应用概率统计 第二十五卷
第六期 2009年12月

Chinese Journal of Applied Probability
and Statistics Vol.25 No.6 Dec. 2009

两个半相依模型回归系数的改进估计 *

马铁丰 王松桂

(西南财经大学统计学院, 成都, 610074) (北京工业大学应用数理学院, 北京, 100022)

摘 要

对于两个半相依回归系统的未知回归系数, 本文首先借鉴文献中给出的两步协方差改进估计的方法给出两种两步协方差改进估计序列, 并给出其与两步估计等价的条件和均方误差意义下的优良性; 其次, 我们对文献中给出的一种两步估计作简单改进, 使得改进后的估计在更大的参数空间内优于最小二乘估计. 再次, 本文另辟蹊径, 构造了一种新的估计, 同样地，此估计也具有更好的小样本性质. 本文最后一节讨论了Pitman准则下两步估计的优良性.

关键词: 半相依模型, 协方差改进估计, 两步估计, Pitman准则.

学科分类号: O212.1.

§1. 引 言

考虑含有两个相依线性回归方程的线性回归系统:

$$y_i = X_i\beta_i + \varepsilon_i, \qquad i = 1, 2, \tag{1.1}$$

这里y_i为$n\times 1$观测向量, X_i为$n\times p_i$列满秩设计阵, β_i为$p_i\times 1$未知回归系数, ε_i是随机误差向量. 假定$(\varepsilon_1, \varepsilon_2)$的行独立同分布, 每一行服从二维正态分布$N_2(0,\Sigma)$, 其中$\Sigma = (\sigma_{ij})$为2阶未知正定阵, 并且$\sigma_{12}\neq 0$.

记$y = (y_1', y_2')'$, $X = \text{diag}(X_1, X_2)$, $\beta = (\beta_1', \beta_2')'$, $\varepsilon = (\varepsilon_1', \varepsilon_2')'$. 易知$\varepsilon$服从$2n$维正态分布$N(0, \Sigma\otimes I)$, 其中符号"$\otimes$"表示矩阵的Kronecker乘积, I为n阶单位阵. 在Σ已知时, β的Gauss-Markov估计为

$$\beta^* = (X'(\Sigma^{-1}\otimes I)X)^{-1}X'(\Sigma^{-1}\otimes I)y. \tag{1.2}$$

在实际应用中, Σ往往是未知的, 这时较为简单的一种处理就是忽略回归误差向量间的相关性的最小二乘估计:

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 \end{pmatrix}. \tag{1.3}$$

---

*西南财经大学"211工程三期"统计学重点学科建设项目资助.

本文2006年11月3日收到.

显然此估计没有充分利用样本信息, 因此需寻求更为有效的估计. 另外一种兼顾误差向量间的相关性和可行性的β_i的估计就是Zellner[1, 2]提出的两步估计

$$\widetilde{\beta} = (X'(S^{-1} \otimes I)X)^{-1}X'(S^{-1} \otimes I)y, \tag{1.4}$$

其中$S = [1/(n-r)] \cdot \widetilde{Y}'\widetilde{N}\widetilde{Y}$, 其中$\widetilde{Y} = (y_1, y_2)$, $\widetilde{N} = I - \widetilde{X}(\widetilde{X}'\widetilde{X})^{-}\widetilde{X}'$, $\widetilde{X} = (X_1, X_2)$, $r = \mathrm{rk}(\widetilde{X})$.

不失一般性, 本文只考虑β_1的估计, β_2的情况类似. 由于两步估计是非线性的, 因此讨论其统计性质比较困难, 许多学者寻求一些其他的方法, Revankar[3, 4]给出一种两步协方差改进估计

$$\widehat{\beta}_1(S) = \widehat{\beta}_1 - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2. \tag{1.5}$$

此估计也是王松桂[5]提出的两步协方差改进估计的$m=2$时的特例. Revankar[3]给出了β_1的两步协方差改进估计$\widehat{\beta}_1(S)$的协方差阵

$$\mathrm{Cov}\,(\widehat{\beta}_1(S)) = \sigma_{11}(X_1'X_1)^{-1} - \sigma_{11}\Big[\rho^2 - \frac{1-\rho^2}{n-r-2}\Big](X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}, \tag{1.6}$$

这里$\rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$.

显然当

$$\rho^2 > \frac{1}{n-r-1} \tag{1.7}$$

时, $\widehat{\beta}_1(S)$优于最小二乘估计$\widehat{\beta}_1$.

在下面的行文中记

$$P_i = X_i(X_i'X_i)^{-1}X_i', \qquad N_i = I - P_i. \tag{1.8}$$

Liu[6]在此基础上给出了一种改进形式

$$\widetilde{\beta}_1(S) = \widehat{\beta}_1 - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{s_{12}^2}{s_{11}s_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1. \tag{1.9}$$

本文第二节在(1.5)和(1.9)的基础上给出一种协方差改进估计序列, 并给出其与两步估计(1.4)等价的条件. 第三节给出了(1.9)的一种改进形式, 新估计不仅与(1.9)同样在均方误差意义下同时优于(1.5)和最小二乘估计, 而且其具有这样优良性的条件与(1.9)相比要更弱. 在第四节, 我们另辟蹊径构造出一种新的估计, 其优于最小二乘估计(1.3)的条件要弱于(1.5)优于最小二乘估计的条件. 本文最后一节讨论了在Pitman意义下两步协方差改进估计(1.5)的优良性, 给出了优于最小二乘估计的条件.

§2. 一种两步协方差改进估计序列

首先给出所需引理.

引理 2.1[7] 设T_1, T_2分别为k_1, k_2维统计量, 且$\mathsf{E}(T_1)=\eta$, $\mathsf{E}(T_2)=0$,

$$\mathsf{Cov}\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}. \tag{2.1}$$

这里V为已知正定阵, 则在估计类$\Omega=\{T=A_1T_1+A_2T_2,\ A_1$和$A_2$分别为$k_1\times k_1$, $k_1\times k_2$非随机阵$\}$中, η的BLUE估计为:

$$\eta^* = T_1 - V_{12}V_{22}^{-1}T_2. \tag{2.2}$$

为行文方便, 记β_1的两步协方差改进估计为

$$\widetilde{\beta}_1^{(2)}(S) = \widehat{\beta}_1(S) = \widehat{\beta}_1 - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2. \tag{2.3}$$

类似的, 记

$$\widetilde{\beta}_1^{(3)}(S) = \widetilde{\beta}_1(S) = \widehat{\beta}_1 - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{s_{12}^2}{s_{11}s_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1. \tag{2.4}$$

根据文献中讨论半相依模型参数估计问题的方法[8–10], 我们首先在Σ已知前提下讨论改进方法

$$\widetilde{\beta}_1^{(3)}(\Sigma) = \widehat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1 \tag{2.5}$$

的两步估计.

取$T_1=\widetilde{\beta}_1^{(3)}(\Sigma)$, $T_2=N_2N_1N_2y_2$, 显然有$\mathsf{E}(T_2)=0$, 则由T_1、T_2的相关性应用引理2.1得到一改进估计

$$\begin{aligned}\widetilde{\beta}_1^{(4)}(\Sigma) &= \widehat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1 \\ &\quad - \frac{\sigma_{12}^3}{\sigma_{11}\sigma_{22}^2}(X_1'X_1)^{-1}X_1'N_2N_1N_2y_2.\end{aligned} \tag{2.6}$$

取$T_1=\widetilde{\beta}_1^{(4)}(\Sigma)$, $T_2=N_2N_1N_2N_1y_1$, 再次应用引理2.1得到改进估计

$$\begin{aligned}\widetilde{\beta}_1^{(5)}(\Sigma) &= \widehat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1 \\ &\quad - \frac{\sigma_{12}^3}{\sigma_{11}\sigma_{22}^2}(X_1'X_1)^{-1}X_1'N_2N_1N_2y_2 + \frac{\sigma_{12}^4}{\sigma_{11}^2\sigma_{22}^2}(X_1'X_1)^{-1}X_1'N_2N_1N_2N_1y_1.\end{aligned} \tag{2.7}$$

以此类推, 我们得到k步改进估计.

(a) 当k为奇数时

$$\begin{aligned}\widetilde{\beta}_1^{(k)}(\Sigma) &= \widehat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1 \\ &\quad + \cdots - \frac{\sigma_{12}^{k-2}}{\sigma_{11}^{(k-3)/2}\sigma_{22}^{(k-1)/2}}(X_1'X_1)^{-1}X_1'N_2N_1\cdots N_2y_2 \\ &\quad + \frac{\sigma_{12}^{k-1}}{\sigma_{11}^{(k-1)/2}\sigma_{22}^{(k-1)/2}}(X_1'X_1)^{-1}X_1'N_2N_1\cdots N_2N_1y_1. \end{aligned} \tag{2.8}$$

(b) 当k为偶数时

$$\begin{aligned}\widetilde{\beta}_1^{(k)}(\Sigma) &= \widehat{\beta}_1 - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2y_2 + \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1 \\ &\quad + \cdots - \frac{\sigma_{12}^{k-2}}{\sigma_{11}^{(k-1)/2}\sigma_{22}^{(k-1)/2}}(X_1'X_1)^{-1}X_1'N_2N_1\cdots N_2N_1y_1 \\ &\quad + \frac{\sigma_{12}^{k-1}}{\sigma_{11}^{k/2-1}\sigma_{22}^{k/2}}(X_1'X_1)^{-1}X_1'N_2N_1\cdots N_1N_2y_2. \end{aligned} \tag{2.9}$$

在实际应用中, 由于Σ往往是未知的, 此时我们用Σ的估计S代替上面的Σ得到可行估计$\widetilde{\beta}_1^{(k)}(S)$.

定理 2.1 (a) 当k为大于2的偶数时, $\widetilde{\beta}_1^{(k)}(S) = \widetilde{\beta}_1$当且仅当

$$P_1P_2P_1(N_2N_1)^{(k-2)/2} = 0; \tag{2.10}$$

(b) 当k为大于3的奇数时, $\widetilde{\beta}_1^{(k)}(S) = \widetilde{\beta}_1$当且仅当

$$P_1P_2P_1(N_2N_1)^{(k-3)/2}N_2 = 0, \tag{2.11}$$

其中$P_1 = I - N_1$, $P_2 = I - N_2$.

证明: 仅以k为偶数的情况进行证明, k为奇数的情况类似可得.

两步估计$\widetilde{\beta}_1$和$\widetilde{\beta}_2$满足下面的等式

$$X'(S^{-1}\otimes I)X\begin{pmatrix}\widetilde{\beta}_1\\ \widetilde{\beta}_2\end{pmatrix} = X'(S^{-1}\otimes I)y, \tag{2.12}$$

进一步展开得到

$$\begin{pmatrix} s_{22}X_1'X_1 & -s_{21}X_1'X_2 \\ -s_{12}X_2'X_1 & s_{11}X_2'X_2 \end{pmatrix}\begin{pmatrix}\widetilde{\beta}_1\\ \widetilde{\beta}_2\end{pmatrix} = \begin{pmatrix} s_{22}X_1' & -s_{21}X_1' \\ -s_{12}X_2' & s_{11}X_2' \end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}. \tag{2.13}$$

上式等价于

$$\begin{aligned}\Big(s_{22}X_1'X_1 - \frac{s_{21}^2}{s_{11}}X_1'P_2X_1\Big)\widetilde{\beta}_1 &= s_{22}X_1'y_1 - s_{12}X_1'N_2y_2 + \frac{s_{12}^2}{s_{11}}X_1'N_2N_1y_1 \\ &\quad - \frac{s_{12}^2}{s_{11}}X_1'y_1 + \frac{s_{12}^2}{s_{11}}X_1'N_2P_1y_1, \end{aligned} \tag{2.14}$$

因此若$\widetilde{\beta}_1^{(k)}(S)=\widetilde{\beta}_1$当且仅当$\widetilde{\beta}_1^{(k)}(S)$是(2.14)中$\widetilde{\beta}_1$的解.

将$\widetilde{\beta}_1^{(k)}(S)$代入上式得到

$$\begin{aligned}
&\Big(s_{22}X_1'X_1-\frac{s_{21}^2}{s_{11}}X_1'P_2X_1\Big)\widetilde{\beta}_1^{(k)}(S)\\
=\;& s_{22}X_1'y_1-s_{12}X_1'N_2y_2+\frac{s_{12}^2}{s_{11}}X_1'N_2N_1y_1-\frac{s_{12}^3}{s_{11}s_{22}}X_1'N_2N_1N_2y_2+\cdots\\
&+\frac{s_{12}^{k-2}}{s_{11}^{k/2-1}s_{22}^{k/2-2}}X_1'N_2N_1\cdots N_2N_1y_1-\frac{s_{12}^{k-1}}{s_{11}^{k/2-1}s_{22}^{k/2-1}}X_1'N_2N_1\cdots N_1N_2y_2\\
&-\frac{s_{12}^2}{s_{11}}X_1'P_2P_1y_1+\cdots-\frac{s_{12}^k}{s_{11}^{k/2}s_{22}^{k/2-1}}X_1'P_2P_1N_2N_1\cdots N_2N_1y_1\\
&+\frac{s_{12}^{k+1}}{s_{11}^{k/2}s_{22}^{k/2}}X_1'P_2P_1N_2N_1\cdots N_1N_2y_2\\
=\;& s_{22}X_1'y_1-s_{12}X_1'N_2y_2+\frac{s_{12}^2}{s_{11}}X_1'N_2N_1y_1-\frac{s_{12}^2}{s_{11}}X_1'y_1+\frac{s_{12}^2}{s_{11}}X_1'N_2P_1y_1,
\end{aligned}\tag{2.15}$$

经整理, (2.15)等价于

$$\begin{aligned}
&-\frac{s_{12}^3}{s_{11}s_{22}}X_1'(N_2N_1-P_2P_1)N_2y_2+\frac{s_{12}^4}{s_{11}^2s_{22}}X_1'(N_2N_1-P_2P_1)N_2N_1y_1+\cdots\\
&-\frac{s_{12}^k}{s_{11}^{k/2}s_{22}^{k/2-1}}X_1'P_2P_1N_2N_1\cdots N_2N_1y_1+\frac{s_{12}^{k+1}}{s_{11}^{k/2}s_{22}^{k/2}}X_1'P_2P_1N_2N_1\cdots N_1N_2y_2=0.
\end{aligned}\tag{2.16}$$

因为$N_2N_1-P_2P_1=N_1-P_2N_1-P_2P_1=N_1-P_2$, 所以有$X_1'(N_1-P_2)N_2=0$, 因此上式化简为

$$X_1'P_2P_1N_2N_1\cdots N_2N_1y_1+\frac{s_{12}}{s_{22}}X_1'P_2P_1N_2N_1\cdots N_1N_2y_2=0.\tag{2.17}$$

为表述方便将上式化简为

$$X_1'P_2P_1(N_2N_1)^{(k-2)/2}\Big(y_1+\frac{s_{12}}{s_{22}}N_2y_2\Big)=0.\tag{2.18}$$

因为$y_1+(s_{12}/s_{22})\cdot N_2y_2$为随机向量, 其均值为0, 协方差阵可逆, 于是结合(2.18)式有

$$X_1'P_2P_1(N_2N_1)^{(k-2)/2}=0,\tag{2.19}$$

也就是

$$P_1P_2P_1(N_2N_1)^{(k-2)/2}=0.\tag{2.20}$$

定理得证. □

§3. 一种改进估计

文献[6]给出了估计(1.9)的协方差阵:

$$\begin{aligned}&\mathsf{Cov}\,(\widetilde{\beta}_1(S))\\ =\ &\mathsf{Cov}\,(\widehat{\beta}_1)-\sigma_{11}(X_1'X_1)^{-1}X_1'N_2\Big(\Big(\rho^2-\frac{1-\rho^2}{n-r-2}\Big)I+\theta N_1\Big)N_2X_1(X_1'X_1)^{-1},\end{aligned}\tag{3.1}$$

这里

$$\theta=2\frac{\sigma_{12}}{\sigma_{11}}\mathsf{E}\Big[\frac{s_{12}^3}{s_{11}s_{22}^2}\Big]-\mathsf{E}\Big[\frac{s_{12}^4}{s_{11}^2s_{22}^2}\Big],$$

进而给出了其优于最小二乘估计的条件:

$$n\geq r+\frac{3}{\rho^{8/3}}+4,\tag{3.2}$$

也就是

$$\rho^4\geq\frac{3\sqrt{3}}{(n-r-4)^{3/2}}.\tag{3.3}$$

显然, 对于$n-r>4$, 这个条件比(1.7)要强, 因此这个改进虽然在均方误差意义下可以优于(1.5), 但其所需的条件要相对苛刻一些, 所以有必要进行一些改进, 使其在均方误差和适用范围两个方面都得到提高.

本节我们给出一种改进形式:

$$\widetilde{\beta}_1^e(S)=\widehat{\beta}_1-k_1\frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2+k_2\frac{s_{12}^2}{s_{11}s_{22}}(X_1'X_1)^{-1}X_1'N_2N_1y_1,\tag{3.4}$$

其中$k_1>0$, $k_2>0$.

首先给出后文所需的引理.

引理 3.1 $(n-r)S$服从自由度为$n-r$协方差阵为Σ的Wishart分布, 即

$$(n-r)S\sim W_2(n-r,\Sigma),\tag{3.5}$$

并且s_{ij}与$X_1'y_1$、$X_1'N_2y_2$和$X_1'N_2N_1y_1$相互独立.

证明参见Muirhead[8].

引理 3.2 对于2×2矩阵S有

$$\frac{s_{12}}{s_{11}}=\Big(\frac{\sigma_{22}}{\sigma_{11}}\Big)^{1/2}\Big(\rho+\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_1\Big),\tag{3.6}$$

$$\frac{s_{12}}{s_{22}}=\Big(\frac{\sigma_{11}}{\sigma_{22}}\Big)^{1/2}\Big(\rho+\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_2\Big),\tag{3.7}$$

这里t_i $(i=1,2)$为自由度为$n-r$的t分布随机变量.

证明参见Liu[6].

定理 3.1　当$\rho^4 > \max\{16k_2/[(2k_1-k_2)(n-r-4)], k_1/[k_1+(2-k_1)(n-r-2)]\}$时, $\widetilde{\beta}_1^e(S)$优于$\widehat{\beta}_1$.

证明:　根据引理3.1给出的统计量的独立性容易得到$\widetilde{\beta}_1^e(S)$的协方差阵

$$\begin{aligned}\mathrm{Cov}\,(\widetilde{\beta}_1^e(S)) &= \mathrm{Cov}\,(\widehat{\beta}_1) + k_1^2\sigma_{22}\mathrm{E}\Big(\frac{s_{12}^2}{s_{22}^2}\Big)(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\\ &\quad\cdot k_2^2\sigma_{11}\mathrm{E}\Big(\frac{s_{12}^4}{s_{11}^2s_{22}^2}\Big)(X_1'X_1)^{-1}X_1'N_2N_1N_2X_1(X_1'X_1)^{-1}\\ &\quad - 2k_1\sigma_{12}\mathrm{E}\Big(\frac{s_{12}}{s_{22}}\Big)(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\\ &\quad - 2k_1k_2\sigma_{12}\mathrm{E}\Big(\frac{s_{12}^3}{s_{11}s_{22}^2}\Big)(X_1'X_1)^{-1}X_1'N_2N_1N_2X_1(X_1'X_1)^{-1}. \end{aligned}\tag{3.8}$$

应用引理3.2有

$$\mathrm{E}\Big[\frac{s_{12}}{s_{22}}\Big] = \frac{\sigma_{12}}{\sigma_{22}},\tag{3.9}$$

$$\mathrm{E}\Big[\Big(\frac{s_{12}}{s_{22}}\Big)^2\Big] = \frac{\sigma_{12}}{\sigma_{22}}\rho^2 + \frac{\sigma_{12}}{\sigma_{22}}(1-\rho^2)\frac{1}{n-r-2},\tag{3.10}$$

代入(3.8)有

$$\begin{aligned}&\mathrm{Cov}\,(\widetilde{\beta}_1^e(S))\\ =\ &\mathrm{Cov}\,(\widehat{\beta}_1) + \sigma_{11}\Big((2k_1-k_1^2)\rho^2 - k_1^2\frac{1-\rho^2}{n-r-2}\Big)(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\\ &-\sigma_{11}\Big(2k_1k_2\frac{\sigma_{12}}{\sigma_{11}}\mathrm{E}\Big(\frac{s_{12}^3}{s_{11}s_{22}^2}\Big) - k_2^2\mathrm{E}\Big(\frac{s_{12}^4}{s_{11}^2s_{22}^2}\Big)\Big)(X_1'X_1)^{-1}X_1'N_2N_1N_2X_1(X_1'X_1)^{-1}.\end{aligned}\tag{3.11}$$

因此, $\widetilde{\beta}_1^e(S)$优于$\widehat{\beta}_1$的一个充分条件为

$$\begin{cases}(2-k_1)\rho^2 - k_1\dfrac{1-\rho^2}{n-r-2} > 0;\\ 2k_1\dfrac{\sigma_{12}}{\sigma_{11}}\mathrm{E}\dfrac{s_{12}^3}{s_{11}s_{22}^2} - k_2\mathrm{E}\dfrac{s_{12}^4}{s_{11}^2s_{22}^2} > 0.\end{cases}\tag{3.12}$$

显然, (3.12)中第一式等价于

$$\rho^2 > \frac{k_1}{k_1+(2-k_1)(n-r-2)}.\tag{3.13}$$

下面我们讨论(3.12)的第二式, 首先应用引理3.2有

$$\begin{aligned}\frac{s_{12}^3}{s_{11}s_{22}^2} &= \Big(\frac{\sigma_{11}}{\sigma_{22}}\Big)^{1/2}\Big[\rho^3 + \rho^2\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_1 + 2\rho^2\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_2\\ &\quad + 2\rho\frac{1-\rho^2}{n-r}t_1t_2 + \rho\frac{1-\rho^2}{n-r}t_2^2 + \Big(\frac{1-\rho^2}{n-r}\Big)^{3/2}t_1t_2^2\Big],\end{aligned}\tag{3.14}$$

$$\begin{aligned}\frac{s_{12}^4}{s_{11}^2 s_{22}^2} &= \rho^4 + 2\rho^3\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_1 + 2\rho^3\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_2 + \rho^2\frac{1-\rho^2}{n-r}t_1^2 \\ &+ \rho^2\frac{1-\rho^2}{n-r}t_2^2 + 4\rho^2\frac{1-\rho^2}{n-r}t_1t_2 + 2\rho\Big(\frac{1-\rho^2}{n-r}\Big)^{3/2}t_1t_2^2 \\ &+ 2\rho\Big(\frac{1-\rho^2}{n-r}\Big)^{3/2}t_2t_1^2 + \Big(\frac{1-\rho^2}{n-r}\Big)^2 t_1^2t_2^2, \end{aligned} \tag{3.15}$$

同时有

$$\mathsf{E}t_1^2 = \mathsf{E}t_2^2 = \frac{n-r}{n-r-2}. \tag{3.16}$$

再应用Cauchy-Schwarz不等式有

$$\mathsf{E}t_1t_2 \le (\mathsf{E}t_1^2\mathsf{E}t_2^2)^{1/2} = \frac{n-r}{n-r-2}, \tag{3.17}$$

$$\mathsf{E}t_1t_2^2 \le (\mathsf{E}t_1^2\mathsf{E}t_2^4)^{1/2} = \frac{\sqrt{3}(n-r)^{3/2}}{(n-r-2)\sqrt{n-r-4}}, \tag{3.18}$$

$$\mathsf{E}t_2t_1^2 \le (\mathsf{E}t_2^2\mathsf{E}t_1^4)^{1/2} = \frac{\sqrt{3}(n-r)^{3/2}}{(n-r-2)\sqrt{n-r-4}}, \tag{3.19}$$

$$\mathsf{E}t_1^2t_2^2 \le \frac{3(n-r)^2}{(n-r-2)(n-r-4)}. \tag{3.20}$$

综合考虑(3.14)~(3.20), (3.12)的第二式成立的充分条件为

$$\begin{aligned}&2k_1\Big(\rho^4 + \rho^2\frac{1-\rho^2}{n-r-2}\Big) \\ > \ & k_2\Big[\rho^4 + 6\frac{\rho^2(1-\rho^2)}{n-r-2} + 4\sqrt{3}\frac{\rho(1-\rho^2)^{3/2}}{(n-r-2)\sqrt{n-r-4}} + 3\frac{(1-\rho^2)^2}{(n-r-2)(n-r-4)}\Big]. \end{aligned} \tag{3.21}$$

进一步化简, 上式成立的一个充分条件为

$$\rho^4 > \frac{16k_2}{(2k_1-k_2)(n-r-4)}, \tag{3.22}$$

综合考虑(3.13)和(3.22)有, 当$\rho^4 > \max\{16k_2/[(2k_1-k_2)(n-r-4)], k_1/[k_1+(2-k_1)(n-r-2)]\}$时, $\widetilde{\beta}_1^e(S)$优于$\widehat{\beta}_1$.

定理证毕. □

定理3.1并没有考虑k_1, k_2的取值问题, 下面我们进一步讨论k_1, k_2的取值问题.

首先, 取$k_1=1$, 显然(3.13)变为

$$\rho^2 > \frac{1}{n-r-1}, \tag{3.23}$$

再令$16k_2/[(2k_1-k_2)(n-r-4)] = 1/(n-r-1)^2$, 则$k_2 = 2(n-r-4)/[16(n-r-1)^2 + n-r-4]$, 此时结合定理(3.1)有$\widetilde{\beta}_1^e(S)$优于$\widehat{\beta}_1$的充分条件为

$$\rho^2 > \frac{1}{n-r-1}. \tag{3.24}$$

显然, 此式与(1.7)相同, 也就是说, 此时我们给出的$\widetilde{\beta}_1^e(S)$优于$\widehat{\beta}_1$的条件与$\widehat{\beta}_1(S)$优于$\widehat{\beta}_1$的条件相同, 因此也就弱于条件(3.3). 当然, 这只是k_1, k_2的一种取值, 我们还可以通过调整k_1, k_2的取值, 给出更弱的可行(3.4)优于$\widehat{\beta}_1$的条件.

至此, 我们得出结论: 我们给出的改进可行估计$\widetilde{\beta}_1^e(S)$与$\widetilde{\beta}_1(S)$相比, 不仅同样地在均方误差意义下优于$\widehat{\beta}_1(S)$, 而且此优良性的存在条件可以更弱.

§4. 一种新估计

本节我们给出一种新估计:

$$\widetilde{\beta}_1^b = (X_1'(I+tN_2)X_1)^{-1}\Big(X_1'(I+tN_2)y_1 - k\frac{s_{12}}{s_{22}}X_1'N_2y_2\Big), \tag{4.1}$$

其中$t>0$, $k>0$.

定理 4.1 当$2(1+t)>[1-1/(n-r-2)]k$, $\rho^2>[t^2+t+k^2/(n-r-2)]/[2k(1+t)-k^2+k^2/(n-r-2)]$时, $\widetilde{\beta}_1^b$为比$\widehat{\beta}_1$更有效的估计.

证明: 显然有$\mathsf{E}(\widetilde{\beta}_1^b)=0$, 因此要想比较$\widetilde{\beta}_1^b$和$\widehat{\beta}_1$只须比较两者的协方差阵.

首先给出$\widehat{\beta}_1$的协方差阵

$$\mathsf{Cov}\,(\widehat{\beta}_1) = (X_1'X_1)^{-1}. \tag{4.2}$$

我们又注意到

$$\mathsf{E}\left[\frac{s_{12}}{s_{22}}\right] = \frac{\sigma_{12}}{\sigma_{22}}, \tag{4.3}$$

$$\mathsf{E}\left[\left(\frac{s_{12}}{s_{22}}\right)^2\right] = \frac{\sigma_{12}}{\sigma_{22}}\rho^2 + \frac{\sigma_{12}}{\sigma_{22}}(1-\rho^2)\frac{1}{n-r-2}. \tag{4.4}$$

结合引理3.1立得

$$\begin{aligned}\mathsf{Cov}\,(\widetilde{\beta}_1^b) &= (X_1'(I+tN_2)X_1)^{-1}\Big[\sigma_{11}X_1'(I+2tN_2+t^2N_2)X_1\\ &\quad +k^2\sigma_{22}\Big(\frac{\sigma_{11}}{\sigma_{22}}\rho^2+\frac{\sigma_{11}}{\sigma_{22}}(1-\rho^2)\frac{1}{n-r-2}\Big)X_1'N_2X_1 - 2k\frac{\sigma_{12}^2}{\sigma_{22}}(1+t)X_1'N_2X_1\Big]\\ &\quad \cdot(X_1'(I+tN_2)X_1)^{-1}\\ &= (X_1'(I+tN_2)X_1)^{-1}\Big[\sigma_{11}X_1'(I+tN_2)X_1\\ &\quad +\sigma_{11}\Big(t^2+t+k^2\Big(\rho^2+\frac{1-\rho^2}{n-r-2}\Big)-2k\rho^2(1+t)\Big)X_1'N_2X_1\Big]\\ &\quad \cdot(X_1'(I+tN_2)X_1)^{-1}\\ &= \sigma_{11}(X_1'(I+tN_2)X_1)^{-1}+\sigma_{11}\Big[t^2+t+k^2\Big(\rho^2+\frac{1-\rho^2}{n-r-2}\Big)-2k\rho^2(1+t)\Big]\\ &\quad \cdot(X_1'(I+tN_2)X_1)^{-1}X_1'N_2X_1(X_1'(I+tN_2)X_1)^{-1},\end{aligned} \tag{4.5}$$

又因为$(X_1'(I+tN_2)X_1)^{-1}<(X_1'X_1)^{-1}$, 所以显然$\widetilde{\beta}_1^b$优于$\widehat{\beta}_1$, 当

$$t^2+t+k^2\Big(\rho^2+\frac{1-\rho^2}{n-r-2}\Big)-2k\rho^2(1+t)<0, \tag{4.6}$$

显然上式成立的充分条件为

$$\begin{cases} 2(1+t)>k^2-\dfrac{k^2}{n-r-2}; \\ \rho^2>\dfrac{t^2+t+k^2/(n-r-2)}{2k(1+t)-k^2+k^2/(n-r-2)}. \end{cases} \tag{4.7}$$

定理证毕. □

注 我们可以通过选取适当的t与k使得$\widetilde{\beta}_1^b$优于$\widehat{\beta}_1$的条件要弱于$\widehat{\beta}_1(S)$优于$\widehat{\beta}_1$的条件.

比较(1.7)与(4.7), 显然若

$$\begin{cases} \dfrac{1}{n-r-1}>\dfrac{t^2+t+k^2/(n-r-2)}{2k(1+t)-k^2+k^2/(n-r-2)}; \\ 2k(1+t)-k^2+\dfrac{k^2}{n-r-2}>0, \end{cases} \tag{4.8}$$

则$\widetilde{\beta}_1^b$优于$\widehat{\beta}_1$的条件(1.7)要弱于$\widehat{\beta}_1^{(2)}(S)$优于$\widehat{\beta}_1$的条件(1.7).

显然(4.8)是容易实现的, 也就是满足这样条件的t、s是存在的.

记$s=n-r-2$, 则(4.8)的第一式等价于

$$\begin{aligned} &\frac{1}{s+1}>\frac{t^2+t+k^2/s}{2k(1+t)-k^2+k^2/s} \\ \Leftrightarrow\quad & k^2-(1+t)k+\frac{(t^2+t)(s+1)}{2}<0 \\ \Leftrightarrow\quad & \frac{1+t-\sqrt{(1+t)^2-2(t^2+t)(s+1)}}{2}<k<\frac{1+t+\sqrt{(1+t)^2-2(t^2+t)(s+1)}}{2}, \end{aligned} \tag{4.9}$$

这里要求$(1+t)^2-2(t^2+t)(s+1)>0$, 即$t>1/(2s+1)$.

另外, (4.8)的第二式等价于$k<2s(1+t)/(s-1)$. 因为对于$t>1/(2s+1)$, $s>1$, 显然有

$$\frac{2s(1+t)}{s-1}>\frac{1+t+\sqrt{(1+t)^2-2(t^2+t)(s+1)}}{2},$$

因此若(4.9)成立, 则(4.8)成立.

因此显然有, 当$t>1/(2s+1)$,

$$\frac{1+t-\sqrt{(1+t)^2-2(t^2+t)(s+1)}}{2}<k<\frac{1+t+\sqrt{(1+t)^2-2(t^2+t)(s+1)}}{2}$$

时, (4.7)要弱于(1.7).

§5. Pitman准则下的两步估计

本节我们从Pitman准则的角度讨论两步协方差改进估计的优良性.

设θ为$p\times 1$参数向量, $\widehat{\theta}$为其一估计. 取损失函数$L(\widehat{\theta},\theta)=(\widehat{\theta}-\theta)'D(\widehat{\theta}-\theta)$, 其中$D$为$p\times p$正定阵, 若$\widehat{\theta}_i$, $i=1,2$为θ的两个估计. 定义Pitman度量

$$P(\theta)=\mathsf{P}\{L(\widehat{\theta}_1,\theta)\le L(\widehat{\theta}_2,\theta)\}. \tag{5.1}$$

当$P(\theta)\ge 1/2$, 且至少对一个θ使得不等式成立, 则称$\widehat{\theta}_1$在Pitman意义下优于$\widehat{\theta}_2$.

文献[12]讨论了Σ已知时协方差改进估计的Pitman优良性, 本节讨论两步协方差改进估计的Pitman优良性.

β_1的两步协方差改进估计为

$$\widehat{\beta}_1(S)=(X_1'X_1)^{-1}X_1'y_1-\frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2=\widehat{\beta}_1-\frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2, \tag{5.2}$$

其中s_{12}、s_{22}的定义同第二节.

下面我们讨论$\widehat{\beta}_1(S)$在Pitman准则下的优良性.

记$F_{m,n}(\alpha)$为自由度为m, n的F分布的α分位点.

定理 5.1 若N_2P_1对称, 则当$\rho^2>F_{1,n-r}(1/2)/[n-r+F_{1,n-r}(1/2)]$时, 在Pitman准则下$\widehat{\beta}_1(S)$优于$\widehat{\beta}_1$.

证明: 要想得到在Pitman准则下$\widehat{\beta}_1(S)$优于$\widehat{\beta}_1$, 只需$\mathsf{P}\{L(\widehat{\beta}_1(S),\beta_1)<L(\widehat{\beta}_1,\beta_1)\}>1/2$.

首先将误差作分解:

$$\varepsilon_2=\frac{\sigma_{12}}{\sigma_{11}}\varepsilon_1+\eta_1, \tag{5.3}$$

其中$\eta_1\sim N(0,\sigma_{22}(1-\rho^2)I)$且与$\varepsilon_1$相互独立.

取$D=X_1'X_1$, 则有

$$\begin{aligned}
&\mathsf{P}\{L(\widehat{\beta}_1(S),\beta_1)<L(\widehat{\beta}_1,\beta_1)\}\\
=&\ \mathsf{P}((\widehat{\beta}_1(S)-\beta_1)'X_1'X_1(\widehat{\beta}_1(S)-\beta_1)<(\widehat{\beta}_1-\beta_1)'X_1'X_1(\widehat{\beta}_1-\beta_1))\\
=&\ \mathsf{P}\Big(\Big[\varepsilon_1'X_1(X_1'X_1)^{-1}-\frac{s_{12}}{s_{22}}\Big(\frac{\sigma_{12}}{\sigma_{11}}\varepsilon_1'+\eta_1'\Big)N_2X_1(X_1'X_1)^{-1}\Big]X_1'X_1\\
&\cdot\Big[(X_1'X_1)^{-1}X_1'\varepsilon_1-\frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2\Big(\frac{\sigma_{12}}{\sigma_{11}}\varepsilon_1+\eta_1\Big)\Big]<\varepsilon_1'X_1(X_1'X_1)^{-1}X_1'\varepsilon_1\Big)\\
=&\ \mathsf{P}\Big(\Big(\frac{s_{12}}{s_{22}}\Big)^2\Big(\frac{\sigma_{12}}{\sigma_{11}}\Big)^2\varepsilon_1'N_2P_1N_2\varepsilon_1+\Big(\frac{s_{12}}{s_{22}}\Big)^2\eta_1'N_2P_1N_2\eta_1-2\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\varepsilon_1'P_1N_2\varepsilon_1\\
&+2\Big(\frac{s_{12}}{s_{22}}\Big)^2\frac{\sigma_{12}}{\sigma_{11}}\varepsilon_1'N_2P_1N_2\eta_1-2\frac{s_{12}}{s_{22}}\varepsilon_1'P_1N_2\eta_1<0\Big)\\
=&\ \mathsf{P}\Big(\Big(\Big(\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\Big)^2-2\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\Big)\varepsilon_1'N_2P_1N_2\varepsilon_1+\Big(\frac{s_{12}}{s_{22}}\Big)^2\eta_1'N_2P_1N_2\eta_1\\
&+2\Big(\Big(\frac{s_{12}}{s_{22}}\Big)^2\frac{\sigma_{12}}{\sigma_{11}}-\frac{s_{12}}{s_{22}}\Big)\varepsilon_1'N_2P_1N_2\eta_1<0\Big),
\end{aligned} \tag{5.4}$$

这里因为N_2P_1对称, 所以有$N_2P_1 = N_2P_1N_2$.

记

$$\eta_2 = \Big(\frac{\sigma_{11}}{\sigma_{22}(1-\rho^2)}\Big)^{1/2}\eta_1,$$

显然有$\eta_2 \sim N(0, \sigma_{11}I)$且与$\varepsilon_1$相互独立, 则(5.4)式等价于

$$\begin{aligned}&\mathsf{P}\{L(\widehat{\beta}_1(S), \beta_1) < L(\widehat{\beta}_1, \beta_1)\}\\ =\ &\mathsf{P}\Big(\Big(\Big(\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\Big)^2 - 2\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\Big)\varepsilon_1'N_2P_1N_2\varepsilon_1 + \Big(\frac{s_{12}}{s_{22}}\Big)^2\frac{\sigma_{22}(1-\rho^2)}{\sigma_{11}}\eta_2'N_2P_1N_2\eta_2\\ &+2\frac{\sigma_{22}(1-\rho^2)}{\sigma_{11}}\Big(\Big(\frac{s_{12}}{s_{22}}\Big)^2\frac{\sigma_{12}}{\sigma_{11}} - \frac{s_{12}}{s_{22}}\Big)\varepsilon_1'N_2P_1N_2\eta_2 < 0\Big).\end{aligned} \tag{5.5}$$

因为η_2与ε_1独立同分布, 结合(5.5)式, 要想得到$\mathsf{P}\{L(\widehat{\beta}_1(S), \beta_1) < L(\widehat{\beta}_1, \beta_1)\} > 1/2$只需

$$\mathsf{P}\Big(2\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}} - \Big(\frac{s_{12}}{s_{22}}\frac{\sigma_{12}}{\sigma_{11}}\Big)^2 > \Big(\frac{s_{12}}{s_{22}}\Big)^2\frac{\sigma_{22}(1-\rho^2)}{\sigma_{11}}\Big) > \frac{1}{2}. \tag{5.6}$$

结合引理3.2, 上式等价于

$$\mathsf{P}\Big(2\frac{\sigma_{12}}{\sigma_{11}}\Big(\frac{\sigma_{11}}{\sigma_{22}}\Big)^{1/2}\Big(\rho+\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_2\Big) > \Big(\frac{\sigma_{11}}{\sigma_{22}}\Big(\frac{\sigma_{12}}{\sigma_{11}}\Big)^2+1-\rho^2\Big)\Big(\rho+\frac{\sqrt{1-\rho^2}}{\sqrt{n-r}}t_2\Big)^2\Big) > \frac{1}{2}. \tag{5.7}$$

经整理得到(5.7)的等价式

$$\mathsf{P}\Big(\rho^2 > \frac{1-\rho^2}{n-r}t_2^2\Big) > \frac{1}{2}. \tag{5.8}$$

因为t_2服从自由度为$n-r$的t分布, 因此$t_2 \sim F_{1,n-r}$, 显然当$(n-r)\rho^2/(1-\rho^2) > F_{1,n-r}(1/2)$时, (5.8)式成立.

定理证毕. □

参 考 文 献

[1] Zellner, A., An efficient method of estimating seemingly unrelated regressions and test for aggregation bias, *J. Amer. Statist. Assoc.*, **57**(1962), 348–368.

[2] Zellner, A., Estimates for seemingly unrelated regression equations: some exact finite sample results, *J. Amer. Statist. Assoc.*, **58**(1963), 977–992.

[3] Revankar, N.S., Some finite results in the context of two seemingly unrelated regression equations, *J. Amer. Statist. Assoc.*, **69**(1974), 187–190.

[4] Revankar, N.S., Use of restricted residuals in SUR systems: some finite sample results, *J. Amer. Statist. Assoc.*, **71**(1976), 183–188.

[5] 王松桂, 线性回归系统回归系数的一种新估计, 中国科学, **10**(1988), 1033–1040.

[6] Liu Aiyi, Efficient estimation of two seemingly unrelated regression equations, *Journal of Multivariate Analysis*, **82**(2002), 445–465.

[7] Rao, C.R., *Proceeding of the Fifth Berkelley Symposium on Math. Stat. & Prob.* (Eds, Cam, L.Le & Neyman, J.), **2**(1967), 355–372.

[8] 林春土, m个半相依回归方程组系数的两步估计, 科学通报, **15**(1984), 957.

[9] 刘金山, 王松桂, SUR模型参数的一种两步估计, 自然科学进展, **9**(1999), 1187–1192.

[10] 王立春, 两个半相依回归方程中的Bayes和经验Bayes迭代估计, 中国科学A辑, **35(5)**2005, 585–600.

[11] Rao, C.R., *Linear Statistical Inference and Applications*, Wiley, New York, 1973.

[12] 王松桂, 严利清, 协方差改进法与半相依回归的参数估计, 应用概率统计, **13(3)**(1997), 288–296.

The Improved Estimates of Regression Coefficients in Seemingly Unrelated Regression Model with Two Equations

Ma Tiefeng

(*Statistics College, Southwestern University of Finance and Economics, Chengdu, 610074*)

Wang Songgui

(*College of Applied Sciences, Beijing University of Technology, Beijing, 100022*)

For regression coefficients in seemingly unrelated regression model with two equations, this paper obtains a two-stage covariance improved estimator sequence by using the covariance improved idea for reference. The condition under which this improved estimator is equivalent to the classic two-stage estimator and the optimal property on the term of mean square error are also obtained. And then, by improving the two-stage estimator in the literature, this paper derives a new improved estimator which shows small sample property relative to least square estimator in more wide parameters space. Through a ultimate way, this paper provides a new estimator which have better small sample property. Finally, this paper discusses the optimal property of two-stage estimator under the Pitman criterion.

Keywords: Seemingly unrelated regression model, covariance improved estimation, two-stage estimator, Pitman criterion.

AMS Subject Classification: 62H12, 62F10.