From Nonparametric Density Estimation to Parametric Estimation of Multidimensional Diffusion Processes

Julien Apala N'DRIN [a] and Ouagnina HILI [b]

[a] Laboratory of Applied Mathematics and computer science,
University of Cocody, 22 BP 582 Abidjan 22, Côte d'Ivoire

[b] Laboratory of Mathematics and New Technologies of Information,
National Polytechnique Institute Houphouët-Boigny of Yamoussoukro,
BP 1093 Yamoussoukro, Côte d'Ivoire

Conference on "Multivariate Count Analysis", Besançon, 4-6 july 2018

PLAN

- Introduction

PLAN

- Introduction
- Nonparametric density estimation

PLAN

- Introduction
- Nonparametric density estimation
- Estimation of the parameter

PLAN

- Introduction
- Nonparametric density estimation
- Estimation of the parameter
- Examples and simulations

PLAN

- Introduction
- Nonparametric density estimation
- Estimation of the parameter
- Examples and simulations
- Appendix

Introduction

• Diffusion processes are widely used for modeling purposes in various fields and especially in finance.

• As a diffusion process is Markovian, the maximum likelihood estimation is the natural choice for parameter estimation to get consistent and asymptotical normally estimator when the transition probability density is known (see **Dacunha-Castelle and Florens-Zmirou (1986)**).

• In the discrete case, for most diffusion processes, the transition probability density is difficult to calculate explicitly.

• For this, several methods have been developed such as the approximation of the likelihood function (see **Pedersen (1995) and Yoshida (1992)**), the approximation of the transition density (see Ait-Sahalia (2002)), schemes of approximation of the diffusion (see **Florens-Zmirou (1989)**) or methods based on martingale estimating functions (see **Bibby and Sϕrensen (1995)**).

Introduction

Here, we study the multidimensional diffusion model

$$dX_t = a(X_t, \theta)dt + b(X_t, \theta)dW_t, \quad t \geq 0$$

under the condition that X_t is positive recurrent and exponentially strong mixing.

Assumption : the diffusion process is observed at regular spaced times $t_k = k\Delta$ where Δ is a positive constant.

Purpose of the work : construct an estimator of θ based on minimum Hellinger distance method, using the density of the invariant distribution of the diffusion.

Introduction

• Let f_θ denote the density of the invariant distribution of the diffusion. The estimator of θ is that value (or values) $\widehat{\theta}_n$ in the parameter space Θ which minimizes the Hellinger distance between f_θ and $\widehat{f}_n$ where $\widehat{f}_n$ is a nonparametric density estimator of f_θ.

• The interest for this method of parametric estimation is that the minimum Hellinger distance estimation method gives efficient and robust estimators (see Beran (1977)).

• The minimum Hellinger distance estimators have been used in parameter estimation for independent observations (see **Beran (1977)**), for nonlinear time series models (see **Hili (1995)**) and recently for univariate diffusion processes (see **N'drin and Hili (2013)**).

Introduction

• In section 2, we present the statistical model and some conditions which imply that X_t is positive recurrent and exponentially strong mixing. Consistence and asymptotic normality of the kernel estimator of the density of the invariant distribution are studied in the same section.
• Section 3 defines the minimum Hellinger distance estimator of θ and studies its properties (consistence and asymptotic normality).
• Section 4 is devoted to some examples and simulations.
• Proofs of some results are presented in Appendix.

Nonparametric density estimation

We consider the d-dimensional diffusion process solution of the multivariate stochastic differential equation :

$$dX_t = a(X_t, \theta)dt + b(X_t, \theta)dW_t, \quad t \geq 0, \tag{1}$$

where

- $\{W_t\}_{t\geq 0}$ is a standard l-dimensional Wiener process,
- θ is an unknown parameter which varies in a compact subset Θ of $\mathbb{R}^s$,
- $a : \mathbb{R}^d \times \Theta \to \mathbb{R}^d$ is the drift coefficient
- $b : \mathbb{R}^d \times \Theta \to \mathbb{R}^d \times \mathbb{R}^l$ is the diffusion coefficient.
- The functions a and b are known up to the parameter θ and b is bounded.
- θ_0 is the unknown true value of the parameter.
- For a matrix A, A^t denotes the transpose of A and the notation $|.|$ denotes a vectorial norm or a matricial norm.

Nonparametric density estimation

The process X_t is observed at discrete time $t_k = k\Delta$ where Δ is a positive constant.

• **Assumptions on the model** :

(A_1) : there exists a constant C such that

$$|a(x,\theta) - a(y,\theta)| + |b(x,\theta) - b(y,\theta)| \leq C|x-y|$$

(A_2) : there exist constants $M_0 > 0$ and $r > 0$ such that

$$(a(x,\theta), x) \leq -r\|x\|, \quad \|x\| \geq M_0$$

(A_3) : the matrix function $b(x,\theta)$ is non degenerate, that is

$$\inf_x \inf_{|\lambda|=1} \lambda^t b(x,\theta) b(x,\theta)^t \lambda > 0, \quad \lambda \in \mathbb{R}^d.$$

Nonparametric density estimation

• Assumptions (A_1)-(A_3) ensure the existence of a unique strong solution for the equation (1) and an invariant measure for the process $\{X_t\}$ that admits a density with respect to the Lebesgue measure and the strong mixing property for $\{X_t\}$ with exponential rate (see Bianchi (2007), Pardoux and Veretennikov (2001)).
• α denotes the strong mixing coefficient.
• We assume that the initial value X_0 follows the invariant law, which implies that the process $\{X_t\}$ is strictly stationary.

Nonparametric density estimation

- The kernel estimator $\widehat{f}_n(x)$ of $f_\theta(x)$:

$$\widehat{f}_n(x) = \frac{1}{nb_n^d} \sum_{k=1}^{n} K\left(\frac{x - X_k}{b_n}\right), \quad x \in \mathbb{R}^d.$$

- (b_n) is a sequence of bandwidths such that $b_n \to 0$ and $nb_n^d \to +\infty$ as $n \to +\infty$
- $K : \mathbb{R}^d \to \mathbb{R}$ is a kernel function which satisfies :

(A_4)

(i) There exists $N_1 > 0$ such that $K(.) \leq N_1 < +\infty$,

(ii) $\int K(x)dx = 1$ and $|x|^d K(x) \to 0$ as $|x| \to \infty$,
(A_5) $\int u_i K(u)du = 0$ and $\int u_i^2 K(u)du < \infty$ for $i = 1, ..., d$.

Nonparametric density estimation

• Assumptions on the density of the invariant distribution :
(A_6) $f_\theta(.)$ is twice continuously differentiable with respect to x.
(A_7) $\theta_1 \neq \theta_2$ implies that $f_{\theta_1}(x) \neq f_{\theta_2}(x)$ for all $x \in \mathbb{R}^d$.
• Properties (consistence and asymptotic normality) of the kernel density estimator are examined in the following theorems.
• Proof of the two theorems are given in the Appendix.

almost sure consistency

Theorem 1

Under assumptions $(A_1)-(A_4)$, if the function $f_\theta(x)$ is continuous with respect to x for all $\theta \in \Theta$, then for any positive sequence (b_n) such that $b_n \to 0$ and $nb_n^d \to +\infty$ as $n \to +\infty$, $\widehat{f}_n(x) \to f_\theta(x)$ almost surely.

Theorem 2

Under assumptions $(A_1)-(A_6)$, if (b_n) is such that $nb_n^{d+4} \to 0$ as $n \to +\infty$ then the limiting distribution of $\sqrt{nb_n^d}\left(\widehat{f}_n(x) - f_\theta(x)\right)$ is $N\left(0, \tau^2(x)\right)$ where

$$\tau^2(x) = f_\theta(x) \int_{\mathbb{R}^d} K^2(u)du.$$

Estimation of the parameter

- The minimum Hellinger distance estimator of θ :

$$\widehat{\theta}_n = Arg \min_{\theta \in \Theta} H_2(\widehat{f}_n, f_\theta)$$

where

$$H_2(\widehat{f}_n, f_\theta) = \left\{ \int_{\mathbb{R}^d} |\widehat{f}_n^{1/2}(x) - f_\theta^{1/2}(x)|^2 dx \right\}^{1/2}.$$

- $\mathcal{G}$ denotes the set of squared integrable functions with respect to the Lebesgue measure on $\mathbb{R}^d$.
- The functional $T : \mathcal{G} \to \Theta$ is as follows : let $g \in \mathcal{G}$ and denote : $A(g) = \{\theta \in \Theta : H_2(f_\theta, g) = \min_{\gamma \in \Theta} H_2(f_\gamma, g)\}$ where H_2 is the Hellinger distance. If $A(g)$ is reduced to an unique element, then $T(g)$ is defined as the value of this element. Elsewhere, we choose an arbitrary but unique element of $A(g)$ and call it $T(g)$.

Almost sure consistency

Theorem 3

Assume that assumptions $(A_1) - (A_4)$ and (A_7) hold. If for all $x \in \mathbb{R}^d$, $f_\theta(x)$ is continuous at θ_0, then for any positive sequence (b_n) such that $b_n \to 0$ and $nb_n^d \to +\infty$, $\widehat{\theta}_n$ converges almost surely to θ_0 as $n \to +\infty$.

Proof

By theorem 1, $\widehat{f}_n(x) \to f_{\theta_0}(x)$ almost surely .
Using the inequality $\left(a^{1/2} - b^{1/2}\right)^2 \leq |a - b|$ for $a, b \geq 0$, we get

$$H_2^2(\widehat{f}_n, f_\theta) = \int_{\mathbb{R}^d} |\widehat{f}_n^{1/2}(x) - f_\theta^{1/2}(x)|^2 dx \leq \int_{\mathbb{R}^d} |\widehat{f}_n(x) - f_\theta(x)| dx.$$

Since

$$\int_{\mathbb{R}^d} \widehat{f}_n(x) dx = \int_{\mathbb{R}^d} f_\theta(x) dx = 1,$$

Almost sure consistency

$H_2^2(\widehat{f}_n, f_{\theta_0}) \rightarrow 0$ almost surely (see Devroye and Györfy (1995) and Glick (1974)).
By theorem 1 (see Beran (1977)), $T(f_{\theta_0}) = \theta_0$ uniquely on Θ ; then the functional T is continuous at f_{θ_0} in the Hellinger topology.
Therefore $\widehat{\theta}_n = T(\widehat{f}_n) \rightarrow T(f_{\theta_0}) = \theta_0$ almost surely.
This achieves the proof of the theorem.

Asymptotic normality

Denote

$$g_\theta = f_\theta^{1/2}, \quad \dot{g}_\theta = \frac{\partial g_\theta}{\partial \theta}, \quad \ddot{g}_\theta = \frac{\partial^2 g_\theta}{\partial \theta^t \partial \theta}$$

when these quantities exist. Furthermore, let

$$V_\theta(x) = \left\{ \int_{\mathbb{R}^d} \dot{g}_\theta(x) \dot{g}_\theta^t(x) dx \right\}^{-1} \dot{g}_\theta(x) \quad \text{and} \quad h_\theta(x) = \frac{\dot{g}_\theta(x)}{2 f_\theta^{1/2}(x)}.$$

• To prove asymptotic normality of the estimator of the parameter, we begin with two lemmas.

Asymptotic normality

Lemma 1

Let E_n be a subset of $\mathbb{R}^d$ and denote E_n^c the complementary set of E_n. Assume that

(1) assumptions (A_1)-(A_5) are satisfied,

(2) $h_{\theta_0}(.)$ is twice continuously differentiable with respect to x and

$$\sqrt{n} b_n^2 \int_{E_n} f_{\theta_0}(y) dy \to 0 \quad \text{and}$$

$$\sqrt{n} b_n^2 \int_{E_n} \left| \frac{\partial^2 h_{\theta_0}(y)}{\partial y_i^2} \right| f_{\theta_0}(y) dy \to 0 \quad \text{for} \quad i = 1, \ldots, d,$$

Asymptotic normality

Lemma 1

$$(3)\ \sqrt{n}\int_{E_n^c}\left(\int_{\mathbb{R}^d}|h_{\theta_0}(y+ub_n)|K(u)du\right)f_{\theta_0}(y)dy \to 0 \text{ and}$$

$$\sqrt{n}\int_{E_n^c}|\dot{g}_{\theta_0}(y)|f_{\theta_0}^{1/2}(y)dy \to 0,$$

$$(4)\ \mathbb{E}\left|h_{\theta_0}^{2+\delta}(X_1)\right| < \infty \text{ for some } \delta > 0,$$

then for any positive sequence (b_n) such that $b_n \to 0$, the limiting distribution of

$$\int_{\mathbb{R}^d}\sqrt{n}h_{\theta_0}(x)\widehat{f}_n(x)dx \text{ is } N(0,\Gamma) \text{ where } \Gamma = \frac{1}{4}\int \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx.$$

Asymptotic normality

The proof can be found in the Appendix.

Remark

• The two dimensional stochastic process (see section 4) with invariant density $f_\theta(x, y) = \frac{\sqrt{\beta\sigma}}{\pi} \exp(-\beta x^2 - \sigma y^2)$, $\beta > 0$, $\sigma > 0$ where $\theta = (\beta, \sigma)$, satisfies the conditions of lemma 1 with for example $E_n = [w_n; +\infty[\times[w_n; +\infty[$ a subset of $\mathbb{R}^2$ where $w_n = n$.

Asymptotic normality

Lemma 2

Let G_n be a compact set of $\mathbb{R}^d$ and denote by G_n^c the complementary set of G_n.
Suppose that assumptions (A_1)-(A_6) are satisfied and :

$$(1)\ \frac{1}{n^{1/2}b_n^d}\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left(\int_{\mathbb{R}^d}K^2(t)f_{\theta_0}(x-tb_n)dt\right)dx \to 0$$

(2) p, q and r are such that $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$ and

$$\frac{1}{n^{3/2}b_n^{d+d/p}}\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left(\left[\int_{\mathbb{R}^d}K^2(t)f_{\theta_0}(x-tb_n)dt\right]^{1/q+1/r}\right)dx$$

$\to 0$

Asymptotic normality

Lemma 2

$$(3)\ n^{1/2}b_n^4\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left(1+\left(\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i^2}\right)^2\right)dx\to 0$$

for $i=1,...,d$.

$$(4)\ \ \sqrt{n}\int_{G_n^c}|\dot{g}_{\theta_0}(x)|f_{\theta_0}^{1/2}(x)dx\to 0$$

$$(5)\ \sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\left(\int_{\mathbb{R}^d}K(u)f_{\theta_0}(x+ub_n)du\right)dx\to 0$$

then $R_n=\int_{\mathbb{R}^d}\sqrt{n}h_{\theta_0}(x)\left(\widehat{f}_n^{1/2}(x)-f_{\theta_0}^{1/2}(x)\right)^2dx\xrightarrow{\mathbb{P}}0$ as $n\to+\infty$.

Asymptotic normality

The proof can be found in the Appendix.

Remark

• Let $G_n = [-v_n; v_n] \times [-v_n; v_n]$ a compact set of $\mathbb{R}^2$ where $\{v_n, n \geq 1\}$ is a sequence of positive numbers diverging to infinity.
• Let $b_n = \frac{(log(n))^{1/2}}{n^r}$, $\frac{1}{8} < r < \frac{1}{4}$ and $v_n = (log(n))^q$, $\frac{1}{2} < q < 1$, then the two dimensional stochastic process with invariant density $f_\theta(x, y) = \frac{\sqrt{\beta\sigma}}{\pi} \exp(-\beta x^2 - \sigma y^2)$, $\beta > 0$, $\sigma > 0$ where $\theta = (\beta, \sigma)$, satisfies the conditions of lemma 2.

Asymptotic normality

Theorem 4

Under assumption (A_7) and conditions of lemma 1 and lemma 2, if
(i) for all $x \in \mathbb{R}^d$, g_θ is twice continuously differentiable at θ_0,
(ii) the components of $\dot{g}_{\theta_0}$ and $\ddot{g}_{\theta_0}$ belong to L_2 and if the norms of these components are continuous functions at θ_0,
(iii) θ_0 is in the interior of Θ and $\int \ddot{g}_{\theta_0}(x) g_{\theta_0}(x)dx$ is a non-singular matrix,
then the limiting distribution of $\sqrt{n}[\widehat{\theta}_n - \theta_0]$ is $N(0, \lambda^2)$ where

$$\lambda^2 = \frac{1}{4}\left\{\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}^t_{\theta_0}(x)dx\right\}^{-1}.$$

Asymptotic normality

Proof

Denote $\Gamma_n(\theta) = \sqrt{n}[\widehat{\theta}_n - \theta_0]$,From theorem 2 (see Beran (1977)), we have :

$$\begin{aligned}
\Gamma_n(\theta) &= \sqrt{n}\left\{\int_{\mathbb{R}^d} V_{\theta_0}(x)[\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)]dx\right\} \\
&+ \sqrt{n}\left\{A_n \int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)[\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)]dx\right\} \\
&= \left\{\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx\right\}^{-1} \int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)[\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)]dx \\
&+ A_n \int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)[\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)]dx
\end{aligned}$$

where A_n is a $(m \text{ x } m)$ matrix which tends to 0 as $n \to +\infty$.

Asymptotic normality

We have

$$\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x) = \frac{\widehat{f}_n(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{1/2}(x)} - \frac{\left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)\right)^2}{2f_{\theta_0}^{1/2}(x)}.$$

Denote

$$D_n = \int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)[\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)]dx.$$

We have

Asymptotic normality

$$
\begin{aligned}
D_n &= \int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)\frac{\widehat{f}_n(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{1/2}(x)}dx \\
&- \int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)\frac{\left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)\right)^2}{2f_{\theta_0}^{1/2}(x)}dx \\
&= \int_{\mathbb{R}^d} \sqrt{n}\frac{\dot{g}_{\theta_0}(x)}{2f_{\theta_0}^{1/2}(x)}\widehat{f}_n(x)dx - \frac{1}{2}\int_{\mathbb{R}^d} \sqrt{n}\dot{g}_{\theta_0}(x)f_{\theta_0}^{1/2}(x)dx - R_n \\
&= \int_{\mathbb{R}^d} \sqrt{n}h_{\theta_0}(x)\widehat{f}_n(x)dx - R_n
\end{aligned}
$$

where $R_n = \int_{\mathbb{R}^d} \sqrt{n}h_{\theta_0}(x)\left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)\right)^2 dx.$

Asymptotic normality

By lemma 2, $R_n \to 0$ in probability as $\to \infty$; then, the limiting distribution of $\sqrt{n}[\widehat{\theta}_n - \theta_0]$ is reduced to that of

$$\left\{ \int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x) \dot{g}^t_{\theta_0}(x) dx \right\}^{-1} \int_{\mathbb{R}^d} \sqrt{n} h_{\theta_0}(x) \widehat{f}_n(x) dx$$

since $A_n \to 0$. But

$$\int_{\mathbb{R}^d} \sqrt{n} h_{\theta_0}(x) \widehat{f}_n(x) dx \xrightarrow{\mathcal{L}} N(0, \Gamma) \text{ with } \Gamma = \frac{1}{4} \int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x) \dot{g}^t_{\theta_0}(x) dx$$

from lemma 1. Therefore the limiting distribution of

$$\left\{ \int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x) \dot{g}^t_{\theta_0}(x) dx \right\}^{-1} \int_{\mathbb{R}^d} \sqrt{n} h_{\theta_0}(x) \widehat{f}_n(x) dx \text{ is } N(0, \lambda^2)$$

Asymptotic normality

where

$$\begin{aligned}\lambda^2 &= \left(\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx\right)^{-1} \frac{1}{4}\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx \\ &\times \left\{\left(\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx\right)^{-1}\right\}^t \\ &= \frac{1}{4}\left\{\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx\right\}^{-1}.\end{aligned}$$

This completes the proof of the theorem.

The two dimensional Ornstein-Uhlenbeck

We consider the two dimensional Ornstein-Uhlenbeck process solution of the stochastic differential equation

$$dZ_t = AZ_t dt + dW_t, \quad Z_0 = z_0 \tag{2}$$

where

$$A = \begin{pmatrix} -\beta & 0 \\ 0 & -\sigma \end{pmatrix}, \quad \beta > 0, \quad \sigma > 0$$

Let $Z = (X, Y)$ and $z = (x, y)$, we have :

$$a(z, \theta) = \begin{pmatrix} -\beta & 0 \\ 0 & -\sigma \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad b(z, \theta) = b = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $\theta = (\beta, \sigma)$.

The two dimensional Ornstein-Uhlenbeck

- $a(z,\theta)$ and $b(z,\theta)$ satisfy assumptions (A_1)-(A_3). Therefore, Z_t is exponentially strong mixing and the invariant distribution μ_θ admits a density f_θ with respect to the Lebesgue measure.
- Furthermore (see Jacobsen (2001)), $\mu_\theta = N(0,\Gamma)$, the Gaussian distribution on $\mathbb{R}^2$ with Γ the unique symmetric solution of the equation

$$C + A\Gamma + \Gamma A^t = 0 \ \text{ where } \ C = bb^t. \tag{3}$$

- The solution of the equation (3) is $\Gamma = \begin{pmatrix} \frac{1}{2\beta} & 0 \\ 0 & \frac{1}{2\sigma} \end{pmatrix}$.
- Therefore (see Caumel (2011)), the density of the invariant distribution is

$$f_\theta(x,y) = \frac{\sqrt{\beta\sigma}}{\pi}\exp\left(-\beta x^2 - \sigma y^2\right)$$

The two dimensional Ornstein-Uhlenbeck

- The minimum Hellinger distance estimator of θ is defined by :

$$\widehat{\theta}_n = Arg \min_{\theta \in \Theta} H_2(\widehat{f}_n, f_\theta)$$

$$\text{where } \; H_2(\widehat{f}_n, f_\theta) = \left\{ \int_{\mathbb{R}^2} \left| \widehat{f}_n^{1/2}(x,y) - f_\theta^{1/2}(x,y) \right|^2 dxdy \right\}^{1/2}$$

$$\text{with } \; \widehat{f}_n(x,y) = \frac{1}{nb_n^2} \sum_{k=1}^{n} K_1\left(\frac{x - X_k}{b_n}\right) K_1\left(\frac{y - Y_k}{b_n}\right)$$

$$\text{and } \; f_\theta(x,y) = \frac{\sqrt{\beta\sigma}}{\pi} \exp\left(-\beta x^2 - \sigma y^2\right)$$

where K_1 is a kernel function which satisfies conditions (A_4) and (A_5) such that $K_1(x)K_1(y) = K(x,y)$.

The two dimensional Ornstein-Uhlenbeck

- Let $W = (W^{(1)}, W^{(2)})$, we can write equation (2) as follows :

$$\begin{pmatrix} dX_t \\ dY_t \end{pmatrix} = \begin{pmatrix} -\beta & 0 \\ 0 & -\sigma \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} dt + \begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \end{pmatrix}$$

which gives the the following system

$$\begin{cases} dX_t & = -\beta X_t dt + dW_t^{(1)} \\ dY_t & = -\sigma Y_t dt + dW_t^{(2)} \end{cases}$$

- Thus, $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are two independent univariate Ornstein-Uhlenbeck processes of parameters β and σ respectively.

The two dimensional Ornstein-Uhlenbeck

• We now give simulations for different parameter values using the R language. For each process, we generate sample paths using the package "sde" (see Lacus (2008)) and to compute a value of the estimator, we use the function "nlm" (see Lafaye de Micheaux, Drouilhet and Liquet (2011)) of the R language.

• The kernel function K_1 is the density of the standard normal distribution. We use the bandwidth $b_n = \frac{\sqrt{log(n)}}{n^{0.24}}$ according to conditions on the bandwidth.

• Simulations are based on 1000 observations of the Ornstein-Uhlenbeck process with 200 replications. Simulation results are given in the table1.

The two dimensional Ornstein-Uhlenbeck

$\theta_0 = (\beta_0, \sigma_0)$	$\widehat{\theta} = (\widehat{\beta},\ \widehat{\sigma})$	
	Means	Standard errors
$(0.3,\ 0.7)$	$(0.2985977,\ 0.6998527)$	$(0.01076013,\ 0.01032311)$
$(0.5,\ 2)$	$(0.4954066,\ 1.998997)$	$(0.0341282,\ 0.008429909)$
$(1,\ 2.4)$	$(0.9987882,\ 2.398991)$	$(0.009604874,\ 0.01262858)$
$(1,\ 3)$	$(0.998918,\ 2.999193)$	$(0.008726621,\ 0.01034987)$
$(0.223,\ 0.6)$	$(0.2223449,\ 0.6006928)$	$(0.01048311,\ 0.01224315)$

Table1 : means and standard errors of the minimum Hellinger distance estimator

The two dimensional Ornstein-Uhlenbeck

In table1, θ_0 denotes the true value of the parameter and $\widehat{\theta}$ denotes an estimation of θ_0 given by the minimum Hellinger distance estimator. Simulation results illustrate the good properties of the estimator. Indeed, the means of the estimator are quite close to the true values of the parameter in all cases and the standard errors are low.

Homogeneous Gaussian diffusion process

• The Homogeneous Gaussian diffusion process (see Sϕrensen (2001)), solution of the stochastic differential equation :

$$dX_t = (A + BX_t)dt + \sigma dW_t, \quad X_0 = x_0 \tag{4}$$

• $\sigma > 0$ is known,
• W is a two-dimensional Brownian motion,
• B is a 2×2 matrix with eigenvalues with stricly negative parts,
• A a 2×1 matrix.
• By condition on the matrix B, X has an invariant probability $\mu = N(m, \Gamma)$ where $m = -B^{-1}A$ and Γ the unique symetric solution of the equation

$$C + B\Gamma + \Gamma B^t = 0 \text{ where } C = DD^t \text{ and } D = \sigma I \text{ with } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let $A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ and $B = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{12} & \beta_{22} \end{pmatrix}$ with $\beta_{11} < 0$ and $\beta_{22} < 0$.

Homogeneous Gaussian diffusion process

As in Sϕrensen (2001), we suppose that $\sigma = \sqrt{2}$. In the following, we suppose that $\beta_{11}\beta_{22} - \beta_{12}^2 \neq 0$.
Then we have

$$B^{-1} = \frac{1}{\beta_{11}\beta_{22} - \beta_{12}^2} \begin{pmatrix} \beta_{22} & -\beta_{12} \\ -\beta_{12} & \beta_{11} \end{pmatrix},$$

$$m = \frac{1}{\beta_{11}\beta_{22} - \beta_{12}^2} \begin{pmatrix} \alpha_2\beta_{12} - \alpha_1\beta_{22} \\ \alpha_1\beta_{12} - \alpha_2\beta_{11} \end{pmatrix}$$

and

$$\Gamma = \frac{1}{\beta_{11}\beta_{22} - \beta_{12}^2} \begin{pmatrix} -\beta_{22} & \beta_{12} \\ \beta_{12} & -\beta_{11} \end{pmatrix}.$$

Homogeneous Gaussian diffusion process

Γ is invertible and we have $\Gamma^{-1} = \begin{pmatrix} -\beta_{11} & -\beta_{12} \\ -\beta_{12} & -\beta_{22} \end{pmatrix}$. Hence, the invariant density of μ is

$$\begin{aligned} f_\theta(x) &= \frac{1}{(\sqrt{2\pi})^2\sqrt{|det(\Gamma)|}} \exp\left(-\frac{1}{2}(x-m)^t\Gamma^{-1}(x-m)\right) \\ &= \frac{\sqrt{|\beta_{11}\beta_{22}-\beta_{12}^2|}}{2\pi} \exp\left(\frac{1}{2}\beta_{11}\left(x_1 - \frac{\alpha_2\beta_{12}-\alpha_1\beta_{22}}{\beta_{11}\beta_{22}-\beta_{12}^2}\right)^2\right) \\ &\times \exp\left(\frac{1}{2}\beta_{22}\left(x_2 - \frac{\alpha_1\beta_{12}-\alpha_2\beta_{11}}{\beta_{11}\beta_{22}-\beta_{12}^2}\right)^2\right) \\ &\times \exp\left(\beta_{12}\left(x_1 - \frac{\alpha_2\beta_{12}-\alpha_1\beta_{22}}{\beta_{11}\beta_{22}-\beta_{12}^2}\right)\left(x_2 - \frac{\alpha_1\beta_{12}-\alpha_2\beta_{11}}{\beta_{11}\beta_{22}-\beta_{12}^2}\right)\right). \end{aligned}$$

Homogeneous Gaussian diffusion process

As in Sϕrensen (2001), the true values of the parameter $\theta = (\alpha_1, \alpha_2, \beta_{11}, \beta_{22}, \beta_{12})$ is $\theta_0 = (4, 1, -2, -3, 1)$ and $\sigma = \sqrt{2}$. For simulation, we must write the stochastic differential equation (7) in matrix form as follows :

$$\begin{pmatrix} dX_t^{(1)} \\ dX_t^{(2)} \end{pmatrix} = \begin{pmatrix} 4 - 2X_t^{(1)} + X_t^{(2)} \\ 1 + X_t^{(1)} - 3X_t^{(2)} \end{pmatrix} dt + \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \end{pmatrix}$$

Now, we can simulate a sample path of the Homogeneous Gaussian diffusion using the "yuima" package of R language Lacus (2011). We generate 500 sample paths of the process, each of size 500. The kernel function and the bandwidth are those of the previous example.

We compare the minimum Hellinger distance estimator (MHD) and the estimator obtained in Sϕrensen (2001) by estimating function.

Homogeneous Gaussian diffusion process

Table 2 summarizes results of simulation of means and standard errors of the different estimators.

θ_0	$\widehat{\theta}$(MHD)		$\widehat{\theta}$(Estimating Function)	
	Means	Standard errors	Means	Standard errors
$\alpha_1 = 4$	3.996942	0.0005203	4.0349	0.2904
$\alpha_2 = 1$	1.00776	0.001311968	1.0035	0.2891
$\beta_{11} = -2$	-2.007696	0.001315799	-2.0155	0.1248
$\beta_{22} = -3$	-2.982749	0.002923666	-3.0247	0.1978
$\beta_{12} = 1$	1.009081	0.001513984	1.0078	0.1177

Table 2 : means and standard errors of the estimators

Homogeneous Gaussian diffusion process

Table 2 shows that the two estimators have good behavior. For the two methods, the means of the estimators are close to the true values of the parameter. But the standard errors of the MHD estimator are lower than those of the estimating function estimator.

References

D. Dacunha-Castelle and D. Florens-Zmirou.
Estimation of the coefficients of a diffusion from discrete observations.
Stochastics, 19 : 263–284, 1986.

A. R. Pedersen.
A New approach to maximum likelihood estimation for stochastic differential equations based on discrete observations.
Scand.J.Statist, 22 : 55–71, 1995.

N. Yoshida.
Estimation for diffusion processes from discrete observation.
Journal of Multivariate Analysis , 41(2) : 220–242, 1992.

References

Y. Aït-Sahalia.
Maximum likelihood estimation of discretely sampled diffusions : a closed-form approximation approach.
Econometrica, 70(1) : 223–262, 2002.

D. Florens-Zmirou.
Approximate discrete-time schemes for statistics of diffusion processes.
Statistics, 20(4) : 547–557, 1989.

B. M. Bibby and M. Sørensen.
Martingale estimation functions for discretely observed diffusion processes.
Bernoulli, 1 : 17–39, 1995.

References

R. Beran.
Minimum Hellinger distance estimates for parametric models.
Ann.Stat, 5 : 445–463, 1977.

O. Hili.
On the estimation of nonlinear time series models.
Stochastics Rep, 52 : 207–226, 1995.

J. A. N'drin and O. Hili.
Parameter estimation of one-dimensional diffusion process by minimum Hellinger distance method.
Random Oper. Stoch. Equ., 21(4) : 403–424, 2013.

References

A. Bianchi.
Nonparametric trend coefficient estimation for multidimensional diffusions.
C. R. Acad. Sci. Paris, Ser.I, 345 : 101–105, 2007.

E. Pardoux and Yu. A. Veretennikov.
On the Poisson equation and diffusion approximation.I.
Ann. Probab., 29 : 1061–1085, 2001.

Yu. A. Veretennikov.
On polynomial mixing bounds for stochastic differential equations.
Stochastic Process, 70 : 115–127, 1997.

References

L. Devroye and L. Györfi.
Nonparametric density estimation : the L_1 view.
Wiley, New York, 1985.

N. Glick.
Consistency conditions for probability estimators and integrals of density estimators.
Utilitas Mathematica, 6 : 61–74, 1974.

M. Jacobsen.
Examples of multivariate diffusions : the time-reversibility, a Cox-Ingersoll-Ross type process.
Departement of Statistics and Operations Research, University of Copenhagen, Preprint 2001 – 6.

References

Y. Caumel.
Probabilités et processus stochastiques.
Springer-Verlag, France, 2011.

S. M. Iacus.
Simulation and Inference for Stochastic Differential Equations.
Springer series in statistics, Springer, New York 2008.

P. Lafaye de Micheaux, R. Drouilhet and B. Liquet.
Le logiciel R : Maitriser le langage-Effectuer des analyses statistiques.
Springer-Verlag France, 2011.

References

H. Sørensen.
Discretely observed diffusions : approximation of the continuous-time score function.
Scand.J.Statist, 28 : 113–121, 2001.

S. M. Iacus.
Option pricing and estimation of financial models with R.
First edition. John wiley and Sons, Ltd, 2011.

G. G. Roussas.
Nonparametric estimation in Markov Processes.
Annals of Inst. of Stat. Math, 73–87, 1969.

References

D. Bosq.
Nonparametric Statistics for Stochastic Processes : estimation and prediction, second edition.
Springer-Verlag New York, 1998.

S. Dharmenda and E. Masry.
Minimum complexity regression estimation with weakly dependent observations.
IEEE Transactions on information theory, 42(6) : 2133–2145, 1996.

F. Dominique and F. Aimé.
Calcul des probabilités-Cours, exercices et problèmes corrigés, 2e édition.
Dunod, 1998.

Appendix :Proof of Theorem 1

$$
\begin{aligned}
|\widehat{f}_n(x) - f_\theta(x)| &= \left| \left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right) + \left(\mathbb{E}\widehat{f}_n(x) - f_\theta(x)\right) \right| \\
&\leq \left| \left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right) \right| + \left| \left(\mathbb{E}\widehat{f}_n(x) - f_\theta(x)\right) \right|
\end{aligned}
$$

Step 1 :

$$
\begin{aligned}
\mathbb{E}\widehat{f}_n(x) &= \mathbb{E}\left[\frac{1}{nb_n^d}\sum_{k=1}^{n} K\left(\frac{x - X_k}{b_n}\right)\right] \\
&= \frac{1}{b_n^d}\mathbb{E}K\left(\frac{x - X_1}{b_n}\right) \\
&= \frac{1}{b_n^d}\int_{\mathbb{R}^d} K\left(\frac{x-u}{b_n}\right) f_\theta(u)du \\
&= \frac{1}{b_n^d}\int_{\mathbb{R}^d} K\left(\frac{t}{b_n}\right) f_\theta(x-t)dt \to f_\theta(x).
\end{aligned}
$$

Appendix :Proof of Theorem 1

by theorem 2.1 (see Roussas (1969)). Hence

$$\mathbb{E}\widehat{f}_n(x) - f_\theta(x) \to 0. \tag{5}$$

Step 2 :

$$|\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)| = \frac{1}{nb_n^d}|\sum_{k=1}^{n} Y_k| \text{ where}$$

$$Y_k = K\left(\frac{x - X_k}{b_n}\right) - \mathbb{E}K\left(\frac{x - X_k}{b_n}\right)$$

•$\mathbb{E}(Y_k) = 0$
•$|Y_k| \leq 2N_1$
Then by theorem 2.1 (see N'drin and Hili (2013)), we have for all $\epsilon > 0$

Appendix :Proof of Theorem 1

$$\mathbb{P}\left\{\frac{1}{nb_n^d}|\sum_{k=1}^{n} Y_k| > \epsilon\right\} = \mathbb{P}\left\{\frac{1}{n}|\sum_{k=1}^{n} Y_k| > \epsilon b_n^d\right\}$$

$$\leq 2C\exp\left(-\frac{\epsilon^2 nb_n^{2d}}{2\left(\mathbb{E}|Y_1|^2 + \frac{2\epsilon N_1 b_n^d}{3}\right)}\right).$$

We have

Appendix :Proof of Theorem 1

$$
\begin{aligned}
\mathbb{E}|Y_1|^2 &= \mathbb{E}\left|K\left(\frac{x-X_1}{b_n}\right) - \mathbb{E}K\left(\frac{x-X_1}{b_n}\right)\right|^2 \\
&= \mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right) - \left(\mathbb{E}K\left(\frac{x-X_1}{b_n}\right)\right)^2 \\
&= b_n^d\left(\frac{1}{b_n^d}\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right) - b_n^{2d}\left(\frac{1}{b_n^d}\mathbb{E}K\left(\frac{x-X_1}{b_n}\right)\right)^2 \\
&= b_n^d\left[\frac{1}{b_n^d}\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right) - b_n^d\left(\frac{1}{b_n^d}\mathbb{E}K\left(\frac{x-X_1}{b_n}\right)\right)^2\right] \\
&= b_n^d A_n
\end{aligned}
$$

where

$$A_n \to f_\theta(x)\int_{\mathbb{R}^d} K^2(u)du.$$

Appendix :Proof of Theorem 1

Then

$$\begin{aligned}\mathbb{P}\left\{\frac{1}{nb_n^d}\left|\sum_{k=1}^{n} Y_k\right| > \epsilon\right\} &\leq 2C\exp\left(-\frac{\epsilon^2 nb_n^{2d}}{2\left(b_n^d A_n + \frac{2\epsilon N_1 b_n^d}{3}\right)}\right)\\ &\leq 2C\exp\left(-\frac{\epsilon^2 nb_n^{d}}{2\left(A_n + \frac{2\epsilon N_1}{3}\right)}\right).\end{aligned}$$

Therefore

$$\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x) \to 0 \text{ almost surely,} \tag{6}$$

by the Borel-Cantelli's lemma.
(5) and (6) imply that

$$\widehat{f}_n(x) \to f_\theta(x) \text{ almost surely.}$$

This achieves the proof of the theorem.

Appendix :Proof of Theorem 2

$$\begin{aligned}\sqrt{nb_n^d}\left(\widehat{f}_n(x)-f_\theta(x)\right) &= \sqrt{nb_n^d}\left(\widehat{f}_n(x)-\mathbb{E}\widehat{f}_n(x)\right)\\ &+ \sqrt{nb_n^d}\left(\mathbb{E}\widehat{f}_n(x)-f_\theta(x)\right)\\ &= F_n(x)+\sqrt{nb_n^d}\left(\mathbb{E}\widehat{f}_n(x)-f_\theta(x)\right).\end{aligned}$$

(1)
By making the change of variable $t=\frac{x-u}{b_n}$ and using assumptions (A_4) and (A_5), we get :

Appendix :Proof of Theorem 2

$$
\begin{aligned}
F_n(x) &= \sqrt{nb_n^d}\left\{\int_{\mathbb{R}^d}\frac{1}{b_n^d}K\left(\frac{x-u}{b_n}\right)f_\theta(u)du - f_\theta(x)\right\}\\
&= \sqrt{nb_n^d}\left\{\int_{\mathbb{R}^d}K(t)\left[f_\theta(x-tb_n)-f_\theta(x)\right]dt\right\}\\
&= \sqrt{nb_n^d}\left\{\int_{\mathbb{R}^d}K(t)\left[\frac{1}{2!}\sum_{i,j}^{d}\frac{\partial^2 f_\theta(x)}{\partial x_i\partial x_j}t_i t_j(-b_n)^2 + o(b_n^2)\right]dt\right\}\\
&= \sqrt{nb_n^{d+4}}\left\{\frac{1}{2}\int_{\mathbb{R}^d}\sum_{i=1}^{d}\frac{\partial^2 f_\theta(x)}{\partial x_i^2}t_i^2K(t)dt + o(1)\right\}\\
&= \sqrt{nb_n^{d+4}}\left\{\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 f_\theta(x)}{\partial x_i^2}\int_{\mathbb{R}^d}t_i^2K(t)dt + o(1)\right\}\\
&\to 0 \text{ as } n\to+\infty.
\end{aligned}
$$

Appendix :Proof of Theorem 2

(2)

$$\sqrt{nb_n^d}\left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right) = (nb_n^d)^{-1/2}\sum_{k=1}^{n} Y_k$$

where

$$Y_k = K\left(\frac{x - X_k}{b_n}\right) - \mathbb{E}K\left(\frac{x - X_k}{b_n}\right).$$

We have $\mathbb{E}(Y_k) = 0$ and $|Y_k| \leq 2N_1$.
Let $p = p(n)$, $q = q(n)$ and $r = r(n)$ be positive integers which tend to infinity as $n \to \infty$ such that $r(p+q) \leq n < r(p+q+1)$.
Define U_m and V_m by

$$U_m = \sum_{k=(m-1)(p+q)+1}^{(m-1)(p+q)+p} Y_k, \quad V_m = \sum_{k=(m-1)(p+q)+p+1}^{m(p+q)} Y_k, \quad m = 1, \ldots, r$$

and

Appendix :Proof of Theorem 2

$$V_{r+1} = \sum_{k=r(p+q)+1}^{n} Y_k.$$

We have

$$\sum_{k=1}^{n} Y_k = \sum_{m=1}^{r} U_m + \sum_{m=1}^{r+1} V_m.$$

Step 1 : We prove that $(nb_n^d)^{-1/2} \sum_{m=1}^{r+1} V_m \to 0$ in probability.

By Minkowski's inequality, we have

$$\left(\mathbb{E}\left|\sum_{m=1}^{r+1} V_m\right|^2\right)^{1/2} \leq \sum_{m=1}^{r} \left(\mathbb{E}|V_m|^2\right)^{1/2} + \left(\mathbb{E}|V_{r+1}|^2\right)^{1/2}$$

$$\leq r\left(\mathbb{E}|V_1|^2\right)^{1/2} + \left(\mathbb{E}|V_{r+1}|^2\right)^{1/2}$$

Appendix :Proof of Theorem 2

(i) Using Billingsley's inequality (see Bosq (1998)),

$$
\begin{aligned}
\mathbb{E}|V_1|^2 &= \mathbb{E}\left(\sum_{k=p+1}^{p+q} Y_k\right)^2 = q\mathbb{E}(Y_1^2) + 2\sum_{i<j}\mathbb{E}(Y_iY_j) \\
&\leq q\mathbb{E}(Y_1^2) + 2q\sum_{j=p+1}^{p+q-1}|\mathbb{E}(Y_{p+1}Y_{j+1})| \\
&\leq q\left(\mathbb{E}(Y_1^2) + 32N_1^2\sum_{j=p+1}^{p+q-1}\alpha(j-p)\right) \\
&\leq q\left(\mathbb{E}(Y_1^2) + 32N_1^2\sum_{j=1}^{\infty}\alpha(j)\right) = qC
\end{aligned}
$$

Appendix :Proof of Theorem 2

(ii)

$$\begin{aligned}\mathbb{E}|V_{r+1}|^2 &= \mathbb{E}\left(\sum_{k=r(p+q)+1}^{n} Y_k\right)^2\\ &= (n-r(p+q))\,\mathbb{E}(Y_1^2)+2\sum_{i<j}\mathbb{E}(Y_iY_j)\\ &\leq rC.\end{aligned}$$

Hence,

$$\left(\mathbb{E}\left|(nb_n^d)^{-1/2}\sum_{m=1}^{r+1}V_m\right|^2\right)^{1/2} \leq C(nb_n^d)^{-1/2}(r\sqrt{q}+\sqrt{r})$$

Appendix :Proof of Theorem 2

Therefore, choosing $q = q(n), r = r(n)$ and b_n such that

$$\frac{r\sqrt{q}}{\sqrt{nb_n^d}} \to 0 \text{ as } n \to \infty, \tag{7}$$

we get

$$\mathbb{E}\left|(nb_n^d)^{-1/2}\sum_{m=1}^{r+1} V_m\right|^2 \to 0.$$

which implies that

$$(nb_n^d)^{-1/2}\sum_{m=1}^{r+1} V_m \to 0 \text{ in probability.}$$

Appendix :Proof of Theorem 2

Step 2 : asymptotic normality of $(nb_n^d)^{-1/2}\sum_{m=1}^{r} U_m$.

$U_m,\ m = 1, \ldots, r$ have the same distribution ; so that

$$\prod_{m=1}^{r} \mathbb{E}\exp(itU_m) = \left(\mathbb{E}\exp(itU_1)\right)^r .$$

Let $\left|\mathbb{E}[\exp(it\sum_{m=1}^{r} U_m)] - [\mathbb{E}\exp(itU_1)]^r\right| = F.$

From lemma 4.2 (see Dharmenda and Masry (1996)), we have

$$\begin{aligned} F &= \left|\mathbb{E}[\exp(it\sum_{m=1}^{r} U_m)] - \prod_{m=1}^{r}\mathbb{E}\exp(itU_m)\right| \\ &= \left|\mathbb{E}(\prod_{m=1}^{r}\exp(itU_m)) - \prod_{m=1}^{r}\mathbb{E}\exp(itU_m)\right| \\ &\leq 4(r-1)\alpha(1+q) \leq 4r\alpha(q). \end{aligned}$$

Appendix :Proof of Theorem 2

Setting $\phi_1(t) = \mathbb{E}\exp(itU_1)$. If $q = q(n)$ and $r = r(n)$ are chosen such that

$$r\alpha(q) \to 0 \text{ as } n \to \infty, \tag{8}$$

the charasteristic function of $(nb_n^d)^{-1/2} \sum_{m=1}^{r} U_m$ is $\phi_1^r\left(t(nb_n^d)^{-1/2}\right)$ which is the charasteristic function of $\sum_{m=1}^{r} Z_m$ where Z_m, $m = 1, ..., r$ are independent random variables with distribution that of $(nb_n^d)^{-1/2} U_1$.
We have $\mathbb{E}(Z_m) = 0$ and

Appendix :Proof of Theorem 2

$$
\begin{aligned}
\mathbb{E}\left(\sum_{m=1}^{r} Z_m\right)^2 &= r\mathbb{E}\left(Z_1^2\right) \\
&= (nb_n^d)^{-1} r\mathbb{E}\left(U_1^2\right) \\
&= (nb_n^d)^{-1} r\left[p\mathbb{E}(Y_1^2) + 2\sum_{i<j}^{p}\mathbb{E}(Y_iY_j)\right] \\
&= \frac{rp}{nb_n^d}\mathbb{E}(Y_1^2) + \frac{2r}{nb_n^d}\sum_{i<j}^{p}\mathbb{E}(Y_iY_j).
\end{aligned}
$$

Appendix :Proof of Theorem 2

(i)

$$\frac{rp}{nb_n^d}\mathbb{E}(Y_1)^2 = \frac{rp}{n}\left\{\frac{1}{b_n^d}\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right) - b_n^d\left(\frac{1}{b_n^d}\mathbb{E}K\left(\frac{x-X_1}{b_n}\right)\right)^2\right\}$$
$$\to f_\theta(x)\int_{\mathbb{R}^d} K^2(u)du.$$

(ii) Note that $\alpha(k) \leq \exp(-\lambda k) = \varphi(k)$ with $\lambda > 0$.

$$\sum_{i<j} |\mathbb{E}(Y_i Y_j)| \leq \sum_{j=1}^{p-1} j|\mathbb{E}(Y_1 Y_{j+1})|$$
$$\leq 16N_1^2 \sum_{j=1}^{p-1} j\alpha(j)$$

Then,

Appendix :Proof of Theorem 2

$$\begin{aligned}
\frac{2r}{nb_n^d}\sum_{i<j}|\mathbb{E}(Y_iY_j)| &\leq \frac{32N_1^2 r}{nb_n^d}\sum_{k=1}^{p-1}\sum_{l=k}^{p-1}\alpha(l) \\
&\leq \frac{32N_1^2 r}{nb_n^d}\sum_{k=1}^{p-1}\sum_{l=k}^{p-1}(\varphi(l))^{1/2}(\varphi(k))^{1/2} \\
&\leq \frac{32N_1^2 r}{nb_n^d}\sum_{k=1}^{\infty}(\varphi(k))^{1/2}\sum_{l=1}^{\infty}(\varphi(l))^{1/2} \\
&\leq \frac{32N_1^2 r}{nb_n^d}\left(\sum_{i=1}^{\infty}(\varphi(i))^{1/2}\right)^2 \to 0 \text{ if}
\end{aligned}$$

$$\frac{r}{nb_n^d} \to 0 \text{ as } n \to \infty. \quad (9)$$

Appendix :Proof of Theorem 2

Therefore

$$\mathbb{E}\left(\sum_{m=1}^{r} Z_m\right)^2 \to f_\theta(x)\int_{\mathbb{R}^d} K^2(u)du \ \text{ as } \ n\to\infty.$$

Since the random variables $U_m(m=1,...,r)$ have the same distribution, then by Lyapunov's theorem (see Dominique and Aimé (1998)), the limiting distribution of $(nb_n^d)^{-1/2}\sum_{m=1}^{r} U_m$ is $N(0,\tau^2(x))$ where

$$\tau^2(x) = f_\theta(x)\int_{\mathbb{R}^d} K^2(u)du.$$

Appendix :Proof of Theorem 2

The conditions (7), (8) and (9) are satisfied, for example, with

$$r(n) \sim \log(n), p(n) \sim \frac{n}{\log(n)} - n^{1/4}, q(n) \sim n^{1/4}$$

and

$$b_n^d = \frac{\log(n)}{n^{\lambda}} \text{ with } 0 < \lambda < \frac{3}{4}.$$

This achieves the proof of the theorem.

Asymptotic normality : Proof of lemma 1

Step 1 : we prove that

$$Y_n = \sqrt{n}\left\{\int_{\mathbb{R}^d} h_{\theta_0}(x)\widehat{f}_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\} \to 0 \text{ in probability.}$$

$$\begin{aligned}
\mathbb{E}|Y_n| &= \sqrt{n}\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d} h_{\theta_0}(x)\frac{1}{b_n^d}K\left(\frac{x-y}{b_n}\right)dx - h_{\theta_0}(y)\right| f_{\theta_0}(y)dy \\
&= \sqrt{n}\int_{E_n}\left|\int_{\mathbb{R}^d} h_{\theta_0}(y+ub_n)K(u)du - h_{\theta_0}(y)\right| f_{\theta_0}(y)dy \\
&+ \sqrt{n}\int_{E_n^c}\left|\int_{\mathbb{R}^d} h_{\theta_0}(y+ub_n)K(u)du - h_{\theta_0}(y)\right| f_{\theta_0}(y)dy \\
&= I_1 + I_2.
\end{aligned}$$

With assumptions (A_4) and (A_5), we have

Asymptotic normality : Proof of lemma 1

$$
\begin{aligned}
I_1 &= \sqrt{n}\int_{E_n}\left|\int_{\mathbb{R}^d}\left[h_{\theta_0}(y+ub_n)-h_{\theta_0}(y)\right]K(u)du\right| f_{\theta_0}(y)dy \\
&= \sqrt{n}\int_{E_n}\left|\int_{\mathbb{R}^d}\left[\frac{1}{2!}\sum_{i,j}^{d}\frac{\partial^2 h_{\theta_0}(y)}{\partial y_i \partial y_j}u_i u_j(b_n^2)+o(b_n^2)\right]K(u)du\right| f_{\theta_0}(y)dy \\
&= \sqrt{n}b_n^2\int_{E_n}\left|\int_{\mathbb{R}^d}\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 h_{\theta_0}(y)}{\partial y_i^2}u_i^2 K(u)du+o(1)\right| f_{\theta_0}(y)dy \\
&\leq \sqrt{n}b_n^2\int_{E_n}\left|\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 h_{\theta_0}(y)}{\partial y_i^2}\int_{\mathbb{R}^d}u_i^2 K(u)du+o(1)\right| f_{\theta_0}(y)dy \\
&\leq \sqrt{n}b_n^2\int_{E_n}f_{\theta_0}(y)\left\{\frac{1}{2}\sum_{i=1}^{d}\left|\frac{\partial^2 h_{\theta_0}(y)}{\partial y_i^2}\right|\left(\int_{\mathbb{R}^d}u_i^2 K(u)du\right)+o(1)\right\}dy.
\end{aligned}
$$

Asymptotic normality : Proof of lemma 1

So $I_1 \to 0$ as $n \to +\infty$. Furthermore,

$$
\begin{aligned}
I_2 &= \sqrt{n}\int_{E_n^c}\left|\int_{\mathbb{R}^d} h_{\theta_0}(y+ub_n)K(u)du - h_{\theta_0}(y)\right| f_{\theta_0}(y)dy \\
&\leq \sqrt{n}\int_{E_n^c}\left(\int_{\mathbb{R}^d} |h_{\theta_0}(y+ub_n)K(u)du|\right) f_{\theta_0}(y)dy \\
&+ \sqrt{n}\int_{E_n^c} |h_{\theta_0}(y)| f_{\theta_0}(y)dy \\
&\leq \sqrt{n}\int_{E_n^c}\left(\int_{\mathbb{R}^d} |h_{\theta_0}(y+ub_n)|K(u)du\right) f_{\theta_0}(y)dy \\
&+ \sqrt{n}\int_{E_n^c} |\dot{g}_{\theta_0}(y)| f_{\theta_0}^{1/2}(y)dy \\
&\to 0 \ \text{ as } \ n \to +\infty.
\end{aligned}
$$

Asymptotic normality : Proof of lemma 1

Therefore

$$Y_n = \sqrt{n}\left\{\int_{\mathbb{R}^d} h_{\theta_0}(x)\widehat{f}_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\} = I_1 + I_2 \xrightarrow{\mathbb{P}} 0.$$

Step 2 : asymptotic normality of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{\theta_0}(X_i)$, $\theta_0 \in \mathbb{R}^s$, $s \geq 1$

(i) $\theta_0 \in \mathbb{R}$

Proof is similar to that of theorem 2 ; we use the inequality of Davidov (see Bosq (1998)) instead of that of Billingsley. Note that :

$$\begin{aligned}
\mathbb{E}(h_{\theta_0}(X_1)) &= \int_{\mathbb{R}^d} h_{\theta_0}(x) f_{\theta_0}(x)dx \\
&= \frac{1}{2}\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x) f_{\theta_0}^{1/2}(x)dx \\
&= \frac{1}{4}\int_{\mathbb{R}^d} 2\dot{g}_{\theta_0}(x) g_{\theta_0}(x)dx = 0.
\end{aligned}$$

Asymptotic normality : Proof of lemma 1

and

$$\begin{aligned}\mathbb{E}(h^2_{\theta_0}(X_1)) &= \int_{\mathbb{R}^d} h^2_{\theta_0}(x) f_{\theta_0}(x)dx \\ &= \frac{1}{4}\int_{\mathbb{R}^d} \dot{g}^2_{\theta_0}(x)dx \\ &= \frac{1}{4}\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}^t_{\theta_0}(x)dx.\end{aligned}$$

(ii) $\theta_0 \in \mathbb{R}^s$, $s > 1$

Recall that $X_n \xrightarrow{\mathcal{L}} X$ if and only if $u^t X_n \xrightarrow{\mathcal{L}} u^t X$ for all $u \in \mathbb{R}^s$.

Asymptotic normality : Proof of lemma 1

Let $u \in \mathbb{R}^s$, $Y_i = h_{\theta_0}(X_i) = \frac{\dot{g}_{\theta_0}}{2f_{\theta_0}^{1/2}}(X_i)$ and $T_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$, the real random variables $(u^t Y_i, i \geq 1)$ are strongly mixing with mean zero and variance $u^t \Gamma u$ where Γ is the covariance matrix of Y_1 ; $\Gamma = \mathbb{E}(Y_1 Y_1^t)$.

From (i), $u^t T_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} u^t Y_i \stackrel{\mathcal{L}}{\to} N(0, u^t \Gamma u)$.

Therefore,

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n} h_{\theta_0}(X_k) \stackrel{\mathcal{L}}{\to} N(0,\Gamma) \ \text{ where } \ \Gamma = \frac{1}{4}\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx.$$

This completes the proof of the lemma.

Asymptotic normality : Proof of lemma 2

$$
\begin{aligned}
R_n &= \int_{\mathbb{R}^d} \sqrt{n} h_{\theta_0}(x) \left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x) \right)^2 dx \\
&= \int_{G_n} \sqrt{n} h_{\theta_0}(x) \left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x) \right)^2 dx \\
&+ \int_{G_n^c} \sqrt{n} h_{\theta_0}(x) \left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x) \right)^2 dx \\
&= \int_{G_n} \sqrt{n} h_{\theta_0}(x) \frac{\left(\widehat{f}_n(x) - f_{\theta_0}(x) \right)^2}{\left(\widehat{f}_n^{1/2}(x) + f_{\theta_0}^{1/2}(x) \right)^2} dx \\
&+ \int_{G_n^c} \sqrt{n} h_{\theta_0}(x) \left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x) \right)^2 dx \\
&= R_{n1} + R_{n2}.
\end{aligned}
$$

Asymptotic normality : Proof of lemma 2

$$
\begin{aligned}
|R_{n1}| &\leq \int_{G_n} \sqrt{n}|h_{\theta_0}(x)| \frac{\left(\widehat{f}_n(x) - f_{\theta_0}(x)\right)^2}{\left(\widehat{f}_n^{1/2}(x) + f_{\theta_0}^{1/2}(x)\right)^2} dx \\
&\leq \int_{G_n} \sqrt{n}|h_{\theta_0}(x)| \frac{\left(\widehat{f}_n(x) - f_{\theta_0}(x)\right)^2}{f_{\theta_0}(x)} dx.
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathbb{E}(|R_{n1}|) &\leq \int_{G_n} |h_{\theta_0}(x)| f_{\theta_0}^{-1}(x) \sqrt{n} \mathbb{E}\left(\widehat{f}_n(x) - f_{\theta_0}(x)\right)^2 dx \\
&\leq \int_{G_n} |h_{\theta_0}(x)| f_{\theta_0}^{-1}(x) \sqrt{n} \left[\mathbb{E}\left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right)^2\right] dx \\
&+ \int_{G_n} |h_{\theta_0}(x)| f_{\theta_0}^{-1}(x) \sqrt{n} \left[\left(\mathbb{E}\widehat{f}_n(x) - f_{\theta_0}(x)\right)^2\right] dx
\end{aligned}
$$

Asymptotic normality : Proof of lemma 2

(i) Denote $B_n(x) = \sqrt{n}\mathbb{E}\left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right)^2$, we have

$$
\begin{aligned}
B_n(x) &= \frac{\sqrt{n}}{(nb_n^d)^2}\mathbb{E}\left(\sum_{i=1}^{n} Y_i\right)^2 \quad \text{with} \quad Y_i = K(\frac{x-X_i}{b_n}) - \mathbb{E}K(\frac{x-X_i}{b_n}) \\
&= \frac{\sqrt{n}}{(nb_n^d)^2}\left[n\mathbb{E}(Y_1^2) + 2\sum_{i<j}^{n}\mathbb{E}(Y_iY_j)\right] \\
&\leq \frac{1}{n^{3/2}b_n^{2d}}\left[n\mathbb{E}(Y_1^2) + 2\sum_{i<j}^{n}|\mathbb{E}(Y_iY_j)|\right] \\
&\leq \frac{1}{n^{3/2}b_n^{2d}}\left[n\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right) + 2\sum_{j=1}^{n-1}j|\mathbb{E}(Y_1Y_{j+1})|\right]
\end{aligned}
$$

Asymptotic normality : Proof of lemma 2

Using Davidov's inequality for mixing processes, we get

$$\begin{aligned}\sum_{j=1}^{n-1} j|\mathbb{E}(Y_1 Y_{j+1})| &\leq \sum_{j=1}^{n-1} j\left[2p\,(2\alpha(j))^{1/p}\,(\mathbb{E}|Y_1|^q)^{1/q}\,(\mathbb{E}|Y_{j+1}|^r)^{1/r}\right] \\ &\leq 2p\,(\mathbb{E}|Y_1|^q)^{1/q}\,(\mathbb{E}|Y_1|^r)^{1/r}\sum_{j=1}^{n-1} j\,(2\alpha(j))^{1/p}\end{aligned}$$

Let $D_n(x) = \sum_{j=1}^{n-1} j|\mathbb{E}(Y_1 Y_{j+1})|$ and $\varphi(j) = (\alpha(j))^{1/p}$.
Choose $q \geq 2$ and $r \geq 2$, we obtain

Asymptotic normality : Proof of lemma 2

$$
\begin{aligned}
D_n(x) &\leq 2p(\mathbb{E}|Y_1|^2|Y_1|^{q-2})^{1/q}(\mathbb{E}|Y_1|^2|Y_1|^{r-2})^{1/r}\sum_{j=1}^{n-1} j(2\alpha(j))^{1/p} \\
&\leq C(\mathbb{E}|Y_1|^2)^{1/q+1/r}\sum_{j=1}^{n-1} j(\alpha(j))^{1/p} \\
&\leq C\left[\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right]^{1/q+1/r}\sum_{j=1}^{n-1} j(\varphi(j)) \\
&\leq C\left[\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right]^{1/q+1/r}\left(\sum_{j=1}^{n-1} \varphi^{1/2}(i)\right)^2 \\
&\leq C_1\left[\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right]^{1/q+1/r}.
\end{aligned}
$$

Asymptotic normality : Proof of lemma 2

Hence,

$$\begin{aligned}
B_n(x) &\leq \frac{1}{n^{3/2}b_n^{2d}}\left[n\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right] \\
&+ \frac{C_1}{n^{3/2}b_n^{2d}}\left[\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right]^{1/q+1/r} \\
&\leq \frac{1}{n^{1/2}b_n^{d}}\left[\frac{1}{b_n^d}\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right] \\
&+ \frac{C_1}{n^{3/2}b_n^{d+d/p}}\left[\frac{1}{b_n^d}\mathbb{E}K^2\left(\frac{x-X_1}{b_n}\right)\right]^{1/q+1/r} \\
&\leq \frac{1}{n^{1/2}b_n^{d}}\int_{\mathbb{R}^d} K^2(t)f_{\theta_0}(x-tb_n)dt \\
&+ \frac{C_1}{n^{3/2}b_n^{d+d/p}}\left[\int_{\mathbb{R}^d} K^2(t)f_{\theta_0}(x-tb_n)dt\right]^{1/q+1/r}.
\end{aligned}$$

Asymptotic normality : Proof of lemma 2

(ii) Let $F_n(x) = \sqrt{n}\left(\mathbb{E}\widehat{f}_n(x) - f_{\theta_0}(x)\right)^2$

$$\begin{aligned}
F_n(x) &= \sqrt{n}\left\{\int_{\mathbb{R}^d} K(t)\left[\frac{1}{2!}\sum_{i,j}^{d}\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i \partial x_j} t_i t_j (-b_n)^2 + o(b_n^2)\right] dt\right\}^2 \\
&= n^{1/2} b_n^4 \left\{\frac{1}{2}\int_{\mathbb{R}^d}\sum_{i=1}^{d}\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i^2} t_i^2 K(t) dt + o(1)\right\}^2 \\
&= n^{1/2} b_n^4 \left\{\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i^2}\int_{\mathbb{R}^d} t_i^2 K(t) dt + o(1)\right\}^2 \\
&\leq 2 d n^{1/2} b_n^4 \left\{\frac{1}{4}\sum_{i=1}^{d}\left(\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i^2}\right)^2 \left(\int_{\mathbb{R}^d} t_i^2 K(t) dt\right)^2 + o(1)\right\}.
\end{aligned}$$

Asymptotic normality : Proof of lemma 2

Therefore,
$\mathbb{E}(|R_{n1}|)$

$$\begin{aligned}
&\leq \frac{1}{n^{1/2}b_n^d}\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left(\int_{\mathbb{R}^d}K^2(t)f_{\theta_0}(x-tb_n)dt\right)dx \\
&+ \frac{C_1}{n^{3/2}b_n^{d+d/p}}\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left(\left[\int_{G_n}K^2(t)f_{\theta_0}(x-tb_n)dt\right]^{1/q+1/r}\right. \\
&+ 2dn^{1/2}b_n^4\int_{G_n}|h_{\theta_0}(x)|f_{\theta_0}^{-1}(x)\left\{\frac{1}{4}\sum_{i=1}^{d}\left(\frac{\partial^2 f_{\theta_0}(x)}{\partial x_i^2}\right)^2\left(\int_{\mathbb{R}^d}t_i^2K(t)dt\right)\right. \\
&\to 0 \text{ as } n\to+\infty.
\end{aligned}$$

The last relation implies that

$$R_{n1}\to 0 \text{ in probability as } n\to+\infty. \qquad (10)$$

Asymptotic normality : Proof of lemma 2

Furthermore,

$$
\begin{aligned}
|R_{n2}| &\leq \sqrt{n}\int_{G_n^c} |h_{\theta_0}(x)| \left(\widehat{f}_n^{1/2}(x) - f_{\theta_0}^{1/2}(x)\right)^2 dx \\
&\leq 2\sqrt{n}\int_{G_n^c} |h_{\theta_0}(x)| \left(\widehat{f}_n(x) + f_{\theta_0}(x)\right) dx \\
&\leq 2\sqrt{n}\int_{G_n^c} |h_{\theta_0}(x)| f_{\theta_0}(x) dx + 2\sqrt{n}\int_{G_n^c} |h_{\theta_0}(x)|\widehat{f}_n(x) dx \\
&\leq 2\sqrt{n}\int_{G_n^c} |\dot{g}_{\theta_0}(x)| f_{\theta_0}^{1/2}(x) dx + 2\sqrt{n}\int_{G_n^c} |h_{\theta_0}(x)|\widehat{f}_n(x) dx \\
&\leq 2\sqrt{n}\int_{G_n^c} |\dot{g}_{\theta_0}(x)| f_{\theta_0}^{1/2}(x) dx + R_{n22}.
\end{aligned}
$$

Asymptotic normality : Proof of lemma 2

We have,

$$\begin{aligned}
\mathbb{E}(R_{n22}) &= 2\sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\mathbb{E}(\widehat{f}_n(x))dx \\
&= 2\sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\left\{\int_{\mathbb{R}^d}\frac{1}{nb_n^d}\sum_{i=1}^{n}K\left(\frac{x-y}{b_n}\right)f_{\theta_0}(y)dy\right\}dx \\
&= 2\sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\left\{\int_{\mathbb{R}^d}\frac{1}{b_n^d}K\left(\frac{x-y}{b_n}\right)f_{\theta_0}(y)dy\right\}dx \\
&= 2\sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\left\{\int_{\mathbb{R}^d}K(u)f_{\theta_0}(x+ub_n)du\right\}dx
\end{aligned}$$

Asymptotic normality : Proof of lemma 2

Therefore, if

$$\sqrt{n}\int_{G_n^c}|\dot{g}_{\theta_0}(x)|f_{\theta_0}^{1/2}(x)dx\to 0$$

and

$$\sqrt{n}\int_{G_n^c}|h_{\theta_0}(x)|\left\{\int_{\mathbb{R}^d}K(u)f_{\theta_0}(x+ub_n)du\right\}dx\to 0$$

then

$$R_{n2}\to 0 \;\text{ in probability as }\; n\to+\infty. \tag{11}$$

(10) and (11) imply that

$$R_n=\int_{\mathbb{R}^d}\sqrt{n}h_{\theta_0}(x)\left(\widehat{f}_n^{1/2}(x)-f_{\theta_0}^{1/2}(x)\right)^2dx\to 0$$

in probability as $n\to+\infty$.
This completes the proof of the lemma.