Journal of Statistics Applications & Probability
An International Journal

Estimation of the Mode Function for $\widetilde{\rho}$-mixing Observations

M. Arfi

16 Grampian Court, PH22 1TB Aviemore, Scotland

Received: Dec. 14, 2011; Revised Mar. 7, 2012; Accepted Aug. 11, 2012
Published online: 1 Nov. 2012

Abstract: It is shown that the (empirically determined) mode of the kernel estimate uniformly converges to the conditional mode function under the $\widetilde{\rho}$-mixing condition over an increasing sequence of compact sets which increases to ^{d}.

Keywords: Kernel density estimate; conditional mode; $\widetilde{\rho}$-mixing, sequence of compact sets.

1. Introduction

Let $\{(X_i, Y_i)\}_{i\in \mathrm{N}}$ be a stationary process where (X_i, Y_i) take values in ${}^d\times$ and distributed as (X, Y). Suppose that a segment of data $\{(X_i, Y_i)\}_{i=1}^n$ has been observed. We are interested in predicting Y from the data for a fixed value of X.
Such an approach has been investigated by several authors when the observed data are i.i.d. or when the process is mixing (*see* the survey by Collomb [6] and Györfi *et al.* [9]).

The objective of this paper is to investigate the estimation of the conditional mode function, assuming that it is uniquely defined. Also, to establish the uniform almost sure convergence for the estimate of the conditional mode function, obtained from the conditional density under the $\widetilde{\rho}$-mixing hypothesis.

Besides, most of the results suppose that the data belong to a fixed compact set, this is rather cumbersome for the applications. In our paper we deal with variables belonging to a sequence of compact sets which increases to ^{d}.

Such a subject has been considered by several authors, to name a few, Collomb & *al.* [7] considered the case of the conditional mode function establishing results of strong consistency, Arfi [1] used the mode function to investigate the prediction, Gasser *et al.* [8] studied the nonparametric estimation of the mode of a distribution of random curves and Hermann & Ziegler [10] proposed rates of consistency for a nonparametric estimation of the mode in absence of the smoothness assumptions.

The conditional mode is defined by means of the conditional density $f(y|x)$ of Y, given X, as follows: $\Theta(x) = \arg\max_{y\in} f(y|x)$, and the so-called *empirical mode predictor* is defined as the maximum of $f_n(y|x)$ over $y \in$, where $f_n(y|x)$ is the kernel estimate of $f(y|x)$ defined by:

$$f_n(y|x) = \frac{f_n(x, y)}{g_n(x)};$$

here $g_n(x) > 0$, is the kernel estimate of the density function of X, $g(x)$, and $f_n(x, y)$ is the kernel estimate of the joint density of the pair (X, Y), $f(x, y)$.

These kernel estimates are defined, respectively, as follows:

$$f_n(x, y) = \frac{1}{nh_n^{d+1}} \sum_{i=1}^{n} K_2\left(\frac{y - Y_i}{h_n}\right) K_1\left(\frac{x - X_i}{h_n}\right),$$

and

$$g_n(x) = \frac{1}{nh_n^{d}} \sum_{i=1}^{n} K_1\left(\frac{x - X_i}{h_n}\right);$$

here K_1 (K_2) are two Parzen-Rosenblatt kernels on ^{d} () with K_1 strictly positive and with bounded variation, and K_2 compactly supported; h_n is a sequence of positive numbers such that: $h_n \to 0$ and $nh_n^{d+1} \to \infty$ when $n \to \infty$.

* Corresponding author: e-mail: mounir.arfi@gmail.com

We show that the random function $\Theta_n(x) = \arg\max_{y\in} f_n(y|x)$ converges uniformly over a sequence of compact sets C_n (which increases to ^{d}) to the mode function $\Theta(x)$.

2 Assumptions and main result

Let $(\Omega,\ \mathcal{F},\ P)$ be a probability space and let $(X_i,\ i\in)$ be a sequence of random variables. We write $\mathcal{F}_2 = \sigma(X_i,\ i\in S\subset)$.

Given the σ-algebras $\mathcal{B}$ and $\mathcal{R}$ in $\mathcal{F}$.

Let $\rho(\mathcal{B},\mathcal{R}) = \sup\{\mathrm{corr}(X,Y),\ X\in L_2\ (\mathcal{B}),\ Y\in L_2\ (\mathcal{R})\}$ where $\mathrm{corr}(X,Y) = (EXY - EXEY)/\sqrt{\mathrm{var}X\mathrm{varY}}$.

Bradley [3] introduced the following coefficients of dependence $\widetilde{\rho}(k) = \sup\{\rho(\mathcal{F}_S,\mathcal{F}_T)\},\ k\geq 0$ where the supremum is taken over all finite subsets $S,T\subset$ such that $\mathrm{dist}(S,T)\geq k$.

Obviously,

$$0\leq\widetilde{\rho}(k+1)\leq\widetilde{\rho}(k)\leq 1,\quad k\geq 0\ \text{ and }\ \widetilde{\rho}(0)=1.$$

Definition 2.1. *A random variable sequence ($X_i,\ i\geq 1$) is said to be a $\widetilde{\rho}$-mixing sequence if there exists $k\in$ such that* $\widetilde{\rho}(k)<1$.

Without loss of generality we may assume that $(X_i,\ i\geq 1)$ is such that $\widetilde{\rho}(1)\leq 1$ (*see* Bryc and Smolenski [5]). In the study of $\widetilde{\rho}$-mixing sequences we refer to Bradley [3], [4] for the central limit theorem, Bryc and Smolenski [5] for moment inequalities and almost sure convergence, Peligrad and Gut [11] for almost sure results.

We will make use of the following assumptions:

A1 The process $(X_i)_{i\in}$ is strictly stationary and $\widetilde{\rho}$-mixing.

A2 The joint distribution $P_{(X,Y)}$ of the pair (X,Y) is absolutely continuous with regard to the Lebesgue measure on $^d\times$.

A3 There exists $a>0$, such that $g(x)\geq n^{-a}$, $n\geq 1$, for all $x\in C_n$, where $C_n=\{x: ||x||\leq c_n\}$ such that $c_n\longrightarrow\infty$, $n\to\infty$.

A4 The kernels K_j, $j=1,2$ are Lipschitz of order $\gamma_1>0$, in the sense that:
$\exists L_K<\infty\quad |K_j(u)-K_j(v)|\leq L_K||u-v||^{\gamma_1}\quad j=1,2.$

A5 K_j, $j=1,2$ are bounded and integrate to one with K_1 assumed to be strictly positive.

A6 The mode function $\Theta(.)$ satisfies the following condition on a sequence of compact sets C_n:

$$\forall\epsilon_n>0,\ \ \exists\beta_n>0,(\forall\zeta\quad C_n\to{}^d)$$

$$\text{if}\ \sup_{x\in C_n}|\Theta(x)-\zeta(x)|\geq\epsilon_n,\ \text{then}\ \sup_{x\in C_n}|f(\Theta(x)|x)-f(\zeta(x)|x)|\geq\beta_n.$$

A7 There exists $\xi>2$ and $M<\infty$ such that $E|Y|^\xi<M$.

Theorem 2.1. *We suppose that the assumptions* A1 - A7 *hold. We further assume that the sequence* h_n satisfies:

$$h_n=o(n^{-\tau})\ \ for\ \ 1/2>\tau>\frac{1+cd+a}{\frac{\xi}{\mu}-\gamma_1^2-(d+1)(1+\gamma_1^2)}>0$$

$with\ \ \xi>2\ \ and\ \ \mu\ and\ a\ are\ two\ positive\ constants.\ Then\ we\ have:$

$$\sup_{x\in C_n}|\Theta_n(x)-\Theta(x)|\overset{a.s.}{\longrightarrow}0,\ \ n\to\infty.$$

Remark 2.1. *As sequence c_n defined in the hypotheses, we can choose $c_n=n^c$ where c is a positive constant.*

3 Preliminary results

$$\sup_{x\in C_n}\sup_{y\in}|f_n(y|x)-f(y|x)|\leq$$

$$\frac{1}{\inf_{x\in C_n}g(x)}\times\{\sup_{x\in C_n}\sup_{y\in}|f_n(x,y)-f(x,y)|+$$

$$\sup_{x\in C_n}\sup_{y\in}|f_n(y|x)||g_n(x)-g(x)|\}\leq$$

$$\frac{1}{n^{-a}}\times\left\{\sup_{x\in C_n}\sup_{y\in}|f_n(x,y)-f(x,y)|+\sup_{x\in C_n}\sup_{y\in}|f_n(y|x)||g_n(x)-g(x)|\right\}$$

with

$$\sup_{y\in}|f_n(y|x)|\leq\frac{\widetilde{K}}{h_n}\ \ \text{then}\ \ \frac{1}{n^{-a}}\sup_{y\in}|f_n(y|x)|\leq\frac{\widetilde{K}}{n^{-a}h_n}=\delta_n$$

where δ_n is such that $\delta_n\longrightarrow\infty$ when $n\ \to\infty$ and
$\widetilde{K}=\max\{\sup_{x\in^d}K_1(x),\sup_{x\in^d}K_2(y),1\}$ and we can write

$$\sup_{x\in C_n}\sup_{y\in}|f_n(y|x)-f(y|x)|\leq$$

$$n^a\sup_{x\in C_n}\sup_{y\in}|f_n(x,y)-f(x,y)|+\delta_n\sup_{x\in C_n}|g_n(x)-g(x)|$$

Lemma 3.1. *Under the assumptions* A1 - A5 *, we have:*

$\delta_n\sup_{x\in C_n}|g_n(x)-g(x)|\overset{a.s.}{\longrightarrow}0,\ \ n\to\infty.$

Proof. Consider the following decomposition:

$g_n(x)-g(x)=[g_n(x)-Eg_n(x)]+[Eg_n(x)-g(x)]$

then,

$\sup_{x\in C_n}\delta_n|g_n(x)-g(x)|=\delta_n\sup_{x\in C_n}|g_n(x)-Eg_n(x)|+\delta_n\sup_{x\in R^d}|Eg_n(x)-g(x)|.$

We start by showing that the stochastic part converges to zero almost surely when n approaches infinity and we write

$g_n(x)-Eg_n(x)=\sum_{i=1}^{n}Z_i$

where

$Z_i(x)=\frac{1}{nh_n^d}\left\{K_1\left(\frac{x-X_i}{h_n}\right)-EK_1\left(\frac{x-X_i}{h_n}\right)\right\}.$

$EZ_i=0,\ \ |Z_i|\leq 2\overline{K}_1(nh_n^d)^{-1},\ \ E|Z_i|\leq\tau n^{-1}$ and $EZ_i^2\leq\upsilon^{-2}h_n^{-d}$ where τ and υ are two positive constants.

And, we write

$\sum_{n=1}^{\infty}P\left(\delta_n|g_n(x)-Eg_n(x)|>\epsilon\right)=\sum_{n=1}^{\infty}P\left(\delta_n|\sum_{i=1}^{n}Z_i|>\epsilon\right).$

Now, we write

$$W_{ni}=Z_i\Rightarrow_{[|Z_i|\leq n^\alpha]}\ \text{and}\ V_{ni}=Z_i\Rightarrow_{[|Z_i|>n^\alpha]}\ \ \text{for}\ \alpha>1\ \text{and}\ 1\leq i\leq n.$$

Then,

$$\left|\sum_{i=1}^{n} Z_i\right| \leq \left|\sum_{i=1}^{n}(W_{ni} - EW_{ni})\right| + \left|\sum_{i=1}^{n} V_{ni}\right| + \left|\sum_{i=1}^{n} EW_{ni}\right|. \tag{3.1}$$

We need to show the following

$$\sum_{n=1}^{\infty} P\left(\delta_n \left|\sum_{i=1}^{n}(W_{ni} - EW_{ni})\right| > n^{\alpha}\epsilon/3\right) < \infty \tag{3.2}$$

$$\sum_{n=1}^{\infty} P\left(\delta_n \left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) < \infty \tag{3.3}$$

$$\delta_n n^{-\alpha}\left|\sum_{i=1}^{n} EW_{ni}\right| \longrightarrow 0, n \to \infty. \tag{3.4}$$

We start by showing (3.2).

The Markov inequality leads to:

$$\sum_{n=1}^{\infty} P\left(\delta_n \left|\sum_{i=1}^{n}(W_{ni} - EW_{ni})\right| > n^{\alpha}\epsilon/3\right) \leq c_1 \sum_{n=1}^{\infty}\sum_{i=1}^{n} \delta_n E|W_{ni}|^{\beta}/n^{\alpha\beta} \leq c_2 \sum_{n=1}^{\infty} n^{\delta+\tau d-\alpha\beta} < \infty$$

if we choose $\delta_n = n^{\delta}$ for $\delta > 0$ and $h_n = n^{-\tau}$ for $0 < \tau < 1/4$ where c_1 and c_2 are two positive constants and β such that $\beta > (1 + \delta + \tau d)/\alpha$.

Now, we show (3.3).

Note that

$$\left(\left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) \subset \bigcup_{i=1}^{n} (|Z_i| > n^{\alpha}) \quad \text{hence,}$$

$$\sum_{n=1}^{\infty} P\left(\delta_n \left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) \leq \sum_{n=1}^{\infty} n\delta_n P\left(|Z_i| > n^{\alpha}\right) \leq$$

$$\sum_{n=1}^{\infty} n\delta_n E|Z_i|^{\beta}/n^{\alpha\beta} \leq c_3 \sum_{n=1}^{\infty} n^{\delta\tau d-\beta} < \infty$$

if we choose β, δ_n and h_n as above and where c_3 is a positive constant.

Lastly we show that (3.4) holds.

We can write

$$\delta_n n^{-\alpha}\left|\sum_{i=1}^{n} EW_{ni}\right| \leq \delta_n n^{-\alpha}\left|\sum_{i=1}^{n} EV_{ni}\right| = \delta_n n^{-\alpha}\sum_{i=1}^{n} E|Z_i| \Rightarrow_{[|Z_i|>n^{\delta}]} =$$

$$\delta_n n^{-\alpha} E|Z_i| \Rightarrow_{[|Z_i|>n^{\alpha}]} \longrightarrow 0.$$

Next, we cover C_n by μ_n spheres in the shape of $\left\{x : ||x - x_{nj}|| \leq c_n\mu_n^{-1}\right\}$ for $1 \leq j \leq \mu_n^d$, $c_n \longrightarrow \infty$ and μ_n chosen such that $\mu_n \longrightarrow \infty$ to be defined later and we make the following decomposition.

$$\begin{aligned}\left|\sum_{i=1}^{n} Z_i(x)\right| \leq{}& \frac{1}{nh_n^d}\left|\sum_{i=1}^{n}\left[K_1\left(\frac{x - X_i}{h_n}\right) - K_1\left(\frac{x_{nj} - X_i}{h_n}\right)\right]\right| + \\ & \frac{1}{nh_n^d}\left|\sum_{i=1}^{n} E\left[K_1\left(\frac{x - X_i}{h_n}\right) - K_1\left(\frac{x_{nj} - X_i}{h_n}\right)\right]\right| + \\ & \frac{1}{nh_n^d}\left|\sum_{i=1}^{n}\left[K_1\left(\frac{x_{nj} - X_i}{h_n}\right) - EK_1\left(\frac{x_{nj} - X_i}{h_n}\right)\right]\right|.\end{aligned}$$

The first and the second term in the right-hand side of the inequality above are to be considered similarly and we have:

$$\frac{1}{nh_n^d}\left|\sum_{i=1}^n\left[K_1\left(\frac{x-X_i}{h_n}\right)-K_1\left(\frac{x_{nj}-X_i}{h_n}\right)\right]\right|\leq\frac{L_K}{h_n^{d+\gamma_1}}||x-x_{nj}||^{\gamma_1}\leq\frac{L_K}{h_n^{d+\gamma_1}}c_n^{\gamma_1}\mu_n^{-\gamma_1}=\frac{1}{Logn}$$

where μ_n is chosen such that

$$\mu_n=\frac{L_K^{1/\gamma_1}c_n(Logn)^{1/\gamma_1}}{h_n^{d/\gamma_1+1}}\longrightarrow\infty.$$

Then:

$$\sup_{x\in C_n}\left|\sum_{i=1}^n Z_i(x)\right|\leq\sup_{1\leq j\leq\mu_n^d}\frac{1}{nh_n^d}\left|\sum_{i=1}^n\left[K_1\left(\frac{x_{nj}-X_i}{h_n}\right)-EK_1\left(\frac{x_{nj}-X_i}{h_n}\right)\right]\right|+\frac{2}{Logn}.$$

For all $n\geq n_1(\epsilon)$ and for all $\epsilon_n>0$

$$P\left(\sup_{x\in C_n}\left|\sum_{i=1}^n Z_i(x)\right|>2\epsilon_n\right)\leq\sum_{j=1}^{\mu_n^d}P\left(\frac{1}{nh_n^d}\left|\sum_{i=1}^n\left[K_1\left(\frac{x_{nj}-X_i}{h_n}\right)-EK_1\left(\frac{x_{nj}-X_i}{h_n}\right)\right]\right|>\epsilon_n\right).$$

Now using similar decomposition as in (3.1) μ_n^d times; the use of $(\mu_n^d\delta_n)$ instead of δ_n permit to conclude that:

$$\delta_n\sup_{x\in C_n}|g_n(x)-Eg_n(x)|\overset{a.s.}{\longrightarrow}0,\quad n\to\infty.$$

Now, we show that the deterministic part $(\delta_n\sup_{x\in R^d}|Eg_n(x)-g(x)|)$ converges to zero when n approaches infinity. We write

$$Eg_n(x)-g(x)=\frac{1}{nh_n^d}\int_{\mathrm{R}^d}K_1\left(\frac{u-x}{h_n}\right)g(u)du-g(x),$$

we set $z=h_n^{-1}(u-x)$; then the use of Bochner lemma and a Taylor expansion permit to conclude.

Lemma 3.2. *Under the assumptions of the Theorem 2.1, we have:*

$$n^a\sup_{x\in C_n}\sup_{y\in\mathrm{R}}|f_n(x,y)-f(x,y)|\overset{a.s.}{\longrightarrow}0,\quad n\to\infty.$$

Proof. We write

$$f_n(x,y)-f(x,y)=\sum_{i=1}^n Z_i(x,y)+T_n(x,y),$$

where

$$T_n(x,y)=\frac{1}{nh_n^{d+1}}\sum_{i=1}^n E\left\{K_2\left(\frac{y-Y_i}{h_n}\right)K_1\left(\frac{x-X_i}{h_n}\right)\right\}-f(x,y),$$

and

$$Z_i(x,y)=\frac{1}{nh_n^{d+1}}\left\{K_2\left(\frac{y-Y_i}{h_n}\right)K_1\left(\frac{x-X_i}{h_n}\right)-E\left[K_2\left(\frac{y-Y_i}{h_n}\right)K_1\left(\frac{x-X_i}{h_n}\right)\right]\right\};$$

we have $E(Z_i)=0$, $|Z_i|\leq 2n^{-1}h_n^{-d-1}\widetilde{K}$, $E|Z_i|\leq 2n^{-1}\Gamma\widetilde{K}$ and $EZ_i^2\leq(2\Gamma\widetilde{K})/(n^2h_n^{d+1})$ where Γ is an upperbound of $f(.\,,.)$ and
$\widetilde{K}=\max\left\{\sup_{x\in^d}K_1(x),\sup_{y\in}K_2(y),\ 1\right\}$.

Now, let us write

$$\sum_{n=1}^\infty P\left(n^a\left|f_n(x,y)-f(x,y)\right|>\epsilon\right)=$$

$$\sum_{n=1}^{\infty} P\left(n^a \left|f_n(x,y) - f(x,y)\right| > \epsilon\right) = \sum_{n=1}^{\infty} P\left(n^a \left|\sum_{i=1}^{n} Z_i\right| > \epsilon\right)$$

And we write

$W_{ni} = Z_i \Rightarrow_{[|Z_i| \leq n^{\alpha}]}$ and $V_{ni} = Z_i \Rightarrow_{[|Z_i| > n^{\alpha}]}$ for $\alpha > 1$ and $1 \leq i \leq n$

Then,

$$\left|\sum_{i=1}^{n} Z_i\right| \leq \left|\sum_{i=1}^{n} (W_{ni} - EW_{ni})\right| + \left|\sum_{i=1}^{n} V_{ni}\right| + \left|\sum_{i=1}^{n} EW_{ni}\right| \tag{3.5}$$

We need to show the following

$$\sum_{n=1}^{\infty} P\left(n^a \left|\sum_{i=1}^{n} (W_{ni} - EW_{ni})\right| > n^{\alpha}\epsilon/3\right) < \infty \tag{3.6}$$

$$\sum_{n=1}^{\infty} P\left(n^a \left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) < \infty \tag{3.7}$$

$$n^{a-\alpha} \left|\sum_{i=1}^{n} EW_{ni}\right| \longrightarrow 0, n \to \infty. \tag{3.8}$$

We start by showing (3.6).
The Markov inequality provides:

$$\sum_{n=1}^{\infty} P\left(n^a \left|\sum_{i=1}^{n} (W_{ni} - EW_{ni})\right| > n^{\alpha}\epsilon/3\right) \leq c_1 \sum_{n=1}^{\infty} \sum_{i=1}^{n} n^{a-\alpha\beta} E|W_{ni}|^{\beta}$$

$$\leq c_2 \sum_{n=1}^{\infty} n^{1-(a+1)\beta} < \infty$$

where c_1 and c_2 are two positive constants and β such that $\beta > 2$.
Now, we show (3.7).
Note that

$$\left(\left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) \subset \bigcup_{i=1}^{n} (|Z_i| > n^{\alpha})$$

hence,

$$\sum_{n=1}^{\infty} P\left(n^a \left|\sum_{i=1}^{n} V_{ni}\right| > n^{\alpha}\epsilon/3\right) \leq \sum_{n=1}^{\infty} n^{1+a} P\left(|Z_i| > n^{\alpha}\right) \leq \sum_{n=1}^{\infty} n^{1-a\beta} E|Z_i|^{\beta} \leq$$

$$c_3 \sum_{n=1}^{\infty} n^{-\alpha\beta-1+a} h_n^{-(d+1)} < \infty$$

if we choose $h_n = n^{-\tau/4}$ and where c_3 is a positive constant and β such that $\beta > (a + d\tau + \tau)/\alpha$ with $a > 0$ and $\alpha > 1$.
Lastly we show that (3.8) holds.
We can write:

$$n^{a-\alpha} \left|\sum_{i=1}^{n} EW_{ni}\right| \leq n^{a-\alpha} \left|\sum_{i=1}^{n} EV_{ni}\right| = n^{a-\alpha} E|Z_i| \Rightarrow_{[|Z_i| > n^{\alpha}]} \longrightarrow 0,\ n \to \infty$$

with $\alpha > a$.

Next, we cover C_n by μ_n^d spheres in the shape of $\left\{x : ||x - x_{nj}|| \leq c_n \mu_n^{-1}\right\}$ with $1 \leq j \leq \mu_n^d$ and $\mu_n \to \infty$ to be defined later.

Consider the following decomposition

$$\sum_{i=1}^{n} Z_i(x,y) = \sum_{i=1}^{n} [\Upsilon_i(x,y) - \Upsilon_i(x_{nj},y)] -$$

$$\sum_{i=1}^{n} E[\Upsilon_i(x,y) - \Upsilon_i(x_{nj},y)] \; + \; \sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)],$$

where $\Upsilon_i(.,y) = \frac{1}{nh_n^{d+1}} K_2\left(\frac{y - Y_i}{h_n}\right) K_1\left(\frac{. - X_i}{h_n}\right)$.

The first and the second term of the equality above are to be considered similarly.

By the fact that the kernel K_1 is Lipschitz, we obtain:

$$\sup_{x\in C_n} \sup_{y\in} |\sum_{i=1}^{n} [\Upsilon_i(x,y) - \Upsilon_i(x_{nj},y)] \leq$$

$$\frac{L_K \widetilde{K}}{h_n^{d+1+\gamma_1}} ||x - x_{nj}||^{\gamma_1} \leq$$

$$\frac{L_K \widetilde{K}}{h_n^{d+1+\gamma_1}} c_n^{\gamma_1} \mu_n^{-\gamma_1} = \frac{1}{\mathrm{Log} n}$$

where μ_n is chosen so that: $\quad \mu_n = \frac{L_K^{1/\gamma_1} \widetilde{K}^{1/\gamma_1} c_n (\log n)^{1/\gamma_1}}{h_n^{(d+1+\gamma_1)/\gamma_1}} \to \infty.$

Thus,

$$\sup_{x\in C_n} \sup_{y\in} |\sum_{i=1}^{n} Z_i(x,y)| \leq$$

$$\sup_{1\leq j\leq \mu_n^d} \sup_{y\in} |\sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)]| + \frac{2}{\mathrm{Log} n},$$

and then, for all $n \geq n_1(\epsilon)$ and all $\epsilon > 0$, we have:

$$P\left\{\sup_{x\in C_n} \sup_{y\in \mathrm{R}} |\sum_{i=1}^{n} Z_i(x,y)| > 2\epsilon\right\} \leq$$

$$\sum_{j=1}^{\mu_n^d} P\left\{\sup_{y\in \mathrm{R}} |\sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)]| > \epsilon\right\}. \tag{3.9}$$

For fixed j, set:

$$\sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)] = \Delta_n(x_{nj},y) \quad \text{if} \quad |y| \leq v_n$$

$$\sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)] = \varphi_n(x_{nj},y) \quad \text{if} \quad |y| > v_n$$

where v_n is defined by $v_n = h_n^{-\frac{1}{\mu}}$ with μ being a positive constant.

Then we have:

$$\sup_{y\in} |\sum_{i=1}^{n} [\Upsilon_i(x_{nj},y) - E\Upsilon_i(x_{nj},y)]| \leq \sup_{|y|\leq v_n} |\Delta_n(x_{nj},y)| + \sup_{|y|> v_n} |\varphi_n(x_{nj},y)|.$$

Cover $[-v_n, v_n]$ by l_n spheres B_s with centers t_s and radii less than or equal to h_n^{η}, where $l_n \leq v_n h_n^{-\eta}$ and η is a fixed number. Then using same arguments to those used previously we obtain:

$$\sup_{|y|\leq v_n} |\widetilde{\Delta_n}(x_{nj}, y)| \leq \lambda_0 h_n^{\gamma_1(\eta-1)-(d+1)} \ \ a.s.,$$

where $\widetilde{\Delta_n}(x_{nj}, y) = \Delta_n(x_{nj}, y) - \Delta_n(x_{nj}, t_s)$ and λ_0 is a positive constant.

Furthermore,

$$\omega_n = P\left\{\max_{s=1,\ldots,l_n} |\Delta_n(x_{nj}, t_s)| > \epsilon/2\right\} \leq \sum_{s=1}^{l_n} P\left\{|\Delta_n(x_{nj}, t_s)| > \epsilon/2\right\} \leq$$

$$l_n \sup_{|y|\leq v_n} P\left\{|\Delta_n(x_{nj}, y)| > \epsilon/2\right\}.$$

Then, making similar decomposition to (3.5) and following the same steps of the foregoing proof with the use of $(l_n n^{-a})$ instead of n^{-a} permit to conclude that

$$n^a \sup_{x\in C} \sup_{|y|\leq v_n} |\sum_{i=1}^{n} Z_i(x, y)| \xrightarrow{a.s.} 0.$$

It remains to show that: $n^a \sup_{|y|>v_n} |\varphi_n(x_{nj}, y)| \xrightarrow{a.s.} 0$. We have

$$\sup_{|y|>v_n} |\varphi_n(x_{nj}, y)| \leq \sup_{|y|>v_n} |\sum_{i=1}^{n} \Upsilon_i(x_{nj}, y)| + \sup_{|y|>v_n} |\sum_{i=1}^{n} E\Upsilon_i(x_{nj}, y)|,$$

and by the compactness of the support of K_2,

$$K_2\left(\frac{y-Y}{h_n}\right) \leq \widetilde{K} \Rightarrow_{[|Y|>v_n/2]}.$$

Therefore,

$$n^a \sup_{|y|>v_n} |\sum_{i=1}^{n} \Upsilon_i(x_{nj}, y)| \leq \frac{n^a}{n h_n^{d+1}} \widetilde{K}^2 \sum_{i=1}^{n} \Rightarrow_{[|Y_i|>v_n/2]}. \quad (3.10)$$

We need the use of the following in our proof

$$P(|Y| > v_n/2) \leq (2v_n^{-1})^{\xi}(E|Y|^{\xi}) \quad (3.11)$$

for a certain $\xi > 0$ such that $\xi > \mu\gamma_1(\eta - 1)$.
For all $\epsilon > 0$, we have

$$P\left\{\sup_{|y|>v_n} |\sum_{i=1}^{n} \Upsilon_i(x_{nj}, y)| > \epsilon\right\} \leq \epsilon^{-1} E\left[\sup_{|y|>v_n} |\sum_{i=1}^{n} \Upsilon_i(x_{nj}, y)|\right].$$

Then, using (3.10) and (3.11) we obtain:

$$P\left\{n^a \sup_{|y|>v_n} |\sum_{i=1}^{n} \Upsilon_i(x_{nj}, y)| > \epsilon\right\} \leq \epsilon^{-1} \widetilde{K}^2 n^a h_n^{-d-1} (2v_n^{-1})^{\xi}(E|Y|^{\xi}) =$$

$$\epsilon^{-1} \widetilde{K}^2 n^a h_n^{-d-1+\frac{\xi}{\mu}} 2^{\xi}(E|Y|^{\xi}).$$

Inequality (3.9) implies:

$$P\left\{n^a \sup_{x\in C_n} \sup_{|y|>v_n} |\sum_{i=1}^{n} Z_i(x, y)| > \epsilon\right\} \leq A n^a \mu_n^d h_n^{-d-1+\frac{\xi}{\mu}} (E|Y|^{\xi}),$$

where A is a positive constant.

The choice of ξ and the assumptions of the theorem permit us to conclude that:

$$n^a \sup_{x\in C_n} \sup_{y\in} |\sum_{i=1}^{n} Z_i(x,y)| \xrightarrow{a.s.} 0$$

To complete the proof of Lemma 3.2, we should show that:

$$n^a \sup_{x\in^d} \sup_{y\in} |T_n(x,y)| \longrightarrow 0, \quad n\to\infty.$$

To this end:

$$T_n(x,y) = \frac{1}{nh_n^{d+1}} \sum_{i=1}^{n} E\left\{K_2\left(\frac{y-Y_i}{h_n}\right) K_1\left(\frac{x-X_i}{h_n}\right)\right\} - f(x,y),$$

with

$$E\left\{K_2\left(\frac{y-Y_i}{h_n}\right) K_1\left(\frac{x-X_i}{h_n}\right)\right\} = \iint_{^d\times} K_2\left(\frac{y-v}{h_n}\right) K_1\left(\frac{x-u}{h_n}\right) f_{X,Y}(u,v)dudv.$$

Properties of the Bochner's integral permit to write

$$T_n(x,y) = \frac{1}{h_n^{d+1}} \iint_{^d\times} K_2\left(\frac{y-v}{h_n}\right) K_1\left(\frac{x-u}{h_n}\right) f_{X,Y}(u,v)dudv - f(x,y).$$

Then, if we set $z_1 = (x-u)/h_n$, $z_2 = (y-v)/h_n$, we obtain

$$T_n(x,y) = \iint_{^d\times} K_2(z_2)K_1(z_1)f_{X,Y}(x-z_1h_n, y-z_2h_n)dz_1dz_2 - f(x,y).$$

The conditions made on the kernels K_j and a Taylor expansion with a proper choice of a permit to write

$$n^a \sup_{x\in^d} \sup_{y\in} |T_n(x,y)| \longrightarrow 0, \quad n\to\infty$$

4 Proof of the main result

By the definitions of $\Theta_n(x)$ and $\Theta(x)$, we have

$$|f(\Theta_n(x)|x) - f(\Theta(x)|x)| \le |f_n(\Theta_n(x)|x) - f(\Theta_n(x)|x)| + |f_n(\Theta_n(x)|x) - f(\Theta(x)|x)|$$

$$\le \sup_{y\in} |f_n(y|x) - f(y|x)| + |\sup_{y\in} f_n(y|x) - \sup_{y\in} f(y|x)|$$

$$\le 2\sup_{y\in} |f_n(y|x) - f(y|x)|.$$

Assumption A6 implies that for all $\epsilon_n > 0$ there exists $\beta_n > 0$ such that:

$$P\left(\sup_{x\in C_n} |\Theta_n(x) - \Theta(x)| \ge \epsilon_n\right) \le P\left(\sup_{x\in C_n} \sup_{y\in} |f_n(y|x) - f(y|x)| \ge \beta_n\right),$$

which completes the proof of Theorem 2.1.

Acknowledgement

The author would like to thank the referees for their remarks and criticism.

References

[1] M. Arfi, Nonparametric prediction from ergodic samples, *J. Nonparametric Statist.* **9** (1998), 23-37.

[2] M. Arfi, *Sur la régression non paramétrique d'un processus stationnaire mélangent ou ergodique*, PhD Thesis, University of Paris VI (1996).

[3] R.C. Bradley, Equivalent mixing conditions of random fields, *Technical Report 336, Center for Stochastic Processes,* University of North Carolina, Chapel Hill (1990).

[4] R.C. Bradley, On the spectral density and asymptotic normality of weakly dependent random fields, *J. Theoret. Probab.* **5** (1992), 355-374.

[5] W. Bryc, W. Smolenski, Moment conditions for almost sure convergence of weakly correlated random variables, *Proc. Amer. Math. Soc.* **19** (1993), 629-635.

[6] G. Collomb, Nonparametric regression: an up to date bibliography, *Statistics* **16** (1987), 309-324.

[7] G. Collomb, G. Hardle, S. Hassani, A note on prediction via estimation of the conditional mode function, *J. Statist. Plann. Inference* **15** (1987), 227-236.

[8] T. Gasser, P. Hall, B. Presnell, Nonparametric estimation of the mode of a distribution of random curves, *J. Roy. Statist. Soc* **60** (1998), 381-386.

[9] L. Györfi, W. Härdle, P. Sarda, P. Vieu, Nonparametric curve estimation from time series, *Lecture Notes in Statistics* **60** (1989), Springer-Verlag.

[10] E. Hermann, K. Ziegler, Rates of consistency for a nonparametric estimation of the mode in absence of smothness assumptions, *Statistics & Probability Letters* **68** (2004), 359-368.

[11] M. Peligrad, A. Gut, Almost sure results for a class of dependent random variables, *J. Theoret. Probab.* **12** (1999), 87-104.