ON THE RATE OF CONVERGENCE TO THE NORMAL APPROXIMATION OF LSE IN MULTIPLE REGRESSION WITH LONG MEMORY RANDOM FIELDS

Nikolai N. Leonenko
Department of Mathematics
Kyiv University (National)
Volodymyrska, 64
252601, Kyiv, Ukraine
e-mail: Leonenko@aki.pp.kiev.ua

and

School of Mathematical Sciences
Queensland University of Technology
GPO Box 2434, Brisbane Q 4001, Australia
e-mail: m.leonenko@fsc.qut.edu.au

Emanuele Taufer
Department of Computer and Management Sciences
University of Trento
Via Inama 5, 38100 Trento, Italy
e-mail: etaufer@cs.unitn.it

Abstract: In this paper we study the rate of convergence to the normal approximation of the least squares estimators in a regression model with long memory stationary 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 the distance of Kolmogorov.

Keywords: Least Squares Estimators, random fields, long memory errors, Hermite polynomials, Kolmogorov distance, Normal Approximation.

1.Introduction

Long memory random fields arise in applications in the most disparate areas such as astronomy, economics, hydrology and the telecommunications, for a pretty complete set of references to applications the reader is referred to the book of Beran (1994). Our aim is to 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 memory errors. The same problem, for single input regression has been considered in Leonenko, Sharapov and El-Bassiouny (1999), and here we will use the methods adopted therein. Note that we consider regression on continuous homogeneous random fields; this is of importance in view of the fact that procedures of discretization lead sometimes to loss of information (see, for example, Leonenko, 1999, pp. 14-16).

Statistical problems related with long memory 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 methods for long memory continuous models.

Typeset by $\mathcal{AMS}$-TEX

Leonenko and Benšić (1996, 1998) and Leonenko and Taufer (1999) present Gaussian and non-Gaussian limit distributions of univariate and multivariate 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, Beran and Hampel (1993), Dahlhaus (1995), Robinson and Hidalgo (1997), Deo (1997), Deo and Hurvich (1998) which consider regression models with long memory errors in discrete time. One can also consult Koul and Mukherjee (1993,1994), see also their references, which consider the asymptotic properties of various robust estimates of regression coefficients.

The paper is organized as follows. In Section 2 we will state our model and assumptions exactly and formulate the main result which show 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 $\Re^n$, $n > 1$ 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 $T > 0$. Practical situations often claim that Δ is a sphere 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 $\boldsymbol{g}(x) = [g_1(x), \ldots g_q(x)]'$ is a known vector function whose coordinate functions $g_i(x)$, $i = 1, \ldots, q$ 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. $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_q]$ is an unknown vector of parameters and $\eta(x)$ is an homogeneous random field of errors with $\boldsymbol{E}\eta(x) = 0$ and $\boldsymbol{E}\eta(x)^2 < \infty$. The problem is to estimate the vector of parameters $\boldsymbol{\theta}$ using the observations $\zeta(x)$, $x \in \Delta(T)$, $T \to \infty$.

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 $\boldsymbol{E}\xi(x) = 0$, $\boldsymbol{E}\xi(x)^2 = 1$ and correlation function

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

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 ($\lim_{t\to\infty} \frac{L(ts)}{L(t)} = 1$, for every $s > 0$) bounded on each finite interval.

Under assumption 2 we have

$$\int_{\Re^n} |B(x)|\, dx = \infty$$

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 $\boldsymbol{E}G(\xi(x)) = 0$ and $\boldsymbol{E}G^2(\xi(x)) < \infty$, $x \in \Re^n$.

Note that the marginal distributions of a field $\eta(x)$, $x \in \Re^n$, satisfying Assumption 3 need not be Gaussian.

The LSE of $\boldsymbol{\theta}$ can be found by minimizing

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

with respect to $\boldsymbol{\theta}$. The final form of LSE is given by (the integral is taken with respect to every element of the matrices)

$$\hat{\boldsymbol{\theta}}_T = \boldsymbol{Q}_T^{-1} \int_{\Delta(T)} \boldsymbol{g}(x)\zeta(x)\, dx \quad = \boldsymbol{\theta} + \boldsymbol{Q}_T^{-1} \int_{\Delta(T)} \boldsymbol{g}(x) G(\xi(x))\, dx \tag{2.1}$$

where

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

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

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

and that

$$\begin{aligned} Var(\hat{\boldsymbol{\theta}}_T) &= \boldsymbol{E}[\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}][\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}]' \\ &= \boldsymbol{Q}_T^{-1} \int_{\Delta(T)} \int_{\Delta(T)} \boldsymbol{g}(x)\boldsymbol{g}(x)' \boldsymbol{E} G(\xi(x)) G(\xi(y))\, dxdy\, \boldsymbol{Q}_T^{-1} \end{aligned} \tag{2.2}$$

Let $H_m(u) = (-1)^m \exp\{u^2/2\} \frac{\partial^m}{\partial u^m} \exp\{u^2/2\}$, $u \in \Re^1$, $m = 0, 1, \dots$, be the Chebyshev-Hermite polynomials with the leading coefficient equal to 1. As it is well known, they form a complete orthogonal system in the Hilbert space $L_2(\Re^1, \phi(u)du)$, where $\phi(u) = (2\pi)^{-1/2} \exp\{-u^2\}$, $u \in \Re^1$. Note that $H_0(u) = 1$, $H_1(u) = u$, $H_2(u) = u^2 - 1 \dots$ It is well known that (see, for example, Ivanov and Leonenko 1989, p.55) that if (ξ, η) is a Gaussian vector with $\boldsymbol{E}\xi = \boldsymbol{E}\eta = 0$, $\boldsymbol{E}\xi^2 = \boldsymbol{E}\eta^2 = 1$, $\boldsymbol{E}\xi\eta = \rho$, then for all $m, p > 0$

$$\boldsymbol{E} H_m(\xi) H_p(\eta) = \delta_m^p \rho^m m! \tag{2.3}$$

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 $L_2(\Re^1, \phi(u)du)$:

$$G(u) = \sum_{m \geq 0} \frac{C_m}{m!} H_m(u), \quad C_m = \int_{\Re} G(u) H_m(u) \phi(u) du \tag{2.4}$$

and by Parseval's relation:

$$\boldsymbol{E} G^2(\xi(0)) = \sum_{m \geq 0} \frac{C_m^2}{m!} = \int_{\Re} G^2(u) \phi(u) du < \infty \tag{2.5}$$

Note that $C_0 = \boldsymbol{E} G(\xi(0)) = 0$. From (2.1)-(2.4) we obtain:

$$Var(\hat{\boldsymbol{\theta}}_T) = \boldsymbol{Q}_T^{-1} \sum_{m \geq 1} \boldsymbol{\Psi}_m(T)\, \boldsymbol{Q}_T^{-1} \tag{2.6}$$

where

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

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

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

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

where

$$\boldsymbol{D}_T = \mathrm{diag}[g_1(\mathbf{1}_q T), \ldots, g_q(\mathbf{1}_q T)]$$

and that $\boldsymbol{L}_m(\alpha, n)$ is positive definite matrix ($\mathbf{1}_q$ is a q-vector of ones).

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

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

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

Assumption 5. Let $m = 1$ or $m = 2$. Suppose that the matrix

$$\boldsymbol{F}_{mT}(x,y) = \boldsymbol{D}_T^{-1} \boldsymbol{g}(xT)\boldsymbol{g}(yT)'\, a^m \left(\frac{x-y}{|x-y|}\right) \frac{1}{|x-y|^{m\alpha}} \left[\frac{L^m(|x-y|T)}{L^m(T)} - 1\right] \boldsymbol{D}_T^{-1} \leq \boldsymbol{F}_m(x,y),$$

$0 < \alpha < n/m$, where the sign $'\leq'$ means this relationship between all single elements of the matrix $\boldsymbol{F}_{mT}(x,y)$, and that

$$\int_{\Delta(T)} \int_{\Delta(T)} \boldsymbol{F}_m(x,y)\, dx\, dy < \infty.$$

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 Benŝic, 1998):

$$\lim_{\to\infty} |\boldsymbol{\Psi}_m(T) - \frac{C_m^2}{m!} T^{2n-m\alpha} L^m(T)\, \boldsymbol{D}_T \boldsymbol{L}_m(\alpha, n)\, \boldsymbol{D}_T| = \boldsymbol{o}(1), \qquad 0 < \alpha < n/m$$

where $\boldsymbol{o}(1)$ is the matrix function such that $\lim_{\to\infty} |\boldsymbol{o}(1)(T)| = \mathbf{0}$.

Let

$$\Pi[\boldsymbol{c}, \boldsymbol{d}] = \{u \in \Re^q : c_i \leq u_i \leq d_i, i = 1, \ldots, q\}$$

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

$$\mathcal{K}(\boldsymbol{X}, \boldsymbol{Y}) = \sup_{\boldsymbol{z}} |P(\boldsymbol{X} \in \Pi[\boldsymbol{\infty}, \boldsymbol{z}]) - P(\boldsymbol{Y} \in \Pi[\boldsymbol{\infty}, \boldsymbol{z}])|$$

Let $\boldsymbol{N}$ be a standard normal random q-vector with zero mean and unit covariance matrix and consider the random vector

$$\boldsymbol{\kappa}_T = \boldsymbol{\Psi}_1^{-\frac{1}{2}}(T) \boldsymbol{Q}_T [\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}],$$

where the LSE $\hat{\boldsymbol{\theta}}_T$ are defined in (2.1) and $\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)$ is a nonsingular matrix such that

$$\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T) \boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)' = \boldsymbol{\Psi}_1(T)^{-1}.$$

The main result of this paper describes the rate of convergence to the normal law of the random vector $\boldsymbol{\kappa}$ as $T \to \infty$. The result is presented in the following theorem

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

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

then the following quantity exists:

$$\lim_{T\to\infty} \sup[T^\alpha / L(T)]^{1/3} \mathcal{K}(\boldsymbol{\kappa}, \boldsymbol{N})$$

and is bounded by

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

where $\boldsymbol{L}_1(\alpha, n)$ and $\boldsymbol{L}_2(\alpha, n)$ are defined in Assumption 4 and

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

$$\begin{aligned} c_1(q) &= \sqrt{2/\pi}, \quad \text{if} \qquad q = 1 \\ &= (q-1)\frac{\Gamma\left[(q-1)/2\right]}{\sqrt{2}\,\Gamma(q/2)}, \quad q \geq 2 \end{aligned} .$$

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

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 $\boldsymbol{X}$,$\boldsymbol{Y}$ be two arbitrary random q-vectors and $\boldsymbol{N}$ be a standard Gaussian q-vector such that, for all $\boldsymbol{a}, \boldsymbol{b} \in \Re^q$,*

$$|P(\boldsymbol{X} \in \Pi[\boldsymbol{a}, \boldsymbol{b}]) - P(\boldsymbol{N} \in \Pi[\boldsymbol{a}, \boldsymbol{b}])| \leq K,$$

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

$$\mathcal{K}(\boldsymbol{X} + \boldsymbol{Y}, \boldsymbol{N}) \leq K + P(\boldsymbol{Y} \notin \Pi[-\boldsymbol{1}_q\varepsilon, \varepsilon\boldsymbol{1}_q]) + \varepsilon c_1(q) \tag{3.1}$$

where $c_1(q)$ is defined in Theorem 1 and $\boldsymbol{1}_q$ 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 Leonenko and Woyczynski, 1998).

Lemma 3.2. *Let $\boldsymbol{Y}$ be a random q-vector with mean $\boldsymbol{EY} = \boldsymbol{0}$ and covariance matrix $\boldsymbol{EYY'} = \boldsymbol{\Sigma} = (\sigma_{ij})_{1\leq i,j\leq q}$, and let $Z_i = Y_i/(\kappa_i\sigma_i)$ where $\sigma_i^2 = \sigma_{ii}$, and $\kappa_1, \ldots, \kappa_q > 0$ are some constants. Then*

$$P\left(\max_{1\leq i\leq q} |Z_i| \geq 1\right) \leq \frac{1}{q^2}\left(\sqrt{s} + \sqrt{(qt-s)(q-1)}\right)^2, \tag{3.2}$$

where $t = tr\,\boldsymbol{\Pi}$, $s = \boldsymbol{1}_q'\boldsymbol{\Pi}\boldsymbol{1}_q$, $\boldsymbol{\Pi} = \boldsymbol{EZZ'} = (\pi_{ij})_{1\leq i,j\leq q}$, $\pi_{ij} = \sigma_{ij}/(\sigma_i\sigma_j\kappa_i\kappa_j)$, and $\boldsymbol{1}_q$ is a q-vector of ones.

Remark. Note that

$$0 \leq s = \sum_{i=1}^{q} \pi_{ii} + 2\sum_{i=1}^{q}\sum_{i<j}^{q} \pi_{ij} \leq 2\sum_{i=1}^{q} \pi_{ii} = 2t \tag{3.3}$$

and hence a less tight version of Lemma 3.2 which will allow us to obtain a more compact result can be stated as

$$P\left(\max_{1\le i\le q}|Z_i|\ge 1\right)\le t\,\frac{\left(\sqrt{2}+\sqrt{q(q-1)}\right)^2}{q^2},$$

for later convenience let

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

Proof of Theorem 2.1. The expansion 2.4 implies the following expansion in the Hilbert space $L_2(\Omega)$:

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

We now consider the random vectors

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

In order to apply Lemma 3.1, we represent $\boldsymbol{\kappa}_T$ as

$$\boldsymbol{\kappa}_T=\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)[\boldsymbol{X}_T+\boldsymbol{Y}_T]$$

where

$$\boldsymbol{X}_T=C_1\boldsymbol{\eta}_1(T),\qquad \boldsymbol{Y}_T=\sum_{m\ge 2}\frac{C_m}{m!}\boldsymbol{\eta}_m(T)$$

Note that $\boldsymbol{X}_T$ is a Gaussian random vector with $\boldsymbol{E}\boldsymbol{X}_T=\boldsymbol{0}$ and $\boldsymbol{E}\boldsymbol{X}_T\boldsymbol{X}_T'=\boldsymbol{\Psi}_1(T)$. So we have

$$\mathcal{K}(\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{X}_T,\boldsymbol{N})=0$$

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

$$P(\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{Y}_t\notin\Pi[-\mathbf{1}_q\varepsilon,\varepsilon\mathbf{1}_q])=P(\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{Y}_t\,\frac{1}{\varepsilon}\notin\Pi[-\mathbf{1}_q,\mathbf{1}_q])\le P\left(\max_{1\le i\le q}|Z_i|\ge 1\right)$$

where $\boldsymbol{Z}=\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{Y}_t\,\frac{1}{\varepsilon}$. By using the properties of the trace operator and in view of Lemma 3.2 and formula (3.3) we have that

$$\begin{aligned} t=\mathrm{tr}\boldsymbol{E}\boldsymbol{Z}\boldsymbol{Z}' &=\frac{1}{\varepsilon^2}\mathrm{tr}\left[\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{E}\boldsymbol{Y}_T\boldsymbol{Y}_T'\boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\right]\\ &=\frac{1}{\varepsilon^2}\mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{E}\boldsymbol{Y}_T\boldsymbol{Y}_T'\right]\\ &=\frac{1}{\varepsilon^2}\mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{\Sigma}\right]\end{aligned}$$

In order to evaluate an upper bound for t, note that from Assumption 2 we know that for $r\le m$ and the following relation hold for any element of the matrices:

$$\boldsymbol{\Psi}_r(T)\,\frac{r!}{C_r^2}\le\boldsymbol{\Psi}_m(T)\,\frac{m!}{C_m^2},\quad 1\le m\le r,$$

so that, for $0 < \alpha < n/2$

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

Next, note that $\boldsymbol{\Psi}_1(T)$ is a symmetric positive definite matrix then, there exists an orthogonal matrix $\boldsymbol{P}$ such that $\boldsymbol{\Psi}_1(T) = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'$ where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\Psi}_1(T)$ which, by positive definiteness are all real and positive. Hence we have that $\operatorname{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{\Sigma}\right] = \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}'\boldsymbol{\Sigma}\right] = \operatorname{tr}\left[\boldsymbol{\Lambda}^{-1}\boldsymbol{\Sigma}\right] = \sum_i \lambda_i^{-1}\sigma_{ii}$. Hence if we take constants $k_i \geq \sigma_{ii}$, $i = 1, \ldots, q$ it follows that $\sum_i \lambda_i^{-1}\sigma_{ii} \leq \sum_i \lambda_i^{-1} k_i$. Following this line of reasoning we can obtain the following estimate

$$
t = \operatorname{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{\Sigma}\right] \leq \left[\sum_{m \geq 2} \frac{C_m^2}{m!} \right] \operatorname{tr}\left[\boldsymbol{\Psi}_1(T)^{-1} \frac{2!}{C_2^2} \boldsymbol{\Psi}_2(T) \right]
$$

Taking the limits, we know that, by assumptions 4 and 5 and Lebesgue dominated convergence theorem $\frac{2!}{C_2^2}\boldsymbol{\Psi}_2(T)$, as $T \to \infty$ converges to

$$
T^{2n-2\alpha} L^2(T) \boldsymbol{D}_T \boldsymbol{L}_2(\alpha, n) \boldsymbol{D}_T, \quad 0 < \alpha < n/2.
$$

Similarly, we have that $\boldsymbol{\Psi}_1(t)^{-1}$, as $T \to \infty$, converges to

$$
C_1^{-2} T^{\alpha - 2n} L^{-1}(T) \boldsymbol{D}_T^{-1} \boldsymbol{L}_1^{-1}(\alpha, n) \boldsymbol{D}_T^{-1}
$$

for $0 < \alpha < n$. Hence we obtain the following upper bound for t:

$$
\frac{1}{\varepsilon^2} c(G) \frac{L(T)}{T^\alpha} \operatorname{tr}\left[\boldsymbol{L}_1^{-1}(\alpha, n)\, \boldsymbol{L}_2(\alpha, n) \right], \quad 0 < \alpha < n/2 \tag{3.4}
$$

where

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

Finally, using Lemma 3.1 with $\boldsymbol{X} = \boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{X}_T$ and $\boldsymbol{Y} = \boldsymbol{\Psi}_1^{-\frac{1}{2}}(T)\boldsymbol{Y}_T$ we obtain from (3.4) that for any $\varepsilon > 0$:

$$
\mathcal{K}(\boldsymbol{\kappa}_T, \boldsymbol{N}) \leq \varepsilon c_1(q) + \frac{1}{\varepsilon} c(G) c_2(q) \operatorname{tr}\left[\boldsymbol{L}_1^{-1}(\alpha, n)\, \boldsymbol{L}_2(\alpha, n) \right]
$$

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

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

and substituting into (3.5) we obtain the following

$$
\mathcal{K}(\boldsymbol{\kappa}_T, \boldsymbol{N}) \leq 2 c_1(q)^{2/3} [c(G) c_2(q) \operatorname{tr}\left(\boldsymbol{L}_1^{-1}(\alpha, n)\, \boldsymbol{L}_2(\alpha, n) \right)]^{1/3} \left[\frac{L(T)}{T^\alpha} \right]^{1/3}
$$

4. Extensions and generalizations

As follows from the results of Leonenko and Bensic (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 Kolomogorov'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 1) and the case of radial regression function: $\boldsymbol{g}(x) = \tilde{\boldsymbol{g}}(|x|)$, $x \in \Re^n$. We consider now the case $\Delta(T) = v(T) = \{x \in \Re^n : |x| < T\}$, $T \to \infty$. Thus the random field $\zeta(x) = \boldsymbol{\theta}'\boldsymbol{g}(x) + \eta(x)$ 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 $\boldsymbol{E}\xi(x) = 0$, $\boldsymbol{E}\xi^2(x) = 1$ and covariance function $B(x) = \tilde{B}(|x|) = \operatorname{cov}(\xi(0), \xi(x)) \to 0$ as $|x| \to \infty$, and $\eta(x) = G(\xi(x))$ where $\boldsymbol{E}G(\xi(x)) = 0$, $\boldsymbol{E}G^2(\xi(x)) < \infty$, $x \in \Re^n$.

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

Assumption 8. There exist a $\delta \in (0,1)$ such that any element of the matrix

$$\begin{aligned}\boldsymbol{\Gamma}_T &= T^{-n(1+\delta)}\boldsymbol{D}_T^{-1}\int_{v(T)}\int_{v(T)} \boldsymbol{g}(x)\boldsymbol{g}(y)'\tilde{B}(|x-y|)\,dxdy\,\boldsymbol{D}_T^{-1} \\ &= T^{n(1-\delta)}\tilde{B}(T)\boldsymbol{D}_T^{-1}\int_{v(T)}\int_{v(T)} \boldsymbol{g}(xT)\boldsymbol{g}(yT)'\frac{\tilde{B}(T|x-y|)}{\tilde{B}(T)}\,dxdy\,\boldsymbol{D}_T^{-1} \to \infty\end{aligned}$$

as $T \to \infty$.

Note that if Assumption 2 (with $a(\cdot) \equiv 1$), and Assumptions 4 and 5 (with $m = 1$) hold, then $\boldsymbol{\Gamma}_T \to \infty$ as $T \to \infty$. Thus the random field $\xi(x)$, $x \in \Re^n$ satisfying Assumption 8 is a long memory random field.

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 exists:

$$\lim_{T\to\infty}\sup\left[\frac{q_1(n)2^{n\delta}}{n}\boldsymbol{1}_q'\boldsymbol{\Gamma}_T^{-1}\boldsymbol{1}_q + \tilde{B}(T^\delta)\right]^{-1/3}\mathcal{K}(\boldsymbol{\kappa}, \boldsymbol{N})$$

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

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

Before proving the theorem, we need some preliminaries. For this purpose, let $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ 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 $|\boldsymbol{U}_1 - \boldsymbol{U}_2|$ between $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ is

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

where

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

is the incomplete beta function.

Using Randomization we obtain for every function $f(|x-y|)$, $x\in\Re^n$, $y\in\Re^n$:

$$\begin{aligned}\int_{v(T)}\int_{v(T)} f(|x-y|)dxdy &= |v(T)|^2 \boldsymbol{E} f(|\boldsymbol{U}_1-\boldsymbol{U}_2|)\\ &= T^{2n}|v(1)|^2\int_0^{2T} f(u)\rho_T(u)du\\ &= c(n)T^n\int_0^{2T} z^{n-1}f(z)I_{1-(z/2T)^2}(\frac{n+1}{2},\frac{1}{2})dz, \end{aligned} \tag{4.1}$$

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*:* wee 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 $t = \mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{\Sigma}\right]$ consider

$$\mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\frac{2!}{C_2^2}\boldsymbol{\Psi}_2(T)\right] = \mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\left(\int\int_{A_1}+\int\int_{A_2}\right)\boldsymbol{g}(x)\boldsymbol{g}(y)'\tilde{B}^2(|x-y|)\,dx\,dy\right] \tag{4.2}$$

On the set A_1 we have $B^2(\cdot)\leq 1$ and then, for the first term on the r.h.s. of (4.2), using (4.1) with $f(|x-y|)=\mathbf{1}_{A_1}$ we have the estimate

$$\begin{aligned}&\mathrm{tr}\left[\boldsymbol{\Psi}_1(T)^{-1}\left(\int\int_{A_1}\boldsymbol{g}(x)\boldsymbol{g}(y)'\tilde{B}^2(|x-y|)\,dx\,dy\right)\right]\\ &\leq \mathrm{tr}\left[\boldsymbol{D}_T^{-1}\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{g}(T)\boldsymbol{g}(T)'\int\int_{A_1} dx\,dy\,\boldsymbol{D}_T^{-1}\boldsymbol{D}_T\right]\\ &= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}\mathrm{tr}\left[\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{D}_T\boldsymbol{D}_T^{-1}\boldsymbol{g}(T)\boldsymbol{g}(T)'\boldsymbol{D}_T^{-1}\right]\\ &= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}\mathrm{tr}\left[\frac{1}{q}\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{D}_T\boldsymbol{D}_T^{-1}\boldsymbol{g}(T)\mathbf{1}_q'\mathbf{1}_q\boldsymbol{g}(T)'\boldsymbol{D}_T^{-1}\right]\\ &= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}\mathrm{tr}\left[\frac{1}{q}\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{D}_T\boldsymbol{J}_q\boldsymbol{J}_q\right]\\ &= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}\mathrm{tr}\left[\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{D}_T\boldsymbol{J}_q\right]\\ &= c(n)T^{n(1+\delta)}\frac{2^{n\delta}}{n}\mathbf{1}_q'\boldsymbol{D}_T\boldsymbol{\Psi}_1(T)^{-1}\boldsymbol{D}_T\mathbf{1}_q\\ &= c(n)\frac{2^{n\delta}}{n}C_1^{-2}\mathbf{1}_q'\boldsymbol{\Gamma}_T^{-1}\mathbf{1}_q\end{aligned}$$

Where $\boldsymbol{J}_q$ is a $q\times q$ matrix of ones. As far as the second term in the r.h.s. of (4.2) 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

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

Using Lemma 3.1 in the same fashion as in the proof of Theorem 2.1 we obtain that

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

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

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

References

Beran, J. (1992). Statistical methods for data with long-range dependence. *Statistical Science* **7**, 4, 404-427.

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

Dahlhaus, R. (1995). Efficient location and regression estimation for long range dependent regression models. *Ann. Stat.* **23**, 1029-1047.

Deo, R.S. (1997) Asymptotic theory for certain regression models with long memory errors. *J.Time Series Anal.* **18**, 385-393.

Deo, R.S. and Hurvic, C.M. (1998) Linear trend with fractional integrated errors. *J.Time Series Anal.* **19**, 379-397.

Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for non linear functionals of Gaussian fields. *Z. Wahrsch. verw. Gebiete* **50**, 27-52.

Giraitis, L. and Koul, H. (1997). Estimation of dependence parameter in linear regression with long range dependent errors. *Stoch. Proc. Appli.* **71**, 207-224.

Ivanov, A.V. and Leonenko, N.N. (1989) *Statistical Analysis of Random Fields* . Kluwer, Dordrecht.

Koul, H.L. and Mukherjee, K. (1993). Asymptotics of R-, MD- and LAD-estimators in linear regression models with long range dependence. *Probab. Theory and Related Fields* **95**, 535-553.

Koul, H.L. and Mukherjee, K. (1994). Regression quantiles and relate processes under long range dependent errors. *J.Multiv. Anal.* **51**, 318-337.

Künsch, H.R., Beran, J., Hampel, F. (1993). Contrasts under long range correlations. *Ann. Stat.* **21**, 943-964.

Leonenko, N.N. and Šilac-Benšić, M. (1996a). Asymptotic properties of the LSE in a regression model with long-memory Gaussian and non-Gaussian stationary errors. *Random Oper. and Stoch. Equ.* **4**, 1, 17-32.

Leonenko, N.N. and Šilac-Benšić, M. (1996b). On estimation of regression coefficients in the case of long-memory noise. *Theory of Stoch. Processes* **18**, 3-4.

Leonenko, N.N. and Šilac-Benšić, M. (1998). On estimation of regression of long-memory random fields observed on the arrays. *Random Oper. and Stoch. Equ.* **6**, 1, 61-76.

Leonenko, N.N. and Anh, V.V. (1998). On the rate of convergence to Rosenblatt distribution for additive functionals of stochastic processes with long range dependence. Submitted for publication.

Leonenko, N.N. and Woyczynski, W.A. (1998). Exact Parabolic Asymptotics for Singular n-D Burgers' Random Fields: Gaussian Approximation. *Stoch. Proc. and Applic.* **76**, 141-165.

Leonenko, N.N., Sharapov, M.M., El-Bassiouny, A.H. (1998). On the exactness of Normal Approximation of LSE of Regression Coefficient of Long Memory Random Fields. Submitted for publication.

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

Leonenko, N.N. and Taufer, E. (1999). Asymptotic Properties of LSE in Multivariate Continuous Regression with Long Memory Stationary Errors. Submitted for publication.

Robinson, P.M. and Hidalgo, F.J. (1997). Time series regression with long range dependence. *Ann. Stat.* **25**, 77-104.

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

Taqqu, M.S. (1979). Convergence of iterated processes of arbitrary Hermite rank. *Z. Wahrsch. verw. Gebiete***50**, 53-58.

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Ann. Stat.* **6**, 2, 791-807.

Yajima, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Ann. Stat.* **19**, 1, 158-177.