{x}=\boldsymbol{t}_i$时达到）

$$i=2,3,\cdots,p$$

❖ (4)设$\boldsymbol{A}$是p阶对称矩阵，$\boldsymbol{B}$是p阶正定矩阵，$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_p$是$\boldsymbol{B}^{-1}\boldsymbol{A}$的$p$个特征值，相应的一组特征向量是$\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_p$，满足$\boldsymbol{t}_i' \boldsymbol{B} \boldsymbol{t}_j = 0, 1 \leq i \neq j \leq p$，则

(i)

$$\max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}' \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}' \boldsymbol{B} \boldsymbol{x}} = \mu_1$$ （当$\boldsymbol{x} = \boldsymbol{t}_1$时达到）

$$\min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}' \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}' \boldsymbol{B} \boldsymbol{x}} = \mu_p$$ （当$\boldsymbol{x} = \boldsymbol{t}_p$时达到）

(ii)

$$\max_{\substack{\boldsymbol{x}' \boldsymbol{B} \boldsymbol{t}_k = 0 \\ k = 1, \cdots, i-1 \\ \boldsymbol{x} \neq \boldsymbol{0}}} \frac{\boldsymbol{x}' \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}' \boldsymbol{B} \boldsymbol{x}} = \mu_i$$ （当$\boldsymbol{x} = \boldsymbol{t}_i$时达到）, $i = 2, 3, \cdots, p$