COMMUNICATIONS IN MATHEMATICAL RESEARCH
29(1)(2013), 32–40

Strong Consistency of M Estimator in Linear Model for $\tilde{\varphi}$-mixing Samples*

WANG XUE-JUN, HU SHU-HE, LING JI-MIN, WEI YUN-FEI AND CHEN ZHU-QIANG
(*School of Mathematical Science, Anhui University, Hefei, 230601*)

Communicated by Wang De-hui

Abstract: The strong consistency of M estimator of regression parameter in linear model for $\tilde{\varphi}$-mixing samples is discussed by using the classic Rosenthal type inequality. We get the strong consistency of M estimator under lower moment condition, which generalizes and improves the corresponding ones for independent sequences.

Key words: $\tilde{\varphi}$-mixing sample, M estimator, strong consistency
2000 MR subject classification: 62J05, 62F12
Document code: A
Article ID: 1674-5647(2013)01-0032-09

1 Introduction

Consider the following linear model:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta}_0 + e_i, \qquad i = 1, 2, \cdots, n, \ n \in \mathbf{N}, \tag{1.1}$$

where $\boldsymbol{x}_i$ $(i = 1, 2, \cdots, n)$ is a known p-dimensional vector, $\boldsymbol{\beta}_0$ is an unknown p-dimensional regression parametric vector, and $e_1, e_2, \cdots, e_n$ are random errors. Let f be a convex function on $\mathbf{R}$. The M estimator of $\boldsymbol{\beta}_0$ is $\hat{\boldsymbol{\beta}}_n$ satisfying the following equation:

$$\sum_{i=1}^{n} f(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n) = \min_{\boldsymbol{\beta}\in\mathbf{R}^p} \sum_{i=1}^{n} f(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}). \tag{1.2}$$

The scope of M estimator is very wide containing the least squares estimator, maximum likelihood estimator *etc.* Since Huber[1] studied the M estimator of regression parameter in linear model, many authors have shown great interest in this field and obtained many useful results; see, for example, [2–5].

Throughout this paper, we use the following notations: Let

$$\boldsymbol{\alpha} = (a_1,\ a_2,\ \cdots,\ a_p)'$$

*Received date: April 27, 2010.
Foundation item: The NSF (11201001, 11171001, 11126176) of China, the NSF (1208085QA03) of Anhui Province, Provincial Natural Science Research Project (KJ2010A005) of Anhui Colleges, Doctoral Research Start-up Funds Projects of Anhui University and the Students' Innovative Training Project (2012003) of Anhui University.

be a p-dimensional vector, and

$$\|\boldsymbol{\alpha}\|^2 \doteq \sum_{i=1}^{p} a_i^2 = \boldsymbol{\alpha}'\boldsymbol{\alpha}, \qquad |\boldsymbol{\alpha}| \doteq \max_{1\le i\le p} |a_i|.$$

Denote

$$\boldsymbol{S}_n = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'.$$

Assume that $\boldsymbol{S}_n^{-1}$ exists, and

$$d_n \doteq \max_{1\le i\le n} \boldsymbol{x}_i' \boldsymbol{S}_n^{-1} \boldsymbol{x}_i.$$

Let f be a convex function, ψ_- and ψ_+ denote the left derivative and right derivative of function f, respectively. $a_n = O(b_n)$ denotes that there exists a positive constant C such that $\left|\dfrac{a_n}{b_n}\right| \le C$ for all sufficiently large n. C and C_i $(i \ge 1)$ are positive constants which may be different in various places.

As for the sufficient condition for the strong consistency of M estimator, Chen and Zhao[2] obtained the following result:

Theorem 1.1[2] *Let $e_1, e_2, \cdots, e_n, \cdots$ be a sequence of independent random variables with identical distribution, and f be a convex function satisfying the following two conditions*:

(1) *There exist constants $l_1 > 0$ and $l_2 > 0$ such that*

$$E(f(e_1+u) - f(e_1)) \ge l_1 u^2, \qquad |u| < l_2;$$

(2) *There exists a constant $\Delta > 0$ such that*

$$E|\psi_+(e_1 \pm \Delta)|^m \le h_m < \infty, \qquad m = 1, 2, \cdots$$

If there exists a δ with $0 < \delta \le 1$ such that $d_n = O(n^{-\delta})$, then, $\hat{\boldsymbol{\beta}}_n$ is the strong consistency estimator of $\boldsymbol{\beta}_0$.

Yang[3] improved the result of Theorem 1.1 and obtained the following Theorem 1.2 and Theorem 1.3.

Theorem 1.2[3] *Let $e_1, e_2, \cdots, e_n, \cdots$ be a sequnce of independent random variables with identical distribution, and f be a convex function satisfying the following two conditions*:

(1) *There exist constants $l_1 > 0$ and $l_2 > 0$ such that*

$$E(f(e_1+u) - f(e_1)) \ge l_1 u^2, \qquad |u| < l_2;$$

(2) *There exist constants $h_0 > 0$, $\Delta > 0$ and $0 < \delta \le 1$ such that*

$$E|\psi_+(e_1 \pm \Delta)|^{2/\delta} \le h_0 < \infty$$

and

$$d_n = O(n^{-\delta}).$$

Then, $\hat{\boldsymbol{\beta}}_n$ is the strong consistency estimator of $\boldsymbol{\beta}_0$.

Theorem 1.3[3] *Let $e_1, e_2, \cdots, e_n, \cdots$ be a sequence of independent random variables, and f be a convex function satisfying the following two conditions*:

(1) *There exist constants* $l_1 > 0$ *and* $l_2 > 0$ *such that*

$$E(f(e_i + u) - f(e_i)) \geq l_1 u^2, \qquad i = 1, 2, \cdots, n, \ |u| < l_2;$$

(2) *There exist constants* $h_0 > 0$, $\Delta > 0$, $0 < \delta \leq 1$ *and* $t > \dfrac{2}{\delta}$ *such that*

$$d_n = O(n^{-\delta}),$$

$$\sup_i E|\psi_+(e_i \pm \Delta)|^t \leq h_0 < \infty.$$

Then, $\hat{\boldsymbol{\beta}}_n$ *is the strong consistency estimator of* $\boldsymbol{\beta}_0$.

Unfortunately, $e_1, e_2, \cdots, e_n, \cdots$ are not independent in most cases. So it is valuable to extend the result for independent samples to the case of dependent samples. Some authors have shown great interests in this field and obtained some valuable results. Wu[4–6] studied the strong consistency of M estimator of regression parameter in linear model for ρ-mixing, ϕ-mixing and ψ-mixing samples (see [4]), the strong consistency of M estimator of regression parameter in linear model for $\tilde{\rho}$-mixing samples (see [5]), the strong consistency of M estimator for NA samples (see Theorem 3.3.1 of [6]), and so forth. In this paper, we investigate the strong consistency of M estimator of regression parameter in linear model for a class of $\tilde{\varphi}$-mixing samples and greatly reduce the moment condition of $|\psi_+(e_i \pm \Delta)|$.

Let $\{X_n, \ n \geq 1\}$ be a sequence of random variables defined on a fixed probability space $(\Omega, \ \mathcal{F}, \ P)$. Write

$$\mathcal{F}_S = \sigma(X_i, \ i \in S \subset \mathbf{N}).$$

Given σ-algebras $\mathcal{B}$, $\mathcal{R}$ in $\mathcal{F}$, let

$$\varphi(\mathcal{B}, \ \mathcal{R}) = \sup\{|P(B|A) - P(B)|; \ A \in \mathcal{B}, \ P(A) > 0, \ B \in \mathcal{R}\}.$$

Define the $\tilde{\varphi}$-mixing coefficients by

$$\tilde{\varphi}(k) = \sup\{\varphi(\mathcal{F}_S, \ \mathcal{F}_T) : \text{finite subsets } S, T \subset \mathbf{N} \text{ such that } \mathrm{dist}(S, \ T) \geq k\}, \qquad k \geq 0.$$

Obviously,

$$0 \leq \tilde{\varphi}(k+1) \leq \tilde{\varphi}(k) \leq 1, \qquad \tilde{\varphi}(0) = 1.$$

Definition 1.1 *A sequence of random variables* $\{X_n, \ n \geq 1\}$ *is said to be* $\tilde{\varphi}$*-mixing if there exists a* $k \in \mathbf{N}$ *such that* $\tilde{\varphi}(k) < 1$.

It is easily seen that independent sequence is the special case of $\tilde{\varphi}$-mixing sequence. The concept of $\tilde{\varphi}$-mixing was introduced by Wu and Lin[7]. Some authors have studied the concept and got some valuable results, for example, Wu and Lin[7], Wang and Hu[8], Wu[9], and so forth.

The following lemma plays an important role to prove the main result of this paper. The proof is similar to which of Lemma 5.1.1 in [6], so we omit the details.

Lemma 1.1 *Let* $\{X_n, \ n \geq 1\}$ *be a sequence of* $\tilde{\varphi}$*-mixing random variables with*

$$EX_n = 0, \quad E|X_n|^q < \infty, \qquad n = 1, 2, \cdots, \ q \geq 2.$$

Then there exists a positive constant C *depending only on* $\tilde{\varphi}(\cdot)$ *and* q *such that*

$$E\left|\sum_{i=1}^n X_i\right|^q \leq C\left[\sum_{i=1}^n E|X_i|^q + \left(\sum_{i=1}^n EX_i^2\right)^{q/2}\right], \qquad n \geq 1.$$

2 Main Result and Its Proof

Theorem 2.1 *Let $e_1, e_2, \cdots, e_n, \cdots$ be a sequence of $\tilde{\varphi}$-mixing random variables, and f be a convex function satisfying the following two conditions*:

(1) *There exist constants $l_1 > 0$ and $l_2 > 0$ such that*

$$E(f(e_i + u) - f(e_i)) \geq l_1 u^2, \qquad i = 1, 2, \cdots, n, \ |u| < l_2;$$

(2) *There exist constants $h_0 > 0$, $\Delta > 0$, $0 < \delta \leq 1$ and $t > 1 + \frac{1}{\delta}$ such that*

$$d_n = O(n^{-\delta}), \tag{2.1}$$

$$\sup_i E|\psi_+(e_i \pm \Delta)|^t \leq h_0 < \infty. \tag{2.2}$$

Then, $\hat{\boldsymbol{\beta}}_n$ is the strong consistency estimator of $\boldsymbol{\beta}_0$.

Proof. Without loss of generality, we assume that $\boldsymbol{\beta}_0 = \mathbf{0}$. Let

$$\boldsymbol{x}_{ni} = \boldsymbol{S}_n^{-\frac{1}{2}} \boldsymbol{x}_i, \quad \boldsymbol{\beta}_{n0} = \boldsymbol{S}_n^{\frac{1}{2}} \boldsymbol{\beta}_0, \qquad 1 \leq i \leq n.$$

Then the model (1.1) can be replaced by

$$y_i = \boldsymbol{x}'_{ni} \boldsymbol{\beta}_{n0} + e_i, \qquad i = 1, 2, \cdots, n, \ n \in \mathbf{N}, \tag{2.3}$$

and

$$\sum_{i=1}^n \boldsymbol{x}_{ni} \boldsymbol{x}'_{ni} = \boldsymbol{I}_p, \qquad \sum_{i=1}^n \|\boldsymbol{x}_{ni}\|^2 = p, \qquad d_n = \max_{1 \leq i \leq n} \|\boldsymbol{x}_{ni}\|^2. \tag{2.4}$$

Let $\boldsymbol{\beta}_n^*$ be the M estimator of $\boldsymbol{\beta}_{n0}$ under the model (2.3) based on the function f. Then

$$\hat{\boldsymbol{\beta}}_n = \boldsymbol{S}_n^{-\frac{1}{2}} \boldsymbol{\beta}_n^*.$$

Take

$$\varepsilon = \frac{l_1}{2}, \qquad 0 < \eta < \min\{1, \ \Delta, \ l_2\}, \qquad \tilde{b}_n = \Big[\frac{\eta}{\sqrt{d_n p}}\Big], \tag{2.5}$$

$$\boldsymbol{D}_n = \{\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_p)' : -\tilde{b}_n \leq \beta_i \leq \tilde{b}_n, \ 1 \leq i \leq p\}. \tag{2.6}$$

Let m be a positive integer such that $m > \frac{4}{\delta} - 3$. Each side of the hypercube $\boldsymbol{D}_n$ can be divided into $2\tilde{b}_n^{m+1}$ equal parts. Thus, the hypercube $\boldsymbol{D}_n$ can be divided into $(2\tilde{b}_n^{m+1})^p$ small hypercubes denoted as $\{\boldsymbol{B}_j : 1 \leq j \leq N_n\}$, $N_n = (2\tilde{b}_n^{m+1})^p$. The length of $\boldsymbol{B}_j$ is $\tilde{b}_n^{-m}$, and the center is denoted as $\boldsymbol{b}_j$. Denote

$$\phi_{ni}(\boldsymbol{\beta}) = f(e_i) - f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{\beta}), \qquad R_{ni}(\boldsymbol{\beta}) = \phi_{ni}(\boldsymbol{\beta}) - E\phi_{ni}(\boldsymbol{\beta}),$$

$$E_n = \Big\{ \sup_{\boldsymbol{\beta} \in \boldsymbol{D}_n} \Big| \sum_{i=1}^n R_{ni}(\boldsymbol{\beta}) \Big| \geq \varepsilon \tilde{b}_n^2 \Big\},$$

$$E_{n1} = \Big\{ \sup_{1 \leq j \leq N_n} \Big| \sum_{i=1}^n R_{ni}(\boldsymbol{b}_j) \Big| \geq \frac{\varepsilon \tilde{b}_n^2}{2} \Big\},$$

$$E_{n2} = \Big\{ \sup_{1 \leq j \leq N_n} \sup_{\boldsymbol{\beta} \in \boldsymbol{B}_j} \Big| \sum_{i=1}^n (R_{ni}(\boldsymbol{\beta}) - R_{ni}(\boldsymbol{b}_j)) \Big| \geq \frac{\varepsilon \tilde{b}_n^2}{2} \Big\}.$$

According to the proof of Theorem 3.1 in [2], we can see that

$$\{\|\hat{\boldsymbol{\beta}}_n\| \geq \varepsilon_0\} \subseteq E_{n1} \cup E_{n2}, \qquad \varepsilon_0 > 0.$$

Thus, in order to prove Theorem 2.1, we only need to show

$$\sum_{n=1}^{\infty} P(E_{n1}) < \infty, \tag{2.7}$$

$$\sum_{n=1}^{\infty} P(E_{n2}) < \infty. \tag{2.8}$$

Firstly, we prove (2.8). Denote

$$g(x) = \max\{|\psi_+(x + \mathit{\Delta})|,\ |\psi_+(x - \mathit{\Delta})|\}.$$

By (2.2) we have

$$\sup_i Eg^t(e_i) \le 2h_0 < \infty. \tag{2.9}$$

Since $t > 1 + \dfrac{1}{\delta} > 1$, it follows that

$$\sup_i Eg(e_i) < \infty.$$

By the convexity of f, we can see that

$$|f(a) - f(b)| \le |b - a| g(x), \qquad a, b \in (x - \mathit{\Delta},\ x + \mathit{\Delta}).$$

If $\boldsymbol{\beta} \in \boldsymbol{D}_n$, then by (2.5) and (2.6) we have

$$|\boldsymbol{x}'_{ni}\boldsymbol{b}_j| \le \sqrt{d_n p}\tilde{b}_n \le \eta \le \min\{1,\ \mathit{\Delta}\}, \tag{2.10}$$

$$|\boldsymbol{x}'_{ni}\boldsymbol{\beta}| \le \sqrt{d_n p}\tilde{b}_n \le \eta \le \min\{1,\ \mathit{\Delta}\}, \tag{2.11}$$

which imply that for $\boldsymbol{\beta} \in \boldsymbol{B}_j$,

$$\begin{aligned} & |f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{b}_j) - f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{\beta})| \\ & \le g(e_i)|\boldsymbol{x}'_{ni}(\boldsymbol{\beta} - \boldsymbol{b}_j)| \\ & \le g(e_i)\sqrt{d_n p}\|\boldsymbol{\beta} - \boldsymbol{b}_j\| \\ & \le g(e_i)\sqrt{d_n p}\tilde{b}_n^{-m}, \qquad 1 \le j \le N_n, \end{aligned} \tag{2.12}$$

and

$$\begin{aligned} & \sup_{1\le j\le N_n} \sup_{\boldsymbol{\beta}\in\boldsymbol{B}_j} \left| \sum_{i=1}^{n} (R_{ni}(\boldsymbol{\beta}) - R_{ni}(\boldsymbol{b}_j)) \right| \\ \le & \sup_{1\le j\le N_n} \sup_{\boldsymbol{\beta}\in\boldsymbol{B}_j} \sum_{i=1}^{n} |f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{b}_j) - f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{\beta})| \\ & + \sup_{1\le j\le N_n} \sup_{\boldsymbol{\beta}\in\boldsymbol{B}_j} \sum_{i=1}^{n} E|f(e_i - \boldsymbol{x}'_{ni}\boldsymbol{b}_j) - f(e_i - \boldsymbol{x}'_{ni}\beta)| \\ \le & \sqrt{d_n p}\tilde{b}_n^{-m} \sum_{i=1}^{n} g(e_i) + \sqrt{d_n p}\tilde{b}_n^{-m} \sum_{i=1}^{n} Eg(e_i) \\ \le & \sqrt{d_n p}\tilde{b}_n^{-m} \sum_{i=1}^{n} g(e_i) + n\sqrt{d_n p}\tilde{b}_n^{-m} \sup_i Eg(e_i). \end{aligned} \tag{2.13}$$

It follows from (2.1) and (2.4)–(2.6) that

$$d_n \ll n^{-\delta}, \qquad \tilde{b}_n \ll n^{\frac{\delta}{2}}, \qquad \sum_{i=1}^{n} |\boldsymbol{x}'_{ni}\boldsymbol{b}_j|^2 \ll n^{\delta}. \tag{2.14}$$

By (2.9) and $t > 1 + \dfrac{1}{\delta} > 1$ we can get

$$\sup_i Eg(e_i) < \infty.$$

Note that $m > \dfrac{4}{\delta} - 3$, which implies that

$$1 - \frac{\delta(m+3)}{2} < -1$$

and

$$n\tilde{b}_n^{-m-2}\sqrt{d_n p}\sup_i Eg(e_i) \ll n^{1-\frac{\delta(m+3)}{2}} \to 0, \qquad n \to \infty.$$

Therefore, for all n large enough,

$$n\tilde{b}_n^{-m}\sqrt{d_n p}\sup_i Eg(e_i) < \frac{\varepsilon \tilde{b}_n^2}{4}. \tag{2.15}$$

By Markov's inequality and (2.13)–(2.15), we have for all n large enough

$$\begin{aligned}
P(E_{n2}) &= P\bigg(\sup_{1\le j\le N_n}\sup_{\beta\in \boldsymbol{B}_j}\bigg|\sum_{i=1}^n (R_{ni}(\boldsymbol{\beta}) - R_{ni}(\boldsymbol{b}_j))\bigg| \ge \frac{\varepsilon\tilde{b}_n^2}{2}\bigg)\\
&\le P\bigg(\tilde{b}_n^{-m}\sqrt{d_n p}\sum_{i=1}^n g(e_i) + \tilde{b}_n^{-m}\sqrt{d_n p}\sum_{i=1}^n Eg(e_i) \ge \frac{\varepsilon\tilde{b}_n^2}{2}\bigg)\\
&\le P\bigg(\tilde{b}_n^{-m}\sqrt{d_n p}\sum_{i=1}^n g(e_i) \ge \frac{\varepsilon\tilde{b}_n^2}{4}\bigg)\\
&\le C_1\tilde{b}_n^{-(2+m)}\sqrt{d_n}\sum_{i=1}^n Eg(e_i)\\
&\le C_2 n\tilde{b}_n^{-(2+m)}\sqrt{d_n}\\
&\le C_3 n^{1-\frac{(m+3)\delta}{2}},
\end{aligned}$$

which together with $m > \dfrac{4}{\delta} - 3$ implies that

$$\sum_{n=1}^\infty P(E_{n2}) < \infty.$$

This completes the proof of (2.8).

Now we prove (2.7). We discuss it for two cases.

1) If $0 < \delta < 1$, we denote

$$r_n = \varepsilon\tilde{b}_n^{2\tau}\ln n,$$

where

$$\tau = \frac{1}{\delta(t-1)} < 1.$$

Let

$$\begin{aligned}
\xi_{ni}^{(1)}(\boldsymbol{b}_j) &= \phi_{ni}(\boldsymbol{b}_j)I(|\phi_{ni}(\boldsymbol{b}_j)| \le r_n),\\
\xi_{ni}^{(2)}(\boldsymbol{b}_j) &= \phi_{ni}(\boldsymbol{b}_j)I(|\phi_{ni}(\boldsymbol{b}_j)| > r_n),\\
\eta_{ni}^{(k)}(\boldsymbol{b}_j) &= \xi_{ni}^{(k)}(\boldsymbol{b}_j) - E\xi_{ni}^{(k)}(\boldsymbol{b}_j), \qquad k = 1, 2.
\end{aligned}$$

It is easy to see that

$$\sum_{i=1}^{n} R_{ni}(\boldsymbol{b}_j) = \sum_{i=1}^{n} \eta_{ni}^{(1)}(\boldsymbol{b}_j) + \sum_{i=1}^{n} \eta_{ni}^{(2)}(\boldsymbol{b}_j),$$

which yields that

$$\begin{aligned} E_{n1} &\subseteq \left\{ \sup_{1\le j\le N_n} \left| \sum_{i=1}^{n} \eta_{ni}^{(1)}(\boldsymbol{b}_j) \right| \ge \frac{\varepsilon \tilde{b}_n^2}{4} \right\} \bigcup \left\{ \sup_{1\le j\le N_n} \left| \sum_{i=1}^{n} \eta_{ni}^{(2)}(\boldsymbol{b}_j) \right| \ge \frac{\varepsilon \tilde{b}_n^2}{4} \right\} \\ &\doteq E_{n11} + E_{n12}. \end{aligned}$$

Take

$$q > \max\left\{ \frac{\dfrac{\delta p(m+1)}{2} - 2\delta\tau + \delta + 1}{\delta(1-\tau)},\ p(m+1) + \frac{2}{\delta},\ 2 \right\}.$$

By Lemma 1.1 we have

$$\begin{aligned} E\left| \sum_{i=1}^{n} \eta_{ni}^{(1)}(\boldsymbol{b}_j) \right|^q &\le C\left[\sum_{i=1}^{n} E|\eta_{ni}^{(1)}(\boldsymbol{b}_j)|^q + \left(\sum_{i=1}^{n} E|\eta_{ni}^{(1)}(\boldsymbol{b}_j)|^2 \right)^{q/2} \right] \\ &\le C\left[\sum_{i=1}^{n} E|\phi_{ni}(\boldsymbol{b}_j)|^q I(|\phi_{ni}(\boldsymbol{b}_j)| \le r_n) + \left(\sum_{i=1}^{n} E|\eta_{ni}^{(1)}(\boldsymbol{b}_j)|^2 \right)^{q/2} \right] \\ &\le C\left[\sum_{i=1}^{n} E|\phi_{ni}(\boldsymbol{b}_j)|^q r_n^{q-2} + \left(\sum_{i=1}^{n} E|\phi_{ni}(\boldsymbol{b}_j)|^2 \right)^{q/2} \right] \\ &\le C\left[r_n^{q-2} \sum_{i=1}^{n} |\boldsymbol{x}_{ni}' \boldsymbol{b}_j|^2 E g^2(e_i) + \left(\sum_{i=1}^{n} |\boldsymbol{x}_{ni}' \boldsymbol{b}_j|^2 E g^2(e_i) \right)^{q/2} \right] \\ &\le C\left[n^{\tau\delta(q-2)+\delta} (\ln n)^{q-2} + n^{\delta q/2} \right]. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{n=1}^{\infty} P(E_{n11}) &\le C \sum_{n=1}^{\infty} \tilde{b}_n^{-2q} \sum_{j=1}^{N_n} E\left| \sum_{i=1}^{n} \eta_{ni}^{(1)}(\boldsymbol{b}_j) \right|^q \\ &\le C \sum_{n=1}^{\infty} \left[n^{\frac{\delta(m+1)p}{2} - \delta q + \tau\delta(q-2) + \delta} (\ln n)^{\tau(q-2)} + n^{\frac{\delta(m+1)p}{2} - \frac{\delta q}{2}} \right] \\ &< \infty. \end{aligned}$$

Finally, we prove that

$$\sum_{n=1}^{\infty} P(E_{n12}) < \infty.$$

It follows from (2.14) that

$$\begin{aligned} \left| \sum_{i=1}^{n} E\xi_{ni}^{(2)}(\boldsymbol{b}_j) \right| &\le \sum_{i=1}^{n} E|\xi_{ni}^{(2)}(\boldsymbol{b}_j)| \\ &\le \sum_{i=1}^{n} E|\phi_{ni}(\boldsymbol{b}_j)| I(|\phi_{ni}(\boldsymbol{b}_j)| > r_n) \\ &= \sum_{i=1}^{n} E|\phi_{ni}(\boldsymbol{b}_j)|^t |\phi_{ni}(\boldsymbol{b}_j)|^{1-t} I(|\phi_{ni}(\boldsymbol{b}_j)| > r_n) \end{aligned}$$

$$\begin{aligned}
&\le \sup_{1\le j\le N_n}\sum_{i=1}^{n}|\boldsymbol{x}'_{ni}\boldsymbol{b}_j|^t Eg^t(e_i)r_n^{1-t}\\
&\le \tilde{b}_n^2\tilde{b}_n^{-2\tau(t-1)},
\end{aligned}$$

which implies that for all n large enough

$$\sup_{1\le j\le N_n}\left|\sum_{i=1}^{n}E\xi_{ni}^{(2)}(\boldsymbol{b}_j)\right|\le\frac{\varepsilon\tilde{b}_n^2}{8}.$$

Since

$$|\phi_{ni}(\boldsymbol{b}_j)|\le g(e_i)|\boldsymbol{x}'_{ni}\boldsymbol{b}_j|,\qquad |\boldsymbol{x}'_{ni}\boldsymbol{b}_j|\le 1,$$

it follows that

$$\begin{aligned}
\sup_{1\le j\le N_n}\left|\sum_{i=1}^{n}\xi_{ni}^{(2)}(\boldsymbol{b}_j)\right| &\le \sup_{1\le j\le N_n}\sum_{i=1}^{n}|\xi_{ni}^{(2)}(\boldsymbol{b}_j)|\\
&\le \sup_{1\le j\le N_n}\sum_{i=1}^{n}|\phi_{ni}(\boldsymbol{b}_j)|I(|\phi_{ni}(\boldsymbol{b}_j)|>r_n)\\
&\le \sup_{1\le j\le N_n}\sum_{i=1}^{n}|\phi_{ni}(\boldsymbol{b}_j)|^t|\phi_{ni}(\boldsymbol{b}_j)|^{1-t}I(|\phi_{ni}(\boldsymbol{b}_j)|>r_n)\\
&\le \sup_{1\le j\le N_n}\sum_{i=1}^{n}|\boldsymbol{x}'_{ni}\boldsymbol{b}_j|^t g^t(e_i)r_n^{1-t}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{n=1}^{\infty}P(E_{n12}) &\le \sum_{n=1}^{\infty}P\left(\sup_{1\le j\le N_n}\sum_{i=1}^{n}|\boldsymbol{x}'_{ni}\boldsymbol{b}_j|^2 g^t(e_i)r_n^{1-t}\ge\frac{\varepsilon\tilde{b}_n^2}{8}\right)\\
&\le C\sum_{n=1}^{\infty}\tilde{b}_n^{-2}\sup_{1\le j\le N_n}\sum_{i=1}^{n}|\boldsymbol{x}'_{ni}\boldsymbol{b}_j|^2 Eg^t(e_i)r_n^{1-t}\\
&\le C\sum_{n=1}^{\infty}n^{-1}(\ln n)^{1-t}\\
&<\infty.
\end{aligned}$$

2) If $\delta=1$, we can also get

$$\sum_{n=1}^{\infty}P(E_{n12})<\infty$$

by using the similar method of the proof of the Theorem in [5].

The proof is completed.

Remark 2.1 We have pointed out that $\tilde{\varphi}$-mixing sequence contains independent sequence as a special case, and thus Theorem 2.1 generalizes the result for independent sequence. On the other hand, the condition (2.2) reduces the moment condition of $|\psi_+(e_i\pm\varDelta)|$ (for $0<\delta\le 1$, $\dfrac{2}{\delta}\ge 1+\dfrac{1}{\delta}$), which greatly improves and extends the results of Chen and Zhao[2] and Yang[3].

References

[1] Huber P J. Robust regression. *Ann. Statist.*, 1973, **1**: 799–821.

[2] Chen X R, Zhao L C. M Method in Linear Model (in Chinese). Shanghai: Science and Technology Press of Shanghai, 1996.

[3] Yang S C. Strong consistence of M estimator of regression parametric in linear model (in Chinese). *Acta Math. Sinica*, 2002, **45**: 21–28.

[4] Wu Q Y. Strong consistency of M estimator in linear model for ρ-mixing, ϕ-mixing, ψ-mixing samples (in Chinese). *Math. Appl.*, 2004, **17**: 393–397.

[5] Wu Q Y. Strong consistency of M estimator in linear model for $\tilde{\rho}$-mixing samples (in Chinese). *Acta Math. Sinica*, 2005, **25**: 41–46.

[6] Wu Q Y. Probability Limit Theory for Mixing Sequences (in Chinese). Beijing: Science Press of China, 2006.

[7] Wu Q Y, Lin L. Convergence properties of $\tilde{\varphi}$-mixing random sequences (in Chinese). *J. Engrg. Math.*, 2004, **21**: 75–80.

[8] Wang X J, Hu S H. Large deviations for the partial sums of a class of random variable sequences (in Chinese). *Math. Appl.*, 2007, **20**: 541–547.

[9] Wu Q Y. Almost sure convergence for $\tilde{\varphi}$-mixing random variable sequences. *Math. Appl.*, 2008, **21**: 629–634.