第二章　随机向量

Tianxiao Pang

Zhejiang University

September 21, 2023

内容

内容

内容

内容

内容

在第一章, 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.

在第一章, 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.

若无特殊说明, 约定:

在第一章, 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.

若无特殊说明, 约定:

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

均值向量与协方差矩阵

设$\boldsymbol{X}=(X_1,\cdots,X_n)'$为$n\times 1$随机向量, 定义$\boldsymbol{X}$的数学期望(均值向量)为

$$\mathsf{E}(\boldsymbol{X})=(\mathsf{E}(X_1),\cdots,\mathsf{E}(X_n))'.$$

设$\boldsymbol{X} = (X_1, \cdots, X_n)'$为$n \times 1$随机向量，定义$\boldsymbol{X}$的数学期望(均值向量)为

$$\mathsf{E}(\boldsymbol{X}) = (\mathsf{E}(X_1), \cdots, \mathsf{E}(X_n))'.$$

类似地，设$\boldsymbol{X} = (X_{ij})_{n\times m}$为一个$n \times m$的随机矩阵，定义$\boldsymbol{X}$的数学期望为

$$\mathsf{E}(\boldsymbol{X}) = (\mathsf{E}(X_{ij}))_{n\times m}.$$

定理（2.1.1）

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

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

定理（2.1.1）

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

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

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

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

定理（2.1.1）

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

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

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

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

求均值得

$$\mathrm{E}(Y_i) = \sum_{j=1}^{n} a_{ij} \mathrm{E}(X_j) + \mathrm{E}(b_i),$$

左边是$\mathrm{E}(\boldsymbol{Y})$的第$i$个元素，右边是$\boldsymbol{A}\mathrm{E}(\boldsymbol{X}) + \mathrm{E}(\boldsymbol{b})$的第$i$个元素，结论得证．

定义n维随机向量$\boldsymbol{X}$的协方差矩阵为

$$\mathsf{Cov}(\boldsymbol{X}) = \mathsf{E}[(\boldsymbol{X} - \mathsf{E}(\boldsymbol{X}))(\boldsymbol{X} - \mathsf{E}(\boldsymbol{X}))'].$$

定义n维随机向量$\boldsymbol{X}$的协方差矩阵为

$$\mathrm{Cov}(\boldsymbol{X}) = \mathrm{E}[(\boldsymbol{X} - \mathrm{E}(\boldsymbol{X}))(\boldsymbol{X} - \mathrm{E}(\boldsymbol{X}))'].$$

推论 (2.1.1)

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

定理 (2.1.2)

设$\boldsymbol{X}$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矩阵.

定理（2.1.2)

设$\boldsymbol{X}$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矩阵.

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$\boldsymbol{c}$, 注意到$\boldsymbol{c}'\boldsymbol{X}$是一个随机变量, 所以

定理 (2.1.2)

设$\boldsymbol{X}$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矩阵.

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$\boldsymbol{c}$, 注意到$\boldsymbol{c}'\boldsymbol{X}$是一个随机变量, 所以

$$\begin{aligned}
0 \le \mathsf{Var}(\boldsymbol{c}'\boldsymbol{X}) &= \mathsf{E}[\boldsymbol{c}'\boldsymbol{X} - \mathsf{E}(\boldsymbol{c}'\boldsymbol{X})]^2 \\
&= \mathsf{E}[(\boldsymbol{c}'\boldsymbol{X} - \mathsf{E}(\boldsymbol{c}'\boldsymbol{X}))(\boldsymbol{c}'\boldsymbol{X} - \mathsf{E}(\boldsymbol{c}'\boldsymbol{X}))'] \\
&= \boldsymbol{c}'\mathsf{E}[(\boldsymbol{X} - \mathsf{E}(\boldsymbol{X}))(\boldsymbol{X} - \mathsf{E}(\boldsymbol{X}))']\boldsymbol{c} \\
&= \boldsymbol{c}'\mathsf{Cov}(\boldsymbol{X})\boldsymbol{c},
\end{aligned}$$

得证.

定理（2.1.3)

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

$$\mathsf{Cov}(\boldsymbol{Y}) = \boldsymbol{A}\,\mathsf{Cov}(\boldsymbol{X})\boldsymbol{A}'.$$

定理（2.1.3）

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

$$\mathit{Cov}(\boldsymbol{Y}) = \boldsymbol{A}\mathit{Cov}(\boldsymbol{X})\boldsymbol{A}'.$$

证明：根据协方差矩阵的定义，

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

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

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

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

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

定理 (2.1.4)

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

$$Cov(\boldsymbol{AX},\boldsymbol{BY})=\boldsymbol{A}Cov(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{B}'.$$

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

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

定理 (2.1.4)

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

$$Cov(\boldsymbol{AX},\boldsymbol{BY}) = \boldsymbol{A}Cov(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{B}'.$$

证明:

$$\begin{aligned}\mathsf{Cov}(\boldsymbol{AX},\boldsymbol{BY}) &= \mathsf{E}[(\boldsymbol{AX}-\mathsf{E}(\boldsymbol{AX}))(\boldsymbol{BY}-\mathsf{E}(\boldsymbol{BY}))'] \\ &= \boldsymbol{A}\mathsf{E}[(\boldsymbol{X}-\mathsf{E}(\boldsymbol{X}))(\boldsymbol{Y}-\mathsf{E}(\boldsymbol{Y}))']\boldsymbol{B}' \\ &= \boldsymbol{A}\mathsf{Cov}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{B}'.\end{aligned}$$

随机向量的二次型

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

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

为$\boldsymbol{X}$的二次型, 称$\boldsymbol{A}$为此二次型的二次型矩阵.

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

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}X_iX_j$$

为$\boldsymbol{X}$的二次型, 称$\boldsymbol{A}$为此二次型的二次型矩阵.

定理 (2.2.1)

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

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

证明: 因为

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

证明: 因为

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

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

$$
\mathsf{E}[(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})]=\mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}).
$$

证明: 因为

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

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

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

利用迹的性质($\mathsf{tr}(\boldsymbol{A}\boldsymbol{B})=\mathsf{tr}(\boldsymbol{B}\boldsymbol{A})$以及求迹和求数学期望可交换次序), 可知

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

补充:

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

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

推论（2.2.1)

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

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

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

(3) 若$\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{I}_n$, 则$E(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}) = tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$.

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

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

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

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

解: 记$Q = (n-1)S^2$. 把Q表示成$\boldsymbol{X} = (X_1, \cdots, X_n)'$的一个二次型.

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

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

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

$$\mathsf{E}(\boldsymbol{X}) = \mu\mathbf{1}_n, \quad \mathsf{Cov}(\boldsymbol{X}) = \sigma^2\boldsymbol{I}_n, \quad \overline{X} = \frac{1}{n}\mathbf{1}_n'\boldsymbol{X},$$

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

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

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

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

应用定理2.2.1得

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

又易知

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

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

来推导二次型$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$的方差公式.

来推导二次型$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$的方差公式.

定理 (2.2.2)

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

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

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

$$\boldsymbol{X} = (X_1, \cdots, X_n)', \;\; \boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)'.$$

则

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

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

证明: 首先注意到

$$\mathsf{Var}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}) = \mathsf{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})^2 - [\mathsf{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})]^2, \tag{2.2.2}$$

证明：首先注意到

$$\mathrm{Var}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}) = \mathrm{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})^2 - [\mathrm{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})]^2, \tag{2.2.2}$$

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

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

所以重点是计算(2.2.2)中的第一项.

证明：首先注意到

$$\mathrm{Var}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}) = \mathrm{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})^2 - [\mathrm{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})]^2, \tag{2.2.2}$$

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

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

所以重点是计算(2.2.2)中的第一项. 将$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$改写为

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

将其平方得到

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

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

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

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

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

下面逐个计算上式所含的每个均值. 由

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

及Z_i的独立性可得

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

便有下列结果:

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

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

而

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

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

$$\mathsf{E}(\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{A}\boldsymbol{Z})=\sum_i\sum_j\sum_k b_i a_{jk}\mathsf{E}(Z_iZ_jZ_k).$$

因为

$$\mathsf{E}(Z_iZ_jZ_k)=\begin{cases} m_3, & \text{若} i=j=k, \\ 0, & \text{其它}, \end{cases}$$

所以

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

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.

定义

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

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

其中$\boldsymbol{x} = (x_1, \cdots, x_n)' \in \mathbb{R}^n$, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)'$, $\boldsymbol{\Sigma}$是对称正定矩阵, 则称$\boldsymbol{X}$为n维正态随机向量, 记为$N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$或$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

事实上, $\boldsymbol{\mu}$和$\boldsymbol{\Sigma}$分别是$\boldsymbol{X}$的均值向量和协方差矩阵.

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

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

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

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

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

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

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

其中J为向量变换的Jocobi行列式:

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

所以

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

所以

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

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

$$\mathsf{E}(\boldsymbol{Y})=\boldsymbol{0},\ \ \mathsf{Cov}(\boldsymbol{Y})=\boldsymbol{I}_n,$$

由此可知

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

所以

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

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

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

由此可知

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

注: 称$N(\mathbf{0},\boldsymbol{I}_n)$为$n$元/$n$维标准正态分布.

设$\boldsymbol{X}$的协方差矩阵具有如下分块对角形式:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{2.3.2}$$

这里$\boldsymbol{\Sigma}_{11}$为$m \times m$矩阵. 将$\boldsymbol{X}, \boldsymbol{x}$和$\boldsymbol{\mu}$也作相应的分块:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \tag{2.3.3}$$

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

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

其中

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

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

其中

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

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

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

定理 (2.3.1)

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

其中

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

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

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

定理 (2.3.1)

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

注: 对多元正态随机向量$\boldsymbol{X}$来说, $\boldsymbol{X}_1$与$\boldsymbol{X}_2$不相关($\mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$), 可以推出$\boldsymbol{X}_1$与$\boldsymbol{X}_2$独立.

定理（2.3.2）

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

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

定理（2.3.2）

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

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

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

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

J为变换的Jocobi行列式：

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

定理（2.3.2）

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

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

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

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

J为变换的Jocobi行列式:

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

注意到

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

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

所以

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

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

由于$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\Sigma}^{-1/2})' = \boldsymbol{I}_n$, 所以有

由于$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\Sigma}^{-1/2})' = \boldsymbol{I}_n$, 所以有

推论 (2.3.1)

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

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

由于$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\Sigma}^{-1/2})' = \boldsymbol{I}_n$, 所以有

推论 (2.3.1)

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

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

这个推论表明: 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($\boldsymbol{X}$)变换为多元标准正态随机向量($\boldsymbol{Z}$).

推论 (2.3.2)

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

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

推论 (2.3.2)

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

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

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

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

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

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

定理 (2.3.3)

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

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

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

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

证明: 在定理2.3.2中取

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

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

定理 (2.3.3)

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

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

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

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

证明: 在定理2.3.2中取

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

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

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

于是

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

于是

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

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

此外, 可知$\boldsymbol{X}_1$与$\boldsymbol{X}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{X}_1$独立.

类似地，可证明$\boldsymbol{X}_2 \sim N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$，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.

下面的定理是定理2.3.2的改进版本.

下面的定理是定理2.3.2的改进版本.

定理 (2.3.4)

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

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

下面的定理是定理2.3.2的改进版本.

定理（2.3.4)

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

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

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

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

应用定理2.3.2得

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

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

推论 (2.3.3)

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

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

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.

推论（2.3.3）

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

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

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

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \boldsymbol{c}'\boldsymbol{X} \sim$$

推论 (2.3.3)

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

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

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

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \boldsymbol{c}'\boldsymbol{X} \sim N\Big(\mu, \frac{\sigma^2}{n}\Big),$$

其中$\boldsymbol{X} = (X_1, \cdots, X_n)' \sim N(\mu\mathbf{1}_n, \sigma^2\boldsymbol{I}_n)$, $\boldsymbol{c} = (\frac{1}{n}, \cdots, \frac{1}{n})' = \frac{1}{n}\mathbf{1}_n$.

推论 (2.3.4)

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

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

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

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

定理 (2.4.1)

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

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

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

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

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

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

把问题转化为求$\boldsymbol{Y}$的二次型的方差, 这里的二次型矩阵是$\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$. 注意到

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

应用定理2.2.2便可得到第一条结论.

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

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

把问题转化为求$\boldsymbol{Y}$的二次型的方差, 这里的二次型矩阵是$\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$. 注意到

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

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

定义 (2.4.1)

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

χ^2分布具有下述性质:

定义 (2.4.1)

设$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{I}_n)$. 称随机变量$Y = \boldsymbol{X}'\boldsymbol{X}$的分布为自由度为$n$, 非中心参数为$\lambda = \boldsymbol{\mu}'\boldsymbol{\mu}$的$\chi^2$分布, 记为$Y \sim \chi^2(n, \lambda)$. 当$\lambda = 0$时, 称$Y$的分布为中心$\chi^2$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$Y_i \sim \chi^2(n_i, \lambda_i), i = 1, \cdots, k$, 且相互独立, 则

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

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

(2) $\boldsymbol{Y} \sim \chi^2(n, \lambda)$, 则$E(\boldsymbol{Y}) = n + \lambda$, $Var(\boldsymbol{Y}) = 2n + 4\lambda$.

证明: (1) 可用特征函数方法证明.
(2) 设$Y \sim \chi^2(n,\lambda)$, 依定义,

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

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

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

因为

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

所以$\mathsf{E}(Y) = n + \lambda$.

此外, 经简单计算可得

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

于是有

$$\begin{aligned}\mathsf{Var}(X_i^2)&=\mathsf{E}(X_i^4)-[\mathsf{E}(X_i^2)]^2=3-1=2,\ \ i=1,\cdots,n-1,\\ \mathsf{Var}(X_n^2)&=\mathsf{E}(X_n^4)-[\mathsf{E}(X_n^2)]^2=2+4\lambda.\end{aligned}$$

把上述结果代入(2.4.1)即可证明第二条结论.

此外, 经简单计算可得

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

于是有

$$\begin{aligned} \mathsf{Var}(X_i^2) &= \mathsf{E}(X_i^4) - [\mathsf{E}(X_i^2)]^2 = 3 - 1 = 2, \ \ i = 1, \cdots, n-1, \\ \mathsf{Var}(X_n^2) &= \mathsf{E}(X_n^4) - [\mathsf{E}(X_n^2)]^2 = 2 + 4\lambda. \end{aligned}$$

把上述结果代入(2.4.1)即可证明第二条结论.

注: 令$\boldsymbol{X} = (X_1, \cdots, X_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{I}_n)$, 其中$\boldsymbol{\mu} = (0, \cdots, 0, \sqrt{\lambda})'$. 则$\boldsymbol{X}'\boldsymbol{X} \sim \chi^2(n, \lambda)$, 即$Y \overset{d}{=} \boldsymbol{X}'\boldsymbol{X}$. 因此, 利用正态随机向量二次型的数学期望和方差公式, 可得

$$\begin{aligned} \mathsf{E}(Y) &= \mathsf{E}(\boldsymbol{X}'\boldsymbol{X}) = n + \lambda, \\ \mathsf{Var}(\boldsymbol{Y}) &= \mathsf{Var}(\boldsymbol{X}'\boldsymbol{X}) = 2n + 4\lambda. \end{aligned}$$

推论（2.4.1)

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

推论 (2.4.1)

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

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

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

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

推论 (2.4.2)

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

推论（2.4.2）

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

推论（2.4.3）

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

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

对于正态随机向量的一般二次型, 有下面的定理.

对于正态随机向量的一般二次型, 有下面的定理.

定理 (2.4.3)

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

对于正态随机向量的一般二次型, 有下面的定理.

定理 (2.4.3)

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

证明: 先证充分性. 设$\boldsymbol{A}$幂等、对称且$\mathsf{rk}(\boldsymbol{A}) = r$. 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$\boldsymbol{Q}$使得

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

对于正态随机向量的一般二次型, 有下面的定理.

定理 (2.4.3)

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

证明: 先证充分性. 设$\boldsymbol{A}$幂等、对称且$\text{rk}(\boldsymbol{A}) = r$. 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$\boldsymbol{Q}$使得

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

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

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

其中$\boldsymbol{Y}_1 : r \times 1$, $\boldsymbol{Q}_1 : r \times n$. 于是$\boldsymbol{A} = \boldsymbol{Q}_1'\boldsymbol{Q}_1, \boldsymbol{Y}_1 \sim N_r(\boldsymbol{Q}_1\boldsymbol{\mu}, \boldsymbol{I}_r)$且

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

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

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

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

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

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

其中$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_t)$, $\lambda_1, \cdots, \lambda_t$非零.

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

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

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

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

其中$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_t)$, $\lambda_1, \cdots, \lambda_t$非零. 只需证明

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

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

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

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

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

其中$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_t)$, $\lambda_1, \cdots, \lambda_t$非零. 只需证明

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

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

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

则可得

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

这里$\boldsymbol{Y} = (Y_1, \cdots, Y_n)'$, $Y_j \sim N(c_j, 1), j = 1, \cdots, t$, 且相互独立. 可算出$\lambda_j Y_j^2$的特征函数为

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

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$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$的特征函数为

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

再来计算$\chi^2(r, \lambda)$的特征函数. 记$u = u_1^2 + u_2^2 + \cdots + u_r^2$, 其中$u_j \sim N(0, 1), j \le r - 1$, $u_r \sim N(\sqrt{\lambda}, 1)$. 则$u \sim \chi^2(r, \lambda)$, $\lambda = \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$. 不难求得$u$的特征函数为

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

依假设, $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} \sim \chi^2(r, \lambda)$, 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, $\lambda_j = 1, j = 1, \cdots, t$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$\boldsymbol{A}$是$n\times n$对称幂等矩阵. 由对称性可知, 存在正交矩阵$\boldsymbol{Q}$使得

$$\boldsymbol{A}=\boldsymbol{Q}\text{diag}(\lambda_1,\cdots,\lambda_n)\boldsymbol{Q}',$$

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

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

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

推论 (2.4.4)

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

推论（2.4.4）

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

推论（2.4.5）

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

推论（2.4.4）

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

推论（2.4.5）

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

推论（2.4.6）

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

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, 这些结果是判断一个随机变量是否服从χ^2分布的有效工具.

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

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

为σ^2的无偏估计, 下证

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

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

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

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$\boldsymbol{X}=(X_1,\cdots,X_n)'\sim N_n(\mu\mathbf{1}_n,\sigma^2\boldsymbol{I}_n)$, 所以

$$\boldsymbol{Y}=\frac{\boldsymbol{X}-\mu\mathbf{1}_n}{\sigma}\sim$$

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

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

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$\boldsymbol{X}=(X_1,\cdots,X_n)'\sim N_n(\mu\mathbf{1}_n,\sigma^2\boldsymbol{I}_n)$, 所以

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

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

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

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$\boldsymbol{X}=(X_1,\cdots,X_n)'\sim N_n(\mu\mathbf{1}_n,\sigma^2\boldsymbol{I}_n)$，所以

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

记$\boldsymbol{C}=\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$，则易知$\boldsymbol{C}$是对称幂等矩阵，$\mathrm{rk}(\boldsymbol{C})=\mathrm{tr}(\boldsymbol{C})=$

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

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

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$\boldsymbol{X}=(X_1,\cdots,X_n)'\sim N_n(\mu\mathbf{1}_n,\sigma^2\boldsymbol{I}_n)$, 所以

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

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

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

补充:

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

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

利用性质$\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$, 有

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

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

定理 (2.4.4)

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

(1) $\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} \sim \chi^2(r-s, \lambda_2)$, $\lambda_2 = \boldsymbol{\mu}'\boldsymbol{A}_2\boldsymbol{\mu}$;

(2) $\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立;

(3) $\boldsymbol{A}_1\boldsymbol{A}_2 = \boldsymbol{0}$.

定理 (2.4.4)

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

(1) $\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} \sim \chi^2(r-s, \lambda_2)$, $\lambda_2 = \boldsymbol{\mu}'\boldsymbol{A}_2\boldsymbol{\mu}$;

(2) $\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立;

(3) $\boldsymbol{A}_1\boldsymbol{A}_2 = \boldsymbol{0}$.

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

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{2.4.5}$$

定理 (2.4.4)

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

(1) $\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} \sim \chi^2(r-s, \lambda_2)$, $\lambda_2 = \boldsymbol{\mu}'\boldsymbol{A}_2\boldsymbol{\mu}$;

(2) $\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立;

(3) $\boldsymbol{A}_1\boldsymbol{A}_2 = \boldsymbol{0}$.

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

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{2.4.5}$$

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

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

由于$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}$的矩阵形式为(2.4.5), 所以必有

由于$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}$的矩阵形式为(2.4.5), 所以必有

$$\boldsymbol{P}'\boldsymbol{A}_1\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{P}'\boldsymbol{A}_2\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

其中$\boldsymbol{B}_1$和$\boldsymbol{B}_2$为$r \times r$对称矩阵.

由于$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}$的矩阵形式为(2.4.5), 所以必有

$$\boldsymbol{P}'\boldsymbol{A}_1\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{P}'\boldsymbol{A}_2\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

其中$\boldsymbol{B}_1$和$\boldsymbol{B}_2$为$r \times r$对称矩阵. 由于$\boldsymbol{A}_1^2 = \boldsymbol{A}_1$, 因此$\boldsymbol{B}_1^2 = \boldsymbol{B}_1$. 故存在$r \times r$正交矩阵$\boldsymbol{Q}$使得

$$\boldsymbol{Q}'\boldsymbol{B}_1\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_s & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad s \le r.$$

由于$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}$的矩阵形式为(2.4.5), 所以必有

$$\boldsymbol{P}'\boldsymbol{A}_1\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{P}'\boldsymbol{A}_2\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

其中$\boldsymbol{B}_1$和$\boldsymbol{B}_2$为$r \times r$对称矩阵. 由于$\boldsymbol{A}_1^2 = \boldsymbol{A}_1$, 因此$\boldsymbol{B}_1^2 = \boldsymbol{B}_1$. 故存在$r \times r$正交矩阵$\boldsymbol{Q}$使得

$$\boldsymbol{Q}'\boldsymbol{B}_1\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_s & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad s \leq r.$$

记

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

则$\boldsymbol{S}'$为正交矩阵, 且使

$$\boldsymbol{S}'\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}'\boldsymbol{A}_1\boldsymbol{S} + \boldsymbol{S}'\boldsymbol{A}_2\boldsymbol{S}$$

形为

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

形为

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

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

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{S}'\boldsymbol{A}\boldsymbol{S}\boldsymbol{Y} = \sum_{i=1}^{r} Y_i^2,$$

$$\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{S}'\boldsymbol{A}_1\boldsymbol{S}\boldsymbol{Y} = \sum_{i=1}^{s} Y_i^2,$$

$$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{S}'\boldsymbol{A}_2\boldsymbol{S}\boldsymbol{Y} = \sum_{i=s+1}^{r} Y_i^2.$$

因为$Y_1, \cdots, Y_n$相互独立, 所以$\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立. 因为$\boldsymbol{S}'\boldsymbol{A}_2\boldsymbol{S}$是对称幂等矩阵, 秩为$r-s$, 所以

$$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{S}'\boldsymbol{A}_2\boldsymbol{S}\boldsymbol{Y} \sim \chi^2(r-s, \lambda_2)$$

且$\lambda_2 = (\boldsymbol{S}'\boldsymbol{\mu})'\boldsymbol{S}'\boldsymbol{A}_2\boldsymbol{S}(\boldsymbol{S}'\boldsymbol{\mu}) = \boldsymbol{\mu}'\boldsymbol{A}_2\boldsymbol{\mu}$. 又

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

所以定理得证.

推论（2.4.7）

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

推论 (2.4.7)

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

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

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

即$\boldsymbol{A}$对称幂等. 由定理2.4.4(2)知$\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立.

推论 (2.4.7)

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

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

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

即$\boldsymbol{A}$对称幂等. 由定理2.4.4(2)知$\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立.

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

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

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

推论 (2.4.8)

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

(1) $\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} \sim \chi^2(r-s, \lambda_2)$;

(2) $\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立;

(3) $\boldsymbol{A}_1\boldsymbol{\Sigma}\boldsymbol{A}_2 = \mathbf{0}$,

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

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

推论 (2.4.8)

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

(1) $\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X} \sim \chi^2(r - s, \lambda_2)$;

(2) $\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}_2\boldsymbol{X}$相互独立;

(3) $\boldsymbol{A}_1\boldsymbol{\Sigma}\boldsymbol{A}_2 = \mathbf{0}$,

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

接下来, 将建立二次型$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$, $\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$和线性型$\boldsymbol{C}\boldsymbol{X}$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$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{I}_n)$, $\boldsymbol{A}$为$n \times n$实对称矩阵, $\boldsymbol{C}$为$m \times n$实矩阵, 若$\boldsymbol{C}\boldsymbol{A} = \boldsymbol{0}$, 则$\boldsymbol{C}\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$相互独立.

定理（2.4.5）

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

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

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

其中, $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_r)$, λ_i, $i = 1, \cdots, r$, 为$\boldsymbol{A}$的非零特征根, r为$\boldsymbol{A}$的秩.

定理 (2.4.5)

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

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

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

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

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

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

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

且$\boldsymbol{Y}_1$与$\boldsymbol{Y}_2$相互独立.

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

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

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Y} = \boldsymbol{Y}_1'\boldsymbol{\Lambda}\boldsymbol{Y}_1, \tag{2.4.6}$$

$$\boldsymbol{C}\boldsymbol{X} = \boldsymbol{C}\boldsymbol{Q}\boldsymbol{Y} \triangleq \boldsymbol{D}\boldsymbol{Y}, \tag{2.4.7}$$

这里$\boldsymbol{D} = \boldsymbol{C}\boldsymbol{Q}$.

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

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

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Y} = \boldsymbol{Y}_1'\boldsymbol{\Lambda}\boldsymbol{Y}_1, \tag{2.4.6}$$

$$\boldsymbol{C}\boldsymbol{X} = \boldsymbol{C}\boldsymbol{Q}\boldsymbol{Y} \triangleq \boldsymbol{D}\boldsymbol{Y}, \tag{2.4.7}$$

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

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

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

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

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Y} = \boldsymbol{Y}_1'\boldsymbol{\Lambda}\boldsymbol{Y}_1, \tag{2.4.6}$$

$$\boldsymbol{C}\boldsymbol{X} = \boldsymbol{C}\boldsymbol{Q}\boldsymbol{Y} \triangleq \boldsymbol{D}\boldsymbol{Y}, \tag{2.4.7}$$

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

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

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

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

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

推论 (2.4.10)

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

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

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

事实上, 若记$\boldsymbol{X} = (X_1, \cdots, X_n)'$, $\mathbf{1}_n = (1, \cdots, 1)'$. 则

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

例2.4.1已告诉我们

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

其中$\boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$. 可以验证$\frac{1}{n}\mathbf{1}_n' \times \sigma^2 \boldsymbol{I}_n \times \boldsymbol{C} = \mathbf{0}$, 所以由推论2.4.10知$\overline{X}$与$S^2$相互独立.

定理 (2.4.6)

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

定理 (2.4.6)

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

证明: 由$\boldsymbol{A}$和$\boldsymbol{B}$的对称性及$\boldsymbol{AB} = \boldsymbol{0}$, 得$\boldsymbol{BA} = \boldsymbol{AB}$, 即$\boldsymbol{A}, \boldsymbol{B}$可交换.

定理 (2.4.6)

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

证明: 由$\boldsymbol{A}$和$\boldsymbol{B}$的对称性及$\boldsymbol{AB} = \boldsymbol{0}$, 得$\boldsymbol{BA} = \boldsymbol{AB}$, 即$\boldsymbol{A}, \boldsymbol{B}$可交换. 所以可用同一正交阵将这两个矩阵对角化, 即存在正交阵$\boldsymbol{Q}$使得

$$\boldsymbol{Q}'\boldsymbol{AQ} = \boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1^{(1)}, \cdots, \lambda_n^{(1)}),$$
$$\boldsymbol{Q}'\boldsymbol{BQ} = \boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_1^{(2)}, \cdots, \lambda_n^{(2)}).$$

定理（2.4.6)

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

证明：由$\boldsymbol{A}$和$\boldsymbol{B}$的对称性及$\boldsymbol{AB} = \boldsymbol{0}$，得$\boldsymbol{BA} = \boldsymbol{AB}$，即$\boldsymbol{A}, \boldsymbol{B}$可交换. 所以可用同一正交阵将这两个矩阵对角化，即存在正交阵$\boldsymbol{Q}$使得

$$\boldsymbol{Q}'\boldsymbol{AQ} = \boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1^{(1)}, \cdots, \lambda_n^{(1)}),$$
$$\boldsymbol{Q}'\boldsymbol{BQ} = \boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_1^{(2)}, \cdots, \lambda_n^{(2)}).$$

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

定理（2.4.6)

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

证明：由$\boldsymbol{A}$和$\boldsymbol{B}$的对称性及$\boldsymbol{AB} = \boldsymbol{0}$，得$\boldsymbol{BA} = \boldsymbol{AB}$，即$\boldsymbol{A}, \boldsymbol{B}$可交换. 所以可用同一正交阵将这两个矩阵对角化，即存在正交阵$\boldsymbol{Q}$使得

$$\begin{aligned} \boldsymbol{Q}'\boldsymbol{AQ} &= \boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1^{(1)}, \cdots, \lambda_n^{(1)}), \\ \boldsymbol{Q}'\boldsymbol{BQ} &= \boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_1^{(2)}, \cdots, \lambda_n^{(2)}). \end{aligned}$$

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

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

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

$$\begin{aligned} \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} &= \boldsymbol{Y}'\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{\Lambda}_1\boldsymbol{Y}, \\ \boldsymbol{X}'\boldsymbol{B}\boldsymbol{X} &= \boldsymbol{Y}'\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{Q}\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{\Lambda}_2\boldsymbol{Y}, \end{aligned}$$

所以由(2.4.8)可知$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$依赖于$\boldsymbol{Y}$的不同分量. 所以$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$与$\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$相互独立.

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

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

推论 (2.4.11)

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

补充:

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

证明: 充分性. 首先可知存在正交阵$\boldsymbol{P}_1$使得

$$\boldsymbol{P}_1'\boldsymbol{A}\boldsymbol{P}_1 = \mathrm{diag}(\lambda_1 \boldsymbol{I}_{r_1}, \cdots, \lambda_s \boldsymbol{I}_{r_s}) =: \boldsymbol{D},$$

其中$r_1 + \cdots + r_s = n$. 又因为$\boldsymbol{AB} = \boldsymbol{BA}$, 所以有

$$\boldsymbol{P}_1'\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_1'\boldsymbol{B}\boldsymbol{P}_1 = \boldsymbol{P}_1'\boldsymbol{B}\boldsymbol{P}_1\boldsymbol{P}_1'\boldsymbol{A}\boldsymbol{P}_1, \text{ 即} \boldsymbol{DC} = \boldsymbol{CD},$$

其中$\boldsymbol{C} = \boldsymbol{P}_1'\boldsymbol{B}\boldsymbol{P}_1$. 因此, $\boldsymbol{C}$必为分块对角矩阵:

$$\boldsymbol{C} = \mathrm{diag}(\boldsymbol{C}_{r_1}, \cdots, \boldsymbol{C}_{r_s}).$$

注意到$\boldsymbol{C}$是对称矩阵, 所以存在正交矩阵

$$\boldsymbol{P}_2 = \mathrm{diag}(\boldsymbol{P}_{r_1}, \cdots, \boldsymbol{P}_{r_s})$$

使得$\boldsymbol{P}_2'\boldsymbol{C}\boldsymbol{P}_2$为对角阵. 令$\boldsymbol{Q} = \boldsymbol{P}_1\boldsymbol{P}_2$, 则可同时使得$\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}$与$\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{Q}$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, 常常需要求似然函数的极值, 这就要用到矩阵微商. 本节讨论一些常用的结果.

矩阵微商

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

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

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

为y对$\boldsymbol{X}$的微商．

接下来介绍两个最常用的矩阵微商结论.

接下来介绍两个最常用的矩阵微商结论.

例2.5.1 设$\boldsymbol{a}, \boldsymbol{x}$均为$n \times 1$向量, $y = \boldsymbol{a}'\boldsymbol{x}$, 则$\frac{\partial y}{\partial \boldsymbol{x}} = \boldsymbol{a}$.

接下来介绍两个最常用的矩阵微商结论.

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

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

此外, 可知

$$\frac{\partial \boldsymbol{x}'\boldsymbol{a}}{\partial \boldsymbol{x}} = \frac{\partial y}{\partial \boldsymbol{x}} = \boldsymbol{a}.$$

例2.5.2 设$\boldsymbol{A}$为$n \times n$矩阵, $\boldsymbol{x}$为$n \times 1$向量, $y = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$, 则$\frac{\partial y}{\partial \boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}'\boldsymbol{x}$.

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

$$\frac{\partial y}{\partial x_1} = \sum_{i=1}^{n} a_{i1} x_i + \sum_{j=1}^{n} a_{1j} x_j,$$

$$\frac{\partial y}{\partial x_2} = \sum_{i=1}^{n} a_{i2} x_i + \sum_{j=1}^{n} a_{2j} x_j,$$

$$\vdots$$

$$\frac{\partial y}{\partial x_n} = \sum_{i=1}^{n} a_{in} x_i + \sum_{j=1}^{n} a_{nj} x_j.$$

因此, 可以看出

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

若$\boldsymbol{A}$为$n \times n$对称矩阵, 则

$$\frac{\partial \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = 2\boldsymbol{A}\boldsymbol{x}.$$