République Algérienne Démocratique et Populaire
Ministère de l'enseignement supèrieur et de la recherche scientifique

Université de Saida - Dr Moulay Tahar.
Faculté des Sciences.
Département de Mathématiques.

Mémoire présenté en vue de l'obtention du diplome de

Master Académique

Filière : MATHÉMATIQUES
Spécialité: Analyse Stochastique Statistique des Processus et Applications

par
Khatir **N**adia[1]

Sous la direction de

Dr.Mme. F.Benziadi

Thème:

Conditional models with simple functional index

Soutenue le 14/06/2022 devant le jury composé de

Mme. Dr. Ait ouali .N	Université de Saida Dr. Moulay Tahar	Présidente
Mme. Dr. Benziadi.F	Université de Saida Dr. Moulay Tahar	Encadreur
Mme. Pr. Rahmani.S	Université de Saida Dr. Moulay Tahar	Examinatrice

Année univ.: 2021/2022

[1]e-mail: nadiakhatir259@gmail.com

⁂ *Remerciements* ⁂

Nous remercions Dieu (sans Lui rien n'aurait pu être possible) tout puissant de nous avoir accordé santé, courage et la volonté pour accomplir ce modeste travail.

Je tiens avant tout à exprimer ma reconnaissance à **Dr.FATIMA Benziadi** pour avoir accepté de m'encadrer dans cette étude. Je la remercie pour son implication, son soutien et ses encouragements tout au long de ce travail.

J'adresse mes sincères remerciements à tous Les membres du laboratoire des Modèles Stochastiques, Statistique et Applications et à la tête le directeur

Pr.A. Kandouci.

Aux membres du jury
Présidente du Jury : **Mme. Dr.Ait ouali .N**
Examinatrice : **Mme. Pr.Rahmani.S**

Messieurs, vous nous faites un grand honneur en acceptant de juger ce travail.

Je remercie ma famille pour son soutien. Merci à ma mère et mon père. Enfin, merci à RABHA de m'avoir si bien soutenu tout au long de mes années de la licence et de master.

⁂ *Dédicace* ⁂

Quoi de plus merveilleux que de pouvoir partager les meilleurs moments de ma vie avec les personnes que j'aime.

J'ai le grand paisir de dédier ce travail.

A mon père Ahmed , mon example éternel, celui qui s'est toujours sacrifié pour me voir réussir.

A ma mère Arbia pour son amour, ses sacrificeset sa tendresse.

A ma soeur et mes Frères Bouteina, Mohamed Amine, Farouk Abd El Ileh.

A mes trés chers amie Rabha.

A mes meilleurs amis : Abla, Fatima, Bouchra, Alaa, Aimen, Abd El Madjid, ...

Et pour finir, merci à toutes les personnes que j'ai oubliées de citer et qui m'ont permis de mener à bien ce mémoire.

Contents

Introduction

Over the last two decades, an immense innovation on measuring instruments has emerged and realized that enabling several objects to be monitored continuously, such as stock market indices, pollution, climatology, satellite images,..., this technological development required the modernization of statistical methods as tools for analysis and control. Thus, a new branch of statistics, called functional statistics, has been developed to treat observations as functional random elements. The first contributions on the subject were devoted to the study of parametric models (see, e.g. Ramsay and Silverman 2005,[39]), such as parametric models are models in which the vector of parameters is a vector in finite dimensional space). Our interest in this case is estimating the vector of parameters. In parametric models, the researcher assumes completely the form of the model and its assumptions. However, statistical analysis via linear models is based on a preliminary knowledge of the nature of covariability between observations, which is very difficult to verify in functional statistic, contrary to the classical statistic where graphic tools are available such as the scatterplot which gives an overview on the relation between the observations. This justifies the importance of modeling functional data by nonparametric methods.

Nonparametric processing of functional data is much more recent than parametric analysis. The term nonparametric does not mean that such models are completely lack parameters, but that the number of the parameters are flexible and not fixed periori. The primary interest is in estimating that infinite-dimensional vector of parameters. In nonparametric regression models, the relationship between the explanatory variables and response is unknown.

Semi-parametric modeling is a hybrid of the parametric and nonparametric approaches of statistical models. It may appear at first that semi-parametric model include nonparametric model, however, It is considered to be "smaller" than a completely nonparametric model because we are often interested only in the finite-dimensional component. By contrast, in nonparametric models, the primary interest is in estimating the infinite dimensional parameter. In result, the estimation is statistically harder in nonparametric models compared to semi-parametric models. While parametric models are being easy to understand and easy to work with, they fail to give a fair representation of what is happening in the real world.

Semi-parametric models allow you to have the best of both worlds, a model that is understandable and offering a fair representation of the messiness that is involved in real life. Beside their contribution in reducing the dimensionality of the covariate's space and therefore damping the curse of dimensionality effect in the estimation procedure, we briefly recall the method SIR introduced by Li (1991) and Duan and Li (1991), this method makes it possible to estimate the parametric part of the semi-parametric models considered without having to estimate the functional part ni to specify the law of the error ε. One possibility is to be interested to the conditional distribution of y (dependent variable knowing x multidimensional explanatory variable). Dimension reduction methods assume that we can replace x by a vector of lower dimension $(\grave{x}\theta_1, ..., \grave{x}\theta_K)$, with K $<$ p, without losing information on the link between y and x. We then seek to estimate a basis of the effective dimension reduction subspace (generated by the vectors θ_k, $k = 1, ..., K$). The basic idea of SIR methods is to exchange the role of x and y (in order to reduce the dimension of the problem) and to study the conditional moments of x knowing y. SIR methods are based on a geometric property of the curve of inverse regression which relies on a crucial condition of linearity of the distribution of the covariate. For a recent survey on semi-parametric literature for infinite dimensional variables, the reader can be referred to Vieu (2018,[48]) and Goia and Vieu (2014,[26]). Semi-parametric models include, but not limited to, functional single-index models (see, e.g. Li et al. 2010; Goia and Vieu 2015,[27]), projection pursuit models (see, e.g. Bali et al. 2011[8]; Chen et al. 2011[10]; Ferraty et al. 2013[16]), partial linear models (see, e.g. Aneiros-Prez and Vieu 2011[3]; Lian 2011[32]; Maity and Huang 2012,[35]; Aneiros-Pérez and

Vieu 2015[4]) and functional sliced inverse regression (see Ferré and Yao 2005). For a broader overview on functional data analysis, the reader can be referred to Aneiros et al. (2019,[2]), Hsing and Eubank (2015,[28]), Cuevas (2014,[11]), Ferraty and Vieu (2006,[25]) and Ramsay and Silverman (2002)[38]. The extension of the single-index model to the functional data framework was first introduced in Ferraty et al. (2003,[17]) to estimate, semi-parametrically, the regression operator where the response variable is real-valued and the covariate is a functional random variable. The single functional index model (SFIM) assumes that a functional explanatory variable acts on a scalar response only through its projection on one functional direction. The SFIM was intensively extended to estimate several statistical parameters describing the shape of the conditional distribution. For instance, Ait-Saïdi et al. (2008,[1]) used SFIM to estimate the regression operator and suggested to use a cross-validation procedure to estimate the unknown link function as well as the unknown single functional index. Furthermore,Attaoui (2014,[7]) studied the estimation of the conditional density Goia and Vieu (2015,[27]) introduced a semi-parametric methodology, which approximates the unknown regression operator through a single index approach, taking possible structure changes into account.
The study of a variable Y conditioned by a variable X is a very important subject in statistics. In nonparametric stat, regression is the main tool for study that kind of variables. However, this tool is not very suitable for certain situations, like the conditional density actually contains more information than the regression function, which is simply the conditional expectation and this letter is affected by the existence of outliers in the sample that we are studying. The purpose of this dissertation is to study some conditional models in the case where the explanatory variable is functional or infinitely dimensioned in single functional index. The memory is orgnised as follows, the first chapter presents the non-parametric estimation of conditional models for functional variables. In next chapters , we treat the uniform almost complete convergence of the kernel estimators of the conditional models in functional index model :conditional density function,conditional distribution function,and their application the conditional mode (resp.the conditional quantile). Finally, the last chapter is devoted on the estimation of the functional single index. We end this work with a

general conclusion and some future perspectives.

Chapter 1

Estimation of conditional models

1.1 Parametric estimation:

The parametric statistics remonstrates with Fisher 1920, it is the "classical" framework of statistics. The statistical model is described there by a finite number of parameters. Typically $\{P_\theta, \theta \in \mathbb{R}^p\}$ is the statistical model that describes the distribution of observed random variables. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X an v.a. from $(\Omega, \mathcal{A})$ into (E, ε). The data of a statistical model is the data of a family of probabilities on (E, ε), $\{P_\theta, \theta \in \Theta\}$.
Given the model, we then assume that the law of X belongs to the model $\{P_\theta, \theta \in \Theta\}$.

1.1.1 Definitions

Definition 1.1.1. *Let $g : \Theta \mapsto \mathbb{R}^k$. We call estimator of $g(\theta)$ in view of observation X, any application $T : \Omega \to \mathbb{R}^k$ of the form $T = h(X)$ where $h : E \to \mathbb{R}^k$ measurable.*

An estimator must not depend on the quantity $g(\theta)$ that we seek to to estimate. We introduce the following properties of an estimator:

Definition 1.1.2. *T is an unbiased estimator of $g(\theta)$ if for all $\theta \in \Theta$, $E_\theta[T] = g(\theta)$. Otherwise, we say that the estimator T is biased and we calls bias the quantity $E_\theta[T] - g(\theta)$.*

Generally X is a vector $(X_1, ..., X_n)$ of observations (n being the number of them). An important example is the case where $X_1, ..., X_n$ form an n-sample, i.e. when $X_1, ..., X_n$ are i.i.d. We can then look at the asymptotic properties of the estimator, stretch the number of observations n towards $+\infty$ .In this case, it is natural to note $T = T_N$ as dependent on n. We then have the following definition:

Definition 1.1.3. *T_n is a consistent estimator of $g(\theta)$ if for all $\theta \in \Theta$, T_n converges in probability to $g(\theta)$ under P_θ when $n \to \infty$.*

We define the quadratic risk of the estimator in the case where $g(\theta) \in \mathbb{R}$.

Definition 1.1.4. *Let T_n be an estimator of $g(\theta)$. The quadratic risk of T_n is defined by*

$$R(T_n, g(\theta)) = E_\theta[(T_n - g(\theta))^2]$$

The quadratic risk is the sum of the variance and the square through the estimator.

1.1.2 Estimation methods

There are several methods of parametric estimation, in this section we will see estimation by the method of moments and estimation by maximum likelihood.

Estimation by the method of moments

X is the vector formed by an n-sample $X_1, ..., X_n$. The X_i are values in a set $\mathcal{X}$. Let $f = (f_1, ..., f_k)$ be an application from $\mathcal{X}$ to $\mathbb{R}^k$ such as the application

$$\Phi : \Theta \to \mathbb{R}^K$$

$$\theta \mapsto E_\theta[f(X_1)]$$

either injective. We define the estimator $\hat{\theta}_n$ as the solution in Θ (when it exists) of the equation

$$\Phi(\theta) = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

Often, when $\mathcal{X} \subset \mathbb{R}$ the function on takes $f_i(x) = x^i$ and Φ corresponds therefore at the ith moment of the variable X_1 under P_θ. This choice justifies the name given to the method.Here is an example of estimators built on this method:

Uniform Law:

Here $k = 1$, Q_θ is the uniform law on $[0, \theta]$ with $\theta > 0$. We only have for all θ,$E_\theta[X_1] = \frac{\theta}{2}$, we can therefore take for example $\Phi(\theta) = \frac{\theta}{2}$ and $f = Id : \mathbb{R} \to \mathbb{R}$. The estimator obtained by the method of moments is then $\hat{\theta}_n = 2\overline{X}_n$. This estimator is unbiased and consistent.

Estimation by maximum white likelihood

Let $\{E, \varepsilon, \{P_\theta, \theta \in \Theta\}\}$ be a statistical model, where $\Theta \subset \mathbb{R}^k$ (we are in a parametric framework). We assume that there is a σ-finite measure μ which dominates the model, that is to say that $\forall \theta \in \Theta$, P_θ admits a density $p(\theta, .)$ by compared to μ.

Definition 1.1.5. *Let X be an observation. We call likelihood of X the application*

$$\Theta \to \mathbb{R}$$

$$\theta \mapsto p(\theta, X)$$

We call maximum likelihood estimator of θ any element $\hat{\theta}$ of Θ maximizing the likelihood, i.e. verifying

$$\hat{\theta} = \arg\max_{\theta \in \Theta} p(\theta, X)$$

Consider the typical case where $X = (X_1, ..., X_n)$, the X_i forming an n sample of law Q_θ where Q_θ s a law on X of unknown parameter $\theta \in \Theta \subset \mathbb{R}^k$. We further assume that for all $\theta \in \Theta$,Q_θ is absolutely continuous by with respect to a measure ν on X. In this case, noting

$$q(\theta, x) = \frac{dQ_\theta}{d\nu}$$

we have that the likelihood is written

$$p(\theta, X) = \prod_{i=1}^{n} q(\theta, X_i)$$

and so

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n} \log[q(\theta, X_i)]$$

Let's see an example:

Bernoulli's model

Let $Q_{\theta_0} = \mathcal{B}(\theta)$ with $\theta \in]0,1[= \Theta$ and ν the counting measure on $\mathbb{N}$. For all $\theta \in]0,1[$ and $x \in \{0,1\}$

$$q(\theta, x) = \theta^x(1-\theta)^{1-x} = (1-\theta)\exp[x\log(\frac{\theta}{1-\theta})]$$

and so the maximum likelihood estimator must maximize in $[0,1]$

$$\log(\theta^{S_n}(1-\theta)n - S_n) = S_n\log(\frac{\theta}{1-\theta}) + n\log(1-\theta)$$

Which leads to $\hat{\theta}_n = \overline{X}$.

1.2 Nonparametric estimation of the conditional density for functional variables:

Non-parametric statistics is concerned with the estimation from a finite number of observations, of an unknown function $f \in \Theta$, where Θ is a fairly large functional space.

Let X and Y be two continuous random variables .The conditional probability density function of Y given $X = x$ given by this formula

$$f_{Y|X=x} : \mathbb{R} \to [0, \infty[$$

$$y \to \frac{f_{XY}(x,y)}{f_X(x)}$$

with : f_{XY} is the density function, and f_X is the marginal function.

In nonparametric regression setting, it is a well-known fact that for forecasting and

statistical inferences, the conditional density function is very useful. It provides the one of the best tool to estimate some characteristic feature of the dataset, such as the conditional mode. Indeed, this last has received a considerable attention, see, for instance, Collomb et al. (1987), Samanta and Thavaneswaran (1990,[45]), Quintela-del-Rio and Vieu (1997,[28]), Berlinet et al.(1998,[9]), and among others.
Estimation of the conditional density function and its derivatives, in statistics functional, was introduced by Ferraty et al (2006,[20]). These authors obtained the almost complete convergence in the i.i.d.The precision of the leading terms of the squared error of the kernel estimator of the conditional density was obtained by Laksaci (2007,[30]). We refer to Laksaci et al. (2010,[31]) for the question of the choice of smoothing parameter in the estimation of the conditional density with a functional explanatory variable. Ali Laksaci et al. (2013,[2]) used the polynomial method local.

1.2.1 Kernel conditional density estimates

Let $(X_1, Y_1)...(X_n, Y_n)$ be a random independent indentically distributed (i.i.d) sample of the pair (X, Y) ,which is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a functional space. the conditional density estimator is given

$$\hat{f}(Y|X = x) = \frac{\sum_{i=1}^{n} K(h_K^{-1} d(x, X_i)) H(h_H^{-1}(y - Y_i))}{h_H \sum_{i=1}^{n} K(h_K^{-1} d(x, X_i))} \tag{1.1}$$

And d is a semi-metric.with K is a kernel ,and $h_K = h_K(n)$(respg=$h_H(n)$)is a sequence of positive real numbers which tends to 0 when n to infinity (it is also called the smoothing parameter).
Ferraty and al.(2005,[21]) established the almost complete convergence of a kernel estimator of the conditional mode defined by the random variable maximizing the conditional density. Alternatively, Ez zahrioui and Ould-Said (2005, 2006,[14],[15]) estimated the conditional mode by the point which cancels the derivative of the kernel

estimator of the conditional density. Those the latter have focused on the asymptotic normality of the estimator proposed in both cases (i.i.d. and mixing).

1.2.2 The conditional mode:

The mode is very popular in classification, because it is a useful tool for representing groups and also robust sweaters than the average (like the median).Mode estimation is often a direct consequence of density estimation. His importance is due to the fact that it is a natural measure of central tendency, which is not influenced by the tails of the distributions. The mode is the most probable value: for a density f, it is the value for which f admits a maximum (global or local). For a symmetric distribution, it coincides with two other positional parameters, the mean and the median.

Presentation of the model:

Consider $(X_i, Y_i)_{i=1...n}$ a sample of couple of random variables (X, Y), the conditional mode is the value that maximizes the conditional density Y knowing $X = x$ .We assume that there is a compact subset $s \in \mathbb{R}$, where the mode is unique denoted by $\theta(x)$:

$$\theta(x) = \arg\sup_{y \in s} f^x(y)$$

The estimator of this mode $\hat{\theta}(x)$ is defined by the expression :

$$\hat{\theta}(x) = \arg\sup_{y \in s} \hat{f}^x(y)$$

With $\hat{f}^x(y)$is defined by (1.1).

1.3 Nonparametric estimation of the conditional distribution function for functional variables:

Let X and Y be two continuous random variables.The conditional distribution function of Y given $X = x$ is a function $F_{Y|X=x}$ such that:

$$F_{Y|X=x} = \mathbb{P}(Y \in [a,b]|X = x) = \int_a^b f_{Y|X=x}(x)dx \tag{1.2}$$

1.3.1 Kernel estimate of the conditional distribution

the kernel estimator of the conditional distribution function $F_{Y|X=x}$ defined by $\forall y \in \mathbb{R}$.

$$\hat{F}_{Y|X=x} = \frac{\sum_{i=1}^{n} K(h_K^{-1} d(x, X_i)) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_K^{-1} d(x, X_i))} \tag{1.3}$$

Roussas (1969,[40]) treated the estimation of the conditional distribution function by the method kernel using Markovian observations. He established the convergence in probability of the constructed estimator. Stute (1986,[47]) added results on the almost complete convergence of the kernel estimator of the distribution function of a vector random variable conditional on a vector explanatory variable.The estimation of this in a functional framework was introduced by Ferraty et al (2006,[20]).They constructed a dual-kernel estimator for the conditional distribution function and they specified the speed of convergence almost completeness of this estimator when the observations are independent and identically distributed. The case of α -mixing observations has been studied by Ferraty et al (2005,[19]).

1.3.2 Conditional quantile

Several authors have dealt with the estimation of the distribution function conditional as a preliminary study of estimating quantiles conditionals. Let us quote for example, Ezzahrioui and Ould-Saïd (2005,2006,[14],[15]) who studied the asymptotic normality of this estimator in both cases (i.i.d. and α -mixing).

Presentation of the model:

Let X and Y be two continuous random variables,$F^x(y)$ is the conditional distribution function of Y given $X = x$,and $\theta \in]0,1[$.The conditional quantile (just note $t_\theta(x)$) is defined by:

$$t_\theta(x) = \inf\{y \in \mathbb{R} : F(Y|X \geq \theta)\} \tag{1.4}$$

The conditional quantile estimate is given for:

$$\hat{t}_\theta = \hat{F}^{-1}(Y|X) \tag{1.5}$$

1.4 Nonparametric estimation of the regression function for functional variables

We have $Y = r(X) + \varepsilon$,where the response variable Y is real-valued while the explanatory variable X has value in infinite dimensional semimetric space $(\mathcal{F}, d)$,we also assume that the variable ε (corresponding to the residual)fulfills $E(\varepsilon|X) = 0$ in such a way that the regression function defined by : $r(X) = E(Y|X = x)$, $r \in \mathcal{C}(\mathbb{R})$. The first results in functional nonparametric statistics were developed by Ferraty et Vieu (2000,[22]) and they concern the estimation of the regression function with an explanatory variable of fractal dimension. They established the almost complete convergence of a kernel estimator of this model not parametric in the i.i.d.Based on recent developments of small ball probability theory, Ferraty et Vieu (2004,[30]) have generalized these last results to the α -mixing case and exploited the importance of nonparametric modeling of functional data in applying their study to curve discrimination and forecasting.In the framework of α -mixing functional observations, Masry (2005,[34]) showed the asymptotic normality of the estimator of Ferraty and Vieu (2004,[30]) for the regression function.

1.4.1 Kernel estimate of the regression function

The kernel estimator of the regression function $\hat{r}$ is defined by:

$$\hat{r}^x(y) = \frac{\displaystyle\sum_{i=1}^{n} Y_i K\left(\frac{X_i - x}{h}\right)}{\displaystyle\sum_{i=1}^{n} k\left(\frac{X_i - x}{h}\right)}$$

.

Chapter 2

Conditional density with simple functional index

2.1 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 $< .,. >$. Let $(X_i, Y_i)_{1 \leq i \leq n}$,be n copies of independent vectors each having the same distribution as (X, Y).
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 $f(\theta, ., x)$.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 modeling, to estimate the functional index θ if it is unknown.
Naturally, the kernel estimator $\hat{f}(\theta, y, x)$ of $f(\theta, y, x)$ is defined by

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

with the convention $0/0 = 0$, and 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.
we will denote by C and $\grave{C}$ some strictly positive generic constants. In the following, we put, for any $x \in \mathcal{F}$, and $i = 1, ..., n$,

$$K_i(\theta, x) = K(h_K^{-1} < x - X_i, \theta >)$$

and, for all $y \in \mathbb{R}$,

$$H_i(y) = H(h_H^{-1}(y - Y_i)).$$

2.2 Uniform almost complete convergence

In this section we propose to study the uniform almost complete convergence of our estimator $\hat{f}$ for this, 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

$$S_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n^{S_{\mathcal{F}}}} B(x_k, r_n) \quad and \quad \Theta_{\mathcal{F}} \subset \bigcup_{j=1}^{d_n^{\Theta_{\mathcal{F}}}} B(t_j, r_n) \tag{2.1}$$

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, and we need the following assumptions.

2.2.1 Assumptions and results

(U1) There exists a differentiable function $\phi(.)$ with

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

such that $\forall x \in S_{\mathcal{F}}$,and $\forall \theta \in \Theta_{\mathcal{F}}$,

$$0 < C\phi(h) \leq \phi_{\theta,x}(h) \leq \grave{C}\phi(h) < \infty$$

and

$$\exists \eta_0 > 0, \quad \forall \eta < \eta_0 \quad , \grave{\phi}(\eta) < C,$$

(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 $\forall\theta \in \Theta_{\mathcal{F}}$.

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

(U3) The kernel K is a positive bounded function with support $[-1, 1]$ and Lipschitz's condition holds

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

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

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

(U5) For some $\gamma \in (0, 1)$,$\lim_{n\longrightarrow\infty} n^\gamma h_H = \infty$,and for $r_n = o(\frac{\log n}{n})$ 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

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

for some $\beta > 1$.

Comments on the assumptions:

Note that assumptions (U1) and (U2) are, respectively, the uniform version of $\mathbb{P}(| < X - x, \theta > | < h) = \phi_{\theta,x}(h) > 0$ and the conditional density $f(\theta, y, x)$ satisfies the Hölder condition, that is $\forall(x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, $\forall(y_1, y_2) \in \mathcal{C}^2$, with $\mathcal{N}_x$ is a fixed neighborhood of x and $\mathcal{C}$ is a fixed compact subset of $\mathbb{R}$

$$|f(\theta, y_1, x_1) - f(\theta, y_2, x_2)| \le C_{\theta,x}(||x_1 - x_2||^{b_1} + |y_1 - y_2|^{b_2}), b_1 > 0, b_1 > 0$$

in pointwise almost complete convergence.
Assumption (U4) is added by Lipschitz condition. Assumptions (U1)and (U5) are linked with the the topological structure of the functional variable.

Theorem 2.2.1. *Under Assumptions (U1)-(U5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta}\sup_{x\in S_{\mathcal{F}}}\sup_{y\in\mathcal{C}}|\hat{f}(\theta,y,x)-f(\theta,y,x)| = O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{nh_H\phi(h_K)}}\right) \tag{2.2}$$

Remark 2.2.1. *In the particular cases,*

1. *where the functional single-index is fixed we get the following result. Under Assumptions (U1)-(U5), we have, as n goes to infinity*

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

2. *In the α-mixing case, to establish the almost complete convergence we use the Fuck-Nagaev exponential-type inequality.*

Proof of Theorem2.2.1

The proof is based on the following decomposition

$$\hat{f}(\theta,y,x)-f(\theta,y,x) = \frac{1}{\hat{f}_D(\theta,x)}\{(\hat{f}_N(\theta,y,x)-E[\hat{f}(\theta,y,x)])+(E[\hat{f}_N(\theta,y,x)]-f(\theta,y,x))\}$$
$$-\frac{f(\theta,y,x)}{\hat{f}_D(\theta,x)}\{\hat{f}_D(\theta,x)-1\}$$

where

$$\hat{f}_N(\theta,y,x) = \frac{1}{nh_HE[K_1(\theta,x)]}\sum_{i=1}^{n}K_i(\theta,x)H_i(y),\ \hat{f}_D(\theta,x) = \frac{1}{nE[K_1(\theta,x)]}\sum_{i=1}^{n}K_i(\theta,x)$$

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

Lemma 2.2.1. *Under Assumptions (U1), (U3) and (U5), we have as $n\to\infty$*

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

Corollary 2.2.1. *Under the assumptions of Lemma2.2.1 , we have,*

$$\sum_{n=1}^{\infty} \mathbb{P}(\inf_{\theta\in\Theta_{\mathcal{F}}} \inf_{x\in S_{\mathcal{F}}} \hat{f}_D(\theta,x) < 1/2) < \infty$$

Lemma 2.2.2. *Under Assumptions (U1), (U2) and (U4), we have, as n goes to infinity*

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

Lemma 2.2.3. *Under the assumptions of Theorem2.2.1 , we have, as n goes to infinity*

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

2.2.2 Proof

Proof of Lemma2.2.1

For all $x \in S_{\mathcal{F}}$, and for all $\theta \in \Theta_{\mathcal{F}}$ we set

$$k(x) = \arg\min_{k\in\{1\ldots r_n\}} ||x - x_k||$$

and

$$j(\theta) = \arg\min_{j\in\{1\ldots l_n\}} ||\theta - t_j||$$

We consider the following decomposition

$$\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta} |\hat{f}_D(\theta,x) - E[\hat{f}_D(\theta,x)]| \leq \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta} |\hat{f}_D(\theta,x) - \hat{f}_D(\theta,x_{k(x)})|}_{T_1} +$$

$$\underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta_{\mathcal{F}}} |\hat{f}_D(\theta,x_{k(x)}) - \hat{f}_D(t_{j(\theta)},x_{k(x)})|}_{T_2} + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta_{\mathcal{F}}} |\hat{f}_D(t_{j(\theta)},x_{k(x)}) - E[\hat{f}_D(t_{j(\theta)},x_{k(x)})]|}_{T_3} +$$

$$\underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta_{\mathcal{F}}} |E[\hat{f}_D(t_{j(\theta)},x_{k(x)})] - E[\hat{f}_D(\theta,x_{k(x)})]|}_{T_4} + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{\theta\in\Theta_{\mathcal{F}}} |E[\hat{f}_D(\theta,x_{k(x)})] - E[\hat{f}_D(\theta,x)]|}_{T_5}$$

For T1, T2, we use the Hölder continuity condition on K, the Cauchy-Schwartz's inequality and the Bernstein's inequality [1]. With theses arguments we get

$$T_1 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \quad and \quad T_2 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \tag{2.3}$$

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

$$T_4 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \quad and \quad T_5 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \tag{2.4}$$

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

$$\mathbb{P}\left(T_3 > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

$$\leq$$

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

Applying Bernstein's exponential inequality to

$$\Delta_i = \frac{1}{\phi(h_K)}\{K_i(t_{j(\theta)}, x_{K(x)}) - E[K_i(t_{j(\theta)}, x_{K(x)})]\}$$

one get , under (U7),

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

[1] Let $W_1 \ldots W_n$ a sequence of random variables which are independent identically distributed, and $\sigma_n^2 = E(W_j^2)$. If there exists $M = M_n < \infty$ such that $W_1 \leq M$. Then we have $\forall \varepsilon > 0, \mathbb{P}\left(\frac{1}{n}\left|\sum_{j=1}^{n} W_j\right| > \varepsilon\right) \leq 2\exp\left(-\frac{\varepsilon^2 n/2}{\sigma^2+\varepsilon M}\right)$

If $\omega_n = n^{-1}\sigma_n^2 \log n$, such that $\lim_{n\to\infty \omega_n = 0}$ with $M/\sigma_n^2 < \infty$, then we have $\frac{1}{n}\sum_{j=1}^{n} W_j = O_{a.co}(\sqrt{\omega_n})$.

and

$$T_5 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right)$$

Finally the result can be easily deduced from the latter together with 2.3 and 2.4.

Proof of Lemma2.2.2

One has

$$E(\hat{f}_N(\theta, y, x)) - f(\theta, y, x) = E(\frac{1}{nh_H E[K_1(\theta, x)]}\sum_{i=1}^{n} K_i(\theta, x)H_i(y)) - f(\theta, y, x)$$

$$= \frac{1}{h_H E[K_1(\theta, x)]}E(K_1(\theta, x)H_1(y)) - f(\theta, y, x) \tag{2.5}$$

$$= \frac{1}{h_H E[K_1(\theta, x)]}E(K_1(\theta, x))E(H_1| < \theta, X_1 >) - f(\theta, y, x)$$

Moreover, we have:

$E(H_1| < \theta, X_1 >) = \int_{\mathbb{R}} H(h^{-1}(y-z))f(\theta, z, X_1)dz$

$$= \frac{1}{h_H}\int_{\mathbb{R}} H(h^{-1}(y-z))f(\theta, z, X_1)dz - f(\theta, y, x)$$

We pose:

$T = h_H^{-1}(y-z) \Rightarrow z = y - h_H T \quad \text{and} \quad dz = h_H T$

$$= \frac{1}{h_H}\int_{\mathbb{R}} H(T)f(\theta, y - h_H T, X_1)h_H dT - f(\theta, y, x)$$

$$= \int_{\mathbb{R}} H(T)[f(\theta, y - h_H T, X_1) - f(\theta, y, x)]dT$$

$$|E(\hat{f}_N(\theta, y, x)) - f(\theta, y, x)| \leq \int_{\mathbb{R}} H(T)|f(\theta, y - h_H T, X_1) - f(\theta, y, x)|dT$$

Finally, the use of (U2) implies that:

$$\leq C\int_{\mathbb{R}} H(T)(h_H^{b_1} + |t|_{b_2} h_H^{b_2})dT \tag{2.6}$$

Because this inequality is uniform on $(\theta, y, x) \in \Theta_{\mathcal{F}} \times S_{\mathcal{F}} \times S_{\mathbb{R}}$ and because of (U4), 2.2.2 is a direct consequence of 2.5,2.6 and of Corollary 2.2.1.

Proof of Lemma2.2.3

We keep the same notations as in Lemma 2.2.1 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:

$$|\hat{f}_N(\theta, y, x) - E[\hat{f}_N(\theta, y, x)]| \leq \underbrace{|\hat{f}_N(\theta, y, x) - \hat{f}_N(\theta, y, x_{k(x)})|}_{F_1} + \underbrace{|\hat{f}_N(\theta, y, x_{k(x)}) - \hat{f}_N(t_{j(\theta)}, y, x_{k(x)})|}_{F_2}$$

$$+ \underbrace{|\hat{f}_N(t_{j(\theta)}, y, x_{k(x)}) - \hat{f}_N(t_{j(\theta)}, y_{j(y)}, x_{k(x)})|}_{F_3} + \underbrace{|\hat{f}_N(t_{j(\theta)}, y_{j(y)}, x_{k(x)} - E[\hat{f}_N(t_{j(\theta)}, y_{j(y)}, x_{k(x)})]|)}_{F_4}$$

$$+ \underbrace{|E[\hat{f}_N(t_{j(\theta)}, y_{j(y)}, x_{k(x)})] - E[\hat{f}_N(t_{j(\theta)}, y, x_{k(x)})]|}_{F_5} + \underbrace{|E[\hat{f}_N(t_{j(\theta)}, y, x_{k(x)})] - E[\hat{f}_N(\theta, y, x_{k(x)})]|}_{F_6}$$

$$+ \underbrace{|E[\hat{f}_N(\theta, y, x_{k(x)})] - E[\hat{f}_N(\theta, y, x)]|}_{F_7}$$

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

$$F_7 \leq F_1 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H \phi(h_K)}}\right) \quad \textit{and} \quad F_6 \leq F_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H \phi(h_K)}}\right) \tag{2.7}$$

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

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

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

$$F_5 \leq F_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H \phi(h_K)}}\right) \tag{2.8}$$

Finally, the evaluation of the term (F_4) is very close to (T_3) in Lemma 2.2.1. Applying Bernstein's exponential inequality to

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

it follows that

$$F_4 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n h_H \phi(h_K)}}\right) \tag{2.9}$$

So, the Lemma can be easily deduced from 2.7- 2.9.

2.3 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 , 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 $\forall x \in S_{\mathcal{F}}$, the function $f(\theta, ., x)$ is j times continuously differentiable $y \in \mathcal{C}$ and

(U7)

$$f^{(l)}(\theta, M_\theta(x), x) = 0$$

, if $1 \leq l \leq j$ and $f^{(j)}(\theta, ., x)$ is uniformly continuous on $\mathcal{C}$ such that

$$f^{(j)}(\theta, M_\theta(x), x) > C > 0.$$

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

$$\hat{M}_\theta(x) = \arg\sup_{y \in \mathcal{C}} \hat{f}(\theta, y, x)$$

Theorem 2.3.1. *Under the assumptions of Theorem 2.2.1 and if the conditional density $f(\theta, ., x)$ satisfies (U6) and (U7) , we have*

$$\sup_{x \in S_{\mathcal{F}}} |\hat{M}_\theta(x) - M_\theta(x)| = O(h_K^{\frac{b_1}{j}}) - O(h_H^{\frac{b_2}{j}}) + O_{a.co}\left(\left(\frac{\log d_n^{S_{\mathcal{F}}}}{n^{1-\gamma}\phi(h_K)}\right)^{\frac{1}{2j}}\right)$$

Proof of Theorem 2.3.1

By the Taylor expansion of $f(\theta, y, x)$ in neighborhood of $M_\theta(x)$, we get

$$\hat{f}(\theta, \hat{M}_\theta(x), x) = f(\theta, M_\theta(x), x) + \frac{f^{(j)}(\theta, M_\theta^*(x), x)}{j!}(\hat{M}_\theta(x) - M_\theta(x))^j$$

where $M_\theta^*(x)$ is between $M_\theta(x)$ and $\hat{M}_\theta(x)$.
Combining the last equality with the fact that

$$|\hat{f}(\theta, \hat{M}_\theta(x), x) - f(\theta, M_\theta(x), x)| \leq 2\sup_{y \in \mathcal{C}} |\hat{f}(\theta, y, x) - f(\theta, y, x)|$$

allow to write

$$\sup_{x \in S_{\mathcal{F}}} |\hat{M}_\theta(x) - M_\theta(x)|^j \leq \frac{j!}{f^{(j)}(\theta, M_\theta^*(x), x)} \sup_{x \in S_{\mathcal{F}}} \sup_{y \in \mathcal{C}} |\hat{f}(\theta, y, x) - f(\theta, y, x)|$$

Using the second part of (U7) we obtain that,

$$\exists c > 0, \qquad \sum_{n=1}^{\infty} \mathbb{P}(f^{(j)}(\theta, M_\theta^*, x) < c) < \infty$$

So, we would have

$$|\hat{M}_\theta(x) - M_\theta(x)|^j = O_{a.co}\left(\sup_{y\in\mathcal{C}} |\hat{f}(\theta, y, x) - f(\theta, y, x)|\right)$$

Lemma 2.3.1. *Under the hpotheses of Theorem 2.2.1 thus we have,*

$$\lim_{n\to\infty} \hat{M}_\theta(x) - M_\theta(x) = 0 \quad a.co$$

Proof of Lemma 2.3.1

Because the continuity of the function $f(\theta, y, x)$ we have, for all $\varepsilon > 0$, $\exists \eta(\varepsilon) > 0$ such that

$$|f(\theta, y, x) - f(\theta, M_\theta(x), x)| \le \eta(\varepsilon) \Rightarrow |y - M_\theta(t)| \le \varepsilon$$

Therefore, for $y = \hat{M}_\theta(x)$,

$$\mathbb{P}\left(|\hat{M}_\theta(x) - M_\theta(x)| > \varepsilon\right) \le \mathbb{P}(|f(\theta, \hat{M}_\theta(x), x) - f(\theta, M_\theta(x), x)| > \eta(\varepsilon)\Big)$$

Then according to lemma, $\hat{M}_\theta - M_\theta$ go almost completely to 0, as n goes to infinity.

2.4 Application to prediction

Let us now define the application framework of our results to prediction problem. For each $n \in \mathbb{N}^*$, let $(X_i(t))_{t\in\mathbb{R}}$ $i = 1, ..., n$ 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) 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 $\hat{M}_\theta(x_{new})$, is the best approximation of y_{new} given x_{new}. Applying the result in the above theorem, we obtain the following result.

Corollary 2.4.1. *Under the assumptions of Theorem 2.3.1, we have as n goes to infinity*

$$\lim_{n\to\infty} \hat{M}_\theta(x_{new}) - M_\theta(x_{new}) = 0 \quad a.co$$

Chapter 3

Conditional distribution function with simple functional index

3.1 Model

Let $(X_i, Y_i)_{1 \leq i \leq n}$ be n random variables, identically distributed as the random pair (X, Y) with values in $\mathcal{F} \times \mathbb{R}$.

$$\forall y \in \mathbb{R}. \quad F(\theta, y, x) = \mathbb{P}(Y \leq y| < X, \theta > = < x, \theta >)$$

We introduce a kernel type estimator $\hat{F}(\theta, y, x)$ of $F(\theta, y, x)$ as follows:

$$\hat{F}(\theta, y, x) = \frac{\sum_{i=1}^{n} K(h_k^{-1}(< x - X_i, \theta >))H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_k^{-1}(< x - X_i, \theta >))} \tag{3.1}$$

3.2 Uniform almost complete convergence

In this section we propose to study the uniform almost complete convergence of our estimator defined above (3.1) for this, we need the following assumptions.

3.2.1 Assumptions and results

(A1) and (A3) are the same assumptions (U1) and (U3) as in the second chapter.
(A2) $\forall (y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}, \forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}$ and $\forall \theta \in \Theta_{\mathcal{F}}$

$$|F(\theta, y_1, x_1) - F(\theta, y_2, x_2)| \leq C_{(x,\theta)}(||x_1 - x_2||^{b_1} + |y_1 - y_2|^{b_2})$$

(A4) For $r_n = O\left(\frac{\log n}{n}\right)$ the sequences $d_n^{S_{\mathcal{F}}}$ and $d_n^{\Theta_{\mathcal{F}}}$ satisfy:

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

and

$$\sum_{n=1}^{\infty} n^{1/2b_2} (d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}})^{1-\beta} < \infty for \beta > 1.$$

(A5) H is such that, for all $(y_1, y_2) \in \mathbb{R}^2$,$|H(y_1) - H(y_2)| \leq C|y_1 - y_2|$, $\int |t|^{b_2} H^{(1)}(t)dt < \infty$.

Comments on the assumptions:

Note that Assumptions (A1) and (A2) are respectively the uniform version of $(X \in B_\theta(x, h)) = \phi_{\theta,x}(h) > 0$ and $\forall (y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}$, $\forall (x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$,

$$|F(\theta, y_1, x_1) - F(\theta, y_2, x_2)| \leq C_{\theta,x}(||x_1 - x_2||^{b_1} + |y_1 - y_2|^{b_2}), b_1 > 0, b_2 > 0$$

in pointwise almost complete convergence.
Assumptions (A1) and (A4) are linked with the the topological structure of the functional variable.

Theorem 3.2.1. *Under Assumptions (A1)-(A5), as n goes to infinity, we have*

$$\sup_{\theta \in \Theta} \sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathbb{R}}} |\hat{F}(\theta, y, x) - F(\theta, y, x)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \tag{3.2}$$

Remark 3.2.1. *In the particular case, where the functional single-index is fixed we get the following result. Under Assumptions (A1)-(A5), as n goes to infinity, we have*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|\hat{F}(y,x)-F(y,x)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}}}{n\phi(h_K)}}\right) \tag{3.3}$$

Proof of Theorem3.2.1

Let $\hat{F}_N(\theta,y,x)$ (resp.$\hat{F}_D(\theta,x)$) be defined as:

$$\hat{F}_N(\theta,y,x) = \frac{1}{nE[K_1(\theta,x)]}\sum_{i=1}^{n}K_i(\theta,x)H_i(y) \text{ (resp.} \hat{F}_D(\theta,x) = \frac{1}{nE[K_1(\theta,x)]}\sum_{i=1}^{n}K_i(\theta,x))$$

This proof is based on the following decomposition:

$$\hat{F}(\theta,y,x)-F(\theta,y,x) = \frac{1}{\hat{F}_D(\theta,x)}(\hat{F}_N(\theta,y,x)-E[\hat{F}_N(\theta,y,x)]) - (F(\theta,y,x)-E[\hat{F}_N(\theta,y,x)])+ \frac{F(\theta,y,x)}{\hat{F}_D((\theta,x)}\{1-\hat{F}_D((\theta,x))\}$$

and on the following intermediate results.

Lemma 3.2.1. *Under Assumptions (A1),(A3)and(A4), we have as $n\to\infty$*

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

Corollary 3.2.1. *Under the assumptions of Lemma 3.2.1, we have*

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

Lemma 3.2.2. *Under Assumptions (A1), (A2)and(A4), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|F_N(\theta,y,x)-E(\hat{F}_N(\theta,y,x))| = O(h_K^{b_1}) + O(h_H^{b_2}) \tag{3.4}$$

Lemma 3.2.3. *Under the assumptions of Theorem 3.2.1 as n goes to infinity, we have*

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

3.2.2 Proof

Proof of Lemma3.2.1

To demonstrate this lemma we do the same work as the proof of lemma (2.2.1) in the second chapter ,instead of density function 'f' we put the distribution function 'F'.

Proof of Lemma3.2.2

One has

$$E(\hat{F}_N(\theta,y,x))-F(\theta,y,x)=E(\frac{1}{E(K_1(x,\theta)H_i(y))})-F(\theta,y,x)$$

$$E(\hat{F}_N(\theta,y,x))-F(\theta,y,x)=\frac{1}{E[K_1(x,\theta)]}E[K_1(x,\theta)]E(H_1(y)|<X_1,\theta>)-F(\theta,y,x) \tag{3.5}$$

Moreover, we have:

$$E(H_1(y)|<X_1,\theta>)=\int_{\mathbb{R}}H(h_H^{-1}(y-z))f(\theta,z,X_1)dz$$

now, integrating by parts .

$$I=\int_{\mathbb{R}}H(h_H^{-1}(y-z))f(\theta,z,X_1)dz$$

we pose:

$$U=H(h_H^{-1}(y-z))\Rightarrow dU=-h_H^{-1}H^{(1)}(h_H^{-1}(y-z))$$

$$dV=f(\theta,z,X_1)\Rightarrow V=F(\theta,z,X_1)$$

$$I = [F(\theta, z, X_1)H(h_H^{-1}(y-z))]_{-\infty}^{+\infty} + \frac{1}{h_H}\int_{-\infty}^{+\infty} H^{(1)}(h_H^{-1}(y-z))F(\theta, z, X_1)dz$$

So

$$E(H_1(y)| < X_1, \theta >) = \frac{1}{h_H}\int_{-\infty}^{+\infty} H^{(1)}(h_H^{-1}(y-z))F(\theta, z, X_1)dz$$

we pose:

$$T = h_H^{-1}(y-z) \Rightarrow z = y - h_H T \quad and \quad dz = h_H T$$

we obtain:

$$E(H_1(y)| < X_1, \theta >) = \int_{\mathbb{R}} H^{(1)}(T)F(\theta, y - h_H T, X_1)dT$$

Thus, we have:

$$|E(H_1(y)| < X_1, \theta >) - F(\theta, t, x)| \leq \int_{\mathbb{R}} H^{(1)}(T)|F(\theta, y - h_H T, X_1) - F(\theta, y, x)|dT$$

Finally, the use of (A2) implies that:

$$|\mathrm{E}(\mathrm{H}_1(y)| < X_1, \theta >) - F(\theta, t, x)| \leq C_{\theta,x} \int_{\mathbb{R}} H^{(1)}(T)(h_H^{b_1} + |t|^{b_2} h_H^{b_2})dT \quad (3.6)$$

Because this inequality is uniform on $(\theta, y, x) \in \Theta_{\mathcal{F}} \times S_{\mathcal{F}} \times S_{\mathbb{R}}$ and because of (A5), (3.4) is a direct consequence of (3.5),(3.2.2) and of Corollary (3.2.1).

Proof of Lemma3.2.3

We keep the notation of the Lemma 3.2.1 and we use the compact of $S_{\mathbb{R}}$, we can write that, for some $t_1, ..., t_{z_n} \in S_{\mathbb{R}}$, $S_{\mathbb{R}} \subset \bigcup_{m=1}^{z_n}(y_m - l_n, y_m + l_n)$ with $l_n = n^{-1/2b_2}$ and $z_n \leq Cn^{-1/2b_2}$. Taking $m(y) = \arg \min_{1,\ldots,z_n} |y - t_{m|}$.

Thus, we have the following decomposition:

$$|\hat{F}_N(\theta,y,x)-E(\hat{F}_N(\theta,y,x))| = \underbrace{|\hat{F}_N(\theta,y,x)-\hat{F}_N(\theta,y,x_{k(x)})|}_{\Gamma_1}+\underbrace{|\hat{F}_N(\theta,y,x_{k(x)})-E(\hat{F}_N(\theta,y,x_{k(x)}))|}_{\Gamma_2}$$

$$+\underbrace{2|\hat{F}_N(t_{j(\theta)},y,x_{k(x)})-\hat{F}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})|}_{\Gamma_3}+\underbrace{2|E(\hat{F}_N(t_{j(\theta)},y,x_{k(x)}))-E(\hat{F}_N(t_{j(\theta)},y_{m(y)},x_{k(x)}))|}_{\Gamma_4}$$

$$+\underbrace{|E(\hat{F}_N(\theta,y,x_{k(x)}))-E(\hat{F}_N(\theta,y,x))|}_{\Gamma_5}$$

Concerning Γ_1 we have

$$|\hat{F}_N(\theta,y,x)-\hat{F}_N(\theta,y,x_{k(x)})| \leq \frac{1}{n}\sum_{i=1}^{n}\left|\frac{1}{K_1(\theta,x)}K_i(\theta,x)H_i(y)-\frac{1}{K_1(\theta,x_{k(x)})}K_i(\theta,x_{k(x)})H_i(y)\right|$$

We use the Holder continuity condition on K, the Cauchy-Schwartz inequality, the Bernstein's inequality and the boundness of H (assumption (A5)). This allows us to get:

$$|\hat{F}_N(\theta,y,x)-\hat{F}_N(\theta,y,x_{k(x)})| \leq \frac{C}{\phi(h_K)}\frac{1}{n}\sum_{i=1}^{n}|K_i(\theta,x)H_i(y)-K_i(\theta,x_{k(x)})H_i(y)|$$

$$\leq \frac{C}{\phi(h_K)}\frac{1}{n}\sum_{i=1}^{n}|H_i(y)||K_i(\theta,x)-K_i(\theta,x_{k(x)})|$$

$$\leq \frac{\acute{C}r_n}{\phi(h_K)}$$

Concerning Γ_2, the monotony of the functions $E(\hat{F}_N(\theta,.,x))$ and $\hat{F}_N(\theta,.,x)$ permits to write $\forall m \leq z_n$, $\forall x \in S_{\mathcal{F}}$, $\forall\theta \in \Theta_{\mathcal{F}}$

$$\hat{F}_N(\theta,y_{m(y)}-l_n,x_{k(x)}) \leq \sup_{y\in(y_{m(y)}-l_n,y_{m(y)}+l_n)}\hat{F}_N(\theta,y,x) \leq \hat{F}_N(\theta,y_{m(y)}+l_n,x_{k(x)})$$

$$E(\hat{F}_N(\theta,y_{m(y)}-l_n,x_{k(x)})) \leq \sup_{y\in(y_{m(y)}-l_n,y_{m(y)}+l_n)}E(\hat{F}_N(\theta,y,x)) \leq E(\hat{F}_N(\theta,y_{m(y)}+l_n,x_{k(x)}))$$

Next, we use the Holder's condition on $F(\theta,y,x)$ and we show that, for any $y_1,y_2 \in S_{\mathbb{R}}$ and for all $x \in S_{\mathcal{F}}$, $\theta \in \Theta_{\mathcal{F}}$

$$|E(\hat{F}_N(\theta, y_1, x)) - E(\hat{F}_N(\theta, y_2, x))| = \frac{1}{E(K_1(x,\theta))} |E(K_1(x,\theta))F(\theta, y_1 X_1) - E(K_1(x,\theta))F(\theta, y_1 X_2)|$$

$$\leq C|y_1 - y_2|^{b_2}$$

(3.7)

Now, we have, for all $\eta > 0$

$$\mathbb{P}\left(\left|\hat{F}_N(\theta, y, x_{k(x)}) - E(\hat{F}_N(\theta, y, x_{k(x)}))\right| > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

$$=$$

$$\mathbb{P}\left(\max_{j\in\{1\ldots d_n^{\Theta_{\mathcal{F}}}\}} \max_{k\in\{1\ldots d_n^{S_{\mathcal{F}}}\}} \max_{1\leq m\leq z_n} \left|\hat{F}_N(\theta, y, x_{k(x)}) - E(\hat{F}_N(\theta, y, x_{k(x)})\right| > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

$$\leq$$

$$z_n d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}} \max_{j\in\{1\ldots d_n^{\Theta_{\mathcal{F}}}\}} \max_{k\in\{1\ldots d_n^{S_{\mathcal{F}}}\}} \max_{1\leq m\leq z_n} \mathbb{P}\left(\left|\hat{F}_N(\theta, y, x_{k(x)}) - E(\hat{F}_N(\theta, y, x_{k(x)}))\right|\right.$$

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

$$\leq 2 z_n d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}} \exp\left(-C\eta^2 \log d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}}\right)$$

Choising $z_n = O(l_n^{-1}) = O(n^{\frac{1}{2b_2}})$, we get:

$$\mathbb{P}\left(\left|\hat{F}_N(\theta, y, x_{k(x)}) - E(\hat{F}_N(\theta, y, x_{k(x)}))\right| > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) \leq \acute{C} z_n (d_n^{S_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}})^{1-C\eta^2}$$

putting $C\eta^2 = \beta$ and using (A4), we get:

$$\Gamma_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

Concerning Γ_3 and Γ_4, using Lipschitz's condition on the kernel H, one can write

$$\left|\hat{F}_N(t_{j(\theta)}, y, x_{k(x)}) - \hat{F}_N(t_{j(\theta)}, y_{m(y)}, x_{k(x)})\right| \leq C\frac{1}{n\phi(h_K)}\sum_{i=1}^{n} K_i(t_{j(\theta)}, x_{k(x)})|H_i(y) - H_i(y_{m(y)})|$$

$$\leq \frac{Cl_n}{nh_H\phi(h_K)}\sum_{i=1}^{n} K_i(t_{j(\theta)}, x_{k(x)})$$

Once again a standard exponential inequality for a sum of bounded variables allows us to write

$\hat{F}_N(t_{j(\theta)}, y, x_{k(x)}) - \hat{F}_N(t_{j(\theta)}, y_{m(y)}, x_{k(x)}) = O(\frac{l_n}{h_H}) + O_{a.co}\left(\frac{l_n}{h_H}\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)$

Now, the fact that $\lim_{n\to\infty} n^{\gamma}h_H = \infty$ and $l_n = n^{-\frac{1}{2b_2}}$ imply that:

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

then:

$$\Gamma_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

Hence, for n large enough, we have

$$\Gamma_3 \leq \Gamma_4 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right).$$

Concerning Γ_5, we have

$E(\hat{F}_N(\theta, y, x_{k(x)})) - E(\hat{F}_N(\theta, y, x)) \leq \sup_{x\in S_{\mathcal{F}}}\left|\hat{F}_N(\theta, y, x) - \hat{F}_N(\theta, y, x_{k(x)})\right|,$

then following similar proof used in the study of Γ_1 and using the same idea as for $E(\hat{F}_N(\theta, y, x_{k(x)})) - E(\hat{F}_N(\theta, y, x))$ we get, for n tending to infinity

$$\Gamma_5 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\phi(h_K)}}\right)$$

3.3 The conditional quantile in functional single-index model

In this part we investigate the asymptotic properties of the conditional quantile function of a scalar response and functional covariate when the observations are from a single functional index model and data are independent and identically distributed (i.i.d.). We will consider the problem of the estimation of the conditional quantiles. Saying that, we are implicitely assuming the existence of a regular version for the conditional distribution of Y given $< X, \theta >=< x, \theta >$. From the conditional distribution function $F(\theta, ., x)$, , it is easy to give the general definition of the α -order quantile:

$$t_\theta(\alpha) = \inf\{t \in \mathbb{R}; F(\theta, t, x) \geq \alpha\}, \forall \alpha \in (0, 1)$$

In order to simplify our framework and to focus on the main interest of our part , we assume that $F(\theta, ., x)$ is strictly increasing and continuous in a neighborhood of $t_\theta(\alpha)$. This is insuring unicity of the conditional quantile $t_\theta(\alpha)$ which is defined by:

$$t_\theta(\alpha) = F^{-1}(\theta, \alpha, x), . \tag{3.8}$$

$\hat{t_\theta}(\alpha)$ is an estimator of $t_\theta(\alpha)$,defined as:

$$\hat{t_\theta}(\alpha) = \hat{F}^{-1}(\theta, \alpha, x) \tag{3.9}$$

Or as:

$$\hat{F}(\theta, \hat{t_\theta}(\alpha), x) = \alpha$$

To insure existence and unicity of this quantile, we will assume that

(A6) $F(\theta, ., x)$ is some point strictly increasing, In order to insure unicity of $\hat{t}_\theta(\alpha)$ we will make the following, quite unrestrictive, assumption:

(A7) H is strictly increasing,

More precisely, we will suppose that there exists some integer $j > 0$ such that:

(A8) $F^{(l)}(\theta, t_\theta(\alpha), x) = 0$, if $1 \leq l < j$, and $F^{(j)}(\theta, ., x)$ is uniformly continuous on $S_{\mathbb{R}}$,such that, $|F^{(j)}(\theta, t_\theta(\alpha), x)| > C > 0$

Theorem 3.3.1.

$$\sup_{x\in S_{\mathcal{F}}} |\hat{t}_\theta(\alpha) - t_\theta(\alpha)| = O(h_K^{\frac{b_1}{j}} + h_H^{\frac{b_2}{j}}) + O_{a.co}\left(\left(\frac{\log d_n^{S_{\mathcal{F}}}}{n\phi_x(h_K)}\right)^{1/2j}\right)$$

proof of Theorem 3.3.1

Let us write the following Taylor expansion of the function $\hat{F}(\theta, ., x)$:

$$\hat{F}(\theta, t_\theta(\alpha), x) - \hat{F}(\theta, \hat{t_\theta}(\alpha), x) = \sum_{l=1}^{j-1} \frac{(t_\theta(\alpha) - \hat{t_\theta}(\alpha))^l}{l!} \hat{F}^{(l)}(\theta, t_\theta(\alpha), x) + \frac{(t_\theta(\alpha) - \hat{t_\theta})^j}{j!} \hat{F}^{(j)}(\theta, t^*, x)$$

where t^* is some point between $t_\theta(\alpha)$ and $\hat{t_\theta}(\alpha)$. It suffices now to use the first part of condition (A8) to be able to rewrite this expression as:

$$\hat{F}(\theta, t_\theta(\alpha), x) - \hat{F}(\theta, \hat{t_\theta}(\alpha), x) = \sum_{l=1}^{j-1} \frac{(t_\theta(\alpha) - \hat{t_\theta}(\alpha))^l}{l!} \left(\hat{f}^{(l-1)}(\theta, t_\theta(\alpha), x) - f^{(l-1)}(\theta, t_\theta(\alpha), x)\right)$$

$$+ \frac{(t_\theta(\alpha) - \hat{t_\theta})^j}{j!} \hat{f}^{(j-1)}(\theta, t^*, x)$$

As long as we could be able to check that

$$\exists \tau > 0, \sum_{n=1}^{n=\infty} (f^{(j-1)}(\theta, t^*, x) < \tau) < \infty,$$

we would have

$$(t_\theta(\alpha)-\hat{t_\theta}(\alpha))^j = O\left(\hat{F}(\theta,t_\theta(\alpha),x)-F(\theta,t_\theta(\alpha),x)\right)+O\left(\sum_{l=1}^{j-1}(t_\theta(\alpha)-\hat{t_\theta}(\alpha))\right)^l\left(\hat{f}^{(l-1)}(\theta,t_\theta(\alpha),x)\right.$$

$$\left.-f^{(l-1)}(\theta,t_\theta(\alpha),x)\right)\ a.co$$

So we have

$$\left(t_\theta(\alpha)-\widehat{t}_\theta(\alpha)\right)^j = O_{a.co}\left(\hat{F}(\theta,t_\theta(\alpha),x)-F(\theta,t_\theta(\alpha),x)\right)$$

Lemma 3.3.1. *If the conditions of Theorem 3.2.1 hold together with (A6) and (A7), then we have:*

$$\lim_{n\to\infty}\hat{t_\theta}(\alpha)-t_\theta(\alpha)=0 \quad a.co$$

Chapter 4

On the estimation of the functional single index

4.1 The method of M-estimators

M-estimators were introduced by Huber (1964,[29]), where he proposed to generalize the maximum likelihood estimator by considering it as a minimization problem of a certain function. We consider an i.i.d.of n real random variables $Y_1, ..., Y_n$ from the same variable Y Huber stands in a general framework of estimators $\hat{\theta}$ defined by

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}} \sum_{i=1}^{n} \rho(Y_i, \theta)$$

where the function $\rho : \mathbb{R}^2 \to \mathbb{R}$ is assumed to be measurable. Different choices of the function ρ lead to different estimators of functionals of the distribution of Y (mode, median and quantile). Intuitively, the family of M-estimators can be seen as a generalization of one of the definitions of the mean. Indeed, the expectation $E(Y)$ of a random variable Y can be defined as the solution of the following minimization problem:

$$E(Y) = \arg\min_{t \in \mathbb{R}} E[(Y - t)^2]$$

This naturally allows us to propose the following estimator of the mean

$$\hat{E}(Y) = \arg\min_{t\in\mathbb{R}} \sum_{i=1}^{n}(Y_i - t)^2$$

Huber in his study distinguished two classes of M-estimator depending on the convexity or non-convexity of the function ρ. In both cases this function must check certain terms. For the case of regression estimation, in a simple index model, Delecroix and Hristache (1999,[12]) propose to estimate θ by the method of M-estimators. The idea of this method is therefore to define an estimator $\hat{\theta}$ of θ, solution of the problem of next maximization

$$\theta = \arg\max_{\theta\in\Theta} E[\psi(Y, E[Y|X\theta])|X = x]$$

The nonparametric estimator $\hat{\theta} = \hat{\theta_\psi}$ is, by analogy, a solution of the problem of maximization

$$\hat{\theta}_\psi = \arg\max_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n}\psi(Y_i, \hat{r}(X_i, \theta))$$

where ψ is a function defined on $\mathbb{R}^2$ and with values in $\mathbb{R}$ satisfying a certain number of conditions and $r(X\theta) = r_\theta(X)$ is the assumed unknown regression function. Several choices of the function ψ can be considered, for more detail on these choice the reader can consult Serfling (1980,[46]).

Delecroix and Hristache (1999[12]) show that $\hat{\theta_\psi}$ is almost surely consistent and asymptotically normal in the case where ψ is the log-likelihood of a density belonging to an exponential family. For the estimation of $\hat{\theta_\psi}$ by the M-estimate method, we can cite Sherman (1994,[35]), Xia and Li (1999,[49]) and Xia et al. (1999,[50]). It exists other estimation methods based on M-estimators in the literature. HAS as an example the pseudo-maximum likelihood method which essentially based on conditional density estimation.

4.1.1 The pseudo-maximum likelihood method

As indicated in the previous paragraph, the pseudo-maximum likelihood method is a special case of robust estimation in simple index model. Indeed, Delecroix et al. (2003,[13]) propose to estimate the density conditional considering the following model

$$f_{Y|X}(x, y) = f_\theta(x\theta, y)$$

where $f_\theta(t, y)$ represents the conditional density of Y knowing $x\theta = t$ evaluated at the point y. The idea is to take the likelihood function

$$\prod_{i=1}^{n} f_\theta(X_i\theta, Y_i) f_X(X_i)$$

and the log-likelihood

$$\sum_{i=1}^{n} \log f_\theta(X_i\theta, Y_i) + \sum_{i=1}^{n} \log f_X(X_i)$$

Since the term $\sum_{i=1}^{n} f_X(X_i)$ does not depend on θ, the estimator of the maximum of likelihood could be defined, if f_θ were known, by maximizing the first term $\sum_{i=1}^{n} \log f_\theta(X_i\theta, Y_i)$. As f_θ is unknown, they define the following M-estimator:

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \frac{1}{n} \sum_{i=1}^{n} \log(\hat{f}_\theta(X_i\theta, Y_i)) \tag{4.1}$$

where $\hat{f}_\theta$ represents the kernel estimator of f_θ.

General Conclusion and perspectives

We were interested specifically in this work to single-index conditional models that treat the case of functional variables in which "response" variable is true while the explanatory variable is functional in the i.i.d case. The objective was the semi-parametrical estimation of the conditional density function as well as the conditional distribution function by the kernel method, and their application the conditional mode(resp.the conditional quantile).

The richness of this functional statistical research area offers many perspectives both theoretically and practically, let us cite:

- The asymptotic normality of our estimators can allow us to test and build confidence intervals.
- We can also consider an asymptotic study for our esitmators in the ergodic case.
- Another possible perspective is to assume that not only the explanatory variable is functional but also the variable of interest.

Bibliography

[1] Ait-Saïdi, A., Ferraty, F., Kassa, R., and Vieu, P. (2008), 'Cross-Validated Estimations in the SingleFunctional Index Model', Statistics, 42, 475494.

[2] Aneiros, G., Cao, R., Fraiman, R., Genest, C., and Vieu, V. (2019), 'Recent Advances in Functional Data Analysis and High-Dimensional Statistics', Journal of Multivariate Analysis, 170, 39.

[3] Aneiros-Prez, G., and Vieu, P. (2011), 'Automatic Estimation Procedure in Partial Linear Model with Functional Data', Statistical Papers, 52(4), 751771.

[4] Aneiros-Pérez, G., and Vieu, P. (2015), 'Partial Linear Modelling with Multi-Functional Covariates', Computational Statistics, 30(3), 647671.

[5] Asser, T., Hall, p. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. J. Roy. Statist. Soc. Ser. B, 60, 681691.

[6] 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

[7] Attaoui, S. (2014), 'On the Nonparametric Conditional Density and Mode Estimates in the Single Functional Index Model with Strongly Mixing Data', Sankhya A, 76(2), 356378.

[8] Bali, L., Boente, G., Tyler, D., and Wang, J.L. (2011), 'Robust Functional Principal Components: A Projection-Pursuit Approach', Annals of Statistics, 39, 28522882.

[9] Berlinet, A., Gannoun, A. and Matzner-Lober, E. (1998). Normalité asymptotique d'estimateurs convergents du mode conditionnel. Canad. J of Statist., 26, 365380.

[10] Chen, D., Hall, P., and Müller, H.G. (2011), 'Single and Multiple Index Functional Regression Models with Nonparametric Link', The Annals of Statistics, 39, 34583486.

[11] Cuevas, A. (2014), 'A Partial Overview of the Theory of Statistics with Functional Data', Journal of Statistical Planning and Inference, 147, 123.

[12] Delecroix, M, Härdle, W. Hristache, M., (1999). M-estimateurs semiparamétriques dans les modèles à direction révélatrice unique. Bull. Belg. Math. Soc. Simon Stevin, 6, No 2, 161-185.

[13] Delecroix, M, Härdle, W. Hristache, M., (2003). Ecient estimation in conditional single-index regression. J. Multivariate Anal. 86,213-226.

[14] Ezzahrioui, M., Ould-Said, E. (2005). Asymptotic normality of nonparametric estimators of the conditional mode for functional data. Technical report, No.249, LMPA, Univ. Littoral Cote d'Opale.

[15] Ezzahrioui, M., Ould-Said, E. (2006). On the asymptotic properties of a nonparametric estimator of the conditional mode for functional dependent data. Preprint, LMPA No 277, Univ. du Littoral Cote d'Opale.

[16] Ferraty, F., Goia, A., Salinelli, E., and Vieu, P. (2013), 'Functional Projection Pursuit Regression', Test, 22(2), 293320.

[17] Ferraty, F., Peuch, A., and Vieu, P. (2003), 'Modèle à Indice Fonctionnel Simple', Comptes Rendus Mathematique, 336, 10251028.

[18] Ferraty, F.Laksaci, A., Vieu,P.(2006) Estimation of some characteristics of the conditional distribution in non parametric functional models. Statistical.Inference for Stochastic processes.9, 47-76.

[19] Ferraty, F., Rabhi, A. and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nino phenomenon. Sankhya. 67 , 378-398.

[20] Ferraty, F.Laksaci, A., Vieu,P.(2006) Estimation of some characteristics of the conditional distribution in non parametric functional models. Statistical.Inference for Stochastic processes.9, 47-76

[21] Ferraty, F., Rabhi, A. and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nino phenomenon. Sankhya. 67 , 378-398.

[22] Ferraty,F.,Vieu,P.(2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-norm´es. C.R.Acad.Sci.,Paris.330, No 2 ,139-142.

[23] Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist. 16, 111-125.

[24] Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist. 16, 111-12.

[25] Ferraty, F., and Vieu, P. (2006), Nonparametric Functional Data Analysis: Theory and Practice, NewYork: Springer.

[26] Goia, A., and Vieu, P. (2014), 'Some Advances on Semi-Parametric Functional Data Modelling', in Contributions in Infinite-Dimensional Statistics and Related Topics, Bologna: Esculapio, pp. 135140.

[27] Goia, A., and Vieu, P. (2015), 'A Partitioned Single Functional Index Model', Computational Statistics, 30(3), 673692.

[28] Hsing, T., and Eubank, R. (2015), Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators, Chichester: Wiley.

[29] Huber, P. J., (1964). Robust estimation of a location parameter. Ann. Math. Statist. 35, 73-101.

[30] Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur a noyau de la densité conditionnelle avec variable explicative fonctionnelle. Ann. I.S.U.P. 51, 69-80.

[31] Laksaci, A., Madani, F., Rachdi, M. (2010). Kernel conditional density estimation when the regressor is valued in a semi metric space. International Statistical Review. (In press).

[32] Li, Y., Wang, N., and Carroll, R.J. (2010), Generalized Functional Linear Models with Semiparametric Single-Index Interactions, Journal of the American Statistical Association, 105(490), 621633.

[33] Ling, N., and Vieu, P. (2018), 'Nonparametric Modelling for Functional Data: Selected Survey and Tracks for Futur', Statistics, 52, 934949.

[34] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. Stoch. Proc. and their Appl., 115, 155-177.

[35] Maity, A., and Huang, J.Z. (2012), 'Partially Linear Varying Coefficient Models Stratified by a Functional Covariate', Statistics Probability Letters, 82(10), 18071814.

[36] Quintela-Del-Rio, A. and Vieu, PH (1997). A nonparametric conditional mode estimate. J. Nonparametr. Stat., 8, 253268.

[37] Roussas, G.(1969) Nonparametric estimation of the transition distribution function Markov process. Ann. Math.Staist.40 1386-140

[38] Ramsay, J., and Silverman, B. (2002), Applied Functional Data Analysis: Methods and Case Studies, New York: Springer.

[39] Ramsay, J., and Silverman, B. (2005), Functional Aata Analysis (2nd ed.), New York: Springer

[40] Roussas, G.(1969) Nonparametric estimation of the transition distribution function Markov process. Ann. Math.Staist.40 1386-140.

[41] Roussas, G.G. (1968). On some properties of nonparametric estimates of probability density functions. Bull. Soc. Math. Greece (N.S.) 9, 2943.

[42] Rosenblatt, M. (1969). Conditional probability density and regression estimates. In Multivariate Analysis II. (P. R. Krishnaiah, ed.). Academic Press, New York N.Y., pp. 25-31.

[43] Ramsay, J., and Silverman, B. (2005), Functional Aata Analysis (2nd ed.), New York: Springer.

[44] Roussas, G.(1969) Nonparametric estimation of the transition distribution function Markov process. Ann. Math.Staist.40 1386-1400.

[45] Samanta, M. and Thavaneswaran, A. (1990). Nonparametric estimation of the conditional mode. Comm. Statist. Theory Methods, 19, 45154524.

[46] Serfing, R. (1980). Approximation Theorems of Mathematical Statistics. Wiley.

[47] Stute, W.(1986) On almost sure convergence of conditional empirical distribution functions.Ann.Probab. 14 891-901.

[48] Vieu, P. (2018), 'On Dimension Reduction Models for Functional Data', Statistics Probability Letters, 136, 134138.

[49] Xia, X, Tong, H Li, W. K. (1999). On extended partially linear single-index models. Biometrika, 86, 831-842.

[50] Xia, X Li, W. K. (1999). On the estimation and testing of functional-coecient linear models. Stat. Sinica, 9, 735-757.

[51] Youndjé,E. (1993) Estimation nonparamétrique de la densité conditionnelle par la méthode du noyau (in french). PhD thesis, University of Rouen.

.