应用概率统计　第二十二卷　　Chinese Journal of Applied Probability
第二期　2006 年 5 月　　and Statistics Vol. 22 No. 2 May 2006

多元回归模型中最大值点的 BRPA 估计的收敛速度 *

吴耀华　　龚　隽
(中国科学技术大学统计与金融系，合肥，　230026)

摘　　要

BRPA 估计是 Changchien (1990) 提出的一种具有良好性质的回归函数最大值点的估计，Chen, Huang and Huang (1996), Bai and Huang (1999), 吴 and 王 (2000) 和 Bai, Chen and Wu (2003) 分别讨论了 BRPA 的极限性质．本篇文章中，我们在很一般的条件下研究了 x 为多维向量时 BRPA 估计的收敛速度，推广了 Bai, Chen and Wu (2003) 的结果．

关 键 词：极值估计，相伴次序统计量，收敛速度.
学 科 分 类 号：O212.7.

§1.　引言与结果

考虑一般的回归模型

$$Y=\theta(x)+\epsilon, \tag{1.1}$$

其中 $\theta(x)$ 是定义在 $\mathbb{R}^k$ 中有界开集 C 上的一个 L 可测函数，有唯一的最大值点 x_0. 现有设计点列 $X_1,X_2,\cdots,X_n$, 其对应的观测值是 $Y_1,Y_2,\cdots,Y_n$, 即

$$Y_i=\theta(X_i)+\epsilon_i,\qquad i=1,\cdots,n. \tag{1.2}$$

其中 $\{X_i\}$ 可以是固定的设计点列，也可以是从 C 上具有密度函数 g 的总体中抽出的 i.i.d. 样本，$\{\epsilon_i\}$ 是服从分布 F 的 i.i.d. 随机误差．当 $\{X_i\}$ 为随机时，有 $\{\epsilon_i\}$ 与 $\{X_i\}$ 独立，而且假定存在 $M>0$ 使得对所有的 $x\in C$ 成立 $g(x)>M$. 我们现在需要利用数据 $\{(X_i,Y_i):i=1,\cdots,n\}$ 对 x_0 作出估计．

比较经典的做法是利用非参数的方法得到 $\theta(x)$ 的估计 $\widehat{\theta}(x)$, 而后利用 $\widehat{\theta}(x)$ 在 C 上的最大值作为 x_0 的估计．这方面的工作可以参看 Ibragimov and Khas'minskii (1980) 和 Müller (1989).

近年来，一种被称为最优 r 点平均 (BRPA) 的估计 x_0 的方法得到了应用工作者的青睐．它的作法是：$Y_{(1)}\le\cdots\le Y_{(n)}$ 是 $Y_1,\cdots,Y_n$ 的次序统计量，而 $Y_{(i)}$ 在样本中对应的 X 为 $X_{[i]}$ (相伴次序统计量), 则 x_0 的 BRPA 估计定义为

$$\widehat{x}(r)=\frac{1}{r}\sum_{i=1}^{r}X_{[n+1-i]}, \tag{1.3}$$

* 本文得到国家自然科学基金 (批号：　10171094, 10471136), 教育部博士点基金与中国科学院和中国科学技术大学创新基金的资助.

本文 2005 年 4 月 16 日收到.

即最大的 r 个 Y 观测值所对应的 X 值的平均．这个估计在 $r=1$ 时的特例是由 Changchien (1990) 提出的．Chen, Huang and Huang (1996) 具体提出了 BRPA 估计．

以 $\xi_{s,t}$ 记 F 的 $1-t/s$ 分位点，其中 $t<s$. 令

$$B(s,t)=\xi_{s,t}-\xi_{s,t+1}, \tag{1.4}$$

特别的，记

$$B(s)=B(s,1). \tag{1.5}$$

Bai and Huang (1999) 和吴 and 王 (2000) 分别证明了 $k=1$ 与 $k>1$ 时 BRPA 估计相合的充要条件是：当 $n\to\infty$ 时 $B(n)\to 0$. Chen, Huang and Huang (1996) 和 Bai, Chen and Wu (2003) 得到了 $k=1$ 时 BRPA 估计的收敛速度．我们这篇文章将对 $k>1$ 的情形给出 BRPA 估计的收敛速度．

为叙述方便，我们引入下面的记号与条件：

对 $\mathbb{R}^k$ 中任意 L 可测集 A 和点 x,y, 以 $|A|$ 表示 A 的 L 测度， $d(x,y)$ 表示欧氏距离．

(A1) 对每个 t, $B(s,t)$ 是 s 的非增函数．

(A2) 存在 $\alpha\ge 0$, 使得 $B(s)=s^{-\alpha}L(s)$, 这里 $L(s)$ 满足对任意固定的 $a>0$, 当 $s\to\infty$ 时， $L(as)/L(s)\to 1$. 且对任给的 $c>0$ 有

$$0<\liminf_{s\to\infty}\frac{L(s^c)}{L(s)}\le\limsup_{s\to\infty}\frac{L(s^c)}{L(s)}<\infty. \tag{1.6}$$

(B1) $\theta(x)$ 在 C 上有唯一的极大值点 x_0.

(B2) 存在一 $\delta_0>0$, 对所有的 $\delta\in(0,\delta_0]$, 令 $A_\delta=\{x\in C:\theta(x_0)-\theta(x)\le\delta\}$, 则存在凸开集 $U\subset C$, 使得 $U\subset A_\delta\subset\overline{U}$. 令 $D_1=\sup\limits_{x\in U}\inf\limits_{y\in\partial U}d(x,y)$, $D_2=\inf\limits_{x\in\overline{U}}\sup\limits_{y\in\overline{U}}d(x,y)$, 则存在 $\tau,c_1,c_2>0$ 使得

$$c_1\delta^\tau\le D_1\le D_2\le c_2\delta^\tau.$$

定理 对模型 (1.1), 在假设 (A1) 、 (A2) 、 (B1) 和 (B2) 下，有

$$\theta(x_0)-\theta(\widehat{x}_0(r))=O_p(B(n)^{1/(\alpha k\tau+1)}) \tag{1.7}$$

和

$$d(x_0,\widehat{x}_0(r))=O_p(B(n)^{\tau/(\alpha k\tau+1)}). \tag{1.8}$$

注 1 我们知道任意一列以 0 为极限的数列 $\{a_n\}$ 总可以表示为 $a_n=n^{-\alpha}L(n)$, 其中 $\alpha\ge 0$, 且 $L(n)$ 为一慢变化函数．而对 $L(s)$ 为一些常见的慢变化函数如： $\log s$, $\log\log s$, 1, $(\log s)^{-1}$, $(\log\log s)^{-1}$ 时，均有 (A2) 成立，由 Bai and Huang (1999) 和吴 and 王 (2000) 的结论知我们的结果对大部分常见的相合 BRPA 估计都成立．

注 2 条件 (B2) 主要是为了保证 x_0 不是回归函数 $\theta(x)$ 的一个孤立点，并不要求 $\theta(x)$ 在 x_0 连续．例如：

$$\theta(x_1,x_2)=\begin{cases}0 & \text{如果 } x_1<0;\\ 1-(x_1^2+x_2^2)^{1/2} & \text{其他,}\end{cases}$$

则 $x_0=(0,0)$, $\tau=1$, 但 $\theta(x)$ 在 $(0,0)$ 点不连续.

注 3　当 $k=1$ 时，我们的结果就是 Bai, Chen and Wu (2003) 的结论.

§2.　定理的证明

不失一般性，设 $|C|=1$, $g(x)\equiv 1$, 此处及以下我们以 $|A|$ 记集合 A 的体积，$\#(A)$ 记集合 A 中包含的点数. 因为 (1.8) 可由 (1.7) 和 (B2) 推得，故只需证明 (1.7) 成立即可. 为证 (1.7), 我们只需证明

$$\mathsf{P}(\theta(x_0)-\theta(\widehat{x}_0(r))\geq C_nB(n)^{1/(\alpha k\tau+1)})\to 0 \qquad \text{当 } n\to\infty \text{ 时}, \tag{2.1}$$

对任意趋于 ∞ 的正数列 $\{C_n\}$ 成立.

取 C_n 使得 $C_nB(n)^{1/(\alpha k\tau+1)}\to 0$ 当 $n\to\infty$. 对 $j\geq 0$, 定义

$$\delta_j=\frac{\delta_0}{2^j}, \tag{2.2}$$

$$U_j\subset A_j=\{x\in C:\theta(x_0)-\theta(x)\leq\delta_j\}\subset\overline{U}_j, \tag{2.3}$$

$$\Delta_j=|A_j|. \tag{2.4}$$

这里 U_j 如 (B2) 中所示. 记满足 $X_i\in A_j$ 的 (X_i,Y_i,ϵ_i) 为

$$\{(X_1^{(j)},Y_1^{(j)},\epsilon_1^{(j)}),\cdots,(X_{n_j}^{(j)},Y_{n_j}^{(j)},\epsilon_{n_j}^{(j)})\}. \tag{2.5}$$

由 (B2) 知存在正常数 $\widetilde{c}_1,\widetilde{c}_2$ 使得 $\widetilde{c}_1n\delta_j^{k\tau}\leq n_j\leq\widetilde{c}_2n\delta_j^{k\tau}$.

令 $m(n)=\max\limits_{1\leq j\leq n}\{j:C_nB(n)^{1/(\alpha k\tau+1)}\leq\delta_j\}$, 简记 $m=m(n)$. 对 $0\leq j\leq m$, 令 $k_{nj}=[2^{k\tau(m-j)}K_n]$, 我们要求 $K_n\to\infty$ 并且 $k_{nj}\leq n_j$. 因为

$$n_j\geq\widetilde{c}_1n\delta_j^{k\tau}=\widetilde{c}_1n(\delta_m2^{m-j})^{k\tau}\approx\widetilde{c}_1n\delta_m^{k\tau}\frac{k_{nj}}{k_{nm}}\geq\frac{\widetilde{c}_1}{\widetilde{c}_2}n_m\frac{k_{nj}}{k_{nm}},$$

只要 $n_m/k_{nm}\to\infty$ 即可，等价于

$$\frac{n(C_nB(n)^{1/(\alpha k\tau+1)})^{k\tau}}{K_n}\to\infty. \tag{2.6}$$

我们可取 $K_n=C_n^{(k\tau+1)/(\alpha k\tau+1)}$ 使条件满足.

设 $\epsilon_{(1)}^{(j)}\leq\epsilon_{(2)}^{(j)}\leq\cdots\leq\epsilon_{(n_j)}^{(j)}$ 是 $\epsilon_1^{(j)},\epsilon_2^{(j)},\cdots,\epsilon_{n_j}^{(j)}$ 的次序统计量，$Y_{(j)}^{(j)}\leq Y_{(2)}^{(j)}\leq\cdots\leq Y_{(n_j)}^{(j)}$ 是 $Y_1^{(1)},Y_2^{(j)},\cdots,Y_{n_j}^{(j)}$ 的次序统计量. $X_{[1]}^{(j)},\cdots,X_{[n_j]}^{(j)}$ 分别对应 $Y_{(1)}^{(j)},\cdots,Y_{(n_j)}^{(j)}$. 令 $\epsilon_{l_i}^{(j)}=\epsilon_{(i)}^{(j)}$, $i=1,\cdots,n_j$. 对 $j\geq 1$, 定义如下事件:

$$R_{nj}=\{\epsilon_{(n_j)}^{(j)}-\epsilon_{(n_j-k_{nj})}^{(j)}\geq\delta_{j+2}\}, \tag{2.7}$$

$$Q_{nj}=\{\#\{i:i\leq k_{nj},X_{l_{n_j}-i}^{(j)}\in U_{j+2}\}<r\}, \tag{2.8}$$

$$D_{nj}=\{X_{[n_j]}^{(j)},\cdots,X_{[n_j-r+1]}^{(j)}\subset U_{j+1}\}. \tag{2.9}$$

当事件 $R_{nj}^c \cap Q_{nj}^c$ 发生时，即至少有 r 个 $i\ (\le k_{nj})$ 使得 $X_{l_{n_j}-i}^{(j)} \in U_{j+2}$，那么对 $\forall X_{i'}^{(j)} \overline{\in} U_{j+1}$ 有

$$\begin{aligned} Y_{l_{n_j}-i}^{(j)} &= \theta(X_{l_{n_j}-i}^{(j)}) + \epsilon_{(n_j-i)}^{(j)} \ge \theta(x_0) - \delta_{j+2} + \epsilon_{(n_j)}^{(j)} - \delta_{j+2} \\ &> \theta(X_{i'}^{(j)}) + \epsilon_{i'}^{(j)} = Y_{i'}^{(j)}. \end{aligned} \tag{2.10}$$

上式也就意味着 $R_{nj}^c \cap Q_{nj}^c \subset D_{nj}$. 定义

$$R_{n0} = \{\epsilon_{(n)} - \epsilon_{(n-k_{n0})} \ge \delta_2\}, \tag{2.11}$$

$$Q_{n0} = \{\#\{i : i \le k_{n0}, X_{l_n-i} \in U_2\} < r\}, \tag{2.12}$$

$$D_{n0} = \{X_{[n]}, \cdots, X_{[n-r+1]} \in U_1\}. \tag{2.13}$$

同样有 $R_{n0}^c \cap Q_{n0}^c \subset D_{n0}$. 注意到

$$\{\theta(x_0) - \theta(\widehat{x}_0(r)) \le C_n B(n)^{1/(\alpha k\tau+1)}\} \supset \bigcap_{j=0}^{m} D_{nj}, \tag{2.14}$$

所以有

$$\begin{aligned} &\mathsf{P}(\theta(x_0) - \theta(\widehat{x}_0(r)) \ge C_n B(n)^{1/(\alpha d\tau+1)}) \\ \le\ & \sum_{j=0}^{m} \mathsf{P}(D_{nj}^c) \le \sum_{j=0}^{m} \mathsf{P}(R_{nj}) + \sum_{j=0}^{m} \mathsf{P}(Q_{nj}) \equiv \gamma_n + \varrho_n. \end{aligned} \tag{2.15}$$

故我们只需证 $\lim\limits_{n\to\infty} \gamma_n = \lim\limits_{n\to\infty} \varrho_n = 0$.

应用 Hoeffding 不等式可得

$$\begin{aligned} \mathsf{P}(Q_{nj}) &= \mathsf{P}\Big(B_i\Big(k_{nj}, \frac{\Delta_{j+2}}{\Delta_j}\Big) < r\Big) \\ &\le \exp\Big(-\frac{2c^2 k_{nj}\Delta_{j+2}}{\Delta_j}\Big(\frac{\Delta_{j+2}}{\Delta_j} - \frac{r}{k_{nj}}\Big)^2\Big) \qquad (\exists\, c > 0) \\ &\le \exp(-\widetilde{c}2^{k\tau(m-j)}K_n) \qquad (\exists\, \widetilde{c} > 0). \end{aligned} \tag{2.16}$$

于是有

$$\begin{aligned} \varrho_n &\le \sum_{j=0}^{m} \exp(-\widetilde{c}2^{k\tau(m-j)}K_n) \\ &\le \sum_{j=0}^{\infty} \exp(-\widetilde{c}2^{k\tau j}K_n) \to 0 \qquad (n\to\infty). \end{aligned} \tag{2.17}$$

由 m 的定义知

$$\delta_{m+1} < C_n B(n)^{1/(\alpha k\tau+1)} \le \delta_m. \tag{2.18}$$

立得对任意的 $0 \le j \le m$,

$$\delta_{j+1} < 2^{m-j} C_n B(n)^{1/(\alpha k\tau+1)} \le \delta_j. \tag{2.19}$$

于是

$$\begin{aligned}
\mathsf{P}(R_{nj}) &= \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge\delta_{j+2})\\
&\le \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge 2^{m-2-j}C_nB(n)^{1/(\alpha k\tau+1)})\\
&\le \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge \tilde{c}2^{m-2-j}C_n^{\alpha k\tau+1}B(n))\\
&= \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge \tilde{c}2^{m-2-j}C_n^{\alpha k\tau+1}n^{-\alpha}L(n))\\
&\le \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge c2^{(m-j)(k\tau+1)}C_n^{\alpha k\tau+1}n_j^{-\alpha}L(n_j))\\
&\le \mathsf{P}(\epsilon^{(j)}_{(n_j)}-\epsilon^{(j)}_{(n_j-k_{nj})}\ge ck_{nj}^{1+k\tau}B(n_j))\\
&\le \mathsf{P}(\epsilon^{(j)}_{(n_j)}\ge\xi_{k_{nj}^{k\tau}n_j,1})+\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{k_{nj}^{k\tau}n_j,1}-ck_{nj}^{1+k\tau}B(n_j)),
\end{aligned}$$

这里 $\tilde{c}$, c 是正常数. 所以

$$\begin{aligned}
\gamma_n &\le \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j)}\ge\xi_{k_{nj}^{k\tau}n_j,1})+\sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{k_{nj}^{k\tau}n_j,1}-ck_{nj}^{1+k\tau}B(n_j))\\
&\equiv \gamma_{n1}+\gamma_{n2}. \qquad (2.20)
\end{aligned}$$

下面我们估计 γ_{n1} 和 γ_{n2}.

$$\begin{aligned}
\gamma_{n1} &= \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j)}\ge\xi_{k_{nj}^{k\tau}n_j,1})=\sum_{j=0}^m(1-(\mathsf{P}(\epsilon^{(j)}_i<\xi_{k_{nj}^{k\tau}n_j,1}))^{n_j})\\
&\le \sum_{j=0}^m\Big(1-\Big(1-\frac{1}{k_{nj}^{k\tau}n_j}\Big)^{n_j}\Big)\le\sum_{j=0}^m\frac{1}{k_{nj}^{k\tau}}\le\sum_{j=0}^m\Big(\frac{1}{2^{(m-1-j)k\tau}K_n}\Big)^{k\tau}\\
&= K_n^{-k\tau}\sum_{j=0}^m2^{-(k\tau)^2(m-1-j)}\le K_n^{-k\tau}\sum_{j=-1}^{\infty}2^{-(k\tau)^2j}\to0 \qquad (n\to\infty). \qquad (2.21)
\end{aligned}$$

利用假设 (A1)(A2), 当 n 充分大时有

$$\begin{aligned}
\gamma_{n2} &= \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{k_{nj}^{k\tau}n_j,1}-ck_{nj}^{1+k\tau}B(n_j))\\
&= \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{k_{nj}^{k\tau}n_j,1}-\xi_{k_{nj}^{k\tau}n_j,k_{nj}^{k\tau}}+\xi_{n_j,1}-ck_{nj}^{1+k\tau}B(n_j))\\
&\le \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le k_{nj}^{k\tau}B(k_{nj}^{k\tau}n_j)+\xi_{n_j,1}-ck_{nj}^{1+k\tau}B(n_j))\\
&\le \sum_{j=0}^m\mathsf{P}\Big(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{n_j,1}-\frac{c}{2}k_{nj}^{1+k\tau}B(n_j)\Big)\\
&\le \sum_{j=0}^m\mathsf{P}(\epsilon^{(j)}_{(n_j-k_{nj})}\le\xi_{n_j,\frac{c}{2}k_{nj}^{1+k\tau}})\le\sum_{j=0}^m\mathsf{P}\Big(\sum_{i=1}^{n_j}I(\epsilon^{(j)}_i\le\xi_{n_j,\frac{c}{2}k_{nj}^{1+k\tau}})\ge n_j-k_{nj}\Big)\\
&\le \sum_{j=0}^m\mathsf{P}\Big(\sum_{i=1}^{n_j}[I(\epsilon^{(j)}_i\le\xi_{n_j,\frac{c}{2}k_{nj}^{1+k\tau}})-\mathsf{E}I(\epsilon^{(j)}_i\le\xi_{n_j,\frac{c}{2}k_{nj}^{1+k\tau}})]\ge\frac{c}{2}k_{nj}^{1+k\tau}-k_{nj}\Big)\\
&\le \sum_{j=0}^m\exp(-c'k_{nj}^{1+k\tau})\to0, \qquad (\exists\, c'>0). \qquad (2.22)
\end{aligned}$$

结合 (2.15), (2.17), (2.20), (2.21) 和 (2.22) 证得 (2.1) 成立，定理证毕.

参 考 文 献

[1] Bai, Z.D., Chen, Z.H. and Wu, Y.H., Convergence rate of the best-r-points-average estimator of the maximizer of a nonparametric regression function, *J. Multivariate Anal.*, **84**(2003), 319–334.

[2] Bai, Z.D. and Huang, M.-N.L., On the consistency of the best-r-points-average estimator for the maximizer of a nonparametric regression function, *Sankhyā: The Indian Journal of Statistics*, **61, Ser. A, Pt. 2**(1999), 208–217.

[3] Changchien, G.M., Optimization of blast furnace burden distribution, In: Chao, M.T. and Cheng, P.E. Eds., *Proceedings of the 1990 Taipei Symposium in Statistics, June 28–30, 1990*, 63–78, 1990.

[4] Chen, H., Huang, M.-N.L. and Huang, W.-J., Estimation of the location of the maximum of a regression function using extreme order statistics, *J. Multivariate Anal.*, **57**(1996), 191–214.

[5] David, H.A., *Order Statistics*, 2nd edition, John Wiley, New York, 1981.

[6] Gnedenko, B., Sur la distribution limite terme maximum d'une serie aleatoire, *Ann. of Math.*, **44**(1943), 423–453.

[7] Gumbel, E.J., *Statistics of Extremes*, Columbia University Press, New York, 1958.

[8] Müller, H.G., Adaptive nonparametric peak estimation, *Ann. Statist.*, **17**(1989), 1053–1069.

[9] Reiss, R.-D., *Approximate Distributions of Order Statistics with Applications to Nonparametric Statistics*, Springer-Verlag, New York, 1989.

[10] Resnick, S.I., *Extreme Values, Regular Variation, and Point Processes*, Springer-Verlag, New York, 1987.

[11] Weissman, I., Multivariate extremal processes generated by independent non-identically distributed random variables, *J. Appl. Prob.*, **12**(1975), 477–487.

[12] Weissman, I., Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Stat. Assoc.*, **73**(1978), 812–815.

[13] 吴耀华, 王小明, 多元回归函数最大值点 BRPA 估计的相合性, 应用概率统计, **16(3)**(2000), 229–302.

Convergence Rate of the Best-r-Points-Average Estimator of the Maximizer of a Multivariate Regression Function

WU YAOHUA GONG JUN

(*Department of Statistics and Finance, University of Science and Technology of China, Hefei, 230026*)

BRPA estimator, proposed in Changchien (1990), has certain merits over the estimators derived through the estimation of the regression function. Some of the properties of the BRPA estimator have been studied in Chen, Huang and Huang (1996), Bai and Huang (1999), Wu and Wang (2000) and Bai, Chen and Wu (2003). In this article, we give the convergence rate of the BRPA estimator of the multivariate regression function under some mild conditions and popularize the result of Bai, Chen and Wu (2003).