

第二章 随机向量

Tianxiao Pang

Zhejiang University

September 16, 2018

内容

① 均值向量与协方差矩阵

内容

- 1 均值向量与协方差矩阵
- 2 随机向量的二次型

内容

- ① 均值向量与协方差矩阵
- ② 随机向量的二次型
- ③ 正态随机向量

内容

- 1 均值向量与协方差矩阵
- 2 随机向量的二次型
- 3 正态随机向量
- 4 正态随机向量的二次型

内容

- 1 均值向量与协方差矩阵
- 2 随机向量的二次型
- 3 正态随机向量
- 4 正态随机向量的二次型
- 5 矩阵微商

在第一章我们看到, 当用矩阵形式来表示一个线性模型时, 观测向量和误差向量都是随机向量, 因此我们有必要了解随机向量, 特别是正态随机向量的一些基本性质.

在课件中, 若无特殊说明, 我们约定:

所有的矩阵和向量都是实矩阵和实向量. 大写字母 $\mathbf{A}, \mathbf{B}, \dots$ 表示矩阵或列向量; 小写字母 $\mathbf{a}, \mathbf{b}, \dots$ 表示列向量; 矩阵 \mathbf{A} 的秩记为 $\text{rk}(\mathbf{A})$; 称方阵 $\mathbf{A} = (a_{ij})_{n \times n}$ 的对角线元素之和 $\sum_{i=1}^n a_{ii}$ 为 \mathbf{A} 的迹(trace), 记为 $\text{tr}(\mathbf{A})$; 若 \mathbf{A} 为正定对称方阵, 则记为 $\mathbf{A} > \mathbf{0}$; 若 \mathbf{A} 为非负定对称方阵, 则记为 $\mathbf{A} \geq \mathbf{0}$; 而 $\mathbf{A} > \mathbf{B}$ 表示 $\mathbf{A} - \mathbf{B} > \mathbf{0}$, $\mathbf{A} \geq \mathbf{B}$ 表示 $\mathbf{A} - \mathbf{B} \geq \mathbf{0}$.

均值向量与协方差矩阵

设 $\mathbf{X} = (X_1, \dots, X_n)'$ 为 $n \times 1$ 随机向量, 称

$$\mathbf{E}(\mathbf{X}) = (\mathbf{E}(X_1), \dots, \mathbf{E}(X_n))'$$

为 \mathbf{X} 的均值向量.

定理 (2.1.1)

设 \mathbf{A} 是 $m \times n$ 非随机矩阵, \mathbf{X} 和 \mathbf{b} 分别是 $n \times 1$ 和 $m \times 1$ 随机向量, 记 $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$, 则

$$\mathbf{E}(\mathbf{Y}) = \mathbf{A}\mathbf{E}(\mathbf{X}) + \mathbf{E}(\mathbf{b}).$$

证明: 设 $\mathbf{A} = (a_{ij})$, $\mathbf{b} = (b_1, \dots, b_m)'$, $\mathbf{Y} = (Y_1, \dots, Y_m)'$. 于是

$$Y_i = \sum_{j=1}^n a_{ij} X_j + b_i, \quad i = 1, \dots, m.$$

求均值得

$$\mathbb{E}(Y_i) = \sum_{j=1}^n a_{ij} \mathbb{E}(X_j) + \mathbb{E}(b_i), \quad i = 1, \dots, m.$$

得证.

定义 n 维随机向量 \mathbf{X} 的协方差矩阵为

$$\text{Cov}(\mathbf{X}) = \text{E}[(\mathbf{X} - \text{E}(\mathbf{X}))(\mathbf{X} - \text{E}(\mathbf{X}))'].$$

推论 (2.1.1)

$$\text{tr}[\text{Cov}(\mathbf{X})] = \sum_{i=1}^n \text{Var}(X_i).$$

定理 (2.1.2)

设 \mathbf{X} 为任意的 $n \times 1$ 随机向量, 则它的协方差矩阵是非负定对称阵.

证明: 对称是显然的. 下证非负定性. 对任意的非随机 $n \times 1$ 向量 \mathbf{c} , 注意到 $\mathbf{c}'\mathbf{X}$ 是一个随机变量, 所以

$$\begin{aligned}
 0 \leq \text{Var}(\mathbf{c}'\mathbf{X}) &= \text{E}[(\mathbf{c}'\mathbf{X} - \text{E}(\mathbf{c}'\mathbf{X}))^2] \\
 &= \text{E}[(\mathbf{c}'\mathbf{X} - \text{E}(\mathbf{c}'\mathbf{X}))(\mathbf{c}'\mathbf{X} - \text{E}(\mathbf{c}'\mathbf{X}))'] \\
 &= \mathbf{c}'\text{E}[(\mathbf{X} - \text{E}(\mathbf{X}))(\mathbf{X} - \text{E}(\mathbf{X}))']\mathbf{c} \\
 &= \mathbf{c}'\text{Cov}(\mathbf{X})\mathbf{c}.
 \end{aligned}$$

得证.

定理 (2.1.3)

设 \mathbf{A} 是 $m \times n$ 矩阵, \mathbf{X} 是 $n \times 1$ 随机向量, $\mathbf{Y} = \mathbf{A}\mathbf{X}$, 则

$$\text{Cov}(\mathbf{Y}) = \mathbf{A} \text{Cov}(\mathbf{X}) \mathbf{A}'.$$

证明: 根据协方差矩阵的定义,

$$\begin{aligned} \text{Cov}(\mathbf{Y}) &= \text{E}[(\mathbf{Y} - \text{E}(\mathbf{Y}))(\mathbf{Y} - \text{E}(\mathbf{Y}))'] \\ &= \text{E}[(\mathbf{A}\mathbf{X} - \text{E}(\mathbf{A}\mathbf{X}))(\mathbf{A}\mathbf{X} - \text{E}(\mathbf{A}\mathbf{X}))'] \\ &= \mathbf{A} \text{E}[(\mathbf{X} - \text{E}(\mathbf{X}))(\mathbf{X} - \text{E}(\mathbf{X}))'] \mathbf{A}' \\ &= \mathbf{A} \text{Cov}(\mathbf{X}) \mathbf{A}'. \end{aligned}$$

设 \mathbf{X} 和 \mathbf{Y} 分别是 $n \times 1$ 和 $m \times 1$ 随机向量, 我们定义

$$\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{E}[(\mathbf{X} - \mathbf{E}(\mathbf{X}))(\mathbf{Y} - \mathbf{E}(\mathbf{Y}))'].$$

定理 (2.1.4)

设 \mathbf{X} 和 \mathbf{Y} 分别是 $n \times 1$ 和 $m \times 1$ 随机向量, \mathbf{A} 和 \mathbf{B} 分别是 $p \times n$ 和 $q \times m$ 非随机矩阵, 则

$$\text{Cov}(\mathbf{AX}, \mathbf{BY}) = \mathbf{ACov}(\mathbf{X}, \mathbf{Y})\mathbf{B}'.$$

证明:

$$\begin{aligned} \text{Cov}(\mathbf{AX}, \mathbf{BY}) &= \mathbf{E}[(\mathbf{AX} - \mathbf{E}(\mathbf{AX}))(\mathbf{BY} - \mathbf{E}(\mathbf{BY}))'] \\ &= \mathbf{AE}[(\mathbf{X} - \mathbf{E}(\mathbf{X}))(\mathbf{Y} - \mathbf{E}(\mathbf{Y}))']\mathbf{B}' \\ &= \mathbf{ACov}(\mathbf{X}, \mathbf{Y})\mathbf{B}'. \end{aligned}$$

随机向量的二次型

假设 $\mathbf{X} = (X_1, \dots, X_n)'$ 为 $n \times 1$ 随机向量, $\mathbf{A} = (a_{ij})$ 为 $n \times n$ 对称矩阵, 则称随机变量

$$\mathbf{X}'\mathbf{A}\mathbf{X} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} X_i X_j$$

为 \mathbf{X} 的二次型.

定理 (2.2.1)

设 $E(\mathbf{X}) = \boldsymbol{\mu}$, $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$, 则

$$E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} + \text{tr}(\mathbf{A}\boldsymbol{\Sigma}). \quad (2.2.1)$$

证明: 因为

$$\begin{aligned} \mathbf{X}'\mathbf{A}\mathbf{X} &= (\mathbf{X} - \boldsymbol{\mu} + \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu} + \boldsymbol{\mu}) \\ &= (\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) + \boldsymbol{\mu}'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) \\ &\quad + (\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}\boldsymbol{\mu} + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}. \end{aligned}$$

由定理2.1.1, 上式第二项与第三项的数学期望为零. 因此, 为证(2.2.1), 只需证明

$$\mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})] = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}).$$

利用迹的性质($\text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$)以及求迹和求期望可交换次序), 可知

$$\begin{aligned} \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})] &= \mathbb{E}[\text{tr}((\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}))] \\ &= \mathbb{E}\{\text{tr}[\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})']\} \\ &= \text{tr}\{\mathbf{A}\mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})']\} \\ &= \text{tr}(\mathbf{A}\boldsymbol{\Sigma}). \end{aligned}$$

补充:

$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ 的证明: 假设 \mathbf{A} 是 $n \times m$ 矩阵, \mathbf{B} 是 $m \times n$ 矩阵, 那么

$$\begin{aligned}
 \text{tr}(\mathbf{AB}) &= \sum_{i=1}^n (\mathbf{AB})_{ii} \\
 &= \sum_{i=1}^n \sum_{j=1}^m a_{ij} b_{ji} \\
 &= \sum_{j=1}^m \sum_{i=1}^n b_{ji} a_{ij} \\
 &= \sum_{j=1}^m (\mathbf{BA})_{jj} \\
 &= \text{tr}(\mathbf{BA}).
 \end{aligned}$$

推论 (2.2.1)

(1) 若 $\mu = \mathbf{0}$, 则 $E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{tr}(\mathbf{A}\Sigma)$.

(2) 若 $\Sigma = \sigma^2 \mathbf{I}$, 则

$$E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \mu' \mathbf{A} \mu + \sigma^2 \text{tr}(\mathbf{A}) = \mu' \mathbf{A} \mu + \sigma^2 \sum_{i=1}^n a_{ii}.$$

(3) 若 $\mu = \mathbf{0}, \Sigma = \mathbf{I}$, 则 $E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{tr}(\mathbf{A})$.

例2.2.1 假设一维总体的均值为 μ , 方差为 σ^2 , X_1, \dots, X_n 为从总体中抽取的一个样本, 试求样本方差 S^2 的均值,

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

解: 记 $Q = (n-1)S^2$. 我们把 Q 表示成 $\mathbf{X} = (X_1, \dots, X_n)'$ 的一个二次型. 用 $\mathbf{1}_n$ 表示所有元素为1的 n 维向量. 则

$$E(\mathbf{X}) = \mu \mathbf{1}_n, \text{Cov}(\mathbf{X}) = \sigma^2 \mathbf{I}_n, \bar{X} = \frac{1}{n} \mathbf{1}_n' \mathbf{X},$$

$$\mathbf{X} - \bar{X} \mathbf{1}_n = \mathbf{X} - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n' \mathbf{X} = (\mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n') \mathbf{X} =: \mathbf{C} \mathbf{X}.$$

其中 $C = I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n'$ 是一个对称幂等矩阵. 于是

$$Q = \sum_{i=1}^n (X_i - \bar{X})^2 = (\mathbf{X} - \bar{X} \mathbf{1}_n)' (\mathbf{X} - \bar{X} \mathbf{1}_n) = \mathbf{X}' C \mathbf{X}.$$

应用定理2.2.1得

$$E(Q) = E(\mathbf{X})' \cdot C \cdot E(\mathbf{X}) + \sigma^2 \text{tr}(C) = \mu^2 \mathbf{1}_n' C \mathbf{1}_n + \sigma^2 \text{tr}(C).$$

又易知

$$C \mathbf{1}_n = \mathbf{0}, \text{tr}(C) = n - 1,$$

所以 $E(Q) = (n - 1)\sigma^2$, 即 $E(S^2) = \sigma^2$.

现在我们来推导二次型 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 的方差公式.

定理 (2.2.2)

设随机变量 $X_i, i = 1, \dots, n$ 相互独立, 且 $E(X_i) = \mu_i$. 假设 $X_i - \mu_i, i = 1, \dots, n$ 是同分布的且

$$\text{Var}(X_i) = \sigma^2, m_r = E(X_i - \mu_i)^r, r = 3, 4.$$

$\mathbf{A} = (a_{ij})$ 为 $n \times n$ 对称矩阵. 记

$$\mathbf{X} = (X_1, \dots, X_n)', \quad \boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'.$$

则

$$\text{Var}(\mathbf{X}'\mathbf{A}\mathbf{X}) = (m_4 - 3\sigma^4)\mathbf{a}'\mathbf{a} + 2\sigma^4 \text{tr}(\mathbf{A}^2) + 4\sigma^2 \boldsymbol{\mu}'\mathbf{A}^2\boldsymbol{\mu} + 4m_3 \boldsymbol{\mu}'\mathbf{A}\mathbf{a},$$

其中 $\mathbf{a} = (a_{11}, \dots, a_{nn})'$, 即 \mathbf{A} 的对角元组成的列向量.

证明: 首先注意到

$$\text{Var}(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{E}(\mathbf{X}'\mathbf{A}\mathbf{X})^2 - [\text{E}(\mathbf{X}'\mathbf{A}\mathbf{X})]^2, \quad (2.2.2)$$

由定理2.2.1以及 $\text{E}(\mathbf{X}) = \boldsymbol{\mu}$ 和 $\text{Cov}(\mathbf{X}) = \sigma^2 \mathbf{I}_n$ 可推得

$$\text{E}(\mathbf{X}'\mathbf{A}\mathbf{X}) = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} + \sigma^2 \text{tr}(\mathbf{A}). \quad (2.2.3)$$

所以我们的问题主要是计算(2.2.2)中的第一项. 将 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 改写为

$$\mathbf{X}'\mathbf{A}\mathbf{X} = (\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) + 2\boldsymbol{\mu}'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu},$$

将其平方得到

$$\begin{aligned} (\mathbf{X}'\mathbf{A}\mathbf{X})^2 &= [(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})]^2 + 4[\boldsymbol{\mu}'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})]^2 + (\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu})^2 \\ &\quad + 2\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}[(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) + 2\boldsymbol{\mu}'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})] \\ &\quad + 4\boldsymbol{\mu}'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}). \end{aligned}$$

令 $\mathbf{Z} = \mathbf{X} - \boldsymbol{\mu}$, 则 $E(\mathbf{Z}) = \mathbf{0}$. 再次利用定理2.2.1得

$$\begin{aligned} E(\mathbf{X}'\mathbf{A}\mathbf{X})^2 &= E(\mathbf{Z}'\mathbf{A}\mathbf{Z})^2 + 4E(\boldsymbol{\mu}'\mathbf{A}\mathbf{Z})^2 + (\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu})^2 \\ &\quad + 2\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}(\sigma^2\text{tr}(\mathbf{A})) + 4E(\boldsymbol{\mu}'\mathbf{A}\mathbf{Z}\mathbf{Z}'\mathbf{A}\mathbf{Z}). \quad (2.2.4) \end{aligned}$$

下面逐个计算上式所含的每个均值. 由

$$(\mathbf{Z}'\mathbf{A}\mathbf{Z})^2 = \sum_i \sum_j \sum_k \sum_l a_{ij}a_{kl}Z_iZ_jZ_kZ_l$$

及 Z_i 的独立性导出的事实

$$E(Z_iZ_jZ_kZ_l) = \begin{cases} m_4, & \text{若 } i = j = k = l, \\ \sigma^4, & \text{若 } i = j \neq k = l; i = k \neq j = l; i = l \neq j = k, \\ 0, & \text{其它,} \end{cases}$$

便有下列结果:

$$\begin{aligned}
E(\mathbf{Z}'\mathbf{A}\mathbf{Z})^2 &= m_4\left(\sum_{i=1}^n a_{ii}^2\right) + \sigma^4\left(\sum_{i \neq k} a_{ii}a_{kk} + \sum_{i \neq j} a_{ij}^2 + \sum_{i \neq j} a_{ij}a_{ji}\right) \\
&= m_4\mathbf{a}'\mathbf{a} + \sigma^4\left\{[\text{tr}(\mathbf{A})]^2 - \mathbf{a}'\mathbf{a} + 2[\text{tr}(\mathbf{A}^2) - \mathbf{a}'\mathbf{a}]\right\} \\
&= (m_4 - 3\sigma^4)\mathbf{a}'\mathbf{a} + \sigma^4\{[\text{tr}(\mathbf{A})]^2 + 2\text{tr}(\mathbf{A}^2)\}, \quad (2.2.5)
\end{aligned}$$

而

$$\begin{aligned}
E(\boldsymbol{\mu}'\mathbf{A}\mathbf{Z})^2 &= E(\mathbf{Z}'\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{A}\mathbf{Z}) \\
&= \sigma^2 \cdot \text{tr}(\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{A}) \\
&= \sigma^2 \cdot \boldsymbol{\mu}'\mathbf{A}^2\boldsymbol{\mu}. \quad (2.2.6)
\end{aligned}$$

最后, 若记 $\mathbf{b} = \mathbf{A}\boldsymbol{\mu}$, 则

$$\mathbf{E}(\boldsymbol{\mu}' \mathbf{A} \mathbf{Z} \mathbf{Z}' \mathbf{A} \mathbf{Z}) = \sum_i \sum_j \sum_k b_i a_{jk} \mathbf{E}(Z_i Z_j Z_k).$$

因为

$$\mathbf{E}(Z_i Z_j Z_k) = \begin{cases} m_3, & \text{若 } i = j = k, \\ 0, & \text{其它,} \end{cases}$$

所以

$$\mathbf{E}(\boldsymbol{\mu}' \mathbf{A} \mathbf{Z} \mathbf{Z}' \mathbf{A} \mathbf{Z}) = m_3 \sum_i b_i a_{ii} = m_3 \mathbf{b}' \mathbf{a} = m_3 \boldsymbol{\mu}' \mathbf{A} \mathbf{a}. \quad (2.2.7)$$

将(2.2.5)-(2.2.7)代入(2.2.4), 再将(2.2.3)和(2.2.4)代入(2.2.2)便可得到需要证明的结果.

正态随机向量

定义

设 n 维随机向量 $\mathbf{X} = (X_1, \dots, X_n)'$ 具有密度函数

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (2.3.1)$$

其中 $\mathbf{x} = (x_1, \dots, x_n)'$, $-\infty < x_i < \infty$, $i = 1, \dots, n$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$, Σ 是对称正定矩阵, 则称 \mathbf{X} 为 n 维正态随机向量, 记为 $N_n(\boldsymbol{\mu}, \Sigma)$ 或 $N(\boldsymbol{\mu}, \Sigma)$.

事实上, μ 和 Σ 分别是 \mathbf{X} 的均值向量和协方差矩阵. 记 $\Sigma^{1/2}$ 为 Σ 的平方根阵, $\Sigma^{-1/2}$ 为 $\Sigma^{1/2}$ 的逆矩阵. 定义

$$\mathbf{Y} = \Sigma^{-1/2}(\mathbf{X} - \mu),$$

则有 $\mathbf{X} = \Sigma^{1/2}\mathbf{Y} + \mu$. 于是 \mathbf{Y} 的密度函数为

$$g(\mathbf{y}) = f(\Sigma^{1/2}\mathbf{y} + \mu)|J|,$$

其中 J 为向量变换的 Jacobi 行列式:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = \det(\Sigma^{1/2}) = (\det \Sigma)^{1/2}.$$

所以

$$g(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{1}{2} \mathbf{y}' \mathbf{y} \right\} = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-y_i^2/2}.$$

这表明 \mathbf{Y} 的 n 个分量相互独立, 服从 $N(0, 1)$. 因此

$$\mathbf{E}(\mathbf{Y}) = \mathbf{0}, \text{Cov}(\mathbf{Y}) = \mathbf{I}_n,$$

由此可知

$$\mathbf{E}(\mathbf{X}) = \boldsymbol{\mu}, \text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}.$$

注: 称 $N(\mathbf{0}, \mathbf{I})$ 为多元标准正态分布.

设 \mathbf{X} 的协方差矩阵具有如下分块对角矩阵形式:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} \end{pmatrix}, \quad (2.3.2)$$

这里 Σ_{11} 为 $m \times m$ 矩阵. 将 \mathbf{X} 和 $\boldsymbol{\mu}$ 也分块为

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad (2.3.3)$$

这里 \mathbf{X}_1 和 $\boldsymbol{\mu}_1$ 均为 $m \times 1$ 向量. 则(2.3.1)可写为

$$f(\mathbf{x}) = f_1(\mathbf{x}_1)f_2(\mathbf{x}_2),$$

其中

$$f_1(\mathbf{x}_1) = \frac{1}{(2\pi)^{m/2}(\det \boldsymbol{\Sigma}_{11})^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) \right\},$$

$$f_2(\mathbf{x}_2) = \frac{1}{(2\pi)^{(n-m)/2}(\det \boldsymbol{\Sigma}_{22})^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right\}.$$

这表明 $\mathbf{X}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii}), i = 1, 2$, 且相互独立.

定理 (2.3.1)

- (a) 设 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 且 \mathbf{X} 和 $\boldsymbol{\mu}$ 分别具有分块形式(2.3.3), 而 $\boldsymbol{\Sigma}$ 具有分块形式(2.3.2), 则 $\mathbf{X}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii}), i = 1, 2$, 且相互独立.
- (b) 若 $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_n$, 且 $\mathbf{X} = (X_1, \cdots, X_n)'$, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)'$, 则 $X_i \sim N(\mu_i, \sigma^2), i = 1, \cdots, n$, 且相互独立.

注: 对多元正态分布来说, \mathbf{X}_1 与 \mathbf{X}_2 不相关($\text{Cov}(\mathbf{X}_1, \mathbf{X}_2) = \boldsymbol{\Sigma}_{12} = \mathbf{0}$), 可以推出 \mathbf{X}_1 与 \mathbf{X}_2 独立.

定理 (2.3.2)

设 n 维随机向量 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} 为 $n \times n$ 非随机可逆矩阵,
 \mathbf{b} 为 $n \times 1$ 向量, 记 $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$, 则

$$\mathbf{Y} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

证明: $\mathbf{X} = \mathbf{A}^{-1}(\mathbf{Y} - \mathbf{b})$, 所以 \mathbf{Y} 的密度函数为

$$g(\mathbf{y}) = f(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}))|J|,$$

J 为变换的Jacobi行列式:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = (\det \mathbf{A})^{-1}.$$

注意到

$$(\det \boldsymbol{\Sigma})^{1/2} |J|^{-1} = (\det \boldsymbol{\Sigma} (\det \mathbf{A})^2)^{1/2} = (\det(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'))^{1/2},$$

$$\begin{aligned}
& (\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}) - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}) - \boldsymbol{\mu}) \\
= & (\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))' (\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')^{-1} (\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b})),
\end{aligned}$$

所以

$$\begin{aligned}
g(\mathbf{y}) &= \frac{1}{(2\pi)^{n/2} (\det \boldsymbol{\Sigma})^{1/2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2} (\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}) - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}) - \boldsymbol{\mu}) \right\} |J| \\
&= \frac{1}{(2\pi)^{n/2} (\det(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'))^{1/2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2} (\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))' (\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')^{-1} (\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b})) \right\}.
\end{aligned}$$

这正是 $N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$ 的密度函数. 得证.

由于 $\Sigma^{-1/2}\Sigma(\Sigma^{-1/2})' = \mathbf{I}$, 所以我们有

推论 (2.3.1)

设 $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$, 则

$$\mathbf{Y} = \Sigma^{-1/2}\mathbf{X} \sim N(\Sigma^{-1/2}\boldsymbol{\mu}, \mathbf{I}), \quad \mathbf{Z} = \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \mathbf{I}).$$

这个推论表明: 我们可以用一个线性变换把诸分量相关且方差不等的多元正态随机向量变换为多元标准正态随机向量.

推论 (2.3.2)

设 $\mathbf{X} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$, \mathbf{Q} 为 $n \times n$ 正交阵, 则

$$\mathbf{Q}\mathbf{X} \sim N(\mathbf{Q}\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n).$$

这个推论表明: 诸分量相互独立且具有等方差的正态随机向量, 经过正交变换后, 变为诸分量仍然相互独立且具有等方差的正态随机向量.

定理 (2.3.3)

设 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 将 $\mathbf{X}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$ 分块为

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

其中 \mathbf{X}_1 和 $\boldsymbol{\mu}_1$ 为 $m \times 1$ 向量. 而 $\boldsymbol{\Sigma}_{11}$ 为 $m \times m$ 矩阵. 则

$$\mathbf{X}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}).$$

证明: 在定理2.3.2中取

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_{n-m} \end{pmatrix}, \mathbf{b} = \mathbf{0},$$

则 $\mathbf{Y} = \mathbf{A}\mathbf{X} \sim N(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$. 由于

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22.1} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}.$$

于是

$$\begin{aligned} \mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_1 \end{pmatrix} \\ &\sim N\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22.1} \end{pmatrix}\right). \end{aligned}$$

由定理2.3.1(a)知 $\mathbf{X}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$. 得证.

类似地, 可证明 $\mathbf{X}_2 \sim N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$, 以及更一般的结论: 对任意的 $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$(X_{i_1}, \cdots, X_{i_k})' \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0),$$

这里

$$\boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{i_1} \\ \vdots \\ \mu_{i_k} \end{pmatrix}, \quad \boldsymbol{\Sigma}_0 = \begin{pmatrix} \sigma_{i_1 i_1} & \cdots & \sigma_{i_1 i_k} \\ \vdots & \ddots & \vdots \\ \sigma_{i_k i_1} & \cdots & \sigma_{i_k i_k} \end{pmatrix}.$$

即正态随机向量的任意维数的子向量仍是正态随机向量.

下列定理是定理2.3.2的改进版本.

定理 (2.3.4)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} 是 $m \times n$ 矩阵, 且秩为 $m(< n)$, 则

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \sim N_m(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

证明: 将 \mathbf{A} 扩充为 $n \times n$ 可逆矩阵

$$\mathbf{C} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}.$$

应用定理2.3.2得

$$\mathbf{Z} = \mathbf{C}\mathbf{X} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \mathbf{X} \sim N\left(\begin{pmatrix} \mathbf{A}\boldsymbol{\mu} \\ \mathbf{B}\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' & \mathbf{A}\boldsymbol{\Sigma}\mathbf{B}' \\ \mathbf{B}\boldsymbol{\Sigma}\mathbf{A}' & \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}' \end{pmatrix}\right).$$

再应用定理2.3.3知 $\mathbf{Y} = \mathbf{A}\mathbf{X} \sim N_m(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$. 得证.

推论 (2.3.3)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{c} 是 $n \times 1$ 非零向量, 则

$$\mathbf{c}'\mathbf{X} \sim N(\mathbf{c}'\boldsymbol{\mu}, \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}).$$

即多元正态随机向量的任意非退化线性组合是一元正态随机变量.

例2.3.2 设 X_1, \dots, X_n 为从 $N(\mu, \sigma^2)$ 中抽取的一个样本. 则

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \mathbf{c}'\mathbf{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

其中 $\mathbf{X} = (X_1, \dots, X_n)'$, $\mathbf{c} = (\frac{1}{n}, \dots, \frac{1}{n})'$.

推论 (2.3.4)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$, $\boldsymbol{\Sigma} = (\sigma_{ij})$, 则

$$X_i \sim N(\mu_i, \sigma_{ii}), \quad i = 1, \dots, n.$$

即多维正态随机向量的任一分量为正态随机变量. 反之不成立(略).

正态随机向量的二次型

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} 为 $n \times n$ 实对称矩阵. 本节的目的是研究二次型 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 的性质. 下面我们总假定 $\boldsymbol{\Sigma} > \mathbf{0}$.

定理 (2.4.1)

(1) 设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} 为 $n \times n$ 实对称矩阵, 则

$$\text{Var}(\mathbf{X}'\mathbf{A}\mathbf{X}) = 2\text{tr}(\mathbf{A}\boldsymbol{\Sigma})^2 + 4\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu};$$

(2) 设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$, \mathbf{A} 为 $n \times n$ 实对称矩阵, 则

$$\text{Var}(\mathbf{X}'\mathbf{A}\mathbf{X}) = 2\sigma^4 \text{tr}(\mathbf{A}^2) + 4\sigma^2 \boldsymbol{\mu}'\mathbf{A}^2\boldsymbol{\mu}.$$

证明: (1) 记 $\mathbf{Y} = \Sigma^{-1/2}\mathbf{X}$, 则 $\mathbf{Y} \sim N(\Sigma^{-1/2}\boldsymbol{\mu}, \mathbf{I}_n)$, 所以 \mathbf{Y} 的各分量相互独立, 且

$$\text{Var}(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{Var}(\mathbf{Y}'\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}\mathbf{Y}).$$

注意到

$$m_3 = \mathbb{E}[Y_i - \mathbb{E}(Y_i)]^3 = 0, \quad m_4 = \mathbb{E}[Y_i - \mathbb{E}(Y_i)]^4 = 3.$$

应用定理2.2.2便可得到第一条结论.

(2) 这是(1)的特殊情况, 易证.

定义 (2.4.1)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$. 随机变量 $Y = \mathbf{X}'\mathbf{X}$ 的分布称为自由度为 n , 非中心参数为 $\lambda = \boldsymbol{\mu}'\boldsymbol{\mu}$ 的 χ^2 分布, 记为 $Y \sim \chi^2(n, \lambda)$. 当 $\lambda = 0$ 时, 称 Y 的分布为中心 χ^2 分布, 记为 $Y \sim \chi^2(n)$.

χ^2 分布具有下述性质:

定理 (2.4.2)

(1) 可加性: 设 $Y_i \sim \chi^2(n_i, \lambda_i), i = 1, \dots, k$, 且相互独立, 则

$$Y_1 + \dots + Y_k \sim \chi^2(n, \lambda),$$

这里 $n = \sum_{i=1}^k n_i, \lambda = \sum_{i=1}^k \lambda_i$.

(2) $E(\chi^2(n, \lambda)) = n + \lambda, \text{Var}(\chi^2(n, \lambda)) = 2n + 4\lambda$.

证明: (1) 可用特征函数方法证明.

(2) 设 $Y \sim \chi^2(n, \lambda)$, 依定义,

$$Y \stackrel{d}{=} X_1^2 + \cdots + X_{n-1}^2 + X_n^2,$$

其中 $X_i \sim N(0, 1), i = 1, \cdots, n-1, X_n \sim N(\sqrt{\lambda}, 1)$, 且相互独立. 于是

$$E(Y) = \sum_{i=1}^n E(X_i^2), \text{Var}(Y) = \sum_{i=1}^n \text{Var}(X_i^2). \quad (2.4.1)$$

因为

$$E(X_i^2) = \text{Var}(X_i) + [E(X_i)]^2 = \begin{cases} 1, & i = 1, \cdots, n-1, \\ 1 + \lambda, & i = n, \end{cases}$$

所以 $E(\chi^2(n, \lambda)) = E(Y) = n + \lambda$.

此外, 经简单计算可得

$$E(X_i^4) = 3, i = 1, \dots, n-1; E(X_n^4) = \lambda^2 + 6\lambda + 3.$$

于是有

$$\text{Var}(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = 3 - 1 = 2, \quad i = 1, \dots, n-1,$$

$$\text{Var}(X_n^2) = E(X_n^4) - [E(X_n^2)]^2 = 2 + 4\lambda.$$

把上述结果代入(2.4.1)即可证明第二条结论.

推论 (2.4.1)

设 $\mathbf{X} \sim N_n(\mathbf{0}, \Sigma)$, Σ 为正定矩阵, 则 $\mathbf{X}'\Sigma^{-1}\mathbf{X} \sim \chi^2(n)$.

证明: 记 $\mathbf{Y} = \Sigma^{-1/2}\mathbf{X}$, 则可知 $\mathbf{Y} \sim N_n(\mathbf{0}, \mathbf{I}_n)$. 又

$$\mathbf{X}'\Sigma^{-1}\mathbf{X} = (\Sigma^{-1/2}\mathbf{X})'\Sigma^{-1/2}\mathbf{X} = \mathbf{Y}'\mathbf{Y},$$

所以 $\mathbf{X}'\Sigma^{-1}\mathbf{X} \sim \chi^2(n)$.

推论 (2.4.2)

设 $X \sim \chi^2(n)$, 则 $E(X) = n$, $Var(X) = 2n$.

推论 (2.4.3)

设 X_1, \dots, X_k 相互独立, 且 $X_i \sim \chi^2(n_i)$, $i = 1, \dots, k$. 则

$$X_1 + \dots + X_k \sim \chi^2(n_1 + \dots + n_k).$$

对于正态随机向量的一般二次型, 我们有下面的定理.

定理 (2.4.3)

$\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, \mathbf{A} 为 $n \times n$ 实对称矩阵, 则 $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2(r, \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu})$
 $\Leftrightarrow \mathbf{A}$ 幂等且 $\text{rk}(\mathbf{A}) = r$.

证明: 先证充分性. 设 \mathbf{A} 幂等、对称且 $\text{rk}(\mathbf{A}) = r$. 因为对称幂等矩阵的特征根只能为0或1(易证), 于是存在正交矩阵 \mathbf{Q} 使得

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

令 $\mathbf{Y} = \mathbf{Q}'\mathbf{X}$, 则 $\mathbf{Y} \sim N_n(\mathbf{Q}'\boldsymbol{\mu}, \mathbf{I}_n)$. 对 \mathbf{Y} 和 \mathbf{Q}' 做分块

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}, \quad \mathbf{Q}' = \begin{pmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{pmatrix},$$

其中 $\mathbf{Y}_1 : r \times 1$, $\mathbf{Q}_1 : r \times n$. 于是

$$\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{Y}' \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{Y} = \mathbf{Y}_1' \mathbf{Y}_1 \sim \chi^2(r, \lambda),$$

其中 $\lambda = (\mathbf{Q}_1\boldsymbol{\mu})'\mathbf{Q}_1\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{Q}_1'\mathbf{Q}_1\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

再证必要性. 设 $\text{rk}(\mathbf{A}) = t$. 因 \mathbf{A} 对称, 故存在正交矩阵 \mathbf{Q} 使得

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

其中 $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_t)$, $\lambda_1, \dots, \lambda_t$ 非零. 我们只需证明

$$\lambda_i = 1, i = 1, \dots, t, \text{ 且 } t = r.$$

令 $\mathbf{Y} = \mathbf{Q}'\mathbf{X}$, 则 $\mathbf{Y} \sim N_n(\mathbf{Q}'\boldsymbol{\mu}, \mathbf{I}_n)$. 记

$$\mathbf{c} = \mathbf{Q}'\boldsymbol{\mu} = (c_1, \dots, c_n)',$$

则可得

$$\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{Y}' \begin{pmatrix} \boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{Y} = \sum_{j=1}^t \lambda_j Y_j^2, \quad (2.4.2)$$

这里 $\mathbf{Y} = (Y_1, \dots, Y_t)'$, $Y_j \sim N(c_j, 1), j = 1, \dots, t$, 且相互独立. 依特征函数的定义可算出 $\lambda_j Y_j^2$ 的特征函数为

$$g_j(z) = (1 - 2i\lambda_j z)^{-1/2} \exp \left\{ \frac{i\lambda_j z}{1 - 2i\lambda_j z} c_j^2 \right\}.$$

利用独立随机变量之和的特征函数等于它们的特征函数之积, 由(2.4.2)得 $\mathbf{X}' \mathbf{A} \mathbf{X}$ 的特征函数为

$$\prod_{j=1}^t (1 - 2i\lambda_j z)^{-1/2} \exp \left\{ \frac{i\lambda_j z}{1 - 2i\lambda_j z} c_j^2 \right\}. \quad (2.4.3)$$

再来计算 $\chi^2(r, \lambda)$ 的特征函数. 设 $u \sim \chi^2(r, \lambda)$, $\lambda = \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}$. 记 $u = u_1^2 + u_2^2 + \dots + u_r^2$, 其中 $u_j \sim N(0, 1), j \leq r-1, u_r \sim N(\sqrt{\lambda}, 1)$. 不难得到 u 的特征函数为

$$(1 - 2iz)^{-r/2} \exp \left\{ \frac{i\lambda z}{1 - 2iz} \right\}. \quad (2.4.4)$$

依假设, $\mathbf{X}' \mathbf{A} \mathbf{X} \sim \chi^2(r, \lambda)$, 于是(2.4.3)与(2.4.4)应该相等. 比较两者的奇点和个数可知, $\lambda_j = 1, j = 1, \dots, t$ 且 $t = r$. 必要性得证.

补充:

对称幂等矩阵的特征根只能为0或1的证明: 设 \mathbf{A} 是 $n \times n$ 对称幂等矩阵. 由对称性可知, 存在正交矩阵 \mathbf{Q} 使得

$$\mathbf{A} = \mathbf{Q} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{Q}',$$

其中 $\lambda_i, i = 1, \dots, n$ 为 \mathbf{A} 的特征根. 另一方面, 由幂等性可知

$$\mathbf{A} = \mathbf{Q} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{Q}' = \mathbf{Q} \text{diag}(\lambda_1^2, \dots, \lambda_n^2) \mathbf{Q}' = \mathbf{A}^2.$$

所以 $\lambda_i = \lambda_i^2, i = 1, \dots, n$, 即 λ_i 非0即1.

推论 (2.4.4)

设 \mathbf{A} 为 $n \times n$ 对称矩阵, $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{I}_n)$, 那么 $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2(k)$ (即中心 χ^2 分布) $\Leftrightarrow \mathbf{A}$ 幂等, $rk(\mathbf{A}) = k$, $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$.

推论 (2.4.5)

设 \mathbf{A} 为 $n \times n$ 对称矩阵, $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_n)$, 那么 $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2(k) \Leftrightarrow \mathbf{A}$ 幂等, $rk(\mathbf{A}) = k$.

推论 (2.4.6)

设 \mathbf{A} 为 $n \times n$ 对称矩阵, $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$, 那么 $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2(k, \lambda)$, $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} \Leftrightarrow \mathbf{A}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{A}$, $rk(\mathbf{A}) = k$.

定理2.4.3以及推论把判定正态随机向量的二次型服从 χ^2 分布的问题转化为研究相应的二次型矩阵的问题, 而后者往往容易处理. 因此, 这些结果是判定 χ^2 分布很有效的工具.

例2.4.1 设 X_1, \dots, X_n 为来自 $N(\mu, \sigma^2)$ 的样本, 例2.2.1已证明

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

为 σ^2 的无偏估计, 下证

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

证明: 因为 $\mathbf{X} = (X_1, \dots, X_n)' \sim N_n(\mu \mathbf{1}_n, \sigma^2 \mathbf{I}_n)$, 所以

$$\mathbf{Y} = \frac{\mathbf{X} - \mu \mathbf{1}_n}{\sigma} \sim N(\mathbf{0}, \mathbf{I}_n).$$

记 $\mathbf{C} = \mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n'$, 则易知 \mathbf{C} 是对称幂等矩阵, $\text{rk}(\mathbf{C}) = \text{tr}(\mathbf{C}) = n-1$. 又 $(\mu \mathbf{1}_n)' \mathbf{C} (\mu \mathbf{1}_n) = 0$, 所以由推论2.4.5,

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\mathbf{X}' \mathbf{C} \mathbf{X}}{\sigma^2} = \mathbf{Y}' \mathbf{C} \mathbf{Y} \sim \chi^2(n-1).$$

补充:

若 \mathbf{A} 为对称幂等矩阵, $\text{rk}(\mathbf{A}) = \text{tr}(\mathbf{A})$ 的证明: 设 $\text{rk}(\mathbf{A}) = r$, 那么存在正交矩阵 \mathbf{Q} 使得

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

利用性质 $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, 有

$$\text{tr}(\mathbf{A}) = \text{tr}\left(\mathbf{Q} \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{Q}'\right) = \text{tr}\left(\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\right) = r.$$

所以 \mathbf{A} 为对称幂等矩阵时, $\text{rk}(\mathbf{A}) = \text{tr}(\mathbf{A})$.

定理 (2.4.4)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, \mathbf{A} 为 $n \times n$ 实对称矩阵, $\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{X}'\mathbf{A}_1\mathbf{X} + \mathbf{X}'\mathbf{A}_2\mathbf{X} \sim \chi^2(r, \lambda)$, $\mathbf{X}'\mathbf{A}_1\mathbf{X} \sim \chi^2(s, \lambda_1)$, $\mathbf{A}_2 = \mathbf{A} - \mathbf{A}_1 \geq \mathbf{0}$, 其中 $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$, $\lambda_1 = \boldsymbol{\mu}'\mathbf{A}_1\boldsymbol{\mu}$. 则

- (1) $\mathbf{X}'\mathbf{A}_2\mathbf{X} \sim \chi^2(r-s, \lambda_2)$, $\lambda_2 = \boldsymbol{\mu}'\mathbf{A}_2\boldsymbol{\mu}$;
- (2) $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 与 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 相互独立;
- (3) $\mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$.

证明: 因为 $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2(r, \lambda)$, 由定理2.4.3知 \mathbf{A} 幂等且 $\text{rk}(\mathbf{A}) = r$. 于是, 存在 $n \times n$ 正交阵 \mathbf{P} 使得

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \quad (2.4.5)$$

因为 $\mathbf{A} - \mathbf{A}_1 \geq \mathbf{0}$, $\mathbf{A} - \mathbf{A}_2 \geq \mathbf{0}$ (因为 \mathbf{A}_1 为对称幂等矩阵, 因此是非负定矩阵), 所以

$$\mathbf{P}'(\mathbf{A} - \mathbf{A}_1)\mathbf{P} \geq \mathbf{0}, \quad \mathbf{P}'(\mathbf{A} - \mathbf{A}_2)\mathbf{P} \geq \mathbf{0}.$$

由于 $P'AP$ 的矩阵形式为(2.4.5), 所以必有

$$P' A_1 P = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P' A_2 P = \begin{pmatrix} B_2 & 0 \\ 0 & 0 \end{pmatrix},$$

其中 B_1 和 B_2 为 $r \times r$ 矩阵. 由于 $A_1^2 = A_1$, 因此 $B_1^2 = B_1$. 故存在正交矩阵 $r \times r$ 正交矩阵 Q 使得

$$Q' B_1 Q = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}.$$

记

$$S' = \begin{pmatrix} Q' & 0 \\ 0 & I_{n-r} \end{pmatrix} P',$$

则 S' 为正交矩阵, 且使

$$S' A S = S' A_1 S + S' A_2 S$$

形为

$$\begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-s} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-s} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

作变换 $\mathbf{Y} = \mathbf{S}'\mathbf{X}$, 则 $\mathbf{Y} \sim N_n(\mathbf{S}'\boldsymbol{\mu}, \mathbf{I}_n)$. 于是

$$\begin{aligned} \mathbf{X}'\mathbf{A}\mathbf{X} &= \mathbf{Y}'\mathbf{S}'\mathbf{A}\mathbf{S}\mathbf{Y} = \sum_{i=1}^r Y_i^2, \\ \mathbf{X}'\mathbf{A}_1\mathbf{X} &= \mathbf{Y}'\mathbf{S}'\mathbf{A}_1\mathbf{S}\mathbf{Y} = \sum_{i=1}^s Y_i^2, \\ \mathbf{X}'\mathbf{A}_2\mathbf{X} &= \mathbf{Y}'\mathbf{S}'\mathbf{A}_2\mathbf{S}\mathbf{Y} = \sum_{i=s+1}^r Y_i^2. \end{aligned}$$

因为 Y_1, \dots, Y_n 相互独立, 所以 $\mathbf{X}' \mathbf{A}_1 \mathbf{X}$ 与 $\mathbf{X}' \mathbf{A}_2 \mathbf{X}$ 相互独立. 再依非中心 χ^2 分布定义可知 $\mathbf{X}' \mathbf{A}_2 \mathbf{X} \sim \chi^2(r-s, \lambda_2)$. 又

$$\mathbf{A}_1 \mathbf{A}_2 = \mathbf{S} \begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}' \mathbf{S} \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-s} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}' = \mathbf{0},$$

所以定理得证.

推论 (2.4.7)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, $\mathbf{A}_1, \mathbf{A}_2$ 实对称, $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 和 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 都服从 χ^2 分布, 则它们相互独立 $\Leftrightarrow \mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$.

证明: 充分性. 令 $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$. 由 $\mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$ 可知 $\mathbf{A}_2\mathbf{A}_1 = (\mathbf{A}_1\mathbf{A}_2)' = \mathbf{0}$. 因此由 $\mathbf{A}_1, \mathbf{A}_2$ 的幂等性得

$$\mathbf{A}^2 = (\mathbf{A}_1 + \mathbf{A}_2)^2 = \mathbf{A}_1^2 + \mathbf{A}_2^2 + \mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1 = \mathbf{A}_1 + \mathbf{A}_2 = \mathbf{A},$$

即 \mathbf{A} 幂等. 由定理 2.4.4 知 $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 与 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 相互独立.

必要性. 若 $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 与 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 相互独立, 则由 χ^2 分布的可加性知 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 服从 χ^2 分布 (这里 $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$), 再由定理 2.4.4(3) 知 $\mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$.

上述两个结论很容易推广到 $\text{Cov}(\mathbf{X}) = \Sigma > \mathbf{0}$ 的情形.

推论 (2.4.8)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$, $\Sigma > \mathbf{0}$, $\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{X}'\mathbf{A}_1\mathbf{X} + \mathbf{X}'\mathbf{A}_2\mathbf{X} \sim \chi^2(r, \lambda)$, $\mathbf{X}'\mathbf{A}_1\mathbf{X} \sim \chi^2(s, \lambda_1)$, $\mathbf{A}_2 = \mathbf{A} - \mathbf{A}_1 \geq \mathbf{0}$, 则

- (1) $\mathbf{X}'\mathbf{A}_2\mathbf{X} \sim \chi^2(r-s, \lambda_2)$;
- (2) $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 与 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 相互独立;
- (3) $\mathbf{A}_1\Sigma\mathbf{A}_2 = \mathbf{0}$,

其中 $\lambda, \lambda_1, \lambda_2$ 为非中心参数, 不再精确写出.

推论 (2.4.9)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$, $\Sigma > \mathbf{0}$, $\mathbf{A}_1, \mathbf{A}_2$ 实对称, $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ 和 $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ 都服从 χ^2 分布, 则它们相互独立 $\Leftrightarrow \mathbf{A}_1\Sigma\mathbf{A}_2 = \mathbf{0}$.

接下来, 我们将建立二次型 $\mathbf{X}'\mathbf{A}\mathbf{X}$, $\mathbf{X}'\mathbf{B}\mathbf{X}$ 和线性型 $\mathbf{C}\mathbf{X}$ 相互独立的条件, 这些结果在线性模型的参数估计和假设检验中将有重要应用.

定理 (2.4.5)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, \mathbf{A} 为 $n \times n$ 实对称矩阵, \mathbf{C} 为 $m \times n$ 实矩阵, 若 $\mathbf{CA} = \mathbf{0}$, 则 \mathbf{CX} 与 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 相互独立.

证明: 因为 \mathbf{A} 为实对称矩阵, 所以存在正交阵 \mathbf{Q} 使得

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

其中, $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_r)$, $\lambda_i, i = 1, \dots, r$ 为 \mathbf{A} 的非零特征根, r 为 \mathbf{A} 的秩. 把 \mathbf{Q} 分块成 $\mathbf{Q} = (\mathbf{Q}_1: \mathbf{Q}_2)$, 其中 \mathbf{Q}_1 是 $n \times r$ 矩阵. 作正交变换

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} = \mathbf{Q}' \mathbf{X}.$$

于是 $\mathbf{Y}_i = \mathbf{Q}_i' \mathbf{X}, i = 1, 2$. 易知 $\mathbf{Y} \sim N_n(\mathbf{Q}' \boldsymbol{\mu}, \mathbf{I}_n)$, 所以

$$\mathbf{Y}_1 \sim N_r(\mathbf{Q}'_1 \boldsymbol{\mu}, \mathbf{I}_r), \quad \mathbf{Y}_2 \sim N_{n-r}(\mathbf{Q}'_2 \boldsymbol{\mu}, \mathbf{I}_{n-r}),$$

且 \mathbf{Y}_1 与 \mathbf{Y}_2 相互独立. 注意到

$$\mathbf{X}' \mathbf{A} \mathbf{X} = \mathbf{X}' \mathbf{Q}_1 \boldsymbol{\Lambda} \mathbf{Q}'_1 \mathbf{X} = \mathbf{Y}'_1 \boldsymbol{\Lambda} \mathbf{Y}_1, \quad (2.4.6)$$

$$\mathbf{C} \mathbf{X} = \mathbf{C} \mathbf{Q} \mathbf{Y} \triangleq \mathbf{D} \mathbf{Y}, \quad (2.4.7)$$

这里 $\mathbf{D} = \mathbf{C} \mathbf{Q}$. 由于 $\mathbf{C} \mathbf{A} = \mathbf{0}$, 所以

$$\mathbf{0} = \mathbf{C} \mathbf{A} \mathbf{Q} = \mathbf{C} \mathbf{Q} \mathbf{Q}' \mathbf{A} \mathbf{Q} = \mathbf{D} \mathbf{Q}' \mathbf{A} \mathbf{Q} = \mathbf{D} \begin{pmatrix} \boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

把 \mathbf{D} 分块成 $\mathbf{D} = (\mathbf{D}_1 : \mathbf{D}_2)$, 其中 \mathbf{D}_1 是 $m \times r$ 矩阵. 则由上式可推知 $\mathbf{D}_1 = \mathbf{0}$, 代入(2.4.7)可知

$$\mathbf{C} \mathbf{X} = \mathbf{D} \mathbf{Y} = \mathbf{D}_2 \mathbf{Y}_2.$$

再由 \mathbf{Y}_1 和 \mathbf{Y}_2 的独立性, 结合上式与(2.4.6), 可知 $\mathbf{C} \mathbf{X}$ 与 $\mathbf{X}' \mathbf{A} \mathbf{X}$ 相互独立.

例2.4.2 设 X_1, \dots, X_n 为来自 $N(\mu, \sigma^2)$ 的样本, 则样本均值 \bar{X} 与样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 相互独立.

事实上, 若记 $\mathbf{X} = (X_1, \dots, X_n)'$, $\mathbf{1}_n = (1, \dots, 1)'$, 以上向量均是 n 维. 则

$$\bar{X} = \frac{1}{n} \mathbf{1}_n' \mathbf{X}.$$

例2.4.1已告诉我们

$$(n-1)S^2 = \mathbf{X}' \mathbf{C} \mathbf{X},$$

其中 $\mathbf{C} = \mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n'$. 可以验证 $\mathbf{1}_n' \mathbf{C} = \mathbf{0}$, 所以由定理2.4.5知 \bar{X} 与 S^2 相互独立.

定理 (2.4.6)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, \mathbf{A} 和 \mathbf{B} 都为 $n \times n$ 对称矩阵, 且 $\mathbf{AB} = \mathbf{0}$, 则二次型 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 与 $\mathbf{X}'\mathbf{B}\mathbf{X}$ 相互独立.

证明: 由 $\mathbf{AB} = \mathbf{0}$ 以及 \mathbf{A} 和 \mathbf{B} 的对称性, 立得 $\mathbf{BA} = \mathbf{AB} = \mathbf{0}$, 即 \mathbf{A}, \mathbf{B} 可交换. 所以可用同一正交阵将这两个矩阵对角化(见王松桂, 《矩阵不等式》, page 11), 即存在正交阵 \mathbf{Q} 使得

$$\begin{aligned}\mathbf{Q}'\mathbf{A}\mathbf{Q} &= \boldsymbol{\Lambda}_1 = \text{diag}(\lambda_1^{(1)}, \dots, \lambda_n^{(1)}), \\ \mathbf{Q}'\mathbf{B}\mathbf{Q} &= \boldsymbol{\Lambda}_2 = \text{diag}(\lambda_1^{(2)}, \dots, \lambda_n^{(2)}).\end{aligned}$$

由 $\mathbf{AB} = \mathbf{0}$, 可推得 $\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2 = \mathbf{0}$, 即

$$\lambda_i^{(1)} \text{ 和 } \lambda_i^{(2)} \text{ 至少有一个为 } 0, i = 1, \dots, n. \quad (2.4.8)$$

令 $Y = Q'X$, 则 $Y \sim N(Q'\mu, I_n)$, 于是 Y 的所有分量都相互独立. 另一方面, 由于

$$\begin{aligned}X'AX &= X'Q\Lambda_1Q'X = Y'\Lambda_1Y, \\X'BX &= X'Q\Lambda_2Q'X = Y'\Lambda_2Y,\end{aligned}$$

所以由(2.4.8)可知 $X'AX$ 与 $X'BX$ 依赖于 Y 的不同分量. 所以 $X'AX$ 与 $X'BX$ 相互独立.

这个定理的逆也是对的, 即设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$, \mathbf{A} 和 \mathbf{B} 都为 $n \times n$ 实对称矩阵, 若 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 与 $\mathbf{X}'\mathbf{B}\mathbf{X}$ 相互独立, 则 $\mathbf{AB} = \mathbf{0}$. 证明略. 此外, 定理2.4.6可推广为

推论 (2.4.10)

设 $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$, \mathbf{A}, \mathbf{B} 对称. 若 $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{0}$, 则 $\mathbf{X}'\mathbf{A}\mathbf{X}$ 和 $\mathbf{X}'\mathbf{B}\mathbf{X}$ 相互独立.

补充:

结论: 设 A, B 为 $n \times n$ 实对称矩阵, 则存在正交阵 Q 使得 $Q' A Q$ 与 $Q' B Q$ 为对角阵, 当且仅当 $AB = BA$.

证明: 充分性. 首先可知存在正交阵 P_1 使得

$$P_1' A P_1 = \text{diag}(\lambda_1 I_{r_1}, \dots, \lambda_s I_{r_s}) =: D,$$

其中 $r_1 + \dots + r_s = n$. 又因为 $AB = BA$, 所以有

$$P_1' A P_1 P_1' B P_1 = P_1' B P_1 P_1' A P_1, \text{ 即 } DC = CD,$$

其中 $C = P_1' B P_1$. 因此, C 必为分块对角矩阵:

$$C = \text{diag}(C_{r_1}, \dots, C_{r_s}).$$

注意到 C 是对称矩阵, 所以存在正交矩阵

$$P_2 = \text{diag}(P_{r_1}, \dots, P_{r_s})$$

使得 $P_2' C P_2$ 为对角阵. 令 $Q = P_1 P_2$, 则可同时使 $Q' A Q$ 与 $Q' B Q$ 为对角阵.

必要性: 注意到

$$Q'ABQ = Q'AQ \cdot Q'BQ = Q'BQ \cdot Q'AQ = Q'BAQ,$$

左乘 Q 并右乘 Q' 可得 $AB = BA$.

矩阵微商

在统计学中, 为了获得参数的极大似然估计, 我们常常需要求似然函数的极值, 这就要用到矩阵微商. 本节我们讨论一些常用的结果.

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

$$\frac{\partial y}{\partial \mathbf{X}} \triangleq \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \frac{\partial y}{\partial x_{m2}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{pmatrix}_{m \times n}$$

为 y 对 \mathbf{X} 的微商.

下面我们都假设矩阵 \mathbf{X} 中的 mn 个变量 $x_{ij}, i = 1, \dots, m, j = 1, \dots, n$ 是独立变量. 接下来我们介绍两个最常用的矩阵微商结论.

例2.5.1 设 \mathbf{a}, \mathbf{x} 均为 $n \times 1$ 向量, $y = \mathbf{a}' \mathbf{x}$, 则 $\frac{\partial y}{\partial \mathbf{x}} = \mathbf{a}$. 事实上, 因为 $y = \sum_{i=1}^n a_i x_i$, 所以

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{a}.$$

例2.5.2 设 A 为 $n \times n$ 对称矩阵, x 为 $n \times 1$ 向量, $y = x'Ax$, 则 $\frac{\partial y}{\partial x} = 2Ax$. 事实上, 由于 $y = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j$, 所以

$$\begin{aligned}\frac{\partial y}{\partial x_1} &= \sum_{i=1}^n a_{i1}x_i + \sum_{j=1}^n a_{1j}x_j, \\ \frac{\partial y}{\partial x_2} &= \sum_{i=1}^n a_{i2}x_i + \sum_{j=1}^n a_{2j}x_j, \\ &\vdots \\ \frac{\partial y}{\partial x_n} &= \sum_{i=1}^n a_{in}x_i + \sum_{j=1}^n a_{nj}x_j.\end{aligned}$$

因此, 可以看出

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{i1}x_i + \sum_{j=1}^n a_{1j}x_j \\ \sum_{i=1}^n a_{i2}x_i + \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{i=1}^n a_{in}x_i + \sum_{j=1}^n a_{nj}x_j \end{pmatrix} = 2\mathbf{A}\mathbf{x}.$$