1996年12月
第26卷第6期

西北大学学报(自然科学版)
Journal of Northwest University (Natural Science Edition)

Dec. 1996
Vol. 26 No. 6

② 467-470

多元线性模型系数的广义根方改进估计类*

O212.4

夏正茂　梁家荣

(西北大学数学系，710069，西安；第一作者28岁，男，工程师)

A

摘　要　从矩阵变换理论出发，对多元线性模型的系数提出了广义根方估计。证明了它优于系数的LS估计，根方改进估计，且是β的线性可容性估计等性质。

关键词　多元线性模型；广义根方改进估计；可容许估计；均方误差

分类号　O213.9

广义根方估计

对于多元线性模型

$$\begin{cases} Y_{n\times q}=X_{n\times p}B_{p\times q}+\varepsilon_{n\times q} \\ r(X)=P,\varepsilon\text{的行向量互不相关,} \\ \text{均值为零,有共同协方差阵 } V>0 \end{cases} \tag{1}$$

由文献1，上述模型(1)可化为：

$$\begin{cases} \mathrm{vec}(Y)=(I\otimes X)\mathrm{vec}(B)+\mathrm{vec}(\varepsilon) \\ cov(\mathrm{vec}(\varepsilon))=V\otimes I_n, E(\mathrm{vec}(\varepsilon))=0 \end{cases} \tag{2}$$

令$\beta=\mathrm{vec}(B)$，易知，β的LS估计为：

$$\hat{\beta}=\mathrm{vec}((X'X)^{-1}X'Y);$$

$$\mathrm{cov}(\hat{\beta})=\mathrm{cov}(\mathrm{vec}(\hat{B}))=V\otimes(X'X)^{-1}$$

(1)的典测形式为：

$$\begin{cases} Y=ZA+\varepsilon \\ \varepsilon\text{的行向量互不相关,均值为零,} \\ \text{有共同的协方差阵 } V>0 \end{cases} \tag{3}$$

其中，$Z=XQ$，$A=Q'B$，Q为正交矩阵，使$Q'X'XQ=\Lambda\triangleq\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_p)$，$\lambda_i$是设计矩阵$(X'X)$的特征根。与(1)同法(3)可化为：

$$\begin{cases} \mathrm{vec}(Y)=(I_q\otimes Z)\mathrm{vec}(A)+\mathrm{vec}(\varepsilon) \\ \mathrm{cov}(\mathrm{vec}(\varepsilon))=V\otimes I_n, \\ \mathrm{E}(\mathrm{vec}(\varepsilon))=0 \end{cases} \tag{4}$$

令$\alpha=\mathrm{vec}(A)$，则α的LS估计为：

$$\hat{\alpha}=\mathrm{vec}(\hat{A})=(I_q\otimes\Lambda^{-1}Z')\mathrm{vec}(Y)$$

$$\mathrm{cov}(\hat{\alpha})=V\otimes\Lambda^{-1}$$

由于$\hat{\beta}$是β的无偏估计，显然：

$$\mathrm{MSE}(\hat{\beta})=\mathrm{tr}(\mathrm{cov}(\hat{\beta}))+\|E\hat{\beta}-\beta\|^2=\mathrm{tr}V\otimes\mathrm{tr}(X'X)^{-1}=\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{-1} \tag{5}$$

当设计矩阵是病态时，$\mathrm{MSE}(\hat{\beta})$较大，此时，$\hat{\beta}$不是$\beta$的较好估计。近30多年来，国内外学者提出了众多估计，Zellner在文献2中首先提出了根方改进估计，文献3做了介绍，并讨论了其优良性，但该估计不能对所有特征向量做有效改进。本文首次在多元线性模型中，提出了广义根方估计类，克服了这一缺点，并讨论了其优良性及可容许性。

* 收稿日期：1995-10-24

1 广义根方估计类

定义 1 对模型(4),称 $\gamma=\{\hat{\alpha}(k_1,k_2,\cdots,k_p)=(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)-1}Z')\mathrm{vec}(Y)\mid k_i\in R\}$ (6)

为其系数 α 的广义根方估计类,其中 $\Lambda^{(k_1,k_2,\cdots,k_p)-1}\triangleq\begin{bmatrix}\lambda_1^{k_1-1} & & \\ & \lambda_2^{k_2-1} & \\ & & \ddots,\lambda_p^{k_p-1}\end{bmatrix}$

显然,当 $k_1=k_2=\cdots=k_p$ 时,γ 为 α 的根方估计类;γ 也是 α 的一个有偏估计类。为了讨论 γ 的均方误差,我们先引进如下的

引理 1[4] 设 X 是 $n\times1$ 随机向量,A 是 $n\times n$ 对称阵,若 $E(X)=\theta,D(X)=\sum=[\sigma_{ij}]$,则 $E(X'AX)=\mathrm{tr}(A\sum)+\theta'A\theta$。

引理 2[3] 对于线性模型(1),若 β^* 是 β 的任一估计,则 $\mathrm{MSE}(\beta^*)=\mathrm{trcov}(\beta^*)+\|E\beta^*-\beta\|^2$

对 γ 中任一估计 $\hat{\alpha}(k_1,k_2,\cdots,k_p)$。

$$\begin{aligned}\mathrm{cov}(\hat{\alpha}(k_1,k_2,\cdots,k_p))&=\mathrm{cov}((I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)-1}Z')\mathrm{vec}(Y))\\&=\mathrm{cov}((I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})(I_q\otimes\Lambda^{-1}Z')\mathrm{vec}(Y))\\&=\mathrm{cov}((I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\hat{\alpha})\\&=(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\mathrm{cov}(\hat{\alpha})(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\\&=(I_q\otimes\Lambda^{2(k_1,k_2,\cdots,k_p)})(V\otimes\Lambda^{-1})(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\\&=V\otimes\Lambda^{2(k_1,k_2,\cdots,k_p)-1}\end{aligned}\tag{7}$$

这里记 $\Lambda^{n(k_1,k_2,\cdots,k_p)+m}=\begin{bmatrix}\lambda_1^{nk_1+m} & & \\ & \lambda_2^{nk_2+m} & \\ & & \ddots,\lambda_p^{nk_p+m}\end{bmatrix}$($n,m$ 是实数)令 $A\triangleq Q'B=(\sigma_{ij})$由(7)式得

引理 3 γ 中任一估计 $\hat{\alpha}(k_1,k_2,\cdots,k_p)$的均方误差为:$\mathrm{MSE}(\hat{\alpha}(k_1,k_2,\cdots,k_p))=\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{2k_i-1}+\sum_{j=1}^{p}\sum_{i=1}^{q}\sigma_{ij}^2(\lambda_i^{k_i}-1)^2$。

证明 $\hat{\alpha}(k_1,k_2,\cdots,k_p)=(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})(I_q\otimes\Lambda^{-1}Z')\mathrm{vec}(Y)=(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\hat{\alpha}$

由引理 1 和(7)可得:

$$\begin{aligned}\mathrm{MSE}(\hat{\alpha}(k_1,k_2,\cdots,k_p))&=\mathrm{tr}(\mathrm{cov}(\hat{\alpha}(k_1,k_2,\cdots,k_p)))+\|E(\hat{\alpha}(k_1,k_2,\cdots,k_p)-\alpha\|^2\\&=\mathrm{tr}(V\otimes\Lambda^{2(k_1,k_2,\cdots,k_p)-1})+E\|(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}-I_q\odot I_p)\alpha\|^2\\&=\mathrm{tr}V\cdot\mathrm{tr}(\Lambda^{2(k_1,k_2,\cdots,k_p)-1})+E\|(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}-I_q\otimes I_p)\mathrm{vec}(Q'B)\|^2\\&=(\mathrm{tr}V)\sum_{i=1}^{p}\lambda_i^{2k_i-1}+\sum_{i=1}^{p}\sum_{j=1}^{q}\sigma_{ij}^2(\lambda_i^{k_i}-1)^2\end{aligned}\tag{8}$$

2 广义根方估计的性质

定理 1 在均方误差意义下,对任一根方估计 $\hat{\alpha}(k)$总在广义估计 γ 中,存在一个 $\hat{\alpha}(k_1,k_2,\cdots,k_p)$使 $\mathrm{MSE}(\hat{\alpha}(k_1,k_2,\cdots,k_p))<\mathrm{MSE}(\hat{\alpha}(k))$

证明 由引理 3 易知

$$\mathrm{MSE}(\hat{\alpha}(k))=\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{2k-1}+\sum_{i=1}^{p}\sum_{j=1}^{q}(\lambda_i^{k}-1)^2\sigma_{ij}^2\tag{9}$$

设 $\mathrm{MSE}(\hat{\alpha}(k))$在 $k=k_l$ 时取得最小值,显然 $o<k_l<\frac{1}{2}$[3] 对 λ_i作如下分类,并取相应的 k_i,满足

$$\begin{cases} \lambda_i<1\text{ 时}, k_i=k_i \\ \lambda_i=1\text{ 时}, k_i=0 \\ 1<\lambda_i<2^{\frac{1}{k_i}}\text{时}, \log_{\lambda_i}(2-\lambda_i^{k_i})<k_i<0 \\ \lambda_i\geq 2^{\frac{1}{k_i}}\text{时}, k_i<0 \end{cases} \tag{10}$$

显然，只须对 $\lambda_i>1$ 情况讨论，此时对 k_i 有

$$\lambda_i^{2k_i}-\lambda_i^{2k_i}<0,(\lambda_i^{k_i}-1)^2-(\lambda_i^{2k_i}-1)^2<0 \tag{11}$$

从而 $\mathrm{MSE}(\hat{a}(k_1,k_2,\cdots,k_p))-\mathrm{MSE}(\hat{a}(k))$

$$=\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{2k_i-1}-\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{2k_i-1}+\sum_{i=1}^{p}\sum_{j=1}^{q}((\lambda_i^{k_i}-1)^2\sigma_{ij}^2-(\lambda^{k_i}-1)^2\sigma_{ij}^2)$$

$$=\mathrm{tr}V\sum_{\lambda_i>1}(\lambda_i^{2k_i}-\lambda_i^{2k_i-1})+\sum_{\lambda_i>1}\sum_{j=1}^{q}[(\lambda_i^{k_i}-1)^2-(\lambda_i^{k_i}-1)^2]\sigma_{ij}^2<0$$

结论成立。

又 $\frac{\partial}{\partial k_i}(\mathrm{MSE}(\hat{a}(k_1,k_2,\cdots,k_p))=\frac{\partial}{\partial k_i}(\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{2k_i-1})+\frac{\partial}{\partial k_i}(\sum_{i=1}^{p}\sum_{j=1}^{q}\sigma_{ij}^2(\lambda_i^{k_i}-1)^2)$

$$=2(\mathrm{tr}V)\lambda_i^{2k_i-1}\ln\lambda_i+2\sum_{j=1}^{q}\sigma_{ij}^2(\lambda_i^{k_i}-1)\lambda_i^{k_i}\ln\lambda_i$$

令 $\frac{\partial}{\partial k_i}(\mathrm{MSE}(\hat{a}(k_1,k_2,\cdots,k_p)))=0$。解得

$$k_i=\log_{\lambda_i}\frac{\sum_{j=1}^{q}\sigma_{ij}^2\lambda_i}{\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2}$$

由于 a 的 LS 估计，$\hat{a}$ 满足 $\mathrm{MSE}(\hat{a})=\mathrm{tr}V\sum_{i=1}^{p}\lambda_i^{-1}$，

将 k_i 代入 $\mathrm{MSE}(k_1,k_2,\cdots,k_p)$ 得

$$\mathrm{MSE}(k_1,k_2,\cdots,k_p)=\sum_{i=1}^{p}\frac{\mathrm{tr}V}{\lambda_i}\frac{\lambda_i^2(\sum_{j=1}^{q}\sigma_{ij}^2)^2}{(\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2)^2}+\sum_{i=1}^{p}\sum_{j=1}^{q}\sigma_{ij}^2\frac{(\mathrm{tr}V)^2}{(\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2)^2}$$

$$=\sum_{i=1}^{p}\mathrm{tr}V\frac{\lambda_i(\sum_{j=1}^{q}\sigma_{ij}^2)^2}{(\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2)^2}+\sum_{i=1}^{p}\frac{(\mathrm{tr}V)^2\sum_{j=1}^{q}\sigma_{ij}^2}{(\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2)^2}$$

$$=\sum_{i=1}^{p}\mathrm{tr}V\frac{\lambda_i(\sum_{j=1}^{q}\sigma_{ij}^2)^2+\mathrm{tr}V\sum_{j=1}^{q}\sigma_{ij}^2}{(\mathrm{tr}V+\lambda_i\sum_{j=1}^{q}\sigma_{ij}^2)^2}$$

$$=\sum_{i=1}^{p}\frac{(\mathrm{tr}V)}{\lambda_i}\frac{(\sum_{j=1}^{q}\sigma_{ij}^2)^2+\frac{\mathrm{tr}V}{\lambda_i}(\sum_{j=1}^{q}\sigma_{ij}^2)}{(\frac{\mathrm{tr}V}{\lambda_i})^2+2\frac{\mathrm{tr}V}{\lambda_i}\sum_{j=1}^{q}\sigma_{ij}^2+(\sum_{j=1}^{q}\sigma_{ij}^2)^2}$$

$$<\sum_{i=1}^{p}\frac{\mathrm{tr}V}{\lambda_i}$$

故 $\mathrm{MSE}(\hat{a}(k_1,k_2,\cdots,k_p))<\mathrm{MSE}(\hat{a}(k))$。

由以上讨论，我们有如下的

定理 2　在均方误差损失意义，广义根方改进估计类优于 LS 估计。

由定理 1，定理 2，无论对 LS 估计，还是任一根方估计，总可以找到一组实数 $(k_1,k_2,\cdots,k_p)$ 使 MSE

$(\alpha(k_1,k_2,\cdots,k_p))$优于它们,这就说明,广义根方估计有较强的抗干扰性。

下面讨论γ的可容许性。为此,先引进一些有用的结论。

对线性模型$\begin{cases} Y_{n\times 1}=X_{n\times p}\beta_{n\times 1}+e_{n\times 1} \\ Y(X)=p\leqslant n, E(e)=0, \mathrm{cov}(e)=V>O \end{cases}$ (12)

令$m=XV^{-1}X$,$\hat{\beta}$是β的LS估计,有:

引理4[1]　$A\hat{\beta}\sim C\beta$的充要条件为:$Am^{-1}A'\leqslant Am^{-1}C$。

定理3　广义根方估计$\hat{a}(k_1,k_2,\cdots,k_p)$中$k_i$满足:

(i) $0<\lambda_i\leqslant 1$, $k_i\in[0,+\infty)$

(ii) $\lambda_i>1$, $k_i\in(-\infty,0)$

则广义根方改进估计$\hat{a}(k_1,k_2,\cdots,k_p)$是$\alpha$的可容性估计。

证明　模型(4)与引理4同意义下的m为

$$\begin{aligned} m &=(I_q\otimes\times Q)'(V\otimes I_n)^{-1}(I_q\otimes\times Q) \\ &=V^{-1}\otimes(Q'X'XQ) \\ &=V^{-1}\otimes\Lambda \end{aligned}$$

由于$\hat{a}(k_1,k_2,\cdots,k_p)=(I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)})\hat{a}$,故$A=I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}$,$C=I_{p\times p}$,从而

$$\begin{aligned} &Am^{-1}C-Am^{-1}A=[I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}][V\otimes\Lambda]^{-1}-[I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}][V\otimes\Lambda^{-1}][I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}] \\ &=[I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}][V\otimes\Lambda^{-1}][I_q\otimes\Lambda^{-(k_1,k_2,\cdots,k_p)}-I_q\otimes I_p][I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}] \end{aligned}$$

由Kronecker积性质,$I_q\otimes\Lambda^{(k_1,k_2,\cdots,k_p)}>0$且对称,又

$$\begin{aligned} &V\otimes\Lambda^{-1}[I_q\otimes\Lambda^{-(k_1,k_2,\cdots,k_p)}-I_q\otimes I_p] \\ &=V\otimes\Lambda^{-1-(k_1,k_2,\cdots,k_p)}-V\otimes\Lambda^{-1} \\ &=V\otimes[\Lambda^{-1-(k_1,k_2,\cdots,k_p)}-\Lambda^{-1}] \end{aligned}$$

由条件,当$0<\lambda_i\leqslant 1$时,$\lambda_i^{-1-k_i}\geqslant\lambda_i^{-1}$;当$\lambda_i>1$时,$\lambda_i^{-1-k_i}>\lambda_i^{-1}$,从而

$V\otimes\Lambda^{-1}[I_q\otimes\Lambda^{-(k_1,k_2,\cdots,k_p)}-I_q\otimes I_p]>0$。由矩阵不等式性质得:$Am^{-1}C-Am^{-1}A\geqslant 0$,由引理4,结论成立,证毕。综上所述,有

推论　令$\mathscr{F}$是γ的一个子集,$\mathscr{F}=\left\{\hat{a}(k_1,k_2,\cdots,k_p)\in\gamma \;\middle|\; \begin{matrix} 0<\lambda_i\leqslant 1\text{时}, k_i\in[0,+\infty) \\ \lambda_i>1\text{时}, k_i\in(-\infty,0) \end{matrix}\right\}$

$\mathscr{F}$是α的可容许广义根方估计类,它优于α的LSE估计,和根方改进估计。

$\mathscr{F}$在实际应用中起较大作用,但对$(k_1,k_2,\cdots,k_p)$的选择,目前还没有较好的方法,只是凭经验,尽管有一些准则可行(如$Q(c)$准则),加上计算机的广泛应用,人们采用类比法也能得到较好的估计效果,其主观性还是较大,此问题有待进一步研究。

本文得到导师林伯明副教授的悉心指导,西北工业大学朱燕堂教授、王潮杰教授提出了宝贵意见,作者对此表示衷心感谢。

参 考 文 献

1 王松桂. 线性模型的理论及应用. 合肥:安徽教育出版社,1987

2 Zellner A. J. Amer Statist Assoc,1962,57:348~369

3 夏结来. 回归系数的根方有偏估计及其应用. 数理统计及应用概率,1988,3(1):21~30

4 Fseber G A. Linear Regression Analysis. John Viley & Sons, 1977

5 Wang Songui. Corariance improvement estimation of the parameter in selming unrelated regression models. The proceeding of the second Japan-china sympasiam on statistics. Japan. 1986

6 倪国熙. 常用矩阵的理论和方法. 上海:上海科学技术出版社,1987

7 Gotz Trenkler. Mean square error matrix imparision of estimators in linear regression, Comm Statist. Theory. Meth. 1985, 14(10):14~20

责任编辑　张素敏

(下转第480页)