ANNALES DE L'I. H. P., SECTION B

B. BERCU

Sur l'estimateur des moindres carrés généralisé d'un modèle ARMAX. Application à l'identification des modèles ARMA

Annales de l'I. H. P., section B, tome 27, n° 3 (1991), p. 425-443

<http://www.numdam.org/item?id=AIHPB_1991__27_3_425_0>

© Gauthier-Villars, 1991, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (http://www.elsevier.com/locate/anihpb) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Ann. Inst. Henri Poincaré,

Vol. 27, n° 3, 1991, p. 425-443. *Probabilités et Statistiques*

Sur l'estimateur des moindres carrés généralisé d'un modèle ARMAX. Application à l'identification des modèles ARMA

par

B. BERCU

Bât. n° 425, Mathématiques,
Université de Paris-Sud,
91405 Orsay Cedex, France

RÉSUMÉ. — Pour un modèle **ARMAX** vectoriel complexe, on étudie, sous l'hypothèse de passivité, la consistance forte de l'estimateur des moindres carrés généralisé pour un schéma récursif **AML**, en précisant les vitesses de convergence presque sûre. On travaille également sur les erreurs de prédiction liées à un tel modèle.

On donne ensuite, pour un modèle **ARMA** vectoriel complexe, le comportement asymptotique presque sûr des trajectoires. On prouve, dans les cas stable et instable, la consistance forte et dans le cas explosif régulier celle de la partie autorégressive, en précisant les vitesses de convergence presque sûre.

ABSTRACT. — For a complex multivariate **ARMAX** model, we study the strong consistency of the extended least squares estimator for a recursive **AML** scheme, on the hypothesis of passivity. We precise the almost sure rates of convergence. We also work on the prediction errors related to such a model.

Afterwards, we give the almost sure asymptotic behaviour of the trajectories of a complex multivariate **ARMA** model. We prove the strong consistency on the stable and unstable cases. On the explosive regular case, we prove the strong consistency of the autoregressive part. We give the almost sure rates of convergence.

I. MODÈLE ARMAX VECTORIEL

I.1. Introduction

De nombreux travaux ont été réalisés sur les modèles **ARMAX**, notamment par Caines [1], Duflo [3], Goodwin, Ramadge et Caines [5], Lai et Wei [6], [8], [9] et [10] et Solo [12].

Ces auteurs ont souvent travaillé dans un cadre unidimensionnel et ont peu travaillé sur les erreurs de prédiction. Après avoir rappelé certains résultats qui découlent de leurs travaux, on donne, sous l'hypothèse de passivité, des critères de consistance forte de l'estimateur des moindres carrés généralisé pour un schéma récursif **AML** et des résultats sur les erreurs de prédiction qui jouent un rôle fondamental dans tous les problèmes de poursuite.

Soit $(\Omega, \mathscr{A}, P)$ un espace de probabilité muni d'une filtration $\mathbf{F}=(\mathscr{F}_n)_{n\geqq 0}$. On considère le modèle **ARMAX** vectoriel complexe d'ordre (p, q, r) défini pour tout $n\geqq 0$ par

$$A(R)\,Y_n = B(R)\,U_n + C(R)\,\varepsilon_n$$

• A, B et C sont des polynômes matriciels complexes avec pour $z\in\mathbf{C}$

$$A(z)=I_{d_1}-A_1 z-\ldots-A_p z^p,$$
$$B(z)=B_1 z+\ldots+B_q z^q,$$
$$C(z)=I_{d_1}+C_1 z+\ldots+C_r z^r,$$

et $A_1, \ldots, A_p, C_1, \ldots, C_r$ sont des matrices carrées complexes d'ordre d_1, $B_1, \ldots, B_q$ sont des matrices rectangulaires complexes d'ordre $d_1\times d_2$ et I_{d_1} est la matrice identité d'ordre d_1.

• Le bruit $\varepsilon=(\varepsilon_n)$ est une suite de vecteurs aléatoires de dimension d_1 adaptée à $\mathbf{F}$ avec $E(\varepsilon_{n+1}\,|\,\mathscr{F}_n)=0$ et $E(\varepsilon_{n+1}\,{}^*\varepsilon_{n+1}\,|\,\mathscr{F}_n)=\Gamma$ où Γ est une matrice de covariance d'ordre d_1 déterministe.

• Le contrôle $U=(U_n)$ est une suite de vecteurs aléatoires de dimension d_2 adaptée à $\mathbf{F}$.

• L'état initial $(Y_{1-p}, \ldots, Y_0, U_{1-q}, \ldots, U_0, \varepsilon_{1-r}, \ldots, \varepsilon_0)$ est $\mathscr{F}_0$ mesurable.

• R est l'opérateur retard sur les suites vectorielles.

NOTATION. — Pour deux vecteurs complexes a et b de même dimension, on utilise ${}^*a={}^t\bar{a}$, $\langle a, b\rangle={}^*ab$ et $\|a\|^2=\langle a, a\rangle$.

Pour une matrice symétrique A, $\lambda_{\max}(A)$ et $\lambda_{\min}(A)$ sont sa plus grande et sa plus petite valeurs propres. Pour un entier naturel d, I_d est la matrice identité d'ordre d. I est la matrice identité d'ordre $d_1(p+r)+d_2 q$.

Soit ${}^*\theta=(A_1, \ldots, A_p, B_1, \ldots, B_q, C_1, \ldots, C_r)$. On cherche à estimer θ constitué par $(p+r)$ matrices carrées d'ordre d_1 et q matrices rectangulaires d'ordre $d_1 \times d_2$.

Pour tout $n \geqq 0$, soit ${}^t\Psi_n=({}^tY_n^p, {}^tU_n^q, {}^t\varepsilon_n^r)$ où

- ${}^tY_n^p=({}^tY_n, \ldots, {}^tY_{n-p+1})$,
- ${}^tU_n^q=({}^tU_n, \ldots, {}^tU_{n-q+1})$,
- ${}^t\varepsilon_n^r=({}^t\varepsilon_n, \ldots, {}^t\varepsilon_{n-r+1})$.

On montre facilement que la relation **ARMAX** ci-dessus, peut s'écrire, pour tout $n \geqq 0$, $Y_{n+1}={}^*\theta\Psi_n+\varepsilon_{n+1}$.

Pour tout $n \geqq 0$, soit $\hat{\varepsilon}_n$ un estimateur de ε_n, Φ_n de Ψ_n et $\hat{\theta}_n$ de θ et soit $\beta_n={}^*\theta\Psi_n-{}^*\hat{\theta}_n\Phi_n$ l'erreur de prédiction à l'instant n. On utilise

$$Q_n=\sum_{k=0}^{n}\Psi_k{}^*\Psi_k+I \qquad \text{et} \qquad S_n=\sum_{k=0}^{n}\Phi_k{}^*\Phi_k+I.$$

Soit q_n et s_n les traces des matrices Q_n et S_n. On utilise également

$$f_n={}^*\Phi_n S_n^{-1}\Phi_n \qquad \text{et} \qquad g_n={}^*\Phi_n S_{n-1}^{-1}\Phi_n.$$

MÉTHODE **AML**. — On travaille sur un schéma récursif **PO** si l'on utilise pour estimer ε_n, l'erreur *a posteriori*

$$\hat{\varepsilon}_n=Y_n-{}^*\hat{\theta}_n\Phi_{n-1} \qquad \text{où} \qquad {}^t\Phi_n=({}^tY_n^p, {}^tU_n^q, {}^t\hat{\varepsilon}_n^r).$$

On appelle méthode récursive **AML**, abréviation anglosaxonne de Approximate **M**aximum **L**ikelihood, l'étude de l'estimateur des moindres carrés généralisé pour le schéma **PO**, défini pour tout $n \geqq 0$ par

$$\hat{\theta}_{n+1}=\hat{\theta}_n+S_n^{-1}\Phi_n{}^*(Y_{n+1}-{}^*\hat{\theta}_n\Phi_n).$$

I.2. Consistance

THÉORÈME 1. — *On suppose que dans le modèle* **ARMAX**, *le bruit* ε *a un moment d'ordre* >2 *fini. On suppose que le polynôme matriciel* C *est passif.*

Soit $F_n=\sum_{k=1}^{n} f_k$. *On a toujours p.s.*

$$\| S_n^{1/2}(\hat{\theta}_{n+1}-\theta)\|^2=O(F_n)$$

$$\sum_{k=1}^{n} \|\Phi_k - \Psi_k\|^2 = O(\mathrm{F}_n)$$

$$\sum_{k=1}^{n} \|{}^*\theta\Psi_k - {}^*\hat{\theta}_{k+1}\Phi_k\|^2 = O(\mathrm{F}_n)$$

$$\sum_{k=1}^{n} \|{}^*(\hat{\theta}_{k+1} - \theta)\Phi_k\|^2 = O(\mathrm{F}_n)$$

$$\sum_{k=1}^{n} \|\hat{\varepsilon}_k - \varepsilon_k\|^2 = O(\mathrm{F}_n)$$

$$\sum_{k=1}^{n} (1 - f_k)\|\beta_k\|^2 = O(\mathrm{F}_n)$$

THÉORÈME 2. — *On se place dans le cadre du théorème* 1 *mais on suppose seulement que le bruit* ε *a un moment d'ordre* 2 *fini. Soit* f *une fonction réelle, strictement positive et croissante avec*

$$\lim_{x \to +\infty} f(x) = +\infty \qquad et \qquad \int_a^{+\infty} f^{-1}(x)\,dx < +\infty$$

pour un $a > 0$. *Alors, sur* $\{\mathrm{F}_n \to +\infty\}$, *le théorème* 1 *reste valable en remplaçant partout* $O(\mathrm{F}_n)$ *par* $o(f(\mathrm{F}_n))$, *sauf dans la première relation.*

Remarque. — On peut prendre par exemple, pour $\alpha > 0$, $f(x) = x^{1+\alpha}$ ou $f(x) = x(\log x)^{1+\alpha}$.

Remarque. — Soit pour $n \geqq 1$, $\mathrm{K}_n = \inf\{n, \log s_n\}$. Comme $\mathrm{F}_n = O(\mathrm{K}_n)$, les théorèmes ci-dessus restent valables en remplaçant partout F_n par K_n.

COMMENTAIRE. — Les deux premières relations ont été prouvées par Lai et Wei, dans un cadre unidimensionnel [8] puis dans un cadre multidimensionnel [9] avec une vitesse de convergence en $\log s_n$. Les trois suivantes en découlent. La dernière donne un résultat utile sur les erreurs de prédiction.

COROLLAIRE 1. — *On se place dans le cadre du théorème* 1.

(*a*) *L'estimateur des moindres carrés généralisé est fortement consistant sur* $\mathrm{I}_1 \cup \mathrm{I}_2$ *où*

$$\mathrm{I}_1 = \left\{ \lim_{n \to +\infty} \frac{\lambda_{\min}(\mathrm{S}_n)}{n} = +\infty \right\} \qquad et \qquad \mathrm{I}_2 = \{\log s_n = o(\lambda_{\min}(\mathrm{S}_n))\}$$

et, sur $\mathrm{I}_1 \cup \mathrm{I}_2$, *on a*

$$\|\hat{\theta}_{n+1} - \theta\|^2 = O\left(\frac{\mathrm{K}_n}{\lambda_{\min}(\mathrm{S}_n)}\right) \qquad où\ \mathrm{K}_n = \inf\{n, \log s_n\}.$$

(*b*) *De plus, si on a* $\lambda_{\max}(Q_n) \to +\infty$ *et* $\log(\lambda_{\max}(Q_n)) = o(\lambda_{\min}(Q_n))$ *p.s. alors l'estimateur des moindres carrés généralisé est fortement consistant sur* $J_1 \cup J_2$ *où*

$$J_1 = \left\{ \lim_{n \to +\infty} \frac{\lambda_{\min}(Q_n)}{n} = +\infty \right\} \qquad et \qquad J_2 = \{ \log q_n = o(\lambda_{\min}(Q_n)) \}$$

et, sur $J_1 \cup J_2$, *on a*

$$\| \hat{\theta}_{n+1} - \theta \|^2 = O\left(\frac{L_n}{\lambda_{\min}(Q_n)} \right) \qquad où \quad L_n = \inf\{ n, \log q_n \}.$$

Corollaire 2. — *On se place dans le cadre du théorème* 2.

Les résultats du corollaire 1 restent valables en remplaçant K_n par $f(K_n)$ et L_n par $f(L_n)$.

Corollaire 3. — *On se place dans le cadre du théorème* 1 *avec*

$$P_n = \sum_{k=1}^{n} \| \beta_k \|^2.$$

(*a*) *Si* $\lambda_{\max}(S_n) \to +\infty$ *et* $\log(s_n) = o(\lambda_{\min}(S_{n-1}))$ *p.s.* $s_n^{-1} P_n \to 0$ *p.s.*

(*b*) *Si* $\lambda_{\max}(Q_n) \to +\infty$ et $\log(q_n) = o(\lambda_{\min}(Q_{n-1}))$ *p.s.* $q_n^{-1} P_n \to 0$ *p.s.*

Corollaire 4. — *On se place dans le cadre du théorème* 2.

Les résultats du corollaire 3 restent valables en remplaçant $\log(s_n)$ par $f(\log(s_n))$ et $\log(q_n)$ par $f(\log(q_n))$ et tous les o par des O.

Remarque. — On utilise pour prouver ces corollaires, les propriétés de transfert suivantes

- Par le théorème 1, on a toujours $\lambda_{\max}(Q_n) = O(\lambda_{\max}(S_n))$ p.s.
- De plus, si l'on suppose que $\lambda_{\max}(Q_n) \to +\infty$ p.s. alors les suites $(\lambda_{\max}(S_n))$, $(\lambda_{\max}(Q_n))$ et les suites (s_n), (q_n) sont équivalentes.
- Si $\lambda_{\max}(Q_n) \to +\infty$ et $\log(\lambda_{\max}(Q_n)) = o(\lambda_{\min}(Q_n))$ p.s. alors les suites $(\lambda_{\min}(S_n))$, $(\lambda_{\min}(Q_n))$ sont du même ordre.

Remarque. — On peut montrer, ce qui nous sera utile dans le cadre **ARMA**, que si $\lambda_{\max}(Q_n) \to +\infty$ et $\log(\lambda_{\max}(Q_n)) = o(\lambda_{\min}(Q_n))$ p.s. alors il existe deux constantes a et b strictement positives telles que, pour n assez grand

$$a\, Q_n \leq S_n \leq b\, Q_n.$$

II. IDENTIFICATION DES MODÈLES ARMA

II.1. Introduction

On reprend, pour les modèles **ARMA**, les résultats obtenus par Duflo, Senoussi et Touati [4] et Lai et Wei [7], sur les modèles autorégressifs. On précise comment l'on passe facilement du cadre **AR** au cadre **ARMA**. On étudie, en supposant que la matrice de covariance du bruit associé au modèle est inversible, le comportement asymptotique presque sûr des trajectoires. On prouve, dans les cas stable et instable, la consistance forte et dans le cas explosif régulier, celle de la partie autorégressive, en précisant les vitesses de convergence presque sûre.

On se place dans le cadre **ARMA** avec $q=0$. On considère le modèle **ARMA** vectoriel complexe d'ordre (p, r) et de dimension d, défini pour tout $n \geqq 0$ par

$$\mathrm{A}(\mathrm{R})\,\mathrm{Y}_n = \mathrm{C}(\mathrm{R})\,\varepsilon_n.$$

• A et C sont des polynômes matriciels complexes avec pour $z \in \mathbf{C}$

$$\begin{gathered}\mathrm{A}(z) = \mathrm{I}_d - \mathrm{A}_1 z - \ldots - \mathrm{A}_p z^p,\\ \mathrm{C}(z) = \mathrm{I}_d + \mathrm{C}_1 z + \ldots + \mathrm{C}_r z^r,\end{gathered}$$

et $\mathrm{A}_1, \ldots, \mathrm{A}_p, \mathrm{C}_1, \ldots, \mathrm{C}_r$ sont des matrices carrées complexes d'ordre d. On suppose que l'on a une représentation **ARMA** irréductible, les polynômes A et C sont premiers entre eux à gauche.

• On suppose dans toute notre étude A_p inversible.

• Le bruit $\varepsilon = (\varepsilon_n)$ est une suite de vecteurs aléatoires de dimension d, adaptée à $\mathbf{F}$, avec $\mathrm{E}(\varepsilon_{n+1} \mid \mathscr{F}_n) = 0$ et $\mathrm{E}(\varepsilon_{n+1}{}^* \varepsilon_{n+1} \mid \mathscr{F}_n) = \Gamma$ où Γ est une matrice de covariance d'ordre d inversible. De plus, on suppose que ε est soit un bruit blanc, soit un bruit admettant un moment d'ordre >2 fini.

Soit $\mathrm{P_A}$ le polynôme défini pour $z \in \mathbf{C}$ par $\mathrm{P_A}(z) = \det(\mathrm{A}(z))$. Soit $\underline{\alpha}$ et $\bar{\alpha}$ le plus petit et le plus grand des modules des racines du polynôme $\mathrm{P_A}$.

• Le modèle est stable si $\underline{\alpha} > 1$.

• Le modèle est instable si $\underline{\alpha} = 1$.

• Le modèle est explosif si $\bar{\alpha} < 1$.

• Le modèle est régulier s'il n'existe aucun $z \in \mathbf{C}$ avec $|z| < 1$ tel que $\mathrm{A}(z)$ ait un noyau de dimension $\geqq 2$.

Remarque. — Comme le montrent Duflo, Senoussi et Touati [4] dans le cadre **AR** vectoriel, l'hypothèse de régularité est inévitable si l'on cherche à obtenir des résultats de consistance forte dans le cas explosif. Elle est donc, *a fortiori*, inévitable dans le cadre **ARMA** vectoriel explosif mais elle est toujours réalisée dans le cadre **ARMA** scalaire explosif.

II.2. Passage ARMA-AR

Soit B la matrice compagne du polynôme matriciel A. B est une matrice carrée complexe d'ordre s où $s=dp$ avec

$$B=\begin{pmatrix} A_1 & A_2 & A_3 & \dots & A_p \\ I_d & 0 & 0 & \dots & 0 \\ 0 & I_d & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & I_d & 0 \end{pmatrix}.$$

Soit J la matrice rectangulaire de dimension $s\times d$ dont le bloc supérieur est la matrice I_d et dont les autres termes sont nuls. Pour tout $1\leqq i\leqq r$, soit $\mathscr{C}_i$ la matrice carrée complexe d'ordre s dont la première colonne correspond à JC_i et dont les autres colonnes sont nulles.

En posant, pour tout $n\geqq 0$, ${}^tX_n={}^tY_n^p$ avec ${}^tY_n^p=({}^tY_n, \dots, {}^tY_{n-p+1})$ et $\xi_n=J\,\varepsilon_n$, on se ramène à l'étude du modèle **ARMA** vectoriel complexe d'ordre $(1, r)$ et de dimension s, défini pour $n\geqq 0$ par

$$X_{n+1}=BX_n+\xi_{n+1}+\mathscr{C}_1\,\xi_n+\dots+\mathscr{C}_r\,\xi_{n-r+1}.$$

Soit $Z=(Z_n)$ le processus défini pour $n\geqq 0$ par

$$Z_n=X_n+D_1\,\xi_n+\dots+D_r\,\xi_{n-r+1}.$$

On montre facilement la relation $Z_n=BZ_{n-1}+D\xi_n$, si l'on a choisi

$$D=D_1+I_s \qquad \text{et} \qquad D_r=B^{-1}\mathscr{C}_r$$

et pour tout $1\leqq i\leqq (r-1)$, $D_i=B^{-1}(\mathscr{C}_i+D_{i+1})$.

Remarque. — Comme A_p est inversible, B est inversible car les valeurs propres de B sont les racines de l'équation

$$\det(A_p+A_{p-1}\,\lambda+\dots+A_1\,\lambda^{p-1}-\lambda^p\,I_d)=0$$

et les valeurs propres non nulles de B sont les inverses des racines du polynôme P_A.

NOTATION. — Pour $n\geqq 0$, on utilise $R_n=\sum_{k=0}^{n} X_k{}^*X_k+I_s$ et $r_n=\operatorname{trace}(R_n)$.

II.3. Comportement des trajectoires

THÉORÈME 3. — *Soit* ν *le maximum des ordres des racines de* P_A *de module* $\underline{\alpha}$. *On a p.s.*

- Cas stable :

$$\|X_n\| = o(n^{1/2}) \qquad et \qquad r_n = O(n).$$

De plus, on a $\lim_{n \to +\infty} {}^*X_n R_n^{-1} X_n = 0$.

- Cas instable :

$$\overline{\lim}\,(n^{2\nu-1} \log\log(n))^{-1/2} \|X_n\| \leqq \mathrm{Cte} \qquad et \qquad r_n = O(n^{2\nu} \log\log(n)).$$

De plus, si le bruit ε *admet un moment d'ordre* >2 *fini, on a*

$$\lim_{n \to +\infty} {}^*X_n R_n^{-1} X_n = 0.$$

- Cas explosif :

$$r_n = O(\underline{\alpha}^{-2n} n^{2\nu-2})$$

$(\underline{\alpha}^{2n} n^{-2\nu+2})\, r_n$ converge vers une v.a. strictement positive et l'on a

$$\lim_{n \to +\infty} \frac{1}{n} \log(r_n) = -2 \log(\underline{\alpha}).$$

COMMENTAIRE. — Comme $Z = (Z_n)$ est un processus **AR** commandable, le comportement asymptotique presque sûr de ces trajectoires est entièrement décrit par Duflo, Senoussi et Touati [4] ou Lai et Wei [7]. On peut en déduire facilement les résultats annoncés pour le comportement des trajectoires du processus **ARMA** $X = (X_n)$.

II.4. Consistance

Soit ${}^*\theta = (A, C)$ avec $A = (A_1, \ldots, A_p)$ et $C = (C_1, \ldots, C_r)$ et soit $\hat{\theta}_n$ l'estimateur des moindres carrés généralisé de θ. ${}^*\theta_n = (\hat{A}_n, \hat{C}_n)$ où $\hat{A}_n$ estime la partie **AR** A et $\hat{C}_n$ estime la partie **MA** C. On travaille, comme dans le cadre **ARMAX**, avec l'erreur *a posteriori*.

NOTATION. — Pour tout $n \geqq 0$, on utilise

- $Q_n = \sum_{k=0}^{n} \Psi_k {}^*\Psi_k + I_{d(p+r)}$ avec ${}^t\Psi_n = ({}^tX_n, {}^t\varepsilon_n^r)$ et $q_n = \mathrm{trace}(Q_n)$.
- $S_n = \sum_{k=0}^{n} \Phi_k {}^*\Phi_k + I_{d(p+r)}$ avec ${}^t\Phi_n = ({}^tX_n, {}^t\hat{\varepsilon}_n^r)$ et $s_n = \mathrm{trace}(S_n)$.

THÉORÈME 4. — *On suppose que le polynôme matriciel C est passif et que le bruit* ε *admet un moment d'ordre* >2 *fini.*

• Cas stable et instable :

$\hat{\theta}_n$ *est un estimateur fortement consistant de* θ *et*

$$\|\hat{A}_n - A\|^2 + \|\hat{C}_n - C\|^2 = O\left(\frac{\log(n)}{n}\right) \quad \text{p.s.}$$

De plus, on a $\lim_{n \to +\infty} {}^*\Phi_n S_n^{-1} \Phi_n = 0$ *p.s.*

• Cas explosif régulier :

$\hat{A}_n$ *est un estimateur fortement consistant de* A *et l'on a*

$$\|(\hat{A}_n - A) B^n\|^2 = O(n) \quad p.s.$$

COMMENTAIRE. — Le cas stable a déjà été traité par de nombreux auteurs, alors que les cas instable et explosif ont été peu explorés. Dans les cas stable et instable, on trouve les mêmes résultats de consistance forte que dans le cadre autorégressif étudié dans [4] et [7]. Par contre, dans le cas explosif régulier, on trouve un O à la place d'un o.

Remarque. — Si l'on suppose seulement que le bruit ε a un moment d'ordre 2 fini, comme dans le théorème 2, on a dans les cas stable et instable

$$\|\hat{A}_n - A\|^2 + \|\hat{C}_n - C\|^2 = O\left(\frac{f(\log(n))}{n}\right) \quad \text{p.s.}$$

et dans le cas explosif régulier

$$\|(\hat{A}_n - A) B^n\|^2 = O(f(n)) \quad \text{p.s.}$$

COROLLAIRE 5. — *On se place dans le cadre du théorème* 4. *On utilise, pour estimer* Γ, *les estimateurs*

• $\hat{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} (Y_k - {}^*\hat{\theta}_{k-1} \Phi_{k-1})\, {}^*(Y_k - {}^*\hat{\theta}_{k-1} \Phi_{k-1})$,

• $\tilde{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} \hat{\varepsilon}_k\, {}^*\hat{\varepsilon}_k$.

Alors, dans les cas stable et instable, $\hat{\Gamma}_n$ *et* $\tilde{\Gamma}_n$ *sont des estimateurs fortement consistants de* Γ.

Remarque. — Soit, pour tout $n \geqq 0$, $\beta_n = {}^*\theta \Psi_n - {}^*\hat{\theta}_n \Phi_n$ l'erreur de prédiction à l'instant n et $f_n = {}^*\Phi_n S_n^{-1} \Phi_n$. On a

• $\hat{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} (\varepsilon_k + \beta_{k-1})\, {}^*(\varepsilon_k + \beta_{k-1})$,

- $\tilde{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} (1-f_{k-1})^2 (\varepsilon_k + \beta_{k-1}) * (\varepsilon_k + \beta_{k-1})$.

III. LES DÉMONSTRATIONS ARMAX

III.1. Propriétés AML

Nous donnons une série de résultats, connus ou non, qui nous seront utiles par la suite.

UTILITAIRE 1. — Pour tout $n \geqq 0$, soit $f_n = {}^*\Phi_n S_n^{-1} \Phi_n$ $g_n = {}^*\Phi_n S_{n-1}^{-1} \Phi_n$ et soit $F_n = \sum_{k=1}^{n} f_k$.

(*a*) $(1-f_n) = (1+g_n)^{-1}$ donc $0 \leqq f_n \leqq 1$.

(*b*) $\hat{\varepsilon}_{n+1} = (1-f_n)(\beta_n + \varepsilon_{n+1})$ et $\hat{\theta}_{n+1} = \hat{\theta}_n + S_{n-1}^{-1} \Phi_n {}^*\hat{\varepsilon}_{n+1}$.

(*c*) $F_n \leqq d \inf \{ n, \log s_n \}$ avec $d = pd_1 + qd_2 + rd_1$.

PASSIVITÉ. — Soit A un polynôme matriciel complexe défini pour $z \in \mathbf{C}$ par $A(z) = A_0 + A_1 z + \ldots + A_n z^n$ où $A_0, A_1, \ldots, A_n$ sont des matrices carrées complexes d'ordre d.

- A est causal si $\det A(z) \neq 0$ pour tout $|z| \leqq 1$.
- A est positif si, pour tout $z \in \mathbf{C}$ avec $|z| = 1$, $A(z) + {}^*A(z) > 0$.
- A est passif si A est causal et $B = \left(A^{-1} - \frac{1}{2} I_d \right)$ est positif.

UTILITAIRE 2. — Soit A un polynôme matriciel complexe de dimension d. On suppose que A est passif. Soit $u = (u_n)$ une suite de vecteurs complexes de $\mathbf{C}^d$ et soit $v = (v_n)$ et $w = (w_n)$ les suites définies pour tout $n \geqq 1$ par $u_n = A(R) v_n$ et $w_n = B(R) u_n$.

(*a*) Il existe quatre constantes positives a_1, $a_1(u)$, b_1 et $b_1(u)$ avec

$$2 \operatorname{Re} \left(\sum_{k=1}^{n} {}^*u_k w_k \right) + a_1(u) \geqq a_1 \sum_{k=1}^{n} \| u_k \|^2$$

et

$$2 \operatorname{Re} \left(\sum_{k=1}^{n} {}^*u_k w_k \right) + b_1(u) \geqq b_1 \sum_{k=1}^{n} \| v_k \|^2.$$

(*b*) Soit $c=(c_n)$ une suite réelle, strictement positive et croissante. On a

$$2\,\mathrm{Re}\left(\sum_{k=1}^{n} c_k^{-1}\,({}^*u_k\,w_k)\right)+\frac{a_1(u)}{c_1}\geqq a_1\sum_{k=1}^{n} c_k^{-1}\,\|u_k\|^2$$

et

$$2\,\mathrm{Re}\left(\sum_{k=1}^{n} c_k^{-1}\,({}^*u_k\,w_k)\right)+\frac{b_1(u)}{c_1}\geqq b_1\sum_{k=1}^{n} c_k^{-1}\,\|v_k\|^2.$$

Commentaire. – L'Utilitaire 2 est prouvé, grâce à des outils d'Analyse de Fourier par Duflo [3]. Il clarifie les notions qui découlent, dans notre cadre, du lemme de Positivité. *Voir* Caines [1], Goodwin, Ramadge et Caines [5] et Solo [12].

III.2. Consistance

Soit $V_n={}^*\breve{\theta}_n\,S_{n-1}\,\breve{\theta}_n$ et $\alpha_n=-{}^*\breve{\theta}_n\,\Phi_{n-1}$. Par les Utilitaires 1 et 2, on montre l'égalité classique similaire à celle de Caines [1], Goodwin, Ramadge et Caine [5] et Lai et Wei [9]

$$\begin{aligned}V_{n+1}-V_n=\alpha_{n+1}\,{}^*\alpha_{n+1}-f_n(1-f_n)(\beta_n+\varepsilon_{n+1})\,{}^*(\beta_n+\varepsilon_{n+1})&\\+2f_n\,\varepsilon_{n+1}\,{}^*\varepsilon_{n+1}-(\breve{\varepsilon}_{n+1}\,{}^*\alpha_{n+1})&\\-{}^*(\breve{\varepsilon}_{n+1}\,{}^*\alpha_{n+1})+(\varepsilon_{n+1}\,{}^*\gamma_{n+1})+{}^*(\varepsilon_{n+1}\,{}^*\gamma_{n+1})&\end{aligned}$$

où l'on a posé

$$\breve{\theta}_n=\hat{\theta}_n-\theta\qquad\text{et}\qquad\breve{\varepsilon}_n=\hat{\varepsilon}_n-\varepsilon_n$$

et

$$\gamma_{n+1}={}^*\breve{\theta}_n\,\Phi_n+f_n\,\beta_n=-\alpha_{n+1}-f_n\,\varepsilon_{n+1}.$$

Soit v_n la trace de la matrice V_n. Par commutativité de la trace, on tire alors

$$\begin{aligned}v_{n+1}-v_n=\|\alpha_{n+1}\|^2-f_n(1-f_n)\,\|\beta_n+\varepsilon_{n+1}\|^2&\\+2f_n\,\|\varepsilon_{n+1}\|^2-2\,\mathrm{Re}\,({}^*\alpha_{n+1}\,\breve{\varepsilon}_{n+1})+2\,\mathrm{Re}\,({}^*\gamma_{n+1}\,\varepsilon_{n+1}).&\end{aligned}$$

En sommant l'égalité ci-dessus, on montre facilement que l'on a

$$\begin{aligned}v_{n+1}-v_1=\sum_{k=1}^{n}\|\alpha_{k+1}\|^2-\sum_{k=1}^{n} f_k(1-f_k)\,\|\beta_k+\varepsilon_{k+1}\|^2&\\+2\sum_{k=1}^{n} f_k\,\|\varepsilon_{k+1}\|^2-2\,\mathrm{Re}\left(\sum_{k=1}^{n}{}^*\alpha_{k+1}\,\breve{\varepsilon}_{k+1}\right)+2\,\mathrm{Re}\,(M_{n+1})&\end{aligned}$$

avec $M_{n+1}=\sum_{k=1}^{n} {}^*\gamma_{k+1}\,\varepsilon_{k+1}$.

$M=(M_n)$ est une martingale de processus croissant $\langle M\rangle=(\langle M\rangle_n)$ avec, pour tout $n\geqq 2$, $\langle M\rangle_n\leqq\lambda_{max}(\Gamma)\,\Delta_n$ où $\Delta_n=\sum_{k=2}^{n}\|\gamma_k\|^2$. Par suite, grâce à la loi forte des grands nombres, il existe une suite réelle convergente $\eta=(\eta_n)$ qui tend vers 0 si $\langle M\rangle_\infty=\infty$ et qui vérifie l'inégalité

$$\mathrm{Re}(M_n)\leqq\eta_n\,\Delta_n\quad \text{p. s.}$$

Soit $h_n=\check{\varepsilon}_n-\frac{1}{2}\alpha_n$. Par ce qui précède, on montre facilement que

$$v_{n+1}-v_1\leqq -2\,\mathrm{Re}\left(\sum_{k=1}^{n} {}^*\alpha_{k+1}\,h_{k+1}\right) +2(1+2\,\eta_{n+1})\sum_{k=1}^{n} f_k\|\varepsilon_{k+1}\|^2+4\,\eta_{n+1}\sum_{k=1}^{n}\|\alpha_{k+1}\|^2.$$

Mais par l'Utilitaire 2, il existe une constante l strictement positive et une v. a positive finie L avec

$$2\,\mathrm{Re}\left(\sum_{k=1}^{n} {}^*\alpha_{k+1}\,h_{k+1}\right)+L\geqq l\sum_{k=1}^{n}\|\alpha_{k+1}\|^2\quad \text{p. s.,}$$

donc

$$v_{n+1}-v_1\leqq L+(4\,\eta_{n+1}-l)\sum_{k=1}^{n}\|\alpha_{k+1}\|^2+2(1+2\,\eta_{n+1})\sum_{k=1}^{n} f_k\|\varepsilon_{k+1}\|^2.$$

Finalement, on a $v_{n+1}\leqq K+T_{n+1}$ où K est une v. a. positive finie et

$$T_{n+1}=3\sum_{k=1}^{n} f_k\|\varepsilon_{k+1}\|^2.$$

$f=(f_n)$ et $\varepsilon=(\varepsilon_n)$ sont deux suites adaptées à F. Si ε admet un moment d'ordre >2 fini, on a $T_n=O(F_{n-1})$ où $F_n=\sum_{k=1}^{n} f_k$. *Voir* Chow [2] ou Lai et Wei [6]. Par suite, on a

$$\|S_n^{1/2}(\hat{\theta}_{n+1}-\theta)\|^2\leqq v_{n+1}\quad \text{donc}\quad \|S_n^{1/2}(\hat{\theta}_{n+1}-\theta)\|^2=O(F_n).$$

On a

$$\sum_{k=1}^{n}\|\alpha_{k+1}\|^2=O(F_n)\quad \text{donc}\quad \sum_{k=1}^{n}\|{}^*(\hat{\theta}_{k+1}-\theta)\,\Phi_k\|^2=O(F_n).$$

De même

$$2\,\mathrm{Re}\left(\sum_{k=1}^{n} {}^{*}\alpha_{k+1}\,h_{k+1}\right)=O(\mathrm{F}_n) \quad \text{donc} \quad \sum_{k=1}^{n} \|\check{\varepsilon}_{k+1}\|^2=O(\mathrm{F}_n).$$

Or

$$\|\Phi_n-\Psi_n\|^2=\sum_{k=1}^{r} \|\check{\varepsilon}_{n-k+1}\|^2 \quad \text{donc} \quad \sum_{k=1}^{n} \|\Phi_k-\Psi_k\|^2=O(\mathrm{F}_n).$$

Comme

$$\sum_{k=1}^{n} \|{}^{*}\theta\Psi_k-{}^{*}\hat{\theta}_{k+1}\,\Phi_k\|^2\leqq 2\,\|\theta\|^2 \sum_{k=1}^{n} \|\Phi_k-\Psi_k\|^2+2\sum_{k=1}^{n} \|\alpha_{k+1}\|^2,$$

on obtient $\sum_{k=1}^{n} \|{}^{*}\theta\Psi_k-{}^{*}\hat{\theta}_{k+1}\,\Phi_k\|^2=O(\mathrm{F}_n)$.

Soit $\beta_n={}^{*}\theta\Psi_n-{}^{*}\hat{\theta}_n\,\Phi_n$ l'erreur de prédiction à l'instant n. On tire, grâce à l'Utilitaire 1, l'égalité $(1-f_n)\,\beta_n=\check{\varepsilon}_{n+1}+f_n\,\varepsilon_{n+1}$. Donc par suite, on a

$$\sum_{k=1}^{n} (1-f_k)^2\,\|\beta_k\|^2=O(\mathrm{F}_n).$$

En reprenant la démonstration ci-dessus, on montre également que

$$\sum_{k=1}^{n} f_k(1-f_k)\,\|\varepsilon_{k+1}+\beta_k\|^2=O(\mathrm{F}_n)$$

donc

$$\sum_{k=1}^{n} f_k(1-f_k)\,\|\beta_k\|^2=O(\mathrm{F}_n)$$

et finalement $\sum_{k=1}^{n} (1-f_k)\,\|\beta_k\|^2=O(\mathrm{F}_n)$.

On a donc achevé la démonstration du théorème 1. La démonstration du théorème 2 est similaire à celle du théorème 1, en utilisant le lemme de Kronecker et l'Utilitaire 2.

IV. LES DÉMONSTRATIONS ARMA

IV.1. Commandabilité

Soit M la matrice carrée complexe d'ordre e avec $e=d(p+r)$,

$$M=\begin{pmatrix} B & C \\ 0 & E \end{pmatrix}$$

où C est la matrice rectangulaire complexe d'ordre $s \times dr$ dont les colonnes sont constituées par $JC_1, JC_2, \ldots, JC_r$ et E est la matrice carrée réelle d'ordre $dr \times dr$ dont les d premières lignes et les d dernières colonnes sont nulles et dont le bloc inférieur gauche correspond à la matrice identité d'ordre $d(r-1)$. Si, pour tout $n \geqq 0$, ${}^t\mathscr{E}_n=({}^tJ\,\varepsilon_n, {}^t\varepsilon_n, 0, \ldots, 0)$, on a

$$\Psi_{n+1}=M\,\Psi_n+\mathscr{E}_{n+1}.$$

$\Psi=(\Psi_n)$ est donc un processus autorégressif de dimension e. On suppose que Γ est inversible et que l'on a une représentation **ARMA** irréductible. Alors, grâce au théorème de transfert d'excitation, *voir* Duflo [3] ou Lai et Wei [9], il existe une constante ρ strictement positive, telle que

$$\underline{\lim}\ \frac{1}{n}\,\lambda_{\min}(Q_n) \geqq \rho \quad \text{p. s.}$$

Par suite, $\Psi=(\Psi_n)$ est un modèle autorégressif commandable. Soit T la matrice rectangulaire complexe, $T=(I_s\ D_1\,J \ldots D_r\,J)$. On tire que $Z=(Z_n)$ est également un modèle autorégressif commandable, car on a $Z_n=T\,\Psi_n$.

IV.2. Comportement des trajectoires

En reprenant les résultats obtenus par Duflo, Senoussi et Touati [4] et par Lai et Wei [7], sur le processus $Z=(Z_n)$, on obtient facilement la démonstration du théorème 3.

En effet, grâce à la relation liant $Z=(Z_n)$ et $X=(X_n)$ et grâce à l'inégalité de Cauchy-Schwarz, les résultats sur les normes et les traces associées à ces processus sont les mêmes. La seule difficulté consiste à montrer que dans les cas stable et instable

$$\lim_{n \to +\infty} {}^*X_n\,R_n^{-1}\,X_n=0 \quad \text{p. s.}$$

Le cas stable est simple. Pour n assez grand, si $h_n = {}^*X_n R_n^{-1} X_n$, on a

$$n = O(\lambda_{\min}(Q_n)) \qquad \text{et} \qquad \lambda_{\min}(Q_n) \leqq \lambda_{\min}(R_n)$$

donc $n = O(\lambda_{\min}(R_n))$.

Or $h_n \leqq \lambda_{max}(R_n^{-1}) \| X_n \|^2$ donc $h_n \leqq (\lambda_{\min}(R_n))^{-1} \| X_n \|^2$. Par suite $h_n = O\left(\frac{\| X_n \|^2}{n}\right)$ et comme $\| X_n \|^2 = o(n)$, on a $\lim_{n \to +\infty} h_n = 0$ p. s.

On étudie maintenant le cas instable. $\Psi = (\Psi_n)$ est un modèle autogréssif commandable. Soit $g_n = {}^*\Psi_n Q_n^{-1} \Psi_n$. Par les travaux déjà effectués dans le cadre autorégressif, *voir* [4] et [7], on a

$$\lim_{n \to +\infty} g_n = 0 \quad \text{p. s.}$$

On a

$$Q_n = \begin{pmatrix} R_n & K_n \\ {}^*K_n & L_n \end{pmatrix}$$

avec

$$K_n = \sum_{k=0}^{n} X_k {}^*\varepsilon_k^r \qquad \text{et} \qquad L_n = \sum_{k=0}^{n} \varepsilon_k^r {}^*\varepsilon_k^r + I_{dr}.$$

De plus, on a

$$g_n = \alpha_n + \beta_n + \gamma_n$$

avec

$$\alpha_n = {}^*\begin{pmatrix} X_n \\ 0 \end{pmatrix} Q_n^{-1} \begin{pmatrix} X_n \\ 0 \end{pmatrix} \qquad \text{et} \qquad \beta_n = {}^*\begin{pmatrix} 0 \\ \varepsilon_n^r \end{pmatrix} Q_n^{-1} \begin{pmatrix} 0 \\ \varepsilon_n^r \end{pmatrix}$$

et

$$\gamma_n = 2\,\mathrm{Re}\, {}^*\begin{pmatrix} X_n \\ 0 \end{pmatrix} Q_n^{-1} \begin{pmatrix} 0 \\ \varepsilon_n^r \end{pmatrix}.$$

On a $\alpha_n \leqq 2 g_n + 2\beta_n$ et comme g_n et β_n convergent vers 0 p. s., on tire

$$\lim_{n \to +\infty} \alpha_n = 0 \quad \text{p. s.}$$

Soit $E_n = L_n - {}^*K_n R_n^{-1} K_n$. Comme la matrice Q_n est inversible, E_n l'est également car on a $\det(Q_n) = \det(R_n) \det(E_n)$. Par suite, on peut choisir pour inverser Q_n, *voir* Rao [11], la matrice

$$Q_n^{-1} = \begin{pmatrix} R_n^{-1} + F_n E_n^{-1} {}^*F_n & -F_n E_n^{-1} \\ -E_n^{-1} {}^*F_n & E_n^{-1} \end{pmatrix} \qquad \text{avec} \quad F_n = R_n^{-1} K_n.$$

Mais E_n est définie positive et

$$\alpha_n = {}^*X_n R_n^{-1} X_n + {}^*X_n F_n E_n^{-1} {}^*F_n X_n$$

donc, comme $\lim_{n \to +\infty} \alpha_n = 0$, on a $\lim_{n \to +\infty} h_n = 0$ p. s.

On a donc achevé la démonstration du théorème 3. On prouve maintenant le théorème 4.

IV.3. Consistance

IV.3.1. *Cas stable et instable*

$\Psi = (\Psi_n)$ est un modèle autorégressif commandable. Par suite, grâce au théorème 3, on peut trouver $\mu > 0$ tel que

$$q_n = O(n^{2\mu}) \quad \text{donc } \log q_n = O(\log(n)).$$

Mais $n = O(\lambda_{\min}(Q_n))$, donc $\log(\lambda_{\max}(Q_n)) = o(\lambda_{\min}(Q_n))$ p. s. Donc, par le corollaire 1, on a

$$\| \hat{\theta}_n - \theta \|^2 = O\left(\frac{\log(n)}{n}\right)$$

donc

$$\| \hat{A}_n - A \|^2 + \| \hat{C}_n - C \|^2 = O\left(\frac{\log(n)}{n}\right) \quad \text{p. s.}$$

De plus, comme $\log(\lambda_{\max}(Q_n)) = o(\lambda_{\min}(Q_n))$ p. s., on peut trouver deux constantes a et b strictement positives telles que, pour n assez grand

$$a Q_n \leqq S_n \leqq b Q_n$$

donc

$$a S_n^{-1} \leqq Q_n^{-1} \qquad \text{et} \qquad a f_n \leqq {}^*\Phi_n Q_n^{-1} \Phi_n.$$

Soit $\Delta_n = \Psi_n - \Phi_n$. On a ${}^*\Phi_n Q_n^{-1} \Phi_n \leqq 2(g_n + {}^*\Delta_n Q_n^{-1} \Delta_n)$. Mais

$${}^*\Delta_n Q_n^{-1} \Delta_n \leqq \lambda_{max}(Q_n^{-1}) \| \Delta_n \|^2$$

donc ${}^*\Delta_n Q_n^{-1} \Delta_n \leqq (\lambda_{\min}(Q_n))^{-1} \| \Delta_n \|^2$.

Or, on a

$$\sum_{k=1}^{n} \| \Delta_k \|^2 = O(\log(n)) \quad \text{donc} \quad {}^*\Delta_n Q_n^{-1} \Delta_n = O\left(\frac{\log(n)}{n}\right).$$

Par suite, on a $\lim_{n\to+\infty} {}^*\Delta_n Q_n^{-1}\Delta_n=0$ p. s. et comme $\lim_{n\to+\infty} g_n=0$, on a $\lim_{n\to+\infty} {}^*\Phi_n Q_n^{-1}\Phi_n=0$ p. s. Finalement $\lim_{n\to+\infty} f_n=0$ p. s.

IV.3.2. *Cas explosif régulier*

Soit $T_n=B^{-n}X_n$ et $G_n=B^{-n}R_{n-1}{}^*B^{-n}$.

On a les mêmes résultats que Duflo, Senoussi et Touati [4] dans le cadre autorégressif, à savoir que T_n converge presque sûrement vers une variable aléatoire finie T dont la loi ne charge aucun hyperplan et que G_n converge presque sûrement vers une matrice aléatoire G, inversible sous l'hypothèse de régularité.

Soit $D_n=\operatorname{diag}(B^n, v(n))$ avec $n^{1/2}=o(v(n))$.

On a

$$S_n=\begin{pmatrix} R_n & \hat{K}_n \\ {}^*\hat{K}_n & \hat{L}_n\end{pmatrix}$$

avec

$$\hat{K}_n=\sum_{k=0}^{n} X_k{}^*\hat{\varepsilon}_k^r \qquad \text{et} \qquad \hat{L}_n=\sum_{k=0}^{n}\hat{\varepsilon}_k^r{}^*\hat{\varepsilon}_k^r+I_{dr}.$$

Par suite, on a $D_n^{-1}S_{n-1}{}^*D_n^{-1}=\begin{pmatrix} G_n & U_n \\ {}^*U_n & V_n\end{pmatrix}$ où l'on a posé

$$U_n=\frac{1}{v(n)}B^{-n}\hat{K}_{n-1} \qquad \text{et} \qquad V_n=\frac{1}{v(n)^2}\hat{L}_{n-1}.$$

Mais, par le théorème 1, on a $\sum_{k=0}^{n}\|\varepsilon_k-\hat{\varepsilon}_k\|^2=O(n)$ et comme $\|\varepsilon_n\|^2=o(n)$, on a $\|\hat{\varepsilon}_n\|^2=O(n)$ et

$$\operatorname*{Sup}_{k\leq n}\|\hat{\varepsilon}_k\|^2=O(n).$$

Par suite, comme $n^{1/2}=o(v(n))$, les matrices aléatoires U_n et V_n convergent presque sûrement vers 0, donc la matrice $D_n^{-1}S_{n-1}{}^*D_n^{-1}$, converge presque sûrement vers la matrice

$$\begin{pmatrix} G & 0 \\ 0 & 0\end{pmatrix}.$$

Par l'Utilitaire 1, on a $S_n \hat{\theta}_{n+1} = S_{n-1} \hat{\theta}_n + \Phi_n {}^*Y_{n+1}$. Donc, si l'on pose $\breve{\theta}_n = \hat{\theta}_n - \theta$, on a, pour $n \geqq 1$,

$$\hat{\theta}_n = S_{n-1}^{-1} (N_n + \hat{\theta}_0) \qquad \text{et} \qquad \breve{\theta}_n = S_{n-1}^{-1} (M_n + \breve{\theta}_0)$$

avec

$$N_n = \sum_{k=1}^{n} \Phi_{k-1} {}^*Y_k \qquad \text{et} \qquad M_n = \sum_{k=1}^{n} \Phi_{k-1} {}^*(\varepsilon_k + {}^*\theta\Delta_{k-1}).$$

Par suite, on a $D_n^{-1} S_{n-1} {}^*D_n^{-1} {}^*D_n \breve{\theta}_n = D_n^{-1} (M_n + \breve{\theta}_0)$, et l'on tire

$$\| (\hat{A}_n - A) B^n \|^2 = O(n)$$

car

$$\left\| B^{-n} \sum_{k=1}^{n} X_{k-1} {}^*\varepsilon_k \right\|^2 \leqq \sup_{k \leqq n} \| \varepsilon_k \|^2 \left(\left\| \sum_{k=1}^{n} B^{-n} X_{k-1} \right\|^2 \right),$$

$$\sup_{k \leqq n} \| \varepsilon_k \|^2 = o(n) \qquad \text{et} \qquad \left\| \sum_{k=1}^{n} B^{-n} X_{k-1} \right\|^2 = O(1)$$

et car

$$\left\| B^{-n} \sum_{k=1}^{n} X_{k-1} {}^*\Delta_{k-1} \theta \right\|^2 \leqq \sup_{k \leqq n} \| \Delta_k \|^2 \left(\left\| \sum_{k=1}^{n} B^{-n} X_{k-1} \right\|^2 \right) \| \theta \|^2$$

et

$$\sup_{k \leqq n} \| \Delta_k \|^2 = O(n).$$

RÉFÉRENCES

[1] P. E. Caines, *Linear Stochastic Systems*, Wiley, 1988.

[2] Y. S. Chow, Local Convergence of Martingales and the Law of Large Numbers, *Ann. Math. Stat.*, vol. **36**, 1965, p. 493-507.

[3] M. Duflo, *Méthodes Récursives Aléatoires*, Masson, 1990.

[4] M. Duflo, R. Senoussi et A. Touati, Propriétés asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autorégressif vectoriel, *Ann. Inst. H. Poincaré*, vol. **27**, 1991, p. 1-25.

[5] G. C. Goodwin, P. J. Ramadge et P. E. Caines, Discrete Time Stochastic Adaptive Control, *S.I.A.M. J. Control and Optimization*, vol. **19**, 1981, p. 829-853.

[6] T. L. Lai et C. Z. Wei, Least Squares Estimates in Stochastic Regression Models with Applications to Identification and Control of Dynamic Systems, *Ann. Statistics*, vol. **10**, 1982, p. 154-166.

[7] T. L. Lai et C. Z. Wei, Asymptotic Properties of General Autoregressive Models and Strong Consistency of Least Squares Estimates of Their Parameters, *J. Mult. Anal.*, vol. **13**, 1983, p. 1-23.

[8] T. L. LAI et C. Z. WEI, Extended Least Squares and Their Applications to Adaptive Control and Prediction in Linear Systems, *I.E.E.E. trans. automatic control.*, vol. **AC-31,** n° 10, 1986, p. 898-906.
[9] T. L. LAI et C. Z. WEI, On the Concept of Excitation in Least Squares Identification and Adaptive Control, *Stochastics*, vol. **16**, 1986, p. 227-254.
[10] T. L. LAI et C. Z. WEI, Asymptotically Efficient Self-Tuning Regulators, *S.I.A.M. J. Control Opt.*, vol. **25**, n° 2, 1987, p. 466-481.
[11] C. R. RAO, *Linear Statistical Inference and Its Appl.*, Wiley, 1965.
[12] V. SOLO, The Convergence of AML, *I.E.E.E. Trans. Automatic Control.*, vol. **AC-24**, n° 6, 1979, p. 958-962.

(*Manuscrit reçu le* 17 *décembre* 1990.)