International Journal of Statistics: Advances in Theory and Applications
Vol. 1, Issue 1, 2020, Pages 75-92
Published Online on June 19, 2019

http://jyotiacademicpress.org

EXPONENTIAL GEOMETRIC ERGODICITY FOR NONLINEAR TIME SERIES

THOMAS R. BOUCHER

Department of Mathematics
Texas A&M-Commerce
TX 75428
USA
e-mail: thomas.boucher@tamuc.edu

Abstract

Conditions are explored under which geometric ergodicity of nonlinear autoregressive time series of order p with additive errors can be extended to hold for an exponentially bounded class of functions of the time series. This immediately extends laws of large numbers and central limit theorems to the larger collection of functions of the series.

1. Introduction

Consider a nonlinear autoregressive time series $\{Y_t\}_{t \geq 0}$ of order p with additive errors defined by $Y_t = f(Y_{t-1}, \ldots, Y_{t-p}) + \epsilon_t$, where $f : \mathbb{R}^p \to \mathbb{R}$ is a nonlinear function, p a positive integer, and $\{\epsilon_t\}$ are mean zero, i.i.d. random variables. Ergodicity of the time series follows from ergodicity of the associated general state Markov chain $X_t = (Y_t, \ldots, Y_{t-p+1})'$, which can be expressed $X_t = \phi(X_{t-1}) + \xi_t$, with $\phi(X_{t-1}) =$

2020 Mathematics Subject Classification: 60, 62.
Keywords and phrases: ergodicity, nonlinear time series, Markov chain.
Received February 17, 2017; Revised March 26, 2017

$(f(Y_{t-1}, \ldots, Y_{t-p}), Y_{t-1}, \ldots, Y_{t-p+1})'$ and $\xi_t = (\epsilon_t, 0, \ldots, 0)'$. Ergodicity of the Markov chain can in turn be established by showing there exists a function $V : \mathbb{R}^p \to \mathbb{R}$ so that the chain $\{X_t\}$ satisfies a stochastic drift criterion, such as $E[V(X_1)|X_0 = x] \le \rho V(x)$ for some $\rho < 1$ when x is large (a detailed treatment of stochastic drift criteria is in [5]). Conditions for $\{X_t\}$ satisfying the stochastic drift criterion will follow from conditions on f and $\{\epsilon_t\}$. It will be assumed that $V(x) \ge 1$ and $V(x) \to \infty$ as $\|x\|$ does, with $\|\cdot\|$ denoting the Euclidean norm.

Let π denote the stationary distribution of the ergodic Markov chain, let $E_\pi[\cdot]$ denote expectation with respect to π, and $E_x[\cdot] = E[\cdot|X_0 = x]$. The chain is said to be V-geometrically ergodic when the convergence to π occurs at a geometric rate when normalized by the function V,

$$\sup_x \sup_{|g| \le V} \frac{|E_x[g(X_n)] - E_\pi[g(X_n)]|}{V(x)} \le R\rho^n, \quad R < \infty, \quad 0 < \rho < 1. \tag{1}$$

Also, (1) implies central limit theorems hold for functions g with $|g| \le \sqrt{V}$ and laws of large numbers hold for functions g with $|g| \le V$ (see Theorems 16.1.5 and 17.0.1 in [5]). The purpose of this paper is to provide conditions under which the function V in (1) can be embedded in an exponential $V' = e^{sV}$ or $V' = e^{V^s}$, $s > 0$, where V' also satisfies (1). The benefit of this is that the class of functions g for which laws of large numbers and central limit theorems immediately follow is extended to the class of functions $|g| \le V'$, $|g| \le \sqrt{V'}$, respectively.

2. Results

It will be assumed that $\{X_t\}$ is a *T-chain*, meaning there is a probability distribution $\{a(n)\}$ on the nonnegative integers and a kernel $T(x, A)$ which is a lower semicontinuous function in $x \in \mathbb{R}^p$ for fixed measurable $A \subset \mathbb{R}^p$ with $\sum_n a(n)P(X_n \in A|X_0 = x) \ge T(x, A)$.

It will also be assumed that $\{X_t\}$ is aperiodic and ψ-irreducible. A chain is said to be ψ-irreducible if there exists a measure ψ so that $\psi(A) > 0$ implies $P(\tau_A < \infty | X_0 = x) > 0$ for all x. A chain is said to be aperiodic if there exists a measure ψ with $P(X_1 \in C | X_0 = x) > \nu(C)$ for all $x \in A, \psi(A) > 0$, for all measurable sets C. Aperiodicity and ψ-irreduciblity are standard notions in the study of general state Markov chains; more details can be found in [5], for example.

In some cases, ergodicity may be easier to prove through analysis of the transitions of the k-step chain $\{X_{tk}\}$ with k a positive integer, rather than through analysis of the single-step transitions of $\{X_t\}$. The k-step chain inherits the T-chain property from $\{X_t\}$ if a stronger condition is put on $\{X_t\}$; when $\{X_t\}$ is weak Feller in addition to being ψ-irreducible and aperiodic, ψ being Lebesgue measure, then so is $\{X_{tk}\}$, implying that $\{X_{tk}\}$ is a T-chain. Weak Feller chains map bounded continuous functions to bounded continuous functions, a condition used to verify a chain is a weak Feller chain.

Lemma 1. *Consider a Markov chain $\{X_t\}$. If $\{X_t\}$ is weak Feller, aperiodic, and ψ-irreducible for a measure ψ whose support has a non-empty interior, then $\{X_{tk}\}$ is a ψ-irreducible, aperiodic T-chain for all integers $k \geq 1$.*

Proof. Pick an integer $k \geq 1$. By the weak Feller assumption $E_x[g(X_1)]$ is a bounded continuous function; by induction so is $E_x[g(X_k)]$, implying that $\{X_{tk}\}$ is a weak Feller chain. It is known that $\{X_t\}$ being ψ-irreducible and aperiodic implies $\{X_{tk}\}$ is, and since the support of ψ has non-empty interior, we have by [5], Proposition 5.4.5 and Theorem 6.2.9 that $\{X_{tk}\}$ is a T-chain. □

The following proposition combines the weak Feller assumption and the k-step approach to yield conditions under which the function V in (1) can be embedded in an exponential $V' = e^{sV}$ or $V' = e^{V^s}$, $s > 0$, where V' also satisfies (1). These conditions are applied in Proposition 2 to extend the exponential geometric ergodicity to nonlinear time series.

Proposition 1. *Assume* $\{X_t\}$ *is a* ψ*-irreducible, aperiodic general state weak Feller chain on* $\mathbb{R}^p$, *with the support of* ψ *having a non-empty interior. Suppose* $V \geq 1$ *has* $V(x) \to \infty$ *as* $\| x \| \to \infty$, *is unbounded off compact sets and bounded on them, there is a function* h *bounded on compact sets with* $h(x)/V(x) \to 0$ *as* $\| x \| \to \infty$, *there is a collection of random variables* $\epsilon_1, \ldots, \epsilon_k$ *and function* $\tau(\epsilon_1, \ldots, \epsilon_k, x)$ *with* $E[\tau(\epsilon_1, \ldots, \epsilon_k, x)]/V(x) \to 0$ *as* $\| x \| \to \infty$ *such that*

$$\sup_{\|x\| > M} V(X_k) < \gamma V(x) + h(x) + \tau(\epsilon_1, \ldots, \epsilon_k, x), \tag{2}$$

$$\sup_{\|x\| \leq M} V(X_k) < NV(x) + h(x) + \tau(\epsilon_1, \ldots, \epsilon_k, x), \tag{3}$$

for some integer $k > 0$, *some* $0 < \gamma < 1$, *some* $M, N < \infty$. *Then*

(i) *If* $E[e^{|\tau(\epsilon_1, \ldots, \epsilon_k, x)|^q}] < \infty$ *for some* $q > 0$ *there exists* $0 < s < \min(q, 1)$, *and* $V'(x) = e^{[V(x)]^s}$ *such that* $\{X_t\}$ *is* V'*-geometrically ergodic.*

(ii) *If* $E[e^{q|\tau(\epsilon_1, \ldots, \epsilon_k, x)|}] < \infty$ *for some* $q > 0$ *there exists* $0 < s < \min(q, 1)$, *and* $V'(x) = e^{sV(x)}$ *such that* $\{X_t\}$ *is* V'*-geometrically ergodic.*

Proof. (i) Get k from the assumptions. The proof applies Theorem 4 in [3] to $\{X_{tk}\}$, which theorem states that if $\{X_{tk}\}$ is an aperiodic, ψ-irreducible T-chain, if V is locally bounded with $V(x) \to \infty$ as $\| x \| \to \infty$, if there exists a random variable $W(x)$ such that $V(X_k) \leq W(x)$

whenever $X_0 = x$, if $|\log[W(x)/V(x)]| + e^{([W(x)]^r - [V(x)]^r)}$ is uniformly integrable for some $r > 0$, and if $\limsup_{\|x\|\to\infty} E_x[\log(W(x)/V(x))] < 0$, then there exists $s > 0$ and $V'(x) = e^{[V(x)]^s}$ such that $\{X_{tk}\}$ is V'-geometrically ergodic. That $\{X_{tk}\}$ is a ψ-irreducible, aperiodic T-chain follows from the weak Feller assumption on $\{X_t\}$ and Lemma 1.

Finally, Lemma 2 in [1] adds that if for some integer $0 < k < \infty$ and all $M < \infty$ it holds that $\limsup_{V(x)\to\infty} E_x[V(X_k)]/V(x) < 1$, $\sup_{V(x)\le M} E_x[V(X_k)] < \infty$, $\sup_x E_x[V(X_1)]/V(x) < \infty$, and the sets $\{x : V(x) \le M\}$ are petite, then $\{X_t\}$ is V-uniformly ergodic. This will be used to show V'-geometric ergodicity of $\{X_t\}$ follows from that of $\{X_{tk}\}$.

From the assumption (2), $V(X_k) < \gamma V(x) + h(x) + \tau(\epsilon_1, \ldots, \epsilon_k, x)$ for $\|x\| > M$ and there exists $\epsilon > 0$ so that $\limsup_{\|x\|\to\infty} E_x[V(X_k)]/V(x) < 1 - \epsilon$ and $\gamma + \epsilon < 1$. Define $W_k(x) := V(X_k) + \epsilon V(x) \ge V(X_k)$. Then by Jensen's inequality $\limsup_{\|x\|\to\infty} E_x[\log(W_k(x)/V(x))] < 0$. Note $V \ge 1$ implies $W_k(x)/V(x) > \epsilon > 0$ so that $\log(W_k(x)/V(x))$ is defined for all x. Pick $\delta > 0$ and $Q = Q(\delta) < \infty$ so that $W_k(x)/V(x) > Q$ implies $[\log(W_k(x)/V(x))]^{1+\delta} < W_k(x)/V(x)$. Then $\sup_x E_x[|\log(W_k(x)/V(x))|^{1+\delta}] < \sup_x E_x[V(X_k)]/V(x) + \epsilon + [\log(Q)]^{1+\delta} < \infty$, implying $|\log(W_k(x)/V(x))|$ is uniformly integrable.

Get q from the assumptions and choose r so that $r^{1+\delta} < \min(q, 1)$. Since V, h are assumed bounded on compact sets and $Ee^{|\tau(\epsilon_1, \ldots, \epsilon_k, x)|^q} < \infty$ then from the assumptions (2), (3) $\|x\| \le M$ implies $E_x(e^{[W_k(x)]^r - [V(x)]^r})^{1+\delta} \le E(e^{[(N+\epsilon)V(x) + h(x) + |\tau(\epsilon_1, \ldots, \epsilon_k, x)|]^r - [V(x)]^r})^{1+\delta}$ is bounded. Since $Ee^{|\tau(\epsilon_1, \ldots, \epsilon_k, x)|^q} < \infty$, $\gamma + \epsilon < 1$ and $h(x)/V(x) \to 0$ as $\|x\| \to \infty$ then when $\|x\| > M$ it holds that

$E_x(e^{[W_k(x)]^r - [V(x)]^r})^{1+\delta} \le E(e^{((\gamma+\epsilon)[V(x)] + [h(x)] + |\tau(\epsilon_1,\ldots,\epsilon_k,x)|)^r - [V(x)]^r})^{1+\delta} < \infty.$

Thus, $\sup_x E_x(e^{[W_k(x)]^r - [V(x)]^r})^{1+\delta} < \infty$ and $e^{[W_k(x)]^r - [V(x)]^r}$ is uniformly integrable.

Since $|\log(W_k(x)/V(x))|$ and $e^{[W_k(x)]^r - [V(x)]^r}$ are each uniformly integrable, so is the sum. The function $W_k(x)$ satisfies the conditions of Theorem 4 in [3] stated above; thus there exist $0 < s < q$ and $V'(x) = e^{[V(x)]^s}$ such that $\{X_{tk}\}$ is V'-uniformly ergodic. Since V is bounded on compact sets the set $A = \{x : \|x\| \le M\}$ is compact and thus petite since $\{X_{tk}\}$ is a T-chain. It also follows from assumptions (2), (3) that $\sup_{x \in A^C} E_x[e^{V(X_k)^s}]/e^{V(x)^s} < 1$, $\sup_{x\in A} E_x[e^{V(X_k)^s}] < \infty$, and $\sup_x E_x[e^{V(X_k)^s}]/e^{V(x)^s} < \infty$; then by Lemma 2 in [1], $\{X_t\}$ is V'-geometrically ergodic as well.

(ii) Similarly, from ([3], Theorem 3), if $|W(x) - V(x)| + e^{r[W(x)-V(x)])}$ is uniformly integrable for some $r > 0$, and if $\limsup_{\|x\|\to\infty} E_x[W(x) - V(x)] < 0$, then there exists $s > 0$ and $V'(x) = e^{sV(x)}$ such that $\{X_t\}$ is V'-geometrically ergodic.

As in (i) the assumption (2) implies there exists $\epsilon > 0$ with $\gamma + \epsilon < 1$, γ from (2), so that $\limsup_{\|x\|\to\infty} E_x[V(X_k)]/V(x) < 1 - \epsilon$. Define $W_k(x) := V(X_k)$. Then $\limsup_{\|x\|\to\infty} E_x[W_k(x) - V(x)] < -\epsilon V(x) < 0$. The assumption $h(x)/V(x) \to 0$ as $\|x\| \to \infty$ implies that for large enough M, $\|x\| > M$ implies $h < \epsilon V$, implying in turn $W_k(x) - V(x) < (\gamma + \epsilon < 1)V(x) + |\tau(\epsilon_1, \ldots, \epsilon_k, x)|^q < |\tau(\epsilon_1, \ldots, \epsilon_k, x)|^q$. Pick $\delta > 0$ so that $y > (1+\delta)\ln y$ for $y > 0$. Since by assumption $E[e^{q|\tau(\epsilon_1,\ldots,\epsilon_k,x)|}] < \infty$, then $E[|\tau(\epsilon_1, \ldots, \epsilon_k, x)|^{(1+\delta)q}] < \infty$ from which it follows that $\sup_x E[(W_k(x) - V(x))^{1+\delta}] < \infty$; therefore $W_k(x) - V(x)$ is uniformly integrable. The assumptions and similar arguments also imply $\sup_x E_x[e^{r(1+\delta)(W_k(x)-V(x))}] < \infty$ for $r < \min(q, 1)$. Thus, $e^{r[W_k(x) - V(x)]}$ is uniformly integrable.

The function $W_k(x)$ satisfies the conditions for [3], Theorem 3 stated above. Again, that $\{X_{tk}\}$ is a T-chain follows from the weak Feller assumption on $\{X_t\}$ and Lemma 1. The set $A = \{x : \|x\| \leq M\}$ is petite. Thus there exist $0 < s < q$ and $V'(x) = e^{sV(x)}$ such that $\{X_{tk}\}$ is V'-uniformly ergodic. It follows from the assumptions (2), (3) that $\sup_{x\in A^C} E_x[e^{sV(X_k)}]/e^{sV(x)} < 1$, $\sup_{x\in A} E_x[e^{sV(X_k)}] < \infty$, and $\sup_x E_x[e^{sV(X_k)}] / e^{sV(x)} < \infty$; then by Lemma 2 in [1], $\{X_t\}$ is V'-geometrically ergodic.

□

The following proposition gives conditions on the time series that guarantee the assumptions of Proposition 1 are satisfied; exponential geometric ergodicity of the time series then follows from Proposition 1. Applications often involve $f(\cdot)$ being piecewise continuous or well approximated by a linear or piecewise linear function. For time series with additive errors, and using norm-like test functions V, the conditions for geometric ergodicity with an exponential V are stated as an appropriate exponential stability condition on the errors $\{\epsilon_t\}$ in addition to an appropriate stability condition on the skeleton $x_t = \phi(x_{t-1})$. Let $\rho(\cdot)$ denote the eigenvalue of maximum modulus of a matrix.

Proposition 2. *Suppose the distribution of* $\{\epsilon_t\}$ *is absolutely continuous with respect to Lebesgue measure and* $E_e^{|\epsilon_t|^q} < \infty$ *(or* $Ee^{q|\epsilon_t|} < \infty$*) for some* $q > 0$.

(i) *If f is sublinear, bounded on compact sets, finite at each x, and for some positive integer* k, $\limsup_{\|x\|\to\infty}\|\phi^{(k)}(x)\| / \|x\|^{1/p} < 1$, *where* $\phi^{(k)}(x)$ *is the k-fold composition of* ϕ *with itself, or*

(ii) *If f is continuous and everywhere differentiable,* $\phi(\cdot)$ *has Jacobian* $J(\cdot)$ *bounded on compact sets, finite at each x and if* $X_0 = x$ *implies* $X_k = J(x)x + c(x) + \xi_k$, k *a positive integer, where* $c(x) = O(\|x\|^r)$, $r < 1$, *and* $c(x)$ *is Lipschitz and finite at each x, and* $\limsup_{\|x\|\to\infty} \rho(J(x)) < 1$, *or*

(iii) *If there is a collection* $A_1, \ldots, A_m$ *of* $p \times p$ *matrices, and regions* $R_1, \ldots, R_m$ *that partition* $\mathbb{R}^p$ *such that* $X_t = \sum_{i=1}^m A_i I(X_{t-1} \in R_i) X_{t-1} + c(x) + \xi_t$, *where* $c(x) = O(\|x\|^r)$, $r < 1$, *and* $c(x)$ *is Lipschitz and finite at each x, and if* $\max_{i \in 1, \ldots, m} \rho(A_i) < 1$, *then*

there exists a norm $\|\cdot\|_v$, $V(x) = 1 + \|x\|_v^q$ *for some* $0 < s < q < 1$, *and* $V'(x) = e^{(V(x))^s}$ *(or* $V'(x) = e^{sV(x)}$*) such that* $\{X_t\}$ *is* V'*-geometrically ergodic.*

Proof. The assumptions on ϵ_t imply $\{X_t\}$ has distribution ν absolutely continuous with respect to Lebesgue measure λ and each set of assumptions implies $\phi(x)$ is finite at each $x \in \mathbb{R}^p$. Also, for $A \in \mathcal{B}(\mathbb{R}^p)$ since $\lambda(A) > 0$ implies $\lambda(A + \phi(x)) > 0$ by translation invariance of Lebesgue measure, then since ν is absolutely continuous with respect to λ and $\phi(x)$ is finite it holds that $P(x, A) = \int_A \nu(y - \phi(x)) \lambda(dy) > 0$, so that $\{X_t\}$ is λ-irreducible. For $C \in \mathcal{B}(\mathbb{R}^p)$ let $A = \{x \in \mathcal{B}(\mathbb{R}^p) : \nu(C + \phi(x)) > 0\}$. Suppose $\lambda(A) = 0$. Then $P(x, C) = 0$ for $x \in A^C$, which by λ-irreducibility of $\{X_t\}$ implies $\lambda(C) = 0$, which violates the assumption $\lambda(A) = 0$ when $C = A^C$, so that by contradiction $\lambda(A) > 0$ which implies $P(x, C) > 0$ and so $\{X_t\}$ is aperiodic.

(i) Consider $\{x_n\}$ with $x_n = (x_1^n, \dots, x_p^n)' \to x = (x_1, \dots, x_p)'$ then since f is bounded on compact sets there exists $C < \infty$ so that $\|\phi(x) - \phi(x_n)\| = \|(f(x), x_1, \dots, x_{p-1})' - (f(x_n), x_1^n, \dots, x_{p-1}^n)\| \le C\|x - x_n\|$. By absolute continuity $\lambda(\{x : \|x - x_n\| < \delta\}) \to 0$ implies $\nu(\{x : \|x - x_n\| < \delta\}) \to 0$. Then if g is a bounded continuous function, by continuity of g, absolute continuity, and bounded convergence $|E_x[g(X_1)] - E_{x_n}[g(X_1)]| = |\int g(y)\nu(dy - \phi(x)) - \int g(y)\nu(dy - \phi(x_n))| = |\int g(z + \phi(x))\nu(dz) - \int g(z + \phi(x_n))\nu(dz)| \to 0$ as $\|x - x_n\| \to 0$ so that $E_x[g(X_1)]$ is a bounded continuous function and $\{X_t\}$ is a weak Feller chain with the irreducibility measure ψ being Lebesgue measure on $\mathbb{R}^p$; thus its support has a nonempty interior. By Lemma 1 then, $\{X_{tk}\}$ is an aperiodic, λ-irreducible T-chain for integers $k \ge 1$.

Let $\phi^{(n)}(x)$ denote the n-fold composition of ϕ with itself. By the contraction theorem, for all y with $|y| < 1$ it holds that $\phi^{(n)}(y) \to \alpha < 1$ as $n \to \infty$, so that for $\epsilon > 0$ there exists $N < \infty$ with $n \ge N$ implying $\phi^{(n)}(y) < 1 - \epsilon$. The assumption $\limsup_{\|x\| \to \infty} \|\phi^{(k)}(x)\| / \|x\|^{1/p} < 1$ implies there is a positive integer k and $M < \infty$ so that $\|\phi^{(k)}(x)\| / \|x\|^{1/p} < 1$ for $\|x\| > M$. Since f is sublinear so is ϕ and note from sublinearity of ϕ that at each x with $\|x\| > M$, $\phi^{(n+k)}(x) / \|x\|^{1/p} = \phi^{(n)}(\phi^{(k)}(x) / \|x\|^{1/p}) < 1 - \epsilon$ for integers $n \ge N$. By subadditivity of ϕ, $\phi(X_{n+k}) \le \phi^{(n+k)}(x) + \sum_{i=0}^{n+k-1} E[\phi^{(i+1)}(\xi_{n+k-i})]$. Define $\tau(\epsilon_1, \dots, \epsilon_{n+k+1}) = \sum_{i=0}^{n+k-1} E[\phi^{(i+1)}(\xi_{n+k-i}) + \xi_{n+k+1}]$ and note at each x with $\|x\| > M$, for integers $n \ge N$, that $\|X_{n+k+1}\| < (1 - \epsilon)\|x\| + \tau(\epsilon_1, \dots, \epsilon_{n+k+1})$. Let $V(x) = 1 + \|x\|$ and assumption (2) of Proposition 1 is satisfied. Since f is assumed bounded on compact sets there exists $N < \infty$ so that $\|X_{n+k+1}\| < N\|x\| + \tau(\epsilon_1, \dots,$

ϵ_{n+k+1}) and assumption (3) of Proposition 1 is satisfied. It was established above that $\{X_{tk'}\}$ is an aperiodic, λ-irreducible T-chain, where $k' = n + k$. Then $\{X_t\}$ is V'-geometrically ergodic by Proposition 2 with $V'(x) = e^{(V(x))^s}$ (or $V'(x) = e^{sV(x)}$) for $0 < s < q$.

(ii) Consider $\{x_n\}$ with $x_n = (x_1^n, \ldots, x_p^n)' \to x = (x_1, \ldots, x_p)'$ then since $J(\cdot)$ is bounded on compact sets and $c(\cdot)$ is Lipschitz there exists $C < \infty$ so that $\|\phi(x) - \phi(x_n)\| = \|J(x)x - c(x) - [J(x_n)x_n - c(x_n)]\| \leq C\|x - x_n\|$. By absolute continuity $\lambda(\{x : \|x - x_n\| < \delta\}) \to 0$ implies $\nu(\{x : \|x - x_n\| < \delta\}) \to 0$. Then if g is a bounded continuous function, by continuity of g, absolute continuity, and bounded convergence $|E_x[g(X_1)] - E_{x_n}[g(X_1)]| = |\int g(y)\nu(dy - \phi(x)) - \int g(y)\nu(dy - \phi(x_n))| = |\int g(z + \phi(x))\nu(dz) - \int g(z + \phi(x_n))\nu(dz)| \to 0$ as $\|x - x_n\| \to 0$ so that $E_x[g(X_1)]$ is a bounded continuous function and $\{X_t\}$ is a weak Feller chain with the irreducibility measure ψ being Lebesgue measure on $\mathbb{R}^p$; thus its support has a nonempty interior. By Lemma 1 then, $\{X_{tk}\}$ is an aperiodic, λ-irreducible T-chain for integers $k \geq 1$.

It is known ([1], Lemma 4 for instance) the assumptions imply there exists $M < \infty$ and a norm $\|\cdot\|_v$ with $\|x\|_v \to \infty$ as $\|x\| \to \infty$, and $\rho < 1$ so that $\|J(x)x\|_v \leq \rho\|x\|_v$ for $\|x\| > M$, and if $c(x) = O(\|x\|^r)$, then $c(x) = O(\|x\|_v^r)$. Suppose w.l.o.g. that $q < 1$. Define $V(x) = 1 + \|x\|_v^q$, then since $\limsup_{\|x\| \to \infty} \rho(J(x)) < 1$ there is $M < \infty$ with $\|x\| > M$ implying $V(X_k) < \rho(1 + \|x\|_v^q) + (1 - \rho) + \|c(x)\|_v^q + \|\xi_k\|_v^q$. Since $J(\cdot)$ is assumed bounded on compact sets there is $N < \infty$ with $\|x\| \leq M$ implying $V(X_k) < N(1 + \|x\|_v^q) + (1 - N) + \|c(x)\|_v^q + \|\xi_k\|_v^q$ and thus $\{X_t\}$ is

V'-geometrically ergodic by Proposition 2 with $\gamma = \rho$, $h(x) = \max((1 - N) + \|c(x)\|_v^q, (1 - \rho) + \|c(x)\|_v^q)$, $\tau(\epsilon_1, \dots, \epsilon_k, x) = \|\xi_k\|_v^q = \|(\epsilon_k, 0, \dots, 0)\|_v^q$, $V'(x) = e^{(V(x))^s}$ (or $V'(x) = e^{sV(x)}$).

(iii) Consider $\{x_n\}$ with $x_n = (x_1^n, \dots, x_p^n)' \to x = (x_1, \dots, x_p)'$ then since $c(\cdot)$ is Lipschitz and $\max_{i \in 1, \dots, m} \rho(A_i) < 1$ there exists $C < \infty$ so that $\|\phi(x) - \phi(x_n)\| = \|\sum_{i=1}^m A_i I((x) \in R_i)x + c(x) - [\sum_{i=1}^m A_i I((x_n) \in R_i) x_n + c(x_n)]\| \le C\|x - x_n\|$. By absolute continuity $\lambda(\{x : \|x - x_n\| < \delta\}) \to 0$ implies $\nu(\{x : \|x - x_n\| < \delta\}) \to 0$. Then if g is a bounded continuous function, by continuity of g, absolute continuity, and bounded convergence $|E_x[g(X_1)] - E_{x_n}[g(X_1)]| = |\int g(y)\nu(dy - \phi(x)) - \int g(y)\nu(dy - \phi(x_n))| = |\int g(z + \phi(x))\nu(dz) - \int g(z + \phi(x_n))\nu(dz)| \to 0$ as $\|x - x_n\| \to 0$ so that $E_x[g(X_1)]$ is a bounded continuous function and $\{X_t\}$ is a weak Feller chain with the irreducibility measure ψ being Lebesgue measure on $\mathbb{R}^p$; thus its support has a nonempty interior. By Lemma 1 then, $\{X_{tk}\}$ is an aperiodic, λ-irreducible T-chain for integers $k \ge 1$.

Similar to (ii) the assumptions imply there exists $M < \infty$ and a norm $\|\cdot\|_v$ with $\|x\|_v \to \infty$ as $\|x\| \to \infty$, and $\rho < 1$ so that $\|A_i x\|_v \le \rho\|x\|_v$ for each i, and if $c(x) = O(\|x\|^r)$, then $c(x) = O(\|x\|_v^r)$. Suppose w.l.o.g. that $q < 1$. Define $V(x) = 1 + \|x\|_v^q$, then since $\max_{i \in 1, \dots, m} \rho(A_i) < 1$, $V(X_k) < \rho(1 + \|x\|_v^q) + (1 - \rho) + \|c(x)\|_v^q + \|\xi_k\|_v^q$. Then $\{X_t\}$ is V'-geometrically ergodic by Proposition 2 with $\gamma = \rho$, $h(x) = 1 + \|c(x)\|_v^q$, $\tau(\epsilon_1, \dots, \epsilon_k, x) = \|\xi_k\|_v^q$, $V'(x) = e^{(V(x))^s}$ (or $V'(x) = e^{sV(x)}$). □

3. Applications

If f is continuous and everywhere differentiable a Taylor expansion of $\phi(X_0)$ around $x = (x_1, \ldots, x_p)'$ yields $X_1 = J(x)X_0 + c(x, X_0) + \xi_t$, where $J(x)$ is the Jacobian of $\phi(\cdot)$ evaluated at x, $\xi_t = \epsilon_t(1, 0, \ldots, 0)$ and $c(x, X_0) = \phi(x) - J(x)x + R_\Phi(x, X_0)$, with $R_\phi(x, X_0)$ being the remainder of the Taylor expansion. Conditioning on $X_0 = x$ gives $E_x[R_\phi(x, X_0)] = 0$ and $E_x[X_1] = J(x)x + c(x) + \xi_t$ with $c(x) = \phi(x) - J(x)x$. With an appropriate condition on $c(x)$ that guarantees $c(x)$ is small when $\|x\|$ is large, a condition for stability would then be $\limsup_{\|x\|\to\infty} \rho(J(x)) < 1$, as is stated in Proposition 2(ii).

Example. Consider the EXPAR(1) process $X_t = (\alpha + \beta \exp\{-X_{t-1}^2\}) X_{t-1} + \epsilon_t$. The function $f(x) = (\alpha + \beta \exp\{-x^2\})x$ has derivative $f'(x) = \alpha + \beta \exp\{-x^2\}(1 - 2x^2)$, so $c(x) = f(x) - f'(x) = [\alpha + \beta \exp\{-x^2\} - (\alpha + \beta \exp\{-x^2\}(1 - 2x^2))]x$ satisfies the condition $c(x) = O(\|x\|^r)$ for some $r < 1$. Suppose $E^{|\epsilon_t|^q} < \infty$ for some $q > 0$. For geometric ergodicity, it is required that $\limsup_{\|x\|\to\infty} |f'(x)| < 1$ which is true if $|\alpha| < 1$. As a matter of convenience, geometric ergodicity could then be shown with function $V(x) = 1 + |x|$ since it is simple enough to show that $E[V(X_1) | X_0 = x] < \rho V(x)$ for some $0 < \rho < 1$ for large enough x. However, this may limit the class of functions $g \le V$ for which it immediately follows that laws of large numbers and central limit theorems hold. Rather than having to prove geometric ergodicity directly using a less tractable function, using Proposition 2 geometric ergodicity holds for $V' = e^{sV} = e^{s(1+|x|)}$, $s < q$, so that moments of all orders exist, and asymptotic results for partial sums of functions g with $|g| \le V$. □

In particular, requiring that $\{Y_t\}$ be asymptotically threshold-like allows us to exchange the smoothness conditions on f for stability conditions on the piecewise linear part. Suppose $X_t = \sum_{i=1}^{m} A_i I(X_{t-1} \in R_i) X_{t-1} + c(x) + \xi_t$ and with conditions on $c(x)$ as in Proposition 2 above. Then $\{X_t\}$ is asymptotically piecewise linear on each of the regions R_i.

Example. The threshold autoregressive process of order p, delay d, $d \leq p$ is

$$Y_t = \phi_1^{(i)} Y_{t-1} + \cdots + \phi_p^{(i)} Y_{t-p} + \epsilon_t, \quad Y_{t-d} \in I_i, \quad i = 1, \ldots, s,$$

where $\{I_i\}$, $i = 1, \ldots, s$ forms a partition of $\mathbb{R}$. Embed Y_t in the chain $X_t = \sum_{i=1}^{m} A_i I(X_{t-1} \in R_i) X_{t-1}$ with $X_t = (Y_t, \ldots, Y_{t-p+1})$, the R_i being regions in $\mathbb{R}^p$ determined by the partition $\{I_i\}$, $i = 1, \ldots, s$, and A_i the matrix with $\phi_1^{(i)}, \ldots, \phi_p^{(i)}$ in the first row, 1 on the subdiagonal, and 0 elsewhere. Suppose the ϵ_t have a continuous density that is positive everywhere and $\max_{i \in 1, \ldots, m} \rho(A_i) < 1$. Ergodicity is often demonstrated with a norm-like V such as $V(x) = 1 + \|x\|$. To enable geometric ergodicity with an exponential function $e^{sV(x)}$ or $e^{[V(x)]^s}$, an appropriately strong moment condition on the error distribution is required, a stronger error condition than is usually given when the concern is simply to show ergodicity. For example, if the ϵ_t also have $Ee^{q|\epsilon_t|} < \infty$ for some $q > 0$, then the assumptions of Proposition 2 are satisfied and the process is geometrically ergodic with $V'(x) = e^{sV(x)}$ for some $0 < s < q$. □

As a more substantial application of the results they are applied to the threshold autoregressive process

$$Y_t = \phi_1^{(i)} Y_{t-1} + \cdots + \phi_p^{(i)} Y_{t-p} + \epsilon_t, \quad Y_{t-d} \in I_i, \quad i = 1, \ldots, s,$$

where $\{R_i\}_{i=1}^s$ is the partition of $\mathbb{R}^p$ described in the example above. Suppose d, s, R_i and p are known. If $\epsilon_t \sim N(0, \sigma^2)$, then the maximum likelihood (and conditional least squares) estimators are

$$\hat{\phi}^{(i)} = \left[\sum_{t=1}^{n} X_{t-1} X'_{t-1} I(X_{t-1} \in R_i)\right]^{-1} \sum_{t=1}^{n} X_{t-1} Y_t I(X_{t-1} \in R_i), \quad i = 1, \ldots, s.$$

$$\hat{\sigma}^2 = \frac{1}{n - sp} \sum_{i=1}^{s} \left(X_t - \hat{\phi}^{(i)} X_{t-1} I(X_{t-1} \in R_i)\right)^2.$$

Asymptotic properties of these estimators have been established (see [2], [4], [6], for example). For threshold autoregressive processes asymptotic results for the $\hat{\phi}_j^{(i)}$ are known to exist when $E\epsilon_t^2 < \infty$, while asymptotic results for $\hat{\sigma}^2$ are known to exist when $E\epsilon_t^4 < \infty$. However, since the ϵ_t are often assumed i.i.d. $N(0, \sigma^2)$ there is little harm in supposing $Ee^{|\epsilon_t|^q} < \infty$ or $Ee^{q|\epsilon_t|} < \infty$ for some $q > 0$, in which case Proposition 2 implies $\{X_t\}$ is V'-geometrically ergodic with $V'(x) = e^{sV(x)}$ for $0 < s < q$.

In this example, limit theorems for the $TAR(p; d; s)$ process are derived using the results in this paper. The parameter estimates often involve vector-valued functions, so a multivariate version of the central limit theorem for V-geometrically ergodic Markov chains ([5], Theorem 17.0.1) is needed. Let $g : \mathbb{R}^p \to \mathbb{R}^p$ be a vector-valued function and let $V \geq 1$ be a real-valued function with the Markov chain being V-geometrically ergodic. Suppose g, V are such that for a vector a there exist constants $K_1 = K_1(a)$, $K_2 = K_2(a)$ so that $(a'g)'(a'g) \leq K_1 V' + K_2$. Then by ([5], Theorem 17.0.1), the partial sum $S_n(a'g)$ obeys the central limit theorem; but $S_n(a'g) = a'S_n(g)$, implying that $a'S_n(g)$ obeys the central limit theorem. By the Cramer-Wald device then, $S_n(g)$ obeys the central limit theorem.

Proposition 2 is used to imply geometric ergodicity with $V'(x) = e^{sV(x)}$ for $s < q$. The multivariate extension of the central limit theorem then implies limit theorems for vector-valued functions.

Proposition 3. *Suppose* ϵ_t *has a density which is continuous and everywhere positive and* $Ee^{|\epsilon_t|^q} < \infty$ *or* $Ee^{q|\epsilon_t|} < \infty$ *for some* $q > 0$. *Then the least squares estimators of the* $TAR(p;\, d;\, s)$ *process are consistent and obey the central limit theorem*, *i.e.*,

(i) $\hat{\phi}^{(i)} \to \phi^{(i)}$ *with probability one.*

(ii) $\hat{\sigma}^2 \to \sigma^2$ *with probability one.*

(iii) $\sqrt{n_i}(\hat{\phi}^{(i)} - \phi^{(i)}) \to N(0, \sigma^2 \Gamma(i)^{-1})$ *in distribution*, *where* $\Gamma(i) = E(X_t X_t^T I(X_t \in R_i))$.

Proof. (i) By Proposition 2, we have that $\{X_t\}$ is V'-geometrically ergodic with $V'(x) = e^{[V(x)]^s}$ or $V'(x) = e^{s[V(x)]}$, $0 < s < \min(q, 1)$ and $V(x) \to \infty$ as $\|x\|$ does. Let $n_i = \sum_{t=1}^{n} I(X_{t-1} \in R_i)$. Since $Y_t = \phi^{(i)} X_{t-1} + \epsilon_t$, then

$$\hat{\phi}^{(i)} = \phi^{(i)} + \left[\frac{1}{n_i}\sum_{t=1}^{n} X_{t-1} X_{t-1}' I(X_{t-1} \in R_i)\right]^{-1} \frac{1}{n_i}\sum_{t=1}^{n} X_{t-1}\epsilon_t I(X_{t-1} \in R_i).$$

Let $g_k((x_1, \ldots, x_p)') = x_k$. Then $|g_k(x)g_j(x)| < V'(x)$ for large x which implies ([5], Theorem 17.0.1) the LLN applies to $\frac{1}{n_i} S_n[g_k(X_{t-1})g_j(X_{t-1}) I(X_{t-1} \in R_i)]$. By V'-geometric ergodicity of $\{X_t\}$, $E[X_{t-1}X'_{t-1}I(X_{t-1} \in R_i)] = \Gamma(i)$ and we have $\frac{1}{n_i} S_n[g_k(X_{t-1})g_j(X_{t-1})I(X_{t-1} \in R_i)] \to \Gamma(i)_{kj}$ almost surely, from which it follows that $\frac{1}{n_i} \sum_{t=1}^{n} X_{t-1}X'_{t-1}I(X_{t-1} \in R_i)$ converges to $\Gamma(i)$ almost surely.

Likewise, the V'-geometric ergodicity of $\{X_t\}$ implies the SLLN for $\frac{1}{n_i} \sum X_{t-1}I(X_{t-1})$. Since $E(X_{t-1}\epsilon_t) = E(X_{t-1})E(\epsilon_t) = 0$, we have that $\hat{\phi}^{(i)} \to \phi^{(i)}$ with probability one by Slutsky's theorem.

(ii) By similar arguments, the SLLN implies $\hat{\sigma}^2 \to \sigma^2$ with probability one.

(iii) Note that

$$\sqrt{n_i}(\hat{\phi}^{(i)} - \phi^{(i)}) = \left[\frac{1}{n_i} \sum_{t=1}^{n} X_{t-1}X'_{t-1}I(X_{t-1} \in R_i)\right]^{-1} \frac{1}{\sqrt{n_i}} \sum_{t=1}^{n} X_{t-1}\epsilon_t I(X_{t-1} \in R_i).$$

Also, by V'-geometric ergodicity of $\{X_t\}$ the SLLN implies that $\frac{1}{\sqrt{n_i}} \sum \Gamma^{-1}(i)X_{t-1}\epsilon_t I(X_{t-1})$ has limiting variance

$$\Gamma(i)^{-1} E[(X_{t-1}I(X_{t-1} \in R_i)\epsilon_t)(X_{t-1}I(X_{t-1} \in R_i)\epsilon_t)']\Gamma(i)^{-1} = \sigma^2\Gamma(i)^{-1}.$$

Clearly, for $g(x) = xI(x \in R_i)$, $a'g \leq K_1V' + K_2$ for suitable $K_1, K_2 < \infty$ given a vector a and so $(1/\sqrt{n_i})\sum \Gamma^{-1}(i)X_{t-1}\epsilon_t I(X_{t-1})$ obeys the CLT, converging to $N(0, \sigma^2\Gamma^{-1}(i))$. Since $\left[\frac{1}{n_i} \sum_{t=1}^{n} X_{t-1}X'_{t-1}I(X_{t-1} \in R_i)\right]^{-1} \to \Gamma^{-1}(i)$ this implies $\sqrt{n_i}(\hat{\phi}^{(i)} - \phi^{(i)})$ converges to $N(0, \sigma^2\Gamma_p^{-1})$ by Slutsky's once again. □

The TAR and EXPAR models discussed in the examples are contained in the more general functional-coefficient autoregressive (FCAR) model

$$Y_t = a_1(Y_{t-d})Y_{t-1} + a_2(Y_{t-d})Y_{t-2} + \dots + a_p(Y_{t-d})Y_{t-p} + \sigma(Y_{t-d})\epsilon_t.$$

The coefficients $a_i(\ )$ are unknown functions and are estimated using nonparametric methods such as kernel-weighted local linear regression. Ergodicity conditions for FCAR can be established by analyzing the associated chain $X_t = \sum_{i=1}^{m} A(X)X_{t-1} + \xi_t$ with $X_t = (Y_t, \dots, Y_{t-p+1})$, $A(X)$ the matrix with $a_1(X), \dots, a_p(X)$ in the first row, 1 on the subdiagonal, and 0 elsewhere, and $\xi_t = (\epsilon_t, 0, \dots, 0)'$. One condition for ergodicity ([4], Theorem 8.1) is that the matrix $A(x)$ have $\sup_x \rho(A(x)) < 1$, which is also the condition implied by Proposition 2 (iii). However, this is stronger than is necessary; for example, using Proposition 2, ergodicity will follow if there is an $M < \infty$ with $\sup_{x>M} \rho(A(x)) < 1$. Laws of large numbers and central limit theorems for the estimators of $a_1(X), \dots, a_p(X)$ can be proved assuming a hyperbolic mixing condition on the process such as $\sum n^c \alpha(n)^{1-2/\delta}$ for some $\delta > 2$ and $c > 1 - 2/\delta$ (for example, see [4], Theorems 8.2 and 8.3). With geometric ergodicity as implied by Proposition 2, a geometric rate of mixing is assured, and the exponential geometric ergodicity implied by Proposition 2 enlarges the collection of functions of the process to which laws of large numbers and central limit theorems directly apply.

References

[1] T. R. Boucher and D. B. H. Cline, Stability of cyclic threshold autoregressive time series models, Statist. Sinica 17(1) (2007), 43-62.

[2] K. S. Chan, Consistency and limiting distributions of a least squares estimator of a threshold autoregressive model, Annals of Statistics 21 (1993), 520-533.

[3] D. B. H. Cline and H. Pu, Stability of nonlinear time series: What does noise have to do with it? Lecture Note - Monograph series, Selected Proceedings of the Symposium on Inference for Stochastic Processes 37 (2001), 151-170.

[4] J. Fan and Q. Yao, Nonlinear Time Series: Nonparametric and Parametric Methods, Springer-Verlag, New York, 2003.

[5] S. P. Meyn and R. L. Tweedie, Markov Chains and Stochastic Stability, Springer, London, 1993.

[6] H. Tong, Nonlinear Time Series Analysis: A Dynamical System Approach, Oxford University Press, London, 1990.

■