Consistency of the recursive nonparametric regression estimation for dependent functional data.

Aboubacar Amiri, Baba Thiam

Consistency of the recursive nonparametric regression estimation for dependent functional data

Aboubacar Amiri*and Baba Thiam †

April 18, 2013

Abstract

We consider the recursive estimation of a regression functional where the explanatory variables take values in some functional space. We prove the almost sure convergence of such estimates for dependent functional data. Also we derive the mean quadratic error of the considered class of estimators. Our results are established with rates and asymptotic appear bounds, under strong mixing condition.

Keywords: Functional data, recursive kernel estimators, regression function, quadratic mean error, almost sure convergence.
Classcode: 62G05, 62G07, 62G08, 62G20, 62L12.

1 Introduction

In this paper we study the regression model of a scalar response variable given a functional covariate. Functional data analysis is a problem of considerable interest in statistics and has been found to be useful in many practical fields, including climatology, economics, linguistics, medicine,... The statistical study of this kind of data is the subject of many papers in parametric and nonparametric statistics. For background material on this subject we highlight the works of Ramsay and Dalzell [17], Ramsay and Silverman [18, 19]. Since these pioneer contributions, the literature on this topic is still growing. A survey of the nonparametric functional regression appears in Ferraty et al. [11], while more recent results are collected in the book by Ferraty and Vieu [10]. There are several ways to study the link between a response variable given an explanatory variable. For example, one of the most studied models is the regression model when the response variable Y is real and the explanatory variable $\mathcal{X}$ belongs to some functional space $\mathcal{E}$. Then, the regression model writes $Y = r(\mathcal{X}) + \varepsilon$, where $r : \mathcal{E} \to \mathbb{R}$ is an operator and ε is an error random variable. Many works have been done around this model when the operator r is supposed to be linear, contributing to the popularity of the so-called functional linear model. We refer the reader for instance to the works of Cardot et al. [4] or Crambes et al. [5] for different methods to estimate r in this linear context. Another way is to estimate r by a nonparametric approach. The first results on this context were obtained by Ferraty and Vieu [8]. They established the almost complete convergence of a kernel estimator of the regression function in the i.i.d case. The study on their Nadaraya-Watson type estimator is extended to several directions. Dabo-Niang and Rhomari [6] stated the L^p-convergence of the kernel estimator, while Delsol [7] gave the L^p-convergence with asymptotic appear bound. The asymptotic normality of the same estimator has been obtained by Masry [15] under strong mixing conditions and extended by Delsol [7]. Ling and Wu [14] stated the almost sure convergence of the kernel estimator under strong mixing conditions. Functional data appear in many practical situations, as soon as one is interested on a continuous phenomenon. To consider such data as objects belonging to some functional space

*EQUIPPE
†EQUIPPE, Université Charles De Gaulle, Lille 3, Maison de la Recherche, Domaine universitaire du Pont de Bois, BP 60149, 59653 Villeneuve d'ascq cedex, France. baba.thiam@univ-lille3.fr

brings more precisions on the studied phenomenon. However, the computation of the estimators can be time consuming in this context, the use of recursive methods remains a good alternative to the classical ones. By 'recursive', we mean that the estimator calculated from the first n observations, say f_n, is only a function of f_{n-1} and the n^{th} observation. In this way, the estimator can be updated with each new observation added to the database.
The purpose of this paper is to apply recursive methods to functional data. Recursive estimation is achieved with the use of recursive estimators, typically kernel ones. For informations on nonparametric recursive methods, the reader is referred to the books by Gyorfi et al. [13], or the recent works of Vilar and Vilar [21], Wang and Liang [22], Quintela-Del-Rio [16], Amiri [1] and the references there in. The first results concerning the recursive kernel estimator of the regression function with functional explanatory variable were obtained by Amiri et al. [2]. They established the mean square error, the almost sure convergence with rates and a central limit theorem for a class of recursive kernel estimates of the regression function when the explanatory variable is functional and the observations are i.i.d. The main goal of this paper is the extension of a few of the results obtained by Amiri et al. [2] to dependent data. The rest of the paper proceeds as follows. We will present the regression model on section 2. On section 3, we give assumptions and results on the strong consistency and mean quadratic error for the recursive regression estimate. Section 4 is devoted to the proofs of our results.

2 Recursive regression estimate for curves

Let us consider a random process $Z_t = (\mathcal{X}_t, Y_t), t \in \mathbb{N}$, where Y_t is a scalar random variable and $\mathcal{X}_t$ takes values in some functional space $\mathcal{E}$ endowed with a semi-norm $\|\cdot\|$. Assume the existence of an operator r satisfying $r(\chi) := \mathbb{E}\left(Y_t|\mathcal{X}_t = \chi\right), \quad \chi \in \mathcal{E}$, for all $t \in \mathbb{N}$. To estimate r, one can consider the family of recursive estimators indexed by a parameter $\ell \in [0,1]$ introduced in Amiri et al. [2] and defined by

$$r_n^{[\ell]}(\chi) := \frac{\sum_{i=1}^{n} \frac{Y_i}{F(h_i)^\ell} K\left(\frac{\|\chi - \mathcal{X}_i\|}{h_i}\right)}{\sum_{i=1}^{n} \frac{1}{F(h_i)^\ell} K\left(\frac{\|\chi - \mathcal{X}_i\|}{h_i}\right)},$$

where K is a kernel, (h_n) a sequence of bandwidths and F is the cumulative distribution function of the random variable $\|\chi - \mathcal{X}\|$. This family of estimators is a recursive modification of the Nadaraya-Watson type estimator of Ferraty and Vieu [10] and can be computed recursively by

$$r_{n+1}^{[\ell]}(\chi) = \frac{\left[\sum_{i=1}^{n} F(h_i)^{1-\ell}\right] \varphi_n^{[\ell]}(\chi) + \left[\sum_{i=1}^{n+1} F(h_i)^{1-\ell}\right] Y_{n+1} K_{n+1}^{[\ell]}\left(\|\chi - \mathcal{X}_{n+1}\|\right)}{\left[\sum_{i=1}^{n} F(h_i)^{1-\ell}\right] f_n^{[\ell]}(\chi) + \left[\sum_{i=1}^{n+1} F(h_i)^{1-\ell}\right] K_{n+1}^{[\ell]}\left(\|\chi - \mathcal{X}_{n+1}\|\right)},$$

with

$$\varphi_n^{[\ell]}(\chi) = \frac{\sum_{i=1}^{n} \frac{Y_i}{F(h_i)^\ell} K\left(\frac{\|\chi - \mathcal{X}_i\|}{h_i}\right)}{\sum_{i=1}^{n} F(h_i)^{1-\ell}}, \quad f_n^{[\ell]}(\chi) = \frac{\sum_{i=1}^{n} \frac{1}{F(h_i)^\ell} K\left(\frac{\|\chi - \mathcal{X}_i\|}{h_i}\right)}{\sum_{i=1}^{n} F(h_i)^{1-\ell}}, \tag{1}$$

and $K_i^{[\ell]}(\cdot) := \frac{1}{F(h_i)^\ell \sum_{j=1}^{i} F(h_j)^{1-\ell}} K\left(\frac{\cdot}{h_i}\right)$. The recursive property of this class of regression estimators offers many advantages and is clearly useful in sequential investigations and also for a large sample size. Indeed, this kind of estimators are of easy implementation and interpretation, fast to compute and they do not require extensive storage of data. The weak and strong consistency of this family of estimators was studied by Amiri et al. [2] in the framework of the independent case.

3 Assumptions and main results

3.1 Assumptions

In the same spirit as Masry [15], we suppose throughout the paper the existence of nonnegative functions f_1 and ϕ such that $\phi(0) = 0$ and $F(h) = \mathbb{P}[\|\chi - \mathcal{X}\| \leq h] = \phi(h) f_1(\chi)$, for h on a neighborhood of zero. Then ϕ is an increasing function of h and $\phi(h) \to 0$ as $h \to 0$. The function f_1 is referred to as a functional probability density (see Gasser et al. [12] for more details). We will assume that the following assumptions hold.

(H1) The operators r and σ_ε^2 are continuous on a neighborhood of χ. Moreover, the function $\zeta(t) := \mathbb{E}[\{r(\mathcal{X}) - r(\chi)\} \ / \ \|\mathcal{X} - \chi\| = t]$ is assumed to be derivable at $t = 0$.

(H2) K is nonnegative bounded kernel with support on the compact $[0, 1]$ such that $\inf_{t\in[0,1]} K(t) > 0$.

(H3) For any $s \in [0, 1], \tau_h(s) := \frac{\phi(hs)}{\phi(h)} \to \tau_0(s) < \infty$ as $h \to 0$.

(H4) (i) $h_n \downarrow 0,\ n\phi(h_n) \to \infty,\ A_{n,\ell} := \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \frac{h_i}{h_n}\left[\frac{\phi(h_i)}{\phi(h_n)}\right]^{1-\ell} \to \alpha_{[\ell]} > 0$ as $n \to \infty$.

(ii) $\forall r \leq 2,\ B_{n,r} := \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\left[\frac{\phi(h_i)}{\phi(h_n)}\right]^{r} \to \beta_{[r]} > 0$, as $n \to \infty$.

(iii) For any $\mu > 0$, $\lim_{n\to\infty} \dfrac{(\ln n)^{3+\frac{2}{\mu}}}{n\phi(h_n)} = 0$.

(H5) (i) $(\mathcal{X}_t)_{t\in\mathbb{N}}$ is a strong mixing process with $\alpha_{\mathcal{X}}(k) \leq ck^{-\rho}$, $k \geq 1$, for some $c > 0$ and $\rho > 2$.

(ii) There exist non negative functions ψ and f_2 such that $\psi(h) \to 0$ as $h \to 0$, the ratio $\dfrac{\psi(h)}{\phi(h)^2}$ is bounded and $\sup_{i\neq j} \mathbb{P}[(\mathcal{X}_i, \mathcal{X}_j) \in \mathcal{B}(\chi, h_i) \times \mathcal{B}(\chi, h_j)] \leq \psi(h_i)\psi(h_j) f_2(\chi)$.

(H6) There exist $\lambda > 0$ and $\mu > 0$ such that $\mathbb{E}[\exp(\lambda|Y|^\mu)] < \infty$.

Since this paper is a generalization to dependent case of the results in Amiri et al. [2], several of the assumptions are the same as those used in the earlier reference. The reader is then referred to this last for more comments on assumptions. Let us mention that the decrease of the sequence $(h_n)_{n\in\mathbb{N}}$ is particular to the recursive estimators and for dependent data. The technical condition (H4)(iii) is unrestrictive and is easily satisfied by the popular choices of ϕ and h_n given by $\phi(h_n) \sim n^{-\xi}$, with $0 < \xi < 1$. Assumption (H5)(i) is the classically strong mixing condition which, is well known to be satisfied by linear or stationary ARMA processes. In order to simplify the presentation, we assume the strong mixing coefficient to be arithmetic, but the main results can be obtained under several conditions on this coefficient. Assumption (H5)(ii) plays a crucial role in our calculus, when we show the negligibility of some covariance terms. It has been used by Masry [15] in the non recursive case. Finally, as developped in Amiri et al. [2], assumption (H6) implies that

$$\mathbb{E}\left(\max_{1\leq i\leq n} |Y_i|^p\right) = O[(\ln n)^{p/\mu}], \forall p \geq 1, n \geq 2. \tag{2}$$

3.2 Main results

For convenience, let us introduce the following notations:

$$\begin{aligned} M_0 &= K(1) - \int_0^1 (sK(s))'\tau_0(s)ds, \quad M_1 = K(1) - \int_0^1 K'(s)\tau_0(s)ds \\ M_2 &= K^2(1) - \int_0^1 (K^2(s))'\tau_0(s)ds. \end{aligned}$$

In the following theorem, we establish the almost sure convergence of the proposed recursive kernel estimator of the regression function.

Theorem 1 *Assume that (H1)-(H6) hold. If* $\lim_{n\to+\infty} nh_n^2 = 0$, *then*

$$\limsup_{n\to\infty}\left[\frac{n\phi(h_n)}{\ln\ n}\right]^{1/2}\left[r_n^{[\ell]}(\chi)-r(\chi)\right]\leq\frac{2}{M_1}\left[1+V_\ell(\chi)\right]\ \ a.s.$$

where

$$V_\ell(\chi)=\frac{\beta_{[1-2\ell]}}{\beta^2_{[1-\ell]}}\frac{\sigma^2_\varepsilon(\chi)}{f_1(\chi)}M_2, \tag{3}$$

for all χ *such that* $f_1(\chi)>0$.

Theorem 1 is an extension of Ferraty and Vieu's [9] result on functional kernel-type estimate to the general family of recursive estimators $r_n^{[\ell]}(\chi)$. A similar result is also obtained by Ling and Wu [14] for a truncated version of the Nadaraya-Watson type estimator, under the condition $\mathbb{E}|Y|<\infty$, which is weaker than assumption (H6). However, Theorem 1 establishes the rate of convergence with exact appear bound, while Ling and Wu's [14] result tells only the rate of convergence in function of the variances of the numerator and denominator of the estimator. As we will see in the proofs below, assumption (H6) will be necessary, for the study of the covariance terms and also when we shall prove the cancellation of the residual term between the estimator and its truncated version. Finally, let us mention that compared with the result in Amiri et al. [2], as in the multivariate framework, it is difficult to obtain the optimal rate $\left[\frac{n\phi(h_n)}{\ln\ln\ n}\right]^{1/2}$ in the dependent case.

The mean square error of $r_n^{[\ell]}(x)$ is given in Theorem 2 below.

Theorem 2 *Under assumptions (H1)-(H6),*

$$\mathbb{E}\left[\left(r_n^{[\ell]}(\chi)-r(\chi)\right)^2\right]=\left[\left(\frac{\zeta'(0)\alpha_{[\ell]}M_0h_n}{\beta_{[1-\ell]}M_1}\right)^2+\frac{\beta_{[1-2\ell]}M_2\sigma^2_\varepsilon(\chi)}{\beta^2_{[1-\ell]}M_1^2f_1(\chi)n\phi(h_n)}\right]\left[1+o(1)\right]$$

for all χ *such that* $f_1(\chi)>0$.

Theorem 2 is an extension to functional data of the result of Amiri [1] in finite dimensional setting. Also, our result generalizes the works of Bosq and Cheze-Payaud [3] to functional and recursive setting. Finally, in counterpart of the almost sure convergence, Theorem 2 gives the same rate of convergence and asymptotic constants as those obtained for the iid case in Amiri et al. [2].

4 Proofs

In the sequel, and through the paper, c will denote a constant whose value is unimportant and may vary from line to line. Also, we set

$$K_i(\chi)=K\left(\frac{\|\chi-\mathcal{X}_i\|}{h_i}\right).$$

Finally, for convenience we will use the following decomposition

$$r_n^{[\ell]}(\chi)-r(\chi)=\frac{\tilde{\varphi}_n^{[\ell]}(\chi)-r(\chi)f_n^{[\ell]}(\chi)}{f_n^{[\ell]}(\chi)}+\frac{\varphi_n^{[\ell]}(\chi)-\tilde{\varphi}_n^{[\ell]}(\chi)}{f_n^{[\ell]}(\chi)}, \tag{4}$$

where $\tilde{\varphi}_n^{[\ell]}(\chi)$ is a truncated version of $\varphi_n^{[\ell]}(\chi)$ defined by

$$\tilde{\varphi}_n^{[\ell]}(\chi)=\frac{1}{\sum_{i=1}^n F(h_i)^{1-\ell}}\sum_{i=1}^n\frac{Y_i}{F(h_i)^\ell}\mathbb{1}_{\{|Y_i|\leq b_n\}}K\left(\frac{\|\chi-\mathcal{X}_i\|}{h_i}\right), \tag{5}$$

b_n being a sequence of real numbers which goes to $+\infty$ as $n\to\infty$.

4.1 Preliminary lemmas

In order to prove the main results, we need the following lemmas.

Lemma 1 *Under assumptions (H1)-(H4), we have*

$$\frac{\mathbb{E}\left[\varphi_n^{[\ell]}(\chi)\right]}{\mathbb{E}\left[f_n^{[\ell]}(\chi)\right]} - r(\chi) = h_n\zeta'(0)\frac{\alpha_{[\ell]}}{\beta_{[1-\ell]}}\frac{M_0}{M_1}\left[1+o(1)\right].$$

Proof. See Amiri et al. [2], since the bias term is not depending to the mixing structure. □

Lemma 2 *Under assumptions (H1)-(H6), we have*

$$\begin{aligned}
Var\left[f_n^{[\ell]}(\chi)\right] &= \frac{\beta_{[1-2\ell]}}{\beta_{[1-\ell]}^2}\frac{M_2}{f_1(\chi)}\frac{1}{n\phi(h_n)}\left[1+o(1)\right];\\
Var\left[\varphi_n^{[\ell]}(\chi)\right] &= \frac{\beta_{[1-2\ell]}}{\beta_{[1-\ell]}^2}\left[r^2(\chi)+\sigma_\epsilon^2(\chi)\right]\frac{M_2}{f_1(\chi)}\frac{[1+o(1)]}{n\phi(h_n)};\\
Cov\left[f_n^{[\ell]}(\chi),\varphi_n^{[\ell]}(\chi)\right] &= \frac{\beta_{[1-2\ell]}}{\beta_{[1-\ell]}^2}r(\chi)\frac{M_2}{f_1(\chi)}\frac{[1+o(1)]}{n\phi(h_n)},
\end{aligned}$$

for all χ such that $f_1(\chi)>0$.

Proof. The variance term of $f_n^{[\ell]}(\chi)$ can be decomposed in variance and covariance terms as

$$\mathrm{Var}(f_n^{[\ell]}(\chi)) = \left[\sum_{i=1}^n F(h_i)^{1-\ell}\right]^{-2}\left(\sum_{i=1}^n A_{i,i}+\sum_{i\neq j}A_{i,j}\right) := F_1+F_2, \tag{6}$$

where for any integers i and j, $A_{i,j}=F(h_i)^{-\ell}F(h_j)^{-\ell}\mathrm{Cov}\left(K_i(\chi),K_j(\chi)\right)$. Noting that the principal term F_1 in the right-hand side of (6) corresponds to the variance term of $f_n^{[\ell]}$ in the independent case (see Amiri et al. [2] fore more details), and is given by

$$n\phi(h_n)F_1 = \frac{\beta_{[1-2\ell]}}{\beta_{[1-\ell]}^2}\frac{M_2}{f_1(\chi)}\left[1+o(1)\right].$$

Now, let us establish that the covariance term F_2 is negligible. To this end, let c_n be a sequence of real numbers tending to ∞ as $n\to\infty$. We can write

$$F_2 \leq \frac{2\left(\sum_{k=1}^{c_n}\sum_{p=1}^n|A_{k+p,p}|+\sum_{k=c_n+1}^{n-1}\sum_{p=1}^n|A_{k+p,p}|\right)}{\left[\sum_{i=1}^n F(h_i)^{1-\ell}\right]^2} := F_{21}+F_{22}. \tag{7}$$

From assumptions (H2) and (H5)(ii), we have for any $i\neq j$

$$\begin{aligned}
\mathbb{E}\left[K_i(\chi)K_j(\chi)\right] &= \int_{[0,1]\times[0,1]}K(u)K(v)d\mathbb{P}^{\left(\frac{\|\chi-\mathcal{X}_i\|}{h_i},\frac{\|\chi-\mathcal{X}_j\|}{h_j}\right)}(u,v)\\
&\leq \|K\|_\infty^2\mathbb{P}\left[\|\chi-\mathcal{X}_i\|\leq h_i,\|\chi-\mathcal{X}_j\|\leq h_j\right]\\
&\leq c\psi(h_i)\psi(h_j).
\end{aligned} \tag{8}$$

Note that, from the proof of Lemma 2 in Amiri et al. [2] we can write

$$\mathbb{E}\left[K_i(\chi)\right] = \phi(h_i)f_1(\chi)\left[K(1)-\int_0^1 K'(s)\tau_{h_i}(s)ds\right],$$

so, we get

$$|\mathrm{Cov}(K_i(\chi),K_j(\chi))| \leq c\left[\psi(h_i)\psi(h_j)+\phi(h_i)\phi(h_j)\right]. \tag{9}$$

Hence, we deduce that

$$F_{21} \leq \frac{c\sum_{k=1}^{c_n}\sum_{p=1}^{n}\left[\frac{\psi(h_{k+p})\psi(h_p)}{\phi(h_{k+p})^{\ell}\phi(h_p)^{\ell}}+\phi(h_{k+p})^{1-\ell}\phi(h_p)^{1-\ell}\right]}{[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}]^2} := F_{211}+F_{212}.$$

Now, Assumption (H5)(ii) ensures that the ratio $\psi(h_i)/\phi(h_i)$ is bounded and since ϕ is increasing, we get

$$F_{211} \leq c\left[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}\right]^{-2}\sum_{k=1}^{c_n}\sum_{p=1}^{n}\phi(h_p)^{2-2\ell} \leq c\frac{B_{n,2-2\ell}}{B_{n,1-\ell}^2}\frac{c_n}{n}.$$

Hence,

$$n\phi(h_n)F_{211} = O\left(\phi(h_n)c_n\right). \tag{10}$$

Now for the second term F_{212}, again, using the fact that ϕ is an increasing function, we get

$$F_{212} \leq c\left[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}\right]^{-2}\sum_{k=1}^{c_n}\sum_{p=1}^{n}\phi(h_p)^{2-2\ell} \leq c\frac{B_{n,2-2\ell}}{B_{n,1-\ell}^2}\frac{c_n}{n},$$

so that

$$n\phi(h_n)F_{212} = O\left(\phi(h_n)c_n\right). \tag{11}$$

From (10) and (11) we deduce

$$n\phi(h_n)F_{21} = O\left(\phi(h_n)c_n\right). \tag{12}$$

Next, for the second term F_{22} in (7), we have from Billingsley's inequality,

$$\begin{aligned} F_{22} &\leq c\left[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}\right]^{-2}\sum_{k=c_n+1}^{n-1}\sum_{p=1}^{n}k^{-\rho}\phi(h_{k+p})^{-\ell}\phi(h_p)^{-\ell} \\ &\leq c\left[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}\right]^{-2}\frac{c_n^{1-\rho}}{\rho-1}\sum_{p=1}^{n}\phi(h_n)^{-\ell}\phi(h_p)^{-\ell} \leq c\frac{B_{n,-\ell}}{B_{n,1-\ell}^2}\frac{c_n^{1-\rho}}{n\phi(h_n)^2}. \end{aligned}$$

Therefore

$$n\phi(h_n)F_{22} = O\left(\frac{c_n^{1-\rho}}{\phi(h_n)}\right). \tag{13}$$

If we choose $c_n = \lfloor\phi(h_n)^{-\frac{2}{\rho}}\rfloor$, we deduce from (12) and (13) that

$$n\phi(h_n)F_2 = O\left(\phi(h_n)^{\frac{\rho-2}{\rho}}\right) = o(1) \text{ as long as } \rho > 2,$$

and the first part of Lemma 2 follows. Now, as in the proof of the first part of Lemma 2, the variance term of $\varphi_n^{[\ell]}$ is decomposed as follows

$$\mathrm{Var}(\varphi_n^{[\ell]}(\chi)) = \left[\sum_{i=1}^{n}F(h_i)^{1-\ell}\right]^{-2}\left(\sum_{i=1}^{n}A_{i,i}+\sum_{i\neq j}A_{i,j}\right) := I_1+I_2,$$

where here $A_{i,j}$ denotes for any integers i and j as follows $A_{i,j} = F(h_i)^{-\ell}F(h_j)^{-\ell}\mathrm{Cov}\left(Y_iK_i(\chi), Y_jK_j(\chi)\right)$. The study of the term I_1 is treated in the same manner as in the independent case (see Amiri et al. [2] for more details) which gives

$$n\phi(h_n)I_1 = \frac{\beta_{[1-2\ell]}}{\beta_{[1-\ell]}^2}\frac{r(\chi)}{f_1(\chi)}M_2\left[1+o(1)\right].$$

Now, for the second term I_2, we always consider a sequence of real numbers c_n which goes to ∞ as $n \to \infty$ and we write

$$I_2 \leq \frac{2\left(\sum_{k=1}^{c_n}\sum_{p=1}^{n}|A_{k+p,p}| + \sum_{k=c_n+1}^{n-1}\sum_{p=1}^{n}|A_{k+p,p}|\right)}{\left[\sum_{i=1}^{n}F(h_i)^{1-\ell}\right]^2} := I_{21} + I_{22}. \tag{14}$$

The term I_{22} is treated exactly as F_{22} in the proof of the first part of this Lemma previously, by substituting the Billingsley lemma with the Davydov lemma. Then, setting $b_n = (\delta \ln n)^{1/\mu}$, using (2) and with the help of (H5), we get

$$\begin{aligned} I_{22} &\leq c(\ln n)^{2/\mu}\left[\sum_{i=1}^{n}\phi(h_i)^{1-\ell}\right]^{-2}\frac{c_n^{1-\rho/2}}{(\rho/2)-1}\phi(h_n)^{-2\ell}\sum_{p=1}^{n}\left(\frac{\phi(h_p)}{\phi(h_n)}\right)^{-\ell} \\ &\leq c(\ln n)^{2/\mu}\frac{B_{n,-\ell}}{B_{n,1-\ell}^2}\frac{1}{n\phi(h_n)^2}c_n^{1-\rho/2}. \end{aligned}$$

Therefore,

$$n\phi(h_n)I_{22} = O\left((\ln n)^{2/\mu}c_n^{1-\rho/2}(\phi(h_n))^{-1}\right). \tag{15}$$

For the second term I_{21}, observe that for any integers i and j,

$$|\mathrm{Cov}(Y_iK_i(\chi), Y_jK_j(\chi))| \leq |\mathbb{E}\left[Y_iY_jK_i(\chi)K_j(\chi)\right]| + |\mathbb{E}\left[Y_iK_i(\chi)\right]|\,|\mathbb{E}\left[Y_jK_j(\chi)\right]|.$$

Now, from assumptions (H1) and (H2) and conditioning on $\mathcal{X}$, one have

$$\mathbb{E}\left[Y_iK_i(\chi)\right] = M_1F(h_i)\left[r(\chi)+\gamma_i\right] \leq c\phi(h_i),$$

where γ_i goes to zero as $i \to \infty$. Using Cauchy-Schwartz' inequality, choosing $b_n = (\ln n)^{1/\mu}$ and (2), we get

$$\begin{aligned} |\mathrm{Cov}(Y_iK_i(\chi), Y_jK_j(\chi))| &\leq \mathbb{E}^{1/2}\left[Y_i^2Y_j^2\right]\mathbb{E}^{1/2}\left[K_i^2(\chi)K_j^2(\chi)\right] \\ &\quad + |\mathbb{E}\left[Y_iK_i(\chi)\right]|\,|\mathbb{E}\left[Y_jK_j(\chi)\right]| \\ &\leq c\left[(\ln n)^{2/\mu}\psi(h_i)^{1/2}\psi(h_j)^{1/2} + \phi(h_i)\phi(h_j)\right]. \end{aligned}$$

The rest of the proof for I_{21} is the same as the one for F_{21} which implies that $I_{21} \leq c\frac{c_n}{n}\left[(\ln n)^{2/\mu}+1\right]$. Hence,

$$n\phi(h_n)I_{21} = O\left(c_n\phi(h_n)(\ln n)^{2/\mu}\right), \tag{16}$$

and the result of the second part of Lemma 2 follows from (15) and (16) with the choice $c_n = \lfloor \phi_n^{-\frac{4}{\rho}} \rfloor$. Next, to treat the last part of Lemma 2, it suffices to decompose the term $n\phi(h_n)\mathrm{Cov}\left(\phi_n^{[\ell]}(\chi), f_n^{[\ell]}(\chi)\right)$ by the principal and covariance terms and use the same procedure as in the proof of the second part of Lemma 2. □

Lemma 3 *Set*

$$N = \left[\frac{n\phi(h_n)}{\ln\ n}\right]^{1/2}\left\{\tilde{\varphi}_n^{[\ell]}(\chi) - r(\chi)f_n^{[\ell]}(\chi) - \mathbb{E}\left[\tilde{\varphi}_n^{[\ell]}(\chi) - r(\chi)f_n^{[\ell]}(\chi)\right]\right\},$$

where $\tilde{\varphi}$ is defined in (5)*. Under assumptions (H1)-(H6), we have*

$$\overline{\lim_{n\to\infty}}\ N \leq 2\left[1 + V_\ell(\chi)\right]\ \ a.s.,$$

where V_ℓ is defined in (3)*.*

Proof. Set $W_{n,i} = \dfrac{K_i(\chi)\left[Y_i \mathbb{1}_{\{|Y_i|\le b_n\}} - r(\chi)\right]}{f_1(\chi) B_{n,1-\ell}\phi(h_n)^{1-\ell}\phi(h_i)^{\ell}}$, where $Z_{n,i} = W_{n,i} - \mathrm{E}W_{n,i}$. To prove Lemma 3, we use the blocks decomposition technique. Let p_n and q_n be some sequences of real numbers defined by $p_n = \lfloor p_0 \ln n \rfloor$ with $p_0 > 0$ and $q_n = \lfloor \frac{n}{2p_n} \rfloor$. Set

$$S'_n = \sum_{j=1}^{q_n} V_n(2j-1), \quad S''_n = \sum_{j=1}^{q_n} V_n(2j) \text{ and } S'''_n = \frac{1}{n} \sum_{k=2p_nq_n+1}^{n} Z_{n,k},$$

with $V_n(j) = \frac{1}{n} \sum_{k=(j-1)p_n+1}^{jp_n} Z_{n,k}, \quad j = 1, \ldots, 2q_n$. Then we have $N = S'_n + S''_n + S'''_n$. Observe that the third term S'''_n is negligible so that, to prove the strong consistency of N, it suffices to check the almost sure convergence for $S'_n + S''_n$. For any $\varepsilon > 0$,

$$\mathbb{P}\left(|S'_n + S''_n| > \varepsilon\right) \le \mathbb{P}\left(|S'_n| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(|S''_n| > \frac{\varepsilon}{2}\right).$$

We just treat S'_n, the term S''_n being similar. Since K is bounded and ϕ is non decreasing, we get for n large enough $|V_n(j)| \le \frac{2\|K\|_\infty p_n b_n}{f_1(\chi) B_{n,1-\ell} n\phi(h_n)}$. Using Rio's [20] coupling lemma, the random variables $V_n(j)$ can be approximated by independent and identically distributed random variables $V_n^*(j)$ such that

$$\mathbb{E}\left|V_n(2j-1) - V_n^*(2j-1)\right| \le \frac{4\|K\|_\infty p_n b_n}{f_1(\chi) B_{n,1-\ell} n\phi(h_n)} \alpha(p_n).$$

Since $p_n q_n \le n$, it follows that

$$\begin{aligned} \sum_{j=1}^{q_n} \mathbb{E}\left|V_n(2j-1) - V_n^*(2j-1)\right| &\le \frac{4\|K\|_\infty p_n q_n b_n}{f_1(\chi) B_{n,1-\ell} n\phi(h_n)} \alpha(p_n) \\ &\le \frac{4\|K\|_\infty b_n}{f_1(\chi) B_{n,1-\ell} \phi(h_n)} \alpha(p_n). \end{aligned}$$

Therefore, for any $\varepsilon, \kappa > 0$, Markov's inequality leads to

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{j=1}^{q_n} \left[V_n(2j-1) - V_n^*(2j-1)\right]\right| > \frac{\varepsilon\kappa}{2(1+\kappa)}\right) &\le \frac{8(1+\kappa)}{\varepsilon\kappa} \frac{\|K\|_\infty b_n \alpha(p_n)}{f_1(\chi) B_{n,1-\ell}\phi(h_n)} \\ &\le \frac{8(1+\kappa)}{\varepsilon\kappa} \frac{\|K\|_\infty b_n \gamma e^{-\rho p_0 \ln n}}{f_1(\chi) B_{n,1-\ell}\phi(h_n)}. \end{aligned} \tag{17}$$

Next setting $\varepsilon_n = \varepsilon\sqrt{\dfrac{\ln n}{n\phi(h_n)}}$ with $\varepsilon > 0$ and $\lambda_n = \sqrt{n\phi(h_n)\ln n}$, we have from (H4)(iii), $|\lambda_n V_n^*(j)| \le \frac{2\|K\|_\infty p_n b_n}{f_1(\chi) B_{n,1-\ell}} \sqrt{\dfrac{\ln n}{n\phi(h_n)}} \to 0$, therefore for n large enough, $|\lambda_n V_n^*(j)| \le \frac{1}{2}$. It follows that

$$\exp\left\{\pm\lambda_n V_n^*(j)\right\} \le 1 \pm \lambda_n V_n^*(j) + \left[\lambda_n V_n^*(j)\right]^2.$$

From Markov's inequality, we get

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{j=1}^{q_n} V_n^*(2j-1)\right| > \frac{\varepsilon_n}{2(1+\kappa)}\right) &\le \mathbb{P}\left(\exp\left[\sum_{j=1}^{q_n} \lambda_n V_n^*(2j-1)\right] > \exp\left[\frac{\lambda_n\varepsilon_n}{2(1+\kappa)}\right]\right) \\ &\quad + \mathbb{P}\left(\exp\left[-\sum_{j=1}^{q_n} \lambda_n V_n^*(2j-1)\right] > \exp\left[\frac{\lambda_n\varepsilon_n}{2(1+\kappa)}\right]\right) \\ &\le 2\exp\left[\frac{-\lambda_n\varepsilon_n}{2(1+\kappa)} + \lambda_n^2 \sum_{j=1}^{q_n} \mathbb{E}V_n^{*2}(2j-1)\right]. \end{aligned}$$

Since, $\sum_{j=1}^{q_n} \mathbb{E}V_n^{*2}(2j-1) \leq \frac{1}{n^2}\left[\sum_{k=1}^{n}\mathrm{Var}(Z_{n,k}) + \sum_{k\neq k'}\mathrm{Cov}(Z_{n,k},Z_{n,k'})\right]$, we will assume for the moment that

$$\frac{\phi(h_n)}{n}\sum_{k=1}^{n}\mathrm{Var}(Z_{n,k}) = V_\ell(\chi)\left[1+o(1)\right] \tag{18}$$

$$\frac{\phi(h_n)}{n}\sum_{k\neq k'}\mathrm{Cov}(Z_{n,k},Z_{n,k'}) = o(1), \tag{19}$$

where V_ℓ is defined in (3). It follows from (18) and (19) that, for n large enough,

$$\lambda_n^2\sum_{j=1}^{q_n}\mathbb{E}V_n^{*2}(2j-1) \leq V_\ell(\chi)\ln n\left[1+o(1)\right].$$

Therefore

$$\mathbb{P}\left(\left|\sum_{j=1}^{q_n}V_n^*(2j-1)\right| > \frac{\varepsilon_n}{2(1+\kappa)}\right) \leq 2e^{\left[\frac{-\varepsilon}{2(1+\kappa)}+V_\ell(\chi)(1+o(1))\right]\ln n} \tag{20}$$

Now, combining (17) and (20), we get

$$\begin{aligned}
\mathbb{P}\left(|S_n'| > \frac{\varepsilon_n}{2}\right) &\leq \mathbb{P}\left(\left|\sum_{j=1}^{q_n}V_n(2j-1)-V_n^*(2j-1)\right| > \frac{\varepsilon_n\kappa}{2(1+\kappa)}\right)\\
&\quad +\mathbb{P}\left(\left|\sum_{j=1}^{q_n}V_n^*(2j-1)\right| > \frac{\varepsilon_n\kappa}{2(1+\kappa)}\right)\\
&\leq \gamma\frac{8(1+\kappa)}{\varepsilon\kappa}\frac{\|K\|_\infty b_n n^{1-p_0\rho}}{f_1(\chi)B_{n,1-\ell}\sqrt{\ln n}\sqrt{n\phi(h_n)}}\\
&\quad +2\exp\left\{\left[-\frac{\varepsilon}{2(1+\kappa)}+V_\ell(\chi)\right]\ln n\right\}.
\end{aligned} \tag{21}$$

Next, with the choice of $b_n = (\delta\ln n)^{1/\mu}$, the conclusion follows from the application of the Borel-Cantelli's lemma whenever $p_0 > \frac{2}{\rho}$ and $\varepsilon > 2(1+\kappa)\left[1+V_\ell(\chi)\right]$, which implies that

$$\overline{\lim_{n\to\infty}}\left[\frac{n\phi(h_n)}{\ln\ n}\right]^{1/2}N \leq 2(1+\kappa)\left[1+V_\ell(\chi)\right]\ \text{a.s.},$$

for all positive κ and Lemma 3 follows. To complete the proof, let us prove (18) and (19). We can write

$$\frac{\phi(h_n)}{n}\sum_{k=1}^{n}\mathrm{Var}(Z_{n,k}) = \frac{\sum_{k=1}^{n}\phi(h_k)^{-2\ell}\mathrm{Var}\left(K_i(\chi)\left[Y_i\mathbb{1}_{\{|Y_i|\leq b_n\}}-r(\chi)\right]\right)}{f_1^2(\chi)B_{n,1-\ell}^2 n\phi(h_n)^{1-2\ell}}.$$

Following the same lines of the proof of Lemma 5 in Amiri et al. [2], one can prove that

$$\sum_{k=1}^{n}\phi(h_k)^{-2\ell}\mathrm{Var}\left(K_k(\chi)\left[Y_k\mathbb{1}_{\{|Y_k|\leq b_n\}}-r(\chi)\right]\right) \sim n\phi(h_n)^{1-2\ell}\beta_{[1-2\ell]}\sigma_\varepsilon^2(\chi)M_2,$$

therefore (18) follows. Next, about the covariance term in (19), for any integers $i \neq j$, let

$$A_{i,j} = F(h_i)^{-\ell}F(h_j)^{-\ell}\left|\mathrm{Cov}\left(K_i(\chi)\left[Y_i\mathbb{1}_{\{|Y_i|\leq b_n\}}-r(\chi)\right], K_j(\chi)\left[Y_j\mathbb{1}_{\{|Y_j|\leq b_n\}}-r(\chi)\right]\right)\right|.$$

Then, we have

$$\begin{aligned}\frac{1}{n^2}\sum_{k\neq k'}\mathrm{Cov}(Z_{n,k},Z_{n,k'}) &\leq \frac{2\left[\sum_{k=1}^{c_n}\sum_{p=1}^{n}|A_{k+p,p}|+\sum_{k=c_n+1}^{n-1}\sum_{p=1}^{n}|A_{k+p,p}|\right]}{f_1^2(\chi)B_{n,1-\ell}^2 n^2\phi(h_n)^{2-2\ell}}\\ &:= J_1+J_2.\end{aligned}$$

Using Billingsley's inequality, one can prove that $n\phi(h_n)J_2 = O\left(b_n^2 c_n^{1-\rho}\phi(h_n)^{-1}\right)$. Next, since

$$\begin{aligned}|A_{k+p,p}| &\leq (b_n+|r(\chi)|)^2\left[\mathbb{E}\left(K_{k+p}(\chi)K_p(\chi)\right)+\mathbb{E}(K_{k+p}(\chi))\mathbb{E}(K_p(\chi))\right]\\ &\leq c\,(b_n+|r(\chi)|)^2\left[\psi(h_{k+p})\psi(h_p)+\phi(h_{k+p})\phi(h_p)\right]\phi(h_{k+p})^{-\ell}\phi(h_p)^{-\ell}.\end{aligned}$$

Therefore, as in the proof of the first part of Lemma 2, we get

$$n\phi(h_n)J_1 = O(b_n^2\phi(h_n)c_n),$$

which together with the choice $c_n = \lfloor\phi(h_n)^{-\frac{2}{\rho}}\rfloor$ imply (19) as long as $\rho > 2$. □

4.2 Proofs of the main results

4.2.1 Proof of Theorem 1

Let us consider the decomposition (4). For the residual term $\frac{\varphi_n^{[\ell]}(\chi)-\tilde{\varphi}_n^{[\ell]}(\chi)}{f_n^{[\ell]}(\chi)}$, following the same lines of proof in Amiri et al. [2] by replacing $\frac{\ln\ln n}{nF(h_n)}$ by $\frac{\ln n}{n\phi(h_n)}$, one can show that

$$\left[\frac{n\phi(h_n)}{\ln n}\right]^{1/2}\left|\varphi_n^{[\ell]}(\chi)-\tilde{\varphi}_n^{[\ell]}(\chi)\right| \to 0 \text{ a.s, when } n\to\infty. \tag{22}$$

For the principal term in (4), we can write

$$\begin{aligned}\tilde{\varphi}_n^{[\ell]}(\chi)-r(x)f_n^{[\ell]}(\chi) &= \left\{\tilde{\varphi}_n^{[\ell]}(\chi)-r(\chi)f_n^{[\ell]}(\chi)-\mathbb{E}\left[\tilde{\varphi}_n^{[\ell]}(\chi)-r(\chi)f_n^{[\ell]}(\chi)\right]\right\}\\ &\quad+\left\{\mathbb{E}\left[\tilde{\varphi}_n^{[\ell]}(\chi)-r(\chi)f_n^{[\ell]}(\chi)\right]\right\}.\end{aligned} \tag{23}$$

Noting that, from Lemma 3 in Amiri et al. [2], we have $\mathbb{E}\left(f_n^{[\ell]}(\chi)\right) = M_1[1+o(1)]$ and it can be shown as the same lines of the proof of Lemma 3 that

$$f_n^{[\ell]}(\chi)-\mathbb{E}\left(f_n^{[\ell]}(\chi)\right) = O\left(\sqrt{\frac{\ln n}{n\phi(h_n)}}\right).$$

Therefore, Theorem 1 follows from the combination of Lemmas 1 and 3, since from Lemma 1, if $\lim_{n\to+\infty} nh_n^2 = 0$, then $\overline{\lim_{n\to\infty}}\left[\frac{n\phi(h_n)}{\ln n}\right]^{1/2}\left\{\mathbb{E}\left[\tilde{\varphi}_n^{[\ell]}(\chi)-r(\chi)f_n^{[\ell]}(\chi)\right]\right\} = 0$.

4.3 Proof of Theorem 2

The mean square error of $r_n^{[\ell]}$ can be decomposed as follow:

$$\begin{aligned}\mathbb{E}\left[\left(r_n^{[\ell]}(\chi)-r(\chi)\right)^2\right] &= \frac{r^2(\chi)\mathrm{Var}\left(f_n^{[\ell]}(\chi)\right)}{\mathbb{E}^2(f_n^{[\ell]}(\chi))}-\frac{2r(\chi)\mathrm{Cov}\left(\phi_n^{[\ell]}(\chi),f_n^{[\ell]}(\chi)\right)}{\mathbb{E}^2(f_n^{[\ell]}(\chi))}\\ &\quad+\frac{\mathrm{Var}\left(\varphi_n^{[\ell]}(\chi)\right)}{\mathbb{E}^2(f_n^{[\ell]}(\chi))}+\frac{\mathbb{E}\varphi_n^{\ell}(\chi)-r(\chi)\mathbb{E}f_n^{[\ell]}(\chi)}{\mathbb{E}^2 f_n^{[\ell]}(\chi)}\\ &\quad+o(h_n^2)+o\left(1/\left(n\phi(h_n)\right)\right).\end{aligned}$$

Theorem 2 follows from Lemmas 1 - 2. □

References

[1] A. Amiri, *Recursive regression estimators with application to nonparametric prediction*, J. Nonparametr. Stat. 24(1) (2012) , pp. 169–186.

[2] A. Amiri, C. Crambes, B. Thiam, *Recursive estimation of nonparametric regression with functional covariate*, under review, arXiv:1211.2780 (2012).

[3] D. Bosq, N. Cheze-Payaud, *Optimal asymptotic quadratic error of nonparametric regression function estimates for a continuous-time process from sampled-data*, Statistics, 32(3) (1999), pp. 229–247.

[4] H. Cardot, F. Ferraty, P. Sarda, *Splines estimators for the functional linear model*, Statistica Sinica, 13 (2003), pp. 571–591.

[5] C. Crambes, A. Kneip, P. Sarda, *Smoothing splines estimators for functional linear regression*, Ann. Statist., 37 (2009), pp. 35-72.

[6] S. Dabo-Niang, N. Rhomari, *Estimation non paramétrique de la régression avec variable explicative dans un espace métrique*, C. R. Math. Acad. Sci. Paris, 336 (2003), pp.75–80.

[7] L. Delsol, *Advances on asymptotic normality in nonparametric functional time series analysis*, Statistics 43(1) (2009), pp. 13–33.

[8] F. Ferraty, P. Vieu, *Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés*, C. R. Math. Acad. Sci. Paris, 330 (2000), pp. 139–142.

[9] F. Ferraty, P. Vieu, *Nonparametric model for functional data with application in regression, time series prediction and cure discrimination*, J. Nonparametr. Stat. 16 (2004), pp. 111–125.

[10] F. Ferraty, P. Vieu, *Nonparametric modelling for functional data. methods, theory, applications and implementations*, Springer-Verlag, London (2006).

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

[12] T. Gasser, P. Hall, B. Presnel, *Nonparametric estimation of the mode of a distribution of random curves*, J. Roy. Statist. Soc. Ser. B, 60 (1998), pp. 681–691.

[13] L. Gyorfi, M. Kholer, A. Krzyzak, H. Walk, *A distribution-free theory of nonparametric regression*, New-York: Springer-Verlag (2002).

[14] N. Ling, Y. Wu, *Consistency of modified kernel regression estimation for functional data*, Statistics, 46(2) (2012), pp. 149–158

[15] E. Masry, *Nonparametric regression estimation for dependent functional data: asymptotic normality*, Stoch. Proc. and their Appl., 115 (2005), pp. 155–177.

[16] A. Quintela-Del-Rio, *Recursive kernel hazard estimation of strong mixing data*, Far East J. Theor. Stat. 30(2) (2010), pp. 89–105.

[17] J.O. Ramsay, C.J. Dalzell, *Some tools for functional data analysis (with discussion)*, J. Roy. Statist. Soc., Ser. B, (1991), pp. 539–572.

[18] J.O. Ramsay, B.W. Silverman, *Applied functional data analysis.* Springer-Verlag, New-York (2002).

[19] J.O. Ramsay, B.W. Silverman, *Functional data analysis (2^{nd} Ed.)* Springer-Verlag, New-York (2006).

[20] E. Rio *Théorie asymptotique des processus aléatoires faiblement dépendants*, Springer 2000.

[21] J. Vilar, J. Vilar, *Recursive local polynomial regression under dependence conditions*, TEST 9 (1), pp. 209–232.

[22] L. Wang, L., H. Y. Liang, *Strong uniform convergence of the recursive regression estimator under φ-mixing conditions*, Metrika 59 (3) (2004), pp. 245–261.