应用数学和力学编委会编
重庆出版社出版

多元线性模型回归系数的压缩LS估计*

陈世基　　曾志斌

（福州 福建师范大学数学系）（福州 福建省统计局）

（林宗池推荐，1993年1月4日收到）

摘　　要

本文采用压缩LS估计$\hat{B}(k)$来估计设计矩阵呈病态的多元线性模型的回归系数B。通过k值的选取，可使$\hat{\beta}(k)=\mathrm{Vec}(\hat{B}(k))$的均方误差MSE小于$\beta=\mathrm{Vec}(B)$的LS估计$\beta^*$的MSE。证明了$\hat{\beta}(k)$具有可容许性、抗干扰性和有效性，并给出了实际应用中选取k值的方法。

关键词　多元线性模型　最小二乘估计　压缩LS估计　均方误差

一、引　　言

多元线性模型

$$\left.\begin{array}{l} Y=XB+\varepsilon \\ \varepsilon\text{ 的行向量互不相关，均值为零，有共同协差阵 } V \end{array}\right\} \tag{1.1}$$

这里，$Y=(y_{ij})_{n\times q}=(y_1,y_2,\cdots,y_q)$为$n\times q$的随机观测矩阵，$X$为$n\times p$的设计矩阵，$R(X)=p$，$B=(\beta_{ij})_{p\times q}=(\beta_1,\beta_2,\cdots,\beta_q)$为$p\times q$的未知参数矩阵，$\varepsilon=(\varepsilon_{ij})_{n\times q}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_q)$为$n\times q$的随机误差矩阵，$V=(v_{ij})_{q\times q}$为$q\times q$的未知正定矩阵。

易知模型(1.1)的典则形式为

$$\left.\begin{array}{l} Y=ZA+\varepsilon \\ \varepsilon\text{ 的行向量互不相关，均值为零，有共同协差阵 } V \end{array}\right\} \tag{1.2}$$

这里，$Z=XQ$，$A=Q'B$，Q为$p\times p$正交矩阵，使得$Q'X'XQ=\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_p)$，称$A$为典则参数矩阵。于是模型(1.1)，(1.2)可分别化成如下的一元线性模型

$$\left.\begin{array}{l} \mathrm{Vec}(Y)=(I_q\otimes X)\mathrm{Vec}(B)+\mathrm{Vec}(\varepsilon) \\ \mathrm{Cov}(\mathrm{Vec}(\varepsilon))=V\otimes I_n,\ E(\mathrm{Vec}(\varepsilon))=0 \end{array}\right\} \tag{1.1$'$}$$

$$\left.\begin{array}{l} \mathrm{Vec}(Y)=(I_q\otimes Z)\mathrm{Vec}(A)+\mathrm{Vec}(\varepsilon) \\ \mathrm{Cov}(\mathrm{Vec}(\varepsilon))=V\otimes I_n,\ E(\mathrm{Vec}(\varepsilon))=0 \end{array}\right\} \tag{1.2$'$}$$

记$\beta=\mathrm{Vec}(B)$，$\alpha=\mathrm{Vec}(A)=(I_q\otimes Q')\mathrm{Vec}(B)$，称$\alpha$为典则参数。易得$B$和$A$的LS估计分别为

* 福建省自然科学基金资助项目

$$B^*=(X'X)^{-1}X'Y \tag{1.3}$$

$$A^*=\Lambda^{-1}Z'Y \tag{1.4}$$

相应地，β和α的LS估计分别为

$$\beta^*=\mathrm{Vec}(B^*)=[I_q\otimes(X'X)^{-1}X']\mathrm{Vec}(Y) \tag{1.3$'$}$$

$$\alpha^*=\mathrm{Vec}(A^*)=(I_q\otimes\Lambda^{-1}Z')\mathrm{Vec}(Y) \tag{1.4$'$}$$

且$E(\beta^*)=\beta$，$\mathrm{Cov}(\beta^*)=V\otimes(X'X)^{-1}$，此时$\beta^*$的均方误差为

$$\mathrm{MSE}(\beta^*)=(\mathrm{tr}V)\sum_{i=1}^{p}\lambda_i^{-1} \tag{1.5}$$

这里$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p$为$X'X$的特征根，因为$X'X>0$，所以$\lambda_i>0$，$i=1, 2, \cdots, p$.

从(1.5)可知，若$X'X$至少有一特征根很小，则$\mathrm{MSE}(\beta^*)$就很大，此时虽然LS估计β^*的方差在线性无偏估计类中最小，但其值很大，使得估计精度较差，表现出相当的不稳定。文[1]指出$X'X$有多少个特征根接近于零，回归自变量之间就存在多少个近似的线性关系。文[1]称之为复共线性，此时称设计矩阵X呈病态，可见导致LS估计的性质变坏之原因是复共线性的存在。为改进LS估计，[2]、[3]分别提出了多元广义岭估计和多元根方估计。本文引入压缩LS估计$\hat{B}(k)$，给出了$\hat{\beta}(k)=\mathrm{Vec}(\hat{B}(k))$优于$\beta^*$的充要条件并讨论了$\hat{\beta}(k)$的一些统计优良性质。

二、压缩LS估计的定义及其优于LS估计的充要条件

定义1 若存在$1\leqslant r<p$使$\lambda_{r+1}<1\leqslant\lambda_r$，记

$$K=\mathrm{diag}\left(\frac{\lambda_1-1+k}{\lambda_1},\cdots,\frac{\lambda_r-1+k}{\lambda_r},k\lambda_{r+1},\cdots,k\lambda_p\right) \tag{2.1}$$

这里$k\in[\lambda_p,1)$为常数，则称

$$\hat{B}(k)=QKQ'B^* \tag{2.2}$$

为模型(1.1)回归系数B的压缩LS估计。

由于$X'X=Q\Lambda Q'$，所以从(2.2)即可得到

$$\hat{B}(k)=QKQ'(X'X)^{-1}X'Y=QK\Lambda^{-1}Q'X'Y=(QK^{-1}\Lambda Q')^{-1}X'Y$$

与LS估计B^*比较，压缩LS估计是把$X'X=Q\Lambda Q'$换成$QK^{-1}\Lambda Q'$而得到的。直观上这样做的理由是明显的。当X呈病态时，$X'X$的特征根至少有一个非常接近于零，而$QK^{-1}\Lambda Q'$的特征根为

$$\frac{\lambda_1^2}{\lambda_1-1+k},\ \frac{\lambda_2^2}{\lambda_2-1+k},\ \cdots,\ \frac{\lambda_r^2}{\lambda_r-1+k},\ k^{-1},\ \cdots,\ k^{-1}$$

它们都不小于1，从而“打破”原来设计矩阵的复共线性，达到降低估计量的均方误差之目的。对于不同的k，(2.2)给出了不同的估计，可见压缩LS估计是一个很大的估计类。从(2.1)可以看出，k保证了K的对角元不至于在$\lambda=1$处出现跳跃，这对于估计的稳定性是必要的，于是通常把定义1中的k称为平稳参数。

记$\hat{\beta}(k)=\mathrm{Vec}(\hat{B}(k))=(I_q\otimes QKQ')\beta^*$，为了叙述上的方便，也把$\hat{\beta}(k)$称为$\beta=\mathrm{Vec}(B)$的压缩LS估计。

定理1 $\hat{\beta}(k)$是β的压缩有偏估计。

证明 $\|\hat{\beta}(k)\|^2=\|(I_q\otimes Q)(I_q\otimes K)(I_q\otimes Q')\beta^*\|^2=\|(I_q\otimes K)(I_q\otimes Q')\beta^*\|^2$

$$<\|(I_q\otimes Q')\beta^*\|^2=\|\beta^*\|^2$$

$$E(\hat{\beta}(k))=(I_q\otimes QKQ')\beta\neq\beta,\qquad(\beta\neq 0)$$

否则$QKQ'=I_p$，即$K=I_p$，这与(2.1)矛盾，所以$\hat{\beta}(k)$是β的压缩型有偏估计。

引理1 若$D>0$，$d>0$，则$dD-xx'>0$的充要条件是$x'D^{-1}x<d$，这里D为m阶矩阵，d为实数，x为$m\times 1$向量。

证明 依给定的条件有

$$\begin{pmatrix} I & 0 \\ -x'D^{-1} & 1 \end{pmatrix}\begin{pmatrix} D & x \\ x' & d \end{pmatrix}\begin{pmatrix} I & -D^{-1}x \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} D & 0 \\ 0 & d-x'D^{-1}x \end{pmatrix}$$

$$\begin{pmatrix} I & -d^{-1}x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} D & x \\ x' & d \end{pmatrix}\begin{pmatrix} I & 0 \\ -d^{-1}x' & 1 \end{pmatrix}=\begin{pmatrix} D-d^{-1}xx' & 0 \\ 0 & d \end{pmatrix}$$

所以 $\begin{pmatrix} D & 0 \\ 0 & d-x'D^{-1}x \end{pmatrix}$与$\begin{pmatrix} D-d^{-1}xx' & 0 \\ 0 & d \end{pmatrix}$合同

由于D为正定矩阵，d为正实数，所以

$$\begin{pmatrix} D & 0 \\ 0 & d-x'D^{-1}x \end{pmatrix}$$

正定的充要条件是$d-x'D^{-1}x>0$。

$$\begin{pmatrix} D-d^{-1}xx' & 0 \\ 0 & d \end{pmatrix}$$

正定的充要条件是$D-d^{-1}xx'>0$。于是$dD-xx'>0$的充要条件是$x'D^{-1}x<d$。

定理2 $\mathrm{MSEM}(\hat{\beta}(k))<\mathrm{MSEM}(\beta^*)$ 的充要条件是 $\beta'R\beta<1$。这里$\mathrm{MSEM}(\hat{\beta}(k))$，$\mathrm{MSEM}(\beta^*)$分别为$\hat{\beta}(k)$和$\beta^*$的均方误差矩阵，$R=V^{-1}\otimes QCQ'$，

$$C=\mathrm{diag}\left(\frac{\lambda_1-k\lambda_1}{2\lambda_1-1+k},\cdots,\frac{\lambda_r-k\lambda_r}{2\lambda_r-1+k},\ \frac{\lambda_{r+1}-k\lambda_{r+1}^2}{1+k\lambda_{r+1}},\cdots,\frac{\lambda_p-k\lambda_p^2}{1+k\lambda_p}\right)$$

证明 令$G=I_q\otimes QKQ'$，则$\hat{\beta}(k)=G\beta^*$，依[4]有

$$\begin{aligned}\mathrm{MSEM}(\hat{\beta}(k))&=\mathrm{Cov}(G\beta^*)+[E(G\beta^*)-\beta][E(G\beta^*)-\beta]'\\&=G[V\otimes(X'X)^{-1}]G'+(G-I_{pq})\beta\beta'(G-I_{pq})'\\&=V\otimes QK\Lambda^{-1}KQ'+[I_q\otimes Q(K-I_p)Q']\\&\quad\cdot\beta\beta'[I_q\otimes(K-I_p)Q']\end{aligned}$$

于是

$$\begin{aligned}&\mathrm{MSEM}(\beta^*)-\mathrm{MSEM}(\hat{\beta}(k))\\&=V\otimes(X'X)^{-1}-V\otimes QK\Lambda^{-1}KQ'-[I_q\otimes Q(K-I_p)Q']\beta\beta'[I_q\otimes Q(K-I_p)Q']\\&=V\otimes Q(\Lambda^{-1}-K\Lambda^{-1}K)Q'-[I_q\otimes Q(K-I_p)Q']\beta\beta'[I_q\otimes Q(K-I_p)Q']\\&=V\otimes Q\Delta_1Q'-(I_q\otimes Q\Delta_2Q')\beta\beta'(I_q\otimes Q\Delta_2Q')\end{aligned}$$

这里

$$\Delta_1=\Lambda^{-1}-K\Lambda^{-1}K=\mathrm{diag}\left(\frac{2(1-k)}{\lambda_1^2}-\frac{(1-k)^2}{\lambda_1^3},\cdots,\right.$$

$$\left.\frac{2(1-k)}{\lambda_r^2}-\frac{(1-k)^2}{\lambda_r^3},\ \frac{1-k^2\lambda_{r+1}^2}{\lambda_{r+1}},\cdots,\frac{1-k^2\lambda_p^2}{\lambda_p}\right)$$

$$\Delta_2=I_p-K=\mathrm{diag}\left(\frac{1-k}{\lambda_1},\cdots,\frac{1-k}{\lambda_r},1-k\lambda_{r+1},\cdots,1-k\lambda_p\right)$$

依引理1，$V\otimes Q\Delta_1Q'-(I_q\otimes Q\Delta_2Q')\beta\beta'(I_q\otimes Q\Delta_2Q')>0$的充要条件是

$$\beta'(I_q\otimes Q\Delta_2Q')(V\otimes Q\Delta_1Q')^{-1}(I_q\otimes Q\Delta_2Q')\beta<1$$

即　$$\beta'(V^{-1}\otimes Q\Delta_2\Delta_1^{-1}\Delta_2Q')\beta<1$$

记

$$C=\Delta_2\Delta_1^{-1}\Delta_2=\mathrm{diag}\left(\frac{\lambda_1-k\lambda_1}{2\lambda_1-1+k},\cdots,\frac{\lambda_r-k\lambda_r}{2\lambda_r-1+k},\frac{\lambda_{r+1}-k\lambda_{r+1}^2}{1+k\lambda_{r+1}},\cdots,\frac{\lambda_p-k\lambda_p^2}{1+k\lambda_p}\right)\tag{2.3}$$

$$R=V^{-1}\otimes QCQ'\tag{2.4}$$

因此，$\mathrm{MSEM}(\beta^*)-\mathrm{MSEM}(\hat{\beta}(k))>0$的充要条件是：

$$\beta'R\beta<1\tag{2.5}$$

根据[4]可得如下的推论

推论1　当且仅当$\beta'R\beta<1$时，有

$$\mathrm{MSE}(\hat{\beta}(k))<\mathrm{MSE}(\beta^*)$$

三、压缩LS估计的性质

3.1　可容许性

定理3　在线性估计类中，$\hat{\beta}(k)$是β的可容许估计.

证明　令$G=I_q\otimes QKQ'$，$W=(I_q\otimes X')(V\otimes I_n)^{-1}(I_q\otimes X)=V^{-1}\otimes X'X$则

$$\begin{aligned}GW^{-1}-GW^{-1}G'&=(I_q\otimes QKQ')[V\otimes(X'X)^{-1}]-(I_q\otimes QKQ')[V\otimes(X'X)^{-1}](I_q\otimes QKQ')\\&=V\otimes QKQ'(X'X)^{-1}-V\otimes QKQ'(X'X)^{-1}QKQ'\\&=V\otimes QK\Lambda^{-1}Q'-V\otimes QK\Lambda^{-1}KQ'\\&=(I_q\otimes Q)[V\otimes(K\Lambda^{-1}-K\Lambda^{-1}K)](I_q\otimes Q')\end{aligned}$$

由于

$$K\Lambda^{-1}-K\Lambda^{-1}K=\mathrm{diag}\left(\frac{(\lambda_1-1+k)(1-k)}{\lambda_1^3},\ \cdots,\ \frac{(\lambda_r-1+k)(1-k)}{\lambda_r^3},\ k(1-k\lambda_{r+1}),\ \cdots,\ k(1-k\lambda_p)\right)\geqslant0$$

且$V>0$，所以$GW^{-1}-GW^{-1}G'\geqslant0$. 依[4]则$\hat{\beta}(k)$是$\beta$的可容许估计.

3.2　抗干扰性

定义2　对于矩阵N，称由$\|N\|=\sqrt{\lambda_{\max}(N'N)}$定义的范数$\|\cdot\|$为$N$的谱范数，其中$\lambda_{\max}(N'N)$为$N'N$的最大特征根.

由定义2可得，当N为非负定矩阵时，则有$\|N\|=\lambda_{\max}(N)$.

定义3　对于估计$\bar{\beta}=H\beta^*$，若H可逆且模型(1.1)的设计矩阵X列满秩，则称

$$\mathrm{cond}(\bar{\beta})=\|H^{-1}(I_q\otimes X'X)\|\cdot\|H(I_q\otimes X'X)^{-1}\|$$

为估计$\hat{\beta}$的条件数。

由于$\hat{\beta}(k)=(I_q\otimes QKQ')\beta^*$，$X'X=Q\Lambda Q'$，所以

$$\mathrm{cond}(\hat{\beta}(k))=\|I_q\otimes QK^{-1}\Lambda Q'\|\|I_q\otimes QK\Lambda^{-1}Q'\|=\|I_q\otimes K^{-1}\Lambda\|\|I_q\otimes K\Lambda^{-1}\|$$

同理

$$\mathrm{cond}(\beta^*)=\|I_q\otimes X'X\|\|(I_q\otimes X'X)^{-1}\|=\|I_q\otimes\Lambda\|\|I_q\otimes\Lambda^{-1}\|=\lambda_1/\lambda_r$$

模型(1.1)′的正则方程为

$$(I_q\otimes X')(V\otimes I_n)^{-1}(I_q\otimes X)\beta=(I_q\otimes X')(V\otimes I_n)^{-1}\mathrm{Vec}(Y) \tag{3.1}$$

由于$V>0$，所以方程(3.1)可化为

$$(I_q\otimes X'X)\beta=(I_q\otimes X')\mathrm{Vec}(Y) \tag{3.2}$$

依[5]，当$\mathrm{cond}(I_q\otimes X'X)\|\Delta(I_q\otimes X'X)\|\|I_q\otimes X'X\|^{-1}<1$时

必有

$$\frac{\|\Delta\beta\|}{\|\beta\|}\leqslant\frac{\mathrm{cond}(I_q\otimes X'X)}{1-\mathrm{cond}(I_q\otimes X'X)\|\Delta(I_q\otimes X'X)\|\|I_q\otimes X'X\|^{-1}}$$

$$\left[\frac{\|\Delta(I_q\otimes X'X)\|}{\|I_q\otimes X'X\|}+\frac{\|\Delta((I_q\otimes X')\mathrm{Vec}(Y))\|}{\|(I_q\otimes X')\mathrm{Vec}(Y)\|}\right]$$

这里 $\dfrac{\|\Delta\beta\|}{\|\beta\|}$，$\dfrac{\|\Delta(I_q\otimes X'X)\|}{\|I_q\otimes X'X\|}$，$\dfrac{\|\Delta((I_q\otimes X')\mathrm{Vec}(Y))\|}{\|(I_q\otimes X')\mathrm{Vec}(Y)\|}$

分别表示β，$I_q\otimes X'X$，$(I_q\otimes X')\mathrm{Vec}(Y)$的相对误差。

由于$\mathrm{cond}(I_q\otimes X'X)=\mathrm{cond}(\beta^*)$，所以$\beta^*$的条件数反映了$\beta^*$抗干扰的程度，于是自然希望新估计$\bar{\beta}$既能改进LS估计，同时又不降低$\bar{\beta}^*$的抗干扰能力，这就得要求$\mathrm{cond}(\bar{\beta})\leqslant\mathrm{cond}(\beta^*)$。通常称满足这个条件的估计$\bar{\beta}$具有相对于$\beta^*$更强的抗干扰性。

定理4 对线性模型(1.1)′，总有$\mathrm{cond}(\hat{\beta}(k))\leqslant\mathrm{cond}(\beta^*)$，即压缩LS估计改进了LS估计的抗干扰性。

证明 由于

$$\|I_q\otimes K^{-1}\Lambda\|=\max\left(\frac{\lambda_1^2}{\lambda_1-1+k},\cdots,\frac{\lambda_r^2}{\lambda_r-1+k},k^{-1}\right)\leqslant\lambda_1k^{-1}\leqslant\frac{\lambda_1}{\lambda_r}$$

$$\|I_q\otimes K\Lambda^{-1}\|=\max\left(\frac{\lambda_1-1+k}{\lambda_1^2},\cdots,\frac{\lambda_r-1+k}{\lambda_r^2},k\right)\leqslant1$$

所以 $$\mathrm{cond}(\hat{\beta}(k))=\|I_q\otimes K^{-1}\Lambda\|\|I_q\otimes K\Lambda^{-1}\|\leqslant\lambda_1/\lambda_r=\mathrm{cond}(\beta^*)$$

即$\hat{\beta}(k)$具有相对于β^*更强的抗干扰性。

3.3 有效性

讨论一个估计量的优劣通常是以它的均方误差来衡量的。推论1已指出，在MSE意义下，$\hat{\beta}(k)$优于β^*，也可以用相对效率（简称效率）来衡量一个估计量的优劣。对于一元线性模型回归系数的无偏估计，[6]、[7]给出了两种不同形式的相对效率。对于本文所讨论的有偏估计，笔者引入如下的效率：

定义4 对于模型(1.1)′，若$\bar{\beta}$为β的线性有偏估计，则称

$$\rho(\bar{\beta})=1-\frac{\mathrm{MSE}(\bar{\beta})}{\mathrm{MSE}(\beta^*)} \tag{3.3}$$

为$\bar{\beta}$相对于β^*的效率。

若 $\mathrm{MSE}(\bar{\beta})\leqslant\mathrm{MSE}(\beta^*)$，则$\bar{\beta}$的效率不小于零，因此 $0\leqslant\rho(\bar{\beta})\leqslant 1$. 效率越是接近于1，$\bar{\beta}$改进$\beta^*$的程度就越高. 下面的定理将给出压缩LS估计$\hat{\beta}(k)$相对于LS估计$\beta^*$的效率的下界.

定理5　当$\lambda_p[1+(\mathrm{tr}V)^{-1}\alpha'\alpha]\leqslant 1$时，有

$$\rho(\hat{\beta}(k))\geqslant\frac{1-\lambda_p-\lambda_p(\mathrm{tr}V)^{-1}\alpha'\alpha}{1+\lambda_p(p-1)} \tag{3.4}$$

证明　依定义1，模型(1.2)′的回归系数α的压缩LS估计为$\hat{\alpha}(k)=(I_q\otimes K)\alpha^*$，且$\hat{\beta}(k)=(I_q\otimes Q)\hat{\alpha}(k)$，则

$$\begin{aligned}\mathrm{MSE}(\hat{\beta}(k))&=\mathrm{MSE}(\hat{\alpha}(k))=\mathrm{trCov}(\hat{\alpha}(k))+\|E(\hat{\alpha}(k))-\alpha\|^2\\&=\mathrm{tr}(V\otimes K\Lambda^{-1}K)+\alpha'[I_q\otimes(K-I_p)]^2\alpha\end{aligned}$$

又$\mathrm{MSE}(\beta^*)=\mathrm{MSE}(\alpha^*)=\mathrm{tr}(V\otimes\Lambda^{-1})$，于是

$$\begin{aligned}\rho(\hat{\beta}(k))&=1-\frac{\mathrm{MSE}(\hat{\beta}(k))}{\mathrm{MSE}(\beta^*)}\\&=1-\frac{\mathrm{tr}V\,\mathrm{tr}(K\Lambda^{-1}K)+\alpha'[I_q\otimes(K-I_p)^2]\alpha}{\mathrm{tr}V\cdot\mathrm{tr}\Lambda^{-1}}\\&\geqslant 1-\frac{\mathrm{tr}V\cdot\mathrm{tr}(K\Lambda^{-1})+\alpha'\alpha}{\mathrm{tr}V\cdot\mathrm{tr}\Lambda^{-1}}\\&=1-\frac{\sum_{i=1}^{r}(\lambda_i-1+k)\lambda_i^{-2}+(p-r)k+(\mathrm{tr}V)^{-1}\alpha'\alpha}{\sum_{i=1}^{p}\lambda_i^{-1}}\\&\geqslant 1-\frac{\sum_{i=1}^{r}\lambda_i^{-1}+(p-r)+(\mathrm{tr}V)^{-1}\alpha'\alpha}{\sum_{i=1}^{r}\lambda_i^{-1}+(p-r)+\lambda_p^{-1}-1}\end{aligned}$$

由于当$a_1\leqslant a_2$时，$(x+a_1)/(x+a_2)$是$x(x>-a_2)$的增函数，而由已知条件有$(\mathrm{tr}V)^{-1}\alpha'\alpha\leqslant\lambda_p^{-1}-1$，所以

$$\rho(\hat{\beta}(k))\geqslant 1-\frac{r+(p-r)+(\mathrm{tr}V)^{-1}\alpha'\alpha}{r+(p-r)+\lambda_p^{-1}-1}=\frac{1-\lambda_p-\lambda_p(\mathrm{tr}V)^{-1}\alpha'\alpha}{1+\lambda_p(p-1)}$$

可见，当设计矩阵X呈病态（即λ_p很小）且使$\lambda_p(\mathrm{tr}V)^{-1}\alpha'\alpha$相对较小时，$\rho(\hat{\beta}(k))$可以很接近于1，也就是说$\hat{\beta}(k)$可以很好地改进$\beta^*$，且$\rho(\beta(k))$的下界与参数$k$无关. 然而$\rho(\hat{\beta}(k))$的下界含有$\alpha$和$V$，这会给应用带来不便，对于存在较多个复共线关系的大型回归问题，将可以避免这点，且$\hat{\beta}(k)$相对于β^*的效率可以达到很高.

定理6　若$\lambda_{p-m}=\lambda_{p-m+1}=\cdots=\lambda_p\approx 0$，$0\leqslant m\leqslant p-r-1$，则在椭球(2.5)内，必有

$$\rho(\hat{\beta}(k))\geqslant\frac{m(1-k^2\lambda_p^2)}{(p-m-1)\lambda_p+m+1}\approx\frac{m}{m+1}\left(1-\frac{p-m-1}{m+1}\lambda_p\right) \tag{3.5}$$

证明　首先证明在椭球(2.5)内有

$$\alpha'[I_q\otimes(K-I_p)^2]\alpha\leqslant\frac{\mu(1-k^2\lambda_p^2)}{\lambda_p} \tag{3.6}$$

其中μ为V的最大特征根，即$\mu=\lambda_{\max}(V)$。由$\alpha=(I_q\otimes Q')\beta$及(2.3)、(2.4)可知(2.5)与

$$\alpha'[V^{-1}\otimes(K-I_p)^2(I_p-K^2)^{-1}\Lambda]\alpha<1 \tag{3.7}$$

等价．由于$\lambda\geqslant1$时，

$$\frac{\lambda^3}{2\lambda-2k\lambda-(1-k)^2}$$

为λ的增函数，所以

$$\frac{\lambda^3}{2\lambda-2k\lambda-(1-k)^2}\geqslant\frac{1}{1-k^2}$$

又当$\lambda_p\leqslant\lambda\leqslant1$时，$\lambda/(1-k^2\lambda)$为$\lambda$的单调递增函数．于是

$$\frac{\lambda_p}{1-k^2\lambda_p^2}\leqslant\frac{\lambda}{1-k^2\lambda}\leqslant\frac{1}{1-k^2}$$

由此可得

$$(I_p-K^2)^{-1}\Lambda=\operatorname{diag}\left(\frac{\lambda_1^3}{2\lambda_1-2k\lambda_1-(1-k)^2},\cdots,\frac{\lambda_r^3}{2\lambda_r-2k\lambda_r-(1-k)^2},\right.$$

$$\left.\frac{\lambda_{r+1}}{1-k^2\lambda_{r+1}^2},\cdots,\frac{\lambda_p}{1-k^2\lambda_p^2}\right)\geqslant\frac{\lambda_p}{1-k^2\lambda_p^2}I_p$$

又$V\leqslant\mu I_q$，即$V^{-1}\geqslant\mu^{-1}I_q$，所以

$$V^{-1}\otimes(I_p-K^2)^{-1}\Lambda\geqslant\frac{\lambda_p}{\mu(1-k^2\lambda_p^2)}\cdot(I_q\otimes I_p)$$

结合(3.7)式即可得到不等式(3.6)。

现在来证明(3.5)式，利用不等式(3.6)及$\mu^{-1}\mathrm{tr}V\leqslant1$，则

$$\rho(\hat{\beta}(k))=1-\frac{\mathrm{tr}V\cdot\mathrm{tr}(K\Lambda^{-1}K)+\alpha'[I_q\otimes(K-I_p)^2]\alpha}{\mathrm{tr}V\cdot\mathrm{tr}\Lambda^{-1}}$$

$$\geqslant1-\frac{\mathrm{tr}V\cdot\mathrm{tr}(K\Lambda^{-1}K)+\lambda_p^{-1}\mu(1-k^2\lambda_p^2)}{\mathrm{tr}V\cdot\mathrm{tr}\Lambda^{-1}}$$

$$=1-\frac{\mathrm{tr}(K\Lambda^{-1}K)+(\mathrm{tr}V)^{-1}\mu\cdot\lambda_p^{-1}(1-k^2\lambda_p^2)}{\mathrm{tr}\Lambda^{-1}}$$

$$\geqslant1-\frac{\mathrm{tr}(K\Lambda^{-1}K)+\lambda_p^{-1}(1-k^2\lambda_p^2)}{\mathrm{tr}\Lambda^{-1}}$$

$$=1-\frac{\sum\limits_{i=1}^{r}(\lambda_i-1+k)^2\lambda_i^{-3}+k^2\lambda_{r+1}+\cdots+k^2\lambda_{p-m-1}+mk^2\lambda_p+\lambda_p^{-1}}{\sum\limits_{i=1}^{r}\lambda_i^{-1}+\lambda_{r+1}^{-1}+\cdots+\lambda_{p-m-1}^{-1}+(m+1)\lambda_p^{-1}}$$

$$\geqslant 1-\frac{\sum_{i=1}^{r}\lambda_i^{-1}+(p-m-r-1)+mk^2\lambda_p-\lambda_p^{-1}}{\sum_{i=1}^{r}\lambda_i^{-1}+(p-m-r-1)+(m+1)\lambda_p^{-1}}$$

$$\geqslant 1-\frac{p-m-1+mk^2\lambda_p+\lambda_p^{-1}}{p-m-1+(m+1)\lambda_p^{-1}}$$

$$=\frac{m(1-k^2\lambda_p^2)}{(p-m-1)\lambda_p+(m+1)}$$

$$\approx\frac{m}{m+1}\cdot\frac{1-[(p-m-1)/(m+1)]^2\lambda_p^2}{1+[(p-m-1)/(m+1)]\lambda_p}$$

$$=\frac{m}{m+1}\left(1-\frac{p-m-1}{m+1}\lambda_p\right)$$

四、平稳参数k的确定

从定理2及推论1知，当k满足(2.5)式时，$\hat{\beta}(k)$优于β^*。但(2.5)与未知参数β，V有关，实际应用中还必须通过实际数据来确定，此时$\hat{\beta}(k)$实际上是一种非线性估计。

令 $$C_0=\mathrm{diag}\left(\frac{\lambda_1-k\lambda_1}{2\lambda_1-1},\cdots,\frac{\lambda_r-k\lambda_r}{2\lambda_r-1},\ \lambda_{r+1}-k\lambda_{r+1}^2,\ \cdots,\ \lambda_p-k\lambda_p^2\right)$$

v为V的最小特征根，即$v=\lambda_{\min}(V)$，则$C<C_0$，$V^{-1}\leqslant v^{-1}I_q$。由于(2.5)与$\alpha'(V^{-1}\otimes C)\alpha<1$等价，所以$\mathrm{MSE}(\hat{\beta}(k))<\mathrm{MSE}(\beta)^*$成立的一个充分条件是

$$\alpha'(I_q\otimes C_0)\alpha\leqslant v \tag{4.1}$$

由于 $$V^*=\frac{1}{n-p}Y'[I-X(X'X)^{-1}X']Y$$

为V的无偏估计，

$$\alpha^*=(I_q\otimes\Lambda^{-1}Z')\mathrm{Vec}(Y)$$

为α的LS估计，所以以α^*及$v^*=\lambda_{\min}(V^*)$代替(4.1)中的α及v，即

$$\alpha^{*\prime}(I_q\otimes C_0)\alpha^*\leqslant v^* \tag{4.2}$$

从(4.2)可以得到k必须满足$k\geqslant k^*$。于是k的估计值可取为

$$\hat{k}=\begin{cases}\lambda_p, & k^*\leqslant\lambda_p\\ k^*, & \lambda_p<k^*<1\\ 1, & k^*\geqslant 1\end{cases}$$

至于选取k值的其它方法及各种方法之间的比较，尚有待于进一步的探讨。

参 考 文 献

[1] 陈希孺、王松桂，《近代回归分析——原理、方法及其应用》，安徽教育出版社，合肥（1987）。

[2] 曾志斌，多元广义岭估计及K值选取的$Q(C)$准则，福建师范大学学报（自然科学版），6(4)(1990)，25—33。

[3] 曾志斌，多元线性模型回归系数的根方估计，福建师范大学学报（自然科学版），7(3)(1991)，35—40。

[4] 王松桂，《线性模型的理论及其应用》，安徽教育出版社，合肥（1987）.
[5] 孙维广，《矩阵扰动分析》，科学出版社，北京（1987）.
[6] Bloomfield, P. and G. S. Waston, The efficiency of least squares, *Biometrika*, 62 (1975), 121—128.
[7] 刘爱义、王松桂，线性模型中LS估计的一种新的相对效率，应用概率统计，5(2) (1987), 97—104.
[8] Arnold, S. F., *Theory of Linear Models and Multivariate Analisis*, John Wiley, New York (1981).
[9] 杨虎，单参数主成分回归估计，高校应用数学学报，4(1) (1989), 74—80.

The Compression LS Estimate of Regression Coefficient in Multivariate Linear Model

Chen Shi-ji

(*Department of Mathematics, Fujian Normal University, Fuzhou*)

Zeng Zhi-bin

(*Statistics Bureau of Fujian Province, Fuzhou*)

Abstract

In this paper, compression LS estimate $\hat{B}(k)$ of the regression coefficient B is considered when the design matrix present ill-condition in multivariate linear model. The MSE (mean square error) of the estimate $\hat{\beta}(k)=\mathrm{Vec}(\hat{B}(k))$ is less than the MSE of LS estimate β^* of the regression coefficient $\beta=\mathrm{Vec}(B)$ by choosing the parameter k. Admissibility, numerical stability and relative efficiency of $\hat{\beta}(k)$ are proved. The method of determining k value for practical use is also suggested.

Key words multivariate linear model, least square estimate, compression LS estimate, mean square error