The Korean Communications
in Statistics Vol. 8, No. 2, 2001
pp. 565-572

Ergodicity of Nonlinear Autoregression with Nonlinear ARCH Innovations[1]

S.Y. Hwang[2] and I.V. Basawa[3]

Abstract

This article explores the problem of ergodicity for the nonlinear autoregressive processes with ARCH structures in a very general setting. A sufficient condition for the geometric ergodicity of the model is developed along the lines of Feigin and Tweedie(1985), thereby extending classical results for specific nonlinear time series. The condition suggested is in turn applied to some specific nonlinear time series illustrating that our results extend those in the literature.

Keywords : Autoregression, Ergodicity, Nonlinear ARCH.

1. Introduction

In last three decades, a large amount of research has been directed to nonlinear time series models. This is because these models are known to explain nonlinear features such as asymmetry, jump phenomenon and limit cycle which would not have been accounted for by linear ARMA time series. We refer to Tong(1990) for a comprehensive treatment of nonlinear models as alternatives to ARMA processes.

The p-th order nonlinear autoregressions are defined by the difference equation

$$y_t = F_\theta(y_{t-1}, \cdots, y_{t-p}) + \varepsilon_t \qquad (1.1)$$

where $\{\varepsilon_t\}$ is *iid* sequence of innovations and θ denotes a parameter vector of appropriate order. (1.1) embodies threshold autoregression and exponential autoregressive process as special cases.

This article postulates the innovations $\{\varepsilon_t\}$ evolving with nonlinear conditional

1) This work was supported by research grant 2001 from Sookmyung University.
2) Associate professor, Dept. of Statistics, Sookmyung Women's University, Seoul, 140-742 Korea.
E-mail : shwang@sookmyung.ac.kr
3) Dept. of Statistics, University of Georgia, Athens, GA. 30602 USA.

heteroscedastic patterns. Specifically $\{\varepsilon_t\}$ is assumed to follow the nonlinear ARCH(m) process, viz.,

$$\varepsilon_t = \sqrt{h_t} \cdot e_t \tag{1.2}$$

$$h_t = \alpha_0 + G_\alpha(\varepsilon_{t-1}, \cdots, \varepsilon_{t-m}) \tag{1.3}$$

where $\{e_t\}$ is *iid* sequence of random variables with zero mean and unit variance, $G_\alpha(\cdot)$ is a possibly nonlinear function of the innovations $\varepsilon_{t-1}, \cdots, \varepsilon_{t-m}$ and $G_\alpha(\cdot) \geq 0$ with α being a vector of parameters indexing the conditional variance h_t. Accordingly, the model defined by (1.1) through (1.3) can be viewed as a fairly general nonlinear autoregression with nonlinear ARCH errors.

Statistical analysis of the time series models tacitly assumes that the model is ergodic and hence is strictly stationary, from which one can establish the central limit convergence and the ergodic theorem, if applicable. The ergodicity of (1.1) has been examined under various sets of conditions by many authors including Feigin and Tweedie(1985), Chan and Tong(1985); more recently, Lee(1996), An et al.(1997) and Lee(1998). However their methods can not be directly applicable to the general models specified by (1.1) to (1.3). Our objective in the present paper is to develop a set of sufficient conditions under very broad setting. We will tackle the questions of ergodicity for the model by modifying the general conditions for the Markovian time series mainly developed by Feigin and Tweedie(1985), An et al.(1997) and Lee(1998).

The rest of the paper proceeds as follows. Section 2 obtains the geometric ergodicity of the model with brief summaries of the standard terminologies appearing in Markov processes for the quick reference. Section 3 illustrates applications of the main results regarding ergodicity to the specific nonlinear time series examples including two relatively new models in time series literature.

2. Geometric ergodicity of the model

We begin with the formal definition of the term "ergodicity" for a Markov process. Let $\{X_t, t \geq 0\}$ be a first order Markov process taking values in R^k, the k-dimensional Euclidian space. For notation, $P^{(t)}(x, A)$ denotes the homogeneous t-step transition probability of the process, i.e., for A in the Borel-σ-field on R^k

$$P^{(t)}(x, A) = P(X_t \in A \mid X_0 = x) \tag{2.1}$$

$\{X_t\}$ is then called ergodic if there exists a probability measure Π such that for every $x \in R^k$

$$\|P^{(t)}(x, A) - \Pi(A)\|_{TV} \rightarrow 0, \quad as\ t \rightarrow \infty \tag{2.2}$$

where $\|\cdot\|_{TV}$ stands for the total variation norm. Moreover, if the convergence in (2.2) is exponentially decaying (to zero), viz., for some $0 < \rho < 1$ and for every $x \in R^k$

$$\|P^{(t)}(x, A) - \Pi(A)\|_{TV} = o(\rho^t) \tag{2.3}$$

then $\{X_t\}$ is called geometrically ergodic. The measure $\Pi(\cdot)$ in (2.2) or (2.3) is usually referred to as the stationary distribution for $\{X_t\}$.

Returning to the time series $\{y_t\}$ specified by (1.1) to (1.3), observe first that $\{y_t\}$ forms a $p + m = k$, say, order Markov process. To obtain the first order representation, introduce $k \times 1$ vectors X_t, M and V defined as follows.

$$X_t = (y_t, y_{t-1}, \cdots, y_{t-k+1})'$$
$$M(X_{t-1}) = (F_\theta(y_{t-1}), y_{t-1}, y_{t-2}, \cdots, y_{t-k+1})'$$

and

$$V(X_{t-1}) = (h_t^{1/2}, 0, 0, \cdots 0)'$$

The model $\{y_t\}$ can then be written in terms of a first order $k \times 1$ vector Markov process, namely,

$$X_t = M(X_{t-1}) + V(X_{t-1}) \cdot e_t \tag{2.4}$$

where e_t is independent of X_{t-1}.

In order to establish the geometric ergodicity of $\{X_t\}$, and hence of $\{y_t\}$, we impose the following conditions.

$(C0)$ $\{e_t\}$ has a probability density with support $(-\infty, \infty)$.

$(C1)$ There exist nonnegative real valued functions $f(\theta)$ and $g(\theta)$ such that, for given $\Delta > 0$, there exists $\gamma > 0$ (γ may depend on θ, α and Δ) satisfying for all $\|\cdot\| > \gamma$

$$|F_\theta(\,\cdot\,)| \leq f(\theta)\|\cdot\| + \Delta$$
$$|G_a(\,\cdot\,)| \leq g(a)^2\|\cdot\|^2 + \Delta$$

where $\|\cdot\|$ is used to denote the standard Euclidian norm defined in appropriate dimension.

Theorem 1 : Suppose that $f(\theta) + g(a) < 1$. Then $\{y_t\}$ is geometrically ergodic.

Proof : We shall verify the conditions in Theorem 1 of Feigin and Tweedie(1985).
First, it follows from $(C\,0)$ that $\{X_t\}$ is ϕ-irreducible with ϕ being the Lebesgue measure on R^k and that $\{X_t\}$ is clearly a Feller chain (See, An et al(1997) for relevant discussions). It now remains to verify the condition (ii) in works from Feigin and Tweedie(1985). To be more precise, it needs to be shown that for some $\delta > 0$, there exists a non-negative continuous real valued function $\Psi : R^k \rightarrow R$ such that

$$E[\,\Psi(X_t)|X_{t-1} = x_{t-1}] \leq (1-\delta)\,\Psi(x_{t-1}) \tag{2.5}$$

for all sufficiently large $\|x_{t-1}\|$. For (2.5) we proceed as follows. Define the set C_1 of all x_{t-1} sufficiently large

$$C_1 = \{x_{t-1} = (y_{t-1}, \cdots, y_{t-k})'\ ;\ |y_{t-i}| > \gamma,\ \ i = 1, \cdots, k\}$$

It then follows from $(C\,1)$ that on C_1

$$|y_t| \leq \Delta + f(\theta)\ \|X_{t-1}\| + \sqrt{h_t}\cdot e_t \tag{2.6}$$

Moreover, it is seen that on C_1

$$\sqrt{h_t} \leq g(a)\ \|X_{t-1}\| + c \tag{2.7}$$

with $c = (a_0 + \Delta)^{1/2}$.

Collecting (2.6) and (2.7) and using $E|e_t| \leq 1$, some tedious but straightforward manipulation yields that

$$E(y_t|X_{t-1} = x_{t-1}) \leq \{f(\theta) + g(a)\}\ \|x_{t-1}\| + c + \Delta \tag{2.8}$$

Our immediate goal is to choose Ψ satisfying (2.5). Construct the following non-negative continuous real valued function $\Psi : R^k \rightarrow R$

$$\Psi(u_1, \cdots, u_k) = 1 + \| (u_1, \cdots, u_k) \|_\infty \tag{2.9}$$

where $\| \cdot \|_\infty$ is a "modified" infinite norm defined as

$$\| (u_1, \cdots, u_k) \|_\infty = \max \{ \beta_1 |u_1|, \cdots, \beta_k |u_k| \} \tag{2.10}$$

Here β's are strictly positive constants with constraints which will be specified in (2.12) in the course of the proof. It is to be noted that e_t is independent of X_{t-1} and $E|e_t| \le 1$ from which it can be deduced using (2.8) that

$$\begin{aligned} E\ [\ \Psi(X_t) | X_{t-1} = x_{t-1}] \\ &= 1 + E[\ \max \{ \beta_1 |y_t|, \beta_2 |y_{t-1}|, \cdots, \beta_k |y_{t-k+1}| \} \ |X_{t-1} = x_{t-1}] \\ &\le 1 + \max \{ \beta_1 (f(\theta) + g(\alpha))\ \|x_{t-1}\| + \beta_1 (c + \Delta), \beta_2 |y_{t-1}|, \cdots, \beta_k |y_{t-k+1}| \} \end{aligned} \tag{2.11}$$

Exploiting $f(\theta) + g(\alpha) < 1$, pick up $\delta = 1 - f(\theta) - g(\alpha) > 0$. The RHS in (2.11) then attain an upper bound $(1-\delta)[1 + \max \{ \beta_1 |y_{t-1}|, \cdots, \beta_k |y_{t-k}| \}] = (1-\delta) \Psi(x_{t-1})$ for all sufficiently large $\|x_{t-1}\|_\infty$ (call this region C_2) provided that

$$0 < \beta_k < \beta_{k-1} < \cdots < \beta_1$$

and

$$\beta_1 (1-\delta) < \beta_k \tag{2.12}$$

Consequently, the assertion (2.5) holds for all $x_{t-1} \in C_1 \cap C_2$, as desired, completing the proof.

3. Examples

This section discusses some specific nonlinear models where the general conditions in Section 2 reduce to the explicit formulation. To avoid repetition below, $\{e_t\}$, in this section, stands for the arbitrary sequence of *iid* random quantities with zero mean and variance unity. We give the main argument and results, only omitting some details. Comparisons and analogies to classical results are also made. Examples are including various nonlinear ARCH models which are interesting in their own rights.

Ex 3.1. Nonlinear autoregression. The difference equation for the classical p-order nonlinear autoregression (NAR(p), hereafter) is given by

$$y_t = F_\theta(y_{t-1}, \cdots, y_{t-p}) + \varepsilon_t$$

with $\{\varepsilon_t\}$ being a sequence of *iid* random errors. Taking $G_\alpha(\cdot) \equiv 0$ in our model yields NAR(p). The condition (C1) is then equivalent to : For given $\Delta > 0$, there exists nonnegative real valued function $f(\theta)$, $\gamma > 0$ such that $f(\theta) < 1$ and $|F_\theta(\cdot)| \le f(\theta) \|\cdot\| + \Delta$ for all $\|\cdot\| > \gamma$. This condition implies that once $(y_{t-1}, \cdots, y_{t-p})$ drifts to "abnormally" large values, y_t is forced to bounce back to "normal" level. Refer to Tong(1990, Ch.4) for similar discussions.

Ex 3.2. β-ARCH model. Consider the following first order β-ARCH processes first discussed by Guegen and Diebolt(1994) and later studied by Hili(1999).

$$y_t = \theta y_{t-1} + (\alpha_0 + \alpha^2 y_{t-1}^{2\beta})^{1/2} e_t, \quad |\theta| < 1 \tag{3.1}$$

where $0 \le \beta \le 1$. Note that $\beta = 1$ gives standard ARCH(1) processes of Engle(1982). One may choose $f(\theta) = |\theta|$ and $g(\alpha) = |\alpha|$ and thus the condition is simply $|\theta| + |\alpha| < 1$, which agrees with the condition imposed by Guegen and Diebolt(1994). An et al.(1997) investigated the m-th order β-ARCH processes,

$$y_t = \{\alpha_0 + \alpha_1^2(y_{t-1})^{2\beta} + \cdots + \alpha_m^2(y_{t-m})^{2\beta}\} \cdot e_t \tag{3.2}$$

Setting $F_\theta \equiv 0$ and taking

$$G_\alpha(\cdot) = \alpha_1^2 \varepsilon_{t-1}^{2\beta} + \cdots + \alpha_m^2 \varepsilon_{t-m}^{2\beta}$$

in (1.1) to (1.3) yield (3.2). By identifying $f(\theta) = 0$ and $g(\alpha) = \max\{|\alpha_1|, \cdots, |\alpha_m|\}$ the ergodicity condition reads

$$\max\{|\alpha_1|, \cdots, |\alpha_m|\} < 1 \tag{3.3}$$

which is an improvement compared to the earlier results $\sum_{i=1}^{m} \alpha_i^2 < 1$ for the case when $\beta = 1$, obtained by An et al.(1997).

Ex 3.3. Threshold ARCH processes. We now consider the following threshold ARCH(p) processes.

$$\begin{aligned} y_t &= \sum_{i=1}^{p} (\theta_{i1} y_{t-1}^{+} + \theta_{i2} y_{t-1}^{-}) + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} \cdot e_t \\ h_t &= \alpha_0 + \sum_{j=1}^{m} [\alpha_{j1}^2 (\varepsilon_{t-j}^{+})^2 + \alpha_{j2}^2 (\varepsilon_{t-j}^{-})^2] \end{aligned} \tag{3.4}$$

where $a^{\pm}$ denotes $aI_{[a \geq 0]}$ and $aI_{[a < 0]}$ respectively. Note that (3.4) introduces "thresholds" both in the conditional mean and in the conditional variance as well. An obvious choice for $f(\theta)$ and $g(\alpha)$ is

$$f(\theta) = \max{}_{1 \leq i \leq p}\{ \max (|\theta_{i1}|, |\theta_{i2}|)\}$$

and

$$g(\alpha) = \max{}_{1 \leq i \leq m}\{ \max (|\alpha_{j1}|, |\alpha_{j2}|)\}$$

For the special case when $p=1$ and $\alpha_{j1} = \alpha_{j2} = 0$, $1 \leq j \leq m$, threshold ARCH models reduce to the standard threshold AR(1) processes. Ergodicity condition is then equivalent to $\max\{|\theta_{11}|, |\theta_{12}|\} < 1$ which agrees with the condition as in Brockwell et al.(1992). Refer to Hwang and Woo (2001) for the first order model where $p=1$ and $m=1$. Our last example delivers a bounded version of the ARCH processes.

Ex 3.4. Bounded ARCH processes. Set $m=1$ without loss of generality and, consider

$$\begin{aligned} y_t &= F_\theta(y_{t-1}, \cdots, y_{t-p}) + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} \cdot e_t \end{aligned}$$

and take

$$h_t = \alpha_0 + \alpha_1^2 \eta(\varepsilon_{t-1}) \tag{3.5}$$

where

$$\eta(\varepsilon_{t-1}) = \varepsilon_{t-1}^2 I_{[|\varepsilon_{t-1}| \leq d]} + d^2 I_{[|\varepsilon_{t-1}| > d]} \tag{3.6}$$

with d being a prescribed constant. (3.6) is fully motivated by Huber function in the context of M-estimation. It may be noted that the usual ARCH(1) structure can be obtained by releasing d infinity. Due to the boundness of the ARCH structure (which is reasonably the case in practice), the ergodicity condition is rather simple for this example. It is easily seen that one needs only to impose the same conditions as for NAR(p) discussed in Ex 3.1. for the geometric ergodicity of $\{y_t\}$.

Acknowledgements

The authors are grateful to two referees and financial support from 2001 Sookmyung Women's University research grant.

References

[1] An, H.Z., Chen, M. and Huang, F.C.(1997). The geometric ergodicity and existence of moments for a class of non-linear time series models. *Statistics & Probability Letters*, 31, 213-224.

[2] Brockwell, P.J., Liu, J. and Tweedie, R.L.(1992). On the existence of stationary threshold autoregressive moving-average processes. *Journal of Time Series Analysis*, 13, 95-107.

[3] Chan, K.S. and Tong, H.(1985). On the use of the deterministic Lyapunov function for the ergodicity of the stochastic difference equation. *Adv. Appl. Prob.*, 17, 666-678.

[4] Engle, R.F.(1982). Autoregressive conditional heteroscedasticty with estimates of the variance of U.K. inflation, *Econometrica*, 50, 987-1008.

[5] Feigin, P. D. and Tweedie, R. L.(1985). Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis*, 6, 1-14.

[6] Guegen, D. and Diebolt, J.(1994). Probabilistic properties of the β-ARCH model. *Statistica Sinica*, 4, 71-87.

[7] Hili, O.(1999). On the estimation of the β-ARCH models. *Statistics & Probability Letters*, 45, 285-293.

[8] Hwang, S. Y. and Woo, M. J. (2001). Threshold ARCH(1) processes : asymptotic inference. *Statistics & Probability Letters*, 53, 11-20.

[9] Lee, C.(1998). Asymptotics of a class of p-order nonlinear autoregressive processes. *Statistics & Probability Letters*, 40, 171-177.

[10] Lee, O.(1996). Conditions for the non-ergodicity of some Markov chains. *Journal of the Korean Statistical Society*, 25, 303-312.

[11] Tong, H.(1990). *Nonlinear Time Series*. Oxford University Press, Oxford.