第12卷 第2期 工程数学学报 Vol. 12 No. 2
1995年6月 JOURNAL OF ENGINEERING MATHEMATICS Jun. 1995

聚集数据的线性模型参数的改进 Peter—Karsten 估计*

刘维奇 潘晋孝
(山西大学数学系,太原 030006) (华北工学院理学系,太原 030051)

提 要

文献[1]给出了聚集数据的线性模型参数的两种估计方法。本文利用引理2对两种估计作了改进,并研究了其效率问题。

关键词 聚集数据的线性模型,估计效率,最佳线性无偏估计

分类号 62J05

0 引 言

考虑线性模型

$$\begin{cases} y = X\beta + u \\ Eu = 0, \mathrm{Var}(u) = \sigma^2 I_n \end{cases} \tag{0.1}$$

其中 y 是 $n \times 1$ 观测向量;X 是 $n \times p$ 列满秩设计矩阵;β 是 $p \times 1$ 未知参数向量;u 是 $n \times 1$ 随机误差向量;σ^2 是误差方差,未知;$n \geqslant p$。

众所周知,模型(0.1)参数 β 的最小二乘估计(简记为LSE)

$$\hat{\beta} = (X'X)^{-1}X'y \tag{0.2}$$

是参数 β 的最佳线性无偏估计(简记为BLUE)。然而在许多实际问题中我们无法观测到 y,而只能观测到 y 的一部分分量或 y 的某个线性组合,常称因变量 y 的这种数据为聚集数据,这时无法得到 β 的LS估计。一般说来,可观测到向量

$$Z = Ty \tag{0.3}$$

其中 T 为已知 n 阶方阵。针对这种问题,Peter & Karsten[1] 提出两种估计方法

$$b = (X'X)^{-1}X'Ty \tag{0.4}$$

$$\tilde{\beta} = (X'T'TX)^{-1}X'T'Ty \tag{0.5}$$

并在 $\mathrm{Rank}(T) = q \geqslant p$ 且 T 对称幂等的条件下证明这种估计具有某种优良性,基于这两种

* 本文1994年9月26日收到。

估计，文献[3]在前述条件下研究了$\hat{\beta}$，b及$\tilde{\beta}$之间的相对效率问题。

本文利用Rao引理的思想对估计(0.4)和(0.5)作了进一步改进，提出修正估计

$$\bar{b} = (X'X)^{-1}X'T(I_n - P_{T'W})y \tag{0.6}$$

$$\bar{\beta} = (X'T'TX)^{-1}X'T'T(I_n - P_{T'W})y \tag{0.7}$$

其中$W = (TX)^{\perp}$是满足

$$(TX)'W = 0 \tag{0.8}$$

的一个最大列满秩矩阵，这里$A^{\perp}$表示由空间$\mu^{\perp}(A')$的一组基为列组成的矩阵，$P_{T'W}$是向列空间$\mu(T'W)$的正交投影阵，即

$$P_{T'W} = T'W(W'TT'W)^{+}W'T \tag{0.9}$$

$(W'TT'W)^{-}$表示矩阵$W'TT'W$的Moore－Penrose广义逆，熟知$P_{T'W}$与广义逆的选取无关。由W的性质可见W的选取不影响$P_{T'W}$，从而$\bar{\beta}$与W选取无关。

本文仅在条件Rank$(TX) = p$下讨论了估计$\bar{\beta}$和$\bar{b}$的存在性，并且证明了$\bar{\beta}$和$\bar{b}$分别比$\tilde{\beta}$和b有较小的均方误差。另外还讨论了$\bar{\beta}$和$\bar{b}$的效率问题，从效率的角度进一步证明$\bar{\beta}$和$\bar{b}$分别比$\tilde{\beta}$和b更优。更值得注意的是经改进之后的估计$\bar{\beta}$，当T满秩时

$$\bar{\beta} = \hat{\beta}$$

即$\bar{\beta}$还是最小二乘估计的推广，所以改进Peter－Karsten估计比原估计更好。

1　改进Peter－Karsten估计的提出

为了便于讨论，首先引入两个基本引理。

引理1(Rao引理)　设T_1, T_2分别为k_1, k_2维统计量，且$E(T_1) = \theta, E(T_2) = 0$，

$$\mathrm{COV}\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \triangleq V$$

这里V为已知正定阵，则在估计类

$\mathscr{A} = \{\mathrm{T} = \mathrm{A}_1\mathrm{T}_1 + \mathrm{A}_2\mathrm{T}_2$，$\mathrm{A}_1$和$\mathrm{A}_2$分别为$\mathrm{k}_1 \times \mathrm{k}_1, \mathrm{k}_2 \times \mathrm{k}_2$非随机阵$\}$中，$\theta$的*BLU*估计为：

$$\theta^* = \mathrm{T}_1 - \mathrm{V}_{12}\mathrm{V}_{22}^{-1}\mathrm{T}_2 \tag{1.1}$$

其协方差阵为：

$$Cov(\theta^*) = \mathrm{V}_{11} - \mathrm{V}_{12}\mathrm{V}_{22}^{-1}\mathrm{V}_{21} \triangleq \mathrm{V}_{11.2} \tag{1.2}$$

引理证明见文献[2]。由此引理可见$\mathrm{V}_{11} \geqslant \mathrm{V}_{11.2}$，所以$\theta^*$和$\mathrm{T}_1$作为$\theta$的两个无偏估计，$\theta^*$比$\mathrm{T}_1$更有效。

引理2　设$\triangle = diag(\delta_1, \cdots, \delta_p)$，$\delta_1 \geqslant \delta_2 \geqslant \cdots \geqslant \delta_p > 0$，U为$\mathrm{n} \times \mathrm{p}$矩阵，记$\lambda_i(\mathrm{A})$为A的第i个顺序(由大到小)特征根，若$\mathrm{A}_{n\times n} > 0$，则

$$\max_{\mathrm{U}'\mathrm{U}=\triangle} tr(\mathrm{U}'\mathrm{AU}) = \sum_{i=1}^{p}\lambda_i(\mathrm{A})\delta_i \tag{1.3}$$

证明见文献[3,4]。

针对模型(0.1)，我们只能观测到

$$\mathrm{Z} = \mathrm{Ty}$$

其中T是一般n阶方阵，满足条件

$$Rank(\mathrm{TX}) = \mathrm{p} \tag{1.4}$$

首先选取

$$T_1 = \tilde{\beta} = (X'T'TX)^{-1}X'T'Ty \tag{1.5}$$

$$T_2 = W'Ty \tag{1.6}$$

则 $ET_1 = \beta, ET_2 = 0$

$$Cov\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} =$$

$$\sigma^2 \begin{pmatrix} (X'T'TX)^{-1}X'T'TT'TX(X'T'TX)^{-1} & (X'T'TX)^{-1}X'T'TT'W \\ W'TT'TX(X'T'TX)^{-1} & W'TT'W \end{pmatrix}$$

$$\triangleq \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

利用引理 1，构造

$$\begin{aligned} \tilde{\beta}_M &= T_1 - V_{12}V_{22}^{-1}T_2 \\ &= (X'T'TX)^{-1}X'T'T(I_n - T'W(W'TT'W)^{-1}W'T)y \end{aligned} \tag{1.7}$$

但注意到 $W'TT'W$ 未必可逆。如果硬要求 $W'TT'W$ 可逆，那么 W 的存在就成了问题。为此我们避开这个问题，选取广义逆，即令

$$\begin{aligned} \bar{\beta} &= (X'T'TX)^{-1}X'T'T(I_n - T'W(W'TT'W)^{+}W'T)y \\ &= (X'T'TX)^{-1}X'T'T(I_n - P_{T'W})y \end{aligned} \tag{1.8}$$

当 T 为满秩矩阵时，即 y 为可观测时，由 $W = (TX)^{\perp}$

$$T'W = X^{\perp} \tag{1.9}$$

再注意到 $P_x + P_{x^{\perp}} = I_n$，即

$$X(X'X)^{-1}X' + X^{\perp}(X^{\perp\prime}X^{\perp})^{-1}X^{\perp\prime} = I_n \tag{1.10}$$

从而 $\bar{\beta} = \hat{\beta}$，所以新估计 $\bar{\beta}$ 是 $LSE\hat{\beta}$ 的推广。

同理可以给出 b 的改进估计

$$\bar{b} = (X'X)^{-1}X'T(I_n - P_{T'W})y \tag{1.11}$$

这里应注意到 b 不是 β 的无偏估计，上述改进形式是从引理形式上得到的。

2　主要结论

定理 1　$E\bar{\beta} = \beta$ (2.1)

$$Cov(\bar{\beta}) \leqslant Cov(\tilde{\beta}) \tag{2.2}$$

证明　(2.1) 式显然。注意到

$$Cov(\bar{\beta}) = \sigma^2 (X'T'TX)^{-1}X'T'T(I_n - P_{T'W})T'TX(X'T'TX)^{-1}$$

$$Cov(\tilde{\beta}) = \sigma^2 (X'T'TX)^{-1}X'T'TT'TX(X'T'TX)^{-1}$$

由于 $I_n - P_{T'W} \leqslant I_n$，故(2.2) 式成立，定理证毕。

定理 2　$E\bar{b} = Eb$ (2.3)

$$Cov(\bar{b}) \leqslant Cov(b) \tag{2.4}$$

$$MSE(\bar{b}) \leqslant MSE(b) \tag{2.5}$$

证明　(2.3) ~ (2.4) 式类似于定理 1 的证明，由于 $\bar{b}$，b 均不是无偏估计，仅从方差大小无法说明他们的优劣，还需从均方误差的角度来看，注意到

$$MSE(\bar{b}) = E(\bar{b} - \beta)(\bar{b} - \beta)'$$

$$= Cov(\bar{b}) + (E\bar{b} - \beta)(E\bar{b} - \beta)'$$

$$MSE(b) = Cov(b) + (Eb - \beta)(Eb - \beta)'$$

由(2.3)～(2.4)式即得(2.5)式，定理证毕。

定理1和定理2说明估计量$\bar{\bar\beta}$和$\bar{b}$在均方误差意义下改进了估计$\bar\beta$和b。

最后我们来讨论估计$\bar{\bar\beta}$和$\bar{b}$相对于$LSE\hat\beta$的效率问题，这里效率定义为

$$e_1(\bar{\bar\beta}|\hat\beta) = \frac{trCov(\hat\beta)}{trCov(\bar{\bar\beta})} \tag{2.6}$$

由于$\bar{b}$是β的有偏估计，故定义

$$e_2(\bar{b}|\hat\beta) = \frac{trCov(\hat\beta)}{trMSE(\bar{b})} \tag{2.7}$$

从定理1和定理2易见

$$e_1(\bar{\bar\beta}|\hat\beta) \geqslant e_1(\bar\beta|\hat\beta) \tag{2.8}$$

$$e_2(\bar{b}|\hat\beta) \geqslant e_2(b|\hat\beta) \tag{2.9}$$

说明在实际中，用估计$\bar{\bar\beta}$和$\bar{b}$代替$\hat\beta$分别比用估计$\bar\beta$和b代替$\hat\beta$有更高的效率，即所蒙受的损失更少一些。从这个意义上讲，新的估计$\bar{\bar\beta}$和$\bar{b}$分别比原估计更优。下面给出效率(2.6)的一个更精确的下界。

记$\delta_1^2 \geqslant \delta_2^2 \geqslant \cdots \geqslant \delta_p^2$为矩阵$(X'X)$的特征根，则有

$$trCov(\hat\beta) = tr(\sigma^2 (X'X)^{-1}) = \sigma^2 \sum_{i=1}^{p} \delta_i^{-2} \tag{2.9}$$

又记$C = T'TX(X'T'TX)^{-1}$，则

$$Rank(C) = Rank(TX) = p \tag{2.10}$$

设Q是n阶正交阵，满足

$$Q'C'CQ = diag(\lambda_1^2, \cdots, \lambda_p^2, 0, \cdots, 0)$$

其中$\lambda_1^2 \geqslant \lambda_2^2 \geqslant \cdots \geqslant \cdots \geqslant 0$是矩阵$C'C$的特征根，则

$$\begin{aligned} trCov(\bar{\bar\beta}) &= \sigma^2 tr[C'(I_n - P_{T'W})C] \\ &\leqslant \sigma^2 tr(C'C) = \sigma^2 \sum_{i=1}^{p} \lambda_i^2 \end{aligned} \tag{2.11}$$

综合(2.9)和(2.11)式可得

定理3　$$e_1(\bar{\bar\beta}|\hat\beta) \geqslant \frac{\sum_{i=1}^{p} \delta_i^{-2}}{\sum_{i=1}^{p} \lambda_i^2} \tag{2.12}$$

其中$\delta_1^2 \geqslant \cdots \geqslant \delta_p^2 > 0$和$\lambda_1^2 \geqslant \cdots \geqslant \lambda_p^2 > 0$分别是矩阵$X'X$和$(X'T'TX)^{-1}X'T'TT'TX(X'T'TX)^{-1}$的特征根。

从定理3可以看到效率$e_1(\bar{\bar\beta}|\hat\beta)$的下界比[3]中给出的效率$e_1(\bar\beta|\hat\beta)$的下界有明显的提高，事实上，由引理2

$$\sum_{i=1}^{p} \lambda_i^2 = tr(C'C) \leqslant \sum_{i=1}^{p} h_i^2 \gamma_{p-i+1}^{-2} \tag{2.13}$$

其中$h_1^2 \geqslant \cdots \geqslant h_r^2 > 0$是矩阵$T'T$的特征根，$\gamma_1^2 \geqslant \cdots \geqslant \gamma_p^2 > 0$是矩阵$X'T'TX$的特征根。

参考文献

[1] Peter S. and Karsten S. On least squares estimation with a particular linear function of the dependent variable. Economics Letters. 23(1), 1987: 59−64.

[2] 王松桂　线性回归系统回归系数的一种新估计，中国科学(A 辑)，1985(10):1033−1040。

[3] 段清堂　聚集数据的线性模型参数估计的效率问题，数理统计与应用概率，7(2)，1992:161−165。

[4] 刘爱义　王松桂　线性模型最小二乘估计的一种新的相对效率，应用概率统计，5(2)，1989:97−104。

The Modified Peter—Karsten Estimate of the Parameters in the Linear Model with Aggregated Date

Liu Weiqi
(Dept. of Math., Shanxi Univ.)

Pan Jinxiao
(Dept. of Sci., North China Inst. of Tech.)

Abstract

Two estimators of the parameters in the linear model with aggregated data are proposed by Peter & Karsten(1987). In this paper, the modified Peter—Karsten estimate is given and its estimate efficiency is studied.