Modèles tdVARMA$^{(n)}$ à coefficients dépendant du temps : propriétés des estimateurs

Guy Mélard [1], Abdelkamel Alj [2] & Rajae Azrak [3]

[1] *Université libre de Bruxelles, SBS-EM, Ecares, Bruxelles, Belgique, gmelard@ulb.ac.be*
[2] *Université Moulay Ismail, FSJES, Meknès, Maroc, abdelkamel.alj@gmail.com*
[3] *Université Mohammed V - Rabat, Faculté des Sciences juridiques, économiques et sociales, Salé, Maroc, rajae.azrak@gmail.com*

Résumé. Dans un article récent, les modèles VARMA à coefficients dépendant du temps t mais pas de la longueur n de la série, ou tdVARMA, ont été étudiés. Lors des journées de statistique de 2017, une nouvelle théorie asymptotique assez générale a été proposée pour l'estimation paramétrique des processus stochastiques à temps discret, non nécessairement stationnaires et ergodiques. Nous appliquons ici ces résultats aux modèles tdVARMA$^{(n)}$ où les coefficients peuvent (mais ne doivent pas) dépendre de n, et aussi la matrice de covariance des erreurs. Contrairement à l'approche des processus localement stationnaires, les coefficients ne doivent pas dépendre de t/n et il n'est pas exigé qu'ils soient des fonctions lisses du temps. Les aspects numériques et pratiques de l'estimation sont également discutés. Enfin, nous traitons le modèle tdVMA$^{(n)}$ d'ordre 1 qui représente un cas particulier de tdVARMA$^{(n)}$, où nous vérifions que les hypothèses de la théorie sont satisfaites et pour lequel des résultats simulés sont présentés.

Mots-clés. Processus non stationnaire, série chronologique multivariée, modèles variant avec le temps.

Abstract.
In a recent paper, VARMA models with coefficients dependent on time t but not on the time series length n, or tdVARMA, have been studied. During the "Journées de statistique 2017", a new rather general asymptotic theory was proposed for the parametric estimation of stochastic processes in discrete time, not necessarily stationary and ergodic. We apply here these results to tdVARMA$^{(n)}$ models where the coefficients can (but don't need to) depend on n, and also the covariance matrix of the errors. Contrarily to the approach of locally stationary processes, the coefficients do not need to depend on t/n and it is not required that they are smooth functions of time. The numerical and practical aspects of estimation are also discussed. Finally, we treat a particular case of tdVARMA$^{(n)}$ model with a tdVMA$^{(n)}$ of order 1 where we verify that the assumptions of the theory are satisfied and for which simulation results are presented.

Keywords. Non-stationary process, multivariate time series, time-varying models.

1 Introduction

Dans l'article récent de Alj *et al.* (2017), les modèles VARMA à coefficients dépendant du temps t mais pas de la longueur n de la série, ou tdVARMA, ont été étudiés. Lors des journées de statistique de 2017, une nouvelle théorie asymptotique assez générale a été proposée par Azrak et Mélard (2017) pour l'estimation paramétrique des processus stochastiques à temps discret, non nécessairement stationnaires et ergodiques. Elle corrige, améliore et étend aux processus multivariés ce qui a été proposé par Azrak et Mélard (2006). Nous appliquons ici ces résultats aux modèles tdVARMA$^{(n)}$ où les coefficients peuvent (mais ne doivent pas, contrairement à Dahlhaus, 2000) aussi dépendre de n. Contrairement à l'approche des processus localement stationnaires, les coefficients ne doivent pas dépendre de t/n et il n'est pas exigé qu'ils soient des fonctions lisses du temps. Ceci généralise les résultats de Alj *et al.* (2017) à ces modèles tdVARMA$^{(n)}$.

Les aspects numériques et pratiques de l'estimation sont également discutés. L'algorithme AJM2 utilisé est basé sur une modification de l'algorithme de Alj *et al.* (2016). Enfin, nous traitons aussi le modèle tdVMA$^{(n)}(1)$ qui représente un cas particulier de tdVARMA$^{(n)}(p, q)$, avec $p = 0$ et $q = 1$ où nous vérifions que les hypothèses de la théorie sont satisfaites et pour lequel des résultats simulés sont présentés.

2 Modèles tdVARMA$^{(n)}$

Considérons la famille des variables aléatoires $x_t^{(n)}$ avec $t \leq n,\ n \in \mathbf{N}$, à valeurs dans $\mathbf{R}^r$ et définies sur l'espace de probabilité (Ω, F, P_θ). Sa distribution dépend d'un vecteur $\theta = (\theta_1, ..., \theta_m)^T$ de paramètres. Soient $E_\theta(.)$ et $E_\theta(./.)$ l'espérance et l'espérance conditionnelle sous P_θ. La vraie valeur de θ est notée $\theta^0 = (\theta_1^0, ..., \theta_m^0)^T$. Un modèle tdVARMA$^{(n)}$ d'ordre $(p,\, q)$, noté tdVARMA$^{(n)}(p,\, q)$, de moyenne 0, est défini par l'équation

$$x_t^{(n)} = \sum_{i=1}^{p} A_{ti}^{(n)}(\theta) x_{t-i}^{(n)} + g_t^{(n)}(\theta)\epsilon_t + \sum_{j=1}^{q} B_{tj}^{(n)}(\theta) g_{t-j}^{(n)}(\theta)\epsilon_{t-j}, \tag{1}$$

où p et q sont des entiers positifs, $\{\epsilon_t : t \in \mathbf{N}\}$ est un processus bruit blanc constitué de vecteurs aléatoires indépendants, de moyenne 0, de matrice de covariance Σ inversible $r \times r$, et où les coefficients $A_{ti}^{(n)}(\theta)$, $i = 1, ..., p$, $B_{tj}^{(n)}(\theta)$, $j = 1, ..., q$, et $g_t^{(n)}(\theta)$ sont des matrices $r \times r$, fonctions déterministes du temps t, de n et de θ. Notons que Σ est considéré comme un paramètre de nuisance. Pour la théorie, les valeurs initiales $x_t^{(n)}, t < 1$, et ϵ_t, $t < 1$, sont égales à 0. Soit $\{F_t : t \in \mathbf{N}\}$ la suite de sous-sigma-algèbres de F avec F_t généré par $\{\epsilon_u : u = 1, 2, ..., t\}$. Supposons que $\kappa := E\left(\mathrm{vec}(\epsilon_t\epsilon_t^T)\,\mathrm{vec}(\epsilon_t\epsilon_t^T)^T\right)$ ne dépend pas de t, pour la simplicité.

Notons $e_t^{(n)}(\theta)$ le résidu en t, c'est-à-dire $e_t^{(n)}(\theta) := x_t^{(n)} - E_\theta(x_t^{(n)}/F_{t-1})$ et $\Sigma_t^{(n)}(\theta) := E_\theta[e_t^{(n)}(\theta)e_t^{(n)T}(\theta)] = g_t^{(n)}(\theta)\Sigma g_t^{(n)T}(\theta)$, supposée inversible. On a $e_t^{(n)}(\theta^0) = g_t^{(n)}(\theta^0)\epsilon_t$. L'estimation est réalisée en maximisant la quasi-vraisemblance gaussienne, ce qui revient à minimiser la somme de $\alpha_t^{(n)}(\theta) = \log(\det(\Sigma_t^{(n)}(\theta))) + e_t^{(n)T}(\theta)\Sigma_t^{(n)-1}(\theta)e_t^{(n)}(\theta)$ pour $t = 1,\ldots,n$.

Comme pour Alj *et al.* (2017), nous utiliserons les coefficients des expressions autorégressive pure et moyenne mobile pure des processus tdVARMA$^{(n)}$, respectivement $x_t^{(n)} = \sum_{k=1}^{t-1}\pi_{tk}^{(n)}(\theta)x_{t-k}^{(n)} + e_t^{(n)}(\theta)$, $x_t^{(n)} = e_t^{(n)}(\theta) + \sum_{k=1}^{t-1}\psi_{tk}^{(n)}(\theta)e_{t-k}^{(n)}(\theta)$, dont les coefficients sont calculables par récurrence, voir Mélard (1985). Cette dernière relation donne pour $\theta = \theta^0$: $x_t^{(n)} = g_t^{(n)}\epsilon_t + \sum_{k=1}^{t-1}\psi_{tk}^{(n)}g_{t-k}^{(n)}\epsilon_{t-k}$, où $g_t^{(n)} := g_t^{(n)}(\theta^0)$ et $\psi_{tk}^{(n)} := \psi_{tk}^{(n)}(\theta^0)$. Ceci permet surtout d'écrire des expressions pour les dérivées des $e_t^{(n)}(\theta)$ par rapport à θ jusqu'à l'ordre 3, comme par exemple $\partial e_t^{(n)}(\theta)/\partial\theta_i = \sum_{k=1}^{t-1}\psi_{tik}^{(n)}(\theta,\theta^0)g_{t-k}^{(n)}\epsilon_{t-k}$. Posons $\psi_{tik}^{(n)} := \psi_{tik}^{(n)}(\theta^0,\theta^0)$.

Considérons un cas particulier de (1) pour $p = 0$, $q = 1$, un processus tdVMA$^{(n)}(1)$, avec $B_t^{(n)}(\theta) := B_{1t}^{(n)}(\theta)$ et $B_t^{(n)} = B_t^{(n)}(\theta^0)$. Nous avons $\pi_{tk}^{(n)}(\theta) = (-1)^{k+1}\prod_{l=0}^{k-1}B_{t-l}^{(n)}(\theta)$, $k = 1,...,t-1$, et $\psi_{t1}^{(n)}(\theta) = B_t^{(n)}(\theta)$, $\psi_{tk}^{(n)}(\theta) = 0$, $k > 1$. Notons $B_t^{(n)[k-1]} = \prod_{l=1}^{k-1}B_{t-l}^{(n)}$ et $B_t^{(n)[0]} = I_r$. Dès lors, pour $i = 1,...,m$,

$$\psi_{tik}^{(n)} = (-1)^{k+1}\left[\left.\frac{\partial B_t^{(n)[k-1]}(\theta)}{\partial\theta_i}\right|_{\theta=\theta^0} - \left.\frac{\partial B_t^{(n)[k-2]}(\theta)}{\partial\theta_i}\right|_{\theta=\theta^0} B_{t-k+1}^{(n)}\right], \quad k = 1,...,t-1. \quad (2)$$

3 Hypothèses et résultats asymptotiques

Les hypothèses sont similaires à celles de Alj *et al.* (2017), sauf que la plupart des symboles comportent $^{(n)}$. Voici les principales hypothèses :

(i) les matrices $A_{ti}^{(n)}(\theta)$, $B_{tj}^{(n)}(\theta)$ et $g_t^{(n)}(\theta)$ sont trois fois continûment différentiables par rapport à θ, $i = 1,\ldots,p$, $j = 1,\ldots,q$,

(ii) des majorants comme $\sum_{k=\nu}^{t-1}\|\psi_{tik}^{(n)}\|_F^2 < N_1P(\nu)\Phi^{\nu-1}$, $\sum_{k=\nu}^{t-1}\|\psi_{tik}^{(n)}\|_F^4 < N_2P(\nu)\Phi^{\nu-1}$, où $\|.\|_F$ est la norme de Frobenius, avec des constantes positives N_1, N_2, $0 < \Phi < 1$, $P(\nu)$ est un polynôme et $\nu = 1,...,t-1$, $i = 1,\ldots,m$,

(iii) l'existence des moments d'ordre $4+\delta$ des ϵ_t, $\delta > 0$, et donc de κ, et des bornes sur $\|\Sigma_t^{(n)}\|_F$ et $\|\Sigma_t^{(n)-1}\|_F$, et leurs dérivées,

(iv) l'existence de la matrice V strictement définie positive, où $V = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}V_t^{(n)}$ avec des éléments $V_{t,ij}^{(n)}$, $i,j = 1,...,m$, donnés par

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

(v) l'existence de $W = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}W_t^{(n)}$, $W_t^{(n)} = \frac{1}{4}E_{\theta^0}(\partial\alpha_t^{(n)}(\theta)/\partial\theta\ \partial\alpha_t^{(n)}(\theta)/\partial\theta^T)$,

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

L'estimateur du quasi-maximum de vraisemblance est alors convergent en probabilité et asymptotiquement normal, plus précisément $n^{1/2}(\widehat{\theta}_n - \theta^0) \xrightarrow{L} \mathcal{N}(0, V^{-1}WV^{-1})$, où V et W sont définis ci-dessus, sachant que W dépend de κ.

Notons que, même si une partie de la démonstration est similaire à celle du cas sans $^{(n)}$ et si l'annexe technique de Alj *et al.* (2017) est aisément adaptée, les arguments de base sont très différents. En effet, on doit employer ici la théorie de Azrak et Mélard (2017) au lieu du théorème de Klimko et Nelson. On doit aussi invoquer une loi faible des grands nombres et un théorème central limite pour des tableaux de martingale ainsi qu'une loi faible des grands nombres pour un tableau de mixingale, ceci au lieu de résultats de convergence presque sûre pour des suites de martingale.

4 Aspects numériques et pratiques

Contrairement à Alj *et al.* (2017), on ne se limite pas du point de vue pratique aux processus gaussiens, auquel cas $V = W$. Nous avons développé AJM2 sous Matlab comme amélioration de l'algorithme AJM de Alj *et al.* (2016), notamment pour estimer W. Nous avons également ajouté un calcul plus précis des dérivées numériques en employant DERIVEST et la possibilité de générer des ϵ_t selon des lois multivariées de Laplace et de Student.

5 Exemple : tdVMA$^{(n)}$(1)

Rappelons que le modèle tdVMA$^{(n)}$(1) est un cas particulier de modèle tdVARMA$^{(n)}(p,q)$ où $p = 0$ et $q = 1$. La vérification analytique des hypothèses est complexe. Limitons-nous ici à un modèle bivarié dépendant de 3 paramètres et défini par l'équation

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

où $L(t,n) = \frac{t-\frac{n+1}{2}}{n-1}$, et B'^0_{22}, α et β sont fixés. On suppose pour les vraies valeurs que $|B'^0_{11}| < 1$, $|B'^0_{22}|e^{B''^0_{22}/2} < 1$ et $\eta^0_{22} > 0$. Alors, grâce à (2), on peut vérifier toutes les hypothèses (i)-(vi) de la section 3, en employant plusieurs théorèmes de Azrak et Mélard (2017) et dans le calcul de la matrice V, on remplace la limite de la moyenne par une intégrale. On peut obtenir des expressions exactes pour les termes $V_t^{(n)}$ et aussi pour $W_t^{(n)}$, au moins dans le cas de lois multivariées de Laplace et de Student.

Table 1: Résultats d'estimation d'un modèle tdVMA$^{(n)}$(1) avec erreurs de Student à 5 degrés de liberté, $n = 100$. "val." : valeur, "Moy" : moyenne (à travers les 1000 simulations), "Err.-T." : erreur-type des estimations, "Ec.-T." : écart-type (à travers les simulations), "théorique" : valeur théorique (basée sur les vraies valeurs), "% rej. H_0: par. = vraie val." : pourcentages de simulations où l'hypothèse $H_0(\theta_i = \theta_i^0)$ est rejetée à 5%.

Paramètre θ_i	B'_{11}	B''_{22}	η_{22}
Vraies val. θ_i^0	0,8000	0,4000	0,7000
Moy. estimations	0,8030	0,3946	0,7066
Moy. Err.-T. (basée sur V^{-1})	0,0448	0,2287	0,1553
Moy. Err.-T. (basée sur $V^{-1}WV^{-1}$)	0,0447	0,2156	0,1873
Moy. Err.-T. (même DERIVEST)	0,0448	0,2156	0,1874
Ec.-T. des estimations	0,0462	0,2320	0,2394
Err.-T theorique	0,0423	0,2230	0,2894
% rej. H_0: paramètre = vraie val.	7,0	8,8	13,8

6 Résultats de simulation

On emploie le modèle tdVMA$^{(n)}$(1) suivant avec dépendance linéraire pour le coefficient :

$$B_t^{(n)}(\theta) = \begin{pmatrix} 0,8 & 0 \\ 0 & 0,25 + 0,4L(t,n) \end{pmatrix},\ g_t^{(n)} = \begin{pmatrix} 1 & -0,6 \\ -0,6 & e^{0,7L(t,n)} \end{pmatrix},\ \Sigma = \begin{pmatrix} 2 & 0,8 \\ 0,8 & 1 \end{pmatrix}.$$

Avec 1000 simulations de séries de longueur 100, on peut obtenir les moyennes des estimations des paramètres et des erreurs-types et comparer ces dernières aux écarts-types des estimations et aux valeurs tirées des résultats théoriques. Les résultats empiriques sont donnés dans le tableau 1. Ils sont évidemement encore meilleurs pour des séries plus longues.

Bibliographie

Alj, A., Azrak, R., Ley, C., and Mélard, G. (2017), Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, *Scandinavian Journal of Statistics*, 44, 617–635, 2017. DOI: 10.1111/sjos.12268.

Alj, A., Jónasson, K., and Mélard, G. (2016), The exact Gaussian likelihood estimation of time-dependent VARMA models, *Comput. Statist. Data Anal.*, 100 (2016) 633–644. DOI: 10.1016/j.csda.2014.07.006.

Azrak, R. and Mélard, G. (2006), Asymptotic properties of quasi-likelihood estimators for ARMA models with time-dependent coefficients, *Statistical Inference for Stochastic Processes*, 9, 279–330.

Azrak, R. et Mélard, G. (2017), Nouveaux résultats pour des modèles ARMA à coefficients dépendant du temps, Actes des 49èmes Journées de Statistique, Avignon, France, 29 mai-2 juin 2017.

Dahlhaus, R. (2000), A likelihood approximation for locally stationary processes. *Annals of Statistics*, 28, 1762–1794.

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.