DOI:10.13656/j.cnki.gxkx.2011.02.019

广西科学 Guangxi Sciences 2011,18(2):133～135

PA 样本回归函数估计的强相合性*

Strong Consistency of the Regression Weighted Function Estimator for Positively Associated Samples

黎玉芳

LI Yu-fang

(广西师范大学数学科学学院,广西桂林 541004)

(College of Mathematics,Guangxi Normal University,Guilin,Guangxi,541004,China)

摘要:针对非参数回归模型 $Y_i = g(x_i)+\varepsilon_i, 1\leqslant i\leqslant n$,在$\{\varepsilon_i, 1\leqslant i\leqslant n\}$为一致可积的平稳PA相依序列条件下,得到未知函数 $g(x)$ 的权函数估计 $g_n(x) = \sum_{i=1}^{n} w_{ni}(x)Y_i$ 的强相合性.

关键词:权函数估计 强相合性 PA 相依 一致可积

中图法分类号:O211.4 **文献标识码**:A **文章编号**:1005-9164(2011)02-0133-03

Abstract:In the nonparametric regression model $Y_i = g(x_i)+\varepsilon_i, 1\leqslant i\leqslant n$, we study the strong consistency of the nonparametric regression weighted function estimator $g_n(x) = \sum_{i=1}^{n} w_{ni}(x)Y_i$, where $\{\varepsilon_i, 1\leqslant i\leqslant n\}$ is uniform integrabilited and positively associated samples.

Key words:weighted function estimator, strong consistency, positive association, uniform integrability

设集合 A_1 与 A_2 为$\{1,2,\cdots n\}$ 的任何两个不相交的非空子集,称随机变量 $X_1,X_2,\cdots,X_n(n\geqslant 2)$ 是 PA 的,如果

$$\mathrm{Cov}(f_1(X_i,i\in A_1),f_2(X_j,j\in A_2))\geqslant 0,$$

其中 f_1 和 f_2 是任何两个使得协方差存在且对每个变元均非降(或同对每个变元均非升)的函数,称随机变量序列$\{X_i,i\geqslant 1\}$ 是 PA 序列,如果对任何 $n\geqslant 2$,$X_1,X_2,\cdots,X_n$ 是 PA 的.

设 d 是一个正整数,A 是R^d 上的一个紧集.考虑非参数回归模型

$$Y_i = g_n(x_i)+\varepsilon_i, 1\leqslant i\leqslant n,$$

其中 x_i 为已知的设计点列,g 是A 上的未知连续有界回归函数,ε_i 是均值为零的随机误差,Y_i 为可观察的随机变量,它们定义在同一个概率空间(Ω,A,P).基于观察值$\{x_i,Y_i,1\leqslant i\leqslant n\}$,$g(x)$ 的一个估计为

$$g_n(x) = \sum_{i=1}^{n} w_{ni}(x)Y_i,$$

其中 $w_{ni}(x)$ 是依赖 $x_i,\cdots,x_n$ 及n 的可测的权函数,$i = 1,2,\cdots,n$,满足基本假设 A_1:

(a) $\sum_{i=1}^{n} w_{ni}(x)\to 1$,当 $n\to\infty$;

(b) $\sum_{i=1}^{n} |w_{ni}(x)|\leqslant B$,对 $\forall n\geqslant 1$;

(c) $\sum_{i=1}^{n} |w_{ni}(x)|I_{(\|x-x_i\|>a)}\to 0$,当 $n\to\infty$,$\forall a>0$.

近年来,在独立样本和其他相依样本情形下,有关 $g_n(x)$ 性质的研究已取得不少成果[1~4].但是在 PA 样本下的研究却较少[5].本文在下述一致可积的条件下得到 $g_n(x)$ 的强相合性.

如果$\lim\limits_{x\to\infty}\sup\limits_{n\in N}\sum_{k=1}^{k_n}|a_{nk}|E|X_{nk}|I_{(|X_{nk}|\geqslant x)} = 0$,则称随机阵列$\{X_{nk},1\leqslant k\leqslant k_n\uparrow\infty,n\in N\}$ 关于$\{|a_{nk}|,1\leqslant k\leqslant k_n\uparrow\infty,n\in N\}$ 一致可积[6].

取 $k_n = n$,若$\lim\limits_{x\to\infty}\sup\limits_{n\in N}\sum_{k=1}^{n}|a_{nk}|E|X_k|I_{(|X_k|\geqslant x)} = 0$,则称随机序列$\{X_n,n\in N\}$ 关于$\{|a_{nk}|,1\leqslant k\leqslant n,n\in N\}$ 是一致可积.显然,若

$$\lim_{x\to\infty}\sup_{n\in N}\sum_{k=1}^{n}|a_{nk}|^pE|X_k|^pI_{(|X_k|\geqslant x)} = 0, \quad (1)$$

则称随机序列$\{|X_n|^p,n\in N\}$ 关于$\{|a_{nk}|^p,1\leqslant k$

收稿日期:2010-10-11

作者简介:黎玉芳(1977-),女,讲师,主要从事概率统计研究。

*国家自然科学基金资助项目(10661003),广西自然科学基金项目(0832102),广西师范大学骨干教师科研项目(2008)资助。

$\leqslant n, n \in N\}$ 是一致可积.

事实上,由

$$E|X_k|^p I_{(|X_k| \geqslant x)} =$$

$$(\int_0^{x^p} + \int_{x^p}^{\infty}) P(|X_k|^p I_{(|X_k| \geqslant x)} > t) \mathrm{d}t = x^p P(|X_k| \geqslant x) + \int_{x^p}^{\infty} P(|X_k|^p > t) \mathrm{d}t,$$

知(1)式等价于

$$\lim_{x \to \infty} \sup_{n \in N} \sum_{k=1}^{n} |a_{nk}|^p x^p P(|X_k| \geqslant x) = 0, \quad (2)$$

和

$$\lim_{x \to \infty} \sup_{n \in N} \sum_{k=1}^{n} |a_{nk}|^p \int_{x^p}^{\infty} P(|X_k|^p > t) \mathrm{d}t = 0. \quad (3)$$

或者等价于

$$\lim_{x \to \infty} \sup_{n \in N} \sum_{k=1}^{n} |a_{nk}|^p x P(|X_k|^p \geqslant x) = 0 \quad (4)$$

和(3)式. 而条件

$$\lim_{x \to \infty} \sup_{n \in N} \sum_{k=1}^{n} |a_{nk}|^p x \ln^{\delta} x P(|X_k|^p \geqslant x) = 0, \quad \delta > 1 \quad (5)$$

蕴含(1)式. 因为由(5)式知,对 $\forall \varepsilon > 0, \exists M > 0$,当 $x > M$ 时有

$$\sup_{n \in N} \sum_{k=1}^{n} |a_{nk}|^p P(|X_k|^p \geqslant x) < \varepsilon x^{-1} \ln^{-\delta} x.$$

由此得(5)式蕴含(3)式,显然也蕴含(4)式,所以(5)式蕴含(1)式.

1 相关引理

引理 1[4] 若基本假设 A_1 成立,则

$$\lim_{x \to \infty} |Eg_n(x) - g(x)| = 0.$$

引理 2[7] 设$\{\varepsilon_j, j \in N\}$是 PA 随机变量序列,对所有的 $j \in N$,满足 $E\varepsilon_j = 0$, $|\varepsilon_j| \leqslant B < \infty$. 记 $S_n = \sum_{j=1}^{n} \varepsilon_j$, $u(n) = \sup_j \sum_{i:|i-j| \geqslant n} \mathrm{Cov}(\varepsilon_j, \varepsilon_k)$. 假设对某个 $r > 2$, $u(n) = O(n^{-(r-2)/2})$,则存在不依赖于 n 的常数 C,使得对所有的 $n \in N$,

$$\sup_{m \in N \cup \{O\}} E|S_{n+m} - S_n|^r \leqslant Cn^{r/2}.$$

引理 3[5] 设$\{a_j, j \in N\}$是一实数列,记 $a = \sup|a_j|$ 且 $a < \infty$. 又设$\{\varepsilon_j, j \in N\}$是 PA 随机变量序列,对所有的 $j \in N$,满足 $E\varepsilon_j = 0$, $|\varepsilon_j| \leqslant B < \infty$. 记 $S_n = \sum_{j=1}^{n} \varepsilon_j$, $u(n) = \sup_j \sum_{i:|i-j| \geqslant n} Cov(\varepsilon_j, \varepsilon_k)$. 假设对某个 $r > 2$, $u(n) = O(n^{-(r-2)/2})$,则存在不依赖于 n 的常数 C,使

$$E|\sum_{j=1}^{n} a_j \varepsilon_j|^r \leqslant Ca^r n^{r/2}.$$

2 主要结果

以下用 C 记与 n 无关的正常数,且 C 在不同的地方表示不同的值,即使在同一式中也是如此.

定理 1 若基本假设 A_1 成立,且存在 $p > 2, \delta > 2$,使得下列条件满足:

(i) 对每个 n, $\{\varepsilon_i, 1 \leqslant i \leqslant n\}$ 是具有零均值的 PA 相依序列.

(ii) $w_n(x) \triangleq \max_{1 \leqslant i \leqslant n} |w_{ni}(x)| = O(n^{-\alpha})$,其中 $\frac{1}{2} + \frac{1}{p} < \alpha \leqslant 1$.

(iii) $u(n) = O(n^{-(r-2)/2})$,其中 r 满足 $(\alpha - \frac{1}{2} - \frac{1}{p})r > 1$.

(iv) 对(ii)中的 α,若

$$\lim_{y \to \infty} \sup_{n \in N} \sum_{i=1}^{n} |w_{ni}(x)|^{1-1/(\alpha p)} y \ln^{\delta} y P(|\varepsilon_i|^p \geqslant y) = 0.$$

则对 $\forall x \in A$, $\lim_{x \to \infty} g_n(x) = g(x)$, a. s. .

证明 注意到 $|g_n(x) - g(x)| \leqslant |g_n(x) - Eg_n(x)| + |Eg_n(x) - g(x)| = |\sum_{i=1}^{n} w_{ni}(x) \varepsilon_i| + |Eg_n(x) - g(x)|$.

由引理 3,只须证明 $|\sum_{i=1}^{n} w_{ni}(x) \varepsilon_i| \to 0$, a. s. .

令 $\varepsilon_i(1) = -n^{1/p} I_{(\varepsilon_i < -n^{1/p})} + \varepsilon_i I_{(|\varepsilon_i| \leqslant n^{1/p})} + n^{1/p} I_{(\varepsilon_i > n^{1/p})}$, $\varepsilon_i(2) = (\varepsilon_i + n^{1/p}) I_{(\varepsilon_i < -n^{1/p})} + (\varepsilon_i - n^{1/p}) I_{(\varepsilon_i > n^{1/p})}$, $\tilde{\varepsilon}_i(1) = \varepsilon_i(1) - E\varepsilon_i(1)$, $\tilde{\varepsilon}_i(2) = \varepsilon_i(2) - E\varepsilon_i(2)$,显然 $\varepsilon_i = \varepsilon_i(1) + \varepsilon_i(2)$. 记

$$S'_n(x) = \sum_{i=1}^{n} w_{ni}(x) \tilde{\varepsilon}_i(1), S''_n(x) = \sum_{i=1}^{n} w_{ni}(x) \tilde{\varepsilon}_i(2), S_n(x) = \sum_{i=1}^{n} w_{ni}(x) \varepsilon_i = S'_n(x) + S''_n(x).$$

故为证明定理 1,只需证明 $|S'_n(x)| \to 0$, $|S''_n(x)| \to 0$. a. s. .

先证明 $|S'_n(x)| \to 0$, a. s. .

由上述取法知 $\varepsilon_i(1)$, $\varepsilon_i(2)$ 均为 ε_i 的增函数,于是

$$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = \mathrm{Cov}(\varepsilon_i(1) + \varepsilon_i(2), \varepsilon_j(1) + \varepsilon_j(2)) = \mathrm{Cov}(\varepsilon_i(1), \varepsilon_j(1)) + \mathrm{Cov}(\varepsilon_i(1), \varepsilon_j(2)) + \mathrm{Cov}(\varepsilon_i(2), \varepsilon_j(1)) + \mathrm{Cov}(\varepsilon_i(2), \varepsilon_j(2)) \geqslant \mathrm{Cov}(\varepsilon_i(1), \varepsilon_j(1)) = \mathrm{Cov}(\tilde{\varepsilon}_i(1), \tilde{\varepsilon}_j(1)).$$

而 $\{n^{-1/p} \tilde{\varepsilon}_i(1), 1 \leqslant i \leqslant n\}$ 也是 PA 的,且 $E[n^{-1/p} \tilde{\varepsilon}_i(1)] = 0$, $|n^{-1/p} \tilde{\varepsilon}_i(1)| \leqslant 1$. 又 $V(n) \triangleq$

$\sup_j \sum_{i:|i-j|\geqslant n} \mathrm{Cov}(n^{-1/p}\tilde{\varepsilon}_i(1), n^{-1/p}\tilde{\varepsilon}_j(1)) =$
$n^{-2/p} \sup_j \sum_{i:|i-j|\geqslant n} \mathrm{Cov}(\tilde{\varepsilon}_i(1), \tilde{\varepsilon}_j(1)) \leqslant u(n) =$
$O(n^{-(r-2)/2})$.

由引理 2 **和** Markov **不等式,对** $\forall \varepsilon > 0$ **得**

$$P(|S'_n(x)| > \varepsilon) \leqslant CE|S'_n(x)|^r =$$
$$Cn^{r/p}E|\sum_{i=1}^n w_{ni}(x)n^{-1/p}\tilde{\varepsilon}_i(1)|^r \leqslant Cn^{r/p}w_n^r(x)n^{r/2} \leqslant$$
$$Cn^{r/p}n^{-\alpha r}n^{r/p} \leqslant Cn^{-(\alpha-\frac{1}{2}-\frac{1}{p})r}.$$

因为 $(\alpha - \frac{1}{2} - \frac{1}{p})r > 1$**,所以** $\sum_{n=1}^{\infty} P(|S'_n(x)| > \varepsilon) < \infty$. **由** Borel-Canttli **引理知** $|S'_n(x)| \to 0$, a. s..

再证明 $|S''_n(x)| \to 0$, a. s.

记 $T_n(x) = \sum_{i=1}^n w_{ni}(x)\varepsilon_i(2)$**,则** $S''_n(x) = T_n(x) - ET_n(x)$. **又对** $\forall \varepsilon > 0$**,由** Markov **不等式得**

$$\sum_{n=1}^{\infty} P(|T_n(x)| > \varepsilon) \leqslant C\sum_{n=1}^{\infty} E|T_n(x)| \leqslant$$
$$C\sum_{n=1}^{\infty}\sum_{i=1}^{n} |w_{ni}(x)| E|\varepsilon_i(2)| \leqslant$$
$$C\sum_{n=1}^{\infty}\sum_{i=1}^{n} |w_{ni}(x)| E|\varepsilon_i| I_{(|\varepsilon_i| > n^{1/p})} \leqslant$$
$$C\sum_{n=1}^{\infty} n^{-(p-1)/p}\sum_{i=1}^{n} |w_{ni}(x)| E|\varepsilon_i|^p I_{(|\varepsilon_i| > n^{1/p})} \leqslant$$
$$C\sum_{n=1}^{\infty} n^{-(p-1)/p}\sum_{i=1}^{n} |w_{ni}(x)| [n^{1/p}P(|\varepsilon_i|^p > n^{1/p}) +$$
$$\int_n^{\infty} P(|\varepsilon_i|^p > t)\mathrm{d}t] \triangleq I_n. \tag{6}$$

由条件(iv) **知,** $\exists M > 0$**,当** $y > M$ **时,**

$$\sup_{n\in N}\sum_{i=1}^{n} |w_{ni}(x)|^{1-1/(\alpha p)} P(|\varepsilon_i|^p > y) <$$
$$y^{-1}\ln^{-\delta} y.$$

又 $n^{1/p} \to \infty$**,所以** $\exists n_0$ **使得** $n \geqslant n_0$ **时,有** $n^{1/p} > M$. **所以**

$$I_n \leqslant C\sum_{n=1}^{n_0-1} n^{-(p-1)/p}\sum_{i=1}^{n} |w_{ni}(x)| E|\varepsilon_i(2)| +$$
$$C\sum_{n=n_0}^{\infty} n^{-(p-1)/p} w_n^{1/(\alpha p)}(x)\sum_{i=1}^{n} |w_{ni}(x)|^{1-1/(\alpha p)} \cdot$$
$$[n^{1/p}P(|\varepsilon_i|^p > n^{1/p}) + \int_n^{\infty} P(|\varepsilon_i|^p > t)\mathrm{d}t] \leqslant C +$$
$$C\sum_{n=n_0}^{\infty} n^{-(p-1)/p} n^{-\alpha\cdot 1/(\alpha p)}(\ln n^{1/p})^{-\delta} +$$
$$C\sum_{n=n_0}^{\infty} n^{-1}\int_n^{\infty} t^{-1}\ln^{-\delta} t\,\mathrm{d}t \leqslant C + C\sum_{n=n_0}^{\infty}\frac{1}{n\ln^{\delta} n} +$$
$$C\sum_{n=n_0}^{\infty}\frac{1}{n\ln^{\delta-1} n} \leqslant C\sum_{n=1}^{\infty}\frac{1}{n\ln^{\delta-1} n} < \infty, \delta > 2.$$

其中倒数第 2 **个不等式成立是因为当** n **较大时** $\ln^{\delta} n > \ln^{\delta-1} n$.

因此由(6) **式知,** $|T_n(x)| \to 0$, a. s. **且** $E|T_n(x)| \to 0$. **从而** $|S''_n(x)| \to 0$, a. s..

注 1 **文献**[5] **要求对每一个** n, $\sup_{1\leqslant i\leqslant n} E|\varepsilon_i|^p < \infty$, $p > 1$. **而事实上,** $E|\varepsilon_i|^p < \infty$ **意味着**

$$E|\varepsilon_i|^p I_{(|\varepsilon_i| \geqslant y)} \to 0, y \to \infty,$$

从而有 $yP(|\varepsilon_i|^p \geqslant y) \to 0, y \to \infty$. **因此只要权函数选取恰当,就可以得到本文的条件**(iv). **而本文对权函数的其他条件和文献**[5] **基本一致,所以本文的条件弱于文献**[5]**的条件.**

参考文献:

[1] **杨善朝. 基于鞅序列非参数回归函数的估计**[J]. **系统科学与数学**, 1999, 19(1): 79-85.

[2] Roussas G G, Tranand L T, Ioannides D A. Fixed design regression for time series: asymptotic normality [J]. Multivariate Anal, 1992, 40: 162-291.

[3] Yang S C. Uniformly asymptotic normality of the regression weighted estimator for negatively association samples[J]. Stat Prob Lett, 2003, 62: 101-110.

[4] Fan Y. Consistent nonparametric multiple regression for dependent heterogeneous processes: the fix design case [J]. Multivariate Anal, 1990, 33: 72-88.

[5] **黎玉芳, 杨善朝.** PA **样本非参数回归权函数估计的强相合性**[J]. **广西科学**, 2003, 10(3): 171-175.

[6] Ordóñez Cabrera M. Convergence of weighted sums of random variable and uniform integrability concerning the weights[J]. Collect Math, 1994, 43: 121-132.

[7] Birkel T. Moment bounds for associated sequences[J]. Ann Probab, 1988, 16: 1184-1193.

(责任编辑:尹 闯)