

Rates of convergence in some SLLN under weak dependence conditions

PAUL DOUKHAN, OLEG KLESOV and GABRIEL LANG

Dedicated to Ferenc Móricz on the occasion of his 70th birthday

Communicated by L. Kérchy

Abstract. Fazekas and Klesov (2000) found conditions for almost sure convergence rates in the law of large numbers that effectively can be applied if maximal inequalities are available. In the spirit of Móricz (1976), we aim at using those conditions in a weakly dependent framework, and this trick is proved to be quite efficient, first in the standard law of large numbers and second in the nonparametric estimation context where rates of convergence of the density kernel estimates are also obtained.

1. Introduction

One of the favorite topics of Ferenc Móricz is related to the notion of super-additive random sequences and fields. Introduced by Serfling [17] in order to obtain some maximal inequality, this notion is later generalized and extended by many authors, Móricz [13] in the first place. One of the main features of the so-called quasi super-additive structure [15] is the possibility to derive maximal moment inequalities from their counterparts for the underlying sequence or field. These maximal inequalities are applied by Móricz [14] to obtain sharp results on the almost sure convergence of random series and various strong laws of large numbers.

Received December 11, 2008, and in revised form May 10, 2010.

AMS Subject Classifications (2000): 60F15, 60F99, 60G10, 62G07.

Key words and phrases: dependence, mixing strong law of large numbers, rate of convergence, kernel estimators.

We further develop this approach by deriving weighted maximal inequalities from their maximal counterparts for several cases arising in statistical models. In doing so we follow the general idea described in [10] (also see [12]).

2. Basic results

A result from Fazekas and Klesov [10] states that:

Theorem 1. *Let $(S_n)_{n>0}$ be an arbitrary sequence of real valued random variables. Assume that (b_n) is a sequence of nondecreasing constants such that $b_n > 0$ and $b_n \rightarrow \infty$ as $n \rightarrow \infty$. Let $p > 0$. If there exists a nondecreasing sequence $(A_n)_{n>0}$ such that*

$$(1) \quad \mathbb{E} \max_{1 \leq k \leq n} |S_k|^p \leq A_n$$

and

$$(2) \quad \sum_{n=1}^{\infty} \frac{A_n - A_{n-1}}{b_n^p} < \infty,$$

then

$$(3) \quad \lim_{n \rightarrow \infty} \frac{S_n}{b_n} = 0 \quad \text{almost surely.}$$

We will also use another Móricz's result of [13] (see Theorem 2 below) to derive maximal inequalities like (1) from simple moment inequalities.

Let $(\xi_k)_{k \geq 0}$ be a sequence of random variables. Put $S_{i,0} = 0$, $i \geq 0$, and

$$(4) \quad S_{i,j} = \sum_{k=i+1}^{i+j} \xi_k, \quad i \geq 0, \quad j \geq 1.$$

Let a function $g : \{0, 1, 2, \dots\} \times \{1, 2, \dots\} \rightarrow \mathbb{R}^+$ be such that

$$(5) \quad g(i, k) + g(i + k, l) \leq g(i, k + l)$$

for all $i \geq 0$, $k \geq 1$, and $l \geq 1$. An example of such a function is given by

$$g(i, j) = j^a, \quad i \geq 0, \quad j \geq 1,$$

if $a \geq 1$.

Theorem 2. Let $\gamma > 0$. We first assume that $\alpha > 1$ and that

$$(6) \quad \mathbb{E}|S_{i,j}|^\gamma \leq g^\alpha(i, j) \quad \text{for all } i \geq 0, j \geq 1,$$

where g satisfies (4). Then

$$\mathbb{E}(\max_{1 \leq k \leq n} |S_{i,k}|)^\gamma \leq C_{\gamma,\alpha} g^\alpha(i, n) \quad \text{for all } i \geq 0, j \geq 1,$$

where $C_{\gamma,\alpha}$ is some universal constant.

Further if (6) holds with $\alpha = 1$, then

$$\mathbb{E}(\max_{1 \leq k \leq n} |S_{i,k}|)^\gamma \leq C_\gamma (\log 2n)^\gamma g(i, n), \quad i \geq 0, \quad n \geq 1,$$

where C_γ is some universal constant.

We aim at using Theorems 1 and 2 in order to derive rates of the almost sure convergence for partial sums and for Nadaraya–Watson kernel regression nonparametric estimators (see [1]).

Notation. In what follows, we put $\|Z\|_q = (\mathbb{E}|Z|^q)^{1/q}$ for any random variable Z and any constant $q > 0$.

3. Partial sums

Let (X_i) be a centered stationary process. Consider the sequence of its cumulative sums

$$S_n = \sum_{i=1}^n X_i.$$

In this simple case we shall give conditions such that $b_n = n^b$ is suitable for each $b > 1/2$ in Theorem 1. Various inequalities can be provided with $A_n = \mathcal{O}(n^{p/2})$ if $p \geq 2$. These inequalities can be derived from Theorem 2 and they lead to the following result by Theorem 1.

Theorem 3. *Let $p > 2$. If*

$$(7) \quad \mathbb{E}|S_n|^p \leq Cn^{p/2}$$

for some constant $C > 0$ and all $n \geq 1$, then

$$\lim_{n \rightarrow \infty} \frac{S_n}{\sqrt{n} \cdot \ln^b(n)} = 0 \quad \text{a.s.}$$

for all $b > 1/p$.

The case of $p = 2$ is not covered by Theorem 3. Since this case is of a special interest, we consider a corresponding result separately.

Theorem 4. *Let $(X_i)_{i \geq 1}$ be a centered stationary process.*

1. *If the series $\sum_{k=-\infty}^{\infty} \mathbb{E}X_0X_k$ converges, then*

$$(8) \quad \mathbb{E}|S_n|^2 \leq Cn, \quad n \geq 1,$$

for some constant $C > 0$ and the strong law of large numbers (3) holds for $b_n = \sqrt{n} \ln^b(n)$ for all $b > 1$.

2. *If*

$$(9) \quad \mathbb{E}|S_n|^2 \leq Cn^a, \quad n \geq 1,$$

for some constants $a > 1$ and $C > 0$, then the strong law of large numbers (3) holds with $b_n = n^{\frac{a}{2}} \ln^b(n)$ for all $b > 1/2$.

Proof. For the case 1, inequality (8) follows by simple algebra. Thus Theorem 2 implies (1) with $A_n = \mathcal{O}(n \ln(n))$ and Theorem 1 completes the proof.

The proof of case 2 is the same: we start from bound (9), use Theorem 2 to get a maximal inequality, and finally apply Theorem 1 to prove (3). ■

Example 1. (long range dependence). Case 2 occurs if the series of covariances diverges, namely if $\mathbb{E}X_0X_n$ behaves like n^{a-2} , $a \geq 1$. The case of the long range dependence is an example where (9) holds. Long range dependence is considered in detail in the monograph [8]; note that alternative models with long range dependence are considered e.g. in [5], [6], [19], or [18].

Example 2. (independence). The case of independent random variables $(X_i)_{i \geq 1}$ gives the limits of our results. The Marcinkiewicz-Zygmund inequality for independent variables yields $C = C_p \mathbb{E}|X_0|^p$ for some constant C_p which only depends on $p \geq 2$ in inequality (7), see [11].

The following three subsections describe other examples of such weakly dependent situations where a maximal inequality holds for the underlying sequence of random variables and thus one can obtain the rate of convergence in the corresponding strong law of large numbers.

3.1. Strong mixing case. Denote by $(\alpha_i)_{i \in \mathbb{N}}$ the strong mixing coefficient sequence of the centered and stationary sequence $(X_i)_{i \in \mathbb{N}}$.

Proposition 1. ([4]). *Assume that $\|X_0\|_b < \infty$ for some $b > p$. If*

$$\alpha_i = \mathcal{O}(i^{-\alpha}), \quad i \geq 1,$$

with some $\alpha > \frac{p}{2} \cdot \frac{b}{b-p}$, then inequality (7) holds for some constant C which only depends on $p \geq 2$, on $\|X_0\|_b$, and on the mixing sequence.

Proposition 1 corresponds to Proposition 1 of [4] in the special case of $X_i = V_i$, $U_i \equiv 1$, and $q = 2$. In its turn, Proposition 1 of [4] is a particular case of Theorem 2.5 by Rio [16].

3.2. Causal weak dependence. Define the γ coefficient of dependence of a centered sequence $(X_i)_{i \in \mathbb{N}}$ with values in \mathbb{R}^d by

$$\gamma_i = \sup_{k \geq 0} \|\mathbb{E}(X_{i+k} | \mathcal{M}_k)\|_1$$

where \mathcal{M}_k is the σ -algebra generated by $\{X_t, t \leq k\}$.

Proposition 2. ([4]). *Assume that $p \geq 2$. Let $\|X_0\|_b < \infty$ for some $b > p$. If*

$$\gamma_i = \mathcal{O}(i^{-\gamma}), \quad i \geq 1,$$

with $\gamma > \frac{p}{2} \cdot \frac{b-1}{b-p}$, then inequality (7) holds for some constant C which only depends on $p \geq 2$, on $\|X_0\|_b$, and on the sequence (γ_i) .

Proposition 2 is a special case of Proposition 2 in [4] for $X_i = V_i$, $U_i \equiv 1$, and $q = 2$. In its turn, Proposition 2 in [4] is a particular case of Corollary 5.3 in [3].

3.3. Non causal weak dependence. Here we consider non causal weakly dependent stationary sequences of bounded variables and assume that p is an integer.

Let $u \geq 1$ be an integer number. By Λ we denote the set of functions $g : \mathbb{R}^u \rightarrow \mathbb{R}$ such that $\text{Lip } g < \infty$ where

$$\text{Lip } g = \sup_{(x_1, \dots, x_u) \neq (y_1, \dots, y_u)} \frac{|g(y_1, \dots, y_u) - g(x_1, \dots, x_u)|}{|y_1 - x_1| + \dots + |y_u - x_u|}.$$

Further let $k \in \mathbb{Z}$ and $u, v \geq 1$ be two integer numbers. Let an u -tuple (i_1, \dots, i_u) and an v -tuple (j_1, \dots, j_v) be such that

$$(10) \quad i_1 \leq \dots \leq i_u < i_u + k \leq j_1 \leq \dots \leq j_v.$$

Let $g_1: \mathbb{R}^u \rightarrow \mathbb{R}$ and $g_2: \mathbb{R}^v \rightarrow \mathbb{R}$ be two real valued functions of $\Lambda^{(1)} = \{g_1 \in \Lambda, \|g_1\|_\infty \leq 1\}$. Then we put

$$\text{Cov}(g_1, g_2) \equiv \text{Cov}(g_1(X_{i_1}, \dots, X_{i_u}), g_2(X_{j_1}, \dots, X_{j_v})).$$

A sequence $(X_i)_{i \in \mathbb{Z}}$ is said to be λ -weakly dependent (respectively, η -weakly dependent or κ -weakly dependent) if there exists a sequence $(\lambda(i))_{i \in \mathbb{Z}}$ (respectively, $(\eta(i))_{i \in \mathbb{Z}}$ or $(\kappa(i))_{i \in \mathbb{Z}}$) decreasing to zero at infinity such that inequality (11), (12) or (13) holds respectively for all $k \in \mathbb{Z}$, all $u, v \geq 1$, and all u -tuples (i_1, \dots, i_u) and v -tuples (j_1, \dots, j_v) satisfying (10), where

$$(11) \quad |\text{Cov}(g_1, g_2)| \leq (u\text{Lip } g_1 + v\text{Lip } g_2 + uv \text{Lip } g_1 \text{Lip } g_2)\lambda(k),$$

$$(12) \quad |\text{Cov}(g_1, g_2)| \leq (u\text{Lip } g_1 + v\text{Lip } g_2)\eta(k),$$

$$(13) \quad |\text{Cov}(g_1, g_2)| \leq (uv \text{Lip } g_1 \text{Lip } g_2)\kappa(k).$$

Inequality (11) involves an additive part $u\text{Lip } g_1 + v\text{Lip } g_2$ and a multiplicative part $uv \text{Lip } g_1 \text{Lip } g_2$. The two other definitions of the weak dependence involve only one of the additive parts. The monograph [3] details weak dependence concepts, as well as extensive models and results. Following [3] (see Definition 4.1 therein) we define

$$(14) \quad c_{X,p}(r) = \max_{1 \leq l < p} \sup_{\substack{t_1 \leq \dots \leq t_p \\ t_{l+1} - t_l \geq r}} |\text{Cov}(X_{t_1} \cdots X_{t_l}, X_{t_{l+1}} \cdots X_{t_p})|$$

These coefficients are related to the previous ones, see [3]:

Proposition 3. *Assume that a stationary and vector valued sequence $(X_n)_{n \in \mathbb{Z}}$ is weakly dependent (respectively, λ - or κ -weakly dependent) and satisfies $\|X_0\|_b < \infty$ for some $b > p$. Then the coefficient defined by (14) is such that*

$$c_{X,p}(r) \leq \begin{cases} 4p^2 8^{-\frac{b-p}{b-2}} \|X_0\|_b^{\frac{(p-2)b}{b-2}} \kappa(r)^{\frac{b-p}{b-2}}, & \text{under } \kappa\text{-dependence,} \\ p^2 4^{\frac{b}{b-1}} \|X_0\|_b^{\frac{(p-1)b}{b-1}} \lambda(r)^{\frac{b-p}{b-1}}, & \text{under } \lambda\text{-dependence.} \end{cases}$$

Proof. We first consider the proof in the case of a κ -dependent sequence. Consider integers $1 \leq \ell < p$ and $t_1 \leq \dots \leq t_p$ such that $t_{\ell+1} - t_\ell \geq r$. We need to estimate (uniformly with respect to ℓ and $t_1 \leq \dots \leq t_p$) the expression

$$(15) \quad c = |\text{Cov}(X_{t_1} \cdots X_{t_\ell}, X_{t_{\ell+1}} \cdots X_{t_p})| = |\text{Cov}(A, B)|.$$

In order to estimate (15), for some $M > 0$ depending on r and to be defined later, we now set $\bar{X}_j = X_j \vee (-M) \wedge M$. Then we also have

$$c \leq |\text{Cov}(\bar{A}, \bar{B})| + |\text{Cov}(A - \bar{A}, B)| + |\text{Cov}(\bar{A}, B - \bar{B})|,$$

with $\bar{A} = \bar{X}_{t_1} \cdots \bar{X}_{t_\ell}$, $\bar{B} = \bar{X}_{t_{\ell+1}} \cdots \bar{X}_{t_p}$. We write

$$\begin{aligned} |(A - \bar{A})B| &\leq \sum_{i=1}^{\ell} Y_i |X_{t_i} - \bar{X}_{t_i}|, \\ |\bar{A}(B - \bar{B})| &\leq \sum_{i=\ell+1}^p Y_i |X_{t_i} - \bar{X}_{t_i}| \end{aligned}$$

where Y_i is the product of $p - 1$ factors $Z_{i,j} = |X_{t_j}|$ or $|\bar{X}_{t_j}|$ for $1 \leq j \leq p$ and $j \neq i$. It is thus clear from Hölder's inequality that an analogous representation of the centering terms yields

$$|\text{Cov}(A - \bar{A}, B)| + |\text{Cov}(\bar{A}, B - \bar{B})| \leq 2 \sum_{i=1}^p \|X_0\|_b^{p-1} \|X_0 - \bar{X}_0\|_p$$

where $(p - 1)/b + 1/p = 1$. Set $h(x) = x \vee (-M) \wedge M$, then $\text{Lip } h = 1$ and $\|h\|_\infty = M$. Thus $f_\ell(x_1, \dots, x_\ell) = h(x_1) \cdots h(x_\ell)$ is such that $\bar{A} = f_\ell(X_{t_1}, \dots, X_{t_\ell})$, $\bar{B} = f_{p-\ell}(X_{t_{\ell+1}}, \dots, X_{t_p})$ with $\|f_\ell\|_\infty \leq M^\ell$ and $\text{Lip } f_\ell \leq M^{\ell-1}$. The definition of the κ -dependence thus implies

$$|\text{Cov}(\bar{A}, \bar{B})| \leq (\ell(p - \ell)) \text{Lip } f_\ell \text{Lip } f_{p-\ell} \kappa(r) \leq \frac{p^2}{4} M^{p-2} \kappa(r).$$

Now

$$\begin{aligned} \mathbb{E}|X_0 - \bar{X}_0|^p &= \mathbb{E}|X_0 - \bar{X}_0|^p \mathbb{1}_{|X_0| \geq M} \leq 2^p \mathbb{E}|X_0|^p \mathbb{1}_{|X_0| \geq M} \leq 2^p \mathbb{E}|X_0|^b M^{p-b}, \\ \|X_0 - \bar{X}_0\|_p &\leq 2 \|X_0\|_b^{1-\frac{b}{p}} M^{1-\frac{b}{p}}. \end{aligned}$$

Setting $\mu = \|X_0\|_b$, we obtain, with the previous inequalities,

$$\begin{aligned} |\text{Cov}(A - \bar{A}, B)| + |\text{Cov}(\bar{A}, B - \bar{B})| &\leq 4p\mu^{p-1+\frac{b}{p}} M^{1-\frac{b}{p}}, \\ |\text{Cov}(A, B)| &\leq 4p\mu^{p-1+\frac{b}{p}} M^{1-\frac{b}{p}} + \frac{p^2}{4} M^{p-2} \kappa(r) \\ &\leq p^2 \left(2\mu^b M^{p-b} + \frac{1}{4} M^{p-2} \kappa(r) \right). \end{aligned}$$

Choosing M from $8\mu^b M^{2-b} = \kappa(r)$:

$$c = |\text{Cov}(A, B)| \leq 4p^2 8^{-\frac{b-p}{b-2}} \mu^{\frac{(p-2)b}{b-2}} \kappa(r)^{\frac{b-p}{b-2}}.$$

The covariance bound for λ -dependent sequences is analogous to that for η -dependence; it is already considered in Proposition 13.1 of [3]. Indeed, comparing this result with the result for the η -dependence, we conclude that the main term comes from the additive part of the weak dependence inequality. ■

Theorem 4.1 in [3] states that the relation $\|S_n\|_p \leq C\sqrt{n}$ holds if $c_{X,p}(r) = \mathcal{O}(r^{-p/2})$. Thus Proposition 3 implies

Theorem 5. *Assume that $p \geq 2$ is an even integer. Assume that a centered and stationary sequence $(X_i)_{i \in \mathbb{Z}}$ is such that $\|X_0\|_b < \infty$ for some $b > p$ and either*

- *a sequence $(X_i)_{i \in \mathbb{Z}}$ is λ -weakly dependent and $\lambda(i) = \mathcal{O}(i^{-\lambda})$ for some $\lambda \geq \frac{p}{2} \cdot \frac{b-1}{b-p}$, or*
- *a sequence $(X_i)_{i \in \mathbb{Z}}$ is κ -weakly dependent and $\kappa(i) = \mathcal{O}(i^{-k})$ for some $k \geq \frac{p}{2} \cdot \frac{b-2}{b-p}$.*

Then there exists a constant $C > 0$ such that

$$\|S_n\|_p \leq C\sqrt{n}.$$

Remark 1. There is no gain to use η weak dependence instead of λ dependence coefficients.

Remark 2. Doukhan and Wintenberger (2007) consider non-integer moments $p \in (2, 3)$ (see [9], Lemma 4), and the same inequality holds if $\mathbb{E}|X_i|^{p'} < \infty$ with $p' = p + \delta$, and $\lambda(i) = \mathcal{O}(i^{-\lambda})$ with $\lambda > 4 + 2/p'$ for $p \leq p'$ small enough, namely:

$$p \leq 2 + \frac{1}{2}(\sqrt{(p' + 4 - 2\lambda)2 + 4(\lambda - 4)(p' - 2) - 2} + p' + 4 - 2\lambda).$$

4. Functional estimation

In this section we begin with a description of the considered regression estimator and then discuss examples of weakly dependent situations.

4.1. Nadaraya–Watson estimator.

As e.g. in Ango Nze & Doukhan (1996) [1], we will consider a sequence $h \equiv h_n \rightarrow 0$ which also satisfies $nh_n^d \rightarrow \infty$ as $n \rightarrow \infty$.

Consider a stationary \mathbb{R}^d -valued process $(X_k)_{k \in \mathbb{Z}}$ with a marginal density $f: \mathbb{R}^d \rightarrow \mathbb{R}$. Kernel density estimators (of f) are defined as

$$\hat{f}(x) = \frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right).$$

Kernel regression estimators of a stationary $\mathbb{R}^d \times \mathbb{R}$ -valued process $(X_k, Y_k)_{k \in \mathbb{Z}}$ are defined as

$$\hat{r}(x) = \frac{\hat{g}(x)}{\hat{f}(x)}, \quad \hat{g}(x) = \frac{1}{nh_n^d} \sum_{i=1}^n Y_i K\left(\frac{X_i - x}{h_n}\right).$$

Then $\hat{g}(x)$ is a kernel estimator of $g(x) = r(x)f(x)$ where $r(x) = \mathbb{E}(Y_0|X_0 = x)$.

Now if the marginal density h of the couple (X_0, Y_0) ($h: \mathbb{R}^{d+1} \rightarrow \mathbb{R}$) exists then $g(x) = \int y h(x, y) dy$, which motivates the previous Nadaraya–Watson estimator for the plug-in estimation.

The following assumptions will be used below.

The first assumption is a regularity condition of order ρ for the functions of interest (here $\rho > 0$):

- (A1) For the point x of interest: the functions f, g are k -times continuously differentiable around x and $(\rho - k)$ -Hölderian, where $k < \rho$ is the largest possible integer.
- (A2) The function K is Lipschitz, admits a compact support, and satisfies

$$\int_{\mathbb{R}^d} K(u) du = 1,$$

$$\int_{\mathbb{R}^d} u_1^{\ell_1} \cdots u_d^{\ell_d} K(u) du = 0, \text{ if } 0 < \ell_1 + \cdots + \ell_d < k.$$

(A3) For the point x of interest, there exist r and s , with $r \leq s$ such that

1. $\|Y_0\|_s = c < \infty$,
2. $g_r(x) = \int |y|^r h(x, y) dy$ is a function bounded around the point x ,
3. $G(x, x) = \sup_i f_i(x, x)$ is bounded around the point (x, x) , where $f_i(x', x'')$ denotes the joint density of (X_0, X_i) .

Theorem 6. *Assume that*

$$(16) \quad \mathbb{E}|\hat{g}(x) - \mathbb{E}\hat{g}(x)|^p \leq \frac{C}{(\sqrt{nh^d})^p}$$

for a constant C which depends neither on n nor on h . Assume further that $h \equiv h_n$ is chosen at the optimal rate corresponding to the regularity ρ of f :

$$h_n = n^{-\frac{1}{2\rho+d}}.$$

Then for any $\varepsilon > 1/p$

$$\lim_{n \rightarrow \infty} \frac{n^{\frac{\rho}{2\rho+d}}}{\ln^\varepsilon(n)} (\hat{g}(x) - g(x)) = 0, \quad a.s.$$

Proof. Define

$$\xi_i = \xi_{i,h} = \frac{1}{h^d} Y_i K\left(\frac{X_i - x}{h}\right) - \frac{1}{h^d} \mathbb{E}Y_i K\left(\frac{X_i - x}{h}\right).$$

We apply Theorem 2 to this setting (now ξ_i depends on the parameter $h \equiv h_n$). Denote by $\hat{g}_{i,j}$ the kernel estimator based on observations from $i+1$ to $i+j$. Then $S_{i,j} = j(\hat{g}_{i,j}(x) - \mathbb{E}\hat{g}_{i,j}(x))$ is now the analogue of the partial sums $S_{i,j}$ defined by equality (4). Using (16) and the stationarity of inputs we derive from Theorem 2

$$\mathbb{E} \max_{k \leq n} |\hat{g}_{1,k}(x) - \mathbb{E}\hat{g}_{1,k}(x)|^p \leq C(nh^d)^{p/2}.$$

But $(nh^d)^{p/2} = n^{\frac{\rho}{2\rho+d}}$. In order to derive our result we thus simply apply Theorem 1 with $A_n = Cn^{\frac{\rho\rho}{2\rho+d}}$ and $b_n = n^{\frac{\rho}{2\rho+d}} (\ln(n))^{\frac{1+\varepsilon}{p}}$ for some $\varepsilon > 0$. ■

Application. First of all, if $Y_i \equiv 1$, Theorem 6 provides a rate of convergence in the SLLN for the kernel density estimator

$$(17) \quad \lim_{n \rightarrow \infty} \frac{n^{\frac{\rho}{2\rho+d}}}{\ln^\varepsilon(n)} (\widehat{f}(x) - f(x)) = 0, \quad \text{a.s.}$$

Secondly, if one uses both results, then we derive rates of convergence for the regression estimate

$$(18) \quad \lim_{n \rightarrow \infty} \frac{n^{\frac{\rho}{2\rho+d}}}{\ln^\varepsilon(n)} (\widehat{r}(x) - r(x)) = 0, \quad \text{a.s.}$$

The conditions needed for this case are as follows:

(A4) For the point x of interest and for some $q > p$, there exists a constant $c > 0$ with

1. $\|\widehat{f}(x) - \mathbb{E}\widehat{f}(x)\|_q \leq c/\sqrt{nh_n^d}$,
2. $\|\widehat{g}(x) - \mathbb{E}\widehat{g}(x)\|_p \leq c/\sqrt{nh_n^d}$.

4.2. Applications to explicit dependence conditions.

Using results of Doukhan and Lang [4], we get various dependence situations where (A4) holds. Then Theorem 6 implies *a.s.* convergence with standard rates for this regression estimator.

Proposition 4. ([7]). *Assume that (X_i, Y_i) are independent and identically distributed. Let (A2), (A3) hold with $r = q$. Then the moment inequalities (A4) hold.*

Proposition 5. ([7]). *Assume that a sequence (X_i, Y_i) is λ -weakly dependent and (A2), (A3) hold. Assume that $\|Y_0\|_s < \infty$ for some $s > 2p$. If $\lambda(i) = \mathcal{O}(i^{-\lambda})$ and*

$$\lambda > \frac{r(2r(s-p) + 2p-s)}{(r-p)(s-2p)(r-1)}(p-1) \vee \frac{2(d-1)}{d}(q-1),$$

then the moment inequalities (A4) hold.

Proposition 6. ([7]). Assume that (A2), (A3) hold with $r > q$ and a sequence (X_i, Y_i) is α -mixing with $\alpha_i = \mathcal{O}(i^{-\alpha})$. Let either

$$\alpha > \left(\frac{r(q-1)}{r-q} \right) \vee \frac{4sr - 2s - 4r}{(r-2)(s-4)}$$

and $h \sim n^{-a}$ for

$$ad \leq \frac{1 - 2/p}{3 - 2/r},$$

or the numbers p and q be even integers and

$$\alpha > \frac{r}{2} \cdot \frac{s-2p}{s-p} \left(1 - \frac{1}{p} \right).$$

Then the moment inequalities (A4) hold.

Those are the cases for which Rosenthal-type moment inequalities are known. We did not try to get an exhaustive view of this question addressed in various papers cited in the bibliography.

Remark 3. Alternative applications of such estimators are provided for resampling in Ango Nze et al. [2]. Many applications are consequences of the present technique.

References

- [1] P. ANGO NZE and P. DOUKHAN, Non parametric minimax estimation in a weakly dependent framework I: Quadratic properties, *Math. Meth. Statist.*, **5** (1996), 404–423.
- [2] P. ANGO NZE, P. BÜHLMAN and P. DOUKHAN, Weak Dependence beyond Mixing and Asymptotics for Nonparametric Regression, *Annals of Statistics*, **30** (2002), 397–430.
- [3] J. DEDECKER, P. DOUKHAN, G. LANG, J. R. LEÓN, S. LOUHICHI and C. PRIEUR, *Weak dependence: models, theory and applications*, Lecture Notes in Statistics **190**, Springer-Verlag, 2007.
- [4] P. DOUKHAN and G. LANG, Evaluation for moments of a ratio with application to regression estimation, *Bernoulli*, **15** (2009), 1259–1286, [arXiv:0805.0228v2](https://arxiv.org/abs/0805.0228v2).
- [5] P. DOUKHAN, G. LANG and D. SURGAILIS, Asymptotics of weighted empirical processes of linear fields with long-range dependence, *Ann. Inst. Henri Poincaré Probabilités et Statistiques*, **38** (2002), 879–896.
- [6] P. DOUKHAN, G. LANG and D. SURGAILIS, Limit theorems for sums of non linear function of ARFIMA processes with random Hurst exponents and Gaussian innovations, *Lith. J. of Math.*, **47** (2007), 1–25.

- [7] P. DOUKHAN, S. LOUHICHI, A new weak dependence condition and applications to moment inequalities, *Stoch. Proc. Appl.*, **84** (1999), 313–342.
- [8] P. DOUKHAN, G. OPPENHEIM and M. TAQQU (Eds), *Long-range dependence, theory and applications*, Birkhäuser, 2002.
- [9] P. DOUKHAN and O. WINTENBERGER, Invariance principle for new weakly dependent stationary models under sharp moment assumptions, *Probab. Math. Statist.*, **27** (2007), 45–73.
- [10] I. FAZEKAS and O. KLESOV, A general approach to the strong laws of large numbers, *Theory Probab. Appl.*, **45** (2000), 568–583.
- [11] T. FIGIEL, P. HITCZENKO, W. B. JOHNSON, G. SCHECHTMANN and J. ZINN, Extremal properties of Rademacher functions with applications to the Khinchine and Rosenthal inequalities, *Trans. Amer. Math. Soc.*, **349** (1997), 997–1027.
- [12] O. I. KLESOV, The Hájek–Rényi inequality for random fields and the strong law of large numbers, *Theory Probab. Math. Statist.*, **22** (1979), 63–71; English translation in 1980.
- [13] F. MÓRICZ, Moment inequalities and the strong law of large numbers, *Z. Wahrsch. verw. Geb.*, **35** (1976), 299–314.
- [14] F. MÓRICZ, Strong limit theorems for blockwise m -dependent and blockwise weakly orthogonal sequences of random variables, *Proc. Amer. Math. Soc.*, **101** (1987), 709–715.
- [15] F. A. MÓRICZ, R. J. SERFLING and W. F. STOUT, Moment and probability bounds with quasi-superadditive structure for the maximum partial sum, *Ann. Probab.*, **10** (1982), 1032–1040.
- [16] E. RIO, *Théorie asymptotique pour des processus aléatoires faiblement dépendants*, SMAI, Mathématiques et Applications 31, Springer, 2000.
- [17] R. J. SERFLING, Moment inequalities for the maximum cumulative sum, *Ann. Math. Statist.*, **41** (1970), 1227–1234.
- [18] D. SURGAILIS and M.-C. VIANO, Long memory properties and covariance structure of the EGARCH model, *ESAIM P&S*, **6** (2002), 311–329.
- [19] M. S. TAQQU and J. B. LEVY, Using renewal processes to generate long range dependence in high variability, *Dependence in Probability and Statistics*, Birkhäuser, Boston, 1986, 73–89.

P. DOUKHAN, Université Cergy-Pontoise, UMR 8088 Analyse, Géométrie et Modélisation, 2, avenue Adolphe Chauvin, 95302, Cergy-Pontoise Cedex, France.

O. KLESOV, Department of Mathematical Analysis and Probability Theory, National Technical University of Ukraine “KPI”, pr. Peremogy, 37, 03056 Kiev, Ukraine;
e-mail: klesov@math.uni-paderborn.de

G. LANG, AgroParisTech, UMR MIA 518 INRA AgroParisTech, 16 rue Claude Bernard, F-75231 Paris Cedex 05, France.