第2章　矩阵论的预备知识

吴密霞

北京工业大学 统计与数据科学系

E-mail: *wumixia@bjut.edu.cn*

教材

- 吴密霞, 王松桂. 2024. 线性模型引论(第2版), 科学出版社.

目录

线性空间

$\mathbb{S}$是向量的一个集合，若它对向量加法和数乘两种运算具有封闭性，则$\mathbb{S}$是一个线性空间.

例如

- 全体$n \times 1$实向量组成的集合为$\mathbb{R}^n$
- $\mathbf{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k)$的列向量张成的子空间

$$\mathcal{M}(\mathbf{A}) = \{\boldsymbol{x} = \mathbf{A}\boldsymbol{t},\ \boldsymbol{t} \in \mathbb{R}^k\},$$

这里$\boldsymbol{t} = (t_1, t_2, \cdots, t_k)'$, $\mathbf{A}\boldsymbol{t} = \sum_{i=1}^k \boldsymbol{a}_i t_i$. 注： $\dim\mathcal{M}(\mathbf{A}) = \mathrm{rk}(\mathbf{A})$

线性相关

若存在不全为零的实数$\alpha_1, \alpha_2, \cdots, \alpha_k$, 使得

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = \mathbf{0},$$

则称向量组$\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$是线性相关的, 否则称它们是线性无关的.

注 $\mathcal{M}(\mathbf{A}) \subset \mathcal{M}(\mathbf{A}, \mathbf{B})$, 特别地若$\mathbf{B}$的列向量$\boldsymbol{b}_j, j = 1, 2, \cdots, m$ 皆可表为$\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$的线性组合, 则

$$\mathcal{M}(\mathbf{A}) = \mathcal{M}(\mathbf{A}, \mathbf{B})$$

线性空间上的几种运算：

- 向量$\boldsymbol{a}=(a_1,a_2,\cdots,a_n)'$, $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)'$ 的内积:
  $(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{a}'\boldsymbol{b}=\sum_{i=1}^{n}a_ib_i$
- $\boldsymbol{a}$与$\boldsymbol{b}$正交: $(\boldsymbol{a},\boldsymbol{b})=0$, 记为$\boldsymbol{a}\perp\boldsymbol{b}$.
- $\boldsymbol{a}$正交于$\mathbb{S}$: a与子空间S中的每一个向量正交, 记为$\boldsymbol{a}\perp\mathbb{S}$.
- 向量$\boldsymbol{a}$的长度: $\|\boldsymbol{a}\|=(\boldsymbol{a}'\boldsymbol{a})^{1/2}=(\sum_{i=1}^{n}a_i^2)^{1/2}$
- 子空间$\mathbb{S}$的正交补空间: $\mathbb{S}^{\perp}=\{\boldsymbol{x}:\ \boldsymbol{x}\perp\mathbb{S}\}$
- 设$\mathbf{A}$为$n\times k$矩阵, 记$\mathbf{A}^{\perp}$为满足条件$\mathbf{A}'\mathbf{A}^{\perp}=\mathbf{0}$且具有最大秩的矩阵, 则$\mathcal{M}(\mathbf{A}^{\perp})=\mathcal{M}(\mathbf{A})^{\perp}$.

定理2.1.1

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

定理2.1.2

设$\mathbf{A}_{n\times m}$, $\mathbf{H}_{k\times m}$, 则

(1) $\mathbb{S} = \{\mathbf{A}\boldsymbol{x} : \mathbf{H}\boldsymbol{x} = \mathbf{0}\}$是$\mathcal{M}(\mathbf{A})$的子空间,

(2) $\dim(\mathbb{S}) = \mathrm{rk}\begin{pmatrix} \mathbf{A} \\ \mathbf{H} \end{pmatrix} - \mathrm{rk}(\mathbf{H})$.

推论2.1.1

设$\mathcal{M}(\mathbf{A}) \cap \mathcal{M}(\mathbf{B}) = \{\mathbf{0}\}$, 则$\mathcal{M}(\mathbf{A}'\mathbf{B}^{\perp}) = \mathcal{M}(\mathbf{A}')$.

定理2.2.1 (矩阵谱分解)

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

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

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

实对称矩阵具有许多优良性质, 如

- 实对称矩阵的特征值都是实数; 特征向量都是实向量.
- 不同特征值对应的特征向量彼此正交.

推论2.2.1

设$\mathbf{A}$ 是一个$n \times n$为对称矩阵, 则

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

这里$\mathrm{tr}(\cdot)$表示方阵的迹(trace), 即主对角线上的元素和, $\det(\cdot)$表示方阵的行列式.

定理2.2.2 (奇异值分解)

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

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

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$\mathbf{A}'\mathbf{A}$的非零特征根, $\mathbf{P}_1$和$\mathbf{Q}_1$的列向量分别为$\mathbf{A}\mathbf{A}'$和$\mathbf{A}'\mathbf{A}$ 对应于r个非零特征根的标准正交化的特征向量.

广义逆矩阵

广义逆矩阵的研究可以追溯到Moore(1935) 的著名论文, 对任意一个矩阵$\mathbf{A}$, Moore用如下四个条件:

$$\begin{aligned}
\mathbf{AXA} &= \mathbf{A}, \\
\mathbf{XAX} &= \mathbf{X}, \\
(\mathbf{AX})' &= \mathbf{AX}, \\
(\mathbf{XA})' &= \mathbf{XA},
\end{aligned}$$

定义了$\mathbf{A}$的广义逆$\mathbf{X}$.

- 广义逆A^{-}: 满足上述第一个条件$\mathbf{AXA} = \mathbf{A}$ 的广义逆.
- 广义逆A^{+}: 满足上述四个条件, 称作Moore-Penrose广义逆.

广义逆矩阵A^-存在性和构造性问题

定理2.3.1

对任何矩阵$\mathbf{A}$, $\mathbf{A}^-$总是存在的. 设$\mathbf{A}$为$m \times n$矩阵, $\mathrm{rk}(\mathbf{A}) = r$. 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$m \times m$, $n \times n$的可逆阵, 则

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

这里$\mathbf{B}$, $\mathbf{C}$和$\mathbf{D}$为适当阶数的任意矩阵.

广义逆矩阵$\mathbf{A}^+$的唯一性和构造性问题

定理2.3.2

对任何矩阵$\mathbf{A}$, $\mathbf{A}^+$唯一. 设$\mathbf{A}$为$m \times n$矩阵, $\mathrm{rk}(\mathbf{A}) = r$, 若$\mathbf{A}$的奇异值分解为

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

则

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

其中$\mathbf{\Lambda}_r = \mathrm{diag}(\lambda_1, \cdots, \lambda_r)$, $\lambda_i > 0$, $i = 1, 2, \cdots, r$, $\lambda_1^2, \cdots, \lambda_r^2$为$\mathbf{A}'\mathbf{A}$的非零特征根.

广义逆的性质

- $\mathbf{A}^-$ 唯一 $\Longleftrightarrow$ $\mathbf{A}$ 为可逆方阵. 此时$\mathbf{A}^- = \mathbf{A}^{-1}$;
- $\mathrm{rk}(\mathbf{A}^-) \geq \mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}^-\mathbf{A}) = \mathrm{rk}(\mathbf{A}\mathbf{A}^-)$;
- 若$\mathcal{M}(\mathbf{B}) \subset \mathcal{M}(\mathbf{A}), \mathcal{M}(\mathbf{C}) \subset \mathcal{M}(\mathbf{A}')$, 则$\mathbf{C}'\mathbf{A}^-\mathbf{B}$与$\mathbf{A}^-$的选择无关;
- $\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'$与广义逆$(\mathbf{A}'\mathbf{A})^-$的选择无关;
- $\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'\mathbf{A} = \mathbf{A}, \quad \mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}' = \mathbf{A}'$.

- $(\mathbf{A}^+)^+ = \mathbf{A}; \quad (\mathbf{A}^+)' = (\mathbf{A}')^+; \quad \mathrm{rk}(\mathbf{A}^+) = \mathrm{rk}(\mathbf{A})$;
- $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^+\mathbf{A}' = \mathbf{A}'(\mathbf{A}\mathbf{A}')^+$, 特别地, $\boldsymbol{a}^+ = \boldsymbol{a}'/\|\boldsymbol{a}\|^2$;
- $\mathbf{I} \geq \mathbf{A}^+\mathbf{A}; \quad (\mathbf{A}'\mathbf{A})^+ = \mathbf{A}^+(\mathbf{A}')^+$.

对于相容线性方程组

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1.1}$$

这里$\mathbf{A}$是$m \times n$矩阵, 其秩$\mathrm{rk}(\mathbf{A}) = r \leq \min(m, n)$, 则

- 当$r = m = n$时, 方程组(1.1)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, (1.1)有无穷多解,
  - 对任一广义逆$\mathbf{A}^-$, $\boldsymbol{x} = \mathbf{A}^-\boldsymbol{b}$必为解;
  - $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$的通解 为$\boldsymbol{x} = \mathbf{A}^-\boldsymbol{b} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\boldsymbol{z}$, 其中$\mathbf{A}^-$为任一固定的广义逆, $\boldsymbol{z}$为任意向量;
  - 在$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$的解集中, $\boldsymbol{x}_0 = \mathbf{A}^+\boldsymbol{b}$为长度最小者.

分块矩阵的广义逆

- 设$\mathbf{A}$的分块形式:

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

- 若$|\mathbf{A}_{11}|\neq 0$, 则$\mathbf{A}$可被对角块化

$$\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}$$

$$=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_{22.1}\end{pmatrix},$$ (若$\mathbf{A}$可逆，则两边取逆可得$\mathbf{A}$的逆)

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

定理2.3.4（分块矩阵的逆）

设矩阵$\mathbf{A}$可逆.

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

定理2.3.5 (分块矩阵的广义逆)

(1) 若$\mathbf{A}_{11}^{-1}$存在, 则

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-} \\ -\mathbf{A}_{22.1}^{-}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22.1}^{-} \end{pmatrix}.$$

(2) 若$\mathbf{A}$半正定, 即$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \geq \mathbf{0}$, 则

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

其中$\mathbf{A}_{22.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-}\mathbf{A}_{12}$.

幂等阵

定义2.4.1

若方阵$\mathbf{A}_{n\times n}$满足$\mathbf{A}^2=\mathbf{A}$, 则称$\mathbf{A}$为幂等阵(idempotent matrix).

- 幂等阵的特征根只能为0 或1;
- 幂等阵的迹等于秩, 即若$\mathbf{A}^2=\mathbf{A}$, 则$\mathrm{tr}(\mathbf{A})=\mathrm{rk}(\mathbf{A})$;
- 若$\mathbf{A}$为对称幂等阵, 则$\mathbf{A}^+=\mathbf{A}$;
- 对任意的矩阵$\mathbf{A}$, $\mathbf{A}^-\mathbf{A}, \mathbf{A}\mathbf{A}^-, \mathbf{I}-\mathbf{A}^-\mathbf{A}$,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;
- 设$\mathbf{P}_{n\times n}$为对称幂等阵$\mathrm{rk}(\mathbf{P})=r$, 则存在秩为$r$ 的$\mathbf{A}_{n\times r}$, 使

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

正交投影

正交投影

设$\boldsymbol{x} \in \mathbb{R}^n$, $\mathbb{S}$为$\mathbb{R}^n$的一个线性子空间.对$\boldsymbol{x}$作分解

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}, \quad \boldsymbol{y} \in \mathbb{S}, \ \ \boldsymbol{z} \in \mathbb{S}^{\perp},$$

则称$\boldsymbol{y}$ 为$\boldsymbol{x}$在$\mathbb{S}$ 上的正交投影. 若**P**为n阶方阵, 使得对一切$\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y}$ 满足$\boldsymbol{y} = \mathbf{P}\boldsymbol{x}$, 则称**P** 为向$\mathbb{S}$ 的正交投影阵.

定理2.4.5

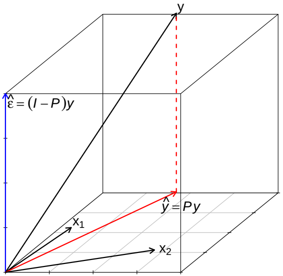
设$\mathbf{A}$为$n \times m$矩阵, $\mathbf{P}_{\mathbf{A}}$为向$\mathcal{M}(\mathbf{A})$的正交投影阵, 则$\mathbf{P}_{\mathbf{A}} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$.

- 正交投影阵$\mathbf{P}_{\mathbf{A}}$唯一. ($\mathbf{P}_{\mathbf{A}} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$与广义逆选择无关)
- **P**为正交投影阵$\Longleftrightarrow$ **P**为对称幂等阵.

定理2.4.7

n阶方阵$\mathbf{P}$为正交投影阵$\Longleftrightarrow$ 对任给$\boldsymbol{y} \in \mathbb{R}^n$,

$$\| \boldsymbol{y} - \mathbf{P}\boldsymbol{y} \| = \inf \| \boldsymbol{y} - \boldsymbol{u} \|, \quad \boldsymbol{u} \in \mathcal{M}(\mathbf{P}). \tag{1.2}$$



记$\lambda_1 \geq \cdots \geq \lambda_n$为$\mathbf{A}$的特征值, $\varphi_1, \cdots, \varphi_n$为对应的标准正交化特征向量.

定理2.5.1 (Rayleigh-Ritz)

设$\mathbf{A}$为$n \times n$对称阵, 则

(1) $\sup\limits_{\boldsymbol{x} \neq \mathbf{0}} \dfrac{\boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \varphi_1'\mathbf{A}\varphi_1 = \lambda_1$; (2) $\inf\limits_{\boldsymbol{x} \neq \mathbf{0}} \dfrac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \varphi_n'\mathbf{A}\varphi_n = \lambda_n$.

- 对任一n阶对称阵$\mathbf{A} = (a_{ij})$, 总有$\lambda_n \leq a_{ii} \leq \lambda_1, i = 1, \cdots, n$.
- $\sup\limits_{\substack{\varphi_i'\boldsymbol{x}=0 \\ i=1,\cdots,k}} \dfrac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \varphi_{k+1}'\mathbf{A}\varphi_{k+1} = \lambda_{k+1}$;
- $\inf\limits_{\substack{\varphi_i'\boldsymbol{x}=0 \\ i=k+1,\cdots,n}} \dfrac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \varphi_k'\mathbf{A}\varphi_k = \lambda_k$.

定理2.5.2 (Courant-Fischer 极大极小定理)

设$\mathbf{A}$为$n \times n$对称阵, $\mathbf{B}$为$n \times k$ 矩阵, 则

$$(1)\quad \inf_{\mathbf{B}} \sup_{\mathbf{B}'\boldsymbol{x}=\mathbf{0}} \frac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \sup_{\boldsymbol{\Phi}'_{(k)}\boldsymbol{x}=\mathbf{0}} \frac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \varphi'_{k+1}\mathbf{A}\varphi_{k+1} = \lambda_{k+1};$$

$$(2)\quad \sup_{\mathbf{B}} \inf_{\mathbf{B}'\boldsymbol{x}=\mathbf{0}} \frac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \inf_{\boldsymbol{\Phi}'_{(k)}\boldsymbol{x}=\mathbf{0}} \frac{(\boldsymbol{x}'\mathbf{A}\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{x}} = \varphi'_{n-k}\mathbf{A}\varphi_{n-k} = \lambda_{n-k},$$

其中$\boldsymbol{\Phi}_k$和$\boldsymbol{\Phi}_{(k)}$分别表示$\boldsymbol{\Phi} = (\varphi_1, \cdots, \varphi_n)$的前$k$列和后$k$列.

定理2.5.3 (Sturm分离定理)

设$\mathbf{A}$为$n \times n$对称阵, 记

$$\mathbf{A}_r = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}, \qquad r = 1, \cdots, n$$

为$\mathbf{A}$的顺序主子式, 则

$$\lambda_{i+1}(\mathbf{A}_{r+1}) \leq \lambda_i(\mathbf{A}_r) \leq \lambda_i(\mathbf{A}_{r+1}), \quad i = 1, 2, \cdots, r.$$

定理2.5.4 (Weyl定理)

设$\mathbf{A}$和$\mathbf{B}$皆为$n \times n$的对称阵, 则

$$\lambda_i(\mathbf{A}) + \lambda_n(\mathbf{B}) \leq \lambda_i(\mathbf{A}+\mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_1(\mathbf{B}), \quad i = 1, \cdots, n.$$

定理2.5.4（Weyl定理）

设A和B皆为$n \times n$的对称阵, 则

$$\lambda_i(\mathbf{A}) + \lambda_n(\mathbf{B}) \le \lambda_i(\mathbf{A}+\mathbf{B}) \le \lambda_i(\mathbf{A}) + \lambda_1(\mathbf{B}), \quad i = 1, \cdots, n.$$

证明 设$\boldsymbol{x}'\boldsymbol{x} = 1$. 令$\boldsymbol{\Phi}_{(i-1)} = (\phi_1, \cdots, \phi_{i-1})$和$\boldsymbol{\Psi}_{(i-1)} = (\psi_1, \cdots, \psi_{i-1})$分别由A和$(\mathbf{A}+\mathbf{B})$的前$i-1$个特征值对应的标准化正交化的特征向量，C是$n \times (i-1)$的矩阵, 则根据定理2.5.1和定理2.5.2有

$$\lambda_i(\mathbf{A})+\lambda_1(\mathbf{B}) \ge \sup_{\boldsymbol{\Phi}'_{(i-1)}\boldsymbol{x}=\mathbf{0}} \boldsymbol{x}'(\mathbf{A}+\mathbf{B})\boldsymbol{x} \ge \inf_{\mathbf{C}} \sup_{\mathbf{C}'\boldsymbol{x}=\mathbf{0}} \boldsymbol{x}'(\mathbf{A}+\mathbf{B})\boldsymbol{x} = \lambda_i(\mathbf{A}+\mathbf{B})$$

$$= \sup_{\boldsymbol{\Psi}'_{(i-1)}\boldsymbol{x}=\mathbf{0}} \boldsymbol{x}'(\mathbf{A}+\mathbf{B})\boldsymbol{x} \ge \sup_{\boldsymbol{\Psi}'_{(i-1)}\boldsymbol{x}=\mathbf{0}} \boldsymbol{x}'\mathbf{A}\boldsymbol{x} + \lambda_n(\mathbf{B}) \ge \lambda_i(\mathbf{A}) + \lambda_n(\mathbf{B}).$$

定理2.5.5 (Poincaré 分离定理)

设$\mathbf{A}_{n\times n}$为对称阵, $\mathbf{P}$为 $n\times k$ 的列正交阵, 即 $\mathbf{P}'\mathbf{P}=\mathbf{I}_k$, 则

$$\lambda_{n-k+i}(\mathbf{A}) \leq \lambda_i(\mathbf{P}'\mathbf{A}\mathbf{P}) \leq \lambda_i(\mathbf{A}), \quad i=1,\cdots,k.$$

定理2.5.5 (Poincaré 分离定理)

设$\mathbf{A}_{n\times n}$为对称阵, P为$n\times k$的列正交阵, 即$\mathsf{P}'\mathsf{P}=\mathbf{I}_k$, 则

$$\lambda_{n-k+i}(\mathbf{A}) \le \lambda_i(\mathsf{P}'\mathbf{A}\mathsf{P}) \le \lambda_i(\mathbf{A}), \quad i=1,\cdots,k.$$

证明 将 P 扩充为正交方阵 $\tilde{\mathsf{P}}=(\mathsf{P},\mathbf{Q})$, 记

$$\mathbf{H}_n=\tilde{\mathsf{P}}'\mathbf{A}\tilde{\mathsf{P}}=\begin{pmatrix}\mathsf{P}'\mathbf{A}\mathsf{P} & \mathsf{P}'\mathbf{A}\mathbf{Q}\\ \mathbf{Q}'\mathbf{A}\mathsf{P} & \mathbf{Q}'\mathbf{A}\mathbf{Q}\end{pmatrix},$$

$\mathbf{H}_k=\mathsf{P}'\mathbf{A}\mathsf{P}$为 $\mathbf{H}_n$ 的 k 阶顺序主子阵. 由Sturm定理得

$$\lambda_i(\mathbf{A})=\lambda_i(\tilde{\mathsf{P}}'\mathbf{A}\tilde{\mathsf{P}}) \ge \lambda_i(\mathsf{P}'\mathbf{A}\mathsf{P})=\lambda_i(\mathbf{H}_k) \ge \lambda_{i+1}(\mathbf{H}_{k+1}) \ge$$

$$\cdots \ge \lambda_{i+(n-k)}(\mathbf{H}_n)=\lambda_{i+n-k}(\tilde{\mathsf{P}}'\mathbf{A}\tilde{\mathsf{P}})=\lambda_{n-k+i}(\mathbf{A}).$$

定理2.5.6 Cauchy-Schwarz不等式

设$\boldsymbol{x}$和$\boldsymbol{y}$ 为任意两个$n \times 1$的向量，则

$$(\boldsymbol{x}'\boldsymbol{y})^2 \leq \boldsymbol{x}'\boldsymbol{x} \cdot \boldsymbol{y}'\boldsymbol{y}, \tag{1.3}$$

等号成立$\Longleftrightarrow$ $\boldsymbol{x}$和$\boldsymbol{y}$线性相关.

定理2.5.6 Cauchy-Schwarz不等式

设$\boldsymbol{x}$和$\boldsymbol{y}$ 为任意两个$n \times 1$的向量，则

$$(\boldsymbol{x}'\boldsymbol{y})^2 \leq \boldsymbol{x}'\boldsymbol{x} \cdot \boldsymbol{y}'\boldsymbol{y}, \tag{1.4}$$

等号成立$\Longleftrightarrow$ $\boldsymbol{x}$和$\boldsymbol{y}$线性相关.

证明 当$\boldsymbol{x}$和$\boldsymbol{y}$至少有一个为零时，结论显然成立. 不妨假设$\boldsymbol{x} \neq \boldsymbol{0}$, 定义

$$\boldsymbol{z} = \boldsymbol{y} - \frac{\boldsymbol{x}'\boldsymbol{y}}{\|\boldsymbol{x}\|^2}\boldsymbol{x},$$

则$\boldsymbol{x}'\boldsymbol{z} = 0$. 于是

$$0 \leq \|\boldsymbol{z}\|^2 = \boldsymbol{z}'\boldsymbol{y} = \|\boldsymbol{y}\|^2 - \frac{\boldsymbol{x}'\boldsymbol{y}}{\|\boldsymbol{x}\|^2}\boldsymbol{x}'\boldsymbol{y} = \|\boldsymbol{y}\|^2 - \frac{(\boldsymbol{x}'\boldsymbol{y})^2}{\|\boldsymbol{x}\|^2},$$

得证$(\boldsymbol{x}'\boldsymbol{y})^2 \leq \boldsymbol{x}'\boldsymbol{x} \cdot \boldsymbol{y}'\boldsymbol{y}$, 等号成立$\Longleftrightarrow \boldsymbol{z} = \boldsymbol{0} \Longleftrightarrow \boldsymbol{x}$和$\boldsymbol{y}$ 成比例.

Cauchy-Schwarz不等式的两种推广形式：

- 设$\mathbf{A}$为$n \times n$半正定对称阵,

$$| \boldsymbol{x}'\mathbf{A}\boldsymbol{y} |^2 \leq (\boldsymbol{x}'\mathbf{A}\boldsymbol{x}) \cdot (\boldsymbol{y}'\mathbf{A}\boldsymbol{y}),$$

  等号成立$\Longleftrightarrow$ $\boldsymbol{x}$和$\boldsymbol{y}$线性相关.

- 设$\mathbf{A}$为$n \times n$正定对称阵,

$$| \boldsymbol{x}'\boldsymbol{y} |^2 \leq (\boldsymbol{x}'\mathbf{A}\boldsymbol{x}) \cdot (\boldsymbol{y}'\mathbf{A}^{-1}\boldsymbol{y}),$$

  等号成立$\Longleftrightarrow$ $\boldsymbol{x}$和$\mathbf{A}^{-1}\boldsymbol{y}$线性相关.

矩阵偏序

设 $\mathbf{A}, \mathbf{B}$为两个 n 阶对称阵.

若 $\mathbf{B}-\mathbf{A} \geq 0$, 即 $\mathbf{B}-\mathbf{A}$ 为半正定阵, 则称 $\mathbf{A}$ 低于$\mathbf{B}$, 记为 $\mathbf{B} \geq \mathbf{A}$ 或 $\mathbf{A} \leq \mathbf{B}$. 类似地, $\mathbf{A} > \mathbf{B}$ 表明 $\mathbf{A}-\mathbf{B}$ 为正定阵.

易证对称阵的这种关系满足下列性质:

- 自反性: $\mathbf{A} \geq \mathbf{A}$;
- 传递性: 若$\mathbf{A} \geq \mathbf{B}, \mathbf{B} \geq \mathbf{C}$, 则$\mathbf{A} \geq \mathbf{C}$;
- 若$\mathbf{A} \geq \mathbf{B}, \mathbf{B} \geq \mathbf{A}$, 则$\mathbf{A} = \mathbf{B}$,

  这种关系被称为Lowner偏序. 因为并非任意两个对称阵都有这种关系, 所以称其为偏序.

定理2.6.1 (单调性)

设$\mathbf{A},\mathbf{B}$为两个n阶对称阵.

(1) 若$\mathbf{A}\geq\mathbf{B}$, 则$\lambda_i(\mathbf{A})\geq\lambda_i(\mathbf{B})$, $i=1,\cdots,n$;

(2) 若$\mathbf{A}>\mathbf{B}$, 则$\lambda_i(\mathbf{A})>\lambda_i(\mathbf{B})$, $i=1,\cdots,n$.

证明 由Weyl定理得

$$\lambda_i(\mathbf{A})=\lambda_n(\mathbf{B}+(\mathbf{A}-\mathbf{B}))\geq\lambda_i(\mathbf{B})+\lambda_n(\mathbf{A}-\mathbf{B})\geq\lambda_i(\mathbf{B}).$$

若$\mathbf{A}\geq\mathbf{B}$, 即$\mathbf{A}-\mathbf{B}\geq\mathbf{0}$, 故$\lambda_n(\mathbf{A}-\mathbf{B})\geq0$, (1) 得证;

若$\mathbf{A}>\mathbf{B}$, 则$\lambda_n(\mathbf{A}-\mathbf{B})>0$, 故$\lambda_n(\mathbf{A}-\mathbf{B})>0$, (2) 得证.

矩阵偏序

推论2.6.1

设$\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$, 则

(1) $\mathrm{tr}(\mathbf{A}) \geq \mathrm{tr}(\mathbf{B})$; (2) $|\mathbf{A}| \geq |\mathbf{B}|$; (3) $\mathrm{rk}(\mathbf{A}) \geq \mathrm{rk}(\mathbf{B})$.

定理2.6.2

设$\mathbf{A}$和$\mathbf{B}$为两个n阶对称阵, $\mathbf{P}$为$n \times k$矩阵.

(1) 若$\mathbf{A} \geq \mathbf{B}$, 则$\mathbf{P}'\mathbf{A}\mathbf{P} \geq \mathbf{P}'\mathbf{B}\mathbf{P}$;

(2) 若$\mathrm{rk}(\mathbf{P}) = k$, $\mathbf{A} > \mathbf{B}$, 则$\mathbf{P}'\mathbf{A}\mathbf{P} > \mathbf{P}'\mathbf{B}\mathbf{P}$.

定理2.6.3

设$\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$, 则$\mathcal{M}(\mathbf{B}) \subset \mathcal{M}(\mathbf{A})$.

定理2.6.4

设$\mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}$, 则下面的命题等价.

(1) $\mathbf{A} \geq \mathbf{B}$;

(2) $\mathcal{M}(\mathbf{B}) \subseteq \mathcal{M}(\mathbf{A})$, 对任意的$\boldsymbol{x} \in \mathcal{M}(\mathbf{A}), \boldsymbol{x}'(\mathbf{A}-\mathbf{B})\boldsymbol{x} \geq 0$;

(3) $\mathcal{M}(\mathbf{B}) \subseteq \mathcal{M}(\mathbf{A}), \lambda_1(\mathbf{B}\mathbf{A}^-) \leq 1$,这里$\lambda_1(\mathbf{B}\mathbf{A}^-)$与$\mathbf{A}^-$的选择无关.

推理2.6.2

设$\mathbf{A} \geq \mathbf{0}$, $\mathbf{B} \geq \mathbf{0}$, 则

(1) 若$\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{B})$, 则$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{B}^+ \geq \mathbf{A}^+$;

(2) 若$\mathbf{B} > \mathbf{0}$, 则$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{B}^{-1} \geq \mathbf{A}^{-1}$; $\mathbf{A} > \mathbf{B} \Leftrightarrow \mathbf{B}^{-1} > \mathbf{A}^{-1}$.

引理2.6.1

设$\mathbf{A}$为$n\times n$实方阵, $\lambda_1(\mathbf{A}), \sigma_1(\mathbf{A})$分别为它的最大特征根和最大奇异值, 则$|\lambda_1(\mathbf{A})| \leq \sigma_1(\mathbf{A})$.

证明 设$\boldsymbol{x}$为$\mathbf{A}$的对应于$\lambda_1(\mathbf{A})$的单位特征向量. 于是

$$(\lambda_1(\mathbf{A}))^2 = (\mathbf{A}\boldsymbol{x})'\mathbf{A}\boldsymbol{x} = \boldsymbol{x}'\mathbf{A}'\mathbf{A}\boldsymbol{x} \leq \lambda_1(\mathbf{A}'\mathbf{A}) = \sigma_1^2(\mathbf{A}).$$

定理2.6.5

设$\mathbf{A}, \mathbf{B}$为两个半正定阵, 则

(1) $\mathbf{A}^2 \geq \mathbf{B}^2 \Longrightarrow \mathbf{A} \geq \mathbf{B}$;

(2) 若$\mathbf{AB} = \mathbf{BA}$, 则$\mathbf{A} \geq \mathbf{B} \Longrightarrow \mathbf{A}^k \geq \mathbf{B}^k \geq \mathbf{0}$, k为任意正整数.

定义2.6.1

设$\mathbf{A} = (a_{ij})$和$\mathbf{B} = (b_{ij})$分别为$m \times n, p \times q$的矩阵, 称

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

为$\mathbf{A}$和$\mathbf{B}$的Kronecker乘积.

这种乘积具有下列性质:

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

$(\mathbf{A} \otimes \mathbf{B})^{-} = \mathbf{A}^{-} \otimes \mathbf{B}^{-}$; $(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2)$.

矩阵的向量化

定义2.6.2 (矩阵的向量化)

矩阵的向量化就是把矩阵按列向量依次排成的向量. 设$\mathbf{A}_{m\times n} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$, 则

$$\mathrm{Vec}(\mathbf{A}) = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

向量化运算具有下列性质:

- $\mathrm{Vec}(\mathbf{A}+\mathbf{B}) = \mathrm{Vec}(\mathbf{A}) + \mathrm{Vec}(\mathbf{B})$; $\mathrm{Vec}(\alpha\mathbf{A}) = \alpha\mathrm{Vec}(\mathbf{A})$, ($\alpha$为数);
- $\mathrm{tr}(\mathbf{AB}) = (\mathrm{Vec}(\mathbf{A}'))'\mathrm{Vec}(\mathbf{B})$; $\mathrm{Vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{Vec}(\mathbf{B})$

假设 $\mathbf{X}$ 为 $n \times m$ 矩阵, $y = f(\mathbf{X})$ 为 $\mathbf{X}$ 的一个实值函数, 矩阵

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

称为 y 对 $\mathbf{X}$ 的微商.

- 设$\boldsymbol{a}, \boldsymbol{x}$均为$n \times 1$向量, $\mathbf{A}_{n \times n}$对称阵, 则

$$\frac{\partial \boldsymbol{a}'\boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{a}, \quad \frac{\partial \boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = 2\mathbf{A}\boldsymbol{x}.$$