线性回归方程系统中参数估计的相对效率比较

王晓平¹, 张志民², 刘朝林²

(1. 宜宾职业技术学院 基础部, 四川 宜宾 644003; 2. 重庆大学 数学与统计学院, 重庆 401331)

摘要: 文章讨论了在线性回归方程系统中, 相同参数在不同准则下的相对效率以及相对效率的性质, 并对各相对效率的上、下界进行了讨论和证明。

关键词: 线性回归模型; 最优线性无偏估计; 相对效率

中图分类号: O212.1 **文献标识码:** A **文章编号:** 1002-6487(2017)15-0023-04

0 引言

统计学家们常常会对参数估计间的有效性作出比较。比如, Bloomfield 和 Watson^[1], Wang 和 Shao^[2]就利用不等式法对线性模型中未知参数的最小二乘估计和最优线性无偏估计进行了相对效率的讨论, Wang 和 Pan^[3]对奇异线性模型中的最小二乘估计和最优无偏估计的相对效率进行了比较。

设由如下两个线性回归方程组成的线性回归方程系统(R):

$$(R) \begin{cases} y_1 = X_1\beta + H_1\beta_1 + e_1 & (1) \\ y_2 = X_2\beta + H_2\beta_2 + e_2 & (2) \end{cases}$$

其中, $i=1, 2$, y_i 是 $n_i \times 1$ 的观测向量, X_i 和 H_i 分别是 $n_i \times p$ 和 $n_i \times t_i$ 的满秩矩阵, 且满足 $\text{Rank}(X_i, H_i) = \text{Rank}(X_i) + \text{Rank}(H_i)$, 符号 $\text{Rank}(\bullet)$ 表示矩阵的秩, β 和 β_i 分别是 $n \times 1$ 和 $t_i \times 1$ 的未知参数, e_i 是 $n_i \times 1$ 的随机向量, 通常假定其服从均值为 0、协方差矩阵为 $\sigma_i I$ 的多元正态分布, σ_i 是未知参数, 向量 e_1 和 e_2 相互独立。为了便于讨论, 设 $Q_i = I - H_i(H_i^T H_i)^{-1} H_i^T$, $T_i = (H_i^T H_i)^{-1} H_i^T X_i$ 以及 $r = \frac{\sigma_1}{\sigma_2}$, $i=1, 2$ 。

一方面, 人们可针对线性回归方程系统(R)得到参数 β 和 $\beta_i (i=1, 2)$ 的最优线性无偏估计, 可记为:

$$\hat{\beta}^*(r) = (X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1} (X_1^T Q_1 y_1 + r X_2^T Q_2 y_2) \quad (3)$$

$$\hat{\beta}_1^* = (H_1 H_1)^{-1} H_1^T y_1 - T_1 \hat{\beta}^*(r) \quad (4)$$

$$\hat{\beta}_2^* = (H_2^T H_2)^{-1} H_2^T y_2 - T_2 \hat{\beta}^*(r) \quad (5)$$

另一方面, 对单个线性回归方程(1), 参数 β 和 β_1 的最优线性无偏估计为:

$$\hat{\beta} = (X_1^T Q_1 X_1)^{-1} X_1^T Q_1 y_1 \quad (6)$$

$$\hat{\beta}_1 = (H_1^T H_1)^{-1} H_1^T y_1 - T_1 \hat{\beta} \quad (7)$$

同理, 对单个线性回归方程(2), 参数 β 和 β_2 的最优线性无偏估计为:

$$\tilde{\beta} = (X_2^T Q_2 X_2)^{-1} X_2^T Q_2 y_2 \quad (8)$$

$$\tilde{\beta}_1 = (H_2^T H_2)^{-1} H_2^T y_2 - T_2 \tilde{\beta} \quad (9)$$

Liu^[4]对 $\hat{\beta}^*(r)$ 、 $\hat{\beta}$ 和 $\tilde{\beta}$ 分别进行了估计, 并讨论了它们的统计性质, 并在均方误差准则下对 $\hat{\beta}^*(r)$ 、 $\hat{\beta}$ 和 $\tilde{\beta}$ 进行了比较。Ma 和 Wang^[5]也在均方误差准则下对 $\hat{\beta}^*(r)$ 、 $\hat{\beta}$ 和 $\tilde{\beta}$ 进行了比较。在本文中, 将针对线性回归系统(R)、线性回归方程(1)、线性回归方程(2)中的相同参数 β 进行讨论, 并分析其不同估计的相对效率。

对参数 β 的不同估计式(3)、式(6)和式(8), 可计算得协方差阵分别为:

$$\text{Cov}(\hat{\beta}^*(r)) = \sigma_1 (X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1} \quad (10)$$

$$\text{Cov}(\hat{\beta}) = \sigma_1 (X_1^T Q_1 X_1)^{-1} \quad (11)$$

$$\text{Cov}(\tilde{\beta}) = \sigma_2 (X_2^T Q_2 X_2)^{-1} \quad (12)$$

由式(10)至式(12)可得:

$$\text{Cov}(\hat{\beta}^*(r)) \leq \text{Cov}(\hat{\beta}), \text{Cov}(\hat{\beta}^*(r)) \leq \text{Cov}(\tilde{\beta}) \quad (13)$$

式(13)表明线性回归系统(R)的协方差阵小于单个线性回归方程的协方差阵, 表明线性回归系统比单个线性回归方程包含更多的信息, 更能反映实际情况。因此, 如果利用 $\hat{\beta}$ 或 $\tilde{\beta}$ 去代替 $\hat{\beta}^*(r)$, 就可能产生部分损失。为了能较好地定义此种损失, 一些学者对其进行了讨论, 如 Wu^[6], Yang 和 Wang^[7]等。Wang 和 Pan^[3]在矩阵欧几里德范数准则下对奇异线性模型的普通最小二乘估计和最优无偏估计进行了相对效率的比较。Wu^[8]在行列式和矩阵迹准则下讨论了四种相对效率。本文引入两种类型的相对效率, 并重点讨论所引入相对效率的上界和下界。

基金项目: 国家自然科学基金资助项目(11171361)

作者简介: 王晓平(1980—), 女, 黑龙江齐齐哈尔人, 讲师, 研究方向: 统计建模与计算。

张志民(1981—), 男, 河北唐山人, 教授, 博士生导师, 研究方向: 统计模型、风险模型。

刘朝林(1980—), 男, 贵州遵义人, 博士, 副教授, 研究方向: 统计建模与风险分析。

1 相对效率的定义

为了讨论利用 $\hat{\beta}$ 或 $\tilde{\beta}$ 去代替 $\hat{\beta}^*(r)$ 产生的损失, 本文沿用文献[9], 给出如下四种类型的相对效率:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_2}{\|\text{Cov}(\hat{\beta})\|_2}, \quad u_2(\hat{\beta}^*(r)|\tilde{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_2}{\|\text{Cov}(\tilde{\beta})\|_2}$$

$$u_3(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_F}{\|\text{Cov}(\hat{\beta})\|_F}, \quad u_4(\hat{\beta}^*(r)|\tilde{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_F}{\|\text{Cov}(\tilde{\beta})\|_F}$$

其中:

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2} = \sqrt{\text{tr}(A^T A)}$$

称 $\|A\|_2$ 和 $\|A\|_F$ 为矩阵 A 的谱范数和 F 范数。由式(13)可知, $0 < u_i(\bullet) \leq 1 (i=1, 2, 3, 4)$ 。

2 $u_1(\hat{\beta}^*(r)|\hat{\beta})$ 和 $u_2(\hat{\beta}^*(r)|\tilde{\beta})$ 的上界和下界

为了讨论的需要, 给出如下几个引理。

设 A(或 B) 是 $n \times n$ 的非负定矩阵, 即 $A \geq 0, B \geq 0$, $\lambda_1(\bullet) \geq \lambda_2(\bullet) \geq \dots \geq \lambda_n(\bullet)$ 表示矩阵 A(或 B) 的 n 个有序特征值。

引理 1^[9]: 设 A 和 B 是 Hermite 矩阵, $A \geq 0, B \geq 0$, 则有:

$$0 \leq \text{tr}(AB) \leq \lambda_1(B)\text{tr}(A) \leq \text{tr}(A) \cdot \text{tr}(B)$$

引理 2^[9]: 设 A 和 B 是 Hermite 矩阵, 则有:

$$\lambda_i(A+B) \leq \lambda_j(A) + \lambda_k(B), \quad i \geq j+k-1$$

$$\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-n}(A+B), \quad j+k \geq n$$

引理 3^[9]: 设 A 和 B 是 Hermite 矩阵, $A \geq 0, B \geq 0$, 则有:

$$\text{tr}(A^{-1}B) \geq \frac{\text{tr}(B)}{\lambda_1(A)} \geq \frac{\text{tr}(B)}{\text{tr}(A)}$$

引理 4^[9]: (Poincare 定理) 设 A 是 $n \times n$ 的对称矩阵, B 是 $n \times k$ 的正交矩阵, 则有:

$$\lambda_{n-k+i}(A) \leq \lambda_i(B^T A B) \leq \lambda_i(A), \quad i=1, 2, \dots, k$$

首先, 有如下结论:

定理 1: 设 $\hat{\beta}^*(r)$ 和 $\hat{\beta}$ 是由式(3)和式(6)确定的估计量, 其相对效率定义为:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_2}{\|\text{Cov}(\hat{\beta})\|_2}$$

则有:

$$\frac{\theta_p}{\min\{\theta_m + r\eta_n; m+n \leq p+1\}} \leq u_1(\hat{\beta}^*(r)|\hat{\beta}) \leq \frac{\theta_p}{\theta_p + r\eta_p}$$

其中, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0, \eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 对应的 p 个有序特征值。

证明: 由 $u_1(\hat{\beta}^*(r)|\hat{\beta})$ 的定义, 可得:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_2}{\|\text{Cov}(\hat{\beta})\|_2} = \frac{\|(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}\|_2}{\|(X_1^T Q_1 X_1)^{-1}\|_2} \quad (14)$$

又因 $X_1^T Q_1 X_1 > 0, X_2^T Q_2 X_2 > 0$, 故有:

$$\|(X_1^T Q_1 X_1)^{-1}\|_2 = \sqrt{\lambda_1[(X_1^T Q_1 X_1)^{-1}]} = \theta_p^{-1} \quad (15)$$

$$\|(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}\|_2 = \sqrt{\lambda_1[(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}]} = \lambda_p^{-1}(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \quad (16)$$

由引理 2 可得:

$$\lambda_i(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \geq \lambda_s(X_1^T Q_1 X_1) + r\lambda_t(X_2^T Q_2 X_2) = \theta_s + r\eta_t, \quad s+t \leq i+p \quad (17)$$

$$\lambda_i(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \leq \lambda_m(X_1^T Q_1 X_1) + r\lambda_n(X_2^T Q_2 X_2) = \theta_m + r\eta_n, \quad m+n \leq i+1 \quad (18)$$

则:

$$\max\{\theta_s + r\eta_t; s+t \leq i+p\} \leq \lambda_i(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \leq \min\{\theta_m + r\eta_n; m+n \leq i+1\} \quad (19)$$

又因:

$$\lambda_1(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \leq \theta_1 + r\eta_1 \quad (20)$$

$$\lambda_p(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \geq \theta_p + r\eta_p \quad (21)$$

故有:

$$[\min\{\theta_m + r\eta_n; m+n \leq i+1\}]^{-1} \leq \lambda_p^{-1}(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2) \leq (\theta_p + r\eta_p)^{-1} \quad (22)$$

由式(15)和式(22), 可得:

$$\frac{\theta_p}{\min\{\theta_m + r\eta_n; m+n \leq i+1\}} \leq u_1(\hat{\beta}^*(r)|\hat{\beta}) \leq \frac{\theta_p}{\theta_p + r\eta_p} \quad (23)$$

证毕。

事实上, 若 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是正定的数量矩阵, 则可得:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\theta_p}{\theta_p + r\eta_p} \quad (24)$$

定理 2: 设 $\hat{\beta}^*(r)$ 和 $\hat{\beta}$ 是由式(3)和式(6)确定的估计量, 若 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 则:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\min_{1 \leq i \leq p} \theta_i}{\min_{1 \leq i \leq p} (\theta_i + r\eta_i)} \quad (25)$$

其中 $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0, \eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值。

证明: 因 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 则存在一个正交阵 P, 使得:

$$X_1^T Q_1 X_1 = P^T \Lambda P, \quad \Lambda = \text{diag}(\theta_1, \theta_2, \dots, \theta_p) \quad (26)$$

$$X_2^T Q_2 X_2 = P^T \Delta P, \quad \Delta = \text{diag}(\eta_1, \eta_2, \dots, \eta_p) \quad (27)$$

则:

$$(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1} = P^T (\Lambda + r\Delta)^{-1} P \quad (28)$$

又因:

$$\|(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}\|_2 = \sqrt{\lambda_1[(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}]}$$

$$= \lambda_p^{-1}(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) = \lambda_p^{-1}(P^T(\Lambda + r\Delta)P) = \lambda_p^{-1}(\Lambda + r\Delta) = \left[\min_{1 \leq i \leq p} (\theta_i + r\eta_i) \right]^{-1} \quad (29)$$

同时,

$$\| (X_1^T Q_1 X_1)^{-1} \|_2 = \left[\min_{1 \leq i \leq p} \theta_i \right]^{-1} \quad (30)$$

将式(29)和式(30)带入 $u_1(\hat{\beta}^*(r)|\hat{\beta})$ 可证得:

$$u_1(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\min_{1 \leq i \leq p} \theta_i}{\min_{1 \leq i \leq p} (\theta_i + r\eta_i)} \quad (31)$$

证毕。

利用同样的方法,可证明关于 $u_2(\hat{\beta}^*(r)|\hat{\beta})$ 的上界和下界。

定理3: 设 $\hat{\beta}^*(r)$ 和 $\tilde{\beta}$ 是由式(3)和式(8)给出的估计量, 由 $\tilde{\beta}$ 代替 $\hat{\beta}^*(r)$ 产生的相对效率定义为:

$$u_2(\hat{\beta}^*(r)|\tilde{\beta}) = \frac{\| \text{Cov}(\hat{\beta}^*(r)) \|_2}{\| \text{Cov}(\tilde{\beta}) \|_2}$$

则有:

$$r \cdot \frac{\eta_p}{\min\{\theta_m + r\eta_n; m+n \leq i+1\}} \leq u_2(\hat{\beta}^*(r)|\tilde{\beta}) \leq r \cdot \frac{\eta_p}{\theta_p + r\eta_p} \quad (32)$$

其中, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值。

定理4: 设 $\hat{\beta}^*(r)$ 和 $\tilde{\beta}$ 是由式(3)和式(8)给出的估计量, 若 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 则有:

$$u_2(\hat{\beta}^*(r)|\tilde{\beta}) = r \cdot \frac{\min_{1 \leq i \leq p} \eta_i}{\min_{1 \leq i \leq p} (\theta_i + r\eta_i)} \quad (33)$$

其中, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值。

3 $u_3(\hat{\beta}^*(r)|\hat{\beta})$ 和 $u_4(\hat{\beta}^*(r)|\tilde{\beta})$ 的上、下界

下文将讨论 $u_3(\hat{\beta}^*(r)|\hat{\beta})$ 和 $u_4(\hat{\beta}^*(r)|\tilde{\beta})$ 的上、下界。

定理5: 设 $\hat{\beta}^*(r)$ 和 $\hat{\beta}$ 是由式(3)和式(6)确定的估计量, 其相对效率定义为:

$$u_3(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\| \text{Cov}(\hat{\beta}^*(r)) \|_F}{\| \text{Cov}(\hat{\beta}) \|_F}$$

则有:

$$\sqrt{\frac{\sum_{i=1}^p [\min\{\theta_m + r\eta_n; m+n \leq i+1\}]^{-1}}{(\theta_1 + r\eta_1) \cdot \sum_{i=1}^p \theta_i^{-2}}} \leq u_3(\hat{\beta}^*(r)|\hat{\beta}) \leq \sqrt{\frac{\sum_{i=1}^p [\max\{\theta_s + r\eta_t; s+t = i+p\}]^{-1}}{(\theta_p + r\eta_p) \cdot \sum_{i=1}^p \theta_i^{-2}}}$$

其中 $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是

$X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值。

证明: 由 $u_3(\hat{\beta}^*(r)|\hat{\beta})$ 的定义可得:

$$u_3(\hat{\beta}^*(r)|\hat{\beta}) = \frac{\| \text{Cov}(\hat{\beta}^*(r)) \|_F}{\| \text{Cov}(\hat{\beta}) \|_F} = \frac{\| \sigma_1(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1} \|_F}{\| \sigma_1(X_1^T Q_1 X_1)^{-1} \|_F} = \frac{\| (X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1} \|_F}{\| (X_1^T Q_1 X_1)^{-1} \|_F} \quad (34)$$

因 $X_1^T Q_1 X_1 \geq 0$, 故存在正交阵 P , 使得:

$$P^T (X_1^T Q_1 X_1) P = \text{diag}(\theta_1, \theta_2, \dots, \theta_p), \theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0 \quad (35)$$

$$\text{则: } \| (X_1^T Q_1 X_1)^{-1} \|_F^2 = \text{tr}[(X_1^T Q_1 X_1)^{-1}] = \sum_{i=1}^p \theta_i^{-2} \quad (36)$$

又由于 $X_1^T Q_1 X_1 + r X_2^T Q_2 X_2 > 0$ ($r = \frac{\sigma_1}{\sigma_2}$), 故:

$$\| (X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1} \|_F^2 = \text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}]^2 \quad (37)$$

由引理1和引理3可得:

$$\frac{\text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}]}{\lambda_1(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)} \leq \text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}]^2 \leq \frac{\text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}]}{\lambda_p(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)} \quad (38)$$

其中:

$$\text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}] = \sum_{i=1}^p \lambda_i^{-1}(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \quad (39)$$

由引理2:

$$\lambda_i(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \geq \lambda_s(X_1^T Q_1 X_1) + r \lambda_t(X_2^T Q_2 X_2) = \theta_s + r\eta_t, \quad s+t = i+p \quad (40)$$

$$\lambda_i(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \leq \lambda_m(X_1^T Q_1 X_1) + r \lambda_n(X_2^T Q_2 X_2) = \theta_m + r\eta_n, \quad m+n \leq i+1 \quad (41)$$

则:

$$\max\{\theta_s + r\eta_t; s+t = i+p\} \leq \lambda_i(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \leq \min\{\theta_m + r\eta_n; m+n \leq i+1\} \quad (42)$$

因此:

$$\sum_{i=1}^p [\min\{\theta_m + r\eta_n; m+n \leq i+1\}]^{-1} \leq \text{tr}[(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2)^{-1}] \leq \sum_{i=1}^p [\max\{\theta_s + r\eta_t; s+t = i+p\}]^{-1} \quad (43)$$

且

$$\lambda_1(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \leq \theta_1 + r\eta_1 \quad (44)$$

$$\lambda_p(X_1^T Q_1 X_1 + r X_2^T Q_2 X_2) \geq \theta_p + r\eta_p \quad (45)$$

由式(36)和式(43)至式(45), 可得:

$$\sqrt{\frac{\sum_{i=1}^p [\min\{\theta_m + r\eta_n; m+n \leq i+1\}]^{-1}}{(\theta_1 + r\eta_1) \cdot \sum_{i=1}^p \theta_i^{-2}}} \leq u_3(\hat{\beta}^*(r)|\hat{\beta}) \leq \sqrt{\frac{\sum_{i=1}^p [\max\{\theta_s + r\eta_t; s+t = i+p\}]^{-1}}{(\theta_p + r\eta_p) \cdot \sum_{i=1}^p \theta_i^{-2}}}$$

证毕。

定理6: 设 $\hat{\beta}^*(r)$ 和 $\hat{\beta}$ 是由式(3)和式(6)确定的估计量, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值, 若 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 则:

$$u_3(\hat{\beta}^*(r)\hat{\beta}) = \sqrt{\frac{\sum_{i=1}^p (\theta_i + r\eta_i)^{-2}}{\sum_{i=1}^p \theta_i^{-2}}}$$

证明: 因 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 故存在正交矩阵 P 使得:

$$X_1^T Q_1 X_1 = P^T \Lambda P, \Lambda = \text{diag}(\theta_1, \theta_2, \dots, \theta_p)$$

$$X_2^T Q_2 X_2 = P^T \Delta P, \Delta = \text{diag}(\eta_1, \eta_2, \dots, \eta_p)$$

故有:

$$(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1} = P^T (\Lambda + r\Delta) P$$

又因:

$$\begin{aligned} & \|(X_1^T Q_1 X_1 + rX_2^T Q_2 X_2)^{-1}\|_F^2 \\ &= \text{tr}[P^T (\Lambda + r\Delta)^{-1} (\Lambda + r\Delta)^{-1} P] = \text{tr}[(\Lambda + r\Delta)^{-2}] \\ &= \text{tr} \begin{pmatrix} \frac{1}{(\theta_1 + r\eta_1)^2} & & \\ & \ddots & \\ & & \frac{1}{(\theta_p + r\eta_p)^2} \end{pmatrix} = \sum_{i=1}^p \frac{1}{(\theta_i + r\eta_i)^2} \end{aligned}$$

所以:

$$u_3(\hat{\beta}^*(r)\hat{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_F}{\|\text{Cov}(\hat{\beta})\|_F} = \sqrt{\frac{\sum_{i=1}^p (\theta_i + r\eta_i)^{-2}}{\sum_{i=1}^p \theta_i^{-2}}}$$

证毕。

用上述类似的方法, 可以得到关于 $u_4(\hat{\beta}^*(r)\tilde{\beta})$ 的上界和下界, 即:

定理7: 设 $\hat{\beta}^*(r)$ 和 $\tilde{\beta}$ 是由式(3)和式(8)给出的估计量, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值, 由 $\tilde{\beta}$ 代替 $\hat{\beta}^*(r)$ 产生的相对效率定义为:

$$u_4(\hat{\beta}^*(r)\tilde{\beta}) = \frac{\|\text{Cov}(\hat{\beta}^*(r))\|_F}{\|\text{Cov}(\tilde{\beta})\|_F}$$

则有:

$$\sqrt{\frac{\sum_{i=1}^p [\min\{\theta_m + r\eta_n; m+n \leq i+p\}]^{-1}}{(\theta_1 + r\eta_1) \cdot \sum_{i=1}^p \eta_i^{-2}}} \leq u_4(\hat{\beta}^*(r)\tilde{\beta}) \leq r \cdot \sqrt{\frac{\sum_{i=1}^p [\max\{\theta_s + r\eta_t; s+t=i+p\}]^{-1}}{(\theta_p + r\eta_p) \cdot \sum_{i=1}^p \eta_i^{-2}}}$$

$$\sqrt{\frac{\sum_{i=1}^p [\max\{\theta_s + r\eta_t; s+t=i+p\}]^{-1}}{(\theta_p + r\eta_p) \cdot \sum_{i=1}^p \eta_i^{-2}}}$$

定理8: 设 $\hat{\beta}^*(r)$ 和 $\tilde{\beta}$ 是由式(3)和式(8)给出的估计量, $\theta_1 \geq \theta_2 \geq \dots \geq \theta_p > 0$ 、 $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p > 0$ 分别是 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 的 p 个有序特征值, 若 $X_1^T Q_1 X_1$ 和 $X_2^T Q_2 X_2$ 是可交换的, 则有:

$$u_4(\hat{\beta}^*(r)\tilde{\beta}) = r \cdot \sqrt{\frac{\sum_{i=1}^p (\theta_i + r\eta_i)^{-2}}{\sum_{i=1}^p \eta_i^{-2}}}$$

4 结束语

本文针对线性回归方程系统中相同参数的三种最优线性无偏估计, 提出了在谱范数和 F 范数准则下不同无偏估计的相对效率, 分析并证明了在两种不同准则下四种相对效率的上界和下界。

参考文献:

- [1] Bloomfield P, Watson G S. The Inefficiency of Least Squares[J]. Biometrika, 1975, (62).
- [2] Wang S G, Shao J. Constrained Kantorovich Inequalities and Relative Efficiency of Least Squares[J]. Journal of Multivariate Analysis, 1992, (42).
- [3] Wang L, Pan G. The Efficiency Comparisons Between OLSE and BLUE in a Singular Linear Model[J]. Journal of Inequalities and Applications, 2013.
- [4] Liu A Y. Estimation of the Parameters in Two Linear Models With Only Some of the Parameter Vectors Identical[J]. Statistics & Probability Letters, 1996, (29).
- [5] Ma T F, Wang S G. Estimation of the Parameters in a Two Linear Regression Equations System With Identical Parameter Vectors[J]. Statistics & Probability Letters, 2009, (79).
- [6] Wu J B. The Relative Efficiency of Liu-Type Estimator in a Partially Linear Model[J]. Applied Mathematics and Computation, 2014, (243).
- [7] Yang H, Wang L T. An Alternative Form of the Watson Efficiency[J]. Journal of Statistical Planning and Inference, 2009, (139).
- [8] Wu J B. Some Properties of Relative Efficiency of Estimators in a Two Linear Regression Equations System With Identical Parameter Vectors[J]. Journal of Inequalities and Applications, 2014.
- [9] Chen Y R. Two Kinds of Relative Efficiencies of Mixed Estimator With Respect to LS Estimator in Linear Regression Model[J]. Journal of Mathematics, 2007, (27).

(责任编辑/亦 民)

控制单个随机解释变量条件下 被解释变量平均改变量的估计方法

王义闹¹, 吴利丰²

(1.温州大学 数学与信息科学学院, 浙江 温州 325035; 2.河北工程大学 经济管理学院, 河北 邯郸 056038)

摘要:文章在将两个相关随机变量中的一个适当分解为两个随机变量的基础上,对解释变量完全是随机变量的不完全多重共线问题,给出了用一个原解释变量以及与其不相关随机解释变量线性表示被解释变量的表示方法。进而研究了单独控制一个原解释变量改变一个单位条件下被解释变量的平均改变量(单控改变量),给出了估计单控改变量的方法——剔除相关变量法,证明了估计量的无偏性和一致性,并证明了估计量方差依概率收敛到有效估计量的方差。

关键词:多重共线性;分解随机变量;被解释变量;剔除相关变量法

中图分类号: O211.6 **文献标识码:** A **文章编号:** 1002-6487(2017)15-0027-05

0 引言

对不完全多重共线问题,在随机误差项同方差、不相关、解释变量与随机误差项不相关的条件下,随着多重共线严重程度的增加,参数的最小二乘估计量的方差急剧增加。为了降低参数估计量的均方误差,Hoerl A E^[1]在1962年首先放弃无偏性要求,提出了估计模型参数的岭回归方法,并与Kennard^[2]于1970年给予了详细讨论;W. F. Massy于1965年提出了主成分回归等有偏估计方法,较好地解决了参数估计问题,估计了当某解释变量改变一个单位,并且其他解释变量保持不变的条件下,被解释变量的平均改变量。文献[2-7]在总结前人研究成果的基础上,提出了一些解决多重共线性问题的思想方法,这些思想方法的目标仍是较好估计原模型中某个或某些解释变量的系数。

在所有解释变量都是可控的前提下,本文由上述方法较好地估计模型参数后,就可以通过控制解释变量达到控

制被解释变量平均改变量的目的。然而在社会、经济等问题中,有些解释变量是可控的,有些解释变量是不可控的。例如文献[8]中显示,中国粮食产量受农业化肥施用量、粮食播种面积、成灾面积的影响。其中成灾面积显然是不可控的。即使较好地估计了模型参数,也不能通过适当控制农业化肥施用量、粮食播种面积并保持成灾面积不变来达到预期的粮食产量。我们应该研究的是,在适当调整农业化肥施用量、粮食播种面积的条件下,粮食产量的平均改变量。本文试图研究最简单的一种情况:控制某解释变量改变一个单位的同时,虽然不能控制其他解释变量保持不变,但解释变量间的相关性依然保持不变,由样本估计的变量之间的定量关系外延一定范围仍成立。进而研究某解释变量改变一个单位条件下被解释变量的平均改变量。

当解释变量全部为随机变量时,不完全多重共线问题中的被解释变量由一组线性相关随机解释变量和一个与随机解释变量不相关的随机误差项所表示。本文关心的

基金项目:教育部人文社会科学研究规划项目(15YJA630017)

作者简介:王义闹(1963—),男,河北安平人,教授,研究方向:计量经济。

吴利丰(1983—),男,河北邯郸人,博士,副教授,研究方向:灰色系统理论。

Comparison The Relative Efficiencies of Parameter Estimators in Linear Regression Equations System

Wang Xiaoping¹, Zhang Zhimin², Liu Chaolin²

(1. Basic Education Department, Yibin Vocational Technical College, Yibin Sichuan 644003, China;

2. College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China)

Abstract: This paper discusses the relative efficiency and its property of identical parameter vectors under the different criterion in a linear regression equation system, and also discusses and verifies the upper and lower bound of the respective relative efficiency.

Key words: linear regression model; optimal linear unbiased estimator; relative efficiency