Chin. Ann. of Math.
3 (2) 1982

A HYPOTHESIS TESTING PROBLEM IN THE LINEAR MODEL

ZHANG YAOTING BIAN GUORUI
(*Wuhan University*) (*Fudan University*)

§ 1. Introduction

In this paper, we discuss the following problem.

Let $y_{(1)}, y_{(2)}, \cdots, y_{(n)}$ be independently and normally distributed with means and covariance matrices given by the linear model, and

$$\underset{n\times k}{Y}=\begin{pmatrix} y'_{(1)} \\ \vdots \\ y'_{(n)} \end{pmatrix}.$$

We can rewrite the above assumption as follows

$$E(Y)=\underset{n\times p}{A}\ \underset{p\times k}{\theta}+\underset{n\times q}{B}\ \underset{q\times k}{\eta}=(AB)\begin{pmatrix}\theta\\ \eta\end{pmatrix},$$

$$E\begin{pmatrix} y_{(1)}-Ey_{(1)} \\ \vdots \\ y_{(n)}-Ey_{(n)} \end{pmatrix}((y_{(1)}-Ey_{(1)})'\cdots(y_{(n)}-Ey_{(n)})')=I_n\otimes V. \tag{1}$$

On the basis of observations $y_{(1)}, \cdots, y_{(n)}$, we are interested in testing the following hypothesis, H_0: there is a matrix $\underset{q\times p}{C}$ such that $\eta=C\theta$.

§ 2. Canonical Form

Lemma 1. *Suppose* rk $\underset{n\times p}{A}=r$, rk $\underset{n\times q}{B}=s$ *and* $\mathscr{L}(A)\cap\mathscr{L}(B)=0$, *where* $\mathscr{L}(A)$ *denotes the subspace generated by the column vectors of* A. *Then there are matrices* Γ *and* G *such that*

$$\Gamma(AB)G=\begin{pmatrix} I_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{matrix} r \\ p-r \\ s \\ n-(p+s) \end{matrix}$$
$$\qquad\qquad\qquad r \quad p-r \quad s \quad q-s$$

where Γ *is an orthogonal matrix and* G *is a nonsingular matrix with the following form*

Manuscript received Dec. 23, 1980. revisecl Mar. 21, 1981.

$$G=\begin{pmatrix} G_{11} & 0 \\ G_{12} & G_{22} \end{pmatrix}\begin{matrix} p \\ q. \end{matrix}$$
$$\quad\; p \quad\; q$$

The proof is simple, so it is omitted.

Lemma 2. *For any matrix* $\underset{n\times m}{C}$

and
$$(I_n+CC')^{-1}=I_n-C(I_m+C'C)^{-1}C',$$
$$C(I_m+C'C)^{-1}=(I_n+CC')^{-1}C.$$

Lemma 3. *Suppose* $\lambda_1\geqslant\cdots\geqslant\lambda_m$ *are the characteristic roots of a symmetric matrix* $\underset{m\times m}{A}$, *and* $x_1, \cdots, x_k$ *are mutually orthogonal vectors.* Then

$$\inf_{x_1,\cdots,x_k}\sum_{i=1}^{k}\frac{x_i'Ax_i}{x_i'x_i}=\sum_{j=m-k+1}^{m}\lambda_j.$$

These results can be found in [1] or [2].

Lemma 4. *Suppose* $\underset{(p+q)\times(p+q)}{A}$ *is a non-negative definite matrix, and* $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{p+q}$ *are the characteristic roots of* A. *We arrange the characteristic vectors* $t_1, \cdots, t_{p+q}$, *corresponding the characteristic roots* $\lambda_1,\cdots, \lambda_{p+q}$ *respectively, in a matrix* T.

$$T=(t_1, t_2, \cdots, t_{p+q})=\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\begin{matrix} p \\ q \end{matrix} \quad \text{and} \ |T_{22}|\neq 0.$$
$$\qquad\qquad\qquad p \quad\; q$$

Then

$$\inf_{B\in\mathscr{B}}\operatorname{tr}\,(I_q+BB')^{-1/2}(-B\,I_q)A\begin{pmatrix}-B'\\ I_q\end{pmatrix}(I_q+BB')^{-1/2}=\sum_{j=p+1}^{p+q}\lambda_j \qquad (2)$$

where $\mathscr{B}=\{B\,|\,B$ *is a* $p\times q$ *matrix*$\}$, *and the minimum is attained when* $B=-(T_{22}')^{-1}T_{12}'$.

The result is obtained by Lemma 3.

To simplify the following derivations, by Lemma 1, we can reduce the linear model and the hypothesis to the canonical form

$$E(\underset{n\times k}{Y})=E\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix}=\begin{pmatrix}\theta\\ \eta\\ 0\end{pmatrix}\begin{matrix} p\\ q\\ n-(p+q),\end{matrix} \qquad (3)$$

where the row vectors of Y are independently normally distributed with the same covariance matrix V. H_0: $\exists\, \underset{q\times p}{C}$ such that $\eta=C\theta$.

Thus, we shall study the canonical form only.

§ 3. Main Theorems

(3A) $V=\sigma^2 I_k$.

When H_0 is true, the likelihood function

$$L(Y;\theta, C, \sigma^2)=(2\pi)^{-\frac{nk}{2}}(\sigma^2)^{-\frac{nk}{2}}\exp\left\{-\frac{1}{2\sigma^2}\operatorname{tr}[(y_1-\theta)'(y_1-\theta)\right.$$

$$+(y_2-C\theta)'(y_2-C\theta)+y_3'y_3]\Big\}.$$

Then
$$0=\frac{\partial \ln L}{\partial \theta}=-\frac{1}{2\sigma^2}(2(y_1-\theta)+2C'(y_2-C\theta)),$$

$$0=\frac{\partial \ln L}{\partial \sigma^2}=-\frac{nk}{2}\cdot\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}\mathrm{tr}[(y_1-\theta)'(y_1-\theta)$$
$$+(y_2-C\theta)'(y_2-C\theta)+y_3'y_3].$$

In other words
$$(I_p+C'C)\theta=y_1+C'y_2,$$
$$\sigma^2=\frac{1}{nk}\mathrm{tr}[(y_1-\theta)'(y_1-\theta)+(y_2-C\theta)'(y_2-C\theta)+y_3'y_3],$$

and C is chosen such that it makes $\mathrm{tr}[(y_1-\theta)'(y_1-\theta)+(y_2-C\theta)'(y_2-C\theta)+y_3'y_3]$ as small as possible. Using Lemma 4 and $P(|T_{22}|=0)=0$, we obtain the maximum likelihood estimators

$$\hat\theta=(I_p+\hat C'\hat C)^{-1}(I_p\hat C)\begin{pmatrix}y_1\\ y_2\end{pmatrix}=(I_p+\hat C'\hat C)^{-1}(y_1+\hat C'y_2),$$
$$\hat C=-(T_{22}')^{-1}T_{12}', \tag{4}$$
$$\hat\sigma^2=\frac{1}{nk}\Big(\sum_{p+1}^{p+q}\lambda_j^*+\sum_1^k d_j\Big),$$

where $d_1, \cdots, d_k$ are the eigenvalues of $y_3'y_3$, $\lambda_1^*\geqslant\lambda_2^*\geqslant\cdots\geqslant\lambda_{p+q}^*$ are the eigenvalues of $\begin{pmatrix}y_1y_1' & y_1y_2'\\ y_2y_1' & y_2y_2'\end{pmatrix}$ in descending order, and their corresponding eigenvectors as columns (from left to right) are arranged in the orthogonal matrix

$$\begin{pmatrix}T_{11} & T_{12}\\ T_{21} & T_{22}\end{pmatrix}\begin{matrix}p\\ q\end{matrix}.$$
$$\begin{matrix}p & \quad q\end{matrix}$$

Because the non-zero eigenvalues of $\begin{pmatrix}y_1y_1' & y_1y_2'\\ y_2y_1' & y_2y_2'\end{pmatrix}$ are the same as those of the matrix $y_1'y_1+y_2'y_2$, we can rewrite $\hat\sigma^2=\frac{1}{nk}\Big(\sum_{p+1}^{k}\lambda_j+\sum_1^k d_j\Big)$, where $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{p+1}\geqslant\cdots\geqslant\lambda_k$ are eigenvalues of $y_1'y_1+y_2'y_2$ in descending order.

Thus, it is easy to verify the following thorem.

Theorem 1. *The generalized likelihood ratio test of* H_0 *is*

$$\Lambda^*=\left(\frac{\sum_1^k d_j}{\sum_{p+1}^k\lambda_j+\sum_1^k d_j}\right)^{\frac{kn}{2}}, \tag{5}$$

or equivalently, the statistic of the test is

$$\Lambda=\sum_{p+1}^k\lambda_j\Big/\sum_1^k d_j, \tag{6}$$

where $\sum_{1}^{k} d_j$ *and* $\sum_{p+1}^{k} \lambda_j$ *are independently distributed.*

The distribution of $\sum_{1}^{k} d_j$ is $\frac{1}{\sigma^2}\chi^2(k(n-p-q))$, and the joint density of $\lambda_1, \cdots, \lambda_k$ is known in [4] and [5].

If $V=\sigma^2\Omega$ and Ω is known, it can be transformed to the case $V=\sigma^2 I$. Then the problem may be treated in a similar way.

(3B) The general case.

When H_0 is true, the least square estimator of θ is

$$\hat{\theta}_L=(I_p+C'C)^{-1}(y_1+C'y_2).$$

The residual matrix is

$$\tilde{Y}'\tilde{Y}=(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1)+y_3'y_3.$$

Since $$\min_C \operatorname{tr}_T \tilde{Y}'\tilde{Y}=\operatorname{tr} y_3'y_3+\min_C \operatorname{tr}(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1),$$

if C is chosen such that $\operatorname{tr}(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1)$ reaches the minimum, we obtain the generalized least square estimator $\hat{C}$ of C

$$\hat{C}=-(T'_{22})^{-1}T'_{12},$$

where T_{ij} is the same as in (4).

Solving the likelihood equation, we obtain

$$\hat{V}=\frac{1}{n}[(y_1-\theta)'(y_1-\theta)+(y_2-C\theta)'(y_2-C\theta)+y_3'y_3].$$

Thus, the estimators of θ, C, V are

$$\begin{aligned}
&\hat{\theta}=(I_p+\hat{C}'\hat{C})^{-1}(y_1+\hat{C}'y_2),\\
&\hat{C}=-(T'_{22})^{-1}T'_{12} \qquad\qquad (7)\\
&\hat{V}=\frac{1}{n}(y_2'T_{22}T'_{22}y_2+y_3'y_3).
\end{aligned}$$

Comparing (4) with (7), it is trivial that the least square estimators of θ and C are the same as the maximum likekihood estimators, when H_0 is true.

Let $$Z=\left[I_{p+q}-\begin{pmatrix} I_p \\ C\end{pmatrix}(I_p+C'C)^{-1}(I_p\ C')\right]\begin{pmatrix} y_1 \\ y_2\end{pmatrix}.$$

Then the hypothesis H_{00}: $\eta=C\theta$, where C is a given constant matrix, is equivalent to H_{00}^*: $EZ=0$ for random matrix $\begin{pmatrix} Z \\ y_3\end{pmatrix}$.

It is easy to prove that the hypothesis H_{00}^* is invariant under the group of full rank linear transforms

$$\begin{pmatrix} Z'Z \\ y_3'y_3\end{pmatrix}\xrightarrow{|D|\neq 0}\begin{pmatrix} D'Z'ZD \\ D'y_3'y_3D\end{pmatrix}$$

and the partitioned orthogonal transform group

$$\begin{pmatrix} Z \\ y_3\end{pmatrix}\longrightarrow\Gamma\begin{pmatrix} Z \\ y_3\end{pmatrix}, \qquad \Gamma=\begin{pmatrix}\Gamma_{p+q} & 0\\ 0 & \Gamma_{n-p-q}\end{pmatrix},$$

the Γ_{p+q} and Γ_{n-p-q} are orthogonal matrices with order $p+q$ and $n-p-q$ respectively.

The maximal invariant statistic under the above two groups are the eigenvalues of $|Z'Z-\lambda y_3'y_3|=0$.

If the Lawley's trace test [6] is used, then the critical region is

$$\operatorname{tr}(Z'Z(y_3'y_3)^{-1})\geqslant a. \tag{8}$$

It is known that the Lawley's trace test is invariant under the above two groups.

Notice that

$$\operatorname{tr}\,(Z'Z(y_3'y_3)^{-1})=\operatorname{tr}[(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1)(y_3'y_3)^{-1}],$$

we can rewrite (8) as

$$\operatorname{tr}[(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1)(y_3'y_3)^{-1}\geqslant a. \tag{9}$$

Next, we consider the problem about testing H_0: $\exists\, C$ such that $\eta=C\theta$.

The critical region is

$$\min_C \operatorname{tr}[(y_2-Cy_1)'(I_q+CC')^{-1}(y_2-Cy_1)(y_3'y_3)]\geqslant a_1, \tag{10}$$

because it can be tested about H_{00}: $\eta=C\theta$ for every C. By Lemma 4,

$$\min_C \operatorname{tr}[(y_2-Cy_1)'(I_q+CC')(y_2-Cy_1)(y_3'y_3)]=\sum_{p+1}^{p+q}\beta_{*j},$$

where $\beta_{*1}\geqslant\beta_{*2}\geqslant\cdots\geqslant\beta_{*p+q}$ are the eigenvalues of matrix $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(y_3'y_3)^{-1}(y_1'\ y_2')$, since the nonzero eigenvalues of matrix $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(y_3'y_3)^{-1}(y_1'\ y_2')$ are the same as those of the matrix $(y_3'y_3)^{-1}(y_1'\ y_2')\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, i. e., $(y_3'y_3)^{-1}(y_1'y_1+y_2'y_2)$.

If $\beta_1\geqslant\beta_2\geqslant\cdots\geqslant\beta_k$ are the eigenvalues of $|y_1'y_1+y_2'y_2-\lambda y_3'y_3|=0$ arranged in descending order, we can rewrite (10) as

$$\sum_{p+1}^{p+q}\beta_j\geqslant a_1. \tag{11}$$

The joint distribution can be found in [4, 5]. we shall discuss the limiting distributions of Λ and $\sum_{p+1}^{k}\beta_j$ in other papers.

References

[1] Rao, C. R., *Linear Statistcal Inference and its Applications*, John Wiley and Sons, New York, (1973).

[2] Chan, N. N., Estimating linear functional relationships, *International Conference in Statistics in Tokyo*, (1979).

[3] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York, (1958).

[4] James, A. T., Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.*, **35** (1964), 475—501.

[5] Constantine, A. G., Noncentral distribution problems in multivariate analysis, *Ann. Math. Statist.*, **34** (1963), 1270—1285.

[6] Lawley, D. N., A generalization of Fisher's Z test, *Biomtrika*, **30** (1938), 180—187.

线性模型中的一个假设检验问题

张尧庭　　卞国瑞
(武汉大学)　(复旦大学)

摘　　要

本文讨论了多元线性模型中的一个假设检验问题. 假定

$$\begin{cases} \underset{n\times k}{E(Y)}=A\theta+B\eta, \\ Y \text{ 的各行独立、正态、同协差阵 } V. \end{cases}$$

现在要检验假设 H_0: 存在矩阵 C 使 $\theta=C\eta$ 是否成立. 首先可将问题化为法式的形式, 对法式分两种情况进行讨论:

(一) $V=\sigma^2 I$, σ^2 未知. 此时可求出 θ, C, σ^2 的最大似然估计(当 H_0 成立时)是

$$\begin{cases} \hat{\theta}=(I_p+\hat{C}'\hat{C})^{-1}(y_1+\hat{C}'y_2), \\ \hat{C}=-(T'_{22})^{-1}T'_{12}, \\ \hat{\sigma}^2=\dfrac{1}{nk}\Big(\sum\limits_{j=p+1}^{p+q}\lambda_j^*+\sum\limits_{j=1}^{k}d_j\Big), \end{cases}$$

其中 y_1, y_2 是法式

$$E\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} \theta \\ \eta \\ 0 \end{pmatrix}\begin{matrix} p \\ q \\ n-(p+q) \end{matrix}$$

中的资料阵 y_1, y_2, d_1, …, d_k 是 y'_3y_3 的全部特征根, $\lambda_1^*\geqslant\cdots\geqslant\lambda_{p+q}^*$ 是 $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(y'_1\ y'_2)$ 的全部特征根, 相应特征向量依 λ_i^* 的大小顺序从左到右排成矩阵 T, T 的分块子阵是 T_{ij}, 即

$$T=\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\begin{matrix} p \\ q. \end{matrix}$$
$$\quad p \qquad q$$

对 H_0 的广义似然比检验是

$$\Lambda=\sum_{j=p+1}^{k}\lambda_j\Big/\sum_{j=1}^{k}d_j,$$

$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_k$ 是 $y'_1y_1+y'_2y_2$ 的全部特征根.

(二)一般情形 V 未知. 此时 θ, C 的估计量同前, 可求出

$$\hat{V}=\frac{1}{n}(y'_2T_{22}T'_{22}y_2+y'_3y_3).$$

H_0 相应的 Lawley 不变检验是

$$\sum_{j=p+1}^{k}\beta_j\geqslant a_1,$$

其中 $\beta_1\geqslant\beta_2\geqslant\cdots\geqslant\beta_k$ 是 $y'_1y_1+y'_2y_2$ 的相对于 y'_3y_3 的全部特征根.

有关 Λ 的以及 $\sum\limits_{j=p+1}^{k}\beta_j$ 的极限分布将在另外的文章中讨论.