Harris recurrent Markov chains and nonlinear monotone cointegrated models

Carlos Fernández

LTCI, Telecom Paris

fernandez@telecom-paris.fr

July 4, 2025

Nonlinear monotone cointegrated models

Two time series X_t and Z_t are said to be *monotonically nonlinearly cointegrated* if they are both nonstationary and there exists a function f_0 and a stationary process W_t such that

$$Z_t = f_0(X_t) + W_t$$

- f_0 is a monotone non-increasing function.
- $\{W_t\}$ is an unobserved process such that $E(W_t|X_t) = 0$.
- $\{X_t\}$ is a Harris recurrent Markov chain (positive or β-null recurrent).

Ex: seasonal relationship between average monthly temperatures and ice cream consumption, real expenditure and household income (Engel's curve).

For the general theory without shape constraints for f_0 in the markovian case, see Karlsen et al. (2007); Cai and Tjøstheim (2015); Tjøstheim (2020) and Wang and Phillips (2009a,b); Wang (2015).

Outline

1. Harris recurrent Markov chains
2. Estimating f_0
3. Perspectives

Outline

1. Harris recurrent Markov chains
   - Positive and null recurrence
2. Estimating f_0
3. Perspectives

General framework : Harris recurrent Markov chains

$X = (X_n)_{n\in\mathbb{N}}$, a ϕ-irreducible, strongly aperiodic, time-homogeneous Markov chain, valued in a (countable generated) measurable space $(E, \mathcal{E})$ with transition kernel $\Pi(x, dy)$ and initial distribution ν.

Notations: $\mathbb{P}_\nu$ (respectively, $\mathbb{P}_x$ for x in E) the probability measure such that $X_0 \sim \nu$ (resp., conditioned upon $X_0 = x$).

Main references: Meyn et al. (2009); Asmussen (2010); Douc et al. (2018).

General framework : Harris recurrent Markov chains

- An irreducible Markov Chain is **Harris recurrent** if for any set A such that $\phi(A) > 0$, the probability of returning to A is equal to one, no matter the starting point: $\forall x \in E, \mathbb{P}_x(\exists n \geqslant 1 : X_n \in A) = 1$.
- For any set $A \in \mathcal{E}$, the random variables

$$\tau_A = \inf\{n \geqslant 1 : X_n \in A\}.$$
$$\sigma_A = \inf\{n \geqslant 0 : X_n \in A\}.$$

  are called the first return and first visit times on A, respectively. The subsequent return and visit times are denoted by $\tau_A(j)$ and $\sigma_A(j)$, and are defined inductively.
- The number of times a Markov Chain visits a set A, up to time n, is denoted by $T_n(A) = \sum_{t=0}^{n} \mathbb{I}\{X_t \in A\}$. This sequence is known as the occupation time sequence.

General framework : Harris recurrent Markov chains

- A Harris Markov chain X is said to be **atomic** when it possesses an accessible atom, *i.e.* a measurable set $\boldsymbol{\alpha}$ such that $\phi(\boldsymbol{\alpha}) > 0$ and $\Pi(x, .) = \Pi(y, .)$ for all x, y in $\boldsymbol{\alpha}$.
- The sample paths of an atomic Markov Chain may be divided into independent blocks of random length corresponding to consecutive visits to $\boldsymbol{\alpha}$: $\mathcal{B}_0 = \left(X_0, X_1, \ldots, X_{\tau_{\boldsymbol{\alpha}}(1)}\right)$, $\mathcal{B}_1 = \left(X_{\tau_{\boldsymbol{\alpha}}(1)+1}, \ldots, X_{\tau_{\boldsymbol{\alpha}}(2)}\right)$, ... $\mathcal{B}_j = \left(X_{\tau_{\boldsymbol{\alpha}}(j)+1}, \ldots, X_{\tau_{\boldsymbol{\alpha}}(j+1)}\right)$, ... Among them, $\{\mathcal{B}_j\}_{j\geqslant 1}$ are i.i.d. with common law $\mathcal{L}_{\mathbb{P}_{\boldsymbol{\alpha}}}\left(X_0, X_1, \ldots, X_{\tau_{\boldsymbol{\alpha}}(1)}\right)$. These blocks are called *regenerations*.
- Denote by $T\,(n)$ the number of complete regenerations up to time n, i.e. $T(n) = max\,(0, T_{\boldsymbol{\alpha}}\,(n) - 1)$.

Example 1: Symmetric random walk

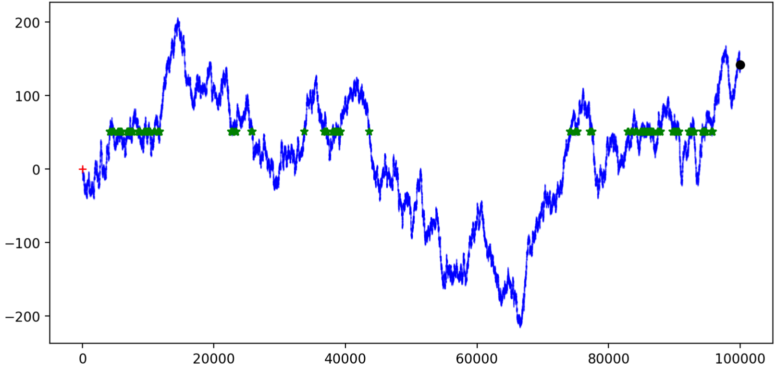


Figure: Simple symmetric random walk (SSRW), $n = 10^5$. This chain has 774 regeneration blocks.

$$X_n = \sum_{k=1}^{n} Y_i \quad , \quad P\left(Y_i = 1\right) = P\left(Y_i = -1\right) = \frac{1}{2}.$$

The split chain

- A Markov chain satisfies the minorization condition $M(m_0, s, \lambda)$ if there exists $m_0 \in \mathbb{N}$, $s \in \mathcal{E}^+$ and a positive non-trivial measure λ such that

$$\Pi^{m_0}(x, A) \geqslant s(x)\,\lambda(A) \quad \forall x \in S, A \in \mathcal{E}. \tag{1}$$

- A function $s \in \mathcal{E}^+$ that satisfies (1) is called a *small function*. A set $S \in \mathcal{E}$ is said to be **small** for X if the function $\mathbb{I}_S$ is small.
- Any Harris recurrent Markov chain that can be "made" atomic. See Nummelin (1978, 1984). This construction is known as the *split chain*.
- From now on, we assume that all Markov chains satisfy the minorization condition $M(1, s, \lambda)$ and we will denote by $\boldsymbol{\alpha}$ the atom of the split chain.

Positive and null recurrence

- A strongly aperiodic *Harris recurrent Markov Chain* admits a unique (up to a multiplicative constant) invariant measure, that is, a measure π that satisfies

$$\pi(B) = \int_E \Pi(x, B)\, d\pi(x).$$

- This measure has the form

$$\pi(B) = \mathbb{E}_{\boldsymbol{\alpha}}\left(\sum_{i=1}^{\tau_{\boldsymbol{\alpha}}} \mathbb{I}\{X_i \in B\}\right), \quad \forall B \in \mathcal{E}. \tag{2}$$

- For any π-integrable function g,

$$\int g d\pi = \mathbb{E}_{\boldsymbol{\alpha}}\left(\sum_{i=1}^{\tau_{\boldsymbol{\alpha}}} g(X_i)\right).$$

- For any small set C, $0 < \pi(C) < +\infty$.

Positive and null recurrence

- An irreducible Harris recurrent Markov Chain is called *positive recurrent* if the invariant measure defined in (2) is finite; if it is only σ-finite, then the chain is called *null-recurrent* (i.e. τ_α is finite with probability 1, but its expectation is infinite).
- A Markov chain is called β-null recurrent if is Harris recurrent and there exists a small function h, an initial measure ν, a constant $\beta \in (0,1)$ and a slowly varying function L_h such that

$$\mathbb{E}_\nu \left[\sum_{t=0}^{n} h\left(X_t\right) \right] \sim \frac{1}{\Gamma\left(1+\beta\right)} n^{\beta} L_h\left(n\right) \tag{3}$$

  as n goes to $+\infty$. See Karlsen and Tjostheim (2001); Karlsen et al. (2007).
- β-null recurrent chains are also called β-regular chains. See Chen (1999, 2000).

Positive and null recurrence

In the atomic case, positive and β-null recurrence are characterized via the first time of return $(\tau_{\boldsymbol{\alpha}})$ to the atom.

- An atomic Harris recurrent Markov chain is positive recurrent if and only if $E_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}} < \infty$ (Kac's theorem).
- Atomic $\beta -$ *null recurrent* chains are characterized by the fact that the time of first return satisfies

$$P\left(\tau_{\boldsymbol{\alpha}} > n\right) = \frac{1}{n^{\beta} L\left(n\right)}, \tag{4}$$

where $L(n)$ is a slowly varying function. In this case,

$$\beta = \sup\left\{p \geqslant 0 : \mathbb{E}_{\boldsymbol{\alpha}}\left[\tau_{\boldsymbol{\alpha}}^{p}\right] < \infty\right\}.$$

Estimation in Harris recurrent Markov Chains

Let g be a π-integrable function, and assume $g(\mathcal{B}_1)$ has finite second moment.

$$S_n(g) = \sum_{k=0}^{n} g(X_k) = g(\mathcal{B}_0) + \sum_{j=1}^{T(n)} g(\mathcal{B}_j) + \sum_{i=\tau_\alpha(T(n)+1)+1}^{n} g(X_i)$$

- $g(\mathcal{B}_j) = \sum_{i=1+\tau_A(j)}^{\tau_A(j+1)} g(X_i)$,
- $g(\mathcal{B}_j)$, $j \geqslant 1$ are i.i.d. and $\mathbb{E}g(\mathcal{B}_j) = \int g d\pi$.
- $\frac{S_n(g)}{T(n)} \xrightarrow{\text{a.s.}} \int g d\pi$,
- $\sqrt{T(n)}\left(S_n(g) - \int g d\pi\right) \to N$.

Estimation in Harris recurrent Markov Chains

Define $u(n) = n$ if $\mathbf{X}$ is positive recurrent and equal to $n^{\beta}L(n)$ if is β-null recurrent.

- If $\mathbf{X}$ is positive recurrent, $\frac{T(n)}{u(n)}$ converges almost surely to $\frac{1}{\mathbb{E}\tau_{\alpha}}$.
- If $\mathbf{X}$ is β-null recurrent, $\frac{T(n)}{u(n)}$ converges in distribution to a Mittag-Leffler random variable with index β, and there is no deterministic sequence $a(n)$ such that $\frac{T(n)}{a(n)}$ converges (almost surely or in probability) to a non-zero constant. See Chen (1999).

Outline

1 Harris recurrent Markov chains

2 Estimating f_0

- The estimator
- Localized β-null recurrent Markov chains
- Consistency
- Rate of convergence
- Simulations

3 Perspectives

The problem

Estimate f_0 under

$$Z_t = f_0(X_t) + W_t$$

- f_0 is a monotone non-increasing function.
- $\{W_t\}$ is an unobserved process such that $E(W_t|X_t) = 0$.
- $\{X_t\}$ is a Harris recurrent Markov chain (positive or β-null recurrent).

The estimator

Let C be a small set that contains the point of interest x_0. Having observed $\{(X_t, Z_t)\}_{t=0}^n$, we denote by $T_n(C)$ the number of times that $\mathbf{X}$ visited C up to time n and by $\sigma_C(i)$ the time of the i-th visit. We consider the nonparametric LSE defined as the minimizer of

$$f \mapsto \sum_{i=1}^{T_n(C)} \left(Z_{\sigma_C(i)} - f\left(X_{\sigma_C(i)}\right)\right)^2 \tag{5}$$

over the set of non-increasing functions f on $\mathbb{R}$.

The estimator

Let m be the number of unique values of $X_{\sigma_C(1)}, \ldots, X_{\sigma_C(T_n(C))}$, and $Y_1 < \cdots < Y_m$ be the corresponding order statistics. Then, $\widehat{f}_n(Y_k)$ is the left-hand slope at $\sum_{i=1}^{T_n(C)} \mathbb{I}\{X_{\sigma_C(i)} \leqslant Y_k\}$ of the least concave majorant of the set of points

$$\left\{(0,0),\ \left(\sum_{i=1}^{T_n(C)} \mathbb{I}\{X_{\sigma_C(i)} \leqslant Y_k\}, \sum_{i=1}^{T_n(C)} Z_{\sigma_C(i)} \mathbb{I}\{X_{\sigma_C(i)} \leqslant Y_k\}\right)_{k=1,\ldots,m}\right\}.$$

Localization

We denote by F_n the process defined by

$$F_n(y) = \frac{1}{T_n(C)} \sum_{i=1}^{T_n(C)} \mathbb{I}\{X_{\sigma_C(i)} \leqslant y\} = \frac{1}{T_n(C)} \sum_{t=0}^{n} \mathbb{I}\{X_t \leqslant y, X_t \in C\}$$

for all $y \in \mathbb{R}$, which is a localized version of the empirical distribution function of the X_t's.

For each $y \in \mathbb{R}$, F_n converges almost surely to the distribution function F supported on C and defined by

$$F\left(y\right) = \frac{\pi\left(C \cap (-\infty, y]\right)}{\pi\left(C\right)}.$$

Localization

Lemma

Assume that

- *The functions* f_0 *and* F *are differentiable in* C*, and their derivatives are bounded, in absolute value, above and away from zero in* C*.*
- $\ell_C(\mathcal{B}_1) = \sum_{t\in\mathcal{B}_1} \mathbb{I}\{X_t \in C\}$ *has finite second moment.*

Then, for all sufficiently small $\varepsilon > 0$ *we have,*

$$T_n(C) \sup_{|y-x_0|\leqslant\varepsilon} \left|F_n(y) - F(y)\right|^2 = O_p\left(1\right)$$

$$T_n(C) \sup_{|p-F(x_0)|\leqslant\varepsilon} \left|F_n^{-1}(p) - F^{-1}(p)\right|^2 = O_p\left(1\right).$$

Consistency

Theorem

Suppose that

- *For each n, the random variables $W_1, \ldots, W_n$ are conditionally independent given $\mathcal{F}_n$, $\mathbb{E}(W_t|\mathcal{F}_n) = 0$ and $\operatorname{Var}(W_t|\mathcal{F}_n) \leqslant \sigma^2$ for some $\sigma > 0$.*
- *F is locally continuous and strictly increasing in C.*
- *f_0 is non-increasing, and f_0 is locally strictly decreasing.*
- *f_0 continuous in x_0.*

Then, as $n \to \infty$, one has

$$\widehat{f}_n(x_0) = f_0(x_0) + o_P(1).$$

Rate of convergence

Theorem

If in addition to the hypothesis of previous two theorems the following holds:

- *There exists a constant K and a neighborhood V of 0, such that*

$$\mathbb{E}_\lambda \left(\sum_{t=0}^{\tau_{\breve{\alpha}}} \left(\mathbb{I}_C \{X_t \leqslant x_0 + \gamma\} - \mathbb{I}_C \{X_t \leqslant x_0 - \gamma\} \right) \right) \leqslant K\gamma \quad \forall \gamma \in V.$$

Then, as $n \to \infty$, one has

$$\widehat{f}_n(x_0) = f_0(x_0) + O_P(u(n)^{-1/3}),$$

- In the positive recurrent case, $u(n) = n$, hence we obtain the same rate $n^{-1/3}$ as in the i.i.d. case. In the β-null recurrent case, the rate of convergence is $n^{-\beta/3} L^{-1/3}(n)$.
- The rate of convergence does not depend on the choice of C.

Simulations

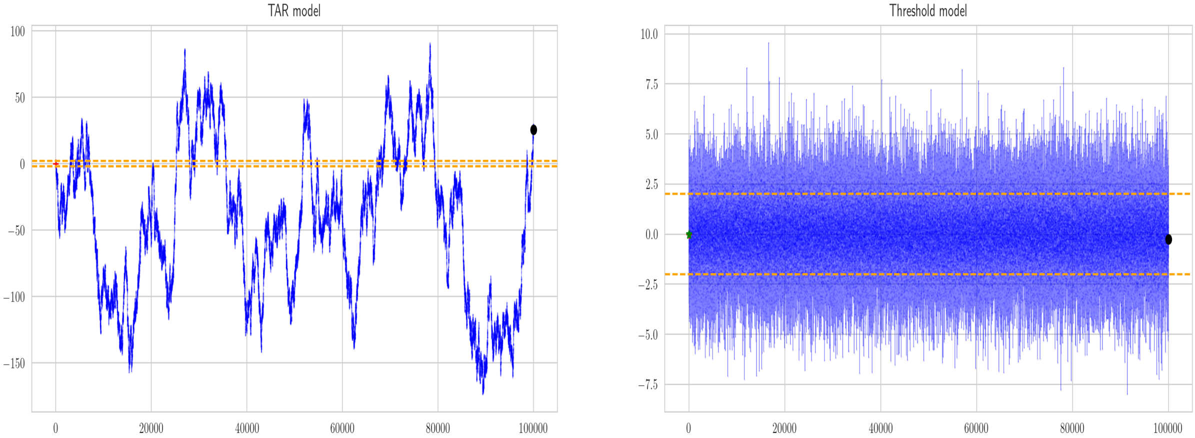
Consider the function

$$f_0(x) = \begin{cases} x^2 + 1 & \text{if } x < 0 \\ \cos(x) & \text{if } 0 \leqslant x \leqslant \frac{\pi}{2} \\ -e^x + e^{\frac{\pi}{2}} & \text{if } x > \frac{\pi}{2} \end{cases}$$

and the following Markov chains.

- Threshold autoregressive (TAR) model with $X_0 = 0$ and $X_n = \alpha_1 X_{n-1} \mathbb{I}\{X_{n-1} \in S\} + X_{n-1} \mathbb{I}\{X_{n-1} \notin S\} + \varepsilon_n$ for $n \geqslant 1$, where $\{\varepsilon_n\}_{n \geqslant 1}$ is an i.i.d sequence of standard Gaussian random variables, $\alpha = 0.5$ and $S = (-\infty, 1]$. β-null recurrent model with $\beta = 0.5$.
- Threshoold model with $X_0 = 0$ and $X_n = A_n \mathbb{I}\{X_{n-1} \geqslant 0\} + B_n \mathbb{I}\{X_{n-1} < 0\}$ for $n \geqslant 1$ where A_n is an i.i.d. sequence of standard Gaussian random variables, B_n is an i.i.d. sequence of centered Gaussian random variables with $\sigma = 2$. A_n and B_n are independent of each other. This is a positive recurrent Markov chain.

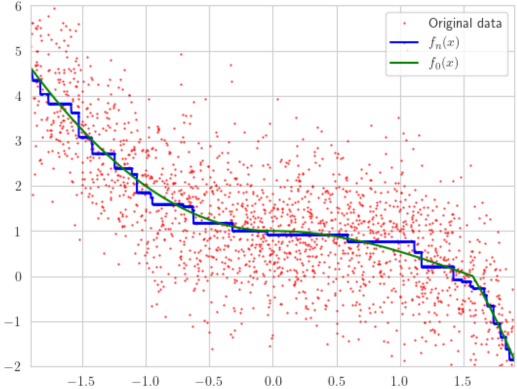
Simulations



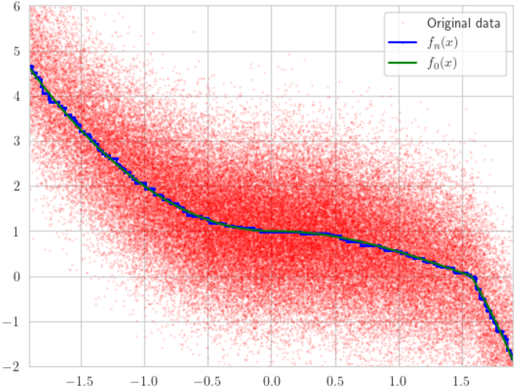
(a) TAR model (b) Threshold model

Figure: Realizations of both chains with $n = 10^5$ points. The orange lines indicate the small set $C = [-2, 2]$.

Simulations



(a) TAR model



(b) Threshold model

Figure: Estimation of f_0 using $\widehat{f}_n$ with $C = [-2, 2]$. For each chain, we have simulated a realization with $n = 10^5$ points. The sequence W_t is taken as an i.i.d. sample of standard Gaussian random variables.

Outline

1 Harris recurrent Markov chains

2 Estimating f_0

3 Perspectives

Perspectives

- Relax the conditional independence assumption between W_t and X_t.
- Find the limit distribution of the $\widehat{f}_n$ estimator.
- Extend the localization approach used here to other shape constrained estimation problems in Markov chains. For example, concave, or log-concave functions.

References I

Asmussen, S. (2010). *Applied Probability and Queues*. Stochastic Modelling and Applied Probability. Springer, 2nd edition.

Cai, B. and Tjøstheim, D. (2015). Nonparametric regression estimation for multivariate null recurrent processes. *Econometrics*, 3(2):265–288.

Chen, X. (1999). How often does a harris recurrent markov chain recur? *The Annals of Probability*, 27.

Chen, X. (2000). On the limit laws of the second order for additive functionals of harris recurrent markov chains. *Probability Theory and Related Fields*, 116.

Douc, R., Moulines, E., Priouret, P., and Soulier, P. (2018). *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer.

Karlsen, H. A., Myklebust, T., and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration type model. *The Annals of Statistics*, 35(1):252–299.

Karlsen, H. A. and Tjostheim, D. (2001). Nonparametric estimation in null recurrent time series. *The Annals of Statistics*, 29.

References II

Meyn, S., Tweedie, R., and Glynn, P. (2009). *Markov chains and stochastic stability*. Cambridge Mathematical Library. Cambridge University Press, 2 edition.

Nummelin, E. (1978). A splitting technique for Harris recurrent chains. *Z. Wahrsch. Verw. Gebiete*, 43:309–318.

Nummelin, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Tracts in Mathematics 83. Cambridge University Press.

Tjøstheim, D. (2020). Some notes on nonlinear cointegration: A partial review with some novel perspectives. *Econometric Reviews*, 39(7):655–673.

Wang, Q. (2015). *Limit Theorems for Nonlinear Cointegrating Regression*. World Scientific.

Wang, Q. and Phillips, P. C. B. (2009a). Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, 25(3):710–738.

Wang, Q. and Phillips, P. C. B. (2009b). Structural nonparametric cointegrating regression. *Econometrica*, 77(6):1901–1948.