Nonparametric kernel regression estimation for functional stationary ergodic data: Asymptotic properties

Naâmane Laib, Djamal Louani *

L.S.T.A., Paris 6 University, France

ARTICLE INFO

Article history:
Received 7 July 2008
Available online 16 June 2010

AMS 1991 subject classifications:
Primary 62G07
secondary 62M09

Keywords:
Asymptotic normality
Consistency
Ergodic processes
Functional dependent data
Martingale difference
Regression estimation

ABSTRACT

The aim of this paper is to study asymptotic properties of the kernel regression estimate whenever functional stationary ergodic data are considered. More precisely, in the ergodic data setting, we consider the regression of a real random variable $Y$ over an explanatory random variable $X$ taking values in some semi-metric abstract space. While estimating the regression function using the well-known Nadaraya–Watson estimator, we establish the consistency in probability, with a rate, as well as the asymptotic normality which induces a confidence interval for the regression function usable in practice since it does not depend on any unknown quantity. We also give the explicit form of the conditional bias term. Note that the ergodic framework is more convenient in practice since it does not need the verification of any condition as in the mixing case for example.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Various nonparametric estimators of the regression function have been proposed in the literature when the explanatory random variables $\{X_i\}$’s take their values in a finite dimensional space where the Lebesgue measure plays an important role. There is an extensive literature dealing with limit properties of these estimators and other related issues as the optimal bandwidth selection in both independent and dependent cases. For an overview, one may refer to [24,4,17] and the references therein.

Asymptotic issues for functional data have recently received an increasing interest, one may refer to [20,13,3,10,11,22,23,21,19,9,2,8,7] and to the recent monograph by Ferraty and Vieu [12] and the references therein.

To formulate the functional regression estimate problem, let $(X_i, Y_i)_{i \in \mathbb{N}}$ be a sequence of pairs of random elements where $Y_i$ is a real-valued random variable and $X_i$ takes its values in some semi-metric abstract space $(\mathcal{E}, d(\cdot, \cdot))$. This covers the case of semi-normed spaces of possibly infinite dimension (e.g., Hilbert or Banach spaces) with the norm $\| \cdot \|$ and the distance $d(x, y) = \|x - y\|$. Assume, for $k = 1, 2$, that $\mathbb{E}(|Y_1|^k) < \infty$ and that, for a fixed $x \in \mathcal{E}$, the conditional mean function $r(x) := \mathbb{E}(Y_1 | X_1 = x)$ and the conditional variance $W_2(x) := \mathbb{E}((Y_1 - r(x))^2 | X_1 = x)$ of $Y_1$ given $X_1 = x$ exist.

The Nadaraya–Watson type estimator of $r$ has been introduced by Ferraty and Vieu [10]. It is defined by

$$
\hat{r}_n(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{d(x, X_i)}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{h}\right)} := \hat{r}_{n, 2}(x) \hat{r}_{n, 1}(x).
$$

(1.1)

* Corresponding author.
E-mail addresses: naamane.laib@upmc.fr (N. Laib), djamal.louani@upmc.fr (D. Louani).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.05.010
when the denominator is not equal to 0. Here, $K$ is a real-valued kernel function, $h := h_n$ is the bandwidth parameter (which goes to 0 as $n$ goes to infinity) and

$$
\hat{r}_{nj}(x) = \frac{1}{nE(\Delta_1(x))} \sum_{i=1}^{n} Y_j^{-1} \Delta_i(x), \quad \text{for } j = 1, 2,
$$

(1.2)

where

$$
\Delta_i(x) = K \left( \frac{d(x, X_i)}{h} \right).
$$

Rates of almost sure uniform convergence, over a compact set, of the estimator $\hat{r}_n$ were established in [11] for mixing processes while Masry [19] obtained the mean squared convergence and the asymptotic normality. In the independent functional data case, some asymptotic results including the mean squared convergence, with rates, as well as the asymptotic normality have been obtained by Ferraty et al. [9].

To be more convenient towards a number of applications in practice, we consider in this paper, the regression function estimation when the data are functional and assumed to be sampled from a stationary and ergodic process to allow the maximum possible generality in regard to the dependence setting. Besides the infinite dimensional character of the data, we avoid here the widely used strong mixing condition and its variants to measure the dependency and the very involved probabilistic calculations that it implies. Moreover, the mixing properties of a number of well-known processes are still open questions. Indeed, several models are given in the literature where mixing properties are still to be verified or even fail to hold for the processes they induce. For instance, the AR(1)-GARCH(1,1) process still needs to check whether it satisfies any mixing condition. Examples in which the AR(1) linear real process with discrete valued random innovation is not strongly mixing are given by Chernick [5] and Andrews [1]. In particular, the process $X_t = \rho X_{t-1} + \epsilon_t$, where $\rho \in (0; 1/2]$ and $(\epsilon_t)_{t \in \mathbb{Z}}$ is a sequence of independent Bernoulli random variables, is not strongly mixing since the mixing coefficient $\alpha_n = 1/4$ for every $n \in \mathbb{N}$ (see, [11]).

For the sake of clarity, introduce some details defining the ergodic property of processes. Taking a measurable space $(\mathcal{S}, \mathcal{F})$, denote by $S^\mathbb{N}$ the space of all functions $s : \mathbb{N} \to \mathcal{S}$. If $s_j$ is the value the function $s$ takes at $j \in \mathbb{N}$, define $H_j$ as the $j$-th coordinate map, i.e., $H_j(s) = s_j$, and consider $H_j^{-1}$ to handle its inverse image. Set $\mathcal{F}^\mathbb{N}$ to be the smallest $\sigma$-algebra in $S^\mathbb{N}$ containing all $\sigma$-algebras $H_j^{-1}(\mathcal{F})$, $j \in \mathbb{N}$. A random process $Z = \{Z_j : j \in \mathbb{N} \}$ can be considered as a random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in $(S^\mathbb{N}, \mathcal{F}^\mathbb{N})$. Now a set $B \in \mathcal{F}$ is called invariant if there exists some set $A \in \mathcal{F}^\mathbb{N}$ such that $B = (\{Z_n, Z_{n+1}, \ldots : n \in A\}$ is true for any $n \geq 1$. The process $Z$ is then said ergodic whenever, for any invariant set $B$, we have $\mathbb{P}(B) = 0$ or $\mathbb{P}(\Omega \setminus B) = 0$. It is well known from the ergodic theorem that, for a stationary ergodic process $Z$, we have

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} Z_i = \mathbb{E}(Z_1), \quad \text{almost surely.}
$$

(1.3)

Therefore, the ergodic property in our setting is formulated on the basis of the statement (1.3) and the requirements are considered in conditions (A2) below. We refer to the book of Krenkel [16] for an account of details and results on the ergodic theory.

The main results of the paper state the weak consistency of $\hat{r}_n$, with a convergence rate, as well as its asymptotic normality that induces a confidence interval for the regression function $r$ which may be used in practice since all the constants appearing in the limit law are estimated. We give also the explicit form of the conditional bias term. To prove our results, our methodology is based on the martingale approximation which allows to provide an unified framework for the nonparametric time series analysis enabling one to launch a systematic study for dependent data.

2. Main results

2.1. Notations and hypotheses

In order to state our results, we introduce some notations. Let $\mathcal{F}_i$ be the $\sigma$-field generated by $((X_1, Y_1), \ldots, (X_i, Y_i))$ and $\mathcal{G}_i$ that generated by $((X_1, Y_1), \ldots, (X_i, Y_i), X_{i+1})$. Let $B(x, u)$ be a ball centered at $x \in \mathcal{E}$ with radius $u$. Let $D_i := d(x, X_i)$ so that $D_i$ is a nonnegative real-valued random variable. Working on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $F(U) = \mathbb{P}(D \leq u) := \mathbb{P}(X_i \in B(x, u))$ and $F_i^{-1}(u) = \mathbb{P}(D_i \leq u \mid \mathcal{F}_{i-1}) = \mathbb{P}(X_i \in B(x, u) \mid \mathcal{F}_{i-1})$ be the distribution function and the conditional distribution function, given the $\sigma$-field $\mathcal{F}_{i-1}$, of $(D_i)_{i \geq 1}$, respectively. Denote by $o_{\alpha}(u)$ a real random function $l$ such that $l(u)/u$ converges to zero almost surely as $u \to 0$. Similarly, define $O_{\alpha}^{\mathcal{F}}(u)$ as a real random function $l$ such that $l(u)/u$ is almost surely bounded. From now on, for $j = 1, 2$ and $x \in \mathcal{E}$, set

$$
\bar{r}_{nj}(x) = \frac{1}{nE(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E} \left[ Y_i^{-1} \Delta_i(x) \mid \mathcal{F}_{i-1} \right],
$$

where $\mathbb{E}(X \mid \mathcal{F})$ is the conditional expectation of the random variable $X$ given the $\sigma$-field $\mathcal{F}$.
Our results are stated under some assumptions we gather hereafter for easy reference.

(A1) $K$ is a nonnegative bounded kernel of class $C^1$ over its support $[0, 1]$ and $K(1) > 0$. The derivative $K'$ exists on $[0, 1]$ and satisfies the condition $K'(t) < 0$, $\forall t \in [0, 1]$ and $\int_0^1 (K'(u)) du < \infty$ for $j = 1, 2$.

(A2) For $x \in \mathcal{E}$, there exist a sequence of nonnegative bounded random functionals $(f_{i,j})_{j \geq 1}$, a sequence of random functions $(g_{i,x})_{j \geq 1}$, a deterministic nonnegative bounded functional $f_i$ and a nonnegative real function $\phi$ tending to zero, as its argument tends to 0, such that

(i) $F_i(u) = \phi(u)f_i(x) + o(\phi(u))$ as $u \to 0$.

(ii) For any $i \in \mathbb{N}$, $F_i^{j-1}(u) = \phi(u)f_{i,j}(x) + g_{i,x}(u)$ with $g_{i,x}(u) = o_{a.s.}(\phi(u))$ as $u \to 0$, $g_{i,x}(u)/\phi(u)$ almost surely bounded and $n^{-1}\sum_{i=1}^n f_{i,j}(x) = o_{a.s.}(\phi(u))$ as $n \to \infty$, $j = 1, 2$.

(iii) $n^{-1}\sum_{i=1}^n f_{i,j}(x) \to f_j(x)$, almost surely as $n \to \infty$, for $j = 1, 2$.

(iv) There exists a nondecreasing bounded function $\tau_0$ such that, uniformly in $u \in [0, 1]$,

$$\frac{\phi(hu)}{\phi(h)} = \tau_0(u) + o(1), \quad h \downarrow 0 \quad \text{and, for } 1 \leq j \leq 2 + \delta \text{ with } \delta > 0, \int_0^1 (K'(u)) \tau_0(u) du < \infty.$$  

(A3) (i) The conditional mean of $Y_i$ given the $\sigma$-field $\mathcal{G}_{i-1}$ depends only on $X_i$, i.e., for any $i \geq 1$, $E(Y_i \mid \mathcal{G}_{i-1}) = r(X_i)$ almost surely.

(ii) The conditional variance of $Y_i$ given the $\sigma$-field $\mathcal{G}_{i-1}$ depends only on $X_i$, i.e., for any $i \geq 1$, $E((Y_i - r(X_i))^2 \mid \mathcal{G}_{i-1}) = W_2(X_i)$ almost surely. Moreover, the function $W_2$ is continuous in a neighborhood of $x$, that is,

$$\sup_{u \in \mathcal{G}(x, u) \in \mathcal{E}} |W_2(u) - W_2(x)| = o(1) \quad \text{as } h \to 0.$$  

(A4) (i) $|r(u) - r(v)| \leq c_3d(u, v)^\beta$ for all $(u, v) \in \mathcal{E}^2$ and some $\beta > 0$ and a constant $c_3 > 0$.

(ii) For some $\delta > 0$, $E(|Y_i|^{2+\delta}) < \infty$ and the function $W_{2+\delta}(u) := E(|Y_i - r(X_i)|^{2+\delta} \mid X_i = u)$, $u \in \mathcal{E}$, is continuous in a neighborhood of $x$.

We give now some further notations. For $j \geq 1$, set

$$M_j = K^{j}(1) - \int_0^1 (K'(u)) \tau_0(u) du.$$  

Define the conditional bias of the regression estimate $\hat{r}_n(x)$ as

$$B_n(x) = C_n(x) - r(x),$$

where

$$C_n(x) := \frac{\hat{r}_{n,2}(x)}{\hat{r}_{n,1}(x)}.$$  

Discussions of hypotheses. Conditions used here share some similarities with that used in [9,8]. Condition (A1) is very usual in nonparametric functional estimation literature while hypothesis (A2) plays an important role in the ergodic and functional context of this paper. In the ergodic and finite dimension setting, say $\mathcal{E} \subseteq \mathbb{R}^d$ equipped with the Euclidian norm, (see, [17]), the condition (A2)(ii) appears with the following form

$$\mathbb{P}(X_{i,d} \in B(x, u) \mid \mathcal{F}_{i-1}) = c_{0f,d}(x)u^d \quad \text{a.s. as } u \to 0,$$

where $X_{i,d} = (X_i, \ldots, X_{i-d+1})$ and $c_0$ is a positive constant. It means that the conditional probability of the $d$-dimensional ball, given the $\sigma$-field $\mathcal{F}_{i-1}$, is asymptotically governed by a local dimension when the radius $u$ tends to zero. This assumption may be interpreted in terms of the fractal dimension which allows to obtain rates of convergence and the asymptotic normality for the kernel regression function without assuming the existence of marginal and conditional densities. This is of particular interest for chaotic models where the underlying process does not possess a density. Note that the statement (2.3) holds true whenever the conditional distribution $\mathbb{P}^{\hat{r}_{X_{i,d}}}_{(2)}$ has a continuous conditional density $f_{X_{i,d}}^{\hat{r}_{X_{i,d}}}(x) = f_i(x)$ at any point of the set $(x : f_{i,j}(x) > 0)$ and the constant $c_0$ takes the value $\sigma^{d/2}/\Gamma((d + 2)/2)$, where $\Gamma$ stands as the gamma function. Notice also that taking $g_{i,d}(u) = C_i \zeta(u)$ with $C_i$ a real random variable and $\zeta$ a function such that $\zeta(u)/\phi(u) \to 0$ as $u \to 0$, the condition (A2)(ii) is clearly satisfied while making use of the ergodic theorem. In Examples 1–3 below, we consider the $d$-vector process and the functional processes associated to autoregressive models of order 1 and find out explicit forms of every element in the condition (A2)(ii).

Consider the usual heteroskedastic regression model $Y_i = r(X_i) + \sigma(X_i)e_i$, where the random variables $e_i$’s stand as martingale differences with respect to the $\sigma$-field $\mathcal{G}_i$ generated by the random elements $(X_1, e_1), \ldots, (X_i, e_i), (X_{i+1})$. Clearly, we have $E(Y_i \mid \mathcal{G}_{i-1}) = r(X_i)$ almost surely. Moreover, if in addition we suppose that $E(e_i^2 \mid \mathcal{G}_{i-1}) = 1$ almost surely, then $E((Y_i - r(X_i))^2 \mid \mathcal{G}_{i-1})$ depends only on $X_i$. Therefore, the statements of Condition (A3) are satisfied. Notice that Gyorfi et al. [14] pointed out that their condition, similar to the condition (A3), is necessary to establish the consistency of the partitioning estimate they considered. Hypotheses (A4) stand as regularity conditions that are of usual nature.
Example 1. Consider the $d$-vector AR(1) model defined by

$$X_i = AX_{i-1} + \epsilon_i, \quad i \geq 1,$$

(2.4)

where $X_i$ and $\epsilon_i$ are real $d$-vectors. Suppose that the components of $\epsilon_i$ are independent and that $\epsilon_i$ is independent of the $\sigma$-field $\mathcal{F}_{i-1}$. Furthermore, suppose that $A$ is a diagonal matrix, i.e.,

$$A = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_d \end{pmatrix}.$$

Notice that the model (2.4) is Markovian of 1-order and that the process $(X_i)_{i \geq 1}$ is stationary whenever $\max_{1 \leq i \leq d} |\theta_i| < 1$. Observe, for any $x \in \mathbb{R}^d$ and any $i \geq 1$, that we have, for any $u > 0$,

$$F^x_{\mathcal{F}_{i-1}}(u) = \mathbb{P}(X_i \in B(x, u) \mid \mathcal{F}_{i-1}) = \mathbb{P}(d(X_i, x) \leq u \mid \mathcal{F}_{i-1}),$$

where $B(x, u)$ is the ball of center $x$ and radius $u$. Using the Markov property of the process $(X_i)_{i \geq 1}$, we can see that

$$F^x_{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d(X_i, x) \leq u \mid X_{i-1}).$$

Consider now the following function

$$F_{X_i|X_{i-1}=x}(u) = \mathbb{P}(d(X_i, x) \leq u \mid X_{i-1} = x).$$

Obviously, taking $d$ as the supremum norm, we have

$$F_{X_i|X_{i-1}=x}(u) = \mathbb{P}(x_j - u \leq \theta_j X_{i-1,j} + \epsilon_i \leq x_j + u, \quad 1 \leq j \leq d \mid X_{i-1} = x)$$

$$= \mathbb{P}(x_j - u - \theta_j t_j \leq \epsilon_i \leq x_j + u - \theta_j t_j, \quad 1 \leq j \leq d)$$

$$= \prod_{j=1}^{d} \mathbb{P}(x_j - u - \theta_j t_j \leq \epsilon_i \leq x_j + u - \theta_j t_j)$$

$$= \prod_{j=1}^{d} \left[ G(x_j + u - \theta_j t_j) - G(x_j - u - \theta_j t_j) \right],$$

(2.5)

where $G$ is the cumulative distribution function of $\epsilon_i$. Since $F^x_{\mathcal{F}_{i-1}}(u) = F_{X_i|X_{i-1}=x}(u)$, it follows that

$$F^x_{\mathcal{F}_{i-1}}(u) = \prod_{j=1}^{d} \left[ G(x_j + u - \theta_j X_{i-1,j}) - G(x_j - u - \theta_j X_{i-1,j}) \right].$$

Suppose that the distribution function is twice differentiable and that the second derivative is bounded. Obviously, this hypothesis is fulfilled in the Gaussian case. Observe also, in the neighborhood of $u = 0$, that

$$G(x_j + u - \theta_j X_{i-1,j}) - G(x_j - u - \theta_j X_{i-1,j}) = 2u G'(x_j - \theta_j X_{i-1,j}) + O(u^2)$$

$$= 2u (G'(x_j - \theta_j X_{i-1,j}) + O(u)).$$

Therefore, we have

$$F^x_{\mathcal{F}_{i-1}}(u) = 2^d u^d \prod_{j=1}^{d} \left( G'(x_j - \theta_j X_{i-1,j}) + O(u) \right)$$

$$= 2^d u^d \prod_{j=1}^{d} \left( g_j(x_j - \theta_j X_{i-1,j}) + O(u^{d+1}) \right)$$

$$= (2u)^d g(x - AX_{i-1}) + O(u^{d+1}),$$

where $g_j$ is the density of $\epsilon_{ij}$ and $g$ is the joint density of $\epsilon$.

Assume now that $\mathcal{E}$ is a separable Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$ and let $(e_j)_{j=1}^{\infty}$ be an orthonormal basis of $\mathcal{E}$. Let $x = \sum_{j=1}^{k} x_j e_j$ be a fixed element of $\mathcal{E}$. For $k \in \mathbb{N}^*$, consider the semi-metric $d_k$ defined, for any $(x, y) \in \mathcal{E}^2$, by

$$d_k(x, y) = \left( \sum_{j=1}^{k} \langle x - y, e_j \rangle^2 \right)^{\frac{1}{2}}.$$

(2.6)
We refer to [12], Lemma 13.6 page 213, for the proof that $d_k$ is actually a semi-metric. Suppose also that $X_i = \sum_{j=1}^{\infty} X_i^j \epsilon_j$ is a squared random element of $\mathcal{E}$ and set $X_i = (X_i^1, \ldots, X_i^k)$ and $x = (x_1, \ldots, x_k)$. Moreover, assume that the conditional density function $F_{X_i}^{\mathcal{F}_{i-1}}(x)$ with respect to the Lebesgue measure on $\mathbb{R}^k$ of $X_i$ given the $\sigma$-field $\mathcal{F}_{i-1}$ is a random function with almost sure continuous paths in the neighborhood of $x$ and $F_{X_i}^{\mathcal{F}_{i-1}}(x) > 0$. Notice that the decomposition $X_i = \sum_{j=1}^{\infty} X_i^j \epsilon_j$ may result from the Karhunen–Loeve expansion as for a centered Gaussian process and that $X_i$ is a $\mathbb{R}^k$-valued random vector with independent $N(0, 1)$ components which clearly possesses a density. It follows then, for $u > 0$, that

$$F_{X_i}^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d_k(x, X_i) \leq u|\mathcal{F}_{i-1})$$

$$= \mathbb{P}\left(\sum_{j=1}^{k} (x_j - X_i^j)^2 \leq u^2|\mathcal{F}_{i-1}\right)$$

$$= \mathbb{P}\left(\sum_{j=1}^{k} (x_j - X_i^j)^2 \leq u^2|\mathcal{F}_{i-1}\right)$$

$$= \mathbb{P}\left(\|X_i - x\|_{\text{Eucl}} \leq u|\mathcal{F}_{i-1}\right),$$

where $\|\cdot\|_{\text{Eucl}}$ is the Euclidean norm on $\mathbb{R}^k$. Clearly, from the almost sure continuity of the conditional density paths, we have

$$F_{X_i}^{\mathcal{F}_{i-1}}(u) = \int \int_{\mathbb{R}^k} f_{X_i}^{\mathcal{F}_{i-1}}(t) dt$$

$$= \frac{2\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} u^k F_{X_i}^{\mathcal{F}_{i-1}}(x) + o_{a.s.}(u^k).$$

While $x = (x_1, \ldots, x_k)$ is the projection of $x \in \mathcal{E}$ on the space $\mathbb{R}^k$, it suffices to take $f_{i,1}(x) = f_{X_i}^{\mathcal{F}_{i-1}}(x)$ to satisfy the condition (A2)(ii).

**Example 2.** On the Hilbert space $\mathcal{E}$ equipped with the norm $\|\cdot\|$ associated to the inner product $(\cdot, \cdot)$, consider the Hilbert autoregressive model of order one defined, for $n \geq 1$, by

$$X_n = \rho(X_{n-1}) + \epsilon_n,$$

where $(\epsilon_n)_{n \geq 1}$ is an i.i.d. sequence of Hilbert random variables such that $\epsilon_n$ is independent of $X_{n-1}$, and $\mathbb{E}\|\epsilon_n\|^2 < \infty$ and $\rho$ is a functional operator on $\mathcal{E}$. Taking the semi-metric defined in the statement (2.6), observe that

$$F_{X_i}^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d_k(x, X_i) \leq u|\mathcal{F}_{i-1}) = \mathbb{P}(d_k(x, \rho(X_{i-1}) + \epsilon_i) \leq u|\mathcal{F}_{i-1}).$$

Since we can write $\epsilon_i = \sum_{j=1}^{\infty} \epsilon_j^i \epsilon_j$ and, for any $s \in \mathcal{E}$, $\rho(s) = \sum_{j=1}^{\infty} (\rho(s))^j \epsilon_j$, it follows then that

$$F_{X_i|X_{i-1}=s}(u) := \mathbb{P}(d_k(x, \rho(X_{i-1}) + \epsilon_i) \leq u|X_{i-1} = s)$$

$$= \mathbb{P}\left(\sum_{j=1}^{k} (x_j - (\rho(s))^j + \epsilon_j^i)^2 \leq u^2\right)$$

$$= \mathbb{P}\left(\|\epsilon_i - (\rho(s)) - x\|_{\text{Eucl}} \leq u\right) = \mathbb{P}(\epsilon_i \in B_k((\rho(s)) - x, u)),$$

where $\epsilon_i = (\epsilon_1^i, \ldots, \epsilon_k^i), (\rho(s)) = ((\rho(s))^1, \ldots, (\rho(s))^k)$ and $B_k((\rho(s)) - x, u)$ is the ball in $\mathbb{R}^k$ of center $(\rho(s)) - x$ and radius $u$. Denote by $g$ the density function of $\epsilon_i$. Clearly, we have

$$F_{X_i|X_{i-1}=s}(u) = \int \int_{B_k((\rho(s)) - x, u)} g(t_1, \ldots, t_k) dt_1 \cdots dt_k$$

$$= \int \int_{B_k((\rho(s)) - x, u)} |g(t_1, \ldots, t_k) - g((\rho(s)) - x)| dt_1 \cdots dt_k + C u^k g((\rho(s)) - x).$$

When $g$ is assumed to be a Lipschitz function of order 1 with a constant $C > 0$, we obtain

$$F_{X_i|X_{i-1}=s}(u) = C u^k g((\rho(X_{i-1})) - x) + o(u^k).$$

Therefore,

$$F_{X_i}^{\mathcal{F}_{i-1}}(u) = F_{X_i|X_{i-1}}(u) = C u^k g((\rho(X_{i-1})) - x) + o(u^k). \Box$$
Example 3. Let $\mathcal{C}$ be a separate abstract space equipped with a semi-distance. Consider the autoregressive model of order one defined, for any $i \geq 1$, by

$$X_i = \rho(X_{i-1}) + \epsilon_i,$$

where $\epsilon_i = \eta_i h$ with a real random variable $\eta_i$ independent of $X_{i-1}$ and $h \in \mathcal{C}$ and $\rho$ is a functional operator on $\mathcal{C}$. For $(x, y) \in \mathcal{C}^2$, consider the semi-distance between $x$ and $y$ given by

$$d(x, y) = \left| \int (x(t) - y(t))dt \right|.$$

Observe; for any $u > 0$, that we have

$$F_{X_i}^u(u) = \mathbb{P} (d(X_i) \leq u | \mathcal{F}_{i-1}) = \mathbb{P} (d(X_i) \leq u | X_{i-1}).$$

Consequently, whenever $0 \neq \int h(t)dt < \infty$, we have

$$F_{X_i|X_{i-1}=s}(u) := \mathbb{P} (d(X_i) \leq u | X_{i-1} = s)$$

$$= \mathbb{P} \left( \left| \int x(t) - X_i(t)dt \right| \leq u | X_{i-1} = s \right)$$

$$= \mathbb{P} \left( \left| \int x(t) - \rho(X_{i-1}(t)) - \eta_i h(t)dt \right| \leq u | X_{i-1} = s \right)$$

$$= \mathbb{P} \left( \left| \int (x(t) - \rho(s)(t) - \eta_i h(t))dt \right| \leq u \right)$$

$$= \mathbb{P} \left( \frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \leq \eta_i \leq \frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \right)$$

$$= \Phi \left( \frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \right) - \Phi \left( \frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \right),$$

where $\Phi$ is the cumulative distribution function of $\eta_i$. Assuming now that $0 < \int h(t)dt < \infty$, $\int x(t)dt < \infty$ and $|\int \rho(s)(t)dt| < \infty$ for any $s \in \mathcal{C}$ and taking $\Phi$ as the $\mathcal{N}(0, 1)$ cumulative distribution function, we obtain

$$F_{X_i|X_{i-1}=s}(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}} \exp \left( -\frac{1}{2} \left( \frac{\int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \right)^2 \right) \left( 1 + o(1) \right).$$

Thus,

$$F_{X_i}^u(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}} \exp \left( -\frac{1}{2} \left( \frac{\int x(t)dt - \int \rho(X_{i-1})(t)dt}{\int h(t)dt} \right)^2 \right) \left( 1 + o(1) \right),$$

and the condition $(A2)(ii)$ is satisfied with

$$\phi(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}}. \quad \square$$

2.2. Bias and consistency

The following results give the order of the conditional bias and state the consistency with a rate of convergence of the regression estimate $\hat{r}_n(x)$. Before displaying them, consider the following conditions

(CB1) Suppose, for $i \geq 1$, that $E(r(X_i) - r(x))d(x, X_i, \mathcal{F}_{i-1}) = E(r(X_i) - r(x))d(x, X_i)) =: \psi(d(x, X_i))$, where the function $\psi$ is differentiable at $0$. Moreover, assume that $\psi(0) = 0$ and that $\psi'(0) \neq 0$.

(CB2) For any $i \geq 1$, the random function $f_{i,1}(x)$ appearing in Assumption $(A2)$ is almost surely bounded from above by a deterministic function $b_i(x)$ such that $n^{-1} \sum_{i=1}^n b_i(x) \rightarrow D(X_i) < \infty$, as $n \rightarrow \infty$.

Proposition 1. Under assumptions $(A1)$, $(A2)$, (CB1) and (CB2) and the fact that $f_i(x) > 0$ and $\left| \int_0^1 (sK(s))' \tau_0(s)ds \right| < \infty$, we have

$$B_n(x) = C_n(x) - r(x) = \frac{h\psi'(0)}{M_1} \left[ K(1) - \int_0^1 (sK(s))' \tau_0(s)ds + o_{a.s.}(1) \right] + O_{a.s.} \left( \frac{h}{n\phi(h)} \right) + O_{a.s.} \left( \frac{\log n}{\phi(h)} \right).$$
Proposition 2. Assume that conditions (A1)–(A4) hold true.
(a) If
\[ \frac{n \phi(h)}{\log \log(n)} \to \infty \text{ as } n \to \infty, \tag{2.7} \]
then, for any \( x \in \mathcal{E} \) such that \( f_1(x) > 0 \), we have
\[ \lim_{n \to \infty} \left( \frac{n \phi(h)}{\log \log(n)} \right)^{\frac{1}{2}} (\hat{r}_n(x) - C_n(x)) \overset{p}{=} 0. \]
(b) If in addition
\[ \frac{n \phi(h)h^{2\beta}}{\log \log(n)} \to 0 \text{ as } n \to \infty, \tag{2.8} \]
where \( \beta \) is specified in (A4)(i), then we have
\[ \lim_{n \to \infty} \left( \frac{n \phi(h)}{\log \log(n)} \right)^{\frac{1}{2}} (\hat{r}_n(x) - r(x)) \overset{p}{=} 0, \]
where \( \overset{p}{=} \) stands as the equality in probability.

2.3. Asymptotic normality

Theorem 1 below deals with the asymptotic normality of \( \hat{r}_n(x) \).

Theorem 1. Assume that conditions (A1)–(A4) hold true and that
\[ n \phi(h) \to \infty \text{ as } n \to \infty. \tag{2.9} \]
Then, for any \( x \in \mathcal{E} \) such that \( f_1(x) > 0 \), we have
(i) \( \sqrt{n \phi(h)}(\hat{r}_n(x) - C_n(x)) \overset{d}{\to} \mathcal{N}(0, \sigma^2(x)) \),
where
\[ \sigma^2(x) = \frac{M_2}{M_1^2} \frac{W_2(x)}{f_1(x)} \tag{2.10} \]
and \( \overset{d}{\to} \) denotes the convergence in distribution.
(ii) If in addition we suppose that
\[ h^\beta (n \phi(h))^{1/2} \to 0 \text{ as } n \to \infty, \tag{2.11} \]
where \( \beta \) is specified in the condition (A4), then we have
\[ \sqrt{n \phi(h)}(\hat{r}_n(x) - r(x)) \overset{d}{\to} \mathcal{N}(0, \sigma^2(x)). \]

Remark 1. (i) Notice that the constants \( M_1 \) and \( M_2 \) are strictly positive. Indeed, making use of the condition (A1) and the fact that the function \( \tau_0 \) is nondecreasing, it suffices to perform a simple integration by parts. Consequently, whenever \( W_2(x) > 0 \), we have \( \sigma^2(x) > 0 \).
(ii) Whenever \( \mathcal{E} = \mathbb{R}^d \), the asymptotic variance expression takes the form
\[ \sigma^2(x) = \frac{1}{d} \frac{W_2(x)}{f_d(x)} \frac{\int_0^1 K^2(u)u^{d-1}du}{\left( \int_0^1 K(u)u^{d-1}du \right)^2}, \]
where \( f_d(x) \) is the marginal density of the random vector \( X_{i,d} \) defined in (2.3).

Observe now in Theorem 1 that the limiting variance contains the unknown function \( f_1 \) and that the normalization depends on the function \( \phi \) which is not identifiable explicitly. Moreover, we have to estimate the quantities \( W_2 \) and \( \tau_0 \). Therefore, Corollary 1, below, which is a slight modification of Theorem 1, allows to have usable form of our results in practice.
As usually, the conditional variance $W_2(x)$ is estimated by

$$W_{2,n}(x) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{r}_n(x))^2 K \left( \frac{d(x_i)}{n} \right) = \frac{1}{n} \sum_{i=1}^{n} Y_i^2 K \left( \frac{d(x_i)}{n} \right) - \left( \hat{r}_n(x) \right)^2$$

Making use of the decomposition of $F_n(u)$ in (A2)(i), one may estimate $\tau_0(u)$ by

$$\tau_n(u) = \frac{F_{x,n}(uh)}{F_{x,n}(h)},$$

where

$$F_{x,n}(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{d(x_i) \leq u\}}.$$

Subsequently, for a given kernel $K$, the quantities $M_1$ and $M_2$ are estimated by $M_{1,n}$ and $M_{2,n}$ respectively replacing $\tau_0$ by $\tau_n$ in their respective expressions. Introduce now some further conditions needed to state Corollary 1. Set

(A5) (i) The conditional mean of $Y_i^2$ given the $\sigma$-field $\mathcal{G}_{i-1}$ depends only on $X_i$, i.e., there exists a function $g$ such that, for any $i \geq 1$, $E(Y_i^2 \mid \mathcal{G}_{i-1}) = g(X_i)$ almost surely,

(ii) The conditional variance of $Y_i^2$ given $\mathcal{G}_{i-1}$ depends only on $X_i$, i.e., for any $i \geq 1$, $E((Y_i^2 - g(X_i))^2 \mid \mathcal{G}_{i-1}) = U(X_i)$, almost surely, for some function $U$. Moreover, the function $U$ is continuous in a neighborhood of $x$, that is, $\sup_{[u,d(x),u] \leq h} |U(u) - U(x)| = o(1)$.

**Corollary 1.** Assume that conditions (A1)–(A5) hold true, $K'$ and $(K'^2)'$ are integrable functions and

$$nF_x(h) \longrightarrow \infty \quad \text{and} \quad h^\beta (nF_x(h))^{1/2} \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$ 

Then, for any $x \in \mathcal{E}$ such that $f_1(x) > 0$, we have

$$\frac{M_{1,n}}{\sqrt{M_{2,n}}} \left( \frac{1}{nW_{2,n}(x)} - \frac{1}{nF_{x,n}(h)} \right) \overset{d}{\longrightarrow} \mathcal{N}(0, 1).$$

**Remark 2.** Using Corollary 1, the asymptotic $100(1 - \alpha)\%$ confidence band for the regression function $r$ is given by

$$\hat{r}_n(x) \pm c_{\alpha} \frac{M_{1,n}}{\sqrt{M_{2,n}}} \left( \frac{1}{nW_{2,n}(x)} - \frac{1}{nF_{x,n}(h)} \right),$$

where $c_{\alpha}$ is the upper $\alpha/2$ quantile of the distribution of $\mathcal{N}(0, 1)$.

**Remark 3.** In some cases, the direct use of estimated constants appearing in the limiting law to construct the confidence bands may lead to poor convergence properties. Notice that this inconvenient can be avoided by using the functional version of the usual wild bootstrap method. In the ergodic functional data context, the algorithms available in the literature for regression estimates on functional data, as the one proposed by Ferraty et al. [9], may be considered in the same terms. However, the theoretical justification has to be established.

3. Proofs

In order to establish our results, introduce some additional notations. For $x \in \mathcal{E}$, set

$$Q_n(x) := (\hat{r}_{n,2}(x) - \tau_{n,2}(x)) - r(x)(\hat{r}_{n,1}(x) - \tau_{n,1}(x))$$

and

$$R_n(x) := -B_n(x)(\hat{r}_{n,1}(x) - \tau_{n,1}(x)).$$

Clearly, we have

$$\hat{r}_n(x) - C_n(x) = \frac{Q_n(x) + R_n(x)}{\hat{r}_{n,1}(x)}.$$
The major interest of the decomposition (3.3) comes from the fact that the summands of the term \( Q_n(x) \) form a martingale difference that allow to establish a central limit theorem for the central term \( Q_n(x) \). The proof of Theorem 1 is split up into several lemmas establishing respectively the convergence in probability of \( \hat{r}_{n,1}(x) \) to 1, the fact that \( R_n(x) \), suitably normalized, is actually equal to \( o_p(1) \) and the asymptotic normality of \( Q_n(x) \).

We start with a technical lemma which plays the same role as the classical Böchner’s lemma in the finite dimension space case.

**Lemma 1.** Assume that conditions (A1) and (A2)(i), (A2)(ii) and (A2)(iv) hold true. For any real numbers \( 1 \leq j \leq 2 + \delta \) and \( 1 \leq k \leq 2 + \delta \) with \( \delta > 0 \), as \( n \to \infty \), we have

\[
\begin{align*}
(i) \quad & \frac{1}{\phi(h)} \mathbb{E}[\Delta'_i(x) \mid \mathcal{F}_{i-1}] = M_{fi_1}(x) + O_{a.s.} \left( \frac{g_{i,x}(h)}{\phi(h)} \right), \\
(ii) \quad & \frac{1}{\phi(h)} \mathbb{E}[\Delta'_i(x)] = M_{fi_1}(x) + o(1) \\
\text{and} \quad & (\text{iii}) \quad \frac{1}{\phi^k(h)} (\mathbb{E}(\Delta_i(x)))^k = M_{fi_1}^k(x) + o(1).
\end{align*}
\]

**Remark 4.** The statements of Lemma 1 remain valid for any real numbers \( j \geq 1 \) and \( k \geq 1 \) provided that the condition (A2)(iv) is satisfied with \( j \geq 1 \) instead of \( 1 \leq j \leq 2 + \delta \).

**Proof of Lemma 1.** Observe, for \( j \geq 1 \), that we can write

\[
\mathbb{E}[\Delta'_i(x) \mid \mathcal{F}_{i-1}] = \int_0^h K'_i \left( \frac{u}{h} \right) d\mathbb{P}_{\mathcal{F}_{i-1}} (d(x, X_1) \leq u) = \int_0^1 K'_i(t) d\mathbb{P}_{\mathcal{F}_{i-1}} \left( \frac{d(x, X_1)}{h} \leq t \right). \tag{3.4}
\]

Since the function \( K'_i \) is of class \( C^1 \), obviously we have \( K'_i(t) = K'_i(0) + \int_0^1 (K'_i(u))' \frac{du}{t} \). Subsequently, since \( \mathbb{P}(d(x, X) \leq 0) = 0 \) for any \( x \in \mathcal{E} \), combining this result with the statement (3.4) and applying Fubini’s theorem, we obtain

\[
\begin{align*}
\mathbb{E}[\Delta'_i(x) \mid \mathcal{F}_{i-1}] &= K'_i(0)F_x \mathbb{P}_{\mathcal{F}_{i-1}}(h) + \int_0^1 \left[ \int_0^t (K'_i(u))' \frac{du}{t} \right] d\mathbb{P}_{\mathcal{F}_{i-1}} \left( \frac{d(x, X_1)}{h} \leq t \right) \\
&= K'_i(0)F_x \mathbb{P}_{\mathcal{F}_{i-1}}(h) + \int_0^1 \int_0^1 (K'_i(u))'1_{[u,1]}(t) d\mathbb{P}_{\mathcal{F}_{i-1}} \left( \frac{d(x, X_1)}{h} \leq t \right) du \\
&= K'_i(0)F_x \mathbb{P}_{\mathcal{F}_{i-1}}(h) + F_x \mathbb{P}_{\mathcal{F}_{i-1}}(h)[K'_i(1) - K'_i(0)] - \int_0^1 (K'_i(u))' F_x \mathbb{P}_{\mathcal{F}_{i-1}}(uh) du \\
&= K'_i(1)F_x \mathbb{P}_{\mathcal{F}_{i-1}}(h) - \int_0^1 (K'_i(u))' F_x \mathbb{P}_{\mathcal{F}_{i-1}}(uh) du.
\end{align*}
\]

Using first the condition (A2)(ii) and then the condition (A2)(iv), we may write

\[
\begin{align*}
\mathbb{E}[\Delta'_i(x) \mid \mathcal{F}_{i-1}] &= K'_i(1)f_{i_1}(x)\phi(h) + g_{i,x}(h)) - \int_0^1 (K'_i)'(u)\phi(uh)f_{i_1}(x) + g_{i,x}(uh))\frac{du}{t} \\
&= \phi(h) \left[ K'_i(1)f_{i_1}(x) + g_{i,x}(h)\frac{\phi(h)}{\phi(h)} - \int_0^1 (K'_i)'(u)(\tau_0(u) + o(1)) \left[ f_{i_1}(x) + g_{i,x}(uh)\frac{\phi(h)}{\phi(uh)} \right] \frac{du}{t} \right].
\end{align*}
\]

Since \( K'_i(h)\frac{\partial g_{i,x}}{\partial h} \) is an almost surely bounded function, making use of the dominated convergence theorem, we obtain

\[
\mathbb{E}[\Delta'_i(x) \mid \mathcal{F}_{i-1}] = \phi(h) f_{i_1}(x) \left[ K'_i(1) - \int_0^1 (K'_i)'(u)\tau_0(u) du \right] + O_{a.s.} (g_{i,x}(h)).
\]

The proof of the first part of Lemma 1 is then achieved. Part (ii) follows from part (i) with \( \mathcal{F}_i \) taken as the trivial \( \sigma \)-field. The proof of part (iii) follows while taking \( j = 1 \) in part (ii) and considering the exponent \( k \) in both sides of the equality. \( \square \)

The following lemma describes the asymptotic behavior of the term \( \hat{r}_{n,1}(x) \).

**Lemma 2.** Assume that hypotheses (A1)–(A2) and the condition (2.9) are satisfied. Then, for any \( x \in \mathcal{E} \) such that \( f_1(x) > 0 \), we have

\[
\lim_{n \to \infty} \hat{r}_{n,1}(x) \overset{P}{=} 1.
\]
Proof of Lemma 2. Observe that
\[ r_{n,1}(x) - 1 = R_{1,n}(x) + R_{2,n}(x), \]
where
\[ R_{1,n}(x) := \frac{1}{n\mathbb{E}[\Delta_1(x)]} \sum_{i=1}^{n} (\Delta_i(x) - \mathbb{E}[\Delta_i(x) \mid \mathcal{F}_{i-1}]), \]
\[ R_{2,n}(x) := \frac{1}{n\mathbb{E}[\Delta_1(x)]} \sum_{i=1}^{n} (\mathbb{E}[\Delta_i(x) \mid \mathcal{F}_{i-1}] - \Delta_1(x)) \]
\[ = \frac{1}{n\mathbb{E}[\Delta_1(x)]} \sum_{i=1}^{n} \mathbb{E}[\Delta_i(x) \mid \mathcal{F}_{i-1}] - 1. \]
Combining Lemma 1 with conditions (A2)-(ii) and (A2)-(iii), it is easily seen that \( R_{2,n}(x) = o_s(1) \) as \( n \to \infty \).

To handle the first term, observe that \( R_{n,1}(x) = \sum_{i=1}^{n} \tilde{L}_i(x) \), where \( \{\tilde{L}_n(x)\} \) is a triangular array of martingale differences with respect to the \( \sigma \)-field \( \mathcal{F}_{i-1} \). Combining Burkholder ([15], page 23) and Jensen inequalities, we obtain for any \( \epsilon > 0 \) that there exists a constant \( C_0 > 0 \) such that
\[ \mathbb{P}(|R_{n,1}(x)| > \epsilon) \leq C_0 \frac{\mathbb{E}([\Delta_1(x)^2])}{\epsilon^2\mathbb{E}(\mathbb{E}(\Delta_1(x)))^2} = O\left(\frac{1}{\epsilon^{2n\phi(h)}} + o(1)\right), \]
where the last equality results from Lemma 1. Since \( n\phi(h) \to \infty \) as \( n \to \infty \), we conclude then that \( R_{n,1}(x) = o_p(1) \) as \( n \to \infty \).

Lemma 3. Assume that hypotheses (A1)-(A2), (A3)(i), (A4)(i) and the condition (2.9) are satisfied. Then, for any \( x \in \mathcal{E} \) such that \( f_1(x) > 0 \), we have
\[ B_n(x) = O_p(h^\beta) \]
and
\[ R_n(x) = O_p\left(\frac{h^\beta}{\sqrt{n\phi(h)}}\right). \]

Proof of Lemma 3. First, we evaluate the conditional bias term. Observe that
\[ B_n(x) = \left(\frac{\tilde{T}_{n,2}(x) - r(x)\tilde{T}_{n,1}(x)}{\tilde{T}_{n,1}(x)}\right). \]
Similarly as in Lemma 2, it is easily seen that \( \tilde{T}_{n,1}(x) - 1 = o_p(1) \). Therefore, we have to establish that
\[ \tilde{B}_n(x) = \tilde{T}_{n,2}(x) - r(x)\tilde{T}_{n,1}(x) = O_{a.s.}(h^\beta). \]
Making use of conditions (A3)(i) and (A4)(i) one can easily see that
\[ |\tilde{B}_n(x)| \leq \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[(Y_i - r(x))\Delta_i(x) \mid \mathcal{F}_{i-1}] \]
\[ = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[(Y_i - r(x))\Delta_i(x) \mid \mathcal{F}_{i-1}] \]
\[ = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[(Y_i - r(x))\Delta_i(x) \mid \mathcal{F}_{i-1}] \]
\[ = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[(r(X_i) - r(x))\Delta_i(x) \mid \mathcal{F}_{i-1}] \]
\[ \leq \sup_{u \in [0,h]} \left| r(u) - r(x) \right| \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[(\Delta_i(x) \mid \mathcal{F}_{i-1})] \]
\[ = O_{a.s.}(h^\beta), \]
since the support of the kernel \( K \) is the interval \([0, 1]\).
To prove the second part of Lemma 3, since
\[ R_n(x) = -B_n(x)(\tilde{T}_{n,1}(x) - \tilde{T}_{n,1}(x)), \]
observe that \( \hat{r}_{n,1}(x) - \Gamma_{n,1}(x) \) is a martingale difference. Following now the same steps as in Lemma 4 below, we can establish that
\[
\sqrt{n\phi(h)}(\hat{r}_{n,1}(x) - \Gamma_{n,1}(x)) \xrightarrow{D} \mathcal{N}(0, \rho^2(x)),
\]
where
\[
\rho^2(x) = \frac{M_2}{M_1^2 f_1(x)}.
\]
Therefore, clearly we have
\[
\hat{r}_{n,1}(x) - \Gamma_{n,1}(x) = \mathcal{O}_p\left( \frac{1}{\sqrt{n\phi(h)}} \right).
\]
Combining this statement with the first part of Lemma 3, we obtain the desired result. \(\square\)

The next lemma establishes the asymptotic normality of the process \( Q_n \).

**Lemma 4.** Assume that hypotheses (A1)–(A4) and the condition (2.9) are satisfied. Then, for any \( x \in \mathcal{E} \) such that \( f_1(x) > 0 \), we have
\[
\sqrt{n\phi(h)}Q_n(x) \xrightarrow{D} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } n \to \infty.
\]

**Proof of Lemma 4.** First, we introduce some notations. Set
\[
\eta_{ni} = \left( \frac{\phi(h)}{n} \right)^{1/2} \left( Y_i - r(x) \right) \frac{\Delta_i(x)}{\mathbb{E} \Delta_1(x)}
\]
and defined \( \xi_{ni} := \eta_{ni} - \mathbb{E} \left[ \eta_{ni} \mid \mathcal{F}_{i-1} \right] \). It is easily seen that
\[
(n\phi(h))^{1/2} Q_n(x) = \sum_{i=1}^{n} \xi_{ni},
\]
where, for any fixed \( x \) in \( \mathcal{E} \), the summands in (3.9) form a triangular array of stationary martingale differences with respect to the \( \sigma \)-field \( \mathcal{F}_{i-1} \). This allows us to apply the central limit theorem for discrete-time arrays of real-valued martingales (see, [15], page 23) to establish the asymptotic normality of \( Q_n(x) \). This can be done if we establish the following statements:

(a) \( \sum_{i=1}^{n} \mathbb{E} \left[ \xi_{ni}^2 \mid \mathcal{F}_{i-1} \right] \xrightarrow{p} \sigma^2(x) \)

and

(b) \( n\mathbb{E} \left[ \xi_{ni}^2 I_{|\xi_{ni}| > \varepsilon} \right] = o(1) \) holds for any \( \varepsilon > 0 \) (Lindeberg condition).

**Proof of part (a).** Observe first that
\[
\mathbb{E} \left[ \eta_{ni} \mid \mathcal{F}_{i-1} \right] = \frac{1}{\mathbb{E} \Delta_1(x)} \left( \frac{\phi(h)}{n} \right)^{1/2} \mathbb{E} \left[ (r(X_i) - r(x)) \Delta_i(x) \mid \mathcal{F}_{i-1} \right]
\]
\[
\leq \frac{1}{\mathbb{E} \Delta_1(x)} \left( \frac{\phi(h)}{n} \right)^{1/2} \sup_{u \in B(x,h)} |r(u) - r(x)| \mathbb{E} \left[ \Delta_i(x) \mid \mathcal{F}_{i-1} \right]
\]
\[
= O(h^\beta) \left( \frac{\phi(h)}{n} \right)^{1/2} \left( \frac{f_{11}(x)}{f_1(x)} + O_{a.s.} \left( \frac{g_{1,x}(h)}{\phi(h)} \right) \right).
\]

Making use of the condition (A4)-(i) and Lemma 1, one has
\[
\mathbb{E} \left[ \eta_{ni} \mid \mathcal{F}_{i-1} \right] = \frac{1}{\mathbb{E} \Delta_1(x)} \left( \frac{\phi(h)}{n} \right)^{1/2} \mathbb{E} \left[ (r(X_i) - r(x)) \Delta_i(x) \mid \mathcal{F}_{i-1} \right]
\]
\[
\leq \frac{1}{\mathbb{E} \Delta_1(x)} \left( \frac{\phi(h)}{n} \right)^{1/2} \sup_{u \in B(x,h)} |r(u) - r(x)| \mathbb{E} \left[ \Delta_i(x) \mid \mathcal{F}_{i-1} \right]
\]
\[
= O(h^\beta) \left( \frac{\phi(h)}{n} \right)^{1/2} \left( \frac{f_{11}(x)}{f_1(x)} + O_{a.s.} \left( \frac{g_{1,x}(h)}{\phi(h)} \right) \right).
\]

Thus, by (A2)(ii)-(iii), we have
\[
\sum_{i=1}^{n} \mathbb{E} \left[ \eta_{ni} \mid \mathcal{F}_{i-1} \right] ^2 = O(h^{2\beta}) \left( \frac{\phi(h)}{n} \right)^{1/2} \left( \frac{f_{11}(x)}{f_1(x)} + O_{a.s.} \left( \frac{g_{1,x}(h)}{\phi(h)} \right) \right)^2
\]
\[
= O(h^{2\beta} \phi(h)) \left( \frac{1}{f_1^2(x)} \frac{1}{n} \sum_{i=1}^{n} f_{11}^2(x) + O_{a.s.}(1) \right)
\]
\[
= O_{a.s.}(\phi(h) h^{2\beta}).
\]
The statement (a) follows then if we show that
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{E} \left[ \eta_{ni}^2 | \mathcal{F}_{i-1} \right] \overset{p}{=} \sigma^2(x). \tag{3.12}
\]
To prove (3.12), observe that
\[
\sum_{i=1}^{n} \mathbb{E} \left[ \eta_{ni}^2 | \mathcal{F}_{i-1} \right] = \frac{\phi(h)}{n(\mathbb{E} \Delta_1(x))^2} \sum_{i=1}^{n} \mathbb{E} \left[ (Y_i - r(x))^2 \Delta_i^2(x) | \mathcal{F}_{i-1} \right] = J_{1n} + J_{2n}, \tag{3.13}
\]
where
\[
J_{1n} := \frac{\phi(h)}{n(\mathbb{E} \Delta_1(x))^2} \sum_{i=1}^{n} \mathbb{E} \left[ \Delta_i^2(x) \mathbb{E}[ (Y_i - r(X_i))^2 | \mathcal{G}_{i-1} ] | \mathcal{F}_{i-1} \right]
= \frac{\phi(h)}{n(\mathbb{E} \Delta_1(x))^2} \sum_{i=1}^{n} \mathbb{E} \left[ W_2(X_i) \Delta_i^2(x) | \mathcal{F}_{i-1} \right] \quad \text{by (A3)(ii)}
\]
and
\[
J_{2n} := \frac{\phi(h)}{n(\mathbb{E} \Delta_1(x))^2} \sum_{i=1}^{n} \mathbb{E} \left[ (r(X_i) - r(x))^2 \Delta_i^2(x) | \mathcal{F}_{i-1} \right]. \tag{3.14}
\]
We give now an upper bound for \( \mathbb{E} \left[ W_2(X_i) \Delta_i^2(x) | \mathcal{F}_{i-1} \right] \). Towards this end, we split it up into \( I_{n1} + I_{n2} \) with
\[
I_{n1} := \mathbb{E} \left[ W_2(x) \mathbb{E} \left[ \Delta_i^2(x) | \mathcal{F}_{i-1} \right] \right] \quad \text{and} \quad I_{n2} := \mathbb{E} \left[ (W_2(X_i) - W_2(x)) \Delta_i^2(x) | \mathcal{F}_{i-1} \right]. \tag{3.15}
\]
Making use of (A4)-(ii), one can write
\[
|I_{n2}| \leq \sup_{|x| \leq M} \mathbb{E} \left[ |W_2(u) - W_2(x)| \mathbb{E} \left[ \Delta_i^2(x) | \mathcal{F}_{i-1} \right] \right] = \mathbb{E} \left[ \Delta_i^2(x) | \mathcal{F}_{i-1} \right] \times o(1).
\]
Thus, in view of Lemma 1 part (i), we have
\[
\mathbb{E} \left[ W_2(X_i) \Delta_i^2(x) | \mathcal{F}_{i-1} \right] = (\sigma(1) + W_2(x)) \mathbb{E} \left[ \Delta_i^2(x) | \mathcal{F}_{i-1} \right]
= (\sigma(1) + W_2(x))(M_2 \phi(h) f_{i,1}(x) + o.a.s. (g_{i,1}(h))). \tag{3.16}
\]
Combining again Lemma 1 and conditions (A2)(ii)–(iii), it is easily seen that
\[
\lim_{n \to \infty} J_{1n} = \frac{M_2}{M_1^2} \frac{W_2(x)}{f_1(x)} \quad \text{almost surely.} \tag{3.17}
\]
whenever \( f_1(x) > 0 \). Consider now the term \( J_{2n} \). Making use of conditions (A2)(ii)–(iii) and (A4)-(i) and Lemma 1, one can write
\[
|J_{2n}| = O(t^{2b}) \left( \frac{\phi(h)}{n(\mathbb{E} \Delta_1(x))^2} \right) \sum_{i=1}^{n} \mathbb{E} \left[ \Delta_i^2(x) | \mathcal{F}_{i-1} \right]
= O(t^{2b}) \left( \frac{M_2}{M_1^2} \frac{1}{f_1(x)} + o.a.s. (1) \right) \to 0 \quad \text{almost surely as } n \to \infty. \tag{3.18}
\]
Therefore,
\[
\lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{E} \left[ \eta_{ni}^2 | \mathcal{F}_{i-1} \right] = \lim_{n \to \infty} (J_{1n} + J_{2n}) = \frac{M_2}{M_1^2} \frac{W_2(x)}{f_1(x)} := \sigma^2(x) \quad \text{almost surely}
\]
whenever \( f_1(x) > 0 \). This completes the Proof of part (a).

**Proof of part (b).** The Lindeberg condition results from Corollary 9.5.2 in [6] which implies that \( n \mathbb{E} \left[ \xi_{ni}^2 | (|\xi_{ni}| > \epsilon) \right] \leq 4n \mathbb{E} \left[ \xi_{ni}^2 | (|\eta_{ni}| > \epsilon/2) \right] \]. Let \( a > 1 \) and \( b > 1 \) such that \( \frac{1}{a} + \frac{1}{b} = 1 \). Making use of Hölder and Markov inequalities one can write, for all \( \epsilon > 0 \),
\[
\mathbb{E} \left[ \eta_{ni}^2 | (|\eta_{ni}| > \epsilon/2) \right] \leq \mathbb{E} [\eta_{ni}^2]^{2a/(\epsilon/2)} \times (\epsilon/2)^{2a/b}. \tag{3.19}
\]
Taking $C_0$ a positive constant and $2a = 2 + \delta$ (with $\delta$ as in (A4)(ii)) and using the condition (A4)(ii), we obtain

\[
4n\mathbb{E}[\eta^4_n I(\eta_n > \varepsilon/2)] \leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} n \mathbb{E} \left[ \left| Y_i - r(x) \Delta_i(x) \right|^{2+\delta} \right] \\
\leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} n \mathbb{E} \left[ \left| Y_i - r(x) \Delta_i(x) \right|^{2+\delta} |X_i| \right] \\
\leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} n \mathbb{E} \left[ (\Delta_i(x))^{2+\delta} \right] \\
\leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} n \left( \mathbb{E} \left[ (\Delta_i(x))^{2+\delta} \right] \right) \mathbb{E} \left[ |W_{2+\delta}(X_i)| \right] \\
\leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} n \left( \mathbb{E} \left[ (\Delta_i(x))^{2+\delta} \right] \mathbb{E} \left[ |W_{2+\delta}(x)| \right] + o(1) \right) \\
\leq C_0 \left( \frac{\phi(h)}{n} \right)^{(2+\delta)/2} \left( \frac{M_2 + df_1(x) + o(1)}{M_1^{2+\delta} f_1^2(x) + o(1)} \right) \left( |W_{2+\delta}(x)| + o(1) \right) = O((\phi(h))^{-(2+\delta)/2}), \quad (3.21)
\]

where the last equality follows from Lemma 1. This completes the Proof of part (b) since $n\phi(h) \to \infty$ as $n \to \infty$ and therefore the proof of Lemma 4. \[\square\]

**Proof of Proposition 1.** As a first step, observe that the conditional bias may be decomposed as follows

\[
B_n(x) = \frac{\bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x)}{\bar{r}_{n,1}(x)} + \frac{(\bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x))(\bar{r}_{n,1}(x) - \bar{r}_{n,1}(x)) + (\bar{r}_{n,1}(x) - \bar{r}_{n,1}(x))^2 C_n(x)}{(\bar{r}_{n,1}(x))^2}
\]

where $B^*_n(x)$ stands as the main term while $U_n(x)$ is the residual one. Considering $N_n(x)$ as the numerator in the form of $B^*_n(x)$, we have by (A3)(i) that

\[
N_n(x) = \frac{1}{n\mathbb{E}(\Delta_i(x))} \sum_{i=1}^n \left[ \mathbb{E}(Y_i\Delta_i(x) | F_{i-1}) - r(x)\mathbb{E}(\Delta_i(x) | F_{i-1}) \right] \\
= \frac{1}{n\mathbb{E}(\Delta_i(x))} \sum_{i=1}^n \mathbb{E} \left[ \Delta_i(x)(r(X_i) - r(x)) | F_{i-1} \right].
\]

Observe now, using the condition (CB1), that

\[
A_i := \mathbb{E} \left[ K \left( \frac{d(x, X_i)}{h} \right) (r(X_i) - r(x)) | F_{i-1} \right] = \mathbb{E} \left[ K \left( \frac{d(x, X_i)}{h} \right) \mathbb{E} \left[ (r(X_i) - r(x)) | (d(x, X_i), F_{i-1}) \right] | F_{i-1} \right] \\
= \mathbb{E} \left[ K \left( \frac{d(x, X_i)}{h} \right) \psi(d(x, X_i)) | F_{i-1} \right] = \int_0^1 K(t)\psi(th) dF_{x}^{F_{i-1}}(th).
\]

Since $\psi(0) = 0$, Taylor series expansion of the function $\psi$ up to the order one in the neighborhood of $t = 0$ gives

\[
\psi(th) = th\psi'(0) + o(h).
\]

Therefore, we have

\[
A_i := h\psi'(0) \int_0^1 tK(t) dF_{x}^{F_{i-1}}(th) + o(h) \int_0^1 K(t) dF_{x}^{F_{i-1}}(th) \\
= h\psi'(0) \left[ K(1)F_{x}^{F_{i-1}}(h) - \int_0^1 (sK(s))'F_{x}^{F_{i-1}}(sh) ds \right] + o(h) \left[ K(1)F_{x}^{F_{i-1}}(h) - \int_0^1 K'(s)F_{x}^{F_{i-1}}(sh) ds \right].
\]

Making use of Condition (A2)(ii), we obtain

\[
A_i := K(1)(\phi(h)f_{i,1}(x) + g_{i,x}(h))(h\psi'(0) + o(h)) - h\psi'(0) \int_0^1 (sK(s))'f_{i,1}(x) + g_{i,x}(hs) ds \\
- o(h) \int_0^1 K'(s)(\phi(h)f_{i,1}(x) + g_{i,x}(hs)) ds.
\]
\[ K(1)\phi(h) \left( f_{i,1}(x) + \frac{g_{r,a}(h)}{\phi(h)} \right) (h\psi'(0) + o(h)) - h\phi(h)\psi'(0) \int_0^1 (sK(s))' \phi(h) \left( f_{i,1}(x) + \frac{g_{r,a}(hs)}{\phi(h)} \right) \, ds \]

\[ - o(h)\phi(h) \int_0^1 K'(s) \phi(hs) \left( f_{i,1}(x) + \frac{g_{r,a}(hs)}{\phi(hs)} \right) \, ds. \]

Proceeding as in the proof of Lemma 1, we have

\[ A_i = K(1)h\phi(h) \left( f_{i,1}(x) + \frac{g_{r,a}(h)}{\phi(h)} \right) \psi'(0) + o(1) \]

\[ - h\phi(h)\psi'(0) \int_0^1 (sK(s))'(\tau_0(s) + o(1)) \left( f_{i,1}(x) + \frac{g_{r,a}(hs)}{\phi(hs)} \right) \, ds \]

\[ - o(h)\phi(h) \int_0^1 K'(s)(\tau_0(s) + o(1)) \left( f_{i,1}(x) + \frac{g_{r,a}(hs)}{\phi(hs)} \right) \, ds \]

\[ = \psi'(0)h\phi(h)f_{i,1}(x) \left[ K(1) - \int_0^1 (sK(s))'(\tau_0(s)ds \right] + O_{a.s.}(h\phi_a(h)). \]

Thus, making use of conditions (A2)(ii)-(iii) and Lemma 1, we obtain

\[ N_n(x) = \frac{1}{nE(\Delta_1(x))} \sum_{i=1}^n A_i \]

\[ = \frac{h\psi'(0)}{M_1} \left[ K(1) - \int_0^1 (sK(s))'(\tau_0(s)ds + o_{a.s.}(1) \right]. \]

Considering the residual term, it is easily seen that

\[ U_n(x) = N_n(x) \hat{r}_{n,1}(x) - \hat{r}_{n,1}(x) \]

\[ = \frac{B_n^*(x)}{\hat{r}_{n,1}(x)\hat{r}_{n,1}(x)} (\hat{r}_{n,1}(x) - \hat{r}_{n,1}(x))^2. \]

It is stated in [18], under conditions (A1), (A2) and (CB2) combined with the fact that the kernel \( K \) is bounded, that

\[ \hat{r}_{n,1}(x) - \hat{r}_{n,1}(x) = O_{a.s.} \left( \sqrt{\frac{\log n}{n\phi(h)}} \right). \]

So we conclude the proof by making use of the fact that both \( \hat{r}_{n,1}(x) \) and \( \hat{r}_{n,1}(x) \) converge almost surely to 1. \( \square \)

**Proof of Proposition 2.** (a) Following the decomposition (3.3), since by Lemma 2, \( \hat{r}_{n,1}(x) \Rightarrow 1 \) as \( n \to \infty \), it suffices to show that \( Q_n(x) \) and \( R_n(x) \) converge in probability to zero with a suitable rate.

By Lemma 3, we have \( R_n(x) = O_p \left( \frac{\delta^d}{\sqrt{\log n}} \right) \). Therefore, we obtain

\[ \left( \frac{n\phi(h)}{\log \log (n)} \right)^{\frac{1}{2}} R_n(x) = O_p \left( \frac{h^d}{\log \log (n)} \right) = o_p(1). \]

Since \( Q_n(x) \) is a sum of centered martingale differences, using successively the Burkholder inequality, the \( C_r \)-inequality and the Jensen inequality, it follows that

\[ \text{Var}(Q_n(x)) \leq \frac{C_0}{n^2(E(\Delta_1(x))^2) \sum_{i=1}^n} \mathbb{E} \left[ \Delta_1^2(x) \mathbb{E} \left[ (Y_i - r(x))^2 | X_i \right] \right] \]

\[ = \frac{C_0}{n^2(E(\Delta_1(x))^2) \sum_{i=1}^n} \mathbb{E} \left[ \Delta_1^2(x) W_2(X_i) \right]. \]

Subsequently, by the stationarity and the conditions (A3)(ii), we obtain

\[ \text{Var}(Q_n(x)) \leq \frac{C_0}{n(E(\Delta_1(x))^2)} \left[ \mathbb{E} \left( \Delta_1^2(x)(W_2(X_1) - W_2(x)) \right) + W_2(x)\mathbb{E}(\Delta_1^2(x)) \right] \]

\[ \leq \frac{C_0\mathbb{E}(\Delta_1^2(x))}{n(E(\Delta_1(x))^2)} \mathbb{I}(1) + W_2(x). \]
Making use of Lemma 1, we have
\[ n\phi(h)\text{Var}(Q_n(x)) = O \left( \frac{M_2 W_2(x)}{M_1^2 f_1(x)} \right). \]

Tchebycheff inequality combined with the condition (2.7) give then
\[ \mathbb{P} \left( (n\phi(h))^\frac{1}{2} Q_n(x) > \epsilon (\log \log(n))^\frac{1}{2} \right) = O \left( \frac{1}{(\log \log(n))} \right) \to 0 \quad \text{as} \quad n \to \infty. \]

This means that, as \( n \to \infty \),
\[ Q_n(x) = o_p \left( \left( \frac{\log \log(n)}{n\phi(h)} \right)^\frac{1}{2} \right). \]

This completes the proof of the part (a).

(b) The proof follows easily from the first part of Lemma 3 and the condition (2.7). \( \square \)

**Proof of Theorem 1.** By the decomposition (3.3), the first part of Theorem 1 follows by making use of Lemma 2, the second part of Lemma 3 and subsequently Lemma 4. If in addition we use the first part of Lemma 3, then we obtain the second part of Theorem 1. \( \square \)

The following lemma gives the consistency in probability of estimators \( \hat{r}_n, \hat{g}_n \) and \( W_{2,n} \) which is needed to prove Corollary 1.

**Lemma 5.** Assume that the conditions (A1), (A2), (A3), (A5) and (2.9) hold true. Then, we have
\[
\begin{align*}
&\text{(i) } \lim_{n \to \infty} \hat{r}_n(x) = r(x), \quad \text{(ii) } \lim_{n \to \infty} \hat{g}_n(x) = g(x) \quad \text{and} \quad \text{(iii) } \lim_{n \to \infty} W_{2,n}(x) = W_2(x).
\end{align*}
\]

**Proof of Lemma 5.** The proof of the first two statements uses Lemmas 2 and 3 and arguments similar to those used in the proof of Proposition 2. The statement (iii) is a direct consequence of the results (i)–(ii) and the decomposition (2.12). \( \square \)

**Proof of Corollary 1.** Observe that
\[
\frac{M_{n,1}}{\sqrt{M_{2,n}}} \sqrt{\frac{nF_{x,n}(h)}{W_{2,n}(x)}} \left( \hat{r}_n(x) - r(x) \right) = M_{n,1} \frac{M_2}{\sqrt{M_{2,n}}} \sqrt{\frac{nF_{x,n}(h)W_2(x)}{W_{2,n}(x)n\phi(h)f_1(x)}} \frac{M_1}{\sqrt{M_2}} \sqrt{\frac{n\phi(h)f_1(x)}{W_2(x)}} \left( \hat{r}_n(x) - r(x) \right).
\]

Making use of Theorem 1(ii), it follows that
\[
\frac{M_1}{\sqrt{M_2}} \frac{\sqrt{n\phi(h)f_1(x)}}{W_2(x)} \left( \hat{r}_n(x) - r(x) \right) \Rightarrow \mathcal{N}(0, 1).
\]

Therefore, we have to establish that
\[
\frac{M_{1,n} \sqrt{M_2}}{M_1 \sqrt{M_{2,n}}} \frac{\sqrt{F_{x,n}(h)W_2(x)}}{W_{2,n}(x)\phi(h)f_1(x)} \to 1, \quad \text{as} \quad n \to \infty.
\]

Obviously, from the consistency of the empirical distribution function and the decomposition in (A2)(i), for \( 0 < u \leq 1 \), we have
\[
\frac{F_{x,n}(h)}{\phi(h)f_1(x)} \to 1, \quad \text{as} \quad n \to \infty.
\]

Consequently, by the condition (A2)(iv), it follows that
\[
\frac{F_{x,n}(uh)}{F_{x,n}(h)} = \frac{F_{x,n}(uh)}{\phi(uh)f_1(x)} \times \frac{\phi(h)f_1(x)}{F_{x,n}(h)} \times \frac{\phi(uh)}{\phi(h)} \to \tau_0(u), \quad \text{as} \quad n \to \infty.
\]

Since \( 0 \leq \frac{F_{x,n}(uh)}{F_{x,n}(h)} \leq 1 \), by the dominated convergence theorem, whenever \((K')^i\) is integrable, we have
\[
M_{1,n} \to M_1, \quad \text{as} \quad n \to \infty.
\]

Similarly, whenever \((K^2')^i\) is integrable, we have
\[
M_{2,n} \to M_2, \quad \text{as} \quad n \to \infty.
\]
Moreover, the consistency of the $W_{n,2}$ established in Lemma 5 yields

$$\frac{W_{n,2}(x)}{W_2(x)} \xrightarrow{P} 1, \quad \text{as } n \to \infty.$$  

Clearly, the proof of Corollary 1 is achieved. □

Acknowledgments

The authors would like to thank the referees for their constructive criticisms and suggestions that have led to improve the presentation of this paper.

References