第14卷 第2期　　华侨大学学报自然科学版　　Vol. 14 No. 2
JOURNAL OF HUAQIAO UNIVERSITY
1993年4月　　(NATURAL SCIENCE)　　Apr. 1993

两个半相依回归系统参数的 Stein 估计

陈建伟

(管理信息科学系)

摘要　Zellner 等学者研究了一类相依回归系统(1)中参数的最小二乘估计的改进，这些改进估计都是无编估计且具有许多优良性质，但是当设计阵呈病态时这些改进估计不再被认为是良好估计．本文提出了一种有偏 Stein 估计，证明了在设计阵呈病态时 Stein 估计的优良性质．

关键词　半相依回归系统，两步估计，协方差改进估计，Stein 估计

0　引言

在计量经济研究中，常常要建立线性回归方程组

$$\begin{cases} y_i = x_i\beta_i + \epsilon_i, (i = 1,2), \\ E\epsilon_i = 0; \mathrm{cov}(\epsilon_i, \epsilon_j) = \sigma_{ij} I_n, \end{cases} \tag{1}$$

其中 y_i 为 $n\times 1$ 的观测向量，x_i 是 $n\times p_i$ 的秩为 p_i 设计矩阵，β_i 为 $p_i\times 1$ 的未知参数，回归方程(1)可以写成如下

$$\begin{cases} y = x\beta + \epsilon, \\ E\epsilon = 0; \mathrm{cov}(\epsilon, \epsilon) = (\sum \otimes I), \sum > 0, \end{cases} \tag{2}$$

这里，$y=(y_1^\tau, y_2^\tau)^\tau$，$x=\mathrm{diag}(x_1, x_2)$，$\beta=(\beta_1^\tau, \beta_2^\tau)^\tau$，$\epsilon=(\epsilon_1^\tau, \epsilon_2^\tau)^\tau$，$\sum=(\sigma_{ij})$为二阶正定矩阵，我们称式(1)、(2)为半相依回归系统(seemingly unrelated regression system)，这种模型在经济和工程等领域中都有重要的应用．众所周知，当$\sum$为已知时，β 的 BLU 估计为

$$\beta = [x^\tau(\sum{}^{-1} \otimes I)x]^{-1} x^\tau(\sum{}^{-1} \otimes I)y,$$

记号"$\otimes$"表示矩阵的 *Kroneker* 乘积，但是在实际问题中$\sum$往往是未知的，Zellner[1]等提出两步估计法(two-stage estimator)最近 Wang Songgui[2]提出一种新的有效估计法—两步协方差

本文 1991-12-03 收到．

* 国务院侨办基金资助课题．

改进估计，即回归系统(1)中 β_1 的协方差改进估计为

$$\tilde{\beta}_1=(x_1^{\tau}x_1)^{-1}x_1^{\tau}y_1-(\sigma_{12}/\sigma_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2y_2,$$

其中，$N_i=I-x_i(x_i^{\tau}x_i)^{-1}x_i^{\tau}$，$(i=1,2)$. 当 $\sum$ 未知时用它的估计 $S=(S_{ij})$ 代替，得到两步方差改进估计

$$\tilde{\beta}_1(T)=(x_1^{\tau}x_1)^{-1}x_1^{\tau}y_1-(S_{12}/S_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2y_2,$$

$\tilde{\beta}_1$ 和 $\tilde{\beta}_1(T)$ 的协方差阵分别为

$$\mathrm{cov}\tilde{\beta}_1=\sigma_{11}(x_1^{\tau}x_1)^{-1}-(\sigma_{12}^2/\sigma_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2X_1(x_1^{\tau}x_1)^{-1},$$

$$\mathrm{cov}(\tilde{\beta}_1(T))=\sigma_{11}(x_1^{\tau}x_1)^{-1}-[2\sigma_{12}E(S_{12}/S_{22})-\sigma_{22}E(S_{12}/S_{22})^2](x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2X_1(x_1^{\tau}x_1)^{-1},$$

由于 $\tilde{\beta}_1$ 和 $\tilde{\beta}_1(T)$ 都是无偏估计，故其均方误差为

$$MSE(\tilde{\beta}_1)=\sigma_{11}tr[(x_1^{\tau}x)^{-1}-(\sigma_{12}^2/\sigma_{11}\sigma_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2x_2(x_1^{\tau}x_1)^{-1}]$$

$$\geqslant\sigma_{11}(1-\rho_{12}^2)\sum_{i=1}^{p_1}\lambda_i^{(1)-1},$$

$$MSE(\tilde{\beta}_1(T))\geqslant\sigma_{11}[1-(2\sigma_{12}E(S_{12}/S_{22})-\sigma_{22}E(S_{12}/S_{22})^2)]\sum_{i=1}^{p_1}\lambda_i^{(1)-1},$$

其中，利用不等式 $0\leqslant N_2\leqslant I_n$，$\rho_{12}=(\sigma_{12}/\sigma_{11}\sigma_{22})^{1/2}$，$\lambda_1^{(1)}\geqslant\cdots\geqslant\lambda_p^{(1)}$ 是 $x_1^{\tau}x_1$ 的特征根.

但当 x_1 呈病态时，$\lambda_p^{(1)}$ 接近于零，此时 $\tilde{\beta}_1$、$\tilde{\beta}_1(T)$ 就不有认为是 β_1 的良好估计. 这就需要我们对 $\tilde{\beta}_1$ 和 $\tilde{\beta}_1(T)$ 作改进以减少均误差. 本文提出 $\tilde{\beta}_1$ 和 $\tilde{\beta}_1(T)$ 的一种 Stein 改进估计，即

$$\tilde{\beta}_1(k)=(x_1^{\tau}x_1)^{-1}x_1^{\tau}y_1k-(\sigma_{12}/\sigma_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2x_1y_2k,\ (1>k>0),$$

$$\tilde{\beta}_1(k\cdot T)=(x_1^{\tau}x_1)^{-1}x_1^{\tau}y_1k-(S_{12}/S_{22})(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2x_1y_2k,$$

并证明了当 k 满足一定条件可使 $\tilde{\beta}_1(k)$ 比 $\tilde{\beta}_1$ 具有较小的均方误差，$\tilde{\beta}_1(k\cdot T)$ 比 $\tilde{\beta}_1(T)$ 具有较小的均方误差.

1　Stein 改进估计及其性质

定理1　对于回归系统(1)，Stein 改进估计 $\tilde{\beta}_1(k)$ 具有以下性质

(i) $\tilde{\beta}_1(k)$ 是 β_1 的有编压缩估计；

(ii) 当 $1>k>\dfrac{\|\beta_1\|^2}{\sigma_{11}(1-\rho_{12}^2)\sum_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2}$ 时，$MSE(\tilde{\beta}_1(k))<MSE(\tilde{\beta}_1)$.

证明　(i) 由于 $\tilde{\beta}_1(k)=k\tilde{\beta}_1$，$\tilde{\beta}_1$ 是 β_1 的无偏估计，所以有 $E\tilde{\beta}_1(k)=kE\tilde{\beta}_1=k\beta_1$；$\|\tilde{\beta}_1(k)\|^2<k^2\|\tilde{\beta}_1\|^2<\|\tilde{\beta}_1\|^2$ 即(i)得证. 其(ii)

$$\begin{aligned}&MSE(\tilde{\beta}_1)-MSE(\tilde{\beta}_1(k))\\&=tr(\mathrm{cov}(\tilde{\beta}_1))-tr(\mathrm{cov}(\tilde{\beta}(k)))-(E\tilde{\beta}_1(k)-\beta_1)^{\tau}(E\tilde{\beta}_1(k)-\beta_1)\\&=tr(\mathrm{cov}(\tilde{\beta}_1))-k^2tr(\mathrm{cov}(\tilde{\beta}_1)-(E\tilde{\beta}_1(k)-\beta_1)^{\tau}(E\tilde{\beta}_1(k)-\beta_1)\\&=(1-k^2)tr(\mathrm{cov}(\tilde{\beta}_1)-(k-1)^2\|\beta_1\|^2\\&=-[tr(\mathrm{cov}(\tilde{\beta}_1))+\|\beta_1\|^2]k^2+2\|\beta_1\|^2k-\|\beta_1\|^2+tr(\mathrm{cov}(\tilde{\beta}_1))\\&\triangleq M_1(k),\end{aligned}$$

在式 $M'_1(k)$ 中对 k 求导得

$$M'_1(k)=-2[tr(\operatorname{cov}(\tilde{\beta}_1))+\|\beta_1\|^2]k+2\|\beta_1\|^2,$$

所以当 $\frac{\|\beta_1\|^2}{\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1^1))+\|\beta_1\|^2}<k<1$ 时，$M'(k)<0$，即 $M(k)$ 是单调递减. 又因为 $M(1)=0$，故当 $\frac{\|\beta_1\|^2}{tr(\operatorname{cov}(\tilde{\beta}_1))+\|\beta_1\|^2}<k<1$ 时，$M(k)<0$，即 $MSE(\tilde{\beta}_1)>MSE(\tilde{\beta}_1(k))$. 又因为

$$\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1))=\operatorname{tr}[\sigma_{11}(x_1^Tx_1)^{-1}-(\sigma_{12}^2/\sigma_{22})(x_1^Tx_1)^{-1}x_1^TN_2X_1(x_1^Tx_1)^{-1}]$$

$$\geqslant\sigma_{11}(1-\rho_{12}^2)\operatorname{tr}(x_1^Tx_1)^{-1}=\sigma_{11}(1-\rho_{12}^2)\sum_{i=1}^{p_1}\lambda_i^{(1)-1},$$

（这是因为 $(x_1^Tx_1)^{-1}x_1^TN_2x_1(x_1^Tx_1)^{-1}\leqslant(x_1^Tx_1)^{-1}x_1^Tx_1(x_1^Tx_1)^{-1}=(x_1^Tx_1)^{-1}$），所以

$$\frac{\|\beta_1\|^2}{\sigma_{11}(1-\rho_{12}^2)\sum\limits_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2}>\frac{\|\beta_1\|^2}{\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1^1))+\|\beta_1\|^2},$$

因此当 $(\|\beta_1\|^2)/(\sigma_{11}(1-\rho_{12}^2)\sum\limits_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2)<k<1$ 时，必有 $MSE(\tilde{\beta}_1)>MSE(\tilde{\beta}_1(k))$，即(ii)得证.

在实际应用中，由于 $\sum$ 是未知的，这时我们可用二步 Stein 改进估计

$$\tilde{\beta}_1(k\cdot T)=(x_1^Tx_1)^{-1}x_1^Ty_1k-(S_{12}/S_{22})(x_1^Tx_1)^{-1}x_1^TN_2x_1y_2k.$$

下面我们讨论 $\tilde{\beta}_1(k\cdot T)$ 的优良性质，利用[2]可证明下列引理

引理　设对于回归系统(1) $\epsilon=(\epsilon_i,\epsilon_j)$ 的行向量相互独立且服从 $N(0,\sum)$ 分布，则有

(i) $\tilde{\beta}_1(T)$ 是 β_1 的无偏估计；

(ii) $x_1^Te_1, x_1^TN_ie_i, x_1^TP_2e_1$ 与所有 $S_{ij}(i,j=1,2)$ 相互独立；

(iii) $E(S_{12}/S_{22})=(\sigma_{12}/\sigma_{22})$，$E(S_{12}/S_{22})^2=(\sigma_{12}/\sigma_{22})\rho_{12}^2+(\sigma_{11}/\sigma_{22})(1-\rho_{12}^2)1/(n-r-2)$. 其中：$S_{ij}=\hat{e}_i'\hat{e}_j/(n-r)$，$\hat{e}_i=(I-\tilde{X}(\tilde{X}'\tilde{X})\tilde{X}^T)y_i$，$r=R(\tilde{X})$，$\tilde{X}=(X_1,X_2)$，$A^-$ 表示 A 的广义逆.

对于 Stein 改进估计 $\tilde{\beta}_1(k\cdot T)$ 有如下性质

定理 2　在引理的条件下，则有

(i) $\tilde{\beta}_1(k\cdot T)$ 是 β_1 的有偏压缩估计；

(ii) 当 $1>k>\|\beta_1\|^2/(\sigma_{11}(1-\rho_{12}^2)(\frac{n-r-1}{n-r-2})\sum\limits_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2)$ 时，有 $MSE(\tilde{\beta}_1(k,T))<MSE(\tilde{\beta}_1(T))$.

证明　(i)显然成立. 其(ii)

$$\begin{aligned}&MSE(\tilde{\beta}_1(T))-MSE(\tilde{\beta}_1(k,T))\\&=\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1(T))-\operatorname{tr}(\operatorname{cov}(\tilde{\beta}(k,T)))-(E\tilde{\beta}_1(k,T)-\beta_1)^T(E\tilde{\beta}_1(k,T)-\beta_1)\\&=(1-k^2)\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1(T))-(k\beta_1-\beta_1)^T(k\beta_1-\beta_1)\\&=-[\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1(T)))+\|\beta_1\|^2]k^2+2\|\beta_1\|^2k-\|\beta_1\|^2+tr(\operatorname{cov}(\tilde{\beta}_1(T)))\\&\triangleq M_2(k).\end{aligned}$$

故当 $\frac{\|\beta_1\|^2}{\operatorname{tr}(\operatorname{cov}(\tilde{\beta}(T)_1))+\|\beta_1\|^2}<k<1$ 时 $M_2(k)>0$，即 $MSE(\tilde{\beta}_1(T))>MSE(\tilde{\beta}_1(k,T))$，又因

为

$$\begin{aligned}\operatorname{cov}(\tilde{\beta}_1(T)) &= E((\tilde{\beta}_1(T)-E\tilde{\beta}_1(T))(\tilde{\beta}_1(T)-E\tilde{\beta}_1(T))^{\tau} \\ &= E\{E[(\tilde{\beta}_1(T)-\beta_1)(\tilde{\beta}_1(T)-\beta_1)^{\tau}/s]\} \\ &= E\{\sigma_{11}(x_1^{\tau}x_1)^{-1} - 2(S_{12}/S_{22})\sigma_{12}(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2X_1(x_1^{\tau}x_1)^{-1}) \\ &\quad + (S_{12}/S_{22})^2\sigma_{22}(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2X_1(x_1^{\tau}x_1)^{-1})\} \\ &= \sigma_{11}(x_1^{\tau}x)^{-1} - (2\sigma_{12}E(S_{12}/S_{22}) - \sigma_{22}E(S_{12}/S_{22})^2)(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2X_1(x_1^{\tau}x_1)^{-1}.\end{aligned}$$

而因为

$$\begin{aligned}&2\sigma_{12}E(S_{12}/S_{22}) - \sigma_{22}E(S_{12}/S_{22})^2 \\ &= 2\sigma_{12}(\sigma_{12}/\sigma_{22}) - \sigma_{22}[(\sigma_{12}/\sigma_{22})\rho_{12}^2 + (\sigma_{11}/\sigma_{22})(1-\rho_{12}^2)/(n-r-2)] \\ &= (\sigma_{12}^2/\sigma_{22}) + (\sigma_{11}\sigma_{22}/\sigma_{22})(1-\rho_{12}^2)/(n-r-2) > 0 \\ &(x_1^{\tau}x_1)^{-1}x_1^{\tau}N_2x_1(x_1^{\tau}x_1)^{-1} \leqslant (x_1^{\tau}x_1)^{-1}x_1^{\tau}I_nx_1(x_1^{\tau}x_1)^{-1} \\ &= (x_1^{\tau}x_1)^{-1}(\text{因 } N_2 = I_n - x_2(x_2^{\tau}x_2)^{-1}x_1^{\tau} \leqslant I_n),\end{aligned}$$

所以可得

$$\begin{aligned}&\operatorname{tr}[\operatorname{cov}(\tilde{\beta}_1(T))] \\ &= \operatorname{tr}[\sigma_{11}(x_1^{\tau}x)^{-1} - (\sigma_{12}^2/\sigma_{22} + (\sigma_{11}\sigma_{22}(1-\rho_{12}^2)/\sigma_{22}(n-r-2))(x_1^{\tau}x)^{-1}x_1^{\tau}N_2x_1(x_1^{\tau}x_1)^{-1}] \\ &\geqslant \operatorname{tr}[\sigma_{11}(x_1^{\tau}x)^{-1} - (\sigma_{12}^2/\sigma_{22} + \sigma_{11}\sigma_{12}(1-\rho_{12}^2)/\sigma_{22}(n-r-2)(x_1^{\tau}x)^{-1}] \\ &= [\sigma_{11}(1-\sigma_{12}^2/(\sigma_{11}\sigma_{22}) - \sigma_{11}(1-\rho_{12}^2)/(n-r-2)]\operatorname{tr}(x_1^{\tau}x_1)^{-1} \\ &= \sigma_{11}(1-\rho_{12}^2)(n-r-1)/(n-r-2)\sum_{i=1}^{p_1}\lambda_i^{(1)-1},\end{aligned}$$

于是

$$\frac{\|\beta_1\|^2}{\operatorname{tr}(\operatorname{cov}(\tilde{\beta}_1(T)) + \|\beta_1\|^2} \leqslant \frac{\|\beta_1\|^2}{\sigma_{11}(1-\rho_{12}^2)(\frac{n-r-1}{n-r-2})\sum_{i=1}^{p_1}\lambda_i^{(1)-1} + \|\beta_1\|^2},$$

故当 $\dfrac{\|\beta_1\|^2}{\sigma_{11}(1-\rho_{12}^2)(\frac{n-r-1}{n-r-2})\sum_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2} < k < 1$ 时，必有 $MSE(\tilde{\beta}_1(k,T)) < MSE(\tilde{\beta}_1(T))$.

综合定理 1 与定理 2 得到

推论 对回归系统(1)，在均方程差意义下，当 $(\|\beta_1\|^2)/(\sigma_{11}(1-\rho_{12}^2)^2\sum_{i=1}^{p_1}\lambda_i^{(1)-1}+\|\beta_1\|^2)<k<1$ 时，有偏 Stein 改进估计 $\tilde{\beta}_1(k)$ 和 $\tilde{\beta}_1(T,k)$ 分别优于无偏估计 $\tilde{\beta}_1$ 和 $\tilde{\beta}_1(T)$.

此外对于线性回归系统(1)的第二个方程的系数 β_2，当设计阵 X_2 呈病态时，同样可构造相应的 Stein 改进估计.

参 考 文 献

〔1〕 Zellner, A., *J. Amer. Statist. Asso.*, 57(1962), 348－368.

〔2〕 林春土，两个回归方程系数的两步估计及其有限样本性质，应用数学之报，8，3(1985)，277－321.

〔3〕 刘爱义、王松桂，两个半相依回归系统参数的有编估计，应用概率统计，7，3(1991)，263－274.

〔4〕 陈希孺,近代回归分析,安徽人民出版社,(1987).

Stein Estimate of the Parameters in Two Semingly Unrelated Regression Systems

Chen Jianwei

(Department of Management Information Science)

Abstract One kind of least square estimators of the parameters in two seemingly unrelated regression systems have been improved by Zellner, et al (1). These improved estimators are all unbiased estimators with a lot of advantages, however, they are no longer considered to be well estimators in case the design matrix is ill-conditioning. In this paper, the author proposes one kind of biased Stein estimators; and demonstrates the advantages of Stein estimatiors in case the design matrix is ill-conditioning.

Key words semingly unrelated regression systems, two step estimate, covariance improved, estimation