广义G-M模型参数估计的相对效率*

黄元亮[1]， 陈桂景[2]， 韦来生[3]

(1. 安徽大学计算机系，合肥 230039； 2. 安徽大学数学系，合肥 230039；
3. 中国科技大学商学院，合肥 230036)

摘 要：本文提出了广义G-M模型中参数β的BLUE β^*和LS估计$\hat{\beta}$的一种新的相对效率，对其与其他两种相对效率的关系及其下界进行了研究，并讨论了它与广义相关系数的关系.

关键词：广义G-M模型；相对效率；广义相关系数.
分类号：AMS(1991) 62G/CLC O212.1
文献标识码：A **文章编号**：1000-341X(2000)01-0103-06

1 引 言

对于广义Gauss-Markoff模型(简记为广义G-M型，见[4]中4 I.1)

$$Y = X\beta + \varepsilon, E(\varepsilon) = 0 \quad \mathrm{cov}(\varepsilon) = \sigma^2\Sigma, \tag{1}$$

其中$X_{n\times p}$，$\Sigma_{n\times n}$为已知矩阵，$r(X) = r \leqslant p$，$r(\Sigma) = k \leqslant n$，$\beta_{p\times 1}$，$\sigma^2 > 0$为未知参数，$\varepsilon$为$n\times 1$的随机向量. 当$r(X) = p$，$r(\Sigma) = n$时，参数$\beta$的BLUE β^*和LS估计$\hat{\beta}$的相对效率常用

$$e_1(\hat{\beta}) = \frac{|\mathrm{cov}(\beta^*)|}{|\mathrm{cov}(\hat{\beta})|} \tag{2}$$

和

$$e_2(\hat{\beta}) = \frac{\mathrm{trcov}(\beta^*)}{\mathrm{trcov}(\hat{\beta})}, \tag{3}$$

此处$|A|$表示$\det A$，$\mathrm{tr}A$表示矩阵A的迹.

相对效率e_1对X的依赖性较弱，并对广义G-M模型没有意义，e_2对X的依赖性明显提高，然而e_2的大小仅取决于$\mathrm{cov}(\hat{\beta})$和$\mathrm{cov}(\beta^*)$的对角线上的元素，与其它元素无关，也就是说与$\beta^*$与$\hat{\beta}$是否相关没有关系，相当于假设了$\varepsilon$各分量相互独立，较理论化. 本文引入另一种相对效率：以β^*与$\hat{\beta}$的各分量间的协方差平方和的算术平方根之比(即$\mathrm{cov}(\beta^*)$与$\mathrm{cov}(\hat{\beta})$的欧氏模之比)来定义它们的相对效率

$$e_3(\hat{\beta}) = \frac{\|\mathrm{cov}(\beta^*)\|}{\|\mathrm{cov}(\hat{\beta})\|}. \tag{4}$$

* **收稿日期**：1996-12-20
基金项目：国家自然科学基金资助项目(19671001)
作者简介：黄元亮(1963-)，男，土家族，湖南永顺人，硕士，安徽大学讲师.

在 e_3 的表达式中,$\mathrm{cov}(\beta^*)$,$\mathrm{cov}(\hat{\beta})$ 的每个元素都对 e_3 产生了影响,同时克服了对 X 的依赖不足和应用的局限性,避免了高阶行列式的计算,对于这种相对效率,由协方差阵的非负定性容易验证:对于 β 的任意无偏估计 $\tilde{\beta}_1$ 和 $\tilde{\beta}_2$,若 $\mathrm{cov}(\tilde{\beta}_1) > \mathrm{cov}(\tilde{\beta}) \geqslant 0$,则有 $e_3(\tilde{\beta}_1) < e_3(\tilde{\beta}_2)$,且 $e_3(\beta^*) = 1$,所以 e_3 以 1 为上确界,e_3 的值随着 $\|\mathrm{cov}(\tilde{\beta})\|$ 值的增大而减小,e_3 越大说明以 $\tilde{\beta}$ 代替 β^* 的损失越小.

在 $r(X) = p$,$r(\Sigma) = n$ 的条件下,三种相对效率有如下关系.

定理 1 i) 设 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$ 为 $\mathrm{cov}(\beta^*)$ 的特征根,$\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_p > 0$ 为 $\mathrm{cov}(\hat{\beta})$ 的特征根,若:

$$\frac{\lambda_1}{\mu_1} \leqslant \frac{\lambda_2}{\mu_2} \leqslant \cdots \leqslant \frac{\lambda_p}{\mu_p},\text{且}\frac{\lambda_1 + \lambda_2 - \sqrt{\lambda_1^2 + \lambda_2^2}}{\mu_1 + \mu_2 - \sqrt{\mu_1^2 + \mu_2^2}} \geqslant \frac{\lambda_p}{\mu_p},$$

则 $e_3(\hat{\beta}) \leqslant e_2(\hat{\beta})$;

ii) $e_3(\hat{\beta}) = 1 \Leftrightarrow e_2(\hat{\beta}) = 1 \Leftrightarrow e_1(\hat{\beta}) = 1$;

iii) $e_3(\hat{\beta}) \geqslant e_1(\hat{\beta})$,$e_2(\hat{\beta}) \geqslant e_1(\hat{\beta})$,仅当 $\mathrm{cov}(\hat{\beta})$ 等于 $\mathrm{cov}(\beta^*)$ 时不等式的等号成立.

注意到 $\|A\|^2 = \sum\limits_{i=1} \lambda_i^2(A)$,$\mathrm{tr}(A) = \sum\limits_{i=1} \lambda_i(A)$,$|A| = \prod\limits_{i=1}^{n} \lambda_i(A)$,定理 1 的证明不难,这里略去,(在本文中 $\lambda_i(A)$ 表示矩阵 A 的第 i 个顺序特征值)

由此定理知,e_1 小于 e_2 和 e_3,e_3 在一定条件下小于 e_2,且三者同时取最大值 1,这说明以 $\hat{\beta}$ 代替 β^* 所产生的误差,e_1 反映最灵敏,e_3 其次.

2 相对效率的下界

在模型(1)下,若 $r(X) = p$,由 C.R.Rao 最小二乘统一理论知 β 的 BLUE 为

$$\beta^* = (X'T^{*-}X)^- X'T^- Y,$$

其中,$T = \Sigma + XUX'$,U 为对称阵,且满足 $r(T) = r(\Sigma \vdots X)$,$T^-$ 为 T 的任一 g 逆,则

$$\mathrm{cov}(\beta^*) = \sigma^2(X'T^- X)^- X'T^- \Sigma T^- X(X'T^- X)^-,$$

U 的一种简单取法为 $U = d^2 I$,其中 d^2 满足:若 $\mu(X) \not\subset \mu(\Sigma)$,则 $d^2 \neq 0$,若 $\mu(X) \subset \mu(\Sigma)$,则 $d^2 = 0$,此时 $T = \Sigma + d^2 XX'$,易知,$\mu(X) \subset \mu(T)$,$\mu(\Sigma) \subset \mu(T)$,从而 $X'T^- X$,$X'T^- \Sigma T^- X$ 均与 T^- 的选择无关,故可将 T^- 用 T^+ 代替,则

$$\mathrm{cov}(\beta^*) = \sigma^2(X'T^+ X)^- X'T^+ \Sigma T^+ X(X'T^+ X)^-, \tag{5}$$

而 β 的 LS 估计为 $\hat{\beta}=(X'X)^- XY$,则

$$\mathrm{cov}(\hat{\beta}) = \sigma^2(X'X)^- X'\Sigma X(X'X)^-.$$

下面来讨论 $e_3(\hat{\beta})$ 在 $r(X) = p$ 的条件下的下界,为此先引入下列引理.

引理 1 设 A,B 为 n 阶实对称阵,且 $B \geqslant 0$,则有

$$\lambda_n(B)\lambda_i(A^2) \leqslant \lambda_i(ABA) \leqslant \lambda_1(B)\lambda_i(A^2) \quad i = 1,2,\cdots,n. \tag{6}$$

证明 因为 $\lambda_1(B)I_n - B \geqslant 0$,则

$$A(\lambda_1(B)I_n - B)A \geqslant 0,$$

而 $\lambda_1(B)A^2 = A(\lambda_1(B)I_n - B)A + ABA$,所以 $\lambda_1(B)A^2 \geqslant ABA \geqslant 0$.

由对称阵的特征值的单调性知：

$$\lambda_i(\lambda_1(B)A^2) \geqslant \lambda_i(ABA) \quad i = 1,2,\cdots,n, \tag{7}$$

所以 $\lambda_1(B)\lambda_i(A^2) \geqslant \lambda_i(ABA)$，同理(6)左端不等式成立.

推论 若 $A,B \in R^{n\times n}$ 且 $A>0, B\geqslant 0$，则 $\lambda_1^{-1}(A)\mathrm{tr}B \leqslant \mathrm{tr}(A^{-1}B) \leqslant \lambda_n^{-1}(A)\mathrm{tr}(B)$.

证明 因为 $\mathrm{tr}(A^{-1}B) = \mathrm{tr}(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}})$.

由引理1可知 $\lambda_n(A^{-1})\lambda_i(B) \leqslant \lambda_i(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}) \leqslant \lambda_1(A^{-1})\lambda_i(B)$，则

$$\lambda_n(A^{-1})\sum_{i=1}^{n}\lambda_i(B) \leqslant \sum_{i=1}^{n}\lambda_i(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}) \leqslant \lambda_n^{-1}(A)\sum_{i=1}^{n}\lambda_i(B),$$

即

$$\lambda_1^{-1}(A)\mathrm{tr}B \leqslant \mathrm{tr}(A^{-1}B) \leqslant \lambda_n^{-1}(A)\mathrm{tr}B.$$

引理2 设 $\Delta = \mathrm{diag}(\delta_1,\cdots,\delta_p)$，$\delta_1 \geqslant \cdots \geqslant \delta_p > 0$，$A_{n\times n} > 0$，$U$ 为 $n\times p$ 矩阵则

$$\min_{U'U=\Delta}\mathrm{tr}(U'AU) = \sum_{i=1}^{p}\lambda_{n-i+1}(A)\delta_i. \tag{8}$$

证明见[1].

定理2 在模型(1)下，若 $r(X) = p$，则

$$e_3(\hat{\beta}) \geqslant \frac{\mu_p^2}{\delta_1\mu_1^2} \times \frac{\left(\sum_{i=1}^{p}\lambda_{m-i+1}^2\delta_{p-i+1}^{-2}\right)^{1/2}}{\sum_{i=1}^{n}\delta_i^{-2}\left(\sum_{i=1}^{m}\lambda_n^2\right)^{1/2}},$$

其中 μ_i,λ_i 分别为 T、Σ 的顺序特征值，$m = r(T)$，δ_i 为 X 的顺序奇异值.

证明 因为 $\Sigma \geqslant 0, XX' \geqslant 0$，则 $T \geqslant 0$，从而存在 n 阶正交阵 Q 使得

$$QTQ' = \mathrm{diag}(\mu_1,\cdots,\mu_m,0,\cdots,0) \stackrel{\triangle}{=} \mathrm{diag}(N,0), \tag{9}$$

而 $\mu(X) \subset \mu(T)$，则 $QX = \begin{pmatrix} W \\ 0 \end{pmatrix}$，$W$ 为 $m\times p$ 矩阵，$r(W) = p \leqslant m$，则

$$\mathrm{cov}(\beta^*) = \sigma^2[(W'N^{-1}W)^{-1}W'N^{-1}\Sigma^*N^{-1}W(W'N^{-1}W)^{-1}], \tag{10}$$

其中 $\Sigma^* = N - d^2WW'$.

对于 $m\times p$ 矩阵 W，存在 $m\times p$ 列正交阵 P_1 和 p 阶正交阵 P_2，使得

$$W = P_1\Delta P_2', \Delta = \mathrm{diag}(\delta_1,\cdots,\delta_p). \tag{11}$$

由(10)，(11)得

$$\begin{aligned}\mathrm{cov}(\beta^*) &= \sigma^2P_2\Delta^{-1}(P'_1N^{-1}P_1)^{-1}P'_1N^{-1}\Sigma^*N^{-1}P_1(P'_1N^{-1}P_1)^{-1}\Delta^{-1}P'_2 \\ &\stackrel{\triangle}{=} \sigma^2P_2\Delta^{-1}A^{-1}BA^{-1}\Delta^{-1}P'_2.\end{aligned} \tag{12}$$

同时有

$$\mathrm{cov}(\hat{\beta}) = \sigma^2P_2\Delta^{-1}P_1'\Sigma^*P_1\Delta^{-1}P_2', \tag{13}$$

则

$$\mathrm{tr}(\mathrm{cov}(\beta^*))^2/\sigma^4 = \mathrm{tr}[\Delta^{-1}A^{-1}BA^{-1}\Delta^{-2}A^{-1}BA^{-1}\Delta^{-1}]. \tag{14}$$

由(14)，引理2得

$$\frac{1}{\sigma^4}\operatorname{tr}(\operatorname{cov}(\beta^*))^2 \geqslant \sum_{i=1}^{p}\lambda_{p-i+1}(A^{-1}BA^{-1}\Delta^{-2}A^{-1}BA^{-1})\lambda_i(\Delta^{-2}). \tag{15}$$

由引理 1,(15)得

$$\begin{aligned}\frac{1}{\sigma^4}\operatorname{tr}(\operatorname{cov}(\beta^*))^2 &\geqslant \sum_{i=1}^{p}\lambda_p(\Delta^{-2})\lambda_{p-i+1}^2(A^{-1}BA^{-1})\lambda_i(\Delta^{-2})\\ &= \sum_{i=1}^{p}\lambda_p(\Delta^{-2})\lambda_{p-i+1}^2(B^{\frac{1}{2}}A^{-2}B^{\frac{1}{2}})\lambda_i(\Delta^{-2})\\ &\geqslant \sum_{i=1}^{p}\lambda_p(\Delta^{-2})\lambda_p^2(A^{-2})\lambda_{p-i+1}^2(B)\lambda i(\Delta^{-2})\\ &= \sum_{i=1}^{p}\frac{1}{\delta_1^2}\lambda_p^2[(P_1'N^{-1}P_1)^{-2}]\lambda_{p-i+1}^2(P_1'N^{-1}\Sigma^*N^{-1}P_1)\delta_{p-i+1}^{-2}.\end{aligned}$$

注意到 $P_1'P_1=I_p$, 由 Poicare 分隔定理,上式变成

$$\begin{aligned}\frac{1}{\sigma^4}\operatorname{tr}(\operatorname{cov}(\beta^*))^2 &\geqslant \frac{1}{\delta_1^2}\sum_{i=1}^{p}\lambda_{m-p+1}^{-4}(N^{-1})\lambda_{m-i+1}^2(N^{-1}\Sigma^*N^{-1})\delta_{p-i+1}^{-2}\\ &= \frac{\mu_p^4}{\delta_1^2}\sum_{i=1}^{p}\lambda_m^2(N^{-2})\lambda_{m-i+1}^2(\Sigma^*)\delta_{p-i+1}^{-2}\\ &\geqslant \frac{\mu_p^4}{\delta_1^2\mu_1^4}\sum_{i=1}^{p}\lambda_{m-i+1}^2\delta_{p-i+1}^{-2},\end{aligned} \tag{16}$$

而

$$\begin{aligned}\|\operatorname{cov}(\hat\beta)\| &= \sigma^2\|P_2\Delta^{-2}P_1'\Sigma^*P_1\Delta^{-1}P_2'\| \leqslant \sigma^2\|P_1\Delta^{-1}\|^2\|\Sigma^*\|\\ &= \sigma^2\operatorname{tr}(\Delta^{-2})\|\Sigma^*\|,\end{aligned}$$

则

$$\|\operatorname{cov}(\hat\beta\|^2 \leqslant \sigma^4(\sum_{i=1}^{p}\delta_i^{-2})^2\operatorname{tr}(\Sigma^{*2}) \leqslant \sigma^4(\sum_{i=1}^{p}\delta_i^{-2})^2\sum_{i=1}^{m}\lambda_i^2. \tag{17}$$

由(4),(16),(17)得

$$e_3(\hat\beta) \geqslant \frac{\mu_p^2}{\delta_1\mu_1^2}\times\frac{(\sum_{i=1}^{p}\lambda_{m-i+1}^2\delta_{p-i+1}^{-2})^{1/2}}{\sum_{i=1}^{p}\delta_i^{-2}(\sum_{i=1}^{m}\lambda_i^2)^{1/2}}.$$

3 相对效率与广义相关系数的关系

在模型(1)下,若 $r(X)=p$, 则 β 的 LS 估计 $\hat\beta$ 和 BLUE β^* 的协方差为

$$\operatorname{cov}(\beta^*,\hat\beta) = \sigma^2(X'T^+X)^-X'T^+\Sigma X(X'X)^- \stackrel{\triangle}{=} \sigma^2 C.$$

记 $\operatorname{cov}(\beta^*)=\sigma^2A$, $\operatorname{cov}(\hat\beta)=\sigma^2B$, 则 β^* 与 $\hat\beta$ 的协方差阵为

$$\operatorname{cov}\begin{pmatrix}\beta^*\\ \hat\beta\end{pmatrix} = \sigma^2\begin{pmatrix}A & C\\ C' & B\end{pmatrix}.$$

设 ρ_i^2 为 β^* 与 $\hat\beta$ 的典型相关系数,则 ρ_i^2 满足

$$|B^+C'A^+C-\rho_i^2I| = 0. \tag{18}$$

张氏广义相关系数有

$$r_Z^{(1)}=\frac{1}{p}\sum_{i=1}^{p}\rho_i^2=\frac{1}{p}\mathrm{tr}(B^+C'A^+C),\tag{19}$$

关于 $e_3(\hat\beta)$ 与 $r_Z^{(1)}$ 有

定理 3　在模型(1)下,若 $r(X)=p,r(\Sigma)=k\geqslant p$,则

$$(\frac{\mu_m}{\mu_1})^4\sqrt{p}(\frac{\sigma_m}{\sigma_1})^4(\frac{\lambda_m}{\lambda_1})^2\frac{r_Z^{(1)}}{(\sum_{i=1}^{m}\lambda_i^2)^{\frac{1}{2}}(\sum_{i=1}^{p}\sigma_i^{-2})}\leqslant e_3(\hat\beta)\leqslant\sqrt{p}\,\frac{\lambda_1^{3/2}\sigma_1^4\mu_1^2}{\lambda_m^4\mu_m^2\sigma_p^2}r_Z^{(1)},\tag{20}$$

其中 $\lambda_i=\lambda_i(\Sigma),\mu_i=\lambda_i(T),\delta_i^2=\lambda_i(XX'),m=r(T)$.

证明　由初等不等式$\frac{1}{n}(a_1^2+\cdots+a_n^2)\geqslant(\frac{1}{n}(a_1+\cdots+a_n))^2$,有

$$\|e_3(\hat\beta)\|^2=\mathrm{tr}A^2/\mathrm{tr}B^2\geqslant\frac{1}{p}(\mathrm{tr}A)^2/\mathrm{tr}B^2.\tag{21}$$

由 $r(X)=p,r(\Sigma)=k\geqslant p$,则 A,B 为非奇异的,且 $r(T)=r(\Sigma)$ 而

$$\mathrm{tr}(A)=\mathrm{tr}(A^{-1}\cdot A^2).\tag{22}$$

由引理 1 的推论、(22)得

$$\mathrm{tr}(A)\geqslant\mathrm{tr}(A^{-1})\lambda_p^2(A).\tag{23}$$

再由引理 1 及其推论、(23)得

$$\begin{aligned}\mathrm{tr}(A)&\geqslant\lambda_1^{-1}(CC')\mathrm{tr}(A^{-\frac{1}{2}}CC'A^{-\frac{1}{2}})\lambda_p^2(A)=\lambda_1^{-1}(CC')\lambda_p^2(A)\mathrm{tr}(C'A^{-1}C)\\&\geqslant\lambda_1^{-1}(CC')\lambda_p^2(A)\lambda_p(B)\mathrm{tr}(B^{-1}C'A^{-1}C).\end{aligned}\tag{24}$$

容易验证

$$\lambda_1(CC')\leqslant\mu_1^2\lambda_1^2/(\delta_p^4\mu_m^2)\tag{25}$$

$$\lambda_p(A)\geqslant\lambda_m\delta_1^{-1}\mu_m^2/\mu_1^2,\ \lambda_p(B)\geqslant\lambda_m\sigma_1^{-2}\tag{26}$$

$$\mathrm{tr}B^2\leqslant\sum_{i=1}^{k}\lambda_i^2(\sum_{i=1}^{p}\delta_i^{-2})^2.\tag{27}$$

由(21),(24),(25),(26),(27)得

$$\|e_3(\hat\beta)\|\geqslant\sqrt{p}(\frac{\delta_m}{\sigma_1})^4(\frac{\mu_m}{\mu_1})^4(\frac{\lambda_m}{\lambda_1})^2\frac{r_Z^{(1)}}{(\sum_{i=1}^{m}\lambda_i^2)^{1/2}(\sum_{i=1}^{p}\delta^{-2})}.$$

注意到 $\|e_3(\hat\beta)\|^2\leqslant\frac{(\mathrm{tr}pA)^2}{p\lambda_p^2(B)}$,用类似方法可得(20)右端不等式.

作为定理 3 的特例有

定理 4　在模型(1)下,若 $r(X)=p,r(\Sigma)=n$,则

$$\sqrt{p}\,\frac{\lambda_n\delta_1^{-1}}{\lambda_1(\sum_{i=1}^{p}\delta_i^{-2})^{1/2}}r_Z^{(1)}\leqslant e_3(\hat\beta)\leqslant\sqrt{p}\,\frac{\lambda_1\delta_1}{\lambda_n\delta_n}r_Z^{(1)}.$$

参考文献:

[1]　刘爱义,王松桂. 线性模型中最小二乘估计的一种新的相对效率 [J]. 应用概率统计,1989 年 5 月.

[2] 王松桂. 线性模型的理论及其应用 [M]. 安徽:安徽教育出版社，1987 年.
[3] 倪国熙. 常用的矩阵理论和方法 [M]. 上海:上海科技出版社，1984 年.
[4] RAO C R. *Linear Statistical inferences and its Applications* [M]. John, Wiley, 1971.

The Relative Efficiency of the Parametor Estimate in Generalized G-M Model

HUANG Yuan-liang[1], CHEN Gui-jing[2], WEI Lai-sheng[3]

(1. Dept. of Computer, Anhui University, Hefei 230039; 2. Dept. of Math., Anhui University, Heifei 230039;
3. University of Science and Technology of China, Hefei 230026)

Abstract: In this paper, we give a new relative efficiency in generalized G-M model, study the relative efficiency's lower bound. And discusse the relation between the relative efficiency and generalized relative coefficient.

Keywords: relative efficiency; generalized G-M model; generalized coefficient.