Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 339 (2004) 55–58
COMPTES RENDUS MATHEMATIQUE

Statistique/Probabilités

# Estimation de modèles ARMA à changements de régime récurrents

Christian Francq, Antony Gautier

*Université Lille 3, GREMARS, BP 149, 59653 Villeneuve d'Ascq cedex, France*

Reçu le 20 mai 2003 ; accepté après révision le 15 avril 2004

Disponible sur Internet le 28 mai 2004

Présenté par Paul Deheuvels

**Résumé**

Cette Note considère l'estimation de modèles ARMA à coefficients dépendant du temps. Nous étudions des modèles à changements de régime récurrents mais non-périodiques. Nous donnons des conditions pour la convergence et la normalité asymptotique de deux suites d'estimateurs des moindres carrés. Ces conditions sont rendues explicites pour des modèles à changements de régime Markoviens. ***Pour citer cet article : C. Francq, A. Gautier, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***


**Abstract**

**Estimating ARMA models with recurrent regime changes.** This Note considers estimation of time-varying ARMA models. We focus on models with recurrent but non-periodic changes in regime. Conditions ensuring the consistency and the asymptotic normality of two sequences of least squares estimators are given. These conditions are made explicit when the regime generated process is a Markov chain. ***To cite this article: C. Francq, A. Gautier, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***


## 1. Introduction

Cette Note s'intéresse à l'estimation des paramètres de modèles ARMA à coefficients récurrents mais non-périodiques et fait suite aux travaux de Bibi et Francq [4] ou Azrak et Mélard [2]. Nous obtenons des conditions plus fines de convergence et de normalité asymptotique des estimateurs des moindres carrés (MC) et des moindres carrés quasi-généralisés (MCQG).

## 2. Définitions et hypothèses

On considère une série chronologique $(X_t)_{t=1,2,\ldots}$ présentant des changements de régime à des instants connus. Supposons qu'il existe $d$ régimes qui apparaissent de manière récurrente mais non-périodique et notons $s_t$ le régime

*Adresses e-mail :* francq@univ-lille3.fr (C. Francq), antony.gautier@univ-lille3.fr (A. Gautier).

1631-073X/$ – see front matter 
doi:10.1016/j.crma.2004.04.014

à la date $t$ ; d'où $s_t = k$ quand la série se trouve dans le $k$-ième régime à l'instant $t$, pour $k \in \{1,\dots,d\}$. La suite $(s_t)$ est observée et est la réalisation d'un processus stochastique $(S_t)$. En effet, nous verrons que le comportement asymptotique des estimateurs des MC et des MCQG dépend non seulement des fréquences relatives de chaque régime, mais aussi de la loi de $(S_t)$.

La dynamique de $X_t$ dans chaque régime est décrite par une équation ARMA$(p,q)$

$$X_t - m(s_t) + \sum_{i=1}^{p} a_i(s_t)\bigl\{X_{t-i} - m(s_{t-i})\bigr\} = \epsilon_t + \sum_{i=1}^{q} b_i(s_t)\epsilon_{t-i}, \quad t = 1, 2, \dots, \tag{1}$$

où $p$ et $q$ sont entiers, les vecteurs $\underline{m} = \{m(1),\dots,m(d)\} \in \mathbb{R}^d$, $\underline{a} = \{a_1(1),\dots,a_p(d)\} \in \mathbb{R}^{pd}$ et $\underline{b} = \{b_1(1),\dots,b_q(d)\} \in \mathbb{R}^{qd}$ sont des paramètres inconnus et $X_t - m(s_t) = \epsilon_t = 0$ pour $t = 1 - \max\{p,q\},\dots,0$. La suite $(\epsilon_t)$ est de la forme $\epsilon_t = \sigma(s_t)\eta_t$ où $(\eta_t)$ est une suite de variables aléatoires i.i.d. avec $E\eta_t = E\eta_t^3 = 0$, $E\eta_t^2 = 1$ et $E\eta_t^4 < \infty$. Nous cherchons à estimer le paramètre d'intérêt $\theta = (\underline{m}, \underline{a}, \underline{b})' = [\theta(1),\dots,\theta(\{p+q+1\}d)]'$ et le paramètre de nuisance $\underline{\sigma} = \{\sigma(1),\dots,\sigma(d)\}'$ dont les vraies valeurs sont respectivement notées $\theta_0$ et $\underline{\sigma}_0$. Le vecteur $\theta$ est un élément d'un sous-ensemble ouvert $\Theta$ de $\mathbb{R}^{(p+q+1)d}$ dont $\Theta^*$ est un compact qui contient $\theta_0$. Par ailleurs, toutes les composantes de $\underline{\sigma}$ sont strictement positives et $\underline{\sigma}$ appartient à un sous-ensemble ouvert $\Phi$ de $\mathbb{R}^d$. Toutes les variables aléatoires apparaissant dans cette note sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

La meilleure prévision à horizon 1 au sens des MC, $\widehat{X}_t := E_\theta(X_t | X_{t-1},\dots,X_1)$, est facilement déterminée de façon récurrente en utilisant (1). Les résidus, définis par $e_t(\theta) = X_t - \widehat{X}_t$, peuvent être écrits à l'aide de la décomposition de Wold–Cramér pour les processus non-stationnaires

$$e_t(\theta) = \sum_{i=0}^{t-1} \psi_{t,i}(\theta,\theta_0)\epsilon_{t-i} + c_t(\theta,\theta_0), \quad c_t(\theta,\theta_0) = \sum_{i=0}^{t-1} \varphi_{t,i}(\theta, s_t,\dots,s_{t-i+1})\bigl\{m_0(s_{t-i}) - m(s_{t-i})\bigr\}. \tag{2}$$

**Remarque 1.** Contrairement à $\{\psi_{t,i}(\theta,\theta_0)\}_t = \{\psi_i(\theta,\theta_0,S_t,\dots,S_{t-i+1})\}_t$, le processus $\{c_t(\theta,\theta_0)\}_t = \{c(\theta,\theta_0,S_t,\dots,S_1)\}_t$ n'est pas stationnaire. Ceci entraîne de délicates manipulations techniques qui n'existent pas dans Bibi et Francq [4] où $m \equiv 0$.

Pour toute fonction différentiable $f_t(\theta)$, on note $f_t^{(k_1,\dots,k_j)}(\theta) = \partial^j f_t(\theta)/\partial\theta(k_1)\cdots\partial\theta(k_j)$ où $k_1,\dots,k_j \in \{1,\dots,(p+q+1)d\}$. On définit $\tilde{c}(s_t, s_{t-1},\dots) = \sum_{i=0}^{\infty} \sup_{\theta\in\Theta^*} |\varphi_{t,i}(\theta, s_t,\dots,s_{t-i+1})|$ et $\tilde{c}^{(k_1,\dots,k_j)}(s_t,s_{t-1},\dots) = \sum_{i=0}^{\infty} \sup_{\theta\in\Theta^*} |\varphi_{t,i}^{(k_1,\dots,k_j)}(\theta, s_t,\dots,s_{t-i+1})|$. On note $E_S Y$ l'espérance d'une variable aléatoire $Y$ mesurable par rapport à la tribu engendrée par $S = (S_t)$.

Pour montrer la convergence forte de nos estimateurs, on considère les hypothèses suivantes.

**H1** Le processus $S = (S_t)$ est stationnaire, ergodique et indépendant de $\eta = (\eta_t)$.
**H2** Il existe $\rho \in [0,1[$ et $\kappa > 0$ tels que $E_S \sup_{\theta\in\Theta^*}\{\psi_i(\theta,\theta_0,S_t,\dots,S_{t-i+1})\}^2 \leqslant \kappa\rho^i$, $\forall i$.
**H3** Nous avons $E_S\{\tilde{c}(S_t,S_{t-1},\dots)\}^2 < \infty$.
**H4** Il existe $\rho \in [0,1[$ et $\kappa > 0$ tels que $E_S \sup_{\theta\in\Theta^*}\{\psi_i^{(k)}(\theta,\theta_0,S_t,\dots,S_{t-i+1})\}^2 \leqslant \kappa\rho^i$, $\forall i$, pour $k = 1,\dots,(p+q+1)d$.
**H5** Pour $k = 1,\dots,(p+q+1)d$, $E_S\{\tilde{c}^{(k)}(S_t,S_{t-1},\dots)\}^2 < \infty$.
**H6** Quand $(\underline{a},\underline{b}) \neq (\underline{a}_0,\underline{b}_0)$, il existe $i_0 \geqslant 1$ tel que $E_S\{\psi_{i_0}(\underline{a},\underline{b},\underline{a}_0,\underline{b}_0,S_t,\dots,S_{t-i_0+1})\}^2 > 0$. Quand $\underline{m} \neq \underline{m}_0$, $\liminf_{n\to\infty} n^{-1}\sum_{t=1}^{n}[c\{(\underline{m},\underline{a}_0,\underline{b}_0),\theta_0,s_t,\dots,s_1\}]^2 > 0$.

Les trois hypothèses ci-après sont utilisées pour établir la normalité asymptotique.

**H7** La vraie valeur $\theta_0$ appartient à l'intérieur de $\Theta^*$.

**H8** Il existe $\rho \in [0,1[$ et $\kappa > 0$ tels que, pour $k_1, k_2, k_3 = 1, \ldots, (p+q+1)d$, $E_S \sup_{\theta\in\Theta^*}\{\psi_i^{(k_1)}(\theta, \theta_0, S_t, \ldots, S_{t-i+1})\}^4 \leqslant \kappa\rho^i$, $E_S \sup_{\theta\in\Theta^*}\{\psi_i^{(k_1,k_2)}(\theta, \theta_0, S_t, \ldots, S_{t-i+1})\}^4 \leqslant \kappa\rho^i$, et $E_S \sup_{\theta\in\Theta^*}\{\psi_i^{(k_1,k_2,k_3)}(\theta, \theta_0, S_t, \ldots, S_{t-i+1})\}^4 \leqslant \kappa\rho^i$, $\forall i$.

**H9** Pour $k_1, k_2, k_3 = 1, \ldots, (p+q+1)d$, $E_S\{\tilde{c}^{(k_1)}(S_t, S_{t-1}, \ldots)\}^4 < \infty$, $E_S\{\tilde{c}^{(k_1,k_2)}(S_t, S_{t-1}, \ldots)\}^4 < \infty$, et $E_S\{\tilde{c}^{(k_1,k_2,k_3)}(S_t, S_{t-1}, \ldots)\}^4 < \infty$.

**Remarque 2.** Il est important de noter que, bien qu'engendrée par un processus aléatoire (voir **H1**), la suite $(s_t)$ des changements de régime est supposée connue. Par conséquent, le modèle (1) ne possède qu'un seul processus latent $(\eta_t)$ et le processus $(X_t)$ est non-stationnaire. Le comportement asymptotique des estimateurs considérés est étudié conditionnellement à $(S_t) = (s_t)$, pour $\mathbb{P}$-presque toute suite $(s_t)$.

## 3. Comportement asymptotique des estimateurs

Etant donné $(X_1, \ldots, X_n)$, les estimateurs des MC et des MCQG sont définis par la procédure ci-dessous.

ETAPE 1. Un estimateur des MC de $\theta_0$ est obtenu en résolvant $\hat{\theta}_n = \arg\min_{\theta\in\Theta^*} n^{-1}\sum_{t=1}^n e_t^2(\theta)$.
ETAPE 2. Pour $k = 1, \ldots, d$ et en notant $\mathbb{I}_k(\cdot)$ la fonction indicatrice de l'état $k$, un estimateur $\underline{\hat{\sigma}}_n = \{\hat{\sigma}_n(1), \ldots, \hat{\sigma}_n(d)\}'$ de $\underline{\sigma}_0$ est défini par $\hat{\sigma}_n^2(k) = \{\sum_{t=1}^n \mathbb{I}_k(s_t)\}^{-1}\sum_{t=1}^n \mathbb{I}_k(s_t)e_t^2(\hat{\theta}_n)$.
ETAPE 3. Un estimateur des MCQG est obtenu en résolvant $\hat{\theta}_n^Q = \arg\min_{\theta\in\Theta^*} n^{-1}\sum_{t=1}^n\{\hat{\sigma}_n^2(s_t)\}^{-1}e_t^2(\theta)$.

Pour énoncer nos théorèmes, nous introduisons des notations supplémentaires. Sous **H3** et **H5**,

$$c_{\infty,t}(\theta,\theta_0) = c_{\infty,t}(\theta,\theta_0,s_t,s_{t-1},\ldots) := \sum_{i=0}^{\infty}\varphi_{t,i}(\theta,s_t,\ldots,s_{t-i+1})\big\{m_0(s_{t-i}) - m(s_{t-i})\big\}$$

et $c_{\infty,t}^{(k)}(\theta,\theta_0)$ existent pour $\mathbb{P}$-presque toute suite $(s_t)$. Le théorème suivant fournit le comportement asymptotique des estimateurs des MC.

**Théorème 3.1.** *Soit $\hat{\theta}_n$ un estimateur des MC de $\theta_0$. Sous les hypothèses* **H1–H6**, *pour $\mathbb{P}$-presque toute suite $(s_t)$, $\hat{\theta}_n \to \theta_0$ p.s. quand $n \to \infty$ et la variable $\sqrt{n}(\hat{\theta}_n - \theta_0)$ converge en distribution vers une loi normale centrée et de matrice de variance-covariance $\Sigma := B^{-1}AB^{-1}$ où $B = \Psi + C$, $A = \tilde{\Psi} + \widetilde{C}$, $\Psi = \sum_{i=1}^{\infty}\Psi_i$ et $\tilde{\Psi} = \sum_{i=1}^{\infty}\tilde{\Psi}_i$ sont des matrices dont les termes génériques sont*

$$C(k_1,k_2) = E_S\big\{c_{\infty,t}^{(k_1)}(\theta_0,\theta_0,S_t,\ldots)c_{\infty,t}^{(k_2)}(\theta_0,\theta_0,S_t,\ldots)\big\},$$
$$\widetilde{C}(k_1,k_2) = E_S\big\{\sigma^2(S_t)c_{\infty,t}^{(k_1)}(\theta_0,\theta_0,S_t,\ldots)c_{\infty,t}^{(k_2)}(\theta_0,\theta_0,S_t,\ldots)\big\},$$
$$\Psi_i(k_1,k_2) = E_S\big\{\psi_i^{(k_1)}(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\psi_i^{(k_2)}(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\sigma^2(S_{t-i})\big\},$$
$$\tilde{\Psi}_i(k_1,k_2) = E_S\big\{\sigma^2(S_t)\psi_i^{(k_1)}(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\psi_i^{(k_2)}(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\sigma^2(S_{t-i})\big\}.$$

*Un estimateur de $\Sigma$ faiblement convergent est donné par $\widehat{\Sigma}_n = \widehat{B}_n^{-1}\hat{A}_n\widehat{B}_n^{-1}$ où*

$$\hat{A}_n = n^{-1}\sum_{t=1}^n e_t^2(\hat{\theta}_n)\frac{\partial}{\partial\theta}e_t(\hat{\theta}_n)\frac{\partial}{\partial\theta'}e_t(\hat{\theta}_n) \quad et \quad \widehat{B}_n = n^{-1}\sum_{t=1}^n\frac{\partial}{\partial\theta}e_t(\hat{\theta}_n)\frac{\partial}{\partial\theta'}e_t(\hat{\theta}_n).$$

Asymptotiquement, les estimateurs des MCQG sont plus précis que les estimateurs des MC.

**Théorème 3.2.** *Sous les hypothèses du Théorème* 3.1, *les estimateurs des MCQG $\hat{\theta}_n^Q$ convergent p.s. vers $\theta_0$ et suivent asymptotiquement une loi normale centrée et de matrice de variance-covariance $\Sigma^Q := B^{Q^{-1}}$ où $B^Q = \Psi^Q + C^Q$ et $\Psi^Q = \sum_{i=1}^{\infty}\Psi_i^Q$ sont des matrices dont les termes génériques sont*

$$C^Q(k_1,k_2)=E_S\bigl\{\sigma^{-2}(S_t)c^{(k_1)}_{\infty,t}(\theta_0,\theta_0,S_t,\ldots)c^{(k_2)}_{\infty,t}(\theta_0,\theta_0,S_t,\ldots)\bigr\},$$
$$\Psi^Q_i(k_1,k_2)=E_S\bigl\{\sigma^{-2}(S_t)\psi^{(k_1)}_i(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\psi^{(k_2)}_i(\theta_0,\theta_0,S_t,\ldots,S_{t-i+1})\sigma^2(S_{t-i})\bigr\}.$$

*De plus, $\Sigma-\Sigma^Q$ est définie non-négative. Un estimateur de $\Sigma$ faiblement convergent est donné par $\widehat{\Sigma}^Q_n=\widehat{B}^{Q-1}_n$ où*

$$\widehat{B}^Q_n=n^{-1}\sum_{t=1}^{n}\hat{\sigma}^{-2}_n(s_t)\frac{\partial}{\partial\theta}e_t\bigl(\hat{\theta}^Q_n\bigr)\frac{\partial}{\partial\theta'}e_t\bigl(\hat{\theta}^Q_n\bigr).$$

Les démonstrations des Théorèmes 3.1 et 3.2 se trouvent dans Francq et Gautier [5].

## 4. Application à des modèles à changements de régime Markoviens

On considère le cas où $(S_t)$ est une chaîne de Markov à espace d'états fini $\{1,\ldots,d\}$, stationnaire, irréductible, apériodique et dont les probabilités de transition sont $\mathbb{P}(i,j)=\mathbb{P}(S_t=j|S_{t-1}=i)$. On note $\pi(i)=\mathbb{P}(S_t=i)$, $1\leqslant i\leqslant d$, les probabilités stationnaires et $Q_f$ la matrice $d\times d$ dont l'élément de la $i$-ème ligne et de la $j$-ème colonne est $f(i)\mathbb{P}(j,i)$, où $f$ est une fonction de $\{1,\ldots,d\}$ à valeurs réelles.

**Exemple 1.** Modèle (1) avec $(p,q)=(1,0)$ : les hypothèses **H1–H9** se réduisent à la condition $\rho(Q_{a^4_{01}})<1$, où $\rho(\cdot)$ désigne le rayon spectral d'une matrice. Le cas où $(S_t)$ est i.i.d., obtenu en prenant $\mathbb{P}(i,j)=\pi(j)$, pour $j=1,\ldots,d$, se ramène à la simple condition $\sum_{j=1}^d a^4_{01}(j)\pi(j)<1$.

**Exemple 2.** Modèle (1) avec $(p,q)=(0,1)$ : les hypothèses **H1–H9** se réduisent à $\sup_{\theta\in\Theta^*}\rho(Q_{b^4_1})<1$. Le cas où $(S_t)$ est i.i.d. se ramène à la condition $\sum_{j=1}^d b^4_1(j)\pi(j)<1$ pour tout $\theta\in\Theta^*$.

**Remarque 3.** Azrak et Mélard [2] et Bibi et Francq [4] ne font pas l'hypothèse **H1**, ce qui leur permet de développer une théorie générale, applicable notamment à des modèles à coefficients périodiques. En raison de **H1**, nous ne pouvons traiter le cas où $(s_t)$ est périodique (étudié par exemple dans Adams et Goodwin [1] ou Basawa et Lund [3]). Cependant, dans le cas de changements de régime récurrents mais non-périodiques, nous obtenons des conditions moins contraignantes. En effet, Azrak et Mélard [2] et Bibi et Francq [4] imposent que $\sup_{t\geqslant 1}\sup_{\theta\in\Theta}|\psi_{t,i}(\theta,\theta_0)|\to 0$ suffisamment rapidement quand $i\to\infty$. Appliquée à un AR(1) de période $d$, cette condition est équivalente à $|\prod_{j=1}^d a_{01}(j)|<1$, ce qui n'exclut pas l'existence de coefficients de module supérieur à 1. Cependant, appliquée à l'Exemple 1 avec $(S_t)$ i.i.d., cette condition correspond à $\max_{j=1,\ldots,d}|a_{01}(j)|<1$ (voir Francq et Gautier [5], page 5), ce qui est bien plus restrictif que la condition que nous obtenons.

## Références

[1] G.J. Adams, G.C. Goodwin, Parameter estimation for periodic ARMA models, J. Time Ser. Anal. 16 (1995) 127–146.
[2] R. Azrak, G. Mélard, Asymptotic properties of quasi-likelihood estimators for ARMA models with time-dependent coefficients, document de travail, ISRO et ECARES, 1999.
[3] I.V. Basawa, R. Lund, Large sample properties of parameter estimates for periodic ARMA models, J. Time Ser. Anal. 22 (2001) 651–663.
[4] A. Bibi, C. Francq, Consistent and asymptotically normal estimators for cyclically time-dependent linear models, Ann. Inst. Statist. Math. 55 (2003) 41–68.
[5] C. Francq, A. Gautier, Estimation of time-varying ARMA models and applications to series subject to Markovian changes in regime, document de travail, Université Lille 3, GREMARS, http://gremars.univ-lille3.fr/~gautier/longversion.ps, 2003.