General estimation results for tdVARMA array models

Guy Mélard [1] *

[1]ECARES, Université libre de Bruxelles, and ITSE sprl, Belgium
(gmelard@ulb.ac.be)

November 23-24, 2017
Lm$^{\mathbf{B}}$ laboratoire de **mathématiques** de **Besançon**
Time Series Analysis

*Partly joint with Rajae Azrak [A] & with Abdelkamel Alj [Aℓ]

Outline

1. Introduction
   - Class of processes = tdVARMA$^{(n)}$
   - Illustrations
   - Estimation

Outline

1. Introduction
   - Class of processes = tdVARMA$^{(n)}$
   - Illustrations
   - Estimation
2. Asymptotic results
   - Main theorem
   - A fundamental theorem for the asymptotic theory
   - A theorem for reducing the assumption on moments
   - Convergence for the two covariance matrices V and W

Outline

1. Introduction
   - Class of processes = tdVARMA$^{(n)}$
   - Illustrations
   - Estimation
2. Asymptotic results
   - Main theorem
   - A fundamental theorem for the asymptotic theory
   - A theorem for reducing the assumption on moments
   - Convergence for the two covariance matrices V and W
3. Application to tdVARMA$^{(n)}$ models
   - Proof of main theorem for tdVARMA$^{(n)}$ models
   - Numerical and simulation strategies
   - Theoretical illustrations: tdVAR(1) and tdVMA(1)
   - Simulations

Outline

Time-dependent VARMA processes [Aℓ2012]

Definition of a m-dimensional tdVARMA$^{(n)}(p, q)$ (time dependent VARMA process) = triangular array of random vectors (r.v.) $(x_t^{(n)}, t \in \mathbb{N})$, $n =$ series length, solution of
$x_t^{(n)} = \sum_{k=1}^{p} A_{tk}^{(n)} x_{t-k}^{(n)} + g_t^{(n)} \epsilon_t + \sum_{k=1}^{q} B_{tk}^{(n)} g_{t-k}^{(n)} \epsilon_{t-k}$, where

- $\{\epsilon_t, t \in \mathbb{N}\}$: independent m-dimensional r.v., with 0 mean and covariance matrix $\Sigma > 0$ (nuisance parameter);
- the coefficients $A_{tk}^{(n)}$, $B_{tk}^{(n)}$, and $g_t^{(n)}$ are $m \times m$ matrices;
- their elements are deterministic functions of t (possibly n);
- $\Sigma_t^{(n)} = g_t^{(n)} \Sigma g_t^{(n)T}$: the error covariance matrix;
- Initial values $x_t^{(n)}, \epsilon_t^{(n)}, t < 1$, supposed to be equal to 0 $^{(*)}$.

$^{(*)}$ Only for the asymptotic theory, not in practice

tdVARMA$^{(n)}$ parametric model

- The $r \times 1$ vector θ contains all the parameters of interest to be estimated, those in the $A_{tk}^{(n)}(\theta)$, $B_{tk}^{(n)}(\theta)$, and $g_t^{(n)}(\theta)$ (not Σ);
- Their elements are deterministic functions of these parameters, in addition to t (and possibly n);
- In the simple VARMA case, the elements are the parameters and $g_t^{(n)}(\theta)$ is absent;
- True value $\theta = \theta^0$, so $A_{tk}^{(n)}(\theta^0) = A_{tk}^{(n)}$, $B_{tk}^{(n)}(\theta^0) = B_{tk}^{(n)}$, and $g_t^{(n)}(\theta^0) = g_t^{(n)}$;
- Residuals: $e_t^{(n)}(\theta) = x_t^{(n)} - \sum_{k=1}^{p} A_{tk}^{(n)}(\theta) x_{t-k}^{(n)} - \sum_{k=1}^{q} B_{tk}^{(n)}(\theta) e_{t-k}^{(n)}(\theta)$;
- Hence $e_t^{(n)}(\theta^0) = g_t^{(n)}(\theta^0)\epsilon_t = g_t^{(n)}\epsilon_t$.

tdARMA and tdVARMA evolution

- Start in 1973 ("FARIMAG" models) in M's thesis
- Starting general case in 1977
- A first talk in 1981 with computational results (WLS)
- Exact maximum likelihood (EML) [M1982]
- Azrak's thesis in 1991-1996: AR case with mixing condition
- Submission for ARMA in 1998 + EML algorithm [AM1998]
- Adding explicit dependency on n in 1999-2002
- Paper in SISP [AM2006] (without mixing)
- tdVARMA models in Aℓj's thesis in 2008-2012
- Array CLT [AℓAM2014] and tdVARMA EML [AℓJM2016]
- Paper in SJS without $^{(n)}$ with Ley [AℓALM2017]
- Improvements for the $^{(n)}$ case 2011-2017 [AM20??a]
- Comparison with other approaches [AM20??b]

Example: univariate tdAR$^{(n)}$(1) case

Example: $\theta = (A', A'', \eta)^T$ (with appropriate conditions)

$$x_t^{(n)} = A_t^{(n)}(\theta) x_{t-1}^{(n)} + g_t^{(n)}(\theta)\epsilon_t$$

$$A_t^{(n)}(\theta) = A' + \frac{1}{n-1}\left(t - \frac{n+1}{2}\right) A'',$$

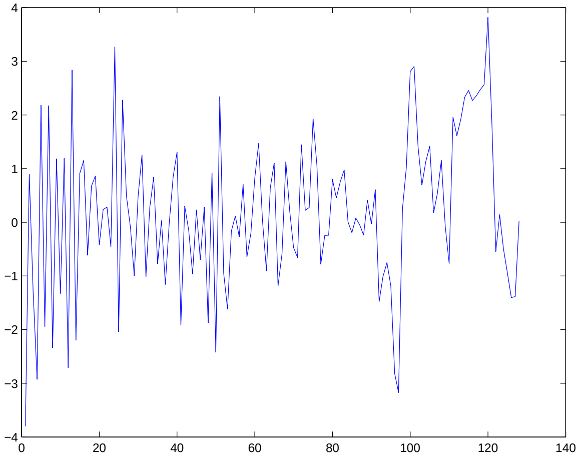
$$g_t^{(n)}(\theta) = \exp\left\{\frac{\eta}{n-1}\left(t - \frac{n+1}{2}\right)\right\}$$

Notes.

- Term $(n+1)/2$ to achieve orthogonality
- Factor $1/(n-1)$ or $1/n$ just for the asymptotics (to restrain the coefficient in a finite interval)
- Not a random sequence x_t but well random array $x_t^{(n)}$
- Similar parametrization for MA coefficients
- Alternative for $g_t(\theta)$: periodic 2-state function $(g, 1/g)$

ULB

Figure: Artificial series ($n = 128$) produced using an tdAR$^{(n)}$(1) process with $A'^0 = 0.15$, $A''^0 = 0.015$, $g_t(\theta^0) = \{6 * 2, 6 * 0.5, ...\}$



Practical time series

- In [AM2006], series from Box & Jenkins (1970), [BJRL2015]
  - Series A ($n = 197$): tdARIMA$^{(n)}$(0,1,1)
  - Series B ($n = 395$): ARIMA(0,1,1) with td$^{(n)}$ error variance
  - Series G airline series ($n = 144$): $\nabla\nabla_{12} \log x_t = (1 - \theta L)(1 - \Theta L^{12})e_t$, or 'airline model', where L is the lag operator.
- In [AℓJM2016]: monthly log returns of IBM stock prices and S&P 500 index (1926-1999) by tdVAR$^{(n)}$(1) and tdVMA$^{(n)}$(3) models
- In [AM20??c] we add:
  - dataset of indices for monthly added value of the Belgian industrial production by branches (26) of activity (1985-1994)
  - dataset of 320 U.S. industrial production time series (January 1986- present)

Main results: Estimation method

- Quasi-maximum likelihood estimator (with $[\alpha_t^{(n)}(\theta)]$):

$$\hat{\theta}^{(n)} = \text{argmin}_{\theta \in \mathbb{R}^r} \sum_{t=1}^{n} \left[\log |\Sigma_t^{(n)}(\theta)| + e_t^{(n)T}(\theta) \Sigma_t^{(n)-1}(\theta) e_t^{(n)}(\theta) \right].$$

- The quasi log-likelihood is computed by an algorithm due to [AℓJM2016], inspired by Jónasson & Ferrando (2008)
- In the univariate tdARMA case: [M1982] & [AM1998]
- The objective function is minimized by numerical optimization
- By-product of the optimization procedure: standard errors obtained by inverting the estimated information matrix (Hessian)

Preliminaries: AR and MA representations

(AR representation) $x_t^{(n)} = e_t^{(n)}(\theta) + \sum_{k=1}^{t-1} \pi_{tk}^{(n)}(\theta) x_{t-k}^{(n)}$ (1)

(MA representation) $x_t^{(n)} = e_t^{(n)}(\theta) + \sum_{k=1}^{t-1} \psi_{tk}^{(n)}(\theta) e_{t-k}^{(n)}(\theta)$ (2)

where the coefficients $\pi_{tk}^{(n)}(\theta)$ and $\psi_{tk}^{(n)}(\theta)$ are obtained by double recurrence (w.r.t. k and t) (see [M1985])

To compute derivatives of $e_t^{(n)}(\theta)$ w.r.t. θ_i we start from (1) and then replace $x_{t-k}^{(n)}$ using (2) for $\theta = \theta^0$:

$$\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_i} = \sum_{k=1}^{t-1} \psi_{tik}^{(n)}(\theta, \theta^0) e_{t-k}^{(n)}(\theta^0), \qquad (3)$$

$$\text{with } \psi_{tik}^{(n)}(\theta, \theta^0) = \sum_{u=1}^{k} \frac{\partial \pi_{tu}^{(n)}(\theta)}{\partial \theta_i} \psi_{t-u,k-u}^{(n)}(\theta^0)$$

Let $\psi_{tik}^{(n)} = \psi_{tik}^{(n)}(\theta^0, \theta^0)$, κ_t = 4th order moment of ϵ_t, $\kappa_t = E((\epsilon_t \epsilon_t^T) \otimes (\epsilon_t \epsilon_t^T))$, and denote the Frobenius norm $\|.\|_F$

Sketch of the assumptions

i $A_{ti}^{(n)}(\theta)$, $B_{tj}^{(n)}(\theta)$ and $g_t^{(n)}(\theta)$ are of class C^3 w.r.t. $\theta \in$ compact set $\Theta \supset \{\theta^0\}$;

ii Upper bounds like $\sum_{k=\nu}^{t-1} \|\psi_{tik}^{(n)}\|_F^2 < N_1 P(\nu)\Phi^{\nu-1}$, $\sum_{k=\nu}^{t-1} \|\psi_{tik}^{(n)}\|_F^4 < N_2 P(\nu)\Phi^{\nu-1}$, ... with positive constants N_1, N_2, $0 < \Phi < 1$, a polynomial $P(\nu)$ (only needed for VARMA), and $\nu = 1, ..., t-1$;

iii Existence of moments of order $4+2\delta$ for ϵ_t's, $\delta > 0$, + bounds on the Frobenius norm of κ_t, & of $\Sigma_t^{(n)}$ and $\Sigma_t^{(n)-1}$, and their derivatives with respect to θ at θ^0;

iv Existence of a strictly positive definite matrix $V = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} V_t^{(n)}$, where $V_{t,ij}^{(n)}$, $i,j = 1, ..., m$, is given by

$$E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}\right)+\frac{1}{2}\,\mathrm{tr}\left[\Sigma_t^{(n)-1}(\theta)\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\theta_j}\right]_{\theta=\theta^0};$$

v A similar existence condition for a positive definite matrix W (outer product of gradient), to be defined, which includes 4th order moment κ_t;

vi That for $i = 1, ..., m$

$$\frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d}\sum_{k=1}^{t-1}\|g_{t-k}^{(n)}\|_F^2\,\|\psi_{tik}^{(n)}\|_F\,\|\psi_{t+d,i,k+d}^{(n)}\|_F = O\left(\frac{1}{n}\right)$$

Main theorem for tdVARMA$^{(n)}$ models

Theorem (AℓAM20??)

Under the (full) assumptions,

- *there exists a sequence of estimators* $\widehat{\theta}^{(n)}$ *such that* $\text{plim}\,\widehat{\theta}^{(n)} = \theta^0$ *when* $n \to \infty$,
- *furthermore*

$$n^{1/2}(\widehat{\theta}^{(n)} - \theta^0) \xrightarrow{L} \mathcal{N}(0, V^{-1}WV^{-1}) \text{ when } n \to \infty.$$

Remarks.

1. For a Gaussian process: $V = W$; otherwise the sandwich formula;
2. In the univariate ARMA case, see [AM2006];
3. If no $^{(n)}$, plim replaced by almost sure convergence, see [AℓALM2017];
4. More on the proof later but parallel to [AℓALM2017] and its Technical Appendix (TA) is used.

New theoretical results [AM20??a]

- No problem if the coefficients don't depend on n, see [AM2006] (tdARMA), [AℓALM2017] (tdVARMA)
- A fundamental theorem for the asymptotic theory in the array context, for the general case
- A theorem for reducing the assumption on moments from 8 to $4 + 2\delta$, $\delta > 0$
- Two theorems to establish convergence for the two covariance matrices V and W involved in the sandwich formula
- Plus Th2.4 = Lemma 1' of [AM2016] = weak version of a result by Hamdoune (1995) - not detailed here

A fundamental theorem for the asymptotic theory

- Purpose: provide an alternative to Klimko-Nelson (1978) theorems for the case where the coefficients depend on n
- Indeed, almost sure convergence is to be replaced by convergence in probability
- We give a direct proof of Theorem 1' in [AM2006]
- This is also proved for vectors, not only scalar processes
- Even with a slight improvement by using an upper bound on $E_{\theta^0}(|\partial\alpha_t^{(n)}(\theta)/\partial\theta_i|^{2+\delta})$, where $\delta > 0$, θ^0 is the true value of the parameter θ and $\alpha_t^{(n)}(\theta)$ it the t-th term of the Gaussian log-likelihood
- This instead of an upper bound on a 4-th power
- The consistency theorem is as follows and there is a further theorem on asymptotic normality

Theorem (AM20??a Th2.1)

Improvement on [AM2006,Theorem 1']

Suppose there exist $C_1 > 0$, $C_2 > 0$, $\delta > 0$, such that for all $t = 1, ..., n$, and uniformly in n:

$$\mathbf{H}_{1.1} \quad E_{\theta^0}\left(\left|\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i}\right|^{2+\delta}\right) \leq C_1, i = 1, ..., m;$$

$$\mathbf{H}_{1.2} \quad E_{\theta^0}\left(\left|\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j} - E_\theta\left(\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}\middle| F_{t-1}^{(n)}\right)\right|^2\right) \leq C_2, i,j = 1, ..., m.$$

Suppose further that

$$\mathbf{H}_{1.3} \quad \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{t=1}^n E_{\theta^0}\left\{\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}\middle| F_{t-1}^{(n)}\right\} = V_{ij} \text{ for } \quad i,j = 1, ..., m, \text{ where}$$

$V = (V_{ij})_{1\leq i,j\leq m}$ is a strictly positive definite matrix of constants;

$$\mathbf{H}_{1.4} \quad \operatorname*{plim}_{n\to\infty} \sup_{\Delta\downarrow 0}(n\Delta)^{-1}\left|\sum_{t=1}^n\left(\left\{\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}\right\}_{\theta=\theta_{ij}^*} - \left\{\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}\right\}_{\theta=\theta^0}\right)\right| <$$

∞, for $i,j = 1, ..., m$, where θ_{ij}^* is a point of the straight line joining θ^0 to every θ, such that $\|\theta - \theta^0\| < \Delta$, $0 < \Delta$, where $\|.\|$ is the Euclidean norm.

Then there exists a sequence of estimators $\widehat{\theta}^{(n)}$ such that $\operatorname{plim} \widehat{\theta}^{(n)} = \theta^0$ *when $n \to \infty$.*

Theorem (AM20??a Th2.2)

Improvement on Theorem 1'. of AM2006

If the assumptions $\mathbf{H}_{1.1} - \mathbf{H}_{1.4}$ *of Theorem 1' are satisfied, as well as* $\mathbf{H}_{1.5}$ *and* $\mathbf{H}_{1.6}$,

$\mathbf{H}_{1.5}$ *for* $i, j = 1, ..., m$

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \left\{ E_{\theta^0}\left(\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_j} \middle| F_{t-1} \right) - E_{\theta^0}\left(\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_j} \right) \right\} = 0,$$

$\mathbf{H}_{1.6}$ *there exists a positive definite matrix* $W = (W_{ij})_{1\leq i,j\leq m}$ *defined by*

$$W_{ij} = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E_{\theta^0}\left(\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_j} \right),$$

then

$$n^{1/2}(\widehat{\theta}^{(n)} - \theta^0) \xrightarrow{L} \mathcal{N}(0, V^{-1}WV^{-1}) \text{ when } n \to \infty.$$

N.B. W is defined here

Sketch of proof of Th2.1 and Th2.2.

It is adapted from the Lehmann and Casella (1998, Section 6.5) proof in the i.i.d. case
+ weak law of large numbers for martingale arrays
+ central limit theorem theorem for martingale arrays with a Lyapunov condition [AℓAM2014]
+ Cramér-Wold device □

A theorem for reducing the assumption on moments

- In [AM2006] we have assumed existence of 8-th moments for the errors
- We have kept that assumption in [AℓALM2017]
- In [AℓALM2017] we make use of a Technical Appendix Lemma 4.11 where that assumption is essential
- However, we are now able to reduce the moment assumption from 8 to $4 + 2\delta$, $\delta > 0$
- This is expressed here in a vector context, e.g. a matrix $\Sigma_t^{(n)}$ instead of $\sigma_t^{(n)2}$

Theorem (AM20??a Th2.3)

Assume that $\alpha_t^{(n)}(\theta)$ *has the form*
$\alpha_t^{(n)}(\theta) = \{x_t^{(n)} - E_\theta(x^{(n)}|F_{t-1}^{(n)})\}^T \Sigma_t^{(n)-1}(\theta)\{x_t^{(n)} - E_\theta(x^{(n)}|F_{t-1}^{(n)})\}$, *for some invertible matrix* $\Sigma_t^{(n)}(\theta)$. *Denote* $e_t^{(n)} = x_t^{(n)} - E_\theta(x^{(n)}|F_{t-1}^{(n)})$ *and* $\|.\|_F$, *the Frobenius norm of a matrix. Suppose that for some* $\delta > 0$ *we have for all* t *and* n

$$\left\| \left. \frac{\partial \Sigma_t^{(n)-1}(\theta)}{\partial \theta_i} \right|_{\theta=\theta^0} \right\|_F^2 \leq K_4, \qquad \left\| \Sigma_t^{(n)-1}(\theta^0) \right\|_F^2 \leq m_2,$$

$$E_{\theta^0}\left(\left| e_t^{(n)T}(\theta) e_t^{(n)}(\theta) \right|^{2+\delta} \right) \leq P_1, \qquad E_{\theta^0}\left(\left| \frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_i} \frac{\partial e_t^{(n)}(\theta)}{\partial \theta_i} \right|^{1+\delta/2} \right) \leq P_2,$$

$i = 1, ..., m$, *for some constants* K_4, m_2, P_1, *and* P_2, *and that* $e_t^{(n)}(\theta^0)$ *and* $\partial e_t^{(n)}(\theta)/\partial \theta_i|_{\theta=\theta^0}$ *are independent. Then, the assumption* $\mathbf{H}_{1.1}$ *is satisfied for that* δ, *which means that there exists a positive constant* C_1 *such that for all* t *and all* n, *and* $i = 1, ..., m$

$$E_{\theta^0}\left\{ \left| \frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i} \right|^{2+\delta} \right\} \leq C_1.$$

Convergence for the two covariance matrices V & W

- We can write:

$$V = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} V_t^{(n)}, \quad W = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} W_t^{(n)}$$

- We can compute numerically $V_t^{(n)}$ and $W_t^{(n)}$ (depending on κ_t) for simple models
- Proving existence of the limits V and W is not that easy
- One way is to use the Cesàro theorem (that if a sequence u_n converges to U, then the Cesàro means $U_n = \frac{1}{n}\sum_{i=1}^{n} u_i$ converges to U)
- This is not always possible, even in some simple examples of [AM2006]
- The following two theorems can thus help us

Theorem (AM20??a Th2.5)

Let $\{u_t^{(n)}, t = 1, ..., n\}$ and $\{v_t^{(n)}, t = 1, ..., n\}$ be two triangular arrays of real numbers such that

- $(1/n)\sum_{t=1}^{n} v_t^{(n)}$ *absolutely converges when $n \to \infty$ and* $\lim_{n\to\infty}(1/n)\sum_{t=1}^{n} v_t^{(n)} = L$, *and that*
- $\{u_t^{(n)}\} \to C > 0$ *when $t \to \infty$, hence $n \to \infty$.*

Then $(1/n)\sum_{t=1}^{n} u_t^{(n)} v_t^{(n)}$ converges when $n \to \infty$ and its limit is LC.

Example (AM2006, Example 3)

tdAR$^{(n)}$(1) model defined by $x_t^{(n)} = A_t^{(n)}(\theta) x_{t-1}^{(n)} + g_t^{(n)} \epsilon_t$, with independent ϵ_t's with 0 mean and finite variance σ^2, and $g_t^{(n)} > 0$, assumed not to depend on the parameters θ, for simplicity. We have to show existence of

$$V_t = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} V_t^{(n)} \text{ where } V_t^{(n)} = \left\{ \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_j} \right\}_{\theta=\theta^0} E(x_{t-1}^{(n)2}),$$

$i, j = 1, ..., m$, see [AℓALM2017]. Assume $A_t^{(n)}(\theta^0) = A^0$ with $|A^0| < 1$. Then (see [AM2006]) $u_t^n = E(x_{t-1}^{(n)2})$ is convergent with limit say $C > 0$. This is true in particular if $g_t^{(n)} = \exp\{\frac{t-(n+1)/2}{n-1}\}$. Therefore, if $\frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_j} \right\}_{\theta=\theta^0}$ is absolutely convergent, $i, j = 1, ..., m$, and converges to a limit L_{ij}, then V does exists by application of [AM20??a Theorem 2.5], and its element V_{ij} is equal to CL_{ij}.
This is the case, in particular, if (see [AM2006]), $A_t^{(n)}(\theta) = \theta_1 + \theta_2 \frac{t-(n+1)/2}{n-1}$ at least when $\theta_1^0 = A^0$ and $\theta_2^0 = 0$.

Theorem (AM20??a Th2.6)

Consider $\lim_{n\to\infty}(1/n)\sum_{t=1}^{n} v_t^{(n)}$. *Assume that there exists a Riemann-integrable function* $V(x)$ *defined on* $[0,1]$ *such that* $V(t/n) = v_t^{(n)}$. *Then* $\lim_{n\to\infty}(1/n)\sum_{t=1}^{n} v_t^{(n)} = \int_0^1 V(x)dx$

Example (AM2006, Examples 2, 3, 4)

Let $g_{t,\theta}^{(n)} = \exp\{\theta(t-(n+1)/2)/(n-1)\}$ for $\theta \geq 0$, $t = 1, ..., n$ (see [AM2006, Examples 2, 3 and 4]). Suppose that, for $\xi \geq 0$ and $\eta > 0$: $v_t^{(n)} = g_{t,\xi}^{(n)}/(1+g_{t,\eta}^{(n)})^2$. Using a variation of Th2.6 (where $(t-(n+1)/2)/(n-1)$ is replaced by x), we have: $\lim_{n\to\infty}(1/n)\sum_{t=1}^{n} v_t^{(n)} = \lim_{n\to\infty} \frac{n-1}{n} \int_{-0.5}^{0.5} \frac{e^{\xi x}}{(1+e^{\eta x})^2} dx$,and a simple primitive for $\xi = 0$ is

$$x + \frac{1}{\eta}\left\{\frac{1}{1+e^{\eta x}} - \log(1+e^{\eta x})\right\} + C.$$

For $\xi > 0$, it is based on the hypergeometric function, see Abramowitz and Stegun (1965, Chapter 15) or Erdélyi (1953)

Proof of main theorem, tdVARMA$^{(n)}$ [AℓAM20??]

- We use most lemmas in [AℓALM2017]'s Technical Appendix = TA, easily generalized in an array context, except TA Lemma 4.11 replaced by [AM20??a, Th2.3]
- Like in [AℓALM2017], we have to use the (full) assumptions to check the conditions of [AM20??a, Th2.1 and 2.2] for all t and uniformly in n (except $\mathbf{H}_{1.6}$ which is assumed)
- $\mathbf{H}_{1.1}$ = bound of $E_{\theta^0}(|\partial\alpha_t^{(n)}(\theta)/\partial\theta_i|^{2+\delta})$: consequence of [AM20??a, Th2.3]
- $\mathbf{H}_{1.2}$ like in [AℓALM2017]
- $\mathbf{H}_{1.3}$ (existence of V as plim): based on [AM20??a, Th2.4]
- $\mathbf{H}_{1.4}$ (3rd order terms): weak law of large numbers for martingale arrays + weak law of large numbers for L_2-mixingale arrays of Meng & Lin (2009)
- $\mathbf{H}_{1.5}$ (expectation vs conditional exp.): [AM20??a, Th2.4]

Numerical and simulation strategies (1)

- Like Jónasson (2008) VARMA estimation Matlab program, AJM in [AℓJM] makes use of Optimization Toolbox fminunc
- Penalties are applied for each evaluation of the log-likelihood where the conditions are not fulfilled
- AJM used by [AℓALM2017], only in the Gaussian case
- It allows computation of the Hessian V at the optimum, not the outer product of gradient W
- Done by numerical divided differences $\Rightarrow$ limited accuracy
- AJM2 is a new version in development, aimed at, in particular but not only, adding the evaluation of W, in addition to the Hessian V
- Hence standard errors based on either $\frac{1}{n}V^{-1}$ or sandwich formula $\frac{1}{n}V^{-1}WV^{-1}$

Numerical and simulation strategies (2)

New program AJM2 aimed at

- adding the evaluation of W, in addition to the Hessian V, hence also the sandwich formula
- W evaluated using estimates of $\partial\alpha_t^{(n)}(\theta)/\partial\theta_i$, for each t, also with limited accuracy
- adding (improved) DERIVEST to provide better accuracy
- simulation procedure changed to permit verification (now inverse of residual computation), plus support of other laws than normal
  - Laplace: normal deviate $\times$ Exp(1) deviate
  - Student with ν df: normal deviate / deviate of sqrt of χ^2_ν/ν

Note. DERIVEST: Found in Aït-Sahalia (2015) Matlab library implementing closed-form MLE for diffusions, due to D'Errico: `hessian` for V, `jacobianest` for W but `hessian` & `jacobianest` had to be modified (either bug, or code unsuited in the case of penalties)

Theoretical illustrations

- Everything is based on MA representations
- Bivariate ($r = 2$) tdVAR(1) and tdVMA(1) are more or less easily handled with linear or exponential functions of time for the coefficients, and $g_t^{(n)}$ exponential
- Cascade of specifications in order to illustrate the assumptions using analytical expressions with a small number of parameters
- Here we go straight to $m = 3$ (1 parameter of each type)
- First compute $\psi_{tik}^{(n)}$ in order to check the bound $\sum_{k=\nu}^{t-1} \|\psi_{tik}^{(n)}\|_F^2 < N_1 P(\nu)\Phi^{\nu-1}$, with $\Phi < 1$ and $P(\nu)$, a polynomial
- Then, investigate the existence of V (2 types of terms)
- Finally, prove the $O(\frac{1}{n})$ triple sum property

Application to a tdVAR$^{(n)}$(1) model (1)

- Bivariate tdVAR$^{(n)}$(1) model [AℓAM20??]:

$$x_t^{(n)} = \underbrace{\begin{pmatrix} A'_{11} & A'^{0}_{12} \\ 0 & A''_{22} L(t,n) \end{pmatrix}}_{= A_t^{(n)}(\theta)} x_{t-1}^{(n)} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & e^{\eta_{22} L(t,n)} \end{pmatrix}}_{= g_t^{(n)}(\theta)} \epsilon_t ,$$

- where $L(t,n) = \frac{t - \frac{n+1}{2}}{n-1}$, and A'^{0}_{12} is fixed
- with conditions (to be given) on the true values A'^{0}_{11}, A''^{0}_{22} and η^{0}_{22} of $\theta = (A'_{11}, A''_{22}, \eta_{22})$
- $\Sigma = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}$
- $E(e_t^{(n)}(\theta^0) e_t^{(n)T}(\theta^0)) = g_t^{(n)} \Sigma g_t^{(n)T} =_{\text{def}} \Sigma_t$

Application to a tdVAR$^{(n)}$(1) model (2)

Checking (ii). tdAR coefficient:

$$A_t^{(n)}(\theta) = \begin{pmatrix} A'_{11} & A'^{0}_{12} \\ 0 & A''_{22} L(t,n) \end{pmatrix}.$$

Let us define $A_t^{(n)(k-1)} = \prod_{l=1}^{k-1} A_{t-l}^{(n)}$, $k > 1$, and $A_t^{(n)(0)} = I_r$.
It can be shown that $\|\psi_{t2k}^{(n)}\|_F^2 = [L(t,n) A_{t,22}^{(n)(k-1)}]^2$.
But $L(t,n) \leq \frac{1}{2}$ and $A_{t,22}^{(n)(k-1)} = (A''^{0}_{22})^{k-1} \prod_{l=1}^{k-1} L(t-l,n)$.
Assume $|A'^{0}_{11}| < 1$, $|A''^{0}_{22}| < 2$. Let $\sqrt{\Phi} = \max\{|A'^{0}_{11}|, \frac{1}{2}|A''^{0}_{22}|\} < 1$.
Hence $\sum_{k=\nu}^{t-1} \|\psi_{t2k}^{(n)}\|_F^2 \leq \sum_{k=\nu}^{t-1} \Phi^{k-1} < N_1 \Phi^{\nu-1}$, with $N_1 = \frac{1}{1-\Phi}$.
More delicate for $\|\psi_{t1k}^{(n)}\|_F^2 = [(A_{t,11}^{(n)(k-1)})^2 + (A_{t,12}^{(n)(k-1)})^2]$.
$\sum_{k=\nu}^{t-1} \|\psi_{t1k}^{(n)}\|_F^2 < N'_1 \Phi^{\nu-1} P_2(\nu)$, $P_2(\nu)$: polynomial of degree 2.

Application to a tdVAR$^{(n)}$(1) model (3)

Checking (iv). Now existence of $V_{ij} = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} V_{t,ij}^{(n)}$
2 types of terms. For $i, j = 1, 2$, only term 1 equal to

$$V_{t,ij}^{(n)} = \left\{ \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_i} \frac{\partial A_t^{(n)}(\theta)}{\partial \theta_j} \right\}_{\theta=\theta^0} \Sigma_t^{(n)-1} E\left(x_{t-1}^{(n)} x_{t-1}^{(n)T}\right). \qquad (1)$$

Difficult case $i = j = 2$. Suppose $\eta_{22}^0 > 0$.
Product of factors 1 & 2 of (1) has element (2,2)
$v_t^{(n)} = L^2(t,n)\exp(-2\eta_{22}^0 L(t,n))$ such that $(1/n)\sum_{t=1}^{n} v_t^{(n)}$ converges to a limit when $n \to \infty$.
Last factor of (1) $u_t^{(n)}$ can be shown to converge to a limit > 0.
Hence application of Th2.5 of [AM20??a] implies existence of V_{22}. Similar for other elements.

Application to a tdVAR$^{(n)}$(1) model (4)

$$V_{t,33}^{(n)} = \frac{1}{2}\,\mathrm{tr}\left[\Sigma_t^{(n)-1}(\theta)\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\eta_{22}}\Sigma_t^{(n)-1}(\theta)\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\eta_{22}}\right]_{\eta_{22}=\eta_{22}^0}.$$

But

$$\Sigma_t^{(n)-1}(\theta)\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\eta_{22}} = \begin{pmatrix} 0 & 0 \\ 0 & 2L(t,n) \end{pmatrix}.$$

Hence

$$V_{t,33}^{(n)} = 2\frac{1}{(n-1)^2}\left(t - \frac{n+1}{2}\right)^2,$$

and, using the variance of a discrete uniform distribution on {1, 2, ..., n} we obtain $V_{33} = \lim_{n\to\infty}(n+1)/(6(n-1)) = 1/6$.

Application to a tdVAR$^{(n)}$(1) model (5)

Checking (vi). Finally, there remains to check that for $i, j = 1, 2$

$$\frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d}\sum_{k=1}^{t-1}\left\|g_{t-k}^{(n)}\right\|_F^2\left\|\psi_{tik}^{(n)}\right\|_F\left\|\psi_{t+d,j,k+d}^{(n)}\right\|_F = O\left(\frac{1}{n}\right).$$

Take $i = j = 2$. First $\|g_{t-k}^{(n)}\|_F^2 = 1 + e^{2\eta_{22}^0 L(t-k,n)} < 1 + e^{\eta_{22}^0}$.
Upper bound of $\|\psi_{t2k}^{(n)}\|_F$: Φ^{k-1}. Thus for $\|\psi_{t+d,2,k+d}^{(n)}\|_F$: Φ^{k+d-1}.
The sum for $k = 1, ..., t-1$ of the product $\Phi^{k-1}\Phi^{k+d-1}$ is bounded by Φ^{d-2} times a constant $1/(1-\Phi^2)$.
By exchanging the two outside sums, we have to find an upper bound of $\sum_{t=1}^{n-1}\sum_{d=1}^{n-t}\Phi^{d-1}$: $\Phi^{-1}\times$ the sum for $t = 1, ..., n-1$ of a constant $1/(1-\Phi)$. Dividing by n^2 we have well $O(1/n)$.
The case $i = j = 1$ is more complex.

Application to a tdVMA$^{(n)}$(1) model (1)

- Bivariate tdVMA$^{(n)}$(1) model [AℓAM20??]:

$$x_t^{(n)} = \underbrace{\begin{pmatrix} B'_{11} & 0 \\ 0 & B'^0_{22} e^{B''_{22} L(t,n)} \end{pmatrix}}_{= B_t^{(n)}(\theta)} e_{t-1}^{(n)} + \underbrace{\begin{pmatrix} 1 & \alpha \\ \beta & e^{\eta_{22} L(t,n)} \end{pmatrix}}_{= g_t^{(n)}(\theta)} \epsilon_t,$$

- where $L(t,n) = \frac{t - \frac{n+1}{2}}{n-1}$, and B'^0_{22} is fixed
- with conditions (to be given) on the true values B'^0_{11}, B''^0_{22} and η^0_{22} of $\theta = (B'_{11}, B''_{22}, \eta_{22})$
- $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$
- $E(e_t^{(n)}(\theta^0) e_t^{(n)T}(\theta^0)) = g_t^{(n)} \Sigma g_t^{(n)T} =_{\text{def}} \Sigma_t$

Application to a tdVMA$^{(n)}$(1) model (2)

Checking (ii). Exponential functions of time for the coefficients are much easier for analytical results on VMA processes.

$(\psi_{t1k}^{(n)})_{11} = (-1)^k (B_{11}'^0)^{k-1}$,

$(\psi_{t2k}^{(n)})_{22} = (-1)^k L(t-k+1,n)(B_{22}'^0)^k e^{\left(B_{22}''^0 \sum_{\ell=0}^{k-1} L(t-\ell,n)\right)}$,

and all other elements are zero.

We assume that the true value of θ satisfies $|B_{11}'^0| < 1$, and $|B_{22}'^0| e^{B_{22}''^0/2} < 1$.

We denote $\sqrt{\Phi} = \max\{|B_{11}'^0|, |B_{22}'^0| e^{B_{22}''^0/2}\} < 1$.

Since $|L(t-\ell,n)| \leq \frac{1}{2}$, $\|\psi_{tik}^{(n)}\|_F^2 < \Phi^k$, $i = 1, 2$.

Note.

In practice, we don't assume zero initial values for the process, but well it is invertible before time 1 $\Rightarrow |B_{22}'^0| e^{B_{22}''^0 L(0,n)} < 1$.

Application to a tdVMA$^{(n)}$(1) model (3)

Checking (iv). Existence of V.
For element $(3, 3)$, slightly more complex here because $g_t^{(n)}$ is non diagonal. We need to use Th2.6 of [AM20??a] and evaluate an integral in order to obtain the value of V_{33}.
Also for that reason, the treatment of V_{ij}, $i, j = 1, 2$ is more complex but Th2.5 of [AM20??a] can again be applied.

Checking (vi). Triple sum is $O\left(\frac{1}{n}\right)$.
Again $\|g_{t-k}^{(n)}\|_F^2$ is bounded by a constant and an upper bound of $\|\psi_{tik}^{(n)}\|_F$ is Φ^{k-1} times a constant, $i = 1, 2$. Then we proceed like in the VAR(1) case.

Common features of tdVAR(1) & tdVMA(1) simulations

- We can obtain exact expressions for the terms $V_t^{(n)}$
- And also for $W_t^{(n)}$ for standard multivariate distributions (normal, Laplace, Student)
- We are able to compare the "theoretical" values of V (and W) to the empirical values through simulation
- tdVAR(1): $A_t^{(n)}(\theta) = \begin{pmatrix} 0.8 & 0.5 \\ 0 & 0.75L(t,n) \end{pmatrix}$,

$$g_t^{(n)} = \begin{pmatrix} 1 & 0 \\ 0 & e^{0.7L(t,n)} \end{pmatrix}, \quad \Sigma = I_2$$

- tdVMA(1): $B_t^{(n)}(\theta) = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.25 + 0.4L(t,n) \end{pmatrix}$,

$$g_t^{(n)} = \begin{pmatrix} 1 & -0.6 \\ -0.6 & e^{0.7L(t,n)} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 2 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

- 1000 simulations of series of length 100

ULB

Simulations for a tdVAR(1) model - normal errors

Table: Estimation results for the tdVAR$^{(n)}$(1) model, with Gaussian errors. "'avg"': average (across the 1000 simulations), "'std err"': standard error of the parameter estimates, "'std dev"': standard deviation (across the 1000 simulations), "'theor"': theoretical (based on true value), "'% rej. H_0: par.=true val."': percentages of simulations rejecting the hypothesis $H_0(\theta_i = \theta_i^0)$ at 5%.

Parameter θ_i	A'_{11}	A''_{22}	η_{22}
True value θ_i^0	0.8000	0.7500	0.7000
Avg estimates	0.7878	0.7329	0.6793
Avg std err (based on V)	0.0527	0.3391	0.2485
Std dev estimates	0.0548	0.3456	0.2606
Theor. std err	0.0530	0.3363	0.2425
% rej. H_0: par.=true val.	5.7	5.6	6.0

Simulations for a tdVAR(1) model - Laplace errors

Table: Estimation results for the tdVAR$^{(n)}$(1) model, with Laplace errors. "‘avg"’: average (across the 1000 simulations), "‘std err"’: standard error of the parameter estimates, "‘std dev"’: standard deviation (across the 1000 simulations), "‘theor"’: theoretical (based on true value), "‘% rej. H_0: par.=true val."’: percentages of simulations rejecting the hypothesis $H_0(\theta_i = \theta_i^0)$ at 5%.

Parameter θ_i	A'_{11}	A''_{22}	η_{22}
True value θ_i^0	0.8000	0.7500	0.7000
Avg estimates	0.7904	0.7238	0.6846
Avg std err (based on V)	0.0522	0.3425	0.2514
Avg std err (based on $V^{-1}WV^{-1}$)	0.0516	0.3136	0.3627
Std dev estimates	0.0582	0.3374	0.3794
Theor. std err	0.0530	0.3363	0.3834
% rej. H_0: par.=true val.	8.3	8.3	6.8

Simulations for a tdVMA(1) model - normal errors

Table: Estimation results for the tdVMA$^{(n)}$(1) model, with normal errors. "'avg"': average (across the 1000 simulations), "'std err"': standard error of the parameter estimates, "'std dev"': standard deviation (across the 1000 simulations), "'theor"': theoretical (based on true value), "'% rej. H_0: par.=true val."': percentages of simulations rejecting the hypothesis $H_0(\theta_i = \theta_i^0)$ at 5%.

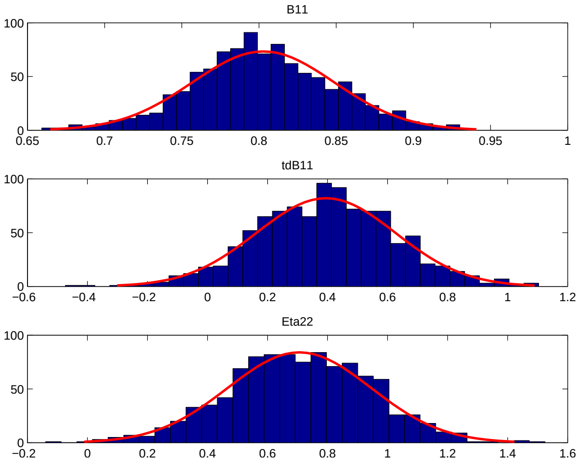
Parameter θ_i	B'_{11}	B''_{22}	η_{22}
True value θ_i^0	0.8000	0.4000	0.7000
Avg estimates	0.8026	0.3994	0.7117
Avg std err (based on V)	0.0437	0.2239	0.1525
Avg std err (based on $V^{-1}WV^{-1}$)	0.0516	0.3136	0.3627
Std dev estimates	0.0461	0.2299	0.1605
Theor. std err	0.0423	0.2223	0.1450
% rej. H_0: par.=true val.	7.0	6.6	6.8

Simulations for a tdVMA(1) model - Student 5 errors

Table: Estimation results for the tdVMA$^{(n)}$(1) model, with Student errors with 5 d.f.. "‘avg"’: average (across the 1000 simulations), "‘std err"’: standard error of the parameter estimates, "‘std dev"’: standard deviation (across the 1000 simulations), "‘theor"’: theoretical (based on true value), "‘% rej. H_0: par.=true val."’: % of simulations rejecting the hypothesis $H_0(\theta_i = \theta_i^0)$ at 5%.

Parameter θ_i	B'_{11}	B''_{22}	η_{22}
True value θ_i^0	0.8000	0.4000	0.7000
Avg estimates	0.8028	0.3946	0.7061
Avg std err (based on V)	0.0442	0.2252	0.1530
Avg std err (based on $V^{-1}WV^{-1}$)	0.0455	0.2153	0.2475
Avg std err (same, DERIVEST)	0.0456	0.2154	0.2475
Std dev estimates	0.0460	0.2319	0.2389
% rej. H_0: par.=true val.	6.5	8.8	4.5

Figure: Histograms of estimates for the 3 parameters of the tdVMA model with 5 d.f. Student errors, compared with normal density



Conclusions

This presentation is mainly based on four papers

- Paper 1 [AℓAM20??]: How to do the asymptotics of tdVARMA$^{(n)}$ models and apply it to simple models, like tdVAR(1) and tdVMA(1) models?
- Paper 2 [AM20??a]: How to improve the fundamental justifications of [AM2006] in the array case and solve the problems of their use (moments, existence of the information matrix
- Paper 3 [AℓALM2017] How to use its Technical Appendix?
- Paper 4 [AℓAJM2016] How to improve its AJM program?

We hope to have answered all these questions.

Thank you. Comments are welcome
References follow

Some references I

Abramowitz, M. & Stegun I. A. (1965) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.

Aït-Sahalia, Y. (2015) User's guide for the Matlab library implementing closed form MLE for diffusions, Dept of Economics & Bendheim Center for Finance, Princeton University.

[Aℓ2012] Alj., A. (2012). Contribution to the estimation of VARMA models with time-dependent coefficients, Ph.D. Thesis, Université libre de Bruxelles.

[AℓALM2017] Alj, A., Azrak, R., Ley, C. & Mélard, G. (2017). Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, *Scandinavian Journal of Statistics*, DOI: 10.1111/sjos.12268.

Some references II

[AℓAM2014] Alj, A., Azrak, R, and Mélard, G. (2014) On conditions in central limit theorems for martingale difference arrays, *Economics Letters* **123** (3), 305-307.

[AℓAM20??] Alj, A., Azrak, R. & Mélard, G. (20??). General estimation results for tdVARMA array models, in preparation.

[AℓJM2016] Alj, A., Jónasson, K., and Mélard, G. (2016) The exact Gaussian likelihood estimation of time-dependent VARMA models, *Comput. Statist. Data Anal.*, **100**, 633-644.

[AM1998] Azrak, R. and Mélard, G. (1998) Exact quasi-likelihood of time-dependent ARMA models, *J. Statist. Plann. Inference* **68**(1), 31−45.

[AM2006] Azrak R. and Mélard, G. (2006). Asymptotic properties of quasi-maximum likelihood estimators for ARMA models with time-dependent coefficients, *Statist. Inference for Stoch. Processes* **9**, No 3, 279-330.

Some references III

[AM20??a] Azrak, R. and Mélard, G. (20??a) Asymptotic properties of conditional least-squares for array time series, in preparation.

[AM20??b] Azrak, R, and Mélard, G. (20??b) Autoregressive models with time-dependent coefficients - A comparison between several approaches, submitted.

[AM20??c] Azrak, R. and Mélard, G. (20??c) ARMA models with time-dependent coefficients with economic examples, in preparation.

Box, G. E. P., Jenkins, G. M., Reinsel G. C. & Ljung, G. M. (2015). *Time series analysis, forecasting and control*, 5th edn. Wiley, New York.

D'Errico, J. (2007) Package DERIVEST, `http://www.mathworks.com/matlabcentral/fileexchange/13490`

Erdélyi, A. (1953) *Higher Transcendental Functions, Vol. 1*, McGraw-Hill, New York.

Some references IV

Hamdoune, S. (1995) Etude des problèmes d'estimation de certains modèles ARMA évolutifs, Thesis presented at Université Henri Poincaré, Nancy 1.
Jónasson, K. (2008) Algorithm 878: Exact VARMA likelihood and its gradient for complete and incomplete data with Matlab, *ACM Trans. Math. Softw.* **35**(1), 6.
Jónasson, K. and Ferrando, S. E. (2008) Evaluating exact VARMA likelihood and its gradient when data are incomplete, *ACM Trans. Math Softw.* **35**(1), 5.
Klimko, L. A. & Nelson, P. I. (1978) On conditional least squares estimation for stochastic processes, *Annals of Statistics* **6**, 629-642.
Lehmann, E. L., Casella, G. (1998) *Theory of Point Estimation*. Springer Verlag, New York.
[M1982] Mélard, G. (1982) The likelihood function of a time-dependent ARMA model, in O. D. Anderson and M. R. Perryman

Some references V

(eds), *Applied Time Series Analysis*, North-Holland, Amsterdam, 229-239.

[M1985] Mélard, G. (1985) *Analyse de données chronologiques*, Coll. Séminaire de mathématiques supérieures - Séminaire scientifique OTAN (NATO Advanced Study Institute) **89**, Presses de l'Université de Montréal, Montréal. Reprint `http://homepages.ulb.ac.be/~gmelard/Montreal1985_2012.pdf`

Meng, Y. & Lin Z. (2009) Maximal inequalities and laws of large numbers for L_q-mixingale arrays. *Statistics and Probability Letters* **79**, 1539-1547.