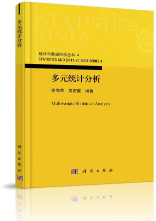
第1章
矩阵运算

李高荣

北京师范大学统计学院

E-mail: *ligaorong@bnu.edu.cn*



微信公众号：BNUlgr

- 扫二维码获取在线课件和相关教学电子资源
- 请遵守电子资源使用协议

- 矩阵是多元统计分析一个十分重要的工具，本章主要介绍多元统计分析中有关矩阵论的一些预备知识。

定义1.1：线性空间

设$\mathcal{H}$为$\mathcal{R}_n$的一个子集，如果它对向量加法和数乘两种运算具有封闭性，即

(1) 对任意$\boldsymbol{x} \in \mathcal{H}$和$\boldsymbol{y} \in \mathcal{H}$，必有$\boldsymbol{x} + \boldsymbol{y} \in \mathcal{H}$;

(2) 对一切实数c和任意$\boldsymbol{x} \in \mathcal{H}$，都有$c\boldsymbol{x} \in \mathcal{H}$。

这时，把满足上面两种运算的子集$\mathcal{H}$称为线性空间。

记全体$n \times 1$实向量组成的集合为$\mathcal{R}_n$。

- 记$\mathcal{S}_0$是由$\mathcal{R}_n$中向量组$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 的一切可能的线性组合构成的集合，即

$$\mathcal{S}_0=\left\{\boldsymbol{x}=\sum_{i=1}^{k}\alpha_i\boldsymbol{a}_i,\quad \alpha_1,\cdots,\alpha_k\text{均为实数}\right\}.$$

- 易证：$\mathcal{S}_0$也是线性空间，称$\mathcal{S}_0$为$\mathcal{R}_n$的一个子空间。
- 若将$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$排成一个$n\times k$矩阵$\mathbf{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k)$，则$\mathcal{S}_0$可表示为$\mathcal{S}_0=\{\boldsymbol{x}=\mathbf{A}\boldsymbol{t},\ \boldsymbol{t}\in\mathcal{R}_k\}$，它是矩阵$\mathbf{A}$的列向量张成的子空间，记为$\mathcal{S}_0=\mathcal{M}(\mathbf{A})$。

定义1.2：线性相关/线性无关

设$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$为$\mathcal{R}_n$中的一组向量，若存在不全为零的实数$\alpha_1, \cdots, \alpha_k$，使得

$$\alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k = 0,$$

则称向量组$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ 线性相关，否则称它们是线性无关的。

- 如果子空间$\mathcal{S}_0$由一组线性无关的向量$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$张成，则称$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$为$\mathcal{S}_0$ 的一组基，k称为$\mathcal{S}_0$的维数，记作$k = \dim(\mathcal{S}_0)$。
- 因此，$\dim(\mathcal{M}(\mathbf{A})) = \mathrm{rank}(\mathbf{A})$。

- 内积： 对$\mathcal{R}_n$中的任意两个n维向量$\boldsymbol{a}'$ 和$\boldsymbol{b}'$，内积定义为：$(\boldsymbol{a},\boldsymbol{b}) = \boldsymbol{a}'\boldsymbol{b} = \sum_{i=1}^{n} a_i b_i$。
- 长度或模： $(\boldsymbol{a}'\boldsymbol{a})^{1/2} = \left(\sum\limits_{i=1}^{n} a_i^2\right)^{1/2}$，记作$\|\boldsymbol{a}\|$。
- 记$\boldsymbol{b} = \dfrac{\boldsymbol{a}}{\|\boldsymbol{a}\|}$， 则$(\boldsymbol{b},\boldsymbol{b}) = 1$， 并称$\boldsymbol{b}$为$\boldsymbol{a}$的标准化后的向量。
- 正交： 若$(\boldsymbol{a},\boldsymbol{b}) = 0$， 则称$\boldsymbol{a}$与$\boldsymbol{b}$正交， 记为$\boldsymbol{a} \perp \boldsymbol{b}$。
- 若$\boldsymbol{a}$与子空间$\mathcal{S}$中的每一个向量正交， 则称$\boldsymbol{a}$正交于$\mathcal{S}$， 记为$\boldsymbol{a} \perp \mathcal{S}$。

定义1.3：正交补空间

设$\mathcal{S}$为一子空间，称子空间$\mathcal{S}^{\perp}=\{\boldsymbol{x}:\boldsymbol{x}\perp\mathcal{S}\}$为$\mathcal{S}$的正交补空间。

定义1.4：正交矩阵

设$\mathbf{P}$为$n\times n$的矩阵，若$\mathbf{P}'\mathbf{P}=\mathbf{I}_n$，则称$\mathbf{P}$为正交矩阵。

1. $\mathbf{A}^{-1}=\mathbf{A}'$;
2. $\mathbf{A}'\mathbf{A}=\mathbf{I}_n$，$\mathbf{A}\mathbf{A}'=\mathbf{I}_n$，即$\mathbf{A}$的所有列向量相互正交，所有行向量也相互正交，各列向量和各行向量的模为1。

- 对于$n\times n$的方阵$\mathbf{A}$，若$\mathbf{A}$ 的列向量$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 是相互正交的，即$\boldsymbol{a}_i'\boldsymbol{a}_j=0$，则对其列向量进行标准化

$$\boldsymbol{p}_i=\frac{\boldsymbol{a}_i}{\|\boldsymbol{a}_i\|},\qquad i=1,\cdots,n,$$

便得到一个正交矩阵：$\mathbf{P}=(\boldsymbol{p}_1,\cdots,\boldsymbol{p}_n)$，显然$\mathbf{P}'\mathbf{P}=\mathbf{I}_n$。

定理1.2.1：

对任意矩阵$\mathbf{A}$，恒有$\mathcal{M}(\mathbf{A})=\mathcal{M}(\mathbf{AA}')$。

- 两种特殊运算：Kronecker乘积与拉直运算。

定义1.5：Kronecker乘积

设$\mathbf{A}=(a_{ij})$ 和$\mathbf{B}=(b_{ij})$ 分别为$m\times n$ 和$p\times q$的矩阵，定义$mp\times nq$的矩阵$\mathbf{C}=(a_{ij}\mathbf{B})$，称为矩阵$\mathbf{A}$ 和$\mathbf{B}$ 的Kronecker 乘积，记为$\mathbf{C}=\mathbf{A}\otimes\mathbf{B}$，即

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

Kronecker乘积性质

(1) (结合律) $(\mathbf{A}\otimes\mathbf{B})\otimes\mathbf{C}=\mathbf{A}\otimes(\mathbf{B}\otimes\mathbf{C})$;

(2) (分配律) $(\mathbf{A}_1+\mathbf{A}_2)\otimes\mathbf{B}=(\mathbf{A}_1\otimes\mathbf{B})+(\mathbf{A}_2\otimes\mathbf{B})$， $\mathbf{A}\otimes(\mathbf{B}_1+\mathbf{B}_2)=(\mathbf{A}\otimes\mathbf{B}_1)+(\mathbf{A}\otimes\mathbf{B}_2)$;

(3) (数量乘法) 对任意实数α和β，有

$$(\alpha\mathbf{A})\otimes(\beta\mathbf{B})=\alpha\beta(\mathbf{A}\otimes\mathbf{B});$$

(4) (矩阵乘法) $(\mathbf{A}_1\otimes\mathbf{B}_1)(\mathbf{A}_2\otimes\mathbf{B}_2)=(\mathbf{A}_1\mathbf{A}_2)\otimes(\mathbf{B}_1\mathbf{B}_2)$;

(5) (矩阵转置) $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$;

(6) (逆矩阵) $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$;

(7) (矩阵的迹) $\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A}) \cdot \mathrm{tr}(\mathbf{B})$;

(8) (行列式) $|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^n \cdot |\mathbf{B}|^m$，其中$\mathbf{A}$和$\mathbf{B}$分别为$m$阶和$n$阶方阵。

定义1.6：矩阵的拉直运算

设矩阵$\mathbf{A} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$是一个$m \times n$矩阵，其中$\boldsymbol{a}_i = (a_{i1}, \cdots, a_{im})'$，且$i = 1, \cdots, n$。把矩阵$\mathbf{A}$按列向量$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$依次排成一个$mn \times 1$的向量，即

$$\mathrm{Vec}(\mathbf{A}) = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix},$$

则称$\mathrm{Vec}(\mathbf{A})$为矩阵$\mathbf{A}$的拉直运算。

(1) $\mathrm{Vec}(\mathbf{A}+\mathbf{B})=\mathrm{Vec}(\mathbf{A})+\mathrm{Vec}(\mathbf{B})$；

(2) 对于任意实数α，有$\mathrm{Vec}(\alpha\mathbf{A})=\alpha\mathrm{Vec}(\mathbf{A})$；

(3) $\mathrm{tr}(\mathbf{AB})=(\mathrm{Vec}(\mathbf{A}'))'\mathrm{Vec}(\mathbf{B})$；

(4) 设$\boldsymbol{a}$和$\boldsymbol{b}$分别为$n\times 1$ 和$m\times 1$的向量，则

$$\mathrm{Vec}(\boldsymbol{a}\boldsymbol{b}')=\boldsymbol{b}\otimes\boldsymbol{a};$$

(5) $\mathrm{Vec}(\mathbf{ABC})=(\mathbf{C}'\otimes\mathbf{A})\mathrm{Vec}(\mathbf{B})$。

定义1.7：对称矩阵

设$\mathbf{A}$为$n\times n$的矩阵，如果$\mathbf{A}'=\mathbf{A}$，则称$\mathbf{A}$为对称矩阵。

- 实对称矩阵$\mathbf{A}$的不同特征值对应的特征向量是正交的。

定义1.8：正定矩阵

设$\mathbf{A}$为$n\times n$的对称矩阵，如果对于一切$n\times 1$的非零向量$\boldsymbol{x}\in\mathbb{R}^n$，都有$\boldsymbol{x}'\mathbf{A}\boldsymbol{x}>0$，则称$\mathbf{A}$为正定矩阵，记作$\mathbf{A}>0$。

对称矩阵$\mathbf{A}$为正定矩阵当且仅当下列条件之一成立：

1. $\mathbf{A}$的所有特征值都大于0；
2. 存在$n \times n$的可逆矩阵$\mathbf{B}$，使得$\mathbf{A} = \mathbf{B}\mathbf{B}'$。

定义1.9：半正定矩阵

设$\mathbf{A}$为$n \times n$的对称矩阵，如果对于一切$n \times 1$的非零向量$\boldsymbol{x} \in \mathbb{R}^n$，都有$\boldsymbol{x}'\mathbf{A}\boldsymbol{x} \geq 0$，则称$\mathbf{A}$为半正定矩阵或非负定矩阵，记作$\mathbf{A} \geq 0$。

1. $\mathbf{A}$的所有特征值大于等于0；
2. 存在$n \times r$的列满秩矩阵$\mathbf{B}$，使得$\mathbf{A} = \mathbf{B}\mathbf{B}'$，这里，$r = \operatorname{rank}(\mathbf{A})$表示矩阵$\mathbf{A}$的秩。

定理1.2.2：Schmidt 三角化分解

设$\mathbf{A}$为$n \times m$矩阵，其中$n \geq m$，则矩阵$\mathbf{A}$可分解为

$$\mathbf{A} = \mathbf{QR},$$

其中$\mathbf{Q}$为$n \times m$的列正交矩阵，$\mathbf{R}$为上三角矩阵，其对角线元素非负；

当$\operatorname{rank}(\mathbf{A}) = m$时，$\mathbf{R}$ 的对角线元素全为正。

定理1.2.3：Cholesky 分解

设$\mathbf{C}$为$n \times n$正定矩阵，即$\mathbf{C} > 0$，则存在一个下三角矩阵$\mathbf{T}$，使得

$$\mathbf{C} = \mathbf{T}\mathbf{T}'.$$

证明： 若$\mathbf{C} > 0$，则存在非奇异阵$\mathbf{A}$ 使得$\mathbf{C} = \mathbf{A}'\mathbf{A}$。对$\mathbf{A}$作Schmidt三角化分解，得$\mathbf{A} = \mathbf{Q}\mathbf{R}$，其中$\mathbf{Q}$为正交矩阵，$\mathbf{R}$为上三角矩阵。于是

$$\mathbf{C} = \mathbf{A}'\mathbf{A} = \mathbf{R}'\mathbf{Q}'\mathbf{Q}\mathbf{R} = \mathbf{R}'\mathbf{R}.$$

记$\mathbf{T} = \mathbf{R}'$，故$\mathbf{T}$为下三角阵，定理得证。

定理1.2.4：矩阵谱分解

设$\mathbf{A}$ 是一个$n \times n$对称矩阵，则存在一个$n \times n$的正交方阵$\mathbf{P}$，使得

$$\mathbf{A} = \mathbf{P}\boldsymbol{\Lambda}_n\mathbf{P}' = \sum_{i=1}^{n} \lambda_i \varphi_i \varphi_i', \tag{1}$$

这里

$$\mathbf{P} = (\varphi_1, \cdots, \varphi_n), \qquad \boldsymbol{\Lambda}_n = \mathrm{diag}(\lambda_1, \cdots, \lambda_n),$$

其中$\lambda_1, \cdots, \lambda_n$为$\mathbf{A}$的特征值，$\varphi_1, \cdots, \varphi_n$为相应的正交标准化的特征向量。

推论1.2.1

设$\mathbf{A}$ 是一个$n\times n$对称矩阵，则

$$\mathrm{tr}(\mathbf{A})=\sum_{i=1}^{n}\lambda_i,\qquad |\mathbf{A}|=\prod_{i=1}^{n}\lambda_i.$$

- 当$\mathrm{rank}(\mathbf{A})=r<n$时，则$\mathbf{A}$仅有$r$个非零特征值：$\lambda_1,\cdots,\lambda_r$。
- 令$\mathbf{\Lambda}_r=\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$，对$\mathbf{P}$进行相应的分块，$\mathbf{P}=(\mathbf{P}_1,\mathbf{P}_2)$，其中$\mathbf{P}_1$为$n\times r$的矩阵，则

$$\mathbf{A}=\mathbf{P}\begin{pmatrix}\mathbf{\Lambda}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{P}'=\mathbf{P}_1\mathbf{\Lambda}_r\mathbf{P}_1'.$$

定理1.2.5：奇异值分解

设$\mathbf{A}$为$m \times n$的矩阵，且$\mathrm{rank}(\mathbf{A}) = r$，则存在两个正交方阵$\mathbf{P}$和$\mathbf{Q}$，使得

$$\mathbf{A} = \mathbf{P}\begin{pmatrix} \mathbf{\Lambda}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{Q}' = \mathbf{P}_1\mathbf{\Lambda}_r\mathbf{Q}_1', \tag{2}$$

这里，$\mathbf{P} = (\mathbf{P}_1, \mathbf{P}_2)$，$\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2)$，$\mathbf{\Lambda}_r = \mathrm{diag}(\lambda_1, \cdots, \lambda_r)$，其中$\lambda_i > 0$ $(i = 1, \cdots, r)$ 为$\mathbf{A}$的奇异值，$\lambda_1^2, \cdots, \lambda_r^2$为$\mathbf{A}'\mathbf{A}$的非零特征值，$\mathbf{P}_1$和$\mathbf{Q}_1$的列向量分别为$\mathbf{A}\mathbf{A}'$和$\mathbf{A}'\mathbf{A}$ 对应于r个非零特征值的标准正交化的特征向量。

- 对于相容线性方程组：

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \tag{3}$$

其中

  - $\mathbf{A}$是$m \times n$ 的矩阵
  - 秩$\mathrm{rank}(\mathbf{A}) = r \le \min(m, n)$

- 当$r = m = n$ 时，方程组有唯一解$\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$。
- 当$\mathbf{A}$ 不可逆或根本不是方阵时，则方程组(3) 有无穷多解。
- 如何用$\mathbf{A}$和$\boldsymbol{b}$通过简单的形式表征方程组(3)的全体解？

定义1.10：广义逆$\mathbf{A}^-$

对一个$m \times n$的矩阵$\mathbf{A}$，一切满足方程组

$$\mathbf{AXA} = \mathbf{A} \tag{4}$$

的矩阵$\mathbf{X}$，称为矩阵$\mathbf{A}$的广义逆，记为$\mathbf{A}^-$。

定理1.2.6

设$\mathbf{A}$为$m \times n$矩阵，$\mathrm{rank}(\mathbf{A}) = r$，若

$$\mathbf{A} = \mathbf{P}\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{Q},$$

这里，$\mathbf{P}$和$\mathbf{Q}$分别为$m \times m$和$n \times n$的可逆矩阵，则

$$\mathbf{A}^- = \mathbf{Q}^{-1}\begin{pmatrix} \mathbf{I}_r & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\mathbf{P}^{-1},$$

其中$\mathbf{B}$、$\mathbf{C}$和$\mathbf{D}$为适当阶数的任意矩阵。

推论1.2.2

对任一矩阵$\mathbf{A}$，有

(1) $\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'$与广义逆$(\mathbf{A}'\mathbf{A})^-$的选择无关；

(2) $\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'\mathbf{A} = \mathbf{A}$ 和$\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}' = \mathbf{A}'$。

定理1.2.7

设$\mathbf{A}x=\boldsymbol{b}$为一相容方程组，则

(1) 对任一广义逆$\mathbf{A}^-$，$x=\mathbf{A}^-\boldsymbol{b}$必为解；

(2) 齐次方程组$\mathbf{A}x=\mathbf{0}$的通解为：$x=(\mathbf{I}-\mathbf{A}^-\mathbf{A})z$，这里$z$为任意的向量，$\mathbf{A}^-$为任意固定的一个广义逆；

(3) $\mathbf{A}x=\boldsymbol{b}$的通解为：$x=\mathbf{A}^-\boldsymbol{b}+(\mathbf{I}-\mathbf{A}^-\mathbf{A})z$，其中$\mathbf{A}^-$为任一固定的广义逆，$z$为任意向量。

- 一般说来广义逆$\mathbf{A}^-$有无穷多个。在这无穷多个$\mathbf{A}^-$中，有一个$\mathbf{A}^-$占有特殊的地位，它就是Moore-Penrose广义逆。

定义1.11：Moore-Penrose广义逆$\mathbf{A}^+$

设$\mathbf{A}$为任一矩阵，若$\mathbf{X}$满足下述四个条件：

$$\mathbf{AXA}=\mathbf{A},\quad \mathbf{XAX}=\mathbf{X},\quad (\mathbf{AX})'=\mathbf{AX},\quad (\mathbf{XA})'=\mathbf{XA}, \tag{5}$$

则称矩阵$\mathbf{X}$为$\mathbf{A}$的Moore-Penrose广义逆，记为$\mathbf{A}^+$。

定理1.2.8

对于任意$m\times n$的矩阵$\mathbf{A}$，$\mathbf{A}^+$是唯一的；若$\mathbf{A}$的奇异值分解为：

$$\mathbf{A}=\mathbf{P}\begin{pmatrix}\mathbf{\Lambda}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q}',$$

这里，$\mathrm{rank}(\mathbf{A})=r$，$\mathbf{P}$和$\mathbf{Q}$分别为$m\times m$和$n\times n$的正交矩阵，则

$$\mathbf{A}^+=\mathbf{Q}\begin{pmatrix}\mathbf{\Lambda}_r^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{P}'.$$

因为$\mathbf{A}^{+}$是一个特殊的$\mathbf{A}^{-}$。因此，它除了具有$\mathbf{A}^{-}$的全部性质外，还有下列性质：

1. $(\mathbf{A}^{+})^{+} = \mathbf{A}$;
2. $(\mathbf{A}^{+})' = (\mathbf{A}')^{+}$;
3. $\mathbf{I} \geq \mathbf{A}^{+}\mathbf{A}$;
4. $\mathrm{rank}(\mathbf{A}^{+}) = \mathrm{rank}(\mathbf{A})$;
5. $\mathbf{A}^{+} = (\mathbf{A}'\mathbf{A})^{+}\mathbf{A}' = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{+}$;
6. $(\mathbf{A}'\mathbf{A})^{+} = \mathbf{A}^{+}(\mathbf{A}')^{+}$;
7. 设$\boldsymbol{a}$为一非零向量，则$\boldsymbol{a}^{+} = \boldsymbol{a}'/\|\boldsymbol{a}\|^{2}$。

Moore-Penrose广义逆$\mathbf{A}^+$

定理1.2.9

设$\mathbf{A}$为$n \times n$的对称矩阵，若$\mathbf{A}$ 的谱分解为：

$$\mathbf{A} = \mathbf{P} \begin{pmatrix} \boldsymbol{\Lambda}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P}',$$

则

$$\mathbf{A}^+ = \mathbf{P} \begin{pmatrix} \boldsymbol{\Lambda}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P}',$$

这里，$\mathrm{rank}(\mathbf{A}) = r$，$\mathbf{P}$为$n \times n$的正交矩阵，$\boldsymbol{\Lambda}_r$ 为$r \times r$的对角矩阵，其对角元素为$\mathbf{A}$的非零特征值。

定义1.12：幂等阵

若方阵$\mathbf{A}$满足$\mathbf{A}^2=\mathbf{A}$，则称$\mathbf{A}$为幂等阵。如果$\mathbf{A}$又是对称矩阵，即$\mathbf{A}'=\mathbf{A}$，则称$\mathbf{A}$为对称幂等阵。

若方阵$\mathbf{A}$为幂等阵，则有下面的性质：

1. 幂等阵$\mathbf{A}$的特征值只能为0或1；
2. $\mathbf{I}-\mathbf{A}$，$\mathbf{A}^-\mathbf{A}$，$\mathbf{A}\mathbf{A}^-$，$\mathbf{I}-\mathbf{A}^-\mathbf{A}$ 和$\mathbf{I}-\mathbf{A}\mathbf{A}^-$ 都是幂等阵。特别，$\mathbf{A}^+\mathbf{A}$，$\mathbf{A}\mathbf{A}^+$，$\mathbf{I}-\mathbf{A}^+\mathbf{A}$ 和$\mathbf{I}-\mathbf{A}\mathbf{A}^+$都是幂等阵；
3. 若$\mathbf{A}$为对称幂等阵，则$\mathbf{A}^+=\mathbf{A}$。

定理1.2.10

设$\mathbf{P}$为$n\times n$对称幂等阵，$\mathrm{rank}(\mathbf{P})=r$，则存在秩为$r$ 的$n\times r$矩阵$\mathbf{A}$，则有

$$\mathbf{P}=\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'.$$

推论1.2.3

设$\mathbf{A}$是一个$n\times n$ 对称幂等矩阵，$\mathrm{rank}(\mathbf{A})=r$，则$\mathbf{A}$的所有非负特征值皆为1，且存在一个$n\times r$列正交矩阵$\mathbf{P}$，使得$\mathbf{A}=\mathbf{P}\mathbf{P}'$。

定义1.13：正交投影

设$\boldsymbol{x} \in \mathcal{R}_n$，$\mathcal{S}$为$\mathcal{R}_n$的一个线性子空间。对$\boldsymbol{x}$作分解

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}, \qquad \boldsymbol{y} \in \mathcal{S}, \qquad \boldsymbol{z} \in \mathcal{S}^{\perp}, \tag{6}$$

则称$\boldsymbol{y}$为$\boldsymbol{x}$在$\mathcal{S}$上的正交投影。若$\mathbf{P}$为n阶方阵，使得对一切$\boldsymbol{x} \in \mathcal{R}_n$，都有式(6)定义的$\boldsymbol{y}$满足

$$\boldsymbol{y} = \mathbf{P}\boldsymbol{x},$$

则称$\mathbf{P}$为向$\mathcal{S}$的正交投影阵。

- 对$\mathcal{R}_n$的任一子空间$\mathcal{S}$，都可以找到矩阵$\mathbf{A}$，使得$\mathcal{S} = \mathcal{M}(\mathbf{A})$。因此，正交投影阵可由矩阵$\mathbf{A}$来表示。

定理1.2.11

设$\mathbf{A}$为$n \times m$矩阵，$\mathbf{P}_\mathbf{A}$为向$\mathcal{M}(\mathbf{A})$的正交投影阵，则$\mathbf{P}_\mathbf{A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$。

定理1.2.12

若$\mathbf{P}_\mathbf{A}$为向$\mathcal{M}(\mathbf{A})$的正交投影阵，则$\mathbf{I} - \mathbf{P}_\mathbf{A}$为向$\mathcal{M}(\mathbf{A})$的正交补空间上的正交投影阵。

- 设$\mathbf{A}$是一个$p \times p$矩阵，对$\mathbf{A}$ 进行如下分块：

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

其中

- $\mathbf{A}_{11}$为$p_1 \times p_1$的方阵
- $\mathbf{A}_{22}$分$p_2 \times p_2$的方阵
- $\mathbf{A}_{12}$为$p_1 \times p_2$的矩阵
- $\mathbf{A}_{21}$为$p_2 \times p_1$的矩阵
- $p_1 + p_2 = p$

定理1.2.13：分块矩阵的逆

设$\mathbf{A}$ 为可逆方阵，若$|\mathbf{A}_{11}| \neq 0$，则

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-1} \\ -\mathbf{A}_{22.1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22.1}^{-1} \end{pmatrix};$$

若$|\mathbf{A}_{22}| \neq 0$，则

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11.2}^{-1} & -\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{pmatrix},$$

其中

$$\mathbf{A}_{22.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}, \quad \mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}.$$

推论1.2.4：分块矩阵的行列式

设$\mathbf{A}$ 为可逆方阵，若$|\mathbf{A}_{11}| \neq 0$，则

$$|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22.1}|; \tag{7}$$

若$|\mathbf{A}_{22}| \neq 0$，则

$$|\mathbf{A}| = |\mathbf{A}_{22}| \cdot |\mathbf{A}_{11.2}|. \tag{8}$$

推论1.2.5

设

$$\mathbf{B} = \begin{pmatrix} 1 & \boldsymbol{b}' \\ \boldsymbol{a} & \mathbf{A} \end{pmatrix},$$

其中$\mathbf{A}$ 为$p \times p$的可逆方阵，$\boldsymbol{a}$和$\boldsymbol{b}$为$p \times 1$的向量，则

$$|\mathbf{B}| = |\mathbf{A} - \boldsymbol{a}\boldsymbol{b}'| = |\mathbf{A}|(1 - \boldsymbol{b}'\mathbf{A}^{-1}\boldsymbol{a}) \tag{9}$$

和

$$(\mathbf{A} \pm \boldsymbol{a}\boldsymbol{b}')^{-1} = \mathbf{A}^{-1} \mp \frac{1}{1 \pm \boldsymbol{b}'\mathbf{A}^{-1}\boldsymbol{a}}\mathbf{A}^{-1}\boldsymbol{a}\boldsymbol{b}'\mathbf{A}^{-1}. \tag{10}$$

定理1.2.14

设$\mathbf{A}$ 和$\mathbf{B}$分别为$n \times p$和$p \times n$ 的矩阵，其中$n > p$。则$\mathbf{AB}$和$\mathbf{BA}$具有相同的非零特征值，相应的重数也相等。如果$\boldsymbol{x}$为$\mathbf{AB}$的对应于非零特征值λ的特征向量，则$Y = \mathbf{B}\boldsymbol{x}$为$\mathbf{BA}$的对应于该特征值$\lambda$的特征向量。

证明： 由式(7)和式(8)，得

$$\begin{vmatrix} \lambda\mathbf{I}_n & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_p \end{vmatrix} = \lambda^{n-p}|\lambda\mathbf{I}_p - \mathbf{BA}| = |\lambda\mathbf{I}_n - \mathbf{AB}|.$$

因此，证得$\mathbf{AB}$的特征值等于$\mathbf{BA}$的p个特征值再加上$n-p$个零，即$\mathbf{AB}$和$\mathbf{BA}$具有相同的非零特征值，相应的重数也相等。设$\boldsymbol{x}$为$\mathbf{AB}$的对应于非零特征值λ的特征向量，则

$$\mathbf{AB}\boldsymbol{x} = \lambda\boldsymbol{x}.$$

上式两边同时左乘$\mathbf{B}$得

$$\mathbf{BA}(\mathbf{B}\boldsymbol{x}) = \lambda\mathbf{B}\boldsymbol{x},$$

即证得$\boldsymbol{Y} = \mathbf{B}\boldsymbol{x}$为$\mathbf{BA}$的对应于该特征值$\lambda$的特征向量。

定义1.14：矩阵对变量的微商

假设$\mathbf{Y}=(y_{ij})$为x的一个$n\times m$的实值函数矩阵，则称矩阵

$$\frac{\partial \mathbf{Y}}{\partial x}=\begin{pmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \cdots & \frac{\partial y_{1m}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \cdots & \frac{\partial y_{2m}}{\partial x} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_{n1}}{\partial x} & \frac{\partial y_{n2}}{\partial x} & \cdots & \frac{\partial y_{nm}}{\partial x} \end{pmatrix}_{n\times m}$$

为矩阵$\mathbf{Y}$对变量x的微商。

定义1.15：变量对矩阵的微商

假设$\mathbf{X} = (x_{ij})$为$n \times m$矩阵，$y = f(\mathbf{X})$为$\mathbf{X}$的一个的实值函数，则称矩阵

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1m}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2m}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y}{\partial x_{n1}} & \frac{\partial Y}{\partial x_{n2}} & \cdots & \frac{\partial y}{\partial x_{nm}} \end{pmatrix}_{n \times m}$$

为变量y对矩阵$\mathbf{X}$的微商。

矩阵微商和变换的雅可比

1. 设$\boldsymbol{a}$和$\boldsymbol{x}$均为$n \times 1$ 向量，$y = \boldsymbol{a}'\boldsymbol{x}$，则$\dfrac{\partial y}{\partial \boldsymbol{x}} = \boldsymbol{a}$；
2. 设$\mathbf{A}$为$n \times n$的对称矩阵，$\boldsymbol{x}$为$n \times 1$的向量，$y = \boldsymbol{x}'\mathbf{A}\boldsymbol{x}$，则$\dfrac{\partial y}{\partial \boldsymbol{x}} = 2\mathbf{A}\boldsymbol{x}$；
3. 设$\mathbf{X} = (x_{ij})$ 为$m \times m$ 的矩阵，则$\dfrac{\partial |\mathbf{X}|}{\partial \mathbf{X}} = |\mathbf{X}|(\mathbf{X}^{-1})'$；
4. 设$\mathbf{A}$，$\mathbf{X}$和$\mathbf{B}$分别为$m \times n$，$n \times p$和$p \times q$的矩阵，则
$$\frac{\partial |\mathbf{AXB}|}{\partial \mathbf{X}} = |\mathbf{AXB}|\mathbf{A}'((\mathbf{AXB})^{-1})'\mathbf{B}';$$
5. 设$\mathbf{A}$，$\mathbf{X}$和$\mathbf{B}$分别为$m \times n$，$n \times p$和$p \times q$的矩阵，则
$$\frac{\partial \mathrm{tr}(\mathbf{XAX}')}{\partial \mathbf{X}} = \mathbf{X}(\mathbf{A} + \mathbf{A}'), \quad \frac{\partial \mathrm{tr}(\mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{A}'\mathbf{B}'.$$

定义1.16：向量对向量的微商

假设$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)'$为$n\times 1$向量，$\boldsymbol{y}=f(\boldsymbol{x})=(y_1,y_2,\cdots,y_m)'$为$\boldsymbol{x}$的一个$m\times 1$的实值函数向量，则称矩阵

$$\frac{\partial \boldsymbol{y}'}{\partial \boldsymbol{x}}=\begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}_{n\times m}$$

为向量$\boldsymbol{y}$对向量$\boldsymbol{x}$的微商。

定义1.17：矩阵对矩阵的微商

假设$\mathbf{X}$为$n \times m$矩阵，$\mathbf{Y}$为$\mathbf{X}$的一个$p \times q$ 的实值函数矩阵，则称矩阵

$$\frac{\partial \mathbf{Y}}{\partial \mathbf{X}} = \frac{\partial \left(\mathrm{Vec}(\mathbf{Y})\right)'}{\partial \left(\mathrm{Vec}(\mathbf{X})\right)}$$

为矩阵$\mathbf{Y}$对矩阵$\mathbf{X}$的微商。

1. 设$\mathbf{X}$为$n \times p$的矩阵，则
$$\frac{\partial \mathbf{X}}{\partial \mathbf{X}} = \mathbf{I}_{np};$$
2. 设$\mathbf{A}$为$n \times n$的可逆矩阵，则
$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} = -\mathbf{A}^{-1} \otimes \left(\mathbf{A}^{-1}\right)';$$
3. 设$\mathbf{A}$，$\mathbf{X}$和$\mathbf{B}$为$m \times n$，$n \times p$和$p \times q$的矩阵，则
$$\frac{\partial \mathbf{AXB}}{\partial \mathbf{X}} = \mathbf{B} \otimes \mathbf{A}'.$$

- 设$\boldsymbol{X} = (X_1, \cdots, X_n)'$是$n \times 1$随机向量，其密度函数为$f(\boldsymbol{x})$。
- 进一步，假设$\boldsymbol{Y} = \boldsymbol{y}(\boldsymbol{X}) = (y_1(\boldsymbol{X}), \cdots, y_n(\boldsymbol{X}))$ 是一一对应的变换，且逆变换$\boldsymbol{X} = \boldsymbol{x}(\boldsymbol{Y})$存在。
- 若偏导数$\dfrac{\partial x_i}{\partial y_j}(i, j = 1, \cdots, n)$存在且连续，则$\boldsymbol{Y} = \boldsymbol{y}(\boldsymbol{X})$的密度函数存在，且可表达为

$$g(\boldsymbol{y}) = f(\boldsymbol{x}(\boldsymbol{Y}))J(\boldsymbol{X} \to \boldsymbol{Y}). \tag{11}$$

- $J(\boldsymbol{X} \to \boldsymbol{Y})$为变换$\boldsymbol{Y} = \boldsymbol{y}(\boldsymbol{X})$的雅可比行列式，其定义为

$$J(\boldsymbol{X} \to \boldsymbol{Y}) = \left| \det\left(\frac{\partial x_i}{\partial y_j} \right) \right|,$$

- 其中式(11)就是两随机向量间变换的雅可比微分表达式，这里$|\det(\cdot)|$表示行列式的绝对值。

- 设$\mathbf{X}=\mathbf{B}\mathbf{Y}$，其中$\mathbf{X}$和$\mathbf{Y}$为$n\times m$的随机矩阵，$\mathbf{B}$为已知的$n\times n$的非奇异阵，则

$$J(\mathbf{X}\to\mathbf{Y})=|\det(\mathbf{B})|^{m}.$$

- 设$\mathbf{X}=\mathbf{B}\mathbf{Y}\mathbf{C}$，其中$\mathbf{X}$和$\mathbf{Y}$为$n\times m$的随机矩阵，$\mathbf{B}$ 和$\mathbf{C}$ 分别为已知的$n\times n$和$m\times m$的非奇异阵，则

$$J(\mathbf{X}\to\mathbf{Y})=|\det(\mathbf{B})|^{m}|\det(\mathbf{C})|^{n}.$$

- 设$\mathbf{X}$为已知的$n\times n$的非奇异阵，则

$$J(\mathbf{X}^{-1}\to\mathbf{X})=|\det(\mathbf{X})|^{-2n}.$$

- 设$\mathbf{X}=\mathbf{BYB}'$，其中$\mathbf{X}$和$\mathbf{Y}$为$m\times m$的随机矩阵，$\mathbf{B}$为已知的$m\times m$非奇异阵，则

$$J(\mathbf{X}\to\mathbf{Y})=\left|\det(\mathbf{B})\right|^{m+1}.$$

- 设$\mathbf{A}=\mathbf{TT}'$，其中$\mathbf{A}$为$m\times m$的随机正定矩阵，$\mathbf{T}$为对角元为正的$m\times m$下三角矩阵，则

$$J(\mathbf{A}\to\mathbf{T})=2^m\prod_{i=1}^{m}t_{ii}^{m+1-i}.$$


STATISTICS
DATA
统计与数据科学丛书 4
STATISTICS AND DATA SCIENCE SERIES 4
多元统计分析
李高荣 吴密霞 编著
Multivariate Statistical Analysis
科学出版社

谢谢，请多提宝贵意见！