International Journal of Statistics and Probability; Vol. 3, No. 1; 2014
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Kernel Gradient of Density of Probability Estimate From Contaminated Associated Observations

K. Djeddour[1] & S. Meddahi[2]

[1] Labo MSTD Faculty of Mathematics, Univ. of Sci. and Tech. of Algiers, Algeria

[2] Faculty of Sciences, Univ. of Boumerdes, Algeria

Correspondence: K. Djeddour, Labo MSTD Faculty of Mathematics, Univ. of Sci. and Tech. of Algiers, Algeria. E-mail: khdjeddour@hotmail.com

Received: November 10, 2013 Accepted: December 18, 2013 Online Published: December 23, 2013

doi:10.5539/ijsp.v3n1p24 URL: http://dx.doi.org/10.5539/ijsp.v3n1p24

Abstract

We consider the estimation of gradient of density function of positive associated random process $(X_i)_i$ from noisy observations. We establish asymptotic expressions for the variance of the gradient of estimator of density of probability. We consider the case of algebraic decay of the tail of the noise characteristic function of ϵ_i.

Keywords: kernel estimator, gradient, associated observations, deconvolution

AMS Subject classification 1991: 62G20

1. Introduction

We consider the problem of estimating the gradient of multivariate probability densities of a stationary process using observations that are corrupted by additive noise. For each integer $p \geq 1$, we consider that there exits a joint probability density $f(x) = f(x_1, ..., x_p)$ for $X_1, ..., X_p$, where $(X_i)_i$ is a real-valued stationary process. Consider the deconvolution problem

$$Y_i = X_i + \epsilon_i, \quad i = 1, 2....$$

Such a model of measurements being contaminated by errors arises in many fields where the measurements cannot be observed directly. The noise process $(\epsilon_i)_i$ consists of independent and identically distributed random variables, and assume furthermore that it is independent of the process $(X_i)_i$, with known marginal density $\overline{h}(x)$. Let $g(x)$ be the joint probability density function of the random variables $Y_1, ..., Y_p$ which is given by

$$g(x) = \int_{\mathbb{R}^p} f(x-u)h(u)du,$$

where $h(u) = \Pi_{j=1}^{p} \overline{h}(u_j)$ and $u = (u_1, ..., u_p)$.

We consider the gradient $\nabla f(x)$ of multivariate density deconvolution when the process $(X_i)_i$ is associated. Our aim is to study the estimate of $\nabla f(x)$ from the noisy observations $(Y_i)_{i=1}^{n}$. This is clearly a multidimensional density deconvolution problem for dependent data.

In Chacon, Duong and Wand (2011), the authors investigate kernel estimators of multivariate density derivative functions using general bandwidth matrix selectors. Given a random sample $X_1, X_2, ..., X_n$ drawn from the same density of probability f, and provide the results for mean integrated square convergence both asymptotically and for finite samples. The influence of the bandwidth matrix on convergence is established.

The deconvolution problem for the estimation of $f(x)$ has been investigated by many authors. We cite the work of Fan (1991), Masry (2001), among others. Most of papers cited above address how to estimate the unknown density and compute the rate of convergence for specific error process. Fan (1991) used kernel density estimator to estimate the unknown density f, as well as its derivatives, for the case of i.i.d observations and $p = 1$. Masry (2003) developed an estimation of the multivariate probability density when the underlying process $(X_i)_i$ is associated $(p \geq 1)$.

We recall the definition of association for collections of random variables.

Definition 1 The sequence $(X_n)_{n\in\mathbb{Z}}$ is said to be positively associated if for every finite subcollection $(X_{i_1},...,X_{i_n})$ and every pair of coordinate-wise non decreasing functions H_1, H_2:

$$cov(H_1(X_{i_1},...,X_{i_n}); H_2(X_{i_1},...,X_{i_n})) \geq 0$$

whenever the covariance is defined.

This definition was introduced by Esary, Proschan, and Walkup (1967). Furthermore, positive association seems to be a natural assumption to model certain clinical trials as those described in Ying and Wei (1994). It is also known, see Pitt (1982), that Gaussian processes are positively associated, if and only if, their covariance function is positive. We note that an important property of associated random variables is that non correlation implies independence; the only alternative frame for this to hold is the Gaussian one. This means that one may hope that dependence will appear in this case only through the covariance structure, and also justifies the study of such processes. Indeed, a covariance is much easier to compute than a mixing coefficient. Unfortunately, a main inconvenience of mixing is that there are only few mixing models for which the mixing coefficients can be explicitly evaluated. We note that association and mixing define two distinct but not disjoint classes of processes.

Example 2 (see Louhichi, 2000) Let (ϵ_i) be a sequence of i.i.d random variables and $(\mu_i)_{i\in\mathbb{Z}}$ a sequence of numbers. Let $X_j^n = \sum_{|i|\leq n}\mu_j\epsilon_{j-i}$; and assume that, there exists X_j such that $\lim_{n\to\infty} X_j^n = X_j$ a.s; $\sup_j E|X_j^n| < \infty$ and $|X_j| \leq \infty$ a.s. The linear process is $X_j = \sum_{i\in\mathbb{Z}}\mu_j\epsilon_{j-i}$. If the sequence $(\mu_i)_{i\in\mathbb{Z}}$ is non negative and $\sum|\mu_j| < \infty$ then $(X_j)_{j\in\mathbb{Z}}$ is associated.

2. Notations and Assumptions

2.1 Notations

We denote the characteristic functions of f, g, h and $\overline{h}$ by ϕ_f, ϕ_g, ϕ_h and $\overline{\phi}_h$ respectively. Then $\phi_g(t) = \phi_f(t)\phi_h(t)$ and $\phi_h(t) = \Pi_{j=1}^p\overline{\phi}_h(t_j)$, where $t = (t_1,...,t_p)$.

Let us consider $\hat{g}_n(x)$ a kernel-type estimate of $g(x)$ that is

$$\hat{g}_n(x) = \frac{1}{h_n}\sum_{i=1}^n K\left(\frac{x-Y_i}{h_n}\right).$$

Let $K(x) = \Pi_{j=1}^p\overline{K}(x_j)$, where $\overline{K}(x)$ be a real-valued, even, bounded density function on the real line satisfying $\overline{K}(x) = O(|x|^{-1-\delta})$ for some $\delta > 0$ and denote its Fourier transform by $\overline{\phi}_K(t)$. Assumptions will be made on $\overline{\phi}_K(t)$ and $\overline{\phi}_h(t)$ which will ensure that $\dfrac{\overline{\phi}_K(t)}{\overline{\phi}_h(t/h)} \in L_1 \cap L_\infty$, where L_1 is the space of Lebesgue integrable functions and L_∞ the space of bounded functions.

For every $h_n > 0$ define the deconvolution kernel

$$\overline{W}_{h_n}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}\frac{\overline{\phi}_K(t)}{\overline{\phi}_h(t/h_n)}dt.$$

Set $W_{h_n}(x) = \Pi_{j=1}^p\overline{W}_{h_n}(x_j)$ where $x = (x_1,...,x_p)$. So that $\phi_K(t) = \Pi_{j=1}^p\overline{\phi}_K(t_j)$. The choice of product type kernel is not essential and is made for sake of simplicity.

The kernel density estimator for estimating the unknown density of X is defined as follows:

Let $(h_n)_{n\geq 1}$ be a sequence of positive numbers such that $h_n \to 0$ as $n \to \infty$; given the observations $(Y_i)_{i=1}^n$, the estimate of $f(x)$ is defined by

$$\widehat{f_n}(x) = \frac{1}{(n-p)h_n^p}\sum_{j=0}^{n-p} W_{h_n}\left(\frac{x-\mathbf{Y}_j}{h_n}\right),$$

where $\mathbf{Y}_j = (Y_{j+1},...,Y_{j+p})$ and it is assumed that $n > p$.

Another expression of $\widehat{f_n}(x)$ is

$$\widehat{f_n}(x) = \frac{1}{(2\pi)^p}\int_{\mathbb{R}^p} e^{-itx}\widehat{\phi_n}(t)\frac{\phi_K(h_nt)}{\phi_h(t)}dt, \quad (1)$$

where $\widehat{\phi}_n(t)$ is the standard estimate of the characteristic function of $\phi_g(t)$:

$$\widehat{\phi}_n(t) = \frac{1}{n-p}\sum_{j=0}^{n-p} e^{it\mathbf{Y}_j}.$$

The difficulty of deconvolution depends on the smoothness of the distribution of the error variable ϵ. By smoothness of the error distribution, we mean the order of characteristic function $\overline{\phi}_h(t)$ of (ϵ_i) as $t \to \infty$. We say that the distribution of ϵ is algebraically decreasing or ordinary smooth of order β if $\overline{\phi}_h(t)$ satisfies:

$$a_0|t|^{-\beta} \leq \overline{\phi}_h(t) \leq a_1|t|^{-\beta} \text{ as } t \to \infty,$$

where a_0, a_1, β are positive real numbers.

2.2 Assumptions

(A1) the probability density $g(x)$ exists.

(A2) The p-dimensional density $g(u, v)$ of the vectors $\mathbf{Y}_0$ and $\mathbf{Y}_l$ for $l \geq p$ exists.

(A3) The process is associated and its covariance function $c_j = cov(X_{j+1}, X_1)$ satisfies $\sum_{j=1}^{\infty} j^{\delta} c_j \leq \infty$ for some $\delta > 1 + \frac{2}{p}$.

We need some lemmas for the proofs of the main results.

3. Some Auxiliary Lemmas

The result from real analysis that is needed here is the following (see for instance Wheeden & Zygmund, 1977, p. 189):

Lemma 3 *Assume that $Q(x)$ is a bounded integrable function on $\mathbb{R}^p$. Let $f \in L_1(\mathbb{R}^p)$ and continuous. Then for almost all $x \in \mathbb{R}^p$ we have*

$$\lim_{n\to\infty}\int_{\mathbb{R}^p}\left[\frac{1}{h_n^p}Q\left(\frac{x-u}{h_n}\right)\right]f(u)du = f(x)\int_{\mathbb{R}^p}Q(u)\,du.$$

Lemma 4 *Assume that $\overline{\phi}_h(t)$ and $\overline{\phi}_K(t)$ satisfies*

i) $|\overline{\phi}_h(t)| > 0$ *for all* $t \in \mathbb{R}$,

ii) $|t|^{\beta}|\overline{\phi}_h(t)| \to |B_1|$ *as* $|t| \to \infty$ *for some* B_1 *and* $\beta > 0$,

iii) $D_1 = \frac{1}{2\pi|B_1|^2}\int_{-\infty}^{\infty}|t|^{2\beta}|\overline{\phi}_K(t)|^2dt < \infty$, $D_2 = \frac{1}{2\pi|B_1|^2}\int_{-\infty}^{\infty}|t|^{2(\beta+1)}|\overline{\phi}_K(t)|^2dt < \infty$ *then* $h^{2\beta}\int_{-\infty}^{\infty}|\overline{W}_{h_n}(x)|^2dx \to D_1$ *and* $h^{2\beta}\int_{-\infty}^{\infty}|\overline{W}'_{h_n}(t)|^2dt \to D_2$ *as* $h_n \to 0$. *Thus*

$$h_n^{2\beta p}\int_{\mathbb{R}^p}\left|g_{n,k}(x)\right|^2 dx \to D_1^{p-1}D_2.$$

Proof. First we have $h^{2\beta}\int_{-\infty}^{\infty}|\overline{W}_{h_n}(x)|^2dx = \frac{h^{2\beta}}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}_K(t)}{\overline{\phi}_h(t/h)}\right|^2dt$. Hence, by ii) $|t|^{\beta}|\overline{\phi}_h(t)| \geq \frac{|B_1|}{2}$. Thus

$$\frac{h^{2\beta}}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}_K(t)}{\overline{\phi}_h(t/h)}\right|^2dt \leq \frac{4}{2\pi|B_1|^2}\int_{-\infty}^{\infty}|t|^{2\beta}|\overline{\phi}_K(t)|^2dt = D_1$$

and $h^{2\beta}\int_{-\infty}^{\infty}|\overline{W}'_{h_n}(t)|^2dt = \frac{h^{2\beta}}{2\pi}\int_{-\infty}^{\infty}\left|\frac{t\overline{\phi}_K(t)}{\overline{\phi}_h(t/h)}\right|^2dt$. Hence by ii)

$$\frac{h^{2\beta}}{2\pi}\int_{-\infty}^{\infty}\left|\frac{t\overline{\phi}_K(t)}{\overline{\phi}_h(t/h)}\right|^2dt \leq \frac{4}{2\pi|B_1|^2}\int_{-\infty}^{\infty}|t|^{2(\beta+1)}|\overline{\phi}_K(t)|^2dt = D_2.$$

□

Lemma 5 *Assume that $\overline{\phi}_h(t)$ and $\overline{\phi}_K(t)$ are twice continuously differentiable with bounded derivatives such that*

i) $|\overline{\phi}_h(t)| > 0$ *for all* $t \in \mathbb{R}$,

ii) $|t|^{\beta}\overline{\phi}_h(t) \to B_1$ *as* $|t| \to \infty$ *for some* B_1 *and* $\beta \geq 1$,

iii) $\int_{-\infty}^{\infty} |t|^{\beta-2}|\overline{\phi}_K(t)|^2 dt < \infty$, $\int_{-\infty}^{\infty} |t|^{\beta-1}|\overline{\phi}'_K(t)|^2 dt < \infty$, $\int_{-\infty}^{\infty} |t|^{\beta}|\overline{\phi}''_K(t)|^2 dt < \infty$. *Then* $h_n^{\beta}|\overline{W}_{h_n}(x)| \leq C_1 \leq \infty$, *where* C_1 *is a constant independent of* h_n, *and* $h_n^{\beta}\left\|\overline{W}_{h_n}(x)\right\|_1 \leq const$.

Proof. By setting $F_{h_n}(t) = \frac{\overline{\phi}_K(t)}{\overline{\phi}_h(t/h_n)}$; we have $F_{h_n}(t) \in L_1$ and is twice continuously differentiable with bounded derivatives; $F''_{h_n}(t) \in L_1$ then by integration by part one has

$$(ix)^2\overline{W}_{h_n}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}F''_{h_n}(t)dt$$

and by the Riemann-Lebesgue lemma $\overline{W}_{h_n}(x) = o(x^{-2})$, so that $\overline{W}_{h_n}(x) \in L_1$. Under such smoothness conditions on $\overline{\phi}_h$ and $\overline{\phi}_K$ we will show that in fact $h_n^{\beta}\int_{-\infty}^{\infty}|F''_{h_n}(t)|dt \leq C_1 < \infty$. From which we obtain a bound for its L_1-norm $||\overline{W}_{h_n}||_1$

$$F''_{h_n}(t) = \frac{\overline{\phi}''_K(t)}{\overline{\phi}_h(t/h_n)} - \frac{2}{h_n}\frac{\overline{\phi}'_h(t/h_n)\overline{\phi}'_K(t)}{\left(\overline{\phi}_h(t/h_n)\right)^2} - \frac{1}{h_n^2}\frac{\overline{\phi}_K(t)\overline{\phi}''_h(t/h_n)}{\overline{\phi}_h(t/h_n)^2} + \frac{1}{h_n^2}\frac{\overline{\phi}_K(t)\left(\overline{\phi}'_h(t/h_n)\right)^2}{\left(\overline{\phi}_h(t/h_n)\right)^3}. \quad (2)$$

Based on (2), we get:

$$(ix)^2\overline{W}_{h_n}(x) = I_1(x) + I_2(x) + I_3(x) + I_4(x)$$

and thus

$$|I_1(x)| \leq \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}''_K(t)}{\overline{\phi}_h(t/h_n)}\right|dt, \quad |I_2(x)| \leq \frac{1}{2\pi}\frac{2}{h_n}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}'_h(t/h_n)\overline{\phi}'_K(t)}{\left(\overline{\phi}_h(t/h_n)\right)^2}\right|dt,$$

$$|I_3(x)| \leq \frac{1}{2\pi}\frac{1}{h_n^2}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}_K(t)\overline{\phi}''_h(t/h_n)}{\overline{\phi}_h(t/h_n)^2}\right|dt, \quad |I_4(x)| \leq \frac{1}{2\pi}\frac{1}{h_n^2}\int_{-\infty}^{\infty}\left|\frac{\overline{\phi}_K(t)\left(\overline{\phi}'_h(t/h_n)\right)^2}{\left(\overline{\phi}_h(t/h_n)\right)^3}\right|dt.$$

By assumptions i), we have

$$|I_1(x)| \leq \frac{1}{2\pi|B_1|h_n^{\beta}}\int_{-\infty}^{\infty}|t|^{\beta}|\overline{\phi}''_K(t)|dt, \quad |I_2(x)| \leq \frac{1}{2\pi|B_1|^2h_n^{\beta}}\int_{-\infty}^{\infty}|t|^{\beta-1}|\overline{\phi}'_K(t)|dt,$$

$$|I_3(x)| \leq \frac{1}{2\pi|B_1|^2h_n^{\beta}}\int_{-\infty}^{\infty}|t|^{\beta-2}|\overline{\phi}_K(t)|dt, \quad |I_4(x)| \leq \frac{1}{2\pi|B_1|^3h_n^{\beta}}\int_{-\infty}^{\infty}|t|^{\beta-2}|\overline{\phi}_K(t)|dt.$$

Thus using iii) we have $h_n^{\beta}|\overline{W}_{h_n}(x)| \leq C_1$ where C_1 is independent of h_n; and $h_n^{\beta}||\overline{W}_{h_n}(x)||_1 \leq const$. □

Lemma 6 *Assume that* $\overline{\phi}_h(t)$ *and* $\overline{\phi}_K(t)$ *are twice continuously differentiable with bounded derivatives such that*

i) $|\overline{\phi}_h(t)| \geq 0$ *for all* $t \in \mathbb{R}$,

ii) $|t|^{\beta}\overline{\phi}_h(t) \to B_1$ *as* $|t| \to \infty$ *for some* B_1 *and* $\beta \geq 1$,

iii) $\int_{-\infty}^{\infty}|u|^{\beta-j}|\overline{\phi}_K(u)|^2du < \infty$ *for* $j = 1, 2$; $\int_{-\infty}^{\infty}|t|^{\beta-j}|\overline{\phi}'_K(t)|^2dt < \infty$ *for* $j = -1, 0, 1$; $\int_{-\infty}^{\infty}|t|^{\beta}|\overline{\phi}''_K(t)|^2dt < \infty$. *Then* $h_n^{\beta}|\overline{W}'_{h_n}(x)| \leq C_2 \leq \infty$, *where* C_2 *is a constant independent of* h_n, *and* $h_n^{\beta}\left\|\overline{W}'_{h_n}(x)\right\|_1 \leq const$.

Proof. Let $F_{h_n}(t) = \frac{t\overline{\phi}_K(t)}{\overline{\phi}_h(t/h_n)}$. Then

$$F''_{h_n}(t) = \frac{(t+1)\overline{\phi}'_K(t) + \overline{\phi}''_K(t)}{\overline{\phi}_h(t/h_n)} - \frac{2}{h_n}\frac{\overline{\phi}_K(t)\overline{\phi}'_h(t/h_n) + t\overline{\phi}'_K(t)\overline{\phi}'_h(t/h_n)(t)}{\left(\overline{\phi}_h(t/h_n)\right)^2} - \frac{1}{h_n^2}\frac{t\overline{\phi}_K(t)\overline{\phi}''_h(t/h_n)}{\left(\overline{\phi}_h(t/h_n)\right)^2} + \frac{2}{h_n^2}\frac{t\overline{\phi}_K(t)\left(\overline{\phi}'_h(t/h_n)\right)^2}{\left(\overline{\phi}_h(t/h_n)\right)^3}.$$

Proceeding in the same way than Lemma 7, we show that $h_n^{\beta}||\overline{W}'_{h_n}(x)||_1 \leq Const$. We conclude, in view of Lemma 7 and lemma 8, that $h_n^{p\beta}||g_{n,k}(x)||_1 \leq const$. □

4. Main Results

Let us now come back to the main purpose of this paper, which is the study of the estimate of the gradient of density of probability under association.

4.1 Estimation of $\nabla \widehat{f_n}(x)$

We have from (1)

$$\frac{\partial \widehat{f_n}(x)}{\partial x_k} = \frac{1}{(2\pi)^p} \int_{\mathbb{R}^p} (-it_k) e^{-itx} \hat{\phi}_n(t) \frac{\phi_K(h_n t)}{\phi_h(t)} dt,$$

where $t = (t_1, ..., t_p)$. Replace $\hat{\phi}_n(t)$ by its expression, we get

$$\frac{\partial \widehat{f_n}(x)}{\partial x_k} = \frac{1}{(2\pi)^p} \frac{1}{(n-p)h_n^{p+1}} \sum_{j=0}^{n-p} \int_{\mathbb{R}^p} (-it_k) e^{-it\left(\frac{x-Y_j}{h_n}\right)} \frac{\phi_K(t)}{\phi_h(t/h_n)} dt.$$

Now let

$$g_{n,k}(x) = \frac{1}{(2\pi)^p} \int_{\mathbb{R}^p} (-it_k) e^{-itx} \frac{\phi_K(t)}{\phi_h(t/h_n)} dt$$

and

$$g_n(x) = (g_{n,1}(x), ..., g_{n,p}(x)).$$

We can write

$$\nabla \widehat{f_n}(x) = \frac{1}{(n-p)h_n^{p+1}} \sum_{j=0}^{n-p} g_n\left(\frac{x - \mathbf{Y}_j}{h_n}\right). \tag{3}$$

4.2 Main Results

Proposition 7 *Assume that $\nabla \widehat{f}(x) \in L_1(\mathbb{R}^p)$ and f continuous, then for all $x \in \mathbb{R}^p$, we have*

$$E\left(\nabla \widehat{f_n}(x)\right) \to \nabla f(x) \text{ as } n \to \infty.$$

Theorem 8 *Let $nh_n^{(2\beta+1)p+1} \to \infty$ as $n \to \infty$. Under assumptions A1-A3 and conditions on kernel function and on distribution of errors in the lemmas 4, 5 and 6; we have:*

$$\lim_{n\to\infty} nh_n^{(2\beta+1)p+1} var\left[\nabla \widehat{f_n}(x)\right] = \sigma^2(x)\, 1_p,$$

where 1_p is the matrix all of whose elements are 1, and

$$\sigma^2(x) = \left(\frac{1}{2\pi|B_1|^2} \int_{-\infty}^{\infty} |t|^{2\beta} |\overline{\phi}_K(t)|^2 dt\right)^{p-1} \left(\frac{1}{2\pi|B_1|^2} \int_{-\infty}^{\infty} |t|^{2(\beta+1)} |\overline{\phi}_K(t)|^2 dt\right) g(x),$$

where B_1 is defined in Lemma 4.

4.3 Special Case of the Main Result

We choose K Gaussian. The distribution of ε is the exponential distribution $h(x) = \lambda e^{-\lambda x}$, then its characteristic function is: $\phi_h(t) = \frac{\lambda}{\lambda - it}$ and $(\phi_h(t))^{-1} = 1 - \frac{i}{\lambda}t$. Here $\beta = 2$. Let

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \nabla f_n(x) = \begin{pmatrix} g_{n,1}(x) \\ g_{n,2}(x) \end{pmatrix}$$

First:

$$I_1 = \int_{-\infty}^{\infty} \left(1 - \frac{it}{\lambda h_n}\right) e^{-itx} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} e^{-\frac{x^2}{2}} \left(1 + \frac{x}{\lambda h_n}\right).$$

Second:

$$I_2 = \int_{-\infty}^{\infty} (-it) \left(1 - \frac{it}{\lambda h_n}\right) e^{-itx} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} e^{-\frac{x^2}{2}} \left(x - \frac{x^2 - 1}{\lambda h_n}\right).$$

Then

$$\begin{aligned} g_{n,1}(x) &= \frac{1}{2\pi}\left(\sqrt{2\pi}e^{-\frac{x_1^2}{2}}\left(x_1 - \frac{x_1^2-1}{\lambda h_n}\right)\right)\left(\sqrt{2\pi}e^{-\frac{x_2^2}{2}}\left(1 + \frac{x_2}{\lambda h_n}\right)\right) \\ &= e^{-\frac{1}{2}(x_1^2+x_2^2)}\left(x_1 - \frac{x_1^2-1}{\lambda h_n}\right)\left(1 + \frac{x_2}{\lambda h_n}\right). \end{aligned}$$

and

$$g_{n,2}(x) = e^{-\frac{1}{2}(x_1^2+x_2^2)}\left(x_2 - \frac{x_2^2-1}{\lambda h_n}\right)\left(1 + \frac{x_1}{\lambda h_n}\right)$$

5. Proof of Main Results

5.1 Proof of Proposition 3

We have from (3):

$$E\nabla \hat{f}_n(x) = \frac{1}{(n-p)h_n^{p+1}}\sum_{j=0}^{n-p} E\left[g_n\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right]$$

and

$$E\left[g_n\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right] = \left(E\left[g_{n,1}\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right], ..., E\left[g_{n,p}\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right]\right),$$

hence

$$\begin{aligned} E\left[g_{n,k}\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right] &= \frac{1}{(2\pi)^p}\int_{\mathbb{R}^p}(-it_k)e^{-itx}E(e^{it\mathbf{Y}_j})\frac{\phi_K(t)}{\phi_h(t/h_n)}dt \\ &= \frac{1}{h_n^{p+1}(2\pi)^p}\int_{\mathbb{R}^p}(-it_k)e^{-itx}\phi_f(t)\phi_K(h_n t)dt. \end{aligned}$$

We use $E(\widehat{\phi_n}(t)) = \phi(t)$.

Now, set $f_k^{'}(x) = \frac{\partial f}{\partial x_k}$, then we can write

$$\begin{aligned} h_n^{p+1}E\left[g_{n,k}\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right] &= \frac{1}{(2\pi)^p}\int_{\mathbb{R}^p} e^{-itx}\phi_{f_k^{'}}(t)\phi_K(h_n t)dt \\ &= f_k^{'}(x) * \frac{1}{h_n}K(\frac{x}{h_n}) = \int_{\mathbb{R}^p}\frac{1}{h_n}K(\frac{u}{h_n})f_k^{'}(x-u)du, \end{aligned}$$

where $*$ is the convolution operator. Then, applying Lemma 3, we get

$$\lim_{n\to\infty} h_n^{p+1}E\left[g_{n,k}\left(\frac{x-\mathbf{Y}_j}{h_n}\right)\right] = f_k^{'}(x) = \frac{\partial f}{\partial x_k}(x).$$

We conclude that $E\nabla\widehat{f_n}(x) \to \nabla f(x)$.

5.2 Proof of Theorem 4

$$\begin{aligned} var\nabla\widehat{f_n}(x) &= \frac{1}{n-p}\sum_{l=-(n-p)}^{n-p}\left(1 - \frac{|l|}{n-p}\right)cov\left(\frac{1}{h_n^{p+1}}g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right), \frac{1}{h_n^{p+1}}g_n\left(\frac{x-\mathbf{Y}_{|l|}}{h_n}\right)\right) \\ &= I_{n,0} + 2\sum_{l=1}^{n-p}\left(1 - \frac{l}{n-p}\right)I_{n,l}. \end{aligned}$$

First part

$$I_{n,0} = \frac{1}{n-p}E\left(\left(\frac{1}{h_n^{p+1}}g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right)\right)^t\left(\frac{1}{h_n^{p+1}}g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right)\right)\right) + O(\frac{1}{n}).$$

i)

$$\begin{aligned} E\left(g_{n,k}^2\left(\frac{x-\mathbf{Y}_0}{h_n}\right)\right) &= \int_{\mathbb{R}^p} g_{n,k}^2\left(\frac{x-u}{h_n}\right)g(u)du \\ &= h_n^{p+1}\int_{\mathbb{R}^p} |g_{n,k}(u)|^2 g(x-uh_n)du \\ &= h_n^{p+1}\left(g(x)\int_{\mathbb{R}^p} |g_{n,k}(u)|^2 du + o(1)\right) \end{aligned}$$

and we conclude using Lemma 4.

ii)

$$\begin{aligned} E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)\right) &= \int_{\mathbb{R}^p} g_{n,k}\left(\frac{x-u}{h_n}\right)g_{n,k'}\left(\frac{x-u}{h_n}\right)g(u)du \\ &= h_n^{p+1}\int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)g(x-uh_n)du \\ &= h_n^{p+1}\left(g(x)\int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)du + o(1)\right). \end{aligned}$$

We can write:

$$\begin{aligned} \int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)du &\leqslant \left|\int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)du\right| \\ &\leqslant \int_{\mathbb{R}^p} \left|g_{n,k}(u)g_{n,k'}(u)\right| du = \| g_{n,k}\times g_{n,k'} \|_1^2 . \end{aligned}$$

By Cauchy-Schwarz's inequality, we get $\| g_{n,k}\times g_{n,k'} \|_1 \leqslant \sqrt{\| g_{n,k}\|_2}\sqrt{\| g_{n,k'}\|_2}$,

thus

$$\int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)du \leqslant \left[\int_{\mathbb{R}^p} \left|g_{n,k}(x)\right|^2 dx\right]\times\left[\int_{\mathbb{R}^p} \left|g_{n,k'}(x)\right|^2 dx\right].$$

We deduce

$$h_n^{2\beta p}\int_{\mathbb{R}^p} \left|g_{n,k}(x)\right|^2 dx \to D_1^{p-1}D_2 \tag{4}$$

and

$$h_n^{2\beta p}\int_{\mathbb{R}^p} \left|g_{n,k'}(x)\right|^2 dx \to D_1^{p-1}D_2 \tag{5}$$

in view of Lemma 4.

$$\begin{aligned} \frac{1}{h_n^{p+1}}\int_{\mathbb{R}^p} g_{n,k}(u)g_{n,k'}(u)du &\leqslant \frac{1}{h_n^{p+1}}\times\sqrt{\left[h_n^{2\beta p}\int_{\mathbb{R}^p}\left|g_{n,k}(x)\right|^2 dx\right]} \\ &\leqslant \frac{1}{h_n^{2\beta p+p+1}}\times\sqrt{\left[h_n^{2\beta p}\int_{\mathbb{R}^p}\left|g_{n,k}(x)\right|^2 dx\right]}\times\sqrt{\left[h_n^{2\beta p}\int_{\mathbb{R}^p}\left|g_{n,k'}(x)\right|^2 dx\right]} \\ &\leqslant \frac{1}{h_n^{2\beta p+p+1}}\left(D_1^{p-1}D_2\right). \end{aligned}$$

and then

$$\lim_{n\to\infty}\left(nh_n^{(2\beta+1)p+1}E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)\right)\right) = \left(D_1^{p-1}D_2\right)g(x).$$

i) and ii) gives

$$\lim_{n\to\infty} nh_n^{p(2\beta+1)+1}I_{n,0} = D_1^{p-1}D_2 g(x)1_p.$$

Second part

$$S = \sum_{j=1}^{n-p}\left(1-\frac{l}{n-p}\right)I_{n,l} \leq \sum_{j=1}^{p-1}|I_{n,l}| + \sum_{j=p}^{\pi_n}|I_{n,l}| + \sum_{j=\pi_n+1}^{n-p}|I_{n,l}| = S_1 + S_2 + S_3$$

with

$$I_{n,l} = \frac{1}{(n-p)h_n^{2p+2}} cov\left(g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right), g_n\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right).$$

Contribution of S_1.

Assume that the vector $\mathbf{Y}_0 = (Y_1, ..., Y_p)$ and $\mathbf{Y}_j = (Y_{j+1}, ..., Y_{j+p})$ have a joint probability density $g(u, v)$ of order $2p$ for $j \geq p$ and let $g(u)$ be the probability density of $\mathbf{Y}_0$. Define the dependence index of the process (Y_i) by

$$d_{p,n} = sup_{u,v} \frac{1}{n} \sum_{j=p}^{n} |g(u, v) - g(u)g(v)|.$$

Assume that

$$sup_{u,v,j\geq p}|g(u, v) - g(u)g(v)| \leq M_2, \tag{6}$$

for some $M_2 < \infty$.

For $1 \leq l \leq p-1$, we note that the vectors $\mathbf{Y}_0 = (Y_1, ..., Y_p)$ and $\mathbf{Y}_l = (Y_{l+1}, ..., Y_{l+p})$ overlap.

Let $g(u^{'}, u^{''}, u^{'''})$ be the joint probability density of $(Y_1, ...Y_{l+p})$ with $u^{'}, u^{''}, u^{'''}$ having dimensions l, $p-l$ and l respectively. Let

$$\begin{aligned} E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right) &= h_n^{p+l+3} \int_{\mathbb{R}^{p+l}} g_{n,k}(u^{'}, u^{''})g_{n,k'}(u^{''}, u^{'''}) \\ & [g(x^{'} - h_n u^{'}, x^{''} - h_n u^{''}, x^{'''} - h_n u^{'''}) - \\ & g(x^{'} - h_n u^{'}, x^{''} - h_n u^{''}).g(x^{''} - h_n u^{''}, x^{'''} - h_n u^{'''})]du^{'}du^{''}du^{'''}. \end{aligned}$$

We deduce then

$$\begin{aligned} E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right) &\leq h_n^{p+l+3} M_2 \int_{\mathbb{R}^{p+l}} g_{n,k}(u^{'}, u^{''})g_{n,k'}(u^{''}, u^{'''})du^{'}du^{''}du^{'''} \\ &\leq h_n^{p+l+3} M_2 \left[\int_{-\infty}^{\infty} |\overline{W}_{h_n}(x)|^2 dx\right]^{p-l-1} \left[\int_{-\infty}^{\infty} |\overline{W}^{'}_{h_n}(x)|^2 dx\right] \\ & \left[\int_{-\infty}^{\infty} |\overline{W}_{h_n}(x)|\right]^{2(l-1)} \left[\int_{-\infty}^{\infty} |\overline{W}^{'}_{h_n}(x)| dx\right]^2. \end{aligned}$$

It follows by Lemmas 4, 5 and 6; that

$$E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right) \leq h_n^{p+l+3} M_2 h_n^{-2\beta l} h_n^{-2\beta(p-l)}.$$

Consequently

$$S_1 \leq O\left(\sum_{l=1}^{p-1} \frac{h_n^{l+1-(2\beta+1)p}}{(n-p)}\right)$$

so that

$$nh_n^{(2\beta+1)p+1} S_1 = O\left(\sum_{l=1}^{p-1} h_n^{l+2}\right).$$

Next we consider S_2.

Let

$$S_2 = \frac{1}{(n-p)h_n^{2p+2}} \sum_{j=p}^{\pi_n} cov\left(g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right), g_n\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right).$$

for $p \leq l \leq \pi_n$, where $\pi_n \to \infty$ such that $\pi_n h_n^{p+1} \to 0$ as $n \to \infty$.

Then

$$\begin{aligned} E\left(g_{n,k}\left(\frac{x-\mathbf{Y}_0}{h_n}\right)g_{n,k'}\left(\frac{x-\mathbf{Y}_l}{h_n}\right)\right) &= \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} g_{n,k}\left(\frac{x-u}{h_n}\right)g_{n,k'}\left(\frac{x-v}{h_n}\right)[g(u,v)-g(u)g(v)]dudv \\ &\leq M_2\left(h_n^{p+1}\int_{\mathbb{R}^p}\left|g_{n,k}(u)\right|du\right).\left(h_n^{p+1}\int_{\mathbb{R}^p}\left|g_{n,k'}(u)\right|du\right) \\ &\leq M_2 h_n^{2p+2}h_n^{-2\beta p}, \end{aligned}$$

where we used Lemmas 5 and 6 and the inequality (6). Hence we get

$$S_2 = o\left(\frac{1}{nh_n^{(2\beta+1)p+1}}\right).$$

We bound S_3 for associated process and $\pi_n + 1 \leq l \leq n-p$.

We use the following Birkel's Lemma

Lemma 9 (Birkel, 1988) *Let V_i; $i \in I$ be a finite collection of associated (PA) random variables. Let I_1 and I_2 be subsets of I and let H_i be functions on $\mathbb{R}^{|j|}$, $j = 1, 2$; with bounded first order partial derivatives. Then*

$$cov\left(H_1(V_i), H_2\left(V_j\right)\right) \leq \sum_i\sum_j \left\|\frac{\partial H_1}{\partial t_i}\right\|_\infty \left\|\frac{\partial H_2}{\partial t_j}\right\|_\infty cov\left(V_i, V_j\right),$$

where $\|\|_\infty$ stands for the sup norm.

In our case H_j are g_n. The derivative can be written: $\frac{\partial g_n(x)}{\partial x_l} = (\frac{\partial g_{n,1}(x)}{\partial x_l}, ..., \frac{\partial g_{n,p}(x)}{\partial x_l})$.

For $k = 1, ..., p$,

$$\frac{\partial g_{n,k}(x)}{\partial x_l} = \frac{1}{(2\pi)^p}\int_{\mathbb{R}^p} t_k t_l e^{-itx}\frac{\phi_K(t)}{\phi_h(t/h_n)}dt,$$

which implies

$$\left|\frac{\partial g_{n,k}(x)}{\partial x_l}\right| \leq \left(\int_{\mathbb{R}^{p-2}}\frac{\phi_K(t)}{\phi_h(t/h_n)}dt\right)\left(\int_{-\infty}^{\infty}\frac{t\phi_K(t)}{\phi_h(t/h_n)}dt\right)^2.$$

Hence, applying Lemma 5 and 6, we get

$$h_n^{\beta p}\left|\frac{\partial g_{n,k}(x)}{\partial x_l}\right| \leq Const,$$

and

$$h_n^{\beta p}\left|\frac{\partial g_n(x)}{\partial x_l}\right| \leq const.$$

For associated process (Y_j), we apply the Lemma 9 above.

$$\begin{aligned} \left|cov\left(g_n\left(\frac{x-\mathbf{Y}_0}{h_n}\right), g_n\left(\frac{x-\mathbf{Y}_{|l|}}{h_n}\right)\right)\right| &\leq \sum_i\sum_j\left(\frac{Const}{h_n^{\beta p}}\right)^2 cov(Y_{j+l}, Y_i) \\ &= \sum_i\sum_j\left(\frac{Const}{h_n^{\beta p}}\right)^2 cov(X_{j+l}, X_i), \end{aligned}$$

where we have used the fact that $cov(Y_{j+l}, Y_i) = cov(X_{j+l}, X_i)$ due to the independence of (X_j) and (ε_j) and i.i.d assumption on the $(\varepsilon_j)'s$. Thus

$$\begin{aligned} S_3 &\leq \frac{Const}{nh_n^{2(\beta p)+2p+2}}\sum_{i=-(p-1)}^{p-1}\sum_{l=\pi_n+1}^{n-p} c_{l+i} \\ &\leq \frac{Const}{nh_n^{p(2\beta+1)+p+2}}\sum_{i=-(p-1)}^{p-1}\sum_{l=\pi_n+i+1}^{n-p+i} c_l \\ &\leq \frac{Const}{\pi^\delta nh_n^{(2\beta+1)p+1+(p+1)}}\sum_{l=\pi_n}^{\infty} l^\delta c_l, \end{aligned}$$

where $c_l = cov(X_{l+1}, X_1)$ is the covariance function of $(X_i)_i$.

We now select $\pi_n = h_n^{-\frac{p+1}{\delta}}$ and conclude by condition (A3) that:

$$nh_n^{(2\beta+1)p+1} S_3 \to 0.$$

The result of Theorem 4 follows from first and second part.

Theorem 4 and Proposition 3 yield the quadratic mean convergence of the estimate $\nabla \widehat{f_n}(x)$.

6. Conclusion

We have studied in this paper the problem of estimating the gradient of multivariate probability densities of a stationary process using observations that are corrupted by additive noise. An important problem in non parametric estimation consists in estimation of the mode, i.e., the location of an isolated maximum of the unknown density. Nonparametric estimation of the mode of a density function via kernel methods may be considered when data is contaminated with associated observations. We can then study the asymptotic properties of these mode estimates.

References

Birkel, T. (1988). On the convergence rate in the central limit theorems for associated processes. *Ann. Probab., 16*(4), 1685-1698. http://dx.doi.org/10.1214/aop/1176991591

Bulinski, A. V. (1996). On the convergence rate in the CLT for positively and negatively dependant random field. In I. A. Ibragimov & A. Y. Zaitov (Eds.), *Probability Theory and Mathematical Statistics*. Gordon and Breach.

Carrol, R. J., & Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc., 83*, 1184-1186. http://dx.doi.org/10.1080/01621459.1988.10478718

Chacon, J. E., Duong, T., & Wand, M. P. (2011). Asymptotics for general multivariate kernel density derivative estimators. *Statistica Sinica, 21*, 807-840. http://dx.doi.org/10.5705/ss.2011.036a

Esary, J., Proschan, F., & Walkup, D. (1967). Association of random variables with applications. *Ann. Math. Stat., 38*, 1466-1476. http://dx.doi.org/10.1214/aoms/1177698701

Fan, J. (1991a). On the rates of convergence for non parametric deconvolution problems. *The annals of statistics., 19*(3), 1257-1272. http://dx.doi.org/10.1214/aos/1176348248

Fan, J. (1991b). Asymptotic normality for deconvolution kernel density. *Sankhya. The Indian Journal of Statistics, Series A, 53*(1), 97-110.

Louhichi, S. (2000). Weak convergence for empirical processes of associated sequences. *Annales de l'Institut Henri Poincare (B) Probability and Statistics, 36*(5), 547-567. http://dx.doi.org/10.1016/S0246-0203(00)00140-0

Masry, E. (1991). Multivariate Probability Density Deconvolution for Stationary Random Processes. *IEEE. Transactions on Information Theory., 37*(4), 1105-1115. http://dx.doi.org/10.1109/18.87002

Masry, E. (2003). Involving Multivariate Kernel Density Estimates from Contamined Associated Observations. *IEEE. Transactions on Information Theory, 49*(11), 2941-2952. http://dx.doi.org/10.1109/TIT.2003.818415

Mendelson, J., & Rice, J. (1982). Deconvolution of microfluorometric histograms with B-splines. *Amer. Statist. Assoc., 77*, 748-753.

Pitt, L. (1982). Positively correlated normal variables are associated. *Ann. Probab., 10*, 496-499. http://dx.doi.org/10.1214/aop/1176993872

Roussas, G. G. (1991). Kernel estimates under association: Strong uniform consistency. *Statist. Probab. Lett., 12*, 215-224. http://dx.doi.org/10.1016/0167-7152(91)90028-P

Roussas, G. G. (2000). Asymptotic normality of the kernel estimate of a probability density function under association. *Statistics and Probability Letters, 50*, 1-12. http://dx.doi.org/10.1016/S0167-7152(00)00072-9

Stefanski, L., & Caroll, R. J. (1990). Deconvolving kernel density estimate. *Statistics, 21*, 169-184. http://dx.doi.org/10.1080/02331889008802238

Wheeden, R. L., & Zygmund, A. (1977). *Measure and Integral*. New York: Marcel Dekker.