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

B. BERCU
M. DUFLO

Moindres carrés pondérés et poursuite

Annales de l'I. H. P., section B, tome 28, nº 3 (1992), p. 403-430
<http://www.numdam.org/item?id=AIHPB_1992__28_3_403_0>

Ann. Inst. Henri Poincaré,

Vol. 28, n° 3, 1992, p. 403-430. *Probabilités et Statistiques*

Moindres carrés pondérés et poursuite

par

B. BERCU et M. DUFLO

U.R.A. n° 743, C.N.R.S., statistique appliquée,
Laboratoire de Statistique, Université Paris-Sud, Bât. 425,
91405 Orsay, France

RÉSUMÉ. — Une loi des grands nombres pondérée sur les martingales vectorielles conduit à introduire un estimateur des moindres carrés pondéré améliorant l'estimateur non pondéré par des majorations en moyennes quadratiques et possédant de bonnes propriétés des erreurs de prédiction; une solution simple du problème de la poursuite d'une trajectoire observable en résulte.

ABSTRACT. — A weighted law of large numbers for vector valued martingales induces the definition of a weighted least squares estimator improving the least squares estimator with bounds for quadratic means and nice properties for the errors of prediction; adaptive tracking problems are then easily solved.

INTRODUCTION

Objectif

Deux questions sont au centre de l'étude des modèles linéaires aléatoires: l'identification et le choix d'un contrôle optimal pour la poursuite d'une trajectoire donnée.

Classification A.M.S. : 62 M, 62 F et J.

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques - 0246-0203
Vol. 28/92/03/403/28/$ 4,80/© Gauthier-Villars

Ce texte s'inscrit dans la longue suite des recherches sur l'estimateur récursif le plus performant. L'estimateur pondéré que nous introduisons cumule les divers avantages des estimateurs du gradient et des moindres carrés, pour les propriétés de consistance et les erreurs de prédiction.

Résumé

On présente dans le paragraphe I, un nouvel estimateur des moindres carrés «pondéré» qui s'avère aussi bon que l'estimateur des moindres carrés pour les problèmes d'identification usuels. On prouve, pour cet estimateur, que les erreurs de prédiction ont les mêmes propriétés que celles de l'estimateur du gradient. Une solution simple du problème de poursuite d'une trajectoire observable en résulte en IV. La consistance est alors obtenue pour des contrôles éventuellement excités. On obtient en outre de bonnes vitesses de convergence presque sûre.

L'outil de base est une loi des grands nombres pondérée relative aux martingales: des résultats, presque sûrs et en moyenne quadratique sont donnés en II et appliqués à la régression en III.

Un estimateur des moindres carrés pondéré généralisé possédant des propriétés analogues est étudié pour les modèles ARMAX passifs dans [3].

Notations

On se place sur $\mathbb{C}^d$ ou sur $\mathbb{C}^\delta$. Pour un vecteur u, on désigne aussi par u la matrice colonne associée. Si R est une matrice, *R est sa conjuguée et le carré de la norme $\| \mathrm{R} \|$ est la trace de *RR. On utilise le produit scalaire hermitien $\langle u, v \rangle = {}^*uv$ et la norme associée $\| . \|$.

Soit $\mathscr{H}$ l'ensemble des matrices $\delta \times \delta$ hermitiennes semi-définies positives. Pour $\mathrm{H} \in \mathscr{H}$, on note Tr H sa trace, Det H son déterminant, $\lambda_{\min} \mathrm{H}$ sa plus petite valeur propre; pour H_1 et H_2 dans $\mathscr{H}$, $\mathrm{H}_1 \leqq \mathrm{H}_2$ signifie que $\mathrm{H}_2 - \mathrm{H}_1 \in \mathscr{H}$.

Les modèles décrits ci-dessous seront toujours définis sur un espace de probabilité $(\Omega, \mathscr{A}, \mathrm{P})$ muni d'une filtration $\mathbb{F} = (\mathscr{F}_n)$; $\mathscr{F}_n$ est la tribu des événements qui se produisent jusqu'à l'instant n.

Pour deux suites de variables aléatoires positives (a_n) et (b_n), $a_n = O(b_n)$ [resp. $o(b_n)$] signifie qu'il existe une suite (c_n) aléatoire bornée [resp. tendant vers 0] telle que $a_n \leqq c_n b_n$.

Enfin, on désigne par $\mathscr{L}$ l'ensemble des fonctions f de $\mathbb{R}^+$ dans $\mathbb{R}^+$ satisfaisant les conditions suivantes:

- f est décroissante et continue; $f(1) \leqq 1$.
- Pour tout $a > 0$, $\displaystyle\int_a^\infty f(t)\, dt < \infty$.

I. UN ESTIMATEUR DES MOINDRES CARRÉS PONDÉRÉ

Définition et propriétés

Pour Γ, matrice de covariance $d \times d$ supposée déterministe et non nulle, un *bruit de covariance* Γ, *adapté à* $\mathbb{F}$, est une suite (ε_n) de vecteurs aléatoires à valeurs dans $\mathbb{C}^d$, adaptée à $\mathbb{F}$, et telle que, pour tout n:

$$\mathrm{E}(\varepsilon_{n+1} \mid \mathcal{F}_n) = 0, \qquad \mathrm{E}(\varepsilon_{n+1}{}^*\varepsilon_{n+1} \mid \mathcal{F}_n) = \Gamma.$$

Un modèle de régression complexe de dimension d associé à ce bruit est une suite (X_n) telle que X_0 soit $\mathcal{F}_0$-mesurable et:

$$\mathrm{X}_n = {}^*\mathrm{R}\,\Phi_{n-1} + \varepsilon_n;$$

(Φ_n) est une suite de vecteurs aléatoires de dimension δ, adaptée à $\mathbb{F}$; R est une matrice $\delta \times d$ à estimer. A l'instant n, Φ_n et X_n sont observés, le bruit est inobservable.

On pose $\mathrm{S}_n = \mathrm{S} + \sum_{k=0}^{n} \Phi_k{}^*\Phi_k$, S matrice hermitienne définie positive et $s_n = \operatorname{tr} \mathrm{S}_n$; on choisit, pour simplifier les constantes rencontrées ci-dessous, S déterministe, $\mathrm{S} \geqq \delta \mathrm{I}$.

Soit $\tau = (\tau_n)$ une suite adaptée à $\mathbb{F}$ de variables aléatoires positives, décroissante et $\leqq 1$; on note τ_∞ sa limite. On propose, pour $n \geqq 1$, l'*estimateur des moindres carrés pondéré par* τ suivant:

$$\hat{\mathrm{R}}_n = \mathrm{T}_{n-1}^{-1} \sum_{k=1}^{n} \tau_{k-1}\, \Phi_{k-1}{}^*\mathrm{X}_k$$

où

$$\mathrm{T}_n = \sum_{k=0}^{n} \tau_k\, \Phi_k{}^*\Phi_k + \mathrm{S}.$$

Sa définition récursive, analogue à celle de l'estimateur non pondéré, en fait un estimateur aisément maniable:
on pose $\hat{\mathrm{R}}_0 = 0$, $\mathrm{T}_{-1} = \mathrm{S}$, et, pour $n \geqq 0$, on a

$$\hat{\mathrm{R}}_{n+1} = \hat{\mathrm{R}}_n + \tau_n \mathrm{T}_n^{-1} \Phi_n{}^*(\mathrm{X}_{n+1} - {}^*\hat{\mathrm{R}}_n \Phi_n),$$

la construction récursive de la suite (T_n^{-1}) étant classique.

Posant $f_n = \tau_n{}^*\Phi_n \mathrm{T}_n^{-1} \Phi_n$, on a l'inégalité classique (voir par exemple [12], p. 50, [18], [23], [28]):

$$f_n \leqq \inf(1, \operatorname{Log}\operatorname{Det} \mathrm{T}_n - \operatorname{Log}\operatorname{Det} \mathrm{T}_{n-1}).$$

La suite τ est une *pondération admissible* lorsque:

$$\Delta = \sum_{n=0}^{\infty} \tau_n f_n \text{ est intégrable.}$$

Nous établirons dans la partie III les théorèmes suivants où

$$\tilde{R}_n = \hat{R}_n - R.$$

On verra en III que les suites τ suivantes sont admissibles, f étant une fonction arbitraire de $\mathscr{L}$ (par exemple, $f(t)=t^{-1-\gamma}$ ou $f(t)=e^{-\gamma t}$, $\gamma>0) : \tau_n = f(\operatorname{Log} s_n)$ ou $\tau_n = f(n)$.

THÉORÈME I.1 (Identification). — *L'estimateur des moindres carrés* $(\hat{R}_n)$, *pondéré par* τ, *possède, lorsque* τ *est admissible, les propriétés suivantes*:

1. *Pour le modèle de régression général décrit ci-dessus,*

(*a*) $\| T_{n-1}^{1/2} \tilde{R}_n \| = O(1)$, p.s.; $E(\| T_{n-1}^{1/2} \tilde{R}_n \|^2) = O(1)$.

(*b*) $E(\sum \alpha_n \| \hat{R}_{n+1} - \hat{R}_n \|^2) < \infty$ *pour la suite* (α_n) *définie par*

$$1/\alpha_n = \sup(\inf[\tau_n, \operatorname{Tr}(T_n^{-1})], \operatorname{Tr}(T_{n-1}^{-1}) - \operatorname{Tr}(T_n^{-1})).$$

La suite (α_n) *est minorée par* 1 *et tend vers* ∞ *si* $\tau_\infty = 0$;

$$\inf(\sup[\tau_n^{-1}, \lambda_{\min} T_n/\delta], \lambda_{\min} T_{n-1}/\delta) \leqq \alpha_n \leqq \sup(\tau_n^{-1}, \lambda_{\min} T_n).$$

2. *Sur* $(\lambda_{\min} T_n \to \infty)$, *l'estimateur est fortement consistant et, p.s.*:

$$\| \tilde{R}_n \| = O(\lambda_{\min} T_{n-1})^{-1/2}.$$

En particulier, sur $(\tau_n \lambda_{\min} S_n \to \infty)$, *l'estimateur est fortement consistant et, p.s.* :

$$\| \tilde{R}_n \| = O(\tau_{n-1}(\lambda_{\min} S_{n-1}))^{-1/2}. \quad \blacksquare$$

THÉORÈME I.2 (Erreurs de prédiction). — *Dans le cadre du théorème précédent, les erreurs de prédiction ont les propriétés suivantes*:

1. *Erreurs* a posteriori

$$E(\sum \tau_n \| {}^*\tilde{R}_{n+1} \Phi_n \|^2) < \infty.$$

2. *Erreurs* a priori

Pour $f_n = \tau_n {}^*\Phi_n T_n^{-1} \Phi_n$,

$$E(\sum \tau_n (1-f_n) \| {}^*\tilde{R}_n \Phi_n \|^2) < \infty.$$

Pour la suite (α_n) *définie dans le théorème* 1, *on a,*

$$E(\sum a_n \| {}^*\tilde{R}_n \Phi_n \|^2) < \infty,$$

avec $a_n = \inf(\tau_n, \alpha_n / \| \Phi_n \|^2)$ *si* $\| \Phi_n \|^2 \neq 0$ *et* $a_n = \tau_n$ *sinon.*

3. *On suppose ici* $1/\tau_n^2 s_n \leqq$ cte, *par exemple* $\tau_n = (\operatorname{Log} s_n)^{-1-\gamma}$ *avec* $\gamma > 0$, *ou* $\tau_n = s_n^{-\gamma}$ *avec* $0 < \gamma \leqq 1/2$.

Pour $\rho_n = \inf(1/\tau_n, \alpha_n) = \inf(1/\tau_n, 1/[\operatorname{Tr}(T_{n-1}^{-1}) - \operatorname{Tr}(T_n^{-1})])$

$$E(\sum (\rho_n / s_n) \| {}^*\tilde{R}_n \Phi_n \|^2) < \infty.$$

En remplaçant ρ_n *par* 1, *il suffit de supposer* $1/(s_n \tau_n) \leqq$ Cte.

Remarque

Pour certains développements, on ajoutera l'une ou l'autre des hypothèses suivantes sur le bruit.

(BB) *Le bruit est un bruit blanc adapté à* $\mathbb{F}$, *suite de vecteurs aléatoires de même loi centrée et de covariance* Γ, *telle que* ε_{n+1} *soit indépendant de* $\mathcal{F}_n$.

(B α) *Le bruit a un moment d'ordre* $a > \alpha$, *c'est-à-dire, p. s.*:

$$\sup_n \mathrm{E}(\|\varepsilon_{n+1}\|^a \mid \mathcal{F}_n) < \infty.$$

(GN) *Le bruit satisfait la loi des grands nombres*:

$$\sigma_n = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k {}^*\varepsilon_k \xrightarrow{\text{p. s.}} \Gamma.$$

(GN) est, en particulier, satisfaite si (BB) ou (B2) l'est.

Commentaires

Comparaison entre l'estimateur pondéré et l'estimateur des moindres carrés usuel

(*A*) *Comparaison avec l'estimateur des moindres carrés lorsque celui-ci est fortement consistant*

Dans le théorème I.1, on déduit la seconde partie de 2 de sa première partie en remarquant que $\lambda_{\min} \mathrm{T}_n \geqq \lambda_{\min} \mathrm{S}_n \tau_n$.

Or, le cadre usuel de consistance ([23], [24], [28], [13], [12]) de l'estimateur des moindres carrés non pondéré correspond à une hypothèse du type $\operatorname{Log} s_n = O(\lambda_{\min} \mathrm{S}_n)$, qui implique, pour toute $f \in \mathcal{L}$, la consistance forte avec vitesse presque sûre:

$$O(f(\operatorname{Log} s_n)\, \lambda_{\min} \mathrm{S}_n)^{-1/2} \text{ [et sous (B2), } O(\operatorname{Log} s_n/\lambda_{\min} \mathrm{S}_n)^{1/2}].$$

Avec l'estimateur pondéré par $\tau_n = f(\operatorname{Log} s_n)$, on obtient, sous les mêmes hypothèses, une vitesse presque sûre un peu plus faible, $O(f(\operatorname{Log} s_n)\, \lambda_{\min} \mathrm{S}_n)^{-1/2}$ pour la fonction f utilisée dans la pondération (dont on a le choix); mais on acquiert une majoration en moyenne quadratique de $\tilde{\mathrm{R}}_n$ et de meilleures propriétés des erreurs de prédiction. ■

(*B*) *Comparaison avec l'estimateur des moindres carrés* (*modèles AR*)

Considérons un modèle autorégressif vectoriel,

$$\mathrm{X}_n = \mathrm{A}_1 \mathrm{X}_{n-1} + \mathrm{A}_2 \mathrm{X}_{n-2} + \ldots + \mathrm{A}_p \mathrm{X}_{n-p} + \varepsilon_n. = {}^*\mathrm{R}\, \Phi_{n-1} + \varepsilon_n,$$

où ${}^t\Phi_{n-1} = ({}^t\mathrm{X}_{n-1}, \ldots, {}^t\mathrm{X}_{n-p})$, ${}^*\mathrm{R} = [\mathrm{A}_1, \ldots, \mathrm{A}_2]$, $\delta = dp$.

Si le bruit satisfait (BB) ou (B2) et si le modèle est « régulier » (en particulier en dimension 1), alors l'estimateur des moindres carrés pondéré est consistant avec, sur les vitesses presque sûres, la même perte qu'en (A) par rapport à celles obtenues en [14] (ou [12], chapitre V) mais avec les avantages évoqués en (A) : on le vérifie facilement.

Rappelons que, pour $p=1$, le cas « singulier » correspond à l'existence pour $A=A_1$ d'un sous-espace propre de dimension >1 associé à une valeur propre de module >1. Alors, on sait que les deux suites $(\lambda_{\min} S_n)$ et $(\operatorname{Log} s_n)$ sont de l'ordre de n ([14], ou [12], p. 217) : il existe un $u \in \mathbb{C}^\delta$ tel que ${}^*u\, S_n\, u = O(n)$.

Pour $\tau_n = (\operatorname{Log} s_n)^{-1-\gamma}$ qui est ici de l'ordre de $n^{-1-\gamma}$:

$$ {}^*u\, T_n\, u = 0\left(\sum_{k=1}^{n} k^{-1-\gamma} |\langle u, \Phi_k \rangle|^2\right) = O\left(\sum_{k=1}^{n} ({}^*u S_k u)^{-1-\gamma} |\langle u, \Phi_k \rangle|^2\right) = O(1); $$

$$ \lambda_{\min} T_n = O(1). $$

L'estimateur pondéré n'est pas plus consistant que l'estimateur des moindres carrés. ■

(C) *Cas des modèles ARX(p, q)*

On considère un modèle $ARX_{d,m}(p, q)$, complexe :

$$ X_n = A_1 X_{n-1} + A_2 X_{n-2} + \ldots + A_p X_{n-p} + B_1 U_{n-1} + \ldots + B_q U_{n-q} + \varepsilon_n, $$

$$ X_n = {}^*R\, \Phi_{n-1} + \varepsilon_n, $$

où ${}^t\Phi_{n-1} = ({}^tX_{n-1}, \ldots, {}^tX_{n-p}, {}^tU_{n-1}, \ldots, {}^tU_{n-q})$, $\delta = dp + mq$.

La suite de vecteurs aléatoires de dimension m, (U_n), est un contrôle adapté à $\mathbb{F}$. Les matrices $d \times d$, $A_1, \ldots, A_p$, et les matrices $d \times m$, $B_1, \ldots, B_q$ sont inconnues;

$$ {}^*R = [A_1, \ldots, A_p, B_1, \ldots, B_q]. $$

Lorsque A_p est inversible et le modèle est irréductible, on sait ([24], ou [12], p. 103-109) choisir des contrôles suffisamment excités pour identifier le modèle par l'estimateur des moindres carrés; la pondération ne s'impose donc pas si l'on se contente d'identifier. ■

Erreurs de prédiction

(D) *Estimateur du gradient*

L'estimateur du gradient sert surtout du fait de la propriété des erreurs *a priori* établie dans [15] que nous retrouvons ici pour l'estimateur pondéré par une pondération admissible. Les démonstrations fondées sur cette relation pourront donc être reprises, on le verra en IV.

Un résultat analogue est prouvé en [17] pour l'estimateur des moindres carrés d'un modèle ARMAX, mais seulement sous l'hypothèse (B2) et avec le contrôle de poursuite qui sera étudié en IV. ■

(*E*) *Covariances empiriques*

Pour estimer la covariance Γ du bruit, on propose les estimateurs empiriques suivants :

– *covariance empirique* a priori :

$$\hat{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - {}^*\hat{R}_{k-1}\Phi_{k-1})\,{}^*(X_k - {}^*\hat{R}_{k-1}\Phi_{k-1});$$

– *covariance empirique* a posteriori :

$$\check{\Gamma}_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - {}^*\hat{R}_k \Phi_{k-1})\,{}^*(X_k - {}^*\hat{R}_k \Phi_{k-1}).$$

On suppose que le bruit satisfait (GN). Posons $\pi_n = -{}^*\tilde{R}_n \Phi_n$. On sait (*cf.* [12], p. 76) que :

$$\hat{\Gamma}_n \to \Gamma \text{ presque sûrement sur } \left\{ \frac{1}{n} \sum_{k=1}^{n} \|\pi_k\|^2 \to 0 \right\}.$$

De plus, lorsque $1/s_n \tau_n^2 \leqq \mathrm{Cte}$, [ou, en ôtant le coefficient ρ_n, $1/s_n \tau_n \leqq \mathrm{Cte}$], p. s. :

$$\sum \| {}^*\hat{R}_n \Phi_n - {}^*\hat{R}_{n+1} \Phi_n \|^2 \rho_n/s_n \leqq \sum \| \hat{R}_n - \hat{R}_{n+1} \|^2 \| \Phi_n \|^2 \rho_n/s_n < \infty,$$

et, si $s_\infty = \infty$, $\sum_{k=1}^{n} \| {}^*\hat{R}_k \Phi_k - {}^*\hat{R}_{k+1} \Phi_k \|^2 \rho_k = o(s_n)$.

En utilisant l'estimateur des moindres carrés pondéré par τ_n *avec, p.s.,* $1/s_n \tau_n \leqq \mathrm{Cte}$, *les conditions* $s_\infty = \infty$ *et* $s_n = O(n)$ *assurent la consistance forte des covariances empiriques* a priori *ou* a posteriori.

Ce résultat n'exige pas de condition sur $\lambda_{\min} S_n$ contrairement à celui que l'on obtient ([13], [12], p. 76) à partir de l'estimateur des moindres carrés. Il s'applique, par exemple, dans le cadre de la poursuite décrit en IV. ■

II. LOI DES GRANDS NOMBRES PONDÉRÉE POUR LES MARTINGALES

On donne ici une loi des grands nombres sur les martingales vectorielles, analogue à la loi des grands nombres classique ([18], [23], [28], [13], [12]-p. 56) mais plus précise.

Soit une martingale vectorielle de carré intégrable adaptée à $\mathbb{F}$, complexe de dimension δ, $M=(M_n)$ dont la variation quadratique prévisible $\langle M\rangle=(\langle M\rangle_n)$, suite croissante à valeurs dans $\mathcal{H}$, est définie par

$$\langle M\rangle_0=0 \text{ et } \langle M\rangle_{n+1}-\langle M\rangle_n=E(M_{n+1}{}^*M_{n+1}-M_n{}^*M_n \mid \mathcal{F}_n).$$

Motivation

Rappelons la « loi des grands nombres » classique relative à une martingale à valeurs dans $\mathbb{R}$, de carré intégrable. Pour $\gamma>0$, on a, p. s. :

$$\langle M\rangle_n^{-1} M_n^2=O(\inf(n, \operatorname{Log}\langle M\rangle_n)^{1+\gamma}).$$

Remarquons que : $\lim_{T\to\infty}[T(\operatorname{Log} T)^{1+\gamma}]^{-1}\int_2^T(\operatorname{Log} t)^{1+\gamma}\,dt=1$. Par suite,

$$V_{n-1}=\sum_{k=1}^{n}(\operatorname{Log}(2+\langle M\rangle_k))^{1+\gamma}(\langle M\rangle_k-\langle M\rangle_{k-1})\geqq\int_{2+\mathrm{Cte}}^{\langle M\rangle_n}(\operatorname{Log} x)^{1+\gamma}\,dx$$

et $(\operatorname{Log}\langle M\rangle_n))^{1+\gamma}\langle M\rangle_n=O(V_{n-1})$.

D'où, p. s. : $V_{n-1}^{-1}M_n^2=O(1)$.

La loi des grands nombres pondérée suivante étend et précise cette remarque lorsque $\delta>1$. ■

Voici d'abord un énoncé général; les corollaires spécifieront les choix convenables de τ_n. Dans ce qui suit la matrice Q est choisie de manière à formuler les résultats le plus simplement possible.

THÉORÈME II.3 (Loi des grands nombres pondérée). — *On donne*

— *une suite aléatoire* $(V_n)_{n\geqq-1}$ *à valeurs dans* $\mathcal{H}$ *croissante et adaptée à* $\mathcal{F}$, $V_{-1}=Q$ *étant choisie déterministe et définie positive*, $Q\geqq I$;

— *une suite de v. a. positives* $\leqq 1$, (τ_n), *décroissant vers* τ_∞, *adaptée à* $\mathbb{F}$ *et satisfaisant* :

$$\textit{pour } f_n=\operatorname{Tr} V_n^{-1/2}(V_n-V_{n-1})V_n^{-1/2},\ \sum_{n=0}^{\infty}\tau_n f_n\leqq\Delta \textit{ avec } E(\Delta)<\infty.$$

On suppose que $\langle M\rangle_{n+1}-\langle M\rangle_n\leqq\tau_n(V_n-V_{n-1})$.

On a alors les propriétés suivantes.

1. $\|V_{n-1}^{-1/2}M_n\|^2$ *converge, p. s., vers une v. a. finie*;

$$E\left(\|V_{n-1}^{-1/2}M_n\|^2+\sum_{k=0}^{n-1}{}^*M_k(V_{k-1}^{-1}-V_k^{-1})M_k\right)$$
$$\leqq E(\Delta)+E(\|M_0\|^2)(\lambda_{\min}Q)^{-1}.$$

2. *Soit* (α_n) *la suite définie par*

$$1/\alpha_n=\sup(\inf(\tau_n, \operatorname{Tr} V_n^{-1}), \operatorname{Tr}(V_{n-1}^{-1}-V_n^{-1}));$$
$$\inf(\sup(\tau_n^{-1}, \lambda_{\min}V_n/\delta), \lambda_{\min}V_{n-1}/\delta)\leqq\alpha_n\leqq\sup(\tau_n^{-1}, \lambda_{\min}V_n);$$

(α_n) *tend vers* ∞ *si* $\tau_\infty = 0$ *et l'on a*:

$$E\left(\sum_{n=0}^{\infty} \alpha_n \| V_n^{-1} M_{n+1} - V_{n-1}^{-1} M_n \|^2\right) < \infty.$$

Démonstration. — (*a*) La preuve de 1 est une application simple du théorème de Robbins-Siegmund à la suite positive,

$$W_n = \| V_{n-1}^{-1/2} M_n \|^2 = {}^*M_n V_{n-1}^{-1} M_n.$$

Les outils utilisés sont les propriétés des matrices hermitiennes dégagées dans la théorie classique:

$$E(W_{n+1} \mid \mathscr{F}_n) \leqq W_n + \zeta_n - \eta_n$$

où $\zeta_n = \operatorname{Tr} V_n^{-1/2} [\langle M \rangle_{n+1} - \langle M \rangle_n] V_n^{-1/2}$, $\eta_n = {}^*M_n [V_{n-1}^{-1} - V_n^{-1}] M_n$.

La matrice $V_{n-1}^{-1} - V_n^{-1}$ est dans $\mathscr{H}$, et $\eta_n \geqq 0$.

On a: $0 \leqq \zeta_n \leqq \tau_n \operatorname{Tr} (V_n^{-1/2} [V_n - V_{n-1}] V_n^{-1/2}) = \tau_n f_n$.

Donc, $\sum_{n=0}^{\infty} \zeta_n \leqq \Delta$: (W_n) converge p. s. et $\sum \eta_n < \infty$, p. s.

De plus: $E\left(W_{n+1} - W_0 + \sum_{k=0}^{n} \eta_k\right) \leqq E(\Delta)$.

(*b*) Montrons 2.

Soit $J_n = (V_{n-1}^{-1} - V_n^{-1})^{1/2}$, racine carrée prise dans $\mathscr{H}$.

On a, d'après 1: $E\left(\sum_{n=0}^{\infty} \| J_n M_n \|^2\right) < \infty$.

Soit (α_n) une suite positive adaptée. Pour tout n, on a:

$$\begin{aligned}\sum_{n=0}^{\infty} \alpha_n \| V_n^{-1} M_{n+1} - V_{n-1}^{-1} M_n \|^2 & \\ \leqq 2 \sum_{n=0}^{\infty} \alpha_n \| V_n^{-1} (M_{n+1} - M_n) \|^2 & \\ + 2 \sum_{n=0}^{\infty} \alpha_n \| (V_{n-1}^{-1} - V_n^{-1}) M_n \|^2 &= 2(G+H).\end{aligned}$$

Or:

$$E(\alpha_n \| V_n^{-1} (M_{n+1} - M_n) \|^2 \mid \mathscr{F}_n) \leqq \operatorname{Tr} V_n^{-1} (V_n - V_{n-1}) V_n^{-1} \alpha_n \tau_n.$$

Si $\alpha_n \operatorname{Tr} V_n^{-1} \leqq \mathrm{Cte}$, cette expression est majorée par $\mathrm{Cte}\, \tau_n f_n$ et la série G converge, p. s. et dans L^1.

Il en va de même si $\alpha_n \tau_n \leqq \mathrm{Cte}$, car, dans $\mathscr{H}$, on a l'inégalité suivante (*cf.* par exemple [12], p. 50):

$$\sum_{n=0}^{\infty} V_n^{-1} (V_n - V_{n-1}) V_n^{-1} \leqq Q^{-1}.$$

Si, $\alpha_n \operatorname{Tr}(V_{n-1}^{-1} - V_n^{-1}) \leqq \text{Cte}$:

$$H = \sum_{n=0}^{\infty} {}^*(J_n M_n)(V_{n-1}^{-1} - V_n^{-1})(J_n M_n)\lambda_n \leqq \text{Cte} \sum_{n=0}^{\infty} \| J_n M_n \|^2;$$

la série H converge p. s. et dans L^1.

Ainsi, dès que $\alpha_n \sup(\inf(\tau_n, \operatorname{Tr} V_n^{-1}), \operatorname{Tr}(V_{n-1}^{-1} - V_n^{-1})) \leqq \text{Cte}$,

$$\sum_{n=0}^{\infty} \alpha_n \| V_n^{-1} M_{n+1} - V_{n-1}^{-1} M_n \|^2 \text{ est intégrable.}$$

Enfin, on remarque que:

$$\lambda_{\min} V_n/\delta \leqq 1/\operatorname{Tr} V_n^{-1} \leqq \lambda_{\min} V_n,$$

et que la suite (V_n^{-1}) décroît toujours dans $\mathscr{H}$ et a donc une limite V_∞^{-1} ce qui assure la convergence de (α_n) vers ∞ lorsque $\tau_\infty = 0$.

Nous utiliserons en fait la transcription suivante du théorème II.3, où la suite (Y_n) n'est pas nécessairement de carré intégrable. Un énoncé commun serait aisé à formuler mais d'une lecture plus difficile.

Proposition II.4. — *Les suites (V_n) et (τ_n) étant définies comme dans le théorème II.3, on considère une suite de vecteurs aléatoires (Y_n) de dimension δ adaptée à $\mathbb{F}$ et un bruit unidimensionnel (ε_n) de variance $\sigma^2 > 0$, adapté à $\mathbb{F}$. On suppose que, pour tout n:*

$$\tau_n Y_n {}^*Y_n \leqq V_n - V_{n-1}.$$

Soit $M_n = \sum_{k=0}^{n-1} \tau_k Y_k \varepsilon_{k+1}$.

(*a*) *Les vecteurs aléatoires $V_{n-1}^{-1/2} M_n$ sont de carrés intégrables.*

(*b*) *Les résultats énoncés dans le théorème II.3 sont valables en remplaçant $\Delta = \sum_{n=0}^{\infty} \tau_n f_n$ par $\sigma^2 \Delta$.*

Démonstration. — (*a*) En comparant les formes quadratiques associées, on obtient, dans $\mathscr{H}$:

$$M_n {}^*M_n \leqq \sum_{k=1}^{n} \tau_{k-1}^2 Y_{k-1} {}^*Y_{k-1} \sum_{k=1}^{n} |\varepsilon_k|^2 \leqq V_{n-1} \sum_{k=1}^{n} |\varepsilon_k|^2;$$

$$\operatorname{Tr} V_{n-1}^{-1/2} M_n {}^*M_n V_{n-1}^{-1/2} \leqq \delta \sum_{k=1}^{n} |\varepsilon_k|^2 \text{ est intégrable.}$$

(*b*) On en déduit facilement :

$$
\begin{aligned}
\mathrm{E}(\mathrm{V}_n^{-1/2}(\mathrm{M}_{n+1}{}^*\mathrm{M}_{n+1}-\mathrm{M}_n{}^*\mathrm{M}_n)\mathrm{V}_n^{-1/2}\mid\mathcal{F}_n)&\\
=\mathrm{E}(\mathrm{V}_n^{-1/2}(\mathrm{M}_{n+1}-\mathrm{M}_n)^*(\mathrm{M}_{n+1}-\mathrm{M}_n)\mathrm{V}_n^{-1/2}\mid\mathcal{F}_n)&\\
=\sigma^2\tau_n^2\mathrm{V}_n^{-1/2}\mathrm{Y}_n{}^*\mathrm{Y}_n\mathrm{V}_n^{-1/2}&\\
\leqq\sigma^2\tau_n\mathrm{V}_n^{-1/2}(\mathrm{V}_n-\mathrm{V}_{n-1})\mathrm{V}_n^{-1/2}&.
\end{aligned}
$$

La démonstration du théorème II.3 s'applique désormais au cadre de la proposition II.4. ■

COROLLAIRE II.5. — *Soit une fonction* $f\in\mathcal{L}$. *Le théorème II.3 et la proposition II.4 s'appliquent dans les deux cas suivants* :

(*a*) $\tau_0=1$, $\tau_n=f(n)$ *pour* $n\geqq 1$, *et* $\Delta=\delta\left(2+\int_1^\infty f(x)\,dx\right)$.

(*b*) *On suppose données une constante c et une suite croissante* (t_n) *de v. a. positives adaptée à* $\mathbb{F}$ *et satisfaisant pour tout n* :

$$\mathrm{Tr}\,\mathrm{V}_n\leqq t_n.$$

On choisit $\tau_n=f(\mathrm{Log}\,t_n)$.

Démonstration. — Ce corollaire résulte de l'inégalité :

$$\mathrm{Tr}(\mathrm{V}_n^{-1/2}(\mathrm{V}_n-\mathrm{V}_{n-1})\mathrm{V}_n^{-1/2})\leqq\inf(\delta,\ \mathrm{Log\,Det}\,\mathrm{V}_n-\mathrm{Log\,Det}\,\mathrm{V}_{n-1}).$$

Prouvons la partie (*b*); la partie (*a*) se démontre plus simplement à partir de la majoration par δ :

$$
\begin{aligned}
\tau_n f_n&\leqq f(\mathrm{Log}\,t_n)(\mathrm{Log\,Det}\,\mathrm{V}_n-\mathrm{Log\,Det}\,\mathrm{V}_{n-1})\\
&\leqq f(\mathrm{Log\,Tr}\,\mathrm{V}_n)(\mathrm{Log\,Det}\,\mathrm{V}_n-\mathrm{Log\,Det}\,\mathrm{V}_{n-1})\\
&\leqq f(\mathrm{Log\,Det}\,\mathrm{V}_n/\delta)(\mathrm{Log\,Det}\,\mathrm{V}_n-\mathrm{Log\,Det}\,\mathrm{V}_{n-1});
\end{aligned}
$$

$$\sum_{n=0}^{\infty}\tau_n f_n\leqq\int_{\mathrm{Log\,Det\,Q}}^{\infty}f(x/\delta)\,dx\leqq\delta\int_{\mathrm{Log\,Det\,Q}/\delta}^{\infty}f(x)\,dx=\Delta<\infty.\ ■$$

III. DÉMONSTRATION DES THÉORÈMES I.1.2

On se place dans le cadre de I. On a :

$$\hat{\mathrm{R}}_n-\mathrm{R}=\tilde{\mathrm{R}}_n=\mathrm{T}_{n-1}^{-1}[\mathrm{M}_n-\mathrm{SR}]$$

où $\mathrm{M}_n=\sum_{k=1}^{n}\tau_{k-1}\Phi_{k-1}{}^*\varepsilon_k$.

Considérons $\mathrm{M}u=(\mathrm{M}_n u)$ avec $u\in\mathbb{C}^d$. Pour ${}^*u\Gamma u\neq 0$, la proposition II.4 s'applique ici à $\mathrm{V}_n=\mathrm{T}_n$, $\sigma^2={}^*u\Gamma u$, et :

$$\sum_{n=0}^{\infty}\tau_n\,\mathrm{Tr}\,\mathrm{T}_n^{-1/2}(\mathrm{T}_n-\mathrm{T}_{n-1})\mathrm{T}_n^{-1/2}=\sum_{n=0}^{\infty}\tau_n^2\,{}^*\Phi_n\mathrm{T}_n^{-1}\Phi_n=\Delta.$$

Or les suites $(T_n^{-1/2})$ et (T_n^{-1}), décroissantes dans $\mathscr{H}$, ont toujours une limite $T_\infty^{-1/2}$ et T_∞^{-1} et sont majorées par une matrice constante. Donc $(T_{n-1}^{-1/2}\,SR)$ et $(T_{n-1}^{-1}\,SR)$ sont des suites bornées et, p. s., convergentes.

Pour $f \in \mathscr{L}$, le corollaire II.5 permet de choisir $\tau_n = f(n)$. En outre, on a $\operatorname{Tr} T_n \leqq \operatorname{Tr} S_n$ et l'on peut choisir $\tau_n = f(\operatorname{Log} s_n)$.

Le théorème I.1 en résulte. Prouvons le théorème I.2.

Pour les erreurs *apriori*, on a :

$$\begin{aligned}(1-f_n)\,\tau_n\,({}^*\tilde{R}_n\,\Phi_n)\,{}^*({}^*\tilde{R}_n\,\Phi_n) &= {}^*(M_n - SR)\,(T_{n-1}^{-1} - T_n^{-1})\,(M_n - SR)\\ &\leqq 2\,{}^*M_n\,(T_{n-1}^{-1} - T_n^{-1})\,M_n + 2\,{}^*(SR)\,(T_{n-1}^{-1} - T_n^{-1})\,(SR).\end{aligned}$$

D'où : $E(\sum (1-f_n)\,\tau_n \,\| {}^*\tilde{R}_n\,\Phi_n \|^2) < \infty$.

Grâce à la relation récursive,

$$\begin{aligned}\tilde{R}_{n+1} &= \tilde{R}_n + \tau_n\,T_n^{-1}\,\Phi_n\,{}^*(\varepsilon_{n+1} - {}^*\tilde{R}_n\,\Phi_n),\\ {}^*\Phi_n\,\tilde{R}_{n+1} &= (1-f_n)\,{}^*\Phi_n\,\tilde{R}_n + f_n\,{}^*\varepsilon_{n+1};\end{aligned}$$

comme $E(\sum \tau_n\,f_n \,\| \varepsilon_{n+1} \|^2) = \operatorname{Tr} \Gamma\, E(\Delta)$, on obtient :

$$E(\sum \tau_n \,\| {}^*\tilde{R}_{n+1}\,\Phi_n \|^2) < \infty.$$

Enfin :

$$\begin{aligned}\| {}^*\tilde{R}_n\,\Phi_n \|^2 &= \| {}^*(\tilde{R}_n - \tilde{R}_{n+1} + \tilde{R}_{n+1})\,\Phi_n \|^2\\ &\leqq 2\,(\| {}^*\tilde{R}_{n+1}\,\Phi_n \|^2 + \| \hat{R}_n - \hat{R}_{n+1} \|^2 \,\| \Phi_n \|^2);\end{aligned}$$

et pour $a_n = \inf(\tau_n, \alpha_n / \| \Phi_n \|^2)$ si $\| \Phi_n \| \neq 0$, $a_n = \tau_n$ sinon,

$$E(\sum a_n \,\| {}^*\tilde{R}_n\,\Phi_n \|^2) < \infty.$$

En particulier, pour $\rho_n = \inf([1/\tau_n], \alpha_n)$, lorsque $1/[s_n\,\tau_n^2] \leqq \mathrm{Cte}$,

$$\rho_n / s_n \leqq 1/(\tau_n\,s_n) \leqq \mathrm{Cte}\,\tau_n,$$

et

$$E(\sum [\rho_n / s_n] \,\| {}^*\tilde{R}_n\,\Phi_n \|^2) < \infty;$$

en remplaçant ρ_n par 1, la condition $1/[s_n\,\tau_n] \leqq \mathrm{Cte}$ suffit. ■

Remarque. — La convergence, p. s., de la série $\sum [\rho_n / s_n] \,\| {}^*\tilde{R}_n\,\Phi_n \|^2$ est assurée dès que $(1/s_n) = O(\tau_n^2)$ p. s. [ou $(1/s_n) = O(\tau_n)$, en remplaçant ρ_n par 1]. ■

IV. POURSUITE

IV.1. Contrôle de poursuite d'un modèle de régression

Généralités sur la poursuite

Dans tout le paragraphe IV, $z = (z_n)$ est une trajectoire prévisible pour $\mathbb{F}$ qu'il s'agit de poursuivre par $X = (X_n)$, modèle contrôlé. Les modèles étudiés sont toujours des modèles vectoriels complexes de dimension d qui se mettent sous la forme :

$$X_{n+1} - z_{n+1} = \pi_n + \varepsilon_{n+1}.$$

La suite (ε_n) est un bruit de dimension d, de covariance Γ, adapté à $\mathbb{F}$ satisfaisant l'hypothèse (GN).

La suite (π_n) des *erreurs de prédiction*, adaptée à $\mathbb{F}$, est fabriquée par les méthodes usuelles de contrôle adaptatif.

Le coût de la poursuite à l'instant n se mesure par la forme quadratique associée à la matrice :

$$C_n = \sum_{k=1}^{n} [X_k - z_k]^* [X_k - z_k].$$

Avec la notation $\sigma_n = \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k {}^*\varepsilon_k$, il résulte de la loi des grands nombres sur les martingales réelles si la suite (π_n) est de carré intégrable et d'un raisonnement analogue à la proposition II.4 sinon, que, pour tout $u \in \mathbb{C}^d$,

$$^*u C_n u = {}^*u \sum_{k=1}^{n} \pi_{k-1} {}^*\pi_{k-1} u (1 + \gamma_n) + n^* u \sigma_n u,$$

où (γ_n) converge p. s., la limite étant nulle lorsque

$$^*u \sum_{n=1}^{\infty} \pi_n {}^*\pi_n u = \infty.$$

Donc, p. s. : $\| C_n - n \sigma_n \| = O\left(\sum_{k=1}^{n} \| \pi_{k-1} \|^2 \right)$ et $\liminf \frac{1}{n} C_n \geqq \Gamma$. Le contrôle est optimal si la borne inférieure est une limite p. s., c'est-à-dire si, et seulement si, $\frac{1}{n} \sum_{k=1}^{n} \| \pi_{k-1} \|^2 \xrightarrow{\text{p. s.}} 0$.

Contrôle de poursuite

Considérons un modèle de régression

$$X_{n+1} = {}^*R \, \Phi_n + {}^*H \, \Psi_n + \varepsilon_{n+1};$$

les vecteurs aléatoires Φ_n et Ψ_n sont observables à l'instant n et dépendent d'un contrôle U_n, dont on a le choix à l'instant n en fonction des observations antérieures; H est connue.

Étant donné un estimateur $\hat{R}_n$ de R, un *contrôle de poursuite* $U(z)$ de la trajectoire z est un contrôle satisfaisant :

$$z_{n+1} = {}^*\hat{R}_n \Phi_n + {}^*H \, \Psi_n;$$
$$X_{n+1} - z_{n+1} = - {}^*\tilde{R}_n \Phi_n + \varepsilon_{n+1}.$$

On suppose que $\hat{R}_n$ est un estimateur des moindres carrés pondéré du modèle $(X_n - {}^*H\,\Psi_{n-1})$.

On définit les covariances empiriques comme dans le commentaire E) de la partie I, pour le modèle $(X_n - {}^*H\,\Psi_{n-1})$.

La covariance empirique *a priori* $\hat{\Gamma}_n$ coïncide ici avec le coût moyen. Si $\check{\Gamma}_n$ est la covariance empirique *a posteriori*,

$$\|\check{\Gamma}_n - \hat{\Gamma}_n\| = O\left(\frac{1}{n}\sum_{k=1}^{n} \|{}^*\hat{R}_{k-1}\,\Phi_{k-1} - {}^*\hat{R}_k\,\Phi_{k-1}\|^2\right)$$

et, d'après le théorème I.1, on a, p. s. :

$$\sum_{k=1}^{n} \|{}^*R_{k-1}\,\Phi_{k-1} - {}^*R_k\,\Phi_{k-1}\|^2\,\alpha_{k-1}/\|\Phi_{k-1}\|^2 < \infty.$$

Les propriétés asymptotiques des covariances empiriques seront ainsi des conséquences simples de celles du coût moyen.

On déduit des théorèmes I.1 et I.2 les propriétés générales suivantes du contrôle de poursuite qui seront appliquées par la suite au modèle $\mathrm{ARX}_{d,m}(p, q)$ mais sont aussi valables pour d'autres modèles de régression.

PROPOSITION IV.6 (Critères d'optimalité et de consistance). — *Dans le cadre général décrit ci-dessus, on utilise le contrôle de poursuite* $U(z)$, *et une pondération* $\tau_n = f(\mathrm{Log}\, s_n)$, *avec* :

$$f(t) = t^{-1-\gamma},\ \gamma > 0, \qquad ou \qquad f(t) = e^{-\gamma t},\quad 0 < \gamma \leqq 1/2.$$

On a les propriétés suivantes :

1. *Le contrôle de poursuite est optimal si, et seulement si, p. s.,*

$$s_\infty = \infty, \qquad s_n = O(n).$$

Alors, p. s. : $\sum \frac{1}{n}\|X_n - z_n - \varepsilon_n\|^2 < \infty$.

2. *Soit* $\lambda = (\lambda_n)$ *une suite positive adaptée croissant vers* ∞. *On suppose que, p. s.,* (λ_{n+1}) *et* (λ_n) *sont du même ordre,* (λ_n/n) *est décroissante et* $[f(\mathrm{Log}\, n)]^{-2} = O(\lambda_n)$.

On fait sur le vecteur de régression les hypothèses suivantes :

— *Les suites* s_n *et n sont, p. s., du même ordre;*

— *La suite* (Φ_n) *est « λ-excitante », c'est-à-dire, p. s.,*

$$\liminf (1/\lambda_n)\,\lambda_{\min} \sum_{k=0}^{n} \Phi_k\,{}^*\Phi_k > 0.$$

Alors, l'estimateur est fortement consistant;

$$\|\hat{R}_n - R\|^2 = O(1/\lambda_n\, f(\mathrm{Log}\, n)),\ \text{p. s.}$$

En outre, p. s. :

$$\sum \|X_n - z_n - \varepsilon_n\|^2\,[1/n f(\mathrm{Log}\, n)] < \infty.$$

On a donc une vitesse pour l'optimalité: *p. s.*,

$$\left\|\frac{1}{n}C_n - \sigma_n\right\| = o\,[f(\operatorname{Log} n)].$$

3. *Soit* (v_n) *une suite positive, croissant vers* ∞, *telle que*, p. s., $\|\Phi_n\|^2 = O(v_n)$.

On suppose, s_n, *p. s., de l'ordre de n et* $v_n = O(n)$.

Alors, *p. s.*:

$$\sum \|X_n - z_n - \varepsilon_n\|^2 \inf(f(\operatorname{Log} n),\ 1/v_n) < \infty.$$
$$\left\|\frac{1}{n}C_n - \sigma_n\right\| = o\,(\sup[1/f(\operatorname{Log} n),\ v_n]/n).$$

Dans le cadre de 2, *si* $v_n = O(\lambda_n)$, *on a*, *p. s.*:

$$\sum \|X_n - z_n - \varepsilon_n\|^2 f(\operatorname{Log} n) < \infty,$$
$$\left\|\frac{1}{n}C_n - \sigma_n\right\| = o\,(1/nf(\operatorname{Log} n)).$$

(*a*) *Dans le cadre de* 1, *les estimateurs empiriques de la covariance* Γ *du bruit*, a priori *et* a posteriori, $\hat{\Gamma}_n$ *et* $\tilde{\Gamma}_n$, *sont fortement consistants.*

Les vitesses obtenues dans les parties 2 *et* 3 *pour la convergence p. s. de* $\left\|\frac{1}{n}C_n - \sigma_n\right\|$ *sont aussi celles de* $\|\hat{\Gamma}_n - \sigma_n\|$ *et de* $\|\tilde{\Gamma}_n - \sigma_n\|$.

Démonstration. — (*a*) Dans le théorème I.2, on a, si $s_n = O(n)$ et $s_\infty = \infty$:

$$\alpha_n \geqq 1, \qquad \rho_n \geqq 1, \qquad \inf(f(\operatorname{Log} n),\ 1/\|\Phi_n\|^2) = O(a_n).$$

D'où la partie 1 et le début de la partie 3.

(*b*) Si (s_n) est une suite du même ordre que n, (τ_n) est du même ordre que $f(\operatorname{Log} n)$. La propriété de l'estimateur résulte de la partie 2 du théorème I. 1.

Dans la partie 3 du théorème I.2, on a ici, $\lambda_n f(\operatorname{Log} n) = O(\alpha_n)$,

$$1/f(\operatorname{Log} n) = O(\rho_n),$$
$$f(\operatorname{Log} n) = O(a_n).$$

D'où la fin des parties 2 et 3.

(*c*) La partie 4 est alors aisée. ■

Commentaire

Les vitesses de convergence presque sûre vers Γ dépendent des hypothèses faites sur le bruit.

Ainsi, par le théorème de Chow, sous l'hypothèse B 2 a, $a > 1$, dès que (t_n) est une suite croissante telle $\sum t_n^{-b} < \infty$ pour tout $b > a$, la suite $\sum_{k=1}^{n} [1/t_k](\varepsilon_k^* \varepsilon_k - \Gamma)$ est, p. s., convergente.

Alors, p. s. : $\|\sigma_n - \Gamma\| = o(t_n/n)$.

IV.2. Poursuite pour un modèle ARX : optimalité

Historique

Rappelons les grandes étapes des travaux réalisés autour du thème de la poursuite. Le cadre en est souvent un modèle ARMAX, alors que nous nous contentons ici de modèles ARX; mais l'extension de notre méthode aux modèles ARMAX est menée à bien en [3].

L'idée du contrôle de poursuite vient de [1]. Avec l'estimateur du gradient, Goodwin, Ramdage et Caines ([16], [5]) prouvent l'optimalité du contrôle de poursuite. Pour le problème scalaire d'ajustement à une cible nulle, Becker, Kumar et Wei ([2], [21]) prouvent la convergence de l'estimateur vers un multiple du paramètre. Kumar et Praly [20] prouvent, dans le cas scalaire, la consistance sous une hypothèse d'excitation de la trajectoire.

Caines [4] force la consistance par un contrôle de poursuite excité. Divers travaux exploitent cette idée (avec l'estimateur du gradient) et obtiennent la consistance et la stabilité du système ([6], [7], [8]). Avec de fortes hypothèses sur le bruit, Chen et Guo [10] modifient cette excitation et parviennent à conjuguer optimalité et consistance.

Avec l'estimateur des moindres carrés, Kumar [19] parvient à des résultats analogues à [2] ou [20], lorsque le bruit est gaussien et pour presque toutes valeurs du paramètre. Lai et Wei ([26], [27]) bâtissent, dans le cas réel et pour un bruit borné, un contrôle de poursuite compliqué, assurant optimalité et consistance.

Puis, Sin et Goodwin [29] prouvent l'optimalité pour un algorithme des moindres carrés modifié. Pour une autre modification, Chen [7] prouve, avec un contrôle de poursuite excité, la stabilité et la consistance.

Avec l'estimateur des moindres carrés et un contrôle compliqué, Chen, Guo et Zhang ([9], [11]) donnent des vitesses presque sûres. Enfin, un article récent de Guo et Chen [17] établit, lorsque le bruit a un moment d'ordre >2, le résultat longtemps attendu : la consistance de l'estimateur des moindres carrés et l'optimalité du contrôle de poursuite avec de bonnes vitesses presque sûres.

L'estimateur pondéré a cependant divers avantages sur l'estimateur des moindres carrés. Il établit l'optimalité – et la consistance pour la poursuite d'une trajectoire assez excitée – sous la seule hypothèse (GN), par exemple pour un bruit blanc sans moment d'ordre >2. Et surtout, en jouant sur le choix des pondérations, on obtient de meilleures vitesses presque sûres pour l'optimalité.

Optimalité

Considérons le modèle $\mathrm{ARX}_{d,m}(p,q)$ introduit en I (commentaire C). On impose l'hypothèse classique suivante :

(AR 1) *$m \leqq d$, la matrice B_1 est de rang m et si B_1^- est la matrice $m \times d$ inverse à gauche de B_1, le polynôme matriciel*

$$D(z) = B_1^- (B_1 + B_2 z + \ldots + B_q z^{q-1})$$

est causal (son déterminant n'a que des zéros de modules > 1).

Soit $\hat{R}_n$ un estimateur des moindres carrés pondéré par une suite τ_n admissible satisfaisant, p. s., $1/s_n = O(\tau_n)$. On note :

$$^*\hat{R}_n = (\hat{A}_{1,n}, \ldots, \hat{A}_{p,n}, \hat{B}_{1,n}, \ldots, \hat{B}_{q,n}), \qquad \tilde{R}_n = \hat{R}_n - R.$$

Pour une trajectoire z prévisible, le contrôle de poursuite $U(z)$ est défini si $m \leqq d$ et si $\hat{B}_{1,n}$ est, p. s., de rang m. C'est en particulier le cas, sous l'hypothèse suivante :

(AR 2) *Pour tout n, la loi de ε_{n+1} conditionnelle à $\mathcal{F}_n$ ne charge aucun sous-espace affine (ce qui est le cas si c'est un bruit blanc dont la loi a une densité).*

Sous l'hypothèse (GN), p. s. :

$$\liminf \frac{1}{n} s_n \geqq \liminf \frac{1}{n} \sum_{k=1}^{n} \|X_k\|^2 \geqq \operatorname{tr} \Gamma > 0; \text{ d'où } s_\infty = \infty.$$

Mais, d'après la dernière phrase de la section III, l'hypothèse $(1/s_n) = O(\tau_n)$ p. s. assure :

$$\sum_{k=1}^{n} \|^*\tilde{R}_k \Phi_k\|^2 = o(s_n), \quad \text{p. s.}$$

D'où, p. s., $\sum_{k=1}^{n} \|X_k - z_k\|^2 = o(s_{n-1}) + O(n)$.

L'hypothèse (AR1) implique que, pour $V_n = U_{n-1}$, R l'opérateur retard, et $A(z) = I - A_1 z - \ldots - A_p z^p$, on a :

$$D(R) V = B_1^- A(R) X - B_1^- \varepsilon.$$

On en déduit, selon un raisonnement classique (*cf.* [12], p. 89) :

$$\sum_{k=1}^{n} \|U_{k-1}\|^2 = O\left(\sum_{k=1}^{n} \|X_k\|^2 + n\right).$$

D'où : $s_{n-1} = O\left(\sum_{k=1}^{n} \|X_k\|^2 + n\right)$.

Or : $\sum_{k=1}^{n} \|X_k\|^2 \leqq 3 \sum_{k=1}^{n} (\|^*\tilde{R}_{k-1} \Phi_{k-1}\|^2 + \|\varepsilon_k\|^2 + \|z_k\|^2)$.

Supposons $\sum_{k=1}^{n} \|z_k\|^2 = O(n)$, p. s.; pour $G_n = \sum_{k=1}^{n} \|X_k\|^2$, on a, p. s.

$$G_n = o(n + G_n) + O(n), \text{ d'où } G_n = O(n) \text{ et } s_n = O(n).$$

Par suite, p. s. : $\sum_{k=1}^{n} \| {}^*\tilde{R}_k \Phi_k \|^2 = o(s_n) = o(n)$.

THÉORÈME IV.7 (Optimalité du contrôle de poursuite). — *On considère un modèle* $\mathrm{ARX}_{d,m}(p, q)$ *satisfaisant les hypothèses* (GN), (AR1) *et* (AR2). *On utilise un estimateur des moindres carrés pondéré par une suite* (τ_n) *admissible telle que,* $1/s_n = O(\tau_n)$, *p. s. On suppose la trajectoire* (z_n) *à poursuivre prévisible et telle que, p. s.,* $\sum_{k=1}^{n} \|z_k\|^2 = O(n)$.

Le contrôle de poursuite U (z) *est optimal*: $\frac{1}{n} C_n \xrightarrow{\text{p. s.}} \Gamma$.

De plus, pour ce contrôle,

$$\sum \frac{1}{n} \|X_n - z_n - \varepsilon_n\|^2 < \infty, p. s.;$$

les covariances empiriques a priori *et* a posteriori *sont des estimateurs fortement consistants de la covariance du bruit.* ■

Optimalité pour un bruit ayant un moment d'ordre >2

Sous l'hypothèse (B2) ces résultats peuvent être renforcés selon une méthode proche de celle de [17]. On suppose toujours ici que l'hypothèse (AR1) est satisfaite.

On peut d'abord supprimer l'hypothèse (AR2), à condition de modifier l'estimateur lorsque $\hat{B}_{1,n}$ est de rang $r<m$.

Soit $C_n = {}^*\hat{B}_{1,n} \hat{B}_{1,n}$ et $(\lambda_i)_{i \leq m}$ ses valeurs propres, $\lambda_{i,n} > 0$ pour $i \leq r$. On considère V_n, matrice orthogonale $m \times m$ dont les r premiers vecteurs colonnes sont $(v_{i,n})_{i \leq r}$, $v_{i,n}$ vecteur propre de C_n associé à $\lambda_{i,n}$; puis W_n matrice orthogonale $d \times d$ dont les r premiers vecteurs colonnes sont $(w_{i,n})_{i \leq r}$, $w_{i,n} = \lambda_{i,n}^{-1/2} \hat{B}_{1,n} v_{i,n}$. On pose

$$\check{B}_{1,n} = \hat{B}_{1,n} \quad \text{si} \quad \lambda_{\min}(\hat{B}_{1,n} {}^*\hat{B}_{1,n}) > 0,$$

$$\check{B}_{1,n} = \hat{B}_{1,n} + n^{-1} s_n^{-1/2} W_n {}^*V_n \quad \text{sinon.}$$

Dans le second cas, $\check{B}_{1,n} v_{i,n} = (\lambda_{i,n}^{1/2} + n^{-1} s_n^{-1/2}) w_{i,n}$; le rang de $\check{B}_{1,n}$ est m. L'estimateur pondéré où $\hat{B}_{1,n}$ est remplacé par $\check{B}_{1,n}$ sera appelé *l'estimateur pondéré normalisé* de R et noté $\check{R}_n$. On a :

$$\|\hat{R}_n - \check{R}_n\|^2 \leq n^{-2};$$

Ainsi, la normalisation ne modifie pas asymptotiquement l'estimation des paramètres.

Sous l'hypothèse (AR2), l'estimateur pondéré coïncide, p. s., avec l'estimateur pondéré normalisé.

On utilise le contrôle de poursuite $U(z)$ pour cet estimateur pondéré normalisé : on a, $z_{n+1} = {}^*\check{R}_n \Phi_n$;

$$X_{n+1} - z_{n+1} = -{}^*\tilde{R}_n \Phi_n + (\hat{B}_{1,n} - \check{B}_{1,n}) U_n + \varepsilon_{n+1} = \pi_n + \varepsilon_{n+1},$$

$$\|X_{n+1} - z_{n+1}\|^2 \leqq 3 \|{}^*\tilde{R}_n \Phi_n\|^2 + 3 [n^{-2} s_n^{-1}] \|U_n\|^2 + 3 \|\varepsilon_{n+1}\|^2.$$

On suppose satisfaites les hypothèses (B2) et (AR1);

$$\liminf [s_n/n] \geqq \operatorname{Tr} \Gamma > 0, \text{ et } s_\infty = \infty.$$

D'après le théorème I.2, $\|{}^*\tilde{R}_n \Phi_n\|^2 = o[1/\tau_n + \|\Phi_n\|^2/\alpha_n]$, où la suite (α_n) tend vers ∞, p. s., puisque $s_\infty = \infty$.

Mais, l'hypothèse (AR1) signifie aussi (*cf.* [5] ou [17]) qu'il existe deux constantes a et λ, $0 < \lambda < 1$, et une variable aléatoire A, telles que :

$$\|U_{n-1}\|^2 \leqq a \left(\sum_{k=0}^{n} \lambda^{n-k} (\|X_k\|^2 + \|\varepsilon_k\|^2) \right) + A.$$

Soit $e_n = \sup_{k \leqq n} (\|z_k\|^2 + \|\varepsilon_k\|^2)$ et $L_n = \sum_{k=0}^{n} \lambda^{n-k} \|X_k\|^2$;

$$\begin{aligned} L_{n+1} = \lambda L_n + \|X_{n+1}\|^2 &\leqq \lambda L_n + o(1/\tau_n) \\ &\quad + O(e_{n+1}) + L_{n+1} O(n^{-2} \tau_n + 1/\alpha_n) \\ &\leqq (\lambda L_n + o(1/\tau_n) + O(e_{n+1}))(1 + O(n^{-2} \tau_n + 1/\alpha_n)). \end{aligned}$$

Pour n assez grand, $\lambda [1 + O(n^{-2} \tau_n + 1/\alpha_n)] < 1$, et :

$$L_n = o(1/\tau_{n-1}) + O(e_n).$$

D'où : $\|U_{n-1}\|^2 = o(1/\tau_{n-1}) + O(e_n)$ et

$$\|\Phi_n\|^2 = o(1/\tau_n) + O(e_{n+1}), \text{ p. s.;}$$

et :

$$\|\pi_n\|^2 \leqq 2 \|{}^*\tilde{R}_n \Phi_n\|^2 + 2 [s_n^{-1} n^{-2}] \|U_n\|^2 \leqq 2 \|{}^*\tilde{R}_n \Phi_n\|^2 + 2 n^{-2} [\|\Phi_n\|^2/\alpha_n].$$

On a donc, comme dans la partie 2 du théorème I.2,

$$E(\sum a_n \|\pi_n\|^2) < \infty, \qquad \text{avec} \quad a_n = \inf(\tau_n, \alpha_n/\|\Phi_n\|^2).$$

Sous l'hypothèse (B2a), $a \geqq 1$, $\sup_{k \leqq n} \|\varepsilon_k\|^2 = o(n^{1/a})$, p. s. Si, p. s., $\sum_{k=1}^{n} \|z_k\|^2 = O(n)$, $s_n = O(n)$; s_n et n sont, p. s., du même ordre. Si, p. s., $\|z_n\|^2 = o(n^{1/a})$, on peut reprendre la partie 3 de la proposition IV.6 pour toute suite v_n croissante telle que,

$$o(1/\tau_n) + e_{n+1} = O(v_n), \qquad v_n = o(n^{1/a}).$$

THÉORÈME IV.8 (Vitesse de l'optimalité sous l'hypothèse (B2). – *On considère un modèle* $\mathrm{ARX}_{d,m}(p,q)$ *satisfaisant l'hypothèse* (AR1), *avec un bruit satisfaisant* (B 2 *a*), $a \geqq 1$. *On suppose que la trajectoire à poursuivre* $z=(z_n)$ *est prévisible et satisfait, p. s.:* $\|z_n\|^2=o(n^{1/a})$ *et* $\sum_{k=1}^{n}\|z_k\|^2=O(n)$.

La pondération est $\tau_n=f(\operatorname{Log} s_n)$ *avec*:

$$f(t)=t^{-1-\gamma},\quad \gamma>0, \qquad ou \qquad f(t)=e^{-\gamma t},\quad 0<\gamma\leqq 1/a.$$

On utilise le contrôle de poursuite associé à l'estimateur pondéré normalisé [*sans normalisation si l'on fait l'hypothèse* (*AR*2)].

(*a*) *Soit* $\sup_{k\leqq n+1}(\|z_k\|^2+\|\varepsilon_k\|^2)+1/f(\operatorname{Log} n)=v_n$.

Le contrôle de poursuite satisfait, p. s.:

$$\|U_n\|^2+\|X_n\|^2=O(v_n),\ \|U_n\|^2+\|X_n\|^2=o(n^{1/a}),$$

$$\sum\|X_n-z_n-\varepsilon_n\|^2/v_n<\infty,$$

$$\left\|\frac{1}{n}C_n-\sigma_n\right\|=o(v_n/n).$$

(*b*) *Les estimateurs empiriques de la covariance satisfont, p. s.*

$$\|\hat{\Gamma}_n-\sigma_n\|=o(V_n/n) \qquad et \qquad \|\check{\Gamma}_n-\sigma_n\|=o(v_n/n). \quad \blacksquare$$

Commentaire

La suite v_n dépend du bruit et de la pondération.

Sous la seule hypothèse (B 2 *a*), on prend $v_n=n^{1/a}$.

Soit un bruit tel que, pour tout λ, $E(\exp[\lambda\varepsilon_{n+1}]\,|\,\mathcal{F}_n)\leqq e^{-c\lambda^2}$, c constante >0, par exemple, un bruit blanc borné ou gaussien.

On déduit du théorème de Borel-Cantelli que $\sup_{k\leqq n}\|\varepsilon_k\|^2=O(\operatorname{Log} n)$.

Donc, dès que la pondération est prise égale à $\tau_n=(\operatorname{Log} s_n)^{-1-\gamma}$, $\gamma>0$, on peut prendre, pour une trajectoire à poursuivre bornée (ou telle que $\|z_n\|^2=O(\operatorname{Log} n)^{1+\gamma}$), $v_n=(\operatorname{Log} n)^{1+\gamma}$.

IV.3. Identification avec optimalité ou « stabilisation »

Les résultats qui suivent seront améliorés en IV.4 sous l'hypothèse (B 2).

On se place dans le cadre du théorème IV.7. Le contrôle utilisé est le contrôle de poursuite U(z); on a donc, p. s.:

$$s_n=O(n) \qquad \text{et} \qquad \liminf s_n/n\geqq \operatorname{tr}\Gamma>0.$$

La suite (s_n) est, p. s., de l'ordre de n.

Désignons par f l'une ou l'autre des fonctions suivantes:

$$f(x)=x^{-1-\gamma} \quad \text{pour } \gamma>0 \qquad \text{ou} \qquad f(x)=e^{-\gamma x} \quad \text{pour } 0<\gamma\leqq 1/2.$$

On utilise dorénavant l'estimateur pondéré par $\tau_n=f(\operatorname{Log} s_n)$.

Grâce à la partie 3 du théorème I.2, on a, p. s.:

$$\sum_{k=0}^{n} \rho_k \| {}^*\tilde{\mathrm{R}}_k \Phi_k \|^2 = o(n).$$

Pour identifier, on ajoute, avec les notations introduites dans (AR1), l'hypothèse suivante:

(AR3) *Irréductibilité*: $\mathrm{B}_1^- \mathrm{B}_q$ *est inversible et les polynômes matriciels* $\mathrm{D}(z)$ *et* $\mathrm{B}_1^- \mathrm{A}(z)$ *sont premiers entre eux à gauche.*

La covariance Γ *du bruit* ε *est en outre supposée inversible.*

Sous les hypothèses (AR1) et (AR3), on peut utiliser le théorème d'excitation (voir [25], [3]). Il existe une constante K et une variable aléatoire finie L telles que, pour tout n, on ait, p. s.:

$$\lambda_{\min} \sum_{k=0}^{n} \Phi_k {}^*\Phi_k \geqq \mathrm{K}\, \lambda_{\min} \left(\sum_{k=0}^{n} \mathrm{H}_{k+1} {}^*\mathrm{H}_{k+1} \right) + \mathrm{L},$$

avec ${}^t\mathrm{H}_n = ({}^t\mathrm{X}_n, \ldots, {}^t\mathrm{X}_{n-m(q-1)-p}, {}^t\varepsilon_n, \ldots, {}^t\varepsilon_{n-m(q-1)})$. ■

Considérons alors la suite des vecteurs (J_n) où J_n est le vecteur obtenu en remplaçant dans l'expression de H_n les composantes X_{n-k} par $\varepsilon_{n-k}+z_{n-k}$. On a, p. s.:

$$\sum_{k=0}^{n} \| \mathrm{J}_{k+1} - \mathrm{H}_{k+1} \|^2 = o(n).$$

Supposons la trajectoire (z_n):

- *satisfaisant, p. s.,* $\sum_{k=1}^{n} \| z_k \|^2 = O(n)$;
- *fortement excitante d'ordre* $m(q-1)+p+1$, *c'est-à-dire telle que*

$$\liminf \lambda_{\min} \frac{1}{n} \sum_{k=m(q-1)+p}^{n} \zeta_k {}^*\zeta_k > 0, \quad \text{p. s.},$$

où ${}^t\zeta_n = ({}^t z_n, \ldots, {}^t z_{n-m(q-1)-p})$;

- *prévisible d'ordre* $m(q-1)+p$, *c'est-à-dire telle que, pour tout* n, z_n *soit mesurable par rapport à* $\mathscr{F}_{n-m(q-1)-p-1}$.

Cette dernière hypothèse est par exemple satisfaite lorsque la trajectoire est indépendante du bruit ε et de l'état initial Φ_0 (la tribu $\mathscr{F}_n$ étant alors la tribu engendrée par Φ_0 et $(\varepsilon_k)_{k \leqq n}$). Alors, p. s.:

$$\liminf [1/n] \lambda_{\min} \sum_{k=0}^{n} \zeta_{k+1} {}^*\zeta_{k+1} > 0,$$

$$\liminf [1/n] \lambda_{\min} \sum_{k=0}^{n} J_{k+1} {}^*J_{k+1} > 0,$$

d'où: $\liminf [1/n] \lambda_{\min} S_n > 0$,

$$\liminf [1/n f(\operatorname{Log} n)] \lambda_{\min} T_n > 0.$$

D'après le théorème I.1, on a donc, p. s.,

$$\| \hat{R}_n - R \|^2 = O(1/nf(\operatorname{Log} n)).$$

On peut dès lors utiliser la partie 3 du théorème I.2, en remarquant que $1/f(\operatorname{Log} n) = O(\rho_n)$.

THÉORÈME IV.9 (Trajectoire excitante). — *On se place sous les hypothèses* (AR1-2-3). *L'estimateur est l'estimateur des moindres carrés pondéré par* $\tau_n = f(\operatorname{Log} s_n)$ *avec* $f(t) = t^{-1-\gamma}$, $\gamma > 0$, *ou* $f(t) = e^{-\gamma t}$, $0 < \gamma \leqq 1/2$.

Le contrôle de poursuite U(z) *d'une trajectoire* $z = (z_n)$ *fortement excitante d'ordre* $m(q-1)+p+1$, *prévisible d'ordre* $m(q-1)+p$ *et pour laquelle, p. s.,* $\sum_{k=1}^{n} \| z_k \|^2 = O(n)$, *a les propriétés suivantes*:

1. U(z) *est optimal. On a, p. s.*:

$$\sum \| X_n - z_n - \varepsilon_n \|^2 [1/nf(\operatorname{Log} n)] < \infty;$$

$$\left\| \frac{1}{n} C_n - \sigma_n \right\| = o(f(\operatorname{Log} n)).$$

2. *L'estimateur* $\hat{R}_n$ *de* R *est fortement consistant et, p. s.*:

$$\| \hat{R}_n - R \|^2 = O(1/nf(\operatorname{Log} n)).$$

3. *Les estimateurs empiriques de la covariance* Γ *du bruit sont fortement consistants et, p. s.,*

$$\| \hat{\Gamma}_n - \sigma_n \| = o(f(\operatorname{Log} n)), \qquad \| \check{\Gamma}_n - \sigma_n \| = o(f(\operatorname{Log} n)). \quad \blacksquare$$

Remarque

Les vitesses $f(\operatorname{Log} n)$ obtenues dans les parties 1 et 3 étaient v_n/n dans le cadre du théorème IV.8.

Sous l'hypothèse (B2 a), $v_n/n = n^{1/a-1}$; lorsque $1 < a < 2$, la pondération $\tau_n = s_n^{-1/2}$ donne $f(\operatorname{Log} n) = n^{-1/2}$ donc de meilleures vitesses que dans le théorème IV.8.

Mais, si, pour améliorer la vitesse de la consistance de $\hat{R}_n$, on prend $\tau_n = (\mathrm{Log}\, s_n)^{-1-\gamma}$, $\gamma > 0$, la vitesse de l'optimalité obtenue dans le théorème IV. 8 est la meilleure.

Pour un bruit gaussien ou borné, on a intérêt à prendre cette dernière pondération. En cumulant les théorèmes IV.8 et IV.9, on obtient, p. s., des vitesses $(\mathrm{Log}\, n)^{1+\gamma}/n$ tant pour $\|\hat{R}_n - R\|^2$ que pour $\left\|\frac{1}{n} C_n - \sigma_n\right\|$, $\|\hat{\Gamma}_n - \sigma_n\|$ ou $\|\check{\Gamma}_n - \sigma_n\|$.

En corollaire, on introduit un contrôle de poursuite *excité*, où la trajectoire à poursuivre est excitée par une trajectoire L, le contrôle choisi étant dès lors $U(z+L)$.

On n'obtient alors que l'optimalité pour la trajectoire de poursuite excitée; il s'agit pourtant d'une bonne propriété de « stabilité ».

COROLLAIRE IV. 10 (Stabilisation et identification). — *Sous les hypothèses* (AR1-2-3), *on considère une trajectoire $z = (z_n)$ prévisible d'ordre $m(q-1)+p$ et pour laquelle, p. s.,*

$$\sum_{k=1}^{n} \|z_k\|^2 = O(n).$$

On donne un bruit $\eta = (\eta_n)$ de dimension m, de covariance Γ_1 inversible, et satisfaisant l'hypothèse (GN);

$$\sigma_n^1 = \frac{1}{n} \sum_{k=1}^{n} \eta_k {}^*\eta_k,$$

On suppose ce bruit η indépendant de ε, de l'état initial et de la trajectoire z. On utilise le même estimateur pondéré que dans le théorème IV.9.

Pour le contrôle $U(z+\eta)$, *on a, p. s., les propriétés suivantes*

1. $\frac{1}{n} C_n \to \Gamma + \Gamma_1$; *et,*

$$\sum \|X_n - z_n - \varepsilon_n - \eta_n\|^2 [1/nf(\mathrm{Log}\, n)] < \infty;$$

$$\left\|\frac{1}{n} C_n - \sigma_n - \sigma_n^1\right\| = o(f(\mathrm{Log}\, n)).$$

2. $\|\hat{R}_n - R\|^2 = O(1/nf(\mathrm{Log}\, n))$.
3. *Les estimateurs empiriques de la covariance Γ du bruit sont fortement consistants et, p. s.,*

$$\|\hat{\Gamma}_n - \sigma_n - \sigma_n^1\| = o(f(\mathrm{Log}\, n)), \qquad \|\check{\Gamma}_n - \sigma_n - \sigma_n^1\| = o(f(\mathrm{Log}\, n)).$$

Commentaires

(*a*) On obtient dans le corollaire un procédé d'excitation identique à celui que l'on utilise avec d'autres estimateurs, permettant d'identifier en préservant la stabilité ou l'optimalité.

(*b*) On a le choix des pondérations selon que l'on veuille accélérer estimation ou optimalité. Par exemple, dans le corollaire IV.10:

– Pour $\tau_n = s_n^{-\gamma}$, $0 < \gamma \leqq 1/2$, et pour des bruits ε et η satisfaisant l'hypothèse B $[2/(1-\gamma)]$, on obtient, p. s.,

$$\left\| \frac{1}{n} C_n - \Gamma - \Gamma_1 \right\| = o(n^{-\gamma}); \qquad \|\tilde{R}_n\|^2 = O(n^{-1+\gamma});$$
$$\|\hat{\Gamma}_n - \Gamma - \Gamma_1\| = o(n^{-\gamma}), \qquad \|\check{\Gamma}_n - \hat{\Gamma}_n\| = o(n^{-\gamma}).$$

– Pour $\tau_n = (\operatorname{Log} s_n)^{-1-\gamma}$, $\gamma > 0$, la vitesse de consistance de l'estimateur est $O([\operatorname{Log} n]^{1+\gamma}/n)^{1/2}$, comme dans le cas stable. Lorsque les bruits ε et η satisfont (B2), les autres vitesses sont plus faibles, $o([\operatorname{Log} n]^{-1-\gamma})$.

Les mêmes remarques s'appliquent au théorème IV.9. ■

IV.4. Optimalité et identification

On se place ici dans le cadre du théorème IV.8, avec les mêmes notations. Le contrôle de poursuite est bâti à partir de l'estimateur pondéré normalisé [sans normalisation sous l'hypothèse (AR2)].

Soit $\lambda = (\lambda_n)$ une suite déterministe, positive, croissant vers ∞. On suppose (λ_n) et (λ_{n+1}) du même ordre, et (λ_n/n) décroissante.

La trajectoire (z_n) sera dite λ-excitante d'ordre $m(q-1)+p+1$ lorsque, pour ${}^t\zeta_n = ({}^t z_n, \ldots, {}^t z_{n-m(q-1)-p})$,

$$\liminf (1/\lambda_n) \lambda_{\min} \sum_{k=m(q-1)+p}^{n} \zeta_k {}^*\zeta_k > 0, \quad \text{p. s.}$$

Alors, si $v_n = O(\lambda_n)$, on peut reprendre la démonstration de IV.3 en utilisant le fait que

$$\liminf (1/\lambda_n) \lambda_{\min} \sum_{k=0}^{n} J_{k+1} {}^*J_{k+1} > 0, \quad \text{p. s.},$$

et

$$\sum_{k=0}^{n} \|J_k - H_k\|^2 = o(v_n).$$

On obtient: $\liminf (1/\lambda_n) \lambda_{\min} S_n > 0$, p. s.

D'après le théorème I.1, on obtient la partie 1 du théorème suivant.

Dans la partie 2, on utilise un contrôle de poursuite λ-*excité*, $U(z+L)$, où L sera défini comme suit:

Pour (λ_n) suite déterministe croissant vers ∞, avec $\lambda_n - \lambda_{n-1} \leqq 1$, et $\eta = (\eta_n)$ un bruit de covariance Γ_1, ayant un moment d'ordre > 2,

$L_n = (\lambda_n - \lambda_{n-1})^{1/2} \eta_n$. On a alors, par le théorème de Chow:

$$\sum_n \lambda_n^{-1} (L_n^* L_n - (\lambda_n - \lambda_{n-1}) \Gamma_1) < \infty, \quad \text{p. s.};$$

d'où, p. s.: $\lambda_n^{-1} \sum_{k=1}^{n} L_k^* L_k \to \Gamma_1$.

D'où, par la proposition IV.6, le théorème suivant:

THÉORÈME IV.11 (Optimalité et identification sous B2). — *On se place sous les hypothèses* (AR1) *et* (AR3), *avec un bruit satisfaisant l'hypothèse* (B2 a), $a > 1$.

La pondération est $\tau_n = f(\operatorname{Log} s_n)$,

$$f(t) = t^{-1-\gamma}, \quad \gamma > 0, \qquad ou \qquad f(t) = e^{-\gamma t}, \quad 0 < \gamma \leqq 1/a.$$

L'estimateur est l'estimateur pondéré $\hat{R}_n$ *et le contrôle est le contrôle de poursuite bâti à partir de l'estimateur pondéré normalisé* [*sans normalisation sous l'hypothèse* (AR2)].

Soit $v_n = 1/f(\operatorname{Log} n) + e_{n+1}$, $e_n = \sup_{k \leqq n} (\| \varepsilon_k \|^2 + \| z_k \|^2)$.

Soit $\lambda = (\lambda_n)$ *une suite déterministe, croissante. On suppose* (λ_n) *et* (λ_{n+1}) *du même ordre*, (λ_n / n) *décroissante, et*:

$$\lambda_n - \lambda_{n-1} \leqq 1, \qquad v_n = O(\lambda_n), \qquad [f(\operatorname{Log} n)]^{-2} = O(\lambda_n).$$

On considère une trajectoire z, *prévisible d'ordre* $m(q-1)+p$, *telle que*, *p. s.*, $\sum_{k=1}^{n} \| z_k \|^2 = O(n)$ *et* $\| z_n \|^2 = o(n^{1/a})$.

1. *Si* z *est* λ*-excitante d'ordre* $m(q-1)+p+1$, *alors pour le contrôle de poursuite* U(z), *on a*, *p. s.*:

$$\| \hat{R}_n - R \|^2 = O(1/\lambda_n f(\operatorname{Log} n)),$$

$$\sum \| X_n - z_n - \varepsilon_n \|^2 f(\operatorname{Log} n) < \infty,$$

$$\left\| \frac{1}{n} C_n - \sigma_n \right\| = o(1/n f(\operatorname{Log} n)),$$

$$\| \hat{\Gamma}_n - \sigma_n \| = o(1/n f(\operatorname{Log} n)), \ \| \check{\Gamma}_n - \sigma_n \| = o(1/n f(\operatorname{Log} n)).$$

2. *Si* z *ne satisfait pas les propriétés données en* 1, *on peut utiliser le contrôle de poursuite excité* U$(z+L)$ *avec* $L_n = (\lambda_n - \lambda_{n-1})^{1/2} \eta_n$, (η_n) *bruit de covariance inversible* Γ_1, *ayant un moment d'ordre* > 2 *et indépendant de* (ε_n), *de l'état initial et de la trajectoire* z, *on a*, *p. s.*:

$$\| \hat{R}_n - R \|^2 = O(1/\lambda_n f(\operatorname{Log} n));$$

$$\sum_n \| X_n - z_n - L_n - \varepsilon_n \|^2 f(\operatorname{Log} n) < \infty;$$

$$\lambda_n^{-1} \sum_{k=1}^{n} (X_k - z_k - \varepsilon_k)^* (X_k - z_k - \varepsilon_k) \to \Gamma_1;$$

$$pour\ \sigma_n^1 = \frac{1}{n}\sum_{k=1}^{n} L_k{}^*L_k,\ p.s.,\ \left\|\frac{1}{n}C_n - \sigma_n - \sigma_n^1\right\| = o(1/nf(\text{Log}\,n));$$

$$\left\|\frac{1}{n}C_n - \sigma_n\right\| = O(\lambda_n/n),$$

$$\|\hat{\Gamma}_n - \sigma_n\| = O(\lambda_n/n),\ \|\check{\Gamma}_n - \sigma_n\| = O(\lambda_n/n).$$

L'optimalité et la forte consistante des covariances empiriques sont ainsi préservées si, p.s., $\lambda_n = o(n)$. ■

Commentaires

(*a*) Pour $\lambda_n = n$, on améliore ainsi les erreurs relatives au contrôle et aux covariances empiriques obtenues dans le théorème IV.9 et le corollaire IV.10.

(*b*) Sous la seule hypothèse (B2 *a*), $v_n = n^{1/a}$ et l'on peut prendre $\lambda_n = n^b$, $1/a \leqq b < 1$.

Pour la pondération $\tau_n = (\text{Log}\,s_n)^{-1-\gamma}$, p.s. :

$$\|\hat{R}_n - R\|^2 = O([\text{Log}\,n]^{1+\gamma} n^{-b}),$$

$$\sum \|X_n - z_n - L_n - \varepsilon_n\|^2 (\text{Log}\,n)^{-1-\gamma} < \infty,$$

où $L_n = 0$ dans la partie 1. Dans la partie 2,

$$n^{-b}\sum_{k=1}^{n} (X_k - z_k - \varepsilon_k)^* (X_k - z_k - \varepsilon_k) \to \Gamma_1.$$

(*c*) Si le bruit est borné, ou si c'est un bruit blanc gaussien, avec la même pondération qu'en (*a*), $v_n = (\text{Log}\,n)^{1+\gamma}$.

Pour $\lambda_n = (\text{Log}\,n)^b$ avec $b \geqq 2 + 2\gamma$,

$$\|\hat{R}_n - R\|^2 = O([\text{Log}\,n]^{1+\gamma-b}),$$

$$\left\|\frac{1}{n}C_n - \sigma_n\right\| = O([\text{Log}\,n]^b/n);$$

les autres résultats sont les mêmes qu'en (*b*).

On voit bien ainsi comment choisir (λ_n), selon que l'on privilégie l'optimalité ou l'identification. ■

Les auteurs remercient le rapporteur de leur avoir signalé l'existence de [17]; ses commentaires et ceux de P. Priouret ont permis d'améliorer cet article.

RÉFÉRENCES

[1] K. J. Aström et B. Wittenmark, On Self Tuning Regulators, *Automatica*, vol. **9**, 1973, p. 185-199.

[2] A. H. BECKER, P. R. KUMAR et C. Z. WEI, Adaptive Control with Stochastic Approximation Algorithm; Geometry of Convergence, *I.E.E.E. trans. automatic control.*, vol. **AC-30**, 1985, p. 330-338.

[3] B. BERCU, Weighted Estimation and Tracking for ARMAX Models, *Prépublication Université Paris-Sud*, 1991.

[4] P. CAINES, Stochastic Adaptive Control: Randomly Varying Parameters and Continually Disturbed Controls, *Control science and technology for the progress of society*, H. AKASHI éd., Pergamon, 1981, p. 925-930.

[5] P. CAINES, Linear Stochastic Systems, *Wiley*, 1988.

[6] P. E. CAINES, S. LAFORTUNE, Adaptive Control with Recursive Identification for Stochastic Linear Systems, *I.E.E.E. Trans. Automatic Control*, vol. **AC-29**, 1984, p. 312-321.

[7] H. F. CHEN, Recursive System Identification and Adaptive Control by Use of the Modified Least Squares Algorithm, *Siam J. Control Opt.*, vol. **22**, 1984, p. 758-776.

[8] H. F. CHEN et P. E. CAINES, The Strong Consistency of the Stochastic Gradient Algorithm of Adaptive Control, *I.E.E.E. Trans Automatic Control.*, vol. **AC-30**, 1985, p. 189-192.

[9] H. F. CHEN et L. GUO, Convergence Rate for Least Squares Identification and Adaptive Control for Stochastic Systems, *Int. J. Control.*, vol. **44**, 1986, p. 1459-1476.

[10] H. F. CHEN et L. GUO, Asymptotically Optimal Adaptive Control with Consistent Parameter Estimates, *S.I.A.M. J. Control and Opt.*, vol. **25**, 1987, p. 558-575.

[11] H. F. CHEN et J. F. ZHANG, Convergence Rates in Stochastic Adaptive Tracking, *Int. J. Control*, vol. **49**, 1989, p. 1915-1935.

[12] M. DUFLO, Méthodes récursives aléatoires, *Masson*, 1990.

[13] M. DUFLO, R. SENOUSSI et A. TOUATI, Sur la loi des grands nombres pour les martngales vectorielles et l'estimateur des moindres carrés d'un modèle de régression, *Ann. Inst. Henri Poincaré, Ser. B.*, vol. **26**, 1990, p. 549-566.

[14] M. DUFLO, R. SENOUSSI et A. TOUATI, Propriétés asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autorégressif vectoriel, *Ann. Inst. Henri Poincaré, Ser. B*, vol. **27**, 1991, p. 1-25.

[15] G. C. GOODWIN, P. J. RAMDAGE et P. E. CAINES, Discrete Time Stochastic Adaptive Control, *SIAM J. Control Opt.*, vol. **19**, 1981, p. 829-853.

[16] G. C. GOODWIN et K. S. SIN, Adaptive Filtering, Prediction and Control, *Prentice Hall*, 1984.

[17] L. GUO et H. F. CHEN, The Aström-Wittenmark Self Tuning Regulator Revisited and ELS-Based Adaptive Trackers, *I.E.E.E. Trans. Automatic Control.*, vol. **36**, 1991, p. 802-812.

[18] H. KAUFMANN, On the Strong Law of Large Numbers for Multivariate Martingales, *Stochastic Processes Appl.*, vol. **26**, 1987, p. 73-85.

[19] P. R. KUMAR, Convergence of Adaptive Control Shemes Using Least Squares Parameter Estimates, *I.E.E.E. Trans. Automatic Control.*, vol. **35**, 1990, p. 416-424.

[20] P. R. KUMAR et L. PRALY, Self Tuning Trackers, *S.I.A.M. J. control Opt.*, vol. **25**, 1987, p. 1053-1071.

[21] P. R. KUMAR et P. VARAIYA, Stochastic Systems: Estimation, Identification and Adaptive Control, *Prentice Hall*, 1986.

[22] T. L. LAI, Asymptotically Efficient Adaptive Control in Stochastic Regression Models, *Adv. Appl. Math.*, vol. **7**, 1986, p. 23-45.

[23] T. L. LAI et C. Z. WEI, Least Squares Estimates in Stochastic Regression Models with Applications to Identification and Control of Dynamic Systems, *Ann. Math. Stat.*, vol. **10**, 1982, p. 154-166.

[24] T. L. LAI et C. Z. WEI, Asymptotic Properties of General Autoregressive Models and Strong Consistency of Least Squares Estimates and their Parameters, *J. Multivariate Analysis*, vol. **13**, 1983, p. 1-23.

[25] 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.

[26] 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**, 1986, p. 898-906.

[27] T. L. LAI et C. Z. WEI, Asymptotically Efficient Self Tuning Regulators, *SIAM J. Control Optimization*, vol. **25**, 1987, p. 466-481.

[28] A. LE BRETON et M. MUSIELA, Consistency Sets of Least Squares Estimates in Stochastic Regression Models, dans « Stochastic Differential Systems », *Lect. Notes Control Information Sci.*, n° 126, N. CHRISTOFEIT, K. HELMES, M. KOHLMAN éd., 1989.

[29] K. S. SIN et G. C. GOODWIN, Stochastic Adaptive Control Using a Modified Least Squares Algorithm, *Automatica, J. I.F.A.C.*, vol. **18**, 1982, p. 315-321.

[30] J. S. WEI, Adaptative Prediction of Least Squares Predictors in Stochastic Regression Models with Application to Time Series, *Ann. Stat.*, vol. **15**, 1987, p. 1667-1682.

(*Manuscrit reçu le* 19 *juin* 1991;
corrigé le 4 *novembre* 1991.)