Nouveaux résultats pour des modèles ARMA à coefficients dépendant du temps

Guy Mélard [1] & Rajae Azrak [2]

[1] *Université libre de Bruxelles et ITSE sprl, Bruxelles, Belgique, gmelard@ulb.ac.be*
[2] *Université Mohamed V, Rabat, Maroc, rajaeazrak@yahoo.fr*

Résumé. Cette contribution est une suite de travaux où nous avons étudié les modèles VARMA à coefficients dépendant du temps, ou modèles tdVARMA. Ici, nous traitons le cas où les coefficients dépendent aussi de la longueur de la série chronologique. Nous proposons des résultats théoriques nouveaux : (i) un théorème fondamental pour la théorie asymptotique ; (ii) un autre qui permet de réduire la supposition sur les moments ; (iii) deux théorèmes pour établir la convergence des deux matrices V et W impliquées dans la matrice de covariance asymptotique de l'estimateur. Nous appliquons ces résultats à certains modèles tdVAR(1). En particulier nous montrons des résultats de simulation pour différents types de distributions (y compris celle de Laplace) et nous comparons les erreurs-types à celles déduites de la théorie.

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

Abstract. This contribution is a follow-up to works where we have studied VARMA models with time-dependent coefficients, or tdVARMA models. Here we focus on the case where the coefficients depend also of the time series length. We provide new theoretical results: (i) a fundamental theorem for the asymptotic theory; (ii) a theorem for reducing the assumption on moments; (iii) two theorems to establish convergence for the two matrices V and W involved in the asymptotic covariance matrix of the estimator. We apply these results to some time-dependent VAR(1) models. In particular we show simulation results for different types of distributions (including the multivariate Laplace) and compare the standard errors to those deduced from the theory.

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

1 Introduction

Cette contribution est une suite de Alj et al. (2015), voir Alj et al. (2017), où nous avons étudié les modèles VARMA à coefficients dépendant du temps, ou modèles tdVARMA. Ici, nous traitons le cas où les coefficients dépendent du temps mais aussi de la longueur de la série chronologique, notée n. Nous les noterons tdVARMA$^{(n)}$. Puisqu'une grand partie de la communication sera illustrée par des modèles univariés, nous donnons ici la définition d'un modèle tdARMA$^{(n)}$.

Définition 1 *Une série chronologique $(x_t^{(n)}, t = 1, \dots n)$ suit un modèle tdARMA$^{(n)}$(p, q) si et seulement si elle vérifie l'équation*

$$x_t^{(n)} = \sum_{k=1}^{p} \phi_{tk}^{(n)}(\beta) x_{t-k}^{(n)} + e_t^{(n)}(\beta) - \sum_{k=1}^{q} \theta_{tk}^{(n)}(\beta) e_{t-k}^{(n)}(\beta),$$

où le vecteur $\beta \in B \subset \mathbb{R}^m$, de vraie valeur β^0 intérieure à l'ouvert B, contient les m paramètres à estimer, $e_t^{(n)}(\beta^0) = g_t^{(n)}(\beta^0)\epsilon_t$, les ϵ_t sont une suite de variables aléatoires indépendantes de moyenne 0 et d'écart-type σ, traité comme paramètre de nuisance (donc non inclus dans β).

Les coefficients $\phi_{tk}^{(n)}(\beta)$, $k = 1, ..., p$, $\theta_{tk}^{(n)}(\beta)$, $k = 1, ..., q$, $g_t^{(n)}(\beta)$, respectivement coefficients autorégressifs et moyenne mobile et facteur de dispersion, sont des fonctions déterministes du temps t, de β, et possiblement de la longueur n de la série. Pour la commodité on suppose des valeurs initiales $x_t^{(n)}$ et ϵ_t, $t < 1$, égales à 0.

Azrak et Mélard (2006) ont étudié ces modèles tdARMA$^{(n)}$ en proposant une méthode d'estimation de type quasi-maximum de vraisemblance. Dans le cas où les coefficients ne dépendent pas de n, donc en omettant les $^{(n)}$, Alj et al. (2015, 2017) ont géneralisé la théorie au cas vectoriel, c'est-à-dire quand x_t et ϵ_t sont des vecteurs $r \times 1$, et $\phi_{tk}(\beta)$, $\theta_{tk}(\beta)$ et $g_t(\beta)$ sont des matrices $r \times r$. Dans ce cas on note Σ la matrice de covariance de ϵ_t et la matrice de covariance du résidu $e_t(\beta)$ est $\Sigma_t(\beta) = g_t(\beta)\Sigma g_t^T(\beta)$.

Dans cette communication, nous revenons sur certains résultats de Azrak et Mélard (2006) pour les compléter. Nous proposons des résultats théoriques nouveaux : (i) un théorème fondamental pour la théorie asymptotique pour remplacer l'approche Klimko-Nelson faible utilisée ; (ii) un théorème qui permet de reduire la supposition sur les moments de 8 à un peu plus de 4 ; (iii) deux théorèmes pour établir la convergence des deux matrices V et W impliquées dans la formule sandwich qui fournit la matrice de covariance asymptotique de l'estimateur. Nous appliquons ces résultats à des modèles tdVAR$^{(n)}$(1). En particulier nous montrons des résultats de simulation pour différents types de distributions (y compris celle de Laplace) et nous comparons les erreurs-types à celles déduites de la théorie. Nous commençons par une illustration destinée à montrer l'intérêt des modèles tdARMA$^{(n)}$. Pour une approche différente, voir Dahlhaus (2000).

2 Illustration des modèles tdARMA$^{(n)}$

Nous avons employé un jeu de données de 320 séries de production industrielle des États-Unis de Proietti et Lü tkepohl (2013), de janvier 1986 à décembre 2016, en réservant l'année 2016 pour la comparer aux prévisions. Nous avons d'abord estimé des modèles à coefficients constants (cARIMA) par une procédure automatique de Tramo-Seats, Gómez & Maravall (2001), généralement des modèles ARIMA saisonniers avec différences ordinaires et saisonnières. Puis nous avons construit des modèles tdARIMA$^{(n)}$ en remplaçant

Table 1 – Pourcentages obtenus dans l'illustration.

Critère	pourcentage
Plus haute statistique $\|t\|$ des paramètres td $> 1,96$	$54,38\%$
SBIC tdARIMA $<$ SBIC cARIMA	$9,06\%$
écart-type résiduel tdARIMA $<$ cARIMA	$61,88\%$
$p(LB)$ tdARIMA $>$ cARIMA	$65,00\%$
MAPE prévisionnel tdARIMA $<$ cARIMA	$37,19\%$

les coefficients constants, par exemple $\theta_1 = \theta_1(\beta)$, par des fonctions linéaires du temps, dans cet exemple $\theta_{t1}^{(n)}(\beta) = \theta_1' + \theta_1''(t-(n+1)/2)/(n-1)$, et ceci pour les ordres $k \leq 12$ seulement. Nous n'avons pas omis les paramètres non significatifs.

Ensuite, nous avons comparé les résultats des modèles tdARIMA par rapport aux modèles cARIMA au moyen des critères suivants :

— la plus grande des statistiques de Student en valeur absolue, $|t|$, des paramètres dépendant du temps (par exemple θ_1'') est-elle supérieure à 1,96 ?
— le critère SBIC du modèle tdARIMA est-il inférieur à celui du modèle cARIMA ?
— l'écart-type résiduel du modèle tdARIMA est-il inférieur à celui du modèle cARIMA ?
— la probabilité de signification du test de Ljung-Box portant sur les autocorrélation résiduelles, $p(LB)$, du modèle tdARIMA est-elle supérieure à celle du modèle cARIMA ?
— le critère MAPE (mean absolute percentage error) prévisionnel sur 2016 du modèle tdARIMA est-il inférieur à celui du modèle cARIMA ?

Nous comptons le pourcentage à travers les séries. Les résultats sont résumés dans le tableau 1. S'ils ne prévoient pas mieux, les modèles tdARIMA sont en partie confirmés.

3 Nouveaux résultats

Azrak and Mélard (2006) ont justifié leur théorie asymptotique pour les modèles tdARMA$^{(n)}$ en invoquant essentiellement une version faible des théorèmes de Klimko et Nelson (1978), c'est-à-dire avec des convergences en probabilité au lieu de convergences presque sûres. Ils en déduisent un théorème (leur théorème 1') pour étudier les propriétés d'un estimateur des moindres carrés conditionnel. Nous commençons par énoncer ce théorème, ici en deux parties, les théorèmes 2.1 et 2.2, pour lesquels il est possible, voir Azrak et Mélard (2017) d'obtenir une démonstration directe, inspirée partiellement de Lehmann et Casella (1998). Pour la commodité, nous ne traitons ici que le cas univarié. Nous gardons la numérotation originale mais le théorème 2.4 est omis.

Considérons un espace probabilisé (Ω, F, P_β), et notons $E_\beta(\cdot)$ l'espérance sous P_β. Pour tout $n \in \mathbb{N}_0$, soit $x^{(n)} = \{x_t^{(n)}; t = 1, ..., n\}$ un processus stochastique. C'est ce qu'on appelle un tableau triangulaire de variables aléatoires (dépendantes). Considérons de même une filtration $\{F_t^{(n)}; t \in \mathbb{N}_0\}$ avec $F_0^{(n)} = \{\emptyset, \Omega\}$ et, pour tout $n \in \mathbb{N}_0$, $F_t^{(n)}$ est générée par un sous-ensemble arbitraire de $\{x_1^{(n)}, x_2^{(n)}, ..., x_t^{(n)}\}$. Supposons l'existence de $x_{t|t-1}^{(n)}(\beta) = E_\beta(x_t^{(n)}|F_{t-1}^{(n)})$, $t = 1, 2, ..., n$, trois fois continûment différentiable en β au voisinage de β^0. Notons $\partial_i(.) = (\partial(.)/\partial\beta_i)$ et de même $\partial_{ij}^2(.)$ pour la dérivée seconde par rapport à β_i et β_j, $i, j = 1, \ldots, m$. Soit $Q^{(n)}(\beta) = Q^{(n)}(\beta, x^{(n)}) = \sum_{t=1}^n \alpha_t^{(n)}(\beta)$ une fonction de pénalité à minimiser par rapport à β, où $\alpha_t^{(n)}(\beta) = \{x_t^{(n)} - x_{t|t-1}^{(n)}(\beta)\}^2/h_{t|t-1}^{(n)}(\beta)$ et $h_{t|t-1}^{(n)}(\beta) = E_\beta(\{x_t^{(n)} - x_{t|t-1}^{(n)}(\beta)\}^2|F_{t-1}^{(n)})$. Notons $\hat{\beta}^{(n)} = \operatorname{argmin}_{\beta \in B} Q^{(n)}(\beta)$ une suite d'estimateurs. Dans ce qui suit, plim représente la convergence en probabilité.

Théorème 2.1
Si les suppositions suivantes sont satisfaites pour tout $t = 1, ..., n$, uniformément en n, et $i, j = 1, ..., m$, où C_1 et C_2 sont deux constantes positives :

$\mathbf{H}_{2.1}$ pour un certain $\delta > 0$: $E_{\beta^0}(|\partial_i \alpha_t^{(n)}(\beta)|^{2+\delta}) \leq C_1$;

$\mathbf{H}_{2.2}$ $E_{\beta^0}(|\partial_{ij}^2 \alpha_t^{(n)}(\beta) - E_\beta(\partial_{ij}^2 \alpha_t^{(n)}(\beta)|F_{t-1}^{(n)})|^2) \leq C_2$;

$\mathbf{H}_{2.3}$ $\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{t=1}^n E_{\beta^0}(\partial_{ij}^2 \alpha_t^{(n)}(\beta)|F_{t-1}^{(n)}) = V_{ij}$, où $V = (V_{ij})_{1 \leq i,j \leq m}$ est une matrice strictement définie positive ;

$\mathbf{H}_{2.4}$ $\operatorname*{plimsup}_{n\to\infty\, \Delta\downarrow 0} \frac{1}{n\Delta} \left| \sum_{t=1}^n (\{\partial_{ij}^2 \alpha_t^{(n)}(\beta)\}_{\beta=\beta_{ij}^*} - \{\partial_{ij}^2 \alpha_t^{(n)}(\beta)\}_{\beta=\beta^0}) \right| < \infty$, où β_{ij}^* est un point de la droite joignant β^0 à tout β tel que $\|\beta - \beta^0\| < \Delta$, $0 < \Delta$;

alors il existe une suite d'estimateurs $\widehat{\beta}^{(n)}$ telle que $\operatorname{plim} \widehat{\beta}^{(n)} = \beta^0$ quand $n \to \infty$.

Théorème 2.2
Si les suppositions $\mathbf{H}_{2.1} - \mathbf{H}_{2.4}$ du théorème 2.1 sont satisfaites, aussi bien que $\mathbf{H}_{2.5}$ et $\mathbf{H}_{2.6}$ pour $i, j = 1, ..., m$,

$\mathbf{H}_{2.5}$ $\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{t=1}^n \{E_{\beta^0}(\partial_i \alpha_t^{(n)}(\beta)\partial_j \alpha_t^{(n)}(\beta)|F_{t-1}) - E_{\beta^0}(\partial_i \alpha_t^{(n)}(\beta)\partial_j \alpha_t^{(n)}(\beta))\} = 0$;

$\mathbf{H}_{2.6}$ il existe $W_{ij} = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^n E_{\beta^0}(\partial_i \alpha_t^{(n)}(\beta)\partial_j \alpha_t^{(n)}(\beta))$, où $W = (W_{ij})_{1 \leq i,j \leq m}$ est une matrice définie positive ;

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

Le théorème suivant est une condition suffisante pour $\mathbf{H}_{2.1}$. Il permet de remplacer la supposition d'existence des moments d'ordre 8 par ceux d'ordre $4 + 2\delta$, $\delta > 0$ dans Azrak et Mélard (2006) (et, moyennant adaptation, dans Alj et al. (2017)).

Théorème 2.3
Avec les notations du paragraphe 1, s'il existe $\delta > 0$, K_4, m_2, P_1, et P_2, tels que pour tout t et n, et $i = 1, ..., m$: $|\partial_i g_t^{(n)-2}(\beta^0)|^2 \leq K_4$, $|g_t^{(n)-2}(\beta^0)|^2 \leq m_2$, $E_{\beta^0}(|e_t^{(n)2}(\beta)|^{2+\delta}) \leq P_1$, $E_{\beta^0}(|\partial_i e_t^{(n)2}(\beta)|^{1+\delta/2}) \leq P_2$, alors $\mathbf{H}_{2.1}$ est satisfait pour ce δ.

Dans Azrak et Mélard (2006, exemples 2, 3 et 4) le théorème qui suit est implicite pour prouver l'existence de la limite V dans $\mathbf{H}_{2.3}$ du théorème 2.1 pour plusieurs modèles tdARMA. C'est valable aussi pour l'existence de la limite W dans $\mathbf{H}_{2.6}$.

Théorème 2.5
Soit $\{u_t^{(n)}, t = 1, ..., n\}$ et $\{v_t^{(n)}, t = 1, ..., n\}$ deux tableaux triangulaires de réels tels que $\frac{1}{n}\sum_{t=1}^{n} v_t^{(n)}$ converge absolument et $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} v_t^{(n)} = L$, et que $\{u_t^{(n)}\}$ converge vers $C > 0$ quand $t \to \infty$, et dès lors $n \to \infty$. Alors $\frac{1}{n}\sum_{t=1}^{n} u_t^{(n)} v_t^{(n)}$ converge quand $n \to \infty$ et sa limite est CL.

Le théorème 2.5 ne suffit pas pour évaluer des limites comme $\frac{1}{n}\sum_{t=1}^{n} v_t^{(n)}$. Le théorème 2.6 peut alors aider.

Théorème 2.6
Considérons $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} v_t^{(n)}$. Supposons qu'il existe une fonction $V(x)$ définie sur $[0, 1]$ et intégrable au sens de Riemann telle que $V(t/n) = v_t^{(n)}$. Alors $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} v_t^{(n)} = \int_0^1 V(x)dx$.

4 Exemple d'application aux modèles tdVAR$^{(n)}$(1)

Considérons le modèle d'équation $x_t^{(n)} = e_t^{(n)}(\beta) + \phi_t^{(n)}(\beta) x_{t-1}^{(n)}$ en dimension $r = 2$, avec des innovations multivariées normale ou Laplace. En notant $L(t, n) = \frac{t - \frac{n+1}{2}}{n-1}$, soit

$$\phi_t^{(n)}(\beta) = \begin{pmatrix} \phi'_{11} + \phi''_{11} L(t,n) & \phi'_{12} \\ 0 & \phi'_{22} + \phi''_{22} L(t,n) \end{pmatrix}, \quad g_t^{(n)}(\beta) = \begin{pmatrix} e^{\eta_{11} L(t,n)} & 0 \\ 0 & e^{\eta_{22} L(t,n)} \end{pmatrix},$$

et une matrice $\Sigma = I_2$. On traite le cas où plusieurs paramètres ont des valeurs fixées : $\phi'_{12} = 0,5$, $\phi'_{22} = \phi''_{11} = \eta_{11} = 0$. Donc $\beta = (\phi'_{11}, \phi''_{22}, \eta_{22})$. On peut montrer que les propriétés asymptotiques sont valables pourvu que $\Phi^{1/2} = \max\{|\phi'^0_{11}|, \frac{1}{2}|\phi''^0_{22}|\} < 1$ et $\eta^0_{22} \geq 0$. On peut obtenir des expressions pour les termes $V_t^{(n)}$ et $W_t^{(n)}$ de $V = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} V_t^{(n)}$ et $W = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} W_t^{(n)}$. On peut aussi obtenir des estimations de V et de W lors de l'estimation numérique par maximum de vraisemblance. On peut alors comparer les valeurs théoriques de V et W aux valeurs empiriques par simulation.

On a développé une stratégie de simulation, par exemple dans le tableau 2 pour une loi de Laplace.

TABLE 2 – Résultats d'estimation d'un modèle tdVAR$^{(n)}(1)$ avec erreurs de Laplace. "val." : valeur, "Moy" : moyenne (à travers les 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 $H_0(\beta_i = \beta_i^0)$ est rejetée à 5%.

Paramètre β_i	$-\phi'_{11}$	$-\phi''_{22}$	η_{22}
Vraies val. β_i^0	0,8000	0,7500	0,7000
Moy. estimations	0,7904	0,7238	0,6846
Moy. Err.-T. (basée sur V^{-1})	0,0522	0,3425	0,2514
Moy. Err.-T. (basée sur $V^{-1}WV^{-1}$)	0,0516	0,3136	0,3627
Ec.-T. des estimations	0,0582	0,3374	0,3794
Err.-T theorique	0,0530	0,3363	0,3834
% rej. H_0 : paramètre = vraie val.	8,3	8,3	6,8

Bibliographie

[1] Alj, A., Ley, C. et Mélard, G. (2015), Propriétés asymptotiques des estimateurs pour des modèles VARMA à coefficients dépendant du temps, avec exemples, *Actes des 47e Journées de Statistique*, Lille, 1-5 juin 2015. `http://papersjds15.sfds.asso.fr/submission_133.pdf`

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

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

[4] Azrak, R. et Mélard, G. (2017), Asymptotic properties of conditional least-squares estimators for array time series. En préparation.

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

[6] Gomez, V., et Maravall, A. (2001), Automatic modelling methods for univariate series. Chapter 7 in Peña, D. Tiao, G. C. and Tsay, R.S., *A Course in Time Series Analysis*, Wiley, New York, pp. 171-201.

[7] Klimko, L. A. et Nelson, P. I. (1978), On conditional least squares estimation for stochastic processes, *Annals of Statistics* 6, 629-642.

[8] Lehmann, E. L. et Casella, G. (1998), *Theory of Point Estimation*, Springer Verlag, New York.

[9] Proietti, T. et Lütkepohl, H. (2013), Does the Box-Cox transformation help in forecasting macroeconomic time series ? *International Journal of Forecasting* 29, 88–99.