STATISTICA, anno LXIII, n. 1, 2003

ON THE RATE OF CONVERGENCE TO THE NORMAL LAW OF LSE IN REGRESSION WITH LONG RANGE DEPENDENCE (*)

N. N. Leonenko, E. Taufer

1. INTRODUCTION

Long range dependence random fields arise in applications in the most disparate areas such as astronomy, economics, hydrology and the telecommunications; the study of random processes and fields with long range dependence presents interesting and challenging probabilistic and statistical problems and recent literature has seen an increasing number of papers developing models for the description and analysis of this phenomenon.

In the present paper we will present some results on the rate of convergence to the normal law of the Least Squares Estimators (LSE) of the regression coefficients in models with multidimensional inputs and long range dependence errors.

The same problem, for single input regression, has been considered in Leonenko *et al.* (2000); this work finds its motivation in the fact that the extension to multiple regression is not immediate and some different tools have to be utilized.

Note that we consider regression on continuous homogeneous random fields, in particular, for $n=1$, we can interpret the parameter of the random field as time. It could be pointed out that even if many of today's data set are available virtually in continuous time, practical applications require to consider the observed phenomena at fixed time points. Notwithstanding, procedures of discretization lead sometimes to loss of information on important parameters (see, for example, Leonenko, 1999, pp. 14-16).

Statistical problems related with long range dependent continuous random processes and fields have been considered in the book by Ivanov and Leonenko (1989), Chambers (1996) considers the problem of estimation of continuous parameters in long memory time series models, in Comte (1996) we find an analysis of different methods of simulation and estimation for long memory continuous models. Leonenko and Benšić (1996a, 1998) and Leonenko and Taufer (2001) present Gaussian and non-Gaussian limit distributions of univariate and multi-

(*) This work was partially supported by the Australian Research Council Grant A69804041.

variate regression for long memory random fields and processes, their results have been obtained by the methods presented in the works of Dobrushin and Major (1979) and Taqqu (1979).

For other results of interest here see Yajima (1988, 1991), Künsch et al. (1993), Dahlhaus (1995), Robinson and Hidalgo (1997), Deo (1997), Deo and Hurvich (1998) which consider regression models with long memory errors in discrete time.

The paper is organized as follows. In Section 2 we will state our model and assumptions exactly and formulate the main result which shows the rate of convergence of Kolmogorov's distance between the distribution of the normalized LSE and the standard normal distribution. The proof of the main result, together with some preparatory lemmas, is given in Section 3. Section 4 contains the discussion for an extension to a wider case which can be done at the price of a slower convergence rate.

We do not take into consideration here the problem of estimation of the dependence index (or Hurst parameter), for this, see Giraitis and Koul (1997) and their references.

2. MAIN RESULTS

Let $\Re^n$, $n \geq 1$, be a n-dimensional Euclidean space, $\Delta \subset \Re^n$ be a bounded and convex subset containing the origin and $\Delta(T)$ be the image of the set Δ under the homotetic transformation with center at the origin and coefficient $T > 0$, that is $\Delta(T) = \{x \in \Re^n : x / T \in \Delta\}$. Practical situations often claim that Δ is a parallelepiped or a ball but we can allow this weaker condition.

Assumption 1. Consider the regression model of the form

$$\zeta(x) = \boldsymbol{\theta}'\boldsymbol{g}(x) + \eta(x), \quad x \in \Re^n$$

where $\boldsymbol{g}(x) = [g_1(x),...,g_q(x)]'$ is a known vector function whose coordinate functions $g_i(x)$, $i = 1,...,q$ form a linearly independent set of real functions positive on Δ and square integrable over the same set for all bounded $\Delta \subset \Re^n$ containing the origin. $\boldsymbol{\theta} = [\theta_1,...,\theta_q]$ is an unknown vector of parameters and $\eta(x)$ is an homogeneous random field of errors with $\mathbf{E}\eta(x) = 0$ and $\mathbf{E}\eta(x)^2 < \infty$.

Assumption 2. Let $\xi(\omega, x) = \xi(x)$, $x \in \Re^n$, $\omega \in \Omega$, be a real valued measurable mean square continuous homogeneous Gaussian random field on the probability space (Ω, F, P) with $\mathbf{E}\xi(x) = 0$, $\mathbf{E}\xi(x)^2 = 1$ and correlation function

$$B(x) = \mathbf{E}\xi(0)\xi(x) = |x|^{-\alpha} L(|x|) a\left(\frac{x}{|x|}\right), \quad 0 < \alpha < n,$$

where $a(\cdot)$ is a continuous function on the n-dimensional sphere $s_{n-1}(1) = \{ x \in \Re^n : |x| = 1\}$, and $L(t) > 0$, $t > 0$ is a slowly varying function at infinity ($\lim_{t\to\infty}(L(ts)/L(t)) = 1$, for every $s > 0$) bounded on each finite interval.

Assumption 3. Let $\eta(x) = G(\xi(x))$, $x \in \Re^n$, where $\xi(x)$ is a random field satisfying Assumption 2, and $G(\cdot)$ is a non-random measurable function such that $\mathbf{E}G(\xi(x)) = 0$ and $\mathbf{E}G^2(\xi(x)) < \infty$, $x \in \Re^n$.

Note that the marginal distributions of a field $\eta(x)$, $x \in \Re^n$, satisfying Assumption 3 need not be Gaussian. Moreover, under Assumption 2 we have $\int_{\Re^n} |B(x)| dx = \infty$. Typical examples of correlation functions satisfying Assumption 2 are the following:

$$B_1(x) = (1 + |x|^2)^{-\alpha/2}, \quad 0 < \alpha < n$$

and

$$B_2(x) = (1 + |x|^\alpha)^{-1}, \quad 0 < \alpha < 2.$$

The first function is known as characteristic function of the multivariate Bessel distribution (see, for example, Fang et al., 1990, p. 69); the second one is known as characteristic function of the multivariate Linnik distribution (see, Anderson, 1992 or Ostrovskii, 1995).

Our aim is to study the rate of convergence to the normal law of the LSE of the vector $\boldsymbol{\theta}$ which can be found by minimizing

$$\int_{\Delta(T)} [\zeta(x) - \boldsymbol{\theta}'\boldsymbol{g}(x)]^2 dx$$

with respect to $\boldsymbol{\theta}$. The final form of LSE is given by (see Leonenko and Šilac-Benšić, 1998)

$$\hat{\boldsymbol{\theta}}_T = \boldsymbol{Q}_T^{-1} \int_{\Delta(T)} \boldsymbol{g}(x)\zeta(x)dx = \boldsymbol{\theta} + \boldsymbol{Q}_T^{-1} \int_{\Delta(T)} \boldsymbol{g}(x)G(\xi(x))dx \tag{1}$$

where the integral is taken with respect to every element of the matrices and

$$\mathbf{Q}_T = \int_{\Delta(T)} \mathbf{g}(x)\mathbf{g}(x)' dx.$$

The existence of $\boldsymbol{Q}_T^{-1}$ follows from linear independence and square integrability of $g_1(x),...,g_q(x)$. It is straightforward to verify that

$$E(\hat{\boldsymbol{\theta}}_T) = \boldsymbol{\theta}$$

and that

$$\begin{aligned}\text{Var}(\hat{\boldsymbol{\theta}}_T) &= E[\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}][\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}]' \\ &= \mathbf{Q}_T^{-1} \int_{\Delta(T)} \int_{\Delta(T)} \mathbf{g}(x)\mathbf{g}(y)'\mathbf{E}[G(\xi(x))G(\xi(y))]dx\ dy\ \mathbf{Q}_T^{-1}.\end{aligned} \tag{2}$$

Let

$$H_m(u) = (-1)^m \exp\left\{\frac{u^2}{2}\right\} \frac{d^m}{du^m} \exp\left\{\frac{-u^2}{2}\right\},$$

$u \in \Re^1$, $m = 0,1,...,$ be the Chebyshev-Hermite polynomials with the leading coefficient equal to 1; they are known to form a complete orthogonal system in the Hilbert space $L_2(\Re^1, \phi(u)du)$, where $\phi(u) = (2\pi)^{-1/2} \exp\{-u^2/2\}$, $u \in \Re^1$. Note that $H_0(u) = 1$, $H_1(u) = u$, $H_2(u) = u^2 - 1 \ldots$

It is well known that (see, for example, Ivanov and Leonenko 1989, p. 55) if (ξ, η) is a Gaussian vector with $\mathbf{E}\xi = \mathbf{E}\eta = 0$, $\mathbf{E}\xi^2 = \mathbf{E}\eta^2 = 1$, $\mathbf{E}\xi\eta = \rho$, then for all $m, p > 0$

$$\mathbf{E}H_m(\xi)H_p(\eta) = \delta_m^p \rho^m m! \tag{3}$$

where δ_m^p is the usual Kronecker's delta.

Under Assumption 3 the function $G(u), u \in \Re^1$ allows the following representation in the Hilbert space $L_2(\Re^1, \phi(u)du)$:

$$G(u) = \sum_{m \geq 0} \frac{C_m}{m!} H_m(u), \;\; C_m = \int_{\Re} G(u)H_m(u)\phi(u)du \tag{4}$$

and by Parseval's relation:

$$\mathbf{E}G^2(\xi(0)) = \sum_{m \geq 0} \frac{C_m^2}{m!} H_m(u) = \int_{\Re} G^2(u)\phi(u)du < \infty \tag{5}$$

Note that $C_0 = \mathbf{E}G^2(\xi(0)) = 0$. From (1) - (4) we obtain:

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}_T) = \mathbf{Q}_T^{-1} \sum_{m \geq 1} \boldsymbol{\Psi}_{m,T} \mathbf{Q}_T^{-1} \tag{6}$$

where

$$\boldsymbol{\Psi}_{m,T} = \frac{C_m^2}{m!} \int_{\Delta T} \int_{\Delta T} \mathbf{g}(x)\mathbf{g}(y)' B^m(x-y)\, dx\, dy$$

Now we need some extra assumptions upon the regression vector function $\mathbf{g}(\cdot)$ and the covariance function $B(\cdot)$.

Assumption 4. Suppose that $g_i(x) > 0$ for all $x \neq 0$, $i = 1,...,q$ and, for $0 < \alpha < n/m$, $m = 1$ or $m = 2$, the following limits exist and are finite:

$$l_m(\alpha, n)_{ij} = \lim_{T \to \infty} \int_\Delta \int_\Delta \frac{g_i(xT) g_j(yT)}{g_i(T\mathbf{1}_n) g_j(T\mathbf{1}_n)}\, a^m \left(\frac{x-y}{|x-y|} \right) |x-y|^{-m\alpha}\, dx\, dy$$

$i, j = 1...q$, and that $\mathbf{L}_m(\alpha, m) = [l_m(\alpha, n)_{ij}]_{(1 \leq i, j \leq q)}$ is a positive definite matrix ($\boldsymbol{1}_n$ is a *n*-vector of ones).

Assumption 5. Let $m = 1$ or $m = 2$. Suppose that there exist a function $F_m(x, y)$ such that

$$\left| \frac{g_i(xT) g_j(yT)}{g_i(T\mathbf{1}_n) g_j(T\mathbf{1}_n)}\, a^m \left(\frac{x-y}{|x-y|} \right) \frac{1}{|x-y|^{m\alpha}} \left(\frac{L^m |x-y| T}{L^m(T)} - 1 \right) \right| \leq F_m(x, y),$$

$i, j = 1...q$, $0 < \alpha < n/m$, and

$$\int_\Delta \int_\Delta F_m(x, y)\, dx\, dy < \infty .$$

Remark 2.1. Consider the case of polynomial radial regression: $\mathbf{g}(x) = \tilde{\mathbf{g}}(|x|) = (|x|^{\mu_1}, ..., |x|^{\mu_q})'$ with $\mu_1, ...\mu_q > 0$. Suppose Assumption 2 hold with $a(\cdot) \equiv 1$ and $B(x) = B_1(x) = (1 + |x|^2)^{-\alpha/2}$. Thus we have $L(|x|) = |x|^\alpha (1 + |x|^2)^{-\alpha/2}$ $0 < \alpha < n$; then Assumptions 4 and 5 hold with

$$l_m(\alpha,n)_{ij} = \frac{1}{n^{(\mu_i+\mu_j)/2}} \int_\Delta \int_\Delta \frac{|x|^{\mu_i} |y|^{\mu_j}}{|x-y|^{m\alpha}} dx\, dy \quad i,j = 1...q\,. \tag{7}$$

Remark 2.2. We write

$$\mathbf{L}_m(\alpha,n) = \lim_{T\to\infty} \mathbf{D}_T^{-1} \int_{\Delta(T)} \int_{\Delta(T)} \mathbf{g}(xT)\mathbf{g}(yT)'\, a^m\left(\frac{x-y}{|x-y|}\right)|x-y|^{-m\alpha} dx\, dy\mathbf{D}_T^{-1} \tag{8}$$

where

$$D_T = diag[g_1(\mathbf{1}_n T),..., g_q(\mathbf{1}_n T)]$$

Remark 2.3. After the transformation $x = x^* T \in \Delta(T)$, $y = y^* T \in \Delta(T)$, $T > 0$, $x^* \in \Delta$, $y^* \in \Delta$, we obtain the following expressions for the matrix $\mathbf{\Psi}_{m,T}$, $0 < \alpha < n/m$:

$$\mathbf{\Psi}_{m,T} = \frac{C_m^2}{m!} T^{2n-m\alpha} L^m(T)\mathbf{D}_T\mathbf{D}_T^{-1} \int_{\Delta(T)} \int_{\Delta(T)} \mathbf{g}(x^*T)\mathbf{g}(y^*T)' \frac{L^m\left(\left|x^*-y^*\right|T\right)}{L^m(T)} \times$$

$$\times a^m\left(\frac{x^*-y^*}{\left|x^*-y^*\right|}\right)\frac{dx^*\, dy^*}{\left|x^*-y^*\right|^{m\alpha}}\mathbf{D}_T^{-1}\mathbf{D}_T, \quad 0 < \alpha < n/m.$$

Let $m = 1$ or $m = 2$ and $C_m \neq 0$. Then, from Assumptions 4 and 5 and Lebesgue dominated convergence theorem we obtain (for details, see Leonenko and Benšić, 1998):

$$\mathbf{\Psi}_{m,T} = \frac{C_m^2}{m!} T^{2n-m\alpha} L^m(T)\mathbf{D}_T\mathbf{L}_m(\alpha,n)\mathbf{D}_T(1+o(1)), \quad 0 < \alpha < n/m \tag{9}$$

as $T \to \infty$.

To complete our preliminaries, we need to define the Kolmogorov's distance between random vectors. To this end, let

$$\mathbf{\Pi}[\mathbf{a},\mathbf{b}] = \{u \in \Re^q : a_i \leq u_i \leq b_i,\ i = 1,...,q\}$$

be a parallelepiped in $\Re^q$ and let $\mathbf{x}$ and $\mathbf{y}$ be two arbitrary q-dimensional random vectors. Introduce the uniform (or Kolmogorov's) distance between distributions of random vectors $\mathbf{x}$ and $\mathbf{y}$ via the formula:

$$K(\mathbf{x},\mathbf{y}) = \sup_{a,b\in\Re^q} |P(\mathbf{x}\in\mathbf{\Pi}[\mathbf{a},\mathbf{b}]) - P(\mathbf{y}\in\mathbf{\Pi}[\mathbf{a},\mathbf{b}])|$$

We are now ready to formulate the main result of this paper. Let $\mathbf{z}$ be a standard normal random *q*-vector with zero mean and unit covariance matrix and consider the random vector

$$\mathbf{k}_T = \mathbf{\Psi}_{1,T}^{-1/2}\mathbf{Q}_T[\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}],$$

where the LSE $\hat{\boldsymbol{\theta}}_T$ are defined in (1) and $\mathbf{\Psi}_{1,T}^{-1/2}$ is a nonsingular matrix such that

$$(\mathbf{\Psi}_{1,T}^{-1/2})(\mathbf{\Psi}_{1,T}^{-1/2})' = \mathbf{\Psi}_{1,T}^{-1}.$$

Theorem 2.1 describes the rate of convergence to the normal law of the random vector $\mathbf{k}_T$ as $T\to\infty$.

Theorem 2.1 Suppose that Assumptions 1-5 hold for $0<\alpha<n/2$, and

$$C_1 = \int_{\Re} uG(u)\phi(u)du \neq 0$$

then the quantity:

$$\lim_{T\to\infty}\sup\left[\frac{T^\alpha}{L(T)}\right]^{1/3} K(\mathbf{k}_T,\mathbf{z})$$

exists and is bounded by

$$2c_1(q)^{2/3}\left[c_2(q)\; c(G)\; tr[\mathbf{L}_1(\alpha,n)^{-1}\mathbf{L}_2(\alpha,n)]\right]^{1/3}$$

where $\mathbf{L}_1(\alpha,n)$ and $\mathbf{L}_2(\alpha,n)$ are defined by (8) and

$$c(G) = C_1^{-2}\left[\int_{\Re} G^2(u)\phi(u)du - C_1^2\right]$$

$$c_1(q) = \sqrt{2/\pi}\text{, if } q=1 \text{ and } c_1(q) = (q-1)\frac{\Gamma[(q-1)/2]}{\sqrt{2}\Gamma(q/2)}\text{, if } q\geq 2$$

$$c_2(q) = \frac{1}{q}\left(1+\sqrt{q-1}\right)^2$$

3. PROOF OF THE MAIN RESULT

Before proving Theorem 2.1 we mention some preliminary results. The following lemma provides an estimate of the Kolmogorov's distance of a sum of random vectors from a standard Gaussian vector. For its proof see Leonenko and Woyczynski (1998).

Lemma 3.1 Let $\mathbf{x}$, $\mathbf{y}$ be two arbitrary random q-vectors and $\mathbf{z}$ be a standard Gaussian q-vector such that, for all $\mathbf{a},\mathbf{b}\in\Re^q$,

$$|P(\mathbf{x}\in\mathbf{\Pi}[\mathbf{a},\mathbf{b}])-P(\mathbf{z}\in\mathbf{\Pi}[\mathbf{a},\mathbf{b}])|\le K,$$

where $K\ge 0$ is a constant. Then, for any $\varepsilon>0$,

$$K(\mathbf{x}+\mathbf{y},\mathbf{z})\le K+P(\mathbf{y}\notin\mathbf{\Pi}[-\mathbf{1}_q\varepsilon,\mathbf{1}_q\varepsilon])+\varepsilon c_1(q) \tag{10}$$

where $c_1(q)$ is defined in Theorem 2.1 and $\mathbf{1}_q$ is a q-vector of ones.

In the proof of Theorem 2.1 we need an estimate on the tails of the maxima of a general second-order random vector's components which is provided by the following Lemma (see Karlin and Studden, 1966).

Lemma 3.2 Let $\mathbf{v}=[V_1,...V_q]$ be a random q-vector with mean $\mathbf{Ev}=0$ and covariance matrix $\mathbf{Evv}'=\mathbf{\Sigma}=(\sigma_{ij})_{1\le i,j\le q}$, and let $W_i=V_i/(k_i\sigma_i)$ where $\sigma_i^2=\sigma_{ii}$, and $k_1,...,k_q>0$ are some constants. Then

$$P\left(\max_{1\le i\le q}|W_i|\ge 1\right)\le\frac{1}{q^2}\left(\sqrt{s}+\sqrt{(qt-s)(q-1)}\right)^2, \tag{11}$$

where $t=tr\,\mathbf{\Pi}$, $s=\mathbf{1}_q'\,\mathbf{\Pi}\mathbf{1}_q$, $\mathbf{\Pi}=\mathbf{Eww}'=(\pi_{ij})_{1\le i,j\le q}$, $\pi_{ij}=\sigma_{ij}/(\sigma_i\sigma_j k_i k_j)$, and $\mathbf{1}_q$ is a q-vector of ones.

Remark 3.1. From Lemma 3.2 we readily obtain an upper bound for $P(\max_{1\le i\le q}|V_i|\ge 1)$ by setting $k_i=1/\sigma_i$, $i=1,...q$.

The purpose of the next Lemma will be made immediately clear in the subsequent Remark.

Lemma 3.3. Let $\mathbf{\Sigma}=(\sigma_{ij})_{1\le i,j\le q}$ be the variance- covariance matrix of a random q-vector $\mathbf{v}$, t its trace and s the sum of its elements. Then

$$0\le s\le qt.$$

Proof: that $s \geq 0$ is obvious since $s = \mathrm{Var}(\mathbf{1}'_q\ \mathbf{v})$. To prove that $s \leq qt$, use the Cauchy-Schwarz inequality to obtain

$$s \leq \sum_{i,j} \left|\sigma_{ij}\right| \leq \sum_i \sigma_i \sum_j \sigma = \left(\sum_i \sigma_i\right)^2 \tag{12}$$

where $\sigma_i^2 = \sigma_{ii}$. Again, by the Cauchy Schwarz inequality we know that $\left(\sum_i a_i b_i\right)^2 \leq \sum_i a_i^2 \sum_i b_i^2$ where a_i, b_i are positive numbers. Set $a_i = \sigma_i$ and $b_i = 1$, $\forall i$. Then

$$\left(\sum_i \sigma_i\right)^2 \leq \sum_i \sigma_i^2 \sum_i 1 = qt .$$

Remark 3.2. From Lemma 3.3 we can obtain a less tight version of Lemma 3.2:

$$P\left(\max_{1 \leq i \leq q} \left|W_i\right| \geq 1\right) \leq \frac{1}{q^2}\left(\sqrt{qt} + \sqrt{qt(q-1)}\right)^2 = \frac{t}{q}(1 + \sqrt{q-1})^2 .$$

For later convenience let

$$c_2(q) = \frac{1}{q}\left(1 + \sqrt{q-1}\right)^2 .$$

Proof of Theorem 2.1. Formula 4 implies the following expansion in the Hilbert space $L_2(\Omega)$:

$$G(\xi(x)) = \sum_{m \geq 1} \frac{C_m}{m!} H_m(\xi(x)) ,$$

we now consider the random vectors

$$\boldsymbol{\eta}_{m,T} = \int_{\Delta(T)} \mathbf{g}(x) H_m(\xi(x)) dx, \quad m = 1, 2...$$

In order to apply Lemma 3.1, we represent $\mathbf{k}_T$ as

$$\mathbf{k}_T = \boldsymbol{\Psi}_{1,T}^{-1/2}\left[\mathbf{x}_T + \mathbf{y}_T\right]$$

where

$$\mathbf{x}_T = C_1 \boldsymbol{\eta}_{1,T},, \qquad \mathbf{y}_T = \sum_{m\geq 2} \frac{C_m}{m!} \boldsymbol{\eta}_{m,T}.$$

Note that $\mathbf{x}_T$ is a Gaussian random vector with $\mathbf{E}\mathbf{x}_T = 0$ and $\mathbf{E}\mathbf{x}_T\mathbf{x}_T' = \boldsymbol{\Psi}_{1,T}$. So we have

$$K(\boldsymbol{\Psi}_{1,T}^{-1/2}\mathbf{x}_T, \mathbf{z}) = 0$$

and we may choose K=0 in Lemma 3.1. We are left with the term

$P(\boldsymbol{\Psi}_{1,T}^{-1/2}\mathbf{y}_T \notin \boldsymbol{\Pi}[-\mathbf{1}_q\varepsilon, \varepsilon\mathbf{1}_q]) = P(\boldsymbol{\Psi}_{1,T}^{-1/2}\mathbf{y}_T \frac{1}{\varepsilon} \notin \boldsymbol{\Pi}[-\mathbf{1}_q, \mathbf{1}_q]) \leq P(\max_{1\leq i\leq q}|V_i| \geq 1)$ where $\mathbf{v} = \boldsymbol{\Psi}_{1,T}^{-1/2}\mathbf{y}_T \frac{1}{\varepsilon}$. We can then find an upper bound for $P(\max_{1\leq i\leq q}|V_i| \geq 1)$ by using Lemma 3.2 as indicated in Remark 3.1. Also, Remark 3.2 shows us that we need only to evaluate the trace of $\mathbf{E}\mathbf{v}\mathbf{v}'$. We have:

$$t = tr\mathbf{E}\mathbf{v}\mathbf{v}' = \frac{1}{\varepsilon^2} tr\left[\boldsymbol{\Psi}_{1,T}^{-1/2}\mathbf{E}\mathbf{y}_T\mathbf{y}_T' \boldsymbol{\Psi}_{1,T}^{-1/2}\right] = \frac{1}{\varepsilon^2} tr\left[\boldsymbol{\Psi}_{1,T}^{-1}\mathbf{E}\mathbf{y}_T\mathbf{y}_T'\right]$$

For notational convenience, denote $\mathbf{E}\mathbf{y}_T\mathbf{y}_T' = \sum_{m\geq 2}\boldsymbol{\Psi}_{m,T}$ as $\boldsymbol{\Sigma}$. Since $|B(x)| \leq 1$ we know that $B^2(x-y) \geq B^m(x-y),\ m \geq 2$. So that, for $0 < \alpha < n/2$,

$$\begin{aligned}\boldsymbol{\Sigma} &= \sum_{m\geq 2} \frac{C_m^2}{m!} \int_{\Delta(T)}\int_{\Delta(T)} \mathbf{g}(x)\mathbf{g}(y)' B^m(x-y)\, dx\, dy \\ &\leq \left[\sum_{m\geq 2} \frac{C_m^2}{m!}\right] \int_{\Delta(T)}\int_{\Delta(T)} \mathbf{g}(x)\mathbf{g}(y)' B^2(x-y)\, dx\, dy = \boldsymbol{\Sigma}^* \quad \text{(say).}\end{aligned}$$

Where the inequality sign refers to any single element of the matrices. Also, we have that $\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} = \sum_{m\geq 2} \frac{C_m^2}{m!} \int_{\Delta(T)}\int_{\Delta(T)} \mathbf{a}'\mathbf{g}(x)\mathbf{g}(y)'\mathbf{a}B^m(x-y)\, dx\, dy$ where $\mathbf{a}$ is a vector of real numbers; hence, by the same reasoning and from the fact that $\boldsymbol{\Psi}_{m,T} = \mathrm{Var}\left(\frac{C_m}{m!}\boldsymbol{\eta}_{m,T}\right)$ is a positive definite matrix it can be seen that $\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} \leq \mathbf{a}'\boldsymbol{\Sigma}^*\mathbf{a}$ for any vector $\mathbf{a}$ of real numbers, i.e. $\boldsymbol{\Sigma}^* - \boldsymbol{\Sigma}$ is a positive definite matrix.

Then, in order to find an upper bound for $t = tr\mathbf{E}\mathbf{v}\mathbf{v}'$ we reason as follows: from positive definiteness there exists an orthogonal matrix $\mathbf{P}$ such that $\mathbf{\Psi}_{1,T} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$ where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{\Psi}_{1,T}$ which are all real and positive. Hence,

$$tr\left[\mathbf{\Psi}_{1,T}^{-1}\mathbf{\Sigma}^* - \mathbf{\Sigma}\right] = tr\left[\mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}'(\mathbf{\Sigma}^* - \mathbf{\Sigma})\right] = tr\left[\mathbf{\Lambda}^{-1}\mathbf{P}'(\mathbf{\Sigma}^* - \mathbf{\Sigma})\mathbf{P}\right] = \sum_{i=1}^{q} \lambda_i^{-1}\mathbf{p}_i'(\mathbf{\Sigma}^* - \mathbf{\Sigma})\mathbf{p}_i \geq 0$$

where $\lambda_i^{-1} \geq 0$ are the diagonal elements of $\mathbf{\Lambda}^{-1}$, $\mathbf{p}_i$ is the i-th column of $\mathbf{P}$. The last inequality follows from positive definiteness of $\mathbf{\Sigma}^* - \mathbf{\Sigma}$. Using this last result we obtain the following estimate

$$t = \frac{1}{\varepsilon^2} tr\left[\mathbf{\Psi}_{1,T}^{-1}\mathbf{\Sigma}\right] = \frac{1}{\varepsilon^2} tr\left[\mathbf{\Psi}_{1,T}^{-1}\mathbf{\Sigma}^*\right] = \frac{1}{\varepsilon^2}\left[\sum_{m\geq 2}\frac{C_m^2}{m!}\right] tr\left[\mathbf{\Psi}_{1,T}^{-1}\frac{2}{C_2^2}\mathbf{\Psi}_{2,T}\right].$$

Using (9) we have that

$$t = \lim_{T\to\infty} tr\left[\frac{2C_1^2}{C_2^2}\frac{T^\alpha}{L(T)}\mathbf{\Psi}_{1,T}^{-1}\mathbf{\Psi}_{2,T}\right] = tr\left[\mathbf{L}_1^{-1}(\alpha,n)\mathbf{L}_2(\alpha,n)\right], \quad 0<\alpha<n/2.$$

Hence we have the following upper bound for t:

$$\frac{1}{\varepsilon^2} c(G)\frac{L(T)}{T^\alpha} tr\left[\mathbf{L}_1^{-1}(\alpha,n)\mathbf{L}_2(\alpha,n)\right], \quad 0<\alpha<n/2 \tag{13}$$

as $T \to \infty$, where

$$c(G) = C_1^{-2}\sum_{m\geq 2}\frac{C_m^2}{m!} = C_1^{-2}\left[\int_{\Re} G^2(u)\phi(u)du - C_1^2\right].$$

Finally, using Lemma 3.1 with $\mathbf{x} = \mathbf{\Psi}_{1,T}^{-1/2}\mathbf{x}_T$ and $\mathbf{y} = \mathbf{\Psi}_{1,T}^{-1/2}\mathbf{y}_T$ we obtain from (13) that for any $\varepsilon > 0$:

$$K(\mathbf{k}_T, \mathbf{z}) \leq \varepsilon c_1(q) + \frac{1}{\varepsilon^2} c(G)c_2(q)\frac{L(T)}{T^\alpha} tr\left[\mathbf{L}_1^{-1}(\alpha,n)\mathbf{L}_2(\alpha,n)\right]. \tag{14}$$

In order to minimize the r.h.s. of the inequality, set

$$\varepsilon = \left[\frac{c(G)c_2(q) tr\left[\mathbf{L}_1^{-1}(\alpha,n)\mathbf{L}_2(\alpha,n)\right]}{c_1(q)}\frac{L(T)}{T^\alpha}\right]^{1/3}$$

and substitute into (14) to obtain the following

$$K(\mathbf{k}_T,\mathbf{z}) \le 2c_1(q)^{2/3}[c(G)c_2(q)tr(\mathbf{L}_1^{-1}(\alpha,n)\mathbf{L}_2(\alpha,n))]^{1/3}\left[\frac{L(T)}{T^{\alpha}}\right]^{1/3}.$$

4. EXTENSIONS AND GENERALIZATIONS

As follows from the results of Leonenko and Šilac-Benšić (1998), the asymptotic normality of the normalized LSE takes place for all $\alpha \in (0,n)$ (see Assumption 1) if $C_1 \neq 0$, whereas Theorem 2.1 gives the convergence rate to Kolmogorov's distance only for $\alpha \in (0,n/2)$.

Nevertheless, our method is applicable also to the broader interval $\alpha \in (0,n)$ at the price of a slower convergence rate.
For simplicity we consider the homogeneous isotropic random field (the function $a(\cdot) \equiv 1$ in Assumption 2) and the case of radial regression function: $\mathbf{g}(x) = \tilde{\mathbf{g}}(|x|),\ x \in \Re^n$.

We consider now the case $\Delta(T) = v(T) = \{x \in \Re^n : |x| < T\},\ T \to \infty$. Thus the random field $\zeta(x) = \boldsymbol{\theta}'\mathbf{g}(x) + \eta(x)$ is observed on the ball $v(T)$.

Assumption 6. Let $\xi(x)$, $x \in \Re^n$, be a real valued mean square continuous homogeneous isotropic Gaussian field with $\mathbf{E}\xi(x) = 0$, $\mathbf{E}\xi(x)^2 = 1$ and correlation function

$$B(x) = \tilde{B}(|x|) = \mathbf{E}\xi(0),\xi(x) \to 0$$

monotonically as $|x| \to \infty$, and $\eta(x) = G(\xi(x))$, where $\mathbf{E}G(\xi(x)) = 0$ and $\mathbf{E}G^2(\xi(x)) < \infty$, $x \in \Re^n$.

Assumption 7. Suppose that for the regression function it holds $\mathbf{g}(x) = \tilde{\mathbf{g}}(|x|),\ x \in \Re^n$ such that $\tilde{g}_i(|x|) > 0$, $i = 1,\ldots,q$, if $|x| \neq 0$, and $\tilde{g}_i(|x|) \le \tilde{g}_i(|y|)$, $i = 1,\ldots,q$, for $|x| \le |y|$.

Assumption 8. There exists a $\delta \in (0,1)$ such that

$$\gamma_{T,ij} = T^{-n(1+\delta)} \int_{v(T)} \int_{v(T)} \frac{g_i(x) g_j(y)}{g_i(T) g_j(T)} \tilde{B}(|x-y|)\, dx\, dy$$

$$= T^{n(1-\delta)} \tilde{B}(T) \int_{v(T)} \int_{v(T)} \frac{g_i(xT) g_j(yT)}{g_i(T) g_j(T)} \frac{\tilde{B}(T|x-y|)}{\tilde{B}(T)}\, dx\, dy \to \infty \quad \forall i,j = 1,...q$$

as $T \to \infty$.

Note that if Assumption 2 (with $a(\cdot) \equiv 1$), and Assumptions 4 and 5 (with $m = 1$) hold, then $\gamma_{T,ij} \to \infty$ as $T \to \infty$ for all $i,j = 1,...,q$. Thus the random field $\xi(x)$, $x \in \Re^n$ satisfying Assumption 8 is a random field with long range dependence.

Remark 4.1. We write

$$\mathbf{\Gamma}_T = [\gamma_{T,ij}]_{1 \le i,j \le q} = T^{-n(1+\delta)} \mathbf{D}_T \int_{v(T)} \int_{v(T)} \mathbf{g}(x) \mathbf{g}(y)' \tilde{B}(|x-y|)\, dx\, dy\, \mathbf{D}_T^{-1}$$

We have the following result:

Theorem 4.1 Suppose that assumptions 6-8 hold, and

$$C_1 = \int_{\Re} u G(u) \phi(u) du \neq 0$$

then the following quantity:

$$\limsup_{T\to\infty} \left[\frac{c(n) 2^{n\delta}}{n} \mathbf{1}_q' \mathbf{\Gamma}_T^{-1} \mathbf{1}_q + \tilde{B}(T^{\delta}) \right]^{-1/3} K(\mathbf{k}_T, \mathbf{z})$$

exists and is bounded by

$$2[\, c(G)\, c_2(q)]^{1/3} c_1(q)^{2/3}$$

where $c_1(q)$ and $c_2(q)$ have been defined in Theorem 2.1, and

$$c_1(n) = \frac{4\pi^n \Gamma^{-2}(n/2)}{n}$$

Before proving the theorem, we need some preliminaries. With this in mind, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two independent random vectors selected in accordance to the uniform law on the ball $v(T) \in \Re^n$. Then (see Ivanov and Leonenko, 1989, p. 25) the density function $\rho_T(u)$ of the Euclidean distance $|\mathbf{u}_1 - \mathbf{u}_2|$ between $\mathbf{u}_1$ and $\mathbf{u}_2$ is

$$\rho_T(u) = T^{-n} n u^{n-1} I_{1-(u/2T)^2}\left(\frac{n+1}{2}, \frac{1}{2}\right), \quad 0 < u < 2T ,$$

where

$$I_\mu(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^u t^{a-1}(1-t)^{b-1} dt , \quad a > 0 , \quad \mu \in [0,1]$$

is the incomplete Beta function. Using randomization we obtain for every function $f(|x-y|), x \in \Re^n, y \in \Re^n$:

$$\begin{aligned} \int_{v(T)} \int_{v(T)} f(|x-y|)\, dx\, dy &= |v(T)|^2\, \mathbf{E} f\left(|\mathbf{u}_1 - \mathbf{u}_2|\right) \\ &= T^{2n} |v(1)|^2 \int_0^{2T} f(u) \rho_T(u)\, du \qquad (15) \\ &= c(n) T^n \int_0^{2T} z^{n-1} f(z) I_{1-(z/2T)^2}\left(\frac{n+1}{2}, \frac{1}{2}\right) dz , \end{aligned}$$

where $c(n) = 4\pi^n \Gamma^{-2}(n/2)/n$ and $|v(T)|$ is the volume of a ball $v(T)$.

Proof of Theorem 4.1: we follow the scheme of proof of Theorem 2.1 including the necessary modifications. Let us introduce the sets

$$A_1 = \{(x,y) : |x-y| < T^\delta\}$$
$$A_2 = \{(x,y) : |x-y| \geq T^\delta\}$$

Following the results of the previous section, in order to find an upper bound for $t = tr[\mathbf{\Psi}_1(T)^{-1} \mathbf{\Sigma}]$ consider

$$tr\left[\mathbf{\Psi}_1(T)^{-1} \frac{2!}{C_2^2} \mathbf{\Psi}_2(T)\right] = tr\left[\mathbf{\Psi}_1(T)^{-1}\left(\iint_{A_1} + \iint_{A_2}\right) \mathbf{g}(x)\mathbf{g}(y)' \tilde{B}(|x-y|)\, dx\, dy\right] \quad (16)$$

Consider first the set A_1, we have $B^2(\cdot) \leq 1$ and then, for the first term on the r.h.s. of (16), using result (15) with $f(|x-y|) = \mathbf{1}_{A_1}$ we have the estimate

$$tr\left[\boldsymbol{\Psi}_1(T)^{-1}\left(\iint_{A_1}\mathbf{g}(x)\mathbf{g}(y)'\tilde{B}^2(|x-y|)dxdy\right)\right]\leq tr\left[\boldsymbol{\Psi}_1(T)^{-1}\tilde{\mathbf{g}}(T)\tilde{\mathbf{g}}(T)'\left(\iint_{A_1}dxdy\right)\right]$$
$$= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}tr\left[\boldsymbol{\Psi}_1(T)^{-1}\tilde{\mathbf{g}}(T)\tilde{\mathbf{g}}(T)'\right] \quad (17)$$

The above result can be manipulated a little further by using the properties of the trace operator. Recall the definition of $\mathbf{D}_T$ and let $\mathbf{1}_q$ and $\mathbf{J}_q$ denote respectively a q vector and a $q\times q$ matrix of ones. Note that $\mathbf{1}_q'\mathbf{1}_q = q$, $\mathbf{1}_q\mathbf{1}_q' = \mathbf{J}_q$ and $\mathbf{J}_q\mathbf{J}_q = q\mathbf{J}_q$. Then we can write

$$\begin{aligned}
tr\left[\boldsymbol{\Psi}_1(T)^{-1}\tilde{\mathbf{g}}(T)\tilde{\mathbf{g}}(T)'\right] &= tr\left[\mathbf{D}_T^{-1}\mathbf{D}_T\boldsymbol{\Psi}_1(T)^{-1}\tilde{\mathbf{g}}(T)\tilde{\mathbf{g}}(T)'\mathbf{D}_T^{-1}\mathbf{D}_T\right]\\
&= tr\left[\frac{1}{q}\mathbf{D}_T\boldsymbol{\Psi}_1(T)^{-1}\mathbf{D}_T\mathbf{D}_T^{-1}\tilde{\mathbf{g}}(T)\mathbf{1}_q'\mathbf{1}_q\tilde{\mathbf{g}}(T)'\mathbf{D}_T^{-1}\right]\\
&= tr\left[\frac{1}{q}\mathbf{D}_T\boldsymbol{\Psi}_1(T)^{-1}\mathbf{D}_T\mathbf{J}_q\mathbf{J}_q\right]\\
&= tr\left[\mathbf{D}_T\boldsymbol{\Psi}_1(T)^{-1}\mathbf{D}_T\mathbf{J}_q\right]\\
&= \mathbf{1}_q'\mathbf{D}_T\boldsymbol{\Psi}_1(T)^{-1}\mathbf{D}_T\mathbf{1}_q\\
&= T^{-n(1+\delta)}C_1^{-2}\mathbf{1}_q'\boldsymbol{\Gamma}_T^{-1}\mathbf{1}_q.
\end{aligned} \quad (18)$$

As far as the second term in the r.h.s. of (16) is concerned note that on the set A_2 we have $\tilde{B}^2(|x-y|)\leq\tilde{B}(T^\delta)\tilde{B}(|x-y|)$ and then

$$\begin{aligned}
&tr\left[\boldsymbol{\Psi}_1(T)^{-1}\left(\iint_{A_2}\mathbf{g}(x)\mathbf{g}(y)'\tilde{B}^2(|x-y|)\,dx\,dy\right)\right]\\
&\leq\tilde{B}(T^\delta)\left[\boldsymbol{\Psi}_1(T)^{-1}\left(\iint_{A_2}\mathbf{g}(x)\mathbf{g}(y)'\tilde{B}^2(|x-y|)\,dx\,dy\right)\right]\\
&\leq\tilde{B}(T^\delta)C_1^{-2}
\end{aligned} \quad (19)$$

Using Lemma 3.1 in the same fashion as in the proof of Theorem 2.1 and results (17)-(19) we obtain that

$$K(\mathbf{k}_T,\mathbf{z})\leq\varepsilon c_1(q)+\frac{1}{\varepsilon^2}c(G)c_2(q)\left[\frac{c(n)2^{n\delta}}{n}\mathbf{1}_q'\boldsymbol{\Gamma}_T^{-1}\mathbf{1}_q+\tilde{B}(T^\delta)\right]$$

In order to minimize the r.h.s. of this inequality, set

$$\varepsilon = \left[\frac{c_2(q)c(G)}{c_1(q)} \mathbf{1}_q' \mathbf{\Gamma}_T^{-1} \mathbf{1}_q + \tilde{B}(T^{\delta}) \right]^{1/3} \left[\frac{c(n)2^{n\delta}}{n} \mathbf{1}_q' \mathbf{\Gamma}_T^{-1} \mathbf{1}_q + \tilde{B}(T^{\delta}) \right]^{1/3}$$

School of Mathematics NIKOLAI N. LEONENKO
Cardiff University

Department of Computer and Management Sciences EMANUELE TAUFER
University of Trento

RIFERIMENTI BIBLIOGRAFICI

D.N. ANDERSON (1992), *A multivariate Linnik distribution,* "Statistics and Probability Letters", 14, 333-336.

J. BERAN (1992), *Statistical methods for data with long-range dependence,* "Statistical Science", 7, 404-427.

J. BERAN (1994), *Statistics for Long-Memory Processes.* Chapman & Hall, New York.

M.J. CHAMBERS (1996), *The estimation of continuous parameter long-memory time series models,* "Econometric Theory", 12, 374-390.

F. COMTE (1996), *Simulation and estimation of long memory continuous time models,* "Journal of Time Series Analysis", 17, 19-36.

R. DAHLHAUS (1995), *Efficient location and regression estimation for long range dependent regression models,* "Annals of Statistics", 23, 1029-1047.

R.S. DEO (1997), *Asymptotic theory for certain regression models with long memory errors,* "Journal of Time Series Analysis", 18, 385-393.

R.S. DEO, AND C.M. HURVICH (1998), *Linear trend with fractional integrated errors*, "Journal of Time Series Analysis", 19, 379-397.

R.L. DOBRUSHIN AND P. MAJOR (1979), *Non-central limit theorems for non linear functionals of Gaussian fields,* "Zeitung für Wahrscheinlichkeit verwandte Gebiete", 50, 27-52.

K.T. FANG, S. KOTZ AND K.W. NG, (1990), *Symmetric Multivariate and Related Distributions*, Chapman & Hall, New York.

L. GIRAITIS, AND H. KOUL (1997), *Estimation of dependence parameter in linear regression with long range dependent errors*, "Stochastic Processes and their Applications", 71, 207-224.

A.V. IVANOV AND N.N. LEONENKO (1989), *Statistical Analysis of Random Fields. Kluwer,* Dordrecht.

S. KARLIN AND W.J. STUDDEN (1966), *Tchebycheff Systems with Applications in Analysis and Statistics*, Interscience, New York.

H.R. KÜNSCH, J. BERAN AND F. HAMPEL (1993), *Contrasts under long range correlations*, "Annals of Statistics", 21, 943-964.

N.N. LEONENKO AND M. ŠILAC-BENŠIĆ (1996A), *Asymptotic properties of the LSE in a regression model with long-memory Gaussian and non-Gaussian stationary errors,* "Random Operators and Stochastic Equations", 4, 17-32.

N.N. LEONENKO AND M. ŠILAC-BENŠIĆ (1996B), *On estimation of regression coefficients in the case of long-memory noise,* "Theory of Stochastic Processes", 18, 3-4.

N.N. LEONENKO AND M. ŠILAC-BENŠIĆ (1998), *On estimation of regression of long-memory random fields observed on the arrays,* "Random Operators and Stochastic Equations", 6, 61-76.

N.N. LEONENKO AND V.V. ANH (1998), *Rate of convergence to the Rosenblatt distribution for additive functionals of stochastic processes with long range dependence,* "Journal of Applied Mathematics and Stochastic Analysis", 14, 27-46.

N.N. LEONENKO AND W.A. WOYCZYNSKI (1998), *Exact parabolic asymptotics for singular n-D Burgers' random fields: Gaussian approximation,* "Stochastic Processes and their Applications", 76, 141-165.

N.N. LEONENKO, M.M. SHARAPOV AND A.H. EL-BASSIOUNY (2000), *On the exactness of Normal approximation of LSE of regression coefficient of long-memory random fields*, "Statistics and Probability Letters", 48, 121-130.

N.N. LEONENKO (1999), *Limit Theorems for Random Fields with Singular Spectrum. Kluwer Academic Publishers,* Dordrecht.

N.N. LEONENKO AND E. TAUFER (2001), *Asymptotic properties of LSE in multivariate continuous regression with Long-Memory stationary errors,* "Metron", 59, 55-72.

I.V. OSTROVSKII (1995), *Analitic and asymptotic properties of multivariate Linnik distribution,* "Math. Phys. Anal. Geom.", 2, 436-455.

P.M. ROBINSON AND F.J. HIDALGO (1997), *Time series regression with long range dependence,* "Annals of Statistics", 25, 77-104.

S.R. SEARLE (1982), *Matrix Algebra Useful for Statistics. Wiley,* New York.

M.S. TAQQU (1979), *Convergence of iterated processes of arbitrary Hermite rank,* "Zeitung für Wahrscheinlichkeit verwandte Gebiete", 50, 53-58.

Y. YAJIMA (1988), *On estimation of a regression model with long-memory stationary errors,* "Annals of Statistics", 16, 791-807.

Y. YAJIMA (1991), *Asymptotic properties of the LSE in a regression model with long-memory stationary errors*, "Annals of Statistics", 19, 158-177.

RIASSUNTO

Convergenza all'approssimazione normale degli stimatori dei minimi quadrati in un modello di regressione multipla con errori fortemente dipendenti

Nel presente lavoro si analizza la convergenza all'approssimazione normale degli stimatori a minimi quadrati in un modello di regressione multipla con errori fortemente dipendenti. Il metodo di studio è basato sull'analisi asintotica di espansioni ortogonali di funzionali non lineari di processi Gaussiani stazionari e sulla distanza di Kolmogorov.

SUMMARY

On the rate of convergence to the normal law of LSE in regression with long range dependence

In this paper we study the rate of convergence to the normal approximation of the least squares estimators in a regression model with long range dependent errors. The method of investigation used is based on the asymptotic analysis of orthogonal expansions of non linear functionals of stationary Gaussian processes and on Kolmogorov's distance.