Strong Representations for LAD Estimators in Linear Models

Gutti Jogesh Babu
Department of Statistics, The Pennsylvania State University, 219 Pond Laboratory, University Park, PA 16802, USA

Summary. Consider the standard linear model \( y_i = z_i \beta + e_i, \ i = 1, 2, \ldots, n, \) where \( z_i \) denotes the \( i \)th row of an \( n \times p \) design matrix, \( \beta \in \mathbb{R}^p \) is an unknown parameter to be estimated and \( e_i \) are independent random variables with a common distribution function \( F \). The least absolute deviation (LAD) estimate \( \hat{\beta} \) of \( \beta \) is defined as any solution of the minimization problem

\[
\sum_{i=1}^{n} |y_i - z_i \hat{\beta}| = \inf \left\{ \sum_{i=1}^{n} |y_i - z_i \beta| : \beta \in \mathbb{R}^p \right\}.
\]

In this paper Bahadur type representations are obtained for \( \hat{\beta} \) under very mild conditions on \( F \) near zero and on \( z_i, \ i = 1, \ldots, n \). These results are extended to the case, when \( \{e_i\} \) is a mixing sequence. In particular the results are applicable when the residuals \( e_i \) form a simple autoregressive process.

1. Introduction

Consider the linear model

\[
y_i = z_i \beta + e_i, \quad i = 1, \ldots, n,
\]

where \( z_i \) denotes the \( i \)th row of an \( n \times p \) design matrix, \( \beta \in \mathbb{R}^p \) is an unknown regression parameter and \( e_i \) are independent random variables with a common distribution function \( F \). The search for robust procedures to estimate \( \beta \), has generated considerable interest in recent years, in developing statistical methods based on least absolute deviations (LAD) using the \( L_1 \)-norm rather than the classical least squares (LS) based on \( L_2 \)-norm (see Koenker and Bassett 1978; Bassett and Koenker 1982; Amemiya 1982; Bloomfield and Steiger 1983; Koenker and Portnoy 1987, and Bai et al. 1987).

It is of interest to note that estimation by the LAD method, that minimizes

\[
\sum_{i=1}^{n} |y_i - z_i \beta|,
\]

was considered by Gauss (1809) and Edgeworth (1887) but was
abandoned in favour of the LS method, in view of computational difficulties. The advent of modern electronic computers and increased interest in robust procedures have renewed the interest in this method.

The special case of (1), when $z_i = 1$ and $p = 1$, (the location parameter case) has been studied extensively during the last few decades. If $F(0) = 1/2$, then $\beta$ is the median of $y_1$. Even though the $L_1$-estimator $\bar{\beta}$ of $\beta$ (the sample median of $y_1, \ldots, y_n$) is a non-linear estimator of $\beta$, Bahadur (1966) has shown, under very mild conditions of $F$, that $\bar{\beta} - \beta$ can be written as

$$
\frac{c}{n} \sum_{i=1}^{n} (I(Y_i \leq 0) - \frac{1}{2}) + r_n,
$$

where $c$ is a positive constant depending only on $F$.

$$
r_n = O(n^{-3/4} \log^{1/2} \log \log n)^{1/4}
$$

with probability 1 and $I(A)$ denotes the indicator of $A$. This result is extended in this paper to the $L_1$-estimator of $\beta$ for the model (1) under mild restrictions on $\max(\|z_i\|: 1 \leq i \leq n)$, where $\| \cdot \|$ denotes the Euclidean norm.

These results are further extended to the case of dependent residuals $\epsilon_i$; that is, when the residuals $\{\epsilon_i\}$ form a mixing process. It is well known that certain linear processes like autoregressive processes are strongly mixing (see Gorodetskii 1977). These processes are described in Sect. 4. In the case of i.i.d. residuals, Koenker and Portnoy (1987) have shown that $P(|r_n| > cn^{-3/4} \log n) \to 0$, as $n \to \infty$, for some $c > 0$, under strong assumptions on $z_i$. Their conditions include

$$
\sum_{i=1}^{n} \|z_i\|^3 = O(n) \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} z_i'z_i = Q + Q_n,
$$

where the largest eigenvalue of $Q_n$ is $O(n^{-1/4})$ and $W'$ denotes the transpose of the matrix $W$.

2. Main Result

We start with some notations and assumptions.

Let $S_n$ denote $\sum_{i=1}^{n} z_i'z_i$. Suppose for some $n_0$, $S_{n_0}$ is nonsingular. It is clear that, in such a case $S_n$ is positive definite for all $n \geq n_0$. By considering

$$
x_{in} = z_i S_n^{-1/2}
$$

and

$$
\beta_0 = S_n^{1/2} \beta
$$

we get the equivalent model

$$
y_i = x_{in} \beta_0 + \epsilon_i, \quad i = 1, 2, \ldots, n
$$

(2)
with

\[ \sum_{i=1}^{n} x_{in} x_{in} = I_p, \]  

(3)

the \( p \times p \) identity vector. Consideration of this equivalent transformed model simplifies the presentation. We note that \( x_{1n}, \ldots, x_{nn}, \) \( n \geq n_0 \) is a triangular array of design vectors. So as \( n \) increases, not only the design matrix changes but also the number of observations \( y_i \) increases. Without loss of generality, we assume that the true parameter \( \beta_0 \) is the zero vector. So \( y_i = e_i \) under this null assumption.

Throughout this paper we assume that \( F \) has a derivative \( f \) in a neighbourhood of 0, \( f(0) > 0, \) \( F(0) = 1/2 \) and for some \( c > 0, \)

\[ |f(y) - f(0)| \leq c|y|^{1/2} \]

holds for all \( y \) in a neighbourhood of 0.

Let

\[ d_n = \max \{ \| x_{in} \| : 1 \leq i \leq n \} = \max \{ (z_i S_n^{-1} x_i')^{1/2} : 1 \leq i \leq n \}. \]

Note that

\[ \sum_{i=1}^{n} x_{in} x_{in} = \sum_{i=1}^{n} \text{trace}(x_{in} x_{in}') = \sum_{i=1}^{n} \text{trace}(x_{in} x_{in}') = \text{trace} \left( \sum_{i=1}^{n} x_{in} x_{in} \right) = p. \]

(4)

As a consequence

\[ d_n \geq (p/n)^{1/2}. \]

(5)

Suppose \( \beta \) satisfies

\[ \sum_{i=1}^{n} |y_i - x_{in} \beta| = \inf \left\{ \sum_{i=1}^{n} |y_i - x_{in} \beta| : \beta \in \mathbb{R}^p \right\} = \inf \left\{ \sum_{i=1}^{n} |y_i - z_i S_n^{-1/2} \beta| : \beta \in \mathbb{R}^p \right\}. \]

(6)

Then it is clear from (6) that for model (1),

\[ \bar{\beta} = S_n^{-1/2} \beta \]

is a solution of

\[ \sum_{i=1}^{n} |y_i - z_i \bar{\beta}| = \inf \left\{ \sum_{i=1}^{n} |y_i - z_i \beta| : \beta \in \mathbb{R}^p \right\}. \]

Further we have

\[ x_{in} \bar{\beta} = z_i \bar{\beta}. \]

It is possible to have more than one solution for (6) like in the location case when \( p = 1 \). Let \( \bar{\beta} \) be any such solution. Then we have
Theorem 1. If \( d_n(\log n)^{1/2} \to 0 \), then with probability 1,

\[
2 f(0) \hat{\beta} = \sum_{i=1}^{n} x_{in} \text{sign} e_i + 0(d_n^{1/2}(\log n)^{3/4}).
\] (7)

This theorem leads to the following corollary for the original model (1).

Corollary. If \( d_n(\log n)^{1/2} \to 0 \), then under the null hypothesis that \( \beta = \beta_1 \), we have with probability 1,

\[
2 f(0) S_n^{1/2}(\hat{\beta} - \beta_1) = \sum_{i=1}^{n} z_i S_n^{-1/2} \text{sign}(y_i - z_i \beta_1) + 0(d_n^{1/2}(\log n)^{3/4})
\]

\[
= \sum_{i=1}^{n} z_i S_n^{-1/2} \text{sign} e_i + 0(d_n^{1/2}(\log n)^{3/4}).
\]

Instead of \( L_1 \)-norm \( |y| \), if \( \rho_\theta(y) = |y| + y I(y < 0), (0 < \theta < 1) \) is used, then the estimate \( \hat{\beta}_0 \) obtained by minimizing

\[
\sum_{i=1}^{n} \rho_\theta(y_i - x_{in} \beta_0)
\]

also has strong (almost sure) representation, similar to (7). The proof of this fact is similar to that of Theorem 1 and so is omitted. Note that in the location parameter case \( \hat{\beta}_0 n^{-1/2} \) corresponds to a \( \theta \)th sample quantile.

If \( m_n \) denotes the sample quantile of \( e_1, \ldots, e_n \), it follows from Theorem 1 that

\[
2 \sqrt{n} f(0) m_n - n^{-1/2} \sum_{i=1}^{n} \text{sign} e_i + 0(n^{-1/4}(\log n)^{3/4})
\]

with probability 1. Except for the power of \( \log n \), the error term matches the error term in Bahadur's (1966) representation of quantiles.

3. Proof of Theorem 1

Babu and Singh (1978) used exponential inequalities for sums of bounded random variables to obtain Bahadur-Kiefer type representations for mixing variables. We essentially adopt the ideas of that paper to derive Theorem 1. The proof is divided into several lemmas.
Lemma 1. Let $Z_1, \ldots, Z_n$ be independent random variables with mean zero and let $|Z_i| \leq b$ for some $b > 0$. Let $V \geq \sum_{i=1}^{n} E(Z_i^2)$. Then for all $0 < s < 1$ and $0 \leq a \leq V/sb$

$$P\left(\sum_{i=1}^{n} Z_i > a\right) \leq \exp(-a^2 s(1-s)/V).$$

Proof. Since $e^y \leq 1 + y + y^2$ for $|y| \leq 1$, we have for any $0 < \mu \leq 1/b$,

$$P\left(\sum_{i=1}^{n} Z_i > a\right) \leq e^{-\mu a} \prod_{i=1}^{n} E(e^{\mu Z_i})$$

$$\leq e^{-\mu a} \prod_{i=1}^{n} (1 + \mu^2 E(Z_i^2))$$

$$\leq \exp(-\mu a + \mu^2 V).$$

The result now follows by taking $\mu = sa/V$.

Lemma 2. Let $\{a_{i\alpha}\}$ be a sequence of real numbers satisfying, $|a_{i\alpha}| \leq \|x_{i\alpha}\|$, for $i=1, \ldots, n$. Then under the conditions of Theorem 1, for any $B$, $k > 0$, there exists an $A > 0$ such that

$$P\left(\left|\sum_{i=1}^{n} a_{i\alpha} Y_i(\beta)\right| > (A d_n)^{1/2} (\log n)^{3/4}\right) = o(n^{-k})$$

uniformly for $\|\beta\| \leq B(\log n)^{1/2}$, where

$$Y_i(\beta) = \text{sign}(e_i - x_{i\alpha} \beta) - \text{sign}(e_i) + 2 F(x_{i\alpha} \beta) - 1.$$

Proof. Since $F$ has a derivative at zero and $F(0) = 1/2$, we have for some $c > 0$, and $\|\beta\| \leq B(\log n)^{1/2}$,

$$\sum_{i=1}^{n} a_{i\alpha}^2 E(Y_i(\beta))^2 \leq 4 \sum_{i=1}^{n} a_{i\alpha}^2 |F(x_{i\alpha} \beta) - \frac{1}{2}|$$

$$\leq c d_n \|\beta\| \leq B c d_n (\log n)^{1/2}.$$

Further, as $|a_{i\alpha} Y_i(\beta)| \leq 2 d_n$, the result follows by letting $A = 4B c$ and applying Lemma 1 with $a = (B d_n)^{1/2} (\log n)^{3/4}$ and $s = 1/2$.

Lemma 3. Under the conditions of Lemma 2, for any $k \geq 2$, $B > 0$, there exists an $A > 0$ such that

$$P\left(\sup_{\|\alpha - \beta\| \leq v_n} \left|\sum_{i=1}^{n} a_{i\alpha} (Y_i(\beta) - Y_i(\alpha))\right| > A t_n\right) = o(n^{-k})$$

uniformly in $\|\beta\| \leq B(\log n)^{1/2}$, where

$$t_n = v_n = d_n \log n.$$
(In the next section, we shall need this lemma with \( t_n = d_n^{1/2} (\log n)^{3/4} \) and \( v_n = d_n \log n \).)

**Proof.** For \( \| \alpha - \beta \| \leq v_n \)

\[
\| \text{sign}(e_i - x_{in} \alpha) - \text{sign}(e_i - x_{in} \beta) \| \leq 2 \gamma_i(\alpha, \beta),
\]

where

\[
\gamma_i(\alpha, \beta) = \begin{cases} 1 & \text{if } e_i \text{ lies between } x_{in} \beta \text{ and } x_{in} \alpha \\ 0 & \text{otherwise.} \end{cases}
\]

Since \( |x_{in}(\alpha - \beta)| \leq \| x_{in} \| v_n \), we have \( \gamma_i(\alpha, \beta) = 0 \), whenever \( |e_i - x_{in} \beta| > 2 v_n \| x_{in} \| \).

Hence

\[
\sup_{\| \alpha - \beta \| \leq v_n} \left| \sum_{i=1}^{n} \alpha_{in}(Y_i(\alpha) - Y_i(\beta)) \right| \leq 2 \sum_{i=1}^{n} |\alpha_{in}| (I(A_i(\beta)) - P(A_i(\beta)))
\]

\[+ 4 \sum_{i=1}^{n} |\alpha_{in}| P(A_i(\beta)), \tag{8}\]

where

\[A_i(\beta) = (|e_i - x_{in} \beta| \leq 2 v_n \| x_{in} \|).\]

Further there exists a \( c > 0 \) such that for \( \| \beta \| \leq B (\log n)^{1/2} \),

\[P(A_i(\beta)) \leq c v_n \| x_{in} \|. \tag{9}\]

Hence by (4), the last sum in (8) is not more than \( 4 c p v_n \) for \( \| \beta \| \leq B (\log n)^{1/2} \).

Let \( D = \max(2k, p c) \). The result now follows by applying Lemma 1 with \( a = D v_n \), \( s = p c / 2D \) to

\[
\sum_{i=1}^{n} |\alpha_{in}| (I(A_i(\beta)) - P(A_i(\beta))), \tag{10}\]

as the variance of the sum above is not more than \( p c v_n d_n \).

**Lemma 4.** Under the assumptions of Lemma 2, for any \( B > 0 \),

\[
\sup_{\| \beta \| \leq B (\log n)^{1/2}} \left| \sum_{i=1}^{n} \alpha_{in} Y_i(\beta) \right| = 0(d_n^{1/2} (\log n)^{3/4})
\]

with probability 1.

**Proof.** The lemma follows from Lemmas 2 and 3 and Borel Cantelli lemma, by dividing the region

\[\{ \beta \in \mathbb{R}^P : \| \beta \| \leq B (\log n)^{1/2} \}\]

into \( M \leq (n + 2)^p/2 \), cells such that the distance between any two points in a cell is not more than \( v_n \).
Lemma 5. Under the conditions of Theorem 1, we have with probability 1,
\[
\limsup_{n \to \infty} \sup_{\|\beta\| = B(\log n)^{1/2}} |f_n(\beta) - \beta| \leq f(0)/2,
\]
where \( B = 8p/f(0) \) and
\[
f_n(\beta) = \sum_{i=1}^{n} |e_i| - |e_i - x_i \beta| - E(|e_i| - |e_i - x_i \beta|).
\]

Proof: The proof is similar to the proof of Lemma 4. The region \( \{\beta: \|\beta\| = B(\log n)^{1/2}\} \) is divided into \( M \leq (B+1)^2(n+2)^{p/2} \) cells such that if \( \alpha \) and \( \beta \) belong to the same cell, then
\[
\|\alpha - \beta\| \leq (\log n/n)^{1/2}.
\]
For \( \alpha \) and \( \beta \) belonging to the same cell, we have by (4), that
\[
|f_n(\alpha) - f_n(\beta)| \leq 2 \sum_{i=1}^{n} \|x_i \alpha\| \|\alpha - \beta\|
\leq 2(p \log n)^{1/2}.
\]
(11)

Since \( F \) has a derivative at 0 and since for any \( y \) and \( z \),
\[
|y| - |y - z| = \int_{0}^{z} (1 - 2 I(y \leq v)) \, dv,
\]
we have by (3),
\[
\text{var}(f_n(\beta)) = 4 \sum_{i=1}^{n} x_i \beta \times x_i \beta \int_{0}^{1} \int_{0}^{1} (F(\min(u, v)) - F(u) F(v)) \, du \, dv
\]
\[
= \beta^T \left( \sum_{i=1}^{n} x_i x_i^T \right) \beta (1 + o(1))
\]
\[
= B^2 (\log n) (1 + o(1))
\]
(13)
uniformly for \( \|\beta\| = B(\log n)^{1/2} \). Since \( M = O(n^{p/2}) \), the result now follows by applying Lemma 1 to \( f_n(\beta) \) for \( \|\beta\| = B(\log n)^{1/2} \), with \( s = 1/2, a = B^2(f(0)/2) \log n \) and using Borel-Cantelli lemma.

Lemma 6. Under the conditions of Theorem 1,
\[
\|\hat{\beta}\| \leq B(\log n)^{1/2}
\]
with probability 1,
for all large \( n \), where \( B = (8p/f(0)) \).
Proof. Since \( F \) is differentiable at 0, using (12) we obtain uniformly for \( \| \alpha \| = B(\log n)^{1/2} \), that

\[
E \left( \sum_{i=1}^{n} (|e_i| - |e_i - x_{in}\alpha|) \right) = \sum_{i=1}^{n} \int_{0}^{x_{in}\alpha} (1 - 2F(v)) \, dv \\
= -\| \alpha \| ^{2} f(0)(1 + o(1)).
\]

So by Lemma 5,

\[
\limsup_{n \to \infty} \sup_{\| \alpha \| = B(\log n)^{1/2}} \sum_{i=1}^{n} (|e_i| - |e_i - x_{in}\alpha|)\| \alpha \| ^{-2} \leq -f(0)/2
\]

with probability 1. Since \( \sum_{i=1}^{n} |e_i - x_{in}\alpha| \) is convex in \( \alpha \), it follows that with probability 1,

\[
\limsup_{n \to \infty} (\log n)^{-1} B^{-2} \left( \sum_{i=1}^{n} |e_i| - \inf_{\| \alpha \| \geq B(\log n)^{1/2}} \left( \sum_{i=1}^{n} |e_i - x_{in}\alpha| \right) \right) \leq -f(0)/2.
\]

This implies, with probability 1, that for all large \( n \)

\[
\sum_{i=1}^{n} |e_i| < \inf_{\| \alpha \| \geq B(\log n)^{1/2}} \sum_{i=1}^{n} |e_i - x_{in}\alpha|.
\]

This completes the proof of the lemma.

Proof of Theorem 1. Clearly, for any \( B > 0 \),

\[
\sum_{i=1}^{n} x_{in}(F(x_{in}\alpha) - F(0)) - f(0) \beta' = 0(d_n^{1/2}(\log n)^{3/4}) \tag{14}
\]

uniformly for \( \| \alpha \| \leq B(\log n)^{1/2} \). Since the true parameter is assumed to be the zero vector, under this null assumption \( y_i = e_i \). If \( \beta \) is a solution of (6), then all the non-zero directional derivatives of

\[
\sum_{i=1}^{n} |e_i - x_{in}\alpha|
\]

at \( \alpha = \beta \) must be non-negative. This implies that

\[
\left| \sum_{i=1}^{n} x_{ijn} \text{sign}(e_i - x_{in}\alpha) \right| \leq \sum_{i=1}^{n} |x_{ijn}| I(e_i = x_{in}\beta), \tag{15}
\]

where \( x_{ijn} \) is the \( j \)-th coordinate of \( x_{in} \). The right hand side of (15) is not more than \( d_n \sum_{i=1}^{n} I(e_i = x_{in}\beta) \) and \( |x_{ijn}| \leq \| x_{in} \| \) for all \( i \) and \( j \). For any \( p + 1 \) distinct integers \( 1 \leq i_j \leq n \), there exist real numbers \( u_1, \ldots, u_{p+1} \), not all of them zero
such that $\sum_{j=1}^{p+1} u_j x_{ij} \beta = 0$. Hence, if $\|\hat{\beta}\| \leq B \log n^{1/2}$ and $e_{ij} = x_{ij} \beta$ for $j = 1, \ldots, p$ + 1, it follows that $\sum_{j=1}^{p+1} u_j e_{ij} = 0$ and $|e_{ij}| \leq B d_n \log n^{1/2}$ for $j = 1, \ldots, p + 1$. Since $F$ is continuous in a neighbourhood of zero, it follows that for any $p + 1$ integers $i_j$,

$$P(e_{ij} = x_{ij} \beta, j = 1, \ldots, p + 1, \|\hat{\beta}\| \leq B \log n^{1/2}) = 0. \quad (16)$$

As a consequence of (15) and (16) we have,

$$P\left(\left|\sum_{i=1}^{n} x_{ijn} \operatorname{sign}(e_i - x_i \beta)\right| \geq (p + 1) d_n, |\beta| \leq B \log n^{1/2}\right) = 0. \quad (17)$$

The theorem now follows from (14), (17) and Lemmas 4 and 6.

4. Dependent Residuals

In this section we consider results similar to Theorem 1, when the residuals $\{e_i\}$ form a mixing sequence of random variables. We start with definitions of $\varphi$-mixing (uniform mixing) and strong mixing sequences.

A sequence $\{Z_n\}$ of random variables is called $\varphi$-mixing if there exists a sequence $\{\varphi_n\}$ such that

$$1 \geq \varphi_1 \geq \varphi_2 \geq \ldots, \lim_{n \to \infty} \varphi_n = 0,$$

and for all $k \geq 1$, $n \geq 1$, $B \in \mathcal{M}_k^n$ and $A \in \mathcal{M}_{k+n}$,

$$|P(A \cap B) - P(A) P(B)| \leq P(B) \varphi_n, \quad (18)$$

where $\mathcal{M}_m$ denotes the $\sigma$-field generated by $Z_i (m \leq i \leq r)$. The process $\{Z_n\}$ is called strong mixing if there exists a sequence $\{\alpha_n\}$ such that

$$1 \geq \alpha_1 \geq \alpha_2 \geq \ldots, \lim_{n \to \infty} \alpha_n = 0,$$

and (18) holds with the right side of the inequality $P(B) \varphi_n$ replaced by $\alpha_n$.

Clearly any $\varphi$-mixing sequence is a strong mixing sequence. Examples of strong mixing sequences include $m$-dependent sequences, certain gaussian processes and linear processes $\{Z_n\}$, where

$$Z_n = \sum_{i=1}^{\infty} b_i \xi_{n-i+1},$$
\{\xi_n\} are i.i.d. random variables and \(b_n \to 0\) at certain rate. Gorodetskiii (1977) has shown that \(\{Z_n\}\) is strongly mixing with \(\alpha_n = O(e^{-\lambda n})\) for some \(\lambda > 0\) when \(b_n = O(e^{-\gamma n})\) for some \(\gamma > 0\). In particular simple autoregressive processes,

\[
Z_n = 0Z_{n-1} + \xi_n
\]

are strongly mixing with \(\alpha_n\) decaying exponentially, if \(|\theta| < 1\), provided \(\xi_1\) has a smooth density.

If \(\{Z_n\}\) is \(\varphi\)-mixing with \(\sum \varphi_n^{1/2} < \infty\), then by (20.35) of Billingsley (1968), we have

\[
\sigma^2 = \text{var} \left( \sum_{i=1}^{n} Z_i \right) = \sum_{i=1}^{n} \text{var}(Z_i) + 2 \sum_{1 \leq i + j \leq n} \text{cov}(Z_i, Z_j) \\
\leq V(1 + 2V \sum_{1 \leq i + j \leq n} \varphi_{i-j}^{1/2} \delta_i \delta_j),
\]

where

\[
V = \sum_{i=1}^{n} \text{var}(Z_i) \quad \text{and} \quad \delta_i = (\text{var}(Z_i)/V)^{1/2}.
\]

So \(\sigma^2 \leq 2Vx^\text{A}x\), where \(a_{ij} = \varphi_{i-j}^{1/2} \geq 0\) is the \((i, j)\)th element of \(A\) and \(x = (\delta_1, \ldots, \delta_n)\). But

\[
x^\text{A}x \leq \lambda = \max(|\lambda_j|; \lambda_j \text{ is an eigenvalue of } A).
\]

As is well known

\[
\lambda \leq \max_i \sum_{j=1}^{n} a_{ij}.
\]

It follows that

\[
\sigma^2 \leq 4 \left( 1 + \sum_{j=1}^{\infty} \varphi_j^{1/2} \right) V\). \quad (19)
\]

This inequality is useful in proving Lemma 7 below.

Throughout this section we shall assume that for real numbers \(u_1, \ldots, u_{p+1}\) not all of them equal to zero, for some \(\varepsilon > 0\) and distinct integers \(i_1, \ldots, i_{p+1}\)

\[
P\left( \sum_{j=1}^{p+1} u_i \varepsilon_{i_j} = 0, |\varepsilon_{i_j}| < \varepsilon \right) = 0. \quad (20)
\]

This condition holds in particular if all the \(p+1\) dimensional distributions of the process \(\{\varepsilon_n\}\) have densities in the neighbourhood of zero vector.

The following theorems extend Theorem 1 to dependent case.

**Theorem 2.** Suppose \(\{e_n\}\) is strictly stationary \(\varphi\)-mixing sequence with

\[
\sum_{n=1}^{\infty} \varphi_n^p < \infty
\]
for some $a \leq 1/2$ and $d_n = O(n^{-\alpha(1+\alpha)^{-\gamma}})$ for some $\varepsilon > 0$. Then (7) holds with probability 1.

Theorem 3. Suppose $\{e_n\}$ is a strictly stationary $\varphi$-mixing sequence with

$$\varphi_n = O(e^{-\lambda n})$$

for some $\lambda > 0$ and $d_n = O(n^{-\varepsilon})$, for some $\varepsilon > 0$. Then (7) holds with probability 1.

Theorem 4. Suppose $\{e_n\}$ is a strictly stationary strong mixing sequence with

$$\alpha_n = O(e^{-\lambda n})$$

for some $\lambda > 0$ and $d_n = O(n^{-\varepsilon - 1/3})$ for some $\varepsilon > 0$. Then, with probability 1, (7) holds when the error term is replaced by $O(d_n^{1/4} \log n^{1/4})$.

The next lemma is the key step for proving Theorems 2 and 3.

Lemma 7. Let $\{Z_n\}$ be a $\varphi$-mixing sequence with $E(Z_n) = 0$. Suppose $|Z_n| \leq b$ for some $b > 0$. Then for any $\eta > 2$, there exist positive constants $T$ and $H$ depending only on $\{\varphi_n\}$ and $\eta$ such that

$$P\left( \sum_{i=1}^{n} Z_i > 8 D \right) \leq T \left( \exp(-yD + Hy^2 + (yb g(n))^{1/2}) \right),$$

where

$$g(n) = n^{\alpha(1+\alpha)} \quad \text{if} \quad \sum \varphi_n^2 < \infty \quad \text{for some} \quad a \leq 1/2$$

$$= \log n \quad \text{if} \quad \varphi_n = O(e^{-\lambda n}) \quad \text{for some} \quad \lambda > 0.$$  

Proof of this result is similar to the proof of Lemma 2.1 of Babu and Singh (1978) and so is omitted. Instead of $k = [n^{2/3}]$ there, we have to take $k = [n^{1/(1+\alpha)}]$.

Lemmas 2 and 3 can be concluded for $\varphi$-mixing case, from Lemma 7 by taking $y = d_n^{-1/2} (\log n)^{1/4}$, $b = d_n$ and $D = Ad_n^{1/2} (\log n)^{3/4}$ for some $A > 0$.

Lemma 5 follows by taking $y$ to be an appropriate constant depending on $f(0)$ and $H$, and by taking $B$ to be a large but fixed constant. Note that Lemma 3 holds with $t_n = d_n^{1/2} (\log n)^{3/4}$. The rest of the proofs of Theorems 2 and 3 are similar to that of Theorem 1.

Theorem 4 can be established on similar lines by adopting Lemma 7 to $\alpha$-mixing sequences. For example see Lemma 3.3 of Babu and Singh (1978).

Acknowledgement. The author would like to thank Professor C.R. Rao for helpful discussions during the preparation of this manuscript.

References

Edgeworth, F.Y.: A new method of reducing observations relating to several quantities. Philosophical Magazine (5th series) 24, 222–223 (1887)

Received November 10, 1988; in revised form July 7, 1989