PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 28, Fasc. 1 (2008), pp. 129–141

BOUNDS FOR $\mathbb{E}\,|S_n|^Q$ FOR SUBORDINATED LINEAR PROCESSES WITH APPLICATION TO M-ESTIMATION

BY

KONRAD FURMAŃCZYK (WARSZAWA)

Abstract. Let $X_j = \sum_{r=0}^{\infty} A_r Z_{j-r}$ be a one-sided m-dimensional linear process, where (Z_n) is a sequence of i.i.d. random vectors with zero mean and finite covariance matrix. The aim of this paper is to prove the moment inequalities of the form

$$\mathbb{E}\,|S_n|^Q \leqslant C n^{Q/2} \tag{0.1}$$

for the sum

$$S_n = \sum_{j=1}^{n} \big(G(X_j) - \mathbb{E}G(X_j)\big), \tag{0.2}$$

where G is a real function defined on $\mathbb{R}^m$. The form of the constant C in (0.1) plays an important role in applications concerning the problems of M-estimation, especially the Ghosh representation.

2000 AMS Mathematics Subject Classification: Primary: 62M10, 62F12; Secondary: 60G35.

Key words and phrases: Linear process, empirical processes, time series, short-range dependence, M-estimation, the Ghosh representation.

1. INTRODUCTION

The one-sided linear process (LP) is defined as follows:

$$X_j = \sum_{r=0}^{\infty} a_r Z_{j-r}, \quad j = 1, 2, \ldots, n,$$

where the innovations (Z_n) are i.i.d. random variables with zero mean and unit variance, and a_r are constant coefficients satisfying the condition

$$\sum_{r=0}^{\infty} a_r^2 < \infty.$$

Linear processes have a wide range of applications in time series analysis. A large class of time series processes can be modelled in such a way, including a subset of the fractional ARIMA processes (see Brockwell and Davis (1987)). In addition, we say that (X_n) has a *short memory* (is *short-range dependent*) if the following condition is satisfied:

$$\sum_{j=1}^{\infty} |\mathbf{Cov}(X_1,\ X_{1+j})| < \infty.$$

It is obvious that if

$$\sum_{j=0}^{\infty} \Big| \sum_{r=0}^{\infty} a_r a_{r+j} \Big| < \infty,$$

then the linear process (X_n) is short-range dependent.

The generalization of the one-sided linear process is the multidimensional linear process (MLP), defined as follows:

$$X_j = \sum_{r=0}^{\infty} \mathbf{A}_r Z_{j-r},$$

where the innovations $(Z_n) = (Z_n^{(1)}, \ldots, Z_n^{(m)})$ are i.i.d. random vectors in $\mathbb{R}^m$ such that

$$\mathbb{E}\,(Z_1) = 0 \quad \text{and} \quad \mathbb{E}\,|Z_1|^2 = 1.$$

The $\mathbf{A}_r$'s are the nonrandom matrices, where $\sum_{r=0}^{\infty} \|\mathbf{A}_r\|^2 < \infty$, $\mathbf{A}_0 = \mathbf{I}$, and

$$|\mathbf{A}_r z| \leqslant \|\mathbf{A}_r\|\,|z| \quad \text{for any } z \in \mathbb{R}^m, \tag{1.1}$$

where $|\cdot|$ is the usual Euclidean norm.

We will also consider the following assumptions:

$$\sum_{r=0}^{\infty} \|\mathbf{A}_r\| < \infty, \tag{a_0}$$

(b_1) the density of the vector Z_1 is the Lipschitz function,

$$\mathbb{E}\,|Z_1|^t < \infty \quad \text{for some } t \geqslant 2. \tag{$\mathrm{b}_2(t)$}$$

In Section 2, we will prove the inequality for the moment bound of the sum of the functionals of linear processes. Namely, we will show that if

$$Y_n = G(X_n) - \mathbb{E}G(X_n), \tag{1.2}$$

then for different classes of functions G

$$\mathbb{E}\Big| \sum_{j=1}^{n} Y_j \Big|^Q \leqslant C n^{Q/2}, \quad n = 1, 2, \ldots, \tag{1.3}$$

where the constant C may depend on Q, G and the distribution of (X_n), but does not depend on n.

The basic idea in the proof of our result is the martingale decomposition of the process (1.2). This decomposition has the following form (see also Ho and Hsing (1996), (1997)):

$$(1.4) \qquad Y_n = \sum_{s=0}^{\infty} U_{n,s}, \quad \text{where } U_{n,s} = \mathbb{E}\{Y_n \mid \mathcal{F}_{n-s}\} - \mathbb{E}\{Y_n \mid \mathcal{F}_{n-s-1}\},$$

where $\mathcal{F}_n := \sigma(\ldots, Z_{n-1}, Z_n)$ is the σ-field generated by the innovations in the "past" $\leqslant n$, (X_n) denotes an m-dimensional process (MLP), and $G\colon \mathbb{R}^m \to \mathbb{R}$ stands for a real function such that $\mathbb{E}\,|Y_1|^Q < \infty$ for some $Q \geqslant 2$.

In Section 3, we will show how to use these moment bounds in order to obtain the Ghosh representation for M-estimation in the case of the short-range dependent observations. We will apply here Andrews and Pollard (1994) results (see Remark 3.2). The basic condition required for the proof of the Ghosh representation is the condition of the asymptotic stochastic equicontinuity (ASE). We obtain this condition from the Pisier result (see Lemma 3.2).

The Ghosh representation is especially useful for the proof of the asymptotic normality of M-estimators, provided the central limit theorem (CLT) holds for the sums S_n. Ho and Hsing (1997), Wu (2002), and Furmańczyk (2007) proved the CLT for the sums (0.2), where (X_n) is a one-dimensional linear process or a multidimensional linear process, in the case of short-range dependence. In Subsection 3.1, we will prove Theorem 3.1 together with some conclusions from this theorem (see Propositions 3.1 and 3.2). At the end of Section 3, we will state a lemma about the consistence of M-estimators.

2. MOMENT INEQUALITIES

Let us put

$$\|X\|_Q = \mathbb{E}^{1/Q}\,|X|^Q$$

and

$$G_s(x) := \mathbb{E}G\Big(\sum_{r=0}^{s-1} \mathbf{A}_r Z_{k-r} + x\Big) = \int G(z+x)\,dF_s(z),$$

where F_s is the distribution function of $\sum_{r=0}^{s-1} \mathbf{A}_r Z_{k-r}$. We will use the following condition:

$$(\mathrm{Lip}) \qquad |G_s(x) - G_s(y)| \leqslant \mathrm{Lip}(G_s)\,|x-y|$$

for each $s = 1, 2, \ldots$ and for all $x, y \in \mathbb{R}^m$, where

$$\sup_{s\geqslant 0} \mathrm{Lip}(G_s) < C \quad \text{for some constant } C.$$

THEOREM 2.1. *Let* $\|G(X_1)\|_Q < \infty$ *for some* $Q \geqslant 2$ *and let* (X_n) *be an* m*-dimensional linear process* (MLP) *satisfying* (a_0) *and the assumption* $(\mathrm{b}_2(t))$ *with* $t = Q$. *Then, if* (Lip) *holds, we have for every* $n \in \mathbb{N}$

$$\mathbb{E}\Big|\sum_{j=1}^{n}\big(G(X_j) - \mathbb{E}G(X_j)\big)\Big|^Q \leqslant Cn^{Q/2} \tag{2.1}$$

with $C = C_{A,Q}C(G)$, *where*

$$C_{A,Q} = C_Q\big(\sum_{r=0}^{\infty}\|\mathbf{A}_r\|\big)^Q, \quad C(G) = \big(\|G(X_1)\|_Q + \sup_{s\geqslant 1}\mathrm{Lip}(G_s)\big)^Q,$$

and C_Q *is a constant dependent on* Q *and* $\|Z_1\|_Q$.

For the proof of Theorem 2.1 we will need the following lemmas:

LEMMA 2.1. *Let* (X_n) *be an* m*-dimensional linear process* (MLP) *and assume that* (Lip) *holds. Then*

$$|U_{k,s}| \leqslant \mathrm{Lip}(G_s)\,\|\mathbf{A}_s\|\,(1 + |Z_{k-s}|) \;\; a.s. \quad \textit{for } s = 1, 2, \ldots \tag{2.2}$$

P r o o f. By the definition of G_s, we obtain

$$U_{k,s} = \mathbb{E}\{G(X_k) \mid \mathcal{F}_{k-s}\} - \mathbb{E}\{G(X_k) \mid \mathcal{F}_{k-s-1}\} = G_s(R_{k,s}) - G_{s+1}(R_{k,s+1}),$$

where $R_{k,s} = X_k - \sum_{r=0}^{s-1}\mathbf{A}_r Z_{k-r}$. From independence of $\mathbf{A}_s Z_{k-s}$ and $R_{k,s+1}$ we have

$$G_{s+1}(x) = \int G_s(x+z)dF_{1,s}(z)$$

and

$$G_{s+1}(R_{k,s+1}) = \int G_s(R_{k,s+1}+z)dF_{1,s}(z),$$

where $F_{1,s}$ is the distribution function of $\mathbf{A}_s Z_{k-s}$. Hence, applying the condition (Lip), we get

$$\begin{aligned}|U_{k,s}| &\leqslant \int |G_s(R_{k,s}) - G_s(R_{k,s+1}+z)|\,dF_{1,s}(z)\\ &\leqslant \mathrm{Lip}(G_s)\int |R_{k,s} - R_{k,s+1} - z|\,dF_{1,s}(z)\\ &\leqslant \mathrm{Lip}(G_s)\big(|R_{k,s} - R_{k,s+1}| + \int |z|\,dF_{1,s}(z)\big)\\ &\leqslant \mathrm{Lip}(G_s)\left(|\mathbf{A}_s Z_{k-s}| + \mathbb{E}\,|\mathbf{A}_s Z_{k-s}|\right).\end{aligned}$$

From (1.1) and the fact that $\mathbb{E}\,|Z_1|^2 = 1$ we have

$$|\mathbf{A}_s Z_{k-s}| \leqslant \|\mathbf{A}_s\|\,|Z_{k-s}| \tag{2.3}$$

and

$$\mathbb{E}\,|\mathbf{A}_s Z_{k-s}| \leqslant \|\mathbf{A}_s\|. \tag{2.4}$$

The relations (2.3) and (2.4) imply the desired result (2.2). ∎

LEMMA 2.2 (Burkholder (1966)). *Let $(Y_n,\ \mathcal{F}_n)$ be a stationary sequence of the martingale differences such that $\mathbb{E}\,|Y_1|^Q < \infty$ for some $Q \geqslant 2$. Then there exists a universal constant $C_Q^{'}$, dependent on Q, such that*

$$\mathbb{E}\Big|\sum_{j=1}^{n} Y_j\Big|^Q \leqslant C_Q^{'} n^{Q/2}\mathbb{E}\,|Y_1|^Q \quad \textit{for every } n \in \mathbb{N}. \tag{2.5}$$

P r o o f. Notice that from the Burkholder inequality (see Stout (1974), 3.3.14) we have

$$\mathbb{E}\Big|\sum_{j=1}^{n} Y_j\Big|^Q \leqslant C_Q^{'}\mathbb{E}\Big|\sum_{j=1}^{n} Y_j^2\Big|^{Q/2}. \tag{2.6}$$

In addition, by the Hölder inequality, we get

$$\sum_{j=1}^{n} 1\cdot Y_j^2 \leqslant \Big(\sum_{j=1}^{n} |Y_j|^Q\Big)^{2/Q} n^{1-2/Q} \quad \text{for } Q > 2.$$

This together with (2.6) imply (2.5). ∎

We now prove Theorem 2.1, the main result of our paper.

P r o o f o f T h e o r e m 2.1. We first use the martingale representation of the form (1.4) for the transformed sequence (Y_n):

$$\sum_{k=1}^{n} Y_k = \sum_{k=1}^{n}\sum_{s=0}^{\infty} U_{k,s} = \sum_{s=0}^{\infty} W_{n,s}, \tag{2.7}$$

where $W_{n,s} := \sum_{k=1}^{n} U_{k,s}$. Since for every $s \geqslant 0$ the sequence $(U_{k,s},\ \mathcal{F}_{k-s})_k$ is a stationary sequence of the martingale differences, $(W_{m,s},\ \mathcal{F}_{m-s})_m$ is a martingale. Applying the Burkholder inequality (see Lemma 2.2), we obtain

$$\mathbb{E}\,|W_{n,s}|^Q \leqslant C_Q^{'} n^{Q/2}\mathbb{E}\,|U_{1,s}|^Q.$$

Moreover, it follows from Lemma 2.1 that

$$|U_{1,s}| \leqslant \mathrm{Lip}(G_s)\,\|\mathbf{A}_s\|\,(1+|Z_{1-s}|) \quad \text{for } s = 1, 2, \ldots \tag{2.8}$$

and that

$$\|W_{n,0}\|_Q \leqslant (C_Q^{'})^{1/Q} n^{1/2}\,\|U_{1,0}\|_Q \leqslant (C_Q^{'})^{1/Q} n^{1/2}\,\|G(X_1)\|_Q.$$

Put $S_n = \sum_{k=1}^{n} Y_k$. Then

$$\|S_n\|_Q \leqslant \|W_{n,0}\|_Q + \sum_{s=1}^{\infty}\|W_{n,s}\|_Q \leqslant C_Q^{1/Q} n^{1/2}\Big(\|G(X_1)\|_Q + \sum_{s=1}^{\infty}\|\mathbf{A}_s\|\,\mathrm{Lip}(G_s)\Big)$$

for some constant C_Q dependent on Q and $\|Z_1\|_Q$, which implies the desired result (2.1). ∎

With reference to Theorem 2.1 and Lemma 2.1 the following property is worth proving.

PROPOSITION 2.1. *The following statements are true*:

(i) *If* G *is Lipschitz*, *then* (Lip) *is satisfied with the constant* $\mathrm{Lip}(G_s)$ *bounded as follows*:

$$\mathrm{Lip}(G_s) \leqslant \mathrm{Lip}(G).$$

(ii) *If* G *is integrable*, *i.e.* $G \in \mathbb{L}^1(\mathbb{R}^m)$ *and* (b_1) *holds*, *then* (Lip) *is satisfied with the constant* $\mathrm{Lip}(G_s)$ *bounded as follows*:

$$\mathrm{Lip}(G_s) \leqslant \mathrm{Lip}(f_1) \int |G(s)|\, ds,$$

where f_1 *is the density of* Z_1.

Proof. The condition (i) is clear. In order to prove (ii), notice that

$$|G_s(x) - G_s(y)| = \left|\int G(z+x) f_s(z) dz - \int G(z+y) f_s(z) dz\right|,$$

where f_s means the density of the random vector $\sum_{r=0}^{s-1} \mathbf{A}_r Z_{k-r}$. Since $f_s = g_s * f_1$ is the convolution of g_s and f_1, where g_s is the density of $\sum_{r=1}^{s-1} \mathbf{A}_r Z_{k-r}$, and f_1 is the density of Z_1, by (b_1) we obtain

$$\begin{aligned}
&\left|\int G(z+x) f_s(z) dz - \int G(z+y) f_s(z) dz\right| \\
&\qquad = \left|\int G(z)\big(f_s(z-x) - f_s(z-y)\big) dz\right| \\
&\qquad = \left|\int G(z)\big(g_s * f_1(z-x) - g_s * f_1(z-y)\big) dz\right| \\
&\qquad \leqslant |x-y|\, \mathrm{Lip}(f_1) \int |G(z)|\, dz. \quad \blacksquare
\end{aligned}$$

3. AN APPLICATION TO M-ESTIMATION

Let $X_1, X_2, \ldots$ be an m-dimensional process (MLP) and let P_θ, where $\theta \in \Theta \subset \mathbb{R}^k$, denote the marginal distribution of X_1. Let, in addition, $\psi\colon \Theta \times \mathbb{R}^m \to \mathbb{R}^k$ be a map and

$$\Psi(\theta) := \mathbb{E}\psi(\theta, X_1) = \int \psi(\theta, x) dP_\theta(x).$$

Moreover, we assume that $\Psi(\theta_0) = 0$ for some $\theta_0 \in \mathrm{int}(\Theta)$. Consider the M-estimator θ_n of a parameter θ_0 as a random function of the form

$$\Psi_n(\theta_n) = o_P(n^{-1/2}),$$

where

$$\Psi_n(\theta) := \frac{1}{n} \sum_{i=1}^{n} \psi(\theta, X_i).$$

Moreover, define an empirical process

$$\upsilon_n\psi(\theta,\cdot) := \sqrt{n}\big(\Psi_n(\theta) - \Psi(\theta)\big) \tag{3.1}$$

indexed by functions ψ. The following conditions will be used in this section:

(C) $$|\theta_n - \theta_0| \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty,$$

where $\xrightarrow{\mathbb{P}}$ denotes the convergence in probability $\mathbb{P}$.

(V) $$\Psi(\theta) = \mathbf{V}(\theta - \theta_0) + o(|\theta - \theta_0|) \quad \text{as } \theta \to \theta_0,$$

where $\mathbf{V}$ is certain nonsingular matrix.

(B) $$\upsilon_n\psi(\theta_0,\cdot) = \mathcal{O}_{\mathbb{P}}(1).$$

(ASE) For all $\varepsilon > 0$ and for every $\eta > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}\Big(\sup_{|\theta-\theta_0|<\delta} |\upsilon_n\psi(\theta,\cdot) - \upsilon_n\psi(\theta_0,\cdot)| > \eta\Big) < \varepsilon. \tag{3.2}$$

REMARK 3.1. Condition (B) holds if $\upsilon_n\psi(\theta_0,\cdot) \xrightarrow{d} \mathcal{N}_k(0,\Sigma)$ as $n \to \infty$.

REMARK 3.2. Under the conditions (C), (V), (B), and (ASE) Andrews and Pollard (1994) proved the *Ghosh representation* for M-estimators provided the observations (X_n) satisfy the strong mixing condition, i.e.

$$\sqrt{n}(\theta_n - \theta_0) = -\mathbf{V}^{-1}\upsilon_n\psi(\theta_0,\cdot) + o_{\mathbb{P}}(1). \tag{3.3}$$

The condition (ASE) (asymptotic stochastic equicontinuity) plays the key role in our applications. Let $\Theta_\delta = \{\theta \in \Theta : |\theta - \theta_0| \leqslant \delta\}$ for some $\delta > 0$.

Assume that the increments of the empirical process satisfy the following condition:

(d_Q) For all $n \geqslant n_0$ and $\theta_1, \theta_2 \in \Theta_\delta$,

$$\|\upsilon_n\psi(\theta_1,\cdot) - \upsilon_n\psi(\theta_2,\cdot)\|_Q \leqslant C\,|\theta_1 - \theta_2|, \tag{3.4}$$

where C is some positive constant.

LEMMA 3.1. *The condition* (d_Q) *for* $Q = k + 1$ *implies* (ASE).

We first prove some auxiliary lemma. Denote by $N(C,\, d,\, r)$ the covering number as the smallest number of closed balls with radius less than or equal to r, whose union covers the set C, where the metric space (C, d) is totally bounded.

LEMMA 3.2. *Let* $\mathbf{X} = (X_t)_{t\in T}$ *be a stochastic process in* $\mathbb{R}^k$, *indexed by the totally bounded metric space* $(T,\, d)$. *Assume that the process* $\mathbf{X}$ *has continuous sample paths*, *its increments are such that*

(i) *for all* $t, s \in T$ *the inequality*

$$\|X_t - X_s\|_Q \leqslant Cd(t,s) \quad \textit{for some } Q \geqslant 1$$

is satisfied, where C *is some constant independent of* t *and* s, *and*

(ii) *the following condition holds*:

$$\int_0^\delta \sqrt[Q]{N(T,d,r)}dr < \infty. \tag{3.5}$$

Then

$$\mathbb{E}\sup_{t\in T}|X_t| \leqslant \mathbb{E}\,|X_{t_0}| + 8C\int_0^\delta \sqrt[Q]{N(T,d,r)}dr, \tag{3.6}$$

where t_0 *is some point of* T *and* $\delta = \sup_{t\in T} d(t,t_0)$.

P r o o f o f L e m m a 3.2. By Theorem 11.1 of Ledoux and Talagrand (1991), putting the Young function $\psi(x) := x^Q$, we have

$$\mathbb{E}\sup_{s,t\in T}|X_s - X_t| \leqslant 8C\int_0^\delta \sqrt[Q]{N(T,d,r)}dr.$$

We also have, for every t_0 in T,

$$\mathbb{E}\sup_{t\in T}|X_t| \leqslant \mathbb{E}\,|X_{t_0}| + \mathbb{E}\sup_{s,t\in T}|X_s - X_t|.$$

Hence we obtain (3.6). ∎

P r o o f o f L e m m a 3.1. Since the metric space (Θ_δ, d) is totally bounded, where $d(\theta_1,\theta_2) = |\theta_1 - \theta_2|$, putting

$$t_0 = \theta_0, \quad t = \theta, \quad T = \Theta_\delta \quad \text{and} \quad X_t := \upsilon_n\psi(t,\cdot) - \upsilon_n\psi(t_0,\cdot)$$

we see that if $\dim(\Theta_\delta) = k$, then $N(T,d,r) = \mathcal{O}\left(r^{-k}\right)$. Consequently, assumption (ii) of Lemma 3.2 holds for $Q = k+1$. Condition (d_Q) for $Q = k+1$ implies that for all $n \geqslant n_0$ and every $t, s \in \Theta_\delta$

$$\|X_t - X_s\|_{k+1} = \|\upsilon_n\psi(t,\cdot) - \upsilon_n\psi(s,\cdot)\|_{k+1} \leqslant C\,|t-s|$$

for some constant C. Thus, assumption (i) of Lemma 3.2 is fulfilled for all $n \geqslant n_0$. Since $X_{t_0} \equiv 0$, by (3.6) we infer that for all $n \geqslant n_0$

$$\begin{aligned} \mathbb{E}\sup_{t\in T}|X_t| &= \mathbb{E}\sup_{d(\theta,\theta_0)<\delta}|\upsilon_n\psi(\theta,\cdot) - \upsilon_n\psi(\theta_0,\cdot)| \\ &\leqslant 8C\int_0^\delta N^{1/(k+1)}(\Theta_\delta,d,r)dr. \end{aligned} \tag{3.7}$$

Hence

$$\limsup_{n\to\infty} \mathbb{E} \sup_{d(\theta,\theta_0)<\delta} |\upsilon_n \psi(\theta,\cdot) - \upsilon_n \psi(\theta_0,\cdot)| \leqslant 8C \int_0^\delta N^{1/(k+1)}(\Theta_\delta, d, r)dr.$$

Therefore, it follows that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbb{E} \sup_{d(\theta,\theta_0)<\delta} |\upsilon_n \psi(\theta,\cdot) - \upsilon_n \psi(\theta_0,\cdot)| < \varepsilon,$$

which yields the desired condition (ASE). ■

3.1. Asymptotic properties of M-estimators. Let us introduce the following notation:

$$\psi_{\theta_1,\theta_2}(x) := \psi(\theta_1, x) - \psi(\theta_2, x),$$

$$\Psi_s(\theta, x) := \mathbb{E}\psi\big(\theta, \sum_{r=0}^{s-1} \mathbf{A}_r Z_{1-r} + x\big), \qquad \Psi_s(x, \theta_1, \theta_2) := \Psi_s(\theta_1, x) - \Psi_s(\theta_2, x).$$

In our further considerations we will apply the following conditions:

(L_Q) There exists a constant L such that

$$\|\psi_{\theta_1,\theta_2}(X_1)\|_Q \leqslant L\,|\theta_1 - \theta_2| \quad \text{for all } \theta_1, \theta_2 \in \Theta_\delta. \tag{3.8}$$

(Lip_1) There exists some constant C such that

$$\sup_{s\geqslant 1} |\Psi_s(x, \theta_1, \theta_2) - \Psi_s(y, \theta_1, \theta_2)| \leqslant C\,|\theta_1 - \theta_2|\,|x - y|$$

for all $\theta_1, \theta_2 \in \Theta_\delta$ and all $x, y \in \mathbb{R}^m$.

THEOREM 3.1. *Suppose that* $\mathbb{E}\,|\psi(\theta, X_1)|^{k+1} < \infty$ *for all* $\theta \in \Theta_\delta$, *the conditions* (B), (C), (V), (a_0), ($\mathrm{b}_2(t)$) *with* $t = k+1$ *are satisfied. Assume also that* (L_{k+1}) *and* (Lip_1) *are fulfilled. Then we obtain the Ghosh representation* (3.3). *Additionally, if* $\Psi_s(\theta_0, \cdot)$ *is Lipschitz, then*

$$\sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} \mathcal{N}_k\big(0, \mathbf{V}^{-1}\Sigma(\mathbf{V}^{-1})^T\big),$$

where the elements of matrix Σ *have the form*

$$\sigma_{i,j} = \mathbb{E}\big(\psi_i(\theta_0, X_1)\psi_j(\theta_0, X_1)\big) + 2\sum_{t=1}^{\infty} \mathbb{E}\big(\psi_i(\theta_0, X_1)\psi_j(\theta_0, X_{1+t})\big) \tag{3.9}$$

and $\psi = (\psi_1, \ldots, \psi_k)$.

P r o o f. In order to get the Ghosh representation, by applying our assumptions, we only need to show that (d_Q) holds for $Q = k+1$ (see Remark 3.2 and Lemma 3.1). It follows from Theorem 2.1 that

$$\|\upsilon_n\psi(\theta_1,\cdot) - \upsilon_n\psi(\theta_2,\cdot)\|_{k+1} \leqslant C^{'}\Big(\|\psi_{\theta_1,\theta_2}(X_1)\|_{k+1} + \sup_{s\geqslant 1}\mathrm{Lip}\,\big(\Psi_s(\cdot,\theta_1,\theta_2)\big)\Big),$$

where $C^{'}$ is some constant independent of ψ, θ_1, θ_2, n. Using (L_{k+1}) and (Lip_1) we state that the condition (d_Q) for $Q = k+1$ holds.

The property of asymptotic normality of θ_n can be obtained from Furmańczyk (2007), Theorem 2.1. ■

REMARK 3.3. Notice that if $\|\mathbf{A}_r\| \sim r^{-\alpha}L(r)$ for some $\alpha > 1$, where $L(\cdot)$ is some slowly varying function, and if $\Psi_s(\theta_0,\cdot)$ is the Lipschitz function for all sufficiently large s, then θ_n is asymptotically normal (see Wu (2002), Corollary 2).

Denote by D_θ and D_x the derivative of some function with respect to θ and x, respectively. Now we give the conditions which imply (Lip_1).

PROPOSITION 3.1. *The condition* (Lip_1) *holds if at least one of the following conditions is satisfied*:

(i) $\sup_{\theta\in\Theta_\delta}\sup_{x,s}|D_\theta D_x\Psi_s(\theta,x)| \leqslant C$ *for some constant* C;

(ii) $\sup_{\theta\in\Theta_\delta}\sup_{x}|D_\theta D_x\psi(\theta,x)| \leqslant C$ *for some constant* C;

(iii) *there exists a function* $L \in \mathbb{L}^1(\mathbb{R}^m)$ *such that*

$$|\psi_{\theta_1,\theta_2}(x)| \leqslant L(x)\,|\theta_1-\theta_2| \quad \textit{for all } \theta_1,\theta_2\in\Theta_\delta \textit{ and } x\in\mathbb{R}^m$$

and (b_1) *holds.*

P r o o f. It is easily seen that from (i) and the theorem about the average value (Lip_1) holds. Similarly, it follows from (ii) that

$$|\psi_{\theta_1,\theta_2}(x) - \psi_{\theta_1,\theta_2}(y)| \leqslant C\,|\theta_1-\theta_2|\,|x-y|,$$

which immediately gives (Lip_1). Now, let us consider the condition (iii). We have

$$\begin{aligned}
&\Psi_s(x,\theta_1,\theta_2) - \Psi_s(y,\theta_1,\theta_2)\\
&\qquad = \mathbb{E}\psi_{\theta_1,\theta_2}\big(\sum_{r=0}^{s-1}\mathbf{A}_r Z_{1-r} + x\big) - \mathbb{E}\psi_{\theta_1,\theta_2}\big(\sum_{r=0}^{s-1}\mathbf{A}_r Z_{1-r} + y\big)\\
&\qquad = \int \psi_{\theta_1,\theta_2}(z) f_s(z-x)dz - \int \psi_{\theta_1,\theta_2}(z) f_s(z-y)dz\\
&\qquad = \int \psi_{\theta_1,\theta_2}(z)\big(g_s * f_1(z-x) - g_s * f_1(z-y)\big)dz,
\end{aligned}$$

where f_s is the density of $\sum_{r=0}^{s-1}\mathbf{A}_r Z_{1-r}$, g_s is the density of $\sum_{r=1}^{s-1}\mathbf{A}_r Z_{1-r}$, and f_1 is the density of Z_1. Since f_s is the convolution of g_s and f_1, by condition (b_1)

the function f_1 is Lipschitz, and

$$\begin{aligned}|\Psi_s(x,\theta_1,\theta_2)-\Psi_s(y,\theta_1,\theta_2)| &\leqslant \mathrm{Lip}(f_1)\,|x-y| \textstyle\int |\psi_{\theta_1,\theta_2}(z)|\,dz \\ &\leqslant \mathrm{Lip}(f_1)\,|x-y|\,|\theta_1-\theta_2| \textstyle\int L(z)dz,\end{aligned}$$

which implies (Lip_1). ■

We will prove the following result for a one-dimensional case of linear processes where $m=1$.

PROPOSITION 3.2. *Suppose that $\mathbb{E}\,|\psi(\theta,X_1)|^2<\infty$ for all $\theta\in\Theta_\delta\subset\mathbb{R}$, conditions* (B), (C), (V) *hold, and* (a_0), (b_1), $(\mathrm{b}_2(t))$ *with $t=2$ are fulfilled for $m=1$. Assume in addition that* (L_2) *holds and*

(i) *there exists a constant C such that*

$$\|\psi_{\theta_1,\theta_2}(\cdot)\|_{tv}\leqslant C\,|\theta_1-\theta_2| \quad \textit{for all } \theta_1,\theta_2\in\Theta_\delta,$$

where $\|\cdot\|_{tv}$ denotes the total variation, and

(b_0) *there exists some constant C such that the condition*

$$\sup_{s\in N}\mathrm{Lip}(F_s)\leqslant C$$

holds, where F_s is the distribution function of $\sum_{r=0}^{s-1}a_rZ_{1-r}$.

Then we obtain the Ghosh representation (3.3). *Additionally, if $\Psi_s(\theta_0,\cdot)$ is Lipschitz, then*

$$\sqrt{n}(\theta_n-\theta_0)\xrightarrow{d}\mathcal{N}(0,\gamma^{-2}\sigma^2),$$

where $\gamma=\gamma(\theta_0)$, $\gamma(\theta)=D_\theta\Psi(\theta)$, and

$$\sigma^2=\mathbb{E}\psi^2(\theta_0,X_1)+2\sum_{r=1}^{\infty}\mathbb{E}\big(\psi(\theta_0,X_1)\psi(\theta_0,X_{1+r})\big).$$

P r o o f. First, we will show that conditions (i) and (b_0) imply (Lip_1). Observe that for any x,y

$$\psi_{\theta_1,\theta_2}(z)F_s(z-x)|_{z=-\infty}^{z=\infty}-\psi_{\theta_1,\theta_2}(z)F_s(z-y)|_{z=-\infty}^{z=\infty}=0.$$

By the formula of integration by parts for Stielties integrals, we have

$$\begin{aligned}&\Psi_s(x,\theta_1,\theta_2)-\Psi_s(y,\theta_1,\theta_2)\\ &\qquad=\textstyle\int\psi_{\theta_1,\theta_2}(z)dF_s(z-x)-\int\psi_{\theta_1,\theta_2}(z)dF_s(z-y)\\ &\qquad=-\textstyle\int\big(F_s(z-x)-F_s(z-y)\big)d\psi_{\theta_1,\theta_2}(z),\end{aligned}$$

where F_s is the distribution function of $\sum_{r=0}^{s-1} \mathbf{A}_r Z_{1-r}$. Therefore, from (b_0) we get $\mathrm{Lip}(F_s) \leqslant C$ and

$$|\Psi_s(x, \theta_1, \theta_2) - \Psi_s(y, \theta_1, \theta_2)| \\ \leqslant \mathrm{Lip}(F_s)\, |x - y|\, \|\psi_{\theta_1, \theta_2}(\cdot)\|_{tv} \leqslant C\, |x - y|\, |\theta_1 - \theta_2|.$$

By Theorem 3.1 we obtain our assertion. ■

Some remarks on the consistence of M-estimators

LEMMA 3.3. *Suppose the conditions* (a_0), (b_1) *are satisfied, and* $(\mathrm{b}_2(t))$ *holds with* $t = k + 1$. *Assume also that* (Lip_1) *is fulfilled. Then for each* $\delta > 0$

$$\sup_{\theta \in \Theta_\delta} |\Psi_n(\theta) - \Psi(\theta)| \xrightarrow{\mathbb{P}} 0.$$

Proof of Lemma 3.3. Obviously, we have

$$\sup_{\theta \in \Theta_\delta} |\Psi_n(\theta) - \Psi(\theta)| \leqslant \frac{1}{\sqrt{n}} |\upsilon_n \psi(\theta_0, \cdot)| + \frac{1}{\sqrt{n}} \sup_{d(\theta, \theta_0) < \delta} |\upsilon_n \psi(\theta, \cdot) - \upsilon_n \psi(\theta_0, \cdot)|.$$

Hence, reasoning as in the proof of Lemma 3.1 (see (3.7)), we have

$$\mathbb{E} \sup_{d(\theta, \theta_0) < \delta} |\upsilon_n \psi(\theta, \cdot) - \upsilon_n \psi(\theta_0, \cdot)| = \mathcal{O}(1),$$

and the proof is completed. ■

REMARK 3.4. One can be seen that if all the assumptions of Lemma 3.3 are satisfied and

$$\inf_{\theta \in \Theta \setminus \Theta_\delta} |\Psi(\theta)| > 0 \quad \text{for every } \delta > 0,$$

then condition (C) is fulfilled (see Van der Vaart (1998), Theorem 5.9).

Acknowledgements. The paper is a part of the author's Ph.D. Thesis written under the supervision of Professor Wojciech Niemiro at Warsaw University. I would like to thank him for his help. I am also grateful to Professor Jan Mielniczuk for his comments and remarks.

REFERENCES

[1] D. W. K. Andrews and D. Pollard, *An introduction to functional central limit theorems for dependent stochastic processes*, Internat. Statist. Review 62 (1994), pp. 119–132.
[2] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York 1968.
[3] P. J. Brockwell and R. A. Davis, *Time Series*: *Theory and Methods*, Springer, New York–Berlin–Heidelberg 1987.

[4] D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. 37 (1966), pp. 1497–1504.
[5] K. Furmańczyk, *Some remarks on the central limit theorem for functionals of linear processes under short-range dependence*, Probab. Math. Statist. 27 (2007), pp. 235–245.
[6] L. Giraitis, *Central limit theorem for functionals of linear process*, Liet. Mat. Rink. 25 (1986), pp. 43–57.
[7] C. Ho and T. Hsing, *On the asymptotic expansion of the empirical process of long memory moving averages*, Ann. Statist. 24 (1996), pp. 992–1024.
[8] C. Ho and T. Hsing, *Limit theorems for functional of moving averages*, Ann. Probab. 25 (1997), pp. 1636–1669.
[9] H. L. Koul and D. Surgailis, *Asymptotic expansion of M-estimators with long-memory errors*, Ann. Statist. 25 (1997), pp. 818–850.
[10] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, Berlin 1991.
[11] V. V. Petrov, *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*, Oxford University Press, Oxford 1995.
[12] A. N. Shiryaev, *Probability*, 2nd edition, Springer, 1996.
[13] W. F. Stout, *Almost Sure Convergence*, Wiley, New York 1974.
[14] A. Van der Vaart, *Asymptotic Statistics*, Cambridge University Press, Cambridge 1998.
[15] W. B. Wu, *Central limit theorems for functionals of linear processes and their applications*, Statist. Sinica 12 (2002), pp. 635–649.

Department of Applied Mathematics
Warsaw University of Life Sciences (SGGW)
ul. Nowoursynowska 159
02-776 Warszawa, Poland
E-mail: konfur@wp.pl

Received on 27.12.2006;
revised version on 30.9.2007