《回归分析》作业参考答案 第一章

sword nalzok

2017 级统计系 2016 级统计系

2019年9月23日

1 第一节

- 1. 矩阵 A 为 n 阶实对称矩阵,证明:
 - (iii) 矩阵 A 有 n 个实特征值(记作 $\lambda_1, \lambda_2, \dots, \lambda_n$),而且 n 个特征向量可正交化;
 - (iv) 存在正交阵 P 使得

$$A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P' \triangleq P \Lambda P'$$

(v) 对任意正整数 s,有

$$A^{s} = P \begin{pmatrix} \lambda_{1}^{s} & & \\ & \ddots & \\ & & \lambda_{n}^{s} \end{pmatrix} P' \triangleq P \Lambda^{s} P'$$

从而 $\operatorname{tr}(A^s) = \sum_{i=1}^n \lambda_i^s$;

- (vi) A 是非奇异的当且仅当 $\lambda_i \neq 0$ ($\forall i$),且矩阵 A^{-1} 的特征值为 λ_i^{-1} , $i=1,2,\cdots,n$;
- (vii) 矩阵 $I_n + cA$ 的特征值为 $1 + c\lambda_i$, $i = 1, 2, \dots, n$;
- (viii) 半正定矩阵的特征值均非负,从而半正定矩阵的迹非负。
- **证明.** (iii) 由代数学基本定理,知矩阵 A 的特征多项式在复数域中有且仅有 n 个根。又由 (i) 得 A 的复特征值均为实数,故矩阵 A 有 n 个实特征值。下面,借用 (iv) 的结论,矩阵 A 可相似对角化,则 A 各个特征值的几何重数等于代数重数,所以 A 有 n 个线性无关的特征向量,从而 n 个特征向量可正交化。
- (iv) 对实对称矩阵 A 的阶数 n 作数学归纳法。 n=1 时,A 本身即为对角阵,再取 P 为一阶单位阵,结论成立。 假设任一 n-1 阶实对称矩阵都能正交相似于一对角阵,下面来看 n 阶的情形。

设 λ_1 为 A 的一个特征值, η_1 为对应的单位特征向量,则 η_1 可扩成一个标准正交基 η_1,\cdots,η_n ,并记矩阵 $T=(\eta_1,\cdots,\eta_n)$ 。从而

$$AT = T \begin{pmatrix} \lambda & \alpha \\ 0 & B \end{pmatrix}$$

又因为

$$\begin{pmatrix} \lambda & \alpha \\ 0 & B \end{pmatrix}' = (T'AT)' = T'AT = \begin{pmatrix} \lambda & \alpha \\ 0 & B \end{pmatrix}$$

所以 $\alpha=0$,矩阵 B 为实对称矩阵。由归纳假设,存在 n-1 阶正交阵 Q,使得 $B=Q'\Lambda_1Q$,其中 Λ_1 为对角阵。令 $Q_1={\rm diag}(1\,,Q)$,就有

$$T'AT = Q' \begin{pmatrix} \lambda & 0 \\ 0 & \Lambda_1 \end{pmatrix} Q$$

再取 $P=Q_1T'$,即可证得 n 阶实对称矩阵亦正交相似于一对角矩阵。

根据数学归纳法,结论成立。

- (v) 注意到 PP'=I,从而 $A^s=(P\Lambda P')^s=P\Lambda P'P\Lambda P'\cdots P\Lambda P'=P\Lambda^s P'$ 。又因为相似矩阵有相同的迹,故 $\operatorname{tr}(A^s)=\operatorname{tr}(\Lambda^s)=\sum_{i=1}^n \lambda_i^s$ 。
- (vi) 矩阵 A 非奇异 \Leftrightarrow 矩阵 Λ 非奇异 $\Leftrightarrow \lambda_i \neq 0 \ (\forall i)$ 。 又 $A^{-1} = (P\Lambda P')^{-1} = P\Lambda^{-1}P'$ 及相似矩阵有相同的特征值,得矩阵 A^{-1} 的特征值为 λ_i^{-1} , $i=1,2,\cdots,n$ 。
- (vii) 注意到 $I_n+cA=PP'+P(c\Lambda)P'=P(I_n+c\Lambda)P'$,与 (vi) 类似即可得证。
- (viii) 任取非零向量 x,记向量 y=Px,则由 P 满秩知 y 亦为非零向量。 从而

$$x'\Lambda x = x'P'APx = y'Ay > 0$$

将 x 取遍标准单位向量即可证得结论。

2. 设 A 为实对称矩阵,用拉格朗日乘数法证明:

$$\max_{x'x\neq 0}\frac{x'Ax}{x'x}=\lambda_{\max}(A)\quad \ \ \, \mathop{\Sigma}\limits_{x'x\neq 0}\frac{x'Ax}{x'x}=\lambda_{\min}(A)$$

证明. 不失一般性,考虑 x 为单位向量的情形即可。取单位向量 $e=(x_1,\cdots,x_n)'$,则 e 的取值范围为单位球面(从而是有界闭集)。易见 e'Ae 为 e 的连续函数,而连续函数在有界闭集上必取得最大、最小值,从而结论中的最值运算是有意义的。

由 A 实对称,知存在正交阵 P 与对角阵 Λ ,使得 $A = P'\Lambda P$ 。从而

$$e'Ae = e'P'\Lambda Pe \triangleq u'\Lambda u$$

由于正交变换保持距离且为可逆变换,所以向量 y 仍为单位向量且取值范围为单位球面。

记
$$y = (y_1, \dots, y_n)'$$
, $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, 则

$$y'\Lambda y = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2$$

令 $f(y) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2$,于是问题转化为求解

$$\max_{y} f(y)$$
 s.t. $\sum_{i=1}^{n} y_i^2 = 1$

$$\min_{y} f(y)$$
 s.t. $\sum_{i=1}^{n} y_i^2 = 1$

应用拉格朗日乘数法,构造拉格朗日函数

$$L(y, \mu) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 - \mu (y_1^2 + \dots + y_n^2 - 1)$$

上式两端对各分量求一阶偏导,得

$$\frac{\partial L}{\partial y_i} = 2\lambda_i y_i - 2\mu y_i = 0 \qquad i = 1, 2, \dots, n$$

$$\frac{\partial L}{\partial \mu} = y_1^2 + \dots + y_n^2 - 1 = 0$$

解得 $\mu=\lambda_i\,,y=\epsilon_i\,,f(\epsilon_i)=\lambda_i$,其中 ϵ_i 为第 i 个分量为 1,其余分量为 0 的单位向量。由此,遍历 A 的特征值,知结论成立。

- 3. 设向量 $\beta = (\beta_1, \beta_2, \dots, \beta_n)'$, 证明:
 - (i) 对任意 n 维向量 a,有

$$\frac{\partial \beta' a}{\partial \beta} = a$$

对任意 n 列矩阵 A,有

$$\frac{\partial A\beta}{\partial \beta'} = A$$

(ii) 若 A 为对称矩阵,则

$$\frac{\partial \beta' A \beta}{\partial \beta} = 2A\beta$$

证明. (i) 设 $a=(a_1\,,a_2\,,\cdots\,,a_n)'$,则 $\beta'a=a_1\beta_1+a_2\beta_2+\cdots+a_n\beta_n$ 。 于是

$$\frac{\partial \beta' a}{\partial \beta} = \left(\frac{\partial}{\partial \beta_1}, \frac{\partial}{\partial \beta_2}, \cdots, \frac{\partial}{\partial \beta_n}\right)' (a_1 \beta_1 + a_2 \beta_2 + \cdots + a_n \beta_n) = (a_1, a_2, \cdots, a_n)' = a$$

又设 $A=(b_{ij})_n$,则

$$\frac{\partial A\beta}{\partial \beta'} = \frac{\partial}{\partial \beta'} \begin{pmatrix} b_{11}\beta_1 + \dots + b_{1n}\beta_n \\ b_{21}\beta_1 + \dots + b_{2n}\beta_n \\ \vdots \\ b_{n1}\beta_1 + \dots + b_{nn}\beta_n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial}{\partial \beta_1} \\ \frac{\partial}{\partial \beta_2} \\ \vdots \\ \frac{\partial}{\partial \beta_n} \end{pmatrix} (b_{11}\beta_1 + \dots + b_{1n}\beta_n, \dots, b_{n1}\beta_1 + \dots + b_{nn}\beta_n)$$

$$= A$$

(ii) 对任意的 n 阶方阵 $A=(b_{ij})_n$,有

$$\frac{\partial \beta' A \beta}{\partial \beta} = \frac{\partial}{\partial \beta} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \beta_{i} \beta_{j} \right)$$

$$= \left(\frac{\partial}{\partial \beta_{1}}, \frac{\partial}{\partial \beta_{2}}, \cdots, \frac{\partial}{\partial \beta_{n}} \right)' \left(\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \beta_{i} \beta_{j} \right)$$

$$= (A + A') \beta$$

特别地,当 A 为对称矩阵时,有

$$\frac{\partial \beta' A \beta}{\partial \beta} = 2A\beta$$

4. 设四分块矩阵

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

满足 $A 与 D - CA^{-1}B$ 均可逆。给出上述四分块矩阵的逆矩阵。

解. 记矩阵 $M = D - CA^{-1}B$,则

$$\begin{pmatrix} A^{-1} & \\ & M^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -M^{-1}B \\ & \mathbf{I} \end{pmatrix} \begin{pmatrix} I & \\ -A^{-1}C & \mathbf{I} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathbf{I}$$

从而

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} \\ M^{-1} \end{pmatrix} \begin{pmatrix} I & -M^{-1}B \\ I \end{pmatrix} \begin{pmatrix} I \\ -A^{-1}C & I \end{pmatrix}$$
$$= \begin{pmatrix} A^{-1} + A^{-1}M^{-1}BA^{-1}C & -A^{-1}M^{-1}B \\ -M^{-1}A^{-1}C & M^{-1} \end{pmatrix}$$

- 5. (i) 举例说明矩阵 A 的广义逆 A^- 不唯一;
 - (ii) 给出计算矩阵 A 广义逆 A^- 的一般公式。
 - 解. (i) 先回顾矩阵广义逆的定义。

定义. 设 A 是一个 $s \times n$ 矩阵,矩阵方程 AXA = A 的通解称为 A 的广义逆矩阵,简称为 A 的广义逆,记作 A^- 。

广义逆矩阵可以不唯一。事实上,对任意非单位阵的幂等阵 A,有

$$AAA = A$$
$$AIA = A$$

成立。从而 A 的广义逆可以是 A 和 I,不唯一。

(ii) 记矩阵 A 的秩为 r, 对 A 作满秩分解

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

则矩阵 A 的广义逆的一般形式为

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

其中 B,C,D 分别是任意 $r \times (s-r),(n-r) \times r,(n-r) \times (s-r)$ 阶矩阵。证明请参阅丘维声著《高等代数》第一版上册 251-252 页。

- 6. 对任意 $n \times p$ 阶矩阵 X, 证明:
 - (i) 无论 $(X'X)^-$ 如何变化, $X(X'X)^-X'$ 保持不变;
 - (ii) $X(X'X)^-X'$ 是一个从 \mathbb{R}^n 到 $\mathcal{M}(X)$ 的投影矩阵。

证明. (i) 记矩阵 X 的秩为 r。由线性代数的知识,知将 X 作一系列初等列变换,可得到一个列阶梯矩阵,即

$$X = AQ$$

其中矩阵 A 为列阶梯矩阵(前 r 列为标准单位向量,后 p-r 列为 0),Q 为可逆矩阵。再适当交换矩阵 A 的行,得

$$PA = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \triangleq \Lambda$$

其中,P 为若干行交换矩阵相乘得到的矩阵(因而是正交阵)。于是,我们得到了矩阵 X 的一个满秩分解 $X=P'\Lambda Q$ 。从而 $X'X=Q'\Lambda Q$ 。由第一章习题一 5(2) 的结论, $(X'X)^-$ 的一般形式为

$$(X'X)^{-} = Q^{-1} \begin{pmatrix} I_r & B \\ C & D \end{pmatrix} (Q')^{-1}$$

其中 B,C,D 分别是任意 $r \times (p-r), (n-r) \times r, (n-r) \times (p-r)$ 阶矩阵。故

$$X(X'X)^{-}X' = P'\Lambda P$$

而与 B, C, D 无关,即 $X(X'X)^-X'$ 与 $(X'X)^-$ 的选取无关。

(ii) 由 (1) 知 $X(X'X)^-X' = P'\Lambda P$,容易验证 $X(X'X)^-X'$ 是一个对称 幂等阵,且对 $\forall \alpha \in \mathbb{R}^n$, $X(X'X)^-X'\alpha = X((X'X)^-X'\alpha) \in \mathcal{M}(X)$,从而结论成立。

7. 令矩阵 $\Sigma = (1 - \rho)I_n + \rho I_n I'_n$,其中 I_n 为元素全为 I 的 n 维向量。求 Σ 的特征值和对应的特征向量。又问 ρ 取何值时 Σ (半) 正定?

解. 由第一章习题一 1(vii) 的结论,只需考察矩阵 $1_n1'_n$ 的特征值与特征向量即可。因为 $1_n1'_n$ 的秩为 1 且易见行和 n 为其特征值,所以矩阵 $1_n1'_n$ 的特征值为 n (1 重) 和 0 (n-1 重)。同时容易求出对应的特征值为 1_n 和 $\epsilon_1-\epsilon_j$, $j=2,\cdots,n$ 。从而矩阵 Σ 的特征值为 $1+(n-1)\rho$ (1 重) 和 $1-\rho$ (n-1 重),所对应的特征向量即为上述特征向量。

由于 Σ 是对称矩阵,所以判断 Σ 是否(半)正定只需考察其特征值。于是容易得出, $\rho=-1/(n-1)$ 或 $\rho=1$ 时, Σ 严格半正定; $-1/(n-1)<\rho<1$ 时, Σ 正定。

- 8. 在 p 维向量空间 \mathbb{R}^p 中,超平面是由线性方程组 $H_{(A,b)}=\{x\,|\,Ax=b\}$ 所定义的点集。
 - (i) 取定点 $x_0 \in H_{(A,b)}$, 证明: $H_{(A,b)} = \{x \mid (x x_0) \perp \mathcal{M}(A')\}$;
 - (ii) 若矩阵 A 行满秩,对 $\forall x \in \mathbb{R}^p$,计算欧氏距离 $\mathrm{d}(x\,,H_{(A,b)})$ 。
 - **解.** 不妨设 A 为 $m \times p$ 阶矩阵。
 - (i) 记集合 $S_1 = \{x \mid Ax = b\}$, $S_2 = \{x \mid (x x_0) \perp \mathcal{M}(A')\}$ 。要证明 S_1 与 S_2 相互包含。

一方面,任取 $x \in S_1$,有 Ax = b。对 $\forall \alpha \in \mathcal{M}(A')$,存在向量 $\beta \in \mathbb{R}^m$,使得 $\alpha = A'\beta$ 。 从而

$$(\alpha, x - x_0) = \beta' A(x - x_0)$$
$$= \beta'(b - b) = 0$$

即 $(x-x_0) \perp \alpha$ $(\forall \alpha \in \mathcal{M}(A'))$ 。 所以 $x \in S_2$, $S_1 \subset S_2$ 。

另一方面,任取 $y\in S_2$,有 $(y-x_0)\perp \mathcal{M}(A')$ 。取 \mathbb{R}^m 中一个基 $\gamma_1\,,\cdots\,,\gamma_m$,令矩阵 $T=(\gamma_1\,,\cdots\,,\gamma_m)$,则 T 为 m 阶可逆矩阵。由于

$$0 = (A'\gamma_1, y - x_0) = \gamma'_1(Ay - b)$$
$$0 = (A'\gamma_2, y - x_0) = \gamma'_2(Ay - b)$$

. . .

$$0 = (A'\gamma_m, y - x_0) = \gamma'_m(Ay - b)$$

从而 T'(Ay-b)=0, Ay-b=0, 故 $y\in S_1$, $S_2\subset S_1$ 。

综上, S_1 与 S_2 相互包含,结论成立。事实上, $H_{(A,b)}$ 为 \mathbb{R}^p 上的线性流形, $H_{(A,b)}-x_0$ 为 \mathbb{R}^p 的线性子空间。从 $H_{(A,b)}-x_0$ 的定义上可以看出 $H_{(A,b)}-x_0$ 是 $\mathcal{M}(A')$ 的正交补空间,于是 $H_{(A,b)}=\{x\,|\,(x-x_0)\perp\mathcal{M}(A')\}$ 。

- (ii) 由于矩阵 A 行满秩,所以方程组 Ax = b 必有解,从而 $H_{(A,b)}$ 不为空集。记方程组 Ax = 0 的一个基解矩阵为 U,则 $\mathrm{d}(x\,,H_{(A,b)})$ 即为向量 $x-x_0$ 到 $\mathcal{M}(U)$ 的距离。若记 y^* 为关于 y 的方程组 $U'Uy = U'(x-x_0)$ 的解,则 $\mathrm{d}(x\,,H_{(A,b)}) = \|x-x_0-Uy^*\|_2$ 。
- 9. 设矩阵 A 非奇异,矩阵 D 为一方阵,证明:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|$$

证明. 由初等变换,易见

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D - CA^{-1}B \end{pmatrix}$$

从而结论成立。

- 10. 设矩阵 A, B 均为半正定矩阵且 A < B。
 - (i) 证明: $AB^{-}A \leq A$;
 - (ii) AB-A 依赖于广义逆的选择吗?
 - **证明.** (i) 事实上, B^- 不一定是对称矩阵,从而无法保证 AB^-A 是对称矩阵,偏序关系无从谈起。
 - (ii)

11. 举例说明:存在这样的矩阵 Σ , Σ 非半正定,但可以找到正交阵 $(P\ Q)$,使得 $P'\Sigma P=\Lambda_r$, $Q'\Sigma Q=0$ 。

证明. 令

$$\Sigma = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad P = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \qquad Q = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

容易验证,矩阵 Σ 的特征值为 0 和 1, $(P\ Q)$ 为正交阵且 $P'\Sigma P=1$, $Q'\Sigma Q=0$ 。

- 12. 设矩阵 A, B 均为半正定矩阵,且 A < B。证明:
 - (i) $tr(A) \leq tr(B)$;
 - (ii) 矩阵 A 的最大特征值不大于矩阵 B 的最大特征值;
 - (iii) 矩阵 A 的最小特征值不大于矩阵 B 的最大特征值;
 - (iv) $|A| \leq |B|$;
 - (v) $\mathcal{M}(A) \subset \mathcal{M}(B)$;
 - (vi) 设 P_A , P_B 分别是 A, B 的正交投影矩阵,则 $P_A \leq P_B$ 。

- **证明.** (i) 由 $A \leq B$ 知 $\forall x \in \mathbb{R}^m$,有 $x'Ax \leq x'Bx$ 。从而将 x 取遍 ϵ_i , $i=1,2,\cdots,m$,就得到 A 对角线元素的值均不大于 B 对角线元素的值。因此 $\operatorname{tr}(A) \leq \operatorname{tr}(B)$ 。
- (ii) 易见

$$\lambda_{\max}(A) = \max_{x'x \neq 0} \frac{x'Ax}{x'x} \le \max_{x'x \neq 0} \frac{x'Bx}{x'x} = \lambda_{\max}(B)$$

(iii) 同样地,有

$$\lambda_{\min}(A) = \min_{x'x \neq 0} \frac{x'Ax}{x'x} \le \min_{x'x \neq 0} \frac{x'Bx}{x'x} = \lambda_{\min}(B)$$

(iv) 不妨设矩阵 A,B 均正定。由 $A \leq B$ 知 $\forall x \in \mathbb{R}^m$ 且 $x \neq 0$,有 $0 < x'Ax \leq x'Bx$ 。从而

$$\begin{split} 1 &\geq \frac{x'Ax}{x'Bx} \\ &= \frac{y'B^{-\frac{1}{2}}AB^{-\frac{1}{2}}y}{y'y} \qquad (y = B^{\frac{1}{2}}x) \end{split}$$

因此 $\lambda_{\max}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})\leq 1$ 。又易见矩阵 $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$ 为正定矩阵,从而其特征值介于 0 和 1 之间(不包含 0)。因此 $|B^{-\frac{1}{2}}AB^{-\frac{1}{2}}|\leq 1$,即 $|A|\leq |B|$ 。

(v) 对 $x \in \mathcal{M}(B)^{\perp}$,有

$$Bx = 0 \Rightarrow x'Bx = 0$$

$$\Rightarrow x'Ax \le x'Bx = 0$$

$$\Rightarrow x'Ax = 0$$

$$\Rightarrow (A^{\frac{1}{2}}x)'(A^{\frac{1}{2}}x) = 0$$

$$\Rightarrow (A^{\frac{1}{2}}x) = 0$$

$$\Rightarrow Ax = 0$$

$$\Rightarrow \mathcal{M}(B)^{\perp} \subset \mathcal{M}(A)^{\perp}$$

从而 $\mathcal{M}(A) \subset \mathcal{M}(B)$ 。

(vi) 对 $\forall \xi \in \mathbb{R}^m$,存在 $\xi_1 \in \mathcal{M}(A)$ 和 $\xi_2 \in \mathcal{M}(A)^{\perp}$ 使得 $\xi = \xi_1 + \xi_2$ 。从 而

$$\xi' P_A \xi = \xi' P_A P_A \xi = (P_A \xi)'(P_A \xi) = \xi_1' \xi_1$$

于是

$$\xi' P_B \xi = \xi_1' P_B \xi_1 + \xi_1' P_B \xi_2 + \xi_2' P_B \xi_1 + \xi_2' P_B \xi_2$$

$$= (P_B \xi)' (P_B \xi) + (P_B \xi_1)' \xi_2 + \xi_2' (P_B \xi_1) + \xi_2' P_B \xi_2$$

$$= \xi_1' \xi_1 + \xi_1 \cdot \xi_2 + \xi_2' \xi_1 + \xi_2' P_B \xi_2$$

$$= \xi_1' \xi_1 + \xi_2' P_B \xi_2$$

$$\geq \xi_1' \xi_1 = \xi' P_A \xi$$

从而 $P_A \leq P_B$ 。

13. 设矩阵 A,B 均为半正定矩阵,且 $\mathcal{M}(A)\subset\mathcal{M}(B)$ 。举例说明 $A\leq B$ 未必成立。

证明. 令 $A = \operatorname{diag}(1,3,0,0)$, $B = \operatorname{diag}(1,1,1,0)$,则易见 A 每一列都可由 B 线性表出,从而 $\mathcal{M}(A) \subset \mathcal{M}(B)$ 。 取向量 x = (0,1,0,0)',则 x'Ax > x'Bx,即 $A \leq B$ 不成立。