Geometric Ergodicity and Mixing Properties of Autoregressive Models

Eckhard Liebscher

University of Applied Sciences Merseburg

Germany

contents:

1. Geometric ergodicity of Markov chains
2. Mixing conditions
3. Introduction to AR-models
4. Vector-ARCH-models
5. Scalar-ARCH-models
6. Threshold-models
Why is it interesting to know that the time series is geometrically ergodic or mixing?

- inequalities for covariances, for moments of sums
- laws of large numbers
- central limit theorem
- consistency and asymptotic normality of statistical estimators
Geometric ergodicity of Markov chains

$X_0, X_1, \ldots$ homogeneous Markov chain with state space $(E, \mathcal{F})$, 
$\mathcal{F}$ $\sigma$-algebra of Borel sets of $E$

$P(x, A) = \mathbb{P}\{X_{t+1} \in A \mid X_t = x\}$ transition probabilities for 
$x \in E, A \in \mathcal{F}$

$P^n(x, A) = \mathbb{P}\{X_{t+n} \in A \mid X_t = x\}$ $n$-step transition probabilities

$\pi$ stationary distribution of $\{X_t\}$

$\pi(A) = \int_E P(x, A) \pi(dx)$
Geometric ergodicity of Markov chains

\[ \{X_t, t = 0, 1, \ldots\} = \{X_0, X_1, \ldots\} \] homogeneous Markov chain with state space \((E, \mathcal{F})\)

\(P(x, A)\) transition probabilities for \(x \in E, A \in \mathcal{F}\)

\(P^n(x, A)\) \(n\)-step transition probabilities

\[ \|\sigma\| := \sup_{A \in \mathcal{F}} \sigma(A) - \inf_{A \in \mathcal{F}} \sigma(A) \] total variation norm

**Definition:** Assume that \(\pi\) is a stationary distribution, \(Q : E \rightarrow [0, +\infty]\) is a measurable function with \(\int_E Q(x) \pi(dx) < +\infty\). The Markov chain \(\{X_n\}\) is called \(Q\)-geometrically ergodic if

\[ \|P^n(x, \cdot) - \pi\| \leq (a + bQ(x)) \gamma^n \] for all \(x \in E\)

with constants \(a, b > 0, \gamma \in (0, 1)\).
Geometric ergodicity of Markov chains

\[ \nu \ldots \text{initial distribution: } \nu(A) = \mathbb{P}\{X_0 \in A\} \]

\{X_n\} \text{ } Q\text{-geometrically ergodic } \implies

\[ |P_\nu\{X_n \in A\} - \pi(A)| = \left| \int_E (P^n(x, A) - \pi(A)) \nu(dx) \right| \]
\[ \leq \frac{1}{2} \int_E \|P^n(x, .) - \pi\| \nu(dx) \]
\[ \leq \frac{1}{2} \left( a + b \int_E Q(x) \nu(dx) \right) \gamma^n, \quad \gamma \in (0, 1) \]

\implies \{X_n\} \text{ } Harris \text{ } ergodic \implies

\[ \frac{1}{n} \sum_{k=1}^{n} g(X_k) \longrightarrow \int_E g(x) \pi(dx) \quad \text{a.s.} \]

for any function \( g : \mathbb{R} \to \mathbb{R}, \int_E g(x) \pi(dx) < +\infty. \)
Auxiliary theorem

**Theorem:** Chan (1989), Mokkadem (1990)

Let $\mu$ be a $\sigma$-finite measure with $\mu(E) > 0$. Assume that

(i) for any compact $C \subset E$, and for any set $N \in \mathcal{F}$ with $\mu(N) = 0$, there is an integer $m$ such that $P^m(x, N) = 0$ for all $x \in C$.

(ii) Further, for any compact set $C \subset E$ and any set $A \in \mathcal{F}$ with $\mu(A) > 0$, there is an integer $\bar{m}$ with $\inf_{x \in C} P^{\bar{m}}(x, A) > 0$. 
**Auxiliary theorem**

**Theorem:** Let $\mu$ be a $\sigma$-finite measure, $\mu(E) > 0$. Assume that

(i) for any compact $C \subset E$, and for any set $N \in \mathcal{F}$ with $\mu(N) = 0$, there is an integer $m$ such that $P^m(x, N) = 0$ for all $x \in C$.

(ii) Further, for any compact set $C \subset E$ and any set $A \in \mathcal{F}$ with $\mu(A) > 0$, there is an integer $\bar{m}$ with $\inf_{x \in C} P^{\bar{m}}(x, A) > 0$.

(iii) There exist a measurable function $Q : E \to [0, +\infty)$, a compact set $K$ with $\mu(K) > 0$, constants $A, \varepsilon, M, \bar{M} > 0$ and $\bar{\gamma} \in (0, 1)$ such that $M \leq Q(x) \leq \bar{M}$ for all $x \in K$,

\[
\mathbb{E}(Q(X_{n+1}) \mid X_n = x) \leq \bar{\gamma}Q(x) - \varepsilon \quad \text{for all } x \in E \setminus K,
\]

\[
\mathbb{E}(Q(X_{n+1}) \mid X_n = x) \leq A \quad \text{for all } x \in K,
\]

Then $\{X_n\}$ is $Q$-geometrically ergodic with stationary distribution $\pi$ and $\int_E Q(x) \pi(dx) < +\infty$. Moreover, $\pi$ is equivalent to $\mu$.

▷ $Q$ . . . test/Lyapunov function, later $\mu$ Lebesgue m., $Q(x) = \|x\|_7$
Mixing coefficients

$X_1, X_2, \ldots$ sequence of random variables, $\mathcal{F}^b_a := \sigma(X_t, a \leq t \leq b)$, $\mathcal{L}^2(\mathcal{F}^b_a)$ set of $\mathcal{F}^b_a$-measurable random variables with finite and positive variance.

\[ \alpha_m := \sup_{k=1,2,\ldots} \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{m+k}^\infty, B \in \mathcal{F}_1^k \right\} \]

\[ \beta_m := \sup_{k=1,2,\ldots} \mathbb{E} \left( \sup \left\{ |\mathbb{P}(A|\mathcal{F}_1^k) - \mathbb{P}(A)| : A \in \mathcal{F}_{m+k}^\infty \right\} \right) \]

\[ \varphi_m := \sup_{k=1,2,\ldots} \sup \left\{ |\mathbb{P}(A|B) - \mathbb{P}(A)| : A \in \mathcal{F}_{m+k}^\infty, B \in \mathcal{F}_1^k, \mathbb{P}(B) > 0 \right\} \]

\[ \rho_m := \sup_{k=1,2,\ldots} \sup \left\{ \text{cov}(X, Y) : X \in \mathcal{L}^2(\mathcal{F}_{m+k}^\infty), Y \in \mathcal{L}^2(\mathcal{F}_1^k) \right\} \]

\[ 2\alpha_m \leq \beta_m \leq \varphi_m, \quad 4\alpha_m \leq \rho_m \leq 2\sqrt{\varphi_m} \quad \text{for } m = 1, 2, \ldots \]

monograph by Doukhan 1994
Mixing properties

\(\alpha\)-mixing: \(\alpha_m \to 0\) as \(m \to \infty\)

\(\beta\)-mixing, absolutely regular: \(\beta_m \to 0\) as \(m \to \infty\)

\(\varphi\)-mixing: \(\varphi_m \to 0\) as \(m \to \infty\)

\(\rho\)-mixing: \(\rho_m \to 0\) as \(m \to \infty\)

\(\psi\)-mixing: \(\psi_m \to 0\) as \(m \to \infty\)

generalisation weak mixing: Doukhan, Dedecker, Louhichi 1999..
Geometric ergodicity and mixing - stationary case

**Theorem:** Assume that the Markov chain \( \{X_n\} \) is stationary.

\( \{X_n\} \) is geometrically ergodic \( \iff \) \( \{X_n\} \) is absolutely regular with

\[ \beta_n = O(\gamma^n) \quad \text{for some } \gamma \in (0, 1). \]
Geometric ergodicity and mixing - nonstationary case

ν...initial distribution: \( \nu(A) = \mathbb{P}\{X_0 \in A\} \)

**Lemma:**
\[
\beta_n \leq 3 \int_E \| P^m(x, .) - \pi \| \nu(dx) + \int_E \| P^m(x, .) - \pi \| \pi(dx) \text{ with } m := \lfloor n/2 \rfloor .
\]

**Theorem:** Suppose that \( \{X_t\} \) is \( Q \)-geometrically ergodic and \( \int_E Q(x) \nu(dx) < +\infty \). Then \( \{X_t\} \) is absolutely regular with
\[
\beta_n = O(\gamma^n) \text{ for some } \gamma \in (0, 1).
\]
Linear autoregressive sequences

\{X_t, t = 1, 2, \ldots\} sequence of random variables satisfying

\[ X_{t+1} = aX_t + b + \varepsilon_{t+1}, \]

\{\varepsilon_t\} sequence of i.i.d. random variables with \( E|\varepsilon_t| < +\infty, a, b \in \mathbb{R}, \)
\( \varepsilon_t \) has a density which is a.e. positive on \( \mathbb{R}, \)

**Theorem:** Assume that \( |a| < 1, \) either

(i) \( \{X_t\} \) stationary or

(ii) \( E|X_1| < +\infty \)

\[ \implies \{X_t\} \text{ is geometrically ergodic and absolutely regular:} \]

\[ \beta_n = O(\gamma^n) \text{ for some } \gamma \in (0, 1). \]
Nonlinear autoregressive sequences

\[ \{X_t, t = 1, 2, \ldots\} \text{ sequence of random variables satisfying} \]
\[ X_{t+1} = g(X_t) + \varepsilon_{t+1}, \]
\[ \{\varepsilon_t\} \text{ sequence of i.i.d. random variables with } E|\varepsilon_t| < +\infty, \]
\[ \varepsilon_t \text{ has a density which is a.e. positive on } \mathbb{R}, \]
\[ g: \mathbb{R} \to \mathbb{R} \text{ measurable function, bounded on compact sets}. \]

**Theorem:** Assume that

for some \( R > 0 \),
\[ \sup_{x:|x| \geq R} \left| \frac{g(x)}{x} \right| < 1, \]

either

(i) \( \{X_t\} \) stationary or (ii) \( E|X_1| < +\infty \)

\[ \implies \{X_t\} \text{ is geometrically ergodic and absolutely regular:} \]
\[ \beta_n = O(\gamma^n) \text{ for some } \gamma \in (0, 1). \]

\[ \text{\textcopyright Tjøstheim (1990)} \]
Linear vector-AR-model

\[ \{X_t, t = 1, 2, \ldots \} \] stationary time series satisfying

\[ X_{t+1} = AX_t + \eta_{t+1} \quad \text{for} \ t = 1, 2, \ldots, \]

where \( X_t \in \mathbb{R}^p, A \in \mathbb{R}^{p \times p}, \eta_2, \eta_3, \ldots \in \mathbb{R}^p \) independent of \( X_1, \mathbb{E} \eta_t = 0, \)
\( \text{cov}(\eta_t) = \Sigma, \eta_t \) has a density \( f_\eta \) with \( f_\eta(x) > 0 \) for \( x \in \mathbb{R}^p, \)

- spectral radius of a matrix \( B: \)

\[ \rho(B) = \max\{|\lambda| : \lambda \text{ is eigenvalue of } B\} \]

**Theorem:** Assume that \( \rho(A) < 1. \) Then the time series \( \{X_t\} \) is absolutely regular with

\[ \beta_n = O(\gamma^n) \quad \text{for some } \gamma \in (0, 1). \]

\[ \text{Tjøstheim (1990)} \]

**Task:** Generalisation to nonlinear models
Nonlinear Vector-ARCH(1)-models

\(X_1, X_2, \ldots\) time series satisfying

\[
X_{t+1} = G(X_t) + \Sigma(X_t) \cdot \eta_{t+1} \quad (t = 0, 1, 2, \ldots),
\]

where \(X_t \in \mathbb{R}^p, \eta_2, \eta_3, \ldots \in \mathbb{R}^p\) independent of \(X_1\), \(E\eta_t = 0\), \(\text{cov}(\eta_t) = I\), \(\eta_t\) has a density \(f_\eta\) with \(f_\eta(x) > 0\) for \(x \in \mathbb{R}^p\), \(G : \mathbb{R}^p \to \mathbb{R}^p\), \(\Sigma : \mathbb{R}^p \to \mathbb{R}^{p \times p}\) measurable, \(\Sigma\) regular for all arguments.

\(\triangleright\) \(\{X_t\}\) Markov process

\[
E(X_{t+1} \mid X_t = x) = G(x), \quad \text{cov}(X_{t+1} \mid X_t = x) = \Sigma(x)\Sigma(x)^T
\]

Assumptions: \(||\cdot||\) Euclidean vector norm

\[
G(x) = A(x) \cdot x + o(||x||), \quad \Sigma(x) = o(||x||)
\]

as \(||x|| \to \infty\), and for each compact set \(C \subset \mathbb{R}^p\), there are constants \(\kappa_1, \kappa_2 > 0\) such that

\[
||\Sigma(x)^{-1}|| \leq \kappa_1, \quad |\det(\Sigma(x))| \leq \kappa_2 \quad \text{for all} \ x \in C.
\]
Joint spectral radius

- spectral radius of a matrix $B$:
  \[ \rho(B) = \max\{|\lambda| : \lambda \text{ is eigenvalue of } B\} \]

- Given a bounded set $\mathcal{A} \subset \mathbb{R}^{n \times n}$.

**Joint spectral radius of $\mathcal{A}$**

\[
\rho(\mathcal{A}) = \limsup_{m \to \infty} \left( \sup_{A \in \mathcal{A}^m} \|A\| \right)^{1/m} = \sup_{m \geq 1} \left( \sup_{A \in \mathcal{A}^m} \rho(A) \right)^{1/m}
\]

$\mathcal{A}^m = \{A_1A_2 \ldots A_m : A_i \in \mathcal{A}, \ i = 1, \ldots, m\}$, $\|\cdot\|$ any matrix norm

$\rho(\mathcal{A})$ does not depend on the choice of this norm

*Gripenberg 1996, Blondel/Nesterov 2004*
Joint spectral radius - Example

\[ A = \{ A_1, A_2 \} \]

\[ A_1 = \begin{pmatrix} -1 & -0.3 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -0.7 & -0.4 \\ 1 & 0 \end{pmatrix}. \]

\[ \rho(A_1) = 0.547722557505, \quad \rho(A_2) = 0.632455532034 \]

\[ \rho(A_1 A_2^2 A_1^2 A_2^3) = 0.681266 \]

Gripenberg’s MATLAB program \( \rightarrow \) \( \rho(A) = 0.70712. \)
Vector-ARCH(1)-models

\[ X_{t+1} = G(X_t) + \Sigma(X_t) \cdot \eta_{t+1} \quad (t = 0, 1, 2, \ldots) \]

\[ G(x) = A(x) \cdot x + o(\|x\|) \]

\[ A_R = \{ A(x) : x \in \mathbb{R}^p, \|x\| \geq R \} \]

**Theorem:** Assume that \( G \) and \( \Sigma \) are bounded on compact sets, there is a \( R > 0 \) such that \( A_R \) is a bounded set and

\[ \rho(A_R) < 1. \]

(i) Then the Markov chain \( \{X_t\} \) is \( Q \)-geometrically ergodic with \( Q(x) = \|x\| \).

(ii) If \( \{X_t\} \) is stationary, then the time series \( \{X_t\} \) is absolutely regular with \( \beta_k = O(\gamma^k), \ \gamma \in (0, 1) \).

(iii) If the distribution of \( X_1 \) is not a stationary one and \( \mathbb{E}\|X_1\| < +\infty \), then \( \{X_t\} \) is absolutely regular with \( \beta_k = O(\gamma^k), \ \gamma \in (0, 1) \).
A sufficient condition

Given a bounded set \( \mathcal{A} \subset \mathbb{R}^{n \times n} \).

**Theorem:**

\[
\inf_{S \in \mathbb{R}^{n \times n}} \sup_{A \in \mathcal{A}^m} \| S^{-1} AS \| < 1 \text{ for some matrix norm } \| . \| \text{ and some } m \geq 1.
\]

\[
\implies \rho(\mathcal{A}) < 1.
\]

**Remark:** \( \| S^{-1} AS \| \) is a matrix norm of \( A \).

**Algorithm:**

Choose \( m \), matrix norm \( \| . \| \),

Minimise \( \sup_{A \in \mathcal{A}^m} \| S^{-1} AS \| \) w.r.t. \( S \in \mathbb{R}^{n \times n} \).
Example

\( \mathcal{A} = \{ A_1, A_2 \} \)

\[
A_1 = \begin{pmatrix} -1 & -0.3 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -0.7 & -0.4 \\ 1 & 0 \end{pmatrix}.
\]

\( \rho(A_1) = 0.547722557505, \quad \rho(A_2) = 0.632455532034, \quad \rho(\mathcal{A}) = 0.70712 \)

\[
\min_{S} \sup_{A \in \mathcal{A}} \| S^{-1} AS \| = 0.756151687595
\]

\[
\min_{S} \sup_{A \in \mathcal{A}^2} \| S^{-1} AS \| = 0.713628235883
\]
Example of a transient time series

$$X_{t+1} = A_1 X_t I(X_t \in R_1) + A_2 X_t I(X_t \in R_2) + \varepsilon_t$$

where $$X_t \in \mathbb{R}^2$$, $$R_1 = \mathbb{R} \times [0, \infty)$$, $$R_2 = \mathbb{R} \times (-\infty, 0)$$,

$$A_1 = \begin{pmatrix} 2.1 & 0.6 \\ -2.4 & -0.3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.6 & -1.1 \\ 0.9 & -0.1 \end{pmatrix},$$

$$\bar{A} : = A_2 A_1 = \begin{pmatrix} 3.9 & 0.69 \\ 2.13 & 0.57 \end{pmatrix}.$$

$$\rho(A_1) = 0.9, \quad \rho(A_2) = 0.9643650761, \quad \text{but} \quad \rho(\bar{A}) = 4.294593406.$$

$$\mathbb{E}e \| \varepsilon_t \|^2 \leq C_0 < +\infty.$$

$$\implies \{X_t\} \text{ is transient.}$$
Scalar-ARCH-models

$\{X_t, t = 1, 2, \ldots\}$ time series satisfying

$$X_{t+1} = g(X_t, \ldots, X_{t-p+1}) + \sigma(X_t, \ldots, X_{t-p+1}) \cdot \varepsilon_{t+1} \quad \text{for } t = p, p + 1, \ldots$$

$\{\varepsilon_t\}_{t=p+1,p+2,\ldots}$ is a sequence of i.i.d. random variables independent of $X_1, \ldots, X_p$ with $\mathbb{E}\varepsilon_t \equiv 0$, $\text{var}(\varepsilon_t) \equiv 1$.

$$g(u) = \sum_{i=1}^{p} b_i(u) \cdot u_i + o(\|u\|), \quad \sigma(u) = o(\|u\|) \quad \text{as } \|u\| \to \infty$$

where $b_1, \ldots, b_p : \mathbb{R}^p \to \mathbb{R}$ are bounded measurable functions.
Scalar ARCH-models

Construction of the Markov chain: $\tilde{X}_t := (X_t, \ldots, X_{t-p+1})^T$.

model $\rightarrow$

$$
\begin{pmatrix}
X_{t+1} \\
X_t \\
\vdots \\
X_{t-p}
\end{pmatrix}
= 
\begin{pmatrix}
g(\tilde{X}_t) \\
g(\tilde{X}_t) \\
\vdots \\
g(\tilde{X}_t)
\end{pmatrix}
+ 
\begin{pmatrix}
\sigma(\tilde{X}_t) & 0 & \cdots & 0 \\
0 & 0 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
0 \\
\vdots \\
0
\end{pmatrix}
$$

$\tilde{X}_{t+1} = G(\tilde{X}_t) + \Sigma(\tilde{X}_t)\eta_{t+1}$

with $G(u) = \left(g(u), u_1, \ldots, u_{p-1}\right)^T$, $u = (u_1, \ldots, u_{p-1})^T$,

$G(u) = A(u)u + o(\|u\|)$ \ as $\|u\| \rightarrow \infty$,

$$
A(u) = 
\begin{pmatrix}
b_1(u) & b_2(u) & \cdots & b_{p-1}(u) & b_p(u) \\
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & 0 & 0 \\
0 & \cdots & 0 & 1 & 0
\end{pmatrix}
$$
Scalar ARCH-models

\[ X_{t+1} = g(X_t, \ldots, X_{t-p+1}) + \sigma(X_t, \ldots, X_{t-p+1}) \cdot \varepsilon_{t+1} \quad \text{for } t = p, p + 1, \ldots \]

\[ A_R = \{ A(u) : \|u\| \geq R, u \in \mathbb{R}^p \} \text{ for } R > 0. \]

\[ A(u) = \begin{pmatrix}
  b_1(u) & b_2(u) & \ldots & b_{p-1}(u) & b_p(u) \\
  1 & 0 & \ldots & 0 & 0 \\
  0 & 1 & \ldots & \vdots & \vdots \\
  \vdots & \vdots & \ddots & 0 & 0 \\
  0 & \ldots & 0 & 1 & 0
\end{pmatrix}, \]

\[ \Sigma(u) = \text{diag}(\sigma(u), 0, \ldots, 0), \]

**Theorem:** Assume that \( G \) and \( \Sigma \) are bounded on compact sets, there is a \( R > 0 \) such that \( A_R \) is a bounded set and

\[ \rho(A_R) < 1. \]

\( \implies \) the Markov chain \( \{\tilde{X}_t\} \) is \( Q \)-geometrically ergodic with \( Q(x) = \|x\| \).
Threshold models - TARCH\((l, p, d)\)-model

Let \(r_1, r_2, \ldots, r_{l-1}\) be thresholds such that \(r_1 < r_2 < \ldots < r_{l-1}\), \(r_0 = -\infty\), \(R_j = (r_{j-1}, r_j]\) for \(j = 1, \ldots, l - 1\), and \(R_l = (r_{l-1}, +\infty)\).

\[
X_t = \sum_{j=1}^{l} \left( a_0^{(j)} + \sum_{i=1}^{p} a_i^{(j)} \cdot X_{t-i} \right) \cdot 1_{R_j} (X_{t-d}) + w(X_{t-d}) \varepsilon_t,
\]

for \(t = p, p + 1, \ldots\), \(\{\varepsilon_t\}\) i.i.d. random variables, \(w : \mathbb{R} \to [w, +\infty)\) measurable, \(w > 0\).  \(w(x) = o(\|x\|)\) as \(\|x\| \to \infty\).
Threshold models - \textsc{TARCH}(l, p, d)-model

\[
X_t = \sum_{j=1}^{l} \left( a_0^{(j)} + \sum_{i=1}^{p} a_i^{(j)} \cdot X_{t-i} \right) \cdot 1_{R_j} (X_{t-d}) + w(X_{t-d}) \varepsilon_t,
\]

\[
\bar{A}_j = \begin{pmatrix}
 a_1^{(j)} & a_2^{(j)} & \ldots & a_{p-1}^{(j)} & a_p^{(j)} \\
 1 & 0 & \ldots & 0 & 0 \\
 0 & 1 & \ldots & : & : \\
 : & : & \ldots & 0 & 0 \\
 0 & \ldots & 0 & 1 & 0
\end{pmatrix}, \quad \mathcal{A} = \{\bar{A}_1, \ldots, \bar{A}_l\}.
\]

\textbf{Theorem:} Assumption on the density of \( \varepsilon_t \). Suppose that

- either \( p > 1 \) and \( \rho(\mathcal{A}) < 1 \)
- or \( p = 1, \ |a_1^{(1)}| < 1 \) and \( |a_1^{(l)}| < 1 \).

\[ \Rightarrow \{\tilde{X}_t\} \text{ is } Q\text{-geometrically ergodic with } Q(u) = \|u\|. \]
Further development - literature

- applications to other types of models: EXPARCH-models


- TAR models: Boucher/Cline 2007
  smooth TAR models: Dueker/Psaradakis/Sola/Spagnola 2007
  TARMA models: Ling/Tong/Li 2007

http://www.iks.hs-merseburg.de/%7Eeliebsch/folrum.pdf
AR models with strong heteroscedasticity

\[
X_{t+1} = g(X_t, \ldots, X_{t-p+1}) + \sigma(X_t, \ldots, X_{t-p+1}) \cdot \varepsilon_{t+1} \quad \text{for } t = p, p + 1, \ldots
\]

\[
\{\varepsilon_t\}_{t=p+1,p+2,\ldots} \text{ is a sequence of i.i.d. random variables independent of } X_1, \ldots, X_p \text{ with } \mathbb{E}\varepsilon_t \equiv 0, \text{ var}(\varepsilon_t) \equiv 1
\]

\[
g(u) = \sum_{i=1}^{p} b_i(u) \cdot u_i + o(\|u\|),
\]

\[
\sigma^2(u_1, \ldots, u_p) = d_1u_1^2 + \ldots + d_pu_p^2 + \bar{\sigma}^2(u),
\]

where \( \bar{\sigma}(u) = o(\|u\|) \) as \( \|u\| \to \infty \), \( b_1, \ldots, b_p : \mathbb{R}^p \to \mathbb{R} \) bounded measurable functions.
**AR models with strong heteroscedasticity**

\[ X_{t+1} = g(X_t, \ldots, X_{t-p+1}) + \sigma(X_t, \ldots, X_{t-p+1}) \cdot \varepsilon_{t+1} \quad \text{for } t = p, p + 1, \ldots \]

**Theorem:** Assumption on the density of \( \varepsilon_t \). Suppose that there is some matrix \( S \in \mathbb{C}^{p \times p} \) such that

\[ \rho(S^*BS) < 1 \]

where \( B := A^T (S^*)^{-1} S^{-1} A + s_1 D, s_1 := \|S^{-1}e_1\|^2, e_1 = (1, 0, \ldots, 0)^T, D = \text{diag}(d_1, \ldots, d_p) \). \( S^* \) denotes the Hermitian transpose of \( S \).

\[ \implies \text{Then } \{\tilde{X}_t\} \text{ is } Q\text{-geometrically ergodic with } Q(u) = \|u\|^2. \]

Masry/Tjøstheim 1995, Lu 1998