Zbornik radova Filozofskog fakulteta u Nišu
Serija Matematika 6:2 (1992), 303–306

FILOMAT '92, Niš, October 8–10, 1992

NONLINEAR APPROXIMABILITY OF SOME RANDOM FIELDS

Zoran Ivković

Abstract. *We introduce the notion of the (global) nonlinear approximability of Hermite polynomial of Gausian martingale* $\{\eta(t), t \in (0,\infty)^d\}$ *in* [1]. *In this paper we generalize this notion to random fields which are nonlinear transformation of* $\{\eta(t)\}$. *Also, some proofs from* [1] *are improved.*

1. Consider a real second-order continuous random process $\{\xi(t), t \in (0,\infty)\}$ and the conditional expectation $\hat{\xi}(s) = \mathbb{E}(\xi(t) \mid \xi(u), u \le s)$, $s < t$, as a nonlinear prediction (approximation) of $\xi(t)$ by $\xi(u), u \le s$. The mean square error of this prediction is $\mathbb{E}(\xi(t) - \hat{\xi}(s))^2 = \|\xi(t) - \hat{\xi}(s)\|^2$. Let $a = t_0 < t_1 < \cdots < t_n = b$ be a partition $\mathcal{P}$ of interval $(a,b) \subset (0,\infty)$. Consider a sum $\mathcal{S}(\mathcal{P}) = \sum_1^n (\xi(t_k) - \hat{\xi}(t_{k-1}))$. Suppose that there exists the random variable $\int_a^b \partial\xi$ - the mean square limit of $\mathcal{S}(\mathcal{P}_n)$ when the partitions become thiner and thiner, $\mathcal{P}_1 \succ \mathcal{P}_2 \succ \dots$ and $\max_{1 \le i \le n} \mid t_i - t_{i-1} \mid \to 0$. We call $\| \int_a^b \partial\xi\|^2$ the nonlinear approximability (in fact predictability) of $\{\xi(t)\}$ on (a,b) . A trivial example is $\xi(t) = W(t)$ - standard Wiener process where $\| \int_a^b \partial\xi\|^2 = b - a$. The example $\xi(t) = W(t)^3$ is less trivial, and $\|W(t)^3 - \widehat{W(t)^3}\|^2 = 27t^2(t-s) - 18t(t-s) + 6(t-s)^3$ and $\| \int_a^b \partial W^3\|^2 = 9(b^3 - a^3)$.

We should note that the approximability may by 0. This is the case if the error $\|\xi(t) - \hat{\xi}(s)\|^2$ is infinitesimal, $s \uparrow t$, of higher order then $t - s$. For instance, the purely nondeterministic process $\{\xi(t)\}$ defined by $\xi(t) = \int_0^t (t-u)dW(u)$ posseses the error $\|\xi(t)\hat{\xi}(s)\|^2 = (t-s)^3/3$, and nonlinear approximability is 0 on any interval (a,b).

In the sequel we shall use the notion of one dimensional Hermite polynomial $H_p(\xi)$ of the defree p in the Gaussian variable ξ. We need the following properties of Hermite polynomials : 1. the inner product is

$$\langle H_p(\xi), H_p(\eta)\rangle = \mathbb{E}(H_p(\xi)H_p(\eta)) = p!\langle \xi, \eta\rangle^p; \tag{1}$$

2. If $\{\xi, \eta_t, t \in T\}$ is set of Gaussian random variables then

$$\mathbb{E}(H_p(\xi) \mid \eta_t, t \in T) = H_p(\hat{\xi}), \quad \hat{\xi} = \mathbb{E}(\xi \mid \eta_t, t \in T), \quad [2] \tag{2}$$

1991 Mathematics subject classification: 60G12.

Supported by Grant 0401A of FNS through Math. Inst. SANU

2. We considered in [1] the random field $\{H_p(t), t \in (0,\infty)^d\}$, $H_p(t) = H_p(\eta(t))$, where $\{\eta(t)\}$ is a mean square continuous Gaussian martingale $\eta(t_1,\dots,t_{i-1},0,$ $t_{i+1},\dots,t_d) = 0$, $F(t) = \|\eta(t)\|^2$, in order to define nonlinear approximability. Martingale $\{\eta(t)\}$ generates the random measure $d\eta$ in the standard manner. Denote by $\eta(\Delta)$ the standard measure of a d-dimensional interval $\Delta = (t-h,t)$ and $\Delta F = \|\eta(\Delta)\|^2$.

Denote $A(t,h) = \cap_{j=1}^n \{s = (s_1,\dots,s_d), s_i < t_j, i \neq j,, s_j < t_j - h_j\}$, $h_j > 0$. Consider the nonlinear approximation of $H_p(t)$ by $\{\eta(s), s \in A(t,h)\} : \hat{H}_p(t,h) = (H_p(t) \mid \eta(s), s \in A(t,h))$. By (2)
$\hat{H}_p(t,h) = H_p\big(\sum_i \eta(t_1,\dots,t_i-h_i,\dots,t_d) - \sum_{i,j}\dots - (-1)^d\eta(t_1-h_1,\dots,t_d-h_d)\big) = H_p(\eta(t)-\eta(s))$. Let $\mathcal{P}$ be a partition of interval $(0,t)$ into subintervals (t_k-h_k,t_k). Consider

$$\mathcal{S}(\mathcal{P}) = \sum_k \big[H_p(t_k) - \hat{H}_p(t_k,h_k)\big] = \sum_k \big[H_p(t_k) - H_p(\eta(t_k)-\eta(\Delta))\big].$$

Let $\mathcal{P}' \succ \mathcal{P}$ means that subintervals of $\mathcal{P}$ are in $\mathcal{P}'$.

Proposition 1. *If $\mathcal{P}' \succ \mathcal{P} \succ \dots$ is a sequence of thiner and thiner partitions with $\max_{\Delta_{nk}\in\mathcal{P}_n} \|\eta(\Delta_{nk})\|^2 \to 0$, when $n \to \infty$, then $\mathcal{S}(\mathcal{P}_n)$ converges in the mean square to the random variable $\int_0^t \partial H_p$ and $\|\int_0^t \partial H_p\|^2 = p!p\int_0^t F^{p-1}(u)dF(u)$.*

Proof. Let $(t-h,t)$ and $(t'-h',t')$ be two arbitrary intervals in $(0,\infty)^d$. Then by (1)

$$\begin{gathered}\big\langle H_p(t) - \hat{H}_p(t,h), H_p(t') - \hat{H}_p(t',h')\big\rangle = \\ \big\langle H_p(t) - H_p(\eta(t)-\eta(\Delta)), H_p(t') - H_p(\eta(t')-\eta(\Delta'))\big\rangle = \\ p!\big[\langle\eta(t),\eta(t')\rangle^p - (\langle\eta(t),\eta(t')\rangle - \langle\eta(t),\eta(\Delta')\rangle)^p - (\langle\eta(t),\eta(t')\rangle - \langle\eta(t'),\eta(\Delta)\rangle)^p + \\ (\langle\eta(t),\eta(t')\rangle - \langle\eta(t),\eta(\Delta')\rangle - \langle\eta(t'),\eta(\Delta)\rangle + \langle\eta(\Delta),\eta(\Delta')\rangle)^p\big] = \\ = p!\big[p\langle\eta(t),\eta(t')\rangle^{p-1}\langle\eta(\Delta),\eta(\Delta')\rangle + o(\|\eta(\Delta)\|^2) + o(\|\eta(\Delta')\|^2)\big].\end{gathered}$$

If $(t-h,t)$ and $(t'-h',t')$ are disjoint, then $\langle\eta(\Delta),\eta(\Delta')\rangle = 0$ and $\langle H_p(t) - \hat{H}_p(t,h), H_p(t') - \hat{H}_p(t',h')\rangle = o(\|\eta(\Delta)\|^2) + o(\|\eta(\Delta')\|^2)$. If $(t-h,t) \subset (t'-h',t')$, then $\langle\eta(t),\eta(t')\rangle = \|\eta(t)\|^2$, $\langle\eta(\Delta),\eta(\Delta')\rangle = \|\eta(\Delta)\|^2$, and $\langle H_p(t) - \hat{H}_p(t,h), H_p(t') - \hat{H}_p(t',h')\rangle = p!p\|\eta(t)\|^{2(p-1)}\|\eta(\Delta)\|^2 + o(\|\eta(\Delta)\|^2) = p!pF^{p-1}(t)\Delta F(t) + o(\Delta F(t))$.

So, for $m > n$ and $\mathcal{P}_m \prec \mathcal{P}_n$, we have $\langle\mathcal{S}(\mathcal{P}_m),\mathcal{S}(\mathcal{P}_n)\rangle = \sum_{\Delta_{nk}\in\mathcal{P}_n} p!p$ $F(t_k)\Delta F(t_k) + o(\Delta F(t_k))$. Hence, $\langle\mathcal{S}(\mathcal{P}_m),\mathcal{S}(\mathcal{P}_n)\rangle \to p!p\int_0^t F^{p-1}(u)dF(u)$, $m,n \to \infty$

We conclude that, according to well known crtieria of the mean square convergence, $\mathcal{S}(\mathcal{P}_m)$ converges in mean square to a random variable $\int_0^t \partial H_p$, $\|\int_0^t \partial H_p\|^2 = p!p\int_0^t F^{p-1}(u)dF(u)$. □

We call the last expression nonlinear approximability of the field $\{H_p(t)\}$ by $\{\eta(u), u \in (0,t)\}$. One can substitute the interval $(0,t)$ by any Borel set in $(0,\infty)^d$ via standard procedure.

3. The definition of nonlinear approximability $\{H_p(t)\}$ enables us to define nonlinear approximability for a larger class of random fields, which are nonlinear transformations of the Gaussian martingale $\{\eta(t)\}$.

Let $\mathcal{H}(t)$ be a Hilbert space — the mean linear closure of polynomials in $\{\eta(t)\}$ centered at expectation: $\mathcal{H}(t) = Cl\{P_n(\eta(s)), s \leq t, n \in \mathbb{N}\}$. The space $\mathcal{H}(t)$ is the ortogonal sum of spaces $\mathcal{H}_p(t)$, $p = 1, 2, \dots$ - the mean square closure of Hermite polynomials $\{H_p(s), s \leq t\}$. Define $\mathcal{H} = \overline{\bigvee_t \mathcal{H}_t}$, $\mathcal{H}_p = \overline{\bigvee_t \mathcal{H}_p(t)}$. It is obvious that

$$\mathcal{H}(t) = \mathbb{E}(\mathcal{H} \mid \eta(s), s \leq t) = \oplus \sum_{p=1}^{\infty} \mathbb{E}(\mathcal{H} \mid \eta(s), s \leq t) = \oplus \sum_{p=1}^{\infty} \mathcal{H}_p(t).$$

Let for the random field $\{\xi(t), t \in (0,\infty)^d\}$ we have $\xi(t) \in \mathcal{H}(t)$, $t \in (0,\infty)^d$. It is easy to show that $dH_p(t)$ is the ortogonal random measure and that any $\xi \in \mathcal{H}_p(t)$ has the representation $\xi = \int_0^t f(u)dH_p(u)$, $\int_0^t f^2(u)p!dF^p(u) < \infty$. In that way the field $\{\xi(t)\}$ has the representation

$$\xi(t) = \sum_{p=1}^{\infty} \int_0^t g_p(t,u)dH_p(u)$$

where nonrandom functions $g_p(t,.)$ satisfy

$$||\xi(t)||^2 = \sum_{p=1}^{\infty} p! \int_0^t g_p^2(t,u)dF^p(u)$$

We define nonlinear approximation of $\{\xi(t)\}$ by $\{\eta(s), s \leq t\}$ using the sum

$$\mathcal{S}(\mathcal{P}) = \sum_{t_k \in \mathcal{P}_m} [\xi(t_k) - \hat{\xi}(t_k, h_k)] \tag{3}$$

as it was done in point **2**.

Proposition 2. *If the functions $g_p(t,u)$ are continuous then there exists the limit $\int_0^t \partial\xi$ of $\mathcal{S}(\mathcal{P}_m)$ and*

$$|| \int_0^t \partial\xi ||^2 = \sum_{p=1}^{\infty} || \int_0^t g_p(t,u)\partial H_p(u)||^2 = \sum_{p=1}^{\infty} p!p \int_0^t g_p^2(u,u)F^{p-1}(u)dF(u).$$

Proof. Consider $\xi(t_k)-\hat{\xi}(t_k,h_k)$ as in (3):

$$\xi(t_k)-\hat{\xi}(t_k,h_k)=$$

$$\sum_{p=1}^{\infty}[\int_0^t g_p(t,u)dH_p(u)-\mathbb{E}(\int_0^t g_p(t,u)dH_p(u)\mid\eta(\Delta)\in A(t_k,h_k))]=$$

$$=\sum_{p=1}^{\infty}[\int_{\Delta_k} g_p(t,u)dH_p(u)-\mathbb{E}(\int_{\Delta_k} g_p(t,u)dH_p(u)\mid\eta(\Delta)\in A)]=$$

$$=\sum_{p=1}^{\infty}[g_p(t_k,u_k)\Delta_k H_p-\mathbb{E}(g_p(t_k,u_k)\Delta_k H_p\mid\eta\in A)]$$

$$=\sum_{p=1}^{\infty}[g_p(t_k,u_k)[H_p(t_k)-\hat{H}_p(t_k,h_k)]]$$

where $u_k\in\Delta_k$. Hence,

$$\mathcal{S}(\mathcal{P}_m)=\sum_k\sum_{p=1}^{\infty}[g_p(t_k,u_k)[H_p(t_k)-\hat{H}_p(t_k,h_k)]]\to\sum_{p=1}^{\infty}\int_0^t g_p(u,u)\partial H_p(u),$$

in the mean square, and

$$\|\int_0^t\partial\xi\|^2=\sum_{p=1}^{\infty}pp!\int_0^t g_p^2(u,u)F^{p-1}(u)dF(u).\quad\Box$$

References

[1] Z. A. Ivković, *A stochastic measure and nonlinear approximation of some random fields*, Teor. Prob. and its Appl. (Moscow) **37** (1992), 594-598.

[2] Z. Ivković and Z. Lozanov, *Hermite polynomials of Gaussian processes*, Rev. of Res. Math. Ser., Fac. of Sci. **12** (1982), 105-115.

Matematički fakultet,
Studentski trg 16,
11000 Beograd

Zbornik radova Filozofskog fakulteta u Nišu
Serija Matematika 6:2 (1992), 307–320

FILOMAT '92, Niš, October 8–10, 1992

ON SOME UNIVARIATE NON-LINEAR TIME SERIES MODELS

JOVAN MALIŠIĆ

ABSTRACT. *This paper is a short review on some results in the theory of univariate non-linear time series models. Bilinear models, threshold models, exponential models, Gamma models, AREX models and other non-linear and non-Gaussian models are considered. Some of non-linearity and non-Gaussianity tests are studied.*

1. Introduction

The mathematical theory of time series has been an area of considerably activity of statisticians in recent decades. If one studies a time series data the first step in the analysis is the selection of a suitable statistical model (or a class of models) for the data. Two very simple types of time series are the sequences of independent and identically distributed (i.i.d.) and sequences of uncorrelated random variables. A good method of describing a time series is to express it as a function of such more elementary time series. In some applications it is possible to express the observed time series as linear combinations of uncorrelated or of i.i.d. random variables.

A lot of current papers on time series are based on stationary and linear models, such as the moving average (MA) models, autoregressive (AR) models or autoregressive-moving average (ARMA) models.

The reasons for the linearity assumptions, firstly, are that these models often provide a good approximation to the real time series data and, secondly, they can easily be analyzed. However, there are some situations for which linear models may not adequately describe the underlying random mechanism. Therefore one may ask if there exists a more general class of models which can provide a better fit to reality. This leads us to consider some classes of non-linear time series models.

ARMA models are rarely considered in respect their marginal distributions. But, if one studies ARMA models in respect of their marginal distributions, for statistical convenience, the distribution in such a model is often assumed to be a Gaussian distribution. In this case, the independent innovation sequence needs to be Gaussian, as well. However, the problem is much harder if we consider non-Gaussian distributions.

In the present paper we shall consider some simple univariate time series models: linear and non-linear, Gaussian and non-Gaussian.

2. Linear processes

1991 Mathematics subject classification: 62M20.

The time series $\{X_t,\ t \in D = \{0, \pm 1, \dots\}\}$, defined by the difference equation:

$$X_t = \sum_{j=-q}^{q} a_j \varepsilon_{t-j}, \tag{2.1}$$

where q is a nonnegative integer, a_j, $j = \overline{-q, q}$ are real numbers with $a_{-q} \neq 0$ and ε_t are uncorrelated random variables with $E\varepsilon_t = 0$, $D\varepsilon_t = E\varepsilon_t^2 = \sigma^2 > 0$, for all $t \in D$, is a two-sided finite moving average time series.

The time series defined by

$$X_t = \sum_{j=0}^{q} a_j \varepsilon_{t-j}, \tag{2.2}$$

where $a_0 \neq 0$, $a_q \neq 0$, are called one-sided moving average of order q (or $MA(q)$ model).

We shall call the time series defined by

$$X_t = \sum_{j=-\infty}^{+\infty} a_j \varepsilon_{t-j} \quad \text{or} \quad X_t = \sum_{j=0}^{+\infty} a_j \varepsilon_{t-j}, \tag{2.3}$$

a infinite moving average time series (two-sided and one-sided MA models, respectively).

Moving average processes have many interesting properties. Among other things, it is clear to see that such a model is stationary. The serial dependency of the $MA(q)$ models (i.e. the covariances or autocorrelation coefficients) stops at the lag q. If we consider, for simplicity, the first order moving average

$$X_t = \varepsilon_t + a\varepsilon_{t-1},\ t \in D,$$

then

$$\rho(1) \equiv \operatorname{corr}(X_t, X_{t+1}) = \frac{a}{a+1},$$

achieves a maximum of 0.5, for $a = 1$, and a minimum of -0.5, for $a = -1$, for any $\rho \in (0, 0.5)$ (or for any $\rho \in (-0.5, 0)$) there are two values of a yielding such ρ, one in the interval $(0, 1)$ (or in $(-1, 0)$) and one in the interval $(1, +\infty)$ (or in $(-\infty, -1)$).

Other simple models, frequently used in the time series analysis, are autoregressive processes. This is the case where we have a chain dependency on $X_{t-1}, \dots, X_{t-p}$. In others words, define

$$X_t = \sum_{j=0}^{p} a_j X_{t-j} = \varepsilon_t,\ t \in D, \tag{2.4}$$

where a_j are some real constants, $a_0 \neq 0$, $a_p \neq 0$ and ε_t are uncorrelated $(0; \sigma^2)$ random variables. Then the process X_t is said to be an autoregressive process of order p (or $AR(p)$).

One of the important problems for such model is to find conditions under which the process is a stationary process. Let the roots of the characteristics equation

$$\sum_{j=0}^{p} a_j r^{p-j} = 0,$$

be less than one in absolute values and let the weights u_j, $j = 0, 1, 2, \dots$ be defined by the solution of the homogeneous difference equation

$$u_j + a_1 u_{j-1} + \cdots + a_p u_{j-p} = 0,$$

subject to the some boundary conditions. Then X_t is the mean square limit

$$X_t = \sum_{j=0}^{\infty} u_p \varepsilon_{t-j}, \tag{2.5}$$

and X_t is a stationary process.

The process $\{X_t\ ,\ t \in D\}$ defined by

$$\sum_{j=0}^{p} a_j X_{t-j} = \sum_{i=0}^{q} b_i \varepsilon_{t-i}, \tag{2.6}$$

where $a_0 = 1$, $a_q \neq 0$, and where $\{\varepsilon_t,\ t \in D\}$ is a sequence of uncorrelated $(0; \sigma^2)$ random variables, is called autoregressive-moving average time series of order (p, q) (or $ARMA(p, q)$. It is easy to find conditions for stationarity: the $ARMA(p, q)$ process is stationary if and only if

$$r^p + a_1 r^{p-1} + \cdots + a_{p-1} r + a_p \neq 0,$$

for $|r| \geq 1$.

The process is said to be a linear process if it can be written in the form

$$X_t = \sum_{j=-\infty}^{\infty} c_j \varepsilon_{t-j}, \tag{2.7}$$

where c_j are some real constants, $\Sigma c_k^2 < \infty$ and ε_t is a white noise, i.e. the sequence of uncorrelated $(0; \sigma^2)$ random variables.

Linear processes have a lot of simple and "good" properties. For example, if the sequence ε_t is a Gaussian sequence, then X_t is Gaussian, as well.

The statistical analysis of a linear process is relatively simple. We can estimate the parameters in a linear model by using the method of moments, the method of maximum likelihood or the method of least squares and we can find the properties of these estimates.

3. On some non-linear autoregressive processes

One of the first non-linear models, studied in the statistical literature, was the Jones's model (Jones (1977a)*, (1977b)*, (1978)*[1]) defined by

$$X_t = \lambda(X_{t-1}) + \varepsilon_t, \tag{3.1}$$

where $\lambda(\cdot)$ is a given non-linear function and ε_t is a strict white noise, i.e. a sequence of i.i.d. random variables with $E\varepsilon_t = 0$ and $D\varepsilon_t = \sigma^2 > 0$, for all t. Following Jones we call the function $\lambda(\cdot)$ the autoregression function of the process X_t and the process X_t non-linear autoregressive process of the first order ($NLAR(1)$ process).

One of the important problems is to find conditions, under which the Jones's process is stationary. Following Anděl (1984*, 1989a, 1989b) we can find some.

[1] The papers cited in the work of Anděl (1989b) (see [4] in references) are marked as $(\dots)^*$.

"negative" and some "positive" conditions about stationarity and ergodicity. For instance,

1. if $P\{X_0 > 0,\ X_1 > 1,\ X_2 > 2, \dots\} > 0$, then the process X_t is a nonstationary process;

2. if $\lambda(\cdot)$ is a continuous function and if there exist constants $k > 0,\ \delta > 0$, such that

$$E\left(|\lambda(x) + \varepsilon_t| - |x|\right) \le -\delta \quad \text{for} \quad |x| > k,$$

then the process X_t is ergodic.

Jones (1977b)* studied the case in which the function λ depends on an unknown parameter θ:

$$X_t = \lambda(X_{t-1}; \theta) + \varepsilon_t. \tag{3.2}$$

In the special case, where the function λ depends on the parameter θ linearly, i.e.

$$X_t = \theta\lambda(X_{t-1}) + \varepsilon_t, \tag{3.3}$$

Jones (1978)* investigated conditions for stationarity and gave the properties of the least squares estimator $\widehat{\theta}_n$ for θ, given $X_0, X_1, \dots, X_n$:

$$\widehat{\theta}_n = \left[\sum_{k=1}^{n} X_k\lambda(X_{k-1})\right] \Big/ \left[\sum_{k=1}^{n} \lambda^2(X_{k-1})\right].$$

Andel (1989b) considered the model

$$X_t = X_{t-1}^{1/2} + \varepsilon_t, \tag{3.4}$$

where ε_t is a strict white noise such that $P(\varepsilon_t = 0) = P(\varepsilon_t = 1) = 0.5$. He generalized the model (3.4) to the model

$$X_t = \lambda \cdot X_{t-1}^{1/h} + \varepsilon_t, \tag{3.5}$$

where $P(\varepsilon_t = 0) = p$, $P(\varepsilon_t = c) = 1 - p$, $\lambda > 0$, $h > 0$, $p \in (0,1)$, $c > 0$. It was shown by Andel how to calculate the values of the stationary distribution function F of X_t and its moments.

Doukhan and Chindes (1980) considered the Markovian model

$$X_{n+1} = f(X_n) + \varepsilon_n, \tag{3.6}$$

where f is an unknown function and $\{\varepsilon_n\}$ is a white noise of unknown law. The estimates of the density of the transition probability were established, as well.

A general non-linear model of the type

$$X_n = \varphi(X_{nt-1}) + \varepsilon_{n-1}, \tag{3.7}$$

was studied by Mokkadem (1985), as well. He obtained some sufficient conditions for ergodicity and geometric ergodicity of the process.

4. Bilinear models

In some econometric situations the quotient

$$\frac{\Delta X_t}{X_{t-1}} = \frac{X_t - X_{t-1}}{X_{t-1}}$$

has a stationary behaviour, i.e. we can set

$$\frac{X_t - X_{t-1}}{X_{t-1}} = \varepsilon_t \quad \text{or} \quad \frac{X_t - X_{t-1}}{X_{t-1}} = \varepsilon_t + b\varepsilon_{t-1},$$

where ε_t is a strict white noise. From here we can write

$$X_t = X_{t-1} + \varepsilon_t X_{t-1},$$

or

$$X_t = X_{t-1} + \varepsilon_t X_{t-1} + b\varepsilon_t X_{t-1}.$$

These two models are two examples of so-called bilinear models.

The general bilinear model $BL(p, q, r, s)$ is the model defined by

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{k=0}^{q} b_k \varepsilon_{t-k} + \sum_{l=0}^{r} \sum_{m=0}^{s} \beta_{lm} \varepsilon_{t-l} X_{t-m}, \tag{4.1}$$

where ε_t is a strict white noise and $b_0 = 1$ (Granger and Andersen (1978)*, Subba Rao (1981)*, Tong (1981)*, Pham (1985a*, 1985b*), Andel (1989)*, etc.). The innovation series at time t is

$$\eta_t = \varepsilon_t \left(b_0 + \sum_{m=1}^{s} \beta_{0m} X_{t-m} \right).$$

If $\beta_{0m} \neq 0$, for some m, then the variance of η_t is

$$d\eta_t = E\eta_t^2 = E\varepsilon_t^2 \cdot E(b_0 + \Sigma\beta_{0m} X_{t-m})^2 \neq \text{const.}.$$

Such a BL model is usually called the BL model with heterogeneous errors. If $\beta_{0m} = 0$, for all m, then the corresponding model is a BL model with homogeneous errors.

The BL model defined by

$$X_t = \sum_{j=1}^{P} \sum_{k=1}^{Q} \beta_{jk} \varepsilon_{t-j} X_{t-k} + \varepsilon_t, \tag{4.2}$$

is so-called completely BL model. If $\beta_{jk} = 0$, for all $j > k$, we have super diagonal completely BL model, if $\beta_{jk} = 0$, for all $j \neq k$, we have diagonal completely BL model and if $\beta_{jk} = 0$, for all $j < k$, we have subdiagonal model. It can be shown that the superdiagonal BL model is stationary in wide sense if and only if the characteristic equation

$$\sigma^2 \sum_{j=1}^{P} \sum_{k=1}^{Q} \beta_{jk}^2 z^k - 1 = 0$$

has no roots inside the unit circle (Andel (1989b)). For subdiagonal and for diagonal models the results are similar to that of superdiagonal BL model.

It should be noted that the statistical analysis of the general $BL(p, q, r, s)$ model is very difficult, because of the fact that we have products of members of the sequences ε_t and X_t. In order to facilitate the problem we must consider some special cases.

Tong (1981)* considered the first-order BL model

$$X_t = aX_{t-1} + \varepsilon_t + b\varepsilon_t X_{t-1}, \tag{4.3}$$

where ε_t is a Gaussian white noise with $E\varepsilon_t = 0$, $D\varepsilon_t = \sigma^2$, for all $t \in D$. He established the ergodicity conditions, proved that not all moments exist and showed that the normality assumption of ε_t can be replaced by an assumption of an absolutely continuous distribution with finite mean.

Pham and Tran (1981)* obtained some properties of the first-order BL model

$$X_t = aX_{t-1} + \varepsilon_t + \beta\varepsilon_{t-1}X_{t-1}. \tag{4.4}$$

They established necessary and sufficient conditions for the existence of a strictly stationary process satisfying (4.4) and studied the invertibility problem, as well. The estimates of the parameters were obtained by a modified least squares method.

Subba Rao (1981)* studied the general BL models $BL(p,0,p,1)$ and $BL(p,0,p,s)$. He derived conditions for stationarity and invertibility. The expressions for covariances and estimations of the parameters of the models were considered, as well, in this paper.

Kim, Billiard and Basawa (1990) considered the problem of estimation of the parameter b in the first-order diagonal BL model

$$X_t = \varepsilon_t + b\varepsilon_{t-1}X_{t-1}, \tag{4.5}$$

where ε_t is a Gaussian white noise with zero mean and variance σ^2. The asymptotic normality of the moment estimator of b was established for both cases when σ^2 is known and σ^2 is unknown.

Liu and Chen (1990) considered the $BL(p,0,1,1)$ model

$$X_t = \sum_{j=1}^{n} a_j X_{t-1} + \varepsilon_t + b\varepsilon_{t-1}X_{t-1}, \tag{4.6}$$

where ε_t is i.i.d. $(0,\sigma^2)$ random sequence. They showed that their model could be expressed as an $ARMA$ model and using this expression they proved that the moment estimators of $\{a_j\}$ were unique and consistent. If ε_t is a Gaussian $\mathcal{N}(0;\sigma^2)$ sequence or if its first few order moment structures are known a prior using up to the third-order moment, the parameters b and σ^2 can also be estimated consistently and these estimates are shown to be asymptotically normal.

Liu and Brockwell (1988) derived a sufficient condition for the existence of a strictly stationary solution of a general BL model. Their condition reduces to the condition of Pham and Tran (1981)* and Bhaskara Rao et al (1983), in the special case which they consider. A solution is constructed which is shown to be causal, stationary and ergodic.

The problem of the identification of some BL time series models is studied by Kumar (1986), Li (1984), Ozaki and Oda (1977)*, Priestley (1980), Andel (1988), etc.

In the paper of Stensholt and Tjostheim (1987)* the definition of multiple BL model is given. Some sufficient conditions for the existence of strictly stationary solutions are obtained. The multiple BL time series models were studied by many other authors, as well.

5. Threshold models

The threshold autoregressive (TAR) process $\{X_t,\ t \in D\}$ of the first order with one threshold r in the process is given by the equation

(5.1) $$X_t = \begin{cases} \alpha X_{t-1} + \varepsilon_t & \text{if } X_{t-1} \le r, \\ \beta X_{t-1} + \varepsilon_t & \text{if } X_{t-1} > r, \end{cases}$$

where α and β are autoregressive parameters (usually, real numbers), r is a real constant and $\{\varepsilon_t,\ t \in D\}$ is an error sequence. This is so-called $TAR(1)$ process or more precisely, $SETAR(2;1,1)$ process (i.e. self-exciting threshold autoregressive process) first proposed by Tong (1978). The sequence X_t in (5.1) consists of two submodels - piecewise linear autoregressive processes. Model (5.1) includes the usual $AR(1)$ model as a special case by setting r equal to $(-\infty)$.

Tong and Lim (1980) consider several examples of TAR model which provide better fits to the real cyclical data then linear models.

The paper of Petruccelli and Woolford (1984)* deals with the $SETAR(2;1,1)$ model with $r = 0$ and i.i.d. sequence ε_t, having mean 0, positive variance σ^2 and strictly positive density on R. Their model may be written in the form

(5.2) $$X_t = \alpha X_{t-1}^- + \beta X_{t-1}^+ + \varepsilon_t,$$

where $Z^- = \min\{Z, 0\}$, $Z^+ = \max\{Z, 0\}$. They proved that X_t is ergodic if and only if α and β satisfy the conditions:

$$\alpha < 1, \quad \beta < 1, \quad \alpha\beta < 0.$$

Under the assumption of a finite absolute moment of order $2 + \delta$, for some $\delta > 0$, in the innovation sequence ε_t (so that the stationary distribution for X_t has finite second-order moment) it is shown that the least squares estimators for α and β are

$$\widehat{\alpha} = \sum_1^n Z_n Z_{n-1}^+ \Big/ \sum_1^n {Z_{n-1}^+}^2,$$

$$\widehat{\beta} = \sum_1^n Z_n Z_{n-1}^- \Big/ \sum_1^n {Z_{n-1}^-}^2,$$

and the corresponding estimator for σ^2:

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_1^n \left(Z_n - \widehat{\alpha} Z_{n-1}^+ - \widehat{\beta} Z_{n-1}^- \right)^2.$$

In the ergodic case the estimators $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}^2$ are consistent estimators of α, β and σ^2, respectively. If the error sequence has Gaussian distribution then $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}^2$ are also the maximum likelihood estimators.

These results were generalized by Chan et al (1988)* to the $SETAR$ models of the first-order with more than one threshold.

Petruccelli (1986) studied the $SETAR(2;,1)$ model

(5.3) $$X_t = \begin{cases} \alpha X_{t-1} + \varepsilon_t, & \text{if } X_{t-1} \le r, \\ \beta X_{t-1} + \eta_t, & \text{if } X_{t-1} > r, \end{cases}$$

where α and β are real constants, ε_t and η_t are i.i.d. random variables with $E\varepsilon_t = E\eta_t = 0$, $D\varepsilon_t = \sigma_\varepsilon^2$, $D\eta_t = \sigma_\eta^2 = \sigma_\varepsilon^2 > 0$ and $\{\varepsilon_t\}$ and $\{\eta_t\}$ are independent. He considers the least squares estimators of α, β and r and proves their consistency. Also, he proves inconsistency of estimators in the case $\sigma_\varepsilon^2 \ne \sigma_\eta^2$.

More general threshold autoregressive model with more than one threshold first proposed by Tong (1983)*. Let $B_1 \cup B_2 \cup \ldots \cup B_k = R$, $B_i \cap B_j = \varnothing$ and let the process X_t is defined by

$$(5.4) \quad X_t = \begin{cases} a_{11}X_{t-1} + a_{21}X_{t-2} + \cdots + a_{n_1 1}X_{t-n_1} + \varepsilon_t, & \text{if } y_{t-1} \in B_1, \\ a_{12}X_{t-1} + a_{22}X_{t-2} + \cdots + a_{n_2 2}X_{t-n_2} + \varepsilon_t, & \text{if } y_{t-1} \in B_2, \\ \cdots \quad \cdots & \cdots \\ a_{1k}X_{t-1} + a_{2k}X_{t-2} + \cdots + a_{n_k k}X_{t-n_k} + \varepsilon_t, & \text{if } y_{t-1} \in B_k. \end{cases}$$

Usually $B_1 = (-\infty, r_1)$, $B_2 = [r_1, r_2), \ldots$, $B_k = [r_{k-1}, \infty)$. Here $r_1, r_2, \ldots, r_{k-1}$ are given real numbers (called thresholds). Then X_t is a TAR process with k thresholds and instrumental time series Y_t. If $Y_t = X_t$, then X_t is $SETAR$ process.

Chan and Tong (1986a)* studied the problem of estimating the threshold in the model

$$(5.5) \qquad X_t = \begin{cases} a_0 + a_1X_{t-1} + \cdots + a_pX_{t-p} + \varepsilon_t, & \text{if } X_{t-d} \le r, \\ b_0 + b_1X_{t-1} + \cdots + b_pX_{t-p} + \varepsilon_t, & \text{if } X_{t-d} > r, \end{cases}$$

where $r \in R$, $d \in \{1, 2, \ldots\}$, $p \in \{0, 1, 2, \ldots\}$ and error sequence ε_t is a sequence of i.i.d. random variables such that ε_t and X_s, $s < t$ are independent.

Andel et al (1984)* considered the model (5.1) with $r = 0$, $\alpha = -\beta \in (0, 1)$, $\varepsilon_t : \mathcal{N}(0; \sigma^2)$ and established that there exists a unique stationary distribution of X_t, with density

$$g(x) = \sqrt{2(1-\alpha^2)/\pi}\, \exp\left\{-(1-\alpha^2)x^2/2\right\} \phi(-\alpha x),$$

where ϕ is the distribution function of $\mathcal{N}(0; 1)$. Andel and Barton (1986)* found an explicit formula for the stationary density in the model (5.1) where ε_t has a Cauchy distribution.

Chen and Tsay (1991) establish a necessary and sufficient condition for the geometrical ergodicity for the general $SETAR(2; 1, 1)$ model.

Wecker (1981)* considers a threshold moving average model defined by

$$(5.6) \qquad X_t = \varepsilon_t + \left(b_1\varepsilon_{t-1}^+ + \cdots + b_m\varepsilon_{t-m}^+\right) + \left(c_1\varepsilon_{t-1}^- + \cdots + c_m\varepsilon_{t-m}^-\right),$$

where $x_t^+ = \max\{x_t, 0\}$, $x_{t-1}^- = \min\{x_t, 0\}$ and b_k, c_k are given coefficients. The problem of stationarity and invertibility are studied as well.

Fuchs (1984)* generalized Wecker's results to the multidimensional case.

Andel and Fuchs (1987)* found the explicit form for the density function in the threshold moving average model

$$(5.7) \qquad X_t = \begin{cases} \varepsilon_t + \beta\varepsilon_{t-1} & \text{if } \varepsilon_{t-1} \le 0, \\ \varepsilon_t - \beta\varepsilon_{t-1} & \text{if } \varepsilon_{t-1} > 0, \end{cases}$$

with $\varepsilon_t : \mathcal{N}(0; 1)$.

Chan (1990) considers the problem of determining whether a TAR model fits a stationary time series significantly better than an AR model does. A test statistic which is equivalent to the conditional likelihood ratio test statistic, when the noise is normally distributed, is proposed in the paper.

The likelihood ratio test statistics for TAR model are also considered in the joint paper of Chan and Tong (1990).

6. Linearity and non-linearity tests

The higher order spectra have very important role in the analysis of non-linear time series. Their usefulness is further strengthened by the fact that no assumption on the underlying model governing the time series is necessary.

If the stationary time series is linear, the second order spectra contains a lot of useful informations present in the series. If the series is non-linear, the second-order spectra will not adequately characterize the series. It means that one has to consider the higher order spectra to study the possible departures from linearity.

Let $\{X_t,\ t \in D\}$ be a real stationary linear process with mean $EX_t = m$, spectral density $g(\omega)$, finite third-order cumulant

$$c(u,v) = E\left[(X_t - m)(X_{t+u} - m)(X_{t+v} - m)\right], \tag{6.1}$$

and bispectral density function

$$g(\omega_1,\omega_2) = \frac{1}{(2\pi)^2}\sum_{-\infty}^{+\infty}\sum_{-\infty}^{+\infty} c(t_1,t_2)e^{-(t_1\omega_1 + t_2\omega_2)}, \quad -\pi \le \omega_1,\ \omega_2 \le \pi. \tag{6.2}$$

One can show that for Gaussian process $g(\omega_1,\omega_2) = 0$ for all ω_1 and ω_2. However, it may happen that the process is linear but not necessary Gaussian (Suba Rao and Gabr (1980)*). In such a situation the ratio

$$r_{ij} = \mu_3^2\,/(2\pi\sigma^2)$$

is constant. In other words, the fact that r_{ij} is a constant for all i and j over a specific range of values is a good test for adequacy of linearity. (For more details see: Subba Rao and Gabr (1980)* or Andel (1989b)).

Hinich (1982)* pointed out that if the relationship between X_t and the innovation sequence ε_t is non-linear, then X_t is non-Gaussian sequence even if ε_t is Gaussian. He considered the Fourier transform of the third moments $c(u,v)$ as a simple estimator of the bispectral density. This sample bispectrum is used to construct a statistic to test whether the bispectral density is non-zero. A rejection of the null hypothesis implies a rejection of the Gaussianity. Another test statistic is presented for testing linearity.

Keenan (1985)* adopted the Tukey's idea (Tukey (1949)) of one degree of freedom test for non-additivity to derive a time-domain statistics, e.g. bispectrum, for discriminating between linear and non-linear models. More precisely, Keenan tests the null-hypothesis

$$H_0: \quad X_t = m + \sum_{k=0}^{n} a_k\varepsilon_{t-k},$$

against

$$H_1: \quad X_1 = m + \sum_{k=0}^{n} a_k\varepsilon_{t-k} + \sum_{i=0}^{P}\sum_{j=0}^{Q} b_{ij}\varepsilon_{t-i}\varepsilon_{t-j}.$$

Tsay (1986)*, also, generalized the idea of Tukey's one degree of freedom for non-additivity test. The starting point in his paper is the following representation of a stationary time series X_t:

$$X_t = m + \sum_{i=-\infty}^{\infty} b_i \varepsilon_{t-i} + \sum_{i,j=-\infty}^{\infty} b_{ij} \varepsilon_{t-i}\varepsilon_{t-j} + \sum_{i,j,k=-\infty}^{\infty} b_{ijk}\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-k} + \dots ,$$

where $m = EX_t$, ε_t is a white noise and the fact that if any of the higher order coefficients $b_{ij}, b_{ijk}, \dots$ is non-zero then X_t is non-linear.

Wecker (1981)* tests the hypothesis symmetry

$$H_0 : \quad b_j = c_j \quad \text{for all } j$$

in the threshold moving average model (5.6). His test statistic is

$$T = 2\log[(\widehat{\sigma}_s / \widehat{\sigma}_a)^n],$$

where n is the sample size and $\widehat{\sigma}_s$ and $\widehat{\sigma}_a$ are standard deviations of the innovation sequence under the hypothesis symmetry (H_0) and asymmetry, respectively. If H_0 holds, then T is asymptotically χ^2_m distributed.

Chan and Tong (1986b)* studied the threshold model

$$(6.3) \qquad X_t = \begin{cases} a_0 + \sum_{k=1}^{p} a_k X_{t-k} + \varepsilon_t, & X_{t-1} \le r, \\ b_0 + \sum_{k=1}^{p} b_k X_{t-k} + \varepsilon_k, & X_{t-1} > r. \end{cases}$$

If the hypothesis $H_0 : a_j = b_j$ for all $j = \overline{0,p}$ holds, then X_t is a linear time series. The test is based on the statistic

$$\lambda = \left(\widehat{\sigma}^2_{NL} / \widehat{\sigma}^2_L\right)^{n/2},$$

where n is the sample size. $\widehat{\sigma}^2_{NL}$ and $\widehat{\sigma}^2_L$ are calculated under the linearity hypothesis H_0 and non-linearity hypothesis H_1, respectively.

Saikkonen and Luukkonen (1988) consider Lagrange multiplier tests of linear $ARMA$ models against bilinear and exponential autoregressive alternatives.

Luukkonen et al (1988a)* study the use of various Lagrange multiplier tests in testing linearity against some non-linear alternatives in univariate models of time series.

Luukkonen et al (1988b)* present three tests for testing linearity against smooth transitions autoregressive model when the innovation sequence is $\mathcal{N}(0;\sigma^2)$ distributed and discuss their properties (among other things, the power of the tests).

Chan (1990) considers the problem of determining whether a threshold AR model fits a stationary time series significantly better that an AR model does. A test statistic which is equivalent to the conditional likelihood ratio test statistic when the noise in normally distributed is proposed. (See, also, Chan and Tong (1990)).

An Hong-zhi and Cheng Bing (1991) introduce a Kolmogorov-Smirnov type statistic to test non-linearity in time series and study its asymptotic distribution.

7. Exponential, gamma and other non-Gaussian processes

One of the simplest models for creating dependent random variables is $AR(1)$ model, model defining by the stohastic difference equation

$$X_t = \rho X_{t-1} + \xi_t, \quad t \in D,$$

where $\rho =$ const. ($|\rho| < 1$) and ξ_t is the sequence of i.i.d. random variables. If X_t is normally distributed so the innovation sequence ξ_t is: if $X_t : \mathcal{N}(m;\sigma^2)$ (for all t) then $\xi_t : \mathcal{N}\left((1-\rho)m;\sigma^2(1-\rho^2)\right)$.

A natural question is: Is such result valid for non-Gaussian processes? For instance, in the exponential case $\mathcal{E}(\lambda)$, $\lambda > 0$.

In Gaver and Lewis (1980) it is shown that if $X_t : \mathcal{E}(\lambda)$, $\lambda > 0$, then ξ_t is the following probability mixture

$$\xi_t = \begin{cases} 0 & \text{with probability (w.p.) } \rho \\ \varepsilon_t & \text{w. p. } 1-\rho \end{cases},$$

where ε_t is a sequence of i.i.d. $\mathcal{E}(\lambda)$ random variables. Then we have, so-called, $EAR(1)$ model

$$X_t = \begin{cases} \rho X_{t-1} & \text{w.p. } \rho, \\ \rho X_{t-1} + \varepsilon_t & \text{w.p. } 1-\rho. \end{cases} \tag{7.1}$$

Another exponential first-order AR model is obtained by interchanging the i.i.d. variables ε_t and X_{t-1}. So

$$X_t = \begin{cases} X_{t-1} + \alpha\varepsilon_t & \text{w.p. } 1-\alpha, \\ \alpha\varepsilon_t & \text{w.p. } \alpha, \end{cases} \tag{7.2}$$

so-called $TEAR(1)$ model, discussed by Lawrance and Lewis (1981).

A new exponential model is the modification of the $TEAR(1)$ model in which the X_{t-1} is scaled by a coefficient β. This gives $NEAR(1)$ model (Lawrance and Lewis (1981)):

$$X_t = \begin{cases} \beta X_{t-1} + \varepsilon_t & \text{w.p. } \alpha, \\ \varepsilon_t & \text{w.p. } 1-\alpha. \end{cases} \tag{7.3}$$

A new more general $AR(1)$ type model of the sequences of random variables is said to be an $AREX(1)$ model introduced in Mališić (1987) if it is defined by the following difference equation

$$X_t = \begin{cases} \xi_t & \text{w.p. } p_0, \\ \alpha X_{t-1} & \text{w.p. } p_1 \\ \beta X_{t-1} + \xi_t & \text{w.p. } q_1, \end{cases} \tag{7.4}$$

where $0 < \alpha, \beta < 1$, $0 \le p_0, p_1, q_1 \le 1$, $p_0 + p_1 + q_1 = 1$, and ξ_t is a sequence of random variables chosen such that $X_t : \mathcal{E}(\lambda)$, $\lambda > 0$, for all t.

Model (7.4) belongs to the class of AR models with random coefficients (or doubly stohastic models) introduced and studied by Andel (1976)*, Nichols and Quinn (1982)*, Tjostheim (1986a)*, etc. Model (7.4) gives

$$X_t = A_t X_{t-1} + B_t \xi_t, \quad t \in D, \tag{7.5}$$

where joint distribution of (A_t, B_t) is the following: $P\{(A_t, B_t) = (0,1)\} = p_0$, $P\{(A_t, B_t) = (\alpha, 0)\} = p_1$, $P\{(A_t, B_t) = (\beta, 1)\} = q_1$. It can be shown that there exists a stationary solutuon X_t if and only if

$$0 \le p_0\beta + p_1 \le \alpha \le \beta < 1.$$

Some special cases of the $AREX(1)$ model are:

1° $p_0 = 0$, $\alpha = \beta = p_1 = \rho \Rightarrow EAR(1)$;

2° $p_0 = 1-\alpha$, $p_1 = 0 \Rightarrow NEAR(1)$, etc.

The problem of the parameter estimations, predictions or mixed exponential distributions with negative weights in $AREX(1)$ models are studied by Popović (1988), Jevremović (1991) and Jevremović and Mališić (1992).

In Gaver and Lewis (1980) and in Lawrance (1982) it was shown that there exists an innovation sequence ξ_t such that the sequence X_t generated by $AR(1)$ scheme, is Gamma distributed. Some new results are given in Sim (1990).

Ed McKenzie (1985a, 1985b, 1986) introduced and discussed in details an integer-valued discrete analogue of the $AR(1)$ process. See, also, the paper of Ed McKenzie (1987) on negative binomial autoregression, the paper of Ed McKemzie (1986) on negative binomial and geometric autoregression and some other papers from References.

Some other interesting problems for higher order autoregressive models are studied in Lawrance and Lewis (1985), in Smith (1986), in Billard and Mohamed (1991), in Karlsen and Tjostheim (1988), in Dewald and Lewis (1985), in Alzaid and Al-Osh (1990), etc.

REFERENCES[1)]

[1] ALZAID,A.A. AND AL-OSH,M.A. (1990), *An integer-valued p-th order autoregressive structure* ($INAR(p)$) *process*, J. Appl. Prob. 27, 314-324.

[2] ANDEL,J. (1988), *On* $AR(1)$ *process with exponential white noise*, Commun. Statist. - Theory Meth. 17 (5), 1481-1495.

[3] ANDEL,J. (1989a), *Non-negative autoregressive process*, J. Time Series Anal., Vol 10., $N^{\underline{o}}$ 1, 1-11.

[4] ANDEL,J. (1989b), *On nonlinear models in time series*, Statistics 20 (4), 615-632.

[5] AN HONG-ZHI AND CHENG BING (1991), *A Kolmogorov-Smirnov type statistic with application to test for nonlinearity in time series*, Intern. Statist. Review 59, $N^{\underline{o}}$ 3, 287-307.

[6] BHASKARA RAO,M. ET AL (1983), *On the existence of some bilinear time series models*, J. Time Series Anal., Vol. 4, $N^{\underline{o}}$ 2, 95-110.

[7] BILLARD,L. AND MOHAMED,F.Y. (1991), *Estimation of the parameters of an* $EAR(p)$ *process*, J. Time Series Anal., Vol. 12, $N^{\underline{o}}$ 3, 179-192.

[8] CHAN,K.S. (1990), *Testing for threshold autoregression*, The Ann. of Statist. 18, $N^{\underline{o}}$ 4, 1886-1894.

[9] CHAN,K.S. AND TONG,H. (1990), *On likelihood ratio tests for threshold autoregressions*, J. R. Statist. Soc. B., Vol. 52, $N^{\underline{o}}$ 3, 469-476.

[10] CHEN,R. AND TSAY,R.S. (1991), *On the ergodicity of* $TAR(1)$ *processes*, Ann. Appl. Probab., Vol. 1, $N^{\underline{o}}$ 4, 613-634.

[11] DEWALD,L.S. AND LEWIS,P.A.W. (1985), *A new second-order autoregressive time series model -* $NLAR(2)$, IEEE Trans. Inf. Theory, Vol. IT-31, —no 5, 645-651.

[12] DOUKHAN,P. ET CHINDES,M. (1980), *Estimations dans le processus* "$X_{n+1} = f(X_n) + e_n$", C.R. Acad. Sc. Paris, T. 291, Ser. A, 61-64.

1) *The papers cited in this work and not cited in* [4] (*Andel* (1989*b*)).

[13] Gaver,D.P. and Lewis,P.A.W. (1980), *First-order autoregressive Gamma sequences and point processes*, Adv. Appl. Prob. 12, 727-745.

[14] Jevremović,V. (1991), *A note on mixed exponential distribution with negative weights*, Statistics and Prob. Lett. 11, 259-265.

[15] Jevremović,V. and Mališić,J. (1992), *On some moving average processes with exponentially distributed innovations*, (to appear).

[16] Karlsen,H. and Tjostheim,D. (1988), *Consistent estimates for the $NEAR$ (2) and $NLAR(2)$ time series models*, J. R. Statist. Soc. B, Vol. 50, $N^{\underline{o}}$ 2, 313-320.

[17] Kim,W.K. et al (1990), *Estimation for the first-order diagonal bilinear time series model*, J. Time Series Anal., Vol. 11, $N^{\underline{o}}$ 3, 215-229.

[18] Kumar,K. (1986), *On the identification of some bilinear time series models*, J. Time Series. Anal. , Vol. 7, $N^{\underline{o}}$ 2, 117-122.

[19] Lawrance,A.J. (1982), *The innovation distribution of a Gamma distributed autoregressive process*, Scand. J. Statist. 9, 234-236.

[20] Lawrance,A.J. and Lewis,P.A.W. (1981), *A new autoregressive time series model in exponential variables* ($NEAR(1)$, Adv. Appl. Prob. 13, 826-845.

[21] Lawrance,A.J. and Lewis,P.A.W. (1985), *Modelling and residual analysis of nonlinear autoregressive time series in exponential variables*, J.R. Statist. Soc. B, Vol. 47, $N^{\underline{o}}$ 2, 165-202.

[22] Li,W.K. (1984), *On the autocorrelation structure and identification of some bilinear time series*, J. Time Series Anal., Vol 5, $N^{\underline{o}}$ 2, 172-181.

[23] Liu,J. and Brockwell,P.J. (1988), *On the general bilinear time series model*, J. Appl. Prob. 25, 553-564.

[24] Liu,J. and Chen,Z.G. (1990), *Bilinear $(p, 0, 1, 1)$ model and its consistent identification*, Tech. Rep. 90, Dept. Stat., Univ. of British Columbia, Canada.

[25] Mališić,J. (1987), *On exponential autoregressive time series models*, Math. Statist. Probab. Theory, Vol. B, eds P.Bauer et al, 147-153.

[26] McKenzie,Ed (1985a), *Contribution to the discussion of "Modelling and residual analysis of nonlinear autoregressive time series in exponential variables" by A.J.Lawrance and P.A.W.Lewis*, J.R. Statist. Soc. B, Vol. 47, 187-188.

[27] McKenzie,Ed (1985b), *An autoregressive process for Beta random variables*, Manag. Sciences. Vol. 31, $N^{\underline{o}}$ 8, 988-997.

[28] McKenzie,Ed (1986), *Autoregressive moving average processes with negative binomial and geometric marginal distributions*, Adv. Appl. Prob. 18, 679-705.

[29] McKenzie,Ed (1987), *Innovation distributions for Gamma and negative binomial autoregressions*, Scand. J. Statist. 14, 1-7.

[30] Mokkadem,A. (1985), *Le modele non lineare $AR(1)$ general. Egrodicite et ergodicite geometrique*, C.R. Acad. Sci. Paris, T. 301, Ser. 1, $N^{\underline{o}}$ 19, 889-892.

[31] Petruccelli,J.D. (1986), *On the consistency of least squares estimators for a threshold $AR(1)$ model*, J. Time Series Anal., Vol. 7, $N^{\underline{o}}$ 4, 269-278.

[32] Popović,B.Č (1988), *Some predictions of the first order exponential time series*, Zbornik rad. Fil. fak. Niš, Ser. Mat. 2, 47-53.

[33] Priestley,M.B. (1980), *State-dependent models: general approach to non-linear time series analysis*, J. Time Series Anal., Vol. 1, $N^{\underline{o}}$ 1, 47-71.

[34] Saikkonen,P. and Luukkonen,R. (1988), *Legrange multiplier tests for testing non-linearities in the time series models*, Scand. J. Statist. 15, 55-68.

[35] Sim,C.H. (1990), *First-order autoregressive models for Gamma and exponential processes*, J. Appl. Prob. 27, 325-332.

[36] Smith,R.L. (1986), *Maximum likelihood estimation for the $NEAR(1)$ model*, J.R. Statist. Soc. A, Vol. 48, $N^{\underline{o}}$ 2, 251-257.

[37] Tong,H. and Lim,K.S. (1980), *Threshold autoregression, limitcycles and cyclical data*, J.R. Statist. Soc. B, Vol. 42, $N^{\underline{o}}$ 3, 245-292.

[38] Tukey,J.W. (1949), *One degree of freedom for non-additivity*, Biometrika 5, $N^{\underline{o}}$ 2, 232-242.

Faculty of Mathematics, University of Belgrade, Yu