线性函数关系的最大似然估计与似然比检验
——关于左$\mathcal{O}(n)$不变分布的场合*

卞国瑞　　张尧庭
（复旦大学）　（武汉大学）

本文在假定误差分布是左$\mathcal{O}(n)$不变的条件下，导出了关于多元线性模型的均值向量之间的线性关系存在性的似然比检验，以及相应参数的最大似然估计量．这些统计量的形式与分布函数的形式无关，因而是稳健的．

§1．问题与记号

在多元线性模型中，检验其均值向量之间是否存在线性关系并估计之，这是一个很有兴趣的问题，已为不少作者所讨论．我们在文〔1〕、〔2〕中，曾将这一问题简化为下面两种法式形式：

模型Ⅰ：

$$\underset{n\times k}{Y}=\begin{pmatrix}Y_1\\Y_2\\Y_3\end{pmatrix}\begin{matrix}p\\r\\n-p-r\end{matrix}=M+\varepsilon,\quad M=\begin{pmatrix}\theta\\\eta\\o\end{pmatrix}\begin{matrix}p\\r\\n-p-r\end{matrix},\tag{1.1}$$

欲检验的假设是

$$H_0:\qquad \exists \underset{r\times p}{C}\ 使得\eta=c\theta\tag{1.2}$$

模型Ⅱ：

$$\underset{n\times k}{Y}=\begin{pmatrix}Y_1\\Y_2\end{pmatrix}\begin{matrix}p\\n-p\end{matrix}=M+\varepsilon,\qquad M=\begin{pmatrix}\theta\\o\end{pmatrix}\begin{matrix}p\\n-p\end{matrix},\tag{1.3}$$

欲检验的假设是

$$H:\quad \exists\underset{r\times p}{\Gamma},\ \Gamma\Gamma'=I_r,\quad (r<p)\ 使得\Gamma\theta=0\tag{1.4}$$

在文〔1〕、〔2〕中，在假定$\varepsilon\sim N(0, I_n, \Sigma)$的条件下关于上述两个模型导出了$H_0$的检验统计量，并在$H_0$为真的条件下获得了这些未知参数的最大似然估计．在文〔3〕中，我们在假定ε的分布为左$\mathcal{O}(n)$不变的条件下，议论了有关H_0的不变检验，并讨论了这些检验的渐近性质．本文将在假定ε为左$\mathcal{O}(n)$不变的条件下，导出H_0的似然比检验，并在H_0为真的条件下，求出相应参数的最大似然估计，这些估计量的形式与ε的分布的具体形式无关，因此这些估计量是稳健的．

记$\mathcal{O}(n)$是$n\times n$正交矩阵的集合，$\mathscr{S}(k)$是$k\times k$正定矩阵的集合，$Gl(k)$是$k\times k$非奇异矩阵的集合．$\mathscr{L}(X)$表示随机矩阵$\underset{n\times k}{X}$的分布函数．如果$\mathscr{L}(\Gamma X)=\mathscr{L}(X)$对一切$\Gamma\in$

* 1984年12月25日收到．

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

(n) 成立，则称 X 是左 $\mathcal{O}(n)$ 不变的．

如果 $\underset{n\times k}{X}$ 是左 $\mathcal{O}(n)$ 不变的，且存在密度函数，则 X 的密度函数可表示为 $q(X'X)$，此时若定义 $Y=M+X\Sigma^{1/2}$，$\Sigma=\Sigma^{1/2}\Sigma^{1/2\prime}>0$，则 Y 的密度函数是

$$f(Y;M,\Sigma)=|\Sigma|^{-n/2}q(\Sigma^{-1/2}(Y-M)'(Y-M)\Sigma^{-1/2\prime}) \tag{1.5}$$

记 Q_M 是左 $\mathcal{O}(n)$ 不变分布的一个子类，定义如下：

$$Q_M=\{q\mid q(VA)=q(AV),\ \forall A\in Gl(k),\ V\in\mathcal{S}(k),\ q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))\ \text{关于}\lambda_i,\ i=1,\cdots,k\ \text{单调不增且可微}\} \tag{1.6}$$

记 $x_{(1)},\cdots,x_{(n)}$ 是 $n\times k$ 随机矩阵 X 的行向量．$\mathrm{Vec}\,X=(x'_{(1)},x'_{(2)},\cdots,x'_{(n)})'$，如果 $\mathcal{L}(\Gamma\mathrm{Vec}\,X)=\mathcal{L}(\mathrm{Vec}\,X)$ 对一切 $\Gamma\in\mathcal{O}(nk)$ 成立．此时说 $\mathrm{Vec}\,X$ 的分布为球对称分布．显然，如果 $\mathrm{Vec}\,X$ 为球对称的，则 X 是左 $\mathcal{O}(n)$ 不变的；如果 $\mathrm{Vec}\,X$ 为球对称的，且存在密度函数，则 X 的密度函数可写为 $q(\mathrm{tr}\,X'X)$；此时若定义 $Y=M+X\Sigma^{1/2}$，$\Sigma=\Sigma^{1/2}\Sigma^{1/2\prime}>0$，则 Y 的密度函数为

$$f(Y;M,\Sigma)=|\Sigma|^{-n/2}q(\mathrm{tr}\,\Sigma^{-1}(Y-M)'(Y-M)) \tag{1.7}$$

记

$$Q_L=\{q\mid q(\mathrm{tr}\,X'X)\ \text{为随机矩阵}\ \underset{n\times k}{X}\ \text{的密度函数},\ q(x)\ \text{关于}\ x\ \text{单调不增且可微}\} \tag{1.8}$$

显然，$Q_L\subset Q_M$

§2．几 个 引 理

引理 1 设 $q(\cdot)$ 为一左 $\mathcal{O}(n)$ 不变随机矩阵 $\underset{n\times k}{X}$ 的密度函数，满足条件．

（i）$q(\cdot)$ 为单调不增函数，即若 $\Sigma_1\geqslant\Sigma_2\geqslant0$，则有 $q(\Sigma_1)\leqslant q(\Sigma_2)$；

（ii）$q(\cdot)$ 连续，即 $\lim\limits_{\Sigma\to\Sigma_0}q(\Sigma)=q(\Sigma_0)$，$\forall\,\Sigma_0\geqslant0$ 则必存在正定矩阵 Σ^*，使得 $|\Sigma|^{n/2}q(\Sigma)$ 在 Σ^* 处达到最大值．

（证明见〔4〕）

引理 2 假定 $q(\cdot)$ 是一左 $\mathcal{O}(n)$ 不变随机矩阵 $\underset{n\times k}{X}$ 的密度函数，$n\geqslant k$，$q\in Q_M$，则

$$\lim_{\lambda\to\infty}\lambda^{nk/2}q(\lambda I)=0 \tag{2.1}$$

证 记 $\lambda_1>\lambda_2>\cdots>\lambda_k>0$ 为 $V=X'X$ 之特征根，则 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 之联合密度函数为

$$C(k,n)C(k,k)\Big(\prod_{i=1}^{k}\lambda_i^{\frac{n-k-1}{2}}\Big)q(\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_k))\prod_{i<j}(\lambda_i-\lambda_j),$$

$$\text{当}\ \lambda_1>\lambda_2>\cdots>\lambda_k>0 \tag{2.2}$$

其中 $C(k,n)=\pi^{nk/2}/\Gamma_k(\frac{n}{2})$，$\Gamma_k(\frac{n}{2})=\pi^{\frac{k(k-1)}{4}}\prod\limits_{j=1}^{k}\Gamma\Big(\frac{n-j+1}{2}\Big)$

因为

$$C(k,n)C(k,k)\int\cdots\int_{0<\lambda_k<\cdots<\lambda_1<\infty}\Big(\prod_{j=1}^{k}\lambda_j^{\frac{n-k-1}{2}}q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))\cdot\prod_{i<j}(\lambda_i-\lambda_j)d\lambda_1\cdots d\lambda_k=1,$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

所以

$$\lim_{y\to\infty}\int\cdots\int_{y<\lambda_k<\cdots<\lambda_1<2y}\Big(\prod_{j=1}^{k}\lambda_j^{\frac{n-k-1}{2}}\Big)q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))\cdot\prod_{i<j}(\lambda_i-\lambda_j)\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_k=0\,. \tag{2.3}$$

由于$q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))$ 关于λ_i，$i=1,\cdots,k$ 单调不增，得

$$\begin{aligned}&\int\cdots\int_{y<\lambda_k<\cdots<\lambda_1<2y}\Big(\prod_{j=1}^{k}\lambda_j^{\frac{n-k-1}{2}}\Big)q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))\prod_{i<j}(\lambda_i-\lambda_j)\cdot\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_k\\&\geqslant y^{\frac{(n-k-1)k}{2}}q(2yI)\int\cdots\int_{y<\lambda_k<\cdots<\lambda_1<2y}\prod_{i<j}(\lambda_i-\lambda_j)\cdot\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_k\\&=c\cdot2^{-\frac{nk}{2}}(2y)^{\frac{nk}{2}}q(2yI)\,,\end{aligned} \tag{2.4}$$

其中

$$c=\int\cdots\int_{0<x_k<\cdots<x_1<1}\begin{vmatrix}1&1&\cdots&1\\x_1&x_2&\cdots&x_k\\\cdots&\cdots&\cdots&\cdots\\x_1^{k-1}&x_2^{k-1}&\cdots&x_k^{k-1}\end{vmatrix}\mathrm{d}x_1\cdots\mathrm{d}x_k\,.$$

联合（2.3）.（2.4）式即得引理立结论.

引理 3 假定引理 2 的条件成立，则

$$h(\lambda_1,\cdots,\lambda_k)=\Big(\prod_{j=1}^{k}\lambda_j^{\frac{n}{2}}\Big)q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k)) \tag{2.5}$$

的最大值，在某有限的a_0I处达到，且a_0为方程

$$q'(xI)+\frac{n}{2x}q(xI)=0\,,\quad x>0 \tag{2.6}$$

之解．其中

$$q'(xI)=\frac{\partial q(\mathrm{diag}(x_1,\cdots,x_k))}{\partial x_1}\Big|_{x_1=\cdots=x_k=x}$$

证 由于$q(VA)=q(AV)$，$\forall A\in Gl(k)$，$V\in\mathscr{S}(k)$，所以$q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))$ 为λ_1，$\cdots$，λ_k的对称函数，因为对称函数$h(\lambda_1,\cdots,\lambda_k)$ 的最大值点一定具有形式：$\lambda_1=\lambda_2=\cdots=\lambda_k$，而$h(0,\cdots,0)=0$，$h(\lambda,\cdots,\lambda)\geqslant0$，由引理 2 知

$$\lim_{\lambda\to\infty}h(\lambda,\cdots,\lambda)=\lim_{\lambda\to\infty}\lambda^{\frac{nk}{2}}q(\lambda I)=0$$

所以$h(\lambda,\cdots,\lambda)$ 的最大值必在有限的a_0处达到.

由于q 可微，得

$$\begin{aligned}\mathrm{d}\log h(\lambda_1,\cdots,\lambda_k)&=\frac{n}{2}\sum_{j=1}^{k}\frac{1}{\lambda_j}\mathrm{d}\lambda_j\\&+\frac{1}{q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))}\sum_{j=1}^{k}\frac{\partial q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))}{\partial\lambda_j}\mathrm{d}\lambda_j\,.\end{aligned}$$

因$q(\mathrm{diag}(\lambda_1,\cdots,\lambda_k))$ 关于$\lambda_1,\cdots,\lambda_k$对称，得

$$\frac{\partial q}{\partial\lambda_i}\Big|_{\lambda_1=\cdots=\lambda_k=\lambda}=\frac{\partial q}{\partial\lambda_j}\Big|_{\lambda_1=\cdots=\lambda_k=\lambda}\,,\quad i,j=1,\cdots,k\,.$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

综上所述，得a_0满足方程（2.6）.

引理 4　假定引理 2 的条件成立，$W>0$，记

$$L(\Sigma)=|\Sigma|^{-n/2}q(\Sigma^{-1}W),\quad \Sigma\in\quad(k) \tag{2.7}$$

则

$$\sup_{\Sigma\in\mathscr{S}(k)}L(\Sigma)=L\left(\frac{1}{a_0}W\right)=|W|^{-n/2}a_0^{nk/2}q(a_0I)\ , \tag{2.8}$$

其中a_0由引理 3 决定.

证　作变换$\widetilde{\Sigma}=W^{\frac{1}{2}}\Sigma^{-1}W^{\frac{1}{2}}$，得

$$L(\Sigma)=|W|^{-\frac{n}{2}}|\widetilde{\Sigma}|^{\frac{n}{2}}q(\widetilde{\Sigma}).$$

记λ_1，…，λ_k是$\widetilde{\Sigma}$的特征根，则

$$|\widetilde{\Sigma}|^{\frac{n}{2}}q(\Sigma)=\prod_{j=1}^{k}\lambda_j^{\frac{n}{2}}q(\mathrm{diag}(\lambda_1,\ \cdots,\ \lambda_k)).$$

由引理 3，当$\Sigma=\lambda_0I$时上式达到最大，所以，当$\Sigma=\frac{1}{a_0}W$时，(2.7）达到最大，且（2.8）成立.

作为引理 4 的特例，有

引理 5　假定$q(\cdot)$为随机矩阵$\underset{n\times k}{X}$的密度函数，$n\geqslant k$，且$q\in Q_L$，$W>0$，记

$$L(\Sigma)=|\Sigma|^{-\frac{n}{2}}q(tr\Sigma^{-1}W),\quad \Sigma\in\mathscr{S}(k)$$

$$h(\lambda)=\lambda^{-\frac{nk}{2}}q\left(\frac{1}{\lambda}trW\right),\qquad \lambda>0$$

则（i）

$$\sup_{\Sigma\in\mathscr{S}(k)}L(\Sigma)=L\left(\frac{k}{\beta_0}W\right)=k^{-\frac{nk}{2}}\beta_0^{\frac{nk}{2}}q(\beta_0)|W|^{-\frac{n}{2}}, \tag{2.9}$$

（ii）

$$\sup_{\lambda>0}h(\lambda)=h\left(\frac{1}{\beta_0}\sum_{j=1}^{k}\lambda_j\right)=\left(\sum_{j=1}^{k}\lambda_j\right)^{-nk/2}\beta_0^{\frac{nk}{2}}q(\beta_0), \tag{2.10}$$

其中β_0是$x^{\frac{nk}{2}}q(x)$的最大值点，它也是下面方程的解：

$$q'(x)+\frac{nk}{2x}q(x)=0\ ,$$

λ_1，…，λ_k是W的特征根.

引理 6　设A是一$(p+r)\times(p+r)$非负定矩阵，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{p+r}$是$A$的特征根，$\gamma_1$，$\gamma_2$，…，$\gamma_{p+r}$是分别对应于$\lambda_1$，…，$\lambda_{p+r}$的$A$的特征向量，记正交矩阵

$$T=(\gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{q+r})=\begin{array}{cc} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} & \begin{array}{c} p \\ r \end{array} \\ \begin{array}{cc} \quad p & \qquad r \end{array} & \end{array}$$

且$|T_{22}|\neq 0$.

则（i）

$$\inf_{B\in\mathscr{B}}\mathrm{tr}(I_r+BB')^{-\frac{1}{2}}(-B,\ I_r)A\begin{pmatrix}-B'\\ I_r\end{pmatrix}(I_r+BB')^{-\frac{1}{2}}=\sum_{j=p+1}^{p+r}\lambda_j\ , \tag{2.11}$$

（ii）

$$\inf_{B\in\mathscr{B}}\det\left[I_r+(I_r+BB')^{-\frac{1}{2}}(-B,\ I_r)A\begin{pmatrix}-B'\\ I_r\end{pmatrix}(I_r+BB')^{-\frac{1}{2}}\right]$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd.　All rights reserved.

$$=\prod_{j=p+1}^{p+r}(1+\lambda_j)\quad, \tag{2.12}$$

且这些极小值在$B_0=-(T'_{22})^{-1}T'_{12}$处达到．其中$\mathscr{B}=\{B \mid B$是$r\times p$矩阵$\}$．

引理 7 设A是一$n\times n$对称矩阵，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$是$A$的特征根，$\gamma_1$，…，$\gamma_n$是分别对应于$\lambda_1$，…，$\lambda_n$的$A$的特征向量，则

(i) $$\min_{\Gamma\in\mathscr{D}}\operatorname{tr}\Gamma' A\Gamma=\sum_{j=n-r+1}^{n}\lambda_j\quad, \tag{2.13}$$

(ii) $$\min_{\Gamma\in\mathscr{D}}\det(I_r+\Gamma' A\Gamma)=\prod_{j=n-r+1}^{n}(1+\lambda_j)\quad, \tag{2.14}$$

且这些最小值在$\Gamma=(\gamma_{n-r+1},\cdots,\gamma_n)$处达到．其中$\mathscr{D}=\{\Gamma \mid \Gamma$是$n\times r$矩阵，$\Gamma'\Gamma=I_r\}$．

§3．关于模型 I 的主要结果

定理3.1 假定模型（1.1）成立，$n\geqslant p+r+k$，ε是左$\mathcal{O}(n)$不变的，Y的密度函数是

$$f(Y;M,\Sigma)=|\Sigma|^{-\frac{n}{2}}q(\Sigma^{-\frac{1}{2}}(Y-M)'(Y-M)\Sigma^{-\frac{1}{2}}),\quad \Sigma\in\mathscr{S}(k) \tag{3.1}$$

且$q(\cdot)$满足引理 1 的条件，则当由（1.2）确定的假设H_0成立时，θ,c,Σ的最大似然估计是

$$\hat{\theta}=(I_p+\hat{c}'\hat{c})^{-1}(Y_1+\hat{c}'Y_2),\qquad \hat{c}=-(T'_{22})^{-1}T'_{12} \tag{3.2}$$

$$\hat{\Sigma}=\hat{W}^{\frac{1}{2}}(\hat{W}^{\frac{1}{2}}\Sigma^*\hat{W}^{\frac{1}{2}\prime})^{-\frac{1}{2}}\hat{W}(\hat{W}^{\frac{1}{2}}\Sigma^*\hat{W}^{\frac{1}{2}\prime})^{-\frac{1}{2}}\hat{W}^{\frac{1}{2}}\quad,$$

$$\hat{W}=(T'_{12}Y_1+T'_{22}Y_2)'(T'_{12}Y_1+T'_{22}Y_2)+Y'_3Y_3\quad,$$

其中（i）Σ^*由引理 1 所定义；

（ii）T_{12}，T_{22}由下决定，记$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{p+r}$是$\begin{pmatrix}Y_1\\Y_2\end{pmatrix}(Y'_3\ \ Y_3)^{-1}(Y'_1\ \ Y'_2)$之特征根，$\gamma_1$，$\gamma_2$，…，$\gamma_{p+r}$分别是对应于$\lambda_1$，$\lambda_2$，…，$\lambda_{p+r}$的特征向量．

$$T=(\gamma_1,\gamma_2,\cdots,\gamma_{p+r})=\begin{array}{cc} \begin{pmatrix}T_{11} & T_{12}\\ T_{21} & T_{22}\end{pmatrix} & \begin{array}{c}p\\ r\end{array}\\ \begin{array}{cc}\ \ p\ \ & \ \ r\ \ \end{array} & \end{array} \tag{3.3}$$

证 由（3.1）似然函数是

$$L(\theta,c,\Sigma;Y)=|\Sigma|^{-\frac{n}{2}}q(\Sigma^{-\frac{1}{2}}[(Y_1-\theta)'(Y_1-\theta)+(Y_2-c\theta)'(Y_2-c\theta)+Y'_3Y_3]\Sigma^{-\frac{1}{2}})$$

由引理 1 得

$$\sup_{\Sigma\in\mathscr{S}(k)}L(\theta,c,\Sigma,Y)=|(Y_1-\theta)'(Y_1-\theta)+(Y_2-c\theta)'(Y_2-c\theta)+Y'_3Y_3|^{-\frac{n}{2}}|\Sigma^*|^{\frac{n}{2}}q(\Sigma^*) \tag{3.4}$$

此极大值当Σ_0满足下式时达到

$$\Sigma^*=\Sigma_0^{-\frac{1}{2}}[(Y_1-\theta)'(Y_1-\theta)+(Y_2-c\theta)'(Y_2-c\theta)+Y'_3Y_3]\Sigma_0^{-\frac{1}{2}}\triangleq\Sigma_0^{-\frac{1}{2}}W\Sigma_0^{-\frac{1}{2}}, \tag{3.5}$$

其中$W=(Y_1-\theta)'(Y_1-\theta)+(Y_2-c\theta)'(Y_2-c\theta)+Y'_3Y_3$．由（3.5）得

$$\Sigma_0=W^{\frac{1}{2}}(W^{\frac{1}{2}}\Sigma^*W^{\frac{1}{2}\prime})^{-\frac{1}{2}}W(W^{\frac{1}{2}}\Sigma^*W^{\frac{1}{2}\prime})^{-\frac{1}{2}}W^{\frac{1}{2}} \tag{3.6}$$

记 $$\theta_0=(I_p+c'c)^{-1}(Y_1+c'Y_2)\quad, \tag{3.7}$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

则 $W=(Y_1-\theta_0)'(Y_1-\theta_0)+(Y_2+c\theta_0)'(Y_2-c\theta_0)+Y_3'Y_3+(\theta-\theta_0)'(\theta-\theta_0)$

$$\triangleq W_1+(\theta-\theta_0)'(\theta-\theta_0). \tag{3.8}$$

所以当$\theta=\theta_0$时$|W|^{-n/2}$达到最大，且最大值是$|W_1|^{-n/2}$，而

$$W_1=(Y_1-\theta_0)'(Y_1-\theta_0)+(Y_2-c\theta_0)'(Y_2-c\theta_0)+Y_3'Y_3$$

$$=Y_3'Y_3+(Y_1'\ \ Y_2')[I_{p+r}-\begin{bmatrix}I_p\\C\end{bmatrix}(I_p+c'c)^{-1}(I_p c')]\begin{bmatrix}Y_1\\Y_2\end{bmatrix}$$

$$=(Y_3'Y_3)^{1/2}(I_k+(Y_3'Y_3)^{-1/2}(Y_1'\ \ Y_2')\begin{bmatrix}-c'\\I_r\end{bmatrix}(I_r+cc')^{-1}(-c\,I_r)\begin{bmatrix}Y_1\\Y_2\end{bmatrix}$$

$$(Y_3'Y_3)^{-1/2'})(Y_3'Y_3)^{1/2'}\ . \tag{3.9}$$

由引理6及$P(|T_{22}|=0)=0$ 得到：当

$$\hat{c}=-(T_{22}')^{-1}T_{12}' \tag{3.10}$$

时$|W_1|$达到极小．且极小值为

$$\prod_{j=p+1}^{p+r}(1+\lambda_j)|Y_3'Y_3|. \tag{3.11}$$

综合（3.6),(3.7),(3.9）和（3.10）即得θ，c，Σ的最大似然估计是（3.2)．且$L(\theta,c,\Sigma;Y)$的极大值是

$$L(\hat{\theta},\hat{c},\hat{\Sigma};Y)=|\Sigma^*|^{n/2}q(\Sigma^*)\Big(\prod_{j=p+1}^{p+r}(1+\lambda_j)\Big)^{-n/2}|Y_3'Y_3|^{-n/2} \tag{3.12}$$

定理3.2 假定模型（1.1）成立，$n>p+r+k$，Y的密度函数是

$$f(Y;M,\Sigma)=|\Sigma|^{-n/2}q(\Sigma^{-1}(Y-M)'(Y-M)),\qquad \Sigma\in\mathcal{S}(k)$$

且 $q\in Q_M$，则当由（1.2）确定的假设H_0成立时，θ,c,Σ的最大似然估计是

$$\hat{\theta}=(I_p+\hat{c}'\hat{c})^{-1}(Y_1+\hat{c}'Y_2),\quad \hat{c}=-(T_{22}')^{-1}T_{12}'\ , \tag{3.13}$$

$$\hat{\Sigma}=\frac{1}{a_0}((T_{12}'Y_1+T_{22}'Y_2)'(T_{12}'Y_1+T_{22}'Y_2)+Y_3'Y_3)$$

其中T_{12}．T_{22}的定义同定理1，a_0由引理3决定，如果$q\in Q_L$，则a_0用相应的β_0/k代替，β_0由引理5决定．

证明类似于定理1的证明，仅需将引理4代替引理1，因而略去．

定理3.3 假定定理1的条件成立，关于由（1.2）确定的假设

$$H_0:\ \exists\ \underset{r\times p}{c}\ 使得\eta=c\theta$$

其似然比检验域为

$$\prod_{j=p+1}^{p+r}(1+\lambda_j)\geqslant c_a \tag{3.14}$$

或等价地为

$$\prod_{j=p+1}^{k}(1+\lambda_j^*)\geqslant c_a' \tag{3.15}$$

其中$\lambda_1^*\geqslant\lambda_2^*\geqslant\cdots\geqslant\lambda_k^*$是$(Y_3'Y_3)^{-1}(Y_1'Y_1+Y_2'Y_2)$的特征根，$c_a$的选取使其水平为$a$．

证 由定理1的（3.12)，得到

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

且$q\in Q_L$，则当$\eta=c\theta$成立时，θ，c，σ^2的最大似然估计量是

$$
\begin{aligned}
&\hat{\theta}=(I_p+\hat{c}'\hat{c})^{-1}(Y_1+cY_2\ ,\\
&c=-(\Gamma_{22}')\Gamma_{12}'\ ,\\
&\sigma^2=\frac{1}{\beta_0}\Big(\prod_{j=p+1}^{p+r}\beta_j+\sum_{j=1}^{k}d_j\Big)\ ,
\end{aligned}
\tag{3.19}
$$

其中β_0由引理 5 决定；d_1，…，d_k是$Y_3'Y_3$之特征根；$\beta_1>\beta_2>\cdots>\beta_{p+r}$是$\begin{pmatrix}Y_1\\Y_2\end{pmatrix}(Y_1'\ \ Y_2')$的特征根，$\gamma_1,\gamma_2$，…，$\gamma_{p+r}$是分别对应于$\beta_1$，…，$\beta_{p+r}$的特征向量，记

$$
\Gamma=(\gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{p+r})=\begin{array}{cc}
\begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix} & \begin{array}{c}p\\ r\end{array}\\
\begin{array}{cc}\ \ p\ \ & \ \ \ r\ \ \end{array} &
\end{array}
\tag{3.20}
$$

（证明与定理 1 的类似，略去）

定理3.5 假定定理3.4的条件成立，关于由（1.2）确定的假设

$$H_0:\quad \exists \underset{r\times p}{c}\quad 使得\quad \eta=c\theta$$

其似然比检验否定域是

$$\frac{\sum\limits_{j=p+1}^{p+r}\beta_j}{\sum\limits_{j=1}^{k}d_j}>c_\alpha \tag{3.21}$$

或等价地是

$$\frac{\sum\limits_{j=p+1}^{k}\beta_j^*}{\sum\limits_{j=1}^{k}d_j}>c_\alpha \tag{3.22}$$

其中β_j，$j=1$，…，$p+r$，d_j，$j=1$，…，k的意义同定理3.4；$\beta_1^*>\beta_2^*>\cdots>\beta_k^*$是$Y_1'Y_1+Y_2'Y_2$的特征根.

（证明类似于定理3.3的证明，略去）

§ 4 . 关 于 模 型 Ⅱ 的 主 要 结 果

定理4.1 假定模型（1.3）成立，$n\geq p+k$，ε是左$\mathcal{O}$（n）不变的，且存在密度函数，Y的密度函数是

$$f(Y;\theta,\Sigma)=|\Sigma|^{n/2}q(\Sigma^{-1/2}[(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2]\Sigma^{-1/2'}),\ \Sigma\in\mathscr{S}(k) \tag{4.1}$$

且q满足引理 1 的条件，则当由（1.4）确定的假设H_0成立时，θ，Γ，Σ的最大似然估计量是

$$
\begin{aligned}
&\hat{\theta}=(I_p-\hat{\Gamma}'\hat{\Gamma})Y_1\ ,\qquad \hat{\Gamma}=(\gamma_{p-r+1},\ \cdots,\ \gamma_p)'\ ,\\
&\hat{\Sigma}=W^{1/2}(W^{1/2}\Sigma^*W^{1/2'})^{-1/2}W(W^{1/2}\Sigma^*W^{1/2'})^{-1/2}W^{1/2'}\\
&W=Y_1'\hat{\Gamma}'\hat{\Gamma}Y_1+Y_2'Y_2
\end{aligned}
\tag{4.2}
$$

其中Σ^*由引理 1 所确定，$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_p$是$Y_1(Y_2'Y_2)^{-1}Y_1'$之特征根，$\gamma_1$，$\gamma_2$，…，$\gamma_p$是分别

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

$$\sup_{\substack{H_0\text{为真}\\ \Sigma\in\ (k)}} L(\theta,\ c,\ \Sigma;\ y) = |\Sigma^*|^{n/2} q(\Sigma^*) |Y_3'Y_3|^{-n/2} \Big(\prod_{j=p+1}^{p+r} (1+\lambda_j) \Big)^{-n/2} .$$

类似地由引理 1 得到

$$\sup_{\substack{\theta,\ \eta\\ \Sigma\in\mathscr{S}(k)}} L(\theta, \eta, \Sigma;\ Y) = |\Sigma^*|^{n/2} q(\Sigma^*) |Y_3'Y_3|^{-n/2}$$

因此似然比检验域是

$$\Lambda = \frac{\sup\limits_{\substack{H_0\text{为真}\\ \in\mathscr{S}(k)}} L(\theta, c, \Sigma;\ Y)}{\sup\limits_{\substack{\theta,\eta\\ \Sigma\in\mathscr{S}(k)}} L(\theta, \eta, \Sigma;\ Y)}$$

$$= \frac{|\Sigma^*|^{n/2} q(\Sigma^*) |Y_3'Y_3|^{-n/2} \Big(\prod\limits_{j=p+1}^{p+r} (1+\lambda_j) \Big)^{-n/2}}{|\Sigma^*|^{n/2} q(\Sigma^*) |Y_3'Y_3|^{-n/2}}$$

$$= \Big(\prod_{j=p+1}^{p+r} (1+\lambda_j) \Big)^{-n/2} < c_0\ , \tag{3.16}$$

这就等价于

$$\prod_{j=p+1}^{p+r} (1+\lambda_j) > c_a\ ,$$

即（3.14）式成立．由于 $\begin{matrix} Y_1 \\ Y_2 \end{matrix}$ $(Y_3'Y_3)^{-1}(Y_1'Y_2')$ 与 $(Y_3'Y_3)^{-1}(Y_1'Y_1+Y_2'Y_2)$ 有相同的非零特征根，所以（3.14）与（3.15）等价．

例3.1 设模型（1.1）成立，Y 服从于矩阵变量 t—分布，即 Y 的密度函数是

$$f(Y;\ M,\ \Sigma) = c\,|\Sigma|^{-n/2} |I_k + \Sigma^{-1}[(Y_1-\theta)'(Y_1-\theta) + (Y_2-\eta)'(Y_2-\eta) + Y_3'Y_3]|^{-\frac{m+n}{2}},\quad n\geqslant k+r+p,\ \ \Sigma\in\mathscr{S}(k), \tag{3.17}$$

其中 c 是正则化常数．

显然，$q(\cdot) = f(\cdot;\ o,\ I)\in Q_M$，但 $q(\cdot)\notin Q_L$.

$$q(\mathrm{diag}(\lambda_1,\ \cdots,\ \lambda_k)) = c\prod_{j=1}^{k} (1+\lambda_j)^{-\frac{m+n}{2}}$$

相应的方程（2.6）的介是 $a_0 = n/m$．因此 θ，c，Σ 的最大似然估计是

$$\theta = (I_p + c'c)^{-1}(Y_1 + c'Y_2),\qquad c = -(T_{22}')^{-1}T_{12}'\ ,$$

$$\Sigma = \frac{m}{n}[(T_{12}'Y_1 + T_{22}'Y_2)'(T_{12}'Y_1 + T_{22}'Y_2)Y_3'Y_3]\ .$$

关于假设（1.2）的似然比检验域是（3.14）或（3.15）．

定理3.4 假定模型（1.1）成立，Y 的密度函数是

$$f(Y;\ M,\ \sigma^2 I) = (\sigma^2)^{-nk/2} q\Big(\frac{1}{\sigma^2}\mathrm{tr}[(Y_1-\theta)'(Y_1-\theta) + (Y_2-\eta)'(Y_2-\eta) + Y_3'Y_3]\Big),\qquad \sigma^2 > 0, \tag{3.18}$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

对应于λ_1，λ_2，…，λ_p的特征向量．

证 由（4.1）得似然函数是

$$L(\theta, \Sigma; Y) = |\Sigma|^{-n/2} q(\Sigma^{-1/2}[(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2]\Sigma^{-1/2'})$$

其中θ满足约束条件$\Gamma\theta = 0$与$\Gamma\Gamma' = I_r$．

运用引理 1 容易验证

$$\sup_{\Sigma \in \mathscr{S}(k)} L(\theta, \Sigma; Y) = |(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2|^{-n/2} |\Sigma^*|^{n/2} q(\Sigma^*) \tag{4.3}$$

其中极大值当Σ_0满足下式时达到，

$$\Sigma^* = \Sigma_0^{-1/2}[(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2]\Sigma_0^{-1/2'} \triangleq \Sigma_0^{-1/2} W \Sigma_0^{-1/2} W \Sigma_0^{-1/2'} .$$

解之得

$$\Sigma_0 = W^{1/2}(W^{1/2}\Sigma^* W^{1/2'})^{-1/2} W (W^{1/2}\Sigma^* W^{1/2'})^{-1/2} W^{1/2'} , \tag{4.4}$$

其中

$$W = (Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2 . \tag{4.5}$$

记

$$\begin{aligned} f(\theta) &= |W| = |(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2| \\ &= |Y_2'Y_2| \, |I_p + (Y_1-\theta)(Y_2'Y_2)^{-1}(Y_1-\theta)'| \end{aligned} \tag{4.6}$$

为在约束条件$\Gamma\theta = 0$，$\Gamma\Gamma' = I_r$之下求$f(\theta)$之极小值，采用Lagrange乘子法，记η是Lagrange乘子，

$$h(\theta, \eta) = \log|I_p + (Y_1-\theta)(Y_2'Y_2)^{-1}(Y_1-\theta)'| + 2tr\eta\Gamma\theta,$$

$$\frac{\partial h(\theta,\eta)}{\partial \eta} = 2\theta'\Gamma' = 0 ,$$

$$\frac{\partial h(\theta,\eta)}{\partial \theta} = -2(I_p + (Y_1-\theta)(Y_2'Y_2)^{-1}(Y_1-\theta)')^{-1}(Y_1-\theta)(Y_2'Y_2)^{-1} + 2\Gamma'\eta' = 0 ,$$

由此得

$$\begin{aligned} &\Gamma\theta = 0 , \\ &\eta\Gamma = (Y_2'Y_2)^{-1}(Y_1-\theta)'(I_p + (Y_1-\theta)(Y_2'Y_2)^{-1}(Y_1-\theta)')^{-1} \\ &\quad = (Y_2'Y_2)^{-1}(I_k + (Y_1-\theta)'(Y_1-\theta)[(Y_1-\theta)'(Y_1-\theta) + \\ &\qquad Y_2'Y_2]^{-1})(Y_1-\theta)' . \end{aligned} \tag{4.7}$$

利用$\Gamma\Gamma' = I_r$得

$$\eta = (Y_2'Y_2)^{-1}(I_k + (Y_1-\theta)'(Y_1-\theta)[Y_2'Y_2 + (Y_1-\theta)'(Y_1-\theta)]^{-1})Y_1'\Gamma' .$$

代入（4.7）得

$$Y_1'\Gamma'\Gamma = (Y_1-\theta)' .$$

所以

$$\theta_0 = (I_p - \Gamma'\Gamma)Y_1 , \tag{4.8}$$

$$\inf_{\substack{\Gamma\theta=0 \\ \Gamma\Gamma'=I_r}} f(\theta) = f(\theta_0) = |Y_1'\Gamma'\Gamma Y_1 + Y_2'Y_2| . \tag{4.9}$$

再由引理 7 得到

$$\inf_{\substack{\Gamma \\ \Gamma\Gamma'=I_r}} |Y_1'\Gamma'\Gamma Y_1 + Y_2'Y_2| = |Y_1'\Gamma'\Gamma Y_1 + Y_2'Y_2| = |Y_2'Y_2| \prod_{j=p-r+1}^{p} (1+\lambda_j) , \tag{4.10}$$

其中$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p$是$Y_1(Y_2'Y_2)^{-1}Y_1'$的特征根，$\gamma_1$ γ_2，…，γ_p是分别与λ_1，…，λ_p相对应的

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

特征向量，

$$\Gamma=(\gamma_{p-r+1}, \cdots, \gamma_p)' \quad . \tag{4.11}$$

综合 (4.4), (4.5), (4.8), (4.10), (4.11) 即得定理之全部结论.

综合 (4.3), (4.9), (4.10), 得到

$$\sup_{\substack{\Sigma \quad (k) \\ \Gamma\theta=0 \\ \Gamma\Gamma'=I_r}} L(\theta, \Sigma; Y)=|\Sigma^*|^{n/2}q(\Sigma^*)|Y_2'Y_2|^{n/2}\Big(\prod_{j=p-r+1}^{p}(1+\lambda_j)\Big)^{-n/2} . \tag{4.12}$$

定理4.2 假定模型 (1.3) 成立，$n\geqslant p+k$，且Y的密度函数是

$$f(Y;\theta,\Sigma)=|\Sigma|^{n/2}q(\Sigma^{-1}[(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2]), \qquad \Sigma\in\mathscr{S}(k) \tag{4.13}$$

和$q(\cdot)\in Q_M$，则当由 (1.4) 确定的假设H_0成立时，θ, Γ, Σ的最大似然估计是

$$\begin{aligned}\theta&=(I_p-\Gamma'\Gamma)Y_1 \quad ,\\ \Gamma&=(\gamma_{p-r+1}\ ,\ \gamma_{p-r+2}\ ,\ \cdots,\ \gamma_p)' \quad ,\\ \Sigma&=\frac{1}{a_0}(Y_1'\Gamma'\Gamma Y_1+Y_2'Y_2) \ ,\end{aligned} \tag{4.14}$$

其中γ_j, $j=p-r+1, \cdots, p$的意义同定理4.1；a_0由引理 3 决定；如果进一步假定$q(\cdot)\in Q_L$，则a_0用相应的β_0/k代替，β_0由引理 5 决定.

定理4.3 假定定理 1 的条件成立，关于由 (1.4) 确定的假设

$$H_0: \exists \underset{r\times p}{\Gamma}, \ \Gamma\Gamma'=I_r, \text{ 使得 } \Gamma\theta=0 \qquad (r\leqslant p),$$

其似然比检验否定域是

$$\prod_{j=p-r+1}^{p}(1+\lambda_j)>c_\alpha \quad , \tag{4.15}$$

或等价地是

$$\prod_{j=p-r+1}^{k}(1+\lambda_j^*)>c_\alpha \ , \tag{4.16}$$

其中$\lambda_1^*\geqslant\lambda_2^*\geqslant\cdots\geqslant\lambda_k^*$是 $(Y_2'Y_2)^{-1}(Y_1'Y_1)$ 的特征根.

定理4.4 假定模型 (1.3) 成立，$n\geqslant p+k$，Y的密度函数是

$$f(Y;\theta,\sigma^2)=(\sigma^2)^{-nk/2}q(\frac{1}{\sigma^2}\mathrm{tr}[(Y_1-\theta)'(Y_1-\theta)+Y_2'Y_2]), \quad \sigma^2>0, \ q\in Q_L, \tag{4.17}$$

则当由 (1.4) 确定的假设H_0成立时，θ, Γ, σ^2的最大似然估计是

$$\begin{aligned}\theta&=(I_p-\Gamma'\Gamma)Y_1 \ ,\\ \Gamma&=(\gamma_{p-r+1}, \ \cdots, \ \gamma_p)' \quad ,\\ \sigma^2&=\frac{1}{\beta_0}\Big(\sum_{j=p-r+1}^{p}\beta_j+\sum_{j=1}^{k}d_j\Big) \ ,\end{aligned} \tag{4.18}$$

其中：β_0由引理 5 决定；$d_1, \cdots, d_k$ 是$Y_2'Y_2$的特征根；$\beta_1\geqslant\beta_2\geqslant\cdots\geqslant\beta_p$是$Y_1Y_1'$的特征根，$\gamma_1, \cdots, \gamma_p$是分别与$\beta_1, \cdots, \beta_p$相对应的特征向量.

定理4.5 假定定理4.4的条件成立，关于由 (1.4) 确定的假设

$$H_0: \exists \underset{r\times p}{\Gamma}, \ \Gamma\Gamma'=I_r \text{ 使得 } \Gamma\theta=0 \qquad (r\leqslant p),$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

其似然比检验否定域是

$$\frac{\sum_{j=p-r+1}^{p}\beta_j}{\sum_{j=1}^{k}d_j}>c_{1\alpha} \tag{4.19}$$

或等价地是

$$\frac{\sum_{j=p-r+1}^{k}\beta_j^*}{\sum_{j=1}^{k}d_j}>c_{1\alpha} \tag{4.20}$$

其中$\beta_1^*\geqslant\beta_2^*\geqslant\cdots\geqslant\beta_k^*$是$Y_1'Y_1$的特征根.

参 考 文 献

〔1〕Zhang Yaoting and Bian Gaorui, A hypothesis testing problem in the linear models, 数学年刊 Vol.3, No.2, (1982)

〔2〕张尧庭，卞国瑞，线性模型的泛线性假设检验，数学物理学报，Vol.1, No.3—4, (1981)

〔3〕Bian Guorui and Zhang Yaoting, Invariant tests of existness of the linear relationship with left $\mathcal{O}(n)$—invariant errors, (1984), (已投数学年刊)

〔4〕张尧庭，方开泰，陈汉峰，矩阵椭球等高分布族，(1984)，数学物理学报，5(1985)，3,341—353.

Maximum Likelihood Estimators and Likelihood Ratio Tests of Linear Functional Relationship with Left $\mathcal{O}(n)$—Invariant

Bian Guorui (Fudan Vniversity)

Zhang Yaoting (Wuhan Vniversity)

Abslract

In this paper, the likelihood ratio tesis of existness of linear functional relationship among row vectors of the mean matrix and the maximum likelihood estimators of those corresponding parameters are discussed about two kinds of canonical forms of multivariate linear models with the left $\mathcal{O}(n)$-invariant error errors. The distributions of those esitmators are independent of the forms of the error's distributions, so they are robust.

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.