Annals of Operations Research 9(1987)421 – 436

ON A CLASS OF THRESHOLD AR(k) PROCESSES

J.D. PETRUCCELLI and S.W. WOOLFORD

Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, USA

Abstract

We consider the model

$$Z_t = \sum_{i=1}^{k} \phi(i,j) Z_{t-i} + a_t(j) \text{ when } [Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k}]' \in R(j),$$

where $\{R(j); 1 \leqslant j \leqslant \ell\}$ is a partition of $\mathbb{R}^k$, and for each $1 \leqslant j \leqslant \ell, \{a_t(j); t \geqslant 0\}$ are i.i.d. zero-mean random variables, having a strictly positive density. Sufficient conditions are obtained for this process to be transient. In addition, for a particular class of such models, necessary and sufficient conditions for ergodicity are obtained. Least-squares estimators of the parameters are obtained and are, under mild regularity conditions, shown to be strongly consistent and asymptotically normal.

Keywords and phrases

Nonlinear time series, SETAR models.

1. Introduction

The framework for the study of self-exciting threshold autoregressive (SETAR) processes was established in Tong and Lim [8]. The usefulness of these nonlinear time series models was further discussed in Tong [7], while an application of a SETAR model to the Canadian lynx data was discussed in Lim and Tong [3]. In recent papers (see Petruccelli and Woolford [6] and Chan et al. [2]), necessary and sufficient conditions for the ergodicity of the SETAR(2; 1. 1) and the SETAR(k; 1, . . . , 1) were established. In particular, it was shown that these simplest of SETAR models were ergodic over a surprisingly broad set of parameter values.

This paper deals with a generalization of the SETAR(ℓ; $k, \ldots, k$) model as defined in Tong and Lim [8]. In particular, we suppose there is an integer ℓ and a partition $\{R(j)\}_{j=1}^{\ell}$ of $\mathbb{R}^k$, i.e.

$$R(j) \subset \mathbb{R}^k,\ 1 \leqslant j \leqslant \ell;\ R(i) \cap R(j) = \phi,\ 1 \leqslant i \neq j \leqslant \ell;\ \bigcup_{j=1}^{\ell} R(j) = \mathbb{R}^k,$$

such that our model is, for $t \geqslant 0$,

$$Z_t = \sum_{i=1}^{k} \phi(i,j) Z_{t-i} + a_t(j) \quad \text{if} \quad [Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k}]' \in R(j). \tag{1.1}$$

Here, $\{\phi(i, j);\ 1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant \ell\}$ are real constants, and for each $1 \leqslant j \leqslant \ell$, $\{a_t(j);\ t \geqslant 0\}$ is a sequence of independent, identically distributed (i.i.d.) random variables, each having a strictly positive density $f_j(\cdot)$ on $\mathbb{R}$ and zero mean. Additionally, we assume $\{a_t(i);\ t \geqslant 0\}$ is independent of $\{a_t(j);\ t \geqslant 0\}$, if $1 \leqslant i \neq j \leqslant \ell$.

It will be convenient to regard (1.1) as a vector-valued Markov chain with state space $\mathbb{R}^k$. Thus, we define

$$Z_t' = [Z_t, Z_{t-1}, \ldots, Z_{t-k+1}], \qquad t \geqslant 0$$

$$A_t(j)' = [a_t(j), 0, \ldots, 0], \qquad t \geqslant 0$$

$$\underset{k \times k}{\Phi}(j) = \left[\begin{array}{cccc|c} \phi(1,j) & \phi(2,j) & \ldots & \phi(k-1,j) & \phi(k,j) \\ \hline & & & & 0 \\ & & & & \vdots \\ & & \underset{(k-1) \times (k-1)}{I} & & 0 \\ & & & & \vdots \\ & & & & 0 \end{array}\right].$$

Then (1.1) is equivalent to the k-dimensional Markov chain

$$Z_t = \sum_{j=1}^{\ell} \left\{ \Phi(j) Z_{t-1} + A_t(j) \right\} I_{R(j)}(Z_{t-1}), \tag{1.2}$$

where $I_A(\cdot)$ is the indicator function of the set A.

We note here that the class of SETAR(ℓ; $k, \ldots, k$) models is obtained from models (1.1) and (1.2) by taking $R(j) = \mathbb{R}^{d-1} \times (r_{j-1}, r_j] \times \mathbb{R}^{k-d}$ for some $1 \leqslant d \leqslant k$, and $-\infty = r_0 < r_1 < \ldots < r_\ell = \infty$.

In sect. 2, we consider some general results for models (1.2) and establish the status sets for $\{Z_t\}$ (see Tweedie [9]). In sect. 3, we consider a particular class of models of the form (1.2) and give necessary and sufficient conditions on the matrices $\{\Phi(j); 1 \leqslant j \leqslant \ell\}$ for models in this class to be ergodic. A consequence of this work is a simple (and, we believe, new) proof of the conditions for ergodicity of an AR(k) process.

Under the assumptions of ergodicity of $\{Z_t\}$ and finiteness of $\sigma^2(j) \equiv E(a_t(j)^2)$, $1 \leqslant j \leqslant \ell$, we establish the strong consistency of the least-squares estimators for $\{(\phi(i,j); 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\}$ and $\{\sigma^2(j); 1 \leqslant j \leqslant \ell\}$ in sect. 4. In addition, a central limit theorem is shown to hold for the estimators of $\{(\phi(i,j); 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\}$. Finally, in sect. 5, we give some concluding remarks.

2. Preliminaries

The transition density for the process (1.2) is

$$p(x,y) = \prod_{m=1}^{k-1} I(y_m = x_{m+1}) \sum_{j=1}^{\ell} I_{R(j)}(x) f_j(y_k - \sum_{i=1}^{k} \phi(i,j) x_{k-i+1}), \quad (2.1)$$

where

$$y' = [y_k, y_{k-1}, \ldots, y_1], \quad x' = [x_k, x_{k-1}, \ldots, x_1].$$

Using the definitions of Orey [5], we also note that $\{Z_t, t \geqslant 0\}$ is μ-irreducible and aperiodic for μ taken to be the product Lebesque measure on $\mathbb{R}^k$. As in Chan et al. [2], the transition law $P(x, \cdot)$ corresponding to (2.1) is not in general strongly continuous (see Tweedie [9]). Moreover, lemma 2.1 in the paper by Chan et al. [2] can not be used to here establish that the relatively compact sets of positive Lebesque measure in $\mathbb{R}^k$ are status sets (Tweedie [9]), a result that will be used in proving the ergodicity conditions in sect. 3.

However, as the following lemma shows, we obtain the desired result for the relatively compact sets if we assume lower semicontinuity of the densities $\{f_j(\cdot); 1 \leqslant j \leqslant \ell\}$:

LEMMA 2.1

If $f_j(\cdot)$ is lower semicontinuous for each $1 \leqslant j \leqslant \ell$, then the relatively compact sets of positive measure in $\mathbb{R}^k$ are status sets for $\{Z_t\}$.

Proof

From (2.1) we have, for $A \in \mathcal{B}^k$ (the Borel sets in $\mathbb{R}^k$):

$$p(x, A) = \sum_{j=1}^{\ell} I_{R(j)}(x) \int_{A_1} f_j(u - \sum_{i=1}^{k} \phi(i,j) x_{k-i+1})\, \mathrm{d}u\,,$$

where $A_1 = \{y : [y, x_k, x_{k-1}, \ldots, x_2]' \in A\}$. Now

$$p(x, A) \geqslant \int_{A_1} \min_{1 \leqslant j \leqslant \ell} f_j(u - \sum_{i=1}^{k} \phi(i,j) x_{k-i+1})\, \mathrm{d}u \equiv \tilde{p}(x, A), \text{ say.}$$

By proposition 5.5 of Tweedie [9], the result will follow if

$$\int_{\mathbb{R}^k} \tilde{p}(x, \mathrm{d}y)\, g(y)$$

is lower semicontinuous in x, whenever $g(\cdot)$ is a bounded lower semicontinuous function.

Since $f_j(\cdot)$ is lower semicontinuous,

$$f_j(u - \sum_{i=1}^{k} \phi(i,j) x_{k-i+1})$$

and hence

$$\min_{1 \leqslant j \leqslant \ell} f_j(u - \sum_{i=1}^{k} \phi(i,j) x_{k-i+1})$$

is lower semicontinuous in $x = [x_k, \ldots, x_1]$. Fatou's lemma then implies, for any $A \in \mathcal{B}^k$,

$$\varliminf_{x \to x_0} \tilde{p}(x, A) \geqslant \tilde{p}(x_0, A)\,,$$

which shows $\tilde{p}(\cdot\,, A)$ to be lower semicontinuous. For $g(\cdot)$ bounded and lower semicontinuous on $\mathbb{R}^k$, take $g_n(\cdot)$ to be an increasing sequence of simple functions on $\mathbb{R}^k$ which converge a.e. to $g(\cdot)$. Then

$$\varliminf_{x \to x_0} \int_{\mathbb{R}^k} \widetilde{p}(x, y) g(y)\, \mathrm{d}y \geqslant \varliminf_{x \to x_0} \int_{\mathbb{R}^k} \widetilde{p}(x, y) g_n(y)\, \mathrm{d}y$$

$$\geqslant \int_{\mathbb{R}^k} \widetilde{p}(x_0, y) g_n(y)\, \mathrm{d}y. \tag{2.2}$$

Since (2.2) holds for all n, the result follows. □

In view of lemma 2.1, we shall assume throughout the rest of this paper that $f_j(\cdot)$ is lower semicontinuous for each $1 \leqslant j \leqslant \ell$.

Using lemma 2.1 above and theorem 9.1(i) of Tweedie [9] with $g(x) = \|x\| = (\langle x, x\rangle)^{1/2}$ (where $\langle \cdot, \cdot \rangle$ is the usual inner product on the complex vector space $\mathbb{C}^k$), one obtains that $\{Z_t\}$ is ergodic if the maximum eigenvalue of $\Phi(j)' \Phi(j)$ is less than 1, for $1 \leqslant j \leqslant \ell$. In fact, Tong and Lim [8] state this sufficient condition. However, this condition is vacuous since it is easy to show that the largest eigenvalue of $\Phi(j)' \Phi(j)$ is at least 1. This fact is alluded to in Tong [7].

General conditions which ensure the ergodicity of $\{Z_t\}$ for model (1.2) seem to be difficult to obtain. However, in the next theorem we establish sufficient conditions for $\{Z_t\}$ to be transient.

THEOREM 2.1

If the smallest eigenvalue of $\Phi(j)' \Phi(j)$ exceeds 1 for all $1 \leqslant j \leqslant \ell$, then $\{Z_t\}$ is transient.

Proof

We first note that $\lambda(j)$, the smallest eigenvalue of $\Phi(j)' \Phi(j)$, satisfies

$$\lambda(j) = \min_{\|x\| = 1} \langle \Phi(j)x, \Phi(j)x \rangle = \min_{\|x\| = 1} \|\Phi(j)x\|^2 .$$

For $Z_{t-1} \in R(j)$,

$$\|\Phi(j)Z_{t-1}\| - \|A_t(j)\| \leqslant \|Z_t\| \leqslant \|\Phi(j)Z_{t-1}\| + \|A_t(j)\| ,$$

which implies that

$$\|\Phi(j)Z_{t-1}\| - E(\|A_t(j)\|) \leqslant E(\|Z_t\| \,|Z_{t-1})$$

$$\leqslant \|\Phi(j)Z_{t-1}\| + E(\|A_t(j)\|).$$

Hence

$$E(|\,\|Z_t\| - E(\|Z_t\|\,|Z_{t-1})\|Z_{t-1}) \leqslant 2E(\|A_t(j)\|) < K < \infty,$$

for some constant K independent of j. Now let $1 < \eta < \min_{1\leqslant j\leqslant \ell}\lambda(j)^{1/2}$. Then there is an $M_1 > 0$ such that $\|Z_{t-1}\| > M_1$ implies

$$E(\|Z_t\|\,|Z_{t-1}) > \eta\|Z_{t-1}\|.$$

But by Markov's inequality

$$P(\|Z_t\| < 2^{-1}(\eta+1)\|Z_{t-1}\|\,|Z_{t-1}) \leqslant 2K/(\eta-1)\|Z_{t-1}\|.$$

From this point, the proof follows that of lemma 2.2 of Petruccelli and Woolford [6], to show that $P(\|Z_t\| \to \infty | Z_0) > 0$. □

3. A specific class of models

The condition for models to belong to the class of models we are considering in this section is that all but one of the $R(j)$, $1 \leqslant j \leqslant \ell$, be a bounded subset of $\mathbb{R}^k$. Specifically, we will assume throughout this section that there is an $M > 0$ such that

$$\bigcup_{j=1}^{\ell-1} R(j) \subset \{x \in \mathbb{R}^k : \|x\| \leqslant M\}.$$

Among models in this class are models having $R(j) = \{x \in \mathbb{R}^k : r_{j-1} \leqslant \|x\| < r_j\}$, for $1 \leqslant j \leqslant \ell$, where $0 = r_0 < r_1 < \ldots < r_\ell = \infty$. These models are natural extensions of SETAR(ℓ; 1, . . . , 1) models, considered in Chan et al. [2].

For models in this class, we obtain necessary and sufficient conditions for ergodicity in the following theorem.

THEOREM 3.1

The process $\{Z_t\}$, defined by (1.2) with $R(j)$ a bounded subset of $\mathbb{R}^k$, $1 \leqslant j \leqslant \ell - 1$, is ergodic if and only if

$$\rho(\Phi(\ell)) < 1,$$

where $\rho(A)$ is the spectral radius of A.

The proof of theorem 3.1 is divided into the following three lemmas:

LEMMA 3.1

If $\rho(\Phi(\ell)) < 1$, then Z_t is ergodic.

Proof

We begin by noting that to each distinct eigenvalue $\lambda(i)$ of $\Phi(\ell)$ there corresponds a unique (up to scalar multiplicity) eigenvector $w(i, 1)$, $1 \leqslant i \leqslant m$. Here, m is the number of distinct eigenvalues of $\Phi(\ell)$. It is easy to show that for any $c > 0$ there is a matrix $W(c)$ such that

$$W(c)^{-1}\,\Phi(\ell)'\,W(c) = J(c) = \begin{bmatrix} J_1(c) & & & 0 \\ & J_2(c) & & \\ & & \ddots & \\ 0 & & & J_m(c) \end{bmatrix} \tag{3.1}$$

where $J_i(c)$ is the same as the Jordan block corresponding to $\lambda(i)$ except for values of c (instead of 1) on the superdiagonal.

Equation (3.1) yields a system of equations analogous to those which define the generalized eigenvectors of $\Phi(\ell)'$ (see Noble [4], p.364). In particular, the columns of $W(c)$ are independent vectors defined by

$$(\Phi(\ell)' - \lambda(i)I)\,w(i,j) = c\,w(i,j-1), \qquad j = 2, \ldots, \alpha_i,$$

where α_i is the algebraic multiplicity of $\lambda(i)$.

We now define

$$g(x) = \sum_{i=1}^{m} \sum_{j=1}^{\alpha_i} |\langle x, w(i,j)\rangle|, \qquad x \in \mathbb{R}^k$$

and take

$$A = \{x \in \mathbb{R}^k : \|x\| \leqslant M_1\}$$

for some $M_1 > M$ to be determined. Then, for $x \notin A$

$$L(x) = E(g(Z_t)|Z_{t-1} = x) \leqslant K\,E(|a_1(\ell)|) + \max_{1 \leqslant i \leqslant m} [|\lambda(i)| + c]\,g(x)$$

for some $0 < K < \infty$. However, since $\rho(\Phi(\ell)) < 1$, we can choose $c > 0$ small enough so that $\max_{1 \leq i \leq m} [|\lambda(i)| + c] < 1$. Hence, for all $\epsilon > 0$, there is an $M_1 > M$ such that

$$L(x) \leqslant g(x) - \epsilon, \quad x \notin A.$$

For $x \in A$, we see that $L(x) \leqslant B < \infty$. Hence, the ergodicity of $\{Z_t\}$ follows from theorem 9.1(i) of Tweedie [9]. □

LEMMA 3.2

If $\rho(\Phi(\ell)) > 1$, then $\{Z_t\}$ is transient.

Proof

Suppose $\Phi(\ell)'w = \lambda w$ for some vector $w \in \mathbb{C}^k$ with $\|w\| = 1$ and some scalar $\lambda \in \mathbb{C}$ with $|\lambda| > 1$. Define the Markov chain $\{V_t; t \geqslant 0\}$ on $[0, \infty)$, by

$$V_t = |\langle Z_t, w\rangle|.$$

Suppose $V_{t-1} > M$. This implies $Z_{t-1} \in R(\ell)$ and so

$$E(V_t | V_{t-1}) = \int_{\mathbb{R}} |\langle u + \Phi(\ell) Z_{t-1}, w\rangle| f_\ell(u)\, du$$

$$\geqslant |\lambda| V_{t-1} - |w_1| E(|a_1(\ell)|),$$

where $u' = [u, 0, \ldots, 0]$ and w_1 is the first component of w. Hence, for $V_{t-1} > M_1 > M$,

$$E(V_t | V_{t-1}) > \eta V_{t-1}$$

for some $1 < \eta < |\lambda|$. Also, for $V_{t-1} > M$

$$E(|V_t - E(V_t | V_{t-1})| \, V_{t-1}) \leqslant 2E(|\langle A_t(\ell), w\rangle|) < B < \infty,$$

independent of t. Thus, following the argument of lemma 2.2 of Petruccelli and Woolford [6], we obtain that

$$P(V_t \to \infty | V_0) > 0.$$

This, in turn, implies that

$$P(\|Z_t\| \to \infty | Z_0) > 0. \qquad \square$$

LEMMA 3.3

If $\rho(\Phi(\ell)) = 1$, then $\{Z_t\}$ is not ergodic.

Proof

Let λ and w be as in the proof of the previous lemma, except that here $|\lambda| = 1$. Consider the process

$$U_t = \langle Z_t, w \rangle .$$

Then, if $|U_{t-1}| > M$,

$$U_t = \bar{\lambda} U_{t-1} + \langle A_t(\ell), w \rangle , \tag{3.2}$$

where $\bar{\lambda}$ is the complex conjugate of λ. By considering the real and imaginary parts of U_t, we can define the Markov chain $\{Y_t\}$ on $\mathbb{R}^2$ by

$$Y_t' = [\mathrm{Re}\,(U_t),\ \mathrm{Im}(U_t)] .$$

Then (3.2) may be written, for $\|Y_{t-1}\| > M$, as

$$Y_t = \Delta Y_{t-1} + \epsilon_t ,$$

where, for some θ, Δ is the orthogonal matrix

$$\Delta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} ,$$

and $E(\epsilon_t) = \mathbf{0}$.

The result will follow from theorem 9.1(ii) of Tweedie [9] by defining $g(x) = \|x\|$, for $x \in \mathbb{R}^2$, and $A = \{x \in \mathbb{R}^2 : \|x\| \leqslant M\}$. Then, by Jensen's inequality,

$$E(g(Y_t)| Y_{t-1} = x) = E(g(\Delta x + \epsilon_t)) \geqslant g(E(\Delta x + \epsilon_t)) = g(x)$$

for $x \notin A$. Since Δ is orthogonal

$$E(|g(Y_t) - g(x)\| Y_{t-1} = x) \leqslant E(\|\epsilon_t\|) \leqslant \beta < \infty$$

for $x \notin A$, whereas for $x \in A$ it is clear that

$$E(|g(Y_t) - g(x)\| Y_{t-1} = x) \leqslant \beta < \infty .$$

Finally, $g(x) > \sup_{y \in A} g(y)$ for $x \notin A$ and so Tweedie's conditions are satisfied and we conclude that $\{Y_t\}$ is not ergodic. From this it follows that $\{Z_t\}$ is not ergodic. □

Remarks

(1) It is interesting to note that the condition that $\rho(\Phi(\ell)) < 1$ is equivalent to requiring that the roots of the backshift polynomial for (1.1) when $Z_{t-1} \in R(\ell)$,

$$1 - \sum_{i=1}^{k} \phi(i, \ell)B^i = 0,$$

lie outside the unit circle. The classical AR(k) process is obtained from model (1.1) by taking $\ell = 1$. In this case, theorem 3.1 gives what we believe to be a new proof of the ergodicity conditions for AR(k) processes.

(2) In lemma 3.2 we have proved something stronger than what is stated in theorem 3.1: that not only is $\{Z_t\}$ not ergodic when $\rho(\Phi(\ell)) > 1$, but that it is in fact transient. The reader should be aware that while in lemma 3.3 we have proved $\{Z_t\}$ to be what Tweedie [9] calls "null" (by which he means "not ergodic"), we have not proved it "null recurrent" (by which we mean "neither ergodic nor transient"). However, we conjecture the latter to be the case.

THEOREM 3.2

Assume $E(|a_t(i)|^k) < \infty$, $1 \leqslant i \leqslant \ell$, and some integer $k > 0$. Then, if $\rho(\Phi(\ell)) < 1$, the invariant probability distribution for the chain $\{Z_t\}$ has a finite kth moment and the model is geometrically ergodic.

Proof

For $x \in \mathbb{R}^k$, let $g(x)$ be defined as in lemma 3.1. Then, for x large and some $0 < K < \infty$,

$$E(g(Z_t)|Z_{t-1} = x) \leqslant KE(|a_1(\ell)|) + \max_{1 \leqslant i \leqslant m} [|\lambda(i)| + c]\, g(x).$$

Since $\rho(\Phi(\ell)) < 1$, we can choose $c > 0$ so that

$$\xi \equiv \max_{1 \leqslant i \leqslant m} [|\lambda(i)| + c] < 1.$$

But $g(x) \to \infty$ as $\|x\| \to \infty$, so there is an $M > 0$ such that for $\|x\| > M$

$$E(g(Z_t)|Z_{t-1} = x) \leqslant (1 - \epsilon) g(x)$$

for some $1 > \epsilon > 0$. The result then follows from Tweedie [10]. □

4. Estimation of model parameters

Throughout this section, we make the following assumptions.

A1: $\{Z_t\}$ is ergodic and its stationary distribution has a finite second moment, and

A2: $\sigma^2(j) \equiv E(a_t(j)^2) < \infty$, $1 \leqslant j \leqslant \ell$.

In order to facilitate the following discussion, we reformulate our model. Let

$$Z_m(i,j) \equiv Z_m I(Z_{j-1} \in R(i)) \equiv Z_m I(i,j)$$

for $1 \leqslant i \leqslant \ell$, $m \geqslant 0$ and $1 \leqslant j - m \leqslant k$. Then we can rewrite our model, for each $n \geqslant 0$, as

$$Y = X\phi + A, \tag{4.1}$$

where

$$Y' = [Z_1, \ldots, Z_n], \tag{4.2}$$

$$\underset{n \times k\ell}{X} = \begin{bmatrix} Z_0(1,1), \ldots, Z_{-k+1}(1,1), \ldots, Z_0(\ell,1), \ldots, Z_{-k+1}(\ell,1) \\ \vdots \\ Z_{n-1}(1,n), \ldots, Z_{n-k}(1,n), \ldots, Z_{n-1}(\ell,n), \ldots, Z_{n-k}(\ell,n) \end{bmatrix} \tag{4.3}$$

$$\phi' = [\phi(1,1), \ldots, \phi(k,1), \ldots, \phi(1,\ell), \ldots, \phi(k,\ell)] \tag{4.4}$$

and

$$A' = \left[\sum_{i=1}^{\ell} a_1(i) I(i,1), \ldots, \sum_{i=1}^{\ell} a_n(i) I(i,n) \right]. \tag{4.5}$$

We further define $J(i) = \{1 \leqslant t \leqslant n: Z_{t-1} \in R(i)\}$, $1 \leqslant i \leqslant \ell$, and let $n(i)$ be the cardinality of $J(i)$. Then, we note that $E(A) = \mathbf{0}$ and

$$E(A'A) = \sum_{i=1}^{\ell} \sigma^2(i) E(n(i)).$$

Assuming that the partition $\{R(\cdot)\}$ is known, the least-squares estimators for the parameters ϕ are given by

$$\hat{\phi} = (X'X)^{-1} X'y .$$

We note that $X'X = \operatorname{diag}(B_1, \ldots, B_\ell)$, where

$$B_i = \begin{bmatrix} \sum_{m=1}^{n} Z_{m-1}^2(i,m), \ldots & \sum_{m=1}^{n} Z_{m-1}(i,m) Z_{m-k}(i,m) \\ \vdots & \\ \sum_{m=1}^{n} Z_{m-1}(i,m) Z_{m-k}(i,m), \ldots & \sum_{m=1}^{n} Z_{m-k}^2(i,m) \end{bmatrix} . \tag{4.7}$$

Under the assumption of ergodicity of $\{Z_t\}$, we define $Z' = [Z(1), \ldots, Z(k)]$ to be a random vector having the same distribution as the stationary distribution of $\{Z_t\}$. Then, as $n \to \infty$,

$$n(i)/n \to \pi(i) \equiv P(Z \in R(i)) \text{ a.s.}, \qquad 1 \leqslant i \leqslant \ell,$$

and

$$n^{-1} \sum_{p=1}^{n} Z_{p-m}(i,p) Z_{p-j}(i,p) \to E[Z(m)Z(j)I(Z \in R(i))]$$

$$\equiv \rho(m,j,i) \text{ a.s.}, \qquad 1 \leqslant i \leqslant \ell,\ 0 \leqslant m-j \leqslant k.$$

Hence, as $n \to \infty$,

$$n^{-1}B_i \to H_i \equiv \{\rho(m,j,i),\ 1 \leqslant m \leqslant k, 1 \leqslant j \leqslant k\} \text{ a.s.},\ 1 \leqslant i \leqslant \ell. \qquad (4.8)$$

However, under our assumptions, the matrix $H = \text{diag}\,(H_1, \ldots, H_\ell)$ is positive definite. Thus, for large n, $X'X$ is invertible and (4.6) holds.

The next two theorems establish the strong consistency and asymptotic normality of the estimators (4.6).

THEOREM 4.1

Under assumptions A1 and A2, as $n \to \infty$,

$$\hat{\phi} \to \phi \text{ a.s.}$$

and

$$\hat{\sigma}^2(j) \equiv n(j)^{-1} \sum_{t \in J(j)} (Z_t - \sum_{m=1}^{k} \hat{\phi}(m,j) Z_{t-m})^2 \to \sigma^2(j) \text{ a.s.} \qquad (4.9)$$

for $1 \leqslant j \leqslant \ell$.

Proof

Using (4.1) and (4.6), we obtain

$$\hat{\phi} = \phi + \left(\frac{X'X}{n}\right)^{-1} \frac{X'A}{n}\ .$$

But $X'A/n$ is a vector whose terms are of the form

$$n^{-1} \sum_{m=1}^{n} Z_{m-p}(j,m) \sum_{i=1}^{\ell} a_m(i) I(i,m), \qquad 1 \leqslant p \leqslant k,\ 1 \leqslant j \leqslant \ell.$$

Hence, as $n \to \infty$, $X'A/n \to 0$ a.s. and $(X'X/n)^{-1} \to H^{-1}$ a.s. and so $\hat{\phi} \to \phi$ a.s.

Expanding (4.9) yields

$$\hat{\sigma}^2(j) = n(j)^{-1}\Bigg[\sum_{t\in J(j)}\Bigg\{a_t(j)^2 + \sum_{m=1}^{k}(\phi(m,j) - \hat{\phi}(m,j))^2 Z_{t-m}^2$$

$$+ 2\sum_{m=1}^{k} a_t(j)(\phi(m,j) - \hat{\phi}(m,j))Z_{t-m}$$

$$+ \sum_{m=1}^{k}\sum_{n=1}^{k}(\phi(m,j) - \hat{\phi}(m,j))(\phi(n,j) - \hat{\phi}(n,j))Z_{t-m}Z_{t-n}\Bigg\}\Bigg].$$

But $n(j)^{-1}\ \Sigma_{t\in J(j)}a_t(j)^2 \to \sigma^2(j)$ a.s., as $n \to \infty$, and all other terms converge almost surely to zero. Hence, the result follows. □

Remark

We note that the above result implies that

$$n^{-1}A'A \to \sum_{i=1}^{\ell} \pi(i)\,\sigma^2(i)\ \text{a.s., as } n \to \infty.$$

THEOREM 4.2

Under assumptions A1 and A2, as $n \to \infty$,

$$\sqrt{n}\,\Delta' \Sigma^{-1} (\hat{\phi} - \phi) \xrightarrow{\mathcal{L}} N_{k\ell}(\mathbf{0}, I_{k\ell}),$$

where $H = \Delta'\Delta$, I_n is an $n \times n$ identity matrix, $\Sigma^{-1} = \text{diag}\,(\Sigma(1)^{-1}, \ldots, \Sigma(\ell)^{-1})$, with $\Sigma(j) = \sigma(j)I_k$ and $\xrightarrow{\mathcal{L}}$ is convergence in law.

Proof

Using (4.6), we obtain, for $\lambda \in \mathbb{R}^{k\ell}$,

$$\sqrt{n}\,\lambda' \left(\frac{X'X}{n}\right)(\hat{\phi} - \phi) = n^{-1/2}X'A = n^{-1/2}\sum_{s=1}^{n} \xi_s,$$

where

$$\xi_s = \sum_{j=1}^{\ell}\sum_{m=1}^{k} \lambda_{k(j-1)+m} Z_{s-m}(j,s)\,a_s(j).$$

However, since $E(\xi_s \mid \xi_{s-1}, \ldots, \xi_1) = 0$ and

$$\lim_{n \to \infty} E(\xi_s^2) = \lambda' \sum H \sum \lambda,$$

$\{\xi_s\}$ is a martingale sequence. Hence, by theorem 23.1 in Billingsley [1],

$$n^{-1/2} \sum_{s=1}^{n} \xi_s \xrightarrow{\mathcal{L}} N(\mathbf{0}, \lambda' \sum H \sum \lambda).$$

But H is positive definite and so there is a matrix Δ such that $H = \Delta\Delta'$. Consequently,

$$n^{-1/2} \lambda' \Delta^{-1} \sum^{-1} H(\hat{\phi} - \phi) \xrightarrow{\mathcal{L}} N_{k\ell}(\mathbf{0}, \lambda'\lambda).$$

Since $\Delta^{-1} \Sigma^{-1} H = \Delta^{-1} H \Sigma^{-1} = \Delta' \Sigma^{-1}$, the result follows. □

5. Concluding remarks

We note that we can define a slightly more general model than that defined by (1.2). In particular, we can define

$$Z_t = \sum_{j=1}^{\ell} \{c(j) + \Phi(j) Z_{t-1} + A_t(j)\} I_{R(j)}(Z_{t-1}), \tag{5.1}$$

where, for $1 \leqslant j \leqslant \ell$, $c(j)' = [\phi(0,j), 0, \ldots, 0]$, with $\phi(0,j)$ a constant, and all other quantities are as defined in sect. 1. This model is the natural extension of the AR(1) model studied by Chan et al. [2]. In addition, the SETAR($\ell; k, \ldots, k$) process by Tong and Lim [8] is a special case of (5.1).

Under the assumption that $\{R(j)\}$ is an arbitrary partition of $\mathbb{R}^k$, it is not difficult to show that theorem 2.1 still holds for $\{Z_t\}$ as defined by 5.1). If we make the additional assumption that all but one of the $R(j)$ be a bounded subset of $\mathbb{R}^k$, as in sect. 3, then it is easy to show that lemma 3.1 and lemma 3.2 also hold for $\{Z_t\}$ as defined by (5.1). It is conjectured that lemma 3.3 also holds in this case, but we have been unable to obtain a proof of this.

References

[1] P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968).
[2] K.S. Chan, J.D. Petruccelli, H. Tong and S.W. Woolford, A multiple-threshold AR(1) model, J. Appl. Prob. 22(1985)267.
[3] K.S. Lim and H. Tong, A statistical approach to difference-delay equation modelling in ecology – two case studies, Journal of Time Series Analysis 4(1983)239.
[4] B. Noble, *Applied Linear Algebra* (Prentice-Hall, Englewood Cliff, New Jersey, 1969).
[5] S. Orey, *Limit Theorems for Markov Chain Transition Probabilities* (Van Nostrand Reinhold, New York, 1971).
[6] J.D. Petruccelli and S.W. Woolford, A threshold AR(1) model, J. Appl. Prob. 21(1984)270.
[7] H. Tong, *Threshold Models in Nonlinear Time Series Analysis*, Lecture Notes in Statistics (Springer-Verlag, New York, 1983).
[8] H. Tong and K.S. Lim, Threshold autoregression, limit cycles and cyclical data, J. Roy. Statist. Soc. B42(1980)245.
[9] R.L. Tweedie, Criteria for classifying general Markov chains, Adv. Appl. Prob. 8(1976)737.
[10] R.L. Tweedie, The existence of moments for stationary Markov chains, J. Appl. Prob. 20 (1983)191.