Journal of Mathematical Inequalities
Volume 11, Number 4 (2017), 965–982 doi:10.7153/jmi-2017-11-73

ALMOST SURE CONVERGENCE FOR END SEQUENCES AND ITS APPLICATION TO M ESTIMATOR IN LINEAR MODELS

XIN DENG, SHIJIE WANG, RUI WANG, XIUJUAN XIE AND XUEJUN WANG*

(*Communicated by J. Pečarić*)

Abstract. In the paper, an almost sure convergence result for weighted sums of extended negatively dependent random variables is obtained. By using the almost sure convergence result, we further study the strong consistency of M estimator of the regression parameter in linear models based on extended negatively dependent random errors under some mild conditions.

1. Introduction

1.1. Brief review

Consider the linear model

$$Y_i = x_i'\beta + e_i, \quad i = 1, \ldots, n,\ n \geqslant 1, \tag{1.1}$$

where $x_1, x_2, \cdots, x_n$ are $p \times 1$ known design vectors, $e_1, e_2, \cdots, e_n$ are random errors, and β is a $p \times 1$ unknown parameter vector. Suppose that ρ is some suitably chosen function on $\mathbf{R}^1$. The M estimator of β is defined by $\hat{\beta}_n$ as follows:

$$\sum_{i=1}^{n} \rho(Y_i - x_i'\hat{\beta}_n) = \min_{\beta \in \mathbf{R}^p} \sum_{i=1}^{n} \rho(Y_i - x_i'\beta). \tag{1.2}$$

M estimators are very generally-used and important examples include maximum likelihood estimators (MLE) with $\rho(x) = -\ln f(x)$, where f is the common density function of e_i, Huber's estimators with $\rho(x) = x^2 I(|x| \leqslant c)/2 + (c|x| - c^2/2) I(|x| > c)$, $c > 0$, the regression quantile estimators with $\rho(x) = \alpha x^+ + (1-\alpha)(-x)^+$, $0 < \alpha < 1$, where $x^+ = \max(x, 0)$, and the $\mathscr{L}^q$ regression estimators with $\rho(x) = |x|^q$, $1 \leqslant q \leqslant 2$.

After Huber (1973) studied M estimators, many statisticians have been interested in studying this topic. A series of useful results were established. For the details on

Mathematics subject classification (2010): 62F12, 60F15.
Keywords and phrases: Almost sure convergence, strong consistency, M estimator, extended negatively dependent random error.
Supported by the National Natural Science Foundation of China (11501004, 11501005), the Natural Science Foundation of Anhui Province (1508085J06), the Key Projects for Academic Talent of Anhui Province (gxbjZD2016005), the Provincial Natural Science Research Project of Anhui Colleges (KJ2015A018), the Quality Engineering Project of Anhui Province (2016jyxm0047), the Quality Improvement Projects for Undergraduate Education of Anhui University (ZLTS2015035, ZLTS2015138).
* The corresponding author.

M estimators, one can refer to Chen and Zhao (1995,1996), He and Shao (1996), Yang (2002), Collins and Szatmari (2004), Georgios (2005), Wu (2006a), Djalil and Didier (2007), Seija et al. (2007), and so forth. Huber (1973) commented that the restriction about the assumption of independence is serious. Recently, Wu and Jiang (2011) relaxed the independent assumption in the classical M estimators theory and established the strong consistent estimator of β in the linear model (1.1) with negatively dependent random errors. Inspired by the literatures above, we will study the M estimator in the linear model (1.1) with extended negatively dependent random errors, which include independent random errors and negatively dependent random errors as special cases.

1.2. Concept of extended negatively dependence structure

In this section, we will present the extended negatively dependence structure introduced in Liu (2009).

DEFINITION 1.1. We call random variables $\{X_n, n \geqslant 1\}$ extended negatively dependent (END, in short) if there exists a constant $M > 0$ such that both

$$P(X_1 > x_1, X_2 > x_2, \cdots, X_n > x_n) \leqslant M \prod_{i=1}^{n} P(X_i > x_i) \tag{1.3}$$

and

$$P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n) \leqslant M \prod_{i=1}^{n} P(X_i \leqslant x_i) \tag{1.4}$$

hold for each $n \geqslant 1$ and all real numbers $x_1, x_2, \cdots, x_n$.

The concept of END sequence was introduced by Liu (2009). In the case $M = 1$, the notion of END random variables reduces to the well-known notion of so-called negatively dependent (ND, in short) random variables, which was introduced by Lehmann (1966) (cf. also Joag-Dev and Proschan, 1983). For more details about ND random variables, one can refer to Volodin (2002), Zarei and Jabbari (2011), Sung (2012), Shen (2011a, 2013), Wang and Si (2015) among others. As it mentioned in Liu (2009), the END structure is substantially more general than ND structure, it can reflect not only a negative dependence structure but also a positive one.

Some applications for END sequence have been found. See, for example, Liu (2009) obtained the precise large deviations for dependent random variables with heavy tails. Liu (2010) studied the sufficient and necessary conditions of moderate deviations for dependent random variables with heavy tails. Chen et al. (2010) established the strong law of large numbers for extend negatively dependent random variables and showed applications to risk theory and renewal theory. Chen et al. (2011) obtained the precise large deviations of random sums in presence of negative dependence and consistent variation. Shen (2011b) presented some probability inequalities for END sequences and gave some applications. Wu and Guan (2012), Wang et al. (2013, 2014), Qiu et al. (2013), Wu et al. (2014) and Hu et al. (2015) obtained complete convergence results for sums or weighted sums of END random variables. Wang and Wang (2013) established the precise large deviations for random sums of END real-valued random variables with consistent variation. Shen (2014) studied the asymptotic approximation of inverse moments for nonnegative END random variables. Wang et al. (2015) and Yang et al. (2016) investigated the complete consistency for the estimator of regression models based on END errors, and so forth. Since END random variables are much weaker than independent random variables, negatively associated random variables and ND random variables, studying the limit behavior of END sequence is of interest.

The main purpose of the paper is to study the almost sure convergence for weighted sums of END random variables. By using the almost sure convergence, we will further study the strong consistency of M estimator in the linear model (1.1) with END random errors, which generalizes the corresponding one with ND random errors.

The following concept of stochastic domination will be used in the paper.

DEFINITION 1.2. A sequence of random variables $\{X_n, n \geqslant 1\}$ is said to be stochastically dominated by a random variable X if there exists a positive constant C such that

$$P(|X_n| > x) \leqslant CP(|X| > x)$$

for all $x \geqslant 0$ and $n \geqslant 1$.

This work is organized as follows: some important lemmas will be presented in Section 2. The almost sure convergence for weighted sums of END random variables will be provided in Section 3 and the strong consistency of M estimator in a linear model with END errors will be studied in Section 4, respectively.

For the sake of convenience, throughout this article, $\lceil x \rceil$ denotes the maximum integer not exceeding x, $x^+ = \max\{x, 0\} \geqslant 0$, $x^- = \max\{-x, 0\} \geqslant 0$. And we always assume that ρ is a nonmonotonic convex function on $\mathbf{R}^1$ which ensures the existence of a M estimator for the linear model (1.1) (see Chen and Zhao, 1996). ψ_- and ψ_+ stand for left and right derivatives of ρ, respectively. Choose increasing function ψ such that $\psi_-(u) \leqslant \psi(u) \leqslant \psi_+(u)$ for all $u \in \mathbf{R}^1$. The function $\psi(\cdot)$ plays an important role in the study of the strong consistency of $\hat{\beta}_n$. Let a denote $p \times 1$ column vector and a' denote its transpose. Write

$$\| a \|^2 \triangleq \sum_{i=1}^{p} a_i^2 = a'a, |a| \triangleq \max_{1 \leqslant i \leqslant p} |a_i|.$$

Let $S_n \triangleq x_1 x_1' + x_2 x_2' + \cdots + x_n x_n'$, suppose that S_n^{-1} exists, and hence, S_n is a positive definite matrix. Let $d_n = \max_{1 \leqslant i \leqslant n} x_i' S_n^{-1} x_i$.

2. Preliminary lemmas

To prove the main results of the paper, we need the following useful lemmas.

LEMMA 2.1. (cf. Liu, 2010) *Let random variables $X_1, X_2, \cdots, X_n$ be END.*

(i) If $f_1, f_2, \cdots, f_n$ are all nondecreasing (or nonincreasing) functions, then the random variables $f_1(X_1), f_2(X_2), \cdots, f_n(X_n)$ are END.

(ii) For each $n \geqslant 1$, there exists a constant $M > 0$ such that

$$E\left(\prod_{j=1}^{n} X_j^+\right) \leqslant M \prod_{j=1}^{n} E X_j^+. \tag{2.1}$$

LEMMA 2.2. (cf. Shen, 2011b) *Let $\{X_n, n \geqslant 1\}$ be a sequence of END random variables and $\{t_n, n \geqslant 1\}$ be a sequence of nonnegative numbers (or nonpositive numbers), then for each $n \geqslant 1$, there exists a constant $M > 0$ such that*

$$E\left(\prod_{j=1}^{n} e^{t_i X_i}\right) \leqslant M \prod_{j=1}^{n} E e^{t_i X_i}. \tag{2.2}$$

LEMMA 2.3. *Let $t > 0$ and $\{X_n, n \geqslant 1\}$ be a sequence of END random variables with $EX_i = 0$ and $|X_i| \leqslant b_i$ a.s. $(i = 1, 2, \ldots)$, where b_i is a positive constant. Denote $B_n = \sum_{i=1}^n EX_i^2$. Assume that $t \cdot \max_{1 \leqslant i \leqslant n} b_i \leqslant 1$. Then there exists a constant $M > 0$ such that*

$$P\left(\left|\sum_{i=1}^{n} X_i\right| > u\right) \leqslant 2M \exp\{-tu + t^2 B_n\}, \quad \text{for all } u > 0.$$

Proof. It follows by $|tX_i| \leqslant 1$ $a.s.$ that

$$e^{tX_i} = \sum_{k=0}^{\infty} \frac{(tX_i)^k}{k!} \leqslant 1 + tX_i + (tX_i)^2 \sum_{k=2}^{\infty} \frac{1}{k!} \leqslant 1 + tX_i + t^2 X_i^2 \ \ a.s.,$$

which together with $EX_i = 0$ yields that

$$E(e^{tX_i}) \leqslant 1 + tEX_i + t^2 EX_i^2 \leqslant e^{t^2 EX_i^2}.$$

By the inequality above, Markov's inequality and Lemma 2.2, we can see that for any $u > 0, t > 0$, there exists a positive constant M such that

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i > u\right) &= P(e^{t\sum_{i=1}^n X_i} > e^{tu}) \leqslant e^{-tu} E e^{t\sum_{i=1}^n X_i} \\ &\leqslant M e^{-tu} \prod_{i=1}^{n} E e^{tX_i} \leqslant M e^{-tu} \prod_{i=1}^{n} e^{t^2 EX_i^2} \\ &= M \exp\{-tu + t^2 B_n\}. \end{aligned}$$

Let $-X_i$ take the place of X_i in the above inequality, we have

$$P\left(\sum_{i=1}^{n} (-X_i) > u\right) = P\left(\sum_{i=1}^{n} X_i < -u\right) \leqslant M \exp\{-tu + t^2 B_n\}.$$

Therefore,

$$\begin{aligned} P\left(\left|\sum_{i=1}^{n} X_i\right| > u\right) &= P\left(\sum_{i=1}^{n} X_i > u\right) + P\left(\sum_{i=1}^{n} X_i < -u\right) \\ &\leqslant 2M \exp\{-tu + t^2 B_n\}. \end{aligned}$$

The proof is completed. □

The last one is a basic property for stochastic domination. For the proof, one can refer to Wu (2006b), or Shen et al. (2015).

LEMMA 2.4. *Let $\{X_n, n \geqslant 1\}$ be a sequence of random variables which is stochastically dominated by a random variable X. For any $\alpha > 0$ and $b > 0$, the following two statements hold:*

$$E|X_n|^\alpha I(|X_n| \leqslant b) \leqslant C_1 \left[E|X|^\alpha I(|X| \leqslant b) + b^\alpha P(|X| > b)\right], \tag{2.3}$$

$$E|X_n|^\alpha I(|X_n| > b) \leqslant C_2 E|X|^\alpha I(|X| > b), \tag{2.4}$$

where C_1 and C_2 are positive constants.

3. Almost sure convergence for weighted sums of END random variables

In this section, we will study the almost sure convergence for weighted sums of END random variables, which can be applied to establish the strong consistency of M estimator in a linear model with END errors.

Our main result is as follows.

THEOREM 3.1. *Let $\{X_n, n \geqslant 1\}$ be a sequence of END identically distributed random variables with mean zero, and for some $0 < \beta \leqslant 1$,*

$$E|X_1|^{1/\beta} < \infty. \tag{3.1}$$

Let $\{a_{nk}, 1 \leqslant k \leqslant n, n \geqslant 1\}$ be an array of constants such that

$$|a_{nk}| \leqslant cn^{-\beta} \text{ for } 1 \leqslant k \leqslant n,\ n \geqslant 1, \tag{3.2}$$

where c is a positive constant. Assume further that there exists a constant $\alpha > 0$ such that

$$\sum_{k=1}^{n} |a_{nk}|^p \leqslant cn^{-\alpha} \text{ for all } n, \tag{3.3}$$

where $p = \min(\frac{1}{\beta}, 2)$. Then,

$$T_n \triangleq \sum_{k=1}^{n} a_{nk} X_k \to 0 \ \ a.s.\ n \to \infty. \tag{3.4}$$

Proof. It is easily seen that $T_n = \sum_{k=1}^n a_{nk}^+ X_k - \sum_{k=1}^n a_{nk}^- X_k$, so, without loss of generality, we assume that $a_{nk} > 0$ for all $1 \leqslant k \leqslant n, n \geqslant 1$.

Let $\varepsilon > 0$ be given, $N = \lceil \frac{2}{\alpha} + 1 \rceil$, $A_{nk} = a_{nk}^{-1} n^{-\alpha\beta/2}$, $B_k = \frac{\varepsilon k^\beta}{Nc}$. For $k \leqslant n$, let

$$\begin{aligned}
X_{nk}^{(1)} &= X_k I_{(X_k \leqslant A_{nk})} + A_{nk} I_{(X_k > A_{nk})}, \\
X_{nk}^{(2)} &= X_k I_{(X_k > B_k)}, \\
X_{nk}^{(3)} &= X_k - X_{nk}^{(1)} - X_{nk}^{(2)} = (X_k - A_{nk}) I_{(A_{nk} < X_k \leqslant B_k)} - A_{nk} I_{(X_k > B_k)},
\end{aligned}$$

and

$$T_n^{(i)} = \sum_{k=1}^{n} a_{nk} X_{nk}^{(i)}, \ \ i = 1, 2, 3.$$

Then,

$$T_n = \sum_{i=1}^{3} T_n^{(i)}. \tag{3.5}$$

Let $Y_{nk} = n^{\alpha\beta/2} a_{nk} X_{nk}^{(1)}$. For fixed $n \geqslant 1$, it is easily seen that $\{Y_{nk}, 1 \leqslant k \leqslant n\}$ are also END random variables by Lemma 2.1. And $Y_{nk} \leqslant n^{\alpha\beta/2} a_{nk} A_{nk} = 1, EY_{nk} = n^{\alpha\beta/2} a_{nk} EX_{nk}^{(1)} \leqslant n^{\alpha\beta/2} a_{nk} EX_k = 0$. It is easy to verify

$$e^y \leqslant 1 + y + |y|^p \ \text{ for } y \leqslant 1,\ 1 \leqslant p \leqslant 2,$$

which implies that

$$e^{Y_{nk}} \leqslant 1 + Y_{nk} + |Y_{nk}|^p.$$

It follows by the inequality above and the fact $EY_{nk} \leqslant 0$ that

$$Ee^{Y_{nk}} \leqslant 1 + EY_{nk} + E|Y_{nk}|^p \leqslant 1 + E|Y_{nk}|^p \leqslant e^{E|Y_{nk}|^p}. \tag{3.6}$$

Noting that $|X_{nk}^{(1)}| \leqslant |X_k|$, $p \leqslant 1/\beta$, $E|X_1|^p < \infty$, we have by Lemma 2.2, (3.3) and (3.6) that

$$\begin{aligned} E\exp(n^{\alpha\beta/2} T_n^{(1)}) &= E\exp\left(\sum_{k=1}^{n} Y_{nk}\right) \leqslant M \prod_{k=1}^{n} E\exp(Y_{nk}) \leqslant M \prod_{k=1}^{n} \exp(E|Y_{nk}|^p) \\ &= M \prod_{k=1}^{n} \exp\left(n^{\alpha\beta p/2} a_{nk}^p E|X_{nk}^{(1)}|^p\right) \leqslant M\exp(n^{\alpha/2} E|X_1|^p \sum_{k=1}^{n} a_{nk}^p) \\ &\leqslant M\exp(cE|X_1|^p n^{-\alpha/2}) \leqslant c_1, \end{aligned}$$

where c_1 is a positive constant. It follows by Markov's inequality and the inequality above that

$$\begin{aligned} \sum_{n=1}^{\infty} P(T_n^{(1)} \geqslant \varepsilon) &\leqslant \sum_{n=1}^{\infty} \exp(-\varepsilon n^{\alpha\beta/2}) E\exp(n^{\alpha\beta/2} T_n^{(1)}) \\ &\leqslant c_1 \sum_{n=1}^{\infty} \exp(-\varepsilon n^{\alpha\beta/2}) < \infty. \end{aligned}$$

By the Borel-Cantelli lemma,

$$P(T_n^{(1)} \geqslant \varepsilon, i.o.) = 0.$$

Thus,

$$P\left(\limsup_{n\to\infty} T_n^{(1)} \geqslant \varepsilon\right) \leqslant P(T_n^{(1)} \geqslant \varepsilon, i.o.) = 0.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$P(\limsup_{n\to\infty} T_n^{(1)} > 0) = 0,$$

i.e.,

$$\limsup_{n\to\infty} T_n^{(1)} \leqslant 0 \ \ a.s. \tag{3.7}$$

It follows by (3.1) that

$$\begin{aligned}
\sum_{k=1}^{\infty} P(|X_k| > B_k) &= \sum_{k=1}^{\infty} P(|X_1| > B_k) \\
&= \sum_{k=1}^{\infty} \sum_{i=k}^{\infty} P\left(\frac{\varepsilon}{Nc} i^{\beta} < |X_1| \leqslant \frac{\varepsilon}{Nc}(i+1)^{\beta}\right) \\
&= \sum_{i=1}^{\infty} \sum_{k=1}^{i} EI\left(\frac{\varepsilon}{Nc} i^{\beta} < |X_1| \leqslant \frac{\varepsilon}{Nc}(i+1)^{\beta}\right) \\
&= \sum_{i=1}^{\infty} E\left(iI\left(\frac{\varepsilon}{Nc} i^{\beta} < |X_1| \leqslant \frac{\varepsilon}{Nc}(i+1)^{\beta}\right)\right) \\
&\leqslant \left(\frac{Nc}{\varepsilon}\right)^{1/\beta} \sum_{i=1}^{\infty} E|X_1|^{1/\beta} I\left(\frac{\varepsilon}{Nc} i^{\beta} < |X_1| \leqslant \frac{\varepsilon}{Nc}(i+1)^{\beta}\right) \\
&\leqslant \left(\frac{Nc}{\varepsilon}\right)^{1/\beta} E|X_1|^{1/\beta} < \infty.
\end{aligned}$$

By the Borel-Cantelli lemma,

$$P(|X_k| > B_k, i.o.) = 0.$$

Therefore,

$$\sum_{k=1}^{\infty} X_k^2 I(|X_k| > B_k) < \infty \ \ a.s..$$

Thus, by the Schwarz's inequality and (3.3),

$$\begin{aligned}
0 \leqslant T_n^{(2)} &= \sum_{k=1}^{n} a_{nk} X_{nk}^{(2)} \leqslant \left(\sum_{k=1}^{n} a_{nk}^2\right)^{1/2} \left(\sum_{k=1}^{n} X_k^2 I_{(X_k > B_k)}\right)^{1/2} \\
&\leqslant \left(c^{2-p} \sum_{k=1}^{n} a_{nk}^p\right)^{1/2} \left(\sum_{k=1}^{n} X_k^2 I_{(X_k > B_k)}\right)^{1/2} \\
&\leqslant \sqrt{c^{3-p}} n^{-\alpha/2} \left(\sum_{k=1}^{n} X_k^2 I_{(X_k > B_k)}\right)^{1/2} \\
&\to 0 \ \ a.s., \ n \to \infty.
\end{aligned} \tag{3.8}$$

Now we prove $\limsup_{n\to\infty} T_n^{(3)} \leqslant 0$ $a.s.$. It follows by (3.2) that

$$a_{nk}(X_k - A_{nk})I_{(A_{nk}<X_k\leqslant B_k)} \leqslant a_{nk}B_k \leqslant \varepsilon/N.$$

Therefore,

$$\begin{aligned}\{T_n^{(3)} \geqslant \varepsilon\} &= \left\{\sum_{k=1}^n a_{nk}(X_k - A_{nk})I_{(A_{nk}<X_k\leqslant B_k)} - \sum_{k=1}^n a_{nk}A_{nk}I_{(X_k>B_k)} \geqslant \varepsilon\right\} \\ &\subset \left\{\sum_{k=1}^n a_{nk}(X_k - A_{nk})I_{(A_{nk}<X_k\leqslant B_k)} \geqslant \varepsilon\right\} \\ &\subset \{\text{There exist at least } N \text{ subscripts } k \text{ such that } X_k > A_{nk}\}.\end{aligned}$$

Noting that $\beta p \leqslant 1$ and $E|X_1|^p < \infty$, we have by (3.3) that

$$\begin{aligned}P(T_n^{(3)} \geqslant \varepsilon) &\leqslant \sum_{1\leqslant k_1<k_2<\cdots<k_N\leqslant n} P(X_{k_1} > A_{nk_1}, X_{k_2} > A_{nk_2}, \cdots, X_{k_N} > A_{nk_N}) \\ &\leqslant M \sum_{1\leqslant k_1<k_2<\cdots<k_N\leqslant n} \prod_{i=1}^N P(X_{k_i} > A_{nk_i}) \\ &\leqslant M \sum_{1\leqslant k_1<k_2<\cdots<k_N\leqslant n} \prod_{i=1}^N P(|X_{k_i}| > A_{nk_i}) \\ &\leqslant M\left(\sum_{k=1}^n P(|X_k| > A_{nk})\right)^N \leqslant M\left(\sum_{k=1}^n A_{nk}^{-p}E|X_1|^p\right)^N \\ &= M\left(\sum_{k=1}^n a_{nk}^p n^{\alpha\beta p/2}E|X_1|^p\right)^N \leqslant M(cE|X_1|^p)^N n^{-\alpha N/2}.\end{aligned}$$

Choosing $N = \lceil \frac{2}{\alpha} + 1 \rceil$ such that $\alpha N/2 > 1$, we have $\sum_{n=1}^{\infty} P(T_n^{(3)} \geqslant \varepsilon) < \infty$ for any given $\varepsilon > 0$. Similar to the proof of (3.7), we have

$$\limsup_{n\to\infty} T_n^{(3)} \leqslant 0 \ \ a.s.$$

Combining (3.5), (3.7) and (3.8), we get

$$\limsup_{n\to\infty} T_n \leqslant 0 \ \ a.s. \tag{3.9}$$

Replacing X_i by $-X_i$ in (3.9), we can get

$$\liminf_{n\to\infty} T_n \geqslant 0 \ \ a.s. \tag{3.10}$$

The desired result (3.4) follows from (3.9) and (3.10) immediately. This completes the proof of the theorem. □

Similar to the proof of Theorem 3.1, we can get the following result by using Lemma 2.4. The details are omitted.

THEOREM 3.2. *Let $\{X_n, n \geqslant 1\}$ be a sequence of END random variables, which is stochastically dominated by a random variable X. Suppose that $E|X|^{1/\beta} < \infty$ for some $0 < \beta \leqslant 1$ and $EX_n = 0$. If conditions (3.2) and (3.3) are satisfied, then (3.4) holds.*

4. Strong consistency of M estimator in a linear model with END errors

In Section 3, we give an almost sure convergence result for weighted sums of END random variables. By using the almost sure convergence result, we will further study the strong consistency of M estimator in a linear model with END errors. The main idea is inspired by Wu and Jiang (2011). Our main results are as follows.

THEOREM 4.1. *In the model (1.1), assume that $e, e_1, e_2, \cdots$ are END random errors with identical distribution. There exist positive constants Δ, c_1, c_2, and $\delta \in (0,1]$ such that the following conditions are satisfied:*

$$\psi(u \pm t) - \psi(u) \text{ is monotonic on } u \in \boldsymbol{R}^1, \text{ and } \psi(u \pm t) - \psi(u) \leqslant c_1,$$

$$\text{for } t \in (0, \Delta), u \in \boldsymbol{R}^1, \tag{4.1}$$

or

$$|\psi(u)| \leqslant c_1, \text{ for } u \in \boldsymbol{R}^1; \tag{4.1$'$}$$

$$E\psi(e) = 0, \text{ and } |E\psi(e+u)| \geqslant c_2|u|, \text{ for } |u| < \Delta; \tag{4.2}$$

$$d_n \leqslant c_1 n^{-\delta}; \tag{4.3}$$

$$E|\psi(e)|^{1/\delta} < \infty \text{ when } 0 < \delta < 1, \text{ and } E|\psi(e)|^{\alpha} < \infty \text{ for some } \alpha > 1 \text{ when } \delta = 1. \tag{4.4}$$

Then, $\hat{\beta}_n$ is a strongly consistent estimator of β.

Proof. Let $\hat{\beta}_n$ be the minimizer of (1.2) and β_0 be the true parameter. Let $x_{ni} = S_n^{-1/2} x_i$, $\beta_{n0} = S_n^{1/2} \beta_0$, $1 \leqslant i \leqslant n$. The model (1.1) can be rewritten as

$$Y_i = x'_{ni}\beta_{n0} + e_i, \ 1 \leqslant i \leqslant n, \tag{4.5}$$

and we have

$$\sum_{i=1}^{n} x_{ni}x'_{ni} = I_p, \quad \sum_{i=1}^{n} \|x_{ni}\|^2 = p, \ \ d_n = \max_{1 \leqslant i \leqslant n} \|x_{ni}\|^2, \tag{4.6}$$

where I_p is the $p \times p$ identity matrix.

Let $\hat{\beta}_{n0}$ be an M estimator of β_{n0} in the model (4.5), it follows that $\hat{\beta}_{n0} = S_n^{1/2}\hat{\beta}_n$. Without loss of generality, we can suppose that the true parameter $\beta_0 = 0$ in the model (1.1), i.e., $\beta_{n0} = 0$ in the model (4.5). Let

$$\sum_{i=1}^{n} \rho(e_i - x'_i\hat{\beta}_{n0}) = \min_{\beta \in \mathbf{R}^p} \sum_{i=1}^{n} \rho(e_i - x'_i\beta). \tag{4.7}$$

Denote the unit sphere $U = \{\beta : \beta \in \mathbf{R}^p, \|\beta\| = 1\}$. Let $\varepsilon > 0$ be any given constant. Without loss of generality, it can be assumed that $2\sqrt{c_1}\varepsilon < \Delta$. Define

$$D_n(\beta) = \sum_{i=1}^{n} \{\rho(e_i - x'_{ni}\beta) - \rho(e_i)\}, \ \ \beta \in \mathbf{R}^p.$$

Then, $D_n(\cdot)$ is a convex function and $D_n(0) = 0$.

Let $\omega_{ni} = -\varepsilon n^{\delta/2} x_{ni}, r \in U$, by the definition of ψ, we have

$$\begin{aligned} D_n(\varepsilon n^{\delta/2} r) &= \sum_{i=1}^{n} \{\rho(e_i - \varepsilon n^{\delta/2} x'_{ni} r) - \rho(e_i)\} \\ &= \sum_{i=1}^{n} \int_0^{\omega_{ni} r} \{\psi(e_i + t) - \psi(e_i)\} dt + \sum_{i=1}^{n} \omega'_{ni} r \psi(e_i) \\ &\triangleq I_{1n}(r) + I_{2n}(r). \end{aligned}$$

Hence,

$$\inf_{r\in U} D_n(\varepsilon n^{\delta/2} r) \geqslant \inf_{r\in U} I_{1n}(r) + \inf_{r\in U} I_{2n}(r) \geqslant \inf_{r\in U} I_{1n}(r) - \sup_{r\in U} |I_{2n}(r)|. \tag{4.8}$$

We can divide U into N equal parts, $U_1, U_2, \cdots, U_N$, such that the diameter of each part is less than n^{-2}, and $N \leqslant (2n^2 + 1)^p$. Let T_j be the smallest close convex set covering U_j. For a fixed T_j, there are following three cases.

(i) $\omega'_{ni} r \geqslant 0$ for all $r \in T_j$, then there exists a $r_{ij} \in T_j$ such that $\omega'_{ni} r_{ij} = \inf_{r\in T_j}\{\omega'_{ni} r\}$;

(ii) $\omega'_{ni} r \leqslant 0$ for all $r \in T_j$, then there exists a $r_{ij} \in T_j$ such that $\omega'_{ni} r_{ij} = \sup_{r\in T_j}\{\omega'_{ni} r\}$;

(iii) $\omega'_{ni} r_1 > 0$ for some $r_1 \in T_j$, and $\omega'_{ni} r_2 < 0$ for some $r_2 \in T_j$, then there exists a $r_{ij} \in T_j$ such that $\omega'_{ni} r_{ij} = 0$.

Let

$$R_i(t) = \psi(e_i + t) - \psi(e_i), \ \ \tilde{R}_i(t) = R_i(t) - ER_i(t).$$

By the monotonicity of ψ, we know that $R_i(t)$ and $ER_i(t)$ are increasing on t, combining the selection of r_{ij} and $U \subset \cup_{j=1}^{N} T_j$, we get

$$\begin{aligned} \inf_{r\in U} I_{1n}(r) &= \inf_{r\in U} \sum_{i=1}^{n} \int_0^{\omega'_{ni} r} R_i(t) dt \geqslant \inf_{1\leqslant j\leqslant N} \inf_{r\in T_j} \sum_{i=1}^{n} \int_0^{\omega'_{ni} r} R_i(t) dt \\ &\geqslant \inf_{1\leqslant j\leqslant N} \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} R_i(t) dt = \inf_{1\leqslant j\leqslant N} \left\{ \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} \tilde{R}_i(t) dt + \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} ER_i(t) dt \right\} \\ &= \min_{1\leqslant j\leqslant N} \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} ER_i(t) dt \left[1 + \frac{\sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} \tilde{R}_i(t) dt}{\sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} ER_i(t) dt} \right] \\ &\geqslant \min_{1\leqslant j\leqslant N} \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} ER_i(t) dt \left[1 - \frac{\left| \sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} \tilde{R}_i(t) dt \right|}{\sum_{i=1}^{n} \int_0^{\omega'_{ni} r_{ij}} ER_i(t) dt} \right]. \end{aligned} \tag{4.9}$$

Let $r \in U_j$ and $r_{ij} \in T_j$. By (4.6) and the definitions of U_j and T_j, for sufficiently large n,

$$\begin{aligned}
-\sum_{i=1}^{n}(x'_{ni}r_{ij})^2+\sum_{i=1}^{n}(x'_{ni}r)^2 &\leqslant \left|\sum_{i=1}^{n}((x'_{ni}r_{ij})^2-(x'_{ni}r)^2)\right| = \left|\sum_{i=1}^{n}(r_{ij}-r)'x_{ni}x'_{ni}(r_{ij}+r)\right| \\
&\leqslant \sum_{i=1}^{n}\|r_{ij}-r\|\|x_{ni}\|^2(\|r_{ij}-r\|+2\|r\|) \\
&\leqslant \sum_{i=1}^{n}n^{-2}(n^{-2}+2)\|x_{ni}\|^2 \leqslant 3n^{-2}p < 1/2,
\end{aligned}$$

which combing (4.6) yields that

$$\begin{aligned}
\sum_{i=1}^{n}(x'_{ni}r_{ij})^2 &> \sum_{i=1}^{n}(x'_{ni}r)^2-1/2 = \sum_{i=1}^{n}r'x_{ni}x'_{ni}r-1/2 \\
&= r'\sum_{i=1}^{n}x_{ni}x'_{ni}r-1/2 = \|r\|^2-1/2 \\
&= 1/2, \quad 1\leqslant j\leqslant N. \qquad (4.10)
\end{aligned}$$

By (4.3), (4.6) and the selection of ε, for sufficiently large n, and for $i=1,2,\ldots,n$ and $j=1,2,\ldots,N$,

$$\begin{aligned}
|\omega'_{ni}r_{ij}| &= |\varepsilon n^{\delta/2}x'_{ni}r_{ij}| \leqslant \varepsilon n^{\delta/2}\|x_{ni}\|(\|r_{ij}-r\|+\|r\|) \\
&\leqslant \varepsilon n^{\delta/2}d_n^{1/2}(n^{-2}+1) \leqslant 2\sqrt{c_1}\varepsilon < \Delta, \qquad (4.11)
\end{aligned}$$

and by (4.2) and (4.10),

$$\begin{aligned}
\min_{1\leqslant j\leqslant N}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt &\geqslant \min_{1\leqslant j\leqslant N}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} c_2t\,dt \\
&\geqslant \frac{c_2}{2}\min_{1\leqslant j\leqslant N}\sum_{i=1}^{n}(\omega'_{ni}r_{ij})^2 \\
&= \frac{c_2\varepsilon^2n^\delta}{2}\min_{1\leqslant j\leqslant N}\sum_{i=1}^{n}(x'_{ni}r_{ij})^2 \\
&\geqslant c_2\varepsilon^2n^\delta/4. \qquad (4.12)
\end{aligned}$$

For fixed $j=1,2,\ldots,N$, let $Y_{ni}=\int_0^{\omega'_{ni}r_{ij}}\tilde{R}_i(t)dt$.

(i) If (4.1) holds, then $\psi(e_i+t)-\psi(e_i)$ is monotonic on e_i, and thus for fixed $n\geqslant 1$, $\{Y_{ni}, i\leqslant n\}$ are monotonic on e_i, which implies that for fixed $n\geqslant 1$, $\{Y_{ni}, 1\leqslant i\leqslant n\}$ are END random variables with $EY_{ni}=0$ by Lemma 2.1. It follows by (4.1) and (4.11) that

$$|Y_{ni}| \leqslant 2c_1|\omega'_{ni}r_{ij}| < 2c_1\Delta \triangleq c_3.$$

By (4.1) and (4.11) again, we have

$$\begin{aligned}
B_n &\triangleq \sum_{i=1}^{n} EY_{ni}^2 = \sum_{i=1}^{n} E\left(\int_0^{\omega'_{ni}r_{ij}} \tilde{R}_i(t)dt\right)^2 \\
&= \sum_{i=1}^{n} E\left(\int_0^{\omega'_{ni}r_{ij}} R_i(t)dt - E\left(\int_0^{\omega'_{ni}r_{ij}} R_i(t)dt\right)\right)^2 \\
&\leqslant \sum_{i=1}^{n} E\left(\int_0^{\omega'_{ni}r_{ij}} R_i(t)dt\right)^2 = \sum_{i=1}^{n} E\left(\int_0^{\omega'_{ni}r_{ij}} R_i(t)dt \int_0^{\omega'_{ni}r_{ij}} R_i(t)dt\right) \\
&\leqslant c_1 \sum_{i=1}^{n} |\omega'_{ni}r_{ij}| \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt \leqslant \frac{c_3}{2} \sum_{i=1}^{n} \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt. \qquad (4.13)
\end{aligned}$$

Let $u = \frac{1}{2}\sum_{i=1}^{n} \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt, t = 1/(2c_3)$. It follows that $tc_3 \leqslant 1$. Applying Lemma 2.3 for $\{Y_{ni}, 1 \leqslant i \leqslant n\}$, we have by (4.12), (4.13) and $N \leqslant (2n^2+1)^p$ that

$$\begin{aligned}
&P\left(\bigcup_{j=1}^{N}\left(\left|\sum_{i=1}^{n} Y_{ni}\right| \geqslant \frac{1}{2}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right)\right) \\
\leqslant& \sum_{j=1}^{N} P\left(\left|\sum_{i=1}^{n} Y_{ni}\right| \geqslant \frac{1}{2}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right) \\
\leqslant& 2M\sum_{j=1}^{N} \exp\left(-\frac{1}{4c_3}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt + \frac{1}{4c_3^2}B_n\right) \\
\leqslant& 2M\sum_{j=1}^{N} \exp\left(-\frac{1}{4c_3}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt + \frac{1}{8c_3}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right) \\
\leqslant& 2M\sum_{j=1}^{N} \exp\left(-\frac{c_2\varepsilon^2}{32c_3}n^{\delta}\right) \leqslant 2M(2n^2+1)^p \exp\left(-\frac{c_2\varepsilon^2}{32c_3}n^{\delta}\right),
\end{aligned}$$

which implies that

$$\sum_{n=1}^{\infty} P\left(\bigcup_{j=1}^{N}\left(\left|\sum_{i=1}^{n} Y_{ni}\right| \geqslant \frac{1}{2}\sum_{i=1}^{n}\int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right)\right) < \infty. \qquad (4.14)$$

(ii) If $(4.1')$ hold, i.e., $|\psi(u)| \leqslant c_1$. Since $Y_{ni} = \int_0^{\omega'_{ni}r_{ij}} (\psi(e_i+t) - E\psi(e_i+t))dt - \int_0^{\omega'_{ni}r_{ij}} (\psi(e_i) - E\psi(e_i))dt \triangleq Y_{ni}^{(1)} - Y_{ni}^{(2)}$, it follows that for fixed $n \geqslant 1$, $\{Y_{ni}^{(1)}, 1 \leqslant i \leqslant n\}$ and $\{Y_{ni}^{(2)}, 1 \leqslant i \leqslant n\}$ are both monotonic on e_i. Hence, for fixed $n \geqslant 1$, it can be checked that $\{Y_{ni}^{(1)}, 1 \leqslant i \leqslant n\}$ and $\{Y_{ni}^{(2)}, 1 \leqslant i \leqslant n\}$ are both END random variables with $EY_{ni}^{(j)} = 0, j = 1,2$. It follows by $(4.1')$ and (4.11) that

$$|Y_{ni}^{(j)}| \leqslant 2c_1|\omega'_{ni}r_{ij}| < 2c_1\Delta \triangleq c_3, \quad j = 1,2.$$

By $(4.1')$ and (4.11) again, we have

$$
\begin{aligned}
B_n^{(j)} &\triangleq \sum_{i=1}^n E(Y_{ni}^{(j)})^2 \leqslant \sum_{i=1}^n 4c_1^2|\omega'_{ni}r_{ij}|^2 = \sum_{i=1}^n 4c_1^2\varepsilon^2 n^\delta |x'_{ni}r_{ij}|^2 \\
&\leqslant 4c_1^2\varepsilon^2 n^\delta \sum_{i=1}^n \|x_{ni}\|^2\|r_{ij}\|^2 \leqslant 4c_1^2\varepsilon^2 n^\delta \sum_{i=1}^n \|x_{ni}\|^2(\|r_{ij}-r\|+\|r\|)^2 \\
&\leqslant 4c_1^2\varepsilon^2 n^\delta \sum_{i=1}^n \|x_{ni}\|^2(n^{-2}+1)^2 \leqslant 16pc_1^2\varepsilon^2 n^\delta. \qquad (4.15)
\end{aligned}
$$

Choose $0<\varepsilon_1\leqslant 1$ such that $A \triangleq \frac{c_2\varepsilon_1}{2\times 16^2 pc_1^2} \leqslant \frac{1}{c_3}$. Let $u=\frac{1}{16}c_2\varepsilon^2 n^\delta>0, t=A>0$. It follows that $tc_3\leqslant 1$. Applying Lemma 2.3 for $\{Y_{ni}^{(1)}, 1\leqslant i\leqslant n\}$ and $\{Y_{ni}^{(2)}, 1\leqslant i\leqslant n\}$, we have by (4.12) and (4.15) that

$$
\begin{aligned}
&P\left(\bigcup_{j=1}^N \left(\left|\sum_{i=1}^n Y_{ni}\right| \geqslant \frac{1}{2}\sum_{i=1}^n \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right)\right) \\
&\leqslant \sum_{j=1}^N P\left(\left|\sum_{i=1}^n Y_{ni}\right| \geqslant \frac{c_2\varepsilon^2 n^\delta}{8}\right) \\
&\leqslant \sum_{j=1}^N \left[P\left(\left|\sum_{i=1}^n Y_{ni}^{(1)}\right| \geqslant \frac{c_2\varepsilon^2 n^\delta}{16}\right) + P\left(\left|\sum_{i=1}^n Y_{ni}^{(2)}\right| \geqslant \frac{c_2\varepsilon^2 n^\delta}{16}\right)\right] \\
&\leqslant \sum_{j=1}^N 4M\exp(-tu+t^2B_n) \leqslant \sum_{j=1}^N 4M\exp\left(-\frac{c_2^2\varepsilon_1\varepsilon^2}{4\times 16^3 pc_1^2}n^\delta\right) \\
&\leqslant 4M(2n^2+1)^p\exp\left(-\frac{c_2^2\varepsilon_1\varepsilon^2}{4\times 16^3 pc_1^2}n^\delta\right).
\end{aligned}
$$

Thus, (4.14) is also established. By Borel-Cantelli Lemma,

$$
P\left(\bigcup_{j=1}^N \left(\left|\sum_{i=1}^n Y_{ni}\right| \geqslant \frac{1}{2}\sum_{i=1}^n \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right), i.o.\right) = 0,
$$

thus,

$$
P\left(\bigcap_{j=1}^N \left(\left|\sum_{i=1}^n Y_{ni}\right| < \frac{1}{2}\sum_{i=1}^n \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt\right), \text{for sufficiently large } n\right) = 1.
$$

By the definition of Y_{ni}, with probability one ($wp1$, in short), for sufficiently large n, for $\forall\, j=1,2,\cdots,N$,

$$
\frac{\left|\sum_{i=1}^n \int_0^{\omega'_{ni}r_{ij}} \tilde{R}_i(t)dt\right|}{\sum_{i=1}^n \int_0^{\omega'_{ni}r_{ij}} ER_i(t)dt} < \frac{1}{2}.
$$

So, $wp1$, for sufficiently large n,

$$\max_{1\leqslant j\leqslant N}\frac{\left|\sum_{i=1}^{n}\int_{0}^{\omega_{ni}'r_{ij}}\tilde{R}_i(t)dt\right|}{\sum_{i=1}^{n}\int_{0}^{\omega_{ni}'r_{ij}}ER_i(t)dt}<\frac{1}{2}. \tag{4.16}$$

Substitute the above inequality and (4.12) in (4.9), we can see that $wp1$, for sufficiently large n,

$$\inf_{r\in U} I_{1n}(r)\geqslant\frac{1}{8}c_2\varepsilon^2 n^{\delta}. \tag{4.17}$$

Denote

$$x_{ni}=\begin{pmatrix}x_{ni1}\\ x_{ni2}\\ \vdots\\ x_{nip}\end{pmatrix},\quad r=\begin{pmatrix}r_1\\ r_2\\ \vdots\\ r_p\end{pmatrix},\quad a_{ni}=\begin{cases}\frac{x_{nij}}{n^{\delta/2}}, & i\leqslant n,\\ 0, & i>n\end{cases}\quad \text{for fixed } j=1,2,\dots,p.$$

It follows by (4.3) and (4.6) that for fixed $j=1,2,\dots,p$, $|x_{nij}|\leqslant\|x_{ni}\|\leqslant d_n^{1/2}\leqslant\sqrt{c_1}n^{-\delta/2}$, and $\sum_{i=1}^{n}x_{nij}^2=1$, thus

$$|a_{ni}|\leqslant\sqrt{c_1}n^{-\delta},\ \ n\geqslant 1,\ 1\leqslant i\leqslant n.$$

(i) When $\delta\leqslant 1/2$, i.e., $\frac{1}{\delta}\geqslant 2$, it follows that $p\triangleq\min(\frac{1}{\delta},2)=2$ and

$$\sum_{i=1}^{n}|a_{ni}|^p=\sum_{i=1}^{n}a_{ni}^2=n^{-\delta}\sum_{i=1}^{n}x_{nij}^2=n^{-\delta}.$$

(ii) When $\frac{1}{2}<\delta<1$, we have $1<1/\delta<2$. Thus, $p\triangleq\min(\frac{1}{\delta},2)=\frac{1}{\delta}$, and $\frac{1}{2}(\frac{1}{\delta}-1)>0$, by Hölder's inequality,

$$\begin{aligned}\sum_{i=1}^{n}|a_{ni}|^p&=\sum_{i=1}^{n}|a_{ni}|^{1/\delta}=n^{-\frac{1}{2}}\sum_{i=1}^{n}|x_{nij}|^{1/\delta}\\ &\leqslant n^{-\frac{1}{2}}\left(\sum_{i=1}^{n}|x_{nij}|^{\frac{1}{\delta}\cdot 2\delta}\right)^{\frac{1}{2\delta}}n^{\frac{2\delta-1}{2\delta}}=n^{-\frac{1}{2}(\frac{1}{\delta}-1)}.\end{aligned}$$

(iii) When $\delta=1$, without loss of generality, we can assume that $1<\alpha<2$ in (4.4). Thus, $p\triangleq\min(\alpha,2)=\alpha$, by the Hölder's inequality,

$$\begin{aligned}\sum_{i=1}^{n}|a_{ni}|^p&=\sum_{i=1}^{n}|a_{ni}|^{\alpha}=n^{-\frac{\alpha}{2}}\sum_{i=1}^{n}|x_{nij}|^{\alpha}\\ &\leqslant n^{-\frac{\alpha}{2}}\left(\sum_{i=1}^{n}|x_{nij}|^{\alpha\frac{2}{\alpha}}\right)^{\frac{\alpha}{2}}n^{\frac{2-\alpha}{2}}=n^{-(\alpha-1)}.\end{aligned}$$

Because $\psi(e_i)$ is increasing on e_i, by Lemma 2.1, $\{\psi(e_i); i \geqslant 1\}$ is also a sequence of END random variables. By applying Theorem 3.1 for $\{\psi(e_i), i \geqslant 1\}$ and $\{a_{ni}, 1 \leqslant i \leqslant n, n \geqslant 1\}$, we have by (4.4) that for fixed $j = 1, 2, \dots, p$,

$$\sum_{i=1}^{n} a_{ni}\psi(e_i) = n^{-\delta/2}\sum_{i=1}^{n} x_{nij}\psi(e_i) \to 0 \ \ a.s.,\ n \to \infty.$$

Thus,

$$\begin{aligned}
n^{-\delta}\sup_{r\in U}|I_{2n}(r)| &= n^{-\delta}\sup_{r\in U}\left|\sum_{i=1}^{n}\omega'_{ni} r\psi(e_i)\right| = n^{-\delta}\sup_{r\in U}\left|\sum_{i=1}^{n}\varepsilon n^{\delta/2}x'_{ni} r\psi(e_i)\right| \\
&= \varepsilon n^{-\delta/2}\sup_{r\in U}\left|\sum_{i=1}^{n}\sum_{j=1}^{p} x_{nij} r_j \psi(e_i)\right| = \varepsilon n^{-\delta/2}\sup_{r\in U}\left|\sum_{j=1}^{p}\left(\sum_{i=1}^{n} x_{nij}\psi(e_i)\right) r_j\right| \\
&\leqslant \varepsilon n^{-\delta/2}\sup_{r\in U}\sqrt{\sum_{j=1}^{p}\left(\sum_{i=1}^{n} x_{nij}\psi(e_i)\right)^2}\sqrt{\sum_{j=1}^{p} r_j^2} \\
&= \varepsilon\sqrt{\sum_{j=1}^{p}\left(n^{-\delta/2}\sum_{i=1}^{n} x_{nij}\psi(e_i)\right)^2} \to 0 \ \ a.s.,\ n \to \infty,
\end{aligned}$$

which implies that $wp1$, for sufficiently large n,

$$\sup_{r\in U}|I_{2n}(r)| \leqslant \frac{c_2\varepsilon^2}{16} n^{\delta}.$$

Substitute the above inequality and (4.17) in (4.8), we have $wp1$, for sufficiently large n,

$$\inf_{r\in U} D_n(\varepsilon n^{\delta/2} r) \geqslant c_2\varepsilon^2 n^{\delta}/16 > 0. \tag{4.18}$$

Closed surface $S = \left\{\beta : \beta \in \mathbf{R}^p, \|\beta\| = \varepsilon n^{\delta/2}\right\}$ divides $\mathbf{R}^p$ into two parts,

$$A = \left\{\beta : \beta \in \mathbf{R}^p, \|\beta\| < \varepsilon n^{\delta/2}\right\}$$

and

$$B = \left\{\beta : \beta \in \mathbf{R}^p, \|\beta\| \geqslant \varepsilon n^{\delta/2}\right\}.$$

Because $D_n(\cdot)$ is a convex function, $D_n(0) = 0$, 0 is a interior point of A, and $\inf_{\beta\in S} D_n(\beta) = \inf_{r\in U} D_n(\varepsilon n^{\delta/2} r) > 0 = D_n(0)$ from (4.18). Thus, for all $\beta \in B, D_n(\beta) > 0$. On the other hand, it follows by the definition (4.7) of $\hat{\beta}_{n0}$ that $D_n(\hat{\beta}_{n0}) \leqslant 0$. Thus, $\hat{\beta}_{n0} \in A$, i.e. $wp1$, for sufficiently large n,

$$\|\hat{\beta}_{n0}\| < \varepsilon n^{\delta/2}, \ \text{ for any given } \varepsilon > 0,$$

which implies that $wp1$,

$$n^{-\delta/2}\|\hat{\beta}_{n0}\| \to 0, \quad n \to \infty. \tag{4.19}$$

Let λ_n be the smallest eigenvalue of S_n which is a positive definite matrix. Then $\lambda_n^{1/2}$ is the smallest eigenvalue of $S_n^{1/2}$, and λ_n^{-1} is the maximum eigenvalue of S_n^{-1}. By (3.10) of Chen and Zhao (1996) and (4.3), for fixed n_0,

$$\begin{aligned}
\lambda_{n_0}\lambda_n^{-1} &\leqslant \mathrm{tr}(S_{n_0}S_n^{-1}) = \sum_{i=1}^{n_0} \mathrm{tr}(x_i(x_i'S_n^{-1})) \\
&= \sum_{i=1}^{n_0} \mathrm{tr}((x_i'S_n^{-1})x_i) = \sum_{i=1}^{n_0} x_i'S_n^{-1}x_i \\
&\leqslant n_0 d_n \leqslant n_0 c_1 n^{-\delta}.
\end{aligned}$$

Therefore, let $c_4 = \sqrt{\frac{n_0 c_1}{\lambda_{n_0}}} > 0$, we have $1 \leqslant c_4 \lambda_n^{1/2} n^{-\delta/2}$. Combining $\|S_n^{1/2}\hat{\beta}_n\| \geqslant \lambda_n^{1/2}\|\hat{\beta}_n\|$ and (4.19), we can get that $wp1$,

$$\begin{aligned}
\|\hat{\beta}_n\| &\leqslant c_4 n^{-\delta/2}\|\lambda_n^{1/2}\hat{\beta}_n\| \leqslant c_4 n^{-\delta/2}\|S_n^{1/2}\hat{\beta}_n\| \\
&= c_4 n^{-\delta/2}\|\hat{\beta}_{n0}\| \to 0, \quad n \to \infty,
\end{aligned}$$

i.e. $wp1$,

$$\hat{\beta}_n \to 0, \quad n \to \infty.$$

This completes the proof of the theorem. □

By using Lemma 2.4 and Theorem 3.2, we can get the following result for END random errors, which is stochastically dominated by a random variable. The proof is similar to that of Theorem 4.1, so the details are omitted.

THEOREM 4.2. *In the model (1.1), assume that $e_1, e_2, \cdots$ are END random errors, which are stochastically dominated by a random variable e. There exist positive constants $\Delta, c, c_1, c_2, \delta \in (0,1]$ such that (4.1) (or $(4.1')$), (4.3), (4.4) hold, and*

$$E\psi(e_k) = 0, \quad \text{and } |E\psi(e_k + u)| \geqslant c_2|u|, \quad \text{for } |u| < \Delta, k \geqslant 1.$$

Then $\hat{\beta}_n$ is a strongly consistent estimator of β.

Acknowledgements. The authors are most grateful to the Editor and two anonymous referees for careful reading of the manuscript and valuable suggestions which helped in improving an earlier version of this paper.

REFERENCES

[1] X. R. CHEN, L. C. ZHAO, *Strong consistency of M-estimates of multiple regression coefficients*, Systems Science and Mathematical Sciences, **8** (1), 82–87, 1995.

[2] X. R. CHEN, L. C. ZHAO, *M-methods in Linear Model*, Shanghai: Scientific and Technical Publishers, 1996.

[3] Y. CHEN, A. CHEN, K. W. NG, *The strong law of large numbers for extend negatively dependent random variables*, Journal of Applied Probability, **47**, 908–922, 2010.

[4] Y. CHEN, K. C. YUEN, K. W. NG, *Precise large deviations of random sums in presence of negative dependence and consistent variation*, Methodology and Computing in Applied Probability, **13** (4), 821–833, 2011.

[5] J. R. COLLINS, V. D. SZATMARI, *Maximal asymptotic biases of M estimators of location with preliminary scale estimates*, Communications in Statistics – Theory and Methods, **33** (8), 1877–1866, 2004.

[6] C. DJALIL, C. DIDIER, *On the strong consistency of asymptotic M Estimators*, Journal of Statistical Planning and Inference, **137**, 2774–2783, 2007.

[7] X. M. HE, Q. M. SHAO, *A general bahadur representation of M Estimator and its application to linear regression with nonstochastic designs*, The Annals of Statistics, **24** (6), 2608–2630, 1996.

[8] P. GEORGIOS, *Application of M estimators to cross-section effect models*, Communications in Statistics-Simulation and Computation, **34** (3), 601–616, 2005.

[9] T. C. HU, K. L. WANG, A. ROSALSKY, *Complete convergence theorems for extended negatively dependent random variables*, Sankhyā A: The Indian Journal of Statistics, **77** (1), 1–29, 2015.

[10] P. J. HUBER, *Robust regression: asymptotics, conjectures and Monte Carlo*, The Annals of Statistics, **1** (5), 799–821, 1973.

[11] L. LIU, *Precise large deviations for dependent random variables with heavy tails*, Statistics and Probability Letters, **79**, 1290–1298, 2009.

[12] L. LIU, *Necessary and sufficient conditions for moderate deviations of dependent random Variables with heavy tails*, Science in China Series A: Mathematics, **53** (6), 1421–1434, 2010.

[13] E. L. LEHMANN, *Some concepts of dependence*, The Annals of Mathematical Statistics, **37**, 1137–1153, 1966.

[14] K. JOAG-DEV, F. PROSCHAN, *Negative association of random variables with applications*, The Annals of Statistics, **11** (1), 286–295, 1983.

[15] D. H. QIU, P. Y. CHEN, R. G. ANTONINI, A. VOLODIN, *On the complete convergence for arrays of rowwise extended negatively dependent random variables*, Journal of the Korean Mathematical Society, **50** (2), 379–392, 2013.

[16] S. SEIJA, T. SARA, O. HANNU, *Symmetrized M estimators of multivariate scatter*, Journal of Multivariate Analysis, **98**, 1611–1629, 2007.

[17] A. T. SHEN, *Some strong limit theorems for arrays of rowwise negatively orthant-dependent random variables*, Journal of Inequalities and Applications, Volume 2011, Article ID 93, 10 pages, 2011a.

[18] A. T. SHEN, *Probability inequalities for END sequence and their applications*, Journal of Inequalities and Applications, Volume 2011, Article ID 98, 12 pages, 2011b.

[19] A. T. SHEN, *On the strong convergence rate for weighted sums of arrays of rowwise negatively orthant dependent random variables*, RACSAM, **107** (2), 257–271, 2013.

[20] A. T. SHEN, *On asymptotic approximation of inverse moments for a class of nonnegative random variables*, Statistics – A Journal of Theoretical and Applied Statistics, **48** (6), 1371–1379, 2014.

[21] A. T. SHEN, Y. ZHANG, A. VOLODIN, *Applications of the Rosenthal-type inequality for negatively superadditive dependent random variables*, Metrika, **78** (3), 295–311, 2015.

[22] S. H. SUNG, *Complete convergence for weighted sums of negatively dependent random variables*, Statistical Papers, **53**, 73–82, 2012.

[23] A. VOLODIN, *On the Kolmogorov exponential inequality for negatively dependent random variables*, Pakistan Journal of Statistics, **18**, 249–254, 2002.

[24] S. J. WANG, W. S. WANG, *Precise large deviations for random sums of END real-valued random variables with consistent variation*, Journal of Mathematical Analysis and Applications, **402**, 660–667, 2013.

[25] X. J. WANG, Z. Y. SI, *Complete consistency of the estimator of nonparametric regression model under ND sequence*, Statistical Papers, **56** (3), 585–596, 2015.

[26] X. J. WANG, S. J. WANG, S. H. HU, J. M. LING, Y. F. WEI, *On complete convergence of weighted sums for arrays of rowwise extended negatively dependent random variables*, Stochastics – An International Journal of Probability and Stochastic Processes, **85** (6), 1060–1072, 2013.

[27] X. J. WANG, X. Q. LI, S. H. HU, X. H. WANG, *On complete convergence for an extended negatively dependent sequence*, Communications in Statistics – Theory and Methods, **43** (14), 2923–2937, 2014.

[28] X. J. WANG, L. L. ZHENG, C. XU, S. H. HU, *Complete consistency for the estimator of nonparametric regression models based on extended negatively dependent errors*, Statistics: A Journal of Theoretical and Applied Statistics, **49** (2), 396–407, 2015.

[29] Q. Y. WU, *Strong consistency of M estimator in linear model for negatively associated samples*, Journal of Systems Science and Complexity, **19** (4), 592–600, 2006a.

[30] Q. Y. WU, *Probability Limit Theory for Mixing sequences*, Science Press of China, Beijing, 2006b.

[31] Q. Y. WU, Y. Y. JIANG, *The strong consistency of M estimator in linear model for negatively dependent random samples*, Communications in Statistics – Theory and Methods, **40**, 476–491, 2011.

[32] Y. F. WU, M. GUAN, *Convergence properties of the partial sums for sequences of END random variables*, Journal of the Korean Mathematical Society, **49** (6), 1097–1110, 2012.

[33] Y. F. WU, M. O. CABRERA, A. VOLODIN, *Complete convergence and complete moment convergence for arrays of rowwise END random variables*, Glasnik Matematicki, **49** (69), 449–468, 2014.

[34] S. C. YANG, *Strong consistency of M estimator in linear model*, Acta Mathematica Sinica, **45** (1), 21–28, 2002.

[35] W. Z. YANG, H. Y. XU, L. CHEN, S. H. HU, *Complete consistency of estimators for regression models based on extended negatively dependent errors*, Statistical Papers, DOI 10.1007/s00362-016-0771-x, in press, 2016.

[36] H. ZAREI, H. JABBARI, *Complete convergence of weighted sums under negative dependence*, Statistical Papers, **52**, 413–418, 2011.

(Received November 10, 2015)

Xin Deng
School of Mathematical Sciences
Anhui University
Hefei, 230601, P. R. China

Shijie Wang
School of Mathematical Sciences
Anhui University
Hefei, 230601, P. R. China

Rui Wang
School of Mathematical Sciences
Anhui University
Hefei, 230601, P. R. China

Xiujuan Xie
School of Mathematical Sciences
Anhui University
Hefei, 230601, P. R. China

Xuejun Wang
School of Mathematical Sciences
Anhui University
Hefei, 230601, P. R. China