Asymptotic distributions of least square estimations in a regression model with singular errors*

Mirta Benšić†

Abstract. *We study some problems of the parameter inference which are in connection with wide sense stationary long memory processes. Here we present the asymptotic behaviour of the corelation matrix and the limit distributions of the LSE for the regression coefficients in some types of linear models with singular Gaussian and non-Gaussian errors.*

Key words: *least squares, asymptotic distributions, regression*

Sažetak Asimptotske distribucije procjenitelja dobivenog metodom najmanjih kvadrata u regresionom modelu sa singularnim greškama. *U ovom radu proučeni su neki problemi procjene parametara vezani uz stacionarne procese u širem smislu koji dugo pamte. Opisano je asimptotsko ponašanje korelacione matrice i granične distribucije procjenitelja dobivenog metodom najmanjih kvadrata za regresione koeficijente u nekim tipovima linearnih modela s Gaussovim i ne-Gaussovim singularnim greškama.*

Ključne riječi: *problemi najmanjih kvadrata, asimptotske distribucije, regresija*

1. Introduction

Estimation problems for linear regression parameters are of great interest in statistical applications and they are widely treated in the mathematical and statistical literature (see for example [1], [4], [2]). Among them, the models with discrete time parameter and independent errors and so called "short memory"

*The lecture presented at the Mathematical Colloquium in Osijek organized by Croatian Mathematical Society - Division Osijek, November 10, 1995. The full material presented on this lecture is a part of the paper [7].

†Faculty of Economics, Department of Mathematics, Gajev trg 7, HR-31 000 Osijek, Croatia, e-mail: bensic@oliver.efos.hr

models were studied in a great deal. These results guarantee that the LSE often has very desirable statistical properties.

Nowadays the models with "long memory" errors have been researched. The papers [3], [6], [9], [11], [12] indicate that in these models the LSE may be very good, too.

In this paper we present the asymptotic theory of the LSE of regression coefficients of the continuous time stationary processes with unbounded spectral density (singular noise).

2. The model

Let us consider the continuous time regression model of the form

$$\nu(t) = a^\tau g(t) + \eta(t), \;\; t \in R,$$

where $g(t) = [g_1(t), \dots, g_n(t)]^\tau$ is the known vector function, $g : R \to R^n$, whose coordinate functions $g_i(t)$ form a linearly independent set of real functions, being positive on $[0, T]$ and square integrable over the same interval for all $T > 0$ and $i = 1, \dots, n$. $a = [a_1, \dots, a_n]^\tau$ is the unknown vector of parameters. $\eta(t), \; t \in R$, is the stationary stochastic process of errors which fulfills the following conditions:

1. $E\eta(t) = 0$, $E\eta^2(t) < \infty$.
2. $\eta(t)$, $t \in R$, allows a representation:

   $$\eta(t) = G(\xi(t)),$$

   where:

   (a) $\xi(t)$ is a measurable, m.s. continuous, stationary Gaussian process with $E(\xi(t)) = 0$, $E(\xi^2(t)) = 1$;

   (b) The covariance function of $\xi(t)$ has the following form:

   $$B(\mid t \mid) = cov(\xi(0), \xi(t)) = \frac{L(\mid t \mid)}{\mid t \mid^\alpha}, \;\; 0 < \alpha < 1,$$

   where $L(t)$ is a real slowly varying function for large values of t (see e.g. [7]), bounded on each finite interval;

   (c) The covariance function $B(t)$ has a spectral density $f(\mid \lambda \mid)$ which is a decreasing function for $\mid \lambda \mid \geq \lambda_0$, $\lambda_0 \in R_+$;

   (d) $G(t)$ is a real, possible nonlinear Borel function such that $EG^2(\xi(0)) < \infty$.

It can be observed that in the case of assumption 2.(d) the function $G(t)$ allows the expansion

$$G(t) = \sum_{i=0}^{\infty} \frac{C_i}{i!} H_i(t), \quad C_i = \int_{-\infty}^{\infty} G(t)\, H_i(t)\, \phi(t)\, dt, \;\; i \in \mathbf{N}_0,$$

where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

and $\{H_i(t) : i \in \mathbf{N}_0\}$ are Hermite polynomials with a unit leading coefficient forming a complete orthogonal system in the Hilbert space $L_2(R, \phi(t)dt)$.

3. The integer $m \geq 1$ exists such that $C_1 = C_2 = \ldots = C_{m-1} = 0$, $C_m \neq 0$. We say that $m = rang\, G$ (see e.q. [1,2]).
4. For the vector function $g(t)$, given α and m, the following limits exist and are finite:

$$l = \lim_{T\to\infty} d(T)^{-1} \int_0^1 \int_0^1 g(Tu)\, g(Tv)^{\tau} \frac{du\, dv}{|\, u-v\,|^{m\alpha}}\, d(T)^{-1},$$

$$L = \lim_{T\to\infty} d(T)^{-1}\, V_1(T)\, d(T)^{-1},$$

$$V_1(T) = \int_0^1 g(uT)\, g(Tu)^{\tau}\, du = \frac{1}{T} V(T),$$

$$d(T) = \mathrm{diag}\,(g_1(T), \ldots, g_n(T)).$$

Also let l be the regular matrix.

We can see that our model is the model with stationary errors and the long-term dependence.

The problem we have been dealing with is estimating by the vector parameter a using observations of the random process $\nu(t)$ during the time interval $[0, T]$. For this we have chosen the LS estimator. Here we have given some of its asymptotic properties.

3. Asymptotic properties of the LSE of regression coefficients

If the assumptions 1 and 2 are true for the process

$$\nu(t) = a^{\tau} g(t) + \eta(t), \;\; t \in R,$$

then the LS estimator for the vector parameter a on the continuous time interval $[0,T]$ has the following form:

$$\hat{a}(T) = V(T)^{-1} \int_0^T g(t)\,\nu(t)\,dt, \quad V(T) = \int_0^T g(t)\,g(t)^\tau\,dt.$$

It is to be noted that the matrix $V(T)$ under our conditions is a real, symmetric, positive definite matrix for all $T > 0$.

It is obvious that the estimator $\hat{a}(T)$ is unbiased. Also for the correlation matrix $\mathcal{D}(\hat{a}(T))$ the following theorem holds:

Theorem 1. *If the assumptions 1, 2, 3 and 4 hold for the process*

$$\nu(t) = a^\tau\, g(t) \,+\, \eta(t), \quad t \in R,$$

with $0 < \alpha < \frac{1}{m}$*, then for the correlation matrix* $\mathcal{D}(\hat{a}(T))$ *of the LS estimator of an unknown vector parameter* a *the following holds:*

$$\lim_{T\to\infty} ||d(T)^{-1}\,V(T)\,\mathcal{D}(\hat{a}(T))\,V(T)\,d(T)^{-1}\,L^{-m}(T)\,T^{m\alpha-2}\,\frac{m!}{C_m^2} - l|| \,=\, 0. \quad (1)$$

This theorem makes an analysis of asymptotic behaviour of the correlation matrix $\mathcal{D}(\hat{a}(T))$ easier. Therefore, by using it we can prove that LSE in the case of polynomial regression under our conditions is one consistent estimator in the sense that the correlation matrix tends to zero matrix (mean-square consistent, see eg. [5]). The following corollary and example confirm this.

Corollary 1. *If the assumptions 1, 2, 3 and 4 are valid for the process*

$$\nu(t) = a^\tau\, g(t) \,+\, \eta(t), \quad t \in R,$$

with $0 < \alpha < \frac{1}{m}$ *and let the matrix* L *be regular. Then for the correlation matrix* $\mathcal{D}(\hat{a}(T))$ *of the LS estimator of an unknown vector parameter* a *the following holds:*

$$\lim_{T\to\infty} ||d(T)\,\mathcal{D}(\hat{a}(T))\,d(T)\,L^{-m}(T)\,T^{m\alpha}\,\frac{m!}{C_m^2} - L^{-1}\,l\,L^{-1}|| \,=\, 0. \quad (2)$$

Example 1. *If we consider the model with* $g(t) = [t^{\nu_1}, \ldots, t^{\nu_n}]^\tau$, $0 < \nu_1 < \ldots < \nu_n$*, using* Corollary *1., we can state that* $\hat{a}(T)$ *is one mean-square consistent estimator for the vector parameter* a*.*

We can also give one representation for the asymptotic distribution of the process

$$X(T) = R(T)^{-1}(\hat{a}(T) - a),$$

$$R(T) = \frac{C_m}{\sqrt{m!}} L^{m/2}(T) T^{1-m/2} V(T)^{-1} d(T) l_1,$$

where l_1 means the regular matrix such that $l = l_1 l_1^\tau$. The main result, described by *Theorem* 3., is proved using the reduction theorem (*Theorem* 2.) which makes it possible to describe an asymptotic distribution for the process

$$X(T) = R(T)^{-1}(\hat{a}(T) - a),$$

in our model with that of the process

$$X_m(T) = \frac{C_m}{m!} R(T)^{-1} V(T)^{-1} \int_0^T g(t)\, H_m(\xi(t))\, dt, \tag{3}$$

where $m = rank\, G$. In fact the following theorems hold:

Theorem 2. *If we suppose that the assumptions 1, 2, 3 and 4 are fulfilled with* $0 < \alpha < \frac{1}{m}$*, then it holds:*

$$\lim_{T\to\infty} E||X(T) - X_m(T)||^2 = 0.$$

This means that if one of the asymptotic distributions for $X(T)$ *or for* $X_m(T)$ *exists, then the other one exists as well, and they are the same.*

Theorem 3. *Let us suppose that the assumptions 1, 2, 3 and 4 hold, that there exists the vector function* $g'(u)$ *for which*

$$||d(T)^{-1}\, g(uT) - g'(u)|| \to 0, \;\; as \;\; T \to \infty, \tag{4}$$

uniformly for $u \in [0,1]$ *and*

$$\int_{\mathbb{R}^m} || \int_0^1 g'(u)\, e^{iu(\lambda_1+\cdots+\lambda_m)} du||^2 \frac{d\lambda_1 \cdots d\lambda_m}{|\lambda_1 \cdots \lambda_m|^{1-\alpha}} < \infty, \tag{5}$$

where $m = rang\, G$*. Then the process* $X(T)$ *converges in distribution to the random vector*

$$\kappa = C \int_{\mathbb{R}^m} \left[\int_0^1 g'(u)\, e^{iu(\lambda_1+\cdots+\lambda_m)}\, du \right] \frac{W(d\lambda_1)\cdots W(d\lambda_m)}{|\lambda_1 \cdots \lambda_m|^{\frac{1-\alpha}{2}}},$$

$$C = \frac{\sqrt{\alpha^m}}{\sqrt{m!\, 2^m\, \Gamma^m(1+\alpha)\, \cos^m \frac{\alpha\pi}{2}}}\, l_1^{-1}.$$

$\int_{\mathbb{R}^m} W(d\lambda_1)\cdots W(d\lambda_m)$ *means a multiple Wiener-Itô integral in the sense of the book Major [8].*

Remark 1. *If the model satisfies all conditions of this theorem then the correlation matrix for* κ *can be calculated and it is*

$$\mathcal{D}(\kappa) = \frac{1}{(m!)^2}\, I.$$

Also, from the integral representation for κ it can be stated that for $m = 1$ κ has a multidimensional normal distribution with the correlation matrix which guarantees that the marginal distributions are independent and identical.

For $m = 2$ the integral representation guarantees that each component of κ has a Rosenblat distribution.

It is to be noted that the polynomial regression with the errors upon our assumptions satisfies all requirements from this section.

References

[1] T. W. Anderson, *The Statistical Analysis of Time Series, John Wiley*, New York, 1972.

[2] P. J. Brockwell, R. A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1991.

[3] R. Dahlhaus, *Efficient location and regression estimation for long range dependence regression models*, Preprint N660, University of Heidelberg, 1992. dependent stationary Gaussian time series, Ann. Statist. **14**(1986), 517–532.

[4] E. J. Hannan, *Multiple Time Series*, John Wiley, New York, 1970.

[5] A. V. Ivanov, N. N. Leonenko, *Statistical Analysis of Random Fields*, Kluwer, London, 1989.

[6] H. Künsch, J. Beran, F. Hampel, *Contrasts under long-range correlations*, Ann. Statist. **21**(1993), 943–964.

[7] N. N. Leonenko, M. Šilac-Benšić, *On estimation of regression coefficients in the case of long-memory noise*, Theory of random processes (in preparation)

[8] P. Major, *Multiple Wiener–Itô Integrals*, Lecture Notes in Mathematics, 849, Springer-Verlag, Berlin, 1981.

[9] P. M. Robbinson, *Semiparametric analysis of long-memory time series*, Ann. Statist. **22**(1994), 515–539.

[10] M. Rosenblatt, *Stationary Sequence and Random Fields*, Birkhäuser, Boston, 1985.

[11] Y. Yajima, *On estimation of regression model with long-memory stationary errors*, Ann. Statist. **16**(1988), 791–807.

[12] Y. Yaima, *Asymptotic properties of the LSE in a regression model with long-memory stationary errors*, Ann. Statist. **19**(1991), 158–177.