Estimation des modèles autorégressifs multivariés à coefficients aléatoires

Nassim TOUCHE [1] & Abdelhakim AKNOUCHE [2]

[1] *Faculté des Sciences Exactes, Université de Bejaia, Algérie, touche.nassim@gmail.com.*

[2] *Faculté de Mathématiques, Université des sciences et de Technologie Houari Boumediene, Alger, Algérie, aknouche_ab@yahoo.com.*

Résumé. Nous proposons deux variantes de la méthode des moindres carrés pondérés en quatre étapes pour l'estimation d'un modèle autorégressif multivarié à coefficients aléatoires (*MRCA*). Les deux estimateurs proposés, qui diffèrent par les pondérations utilisées, jouissent des propriétés de convergence et de normalité asymptotique sous des hypothèses assez faibles, notamment sans aucune condition sur les moments du processus observé. De plus, ils sont tous deux asymptotiquement plus efficaces que la méthode des moindres carrés ordinaire en deux étapes déjà connu pour ce modèle. Application sur des données simulées et des données réelles illustre les méthodes proposées.

Mots-clés. Modèle RCA multivarié, modèle $ARCH$ multivarié, méthode OLS en deux étapes, méthode WLS en quatre étapes, estimateur de type Schick.

Abstract. We propose two variants of the four-stage weighted least squares method for estimating the multivariate random coefficient autoregressive model (*MRCA*). The proposed estimators, which differ from the used weights, are consistent and asymptotically normal under quite mild assumptions, in particular without any moment conditions on the observed process. Moreover, both estimates are asymptotically more efficient than the two-stage ordinary least squares method already known for this model. Application on both simulated and real data illustrates the proposed methods.

Keywords. Multivariate RCA, multivariate $ARCH$ models, two-stage OLS estimate, four-stage WLS estimate, Schick-type estimate.

1 Introduction

Considérons le modèle autorégressif multivarié à coefficients aléatoires d'ordre 1 ($MRCA\,(1)$),

$$y_t = (\Phi_0 + B_t)\, y_{t-1} + \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{1.1}$$

soumis aux hypothèses suivantes :

A0 Φ_0 *est une* $m \times m$*-matrice de nombres réelles.*

A1 $\{\varepsilon_t, t \in \mathbb{Z}\}$ *est une suite de* $m \times 1$*-vecteurs aléatoires indépendants identiquement distribués* (iid) *de moyenne nulle et de matrice de covariance* G_0*, définie positive.*

A2 $\{B_t, t \in \mathbb{Z}\}$ *est une suite de* $m \times m$*-matrices aleatoires iid de moyenne nulle avec hyper-matrice de variance* $E\left(vec\left(B_1\right) vec'\left(B_1\right)\right) = \Sigma_0$ *qui est définie positive.*

A3 *Les suites* $\{\varepsilon_t, t \in \mathbb{Z}\}$ *et* $\{B_t, t \in \mathbb{Z}\}$ *sont indépendantes.*

Nous faisons aussi l'hypothèse de stabilité suivante :

A4 *L'équation* (1.1) *admet une unique solution strictement stationnaire et ergodique.*

Une condition suffisante pour que **A4** soit satisfaite peut être explicitée en termes des entrées $\{(\varepsilon_t, B_t), t \in \mathbb{Z}\}$ de l'équation (1.1). En effet, si γ désigne le plus grand exposant de Lyapunov associé a (1.1) (Bougerol et Picard, 1992), alors d'après Bougerol et Picard (1992), **A4** est verifiée si $\gamma < 0$.

Sous **A0**-**A3**, l'espérance et la variance conditionnelles du modèle (1.1) sont données respectivement par

$$\begin{aligned} E\left(y_t/\mathcal{F}_{t-1}\right) &= \Phi_0 y_{t-1} = \left(y'_{t-1} \otimes I\right) vec\left(\Phi_0\right) := \mu_t\left(\Phi_0\right). \\ Var\left(y_t/\mathcal{F}_{t-1}\right) &= \left(y'_{t-1} \otimes I\right) \Sigma_0 \left(y_{t-1} \otimes I\right) + G_0 := H_t\left(\Sigma_0, G_0\right), \end{aligned}$$

I étant la matrice d'identité d'ordre m et $\mathcal{F}_t$ la tribu générée par $\{(\varepsilon_s, B_s), s \leq t\}$. Ainsi, (1.1) se réduit à un modèle $ARCH$ multivarié lorsque $\Phi_0 = 0$ et à un modèle AR vectoriel pour $\Sigma_0 = 0$.

Malgré la littérature abondante et continue concernant les modèles autorégressifs à coefficients aléatoires univariés (ex. Aue et Horvath, 2011, Truquet et Yao, 2012, Aknouche, 2013), il semble que peu de travaux ont été consacrés au cas mulivarié. Nicholls et Quinn (1982, Chapitre 7) sont les premiers à avoir proposé un estimateur des moindres carrés ordinaire en deux étapes ($2SLS$) pour le modèle (1.1). La première étape estime la matrice autorégressive Φ_0 tandis que la seconde estime les variances Σ_0 et G_0 des bruits du modèle. L'estimateur proposé est consistant et asymptotiquement normal, mais sous une condition de moment d'ordre huit, très contraignante sur le processus observé. D'autre part, Praskova et Vanecek (2011) ont proposé pour le modèle (1.1) une généralisation de l'estimateur de Schick (1996). Bien que leur estimateur possède certains avantages par rapport à l'estimateur $2SLS$, il présente cependant l'inconvénient de ne converger que sous l'hypothèse du moment d'ordre deux du processus observé. De plus, leur méthode n'estime que la matrice autorégressive Φ_0 et non les variances G_0 et Σ_0 du modèle. Dans ce travail nous proposons deux variantes de la méthode des moindres carrés pondérés en quatre étapes ($4SWLS$) pour l'estimlation des paramètres du modèle (1.1). Les estimateurs proposés sont consistants et asymptotiquement Gaussiens sans aucune condition sur les moments du processus observé. De plus, ils sont plus efficaces que l'estimateur $2SLS$ et, trivialement, s'adaptent au cas $MRCA\left(p\right)$ d'ordre supérieur.

2 Estimateurs $4SWLS$ pour un modèle $MRCA$

Supposons qu'une tranche $y_1, ..., y_n$ de la solution de (1.1) soit observée, partant de laquelle nous désirons estimer Φ_0, G_0 et Σ_0. Nous suivons ici la démarche des moindres carrés pondérés en quatre étapes, laquelle est basée sur deux régressions principales : La première concernant la moyenne conditionnelle du modèle, à savoir

$$y_t = E\left(y_t/\mathcal{F}_{t-1}\right) + u_t = \Phi_0 y_{t-1} + u_t, \qquad u_t = B_t y_{t-1} + \varepsilon_t, \tag{2.1}$$

tandis que la seconde est issue de la variance conditionnelle

$$u_t u_t' = G_0 + \left(y_{t-1}' \otimes I\right) \Sigma_0 \left(y_{t-1} \otimes I\right) + \Xi_t, \tag{2.2a}$$

où $\Xi_t = u_t u_t' - E\left(u_t u_t'/\mathcal{F}_{t-1}\right)$ et alors $E\left(\Xi_t/\mathcal{F}_{t-1}\right) = 0$. Sous forme vectorielle (2.2a) devient

$$v_t = W_{t-1}\theta_0 + e_t, \tag{2.2b}$$

où

$$\begin{aligned} \sigma_0 &= vech\left(\Sigma_0\right),\ g_0 = vech\left(G_0\right),\ \theta_0 = \left(g_0', \sigma_0'\right)',\ W_{t-1} = \left[\begin{array}{cc} I_{m(m+1)/2} & Z_{t-1}' \end{array}\right], \\ e_t &= vech\left(\Xi_t\right),\ v_t = vech\left(u_t u_t'\right),\ Z_{t-1}' = H\left[\left(y_{t-1} \otimes I\right) \otimes \left(y_{t-1} \otimes I\right)\right]' K', \end{aligned}$$

H ($\frac{m(m+1)}{2} \times m^2$) et K' ($m^4 \times \frac{m^2\left(m^2+1\right)}{2}$) étant respectivement les matrices d'élimination et de duplication donnant la relation entre les opérateurs vec et $vech$. Posons $\lambda_0 = \left(\phi_0', \theta_0'\right)'$ où $\phi_0 = vec\left(\Phi_0\right)$.

Nous proposons deux algorithmes d'estimation de λ_0 fondés sur les régressions (2.1) et (2.2). La démarche étant la même, mais la différence réside dans les pondérations utilisées.

2.1 Pondération via la variance conditionnelle

Le principe de la première méthode repose sur les quatre étapes suivantes effectuées séquentiellement.

Étape 1 En se basant sur la régression (2.1), la matrice autoregressive Φ_0 est estimée via la méthode des moindres carrés pondérés :

$vec\left(\widehat{\Phi}_{1WLS}\right) = \arg\min_{\Phi} \sum_{t=1}^{n} (y_t - \Phi y_{t-1})' \widetilde{H}_t^{-1} (y_t - \Phi y_{t-1})\,,$

où la pondération est l'inverse de la variance conditionnelle, $\widetilde{H}_t^{-1} = H_t^{-1}\left(\widetilde{\Sigma}, \widetilde{G}\right)$, évaluée à des matrices $\widetilde{\Sigma} > 0, \widetilde{G} > 0$ définies positives arbitrairement fixées.

Étape 2 Ensuite, les matrices G_0 et Σ_0 sont alors estimées en se basant sur (2.2b) via la méthode des moindres carrés pondérés :

$\widehat{\theta}_{1WLS} = \arg\min_\theta \sum_{t=1}^{n} \left(\widehat{v}_{1t} - W_{t-1}\theta\right)' \left(\widetilde{h}_t \widetilde{h}_t'\right)^{-1} \left(\widehat{v}_{1t} - W_{t-1}\theta\right),$

avec comme pondération le "carré" de la matrice de pondération précédente $\widetilde{h}_t \widetilde{h}_t'$ avec $\widetilde{h}_t = vech\left(\widetilde{H}_t\right)$. L'inobservable régressant v_t dans (2.2) étant remplacé par le résidu,

$\widehat{v}_{1t} = \widehat{v}_{1t}\left(\widehat{\Phi}_{1WLS}\right)$, de (2.1), lequel est obtenu à partir de l'estimateur $\widehat{\Phi}_{1WLS}$ de l'étape précédente.

Étape 3 Φ_0 est de nouveau estimée via la même méthode que dans la première étape, $vec\left(\widehat{\Phi}_{2WLS}\right) = \arg\min_{\Phi} \sum_{t=1}^{n} \left(y_t - \Phi y_{t-1}\right)' \widehat{H}_t^{-1} \left(y_t - \Phi y_{t-1}\right),$

mais en remplaçant la matrice de pondération $H_t\left(\widetilde{\Sigma}, \widetilde{G}\right)$ par $\widehat{H}_t = H_t(\widehat{\Sigma}_{1WLS}, \widehat{G}_{1WLS})$, laquelle est évaluée aux matrices $\widehat{\Sigma}_{1WLS}$ et $\widehat{G}_{1WLS}$ estimées dans la deuxième étape.

Étape 4 Enfin, les matrices de variance G_0 et Σ_0 sont de nouveau estimées en se basant toujours sur (2.2b) via la méthode des moindres carrés pondérés,

$\widehat{\theta}_{2WLS} = \arg\min_{\theta} \sum_{t=1}^{n} \left(\widehat{v}_{2t} - W_{t-1}\theta\right)' \left(\widehat{h}_t\widehat{h}_t'\right)^{-1} \left(\widehat{v}_{2t} - W_{t-1}\theta\right),$

mais dont la pondération est l'inverse du "carré" de la variance conditionnelle $\left(\widehat{h}_t\widehat{h}_t'\right)^{-1}$ avec $\widehat{h}_t = vech\left(\widehat{H}_t\right)$ est évaluée aux matrices $\widehat{\Sigma}_{1WLS}$ et $\widehat{G}_{1WLS}$ estimées à l'étape 2. Le régressant v_t dans (2.2b) est maintenant remplacé par le résidu $\widehat{v}_{2t}$ de (2.1), obtenu à partir de l'estimateur $\widehat{\Phi}_{2WLS}$ de l'étape 3.

Une méthode similaire a été proposée par Aknouche (2015) dans le cadre des modèles RCA univariés. En raison de la linéarité de la moyenne et la variance conditionnelles par rapports aux paramètres à estimer, la première méthode s'écrit en termes plus exacts comme suit.

Algorithme 2.1 *Étant donné deux matrices* $\widetilde{\Sigma} > 0$ *et* $\widetilde{G} > 0$ *définies positives arbitrairement fixées. La méthode consiste en un quadruplet* $\left(\widehat{\Phi}_{1WLS}, \widehat{\theta}_{1WLS}, \widehat{\Phi}_{2WLS}, \widehat{\theta}_{2WLS}\right)$ *d'estimateurs calculés selon les quatre étapes séquentielles suivantes.*

$$\begin{aligned}
vec\left(\widehat{\Phi}_{1WLS}\right) &= \left(\sum_{t=1}^{n} \left(y_{t-1} \otimes I\right) \widetilde{H}_t^{-1} \left(y_{t-1}' \otimes I\right)\right)^{-1} \sum_{t=1}^{n} \left(y_{t-1} \otimes I\right) \widetilde{H}_t^{-1} y_t \qquad (2.3) \\
\widehat{\theta}_{1WLS} &= \left(\sum_{t=1}^{n} W_{t-1}' \left(\widetilde{h}_t\widetilde{h}_t'\right)^{-1} W_{t-1}\right)^{-1} \sum_{t=1}^{n} W_{t-1}' \left(\widetilde{h}_t\widetilde{h}_t'\right)^{-1} \widehat{v}_{1t} \\
vec\left(\widehat{\Phi}_{2WLS}\right) &= \left(\sum_{t=1}^{n} \left(y_{t-1} \otimes I\right) \widehat{H}_t^{-1} \left(y_{t-1}' \otimes I\right)\right)^{-1} \sum_{t=1}^{n} \left(y_{t-1} \otimes I\right) \widehat{H}_t^{-1} y_t \qquad (2.4) \\
\widehat{\theta}_{2WLS} &= \left(\sum_{t=1}^{n} W_{t-1}' \left(\widehat{h}_t\widehat{h}_t'\right)^{-1} W_{t-1}\right)^{-1} \sum_{t=1}^{n} W_{t-1}' \left(\widehat{h}_t\widehat{h}_t'\right)^{-1} \widehat{v}_{2t}
\end{aligned}$$

où $\widetilde{H}_t = H_t\left(\widetilde{\Sigma}, \widetilde{G}\right)$, $\widetilde{h}_t = vech(\widetilde{H}_t)$, $\widehat{u}_{1t} = y_t - \widehat{\Phi}_{1WLS} y_{t-1}$, $\widehat{v}_{1t} = vech\left(\widehat{u}_{1t}\widehat{u}_{1t}'\right)$, $\widehat{H}_t = H_t\left(\widehat{\Sigma}_{1WLS}, \widehat{G}_{1WLS}\right)$, $\widehat{h}_t = vech(\widehat{H}_t)$, $\widehat{u}_{2t} = y_t - \widehat{\Phi}_{2WLS} y_{t-1}$ *et* $\widehat{v}_{2t} = vech\left(\widehat{u}_{2t}\widehat{u}_{2t}'\right)$.

Un avantage numérique de la méthode proposée et spécialement comparée à la méthode du quasi-maximum de vraisemblance est que les estimateurs fournis ont une forme explicite ne nécessitant point le recours à des routines d'optimisation numériques ô combien encombrantes, spécialement dans le présent cas multivarié.

La consistance et la normalité asymptotique de l'estimateur proposé sont établies sans aucune condition sur les moments du processus observé. Considérons l'hypothèse de moment suivante sur les bruits du modèle (1.1).

A5 $E\left(\varepsilon_1^{\otimes 4}\right) < \infty$, $E\left(vec\left(B_1\right)^{\otimes 4}\right) < \infty$.

Les matrices limites suivantes apparaissent dans le résultat de normalité asymptotique figurant dans le Théorème 2.1 ci-dessous. Elles existent bien entendu grâce à **A0-A5**.

$$\begin{aligned}
A\left(\lambda_0, \widetilde{\theta}\right) &= E\left(\left(y_{t-1} \otimes I\right) \widetilde{H}_t^{-1}\left(y_{t-1}' \otimes I\right)\right) \text{ et } A\left(\lambda_0, \theta_0\right) = E\left(\left(y_{t-1} \otimes I\right) H_t^{-1}\left(y_{t-1}' \otimes I\right)\right). \\
B\left(\lambda_0, \widetilde{\theta}\right) &= E\left(\left(y_{t-1} \otimes I\right) \widetilde{H}_t^{-1} H_t^{-1} \widetilde{H}_t^{-1}\left(y_{t-1}' \otimes I\right)\right) \text{ et } B\left(\lambda_0, \theta_0\right) = A\left(\lambda_0, \theta_0\right). \\
C\left(\lambda_0, \widetilde{\theta}\right) &= A^{-1}\left(\lambda_0, \widetilde{\theta}\right) B\left(\lambda_0, \widetilde{\theta}\right) A^{-1}\left(\lambda_0, \widetilde{\theta}\right) \text{ et } C\left(\lambda_0, \theta_0\right) = A^{-1}\left(\lambda_0, \theta_0\right).
\end{aligned}$$

$$\begin{aligned}
D\left(\lambda_0, \widetilde{\theta}\right) &= E\left(W_{t-1}'(\widetilde{h}_t \widetilde{h}_t')^{-1} W_{t-1}\right) \text{ et } D\left(\lambda_0, \theta_0\right) = E\left(W_{t-1}'(h_t h_t')^{-1} W_{t-1}\right) \\
F\left(\lambda_0, \widetilde{\theta}\right) &= E\left(W_{t-1}'(\widetilde{h}_t \widetilde{h}_t')^{-1} E\left(e_t e_t' / \mathcal{F}_{t-1}\right)(\widetilde{h}_t \widetilde{h}_t')^{-1} W_{t-1}\right), \\
L\left(\lambda_0, \widetilde{\theta}\right) &= D^{-1}\left(\lambda_0, \widetilde{\theta}\right) F\left(\lambda_0, \widetilde{\theta}\right) D^{-1}\left(\lambda_0, \widetilde{\theta}\right).
\end{aligned}$$

Théorème 2.1 *i) Sous **A0-A4**,*

$$\widehat{\Phi}_{1WLS} \underset{n\to\infty}{\overset{p.s.}{\to}} \Phi_0, \quad \widehat{\Phi}_{2WLS} \underset{n\to\infty}{\overset{p.s.}{\to}} \Phi_0, \quad \widehat{\theta}_{1WLS} \underset{n\to\infty}{\overset{p.s.}{\to}} \theta_0, \quad \widehat{\theta}_{2WLS} \underset{n\to\infty}{\overset{p.s.}{\to}} \theta_0.$$

*iii) Si de plus **A5** est vérifiée alors*

$$\begin{aligned}
&\sqrt{n}\left(\widehat{\Phi}_{1WLS} - \Phi_0\right) \underset{n\to\infty}{\overset{L}{\to}} N\left(0, C\left(\lambda_0, \widetilde{\theta}\right)\right), \ \sqrt{n}\left(\widehat{\Phi}_{2WLS} - \Phi_0\right) \underset{n\to\infty}{\overset{L}{\to}} N\left(0, A^{-1}\left(\lambda_0, \theta_0\right)\right). \\
&\sqrt{n}\left(\widehat{\theta}_{1WLS} - \theta_0\right) \underset{n\to\infty}{\overset{L}{\to}} N\left(0, L\left(\lambda_0, \widetilde{\theta}\right)\right), \quad \sqrt{n}\left(\widehat{\theta}_{1WLS} - \theta_0\right) \underset{n\to\infty}{\overset{L}{\to}} N\left(0, L\left(\lambda_0, \theta_0\right)\right).
\end{aligned}$$

Notons que la covariance asymptotique entre $\widehat{\Phi}_{1WLS}$ et $\widehat{\theta}_{1WLS}$ est nulle dans le cas de normalité du modèle.

2.2 Pondération par le résidu de la régression de variance conditionnelle

Nous proposons un deuxième algorithme d'estimation du modèle (1.1) dont le principe est le même que le premier, mais qui en diffère par le fait que les pondérations utilisées dans la deuxième et quatrième étapes dépendent de l'inverse du "carré" de l'erreur e_t de régression de variance conditionnelle (2.2b) plutôt que du carré, $vech\left(H_t\right) vech'\left(H_t\right)$, de la variance conditionnelle H_t. Il s'avère que l'algorithme proposé peut être vu comme extension de la méthode de Schick (1996) en deux paliers. D'abord, en intégrant l'estimation des variances du modèle, puis en généralisant cela au cas multivarié.

Algorithme 2.2 *Étant donné deux matrices* $\widetilde{\Sigma} > 0, \widetilde{G} > 0$*, définies positives arbitrairement fixées. L'algorithme consiste en le quadruplet* $\left(\widehat{\Phi}^*_{1WLS}, \widehat{\theta}^*_{1WLS}, \widehat{\Phi}^*_{2WLS}, \widehat{\theta}^*_{2WLS}\right)$ *d'estimateurs suivants.*

Étape 1 *Calculer vec* $\left(\widehat{\Phi}^*_{1WLS}\right)$ *d'après* (2.3).

Étape 2 $\widehat{\theta}^*_{1WLS} = \arg\min_\theta \sum_{t=1}^n (\widehat{v}_{1t} - W_{t-1}\theta)' (\widetilde{e}_t \widetilde{e}'_t)^{-1} (\widehat{v}_{1t} - W_{t-1}\theta)$,
où $\widetilde{e}_t = \widehat{v}^*_{1t} - W_{t-1}\widetilde{\theta}$, $\widehat{u}^*_{1t} = y_t - \widehat{\Phi}^*_{1WLS} y_{t-1}$ *et* $\widehat{v}^*_{1t} = vech\left(\widehat{u}^*_{1t}\widehat{u}^{*\prime}_{1t}\right)$.

Étape 3 *Calculer vec* $\left(\widehat{\Phi}^*_{2WLS}\right)$ *selon* (2.4) *tout en remplaçant* $\widehat{\theta}_{1WLS}$ *par* $\widehat{\theta}^*_{1WLS}$.

Étape 4 $\widehat{\theta}_{2WLS} = \arg\min_\theta \sum_{t=1}^n (\widehat{v}^*_{2t} - W_{t-1}\theta)' (\widehat{e}_t \widehat{e}'_t)^{-1} (\widehat{v}^*_{2t} - W_{t-1}\theta)$,
où $\widehat{e}_t = \widehat{v}^*_{2t} - W_{t-1}\widehat{\theta}^*_{1WLS}$, $\widehat{u}^*_{2t} = y_t - \widehat{\Phi}^*_{2WLS} y_{t-1}$ *et* $\widehat{v}^*_{2t} = vech\left(\widehat{u}^*_{2t}\widehat{u}^{*\prime}_{2t}\right)$.

Nous montrons sous **A0**-**A5** que l'estimateur fourni par l'Algorithme 2.2 est consistant et asymptotiquement Gaussien.

Enfin, nous montrons aussi que seulement sous l'hypothèse de Normalité du modèle, les deux estimateurs proposés ont le même niveau d'efficacité asymptotique. Une étude de simulation est alors conduite pour comparer les deux estimateurs en échantillons finis, tandis qu'une application réelle illustre le travail.

Bibliographie

[1] Aknouche, A. (2015). Quadratic random coefficient autoregression with linear in parameters volatility. *Statistical Inference for Stochastic Processes,* **18***,* 99-125.

[2] Aknouche, A. (2013). Two-stage weighted least squares estimation of nonstationary random coefficient autoregressions. *Journal of Time Series Econometrics*, **5**, 25-47.

[3] Aue, A. & Horváth, L. (2011). Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Statistica Sinica*, **21**, 973-999.

[4] Bougerol, P. and Picard, N. (1992). Strict stationarity of generalized autoregressive processes. *Annals of Probability*, **20**, 1714-1730.

[5] Nicholls, D.F. & Quinn, B.G. (1982). *Random coefficient autoregressive model*: *An introduction*. Springer Verlag, New York.

[6] Praskova Z. & Vanecek P. (2011). On a class of estimators in a multivariate $RCA(1)$ model. *Kybernetika*, **47**, 501-518.

[7] Schick, A. (1996). $\sqrt{n}$-consistent estimation in a random coefficient autoregressive model. *Australian Journal of Statistics*, **38**, 155-60.

[8] Truquet, L. & Yao, J. (2012). On the quasi-likelihood estimation for random coefficient autoregressions. *Statistics*, **46**, 505-521.