

广义压缩估计的若干性质

张 钦

(淮阴工业专科学校 淮阴 223001)

摘 要 针对人们处理较多自变量的大型回归问题时设计矩阵难免存在近似的线性关系的问题,讨论了文义压缩最小二乘估计(即 GSLS 估计)的一些性质,如相似性、优良性、精度等,认为 GSLS 估计是一个非常实用的估计类.

关键词 广义压缩最小二乘估计(GSLS) 相合性 优良性 精度

中图号 O213

1993 年王力群^[1]引进了线性回归模型中回归系数的一个估计类——广义压缩最小二乘估计(General Shurnken Least Squares, 简记为 GSLS). 许多常用的估计如岭估计、广义岭估计、主成分估计和压缩估计都属于这个估计类. 文[1]中已经讨论了该估计类中估计的相容性问题以及在矩阵均方准则下估计的比较问题. 本文继续文[1]的工作, 讨论该估计类的其他性质.

设线性回归模型

$$y = X\beta + \varepsilon \quad E\varepsilon = 0 \quad \text{cov}(\varepsilon) = \sigma^2 I \quad (1)$$

式中: X —— $n \times p$ 的设计阵; ε —— $n \times 1$ 随机向量; I ——单位矩阵; β, σ ——未知参数, $\beta \in \mathbf{R}, 0 < \sigma < \infty$. 此外我们还假定 $2 \leq \text{rank } X = p \leq n$.

本文的有关符号及约定详见文[3]及文[4].

1 GSLS 估计的基本性质及优良性

根据 GSLS 估计的定义, 该估计类显然是一个有偏估计类. 以 $\|\cdot\|$ 记欧氏模, 则具有下列基本性质:

- 对任意 A , 若 $\|\beta\| \neq 0$, 总有 $\|\beta(A)\| \leq \|\hat{\beta}\|$;
- 存在 A , 使 $\text{MSE}(\beta(A)) < \text{MSE}(\hat{\beta})$ 成立.

定理 1 设 β_i^* 为回归系数 β 的保留前 i 个主成分 Massy 的主成分估计, 则有

$$\beta(A) = \sum_{i=1}^p w_i \beta_i^*$$

其中
$$w_i = \begin{cases} a_i - a_{i+1} & 1 \leq i < p \\ a_i & i = p \end{cases}$$

定理 2 令

$$\phi = \text{diag}\left(\frac{1 - a_1}{1 + a_1} \lambda_1, \dots, \frac{1 - a_p}{1 + a_p} \lambda_p\right)$$

则在椭圆 $\beta' P \phi P' \beta \leq \sigma^2$ 内, 有 $\text{MSE}(\beta(A)) \leq \text{MSE}(\hat{\beta})$ 成立.

定理 2 说明了由于 LS 在线性估计类中是可容许估计, 因此我们不可能找到一致地优于 LS 估计的线性有偏估计, 线性有偏估计只能局部地改进 LS 估计.

2 GSLS 估计的有效性

笔者已经用均方误差作为比较 $\beta(A)$ 的 $\hat{\beta}$ 的标准, 讨论了 $\beta(A)$ 的优良性, 证明了在椭圆 $\beta' P \phi P' \beta \leq \sigma^2$ 内 $\beta(A)$ 比 $\hat{\beta}$ 有较小的均方误差.

比较两个估计的另一个重要的标准是相对效率(简称效率),文献[3]给出了线性模型参数估计的效率的一些基本结果.对于本文讨论的有偏估计,引入如下的效率定义:

定义 1 对模型(1)设 $\bar{\beta}$ 为 β 的线性有偏估计,用 $1 - [\text{MSE}(\bar{\beta})/\text{MSE}(\hat{\beta})]$ 来表示 $\bar{\beta}$ 对 $\hat{\beta}$ 的效率,记为 $\rho(\bar{\beta})$.

如果 $\text{MSE}(\bar{\beta}) \leq \text{MSE}(\hat{\beta})$, 则 $\bar{\beta}$ 的效率不小于 0, 因而 $0 \leq \rho(\bar{\beta}) \leq 1$. 效率越接近 1, $\bar{\beta}$ 改进 $\hat{\beta}$ 的程度就越高. 下面的定理给出了 $\beta(A)$ 相对于 LS 估计 $\hat{\beta}$ 的效率的下界.

定理 3 如果 $\lambda_{p-k} \approx \lambda_{p-k+1} \approx \cdots \approx \lambda_p \approx 0$, 则在椭圆 $\beta' P \Phi P' \beta \leq \sigma^2$ 内, 有

$$\rho(\beta(A)) \geq \frac{k(1-M^2)}{k+1+(P-k-\lambda_p)}$$

其中 $M = \max \{a_1, a_2, \cdots, a_p\}$, 且 $M \neq 1$.

从定理的结论可以看出,若 M 的取值越小,即矩阵 $A = \text{diag}(a_1, \cdots, a_p)$ 的元素取值越小,则效率越接近于 1, $\beta(A)$ 改进 $\hat{\beta}$ 的程度就越高. $\beta(A)$ 一般不等于 $\hat{\beta}$, 那么两者之间究竟有多大的偏差? 笔者根据文[2], 给出精度的一个定义, 并由此得到精度的上界.

定义 2 式 $\|\text{cov}(\bar{\beta}) - \text{cov}(\hat{\beta})\|$ 称为估计 $\bar{\beta}$ 代替 $\hat{\beta}$ 的精度.

定理 4 对于模型(1)的 GSLS 估计, 有 $\|\text{cov}(\bar{\beta}) - \text{cov}(\hat{\beta})\| \leq \sigma^2 \frac{\lambda_p}{\lambda_1}$.

若 λ_p 越小, 即矩阵 X 越病态, 则精度越好, 也就是 $\beta(A)$ 与 $\hat{\beta}$ 的偏差越小.

3 GSLS 估计的相合性

对于多元线性回归模型

$$y_i = \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i \quad i = 1, 2, \cdots, n \tag{2}$$

以下记

$$\begin{aligned} X_n &= (x_{ij})_{n \times p} & \beta &= (\beta_1, \cdots, \beta_p)' \\ \varepsilon_n &= (e_1, \cdots, e_n)' & Y_n &= (y_1, \cdots, y_n)' & S_n &= X'X \end{aligned}$$

且假定 n 充分大, X_n 满秩(即 S_n^{-1} 存在).

回归系数的 LS 估计为 $\hat{\beta}_n = S_n^{-1} X_n' Y_n$

式(2)的紧凑形式为 $Y_n = Z_n \alpha + \varepsilon_n$

其中 $Z_n = X_n P$, $\alpha = P' \beta$, 而 P 为正交矩阵, 使得 $P' S_n P = Z_n' Z_n \triangleq A_n = \text{diag}(\lambda_1, \cdots, \lambda_p)$. 这样又可给出 GSLS 估计的定义.

对于 $\hat{\beta}_n$ 的相合性的研究已相当深入, 文献[4]分别给出其作为 β 的弱、强及 r 阶估计的条件, 特别当随机误差 ε_n 满足 GM 条件时, 文献[4]指出 $\hat{\beta}_n$ 为 β 的弱相合估计, 当且仅当

$$\lim_{n \rightarrow \infty} S_n^{-1} = 0$$

下面讨论 $\beta_n(A_n)$ 的相合性.

引理 1 设 S_n 为正定矩阵序列, 则 $\lim_{n \rightarrow \infty} S_n^{-1} = 0$, 当且仅当 $\lim_{n \rightarrow \infty} \text{tr} S_n^{-1} = 0$.

定理 5 对模型(2), 设 ε_n 满足 GM 条件:

$$E \varepsilon_n = 0 \quad \text{cov}(\varepsilon_n) = \sigma^2 I_n$$

则有下面的两个等价命题:

- a. $\lim_{n \rightarrow \infty} S_n^{-1} = 0$, 且 $\lim_{n \rightarrow \infty} A_n = I$;
- b. $\beta_n(A_n)$ 是 β 的均方相合估计.

因为由均方收敛就一定得依概率收敛, 所以在定理 5 的条件(a)下就有 $\beta_n(A_n)$ 是 β 弱相合估计这个结论.

下面给出 r 阶平均相合性的充分条件. 为避免重复, 以下总假定在一般线性模型(2)中, 随机误差 $e_1, e_2, \cdots$ 相互独立, 又 $r \in (0, 2)$ 为一给定的数, 且 $\lim_{n \rightarrow \infty} A_n = I$.

定理 6 以 F_i 记 e_i 的分布, 并令

$$\alpha_n(c) = \max_{1 \leq i \leq n} \int_{|x| \geq c} |x|^r dF_i(x)$$

设 $\{e_i |^r\}$ 一致可积,即

$$\lim_{c \rightarrow \infty} \{\sup \alpha_n(c)\} = 0$$

且当 $r \geq 1$ 时,进一步假定 $Ee_1 = Ee_2 = \cdots = 0$,当 $0 < r < 1$ 时,假定每个 e_i 的分布关于原点对称,则当试验点列 $\{w_i\}$ 满足条件 $S_n^{-1} = O(n^{(r-2)/r})$ 时,对任意指定的 $a > 1$,有

$$E \parallel \beta_n(A_n) - \beta \parallel^r = O \{n^{-(a-1)(2-a)/(2a)} + (\alpha_n(n^{1/(ar)}))^{r/2}\}$$

对任意给定的 $a > 1$,由定理 2 的条件及文献[4]中的定理 5.2 知

$$E \parallel \beta_n(A_n) - \beta \parallel^r = O \{n^{-(a-1)(2-a)/(2a)} + (\alpha_n(n^{1/(ar)}))^{r/2}\}$$

定理7 对模型(2),记

$$S_n^{-1} = \begin{bmatrix} S_n^{-1}(1,1) & & & \\ & * & & \\ & & \ddots & \\ & & & * \\ & & & & S_n^{-1}(p,p) \end{bmatrix}$$

设函数 g 定义为 $(0, \infty)$,满足条件:

- a. 当 x 递减趋于 0 时, $0 < g(x)/x$ 且递增趋于无穷大.
 - b. 对某个 $B > 0$, 有 $\int_0^B \frac{1}{g^2(x)} dx < \infty$.
 - c. $\{g(S_n^{-1}(t,t))\varepsilon_n\}$ 为收敛系统,即对任一实数列 $\{a_i, i \geq 0\}$ 若 $\sum_1^\infty a_i^2 < \infty$, 则有 $\sum_1^\infty a_i \varepsilon_i$ 为 a. s. 收敛.
- 又设 $\lim_{n \rightarrow \infty} S_n^{-1}(t,t) = 0, t = 1, 2, \cdots, p$;
- d. $\lim_{n \rightarrow \infty} A_n = I$, 则

$$\lim_{n \rightarrow \infty} \beta_n(A_n) = \beta \quad \text{a. s. 收敛}$$

参 考 文 献

1 王力群. 广义压缩最小二乘估计. 应用概率统计, 1993 (3)
2 杨虎, 王松桂. 条件数、谱范数与估计精度. 应用概率统计, 1993 (4)
3 王松桂. 线性模型的理论及其应用. 合肥 安徽教育出版社, 1987. 5 ~ 16
4 陈希孺, 陈桂景. 线性模型的参数估计理论. 北京 科学出版社, 1985. 63 ~ 65

Some Qualities of Generalized Shrunken Least Square Estimate

Zhang Qin

(Huaiyin Industrial College, Huaiyin 223001)

Abstract Although the regression coefficient of the LS estimate has many good qualities, the large regression problem will probably be manipulated with the rapid development of modern electronic computing technology. At present, because of many self-variables, sometimes it may be approximate linear relation for the design matrix. Many scholars have given a lot of improved methods of LS estimate to solve the problem, but the GSLS estimate is still useful by now. For this reason, the optimality, consistency accuracy, etc. of GSLS estimate are discussed. It is considered to be a practical estimate.

Key words GSLS ; consistency ; optimality ; accuracy