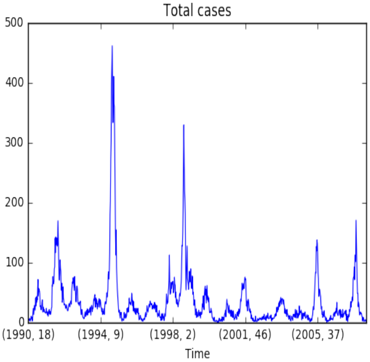
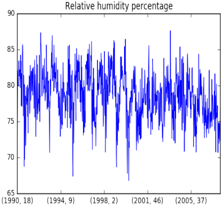
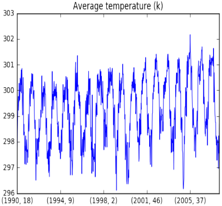
Time series, Exogeneity and Random Environments

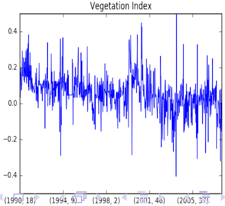
Lionel Truquet (CREST-ENSAI, Rennes),

Real data example: Weekly number of Dengue cases in San Juan (Porto Rico)









Exogenous covariates and non-linear times series

- Let (Y_t) be the time series of interest (e.g. weekly number of cases for the Dengue) and (X_t) the exogenous covariates (e.g. humidity, temperature, vegetation..). Explaining or predicting Y_t using Y_{t-j}, X_{t-j} for $j \geq 1$ are natural questions.
- Defining and studying non-linear autoregressive processes (including discrete-valued time series) that incorporate exogenous regressors is a quite recent topic (at least if we are interested in getting theoretical guarantees for stationarity, consistency of MLE...).
- A few recent contributions: dynamic binary choice models [de Jong and Woutersen (2011)], count autoregressions [Agosto, Cavaliere, Kristensen and Rahbek (2016)], GARCH processes [Pedersen and Rahbek (2018), Francq and Thieu (2019)], beta-autoregressions [Gorgi and Koopman (2020)].
- No general approach available for studying general time series models with exogenous covariates.

What do we call exogenous covariates ?

- Several notions of exogeneity can be found in the econometric literature [Engle, Hendry and Richard (1983)].
  - Strict exogeneity [Sims (1972), Chamberlain (1982)]. Y_t independent to $X_t, X_{t+1}, \ldots$ conditional on $(Y_{t-j}, X_{t-j})_{j\geq 1}$.
  - Grander non causality. X_t is independent of $Y_t, Y_{t-1}, \ldots$ conditional on $X_{t-1}, X_{t-2}, \ldots$.
  - Under an additional regularity condition, Granger non causality is equivalent to strict exogeneity.
  - There exists a weak exogeneity notion [Engle, Hendry and Richard (1983)]. The notion is relative to parameter estimation. Other concepts such as strong exogeneity or super exogeneity are also defined.

Observation-driven models

- 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 practioners 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.

Markov chains in random environments (1)

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

- A direct link between the notion of strict exogeneity [econometrics] and that of random environments [probability].
- Conditional on X, the process $\left(\lambda_t\right)_t$ is a non-homogeneous Markov chain with (random) transition kernels

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

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

Markov chains in random environments (2)

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

- Seminal contributions [Cogburn (1984), Orey (1991), Kifer (1998), Stenflo (2001)] but the state-space of the Markov chains is either discrete or satisfies Doeblin's type condition or uniform contraction conditions.
- If a stationary solution exists, the (conditional) marginal distribution $\lambda_t|X$, 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

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

Outline

1. Stationarity result from contractivity assumptions

2. A semi-contractivity result

Random map dynamic

It is always possible to express the dynamic using random maps. For instance, in $\mathbb{R}$, $Y_t = F_{\lambda_t}^{-1}(U_t)$ where U_t is uniform and $F_s(x) = \int_{-\infty}^{x} p\,(dz|s)$.

- A1 $(U_t)_{t\in\mathbb{Z}}$ is a sequence of uniformly distributed random variable over $[0,1]$, independent from the stationary and ergodic sequence $(X_t)_{t\in\mathbb{Z}}$.
- A2 There exists $s \in \mathbb{R}$ s.t. for any x, $f\left(s, F_s^{-1}(U_0), x\right)$ is integrable.
- A3 Contractivity condition: there exists a nonnegative function κ such that $\mathbb{E}\log^+ \kappa(X_0) < \infty$ and $\mathbb{E}\log\kappa(X_0) < 0$, such that for all $s, s' \in \mathbb{R}$ (or a subset of it) and $x \in \mathbb{R}^d$ (or a subset of it),

$$\mathbb{E}\left[\left|f\left(s, F_s^{-1}(u), x\right) - f\left(s', F_{s'}^{-1}(u), x\right)\right|\right] \leq \kappa(x)|s - s'|.$$

A3 is satisfied as soon as $\int_0^1 \left|F_s^{-1}(u) - F_{s'}^{-1}(u)\right| du \leq |s - s'|$ and

$$|g(s,y,x) - g(s',y',x)| \leq \alpha(x)|s - s'| + \beta(x)|y - y'|$$

with $\mathbb{E}\log\left(\alpha(X_0) + \beta(X_0)\right) < 0$.

Result

Theorem (Debaly and T. (2019))

Under **A1-A3**, *and if*

$$\mathbb{E}\left[\log^+ \int \left| f\left(s, F_s^{-1}(u), X_0\right)\right| du\right] < \infty,$$

there 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. Moreover* $\mathbb{E}\left[|\lambda_t|\,|X\right] < \infty$ *a.s.*

Example: $Y_t|X, Y_{t-1}, \ldots \sim$ Poisson(λ_t), $\lambda_t = f\left(\lambda_{t-1}, Y_{t-1}, X_{t-1}\right)$.

For instance, if $f(s, y, x) = \gamma(x) + \alpha(x)s + \beta(x)y$:

- model with interaction $\beta(x) = \beta_1 + \beta_2 x$,
- model with threshold $\beta(x) = \beta_0 + \sum_{j=1}^{k} \beta_j \mathbb{1}_{x\in A_j}$.

Outline

1. Stationarity result from contractivity assumptions

2. A semi-contractivity result

More refined models

- One can define more complex models, e.g. threshold models, without the full contractivity condition, e.g.

$$f(s,y,x) = \sum_{j=1}^{k} \left[\gamma_j(x) + \alpha_j(x)s + \beta_j(x)y\right] \mathbb{1}_{y \in A_j(x)}.$$

- In the later case, $y \mapsto f(s,y,x)$ is not necessarily continuous.
- For deterministic environments, [Doukhan et al. (2012)] or [Wang et al. (2014)] already studied this problem for Poisson autoregressions.

Another set of Assumptions

B1 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 $x \in \mathbb{R}^d$,

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

B2 There exist functions $\delta_j : \mathbb{R}^d \to \mathbb{R}_+$, $1 \leq j \leq 3$ s.t. $\mathbb{E}\log^+ \delta_j(X_0) < \infty$, $\mathbb{E}\log(\delta_1(X_0) + \delta_2(X_0)) < 0$ and

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

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

B3 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 (Doukhan, Neumann, T. (2020))

Let Assumptions **B1-B3** *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.*

- The convergence of the backward iterations (for defining the process $\lambda_t|X$) is obtained in Wasserstein metric. Coupling methods play an important rule in the proof.
- A first example concerns Poisson autoregressions, $p(k|s) = \exp(-s)s^k/k!$ and the structural assumption on f (with $i = 1$).
- Another example concerns binary time series $p(1|s) = 1 - p(0|s) = F(s)$ with

$$F(s) = \begin{cases} (1+\exp(-s))^{-1} & \text{logistic model}, \\ (2\pi)^{-1/2}\int_{-\infty}^{s}\exp\left(-u^2/2\right)du & \text{probit model} \end{cases}$$

When

$$\lambda_t = \kappa\left(X_{t-1}\right)\lambda_{t-1} + \delta_1\left(X_{t-1}\right)Y_{t-1} + \delta_2\left(X_{t-1}\right),$$

only $\mathbb{E}\log\kappa\left(X_0\right) < 0$ is necessary (up to logarithmic integrability conditions).

To take away...

- Many other models satisfy the assumptions (Categorical time series, GARCH type processes, Negative binomial count autoregressions...), but **B3** is the most delicate assumption to check.
- The approach based on the "Markov chains in random environments" setup is interesting for extending the classical theory of non-linear autoregressive time series. Extensions to higher-order autoregressive processes may be possible.
- Possible extensions to seasonal models (as for the Dengue example) or other forms of nonstationarities (in a sense to precise) could be interesting.

References

- Fokianos, K. and Truquet, L. (2019) *On categorical time series models with covariates.* Stochastic Process. Appl. **129** 3446-3462.
- Truquet, L. (2020) *Coupling and perturbation techniques for categorical time series.* Bernoulli, **26**(4), 3249-32.
- Debaly, Zinsou Max and Truquet, L. (2019) *Iterations of dependent random maps and exogeneity in nonlinear dynamics.* arXiv:1908.00845.
- Doukhan, P., Neumann, M.H. and Truquet, L. (2020) *Stationarity and ergodic properties of some observation-driven models in random environments.* arXiv:2007.07623.