Estimation of the Mode Function for $\tilde{\rho}$-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 $\tilde{\rho}$-mixing condition over an increasing sequence of compact sets which increases to $d$.

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

1. Introduction

Let $\{(X_i, Y_i)\}_{i \in \mathbb{N}}$ 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 $\tilde{\rho}$-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: $\Theta(x) = \arg\max_{y \in \Theta} f(y|x)$, and the so-called empirical mode predictor is defined as the maximum of $f_n(y|x)$ over $y \in \Theta$, 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:

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

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 \( \Theta_n(x) = \arg \max_{y \in C_n} f_n(y|x) \) 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 \((\Omega, \mathscr{F}, P)\) be a probability space and let \((X_i, i \in \mathbb{N})\) be a sequence of random variables. We write \( \mathscr{F}_2 = \sigma(X_i, i \in S) \).

Given the \( \sigma \)-algebras \( \mathcal{B} \) and \( \mathcal{R} \) in \( \mathcal{F} \).

Let \( \rho(\mathcal{B}, \mathcal{R}) = \sup \{ \text{corr}(X,Y), X \in L_2(\mathcal{B}), Y \in L_2(\mathcal{R}) \} \)

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

Obviously,

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

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

Without loss of generality we may assume that \((X_i, i \geq 1)\) is such that \( \rho(1) \leq 1 \) (see Bryc and Smolenski [5]). In the study of \( \rho \)-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 \((X_i)_{i \in \mathbb{N}}\) is strictly stationary and \( \bar{\rho} \)-mixing.

**A2** The joint distribution \( P_{(X,Y)}(x,y) \) 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 \( c_n \to \infty, n \to \infty \).

**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 \( \Theta(\cdot) \) satisfies the following condition on a sequence of compact sets \( C_n : \)

\[
\forall \epsilon_n > 0, \quad \exists \beta_n > 0, \quad (\forall \zeta, \quad C_n \to d)
\]

if \( \sup_{x \in C_n} |\Theta(x) - \zeta(x)| \geq \epsilon_n \), 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:

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

with \( \xi > 2 \) and \( \mu \) and \( a \) are two positive constants. Then we have:
\[ \sup_{x \in C_n} |\Theta_n(x) - \Theta(x)| \xrightarrow{a.s.} 0, \ n \to \infty. \]

**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

\[ \sup_{x \in C_n} \sup_{y \in C_n} |f_n(y|x) - f(y|x)| \leq \frac{1}{\inf_{x \in C_n} g(x)} \times \left\{ \sup_{x \in C_n} \sup_{y \in C_n} |f_n(x, y) - f(x, y)| + \sup_{x \in C_n} \sup_{y \in C_n} |f_n(y|x)||g_n(x) - g(x)| \right\} \]

with

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

where \( \delta_n \) is such that \( \delta_n \to 0 \) when \( n \to \infty \) and \( \tilde{K} = \max \{ \sup_{x \in \mathbb{R}^d} K_1(x), \sup_{x \in \mathbb{R}^d} K_2(y), 1 \} \) and we can write

\[ \sup_{x \in C_n} \sup_{y \in C_n} |f_n(x, y) - f(x, y)| + \delta_n \sup_{x \in C_n} |g_n(x) - g(x)| \]

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

\[ \delta_n \sup_{x \in C_n} |g_n(x) - g(x)| \xrightarrow{a.s.} 0, \ n \to \infty. \]

**Proof.** Consider the following decomposition:

\[ g_n(x) - g(x) = [g_n(x) - E g_n(x)] + [E g_n(x) - g(x)] \]

then,

\[ \sup_{x \in C_n} \delta_n |g_n(x) - g(x)| = \delta_n \sup_{x \in C_n} |g_n(x) - E g_n(x)| + \delta_n \sup_{x \in \mathbb{R}^d} |E g_n(x) - g(x)|. \]

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

\[ g_n(x) - E g_n(x) = \sum_{i=1}^{n} Z_i \]

where

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

\[ E Z_i = 0, \ |Z_i| \leq 2 \tilde{K}_1(nh_n^d)^{-1}, \ E |Z_i| \leq \tau n^{-1} \text{ and } E Z_i^2 \leq v^{-2} h_n^{-d} \] where \( \tau \) and \( v \) are two positive constants.

And, we write

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

Now, we write

\[ W_{n,i} = Z_i \Rightarrow |Z_i| \leq \alpha n \] and \[ V_{n,i} = Z_i \Rightarrow |Z_i| > \alpha n \] for \( \alpha > 1 \) and \( 1 \leq i \leq n \).
Then,
\[ \left| \sum_{i=1}^{n} Z_i \right| \leq \left| \sum_{i=1}^{n} (W_{ni} - EW_{ni}) \right| + \left| \sum_{i=1}^{n} V_{ni} \right| + \left| \sum_{i=1}^{n} EW_{ni} \right|. \]  
(3.1)

We need to show the following
\[ \sum_{n=1}^{\infty} P \left( \delta_n \sum_{i=1}^{n} (W_{ni} - EW_{ni}) > n^\alpha \epsilon / 3 \right) < \infty \]  
(3.2)

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

\[ \delta_n n^{-\alpha} \sum_{i=1}^{n} EW_{ni} \rightarrow 0, n \rightarrow \infty. \]  
(3.4)

We start by showing (3.2).

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

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 \( \beta \) such that \( \beta > (1 + \delta + \tau d) / \alpha \).

Now, we show (3.3).

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

hence,
\[ \sum_{n=1}^{\infty} P \left( \delta_n \sum_{i=1}^{n} V_{ni} > n^\alpha \epsilon / 3 \right) \leq \sum_{n=1}^{\infty} n \delta_n P (|Z_i| > n^\alpha) \leq \sum_{n=1}^{\infty} n \delta_n E|Z_i|^{\beta / n^\alpha \beta} \leq c_3 \sum_{n=1}^{\infty} n^{\delta + \tau d - \beta} < \infty \]

if we choose \( \beta, \delta_n \), and \( h_n \) as above and where \( c_3 \) is a positive constant.

Lastly we show that (3.4) holds.

We can write
\[ \delta_n n^{-\alpha} \sum_{i=1}^{n} EW_{ni} \leq \delta_n n^{-\alpha} \sum_{i=1}^{n} EV_{ni} = \delta_n n^{-\alpha} \sum_{i=1}^{\infty} E|Z_i| \Rightarrow ||Z_i| > n^\alpha \| = \delta_n n^{-\alpha} E|Z_i| \Rightarrow ||Z_i| > n^\alpha \| \rightarrow 0. \]

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

\[ \sum_{i=1}^{n} Z_i(x) \leq \frac{1}{nh_n^d} \sum_{i=1}^{n} \left[ K_1 \left( \frac{x - X_i}{h_n} \right) - K_1 \left( \frac{x_{nj} - X_i}{h_n} \right) \right] + \]
\[ \frac{1}{nh_n^d} \sum_{i=1}^{n} E \left[ K_1 \left( \frac{x - X_i}{h_n} \right) - K_1 \left( \frac{x_{nj} - X_i}{h_n} \right) \right] + \]
\[ \frac{1}{nh_n^d} \sum_{i=1}^{n} \left[ K_1 \left( \frac{x_{nj} - X_i}{h_n} \right) - E K_1 \left( \frac{x_{nj} - X_i}{h_n} \right) \right]. \]
The first and the second term in the right-hand side of the inequality above are to be considered similarly and we have:

\[
\frac{1}{nh_n^d} \sum_{i=1}^{n} \left[ K_1 \left( \frac{x - X_i}{h_n} \right) - K_1 \left( \frac{x_{nj} - X_i}{h_n} \right) \right] \leq \frac{L_K}{h_n^{d + \gamma_1}} \left| x - x_{nj} \right|^{\gamma_1} \leq \frac{L_K}{h_n^{d + \gamma_1}} c_1 n^{\gamma_1} = \frac{1}{\text{Log}n}
\]

where \(\mu_n\) is chosen such that

\[
\mu_n = \frac{L_1^{1/\gamma_1} c_n (\text{Log}n)^{1/\gamma_1}}{h_n^{d/\gamma_1 + 1}} \rightarrow \infty.
\]

Then:

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

For all \(n \geq \epsilon_1(n)\) and for all \(\epsilon_n > 0\)

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

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

\[
\delta_n \sup_{x \in C_n} |g_n(x) - Eg_n(x)| \overset{a.s.}{\rightarrow} 0, \ n \rightarrow \infty.
\]

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

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

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:

\[
n^n \sup_{x \in C_n, y \in R} |f_n(x, y) - f(x, y)| \overset{a.s.}{\rightarrow} 0, \ n \rightarrow \infty.
\]

**Proof.** We write

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

where

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

and

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

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

Now, let us write

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

And we write

\[ W_{ni} = Z_i \Rightarrow \| Z_i \| \leq n^a \]
\[ V_{ni} = Z_i \Rightarrow \| Z_i \| > n^a \]

for \( \alpha > 1 \) and \( 1 \leq i \leq n \).

Then,

\[ \sum_{i=1}^{n} Z_i \leq \sum_{i=1}^{n} (W_{ni} - EW_{ni}) + \sum_{i=1}^{n} V_{ni} + \sum_{i=1}^{n} EW_{ni} \] (3.5)

We need to show the following

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

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

\[ n^{a-\alpha} \left| \sum_{i=1}^{n} EW_{ni} \right| \rightarrow 0, n \rightarrow \infty. \] (3.8)

We start by showing (3.6).

The Markov inequality provides:

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

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

where \( c_1 \) and \( c_2 \) are two positive constants and \( \beta \) such that \( \beta > 2 \).

Now, we show (3.7).

Note that

\[ \left( \sum_{i=1}^{n} V_{ni} \right) > n^a \epsilon/3 \subseteq \bigcup_{i=1}^{n} (\| Z_i \| > n^a) \]

hence,

\[ \sum_{n=1}^{\infty} P\left( n^a \left| \sum_{i=1}^{n} V_{ni} \right| > n^a \epsilon/3 \right) \leq \sum_{n=1}^{\infty} n^{1+a} P(\| Z_i \| > n^a) \leq \sum_{n=1}^{\infty} n^{1-a\beta} E|Z_i|^\beta \leq c_3 \sum_{n=1}^{\infty} n^{-\alpha\beta-1+a} h_n^{-(d+1)} < \infty \]

if we choose \( h_n = n^{-r/4} \) and where \( c_3 \) is a positive constant and \( \beta \) such that \( \beta > (a + d\tau + \tau)/\alpha \) with \( a > 0 \) and \( \alpha > 1 \).

Lastly we show that (3.8) holds.

We can write:

\[ n^{a-\alpha} \sum_{i=1}^{n} EW_{ni} \leq n^{a-\alpha} \left| \sum_{i=1}^{n} EV_{ni} \right| = n^{a-\alpha} E|Z_i| \Rightarrow (\| Z_i \| > n^a) \rightarrow 0, n \rightarrow \infty \]

with \( \alpha > a \).
Next, we cover \( C_n \) by \( \mu_n^d \) spheres in the shape of \( \{ x : ||x - x_{nj}|| \leq c_n \mu_n^{-1} \} \) 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 \( \Upsilon_i(\cdot, y) = \frac{1}{nh_n^{d+1}} K_2 \left( \frac{y - Y_n}{h_n} \right) K_1 \left( \frac{-X_i}{h_n} \right) \).

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:

\[
\sup_{x \in C_n} \sup_{y \in R} \left| \sum_{i=1}^{n} [\Upsilon_i(x, y) - \Upsilon_i(x_{nj}, y)] \right| \leq \\
\frac{L_K \bar{K}}{h_n^{d+1+\gamma_1}} ||x - x_{nj}|| \gamma_1 \leq \\
\frac{L_K \bar{K}}{h_n^{d+1+\gamma_1}} \mu_n^{-\gamma_1} \frac{1}{\log n},
\]

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

Thus,

\[
\sup_{x \in C_n} \sup_{y \in R} \left| \sum_{i=1}^{n} Z_i(x, y) \right| \leq \\
\sup_{1 \leq j \leq \mu_n^d} \sup_{y \in R} \left| \sum_{i=1}^{n} [\Upsilon_i(x_{nj}, y) - E\Upsilon_i(x_{nj}, y)] \right| + \frac{2}{\log n},
\]

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

\[
P \left\{ \sup_{x \in C_n} \sup_{y \in R} \left| \sum_{i=1}^{n} Z_i(x, y) \right| > 2\epsilon \right\} \leq \\
\sum_{j=1}^{\mu_n^d} P \left\{ \sup_{y \in R} \left| \sum_{i=1}^{n} [\Upsilon_i(x_{nj}, y) - E\Upsilon_i(x_{nj}, y)] \right| > \epsilon \right\}.
\]

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 \( v_n = \frac{1}{h_n^{d}} \) with \( \mu \) being a positive constant.

Then we have:

\[
\sup_{y \in R} \left| \sum_{i=1}^{n} [\Upsilon_i(x_{nj}, y) - E\Upsilon_i(x_{nj}, y)] \right| \leq \sup_{|y| \leq v_n} |\Delta_n(x_{nj}, y)| + \sup_{|y| > v_n} |\varphi_n(x_{nj}, y)|.
\]
Cover \([-v_n, v_n]\) by \(l_n\) spheres \(B_s\) with centers \(t_s\) and radii less than or equal to \(h_n\), where \(l_n \leq v_n h_n^{-\eta}\) and \(\eta\) is a fixed number. Then using same arguments to those used previously we obtain:

\[
\sup_{|y| \leq v_n} |\Delta_n(x_nj, y)| \leq \lambda_0 h_n^{-(\eta-1/2)} a.s.,
\]

where \(\Delta_n(x_nj, y) = \Delta_n(x_nj, \tilde{Y}) - \Delta_n(x_nj, t_s)\) and \(\lambda_0\) 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

\[
n^a \sup_{x \in C} \sup_{|y| \leq v_n} \left| \sum_{i=1}^{n} Z_i(x, y) \right| \overset{a.s.}{\to} 0.
\]

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

\[
\sup_{|y| > v_n} |\varphi_n(x_nj, y)| \leq \sup_{|y| > v_n} \left| \sum_{i=1}^{n} T_i(x_nj, y) \right| + \sup_{|y| > v_n} \left| \sum_{i=1}^{n} \epsilon Y_i(x_nj, y) \right|,
\]

and by the compactness of the support of \(K_2\),

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

Therefore,

\[
n^a \sup_{|y| > v_n} \left| \sum_{i=1}^{n} T_i(x_nj, y) \right| \leq \frac{n^a}{n h_{d+1}} \tilde{K}^2 \sum_{i=1}^{n} \Rightarrow |Y| > v_n / 2 \cdot
\]

(3.10)

We need the use of the following in our proof

\[
P(|Y| > v_n / 2) \leq (2v_n^{-1}) \xi (E|Y| \xi)
\]

(3.11)

for a certain \(\xi > 0\) such that \(\xi > \mu \gamma_1 (\eta - 1)\).

For all \(\epsilon > 0\), we have

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

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

\[
P \left\{ n^a \sup_{|y| > v_n} \left| \sum_{i=1}^{n} T_i(x_nj, y) \right| > \epsilon \right\} \leq \epsilon^{-1} \tilde{K}^2 n^a h_n^{-d-1} (2v_n^{-1}) \xi (E|Y| \xi) = \epsilon^{-1} \tilde{K}^2 n^a h_n^{-d-1+\frac{\xi}{2}} 2\xi (E|Y| \xi).
\]

Inequality (3.9) implies:

\[
P \left\{ n^a \sup_{x \in C, |y| > v_n} \left| \sum_{i=1}^{n} Z_i(x, y) \right| > \epsilon \right\} \leq An^a \mu_d h_n^{-d-1+\frac{\xi}{2}} (E|Y| \xi),
\]

where \(A\) is a positive constant.
The choice of $\xi$ and the assumptions of the theorem permit us to conclude that:

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

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

$$n^a \sup_{x \in \mathbb{R}} \sup_{y \in \mathbb{R}} |T_n(x, y)| \longrightarrow 0, \ n \rightarrow \infty.$$  

To this end:

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

with

$$E \left\{ K_2 \left( \frac{y - Y_i}{h_n} \right) K_1 \left( \frac{x - X_i}{h_n} \right) \right\} = \int_{\mathbb{R}} \int_{\mathbb{R}} K_2 \left( \frac{y - v}{h_n} \right) K_1 \left( \frac{x - u}{h_n} \right) f(x, y, u, v) \, du \, dv.$$  

Properties of the Bochner's integral permit to write

$$T_n(x, y) = \frac{1}{n^d+1} \int_{\mathbb{R}} \int_{\mathbb{R}} K_2 \left( \frac{y - v}{h_n} \right) K_1 \left( \frac{x - u}{h_n} \right) f(x, y, u, v) \, du \, dv - f(x, y).$$

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

$$T_n(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} K_2(z_2) K_1(z_1) f(x, y, x-z_1h_n, y-z_2h_n) \, dz_1 \, dz_2 - f(x, y).$$

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

$$n^a \sup_{x \in \mathbb{R}} \sup_{y \in \mathbb{R}} |T_n(x, y)| \longrightarrow 0, \ n \rightarrow \infty.$$

4 Proof of the main result

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

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

$$\leq \sup_{y \in \mathbb{R}} |f_n(y|x) - f(y|x)| + \sup_{y \in \mathbb{R}} |f_n(y|x) - f(y|x)|$$

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

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

$$P \left( \sup_{x \in C_n} |\Theta_n(x) - \Theta(x)| \geq \epsilon_n \right) \leq P \left( \sup_{x \in C_n} \sup_{y \in \mathbb{R}} |f_n(y|x) - f(y|x)| \geq \beta_n \right),$$

which completes the proof of Theorem 2.1.

Acknowledgement

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