第二章、多元正态分布及参数的估计

September 25, 2018

2.0 引入

1. 多元统计的主要研究内容

分类（判别分析），建模（多元回归，变量选择），降维（公因子，压缩惩罚），典型相关分析（组间）

2. 多元数据的直觉

两两之间的关系（ACF, 散点图），轮廓图，雷达图

3. 多元分析的理论基础

数据的估计与推断离不开理论的支撑，这就要求数据分布有一定的假设

一、随机向量的联合分布,边缘分布,条件分布

1. 多元数据

a. 随机向量：若X_i 都为随机变量，则

称$\boldsymbol{X} = (X_1, \cdots, X_p)'$为$p$维随机向量。

b. 随机矩阵：$\boldsymbol{X}_{n\times p} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)'$, 其中$\boldsymbol{X}_i$ 为p维随机向量。

2. 多元分布

(i) 分布函数：$F(\boldsymbol{X} \le \boldsymbol{x}) = P(X_1 \le x_1, \ldots, X_p \le x_p)$, $\boldsymbol{x} = (x_1, \ldots, x_p)' \in \mathcal{R}^p$为$\boldsymbol{X}$。

(ii) 密度函数：设$F(\boldsymbol{x})$ 为$\boldsymbol{X}$的分布函数，若存在非负函数$f(\boldsymbol{x})$，使得对任意$\boldsymbol{x} \in \mathcal{R}^p$, 都有

$$F(\boldsymbol{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_p} f(x_1, x_2, \ldots, x_p)\, dx_1 dx_2 \ldots dx_p,$$

则称$f(\boldsymbol{x})$ 为$\boldsymbol{X}$的密度函数。

3. 边际分布

设$F(\boldsymbol{x})$为$\boldsymbol{X}$的分布函数, $\boldsymbol{X}_1 = (X_{i_1}, \ldots, X_{i_q})'$, $q < p$为$\boldsymbol{X}$的子向量。则对任意$\boldsymbol{x}_1 = (x_{i_1}, \ldots, x_{i_q})'$,

$$P(\boldsymbol{X}_1 \le \boldsymbol{x}_1) = F(\infty, \ldots, x_{i_1}, \infty, \ldots, x_{i_q}, \ldots, \infty).$$

4. 条件分布

设$\boldsymbol{X} = (\boldsymbol{X}_1', \boldsymbol{X}_2')'$, 则在给定$\boldsymbol{X}_2$条件下，$\boldsymbol{X}_1$的条件分布为

$$f(\boldsymbol{x}_1|\boldsymbol{x}_2) = f(\boldsymbol{x}_1, \boldsymbol{x}_2)/f_{X_2}(\boldsymbol{x}_2).$$

5. 相互独立

设$(\boldsymbol{X}_1', \boldsymbol{X}_2')'$的联合密度函数，与边际密度函数满足：

$$f(\boldsymbol{x}_1, \boldsymbol{x}_2) = f_{x_1}(\boldsymbol{x}_1)f_{x_2}(\boldsymbol{x}_2).$$

进一步，若

$$f(\boldsymbol{x}) = \prod_{i=1}^{p} f_i(x_i),$$

则称$X_1, \dots, X_p$相互独立。

二、随机向量的数字特征

设$X=(X_1,\cdots,X_p)'$,$Y=(Y_1,\cdots,Y_q)'$ 是两个随机向量.

1、随机向量X的均值向量

$$E(X)=[E(X_1),\cdots,E(X_p)]'=(\mu_1,\cdots,\mu_p)'$$

2、随机向量X的协方差阵

$$D(X)=E[(X-E(X))(X-E(X))']=(\sigma_{ij})_{(p\times p)}\overset{def}{=}\Sigma$$

3、随机向量X和Y的协方差阵

$$\mathrm{Cov}(X,Y)=E[(X-E(X))(Y-E(Y))']=(\sigma_{ij})_{(p\times q)}$$

如果$\mathrm{Cov}(X,Y)=0$,则称X和Y是不相关的.

$\boldsymbol{\rho} = (\rho_{ij})_{(p \times p)}$ 为X 的**相关（系数矩）阵**,其中

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

若记$V^{1/2} = diag(\sqrt{\sigma_{11}}, \cdots, \sqrt{\sigma_{pp}})$ 为**标准差矩阵**, 协方差矩阵和相关阵的关系如下,

$$\boldsymbol{\rho} = V^{-1/2}\Sigma V^{-1/2}$$

或

$$\Sigma = V^{1/2}\boldsymbol{\rho} V^{1/2}$$

其中：$V^{-1/2} = (V^{1/2})^{-1} = diag(1/\sqrt{\sigma_{11}}, \cdots, 1/\sqrt{\sigma_{pp}})$.

三、均值向量和协方差阵性质

性质1 设X和Y是随机向量,A, B 是常数矩阵,则

$$
\begin{aligned}
\mathsf{E}(AX) &= A\mathsf{E}(X) \\
\mathsf{E}(AXB) &= AE(X)B \\
\mathsf{Var}(AX) &= A\mathsf{Var}(X)A' \\
\mathrm{Cov}(AX, BY) &= A\mathrm{Cov}(X, Y)B'.
\end{aligned}
$$

性质2 若X和Y相互独立,则$\mathrm{Cov}(X, Y) = 0$,反之未必成立.

性质3 随机向量$X = (X_1, \cdots, X_p)$的协方差矩阵$\mathsf{Var}(X) = \Sigma$是对称非负定矩阵.

性质4 $\Sigma = L^2$, 其中L为非负定矩阵. 特别当$\Sigma > 0$ 时, $L > 0$. 矩阵L也称为矩阵Σ的**平方根矩阵**.

例1：已知$\boldsymbol{X} = (X_1, X_2)'$, $\mathrm{E}\boldsymbol{X} = (\mu_1, \mu_2)'$,
$\mathrm{Cov}\boldsymbol{X} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$. 求$Y_1 = 2X_1 - X_2$, $Y_2 = X_1 + X_2$ 的均值向量和协方差阵？

2.2 多元正态分布的定义与基本性质

注意：多元正态分布的四种等价定义.

定义2.2.1 设$U=(U_1,U_2,\cdots,U_q)'$为随机向量, $U_1,\cdots,U_q$ 相互独立且同$N(0,1)$分布, 设μ 为p 维常数向量, A 为$p\times q$常数矩阵,则称$X=AU+\mu$的分布为p**元正态分布**,或称X为p**维正态随机向量**,记为$X\sim N_p(\mu,AA')$.

性质1　设$U=(U_1,U_2,\cdots,U_q)'$为随机向量,$U_1,\cdots,U_q$相互独立且同$N(0,1)$分布,则$X=AU+\mu$的特征函数为

$$\Phi_X(t)=\exp\left[it'\mu-\frac{1}{2}t'AA't\right]$$

其中$t=(t_1,\cdots,t_p)'\in\mathbb{R}^p$.

由概率论知识,

- 标准正态分布的特征函数,

$$\Phi_{U_i}(t_i)=\exp\left[-\frac{1}{2}t_i^2\right].$$

- 随机向量X的特征函数,

$$\Phi_X(t)=\exp\left[it'X\right].$$

其中$t=(t_1,\cdots,t_q)'\in\mathbb{R}^p$.

定义2.2.2 若p维随机向量X的特征函数为

$$\Phi_X(t) = \exp\left[it'\mu - \frac{1}{2}t'\Sigma t\right]$$

则称X服从p元正态分布,记为$X \sim N_p(\mu, \Sigma)$

性质2　设$X \sim N_p(\mu, \Sigma)$,B为$s \times p$常数矩阵,d为s维常数向量,令$Z = BX + d$,则$Z \sim N_s(B\mu + d, B\Sigma B')$.

性质2指出正态随机向量的任意线性变换仍服从正态分布.

推论　设$X = \begin{array}{cc} \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} & \begin{array}{l} r \\ p-r \end{array} \end{array} \sim N_p(\mu, \Sigma)$,将$\mu, \Sigma$剖分为

$$\mu = \begin{array}{cc} \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix} & \begin{array}{l} r \\ p-r \end{array} \end{array}, \Sigma = \begin{array}{cc} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} & \begin{array}{l} r \\ p-r \end{array} \end{array}$$

则$X^{(1)} \sim N_r(\mu^{(1)}, \Sigma_{11}), X^{(2)} \sim N_{p-r}(\mu^{(2)}, \Sigma_{22})$.

性质3 若$X \sim N_p(\mu, \Sigma)$,则$E(X) = \mu$, $\mathsf{Var}(X) = \Sigma$

性质4 设$X = (X_1, X_2, \cdots, X_p)'$为$p$维随机向量,则$X$ 服从p元正态分布$\Longleftrightarrow$对任一p维实向量a,(a为非零向量), $\xi = a'X$ 是一维正态随机变量.

(利用特征函数来证明)

定义2.2.3 若p维随机向量X的任意线性组合 $a'X$ (a为非零向量)均服从一元正态分布, 则称X 为p **维正态随机向量**.

性质5 若$X \sim N_p(\mu, \Sigma)$,且$\Sigma > 0$(正定),则X的联合密度函数是

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

- 多元标准正态随机变量$U = (U_1, U_2, \cdots, U_q)'$的密度函数，记: $u = (u_1, \cdots, u_q)$,
  $$f(u) = (2\pi)^{-\frac{q}{2}}\Pi_{i=1}^{q} \exp\{-\frac{1}{2}u_i^2\} = (2\pi)^{-\frac{q}{2}} \exp\{-\frac{1}{2}u'u\}.$$
- 多元正态随机变量$X = AU + \mu$, 采用积分变换，记Jacobian行列式为$J(u \to x)$, 由U的密度函数，得出X的密度函数.
  $$f(x) = f\left(u(A^{-1}(x-\mu))\right) J(u \to x),$$
  其中
  $$J(u \to x) = |AA'|^{1/2} = |\Sigma|^{1/2}.$$

定义2.2.4 若p维随机向量X的联合密度函数为

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

其中μ为p维常数向量，Σ 是p阶正定矩阵；也称X为p**维正态随机向量**.

注意：**定义2.2.4** 要求$\Sigma > 0$，但前面三种定义只要求$\Sigma \geq 0$.

例2、(二元正态分布密度函数): 设$X=(X_1,X_2)\sim N_2(\mu,\Sigma)$,则

$$\mu=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix},\ \Sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix},$$

$$\Sigma^{-1}=\frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_1^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}.$$

代入

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

得二元正态分布密度函数。

二元正态密度函数图: 绘制二元正态密度函数的图形及其相应的等高线图形. R中的graphics包中的perps函数用于作3D图形

(一)赋值

```
mu1<-0; mu2<-0;s11<-10;s12<-15;s22<-10;rho<-0.5;
x1<-seq(-10,10,length=41) # generating the vector series x1
x2<-x1 # copying x1 to x2
```

(二)写出密度函数

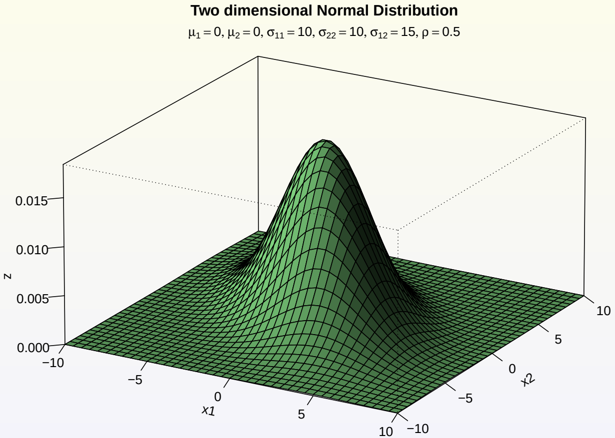
```
f<-function(x1,x2)
{
term1<-1/(2*pi*sqrt(s11*s22*(1-rho^2)))
term2<--1/(2*(1-rho^2))
term3<-(x1-mu1)^2/s11
term4<-(x2-mu2)^2/s22
term5<--2*rho*((x1-mu1)*(x2-mu2))/(sqrt(s11)*sqrt(s22))
2
term1*exp(term2*(term3+term4-term5))
}
```

(三)计算密度函数值

```
z<-outer(x1,x2,f) # calculating the density values
```

(四)作图

```
persp(x1, x2, z)
```



$$f(\mathbf{x}) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho^2)}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_{11}} - 2\rho\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}} + \frac{(x_2-\mu_2)^2}{\sigma_{22}}\right]\right\}$$

等高面： 称$\{\boldsymbol{x}: f(\boldsymbol{x})=c\}$ 为$\boldsymbol{X}$的等高面，等价于

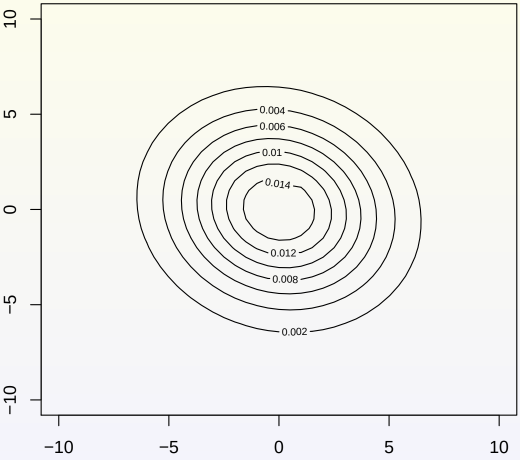
$$\{\boldsymbol{x}:(\boldsymbol{x}-\mu)'\Sigma^{-1}(\boldsymbol{x}-\mu)=d\}.$$

特别地，当$\boldsymbol{X}$为二维正态随机变量时，$f(\boldsymbol{x})=c$等价于

$$\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2=d(1-\rho^2),$$

其中，$\rho=\mathrm{Cov}(X_1,X_2)$.

R中的graphics包中的contour函数。 例如，$contour(x1,x2,z)$可得下面图形。


0.004
0.006
0.01
0.014
0.012
0.008
0.002
10
5
0
−5
−10
−10
−5
0
5
10

2.3 条件分布和独立性

注意：本节讨论的$\Sigma > 0$的情形.

一、独立性

定理2.3.1 若p维随机向量$X \sim N_p(\mu, \Sigma)$

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \sim N_p\left(\begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

则

$$X^{(1)}\text{与}X^{(2)}\text{独立} \Longleftrightarrow \Sigma_{12} = \mathbf{0} \ \text{即}X^{(1)}\text{与}X^{(2)}\text{互不相关.}$$

证明思想：

$\Rightarrow$ 独立推出不相关，即$\Sigma_{12} = \mathbf{0}$;

$\Leftarrow$ $\Sigma_{12} = \mathbf{0}$，由此推出，联合密度等边际密度的乘积.

$$
\begin{aligned}
& f(x^{(1)}, x^{(2)}) \\
= \; & \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)' \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}^{-1} (x-\mu)\right] \\
= \; & f_1(x^{(1)}) f_2(x^{(2)}).
\end{aligned}
$$

推论1 设$r_i \geq 0(i = 1, \cdots, k)$,且$r_1 + r_2 + \cdots + r_k = p$,有

$$X = \begin{array}{c} \left[\begin{array}{c} X^{(1)} \\ \vdots \\ X^{(k)} \end{array}\right] \end{array} \begin{array}{c} r_1 \\ \vdots \\ r_k \end{array} \sim N_p\left(\left[\begin{array}{c} \mu^{(1)} \\ \vdots \\ \mu^{(k)} \end{array}\right], \left[\begin{array}{ccc} \Sigma_{11} & \cdots & \Sigma_{1k} \\ \vdots & & \vdots \\ \Sigma_{k1} & \cdots & \Sigma_{kk} \end{array}\right]_{p\times p}\right)$$

则$X^{(1)}, \cdots, X^{(k)}$相互独立$\Longleftrightarrow \Sigma_{ij} = 0$(一切$i \neq j$).

推论2 设$X=(X_1,\cdots,X_p)'\sim N_p(\mu,\Sigma)$,若$\Sigma$为对角阵,则$X_1,\cdots,X_p$相互独立.

$$X=\begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}\sim N_p\left(\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix},\begin{bmatrix} \sigma_{11} & \cdots & 0 \\ \vdots & \sigma_{ii} & \vdots \\ 0 & \cdots & \sigma_{pp} \end{bmatrix}\right)$$

例3、设三维随机向量$\boldsymbol{X} \sim N_3(\mu, 2I_3)$, $\mu = (2,0,0)'$,

$$A = \begin{pmatrix} 0.5 & -1 & 0.5 \\ -0.5 & 0 & -0.5 \end{pmatrix}, d = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

试求$Y = A\boldsymbol{X} + d$ 的分布?

例4、设$(X_1, X_2, X_3)' \sim N_3(\mu, \Sigma)$, $\mu = (2, -3, 1)'$,

$$\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

(1) 试求$3X_1 - X_2 + 2X_3$ 的分布;

(2) 求二维向量$a = (a_1, a_2)'$, 使得X_3 与$X_3 - a'(X_1, X_2)'$ 相互独立。

二、条件分布

首先考虑二元正态随机向量(X_1, X_2)的条件分布,有条件密度的定义知, 给定X_2时, X_1的条件密度为

$$f_1(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

所以可以推得

$$(X_1|X_2 = x_2) \sim N_1\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1-\rho^2)\right)$$

协方差矩阵

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

根据分块矩阵求逆公式:

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11.2}^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11.2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

其中:

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

定理2.3.2 设$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \begin{matrix} r \\ p-r \end{matrix} \sim N_p(\mu, \Sigma)(\Sigma > 0)$,则当$X^{(2)} = x^{(2)}$给定时,$X^{(1)}$的条件分布为

$$(X^{(1)}|X^{(2)} = x^{(2)}) \sim N_r(\mu_{1\cdot 2}, \Sigma_{11\cdot 2})$$

其中

$$\begin{aligned} \mu_{1\cdot 2} &= \mu^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(x^{(2)} - \mu^{(2)}), \\ \Sigma_{11\cdot 2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \end{aligned}$$

证明思路：

$$Z=\begin{bmatrix} Z^{(1)} \\ Z^{(2)} \end{bmatrix} = \begin{bmatrix} X^{(1)}-\Sigma_{12}\Sigma_{22}^{-1}X^{(2)} \\ X^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} I_r & -\Sigma_{12}\Sigma_{22}^{-1} \\ O & I_{p-r} \end{bmatrix} \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \widehat{=} BX.$$

由前面的性质,知Z的分布为仍为正态分布

$$Z=N_p\left(\begin{bmatrix} \mu^{(1)}-\Sigma_{12}\Sigma_{22}^{-1}\mu^{(2)} \\ \mu^{(2)} \end{bmatrix}, \begin{bmatrix} \Sigma_{11.2} & O \\ O & \Sigma_{22} \end{bmatrix}\right).$$

推论 在定理2.3.2条件下可得:

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

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

(3)$X^{(2)}|X^{(1)}\sim N_{p-r}(\mu_{2\cdot1},\Sigma_{22\cdot1})$,其中

$$\mu_{2\cdot1}=\mu^{(2)}+\Sigma_{21}\Sigma_{11}^{-1}(x^{(1)}-\mu^{(1)}),$$

$$\Sigma_{22\cdot1}=\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

拓展问题:

$$假设X = \begin{array}{c} \begin{bmatrix} X^{(1)} \\ X^{(2)} \\ X^{(3)} \end{bmatrix} \end{array} \begin{array}{c} r_1 \\ r_2 \\ r_3 \end{array} \sim N_p(\mu, \Sigma)(\Sigma > 0), 其中$$

$$\mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \\ \mu^{(3)} \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}$$

- $X^{(1)}|(X^{(2)}, X^{(3)})$服从什么分布，均值向量和协方差矩阵的形式是什么？
- $X^{(2)}|(X^{(1)}, X^{(3)})$服从什么分布，均值向量和协方差矩阵的形式是什么？

三、条件数字特征

1.条件期望,回归系数,偏相关系数

设

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \sim N_p\left(\begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

又已知$X^{(2)} = x^{(2)}$给定时,$X^{(1)}$的条件分布为

$$(X^{(1)}|X^{(2)} = x^{(2)}) \sim N_r(\mu_{1\cdot 2}, \Sigma_{11\cdot 2})$$

称

$$\mu_{1\cdot 2} = \mu^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(x^{(2)} - \mu^{(2)}),$$

为**条件期望**,记为$\mathrm{E}(X^{(1)}|X^{(2)} = x^{(2)})$;并称$\mu_{1\cdot 2}$ 为$X^{(1)}$ 对$X^{(2)}$的**回归**,称$\Sigma_{12}\Sigma_{22}^{-1} \overset{def}{=} B$为**回归系数**.

记

$$\Sigma_{11\cdot 2} = (\sigma_{ij\cdot r+1,\cdots,p})_{r\times r}(i,j=1,\cdots,r).$$

称

$$\rho_{ij\cdot r+1,\cdots,p} = \frac{\sigma_{ij\cdot r+1,\cdots,p}}{\sqrt{\sigma_{ii\cdot r+1,\cdots,p}}\sqrt{\sigma_{jj\cdot r+1,\cdots,p}}}$$

为偏相关系数，即当$X^{(2)}=(X_{r+1},\cdots,X_p)'$ 给定时，X_i 与X_j $(i,j=1,2,\cdots,r)$ 的**偏相关系数**.

2.全相关系数

设$Z=\begin{bmatrix} X \\ Y \end{bmatrix}\begin{matrix} p \\ 1 \end{matrix} \sim N_{p+1}\left(\begin{bmatrix} \mu_X \\ \mu_y \end{bmatrix},\begin{bmatrix} \Sigma_{XX} & \Sigma_{Xy} \\ \Sigma_{yX} & \sigma_{yy} \end{bmatrix}\right)$,则称

$$R=\left(\frac{\Sigma_{yX}\Sigma_{XX}^{-1}\Sigma_{Xy}}{\sigma_{yy}}\right)^{1/2}$$

为Y与$X=(X_1,\cdots,X_p)'$的**全相关系数**.

3.最佳预测

在定理2.3.2条件下,考虑$r=1$,
记$X^{(1)}=Y,g(x^{(2)})=E(Y|X^{(2)}=x^{(2)})$,则对任意的函数$\phi(\cdot)$,可以证明:

$$E[(Y-g(x^{(2)}))^2]\le E[(Y-\phi(x^{(2)}))^2]$$

则在均方差最小的准则下,条件期望$g(x^{(2)})$是对Y的最佳预测函数.

2.4 随机阵的正态分布

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} X'_{(1)} \\ X'_{(2)} \\ \vdots \\ X'_{(n)} \end{bmatrix} = (X_1, X_2, \cdots, X_p)$$

- $X'_{(i)}$ 是来自p元总体的第i个观测值, $i = 1, \cdots n$.
- X_j 是随机向量第j 个分量的一个样本, $j = 1, \cdots, p$.
- $\mathbf{X}$ 是一个样本观测矩阵, 是一个随机矩阵.

一、拉直运算和克罗内克(Kronecker)积

拉直运算是指将矩阵拉成一个长向量, 通过它来建立矩阵和各量之间的关系. 记

$$\mathrm{Vec}\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = (x_{11}, \cdots x_{n1}, \cdots, x_{1p}, \cdots, x_{np})'.$$

相应的,

$$\mathrm{Vec}\mathbf{X}' = \begin{bmatrix} X_{(1)} \\ X_{(2)} \\ \vdots \\ X_{(n)} \end{bmatrix} = (x_{11}, x_{12}, \cdots, x_{n1}, \cdots, x_{np})'.$$

代数中的**克罗内克积**

矩阵$A=(a_{ij})$ 和B 分别为$n\times p$ 和$m\times q$ 的矩阵，A 和B 的克罗内克积$A\bigotimes B$ 定义为

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

$A\otimes B$ 为$mn\times pq$ 矩阵，在多元统计中被称为矩阵的直积.

二、随机阵的正态分布

设$X_{(i)}$ 为来自p 元正态总体$N_p(\mu,\Sigma)$的简单随机样本(独立同分布), $i=1,\cdots,n$, 随机阵$\mathbf{X}$ 利用按行拉直运算可知,

$$\mathrm{Vec}(\mathbf{X}')\sim N_{np}(\mathbf{I}_n\otimes\mu,\mathbf{I}_n\otimes\Sigma).$$

$$\mathbf{I}_n\otimes\mu=\begin{bmatrix}\mu\\ \vdots\\ \mu\end{bmatrix},\ \mathbf{I}_n\otimes\Sigma=\begin{bmatrix}\Sigma & \cdots & O\\ \vdots & \vdots & \vdots\\ O & \cdots & \Sigma\end{bmatrix}$$

如果随机矩阵$\mathbf{X}$ 按行拉直后, $\mathrm{Vec}(\mathbf{X}') \sim N_{np}(\mathbf{I}_n \otimes \mu, \mathbf{I}_n \otimes \Sigma)$, 则称$\mathbf{X}$服从矩阵正态分布,记作

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

其中,

$$M = \begin{bmatrix} \mu_1 & \cdots & \mu_p \\ \vdots & \vdots & \vdots \\ \mu_1 & \cdots & \mu_p \end{bmatrix} = \mathbf{1}_n \mu'.$$

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

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

2.5 多元正态分布的参数估计

考虑p元正态总体$\mathbf{X} \sim N_p(\mu, \Sigma)$, 其观测数据阵为

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

(1)$x_{(i)} = (x_{i1}, \cdots, x_{ip})'$ 为随机向量$\mathbf{X}$ 的第j个观察值向量, $i = 1, \cdots, n$. 一般假设为简单随机样本.

(2) $x_j = (x_{1j}, \cdots, x_{nj})'$ 为随机向量$\mathbf{X}$ 的第j个分量X_j的n个观察值, $j = 1, \cdots, p$.

在符号不会混淆的情况下，有时$x_{(j)}$ 记为$X_{(j)}$, x_i记为X_i.

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} = (X_{(1)}, \ldots, X_{(n)})' = \begin{pmatrix} X'_{(1)} \\ \vdots \\ X'_{(n)} \end{pmatrix}.$$

or,

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} = (X_1, \ldots, X_p).$$

一、多元正态总体样本的基本统计量

(1)样本均值向量$\overline{X}$:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_{(i)} = (\overline{x}_1, \cdots, \overline{x}_p)' = \frac{1}{n} X' \mathbf{1}_n,$$

其中

$$\overline{x}_j = \frac{1}{n}\sum_{\alpha=1}^{n} x_{\alpha j} \quad (j = 1, 2, \cdots, p).$$

(2)样本离差阵(又称交叉乘积阵)A:

$$\begin{aligned} A &= \sum_{\alpha=1}^{n}(X_{(\alpha)}-\overline{X})(X_{(\alpha)}-\overline{X})' = X'X-n\overline{X}\,\overline{X}' \\ &= X'[I_n-\frac{1}{n}\mathbf{1_n}\mathbf{1_n'}]X \overset{def}{=} (a_{ij})_{p\times p} \end{aligned}$$

其中

$$a_{ij}=\sum_{\alpha=1}^{n}(x_{\alpha i}-\overline{x}_i)(x_{\alpha j}-\overline{x}_j)(i,j=1,2,\cdots,p)$$

(3)样本协方差阵S: 注意

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

其中

$$s_{ii} = \frac{1}{n-1}\sum_{\alpha=1}^{n}(x_{\alpha i} - \overline{x}_i)^2 (i = 1, 2, \cdots, p)$$

称为变量X_i的**样本方差**;

样本方差的平方根$\sqrt{s_{ii}}$ 称为变量X_i 的**样本标准差**.

(4)样本相关阵$\widehat{R}$:

$$\widehat{R} = (r_{ij})_{p\times p},$$

其中

$$r_{ij} = \widehat{\rho}_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}} \overset{\text{or}}{=} \frac{a_{ij}}{\sqrt{a_{ii}}\sqrt{a_{jj}}}$$

样本相关系数r_{ij}为总体相关系ρ_{ij}的估计.

$$\rho_{ij} = \frac{Cov(X_i, X_i)}{\sqrt{\mathsf{Var}(X_i)\mathsf{Var}(X_j)}}.$$

(5) 正态参数μ, Σ的矩估计

- 样本均值向量$\overline{X}$可作为总体均值μ的估计,

$$\widehat{\mu} = \overline{X}.$$

- 样本协方差阵S 可以作为总体协方差阵Σ的估计.

$$\widehat{\Sigma} = A/n.$$

- 样本相关阵$\widehat{R}$ 可以作为总体相关阵R的估计.

二、μ, Σ的最大似然估计

1.似然函数

$$
\begin{aligned}
& L(\mu, \Sigma) \\
= & \prod_{i=1}^{n} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(X_{(i)}-\mu)'\Sigma^{-1}(X_{(i)}-\mu)\right] \\
= & \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{\mathrm{tr}\left[-\frac{1}{2}\Sigma^{-1}\sum_{i=1}^{n}(X_{(i)}-\mu)(X_{(i)}-\mu)'\right]\right\} \\
= & \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{\mathrm{tr}\left[-\frac{1}{2}\Sigma^{-1}(A+n(\overline{X}-\mu)(\overline{X}-\mu)')\right]\right\}.
\end{aligned}
$$

2.讨论$\ln L(\mu,\Sigma)$的最大值点

$$\begin{aligned}
\ln L(\mu,\Sigma) &= -\frac{np}{2}\ln 2\pi-\frac{n}{2}\ln|\Sigma| \\
&\quad -\frac{1}{2}tr[\Sigma^{-1}\sum_{i=1}^{n}(X_{(i)}-\mu)(X_{(i)}-\mu)'] \\
&= C-\frac{1}{2}tr[\Sigma^{-1}A+n\Sigma^{-1}(\overline{X}-\mu)(\overline{X}-\mu)'] \\
&= C-\frac{1}{2}tr(\Sigma^{-1}A)-\frac{n}{2}[(\overline{X}-\mu)\Sigma^{-1}(\overline{X}-\mu)'] \\
&\leq C-\frac{1}{2}tr(\Sigma^{-1}A)
\end{aligned}$$

以上等式当且仅当$\mu=\overline{X}$ 时成立,即对于固定的$\Sigma>0$,有

$$\ln L(\overline{X},\Sigma)=\max_{\mu}\ln L(\mu,\Sigma)$$

$$
\begin{aligned}
& \ln L(\overline{X}, \Sigma) \\
= & -\frac{np}{2}\ln 2\pi - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\mathrm{tr}(\Sigma^{-1}A) \\
= & -\frac{np}{2}\ln 2\pi - \frac{n}{2}\left[\ln|\Sigma| + \mathrm{tr}\left(\Sigma^{-1}\frac{A}{n}\right)\right] \\
= & C_1 - \frac{n}{2}\left[\mathrm{tr}\left(\Sigma^{-1}\frac{A}{n}\right) - \ln\left|\Sigma^{-1}\frac{A}{n}\right| + \ln\left|\frac{A}{n}\right|\right] \\
= & C_1 - \frac{n}{2}\left[\mathrm{tr}\left(\Sigma^{-1/2}\frac{A}{n}\Sigma^{-1/2}\right) - \ln\left|\Sigma^{-1/2}\frac{A}{n}\Sigma^{-1/2}\right| + \ln\left|\frac{A}{n}\right|\right] \\
\leq & C_1 - \frac{np}{2} - \frac{n}{2}\ln\left|\frac{A}{n}\right|
\end{aligned}
$$

注意到：$\ln|\Sigma| = -\ln|\Sigma^{-1}|$,其中最后一个不等式利用引理2.5.1（见下一页）.

引理2.5.1 设B 为p 阶正定矩阵, 则

$$\mathbf{tr}B - \ln|B| \geq p,$$

且等号成立的充分必要条件是$B = I_p$.

- 在本定理证明中，$B = \Sigma^{-1/2}\frac{A}{n}\Sigma^{-1/2}$ 是正定矩阵.
- 由引理2.5.1知当且仅当$B = \Sigma^{-1/2}\frac{A}{n}\Sigma^{-1/2} = I_p$ 时等号成立, 即

$$\widehat{\Sigma} = \frac{A}{n}.$$

- 似然函数的最大值是

$$L\left(\overline{X}, \frac{A}{n}\right) = (\frac{1}{2\pi e})^{np/2}|A|^{-n/2}.$$

定理2.5.1 设$X_{(i)}(i=1,\cdots,n)$为p元正态总体$N_p(\mu,\Sigma)$的随机样本,$n>p$,则μ,Σ 的最大似然估计为

$$\widehat{\mu}=\overline{X},\ \widehat{\Sigma}=\frac{1}{n}A.$$

三、最大似然估计量的性质

定理2.5.2 设$\overline{X}$和A分别为p元正态总体$N_p(\mu,\Sigma)$的均值向量和样本离差阵,则

1. $\overline{X} \sim N_p(\mu, \frac{1}{n}\Sigma)$.
2. $A \overset{d}{=} \sum_{t=1}^{n-1} Z_t Z_t'$,
   其中：$Z_1, \cdots, Z_{n-1}$ 独立同分布于$N_p(0,\Sigma)$.
3. $\overline{X}$ 与A 相互独立.
4. $P(A>0)=1 \Longleftrightarrow n>p$.

2和3的证明

- 设Γ 是n 阶标准正交矩阵($\Gamma'\Gamma = \Gamma\Gamma' = I$), 具有以下形式

$$\Gamma = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & & \vdots \\ r_{(n-1)1} & \cdots & x_{(n-1)n} \\ 1/\sqrt{n} & \cdots & 1/\sqrt{n} \end{bmatrix} = (r_{ij})_{n\times n}.$$

- 记

$$Z = \begin{bmatrix} Z'_{(1)} \\ \vdots \\ Z'_{(n)} \end{bmatrix} = \Gamma \begin{bmatrix} X'_{(1)} \\ \vdots \\ X'_{(n)} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n r_{1i}X'_{(i)} \\ \vdots \\ \sum_{i=1}^n r_{ni}X'_{(i)} \\ 1/\sqrt{n}\sum_{i=1}^n X'_{(i)} \end{bmatrix}.$$

即

$$Z_{(i)} = \sum_{i=1}^{n} r_{ij} X_{(j)} = (X_{(1)}, \cdots, X_{(n)}) \begin{bmatrix} r_{i1} \\ \vdots \\ r_{in} \end{bmatrix}.$$

计算得

$$\begin{aligned} \mathsf{E}\left[Z_{(i)}\right] &= \mathsf{E}\left[\sum_{j=1}^{n} X_{(j)} r_{ij}\right] = \sum_{j=1}^{n} r_{ij}\mu = 0. \\ \mathsf{Var}\left[Z_{(i)}\right] &= \mathsf{Var}\left[\sum_{j=1}^{n} X_{(j)} r_{ij}\right] = \sum_{j=1}^{n} r_{ij}^2 \Sigma = \Sigma. \\ \mathsf{Cov}\left[Z_{(i)}, Z_{(k)}\right] &= \sum_{j=1}^{n} r_{ij} r_{kj} \Sigma = 0. \end{aligned}$$

$Z_{(i)}$ 是p维正态随机向量的线性组合，因此，仍服从正态分布.

- $Z_{(i)} \sim N_p(0, \Sigma),\ i = 1, \ldots, n-1$;
- $Z_{(n)} = \sqrt{n}\overline{X} \sim N(\sqrt{n}\mu, \Sigma)$;

由于Γ的正交性，

$$
\begin{aligned}
Z'Z &= \sum_{i=1}^{n} Z_{(i)} Z'_{(i)} = \sum_{i=1}^{n-1} Z_{(i)} Z'_{(i)} + Z_{(n)} Z'_{(n)}, \\
Z'Z &= X'\Gamma' \cdot \Gamma X = X'X, \\
X'X &= \sum_{i=1}^{n} X_{(i)} X'_{(i)}.
\end{aligned}
$$

- 由此可得(2)，

$$\sum_{i=1}^{n-1} Z_{(i)}Z'_{(i)} = \sum_{i=1}^{n} X_{(i)}X'_{(i)} - Z_{(n)}Z'_{(n)} = \sum_{i=1}^{n} X_{(i)}X'_{(i)} - n\overline{X}\overline{X}' = A.$$

离差阵A可以表示如下，

$$A = \sum_{i=1}^{n-1} Z_{(i)}Z'_{(i)}.$$

- 因为$Z_{(i)}$是独立的，因此，$Z_{(n)} = \sqrt{n}\overline{X}$ 与A独立, 即得(3).

4.证明 由上面知$A=\sum_{i=1}^{n-1} Z_{(i)}Z'_{(i)} \hat{=} BB'$, B为$p\times n-1$的矩阵.

- $\Pr(A>0)=1 \Rightarrow n>p$.

  A a.s为正定矩阵, 则A的秩是p, $B_{p\times n-1}$ 的秩也为p,
  即$p\leq n-1$, 或$p<n$.

- $n>p \Rightarrow \Pr(A>0)=1$. 只要证明

$$\Pr(B\text{的前}p\text{列为线性相关})=0.$$

（下面续）

$p < n$, 则有$p \leq n-1$.

(1) $A = BB'$, 其中矩阵$B = (Z_{(1)}, \cdots, Z_{(n-1)})$. $R(B) = p$.

(2) 前p 列为独立同分布，服从正态分布的随机向量.

(3) 由正态分布的性质知$Z_{(1)}, \cdots, Z_{(p)}$ 的线性组合也服从正态分布, 即对于任意非零的向量$(\beta_1, \cdots, \beta_p)$,

$$\sum_{i=1}^{p} \beta_i Z_{(i)} \sim \text{正态分布}.$$

(4) 由$\sum_{i=1}^{p} \beta_i Z_{(i)}$为正态分布, 因此,

$$\Pr\left(\sum_{i=1}^{p} \beta_i Z_{(i)} = 0\right) = 0 \quad \Rightarrow \quad (Z_{(1)}, \cdots, Z_{(p)})\text{线性无关}.$$

统计意义下线性无关.

μ 和Σ 的估计所具有的一些性质.

1. 无偏性$\mathsf{E}[\overline{X}] = \mu$, $\mathsf{E}[S] = \frac{1}{n-1}\mathsf{E}[A] = \Sigma$.
2. $\overline{X}$ 和$S(\widehat{\Sigma})$是充分统计量.
3. 可以证明$\overline{X}$ 和$\widehat{\Sigma}$ 是极大极小估计.
4. 利用大数定理证明相合性.
5. 可以证明渐近正态性.
6. 其他

四、函数参数的最大似然估计 设参数向量θ的变化范围

是$\Theta \in \mathbb{R}_k$.$L(\Theta)$是似然函数. 设$w = g(\theta)$ 是Θ 到Θ^* 的Borel 可测映射, 这里Θ^* 是$\mathbb{R}_k$的子集.

对任何$w \in \Theta^*$,令

$$M(w) = \sup_{\theta: g(\theta)=w} L(\theta).$$

定义2.5.1 称$M(w)$为函数$g(\theta)$诱导出的**似然函数**.

定义2.5.2 若$\widehat{w}$满足$M(\widehat{w}) = \sup_w M(w)$, 则称$\widehat{w}$是$g(\theta)$的**最大似然估计**.

定理2.5.3 若$\widehat{\theta}$是θ的**最大似然估计**,则$\widehat{w} = g(\widehat{\theta})$是$g(\theta)$的**最大似然估计**.

例2.5.1 设$X=(X_1,X_2,\cdots,X_p)'\sim N_p(\mu,\Sigma)$, X_i,X_j的相关系数

$$\rho_{ij}=\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$$

的最大似然估计为

$$r_{ij}=\frac{a_{ij}}{\sqrt{a_{ii}}\sqrt{a_{jj}}}$$

其中：a_{ij} 为离差阵A的第i 行，第j列的元素.

例2.5.2 设

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \sim N_p\left(\begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

则$X^{(1)}$对$X^{(2)}$的回归系数阵$B = \Sigma_{12}\Sigma_{22}^{-1}$ 的极大似然估计为

$$\widehat{B} = \widehat{\Sigma}_{12}\widehat{\Sigma}_{22}^{-1} = A_{12}A_{22}^{-1}$$

及$X^{(2)}$给定时$X^{(1)}$的条件协方差阵$\Sigma_{11\cdot 2}$ 的极大似然估计量为

$$\widehat{\Sigma}_{11\cdot 2} = \frac{1}{n}(A_{11} - A_{12}A_{22}^{-1}A_{21}).$$

Definition The random vector $\mathbf{X}$ is said to have an elliptical distribution with parameters μ and Σ if its characteristic function can be expressed as

$$\mathsf{E}[\exp(it'\mathbf{X})] = \exp(it'\mu)\phi(t'\Sigma t),$$

for some function $\phi : \mathbb{R} \to \mathbb{R}$ and where Σ is given by

$$\Sigma = AA'$$

for some $n \times m$ matrix A. $\mathbf{X} \sim Ell_n(\mu, \Sigma, \phi)$.

Property 1

An $n-$dimensional random vector $\mathbf{X}$ has $Ell_n(\mu, \Sigma, \phi)$ distribution if, and only if, for any vector $\alpha \in \mathbb{R}^n$, one has

$$\alpha' \mathbf{X} \sim Ell_1(\alpha' \mu, \alpha' \Sigma \alpha, \phi)$$

Property 2 For $k = 1, \ldots, n$,

$$X_k \sim Ell_1(\mu_k, \sigma_k^2, \phi)$$

Let

$$S = \sum_{k=1}^{n} X_k = \mathbf{e}'\mathbf{X}$$

where $\mathbf{e} = (1, \ldots, 1)$, it follows that

$$\mathbf{X} \sim Ell_n(\mu, \Sigma, \phi) \Rightarrow Ell_1(\mathbf{e}'\mu, \mathbf{e}'\Sigma\mathbf{e}, \phi)$$

where $\mathbf{e}'\mu = \sum_{k=1}^{n} \mu_k$ and $\mathbf{e}'\Sigma\mathbf{e} = \sum_{k=1}^{n} \sum_{l=1}^{n} \sigma_{kl}$.

Property 3 For any $m \times n$ matrix $\mathbf{B}$, any vector $\mathbf{c} \in \mathbb{R}^m$ and random vector $\mathbf{X} \sim Ell_n(\mu, \Sigma, \phi)$, we have

$$\mathbf{BX} + \mathbf{c} \sim Ell_n(\mathbf{B}\mu + \mathbf{c}, \mathbf{B\Sigma B}', \phi).$$

An elliptically distributed random vector $\mathbf{X} \sim Ell_n(\mu, \Sigma, \phi)$. A necessary condition for $\mathbf{X}$ to have a density is that $\operatorname{rank}(\Sigma) = n$. If $\mathbf{X} \sim Ell_n(\mu, \Sigma, \phi)$ has a density, then it will be of the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c}{\sqrt{|\Sigma|}} g((\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)).$$

- $g_n(u) = \exp(-u/2)$, $c_n = (2\pi)^{-n/2} \Rightarrow \mathbf{X} \sim N_n(\mu, \Sigma)$.
- $g_n(u) = \left(1 + \frac{u}{m}\right)^{(n+m)/2}$, $c_n = \frac{\Gamma((n+m)/2)}{(m\pi)^{n/2}\Gamma(m/2)}$. $\Rightarrow t_n^m(\mu, \Sigma)$.

Some families of elliptical distributions with their characteristic generator and/or density generator.

- Cauchy, $g_n(u) = (1+u)^{-(n+1)/2}$;
- Exponential power, $g_n(u) = \exp(-ru^s)$, $r, s > 0$;
- Laplace, $g_n(u) = \exp(-|u|)$;
- Logistic, $g_n(u) = \frac{\exp(-u)}{(1+\exp(-u))^2}$;
- Normal, $g_n(u) = \exp(-u/2)$;
- Stable laws, $\phi(u) = \exp\left(-ru^{s/2}\right)$, $0 < s \leq 2$, $r > 0$;
- Student, $g_n(u) = \left(1 + \frac{u}{m}\right)^{(n+m)/2}$, $m > 0$ an integer.