Tang and Wang *Advances in Difference Equations* 2013, **2013**:200
http://www.advancesindifferenceequations.com/content/2013/1/200

Advances in Difference Equations
a SpringerOpen Journal

RESEARCH

Open Access

On nonergodicity for nonparametric autoregressive models

Mingtian Tang and Yunyan Wang*

*Correspondence:
yunyanwang0329@yahoo.com
School of Science, Jiangxi University of Science and Technology, Ganzhou, 341000, China

Abstract

In this paper, we introduce a class of nonlinear time series models with random time delay under random environment, sufficient conditions for nonergodicity of these models are developed. The so-called Markovnization methods are used, that is, proper supplementary variables are added to a non-Markov process, then a new Markov process can be obtained.

MSC: 60J05; 60J10; 60K37

Keywords: Markov chains; random delay; random environment; nonergodic

1 Introduction

By virtue of their superduper properties, stable (ergodic or recurrent) stochastic processes are very popular among many researchers, so there has been a large literature devoted to the stable (ergodic or recurrent) or even stationary stochastic processes. For instance, Jeantheau [1] and Tjøstheim [2] established consistency of the estimator they proposed under stationarity and ergodicity conditions (see also [3–5]). Fernandes and Grammig [6] established conditions for the existence of higher-order moments, strict stationarity, geometric ergodicity and β-mixing property with exponential decay. This owes a great deal to the beautiful properties of stable processes, such as an ergodic Markov chain has an invariant probability measure which is finite, a recurrent stochastic process re-visits an arbitrary point in its image an infinite number of times. Just because of this, many researchers often like to target ergodicity or recurrence as their assumptions in their papers or books.

However, in this colorful world, lots and lots of phenomena exhibit instability behavior, for example, David [7] argued that an important lesson from economic history was that economies exhibited nonergodic behavior along many dimensions. Margolin and Barkai [8] indicated that time series of many systems exhibited intermittency, that is to say, at random times the system will switch from state on (or up) to state off (or down) and *vice versa*. One method to characterize such time series is using time average correlation functions to exhibit a nonergodic behavior.

Hence more and more researchers become increasingly interested in these instable processes. Recently, some problems of nonergodic stochastic processes have been studied by many authors. Basawa and Koul [9], Basawa and Brockwell [10], Basawa and Scott [11] and Feigin [12] studied asymptotic inference problems for parameters of nonergodic stochastic processes. Budhiraja and Ocone [13] proved an asymptotic stability result for discrete time systems in which the signal was allowed to be nonergodic. Durlauf [14] considered nonergodic economic growth. Goodman and Massey [15] generalized Jackson's theorem so that the large-time behavior can be described for any nonergodic N-node Jackson network system. Griffeath [16] developed limit theorems for nonergodic set-valued Markov processes. Jacod [17] constructed the estimators for drift and diffusion coefficients of a multidimensional diffusion process and obtained consistent results without any kind of ergodicity or even recurrence assumption on the diffusion process.

Ergodicity criteria with drift functions for Markov processes have been studied by many authors. For instance, see Cline [18] and Tweedie [19–21] and the references therein. As for nonergodicity criteria for Markov processes, the readers are referred to [22–24]. Sheng *et al.* [25] also developed some sufficient conditions for nonergodicity of some time series models.

However, the processes considered by many researchers do not reflect the factors of the interference in a system and the system itself influenced by sudden environmental change. On the other hand, the time delay in the models studied is usually a fixed constant. In this paper, we popularize general nonparametric autoregressive models through introducing random environment and at the same we turn to a random time delay instead of a fixed time delay.

The remainder of the paper is organized as follows. Section 2 introduces the nonparametric autoregressive model with random time delay under random environment. Section 3 develops some useful lemmas and gives some sufficient conditions for nonergodicity of the proposed model as our main results. All the proofs are collected in Section 4.

2 The nonparametric AR model with random time delay under random environment

In this section, we first give some notations which will be used throughout the paper. In what follows, we always have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a finite set $\mathrm{E} = \{1, 2, \dots, r\}$ (r is a positive integer number), with $\mathcal{H}$ as the σ-algebra generated by all sets of E. We also let R^m be an m-dimensional real space and $\mathcal{B}_m$ be the σ-algebra generated by all Boreal subsets of R^m.

The nonparametric autoregressive model with random time delay under random environment is defined as follows:

$$X_{t+1} = F(X_t, X_{t-1}, \dots, X_{t-Z_{t+1}}) + \varepsilon_{t+1}(Z_{t+1}), \tag{2.1}$$

where $\{Z_t, t \geq 1\}$ is an irreducible and aperiodic Markov chain defined on $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in E; $\forall j \in \mathrm{E}$, $F : R^{j+1} \to R$ is a Borel measurable mapping; $\varepsilon_t(Z_t) = \sum_{i=1}^r \varepsilon_t(i) I_{\{i\}}(Z_t)$, where $I_{\{i\}}(Z_t)$ denotes the indicator function of a single element set $\{i\}$; $\{\varepsilon_t(1)\}, \{\varepsilon_t(2)\}, \dots, \{\varepsilon_t(r)\}$ are sequences of i.i.d. random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $(R, \mathcal{B})$. In what follows, the density function of $\{\varepsilon_t(i)\}$, $i \in \mathrm{E}$, is denoted as $\Phi(\cdot)$.

In this paper, we aim to obtain some sufficient conditions for nonergodicity of the proposed new model. Since the nonparametric AR model with random time delay under random environment defined by (2.1) itself does not have the Markovian property, we consider the following sequence:

$$(X_t, X_{t-1}, \dots, X_{t-r}, Z_t) = \bigl(F(X_{t-1}, X_{t-2}, \dots, X_{t-Z_t}), X_{t-1}, \dots, X_{t-r}, Z_t\bigr), \quad t \leq 1.$$

For simplicity, let

$$Y_t = (X_t, X_{t-1}, \ldots, X_{t-r}, Z_t), \quad t \geq 1.$$

From (2.1) and under mild conditions (see Lemma 1 below), it is easy to show that the sequence $\{Y_t\}$ is a Markov chain on $R^{r+1} \times \mathrm{E}$ with the following transition probability:

$$\mathbb{P}\bigl(Y_{t+1} \in A \times \{j\} | Y_t = \hat{x} \times \{i\}\bigr) = p_{ij} \prod_{q=0}^{r-1} I_{A_{q+1}}(x_q) \int_{A_0} \Phi_j\bigl(y_0 - F(x_0, x_1, \ldots, x_j)\bigr)\, dy_0, \quad (2.2)$$

where $A = A_0 \times A_1 \times \cdots \times A_r \subset R^{r+1}$ and $A_i \subset R$; $I_A(\cdot)$ denotes the indicator function of A; $p_{ij} = \mathbb{P}(Z_{t+1} = j | Z_t = i)$ is the transition function of the Markov chain $\{Z_t, t \geq 1\}$.

Let $P^{(t)}((\hat{x}, i), A \times \{j\}) = \mathbb{P}(Y_{s+t} \in A \times \{j\} | Y_s = (\hat{x}, i))$ be the t-step transition probability of $\{Y_t\}$, then by the property of conditional probability and inductive approach we have: when $2 \leq t \leq r+1$,

$$\begin{aligned} &P^{(t)}\bigl((\hat{x}, i), A \times \{j\}\bigr) \\ &\quad = \prod_{q=0}^{r-t} I_{A_{q+t}}(x_q) \sum_{k_1 k_2 \cdots k_{t-1} \in \mathrm{E}} p_{ik_1} p_{k_1 k_2} \cdots p_{k_{t-1} j} \int_{A_{t-1}} \Phi_{k_1}\bigl(y_{t-1} - F(U_t)\bigr)\, dy_{t-1} \\ &\qquad \cdot \int_{A_{t-2}} \Phi_{k_2}\bigl(y_{t-2} - F(U_{t-1})\bigr)\, dy_{t-2} \cdots \int_{A_1} \Phi_{k_{t-1}}\bigl(y_1 - F(U_2)\bigr)\, dy_1 \\ &\qquad \cdot \int_{A_0} \Phi_j\bigl(y_0 - F(U_1)\bigr)\, dy_0, \end{aligned}$$

when $t \geq r+2$,

$$\begin{aligned} &P^{(t)}\bigl((\hat{x}, i), A \times \{j\}\bigr) \\ &\quad = \sum_{k_1 k_2 \cdots k_{t-1} \in \mathrm{E}} p_{ik_1} p_{k_1 k_2} \cdots p_{k_{t-1} j} \int_R \Phi_{k_1}\bigl(y_{t-1} - F(U_t)\bigr)\, dy_{t-1} \\ &\qquad \cdot \int_R \Phi_{k_2}\bigl(y_{t-2} - F(U_{t-1})\bigr)\, dy_{t-2} \cdots \int_R \Phi_{k_{t-(r+1)}}\bigl(y_{r+1} - F(U_{r+2})\bigr)\, dy_{r+1} \\ &\qquad \cdot \int_{A_r} \Phi_{k_{t-r}}\bigl(y_r - F(U_{r+1})\bigr)\, dy_r \cdots \int_{A_0} \Phi_j\bigl(y_0 - F(U_1)\bigr)\, dy_0, \end{aligned}$$

where

$$\begin{aligned} &U_t = (x_0, x_1, \ldots, x_{k_1}), \\ &U_{t-s} = (y_{t-s}, y_{t-s+1}, \ldots, y_{t-1}, x_0, \ldots, x_{k_{s+1}-s}), \quad 1 \leq s \leq k_{s+1}, \\ &U_{t-s} = (y_{t-s}, y_{t-s+1}, \ldots, y_{t-s+k_{s+1}}), \quad k_{s+1} < s \leq t-1. \end{aligned}$$

3 Main results

This section gives the main results of the new model described in Section 2. We need the following conditions.

Assumption 1 *$\{\varepsilon_t(1)\},\dots,\{\varepsilon_t(r)\}$ are mutually independent, and $\forall i\in \mathrm{E}$, $\varepsilon_{t+1}(i)$ is independent of $\{X_s, s\le t\}$. Moreover, for each i, $E(\varepsilon_t(i))$ is a constant independent of t and $E[\varepsilon_t(i)]^2<\infty$.*

Assumption 2 *$\{Z_t, t\ge 1\}$ is independent of the initial random variable X_0.*

Assumption 3 *$\forall i\in \mathrm{E}$, $\{Z_t, t\ge 1\}$ is independent of $\varepsilon_{t+1}(i)$.*

Assumption 1 assures the stationarity of $\{\varepsilon_t(i)\}$, $i\in \mathrm{E}$. Assumption 2 and Assumption 3 guarantee the Markov property of $\{Y_t\}$. These are the basic conditions we know that guarantee the following lemmas can be used properly throughout the paper.

Lemma 1 *Suppose that Assumptions* 1-3 *hold, then the sequence $\{Y_t\}$ is a time-homogeneous Markov chain defined on $(\Omega,\mathcal{F},\mathbb{P})$ with state space $(R^{r+1}\times \mathrm{E},\mathcal{B}_{r+1}\times\mathcal{H})$.*

The irreducibility and aperiodicity in Lemma 2 are standard and can be found in Meyn and Tweedie [26] and Tong [27]. The two concepts are very useful to derive the nonergodicity of the sequence $\{Y_t\}$. But before we state the results about the irreducibility and aperiodicity of the sequence $\{Y_t\}$, we need the following condition about the density function of $\varepsilon_t(i)$, $i\in \mathrm{E}$.

Assumption 4 *The density function $\Phi_i(\cdot)$ of $\varepsilon_t(i)$, $i\in \mathrm{E}$ is strictly positive everywhere, i.e., $\forall i\in \mathrm{E}$, $\Phi_i(\cdot)>0$.*

Lemma 2 *Under Assumptions* 1-4, *the Markov chain $\{Y_t\}$ is $\mu_{r+1}\times\varphi$ irreducible and aperiodic, where φ is a measure on $(\mathrm{E},\mathcal{H})$, μ_{r+1} is a Lebesgue measure on $(R^{r+1},\mathcal{B}_{r+1})$ satisfying $\mu_{r+1}\times\varphi(A\times B)>0$ if $\mu_{r+1}(A)>0$, $A\in\mathcal{B}_{r+1}$, $B\in\mathcal{H}$.*

Remark 1 Obviously, if the Markov chain $\{Y_t\}$ is φ_1-irreducible, for any nontrivial and σ-finite measure φ_2 which is absolutely continuous with respect to φ_1, then $\{Y_t\}$ is also φ_2-irreducible. So, we need a normal irreducibility which can define the range of the chain much more completely than some more arbitrary irreducibility measures one may construct initially. Fortunately, Sheng *et al.* [25] and Meyn and Tweedie [26] proved that if $\{Y_t\}$ is a $\mu_{r+1}\times\varphi$ irreducible Markov chain, then there exists a maximal irreducibility measure Q. In this paper, we use those subsets whose maximal irreducibility measure is greater than zero, so here we denote $(\mathcal{B}_{r+1}\times\mathcal{H})^+=\{A\in\mathcal{B}_{r+1}\times\mathcal{H}: Q(A)>0\}$.

Our main results are as follows.

Theorem 1 *Suppose that Assumptions* 1-4 *hold and if there exist a non-negative measurable function g on $(R^{r+1}\times \mathrm{E},\mathcal{B}_{r+1}\times\mathcal{H})$, a set $A\times \mathrm{M}\in\mathcal{B}_{r+1}\times\mathcal{H}$, and a non-negative measurable function $h(\cdot)$ on $[0,\infty)$ satisfying*

$$\max_{j\in \mathrm{E}}\left(\int_R h(u)\Phi_j(u)\,du\right)<\infty,$$

such that

(1) *For* $\forall x_0, x_1, \ldots, x_r, y_0 \in R$, $\forall i, j \in \mathrm{E}$, *we have*

$$\bigl|g\bigl((x_0, x_1, \ldots, x_r), i\bigr) - g\bigl((y_0, x_1, \ldots, x_r), j\bigr)\bigr| \le h\bigl(|x_0 - y_0|\bigr);$$

(2) *For* $(A \times \mathrm{M})^c \in (\mathcal{B}_{r+1} \times \mathcal{H})^+$, *we have*

$$g(\hat{x}, i) \ge \sup_{(\hat{y}, j) \in A \times \mathrm{M}} g(\hat{y}, j); \tag{3.1}$$

(3) *For* $\forall(\hat{x}, j) \in (A \times \mathrm{M})^c$, *we have*

$$g\bigl(\bigl(F(x_0, x_1, \ldots, x_j), y_1, \ldots, y_r\bigr), j\bigr) \ge g(\hat{x}, j) + \max_{j \in \mathrm{E}} \biggl(\int_R h(u)\Phi_j(u)\,du\biggr), \tag{3.2}$$

and there exist $B \times \mathrm{I} \subset (A \times \mathrm{M})^c$, $B \times \mathrm{I} \in (\mathcal{B}_{r+1} \times \mathcal{H})^+$, $\forall(\hat{x}, i) \in B \times \mathrm{I}$, *such that*

$$g\bigl(\bigl(F(x_0, x_1, \ldots, x_j), y_1, \ldots, y_r\bigr), j\bigr) > g(\hat{x}, j) + \max_{j \in \mathrm{E}} \biggl(\int_R h(u)\Phi_j(u)\,du\biggr). \tag{3.3}$$

Then the Markov chain $\{Y_t\}$ *is nonergodic. Moreover, whatever a probability distribution function of* $\{X_t\}$ *is, its probability distribution function will never converge to some probability distribution function.*

Remark 2 Conditions (2) and (3) in the above Theorem 1 can be substituted for (4) or (5) as follows:

(4) For $A \times \mathrm{M}, (A \times \mathrm{M})^c \in (\mathcal{B}_{r+1} \times \mathcal{H})^+$, (3.1) holds, and $\forall(\hat{x}, i) \in (A \times \mathrm{M})^c$, (3.2) holds.

(5) For $A \times \mathrm{M} \in (\mathcal{B}_{r+1} \times \mathcal{H})^+$ and $\forall(\hat{x}, i) \in R^{r+1} \times \mathrm{E}$, (3.2) holds, where when $(\hat{x}, i) \in A \times \mathrm{M}$, (3.3) holds.

That is, under conditions (1) and (4) or (1) and (5), we can also show that $\{Y_t\}$ is nonergodic, and the method of the proof is similar to that under conditions (1)-(3).

4 Proofs

Proof of Lemma 1 $\forall \hat{x} = (x_0, x_1, \ldots, x_r), \hat{x}_s \in R^{r+1}$, and $\forall i, i_s \in \mathrm{E}$, where s is an integer number satisfying $0 \le s < t$, we have

$$\begin{aligned}
&\mathbb{P}\bigl\{Y_{t+1} \in A \times \{j\} | Y_t = (\hat{x}, i), Y_s = (\hat{x}_s, i_s), 0 \le s < t\bigr\} \\
&\quad = \mathbb{P}\bigl\{(X_{t+1}, X_t, \ldots, X_{t+1-r}) \in A, Z_{t+1} = j | Y_t = \hat{x}, Z_t = i, Y_s = (\hat{x}_s, i_s), 0 \le s < t\bigr\} \\
&\quad = \mathbb{P}\bigl\{F(X_t, X_{t-1}, \ldots, X_{t-Z_{t+1}}) + \varepsilon_{t+1}(Z_{t+1}) \in A_0, \\
&\qquad X_t \in A_1, X_{t-1} \in A_2, \ldots, X_{t+1-r} \in A_r, Z_{t+1} = j | \\
&\qquad (X_t, X_{t-1}, \ldots, X_{t-r}) = (x_0, x_1, \ldots, x_r), Z_t = i, Y_s = (\hat{x}_s, i_s), 0 \le s < t\bigr\} \\
&\quad = \mathbb{P}\bigl\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r, Z_{t+1} = j | \\
&\qquad (X_t, X_{t-1}, \ldots, X_{t-r}) = (x_0, x_1, \ldots, x_r), Z_t = i\bigr\} \\
&\quad = \mathbb{P}\bigl\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r\bigr\} P\{Z_{t+1} = j | Z_t = i\} \\
&\quad = p_{ij} \cdot \mathbb{P}\bigl\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r\bigr\},
\end{aligned}$$

where the last equation follows from the definition of the (2.1) model, Assumption 1 and the notation $p_{ij} = \mathbb{P}\{Z_{t+1} = j|Z_t = i\}$.

On the other hand,

$$\begin{aligned}
&\mathbb{P}\big\{Y_{t+1} \in A \times \{j\}|Y_t = (\hat{x}, i)\big\} \\
&\quad= \mathbb{P}\big\{(X_{t+1}, X_t, \ldots, X_{t+1-r}) \in A, Z_{t+1} = j|(X_t, X_{t-1}, \ldots, X_{t-r}) = (x_0, x_1, \ldots, x_r), Z_t = i\big\} \\
&\quad= \mathbb{P}\big\{F(X_t, X_{t-1}, \ldots, X_{t-Z_{t+1}}) + \varepsilon_{t+1}(Z_{t+1}) \in A_0, \\
&\qquad X_t \in A_1, X_{t-1} \in A_2, \ldots, X_{t+1-r} \in A_r, Z_{t+1} = j| \\
&\qquad (X_t, X_{t-1}, \ldots, X_{t-r}) = (x_0, x_1, \ldots, x_r), Z_t = i\big\} \\
&\quad= \mathbb{P}\big\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r, Z_{t+1} = j| \\
&\qquad (X_t, X_{t-1}, \ldots, X_{t-r}) = (x_0, x_1, \ldots, x_r), Z_t = i\big\} \\
&\quad= \mathbb{P}\big\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r\big\} P\{Z_{t+1} = j|Z_t = i\} \\
&\quad= p_{ij} \cdot \mathbb{P}\big\{F(x_0, x_1, \ldots, x_j) + \varepsilon_{t+1}(j) \in A_0, x_0 \in A_1, x_1 \in A_2, \ldots, x_{r-1} \in A_r\big\}.
\end{aligned}$$

Hence the sequence $\{Y_t\}$ is a Markov chain, and its time-homogeneity follows from the stationarity of $\varepsilon_{t+1}(j)$, $j \in \mathrm{E}$. This completes the proof. □

Proof of Lemma 2 Suppose that $A \times B \in \mathcal{B}^{r+1} \times \mathcal{H}$ and $\mu_{r+1} \times \varphi(A \times B) > 0$. Since $\{Z_t\}$ is irreducible, $\forall i, j \in \mathrm{E}$, $\exists s > 0$, such that

$$p_{ij}^{(t)} = \mathbb{P}(Z_{t+s} = j|Z_s = i) > 0, \quad \forall t \geq s,$$

that is, $\exists k_1, k_2, \ldots, k_{t-1} \in \mathrm{E}$, such that

$$p_{ik_1} p_{k_1 k_2} \cdots p_{k_{t-1} j} > 0.$$

Then from the t-step transition probability of $\{Y_t\}$, $\forall (\hat{x}, i) \in R^{r+1} \times \mathrm{E}$, we have

$$P^{(t)}(\hat{x}, i; A, j) > 0,$$

so $\{Y_t\}$ is $\mu_{r+1} \times \varphi$ irreducible, and the aperiodicity of $\{Y_t\}$ follows from Tong [27]. This completes the proof. □

For $x, y \in \Omega$, $z \in [0, 1)$, let

$$\psi_g(x, z) = \frac{1}{1 - z} \left(z^{g(x)} - \int_\Omega P(x, dy) z^{g(y)} \right).$$

In order to deal with the proofs of Theorem 1, we need the following propositions.

Lemma 3 [25] *Suppose* $\{X_t\}$ *is a* φ*-irreducible Markov chain on the state space* $(\Omega, \mathcal{F})$. *If there exist constants* $N > 0$, $0 < C < 1$, *a set* A *and a nonnegative measurable function* $g(x)$ *which satisfies* $\forall x \in \Omega$, $\int_\Omega P(x, dy) g(y) < +\infty$, *such that*

(1) $\psi_g(x, z) \geq -N$, $x \in \Omega$, $z \in [C, 1)$;

(2) *One of the following* (i), (ii), (iii) *hold*:

(i) *For* $A^c \in \mathcal{F}^+$, $\forall x \in A^c$, *we have*

$$g(x) \geq \sup_{y\in A} g(y), \qquad \int_\Omega P(x,dy)\big[g(y)-g(x)\big] \geq 0;$$

furthermore, there exist $B \subset A^c$, $B \in \mathcal{F}^+$ *such that*

$$\int_\Omega P(x,dy)\big[g(y)-g(x)\big] > 0, \quad x \in B.$$

(ii) *For* $A \in \mathcal{F}^+$, $A^c \in \mathcal{F}^+$, $\forall x \in A^c$, *we have*

$$g(x) \geq \sup_{y\in A} g(y), \qquad \int_\Omega P(x,dy)\big[g(y)-g(x)\big] > 0.$$

(iii) *For* $A \in \mathcal{F}^+$, *we have*

$$\int_\Omega P(x,dy)\big[g(y)-g(x)\big] \geq 0, \quad x \in \Omega,$$
$$\int_\Omega P(x,dy)\big[g(y)-g(x)\big] > 0, \quad x \in A.$$

Then $\{X_t\}$ *is nonergodic.*

Remark 3 Condition (1) in Lemma 3 is usually called the Kaplan condition; readers can consult Kaplan [24] for details. Sheng *et al.* [25] found a class of functions satisfying this condition. That is, if there exists a constant $N > 0$ such that

$$\int_{g(y)<g(x)} P(x,dy)\big[g(y)-g(x)\big] \geq -N, \quad x \in \Omega,$$

then $g(x)$ is the function wished.

Remark 4 Generally, the function $\int_\Omega P(x,dy)[g(y)-g(x)]$, which we use frequently in Lemma 3, goes by the name of g-drift of the point x and is often expressed as $\gamma_g(x)$. In addition, if

$$\int_\Omega P(x,dy)g(y) < \infty, \quad x \in \Omega,$$

we have

$$\lim_{z\to 1} \psi_g(x,z) = -g(x) + \int_\Omega P(x,dy)g(y) = \gamma_g(x).$$

In fact, $\frac{dz^{g(y)}}{dz} = g(y)z^{g(y)-1}$, so when $z \in [\frac{1}{2},1)$, $\frac{dz^{g(y)}}{dz} \leq 2g(y)$.

Proof of Theorem 1 By Lemma 1 and Lemma 2 we know that $\{Y_t\}$ is a $\mu_{r+1} \times \varphi$ irreducible, aperiodic and time-homogeneous Markov chain with state space $(R^{r+1} \times \mathrm{E}, \mathcal{B}_{r+1} \times \mathcal{H})$. So, we only need to check the conditions in Lemma 3. For simplicity, let

$$C = \max_{j\in \mathrm{E}} \left(\int_R h(u)\Phi_j(u)\,du \right),$$
$$\hat{v}_u = \bigl(F(x_0, x_1, \ldots, x_j) + u, x_0, \ldots, x_{r-1}\bigr),$$
$$\hat{v} = \bigl(F(x_0, x_1, \ldots, x_j), x_0, \ldots, x_{r-1}\bigr).$$

Step 1: To show that $\int_{R^{r+1}\times \mathrm{E}} P((\hat{x}, i), d(\hat{y}, j))g(\hat{y}, j) < \infty$. For $\hat{x} = (x_0, x_1, \ldots, x_r) \in R^{r+1}$ and $\hat{y} = (y_0, y_1, \ldots, y_r) \in R^{r+1}$, $i, j \in \mathrm{E}$, by (2.2), we have

$$P\bigl((\hat{x}, i), d(\hat{y}, j)\bigr) = p_{ij} \cdot \prod_{q=0}^{r-1} \delta(y_{q+1} - x_q)\Phi_j\bigl(y_0 - F(x_0, \ldots, x_j)\bigr)\,dy_0,$$

where $\delta(\cdot)$ is a δ-function, that is, $\delta(x - y) = 1$ if $x = y$ and zero otherwise.

So, by the integral transformation theorem, we can get

$$\begin{aligned}
&\int_{R^{r+1}\times \mathrm{E}} P\bigl((\hat{x}, i), d(\hat{y}, j)\bigr)g(\hat{y}, j) \\
&\quad = \sum_{j\in \mathrm{E}} \int_{R^{r+1}} P\bigl((\hat{x}, i), d(\hat{y}, j)\bigr)g(\hat{y}, j) \\
&\quad = \sum_{j\in \mathrm{E}} p_{ij} \int_R \Phi_j\bigl(y_0 - F(x_0, x_1, \ldots, x_j)\bigr)g(y_0, x_0, \ldots, x_{r-1}, j)\,dy_0 \\
&\quad = \sum_{j\in \mathrm{E}} p_{ij} \int_R g(\hat{v}_u, j)\Phi_j(u)\,du \\
&\quad \le \sum_{j\in \mathrm{E}} p_{ij} \int_R h(u)\Phi_j(u)\,du + \sum_{j\in \mathrm{E}} p_{ij} g(\hat{v}, i) \\
&\quad \le C + g(\hat{v}, i) < \infty,
\end{aligned}$$

where the second-to-last line follows from condition (1) in the theorem.

Step 2: To show that when $(\hat{x}, i) \in (A \times \mathrm{M})^c$, $\gamma_g(\hat{x}, i) \ge 0$ and when $(\hat{x}, i) \in B \times \mathrm{I}$, $\gamma_g(\hat{x}, i) > 0$. In fact,

$$\begin{aligned}
\gamma_g(\hat{x}, i) &= \int_{R^{r+1}\times \mathrm{E}} P\bigl((\hat{x}, i), d(\hat{y}, j)\bigr)\bigl[g(\hat{y}, j) - g(\hat{x}, i)\bigr] \\
&= \sum_{j\in \mathrm{E}} \int_{R^{r+1}} P\bigl((\hat{x}, i), d(\hat{y}, j)\bigr)\bigl[g(\hat{y}, j) - g(\hat{x}, i)\bigr] \\
&= \sum_{j\in \mathrm{E}} \int_R p_{ij}\Phi_j\bigl(y_0 - F(x_0, x_1, \ldots, x_j)\bigr)\bigl[g(y_0, x_0, \ldots, x_{r-1}, j) - g(\hat{x}, i)\bigr]\,dy_0 \\
&= \sum_{j\in \mathrm{E}} p_{ij} \left(\int_R g(\hat{v}_u, j)\Phi_j(u)\,du - g(\hat{x}, i) \right)
\end{aligned}$$

$$\begin{aligned}&\geq \sum_{j\in \mathrm{E}} p_{ij}\Bigl(-\int_R h(u)\Phi_j(u)\,du + \bigl(g(\hat{\nu},i)-g(\hat{x},i)\bigr)\Bigr)\\&\geq \bigl(g(\hat{\nu},i)-g(\hat{x},i)-C\bigr),\end{aligned}$$

so by condition (3) we get the results wanted.

Step 3: We show that $\psi_g((\hat{x},i),z)$ has uniformly lower bound. Note that $\hat{\nu}_u = (F(x_0,x_1,\ldots,x_j)+u,y_1,\ldots,y_r)$, $\hat{\nu} = (F(x_0,x_1,\ldots,x_j),y_1,\ldots,y_r)$, we have

$$\begin{aligned}&-\psi_g\bigl((\hat{x},i),z\bigr)\\&\quad=\Bigl(\int_{R^{r+1}\times \mathrm{E}} P\bigl((\hat{x},i),d(\hat{y},j)\bigr)\bigl[z^{g(\hat{y},j)}-z^{g(\hat{x},i)}\bigr]\Bigr)\Big/(1-z)\\&\quad=\frac{1}{1-z}\sum_{j\in \mathrm{E}}\int_{R^{r+1}} P\bigl((\hat{x},i),d(\hat{y},j)\bigr)\bigl[z^{g(\hat{y},j)}-z^{g(\hat{x},i)}\bigr]\\&\quad=\frac{1}{1-z}\sum_{j\in \mathrm{E}} p_{ij}\Bigl(\int_R z^{g(\hat{\nu}_u,j)}\Phi_j(u)\,du - z^{g(\hat{x},i)}\Bigr)\\&\quad=\frac{1}{1-z}\sum_{j\in \mathrm{E}} p_{ij}\Bigl(\int_R z^{g(\hat{\nu}_u,j)}\bigl[1-z^{g(\hat{\nu},i)-g(\hat{\nu}_u,j)}\bigr]\Phi_j(u)\,du + z^{g(\hat{\nu},i)}-z^{g(\hat{x},i)}\Bigr)\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{1-z}\int_R z^{g(\hat{\nu}_u,j)}\bigl[1-z^{g(\hat{\nu},i)-g(\hat{\nu}_u,j)}\bigr]\Phi_j(u)\,du\\&\quad\leq \frac{1}{1-z}\int_R \bigl[1-z^{h(|u|)}\bigr]\Phi_j(u)\,du\\&\quad\leq 1+\int_R h\bigl(|u|\bigr)\Phi_j(u)\,du\\&\quad\leq 1+C<\infty,\end{aligned}$$

where the third line comes from the inequality $\frac{1-z^x}{1-z}\leq 1+x$, $z\in[0,1)$.

From condition (2) of this theorem, we can get when $(\hat{x},i)\in(A\times \mathrm{M})^c$, $z^{g(\hat{\nu},i)}-z^{g(\hat{x},i)}\leq 0$; when $(\hat{x},i)\in A\times \mathrm{M}$, if $g(\hat{\nu},i)>g(\hat{x},i)$, $\forall z\in[0,1)$, we have $z^{g(\hat{\nu},i)}-z^{g(\hat{x},i)}\leq 0$, if $g(\hat{\nu},i)\leq g(\hat{x},i)$, we have

$$\begin{aligned}\bigl[z^{g(\hat{\nu},i)}-z^{g(\hat{x},i)}\bigr]/(1-z) &= z^{g(\hat{\nu},i)}\bigl[1-z^{g(\hat{x},i)-g(\hat{\nu},i)}\bigr]/(1-z)\\&\leq \bigl[1-z^{g(\hat{x},i)-g(\hat{\nu},i)}\bigr]/(1-z)\\&\leq 1+g(\hat{x},i)-g(\hat{\nu},i)\\&< 1+2\inf_{(\hat{y},j)\in(A\times \mathrm{M})^c} g(\hat{y},j),\end{aligned}$$

where the last inequality lies in the fact that when $(\hat{x},i)\in A\times \mathrm{M}$, $g(\hat{x},i)\leq \inf_{(\hat{y},j)\in(A\times \mathrm{M})^c} g(\hat{y},j)<\infty$.

Therefore, for all $(\hat{x}, i) \in R^{r+1} \times \mathrm{E}$ and $z \in [0, 1)$,

$$\begin{aligned}&\left(\int_{R^{r+1}\times \mathrm{E}} P\big((\hat{x}, i), d(\hat{y}, j)\big)\big[z^{g(\hat{y},j)} - z^{g(\hat{x},i)}\big]\right) \Big/ (1 - z)\\ &\quad < \sum_{j\in \mathrm{E}} p_{ij}\Big(2 + C + 2 \inf_{(\hat{y},j)\in(A\times \mathrm{M})^c} g(\hat{y}, j)\Big)\\ &\quad \le 2 + C + 2C',\end{aligned}$$

where $C' = \inf_{(\hat{y},j)\in(A\times \mathrm{M})^c} g(\hat{y}, j)$. In other words, for all $(\hat{x}, i) \in R^{r+1} \times \mathrm{E}$ and $z \in [0, 1)$, we have

$$\psi_g\big((\hat{x}, i), z\big) > -r\big(2 + C + 2C'\big),$$

so by Lemma 3 we know that $\{Y_t\}$ is nonergodic.

Next we will prove whatever an initial probability distribution function of $\{X_t\}$ is, its probability distribution function will never converge to some probability distribution function. We will use *reductio ad absurdum* to prove it. Suppose that $A \in \mathcal{B}_{r+1}$ and there exists a probability distribution π such that

$$\lim_{t\to\infty} \big\| P(X_t \in A | X_0 = \hat{x}) - \pi(A)\big\|_\tau = 0, \tag{4.1}$$

where $\|\cdot\|_\tau$ is the total variation norm. As a matter of fact, due to the equivalence of norm, we can use any norm here. Since

$$\begin{aligned}&\mathbb{P}(X_t \in A | X_0 = \hat{x})\\ &\quad = \sum_{j\in \mathrm{E}} \mathbb{P}(X_t \in A, Z_t = j | X_0 = \hat{x})\\ &\quad = \sum_{j\in \mathrm{E}}\sum_{i\in \mathrm{E}} \mathbb{P}(X_t \in A, Z_t = j | X_0 = \hat{x}, Z_0 = i)\mathbb{P}(Z_0 = i | X_0 = \hat{x})\\ &\quad = \sum_{i\in \mathrm{E}} \mathbb{P}\big(Y_t \in A \times \mathrm{E} | Y_0 = (\hat{x}, i)\big)\mathbb{P}(Z_0 = i | X_0 = \hat{x})\\ &\quad = \sum_{i\in \mathrm{E}} P^{(t)}\big((\hat{x}, i), A \times \mathrm{E}\big)\mathbb{P}(Z_0 = i | X_0 = \hat{x}).\end{aligned}$$

Define $\pi^*(A \times \mathrm{E}) = \pi(A)$, and it is well known that

$$\pi^*(A \times \mathrm{E}) = \sum_{j\in \mathrm{E}}\sum_{i\in \mathrm{E}} \pi^*\big(A \times \{j\}\big)P(Z_0 = i | X_0 = \hat{x});$$

and therefore we have

$$\lim_{t\to\infty} \big\| P^{(t)}\big((\hat{x}, i), A \times \mathrm{E}\big) - \pi^*(A \times \mathrm{E})\big\|_\tau = 0,$$

but this conflicts with the nonergodicity of the Markov chain $\{Y_t\}$. So, there is no probability distribution function π such that (4.1) holds. This completes the proof. □

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors jointly worked on the results and they read and approved the final manuscript.

Authors' information
School of Science, Jiangxi University of Science and Technology, No. 86, Hongqi Ave., Ganzhou, 341000, Jiangxi, P.R. China.

Acknowledgements
The authors would like to thank the editor and anonymous referees for their valuable suggestions, which greatly improved our paper. This research is supported by the NSF of Jiangxi Province (No. 20132BAB211005), the SF of Jiangxi Provincial Education Department (No. GJJ12356), Key Scientific and Technological Research Project of Department of Education of Henan Province (No. 12B110006), and Foundation of Jiangxi University of Science and Technology (No. jxxj12064).

Received: 25 April 2013 Accepted: 19 June 2013 Published: 5 July 2013

References

1. Jeantheau, T: Strong consistence of estimators for multivariate ARCH models. Econom. Theory **14**, 70-86 (1998)
2. Tjøstheim, D: Estimation in nonlinear time series models. Stoch. Process. Appl. **21**, 251-273 (1986)
3. Fan, J, Yao, Q: Nonlinear Time Series: Nonparametric and Parametric Methods. Springer Series in Statistics. Springer, New York (2003)
4. Hafner, CM, Preminger, A: On asymptotic theory for multivariate GARCH models. J. Multivar. Anal. **100**, 2044-2054 (2009)
5. Lintona, O, Sancetta, A: Consistent estimation of a general nonparametric regression function in time series. J. Econom. **152**, 70-78 (2009)
6. Fernandes, M, Grammig, J: A family of autoregressive conditional duration models. J. Econom. **130**, 1-23 (2006)
7. David, PA: Path-dependence: putting the past into the future of economics. Institute for Mathematical Studies in the Social Sciences. Stanford University (1988)
8. Margolin, G, Barkai, E: Nonergodicity of blinking nanocrystals and other Levy-walk processes. Phys. Rev. Lett. **4**, 1-4 (2005)
9. Basawa, IV, Koul, HL: Asymptotically minimax tests of composite hypotheses for nonergodic type processes. Stoch. Process. Appl. **14**, 41-54 (1983)
10. Basawa, IV, Brockwell, PJ: Asymptotic conditional inference for regular nonergodic models with an application to autoregressive processes. Ann. Stat. **12**, 161-171 (1984)
11. Basawa, IV, Scott, DJ: Asymptotic Optimal Inference for Nonergodic Models. Springer Lecture Notes in Statistics, vol. 17. Springer, New York (1983)
12. Feigin, PD: Conditional exponential families and a representation theorem for asymptotic inference. Ann. Stat. **9**, 597-603 (1981)
13. Budhiraja, A, Ocone, D: Exponential stability in discrete-time filtering for non-ergodic signals. Stoch. Process. Appl. **82**, 245-257 (1999)
14. Durlauf, ST: Nonergodic economic growth. Rev. Econ. Stud. **60**, 349-366 (1993)
15. Goodman, JB, Massey, WA: The nonergodic Jackson network. J. Appl. Probab. **21**, 860-869 (1984)
16. Griffeath, D: Limit theorems for nonergodic set-valued Markov processes. Ann. Probab. **6**, 379-387 (1978)
17. Jacod, J: Parametric inference for discretely observed nonergodic diffusions. Bernoulli **12**, 383-401 (2006)
18. Cline, DBH: Regular variation of order 1 nonlinear AR-ARCH models. Stoch. Process. Appl. **117**, 840-861 (2008)
19. Tweedie, RL: Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. Math. Proc. Camb. Philos. Soc. **78**, 125-136 (1975)
20. Tweedie, RL: Criteria for ergodicity, exponential ergodicity and strong ergodicity of Markov processes. J. Appl. Probab. **18**, 122-130 (1981)
21. Tweedie, RL: Drift conditions and invariant measures for Markov chains. Stoch. Process. Appl. **92**, 345-354 (2001)
22. Choi, BD, Kim, B: Non-ergodicity criteria for denumerable continuous time Markov processes. Oper. Res. Lett. **32**, 574-580 (2004)
23. Kima, B, Leeb, I: Tests for nonergodicity of denumerable continuous time Markov processes. Comput. Math. Appl. **55**, 1310-1321 (2008)
24. Kaplan, M: A sufficient condition for nonergodicity of a Markov chain. IEEE Trans. Inf. Theory **25**, 470-471 (1979)
25. Sheng, ZH, Wang, T, Liu, DL: Stability Analysis of Nonlinear Time Series Models - Ergodic Theory and Applications. Science Press, Beijing (1993)
26. Meyn, SP, Tweedie, RL: Markov Chains and Stochastic Stability. Springer, New York (1993)
27. Tong, H: Nonlinear Time Series: A Dynamical System Approach. Oxford University Press, Oxford (1990)

doi:10.1186/1687-1847-2013-200
Cite this article as: Tang and Wang: **On nonergodicity for nonparametric autoregressive models.** *Advances in Difference Equations* 2013 **2013**:200.