Statistica Sinica **8**(1998), 1205-1217

ON THE GEOMETRIC ERGODICITY OF A NON-LINEAR AUTOREGRESSIVE MODEL WITH AN AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTIC TERM

Zudi Lu

Academia Sinica, Beijing

Abstract: In this paper, the geometric ergodicity of a non-linear AR model with an ARCH term is discussed. Two non-vacuous and mild sufficient conditions are given. The results obtained modify the vacuous part and reduce the restriction of Masry and Tjϕstheim (1995)'s conditions, and lay a foundation for statistical inference of the model (e.g. Mckeague and Zhang (1994) and Masry and Tjϕstheim (1995)). It is worth pointing out that the geometric ergodicity of the general β–ARCH(p) model which could not be solved in Guegan and Diebolt (1994) may be easily derived from our results. Compared with Nze (1992), the conditions of this paper may guarantee the existence of the second moments for the stationary solution. A conjecture is also given.

Key words and phrases: Autoregression, β–ARCH(p), conditional heteroscedasticity, geometric ergodicity, Markov chain, nonlinear AR model with ARCH term.

1. Introduction

Consider the non-linear autoregressive (AR) model with autoregressive conditional heteroscedastic (ARCH) term:

$$X_t = f(X_{t-1},\dots,X_{t-p}) + \varepsilon_t[h(X_{t-1},\dots,X_{t-q})]^{1/2}, \tag{1.1}$$

where $f : R^p \to R^1$ is a Borel measurable function on R^p, $h : R^q \to R^1$ is a positive Borel measurable function on R^q, $\{X_t\}$ and $\{\varepsilon_t\}$ are two sequences of random variables defined on the same probability space $(\Omega,\mathcal{F},\mathcal{P})$ and $\{\varepsilon_t\}$ consists of i.i.d. random variables with $E\varepsilon_t = 0$ and $E\varepsilon_t^2 = 1$ for which ε_t is independent of $\{X_s,\ s < t\}$. This model is quite general; for example, if $h \equiv \sigma^2$ (positive constant) (1.1) is the usual non-linear autoregressive model (Tong (1990)), and if $f \equiv 0$, (1.1) is the general ARCH model first proposed by Engle (1982). This model includes two terms in which f and h are the conditional mean and conditional variance, respectively, of X_t given the past information $\{X_s,\ s < t\}$, and thus has the advantages of both the AR model and the ARCH model: the conditional mean f describes the prediction value and the conditional variance h measures the risk of this prediction based on past information.

Recently, model (1.1) has received much attention in the literature. On the nonparametric statistical inference of the functions f and h, Auestad and Tjϕstheim (1990), Mckeague and Zhang (1994), Masry and Tjϕstheim (1995) and the references therein contain some detailed discussions for different nonparametric methods. In those publications, the asymptotic properties, such as (strong) consistency and asymptotic normality, of the nonparametric estimators of the functions f and h have been studied based on the geometric ergodicity (see (2.2) for the definition) of model (1.1) from which the β-mixing (hence α-mixing) can be deduced. By Mckeague and Zhang (1994) and Masry and Tjϕstheim (1995), it is easily seen that the geometric ergodicity of model (1.1) is very important for statistical inference. In Mckeague and Zhang (1994), when $p = q = 1$, a sufficient condition for the geometric ergodicity is given; that is f and h are bounded on compact sets and there exists a constant C such that

$$\sup_{|x|>C} (|f(x)|/|x|) < 1 \quad \text{and} \quad \sup_{|x|>C} h(x) < \infty. \tag{1.2}$$

Under this condition, the conditional variance function h is bounded on the whole real line. Nze (1992) considered the geometric ergodicity of model (1.1) and gave a sufficient condition (Proposition 3), which, as one of the referees pointed out, may not guarantee the existence of the second moments. However, it is well known (e.g. Masry and Tjϕstheim (1995)) that the existence of the second moment for X_t is necessary for the nonparametric kernel estimation of h. In Masry and Tjϕstheim (1995), a mild condition (not only for the general p and q but also for the unbounded h) is given, which we state as follows:

Set $r = \max(p, q)$. For any $y = (y_1, \ldots, y_r)' \in R^r$, let $\tilde{f}(y) = f(y_1, \ldots, y_p)$, $\tilde{h}(y) = h(y_1, \ldots, y_q)$; then (1.1) can be expressed as

$$X_t = \tilde{f}(X_{t-1}, \ldots, X_{t-r}) + \varepsilon_t[\tilde{h}(X_{t-1}, \ldots, X_{t-r})]^{1/2}.$$

Hence, without loss of generality, assume that $p = q = r$ in (1.1).

Assumption 0. $(p = q)$

(a) The i.i.d. random variables $\{\varepsilon_t\}$ have a probability density function which is positive over R^1.

(b) The functions f and h are non-periodic, bounded on compact sets, and $h(y) > 0$ for all $y \in R^p$.

(c) There exist vectors $a = (a_1, \ldots, a_p)' \in R^p$ and $d = (d_1, \ldots, d_q)' \in R^q$ (which may both be zero) with $d_i \geq 0, i = 1, \ldots, q$ such that, as $\|y\|$ (Euclidean norm) $\to \infty$,

$$f(y) = \sum_1^p a_i y_i + o(\|y\|) \quad \text{and} \quad h(y) = \sum_1^q d_i y_i^2 + o(\|y\|^2). \tag{1.3}$$

Moreover, the p-dimensional square matrix A defined by $\mathbf{0}$ if $a = 0$, and by

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{1.4}$$

otherwise, satisfies $\rho(A'A) + \max_i(d_i) < 1$ if $d \neq 0$, and $\rho(A) < 1$ if $d = 0$. Here $\rho(\cdot)$ denotes the spectral radius.

Theorem 0. (Lemma 3.1 of Masry and Tjϕstheim (1995)) *Under the above Assumption* 0*, the nonlinear autoregressive model with ARCH term* (1.1) *is strongly mixing* (*i.e.,* α*-mixing*) *with mixing coefficient* $\alpha(k) \sim e^{-\beta k}$ *for some* $\beta > 0$.

Remark 0. The description of this theorem differs from that of Lemma 3.1 of Masry and Tjϕstheim (1995), but the facts are the same. Also, in the proof of Masry and Tjϕstheim (1995), the geometric ergodicity of model (1.1) is obtained, in fact, under Assumption 0, from which the strong mixing is deduced.

Also, there are some other publications in the literature on the geometric ergodicity of model (1.1). When $h \equiv \sigma^2$ (positive constant), Chan and Tong (1985), Tjϕstheim (1990, 1994), An and Huang (1994) and the references therein gave many sufficient conditions; when $f \equiv 0$, Lu (1994, 1995) considered many ARCH type models and obtained two general theorems that cover many ARCH type models which have appeared in the econometric literature. In addition, when $p = q = 1$ and h is bounded away from zero and ∞, Bhattacharya and Lee (1995) conducted a detailed study of the geometric ergodicity for every case of the limits of $f(x)/x$ as $x \to \pm\infty$. Of course, these models considered are all special cases of model (1.1). Lemma 3.1 of Masry and Tjϕstheim (1995) and Proposition 3 of Nze (1992) seem quite general for the models considered.

However, as stated above, the conditions of Nze (1992) may not guarantee the existence of the second moments; it is also worth pointing out that in Theorem 0, if $d \neq 0$, the sufficient condition for the geometric ergodicity, that is $\rho(A'A) + \max_i(d_i) < 1$, is vacuous when $p = q > 1$ (but it is non-vacuous when $p = q = 1$). In fact, for example, if $p = q = 2$, the matrix A in (1.4) and AA' are

$$A = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad AA' = \begin{pmatrix} a_1^2 + a_2^2 & a_1 \\ a_1 & 1 \end{pmatrix}. \tag{1.5}$$

So the characteristic function of AA' is

$$c(x) = |xI_2 - AA'| = x^2 - (1 + a_1^2 + a_2^2)x + a_2^2. \tag{1.6}$$

If $\rho(AA') = \rho(A'A) < 1$, then it is obvious that $c(1) > 0$. This contradicts $c(1) = -a_1^2 \leq 0$ by (1.6). Hence for any $a = (a_1, a_2)' \in R^2$, $\rho(A'A) \geq 1$ and moreover, $\rho(A'A) + \max_i(d_i) \geq 1$ (since $d_i \geq 0$ in (1.3)). For general $p > 2$, the situations are similar. Hence, the condition of Theorem 0 is vacuous when $d \neq 0$ and $p > 1$.

In this paper, we will retain the Masry and Tjϕstheim's (1995) decompositions of the functions f and h in (1.3), but relax the restriction on the functions f and h in (b) of Assumption 0. First, we state the following assumption which is similar to the one in the above:

Assumption 1. ($p = q$ in (1.1))
(1) The i.i.d. random variables $\{\varepsilon_t\}$ have a probability density function ψ which is positive and lower-semicontinuous over R^1.
(2) The functions f and h are bounded on any bounded Borel measurable set of R^p, and $h(y)$ is also either (i) continuous and $h(y) > 0$ for all $y \in R^p$, or (ii) $h(y) \geq d_0 > 0$ for some positive constant d_0 and for all $y \in R^p$.
(3) The functions f and h have decompositions of the form (1.3).

In the following, we discuss the geometric ergodicity of model (1.1) under Assumption 1. The main tools are the drift criterion for the geometric ergodicity of Markov chains. In Section 2, the Markovian represention of model (1.1) is set up and the irreducibility, aperiodicity and the small set property of this Markov chain are proved under Assumption 1. The drift criterion is also stated. In Section 3, two main theorems will be given. Theorem 1 gives a non-vacuous sufficient condition for the geometric ergodicity and overcomes the difficulty in Theorem 0 when $d \neq 0$. Theorem 2 is similar to Theorem 0 for the case $d = 0$, but removes the restriction of the non-periodicity on the functions f and h and covers the case of the non-linear autoregressive model (Tjϕstheim (1994) and Theorem 1 of An and Huang (1994)). Applying Theorem 1 to the β–ARCH(p) model, the geometric ergodicity of this model easily follows for the general p and $0 \leq \beta \leq 1$, which could not be obtained for $p \geq 2$ in Guegan and Diebolt (1994). Finally, in Section 4, some further discussions are given and a conjecture is made on a sufficient condition for the geometric ergodicity of the model (1.1).

2. Markovian Representation and Preliminaries

In model (1.1) ($p = q$), let

$$\begin{aligned} Y_t &= (X_t, X_{t-1}, \ldots, X_{t-p+1})', \\ T_1(Y_{t-1}, \varepsilon_t) &= f(Y_{t-1}) + \varepsilon_t (h(Y_{t-1}))^{1/2}, \quad T_i(Y_{t-1}, \varepsilon_t) = X_{t-i+1}, i = 2, \ldots, p, \\ T(Y_{t-1}, \varepsilon_t) &= (T_1(Y_{t-1}, \varepsilon_t), T_2(Y_{t-1}, \varepsilon_t), \ldots, T_p(Y_{t-1}, \varepsilon_t))'. \end{aligned}$$

Then it can be expressed as

$$X_t = \gamma' Y_t, \tag{2.1a}$$

$$Y_t = T(Y_{t-1}, \varepsilon_t), \tag{2.1b}$$

where $\gamma = (1, 0, \ldots, 0)' \in R^p$. Since ε_t is independent of Y_{t-1}, it follows from (2.1b) that $\{Y_t\}$ is a Markov chain whose state space is R^p, the p-dimensional Euclidean space. The Markov process $\{Y_t\}$ is said to be geometrically ergodic, if there exist a probability measure π on $(R^p, \mathcal{B}^p)$, a positive constant $\rho < 1$ and a π-integrable non-negative measurable function $J(x)$ such that as $n \to \infty$

$$\|P^n(x, \cdot) - \pi(\cdot)\|_\tau \le \rho^n J(x), \; x \in R^p, \tag{2.2}$$

where $P^n(\cdot, \cdot)$ is the n-step transition probability of $\{Y_t\}$ and $\|\cdot\|_\tau$ denotes the total variation norm.

Under Assumption 1, we first prove that the Markov chain $\{Y_t\}$ is irreducible with respect to the Lebesgue measure μ_p of R^p and is aperiodic. (For the definitions of irreducibility and aperiodicity of a Markov chain, see Chan (1990) in Appendix of Tong (1990), 448-466.)

Lemma 1. *Under Assumption* 1*, the Markov chain* $\{Y_t\}$ *defined by* (2.1b) *is* μ_p*-irreducible and aperiodic.*

Proof. For any $A \in \mathcal{B}^p$ with $\mu_p(A) > 0$, $y = (y_1, \ldots, y_p)' \in R^p$, it is easily proved that

$$P^p(y, A) = \int_A \prod_{j=0}^{p-1} q_j(x, y)\mu_p(dx), \tag{2.3}$$

where

$$\begin{aligned} q_0(x, y) &= \psi\Big((x_p - f(y_1, \ldots, y_p))[h(y_1, \ldots, y_p)]^{-1/2}\Big) \times [h(y_1, \ldots, y_p)]^{-1/2}, \\ q_j(x, y) &= \psi\Big((x_{p-j} - f(x_{p-j+1}, \ldots, x_p, y_1, \ldots, y_{p-j})) \\ &\qquad\qquad \times [h(x_{p-j+1}, \ldots, x_p, y_1, \ldots, y_{p-j})]^{-1/2}\Big) \\ &\qquad \times [h(x_{p-j+1}, \ldots, x_p, y_1, \ldots, y_{p-j})]^{-1/2}, \qquad j = 1, \ldots, p-1 \\ \mu_p(dx) &= dx_1 dx_2 \cdots dx_p, \end{aligned}$$

which shows that $P^p(y, A) > 0$ for any $y \in R^p$ and $A \in \mathcal{B}^p$ with $\mu_p(A) > 0$. By the definition of irreducibility and aperiodicity, the proof of this Lemma is completed.

In the following, we prove that for the Markov chain $\{Y_t\}$, any bounded set K in $\mathcal{B}^p$ with a positive μ_p-measure is a small set. (For the definition of a small set, also see Chan (1990), in Appendix of Tong (1990), 448-466.)

Lemma 2. *If the conditions of Lemma* 1 *hold, then any bounded set* K *in* R^p *with positive Lebesgue measure is a small set of the Markov chain* $\{Y_t\}$*.*

Proof. By Proposition 2.11 of Nummelin (1984), if we can prove that for any $A \in \mathcal{B}^p$ with positive μ_p Lebesgue measure,

$$\inf_{y \in K} P^p(y, A) > 0, \tag{2.4}$$

then the lemma is proved.

In fact, if (2.4) was not true, then there would exist a set $A \in \mathcal{B}^p$ with $\mu_p(A) > 0$ for which

$$\inf_{y \in K} P^p(y, A) = 0. \tag{2.5}$$

By the definition of the infimum, there would exist a sequence $\{\overline{y}_n\} \subset K$, where $\overline{y}_n = (\overline{y}_{1,n}, \ldots, \overline{y}_{p,n})'$, such that $\lim_{n\to\infty} P^p(\overline{y}_n, A) = 0$. Hence, by Fatou's Lemma, it easily follows from (2.3) that

$$\int_A \underline{\lim}_{n\to\infty} \prod_{j=0}^{p-1} q_j(x, \overline{y}_n)\mu_p(dx) \leq \lim_{n\to\infty} P^p(\overline{y}_n, A) = 0. \tag{2.6}$$

Set $a(x) = \underline{\lim}_{n\to\infty} \prod_{j=0}^{p-1} q_j(x, \overline{y}_n)$. If it can be proved that

$$a(x) > 0 \quad \text{for any} \quad x \in R^p, \tag{2.7}$$

it will result in a contradiction to (2.6). The remainder is devoted to the proof of (2.7).

For any fixed $x \in R^p$, there is a subsequence $\{\overline{y}_{n_l}\}$ of $\{\overline{y}_n\}$ such that

$$a(x) = \lim_{l\to\infty} \prod_{j=0}^{p-1} q_j(x, \overline{y}_{n_l}). \tag{2.8}$$

Since $\{\overline{y}_{n_l}\} \subset \{\overline{y}_n\} \subset K$ (bounded) and f and h are bounded on K, there is a subsequence of $\{\overline{y}_{n_l}\}$, denoted still by $\{\overline{y}_{n_l}\}$, such that $\lim_{l\to\infty} \overline{y}_{n_l} = \overline{y}_0 = (\overline{y}_{1,0}, \ldots, \overline{y}_{p,0})'$,

$$\lim_{l\to\infty} f(x_{p-j+1}, \ldots, x_p, \overline{y}_{1,n_l}, \ldots, \overline{y}_{p-j,n_l}) = L_{1j},$$
$$\lim_{l\to\infty} h(x_{p-j+1}, \ldots, x_p, \overline{y}_{1,n_l}, \ldots, \overline{y}_{p-j,n_l}) = L_{2j}$$

exist. We consider two cases. In case 1, assume that 2(i) of Assumption 1 holds. Then it follows from (2.8) and the lower-semicontinuity of $\psi(t)$ that $|L_{1j}| < \infty$

and

$$
\begin{aligned}
a(x) &\geq \prod_{j=0}^{p-1} \underline{\lim}_{l\to\infty} q_j(x, \overline{y}_{n_l}) \\
&\geq \prod_{j=0}^{p-1} \psi((x_{p-j} - L_{1j})[h(x_{p-j+1}, \ldots, x_p, \overline{y}_{1,0}, \ldots, \overline{y}_{p-j,0})]^{-1/2}) \\
&\qquad \times [h(x_{p-j+1}, \ldots, x_p, \overline{y}_{1,0}, \ldots, \overline{y}_{p-j,0})]^{-1/2} \\
&> 0 \qquad (2.9)
\end{aligned}
$$

for any $x \in R^p$; and secondly assume that 2(ii) in Assumption 1 holds, then $|L_{1j}| < \infty$, $d_0 \leq L_{2j} < \infty$ and

$$
a(x) \geq \prod_{j=0}^{p-1} \underline{\lim}_{l\to\infty} q_j(x, \overline{y}_{n_l}) \geq \prod_{j=0}^{p-1} \psi((x_{p-j} - L_{1j}) L_{2j}^{-1/2}) / L_{2j}^{-1/2} > 0 \qquad (2.10)
$$

for any $x \in R^p$. The proof is completed.

Remark 1. If the continuity of $h(y)$ is not assumed, the assumption that $h(y) \geq d_0 > 0$ for all $y \in R^p$ (see 2) of Assumption (1) is essential to ensure that $L_{2j} > 0$ uniformly. Otherwise, L_{2j} may equal 0. To ensure that the last inequality of (2.10) holds, it is necessary to suppose that $\lim_{x\to\infty} x\psi(x) > 0$. However, the latter inequality is not satisfied if ψ is a normal density function. Usually, the assumption that $h(y) \geq d_0 > 0$ for all $y \in R^p$ holds (Engle (1982)). This is not pointed out in Lemma 3.1 of Masry and Tjϕstheim (1995). (cf. (b) of Assumption 0)

The following lemma is crucial in the proof of the main theorems (see the next section), which is usually called the drift criterion for geometric ergodicity.

Lemma 3. *Let $\{Y_t\}$ be aperiodic and irreducible. Suppose that there exists a small set C, a non-negative measurable function g, and constants $0 < r < 1$, $\gamma > 0$, and $B > 0$ such that*

$$
E(g(Y_{t+1})|Y_t = y) < rg(y) - \gamma, \quad y \notin C; \qquad (2.11)
$$

$$
E(g(Y_{t+1})|Y_t = y) < B, \qquad y \in C. \qquad (2.12)
$$

Then $\{Y_t\}$ is geometrically ergodic. If g is also bounded away from 0 and ∞ on C, then $E_\pi g(Y_t) < \infty$, where E_π means the expectation with respect to the limiting probability measure π in (2.2).

Proof. For the proof of this lemma, see Nummelin and Tuominen (1984), Theorem 3.1, Page 196 and Tweedie (1983), Theorem 3, Page 194.

3. Main Theorems

The main purpose of this section is to derive a set of sufficient conditions for the geometric ergodicity of model (1.1). First, we give the following theorem which is a non-vacuous sufficient condition and overcomes the difficulty encountered by Theorem 0 (Lemma 3.1 of Masry and Tjϕstheim (1995)) for the case $d \neq 0$.

Theorem 1. *For the Markov chain* $\{Y_t\}$ *in* (2.1b) *defined by* (1.1) ($p = q$), *if Assumption* 1 *holds and*

$$\Big(\sum_{i=1}^{p} |a_i|\Big)^2 + \sum_{i=1}^{q} d_i < 1, \tag{3.1}$$

then $\{Y_t\}$ *is geometrically ergodic* (*hence so is model* (1.1)) *and* $E_\pi(X_t^2) < \infty$, *where* π *is the limiting probability measure* (*see* (2.2)).

Proof. By Lemma 1, $\{Y_t\}$ is irreducible and aperiodic, so we focus on finding a non-negative test function $g(x)$ for which the conditions of Lemma 3 hold.

For $y = (y_1, \ldots, y_p)' \in R^p$, take the test function $g(y)$ in the form

$$g(y) = 1 + y_1^2 + b_1 y_2^2 + \cdots + b_{p-1} y_p^2, \tag{3.2}$$

where $b_i, i = 1, \ldots, p-1$, are positive constants to be specified later.

By (1.1) and (1.3), as $\|y\| \to \infty$,

$$\begin{aligned}
E(g(Y_{t+1})|Y_t = y) &= 1 + E(X_{t+1}^2|Y_t = y) + b_1 y_1^2 + \cdots + b_{p-1} y_{p-1}^2 \\
&= 1 + f(y)^2 + h(y) + b_1 y_1^2 + \cdots + b_{p-1} y_{p-1}^2 \\
&= 1 + \sum_{i=1}^{p-1} (a_i^2 + d_i + b_i) y_i^2 + (a_p^2 + d_p) y_p \\
&\quad + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} a_i a_j y_i y_j + o(\|y\|) \sum_{i=1}^{p} a_i y_i + o(\|y\|^2) \\
&\leq 1 + \sum_{i=1}^{p-1} (a_i^2 + d_i + b_i) y_i^2 + (a_p^2 + d_p) y_p \\
&\quad + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} |a_i a_j| (y_i^2 + y_j^2) + o(\|y\|^2) \\
&= 1 + \Big(a_1^2 + d_1 + b_1 + |a_1| \sum_{j \neq 1} |a_j|\Big) y_1^2 \\
&\quad + \sum_{i=2}^{p-1} \Big[(a_i^2 + d_i + b_i + |a_i| \sum_{j \neq i} |a_j|)/b_{i-1}\Big] b_{i-1} y_i^2 \\
&\quad + \Big[(a_p^2 + d_p + |a_p| \sum_{j \neq p} |a_j|)/b_{p-1}\Big] b_{p-1} y_p^2 + o(\|y\|^2).
\end{aligned} \tag{3.3}$$

If we choose the positive constants b_i's to satisfy

$$\begin{aligned}
&a_1^2 + d_1 + b_1 + |a_1| \sum_{j\neq 1} |a_j| < 1, \\
&\Big(a_i^2 + d_i + b_i + |a_i| \sum_{j\neq i} |a_j|\Big)/b_{i-1} < 1, \qquad i = 2, 3, \ldots, p-1, \\
&\Big(a_p^2 + d_p + |a_p| \sum_{j\neq p} |a_j|\Big)/b_{p-1} < 1,
\end{aligned}$$

that is,

$$\begin{aligned}
a_p^2 + d_p + |a_p| \sum_{j\neq p} |a_j| < b_{p-1} < 1 - \sum_{i=1}^{p-1} \Big(a_i^2 + d_i + |a_i| \sum_{j\neq i} |a_j|\Big), \\
a_{i+1}^2 + d_{i+1} + b_{i+1} + |a_{i+1}| \sum_{j\neq i+1} |a_j| < b_i < 1 - \sum_{k=1}^{i} \Big(a_k^2 + d_k + |a_k| \sum_{j\neq k} |a_j|\Big), \\
i = 1, 2, \ldots, p-2,
\end{aligned}$$

(since (3.1) holds, such a choice of b_i's is possible) and set

$$\begin{aligned}
\rho = \max\Big\{&a_1^2 + d_1 + b_1 + |a_1| \sum_{j\neq 1} |a_j|, (a_i^2 + d_i + b_i + |a_i| \sum_{j\neq i} |a_j|)/b_{i-1}, i = 2, \ldots, p-1, \\
&(a_p^2 + d_p + |a_p| \sum_{j\neq p} |a_j|)/b_{p-1}\Big\};
\end{aligned}$$

then $0 < \rho < 1$,

$$\begin{aligned}
E(g(Y_{t+1}|Y_t = y) &\leq \rho(1 + y_1^2 + b_1 y_2^2 + \cdots + b_{p-1} y_p^2) + o(\|y\|^2) \\
&= (\rho + o(1))[1 + y_1^2 + b_1 y_2^2 + \cdots + b_{p-1} y_p^2] - 1, \qquad (3.4)
\end{aligned}$$

where $o(1) \to 0$ as $\|y\| \to \infty$. Choose $c > 0$ large enough, such that $\rho + o(1) < r_0 < 1$, as $\|y\| > c$. Set $C = \{\|y\| \leq c\}$. Then C is a bounded set with $\mu_p(C) > 0$. It follows from lemma 2 that C is a small set of $\{Y_t\}$. Then (2.11) and (2.12) obviously hold. Hence by Lemma 3, $\{Y_t\}$ is geometrically ergodic (so is $\{X_t\}$) and $E_\pi(X_t^2) < \infty$.

Example 1. Consider the following β-ARCH(p) model (see (4.1) of Guegan and Diebolt (1994))

$$X_t = aX_{t-1} + \Big[a_0 + (\sum_{i=1}^{p} (a_i^+ X_{t-i}^+ + a_i^- X_{t-i}^-))^{2\beta}\Big]^{1/2} \varepsilon_t, \qquad (3.5)$$

where $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$, a is a real constant, and β, a_0, a_i^+ and a_i^- are positive constants. If $a = 0$ and $p = 1$, the probabilistic properties of the β–ARCH(1) model were discussed by Guegan and Diebolt (1994); but for general p, as they have pointed out, their results cannot be directly extended to model (3.5) for $p \geq 2$. However, if $\beta < 1$ and $|a| < 1$, it is easily seen from the above Theorem 1 that under 1) of Assumption 1, model (3.5) is geometrically ergodic with $E_\pi(X_t^2) < \infty$; and if $\beta = 1$ and $a^2 + (\sum_{i=1}^p \max(a_i^+, a_i^-))^2 < 1$, model (3.5) is still geometrically ergodic with $E_\pi(X_t^2) < \infty$. The second part may be obtained by using Theorem 1 (see also the following Remark 2) and the facts that corresponding to (1.3), $f(x) = ax_1$ and

$$h(x) = a_0 + \Big(\sum_{i=1}^p (a_i^+ x_i^+ + a_i^- x_i^-)\Big)^2 \leq \Big(\sum_{j=1}^p \delta_j\Big) \sum_{i=1}^p \delta_i x_i^2 + o(\|x\|^2)$$

as $\|x\| \to \infty$, where $\delta_i = \max(a_i^+, a_i^-)$.

If $f \equiv 0$, then (1.1) is the pure ARCH model. From Theorem 1 we have

Corollary 1. *For the pure ARCH model* (1.1) ($f \equiv 0$)*, if Assumption* 1 *holds and*

$$\sum_{i=1}^q d_i < 1, \tag{3.6}$$

then model (1.1) ($f \equiv 0$) *is geometrically ergodic and* $E_\pi(X_t^2) < \infty$.

Remark 2. The result of this Corollary can be derived from Lu (1995), who gave a more general result. And in fact, as shown in Lu (1995), if the equality sign in the decomposition of h in (1.3) is replaced by the inequality sign " $\leq$ " , the conclusion of Theorem 1 still holds. This is easily seen from the proof. In addition, the existence of the second moment for X_t is obtained in Theorem 1; however, Proposition 3 of Nze (1992) may not guarantee the existence of second moments.

If $d = \mathbf{0}$ in (1.3), we may still, from Theorem 1, deduce a sufficient condition for the geometric ergodicity of (1.1), but the following theorem gives a milder sufficient condition for this case.

Theorem 2. *For the Markov chain* $\{Y_t\}$ *in* (2.1b) *defined by* (1.1) ($p = q$)*, if Assumption* 1 *with* $d_i = 0$*,* $i = 1, \ldots, p$ *in* (1.3) *holds and*

$$\rho(A) < 1, \tag{3.7}$$

then $\{Y_t\}$ *is geometrically ergodic* (*hence so is model* (1.1)) *and* $E_\pi(|X_t|) < \infty$.

Proof. It may be proved similarly to the corresponding part of Lemma 3.1 of Masry and Tjϕstheim (1995) (i.e. Theorem 0 in the introduction) and hence the proof is omitted.

If $h \equiv \sigma^2$ (positive constant) in (1.1), from Theorem 2 we easily get the following corollary which can be seen in Tjϕstheim (1994) and An and Huang (1994).

Corollary 2. *For the non-linear autoregressive model* (1.1) ($h \equiv \sigma^2$)*, if Assumption* 1 *holds and* (3.7) *is satisfied, then model* (1.1) ($h \equiv \sigma^2$) *is geometrically ergodic and* $E_\pi(|X_t|) < \infty$.

Remark 3.

(a) From the proof of Theorem 2, it follows that only the finiteness of the absolute first moment of ε_t is used. Hence, the requirement $E(\varepsilon_t^2) = 1$ in (1.1) can be weakened to $E(|\varepsilon_t|) < \infty$ for this theorem.

(b) In Theorem 0 (i.e. Lemma 3.1 of Masry and Tjϕstheim (1995)), the functions f and h are assumed to be non–periodic in (b) of Assumption 0, but this condition is not necessary in Theorem 2.

(c) It is well known that for the linear AR(p) model: $X_t = a_1X_{t-1} + \cdots + a_pX_{t-p} + \varepsilon_t$ (for the assumption on $\{\varepsilon_t\}$, see (1.1)), (3.7) is a necessary and sufficient condition for the stationarity. Hence, under the setting of this theorem, condition (3.7) is quite mild.

(d) In Mckeague and Zhang (1994) and Bhattacharya and Lee (1995), h is assumed to be bounded. Similar to Theorem 2, if h is unbounded with $h(y) = o(\|y\|^2)$ as $\|y\| \to \infty$, the results of Mckeague and Zhang (1994) and Bhattacharya and Lee (1995) still hold.

4. Some Further Discussion

The sufficient condition (3.1) in Theorem 1 is not the weakest possible for the geometric ergodicity. This can be seen from Theorem 2 as $d = \mathbf{0}$. Specifically, let $p = 2$, $f(y_1, y_2) = a_1y_1 + a_2y_2$, $h(y_1, \ldots, y_q) \equiv \sigma^2$, Then it follows from Theorem 1 that a sufficient condition is $|a_1|+|a_2| < 1$; but it is well known that the weakest condition is $|a_1| + a_2 < 1$ and $|a_2| < 1$. However, when $p = 1$, $f(y_1) = a_1y_1$, $h(y_1, \ldots, y_q) = d_0 + d_1y_1^2 + \cdots + d_qy_q^2$ with $d_i \geq 0, i = 1, \ldots, q$ and $d_0 > 0$, we have

Theorem 3. *Consider the following model:*

$$X_t = a_1X_{t-1} + \varepsilon_t(d_0 + d_1X_{t-1}^2 + \cdots + d_qX_{t-q}^2)^{1/2}, \tag{4.1}$$

where $d_i \geq 0$, $i = 1, \ldots, q$, $d_0 > 0$ *and* $\{\varepsilon_t\}$ *satisfies the assumption right after* (1.1)*. If* (1) *of Assumption* 1 *holds, then the necessary and sufficient condition for the geometric ergodicity of* (4.1) *with* $E_\pi(X_t^2) < \infty$ *is*

$$a_1^2 + \sum_1^q d_i < 1. \tag{4.2}$$

Proof. The sufficiency is obvious from Theorem 1.

For the necessity, on account of the geometric ergodicity of (4.1) with $E_\pi(X_t^2) < \infty$, it follows from (4.1) that

$$\begin{aligned} E_\pi(X_t^2) &= a_1^2 E_\pi(X_{t-1}^2) + d_0 + d_1 E_\pi(X_{t-1}^2) + \cdots + d_q E_\pi(X_{t-q}^2) \\ &= (a_1^2 + d_1 + d_2 + \cdots + d_q) E_\pi(X_t^2) + d_0, \end{aligned}$$

which shows that (4.2) holds.

Remark 4.

(a) If $a_1 X_{t-1}$ in (4.1) is replaced by some $a_{i_o} X_{t-i_o}$ and a_1^2 in (4.2) by $a_{i_o}^2$, the conclusion of Theorem 3 still holds.

(b) Just as Remark 3(c) shows, if $f(y) = a_{i_o} y_{i_o} + o(\|y\|)$ for some $1 \le i_o \le p$ and $h(y) = \sum_1^q d_i y_i^2 + o(\|y\|^2)$ $(\|y\| \to \infty)$ in (1.1), the sufficient condition (4.2) with a_1^2 replaced by $a_{i_o}^2$ is quite weak. As $\rho(A) = |a_{i_o}|$ (see (1.4) for A), this leads us to pose the following conjecture.

Conjecture: For the Markov chain $\{Y_t\}$ defined by (2.1b), if Assumption 1 holds and $\rho(A)^2 + \sum_{i=1}^q d_i < 1$, then $\{Y_t\}$ is geometrically ergodic (hence so is model (1.1)) and $E_\pi(X_t^2) < \infty$.

If this conjecture is true, it subsumes many known results on the geometric ergodicity of model (1.1) in the literature. It needs further study.

Acknowledgement

I thank Prof. Hong-zhi An for introducing me to the ARCH model and Prof. Dag Tjϕstheim for his kindly sending me a preprint of his paper (with E. Masry) and his encouraging comments on this paper. I would like to express my especial gratitude to two referees, an Associate Editor, Prof. C. F. Jeff Wu and Prof. Ching-Shui Cheng for their useful comments, suggestions and great helps which substantially improved the early version of this paper. Thanks are also due to Dr. Li-xing Zhu for his careful reading and help and to my two advisors Prof. Ping Cheng and Prof. Guo-ying Li for their helps.

References

An, H. Z. and Huang, F. C. (1994). The geometric ergodicity of nolinear autoregressive model. Technical Report, Institute of Applied Mathematics, Academia Sinica.

Auestad, B. and Tjϕstheim, D. (1990). Identification of nonlinear time series: First order characterization and order determination. *Biometrika* **77**, 669–687.

Bhattacharya, R. N. and Lee, C. (1995). Ergodicity of nonlinear first order autoregressive models. *J. Theoret. Probab.* **8**, 207–219.

Chan, K. S. (1990). Deterministic stability, stochastic stability, and ergodicity. In *Nonlinear Time Series: A Dynamical System Approach* 1990 (Edited by H. Tong), 448-466. Oxford University Press, Oxford.

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

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.

Guegan, D. and Diebolt, J. (1994). Probabilistic properties of the β-ARCH model. *Statist. Sinica* **4**, 71-87.

Lu, Z. D. (1996). Geometric ergodicity of a general ARCH type model with applications to some typical models. *Chinese Sci. Bull.* **41**, 1630. (digest, in Chinese)

Lu, Z. D. (1996). A note on geometric ergodicity of autoregressive conditional heteroscedasticity (ARCH) model. *Statist. Probab. Lett.* **30**, 305-311.

Masry, E. and Tjϕstheim, D. (1995). Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality. *Econometric Theory* **11**, 258-289.

Mckeague, I. W. and Zhang, M. J. (1994). Identification of nonlinear time series from first order cumulative characteristics. *Ann. Statist.* **22**, 495-514.

Nummelin, E. (1984). *General Irreducible Markov Chains and Non-negative Operations.* Cambridge University Press, Cambridge.

Nummelin, E. and Tuominen, P. (1984). Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.* **12**, 187-202.

Nze, P. A. (1992). Critéres d'ergodicité de quelques modéles à représentation markovienne. *C. R. Acad. Sci. Paris Sér. I Math.* **315**, 1301-1304.

Tjϕstheim, D. (1990). Nonlinear time series and Markov chains. *Adv. Appl. Probab.* **22**, 587-611.

Tjϕstheim, D. (1994). Nonlinear time series: a selective review. *Scand. J. Statist.* **21**, 97-130.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach.* Oxford University Press, Oxford.

Tweedie, R. L. (1983). The existence of moments for stationary Markov chains. *J. Appl. Probab.* **20**, 191-196.

Institute of Systems Science, Academia Sinica, Beijing 100080.

E-mail: zdlu@iss04.iss.ac.cn

(Received September 1995; accepted September 1997)