Statistical analysis of causal affine processes.

From articles with Y. Boularouk (Algiers), W. Kengne (Cergy) and O. Wintenberger (Paris 6)

Jean-Marc Bardet, SAMM, Université Paris 1 Panthéon-Sorbonne
bardet@univ-paris1.fr

CIRM 2017, Mathematical Methods of Modern Statistics

13 July 2017

Examples: Causal AR[∞] and ARCH(∞) models

With $(\xi_t)_{t\in\mathbb{Z}}$ a sequence of centered i.i.d.r.v.,

- AR(∞) processes $X_t = \sum_{i=1}^{\infty} a_i X_{t-i} + \xi_t$

  $\Longrightarrow$ Causal ARMA(p,q) processes $X_t + \sum_{i=1}^{p} a_i X_{t-i} = \xi_t + \sum_{i=1}^{q} b_i \xi_{t-i}$.

- ARCH(∞) processes, (Robinson, 1991), with $b_0 > 0$ and $b_j \geq 0$

$$\begin{cases} X_t &= \sigma_t \xi_t, \\ \sigma_t^2 &= b_0 + \sum_{j=1}^{\infty} b_j X_{t-j}^2. \end{cases}$$

$\Longrightarrow$ GARCH(p,q) processes, (Bollersev, 1986), with $b_0 > 0$, b_j, $c_j \geq 0$

$$\begin{cases} X_t &= \sigma_t \xi_t, \\ \sigma_t^2 &= b_0 + \sum_{j=1}^{p} b_j X_{t-j}^2 + \sum_{j=1}^{q} c_j \xi_{t-j}^2 \end{cases}$$

A common frame for studying time series

A common class of models for AR, ARMA, ARCH and GARCH processes:

Causal affine models

$$X_t = M(X_{t-1}, X_{t-2}, \ldots)\,\xi_t + f(X_{t-1}, X_{t-2}, \ldots), \quad \forall\, t \in \mathbb{Z}, \text{ a.s..}$$

- $M(\cdot)$ and $f(\cdot)$ are real valued function on $\mathbb{R}^{\mathbb{N}}$;
- $(\xi_t)_{t\in\mathbb{Z}}$ sequence of i.i.d.r.v. with $\mathbb{E}(\xi_0) = 0$ and $\mathbb{E}\big(|\xi_0|^r\big) < \infty$, $r \geq 1$.

Extensions of univariate ARCH models

- TGARCH(∞) processes, (Zakoïan, 1994), with $b_0,\ b_j^+,\ b_j^- \geq 0$

$$\begin{cases} X_t = \sigma_t \, \xi_t, \\ \sigma_t = b_0 + \sum_{j=1}^{\infty} \left[b_j^+ \max(X_{t-j}, 0) - b_j^- \min(X_{t-j}, 0) \right] \end{cases} .$$

- APARCH(δ, p, q) processes, (Ding *et al.*, 1993)

$$\begin{cases} X_t = \sigma_t \, \zeta_t, \\ \sigma_t^\delta = \omega + \sum_{j=1}^{p} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \end{cases}$$

with $\delta \geq 1$, $\omega > 0$, $-1 < \gamma_i < 1$ and $\alpha_i,\ \beta_j \geq 0$.

Combinations of models

- ARMA-GARCH processes, (Ding *et al.*, 1993, Ling and McAleer, 2003)

$$\begin{cases} X_t &= \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \\ \varepsilon_t &= \sigma_t \, \zeta_t, \quad \text{with} \quad \sigma_t^2 = c_0 + \sum_{i=1}^{p'} c_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q'} d_j \sigma_{t-j}^2 \end{cases}$$

- ARMA-APARCH processes, (Ding *et al.*, 1993)

$$\begin{cases} X_t &= \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \\ \varepsilon_t &= \sigma_t \, \zeta_t, \quad \text{with} \quad \sigma_t^\delta = \omega + \sum_{j=1}^{p'} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^{q'} \beta_j \sigma_{t-j}^\delta \end{cases}$$

Existence and stationarity of causal affine models

$$X_t = M(X_{t-1}, X_{t-2}, \ldots)\,\xi_t + f(X_{t-1}, X_{t-2}, \ldots), \quad \forall\, t \in \mathbb{Z},$$

We will assume that f and M satisfy Lipschitzian conditions:

$$\begin{cases} |f(x) - f(y)| & \leq & \sum_{j=1}^{\infty} \alpha_j^{(0)}(f)|x_j - y_j| \\ |M(x) - M(y)| & \leq & \sum_{j=1}^{\infty} \alpha_j^{(0)}(M)|x_j - y_j|. \end{cases}$$

for $x = (x_j)_{j\in\mathbb{N}}$ and $y = (y_j)_{j\in\mathbb{N}}$ two sequences of $\mathbb{R}^{\mathbb{N}}$.

Proposition (from Doukhan and Wintenberger, 2007)

If $\sum_{j=1}^{\infty} \alpha_j^{(0)}(f) + \big(\mathbb{E}(|\xi_0|^r)\big)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M) < 1,$ *there exists a unique causal (X_t is independent of $(\xi_i)_{i>t}$ for $t \in \mathbb{Z}$) solution $(X_t)_{t\in\mathbb{Z}}$ which is strictly stationary, ergodic and such as $\mathbb{E}(|X_0|^r) < \infty$.*

Examples

Conditions on stationarity become:

- Causal AR[∞]:
  $X_t = \sum_{j=0}^{\infty} a_j \xi_{t-j} \Longrightarrow \sum_{j=0}^{\infty} |a_j| < 1$;
- Causal ARCH[∞]:
  $X_t = \xi_t \sqrt{c_0 + \sum_{j=1}^{\infty} c_j X_{t-j}^2} \Longrightarrow \left(\mathbb{E}\left[|\xi_0|^r\right]\right)^{1/r} \sum_{j=1}^{\infty} c_j < 1$;
- Causal TARCH[∞]:
  $$X_t = \xi_t \left(b_0 + \sum_{j=1}^{\infty} \left[b_j^+ \max(X_{t-j},0) - b_j^- \min(X_{t-j},0)\right]\right)$$
  $$\Longrightarrow \left(\mathbb{E}\left[|\xi_0|^r\right]\right)^{1/r} \sum_{j=1}^{\infty} \max\left(b_j^-, b_j^+\right) < 1;$$

Gaussian QMLE of causal affine model

Let $(X_1, \ldots, X_n)$ an observed trajectory of:

$$X_t = M_{\theta^*}(X_{t-1}, X_{t-2}, \ldots)\, \xi_t + f_{\theta^*}(X_{t-1}, X_{t-2}, \ldots), \quad \forall\ t \in \mathbb{Z}$$

- With $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \ldots)$, $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \ldots)$,

  Gaussian conditional log-density: $q_t(\theta) = -\frac{1}{2}\Big[\frac{(X_t - f_\theta^t)^2}{(M_\theta^t)^2} + \log(M_\theta^t)^2\Big]$

- Let $\widehat{f}_\theta^t = f_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$ and $\widehat{M}_\theta^t = M_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$

$$\widehat{q}_t(\theta) = -\frac{1}{2}\Big[\frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} + \ \log\big((\widehat{M}_\theta^t)^2\big)\Big].$$

$\Longrightarrow$ Gaussian QMLE: $\widehat{\theta}_n = \arg\max_{\theta \in \Theta} \widehat{L}_n(\theta)$ with $\widehat{L}_n(\theta) = \sum_{t=1}^n \widehat{q}_t(\theta)$.

Assumptions and strong consistency

We assume:

- **C0**: $r \geq 2$ and $\mathbb{E}(\xi_0^2) = 1$;
- **C1**: Θ is a compact set included in

$$\Theta(r) = \Big\{ \theta \in \mathbb{R}^d \ \Big/ \ \sum_{j=1}^{\infty} \alpha_j^{(0)}(f_\theta) + \big(\mathbb{E}(|\xi_0|^r)\big)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M_\theta) < 1 \Big\}.$$

- **C2**: $\exists \underline{M} > 0$ such that $M_\theta(x) \geq \underline{M}$ for all $\theta \in \Theta$, $x \in \mathbb{R}^{\mathbb{N}}$.
- **C3**: M_θ and f_θ are such that for all $\theta_1, \ \theta_2 \in \Theta$, then:

$$\big(M_{\theta_1} = M_{\theta_2} \quad \text{and} \quad f_{\theta_1} = f_{\theta_2}\big) \quad \Longrightarrow \quad \theta_1 = \theta_2$$

- $\mathbf{A}_0(K, \Theta)$; There exists $\big(\alpha_j^{(0)}(K,\Theta)\big)_j$ such that $\forall x, \ y \in \mathbb{R}^{\mathbb{N}}$

$$\sup_{\theta \in \Theta} \big| K_\theta(x) - K_\theta(y) \big| \leq \sum_{j=1}^{\infty} \alpha_j^{(0)}(K,\Theta) |x_j - y_j|,$$

$$\text{with } \sum_{j=1}^{\infty} \alpha_j^{(0)}(K,\Theta) < \infty.$$

Strong consistency

Theorem (Bardet and Wintenberger, 2009)

Assume $r \geq 2$, $\Theta \subset \Theta(2)$, Conditions C0-3 and $A_0(f,\Theta)$ and $A_0(M,\Theta)$ with

$$\alpha_j^{(0)}(f,\Theta) + \alpha_j^{(0)}(M,\Theta) = O\big(j^{-\ell}\big) \quad \text{for some } \ \ell > 2,$$

then the QMLE $\widehat{\theta}_n$ is strongly consistent, i.e. $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta^*$.

Assumptions on the derivatives

$\mathbf{A}_k(K,\Theta)$: Function $\theta \to K_\theta(\cdot) \in \mathcal{C}^k(\Theta)$ satisfies $\sup_{\theta\in\Theta} \left\| \partial^k_\theta K_\theta(0) \right\| < \infty$ and there exists $\left(\alpha_j^{(k)}(K,\Theta)\right)_j$ a sequence such that $\forall x,\ y \in \mathbb{R}^{\mathbb{N}}$

$$\sup_{\theta\in\Theta} \left\| \partial^k_\theta K_\theta(x) - \partial^k_\theta K_\theta(y) \right\| \leq \sum_{j=1}^{\infty} \alpha_j^{(k)}(K,\Theta) |x_j - y_j|,$$

$$\text{with } \sum_{j=1}^{\infty} \alpha_j^{(0)}(K,\Theta) < \infty.$$

Asymptotic normality

Theorem (Bardet and Wintenberger, 2009)

Under conditions of SLLN, and if $r \geq 4$, *if* $\theta^* \in \overset{\circ}{\Theta} \cap \Theta(4)$ *and if* $\mathbf{A}_k(K, \Theta)$ *for* $k = 1, 2$ *and*

$$\alpha_j^{(1)}(f, \Theta) + \alpha_j^{(1)}(M, \Theta) = O\big(j^{-\ell'}\big) \quad \text{for some} \quad \ell' > 3/2, \tag{1}$$

then the QMLE $\widehat{\theta}_n$ *is asymptotically normal, i.e., there exists matrix* $F(\theta^*)^{-1}$ *and* $G(\theta^*)$ *such that*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}_d\big(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\big). \tag{2}$$

- Could be applied to all cited processes ARMA, ARCH, APARCH,...
- But requires $r \geq 4$ and not very robust.

Laplacian QMLE of causal affine model

Let $(X_1, \ldots, X_n)$ an observed trajectory of:

$$X_t = M_{\theta^*}(X_{t-1}, X_{t-2}, \ldots)\, \xi_t + f_{\theta^*}(X_{t-1}, X_{t-2}, \ldots), \quad \forall\ t \in \mathbb{Z}$$

If $(\xi_t)_t$ standardized Laplacian i.r.v., *i.e.* $f_\xi(t) = \frac{1}{2} \exp(-|t|)$, then

$$\widehat{q}_t(\theta) = -\log|\widehat{M}_\theta^t| - |\widehat{M}_\theta^t|^{-1}|X_t - \widehat{f}_\theta^t|.$$

with $\widehat{f}_\theta^t = f_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$ and $\widehat{M}_\theta^t = M_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$

$\Longrightarrow$ Laplacian QMLE: $\widehat{\theta}_n = \arg\max_{\theta \in \Theta} \widehat{L}_n(\theta)$ with $\widehat{L}_n(\theta) = \sum_{t=1}^{n} \widehat{q}_t(\theta)$.

Strong consistency

Theorem (Bardet and Boularouk, 2016)

Assume $r \geq 1$, $\mathbb{E}(|\xi_0|) = 1$, $\Theta \subset \Theta(r)$, *conditions C1-3 and* $A_0(f,\Theta)$ *and* $A_0(M,\Theta)$ *with*

$$\alpha_j^{(0)}(f,\Theta) + \alpha_j^{(0)}(M,\Theta) = O\big(j^{-\ell}\big) \quad \text{for some} \quad \ell > \frac{2}{\min(r\,,\,2)},$$

then the QMLE $\widehat{\theta}_n$ *is strongly consistent, i.e.* $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta^*$.

Sketch of the proof

The proof proceed in two (classical) steps.

1. $L(\theta) := \mathbb{E}[L_n(\theta)]$ has a unique maximum in θ^*.

2. A uniform law of large numbers on $(\widehat{q}_t)_{t\in\mathbb{N}^*}$ is established by proving

$$\frac{1}{n}\sup_{\theta\in\Theta}\left|\widehat{L}_n(\theta) - L(\theta)\right| \xrightarrow[n\to\infty]{a.s.} 0.$$

This is obtained from Kounias and Weng (1969): $\exists s \in (0,1]$ such as

$$\sum_{t\geq 1}\frac{1}{t^s}\,\mathbb{E}\Big[\sup_{\theta\in\Theta}\left|\widehat{q}_t(\theta) - q_t(\theta)\right|^s\Big] < \infty.$$

Both those results lead to the strong consistency of $\widehat{\theta}_n$, Pfanzagl (1969).

Asymptotic normality

With
$$\begin{cases} \Gamma_F = \left(\mathbb{E}\left[(M^0_{\theta^*})^{-2} \left(\frac{\partial f^0_\theta}{\partial \theta_i} \right)_{\theta^*} \left(\frac{\partial f^0_\theta}{\partial \theta_j} \right)_{\theta^*} \right] \right)_{1 \le i,j \le d} \\ \Gamma_M = \left(\mathbb{E}\left[\left(\frac{\partial \log(M^0_\theta)}{\partial \theta_i} \right)_{\theta^*} \left(\frac{\partial \log(M^0_\theta)}{\partial \theta_j} \right)_{\theta^*} \right] \right)_{1 \le i,j \le d} \end{cases}$$

Theorem (Bardet and Boularouk, 2016)

Under conditions of SLLN, and if $r \geq 2$, if $\theta^* \in \overset{\circ}{\Theta} \cap \Theta(2)$ and if $\mathbf{A}_k(K, \Theta)$ for $k = 1, 2$. Then, if the distribution function of ζ_0 is symmetric, $\mathcal{C}^1(\mathcal{V}_0)$ with derivative $g(0)$ and if Γ_F or Γ_M are definite positive symmetric matrix, then

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}_d\Big(0, \big(\Gamma_M + 2g(0)\,\Gamma_F\big)^{-1}\big((\sigma^2_\zeta - 1)\,\Gamma_M + \Gamma_F\big)\big(\Gamma_M + 2g(0)\,\Gamma_F\big)^{-1}\Big).$$

Comparison with existent results

- ARMA(p, q): Same results than in Davis and Dunsmuir (1997) for LAD estimator.
- ARCH(p): Same results than in Peng and Yao (2003).
- GARCH(p, q): Same results than in Berkes and Horvath (2004) and Franq and Zakoian (2013).
- ARCH(∞), APARCH(δ, p, q), ARMA-GARCH,...: New results.

Sketch of the proof

Let $v = \sqrt{n}(\theta - \theta^*) \in \mathbb{R}^d$. Maximizing $\widehat{L}_n(\theta)$ is equivalent to maximizing

$$W_n(v) = -\sum_{t=1}^{n} \big(q_t(\theta^* + n^{-1/2} v) - q_t(\theta^*) \big)$$
$$= \sum_{t=1}^{n} \log \Big(\frac{(M^t_{\theta^*+n^{-1/2}v})^{-1}}{(M^t_{\theta^*})^{-1}} \Big) + |X_t - f^t_{\theta^*}| \big((M^t_{\theta^*})^{-1} - (M^t_{\theta^*+n^{-1/2}v})^{-1} \big)$$
$$+ \sum_{t=1}^{n} (M^t_{\theta^*+n^{-1/2}v})^{-1} \big(|X_t - f^t_{\theta^*}| - \big| X_t - f^t_{\theta^*+n^{-1/2}v} \big| \big)$$

Theorem (Extension of Theorem 2, Davis and Dunsmuir, 1997)

Let $(Z_t)_{t\in\mathbb{Z}}$ be i.i.d.r.v such as $Var(Z_0) = \sigma^2 < \infty$ with symmetric distribution function $\mathcal{C}^1(\mathcal{V}_0)$ with derivative $g(0)$.
Denote $\mathcal{F}_t = \sigma(Z_t, Z_{t-1}, \cdots)$ for $t \in \mathbb{Z}$ and let $(Y_t)_{t\in\mathbb{Z}}$ and $(V_t)_{t\in\mathbb{Z}}$ two stationary processes adapted to $(\mathcal{F}_t)_t$ and such as $\mathbb{E}\big[Y_0^2 V_0^2\big] < \infty$. Then

$$\sum_{t=1}^{n} V_{t-1} \big(|Z_t - n^{-1/2} Y_{t-1}| - |Z_t| \big) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\Big(g(0)\, \mathbb{E}\big[V_0 Y_0^2\big] \,,\, \mathbb{E}\big[V_0^2 Y_0^2\big] \Big)$$

Simulation results

		n	$\mathcal{L}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{L}$ $\widehat{\theta}_n^{LQL}$	$\mathcal{N}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{N}$ $\widehat{\theta}_n^{LQL}$	t_3 $\widehat{\theta}_n^{GQL}$	t_3 $\widehat{\theta}_n^{LQL}$	$\mathcal{U}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{U}$ $\widehat{\theta}_n^{LQL}$	$\mathcal{M}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{M}$ $\widehat{\theta}_n^{LQL}$
ARMA(1,1)	θ	100	0.106	0.091	0.114	0.117	0.113	0.090	0.112	0.059	0.110	0.078
		1000	0.031	0.024	0.032	0.032	0.036	0.027	0.031	0.014	0.031	0.023
		5000	0.014	0.010	0.014	0.015	0.016	0.011	0.016	0.012	0.013	0.010
	ϕ	100	0.119	0.102	0.121	0.128	0.123	0.102	0.120	0.067	0.121	0.090
		1000	0.037	0.028	0.036	0.036	0.040	0.030	0.036	0.017	0.036	0.027
		5000	0.016	0.012	0.014	0.016	0.017	0.013	0.016	0.007	0.014	0.006
ARCH(1)	ω	100	0.068	0.061	0.048	0.049	0.254	0.085	0.035	0.025	0.062	0.052
		1000	0.020	0.018	0.015	0.015	0.134	0.049	0.011	0.016	0.036	0.018
		5000	0.010	0.009	0.006	0.006	0.115	0.044	0.005	0.015	0.031	0.008
	α	100	0.161	0.155	0.141	0.142	0.979	0.418	0.102	0.064	0.484	0.423
		1000	0.063	0.058	0.043	0.043	0.852	0.169	0.029	0.033	0.157	0.133
		5000	0.016	0.014	0.012	0.012	0.378	0.109	0.013	0.031	0.087	0.062
GARCH(1,1)	α_0	100	0.112	0.105	0.095	0.100	0.211	0.126	0.081	0.047	0.134	0.114
		1000	0.036	0.032	0.028	0.028	0.098	0.058	0.023	0.018	0.066	0.051
		5000	0.016	0.014	0.012	0.012	0.055	0.043	0.010	0.015	0.040	0.023
	α_1	100	0.162	0.157	0.149	0.150	0.453	0.364	0.115	0.070	0.507	0.429
		1000	0.061	0.056	0.449	0.449	0.333	0.150	0.030	0.033	0.160	0.136
		5000	0.029	0.026	0.020	0.020	0.193	0.095	0.013	0.030	0.086	0.058
	β	100	0.225	0.209	0.188	0.190	0.499	0.429	0.163	0.105	0.483	0.390
		1000	0.060	0.055	0.051	0.051	0.285	0.174	0.044	0.022	0.170	0.169
		5000	0.027	0.024	0.022	0.022	0.180	0.075	0.019	0.009	0.072	0.075

Simulation results

		n	$\mathcal{L}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{L}$ $\widehat{\theta}_n^{LQL}$	$\mathcal{N}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{N}$ $\widehat{\theta}_n^{LQL}$	t_3 $\widehat{\theta}_n^{GQL}$	t_3 $\widehat{\theta}_n^{LQL}$	$\mathcal{U}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{U}$ $\widehat{\theta}_n^{LQL}$	$\mathcal{M}$ $\widehat{\theta}_n^{GQL}$	$\mathcal{M}$ $\widehat{\theta}_n^{LQL}$
ARMA(1, 1) -GARCH(1, 1)	θ	100	0.120	0.097	0.107	0.107	0.121	0.098	0.097	0.067	0.123	0.087
		1000	0.035	0.024	0.028	0.028	0.048	0.029	0.024	0.015	0.035	0.026
		5000	0.016	0.010	0.012	0.012	0.023	0.012	0.015	0.011	0.011	0.007
	ϕ	100	0.135	0.109	0.117	0.119	0.141	0.116	0.110	0.077	0.132	0.102
		1000	0.044	0.030	0.033	0.033	0.063	0.035	0.029	0.023	0.053	0.046
		5000	0.020	0.014	0.015	0.015	0.029	0.015	0.013	0.012	0.019	0.014
	α_0	100	0.104	0.096	0.085	0.084	0.158	0.129	0.073	0.055	0.131	0.116
		1000	0.031	0.028	0.025	0.025	0.241	0.060	0.021	0.019	0.053	0.046
		5000	0.014	0.012	0.010	0.010	0.052	0.042	0.009	0.016	0.036	0.019
	α_1	100	0.179	0.177	0.166	0.167	0.469	0.385	0.134	0.107	0.494	0.405
		1000	0.064	0.060	0.045	0.045	0.328	0.161	0.031	0.046	0.160	0.137
		5000	0.031	0.027	0.020	0.020	0.182	0.096	0.013	0.038	0.090	0.062
	β	100	0.302	0.269	0.252	0.233	0.604	0.497	0.217	0.164	0.553	0.472
		1000	0.057	0.051	0.051	0.051	0.312	0.187	0.045	0.049	0.165	0.170
		5000	0.025	0.022	0.020	0.020	0.199	0.073	0.062	0.066	0.019	0.025
ARMA(1, 1) -APARCH(1, 1)	θ	100	0.110	0.086	0.096	0.101	0.112	0.090	0.097	0.068	0.125	0.091
		1000	0.029	0.021	0.023	0.024	0.031	0.021	0.022	0.014	0.033	0.024
		5000	0.013	0.008	0.010	0.010	0.014	0.009	0.010	0.006	0.015	0.011
	ϕ	100	0.138	0.114	0.121	0.126	0.128	0.107	0.111	0.086	0.146	0.107
		1000	0.040	0.027	0.032	0.032	0.041	0.028	0.029	0.026	0.043	0.030
		5000	0.018	0.012	0.012	0.012	0.020	0.012	0.013	0.014	0.019	0.013
	ω	100	0.198	0.192	0.199	0.210	0.254	0.262	0.221	0.170	0.290	0.272
		1000	0.079	0.067	0.056	0.056	0.226	0.218	0.044	0.045	0.142	0.129
		5000	0.033	0.028	0.025	0.025	0.209	0.207	0.017	0.029	0.061	0.056
	α	100	0.206	0.201	0.183	0.184	0.464	0.449	0.167	0.131	0.352	0.327
		1000	0.060	0.053	0.041	0.041	0.447	0.432	0.029	0.043	0.143	0.134
		5000	0.025	0.023	0.018	0.018	0.421	0.414	0.012	0.027	0.071	0.059
	γ	100	0.413	0.386	0.346	0.356	0.439	0.426	0.310	0.233	0.613	0.601
		1000	0.105	0.094	0.071	0.070	0.101	0.092	0.057	0.041	0.217	0.220
		5000	0.042	0.039	0.029	0.029	0.045	0.038	0.024	0.018	0.086	0.089
	β	100	0.297	0.282	0.255	0.238	0.312	0.288	0.186	0.145	0.476	0.468

The problem of change detection

The problem of multiple change detection is the following:

$$X_t = M^t_{\theta_k} \xi_t + f^t_{\theta_k} \quad \text{for } t \in T_k := \{t_{k-1}+1, \ldots, t_k\}$$

for $k = 1, \ldots, K$ and:

- $(X_1, \cdots, X_n)$ is observed;
- $t_k = [n\,\tau_k]$ with $\tau_0 = 0 < \tau_1 < \cdots < \tau_{K+1} = 1$;
- $\theta_k \neq \theta_{k+1}$ for $k = 1, \ldots, K-1$;
- f_θ, M_θ are known, $K, (\tau_k)_k, (\theta_k)_k$ and the law of ξ are unknown.

Some existing results

- Davis *et al.* (1995): detecting changes in AR process;
- Lavielle and Moulines (2000) : detecting changes in the mean of a random process using least-squares contrast;
- Lavielle and Ludena (2000): detecting changes in the spectral density of long-memory processes using Whittle contrast;
- Kokoszka and Leipus (2000): Change-point estimation in ARCH models.
- Berkes *et al.* (2004): Sequential CUSUM procedure for GARCH processes.
- Davis *et al.* (2008): MDE detection for a class of nonlinear time-series.

Remark: all these papers supposed the independence of the process on both sides of the break point.

Definition of the penalized contrast

For fixed $K, \underline{t}, \underline{\theta}$, define the penalized contrast function $\widehat{J}_n$ by:

$$\widehat{J}_n(K, \underline{t}, \underline{\theta}) = \sum_{k=1}^{K} QLIK(\{t_{k-1}+1, \ldots, t_k\}, \theta_k) + \kappa_n\, K,$$

with $\kappa_n \to \infty$ $(n \to \infty)$ and

$$QLIK(\{t_{k-1}+1, \ldots, t_k\}, \theta_k) := \sum_{t=t_{k-1}+1}^{t_k} \big(X_t - \widehat{f}^{\,t}_{\theta_k}\big)' \big(\widehat{M}^{t}_{\theta_k}\big)^{-2} \big(X_t - \widehat{f}^{\,t}_{\theta_k}\big) + 2 \log\big(\widehat{M}^{t}_{\theta_k}\big).$$

QLIK is a Quasi-Likelihood contrast

Definition of the estimator

Define the estimator $(\widehat{K}, \widehat{\underline{t}}_n, \widehat{\underline{\theta}}_n)$ of $(K, \underline{t}, \underline{\theta})$ by:

$$(\widehat{K}, \widehat{\underline{t}}_n, \widehat{\underline{\theta}}_n) := \underset{(K,\underline{t},\underline{\theta}) \in \{0,\ldots,K_{max}\} \times \mathcal{F}_K \times \Theta_K}{\text{Argmin}} \widehat{J}_n(K, \underline{t}, \underline{\theta})$$

and the estimator of breaks is: $\quad \widehat{\underline{\tau}}_n = \dfrac{\widehat{\underline{t}}_n}{n}.$

Remark: if K and $\widehat{\underline{t}}$ are known then $\widehat{\underline{\theta}}_n$ is the QMLE of $\underline{\theta}^*$.

Consistency:

Theorem

If $r \geq 2$, under conditions on Lipshitzian coefficients and κ_n, if $K_{\max} \geq K^*$ then:

$$\big(\widehat{K}_n, \widehat{\underline{\tau}}_n, \widehat{\underline{\theta}}_n\big) \xrightarrow[n\to\infty]{\mathcal{P}} \big(K^*, \underline{\tau}^*, \underline{\theta}^*\big)$$

Idea of the proof: • in each T_k, control a distance between $QLIK_k$ of $(X_t)_{t\in T_k}$ (non-stationary) and $\overline{QLIK}_k$ obtained from a stationary solution of $\mathcal{M}_{T_k}\big(f_{\theta_k}, M_{\theta_k}\big)$.
• use the asymptotic behavior of $\overline{QLIK}_k$ (note $\overline{QLIK}_k \sim C_k\, n\ a.s.$).
• since $\widehat{J}_n(K, \underline{t}, \underline{\theta}) = \sum_{k=1}^{K} \big(QLIK_k - \overline{QLIK}_k\big) + \overline{QLIK}_k + \kappa_n K$

$$\Longrightarrow \ \big|QLIK_k - \overline{QLIK}_k\big| = o(\kappa_n).$$

Rates of convergence:

Theorem

If $r \geq 4$ and under conditions on Lipshitzian coefficients and κ_n, if $K_{\max} \geq K^*$ then

$$\lim_{\delta \to \infty} \lim_{n \to \infty} \mathbb{P}(\|\underline{\widehat{t}}_n - \underline{t}^*\|_m > \delta) = 0.$$

Remark: Same convergence rate as if (X_t) a sequel of independent r.v.

Theorem

If $r \geq 4$, $K_{\max} \geq K^*$, $\kappa_n = \sqrt{n}$ and under conditions on Lipshitzian coefficients, for all $j = 1, \cdots, K^*$,

$$\sqrt{n_j^*}\,\big(\widehat{\theta}_n(\widehat{T}_j) - \theta_j^*\big) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_d\big(0, G(\theta_j^*)\big),$$

Results of simulation for AR(1) process

Model			$\widehat{K}_n = K^*$	$\widehat{K}_n < K^*$	$\widehat{K}_n > K^*$
scenario A_0 ($K^* = 1$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.74	0.00	0.26
		$\kappa_n = \log n$	0.50	0.00	0.50
		$\kappa_n = \sqrt{n}$	0.94	0.00	0.06
		MDL procedure	0.95	0.00	0.05
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.81	0.00	0.20
		$\kappa_n = \log n$	0.43	0.00	0.57
		$\kappa_n = \sqrt{n}$	1.00	0.00	0.00
		MDL procedure	0.97	0.00	0.03
scenario A_1 ($K^* = 2$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.52	0.06	0.42
		$\kappa_n = \log n$	0.40	0.04	0.56
		$\kappa_n = \sqrt{n}$	0.23	0.77	0.00
		MDL procedure	0.44	0.56	0.00
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.78	0.00	0.22
		$\kappa_n = \log n$	0.40	0.00	0.60
		$\kappa_n = \sqrt{n}$	0.38	0.62	0.00
		MDL procedure	0.87	0.13	0.00
scenario A_2 ($K^* = 2$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.48	0.00	0.52
		$\kappa_n = \log n$	0.17	0.00	0.83
		$\kappa_n = \sqrt{n}$	0.29	0.71	0.00
		MDL procedure	0.56	0.44	0.00
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.76	0.00	0.24
		$\kappa_n = \log n$	0.06	0.00	0.94
		$\kappa_n = \sqrt{n}$	0.57	0.43	0.00

Results of simulation for AR(1) process

Model			$\widehat{K}_n = K^*$	$\widehat{K}_n < K^*$	$\widehat{K}_n > K^*$
scenario A_3 ($K^* = 3$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.45	0.32	0.23
		$\kappa_n = \log n$	0.37	0.26	0.37
		$\kappa_n = \sqrt{n}$	0.00	1.00	0.00
		MDL procedure	0.01	0.99	0.00
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.61	0.13	0.26
		$\kappa_n = \log n$	0.39	0.00	0.61
		$\kappa_n = \sqrt{n}$	0.00	1.00	0.00
		MDL procedure	0.20	0.80	0.00
scenario A_4 ($K^* = 3$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.53	0.12	0.35
		$\kappa_n = \log n$	0.28	0.06	0.66
		$\kappa_n = \sqrt{n}$	0.04	0.96	0.00
		MDL procedure	0.09	0.91	0.00
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.75	0.00	0.25
		$\kappa_n = \log n$	0.12	0.00	0.88
		$\kappa_n = \sqrt{n}$	0.06	0.94	0.00
		MDL procedure	0.54	0.46	0.00

Table: Frequencies of the number of breaks estimated after 100 replications for AR(1) process following scenarios $\mathbf{A}_0$-$\mathbf{A}_4$.

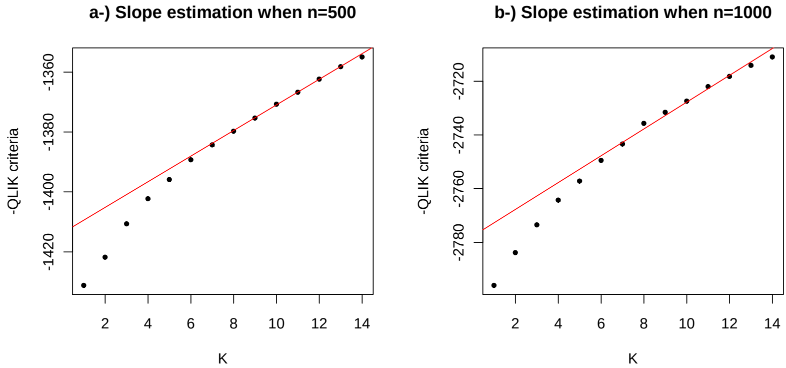
An algorithm for choosing an adaptive κ_n

Problem: if n not very large (typically $n = 1000$), $\widehat{K}_n \neq K^*$ very often with a theoretical a priori choice of κ_n (typically $\kappa_n = \sqrt{n}$).

$\Longrightarrow$ Adaptive choice of κ_n following a model selection procedure (Lebarbier, 2005, Arlot and Massart, 2009):

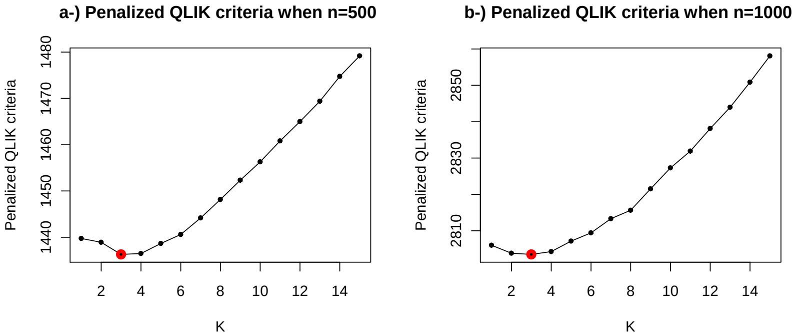
The slope heuristic procedure

Adaptive choice $\widehat{\kappa}_n$



The slope heuristic procedure for a AR(1) process:
$n = 500$, $\widehat{\kappa}_n \simeq 2 * 4.3 \simeq 8.6$ and $n = 1000$, $\widehat{\kappa}_n \simeq 2 * 5 \simeq 10$

Simulations: estimation of $\widehat{K}$



The penalized QLIK criteria for a AR(1) process

Results of simulation for GARCH(1, 1) process

Model			$\widehat{K}_n = K^*$	$\widehat{K}_n < K^*$	$\widehat{K}_n > K^*$
scenario G_0 ($K^* = 1$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.44	0.00	0.56
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.58	0.00	0.42
		MDL procedure	0.51	0.00	0.49
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.60	0.00	0.40
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.75	0.00	0.25
		MDL procedure	0.63	0.00	0.37
scenario G_1 ($K^* = 2$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.42	0.12	0.46
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.52	0.05	0.00
		MDL procedure	0.55	0.35	0.10
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.65	0.00	0.35
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.74	0.10	0.00
		MDL procedure	0.67	0.09	0.24
scenario G_2 ($K^* = 2$)	$n = 500$	$\kappa_n = \hat{\kappa}_n$	0.52	0.20	0.28
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.39	0.44	0.17
		MDL procedure	0.44	0.40	0.16
	$n = 1000$	$\kappa_n = \hat{\kappa}_n$	0.56	0.10	0.34
		$\kappa_n = \log n$	0.00	0.00	1.00
		$\kappa_n = \sqrt{n}$	0.42	0.48	0.10

Let $\mathcal{M}$ a family of parametric causal affine models.

For $m \in \mathcal{M}$ with parameter $\theta^{(m)}$, define the criterion:

$$\widehat{Crit}(m) = -2 \times \widehat{L}_n(\widehat{\theta}^{(m)}) + \kappa_n \times \#(\theta^{(m)}),$$

with $\kappa_n \to \infty$ $(n \to \infty)$ and

$$\widehat{m} = \underset{m \in \mathcal{M}}{\text{Argmin}} \left\{ \widehat{Crit}(m) \right\}$$

Two open problems:

- Asymptotic legitimacy of the BIC criterion ($\kappa_n = \log n$) and its consistency ?
- Consistency of the slope heuristic procedure?

Conclusion

1. Laplacian vs Gaussian:
   - Strong consistency obtained for $r = 1$ (Laplacian) vs $r = 2$ (Gaussian);
   - Asympto. normality obtained for $r = 2$ (Laplacian) vs $r = 4$ (Gaussian);
   - Simulations show more accurate and robust results for Laplacian;
   - Confidence intervals require the estimation of $g(0)$ and σ_ξ for Laplacian.
2. Change-point detection:
   - Convergence rates almost the same than for independent r.v.!
   - Slope heuristic provides best numerical results
   - Online detection also possible
3. Model selection: on going work!

References

- Arlot and Massart (2009). Data-driven Calibration of Penalties for Least-Squares Regression. *Journal of Mach. Learn.*
- Bardet and Boularouk (2017). Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes. *Electr. Journal. Statist.*
- Bardet and Kengne (2014). Monitoring procedure for parameter change in causal time series. *Journal of Multi. Analysis*
- Bardet, Kengne and Wintenberger (2012). Detecting multiple change-points in general causal time series using penalized quasi-likelihood. *Elect. Journal Statis.*
- Bardet and Wintenberger (2009). Asymptotic normality of the Quasi-Maximum Likelihood Estimator for multidimensional causal processes. *Ann. Statist.*
- Doukhan and Wintenberger (2008). Weakly dependent chains with infinite memory. *Stoch. Proc. and their Appli.*
- Francq, Lepage and Zakoian (2011) Two-stage non Gaussian QML estimation of GARCH models and testing the efficiency of the Gaussian QMLE. *J. Econometrics*.
- Straumann (2005). *Estimation in conditionally heteroscedastic time series models*. Springer, New York.