Stability Properties of some Markov Chain Models in Random Environments

Lionel Truquet [1]

[1]CREST-ENSAI

Outline

1. Markov chains, strict exogeneity and random environments
   - Motivation and general setup
   - Existence of stationary measures via a coupling method
   - Ergodic properties

2. Observation-driven models and random environments
   - Model formulation
   - Existence of stationary solutions under semi-contractivity conditions
   - Some examples

Outline

1. Markov chains, strict exogeneity and random environments
   - Motivation and general setup
   - Existence of stationary measures via a coupling method
   - Ergodic properties
2. Observation-driven models and random environments
   - Model formulation
   - Existence of stationary solutions under semi-contractivity conditions
   - Some examples

Motivation: time series models with strictly exogenous covariates

- Our aim is to find explicit conditions that guarantee existence of stationary processes $Y := (Y_t)_{t\in\mathbb{Z}}$ defined conditionally on another stationary stochastic process $X := (X_t)_{t\in\mathbb{Z}}$.

- We will discuss the case of (conditional) Markov chains models satisfying

$$P\left(Y_t \in A | X, Y_{t-1}, Y_{t-2}, \ldots\right) = P_{X_{t-1}}\left(Y_{t-1}, A\right), \quad t \in \mathbb{Z}. \qquad (1)$$

  where $(x, y, A) \mapsto P_x(y, A)$ is a probability kernel from $F \times E$ to E, E, F Polish spaces.

- Conditional independence property: $\left(X_{t+j}\right)_{j\geq 0}$ **is independent of** Y_t **conditionally on** $\left(Y_{t-j}, X_{t-j}\right)_{j\geq 1}$.

- In econometrics, the latter independence condition is often called **strict exogeneity** ([Sims (1972), Chamberlain (1982)]).

- In probability theory, (1) refers to **Markov chain in random environments**. See [Cogburn (1984), Orey (1991), Kifer (1995), Stenflo (2001)]. Discrete state spaces or very strong assumptions are mainly used for existence of stationary laws.

How to construct stationary laws for MCRE ?

- $P\left(Y_t \in A | X, Y_{t-1}, Y_{t-2}, \ldots\right) = P_{X_{t-1}}\left(Y_{t-1}, A\right), \quad t \in \mathbb{Z}.$

- If a stationary solution exists, the (conditional) marginal distribution $Y_t|X$, denoted by π_t, satisfies the invariance equations $\pi_t P_{X_t} = \pi_{t+1}$ a.s.

- Since $\pi_t = \pi_{t-1} P_{X_{t-1}} = \pi_{t-2} P_{X_{t-1}} P_{X_t} = \cdots = \pi_{t-n} P_{X_{t-n}} \cdots P_{X_{t-1}}$, natural candidates for π_t are given by the almost sure limits of the backward iterations of the chain

$$\lim_{n\rightarrow\infty} \mu P_{X_{t-n}} \cdots P_{X_{t-1}}.$$

- Studying the **almost sure limits of the backward iterations** of such time-inhomogeneous Markov chains ($t = 0$ is sufficient) is one possibility to construct stationary laws (with a topology to find...).

Constructive results for MCRE: Kifer (1995)

Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process.

Theorem 1

Suppose that there exist a positive integer N, a probability kernel $(x,A)\mapsto \nu_x(A)$ from F^N to E and a measurable mapping $\eta : F^N \to (0,\infty)$ such that a.s.,

$$P_{X_{-N}}\cdots P_{X_{-1}}(y,A) \geq \eta\left(X_{-N},\ldots,X_{-1}\right)\nu_{X_{-N},\ldots,X_{-1}}(A), \quad (y,A)\in E\times\mathcal{B}(E).$$

There then exists a random probability measure $\pi_{X_{-1}^-}$ and two random variables $L:\Omega\to(0,\infty)$ and $\kappa:\Omega\in(0,1)$ such that a.s.

$$\sup_{y\in E}\sup_{A\in\mathcal{B}(E)}\left|\delta_y P_{X_{-n}}\cdots P_{X_{-1}}(A)-\pi_{X_{-1}^-}(A)\right|\leq L\kappa^n.$$

- The integer N can be also a random variable.
- Existence and uniqueness of a stationary process $(Y_t,X_t)_{t\in\mathbb{Z}}$ easily follows from this result. Moreover, $(X_t)_{t\in\mathbb{Z}}$ ergodic implies $(Y_t,X_t)_{t\in\mathbb{Z}}$ ergodic.
- This random Doeblin's type condition (uniform minorization of the transition probabilities) is mainly interesting for bounded state spaces E.

Constructive results for MCRE: Lovas and Rásonyi (2021)

- In order to relax the uniform minorization condition, drift type conditions can be used.

- Assume the existence of $V : E \to (0, \infty)$ s.t. for any $x \in F$, $P_x V \le \lambda(x) V + b(x)$ (**drift condition**) with a long-time contractivity condition:

$$\limsup_n \mathbb{E}^{1/n} \left[b(X_0) \prod_{k=1}^{n} \lambda(X_k) \right] < 1.$$

- Assume furthermore the **minorization condition on a level set** $\{V \le R(x)\}$, i.e. $P_x(y, A) \ge \eta(x) \nu_x(A)$ when $V(y) \le R(x)$, with the smallness condition:

$$\lim_{n\to\infty} \mathbb{E}^{1/n^\theta} \left[(1 - \eta(X_0))^n \right] = 0 \text{ for some } \theta \in (0, 1).$$

- If $R(x)$ is "large enough", one can derive existence and uniqueness of a compatible stationary process as well as some weak laws of large numbers for the MCRE $(Y_n^y)_{n\ge0}$ arbitrarily initialized with $Y_0^y = y$ and a convergence rate for $\mathbb{P}_{Y_n^y}$ towards a universal measure not depending on y.

Outline

1. Markov chains, strict exogeneity and random environments
   - Motivation and general setup
   - Existence of stationary measures via a coupling method
   - Ergodic properties
2. Observation-driven models and random environments
   - Model formulation
   - Existence of stationary solutions under semi-contractivity conditions
   - Some examples

Assumptions used in the rest of the talk

A1 The environment $(X_t)_{t\in\mathbb{Z}}$ is stationary and **ergodic**.

A2 There exist three measurable mappings $V : E \to (0,\infty)$ and $\lambda, b : F \to (0,\infty)$ s.t. for all $x \in F$, $P_x V \leq \lambda(x) V + b(x)$. Moreover

$$\mathbb{E}\log^+(b(X_0)), \quad \mathbb{E}\log^+(\lambda(X_0)), \quad \mathbb{E}\log(\lambda(X_0)) < 0.$$

A3 There exist a measurable mapping $\eta : (0,\infty) \times F \to (0,1)$ such that for any $R > 0$, one can find a probability kernel ν_R from F to E such that

$$P_x(y, A) \geq \eta(R, x)\nu_R(x, A), \quad (x, y, A) \in F \times V^{-1}([0, R]) \times \mathcal{B}(E).$$

Main result

Theorem 2

Assume **A1-A3**. *The following assertions hold true.*

1. *The sequence* $\left(\delta_z P_{X_{-n}} \cdots P_{X_{-1}}\right)_{n\geq 0}$ *is converging* $\mathbb{P}-$*almost surely in total variation towards a random probability measure* $\pi_{X_{-1}^{-}}$ *not depending on* z. *Moreover, for some random variables* $L:\Omega\to(0,\infty)$ *and* $\kappa:\Omega\to(0,1)$ *s.t.* $\mathbb{P}-$*a.s.,*
$$d_{TV}\left(\delta_z P_{X_{-n}} \cdots P_{X_{-1}}, \pi_{X_{-1}^{-}}\right) \leq L\left(1+V(z)\right)\kappa^n. \quad (2)$$
2. *For any* $t\in\mathbb{Z}$, *if* $\pi_t=\pi_{X_t^{-}}$, *we have* $\pi_{t-1}P_{X_t}=\pi_t$ *a.s.*
3. *If* $(\nu_t)_{t\in\mathbb{Z}}$ *is a sequence of identically distributed random probability measures such that* $\nu_{t-1}P_{X_t}=\nu_t$ *a.s., then* $\nu_0=\pi_{X_0^{-}}$ *a.s.*
4. *We have for any* $t\in\mathbb{Z}$, $\pi_t V<\infty$ *a.s.*

Existence and uniqueness of a stationary law easily follows from this result. It can be also extended when the drift/small set conditions are obtained after iteration.

Key lemma for the proof of Theorem 2

Proposition 1

Assume **A1-A3**. *There exist two random variables* $L : \Omega \to (0, \infty)$ *and* $\kappa : \Omega \to (0, 1)$ *s.t.* $\mathbb{P}-$*a.s.,*

$$d_{TV}\left(\delta_{y'} P_{X_{-n}} \cdots P_{X_{-1}}, \delta_y P_{X_{-n}} \cdots P_{X_{-1}}\right) \leq L\left(1 + V(y) + V(y')\right) \kappa^n.$$

- To get an upper bound of the total variation distance is to construct a coupling $(Y_t, Y_t')_{t \geq -n}$ of two chains with $Y_{-n} = y$, $Y'_{-n} = y'$ and conditionally on X, both processes are time-inhomogeneous Markov chains with transition kernels $P_{X_{-n}}, P_{X_{-n+1}}, \ldots$.
- There exist some related bounds in the literature but with quite stringent conditions on the drift parameters (e.g. [Douc, Moulines and Rosenthal (2004)]).

Prerequisites: a standard coupling scheme for homogeneous Markov chains

- When $PV \leq \lambda V + b$ and $P(y, A) \geq \eta\nu(A)$ if $V(y) \leq R$, [Rosenthal (1995)] uses a specific coupling scheme for approximating the invariant probability measure by the marginal law of the chain in the context of geometric ergodicity.

- Set $Y_0 = y, Y_0' = y'$.
  - On the event $\{Y_{t-1} = Y_{t-1}'\}$, set
$$\mathbb{P}\left(Y_t \in A, Y_t' \in A' | Y_{t-1}, Y_{t-1}'\right) = P\left(Y_{t-1}, A \cap A'\right).$$
  - On the event $\{Y_{t-1} \neq Y_{t-1}', V(Y_{t-1}) \vee V(Y_{t-1}') > R\}$, we set
$$\mathbb{P}\left(Y_t \in A, Y_t' \in A' | Y_{t-1}, Y_{t-1}'\right) = P\left(Y_{t-1}, A\right) P\left(Y_{t-1}', A'\right).$$
  - On the event $\{Y_{t-1} \neq Y_{t-1}', V(Y_{t-1}) \vee V(Y_{t-1}') \leq R\}$, set
$$\begin{aligned}\mathbb{P}\left(Y_t \in A, Y_t' \in A' | Y_{t-1}, Y_{t-1}'\right) &= \eta \cdot \nu\left(A \cap A'\right) \\ &+ (1-\eta) Q\left(Y_{t-1}, A\right) Q\left(Y_{t-1}', A'\right),\end{aligned}$$

with $Q(y, A) = \frac{P(y,A) - \eta \cdot \nu(A)}{1-\eta}$.

Prerequisites: a standard coupling scheme for homogeneous Markov chains

- Let T_i, $i \geq 1$, the successive random times such that $V(Y_{T_i}) + V\left(Y'_{T_i}\right) \leq R$ a.s.

- For an arbitrary integer $m < n$,

$$\mathbb{P}(Y_n \neq Y'_n) \leq \mathbb{P}(T_m \geq n) + \mathbb{P}(T_m < n, Y_n \neq Y'_n).$$

  On the event $\{T_m < n\}$, we have a probability smaller than $(1-\eta)^m$ to not get a coalescence of the two paths, we deduce that

$$\mathbb{P}(Y_0 \neq Y'_0) \leq \mathbb{P}(T_m \geq n) + (1-\eta)^m.$$

- It then remains to bound the probability $\mathbb{P}(T_m \geq n)$

Prerequisites: a standard coupling scheme for homogeneous Markov chains

Lemma 1

Set $\rho_j = T_j - T_{j-1}$, $\zeta = \frac{2}{1+\lambda}$, $R = \frac{2b+2}{1-\lambda}$ *and* $D = 1 + \frac{b+\lambda R}{1-\lambda}$. *We have the two following bounds.*

1. *If* $V(y) + V(y') > R$, *we have* $\mathbb{E}\left(\zeta^{\rho_1}\right) \leq V(y) + V(y')$.
2. *For any* $j \geq 2$,
$$\mathbb{E}\left(\zeta^{\rho_j} | \mathcal{F}_{T_{j-1}}\right) \leq D\zeta.$$

We then get $\mathbb{P}\left(T_m \geq n\right) \leq D^m \left(V(y) + V(y')\right) \zeta^{-n+m}$. Optimizing w.r.t. m, for some explicit constants:

$$d_{TV}\left(\delta_y P^n, \delta_{y'} P^n\right) \leq L(1 + V(y) + V(y'))\kappa^n.$$

Extension to random environments

- We study the effect of the same coupling for MCRE conditionally on the environment (at time $t = 0$ and starting with $Y_{-n} = y, Y'_{-n} = y'$).

- A first issue is to choose the radius R of the level set for V.

- We will try to avoid the "storms". For many time indices, the drift parameters can be large and it is complicated to control the return time of the chain in a level set $\{V \leq R\}$. The probability $\eta(R, x)$ to stick the path can be also arbitrarily small.

- The main idea is to to define some random times τ_i only depending on the environment and such that $\left(Y_{\tau_i}, Y'_{\tau_i}\right)_i$ have some drift parameters under control. At the same time, the probability to stick the paths at time $\tau_i + 1$ should be kept under control.

- Notations $\mathbb{P}_\omega$ and $\mathbb{E}_\omega$ are used to stress the "conditionally on the environment" expectations.

A key intermediate result

Lemma 2

There exist two positive real numbers C_1, C_2 and an increasing sequence of random times $(\tau_i)_{i\in\mathbb{Z}}$, $\tau_i : \Omega \to \mathbb{Z}$ such that the following statements are valid.

1. $\tau_{-1} \leq -1$, $\tau_0 \geq 0$ *and for* $i \in \mathbb{Z}$, $\tau_i - \tau_{i-1} \geq C_1$, $\mathbb{P}-$*a.s.*
2. *If* $\omega \in \Omega$, *We then have*
$$\mathbb{E}_\omega\left[V\left(Y_{\tau_i(\omega)}\right)|Y_{\tau_{i-1}(\omega)}\right] \leq (1 - 1/C_1)\, V\left(Y_{\tau_{i-1}(\omega)}\right) + C_1,$$
$$\mathbb{E}_\omega\left[V\left(Y'_{\tau_i(\omega)}\right)|Y'_{\tau_{i-1}(\omega)}\right] \leq (1 - 1/C_1)\, V\left(Y'_{\tau_{i-1}(\omega)}\right) + C_1.$$
3. *Setting* $R = 2C_1(2C_1+1)$, *we have* $\eta\left(R, X_{\tau_i}\right) \geq 1/C_2$, $\mathbb{P}-$*a.s.*
4. $\lim_{i\to\infty}\tau_i = \infty$ *and* $\lim_{i\to-\infty}\tau_i = -\infty$ *a.s. Moreover if* $L_n = \sup\left\{i \geq 1 : \tau_{-i} \geq -n\right\}$, *then*
$$\lim_{n\to\infty}\frac{L_n}{n} > 0 \quad \mathbb{P}\text{-a.s.}$$

Ideas for the proof of Lemma 2

- We have

$$\mathbb{E}_\omega\left(V(Y_t)|Y_{t-j}\right) \leq \prod_{s=1}^{j}\lambda\left(X_{t-s}(\omega)\right)V\left(Y_{t-j}\right)+b\left(X_{t-1}(\omega)\right) + \sum_{k\geq 2}\prod_{s=1}^{k-1}\lambda\left(X_{t-s}(\omega)\right)b\left(X_{t-k}(\omega)\right).$$

- Choose then $C_1 > 0$ in order to get $\mathbb{P}\left(X \in A_{1,C_1}\right) > 0$, A_{1,C_1} being the set of $x \in F^{\mathbb{Z}}$ s.t.

$$\sup_{j\geq C_1}\prod_{i=1}^{j}\lambda(x_{-i}) \leq 1 - 1/C_1 \text{ and}$$

$$b(x_{-1}) + \sum_{i\geq 2}\prod_{k=1}^{i-1}\lambda(x_{-k})b(x_{-i}) \leq C_1.$$

- Choose next $C_2 > 0$ such that $\mathbb{P}\left(X \in A_{1,C_1}\cap A_{2,C_2}\right) > 0$ with

$$A_{2,C_2} = \left\{x \in F^{\mathbb{Z}} : \eta\left(2C_1(2C_1+1), x_0\right) \geq 1/C_2\right\}.$$

Ideas for the proof of Lemma 2

- Setting $A_C = A_{1,C_1} \cap A_{2,C_2}$, from the ergodic theorem (set $\theta x = (x_{i+j})_{j\in\mathbb{Z}}$):

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}_{A_C}\left(\theta^t X\right) = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}_{A_C}\left(\theta^{-t} X\right) = \mathbb{P}\left(X \in A_C\right) > 0.$$

- Denote by $0 \leq \widetilde{\tau}_0 < \widetilde{\tau}_1 < \cdots$ and $-1 \geq \widetilde{\tau}_{-1} > \widetilde{\tau}_{-2} > \cdots$ the successive time points t such that $\theta^t X \in A_C$ or $\theta^{-t} X \in A_C$.
- Take $\tau_i = \widetilde{\tau}_{1+(i+1)C_1}$ for $i \geq 0$ and $\tau_{-i} = \widetilde{\tau}_{1-(i-1)C_1}$ for $i \geq 1$.
- The last point of the result is a consequence of the ergodic theorem. Extension to stationary non-ergodic environments is possible, in this case C_1, C_2 and R are random...

End of the proof for Proposition 1

- We denote by $T_{\omega,i}$, $i \geq 1$, the successive return times of the Markov chain $(Z_{\omega,i}, Z'_{\omega,i}) := \left(Y_{\tau_i(\omega)}, Y'_{\tau_i(\omega)}\right)$ in the set $\left\{(y,y') \in E^2 : V(y) + V(y') \leq R := 2C_1(2C_1+1)\right\}$.

- We get the bound

$$\mathbb{P}_\omega\left(Y_0 \neq Y_0'\right) \leq \inf_{1 \leq m \leq L_n(\omega)} \left\{(1 - 1/C_2)^m + D^m\left(1 + V(y) + V(y')\right)\zeta^{m - L_n(\omega)}\right\}$$

- The required bound is obtained if $m \sim L_n(\omega)/k$ with $(D\zeta)^{1/k}/\zeta < 1$. Finally, we remember that $L_n(\omega)$ is of order n.

Outline

1. Markov chains, strict exogeneity and random environments
   - Motivation and general setup
   - Existence of stationary measures via a coupling method
   - Ergodic properties

2. Observation-driven models and random environments
   - Model formulation
   - Existence of stationary solutions under semi-contractivity conditions
   - Some examples

Ergodic properties of the unique stationary solution

- A direct proof of ergodicity of the unique stationary solution can be obtained from our main result.
- To this end, we work on the canonical space $E^{\mathbb{N}} \times F^{\mathbb{Z}}$ and denote by γ the probability distribution of the pair $((Y_t)_{t\in\mathbb{N}}, X)$.
- We denote by γ^ω the distribution of $(Y_t)_{t\in\mathbb{N}}$ conditionally on the environment path. We simply have

$$\gamma\left(A \times B\right) = \int_B \gamma^\omega(A) d\mathbb{P}(\omega).$$

- Denoting by $\tau := \left(\theta_*, \theta\right).(y,\omega) = (\theta_* y, \theta\omega)$, with $\theta_* y = (y_{t+1})_{t\in\mathbb{N}}$. Ergodicity of τ for γ means $\gamma(I) \in \{0,1\}$ if $\tau^{-1} I = I$.

A key lemma entailing ergodicity

Lemma 3

For $\mathbb{P}-$almost $\omega \in F^{\mathbb{Z}}$, there exists a sequence $n_i = n_i(\omega) \to \infty$ s.t. for any $A \in \mathcal{B}\left(E^{\mathbb{N}}\right)$,

$$\lim_{i\to\infty} \sup_{B\in\mathcal{B}(E^{\mathbb{N}})} \left|\gamma^{\omega}\left(A \cap \theta_*^{-n_i(\omega)} B\right) - \gamma^{\omega}(A)\gamma^{\omega}\left(\theta_*^{-n_i(\omega)} B\right)\right| = 0.$$

Proof of the lemma. Take a cylinder set $A = \prod_{i=0}^{k} A_i \times E \times E \cdots$. It can be easily shown that

$$\begin{aligned} & \left|\gamma^{\omega}\left(A \cap \theta_*^{-n} B\right) - \gamma^{\omega}(A)\gamma^{\omega}\left(\theta_*^{-n} B\right)\right| \\ \leq \quad & \int_E \pi_{\omega_{k-1}^-}(dy_k)\left(1 + V(y_k)\right) L\left(\theta^n\omega\right)\kappa\left(\theta^n\omega\right)^{n-k}. \end{aligned}$$

Define $n_1 < n_2 < \cdots$ such that $L \circ \theta^{n_i} \leq c$ and $\kappa \circ \theta^{n_i} \leq 1 - 1/c$ $\mathbb{P}-$a.s. with a constant $c > 0$ such that $\mathbb{P}\left(L < c, \kappa < 1 - 1/c\right) > 0$. For an arbitrarily $A \in \mathcal{B}\left(E^{\mathbb{N}}\right)$, one can approximate A by a finite union of disjoints cylinder sets (for γ^{ω}).

Example of an autoregressive process with threshold

- Consider an $\mathbb{R}^d-$valued stationary and ergodic process $(X_t)_{t\in\mathbb{Z}}$ independent from a real-valued i.i.d. sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$.
- For $a_i, b_i, r : \mathbb{R}^d \to \mathbb{R}$, assume that

$$\begin{aligned} Y_t &= [b_1(X_{t-1}) + a_1(X_{t-1})Y_{t-1}]\,\mathbb{1}_{Y_{t-1}\leq r(X_{t-1})} \\ &+ [b_2(X_{t-1}) + a_2(X_{t-1})Y_{t-1}]\,\mathbb{1}_{Y_{t-1}> r(X_{t-1})} + \varepsilon_t. \end{aligned}$$

- Set $\lambda(x) = \max(|a_1(x)|, |a_2(x)|)$.

Proposition 2

Assume that $\mathbb{E}|\varepsilon_0| < \infty$, the distribution ε_0 has a positive density f lower-bounded on any compact subset of $\mathbb{R}$,

$$\mathbb{E}\log^+ a_i(X_0), \mathbb{E}\log^+ b_i(X_0) < \infty, \quad \mathbb{E}\log\lambda(X_0) < 0.$$

There then exists a unique stationary and ergodic solution $((Y_t, X_t))_{t\in\mathbb{Z}}$ satisfying this dynamic.

Outline

Outline

Observation-driven models

$$Y_t|(\lambda_{t-j}, Y_{t-j-1})_{j\geq 0} \sim p\left(\cdot|\lambda_t\right), \quad \lambda_t = f\left(\lambda_{t-1}, Y_{t-1}\right).$$

- $(s, A) \in F \times \mathcal{B}(E) \mapsto p(A|s)$ is a probability kernel and E, F Borel subsets of $\mathbb{R}^k, \mathbb{R}^\ell$.
- Both processes $(\lambda_t)_{t\geq 0}$ and $(Y_t, \lambda_t)_{t\geq 0}$ form a Markov chain
- Such examples contain many time series models used in Econometrics (e.g. GARCH models $Y_t = \varepsilon_t\sqrt{\lambda_t}$ with $\nu_s = \frac{1}{\sqrt{s}} f_\varepsilon\left(\frac{\cdot}{\sqrt{s}}\right)$).

Irreducibility problem for discrete-valued time series

$$Y_t|\lambda_t \sim \mathcal{P}(\lambda_t), \quad \lambda_t = c + b\lambda_{t-1} + aY_{t-1}.$$

- Recursions imply that $\lambda_t = b^t\lambda_0 + \sum_{j=0}^{t-1} b^j\,(c + aY_{t-j-1})$.
- Since the Y_t's take integer values, $\lambda_t|\lambda_0 = s$ and $\lambda_t|\lambda_0 = s'$ can have disjoint discrete supports (e.g. $s, a, b, c \in \mathbb{Q}$ and $s' \notin \mathbb{Q}$).
- Small set conditions on the Markov chain λ_t are not possible.
- There exist alternative criteria for studying existence and uniqueness of stationary probability measures. See Douc, Doukhan and Moulines [SPA, 2013], Doukhan and Neummann [JoAP, 2019]

Observation-driven models with covariates

- We consider models with strictly exogenous regressors defined by conditional distributions

$$\mathbb{P}\left(Y_t \in A | (X_s, Y_u, \lambda_{u-1}); s \in \mathbb{Z}, u \leq t-1\right) = p\left(A|\lambda_t\right),$$

$$\lambda_t = f\left(\lambda_{t-1}, Y_{t-1}, X_{t-1}\right).$$

- This class of models, called observation-driven, are widely used by the practitioners but probabilistic guarantees (e.g. existence of stationary paths) have been mainly obtained without exogenous regressors.

- Examples of one-parameter probability distributions p are
  - Poisson, $p(k|s) = \exp(-s)s^k/k!$,
  - Bernoulli of parameter $F(s) = (1+\exp(-s))^{-1}$ (logistic link function) or $F(s) = (2\pi)^{-1/2}\int_{-\infty}^{s} \exp(-u^2/2)du$ (probit link function),
  - $p(A|s) = \int_A s^{-1/2} f\left(s^{-1/2}u\right) du$, corresponding to a GARCH process $Y_t = \varepsilon_t \sqrt{\lambda_t}$ and f probability density of ε.

The latent process MCRE

For a stationary and ergodic process X, our aim is to study processes defined by

$$\mathbb{P}\left(Y_t \in A|(X_s, Y_u, \lambda_{u-1}); s \in \mathbb{Z}, u \leq t-1\right) = p\left(A|\lambda_t\right),$$

$$\lambda_t = f\left(\lambda_{t-1}, Y_{t-1}, X_{t-1}\right).$$

- Conditional on X, the process $(\lambda_t)_t$ is a non-homogeneous Markov chain with (random) transition kernels

$$P_{X_t} h(s) = \int h \circ f\left(s, y, X_t\right) p\left(dy|s\right).$$

- The bivariate process $(Y_t, \lambda_t)_t$ is also a Markov chain in random environments.

Outline

1. Markov chains, strict exogeneity and random environments
   - Motivation and general setup
   - Existence of stationary measures via a coupling method
   - Ergodic properties

2. Observation-driven models and random environments
   - Model formulation
   - Existence of stationary solutions under semi-contractivity conditions
   - Some examples

Structure of the latent process λ_t

- Our aim is to define complex models such as threshold models, the mapping $y \mapsto f(s, y, x)$ is not necessarily continuous.

- For deterministic environments, [Doukhan, Douc & Moulines (2013)], [Wang, Liu, Yao, Davis & Li (2014)] or [Doukhan & Neumann (2019)] already studied this problem of threshold for Poisson autoregressions.

- For deterministic environments, the Markov chain $(\lambda_t)_t$ does not satisfy the standard irreducibility assumption when $(Y_t)_t$ is discrete. Techniques based on coupling or the theory of $T-$chains have been used.

Assumptions

A1 There exists a measurable function $\kappa : \mathbb{R}^d \to \mathbb{R}_+$ such that $\mathbb{E}\log^+ \kappa(X_0) < \infty$, $\mathbb{E}\log\kappa(X_0) < 0$ and for all $y \in \mathbb{R}$, $s, s' \in \mathbb{R}$ and $x \in \mathbb{R}^d$,

$$|f(s, y, x) - f(s', y, x)| \leq \kappa(x)|s - s'|.$$

A2 There exist three measurable functions γ, δ, V and $\alpha \in (0, 1]$ such that $\mathbb{E}\log^+ \delta(X_0) < \infty$, $\mathbb{E}\log^+ \gamma(X_0) < \infty$, $\mathbb{E}\log\gamma(X_0) < 0$, $V(s) \geq |s|^\alpha$ for $s \in L$ and

$$P_x V(s) \leq \gamma(x)V(s) + \delta(x).$$

A3 There exists a polynomial function ϕ, with positive coefficients, vanishing at 0 and such that for every $(s, s') \in \mathbb{R}^2$,

$$d_{TV}\left(p(\cdot|s), p(\cdot|s')\right) \leq 1 - \exp\left(-\phi\left(|s - s'|\right)\right).$$

Result

Theorem 3 (Doukhan, Neumann, T. (2023))

Let Assumptions **A1-A3** *hold true. There then exists a stationary and ergodic process* $(Y_t, \lambda_t, X_t)_{t\in\mathbb{Z}}$ *solution of the recursions. The distribution of such process is unique.*

Assumption **A2** is satisfied with $V(s) = 1 + |s|$ if there exist functions $\delta_j : \mathbb{R}^d \to \mathbb{R}_+$, $1 \leq j \leq 3$ s.t. $\mathbb{E}\log^+ \delta_j(X_0) < \infty$, $\mathbb{E}\log\left(\delta_1(X_0) + \delta_2(X_0)\right) < 0$ and

$$|f(s, y, x)| \leq \delta_1(x)|s| + \delta_2(x)|y|^i + \delta_3(x)$$

and $\int |y|^i p(dy|s) \leq |s| + Cte$.

Sketch of the proof. Maximal coupling

- The first step is to adapt the proof of [Doukhan & Neumann (2019)] based on the maximal coupling.

- We define two processes $((Y_t, \lambda_t))_{t\geq 0}$ and $((Y'_t, \lambda'_t))_{t\geq 0}$ and a probability measure $\overline{\mathbb{P}}_\omega$ such that $\lambda_0 = s$, $\lambda'_0 = s'$ and for $t \geq 0$,

$$\overline{\mathbb{P}}_\omega\left(Y_t \neq Y'_t | \lambda_t, \lambda'_t\right) = d_{TV}\left[p\left(\cdot|\lambda_t\right), p\left(\cdot|\lambda'_t\right)\right].$$

  We then define

$$\lambda_{t+1} = f\left(\lambda_t, Y_{t-1}, X_{t-1}(\omega)\right), \quad \lambda'_{t+1} = f\left(\lambda'_t, Y'_{t-1}, X_{t-1}(\omega)\right).$$

- For deterministic environments, the drift condition allows to control the tail probability of the return times of the process $(\lambda_t, \lambda'_t)_{t\geq 0}$ in the center (say a ball C) of the state space.

- When at a given time t, $(\lambda_t, \lambda'_t) \in C$, Assumptions **A1** and **A3** ensure a positive lower bound for the probability of fastening the paths, i.e. of the event $Y_{t+i} = Y'_{t+i}$ for $i \geq 0$ which in turn provides a decreasing upper bound for $\left|\lambda_{t+i} - \lambda'_{t+i}\right|$.

Sketch of the proof. Subsampling the chain for stabilizing the environment effect

- We carefully adapt the previous argument by studying the effect of the coupling near "favorable" random time points $0 < \tau_1(\omega) < \tau_2(\omega) < \cdots$ only depending on the covariate process $(X_t(\omega))_t$.

- These random time points are chosen so that the sub-Markov chain $\left(\lambda_{\tau_i(\omega)}, \lambda'_{\tau_i(\omega)}\right)_i$ satisfies a drift condition with non-random constants.

- Moreover the random time points are chosen to get a non-random lower bound for fastening the paths, i.e. the probability of getting an equality $Y_{\tau_i(\omega)+j} = Y'_{\tau_i(\omega)+j}$ for $j \geq 0$, when the $\left(\lambda_{\tau_i(\omega)}, \lambda'_{\tau_i(\omega)}\right)$ goes back to the center of the state space.

$$\mathcal{W}_1(\mu,\nu) = \inf\left\{\int (|s-s'|\wedge 1)\,\gamma(ds,ds')\right\},$$

where the infimum is on the set of probability measures γ with marginals μ and ν.

Proposition 3

There exist $C>0$ and $\rho\in(0,1)$ s.t. for all s,s',

$$\begin{aligned} & \mathcal{W}_1\left(\delta_s P_{X_0(\omega)}\cdots P_{X_{n-1}(\omega)}, \delta_{s'} P_{X_0(\omega)}\cdots P_{X_{n-1}(\omega)}\right) \\ \leq\ & C\left(1+V(s)+V(s')\right)\rho^{\sqrt{M_n(\omega)}}, \end{aligned}$$

where $M_n(\omega)$ denotes the number of random points $\tau_i(\omega)$ between time $t=0$ and time $t=n$.

Sketch of the proof. Almost sure convergence of the backward iterations

Recall the definition of the two Markov kernels,

$$P_{X_t(\omega)}h(s) = \int h\left(f\left(s, y, X_t(\omega)\right)\right) p(dy|s),$$

$$R_{X_t(\omega)}\overline{h}(y, s) = \int \overline{h}\left(y', f\left(s, y, X_t(\omega)\right)\right) p\left(dy' | f\left(s, y, X_t(\omega)\right)\right).$$

Proposition 4

Let Assumptions **A1-A3** *hold true.*

1. *There then exists a unique process* $(\pi_t)_{t\in\mathbb{Z}}$ *of identically distributed random probability measures such that and such that* $\pi_t P_{X_t} = \pi_{t+1}$ *a.s. Moreover, almost surely, for any* s,

$$\lim_{n\to\infty} \mathcal{W}_1\left(\delta_s P_{X_{t-n}} \cdots P_{X_{t-1}}, \pi_t\right) = 0.$$

2. *As a consequence,* $\nu_t(dy, ds) = p(dy|s)\pi_t(ds)$ *is the unique process of identically distributed random measures s.t.* $\nu_t R_{X_t} = \nu_{t+1}$ *a.s.*

Existence of a unique stationary path and ergodic properties can be obtained.

Outline

Poisson autoregressions

$$p(k|s) = \exp(-s)s^k/k!, \quad \lambda_t = f\left(\lambda_{t-1}, Y_{t-1}, X_{t-1}\right).$$

- The result applies as soon as

$$|f(s,y,x) - f(s',y,x)| \leq \kappa(x)\left|s - s'\right|,$$

$$|f(s,y,x)| \leq \kappa(x)|s| + \widetilde{\kappa}(x)y + \gamma(x)$$

and the required conditions on log-moments hold true.

- The example $f(s,y,x) = \kappa(x)s + \widetilde{\kappa}_1(x)y\mathbb{1}_{y\leq c(x)} + \widetilde{\kappa}_2(x)y\mathbb{1}_{y>c(x)} + \delta(x)$ generalizes the threshold Poisson models discussed in previous references.

$$p\,(1|s) = F(s), \quad \lambda_t = f\,(\lambda_{t-1}, Y_{t-1}, X_{t-1})\,.$$

- The results apply to the logistic (i.e. $F(s) = (1 + \exp(-s))^{-1}$) or the probit (i.e. F c.d.f. of $\mathcal{N}(0,1)$).
- For the simple model $f(s,y,x) = \kappa(x)s + \widetilde{\kappa}(x)y + \delta(x)$, only the condition $\mathbb{E}\log|\kappa(X_0)| < 0$ is necessary (up to existence of others log-moments).
- The result applies to models used in econometrics [Kauppi and Saikkonen (2008)], [Russell and Engle (2005)], [Rydberg and Shephard (2003)] and extend or sharpen existing results for such models [Fokianos and Moyssiadis (2014)], [Fokianos and Truquet (2019)], [Truquet (2020)].

GARCH-type processes

$$Y_t = \varepsilon_t \sqrt{\lambda_t}, \quad \lambda_t = f\left(\lambda_{t-1}, Y_{t-1}, X_{t-1}\right).$$

- The ε_t's are i.i.d. $(0,1)$. The probability density of ε_0 is non-decreasing on $(-\infty, 0]$ and non-increasing on $[0, \infty)$.
- The mapping f is lower-bounded by a positive constant and satisfies the structural assumptions,

$$\left|f(s,y,x) - f(s',y,x)\right| \leq \kappa(x)\left|s - s'\right|,$$

$$|f(s,y,x)| \leq \kappa(x)|s| + \widetilde{\kappa}(x)y^2 + \gamma(x).$$

Extensions and perspectives

- The approach based on the "Markov chains in random environments" setup and the control of the backward iterations is interesting for extending the classical theory of non-linear autoregressive time series.
- Other type of results could be possible (e.g. using other coupling techniques for time-inhomogeneous Markov chains) .
- Mixing type conditions for $(X_t, Y_t)_t$ have been only derived for Doeblin's type chain. General case (?)
- Quenched central limit theorems (?)

References

Doukhan, P., Neumann, M.H. and Truquet, L. *Stationarity and ergodic properties of some observation-driven models in random environments.* The Annals of Applied Probability, **33**, 6B, 5145–5170, 2023.

Kifer, Y. (1996) *Perron-Frobenius theorem, large deviations, and random perturbations in random environments*. Mathematische Zeitschrift, **222**, 677–698, Springer-Verlag.

Lovas, A. and Rásonyi, M. *Markov chains in random environment with applications in queuing theory and machine learning*, Stochastic Processes and their Applications, **137**, 294–326, 2021.

Truquet, L. *Ergodic properties of some Markov chains models in random environments.* arXiv:2108.06211v1.

Truquet L. *Strong mixing properties of discrete-valued time series with exogenous covariates*, Stochastic Processes and their Applications, **160**, 294–317, 2023.