

ON THE APPROXIMATION OF TIME SERIES BY THRESHOLD AUTOREGRESSIVE MODELS

By JOSEPH D. PETRUCCELLI

Worcester Polytechnic Institute

SUMMARY. It is shown that threshold autoregressive models can approximate a general class of time series processes almost surely. This class includes exponential autoregressive and invertible bilinear processes.

1. INTRODUCTION

In recent years a number of nonlinear alternatives to the classical linear time series models have been developed. The most notable among these are the bilinear (Granger and Andersen, 1978 ; Subba Rao, 1981), the threshold autoregressive (TAR) (Tong and Lim, 1980 ; Tong, 1983), and the exponential autoregressive (Ozaki, 1981 ; Haggen and Ozaki, 1981) classes of models. In addition to these specific classes, Priestley (1980) has introduced a generalized class of nonlinear models, the state dependent models, (SDMs) which includes each of the above classes as a special case.

In this note we show that threshold autoregressive models can almost surely approximate a general class of time series processes, including both exponential autoregressive and invertible bilinear processes.

While Haggen *et al.* (1984) give a passing reference to TAR models as approximations to SDMs, and Tong (1983) provides an heuristic justification of TAR models as approximations to more general models, we feel the question of such approximations has not been adequately addressed.

The importance of knowing that TAR models can approximate a wide class of processes lies in their tractability and interpretability. In addition, by knowing the ranges of possible behaviour of TAR processes, we can infer the behaviour of the processes they approximate.

These results also help to explain why tests designed to detect TAR-type nonlinearity (see, eg. Chan and Tong, 1987 ; Petruccielli and Davies, 1986) often perform well in detecting other types of nonlinearity such as exponential autoregressive or bilinear.

Paper received. December 1988 ; revised January 1991.

AMS (1985) subject classification. 62M10.

Key words and phrases. Threshold models, approximation.

The approximation theorems are stated and proved in Section 3. In Section 2 a brief review of the main classes of nonlinear models is given. In Section 4 TAR approximations to specific classes of nonlinear models are considered.

2. NONLINEAR MODELS

The following are the nonlinear models we will consider. In all cases $\{a_t, t \geq 0\}$ is an i.i.d white noise process, and $\{Z_t\}$ are observed values of the process.

(i) *Bilinear models* (Granger and Andersen, 1978 ; Subba Rao, 1981).

$$Z_t = \mu + \sum_{i=1}^p \varphi_i Z_{t-i} + \sum_{j=1}^q \theta_j a_{t-j} + \sum_{k=1}^q \sum_{l=1}^p \beta_{kl} a_{t-k} Z_{t-l} + a_t,$$

where $\mu, \{\varphi_i\}, \{\theta_j\}, \{\beta_{kl}\}$ are constants.

(ii) *Threshold AR (TAR) models* (Tong and Lim, 1980 ; Tong, 1983).

$$Z_t = \mu^{(i)} + \sum_{j=1}^p \varphi_j^{(i)} Z_{t-j} + a_t,$$

$$(Z_{t-1}, \dots, Z_{t-p}) \in R^{(i)}, i = 1, \dots, r$$

where $\{R^{(i)}, 1 \leq i \leq r\}$ is a partition of \mathbb{R}^p , and $\{\mu^{(i)}\}, \{\varphi_j^{(i)}\}$ are constants.

(iii) *Exponential AR (EAS) models* (Ozaki, 1981 ; Haggan and Ozaki, 1981).

$$Z_t = \sum_{i=1}^p (\theta_i + \pi_i e^{-\gamma Z_{t-1}^2}) Z_{t-i} + a_t$$

where $\gamma, \{\theta_i\}, \{\pi_i\}$ are constants.

(iv) *State dependent models* (SDMs) (Priestley 1980, 1981 ; Haggan et al. 1984).

$$Z_t = \mu(Z_{t-1}) + \sum_{i=1}^p \varphi_i(Z_{t-1}) Z_{t-i} + \sum_{j=1}^q \theta_j(Z_{t-1}) a_{t-j} + a_t$$

where $Z_{t-1} = (Z_{t-1} \dots Z_{t-p}, a_{t-1}, \dots, a_{t-q})$ is the state vector of the process at time $t-1$, and $\mu(\cdot), \{\varphi_i(\cdot)\}, \{\theta_j(\cdot)\}$ are suitably smooth (usually analytic) real-valued functions.

By defining the functions $\mu(\cdot), \{\varphi_i(\cdot)\}$ and $\{\theta_j(\cdot)\}$ appropriately models (i)–(iii) are seen to be special cases of the SDM (see Priestley, 1980).

3. APPROXIMATION RESULTS

Suppose, to begin with, that $f: R^k \rightarrow R^k$ is a continuous function and consider the k -dimensional process

$$Z_t = f(Z_{t-1}) + A_t, \quad t \geq 1 \quad \dots (3.1)$$

where $\{A_t\}$ are i.i.d. zero mean k -dimensional random vectors. Univariate processes of the form

$$Z_t = g(Z_{t-1}, \dots, Z_{t-k}) + a_t, \quad \dots (3.2)$$

where $g: R^k \rightarrow R$ and $\{a_t\}$ are i.i.d. zero mean random variables, may be written in the form (3.1) by using the state space representation

$$\begin{aligned} Z'_t &= (Z_t, \dots, Z_{t-k+1}) \\ [f(Z_{t-1})]' &= (g(Z_{t-1}, \dots, Z_{t-k}), Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}) \\ A'_t &= (a_t, 0, \dots, 0). \end{aligned}$$

Let $f_n: R^k \rightarrow R^k$, $n = 1, 2, \dots$, be a sequence of functions approximating f uniformly on any compact set in R^k . That is for each ϵ , $L > 0$, there is an $M = M(\epsilon, L)$ such that $n > M$ implies $\|f_n(x) - f(x)\| < \epsilon$ whenever $\|x\| \leq L$. Here $\|\cdot\|$ is the usual norm on R^k . Assume (Ω, \mathcal{B}, P) is the underlying probability space throughout. We then have

Theorem 3.1. *Let $Z_0 \in R^k$ be the starting point of the process defined by (3.1) and consider the process $\{Z_{t,n}, t \geq 0\}$ defined by $Z_{0,n} = Z_0$, $Z_{t,n} = f_n(Z_{t-1,n}) + A_t$, $t \geq 1$, where A_t is the white noise process driving (3.1) and f_n is as above. Then for each $\omega \in \Omega$, $Z_{t,n}(\omega) \rightarrow Z_t(\omega)$ as $n \rightarrow \infty$, $t = 1, 2, \dots$.*

Proof. For each $\omega \in \Omega$, $Z_{1,n}(\omega) = f_n(Z_0(\omega)) + A_1(\omega) \xrightarrow{n \rightarrow \infty} f(Z_0(\omega)) + A_1(\omega) = Z_1(\omega)$.

Suppose $Z_{k,n} \xrightarrow[n \rightarrow \infty]{a.s.} Z_k$, $k = 1, 2, \dots, t-1$. Then, for every $\omega \in \Omega$,

$$\begin{aligned} &\|Z_{t,n}(\omega) - Z_t(\omega)\| \\ &= \|f_n(Z_{t-1,n}(\omega)) + A_t(\omega) - (f(Z_{t-1}(\omega)) + A_t(\omega))\| \\ &= \|f_n(Z_{t-1,n}(\omega)) - f(Z_{t-1,n}(\omega))\| \\ &\quad + \|f(Z_{t-1,n}(\omega)) - f(Z_{t-1}(\omega))\|. \quad \dots (3.3) \end{aligned}$$

By the continuity of f , there is, for each $\epsilon > 0$, a $\delta > 0$ such that $\|x - Z_{t-1}(\omega)\| < \delta$ implies $\|f(x) - f(Z_{t-1}(\omega))\| < \epsilon/2$.

By assumption there is an $N = N(\omega, \delta)$ such that $n > N$ implies $\|Z_{t-1,n}(\omega) - Z_{t-1}(\omega)\| < \delta$.

Let $L = \|Z_{t-1}(\omega)\| + \delta$. Then for $M(\epsilon/2, L)$ as defined above and for $n > \max\{N, M(\epsilon/2, L)\}$ we have

$$\|f_n(Z_{t-1, n}(\omega)) - f(Z_{t-1, n}(\omega))\| < \epsilon/2.$$

Thus (3.3) is bounded above by ϵ , proving the theorem.

Remark. In particular, we may choose piecewise linear f_n to satisfy the conditions of Theorem 3.1. Thus TAR models may be used to approximate any process of the form (3.1) when f is continuous.

In order to extend the result of Theorem 3.1 to processes defined by arbitrary measurable f we restrict consideration to univariate models of the form (3.2) and to approximation by TAR processes.

Lemma 3.1. *Consider the $\{Z_t, t \geq 0\}$ process defined by (3.2) with g measurable and bounded, with $\{a_t\}$ having a density on \mathbf{R} and with fixed starting values Z_{1-k}, \dots, Z_0 . Then for each $\epsilon > 0$ there is a sequence of TAR processes $\{Z_{t, n}, t \geq 0\}$, $n = 1, 2, \dots$, such that $Z_{t, n} \rightarrow Z_t^*$ almost surely as $n \rightarrow \infty$, $t = 1, 2, \dots$, where $P(Z_l^* \neq Z_l, \text{ some } 1-k \leq l \leq t) < \epsilon t$.*

Proof. Suppose that $|g|$ is bounded by M and choose $R = R(Z_0, \epsilon)$ such that $P(|a_t| > R/\sqrt{k} - M) < \epsilon/2k$ and $\|Z_0\| < R$. It can be shown that for $r \geq 1$, the r -step transition probabilities for the process, given Z_0 , are absolutely continuous with respect to Lebesgue measure on \mathbf{R}^k . Thus there is a $\delta > 0$ such that for any set $A \subset \mathbf{R}^k$ of Lebesgue measure less than δ ,

$$P(Z_{r-1} \in A | Z_0) < \epsilon/2, r = 2, \dots, t.$$

By Lusin's Theorem (see, e.g. Rubin, 1966) we may find a function g_ϵ continuous on $C(R) = \{x : \|x\| \leq R\}$ such that $|g_\epsilon| \leq M$ on $C(R)$ and such that $g_\epsilon = g$ on $C(R)$ except on a set $B(R, \epsilon)$ of Lebesgue measure less than δ . We may, and do, choose g_ϵ such that $g_\epsilon(Z_0) = g(Z_0)$ and such that $g_\epsilon \equiv 0$ on $C^c(R)$, the complement of $C(R)$. Let $\{g_{\epsilon, n}, n \geq 1\}$ be a sequence of piecewise linear functions on \mathbf{R}^k converging uniformly to g_ϵ . Define the processes

$$Z_t^* = g_\epsilon(Z_{t-1}^*, \dots, Z_{t-k}^*) + a_t$$

$$Z_{t, n}^* = g_{\epsilon, n}(Z_{t-1, n}^*, \dots, Z_{t-k, n}^*) + a_t, t \geq 0$$

with $Z_{j, n}^* = Z_j^* = Z_j$, $j = -k+1, -k+2, \dots, 0$.

By Theorem 3.1

$$Z_{t, n}^* \xrightarrow[n \rightarrow \infty]{a.s.} Z_t^*, t = 1, 2, \dots$$

Let $A(t) = \{Z_i^* = Z_i, i = 1-k, \dots, t\}, t \geq 1$. Then $Z_1^* = Z_1$ so $P(A(1)) = 1$. Now for $t > 1$,

$$\begin{aligned} P(A^c(t) \cap A(t-1)) &= P(Z_i^* \neq Z_i, Z_i^* = Z_i, i \leq t-1) \\ &\leq P(Z_{t-1} \in B(R, \epsilon)) + P(Z_{t-1} \in C^c(R)) \\ &\leq \epsilon/2 + \sum_{i=1}^k P(|Z_{t-i}| > R/\sqrt{k}) \\ &\leq \epsilon/2 + \sum_{i=1}^k P(|a_{t-i}| > R/\sqrt{k} - M) \leq \epsilon. \end{aligned}$$

Thus $P(A(2)) = P(A(2) \cap A(1)) \geq 1 - \epsilon$. Suppose $P(A(r)) \geq 1 - r\epsilon, r = 1, \dots, t-1$. Then

$$P(A^c(t) | A(t-1)) = P(A^c(t) \cap A(t-1)) / P(A(t-1)) > \epsilon / (1 - (t-1)\epsilon)$$

so

$$P(A(t) | A(t-1)) \geq 1 - \epsilon / (1 - (t-1)\epsilon) = (1 - t\epsilon) / (1 - (t-1)\epsilon)$$

and

$$P(A(t)) = P(A(t) \cap A(t-1)) \geq (1 - (t-1)\epsilon) [(1 - t\epsilon) / (1 - (t-1)\epsilon)] = 1 - t\epsilon$$

Theorem 3.2. Suppose the function g defining process (3.2) is an arbitrary measurable function. Suppose also that $\{a_t\}$ has a density on \mathbf{R} and that the process $\{Z_t\}$ has fixed starting values Z_{1-k}, \dots, Z_0 . Then there is a sequence of TAR processes $\{Z_{i,n}^*(N), t, n, N \geq 1\}$, each having these same fixed starting values, and a sequence of integers $\{n_N\}$ such that $Z_{i,n_N}^*(N) \rightarrow Z_i$ almost surely as $N \rightarrow \infty$, for all $t \geq 1$.

Proof. For each $N \geq 0$ let

$$g_N(\cdot) = g(\cdot) I_{[0, N]}(|g(\cdot)|).$$

Then g_N is bounded and measurable. Let $\{Z_t, t \geq 0\}$ be defined by (3.2) and define

$$Z_t(N) = g_N(Z_{t-1}(N), \dots, Z_{t-k}(N)) + a_t, t \geq 1$$

where the startup values $(Z_0(N), \dots, Z_{1-k}(N)) = (Z_0, \dots, Z_{1-k})$.

Then clearly, for $t \geq 1$,

$$Z_t(N) = Z_t, N > \max_{1 \leq i \leq t} \{|g(Z_{i-1}, \dots, Z_{i-k})|\}$$

so that

$$Z_t(N) \xrightarrow{a.s.} Z_t \text{ as } N \rightarrow \infty.$$

But by the proof of Lemma 3.1 and the above, for $\epsilon_N \downarrow 0$ as $N \rightarrow \infty$ we may find a sequence of piecewise linear functions $\{g_n^N(\cdot); n, N \geq 1\}$ such that if

$$Z_{i,n}^*(N) = g_n^N(Z_{i-1,n}^*(N), \dots, Z_{i-k,n}^*(N)) + a_t$$

where $(Z_{0,n}^*(N), \dots, Z_{1-k,n}^*(N)) = (Z_0, \dots, Z_{1-k})$, then $Z_{t,n}^*(N) \xrightarrow{\text{a.s.}} Z_t^*(N)$ as $n \rightarrow \infty$ where $P(Z_t^*(N) \neq Z_t(N)) < \epsilon_N/3$. Let $\epsilon > 0$ and let $n_N(t)$ be such that

$$P(|Z_{t,n_N(t)}^*(N) - Z_t^*(N)| > \epsilon/3) < \epsilon_N/3.$$

Thus

$$P(|Z_{t,n_N(t)}^*(N) - Z_t(N)| > 2\epsilon/3) < 2\epsilon_N/3.$$

Now there is an $M = M(\epsilon_N)$ such that $N > M$ implies

$$P(|Z_t(N) - Z_t| < \epsilon/3) < \epsilon_N/3.$$

Thus $N > M$ implies

$$P(|Z_{t,n_N(t)}^*(N) - Z_t| > \epsilon) < \epsilon_N.$$

That is $Z_{t,n_N(t)}^*(N) \xrightarrow{P} Z_t$. Now for $t \geq 1$ we may take $\{n_N(t+1)\}$ to be a subsequence of $\{n_N(t)\}$. Then define a subsequence $\{n_N\}$ to be the resulting diagonal subsequence. We then have the approximating sequence $Z_{t,n_N}^*(N) \xrightarrow{P} Z_t$, for all $t \geq 1$. Hence there is a subsequence which converges a.s. to Z_t .

4. APPROXIMATION TO SPECIFIC CLASSES OF PROCESSES

Approximation to EAR models. Clearly EAR models are a special case of (3.2), so that Theorem 3.1 guarantees their almost sure approximation by TAR models.

Approximation to ARMA models with observation dependent coefficients. Consider models of the form

$$Z_t = \sum_{i=1}^p \varphi_i(Z_{t-1})Z_{t-i} + a_t - \sum_{j=1}^q \theta_j(Z_{t-1})a_{t-j} \quad \dots \quad (4.1)$$

where $Z_t = (Z_t, Z_{t-1}, \dots, Z_{t-k})$. Define (4.1) to be invertible if it has the mean square convergent representation

$$Z_t = \sum_{i=1}^{\infty} \psi_i(Z_{t-1}^{(i)})Z_{t-i} + a_t,$$

where $Z_t^{(i)} = (Z_t, Z_{t-1}, \dots, Z_{t-k-i+1})$ (4.2)

That is,
$$\sum_{i=1}^n \psi_i(Z_{t-1}^{(i)})Z_{t-i} + a_t \xrightarrow{\text{m.s.}} Z_t \quad \dots \quad (4.3)$$

as $n \rightarrow \infty, t \geq 1$. This implies convergence in probability and hence almost sure convergence of a subsequence. But TAR processes, being almost surely dense in processes of the form (4.3) are therefore almost surely dense in invertible processes of the form (4.1).

As bilinear processes are special cases of (4.1), we get almost sure TAR approximation to invertible bilinear processes.

Approximation to SDMs. If, in the proof of Theorem 3.1, we take $\mathbf{Z}_t' = (Z_t, \dots, Z_{t-k+1}, a_t, \dots, a_{t-l+1})$ and we define the process generating the \mathbf{Z}_t in state space form to be

$$\mathbf{Z}_t = f(\mathbf{Z}_{t-1}) + \mathbf{A}_t \quad \dots \quad (4.4)$$

where

$$\mathbf{A}_t' = (a_t, 0, \dots, 0)$$

$$[f(\mathbf{Z}_{t-1})]' = (g(\mathbf{Z}_{t-1}), Z_{t-1}, \dots, Z_{t-k}, a_{t-1}, \dots, a_{t-l})$$

for some $g: \mathbf{R}^{2k} \rightarrow \mathbf{R}$ continuous, then the results of the theorem still hold.

Thus piecewise linear processes of the form

$$Z_t = \mu^{(v)} + \sum_{i=1}^p \varphi_i^{(v)} Z_{t-i} + a_t - \sum_{j=1}^q \theta_j^{(v)} a_{t-j}, \quad \dots \quad (4.5)$$

$$\mathbf{Z}_{t-1} \in R^{(v)}, v = 1, \dots, r$$

where $\{R^{(v)}\}$ is a partition of \mathbf{R}^{2k} , are almost surely dense in the set of processes of the form (4.4). But these processes contain Priestley's SDMs as a subclass (see Priestley, 1980). Thus the processes defined in (4.5), which may be termed threshold ARMA, or TARMA, processes, may be used to approximate SDMs almost surely.

The practical difficulty in fitting processes such as (4.5) lies in determining when the vector of both observations and error terms lies in a given region.

Approximation of heteroscedastic processes. A referee has suggested that by allowing the distributions of error terms in the various regions of the approximating TAR model to differ, we might be able to approximate a wider class of processes. That is, the TAR models would be of the form

$$Z_t = \mu^{(i)} + \sum_{j=1}^p \varphi_j^{(i)} Z_{t-j} + a_t^{(i)}, (Z_{t-1}, \dots, Z_{t-p}) \in R^{(i)} \quad \dots \quad (4.6)$$

where for each i , $\{a_t^{(i)}\}$ would be i.i.d. but $a_t^{(i)}$ and $a_t^{(j)}$ might have different distributions, $i \neq j$.

Indeed, one interesting class of models that can be approximated by models of the form (4.6) are heteroscedastic versions of (3.1) and (3.2):

$$\mathbf{Z}_t = f(\mathbf{Z}_{t-1}) + \Sigma(\mathbf{Z}_{t-1})\mathbf{A}_t \quad \dots \quad (4.7)$$

and

$$Z_t = g(Z_{t-1}, \dots, Z_{t-k}) + \sigma(Z_{t-1}, \dots, Z_{t-k}) a_t \quad \dots \quad (4.8)$$

where $\Sigma: \mathbf{R}^k \rightarrow \mathbf{R}^k$ and $\sigma: \mathbf{R}^k \rightarrow (0, \infty)$. Then, assuming Σ is continuous, and $\{\Sigma_n\}$ is a sequence of functions approximating Σ uniformly on any compact

set in \mathbf{R}^k , an analogous version of Theorem 3.1 holds with the approximating sequence being

$$\mathbf{Z}_{t,n} = f_n(\mathbf{Z}_{t-1,n}) + \Sigma_n(\mathbf{Z}_{t-1,n})\mathbf{A}_t.$$

In particular piecewise linear f_n and piecewise constant Σ will give a multivariate version of (4.6) as approximants.

Similar modifications to the proofs of Lemma 3.1 and Theorem 3.2 show that TAR processes of the form (4.6) can approximate models of the form (4.8) with g and σ measurable functions in exactly the sense given by that lemma and theorem.

Acknowledgement. The author wishes to thank Neville Davies and John Pemberton for suggesting the problem of approximating nonlinear processes by TAR models.

REFERENCES

- CHAN, W. S., and TONG, H. (1986). On tests for non-linearity in time series analysis. *J. Forecasting* **5**, 217-228.
- GRANGER, C. W. J. and ANDERSEN, A. P. (1978). *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht: Göttingen.
- HAGGAN, V., HERAVI, S. N. and PRIESTLEY, M. B. (1984). A study of the application of state dependent models in non-linear time series analysis. *J. Time Series Anal.* **2**, 69-102.
- HAGGAN, V. and OZAKI, T. (1981). Modelling non-linear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, **68**, 189-196.
- OZAKI, T. (1981). Non-linear threshold autoregressive models for non-linear random vibrations. *J. Appl. Prob.*, **18**, 443-451.
- PETRUCCELLI, J. D. and DAVIES, N. (1986). A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika*, **73**, 687-694.
- PRIESTLEY, M. B. (1980). State-dependent models: A general approach to non-linear time series analysis. *J. Time Series Anal.*, **1**, 47-71.
- (1981). On the fitting of general non-linear time series models. In *Time Series Analysis: Theory and Practice I*, edited by O. D. Anderson. North-Holland: Amsterdam, 717-731.
- RUDDIN, W. (1966). *Real and Complex Analysis*, McGraw-Hill, New York.
- SUBBA, RAO, T. (1981). On the theory of bilinear time series models. *J. R. Statist. Soc. (B)*, **43**, 244-255.
- TONG, H. (1983). *Threshold Models in Non-Linear Time Series Analysis*. Lecture Notes in Statistics, Springer-verlag, New York.
- TONG, H. and LIM, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. Roy Statist. Soc., (B)* **42**, 245-292.

DEPARTMENT OF MATHEMATICAL SCIENCES
WORCESTER POLYTECHNIC INSTITUTE
WORCESTER, MASSACHUSETTS
USA.