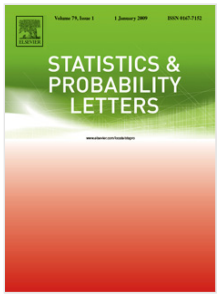
A note on the conditional density estimate in single functional index model

Said Attaoui, Ali Laksaci, Elias Ould Said

Accepted Manuscript



A note on the conditional density estimate in single functional index model

Said Attaoui, Ali Laksaci, Elias Ould Said

PII: S0167-7152(10)00267-1
DOI: 10.1016/j.spl.2010.09.017
Reference: STAPRO 5803

To appear in: *Statistics and Probability Letters*

Received date: 24 March 2010
Revised date: 17 September 2010
Accepted date: 22 September 2010

Please cite this article as: Attaoui, S., Laksaci, A., Ould Said, E., A note on the conditional density estimate in single functional index model. *Statistics and Probability Letters* (2010), doi:10.1016/j.spl.2010.09.017

A note on the conditional density estimate in single functional index model

Said Attaoui, Ali Laksaci
Département de Mathématiques
Université Djillali Liabès
BP 89, 22000, Sidi Bel Abbès, Algérie
e-mail:(s_attaoui or alilak)@yahoo.fr

Elias Ould Said[1,2],*
[1] Univ. Lille Nord de France
F-59000 Lille, France
[2] ULCO, LMPA, F-62228, Calais, France
e-mail: ouldsaid@lmpa.univ-littoral.fr

September 17, 2010

Abstract

In this paper, we consider estimation of the conditional density of a scalar response variable Y given a Hilbertian random variable X when the observations are linked with a single-index structure. We establish the pointwise and the uniform almost complete convergence (with the rate) of the kernel estimate of this model. As an application, we show how our result can be applied in the prediction problem via the conditional mode estimate. Finally, the estimation of the functional index via the pseudo-maximum likelihood method is also discussed but not attacked.

Key words: Conditional single-index, conditional density, nonparametric estimation, semiparametric estimation, semi-metric choice.

1 Introduction

For the past two decades, the single-index model, a special case of projection pursuit regression, has proven to be an efficient way of coping with the high dimensional problem in nonparametric regression. Here we deal with single-index modeling when the explanatory variable is functional. More precisely, we consider the problem of estimating the conditional density of a real variable Y given a functional variable X when the explanation of Y given X is done through its projection on one functional direction.

The conditional density plays an important role in nonparametric prediction, because the several prediction tools in nonparametric statistic, such as the conditional mode, the conditional median or the conditional quantiles, are based on the preliminary estimate of this functional parameter. Nonparametric estimation of the conditional density has been widely studied, when the data is real. The first related result in nonparametric functional statistic was obtained by Ferraty *et al.* (2006). They established the almost complete consistency in the independent and identically distributed (i.i.d.) random variables of the kernel estimator of the conditional probability density. The asymptotic normality of this kernel estimator has been studied in the dependent data by Ezzahrioui and Ould Saïd (2010).

The single-index approach is widely applied in econometrics as a reasonable compromise between nonparametric and parametric models. Such kind of modelization is intensively studied in the multivariate case. Without pretend to exhaustivity, we quote for example Härdle *et al.* (1993), Hristache *et al.* (2001). Based on the regression function, Delecroix *et al.* (2003) studied the estimation of the single-index and established some asymptotic properties. The

*corresponding author

literature is strictly limited in the case where the explanatory variable is functional (that is a curve). The first asymptotic properties in the fixed functional single-model were obtained by Ferraty *et al.* (2003). They established the almost complete convergence, in the i.i.d. case, of the link regression function of this model. Their results were extended to dependent case by Aït Saidi *et al.* (2005). Aït Saidi *et al.* (2008) studied the case where the functional single-index is unknown. They proposed an estimator of this parameter, based on the cross-validation procedure.
The goal of this paper is to study the estimation of the conditional density in the single functional index model. We construct an estimator of this model by a kernel method and we prove, under general conditions, its pointwise and uniform almost complete convergence (with rate). This extend, to functional case, the work of Delecroix *et al.* (2003) and complete the results of Ferraty *et al.* (2006) and Ferraty *et al.* (2010) by studying the conditional density in the semi-parametric functional settings. In practice, this study has great importance, because, it permit us to construct a prediction method based on the conditional mode estimator. Moreover, in the case where the functional single index is unknown, our estimate can be used to estimate this parameter via the pseudo-maximum likelihood estimation method. Noting that the estimation of the functional single-index has great interest on the semi-metric choice in nonparametric functional data analysis but it has been not attacked in this paper.
The paper is organized as follows. We present our model in Section 2. In Section 3 we introduce notations, assumptions and state the main results. Section 4 is devoted to some discussions and comments on the impact of our study in the prediction problem. The proofs of the results are relegated to the last section.

2 Model

Let (X, Y) be a couple of random variables taking its values in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a Hilbertian space with scalar product $< \cdot, \cdot >$. Let $(X_i, Y_i)_{1 \leq i \leq n}$, be n copies of independent vectors each having the same distribution as (X, Y). Assume that the explanation of Y given X is done through a fixed functional index θ in $\mathcal{F}$. In the sense that, there exists an θ in $\mathcal{F}$ (unique up to a scale normalization factor) such that

$$\mathbb{E}[Y|X] = \mathbb{E}[Y| < \theta, X >]. \tag{1}$$

The identifiability of this model has been studied in Ferraty et al. (2003). They supposed that the regression operator $r(x) = \mathbb{E}[Y|X = x]$ is differentiable and θ such that $< \theta, e_1 >= 1$, where e_1 is the first vector of an orthonormal basis of $\mathcal{F}$.
Noting that, similarly to the multivariate case, the single functional index approach is very efficient way to reduce the effect of the infinite dimensional feature of the nonparametric estimation in functional statistic. The main aim of this work is the estimation of the conditional density of Y given $< \theta, x >$, denoted by $f(\theta, \cdot, x)$. It is well known that, in nonparametric statistics, this latter provides an alternative approach to study the links between Y and X and it can be also used, in single index modelling, to estimate the functional index θ if it is unknown.

Naturally, the kernel estimator $\widehat{f}(\theta, y, x)$ of $f(\theta, y, x)$ is defined by

$$\widehat{f}(\theta, y, x) = \frac{h_H^{-1} \sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >)) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >))}, \qquad \forall y \in \mathbb{R}$$

with the convention $0/0 = 0$. The functions K and H are kernels and $h_K := h_{K,n}$ (resp. $h_H := h_{H,n}$) is a sequence of positive real numbers which goes to zero as n tends to infinity. Note that a similar estimate was already introduced in the special case where X is a real random variable by Delecroix *et al.* (2003).

3 Main results

All along the paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. In the following, we put, for any $x \in \mathcal{F}$, and $i = 1, \ldots, n$,

$$K_i(\theta, x) := K(h_K^{-1} < x - X_i, \theta >) \text{ and, for all } y \in \mathbb{R},\ H_i(y) := H(h_H^{-1}(y - Y_i))$$

3.1 Pointwise almost complete convergence

In he following x is a fixed point in $\mathcal{F}$, $\mathcal{N}_x$ is a fixed neighborhood of x and $\mathcal{C}$ is a fixed compact subset of $\mathbb{R}$. In order to establish the almost complete (a.co.)[1]convergence of our estimate we need the following assumptions:

(H1) $\mathbb{P}(|<X-x,\theta>|<h) =: \phi_{\theta,x}(h) > 0$.

(H2) The conditional density $f(\theta,y,x)$ satisfies the Hölder condition, that is:
$\forall (y_1,y_2)\in\mathcal{C}^2$, $\forall (x_1,x_2)\in\mathcal{N}_x\times\mathcal{N}_x$,

$$|f(\theta,y_1,x_1)-f(\theta,y_2,x_2)|\leq C_{\theta,x}\left(\|x_1-x_2\|^{b_1}+|y_1-y_2|^{b_2}\right),\quad b_1>0,\, b_2>0.$$

(H3) K is a positive bounded function with support $[-1,1]$.

(H4) H is bounded function, such that

$$\int H(t)dt=1,\qquad \int |t|^{b_2}H(t)dt<\infty \text{ and } \int H^2(t)dt<\infty,$$

(H5) The bandwidths h_K and h_H satisfy

$$\lim_{n\to\infty}\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}=0.$$

Comments on the assumptions
Our assumptions are very standard for this kind of model. Assumptions (H1) and (H3) are the same as those used in Ferraty *et al.* (2003). Assumption (H2) is a regularity conditions which characterize the functional space of our model and is needed to evaluate the bias term of our asymptotic results. Assumptions (H4) and (H5) are technical conditions and are also similar to those in Ferraty *et al.* (2006).

Our first main result is given in the following theorem.

Theorem 3.1 *Under Assumptions (H1)-(H5), and for any fixed y, we have, as n goes to infinity*

$$\left|\widehat{f}(\theta,y,x)-f(\theta,y,x)\right| = O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co.}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}}\right). \tag{2}$$

In the particular case, where the real random variable $Z:=<X,\theta>$ has continuous density we can reformulate the general result given in Theorem 3.1 in the following way

Corollary 3.1 *Under Assumptions (H2)-(H5) and if the density of Z does not vanish and for any fixed y, we have, as n goes to infinity*

$$\left|\widehat{f}(\theta,y,x)-f(\theta,y,x)\right| = O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co.}\left(\sqrt{\frac{\log n}{nh_Hh_K}}\right).$$

Proof of Theorem 3.1.
The proof is based on the following decomposition

$$\begin{aligned}\widehat{f}(\theta,y,x)-f(\theta,y,x) &= \frac{1}{\widehat{f}_D(\theta,x)}\left\{\left(\widehat{f}_N(\theta,y,x)-\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right]\right)+\left(\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right]-f(\theta,y,x)\right)\right\}\\ &\quad -\frac{f(\theta,y,x)}{\widehat{f}_D(\theta,x)}\left\{\widehat{f}_D(\theta,x)-1\right\}\end{aligned}$$

[1]We say that a sequence Z_n converges *a.co.* to Z if and only if, for any $\epsilon>0$, $\sum_n \mathbb{P}\left(|Z_n-Z|>\epsilon\right)<\infty$.

where

$$\widehat{f}_N(\theta,y,x)=\frac{1}{nh_H\mathbb{E}\left[K_1(\theta,x)\right]}\sum_{i=1}^{n}K_i(\theta,x)H_i(y),\qquad \widehat{f}_D(\theta,x)=\frac{1}{n\mathbb{E}\left[K_1(\theta,x)\right]}\sum_{i=1}^{n}K_i(\theta,x).$$

So, the proof is a direct consequence of the following results

Lemma 3.1 *(see Aït-Saidi et al., 2005) Under Assumptions (H1), (H3) and (H5), as n goes to infinity, we have*

$$\left|\widehat{f}_D(\theta,x)-1\right|=O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right). \tag{3}$$

Furthermore, we have

$$\sum_{n=1}^{\infty}\mathbb{P}\Big(\left|\widehat{f}_D(\theta,x)\right|\leq 1/2\Big)<\infty. \tag{4}$$

Lemma 3.2 *Under Assumptions (H1)-(H5), as n goes to infinity, we have*

$$\left|\widehat{f}_N(\theta,y,x)-\mathbb{E}\Big[\widehat{f}_N(\theta,y,x)\Big]\right|=O_{a.co.}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}}\right). \tag{5}$$

Lemma 3.3 *Under Assumptions (H1)-(H5), as n goes to infinity, we have*

$$\left|\mathbb{E}\Big[\widehat{f}_N(\theta,y,x)\Big]-f(\theta,y,x)\right|=O(h_K^{b_1})+O(h_H^{b_2}). \tag{6}$$

3.2 Uniform almost complete convergence

This section is devoted to derivation of the the uniform version of Theorem 3.1. The study of the uniform consistency is motivated by the fact that the latter is an indispensable tool for studying the asymptotic properties of all estimate of the functional index if is unknown. Noting that, in the multivariate case, the uniform consistency is a standard extension of the pointwise one, however, in our functional case, it requires some additional tools and topological conditions (see Ferraty *et al.*, 2009, for more discussion on the uniform convergence in nonparametric functional statistics). Thus, in addition to the conditions introduced in the previous section, we need the following ones. Firstly, we suppose that $\mathcal{C}$ is subset compact of $\mathbb{R}$ and $S_{\mathcal{F}}$ (resp. $\Theta_{\mathcal{F}}$, the space of parameters) are such that

$$S_{\mathcal{F}}\subset\bigcup_{k=1}^{d_n^{S_{\mathcal{F}}}}B(x_k,r_n)\qquad\text{and}\quad\Theta_{\mathcal{F}}\subset\bigcup_{j=1}^{d_n^{\Theta_{\mathcal{F}}}}B(t_j,r_n) \tag{7}$$

with x_k (resp. t_j) $\in\mathcal{F}$ and $r_n,d_n^{S_{\mathcal{F}}},d_n^{\Theta_{\mathcal{F}}}$ are sequences of positive real numbers which tend to infinity as n goes to infinity.
Furthermore, we need the following assumptions:

(U1) There exists a differentiable function $\phi(.)$ such that $\forall x\in S_{\mathcal{F}}$, and $\forall\theta\in\Theta_{\mathcal{F}}$,

$$0<C\phi(h)\leq\phi_{\theta,x}(h)\leq C'\phi(h)<\infty\ \text{ and }\ \exists\eta_0>0,\ \forall\eta<\eta_0,\ \phi'(\eta)<C,$$

(U2) The conditional density is such that $\forall(y_1,y_2)\in\mathcal{C}\times\mathcal{C},\ \forall(x_1,x_2)\in S_{\mathcal{F}}\times S_{\mathcal{F}}$, and $\forall\theta\in\Theta_{\mathcal{F}}$,

$$|f(\theta,y_1,x_1)-f(\theta,y_2,x_2)|\leq C\left(\|x_1-x_2\|^{b_1}+|y_1-y_2|^{b_2}\right),$$

(U3) The kernel K satisfy (H3) and Lipschitz's condition holds

$$|K(x) - K(y)| \leq C\|x - y\|,$$

(U4) H is a bounded Lipschitz continuous function, such that

$$\int H(t)dt = 1, \qquad \int |t|^{b_2} H(t)dt < \infty \text{ and } \int H^2(t)dt < \infty,$$

(U5) For some $\gamma \in (0,1)$, $\lim_{n\to+\infty} n^{\gamma} h_H = \infty$, and for $r_n = O\left(\frac{\log n}{n}\right)$ the sequences $d_n^{S_{\mathcal{F}}}$ and $d_n^{\Theta_{\mathcal{F}}}$ satisfy:

$$\frac{(\log n)^2}{nh_H\,\phi(h_K)} < \log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}} < \frac{nh_H\,\phi(h_K)}{\log n},$$

and

$$\sum_{n=1}^{\infty} n^{(3\gamma+1)/2} (d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}})^{1-\beta} < \infty, \text{ for some } \beta > 1.$$

Remark 3.1 *Note that Assumptions (U1) and (U2) are, respectively, the uniform version of (H1) and (H2). Assumption (U4) is condition (H4) added by contion Lipschitz condition. Assumptions (U1) and (U5) are linked with the the topological structure of the functional variable. For examples of subsets such as (7) see Ferraty et al. (2009).*

Theorem 3.2 *Under Assumptions (U1)-(U5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{x\in S_{\mathcal{F}}} \sup_{y\in\mathcal{C}} \left|\widehat{f}(\theta,y,x) - f(\theta,y,x)\right| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{nh_H\phi(h_K)}}\right). \tag{8}$$

In the particular case, where the functional single-index is fixed we get the following result.

Corollary 3.2 *Under Assumptions (U1)-(U5), we have, as n goes to infinity*

$$\sup_{x\in S_{\mathcal{F}}} \sup_{y\in\mathcal{C}} \left|\widehat{f}(\theta,y,x) - f(\theta,y,x)\right| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}}}{nh_H\phi(h_K)}}\right).$$

Proof. Clearly Theorem 3.2 and Corollary 3.2 can be deduced from the following intermediate results which are uniform version of Lemmas **3.1-3.3**.

Lemma 3.4 *Under Assumptions (U1), (U3) and (U5), we have as $n \to \infty$*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{x\in S_{\mathcal{F}}} \left|\widehat{f}_D(\theta,x) - 1\right| = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right).$$

Corollary 3.3 *Under the assumptions of Lemma 3.4, we have,*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{F}}} \inf_{x\in S_{\mathcal{F}}} \widehat{f}_D(\theta,x) < \frac{1}{2}\right) < \infty.$$

Lemma 3.5 *Under Assumptions (U1), (U2) and (H4), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{x\in S_{\mathcal{F}}} \sup_{y\in\mathcal{C}} \left|f_N(\theta,y,x) - \mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right]\right| = O(h_K^{b_1}) + O(h_H^{b_2}).$$

Lemma 3.6 *Under the assumptions of Theorem 3.2, we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{x\in S_{\mathcal{F}}} \sup_{y\in\mathcal{C}} \left|\widehat{f}_N(\theta,y,x) - \mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{nh_H\phi(h_K)}}\right).$$

4 Some applications and comments

4.1 The conditional mode in functional single-index model

Let us now study the estimation of the conditional mode in the functional single-index model. Our main aim, here, is to establish the a.co. convergence of the kernel estimator of the conditional mode of Y given θX, denoted by $M_\theta(x)$, uniformly on fixed subset $S_{\mathcal{F}}$ of $\mathcal{F}$. For this, we assume that $M_\theta(x)$ satisfies on S the following uniform uniqueness property

(U6) $\forall \epsilon_0 > 0\ \exists \eta > 0,\ \forall \upsilon : S_{\mathcal{F}} \to \mathcal{C}$,

$$\sup_{x\in S_{\mathcal{F}}} |M_\theta(x) - \upsilon(x)| \geq \epsilon_0 \ \Rightarrow\ \sup_{x\in S_{\mathcal{F}}} |f(\theta, \upsilon(x), x) - f(\theta, M_\theta(x), x)| \geq \eta.$$

Moreover, we also suppose that there exists some integer $j > 1$ such that $\forall x \in S$ the function $f(\theta, \cdot, x)$ is j times continuously differentiable w.r.t. y on $\mathcal{C}$ and

(U7)
$$\left\{ \begin{array}{c} f^{(l)}(\theta, M_\theta(x), x) = 0 \text{ if } 1 \leq l < j \\ \text{and } \ f^{(j)}(\theta, \cdot, x) \text{ is uniformly continuous on } \mathcal{C} \\ \text{such that } \ |f^{x(j)}(\theta, \cdot, x)| > C > 0 \end{array} \right.$$

where $f^{(j)}(\theta, \cdot, x)$ is the j^{th} order derivative of the conditional density $f(\theta, \cdot, x)$.
We estimate the conditional mode $M_\theta(x)$ with a random variable $\widehat{M_\theta(x)}$ such that

$$\widehat{M_\theta(x)} = \arg\sup_{y\in\mathcal{C}} \widehat{f}(\theta, y, x).$$

From Corollary 3.2 we derive the following result.

Corollary 4.1 *Under the assumptions of Theorem 3.2 and if the conditional density $f(\theta, \cdot, x)$ satisfies (U6) and (U7), we have*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{M_\theta(x)} - M_\theta(x)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}}}{n^{1-\gamma}\phi(h_K)}}\right).$$

4.2 Application to prediction

Let us now define the application framework of our results to prediction problem. For each $n \in \mathbb{N}^*$, let $(X_i(t))_{t\in\mathbb{R}}$ $i = 1, \ldots, n$ be a Hilbertian random variable. For each curve $(X_i(t))_{t\in\mathbb{R}}$, we have a real response variable Y_i. We suppose that the observations $(X_i, Y_i)_{1\leq i\leq n}$ are generated with single-index structure. The prediction aim is to evaluate y_{new} given $(X_{n+1}(t))_{t\in\mathbb{R}} = x_{new}$. The estimation of the conditional mode in functional single-index model shows that the random variable $\widehat{M_\theta(x_{new})}$, is the best approximation of y_{new}, given x_{new}. Applying the result in the above corollary, we obtain the following result.

Corollary 4.2 *Under the assumptions of Corollary 4.1, we have as n goes to infinity*

$$\widehat{M_\theta(x_{new})} - M_\theta(x_{new}) \to 0 \quad a.co.$$

4.3 On the estimation of the functional single index

Another way to highlight the interest of our study is to show how the conditional density estimate can be used to derive an estimate of the functional single index if the latter is unknown. The estimation of the functional single index has been extensively studied in the multivariate case. In the functional case, Aït-Saidi *et al.* (2008) adopt the leave-out-one-curve cross-validation procedure. Their ideas can be combined with those of Laksaci *et al.*

(2010), on the smoothing parameter choice in conditional density estimation, in order to construct an estimator of θ asymptotically optimal. Of course the convergence of this estimator to the index θ request some additional model assumptions. Alternatively, this parameter can be estimated via the pseudo-maximum likelihood method which is based on the preliminary estimate of the conditional density of Y given X by

$$\widehat{\theta} = \arg\max_{\theta\in\Theta_{\mathcal{F}}} \widehat{L}(\theta)$$

with

$$\widehat{L}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log \widehat{f}(\theta, Y_i, X_i)\tau(Y_i, X_i)$$

where $\tau(Y_i, X_i)$ is a trimming function introduced to guard against small values for the denominators
This method has been studied by Delecroix *et al.* (2003) in the real case where they showed that this technique has minimal variance among all estimators. The asymptotic optimality of this procedure in functional statistic, is an important prospect of the present work. Noting that, technically, the main step is to show that the functional index θ defined by (1) is the unique solution of the following optimization problem

$$\max_{\theta} \mathbb{E}[\log f(\theta, Y, X)\tau(Y, X)].$$

It should be noted that our asymptotic result is pivotal step for the both methods.
As an application, this approach can be used for answering the semi-metric choice question. Indeed, it is well known that, in nonparametric functional statistic, the projection-type semi-metric is very interesting for increasing the concentration property. The functional index model is a particular case of this family of semi-metric, because it is based on the projection on one functional direction. So, the estimation procedures of this direction permit us to compute adaptive semi-metrics in the general context of nonparametric functional data analysis.

5 Appendix

The Lemmas 3.2, 3.3 and 3.5, are special cases of the Lemmas 2.3.4, 2.3.5 in Ferraty et al (2006) and Lemma 14 in Ferraty et al (2010) with $d(x_1, x_2) =< x_1 - x_2, \theta >$ respectively. The proof of Lemma 3.4 and 3.6 are also very close to those of Lemmas 8 and 15 in Ferraty *et al.* (2010) and they are therefore presented in a shorter fashion.
Proof of Lemma 3.4
For all $x \in S_{\mathcal{F}}$, and for all $\theta \in \Theta_{\mathcal{F}}$ we set

$$k(x) = \arg\min_{k\in\{1,\dots r_n\}} \|x - x_k\| \text{ and } j(\theta) = \arg\min_{j\in\{1,\dots \ell_n\}} \|\theta - t_j\|.$$

We consider the following decomposition

$$\begin{aligned}
\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\widehat{f}_D(\theta,x) - \mathbb{E}\left[\widehat{f}_D(\theta,x)\right]\right| \leq & \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\widehat{f}_D(\theta,x) - \widehat{f}_D(\theta,x_{k(x)})\right|}_{T_1} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\widehat{f}_D(\theta,x_{k(x)}) - \widehat{f}_D(t_{j(\theta)},x_{k(x)})\right|}_{T_2} + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\widehat{f}_D(t_{j(\theta)},x_{k(x)}) - \mathbb{E}\left[\widehat{f}_D(t_{j(\theta)},x_{k(x)})\right]\right|}_{T_3} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\mathbb{E}\left[\widehat{f}_D(t_{j(\theta)},x_{k(x)})\right] - \mathbb{E}\left[\widehat{f}_D(\theta,x_{k(x)})\right]\right|}_{T_4} + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{\theta\in\Theta_{\mathcal{F}}} \left|\mathbb{E}\left[\widehat{f}_D(\theta,x_{k(x)})\right] - \mathbb{E}\left[\widehat{f}_D(\theta,x)\right]\right|}_{T_5}.
\end{aligned}$$

For T_1, T_2, we use the Hölder continuity condition on K, the Cauchy-Schwartz's inequality and the Bernstein's inequality. With theses arguments we get

$$T_1 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \text{ and } T_2 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right). \tag{9}$$

Moreover, using the fact that $T_4 \leq T_1$ and $T_5 \leq T_2$ to get, for n tending to infinity

$$T_4 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right) \quad \text{and} \quad T_5 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right). \tag{10}$$

Now, we deal with T_3. For all $\eta > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(T_3 > \eta\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right) \\
&\leq d_n^{S_\mathcal{F}} d_n^{\Theta_\mathcal{F}} \max_{k\in\{1\ldots d_n^{S_\mathcal{F}}\}} \max_{j\in\{1\ldots d_n^{\Theta_\mathcal{F}}\}} \mathbb{P}\left(\left|\widehat{f}_D(t_{j(\theta)}, x_{k(x)}) - \mathbb{E}\left[\widehat{f}_D(t_{j(\theta)}, x_{k(x)})\right]\right| > \eta\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right).
\end{aligned}$$

Applying Bernstein's exponential inequality to $\Delta_i = \frac{1}{\phi(h_K)}\Big\{K_i(t_{j(\theta)}, x_{k(x)}) - \mathbb{E}\left[K_i(t_{j(\theta)}, x_{k(x)})\right]\Big\}$ one get, under (U7),

$$T_3 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right).$$

Finally the result can be easily deduced from the latter together with (9) and (10). ∎

Proof of Lemma 3.6 We keep the same notations as in Lemma 3.4 and we use the compactness of $\mathcal{C}$. We can write

$$\mathcal{C} \subset \bigcup_{k=1}^{z_n} (y_j - \ell_n, y_j + \ell_n)$$

with $\ell_n = n^{-\frac{3}{2}\gamma - \frac{1}{2}}$ and $z_n \leq C\, n^{\frac{3}{2}\gamma + \frac{1}{2}}$. Taking

$$j(y) = \arg \min_{j\in\{1,2,\ldots,z_n\}} |y - t_j|.$$

We get the following decomposition:

$$\begin{aligned}
\left|\widehat{f}_N(\theta,y,x) - \mathbb{E}\Big[\widehat{f}_N(\theta,y,x)\Big]\right| \leq\; & \underbrace{\left|\widehat{f}_N(\theta,y,x) - \widehat{f}_N(\theta,y,x_{k(x)})\right|}_{F_1} + \underbrace{\left|\widehat{f}_N(\theta,y,x_{k(x)}) - \widehat{f}_N(t_{j(\theta)},y,x_{k(x)})\right|}_{F_2} \\
& + \underbrace{\left|\widehat{f}_N(t_{j(\theta)},y,x_{k(x)}) - \widehat{f}_N(t_{j(\theta)},y_{j(y)},x_{k(x)})\right|}_{F_3} + \underbrace{\left|\widehat{f}_N(t_{j(\theta)},y_{j(y)},x_{k(x)}) - \mathbb{E}\left[\widehat{f}_N(t_{j(\theta)},y_{j(y)},x_{k(x)})\right]\right|}_{F_4} \\
& + \underbrace{\left|\mathbb{E}\left[\widehat{f}_N(t_{j(\theta)},y_{j(y)},x_{k(x)})\right] - \mathbb{E}\left[\widehat{f}_N(t_{j(\theta)},y,x_{k(x)})\right]\right|}_{F_5} + \underbrace{\left|\mathbb{E}\left[\widehat{f}_N(t_{j(\theta)},y,x_{k(x)})\right] - \mathbb{E}\left[\widehat{f}_N(\theta,y,x_{k(x)})\right]\right|}_{F_6} \\
& + \underbrace{\left|\mathbb{E}\Big[\widehat{f}_N(\theta,y,x_{k(x)})\Big] - \mathbb{E}\Big[\widehat{f}_N(\theta,y,x)\Big]\right|}_{F_7}.
\end{aligned}$$

Using the same ideas as for T_1, T_2, T_4 and T_5, permit us to get, , for n tending to infinity

$$F_7 \leq F_1 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\,\phi(h_K)}}\right), \qquad \text{and} \qquad F_6 \leq F_2 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\,\phi(h_K)}}\right). \tag{11}$$

Concerning the terms F_3 and F_5, using Lipschitz's condition on the kernel H, permits us to write,

$$\left|\widehat{f}_N(t_{j(\theta)}, y, x_{k(x)}) - \widehat{f}_N(t_{j(\theta)}, y_{j(y)}, x_{k(x)})\right| \leq \frac{\ell_n}{h_H^2\,\phi(h_K)}.$$

Now, the fact that $\lim_{n\to+\infty} n^{\gamma} h_H = \infty$ and choosing $\ell_n = n^{-\frac{3}{2}\gamma-\frac{1}{2}}$ imply that

$$\frac{\ell_n}{h_H^2\,\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H\,\phi(h_K)}}\right).$$

Hence, for n large enough, we have

$$F_5 \leq F_3 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H\,\phi(h_K)}}\right). \tag{12}$$

Finally, the evaluation of the term (F_4) is very close to (T_3) in Lemma 3.4. Applying Bernstein's exponential inequality to

$$\Gamma_i = \frac{1}{h_H\,\phi(h_K)}\left[K_i(x_k)H_i(t_j) - \mathbb{E}\left(K_i(x_k)H_i(t_j)\right)\right],$$

it follows that

$$F_4 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H\,\phi(h_K)}}\right). \tag{13}$$

So, the Lemma can be easily deduced from (11)-(13). ■

Acknowlegdments. The authors would like to thank the editor and the referee for her/his careful reading and relevant remarks which permit us to improve the paper.

References

Aït Saidi, A., Ferraty, F., Kassa, R., 2005. Single functional index model for a time series. R. Roumaine Math. Pures et Appl. 50, 321–330.

Aït Saidi, A., Ferraty, F., Kassa, R., Vieu, P., 2008. Cross-validated estimation in the single functional index model. Statistics 42, 475–494.

Delecroix, M, Härdle, W., Hristache, M., 2003. Efficient estimation in conditional single-index regression. J. Multivariate Anal. 86, 213–226.

Ezzahrioui, M., Ould Saïd, E., 2010. Some asymptotic results of a nonparametric conditional mode estimator for functional time series data. Statist. Neerlandica 64, 171-201 .

Ferraty, F., Laksaci, A., Vieu, P., 2006. Estimating some characteristics of the conditional distribution in nonparametric functional models. Statist. Inf. for Stoch. Proc. 9, 47–76.

Ferraty, F., Laksaci, A., Tadj, A., Vieu, P., 2010. Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plann. and Inf. 140, 335–352.

Ferraty, F., Peuch, A., Vieu, P., 2003. Modèle à indice fonctionnel simple. C. R. Mathématiques Paris 336, 1025–1028.

Ferraty, F., Vieu, P., 2006. Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics, Springer, New York.

Härdle, W., Hall, P., Ichumira, H., 1993. Optimal smoothing in single-index models, Ann. Statist. 21, 157–178.

Hristache, M., Juditsky, A., Spokoiny, V. 2001. Direct estimation of the index coefficient in the single-index model. Ann. Statist. 29, 595–623.

Laksaci, A., Lemdani, M., Ould Saïd, E., 2009. A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors: Consistency and asymptotic normality. Statist. & Probab. Lett. 79, 1065–1073.
Laksaci, A., Madani, F., Rachdi., M. 2010. Bandwidth selection for kernel conditional density estimation when the regressor is valued in a semi-metric space, International Statistical Review (accepted).
Rosenblatt, M., 1969. Conditional probability density and regression estimators. In Multivariate Analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London.