Spécification de modèles ARIMA par la méthode des autocorrélations généralisées *

Jean-Marie Dufour †
Université de Montréal

Première version: Mars 1991
Révisions: Février 2000
Cette version: 13 mars 2002
Compilé: 13 mars 2002, 11:30am

* Cette recherche a bénéficié du support financier de la Chaire de recherche du Canada en économétrie, du Conseil des Arts du Canada (Bourse Killam), du Conseil de recherche en sciences humaines du Canada, du Conseil de recherche en sciences naturelles et en génie du Canada, du Réseau canadien de centres d'excellence (projet MITACS) et du Fonds FCAR du Québec.

† L'auteur est titulaire de la Chaire de recherche du Canada en économétrie. Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre de recherche et développement en économique (C.R.D.E.), et Département de sciences économiques, Université de Montréal. Adresse postale: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre Ville, Montréal, Québec, Canada H3C 3J7. TEL: (514) 343 2400; FAX: (514) 343 5831; courriel: jean.marie.dufour@umontreal.ca. Page Web: http://www.fas.umontreal.ca/SCECO/Dufour .

Table des matières

Liste de définitions, propositions et théorèmes

1. Modèle

Supposons qu'un processus X_t satisfait une équation de la forme

$$\varphi_p(B)\,U_d(B)\,X_t = \bar{\mu} + \theta_q(B)\,u_t\ ,\ t \geq t_0 \tag{1.1}$$

où

$$\begin{aligned}
\varphi_p(B) &= 1 - \varphi_1 B - \cdots - \varphi_p B^p\ ,\ \varphi_p \neq 0 \text{ si } p > 0, \\
\theta_q(B) &= 1 - \theta_1 B - \cdots - \theta_q B^q\ ,\ \theta_q \neq 0 \text{ si } q > 0, \\
U_d(B) &= 1 - U_1 B - \cdots - U_d B^d\ ,\ U_d \neq 0 \text{ si } d > 0, \\
(u_t : \in \mathbb{Z}) &\sim IID\left(0, \sigma^2\right)\ , \sigma^2 > 0, \\
E\left(u_t^4\right) &< \infty, \\
\varPhi_{\bar{p}}(B) &\equiv \varphi_p(B)\,U_d(B) \\
&= 1 - \varPhi_1 B - \cdots - \varPhi_{\bar{p}} B^{\bar{p}}, \\
\bar{p} &= p + d\ ,\ t_0 \leq 1.
\end{aligned}$$

De plus, on suppose que :

$$\begin{gathered}\text{tous les racines du polynôme } \varphi_p(B) \\ \text{sont à l'extérieur du cercle unité ;}\end{gathered} \tag{1.2}$$

$$\begin{gathered}\text{toutes les racines du polynôme } U_d(B) \\ \text{sont sur le cercle unité (lorsque } d > 0) \text{ ;}\end{gathered} \tag{1.3}$$

$$\varPhi_{\bar{p}}(B) \text{ et } \theta_q(B) \text{ n'ont aucun facteur en commun ;} \tag{1.4}$$

$$\bar{\mu} = 0. \tag{1.5}$$

2. Méthode

Dans un premier temps, considérons la régression :

$$X_t = \sum_{\ell=1}^{\bar{p}} \varPhi_{\ell(\bar{p})}^{(0)} X_{t-\ell} + e_{\bar{p},t}^{(0)}, t = \bar{p} + 1,\ \ldots\ ,\ n \tag{2.1}$$

que l'on estime par les moindres carrés ordinaires (MCO). Les estimateurs MCO des coefficients sont $\hat{\varPhi}_{\ell(\bar{p})}^{(0)}, \ell = 1,\ \ldots\ ,\ \bar{p}$. Si X_t est un processus

a) autorégressif pur $(q = 0)$, ou

b) ARMA purement non stationnaire $(p = 0, d \geq 1)$,

les estimateurs MCO sont convergents :

$$\hat{\Phi}^{(0)}_{\ell(\bar{p})} \xrightarrow[n\to\infty]{\mathsf{p}} \Phi_\ell\ ,\ \ell = 1,\ \dots\ ,\ \bar{p}. \tag{2.2}$$

Dans le cas contraire, les estimateurs MCO ne sont pas convergents et les résidus

$$\hat{e}^{(0)}_{\hat{p},t} = X_t - \sum_{\ell=1}^{\bar{p}} \hat{\Phi}^{(0)}_{\ell(\bar{p})} X_{t-\ell}\ ,\ t = \bar{p} + 1,\ \dots\ ,\ n \tag{2.3}$$

ne sont pas des bruits blancs (BB) même asymptotiquement. Les valeurs passées $\hat{e}^{(0)}_{\bar{p},t-j}, j \geq 1$ contiennent de l'information pouvant servir à prédire X_t. Dans un second temps, considérons la régression étendue :

$$X_t = \sum_{\ell=1}^{\bar{p}} \Phi^{(1)}_{\ell(\bar{p})} X_{t-\ell} + \beta^{(1)}_{\ell(\bar{p})} \hat{e}^{(0)}_{\bar{p},t-1} + e^{(1)}_{\bar{p},t}\ ,\ t = \bar{p} + 2,\ \dots\ ,\ n. \tag{2.4}$$

Si

a) $q \leq 1$, ou

b) $p = 0$ et $d \geq 1$,

les estimateurs MCO des coefficients $\Phi^{(1)}_{\ell(\bar{p})}, \ell = 1,\ \dots\ ,\ \bar{p}$, sont convergents :

$$\hat{\Phi}^{(1)}_{\ell(\bar{p})} \xrightarrow[n\to\infty]{\mathsf{p}} \Phi_\ell\ ,\ \ell = 1,\ \dots\ ,\ \bar{p}. \tag{2.5}$$

Dans le cas contraire, ces estimateurs ne sont pas convergents et les résidus

$$\hat{e}^{(1)}_{\bar{p},t} = X_t - \sum_{\ell=1}^{\bar{p}} \hat{\Phi}^{(1)}_{\ell(\bar{p})} X_{t-\ell} - \hat{\beta}^{(0)}_{\ell(\bar{p})} \hat{e}^{(0)}_{\bar{p},t-1}\ ,\ t = \bar{p} + 2,\ \dots\ ,\ n \tag{2.6}$$

ne sont pas des BB même asymptotiquement. Les valeurs passées $\hat{e}^{(1)}_{\bar{p},t-j}, j \geq 1$, contiennent encore de l'information pouvant servir à prédire X_t.

Troisièmement, on considère une seconde régression étendue,

$$X_t = \sum_{\ell=1}^{\bar{p}} \Phi^{(2)}_{\ell(\bar{p})} X_{t-\ell} + \beta^{(2)}_{1(\bar{p})} \hat{e}^{(1)}_{\bar{p},t-1} + \beta^{(2)}_{2(\bar{p})} \hat{e}^{(0)}_{\bar{p},t-2} + e^{(2)}_{\bar{p},t}\ ,\ t = \bar{p} + 3,\ \dots\ ,\ n. \tag{2.7}$$

Si

a) $q \leq 2$, ou

b) $p = 0$ et $d \geq 1$,

les estimateurs MCO des coefficients $\Phi^{(2)}_{\ell(\bar{p})}$, $\ell = 1, \ldots, \bar{p}$, sont convergents :

$$\hat{\Phi}^{(2)}_{\ell(\bar{p})} \xrightarrow[n\to\infty]{\mathsf{p}} \Phi_\ell \,, \; \ell = 1, \ldots, \bar{p}. \tag{2.8}$$

Dans le cas contraire les résidus MCO $\hat{e}^{(2)}_{\bar{p},t}$ ne sont pas des BB même asymptotiquement. On peut, bien sûr, poursuivre ce procédé : après $q+1$ itérations, on obtient des estimateurs convergents de $\Phi_1, \ldots, \Phi_{\bar{p}}$.

En pratique, les valeurs de $\bar{p}$ et q sont inconnues. Considérons la suite de régressions :

$$X_t = \sum_{\ell=1}^{k} \Phi^{(j)}_{\ell(k)} X_{t-\ell} + \sum_{i=1}^{j} \beta^{(j)}_{i(k)} \hat{e}^{(j-i)}_{k,t-i} + e^{(j)}_{k,t} \,, \; t = k+j+1, \ldots, n \,, \tag{2.9}$$

pour $k = 1, 2, \ldots$ et $j = 0, 1, \ldots,$ où

$$\hat{e}^{(j)}_{k,t} = X_t - \sum_{\ell=1}^{k} \hat{\Phi}^{(j)}_{\ell(k)} X_{t-\ell} - \sum_{h=1}^{j} \hat{\beta}^{(j)}_{h(k)} \hat{e}^{(j-h)}_{k,t-h} \,. \tag{2.10}$$

Les coefficients $\Phi^{(j)}_{\ell(k)}$ peuvent être calculés de façon récursive :

$$\hat{\Phi}^{(j)}_{\ell(k)} = \hat{\Phi}^{(j-1)}_{\ell(k+1)} - \left[\hat{\Phi}^{(j-1)}_{\ell-1(k)} \hat{\Phi}^{(j-1)}_{k+1(k+1)} / \hat{\Phi}^{(j-1)}_{k(k)} \right] . \tag{2.11}$$

2.1 Théorème RYTHME DE CONVERGENCE DES ESTIMATEURS MCO D'UNE AUTORÉGRESSION. *Soit X_t un processus non stationnaire qui satisfait les hypothèses* (1.1) *à* (1.5)*, où $d \geq 1$. Alors*

$$\hat{\Phi}^{(j)}_{\ell(d)} = U_\ell + O_p\big(n^{-1}\big) \,, \; \ell = 1, \ldots, d \,, \; \forall j \geq 0 \,,$$

où U_ℓ est une variable aléatoire.

DÉMONSTRATION Voir Tsay and Tiao (1984, Theorem 2.1, sections 5.3, 7) et Tiao and Tsay (1983, Theorem 3.2). ∎

2.2 Théorème CONVERGENCE DES ESTIMATEURS MCO D'UNE AUTORÉGRESSION. *Soit X_t un processus qui satisfait les hypothèses* (1.1) *à* (1.5)*, stationnaire ou non station-*

naire. Si $k \geq \bar{p}$ et $j = q$, ou $k = \bar{p}$ et $j > q$, alors

$$\hat{\varPhi}^{(j)}_{\ell(k)} = \varPhi_\ell + O_p\big(n^{-1/2}\big)\ ,\ \ell = 1,\ \dots\ ,\ k\,.$$

De plus, si $p = 0$ et $k = d$, on peut remplacer $O_p\left(n^{-1/2}\right)$ par $O_p\left(n^{-1}\right)$.

DÉMONSTRATION Voir Tsay and Tiao (1984, Theorem 2.2, sections 5.3, 7) et Tiao and Tsay (1983, Theorem 3.2). ∎

2.3 Remarque Si $k > \bar{p}$ et $j \neq q$, $\hat{\varPhi}^{(j)}_{\ell(k)}$ peut ne pas être convergent.

Définissons

$$W^{(j)}_{k,t} = X_t - \sum_{\ell=1}^{k} \hat{\varPhi}^{j}_{\ell(k)} X_{t-\ell}\ ,\ t = k + j + 1,\ \dots\ ,\ n. \tag{2.12}$$

Lorsque $k = \bar{p}$ et les estimateurs $\hat{\varPhi}^{j}_{\ell(k)}$, $\ell = 1,\ \dots\ ,\ k$, sont convergents,

$$W^{(j)}_{k,t} \xrightarrow[n\to\infty]{\mathsf{p}} \theta_q\left(B\right) u_t\ ,\ t = k + j + 1,\ \dots\ ,\ n. \tag{2.13}$$

Pour n grand, $W^{(j)}_{k,t}$ suit un processus $MA\left(q\right)$. Pour déterminer q, on peut donc examiner les autocorrélations de $W^{(j)}_{k,t}$.

2.4 Définition AUTOCORRÉLATIONS GÉNÉRALISÉES. *On appelle autocorrélation généralisée échantillonnale d'ordre (k, j) l'autocorrélation échantillonnale de délai j de la série $W^{(j)}_{k,t}$, $t = k + j + 1,\ \dots\ ,\ n$, i.e.*

$$\begin{aligned} r^{(k)}_j &= r_j\big(W^{(j)}_k\big) \\ &= \sum_{t=k+j+1}^{n-j} W^{(j)}_{k,t} W^{(j)}_{k,t+j} \Big/ \sum_{t=k+j+1}^{n} \big(W^{(j)}_{k,t}\big)^2. \end{aligned}$$

À cause du Théorème 2 et de la propriété (2.13), on voit alors aisément que

$$\begin{aligned} r^{(k)}_j &\xrightarrow[n\to\infty]{\mathsf{p}} 0\,, && \text{si } k = \bar{p} \text{ et } j > q\,, \\ &\xrightarrow[n\to\infty]{\mathsf{p}} \rho^{(k)}_j \neq 0\,, && \text{si } k = \bar{p} \text{ et } j = q. \end{aligned} \tag{2.14}$$

De façon plus générale,

$$r_j^{(k)} \begin{array}{ll} \xrightarrow[n\to\infty]{\mathsf{p}} 0\,, & \text{si } j - q > k - \bar{p} \geq 0\,, \\ \xrightarrow[n\to\infty]{\mathsf{p}} C\left(k - \bar{p},\, j - q\right)\,, & \text{autrement,} \end{array} \tag{2.15}$$

où $C\left(k - \bar{p}, j - q\right)$ est une constante différente de zéro ou une variable aléatoire non dégénérée à zéro.

On peut calculer des écart-types en se servant de la formule de Bartlett pour un processus $MA\left(q\right)$ ou encore en utilisant la formule $\left(n - k - j\right)^{-1/2}$. Cette dernière tend à rejeter trop souvent.

TABLEAU DES AUTOCORRÉLATIONS GÉNÉRALISÉES

	MA				
AR	0	1	2	3	$\cdots$
0	$r_1^{(0)}$	$r_2^{(0)}$	$r_3^{(0)}$	$r_4^{(0)}$	$\cdots$
1	$r_1^{(1)}$	$r_2^{(1)}$	$r_3^{(1)}$	$r_4^{(1)}$	$\cdots$
2	$r_1^{(2)}$	$r_2^{(2)}$	$r_3^{(2)}$	$r_4^{(2)}$	$\cdots$
3	$r_1^{(3)}$	$r_2^{(3)}$	$r_3^{(3)}$	$r_4^{(3)}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Pour un processus avec $\bar{p} = 1$ et $q = 2$, on devrait trouver un triangle commençant à (1,2).

	MA						
AR	0	1	2	3	4	5	$\cdots$
0	X	X	X	X	X	X	$\cdots$
1	X	X	0	0	0	0	$\cdots$
2	X	X	X	0	0	0	$\cdots$
3	X	X	X	X	0	0	$\cdots$
4	X	X	X	X	X	0	$\cdots$
5	X	X	X	X	X	X	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Références

TIAO, G. C., AND R. S. TSAY (1983) : “Consistency Properties of Least Squares Estimates of Autoregressive Parameters in ARMA Models,” *The Annals of Statistics*, 11, 856–871.

TSAY, R. S., AND G. C. TIAO (1984) : “Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models,” *Journal of the American Statistical Association*, 19, 84–96.