Research & Reviews: Journal of Statistics and Mathematical Sciences

Estimation of Long Memory Linear Processes

Ichaou Mounirou*

Faculty of Management and Economics Sciences, Université de Parakou, Benin

Letter

Received date: 10/05/2016

Accepted date: 24/05/2016

Published date: 28/05/2016

***For Correspondence**

Ichaou Mounirou, Faculty of Management and Economics Sciences, Université de Parakou, Benin

E-mail: ichaou_bassir@yahoo.fr

Keywords: Multivariate Processes, Kernel Density, Hellinger Distance, Linear Process, Parametric Estimation, Long Memory, Multivariate Processes

ABSTRACT

This paper studies asymptotic properties of the minimum distance Hellinger estimates for a stationary multivariate linear gaussien long range dependent process of the form $X_t = \sum_{u=0}^{\infty} A_u(\theta) Z_{t-u}$, where $(Z_t)_{t\in\mathbb{Z}}$ is a sequence of strictly stationary d-dimensional associated random vectors with $E(Z_t) = 0$ and $E\left(\|Zt\|^2\right) < \infty$ and $\{A_u\}$ is a sequence of coefficient matrices with $\sum_{u=0}^{\infty} \|Au\| < \infty$ and $\sum_{u=0}^{\infty} A_u \neq 0_d \times_d$. By means of the properties of the kernel density estimate, the minimum distance Hellinger of this class are shown to be consistent, asymptotically normal and robust.

INTRODUCTION

Let $(X_t)_{t\in\mathbb{Z}}$ bead-variate linear process independent of the form:

$$X_t = \sum_{u=0}^{\infty} A_u(\theta) Z_{t-u} \tag{1}$$

Defined on a probability space (Ω, $\mathfrak{I}$ (F), p), where $\{Z_t\}$ is a sequence of stationary d-variate associated random vectors with $E(Z_t) = 0$, $E\left(\|Zt\|^2\right) < +\infty$ and positive definite covariance matrix $\tilde{a} : d \times d$. Throughout this paper we shall assume that

$$\sum_{u=0}^{\infty} \|Au\| < \infty \tag{2}$$

$$\sum_{u=0}^{\infty} A_u \neq 0_d \times_d \tag{3}$$

where for any d d, d ≥ 2, matrix $A = (a_{ij}(\theta))$ whose components depend on the parameter θ, such as $\|Au\| \sum_{i=1}^{d} \sum_{j=1}^{d} a_{ij}; \sum_{u=0}^{\infty} \|ai\|^2 < \infty$ and $0_{d\times d}$ denotes the d×d zero matrix. Here $\theta \in \Theta$ with $\Theta \subset \mathbb{R}$, with. Let

$$T = \left(\sum_{j=0}^{\infty} A_j\right) r \left(\sum_{j=0}^{\infty} A_j\right)', \tag{4}$$

where the prime denotes transpose, and the matrix $r = (\sigma_{kj})$ with

$$\sigma_{kj} = E(Z_{1k} Z_{1j}) + \sum_{t=2}^{\infty} (E(Z_{1k} Z_{tj}) + E(Z_{1k} Z_{tj})) \tag{5}$$

Further, let $Sn = \sum_{t=1}^{n} X_t, n \geq 1)$ $(S_0 = 0)$.

$\{X_t\}_{t\in\mathbb{Z}}$ is assumed to be gaussian and have long rang dependent process. Fakhre-Zakeri and Lee proved a central theorem for multivariate linear processes generated by independent multivariate random vectors and Fakhre-Zakeri and Lee also derived a functional central limit theorem for multivariate linear processes generated by multivariate random vectors with martingale difference sequence. Tae-Sung Kim, Mi-HwaKo and Sung-Mo Chung [1] prove a central limit theorem for d-variate associated random vectors. The problem is how to estimate θ in order to investigate the fitting of the model to the data? An estimation of θ would have two essential properties: it would be efficient and its distribution would not be greatly pertubated.

{X_t} is a multivariate Gaussian process in dependent with density $f_\theta(.)$. We estimate the parameters in the general multivariate linear processes in (1).

In this paper is to prove a general estimation of the parameter vector θ by the minimum Hellinger distance Method (MHD). The only existing examples of MHD estimates are related to i.i.d. sequences of random variables's [2-4]. For long memory univariate linear processes see Bitty and Hili [5]. The long memory concept appeared since 1950 from the works of Hurst in hydraulics. The process $\{X_t\}_{t\in\mathbb{Z}}$ is said to be a long memory process if in (1), λ is a parameter of long memory, and $1/2<\lambda_{ii}<1$ for $j=1;\dots;d$ and $i=1;\dots;d$.

The paper developers in section 2, some assumptions and lemmas, essentially based on the work of Tae-Sung Kim, Mi-Hwa Ko and Sung-Mo Chung [1] and the work of Theophilos Cacoullos [6]. Our main results arein section 3, based on work of Bitty and Hili [5] which show consistency and the asymptotic properties of the MHD estimators of the parameter θ. We conclude with some examples,

ASSUMPTIONS AND LEMMAS

Parzen [7] gave the asymptotic properties of a class of estimates $f_n(x)$ an univariate density function f(x) on the basis of random sample $X_1,\dots,X_n$ from f(x). Motivated as in Parzen, we consider estimates $f_n(x)$ ofthe density function f(x) of the following form:

$$f_n(x)=\int\frac{1}{h_n^d}K\left(\frac{x-y}{h_n}\right)dF_n(y) \quad (6)$$

$$=\frac{1}{nh_n^d}\sum_{j=1}^{n}K\left(\frac{x-X_j}{h_n}\right), \quad (7)$$

where $F_n(x)$ denotes the empirical distribution function based on the sample of *n* independent observations $X_1,\dots,X_n$ on the random d-dimensional vector X with chosen to satisfy suitable conditions and $\{h_n\}$ is a sequence of positive constants which in the sequel will always satisfy $h_n\to 0$, as $n\to\infty$. We suppose K(y) is a bore scalar function on E_d such that

$$sup_{y\in E_d}|K(y)|<\infty \quad (8)$$

$$\int|K(y)|dy \quad <\infty \quad (9)$$

$$|y|^d K(y)\to 0,\ \text{as}|y|\to\infty \quad (10)$$

where $|y|$ denotes the length of the vector.

And

$$\int K(y)dy=1 \quad (11)$$

$$K(y)=K(-y) \quad \text{for all } y, \quad (12)$$

also K(y) is absolutely integrable (hence f(x) is uniformly continuous).

$$\int y_i K(y)dy=0 \quad (13)$$

and $\int K^3(y)dy<\infty\ (p+4)^{-1}<\alpha<p^{-1}\ and\ 3\le \iota\le 5$ (14)

See Theopilos Cacoullos [6] and Bitty, Hili [5]

Notations and Assumptions: Let $\mathcal{F}=\{f(.,\theta)\}_{\theta\in\Theta}$ be a family of functions whereΘ is a compact parameter set of $\mathbb{R}^q$ such that for all $\theta\in\Theta$, $f(.,\theta):\mathbb{R}^d\to\mathbb{R}$ is a positive integral function. Assume that f(.,θ) satisfies the following assumptions.

(A1): For all $\theta,\mu\in\Theta,\theta\neq\mu$ is a continuous differentiable function at $\theta\in\Theta$.

(A2): (i) $f(x,\theta)$ and $\frac{\partial}{\partial x}f^{\frac{1}{2}}(x,\theta)$ have a zero Lebesgue measure and f (.,θ)is bounded on $\mathbb{R}^d$.

(ii) For $\theta,\mu\in\Theta,\theta\neq\mu$ implies that $\{x/f(x,\theta)\neq f(x,\mu)\}$ is a set of positive Lebesgue mesure, for all $x\in\mathbb{R}^d$

(A3): K the kernel function such that

$$\int_{\mathbb{R}^d}K(u)du=1,\tau^2=\int_{\mathbb{R}^d}K^2(u)du<\infty.$$

(A4): The bandwidths {b_n} satisfy natural conditions, $b_n \to 0, n^{\iota} b_n^d \to \infty$ for $\iota \geq 1$ *when* $n \to \infty$

(A5): There exists a constant β>0 such that $inf_{\theta \epsilon \Theta} inf_{x \epsilon \mathbb{R}^d} f_\theta \geq \beta.$

Let F denote the set of densities with respect to the Lebesgue measure on $\mathbb{R}^d$. Define the functional $T: \mathcal{G} \to \Theta$ in the following:

Let $g \dot{o} \mathcal{G}$. Denote by B(g) the set $B(g) = \{\theta \epsilon \Theta : H_2(f_\theta, g) = min_{\beta\ \epsilon,\ \Theta} H_2(f_\beta, g)\}$ where H_2 is the Hellinger distance.

If B(g) is reduced to a unique element, then define T(g) as the value of this element. Elsewhere, we choose an arbitrary but unique element of these minimums and call it T(g).

Lemma 1: Let $(Zt)t \epsilon \mathbb{Z}$ be a strictly stationary associated sequence of d–dimensional random vectors with E(Zt) = 0, $E(Zt) < +\infty$ and positive definite covariance matrix r as (5). Let (X_t) be a d-variate linear process defined as in (1). Assume that

$$E\left(\|Z1\|^2\right) + \sum_{t=2}^{\infty} \sum_{i=1}^{d} cov(Z_{1i}, Z_{ti}) = \sigma^2 < \infty, \quad (15)$$

then, the linear process(X_t) fulfills the limit central theorem, that is, $Sn = n^{-1/2} S_n \to^{\mathfrak{D}} N(0, T)$, (16)

Where $\to^{\mathfrak{D}}$ denotes the convergence in distribution and N (0, T) indicates an normal distribution with mean zero vector and covariance matrix T defined in (4).

For the proof of lemma 1, see theorem 1.1 of Tae-Sung Kim, Mi-Hwa Ko and Sung-Mo Chung [1]

Lemma 2: To remark 3.2 and theorem 3.5 of Tae-Sung Kim, Mi-Hwa Ko and Sung-Mo Chung [1], we have

$$\left(nh^p\right)^{\iota/2} \left(f_n(x) - f(x)\right) \to^{\mathfrak{D}} N\left(0, f(x) \int K^2(y) dy\right). \quad (17)$$

For the proof of lemma 2, see Tae-Sung Kim, Mi-Hwa Ko and Sung-Mo Chung [1]

Lemma 3: Assume that (A_5) holds. If f_1 is continuous on $\mathbb{R}$ and if for almost all x, h is continuous on Φ, then

(i) for all $g \epsilon, F\ B(g) \neq \emptyset.$

(ii) If B(g) is reduced to an unique element, then t is continuous on g Hellinger topology.

(iii) $T(f_\theta) = \theta$ Uniquely on Θ

Proof: See Lemma 3.1 in Bitty and Hili [5].

Lemma 4: Assume that $g_\theta = f^{1/2}(., \theta)$ satisfies assumptions (A_1),-(A_3). Then, for all sequence of density $\{\hat{f}_n\}_{n \epsilon \mathbb{N}}$ converges to f_θ in the Hellinger topology.

$$T(\hat{f}_n) = T(f_\theta) + \int_{\mathbb{R}^d} \rho_\theta(x) \left[\hat{f}_n^{1/2}(x) - f_\theta^{1/2}(x)\right] dx + a_n \int_{\mathbb{R}^d} \dot{g}_{T(f_\theta)}(x) \left[\hat{f}_n^{1/2}(x) - f_\theta^{1/2}(x)\right] dx,$$

where,

$$\rho_\theta(x) = -\left[\int_{\mathbb{R}^d} \ddot{g}_{T(f_\theta)}(x) f_\theta^{1/2}(x) dx\right]^{-1} \dot{g}_{T(f_\theta)}(x),$$

With a_n $a(q \times q)$ - matrix whose components tends to zero $n \to \infty$

Proof: See Theorem 2 in Beran [2]

Lemma 5: Under assumptions (A_3), if the bandwidth b_n is an theorems 1 and 2, if *f*(.,*θ*) is continuous with a compact support. And if the density *f*(.,*θ*) of the observations satisfies assumptions(A_1) -(A_2). Then $\hat{f}_n(.)$ converges to *f*(.,*θ*) in the Hellinger topology.

Proof of lemma 8

Under assumption (A_2), (A_3) and (A_5) and lemma 2, we have

$$\left(\int_{\mathbb{R}^d} \left|\hat{f}_n^{1/2}(x) - f^{1/2}\right|^2 dx\right)^{1/2} \to 0, \ a.s.\ n \to \infty,$$

Then $\hat{f}_n \to f$, *a.s.* $n \to \infty$ in Hellinger topology

ESTIMATION OF THE PARAMETER

This method has been introduced by Beran [2] for independent samples, developed by Bitty and Hili [5] for linear univariate processes dependent in long memory. The present paper suppose the process independent multivariate with associated random vectors under same condition of Bittyand Hili [5] in long memory. The minimum Hellinger distance estimate of the parameter vector is obtained via a nonparametric estimate of the density of the process (X_t). We define $\hat{\theta}_n$ as the value of $\theta \epsilon \Theta$ which minimizes the Hellinger distance $H_2(\hat{f}_n, f(., \theta)$

i.e: $H_2\left(\hat{f}_n, f\left(., \hat{\theta}_n\right)\right) = min_{\theta \epsilon \Theta} H_2\left(\hat{f}_n, f(., \theta)\right),$

where $\hat{f}_n$ is the nonparametric estimate of $f(.,\theta)$ and

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

There exist many methods of nonparametric estimation in the literature. See for instance Rosenblatt [8] and therein. For computational reasons, we consider the kernel density estimate which is defined in section 2. Before analyzing the optimal properties of $\hat{\theta}_n$ we need some assumptions.

i.e: $H_2\left(\hat{f}_n, f\left(.,\hat{\theta}_n\right)\right) = min_{\theta\in\Theta} H_2\left(\hat{f}_n, f(.,\theta)\right),$

where $\hat{f}_n$ is the nonparametric estimate of $f(.,\theta)$ and

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

Asymptotic properties

Theorem 1 (Almost Sure Consistency): Assume that (A_1)-(A_5) hold. Then, $\hat{\theta}_n$ almost surely converges toθ.

For the proof, see section 3.

Let denote by J_n as: $J_n = (nb_n)^{\iota/2}$

Let us denote by $R_\theta(.) = f^{1/2}(.,\theta)$, $\dot{R}_\theta(.) = \partial f^{1/2}(.,\theta)/\partial\theta$ *and* $\rho(.,\theta)$ the following function.

$$\rho(x,\theta) = \left[\int_{\mathbb{R}^d} \dot{R}_\theta(x)\dot{R}_\theta^t(x)dx\right]^{-1} \dot{R}_\theta(x),$$

Where $\dot{R}_\theta(x)$ is a quantity which exists, and t denotes the transpose.

Condition 1:

We have the (q × q) matrix sequence v_n in lemma 4 and the sequence J_n are such that $J_n v_n$ tend to zero as $n \to \infty$.

Theorem 2 (Asymptotic distribution): Assume that (A1)-(A6) and condition 1 hold. If

(i) $\int_{\mathbb{R}^d} \ddot{R}_\theta(x)R_\theta(x)dx$ is a nonsingular (qq) -matrix,

(ii) $\rho(.,\theta)$ admits a compact support then, we have $J_n\left[\hat{\theta}_n - \theta\right] \to^{\mathcal{L}} N\left(0; \int_{\mathbb{R}} Y(x,\theta)\sum^2(x)Y^t(x,\theta)dx\right)$

For the proof, see section 3.

Appendices

Proof of theorem 1

$$H_2\left(\hat{f}_n(.), f(.,\theta)\right) = \left\{\int_{\mathbb{R}} \left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]^2 dx\right\}^{1/2} \to 0 \ \ a.s.\ when\, n \to \infty.$$

From lemma 3,

$$H_2\left(\hat{f}_n(.), f(.,\theta)\right) = \left\{\int_{\mathbb{R}} \left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]^2 dx\right\}^{1/2} \to 0 \ \ a.s.\ when\, n \to \infty.$$

As $T\left(\hat{f}_n(.)\right) = \hat{\theta}_n$ *and* $T\left(f(.,\theta)\right) = \theta$ uniquely, the remainder of proof follows from the continuity of the functional T(.) in lemma 1.

Proof of theorem 2

From lemma 2 and the proof of theorem 2 of Bitty and Hili [5], we have

$$J_n\left[\hat{\theta}_n - \theta\right] = J_n \int_{\mathbb{R}} \rho(x,\theta)\left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]dx$$
$$+ v_n J_n \int_{\mathbb{R}} \dot{R}(x,\theta)\left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]dx,$$

where an a (d×d)-matrix whose components tend to zero in probability when $n \to \infty$.

Under condition 1, we have

$$W_n(\theta) = v_n J_n \int_{\mathbb{R}} \dot{R}(x,\theta)\left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]dx \to_p 0.$$

So the limiting distribution of $J_n\left[\hat{\theta}_n - \theta\right]$ depends on the limiting distribution of $J_n L_n(\theta)$, With

$$L_n(\theta) = \int_{\mathbb{R}} \rho(x,\theta)\left[\hat{f}_n^{1/2}(x) - f^{1/2}(x,\theta)\right]dx.$$

For $a \geq 0$,$b \geq 0$, we have the algebraic identity

$$a^{1/2} - b^{1/2} = 2^{-1}b^{-1/2}(a-b) - \left[2b^{1/2}(a^{1/2} + b^{1/2})^2\right]^{-1}(a-b)^2.$$

For $a=\hat{f}_n(x)$ and b = f(x,θ), we have

$$W_n(\theta)=2^{-1}v_nJ_n\int_{\mathbb{R}}\dot{R}_\theta(x)f^{-1/2}(x,\theta)[\hat{f}_n(x)-f(x,\theta)]dx-2^{-1}v_nJ_n\left(\int_{\mathbb{R}}\dot{R}_\theta(x)f^{-1/2}\frac{(\hat{f}_n(x)-f(x,\theta))^2}{(\hat{f}_n^{1/2}(x)-\dot{R}_\theta(x)^{1/2})^2}dx\right)$$

$$=D_n(\theta)+E_n(\theta)$$

With

$$D_n(\theta)=2^{-1}v_nJ_n\int_{\mathbb{R}}\dot{R}_\theta(x)f^{-1/2}(x,\theta)\left[\hat{f}_n(x)-f(x,\theta)\right]dx$$

And

$$E_n(\theta)=-2^{-1}v_nJ_n\left(\int_{\mathbb{R}}\dot{R}_\theta(x)f^{-1/2}\frac{(\hat{f}_n(x)-f(x,\theta))^2}{(\hat{f}_n^{1/2}(x)-\dot{R}_\theta(x)^{1/2})^2}dx\right)$$

From assumption (A6), then $inf_x f(x,\theta)\geq\beta>0$,

$$\left|E_n(\theta)\right|\leq 2^{-1}\beta^{-3/2}v_nJ_n\int_{\mathbb{R}}\left|\dot{R}_\theta(x)\right|\left[\hat{f}_n(x)-f(x,\theta)\right]^2dx.$$

For $a\geq 0$,$b>0$, $(a-b)^2\leq 2\left(a^2+b^2\right)$, then

$$\left|E_n(\theta)\right|\leq\beta^{-3/2}v_nJ_n\int_{\mathbb{R}}\left|\dot{R}_\theta(x)\right|\left[\hat{f}_n(x)-f(x,\theta)\right]^2dx+\beta^{-3/2}v_nJ_n\int_{\mathbb{R}}\left|\dot{R}_\theta(x)\right|\left[E\left(\hat{f}_n(x)\right)-f(x,\theta)\right]^2dx$$

$$=E_{n1}(\theta)+E_{n2}(\theta)$$

With

$$E_{n1}(\theta)=\beta^{-3/2}v_nJ_n\int_{\mathbb{R}}\left|\dot{R}_\theta(x)\right|\left[\hat{f}_n(x)-f(x,\theta)\right]^2dx$$

And

$$E_{n2}(\theta)=\beta^{-3/2}v_nJ_n\int_{\mathbb{R}}\left|\dot{R}_\theta(x)\right|\left[E\left(\hat{f}_n(x)\right)-f(x,\theta)\right]^2dx$$

Under assumptions (A1)−(A2) we apply Taylor Lagrange in order 2 and assumption (A4) we have:

$$E\left(\hat{f}_n(x)\right)-f(x,\theta)=\int_{\mathbb{R}}\left[f(x-b_nz,\theta)-f(x,\theta)\right]K(z)dz$$

$$=\int_{\mathbb{R}}\left(-b_nzf'(x,\theta)+2^{-1}b_n^2z^2f''(x,\theta)\right)K(z)dz$$

$$=2^{-1}b_n^2f''(x,\theta)\int_{\mathbb{R}}z^2K(z)dz$$

So

$$sup_x\left|E\left(\hat{f}_n(x)\right)-f(x,\theta)\right|\leq 2^{-1}b_n^2sup\left|f''(x,\theta)\right|\int_{\mathbb{R}}z^2K(z)dz$$

$$=0\left(b_n^2\right)\ \ when\ n\to\infty$$

So

$$E_{n2}(\theta)\to 0\ \ when\ n\to\infty$$

Furthermore, we have $\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)=b_n^{-1}\int_{\mathbb{R}}K\frac{x-y}{b_n}d(F_n(y)-F(y)$

where $F_n(.)$ and F(.) are respectively the empirical distribution function and distribution function of the process.

By integration by part, we have

$$\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)=-b_n^{-1}\int_{\mathbb{R}}K'(z)\left(F_n(x-b_nz)-F(x-b_nz)\right)dz$$

$$sup_x\left|\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right|=b_n^{-1}sup_x\left|F_n(x)-F(x)\right|\int_{\mathbb{R}}\left|K'(z)\right|dz$$

From Ho and Hsing [9,10] (theorem 2.1 and remark 2.2) and assumptions (A2) and (A4), we have

$$n^{\lambda+1/2}L^{-1}(n)sup_x\left|F_n(x)-F(x)\right|\to^{\mathfrak{D}}\left|\Phi\right|sup_x\left(f(x)\right)$$

where Φ is a standard Gaussian random variable and !D denotes convergence in distribution.

So

$$b_n^{-1} sup_x \left|\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right| \int_{\mathbb{R}} \left|K'(z)\right| dz \le \frac{1}{n^{\lambda+1/2} b_n L^{-1}(n)} sup_x \left| n^{\lambda+1/2} b_n L^{-1}(n) F_n(x) - F(x)\right| \times \int_{\mathbb{R}} \left|K'(z)\right| dz$$
$$\simeq \left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-1} sup_x \left|f(x)\right| \left|Ö\right| \int_{\mathbb{R}} \left|K'(z)\right| dz$$

where $\simeq$. For all ξ >0,

$$Prob(\left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-1} sup_x \left|f(x)\right| \left|\Phi\right| > \xi) \le \left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-2} (sup_x \left|f(x)\right|)^2 \times Var\left(\left|\Phi\right|\right) \times \frac{1}{\xi^2}$$

The convergence of $\left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-2} (sup_x \left|f(x)\right|)^2 \times Var(\left|\Phi\right|)$ depends on the convergence of $\left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-2}$.

So under assumptions $b_n \to 0, n \to \infty, nb_n \to \infty \, and \, n^{\iota} b_n \to \infty$ for $3 \le \iota \le 5$, we have

$$\begin{aligned}\left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-2} &= \left[n^{\lambda+1/2} b_n\right]^{-2} L^2(n) \\ &= o\left(\left[n^{\lambda+1/2} b_n\right]^{-2}\right) \\ &\to 0 \;\; when \;\; n \to \infty.\end{aligned}$$

We have $(sup_x \left|f(x)\right|)^2 < \infty \, and \, Var\left(\left|\Phi\right|\right) < \infty,$ so

$$\left(n^{\lambda+1/2} b_n L^{-1}(n)\right)^{-1} sup_x \left|f(x)\right| \left|Ö\right| \to_p 0 \;\; when \, n \to \infty.$$

So $E_{n1} \to^p 0 \;\; when \;\; n \to \infty.$

Then $E_n \to^p 0 \;\; when \;\; n \to \infty.$

$$\begin{aligned}D_n(\theta) &= 2^{-1} v_n J_n \int_{\mathbb{R}} \dot{R}_\theta(x,\theta) f^{-1/2}(x,\theta) \left[\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right] dx + 2^{-1} v_n J_n \int_{\mathbb{R}} \dot{R}_\theta(x,\theta) f^{-1/2}(x,\theta) [E\left(\hat{f}_n(x)\right) - f(x,\theta)] dx \\ &= D_{n1}(\theta) + D_{n2}(\theta)\end{aligned}$$

With

$$D_{n1}(\theta) = 2^{-1} v_n J_n \int_{\mathbb{R}} \dot{R}_\theta(x,\theta) f^{-1/2}(x,\theta) \left[\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right] dx$$

and

$$D_{n1}(\theta) = 2^{-1} v_n J_n \int_{\mathbb{R}} \dot{R}_\theta(x,\theta) f^{-1/2}(x,\theta) \left[E\left(\hat{f}_n(x)\right) - f(x,\theta)\right] dx.$$

Under assumptions (A1) − (A2) we apply Taylor-Lagrange formula in $D_{n2}(\theta) \to 0 \; when \; n \to \infty.$ order 2 and assumption (A4), we have

Furthermore, from propositions 1, 2 and 3, we have

Part (a)

$$J_n\left[\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right] \to^{\mathcal{L}} N\left(0; \sum^2(x)\right)$$

or

$$J_n\left[\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right] \Rightarrow U(x),$$

Where $\sum^2(x)$ and U(x) take values according to the different points of the proof of lemma 3:

$$\sum^2(x) = \begin{cases} f(x,\theta) \int_{\mathbb{R}} K^2 du & in(i) \\ \left|f^{(1)}(x,\theta)\right|^2 & in(ii) \\ \left|f^{(r)}(x,\theta) k_{r-1}\right|^2 & in(iii) \\ \sigma^2(x,c) & in(v) \end{cases}$$

and

$$U(x)=\begin{cases}(-1)^r Z_{r,\lambda} f^{(r)}(x,\theta) & in(iv)\\ \sum_{j=0}^{r-1}\left[\frac{C(\lambda,r-j)}{C^{r-j-1}(\lambda,1)}\right]^{1/2}\times\frac{k_j}{C^{r-j-1}(\lambda,1)}Z_{r-j,\lambda}f^{(r)}(x,\theta). & in(vi)\end{cases}$$

Here $Z_{r,\lambda}$ is the Multiple Wiener-Itô Integral defined in the relation (9) of section 1.1 and $\sigma^2(x,c)$ is defined in the first point of proposition 3. Denote by $Y_1(x,\theta)=\dot{R}_\theta(x)f^{-1/2}(x,\theta)$

$$Y_1(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\to^{\mathcal{L}} N\left(0,Y_1(x,\theta)\sum^2(x)Y_1^t(x,\theta)\right)$$

or

$$Y_1(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\Rightarrow Y_1(x,\theta)U(x)Y_1^t(x,\theta).$$

We deduce that

$$\int_{\mathbb{R}}Y_1(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx\to^{\mathcal{L}} N\left(0,\int_{\mathbb{R}}Y_1(x,\theta)\sum^2(x)Y_1^t(x,\theta)\right)dx.$$

or

$$\int_{\mathbb{R}}Y_1(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx\Rightarrow\int_{\mathbb{R}}Y_1(x,\theta)U(x)Y_1^t(x,\theta)dx$$

Were call that $v_n\to^p 0\ when\ n\to\infty$, then $D_{n1}\to^p 0\ when n\to\infty$. So we conclude that $D\ \ \to\ 0\ when n\to\infty$

Part (b)

$$J_nL_n(\theta)=2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f(x,\theta)^{-1/2}\left[\hat{f}_n(x)-f(x,\theta)\right]dx+2^{-1}J_n\left(\int_{\mathbb{R}}\rho(x,\theta)f(x,\theta)^{-1/2}\frac{(\hat{f}_n(x)-f(x,\theta))^2}{(\hat{f}_n^{1/2}(x)-f(x,\theta)^{1/2})^2}dx\right)$$
$$=D_n'(\theta)+E_n'(\theta)$$

with

$$D_n'(\theta)=2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f(x,\theta)^{-1/2}\left[\hat{f}_n(x)-f(x,\theta)\right]dx$$

and

$$E_n'(\theta)=2^{-1}J_n\left(\int_{\mathbb{R}}\rho(x,\theta)f(x,\theta)^{-1/2}\frac{(\hat{f}_n(x)-f(x,\theta))^2}{(\hat{f}_n^{1/2}(x)-f(x,\theta)^{1/2})^2}dx\right)$$

From part (a), the proof of $E_n'(\theta)$ is the same as the proof of $E_n(\theta)$. Were place $\dot{R}_\theta(x,\theta)$ by (x,θ). Then, $E_n'(\theta)\to^p 0$, $when\ \ n\to\infty$. Hence it suffices to prove that the limiting distribution of $J_n\left[\hat{\theta}_n-\theta\right]$ is the same as the limiting distribution of $D_n'(\theta)$ .Since $\hat{f}_n(x)-f(x,\theta)=\left(\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right)+\left(E\left(\hat{f}_n(x)\right)-f(x,\theta)\right)$,

then

$$D_n'(\theta)=2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f^{-1/2}(x,\theta)\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx+2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f^{-1/2}(x,\theta)\left[E\left(\hat{f}_n(x)\right)-f(x,\theta)\right]dx=G_n(\theta)+G_n'(\theta)$$

with

$$G_n(\theta)=2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f^{-1/2}(x,\theta)\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx$$

and

$$G_n'(\theta)=2^{-1}J_n\int_{\mathbb{R}}\rho(x,\theta)f^{-1/2}(x,\theta)\left[E\left(\hat{f}_n(x)\right)-f(x,\theta)\right]dx$$

From the proof of lemma 3 (part (b)), we have:

$$J_n\left(E\left(\hat{f}_n(x)\right)-f(x,\theta)\right)\to 0,\ when\ n\to\infty,$$

then

$$G_n'(\theta)\to 0,\ \ when\ n\to\infty$$

From propositions 1, 2 and 3, we have:

$$J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\to^{\mathcal{L}} N\left(0;\sum^2(x)\right)$$

or

$$J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\Rightarrow U(x).$$

Denote by $Y(x,\theta)=\rho(x,\theta)f^{-1/2}(x,\theta)$

$$Y(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\to^{\mathcal{L}} N\left(0,Y(x,\theta)\sum^2(x)Y^t(x,\theta)\right)$$

or

$$Y(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]\Rightarrow Y(x,\theta)U(x)Y^t(x,\theta))$$

We deduce that

$$\int_{\mathbb{R}}Y(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx\to^{\mathcal{L}} N\left(0,\int_{\mathbb{R}}Y(x,\theta)\sum^2(x)Y^t(x,\theta)\right)dx.$$

or

$$\int_{\mathbb{R}}Y(x,\theta)J_n\left[\hat{f}_n(x)-E\left(\hat{f}_n(x)\right)\right]dx\Rightarrow\int_{\mathbb{R}}Y(x,\theta)U(x)Y^t(x,\theta)dx$$

So,

$$G_n(\theta)\to^{\mathcal{L}} N\left(0;\int_{\mathbb{R}}Y(x,\theta)\sum^2(x)Y^t(x,\theta)dx.\right)$$

or

$$G_n(\theta)\Rightarrow\int_{\mathbb{R}}Y(x,\theta)U(x)Y^t(x,\theta)dx.$$

Then,

$$D_n^{'}(\theta)\to^{\mathcal{L}} N\left(0;\int_{\mathbb{R}}Y(x,\theta)\sum^2(x)Y^t(x,\theta)dx.\right)$$

or

$$D_n^{'}(\theta)\int_{\mathbb{R}}Y(x,\theta)U(x)Y^t(x,\theta)dx.$$

CONCLUSION

We conclude that we have either an asymptotic normal distribution or an asymptotic process towards the Multiple Wiener-Itô Integral.

REFERENCES

1. Tae-Sung K et al. A central limit theorem for the stationary multivariate linear process generated by associated random vectors. Commun Korean Math Soc. 17 2002;17:95-102.
2. Beran R. Minimum Hellinger distance estimates for parametric models. Ann Stat. 1977;5:445-463.
3. Beran R. An efficient and robust adaptive estimator of location. Ann Stat. 1978;6:292-313.
4. Beran R. Efficient robust estimates in parametric models. Z Walrsch Verw Gebiete. 1985;55:91-108.
5. Bitty AL and Hili O. Hellinger distance estimates of long memory linear processes. CR Acad Sci Paris Ser. 2010;348:445-448.
6. Cacoullos T. Estimation of a multi-variate density. Ann Inst Statist Math. 1966;18:179-189.
7. Prazen E. Estimation of a probability density function and mode. Ann Math Statist.1962;33:1065-1076.
8. Rosenblatt M. Remarks on some non-parametric estimates of a density function. Ann Math Statist. 1985;27:832-837.
9. Ho HC and Hsing T. On the asymptotic expansion of the empirical process of long memory moving averages. Ann Statist. 1996;24:992-1024.
10. Ho HC and Hsing T. Limit theorems for functionals of moving averages. Ann Probab. 1997;25:1636- 1669.