



HAL
open science

Rates of almost sure convergence of families of recursive kernel estimators

Aboubacar Amiri

► **To cite this version:**

Aboubacar Amiri. Rates of almost sure convergence of families of recursive kernel estimators. Annales de l'ISUP, 2010, LIV (3), pp.3-24. <hal-03635287>

HAL Id: hal-03635287

<https://hal.science/hal-03635287v1>

Submitted on 8 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Pub. Inst. Stat. Univ. Paris
LIV, fasc. 3, 2010, 3 à 24

Rates of almost sure convergence of families of recursive kernel estimators

Aboubacar Amiri

Université d'Avignon et des Pays de Vaucluse, Laboratoire d'Analyse Non Linéaire et Géométrie (EA 2151), F-84018 Avignon

Abstract

The purpose of this paper is to establish, by an invariance principle approach, almost-sure convergence of parametric families of recursive kernel density and regression estimators. For these families, rates of convergence are obtained with exact upper bounds.

Key words: Pointwise convergence, Recursive kernel estimators, Density function, Regression function, Rates of convergence.

Index to information contained in this paper

1. Introduction
2. Density estimation
 - 2.1. Notation and assumptions
 - 2.2. Main results
3. Regression estimation
 - 3.1. Notation and assumptions
 - 3.2. Main results
4. Proofs
 - 4.1. Proof of Theorem 2.1
 - 4.2. Proof of Corollary 2.3
 - 4.3. Proof of Theorem 3.1
5. Examples
6. Appendix
7. Acknowledgement

1. Introduction

In this paper, we consider nonparametric sequential estimation for independent and identically distributed (i.i.d.) observations. Nonparametric functional estimation is a central field in statistics with emphasis usually being on density and regression estimators. The literature on this topic is still growing : for information on various methods and results in density estimation, one may refer to the books of Prakasa Rao (1983), Devroye and Györfi (1985), Silverman (1986). Also, a survey of nonparametric regression estimators appears in Collomb (1981, 1985) while more recent results are collected in e.g. the book of Györfi et al (2002). Sequential estimation is achieved with the use of recursive estimators, typically kernel ones. Here 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. These estimators have many advantages : they are of easy implementation and interpretation, fast to compute and they do not require extensive storage of data. In particular cases, they appear also as more efficient than classical estimators. Since for large database, the total re-computation of an estimator can be time-consuming (even for powerful computers), recursive estimators are more adapted to continuously updated data. A typical example could be weather forecasting where the sample size is not fixed in advance.

First kernel recursive estimators have been introduced by Wolverton, Wagner (1969) and Yamato (1972). Next Davis (1973), Deheuvels (1974), Wegman, Davis (1979), Wertz (1985) and Roussas (1992) have independently studied the rates of the almost sure convergence of particular recursive density estimates.

Concerning the regression, popular kernel recursive estimators are those of Devroye and Wagner (1980) and Ahmad-Lin (1976), commonly called recursive and semi-recursive estimators. These two estimators have also been studied by Györfi et al. (2002) in the iid case. The L^1 -convergence of the recursive estimator have been proved by Devroye and Wagner (1980), and Krzyżak (1992) established its rate of L^1 -convergence. Ahmad and Lin (1976) established the uniform consistency (with no rates) of their regression estimator on a compact set. Roussas (1992) studies the exact rates of almost sure convergence of recursive regression estimators.

For both density and regression problems, we derive in this paper exact upper bounds for the almost sure convergence of general families of kernel estimators, recently derived from Amiri (2009). These families include the popular recursive estimators discussed above.

In the next section we present our main assumptions and the results for density estimation while regression is studied in section 3. Proofs are

postponed to the section 4 and examples of application are given at the section 5 to validate the performance of recursive methods.

2. Density estimation

2.1. Notation and assumptions

Let $\{X_t, t \in \mathbb{N}\}$ be a sequence of iid random \mathbb{R}^d -vectors ($d \geq 1$) defined on a probability space (Ω, \mathcal{A}, P) and admitting a density f . Following Amiri (2009), the density f can be estimated by means of the general family of recursive kernel estimators

$$f_n^\ell(x) := \frac{1}{\sum_{i=1}^n h_i^{d(1-\ell)}} \sum_{i=1}^n \frac{1}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right), \quad x \in \mathbb{R}^d \quad (\ell \in [0, 1]),$$

which can be computed recursively by the formula

$$f_n^\ell(x) = \frac{\sum_{i=1}^{n-1} h_i^{d(1-\ell)}}{\sum_{i=1}^n h_i^{d(1-\ell)}} f_{n-1}^\ell(x) + K_n^\ell(x - X_n),$$

with

$$K_i^\ell(\cdot) := \frac{1}{h_i^{d\ell} \sum_{j=1}^i h_j^{d(1-\ell)}} K\left(\frac{\cdot}{h_i}\right), \quad (2.1)$$

where h_n is a sequence of positive constants (bandwidth) and K is the kernel. For $d = 1$, the family $(f_n^\ell(x))$ can be considered as a subfamily of the general family

$$f_n^H(x) := \left(\sum_{i=1}^n h_i H(h_i)\right)^{-1} \sum_{i=1}^n H(h_i) K\left(\frac{x - X_i}{h_i}\right), \quad (2.2)$$

introduced by Deheuvels (1974). To simplify the notation, let us set

$$B_{n,r} := \frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n}\right)^r, \quad \text{for all } r \in \mathbb{R}.$$

By this way, our family of estimators can be rewritten as

$$f_n^\ell(x) = \frac{1}{n B_{n,d(1-\ell)} h_n^{d(1-\ell)}} \sum_{i=1}^n \frac{1}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right). \quad (2.3)$$

To study its almost sure behavior, we consider kernels K satisfying the following conditions.

Assumption \mathcal{K} .

- (i) $K : \mathbb{R}^d \mapsto \mathbb{R}$ is a positive, bounded, symmetric density of probability ;
- (ii) For all $x \in \mathbb{R}^d, \lim_{\|x\| \rightarrow +\infty} \|x\|^d K(x) = 0$;
- (iii) $\int_{\mathbb{R}^d} |v_i v_j| K(v) dv < \infty, \quad i, j = 1, \dots, d.$

Also, we need the following conditions checked by the bandwidth.

Assumption \mathcal{H} .

- (i) $h_n \downarrow 0$ and $nh_n^{d+2} \rightarrow \infty$ as $n \rightarrow \infty$;
- (ii) $B_{n,r} \rightarrow \beta_r < \infty$, as $n \rightarrow \infty$ for all $r \in]-\infty, d+2]$.

Finally the density f is supposed to satisfy the following assumption.

Assumption \mathcal{P} . $f \in C_d^2(b)$, where $C_d^2(b)$ denotes the set of twice continuously differentiable functions $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ such that $\psi^{(2)}$ exists and bounded with $\psi^{(2)}$ standing for any partial derivative of order two of ψ .

The assumptions \mathcal{K} and \mathcal{P} are classical in nonparametric estimation field. The condition $\mathcal{H}(ii)$ is particular to the recursive problem and plays a crucial role in our calculus. The choice $h_n = c \left(\frac{\ln \ln n}{n}\right)^{\frac{1}{d+4}}$, with $c > 0$ where we have noted $\ln \ln(\cdot) := \ln^+ \ln^+(\cdot)$, with $\ln^+(\cdot) := \ln \max(\cdot, e)$, is a typical example of bandwidth satisfying \mathcal{H} . Indeed, h_n is decreasing and one may show that if $0 \leq r \leq d+2$ (resp. if $r < 0$), $B_{n,r}$ is a increasing (resp. decreasing) sequence and superiorly (resp. inferiorly) bounded by a converging sequence. By this way $B_{n,r}$ converges.

2.2. Main results

Now we can establish the almost sure convergence of density estimators.

Theorem 2.1 *Let \mathcal{K} and \mathcal{H} be satisfied.*

- (a) *In addition if for all $\alpha \geq 0$*

$$\frac{nh_n^d}{(\ln \ln n)^{2(\alpha+1)} \ln n} \rightarrow \infty, \text{ as } n \rightarrow \infty \text{ and } \lim_{n \rightarrow \infty} \frac{\ln h_n}{\ln n} < \infty, \quad (2.4)$$

then for all x such that $f(x) > 0$,

$$\overline{\lim}_{n \rightarrow \infty} \sqrt{\frac{nh_n^d}{\ln \ln n}} (f_n^\ell(x) - Ef_n^\ell(x)) = \sqrt{\frac{2\beta_{d(1-2\ell)}}{\beta_{d(1-\ell)}^2}} f(x) \|K\|_2 \text{ a.s.} \quad (2.5)$$

(b) Moreover, if \mathcal{P} holds, the choice $h_n = c \left(\frac{\ln \ln n}{n}\right)^{\frac{1}{d+4}}$, with $c > 0$, implies that for all x such that $f(x) > 0$,

$$\overline{\lim}_{n \rightarrow \infty} \left(\frac{n}{\ln \ln n}\right)^{\frac{2}{d+4}} (f_n^\ell(x) - f(x)) = \sqrt{\frac{2f(x)\|K\|_2^2 c^d \beta_{d(1-2\ell)}}{\beta_{d(1-\ell)}^2} + \frac{c^2 b_2(x) \beta_{d(1-\ell)+2}}{\beta_{d(1-\ell)}}} \quad (2.6)$$

a.s., where $b_2(x) := \frac{1}{2} \left(\sum_{1 \leq i, j \leq d} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \int_{\mathbb{R}^d} v_i v_j K(v) dv \right)$.

In the univariate case ($d = 1$), for $\ell = 1$, we find from (2.5) a result comparable to that stated by Wertz (1985) and for $\ell = \frac{1}{2}$, we find the following result proved by Wegman and Davis (1979) :

$$\overline{\lim}_{n \rightarrow \infty} \sqrt{\frac{nh_n}{\ln \ln n}} (f_n^{DW}(x) - \mathbb{E}f_n^{DW}(x)) = \sqrt{2\beta_{1/2} f(x) \|K\|_2} \text{ a.s.},$$

as $f_n^{DW}(x) = B_{n, \frac{1}{2}} f_n^{1/2}(x)$. Note that $f_n^{DW}(x)$ is an asymptotically biased estimator of $f(x)$ as soon as $\beta_{\frac{1}{2}} \neq 1$. Also our result (2.5) generalizes to d -dimensional random variables the one of Deheuvels (1974) for the general family (2.2) in the univariate case, but here the proof of (2.5) is based on the Theorem 6.2 (see the Appendix) rather the Prohorov lemma. Indeed by adapting the conditions of Deheuvels (1974) to the family (2.3), his result may be rewritten, under some conditions of regularity on K , and for h_n satisfying the conditions

$$\sum_{n=1}^{\infty} h_n^{1-\ell} = \infty, \sum_{n=1}^{\infty} h_n^{1-2\ell} = \infty \text{ and } h_n^{-\ell} = o\left(\frac{\sum_{i=1}^n h_i^{1-2\ell}}{\ln \ln \sum_{i=1}^n h_i^{1-2\ell}}\right)^{\frac{1}{2}}, \quad (2.7)$$

as

$$\overline{\lim}_{n \rightarrow \infty} \left(\text{resp. } \underline{\lim}_{n \rightarrow \infty} \right) \frac{\sum_{i=1}^n h_i^{1-\ell} (f_n^\ell(x) - \mathbb{E}f_n^\ell(x))}{\sqrt{2\|K\|_2^2 \sum_{i=1}^n h_i^{1-2\ell} \ln \ln \sum_{i=1}^n h_i^{1-\ell}}} = \pm \sqrt{f(x)}.$$

It can be shown that (2.7) is easily deduced from (2.4) and the assumption \mathcal{H} .

We close this section by studying the efficiency of our estimators in term of an almost sure converging criterion defined below.

Definition 2.2 Let $\ell_1, \ell_2 \in [0, 1]$. We say that $f_n^{\ell_1}(x)$ is preferable to $f_n^{\ell_2}(x)$ in term of the almost sure converging criterion if for some given K and h_n ,

$$0 \leq \frac{\overline{\lim}_{n \rightarrow \infty} |f_n^{\ell_1}(x) - f(x)|}{\overline{\lim}_{n \rightarrow \infty} |f_n^{\ell_2}(x) - f(x)|} < 1 \text{ a.s.}$$

So we can classify our estimators $f_n^\ell(x)$ according to ℓ in term of the almost sure converging criterion.

Corollary 2.3 *Under the assumptions \mathcal{K}, \mathcal{H} and \mathcal{P} , the choice*

$$h_n = C_n n^{-\nu}, 0 < \nu < \frac{1}{d+2}, C_n \downarrow c > 0$$

implies that for $\ell_1, \ell_2 \in [0, 1]$ $f_n^{\ell_1}$ is preferable to $f_n^{\ell_2}$ if $\ell_1 < \ell_2$.

3. Regression estimation

3.1. Notation and assumptions

Turning to the regression estimation, let $\{(X_t, Y_t), t \in \mathbb{N}\}$ be iid random variables taking values in $\mathbb{R}^d \times \mathbb{R}^{d'}$ ($d \geq 1, d' \geq 1$), and having probability density function f^* with respect to the Lebesgue mesure. We define the regression function as

$$r(x) := \begin{cases} E(m(Y_0) | X_0 = x) = \frac{\int_{\mathbb{R}^{d'}} m(y) f^*(x, y) dy}{f(x)} := \frac{\varphi(x)}{f(x)}, & \text{if } f(x) > 0 \\ Em(Y_0), & \text{if } f(x) = 0, \end{cases}$$

where m is Borelian function of $\mathbb{R}^{d'}$ into \mathbb{R} such that $\omega \mapsto m^2(Y_t(\omega))$ is P -integrable. Note that m is chosen by the statistician leading to multiple choices of estimation. Typical examples of m are identity and polynomial functions to estimate respectively the usual regression and the conditional moments.

As for density estimation we consider the parametric family of kernel regression estimators defined by

$$r_n^\ell(x) := \frac{\varphi_n^\ell(x)}{f_n^\ell(x)} \text{ where } \varphi_n^\ell(x) := \frac{1}{\sum_{i=1}^n h_i^{d(1-\ell)}} \sum_{i=1}^n \frac{m(Y_i)}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right).$$

The estimator $r_n^\ell(x)$ can be computed recursively by

$$r_n^\ell(x) = \frac{\left(\sum_{i=1}^{n-1} h_i^{d(1-\ell)}\right) \varphi_{n-1}^\ell(x) + \left(\sum_{i=1}^n h_i^{d(1-\ell)}\right) m(Y_n) K_n^\ell(x - X_n)}{\left(\sum_{i=1}^{n-1} h_i^{d(1-\ell)}\right) f_{n-1}^\ell(x) + \left(\sum_{i=1}^n h_i^{d(1-\ell)}\right) K_n^\ell(x - X_n)},$$

where $K_i^\ell(\cdot)$ is defined in (2.1). Our family includes the popular kernel recursive estimators say recursive ($\ell = 0$) and semi-recursive ($\ell = 1$).

Next, we introduce an additional assumption to get the almost sure behavior for our family of regression estimators.

Assumption \mathcal{R} .

- (i) $\varphi \in C_d^2(b)$;
- (ii) K satisfies the condition $\int_{\mathbb{R}^d} \|u\|^3 K(u) < \infty$;
- (iii) The function $E(m^2(Y)|X_0 = \cdot) f(\cdot)$ is both continuous and bounded away from zero at x ;
- (iv) There exist $\lambda > 0, \theta > 0$ such that $E \exp(\lambda|m(Y_0)|^\theta) < \infty$.

The assumption \mathcal{R} is classical in regression estimation, see e.g. Bosq and Blanke (2007) for the nonrecursive case. Note, that the assumption $\mathcal{R}(iv)$ is stronger than the condition $E|m(Y)|^{2+\delta} < \infty$, for some $\delta > 0$, used by Roussas (1992), in the particular case $\ell = 1/2$, but it's allow us to generalize his result to all $\ell \in [0, 1]$. If the function m is bounded, the condition $\mathcal{R}(iv)$ can be relaxed.

3.2. Main result

The question addressed in this section is getting the exact rates for the almost sure convergence of our family of regression estimators.

Theorem 3.1 *Under the assumptions $\mathcal{K}, \mathcal{H}, \mathcal{P}$ and \mathcal{R} , and for all x such that $f(x) > 0$, the choice $h_n = c \left(\frac{\ln \ln n}{n}\right)^{\frac{1}{d+4}}$, with $c > 0$ implies that*

$$\overline{\lim}_{n \rightarrow \infty} \left(\frac{n}{\ln \ln n}\right)^{\frac{2}{d+4}} (r_n^\ell(x) - r(x)) = C_0(x, K, \varphi, f, \ell), \quad (3.1)$$

where: $C_0(x, K, \varphi, f, \ell) := \frac{c^{\frac{d}{2}} \sqrt{2\beta_{d(1-2\ell)} f(x) V(x)} \|K\|_2 + c^2 \beta_{d(1-\ell)+2} b_{\varphi, f}(x)}{\beta_{d(1-\ell)} f(x)},$

with: $b_{\varphi, f}(x) := \frac{1}{2} \sum_{i,j=1}^d \left\{ \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) - r(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right\} \int_{\mathbb{R}^d} |z_i z_j| K(z) dz,$

and $V(x)$ is the conditional variance defined by

$$V(x) := E(m^2(Y_0)|X_0 = x) - r^2(x).$$

This theorem generalize the result of Roussas (1992) derived for $r_n^{\frac{1}{2}}(x)$. Note that, for the nonrecursive estimator, the rate of convergence obtained by Bosq and Blanke (2007) is of type $\left(\frac{\ln n}{n}\right)^{\frac{2}{d+4}}$, and without exact upper bound.

Now let us give the proofs of the main results.

4. Proofs

4.1. Proof of Theorem 2.1.

(a) We apply Theorem 6.2 (cf. Appendix) to the variables Z_i defined as:

$$Z_i = \frac{1}{h_i^{d\ell}} \left(K \left(\frac{x - X_i}{h_i} \right) - \mathbb{E}K \left(\frac{x - X_i}{h_i} \right) \right).$$

Next, if we set $V_n = \sum_{i=1}^n \mathbb{E}Z_i^2$ one may write

$$V_n = nB_{n,d(1-\ell)}^2 h_n^{d(1-2\ell)} n h_n^d \text{Var} f_n^\ell(x).$$

Let $S_n = \sum_{i=1}^n Z_i$ and S be a random function defined on $[0, +\infty[$ obtained by setting for $t \in [V_n, V_{n+1}[$, $S(t) = S_n$. From the relation (2.3), we can write

$$f_n^\ell(x) - \mathbb{E}f_n^\ell(x) = \frac{S_n}{nB_{n,d(1-\ell)} h_n^{d(1-\ell)}}. \quad (4.1)$$

Following Amiri (2009), we have under the assumptions \mathcal{K} and \mathcal{H}

$$n h_n^d \text{Var} f_n^\ell(x) \rightarrow \frac{\beta_{d(1-2\ell)}}{\beta_{d(1-\ell)}^2} f(x) \int_{\mathbb{R}^d} K^2(u) du \text{ as } n \rightarrow \infty.$$

Then for all x such that $f(x) > 0$, we have:

$$V_n \sim n h_n^{d(1-2\ell)} \beta_{d(1-2\ell)} f(x) \int_{\mathbb{R}^d} K^2(u) du \rightarrow \infty \text{ as } n \rightarrow \infty. \quad (4.2)$$

Since (2.4) holds then:

$$\frac{n h_n^d}{\ln \left(n h_n^{d(1-2\ell)} \right) \left(\ln \ln \left(n h_n^{d(1-2\ell)} \right) \right)^{2(\alpha+1)}} \rightarrow \infty \text{ as } n \rightarrow \infty,$$

and there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\frac{n h_n^{d(1-2\ell)}}{\ln \left(n h_n^{d(1-2\ell)} \right) \left(\ln \ln \left(n h_n^{d(1-2\ell)} \right) \right)^{2(\alpha+1)}} > \|K\|_\infty^2 h_n^{-2d} \geq Z_n^2.$$

By this way, the event $\left\{ Z_n^2 > \frac{n h_n^{d(1-2\ell)}}{\ln \left(n h_n^{d(1-2\ell)} \right) \left(\ln \ln \left(n h_n^{d(1-2\ell)} \right) \right)^{2(\alpha+1)}} \right\}$ is an impossible event for $n \geq n_0$. It follows by (4.2), that

$$\sum_{n=1}^{\infty} \frac{(\ln \ln V_n)^\alpha}{V_n} \mathbb{E} \left(Z_n^2 \mathbf{1}_{\left\{ Z_n^2 > \frac{V_n}{\ln V_n (\ln \ln V_n)^{2(\alpha+1)}} \right\}} \right) < \infty.$$

Applying Theorem 6.2 and with the help of (4.2), there exists a Brownian motion W such that,

$$\text{for } t \in [V_n, V_{n+1}[, \left| \frac{S(t) - W(t)}{(2t \ln \ln t)^{\frac{1}{2}}} \right| = o\left((\ln \ln t)^{-\frac{\alpha}{2}}\right).$$

Since the Brownian motion verifies the Law of Iterated Logarithm, then

$$\overline{\lim}_{t \rightarrow \infty} \frac{S(t)}{(2t \ln \ln t)^{\frac{1}{2}}} = 1 \text{ a.s.}$$

Hence, by (4.1), (4.2), the definition of $S(t)$ and since $t \in [V_n, V_{n+1}[$ and that $\frac{V_{n+1}}{V_n} \rightarrow 1$, we have $\overline{\lim}_{n \rightarrow \infty} \frac{B_{n,d(1-\ell)} n h_n^{d(1-\ell)} (f_n^\ell(x) - \mathbb{E} f_n^\ell(x))}{(2V_n \ln \ln V_n)^{\frac{1}{2}}} = 1$ a.s. But:

$$\frac{\left(n h_n^{d(1-2\ell)} \ln \ln \left(n h_n^{d(1-2\ell)} \right) \right)^{\frac{1}{2}} B_{n,d(1-\ell)}}{(2V_n \ln \ln V_n)^{\frac{1}{2}}} \rightarrow \frac{\beta_{d(1-\ell)}}{(2\beta_{d(1-2\ell)} f(x) \int_{\mathbb{R}^d} K^2(u) du)^{\frac{1}{2}}},$$

so that:

$$\overline{\lim}_{n \rightarrow \infty} \left(\frac{n h_n^d}{\ln \ln \left(n h_n^{d(1-2\ell)} \right)} \right)^{\frac{1}{2}} (f_n^\ell(x) - \mathbb{E} f_n^\ell(x)) = \sqrt{\frac{2\beta_{d(1-2\ell)}}{\beta_{d(1-\ell)}^2} f(x) \|K\|_2} \text{ a.s.} \quad (4.3)$$

We deduce (2.5) from (4.3) by employing the second part of (2.4) which implies that $\ln \ln \left(n h_n^{d(1-2\ell)} \right) = (\ln \ln n) (1 + o(1))$.

- (b) The choice $h_n = c \left(\frac{\ln \ln n}{n} \right)^{\frac{1}{d+4}}$ fulfils \mathcal{H} and (2.4), consequently the result follows from (2.5) and the result of the bias established in Amiri (2009). ■

4.2. Proof of Corollary 2.3

The choice: $h_n = C_n n^{-\nu}$, $0 < \nu < \frac{1}{d+2}$, $C_n \downarrow c > 0$ satisfies \mathcal{H} and (2.4). Using (2.5) and the result of the bias established in Amiri (2009), we find

$$\overline{\lim}_{n \rightarrow \infty} \sqrt{\frac{n^{1-\nu d}}{\ln \ln n}} (f_n^\ell(x) - f(x)) = \sqrt{\frac{2(1-\nu d(1-\ell))^2}{c^d(1-\nu d(1-2\ell))}} f(x) \|K\|_2 \text{ a.s.}$$

for all x such that $f(x) > 0$. Next, it's easy to show that the asymptotical constant is a strictly increasing function in term of ℓ . ■

4.3. Proof Theorem 3.1

We introduce the truncated version of $\varphi_n^\ell(x)$, defined by

$$\tilde{\varphi}_n^\ell(x) := \frac{1}{\sum_{i=1}^n h_i^{(1-\ell)d}} \sum_{i=1}^n \frac{1}{h_i^{d\ell}} m(Y_i) \mathbf{1}_{\{|m(Y_i)| \leq b_n\}} K\left(\frac{x - X_i}{h_i}\right),$$

where b_n is a real sequence converging to $+\infty$ as $n \rightarrow \infty$. Next, we proceed similarly than Bosq and Blanke (2007), for the nonrecursive case by writing

$$r_n^\ell(x) - r(x) = \frac{\tilde{\varphi}_n^\ell(x) - r(x)f_n^\ell(x)}{f_n^\ell(x)} + \frac{\varphi_n^\ell(x) - \tilde{\varphi}_n^\ell(x)}{f_n^\ell(x)} \quad (4.4)$$

and consider the following decomposition

$$\begin{aligned} \tilde{\varphi}_n^\ell(x) - r(x)f_n^\ell(x) &= (\tilde{\varphi}_n^\ell(x) - r(x)f_n^\ell(x) - \mathbb{E}(\tilde{\varphi}_n^\ell(x) - r(x)f_n^\ell(x))) \\ &\quad + (\mathbb{E}(\tilde{\varphi}_n^\ell(x) - r(x)f_n^\ell(x))) \\ &=: I_1^* + I_2^*. \end{aligned} \quad (4.5)$$

Let us set $Z_i^* = W_{n,i} - \mathbb{E}W_{n,i}$ with

$$W_{n,i} = \frac{1}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right) \mathbf{1}_{\{|m(Y_i)| \leq b_n\}} (m(Y_i) - r(x))$$

and $S_n^* = \sum_{i=1}^n Z_i^*$, one has $I_1^* = \frac{S_n^*}{nB_{n,d(1-\ell)}h_n^{d(1-\ell)}}$. If we set $V_n^* = \sum_{i=1}^n \mathbb{E}Z_i^{*2}$, then

$$\begin{aligned} \frac{V_n^*}{nh_n^{d(1-2\ell)}B_{n,d(1-2\ell)}} &= \frac{nh_n^d}{\left[\sum_{i=1}^n h_i^{d(1-\ell)}\right]^2} \sum_{k=1}^n \left\{ h_k^{-2d\ell} \mathbb{E}K^2\left(\frac{x-X_0}{h_k}\right) [m(Y_0) - r(x)]^2 \right. \\ &\quad \left. + h_k^{-2d\ell} \mathbb{E}K^2\left(\frac{x-X_0}{h_k}\right) m(Y_0) [2r(x) - m(Y_0)] \mathbf{1}_{\{|m(Y_0)| > b_n\}} \right\} \\ &\leq \frac{nh_n^d}{\left[\sum_{i=1}^n h_i^{d(1-\ell)}\right]^2} \sum_{k=1}^n \left[h_k^{d(1-2\ell)} \int_{\mathbb{R}^d} \frac{1}{h_k^d} K^2\left(\frac{x-u}{h_k}\right) V(u) f(u) du \right] \\ &\quad + \frac{\|K\|_\infty^2 \left\{ \mathbb{E}m^2(Y_0) [2r(x) - m(Y_0)]^2 P(|m(Y_0)| > b_n) \right\}^{\frac{1}{2}} nh_n^d \sum_{k=1}^n h_k^{-2d\ell}}{\left[\sum_{i=1}^n h_i^{d(1-\ell)}\right]^2} \\ &\leq \frac{nh_n^d}{B_{n,d(1-\ell)}^2 n^2 h_n^d} \sum_{k=1}^n \left[\left(\frac{h_k}{h_n}\right)^{d(1-2\ell)} \int_{\mathbb{R}^d} \frac{1}{h_k^d} K^2\left(\frac{x-u}{h_k}\right) V(u) f(u) du \right] \\ &\quad + \frac{\|K\|_\infty^2 \left\{ \mathbb{E}m^2(Y_0) [2r(x) - m(Y_0)]^2 P(|m(Y_0)| > b_n) \right\}^{\frac{1}{2}} B_{n,-2d\ell}}{h_n^d B_{n,d(1-\ell)}^2} \\ &=: D_1 + D_2. \end{aligned}$$

The assumptions $\mathcal{R}(ii)$, (iii) , the dominate convergence theorem and Bochner lemma imply that $\int_{\mathbb{R}^d} \frac{1}{h_k^d} K^2\left(\frac{x-u}{h_k}\right) V(u) f(u) du \rightarrow f(x)V(x) \|K\|_2^2$, as $k \rightarrow \infty$. Then the assumption $\mathcal{H}(ii)$ and Toeplitz lemma (see Appendix) imply

that $D_1 \rightarrow f(x)V(x) \|K\|_2^2$, as $n \rightarrow \infty$. Now if $b_n = (\delta \ln n)^{\frac{1}{\theta}}$ avec $\delta > \frac{2}{\lambda}$, by using the Markov inequality we have

$$D_2 = O \left[\frac{\exp \left(-\frac{\lambda b_n^\theta}{2} \right) (\ln n)^{\frac{2}{\theta}} B_{n,-2d\ell}}{h_n^d B_{n,d(1-\ell)}^2} \right] = O \left[\frac{n^{\frac{d}{d+4} - \frac{\lambda\delta}{2}}}{(\ln n)^{\frac{d}{d+4} - \frac{2}{\theta}}} \right] \rightarrow 0,$$

as $n \rightarrow \infty$. Therefore $V_n^* \sim nh_n^{d(1-2\ell)} \beta_{d(1-2\ell)} f(x)V(x) \int_{\mathbb{R}^d} K^2(u)du$, as $n \rightarrow \infty$, for all x such that $f(x) > 0$. Next, one may show that for all $\alpha \geq 0$, the condition

$$\frac{nh_n^d}{(\ln n)^{1+\frac{2}{\theta}} (\ln \ln n)^{2(\alpha+1)}} \rightarrow \infty, \text{ as } n \rightarrow \infty, \text{ and } \lim_{n \rightarrow \infty} \frac{\ln h_n}{\ln n} < \infty \quad (4.6)$$

is easily checked by the choice of bandwidth $h_n = c \left(\frac{\ln \ln n}{n} \right)^{\frac{1}{d+4}}$, with $c > 0$. It follows that (4.6) implies that

$$\frac{nh_n^d (\ln n)^{-\frac{2}{\theta}}}{\ln \left(nh_n^{d(1-2\ell)} \right) \left(\ln \ln \left(nh_n^{d(1-2\ell)} \right) \right)^{2(\alpha+1)}} \rightarrow \infty \text{ as } n \rightarrow \infty,$$

and then the choice $b_n = (\delta \ln n)^{\frac{1}{\theta}}$ implies that there exists $n_0 \in \mathbb{N}$ such that

$$\frac{nh_n^{d(1-2\ell)}}{\ln \left(nh_n^{d(1-2\ell)} \right) \left(\ln \ln \left(nh_n^{d(1-2\ell)} \right) \right)^{2(\alpha+1)}} > \|K\|_\infty^2 h_n^{-2d\ell} (\delta \ln n)^{\frac{2}{\theta}} \geq Z_n^{*2},$$

for all $n \geq n_0$. Now, proceeding as in the proof of the Theorem 2.1(a), one applies the Theorem 6.2 and finds $\frac{S_n^*}{(2V_n^* \ln \ln V_n^*)^{\frac{1}{2}}} \rightarrow 1$, as $n \rightarrow \infty$ a.s. By means of the condition (4.6) and the fact that

$$\frac{\left(nh_n^{d(1-2\ell)} \ln \ln \left(nh_n^{d(1-2\ell)} \right) \right)^{\frac{1}{2}} B_{n,d(1-\ell)}}{(2V_n^*(x) \ln \ln V_n^*(x))^{\frac{1}{2}}} \rightarrow \frac{\beta_{d(1-\ell)}}{(2\beta_{d(1-2\ell)} f(x)V(x) \int_{\mathbb{R}^d} K^2(u)du)^{\frac{1}{2}}},$$

we have

$$\overline{\lim}_{n \rightarrow \infty} \left(\frac{nh_n^d}{\ln \ln n} \right)^{\frac{1}{2}} I_1^* = \frac{(2\beta_{d(1-2\ell)} f(x)V(x) \int_{\mathbb{R}^d} K^2(u)du)^{\frac{1}{2}}}{\beta_{d(1-\ell)}} \text{ a.s.} \quad (4.7)$$

For the term I_2^* of (4.5) we proceed as follow.

$$\begin{aligned} I_2^* &= \frac{1}{\sum_{i=1}^n h_i^{d(1-\ell)}} \sum_{i=1}^n h_i^{-d\ell} \int_{\mathbb{R}^d} K \left(\frac{x-u}{h_i} \right) (r(u) - r(x)) f(u) du \\ &= \frac{1}{\sum_{i=1}^n h_i^{d(1-\ell)}} \sum_{k=i}^n h_i^{d(1-\ell)} \int_{\mathbb{R}^d} K(z) (r(x - h_i z) - r(x)) f(x - h_i z) dz. \end{aligned}$$

The assumptions $\mathcal{P}, \mathcal{R}(i), (ii)$ and Taylor formula give

$$I_2^* = \frac{h_n^2}{2nB_{n,d(1-\ell)}} \sum_{k=1}^n \left(\frac{h_k}{h_n}\right)^{d(1-\ell)+2} \sum_{i,j=1}^d \int_{\mathbb{R}^d} \left\{ \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x - \tau h_k z) - r(x - \tau h_k z) \frac{\partial^2 f}{\partial x_i \partial x_j}(x - \tau h_k z) \right. \\ \left. \times z_i z_j K(z) dz, \text{ with } 0 < \tau < 1. \right.$$

We have by the assumptions $\mathcal{P}, \mathcal{R}(i), \mathcal{K}(ii), (iii), \mathcal{H}(i), (ii)$, Bochner and Toeplitz's lemmas:

$$h_n^{-2} I_2^* \rightarrow \frac{\beta_{d(1-\ell)+2}}{\beta_{d(1-\ell)}} b_{\varphi, f}(x), \quad n \rightarrow \infty. \quad (4.8)$$

Finally concerning the last term of (4.4), one has for all $\varepsilon > 0$,

$$P \left(|\varphi_n^\ell(x) - \tilde{\varphi}_n^\ell(x)| > \varepsilon \left(\frac{\ln \ln n}{n} \right)^{\frac{2}{d+4}} \right) \leq \\ P \left(\sum_{i=1}^n |m(Y_i)| h_i^{-d} \mathbf{1}_{\{|m(Y_i)| > b_n\}} K \left(\frac{x - X_i}{h_i} \right) > \varepsilon \left(\frac{\ln \ln n}{n} \right)^{\frac{2}{d+4}} \sum_{i=1}^n h_i^{d(1-\ell)} \right) \\ \leq P \left(\bigcup_{i=1}^n \{|m(Y_i)| > b_n\} \right) \leq \sum_{i=1}^n P(|m(Y_i)| > b_n) = nP(|m(Y_0)| > b_n) \\ \leq E e^{\lambda |m(Y_0)|^\theta} n^{1-\lambda\delta}.$$

By the condition $\mathcal{R}(iv)$ we have:

$$\sum_{n=1}^{\infty} P \left(|\varphi_n^\ell(x) - \tilde{\varphi}_n^\ell(x)| > \varepsilon \left(\frac{\ln \ln n}{n} \right)^{\frac{2}{d+4}} \right) < \infty,$$

if $\delta > \frac{2}{\lambda}$ and Borel-Cantelli lemma implies that

$$\left(\frac{n}{\ln \ln n} \right)^{\frac{2}{d+4}} |\varphi_n^\ell(x) - \tilde{\varphi}_n^\ell(x)| \rightarrow 0 \text{ a.s.} \quad (4.9)$$

So since $f_n^\ell(x) \rightarrow f(x)$ a.s, (3.1) is from (4.7), (4.8), (4.9). \blacksquare

5. Examples

In order to validate our theoretical results, let us study numerically the efficiency of our family of recursive estimators, for $d = 1$.

5.1. Density estimation

In this section we study, by simulations, the efficiency of $f_n^\ell(x)$. Let us consider a sequence (X_n) of $N(N \in \mathbb{N}^*)$ zero-mean gaussian random variables iid, with variance σ^2 . The goal is to estimate the density $f(x)$ in an interval $I = [a, b]$, on the basis of the k first observations, for each value of k from n to N ($n \leq N$). More precisely, we consider the subdivision ϖ of I defined by:

$$\varpi = \{x_0 = a, x_1, \dots, x_J = b\}, J \in \mathbb{N}^*,$$

and for $k = n, \dots, N$, we compute $f_k^\ell(x_j)$ ($\ell \in \{0, 0.25, 0.5, 0.75, 1\}$) and $f_k^{PR}(x_j)$, at each point x_j of ϖ , where f_k^{PR} is the Parzen-Rozenblatt estimator based on the observations X_1, \dots, X_k . For recursive estimators, we use the recursive formula (2.1). The efficiency of one estimator is evaluated by its time of computation, the Empirical Mean Square Error (EMSE), and the Empirical Maximum Absolute Error (EMAE), defined (for some estimator \widehat{f}_k of f) respectively by:

$$EMSE(\widehat{f}_k) := \frac{1}{J+1} \sum_{j=0}^J \left(\widehat{f}_k(x_j) - f(x_j) \right)^2,$$

and

$$EMAE(\widehat{f}_k) := \max_{0 \leq j \leq J} \left| \widehat{f}_k(t_j) - f(t_j) \right|.$$

For this end, let us recall the following theoretical results concerning the asymptotic mean square error of our estimators (c.f Amiri (2009)).

Theorem 5.1 *Under the assumptions \mathcal{K}, \mathcal{H} and \mathcal{P} , the choice $h_n = C_n n^{-\frac{1}{d+4}}$, $C_n \downarrow c > 0$, implies that:*

(a) for $\ell \in [0, 1]$,

$$n^{\frac{4}{d+4}} E(f_n^\ell(x) - f(x))^2 \rightarrow c^4 \left(\frac{4+d\ell}{2+d\ell} \right)^2 b_2^2(x) + \frac{(4+d\ell)^2 f(x) \|K\|_2^2}{2c^d(4+d)(2+d\ell)},$$

as $n \rightarrow \infty$, for any x such that $f(x) > 0$.

(b) If $f(x) > 0$, the efficiency of $(f_n^\ell(x))$ is increasing according to ℓ in terms of the Mean Square Error (MSE) criterion.

(c) For $d = 1$, if $f(x) > 0$, $f_n^{PR}(x)$ is better than all $f_n^\ell(x)$, $\ell \in [0, 1]$ in terms of the MSE criterion.

The MSE criterion is defined as follow.

Definition 5.2 f_n is preferable to g_n in terms of the MSE criterion if:

$$0 \leq \lim_{n \rightarrow \infty} \frac{E(f_n(x) - f(x))^2}{E(g_n(x) - f(x))^2} < 1.$$

The kernel K considered in our simulations is the gaussian kernel used by many authors. For the smooth parameter, we use the well known choice, defined by $h_k = c_k k^{-\frac{1}{5}}$, where c_k is the empirical standard deviation of the k first observations. This choice is motivated by the Theorem 5.1 and the corollary 2.3. Note however that this choice is optimal only for the MSE criterion.

These practical implementations are done on **R** software for $n = 200$, $N = 500$, $\sigma^2 = 5$, $I = [-10, 10]$, $J = 400$ and $x_j = x_{j-1} + 0.05$, for $j = 1, \dots, 400$.

k	200	250	300	350	400	450	500
$\ell = 0$	0.094	0.083	0.072	0.066	0.060	0.054	0.051
$\ell = 0.25$	0.090	0.079	0.069	0.063	0.057	0.051	0.049
$\ell = 0.5$	0.086	0.076	0.066	0.060	0.055	0.049	0.047
$\ell = 0.75$	0.083	0.074	0.064	0.059	0.053	0.048	0.046
$\ell = 1$	0.081	0.072	0.062	0.057	0.052	0.047	0.045
P-R	0.075	0.067	0.057	0.052	0.048	0.044	0.042

Table 1: $\text{EMSE} \times 10^3$ for density estimators.

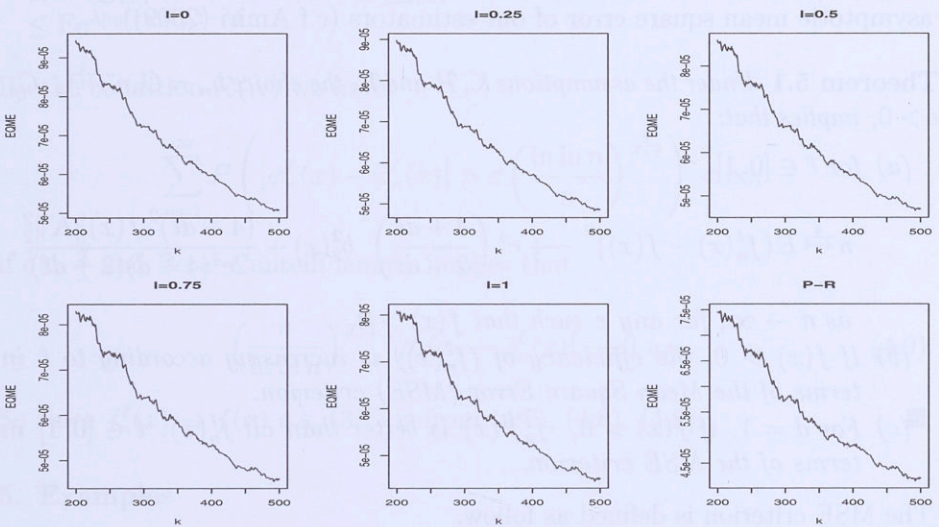


Figure 1: EMSE for density estimators.

Table 1 and Figure 1 display the empirical mean square errors of our recursive estimators for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and the one of Parzen-Rosenblatt (P-R) estimator. In this example, these mean square errors are very close for $k \geq 200$. Also, these results show that the efficiency of our recursive estimators is increasing according to ℓ in terms of the EMSE, and the Parzen-Rosenblatt estimator is preferable to recursive estimators, that confirms the theoretical results (b) and (c) of the Theorem 5.1.

k	200	250	300	350	400	450	500
$\ell = 0$	0.02402	0.02260	0.02161	0.02074	0.02014	0.01928	0.01875
$\ell = 0.25$	0.02357	0.02215	0.02115	0.02026	0.01971	0.01887	0.01843
$\ell = 0.5$	0.02322	0.02182	0.02078	0.01989	0.01936	0.01855	0.01819
$\ell = 0.75$	0.02296	0.02157	0.02052	0.01963	0.01908	0.01830	0.01803
$\ell = 1$	0.02332	0.02197	0.02069	0.01953	0.01890	0.01811	0.01791
P-R	0.02246	0.02153	0.02003	0.01923	0.01888	0.01823	0.01779

Table 2: EAME for density estimators.

Table 2 and Figure 2 display the empirical maximum absolute error of our recursive estimators for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and the one of Parzen-Rozenblatt estimator. In this example, we show that the six estimators can be suggested in order to estimate f because the difference between their EMAE is very small.

Finally, Table 3 represents the CPU-time in seconds of recursive estimators, for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and for the Parzen-Rozenblatt estimator. Here we note however that it takes more time to compute the P-R estimator than to compute the $f_n^\ell(x)$ ones using the recursive relation. Also, we see that the choice of ℓ has non influence. CPU-time is obtained with the following computer characteristics: CPU: Duo T9300; HD : 250GB and Memory : 3GB.

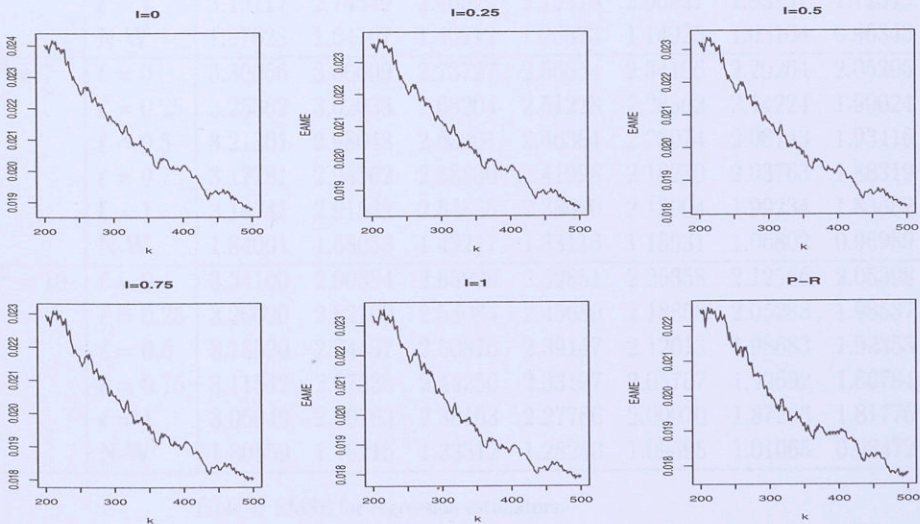


Figure 2: EAME for density estimators.

k	200	250	300	350	400	450	500
$\ell = 0$	0.05	0.14	0.19	0.27	0.39	0.46	0.49
$\ell = 0.25$	0.06	0.16	0.27	0.28	0.39	0.53	0.53
$\ell = 0.5$	0.03	0.14	0.18	0.24	0.33	0.36	0.44
$\ell = 0.75$	0.03	0.18	0.27	0.27	0.40	0.42	0.60
$\ell = 1$	0.06	0.16	0.17	0.24	0.31	0.37	0.38
P-R	0.08	3.24	6.60	10.17	14.34	19.36	24.82

Table 3: CPU-time for density estimators.

5.2. Regression estimation

As for density we suggest to study in this section the efficiency of the regression estimators. We consider a bivariate sequence $Z_t = (X_t, Y_t, 1 \leq t \leq N)$, $N \in \mathbb{N}^*$ of iid random vectors such that $r(x) = E(Y_t | X_t = x)$. X_t is a sequence of zero-mean gaussian random variables iid, with variance σ^2 and (Y_t) is defined by $Y_t = r(X_t) + \varepsilon_t$ where $\varepsilon \sim (0, \kappa^2)$ for $t = 1, \dots, N$. The goal is to estimate the regression function $r(x)$ in an interval $I = [a, b]$, on the basis of the k first observations, for each value of k from n to N . As for density estimation, we consider a subdivision of I and use the same criteria (EMSE, EAME and CPU time) to evaluate the efficiency of estimators.

The following results are obtained with $n = 200$, $N = 500$, $\sigma^2 = 5$, $I = [-4, 4]$, $J = 160$, $r(x) = x^2 + x + 1$ and $x_j = x_{j-1} + 0.05$, $1 \leq j \leq 160$.



Figure 1: Regression estimates.



Figure 2: Absolute error of regression estimates.

	k	200	250	300	350	400	450	500
$\kappa^2 = 0$	$\ell = 0$	3.00745	2.66555	2.38530	2.17022	2.01224	1.87914	1.75416
	$\ell = 0.25$	2.93818	2.59119	2.31165	2.10231	1.94431	1.81494	1.69424
	$\ell = 0.5$	2.87695	2.52915	2.24852	2.04038	1.88330	1.75497	1.63527
	$\ell = 0.75$	2.81949	2.47122	2.18985	1.98311	1.82709	1.69999	1.58144
	$\ell = 1$	2.76545	2.41708	2.13530	1.93014	1.77531	1.64960	1.53230
	N-W	1.53149	1.28392	1.09015	0.96028	0.87034	0.78670	0.71280
$\kappa^2 = 0.5$	$\ell = 0$	3.07512	2.70907	2.43840	2.21256	2.05583	1.91732	1.80315
	$\ell = 0.25$	3.01449	2.64549	2.37334	2.14725	1.99116	1.85300	1.74015
	$\ell = 0.5$	2.95793	2.58623	2.31290	2.08677	1.93150	1.79393	1.68253
	$\ell = 0.75$	2.90489	2.53085	2.25666	2.03070	1.87641	1.73967	1.62984
	$\ell = 1$	2.85512	2.47909	2.20435	1.97876	1.82561	1.68986	1.58168
	N-W	1.63359	1.36286	1.17478	1.02998	0.94075	0.85049	0.78340
$\kappa^2 = 1$	$\ell = 0$	3.19881	2.86802	2.57585	2.37147	2.19237	2.06544	1.95558
	$\ell = 0.25$	3.13751	2.80621	2.51128	2.30806	2.12816	2.00218	1.89367
	$\ell = 0.5$	3.08054	2.74881	2.45137	2.24943	2.06890	1.94404	1.83702
	$\ell = 0.75$	3.02715	2.69523	2.39558	2.19509	2.01409	1.89054	1.78513
	$\ell = 1$	2.97692	2.64509	2.34357	2.14468	1.96338	1.84131	1.73760
	N-W	1.73268	1.49316	1.28177	1.15493	1.04319	0.96071	0.89734
$\kappa^2 = 5$	$\ell = 0$	3.22892	2.96587	2.60917	2.34679	2.20721	2.07679	1.96930
	$\ell = 0.25$	3.36280	2.92713	2.59530	2.37707	2.19007	2.01139	1.89688
	$\ell = 0.5$	3.30156	2.86147	2.52715	2.31107	2.12456	1.94607	1.83352
	$\ell = 0.75$	3.24453	2.80042	2.46398	2.25027	2.06447	1.88643	1.77592
	$\ell = 1$	3.19117	2.74349	2.40529	2.19415	2.00927	1.83191	1.72349
	N-W	1.97628	1.64167	1.40112	1.26373	1.14925	1.03104	0.96343
$\kappa^2 = 7$	$\ell = 0$	3.30056	3.06609	2.73727	2.56654	2.34196	2.20264	2.05205
	$\ell = 0.25$	3.25362	3.02033	2.68204	2.51228	2.28362	2.14224	1.99024
	$\ell = 0.5$	3.21301	2.98043	2.63261	2.46364	2.23074	2.08743	1.93416
	$\ell = 0.75$	3.17781	2.94562	2.58830	2.41998	2.18270	2.03763	1.88319
	$\ell = 1$	3.14741	2.91532	2.54855	2.38080	2.13904	1.99234	1.83682
	N-W	1.84001	1.68053	1.43217	1.33116	1.15931	1.06802	0.96989
$\kappa^2 = 10$	$\ell = 0$	3.34100	2.90384	2.65928	2.52851	2.26358	2.12566	2.05398
	$\ell = 0.25$	3.26020	2.82100	2.58084	2.45680	2.18858	2.05288	1.98537
	$\ell = 0.5$	3.18520	2.74437	2.50876	2.39147	2.12013	1.98683	1.92353
	$\ell = 0.75$	3.11542	2.67338	2.44250	2.33197	2.05767	1.92692	1.86784
	$\ell = 1$	3.05049	2.60763	2.38163	2.27786	2.00070	1.87263	1.81776
	N-W	1.80559	1.48215	1.33512	1.28240	1.09695	1.01065	0.98472

Table 4: EMSE for regression estimators.

Table 4 displays the empirical mean square errors of our recursive regression estimators for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and the one of Nadaraya-Watson (N-W)

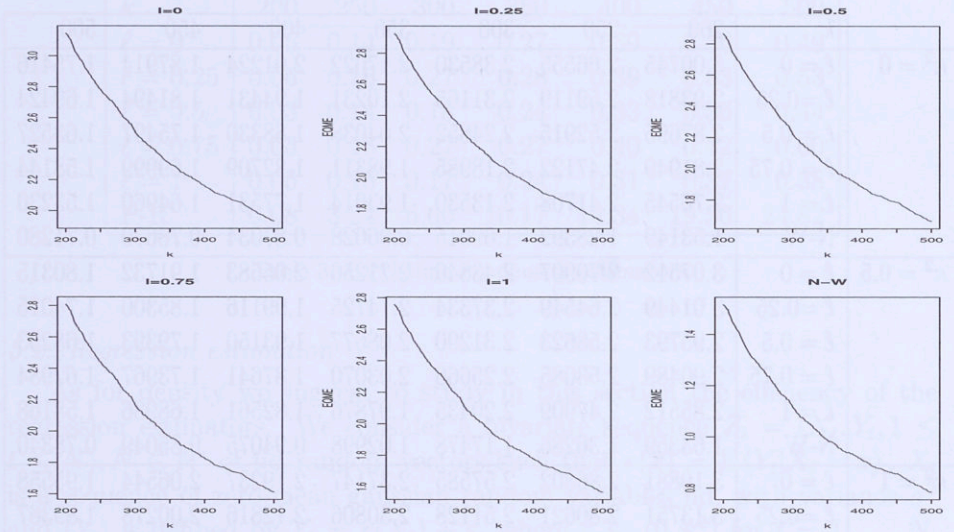


Figure 3: EMSE for regression estimators.

estimator, for $\kappa = 0, 0.5, 1, 5, 7, 10$. These results show that the efficiency of our recursive estimators is increasing according to ℓ in terms of the EMSE, and the Nadaraya-Watson estimator is preferable to recursive estimators. The results of Table 4 are represented at Figure 3 for $\kappa = 0.5$

Table 5 displays the empirical maximum absolute error of our recursive regression estimators for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and the one of Nadaraya-Watson estimator, for $\kappa = 0, 0.5, 1, 5, 7, 10$. In this example, we observe the same behavior of regression estimators as for density estimation.

Figure 4 represents the case $\kappa = 0.5$ of Table 5. Finally, Table 6 represents the CPU-time in seconds of recursive estimators, for $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and for Nadaraya-Watson estimator. We note in this example a big advantage in time of computation point view by replacing the classical kernel by recursive kernel.

	k	200	250	300	350	400	450	500
$\kappa^2 = 0$	$\ell = 0$	4.09858	3.83213	3.61061	3.42434	3.29892	3.17187	3.01985
	$\ell = 0.25$	4.04305	3.76962	3.54374	3.35715	3.23011	3.10135	2.94602
	$\ell = 0.5$	3.98929	3.71029	3.48023	3.29378	3.16580	3.03564	2.87728
	$\ell = 0.75$	3.93741	3.65455	3.42115	3.23412	3.10569	2.97428	2.81394
	$\ell = 1$	3.88759	3.60383	3.36560	3.17814	3.04993	2.91750	2.75508
	N-W	2.74989	2.48017	2.26247	2.09108	1.99919	1.87329	1.72352
$\kappa^2 = 0.5$	$\ell = 0$	4.72216	4.43001	4.19203	3.99124	3.85568	3.74682	3.61955
	$\ell = 0.25$	4.65306	4.35747	4.11601	3.91216	3.77838	3.66920	3.54210
	$\ell = 0.5$	4.59221	4.29144	4.04838	3.84166	3.70901	3.60002	3.47266
	$\ell = 0.75$	4.53663	4.23125	3.98790	3.77818	3.64648	3.53756	3.41000
	$\ell = 1$	4.48656	4.17682	3.93319	3.72249	3.58970	3.48110	3.35380
	N-W	3.26703	3.01214	2.82928	2.64388	2.53107	2.43394	2.34552
$\kappa^2 = 1$	$\ell = 0$	4.96471	4.70513	4.47659	4.29361	4.15159	4.02833	3.88661
	$\ell = 0.25$	4.76298	4.53215	4.27563	4.12990	3.95074	3.85290	3.73843
	$\ell = 0.5$	4.70631	4.47179	4.21082	4.06583	3.88465	3.78494	3.67194
	$\ell = 0.75$	4.65292	4.41511	4.15186	4.00705	3.82354	3.72222	3.61069
	$\ell = 1$	4.60316	4.36401	4.09914	3.95342	3.76712	3.66473	3.55527
	N-W	3.34443	3.13341	2.91629	2.79458	2.62940	2.55343	2.46962
$\kappa^2 = 5$	$\ell = 0$	4.97337	4.66381	4.39176	4.18760	4.02391	3.85424	3.70874
	$\ell = 0.25$	4.91732	4.59799	4.31603	4.11142	3.94858	3.77660	3.62706
	$\ell = 0.5$	4.86498	4.53534	4.24575	4.03990	3.87940	3.70429	3.55122
	$\ell = 0.75$	4.81712	4.47631	4.18150	3.97399	3.81542	3.63711	3.48132
	$\ell = 1$	4.77187	4.42031	4.12276	3.91333	3.75573	3.57469	3.41663
	N-W	3.68481	3.32573	3.05368	2.86946	2.73227	2.54701	2.40723
$\kappa^2 = 7$	$\ell = 0$	4.97394	4.74539	4.50261	4.37304	4.19653	4.08109	3.93051
	$\ell = 0.25$	4.91896	4.69523	4.44663	4.31789	4.13483	4.01538	3.86248
	$\ell = 0.5$	4.88078	4.64928	4.39485	4.26645	4.07865	3.95486	3.80200
	$\ell = 0.75$	4.84819	4.60888	4.34742	4.21935	4.02835	3.90040	3.74650
	$\ell = 1$	4.82042	4.57445	4.30448	4.17700	3.98201	3.85086	3.69624
	N-W	3.62564	3.39146	3.12889	3.02900	2.85088	2.73432	2.62157
$\kappa^2 = 10$	$\ell = 0$	5.02434	4.70932	4.51557	4.38363	4.17785	4.05350	3.94531
	$\ell = 0.25$	4.95799	4.63293	4.44064	4.31099	4.10012	3.97475	3.86640
	$\ell = 0.5$	4.89591	4.56150	4.37124	4.24448	4.02766	3.90290	3.79535
	$\ell = 0.75$	4.83725	4.49501	4.30819	4.18272	3.96062	3.83716	3.73052
	$\ell = 1$	4.78165	4.43303	4.24968	4.12604	3.89903	3.77691	3.67104
	N-W	3.60823	3.29731	3.13773	3.04311	2.85004	2.73057	2.64798

Table 5: EAME for regression estimators.

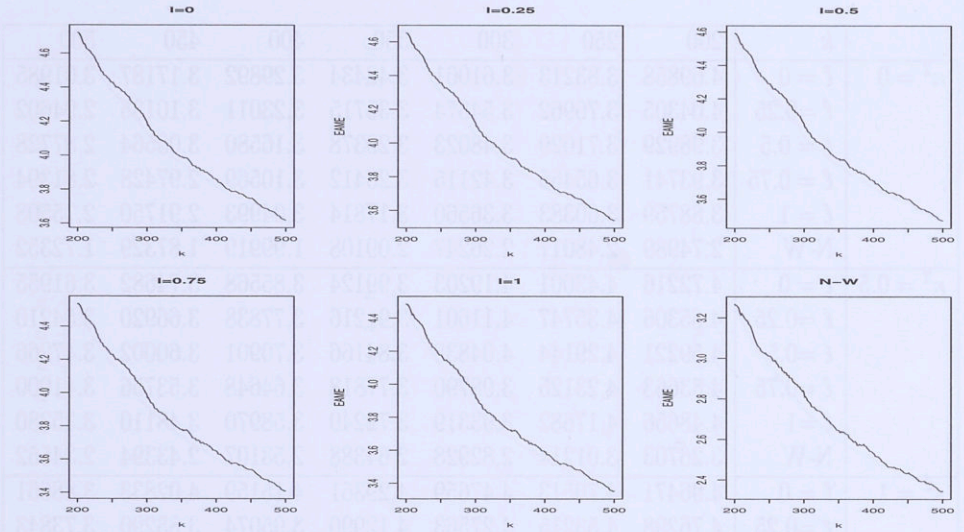


Figure 4: EAME for regression estimators.

k	200	250	300	350	400	450	500
$\ell = 0$	0.17	0.20	0.30	0.39	0.49	0.61	0.71
$\ell = 0.25$	0.10	0.22	0.32	0.43	0.55	0.70	0.85
$\ell = 0.5$	0.11	0.20	0.30	0.38	0.48	0.60	0.70
$\ell = 0.75$	0.11	0.22	0.33	0.42	0.56	0.69	0.81
$\ell = 1$	0.10	0.21	0.28	0.36	0.46	0.55	0.65
N-W	0.24	9.20	18.87	29.65	41.65	54.46	68.50

Table 6: CPU time for regression estimators.

6. Appendix

Let us begin this appendix by recalling the Toeplitz's lemma used in our proofs.

Lemma 6.1 (Toeplitz). Let $(a_{n,k})_{n \geq 1, k \geq 1}$ be a sequence of real number and let $(w_n)_{n \geq 1}$ a sequence which converge to w . Moreover we suppose that:

- (i) For all $k \geq 1$ $\lim_{n \rightarrow \infty} a_{n,k} = 0$;
- (ii) $\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} a_{n,k} = A < \infty$;
- (iii) For all $n \geq 1$, $\sum_{k=1}^{\infty} |a_{n,k}| < \infty$.

Then $\sum_{k=1}^{\infty} a_{n,k} w_k \rightarrow A w$ as $n \rightarrow \infty$.

Now we recall the following theorem, Theorem 3.2 in Jain et al. (1975), that is the main tool in establishing the main results.

Theorem 6.2 Let X_n be a sequence of random variables defined on a probability space (Ω, \mathcal{F}, P) such that for all $n \geq 0$, $E[X_n^2 | \sigma(X_1, \dots, X_{n-1})]$ exists and $E[X_n | \sigma(X_1, \dots, X_{n-1})] = 0$ a.s. Let $S_n = \sum_{i=1}^n X_i$, $S_0 = 0$ and $V_n = \sum_{i=1}^n E[X_i^2 | \sigma(X_1, \dots, X_{i-1})]$ if $n \geq 1$, $V_0 = 0$. For a fixed $\alpha \geq 0$, assume that $V_n \rightarrow \infty$ a.s. and

$$\sum_{k=1}^{\infty} \frac{(\ln \ln V_k)^\alpha}{V_k} E \left(X_k^2 \mathbf{1}_{\left\{ X_k^2 > \frac{V_k}{\ln V_k (\ln \ln V_k)^{2(\alpha+1)}} \right\}} \right) < \infty \text{ a.s.}$$

Let S be a random function defined on $[0, +\infty[$ obtained by setting for $t \in [V_n, V_{n+1}[$, $S(t) = S_n$. Then, redefining $\{S(t), t \geq 0\}$ if necessary on a new probability space, there exists a Brownian motion ξ such that:

$$|S(t) - \xi(t)| = o \left(t^{\frac{1}{2}} (\ln \ln t)^{\frac{1-\alpha}{2}} \right).$$

7. Acknowledgement

References

- [1] **Ahmad**, I.A. and **Lin**, P.E. (1976). Nonparametric sequential estimation of a multiple regression function. *Bull. Math. Stat.* 17, 63-75.
- [2] **Amiri**, A. (2009). Sur une famille paramétrique d'estimateurs séquentiels de la densité pour un processus fortement mélangeant. *C. R. Acad. Sci. Paris, Ser. I* 347, 309-314.
- [3] **Bosq**, D. (1998). Nonparametric statistics for stochastic processes. Estimation and prediction, Lecture notes in statistics 110 2nd ed. Springer-Verlag, New York.
- [4] **Bosq**, D. and **Blanke**, D. (2007). Inference and prediction in large dimensions. *Wiley Series in Prob. and Stat.* ISBN 978-0-470-08147-1.
- [5] **Collomb**, G. (1981). Estimation non-paramétrique de la régression: revue bibliographique. *Int. Stat. Rev.*, Vol. 49, 1, 75-93.
- [6] **Collomb**, G. (1985) Nonparametric regression: an up-to-date bibliography. *Statistics* 16, 2, 309-324.
- [7] **Davis**, H.L. (1973). Strong consistency of a sequential estimator of a probability density function, *Bull. Math. Stat.* 15, 49-54.
- [8] **Deheuvels**, P. (1974). Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité. *C. R. Acad. Sci. Paris* Vol. 278, 1217-1220.

- [9] **Devroye, L., Györfi, L.** (1985). Nonparametric density estimation: the L_1 view, Wiley, New-York, Russian translation: Mir, Moscow, 1988.
- [10] **Devroye, L. and Wagner, T.J.** (1980). On the L_1 convergence of kernel estimators of regression functions with application in discrimination. *Z. Wahrsch. verw. Gebiete* 51, 15-25.
- [11] **Györfi, L., Kholer, M., Krzyżak, A. and Walk, H.** (2002). A distribution-free theory of nonparametric regression, Springer-Verlag New York.
- [12] **Jain, C., Jogdeo, K. and Stout, W.** (1975). Upper and lower functions for martingales and mixing processes, *Ann. Prob.*, 3, 1, 119-145.
- [13] **Krzyżak, A.** (1992). Global convergence of the recursive kernel regression estimates with applications in classification and nonlinear system estimation, *IEEE Trans.*, 38, 4, 1323-1328.
- [14] **Prakasa Rao, B.L.S.** (1983), Nonparametric functional estimation, New-York: Academic Press.
- [15] **Roussas, G.G.** (1992). Exact rates of almost sure convergence of a recursive kernel estimate of a probability density function: Application to regression and hazard rate estimation, *J. Nonparam. Stat.* 1, 171-195.
- [16] **Silverman, B. W.** (1986). Density estimation for statistics and data analysis. Monographs on Stat. and Appl. Prob. Chapman and Hall.
- [17] **Wegman, E. and Davis, H.I.** (1979). Remarks on some recursive estimators of a probability density, *Ann. Stat.*, 7, 2, 316-327.
- [18] **Wertz, W.** (1985). Sequential and recursive estimators of the probability density, *Statistics* 16, 277-295.
- [19] **Wolverton, C. and Wagner, T.J.** (1969). Recursive estimates of probability densities. *IEEE Trans.* 5, 307-308.
- [20] **Yamato, H.** (1972). Sequential estimation of a continuous probability density function and mode. *Bull. Math. Statist. Jap.*, Vol. 14, 1-12.