

第3章 多元正态分布

李高荣

北京师范大学统计学院

E-mail: liqaorong@bnu.edu.cn



1

随机向量

- 随机向量及其分布表示
- 随机向量的数字特征
- 条件分布
- 变量变换

2

多元正态分布

- 多元正态分布的定义及性质
- 条件分布和独立性

3

偏相关系数

4

拉直运算与Kronecker积

5

矩阵多元正态分布

6

多元正态分布样本统计量

7

特征函数*



- 扫描二维码获取在线课件和相关教学电子资源
- 请遵守电子资源使用协议

在多元统计分析中，多元正态分布占有相当重要的地位，主要原因：

- 许多实际问题涉及的随机向量服从正态分布或近似服从正态分布；
- 当样本量很大时，许多统计量的极限分布往往和正态分布有关；
- 对多元正态分布，理论与实践都比较成熟，已有一整套行之有效的统计推断方法。

随机向量及其分布表示

定义3.1: 随机向量

将 p 个随机变量 X_1, X_2, \dots, X_p 的整体称为 p 维随机向量, 记为 $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ 。

定义3.2: 联合分布函数

设 $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ 为 p 维随机向量, 对任意的 p 维实数向量 $\mathbf{x} = (x_1, x_2, \dots, x_p)' \in \mathbb{R}^p$, 称 p 元函数

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, x_2, \dots, x_p) \\ &= \Pr\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p\} \end{aligned}$$

为 \mathbf{X} 的联合分布函数, 并记成 $\mathbf{X} \sim F(\mathbf{x})$, 其中 \mathbb{R}^p 表示 p 维欧氏空间。

- 多维随机向量的统计特性可以通过联合分布函数来完整刻画和描述。

定义3.3: 联合概率分布函数

设 $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ 为 p 维随机向量, 若存在有限个或可列个 p 维向量 $\mathbf{x}_1, \mathbf{x}_2, \dots$, 记 $\Pr(\mathbf{X} = \mathbf{x}_k) = p_k$ ($k = 1, 2, \dots$) 且满足 $p_1 + p_2 + \dots = 1$, 则称 \mathbf{X} 为离散型随机向量, 称 $\Pr(\mathbf{X} = \mathbf{x}_k) = p_k$ ($k = 1, 2, \dots$)为离散型随机向量 \mathbf{X} 的联合概率分布函数。

随机向量及其分布表示

当随机向量 \mathbf{X} 中每一个分量都是连续型随机变量时，我们把 \mathbf{X} 称为连续型随机向量，它可以通过联合概率密度函数来描述。

定义3.4: 联合概率密度函数

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 为 p 维随机向量，若存在非负函数 $f(x_1, \dots, x_p)$ ，使得随机向量 \mathbf{X} 的联合分布函数对一切 $(x_1, \dots, x_p)' \in \mathbb{R}^p$ 均可表示为

$$F(x_1, \dots, x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \dots, u_p) du_1 \cdots du_p,$$

则称 \mathbf{X} 为连续型随机向量，称 $f(x_1, \dots, x_p)$ 为连续型随机向量 \mathbf{X} 的联合概率密度函数，或称为多元密度函数或密度函数。

随机向量及其分布表示

多元密度函数 $f(x_1, x_2, \dots, x_p)$ 具有以下四条性质:

- ① $f(x_1, x_2, \dots, x_p) \geq 0$, 对一切实数 x_1, x_2, \dots, x_p ;
- ② $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_p) dx_1 dx_2 \cdots dx_p = 1$;
- ③ 若 $f(x_1, x_2, \dots, x_p)$ 在 (x_1, x_2, \dots, x_p) 处连续, 则有

$$\frac{\partial^p F(x_1, x_2, \dots, x_p)}{\partial x_1 \partial x_2 \cdots \partial x_p} = f(x_1, x_2, \dots, x_p);$$

- ④ 设 D 是 \mathbb{R}^p 空间的任意一个空间, 则随机向量 X 落在 D 内的概率为

$$\Pr(X \in D) = \int \cdots \int_D f(x_1, x_2, \dots, x_p) dx_1 \cdots dx_p.$$

随机向量及其分布表示

- $X = (X_1, \dots, X_p)'$ 联合分布函数为 $F(x_1, \dots, x_p)$, 其任一个分量 $X^{(m)} = (X_1, \dots, X_m)'$ 也是一个随机向量(其中 $1 \leq m < p$), 也有自己的联合分布函数, 记为 $F_{X^{(m)}}(x_1, \dots, x_m)$, 称为 $X^{(m)}$ 的**边缘分布函数**, 记为

$$\begin{aligned} & \Pr\{X_1 \leq x_1, \dots, X_m \leq x_m\} \\ &= \Pr\{X_1 \leq x_1, \dots, X_m \leq x_m, \mathbf{X}_{m+1} \leq \infty, \dots, X_p \leq \infty\} \\ &= F(x_1, \dots, x_m, \infty, \dots, \infty). \end{aligned}$$

- 由 X 的联合概率密度函数 $f(x_1, \dots, x_p)$, 得到 $X^{(m)}$ 的**边缘概率密度函数**为

$$f_{X^{(m)}}(x_1, \dots, x_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_p) dx_{m+1} \cdots dx_p.$$

随机向量及其分布表示

定义3.5: 相互独立

设 X 和 Y 是两个随机向量, 对一切的 \mathbf{x} 和 \mathbf{y} , 若

$$\Pr(X \leq \mathbf{x}, Y \leq \mathbf{y}) = \Pr(X \leq \mathbf{x}) \Pr(Y \leq \mathbf{y})$$

成立, 则称 X 和 Y 相互独立。若 $F(\mathbf{x}, \mathbf{y})$ 为 (X, Y) 的联合分布函数, $F_X(\mathbf{x})$ 和 $F_Y(\mathbf{y})$ 分别是 X 和 Y 的边缘分布函数, 则 X 和 Y 相互独立当且仅当

$$F(\mathbf{x}, \mathbf{y}) = F_X(\mathbf{x})F_Y(\mathbf{y}).$$

若 $f(\mathbf{x}, \mathbf{y})$ 为 (X, Y) 的联合概率密度函数, $f_X(\mathbf{x})$ 和 $f_Y(\mathbf{y})$ 分别是 X 和 Y 的边缘概率密度函数, 则 X 和 Y 相互独立当且仅当

$$f(\mathbf{x}, \mathbf{y}) = f_X(\mathbf{x})f_Y(\mathbf{y}).$$

思考: 如何定义 X_1, \dots, X_p 的相互独立性?

均值向量

均值向量

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 是 p 维随机向量, 若 $E(X_i) = \mu_i (i = 1, \dots, p)$ 存在, 定义随机向量 \mathbf{X} 的均值为:

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} = \boldsymbol{\mu},$$

其中 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ 是一个 p 维向量, 称为 **均值向量**。

协方差矩阵

定义3.6: 协方差矩阵

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 是 p 维的随机向量, 若 X_i 和 X_j 的协方差 $\text{Cov}(X_i, X_j)$ 存在($i, j = 1, \dots, p$), 则称

$$\begin{aligned}\text{Cov}(\mathbf{X}) &= E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))'] \\ &= \begin{pmatrix} \text{Cov}(X_1, X_1) & \cdots & \text{Cov}(X_1, X_p) \\ \text{Cov}(X_2, X_1) & \cdots & \text{Cov}(X_2, X_p) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_p, X_1) & \cdots & \text{Cov}(X_p, X_p) \end{pmatrix}\end{aligned}$$

为随机向量 \mathbf{X} 的**方差矩阵**或**协方差矩阵**, 记为 $\Sigma = \text{Cov}(\mathbf{X}) = (\sigma_{ij})_{p \times p}$, 其中 $\sigma_{ij} = \text{Cov}(X_i, X_j)$ 。

协方差矩阵

定义3.6(续): 协方差矩阵

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 和 $\mathbf{Y} = (Y_1, \dots, Y_q)'$ 分别是 p 维和 q 维的两个随机向量, 若 X_i 和 Y_j 的协方差 $\text{Cov}(X_i, Y_j)$ 存在($i = 1, \dots, p, j = 1, \dots, q$), 则称

$$\begin{aligned}\text{Cov}(\mathbf{X}, \mathbf{Y}) &= E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{Y} - E(\mathbf{Y}))'] \\ &= \begin{pmatrix} \text{Cov}(X_1, Y_1) & \cdots & \text{Cov}(X_1, Y_q) \\ \text{Cov}(X_2, Y_1) & \cdots & \text{Cov}(X_2, Y_q) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_p, Y_1) & \cdots & \text{Cov}(X_p, Y_q) \end{pmatrix}\end{aligned}$$

为随机向量 \mathbf{X} 和 \mathbf{Y} 的**协方差矩阵**。

- 当 $X = Y$ 和 $p = q$ 时, $\text{Cov}(X, Y) = \text{Cov}(X)$
- 若 $\text{Cov}(X, Y) = \mathbf{0}_{p \times q}$ (其中 $\mathbf{0}_{p \times q}$ 表示 $p \times q$ 的零矩阵), 则称随机向量 X 和 Y 不相关
- 当 X 和 Y 独立时, 则 $\text{Cov}(X, Y) = \mathbf{0}_{p \times q}$, 反之则不成立

相关阵

定义3.7: 相关阵或相关系数矩阵

若 X_i 和 X_j 的协方差 $\text{Cov}(X_i, X_j)$ 存在($i, j = 1, \dots, p$), 称 $\mathbf{R} = (\rho_{ij})_{p \times p}$ 为 X 的**相关阵**, 其中

$$\rho_{ij} = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)}\sqrt{\text{Var}(X_j)}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}, \quad i, j = 1, \dots, p,$$

其中 $\sigma_{ii} = \text{Var}(X_i)$ 为随机变量 X_i 的**方差**, 而 $\sqrt{\sigma_{ii}}$ 为随机变量 X_i 的**标准差**($i = 1, \dots, p$)。

若记 $\mathbf{D}^{1/2} = \text{diag}(\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{pp}})$ 为标准差对角阵, 则

$$\mathbf{R} = \mathbf{D}^{-1/2} \mathbf{\Sigma} \mathbf{D}^{-1/2}, \quad \mathbf{\Sigma} = \mathbf{D}^{1/2} \mathbf{R} \mathbf{D}^{1/2}.$$

- 均值向量和协方差矩阵具有如下的性质：

性质3.1.1

设 \mathbf{X} 和 \mathbf{Y} 是两个随机向量，假定 \mathbf{A} 和 \mathbf{B} 为大小适合运算的两个常数矩阵，则

$$E(\mathbf{AX}) = \mathbf{A}E(\mathbf{X}),$$

$$E(\mathbf{AXB}) = \mathbf{A}E(\mathbf{X})\mathbf{B},$$

$$E(\mathbf{AX} + \mathbf{BY}) = \mathbf{A}E(\mathbf{X}) + \mathbf{B}E(\mathbf{Y}),$$

$$\text{Cov}(\mathbf{AX}) = \mathbf{A}\text{Cov}(\mathbf{X})\mathbf{A}',$$

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

证明： 这里只给出最后一个公式的证明。

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

性质3.1.2

若 \mathbf{X} 和 \mathbf{Y} 是两个相互独立的随机向量, 则 $\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}_{p \times q}$, 其中 p 是 \mathbf{X} 的维数并且 q 是 \mathbf{Y} 的维数。

性质3.1.3

随机向量 $\mathbf{X} = (X_1, \dots, X_p)'$ 的协方差矩阵 $\Sigma = \text{Cov}(\mathbf{X})$ 是对称非负定矩阵。

证明: 因为 $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$, 所以 $\Sigma = \Sigma'$ 。对任给 $\mathbf{a} = (a_1, \dots, a_p)'$, 有

$$\begin{aligned} \mathbf{a}'\Sigma\mathbf{a} &= E[\mathbf{a}'(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))'\mathbf{a}] \\ &= E[(\mathbf{a}'(\mathbf{X} - E(\mathbf{X})))^2] \geq 0, \end{aligned}$$

所以 $\Sigma \geq 0$ ，即证明了 Σ 为非负定矩阵。

性质3.1.4

$\Sigma = \mathbf{L}^2$ ，其中 \mathbf{L} 为非负定矩阵。

证明： 由性质3.1.3可知， Σ 是一个对称非负定矩阵，利用实对称矩阵的对角化定理，存在正交矩阵 Γ ，使得

$$\begin{aligned}\Sigma &= \Gamma \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix} \Gamma' \\ &= \Gamma \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_p} \end{pmatrix} \Gamma' \cdot \Gamma \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_p} \end{pmatrix} \Gamma' = \mathbf{L}^2,\end{aligned}$$

其中 $\mathbf{L} = \Gamma \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_p}) \Gamma'$ ， $\lambda_i \geq 0 (i = 1, \cdots, p)$ ，
且 $\mathbf{L} = \mathbf{L}'$ ，所以 $\mathbf{L} \geq 0$ 。

性质3.1.5

设 X 为 p 维随机向量，期望和协方差矩阵存在，记 $\mu = E(X)$ ， $\Sigma = \text{Cov}(X)$ ， A 为 $p \times p$ 常数矩阵，则

$$E(X'AX) = \text{tr}(A\Sigma) + \mu'A\mu.$$

证明作为练习。

性质3.1.5

设 X 为 p 维随机向量，期望和协方差矩阵存在，记 $\mu = E(X)$ ， $\Sigma = \text{Cov}(X)$ ， A 为 $p \times p$ 常数矩阵，则

$$E(X'AX) = \text{tr}(A\Sigma) + \mu' A \mu.$$

证明作为练习。

证明：一些简单计算有：

$$\begin{aligned} X'AX &= (X - EX)'A(X - EX) + X'AEEX \\ &\quad + (EX)'AX - (EX)'AEX. \end{aligned}$$

两边取数学期望，并简单化简，即完成性质3.1.5的证明。

性质

矩阵性质：

令 \mathbf{A} 是一个 $p \times p$ 对称矩阵， \mathbf{x} 是一个 $p \times 1$ 的向量，则

① $\mathbf{x}'\mathbf{A}\mathbf{x} = \text{tr}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \text{tr}(\mathbf{A}\mathbf{x}\mathbf{x}')$

② $\text{tr}(\mathbf{A}) = \sum_{k=1}^p \lambda_k$ ，其中 λ_k 是矩阵 \mathbf{A} 的特征值。

性质

矩阵性质：

令 \mathbf{A} 是一个 $p \times p$ 对称矩阵， \mathbf{x} 是一个 $p \times 1$ 的向量，则

① $\mathbf{x}'\mathbf{A}\mathbf{x} = \text{tr}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \text{tr}(\mathbf{A}\mathbf{x}\mathbf{x}')$

② $\text{tr}(\mathbf{A}) = \sum_{k=1}^p \lambda_k$ ，其中 λ_k 是矩阵 \mathbf{A} 的特征值。

证明：对(1)，注意到： $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是常数，则 $\mathbf{x}'\mathbf{A}\mathbf{x} = \text{tr}(\mathbf{x}'\mathbf{A}\mathbf{x})$ 。利用结论：对 $m \times p$ 矩阵 \mathbf{B} 和 $p \times m$ 矩阵 \mathbf{C} ，有 $\text{tr}(\mathbf{BC}) = \text{tr}(\mathbf{CB})$ 。令 $\mathbf{B} = \mathbf{x}'$ 且 $m = 1$ ， $\mathbf{C} = \mathbf{A}\mathbf{x}$ ，则完成了(1)的证明。

对(2)，利用谱分解，即 $\mathbf{A} = \mathbf{\Gamma}'\mathbf{\Lambda}\mathbf{\Gamma}$ ，其中 $\mathbf{\Gamma}\mathbf{\Gamma}' = \mathbf{I}$ 和 $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \dots, \lambda_p\}$ ，则

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{\Gamma}'\mathbf{\Lambda}\mathbf{\Gamma}) = \text{tr}(\mathbf{\Lambda}\mathbf{\Gamma}\mathbf{\Gamma}') = \text{tr}(\mathbf{\Lambda}) = \sum_{k=1}^p \lambda_k.$$

- 假设 $f(x_1, \dots, x_p)$ 是随机向量 \mathbf{X} 的联合概率密度函数, 则联合矩为

$$E(X_1^{h_1} \cdots X_p^{h_p}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{h_1} \cdots x_p^{h_p} f(x_1, \dots, x_p) dx_1 \cdots dx_p.$$

- 对于 $m < p$, 变量子集 X_1, \dots, X_m 的联合矩为

$$E(X_1^{h_1} \cdots X_m^{h_m}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{h_1} \cdots x_m^{h_m} f_{\mathbf{X}^{(m)}}(x_1, \dots, x_m) dx_1 \cdots dx_m.$$

- 如果 X_1, \dots, X_p 相互独立, 则

$$E(X_1^{h_1} \cdots X_p^{h_p}) = \prod_{i=1}^p E(X_i^{h_i}).$$

- 二维随机向量 (X, Y) , X 与 Y 的相互关系除了独立性以外, 还有相依关系, 即随机变量的取值往往彼此是有影响的, 这种关系用条件分布能更好地表达出来。
- 设给定二维随机向量 (X, Y) , 假设事件 $A = \{\omega : x_1 \leq X \leq x_2\}$, $B = \{\omega : y_1 \leq Y \leq y_2\}$, 则给定事件 B 的条件下, 事件 A 发生的条件概率为:

$$\begin{aligned}\Pr(A|B) &= \Pr(x_1 \leq X \leq x_2 | y_1 \leq Y \leq y_2) \\ &= \frac{\Pr(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2)}{\Pr(y_1 \leq Y \leq y_2)}\end{aligned}$$

离散型条件概率

如果二维离散随机向量 (X, Y) 的联合分布列为: $p_{ij} = \Pr(X = x_i, Y = y_j), i = 1, 2, \dots; j = 1, 2, \dots$ 。对一切使得 $\Pr(Y = y_j) = \sum_{i=1}^{\infty} p_{ij} = p_{\cdot j} > 0$ 的 y_j , 则称

$$p_{i|j} = \Pr(X = x_i | Y = y_j) = \frac{\Pr(X = x_i, Y = y_j)}{\Pr(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}$$

为在给定 $Y = y_j$ 条件下 X 的**条件分布列**, 其中 $i = 1, 2, \dots$ 。

条件分布函数与条件密度函数

设 (X, Y) 为连续型随机向量，联合密度函数为 $f(x, y)$ ，边际分布函数分别为 $f_X(x)$ 和 $f_Y(y)$ 。对一切 $f_Y(y) > 0$ 的 y ，在给定 $Y = y$ 条件下， X 的**条件分布函数**和**条件密度函数**分别为：

$$F(x|y) = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} du, \quad f(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

对一切 $f_X(x) > 0$ 的 x ，在给定 $X = x$ 条件下， Y 的**条件分布函数**和**条件密度函数**分别为：

$$F(y|x) = \int_{-\infty}^y \frac{f(x, v)}{f_X(x)} dv, \quad f(y|x) = \frac{f(x, y)}{f_X(x)}.$$

条件分布

一般而言, 随机向量 $\mathbf{X} = (X_1, \dots, X_p)'$ 的联合分布函数为 $F(x_1, \dots, x_p)$, 联合密度函数为 $f(x_1, \dots, x_p)$, 则

- 给定 $X_{m+1} = x_{m+1}, \dots, X_p = x_p$ 条件下, $(X_1, \dots, X_m)'$ 的条件分布函数为:

$$\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} \frac{f(u_1, \dots, u_m, x_{m+1}, \dots, x_p)}{f_{\mathbf{X}^{(-m)}}(x_{m+1}, \dots, x_p)} du_1 \cdots du_m.$$

- 给定 $X_{m+1} = x_{m+1}, \dots, X_p = x_p$ 条件下, $(X_1, \dots, X_m)'$ 的条件概率密度函数为:

$$\frac{f(x_1, \dots, x_p)}{f_{\mathbf{X}^{(-m)}}(x_{m+1}, \dots, x_p)},$$

其中 $f_{\mathbf{X}^{(-m)}}(x_{m+1}, \dots, x_p)$ 为 $\mathbf{X}^{(-m)} = (X_{m+1}, \dots, X_p)'$ 的边缘密度函数。

变量变换

令随机向量 $(X_1, \dots, X_p)'$ 的联合密度函数为 $f(x_1, \dots, x_p)$, 定义下面的 p 实值函数:

$$y_i = y_i(x_1, \dots, x_p), \quad i = 1, \dots, p.$$

满足:

- 假设从 x 空间到 y 空间的变换是一一对应的;
- 存在逆变换: $x_i = x_i(y_1, \dots, y_p), i = 1, \dots, p$;
- 定义**Jacobian(雅可比)矩阵** $J(y_1, \dots, y_p)$ 如下:

$$J(y_1, \dots, y_p) = \text{mod} \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots & \frac{\partial x_1}{\partial y_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_p}{\partial y_1} & \frac{\partial x_p}{\partial y_2} & \dots & \frac{\partial x_p}{\partial y_p} \end{vmatrix},$$

其中 mod 是指**模数(modulus)或绝对值**。

多元函数的密度变换公式:

令随机变量 Y_1, \dots, Y_p 定义为:

$$Y_i = y_i(X_1, \dots, X_p), \quad i = 1, \dots, p.$$

则 Y_1, \dots, Y_p 的概率密度函数为:

$$g(y_1, \dots, y_p) = f\left[x_1(y_1, \dots, y_p), \dots, x_p(y_1, \dots, y_p)\right] \\ \times J(y_1, \dots, y_p).$$

应用: 如果给定了 X_1, \dots, X_p 的联合密度函数, 以及 Y 和 X 的函数关系, 则可用上式计算 Y_1, \dots, Y_p 落在一个给定区域 S 中的概率。

- 设 Y_1, \dots, Y_p 为独立同分布(i.i.d.)的随机变量序列
- $Y_i \sim N(0, 1)$, $i = 1, \dots, p$, 则随机向量 $\mathbf{Y} = (Y_1, \dots, Y_p)'$ 的概率密度函数为:

$$\begin{aligned} f(y_1, \dots, y_p) &= \prod_{i=1}^p \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} y_i^2 \right\} \\ &= (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} \mathbf{y}' \mathbf{y} \right\}, \end{aligned}$$

其中 $\mathbf{y} = (y_1, \dots, y_p)'$ 。

设 $\mathbf{Y} \sim N_p(\mathbf{0}, \mathbf{I}_p)$, 则有

- $E(\mathbf{Y}) = \mathbf{0}$, $\text{Cov}(\mathbf{Y}) = \mathbf{I}_p$;
- $E(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \text{tr}(\mathbf{A})$;
- 设 \mathbf{a} 为 p 元向量, \mathbf{A} 为对称矩阵, 则

$$\text{Cov}(\mathbf{a}'\mathbf{Y}, \mathbf{Y}'\mathbf{A}\mathbf{Y}) = 0;$$

- 设 \mathbf{A}, \mathbf{B} 为对称矩阵, 则

$$\text{Cov}(\mathbf{Y}'\mathbf{A}\mathbf{Y}, \mathbf{Y}'\mathbf{B}\mathbf{Y}) = 2\text{tr}(\mathbf{AB}).$$

证明：由 $Y \sim N_p(\mathbf{0}, \mathbf{I}_p)$, $Y' \mathbf{A} Y = \sum_{i,j} a_{ij} Y_i Y_j$ 和 $Y' \mathbf{B} Y = \sum_{k,l} b_{kl} Y_k Y_l$, 以及

$$E(Y_i Y_j Y_k Y_l) = \begin{cases} 3, & i = j = k = l; \\ 1, & i = j \neq k = l \\ & i = k \neq j = l \\ & i = l \neq k = j; \\ 0, & \text{其他.} \end{cases}$$

则

$$\begin{aligned} E(Y' \mathbf{A} Y Y' \mathbf{B} Y) &= 3 \sum_{i=1}^p a_{ii} b_{ii} + \sum_{1 \leq i \neq k \leq p} a_{ii} b_{kk} + \sum_{1 \leq i \neq j \leq p} a_{ij} b_{ij} + \sum_{1 \leq i \neq k \leq p} a_{ik} b_{ki} \\ &= 3 \sum_{i=1}^p a_{ii} b_{ii} + \sum_{1 \leq i \neq k \leq p} a_{ii} b_{kk} + 2 \sum_{1 \leq i \neq j \leq p} a_{ij} b_{ij}. \end{aligned}$$

注意到

$$\sum_{1 \leq i \neq k \leq p} a_{ii} b_{kk} = \sum_{k=1}^p \sum_{i=1}^p a_{ii} b_{kk} - \sum_{i=1}^p a_{ii} b_{ii} = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B}) - \sum_{i=1}^p a_{ii} b_{ii},$$

$$\sum_{1 \leq i \neq j \leq p} a_{ij} b_{ij} = \sum_{j=1}^p \sum_{i=1}^p a_{ij} b_{ij} - \sum_{i=1}^p a_{ii} b_{ii} = \text{tr}(\mathbf{AB}) - \sum_{i=1}^p a_{ii} b_{ii}.$$

则

$$E(\mathbf{Y}' \mathbf{A} \mathbf{Y} \mathbf{Y}' \mathbf{B} \mathbf{Y}) = 2\text{tr}(\mathbf{AB}) + \text{tr}(\mathbf{A})\text{tr}(\mathbf{B}).$$

最后可得到：

$$\begin{aligned} \text{Cov}(\mathbf{Y}' \mathbf{A} \mathbf{Y}, \mathbf{Y}' \mathbf{B} \mathbf{Y}) &= E(\mathbf{Y}' \mathbf{A} \mathbf{Y} \mathbf{Y}' \mathbf{B} \mathbf{Y}) - E(\mathbf{Y}' \mathbf{A} \mathbf{Y})E(\mathbf{Y}' \mathbf{B} \mathbf{Y}) \\ &= 2\text{tr}(\mathbf{AB}). \end{aligned}$$

多元正态分布的定义

- 多元统计分析的主要理论都是建立在多元正态分布总体的基础上，因此多元正态分布是多元统计分析的基础。
- **思考：**如何给出多元正态分布的定义？

多元正态分布的定义

- 多元统计分析的主要理论都是建立在多元正态分布总体的基础上，因此多元正态分布是多元统计分析的基础。
- 思考：**如何给出多元正态分布的定义？

定义3.8: 多元正态分布

设 Y_1, \dots, Y_q 为i.i.d.的随机变量序列, $Y_i \sim N(0, 1), i = 1, \dots, q$, \mathbf{A} 是 $p \times q$ 常数矩阵, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ 为 $p \times 1$ 常数向量。记 $\mathbf{Y} = (Y_1, \dots, Y_q)'$, 则称随机向量 \mathbf{Y} 的线性组合 $\mathbf{X} = \mathbf{A}\mathbf{Y} + \boldsymbol{\mu}$ 为

p 元(维)正态分布

, 记成 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 其中 $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$ 。

多元正态分布的定义

可以从联合密度函数定义多元正态分布，定义如下：

定义3.9: 基于联合密度函数的多元正态分布定义

设 \mathbf{X} 是一个 p 维随机向量，具有联合概率密度函数

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

其中 $\mathbf{x} = (x_1, \dots, x_p)'$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ 是 p 维实向量， Σ 是 p 阶正定矩阵， $|\Sigma|$ 为 Σ 的行列式值， Σ^{-1} 为 Σ 的逆矩阵，则称 \mathbf{X} 服从均值向量为 $\boldsymbol{\mu}$ ，协方差阵为 Σ 的

多元正态分布

，记成 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ 。

作业: 如何推导得到上面的多元正态概率密度函数？

多元正态分布的定义

可以从联合密度函数定义多元正态分布，定义如下：

定义3.9: 基于联合密度函数的多元正态分布定义

设 \mathbf{X} 是一个 p 维随机向量，具有联合概率密度函数

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

其中 $\mathbf{x} = (x_1, \dots, x_p)'$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ 是 p 维实向量， Σ 是 p 阶正定矩阵， $|\Sigma|$ 为 Σ 的行列式值， Σ^{-1} 为 Σ 的逆矩阵，则称 \mathbf{X} 服从均值向量为 $\boldsymbol{\mu}$ ，协方差阵为 Σ 的

多元正态分布

，记成 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ 。

作业: 如何推导得到上面的多元正态概率密度函数?
(参考Anderson (2003)的2.3节。)

常数密度轮廓线

由定义3.9, 多元正态分布密度函数在

$$(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = c^2 \quad (c \text{ 为常数})$$

上是常数。因此, 在 p 维欧氏空间上, 所有满足上式的 \mathbf{x} 称为**常数密度函数轮廓线**, 它是以 $\boldsymbol{\mu}$ 为中心, $\pm c\sqrt{\lambda_i}\boldsymbol{\phi}_i$ ($i = 1, \dots, p$)为轴的椭圆面。

- λ_i ($i = 1, \dots, p$)为 $\boldsymbol{\Sigma}$ 的特征值
- $\boldsymbol{\phi}_i$ 为 λ_i 对应的单位特征向量, $\boldsymbol{\Sigma}\boldsymbol{\phi}_i = \lambda_i\boldsymbol{\phi}_i$
- 椭球的形状和方向由 $\boldsymbol{\Sigma}$ 决定
- 椭球的大小由常数 c^2 决定
- 如果 $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}_p$, 则定义3.9给出的密度函数为球密度函数

多元正态分布的性质

多元正态分布的性质：

性质3.2.1: 均值和协方差阵

设 $X \sim N_p(\mu, \Sigma)$, 则 $E(X) = \mu$, $\text{Cov}(X) = \Sigma$ 。

性质3.2.2: 特征函数

设 $X \sim N_p(\mu, \Sigma)$, 则

$$\varphi_X(\mathbf{t}) = \exp \left\{ i\mathbf{t}'\mu - \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t} \right\},$$

其中 $\varphi_X(\mathbf{t}) = \varphi_X(t_1, \dots, t_p)$ 为多元正态随机向量 X 的**特征函数**, $\mathbf{t} = (t_1, \dots, t_p)'$, $i = \sqrt{-1}$ 。

多元正态分布的性质

证明： 令 $\mathbf{z} = (z_1, \dots, z_q)'$, 标准正态随机向量 \mathbf{Y} 的特征函数为 $\varphi_{\mathbf{Y}}(\mathbf{z}) = \varphi_{\mathbf{Y}}(z_1, \dots, z_q)$, 其中 $\varphi_j(z_j)$ 为 Y_j 的特征函数, $j = 1, \dots, q$, 则

$$\varphi_{\mathbf{Y}}(\mathbf{z}) = \prod_{j=1}^q \varphi_j(z_j) = \prod_{j=1}^q \exp \left\{ -\frac{1}{2} z_j^2 \right\} = \exp \left\{ -\frac{1}{2} \mathbf{z}' \mathbf{z} \right\}.$$

由随机向量特征函数性质, 得 $\mathbf{X} = \mathbf{A}\mathbf{Y} + \boldsymbol{\mu}$ 的特征函数为

$$\begin{aligned} \varphi_{\mathbf{X}}(\mathbf{t}) &= \exp\{\mathbf{i}\mathbf{t}'\boldsymbol{\mu}\} \varphi_{\mathbf{Y}}(\mathbf{A}'\mathbf{t}) \\ &= \exp\{\mathbf{i}\mathbf{t}'\boldsymbol{\mu}\} \exp \left\{ -\frac{1}{2} \mathbf{t}' \mathbf{A} \mathbf{A}' \mathbf{t} \right\} \\ &= \exp \left\{ \mathbf{i}\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2} \mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} \right\}. \end{aligned}$$

多元正态分布的定义

- 由性质3.2.2可知, 特征函数 $\varphi_X(\mathbf{t})$ 只与 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 有关。
- 由随机向量与其特征函数一一对应关系可知, p 元正态分布 $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 又可用特征函数定义如下。

定义3.10: 基于特征函数的多元正态分布定义

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 为 p 维随机向量, 若其特征函数为

$$\varphi_X(\mathbf{t}) = \exp \left\{ i \mathbf{t}' \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} \right\},$$

其中 $\mathbf{t} = (t_1, \dots, t_p)'$, $i = \sqrt{-1}$ 。则称 \mathbf{X} 服从

p 元正态分布

, 记成 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 。

多元正态分布的性质

性质3.2.3: 线性变换

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{B} 为 $q \times p$ 的常数矩阵, $\boldsymbol{\theta}$ 为 $q \times 1$ 常向量, 令 $\mathbf{Z} = \mathbf{B}\mathbf{X} + \boldsymbol{\theta}$, 则 $\mathbf{Z} \sim N_q(\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\theta}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$ 。

多元正态分布的性质

性质3.2.3: 线性变换

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{B} 为 $q \times p$ 的常数矩阵, $\boldsymbol{\theta}$ 为 $q \times 1$ 常向量, 令 $\mathbf{Z} = \mathbf{B}\mathbf{X} + \boldsymbol{\theta}$, 则 $\mathbf{Z} \sim N_q(\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\theta}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$ 。

证明: 令 $\mathbf{t} = (t_1, \dots, t_p)'$, $\mathbf{s} = (s_1, \dots, s_q)'$ 。由定义3.10 知: 随机向量 \mathbf{X} 的特征函数为

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp \left\{ i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t} \right\}.$$

由特征函数的性质，得 $\mathbf{Z} = \mathbf{B}\mathbf{X} + \boldsymbol{\theta}$ 的特征函数为

$$\begin{aligned}\varphi_{\mathbf{Z}}(\mathbf{s}) &= \exp\{\mathbf{i}\mathbf{s}'\boldsymbol{\theta}\} \varphi_{\mathbf{X}}(\mathbf{B}'\mathbf{s}) \\ &= \exp\{\mathbf{i}\mathbf{s}'\boldsymbol{\theta}\} \exp\left\{\mathbf{i}\mathbf{s}'\mathbf{B}\boldsymbol{\mu} - \frac{1}{2}\mathbf{s}'\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'\mathbf{s}\right\} \\ &= \exp\left\{\mathbf{i}\mathbf{s}'(\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\theta}) - \frac{1}{2}\mathbf{s}'\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'\mathbf{s}\right\}.\end{aligned}$$

由定义3.10可知， $\mathbf{Z} \sim N_q(\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\theta}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$ 。

性质3.2.4:

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 为 p 维随机向量, 则 \mathbf{X} 服从 p 元正态分布 $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 的充分必要条件是它的任意一个线性组合 $\mathbf{a}'\mathbf{X}$ 都服从一元正态分布, 其中 $\mathbf{a} \neq \mathbf{0}$ 。

多元正态分布的性质

性质3.2.4:

设 $\mathbf{X} = (X_1, \dots, X_p)'$ 为 p 维随机向量, 则 \mathbf{X} 服从 p 元正态分布 $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 的充分必要条件是它的任意一个线性组合 $\mathbf{a}'\mathbf{X}$ 都服从一元正态分布, 其中 $\mathbf{a} \neq \mathbf{0}$ 。

证明: “ \Rightarrow ”显然。现在证明“ \Leftarrow ”。对任意非零向量 $\mathbf{a} \in \mathbb{R}^p$, 由于 $\mathbf{a}'\mathbf{X}$ 服从一元正态分布, 故 \mathbf{X} 的二阶矩存在, 记 $E(\mathbf{X}) = \boldsymbol{\mu}$, $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$, 则 $E(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\boldsymbol{\mu}$, $\text{Var}(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}$ 。注意到 $\mathbf{a}'\mathbf{X}$ 的特征函数

$$\varphi_{\mathbf{a}'\mathbf{X}}(t) = \exp \left\{ i t \mathbf{a}'\boldsymbol{\mu} - \frac{1}{2} t^2 \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} \right\}, \quad t \in \mathbb{R}.$$

令 $t = 1$, 有 $\varphi_{\mathbf{a}'\mathbf{X}}(1) = \exp \left\{ i \mathbf{a}'\boldsymbol{\mu} - \frac{1}{2} \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} \right\} = \varphi_{\mathbf{X}}(\mathbf{a})$ 。视其为 \mathbf{a} 的函数, 则由定义3.10, 可知 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 。

性质3.2.5: 可加性

设 X_1, \dots, X_k 是相互独立的 k 组 p 维随机向量, $X_i \sim N_p(\mu_i, \Sigma_i), i = 1, \dots, k$, 则

$$\sum_{i=1}^k X_i \sim N_p \left(\sum_{i=1}^k \mu_i, \sum_{i=1}^k \Sigma_i \right).$$

多元正态分布的性质

性质3.2.6: 二次型

设 $X \sim N_p(\mu, \Sigma)$, $\Sigma > 0$, 对于 p 阶矩阵 $A \geq 0$, 则二次型 $(X - \mu)'A(X - \mu)$ 服从中心 χ_m^2 分布的充要条件为 $A\Sigma A\Sigma = A\Sigma$, 且其自由度 $m = \text{tr}(A\Sigma)$ 。特别地, 在 $A = \Sigma^{-1}$ 时, $(X - \mu)'\Sigma^{-1}(X - \mu)$ 服从自由度为 p 的中心 χ^2 分布。

证明留给大家课后作业。

- `install.packages("mvtnorm")` # 安装程序包
- `library(mvtnorm)` # 加载程序包
- `dmvnorm()` 计算多元正态密度函数
- `pmvnorm()` 计算分布函数
- `qmvnorm()` 计算分位数
- `rmvnorm()` 产生多元正态分布的随机数

二元正态分布

二元正态分布

设二维随机向量 $\mathbf{Z} = (X, Y)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 或记为 $(X, Y)' \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 其中 $\sigma_1 > 0, \sigma_2 > 0, |\rho| < 1$, 并且

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} =: \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} > 0.$$

二维正态随机向量 $(X, Y)'$ 的联合密度函数为

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi|\boldsymbol{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \right] \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right) \right\}, \end{aligned}$$

其中 $\mathbf{z} = (x, y)'$ 。

二元正态分布

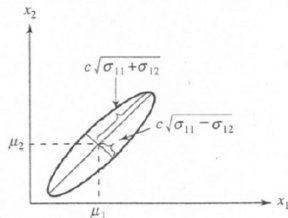
例1: 对二元正态 $N_2 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right)$, 当 $\sigma_{11} = \sigma_{22}$ 时, 求其常数密度轮廓线椭圆。

解: 由 $|\lambda \mathbf{I}_2 - \Sigma| = 0$ 和 $\Sigma \phi = \lambda \phi$, 则易得其特征值和特征向量为

$$\lambda_1 = \sigma_{11} + \sigma_{12}, \phi_1 = (1, 1)' / \sqrt{2}$$

$$\lambda_2 = \sigma_{11} - \sigma_{12}, \phi_2 = (1, -1)' / \sqrt{2}$$

Bivariate Normal Contours ($\sigma_{11} = \sigma_{22}$)



课下习题:

- 假设 $p = 2, \mathbf{X} = (X_1, X_2)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- 其中 $\boldsymbol{\mu} = (1, 2)'$, $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$
- **问题:** 用R语言写程序, 根据上面二元正态分布产生100个随机数, 画出密度轮廓线椭圆。

二元正态分布

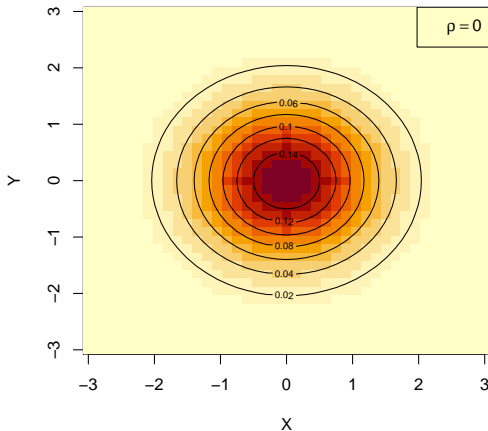
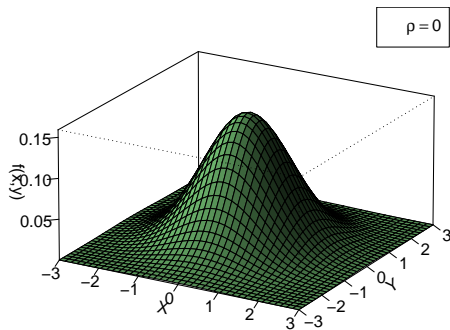


Figure 1: $\sigma_1^2 = 1, \sigma_2^2 = 1, \rho = 0$ 时的二元正态密度函数及其等高线图。

二元正态分布

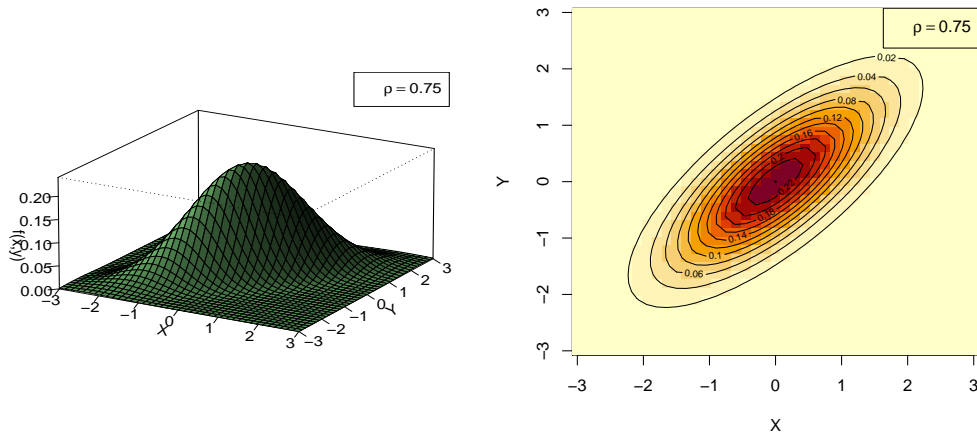


Figure 2: $\sigma_1^2 = 1, \sigma_2^2 = 1, \rho = 0.75$ 时的二元正态密度函数及其等高线图。

二元正态分布

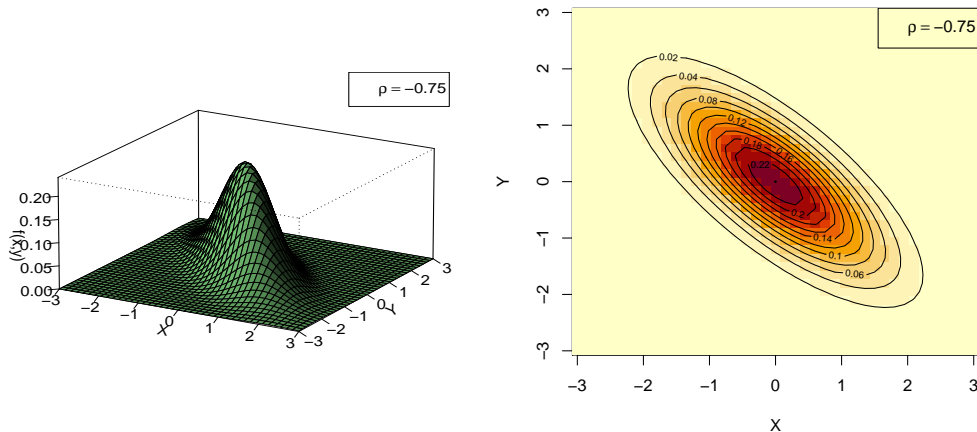


Figure 3: $\sigma_1^2 = 1, \sigma_2^2 = 1, \rho = -0.75$ 时的二元正态密度函数及其等高线图。

设 $X \sim N_p(\mu, \Sigma)$, $p \geq 2$, 将 X, μ 和 Σ 剖分为

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad (1)$$

其中

- $X^{(1)}$ 和 $\mu^{(1)}$ 为 $q \times 1$ 的向量
- Σ_{11} 为 $q \times q$ 矩阵
- $X^{(2)}$ 和 $\mu^{(2)}$ 为 $(p - q) \times 1$ 的向量
- Σ_{22} 为 $(p - q) \times (p - q)$ 矩阵
- $\Sigma_{12} = \Sigma'_{21}$ 为 $q \times (p - q)$ 矩阵

性质3.2.7: 边际分布

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), p \geq 2$, 则

$$\mathbf{X}^{(1)} \sim N_q(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}_{11}), \quad \mathbf{X}^{(2)} \sim N_{p-q}(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}_{22}).$$

性质3.2.7: 边际分布

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), p \geq 2$, 则

$$\mathbf{X}^{(1)} \sim N_q(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}_{11}), \quad \mathbf{X}^{(2)} \sim N_{p-q}(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}_{22}).$$

证明1: 令 $\mathbf{X}^{(1)} = \mathbf{B}^{(1)}\mathbf{X} + \boldsymbol{\theta}^{(1)}$ 和 $\mathbf{X}^{(2)} = \mathbf{B}^{(2)}\mathbf{X} + \boldsymbol{\theta}^{(2)}$, 其中取

$$\mathbf{B}_{q \times p}^{(1)} = (\mathbf{I}_q, \mathbf{0}_{q \times (p-q)}), \quad \boldsymbol{\theta}^{(1)} = \mathbf{0},$$

$$\mathbf{B}_{(p-q) \times p}^{(2)} = (\mathbf{0}_{(p-q) \times q}, \mathbf{I}_{p-q}), \quad \boldsymbol{\theta}^{(2)} = \mathbf{0}.$$

由性质3.2.3, 即可证明性质3.2.7。

证明2: 由分块矩阵的行列式和逆公式, 有

$$|\Sigma| = |\Sigma_{11}| |\Sigma_{22.1}|,$$

$$\Sigma^{-1} = \begin{pmatrix} \mathbf{I}_q & -\Sigma_{11}^{-1}\Sigma_{12} \\ \mathbf{0} & \mathbf{I}_{p-q} \end{pmatrix} \begin{pmatrix} \Sigma_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22.1}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ -\Sigma_{21}\Sigma_{11}^{-1} & \mathbf{I}_{p-q} \end{pmatrix},$$

其中 $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ 。则 \mathbf{X} 的密度函数 $f(\mathbf{x}) = f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ 为:

$$f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = f_1(\mathbf{x}^{(1)})f(\mathbf{x}^{(2)}|\mathbf{x}^{(1)}),$$

其中

$$f_1(\mathbf{x}^{(1)}) = \frac{1}{(2\pi)^{q/2}|\Sigma_{11}|^{1/2}} \exp \left[-\frac{1}{2} \left(\mathbf{x}^{(1)} - \boldsymbol{\mu}^{(1)} \right)' \Sigma_{11}^{-1} \left(\mathbf{x}^{(1)} - \boldsymbol{\mu}^{(1)} \right) \right],$$

$$f(\mathbf{x}^{(2)}|\mathbf{x}^{(1)}) = \frac{1}{(2\pi)^{(p-q)/2} |\Sigma_{22.1}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}^{(2)} - \boldsymbol{\mu}_{2.1})' \Sigma_{22.1}^{-1} (\mathbf{x}^{(2)} - \boldsymbol{\mu}_{2.1}) \right],$$

其中 $\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}^{(2)} + \Sigma_{21} \Sigma_{11}^{-1} (\mathbf{x}^{(1)} - \boldsymbol{\mu}^{(1)})$ 。可知 $f_1(\mathbf{x}^{(1)})$ 是 q 元正态分布 $N_q(\boldsymbol{\mu}^{(1)}, \Sigma_{11})$ 的概率密度函数；此外，如果把 $\mathbf{x}^{(1)}$ 当作一个常数，则 $f(\mathbf{x}^{(2)}|\mathbf{x}^{(1)})$ 是 $p - q$ 元正态分布 $N_{p-q}(\boldsymbol{\mu}_{2.1}, \Sigma_{22.1})$ 的概率密度函数。由边缘概率密度函数的定义， $\mathbf{X}^{(1)}$ 的边缘密度函数为：

$$\int f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) d\mathbf{x}^{(2)} = f_1(\mathbf{x}^{(1)}) \int f(\mathbf{x}^{(2)}|\mathbf{x}^{(1)}) d\mathbf{x}^{(2)} = f_1(\mathbf{x}^{(1)}).$$

由此可证明: $\mathbf{X}^{(1)} \sim N_q(\boldsymbol{\mu}^{(1)}, \Sigma_{11})$ 。

思考： 考虑用特征函数证明性质3.2.7？

定理3.2.1

设 $X \sim N_p(\mu, \Sigma)$, $\Sigma > 0$, 对 X 和 Σ 作剖分, 则

① 给定 $X^{(2)} = \mathbf{x}^{(2)}$ 时, $X^{(1)}$ 的条件分布服从 q 元正态分布, 即

$$(X^{(1)} | X^{(2)} = \mathbf{x}^{(2)}) \sim N_q(\mu_{1.2}, \Sigma_{11.2}),$$

其中 $\mu_{1.2} = \mu^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}^{(2)} - \mu^{(2)})$, $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ 。

② 给定 $X^{(1)} = \mathbf{x}^{(1)}$ 时, $X^{(2)}$ 的条件分布服从 $p - q$ 元正态分布, 即

$$(X^{(2)} | X^{(1)} = \mathbf{x}^{(1)}) \sim N_{p-q}(\mu_{2.1}, \Sigma_{22.1}),$$

其中 $\mu_{2.1} = \mu^{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x}^{(1)} - \mu^{(1)})$, $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ 。

条件分布和独立性

证明：由性质3.2.7的证明，以及条件密度定义及分块矩阵的逆易证。

由性质3.2.7和定理3.2.1可知：

- $\mathbf{X}^{(1)}$ 和 $(\mathbf{X}^{(1)}|\mathbf{X}^{(2)})$ 的分布都服从正态分布，它们的协方差阵分别为 Σ_{11} 和 $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ ；
- 由于 $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \geq 0$ ，故 $\Sigma_{11} \geq \Sigma_{11.2}$ ，等号成立当且仅当 $\Sigma_{12} = \Sigma'_{21} \equiv \mathbf{0}$ ；
- 协方差阵描述变量之间关系及散布程度的， $\Sigma_{11} \geq \Sigma_{11.2}$ 表明：在已知 $\mathbf{X}^{(2)}$ 的条件下，减少了 $\mathbf{X}^{(1)}$ 的散布程度，而且只有当 $\Sigma_{12} = \mathbf{0}$ 时，两者相等；
- $\Sigma_{12} = \mathbf{0}$ 等价于 $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立。独立性表明，即使给定 $\mathbf{X}^{(2)}$ 的条件下，对 $\mathbf{X}^{(1)}$ 的分布没有影响。

条件分布和独立性

在二元正态场合 $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right)$,

则 $X_1|X_2 = x_2 \sim N \left(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}} (x_2 - \mu_2), \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} \right)$ 。

注: $\sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} = \sigma_{11}(1 - \rho^2)$. $|\rho|$ 越大, 方差越小。

例2

假设大学生男生的体重(lbs) 和身高(inches, 英寸) 服从多元正态分布, 均值 $\mu = (175, 71)'$ 和协方差矩阵 $\Sigma = \begin{pmatrix} 550 & 40 \\ 40 & 8 \end{pmatrix}$, 试求给定身高为 $X_2 = x_2$ 条件下体重 X_1 的分布。

解: 由定理3.2.1, $X_1|X_2 = x_2 \sim N(-180 + 5x_2, 350)$, 例如当身高为 $x_2 = 70$ 时, 体重则服从 $N(170, 350)$ 。

定理3.2.2:

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, 对 \mathbf{X} 和 $\boldsymbol{\Sigma}$ 作(1)的剖分, 则 $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立当且仅当 $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ 。

定理3.2.2:

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$, 对 \mathbf{X} 和 $\boldsymbol{\Sigma}$ 作(1)的剖分, 则 $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立当且仅当 $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ 。

证明: “ \Rightarrow ” 当 $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立时, 显然

$$\text{Cov}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) = \boldsymbol{\Sigma}_{12} = \mathbf{0}.$$

“ \Leftarrow ” 当 $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ 时, 由性质3.2.7的证明可知, $\boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22}$, 则

$$f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = f_{\mathbf{X}^{(1)}}(\mathbf{x}^{(1)})f_{\mathbf{X}^{(2)}}(\mathbf{x}^{(2)}).$$

可知, $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立。

思考: 对任何分布, 定理3.2.2成立吗?

定理3.2.3:

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, 对 \mathbf{X} 和 $\boldsymbol{\Sigma}$ 作类似(1)的剖分:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(m)} \end{pmatrix} \begin{matrix} q_1 \\ \vdots \\ q_m \end{matrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1m} \\ \vdots & & \vdots \\ \boldsymbol{\Sigma}_{m1} & \cdots & \boldsymbol{\Sigma}_{mm} \end{pmatrix} \begin{matrix} q_1 \\ \vdots \\ q_m \end{matrix},$$

其中 $q_1 + \cdots + q_m = p$, 则 $\text{Cov}(\mathbf{X}^{(i)}, \mathbf{X}^{(j)}) = \boldsymbol{\Sigma}_{ij} = \mathbf{0} (1 \leq i < j \leq m)$ 都成立, 是 $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)}$ 相互独立的充要条件。

推论3.2.1:

在定理3.2.1的条件下, 若 $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互独立, 即 $\Sigma_{12} = \Sigma'_{21} = \mathbf{0}$, 则存在 q 维实数向量 $\mathbf{a} = (a_1, \dots, a_q)'$ 和 $p - q$ 维的实数向量 $\mathbf{b} = (b_1, \dots, b_{p-q})'$, 使得变量 $\xi = \mathbf{a}'\mathbf{X}^{(1)}$ 和 $\eta = \mathbf{b}'\mathbf{X}^{(2)}$ 相互独立。

推论3.2.2:

设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, $\Sigma > 0$, 若 Σ 为对角矩阵, 则 X_1, \dots, X_p 相互独立。

定理3.2.4:

在定理3.2.1的条件下, 则

(1) $X^{(1)}$ 与 $X^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}X^{(1)}$ 相互独立, 并且

$$X^{(1)} \sim N_q(\mu^{(1)}, \Sigma_{11}),$$

$$X^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}X^{(1)} \sim N_{p-q}(\mu^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\mu^{(1)}, \Sigma_{22.1});$$

(2) $X^{(2)}$ 与 $X^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}X^{(2)}$ 相互独立, 并且

$$X^{(2)} \sim N_{p-q}(\mu^{(2)}, \Sigma_{22}),$$

$$X^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}X^{(2)} \sim N_q(\mu^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\mu^{(2)}, \Sigma_{11.2}).$$

证明： 令

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ -\Sigma_{21}\Sigma_{11}^{-1} & \mathbf{I}_{p-q} \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{pmatrix},$$

则 $\mathbf{Y}^{(1)} = \mathbf{X}^{(1)}$, $\mathbf{Y}^{(2)} = \mathbf{X}^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X}^{(1)}$ 。由于 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, 利用性质3.2.3, 通过简单计算有

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \end{pmatrix} \sim N_p \left(\begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}^{(1)} \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22.1} \end{pmatrix} \right).$$

由定理3.2.2知, $\mathbf{Y}^{(1)} = \mathbf{X}^{(1)}$ 和 $\mathbf{Y}^{(2)} = \mathbf{X}^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X}^{(1)}$ 相互独立, 且

$$\mathbf{X}^{(1)} \sim N_q(\boldsymbol{\mu}^{(1)}, \Sigma_{11}),$$

$$\mathbf{X}^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{X}^{(1)} \sim N_{p-q}(\boldsymbol{\mu}^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}^{(1)}, \Sigma_{22.1}).$$

由定理3.2.1, 可知

$$E(\mathbf{X}^{(1)}|\mathbf{X}^{(2)}) = \boldsymbol{\mu}^{(1)} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})$$

- 这表明 $\mathbf{X}^{(1)}$ 和 $\mathbf{X}^{(2)}$ 之间有回归关系;
- 于是称矩阵 $\boldsymbol{\beta} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$ 为 $\mathbf{X}^{(1)}$ 对 $\mathbf{X}^{(2)}$ 的回归系数;
- $\boldsymbol{\Sigma}_{11.2}$ 为条件协方差矩阵, 它的元素用 $\sigma_{ij \cdot q+1, \dots, p}$ 表示, $i, j = 1, \dots, q$ 。

- **偏相关系数** (partial correlation of coefficient): 由于 $(X^{(1)}|X^{(2)} = \mathbf{x}^{(2)}) \sim N_q(\boldsymbol{\mu}_{1.2}, \boldsymbol{\Sigma}_{11.2})$, 故可定义条件相关系数

$$\rho_{ij \cdot q+1, \dots, p} = \frac{\sigma_{ij \cdot q+1, \dots, p}}{(\sigma_{ii \cdot q+1, \dots, p} \sigma_{jj \cdot q+1, \dots, p})^{1/2}}, \quad i, j = 1, \dots, q$$

为在给定 $X^{(2)} = \mathbf{x}^{(2)}$ 条件下, X_i 与 X_j 的**偏相关系数** ($\text{corr}(X^{(1)}|X^{(2)})$)。

- 显然

$$\mathbf{R}_{11.2} = [\text{diag}(\boldsymbol{\Sigma}_{11.2})]^{-1/2} \boldsymbol{\Sigma}_{11.2} [\text{diag}(\boldsymbol{\Sigma}_{11.2})]^{-1/2}.$$

定义3.12: 残差向量和回归系数

向量 $\epsilon^{(1.2)} = X^{(1)} - \mu^{(1)} - \beta(X^{(2)} - \mu^{(2)})$ 称为 $X^{(1)}$ 对 $X^{(2)}$ 的残差向量, 其中 $\beta = \Sigma_{12}\Sigma_{22}^{-1}$ 为 $X^{(1)}$ 对 $X^{(2)}$ 的回归系数, 它的第 i 行第 $(k - q)$ 列元素用 $\beta_{ik \cdot q+1, \dots, p}$ 表示, $i = 1, \dots, q, k = q + 1, \dots, p$.

定理3.3.1

残差向量 $\epsilon^{(1.2)}$ 与 $X^{(2)}$ 是不相关的。

证明: 由定理3.2.4知, $\epsilon^{(1.2)} = Y^{(1)} - E(Y^{(1)})$, 其中

$$Y^{(1)} = X^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}X^{(2)}.$$

结果可由定理3.2.4得到。

定理3.3.2

对于任意一个向量 \mathbf{a} , 则有 $\text{Var}(\varepsilon_i^{(1.2)}) \leq \text{Var}(X_i - \mathbf{a}'\mathbf{X}^{(2)})$, 其中 $\varepsilon_i^{(1.2)}$ 为 $\boldsymbol{\varepsilon}^{(1.2)}$ 的第 i 个分量。

证明: 由定理3.3.1和简单计算, 有

$$\begin{aligned}\text{Var}(X_i - \mathbf{a}'\mathbf{X}^{(2)}) &= E[X_i - \mu_i - \mathbf{a}'(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})]^2 \\ &= E[\varepsilon_i^{(1.2)} - E(\varepsilon_i^{(1.2)}) + (\boldsymbol{\beta}_{(i)} - \mathbf{a})'(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})]^2 \\ &= \text{Var}(\varepsilon_i^{(1.2)}) + (\boldsymbol{\beta}_{(i)} - \mathbf{a})'\boldsymbol{\Sigma}_{22}(\boldsymbol{\beta}_{(i)} - \mathbf{a}),\end{aligned}$$

其中 $\boldsymbol{\beta}'_{(i)}$ 是 $\boldsymbol{\beta}$ 的第 i 行。因为 $\boldsymbol{\Sigma}_{22}$ 是正定矩阵, 则上式第二项二次型 $(\boldsymbol{\beta}_{(i)} - \mathbf{a})'\boldsymbol{\Sigma}_{22}(\boldsymbol{\beta}_{(i)} - \mathbf{a}) \geq 0$ 。等号当且仅当 $\boldsymbol{\beta}_{(i)} = \mathbf{a}$ 时成立。

注: 定理3.3.2表明 $\mu_i + \boldsymbol{\beta}'_{(i)}(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})$ 是 X_i 的一个最佳线性预测。

定理3.3.3

设 $\mathbf{X} = (X_1, \dots, X_p)' \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$. 令 $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1} = (\omega_{ij})_{i,j=1,\dots,p}$ 为精度矩阵。有下面两个结论:

(1) 精度矩阵 $\boldsymbol{\Omega}$ 的对角线上的元素 ω_{ii} 是其余变量给定后 X_i 的条件方差的倒数, 即

$$\omega_{ii} = [\text{Var}(X_i | X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p)]^{-1}, \quad i = 1, \dots, p.$$

(2) 类似于相关系数矩阵的计算。令

$$\mathbf{C} = \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{p1} & \cdots & c_{pp} \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} \frac{1}{\sqrt{\omega_{11}}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\omega_{pp}}} \end{pmatrix}.$$

定理3.3.3(续)

并且令

$$\mathbf{C} = \mathbf{\Lambda} \mathbf{\Omega} \mathbf{\Lambda}.$$

也就是有

$$c_{ij} = \frac{\omega_{ij}}{\sqrt{\omega_{ii}}\sqrt{\omega_{jj}}}, \quad i, j = 1, \dots, p.$$

可见矩阵 \mathbf{C} 对角线元素为1, 而非对角线元素为 c_{ij} , 是其余变量给定后, X_i 和 X_j 的偏相关系数的相反数, 即

$$\begin{aligned} c_{ij} &= -\rho_{X_i, X_j | X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_p} \\ &= -\rho_{ij.1, \dots, i-1, i+1, \dots, j-1, j+1, \dots, p}, \quad 1 \leq i < j \leq p. \end{aligned}$$

证明：对于(1)，仅证明 $i = 1$ 的情况，即

$$\omega_{11} = [\text{Var}(X_1|X_2, \dots, X_p)]^{-1},$$

其他情况可类似证明。将协方差矩阵 Σ 和精度矩阵 Ω 进行剖分：

$$\Sigma = \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{matrix} 1 \\ p-1 \end{matrix}, \quad \Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{matrix} 1 \\ p-1 \end{matrix},$$

由定理3.2.1，有： $\text{Var}(X_1|X_2, \dots, X_p) = \sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ 。

由 $\Omega = \Sigma^{-1}$ ，根据分块矩阵的逆矩阵公式，有

$$\omega_{11} = (\sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}.$$

由上面两个公式，则证明了 $i = 1$ 的情况。

证明(续): 对于(2), 仅证明 $i = 1, j = 2$ 的情况, 即给定 X_3, \dots, X_p 后, X_1 与 X_2 的偏相关系数为:

$$\rho_{X_1, X_2 | X_3, \dots, X_p} = \rho_{12.3, \dots, p} = -c_{12},$$

其他情况可类似证明。将协方差矩阵 Σ 和精度矩阵 Ω 进行剖分:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}_{p-2}, \quad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}_{p-2},$$

根据分块矩阵的逆矩阵公式, 有

$$\Omega_{11} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}.$$

证明(续): 记

$$\Omega_{11} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}, \quad \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

则

$$\omega_{11} = \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}}, \quad \omega_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}}, \quad \omega_{12} = -\frac{a_{12}}{a_{11}a_{22} - a_{12}a_{21}},$$

在给定 X_3, \dots, X_p 后, X_1 与 X_2 的偏相关系数为

$$\rho_{12.3, \dots, p} = \frac{a_{12}}{\sqrt{a_{11}a_{22}}}.$$

由于

$$c_{12} = \frac{\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}} = -\frac{a_{12}}{\sqrt{a_{11}a_{22}}} = -\rho_{12.3, \dots, p}.$$

精度矩阵

- 由定理3.3.3(2)可知：精度矩阵 Ω 的非对角线元素 $\omega_{ij} = 0$, $\omega_{ij} > 0$ 和 $\omega_{ij} < 0$ 分别表示其余变量给定后 X_i 与 X_j 条件独立, 条件负相关和条件正相关的充要条件。
- 假设 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, 将 \mathbf{X} , $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$ 类似进行剖分：

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \mathbf{X}^{(3)} \end{pmatrix} \begin{matrix} q_1 \\ q_2 \\ q_3 \end{matrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{pmatrix} \begin{matrix} q_1 \\ q_2 \\ q_3 \end{matrix},$$
$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} & \boldsymbol{\Omega}_{13} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} & \boldsymbol{\Omega}_{23} \\ \boldsymbol{\Omega}_{31} & \boldsymbol{\Omega}_{32} & \boldsymbol{\Omega}_{33} \end{pmatrix} \begin{matrix} q_1 \\ q_2 \\ q_3 \end{matrix}.$$

定理3.3.4

在 $\mathbf{X}^{(3)}$ 给定后, $\mathbf{X}^{(1)}$ 与 $\mathbf{X}^{(2)}$ 相互条件独立的充要条件为 $\boldsymbol{\Omega}_{12} = \mathbf{0}$ 。

证明：在 $X^{(3)}$ 给定后， $X^{(1)}$ 与 $X^{(2)}$ 的条件协方差矩阵为

$$\begin{aligned}\Sigma_{12.3} &= \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} - \begin{pmatrix} \Sigma_{13} \\ \Sigma_{23} \end{pmatrix} \Sigma_{33}^{-1} (\Sigma_{31}, \Sigma_{32}) \\ &= \begin{pmatrix} \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} \\ \Sigma_{21} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{22} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{32} \end{pmatrix}.\end{aligned}$$

所以在给定 $X^{(3)}$ 后， $X^{(1)}$ 与 $X^{(2)}$ 相互独立的充要条件是

$$\Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} = \mathbf{0}.$$

由于 $\Omega = \Sigma^{-1}$ ，根据分块矩阵逆矩阵的计算公式，有

$$\begin{aligned}\begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} &= \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} - \begin{pmatrix} \Sigma_{13} \\ \Sigma_{23} \end{pmatrix} \Sigma_{33}^{-1} (\Sigma_{31}, \Sigma_{32}) \\ &= \begin{pmatrix} \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} \\ \Sigma_{21} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{22} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{32} \end{pmatrix}.\end{aligned}$$

由此可见： $\Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} = \mathbf{0}$ 的充要条件是 $\Omega_{12} = \mathbf{0}$ 。

思考题：对于给定的相关系数矩阵 $\mathbf{R} = (\rho_{ij})_{1 \leq i, j \leq p}$ ，考虑下面两种结构：

(1) 当 $i = j$ 时， $\rho_{ij} = 1$ ；当 $i \neq j$ 时， $\rho_{ij} = \rho$ ；

(2) AR(1)自回归结构，即 $\rho_{ij} = \rho^{|i-j|}$ 。

问题：观察当维数 p 变大时，精度矩阵 $\mathbf{\Omega} = \mathbf{R}^{-1}$ 中元素的变化。

```
n = 100; p = 30; rho = 0.5
```

```
corrmat=diag(rep(1-rho,p))+matrix(rho,p,p)
```

```
solve(corrmat)
```

```
ar1mat=rho^outer(1:p,1:p, function(x,y) abs(x-y))
```

```
solve(ar1mat)
```

- 设总体 $\mathbf{X} = (X_1, \dots, X_p)' \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- 假设 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 是来自总体 \mathbf{X} 的样本容量为 n 的随机样本，记为 $n \times p$ 的样本矩阵 \mathbf{X} ，即

$$\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)' = \begin{pmatrix} \mathbf{x}'_1 \\ \vdots \\ \mathbf{x}'_n \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}.$$

- 讨论样本随机矩阵 \mathbf{X} 的分布时，通常把 \mathbf{X} 的行向量拉直构成一个长度为 np 的向量，然后讨论这个 np 维向量的分布。
- 为了讨论 np 维向量的分布，我们首先需要介绍矩阵的拉直运算和克罗内克(Kronecker)积运算。

- **矩阵拉直：** 令 $n \times p$ 阶矩阵 $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_p)$ ，其中 $\mathbf{x}_i = (x_{1i}, \cdots, x_{ni})'$ ， $i = 1, \cdots, p$ 。所谓矩阵的拉直运算就是将矩阵按列拉直为向量， \mathbf{X} 按列拉直后的向量记为 $\text{Vec}(\mathbf{X})$ ，它等于

$$\text{Vec}(\mathbf{X}) = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_p \end{pmatrix} = (x_{11}, x_{21}, \cdots, x_{n1}, \cdots, x_{1p}, x_{2p}, \cdots, x_{np})'.$$

由于 \mathbf{X} 是 $n \times p$ 阶矩阵，所以 $\text{Vec}(\mathbf{X})$ 是 $np \times 1$ 维向量。另外

$$\text{Vec}(\mathbf{X}') = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} = (x_{11}, x_{12}, \cdots, x_{1p}, \cdots, x_{n1}, x_{n2}, \cdots, x_{np})'.$$

- **克罗内克(Kronecker)积**: 令 \mathbf{A} 和 \mathbf{B} 分别是 $n \times p$ 和 $m \times q$ 阶矩阵, 并记 $\mathbf{A} = (a_{ij})$, 其中 a_{ij} 为矩阵 \mathbf{A} 的元素, $i = 1, \dots, n, j = 1, \dots, p$ 。矩阵 \mathbf{A} 和 \mathbf{B} 的Kronecker乘积, 记为 $\mathbf{A} \otimes \mathbf{B}$, 它等于

$$\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B}) = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1p}\mathbf{B} \\ \vdots & & \vdots \\ a_{n1}\mathbf{B} & \cdots & a_{np}\mathbf{B} \end{pmatrix},$$

可知, $\mathbf{A} \otimes \mathbf{B}$ 是 $nm \times pq$ 阶矩阵。

- 利用矩阵拉直运算和Kronecker积的定义和性质, 可知

$$\text{Vec}(\mathbf{X}') \sim N_{np}(\mathbf{1}_n \otimes \boldsymbol{\mu}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}),$$

其中 $\mathbf{1}_n$ 表示元素均为1的 n 维向量, \mathbf{I}_n 表示 $n \times n$ 的单位阵。

- np 维向量 $\text{Vec}(\mathbf{X}')$ 的联合密度函数可表示为

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \prod_{i=1}^n \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right] \\ &= \frac{1}{(2\pi)^{np/2} |\boldsymbol{\Sigma}|^{n/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right] \\ &= \frac{1}{(2\pi)^{np/2} |\boldsymbol{\Sigma}|^{n/2}} \exp \left[-\frac{1}{2} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu} \\ \vdots \\ \mathbf{x}_n - \boldsymbol{\mu} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Sigma} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \boldsymbol{\Sigma} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu} \\ \vdots \\ \mathbf{x}_n - \boldsymbol{\mu} \end{pmatrix} \right]. \end{aligned}$$

定义3.13: 矩阵正态分布

设 \mathbf{X} 为 $n \times p$ 随机矩阵, 若 \mathbf{X} 按行拉直后服从下面的分布

$$\text{Vec}(\mathbf{X}') \sim N_{np}(\mathbf{1}_n \otimes \boldsymbol{\mu}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}),$$

则称 \mathbf{X} 服从**矩阵正态分布**, 记作

$$\mathbf{X} \sim N_{n \times p}(\mathbf{M}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}),$$

其中

$$\mathbf{M} = \begin{pmatrix} \mu_1 & \cdots & \mu_p \\ \vdots & \ddots & \vdots \\ \mu_1 & \cdots & \mu_p \end{pmatrix}.$$

性质3.4.1

设 $\mathbf{X} \sim N_{n \times p}(\mathbf{M}, \mathbf{I}_n \otimes \Sigma)$, \mathbf{A} 为 $k \times n$ 常数矩阵, \mathbf{B} 为 $p \times q$ 常数矩阵, \mathbf{D} 为 $k \times q$ 常数矩阵, 令 $\mathbf{Z} = \mathbf{AXB} + \mathbf{D}$, 则

$$\mathbf{Z} \sim N_{k \times q}(\mathbf{AMB} + \mathbf{D}, (\mathbf{AA}') \otimes (\mathbf{B}'\Sigma\mathbf{B})).$$

证明留作作业, 课下完成。

多元正态分布样本统计量

- 设 $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ 为 p 元正态总体 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 的样本容量为 n 的简单随机样本矩阵, 其中 $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma} > 0$, $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ 和 $n > p$ 。

样本均值向量

设 $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})' (i = 1, \dots, n)$ 为一组随机样本, 则称

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i = \frac{1}{n} \mathbf{X}' \mathbf{1}_n = (\bar{x}_1, \dots, \bar{x}_p)'$$

为 **样本均值向量**, 其中 $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik}, k = 1, \dots, p$ 。

样本离差阵

设 $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ ($i = 1, \dots, n$) 为一组随机样本，则称

$$\begin{aligned}\mathbf{V} &= \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' = \mathbf{X}'\mathbf{X} - n\bar{\mathbf{x}}\bar{\mathbf{x}}' \\ &= \mathbf{X}'\left[\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right]\mathbf{X}\end{aligned}$$

为样本离差阵，其中 \mathbf{X} 为 $n \times p$ 的样本矩阵。

$$\begin{aligned}\sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})' &= \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' + n(\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})' \\ &= \mathbf{V} + n(\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})'\end{aligned}$$

样本协方差阵

设 $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})' (i = 1, \dots, n)$ 为一组随机样本，则称

$$\mathbf{S} = \frac{1}{n-1} \mathbf{V} = (s_{ij})_{p \times p}$$

为**样本协方差阵**。其中 $s_{kk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 (k = 1, \dots, p)$ 称为第 k 个样本变量 \mathbf{x}_k 的**样本方差**； $\sqrt{s_{kk}}$ 称为变量 \mathbf{x}_k 的**样本标准差**。

作用：

- 估计多元正态总体分布的协方差阵
- 对有关总体分布均值向量和协方差阵的假设进行检验

样本相关阵

设 $\mathbf{S} = (s_{ij})_{p \times p}$ 为样本协方差阵, 则称 $\tilde{\mathbf{R}} = (r_{ij})_{p \times p}$ 为样本相关阵, 其中

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}, \quad i, j = 1, \dots, p.$$

若记 $\tilde{\mathbf{D}}^{1/2} = \text{diag}(\sqrt{s_{11}}, \dots, \sqrt{s_{pp}})$ 为样本标准差对角阵, 则

$$\tilde{\mathbf{R}} = \tilde{\mathbf{D}}^{-1/2} \mathbf{S} \tilde{\mathbf{D}}^{-1/2}, \quad \mathbf{S} = \tilde{\mathbf{D}}^{1/2} \tilde{\mathbf{R}} \tilde{\mathbf{D}}^{1/2}.$$

定义：多元特征函数

随机向量 $\mathbf{X} = (X_1, \dots, X_p)'$ 的特征函数称为 **多元特征函数**，它等于

$$\varphi_{\mathbf{X}}(\mathbf{t}) = E[\exp(\mathbf{i}\mathbf{t}'\mathbf{X})],$$

其中 $\mathbf{t} = (t_1, \dots, t_p)'$, $\mathbf{i} = \sqrt{-1}$ 。

特征函数*

下面介绍多元特征函数的一些性质：

性质1'：矩

若随机向量 $\mathbf{X} = (X_1, \dots, X_p)'$ 的特征函数为 $\varphi_{\mathbf{X}}(\mathbf{t})$ ，并且 $E(X_1^{k_1} \dots X_p^{k_p})$ 存在，则

$$E(X_1^{k_1} \dots X_p^{k_p}) = \mathbf{i}^{k_1 + \dots + k_p} \left[\frac{\partial^{k_1 + \dots + k_p} \varphi_{\mathbf{X}}(\mathbf{t})}{\partial t_1^{k_1} \dots \partial t_p^{k_p}} \right]_{t_1 = \dots = t_p = 0}.$$

特别地，

(1) 若 X_j 的均值 $E(X_j)$ 存在，则

$$E(X_j) = \mathbf{i} \left[\frac{\partial \varphi_{\mathbf{X}}(t_1, \dots, t_p)}{\partial t_j} \right]_{t_1 = \dots = t_p = 0};$$

性质1'(续): 矩

(2) 若 X_j 的二阶矩 $E(X_j^2)$ 存在, 则

$$E(X_j^2) = - \left[\frac{\partial^2 \varphi_X(t_1, \dots, t_p)}{\partial t_j^2} \right]_{t_1=\dots=t_p=0};$$

(3) 若 X_j 和 $X_k (j \neq k)$ 的混合矩 $E(X_j X_k)$ 存在, 则

$$E(X_j X_k) = - \left[\frac{\partial^2 \varphi_X(t_1, \dots, t_p)}{\partial t_j \partial t_k} \right]_{t_1=\dots=t_p=0}.$$

特征函数*

性质2': 边际分布

若随机向量 $\mathbf{X} = (X_1, \dots, X_p)'$ 的特征函数为 $\varphi_{\mathbf{X}}(t_1, \dots, t_p)$, 则 $\mathbf{X}^{(m)} = (X_1, \dots, X_m)'$ 的特征函数为 $\varphi_{\mathbf{X}^{(m)}}(t_1, \dots, t_m, 0, \dots, 0)$, 其中 $m < p$ 。

性质3': 唯一性

特征函数与分布函数相互唯一确定。

性质4': 独立性

如果 $\mathbf{X} = (X_1, \dots, X_p)'$ 的分量是相互独立的, 则

$$\varphi_{\mathbf{X}}(t_1, \dots, t_p) = \prod_{i=1}^p E(\exp(it_i X_i)).$$

特征函数*

性质5': 和

假设 $\mathbf{X}_k \in \mathbb{R}^p$, \mathbf{X}_k 的特征函数为 $\varphi_{\mathbf{X}_k}(\mathbf{t})$, $k = 1, \dots, m$ 。如果 $\mathbf{X}_1, \dots, \mathbf{X}_m$ 相互独立, 则 $\mathbf{X}_1 + \dots + \mathbf{X}_m$ 的特征函数为 $\varphi(\mathbf{t})$ 等于 $\mathbf{X}_1, \dots, \mathbf{X}_m$ 的特征函数的乘积, 即

$$\varphi(\mathbf{t}) = \varphi_{\mathbf{X}_1}(\mathbf{t}) \cdots \varphi_{\mathbf{X}_m}(\mathbf{t}).$$

性质6': 密度函数

如果 $\mathbf{X} = (X_1, \dots, X_p)'$ 有密度函数 $f(\mathbf{x})$ 和特征函数 $\varphi_{\mathbf{X}}(\mathbf{t})$, 则

$$f(\mathbf{x}) = \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(-i\mathbf{t}'\mathbf{x}) \varphi_{\mathbf{X}}(\mathbf{t}) dt_1 \cdots dt_p,$$

其中 $\mathbf{x} = (x_1, \dots, x_p)'$ 和 $\mathbf{t} = (t_1, \dots, t_p)'$ 。



谢谢，请多提宝贵意见！