

On Ergodicity of Some TAR(2) Processes *

Naoto Kunitomo
Faculty of Economics, University of Tokyo

December 2000

Summary

We give a set of sufficient conditions for the geometrical ergodicity and the non-explosiveness of the solutions in the second-order threshold autoregressive (TAR) processes. We also discuss some conditions for the geometrical ergodicity of the second-order simultaneous switching autoregressive (SSAR) processes. Unlike the linear autoregressive processes and the first-order TAR processes, the ergodic regions and non-explosive regions become quite complicated even in some special TAR processes.

1. Introduction

In the statistical time series analysis several nonlinear time series models have been proposed in the past decade. In particular, considerable attention has been paid to the class of threshold autoregressive (TAR) processes, which was systematically investigated by Tong (1990) and some applications have been reported. The statistical properties of the first order TAR processes have been investigated first by Petrucci and Woolford (1984), and later by Chen and Tsay (1991) in more details. Unlike the linear autoregressive models, however, the statistical properties of the second order TAR processes have not been fully investigated mainly due to some technical problems involved. It seems that the necessary and sufficient conditions for the ergodicity have been known only for the first order TAR processes. The main purpose of my work is to investigate the basic properties of the second-order TAR processes and the second-order SSAR (simultaneous switching autoregressive) processes.

Let $\{y_t\}$ be a sequence of scalar time series satisfying

$$(1.1) \quad y_t = \begin{cases} a_1 y_{t-1} + a_2 y_{t-2} + \sigma_1 v_t & \text{if } y_{t-d} \geq 0 \\ b_1 y_{t-1} + b_2 y_{t-2} + \sigma_2 v_t & \text{if } y_{t-d} < 0 \end{cases},$$

*My report was based on a revised version of Discussion Paper CIRJE-F-55, Faculty of Economics, University of Tokyo. I thank Dr. Seisho Sato of the Institute of Statistical Mathematics for the help of preparing some figures based on the simulations reported in my work.

where d is a positive (finite) integer parameter, and a_i, b_i, σ_i (> 0) ($i = 1, 2$) are unknown parameters, and $\{v_t\}$ are a sequence of independently and identically distributed (i.i.d.) random variables having an absolutely continuous density $f(v)$ with respect to the Lebesgue measure and $E[v_t] = 0$. We assume that $f(v)$ is continuous and everywhere positive in $\mathbf{R}$. The second-order threshold autoregressive model given by (1.1) will be denoted as TAR(2:d) and the integer-valued parameter d is called the delayed parameter. (We also use the notation as TAR(2) for the second order TAR processes.) Petrucelli and Woolford (1984) have considered the first-order TAR process when $a_2 = b_2 = 0$ and $d = 1$, which is denoted as TAR(1:1). They have shown that the necessary and sufficient conditions for the geometrical ergodicity are given by

$$(1.2) \quad a_1 < 1, b_1 < 1, a_1 b_1 < 1.$$

Chen and Tsay (1991) have extended their results to the TAR(1:d) processes when d takes an arbitrary positive integer. The conditions they have obtained for the geometrical ergodicity include (1.2) as a special case when $d = 1$.

On the other hand, Kunitomo and Sato (1996, 2000), and Sato and Kunitomo (1996) have proposed the class of simultaneous switching autoregressive (SSAR) processes, which can be regarded as a natural extension of the TAR processes for some econometric applications. The second-order SSAR model can be written as

$$(1.3) \quad y_t = \begin{cases} a_1 y_{t-1} + a_2 y_{t-2} + \sigma_1 v_t & \text{if } y_t \geq y_{t-1} \\ b_1 y_{t-1} + b_2 y_{t-2} + \sigma_2 v_t & \text{if } y_t < y_{t-1} \end{cases},$$

where a_i, b_i, σ_i (> 0) ($i = 1, 2$) are unknown parameters, and $\{v_t\}$ are a sequence of independently and identically distributed (i.i.d.) random variables having an absolutely continuous density $f(v)$ with respect to the Lebesgue measure and $E[v_t] = 0$. We also assume that $f(v)$ is continuous and everywhere positive in $\mathbf{R}$. The discrete time series model given by (1.3) will be denoted by SSAR(2). By imposing the restrictions on parameters given by

$$(1.4) \quad \frac{1 - a_1}{\sigma_1} = \frac{1 - b_1}{\sigma_2} = r_1, \frac{a_2}{\sigma_1} = \frac{b_2}{\sigma_2} = r_2,$$

this time series model can be written as

$$(1.5) \quad y_t = \begin{cases} a_1 y_{t-1} + a_2 y_{t-2} + \sigma_1 v_t & \text{if } v_t \geq r_1 y_{t-1} - r_2 y_{t-2} \\ b_1 y_{t-1} + b_2 y_{t-2} + \sigma_2 v_t & \text{if } v_t < r_1 y_{t-1} - r_2 y_{t-2} \end{cases},$$

where r_i ($i = 1, 2$) are unknown parameters. We should note that the class of simultaneous switching autoregressive (SSAR) process is different from the TAR models given by (1.1). The two phases at t are determined by both the past time series and the present innovation at t , we do not have any delayed parameter in the SSAR processes. Then the SSAR(2) process has the Markovian representation with the state vector $\mathbf{y}'_t = (y_t, y_{t-1})$ by using the relation

$$(1.6) \quad y_t = y_{t-1} + [\sigma_1 I(v_t \geq r_1 y_{t-1} - r_2 y_{t-2}) + \sigma_2 I(v_t < r_1 y_{t-1} - r_2 y_{t-2})][v_t - (r_1 y_{t-1} - r_2 y_{t-2})],$$

where $I(\cdot)$ is the indicator function. When $\sigma_1 = \sigma_2 = \sigma$, then this SSAR process becomes the standard $AR(2)$ model by a re-parametrization. Kunitomo and Sato (1996) have shown that even the first order SSAR process (denoted as SSAR(1)) when $a_2 = b_2 = 0$ gives us some explanation and description on an important aspect of the asymmetrical movement of time series in two different (up-ward and down-ward) phases. The ergodicity condition for the SSAR(1) process is the same as (1.2) with the coherency conditions implied by (1.4).

2. Some Preliminaries

The first important property of any statistical time series model is whether it is ergodic or not. For the Markovian time series models with continuous state spaces and discrete time intervals, the geometrical ergodicity and the related concepts have been developed in the nonlinear time series analysis. For the sake of completeness, we mention to its definition and the drift criterion. For the more precise definitions of related concepts including irreducibility, aperiodicity, and ergodicity of the Markov chains with the general state spaces, see Nummelin (1984) or the Appendix of Tong (1990).

Definition 1 (Geometrical Ergodicity) : Let $\{y_t\}$ be an $m \times 1$ Markovian process with the state space of R^m .

(i) $\{y_t\}$ is geometrically ergodic if there exists a probability measure π on $(R^m, \mathcal{B}(R^m))$, a non-negative constant $\rho < 1$ and π -integrable non-negative measurable function $h(\cdot)$ such that

$$(2.1) \quad \|P^n(x, \cdot) - \pi(\cdot)\|_\tau \leq \rho^n h(x),$$

where $\|\cdot\|_\tau$ denotes the total variation norm, $x = (x_i)$ is an $m \times 1$ vector, and $P(x, \cdot)$ is the transition probability.

(ii) $\{y_t\}$ is ψ -irreducible if for any $x \in R^m$, $A \in \mathcal{B}(R^m)$ with $\psi(A) > 0$ ($\psi(\cdot)$ is a σ -finite measure), and

$$(2.2) \quad \sum_{n=1}^{\infty} P^n(x, A) > 0.$$

For the geometrical ergodicity of the Markov Chains with the continuous general states, Tjøstheim (1990) has given a drift criterion, which will be useful for our purpose and thus reported as Lemma 1. It is an extension of the well-known drift criterion on the Markov chain with the general states due to Tweedie (1975).

Lemma 1 (Tjøstheim (1990)) : Assume that $\{y_t\}$ is an aperiodic ψ -irreducible Markov Chain with the state space of R^m and g is a non-negative continuous (measurable) function. Then $\{y_t\}$ is geometrically ergodic if there exist a positive integer h , a small set C satisfying (2.2) with $\psi(C) > 0$, positive constants $\epsilon > 0$, $M < +\infty$, and $r > 1$ such that

$$(2.3) \quad rE[g(y_{t+h})|y_t = y] \leq g(y) - \epsilon \text{ if } y \in C^c,$$

for any positive constant ϵ and

$$(2.4) \quad E[g(y_{t+h})|y_t = y] \leq M \text{ if } y \in C,$$

where C^c is the complement set of C .

We note that when the density function of disturbances $f(v)$ is continuous and everywhere positive in $\mathbf{R}$ we can take a compact set as the small set in Lemma 1. See Theorem A.1.7 of Tong (1990) on the related problems.

Now we introduce another concept on the stability of the solution, which is slightly different from the geometrical ergodicity. Because its conditions are slightly weaker than those in Lemma 1 and they are necessary for the geometrical ergodicity, we use the terminology of *Near Geometrical Ergodicity*. We shall use this concept in this paper due to the technical reason indicated in later sections. However, the detail of its mathematical properties has not been fully explored and is still under investigation.

Let Q be the state space of the stochastic process $\{y_t\}$. Then we partition the state space Q into a finite number of disjoint subspaces Q^i ($i = 1, \dots, k$) such that $Q = \bigcup_{i=1}^k Q^i$ ($\psi(Q^i) > 0$, $Q^i \cap Q^j = \emptyset$; $i \neq j$) and for any t

$$(2.5) \quad 1 = \sum_{i=1}^k I(y_t \in Q^i),$$

where $I(\cdot)$ is the indicator function and k is a positive integer.

Definition 2 (Near Geometrical Ergodicity) : We say that the solution $\{y_t\}$ is near geometrically ergodic if (i) for any i, i' ($i, i' = 1, \dots, k$) and (sufficiently large) positive integers h_j ($j = j(i, i') \geq 1$), there exist a sequence of positive constants ϵ_j and r_j ($r_j > 1$) such that

$$(2.6) \quad r_j E[g(y_{t+h_j}) I(y_{t+h_j} \in Q^{i'}) | y_t = y] \leq g(y) - \epsilon_j \text{ if } y \in Q^i \cap C^c,$$

where C is a small set satisfying $\psi(C) > 0$ in (2.2), and (ii) there exists a positive constant M such that

$$(2.7) \quad E[g(y_{t+h_j}) I(y_{t+h_j} \in Q^{i'}) | y_t = y] \leq M \text{ if } y \in Q^i \cap C$$

for any i, i' ($i, i' = 1, \dots, k$) and positive integers h_j ($h_j \geq 1$).

We notice that there is a decomposition

$$(2.8) \quad I(y_{t+h} \in Q^{i'}) I(y_t \in Q^i) = \sum_{l=1}^k I(y_{t+h} \in Q^{i'}) I(y_{t+h_1} \in Q^l) I(y_t \in Q^i)$$

for any $h > h_1 \geq 1$ and $i, i' = 1, 2, \dots, k$. Then it is apparent that (2.6) is reduced to (2.3) if the term $I(\cdot)$ can be deleted in (2.6) with the common $r = r_j$ and $h = h_j$ ($j \geq 1$) with a finite j 's. Hence the near geometrical ergodicity we introduce in this section coincides with the geometrical ergodicity in the nonlinear time series analysis if we can take the common positive integer $h = h_j$ and the common constant $r = r_j$ ($j \geq 1$). It is a kind of the stability property or non-explosiveness of the sample paths of the underlying stochastic process, whose behavior is near to the geometrical ergodicity.

3. Ergodic conditions for TAR(2) in the leading case

In this section we consider the conditions for the geometrical ergodicity of the TAR(2) processes in a special case. For this purpose we shall utilize the Markovian representation

of the TAR(2) processes. Let $\mathbf{y}_t = (y_t, y_{t-1})'$ be a 2×1 vector of the time series generated by $\{y_t\}$. The second order TAR process given by (1.1) can be represented by

$$(3.1) \quad \mathbf{y}_t = \begin{cases} \mathbf{A}\mathbf{y}_{t-1} + D\sigma_1 v_t & \text{if } \mathbf{e}'_k \mathbf{y}_{t-1} \geq 0 \\ \mathbf{B}\mathbf{y}_{t-1} + D\sigma_2 v_t & \text{if } \mathbf{e}'_k \mathbf{y}_{t-1} < 0 \end{cases},$$

where $\mathbf{e}_1 = (1, 0)'$, $\mathbf{e}_2 = (0, 1)'$ (for $\mathbf{e}_k$ with $k = 1, 2$), and $D = (1, 0)'$ are 2×1 constant vectors, and $\mathbf{A}, \mathbf{B}$ are 2×2 coefficient matrices given by

$$(3.2) \quad \mathbf{A} = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_1 & b_2 \\ 1 & 0 \end{pmatrix}.$$

In the rest of this section we shall consider the leading case for the TAR(2) processes when $b_2 = 0$. This is simply because we can obtain general characterizations on the geometrical ergodic regions in the leading case and the necessary and sufficient conditions for the geometrical ergodicity can be mostly obtainable. We have to stress that even in this leading case our conditions sometimes become quite complicated and non-standard in comparison with the results known for the TAR(1) processes. In order to obtain the conditions for the geometrical ergodicity and state our results, we partition the parameter space of (a_1, a_2) into four different regions given by

$C_1 = \{a_1 \geq 0, a_2 \geq 0\}$, $C_2 = \{a_1 < 0, a_2 \geq 0\}$, $C_3 = \{a_1 < 0, a_2 < 0\}$, and $C_4 = \{a_1 \geq 0, a_2 < 0\}$, respectively.

Because we set $b_2 = 0$, we need to consider two cases when $b_1 \leq 0$ and $b_1 > 0$. Then it is intuitively obvious that in the latter case we have to restrict the conditions for the geometrical ergodicity when $0 < b_1 < 1$ in order to avoid the possible explosion of the sample paths of the solutions. Because we have the density function $f(v)$ over R , this can be proven in the rigorous way.

3.1 TAR(2:1) Processes

First, we consider the TAR(2:1) process when $b_1 > 0$ and $b_2 = 0$. This is the simplest case in the TAR(2) process in terms of our conditions on the coefficients. Although the characteristic roots of $\mathbf{B}$ are b_1 ($0 < b_1 < 1$) and 0, we have some complications due to the behaviors of two characteristic roots of $\mathbf{A}$. Let

$$(3.3) \quad D(\mathbf{A}) = a_1^2 + 4a_2,$$

which is the discriminant of the characteristic equation for the first phase in (3.1) given by

$$(3.4) \quad g_A(\lambda) = \lambda^2 - a_1\lambda - a_2 = 0.$$

Then we can present the necessary and sufficient conditions for the geometrical ergodicity in the present case. Some of proofs and derivations of the propositions in this section are given in Section 5. We consider a set of conditions :

Condition I :

$$C_1 : a_1 + a_2 < 1, 0 < b_1 < 1,$$

$$C_2 : 0 < b_1 < 1 ,$$

$$C_3 : 0 < b_1 < 1 ,$$

$$C_4 : [\text{either } a_1 + a_2 < 1 \text{ (} 0 \leq a_1 < 2, D(A) \geq 0 \text{) or } D(A) < 0], 0 < b_1 < 1 .$$

Proposition 1 : *For the TAR(2:1) process with $b_1 > 0$ and $b_2 = 0$, the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by Condition I.*

Second, we consider the TAR(2:1) process when $b_1 \leq 0$ and $b_2 = 0$. The conditions in this case become far more complicated than in the first case. This is mainly because the stochastic process $\{y_t\}$ can be ergodic when $b_1 < 0$ and its absolute value is greater than 1. In order to deal with some complications involved, we define $\rho(A^k B)$ be the non-zero characteristic root of the matrix $A^k B$ for any positive integer k . In the present case

$$\rho(A^k B) = e_1' A^k b ,$$

where $b' = (b_1, 1)$.

Then we give a set of conditions for the geometrical ergodicity and the near geometrical ergodicity of the sample paths of the solutions. The concept of the near geometrical ergodicity of the solutions has been given in Definition 2 in Section 2. We note that it is not needed for the linear time series processes and TAR(1) processes. We consider a set of conditions :

Condition II :

$$C_1 : a_1 + a_2 < 1 ,$$

$$C_2 : \rho(AB) < 1 ,$$

$$C_3 : \rho(A^j B) < 1 \text{ (} j = 1, 2 \text{),}$$

$$C_4 : \text{either } [a_1 + a_2 < 1 \text{ (} 0 \leq a_1 < 2, D(A) \geq 0 \text{)] or } [D(A) < 0 \text{ and there exists some } j \text{ (} j \geq 1 \text{) such that } 0 \leq \rho(A^j B) < 1] .$$

Proposition 2 : *For the TAR(2:1) process with $b_1 \leq 0$ and $b_2 = 0$, (i) the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by Condition II with C_1, C_2 , and C_3 , and (ii) the sufficient conditions for the near geometrical ergodicity of the solution $\{y_t\}$ are given by Condition II with C_4 .*

As an illustration, we present two figures of the ergodic regions and non-explosive regions for the TAR(2:1) processes, which are based on the simulations of the stochastic processes $\{y_t\}$. Contrary to the ergodic regions for the linear AR(2) models which have been known in the statistical time series analysis (see Brockwell and Davis (1991) for instance), they are often unbounded as we see in these figures. Some of the conditions above can be written more explicitly by using

$$\rho(AB) = a_1 b_1 + a_2 , \rho(A^2 B) = a_1(a_1 b_1 + a_2) + a_2 b_1 .$$

The most complicated region in the TAR(2:1) process is C_4 in Condition II and we have found some strange shapes as the non-explosive regions for the sample paths of the solutions depending upon the parameter values of A and B . Although we cannot

give the complete characterizations of the ergodic regions on the present case, we may conjecture that the conditions are necessary and sufficient for the geometrical ergodicity in all cases. As an immediate corollary of the above two propositions, we can obtain the result originally derived by Petrucci and Woolford (1984) for the TAR(1:1) process.

Corollary 1 : *For the TAR(1:1) model, the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by*

$$(3.5) \quad a_1 < 1, b_1 < 1, a_1 b_1 < 1.$$

3.2 TAR(2:2) Processes

Next, we consider the TAR(2:2) process when $b_1 \leq 0$ and $b_2 = 0$. Contrary to the TAR(2:1) process, this is simpler than the case when $b_2 > 0$ in the TAR(2:2) process. In the present case we can give the necessary and sufficient conditions :

Condition III :

- $C_1 : a_1 + a_2 < 1, \rho(AB^2) < 1,$
- $C_2 : \rho(AB) < 1,$
- $C_3 : \rho(AB) < 1, \rho(AB^2) < 1,$
- $C_4 : \text{either } [a_1 + a_2 < 1 \text{ } (0 \leq a_1 < 2, D(A) \geq 0)] \text{ or } [D(A) < 0, \rho(AB^2) < 1].$

Proposition 3 : *For the TAR(2:2) process with $b_1 \leq 0$ and $b_2 = 0$, the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by Condition III.*

Second, we consider the TAR(2:2) process when $b_1 > 0$ and $b_2 = 0$. It is far more complicated than in the previous case as for the TAR(2:1) process when $b_1 \leq 0$ and $b_2 = 0$. In this case we give a set of conditions for the geometrical ergodicity and the non-explosiveness of the solution.

Condition IV :

- $C_1 : a_1 + a_2 < 1, 0 < b_1 < 1,$
- $C_2 : a_1 + a_2 < 1 \text{ and there exists some } j \text{ } (j \geq 2) \text{ such that } 0 \leq \rho(A^j B) < 1,$
- $C_3 : \rho(A^2 B) < 1, 0 < b_1 < 1,$
- $C_4 : \text{either } [a_1 + a_2 < 1 \text{ } (0 \leq a_1 < 2, D(A) \geq 0)] \text{ or } [D(A) < 0], 0 < b_1 < 1.$

Proposition 4 : *For the TAR(2:2) process with $b_1 > 0$ and $b_2 = 0$, (i) the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by Condition IV with C_1, C_3 , and C_4 , and (ii) the sufficient conditions for the near geometrical ergodicity of the solution $\{y_t\}$ are given by Condition IV with C_2 .*

As an illustration, we present two figures of the ergodic and the non-explosive regions for the TAR(2:2) processes, which are based on simulations. Some of the conditions above can be written more explicitly by using

$$\rho(AB^2) = b_1(a_1 b_1 + a_2).$$

We note that some of the conditions for the TAR(2:2) process in the leading case when $D(\mathbf{A}) < 0$ have been partially discussed by Tong (1990) without the disturbance terms. From Page 70 of Tong (1990) we set $b_1 = -0.9, a_1 = 1.8, a_2 = -0.9$ as Example 1 and $b_1 = -1.1, a_1 = 0.6, a_2 = -0.1$ as Example 2. By using Proposition 3, we can conclude that Example 1 is not geometrically ergodic while Example 2 is geometrically ergodic.

Contrary to the ergodic regions for the linear AR(2) processes, some of them are unbounded as we see in these figures. The most complicated region in the TAR(2:2) process is C_2 in Condition IV although it may be difficult to judge this finding directly from the figure. We have found some strange shapes as the non-explosiveness regions in C_2 depending upon the parameter values of $\mathbf{A}$ and $\mathbf{B}$. It seems that the complications involved in this case is different from the corresponding one in the region C_4 for the TAR(2:1) process when $b_1 \leq 0$. However, we have not succeeded in the complete characterization of this situation except the present conditions we have obtained.

As an immediate corollary of the above two propositions, we have the result obtained by Chen and Tsay (1991) for the first-order threshold model TAR(1:d) when $d = 2$.

Corollary 2 : *For the TAR(1:2) process, the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by*

$$(3.6) \quad a_1 < 1, b_1 < 1, a_1 b_1 < 1, a_1^2 b_1 < 1, a_1 b_1^2 < 1.$$

We should mention that Chen and Tsay (1991) have given the necessary and sufficient conditions for the TAR(1:d) processes with any positive integer-valued parameter d .

4. Discussions

4.1 SSAR(2) Processes

By using (1.3)-(1.5), the SSAR(2) process have the Markovian representation, which is similar to the one given in (3.1) for the TAR(2) processes. Let $\mathbf{y}_t = (y_t, y_{t-1})'$ be a 2×1 vector of time series. Then the SSAR(2) process in (1.3)-(1.5) can be represented by

$$(4.1) \quad \mathbf{y}_t = \begin{cases} \mathbf{A}\mathbf{y}_{t-1} + D\sigma_1 v_t & \text{if } e' \mathbf{y}_t \geq 0 \\ \mathbf{B}\mathbf{y}_{t-1} + D\sigma_2 v_t & \text{if } e' \mathbf{y}_t < 0 \end{cases},$$

where $e = (1, -1)'$, $D = (1, 0)'$, and the coefficient matrices $\mathbf{A}$ and $\mathbf{B}$ are given by (3.2) with the coherency conditions (1.4). We first note that due to the structure of the SSAR processes, it is not possible to investigate the leading case as in Section 3. The determination of phases at time t depends on a finite number of past time series and the present innovation as given by (1.6), the stability of the paths of the solution become quite complex and it depends on the complicated products of matrices $\mathbf{A}^k \mathbf{B}^l$ and $\mathbf{B}^k \mathbf{A}^l$ for $0 \leq k, l$ in general. As the sufficient conditions, we consider the restrictions :

Condition V :

$$(4.2) \quad a_1 + |a_2| < 1, b_1 + |b_2| < 1, (a_1 - |a_2|)(b_1 - |b_2|) < 1, \min\{|a_2|, |b_2|\} < 1.$$

By using the coherency conditions (1.4) on a_i, b_i ($i = 1, 2$) for the SSAR(2) process, the above conditions can be rewritten in the parameter space of σ_i ($i = 1, 2$) and τ_i ($i = 1, 2$) as

Condition V¹ :

$$(4.3) \quad r_1 > |r_2|, r_1 + |r_2| < \frac{1}{\sigma_1} + \frac{1}{\sigma_2}, |r_2| < \min\left\{\frac{1}{\sigma_1}, \frac{1}{\sigma_2}\right\}.$$

Presently it is only possible to show that the first three conditions are necessary for the geometrical ergodicity by investigating possible many cases which would be occurred and we omit their lengthy proofs. The last condition in (4.3) is quite strong one, which is not necessary, but we expect that **Condition V** are sufficient for the geometrical ergodicity of $\{y_t\}$ in the SSAR(2) process. They are slightly weaker than the sufficient conditions

$$(4.4) \quad \rho = \min\{|a_1| + |a_2|, |b_1| + |b_2|\} < 1.$$

This type of conditions for the TAR(p :1) processes ($p \geq 1$) has been obtained by Chan and Tong (1985). However, they are not necessary even for the TAR(2:1) processes as we have shown in Section 3. As an illustration, we present one figure of the non-explosive regions for the SSAR(2) process, which is based on the simulations and drawn in the (r_1, r_2) phase. The geometrically ergodic regions in this case seem to be different from the corresponding ones in the TAR(2) processes and the ergodic region is bounded in the (r_1, r_2) -space due to the coherency conditions given by (1.4). The conditions for the geometrical ergodicity of the SSAR(1) process is considerably simpler than the SSAR(2) process, which can be summarized as the next proposition.

Proposition 5 : *For the SSAR(1) process, the necessary and sufficient conditions for the geometrical ergodicity of $\{y_t\}$ are given by*

$$(4.5) \quad a_1 < 1, b_1 < 1, a_1 b_1 < 1.$$

By using the coherency conditions given by (1.4) for the SSAR(2) process, these conditions can be further re-written as

$$(4.6) \quad 0 < r_1 < \frac{1}{\sigma_1} + \frac{1}{\sigma_2}.$$

From this representation it is clear that (4.2) and (4.3) are natural generalizations of (4.5) and (4.6) in a sense. We should note that the ergodic region in terms of the (r_1, r_2) -space is bounded for the SSAR(1) model.

4.2 TAR(2) Processes

In principle it would be possible to develop the conditions for the geometrical ergodicity and the near geometrical ergodicity for the general case of the TAR(2) processes. However, they become substantially more complicated than in the leading case (i.e. $b_2 = 0$) we have discussed in Section 3. In particular, it is quite tedious to write down the explicit

expressions in terms of the coefficient matrices A and B as the necessary and/or sufficient conditions of the geometrical ergodicity and the near geometrical ergodicity. Thus for the illustrative purpose we have done some simulations. Among many simulations we only present several figures having complex shapes in some regions. However, they have some similar aspects essentially to those in the leading cases when the TAR(2:1) process with $b_1 \leq 0$ and the TAR(2:2) process with $b_1 > 0$ as we have discussed in Section 3. For illustrations we also give some figures suggesting the mathematical complexities involved in our situations. Also we have confirmed that the analytical characterizations of those figures are not easy tasks.

4.3 Concluding Remarks

Since the conditions for the geometrical ergodicity are the basic property of the Markovian stochastic processes, they have some implications for statistical inferences and the modelling procedure of the TAR(2) processes. In general, from our findings in this paper we expect that the higher order TAR(p) processes with an arbitrary delayed parameter d have quite complicated ergodic conditions. Although it is possible to use the least squares estimation method¹ for the consistent estimation of the unknown parameters in the TAR processes, we need a more careful investigation on the properties of the estimation results in empirical studies. Because of our results reported in the previous sections, however, it is far beyond the scope of this paper to have a complete characterization of the stochastic processes of the TAR(p) and the SSAR(p) processes for practical usages and there still remains many statistical problems to be solved.

5. The Method of Proofs

In this section, we discuss our method of derivations and proofs of our results in Section 3. Due to the lack of space, the interested readers should ask the author to send the revised version of the full paper presented. The method of proofs is basically similar to the one developed by Chen and Tsay (1991) for the TAR(1) processes. For the theoretical results on Markov chain with the general state space, see Nummelin (1984), or Tjøstheim (1990). Here we prepare some notations used in this section. Let $\mathcal{F}_t$ be the σ -field generated by a sequence of random variables $\{y_s, s \leq t\}$ and we use the notation for the conditional expectation $E_t[\cdot] = E[\cdot | \mathcal{F}_t]$. Also define a sequence of time dependent phases for the stochastic process $\{y_t\}$ as $Q_t^1 = \{y_t > 0, y_{t-1} > 0\}$, $Q_t^2 = \{y_t \leq 0, y_{t-1} > 0\}$, $Q_t^3 = \{y_t \leq 0, y_{t-1} \leq 0\}$, and $Q_t^4 = \{y_t > 0, y_{t-1} \leq 0\}$.

By using the indicator function $I(\cdot)$, we can decompose 1 into the indicator functions with four different phases as $1 = I(Q_t^1) + I(Q_t^2) + I(Q_t^3) + I(Q_t^4)$. Then we can further decompose $I(Q_t^1)$ as $I(Q_t^1) = I(Q_t^1 Q_{t-1}^1) + I(Q_t^1 Q_{t-1}^4)$, for instance. The most important technical finding in our derivations and proofs of our results lies in the fact that we can ignore many terms when we evaluate the growth condition (2.3) as long as for the leading cases (i.e. $b_2 = 0$).

¹ For the SSAR processes, however, Sato and Kunitomo (1996) have shown that the standard least squares method does not give us any reliable estimation results.

We illustrate the details of our method by using the case of TAR(2:1) when $b_1 > 0$. [1] In this stochastic process we notice that for Q_{i+h}^4 ($h \geq 1$) we have $y_{t+h-1} \leq 0$ and $y_{t+h} = b_1 y_{t+h-1} + v_{t+h} > 0$. It implies that $E_{t+h-1}[|y_{t+h}|I(Q_{i+h}^4)]$ and $|y_{t+h-1}|$ are bounded because we have $b_1 > 0$, $v_{t+h} > -b_1 y_{t+h-1} \geq 0$, and $2|v_{t+h}| > |y_{t+h}|$. Then we have used the relation

$$(5.1) \quad \frac{1}{b_1} E_{t+h-1}[|v_{t+h}|] > |y_{t+h-1}| \geq 0.$$

We notice that the boundedness of the conditional expectations of y_{t+h-1} and y_{t+h} implies the boundedness of the conditional expectation of y_{t+h+1} . Sequentially we can show that $E_t[|y_{t+h+k}|I(Q_{i+h}^4)]$ are bounded for any integer $k \geq 1$. Hence we can find a positive constant $c_{1.1}$ such that $E_t[||y_{t+h+k}||I(Q_{i+h}^4)] \leq c_{1.1}$ for any (positive) integers $h \geq 1$ and $k \geq 1$. By using this boundedness relation, we have several consequences. For instance, since we can decompose $I(Q_{i+h}^1) = I(Q_{i+h}^1 Q_{i+h-1}^1) + I(Q_{i+h}^1 Q_{i+h-1}^4)$ and $I(Q_{i+h}^2) = I(Q_{i+h}^2 Q_{i+h-1}^1) + I(Q_{i+h}^2 Q_{i+h-1}^4)$, the conditional expectations of $E_t[||y_{t+h}||I(Q_{i+h}^1 Q_{i+h-1}^1)]$ and $E_t[||y_{t+h}||I(Q_{i+h}^2 Q_{i+h-1}^1)]$ are bounded for any integer $h \geq 1$.

By this consideration on the present case, we only need to evaluate the conditional expectation terms associated with the four phases on the process :

$$I(Q_{i+h}^1 Q_{i+h-1}^1), I(Q_{i+h}^2 Q_{i+h-1}^1), I(Q_{i+h}^3 Q_{i+h-1}^2 Q_{i+h-2}^1), \text{ and } I(Q_{i+h}^3 Q_{i+h-1}^3).$$

Then we shall consider the ergodic conditions for four regions of the parameter values C_1, C_2, C_3 , and C_4 , separately.

[2] C_1 : In this case we first notice that $0 \leq e'_1 A l < 1$ implies $0 \leq e'_1 A^2 l < 1$ for a 2×1 vector $l = (1, 1)'$. Define the indicator functions by $I_{i+2}^{(1)} = I(y_{t+2} \geq y_{t+1})$ and $I_{i+2}^{(2)} = I(y_{t+2} < y_{t+1})$. Then there exists a positive $c_{1.2}$ such that

$$\begin{aligned} & E_t[||y_{t+2}|| (I(Q_{i+2}^1 Q_{i+1}^1) + I(Q_{i+2}^2 Q_{i+1}^1))] \\ & \leq E_t[(e'_1 A^2 l I_{i+2}^{(1)} + e'_1 A l I_{i+2}^{(2)}) ||y_t|| (I(Q_{i+2}^1 Q_{i+1}^1) + I(Q_{i+2}^2 Q_{i+1}^1))] + c_{1.2}. \end{aligned}$$

Hence we can find positive constants $c_{1.3}$ and $\delta_{1.1}$ ($0 \leq \delta_{1.1} < 1$) such that

$$E_t[||y_{t+2}|| (I(Q_{i+2}^1 Q_{i+1}^1) + I(Q_{i+2}^2 Q_{i+1}^1))] \leq \delta_{1.1} ||y_t|| E_t[I(Q_{i+2}^1 Q_{i+1}^1) + I(Q_{i+2}^2 Q_{i+1}^1)] + c_{1.3}.$$

For the phase $Q_{i+2}^3 Q_{i+1}^2$, we substitute (3.1) for $||y_{t+2}|| = -y_{t+2} I_{i+2}^{(1)} - y_{t+1} I_{i+2}^{(2)}$ and we take positive constants $c_{1.4}$ and $\delta_{1.2}$ ($0 \leq \delta_{1.2} < 1$) such that

$$\begin{aligned} & E_t[||y_{t+2}|| I(Q_{i+2}^3 Q_{i+1}^2)] \\ & = E_t[(b_1 |e'_1 A y_t| I_{i+2}^{(1)} + |e'_1 A y_t| I_{i+2}^{(2)} + |v_{t+2} + b_1 v_{t+1}| I_{i+2}^{(1)} + |v_{t+1}| I_{i+1}^{(2)}) I(Q_{i+2}^3 Q_{i+1}^2)] \\ & \leq \delta_{1.2} ||y_t|| E_t[I(Q_{i+2}^3 Q_{i+1}^2)] + c_{1.4}, \end{aligned}$$

where we have taken a positive constant $\delta_{1.2} = \max\{b_1(a_1 + a_2), a_1 + a_2\}$. Similarly, for the phase $Q_{i+2}^3 Q_{i+1}^3$, we can find a positive constant $c_{1.5}$ such that

$$E_t[||y_{t+2}|| I(Q_{i+2}^3 Q_{i+1}^3)] \leq \max\{b_1^2, b_1\} (-y_t) E_t[I(Q_{i+2}^3 Q_{i+1}^3)] + c_{1.5}.$$

By summarizing the above inequalities on four phases, we can find positive constants $c_{1.6}$ and $\delta_{1.3}$ ($0 \leq \delta_{1.3} < 1$) such that $E_t[||y_{t+2}||] \leq \delta_{1.3} ||y_t|| + c_{1.6}$, which leads to (2.3) in Lemma 1. This is because we can find a sufficiently large M and a compact set $C(M)$

depending on M such that $C(M) = \{\|\mathbf{y}\| \leq M\}$ and the growth condition (2.3) in Lemma 1 can be satisfied.

In the following derivations for other cases we need to use the above type of arguments repeatedly in each case. Because the arguments are quite similar and tedious, however, we shall try to discuss the essential differences and not to repeat the same arguments.

[3] C_2 : By repeating the procedure we have used in [2] and taking the conditional expectations given $\mathcal{F}_t$, we can find a positive integer $h (\geq 2)$ and a positive constant $c_{1.7}$ such that

$$(5.2) \quad E_t[\|\mathbf{y}_{t+h}\|] \leq E_t \left[\|\mathbf{y}_{t+h}\| \left(I(Q_{t+h}^1 Q_{t+h-1}^1 Q_{t+h-2}^1) + I(Q_{t+h}^2 Q_{t+h-1}^1 Q_{t+h-2}^1) \right. \right. \\ \left. \left. + I(Q_{t+h}^3 Q_{t+h-1}^2 Q_{t+h-2}^1) + I(Q_{t+h}^3 Q_{t+h-1}^3 Q_{t+h-2}^3) \right) \right] + c_{1.7}.$$

Since $c_{1.7}$ is bounded, we need to evaluate its first four terms. Let λ_i ($i = 1, 2$) be the characteristic roots of $g_A(\lambda) = 0$ in (3.4). Because $a_1 < 0$ and $a_2 \geq 0$ in C_2 , we have the relation that $\lambda_1 > 0 > \lambda_2$ and $|\lambda_2| > |\lambda_1|$. Then there exists a positive integer h_1 such that $\mathbf{e}'_1 \mathbf{A}^{h_1} \mathbf{y}_t < 0$ given $\mathbf{y}_t \in Q_t^1$ because each component of $\mathbf{e}'_1 \mathbf{A}^{h_1}$ become eventually negative for a sufficiently large h_1 . If we write

$$(5.3) \quad \mathbf{y}_{t+h_1} = \mathbf{e}'_1 \mathbf{A}^{h_1} \mathbf{y}_t + \sigma_1 \left(\sum_{i=1}^{h_1} \mathbf{e}'_1 \mathbf{A}^{i-1} \mathbf{e}_1 v_{t+i} \right)$$

and we denote the second term of (5.4) as w_{t+h_1} , then we have the condition $w_{t+h_1} > -\mathbf{e}'_1 \mathbf{A}^{h_1} \mathbf{y}_t \geq 0$ when $\mathbf{y}_{t+h_1} \in Q_{t+h_1}^1$. Hence there exists a positive constant $c_{1.8}$ such that $E_t[\|\mathbf{y}_{t+h_1}\| I(\Pi_{i=0}^{h_1} Q_{t+i}^1)] \leq c_{1.8}$.

For the last term of (5.3), we use the relation that $\mathbf{y}_{t+h_2} = b_1 \mathbf{y}_{t+h_2-1} + v_{t+h_2}$ when $\mathbf{y}_{t+h_2} \in Q_{t+h_2}^3$ for any positive integer h_2 and we can take a positive $c_{1.9}$ such that

$$E_t[\|\mathbf{y}_{t+h_2}\| I(\Pi_{i=0}^{h_2} Q_{t+h_2-i}^3)] \leq E_t[(b_1^{h_2} I_t^{(1)} + b_1^{h_2-1} I_t^{(2)}) (-\mathbf{y}_t) I(\Pi_{i=0}^{h_2} Q_{t+h_2-i}^3)] + c_{1.9}.$$

By repeating the substitution of each phase on the right hand side of (5.4), we take a sufficiently large $h_3 (\geq h_1 + h_2 + 2)$ and we can reduce the second and the third terms of (5.4) to their first and last terms. By using the condition $0 \leq b_1 < 1$ we can find positive constants $c_{1.10}$ and $\delta_{1.4}$ ($0 \leq \delta_{1.4} < 1$) such that $E_t[\|\mathbf{y}_{t+h_3}\|] \leq \delta_{1.4} \|\mathbf{y}_t\| + c_{1.10}$.

[4] C_3 : In this case we notice that when $\mathbf{y}_{t+3} \in Q_{t+3}^1$ and $\mathbf{y}_{t+2} \in Q_{t+2}^1$, the equation $\mathbf{y}_{t+3} = a_1 \mathbf{y}_{t+2} + a_2 \mathbf{y}_{t+1} + v_{t+3}$ implies that $v_{t+3} > (-a_1) \mathbf{y}_{t+2} + (-a_2) \mathbf{y}_{t+1} \geq 0$. Then by applying the same argument as [2] and [3] to $\mathbf{y}_{t+1}$, we can show that $E_t[\|\mathbf{y}_{t+3}\| I(Q_{t+3}^1 Q_{t+2}^1 Q_{t+1}^1)]$ is bounded. By using the facts that we can decompose

$$I(Q_{t+3}^2 Q_{t+2}^1 Q_{t+1}^1) = I(Q_{t+3}^2 Q_{t+2}^1 Q_{t+1}^1 Q_t^1) + I(Q_{t+3}^2 Q_{t+2}^1 Q_{t+1}^1 Q_t^4),$$

and the conditional expectations $\mathbf{y}_{t+2}$ and $\mathbf{y}_{t+1}$ are bounded for the phase $Q_{t+1}^1 Q_t^1$, we find that $E_t[\|\mathbf{y}_{t+2}\| I(Q_{t+2}^1 Q_{t+1}^1)]$ is bounded and hence $E_t[\|\mathbf{y}_{t+3}\| I(Q_{t+3}^2 Q_{t+2}^1 Q_{t+1}^1)]$ is also bounded. Also by using that $0 < \mathbf{y}_{t+1} = a_1 \mathbf{y}_t + a_2 \mathbf{y}_{t-1} + v_{t+1} < v_{t+1}$ for the phase $Q_{t+3}^2 Q_{t+2}^1 Q_{t+1}^1 Q_t^1$, we have that $E_t[\mathbf{y}_{t+1} I(Q_{t+1}^1 Q_t^1)]$ is bounded. Then $E_t[\|\mathbf{y}_{t+2}\| I(Q_{t+2}^1 Q_{t+1}^1)]$ and $E_t[\|\mathbf{y}_{t+3}\| I(Q_{t+3}^3 Q_{t+2}^2 Q_{t+1}^1 Q_t^1)]$ are bounded. For the term for $I(Q_{t+3}^3 Q_{t+2}^3)$, we can take a positive constant $c_{1.11}$ such that

$$E_t[\|\mathbf{y}_{t+3}\| I(Q_{t+3}^3 Q_{t+2}^3)] \leq E_t[(b_1^3 I_t^{(1)} + b_1 I_t^{(2)}) (-\mathbf{y}_t) I(Q_{t+3}^3 Q_{t+2}^3)] + c_{1.11}.$$

Because $0 < b_1 < 1$, we can find positive constants $c_{1.12}$ and $\delta_{1.5}$ ($0 \leq \delta_{1.5} < 1$) such that $E_t[\|y_{t+3}\|] \leq \delta_{1.5}\|y_t\| + c_{1.12}$.

[5] C_4 : We need to consider the terms involving the phases $Q_{t+h}^1 Q_{t+h-1}^1$ and $Q_{t+h}^2 Q_{t+h-1}^1$ ($h \geq 2$) in particular. Since $a_2 < 0 \leq a_1$ in C_4 , there are two different cases depending on whether the characteristic roots of $g_A(\lambda) = 0$ in (3.4) are real or complex, separately.

When $D(A) \geq 0$ and $0 \leq a_1 + a_2 < 1$, the characteristic roots are real and their absolute values are less than one. In this case we immediately confirm the conditions that $0 \leq e_1' A l < 1$ and $e_1' A^2 l < 1$. Then we have the same inequality as in [2]. When $D(A) < 0$, on the other hand, there exists a positive integer h such that (5.6) holds because the angle θ of two roots is in $(0, \pi)$ or $(\pi, 2\pi)$. Then we can reduce the problem into the one with $Q_{t+h}^3 Q_{t+h-1}^3$ in [3]. Hence we find positive constants $\delta_{1.6}$ ($0 \leq \delta_{1.6} < 1$) and $c_{1.13}$ such that the last inequality holds instead of $\delta_{1.4}$ and $c_{1.10}$ in [3].

For the remaining term involving the phase $Q_{t+h}^3 Q_{t+h-1}^3$, we can use the same argument as [3] for C_2 because of the condition $0 < b_1 < 1$.

[6] *Necessity*: For proving the necessity of our conditions, we modify the similar arguments used by Petrucci and Woolford (1984). Because they become quite lengthy and tedious, we consider the case when $a_i > 0$ ($i = 1, 2$) and $a_1 + a_2 > 1$ as an illustration. In this case we have two real characteristic roots λ_i ($i = 1, 2$), which satisfy the condition $\lambda_1 > 1 > 0 > \lambda_2$. Then we have two cases depending on the relative magnitudes of two roots, that is, (i) $|\lambda_1| > |\lambda_2|$ and (ii) $|\lambda_1| < |\lambda_2|$. By defining a nonsingular 2×2 matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}$$

and transforming the original system as 2×1 vector $x_t = (x_{1t}, x_{2t})' = \Lambda^{-1} y_t$. If we write $U_t = (u_{1t}, u_{2t})' = \Lambda^{-1}(\sigma(t)v_t, 0)'$ with $\sigma(t) = \sigma_1 I(y_{t-1} > 0) + \sigma_2 I(y_{t-1} \leq 0)$, we have a new representation of the process $x_{it} = \lambda_i x_{i,t-1} + u_{it}$ ($i = 1, 2$). For the case of (i), we consider $\{x_{1t}\}$ and it is not difficult to show that its solution explodes with a positive probability by the same method used of Petrucci and Woodford (1984). For the case of (ii), we consider $\{x_{2t}\}$ and it is also not difficult to show that its solution explodes with a positive probability.

REFERENCES

- Brockwell, P. and R. Davis (1991), *Time Series : Theory and Methods*, 2nd Edition, Springer.
- Chan, K.S. and H. Tong (1985), "On the Use of the Deterministic Lyapunov Function for the Ergodicity of Stochastic Difference Equation," *Adv. Appl. Prob.*, 17, 666-678.
- Chen, R. and R.S. Tsay (1991), "On the Ergodicity of TAR(1) Processes," *Annals of Applied Probability*, 1, 613-634.
- Kunitomo, N. and S. Sato (1996), "Asymmetry in Economic Time Series and Simultaneous Switching Autoregressive Model," *Structural Change and Economic Dynamics*, 7,

- Kunitomo, N. and S. Sato (2000), "On Simultaneous Switching Autoregressive Models," in *Nonlinear Statistical Inference*, edited by C.Hsiao, Cambridge University Press, forthcoming.
- Nummelin, E. (1984), "General Irreducible Markov Chains and Non-negative Operators," *Cambridge University Press*.
- Petrucelli, J. and S.W. Woolford (1984), "A Threshold AR(1) Model," *Journal of Applied Probability*, 17, 666-678.
- Sato, S. and N. Kunitomo (1996), "Some Properties of the Maximum Likelihood Estimator in the Simultaneous Switching Autoregressive Model," *Journal of Time Series Analysis*, 17, 287-307.
- Tjøstheim, D. (1990), "Nonlinear Time Series and Markov Chains," *Adv. Appl. Prob.*, 22, 587-611.
- Tweedie, R.L. (1975), "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space," *Stoch. Proc. Appl. Prob.*, 3, 385-403.
- Tong, H. (1990), *Non-linear Time Series*, Oxford University Press, Oxford.

Appendix : Some Figures

In this appendix, we give several figures. All figures reported here are the results of the simulations on the sets of 10,000 realizations of $\{y_t\}$ based on the TAR(2) processes and the SSAR(2) processes without any disturbance terms. We also have checked the non-explosiveness of the sample paths of the solutions and basically confirmed the adequacy of the same regions by the corresponding simulations for the TAR(2) and SSAR(2) processes with disturbances. It was all we could do because the criteria of convergence in simulations are more difficult and subtle when there are noise terms.

All figures for the TAR(2:d) processes are denoted by TAR(2) with the delayed parameter d and drawn in the (a_1, a_2) space while the figure for the SSAR(2) process are drawn in the (r_1, r_2) -space. The shaded areas in figures are the non-explosive regions in our simulations.

Appendix : Some Figures

TAR(2) : $b_1 = 0.6$, $b_2 = 0$, $d = 1$

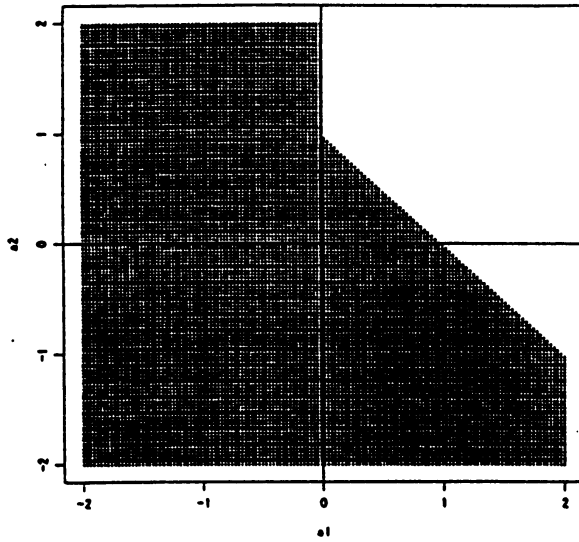


Figure 1

TAR(2) : $b_1 = -0.6$, $b_2 = 0$, $d = 1$

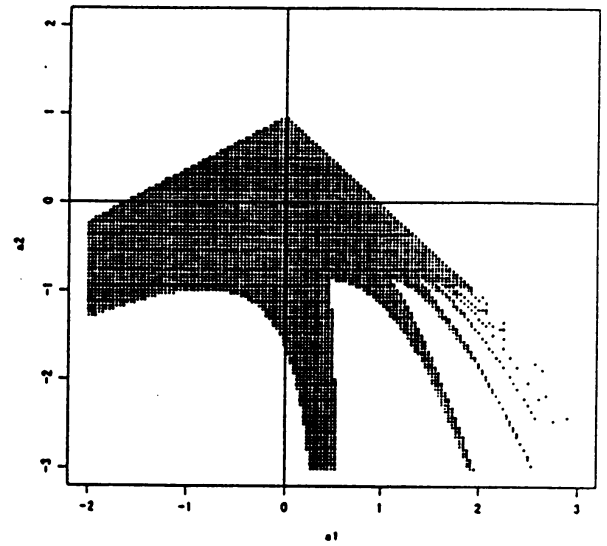


Figure 2

TAR(2) : $b_1 = -0.6$, $b_2 = 0$, $d = 2$

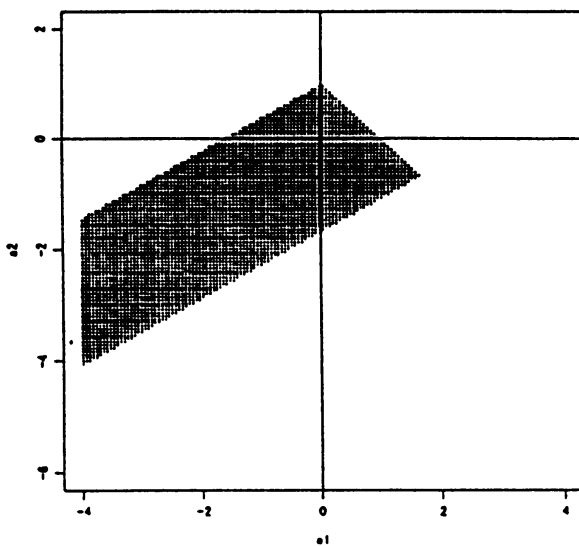


Figure 3

TAR(2) : $b_1 = 0.2$, $b_2 = 0$, $d = 2$

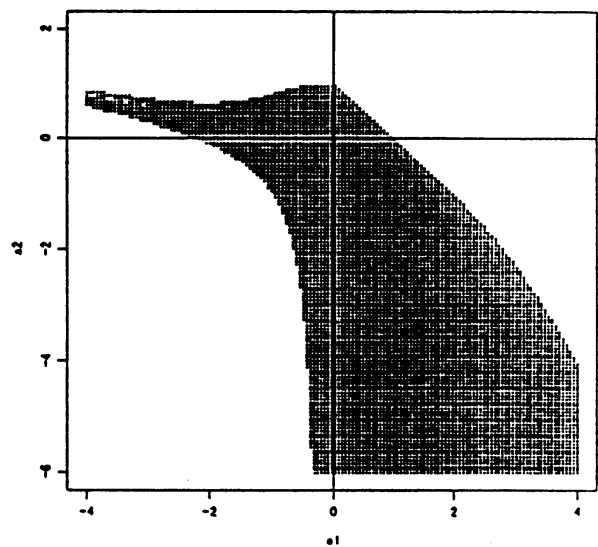


Figure 4

Appendix : Some Figures (Continued)

SSAR(2) : sig1= 1 . sig2= 0.1

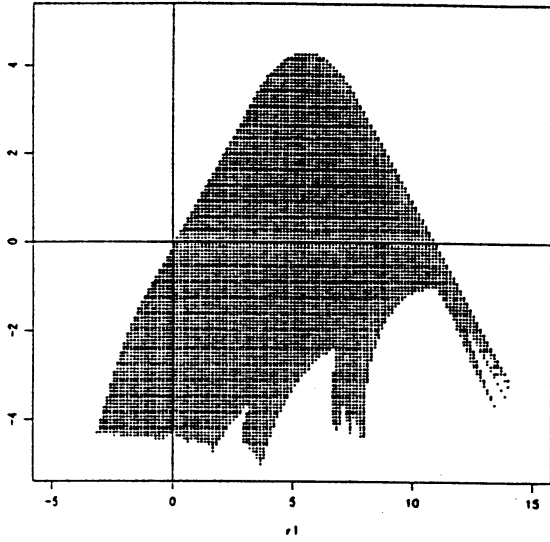


Figure 5

TAR(2) : b1= 0.2 , b2= -0.5 , d= 1

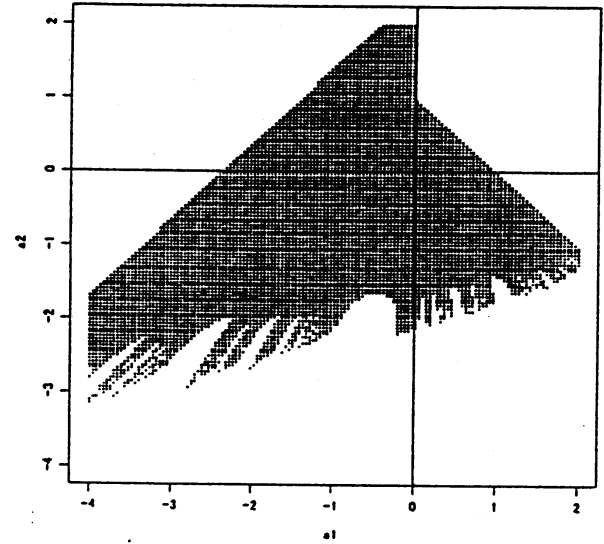


Figure 6

TAR(2) : b1= -0.2 , b2= 0.1 , d= 1

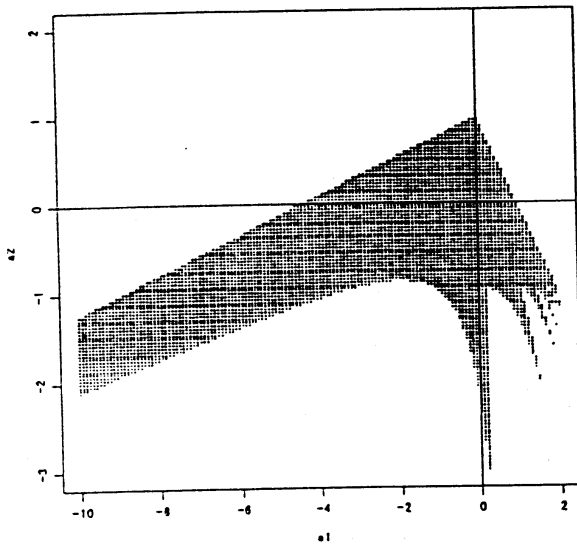


Figure 7