Ann. Inst. Statist. Math.
Vol. 50, No. 1, 87–117 (1998)

SECOND ORDER REPRESENTATIONS OF THE LEAST ABSOLUTE DEVIATION REGRESSION ESTIMATOR*

MIGUEL A. ARCONES

Department of Mathematics, University of Texas, Austin, TX 78712-1082, U.S.A.

(Received April 8, 1996; revised April 28, 1997)

Abstract. We consider exact weak and strong Bahadur-Kiefer representations of the least absolute deviation estimator for the linear regression model. The precise behavior of these representations is obtained under minimal conditions.

Key words and phrases: Least absolute deviation, regression, empirical processes.

1. Introduction

We consider the following linear regression model:

$$Y_i = x_i'\beta + U_i, \qquad i = 1, \dots, n, \tag{1.1}$$

where $\{U_i\}_{i=1}^{\infty}$ is a sequence of independent identically distributed random variables; $\{x_i\}_{i=1}^{\infty}$ is a sequence of p dimensional vectors and β is a p dimensional parameter to be estimated. Usually $x_i' = (1, x_{i,2}, \dots, x_{i,p})$. U will denote a r.v. with the distribution of U_1. We want to estimate β from a sample $Y_1, \dots, Y_n$. A least absolute deviation (LAD) estimator $\hat{\beta}_n$ of β is a random variable such that

$$n^{-1}\sum_{i=1}^{n} |Y_i - x_i'\hat{\beta}_n| = \inf_{b \in \mathbb{R}^d} n^{-1} \sum_{i=1}^{n} |Y_i - x_i'b|, \tag{1.2}$$

where $|v|$ is the Euclidean distance.

This well-known model represents the dependence of two variables y and x, possibly multivariate (see for example Draper and Smith (1981); and Rao and Toutenburg (1995)). The value of x has some influence on the value of y. x is called the predictor variable. Y is called the response variable. β represents the linear relation between the two variables. U is the random error. The problem is

* Research partially supported by NSF Grant DMS–93–02583 and carried out at the Department of Mathematics of the University of Utah.

to estimate β. The usual method is to estimate β is to use the least squares (LS) method. The advantage of this method is the easy computability of the estimator. In contrast, we consider the least absolute deviation (LAD) estimator. The LAD estimator has better robustness properties than the LS estimator. Since using a computer program, it is not difficult (nor long) to compute the LAD estimator, the LAD seems a better choice than the LS estimator. Observe that convexity guarantees the existence of a solution to (1.2). These statistical methods have a long history (see for example Stigler (1986)).

It is known that, under regularity conditions, $\{V_n(\hat{\beta}_n - \beta)\}_{n=1}^{\infty}$ converges in distribution, where V_n is the $p \times p$ matrix determined by

$$V_n^2 = \sum_{j=1}^{n} x_j x_j', \tag{1.3}$$

see Bassett and Koenker (1978), Ruppert and Carroll (1980), Bloomfield and Steiger (1983), Koenker and Portnoy (1987) and Bai *et al.* (1989) and Pollard (1991). In fact, it follows from their work that

$$2F_U'(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j \xrightarrow{\mathrm{P}} 0 \tag{1.4}$$

where $F_U(u) = \mathrm{P}\{U \le u\}$. Here, we will study the rate of convergence of

$$2F_U'(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j, \tag{1.5}$$

both in distribution and almost surely, where $\operatorname{sign}(x) = 1$ if $x > 0$; $\operatorname{sign}(x) = 0$ if $x = 0$ and $\operatorname{sign}(x) = -1$ if $x < 0$. Previous papers in this problem are the ones by Koenker and Portnoy (1987), Babu (1989), Rao and Zhao (1992), Arcones (1996) and He and Shao (1996). Koenker and Portnoy (1987) and He and Shao (1996) imposed conditions such as $n^{-1}\sum_{j=1}^{n} x_j x_j'$ converges to a positive definite matrix and $n^{-1}\sum_{j=1}^{n} |x_j|^{4+\tau}$ is a bounded sequence, for some $\tau > 0$. The problem with these conditions is that they are difficult to justify in a real problem, unless we may choose the values x_j beforehand. In the case $p = 2$, a common choice for the points x_j is to select evenly spaced points, in a region growing to infinity, with some repetitions and having more repetitions in the outer points than in the points in the middle (see for example Draper and Smith (1981)). For example, suppose that $n = 2km + 1$ and $x_j' = (1, -k + m^{-1}(j-1))$, for $1 \le j \le 2km + 1$. Then, $\sum_{j=1}^{n} x_j x_j'$ is of the order $\begin{pmatrix} n & 0 \\ 0 & nk^2 \end{pmatrix}$. Since $k \to \infty$, the conditions assumed by Koenker and Portnoy (1987) and He and Shao (1996) do not hold. In the previous computation, we did not allow repetitions between the x's. A similar computation shows that $\sum_{j=1}^{n} x_j x_j'$ is of the order $\begin{pmatrix} n & 0 \\ 0 & nk^2 \end{pmatrix}$, when the predictor points are $(1, -k + m^{-1}(j-1))$, for $1 \le j \le 2km + 1$ with some repetitions and the number of repetitions is nondecreasing with the distance of the point to the center of the range of the predictors values. Of course, the conditions assumed by Koenker and

Portnoy (1987) and He and Shao (1996) hold if $\{x_j\}$ is a sequence of i.i.d.r.v.'s with finite $(4+\tau)$-th moment. In this case, exact second order expansions for the LAD estimator follows from the work in Arcones (1996), assuming that $E[|x_i|^3] < \infty$. Babu (1989) considered general types of conditions in the sequence of regressors. The problem is that his bound does not give the right rate for some sequences of regressors (see Section 3 below). Koenker and Portnoy (1987) and Babu (1989) only considered LAD estimators. Instead of considering the particular function $h(x) = |x|$, Rao and Zhao (1992) and He and Shao (1996) considered a more general class of regression M-estimators.

There are several reasons to study the rate of convergence of term in (1.5). It is a measure of the differentiability of a M-estimator. It is also useful in the study of L-statistics (see Koenker and Portnoy (1987)). It is also useful in the construction of sequential fixed-width confidence intervals (see Chow and Robbins (1965)). Among other results, we will see that, under regularity conditions,

$$(1.6)\qquad a_n \left| 2F_U'(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^n \operatorname{sign}(U_j)V_n^{-1}x_j \right| = O_{\mathrm{P}}(1)$$

and

$$(1.7)\qquad \limsup_{n\to\infty} a_n(2\log\log n)^{-3/4} \cdot \left| 2F_U'(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^n \operatorname{sign}(U_j)V_n^{-1}x_j \right| = c \qquad \text{a.s.}$$

where $a_n = (\sum_{j=1}^n |V_n^{-1}x_j|^3)^{-1/2}$ and c is a positive constant.

The main tools, that we use, come from empirical processes. We apply a method very similar to the one in Arcones (1994): a central limit theorem (CLT) and a law of the iterated logarithm (LIL) holding uniformly over certain classes of functions. In that reference, the classes of functions are VC classes and the summands are i.i.d. Here, we use the results in Arcones (1995*a*, 1995*b*) which apply to some processes of sums of independent, but non-necessarily identically distributed r.v.'s. Another ingredient in the proofs will be the Talagrand isoperimetric inequality (Ledoux and Talagrand (1991), Theorem 6.17). We will use this inequality as an infinite dimensional Bernstein's inequality.

We will use the notation in empirical processes. We will denote by $\{\epsilon_j\}_{j=1}^\infty$ a sequence of Rademacher r.v.'s independent of the sequence $\{U_j\}_{j=1}^\infty$. c will design a constant that may vary from occurrence to occurrence.

2. The weak Bahadur-Kiefer representation of LAD regression estimators

In this section we study the distributional asymptotic behavior of (1.5). Most of the time, we will assume the following assumptions:

(A.1) U has a density $f_U(u)$ in a neighborhood of 0, $f_U(u)$ is continuous at 0, $f_U(0) \neq 0$ and

$$\mathrm{P}\{U \le x\} = 2^{-1} + x f_U(0) + O(x^2) \qquad \text{as} \quad x \to 0.$$

(A.2) For n large enough, $\sum_{j=1}^{n} x_j x_j'$ has a positive definite square root V_n.
(A.3) $\max_{1\le j\le n} |V_n^{-1} x_j| \to 0$.
(A.4) For each $\tau > 0$,

$$\sum_{j=1}^{n} |V_n^{-1} x_j| I_{a_n |V_n^{-1} x_j| > \tau} \to 0,$$

where $a_n = (\sum_{j=1}^{n} |V_n^{-1} x_j|^3)^{-1/2}$.

It is well known that, under (A.1)–(A.3), we have asymptotic normality of the LAD estimator (see Bassett and Koenker (1978); Bloomfield and Steiger (1983); Bai *et al.* (1989); and Pollard (1991)):

THEOREM 2.1. *Under conditions* (A.1)–(A.3),

$$2f_U(0)V_n(\hat\beta_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j \xrightarrow{\mathrm{P}} 0. \tag{2.1}$$

Observe that (2.1) and the multivariate CLT imply that

$$V_n(\hat\beta_n - \beta) \xrightarrow{d} g, \tag{2.2}$$

where g is a centered Gaussian random vector.

To get the second order expansion of the LAD estimator, we are requiring the extra condition (A.4). This condition is very weak. Observe that if $\{x_j\}_{j=1}^{\infty}$ is a sequence of $\mathbb{R}^p$-valued i.i.d.r.v.'s such that $E[|x_1|^2] < \infty$ and $E[x_1 x_1']$ is a positive definite matrix, then, with probability one, (A.3) and (A.4) hold. By the law of the large numbers, with probability one, $n^{-1}\sum_{j=1}^{n} x_j x_j'$ converges to $E[x_1 x_1']$. By the Marcinkiewicz law of the large numbers, with probability one, $n^{-1/2}\sum_{j=1}^{n} |x_j| \to 0$.

If $n = 2km + 1$ and $x_j' = (1, -k + m^{-1}(j-1))$, for $1 \le j \le 2km + 1$. Then, $\sum_{j=1}^{n} x_j x_j'$ is of the order $\begin{pmatrix} n & 0 \\ 0 & nk^2 \end{pmatrix}$, $\max_{1\le i\le n} |V_n^{-1} x_i|$ is of the order $n^{-1/2}$ and a_n is of the order $n^{1/4}$. So, conditions (A.3) and (A.4) hold in this case. Last assertion is also true when the points are chosen in a similar way for the multivariate case.

Next, we give the heuristic arguments that we will use. Given the definition of $\hat\beta_n$, we should expect that

$$a_n \sum_{j=1}^{n} \operatorname{sign}(U_j - x_j' V_n^{-1}(\hat\beta_n - \beta)) V_n^{-1} x_j \xrightarrow{\mathrm{P}} 0. \tag{2.3}$$

Let

$$W_n(t) := a_n \sum_{j=1}^{n} (\operatorname{sign}(U_j - x_j' V_n^{-1} t) - \operatorname{sign}(U_j) - E[\operatorname{sign}(U_j - x_j' V_n^{-1} t)]) V_n^{-1} x_j \tag{2.4}$$

and let

$$H_n(t) := a_n \sum_{j=1}^{n} \operatorname{sign}(U_j - x_j' V_n^{-1} t) V_n^{-1} x_j.$$

It is easy to see that

$$H_n(t) = W_n(t) + a_n \sum_{j=1}^{n} E[\operatorname{sign}(U_j - x_j' V_n^{-1} t)] V_n^{-1} x_j + a_n \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j$$

and

$$a_n \sum_{j=1}^{n} E[\operatorname{sign}(U_j - x_j' V_n^{-1} t) - \operatorname{sign}(U_j)] V_n^{-1} x_j = -2F_U'(0) a_n t + o(1).$$

Thus, we expect to get that

$$\begin{aligned} 0 = H_n(V_n(\hat{\beta}_n - \beta)) &= W_n(V_n(\hat{\beta}_n - \beta)) \\ &\quad - 2F_U'(0) a_n V_n(\hat{\beta}_n - \beta) + a_n \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j + o_{\mathrm{P}}(1). \end{aligned}$$

The sequence of real numbers $\{a_n\}$ is chosen so that the stochastic process $\{W_n(t) : |t| \le T\}$ converges to a nondegenerate limit, for each $T < \infty$. We will get that

$$2F_U'(0) a_n V_n(\hat{\beta}_n - \beta) - a_n \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j$$

converges in distribution to the limit of $W_n(V_n(\hat{\beta}_n - \beta))$. Next, we make previous arguments rigorous and precise. First, we prove (2.3).

LEMMA 2.1. *Under conditions* (A.1)–(A.4):

$$a_n \left| \sum_{j=1}^{n} \operatorname{sign}(U_j - x_j'(\hat{\beta}_n - \beta)) V_n^{-1} x_j \right| \xrightarrow{\mathrm{P}} 0.$$

PROOF. Let $\delta > 0$ be such that U has a density in $[-\delta, \delta]$. Since $\hat{\beta}_n$ is a solution of (1.2), for each $v \in \mathbb{R}^p$, $v \ne 0$,

$$\begin{aligned} 0 &\le \lim_{t \to 0+} t^{-1} \sum_{j=1}^{n} (|U_j - x_j'(\hat{\beta}_n + tv - \beta)| - |U_j - x_j'(\hat{\beta}_n - \beta)|) \\ &= -\sum_{j=1}^{n} x_j' v \operatorname{sign}(U_j - x_j'(\hat{\beta}_n - \beta)) I_{U_j \ne x_j'(\hat{\beta}_n - \beta)} + \sum_{j=1}^{n} |x_j' v| I_{U_j = x_j'(\hat{\beta}_n - \beta)}. \end{aligned}$$

Taking $v = -V_n^{-1}$ and $v = V_n^{-1}$, we get that

$$a_n \left| \sum_{j=1}^{n} \operatorname{sign}(U_j - x_j'(\hat{\beta}_n - \beta)) V_n^{-1} x_j \right| \le a_n \sum_{j=1}^{n} |V_n^{-1} x_j| I_{U_j = x_j'(\hat{\beta}_n - \beta)}.$$

Given $1 \le i_1 < i_2 < \cdots < i_{p+1} \le n$, there are $\lambda_1, \dots, \lambda_{p+1}$ such that $\sum_{j=1}^{p+1} \lambda_j x_j = 0$ and not all λ's are zero. If $U_{i_j} = z'_{i_j}(\hat\beta_n - \beta)$, for $1 \le j \le p+1$, then $\sum_{j=1}^{p+1} \lambda_j U_{i_j} = 0$. But, since the distribution of U is continuous in $[-\delta, \delta]$, when $|U_{i_j}| < \delta$ for each $1 \le j \le p+1$, $\sum_{j=1}^{p+1} \lambda_j U_{i_j} = 0$ happens with probability zero. So, when $|V_n(\hat\beta_n - \beta)| \le \delta$, with probability one,

$$a_n \sum_{j=1}^{n} |V_n^{-1} x_j| I_{U_j = x'_j(\hat\beta_n - \beta)} \le p \max_{1\le j\le n} a_n |V_n^{-1} x_j| I_{|U_j| \le |x'_j(\hat\beta_n - \beta)|}.$$

So, it suffices to prove that

$$\max_{1\le j\le n} a_n |V_n^{-1} x_j| I_{|U_j| \le M|V_n^{-1} x_j|} \xrightarrow{\mathrm{P}} 0,$$

for each $M < \infty$. Given $\tau > 0$

$$\mathrm{P}\left\{ \max_{1\le j\le n} a_n |V_n^{-1} x_j| I_{|U_j| \le M|V_n^{-1} x_j|} \ge \tau \right\}$$
$$\le \sum_{j=1}^{n} \mathrm{P}\{|U_j| \le M|V_n^{-1} x_j|\} I_{a_n |V_n^{-1} x_j| \ge \tau} \le c \sum_{j=1}^{n} |V_n^{-1} x_j| I_{a_n |V_n^{-1} x_j| \ge \tau}.$$

So, the claim follows. □

LEMMA 2.2. *Under conditions* (A.1)–(A.4),

$$\sup_{|t|\le T} a_n \left| \sum_{j=1}^{n} (E[\mathrm{sign}(U_j - x'_j V_n^{-1} t)] V_n^{-1} x_j + 2 f_U(0) t) \right| \to 0$$

as $n \to \infty$, *for each* $T < \infty$.

PROOF. By condition (A.1)

$$\sup_{|t|\le T} a_n \left| \sum_{j=1}^{n} (E[\mathrm{sign}(U_j - x'_j V_n^{-1} t)] V_n^{-1} x_j + 2 f_U(0) t) \right|$$
$$= \sup_{|t|\le T} a_n \left| \sum_{j=1}^{n} (V_n^{-1} x_j (1 - 2F(x'_j V_n^{-1} t) + 2 f_U(0) x'_j V_n^{-1} t)) \right|$$
$$\le c a_n \sum_{j=1}^{n} |V_n^{-1} x_j|^3 = c \left(\sum_{j=1}^{n} |V_n^{-1} x_j|^3 \right)^{1/2}$$
$$\le c \max_{1\le j\le n} |V_n^{-1} x_j|^{1/2} \left(\sum_{i=1}^{n} |V_n^{-1} x_i|^2 \right)^{1/2} \to 0.$$

Observe that

$$\sum_{j=1}^{n} |V_n^{-1}x_j|^2 = \text{trace}\left(\sum_{j=1}^{n} x_j' V_n^{-2} x_j\right)$$
$$= \text{trace}\left(\sum_{j=1}^{n} x_j x_j' V_n^{-2}\right) = \text{trace}(I_{p\times p}) = p,$$

where $I_{p\times p}$ is the identity matrix $p \times p$. □

We will need to bound tails probabilities for some stochastic processes. We will need the following theorem which is Theorem 12 in Arcones (1995*b*):

THEOREM 2.2. *Let $X_1, \ldots, X_n$ be independent r.v.'s with values in the measurable spaces $(S_1, \mathcal{S}_1), \ldots, (S_n, \mathcal{S}_n)$ respectively. Let T_0 be a parameter set. Let $f_j(\cdot, t) : S_j \to \mathbb{R}$ be a measurable function for each $t \in T_0$ and each $1 \le j \le n$. Let $F_j(x) = \sup_{t\in T_0} |f_j(x,t)|$. Let $a \ge 4$. Let q_0 be an integer. Suppose that for each $q \ge q_0$, there exists a function $\pi_q : T_0 \to T_0$ such that*

$$\pi_{q-1}(t) = \pi_{q-1}(s), \qquad \text{if} \quad \pi_q(t) = \pi_q(s),$$
$$\sum_{j=1}^{n} E[(2^q \Delta_{j,q}(X_j, \pi_q(t))) \wedge (2^{2q}\Delta_{j,q}^2(X_j, \pi_q(t)))] \le 1$$

where

$$\Delta_{j,q}(x, \pi_q(t)) = \sup_{s:\pi_q(s)=\pi_q(t)} |f_j(x,s) - f_j(x, \pi_q(t))|.$$

Then,

$$\mathrm{P}\left\{\sup_{t\in T_0}\left|\sum_{j=1}^{n} \epsilon_j (f(X_j,t) - f(X_j, \pi_{q_0}(t)))\right| \ge M\right\}$$
$$\le \sum_{j=1}^{n} \mathrm{P}\left\{F_j(X_j) \ge M(a+1)^{-1} 2^{-6-q_0} (\log N_{q_0+1})^{-1/2}\right.$$
$$\left.\cdot\left(\sum_{q=q_0+2}^{\infty} 2^{-q}(\log N_q)^{1/2}\right)^{-1}\right\}$$
$$+ \sum_{q=q_0}^{\infty} 4N_q^{-a},$$

for

$$M \ge 2^4(a+1)\sum_{q=q_0+2}^{\infty} 2^{-q}(\log N_q)^{1/2},$$

where N_q is the cardinality of $\pi_q(T_0)$.

Next lemma states the stochastic process in (2.4) is stochastically bounded.

LEMMA 2.3. *Assume conditions* (A.1)–(A.4). *Let*

$$W_n(t) := a_n \sum_{j=1}^{n} (\operatorname{sign}(U_j - x_j' V_n^{-1} t) - \operatorname{sign}(U_j) - E[\operatorname{sign}(U_j - x_j' V_n^{-1} t)]) V_n^{-1} x_j.$$

Then,

$$\sup_{|t| \le T} |W_n(t)| = O_{\mathrm{P}}(1)$$

for each $T < \infty$.

PROOF. Without loss of generality, we may assume that T is a positive integer. We have that

$$\begin{aligned} (2.5) \quad W_n(t) = 2a_n &\sum_{j=1}^{n} (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t} - E[I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t}]) V_n^{-1} x_j \\ &+ a_n \sum_{j=1}^{n} (I_{U_j = x_j' V_n^{-1} t} - \mathrm{P}\{U_j = x_j' V_n^{-1} t\}) V_n^{-1} x_j \\ &- a_n \sum_{j=1}^{n} I_{U_j = 0} V_n^{-1} x_j. \end{aligned}$$

There are $c_1, \delta > 0$ such that $|f_U(u)| \le c_1 |u|$, for $|u| \le \delta$, and $F_U(u)$ has a density in $|u| \le \delta$. Take n such that $\max_{1 \le j \le n} |V_n^{-1} x_j| \le \delta$. Take $M < \infty$ such that $2c_1 |T| < M^2$. We have that

$$\begin{aligned} &\sup_{|t| \le T} \mathrm{P}\left\{ a_n \left| \sum_{j=1}^{n} (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t} - E[I_{U_j \le 0} - I_{U_j < x_j' V_n^{-1} t}]) V_n^{-1} x_j \right| \ge M \right\} \\ &\le M^{-2} \sup_{|t| \le T} a_n^2 \sum_{j=1}^{n} |F_U(x_j' V_n^{-1} t) - F_U(0)| |V_n^{-1} x_j|^2 \\ &\le M^{-2} a_n^2 \sum_{j=1}^{n} c_1 |T| |x_j' V_n^{-1}|^3 = M^{-2} c_1 |T| \le 1/2 \end{aligned}$$

So, by Lemma 1.2.1 in Giné and Zinn (1986),

$$\begin{aligned} &\mathrm{P}\left\{ \sup_{|t| \le T} a_n \left| \sum_{j=1}^{n} (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t} - E[I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t}]) V_n^{-1} x_j \right| \ge 4M \right\} \\ &\le 2\mathrm{P}\left\{ \sup_{|t| \le M} a_n \left| \sum_{j=1}^{n} \epsilon_j (I_{U_j < 0} - I_{U_j \le x_j' V_n^{-1} t}) V_n^{-1} x_j \right| \ge M \right\}. \end{aligned}$$

Now, we apply Theorem 2.2 to $T_0 = \{t \in \mathbb{R}^p : |t| \le T\}$. Given $q \ge 1$, large enough, take a positive integer m such that $4p^{1/2}c_1 m^{-1} \le 2^{-2q} < 4p^{1/2}c_1(m-1)^{-1}$. Given $t = (t^{(1)}, \ldots, t^{(p)})'$, we define $\pi_q(t)$ as follows. If $-T + m^{-1}j_i \le t^{(i)} < -T + m^{-1}(j_i + 1)$, for some integers $j_1, \ldots, j_p$ with $0 \le j_1, \ldots, j_p \le 2Tm$, we define $\pi_q(t) = (-T + m^{-1}j_1, \ldots, -T + m^{-1}j_p)'$. Then, $|\pi_q(t) - t| \le m^{-1}p^{1/2}$ for each $t \in T_0$. We have that

$$\Delta_{j,q}(x, \pi_q(t)) = \sup_{s:\pi_q(s)=\pi_q(t)} |f_j(x,s) - f_j(x, \pi_q(t))|$$
$$\le 2a_n I_{|U_j - V_n^{-1}x_j\pi_q(t)| \le |V_n^{-1}x_j| m^{-1}p^{1/2}} |V_n^{-1}x_j|.$$

So,

$$\sum_{j=1}^n E[\Delta_{j,q}^2(X_j, \pi_q(t))] \le 4c_1 m^{-1}p^{1/2}a_n^2 \sum_{j=1}^n |V_n^{-1}x_j|^3 \le 2^{-2q}.$$

We also have that

$$N_q \le (2Tm+1)^p \le (8Tp^{1/2}c_1 2^{2q} + 2T + 1)^p.$$

Therefore, by Theorem 2.2,

$$\mathrm{P}\left\{ \sup_{|t|\le M} \left| a_n \sum_{j=1}^n \epsilon_i (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t}) V_n^{-1} x_j \right| \ge M \right\}$$
$$\le \sum_{j=1}^n \mathrm{P}\Bigg\{ |U_j| \le T|x_j' V_n^{-1}|,$$
$$a_n |V_n^{-1} x_j| \ge M(a+1)^{-1} 2^{-6-q_0} (\log N_{q_0+1})^{-1/2} \cdot \left(\sum_{q=q_0+2}^{\infty} 2^{-q} (\log N_q)^{1/2} \right)^{-1} \Bigg\}$$
$$+ \sum_{q=q_0}^{\infty} 4N_q^{-a}$$
$$\le \sum_{q=q_0}^{\infty} 4N_q^{-a} + \sum_{j=1}^n c|x_j' V_n^{-1}| I_{a_n |V_n^{-1} x_j| \ge c},$$

for

$$M \ge 2^4(a+1) \sum_{q=q_0+2}^{\infty} 2^{-q} (\log N_q)^{1/2}.$$

So,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathrm{P}\left\{ \sup_{|t|\le T} a_n \left| \sum_{j=1}^n (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t} - E[I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t}]) V_n^{-1} x_j \right| \ge M \right\} = 0.$$

It is easy to see that the proof of Lemma 2.1 gives that

$$\sup_{|t|\le T}\left|a_n\sum_{j=1}^{n} I_{U_j=x_j'V_n^{-1}t}V_n^{-1}x_j\right|\xrightarrow{\mathrm{P}}0.$$

Since $\mathrm{P}\{U=0\}=0$,

$$a_n\sum_{j=1}^{n} I_{U_j=0}V_n^{-1}x_j\xrightarrow{\mathrm{P}}0.$$

Therefore, the claim follows. □

THEOREM 2.3. *Under conditions* (A.1)–(A.4),

$$a_n\left|2f_U(0)V_n(\hat\beta_n-\beta)-\sum_{j=1}^{n}\mathrm{sign}(U_j)V_n^{-1}x_j\right|=O_{\mathrm{P}}(1). \tag{2.6}$$

PROOF. By Lemmas 2.2 and 2.3,

$$\sup_{|t|\le T}\left|2a_nf_U(0)t+a_n\sum_{j=1}^{n}(\mathrm{sign}(U_j-x_j'V_n^{-1}t)-\mathrm{sign}(U_j))V_n^{-1}x_j\right|=O_{\mathrm{P}}(1),$$

for each $T<\infty$. From this and (2.2),

$$2a_nf_U(0)V_n(\hat\beta_n-\beta)+a_n\sum_{j=1}^{n}(\mathrm{sign}(U_j-x_j'(\hat\beta_n-\beta))-\mathrm{sign}(U_j))V_n^{-1}x_j=O_{\mathrm{P}}(1).$$

From this and Lemma 2.1, (2.6) follows. □

Next, we will prove that under some extra conditions, the bound in Theorem 2.3 is attained. We will need the following lemma, which is a particular case of Theorem 1.5 in Arcones (1995*a*).

THEOREM 2.4. *Let* $\{k_n\}_{n=1}^{\infty}$ *be a sequence of positive integers converging to infinity. Let* $X_{n,1},\dots,X_{n,k_n}$ *be independent r.v.'s with values in the measurable spaces* $(S_{n,1},\mathcal{S}_{n,1}),\dots,(S_{n,k_n},\mathcal{S}_{n,k_n})$ *respectively. Let* T_0 *be a parameter set. Let* $f_{n,j}(\cdot,t):S_{n,j}\to\mathbb{R}$ *be a measurable function for each* $t\in T_0$ *and each* $1\le j\le k_n$. *Let* $b>0$. *Suppose that:*

(i) $\sum_{j=1}^{k_n}\mathrm{P}\{F_{n,j}(X_{n,j})\ge\eta\}\to0$, *for each* $\eta>0$, *where* $F_{n,j}(x)=\sup_{t\in T_0}|f_{n,j}(x,t)|$.

(ii) $\lim_{n\to\infty}\mathrm{Cov}(S_n(s,b),S_n(t,b))$ *exists for each* $s,t\in T_0$, *where*

$$S_n(t,b)=\sum_{j=1}^{k_n}f_{n,j}(X_{n,j},t)I_{|f_{n,j}(X_{n,j},t)|\le b}.$$

(iii) *There are positive integers* q_0 *and* n_0; *a function* $\pi_q : T_0 \to T_0$, *for each* $q \geq q_0$; *and a function* $\Delta_{n,j,q}(\cdot, \pi_q(t)) : S_{n,j} \to [0, \infty)$, *for each* $1 \leq j \leq k_n$, *each* $n \geq n_0$, *each* $q \geq q_0$ *and each* $t \in T_0$; *such that*

$$|f_{n,j}(x,t) - f_{n,j}(x, \pi_q t)| \leq \Delta_{n,j,q}(x, \pi_q(t)),$$

for each $x \in S_{n,j}$, *each* $1 \leq j \leq k_n$, *each* $n \geq n_0$, *each* $q \geq q_0$ *and each* $t \in T_0$;

$$\sup_{n \geq n_0} \sum_{j=1}^{k_n} E[(2^q \Delta_{n,j,q}(X_{n,j}, \pi_q(t))) \wedge (2^{2q} \Delta^2_{n,j,q}(X_{n,j}, \pi_q(t))) I_{\Delta_{n,j,q}(X_{n,j}, \pi_q(t)) \leq b}] \leq 1$$

and

$$\sum_{q=q_0}^{\infty} 2^{-q} (\log N_q)^{1/2} < \infty,$$

where N_q *is the cardinality of* $\pi_q(T_0)$.

Then,

$$\left\{ \sum_{j=1}^{n} (f_{n,j}(X_{n,j}, t) - E[f_{n,j}(X_{n,j}, t) I_{|f_{n,j}(X_{n,j},t)| \leq b}]) : t \in T_0 \right\}$$

converges weakly to a Gaussian process $\{Z(t) : t \in T_0\}$ *with mean zero and covariance given by*

$$E[Z(s)Z(t)] = \lim_{n \to \infty} \mathrm{Cov}(S_n(s,b), S_n(t,b)).$$

We also need the following lemma:

LEMMA 2.4. *Assume conditions* (A.1)–(A.4), *plus the following condition:*
(A.5) *There is a function* $B : \mathbb{R}^p \to M(p \times p)$, *where* $M(p \times p)$ *is the set of* $p \times p$ *matrices, such that* $a_n^2 \sum_{j=1}^n |x_j' V_n^{-1} t| \alpha' V_n^{-1} x_j x_j' V_n^{-1} \beta \to \alpha' B(t) \beta$, *for each* $\alpha, \beta, t \in \mathbb{R}^p$.

Then, $\{W_n(t) : |t| \leq T\}$ *converges weakly to the process* $\{W(t) : |t| \leq T\}$ *for each* $T < \infty$, *where*

$$W_n(t) := a_n \sum_{j=1}^{n} (\mathrm{sign}(U_j - x_j' V_n^{-1} t) - \mathrm{sign}(U_j) - E[\mathrm{sign}(U_j - x_j' V_n^{-1} t)]) V_n^{-1} x_j$$

and the limit process $\{W(t) : t \in \mathbb{R}^p\}$ *is a Gaussian process of* $\mathbb{R}^p$*-valued random vectors with mean zero and*

$$E[(\alpha' W(t))(\beta' W(s))] = 2^{-1} \alpha' (B(t) + B(s) - B(t-s)) \beta,$$

for each $s, t, \alpha, \beta \in \mathbb{R}^p$.

PROOF. By the estimations in Lemmas 2.1 and 2.2, it suffices to show that

$$\left\{ a_n \sum_{j=1}^{n} (I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t} - E[I_{U_j \le 0} - I_{U_j \le x_j' V_n^{-1} t}]) V_n^{-1} x_j : |t| \le T \right\}$$

converges weakly. This follows from Theorem 2.4. □

THEOREM 2.5. *Under the conditions* (A.1)–(A.5),

$$a_n(2f_U(0)V_n(\hat{\beta}_n - \beta) - \sum_{i=1}^{n} \operatorname{sign}(U_i) V_n^{-1} x_i) \xrightarrow{d} W(g), \tag{2.7}$$

where $\{W(t) : t \in \mathbb{R}^p\}$ *is the process in Lemma 2.4 and* g *is a* $\mathbb{R}^p$*-valued centered Gaussian random vector, independent of the process* $\{W(t) : t \in \mathbb{R}^p\}$, *with* $E[g(g)'] = (4f_U^2(0))^{-1} I_{p\times p}$, *where* $I_{p\times p}$ *is the* $p \times p$ *identity matrix.*

PROOF. It follows as Theorem 2.3, but using Lemma 2.4 instead of Lemma 2.3. □

Observe that if $\{x_j\}_{j=1}^{\infty}$ is a sequence of $\mathbb{R}^p$-valued i.i.d.r.v.'s with finite third absolute moment, then, with probability one,

$$\begin{aligned}
&n^{-1/2} V_n \to (E[x_1 x_1'])^{1/2} =: V, \\
&n^{-1/4} a_n \to (E[|x_1 V^{-1}|^3])^{-1/2}, \\
&n^{1/2} \sum_{j=1}^{n} |x_j V_n^{-1} t| \alpha' V_n^{-1} x_j x_j' V_n^{-1} \beta \to E[|x_1 V^{-1} t| \alpha' V^{-1} x_1 x_1' V^{-1} \beta],
\end{aligned}$$

for each $\alpha, \beta, t \in \mathbb{R}^p$, and

$$n^{-1/2} \sum_{j=1}^{n} |x_j| I_{|x_j| \ge \tau n^{1/4}} \to 0, \qquad \text{for each} \quad \tau > 0.$$

So, with probability one, conditions (A.2)–(A.5) are satisfied. To obtain the asymptotic normality of the LAD estimator with the rate $n^{1/2}$, it seems natural to require that $E[|x_1|^2] < \infty$. To get the second order representation with a_n of the order $n^{1/4}$, we are requiring that $E[|x_1|^3] < \infty$. Observe that, even if $E[|x_1|^3] = \infty$, by Theorem 2.3, we always have that

$$2F_U'(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j) V_n^{-1} x_j = O_P\left(n^{-3/4} \left(\sum_{j=1}^{n} |x_j|^3 \right)^{1/2} \right).$$

In this case, the order of $(n^{-3/4}(\sum_{j=1}^{n} |x_j|^3)^{1/2})$ is bigger than $n^{1/4}$.

Even if $E[|x_1|^2] < \infty$, it is possible to get that $V_n(\hat{\beta}_n - \beta)$ converges in distribution (see Pollard (1991), and Davis *et al.* (1992)). This happens when the sequence $\{x_i\}$ is in the domain of attraction of a stable distribution (Cauchy innovations for example). But, in this case it is not necessarily true that the limit distribution is normal and the rate of convergence is $n^{1/2}$.

The case $p = 1$ is the case of simple linear regression without a constant term. In this case, conditions (A.3) and (A.4) have a more simple statement. In particular, Theorem 2.5 in the case $p = 1$ says the following:

THEOREM 2.6. *Assume conditions* (A.1)–(A.2), $p = 1$, *and*

$$\text{(A.6)} \qquad \left(\sum_{i=1}^{n}|x_i|^2\right)^{-1/2}\sum_{j=1}^{n}|x_j|I_{|x_j|\geq\tau(\sum_{i=1}^n|x_i|^2)^{-1/4}(\sum_{i=1}^n|x_i|^3)^{1/2}}\to 0,$$

$$\text{as} \quad n\to\infty, \quad \text{for each } \tau > 0.$$

Then,

$$a_n 2f_U(0)V_n(\hat{\beta}_n - \beta) - a_n\sum_{i=1}^{n}\operatorname{sign}(U_i)V_n^{-1}x_i$$

converges in distribution to $W(g)$, *where* $\{W(t) : t \in \mathbb{R}\}$ *is a Brownian motion*, g *is a mean-zero Gaussian random variable*, *independent of* $\{W(t) : t \in \mathbb{R}\}$, *with* $E[g^2] = (4f_U^2(0))^{-1}$, $V_n = (\sum_{j=1}^n|x_j|^2)^{1/2}$ *and* $a_n = (\sum_{j=1}^n|x_j|^2)^{3/4}\cdot(\sum_{j-1}^n|x_j|^3)^{-1/2}$.

PROOF. It suffices to show that (A.6) implies (A.3)–(A.5). Take n such that

$$\left(\sum_{i=1}^{n}|x_i|^2\right)^{-1/2}\sum_{j=1}^{n}|x_j|I_{|x_j|\geq\tau(\sum_{i=1}^n|x_i|^2)^{-1/4}(\sum_{i=1}^n|x_i|^3)^{1/2}}\leq\tau.$$

If $|x_j| \geq \tau(\sum_{i=1}^n|x_i|^2)^{-1/4}(\sum_{i=1}^n|x_i|^3)^{1/2}$, then

$$\left(\sum_{i=1}^{n}|x_i|^2\right)^{-1/2}|x_j|\leq\tau.$$

If $|x_j| < \tau(\sum_{i=1}^n|x_i|^2)^{-1/4}(\sum_{i=1}^n|x_i|^3)^{1/2}$, using that $\sum_{i=1}^n|x_i|^3 \leq (\sum_{i=1}^n|x_i|^2)^{3/2}$, we have that then

$$\left(\sum_{i=1}^{n}|x_i|^2\right)^{-1/2}|x_j|\leq\tau.$$

(A.4) follows from (A.6) directly. (A.5) holds obviously for $m = 1$. □

Suppose that $p = 1$ and $x_j = c_1 j^b$, for each $j \geq 1$, where $c_1, b \in \mathbb{R}$. We have that (A.2)–(A.3) hold only if $b > -1/2$. We also have that (A.2) and (A.6) hold if $b > -1/2$. In this situation to get a second order expansion, we are not restricting the possible sequences.

If $\{x_j\}$ is a sequence of real-valued nondegenerate random variables with finite second moment, then (A.2) and (A.6) hold.

3. The strong Bahadur-Kiefer representation of LAD regression estimators

Here, we will study the almost sure asymptotic behavior of (1.5). Here, we will proceed in two different ways. First, we give a bound which holds under quite general conditions. Secondly, we give an exact bound which holds under more restricted conditions. The difference between the two approaches is to do or not to do a blocking argument. Next theorem gives a sharp bound in the second order expansion of the LAD estimator:

THEOREM 3.1. *Assume that for some sequence of real numbers* $\{b_n\}$:

(A.1) *U has a density $f_U(u)$ in a neighborhood of 0, $f_U(u)$ is continuous at 0, $f_U(0) \neq 0$ and*

$$\mathrm{P}\{U \leq x\} = 2^{-1} + x f_U(0) + O(x^2) \qquad as \quad x \to 0.$$

(A.2) *For n large enough, $\sum_{j=1}^n x_j x_j'$ has a positive definite square root V_n.*

(A.7) $(\log n)^{1/2} \max_{1\leq j\leq n} |V_n^{-1} x_j| \to 0$.

(A.8) *For each $\tau > 0$,*

$$\sum_{n=1}^{\infty} \sum_{i=1}^{n} (\log n)^{1/2} |x_i V_n^{-1}| I_{|x_i V_n^{-1}| \geq b_n} < \infty.$$

(A.9) $b_n \geq (\sum_{i=1}^n |x_i V_n^{-1}|^3)^{1/2} (\log n)^{-1/4}$.

Then, there exists a finite constant c such that

$$\limsup_{n\to\infty} b_n^{-1} (\log n)^{-1} \left| 2 f_U(0) V_n(\hat\beta_n - \beta) - \sum_{j=1}^n \mathrm{sign}(U_j) V_n^{-1} x_j \right| \leq c \qquad a.s.$$

Previous theorem applies to

$$b_n = \left((\log n)^{-1/4} \left(\sum_{j=1}^n |x_j' V_n^{-1}|^3 \right)^{1/2} \right) \vee d_n,$$

where $d_n = \max_{1\leq j\leq n} |V_n^{-1} x_j|$, giving that, with probability one,

$$\left| 2 f_U(0) V_n(\hat\beta_n - \beta) - \sum_{j=1}^n \mathrm{sign}(U_j) V_n^{-1} x_j \right| = O\left(\left((\log n)^{3/4} \left(\sum_{j=1}^n |x_j' V_n^{-1}|^3 \right)^{1/2} \right) \vee (d_n \log n) \right).$$

Under similar conditions, Babu (1989) showed that, with probability one,

$$\left| 2f_U(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j)V_n^{-1}x_j \right| = O(d_n^{1/2}(\log n)^{3/4}).$$

Observe that

$$(\log n)^{3/4} \left(\sum_{j=1}^{n} |x_j' V_n^{-1}|^3 \right)^{1/2} \leq d_n^{1/2}(\log n)^{3/4}$$

and, under condition (A.7),

$$d_n \log n \ll d_n^{1/2}(\log n)^{3/4}.$$

So, the presented bound improves that in Babu (1989).

If $n = 2km + 1$ and $x_j' = (1, -k + m^{-1}(j-1))$, for $1 \leq j \leq 2km + 1$. Then, $b_n = n^{-1/4}(\log n)^{-1/4}$. So,

$$\left| 2f_U(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j)V_n^{-1}x_j \right| = O(n^{-1/4}(\log n)^{3/4}).$$

In this example, d_n is of the order $n^{-1/2}$. So, the obtained bound is of the same order as the one obtained from Babu (1989).

If $p = 1$ and $x_j = j^{-b}$ for some $0 < b < 1/4$, Babu's bound gives that

$$\left| 2f_U(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j)V_n^{-1}x_j \right| = O(n^{-(1-2b)/4}(\log n)^{3/4}).$$

Theorem 3.1 gives that

$$\left| 2f_U(0)V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \operatorname{sign}(U_j)V_n^{-1}x_j \right| = O(n^{-1/4}(\log n)^{3/4}).$$

So, in this case Theorem 3.1 gives a sharper bound than that in Babu (1989).

To prove Theorem 3.1, we will need the following elemental observation (its proof is omitted).

LEMMA 3.1. *Let U be a r.v. Suppose that U has a density $f_U(u)$ in a neighborhood of 0 which is continuous at 0. Then*

$$E[|U - t|] = E[|U|] - tE[\operatorname{sign}(U)] + t^2 f_U(0) + o(t^2),$$

as $t \to 0$.

First, we prove an LIL for the LAD estimator. Next lemma says that certain process goes to zero almost surely.

LEMMA 3.2. *Assume the conditions* (A.1), (A.2) *and* (A.7). *Then,*

$$\sup_{|t|\le T} |\alpha_n(t)| \to 0 \qquad a.s.$$

for each $T < \infty$, *where*

$$\alpha_n(t) := (\log n)^{-1} \sum_{j=1}^{n} (g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) - E[g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t)])$$

and $g(x,t) = |x-t| - |x| + \text{sign}(x)t$.

PROOF. We claim that by the Talagrand inequality (see Theorem 6.17 in Ledoux and Talagrand (1991))

$$\text{(3.1)} \qquad \sum_{n=1}^{\infty} \mathrm{P}\left\{ \sup_{|t|\le T} \left| \sum_{j=1}^{n} (g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) - E[g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t)]) \right| \ge \eta \log n \right\} < \infty,$$

for each $\eta > 0$. Using that $|g(x,t)| \le |x| I_{|x|\le|t|}$, we get that

$$\begin{aligned} &\sup_{|t|\le T} E\left[\left| \sum_{j=1}^{n} (g(U_j, (\log n)^{1/2} z_1' V_n^{-1} t) - E[g(U_j, (\log n)^{1/2} z_1' V_n^{-1} t)]) \right|^2 \right] \\ &\le c \sup_{|t|\le T} \sum_{j=1}^{n} E[|g(U_j, (\log n)^{1/2} z_j' V_n^{-1} t)|^2] \\ &\le c(\log n) \sum_{j=1}^{n} |z_j' V_n^{-1}|^2 \mathrm{P}\{|U| < T(\log n)^{1/2} |z_j' V_n^{-1}|\} = o(\log n). \end{aligned}$$

Thus, we may symmetrize in (3.1). In fact, we have that

$$\begin{aligned} \text{(3.2)} \qquad &\mathrm{P}\left\{ \sup_{|t|\le T} \left| \sum_{j=1}^{n} (g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) - E[g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t)]) \right| \ge 4\eta \log n \right\} \\ &\le 4\mathrm{P}\left\{ \sup_{|t|\le T} \left| \sum_{j=1}^{n} \epsilon_j g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) \right| \ge \eta \log n \right\}. \end{aligned}$$

Now, we apply the Talagrand inequality to the last expression with

$$q = 4K_0, \qquad s = t = 2^{-3}\eta \log n \qquad \text{and} \qquad k = [\log n],$$

where K_0 is the universal constant in that inequality, to get that (3.2) is bounded by

$$2^{2-2\log n} + 4\mathrm{P}\left\{ \sum_{j=1}^{[\log n]} Y_j^* \geq 2^{-3}\eta \log n \right\} + 8\exp\left(-\frac{\eta^2(\log n)^2}{2^{16}K_0 M_n^2}\right), \tag{3.3}$$

where $\{Y_j^* : 1 \leq j \leq n\}$ are the order statistics of

$$\left\{ \sup_{|t|\leq T} |g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t)| : 1 \leq j \leq n \right\}$$

and

$$M_n = E\left[\sup_{|t|\leq T} \left| \sum_{j=1}^{n} \epsilon_j g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) \right| \right].$$

Since $|g(x,t)| \leq |t|$,

$$\sum_{j=1}^{[\log n]} Y_j^* \leq (\log n)^{3/2} |T| \max_{1\leq j\leq n} |x_j' V_n^{-1}| = o(\log n).$$

So, the middle term in the right hand side of (3.3) is zero. It is easy to see that Theorem 2.2 implies that

$$\sup_{|t|\leq T} (\log n)^{-1/2} \left| \sum_{j=1}^{n} \epsilon_j g(U_j, (\log n)^{1/2} x_j' V_n^{-1} t) \right| \xrightarrow{\mathrm{P}} 0.$$

This and the Hoffmann-Jørgensen inequality (see for example Proposition 6.8 in Ledoux and Talagrand (1991)) implies that $M_n = o((\log n)^{1/2})$. Therefore, (3.1) follows. □

From the previous lemma, we get the following:

LEMMA 3.3. *Assume the conditions* (A.1), (A.2) *and* (A.7). *Then*

$$(\log n)^{-1/2} \left(2f_U(0) V_n(\hat{\beta}_n - \beta) - \sum_{j=1}^{n} \mathrm{sign}(U_j) V_n^{-1} x_j \right) \to 0 \qquad a.s. \tag{3.4}$$

PROOF. Let $T < \infty$. By Lemma 3.1,

$$
\begin{aligned}
(3.5)\quad &\sup_{|t|\le T}\left|(\log n)^{-1}\left(\sum_{j=1}^{n}E[|U_j-(\log n)^{1/2}x_j'V_n^{-1}t|-|U_j|]\right)-f_U(0)|t|^2\right| \\
&=\sup_{|t|\le T}\Bigg|(\log n)^{-1}\sum_{j=1}^{n}(E[|U_j-(\log n)^{1/2}x_j'V_n^{-1}t|-|U_j| \\
&\qquad\qquad +\operatorname{sign}(U_j)(\log n)^{1/2}x_j'V_n^{-1}t \\
&\qquad\qquad -f_U(0)(\log n)(x_j'V_n^{-1}t)^2])\Bigg| \\
&\le \sup_{|t|\le T}(\log n)^{-1}\sum_{j=1}^{n}o((\log n)|x_j'V_n^{-1}t|^2)\to 0.
\end{aligned}
$$

Let $G_n(t) = (\log n)^{-1}\sum_{j=1}^{n}(|U_j - (\log n)^{1/2}x_j'V_n^{-1}t| - |U_j|)$ and let $\eta_n = (\log n)^{-1/2}(2f_U(0))^{-1}\sum_{j=1}^{n}\operatorname{sign}(U_j)V_n^{-1}x_j$. By Lemma 3.2 and (3.5),

$$G_n(t) = f_U(0)|t|^2 - 2f_U(0)\eta_n't + \Delta_n(t) = f_U(0)|t-\eta_n|^2 - f_U(0)|\eta_n|^2 + \Delta_n(t),$$

where $\sup_{|t|\le T}|\Delta_n(t)| \to 0$ a.s. for each $T<\infty$. Given $\delta > 0$,

$$\inf_{|t-\eta_n|\ge\delta}(G_n(t)-G_n(\eta_n)) = \inf_{|t-\eta_n|=\delta}(G_n(t)-G_n(\eta_n)) = f_U(0)\delta^2 + o(1) \qquad \text{a.s.}$$

So, (3.4) follows. □

Now, we can proceed as in Section 2.

LEMMA 3.4. *Assume the conditions* (A.1), (A.2), (A.7) *and* (A.8). *Then, with probability one,*

$$b_n^{-1}(\log n)^{-1}\left|\sum_{j=1}^{n}\operatorname{sign}(U_j - x_j'(\hat\beta_n-\beta))V_n^{-1}x_j\right| \to 0.$$

PROOF. By the arguments in the proof of Lemma 2.1, it suffices to prove that

$$\max_{1\le j\le n} b_n^{-1}(\log n)^{-1}|V_n^{-1}x_j|I_{|U_j|\le M(\log n)^{1/2}|V_n^{-1}x_j|} \to 0 \qquad \text{a.s.,}$$

for each $M<\infty$. Given $\tau > 0$

$$
\begin{aligned}
&\mathrm{P}\left\{\max_{1\le j\le n} b_n^{-1}(\log n)^{-1}|V_n^{-1}x_j|I_{|U_j|\le M(\log n)^{1/2}|V_n^{-1}x_j|}\ge\tau\right\} \\
&\le \sum_{j=1}^{n}\mathrm{P}\{|U_j| < M(\log n)^{1/2}|V_n^{-1}x_j|\}I_{b_n^{-1}(\log n)^{-1}|V_n^{-1}x_j|\ge\tau} \\
&\le c(\log n)^{1/2}\sum_{j=1}^{n}|V_n^{-1}x_j|I_{b_n^{-1}(\log n)^{-1}|V_n^{-1}x_j|\ge\tau}.
\end{aligned}
$$

Hence, the claim follows. □

The proof of next lemma is similar to that of Lemma 2.2 and it is omitted.

LEMMA 3.5. *Under conditions* (A.1), (A.2), (A.7) *and* (A.9),

$$\sup_{|t|\le T} b_n^{-1}\left|(\log n)^{-1}\sum_{j=1}^{n} E[\mathrm{sign}(U_j-(\log n)^{1/2}x_j'V_n^{-1}t)]V_n^{-1}x_j+2f_U(0)t\right|\to 0,$$

as $n\to\infty$, *for each* $T<\infty$.

LEMMA 3.6. *Assume the conditions* (A.1), (A.2), (A.7) *and* (A.8). *Then, for each* $T<\infty$, *there exists a finite constant* C_T *such that*

$$\limsup_{n\to\infty}\sup_{|t|\le T}|W_n(t)|<C_T \qquad a.s.,$$

where

$$W_n(t):=b_n^{-1}(\log n)^{-1}\sum_{j=1}^{n}(I_{U_j\le(\log n)^{1/2}x_j'V_n^{-1}t}-I_{U_j\le 0} - E[I_{U_j\le(\log n)^{1/2}x_j'V_n^{-1}t}-I_{U_j\le 0}])V_n^{-1}x_j.$$

PROOF. The proof follows from the arguments in Lemma 3.2. So, we only sketch the proof. We have to prove that

$$\sum_{n=1}^{\infty}\mathrm{P}\left\{\sup_{|t|\le T}\left|\sum_{j=1}^{n}\epsilon_j(I_{U_j\le(\log n)^{1/2}x_j'V_n^{-1}t}-I_{U_j\le 0})V_n^{-1}x_j\right| \ge \eta b_n(\log n)\right\}<\infty. \tag{3.6}$$

Now, we apply the Talagrand inequality to the last expression with

$$q=4K_0,\qquad s=t=2^{-3}\eta b_n\log n\qquad\text{and}\qquad k=[\log n],$$

where K_0. We get that (3.6) is bounded by

$$2^{-2\log n}+\mathrm{P}\left\{\sum_{j=1}^{[\log n]}Y_j^*\ge 2^{-3}\eta b_n\log n\right\}+2\exp\left(-\frac{\eta^2 b_n^2(\log n)^2}{2^{16}K_0M_n^2}\right),$$

where $\{Y_j^* : 1 \le j \le n\}$ are the order statistics of $\{I_{|U_j| \le (\log n)^{1/2} T |x_j' V_n^{-1}|} |V_n^{-1} x_j| : 1 \le j \le n\}$ and

$$M_n = E\left[\sup_{|t| \le T} \left| \sum_{j=1}^n \epsilon_j (I_{U_j \le (\log n)^{1/2} x_j' V_n^{-1} t} - I_{U_j \le 0}) V_n^{-1} x_j \right| \right].$$

We have that

$$\begin{aligned} \mathrm{P}&\left\{ \sum_{j=1}^{[\log n]} Y_j^* \ge 2^{-3} \eta b_n \log n \right\} \\ &\le \mathrm{P}\left\{ \max_{1 \le j \le n} I_{|U_j| \le (\log n)^{1/2} |x_j' V_n^{-1}|} |V_n^{-1} x_j| \ge 2^{-3} \eta b_n \right\} \\ &\le \sum_{j=1}^n c(\log n)^{1/2} |x_j' V_n^{-1}| I_{|V_n^{-1} x_j| \ge \tau b_n}. \end{aligned}$$

So, by condition (A.8),

$$\sum_{n=1}^{\infty} \mathrm{P}\left\{ \sum_{j=1}^{[\log n]} Y_j^* \ge 2^{-3} \eta b_n \log n \right\} < \infty.$$

It is easy to see that Theorem 2.2 implies that

$$\begin{aligned} &\sup_{|t| \le T} (\log n)^{-1/4} \left(\sum_{i=1}^n |x_i V_n^{-1}|^3 \right)^{-1/2} \\ &\qquad \cdot \left| \sum_{j=1}^n \epsilon_j (I_{U_j \le (\log n)^{1/2} x_j' V_n^{-1} t} - I_{U_j \le 0}) V_n^{-1} x_j \right| \xrightarrow{\mathrm{P}} 0. \end{aligned}$$

This limit and the Hoffmann-Jørgensen inequality (see for example Proposition 6.8 in Ledoux and Talagrand (1991)) imply that $M_n = o(b_n (\log n)^{1/2})$. Therefore, (3.6) follows. □

PROOF OF THEOREM 3.1. It follows from the previous lemmas, using the arguments in Theorem 2.3. □

Theorem 3.1 does not give the exact rate of convergence. Following the previous approach, but using a blocking argument, we will get the exact rate. We have the following theorem:

THEOREM 3.2. *Assume that:*

(A.1) *U has a density $f_U(u)$ in a neighborhood of 0, $f_U(u)$ is continuous at 0, $f_U(0) \ne 0$ and*

$$\mathrm{P}\{U \le x\} = 2^{-1} + x f_U(0) + O(x^2) \qquad \textit{as} \quad x \to 0.$$

(A.2) *For n large enough, $\sum_{j=1}^n x_j x_j'$ has a positive definite square root V_n.*

(A.5) *There is a function $B : \mathbb{R}^p \to M(p \times p)$, where $M(p \times p)$ is the set of $p \times p$ matrices, such that $a_n^2 \sum_{j=1}^n |x_j' V_n^{-1} t| \alpha' V_n^{-1} x_j x_j' V_n^{-1} \beta \to \alpha' B(t) \beta$, for each $\alpha, \beta, t \in \mathbb{R}^p$.*

$$\text{(A.10)} \qquad (\log\log n)^{1/2} \max_{1 \le j \le n} |V_n^{-1} x_j| \to 0.$$

$$\text{(A.11)} \qquad \lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{n \le m \le [(1+\delta)n]} |1 - V_n V_m^{-1}| = 0.$$

(A.12) *There are $r, \tau > 0$ such that*

$$\sum_{j=1}^{\infty} \sum_{i=1}^{[\gamma^j]} \exp(-r(\log j)^{3/4} a_{[\gamma^j]}^{-1} |V_{[\gamma^j]}^{-1} x_i|^{-1}) \cdot (\log j)^{1/2} |V_{[\gamma^j]}^{-1} x_i| I_{a_{[\gamma^j]} |V_{[\gamma^j]}^{-1} x_i| \ge \tau (\log j)^{-1/4}} < \infty,$$

for each $\gamma > 1$.

$$\text{(A.13)} \qquad \lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{n \le m \le [(1+\delta)n]} |1 - a_n a_m^{-1}| = 0.$$

Then, with probability one,

$$\left\{ a_n (2 \log\log n)^{-3/4} \left(2 f_U(0) V_n (\hat{\beta}_n - \beta) - \sum_{j=1}^n \text{sign}(U_j) V_n^{-1} x_j \right) \right\}_{n=1}^{\infty}$$

is relatively compact in $\mathbb{R}^p$ and its limit set is

$$\{x(v) : (x(t))_{t \in T^*} \in K_{T^*}\},$$

for T large enough, where $T^* = \{t \in \mathbb{R}^p : |t| \le T\} \cup \{\infty\}$ and K_{T^*} is the unit ball of the reproducing kernel Hilbert space of the Gaussian process $\{W(t) : t \in T^*\}$ which was defined in Theorem 2.4.

Under the conditions in Theorem 3.2, there exists a finite constant c such that

$$\limsup_{n \to \infty} a_n (2 \log\log n)^{-3/4} \left| 2 f_U(0) V_n (\hat{\beta}_n - \beta) - \sum_{j=1}^n \text{sign}(U_j) V_n^{-1} x_j \right| = c \quad \text{a.s.}$$

By the argument in the proof of Lemma 11 in Arcones (1996), if there exists a vector $v \in \mathbb{R}^p$ such that $B(\lambda v)$ is positive definite for each $\lambda > 0$ small enough, then $c > 0$. If $p = 1$, by the argument in Corollary 16 in Arcones (1996),

$$K_{T^*} = \left\{ (\alpha(t))_{t \in T^*} : \int_{-T}^{T} (\alpha'(t))^2 dt + 4(f_U(0))^2 (\alpha(\infty))^2 \le 1 \right\}$$

and

$$\limsup_{n\to\infty} a_n(2\log\log n)^{-3/4}\left|2f_U(0)V_n(\hat{\beta}_n-\beta)-\sum_{j=1}^n \operatorname{sign}(U_j)V_n^{-1}x_j\right|$$
$$=3^{-3/4}(f_U(0))^{-1/2} \qquad \text{a.s.}$$

Since $e^{-x^{-1}} \le x^{-1}$, for $x>0$, condition (A.12) holds if

$$\sum_{j=1}^{\infty}\sum_{i=1}^{[\gamma^j]}(\log j)^{-1/4}a_{[\gamma^j]}|V_{[\gamma^j]}^{-1}x_i|^2 I_{a_{[\gamma^j]}|V_{[\gamma^j]}^{-1}x_i|\ge\tau(\log j)^{-1/4}}<\infty, \tag{3.7}$$

for each $\gamma>1$.

If $\{x_j\}_{j=1}^{\infty}$ is a sequence of $\mathbb{R}^p$-valued i.i.d.r.v.'s with finite third absolute moment and $E[x_1x_1']$ is positive definite matrix, then with probability one, conditions (A.2), (A.5), (A.10)–(A.13) hold with

$$n^{-1/2}V_n \to (E[x_1x_1'])^{1/2} =: V,$$
$$n^{-1/4}a_n \to (E[|x_1V^{-1}|^3])^{-1/2}$$

and

$$a_n^2\sum_{j=1}^n |x_jV_n^{-1}t|\alpha'V_n^{-1}x_jx_j'V_n^{-1}\beta \to E[|x_1V^{-1}t|\alpha'V^{-1}x_1x_1'V^{-1}\beta],$$

for each $\alpha,\beta,t\in\mathbb{R}^p$. Observe that in this situation,

$$\sum_{j=1}^{\infty}\gamma^{j/4}(\log j)^{-1/4}E[|x_i|^2 I_{|x_i|\ge\tau\gamma^{j/4}(\log j)^{-1/4}}]<\infty.$$

If $2\le q<3$ and $\{x_j\}_{j=1}^{\infty}$ is a sequence of $\mathbb{R}^p$-valued i.i.d.r.v.'s with $E[|x_1|^q]<\infty$ and $E[x_1x_1']$ is positive definite matrix, then with probability one, conditions (A.2), (A.5), (A.10)–(A.13) hold with

$$n^{-1/2}V_n \to (E[x_1x_1'])^{1/2} =: V, \qquad a_n = n^{(3p-6)/(4p)}$$

and

$$a_n^2\sum_{j=1}^n |x_jV_n^{-1}t|\alpha'V_n^{-1}x_jx_j'V_n^{-1}\beta \to 0,$$

for each $\alpha,\beta,t\in\mathbb{R}^p$. This implies that

$$\limsup_{n\to\infty} n^{(6-3p)/(4p)}(2\log\log n)^{-3/4}$$
$$\cdot\left|2f_U(0)V_n(\hat{\beta}_n-\beta)-\sum_{j=1}^n \operatorname{sign}(U_j)V_n^{-1}x_j\right| = 0 \qquad \text{a.s.}$$

LEMMA 3.7. *Assume conditions*: (A.1), (A.2), (A.10) *and* (A.11). *Then*,

$$\sup_{|t|\le T} |\alpha_n(t)| \to 0 \qquad a.s. \tag{3.8}$$

for each $T < \infty$, *where*

$$\alpha_n(t) := (\log\log n)^{-1} \sum_{j=1}^{n} (g(U_j, (\log\log n)^{1/2} x_j' V_n^{-1} t) - E[g(U_j, (\log\log n)^{1/2} x_j' V_n^{-1} t)])$$

and $g(u,t) = |u-t| - |u| + \operatorname{sign}(u)t$.

PROOF. Take $\delta > 0$ such that

$$\limsup_{n\to\infty} \sup_{n\le m\le[(1+\delta)n]} |1 - V_n V_m^{-1}| \le 1/2.$$

Take $n_k := [(1+\delta)^k]$. Let $\eta > 0$. By symmetrization and the Lévy inequality,

$$\begin{aligned}
&\mathrm{P}\left\{\sup_{n_k\le n\le n_{k+1}} \sup_{|t|\le T} (\log\log n)^{-1} \left|\sum_{j=1}^{n} (g(U_j, (\log\log n)^{1/2} x_j' V_n^{-1} t) - E[g(U_j, (\log\log n)^{1/2} x_j' V_n^{-1} t)])\right| \ge 4\eta\right\} \\
&\le 4\mathrm{P}\left\{\sup_{n_k\le n\le n_{k+1}} \sup_{|t|\le T} \left|\sum_{j=1}^{n} \epsilon_j g(U_j, (\log\log n)^{1/2} x_j' V_n^{-1} t)\right| \ge \eta \log\log n_k\right\} \\
&\le 4\mathrm{P}\left\{\sup_{n_k\le n\le n_{k+1}} \sup_{|t|\le 2T} \left|\sum_{j=1}^{n} \epsilon_j g(U_j, (\log\log n_k)^{1/2} x_j' V_{n_k}^{-1} t)\right| \ge \eta \log\log n_k\right\} \\
&\le 8\mathrm{P}\left\{\sup_{|t|\le 2T} \left|\sum_{j=1}^{n_{k+1}} \epsilon_j g(U_j, (\log\log n_k)^{1/2} x_j' V_{n_k}^{-1} t)\right| \ge \eta \log\log n_k\right\}.
\end{aligned}$$

By the Talagrand inequality,

$$\sum_{k=1}^{\infty} \mathrm{P}\left\{\sup_{|t|\le 2T} \left|\sum_{j=1}^{n_{k+1}} \epsilon_j g(U_j, (\log\log n_k)^{1/2} x_j' V_{n_k}^{-1} t)\right| \ge \eta \log\log n_k\right\} < \infty.$$

Therefore, from this and the lemma of Borel-Cantelli, (3.8) follows. □

It is easy to see that the proof of Lemma 3.3 gives the following:

LEMMA 3.8. *Assume the conditions* (A.1), (A.2), (A.10) *and* (A.11). *Then*

$$(\log\log n)^{-1/2}\left(2f_U(0)V_n(\hat{\beta}_n-\beta)-\sum_{j=1}^{n}\operatorname{sign}(U_j)V_n^{-1}x_j\right)\to 0 \qquad a.s.$$

We also have that the arguments in Lemma 3.4 give that:

LEMMA 3.9. *Assume the conditions* (A.1), (A.2), (A.11)–(A.13). *Then, with probability one,*

$$a_n(2\log\log n)^{-3/4}\left|\sum_{j=1}^{n}\operatorname{sign}(U_j-x_j'(\hat{\beta}_n-\beta))V_n^{-1}x_j\right|\to 0.$$

LEMMA 3.10. *Assume* (A.1), (A.2), (A.5) *and* (A.10)–(A.13). *Let*

$$\begin{aligned}W_n(t) := {} & a_n(2\log\log n)^{-3/4}\\ & \times\sum_{i=1}^{n}(I_{U_i\le(2\log\log n)^{1/2}x_i'V_n^{-1}t}-I_{U_i\le 0}\\ & \qquad -E[I_{U_i\le(2\log\log n)^{1/2}x_i'V_n^{-1}t}-I_{U_i\le 0}])V_n^{-1}x_i.\end{aligned}$$

Then, with probability one, $\{W_n(t):|t|\le T\}$ *is relatively compact in* $l_\infty(T)$ *and its limit set is the unit ball* K_T *of the reproducing kernel Hilbert space of the Gaussian process* $\{W(t):|t|\le T\}$ *defined in Lemma* 2.8.

PROOF. We will proceed as in Theorem 1 in Arcones (1994). Here, $\{\epsilon_i\}_{i=1}^{\infty}$ will denote a sequence of i.i.d. Rademacher's r.v.'s, independent of sequence $\{U_i\}_{i=1}^{\infty}$. By a standard argument based on the Arzela-Ascoli theorem, we have to prove that

$$\limsup_{n\to\infty}\sum_{l=1}^{m}\lambda_l'W_n(t_l)=\sup_{x\in K_T}\sum_{l=1}^{m}\lambda_l'x(t_l) \tag{3.9}$$

for each $t_1,\dots,t_m\in\mathbb{R}^d$, $\lambda_1,\dots,\lambda_m\in\mathbb{R}^d$ and each m, and

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{\substack{|t_1|,|t_2|\le T\\ |t_1-t_2|\le\delta}}|W_n(t_1)-W_n(t_2)|=0 \qquad \text{a.s.}$$

Condition (3.9) follows exactly as the corresponding part in Theorem 1 in Arcones (1994). So, the proof of this part is omitted. Observe that

$$\sup_{x\in K_T}\sum_{l=1}^{m}\lambda_j'x(t_l)=\lim_{n\to\infty}\operatorname{Var}\left(\sum_{l=1}^{m}(2\log\log n)^{1/2}\lambda_l'W_n(t_l)\right).$$

Let $\gamma > 1$ and let $n_j = [\gamma^j]$. Next, we will see that

$$\limsup_{j\to\infty} \sup_{\substack{|t_1|,|t_2|\le T\\ |t_1-t_2|\le\delta}} |W_{n_j}(t_1) - W_{n_j}(t_2)| \le 2^{13}K_0^{1/2}T^{1/2}(1+f_U(0))^{1/2}\delta^{1/4} + 2^6(1+\tau+r)(\log\delta^{-1})^{-1} \quad \text{a.s.,} \tag{3.10}$$

for each $0 < \delta < 2^{-1}$, where K_0 is the constant in Theorem 6.17 in Ledoux and Talagrand (1991). By symmetrization,

$$\mathrm{P}\left\{ \sup_{\substack{|t_1|,|t_2|\le T\\ |t-s|\le\delta}} |W_{n_j}(t_1) - W_{n_j}(t_2)| \ge 2^{13}K_0^{1/2}T^{1/2}(1+f_U(0))^{1/2}\delta^{1/4} + 2^6(1+\tau+r)(\log\delta^{-1})^{-1} \right\}$$
$$\le 4\mathrm{P}\left\{ \sup_{\substack{|t|,|s|\le T\\ |t_1-t_2|\le\delta}} a_{n_j}\left|\sum_{i=1}^{n_k} \epsilon_i V_{n_j}^{-1} x_i (I_{U_i\le(2\log\log n)^{1/2}x_i'V_n^{-1}t_1} - I_{U_i\le(2\log\log n)^{1/2}x_i'V_n^{-1}t_2})\right| \ge (\log j)^{3/4}(65K_0^{1/2}T^{1/2}(1+f_U(0))^{1/2}\delta^{1/4} + 2^5(1+\tau+r)(\log\delta^{-1})^{-1}) \right\} \tag{3.11}$$

for j large. Let

$$\Delta_j = \sum_{i=1}^{[\gamma^j]} \exp(-r(\log j)^{3/4} a_{[\gamma^j]}^{-1} |V_{[\gamma^j]}^{-1} x_i|^{-1}) \cdot (\log j)^{1/2} |V_{[\gamma^j]}^{-1} x_i| I_{a_{[\gamma^j]}|V_{[\gamma^j]}^{-1}x_i|\ge\tau(\log j)^{-1/4}}.$$

If $\Delta_j \le j^{-3}$, we apply the Talagrand inequality to the last expression with

$$q = K_0\delta^{-1/2}, \qquad k = 4(\log j)(\log\delta^{-1})^{-1},$$
$$t = 2^6K_0^{1/2}T^{1/2}(1+f_U(0))^{1/2}\delta^{1/4}(\log j)^{3/4}$$

and

$$s = 2^4(1+\tau+r)(\log j)^{3/4}(\log\delta^{-1})^{-1}.$$

Observe that

$$(K_0/q)^k = j^{-2},$$

$$a_{n_j} E\left[\sup_{\substack{|t_1|,|t_2|\le T\\ |t_1-t_2|\le\delta}} \left| \sum_{i=1}^{n_j} \epsilon_i V_{n_j}^{-1} x_i (I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_n^{-1} t_1} - I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_n^{-1} t_2}) \right| \right] = O((\log j)^{1/4})$$

and

$$a_{n_j}^2 \sup_{\substack{|t_1|,|t_2|\le T\\ |t_1-t_2|\le\delta}} \sum_{i=1}^{n_j} |x_i' V_n^{-1}|^2 E[(I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_n^{-1} t_1} - I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_n^{-1} t_2})^2] + (8Ms/k)$$
$$\le 4T\delta(1+f_U(0))(\log j)^{1/2},$$

for j large. We also have that

$$\sum_{j=1}^{k} Y_{n_j,j}^* \le 8(\log j) a_{n_j} \sup_{1\le i\le n_j} I_{|U_i|\le (2\log\log n)^{1/2} |x_i' V_n^{-1}| T}$$

where $\{Y_{n_j,i}^*\}_{i=1}^{n_j}$ is a nonincreasing rearrangement of

$$a_{n_j} \sup_{\substack{|t_1|,|t_2|\le T\\ |t-s|\le\delta}} |I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t_1} - I_{U_i\le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t_2}| |V_n^{-1} x_i|, \quad 1\le i\le n_j.$$

So,

$$\mathrm{P}\left\{ \sum_{j=1}^{k} Y_{n_j,j}^* \ge s \right\}$$
$$\le \mathrm{P}\left\{ a_{n_j} \sup_{1\le i\le n_j} I_{|U_i|\le (2\log\log n)^{1/2} |x_i' V_n^{-1}| T} \ge 2(1+\tau+r)(\log j)^{-1/4} \right\}$$
$$\le \sum_{i=1}^{n_j} \mathrm{P}\{ a_{n_j} |x_i' V_n^{-1}| I_{|U_i|\le (2\log\log n)^{1/2} |x_i' V_n^{-1}| T} \ge 2(\tau+r)(\log j)^{-1/4} \}$$
$$\le 2T(f_U(0)+1) \sum_{i=1}^{n_j} (\log j)^{1/2} |x_i' V_{n_j}^{-1}| I_{a_{n_j} |x_i' V_{n_j}^{-1}| \ge 2(\tau+r)(\log j)^{-1/4}}$$
$$\le 2T(f_U(0)+1) j^{1/2} \sum_{i=1}^{n_j} \exp(-r(\log j)^{3/4} a_{n_j}^{-1} |x_i' V_{n_j}^{-1}|^{-1})$$
$$\cdot (\log j)^{1/2} |x_i' V_{n_j}^{-1}| I_{a_{n_j} |x_i' V_{n_j}^{-1}| \ge \tau(\log j)^{-1/4}}$$
$$\le 2T(f_U(0)+1) j^{-5/2}.$$

From these estimations, we get that (3.11) is bounded by

$$16(1+T(f_U(0)+1))j^{-2}, \tag{3.12}$$

for j large.

If $\Delta_j > j^{-3}$, we apply the Talagrand inequality to the last expression with

$$q = K_0\delta^{-1/2}, \qquad k = 4(\log\Delta_j^{-1})(\log\delta^{-1})^{-1},$$
$$t = 2^6 K_0^{1/2} T^{1/2}(1+f_U(0))^{1/2}\delta^{1/4}(\log j)^{3/4}$$

and

$$s = 2^4(1+\tau+r)(\log j)^{3/4}(\log\delta^{-1})^{-1}.$$

Observe that

$$(K_0/q)^k = \Delta_j,$$
$$a_{n_j}E\left[\sup_{\substack{|t_1|,|t_2|\le T\\ |t_1-t_2|\le\delta}}\left|\sum_{i=1}^{n_j}\epsilon_i V_{n_j}^{-1}x_i(I_{U_i\le(2\log\log n_j)^{1/2}x_i'V_n^{-1}t_1} - I_{U_i\le(2\log\log n_j)^{1/2}x_i'V_n^{-1}t_2})\right|\right] = O((\log j)^{1/4})$$

and

$$a_{n_j}^2\sup_{\substack{|t_1|,|t_2|\le T\\ |t-s|\le\delta}}\sum_{i=1}^{n_j}|x_i'V_n^{-1}|^2E[(I_{U_i\le(2\log\log n_j)^{1/2}x_i'V_n^{-1}t_1} - I_{U_i\le(2\log\log n_j)^{1/2}x_i'V_n^{-1}t_2})^2] + (8Ms/k)$$
$$\le 4T\delta(1+f_U(0))(\log j)^{1/2} + O(1)(\log j)(\log\Delta_j^{-1}),$$

for j large. We also have that

$$\sum_{j=1}^{k} Y_{n_j,j}^* \le Y_{n_j,1}^* + kY_{n_j,2}^*$$

where $\{Y_{n_j,i}^*\}_{i=1}^{n_j}$ is as before. So,

$$\mathrm{P}\left\{\sum_{j=1}^{k} Y_{n_j,j}^* \ge s\right\} \le \mathrm{P}\{Y_{n_j,1}^* \ge 2^{-1}s\} + \mathrm{P}\{Y_{n_j,2}^* \ge 2^{-1}k^{-1}s\}.$$

We have that

$$\mathrm{P}\{Y_{n_j,1}^* \ge 2^{-1}s\}$$

$$\le \sum_{i=1}^{n_j} \mathrm{P}\{a_{n_j}|V_{n_j}^{-1}x_i|I_{|U_i|\le(2\log\log n_j)^{1/2}|V_{n_j}^{-1}x_i|T} \ge 4r(\log j)^{3/4}(\log\delta^{-1})^{-1}\}$$

$$\le 2T(f_U(0)+1)\sum_{i=1}^{n_j}(2\log j)^{1/2}|x_i'V_{n_j}^{-1}|I_{a_{n_j}|V_{n_j}^{-1}x_i|\ge 4r(\log j)^{3/4}(\log\delta^{-1})^{-1}}$$

$$\le 2T(f_U(0)+1)\sum_{i=1}^{n_j}\exp(-r(\log j)^{3/4}a_{[\gamma^j]}^{-1}|V_{[\gamma^j]}^{-1}x_i|^{-1})$$
$$\cdot(\log j)^{1/2}|x_i'V_{n_j}^{-1}|\delta^{1/4}I_{a_{n_j}|V_{n_j}^{-1}x_i|\ge\tau(\log j)^{-1/4}}$$
$$\le 2T(f_U(0)+1)\delta^{1/4}\Delta_j$$

and

$$\mathrm{P}\{Y_{n_j,2}^* \ge 2^{-1}k^{-1}s\}$$
$$\le \left(\sum_{i=1}^{n_j}\mathrm{P}\{a_{n_j}|V_{n_j}^{-1}x_i|I_{|U_i|\le(2\log\log n_j)^{1/2}|V_{n_j}^{-1}x_i|T} \ge 2(\tau+r)(\log j)^{3/4}(\log\Delta_j^{-1})^{-1}\}\right)^2$$
$$\le \left(\sum_{i=1}^{n_j}2T(f_U(0)+1)\cdot(\log j)^{1/2}|V_{n_j}^{-1}x_i|I_{a_{n_j}|V_{n_j}^{-1}x_i|\ge 2(\tau+r)(\log j)^{3/4}(\log\Delta_j^{-1})^{-1}}\right)^2$$
$$\le 2^3T^2(f_U(0)+1)^2$$
$$\times\left(\sum_{i=1}^{n_j}(\log j)^{1/2}|V_{n_j}^{-1}x_i|\exp(-r(\log j)^{3/4}a_{n_j}^{-1}|V_{n_j}^{-1}x_i|^{-1})\cdot\Delta_j^{1/2}I_{a_{n_j}|V_{n_j}^{-1}x_i|\ge\tau(\log j)^{-1/4}}\right)^2$$
$$\le 2^4T^2(f_U(0)+1)^2\Delta_j.$$

From all these estimations, we deduce that (3.11) is bounded by

$$2^8T^2(f_U(0)+1)^2\Delta_j + 8j^{-2}, \tag{3.13}$$

for j large. Therefore, (3.10) follows.

By (3.10), to end the proof, it suffices to show that

$$\lim_{\gamma\to 1+}\limsup_{n\to\infty}\sup_{|t|\le T}\sup_{\gamma^j<n\le\gamma^{j+1}}|W_n(t)-W_{[\gamma^j]}(t)|\to 0 \quad \text{a.s.} \tag{3.14}$$

Let $\gamma > 1$ and let $n_j = [\gamma^j]$. We have that, for $n_j < n \le n_{j+1}$

(3.15) $W_n(t) - W_{n_j}(s)$

$$
\begin{aligned}
&= a_n(2\log\log n)^{-3/4} \\
&\quad \cdot \sum_{i=1}^{n} (I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le 0} \\
&\qquad - E[I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le 0}]) V_n^{-1} x_i \\
&\quad - a_{n_j}(2\log\log n_j)^{-3/4} \\
&\quad \cdot \sum_{i=1}^{n_j} (I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t} - I_{U_i \le 0} \\
&\qquad - E[I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t} - I_{U_i \le 0}]) V_{n_j}^{-1} x_i \\
&= a_n(2\log\log n)^{-3/4} \\
&\quad \cdot \sum_{i=1}^{n_j} (I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le 0} \\
&\qquad - E[I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le 0}]) (V_n^{-1} x_i - V_{n_j}^{-1} x_i) \\
&\quad + (a_n(2\log\log n)^{-3/4} - a_{n_j}(2\log\log n_j)^{-3/4}) \\
&\quad \cdot \sum_{i=1}^{n_j} (I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t} - I_{U_i \le 0} \\
&\qquad - E[I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t} - I_{U_i < 0}]) V_{n_j}^{-1} x_i \\
&\quad + a_n(2\log\log n)^{-3/4} \\
&\quad \cdot \sum_{i=n_j+1}^{n} (I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le 0} \\
&\qquad - E[I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i < 0}]) V_n^{-1} x_i \\
&\quad + a_n(2\log\log n)^{-3/4} \\
&\quad \cdot \sum_{i=1}^{n_j} (I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} - I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_n^{-1} t} \\
&\qquad - E[I_{U_i \le (2\log\log n)^{1/2} x_i' V_n^{-1} t} \\
&\qquad - I_{U_i \le (2\log\log n_j)^{1/2} x_i' V_{n_j}^{-1} t}]) V_{n_j}^{-1} x_i \\
&=: I_n(t) + II_n(t) + III_n(t) + IV_n(t).
\end{aligned}
$$

Since $a_n(2\log\log n)^{-3/4} \le 2a_{n_j}(2\log\log n_j)^{-3/4}$ and $|V_{n_j} V_n^{-1}| \le 2$, for j large and $n_j < n \le n_{j+1}$,

$$
\sup_{n_j < n \le n_{j+1}} \sup_{|t| \le T} |I_n(t)| \le 2 \sup_{n_j < n \le n_{j+1}} |V_{n_j} V_n^{-1} - I| \sup_{|t| \le 2T} |W_n(t)|. \tag{3.16}
$$

whose $\lim_{\gamma \to 1+} \limsup_{j \to \infty}$ is zero. By (A.13),

$$
\sup_{n_j < n \le n_{j+1}} \sup_{|t| \le T} |II_n(t)| \tag{3.17}
$$

$$\leq \sup_{n_j<n\leq n_{j+1}} |a_n(2\log\log n)^{-3/4}a_{n_j}^{-1}(2\log\log n_j)^{3/4}-1| \cdot \sup_{|t|\leq 2T}|W_n(t)|,$$

which satisfies the same property.

By the Talagrand inequality (and arguments similar to the used in other parts)

$$\lim_{\gamma\to 1+}\limsup_{j\to\infty}\sup_{n_j<n\leq n_{j+1}}\sup_{|t|\leq T}|III_n(t)|=0 \quad \text{a.s.} \tag{3.18}$$

For j large and $n_j < n \leq n_{j+1}$, $a_n(2\log\log n)^{-3/4} \leq 2a_{n_j}(2\log\log n_j)^{-3/4}$ and $|V_{n_j}V_n^{-1}-1|\leq\delta$. So,

$$\lim_{\gamma\to 1+}\limsup_{j\to\infty}\sup_{n_j<n\leq n_{j+1}}\sup_{|t|\leq T}|IV_n(t)| \leq \lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{\substack{|t|,|s|<T\\|t-s|\leq\delta}}|W_n(t)-W_n(s)|=0 \quad \text{a.s.} \tag{3.19}$$

By (3.15)–(3.19), (3.14) follows. □

PROOF OF THEOREM 3.2. It follows from the previous lemmas, using the arguments in Theorem 2.5. □

REFERENCES

Arcones, M. A. (1994). Some strong limit theorems for M-estimators, *Stochastic Process. Appl.*, **53**, 241–268.

Arcones, M. A. (1995a). Weak convergence for the row sums of a triangular array of empirical processes under bracketing conditions, *Stochastic Process. Appl.* (submitted).

Arcones, M. A. (1995b). M-estimators converging to a stable limit (preprint).

Arcones, M. A. (1996). The Bahadur-Kiefer representation of L_p regression estimators, *Econom. Theory*, **12**, 257–283.

Babu, G. J. (1989). Some strong representations for LAD estimators in linear models, *Probab. Theory Related Fields*, **83**, 547–558.

Bai, Z. D., Chen, X. R., Wu, Y. and Zhao, L. C. (1989). Asymptotic normality of minimum L_1-norm estimates in linear models, *Gujarat Statist. Rev.*, **16**, 13–32.

Bassett, G. and Koenker, R. (1978). Asymptotic theory of least absolute error regression, *J. Amer. Statist. Assoc.*, **73**, 618–622.

Bloomfield, P. and Steiger, W. L. (1983). *Least Absolute Deviations. Theory, Applications, and Algorithms*, Birkhäuser, Boston.

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed sequential confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457–462.

Davis, R. A., Knight, K. and Liu, J. (1992). M-estimation for autoregression with infinite variance, *Stochastic Process. Appl.*, **40**, 145–180.

Draper, N. R. and Smith, H. (1981). *Applied Regression Analysis*, Wiley, New York.

Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes, *Lecture Notes in Math.*, **1221**, 50–112, Springer, New York.

He, X. and Shao, Q–M. (1996). A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs, *Ann. Statist.*, **24**, 2608–2630.

Koenker, R. and Portnoy, S. (1987). L-estimation for linear models, *J. Amer. Statist. Assoc.*, **82**, 851–857.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*, Springer, New York.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators, *Econom. Theory*, **7**, 186–199.

Rao, C. R. and Toutenburg, H. (1995). *Linear Models, Least Squares and Alternatives*, Springer, New York.

Rao, C. R. and Zhao, L. C. (1992). Linear representations of M-estimates in linear models, *Canad. J. Statist.*, **20**, 359–368.

Ruppert, D. and Carroll, R. J. (1980). Trimmed least squares estimation in the linear model, *J. Amer. Statist. Assoc.*, **75**, 828–838.

Stigler, J. M. (1986). *The History of Statistics*, Harvard University Press, Cambridge, Massachusetts.