Journal of the Korean Statistical Society 45 (2016) 1–13

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Uniform in bandwidth rate of convergence of the conditional mode estimate on functional stationary ergodic data

Naâmane Laïb, Djamal Louani *

L.S.T.A., Université de Paris 6 and Université de Reims Champagne Ardenne, France

ARTICLE INFO

Article history:
Received 18 July 2014
Accepted 2 May 2015
Available online 27 May 2015

AMS 2000 subject classifications:
primary 60F10
secondary 62G07

Keywords:
Bandwidth interval
Conditional mode estimation
Ergodic processes
Functional data
Martingale difference
Strong consistency

ABSTRACT

The aim of this paper is to establish the uniform consistency with rate over a bandwidth interval of the kernel conditional mode estimate whenever functional stationary ergodic data are considered. This kind of result is immediately applicable to proving uniform consistency of kernel-type estimators when the bandwidth h is a function of the data or the location x. Notice that our uniform in bandwidth results are the first ones to be established in this setting. Moreover, the ergodic setting offers a more general framework in regards to the practice than the usual mixing structure.

1. Introduction

Let (X, Z) be a $E \times F$-valued random elements, where E and F are some semi-metric abstract spaces. Denote by d_E and d_F semi-metrics associated to spaces E and F respectively. Let $\mathcal{C}$ be a class of real functions defined upon F. Obviously, for any $\varphi \in \mathcal{C}$, $\varphi(Z)$ is a real random variable. Suppose now that we observe a sequence $(X_i, Z_i)_{i\geq 1}$ of copies of (X, Z) that we assume to be stationary and ergodic. For any $x \in E$ and any $\varphi \in \mathcal{C}$, let $g_\varphi(.|x)$ be the conditional density of $\varphi(Z)$ given $X = x$. We assume that $g_\varphi(.|x)$ is unimodal on some compact $S_\varphi \subset \mathbb{R}$. The conditional mode is defined, for any fixed $x \in E$, by

$$\Theta_\varphi(x) = \arg\sup_{y\in S_\varphi} g_\varphi(y|x).$$

Note that, if there exists $\xi > 0$ such that for any $\varphi \in \mathcal{C}$

$$g_\varphi(.|x) \uparrow \quad \text{on } (\Theta_\varphi(x) - \xi, \Theta_\varphi(x)) \quad \text{and} \quad g_\varphi(.|x) \downarrow \quad \text{on } (\Theta_\varphi(x), \Theta_\varphi(x) + \xi), \tag{1}$$

and if we choose $S_\varphi = [\Theta_\varphi(x) - \xi,\ \Theta_\varphi(x) + \xi]$, then the mode $\Theta_\varphi(x)$ is uniquely defined for any φ. The kernel estimator, say $\hat{\Theta}_{\varphi,n}(x)$, of $\Theta_\varphi(x)$ may be defined as the value maximizing the kernel estimator $g_{\varphi,n}(y|x)$ of $g_\varphi(y|x)$, that is,

$$g_{\varphi,n}(\hat{\Theta}_{\varphi,n}(x)|x) = \sup_{y\in S_\varphi} g_{\varphi,n}(y|x).$$

* Corresponding author.
E-mail address: djamal.louani@upmc.fr (D. Louani).

http://dx.doi.org/10.1016/j.jkss.2015.05.001
1226-3192/

Here,

$$g_{\varphi,n}(y|x) = \frac{\sum_{i=1}^{n}\left[K\left(\frac{d(x,X_i)}{h}\right)H\left(\frac{y-\varphi(Z_i)}{h}\right)\right]}{h\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{h}\right)},$$

where K and H are two real valued kernels and $h := h_n$ is a sequence of positive real numbers tending to zero as $n \to \infty$.

In a recent work, Chaouch, Laïb, and Louani (2014) have studied the rate of uniform consistency with respect to the function parameter $\varphi \in \mathcal{C}$, of the conditional mode estimator $\hat{\Theta}_{\varphi,n}(x)$, with application to electricity consumption. The aim of this paper is to establish the uniform consistency (for a fixed φ), relatively to the smoothing parameter h taking values in some set H_n, of the conditional mode estimator $\hat{\Theta}_{\varphi,n}(x)$ when data are assumed to be sampled from a stationary and ergodic processes. More precisely, under suitable conditions upon the rate of convergence of the smoothing parameter h together with some regularity conditions on the distribution of the random element (X, Z), we obtain results of type

$$\sup_{h_n \in H_n} |\hat{\Theta}_{\varphi,n}(x) - \Theta_{\varphi}(x)| = O(\beta_n), \quad \text{a.s.},$$

where β_n is a quantity to be specified later on. Motivations to consider ergodic data are discussed in Laïb (2005) and Laïb and Louani (2010) where details defining the ergodic property of processes together with examples of such processes are also given.

Indexing by a function φ allows to consider simultaneously various situations related to model fitting and time series forecasting. Whenever $Z := \{Z(t) : t \in T\}$ denotes a process defined on some real set T, one may consider the following functionals $\varphi_1(Z) = \sup_{t\in T} Z(t)$ and $\varphi_2(Z) = \inf_{t\in T} Z(t)$ giving extremes of the process Z that are of interest in various domains as, for example, the finance, hydraulics, energy and the weather forecasting. For some weight function W defined on T and some $p > 0$, one may consider the functional $\varphi_{p,W}$ defined by $\varphi_{p,W}(Z) = \int_T W(t)Z^p(t)dt$. Further situation is to consider, for some subset A of T, the functional $Z \to \varphi_\rho(Z) = \inf\{t \in A : Z(t) \geq \rho\}$ for some threshold ρ. Such a case is very useful in threshold and barrier crossing problems encountered in various domains as finance, physical chemistry and hydraulics. Application aspects to electricity consumption are carried out in Chaouch et al. (2014).

In practice, one has to choose a bandwidth sequence h in such a way that the bias and the variance parts are reasonably balanced. The optimal choice of h often yields to quantities that depend on some unknown parameters of the distribution which has to be estimated. This leads to bandwidth sequences depending on the data and the estimation location x. As a consequence, nonuniform results do not apply if one is interested in estimators with such general bandwidth sequences. Our uniform in h results make it possible to establish consistency of estimators when the bandwidth h is allowed to range in an interval which may increase or decrease in length with the sample size. These kinds of results are immediately applicable to proving the uniform consistency of kernel-type estimators when the bandwidth h is a function of the data or the location x, see below for details. Notice that our uniform in bandwidth results are the first ones to be established in this setting.

The modelization of the functional variable is becoming more and more popular since the publication of the monograph of Ramsay and Silverman (1997) on functional data analysis. Note however that the first results dealing with nonparametric models (mainly the regression function) were obtained by Ferraty and Vieu (2000). Since then, an increasing number of papers on this topic has been published. One may refer to the monograph by Ferraty and Vieu (2006) and the recent work by Chaouch et al. (2014) for an overview on the subject and the references therein. Extensions to other regression issues as the time series prediction have been carried out in a number of publications, see for instance Delsol (2009). The general framework of ergodic functional data has been considered by Laïb and Louani (2010, 2011) who stated consistencies with rates together with the asymptotic normality of the regression function estimate.

2. Results

In order to state our results, we introduce some notations. Let $\mathcal{F}_i$ be the σ-field generated by $((X_1, Z_1), \ldots, (X_i, Z_i))$ and $\mathcal{G}_i$ the one generated by $((X_1, Z_1), \ldots, (X_i, Z_i), X_{i+1})$. Let $B(x, u)$ be a ball centered at $x \in E$ with radius u. Let $D_i(x) := d(x, X_i)$ so that $D_i(x)$ is a nonnegative real-valued random variable. Working on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $F_x(u) = \mathbb{P}(D_i(x) \leq u) := \mathbb{P}(X_i \in B(x, u))$ and $F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(D_i(x) \leq u|\mathcal{F}_{i-1}) = \mathbb{P}(X_i \in B(x, u)\ |\mathcal{F}_{i-1})$ be the distribution function and the conditional distribution function, given the σ-field $\mathcal{F}_{i-1}$, of $(D_i(x))_{i\geq 1}$ respectively. Denote by $o_{\text{a.s.}}(u)$ a real random function l such that $l(u)/u$ converges to zero almost surely as $u \to 0$. Similarly, define $O_{\text{a.s.}}(u)$ as a real random function l such that $l(u)/u$ is almost surely bounded.

Our results are stated under some assumptions we gather hereafter for easy reference.

A1 For $x \in E$, there exists a sequence of nonnegative random functionals $(f_{i,1})_{i\geq 1}$ almost surely bounded by a sequence of deterministic quantities $(b_i(x))_{i\geq 1}$ accordingly, a sequence of random functions $(\psi_{i,x})_{i\geq 1}$, a deterministic nonnegative bounded functional f_1 and a nonnegative nondecreasing real function ϕ tending to zero as its argument goes to zero, such that