第 13 卷　第 4 期
1996 年 12 月

工 程 数 学 学 报
JOURNAL OF ENGINEERING MATHEMATICS

Vol. 13 No. 4
Dec. 1996

聚集数据线性模型参数的一种新估计*

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)

摘　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。本文提出了一种新的估计方法,并在理论和数值模拟方面证明了新估计较原估计有更高的精确度。

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

1　引　言

考虑线性回归模型

$$\begin{cases} \boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{u} \\ E\boldsymbol{u} = 0,\ \mathrm{Var}(\boldsymbol{u}) = \sigma^2 I_n \end{cases} \tag{1}$$

其中 $\boldsymbol{y}$ 为 $n\times 1$ 观测向量; $\boldsymbol{X}$ 为 $n\times p$ 列满秩设计矩阵; β 为 $p\times 1$ 未知参数向量; $\boldsymbol{u}$ 为 $n\times 1$ 随机误差向量; σ^2 是误差方差,未知,在下面的讨论中假定 $n > p$。

众所周知, β 的最小二乘估计(LSE)

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{2}$$

是 β 的最佳线性无偏估计(BLUE)。但是在某些场合我们无法观测到 $\boldsymbol{y}$ 而只能观测到其某些分量或某个线性组合,记为

$$\boldsymbol{Z} = \boldsymbol{T}\boldsymbol{y} \tag{3}$$

其中 $\boldsymbol{T}$ 为已知的 n 阶方阵。例如在抽样调查,生物研究中的某些问题以及一般线性模型的缺落值问题等均可归结为这样情形。称因变量 $\boldsymbol{y}$ 的这种数据为聚集数据(aggregated data)。文献[1]给出相应模型参数的两种估计:

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{T}\boldsymbol{y} \tag{4}$$

$$\tilde{\beta} = (\boldsymbol{X}'\boldsymbol{T}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{T}'\boldsymbol{y} \tag{5}$$

并在 $\mathrm{rank}(\boldsymbol{T}) = q \geq p$ 且 $\boldsymbol{T}$ 对称幂等的条件下,证明了这两种估计具有某些优良性。在同样的条件下文献[2]讨论了估计 $\tilde{\beta}$, $\boldsymbol{b}$, $\hat{\beta}$ 之间的相对效率及其界,关于效率可参见[3,4],本文给

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved.　http://

* 本文 1996 年 4 月 1 日收到。

出一种新估计方法, 即

$$\beta^* = (X'P_{T'}X)^{-1}X'P_{T'}y \tag{6}$$

其中 $P_{T'} = T'(TT')^+ T$ 是向量空间 $\mu(T')$ 的正交投影阵, $(TT')^+$ 表示矩阵 TT' 的 Moore-Penrose 广义逆。

考虑到参数 β 的可估性, 假定 $\mathrm{rank}(TX) = p$。本文讨论了估计 β 和 β^* 的存在性, 证明了 β^* 比 β 具有更小的方差, 同时指出, 在某种情况下 β^* 还是 β 和 $\hat{\beta}$ 的推广, 即当 T 可逆时 $\beta^* = \hat{\beta}$, 当 $T'T$ 幂相等时 $\beta^* = \beta$, 所以 β^* 比 β 更优。另外我们还研究了估计 β^* 的效率问题。最后做了实例模拟, 对估计 β, b 和 β^* 作了比较, 进一步证明了 β^* 优于 β 和 b。

2 估计 β^* 的提出及其优良性

为了便于讨论, 首先引入两个引理。

引理 1 $X'T'TX$ 可逆 $\Leftrightarrow X'P_{T'}X$ 可逆

证明 记 $q = \mathrm{rank}(T)$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_q > 0$ 是 TT' 的非零特征根, 则存在正交矩阵 Q 使得

$$TT' = Q\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_q & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}Q' \triangleq Q\begin{pmatrix} \wedge_q & \\ & 0 \end{pmatrix}Q'$$

又可写为

$$TT' = Q\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}\begin{pmatrix} I_q & \\ & 0 \end{pmatrix}\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}Q'$$

故

$$(TT')^- = Q\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}\begin{pmatrix} I_q & B \\ C & D \end{pmatrix}\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}Q'$$

其中 B、C、D 为适当阶数的任意阵, 取 $B = 0$, $C = 0$, $D = I_{n-q}$, 并注意到 $P_{T'}$ 的唯一性, 我们有

$$P_{T'} = T'(TT')^+ T = T'(TT')^- T = T'Q\begin{pmatrix} \wedge_q^{-1/2} & \\ & I_{n-q} \end{pmatrix}\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}Q'T$$

于是

$$X'P_{T'}X = X'T'Q\begin{pmatrix} \wedge_q^{-1/2} & \\ & I_{n-q} \end{pmatrix}\begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}Q'TX$$

由此可见 $X'T'TX$ 可逆 $\Leftrightarrow TX$ 列满秩 $\Leftrightarrow \begin{pmatrix} \wedge_q^{1/2} & \\ & I_{n-q} \end{pmatrix}Q'TX$ 列满秩 $\Leftrightarrow X'P_{T'}X$ 可逆。证毕。

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

$$\max_{U'U=\Delta} \mathrm{tr}(U'AU)^{-1} = \sum_{i=1}^{p} \lambda_{n-i+1}^{-1}(A)\delta_{p-i+1}^{-1} \tag{7}$$

证明 可参见文献[2, 5]。

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved. http://

针对模型(1), 我们只观测到

$$Z = Ty$$

其中 T 是一般 n 阶方阵, 满足条件

$$\operatorname{rank}(TX) = p \tag{8}$$

用 $P_{T'}$ 左乘(1) 式两端得

$$\begin{cases} P_{T'}y = P_{T'}X\beta + P_{T'}u \\ EP_{T'}u = 0,\ \operatorname{Var}(P_{T'}u) = \sigma^2 P_{T'} \end{cases} \tag{9}$$

对于奇异线性模型(9), 由[3: 175, 定理 4.1(1)] 可得 β的 BLUE

$$\beta_R^* = (X'P_{T'}\Sigma^- P_{T'}X)^- X'P_{T'}\Sigma^- P_{T'}y \tag{10}$$

其中 $\Sigma = P_{T'} + P_{T'}XRX'P_{T'}$, R 对称, 且使

$$\operatorname{rank}(\Sigma) = \operatorname{rank}(P_{T'} \mid P_{T'}X)$$

又

$$\begin{aligned} \mu(P_{T'}X) &= \{P_{T'}Xt \mid t \in \mathbf{R}_p\} \\ &= \{P_{T'}Xt \mid Xt \in \mathbf{R}_n\} \subset \{P_{T'}t \mid t \in \mathbf{R}_n\} = \mu(P_{T'}) \end{aligned}$$

我们可取 $R = 0$, 于是

$$\Sigma = \mu P_{T'}$$

此时(10) 式变化为

$$\beta_R^* = (X'P_{T'}X)^- X'P_{T'}y \tag{11}$$

根据(8) 式可知 $X'T'TX$ 可逆, 由引理 1 知 $X'P_{T'}X$ 可逆, 故

$$\beta_R^* = (X'P_{T'}X)^{-1}X'P_{T'}y$$

这就是我们要提出的估计, 记为 β^*。

由表达式易见, 当 T 为满秩矩阵时, 即 y 可观测时, $\beta^* = \hat{\beta}$, 而此时 $\tilde{\beta} \neq \hat{\beta}$, $b \neq \hat{\beta}$, 因此可以认为 β^* 是聚集数据情形 LSE$\hat{\beta}$的推广。从而可见 β^* 优于 $\tilde{\beta}$和 b, 下面讨论 β^* 的性质。

定理 1　$$E\beta^* = \beta \tag{12}$$

$$\operatorname{Var}(\beta^*) = \sigma^2 (X'P_{T'}X)^{-1} \tag{13}$$

由定理 1 可见 β^* 是 β的无偏估计。

定理 2　设 C 为任一 p 维向量, 则

$$\operatorname{Var}(C'\tilde{\beta}) \geq \operatorname{Var}(C'\beta^*) \tag{14}$$

证明　用 $\|a\|$ 表示向量 a 的模, 则有

$$\operatorname{Var}(C'\tilde{\beta}) = \sigma^2 \|T'TX(X'T'TX)^{-1}C\|^2 \tag{15}$$

$$\operatorname{Var}(C'\beta^*) = \sigma^2 \|P_{T'}X(X'P_{T'}X)^{-1}C\|^2 \tag{16}$$

为简便起见, 记 $M = X'P_{T'}X$, 由于

$$\begin{aligned} &\|T'TX(X'T'TX)^{-1}C\|^2 \\ &= \|T'TX(X'T'TX)^{-1}C - P_{T'}XM^{-1}C + P_{T'}XM^{-1}C\|^2 \\ &= \|T'TX(X'T'TX)^{-1}C - P_{T'}XM^{-1}C\|^2 + \|P_{T'}XM^{-1}C\|^2 \\ &\quad + 2C'M^{-1}X'P_{T'}[T'TX(X'T'TX)^{-1} - P_{T'}XM^{-1}]C \end{aligned}$$

注意到 $P_{T'}^2 = P_{T'}$, $P_{T'}T' = T'$, 则上式右端第三项

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved.　http://

$$X'P_{T'}[T'TX(X'T'TX)^{-1} - P_{T'}XM^{-1}]$$

$$= X'P_T T'TX(X'T'TX)^{-1} - X'P_T XM^{-1}$$
$$= I_p - MM^{-1} = 0$$

所以

$$\| T'TX(X'T'TX)^{-1}C \|^2 \geq \| P_T XM^{-1}C \|^2$$

由(15)和(16)即得(14)式。证毕。

由上述定理可见 β^* 比 $\tilde{\beta}$ 有较小的方差,又二者均为无偏估计,所以 β^* 比 $\tilde{\beta}$ 更优。

定理3 若 $T'T$ 为幂等矩阵,则 $\beta^* = \tilde{\beta}$。

证明 $T'T$ 幂等,则其特征根为1或0,从而 TT' 的特征根也为1或0,从而存在正交矩阵 Q_1,使得

$$TT' = Q_1\begin{pmatrix} I_q & \\ & 0 \end{pmatrix}Q_1'$$

那么 TT' 也是幂等矩阵,根据[3:18],$(TT')^+ = TT'$,于是

$$P_{T'} = T'(TT')^+ T = TT' \tag{17}$$

故

$$\beta^* = (X'P_T X)^{-1}X'P_T y = (X'T'TX)^{-1}X'T'Ty = \tilde{\beta}$$

定理证毕。

文献[1]在 $\mathrm{rank}(T) = q \geq p$ 且 T 对称幂等的条件下讨论了 $\tilde{\beta}$ 的优良性。在此条件下由定理3保证 $\beta^* = \tilde{\beta}$,故从某种意义上讲 β^* 是 $\tilde{\beta}$ 的推广或改进。

需要指出的是尽管聚集数据的线性模型可化为奇异型线性模型

$$\begin{cases} Z = TX\beta + v \\ Ev = 0,\ \mathrm{Var}(v) = \sigma^2 T'T,\ T'T \geq 0 \end{cases}$$

根据Rao定理已有 β 的BLUE β_R^* 见[3],但其中存在对 U 的选择问题,从应用角度讲 β^* 比 β_R^* 计算简便易行。

最后进一步讨论估计 β^* 的效率问题,我们使用效率的定义为

$$e(\beta^* \mid \hat{\beta}) = \frac{\mathrm{tr\ Var}\hat{\beta}}{\mathrm{tr\ Var}\beta^*} \tag{18}$$

由Gauss-Markov定理知

$$\mathrm{Var}(\beta^*) \geq \mathrm{Var}(\hat{\beta})$$

经计算

$$e(\beta^* \mid \hat{\beta}) = \frac{\mathrm{tr}\ (X'X)^{-1}}{\mathrm{tr}(X'P_T X)^{-1}} \tag{19}$$

设 $X'X$ 的特征根为 $\varphi_1 \geq \varphi_2 \geq \cdots \geq \varphi_p > 0$,则

$$\mathrm{tr}\ (X'X)^{-1} = \sum_{i=1}^{p}\varphi_i^{-1} \tag{20}$$

设 $X'T'TX$ 的特征根为 $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_p > 0$,且假定 S 是使

$$S'X'T'TXS = \mathrm{diag}(\psi_1, \psi_2, \cdots, \psi_p)$$

的 p 阶正交方阵,由

$$\mathrm{tr}\ (X'P_T X)^{-1} = \mathrm{tr}(X'T'(T'T)^- TX)^{-1}$$

在引理1证明中取 $B = 0, C = 0, D = \lambda_q^{-1}I_{n-q}$,并利用引理2,有

$$\mathrm{tr}\ (X'P_T X)^{-1} = \mathrm{tr}\left[X'T'Q\begin{pmatrix} \wedge_q^{-1} & \\ & \lambda_q^{-1}I_{n-q} \end{pmatrix}Q'TX\right]^{-1}$$

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved. http://

$$= \mathrm{tr}\left[SS'X'T'Q\begin{pmatrix} \wedge_q^{-1} & \\ & \lambda_q^{-1}I_{n-q} \end{pmatrix} Q'TXSS' \right]^{-1}$$

$$= \mathrm{tr}\left[SX'T'Q\begin{pmatrix} \wedge_q^{-1} & \\ & \lambda_q^{-1}I_{n-q} \end{pmatrix} Q'TXS \right]^{-1}$$

$$\leq \sum_{i=1}^{p} \lambda_i \cdot \psi_{p-i+1}^{1} \tag{21}$$

于是由(19) ～(21) 式可得如下定理

定理 4　若 rank(TX) = p , 则

$$e(\beta^* \mid \hat{\beta}) \geq \frac{\sum_{i=1}^{p} \varphi_i^{1}}{\sum_{i=1}^{p} \lambda_i \cdot \varphi_{p-i+1}^{1}} \tag{22}$$

特别当 T 满秩时 $\beta^* = \hat{\beta}$, $e(\beta^* \mid \hat{\beta}) = 1$.

3　模拟比较

通过前面的讨论, 我们得到了聚集数据的线性模型参数的一种新估计 β^* , 并证明了它较原估计 b 及 $\hat{\beta}$ 有更高的精度, 这一点将通过下面的模拟结果得到进一步证明。我们的模拟方法是预先确定设计矩阵 X 和 T, 针对不同的参数 β 及 σ^2 产生多组正态随机数, 对其分别计算观测向量 y, 然后利用 $\hat{\beta}$, b 和 β^* 三种方法估计出参数 β, 最后计算出三种估计值的样本均值和样本方差。下面仅举两例, 具体 X、T 和 y 的数值略, 只将参数值以及估计均值和方差列于下表。

表　估计 $\hat{\beta}$, b 和 β^* 的模型比较表

	β	$\hat{\beta}$		b		β^*	
		均值	方差	均值	方差	均值	方差
例 1	6.0	9.504757	17.38767	-224.331	177.3327	3.315930	1.239337
	-4.2	-4.169478	0.032282	11.93595	0.897657	-4.302840	0.021921
	5.3	5.390732	0.054535	-5.570057	0.119953	5.507424	0.031478
例 2	0.5	0.531868	8.817726	161.4047	9.77787	0.537519	1.734438
	0.6	0.585278	0.260399	-0.366934	0.001627	0.567327	0.000003
	4.9	3.855774	6.294691	-108.4115	55.90733	4.642895	2.072434
	-2.2	-3.119400	0.159775	42.17307	22.80834	-2.031907	0.033165

从表中数据和多次模拟结果可以看出 β^* 比 $\hat{\beta}$ 和 b 更接近于真值, 均方误差也较小, 特别当 n 较大而 rank(T) 较小时这种估计的优越性尤为突出, 所以在实际中用 β^* 代替 $\hat{\beta}$ 所蒙受的损失会更小, 从而估计精度会更高。同时我们也注意到, 比起 β^* 的其它分量, 第一分量的模拟效果较差; 再者, b 是由一般线性模型的最小二乘法得到的估计, 其效果很差, 但就目前

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved.　http://w

而言,当 T 未知时只有这一种估计。为此,对于聚集数据的线性模型的参数估计有待于我们进一步研究。

参考文献

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

2 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, 1992, 7(2): 161～165

3 王松桂. 线性模型的理论及应用. 合肥,安徽教育出版社, 1987: 173～177

4 张信东, 刘兰亭. Lindley 相关性度量与 LS 估计效率. 数理统计与应用概率, 1991, 6(4): 513～526

5 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, 1989, 5(2): 97～104

A New Estimate of the Parameters in the Linear Model with Aggregated Data

Liu Weiqi
(Dept. of Math., Shanxi Univ. Taiyuan, 030006)
Pan Jinxiao
(Dept. of Sci., North China Inst. of Tech. Taiyuan, 030051)

Abstract

Two estimators of the parameters in the linear model with aggregated data are proposed by Peter & Karsten(1987)[1]. In this paper, we find a new estimate according to the better way, also it is proved that the new esttimate have higheraccuracy than the old estimates by using both theoretical method and numerical fitting.

© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved. http://