第13卷第4期 数学研究与评论 Vol. 13 No. 4
1993年11月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Nov. 1993

线性回归模型参数的联立经验 Bayes 估计的收敛速度*

黄养新
(武汉工业大学数理系, 430070)

摘要

本文考虑了线性模型中回归系数 $\beta=(\beta_1,\cdots,\beta_p)'$ 和误差方差 σ^2 的联立经验 Bayes(EB) 估计. 在二次损失下, 利用密度函数及其偏导数的核估计构造出参数 $\theta=(\beta_1,\cdots,\beta_p,\sigma^2)$ 的联立 EB 估计, 在一定条件下证明了 θ 的联立 EB 估计的收敛速度任意接近于 1. 最后, 给出了一个实例.

§1 引言

考虑线性模型

$$y = X\beta + e, \quad e \sim N(0,\sigma^2 I_N), \tag{1}$$

其中 y 为 $N\times 1$ 观测向量, X 为 $N\times p$ 设计阵, $rk(X)=p<N$, e 为 $N\times 1$ 随机误差向量, 回归系数 $\beta=(\beta_1,\cdots,\beta_p)'\in R^p$ 为 $p\times 1$ 参数向量, $\sigma^2(>0)$ 为误差方差.

经验 Bayes 估计作为一种有效的参数估计已得到广泛的讨论; 关于单参数的 EB 估计及其收敛速度这方面已有很多结果[1]; 但对多参数的 EB 估计的收敛速度方面的研究并不多[2]; 本文在平方损失下, 研究了线性模型参数 $\theta=(\beta_1,\cdots,\beta_p,\sigma^2)$ 的联立 EB 估计, 在一定条件下证明了 θ 的 EB 估计的收敛速度可任意接近于 1.

为此, 取损失函数

$$L(d,\theta) = \sum_{i=1}^{p}(d_i-\beta_i)^2/\sigma^2 + (d_{p+1}-\sigma^2)^2/\sigma^{2l}, \tag{2}$$

其中 $l\geqslant 1$ 为正整数, $d=(d_1,\cdots,d_p,d_{p+1})$.

由于 $y\sim N(X\beta,\sigma^2 I_N)$, 因此, 参数为 θ 时, y 的概率密度为

$$\begin{aligned} g(y,\theta)d\mu &= C\sigma^{-N}\exp\{-(y-X\beta)'(y-X\beta)/2\sigma^2\}d\mu \\ &= C\sigma^{-N}\exp\{-(N-p)S^2/2\sigma^2-(\hat\beta-\beta)'X'X(\hat\beta-\beta)/2\sigma^2\}d\mu, \end{aligned}$$

其中 β 的极大似然估计 $\hat\beta=(X'X)^{-1}X'y$, $S^2=(y-X\hat\beta)'(y-X\hat\beta)/(N-p)$, 且 $\hat\beta$ 与 S^2 相互独立. 本文中 C 表示正常数, 且在各处所表示的值可以不同. 记

$$\hat\beta=(\hat\beta_1,\cdots,\hat\beta_p)', \quad T(y)=(\hat\beta_1,\cdots,\hat\beta_p,S^2).$$

设 μ^T 为通过可测变换 $T(y)$ 由勒贝格测度 μ 导出的测度, 由[3]中引理 1.3.3 知, $T(y)$ 是

* 1991年6月5日收到. 1993年5月16日收到修改稿.

完全充分统计量，且它的分布密度为

$$f(t|\theta)d\mu^T(t) = C\sigma^{-N}\exp\{-(N-p)S^2/2\sigma^2 - (\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)/2\sigma^2\}d\mu^T(t). \quad (3)$$

设 θ 的先验分布为 $G(\theta)$，满足条件

$$\int |\beta_i|^2\sigma^{-2}dG(\theta) < \infty \ (i=1,2,\cdots,p), \int |\sigma^2|^2\sigma^{-2l}dG(\theta) < \infty, \int \sigma^{-2l}dG(\theta) < \infty, \quad (4)$$

则 T 的边缘分布密度为 $f(t) = \int f(t|\theta)dG(\theta)$.

引理 1 在损失函数(2)下，回归系数 $\beta=(\beta_1,\cdots,\beta_p)'$ 和误差方差 σ^2 的 Bayes 估计分别为

$$\hat{\mu}_i = \hat{\beta}_i - \frac{N-p}{2}\Big[\sum_{j=1}^{p} b_{ij}f_{\hat{\beta}_j}(t)/f_{S^2}(t)\Big], i=1,2,\cdots,p, \quad (5)$$

$$\hat{\sigma}^2 = -\frac{N-p}{2}\big[f_{S^2}^{(l-1)}(t)/f_{S^2}^{(l)}(t)\big], \quad (6)$$

其中

$$B=(X'X)^{-1} \overset{\triangle}{=} (b_{ij})_{p\times p},\ f_{S^2}^{(i)}(t) = \int (-(N-p)/2\sigma^2)^i f(t|\theta)dG(\theta),$$

$$f_{\hat{\beta}}(t) = \int -\sigma^{-2}X'X(\hat{\beta}-\beta)f(t|\theta)dG(\theta) \overset{\triangle}{=} (f_{\hat{\beta}_1}(t),\cdots,f_{\hat{\beta}_p}(t))'.$$

证明 取决策函数 $d(t)$，则 $d(t)$的 Bayes 风险为

$$R(d) = \iint L(d,\theta)f(t|\theta)dtdG(\theta) = \iint L(d,\theta)f(t|\theta)dG(\theta)dt.$$

欲使 $R(d)$最小，等价于使 $r(d) \overset{\triangle}{=} \int L(d,\theta)f(t|\theta)dG(\theta)$ 最小. 令$\dfrac{\partial r(d)}{\partial d_i}=0$ $(i=1,2,\cdots,p)$，$\dfrac{\partial r(d)}{\partial d_{p+1}}=0$. 于是有

$$\hat{\mu}_i = \int \beta_i\sigma^{-2}f(t|\theta)dG(\theta)\Big/\int\sigma^{-2}f(t|\theta)dG(\theta) \quad (i=1,2,\cdots,p),$$

$$\hat{\sigma}^2 = \int \sigma^{-2(l-1)}f(t|\theta)dG(\theta)\Big/\int \sigma^{-2l}f(t|\theta)dG(\theta).$$

由 $B'f_{\hat{\beta}}(t) = -\hat{\beta}\int\sigma^{-2}f(t|\theta)dG(\theta) + \int\beta\sigma^{-2}f(t|\theta)dG(\theta)$ 及 $f_{S^2}^{(i)}(t)$的表达式便知(5)，(6)式成立.

§2 经验 Bayes 估计

假设给定历史样本(y_i,X_i) $(i=1,2,\cdots,n)$；(y,X)是当前样本. 于是有

$$y_i = X_i\beta^{(i)} + e_i, \quad e_i \sim N(0,\sigma_i^2 I_N),$$

这里 y_i $(i=1,2,\cdots,n)$；y 相互独立且 $y_i\sim N(X_i\beta^{(i)},\sigma_i^2I_N)$，$\beta^{(i)}=(\beta_1^{(i)},\cdots,\beta_p^{(i)})$. 进一步假设$(\beta_1^{(i)},\cdots,\beta_p^{(i)},\sigma_i^2)$ $(i=1,2,\cdots,n)$，$(\beta_1,\cdots,\beta_p,\sigma^2)$是从先验分布 $G(\theta)$中抽取的 i. i. d 样本. 记

$$T_i = (\hat{\beta}_1^{(i)},\cdots,\hat{\beta}_p^{(i)},S_i^2) \quad (i=1,2,\cdots,n),$$

其中 $\hat{\beta}^{(i)}=(\hat{\beta}_1^{(i)},\cdots,\hat{\beta}_p^{(i)})'=(X_i'X_i)^{-1}X_i'y_i$，$S_i^2=(y_i-X_i\hat{\beta}^{(i)})'(y_i-X_i\hat{\beta}^{(i)})/(N-p)$. 这样，$T_1,T_2,\cdots,T_n,T$ 相互独立服从共同的边缘分布密度 $f(t)$.

当作出 θ 的 EB 估计，对 $f(t),f_{\hat{\beta}_j}(t),f_{S^2}^{(l)}$ 用[4]中提出的方法作核估计. $f(t)$的 r 阶混合偏

导数为 $f^{(r)}(t)=\dfrac{\partial^r f(t)}{\partial\hat\beta_1^{r_1}\cdots\partial\hat\beta_p^{r_p}\partial(S^2)^{r_{p+1}}}\triangleq f^{(r)}(r_1,\cdots,r_{p+1};t)$，其中 $r=\sum_{i=1}^{p+1}r_i$，$r_i(i=1,2,\cdots,p+1)$ 为非负整数.

作 $f^{(r)}(t)$ 的核估计如下 $f_n^{(r)}(t)\triangleq f_n^{(r)}(r_1,\cdots,r_{p+1};t)=(nh_n^{p+r+1})^{-1}\sum_{i=1}^{n}K_{r,r_1,\cdots,r_{p+1}}(\frac{T_i-t}{h_n})$，其中 $h_n\geqslant 0$ 且 $\lim_{n\to\infty}h_n=0$，核函数 $K_{r,r_1,\cdots,r_{p+1}}(u)\triangleq K_r(u)\ (u\in R^{p+1})$ 为 $p+1$ 元实值可测函数.且满足条件

$$\text{(a)}\quad\begin{cases}|K_r(u)|\leqslant C, & u\in\prod_{i=1}^{p+1}[0,a_i]\triangleq D,\\ K_r(u)=0, & \text{其它};\end{cases}$$

$$\text{(b)}\quad\frac{1}{k_1!\cdots k_{p+1}!}\int K_r(u)\prod_{i=1}^{p+1}u_i^{K_i}du=\begin{cases}1 & k_i=r_i\,(i=1,2,\cdots,p+1)\\ 0 & \text{其它且}\ 0<\sum_{i=1}^{p+1}k_i\leqslant k\end{cases},$$

其中 $a_i(i=1,2,\cdots,p+1)$ 均为有限正实数，$k>l$ 为正整数.由此构造 $\beta_i\ (i=1,2,\cdots,p)$，$\sigma^2$ 的 EB 估计如下

$$\hat\mu_{in}=\hat\beta_i-\frac{N-p}{2}\Big[\sum_{j=1}^{p}b_{ij}f_{\beta_j n}(t)/f_{S^2 n}(t)\Big]_{h_n^{-1}}\quad(i=1,2,\cdots,p),\tag{7}$$

$$\hat\sigma_n^2=-\frac{N-p}{2}\Big[f_{S^2 n}^{(l-1)}(t)/f_{S^2 n}^{(l)}(t)\Big]_{h_n^{-1}},\tag{8}$$

其中 $[a]_b=\begin{cases}b & a>b\\ a & |a|\leqslant b\\ -b & a<-b\end{cases}$. $f_{\beta_j n}(t)$，$f_{S^2 n}^{(l)}$ 分别表示 $f_{\beta_j}(t)$，$f_{S^2}^{(l)}(t)$ 的核估计.

本文用 E_n 和 E 分别表示对 $(T_1,T_2,\cdots,T_n)$ 和 $(T_1,\cdots,T_n,(T,\theta))$ 的联合分布求期望.

θ 的 Bayes 估计 $\hat\theta=(\hat\mu_1,\cdots,\hat\mu_p,\hat\sigma^2)$ 的 Bayes 风险为

$$\begin{aligned}R_G&=R(\hat\theta,G)=\iint L(\hat\theta,\theta)f(t|\theta)dtdG(\theta)\\&=\iint\Big[\sum_{i=1}^{p}(\hat\mu_i-\beta_i)^2/\sigma^2\Big]f(t|\theta)dtdG(\theta)+\iint\big[(\hat\sigma^2-\sigma^2)^2/\sigma^{2l}\big]f(t|\theta)dtdG(\theta)\\&\triangleq R_G^{(1)}+R_G^{(2)}.\end{aligned}\tag{9}$$

θ 的 EB 估计 $\hat\theta_n(t)=(\hat\mu_{1n},\cdots,\hat\mu_{pn},\hat\sigma_n^2)$ 的全面 Bayes 风险为

$$\begin{aligned}R_n&=R(\hat\theta_n(t),G)=\iint L(\hat\theta_n(t),\theta)f(t|\theta)dtdG(\theta)\\&=\iint E_n\Big[\sum_{i=1}^{p}(\hat\mu_{in}-\beta_i)^2/\sigma^2\Big]f(t|\theta)dtdG(\theta)+\iint E_n\big[(\hat\sigma_n^2-\sigma^2)^2/\sigma^{2l}\big]f(t|\theta)dtdG(\theta)\\&\triangleq R_n^{(1)}+R_n^{(2)}\end{aligned}\tag{10}$$

若在先验分布 $G(\theta)$ 下，$R_n-R_G\to 0\ (n\to 0)$，则称 $\{\hat\theta_n(t)\}$ 为 θ 的渐近最优 EB 估计；若 $R_n-R_G=O(n^{-q})$，$q>0$，则称 θ 的 EB 估计 $\{\hat\theta_n(t)\}$ 的收敛速度为 q.

§3 几个引理

引理 2 若 $R_G^{(1)}<\infty, R_G^{(2)}<\infty$，则

$$R_n^{(1)}-R_G^{(1)}=\sum_{i=1}^{p}E[(\hat{\mu}_{in}-\hat{\mu}_i)^2/\sigma^2],\quad R_n^{(2)}-R_G^{(2)}=E[(\hat{\sigma}_n^2-\hat{\sigma}^2)^2/\sigma^{2l}]. \tag{11}$$

证明 由于

$$E[(\hat{\mu}_{in}-\beta_i)^2/\sigma^2]=E[(\hat{\mu}_{in}-\hat{\mu}_i+\hat{\mu}_i-\beta_i)^2/\sigma^2]=E[(\hat{\mu}_{in}-\hat{\mu}_i)^2/\sigma^2]+E[(\hat{\mu}_i-\beta_i)^2/\sigma^2],$$

因此有 $R_n^{(1)}-R_G^{(1)}=\sum_{i=1}^{p}E[(\hat{\mu}_{in}-\beta_i)^2/\sigma^2]-\sum_{i=1}^{p}E[(\hat{\mu}_i-\beta_i)^2/\sigma^2]=\sum_{i=1}^{p}[(\hat{\mu}_{in}-\hat{\mu}_i)^2/\sigma^2]$.

同理可证(13)的另一式成立. □

引理 3 如果在核估计 $\{f_n^{(r)}(t)\}$ 中，取 $h_n=n^{-1/[2(k+1)+p+1]}$，则对任意 $\varepsilon>0, 0<r\leqslant k$ 有

$$E_n|f_n^{(r)}(t)-f^{(r)}(t)|^2\leqslant Ch_n^{2(k+r-1)}[(f_\varepsilon^{(k+1)}(t))^2+f_\varepsilon(t)], \tag{12}$$

其中 $f_\varepsilon^{(k+1)}(t)=\sup\limits_{0\leqslant v_i\leqslant\varepsilon}|f^{(k+1)}(t+v)|, f_\varepsilon(t)=\sup\limits_{0\leqslant v_i\leqslant\varepsilon}f(t+v),\ v=(v_1,\cdots,v_{p+1})$.

证明 因为

$$E_n|f_n^{(r)}(t)-f^{(r)}(t)|^2=|E_n(f_n^{(r)}(t))-f^{(r)}(t)|^2+\mathrm{Var}(f_n^{(r)}(t)),$$

$$E_n(f_n^{(r)}(t))=h_n^{p+1+r})^{-1}E_n[K_r(\frac{T_1-t}{h_n})]=h_n^{-(p+1+r)}\int K_r(\frac{Z-t}{h_n})f(Z)dZ.$$

令 $(Z-t)/h_n=u$，则 $Z=t+h_nu$，因此 $E_n(f_n^{(r)}(t))=h_n^{-r}\int_D K_r(u)f(t+h_nu)du$，将函数 $f(t+h_nu)$ 在 $t\in R^{p+1}$ 处作 Taylor 展开，并利用核函数的性质，有

$$E_n(f_n^{(r)}(t))=f^{(r)}(t)+\sum_{k_1+\cdots+k_{p+1}=k}h_n^{k+1-r}\int_D\frac{f^{(k+1)}(t+h_nu\xi_{k+1})}{k_1!\cdots k_{p+1}!}K_r(u)\prod_{i=1}^{p+1}u_i^{k_i}du,$$

其中 $0<\xi_{k+1}<1,\ u=(u_1,\cdots,u_{p+1})'$.

如果 $h_nu_i\xi_{k+1}\leqslant h_nu_i\leqslant h_na_i\leqslant\varepsilon$（$n$ 充分大，$i=1,2,\cdots,p+1$），则有

$$|E_n(f_n^{(r)}(t))-f^{(r)}(t)|\leqslant h_n^{k+1-r}f_\varepsilon^{(k+1)}(t)\sum_{k_1+\cdots+k_{p+1}=k}\int_D\frac{|K_r(u)|}{k_1!\cdots k_{p+1}!}\prod_{i=1}^{p+1}u_i^{k_i}du\leqslant Ch_n^{k+1-r}f_\varepsilon^{(k+1)}(t)$$

同理可证明 $\mathrm{Var}(f_n^{(r)}(t))=n^{-1}h_n^{-2(p+1+r)}\mathrm{Var}\{K_r(\frac{T_1-t}{h_n})\}\leqslant Ch_n^{2(k+1-r)}f_\varepsilon(t)$. 综合以上各式，便知(12)式成立. 引理证毕.

记

$$F_i^{(1)}=\sum_{j=1}^{p}b_{ij}f_{\beta_j}(t)/f_{S^2}(t),\quad F_{in}^{(1)}=[\sum_{j=1}^{p}b_{ij}f_{\beta_{jn}}(t)/f_{S^2n}(t)]_{h_n^{-1}}, \tag{13}$$

$$F^{(2)}=f_{S^2}^{(l-1)}(t)/f_{S^2}^{(l)}(t),\quad F_n^{(2)}=[f_{S^2n}^{(l-1)}(t)/f_{S^2n}^{(l)}(t)]_{h_n^{-1}}. \tag{14}$$

则

$$\begin{aligned}|F_i^{(1)}-F_{in}^{(1)}|^2&\leqslant 2\{([F_i^{(1)}]_{h_n^{-1}}-F_{in}^{(1)})^2+(F_i^{(1)}-[F_i^{(1)}]_{h_n^{-1}})^2\}\\&\leqslant C\{|F_i^{(1)}|^2I_{(|F_i^{(1)}|\geqslant h_n^{-1})}+(F_{in}^{(1)}-[F_i^{(1)}]_{h_n^{-1}})^2I_{(|F_i^{(1)}|\leqslant h_n^{-1})}\}\\&\triangleq C(U_{i1}^{(1)}+U_{i2}^{(1)}),\end{aligned} \tag{15}$$

$$|F^{(2)}-F_n^{(2)}|\leqslant C\{|F^{(2)}|^2I_{(|F^{(2)}|\geqslant h_n^{-1})}+(F_n^{(2)}-[F^{(2)}]_{h_n^{-1}})^2I_{(|F^{(2)}|\leqslant h_n^{-1})}\}$$

$$\triangleq C(U_1^{(2)} + U_2^{(2)}). \tag{16}$$

引理 4 (a) 设存在 $\gamma>1$，使得 $E|\beta_i|^{2\gamma}\sigma^{-2}<\infty\ (i=1,\cdots,p)$，$E(\sigma^{2(\gamma-1)})<\infty$，$E(\sigma^{-2})<\infty$，则 $\sum_{i=1}^{p}E(U_{i1}^{(1)}/\sigma^2)\leqslant Ch_n^{2(\gamma-1)}$.

(b) 设存在 $\gamma>1$，使得 $E(\sigma^{2(2\gamma-l)})<\infty$，$E(\sigma^{-2l})<\infty$，则 $E(U_1^{(2)}/\sigma^{2l})\leqslant Ch_n^{2(\gamma-1)}$.

证明 (a) 由于 $E(\sigma^{-2})<\infty$，不妨设 $E(\sigma^{-2})=1$. 记 $d\widetilde{G}(\theta)=\sigma^{-2}dG(\theta)$，则 $d\widetilde{G}(\theta)$是分布密度，从而可得到相应的 $\tilde{f}(t)$，$\widetilde{E}(\cdot)$等. 由 Hölder 不等式知

$$E(U_{i1}^{(1)}/\sigma^2)=\widetilde{E}(U_{i1}^{(1)})=\widetilde{E}[|F_i^{(1)}|^2I_{(|F_i^{(1)}|\geqslant h_n^{-1})}]\leqslant(\widetilde{E}|F_i^{(1)}|^{2\gamma})^{1/\gamma}(\widetilde{E}|F_i^{(1)}h_n|^{\alpha})^{(\gamma-1)/\gamma}$$
$$\leqslant(\widetilde{E}|F_i^{(1)}|^{2\gamma})^{1/\gamma}(\widetilde{E}|F^{(1)}|^{\alpha})^{(\gamma-1)/\gamma}h_n^{\alpha(\gamma-1)/\gamma}.$$

取 $\alpha=2\gamma$，则有 $E(U_{i1}^{(1)}/\sigma^2)\leqslant(\widetilde{E}|F_i^{(1)}|^{2\gamma})h_n^{2(\gamma-1)}$. 又由 $\hat{\mu}_i=\widetilde{E}(\beta_i|T)$，$\hat{\mu}_i=\hat{\beta}_i-[(N-p)/2]F_i^{(1)}$，$\hat{\beta}\sim N(\beta,\sigma^2(X'X)^{-1})$及文献[5]中引理 5，得 $\widetilde{E}|F_i^{(1)}|^{2\gamma}\leqslant C(\widetilde{E}|\hat{\mu}_i|^{2\gamma}+\widetilde{E}|\hat{\beta}_i|)^{2\gamma})\leqslant C(\widetilde{E}|\hat{\mu}_i|^{2\gamma}+\widetilde{E}(\sigma^{2\gamma}))\leqslant C$.

同理可证明(b)成立. 引理证毕.

引理 5 (a) 若对 $0<\delta<1$，$E_{(T,\theta)}\frac{1}{\sigma^2}(\frac{f_t^{(k+1)}(t)}{f_{S^2}(t)})^{2\delta}<\infty$，$E_{(T,\theta)}\frac{1}{\sigma^2}(\frac{f_t(t)}{(f_{S^2}(t))^2})^{\delta}<\infty$，$k>l\geqslant1$，则当取 $h_n=n^{-1/[2(k+1)+p+1]}$ 时，有 $\sum_{i=1}^{p}E(U_{i2}^{(1)}/\sigma^2)\leqslant Ch_n^{2(k\delta-1)}$.

(b) 若对 $0<\delta<1$，$E_{(T,\theta)}\frac{1}{\sigma^{2l}}(\frac{f_t^{(k+1)}(t)}{f_{S^2}^{(l)}(t)})^{2\delta}<\infty$，$E_{(T,\theta)}\frac{1}{\sigma^{2l}}(\frac{f_t(t)}{f_{S^2}^{(l)}(t))^2})^{\delta}<\infty$，$k>l\geqslant1$，则当取 $h_n=n^{-1/[2(k+1)+p+1]}$ 时，有 $E(U_2^{(2)}/\sigma^{2l})\leqslant Ch^{2[(k-l+1)\delta-1]}$

证明 (a) 因为

$$U_{i2}^{(1)}\leqslant\{[F_{in}^{(1)}-F_i^{(1)}]_{2h_n^{-1}}\}^2I_{(|F_i^{(1)}|\leqslant h_n^{-1})}\leqslant(2h_n^{-1})^{2-2\delta}\{[(\sum_{i=1}^{p}b_{ij}f_{\hat{\beta}_jn}(t)/f_{S^2n}(t))$$
$$-(\sum_{i=1}^{p}b_{ij}f_{\hat{\beta}_j}(t)/f_{S^2}(t))]_{2h_n^{-1}}\}^{2\delta}I_{(|F_i^{(1)}|\leqslant h_n^{-1})},$$

由文献[6]中引理 3，本文引理 3 及 Hölder 不等式，有

$$E(U_{i2}^{(1)}/\sigma^2)\leqslant Ch_n^{2\delta-2}E_{(T,\theta)}\{\sigma^{-2}|f_{S^2}(t)|^{-2\delta}(h_n^{2\delta k}+h_n^{-2\delta}h_n^{2\delta k})[(f_t^{(k+1)}(t))^{2\delta}+(f_t(t))^{\delta}]\}$$
$$\leqslant Ch^{2(k\delta-1)}E_{(T,\theta)}\{\sigma^{-2}[(f_t^{(k+1)}(t))^{2\delta}+(f_t(t))^{\delta}]/|f_{S^2}(t)|^{2\delta}\}\leqslant Ch_n^{2(k\delta-1)}.$$

由此便知(21)式成立.

同理可证(b)成立. □

§4 主要定理及证明

定理 1 设 $1/2<\delta<1$，$k>l\geqslant1$ 为整数，若

(a) $E(|\beta_i|^{2k\delta}\sigma^{-2})<\infty\ (i=1,2,\cdots,p)$，$E(\sigma^{2(k\delta-1)})<\infty$，$E(\sigma^{-2})<\infty$；

(b) $E_{(T,\theta)}\sigma^{-2}(\frac{f_t^{(k+1)}(t)}{f_{S^2}(t)})^{2\delta}<\infty$，$E_{(T,\theta)}\sigma^{-2}(\frac{f_t(t)}{(f_{S^2}(t))^2})^{\delta}<\infty$，

则当取 $h_n=n^{-\frac{1}{2(k+1)+p+1}}$ 时，有 $R_n^{(1)}-R_G^{(1)}=O(n^{-\frac{2(k\delta-1)}{2(k+1)+p+1}})$.

证明 由引理 2 及(5),(7),(13),(15)式知

$$R_n^{(1)} - G_G^{(1)} = \sum_{i=1}^{p} E[(\hat{\mu}_{in} - \hat{\mu}_i)^2/\sigma^2] \leqslant C\sum_{i=1}^{p} E(U_{i1}^{(1)}/\sigma^2) + C\sum_{i=1}^{p} E(U_{i2}^{(1)}/\sigma^2).$$

令 $\gamma = k\delta > 1$,由引理 4(a)及引理 5(a)得 $R_n^{(1)} - R_G^{(1)} \leqslant C[h_n^{2(k\delta-1)} + h_n^{2(k\delta-1)}] = Cn^{-\frac{2(k\delta-1)}{2(k+1)+p+1}}$. □

定理 2 设 $1/2<\delta<1$, $k>l\geqslant 1$ 为整数,若

(a) $E(\sigma^{2(2k\delta-l)})<\infty$, $E(\sigma^{-2l})<\infty$;

(b) $E_{(T,\theta)}\sigma^{-2l}\left(\frac{f_t^{(k+1)}(t)}{f_{S^2}^{(l)}(t)}\right)^{2\delta}<\infty$, $E_{(T,\theta)}\sigma^{-2l}\left(\frac{f_t(t)}{(f_{S^2}^{(l)}(t))^2}\right)^{\delta}<\infty$,

则当取 $h_n = n^{-\frac{1}{2(k+1)+p+1}}$ 时,有 $R_n^{(2)} - R_G^{(2)} = O(n^{-\frac{2(k-l+1)\delta-2}{2(k+1)+p+1}})$.

证明 由引理 2 及(6),(8),(14),(16)式知

$$R_n^{(2)} - G_G^{(2)} = E[(\hat{\sigma}_n^2 - \hat{\sigma}^2)^2/\sigma^{2l}] \leqslant CE(U_1^{(2)}/\sigma^{2l}) + CE(U_2^{(2)}/\sigma^{2l}).$$

令 $\gamma = (k-l+1)\delta > 1$,由引理 4(b)及引理 5(b)得

$$R_n^{(2)} - R_G^{(2)} \leqslant C[h_n^{2[(k-l+1)\delta-1]} + h_n^{2[(k-l+1)\delta-1]}] = Cn^{-\frac{2(k-l+1)\delta-2}{2(k+1)+p+1}}.$$ □

由定理 1 和定理 2 立即可得

定理 3 设 $1/2<\delta<1$, $k>l\geqslant 1$ 为整数,若

(a) $E(|\beta_i|^{2k\delta}\sigma^{-2})<\infty$ $(i=1,2,\cdots,p)$, $E(\sigma^{2(2k\delta-l)})<\infty$, $E(\sigma^{-2l})<\infty$;

(b) $E_{(T,\theta)}\sigma^{-2l}\left(\frac{f_t^{(k+1)}(t)}{f_{S^2}^{(l)}(t)}\right)^{2\delta}<\infty$, $E_{(T,\theta)}\sigma^{-2l}\left(\frac{f_t(t)}{f_{S^2}^{(l)}(t))^2}\right)^{\delta}<\infty$,

则当取 $h_n = n^{-\frac{1}{2(k+1)+p+1}}$ 时,有 $R_n - R_G = O(n^{-\frac{2(k-l+1)\delta-2}{2(k+1)+p+1}})$.

由以上定理可见,若对任意整数 $k>l\geqslant 1$ 及任意接近于 1 的 δ,定理条件均成立,则 θ 的 EB 估计的收敛速度可任意接近于 1.

§5 一个例子

在线性模型的 Bayes 分析中,经常采用先验分布为正态—伽马分布,其具体形式如下

$$dG_1(\beta_1,\cdots,\beta_p|\sigma^2) = C\prod_{i=1}^{p}\sigma^{-1}\exp\{-(\beta_i-\alpha_i)^2/2\sigma^2\}d\beta_i,$$

$$dG_2(\sigma^2) = C(\sigma^{-2})^{a+1}\exp\{-\sigma^{-2}b\}d\sigma^2,$$

$$G(\beta_1,\cdots,\beta_p,\sigma^2) = G_1(\beta_1,\cdots,\beta_p|\sigma^2)G_2(\sigma^2),$$

其中 $|\beta_i|<\infty$ $(i=1,2,\cdots,p)$, $\sigma^2>0$, $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_p)'$ 任意给定, $b>0$, $a>k\delta-1$. 因此

$$dG(\theta) = C\prod_{i=1}^{p}(\sigma^{-1-2(a+1)})\exp\{-\sigma^{-2}b - (\beta_i-\alpha_i)^2/2\sigma^2\}d\beta_i d\sigma^2,$$

又 $f(t) = \int f(t|\theta)dG(\theta) = C[2b + (N-p)S^2 + \hat{\beta}'X'X\hat{\beta} + \alpha'\alpha]^{-(2a+N+p)/2}$.

不难验证各定理条件皆成立. 因此,在先验分布为正态—伽马分布时,线性模型中参数 $\theta = (\beta_1,\cdots,\beta_p,\sigma^2)$ 的联立 EB 估计 $\hat{\theta}_n(t) = (\hat{\mu}_{1n},\cdots,\hat{\mu}_{pn},\hat{\sigma}_n)$ 的收敛速度可任意接近于 1.

参 考 文 献

[1] R. S. Singh, *Empirical Bayes estimation in Lebesgue—exponential families with rate near the best possible rate*, Ann. Statist., 7(1979), 890—902.

[2] 韦来生，连续型多参数指数族参数的经验 Bayes 估计的收敛速度，数学学报，30:2(1987)，272—279.

[3] 陈希孺，数理统计引论，科学出版社(1981).

[4] 卢昆亮，密度的混合偏导数的核估计及其收敛速度，系统科学和数学，2:3(1982)，220—226.

[5] 钱伟民，正态分布族的参数的经验 Bayes 估计的收敛速度，应用概率统计，6:2(1990)，113—120.

[6] 赵林城，一类离散分布参数的经验 Bayes 估计的收敛速度，数学研究与评论，1(1981)，59—69.

[7] L. D. Broemeling, *Bayesian analysis of linear models*, Narcel Dekker INC (1985).

The Convergence Rate of Simultaneous Empirical Bayes Estimator for Parameters in Linear Regression Models

Huang Yangxin

(Dept. of Math. & Physics, Wuhan Univ. of Technology)

Abstract

In this paper, the simultaneous empirical Bayes (EB) estimator for regression cofficient $\beta = (\beta_1, \beta_2, \cdots, \beta_p)'$ and error variance σ^2 is considered in linear models. Suppose that the loss function is square, the convergence rate of simultaneous EB estimator about the parameter $\theta = (\beta_1, \cdots, \beta_p, \sigma^2)$ is obtained and this convergence rate can be arbitrarily close to 1 under suitable conditions. At last, an example is given.