Asymptotical analysis for bifurcating autoregressive processes via a martingale approach

B. Bercu B. de Saporta **A. Gégout-Petit**

Université de Bordeaux, IMB UMR 5251 and INRIA Project CQFD

Rennes 27-29 août 2008

Institut de Mathématiques de Bordeaux

Table of contents

Institut de Mathématiques de Bordeaux

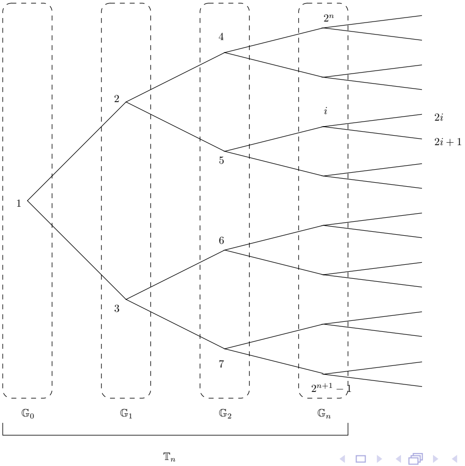
Bifurcating AutoRegressive (BAR) processes

Definition

$$\begin{cases} X_{2n} = a + bX_n + \varepsilon_{2n}, \\ X_{2n+1} = c + dX_n + \varepsilon_{2n+1}. \end{cases}$$

- X_1 is the ancestor
- $(\varepsilon_{2n}, \varepsilon_{2n+1})$ is the driven noise
- Aim : inference on the deterministic coefficient (a, b, c, d)

Associated binary tree



Institut de Mathématiques de Bordeaux

Characteristics of the binary tree

- The two offsprings of n are labelled $2n$ and $2n+1$
- The mother of n is $[n/2]$.
- individual $n \in \mathbb{G}_{r_n}$ with $r_n = \log_2(n)$.
- The ancestors of n are $[n/2], [n/2^2], \ldots, [n/2^{r_n}]$.
- $\mathbb{G}_0 = \{1\}$, $\mathbb{G}_1 = \{2,3\}$, $\mathbb{G}_2 = \{4,5,6,7\}$
- The n-th generation is $\mathbb{G}_n = \{2^n, 2^n + 1, \ldots, 2^{n+1} - 1\}$.
- subtree $\mathbb{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_k$
- Cardinality $|\mathbb{G}_n| = 2^n$, $|\mathbb{T}_n| = 2^{n+1} - 1$

Filtration associated to the tree

Definition

$$\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$$

Remark :

- The filtration is indexed by the generations of the tree.
- The information of $\mathcal{F}_n$ grows exponentially (2^n terms by step !)

Table of contents

Institut de Mathématiques de Bordeaux

Previous works I

stationary processes assumptions

- Cowan,Staudte, Biometrics,1986 Intro
- Huggins, Ann of Stat,1996: MLE large tree
- Huggins, Basawa,J Appl Probab, 1999: MLE small ind trees and Aust. N. Z. J. Stat.2000: MLE large tree and higher order
- Basawa, Zhou,J. Appl. Probab., 2004 expo BAR process Statist. Probab. Lett.,2005 LSE J. Time Ser. Anal.,2005

Previous works II

No stationarity

Guyon, Annals of Applied Probability, 2007 : LSE, iid gaussian hypotheses on $\varepsilon_{2n}, \varepsilon_{2n+1}$, extensively use of the tree structure and Markov chains

Our paper : LSE, relaxed hypotheses on $\varepsilon_{2n}, \varepsilon_{2n+1}$: martingales difference , moments of order 4 (8 for the CLT)

Delmas, Marsalle : Hyp Guyon + possibility for a cell to die.

Table of contents

Institut de Mathématiques de Bordeaux

Assumptions

$$\begin{cases} X_{2n} = a + bX_n + \varepsilon_{2n}, \\ X_{2n+1} = c + dX_n + \varepsilon_{2n+1}. \end{cases}$$

- $\mathbb{E}[X_1^8] < \infty$
- $0 < \max(|b|, |d|) < 1 \quad \text{and} \quad |a| + |c| \neq 0.$

Assumptions

(H.1) $\forall n \geq 0$ and $\forall k \in \mathbb{G}_{n+1}$

$$\mathbb{E}[\varepsilon_k | \mathcal{F}_n] = 0 \qquad \text{and} \qquad \mathbb{E}[\varepsilon_k^2 | \mathcal{F}_n] = \sigma^2 > 0 \qquad \text{a.s.}$$

(H.2) $\forall n \geq 0$ and $\forall k \neq l \in \mathbb{G}_{n+1}$,
- if $[k/2] \neq [l/2]$, ε_k and ε_l are cond. indep. given $\mathcal{F}_n$
- if $[k/2] = [l/2]$, for $\rho < \sigma^2$

$$\mathbb{E}[\varepsilon_k \varepsilon_l | \mathcal{F}_n] = \rho \qquad \text{a.s.}$$

(H.3)

$$\sup_{n \geq 0} \sup_{k \in \mathbb{G}_{n+1}} \mathbb{E}[\varepsilon_k^4 | \mathcal{F}_n] < \infty \qquad \text{a.s.}$$

Stronger assumptions for the CLT

Table of contents

Institut de Mathématiques de Bordeaux

The estimator

We want to estimate $\theta = (a, b, c, d)^t$ by its least squares estimator :

$$(\widehat{\theta}_n) = \begin{pmatrix} \widehat{a}_n \\ \widehat{b}_n \\ \widehat{c}_n \\ \widehat{d}_n \end{pmatrix} = (\mathrm{I}_2 \otimes S_{n-1}^{-1}) \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_{2k} \\ X_k X_{2k} \\ X_{2k+1} \\ X_k X_{2k+1} \end{pmatrix}$$

$$\text{with } S_n = \sum_{k \in \mathbb{T}_n} \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}$$

Table of contents

Institut de Mathématiques de Bordeaux

Under ***(H.1)***, ***(H.2)*** and ***(H.3)***

$$\lim_{n\to+\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \varepsilon_k = 0 \qquad \text{a.s.}$$

$$\lim_{n\to+\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \varepsilon_k^2 = \sigma^2 \qquad \text{a.s.}$$

$$\lim_{n\to+\infty} \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k\in\mathbb{T}_{n-1}} \varepsilon_{2k}\varepsilon_{2k+1} = \rho \qquad \text{a.s.}$$

Lemma

Under ***(H.1)***, ***(H.2)*** and ***(H.3)***

$$\lim_{n\to\infty} \frac{S_n}{|\mathbb{T}_n|} = L \qquad \text{a.s.}$$

Table of contents

An exponential martingale

$$(\widehat{\theta}_n - \theta) = (\mathrm{I}_2 \otimes S_{n-1}^{-1}) \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} \varepsilon_{2k} \\ X_k \varepsilon_{2k} \\ \varepsilon_{2k+1} \\ X_k \varepsilon_{2k+1} \end{pmatrix} = (\mathrm{I}_2 \otimes S_{n-1}^{-1}) M_n$$

M_n is an $\mathcal{F}_n$-martingale with $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$

Table of contents

Convergence's result

We will extensively use the strong law of large number for martingales :

L^2 martingale's convergence Theorem

If M_n is a scalar L^2 $\mathcal{F}_n$-martingale of associated increasing process $< M >_n$. If $\lim_{n\to\infty} < M >_n = +\infty$, then $\frac{M_n}{<M>_n} \to 0$ a.s.

$\widehat{\theta}_n$ convergence

$<M>_n = \Gamma \otimes S_{n-1}$ and we have seen that

$$\lim_{n\to\infty} \frac{S_n}{|\mathbb{T}_n|} = L \qquad \text{a.s.}$$

This convergence implies $\lim_{n\to\infty} <M>_n = +\infty$ and the SLLN for martingales gives :

First convergence result

$$(\widehat{\theta}_n - \theta) = (\mathrm{I}_2 \otimes S_{n-1}^{-1}) M_n = \mathcal{O}\left(\frac{M_n}{<M>_n}\right) \to 0 \quad \text{a.s.}$$

Convergence rate

L^2 martingale's rate of convergence Theorem (Neveu)

If M_n is a scalar L^2 $\mathcal{F}_n$-martingale of associated increasing process $< M >_n$. Under condition on moments, if $\lim_{n\to\infty} < M >_n = +\infty$, then $\left(\frac{M_n}{<M>_n}\right)^2 = \mathcal{O}(\frac{log(<M>_n)}{<M>_n})$ a.s.

Unfortunately, this theorem doesn't apply in our case (M_n is a vectorial, exponential martingale !).

Exponential Martingale

Let Φ_n of dimension 2×2^n and ξ_n of dimension 2^n be given by

$$\Phi_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{2^n} & X_{2^n+1} & \cdots & X_{2^{n+1}-1} \end{pmatrix} \qquad \xi_n = \begin{pmatrix} \varepsilon_{2^n} \\ \varepsilon_{2^n+2} \\ \cdot \\ \varepsilon_{2^{n+1}-2} \\ \varepsilon_{2^n+1} \\ \\ \cdot \\ \varepsilon_{2^{n+1}-1} \end{pmatrix}$$

Then $M_n = \sum_{k=1}^{n} \Psi_{k-1}\xi_k$ where $\Psi_n = \mathrm{I}_2 \otimes \Phi_n$

Rate of convergence for $\widehat{\theta}_n - \theta$

But after some work and calculations (sometimes tricky, sometimes tedious !) we have proved :

Rate of convergence

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log |\mathbb{T}_n|}{|\mathbb{T}_n|} \right) \qquad \text{a.s.}$$

Tools for the convergence of $\widehat{\theta}_n - \theta$

Convergence of $\| \widehat{\theta}_n - \theta \|^2$ is linked to that of

$$\mathcal{V}_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1} (\widehat{\theta}_n - \theta) \text{ by}$$

$$\|\widehat{\theta}_n - \theta\|^2 \leq \frac{\mathcal{V}_n}{\lambda_{\min}(\Sigma_{n-1})}$$

where $\Sigma_n = I_2 \otimes S_n$ and $\lambda_{\min}(S_n) = \mathcal{O}(|\mathbb{T}_n|)$

Proposition

$$\mathcal{V}_n = \mathcal{O}(n)$$

Institut de Mathématiques de Bordeaux

Tools for the convergence of $\widehat{\theta}_n - \theta$

$$\begin{aligned}\mathcal{V}_{n+1} &= M_{n+1}^t \Sigma_n^{-1} M_{n+1} = (M_n + \Delta M_{n+1})^t \Sigma_n^{-1} (M_n + \Delta M_{n+1}), \\ &= \mathcal{V}_n - M_n^t(\Sigma_{n-1}^{-1} - \Sigma_n^{-1}) M_n + 2 M_n^t \Sigma_n^{-1} \Delta M_{n+1} + \Delta M_{n+1}^t \dots\end{aligned}$$

By iteration, we obtain the main decomposition

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1}, \qquad (1)$$

where

$$\mathcal{A}_n = \sum_{k=2}^{n} M_k^t (\Sigma_{k-1}^{-1} - \Sigma_k^{-1}) M_k,$$

$$\mathcal{B}_{n+1} = 2 \sum_{k=2}^{n} M_k^t \Sigma_k^{-1} \Delta M_{k+1} \quad \text{and} \quad \mathcal{W}_{n+1} = \sum_{k=2}^{n} \Delta M_{k+1}^t \Sigma_k^{-1} \Delta M_{k+1}.$$

Institut de Mathématiques de Bordeaux

Tools for the convergence of $\widehat{\theta}_n - \theta$

- Ricatti's Lemma:
  $\Sigma_{n-1}^{-1} - \Sigma_n^{-1} = \Sigma_{n-1}^{-1}\Psi_n\big(\mathrm{I}_2 \otimes (\mathrm{I}_{2^n} - h_n)\big)\Psi_n^t\Sigma_{n-1}^{-1}$
- We write each term into a sum of a square integrable martingale and a rest
- We use the Strong Law of Large number for martingale

Quadratic Strong Law

Result

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |\mathbb{T}_k| (\widehat{\theta}_k - \theta)^t \Lambda (\widehat{\theta}_k - \theta) = 4\sigma^2 \qquad \text{a.s.}$$

with $\Lambda = I_2 \otimes L$.

Tool : Wei's Lemma

Table of contents

Institut de Mathématiques de Bordeaux

Central Limit Theorem. Assumptions

(H.4) For all $n \geq 0$ and for all $k \in \mathbb{G}_{n+1}$

$$\mathbb{E}[\varepsilon_k^4 | \mathcal{F}_n] = \tau^4 \qquad \text{a.s.}$$

Moreover, for all different $k, l \in \mathbb{G}_{n+1}$ with $[k/2] = [l/2]$ and for $\nu^2 < \tau^4$

$$\mathbb{E}[\varepsilon_k^2 \varepsilon_l^2 | \mathcal{F}_n] = \nu^2 \qquad \text{a.s.}$$

(H.5)

$$\sup_{n \geq 0} \sup_{k \in \mathbb{G}_{n+1}} \mathbb{E}[\varepsilon_k^8 | \mathcal{F}_n] < \infty \qquad \text{a.s.}$$

Central Limit Theorem

Theorem

If $(\varepsilon_{2n}, \varepsilon_{2n+1})$ satisfy **H1–H5**, we have

$$\sqrt{|\mathbb{T}_n|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma \otimes L^{-1}).$$

Tool : Application of the CLT for martingale difference sequences (Hamilton). Change of filtration : thanks to **H2**

$M'_n = \sum_{k \leq n} \begin{pmatrix} \varepsilon_{2k} \\ X_k \varepsilon_{2k} \\ \varepsilon_{2k+1} \\ X_k \varepsilon_{2k+1} \end{pmatrix}$ is again a $\mathcal{G}_n$- martingale with

$$\mathcal{G}_n = \sigma\Big(\mathcal{F}_{r_n} \cup \{(X_{2k}, X_{2k+1}),\ k \in \mathbb{G}_{r_n},\ k \leq n\}\Big).$$

Definition

Definition

$$\widehat{\sigma}_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} (\widehat{\varepsilon}_{2k}^2 + \widehat{\varepsilon}_{2k+1}^2)$$

$$\widehat{\rho}_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} (\widehat{\varepsilon}_{2k} \widehat{\varepsilon}_{2k+1})$$

with for $k \in \mathbb{G}_n$:

$$\begin{cases} \widehat{\varepsilon}_{2k} = X_{2k} - \widehat{a}_n - \widehat{b}_n X_k, \\ \widehat{\varepsilon}_{2k+1} = X_{2k+1} - \widehat{c}_n - \widehat{d}_n X_k. \end{cases}$$

Convergence

Theorem

Under **H1–H3**, we have

$$\widehat{\sigma}_n^2 - \sigma_n^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_n|}\right) \quad \text{a.s.}$$

$$\widehat{\rho}_n - \rho_n = \mathcal{O}\left(\frac{n}{|\mathbb{T}_n|}\right) \quad \text{a.s.}$$

with

$$\sigma_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} (\varepsilon_{2k}^2 + \varepsilon_{2k+1}^2) \text{ and } \rho_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} \varepsilon_{2k}\varepsilon_{2k+1}.$$

Institut de Mathématiques de Bordeaux

CLT

Theorem

Under **H1–H5**, we have

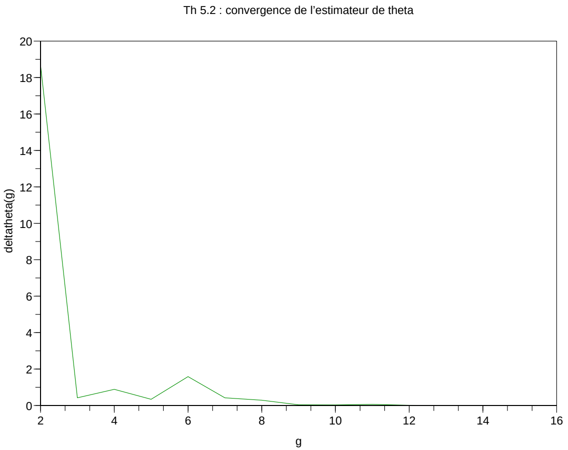
$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\sigma}_n^2 - \sigma^2) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{\tau^4 - 2\sigma^4 + \nu^2}{2}\Big)$$

and

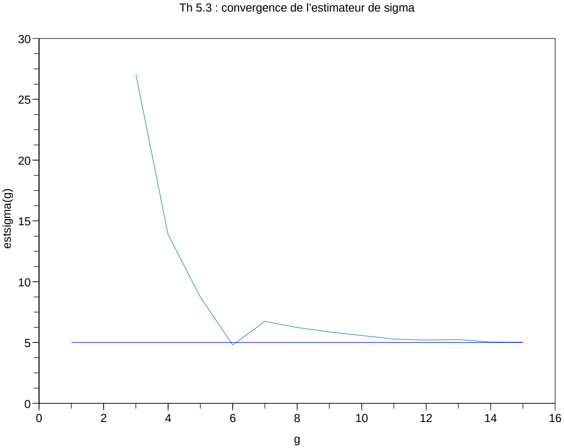
$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \nu^2 - \rho^2).$$

Institut de Mathématiques de Bordeaux

$||\hat{\theta}_n - \theta||^2$'s Convergence

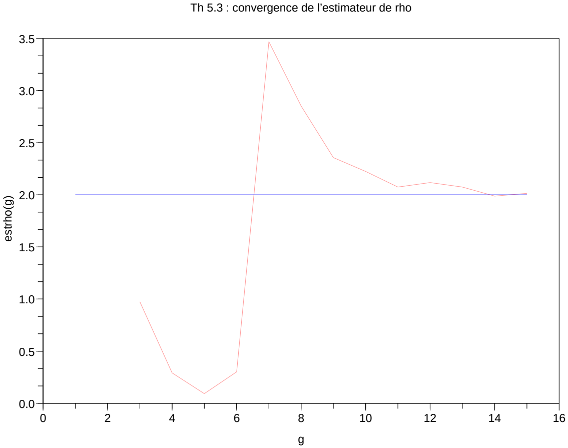


σ's convergence

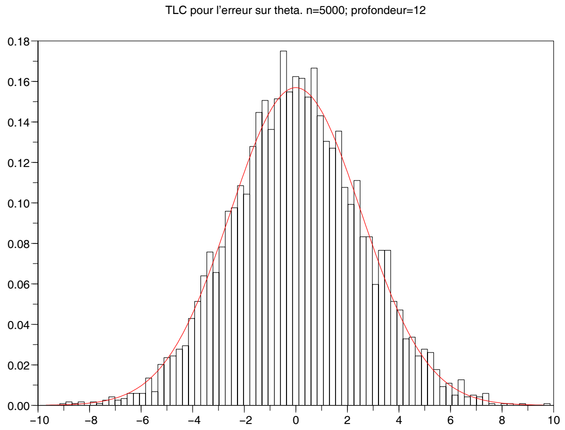


Institut de Mathématiques de Bordeaux

ρ's convergence



CLT for $\hat{a}$



Institut de Mathématiques de Bordeaux

End

Thank you for your attention

Parameter

- $\varepsilon(2k) = A_k Y_k$ and $\varepsilon(2k+1) = A_k Y_k Z_k$
- $A_k = 1$ if $X_k <= 5$, -1 otherwise
- Y_k iid indep from F_k, $P(Y_k = -1) = 1 - P(Y_k = -1) = 5/6$
- Z_k iid indep from F_k and X_k, $Z_k \simeq \mathcal{N}(\rho/\sigma^2; 1-(\rho/\sigma^2)^2)$.
- $a = 3, b = 0.7, c = 2, d = 0.3, \sigma^2 = 5, \rho = 2, X_1 = 1$

Institut de Mathématiques de Bordeaux