Strong Consistency and Convergence Rate of Modified Partitioning Estimate of Nonparametric Regression Function under α -Mixing Sample

YAO Mei, DU Xue Qiao

(Department of Mathematics, Hefei University of Technology, Anhui 230009, China) (E-mail: ymnancy@sina.com)

Abstract In this paper, we study the strong consistency and convergence rate of modified partitioning estimate of nonparametric regression function under the sample $\{(X_i, Y_i), i \geq 1\}$ that is α sequence taking values in $\mathbb{R}^d \times \mathbb{R}^1$ with identical distribution.

Keywords nonparametric regression function; modified partitioning estimate; strong consistency; convergence rate; α -mixing.

Document code A MR(2000) Subject Classification 62G05; 62G08 Chinese Library Classification 0212.7

1. Introduction and some lemmas

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be α -mixing sample sequence from (X, Y) in $\mathbb{R}^d \times \mathbb{R}^1$ with identical distribution and $(E|Y| < \infty)$. The regression function of Y for given X is defined as $m(x) = E(Y|X = x)(x \in \mathbb{R}^d)$.

As far as we know, for methods of estimating m(x) based on samples $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, there are kernel estimate, k-nearest neighbor estimate and partitioning estimate etc.^[1] The bulk of the literature in these areas have focused on large sample properties of modified kernel estimate and modified nearest neighbor estimate for the regression function under i.i.d samples and some dependence samples such as α -mixing, φ -mixing^[2-10]. The literature [11] and [12] only proved the strong consistence of modified partitioning estimate for the regression function under i.i.d samples and φ -mixing samples. However, the strong consistence of modified partitioning estimate for the regression function has not been proved under the α -mixing samples with the better weakly mixing dependent conditions. This paper concerns itself with the strong consistency and convergence rate of modified partitioning estimate for the regression function under α -mixing samples. The result shows that the convergence rate of modified partitioning estimate under φ -mixing samples is similar to that of modified partitioning estimate under φ -mixing samples.

For each $n \ge 1$, let $\mathcal{A}_n = \{A_{ni} : i \ge 1\}$ be a partition of \mathbb{R}^d into a finite or countable collection of Borel subsets. For $x \in \mathbb{R}^d$, let $A_n(x)$ denote the atom of partition that contains the

Received date: 2006-07-06; Accepted date: 2007-05-26

point x.

$$A_n(x) = A_{ni}, \text{ if } x \in A_{ni}.$$

The modified partitioning estimate of regression function is defined as

$$\widetilde{m_n}(x) = \frac{\sum_{i=1}^n I_{A_n(x)}(X_i) \cdot Y_i I(|Y_i| \le b_n)}{\sum_{i=1}^n I_{A_n(x)}(X_i)}, \quad x \in \mathbb{R}^d.$$
(1.1)

Where $I_A(x) = I\{x \in A\}$ is the indicator function of a set $A, b_n \to \infty(n \to \infty)$, and 0/0 is 0 by convention.

For simplicity, we use the following symbols in this paper:

- (1) F denotes the distribution of X;
- (2) S(F) denotes the support of F;
- (3) C > 0 denotes a constant;
- (4) C(x) > 0 denotes a constant depending upon x.

(These constants can be assumed to be different values in their appearance, even within the same expression.)

Lemma 1 Suppose

(i) The partition $\mathcal{A}_n = \{A_{ni}; i \geq 1\}$ are monotonically increasing and satisfy that for each sphere S centered at the origin, there holds

$$\sup_{i:A_{ni}\cap S\neq\emptyset} \operatorname{supdiam} A_{ni} \to 0, \ n \to \infty.$$
(1.2)

(ii) The function g(x) is Borel-measurable and satisfies $\int |g(x)| F(dx) < \infty$.

Then

$$\lim_{n \to \infty} \int_{A_n(x)} |g(x) - g(t)| \frac{F(\mathrm{d}t)}{F(A_n(x))} = 0, \quad \text{a.e. } [F]$$
(1.3)

or

$$\lim_{n \to \infty} \int_{A_n(x)} g(t) \frac{F(\mathrm{d}t)}{F(A_n(x))} = g(x), \quad \text{a.e. } [F]$$
(1.4)

Proof See Lemma 1 in [11].

Lemma 2 Let $\nu_n = \nu_n(x)$ be the volume of the set $A_n(x)$ and the condition (i) in Lemma 1 exists. Then there exists a nonnegative function L(x) with $L(x) < \infty$ such that

$$\frac{v_n}{F(A_n(x))} \to L(x), \quad n \to \infty \quad \text{a.e.} \ [F]$$
(1.5)

Proof Refer to the proof of Lemma 2.2 in [13].

Lemma 3 (Bernstein inequality for α -mixing processes) Let $\{X_i, i \ge 1\}$ be a sequence of α mixing random variables verifying $\alpha(n) = o(\rho^n)$ ($0 < \rho < 1$), $E(X_i) = 0, |X_i| \le 1$. Denote $\gamma = \frac{2}{1-\theta}$, and $\sigma = \sup\{(E|\xi_i|^{\gamma})^{\frac{1}{\gamma}} : i \in N, N = \{1, 2, 3, \ldots\}\}$. If $n^{\frac{1}{2}}\sigma \ge 1$, then there exist constants c_1 and c_2 which depend only on the mixing coefficients such that for $\varepsilon > 0, 0 < \theta < 1$,

638

Strong consistency and convergence rate of modified partitioning estimate under α -mixing sample 639

we have

$$P(|\sum_{i=1}^{n} x_i| > \varepsilon) \le c_1 \theta^{-1} \exp(-c_2 \varepsilon^{\frac{1}{2}}).$$
(1.6)

Proof See Lemma 3.2 in [14].

In the following we will repeatedly make use of Lemmas 1 and 2, and for some r.v.Z with $E|Z| < \infty$, there is

$$\lim_{b \to \infty} E(ZI_{(|z|>b)} | X = x) = 0, \text{ a.e. } [F].$$
(1.7)

For the formulas (1.3)-(1.5) and (1.7), there is an exceptive set on which the related formula may not be true. These exceptive sets sum up to *F*-null set. Without loss of generality, we can suppose that this set is empty.

2. Strong consistency of modified partitioning estimate

Theorem 2.1 Suppose that $\{(X_i, Y_i)\}$ is a α -mixing sequence with $(X_i, Y_i) \stackrel{d}{=} (X, Y), i = 1, 2, \ldots, n, E|Y| < \infty$, and the condition (i) of Lemma 1 exists. If the mixing rate $\alpha(\cdot)$ and $\nu_n = \nu_n(x)$ satisfy the following conditions:

(i)
$$\alpha(n) = o(\rho^n) \ (0 < \rho < 1); \ \alpha(n) \downarrow$$

(ii)

$$\frac{\sqrt{nv_n}}{\sqrt{b_n} \cdot \log n} \to \infty, \quad n \to \infty, \quad \text{a.e.} \ [F]$$
(2.1)

where, $b_n \to \infty$ $(n \to \infty)$. Then

$$\widetilde{m_n}(x) \xrightarrow{\text{a.s.}} m(x), \quad n \to \infty, \quad \text{a.e. } x \in \mathbb{R}^d.$$
 (2.2)

Proof Take $x \in S(F)$. Set

$$m_1(x) = E[Y \cdot I(|Y| > b_n) | X = x],$$

$$g(t) = |m(t) - m(x)|,$$

$$\tilde{g}_b(x) = E[|Y|I(|Y| > b) | X = x].$$

Suppose that $\sum_{i=1}^{n} I_{A_n(x)}(X_i) > 0$. Then we have

$$m_n(x) - m(x) = \frac{S_n(x)}{T_n(x)},$$
(2.3)

where

$$S_n(x) = \sum_{i=1}^n \frac{Z_{ni}(x)}{nF(A_n(x))}, \quad T_n(x) = \sum_{i=1}^n \frac{I_{A_n(x)}(X_i)}{nF(A_n(x))},$$
$$Z_{ni}(x) = I_{A_n(x)}(X_i)(Y_iI(|Y_i| \le b_n) - m(x)).$$

Now we come to prove

$$S_n(x) \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \quad \text{a.e.} [F]$$
 (2.4)

$$T_n(x) \xrightarrow{\text{a.s.}} 1, \quad n \to \infty, \quad \text{a.e.} \ [F]$$

$$(2.5)$$

By $(X_i, Y_i) \stackrel{d}{=} (X, Y), i = 1, 2, \dots, n$, we know that

$$ES_n(x) = \frac{E\{E(Z_{n1}(x)|X_1)\}}{F(A_n(x))}.$$

Since

$$Z_{n1}(x) = I_{A_n(x)}(X_1)(Y_1I(|Y_1| \le b_n) - m(x))$$

= $I_{A_n(x)}(X_1)(Y_1 - m(x) - Y_1I(|Y_1| > b_n)),$

we have

$$|ES_{n}(x)| \leq \left| \frac{E\{I_{A_{n}(x)}(X_{1})[E(Y|X_{1}) - m(x)]\}}{F(A_{n}(x))} \right| + \left| \frac{E(I_{A_{n}(x)}(X_{1})m_{1}(X_{1}))}{F(A_{n}(x))} \right|$$

= $I_{1n}(x) + I_{2n}(x).$ (2.6)

Using Lemma 1, we get

$$\begin{split} I_{1n}(x) &= \Big| \int_{A_n(x)} (m(t) - m(x)) \frac{F(\mathrm{d}t)}{F(A_n(x))} \Big| \\ &\leq \int_{A_n(x)} g(t) \frac{F(\mathrm{d}t)}{F(A_n(x))} \to 0, \quad n \to \infty, \quad \text{a.e. } [F], \end{split}$$

or

$$I_{1n}(x) \to 0, \ n \to \infty, \ \text{a.e.} [F].$$
 (2.7)

On the other hand, we have

$$I_{2n}(x) \le \int_{A_n(x)} |m_1(t)| \frac{F(\mathrm{d}t)}{F(A_n(x))}.$$

Take b > 0. From $b_n \to \infty$ $(n \to \infty)$ and (1.4) of Lemma 1, by a method similar to the proof of the formula (38) in [10], we get

$$\lim_{n \to \infty} \sup \int_{A_n(x)} |m_1(t)| \frac{F(\mathrm{d}t)}{F(A_n(x))} \le \lim_{b \to \infty} \lim_{n \to \infty} \int_{A_n(x)} \widetilde{g_b}(t) \frac{F(\mathrm{d}t)}{F(A_n(x))} = \lim_{b \to \infty} \widetilde{g_b}(x) = 0, \text{ a.e. } [F].$$

Thus, we get

$$I_{2n}(x) \to 0, \ n \to \infty, \ \text{a.e.} [F].$$
 (2.8)

Therefore, we have

$$ES_n(x) \to 0, \quad n \to \infty, \quad \text{a.e.} \ [F].$$
 (2.9)

Note that

$$S_n(x) - ES_n(x) = \sum_{i=1}^n \frac{Z_{ni}(x) - EZ_{ni}(x)}{nF(A_n(x))}$$

For simplicity, we denote

$$\xi_i = \frac{Z_{ni}(x) - EZ_{ni}(x)}{b_n + |m(x)|}.$$

Obviously, $\{\xi_i, i \geq 1\}$ is a $\alpha\text{-mixing}$ sequence with identical distribution and,

$$E\xi_i = 0, \ |\xi_i| = \frac{|Z_{ni}(x) - EZ_{ni}(x)|}{b_n + |m(x)|} \le 1;$$

640

When $\gamma > 2$ is given, $\sigma = \sup\{(E|\xi_i|^{\gamma})^{\frac{1}{\gamma}} : i \in N, N = \{1, 2, 3, \ldots\}\} = 1, n^{\frac{1}{2}}\sigma \ge 1$. Using Lemma 3, for any $\varepsilon > 0$, when n is large enough, we get

$$P(|S_n(x) - ES_n(x)| > \varepsilon) = P\left(\left|\sum_{i=n}^n \xi_i\right| > \frac{\varepsilon n F(A_n(x))}{b_n + |m(x)|}\right)$$
$$\leq c_1 \theta^{-1} \exp\left(-c_2 \frac{\sqrt{\varepsilon n} F^{\frac{1}{2}}(A_n(x))}{\sqrt{b_n + |m(x)|}}\right).$$
(2.10)

Note $p_n = F(A_n(x))$. By Lemma 2, there is C(x) > 0 such that $p_n \ge C(x)v_n$, a.e. [F]. Since $|m(x)| < \infty$ and $b_n \to \infty$ $(n \to \infty)$. From (2.1), for any $\varepsilon > 0$, when n is large enough, we have

$$\begin{aligned} P(|S_n(x) - ES_n(x)| > \varepsilon) &\leq C \exp\left(-\frac{\sqrt{C(x)nv_n}}{\sqrt{b_n}}\right) \\ &\leq C\frac{1}{n^2}, \ \text{a.e.}[F]. \end{aligned}$$

Thus, by the Borel-Cantelli Lemma, we get

$$S_n(x) - ES_n(x) \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \quad \text{a.e.} \ [F].$$
 (2.11)

Therefore, from (2.9) and (2.11), it is trivial to get (2.4). Now we come to prove (2.5). Let $\zeta_i = \frac{1}{2}(I_{A_n(x)}(X_i) - EI_{A_n(x)}(X_i))$. Then $\{\zeta_i, i \ge 1\}$ is also a α – mixing sequence with

$$E\zeta_i = 0, \quad |\zeta_i| \le 1, \quad \text{a.s.}$$

When $\gamma > 2$ is given, $\sigma' = \sup\{(E|\zeta_i|^{\gamma})^{\frac{1}{\gamma}} : i \in N, N = \{1, 2, 3, \ldots\}\} = 1, n^{\frac{1}{2}}\sigma' \geq 1$. Using Lemma 3, for any $\varepsilon > 0$, when n is large enough, we have

$$P(|T_n(x) - 1| > \varepsilon) = P\left(|\sum_{i=n}^n \zeta_i| > \frac{\varepsilon n p_n}{2}\right)$$

$$\leq C \exp(-\sqrt{n p_n})$$

$$\leq C \exp\left(-\sqrt{C(x) n v_n}\right), \text{ a.e. } [F].$$
(2.12)

Since $b_n \to \infty$ $(n \to \infty)$ and (2.1) is valid, we have

$$P(|T_n(x) - 1| > \varepsilon) \le C \exp\left(-\sqrt{C(x)nv_n}\right) \le C\frac{1}{n^2}, \text{ a.e. } [F].$$

Therefore, by the Borel-Cantelli Lemma, we know that (2.5) is valid. By (2.3), (2.4) and (2.5), when $x \in S(F)$ is given, we get

$$\widetilde{m_n}(x) \xrightarrow{\text{a.s.}} m(x), \quad n \to \infty.$$

We shall get (2.2) by Fubini's Theorem since the *F*-measure of S(F) is 1. This completes the proof of the theorem. \Box

3. Convergence rate of modified partitioning estimate

In order to discuss the convergence rate, let m(x) satisfy the following condition: there exist D(x) > 0 and $0 < q \leq 1$, and for a.e. x[F], there exists $\delta = \delta(x) > 0$, such that when

 $v_n = v_n(x) < \delta$, we have

$$\int_{A_n(x)} |m(t) - m(x)| \frac{F(\mathrm{d}t)}{F(A_n(x))} \le D(x) v_n^q, \tag{3.1}$$

where $v_n(x)$ is the volume of the set $A_n(x)$.

Theorem 3.1 Let (X_i, Y_i) ; $i \ge 1$ be a sequence of α -mixing random variables verifying $(X_i, Y_i) \stackrel{d}{=} (X, Y), i = 1, 2, ..., n$; Suppose that for some $r > 1, E|Y|^r < \infty$, the condition (i) of Lemma 1 and the condition (i) of Theorem (2.1) exist and the regression function m(x) satisfies (3.1). Take

$$v_n = \left(\frac{1}{\sqrt{n}}\log^2 n\right)^{\frac{3(r-1)}{4(r+1)(q+1)}}, \text{ a.e. } [F]$$
$$b_n = v_n^{-\frac{q+1}{r-1}}.$$

Then

$$\widetilde{m_n}(x) - m(x) = O(v_n^q), \quad \text{a.s. a.e. } x \in \mathbb{R}^d.$$
(3.2)

Proof We prove that for any positive number sequence $\{M_n, n \ge 1\}$, with $\lim_{n\to\infty} M_n = \infty$, $\lim_{n\to\infty} v_n^q M_n = 0$, we have

$$\lim_{n \to \infty} v_n^{-q} M_n^{-1}(\widetilde{m}_n(x) - m(x)) = 0, \quad \text{a.s.}, \quad \text{a.e.} \ [F].$$
(3.3)

In fact, take $x \in S(F)$. Then it follows from (1.4), (2.3), (2.5) and (3.1)

$$\begin{aligned} |v_n^{-q}M_n^{-1}ES_n(x)| &\leq v_n^{-q}M_n^{-1}I_{1n}(x) + v_n^{-q}M_n^{-1}I_{2n}(x) \\ &\leq v_n^{-q}M_n^{-1}\int_{A_n(x)} g(t)\frac{F(\mathrm{d}t)}{F(A_n(x))} + v_n^{-q}M_n^{-1}\int_{A_n(x)} |m_1(t)|\frac{F(\mathrm{d}t)}{F(A_n(x))} \end{aligned}$$

where

$$\begin{split} v_n^{-q} M_n^{-1} \int_{A_n(x)} g(t) \frac{F(\mathrm{d}t)}{F(A_n(x))} &= v_n^{-q} M_n^{-1} \int_{A_n(x)} |m(t) - m(x)| \frac{F(\mathrm{d}t)}{F(A_n(x))} \\ &\leq v_n^{-q} M_n^{-1} D(x) v_n^q \\ &\leq M_n^{-1} D(x) \to 0, \ n \to \infty, \ \text{a.e.} \ [F] \\ v_n^{-q} M_n^{-1} \int_{A_n(x)} |m_1(t)| \frac{F(\mathrm{d}t)}{F(A_n(x))} &= v_n^{-q} M_n^{-1} \int_{A_n(x)} |E[YI(|Y| > b_n)| X = t] |\frac{F(\mathrm{d}t)}{F(A_n(x))} \\ &\leq v_n^{-q} M_n^{-1} b_n^{1-r} \int_{A_n(x)} E[|Y|^r |X = t] \frac{F(\mathrm{d}t)}{F(A_n(x))} \\ &\leq M_n^{-1} \int_{A_n(x)} E[|Y|^r |X = t] \frac{F(\mathrm{d}t)}{F(A_n(x))}. \end{split}$$

Since $E|Y|^r < \infty$, by Lemma 1, we get

$$v_n^{-q} M_n^{-1} \int_{A_n(x)} |m_1(t)| \frac{F(\mathrm{d}t)}{F(A_n(x))} \le M_n^{-1} \int_{A_n(x)} E[|Y|^r | X = t] \frac{F(\mathrm{d}t)}{F(A_n(x))} \longrightarrow 0, \quad n \to \infty$$

and consequently,

$$v_n^{-q} M_n^{-1} ES_n(x) \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \quad \text{a.e.} \ [F].$$
 (3.4)

On the other hand, using the same method with which we prove (2.10) in Theorem 2.1, from the values of v_n and b_n taken, we get

$$P(v_n^{-q}M_n^{-1}|S_n(x) - ES_n(x)| > \varepsilon) = P(|S_n(x) - ES_n(x)| > \varepsilon v_n^q M_n)$$

$$\leq c_1 \theta^{-1} \exp\left(-c_2 \frac{\sqrt{\varepsilon n M_n v_n^q F(A_n(x))}}{(b_n + |m(x)|)^{\frac{1}{2}}}\right).$$

Write $p_n = F(A_n(x))$. By Lemma 2, there exists C(x) > 0 such that $p_n \ge C(x)v_n$, a.e. [F]. Since $|m(x)| < \infty$ and $b_n \to \infty$ $(n \to 0)$, for any $\varepsilon > 0$ and n large enough, we have by (2.1)

$$P(v_n^{-q}M_n^{-1}|S_n(x) - ES_n(x)| > \varepsilon) \le C \exp\left(-C(x)\frac{\sqrt{nv_n^{q+1}M_n}}{\sqrt{b_n}}\right)$$
$$\le C \exp\left(-M_n^{\frac{1}{2}}\log^2 n\left(\frac{\sqrt{n}}{\log^2 n}\right)^{\frac{5}{8}}\right)$$
$$\le C \exp\left(-M_n^{\frac{1}{2}}\log n\right)$$

and consequently,

$$P(v_n^{-q}M_n^{-1}|S_n(x) - ES_n(x)| > \varepsilon) \le C \exp\left(-M_n^{\frac{1}{2}}\log n\right) \le C\frac{1}{n^2}, \text{ a.e. } [F].$$
(3.5)

Thus, by Borel-Cantelli Lemma, we get

$$v_n^{-q}M_n^{-1}(S_n(x) - ES_n(x)) \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \quad \text{a.e.} \ [F]. \tag{3.6}$$

It follows from (3.4) and (3.6)

$$v_n^{-q} M_n^{-1} S_n(x) \xrightarrow{\text{a.s.}} 0n \to \infty, \text{ a.e. } [F].$$
 (3.7)

By (2.3), (2.5) and (3.7): when $x \in S(F)$ is given, we get

$$\lim_{n \to \infty} v_n^{-q} M_n^{-1}(\widetilde{m}_n(x) - m(x)) = 0, \quad \text{a.s.}.$$

We can get (3.2) by Fubini's Theorem since the F-measure of S(F) is 1.

References

- ALGOET P, GYÖRFI L. Strong universal pointwise consistency of some regression function estimates [J]. J. Multivariate Anal., 1999, 71(1): 125–144.
- CHENG Ping. On the strong consistency and convergence rate of improved kernel estimates for the regression function [J]. J. Systems Sci. Math. Sci., 1983, 3(4): 304–315. (in Chinese)
- HU Shuhe. Consistency of an improved kernel estimate for regression functions [J]. J. Systems Sci. Math. Sci., 1993, 13(2): 141–151. (in Chinese)
- [4] QIAN Weimin, MAMMITZSCH V. Strong uniform convergence for the estimator of the regression function under φ-mixing conditions [J]. Metrika, 2000, 52(1): 45–61.
- [5] ZHAO Xia. Asymptotic distribution of the improved kernel estimates for regression functions [J]. Chinese J. Appl. Probab. Statist., 2001, 17(1): 81–87. (in Chinese)
- [6] ZHENG Zukang. A class of estimators of the parameters in linear regression with censored data [J]. Acta Math. Sinica (N.S.), 1987, 3(3): 231–241.
- HU Shuhe. Distribution-free consistency of kernel estimates of regression functions [J]. Acta Math. Appl. Sinica, 1997, 20(3): 448–455. (in Chinese)
- [8] HONG Shengyan. Asymptotic normality of nearest neighbour regression estimators [J]. Chinese J. Appl. Probab. Statist., 1991, 7(2): 187–191. (in Chinese)

- [9] HU Shuhe. Distribution-free consistency of nearest neighbor kernel estimates of regression functions [J]. Acta Math. Sinica, 1995, 38(4): 559–567. (in Chinese)
- [10] XUE Liugen. Strong consistency of two forms improved estimates of Regression function for dependent sample [J]. Math. Stati. Appl. Prob., 1993, 8(1): 82–93. (in Chinese)
- [11] ZAO Lincheng, DIAO Guoqing. Strong convergence of modified partitioning estimates of nonparametric regression functions [J]. Chinese J. Appl. Probab. Statist., 2002, 18(2): 192–196.
- [12] LING Nengxiang. Strong convergence and its rate of modified partitioning estimation for nonparametric regression function under dependence samples [J]. Northeast. Math. J., 2004, 20(3): 349–354.
- [13] DEVROYE L. On the almost everywhere convergence of nonparametric regression function estimates [J]. Ann. Statist., 1981, 9(6): 1310–1319.
- [14] BOENTE G, FRAIMAN R. Consistency of a nonparametric estimate of a density function for dependent variables [J]. J. Multivariate Anal., 1988, 25(1): 90–99.