C. R. Acad. Sci. Paris, Ser. I 349 (2011) 817–820

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistique/Économie mathématique

Estimation de la matrice de variance asymptotique des estimateurs du QMV de modèles ARMA faibles multivariés

Estimating the asymptotic variance matrix of the QMLE of weak multivariate ARMA models

Yacouba Boubacar Maïnassara

Université Lille III, EQUIPPE, BP 60 149, 59653 Villeneuve d'Ascq cedex, France

INFO ARTICLE

Historique de l'article :
Reçu le 25 avril 2010
Accepté après révision le 31 mai 2011
Disponible sur Internet le 23 juillet 2011

Présenté par Paul Deheuvels

RÉSUMÉ

Dans cette Note, nous considérons le problème de l'estimation de la matrice de variance asymptotique du quasi-maximum de vraisemblance (QMV) des paramètres d'un modèle ARMA multivarié avec innovations linéaires non corrélées mais non nécessairement indépendantes (i.e. VARMA faible). Dans un premier temps, nous proposons une expression des dérivées des résidus en fonction des résidus passés et des paramètres du modèle VARMA. Ceci nous permettra ensuite de donner une expression explicite de la variance asymptotique de l'estimateur du QMV, en fonction des paramètres des polynômes VAR et MA, et des moments d'ordre deux et quatre du bruit. Enfin nous en déduisons un estimateur convergent de la matrice de variance asymptotique.

ABSTRACT

In this Note, we consider the problems of estimating the asymptotic variance of the quasi-maximum likelihood estimator (QMLE) of vector autoregressive moving-average (VARMA) models under the assumption that the errors are uncorrelated but not necessarily independent (i.e. weak VARMA). We first give expressions for the derivatives of the VARMA residuals in terms of the parameters of the models. Secondly we give an explicit expression of the asymptotic variance of the QMLE, in terms of the VAR and MA polynomials, and of the second- and fourth-order structure of the noise. We deduce a consistent estimator of the asymptotic variance of the QMLE.

© 2011 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Introduction

Dans cette Note, nous étudions l'estimation de la matrice de variance asymptotique de l'estimateur du QMV des paramètres d'un modèle VARMA faible de la forme :

$$A_{00}X_t - \sum_{i=1}^{p_0} A_{0i}X_{t-i} = B_{00}\epsilon_t - \sum_{j=1}^{q_0} B_{0i}\epsilon_{t-j}, \quad \forall t \in \mathbb{Z}, \tag{1}$$

Adresse e-mail : yacouba.boubacarmainassara@univ-lille3.fr.

1631-073X/$ – see front matter
doi:10.1016/j.crma.2011.05.015

où $X_t = (X_{1t}, \dots, X_{dt})'$ est un processus vectoriel stationnaire au second ordre de dimension d, à valeurs réelles. Le terme d'erreur $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{dt})'$ est un bruit blanc faible, c'est-à-dire une suite de variables aléatoires centrées ($E\epsilon_t = 0$), non corrélées, avec une matrice de covariance non singulière Σ_0. Nous relâchons l'hypothèse d'indépendance habituelle sur le bruit, contrairement à ce qui est fait habituellement pour l'inférence de ces modèles, ceci permet aux modèles VARMA faibles de couvrir une large classe de processus non linéaires.

Nous proposons une méthode d'estimation de la matrice de variance asymptotique qui consiste à trouver des estimateurs des matrices, impliquées dans la variance asymptotique de l'estimateur du QMV, dans lesquels, les estimateurs des coefficients du modèle VARMA sont isolés de ceux dépendants de la distribution de l'innovation.

2. Hypothèses et notations

Pour l'estimation des paramètres du modèle (1), nous utiliserons la méthode du QMV, qui est la méthode du maximum de vraisemblance gaussien lorsque l'hypothèse de bruit blanc gaussien est relâchée. Les paramètres sont les coefficients des matrices carrées d'ordre d suivantes : A_i, $i \in \{0, \dots, p_0\}$, B_j, $j \in \{0, \dots, q_0\}$. Dans cette Note, nous considérons la matrice Σ comme un paramètre de nuisance. Nous supposons que ces matrices sont paramétrées suivant le vecteur des vraies valeurs des paramètres noté θ_0. Nous notons $A_{0i} = A_i(\theta_0)$, $i \in \{0, \dots, p_0\}$, $B_j = B_j(\theta_0)$, $j \in \{0, \dots, q_0\}$, et $\Sigma_0 = \Sigma(\theta_0)$, où θ_0 appartient à l'espace compact des paramètres $\Theta \subset \mathbb{R}^{k_0}$, et k_0 est le nombre de paramètres inconnus qui est inférieur à $(p_0 + q_0 + 2)d^2$ (i.e. le nombre de paramètres sans aucune contrainte). Pour tout $\theta \in \Theta$, nous supposons aussi que les applications $\theta \mapsto A_i(\theta)$, $i = 0, \dots, p_0$, $\theta \mapsto B_j(\theta)$, $j = 0, \dots, q_0$, et $\theta \mapsto \Sigma(\theta)$ admettent des dérivées continues d'ordre 3. Nous ferons les hypothèses essentielles d'ergodicité et de mélange suivantes : **A1 :** Le processus (ϵ_t) est stationnaire et ergodique ; **A2 :** Il existe un réel $\nu > 0$ tel que $E\|\epsilon_t\|^{4+2\nu} < \infty$ et les coefficients de mélange fort du processus (ϵ_t), $\alpha_\epsilon(k) := \sup_{A \in \sigma(\epsilon_u, \leqslant t), B \in \sigma(\epsilon_u, \geqslant t+k)} |P(A \cap B) - P(A)P(B)|$, vérifient $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\nu}{2+\nu}} < \infty$.

Soit les polynômes $A_\theta(z) = A_0 - \sum_{i=1}^{p_0} A_i z^i$ et $B_\theta(z) = B_0 - \sum_{i=1}^{q_0} B_i z^i$. Nous ferons également les hypothèses **A3 :** $\det A(z) \det B(z) = 0 \Rightarrow |z| > 1$; **A4 :** Pour tout $\theta \in \Theta$ tel que $\theta \neq \theta_0$, soit les fonctions de transfert $A_0^{-1} B_0 B_\theta^{-1}(z) A_\theta(z) \neq A_{00}^{-1} B_{00} B_{\theta_0}^{-1}(z) A_{\theta_0}(z)$ pour un $z \in \mathbb{C}$ ou alors $A_0^{-1} B_0 \Sigma B_0' A_0^{-1'} \neq A_{00}^{-1} B_{00} \Sigma_0 B_{00}' A_{00}^{-1'}$; **A5 :** Nous avons $\theta_0 \in \overset{\circ}{\Theta}$, où $\overset{\circ}{\Theta}$ est l'intérieur du sous espace compact Θ de l'espace des paramètres.

Pour des raisons pratiques, nous écrivons le modèle (1) sous la forme réduite

$$X_t - \sum_{i=1}^{p_0} A_{00}^{-1} A_{0i} X_{t-i} = e_t - \sum_{j=1}^{q_0} A_{00}^{-1} B_{0j} B_{00}^{-1} A_{00} e_{t-j}, \quad e_t = A_{00}^{-1} B_{00} \epsilon_t.$$

En disposant d'observations $X_1, \dots, X_n$ de longueur n, pour $0 < t \leqslant n$ et pour tout $\theta \in \Theta$, nous définissons récursivement les variables aléatoires $\tilde{e}_t(\theta)$ par

$$\tilde{e}_t(\theta) = X_t - \sum_{i=1}^{p_0} A_0^{-1} A_i X_{t-i} + \sum_{j=1}^{q_0} A_0^{-1} B_j B_0^{-1} A_0 \tilde{e}_{t-j}(\theta),$$

où les valeurs initiales inconnues sont remplacées par zéro : $\tilde{e}_0(\theta) = \cdots = \tilde{e}_{1-q_0}(\theta) = X_0 = \cdots = X_{1-p_0} = 0$. Dans [2], noté BMF dans la suite, il est montré que ces valeurs initiales sont asymptotiquement négligeables et, en particulier, que $\tilde{e}_t(\theta_0) - e_t \to 0$ presque sûrement quand $t \to \infty$. La quasi-vraisemblance gaussienne s'écrit

$$\tilde{\mathrm{L}}_n(\theta) = \prod_{t=1}^{n} \frac{1}{(2\pi)^{d/2} \sqrt{\det \Sigma_e}} \exp\left\{ -\frac{1}{2} \tilde{e}_t'(\theta) \Sigma_e^{-1} \tilde{e}_t(\theta) \right\},$$

où $\Sigma_e = \Sigma_e(\theta) = A_0^{-1} B_0 \Sigma B_0' A_0^{-1'}$. Un estimateur du QMV de θ_0 est défini comme toute solution mesurable $\hat{\theta}_n$ de

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} \tilde{\mathrm{L}}_n(\theta) = \arg\min_{\theta \in \Theta} \tilde{\ell}_n(\theta), \quad \tilde{\ell}_n(\theta) = \frac{-2}{n} \log \tilde{\mathrm{L}}_n(\theta).$$

Sous les hypothèses **A1**–**A5**, la convergence forte et la normalité asymptotique de l'estimateur $\hat{\theta}_n$ du QMV des paramètres du modèle (1) ont été établies dans BMF. Pour la normalité asymptotique, ces auteurs ont obtenu

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}\big(0, \Omega := J^{-1} I J^{-1}\big), \quad \text{où } J = J(\theta_0) \quad \text{et } I = I(\theta_0),$$

$$\text{avec} \quad J(\theta) = \lim_{n \to \infty} \frac{\partial^2}{\partial\theta \partial\theta'} \tilde{\ell}_n(\theta) \text{ p.s.}, \qquad I(\theta) = \lim_{n \to \infty} \operatorname{Var} \frac{\partial}{\partial\theta} \tilde{\ell}_n(\theta).$$

Soit L l'opérateur retard. Posons $\mathbf{A}_\theta(L) = I_d - \sum_{i=1}^{p_0} \mathbf{A}_i L^i$ le polynôme VAR et $\mathbf{B}_\theta(L) = I_d - \sum_{i=1}^{q_0} \mathbf{B}_i L^i$ le polynôme MA, où $\mathbf{A}_i = A_0^{-1} A_i$ et $\mathbf{B}_i = A_0^{-1} B_i B_0^{-1} A_0$. Pour $i, j = 1, \dots, d$, définissons les $(d \times d)$ opérateurs matriciels $M_{ij}(L) := \mathbf{B}_\theta^{-1}(L) E_{ij} \mathbf{A}_\theta^{-1}(L) \mathbf{B}_\theta(L)$ et $N_{ij}(L) := \mathbf{B}_\theta^{-1}(L) E_{ij}$, où E_{ij} est une matrice $d \times d$ prenant la valeur 1 à la position (i, j) et 0 ailleurs. Notons $\mathbf{A}_{ij,h}^*$ et $\mathbf{B}_{ij,h}^*$ les matrices $d \times d$ définies par

$$M_{ij}(z)=\sum_{h=0}^{\infty}\mathbf{A}^*_{ij,h}z^h,\qquad N_{ij}(z)=\sum_{h=0}^{\infty}\mathbf{B}^*_{ij,h}z^h,\quad |z|\leqslant 1 \text{ pour } h\geqslant 0.$$

Par convention $\mathbf{A}^*_{ij,h}=\mathbf{B}^*_{ij,h}=0$ quand $h<0$. Soit $\mathbf{A}^*_h=[\mathbf{A}^*_{11,h}:\mathbf{A}^*_{21,h}:\cdots:\mathbf{A}^*_{dd,h}]$ et $\mathbf{B}^*_h=[\mathbf{B}^*_{11,h}:\mathbf{B}^*_{21,h}:\cdots:\mathbf{B}^*_{dd,h}]$ les matrices de tailles $d\times d^3$. Posons la matrice de taille $d\times d^3(p_0+q_0)$

$$\lambda_h(\theta)=\left[-\mathbf{A}^*_{h-1}:\cdots:-\mathbf{A}^*_{h-p_0}:\mathbf{B}^*_{h-1}:\cdots:\mathbf{B}^*_{h-q_0}\right],\tag{2}$$

constituée uniquement des paramètres des polynômes VAR et MA. Cette matrice est bien définie car les coefficients $\mathbf{A}_\theta^{-1}$ et $\mathbf{B}_\theta^{-1}$ décroissent exponentiellement vers zéro.

3. Dérivées des résidus du modèle VARMA

Dans le cas de modèles ARMA univariés, [4] avait défini des dérivées résiduelles par

$$\frac{\partial e_t}{\partial\phi_i}=\upsilon_{t-i},\quad i=1,\dots,p_0\quad\text{et}\quad\frac{\partial e_t}{\partial\beta_j}=u_{t-j},\quad j=1,\dots,q_0,$$

où ϕ_i et β_j sont respectivement les paramètres AR et MA. Définissons les polynômes $\phi_\theta(L)=1-\sum_{i=1}^{p_0}\phi_iL^i$ et $\varphi_\theta(L)=1-\sum_{i=1}^{q_0}\varphi_iL^i$. Notons ϕ^*_h et φ^*_h les coefficients définis par

$$\phi_\theta^{-1}(z)=\sum_{h=0}^{\infty}\phi^*_hz^h,\qquad\varphi_\theta^{-1}(z)=\sum_{h=0}^{\infty}\varphi^*_hz^h,\quad|z|\leqslant 1$$

pour $h\geqslant 0$. Posons $\theta=(\theta_1,\dots,\theta_{p_0},\theta_{p_0+1},\dots,\theta_{p_0+q_0})'$, pour p_0 et q_0 différents de 0. Alors, on peut facilement représenter les dérivées résiduelles univariées par

$$\frac{\partial e_t(\theta)}{\partial\theta}=\big(\upsilon_{t-1}(\theta),\dots,\upsilon_{t-p_0}(\theta),u_{t-1}(\theta),\dots,u_{t-q_0}(\theta)\big)',\quad\text{où}$$

$$\upsilon_t(\theta)=-\phi_\theta^{-1}(L)e_t(\theta)=-\sum_{h=0}^{\infty}\phi^*_he_{t-h}(\theta),\qquad u_t(\theta)=\varphi_\theta^{-1}(L)e_t(\theta)=\sum_{h=0}^{\infty}\varphi^*_he_{t-h}(\theta)$$

avec l'innovation $e_t(\theta)=\varphi_\theta^{-1}(L)\phi_\theta(L)X_t$. Nous énonçons la proposition suivante qui donne une extension au cas multivarié des expressions de ces dérivées résiduelles.

Proposition 1. *Sous certaines hypothèses* **A1–A5**, *nous avons*

$$\frac{\partial e_t(\theta)}{\partial\theta'}=\left[V_{t-1}(\theta):\cdots:V_{t-p_0}(\theta):U_{t-1}(\theta):\cdots:U_{t-q_0}(\theta)\right],\quad\textit{où}$$

$$V_t(\theta)=-\sum_{h=0}^{\infty}\mathbf{A}^*_h\big(I_{d^2}\otimes e_{t-h}(\theta)\big)\quad\textit{et}\quad U_t(\theta)=\sum_{h=0}^{\infty}\mathbf{B}^*_h\big(I_{d^2}\otimes e_{t-h}(\theta)\big).$$

De plus, pour $\theta=\theta_0$ nous avons : $\partial e_t/\partial\theta'=\sum_{h\geqslant 1}\lambda_h(I_{d^2(p_0+q_0)}\otimes e_{t-h})$, où les matrices λ_h sont définies par (2).

4. Expressions explicites des matrices *J* et *I*

Nous donnons des expressions explicites des matrices I et J impliquées dans la variance asymptotique Ω de l'estimateur du QMV. Dans ces expressions, nous isolons tout ce qui est fonction du paramètre θ_0 du modèle VARMA, de ce qui est fonction de la distribution du bruit blanc faible e_t. Nous définissons vec$(\cdot)$ l'opérateur qui consiste à empiler en un vecteur les colonnes d'une matrice en partant de la première colonne jusqu'à la dernière. Le produit de Kronecker de deux matrices A et B est noté par $A\otimes B$ (noté dans la suite $A^{\otimes 2}$ quand $A=B$). Nous notons par

$$\mathcal{M}:=E\big\{\big(I_{d^2(p_0+q_0)}\otimes e'_t\big)^{\otimes 2}\big\}$$

la matrice impliquant les moments d'ordre deux de l'innovation (e_t). Maintenant nous donnons l'expression de $J=J(\theta_0,\Sigma_{e0})$, où $\Sigma_{e0}=\Sigma_e(\theta_0)=A_{00}^{-1}B_{00}\Sigma_0B'_{00}A_{00}^{-1'}$, dans laquelle les termes dépendants de θ_0 (via les matrices λ_h) sont distingués de ceux dépendants des moments d'ordre deux du processus (e_t) (via la matrice $\mathcal{M}$) et des termes dépendants de la variance de l'innovation (via la matrice Σ_{e0}).

Proposition 2. *Sous les hypothèses de la proposition* 1, *nous avons* vec $J=2\sum_{h\geqslant 1}\mathcal{M}\{\lambda'_h\otimes\lambda'_h\}$ vec Σ_{e0}^{-1}, *où les matrices λ_h sont définies par* (2).

Comme pour la matrice J, nous décomposons I en termes impliquant le paramètre θ_0 du modèle VARMA et ceux impliquant la distribution des innovations e_t. Rappelons que

$$I = \lim_{n\to\infty} \operatorname{Var} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Upsilon_t = \sum_{h=-\infty}^{+\infty} \operatorname{Cov}(\Upsilon_t, \Upsilon_{t-h}), \tag{3}$$

où $\Upsilon_t = \frac{\partial}{\partial\theta}\{\log\det \Sigma_e + e_t'(\theta)\Sigma_e^{-1}e_t(\theta)\}_{\theta=\theta_0}$. L'existence de la somme des covariances dans (3) est une conséquence de l'hypothèse **A2** et de l'inégalité de covariance de [3]. Soit les matrices

$$\mathcal{M}_{ij,h} := E\big(\big\{e_{t-h}' \otimes \big(I_{d^2(p_0+q_0)} \otimes e_{t-j-h}'\big)\big\} \otimes \big\{e_t' \otimes \big(I_{d^2(p_0+q_0)} \otimes e_{t-i}'\big)\big\}\big).$$

Les termes dépendants des paramètres du modèle VARMA sont les matrices λ_h définies par (2). Les matrices $\boldsymbol{\Gamma}(i,j) := \sum_{h=-\infty}^{+\infty} \mathcal{M}_{ij,h}$ mettent en jeux les moments d'ordre quatre du bruit blanc faible e_t. Les termes dépendants de la variance de l'innovation sont dans la matrice Σ_{e0}. Notons I_r la matrice identité d'ordre r. Nous énonçons maintenant la proposition suivante pour la matrice $I = I(\theta_0, \Sigma_{e0})$ qui est similaire à la proposition 2.

Proposition 3. *Sous les hypothèses de la proposition* 2*, nous avons*

$$\operatorname{vec} I = 4\sum_{i,j=1}^{+\infty} \boldsymbol{\Gamma}(i,j)\big(\big\{I_d \otimes \lambda_j'\big\} \otimes \big\{I_d \otimes \lambda_i'\big\}\big) \operatorname{vec}\big(\operatorname{vec} \Sigma_{e0}^{-1}\big\{\operatorname{vec} \Sigma_{e0}^{-1}\big\}'\big).$$

5. Estimation de la matrice de variance asymptotique

Étant donné que nous disposons des expressions explicites de I et J, nous nous intéressons à l'estimation de ces matrices. Soit $\hat{e}_t = \tilde{e}_t(\hat{\theta}_n)$ les résidus de l'estimateur du QMV quand $p_0 + q_0 \neq 0$. En vue de la proposition 2, nous définissons un estimateur $\hat{J}_n$ de J par $\operatorname{vec}\hat{J}_n = \sum_{h\geqslant 1} \hat{\mathcal{M}}_n\{\hat{\lambda}_h' \otimes \hat{\lambda}_h'\} \operatorname{vec}\hat{\Sigma}_{e0}^{-1}$, où $\hat{\mathcal{M}}_n = \frac{1}{n}\sum_{t=1}^{n}\{(I_{d^2(p+q)} \otimes \hat{e}_t')^{\otimes 2}\}$. Nous énonçons le théorème suivant qui montre la consistance forte de $\hat{J}_n$.

Théorème 5.1 *(Convergence forte de $\hat{J}_n$). Sous les hypothèses de la proposition* 3*, nous avons $\hat{J}_n \to J$ p.s. quand $n \to \infty$.*

Dans le cas de modèles VARMA forts standard, i.e. quand l'hypothèse d'ergocité **A1** est remplacée par celle que les termes d'erreur (ϵ_t) sont iid, nous avons $I = 2J$, ainsi $\hat{\Omega} := 2\hat{J}^{-1}$ est un estimateur fortement convergent de $\Omega = 2J^{-1}$. Dans le cadre général de modèles VARMA faibles, cet estimateur n'est pas convergent dès que $I \neq 2J$. Dans ce cas, nous avons besoin d'un estimateur de I. Cette matrice I est délicate à estimer car son expression explicite fait intervenir une infinité de moments d'ordre 4 (via $\boldsymbol{\Gamma}(i,j)$). Soit $\mathcal{M}_{nij,h} := n^{-1}\sum_{t=1}^{n-|h|}(\{e_{t-h}' \otimes (I_{d^2(p_0+q_0)} \otimes e_{t-j-h}')\} \otimes \{e_t' \otimes (I_{d^2(p_0+q_0)} \otimes e_{t-i}')\})$. Pour estimer $\boldsymbol{\Gamma}(i,j)$, nous allons pondérer convenablement certains de ces moments empiriques (voir [1], voir aussi [5]). Cette pondération se fait au moyen d'une fonction de poids (ou fenêtre) et d'un paramètre de troncature. Nous considérons une suite de réels $(b_n)_{n\in\mathbb{N}^*}$ telle que $b_n \to 0$ et $nb_n^{(10+4\nu)/\nu} \to \infty$ quand $n \to \infty$, et une fonction de poids $f : \mathbb{R} \to \mathbb{R}$ bornée, à support compact $[-a, a]$ et continue à l'origine avec $f(0) = 1$. Nous avons aussi $b_n \sum_{|h|<n} |f(hb_n)| = O(1)$. Soit $\hat{\mathcal{M}}_{nij,h} := n^{-1}\sum_{t=1}^{n-|h|}(\{\hat{e}_{t-h}' \otimes (I_{d^2(p_0+q_0)} \otimes \hat{e}_{t-j-h}')\} \otimes \{\hat{e}_t' \otimes (I_{d^2(p_0+q_0)} \otimes \hat{e}_{t-i}')\})$. Nous considérons la matrice $\hat{\boldsymbol{\Gamma}}_n(i,j) := \sum_{h=-T_n}^{+T_n} f(hb_n)\hat{\mathcal{M}}_{nij,h}$ et $T_n = [a/b_n]$, où $[x]$ désigne la partie entière de x. En vue de la proposition 3, nous définissons un estimateur $\hat{I}_n$ de I par $\operatorname{vec}\hat{I}_n = 4\sum_{i,j=1}^{+\infty} \hat{\boldsymbol{\Gamma}}_n(i,j)(\{I_d \otimes \hat{\lambda}_i'\} \otimes \{I_d \otimes \hat{\lambda}_j'\}) \operatorname{vec}(\operatorname{vec} \hat{\Sigma}_{e0}^{-1}\{\operatorname{vec} \hat{\Sigma}_{e0}^{-1}\}')$. Maintenant nous énonçons le théorème suivant qui montre la consistance faible de l'estimateur empirique $\hat{I}_n$.

Théorème 5.2. *Sous les hypothèses du théorème* 5.1*, nous avons*

$$\hat{I}_n \to I \textit{ en probabilité}, \quad \textit{quand } n \to \infty.$$

Les théorèmes 5.1 et 5.2 montrent alors que $\hat{\Omega}_n := \hat{J}_n^{-1}\hat{I}_n\hat{J}_n^{-1}$ est un estimateur faiblement convergent de la matrice de covariance asymptotique $\Omega = J^{-1}IJ^{-1}$.

Références

[1] D.W.K. Andrews, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59 (1991) 817–858.
[2] Y. Boubacar Mainassara, C. Francq, Estimating structural VARMA models with uncorrelated but non-independent error terms, Journal of Multivariate Analysis 102 (2011) 496–505.
[3] Y.A. Davydov, Convergence of distributions generated by stationary stochastic processes, Theory of Probability and Applications 13 (1968) 691–696.
[4] A.I. McLeod, On the distribution of residual autocorrelations in Box–Jenkins models, Journal of the Royal Statistical Society B 40 (1978) 296–302.
[5] W.K. Newey, K.D. West, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55 (1987) 703–708.