应用概率统计 第十卷
第二期 1994年5月

Chinese Journal of Applied Probability
and Statistics Vol. 10, No. 2, May, 1994

一类自适应岭估计的小样本性质

严利清 王松桂
(中国科学技术大学,合肥,230026)

摘 要

本文利用比较简洁的方法研究了线性模型中关于回归系数的一类自适应岭估计的精确偏差和均方误差,并通过对均方误差的分析得到了该估计类一致优于最小二乘估计的充分条件.

§1. 引 言

考虑线性模型

$$Y=X\beta+e \tag{1.1}$$

其中Y是n维观察向量,X是$n\times p$列满秩设计阵,β是p维未知回归系数,随机误差$e\sim N(0,\sigma^2I_n)$($\sigma^2>0$未知). 为估计β通常考虑最小二乘估计(LS估计)$\hat{\beta}=(X'X)^{-1}X'Y$. 但是当设计阵病态时$LS$估计的效果往往很不理想,为此Hoerl和Kennard(1970)在$p>2$时提出了著名的岭估计$\hat{\beta}(k)=(X'X+kI)^{-1}X'Y$, $k>0$,来改善LS估计,由于k是非随机常数时,$\hat{\beta}(k)$只能局部地改善LS估计,我们通常选择k依赖于观察向量Y并称这时的岭估计为自适应岭估计. 岭估计的研究和应用一直受到了广泛的重视且它已成为目前最有影响的一种有偏估计. 但由于问题本身的复杂性,自适应岭估计的小样本性质的研究相对来说要少得多. 例如Sawa(1972), Kadiyala(1980)和Firinguetti(1987)讨论了一些比较特殊的自适应岭估计的小样本性质.

为方便起见先把模型(1.1)表示成典则形式.设$\lambda_1\geqslant\cdots\geqslant\lambda_p$是$X'X$的特征根,$\phi_1,\cdots,\phi_p$是相应的标准正交特征向量,记$\Phi=(\phi_1,\cdots,\phi_p)$, $Z=X\Phi$, $\alpha=\Phi'\beta$,则模型(1.1)的典则形式为

$$Y=Z\alpha+e \tag{1.2}$$

由(1.2)得到α的LS估计$\hat{\alpha}=(\hat{\alpha}_1,\cdots,\hat{\alpha}_p)'=(Z'Z)^{-1}Z'Y=\Phi'\hat{\beta}$, σ^2的LS估计

$$\hat{\sigma}^2=\frac{1}{n-p}(Y-Z\hat{\alpha})'(Y-Z\hat{\alpha}).$$

记$\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_p)$,定义$\alpha$的广义岭估计为$\hat{\alpha}(K)=(\Lambda+K)^{-1}\Lambda\hat{\alpha}$,其中

$$K=\mathrm{diag}(k_1,\cdots,k_p),\ k_i\geqslant 0,$$

回归系数β的广义岭估计定义为$\hat{\beta}(K)=\Phi\hat{\alpha}(K)$. 由于

$$E\hat{\beta}(K)=\Phi E\hat{\alpha}(K),\ MSE(\hat{\beta}(K))=E(\hat{\beta}(K)-\beta)'(\hat{\beta}(K)-\beta)=MSE(\hat{\alpha}(K)),$$

故只需在模型(1.2)下考虑典则参数α的岭估计. 下面定义一个本文研究的自适应广义岭估

本文1992年9月26日收到,1993年4月10日收到修改稿.

计，它包含了常见的岭估计，如 Stein 型估计，Vinod 和 Ullah(1981)提出的双 h 类估计，Wang 和 Chow(1990)提出的双 l 类估计，它还包含了几个新的有意义的岭估计.

定义

$$\alpha^*(W)=(\alpha_1^*(W),\ \cdots,\ \alpha_p^*(W))'=(\Lambda+\hat{K})^{-1}\Lambda\hat{\alpha} \tag{1.3}$$

其中 $\hat{K}=\mathrm{diag}(\hat{k}_1,\ \cdots,\ \hat{k}_p)$，$k_i=\dfrac{\lambda_i b_i\hat{\sigma}^2}{\hat{\alpha}'W\hat{\alpha}(n-p)^{-1}+(a_i-b_i)\hat{\sigma}^2}$，

$W=\mathrm{diag}(\eta_1,\ \cdots,\ \eta_p)$常数 a_i，b_i，η_i 都大于零．如果记

$$V=\frac{\hat{\alpha}'W\hat{\alpha}}{(n-p)\hat{\sigma}^2},\quad G_i=\frac{1}{a_i+V},$$

则

$$\alpha_i^*(W)=(1-b_iG_i)\hat{\alpha}_i,\quad i=1,\ 2,\ \cdots,\ p.$$

本文将在第二节给出 $\alpha^*(W)$ 的精确偏差和均方误差，在第三节给出 $\alpha^*(W)$ 在均方误差意义下一致优于 LS 估计的充分条件，在第四节给出几个应用.

§2. $\alpha^*(W)$的精确偏差和均方误差

先给出几个引理

引理 2.1 设随机向量 $Y=(Y_1,\ \cdots,\ Y_n)'\backsim N(\mu,\ I_n)$，记 $\mu=(\mu_1,\ \cdots,\ \mu_n)'$，可测函数 $f(x)$满足 $Ef^2(Y)<\infty$，则

$$E(Y_i-\mu_i)f(Y)=\frac{\partial}{\partial\mu_i}Ef(Y)$$

$$E(Y_i-\mu_i)^2f(Y)=Ef(Y)+\frac{\partial^2}{\partial\mu_i^2}Ef(Y)$$

证明 利用陈希孺(1981)定理 1.2.1 即得.

引理 2.2 设 G_i 是 $\hat{\alpha}$, $\hat{\sigma}^2$ 可测函数 $EG_i^2<\infty$，记

$$\alpha_i^*=(1-G_i)\hat{\alpha}_i,\quad \delta=\frac{1}{2}\sum_{i=1}^{p}\mu_i^2,\quad \mu_i=\sqrt{\lambda_i}\ \alpha_i/\sigma$$

则

$$E\alpha_i^*=\alpha_i-\alpha_i\left(\frac{\partial}{\partial\delta}+1\right)EG_i$$

$$E(\alpha_i^*-\alpha_i)^2=\frac{\sigma^2}{\lambda_i}+\frac{\sigma^2}{\lambda_i}\left\{\mu_i^2\left[\left(\frac{\partial}{\partial\delta}+1\right)^2EG_i^2-2\frac{\partial}{\partial\delta}\left(\frac{\partial}{\partial\delta}+1\right)EG_i\right]+\left(\frac{\partial}{\partial\delta}+1\right)E(G_i^2-2G_i)\right\}$$

证明 因为 $\dfrac{\sqrt{\lambda_i}\ \hat{\alpha}_i}{\sigma}\backsim N\left(\dfrac{\sqrt{\lambda_i}\ \alpha_i}{\sigma},\ 1\right)$，利用引理 2.1 并注意到

$$\frac{\partial}{\partial\mu_i}=\mu_i\frac{\partial}{\partial\delta},\quad \frac{\partial^2}{\partial\mu_i^2}=\mu_i^2\frac{\partial^2}{\partial\delta^2}+\frac{\partial}{\partial\delta}\text{即可.}$$

$$\begin{aligned}E\alpha_i^*&=E(1-G_i)\hat{\alpha}_i=E(\hat{\alpha}_i-G_i(\hat{\alpha}_i-\alpha_i)-G_i\alpha_i)\\&=\alpha_i-\frac{\sigma}{\sqrt{\lambda_i}}E\left(\frac{\sqrt{\lambda_i}\ \hat{\alpha}_i}{\sigma}-\frac{\sqrt{\lambda_i}\ \alpha_i}{\sigma}\right)G_i-\alpha_iEG_i\\&=\alpha_i-\frac{\sigma}{\sqrt{\lambda_i}}\frac{\partial}{\partial\mu_i}EG_i-\alpha_iEG_i=\alpha_i-\alpha_i\left(\frac{\partial}{\partial\delta}+1\right)EG_i\end{aligned}$$

$$E(\alpha_i^*-\alpha_i)^2=E[(\hat{\alpha}_i-\alpha_i)(G_i-1)+\alpha_iG_i]^2$$

将上式展开后利用前面相同方法即可.

由引理2.2知道计算$\alpha^*(W)$的精确偏差和均方误差的关键在于计算出$G_i=\frac{1}{a_i+V}$, $\left(V=\frac{\hat{a}'W\hat{a}}{(n-p)\hat{\sigma}^2}\right)$的精确一、二阶矩. 为此需要下面引理.

引理2.3 设随机变量X_1, $X_2(X_2>0)$的联合矩母函数

$$M(t_1,\ t_2)=E\exp(t_1X_1+t_2X_2),$$

在$-\varepsilon<t_1<\varepsilon$, $-\infty<t_2<\varepsilon(\exists\varepsilon>0)$内存在, 如果对自然数$k$

$$\int_{-\infty}^{0}(-t_2)^{k-1}\left.\frac{\partial^k M(t_1,\ t_2)}{\partial t_1^k}\right|_{t_1=0}dt_2<\infty$$

则$E\left(\frac{X_1}{X_2}\right)^k$存在, 且

$$E\left(\frac{X_1}{X_2}\right)^k=\int_{-\infty}^{0}\frac{(-t_2)^{k-1}}{(K-1)!}\left.\frac{\partial^k M(t_1,\ t_2)}{\partial t_1^k}\right|_{t_1=0}dt_2$$

证明 利用 Fubini 定理即可.

引理2.4 设随机向量$Y\backsim N(\mu,\ I_n)$, 非负定矩阵A, B, 记$X_1=Y'AY$, $X_2=Y'BY$, 则X_1, X_2的联合矩母函数$M(t_1,\ t_2)$在$-\varepsilon<t_1<\varepsilon$, $-\infty<t_2<\varepsilon$内存在$(\exists\varepsilon>0)$, 且

$$M(t_1,\ t_2)=|T|^{-\frac{1}{2}}\exp\left\{\frac{1}{2}\mu'(T^{-1}-I_n)\mu\right\}$$

其中$T=I_n-2t_1A-2t_2B$.

证明 利用正态分布密度函数直接计算得到. 取ε充分小使T为正定矩阵.

下面引进一个函数

$$G(a_i)=\int_0^1\left((1-x)^{\frac{n}{2}-1}(1-C_ix)^{-\frac{n-p}{2}-1}\prod_{k=1}^{p}\left(1-x+\frac{\eta_k}{\lambda_k}x\right)^{-\frac{1}{2}}\exp\left\{-\sum_{k=1}^{p}\frac{\delta_k\frac{\eta_k}{\lambda_k}x}{1+\left(\frac{\eta_k}{\lambda_k}\right)-1)x}\right\}dx\right.$$

其中 $$C_i=1-a_i,\ a_i>0,\ \delta_k=\frac{\lambda_k\alpha_k^2}{2\sigma^2},\ i,\ k=1,\ 2,\ \cdots,\ p,$$

引理2.5 $EG_i=\frac{n-p}{2}G(a_i)$, $EG_i^2=\frac{\partial}{\partial C_i}EG_i=\frac{n-p}{2}\frac{\partial}{\partial C_i}G(a_i)$.

证明 记$P_Z=I-Z(Z'Z)^{-1}Z'$, $A=P_Z$, $B=a_iP_Z+Z(Z'Z)^{-1}W(Z'Z)^{-1}Z'$, $X_1=\sigma^{-2}Y'AY$, $X_2=\sigma^{-2}Y'BY$, 则$G_i=\frac{X_1}{X_2}$, 由引理2.4知. 存在$\varepsilon>0$使得X_1, X_2的矩母函数$M(t_1,\ t_2)$在$-\varepsilon<t_1<\varepsilon$, $-\infty<t_2<\varepsilon$内存在且

$$M(t_1,\ t_2)=|T|^{-\frac{1}{2}}\exp\left\{\frac{1}{2}\mu'(T^{-1}-I_n)\mu\right\},$$

其中 $$\mu=Z\alpha/\sigma,\ T=I_n-(2t_1+2a_it_2]P_Z-2t_2Z\Lambda^{-1}W\Lambda^{-1}Z'.$$

记 $$D=(I-2t_2W\Lambda^{-1})^{-1}2t_2W$$

则 $$T^{-1}=I_n+\frac{2t_1+2a_it_2}{1-2t_1-2a_it_2}P_Z+Z\Lambda^{-1}D\Lambda^{-1}Z'$$

从而 $$\mu'(T^{-1}-I_n)\mu=\alpha'D\alpha/\sigma^2=2t_2\alpha'(W^{-1}-2t_2\Lambda^{-1})^{-1}\alpha/\sigma^2,$$

又因为 $$|T|=(1-2t_1-2a_it_2)^{n-p}\prod_{k=1}^{p}\left(1-2\frac{\eta_k}{\lambda_k}t_2\right)$$

由引理2.3得

$$EG_i=E\frac{X_1}{X_2}=\int_{-\infty}^{0}\frac{\partial}{\partial t_1}M(t_1,\ t_2)\big|_{t_1=0}\,dt_2$$

$$=\int_{-\infty}^{0}(n-p)(1-2a_it_2)^{-\frac{n-p}{2}-1}\prod_{k=1}^{p}\left(1-2\frac{\eta_k}{\lambda_k}t_2\right)^{-\frac{1}{2}}$$

$$\cdot\exp\left\{\sum_{k=1}^{p}\frac{2\delta_k\eta_kt_2}{\lambda_k-2t_2\eta_k}\right\}dt_2$$

$$=\frac{n-p}{2}\int_0^1(1-x)^{\frac{n}{2}-1}(1-C_ix)^{-\frac{n-p}{2}-1}\prod_{k=1}^{p}\left(1-x+\frac{\eta_k}{\lambda_k}x\right)^{-\frac{1}{2}}$$

$$\cdot\exp\left\{-\sum_{k=1}^{p}\frac{\delta_k\frac{\eta_k}{\lambda_k}x}{1+\left(\frac{\eta_k}{\lambda_k}-1\right)x}\right\}dx$$

$$=\frac{n-p}{2}G(a_i)$$

又因为
$$\left(\frac{1}{a_i+V}\right)^2=\frac{\partial}{\partial C_i}\left(\frac{1}{a_i+V}\right),$$

所以
$$EG_i^2=E\left(\frac{1}{a_i+V}\right)^2=E\frac{\partial}{\partial C_i}\left(\frac{1}{a_i+V}\right)=\frac{\partial}{\partial C_i}EG_i=\frac{n-p}{2}\frac{\partial}{\partial C_i}G(a_i)$$

其中
$$C_i=1-a_i.$$

由于 $G_i<a_i^{-1}$ 所以上面运算顺序的交换是合理的.

下面给出本节的主要结论

定理 2.1 对于(1.3)所定义的自适应岭估计 $\alpha^*(W)$,有

$$E\alpha_i^*(W)=\alpha_i-\frac{n-p}{2}b_i\alpha_i\left(\frac{\partial}{\partial\delta_i}+1\right)G(a_i)$$

$$E(\alpha_i^*(W)-\alpha_i)^2=\frac{\sigma^2}{\lambda_i}+\frac{\sigma^2}{\lambda_i}\frac{n-p}{2}b_i\left\{2\delta_i\left[\left(\frac{\partial}{\partial\delta_i}+1\right)^2b_i\frac{\partial}{\partial C_i}-2\frac{\partial}{\partial\delta_i}\left(\frac{\partial}{\partial\delta_i}+1\right)\right]+\left(\frac{\partial}{\partial\delta_i}+1\right)\left(b_i\frac{\partial}{\partial C_i}-2\right)\right\}G(a_i)$$

证明 利用引理 2.2 和 2.5 并注意到 $\frac{\partial}{\partial\delta_i}=\frac{\partial}{\partial\delta}$ 即可.

由于 W 在一般情形下结论比较复杂, 下面在 $W=\Lambda$ 下进步讨论 $\alpha^*(W)$ 的精确矩. 先引进一个函数,

$$g_{uv}^i=g_{uv}(a_i)=\int_0^1(1-t)^{\frac{n}{2}+u-2}(1-C_it)^{-\frac{n-p}{2}-v}e^{-\delta t}dt$$

其中 $C_i=1-a_i^{-1}$, $a_i>0$, $\delta=\sum_{k=1}^{p}\delta_k$, u、v 为自然数.

显然在 $W=\Lambda$ 时, 即 $\lambda_i=\eta_i$, $i=1, 2, \cdots, p$ 时, 有 $g_{11}(a_i)=G(a_i)$. 注意到

$$\frac{\partial}{\partial\delta}g_{uv}^i=g_{u+1,v}^i-g_{uv}^i,$$

即
$$\left(\frac{\partial}{\partial\delta}+1\right)g_{uv}^i=g_{u+1,v}^i,\quad\left(\frac{\partial}{\partial\delta}+1\right)^2g_{uv}^i=g_{u+2v}^i,$$

由定理 2.1 得到

推论 2.1 如果 $W=\Lambda$,则对(1.3)的 $\alpha^*(W)$有

1) $E\alpha_1^*(\Lambda)=\alpha_i-\alpha_i\frac{n-p}{2}b_ig_{21}^i$

2) $E(\alpha_i^*(\Lambda)-\alpha_i)^2=\frac{\sigma^2}{\lambda_i}+\frac{\sigma^2}{\lambda_i}\frac{n-p}{2}b_i\left\{2\delta_i\left(b_i\frac{\partial}{\partial C_i}g_{31}^i-2\frac{\partial}{\partial\delta}g_{2i}\right)+\left(b_i\frac{\partial}{\partial C_i}-2\right)g_{21}^i\right\}$

3) $-\frac{n-p}{p}b_i\leqslant\frac{E\alpha_i^*(\Lambda)-\alpha_i}{\alpha_i}\leqslant 0$

4) $\lim\limits_{\delta\to\infty}E\frac{\alpha_i^*(\Lambda)-\alpha_i}{\alpha_i}=0$

证明 关于3)、4)证明只须注意到两个事实

$$0\leqslant g_{21}^i\leqslant\frac{2}{p},\quad \lim_{\delta\to\infty}g_{21}^i=0$$

在 $W=\Lambda$ 时如果假设 $a_i<2$，则 EG_i 有一个简单的算法，我们将它表成如下定理。

定理 2.2 当 $W=\Lambda$，$a_i<2$ 时，$EG_i=\frac{n-p}{2}g_{11}^i$.

证明 当 $W=\Lambda$ 时，$\frac{n-p}{p}V$ 服从自由度为 $(p,\ n-p)$，非中心参数为 2δ 的 F 分布，利用 F 分布的密度函数得

$$E\frac{1}{(1+V)^k}=\frac{1}{B\left(\frac{n-p}{2},\ k\right)}\int_0^1(1-t)^{\frac{n}{2}-1}t^{k-1}e^{-\delta t}dt,\quad k\text{ 为自然数}$$

其中 $B(r,\ t)=\int_0^1 x^{r-1}(1-x)^{t-1}dx$.

当 $a<2$，$x>0$ 时，

$$\frac{1}{a+x}=\sum_{k=1}^{\infty}\frac{C^k}{(1+x)^k},\quad C=1-a$$

当 $|x|<1$，$l>0$ 时，

$$\frac{l}{(1-x)^{l+1}}=\sum_{k=1}^{\infty}\frac{x^{k-1}}{B(l,\ k)}$$

所以

$$\begin{aligned}EG_i&=E\frac{1}{a_i+V}=E\sum_{k=1}^{\infty}\frac{C_i^{k-1}}{(1+V)^k}=\sum_{k=1}^{\infty}E\frac{C_i^{k-1}}{(1+V)^k}\\&=\sum_{k=1}^{\infty}\frac{C_i^{k-1}}{B\left(\frac{n-p}{2},\ k\right)}\int_0^1 e^{-\delta t}t^{k-1}(1-t)^{\frac{n}{2}-1}dt\\&=\frac{n-p}{2}\int_0^1 e^{-\delta t}(1-t)^{\frac{n}{2}-1}(1-C_it)^{-\frac{n-p}{2}-1}dt=\frac{n-p}{2}g_{11}^i\end{aligned}$$

证毕.

§3. $\alpha^*(W)$ 一致优于 LS 估计的充分条件

关于自适应岭估计一致优于 LS 估计的充分条件的研究，以往的工作是讨论某个特殊类型的岭估计，它们通常含有一个或二个参数，如 Stein 型估计，双 h 类估计等. 本节通过对均方误差的分析得到了含有 $2p$ 个参数的岭估计 $\alpha^*(W)$ 一致优于 LS 估计的充分条件. 先介绍两个引理.

引理 3.1 设随机变量 $X\backsim N(\mu,\ \sigma^2)$，实函数 $f(x)$ 几乎处处一阶可微，如果 $E|f'(X)|<\infty$ 则 $E(X-\mu)f(X)$ 存在

且
$$E(X-\mu)f(X)=\sigma^2Ef'(X)$$

证明 见王松桂(1987)第304页.

引理3.2 对常数 $b>0$, 和与 $\hat{\sigma}^2$ 独立的非负随机变量 a, 有

$$E\frac{\hat{\sigma}^4}{(a+b\hat{\sigma}^2)^2}\leqslant\frac{n-p+2}{n-p}E\frac{\sigma^2\hat{\sigma}^2}{(a+b\hat{\sigma}^2)^2}$$

证明 利用严利清(1993)中引理3.3即可.

定理3.1 对于(1.3)定义的 $\alpha^*(W)$, 如果参数 a_i, b_i 满足

$$\max_i a_i^{-2}b_i\left(\frac{b_i}{\eta_i}(n-p+2)+\frac{4}{\lambda_i}\right)\leqslant 2\min_i a_i^{-2}\sum_{i=1}^{p}\frac{b_i}{\lambda_i} \tag{3.1}$$

则 $\alpha^*(W)$ 在均方误差意义下一致优于 LS 估计, 即

$$MSE(\alpha^*(W))\leqslant MSE(\hat{\alpha})$$

注 使(3.1)成立的必要条件是 $p>2$, 且在 $p>2$ 时, 总可以选取恰当的 a_i, b_i 使(3.1)成立. 而这一点是以前给出的充分条件所不具有的, 见严利清(1993).

定理的证明 记 $g_i=\hat{\alpha}'W\hat{\alpha}+a_i(n-p)\hat{\sigma}^2$, 则

$$\alpha_i^*(W)=\left(1-\frac{b_i(n-p)\hat{\sigma}^2}{g_i}\right)\hat{\alpha}_i$$

因为 $\hat{\alpha}_i\sim N\left(\alpha_i,\ \frac{\sigma^2}{\lambda_i}\right)$, 由引理3.1得

$$\begin{aligned}E(\hat{\alpha}_i-\alpha_i)\frac{\hat{\sigma}^2}{g_i}\hat{\alpha}_i&=E\hat{\sigma}^2E[(\hat{\alpha}_i-\alpha_i)g_i^{-1}\hat{\alpha}_i\,|\,\hat{\sigma}^2]=E\hat{\sigma}^2\frac{\sigma^2}{\lambda_i}E(g_i^{-1}-2\eta_i\hat{\alpha}_i^2g_i^{-2})\,|\,\hat{\sigma}^2]\\&=\frac{\sigma^2}{\lambda_i}E\left(\frac{\hat{\sigma}^2}{g_i}-\frac{2\eta_i\hat{\alpha}_i^2\hat{\sigma}^2}{g_i^2}\right)\end{aligned}$$

由引理3.2得

$$E\hat{\alpha}_i^2\frac{\hat{\sigma}^4}{g_i^2}\leqslant\frac{n-p+2}{n-p}\sigma^2E\frac{\hat{\alpha}_i^2\hat{\sigma}^2}{g_i^2},$$

所以

$$\begin{aligned}E(\alpha_i^*(W)-\alpha_i)^2&=E(\hat{\alpha}_i-\alpha_i-(n-p)b_i\hat{\sigma}^2g_i^{-1}\hat{\alpha}_i)^2\\&=\frac{\sigma^2}{\lambda_i}-2b_i(n-p)E(\hat{\alpha}_i-\alpha_i)\frac{\hat{\sigma}^2}{g_i}\hat{\alpha}_i+b_i^2(n-p)^2E\hat{\alpha}_i^2\frac{\hat{\sigma}^4}{g_i^2}\\&\leqslant\frac{\sigma^2}{\lambda_i}+(n-p)\sigma^2\left\{b_i\left(\frac{b_i}{\eta_i}(n-p+2)+\frac{4}{\lambda_i}\right)E\frac{\hat{\sigma}^2\eta_i\hat{\alpha}_i^2}{g_i^2}-2\frac{b_i}{\lambda_i}E\frac{\hat{\sigma}^2}{g_i}\right\}\\&\leqslant\frac{\sigma^2}{\lambda_i}+(n-p)\sigma^2\left\{\frac{a_0^2}{a_i^2}b_i\left(\frac{b_i}{\eta_i}(n-p+2)+\frac{4}{\lambda_i}\right)E\frac{\hat{\sigma}^2\eta_i\hat{\alpha}_i^2}{g_0^2}-2\frac{b_i}{\lambda_i}E\frac{\hat{\sigma}^2}{g_0}\right\}\end{aligned}$$

这儿 $g_0=\max\limits_i g_i$, $a_0=\max\limits_i a_i$, 上式由于 $g_0\geqslant g_i\geqslant\frac{a_i}{a_0}g_0$. 于是

$$\begin{aligned}&MSE(\alpha^*(W))-MSE(\hat{\alpha})\\&=\sum_{i=1}^{p}\left(E(\alpha_i^*(W)-\alpha_i)^2-\frac{\sigma^2}{\lambda_i}\right)\\&\leqslant(n-p)\sigma^2\left\{a_0^2\max_i a_i^{-2}b_i\left(b_i\eta_i^{-1}(n-p+2)+\frac{4}{\lambda_i}\right)E\frac{\hat{\sigma}^2\hat{\alpha}'W\hat{\alpha}}{g_0^2}-2\sum_{i=1}^{p}\frac{b_i}{\lambda_i}E\frac{\hat{\sigma}^2}{g_0}\right\}\\&\leqslant(n-p)\sigma^2\left\{a_0^2\max_i a_i^{-2}b_i\left(\frac{b_i}{\eta_i}(n-p+2)+\frac{4}{\lambda_i}\right)-2\sum_{i=1}^{p}\frac{b_i}{\lambda_i}\right\}E\frac{\hat{\sigma}^2}{g_0}\end{aligned}$$

这样就证明了(3.1)是 $MSE(\alpha^*(W))\leqslant MSE(\hat{\alpha})$ 成立的充分条件.

§4. 几个例子

本节通过选取恰当的参数 a_i, b_i 和 W 使得(1.3)所定义的 $\alpha^*(W)$ 成为常见的岭估计或有意义的新的岭估计. 从而, 由定理 2.1 和推论 2.1 可以得到它们的精确偏差和均方误差, 由定理 3.1 可以得到该岭估计一致优于 LS 估计的充分条件.

1) 如果取 $W=\Lambda$, $a_i=\dfrac{1-k_2}{n-p}$, $(k_2<1)$, $b_i=\dfrac{k_1}{n-p}$, 则 $\alpha^*(W)$ 就是 Stein 型估计

$$\hat{\alpha}_{(k_1,k_2)}=\left(1-\frac{k_1\hat{\sigma}^2}{\hat{\alpha}'\Lambda\hat{\alpha}+(1-k_2)\hat{\sigma}^2}\right)\hat{\alpha}$$

这时充分条件(3.1)为 $0\leqslant k_1\leqslant\dfrac{2(n-p)}{n-p+2}\left(\lambda_p\sum\limits_{i=1}^{p}\lambda_i^{-1}-2\right)$ 这和王松桂(1987)第 314 页上结果是相同的.

2) 如果取 $a_i=\dfrac{h_1+h_2\lambda_i}{(n-p)\lambda_i}$, $b_i=\dfrac{h_1}{(n-p)\lambda_i}$, 则 $\alpha^*(W)$ 就是 Vinod 和 Ullah(1981)提出的双 h 类估计 $\hat{\alpha}_{(h_1,h_2)}=(\Lambda+\hat{k}I)^{-1}\Lambda\hat{\alpha}$, 其中 $\hat{k}=\dfrac{h_1\hat{\sigma}^2}{\hat{\alpha}'W\hat{\alpha}+h_2\hat{\sigma}^2}$, 充分条件(3.1)为

$$\max_i\frac{1}{(h_1+h_2\lambda_i)^2}\left(\frac{n-p+2}{n-p}\,\frac{h_1}{\eta_i}+4\right)\leqslant\frac{2\lambda_p^2\sum\limits_{i=1}^{p}\frac{1}{\lambda_i^2}}{(h_1+h_2\lambda_p)^2}\tag{4.1}$$

而王松桂(1987)第 315 页给出的充分条件为

$$h_2\geqslant 0,\ 0<h_1\leqslant\frac{2(n-p)}{n-p+2}\min_i\eta_i\left(\lambda_p^2\sum_{i=1}^{p}\lambda_i^{-2}-2\right)\tag{4.2}$$

利用 $\max\limits_i a_ib_i\leqslant\max\limits_i a_i\max\limits_i b_i$ 这个事实可知由(4.2)可以推出(4.1). 即定理 3.1 给出的充分条件在有些情形下比原来的要宽一些.

3) 如果取 $W=I_p$, $a_i=\dfrac{l_1-l_2}{(n-p)\lambda_i}$, $b_i=\dfrac{l_1}{(n-p)\lambda_i}$, 则 $\alpha^*(W)$ 就是 Wang 和 Chow(1990)提出的双 l 类估计 $\hat{\alpha}_{(l_1,l_2)}=(\Lambda+\hat{K})^{-1}\Lambda\hat{\alpha}$, 其中

$$\hat{K}=\mathrm{diag}(\hat{k}_1,\ \cdots,\ \hat{k}_p),\quad \hat{k}_i=\frac{l_1\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}-l_2\hat{\sigma}^2/\lambda_i}.$$

考察充分条件(3.1)后可以得到和 2)相似的结论.

4) 如果取 $a_i=a^{-1}>0$, $b_i=b\lambda_i^{-1}(n-p)^{-1}$, 则得到一个新的岭估计

$$\alpha^*(W)=(\alpha_1^*,\ \cdots,\ \alpha_p^*)$$

其中

$$\alpha_1^*=\left(1-\frac{ab\hat{\sigma}^2/\lambda_i}{a\hat{\alpha}'W\hat{\alpha}+(n-p)\hat{\sigma}^2}\right)\hat{\alpha}_i$$

充分条件(3.1)为

$$\max_i\frac{1}{\lambda_i^2}\left(\frac{n-p+2}{n-p}\,\frac{b}{\eta_i}+4\right)<2\sum_{i=1}^{p}\frac{1}{\lambda_i^2}$$

5) 如果取 $a_i=a^{-1}>0$, $b_i=b\lambda_i(n-p)^{-1}$, 则得到另外一个新的岭估计

$$\alpha^*(W)=(\alpha_1^*,\ \alpha_2^*,\ \cdots,\ \alpha_p^*)$$

其中

$$\alpha_i^*=\left(1-\frac{ab\lambda_i\hat{\sigma}^2}{a\hat{\alpha}'\Lambda\hat{\alpha}+(n-p)\hat{\sigma}^2}\right)\hat{\alpha}_i.$$

充分条件(3.1)成为

$$b<\frac{2(n-p)(p-2)}{n-p+2}\min_i\frac{\eta_i}{\lambda_i^2}.$$

6) 取 $W=\Lambda(\Lambda+kI)^{-1}(\Lambda+kI)^{-1}\Lambda$, 即 $\eta_1=\left(\frac{\lambda_i}{\lambda_i+k}\right)^2$ 这时,

$$\hat{\alpha}'W\hat{\alpha}=\hat{\alpha}'(k)\cdot\hat{\alpha}(k),\ \hat{\alpha}(k)=(\Lambda+kI)^{-1}\Lambda^{-1}\hat{\alpha}$$

是 Hoerl 和 Kennard(1970)提出的简单岭估计($k>0$ 常数), 它只能局部优于 LS 估计. $\alpha^*(W)$ 可以看作对 $\hat{\alpha}(k)$ 的改进, 它一致优于 LS 估计的充分条件是:

$$\max_i a_i^{-2}b_i\left(b_i\left(1+\frac{k}{\lambda_i}\right)^2+\frac{4}{\lambda_i}\right)<2\min_i a_i^{-2}\sum_{i=1}^{p}\frac{b_i}{\lambda_i}$$

由以上例子可以知道(1.3)定义的自适应岭估计 $\alpha^*(W)$ 是一个相当广泛的岭估计类. 需要指出的是, 参数 a_i, b_i 和 W 的选取除了考虑它的偏差和均方误差外, 还应同时考虑到作为一个好的估计的其它一些性质, 如稳健性等. 这一些都还有待于作进一步的研究.

参 考 文 献

[1] 陈希孺,《数理统计引论》,科学出版社,1981.

[2] 王松桂,《线性模型的理论及其应用》,安徽教育出版社,1987.

[3] 严利清,线性模型中几个有偏估计的改进,中国科学技术大学学报(1993).

[4] Firingutti, L., Exact moments of Lawless and Wang's operational ridge regression estimator, *COMMUN. STATIST. Theory Meth.*, **16**(3) (1987), 731—745.

[5] Hoerl, A. E. and R. W. Kennard, Ridge regression, *Technometrics*, **12** (1970), 55—82.

[6] Kadiyala, K., Some finite sample properties of generalized ridge regression estimators, *The Canadian Journal of Statistics*, Vol. **8**, No. 1 (1980), 47—58.

[7] Sawa, T., Finite-Sample properties of the k-class estimation, *Econometrica*, Vol. **40**, No. 4 (1970), 653—680.

[8] Vinod, H. D. and Ullah, A., Recent Advances in Regression Methods, Marcel, Dekker, 1981.

[9] Wang, S. G. and Chow, S. C.)A note on adaptive generalized ridge regression estimator, *STATIST. & PROB. Letters*, Vol. **10**, No. 1 (1990), 17—21.

FINITE-SAMPLE PROPERTIES OF AN ADATIVE RIDFE ESTIMATOR

YAN LIQING WANG SONGGUI

(University of Science and Technology of China, Hefei, 230026)

By using a directed method, we obtained the exact biase and mean squared error of an adaptive ridge estimator in linear regressional model, and got the sufficient condition under which the ridge estimator is uniformly better than the least squared estimator in sense of MSE.