回归分析作业

sword

2019-09-25

目录

1 第一章习题

1.1 习题一

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 $\lambda_1, \lambda_2, \cdots, \lambda_n$ ），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 $\mathrm{I}_n + cA$ 的特征值为 $1 + c\lambda_i$，$i = 1, 2, \cdots, n$ ；

(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 $\eta_1, \cdots, \eta_n$ ，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 $B = Q'\Lambda_1 Q$ ，其中 Λ_1 为对角阵。令 $Q_1 = \mathrm{diag}(1\,, Q)$ ，就有

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

再取 $P = Q_1 T'$ ，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 $PP' = \mathrm{I}$，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 $\mathrm{tr}(A^s) = \mathrm{tr}(\Lambda^s) = \sum_{i=1}^n \lambda_i^s$ 。

(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 $\mathrm{I}_n + cA = PP' + P(c\Lambda)P' = P(\mathrm{I}_n + c\Lambda)P'$ ，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

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

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：

$$\max_{x'x \neq 0} \frac{x'Ax}{x'x} = \lambda_{\max}(A) \quad 及 \quad \min_{x'x \neq 0} \frac{x'Ax}{x'x} = \lambda_{\min}(A)$$

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

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

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 $y=(y_1\,,\cdots\,,y_n)'$ ，$\Lambda=\mathrm{diag}(\lambda_1\,,\cdots\,,\lambda_n)$ ，则

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

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

$$\max_y f(y) \qquad \text{s.t. } \sum_{i=1}^n y_i^2=1$$
$$\min_y f(y) \qquad \text{s.t. } \sum_{i=1}^n y_i^2=1$$

应用拉格朗日乘数法，构造拉格朗日函数

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

上式两端对各分量求一阶偏导，得

$$\frac{\partial L}{\partial y_i}=2\lambda_i y_i-2\mu y_i=0 \qquad i=1\,,2\,,\cdots\,,n$$
$$\frac{\partial L}{\partial \mu}=y_1^2+\cdots+y_n^2-1=0$$

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 $\beta=(\beta_1\,,\beta_2\,,\cdots\,,\beta_n)'$ ，证明：
(1) 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

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

(2) 若 A 为对称矩阵，则

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

证明

(1) 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

$$\begin{aligned}\frac{\partial A\beta}{\partial\beta'}&=\frac{\partial}{\partial\beta'}\begin{pmatrix}b_{11}\beta_1+\cdots+b_{1n}\beta_n\\ b_{21}\beta_1+\cdots+b_{2n}\beta_n\\ \vdots\\ b_{n1}\beta_1+\cdots+b_{nn}\beta_n\end{pmatrix}\\&=\begin{pmatrix}\frac{\partial}{\partial\beta_1}\\ \frac{\partial}{\partial\beta_2}\\ \vdots\\ \frac{\partial}{\partial\beta_n}\end{pmatrix}(b_{11}\beta_1+\cdots+b_{1n}\beta_n\,,\cdots\,,b_{n1}\beta_1+\cdots+b_{nn}\beta_n)\\&=A\end{aligned}$$

(2) 对任意的 n 阶方阵 $A=(b_{ij})_n$，有

$$\begin{aligned}\frac{\partial\beta' A\beta}{\partial\beta}&=\frac{\partial}{\partial\beta}\left(\sum_{i=1}^n\sum_{j=1}^n b_{ij}\beta_i\beta_j\right)\\&=\left(\frac{\partial}{\partial\beta_1}\,,\frac{\partial}{\partial\beta_2}\,,\cdots\,,\frac{\partial}{\partial\beta_n}\right)'\left(\sum_{i=1}^n\sum_{j=1}^n b_{ij}\beta_i\beta_j\right)\\&=(A+A')\beta\end{aligned}$$

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

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

从而

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

5.(1) 举例说明矩阵 A 的广义逆 A^- 不唯一；
(2) 给出计算矩阵 A 广义逆 A^- 的一般公式。

解

(1) 先回顾矩阵广义逆的定义。
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

$$AAA = A$$
$$A\mathrm{I}A = A$$

成立。从而 A 的广义逆可以是 A 和 I ，不唯一。

(2) 记矩阵 A 的秩为 r ，对 A 作满秩分解

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

则矩阵 A 的广义逆的一般形式为

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

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：

(1) 无论 $(X'X)^-$ 如何变化，$X(X'X)^-X'$ 保持不变；

(2) $X(X'X)^-X'$ 是一个从 $\mathbb{R}^n$ 到 $\mathcal{M}(X)$ 的投影矩阵。

证明

(1) 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

$$PA = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix} \triangleq \Lambda$$

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

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

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。

(2) 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 $\Sigma = (1-\rho)\mathrm{I}_n + \rho 1_n 1_n'$ ，其中 1_n 为元素全为 1 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 $1_n 1_n'$ 的特征值与特征向量即可。因为 $1_n 1_n'$ 的秩为 1 且易见行和 n 为其特征值，所以矩阵 $1_n 1_n'$ 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 $H_{(A,b)} = \{x \mid Ax = b\}$ 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

$$
\begin{aligned}
(\alpha, x - x_0) &= \beta' A(x - x_0) \\
&= \beta'(b - b) = 0
\end{aligned}
$$

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

$$
\begin{aligned}
&0 = (A'\gamma_1, y - x_0) = \gamma_1'(Ay - b) \\
&0 = (A'\gamma_2, y - x_0) = \gamma_2'(Ay - b) \\
&\cdots \\
&0 = (A'\gamma_m, y - x_0) = \gamma_m'(Ay - b)
\end{aligned}
$$

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 $H_{(A,b)} = \{x \mid (x - x_0) \perp \mathcal{M}(A')\}$ 。

(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

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

从而结论成立。

10. 设矩阵 A, B 均为半正定矩阵且 $A \leq B$ 。

(i) 证明：$AB^{-}A \leq A$ ；

(ii) $AB^{-}A$ 依赖于广义逆的选择吗？

解

(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 $P'\Sigma P = \Lambda_r\,, Q'\Sigma Q = 0$ 。

解

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 $A \leq B$ 。证明：

(i) $\mathrm{tr}(A) \leq \mathrm{tr}(B)$ ；

(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) $|A| \le |B|$ ；

(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 $P_A \le P_B$ 。

证明

设矩阵 A, B 的阶数为 m 。

(i) 由 $A \le B$ 知 $\forall x \in \mathbb{R}^m$ ，有 $x'Ax \le x'Bx$ 。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 $\mathrm{tr}(A) \le \mathrm{tr}(B)$ 。

(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 $A \le B$ 知 $\forall x \in \mathbb{R}^m$ 且 $x \neq 0$ ，有 $0 < x'Ax \le x'Bx$ 。从而

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

因此 $\lambda_{\max}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}) \le 1$ 。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 $|B^{-\frac{1}{2}}AB^{-\frac{1}{2}}| \le 1$ ，即 $|A| \le |B|$ 。

(v) 对 $x \in \mathcal{M}(B)^\perp$ ，有

$$\begin{aligned} Bx = 0 \quad &\Rightarrow x'Bx = 0 \\ &\Rightarrow x'Ax \le x'Bx = 0 \\ &\Rightarrow x'Ax = 0 \\ &\Rightarrow (A^{\frac{1}{2}}x)'(A^{\frac{1}{2}}x) = 0 \\ &\Rightarrow (A^{\frac{1}{2}}x) = 0 \\ &\Rightarrow Ax = 0 \\ &\Rightarrow \mathcal{M}(B)^\perp \subset \mathcal{M}(A)^\perp \end{aligned}$$

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

$$\begin{aligned}
\xi' P_B \xi &= \xi_1' P_B \xi_1 + \xi_1' P_B \xi_2 + \xi_2' P_B \xi_1 + \xi_2' P_B \xi_2 \\
&= (P_B\xi)'(P_B\xi) + (P_B\xi_1)'\xi_2 + \xi_2'(P_B\xi_1) + \xi_2' P_B \xi_2 \\
&= \xi_1'\xi_1 + \xi_1\xi_2 + \xi_2'\xi_1 + \xi_2' P_B \xi_2 \\
&= \xi_1'\xi_1 + \xi_2' P_B \xi_2 \\
&\geq \xi_1'\xi_1 = \xi' P_A \xi
\end{aligned}$$

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。

解

令 $A = \mathrm{diag}(1\,, 3\,, 0\,, 0)$,$B = \mathrm{diag}(1\,, 1\,, 1\,, 0)$ ，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。

1.2 习题二

1. 设随机变量 X_1, X_2, X_3 有相同的均值 μ，协方差矩阵为

$$\mathrm{Var}(\mathbf{X}) = \sigma^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{4} \\ 0 & \frac{1}{4} & 1 \end{pmatrix}$$

求 $\mathrm{E}(X_1^2 + 2X_1X_2 - 4X_2X_3 + X_3^2)$。

解

由题设，得

$$\begin{aligned} & \mathrm{E}(X_1^2 + 2X_1X_2 - 4X_2X_3 + X_3^2) \\ = & \mathrm{Var}(X_1) + \mathrm{E}^2(X_1) + 2\mathrm{Cov}(X_1, X_2) + 2\mathrm{E}(X_1)E(X_2) \\ & - 4\mathrm{Cov}(X_2, X_3) - 4\mathrm{E}(X_2)E(X_3) + \mathrm{Var}(X_3) + \mathrm{E}^2(X_3) \\ = & \sigma^2 \end{aligned}$$

2. 令 $\mathbf{X} = (X_1, X_2, X_3)'$ 且

$$\mathrm{Var}(\mathbf{X}) = \begin{pmatrix} 5 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 3 \end{pmatrix}$$

(a) 求 $X_1 - 2X_2 + X_3$ 的方差；
(b) 求随机向量 $Y = (Y_1, Y_2)'$ 的协方差阵，其中 $Y_1 = X_1 + X_2$, $Y_2 = X_1 + X_2 + X_3$。

解

(a) 由题设，得

$$\begin{aligned} & \mathrm{Var}(X_1 - 2X_2 + X_3) \\ = & \mathrm{Var}(X_1) + 4\mathrm{Var}(X_2) + \mathrm{Var}(X_3) \\ & - 4\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) - 4\mathrm{Cov}(X_2, X_3) \\ = & 18 \end{aligned}$$

(b) 易见

$$\begin{aligned} & \mathrm{Var}(X_1 + X_2) = 12 \\ & \mathrm{Var}(X_1 + X_2 + X_3) = 21 \\ & \mathrm{Cov}(X_1 + X_2, X_1 + X_2 + X_3) = \mathrm{Var}(X_1 + X_2) + \mathrm{Cov}(X_1 + X_2, X_3) = 15 \end{aligned}$$

从而

$$
\begin{aligned}
\mathrm{Cov}(Y) &= \begin{pmatrix} \mathrm{Var}(X_1+X_2) & \mathrm{Cov}(X_1+X_2\,,X_1+X_2+X_3) \\ \mathrm{Cov}(X_1+X_2\,,X_1+X_2+X_3) & \mathrm{Var}(X_1+X_2+X_3) \end{pmatrix} \\
&= \begin{pmatrix} 12 & 15 \\ 15 & 21 \end{pmatrix}
\end{aligned}
$$

3. 设 $X\,,Y$ 均为随机向量，根据重期望公式

$$\mathrm{E}(U)=\mathrm{E}(\mathrm{E}(U\,|\,V))$$

证明

$$\mathrm{Var}(X)=\mathrm{E}(\mathrm{Var}(X\,|\,Y))+\mathrm{Var}(\mathrm{E}(X\,|\,Y))$$

你能说出上式在概率统计中的解释吗？

证明

由

$$\mathrm{Var}(X\,|\,Y)=\mathrm{E}(X^2\,|\,Y)-\mathrm{E}^2(X\,|\,Y)$$

得

$$\mathrm{E}(\mathrm{Var}(X\,|\,Y)=\mathrm{E}(X^2)-\mathrm{E}(\mathrm{E}^2(X\,|\,Y))$$

又

$$
\begin{aligned}
\mathrm{Var}(\mathrm{E}(X\,|\,Y)) &= \mathrm{E}(\mathrm{E}^2(X\,|\,Y))-\mathrm{E}^2(\mathrm{E}(X\,|\,Y)) \\
&= \mathrm{E}(\mathrm{E}^2(X\,|\,Y))-\mathrm{E}^2(X)
\end{aligned}
$$

综上，知结论成立。形象地来看，$\mathrm{Var}(X\,|\,Y)$ 可分解为剖面上波动的均值加上剖面上均值的波动。

4. 设 $\mathbf{X}\,,\mathbf{Y}$ 是两个随机变量，A 为一纯量矩阵，证明

$$\mathrm{E}(\mathbf{X}'A\mathbf{Y})=\mathrm{tr}(A\mathrm{Cov}(\mathbf{Y}\,,\mathbf{X}))+\mathrm{E}(\mathbf{X})'A\mathrm{E}(\mathbf{Y})$$

证明

由

$$\begin{aligned}
\mathrm{E}(\mathbf{X}'A\mathbf{Y}) = \mathrm{E}[\mathrm{tr}(\mathbf{X}'A\mathbf{Y})] &= \mathrm{E}[\mathrm{tr}(A\mathbf{Y}\mathbf{X}')] \\
&= \mathrm{tr}[\mathrm{E}(A\mathbf{Y}\mathbf{X}')] \\
&= \mathrm{tr}[A\mathrm{E}(\mathbf{Y}\mathbf{X}')] \\
&= \mathrm{tr}[A(\mathrm{Cov}(\mathbf{Y}\,,\mathbf{X}) + \mathrm{E}(\mathbf{Y})\mathrm{E}(\mathbf{X}'))] \\
&= \mathrm{tr}[A\mathrm{Cov}(\mathbf{Y}\,,\mathbf{X})] + \mathrm{tr}[A\mathrm{E}(\mathbf{Y})\mathrm{E}(\mathbf{X}')] \\
&= \mathrm{tr}[A\mathrm{Cov}(\mathbf{Y}\,,\mathbf{X})] + \mathrm{tr}[\mathrm{E}(\mathbf{X}')A\mathrm{E}(\mathbf{Y})] \\
&= \mathrm{tr}(A\mathrm{Cov}(\mathbf{Y}\,,\mathbf{X})) + \mathrm{E}(\mathbf{X}')A\mathrm{E}(\mathbf{Y})
\end{aligned}$$

知结论成立。

5. 设随机向量 $\mathbf{X} = (X_1\,, X_2\,, \cdots\,, X_n)'$ 。令

$$\begin{aligned}
&Y_1 = X_1 \\
&Y_i = X_i - X_{i-1} \qquad i = 2\,,\,,3\,,\cdots\,,n
\end{aligned}$$

若诸 Y_i 相互独立，求 $\mathrm{Var}(\mathbf{X})$ 。

解

设 $\mathrm{Cov}(X_i\,, X_j) = \sigma_{ij}$ 。由题设，得

$$\begin{aligned}
&0 = \mathrm{Cov}(X_1\,, X_2 - X_1) = \sigma_{12} - \sigma_{11} \\
&0 = \mathrm{Cov}(X_1\,, X_3 - X_2) = \sigma_{13} - \sigma_{12} \\
&\cdots \\
&0 = \mathrm{Cov}(X_1\,, X_n - X_{n-1}) = \sigma_{1n} - \sigma_{1(n-1)}
\end{aligned}$$

于是

$$\sigma_{11} = \sigma_{12} = \cdots = \sigma_{1n}$$

接下来，由

$$\begin{aligned}
&0 = \mathrm{Cov}(X_2 - X_1\,, X_3 - X_2) = \sigma_{23} - \sigma_{22} - \sigma_{13} + \sigma_{12} = \sigma_{23} - \sigma_{22} \\
&0 = \mathrm{Cov}(X_2 - X_1\,, X_4 - X_3) = \sigma_{24} - \sigma_{23} - \sigma_{14} + \sigma_{13} = \sigma_{24} - \sigma_{23} \\
&\cdots \\
&0 = \mathrm{Cov}(X_2 - X_1\,, X_n - X_{n-1}) = \sigma_{2n} - \sigma_{2(n-1)} - \sigma_{1n} + \sigma_{1(n-1)} = \sigma_{2n} - \sigma_{2(n-1)}
\end{aligned}$$

得

$$\sigma_{22} = \sigma_{23} = \cdots = \sigma_{2n}$$

由此推知

$$\sigma_{ii} = \sigma_{ij} \qquad i = 1, \cdots, n\ ,\ j = i, \cdots, n$$

因此

$$\mathrm{Var}(\mathbf{X}) = \begin{pmatrix} \sigma_{11} & \sigma_{11} & \sigma_{11} & \cdots & \sigma_{11} \\ \sigma_{11} & \sigma_{22} & \sigma_{22} & \cdots & \sigma_{22} \\ \sigma_{11} & \sigma_{22} & \sigma_{33} & \cdots & \sigma_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{11} & \sigma_{22} & \sigma_{33} & \cdots & \sigma_{nn} \end{pmatrix}$$

6. 设随机变量序列 $X_1, X_2 \cdots, X_n \overset{\text{i.i.d}}{\sim} \mathrm{N}(\mu, \sigma^2)$ 。定义

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad Q = \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2$$

(1) 证明 $\mathrm{Var}(S^2) = 2\sigma^4/(n-1)$ ；

(2) 证明 Q 是 σ^2 的无偏估计；

(3) 求 $\mathrm{Var}(Q)$ ，并由此说明 $n \to \infty$ 时，Q 的功效是 S^2 的 $2/3$ 。

证明

(1) 易见

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

从而

$$\mathrm{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

故 $\mathrm{Var}(S^2) = 2\sigma^4/(n-1)$ 。

(2) 由 $X_{i+1} - X_i \sim \mathrm{N}(0, 2\sigma^2)$ ，从而

$$\mathrm{E}(Q) = \frac{1}{2(n-1)}\sum_{i=1}^{n-1}\mathrm{E}(X_{i+1} - X_i)^2 = \sigma^2$$

即 Q 是 σ^2 的无偏估计。

(3) 不失一般性，令 $\mu=0$ ，则

$$
\begin{aligned}
&\operatorname{Var}\left(\frac{1}{2(n-1)}\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right)\\
=&\frac{1}{4(n-1)^2}\operatorname{Var}\left(X_1^2+2X_2^2+\cdots+2X_{n-1}^2+X_n^2-2X_1X_2-\cdots-2X_{n-1}X_n\right)\\
=&\frac{1}{4(n-1)^2}(\operatorname{Var}(X_1^2)+4\sum_{i=2}^{n-1}\operatorname{Var}(X_i^2)+\operatorname{Var}(X_n^2)-4\operatorname{Cov}(X_1^2,X_1X_2)\\
&-8\sum_{i=2}^{n-1}\operatorname{Cov}(X_i^2,X_{i-1}X_i)-8\sum_{i=2}^{n}\operatorname{Cov}(X_i^2,X_iX_{i+1})-4\operatorname{Cov}(X_n^2,X_{n-1}X_n))
\end{aligned}
$$

其中

$$
\begin{aligned}
\operatorname{Var}(X_i^2)&=\mathrm{E}(X_i^4)-\mathrm{E}^2(X_i^2)=3\sigma^4-\sigma^4=2\sigma^4\\
\operatorname{Cov}(X_i^2,X_{i-1}X_i)&=\mathrm{E}(X_i^3X_{i-1})-\mathrm{E}(X_i^2)\mathrm{E}(X_iX_{i-1})=0\\
\operatorname{Cov}(X_i^2,X_iX_{i+1})&=\mathrm{E}(X_i^3X_{i+1})-\mathrm{E}(X_i^2)\mathrm{E}(X_iX_{i+1})=0
\end{aligned}
$$

因此

$$\operatorname{Var}(Q)=\frac{3\sigma^4}{n-1}$$

故

$$\lim_{n\to\infty}\frac{\operatorname{Var}(Q)}{\operatorname{Var}(S^2)}=\frac{3}{2}$$

即 $n\to\infty$ 时，Q 的功效是 S^2 的 $2/3$ 。

7. 设随机变量序列 $X_1,X_2\cdots,X_n\overset{\text{i.i.d}}{\sim}\mathrm{N}(0,\sigma^2)$ 。令 $\mathbf{X}=(X_1,X_2\cdots,X_n)'$ ，A,B 为任意 n 阶对称矩阵。证明

$$\operatorname{Cov}(\mathbf{X}'A\mathbf{X},\mathbf{X}'B\mathbf{X})=2\sigma^4\operatorname{tr}(AB)$$

证明

易见

$$
\begin{aligned}
&\operatorname{Cov}(\mathbf{X}'A\mathbf{X},\mathbf{X}'B\mathbf{X})\\
=&\mathrm{E}(\mathbf{X}'A\mathbf{X}\mathbf{X}'B\mathbf{X})-\mathrm{E}(\mathbf{X}'A\mathbf{X})\mathrm{E}(\mathbf{X}'B\mathbf{X})\\
=&3\sigma^4\sum_{i=1}^{n}a_{ii}b_{ii}+\sigma^4\left(\sum_{i=1}^{n}\sum_{j\neq i}a_{ii}b_{jj}+\sum_{i=1}^{n}\sum_{j\neq i}a_{ij}b_{ij}+\sum_{i=1}^{n}\sum_{j\neq i}a_{ij}b_{ji}\right)\\
&-\sigma^4\operatorname{tr}(A)\operatorname{tr}(B)\\
=&2\sigma^4\sum_{i=1}^{n}a_{ii}b_{ii}+\sigma^4\left(\sum_{i=1}^{n}\sum_{j\neq i}a_{ij}b_{ij}+\sum_{i=1}^{n}\sum_{j\neq i}a_{ij}b_{ji}\right)\\
=&2\sigma^4\operatorname{tr}(AB)
\end{aligned}
$$

8. 设随机向量 $\mathbf{X}=(X_1,X_2\cdots,X_n)'$ 的各分量有相同的均值 μ 和方差 σ^2，且两两之间的相关系数均为 ρ 。
(1) 求 ρ 的范围;
(2) 构造满足题设的 $\mathbf{X}$;
(3) 记样本方差 $Q=\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2$,
(i) 求 $\mathrm{E}(Q)$;
(ii) 若 $A=a\sum_{i=1}^n X_i^2+b(\sum_{i=1}^n X_i)^2$ 是 σ^2 的无偏估计，求 a,b ，并由此说明，此时

$$A=\sum_{i=1}^n\frac{(X_i-\bar{X})^2}{(1-\rho)(n-1)}$$

证明

(1) 易见 $\mathrm{Var}(\mathbf{X})=\sigma^2((1-\rho)\mathrm{I}_n+\rho 1_n1_n')$ 。从而，为使 $\mathrm{Var}(\mathbf{X})$ 半正定，由第一章习题一 7 的结论，知

$$-\frac{1}{n-1}\le\rho\le 1$$

(2) 令 $\mathbf{X}=(X_1,X_2\cdots,X_n)'\sim\mathrm{N}(\boldsymbol{\mu},\Sigma)$ 即可，其中

$$\begin{aligned}\boldsymbol{\mu}&=\mu 1_n\\ \Sigma&=\sigma^2((1-\rho)\mathrm{I}_n+\rho 1_n1_n')\end{aligned}$$

(3)

(i) 易见

$$\begin{aligned}&\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2\right)\\ =&\frac{1}{n}\mathrm{E}\left(\mathbf{X}'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\mathbf{X}\right)\\ =&\frac{1}{n}\mathrm{tr}\left(\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\mathrm{Cov}(\mathbf{X})\right)+\mathrm{E}(\mathbf{X}')\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\mathrm{E}(\mathbf{X})\\ =&\sigma^2\mathrm{tr}\left((1-\rho)\mathrm{I}_n-\frac{1-\rho}{n}1_n1_n'\right)+\mu^2 1_n'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)1_n\\ =&\sigma^2(1-\rho)\frac{n-1}{n}\end{aligned}$$

(ii) 由

$$\begin{aligned}\sigma^2&=\mathrm{E}\left(a\sum_{i=1}^n X_i^2+b\left(\sum_{i=1}^n X_i\right)^2\right)\\ &=an(\sigma^2+\mu^2)+b[n(\sigma^2+\mu^2)+n(n-1)(\rho\sigma^2+\mu^2)]\end{aligned}$$

解得

$$a = \frac{1}{(n-1)(1-\rho)}$$
$$b = -\frac{1}{n(n-1)(1-\rho)}$$

因此

$$\begin{aligned} A &= \frac{1}{(n-1)(1-\rho)}\sum_{i=1}^{n} X_i^2 - \frac{1}{n(n-1)(1-\rho)}\left(\sum_{i=1}^{n} X_i\right)^2 \\ &= \frac{1}{(n-1)(1-\rho)}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right) \\ &= \sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{(1-\rho)(n-1)} \end{aligned}$$

9. 设随机变量序列 $X_1, X_2 \cdots, X_n, \cdots$ 的各分量有相同的均值 μ 和方差 σ^2 ，且两两之间的相关系数均为 ρ 。

(1) 求 ρ 的范围；

(2) 构造满足题设的序列且 ρ 不能为 0。

解

(1) 任取 $X_1, X_2 \cdots, X_n, \cdots$ 的有限子集 $X_{n_1}, X_{n_2}, \cdots, X_{n_r}$ ，其协方差矩阵为 $\sigma^2((1-\rho)\mathrm{I}_r + \rho 1_r 1_r')$ 。与第一章习题二 8(1) 类似，ρ 需满足

$$-\frac{1}{r-1} \le \rho \le 1 \qquad \forall r \ge 2$$

因此 $0 \le \rho \le 1$ 。

(2) 令随机变量 $X \sim \mathrm{N}(\mu, \sigma^2)$ ， $X_i = X$ $(\forall i)$ 。则易见此序列满足题设且 $\rho = 1$ 。

1.3 习题三

1. 证明：

(1) 设随机变量 $X \sim \mathrm{N}(0\,,1)$ ，则 $M_X(t)=\mathrm{e}^{\frac{t^2}{2}}$ ；

(2) 设随机变量 $X \sim \chi^2(r)$ ，则 $M_X(t)=(1-2t)^{-\frac{r}{2}} \quad t\in[0\,,\frac{1}{2})$ 。

证明

(1) 已知 X 的特征函数为 $\varphi(t)=\mathrm{e}^{-\frac{t^2}{2}}$ ，故

$$M_X(t)=\mathrm{E}\left(\mathrm{e}^{tX}\right)=\mathrm{E}\left(\mathrm{e}^{\mathrm{i}\frac{t}{\mathrm{i}}X}\right)=\varphi\left(\frac{t}{\mathrm{i}}\right)=\mathrm{e}^{\frac{t^2}{2}}$$

(2) 易见

$$\begin{aligned}
\mathrm{E}\left(\mathrm{e}^{tX}\right)&=\int_0^{\infty}\mathrm{e}^{tx}\frac{\left(\frac{1}{2}\right)^{\frac{r}{2}}}{\Gamma\left(\frac{r}{2}\right)}x^{\frac{r}{2}-1}\mathrm{e}^{-\frac{1}{2}x}\,\mathrm{d}x\\
&=\int_0^{\infty}\frac{\left(\frac{1}{2}\right)^{\frac{r}{2}}}{\Gamma\left(\frac{r}{2}\right)}x^{\frac{r}{2}-1}\mathrm{e}^{-\left(\frac{1}{2}-t\right)x}\,\mathrm{d}x\\
&=\frac{\left(\frac{1}{2}\right)^{\frac{r}{2}}}{\left(\frac{1}{2}-t\right)^{\frac{r}{2}}}\\
&=(1-2t)^{-\frac{r}{2}}\qquad t<\frac{1}{2}
\end{aligned}$$

2. 求二项分布、多维正态分布、指数分布、几何分布与泊松分布的矩母函数。

解

(i) 二项分布 $b(n\,,p)$ ：

$$\begin{aligned}
M(t)&=\sum_{k=0}^{n}\mathrm{e}^{tk}C_n^k p^k(1-p)^{n-k}\\
&=\sum_{k=0}^{n}C_n^k(p\mathrm{e}^t)^k(1-p)^{n-k}\\
&=(p\mathrm{e}^t+1-p)^n
\end{aligned}$$

(ii) 多维正态分布 $\mathrm{N}(\boldsymbol{\mu}\,,\Sigma)$ ：

$$M(\mathbf{t})=\varphi\left(\frac{\mathbf{t}}{\mathrm{i}}\right)=\mathrm{e}^{\mathbf{t}'\boldsymbol{\mu}+\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}}$$

(iii) 指数分布 $Exp(\lambda)$ ：

$$\begin{aligned} M(t) &= \int_0^\infty \mathrm{e}^{tx}\lambda\mathrm{e}^{-\lambda x}\,\mathrm{d}x \\ &= \left(1-\frac{t}{\lambda}\right)^{-1} \qquad t<\lambda \end{aligned}$$

(iv) 几何分布 $Ge(p)$ ：

$$M(t)=\sum_{k=0}^{n}\mathrm{e}^{tk}p(1-p)^k=\frac{p}{1-(1-p)\mathrm{e}^t} \qquad t<-\ln(1-p)$$

(v) 泊松分布 $P(\lambda)$ ：

$$M(t)=\sum_{k=0}^{\infty}\mathrm{e}^{tk}\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda\left(1-\mathrm{e}^t\right)}$$

3. 证明：$MX(t)\geq \mathrm{e}^{\mu t}$ 。

证明

由指数函数的下凸性质及 Jensen 不等式，得

$$M_X(t)=\mathrm{E}\left(\mathrm{e}^{tX}\right)\geq \mathrm{e}^{\mathrm{E}(tX)}=\mathrm{e}^{\mu t}$$

4. 设随机变量 X 的矩母函数为 $M_X(t)=(1+\mathrm{e}^t)/2$ 。求 X 的方差。

解

易见

$$\mathrm{Var}(X)=M_X''(0)-[M_X'(0)]^2=\frac{1}{4}$$

5. 设实值随机变量 X 期望存在，且存在常数 a,b ，$a<0<b$ ，使得 $a\leq X\leq b$ 几乎处处成立。证明：对 $\forall\lambda\in\mathbb{R}$ ，有

$$\mathrm{E}(\mathrm{e}^{\lambda X})\leq \mathrm{e}^{\lambda^2(b-a)^2/8}$$

证明

我们需要增加一个条件 $\mathrm{E}(X)=0$ ，否则，可举反例，例如，令 $X\sim \mathrm{U}(a,b), a=-0.1, b=0.11, \lambda=1$ 。增加条件后的证明请参见 https://en.wikipedia.org/wiki/Hoeffding%27s_lemma 。

6. 设随机变量 X 的各阶矩非负，证明：

$$M_X(t) \geq \frac{t^k \mathrm{E}(X^k)}{k!}$$

证明

由 $\mathrm{E}(X^k) \geq 0$ ，得

$$\begin{aligned} M_X(t) &= \mathrm{E}(\mathrm{e}^{tX}) \\ &= \mathrm{E}\left(\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right) \\ &= \sum_{k=0}^{\infty} \frac{t^k \mathrm{E}(X^k)}{k!} \geq \frac{t^k \mathrm{E}(X^k)}{k!} \end{aligned}$$

7. 设离散型随机变量 $X\,,Y$ 均只取值 0 与 1，记 $p_{ij} = P\{X = i\,,Y = j\}$ 。证明：$X\,,Y$ 独立当且仅当 $\mathrm{Cov(X\,,Y)} = 0$ 。

证明

记 $p_{i+} = P\{X = i\}\,,p_{+i} = P\{Y = i\}$ 。

一方面，$X\,,Y$ 独立时，有 $\mathrm{E}(XY) = E(X)E(Y)$ 。从而

$$\mathrm{Cov(X\,,Y)} = \mathrm{E}(XY) - E(X)E(Y) = 0$$

另一方面，$\mathrm{Cov(X\,,Y)} = 0$ 时，有 $p_{11} = p_{1+} \cdot p_{+1}$ 。因此

$$p_{01} = p_{+1} - p_{11} = (1 - p_{1+})p_{+1} = p_{0+}p_{+1}$$
$$p_{00} = p_{0+} - p_{01} = p_{0+}(1 - p_{+1}) = p_{0+}p_{+0}$$
$$p_{10} = p_{+0} - p_{00} = (1 - p_{0+})p_{+0} = p_{1+}p_{+0}$$

于是 $X\,,Y$ 独立。

8. 设多维随机变量 $(X\,,Y\,,Z)$ 的联合密度函数为

$$f(x\,,y\,,z) = \frac{1}{8}(1 + xyz) \qquad -1 \leq x\,,y\,,z \leq 1$$

证明：$(X\,,Y\,,Z)$ 两两独立但不相互独立。

证明

由于 $f(x,y,z)$ 不能分解为形如 $f_1(x), f_2(y), f_3(z)$ 的函数的乘积，故 (X,Y,Z) 不相互独立。下面仅对 (X,Y) 相互独立作出说明，其余类似。事实上，易见 (X,Y) 的联合分布函数为

$$g(x,y) = 1/4 \qquad -1 \le x, y \le 1$$

因此 (X,Y) 服从矩形域上的均匀分布，从而相互独立。

2 第二章习题

1. 设随机向量

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y_1} \\ \boldsymbol{Y_2} \end{pmatrix} \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$$

其中，$\boldsymbol{Y_1}, \boldsymbol{Y_2}$ 是 $\boldsymbol{Y}$ 的分量（也是向量）且

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

证明：$\boldsymbol{Y_1} \perp \boldsymbol{Y_2}$ （即 $\mathrm{Cov}(\boldsymbol{Y_1}, \boldsymbol{Y_2}) = 0$ ）当且仅当 $\Sigma_{12} = 0$ 。

证明

易见

$$\begin{aligned} \mathrm{Cov}(\boldsymbol{Y_1}, \boldsymbol{Y_2}) &= \mathrm{Cov}\left((I\ O)\boldsymbol{Y}, (O\ I)\boldsymbol{Y}\right) \\ &= (I\ O)\,\mathrm{Cov}(\boldsymbol{Y}) \begin{pmatrix} O \\ I \end{pmatrix} \\ &= (I\ O) \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} O \\ I \end{pmatrix} \\ &= \Sigma_{12} \end{aligned}$$

因此 $\mathrm{Cov}(\boldsymbol{Y_1}, \boldsymbol{Y_2}) = 0 \Leftrightarrow \Sigma_{12} = 0$ ，结论成立。

2. 设随机向量 $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$ ，其中

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

求 $Z_1 = Y_1 + Y_2 + Y_3$ 与 $Z_2 = Y_1 - Y_2$ 的联合分布。

解

令矩阵

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

则 $(Z_1, Z_2)' = A\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu_1}, \Sigma_1)$ ，其中

$$\boldsymbol{\mu_1} = A\boldsymbol{\mu} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} \qquad \Sigma_1 = A\Sigma A' = \begin{pmatrix} 10 & 0 \\ 0 & 3 \end{pmatrix}$$

3. 设随机变量序列 $Y_1\,,Y_2\,,\cdots\,,Y_n \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\mu\,,\sigma^2)$ ，证明

$$\bar{Y} \perp S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)$$

证明

记随机向量 $\boldsymbol{Y} = (Y_1\,,Y_2\,,\cdots\,,Y_n)$ ，矩阵 $A = \mathrm{I}_n - \frac{1}{n}1_n1_n'$ 。则

$$\begin{aligned}\bar{Y} &= \frac{1}{n}1_n1_n'\boldsymbol{Y}\\ S^2 &= \frac{1}{n-1}(A\boldsymbol{Y})'(A\boldsymbol{Y})\end{aligned}$$

考察 $(\bar{Y}\,,A\boldsymbol{Y})'$ 的联合分布，其协方差阵为

$$\begin{pmatrix}\frac{1}{n}1_n1_n'\\ A\end{pmatrix}\sigma^2\mathrm{I}_n\begin{pmatrix}\frac{1}{n}1_n1_n' & A\end{pmatrix} = \sigma^2\begin{pmatrix}\frac{1}{n}1_n1_n' & 0\\ 0 & A\end{pmatrix}$$

从而 $\bar{Y}$ 与 $A\boldsymbol{Y}$ 相互独立。于是，作为 $A\boldsymbol{Y}$ 的函数，S^2 与 $\bar{Y}$ 相互独立，协方差阵为 0。

4. 设随机向量 $\boldsymbol{Y} \sim \mathrm{N}(0\,,\mathrm{I}_n)$ ，$\boldsymbol{X} = A\boldsymbol{Y}\,,\boldsymbol{U} = B\boldsymbol{Y}\,,\boldsymbol{V} = C\boldsymbol{Y}$ 。已知

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{X}\,,\boldsymbol{U}) = 0\\ \mathrm{Cov}(\boldsymbol{X}\,,\boldsymbol{V}) = 0\end{aligned}$$

证明 $\boldsymbol{X}$ 与 $\boldsymbol{U}+\boldsymbol{V}$ 相互独立。

证明

由题设，知 $AB' = AC' = 0$ ，从而

$$\begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{U}+\boldsymbol{V}\end{pmatrix} = \begin{pmatrix}A\\ B+C\end{pmatrix}\boldsymbol{Y} \sim \mathrm{N}(0\,,\Sigma)$$

其中

$$\Sigma = \begin{pmatrix}AA' & A(B+C)'\\ (B+C)A' & (B+C)(B+C)'\end{pmatrix} = \begin{pmatrix}AA' & 0\\ 0 & (B+C)(B+C)'\end{pmatrix}$$

于是结论成立。

5. 设随机向量 $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}\,,\Sigma)$ ，其中

$$\Sigma = \sigma^2\begin{pmatrix}1 & \rho & 0\\ \rho & 1 & \rho\\ 0 & \rho & 1\end{pmatrix}$$

已知随机变量 $Y_1+Y_2+Y_3$ 与 $Y_1-Y_2-Y_3$ 相互独立，求 ρ 的值。

解

首先，为使 Σ 半正定，需满足

$$1-\rho^2 \leq 0$$

$$1-2\rho^2 \leq 0$$

解出

$$|\,\rho\,| \leq \frac{\sqrt{2}}{2}$$

其次，易见 $Y_1+Y_2+Y_3$ 与 $Y_1-Y_2-Y_3$ 联合分布的协方差阵为

$$\Sigma = \sigma^2 \begin{pmatrix} 4\rho+3 & -2\rho-1 \\ -2\rho-1 & 3 \end{pmatrix}$$

从而，为使 $Y_1+Y_2+Y_3$ 与 $Y_1-Y_2-Y_3$ 相互独立，$\rho=-1/2$ ，满足题意。

6. 设随机变量序列 $Y_1\,,Y_2\,,\cdots\,,Y_n \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\mu\,,\sigma^2)$ ，证明 $\bar{Y}$ 与 $\sum_{i=1}^{n-1}(Y_i-Y_{i+1})^2$ 相互独立。

证明

记随机向量 $\boldsymbol{Y}=(Y_1\,,Y_2\,,\cdots\,,Y_n)$ ，矩阵

$$A = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

则 $\sum_{i=1}^{n-1}(Y_i-Y_{i+1})^2=\boldsymbol{Y}'A'A\boldsymbol{Y}$ 。由 $\frac{1}{n}1_n1_n'A=0$ 知结论成立。

7. 设随机向量 $\boldsymbol{Y}\sim\mathrm{N}(0\,,\mathrm{I}_n)$ ，矩阵 A 对称。

(1) 证明随机变量 $\boldsymbol{Y}A\boldsymbol{Y}$ 的矩母函数为 $[\det(\mathrm{I}_n-2tA)]^{-1/2}$ ；

(2) 若 A 为秩为 r 的幂等阵，证明上述矩母函数可化简为 $(1-2t)^{-r/2}$ ；

(3) 求随机向量 $\boldsymbol{Z}\sim\mathrm{N}(0\,,\Sigma)$ 的矩母函数。

证明

(a) 记矩阵 A 的特征值为 $\lambda_1, \cdots, \lambda_n$ 。由 A 对称，知存在正交阵 P ，对角阵

$$\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$$

使 $A = P'\Lambda P$ 。

记随机向量 $\boldsymbol{Z} = P\boldsymbol{Y} \sim \mathrm{N}(0, PP') \overset{d}{=} \mathrm{N}(0, \mathrm{I}_n)$ ，$\boldsymbol{Z}$ 的各分量为 Z_i ，则诸 $Z_i^2 \sim \chi_1^2$ 。于是

$$\begin{aligned}
\mathrm{E}\left(\exp\{t\boldsymbol{Y}'A\boldsymbol{Y}\}\right) &= \mathrm{E}\left(\exp\{t\boldsymbol{Z}'\Lambda\boldsymbol{Z}\}\right) \\
&= \mathrm{E}\left(\exp\{t\lambda_1 z_1^2 + \cdots + t\lambda_n z_n^2\}\right) \\
&= \prod_{i=1}^n \mathrm{E}\left(\exp\{t\lambda_i z_i^2\}\right) \\
&= \prod_{i=1}^n (1-2t\lambda_i)^{-1/2}
\end{aligned}$$

而

$$\begin{aligned}
\det(\mathrm{I}_n - 2tA)t^{-1/2} &= [\det\left(P'(\mathrm{I}_n - 2t\Lambda)P\right)]^{-1/2} \\
&= [\det(\mathrm{I}_n - 2t\Lambda)]^{-1/2} \\
&= \prod_{i=1}^n (1-2t\lambda_i)^{-1/2}
\end{aligned}$$

从而结论成立。

(b) 当 A 为秩为 r 的幂等阵时，A 的特征值中有 r 个为 1，其余为 0，从而易见结论成立。

(c) 由于 $\mathrm{N}(0, \Sigma)$ 的特征函数为

$$\varphi(\boldsymbol{t}) = \exp\left\{-\frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right\}$$

故

$$M_{\boldsymbol{Z}}(\boldsymbol{t}) = \varphi\left(\boldsymbol{t}/\mathrm{i}\right) = \exp\left\{\frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right\}$$

8. 设随机向量 $\boldsymbol{Y} \sim \mathrm{N}(\mu 1_n, \Sigma)$ ，其中

$$\Sigma = (1-\rho)\mathrm{I}_n + \rho 1_n 1_n' \qquad \rho > -\frac{1}{n-1}$$

证明

$$\sum_{i=1}^n \frac{(Y_i - \bar{Y})^2}{1-\rho} \sim \chi_{n-1}^2$$

证明

记随机向量 $\boldsymbol{Z}=(\sqrt{1-\rho})^{-1}\boldsymbol{Y}$ ，则

$$\begin{aligned}\sum_{i=1}^{n}\frac{(Y_i-\bar{Y})^2}{1-\rho}&=\boldsymbol{Z}'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{Z}\\&=\boldsymbol{Z}'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{Z}\end{aligned}$$

考察正态随机向量 $\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{Z}$ ，其均值为

$$\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\left(\sqrt{1-\rho}\right)^{-1}\mu1_n=0$$

协方差阵为

$$\frac{1}{1-\rho}\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)((1-\rho)\mathrm{I}_n+\rho1_n1_n')\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)=\mathrm{I}_n-\frac{1}{n}1_n1_n'$$

从而 $\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{Z}\overset{d}{=}\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{U}$ ，其中 $\boldsymbol{U}\sim\mathrm{N}(0,\mathrm{I}_n)$ 。于是

$$\sum_{i=1}^{n}\frac{(Y_i-\bar{Y})^2}{1-\rho}\overset{d}{=}\boldsymbol{U}'\left(\mathrm{I}_n-\frac{1}{n}1_n1_n'\right)\boldsymbol{U}$$

由于 $\mathrm{I}_n-\frac{1}{n}1_n1_n'$ 是秩为 $n-1$ 的对称幂等阵，根据定理 2.5，结论成立。

9. 设随机向量 $\boldsymbol{Y}\sim\mathrm{N}(0,\mathrm{I}_n)$ ，矩阵 A,B 对称幂等，且满足 $AB=BA=0$ 。证明随机变量 $\boldsymbol{Y}'A\boldsymbol{Y}$, $\boldsymbol{Y}'B\boldsymbol{Y}$, $\boldsymbol{Y}'(\mathrm{I}_n-A-A)\boldsymbol{Y}$ 独立服从于卡方分布。

证明

根据定理 2.8，只需证明 $\mathrm{rank}(A)+\mathrm{rank}(B)+\mathrm{rank}(\mathrm{I}_n-A-B)=n$ 。我们先证明 $\mathrm{rank}(A)+\mathrm{rank}(B)=\mathrm{rank}(A+B)$ 。由初等变换，知

$$\begin{aligned}\begin{pmatrix}A&0\\0&B\end{pmatrix}&\xrightarrow{\text{第一行加到第二行}}\begin{pmatrix}A&0\\A&B\end{pmatrix}\\&\xrightarrow{\text{第一列加到第二列}}\begin{pmatrix}A&A\\A&A+B\end{pmatrix}\\&\xrightarrow{\text{第二行左乘 }-A\text{ 加到第一行}}\begin{pmatrix}0&0\\A&A+B\end{pmatrix}\\&\xrightarrow{\text{第二列右乘 }-A\text{ 加到第一列}}\begin{pmatrix}0&0\\0&A+B\end{pmatrix}\end{aligned}$$

从而 $\operatorname{rank}(A)+\operatorname{rank}(B)=\operatorname{rank}(A+B)$ 。
再证明 $\operatorname{rank}(A+B)+\operatorname{rank}(\mathrm{I}_n-A-B)=n$ 。同样地，有

$$
\begin{aligned}
\begin{pmatrix} A+B & 0 \\ 0 & \mathrm{I}_n-A-B \end{pmatrix}
&\xrightarrow{\text{第一行加到第二行}} \begin{pmatrix} A+B & 0 \\ A+B & \mathrm{I}_n-A-B \end{pmatrix} \\
&\xrightarrow{\text{第一列加到第二列}} \begin{pmatrix} A+B & A+B \\ A+B & \mathrm{I}_n \end{pmatrix} \\
&\xrightarrow{\text{第二行左乘 } -(A+B) \text{ 加到第一行}} \begin{pmatrix} 0 & 0 \\ A+B & \mathrm{I}_n \end{pmatrix} \\
&\xrightarrow{\text{第二列右乘 } -(A+B) \text{ 加到第一列}} \begin{pmatrix} 0 & 0 \\ 0 & \mathrm{I}_n \end{pmatrix}
\end{aligned}
$$

从而 $\operatorname{rank}(A+B)+\operatorname{rank}(\mathrm{I}_n-A-B)=n$ 。

10. 设随机向量 $\boldsymbol{Y}=(\boldsymbol{Y_1}',\boldsymbol{Y_2}')'\sim \mathrm{N}(0,\Sigma)$ ，其中 $\boldsymbol{Y_1}$ 的维数为 m ，$\boldsymbol{Y_2}$ 的维数为 n ，Σ 可被分块为

$$
\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}
$$

证明 $\boldsymbol{Y}'\Sigma^{-1}\boldsymbol{Y}-\boldsymbol{Y_1}'\Sigma_{11}^{-1}\boldsymbol{Y}\sim\chi_n^2$ 。

证明

记 $\Sigma=(\Sigma^{\frac12})'(\Sigma^{\frac12})$ ，则 $\boldsymbol{Y}\overset{d}{=}(\Sigma^{\frac12})'\boldsymbol{U}$ ，其中 $\boldsymbol{U}\sim\mathrm{N}(0,\mathrm{I}_n)$ 。于是

$$
\begin{aligned}
\boldsymbol{Y}'\Sigma^{-1}\boldsymbol{Y}-\boldsymbol{Y_1}'\Sigma_{11}^{-1}\boldsymbol{Y} &= \boldsymbol{Y}'\left(\Sigma^{-1}-\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}\right)\boldsymbol{Y} \\
&\overset{d}{=} \boldsymbol{U}'\Sigma^{\frac12}\left(\Sigma^{-1}-\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}\right)(\Sigma^{\frac12})'\boldsymbol{U} \\
&= \boldsymbol{U}'\left(\mathrm{I}-\Sigma^{\frac12}\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}(\Sigma^{\frac12})'\right)\boldsymbol{U}
\end{aligned}
$$

要证明矩阵

$$
\mathrm{I}-\Sigma^{\frac12}\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}(\Sigma^{\frac12})'
$$

对称幂等。易见其对称，接下来，我们先证明幂等性，即

$$
\begin{aligned}
&\left(\mathrm{I}-\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\right)\left(\mathrm{I}-\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\right)\\
=&\mathrm{I}-2\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'+\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}\Sigma\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\\
=&\mathrm{I}-2\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'+\Sigma^{\frac{1}{2}}\begin{pmatrix}\mathrm{I} & \Sigma_{11}^{-1}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\\
=&\mathrm{I}-\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'
\end{aligned}
$$

其次，由上题的证明过程可以发现，对于 n 阶幂等阵 A ，有 $\mathrm{rank}(A)+\mathrm{rank}(\mathrm{I}-A)=n$ 。从而

$$
\begin{aligned}
&\mathrm{rank}\left(\mathrm{I}-\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\right)\\
=&m+n-\mathrm{rank}\left(\Sigma^{\frac{1}{2}}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(\Sigma^{\frac{1}{2}})'\right)\\
=&m+n-m=n
\end{aligned}
$$

综上，结论成立。

11. 设随机向量 $\boldsymbol{Y}\sim\mathrm{N}(0\,,\mathrm{I}_n)$ ，求随机变量

$$(Y_1-Y_2)^2+(Y_2-Y_3)^2+\cdots+(Y_{n-1}-Y_n)^2$$

的方差。

解

记矩阵

$$
A=\begin{pmatrix}
1 & -1 & 0 & \cdots & 0 & 0\\
0 & 1 & -1 & \cdots & 0 & 0\\
0 & 0 & 1 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & 0 & \cdots & 1 & -1
\end{pmatrix}
$$

则

$$(Y_1-Y_2)^2+(Y_2-Y_3)^2+\cdots+(Y_{n-1}-Y_n)^2=\boldsymbol{Y}'A'A\boldsymbol{Y}$$

由第一章习题二第 7 题的结论，知

$$\mathrm{Var}(\boldsymbol{Y}'A'A\boldsymbol{Y}) = 2\mathrm{tr}(A'AA'A) = 6n - 2$$

12. 设 n 维随机向量 $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$ ，证明

$$\mathrm{Var}(\boldsymbol{Y}'A\boldsymbol{Y}) = 2\mathrm{tr}(A\Sigma A\Sigma) + 4\boldsymbol{\mu}'A\Sigma A\boldsymbol{\mu}$$

证明

太复杂惹

13. 证明 $\chi^2_m(\delta^2)$ 随机变量的矩母函数为

$$M_{\chi^2_m(\delta^2)}(t) = (1-2t)^{-\frac{m}{2}} \exp\left\{\frac{t}{1-2t}\delta^2\right\} \qquad t < \frac{1}{2}$$

证明

设随机向量 $\boldsymbol{Z} = (Z_1, \cdots, Z_m)' \sim \mathrm{N}(\boldsymbol{\mu}, \mathrm{I}_m)$ ，其中 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)$ 且 $\|\boldsymbol{\mu}\|_2^2 = \delta^2$ ，则 $\boldsymbol{Z}'\boldsymbol{Z} \sim \chi^2_m(\delta^2)$ 。于是

$$\begin{aligned}
M_{\chi^2_m(\delta^2)}(t) &= \mathrm{E}\left(\exp\left\{t\boldsymbol{Z}'\boldsymbol{Z}\right\}\right) \\
&= \prod_{i=1}^m \mathrm{E}\left(\exp\left\{tZ_i^2\right\}\right) \\
&= \prod_{i=1}^m \int_{\mathbb{R}} \exp\left\{ty^2\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{1}{2}(y-\mu_i)^2\right\} \mathrm{d}y \\
&= \prod_{i=1}^m (1-2t)^{-\frac{1}{2}} \exp\left\{\frac{t}{1-2t}\mu_i^2\right\} \\
&= (1-2t)^{-\frac{m}{2}} \exp\left\{\frac{t}{1-2t}\delta^2\right\} \qquad t < \frac{1}{2}
\end{aligned}$$

14. 设随机向量 $\boldsymbol{X} \sim \mathrm{N}(\mu 1_n, \sigma^2 \mathrm{I}_n)$ ，$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ 。证明

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\boldsymbol{X}'\left(\mathrm{I}_n - \frac{1}{n}1_n 1_n'\right)\boldsymbol{X} \sim \chi^2_{n-1}$$

证明

只需证明矩阵 $\mathrm{I}_n - \frac{1}{n}1_n1_n'$ 是秩为 $n-1$ 的对称幂等阵。其对称性与幂等性是显然的，而秩，类似于第二章第 10 题的做法，有

$$\operatorname{rank}\left(\mathrm{I}_n - \frac{1}{n}1_n1_n'\right) = n - \operatorname{rank}\left(\frac{1}{n}1_n1_n'\right) = n-1$$

15. 设随机向量 $\boldsymbol{Y} = (\boldsymbol{Y_1}', \boldsymbol{Y_2}') \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$ ，其中

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} > 0$$

用密度函数证明

$$\boldsymbol{Y_1}|\boldsymbol{Y_2} \sim \mathrm{N}\left(\boldsymbol{\mu_1} + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{Y_2} - \boldsymbol{\mu_2}), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

证明

记 $\boldsymbol{Y_1}$ 的维数为 m ，$\boldsymbol{Y_2}$ 的维数为 n 。则 $\boldsymbol{Y}$ 的密度函数为

$$f(\boldsymbol{y_1}, \boldsymbol{y_2}) = (2\pi)^{-\frac{m+n}{2}} |\det(\Sigma)|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{y_1}' - \boldsymbol{\mu_1}', \boldsymbol{y_2}' - \boldsymbol{\mu_2}')\Sigma^{-1} \begin{pmatrix} \boldsymbol{y_1} - \boldsymbol{\mu_1} \\ \boldsymbol{y_2} - \boldsymbol{\mu_2} \end{pmatrix} \right\}$$

若记 $M = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ ，则

$$\begin{aligned} \det(\Sigma) &= \det(M) \cdot \det(\Sigma_{22}) \\ \Sigma^{-1} &= \begin{pmatrix} M^{-1} & -M^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{12}M^{-1} & \Sigma_{22}^{-1}\Sigma_{12}M^{-1}\Sigma_{12}\Sigma_{22}^{-1} + \Sigma_{22}^{-1} \end{pmatrix} \end{aligned}$$

而 $\boldsymbol{Y_2}$ 的边际密度函数为

$$f_{\boldsymbol{Y_2}}(\boldsymbol{y_2}) = (2\pi)^{-\frac{n}{2}} |\det(\Sigma_{22})|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{y_2} - \boldsymbol{\mu_2})'\Sigma_{22}^{-1}(\boldsymbol{y_2} - \boldsymbol{\mu_2}) \right\}$$

从而 $\boldsymbol{Y_1}$ 的条件密度函数为

$$\begin{aligned} f_{\boldsymbol{Y_1}|\boldsymbol{Y_2}}(\boldsymbol{y_1}|\boldsymbol{y_2}) &= \frac{f(\boldsymbol{y_1}, \boldsymbol{y_2})}{f_{\boldsymbol{Y_2}}(\boldsymbol{y_2})} \\ &= (2\pi)^{-\frac{m}{2}} |\det(M)|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left(\boldsymbol{y_1} - \boldsymbol{\mu_1} - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{y_2} - \boldsymbol{\mu_2})\right)' \right. \\ &\qquad \left. M^{-1}\left(\boldsymbol{y_1} - \boldsymbol{\mu_1} - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{y_2} - \boldsymbol{\mu_2})\right) \right\} \end{aligned}$$

即 $\boldsymbol{Y_1}|\boldsymbol{Y_2} \sim \mathrm{N}\left(\boldsymbol{\mu_1} + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{Y_2} - \boldsymbol{\mu_2}), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$ 。

16. 设随机向量 $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$ ，$\Sigma \geq 0$ ，证明

$$(\boldsymbol{Y}-\boldsymbol{\mu})'\Sigma^{-}(\boldsymbol{Y}-\boldsymbol{\mu}) \sim \chi_r^2$$

其中 $r = \mathrm{rank}(\Sigma)$ 。

证明

由题设，知存在 r 阶满秩对角阵 Λ_r ，正交阵 P ，使得

$$\Sigma = P' \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} P$$

于是 Σ^{-} 具有下面的形式

$$\Sigma^{-} = P' \begin{pmatrix} \Lambda_r^{-1} & B \\ C & D \end{pmatrix} P$$

其中 B, C, D 是任意阶数合适的矩阵。从而

$$(\boldsymbol{Y}-\boldsymbol{\mu})'\Sigma^{-}(\boldsymbol{Y}-\boldsymbol{\mu}) = (\boldsymbol{Y}-\boldsymbol{\mu})'P' \begin{pmatrix} \Lambda_r^{-1} & B \\ C & D \end{pmatrix} P(\boldsymbol{Y}-\boldsymbol{\mu})$$

记随机向量 $\boldsymbol{Z} = P(\boldsymbol{Y}-\boldsymbol{\mu})$ ，则 $\boldsymbol{Z} \sim \mathrm{N}\left(0, \mathrm{diag}(\Lambda_r, 0)\right)$ 。将 $\boldsymbol{Z}$ 分为 $(\boldsymbol{Z_1'}, \boldsymbol{Z_2'})'$ ，其中 $\boldsymbol{Z_1}$ 的维数为 r ，则 $\boldsymbol{Z_1} \sim \mathrm{N}(0, \Lambda_r)$ ，$\boldsymbol{Z_2} \overset{d}{=} 0$ ，并且

$$\begin{aligned} \boldsymbol{Z}' \begin{pmatrix} \Lambda_r^{-1} & B \\ C & D \end{pmatrix} \boldsymbol{Z} &\overset{d}{=} \boldsymbol{Z_1'}\Lambda_r^{-1}\boldsymbol{Z_1} \\ &= \left(\Lambda^{-\frac{1}{2}}\boldsymbol{Z_1}\right)' \left(\Lambda^{-\frac{1}{2}}\boldsymbol{Z_1}\right) \sim \chi_r^2 \end{aligned}$$