School of Mathematical
and Physical Sciences

Department of Mathematics and Statistics

Preprint MPS-2012-22

18 October 2012

Nonparametric Multivariate L_1-median Regression Estimation with Functional Covariates

by

Mohamed Chaouch and Naâmane Laïb

Nonparametric Multivariate L_1-median Regression Estimation with Functional Covariates

Mohamed Chaouch[1,] and Naâmane Laïb[2]
[1] Centre for the Mathematics of Human Behaviour (CMoHB)

quite interesting in practice. Furthermore, it is widely acknowledged that quantiles are more robust to outliers than regression function.

Conditional quantiles are widely studied when the explanatory variable X lies within a *finite* dimensional space. There are many references on this topic (see Gannoun *et al.* (2003a)).

During the last decade, thanks to progress of computing tools, there is an increasing number of examples coming from different fields of applied sciences for which the *data are curves*. For instance, some random variables can be observed at several different times. This kind of variables, known as *functional variables* (of time for instance) in the literature, allows us to consider the data as curves. The books by Bosq (2000) and Ramsay and Silverman (2005)) propose an interesting description of the available procedures dealing with functional observations whereas Ferraty and Vieu (2006) present a completely non-parametric point of view. These functional approaches mainly rely on generalizing multivariate statistical procedures in functional spaces and have been proved to be useful in various areas such as chemiomertrics (Hastie and Mallows (1993) and Quintela-del Río and Francisco-Fernández (2011)), economy (Kneip and Utikal (2001)), climatology (Besse *et al.* (2000)), biology (Kirkpatrick and Heckman (1989)), Geoscience (Quintela-del Río and Francisco-Fernández (2011)) or hydrology (Chebana and Ouarda (2011)). These functional approaches are generally more appropriate than longitudinal data models or time series analysis when there are, for each curve, many measurement points (Rice (2004)).

In the *univariate* case (i.e. Y R and X

extension of multivariate quantiles based on norm minimization and on the geometry of multivariate data clouds.

In contrast, relative little attention has been paid to the multivariate conditional quantiles (Y R^d and X R^s) and their large sample properties. Cadre (2001) defined the conditional L_1-median and provided its uniform consistency on a compact subsets of R^s. Recently, De Gooijer *et al.* (2006) have introduced a multivariate conditional quantile notion, which extends the definition of unconditional quantiles by Abdous and Theodorescu (1992), to predict tails from bivariate time series. Cheng and De Gooijer (2007) have generalized the notion of geometric quantiles, defined by Chaudhuri (1996), to the conditional setting. They have established a Bahadur-type linear representation of the u-th geometric conditional estimator as well as the asymptotic normality in the i.i.d. case.

The purpose of this paper is to add some new results to the non-parametric estimation of the conditional L_1-median when Y is a random vector with values in R^d while the covariable X take its values in some *infinite* dimensional space F. As far as we know, this problem has not been studied in literature before and the results obtained here are believed to be novel. Moreover, our motivation for studying this type of robust estimator is due to its interest in some practical applications. Note also that, it would be better to predict all components of a vector of random variables simultaneously in order to take into account the correlation between them rather than predicting each of component separately. For instance, in EDF (French electricity company) the estimation of the minimum and the maximum of the electricity power demand represents an important research issue for both economic and security reasons. Because an underestimation of the maximum consumed quantity of electricity (especially in winter) may require importation of el el istoeqea-1(an)-1(eou)-1(sl)-1(y)-280350.481

of the results in Section 3 are relegated to the Appendix.

2 Notations and definitions

Let us consider a random pair (X, Y) where X and Y are two random variables defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We suppose that Y is $\mathbb{R}^d$-valued and X is a *functional random variable (f.r.v.)* takes its values in some infinite dimensional vector space $(\mathcal{F}, d(\cdot,\cdot))$ equipped with a semi-metric $d(\cdot,\cdot)$. Let x be a fixed point in $\mathcal{F}$ and $F(.|x)$ be the conditional cumulative distribution function (cond. c.d.f) of Y given $X = x$. The conditional L_1-median, $\mu : \mathcal{F} \longrightarrow \mathbb{R}^d$, of Y given $X = x$, is defined as the miminizer over u of

$$\arg\min_{u \in \mathbb{R}^d} \mathbb{E}[(\|Y - u\| - \|Y\|) \mid X = x] = \arg\min_{u \in \mathbb{R}^d} \int (\|y - u\| - \|y\|)\, dF(y \mid x). \qquad (1)$$

The general definition (1) does not assume the existence of the firendFdFdFdFd259(()26.01Tf()T()TjETBT10.9091

Notice that $H^x(u)$ is bounded whenever $\mathrm{E}\left[\left| Y - u \right|^{-1} \mid X = x \right] <$

3 Main Results

3.1 Further notations and hypotheses

Let x be a given point in F and V_x a neighbourhood of x. Denote by $B(x, h)$ the ball of center x and radius h, namely $B(x, h) = \{x \quad F : d(x, x\ ) \leq h\}$. For $(\ , u) \quad \mathbb{R} \times \mathbb{R}^d$, denote by $G^x\ (u) = \mathbb{E}\ \ Y - u \quad | X = x\ $, for $x \quad F$. Our hypotheses are gathered here for easy reference.

(H1) K is a nonnegative bounded kernel of class C^1 over its support $[0, 1]$ such that $K(1) > 0$. The derivative K exists on $[0, 1]$ and satisfy the condition $K\ (t) < 0$, for all $t \quad [0, 1]$ and $|\int_0^1 (K^j)\ (t)dt| < \infty$ for $j = 1, 2$.

(H2) For $x \quad F$, there exists a deterministic nonnegative bounded function g and a nonnegative real function φ tending to zero, as its argument tends to 0, such that

(*i*) $F_x(h) := \mathbb{P}(X \quad B(x, h)) = \varphi(r) \cdot g(x) + o(\varphi(h))$ as $h \to 0$.

(*ii*) There exists a nondecreasing bounded function τ_0 such that, uniformly in $s \quad [0, 1]$, $\varphi(hs$

Comments on the Hypotheses

Theorem 3.2 *Assume (H1)-(H2), (H3)(i) and (H4)(i) and condition (10) hold true. Then, we have*

$$\lim_{n\to\infty} \mu_n(x) = \mu(x) \quad a.s. \tag{12}$$

3.3 Asymptotic normality

with $\delta_{k,j} = 1$ if $k = j$ and zero otherwise and $M_{k,j}(Y_i, u) = [\delta_{k,j} - \frac{(Y_i^j - u^j)(Y_i^k - u^k)}{Y_i - u \;\; ^2}]/u$

(ii)

and the matrix $H^x(\mu)$ by

$$H_n^x(\mu_n) = \sum_{i=1}^{n} w_{n,i}(x)\, M\,(Y_i, \mu_n).$$

Making use of the decomposition of $F_x(u)$ in (H

4.1 Simulation example

Let us consider a bi-dimensional vector $\mathbf{Y} = (Y_1, Y_2) \in \mathbb{R}^2$ and $X(t)$ is a Brownian motion trajectories defined on $[0, 1]$. The eigenfunctions of the covariance operator of X are known to be (see Ash and Gardner (1975)), for $j = 1, 2, \ldots$

$$f_j(t) = \sqrt{2}\sin\{(j - 0.5)\pi t\}, \ t \in [0, 1].$$

Let $(f_1(t))_{t \in [0,1]}$ (resp. $(f_2(t))_{t \in [0,1]}$) be the first (resp. the second) eigenfunction corresponding to the first (resp. second) greater eigenvalue of the covariance operator of X. It is well known that $f_1(t)$ and $f_2(t)$ are orthogonal by construction, i.e. $< f_1, f_2 >:= \int_0^1 f_1(t)f_2(t) = 0.$
We modelize then the dependence between $\mathbf{Y}$ and X by the following model:

- $Y^1 = \int_0^1 f_1(t)X(t)\, dt + \epsilon$
- $Y^2 = \int_0^1 f_2(t)X(t)\, dt + \epsilon$

where ϵ is a standard normal random variable.

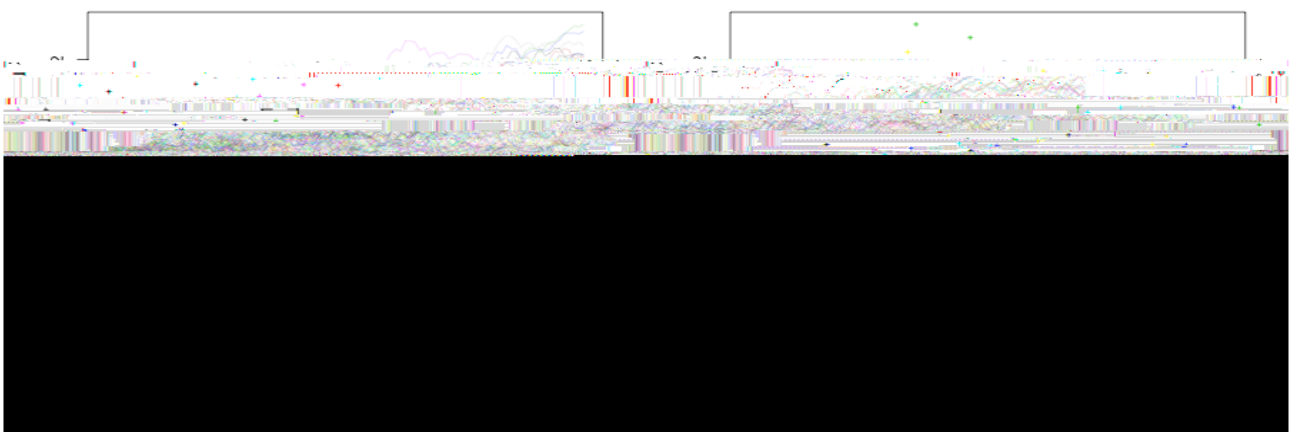

Figure 1: Sample of 200 simulated couples of observations $(X_i, \mathbf{Y}_i)_{i=1,\ldots,200}$. The left box contains the covariates X_i and in the right one we present their associated vectors $\mathbf{Y}_i$.

We have simulated $n = 200, 700$ independent realizations $(X_i, \mathbf{Y}_i)$, $i = 1, \ldots, n.$ To deal with the Brownian random functions $X_i(t)$, their sample were discretized by 100 points equispaced in $[0, 1]$. In Figure 1, we plot a 200 simulated couples $(X_i, \mathbf{Y}_i)_{i=1,\ldots,200}$ as described above. The left box contains the covariates X_i and in the right one we present the associated vectors $\mathbf{Y}_i = (Y_i^1, Y_i^2)$.

We aim to assess, for a fixed curve $X = x$, the performance of the asymptotic conditional confidence ellipsoid given by (18) in finite sample. For that we have first to estimate $\mu(x)$. Three

$\hat{F}^j(\cdot \mid X_i)$ is the conditional distribution function estimator of the component Y^j given $X = X_i$. Ferraty and Vieu (2006), p. 56, have proposed a Nadaraya-Watson kernel estimator of the conditional distribution, $F^j(\cdot \mid X = X_i)$, when covariate takes values in some infinite dimensional space. This estimator is given by

$$\hat{F}^j(y^j \mid X = X_i) = \sum_{k=1}^{160} \mathbb{1}_{\{Y_k^j \le y^j\}} K(d(X_i, X_k)/h_n) \Big/ \sum_{k=1}^{160} K(d(X_i, X_k)/h_n), \quad y^j \in \mathbb{R}.$$

To apply this approach, we used the Ferraty and Vieu's R/routine *funopare.quantile.lcv*[1] to estimate $\hat{\mu}^j(X_i)$. The optimal bandwidth is chosen by the cross-validation method on the k nearest neighbours (see Ferraty and Vieu (2006), p.102 for more details).

(iii) Conditional Multivariate Median (CMM)

The approach that we propose here supposes the covariate X is a curve and the response Y is a vector. For each $i = 1, \ldots, 160$ in the learning sample we take

$$\hat{Y}_i = \hat{\mu}(X_i),$$

where

$$\hat{\mu}(X_i) = \arg\min_{u \in \mathbb{R}^3} \sum_{j=1}^{160} w_{n,j}(X_i) \; Y$$

	CMM				VCCM				NF			
	Mean	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
Moist.	1.301	0.479	1.100	2.202	1.776	0.460	1.879	2.383	7.222	1.663	6.374	11.44
Fat	1.565	0.430	1.500	2.401	2.343	0.925	1.716	2.867	9.758	2.328	8.4	15.24
Prot.	1.125	0.300	0.800	1.437	1.313	0.518	1.182	1.806	2.446	0.787	2.329	3.394
$R(\hat{Y})$	2.638	1.349	2.530	3.623	3.561	1.877	2.909	3.799	12.6	3.523	10.6	19.27

Table 1: Distribution of absolute errors for Moisture, Fat and Protein and global estimation error of the vector **Y**.

The used bandwidth for each curve X_i in the test sample is the one obtained for the nearest curve in the learning sample. Because the spectrometric curves presented in Figure (3) are very smooth, we can choose as semi-metric $d(\cdot,\cdot)$ the L_2 distance between the second derivative of the curves. This choice has been made by Attouch *et al.* (2009) and Ferraty *et al.* (2007) for the same spectrometric curves.

Both (CMM) and (NF) methods take into account the covariance structure between variables of of the vector **Y**. In fact, the correlation coefficients between $Y_1 = moisture$, $Y_2 = fat$ and $Y_3 = protein$ are given by $\rho_{1,2} = -0.988$, $\rho_{1,3} = 0.814$ and $\rho_{2,3} = -0.860$. As we can see moisture,

5 Concluding remarks

In this paper, we have introduced a kernel-based estimator for the L_1-median of a multivariate con-

Since $G^{x}_{n,1}$ is independent of u, it follows from decomposition (21

u

Lemma 5.3 *(i) Under conditions (H1)-(H2) (H3)(i), we have*

$$\sup_{u\in\mathbb{R}^d} |B_n^x(u)| = O_{a.s.}(h^\beta). \tag{23}$$

(ii) If in addition that (H1)-(H2) hold true and condition (10) is satisfied, we have

$$\sup_{u\in\mathbb{R}^d} |R_n^x(u)| = O_{a.s.}\left(h^\beta\sqrt{\frac{\log n}{n\varphi(h)}}\right) \tag{24}$$

Proof of Lemma 5.3. Recall that

$$B_n^x(u) = \overline{G}_{n,2}^x(u) - G^x(u).$$

Conditioning by X and using the definition of $G^x(u)$ and condition (H3)(i), one has

$$\begin{aligned}
|B_n^x(u)| &= \left|\frac{1}{E\Delta_1(x)}E\left\{\Delta_1(x)E[\mathbb{1}_{Y_1 - u} \mid X]\right\} - G^x(u)\right| \\
&= \left|\frac{1}{E\Delta_1(x)}E\left[\Delta_1(x)(G^X(u) - G^x(u))\right]\right| \\
&\leq \sup_{x'\in B(x,h)} |G^{x'}(u) - G^x(u)| = O_{a.s.}(h^\beta).
\end{aligned}$$

The later quantity is independent of u, this leads to $\sup_{u\in\mathbb{R}^d} |B_n^x(u)| = O_{a.s.}(h^\beta)$.

We have

$$
\begin{aligned}
& \sup_{\|u\|\leq n^{\gamma}} |G^{x}_{n,2}(u) - \overline{G}^{x}_{n,2}(u)| \\
\leq\ & \max_{1\leq j\leq k_n^d} \sup_{u \in S_{n,j}} |G^{x}_{n,2}(u) - G^{x}_{n,2}(u_j)| + \max_{1\leq j\leq k_n^d} |G^{x}_{n,2}(u_j) - \overline{G}^{x}_{n,2}(u_j)| \\
+\ & \max_{1\leq j\leq k_n^d} \sup_{u \in S_{n,j}} |\overline{G}^{x}_{n,2}(u) - \overline{G}^{x}_{n,2}(u_j)| := I_{n,1} + I_{n,2} + I_{n,3}.
\end{aligned}
\tag{25}
$$

Observe now that

$$
\begin{aligned}
\sup_{u \in S_{n,j}} |G^{x}_{n,2}(u) - G^{x}_{n,2}(u_j)| &\leq \frac{1}{n\mathrm{E}(\Delta_1(x))} \sum_{i=1}^{n} \sup_{u \in S_{n,j}} \Big| \|Y_i - u\| - \|Y_i - u_j\| \Big| \Delta_i(x) \\
&\leq \frac{1}{n\mathrm{E}(\Delta_1(x))} \sum_{i=1}^{n} \Delta_i(x) \sup_{u \in S_{n,j}} \|u - u_j\| = b_n G^{x}_{n,1},
\end{aligned}
$$

and

$$
\sup_{u \in S_{n,j}} |\overline{G}^{x}_{n,2}(u) - \overline{G}^{x}_{n,2}(u_j)| \leq \mathrm{E}\left(\sup_{u \in S_{n,j}} |G^{x}_{n,2}(u) - G^{x}_{n,2}(u_j)| \right) = b_n.
$$

If we denote by $\alpha_n = \sqrt{n\varphi(h)/\log G}$

In order to apply an exponential type inequality, we have to give an upper bound for $\mathrm{E}\left(|Z_{n,1}(x)|^m\right)$. It follows from the above inequality that

$$\mathrm{E}\left(|Z_{n,1}(x)|^m\right) \leq C \sum_{k=0}^{m} \binom{m}{k} \mathrm{E}\left(\left|\left(Y_1 - u_j\right) \Delta_1(x)\right|^k\right) \left[\mathrm{E}\left(\left|\left(Y_1 - u_j\right) \Delta_1(x)\right|\right)\right]^{m-k}.$$

On the other hand, we have for any $k \geq 2$

$$\begin{aligned} \mathrm{E}\left(\left|\left(Y_1 - u_j\right) \Delta_1(x)\right|^k\right) &= \mathrm{E}\left((\Delta_1(x))^k \mathrm{E}\left(\left|Y_1 - u_j\right|^k \mid X_1\right)\right) \\ &= \mathrm{E}\left((\Delta_1(x))^k G_k^{X_1}(u_j)\right). \end{aligned}$$

that

$$P(|I_{n,2}| \geq \epsilon) \quad \leq \quad 2k_n^d \exp\left(-\epsilon_0^2 \log n \left(\frac{1}{2(1+\epsilon_0\sqrt{v_n})} \right)\right) \quad \leq 2k_n^d n^{-\epsilon_0^2}.$$

One may choose ϵ_0 large enough such that

$$\sum_n P(|I_{n,2}| \geq \epsilon) < \infty.$$

We conclude by Borel-Cantelli lemma and (26) that

$$\alpha_n \sup_{||u|| \leq n^\gamma} |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| = O_{a.s}(\alpha_n \sqrt{v_n}) = O_{a.s}(1).$$

Next, we have

$$\begin{aligned}
\sup_{u \in \mathbb{R}^d} \alpha_n |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| \quad &\leq \quad \sup_{||u|| \leq n^\gamma} \alpha_n |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| + \sup_{||u|| > n^\gamma} \alpha_n |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| \\
&= \quad \sup_{||u|| > n^\gamma} \alpha_n |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| + O_{a.s.}(1),
\end{aligned}$$

in view of the above result. Now, we have

$$\begin{aligned}
&\alpha_n \sup_{u:||u|| \geq n^\gamma} |G_{n,2}^x(u) - \overline{G}_{n,2}^x(u)| \\
\leq \quad &\alpha_n \sup_{u:||u|| \geq n^\gamma} |G_{n,2}^x(u)| + \alpha_n \sup_{u:||u|| \geq n^\gamma} |G^x(u)| + \alpha_n \sup_u |G^x(u) - \overline{G}_{n,2}^x(u)|. \qquad (27)
\end{aligned}$$

The last term in (27) is zero for large n, since conditioning by X, one may write

$$\alpha_n |\overline{G}_{n,2}^x(u) - G^x(u)| \quad = \quad \alpha_n |B_n^x(u)| = O_{a.s.}(h_n^\beta \alpha_n) =_{a.s.} (1)$$

in view Lemma 5.3 (i) whenever condition (10)(ii) is satisfied. For the second term in (27), we have

$$\alpha_n \sup_{||u|| > n^\gamma} G^x(u) \leq \frac{\alpha_n}{n^\gamma} \sup_{||u|| > n^\gamma} ||u|| G^x(u) = o(1),$$

whenever $\gamma > 1/2$ and the condition (11) is satisfied.

Moreover, we have for any $\epsilon > 0$

$$\begin{aligned}
&P\left(\alpha_n \sup_{u:||u|| \geq n^\gamma} |G_{n,2}^x(u)| \geq \epsilon \right) \\
\leq \quad &P\left(\alpha_n \sup_{u:||u|| \geq n^\gamma} \frac{1}{nE(\Delta_1)} \sum_{i:||Y_i - u|| > n^\gamma/2} ||Y_i - u|| \Delta_i(x)|| \geq \epsilon/2 \right) \\
+ \quad &P\left(\alpha_n \sup_{u:||u|| \geq n^\gamma} \frac{1}{nE(\Delta_1)} \sum_{i:||Y_i - u|| \leq n^\gamma/2} ||Y_i - u|| \Delta_i(x)|| \geq \epsilon/2 \right) := J_{n,1} +
\end{aligned}$$

To treat $J_{n,1}$, denote by

$$A_n(\omega) := \{\omega : \alpha_n \sup_{||u||>n^\gamma} \frac{1}{n} \sum_{i=1:||Y_i-u||>n^\gamma/2}^{n} ||Y_i - u||\Delta_i \geq \ /2\}.$$

The event $A_n(\omega)$ is nonempty if and only if there exists at least i_0 ($1 \leq i_0 \leq n$) such that $||Y_{i_0} - u|| > n^\gamma/2$. Thus "$A_n(\omega) = \square$" $\bigcup_{i=1}^{n}\{\omega : ||Y_i - u|| \geq n^\gamma/2\}$. It follows from Markov's inequality, if $\mathrm{E}(||Y_1 - u||) < \infty$, that

$$\mathrm{P}(A_n(\omega) = \square) = O(n^{-(\gamma-1)}) \quad \text{and} \quad \sum_n \mathrm{P}(A_n(\omega) = \square) < \infty,$$

whenever $\gamma > 1$, which implies that $J_{n,1} = o_{a.s.}(1)$ by Borel-Cantelli Lemma.

To deal with $J_{n,2}$, let us denote by

$$B_n(\omega) := \{\omega : \alpha_n \sup_{u:||u||\geq n^\gamma} \frac{1}{n\mathrm{E}(\Delta_1)} \sum_{i:||Y_i-u||\leq n^\gamma/2} ||Y_i - u||\Delta_i(x)|| \geq \ /2\}.$$

$B_n(\omega)$ is nonempty if and only if there exists at least i_0 ($1 \leq i_0 \leq n$

Proof of Theorem 3.2.

We have from the definitions of $\mu(x)$ and $\mu_n(x)$ and the existence and the uniqueness of these quantities that:

$$G^x(\mu(x)) = \inf_{u \quad \mathbb{R}^d} G^x(u) \quad \text{and} \quad G_n^x($$

(

Concerning the first term, observe that

$$\begin{aligned} \left|\widehat{H}_n^x(\xi_n(i)) - \widehat{H}_n^x(\mu)\right| &\le \frac{1}{n\,\mathrm{E}(\Delta_1(x))}\sum_{i=1}^{n} \left\| M(Y_i,\xi_n(j)) - M(Y_i,\mu)\right\| \Delta_i(x) \\ &:= A_n + B_n, \end{aligned} \tag{32}$$

where

$$A_n := \frac{\sqrt{d}}{n\mathrm{E}(\Delta_1(x))}\sum_{i=1}^{n}\frac{\left|\,\|Y_i-\mu\| - \|Y_i-\xi_n(j)\|\,\right|\Delta_i(x)}{\|Y_i-\mu\|\,\|Y_i-\xi_n(j)\|}$$

and

$$B_n := \frac{1}{n\mathrm{E}(\Delta_1(x))}\sum_{i=1}^{n}\Delta_i(x)\frac{\left\|\,\|Y_i-\xi_n(j)\|\,U(Y_i-\mu)\,U^T(Y_i-\mu) - \|Y_i-\mu\|\,U(Y_i-\xi_n(j))\,U^T(Y_i-\xi_n(j))\right\|}{\|Y_i-\mu\|\,\|Y_i-\xi_n(j)\|}.$$

Using Theorem 3.2 and the triangular inequality we can easily see that

We have to show that each term $K_{n,i}$ $(i = 1, 2)$ is asymptotically negligible. We have

$$\|K_{n,1}\|^2 = tr(K_{n,1}^T K_{n,1}) = \sum_{k=1}^{d} \sum_{j=1}^{d}$$

The result may be obtained by applying the Liapounov Central Theorem Limit. For this propose, we have to prove the following Lindeberg condition:

δ

Lemma 5.7 *Under conditions (H1)-(H2) and (H4)(ii), we have*

$$\sigma^2(x) = \lim_{n\to\infty} Var\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} {}^t\tilde{A}_i\right) = \frac{M_2}{M_1^2 g(x)}\ {}^t\Sigma^x(\mu)\ .$$

Proof of Lemma 5.7. Since the random variables $({}^t\tilde{A}_i)_{i=1,\ldots,n}$ are i.i.d. with mean zero, it follows that

$$\sigma^2(x) = \lim_{n\to\infty} Var\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} {}^t\tilde{A}_i\right) = \lim_{n\to\infty} Var({}^t\tilde{A}_1) = \lim_{n\to\infty} \mathrm{E}\left[({}^t A_1)^2\right].$$

On the other hand, making use of the properties of conditional expectation one may write

$$\mathrm{E}\left[\left({}^t A_1\right)^2\right] = \frac{\varphi(h)}{(\mathrm{E}\Delta_1)^2}\,\mathrm{E}\left[\Delta_1\ {}^tU(Y_1-\mu)\right]^2 = \frac{\varphi(h)}{(\mathrm{E}\Delta_1)^2}\,\mathrm{E}\left[\Delta_1^2 W_2^{X_1}(\mu)\right]$$

Making use of the condition (H4)(ii) and the fact that the functions $W_2^x(\cdot)$ is bounded, we obtain

E

$$A_1 := \mathrm{E}\left[K\left(\frac{d(x, X_1)}{h}\right)\psi(d(x, X_1))\right] = \int_0^1 K(t)\psi(th)dF(th),$$

where F is the cumulative distribution function of the real random variable $d(x, X)$. On the other

Write

$$
\begin{aligned}
V_n^x(\mu_n) &= \frac{M_{1,n}}{M_1}\sqrt{\frac{M_2}{M_{2,n}}}\sqrt{nF_{x,n}(h)}\,(n\varphi(h)g(x))^{-1}\,T_n^x(\mu_n)\,[T^x(\mu)]^{-1} \times \frac{M_1}{\sqrt{M_2}}\sqrt{n\varphi(h)g(x)}\,T^x(\mu)\,(\mu_n-\mu) \\
&:= V_{n,1}^x \times V_{n,2}^x. \qquad (33)
\end{aligned}
$$

Making use of Theorem 3.3

Besse, P., C. H., , and Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations. *J. Nonparametr. Statist.*, **27**, 673–687.

Besse, P. and Ramsay, J. O. (1986). Principal components analysis of sampled functions. *Psychometrika*, **51**(2), 285–311.

Bosq, D. (1996). *Nonparametric statistics for stochastic processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Estimation and prediction.

Bosq, D. (2000). *Linear processes in function spaces*, volume 149 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Theory and applications.

Cadre, B. (2001). Convergent estimators for the L_1-median of a Banach valued random variable. *Statistics*, **35**(4), 509–521.

Cardot, H., Crambes, C., and Sarda, P. (2005). Quantile regression when the covariates are functions. *J. Nonparametr. Stat.*, **17**(7), 841–856.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach. *Ann. Statist.*, **20**(2), 897–916.

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.*, **91**(434), 862–872.

Chebana, F. and Ouarda, T. B. M. J. (2011). Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, **22**(1), 63–78.

Cheng, Y. and De Gooijer, J. G. (2007). On the uth geometric conditional quantile. *J. Statist. Plann. Inference*, **137**(6), 1914–1930.

Collomb, G., Härdle, W., and Hassani, S. (1987). A note on prediction via estimation of the conditional mode function. *J. Statist. Plann. Inference*, **15**(2), 227–236.

Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Comm. Statist. Theory Methods*, **41**(7), 1254–1268.

Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Sankhyā*, **67**(2), 378–398.

Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data: inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**(3), 267–286.

Gannoun, A., Saracco, J., and Yu, K. (2003a). Nonparametric prediction by conditional median and quantiles. *J. Statist. Plann. Inference*, **117**(2), 207–223.

Gannoun, A., Saracco, J., Yuan, A., and Bonney, G. E. (2003b). On adaptive transformation-retransformation estimate of conditional spatial median. *Comm. Statist. Theory Methods*, **32**

Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stochastic Process. Appl.*, **115**(1), 155–177.

Quintela-del Río, A. and Francisco-Fernández, M. (2011). Nonparametric functional data estimation applied to ozone data: Prediction and extreme value analysis. *Chemosphere*, **82**.

Quintela-del Río, A., F. F. and Vieu, P. (2011). Analysis of time of occurrence of earthquakes: a functional data approach. *Math. Geosci.*, **43**, 695–719.