7.2 多元正态分布参数的估计与假设检验

一、参数 μ 和 Σ 的估计

二、正态总体均值向量的假设检验

一、参数 μ 和 Σ 的估计

设$X \sim N_p(\mu,\Sigma)$, $(X_1, X_2,\cdots, X_n)^{\mathrm{T}}$为来自总体$N_p(\mu,\Sigma)$的容量为$n$的样本$(n > p)$.

这里每个X_i都为p维向量$(i = 1,2,\cdots,n)$.

即

$$(X_1, X_2,\cdots, X_n)^{\mathrm{T}} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{pmatrix}_{n\times p}$$

为简单计,仅考虑$|\Sigma| \neq 0$的情形,即非退化情形.称

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad (7.5)$$

为样本均值向量。

$$S = \sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})^T \qquad (7.6)$$

为样本离差阵。

定理7.1

若$(X_1, X_2, \cdots, X_n)^T$为取自非退化$p$维正态总体$N_p(\mu,\Sigma)$的容量为$n$的样本, $f(x_i;\mu,\Sigma)$为X_i的分布密度,则

$\bar{X}$是μ的最大似然估计,

S/n是Σ的最大似然估计. 且

$$\max_{\mu,\Sigma}\prod_{i=1}^{n} f(x_i;\ \mu,\Sigma)=\prod_{i=1}^{n} f(x_i;\ \bar{X},S/n)$$

$$=\exp\{-\frac{np}{2}\ \ln(2\pi)-\frac{n}{2}\ \ln\left|\frac{S}{n}\right|-\frac{np}{2}\}$$

定理7.2

若$(X_1, X_2, \cdots, X_n)^T$为来自$p$维正态总体$N_p(\mu, \Sigma)$的容量为$n$的样本,则

$\bar{X}$是μ的最小方差无偏估计,

$S/(n-1)$是Σ的最小方差无偏估计.

定理7.3 ***

若$(X_1, X_2, \cdots, X_n)^T$为来自$p$维正态总体$N_p(\mu,\Sigma)$的容量为$n$的样本，$\bar{X}, S$分别由式(7.5)和式(7.6)给定，则

(1) $\bar{X} \sim N_p(\mu, \frac{1}{n}\Sigma)$;

(2) 存在相互独立同分布于$N_p(0,\Sigma)$的随机变量 $Y_1, \cdots, Y_{n-1}$，使S可表示为

$$S = \sum_{i=1}^{n-1} Y_i Y_i^T; \qquad (7.7)$$

(3) $\bar{X}$与S相互独立.

证明: (1)　因为

$$\bar{X}=\begin{pmatrix}\bar{X}_1 & \bar{X}_2 & \cdots & \bar{X}_p\end{pmatrix}^T=\begin{pmatrix}\frac{1}{n}(X_{11}+X_{21}+\cdots+X_{n1})\\ \frac{1}{n}(X_{12}+X_{22}+\cdots+X_{n2})\\ \vdots \\ \frac{1}{n}(X_{1p}+X_{2p}+\cdots+X_{np})\end{pmatrix}$$

$$=\frac{1}{n}\begin{pmatrix}X_{11} & X_{21} & \cdots & X_{n1}\\ X_{12} & X_{22} & \cdots & X_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ X_{1p} & X_{2p} & \cdots & X_{np}\end{pmatrix}_{p\times n}\begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix}_{n\times 1}$$

$$
= \frac{1}{\boldsymbol{n}} \left[\begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} & \cdots & \boldsymbol{X}_{1\boldsymbol{p}} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} & \cdots & \boldsymbol{X}_{2\boldsymbol{p}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}_{\boldsymbol{n}1} & \boldsymbol{X}_{\boldsymbol{n}2} & \cdots & \boldsymbol{X}_{\boldsymbol{np}} \end{pmatrix}_{\boldsymbol{n}\times \boldsymbol{p}} \right]^{\boldsymbol{T}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{\boldsymbol{n}\times 1}
$$

$$
= \frac{1}{\boldsymbol{n}} \left[\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}_{1\times \boldsymbol{n}} \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} & \cdots & \boldsymbol{X}_{1\boldsymbol{p}} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} & \cdots & \boldsymbol{X}_{2\boldsymbol{p}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}_{\boldsymbol{n}1} & \boldsymbol{X}_{\boldsymbol{n}2} & \cdots & \boldsymbol{X}_{\boldsymbol{np}} \end{pmatrix}_{\boldsymbol{n}\times \boldsymbol{p}} \right]^{\boldsymbol{T}}
$$

所以

$$\bar{X}=\left(\bar{X}_1 \quad \bar{X}_2 \quad \cdots \quad \bar{X}_p\right)^T$$

$$=\left[\frac{1}{n}\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}_{1\times n}\begin{pmatrix}X_{11} & X_{12} & \cdots & X_{1p}\\ X_{21} & X_{22} & \cdots & X_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ X_{n1} & X_{n2} & \cdots & X_{np}\end{pmatrix}_{n\times p}\right]^T$$

可看为$n\times p$维正态变量$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$的线性变换，

由性质3，$\bar{X}$服从正态分布。

(2), (3)的证明

若记$p \times n$矩阵为

$$X^* = (X_1 - \mu, \mathrm{L}\ , X_n - \mu)$$

$$= \begin{pmatrix} X_{11} - \mu_1 & X_{21} - \mu_1 & \cdots & X_{n1} - \mu_1 \\ X_{12} - \mu_2 & X_{22} - \mu_2 & \cdots & X_{n2} - \mu_2 \\ \vdots & \vdots & \ddots & \vdots \\ X_{1p} - \mu_p & X_{2p} - \mu_p & \cdots & X_{np} - \mu_p \end{pmatrix}_{p \times n}$$

取$n\times n$正交阵U，其第n列元素都是$\dfrac{1}{\sqrt{n}}$，

$$U=\begin{pmatrix} u_{11} & u_{12} & \cdots & \dfrac{1}{\sqrt{n}} \\ u_{21} & u_{22} & \cdots & \dfrac{1}{\sqrt{n}} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & \dfrac{1}{\sqrt{n}} \end{pmatrix}_{n\times n}$$

又令

$$\eta=(Y_1,Y_2,\cdots,Y_n)$$
$$=X^{*}U$$

则 $\eta = (Y_1, Y_2, \cdots, Y_n) = X^* U$

$$= (X_1 - \mu, X_2 - \mu, \cdots, X_n - \mu)\begin{pmatrix} u_{11} & u_{12} & \cdots & \frac{1}{\sqrt{n}} \\ u_{21} & u_{22} & \cdots & \frac{1}{\sqrt{n}} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & \frac{1}{\sqrt{n}} \end{pmatrix}_{n \times n}$$

$$Y_j = \sum_{i=1}^{n} u_{ij}(X_i - \mu),$$

$$j = 1, 2, \cdots, n.$$

$Y_1, Y_2, \cdots, Y_n$ 是由 $X_1, X_2, \cdots, X_n$ 的线性变换构成，所以 $Y_1, Y_2, \cdots, Y_n$ 服从 p 维正态分布.

特别地

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) = \sqrt{n}(\bar{X} - \mu)$$

$$Y_j = \sum_{i=1}^{n} u_{ij}(X_i - \mu), \quad j = 1, 2, \cdots, n.$$

$$E(Y_j) = E\sum_{i=1}^{n} u_{ij}(X_i - \mu) = 0$$

$$U=\begin{pmatrix} u_{11} & u_{12} & \cdots & \frac{1}{\sqrt{n}} \\ u_{21} & u_{22} & \cdots & \frac{1}{\sqrt{n}} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & \frac{1}{\sqrt{n}} \end{pmatrix}_{n\times n}$$

$$Y_j=\sum_{i=1}^{n} u_{ij}(X_i-\mu),\quad j=1,2,\cdots,n.$$

$$\operatorname{cov}(Y_j,Y_k)=\operatorname{cov}\left(\sum_{i=1}^{n} u_{ij}X_i,\sum_{l=1}^{n} u_{lk}X_l\right)$$

$$=\sum_{i=1}^{n} u_{ij}u_{ik}\operatorname{cov}(X_i,X_i)=\left(\sum_{i=1}^{n} u_{ij}u_{ik}\right)\Sigma=\delta_{jk}\Sigma,$$

其中 $\delta_{jk}=\begin{cases}1, & j=k,\\ 0, & j\neq k,\end{cases} \qquad =\begin{cases}\Sigma, & j=k,\\ 0 & j\neq k,\end{cases}$

所以$Y_1,Y_2,\cdots,Y_n$相互独立同服从于正态分布$N_p(0,\Sigma)$.

而$\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T = X^* X^{*T}$

$$=(\eta U^{-1})(\eta U^{-1})^T = \eta U^{-1} U \eta^T$$

$$=\eta\eta^T = \sum_{i=1}^{n} Y_i Y_i^T$$

$$\eta = (Y_1, Y_2, \cdots, Y_n) = X^* U \Rightarrow X^* = \eta U^{-1}$$

由（7.6）

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

$$Y_n = \sqrt{n}(\bar{X}-\mu)$$

$$=\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T - n(\bar{X}-\mu)(\bar{X}-\mu)^T$$

$$=\sum_{i=1}^{n} Y_i Y_i^T - Y_n Y_n^T = \sum_{i=1}^{n-1} Y_i Y_i^T$$

则S与$Y_n=\sqrt{n}(\bar{X}-\mu)$独立，

即S与$\bar{X}$独立.

例 7.2 已知 $X=(X_1, X_2)^{\mathrm{T}}$ 服从正态分布 $N_2(\mu, \Sigma)$, 今从中抽取容量为 20 的一个样本, 得样本值 (见表 7.1).

表7. 1: 样本值

序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_1	63	63	70	6	65	9	10	12	20	30	33	27	21	5	14	27	17	53	62	65
x_2	971	892	1125	82	931	112	162	321	315	375	462	352	305	34	29	332	185	703	872	740

229

试求 μ 和 Σ 的最小方差无偏估计值。

解 由定理 7.2 知 μ 的最小方差无偏估计

$$\hat{\mu}=\bar{X}=\frac{1}{20}\begin{bmatrix}63+63+\mathrm{L}\ +65\\971+892+\mathrm{L}\ +740\end{bmatrix}$$

$$=\begin{bmatrix}33.85\\477.50\end{bmatrix}=\begin{bmatrix}\bar{x}_1\\\bar{x}_2\end{bmatrix}.$$

样本离差阵

$$S=\sum_{k=1}^{20}(X_k-\bar{X})(X_k-\bar{X})^T$$

$$=\sum_{k=1}^{20}\left[\begin{bmatrix} x_{k_1} \\ x_{k_2} \end{bmatrix}-\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}\right]\left[\begin{bmatrix} x_{k_1} \\ x_{k_2} \end{bmatrix}-\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}\right]^T$$

$$=\begin{bmatrix} \sum_{k=1}^{20}(x_{k_1}-\bar{x}_1)^2 & \sum_{k=1}^{20}(x_{k_1}-\bar{x}_1)(x_{k_2}-\bar{x}_2) \\ \sum_{k=1}^{20}(x_{k_2}-\bar{x}_2)(x_{k_1}-\bar{x}_1) & \sum_{k=1}^{20}(x_{k_2}-\bar{x}_2)^2 \end{bmatrix}$$

$$=\begin{bmatrix} 10838.55 & 149056.50 \\ 149056.50 & 2135681.00 \end{bmatrix}$$

所以Σ的最小方差无偏估计$\hat{\Sigma}$为

$$\hat{\boldsymbol{\Sigma}}=\frac{1}{20-1}\boldsymbol{S}=\begin{bmatrix} 570.45 & 7845.08 \\ 7845.08 & 112404.26 \end{bmatrix}$$

二、正态总体均值向量μ的假设检验

第 4 章, 我们曾讨论了单个正态总体与两个正态总体均值的有关检验。

现在讨论单个和两个 p 维正态总体均值向量的有关检验.

类似于一维情形, 在此分别对协差阵已知与未知两种情形进行讨论.

1.协差阵 Σ 已知时，均值向量 μ 的检验

设$(X_1, X_2, \cdots, X_n)$为来自p维正态总体$N_p(\mu,\Sigma)$的容量为n的样本,其中Σ已知.要检验假设

$$H_0: \mu = \mu_0（已知）\leftrightarrow H_1: \mu \neq \mu_0,$$

为了检验假设H_0,引入统计量,

$$\eta = n(\bar{X} - \mu_0)^T \Sigma^{-1} (\bar{X} - \mu_0)$$

由定理7.3知,$\bar{X} \sim N_p(\mu, \frac{1}{n}\Sigma)$,又由性质8知,

当H_0成立时,即$\mu = \mu_0$时,$\eta \sim \chi^2(p)$分布.

而当H_0不成立时,即$\mu \neq \mu_0$时,η有偏大的趋势.

因此,对于给定的检验水平α,查χ^2分布表,可得$\chi^2_\alpha(p)$的值使

$$P\{\eta \geq \chi^2_\alpha(p)\} = \alpha.$$

根据样本值可计算出η的值η_0,

若$\eta_0 \geq \chi^2_\alpha(p)$则拒绝$H_0$,

即认为总体均值向量与μ_0有显著差异,

若$\eta_0 < \chi^2_\alpha(p)$,则接受$H_0$,

即认为总体均值向量与μ_0无显著差异.

注：计算$\eta = n(\bar{X}-\mu_0)^T\Sigma^{-1}(\bar{X}-\mu_0)$时，需要计算$\Sigma^{-1}$，计算量会很繁琐。

一个避免繁琐的小技巧：

先计算$\bar{X}-\mu_0$，

令 $$b=\Sigma^{-1}(\bar{X}-\mu_0) \tag{7.8}$$

则 $$\Sigma b=(\bar{X}-\mu_0) \tag{7.9}$$

（7.9）是b满足的线性方程组，由于Σ已知，解方程组可求出b,然后就可求出

$$\eta = n(\bar{X}-\mu_0)^T b$$

2.协差阵 Σ 未知时, 均值向量 μ 的检验

设$(X_1, X_2, \cdots, X_n)$为来自p维正态总体$N_p(\mu, \Sigma)$的容量为n的样本,其中Σ未知.要检验假设

$$H_0: \mu = \mu_0(\text{已知}) \leftrightarrow H_1: \mu \neq \mu_0,$$

由于Σ未知，不能使用统计量

$$\eta = n(\bar{X} - \mu_0)^T \Sigma^{-1} (\bar{X} - \mu_0)$$

但由定理7.2知, $S/(n-1)$是Σ的最小方差无偏估计,从而可用$(n-1)S^{-1}$代替Σ^{-1},这里S^{-1}为样本离差阵S的逆矩阵.

引入统计量

$$F=\frac{(n-p)}{(n-1)p}\,T^2$$

其中$T^2=n(n-1)(\bar{X}-\mu_0)^T S^{-1}(\bar{X}-\mu_0)$,

称T^2为霍太林统计量,可以证明,

当H_0成立时,即$\mu=\mu_0$时,$F\sim F(p,\ n-p)$.

而当H_1成立时,统计量F有偏大的趋势.

因此F可作为检验假设H_0的统计量.

当给定检验水平α时,可查出$F_\alpha(p,\ n-p)$的值,使

$$P\{F\geq F_\alpha(p,\ n-p)\}=\alpha.$$

根据样本值求出F的值F^*,

当$F^* \geq F_\alpha(p, n-p)$时,

则拒绝H_0,即认为总体均值向量与μ_0有显著差异,

当$F^* < F_\alpha(p, n-p)$时,

则接受H_0,即认为总体均值向量与μ_0无显著差异.

3. 两个正态总体均值向量是否相等的检验 (Σ已知)

设$(X_1, X_2, \cdots, X_m)$为来自p维正态总体$N_p(\mu_1, \Sigma)$的容量为m的样本,$(Y_1, Y_2, \cdots, Y_n)$为来自p维正态总体$N_p(\mu_2, \Sigma)$的容量为n的样本,$m > p, n > p$且两个样本相互独立,Σ是两个正态总体共同的协方差$(\Sigma > 0)$.要检验假设

$$H_0: \mu_1 = \mu_2 \leftrightarrow H_1: \mu_1 \neq \mu_2,$$

当Σ已知时,引入统计量

$$\chi^2_{mn} = \frac{mn}{m+n}(\bar{X}-\bar{Y})^T \Sigma^{-1}(\bar{X}-\bar{Y})$$

其中 $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$,

利用定理7.3及χ^2分布的性质可以证明,

当H_0成立时,χ^2_{mn}服从自由度为p的χ^2分布;

当H_1成立时,χ^2_{mn}的值有偏大的趋势,

因此χ^2_{mn}可作为检验假设H_0的统计量.

对于给定的检验水平α,查χ^2分布表可得$\chi_{\alpha}^{2}(p)$的值,使

$$P\{\chi_{mn}^{2} \geq \chi_{\alpha}^{2}(p)\} = \alpha.$$

由样本值计算χ_{mn}^{2}的观测值χ_{0}^{2},当$\chi_{0}^{2} \geq \chi_{\alpha}^{2}(p)$时,拒绝$H_0$,即认为两正态总体的均值向量有显著差异;当$\chi_{0}^{2} < \chi_{\alpha}^{2}(p)$时,接受$H_0$,即认为两正态总体的均值向量无显著差异

4. 两个正态总体均值向量是否相等的检验(Σ未知)

当 Σ 未知时， 检验的统计量为

$$F=\frac{mn(m+n-p-1)}{p(m+n)(m+n-2)}(\bar{X}-\bar{Y})^{T}S^{-1}(\bar{X}-\bar{Y}),$$

其中 $\bar{X}=\frac{1}{m}\sum_{i=1}^{m}X_i,\quad \bar{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i,$

$$S_1=\sum_{k=1}^{m}(X_k-\bar{X})(X_k-\bar{X})^{T},\qquad S_2=\sum_{k=1}^{n}(Y_k-\bar{Y})(Y_k-\bar{Y})^{T}$$

$$S^{-1}=(m+n-2)(S_1+S_2)^{-1},$$

这里S是协差阵Σ的估计量，S_1,S_2分别是$N_p(\mu_1,\Sigma)$与$N_p(\mu_2,\Sigma)$的样本离差阵.

可以证明,当$H_0: \mu_1 = \mu_2$成立时

$$F \sim F(p, m+n-p-1)$$

而当H_1成立时,F有偏大的趋势,因此,对给定的检验水平α,查F分布表,可得$F_\alpha(p, m+n-p-1)$的值,再由样本值求出统计量的观察值F_0,

若$F_0 \geq F_\alpha(p, m+n-p-1)$,拒绝$H_0$,

即认为两正态总体的均值向量有显著差异,

若$F_0 < F_\alpha(p, m+n-p-1)$,则接受$H_0$,

即认为两个正态总体的均值向量无显著差异.

例7.3　设X,Y为两个正态总体，$X \sim N_2(\mu_1,\Sigma)$，$Y \sim N_2(\mu_2,\Sigma)$，Σ未知$(\Sigma > 0)$，今从两个总体中随机抽取两个相互独立的样本(每个总体抽取一个，样本量为4)，得样本值如下(见表7.2)。试在检验水平$\alpha = 0.01$下检验假设

$$H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2 .$$

表7.2

	X_1	X_2	X_3	X_4		Y_1	Y_2	Y_3	Y_4
x_{k_1}	131.5	145	141	150	y_{k_1}	40.5	80	50	90
x_{k_2}	9	12	30	36	y_{k_2}	54	74.5	64.5	60.5

解 本例中$m=n=4$, $p=2$.经计算知

$$\bar{X}=\begin{bmatrix}\bar{x}_1\\ \bar{x}_2\end{bmatrix}=\begin{bmatrix}141.875\\ 21.75\end{bmatrix},\quad \bar{Y}=\begin{bmatrix}\bar{y}_1\\ \bar{y}_2\end{bmatrix}=\begin{bmatrix}65.125\\ 63.375\end{bmatrix},$$

$$S=\frac{1}{m+n-2}(S_1+S_2)$$

$$=\frac{1}{6}(S_1+S_2)=\begin{bmatrix}309.90 & 86.36\\ 86.36 & 124.99\end{bmatrix},$$

$$T^2 \triangleq \frac{mn}{m+n}(\bar{X}-\bar{Y})^T S^{-1}(\bar{X}-\bar{Y}) = 116.7,$$

$$F = \frac{m+n-p-1}{p(m+n-2)} T^2$$

$$= \frac{5}{6\times 2}\times 116.7 = 48.625$$

$$> F_{0.01}(2,5) = 13.27,$$

从而拒绝H_0,可认为两个正态总体均值向量的差异高度显著.

Thank you!