

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

M. DUFLO

R. SENOUSI

A. TOUATI

**Sur la loi des grands nombres pour les martingales
vectorielles et l'estimateur des moindres carrés
d'un modèle de régression**

Annales de l'I. H. P., section B, tome 26, n° 4 (1990), p. 549-566

http://www.numdam.org/item?id=AIHPB_1990__26_4_549_0

© Gauthier-Villars, 1990, 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/>

Sur la loi des grands nombres pour les martingales vectorielles et l'estimateur des moindres carrés d'un modèle de régression

par

M. DUFLO, R. SENOUSI

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

et

A. TOUATI

École Normale Supérieure, Bizerte, Tunisie

RÉSUMÉ. — Pour une martingale vectorielle complexe discrète et de carré intégrable, on donne des conditions de validité de la loi des grands nombres et l'on précise les vitesses de convergence. Ces résultats sont appliqués à l'estimateur des moindres carrés et au prédicteur d'un modèle de régression.

ABSTRACT. — For a discrete complex vector valued and square integrable martingale, we give conditions for the validity of the law large numbers and specify the almost sure rates of convergence. Those criteriums are applied to the least squares estimator and to the predictor of a regression model.

Classification A.M.S. : 62 M, 60 F.

I. LOI DES GRANDS NOMBRES POUR LES MARTINGALES

I.1. Loi des grands nombres

On se place sur \mathbb{C}^d muni du produit scalaire hermitien $\langle u, v \rangle = {}^*u v$ et de la norme associée $\| \cdot \|$; selon l'usage, u désigne indifféremment le vecteur ou sa matrice colonne et *u est sa matrice adjointe. Si A est une matrice $d \times d$, on note

$$\| A \| = \sup [\| A u \| ; \| u \| = 1].$$

Soit (M_n) une martingale centrée de carré intégrable à valeurs dans \mathbb{C}^d adaptée à une filtration $\mathbb{F} = (\mathcal{F}_n)$ et $(\langle M \rangle_n)$ sa variation quadratique prévisible :

$$\langle M \rangle_n = \sum_{k=1}^n \Delta_k \quad \text{où} \quad \Delta_k = E(M_k {}^*M_k - M_{k-1} {}^*M_{k-1} | \mathcal{F}_{k-1}).$$

On note respectivement $\lambda_{\min} \langle M \rangle_n$ et $\lambda_{\max} \langle M \rangle_n$ la plus petite et la plus grande valeur propre de la matrice hermitienne positive $\langle M \rangle_n$.

Pour une martingale réelle et $d=1$, la « loi des grands nombres » date de plus de 20 ans : sur $\{ \langle M \rangle_n \rightarrow \infty \}$, $\langle M \rangle_n^{-1} M_n$ tend, p. s., vers 0 et les vitesses de convergence presque sûres sont connues ([25], [2]-[3], [22]). Le cas vectoriel a donné matière à de nombreux travaux récents, le plus souvent dans le cadre des modèles de régression ([1], [12] à [19] et [30]-[31], [20]). On sait que l'ensemble des trajectoires telles que $\langle M \rangle_n^{-1} M_n$ tend vers 0 est, en général, p. s. contenu dans $\{ \lambda_{\min} \langle M \rangle_n \rightarrow \infty \}$ et un contre-exemple dû à Lai et Robbins ([13], [14], [15]) prouve que la loi des grands nombres peut être mise en défaut lorsque :

$$\lambda_{\min} \langle M \rangle_n = O(\text{Log } n), \quad \lambda_{\max} \langle M \rangle_n = O(n).$$

Nos résultats relatifs aux martingales vectorielles seront formulés dans les deux théorèmes suivants. Il est plus aisé de les formuler en remplaçant $\langle M \rangle_n$ par $Q_n = \langle M \rangle_n + Q$, Q matrice hermitienne définie positive, afin que Q_n^{-1} ait toujours un sens; ils sont aussi valables pour $Q_n = \langle M \rangle_n$ sur $\{ \lim \lambda_{\min} \langle M \rangle_n > 0 \}$.

On note $\xrightarrow{\text{p. s.}}$ pour la convergence presque sûre et $\xrightarrow{\text{P}}$ pour la convergence en probabilité.

THÉORÈME 1. — LOIS DES GRANDS NOMBRES. — Soit

$$\Gamma_\infty = \{ \lambda_{\min} \langle M \rangle_n \rightarrow \infty \}, \quad \Delta_n = \langle M \rangle_n - \langle M \rangle_{n-1}, \quad Q_n = \langle M \rangle_n + Q,$$

Q étant une matrice hermitienne définie positive.

(1) On a la loi forte des grands nombres $Q_n^{-1} M_n \rightarrow 0$ presque sûrement sur

$$E = \left\{ \sum_{k=1}^{\infty} (\lambda_{\min} Q_{k-1})^{-1} (\text{tr } Q_k^{-1/2} \Delta_k Q_k^{-1/2}) < \infty \right\} \cap \Gamma_\infty.$$

Pour qu'une trajectoire de Γ_∞ soit dans E , il suffit que l'une ou l'autre des conditions suivantes soit réalisée :

$$\sum_{k=0}^{\infty} (\lambda_{\min} Q_k)^{-1} < \infty,$$

$$\sum_{k=1}^{\infty} (\lambda_{\min} Q_{k-1})^{-1} (\text{Log Det } Q_k - \text{Log Det } Q_{k-1}) < \infty,$$

ou

$$\sum_{k=1}^{\infty} (\lambda_{\min} Q_{k-1})^{-2} \|\Delta_k\| < \infty.$$

(2) On a : $E(\|Q_n^{-1/2} M_n\|^2) \leq d \inf(n, E(\text{Log } \lambda_{\max} Q_n)) = K_n$;

$$E\left(\sum_{k=0}^{n-1} {}^*M_k [Q_k^{-1} - Q_{k+1}^{-1}] M_k\right) \leq K_n.$$

Si $(\lambda_{\min} Q_n / K_n)$ tend vers ∞ en probabilité, alors on a la loi faible des grands nombres, $Q_n^{-1} M_n \xrightarrow{P} 0$.

THÉORÈME 2. — VITESSE. — Soit a une fonction réelle qui croît vers ∞ si la variable croît vers ∞ et telle que $\int_0^\infty [a(t)]^{-1} dt < \infty$ [par exemple, $a(t) = t^{1+\gamma}$ ou $a(t) = t(\text{Log } t)^{1+\gamma}$ pour $\gamma > 0$]. Avec les notations du théorème 1, on a, p. s. :

$$(1) \quad [a(n)]^{-1/2} Q_n^{-1/2} M_n \rightarrow 0;$$

et, sur $\{ \lambda_{\max} Q_n \rightarrow \infty \}$, $(a(\text{Log } \lambda_{\max} Q_n))^{-1/2} Q_n^{-1/2} M_n \rightarrow 0$.

$$(2) \quad \sum_{n=0}^{\infty} [a(n+1)]^{-1} {}^*M_n (Q_n^{-1} - Q_{n+1}^{-1}) M_n < \infty.$$

$$\sum_{n=0}^{\infty} [a(\text{Log } \lambda_{\max} Q_{n+1})]^{-1} {}^*M_n (Q_n^{-1} - Q_{n+1}^{-1}) M_n < \infty.$$

Commentaires. – 1. Nous avons formulé nos résultats dans \mathbb{C}^d parce que, dans l'utilisation qui en est faite pour les modèles linéaires même réels, il est parfois plus agréable de pouvoir faire des transformations linéaires complexes : l'analyse de modèles linéaires complexes n'est donc pas superflue (voir [7]).

2. Dans le cadre des modèles de régression réels la partie (1) du théorème 1 est donnée dans [20], tandis que la partie (1) du théorème 2 est donné dans [15] et [19]. Nous donnerons une démonstration succincte de ces résultats car leur démonstration est analogue à celles des modèles de régression. On trouve dans [10] un résultat contenu dans la partie (1) du théorème 2, prouvé dans le cadre des martingales. Il reste à généraliser nos résultats au temps continu afin d'améliorer [4], [21] ou [24].

I.2. Démonstration des théorèmes 1 et 2

(a) Posons

$$V_n = \|Q_n^{-1/2} M_n\|^2 = {}^*M_n Q_n^{-1} M_n$$

et

$$\Delta_n = Q_n - Q_{n-1} = E(M_n {}^*M_n - M_{n-1} {}^*M_{n-1} | \mathcal{F}_{n-1}).$$

On a

$$E(V_{n+1} | \mathcal{F}_n) = V_n + \chi_n - \eta_n,$$

avec

$$\begin{aligned} \chi_n &= E({}^*M_{n+1} Q_{n+1}^{-1} M_{n+1} | \mathcal{F}_n) - {}^*M_n Q_{n+1}^{-1} M_n \\ &= \text{tr} Q_{n+1}^{-1} E(M_{n+1} {}^*M_{n+1} - M_n {}^*M_n | \mathcal{F}_n) \\ &= \text{tr} Q_{n+1}^{-1/2} \Delta_{n+1} Q_{n+1}^{-1/2}, \\ \eta_n &= {}^*M_n (Q_n^{-1} - Q_{n+1}^{-1}) M_n. \end{aligned}$$

Les suites (χ_n) et (η_n) sont positives et, d'après Solo [27],

$$\begin{aligned} \text{tr} &= (Q_{n+1}^{-1/2} \Delta_{n+1} Q_{n+1}^{-1/2}) \\ &\leq \inf \{ d, \text{Log Det } Q_{n+1} - \text{Log Det } Q_n, d(\lambda_{\min} Q_{n+1})^{-1} \|\Delta_{n+1}\| \}; \end{aligned}$$

donc

$$E\left(V_n + \sum_{k=0}^{n-1} \eta_k\right) \leq d E[\inf(n, \text{Log } \lambda_{\max} Q_n)].$$

$$\text{Or } V_n = \|Q_n^{-1/2} M_n\|^2 \text{ et } \|Q_n^{-1} M_n\|^2 \leq [\lambda_{\min} Q_n]^{-1} V_n.$$

Soit (ρ_n) une suite strictement positive telle que $[V_n/\rho_n]$ soit tendue et $\lambda_{\min} Q_n/\rho_n \xrightarrow{P} \infty$; alors $\|Q_n^{-1} M_n\|$ tend vers 0 en probabilité.

Comme la suite $V_n/E(V_n)$ est tendue, on obtient la partie (2) du théorème 1.

(b) Considérons une suite adaptée (a_n) croissant vers ∞ . Un corollaire classique du théorème de Robbins-Siegmund [26] établit que sur $\Gamma^1 = \{ \sum a_n^{-1} \chi_n < \infty \}$, p. s.,

$$\begin{aligned} & \sum a_n^{-1} (V_{n+1} - V_n) \text{ converge} \quad \text{et} \quad \sum a_n^{-1} \eta_n < \infty; \\ & \text{sur } \{ a_n \rightarrow a_\infty < \infty \} \cap \Gamma^1, \quad (V_n) \text{ converge}; \\ & \text{sur } \{ a_n \rightarrow \infty \} \cap \Gamma^1, \quad a_n^{-1} V_n \xrightarrow{p. s.} 0 \quad \text{et} \quad a_n^{-1} V_{n+1} \xrightarrow{p. s.} 0. \end{aligned}$$

Sur E, en prenant $a_n = \lambda_{\min} Q_n$, on obtient p. s. :

$$(\lambda_{\min} Q_n)^{-1} \|Q_n^{-1/2} M_n\|^2 \rightarrow 0, \quad \text{d'où} \quad Q_n^{-1} M_n \rightarrow 0;$$

en outre

$$\sum (\lambda_{\min} Q_n)^{-1} \eta_n < \infty.$$

Le théorème 1 est établi.

Soit a une fonction réelle croissante vers ∞ , positive, telle que

$$\int^\infty [a(t)]^{-1} dt < \infty, \quad \text{ce qui équivaut à} \quad \int^\infty [a(t/d)]^{-1} dt < \infty.$$

Comme

$$[a(\text{Log Det } Q_{n+1}/d)]^{-1} \chi_n \leq \int_{\text{Log Det } Q_n}^{\text{Log Det } Q_{n+1}} [a(t/d)]^{-1} dt,$$

en prenant $a_n = a(n)$, ou $a_n = a(\text{Log } \lambda_{\max} Q_n)$, on a :

$$\sum a_{n+1}^{-1} \chi_n < \infty.$$

Si (a_n) tend vers ∞ , $a_n^{-1/2} Q_n^{-1/2} M_n \xrightarrow{p. s.} 0$.

En outre, p. s., $\sum a_{n+1}^{-1} \eta_n < \infty$. Le théorème 2 est établi. ■

I.3. Un contre-exemple

Diverses conjectures restent ouvertes.

Il est sans doute possible de prouver que la suite $\langle M \rangle_n^{-1} M_n$ a toujours une limite presque sûre, ce qui permettrait de substituer à la loi faible du théorème 1-(2) une loi forte. Kumar [11] a ainsi montré par une méthode

bayésienne un résultat de ce type : pour un modèle de régression, si le bruit est blanc et gaussien et si $d=1$, l'estimateur des moindres carrés (vecteur de dimension q) a presque sûrement une limite pour toute valeur du paramètre située hors d'un ensemble de \mathbb{R}^q de mesure de Lebesgue nulle. On sait que la condition $\lambda_{\min} \langle M \rangle_n \rightarrow \infty$ n'est pas suffisante; un contre-exemple est donné par Lai et Robbins ([13]-[14]).

Voici un autre contre-exemple, qui nous semble éclairant; il prouve que l'estimateur des moindres carrés d'un modèle autorégressif explosif d'ordre 1 de dimension 2 ayant un sous-espace propre de dimension >1 peut ne pas être consistant; on montre dans [7] qu'il s'agit là du cas type où la consistance de l'estimateur des moindres carrés d'un modèle autorégressif peut être mise en défaut.

Soit (ε_n) une suite de vecteurs aléatoires indépendants et de même loi réelle de dimension 2, centrée et de covariance I; (X_n) le « modèle autorégressif » de dimension 2 :

$$X_{n+1} = 2X_n + \varepsilon_{n+1}, \quad X_0 \text{ étant réel indépendant de } (\varepsilon_n).$$

Soit $M_n = \sum_{k=1}^n X_{k-1} {}^t \varepsilon_k$; les vecteurs colonnes de M_n sont des martingales

vectérielles de crochet $S_{n-1} = \sum_{k=0}^{n-1} X_k {}^t X_k$. Montrons que, pour $Q_n = S_n + Q$, Q symétrique et définie positive, $Q_{n-1}^{-1} M_n$ a, p. s., une limite non nulle.

Posons $\eta = ({}^t \eta^1, \eta^2) = \sum_{k=1}^{\infty} 2^{-k} \varepsilon_k + X_0$. Il résulte d'un théorème de Paul

Lévy [23] que η ne charge aucun hyperplan. On a :

$$X_n = (X_n^1, X_n^2) = 2^n \eta - R_n \quad \text{avec} \quad R_n = \sum_{k=1}^{\infty} 2^{-k} \varepsilon_{k+n}.$$

Calculs préliminaires

On a

$$\alpha_n = \sup_{k \geq n} k^{-1/2} \|\varepsilon_k\| \xrightarrow{\text{p. s.}} 0,$$

et

$$\|R_n\| n^{-1/2} \leq \text{Cte } \alpha_n \xrightarrow{\text{p. s.}} 0, \quad 2^{-n} X_n \xrightarrow{\text{p. s.}} \eta.$$

En outre $2 R_n = \varepsilon_{n+1} + R_{n+1}$ et, pour tout $m \geq 1$, on a :

$$\begin{aligned} \|R_n\|^2 &\leq \left(2^{-m} \|R_{n+m}\| + \sum_{k=1}^m 2^{-k} \|\varepsilon_{k+n}\| \right)^2 \\ &\leq 2 \left(2^{-2m} \|R_{n+m}\|^2 + \left[\sum_{j=1}^{\infty} 2^{-j} \right] \left[\sum_{k=1}^m 2^{-k} \|\varepsilon_{k+n}\|^2 \right] \right). \end{aligned}$$

Pour

$$N \geq 1, \quad \sum_{n=1}^N \|R_n\|^2 \leq 2^{-2m+1} \sum_{n=1}^N \|R_{n+m}\|^2 + 2 \sum_{n=1}^{N+m} \|\varepsilon_n\|^2;$$

prenant $m = N$,

$$N^{-1} \sum_{n=1}^N \|R_n\|^2 \leq 2^{-2N+1} N^{-1} \sum_{n=1}^N \|R_{n+N}\|^2 + 2 N^{-1} \sum_{n=1}^{2N} \|\varepsilon_n\|^2,$$

le premier terme du membre de droite, ρ_N , tend vers 0, p. s., car

$$P[\rho_N \geq 2^{-N}] \leq \text{cte } 2^{-N};$$

donc, p. s., $\limsup N^{-1} \sum_{n=1}^N \|R_n\|^2 < \infty$. Soit $T_N = \sum_{n=1}^N R_n {}^t R_n$ et $t_N = T_N/N$;

(t_N) est une suite matricielle bornée, p. s., et :

$$4 T_{N-1} = T_N - T_1 + \sum_{n=2}^N \varepsilon_n {}^t \varepsilon_n + W_N + {}^t W_N \quad \text{avec} \quad W_N = \sum_{n=2}^N R_n {}^t \varepsilon_n.$$

La suite $(R_n {}^t \varepsilon_n)$ est une suite de matrices aléatoires centrées de même loi et de carré intégrable. Pour tous u et v de \mathbb{R}^2 et $j \neq k$, la variable aléatoire $\langle u, R_j \rangle \langle v, \varepsilon_j \rangle \langle u, R_k \rangle \langle v, \varepsilon_k \rangle$ est centrée :

$$E([{}^t u W_N v]^2) = (N-1) E([\langle u, R_n \rangle \langle v, \varepsilon_n \rangle]^2) = O(N).$$

Cela implique par une loi des grands nombres sur les suites de carré intégrable, (cf. [6], lemme 1.4.34) que $W_N/N \xrightarrow{p. s.} 0$.

D'où $N^{-1} [4 T_{N-1} - T_N] \rightarrow I$. En outre,

$$\|t_N - t_{N-1}\| = \left\| \frac{1}{N} R_N {}^t R_N - \frac{1}{N(N-1)} \sum_{n=1}^{N-1} R_n {}^t R_n \right\| \leq \frac{1}{N} \|R_N\|^2 + \frac{\|t_{N-1}\|}{N-1};$$

d'où $\|t_N - t_{N-1}\| \xrightarrow{p. s.} 0$ et, p. s., la seule valeur d'adhérence possible de la suite bornée (t_N) est $I/3$.

Étude $Q_{n-1}^{-1} M_n$

On a :

$$\langle X_n, \eta \rangle = 2^n \|\eta\|^2 - \langle R_n, \eta \rangle,$$

$$2^{-2N} {}^t \eta S_N \eta = 2^{-2N} \sum_{n=0}^N (2^n \|\eta\|^2 - \langle R_n, \eta \rangle)^2 \rightarrow \frac{4}{3} \|\eta\|^4.$$

Posons $\zeta = (\eta^2, -\eta^1)$. On a $\langle X_n, \zeta \rangle = -\langle R_n, \zeta \rangle$;

$$N^{-1} {}^t \zeta S_N \zeta \xrightarrow{p.s.} \frac{1}{3} \|\zeta\|^2.$$

Or $2^{-N} N^{-1/2} {}^t \zeta S_N \eta =$

$$2^{-N} N^{-1/2} \sum_{n=0}^N (-2^n \|\eta\|^2 \langle R_n, \zeta \rangle + \langle R_n, \eta \rangle \langle R_n, \zeta \rangle);$$

$$2^{-N} N^{-1/2} \sum_{n=1}^N 2^n n^{1/2} [n^{-1/2} |\langle R_n, \zeta \rangle|] \xrightarrow{p.s.} 0;$$

$$2^{-N} N^{-1/2} {}^t \zeta S_N \eta \xrightarrow{p.s.} 0.$$

Posons $P = [\zeta, \eta]$, matrice dont les vecteurs colonnes sont ζ et η et $\Delta_n = \text{Diag}[n^{1/2}, 2^n]$. On vient d'obtenir :

$$G_n = \Delta_n^{-1} {}^t P Q_n P \Delta_n^{-1} \xrightarrow{p.s.} G = \begin{bmatrix} \|\zeta\|^2/3 & 0 \\ 0 & 4\|\eta\|^4/3 \end{bmatrix}.$$

On écrit

$$Q_{n-1}^{-1} M_n = Q_{n-1}^{-1} ({}^t P)^{-1} \Delta_{n-1} \Delta_{n-1}^{-1} {}^t P M_n;$$

pour $H_n = \Delta_{n-1}^{-1} {}^t P M_n$,

$$Q_{n-1}^{-1} M_n = P \Delta_{n-1}^{-1} G_{n-1}^{-1} H_n;$$

et

$$\lim P^{-1} Q_{n-1}^{-1} M_n = \lim \begin{bmatrix} 1 & 0 \\ 0 & n^{1/2} 2^{-n+1} \end{bmatrix} G_{n-1}^{-1} (n^{-1/2}) H_n.$$

Étudions $H_n = \begin{bmatrix} H_n^1 \\ H_n^2 \end{bmatrix}$;

$$\begin{aligned} (N-1)^{-1/2} H_N^1 &= (N-1)^{-1} \sum_{n=1}^N \langle X_{n-1}, \zeta \rangle {}^t \varepsilon_n \\ &= -(N-1)^{-1} \sum_{n=1}^N \langle R_{n-1}, \zeta \rangle {}^t \varepsilon_n \\ &= -(N-1)^{-1} \sum_{n=1}^N (1/2) \langle \varepsilon_n, \zeta \rangle {}^t \varepsilon_n - (N-1)^{-1} 2^{-1} \sum_{n=1}^N \langle R_n, \zeta \rangle {}^t \varepsilon_n. \end{aligned}$$

Comme on a vu que $W_N/N \xrightarrow{p.s.} 0$, $\frac{1}{N} \sum_{n=1}^N \langle R_n, \zeta \rangle {}^t \varepsilon_n \xrightarrow{p.s.} 0$;

$$N^{-1/2} H_N^1 \rightarrow -(1/2) {}^t \zeta = -(1/2) (\eta^2, -\eta^1);$$

$$\begin{aligned} N^{-1} \|H_N^2\| &\leq 2N^{-1} \sum_{n=1}^N 2^{-(N-n)} |\langle 2^{-n} X_{n-1}, \eta \rangle| \|\varepsilon_n\| \\ &\leq o(N^{-1/2}) \|\eta\| \text{Cte sup}_n \|2^{-n} X_{n-1}\| = o(N^{-1/2}); \end{aligned}$$

$$P^{-1} Q_{n-1}^{-1} M_n \xrightarrow{p.s.} -(3/2) \|\eta\|^{-2} \begin{bmatrix} \eta^2 & -\eta^1 \\ 0 & 0 \end{bmatrix}. \quad \blacksquare$$

II. RÉGRESSION

Les résultats obtenus en I ont une traduction utile aux modèles de régression.

Soit $\mathbb{F} = (\mathcal{F}_n)$ une filtration et un « bruit adapté à la filtration \mathbb{F} », c'est-à-dire une suite $\varepsilon = (\varepsilon_n)$ de vecteurs aléatoires à valeurs dans \mathbb{C}^d , adaptée à \mathbb{F} et telle que, pour tout n :

$$E(\varepsilon_{n+1} | \mathcal{F}_n) = 0, \quad E(\varepsilon_{n+1} {}^* \varepsilon_{n+1} | \mathcal{F}_n) = \Gamma,$$

où Γ est une matrice de covariance déterministe, que nous appellerons « covariance » du bruit. Un « modèle de régression » de dimension d adapté à \mathbb{F} est une suite de la forme :

$$X_{n+1} = {}^* R \Phi_n + \varepsilon_{n+1}$$

où (Φ_n) est une suite de vecteurs aléatoires complexes de dimension q adaptée à \mathbb{F} , de carré intégrable et R une matrice $q \times d$. Notons

$$S_n = \sum_{k=0}^n \Phi_k * \Phi_k; \quad s_n = \sum_{k=0}^n \|\Phi_k\|^2 = \text{tr } S_n.$$

Dès que S_n est inversible, l'estimateur des moindres carrés est

$$\hat{R}_n = S_{n-1}^{-1} \left(\sum_{k=1}^n \Phi_{k-1} * X_k \right).$$

Pour éviter les problèmes d'inversibilité, on remplace, comme en I, S_n par $Q_n = Q + S_n$, Q hermitienne définie positive.

Posant $M_n = \sum_{k=1}^n \Phi_{k-1} * \varepsilon_k$, on a $\check{R}_n = \hat{R}_n - R = Q_{n-1}^{-1} [M_n - QR]$.

Soit $u \in \mathbb{C}^d$: $(\hat{R}_n - R)u = Q_{n-1}^{-1} [M_n - QR]u$; $M_n u = \sum_{k=1}^n \Phi_{k-1} \langle \varepsilon_k, u \rangle$ est

une martingale de variation quadratique $[*u \Gamma u] S_{n-1}$.

Il est d'autre part utile, notamment en contrôle adaptatif, de considérer les « prédicteurs » $(*\check{R}_n \Phi_n)$ et les « erreurs de prédiction » $\pi_n = *(\hat{R}_n - R) \Phi_n = *\check{R}_n \Phi_n$ (voir [8], [31]).

Pour l'analyse de ces erreurs de prédiction, on rappelle que, en notant

$$f_n = *\Phi_n Q_n^{-1} \Phi_n \quad \text{et} \quad g_n = *\Phi_n Q_{n-1}^{-1} \Phi_n; \quad (1-f_n) = (1+g_n)^{-1};$$

$$Q_{n-1}^{-1} - Q_n^{-1} = (1-f_n) Q_{n-1}^{-1} \Phi_n * \Phi_n Q_{n-1}^{-1} = (1+g_n) Q_n^{-1} \Phi_n * \Phi_n Q_n^{-1}.$$

Et, pour $u \in \mathbb{C}^d$:

$$*u \pi_n * \pi_n u = *u * [M_n - QR] Q_{n-1}^{-1} \Phi_n * \Phi_n Q_{n-1}^{-1} [M_n - QR] u$$

$$= (1-f_n)^{-1} *u * [M_n - QR] (Q_{n-1}^{-1} - Q_n^{-1}) [M_n - QR] u.$$

Lorsque $*u \Gamma u$ est nul, $\check{R}_n u$ tend vers 0, p. s. sur $\{\lambda_{\min} Q_n \rightarrow \infty\}$.

Sinon, on applique les théorèmes 1 et 2 à $M_n u$. Les variables aléatoires η_n que nous avons introduites sont ici proportionnelles à

$$*u * M_n (Q_{n-1}^{-1} - Q_n^{-1}) M_n u;$$

et la suite de ces variables est, p. s. du même ordre que

$$((1-f_n) *u \pi_n * \pi_n u).$$

COROLLAIRE 3. — Soit (\hat{R}_n) l'estimateur des moindres carrés de R dans un modèle de régression $X_{n+1} = *R \Phi_n + \varepsilon_{n+1}$. Les notations utilisées sont précisées ci-dessus.

1. Cet estimateur est fortement consistant sur l'ensemble des trajectoires où l'une des trois séries suivantes converge :

$$\sum (\lambda_{\min} Q_n)^{-1}, \quad \sum (\lambda_{\min} Q_n)^{-2} \|\Phi_n\|^2, \\ \sum (\lambda_{\min} Q_n)^{-1} [\text{Log det } Q_n - \text{Log Det } Q_{n-1}].$$

2. L'estimateur est faiblement consistant sous l'une ou l'autre des hypothèses suivantes :

(H1) $\lambda_{\min} S_n/n \xrightarrow{P} \infty;$

(H2) $\lambda_{\min} S_n/E(\text{Log } s_n) \xrightarrow{P} \infty.$

3. Pour toute fonction a réelle croissant vers ∞ , telle que $\int^\infty [1/a]$ soit convergente, on a, p. s.

$$\|S_{n-1}^{1/2} (\hat{R}_n - R)\| = o(a(n))^{1/2};$$

et, sur $\{s_n \rightarrow \infty\}$:

$$\|S_{n-1}^{1/2} (\hat{R}_n - R)\| = o(a[\text{Log } s_{n-1}])^{1/2}.$$

En outre, pour $f_n = * \Phi_n Q_n^{-1} \Phi_n$, et $\pi_n = * (\hat{R}_n - R) \Phi_n$, on a, p. s.

$$\sum [a(n)]^{-1} (1 - f_n) \pi_n * \pi_n < \infty, \\ \sum [a(\text{Log } s_n)]^{-1} (1 - f_n) \pi_n * \pi_n < \infty.$$

Commentaires. — (a) La partie (1) de ce corollaire est due à Le Breton-Musiela [20].

(b) Lai et Wei prouvent $\|S_{n-1}^{1/2} (\hat{R}_n - R)\| = o([\text{Log } s_{n-1}]^{1+\gamma})$ dans [16]; cela revient au même que l'utilisation de $a(\text{Log } s_{n-1})$.

(c) La partie (2) et la partie relative aux erreurs de prédiction sont nouvelles sans hypothèse complémentaire sur le bruit.

(d) On dit que le bruit « un moment conditionnel d'ordre > 2 » s'il existe un $\alpha > 2$ tel que $\sup_n E(\|\varepsilon_n\|^\alpha | \mathcal{F}_{n-1}) < \infty$. On obtient alors le résultat

plus précis suivant.

Pour tout $u \in \mathbb{C}^d$, posant

$$T_n(u) = \|S_{n-1}^{1/2} (\hat{R}_n - R) u\|^2 + \sum_{k=1}^{n-1} (1 - f_k) |\langle u, \pi_k \rangle|^2,$$

et $F_n = \sum_{k=1}^n f_k$, on a p. s. :

$$\lim [F_n]^{-1} T_{n+1}(u) = {}^*u \Gamma u (1 + \rho)$$

où ρ est une v. a. finie, nulle si $(F_n) \rightarrow \infty$.

La démonstration de ce résultat est une adaptation simple de celle donnée dans [31] dans le cas des modèles de régression unidimensionnels, c'est pourquoi nous ne la donnons pas. On en déduit $\|S_{n-1}^{1/2}(\hat{R}_n - R)u\| = O(\inf[n, \text{Log } s_{n-1}]^{1/2})$, relation établie dans [16] pour le $\text{Log } s_{n-1}$ et [20] pour n . En ce qui concerne les erreurs de prédiction, elle implique :

$$\sum_{k=1}^n (1-f_k) \|\pi_k\|^2 = O(\inf[n, \text{Log } s_n]).$$

Ces résultats relatifs aux erreurs de prédiction s'appliquent bien lorsque $(f_n) \rightarrow 0$, ou lorsque (f_n) a p. s. une limite < 1 . Voir [7] pour leur application aux modèles autorégressifs vectoriels.

COROLLAIRE 4. — Avec les notations du corollaire 3, on pose :

$$P_n = \sum_{k=1}^n \pi_k {}^*\pi_k.$$

On suppose ici que le bruit satisfait la loi des grands nombres :

$\frac{1}{n} \sum_{k=1}^n \varepsilon_k {}^*\varepsilon_k \xrightarrow{\text{p. s.}} \Gamma$; c'est le cas, en particulier, si le bruit est blanc ou s'il a

un moment conditionnel d'ordre > 2 .

1. Soit l'« estimateur empirique de la covariance »

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

Si $(P_n/n) \xrightarrow{\text{p. s.}} P$, alors $(\hat{\Gamma}_n) \xrightarrow{\text{p. s.}} \Gamma + P$.

$(\hat{\Gamma}_n)$ est donc fortement consistant si $P=0$; c'est en particulier le cas si $(f_n) \rightarrow 0$ et si $a[\text{Log } s_n] = O(n)$.

2. On a, p. s., $s_n^{-1} P_n \rightarrow 0$ sur l'ensemble

$$\{s_n \rightarrow \infty\} \cap \{ \inf[a(n), a(\text{Log } s_n)] = O(\lambda_{\min} S_{n-1}) \}.$$

3. Si le bruit a un moment conditionnel d'ordre > 2 fini, on a, $s_n^{-1} P_n$ tend, p. s. vers 0 sur :

$$\{s_n \rightarrow \infty\} \cap \{ \text{Log } s_n = o(\lambda_{\min} S_{n-1}) \};$$

l'estimateur empirique de la covariance est fortement consistant si : $s_n = O(n)$, $\text{Log } n = o(\lambda_{\min} S_n)$.

Démonstration. – (a) Reprenant les notations du corollaire 3.

$$\begin{aligned} n *u \hat{\Gamma}_n u &= \sum_{k=1}^n *u (\varepsilon_k - * \check{R}_{k-1} \Phi_{k-1}) * (\varepsilon_k - * \check{R}_{k-1} \Phi_{k-1}) u \\ &= \sum_{k=1}^n *u \varepsilon_k * \varepsilon_k u + *u P_{n-1} u - N_n \end{aligned}$$

où $N_n = \sum_{k=0}^{n-1} 2 \text{Re} \langle u, \varepsilon_{k+1} \rangle \langle \pi_k, u \rangle$; (N_n) est une martingale et $\langle N \rangle_n \leq 4 *u \Gamma u *u P_{n-1} u$. Donc, si $*u \Gamma u = 0$, $*u \hat{\Gamma}_n u$ tend p. s. vers 0; et, si $*u \Gamma u \neq 0$, $N_n = \eta_n *u P_{n-1} u$ où (η_n) converge p. s., la limite étant nulle si $(*u P_n u)$ tend vers ∞ . L'estimateur empirique de la covariance est donc fortement consistant si P_n/n tend p. s. vers 0.

(b) $(1 - f_n) = (1 + g_n)^{-1} \geq [s_n + \text{tr } Q]^{-1} (\lambda_{\min} S_{n-1})$. On déduit dès lors de la partie (3) du corollaire 3

$$\sum s_n^{-1} \pi_n * \pi_n < \infty,$$

d'où la partie (2) du corollaire par le lemme de Kronecker.

(c) Dans le cadre de la partie (3), p. s., il existe un β fini tel que

$$\sum_{k=1}^n (1 - f_k) \|\pi_k\|^2 \leq \beta \text{Log } s_n.$$

Pour tout $\alpha > 0$, si n est assez grand, $\lambda_{\min} S_{n-1} \geq \alpha \text{Log } s_n$; pour n assez grand :

$$\sum_{k=1}^n \alpha s_k^{-1} \text{Log } s_k *u \pi_k * \pi_k u \leq \beta (\text{Log } s_n) + \text{Cte};$$

vu que la suite $(s_n^{-1} \text{Log } s_n)$ est décroissante pour n assez grand, il en résulte facilement :

$$\alpha s_n^{-1} \text{Log } s_n \sum_{k=1}^n *u \pi_k * \pi_k u \leq \beta (\text{Log } s_n) + \text{Cte};$$

cela étant vrai pour tout α , on obtient la partie (3) du corollaire. ■

III. LOIS DU LOGARITHME ITÉRÉ

III.1. Une majoration simple de type « logarithme itéré »

Comme l'on remarqué Wei dans [30], Darwich [5], ou Heunis [9], on peut parfois améliorer les vitesses de convergence presque sûres des estimateurs des moindres carrés lorsque la loi du logarithme itéré s'applique.

Considérons la fonction $h(x) = (2x \text{LL } x)^{1/2}$ où $\text{LL } x$ est le « logarithme itéré » $\text{Log}(\text{Log } x)$. La démonstration de la proposition suivante est à peu près celle de Stout ([28], [29]). Nous nous contenterons ici d'une majoration simple du type « loi du logarithme itéré » qui nous semble bien adaptée aux estimateurs des moindres carrés parce qu'elle n'utilise que des majorations des traces de variations quadratiques prévisibles des martingales.

C'est surtout une borne constante presque sûre qui est utile.

Voici l'un des résultats de ce type relatif aux martingales.

PROPOSITION 5. — Soit (M_n) une martingale adaptée à \mathbb{F} et (s_n^2) une suite adaptée qui tend, p. s., vers ∞ et telle que $\langle M \rangle_n \leq s_{n-1}^2$; $h(x) = [2x \text{LL } x]^{1/2}$.

(1) On suppose $|M_{n+1} - M_n| \leq C s_n^2 [h(s_n^2)]^{-1}$ où C est \mathcal{F}_0 -mesurable et $C < 1$. Alors, p. s. :

$$\limsup |M_n| / h(s_{n-1}^2) \leq 1 + C/2.$$

(2) On suppose $|M_{n+1} - M_n| \leq C_n s_n^2 [h(s_n^2)]^{-1}$ où (C_n) est une suite adaptée tendant vers 0; alors, p. s. :

$$\limsup |M_n| / h(s_{n-1}^2) \leq 1.$$

Démonstration. — Il suffit d'étudier $\limsup M_n / h(s_{n-1}^2)$; en considérant $-M$, on en déduit le résultat relatif à $|M_n|$.

Comme en [28] et [29], l'outil est la surmartingale exponentielle définie pour $\lambda \geq 0$ et $c > 0$ par,

$$Z_n(\lambda) = \exp[\lambda(M_n - M_0) - \Phi_c(\lambda) \langle M \rangle_n]$$

où

$$\Phi_c(\lambda) \leq [\lambda^2/2](1 + \lambda c) \quad \text{lorsque } \lambda c \leq 1.$$

Soit $\alpha > 0$, \mathcal{F}_0 -mesurable,

$$P(\sup_{k \leq n} \lambda(M_k - M_0) \geq \Phi_c(\lambda) \langle M \rangle_n + \alpha \lambda \mid \mathcal{F}_0) \leq P(\sup_{n \geq 0} Z_n(\lambda) \geq e^{\alpha \lambda} \mid \mathcal{F}_0) \leq e^{-\alpha \lambda}$$

(a) Prouvons (1). Soit θ , $1 < \theta < 1/C$, $t_k = \inf\{n; s_n^2 \geq \theta^k\}$; c'est un temps d'arrêt. On considère $N_n = M_{n \wedge t_{k+1}}$, x_0 tel que la fonction $x \rightarrow x/h(x)$ soit croissante sur $[x_0, \infty[$, $x_0 = 6$ par exemple, et $v = \inf\{n; s_n^2 \geq x_0\}$.

Pour k assez grand, $t_k \geq v$ et $H_n = N_{n+v}$ vérifie :

$$|H_{n+1} - H_n| = |M_{n+v+1} - M_{n+v}| \mathbb{1}_{\{t_{k+1} > n+v\}} \leq C \theta^{k+1} / h(\theta^{k+1}) \leq C \theta^{k+1} / h(\theta^k) = d_k.$$

Par conséquent, pour $0 \leq \lambda d_k \leq 1$:

$$P(\sup_{0 \leq j \leq n} [H_j - H_0] \geq (\lambda/2)(1 + \lambda d_k) \langle H \rangle_n + \alpha \mid \mathcal{F}_v) \leq e^{-\lambda \alpha}.$$

Or $\langle H \rangle_n \leq \langle M \rangle_{t_{k+1}} \leq s_{t_{k+1}-1}^2 \langle \theta^{k+1} \rangle$ et $H_0 = M_v$. D'où :

$$P(\sup_{v \leq n \leq t_{k+1}} [M_n - M_v] \geq (\lambda/2)(1 + \lambda d_k) \theta^{k+1} + \alpha \mid \mathcal{F}_v) \leq e^{-\lambda \alpha}.$$

Soit $\delta > 0$ et

$$\begin{aligned} \lambda &= h(\theta^k) \theta^{-k}, & \alpha &= (1 + \delta) h(\theta^k) / 2; \\ \alpha \lambda &= (1 + \delta) \text{Log}(k \text{Log } \theta); & \lambda d_k &= C \theta < 1; \\ (\lambda/2)(1 + \lambda d_k) \theta^{k+1} + \alpha &= h(\theta^k) [1 + \delta + \theta(1 + \theta C)] / 2. \end{aligned}$$

Compte tenu du fait que $t_k \geq v$ et que, lorsque $t_k < n \leq t_{k+1}$,

$$h(s_{n-1}^2) \geq h(s_{t_k}^2) \geq h(\theta^k),$$

on a :

$$P(\sup_{t_k < n \leq t_{k+1}} [h(s_{n-1}^2)]^{-1} (M_n - M_v) \geq [1 + \delta + \theta(1 + \theta C)] / 2 \mid \mathcal{F}_v) \leq (k \text{Log } \theta)^{-1 - \delta}.$$

On en déduit par le lemme de Borel-Cantelli que, p. s. :

$$\limsup [h(s_{n-1}^2)]^{-1} M_n \leq [1 + \delta + \theta(1 + \theta C)] / 2.$$

D'où, $\delta > 0$ et $\theta > 1$ étant arbitraires, p. s. :

$$\limsup [h(s_{n-1}^2)]^{-1} M_n \leq 1 + C/2.$$

(b) La partie (2) en résulte facilement. Posons

$$\zeta_{n+1} = M_{n+1} - M_n \quad \text{et, pour } 0 < C < 1, \quad \chi_{n+1} = \zeta_{n+1} \mathbb{1}_{\{C_n \leq C\}}.$$

On a $|\chi_{n+1}| \leq C s_n^2 [h(s_n^2)]^{-1}$ et $E(\chi_{n+1} \mid \mathcal{F}_n) = 0$; donc, d'après la partie (a) de la démonstration, p. s.,

$$\limsup [h(s_{n-1}^2)]^{-1} \sum_{k=1}^n \chi_k \leq 1 + C/2.$$

Mais, $\limsup \{\zeta_n \neq \chi_n\} = \limsup \{C_{n-1} > C\} = \emptyset$. Comme C est choisi arbitraire avec $0 < C < 1$, on a, p. s. :

$$\limsup_{n \rightarrow \infty} [h(s_{n-1}^2)]^{-1} M_n \leq 1. \quad \blacksquare$$

COROLLAIRE 6. — Soit (ε_n) une suite de variables aléatoires adaptée à \mathbb{F} et telle que, p. s. :

- quel que soit n , $E(\varepsilon_{n+1} | \mathcal{F}_n) = 0$, $E(\varepsilon_{n+1}^2 | \mathcal{F}_n) \leq \sigma^2$;
- pour un $0 < \alpha < 1$, $\sup_n E(|\varepsilon_{n+1}|^{2+2\alpha} | \mathcal{F}_n) < \infty$.

Soit (Y_n) une suite adaptée à \mathbb{F} , et (T_n) une suite adaptée à \mathbb{F} telle que $|Y_n| \leq T_n$. On pose :

$$M_n = \sum_{k=1}^n Y_{k-1} \varepsilon_k, \quad \tau_n = \sum_{k=0}^n T_k^2.$$

On suppose que (τ_n) croît vers ∞ et que

$$\sum |T_n|^{2+2\alpha} \tau_n^{-1-\alpha} < \infty.$$

Alors :

$$\limsup [2 \tau_{n-1} \text{LL} \tau_{n-1}]^{-1/2} |M_n| \leq \sigma \quad (\text{p. s.}).$$

Démonstration. — Soit (C_n) une suite positive adaptée à \mathbb{F} et décroissant vers 0 qu'on déterminera ultérieurement.

Soit $\Gamma_{n+1} = (|Y_n \varepsilon_{n+1}| \leq C_n \sigma^2 \tau_n [h(\sigma^2 \tau_n)]^{-1})$ et G_{n+1} son complémentaire. Posons :

$$\begin{aligned} \xi_{n+1} &= Y_n (\varepsilon_{n+1} \mathbb{1}_{\Gamma_{n+1}} - E[\varepsilon_{n+1} \mathbb{1}_{\Gamma_{n+1}} | \mathcal{F}_n]), \\ \zeta_{n+1} &= Y_n \varepsilon_{n+1} - \xi_{n+1} = Y_n (\varepsilon_{n+1} \mathbb{1}_{G_{n+1}} - E\{\varepsilon_{n+1} \mathbb{1}_{G_{n+1}} | \mathcal{F}_n\}); \\ X_n &= \sum_{k=1}^n \xi_k \quad \text{et} \quad Z_n = \sum_{k=1}^n \zeta_k = M_n - X_n. \end{aligned}$$

(a) (X_n) est une martingale et, comme $E[\xi_n^2 | \mathcal{F}_{n-1}] \leq \sigma^2 T_{n-1}^2$, la proposition 5 s'applique à (X_n) en prenant $s_n^2 = \sigma^2 \tau_n$. Comme,

$$h(s_n^2) = \sigma [2 \tau_n \text{Log}(\text{Log} \tau_n + 2 \text{Log} \sigma)]^{1/2},$$

la suite $(h(s_n^2))$ équivaut à $(\sigma h(\tau_n))$ et, p. s.,

$$\limsup [h(\tau_{n-1})]^{-1} |X_n| \leq \sigma.$$

(b) Soit $N_n = \sum_{k=1}^n \zeta_k [h(s_{k-1}^2)]^{-1}$; (N_n) est une martingale qui converge,

p. s. En effet, le moment conditionnel d'ordre $2 + 2\alpha$ de ε_{n+1} par rapport à \mathcal{F}_n étant majoré par une variable aléatoire finie L indépendante de n ,

$$\begin{aligned} E[\zeta_{n+1}^2 | \mathcal{F}_n] &\leq T_n^2 E[\varepsilon_{n+1}^2 \mathbb{1}_{G_{n+1}} | \mathcal{F}_n] \leq T_n^2 L [|Y_n| [h(s_n^2)] s_n^{-2} C_n^{-1}]^{2\alpha} \\ \langle N \rangle_{n+1} - \langle N \rangle_n &\leq \text{Cte} L T_n^{2+2\alpha} [h(s_n^2)]^{2\alpha-2} [s_n^2]^{2\alpha} C_n^{-2\alpha} \\ &= \text{Cte} L T_n^{2+2\alpha} s_n^{-2-2\alpha} (\text{LL} s_n^2)^{\alpha-1} C_n^{-2\alpha}. \end{aligned}$$

En choisissant $C_n^2 = (LL s_n^2)^{1-1/\alpha}$, la condition du corollaire assure que $\langle N \rangle_\infty < \infty$. Donc, p. s., (N_n) converge; comme les suites $(h(s_n^2))$ et $(\sigma h(\tau_n))$ sont équivalentes, par le lemme de Kronecker, $[h(\tau_{n-1})]^{-1} Z_n \xrightarrow{p. s.} 0$. D'où le résultat annoncé. ■

Exemple. — Si $T_n = an^\nu$ avec $a > 0$ et $\nu \geq 0$, $\tau_{n-1} \leq a^2(n+1)^{2\nu+1}/(2\nu+1)$. D'où, p. s. :

$$\limsup [n^{-\nu-1/2} (LL n)^{-1/2}] |M_n| \leq a \sigma (\nu + 1/2)^{-1/2}.$$

III.2. Application au modèle de régression

Dans le cadre du modèle de régression décrit en II, le corollaire 6 a une application immédiate, à condition de supposer que le bruit a un moment conditionnel d'ordre > 2 fini :

Si p. s., $\|\Phi_n\|^2 \leq T_n$, $\tau_n = \sum_{k=0}^n T_k$, si la suite (τ_n) tend vers ∞ et si $\sum T_n^{2+2\alpha} \tau_n^{-1-\alpha} < \infty$, alors, p. s., pour tout $u \in \mathbb{C}^d$ de norme 1 :

$$\limsup [2 \tau_{n-1} LL \tau_{n-1}]^{-1/2} \|S_{n-1}(\hat{R}_n - R)u\| \leq (q \lambda_{\max} \Gamma)^{1/2}.$$

RÉFÉRENCES

- [1] T. W. ANDERSON et J. TAYLOR, Strong Consistency of Least Squares Estimators in Dynamic Models, *Ann. Stat.*, vol. 7, 1979, p. 484-489.
- [2] Y. S. CHOW, A Martingale Inequality and the Law of Large Numbers, *Proc. Am. Math. Soc.*, vol. 11, 1960, p. 107-111.
- [3] Y. S. CHOW, Local Convergence of Martingales and the Law of Large Numbers, *Ann. Math. Stat.*, vol. 36, 1965, p. 493-507.
- [4] N. CHRISTOPEIT, Quasi-least Squares Estimation in Semi-Martingale Regression Models, *Stochastics*, vol. 16, 1986, p. 255-278.
- [5] A. R. DARWICH, Une loi du logarithme itéré pour les martingales locales multidimensionnelles et son application en régression linéaire stochastique, *C. R. Acad. Sci. Paris*, série I, 1989, p. 387-390.
- [6] D. DACUNHA-CASTELLE et M. DUFLO, *Probabilités et statistiques*, tome 2, Masson, 1983.
- [7] M. DUFLO, R. SENOSSI et A. TOUATI, Propriétés asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autorégressif vectoriel, *Ann. Inst. Henri Poincaré, série B*, vol. 27, n° 1, 1991.
- [8] W. A. FULLER et D. P. HASZA, Properties of Predictors for Autoregressive Time Series, *J. Am. Stat. Assoc.*, vol. 76, 1981, p. 155-161.

- [9] A. J. HEUNIS, Asymptotic Properties of Prediction Error Estimators in Approximate System Identification, *Stochastic*, vol. **24**, 1988, p. 1-43.
- [10] H. KAUFMANN, On the Strong Law of Large Numbers for Multivariate Martingales, *Stoch. Proc. Appl.*, vol. **26**, 1987, p. 73-85.
- [11] P. R. KUMAR, Convergence of Adaptive Control Schemes using Least Squares Parameter Estimates, *I.E.E.E. Trans. Aut. Control.*, 1990, (à paraître).
- [12] T. L. LAI, Asymptotically Efficient Adaptive Control in Stochastic Regression Models, *Adv. Appl. Math.*, vol. **7**, 1986, p. 23-45.
- [13] T. L. LAI et H. ROBBINS, Adaptive Design and Stochastic Approximation, *Ann. Stat.*, vol. **7**, n° 6, 1979, p. 1196-1221.
- [14] T. L. LAI et H. ROBBINS, Consistency and Asymptotic Efficiency of Slope Estimates in Stochastic Approximation Schemes, *Z. Wahr. Verw. Gebiete*, vol. **56**, 1981, p. 329-360.
- [15] T. L. LAI, H. ROBBINS et C. Z. WEI, Strong Consistency of Least Squares Estimates in Multiple Regression, II, *J. Mult. Anal.*, vol. **9**, 1979, p. 343-361.
- [16] T. L. TAI 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.
- [17] T. L. LAI et C. Z. WEI, Asymptotic Properties of Projections with Applications to Stochastic Regression Problems, *J. Mult. Anal.*, vol. **12**, 1982, p. 346-370.
- [18] T. L. TAI et C. Z. WEI, Asymptotic Properties of General Autoregressive Models and Strong Consistency of Least Squares Estimates and their Parameters, *J. Mult. Anal.*, vol. **13**, 1983, p. 1-23.
- [19] T. L. TAI et C. Z. WEI, Asymptotic Properties of Multivariate Weighted Sums with Applications to Stochastic Regression in Linear Dynamic Systems, *Multivariate analysis*, VI, éd. P. R. KRISHNAIAH, North-Holland, 1985, p. 375-393.
- [20] A. LE BRETON et M. MUSIELA, Consistency Sets of Least Squares Estimates in Stochastic Regression Models, *Stoch. diff. systems, Lect. Notes Control Inf. Sci.*, vol. **126**, 1989.
- [21] A. LE BRETON et M. MUSIELA, Laws of Large Number for Semimartingales with Applications to Stochastic Regression, *Probab. Theor. Rel. Fields*, vol. **81**, 1989, p. 275-290.
- [22] D. LÉPINGLE, Sur le comportement asymptotique des martingales locales, *Lect. Notes Prob.*, vol. **649**, Springer, 1978.
- [23] P. LÉVY, Sur les séries dont les termes sont des variables indépendantes, *Studia Math.*, vol. **3**, 1931, p. 119-155.
- [24] A. V. MELNIKOV, The Law of Large Numbers for Multidimensional Martingales, *Soviet Math. Dokl.*, vol. **33**, 1986, p. 131-135.
- [25] J. NEVEU, Bases mathématiques du calcul des probabilités, Masson, 1964 (2^e édition 1970).
- [26] H. ROBBINS et D. SIEGMUND, A Convergence Theorem for Non Negative Almost Supermartingales and some Applications, dans J. S. RUSTAGI, *Optimization methods in statistics*, Academic Press, 1971, p. 233-257.
- [27] V. SOLO, Topics in Advanced Time Series Analysis, *Lect. Notes Math.*, vol. **1215**, Springer-Verlag, 1982.
- [28] W. STOUT, *Almost Sure Convergence*, Academic Press.
- [29] W. STOUT, A Martingale Analogue of Kolmogorov's Law of Iterated Logarithm, *Z. Wahrscheinlichkeitstheorie*, vol. **15**, 1970, p. 279-290.
- [30] C. Z. WEI, Asymptotic Properties of Least Squares Estimates in Stochastic Regression Models, *Ann. Stat.*, vol. **13**, 1985, p. 1498-1508.
- [31] C. Z. WEI, Adaptive Prediction by Least Squares Predictors in Stochastic Regression Models with Applications to Time Series, *Ann. Stat.*, vol. **15**, 1987, p. 1667-1682.

(Manuscrit reçu le 14 novembre 1989;
corrigé le 14 mai 1990.)