

STATISTIQUE ET ANALYSE DES DONNÉES

JEAN-MARIE MONNEZ

Étude d'un modèle de corrélation en approximation stochastique

Statistique et analyse des données, tome 16, n° 3 (1991), p. 143-181

http://www.numdam.org/item?id=SAD_1991__16_3_143_0

© Association pour la statistique et ses utilisations, 1991, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » 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/>

ETUDE D'UN MODELE DE CORRELATION EN APPROXIMATION STOCHASTIQUE

Jean-Marie MONNEZ

Laboratoire de Mathématiques
Université de Nancy I - B.P. 239
54506 - VANDOEUVRE-LES-NANCY Cedex

Résumé : *Dans cet article est étudiée la convergence du processus stochastique (X_n) dans \mathbb{R}^k tel que $X_{n+1} = X_n - a_n h_n(X_n) - a_n U_n$, lorsqu'il existe un modèle de corrélation entre les variables aléatoires U_n . La convergence d'un tel processus a été étudiée dans le cas $h_n = h$ ne dépendant pas de n par la méthode de l'équation différentielle ordinaire. Ici, aucune limite en quelque sens que ce soit n'est imposée à la suite de fonctions (h_n) et le processus ne peut s'écrire de façon naturelle sous la forme $X_{n+1} = X_n - a_n h(X_n) - a_n U_n$, où la fonction h ne dépend pas de n . Des théorèmes de convergence presque sûre et de comportement asymptotique du moment d'ordre deux sont établis, d'une part dans le cas général, en utilisant la méthode de démonstration dite des martingales, d'autre part dans le cas linéaire, où $h_n(X_n) = G_n \cdot (X_n - \theta_n)$, avec des hypothèses plus faibles ; des corollaires dans le cas $a_n = 1/n^\alpha$ sont donnés, ainsi qu'un corollaire d'un théorème de normalité asymptotique de Fabian dans le cas linéaire. L'utilisation d'un processus de ce type dans l'estimation du paramètre d'un modèle de régression linéaire généralisée est étudiée.*

Mots-clés : approximation stochastique, modèle de corrélation, convergence p.s., moment d'ordre deux, normalité asymptotique, régression linéaire généralisée.

Classification AMS : 62 L 20.

Classification STMA : 04 100.

Abstract : *In this paper is studied the convergence of the stochastic approximation process (X_n) in \mathbb{R}^k such that $X_{n+1} = X_n - a_n h_n(X_n) - a_n U_n$, when the random variables U_n are correlated. The convergence study of this type of process has been made in the case $h_n = h$ not dependent on n by the ordinary differential equation method. Here no limit in any sense is required for the sequence of functions (h_n) and one can't write the process in a natural way in the form $X_{n+1} = X_n - a_n h(X_n) - a_n U_n$, where the function h is not dependent on n . The correlation model used here is defined by hypotheses on the process (W_n) in \mathbb{R}^k such that $W_{n+1} = (1 - a_n)W_n + a_n U_n$; this model has been studied by Ljung.*

First an a.s. convergence theorem is established, including a stability hypothesis on the sequence of functions (h_n) extending the classical one ; a deterministic version of the almost supermartingales convergence lemma of Robbins and Siegmund is used in the proof. Then a theorem on the asymptotic behaviour of the moment of order two is given, with a corollary in the case $a_n = 1/n^\alpha$; a lemma of the Chung type is used in the proof. These results are applied to the estimation of the parameter of a generalized linear regression model.

Next the linear case $h_n(X_n) = G_n \cdot (X_n - \theta_n)$ is studied, G_n being a random variable in the space of matrices $k \times k$ and θ_n the parameter to be estimated at time n . The same results as in the general case are established but with less restrictive hypotheses and with different methods of proof. A corollary in the case $a_n = 1/n^\alpha$ and G_n of positive type is given. The a.s. convergence result can be used when the sequence (θ_n) is not convergent but varies slowly with time ; when $a_n = 1/n$ the classical ergodicity hypothesis of the process (G_n) is replaced by a different condition. Finally a corollary of an asymptotic normality theorem of Fabian is given for this process.

Keywords : stochastic approximation, correlation model, a.s. convergence, moment of order two, asymptotic normality, generalized linear regression model.

1 - INTRODUCTION

Dans cet article est étudiée la convergence du processus d'approximation stochastique (X_n) dans \mathbb{R}^k tel que

$$X_{n+1} = X_n - a_n h_n(X_n) - a_n U_n, \quad (1.1)$$

(h_n) étant une suite de fonctions de \mathbb{R}^k dans \mathbb{R}^k , (U_n) une suite de variables aléatoires dans \mathbb{R}^k , (a_n) une suite de nombres réels positifs, lorsqu'il existe une structure de corrélation entre les variables aléatoires U_n .

A l'origine de la définition d'un tel processus, il y a le problème suivant : soit une famille de variables aléatoires $Y_n(x)$ dans \mathbb{R}^p , $x \in \mathbb{R}^k$, $n \in \mathbb{N}$, d'espérance mathématique $M_n(x)$ qui a pour zéro θ_n ; on cherche à estimer θ_n au temps n . Pour cela, on définit le processus (X_n) tel que

$$X_{n+1} = X_n - a_n A_n(X_n) Y_n,$$

Y_n représentant une observation de $Y_n(X_n)$, A_n étant une application mesurable définie sur \mathbb{R}^k à valeurs matricielles $k \times p$, a_n un nombre réel positif. En décomposant Y_n sous la forme $M_n(X_n) + Z_n$, on a

$$X_{n+1} = X_n - a_n A_n(X_n) M_n(X_n) - a_n A_n(X_n) Z_n,$$

ce qui rentre dans le cadre général (1.1).

Dans les processus d'approximation stochastique du type Robbins-Monro, on suppose que la variable aléatoire Y_n est conditionnellement indépendante à X_n fixé de $X_1, Y_1, Y_2, \dots, Y_{n-1}$; c'est l'hypothèse dite d'indépendance ou de Robbins-Monro [16].

L'étude des algorithmes stochastiques avec observations corrélées, où cette hypothèse n'est plus vérifiée, a été rendue nécessaire par leurs applications dans l'identification des systèmes, le filtrage adaptatif, la reconnaissance des formes, etc... Par exemple, en filtrage adaptatif linéaire, on a étudié l'algorithme de Widrow

$$X_{n+1} = X_n - a_n Z_n (Z_n^T X_n - c_n),$$

où c_n représente un signal de référence, Z_n un vecteur de signaux observés, les couples (c_n, Z_n) n'étant pas indépendants ; différents articles sur ce sujet sont rassemblés dans [9]. Un autre exemple concerne l'identification des systèmes linéaires [12] ; Ljung définit pour le vecteur Y_n des mesures effectuées sur le système une dynamique markovienne conditionnellement linéaire,

$$Y_n = A(X_n)Y_{n-1} + B(X_n)W_n,$$

où (W_n) est une suite de variables aléatoires indépendantes de même loi. Une généralisation de cette approche est faite par Métivier et Priouret, qui définissent entre les Y_n une dépendance markovienne [4], [13], [14].

Les algorithmes stochastiques avec observations corrélées qui ont été étudiés peuvent se ramener au type suivant,

$$X_{n+1} = X_n - a_n h(X_n) - a_n U_n, \quad (1.2)$$

la fonction h ne dépendant pas du temps n . Ils peuvent être considérés comme des perturbations stochastiques de l'algorithme déterministe

$$x_{n+1} = x_n - a_n h(x_n),$$

auquel est associée l'équation différentielle ordinaire

$$\frac{dx(t)}{dt} = -h(x(t)).$$

Le comportement asymptotique de X_n est lié à celui des solutions de cette équation différentielle.

La différence essentielle entre les algorithmes du type (1.1) et du type (1.2) est que dans (1.1), on n'impose pas de limite en quelque sens que ce soit à la suite de fonctions (h_n) ; au temps n , on a une fonction h_n dépendant de n et le processus ne peut s'écrire naturellement sous la forme (1.2) faisant intervenir une fonction h ne dépendant pas du temps n .

La démonstration de la convergence fera intervenir la technique dite des martingales, utilisée dans l'étude des processus d'approximation stochastique du type Robbins-Monro, avec des développements spécifiques.

Un exemple d'utilisation d'un processus du type (1.1) concerne l'estimation par approximation stochastique du paramètre d'un modèle de régression. Albert et Gardner ont consacré un livre [1] à ce sujet dans le cas d'observations indépendantes. Pour tout n , soit Y_n une variable aléatoire réelle observable, F_n une fonction connue de \mathbb{R}^k dans \mathbb{R} ; on considère le modèle de régression $Y_n = F_n(\theta) + Z_n$, Z_n étant la variable aléatoire résiduelle et θ le paramètre inconnu à estimer appartenant à \mathbb{R}^k . Soit une application mesurable A_n de \mathbb{R}^k dans \mathbb{R} ; pour estimer θ , on définit le processus d'approximation stochastique (X_n) dans \mathbb{R}^k tel que

$$X_{n+1} = X_n - a_n A_n(X_n)(F_n(X_n) - Y_n),$$

ce qui s'écrit également

$$X_{n+1} = X_n - a_n A_n(X_n)(F_n(X_n) - F_n(\theta)) + a_n A_n(X_n) Z_n$$

et rentre dans le cadre général (1.1) avec $h_n(x) = A_n(x)(F_n(x) - F_n(\theta))$ et $U_n = -A_n(X_n)Z_n$. Dans cet article, nous prendrons comme exemples un modèle linéaire et un modèle linéaire généralisé.

La structure de corrélation entre les variables aléatoires U_n que l'on considère ici est celle utilisée par Ljung dans [10] ; elle est définie par des hypothèses portant sur la suite de variables aléatoires (W_n) dans \mathbb{R}^k telle que

$$W_{n+1} = (1 - a_n)W_n + a_n U_n, \quad W_1 = 0. \quad (1.3)$$

Remarquons que pour $a_n = 1/n$, W_{n+1} est la moyenne d'ordre n des U_i . On peut trouver dans [3] plusieurs ensembles d'hypothèses sur les U_n qui impliquent la convergence presque sûre de W_n vers 0.

Dans le paragraphe 2.1, nous donnons un théorème de convergence presque sûre du processus $X_n + W_n - \theta_n$ vers 0, la suite (θ_n) étant convergente, avec une hypothèse de stabilité sur les fonctions h_n qui étend l'hypothèse classique, et rappelons des résultats de Ljung [10] sur le modèle de corrélation utilisé. La démonstration fait intervenir une version déterministe du lemme de convergence des pseudo-sur-martingales de Robbins et Siegmund [17].

Dans le paragraphe 2.2, nous faisons l'étude du comportement asymptotique du moment d'ordre deux de $\|X_n + W_n - \theta_n\|$; le théorème énoncé est un corollaire pour $a_n = 1/n^\alpha$, avec un résultat du type $\overline{\lim} n^\beta E[\|X_n - \theta_n\|^2] < \infty$, d'une proposition plus générale que nous démontrons. Nous utilisons un lemme du type Chung [5], dont la démonstration, faite dans [6], est rappelée dans un appendice.

Ces deux résultats sont appliqués dans le paragraphe 2.3 à l'estimation du paramètre d'un modèle de régression linéaire généralisée.

Le paragraphe 3 est consacré à l'étude du cas linéaire, où $h_n(X_n)$ prend la forme $G_n(X_n - \theta_n)$, G_n étant une variable aléatoire matricielle carrée d'ordre k . Nous avons dans ce cas un allègement des hypothèses des théorèmes. Les techniques de démonstration utilisées ici avec des développements spécifiques ont été introduites par Albert et Gardner [1].

Nous établissons dans le paragraphe 3.1 un théorème de convergence presque sûre de $X_n + W_n - \theta_n$ vers 0. Ce résultat peut être utilisé dans le cas où la suite (θ_n) n'est pas convergente, mais où le paramètre θ_n varie lentement dans le temps ; dans le cas $a_n = 1/n$, il ne fait pas intervenir l'hypothèse d'ergodicité du processus (G_n) que

l'on trouve dans les résultats classiques, [11] par exemple, et qui est remplacée par une autre condition. Nous utilisons un lemme du type Chung [6], dont la démonstration est rappelée dans l'appendice.

Dans le paragraphe 3.2, nous étudions le comportement asymptotique du moment d'ordre deux de $\|X_n + W_n - \theta_n\|$. Le théorème énoncé est un corollaire dans le cas $a_n = 1/n^\alpha$, G_n de type positif, d'une proposition plus générale que nous démontrons. Une première forme de ce résultat dans le cas d'un processus avec observations non corrélées a été donnée dans [2].

Dans le paragraphe 3.3, nous donnons un théorème de normalité asymptotique du processus $X_n + W_n - \theta_n$, qui est établi comme corollaire d'un théorème de Fabian [7] rappelé dans l'appendice.

2 - CAS GENERAL

On considère les processus (X_n) et (W_n) dans \mathbb{R}^k tels que

$$\begin{aligned} X_{n+1} &= X_n - a_n h_n(X_n) - a_n U_n, \\ W_{n+1} &= (1 - a_n) W_n + a_n U_n, \quad W_1 = 0. \end{aligned}$$

2.1 - Convergence presque sûre

Théorème 1 :

On suppose les hypothèses suivantes vérifiées :

(1) Il existe une suite de réels (θ_n) , deux suites de réels positifs (d_n) et (e_n) , un entier r , une suite croissante d'entiers (n_ℓ) vérifiant $n_{\ell+1} \leq n_\ell + r$, tels que :

- (i) pour tout n et tout x , $\|h_n(x)\| \leq d_n \|x - \theta_n\| + e_n$;
- (ii) pour tout n et tout x , $\langle x - \theta_n, h_n(x) \rangle \geq 0$;
- (iii) pour tout $0 < \varepsilon < 1$,

$$b(\varepsilon) = \inf_{\ell} \inf_{\{x: \|x - \theta_{n_\ell}\| < \frac{1}{\varepsilon}\}} \sum_{j \in I_\ell} \langle x - \theta_j, h_j(x) \rangle > 0,$$

où I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

(2) Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$(\|x_1 - x_2\| < \eta) \Rightarrow (\sup_n \|h_n(x_1) - h_n(x_2)\| < \varepsilon).$$

$$(3) \text{ Pour tout } n, a_n > 0; \sum_1^\infty \min_{j \in I_{\lambda}} a_j = +\infty; \sum_1^\infty a_n^2 d_n^2 < \infty; \sum_1^\infty a_n^2 e_n^2 < \infty.$$

$$(4) \sum_1^\infty \|\theta_{n+1} - \theta_n\| < \infty; \sum_1^\infty a_n \|h_n(X_n + W_n) - h_n(X_n) - W_n\| < \infty \text{ p.s.}$$

Soit $\theta = \lim_{n \rightarrow \infty} \theta_n$; alors, $X_n + W_n \xrightarrow{\text{p.s.}} \theta$.

Remarques et exemple :

1) L'hypothèse $\sum_1^\infty \|\theta_{n+1} - \theta_n\| < \infty$ implique la convergence de la suite (θ_n) .

Certains modèles d'évolution du paramètre θ_n en fonction du temps n ont été étudiés par plusieurs auteurs, dans le cas d'observations indépendantes ; les processus d'approximation stochastique correspondants sont dits dynamiques. On peut trouver une bibliographie sur ce sujet dans [15]. On peut constater que l'étude du modèle linéaire généralisé faite dans le paragraphe 2.3 rentre dans ce cadre (voir la remarque dans 2.3).

2) L'hypothèse (liii) est une extension au cas où $\text{card } I_{\lambda} > 1$ de l'hypothèse de stabilité figurant dans les théorèmes classiques d'approximation stochastique, où I_{λ} est réduit à un seul élément. On peut constater son intérêt dans l'exemple simple suivant, qui relève du cas linéaire.

Soit (Y_n) une suite de variables aléatoires réelles observables et le modèle de régression linéaire

$$Y_n = b_n' \theta + Z_n.$$

Pour tout n , b_n est un vecteur de \mathbb{R}^k connu au temps n , θ le vecteur inconnu des k paramètres réels de la fonction de régression linéaire, Z_n la variable aléatoire réelle résiduelle. On peut estimer θ séquentiellement en utilisant le processus d'approximation stochastique (X_n) dans \mathbb{R}^k ,

$$X_{n+1} = X_n - \frac{1}{n^\alpha} \frac{b_n}{\|b_n\|^2} (b_n' X_n - Y_n),$$

que l'on peut aussi écrire

$$X_{n+1} = X_n - \frac{1}{n^\alpha} \frac{b_n b_n'}{\|b_n\|^2} (X_n - \theta) + \frac{1}{n^\alpha} \frac{b_n}{\|b_n\|^2} Z_n,$$

ce qui rentre dans le cadre général précédent en posant

$$a_n = \frac{1}{n^\alpha}, \quad h_n(x) = \frac{b_n b_n'}{\|b_n\|^2} (x - \theta), \quad U_n = - \frac{b_n}{\|b_n\|^2} Z_n.$$

Pour $k > 1$, la matrice $b_n b_n'$ étant de rang 1, la suite de fonctions (h_n) ne peut vérifier l'hypothèse (liii) pour $\text{card } I_\ell = 1$; par contre, elle peut la vérifier pour $\text{card } I_\ell \geq k$; il suffit de supposer que

$$\lambda = \inf_{\ell} \lambda \left(\sum_{j \in I_\ell} \frac{b_j b_j'}{\|b_j\|^2} \right) > 0,$$

car on aura alors pour $\|x - \theta\| > \varepsilon$,

$$\sum_{j \in I_\ell} \langle x - \theta, \frac{b_j b_j'}{\|b_j\|^2} (x - \theta) \rangle \geq \lambda \|x - \theta\|^2 > \lambda \varepsilon^2.$$

Cette hypothèse concernant le cas linéaire a été introduite et commentée par Albert et Gardner [1]. On remarque que

$$\lambda \left(\sum_{j \in I_\ell} \frac{b_j b_j'}{\|b_j\|^2} \right) = \min_{x \in \mathbb{R}^k} \left(\sum_{j \in I_\ell} \cos^2(b_j, x) \right).$$

On donne l'exemple suivant de cas simple où elle est vérifiée ; d'autres figurent dans [1] : on considère un modèle d'analyse de la variance à un facteur à k modalités, $Y_n = b_n' \theta + Z_n$, avec

$$b_{(\ell-1)k+1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad b_{(\ell-1)k+2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad b_{\ell k} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

On prend $I_\ell = \{(\ell - 1)k + 1, (\ell - 1)k + 2, \dots, \ell k\}$. On a $\|b_n\| = 1$, pour tout ℓ ,

$$\sum_{j \in I_\ell} b_j b_j' = I, \quad \lambda_{\min} = 1.$$

On constate que la fonction $h_n(x) = \frac{b_n b_n'}{\|b_n\|^2} (x - \theta)$ vérifie les hypothèses (1i) et

(2) du théorème 1. Il suffit de supposer en outre

$$\frac{1}{2} < \alpha \leq 1, \sum_1^{\infty} \frac{1}{n^\alpha} \|W_n\| < \infty \text{ p.s.}$$

pour avoir $X_n + W_n \xrightarrow{\text{p.s.}} \theta$.

3) L'hypothèse (4) du théorème est vérifiée si

$$\sum_1^{\infty} \|\theta_n - \theta_{n+1}\| < \infty, \sum_1^{\infty} a_n \|W_n\| < \infty \text{ p.s.}, \sum_1^{\infty} a_n \|h_n(X_n + W_n) - h_n(X_n)\| < \infty \text{ p.s.}$$

Nous rappelons dans la suite deux ensembles (A) et (B) de conditions suffisantes pour que $W_n \xrightarrow{\text{p.s.}} 0$, énoncés par Ljung dans [10] ; on peut en déduire aussi que $\sum_1^{\infty} a_n \|W_n\| < \infty$ p.s. (voir [3]). Le premier ensemble correspond au cas $a_n = \frac{1}{n}$;

on a alors $W_{n+1} = \frac{1}{n} \sum_1^n U_i$, et différentes lois des grands nombres peuvent être utilisées. Des conditions suffisantes sur les fonctions h_n peuvent être données pour que l'on ait en outre $\sum_1^{\infty} a_n \|h_n(X_n + W_n) - h_n(X_n)\| < \infty$ p.s.

On a $W_n \xrightarrow{\text{p.s.}} 0$ si (A) ou (B) est vérifié :

(A) $a_n = \frac{1}{n}$; on suppose qu'il existe un réel positif K et deux réels p et q tels que :

$$0 \leq 2p < q < 1 ; \forall n, m \in \mathbb{N}^*, E[\langle U_n, U_m \rangle] \leq K \frac{n^p + m^p}{1 + |n - m|^q}.$$

(B) On suppose que la suite (a_n) est non croissante et vérifie $L = \overline{\lim} \left| \frac{1}{a_n} - \frac{1}{a_{n-1}} \right| < \infty$; on suppose que $U_n = \sum_{k=0}^n h_{nk} Z_k$, (h_{nk}) étant une famille de réels,

(Z_k) une suite de vecteurs aléatoires indépendants de moyenne nulle et de matrice des covariances identité, et qu'il existe un entier $p > 0$, deux réels $C > 0$ et $\lambda < 1$, une suite non décroissante de réels positifs (α_n) , tels que :

$$\forall k, E[\|Z_k\|^{2p}] < C; \forall n, k, |h_{nk}| < \alpha_n \lambda^{n-k}; \sum_1^{\infty} (a_n \alpha_n^2)^p < \infty.$$

On a $\sum_1^{\infty} a_n \|W_n\| < \infty$ si (A) est vérifié ou si (B) est vérifié avec $L \leq 1$,

$$\sum_1^{\infty} a_n^{3/2} \alpha_n < \infty.$$

Dans la démonstration du théorème, on utilise le lemme suivant, qui est une version déterministe du lemme de convergence des pseudo-sur-martingales de Robbins et Siegmund [17].

Lemme 1 :

Soit $(b_n), (\beta_n), (\gamma_n), (\lambda_n)$ quatre suites de réels non négatifs telles que :

$$(1) \text{ pour tout } n, b_{n+1} \leq (1 + \beta_n)b_n + \gamma_n - \lambda_n;$$

$$(2) \sum_1^{\infty} \beta_n < \infty, \sum_1^{\infty} \gamma_n < \infty.$$

Alors, la suite (b_n) converge et $\sum_1^{\infty} \lambda_n < \infty$.

Démonstration du théorème :

$$\text{On note } R_n = h_n(X_n + W_n) - h_n(X_n) - W_n + \frac{1}{a_n}(\theta_n - \theta_{n+1}).$$

On raisonne à ω fixé, appartenant à l'ensemble $\left\{ \sum_1^{\infty} a_n \|R_n\| < \infty \right\}$.

1) D'après les définitions de X_n, W_n, R_n , en notant $Y_n = X_n + W_n$, on a :

$$Y_{n+1} - \theta_{n+1} = Y_n - \theta_n - a_n h_n(Y_n) + a_n R_n \quad (2.1.1)$$

$$\|Y_{n+1} - \theta_{n+1}\|^2 \leq \|Y_n - \theta_n\|^2 + 2a_n^2 \|h_n(Y_n)\|^2 + 2a_n^2 \|R_n\|^2$$

$$- 2a_n \langle Y_n - \theta_n, h_n(Y_n) \rangle + 2a_n \langle Y_n - \theta_n, R_n \rangle.$$

Comme $|\langle Y_n - \theta_n, R_n \rangle| \leq \|R_n\| (1 + \|Y_n - \theta_n\|^2)$ et, sous (1i), $\|h_n(Y_n)\|^2 \leq 2d_n^2 \|Y_n - \theta_n\|^2 + 2e_n^2$, on a :

$$\begin{aligned} \|Y_{n+1} - \theta_{n+1}\|^2 &\leq \|Y_n - \theta_n\|^2 (1 + 2a_n \|R_n\| + 4a_n^2 d_n^2) + 2a_n \|R_n\| \\ &\quad + 4a_n^2 e_n^2 + 2a_n^2 \|R_n\|^2 - 2a_n \langle Y_n - \theta_n, h_n(Y_n) \rangle. \end{aligned}$$

Sous (1ii), (3), (4), on peut appliquer le lemme 1 :

$$\exists T \geq 0 : \|Y_n - \theta_n\| \rightarrow T ; \sum_1^\infty a_n \langle Y_n - \theta_n, h_n(Y_n) \rangle < \infty. \quad (2.1.2)$$

2) A partir de (2.1.1), on obtient :

$$\begin{aligned} \|Y_{n+1} - Y_n\| &\leq a_n \|h_n(Y_n)\| + a_n \|R_n\| + \|\theta_{n+1} - \theta_n\| \\ &\leq a_n d_n \|Y_n - \theta_n\| + a_n e_n + a_n \|R_n\| + \|\theta_{n+1} - \theta_n\|. \end{aligned}$$

Sous (3), (4), comme $\|Y_n - \theta_n\| \rightarrow T$, on a $\|Y_{n+1} - Y_n\| \rightarrow 0$.

3) On suppose $T \neq 0$. Alors, il existe un réel ε_1 compris entre 0 et 1, un entier $N(\varepsilon_1)$ tels que, pour $n > N(\varepsilon_1)$, $\varepsilon_1 < \|Y_n - \theta_n\| < \frac{1}{\varepsilon_1}$; sous (1iii), ceci implique, à partir d'un rang L :

$$\sum_{j \in I_L} \langle Y_{n_j} - \theta_j, h_j(Y_{n_j}) \rangle \geq b(\varepsilon_1).$$

On en déduit qu'il existe un entier m_L appartenant à I_L tel que :

$$\langle Y_{n_L} - \theta_{m_L}, h_{m_L}(Y_{n_L}) \rangle \geq \frac{b(\varepsilon_1)}{T}.$$

On considère la décomposition :

$$\begin{aligned} \langle Y_{n_L} - \theta_{m_L}, h_{m_L}(Y_{m_L}) \rangle &= \langle Y_{n_L} - \theta_{m_L}, h_{m_L}(Y_{n_L}) \rangle + \langle Y_{n_L} - \theta_{m_L}, h_{m_L}(Y_{m_L}) - h_{m_L}(Y_{n_L}) \rangle \\ &= A + B. \end{aligned}$$

Soit $\varepsilon > 0$; sous (2), à ε correspond un η ; soit $\varepsilon' < \eta$. Comme $Y_{n+1} - Y_n \rightarrow 0$ et $m_\ell - n_\ell < r$, on a, à partir d'un certain rang, $\|Y_{m_\ell} - Y_{n_\ell}\| < \varepsilon'$; donc, $\|h_{m_\ell}(Y_{m_\ell}) - h_{m_\ell}(Y_{n_\ell})\| < \varepsilon$ et

$$|B| \leq (\|Y_{n_\ell} - Y_{m_\ell}\| + \|Y_{m_\ell} - \theta_{m_\ell}\|)\varepsilon < (\varepsilon' + \frac{1}{\varepsilon_1})\varepsilon.$$

Comme $A \geq \frac{b(\varepsilon_1)}{r}$, on en déduit que

$$\langle Y_{n_\ell} - \theta_{m_\ell}, h_{m_\ell}(Y_{m_\ell}) \rangle > \frac{b(\varepsilon_1)}{r} - (\varepsilon' + \frac{1}{\varepsilon_1})\varepsilon > \frac{b(\varepsilon_1)}{2r}$$

pour ε et ε' suffisamment petits.

Sous (3), on a alors :

$$\sum_{\ell} a_{m_\ell} \langle Y_{n_\ell} - \theta_{m_\ell}, h_{m_\ell}(Y_{m_\ell}) \rangle = +\infty. \quad (2.1.3)$$

On décompose cette série en

$$\sum_{\ell} a_{m_\ell} \langle Y_{m_\ell} - \theta_{m_\ell}, h_{m_\ell}(Y_{m_\ell}) \rangle - \sum_{\ell} a_{m_\ell} \langle Y_{m_\ell} - Y_{n_\ell}, h_{m_\ell}(Y_{m_\ell}) \rangle. \quad (2.1.4)$$

A partir de (2.1.1), on obtient :

$$\begin{aligned} Y_{m_\ell} - Y_{n_\ell} &= \sum_{j=n_\ell}^{m_\ell-1} (-a_j h_j(Y_j) + a_j R_j + \theta_{j+1} - \theta_j) \\ \sum_{\ell} a_{m_\ell} |\langle Y_{m_\ell} - Y_{n_\ell}, h_{m_\ell}(Y_{m_\ell}) \rangle| &\leq \sum_{\ell} a_{m_\ell} \sum_{j=n_\ell}^{m_\ell-1} a_j \|h_j(Y_j)\| \|h_{m_\ell}(Y_{m_\ell})\| \\ &\quad + \sum_{\ell} a_{m_\ell} \sum_{j=n_\ell}^{m_\ell-1} (a_j \|R_j\| + \|\theta_{j+1} - \theta_j\|) \|h_{m_\ell}(Y_{m_\ell})\| \\ &= C + D. \end{aligned}$$

Sous (1i) et (3), on a :

$$\begin{aligned}
 C &\leq \sum_{\ell} \sum_{j \in I_{\ell}} a_{m_{\ell}} a_j (d_j \| Y_j - \theta_j \| + e_j) (d_{m_{\ell}} \| Y_{m_{\ell}} - \theta_{m_{\ell}} \| + e_{m_{\ell}}) \\
 &\leq \sum_{\ell} \sum_{j \in I_{\ell}} a_{m_{\ell}} a_j \left(\frac{1}{\varepsilon_1} d_j + e_j \right) \left(\frac{1}{\varepsilon_1} d_{m_{\ell}} + e_{m_{\ell}} \right) < \infty.
 \end{aligned}$$

En outre, comme sous (1i), (3) et (2.1.2), $a_{m_{\ell}} \| h_{m_{\ell}}(Y_{m_{\ell}}) \| \rightarrow 0$, on a, sous (4), $D < \infty$.

De (2.1.3) et (2.1.4), on déduit alors que :

$$\sum_{\ell} a_{m_{\ell}} \langle Y_{m_{\ell}} - \theta_{m_{\ell}}, h_{m_{\ell}}(Y_{m_{\ell}}) \rangle = +\infty.$$

Sous (1ii) et (3), ceci implique :

$$\sum_n a_n \langle Y_n - \theta_n, h_n(Y_n) \rangle = +\infty.$$

Ceci est en contradiction avec (2.1.2). Donc, $T = 0$.

2.2 - Comportement asymptotique du moment d'ordre deux

Théorème 2 :

Soit, pour tout n , $a_n = \frac{1}{n^{\alpha}}$ avec $0 < \alpha \leq 1$; soit un réel $\beta > 1 - \alpha$. On suppose

les hypothèses suivantes vérifiées :

(1) Il existe une suite convergente de réels (θ_n) , deux réels positifs C et λ , un entier positif r tels que :

(i) pour tout n , $h_n(\theta_n) = 0$;

(ii) pour tout n et tout x_1, x_2 , $\| h_n(x_1) - h_n(x_2) \| \leq C(n^{\frac{\alpha-\beta}{2}} \| x_1 - x_2 \| + n^{-\beta})$;

(iii) pour tout n et tout x , $\langle x - \theta_n, h_n(x) \rangle \geq 0$;

$$(iv) \text{ pour tout } \ell \text{ et tout } x, \sum_{j \in I_\ell} \langle x - \theta_j, h_j(x) \rangle \geq \lambda \|x - \theta_{n_\ell}\|^2,$$

où $n_\ell = (\ell - 1)r + 1$, I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

(2) Pour $\alpha = 1$, $\beta r < 2\lambda$.

$$(3) \sup_n n^{\beta + \max(\alpha, \beta)} E [\|W_n\|^2] < \infty.$$

$$(4) \sup_n n^{\alpha + \beta} \|\theta_n - \theta_{n+1}\| < \infty.$$

Alors, $\overline{\lim} n^\beta E [\|X_n - \theta_n\|^2] < \infty$.

Ce théorème est en fait un corollaire, de conclusion facilement interprétable, de la proposition suivante que nous allons démontrer.

Dans l'énoncé suivant, on note

$$\mu_\ell = \min_{j \in I_\ell} a_j, S_n = h_n(X_n + W_n) - h_n(X_n) - W_n,$$

$$g_\ell = \sum_{j \in I_\ell} (a_j^2 d_j^2 + a_j e_j + a_j E [\|S_j\|^2]^{\frac{1}{2}} + \|\theta_{j+1} - \theta_j\|).$$

Proposition 1 :

On suppose les hypothèses suivantes vérifiées :

(1) Il existe une suite convergente de réels (θ_n) , deux suites de réels positifs (d_n) et (e_n) , un réel positif λ , un entier r , une suite croissante d'entiers (n_ℓ) vérifiant $n_{\ell+1} \leq n_\ell + r$, tels que :

(i) pour tout n , $h_n(\theta_n) = 0$;

(ii) pour tout n et tout x_1, x_2 , $\|h_n(x_1) - h_n(x_2)\| \leq d_n \|x_1 - x_2\| + e_n$;

(iii) pour tout n et tout x , $\langle x - \theta_n, h_n(x) \rangle \geq 0$;

$$(iv) \text{ pour tout } \ell \text{ et tout } x, \sum_{j \in I_\ell} \langle x - \theta_j, h_j(x) \rangle \geq \lambda \|x - \theta_{n_\ell}\|^2,$$

où I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

$$(2) \text{ (i) } a_n > 0; \sum_1^\infty a_n^2 d_n^2 < \infty; \sum_1^\infty a_n^2 e_n^2 < \infty; \mu_\ell = o(1); \sum_1^\infty \mu_\ell = +\infty.$$

(ii) Il existe une suite (γ_ℓ) dans \mathbb{R}^+ vérifiant $\sum_1^\infty \gamma_\ell < \infty$ et un réel

$\gamma < 2\lambda$ tels que, pour tout ℓ :

$$\frac{\mu_{\ell+1}}{g_{\ell+1}} \leq \frac{\mu_\ell}{g_\ell} (1 + \gamma \mu_\ell + \gamma_\ell + o(\mu_\ell)).$$

$$(3) \sum_1^\infty \|\theta_{n+1} - \theta_n\| < \infty; \sum_1^\infty a_n E[\|h_n(X_n + W_n) - h_n(X_n) - W_n\|^2]^{\frac{1}{2}} < \infty.$$

Soit, pour tout n , $\ell(n)$ tel que $n \in I_{\ell(n)}$.

$$\text{Alors, } \overline{\lim}_{g_{\ell(n)}} \frac{\mu_{\ell(n)}}{g_{\ell(n)}} E[\|X_n + W_n - \theta_n\|^2] < \infty.$$

Dans la démonstration de la proposition 1, on utilise le lemme suivant [6], dont la démonstration est rappelée dans l'appendice.

Lemme 2 :

Soit (a_n) et (g_n) deux suites de réels positifs, (b_n) et (γ_n) deux suites de réels non négatifs, λ, μ deux réels positifs, γ un réel inférieur à λ , tels que :

$$(1) \text{ pour tout } n, b_{n+1} \leq (1 - \lambda a_n) b_n + \mu g_n;$$

$$(2) a_n = o(1), \sum_1^\infty a_n = \infty;$$

$$(3) \text{ pour tout } n, \frac{a_{n+1}}{g_{n+1}} \leq \frac{a_n}{g_n} (1 + \gamma a_n + \gamma_n + o(a_n));$$

$$(4) \sum_1^\infty \gamma_n < \infty.$$

$$\text{Alors, } \overline{\lim}_{g_n} \frac{a_n}{g_n} b_n \leq \frac{\mu}{\lambda - \gamma}.$$

Démonstration de la proposition 1 :

On note $Y_n = X_n + W_n$. D'après (2.1.1) dans la démonstration du théorème précédent, on a, avec $R_n = h_n(X_n + W_n) - h_n(X_n) - W_n + \frac{1}{a_n}(\theta_n - \theta_{n+1})$:

$$\begin{aligned} E[\|Y_{n+1} - \theta_{n+1}\|^2] &\leq E[\|Y_n - \theta_n\|^2] + 2a_n^2 E[\|h_n(Y_n)\|^2] + 2a_n^2 E[\|R_n\|^2] \\ &\quad - 2a_n E[\langle Y_n - \theta_n, h_n(Y_n) \rangle] + 2a_n E[\langle Y_n - \theta_n, R_n \rangle]. \end{aligned}$$

Comme $|E[\langle Y_n - \theta_n, R_n \rangle]| \leq (1 + E[\|Y_n - \theta_n\|^2]) E[\|R_n\|^2]^{\frac{1}{2}}$ et, sous (1), $\|h_n(Y_n)\|^2 \leq 2d_n^2 \|Y_n - \theta_n\|^2 + 2e_n^2$, on a :

$$\begin{aligned} E[\|Y_{n+1} - \theta_{n+1}\|^2] &\leq E[\|Y_n - \theta_n\|^2](1 + 4a_n^2 d_n^2 + 2a_n E[\|R_n\|^2]^{\frac{1}{2}}) \\ &\quad + 2a_n E[\|R_n\|^2]^{\frac{1}{2}} + 2a_n^2 E[\|R_n\|^2] + 4a_n^2 e_n^2 - 2a_n E[\langle Y_n - \theta_n, h_n(Y_n) \rangle]. \end{aligned}$$

Sous (1iii), (2i), (3), on peut appliquer le lemme 1 :

$$\exists t \geq 0 : E[\|Y_n - \theta_n\|^2] \rightarrow t. \quad (2.2.1)$$

On en déduit qu'il existe un réel positif M tel que :

$$E[\|Y_{n+1} - \theta_{n+1}\|^2] \leq E[\|Y_n - \theta_n\|^2] + b_n - 2a_n E[\langle Y_n - \theta_n, h_n(Y_n) \rangle] \quad (2.2.2)$$

avec

$$b_n = M(a_n^2(d_n^2 + e_n^2) + a_n E[\|R_n\|^2]^{\frac{1}{2}}). \quad (2.2.3)$$

Par conséquent :

$$E[\|Y_{n_{\ell+1}} - \theta_{n_{\ell+1}}\|^2] \leq E[\|Y_{n_{\ell}} - \theta_{n_{\ell}}\|^2] + \sum_{j \in I_{\ell}} b_j - 2\mu_{\ell} \sum_{j \in I_{\ell}} E[\langle Y_j - \theta_j, h_j(Y_j) \rangle].$$

On considère la décomposition :

$$\begin{aligned} \sum_{j \in I_\ell} E [\langle Y_j - \theta_j, h_j(Y_j) \rangle] &= \sum_{j \in I_\ell} E [\langle Y_{n_\ell} - \theta_j, h_j(Y_{n_\ell}) \rangle] \\ + \sum_{j \in I_\ell} E [\langle Y_j - Y_{n_\ell}, h_j(Y_{n_\ell}) \rangle] &+ \sum_{j \in I_\ell} E [\langle Y_j - \theta_j, h_j(Y_j) - h_j(Y_{n_\ell}) \rangle] \\ &= A + B + C. \end{aligned}$$

Sous (1iv), $A \geq \lambda E [\| Y_{n_\ell} - \theta_{n_\ell} \|^2]$; donc :

$$E [\| Y_{n_{\ell+1}} - \theta_{n_{\ell+1}} \|^2] \leq (1 - 2\lambda\mu_\ell) E [\| Y_{n_\ell} - \theta_{n_\ell} \|^2] + \sum_{j \in I_\ell} b_j - 2\mu_\ell B - 2\mu_\ell C.$$

On a : $|B| \leq \sum_{j \in I_\ell} E [\| Y_j - Y_{n_\ell} \|^2]^{\frac{1}{2}} E [\| h_j(Y_{n_\ell}) \|^2]^{\frac{1}{2}}$; sous (1), (3) et (2.2.1), il

existe un réel positif M_1 tel que :

$$E [\| h_j(Y_{n_\ell}) \|^2]^{\frac{1}{2}} \leq (2d_j^2 E [\| Y_{n_\ell} - \theta_j \|^2] + 2e_j^2)^{\frac{1}{2}} \leq M_1(d_j + e_j);$$

en outre, il existe un réel positif M_2 tel que :

$$\begin{aligned} |C| &\leq \sum_{j \in I_\ell} E [\| Y_j - \theta_j \|^2]^{\frac{1}{2}} E [\| h_j(Y_j) - h_j(Y_{n_\ell}) \|^2]^{\frac{1}{2}} \\ &\leq M_2 \sum_{j \in I_\ell} (2d_j^2 E [\| Y_j - Y_{n_\ell} \|^2] + 2e_j^2)^{\frac{1}{2}} \\ &\leq M_2 \sqrt{2} \sum_{j \in I_\ell} (d_j E [\| Y_j - Y_{n_\ell} \|^2]^{\frac{1}{2}} + e_j). \end{aligned}$$

Donc, il existe un réel positif M_3 tel que :

$$\begin{aligned} E [\| Y_{n_{\ell+1}} - \theta_{n_{\ell+1}} \|^2] &\leq (1 - 2\lambda\mu_\ell) E [\| Y_{n_\ell} - \theta_{n_\ell} \|^2] + \sum_{j \in I_\ell} b_j \\ &\quad + M_3\mu_\ell \sum_{j \in I_\ell} ((d_j + e_j) E [\| Y_j - Y_{n_\ell} \|^2]^{\frac{1}{2}} + e_j). \end{aligned}$$

Or, on a, avec $S_n = h_n(Y_n) - h_n(X_n) - W_n$:

$$Y_j - Y_{n_\ell} = \sum_{k=n_\ell}^{j-1} (-a_k h_k(Y_k) + a_k S_k)$$

$$\|Y_j - Y_{n_\ell}\| \leq \sum_{k=n_\ell}^{j-1} (a_k d_k \|Y_k - \theta_k\| + a_k e_k + a_k \|S_k\|).$$

Sous (2.2.1), comme $\text{card } I_\ell$ est borné, il existe un réel positif M_4 tel que :

$$E [\|Y_j - Y_{n_\ell}\|^2] \leq M_4 \sum_{k \in I_\ell} (a_k d_k + a_k e_k + a_k E [\|S_k\|^2]).$$

Comme, pour $j \in I_\ell$, $\mu_\ell(d_j + e_j) \leq a_j(d_j + e_j) \rightarrow 0$, il existe un réel positif M_5 tel que :

$$D = \mu_\ell \sum_{j \in I_\ell} ((d_j + e_j) E [\|Y_j - Y_{n_\ell}\|^2] + e_j)$$

$$\leq M_5 \sum_{j \in I_\ell} (a_j^2 d_j^2 + a_j E [\|S_j\|^2] + a_j e_j)$$

$$\leq M_5 g_\ell.$$

D'après la définition de b_j (2.2.3), il existe un réel positif M_6 tel que :

$$\sum_{j \in I_\ell} b_j + M_3 D \leq M_6 g_\ell.$$

Par conséquent :

$$E [\|Y_{n_{\ell+1}} - \theta_{n_{\ell+1}}\|^2] \leq (1 - 2\lambda\mu_\ell) E [\|Y_{n_\ell} - \theta_{n_\ell}\|^2] + M_6 g_\ell.$$

Sous (2), l'application du lemme 2 donne $\overline{\lim} \frac{\mu_\ell}{g_\ell} E [\|Y_{n_\ell} - \theta_{n_\ell}\|^2] < \infty$. Sous (2.2.2), il existe un réel positif M_7 tel que, pour $n \in I_\ell$:

$$E [\| Y_{n+1} - \theta_{n+1} \|^2] \leq E [\| Y_n - \theta_n \|^2] + M_7 g_\lambda.$$

Comme $\mu_\lambda = o(1)$, on en déduit aisément, en notant, pour tout n , $\lambda(n)$ l'entier tel que $n \in I_{\lambda(n)}$, que :

$$\overline{\lim} \frac{\mu_{\lambda(n)}}{g_{\lambda(n)}} E [\| Y_n - \theta_n \|^2] < \infty.$$

Dans la démonstration du théorème 2, on utilise le lemme suivant, dont la démonstration est immédiate.

Lemme 3 :

Pour $\mu_\lambda = \frac{a}{\lambda^\alpha}$ et $g_\lambda = \frac{b}{\lambda^{\alpha+\beta}}$, avec $a > 0, b > 0, 0 < \alpha \leq 1$, on a, pour $\alpha = 1, \frac{\mu_{\lambda+1}}{g_{\lambda+1}} = \frac{\mu_\lambda}{g_\lambda} (1 + \frac{\beta}{a} \mu_\lambda + o(\mu_\lambda))$ et, pour $0 < \alpha < 1$, pour tout $\gamma > 0$, à partir du rang $L = \min \{ \lambda : \lambda^{1-\alpha} \geq \frac{\beta}{\gamma a}, \frac{\mu_{\lambda+1}}{g_{\lambda+1}} \leq \frac{\mu_\lambda}{g_\lambda} (1 + \gamma \mu_\lambda + o(\mu_\lambda)) \}$.

Démonstration du théorème 2 :

La vérification des hypothèses (1) et (2i) de la proposition 1 est immédiate, avec $\mu_\lambda = \frac{1}{\lambda^\alpha}$. En outre, sous (1ii) :

$$\| S_n \| \leq (Cn^{\frac{\alpha-\beta}{2}} + 1) \| W_n \| + Cn^{-\beta}$$

$$E [\| S_n \|^2]^{\frac{1}{2}} \leq \sqrt{2} (Cn^{\frac{\alpha-\beta}{2}} + 1) E [\| W_n \|^2]^{\frac{1}{2}} + \sqrt{2} Cn^{-\beta}.$$

Sous (3), (4), il existe $D > 0, F > 0$ tels que $\sum_1^\infty a_n E [\| S_n \|^2]^{\frac{1}{2}} < D \sum_1^\infty \frac{1}{n^{\alpha+\beta}} < \infty$ et $g_\lambda < \frac{F}{\lambda^{\alpha+\beta}}$.

Dans la proposition 1, on peut remplacer la suite (g_λ) par toute suite majorante, ici $(\frac{F}{\lambda^{\alpha+\beta}})$; l'application du lemme 3 montre alors que l'hypothèse (2ii) de la proposition est vérifiée lorsque, dans le cas $\alpha = 1, \beta r < 2\lambda$.

On déduit de la proposition que $\overline{\lim} n^\beta E [\| X_n + W_n - \theta_n \|^2] < \infty$. Comme $n^\beta E [\| W_n \|^2] \rightarrow 0$ sous (3), on a $\overline{\lim} n^\beta E [\| X_n - \theta_n \|^2] = 0$.

2.3 - Application à un modèle linéaire généralisé

Soit, pour tout n , Y_n^* une variable aléatoire réelle observable, que l'on décompose selon le modèle de régression linéaire généralisée

$$Y_n^* = f(b_n' \theta) + Z_n,$$

où θ est le vecteur de \mathbb{R}^k des paramètres inconnus, b_n un vecteur de \mathbb{R}^k connu au temps n , f une fonction connue de \mathbb{R} dans \mathbb{R} , Z_n la variable aléatoire réelle résiduelle. Pour estimer θ , on peut utiliser le processus (X_n) tel que

$$X_{n+1} = X_n - a_n b_n (f(b_n' X_n) - Y_n^*),$$

ce que l'on peut aussi écrire

$$X_{n+1} = X_n - a_n b_n (f(b_n' X_n) - f(b_n' \theta) - Z_n).$$

Si l'on note $D_n = b_n' X_n = \langle b_n, X_n \rangle$ et M_n la fonction de \mathbb{R} dans \mathbb{R} définie par $M_n(d) = f(d) - f(\delta_n)$, de zéro $\delta_n = b_n' \theta = \langle b_n, \theta \rangle$, $\langle \cdot, \cdot \rangle$ étant le produit scalaire euclidien usuel dans \mathbb{R}^k , on a :

$$X_{n+1} = X_n - a_n b_n (M_n(D_n) - Z_n).$$

Nous allons étendre la définition de ce processus au cas où M_n est une application de \mathbb{R}^p dans \mathbb{R}^p et retrouver ainsi un processus d'approximation stochastique dynamique [15].

On note $M_n^1, M_n^2, \dots, M_n^p$ les composantes de M_n . Soit p entiers positifs k_1, k_2, \dots, k_p . On suppose que M_n admet un zéro unique δ_n , de composantes $\delta_n^1, \delta_n^2, \dots, \delta_n^p$, telles que, pour $i = 1, 2, \dots, p$, δ_n^i soit le produit scalaire euclidien usuel dans \mathbb{R}^{k_i} de deux vecteurs b_n^i et θ^i : $\delta_n^i = \langle b_n^i, \theta^i \rangle$ (on omet l'indice i du produit

scalaire pour simplifier la notation). On définit, pour $i = 1, 2, \dots, p$, les processus (X_n^i) et (W_n^i) dans \mathbb{R}^k tels que

$$X_{n+1}^i = X_n^i - a_n b_n^i (M_n^i(D_n) + Z_n^i),$$

$$W_{n+1}^i = (1 - a_n)W_n^i + a_n b_n^i Z_n^i, \quad W_1^i = 0,$$

où (Z_n^i) est un processus stochastique réel, (a_n) une suite de réels positifs, D_n a pour composantes $D_n^1, D_n^2, \dots, D_n^p$, avec $D_n^i = \langle b_n^i, X_n^i \rangle$. On note $k = k_1 + k_2 + \dots + k_p$.

Dans \mathbb{R}^k , on a :

$$\begin{pmatrix} X_{n+1}^1 \\ \vdots \\ X_{n+1}^p \end{pmatrix} = \begin{pmatrix} X_n^1 \\ \vdots \\ X_n^p \end{pmatrix} - a_n \begin{pmatrix} b_n^1 M_n^1(D_n) \\ \vdots \\ b_n^p M_n^p(D_n) \end{pmatrix} - a_n \begin{pmatrix} b_n^1 Z_n^1 \\ \vdots \\ b_n^p Z_n^p \end{pmatrix}$$

$$X_{n+1} = X_n - a_n h_n(X_n) - a_n U_n$$

$$W_{n+1} = (1 - a_n)W_n + a_n U_n.$$

On munit \mathbb{R}^k et \mathbb{R}^p du produit scalaire euclidien usuel. On note $R_n = h_n(X_n + W_n) - h_n(X_n) - W_n$, $\theta = (\theta^1, \theta^2, \dots, \theta^p)'$.

Remarque :

On peut interpréter cette étude comme celle d'un modèle d'évolution dans le temps du paramètre δ_n , dont les composantes $\delta_n^i = \langle b_n^i, \theta^i \rangle$ sont des combinaisons linéaires de fonctions connues du temps.

Corollaire 1 :

On suppose les hypothèses suivantes vérifiées :

(1) (i) Pour tout n , il existe deux réels positifs p_n et q_n tels que, pour tout $d \in \mathbb{R}^p$,

$$\|M_n(d)\| \leq p_n \|d - \delta_n\| + q_n.$$

(ii) Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$(\|d_1 - d_2\| < \eta) \Rightarrow (\sup_n \|M_n(d_1) - M_n(d_2)\| < \varepsilon).$$

(iii) Pour tout $0 < \varepsilon < 1$,

$$c(\varepsilon) = \inf_n \left(\inf_{\{\varepsilon < \|d - \delta_n\| < \frac{1}{\varepsilon}\}} \langle d - \delta_n, M_n(d) \rangle > 0.$$

$$(2) (i) \quad b = \sup_{i, n} \|b_n^i\| < \infty.$$

(ii) Il existe un entier r , une suite croissante d'entiers (n_ℓ) vérifiant

$n_{\ell+1} \leq n_\ell + r$, tels que

$$\delta = \inf_{i, \ell} \lambda \left(\sum_{j \in I_\ell} b_j^i b_j^{i'} \right) > 0$$

où I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

$$(3) \text{ Pour tout } n, a_n > 0; \sum_1^\infty \min_{j \in I_\ell} a_j = +\infty; \sum_1^\infty a_n^2 p_n^2 < \infty; \sum_1^\infty a_n^2 q_n^2 < \infty.$$

$$(4) \sum_1^\infty a_n \|R_n\| < \infty \text{ p.s.}$$

Alors, $X_n + W_n \xrightarrow{\text{p.s.}} \theta$.

Corollaire 2 :

Soit, pour tout n , $a_n = \frac{1}{n^\alpha}$, avec $0 < \alpha \leq 1$; soit un réel $\beta > 1 - \alpha$.

Outre l'hypothèse (2) du corollaire 1, avec $n_\ell = (\ell - 1)r + 1$, on suppose les hypothèses suivantes vérifiées :

(1) (i) Il existe un réel positif C tel que, pour tout n et tout $d_1, d_2 \in \mathbb{R}^P$,

$$\|M_n(d_1) - M_n(d_2)\| \leq C(n^{\frac{\alpha-\beta}{2}} \|d_1 - d_2\| + n^{-\beta}).$$

(ii) Il existe un réel positif λ tel que, pour tout n et tout $d \in \mathbb{R}^P$,

$$\langle d - \delta_n, M_n(d) \rangle \geq \lambda \|d - \delta_n\|^2.$$

(2) Pour $\alpha = 1, \beta r < 2\lambda$.

(3) $\sup_n n^{\beta + \max(\alpha, \beta)} E[\|W_n\|^2] < \infty$.

Alors, $\overline{\lim} n^\beta E[\|X_n - \theta\|^2] < \infty$.

Remarque :

Dans le cas du modèle linéaire, on a le processus

$$\begin{aligned} X_{n+1} &= X_n - a_n \frac{b_n}{\|b_n\|^2} (b_n' X_n - Y_n^*) \\ &= X_n - a_n b_n^u (b_n^u X_n - b_n^u \theta - \frac{Z_n}{\|b_n\|}) \end{aligned}$$

avec $b_n^u = \frac{b_n}{\|b_n\|}$. On peut appliquer les corollaires 1 et 2 en remplaçant b_n par b_n^u , Z_n par $\frac{Z_n}{\|b_n\|}$ et en prenant $M_n(d) = d - \delta_n$.

Démonstration des corollaires :

Nous donnons uniquement la vérification des hypothèses (1ii) et (1iii) du théorème 1; le reste est immédiat. On a :

$$\langle x - \theta, h_n(x) \rangle = \sum_{i=1}^p \langle x^i - \theta^i, b_n^i \rangle M_n^i(d_n) = \langle d_n - \delta_n, M_n(d_n) \rangle \geq 0$$

avec $d_n^i = \langle b_n^i, x^i \rangle$, $d_n = (d_n^1, d_n^2, \dots, d_n^p)$.

Pour $\epsilon < \|x - \theta\| < \frac{1}{\epsilon}$:

$$\|d_j - \delta_j\|^2 \leq b^2 \|x - \theta\|^2 < \frac{b^2}{\epsilon^2},$$

$$\begin{aligned} \sum_{j \in I_\delta} \|d_j - \delta_j\|^2 &= \sum_{j \in I_\delta} \sum_{i=1}^p \langle b_j^i, x^i - \theta^i \rangle^2 = \sum_{i=1}^p (x^i - \theta^i)' \sum_{j \in I_\delta} b_j^i b_j^i (x^i - \theta^i) \\ &\geq \delta \sum_{i=1}^p \|x^i - \theta^i\|^2 = \delta \|x - \theta\|^2 > \delta \epsilon^2; \end{aligned}$$

donc, il existe $j \in I_\delta$ tel que $\|d_j - \delta_j\| > \frac{\sqrt{\delta}}{\sqrt{r}} \varepsilon$.

Soit ε_1 tel que $\varepsilon_1 < \frac{\sqrt{\delta}}{\sqrt{r}} \varepsilon$, $\frac{1}{\varepsilon_1} > \frac{b}{\varepsilon}$; ($\varepsilon < \|x - \theta\| < \frac{1}{\varepsilon}$) implique ($\varepsilon_1 < \|d_j - \delta_j\| < \frac{1}{\varepsilon_1}$); d'après l'hypothèse (1iii) du corollaire 1, on a :

$$\sum_{j \in I_\delta} \langle x - \theta, h_j(x) \rangle = \sum_{j \in I_\delta} \langle d_j - \delta_j, M_j(d_j) \rangle \geq c(\varepsilon_1) > 0.$$

3 - CAS LINEAIRE

Dans ce paragraphe, $h_n(X_n)$ prend la forme $G_n(X_n - \theta_n)$. En utilisant des démonstrations spécifiques à ce cas, nous obtenons des résultats de convergence avec des hypothèses plus faibles.

On définit les processus (X_n) et (W_n) dans \mathbb{R}^k :

$$X_{n+1} = X_n - a_n G_n(X_n - \theta_n) - a_n U_n,$$

$$W_{n+1} = (1 - a_n)W_n + a_n U_n, \quad W_1 = 0,$$

où G_n est une variable aléatoire à valeurs matricielles $k \times k$, U_n une variable aléatoire dans \mathbb{R}^k , $\theta_n \in \mathbb{R}^k$, $a_n \in \mathbb{R}^+$.

3.1 - Convergence presque sûre dans le cas linéaire

Théorème 3 :

On suppose que, pour ω appartenant à un ensemble de probabilité 1, les hypothèses suivantes sont vérifiées :

(1) Il existe un réel λ positif, un entier r , une suite croissante d'entiers (n_ℓ) vérifiant $n_{\ell+1} \leq n_\ell + r$, une suite (μ_ℓ) dans \mathbb{R}^+ tels que :

(i) $\mu_\ell = o(1)$; $\sum_1^\infty \mu_\ell = +\infty$;

(ii) pour tout ℓ , $\lambda \left(\sum_{i \in I_\ell} a_i (G_i + G'_i) \right) \geq \lambda \mu_\ell$,

où I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

(2) $\sum_{i \in I_\ell} a_i^2 \|G_i\|^2 = o(1) \mu_\ell$.

(3) $\sum_{i \in I_\ell} \|\theta_{i+1} - \theta_i\| = o(1) \mu_\ell$; $\sum_{i \in I_\ell} a_i \|(G_i - I) W_i\| = o(1) \mu_\ell$.

Alors, $X_n + W_n - \theta_n \xrightarrow{p.s.} 0$.

Remarques :

1) Considérons le processus d'estimation du paramètre d'un modèle de régression linéaire étudié dans le paragraphe 2.1 :

$$\begin{aligned} X_{n+1} &= X_n - \frac{1}{n\alpha} \frac{b_n}{\|b_n\|^2} (b'_n X_n - Y_n) \\ &= X_n - \frac{1}{n\alpha} \frac{b_n b'_n}{\|b_n\|^2} (X_n - \theta) + \frac{1}{n\alpha} \frac{b_n}{\|b_n\|^2} Z_n. \end{aligned}$$

On pose $a_n = \frac{1}{n\alpha}$, $G_n = \frac{b_n b'_n}{\|b_n\|^2}$, $\mu_\ell = \min_{j \in I_\ell} a_j$. On a $X_n - \theta \xrightarrow{p.s.} 0$ sous les hypothèses :

$$\inf_{\ell} \lambda \left(\sum_{i \in I_\ell} \frac{b_i b'_i}{\|b_i\|^2} \right) > 0; \frac{1}{2} < \alpha \leq 1; W_n \xrightarrow{p.s.} 0.$$

Si l'on compare au résultat déjà énoncé, on constate que la condition $\sum_1^\infty \frac{1}{n\alpha} \|W_n\| < \infty$ p.s. n'est plus demandée.

2) Dans le cas $\theta_n = \theta$ ne dépendant pas de n et $a_n = \frac{1}{n}$ (on a alors $W_{n+1} = \frac{1}{n} \sum_1^n U_i$), il existe un résultat classique de convergence presque sûre de X_n

vers θ , sous l'hypothèse d'ergodicité du processus (G_n, U_n) ; on peut appliquer ce résultat en particulier au cas de la régression linéaire avec b_n aléatoire [11].

3) On constate que la suite (θ_n) n'est pas nécessairement convergente, il suffit qu'elle vérifie $\sum_{i \in I_\ell} \|\theta_i - \theta_{i+1}\| = o(1)\mu_\ell$; par exemple, pour $a_n = \frac{1}{n^\alpha}$, $\alpha < 1$, $\mu_\ell = \min_{j \in I_\ell} a_j$, on peut avoir dans le cas réel $\theta_{n+1} = \theta_n + \frac{1}{n}$.

Dans la démonstration du théorème, on prend pour norme matricielle la norme spectrale. On utilise les deux lemmes suivants :

Lemme 4 :

Sous les hypothèses (1), (2), on a $\|\prod_{i \in I_\ell} (I - a_i G_i)\|^2 \leq 1 - (\lambda + o(1))\mu_\ell$.

Démonstration :

On note $H_i = a_i G_i$, $Q_\ell = \prod_{i \in I_\ell} (I - H_i)$.

$$Q_\ell = I - \sum_{i \in I_\ell} H_i + \sum_{m=2}^{n_{\ell+1}-n_\ell} (-1)^m S_m$$

avec $S_m = \sum_{\ell_1 > \ell_2 > \dots > \ell_m} H_{\ell_1} H_{\ell_2} \dots H_{\ell_m}$, $\ell_j \in I_\ell$.

$$Q'_\ell Q_\ell = I - \sum_{i \in I_\ell} (H_i + H'_i) + E_\ell \tag{3.1.1}$$

avec

$$E_\ell = \sum_{m=2}^{n_{\ell+1}-n_\ell} (-1)^m (S_m + S'_m) + \left(\sum_{i \in I_\ell} H'_i \right) \left(\sum_{i \in I_\ell} H_i \right) - \left(\sum_{i \in I_\ell} H'_i \right) \left(\sum_{m=2}^{n_{\ell+1}-n_\ell} (-1)^m S_m \right)$$

$$- \left(\sum_{m=2}^{n_{\ell+1}-n_{\ell}} (-1)^m S'_m \right) \left(\sum_{i \in I_{\ell}} H_i \right) + \left(\sum_{m=2}^{n_{\ell+1}-n_{\ell}} (-1)^m S'_m \right) \left(\sum_{m=2}^{n_{\ell+1}-n_{\ell}} (-1)^m S_m \right).$$

Sous (2), en notant $F_i = H_i + H'_i$, $\left\| \sum_{i \in I_{\ell}} F_i \right\|$ tend vers 0 ; donc, à partir d'un

certain rang, $I - \sum_{i \in I_{\ell}} F_i$ a toutes ses valeurs propres positives et, sous (1), on a :

$$\left\| I - \sum_{i \in I_{\ell}} F_i \right\| = 1 - \lambda \left(\sum_{i \in I_{\ell}} F_i \right) \leq 1 - \lambda \mu_{\ell}. \tag{3.1.2}$$

On note $\alpha_{\ell} = \sum_{i \in I_{\ell}} \|H_i\|$; on a $\|S_m\| \leq \alpha_{\ell}^m$; à partir d'un certain rang, $\alpha_{\ell} \leq \frac{1}{2}$

; alors, sous (3.1.2), on a :

$$\sum_{m \geq 2} \|S_m\| \leq \sum_{m \geq 2} \alpha_{\ell}^m \leq 2\alpha_{\ell}^2$$

$$\tag{3.1.3}$$

$$\|E_{\ell}\| \leq 8\alpha_{\ell}^2 = o(1)\mu_{\ell}. \tag{3.1.4}$$

De (3.1.1), (3.1.2), (3.1.4), on déduit que :

$$\|Q_{\ell}\|^2 \leq 1 - (\lambda + o(1))\mu_{\ell}.$$

Lemme 5 :

Soit (b_n) , (α_n) , (β_n) , (γ_n) , (δ_n) cinq suites de réels non négatifs, λ et Δ des réels positifs tels que :

(1) pour tout n , $b_{n+1} \leq (1 - \lambda\alpha_n + \beta_n)b_n + \gamma_n + \alpha_n\delta_n$;

(2) $\alpha_n = o(1)$, $\sum_1^{\infty} \alpha_n = \infty$, $\sum_1^{\infty} \beta_n < \infty$, $\sum_1^{\infty} \gamma_n < \infty$;

(3) pour tout n , $\delta_n \leq \Delta$.

Alors, $\overline{\lim} b_n \leq \frac{\Delta}{\lambda}$.

La démonstration de ce lemme est donnée en appendice.

Démonstration du théorème 3 :

A partir des définitions de X_n et W_n , on obtient, en notant $Y_n = X_n + W_n - \theta_n$,

$$H_n = a_n G_n, \quad Q_\ell = \prod_{i \in I_\ell} (I - H_i), \quad Q_{i,\ell} = \prod_{j=i+1}^{n_{\ell+1}-1} (I - H_j), \quad \alpha_\ell = \sum_{i \in I_\ell} \|H_i\|,$$

$$R_n = (G_n - I) W_n + \frac{1}{a_n} (\theta_n - \theta_{n+1}) :$$

$$Y_{n+1} = (I - H_n) Y_n + a_n R_n$$

$$Y_{n_{\ell+1}} = Q_\ell Y_{n_\ell} + \sum_{i \in I_\ell} a_i Q_{i,\ell} R_i. \quad (3.1.5)$$

On a

$$Q_{i,\ell} = I - \sum_{j=i+1}^{n_{\ell+1}-1} H_j + \sum_{m=2}^{n_{\ell+1}-i-1} (-1)^m S_{im}$$

avec

$$S_{im} = \sum_{\substack{\ell_1 > \ell_2 > \dots > \ell_m \\ \in [i+1, n_{\ell+1}-1]}} H_{\ell_1} H_{\ell_2} \dots H_{\ell_m}.$$

Sous (2), $\left\| \sum_{j=i+1}^{n_{\ell+1}-1} H_j \right\| = o(1)$; $\sum_{m=2}^{n_{\ell+1}-i-1} \|S_{im}\| \leq \sum_{m \geq 2} \alpha_\ell^m \leq 2\alpha_\ell^2 = o(1)$; donc,

en notant $E_\ell = \|Y_{n_\ell}\|$, on a :

$$E_{\ell+1} \leq \|Q_\ell\| E_\ell + \sum_{i \in I_\ell} a_i \|R_i\| (1 + o(1)).$$

Sous (1), (2), (3), d'après le lemme 4, il existe $\nu > 0$ tel que :

$$E_{\ell+1} \leq (1 - \nu \mu_\ell) E_\ell + o(1) \mu_\ell.$$

Par application du lemme 5, avec $b_\ell = E_\ell$, $\alpha_\ell = \mu_\ell$, $\beta_\ell = 0$, $\gamma_\ell = 0$, on conclut que E_ℓ tend vers 0.

Sous (2), (3), il existe $M > 0$ tel que, pour $n \in I_\ell$:

$$\| Y_{n+1} \| \leq M \| Y_n \| + o(1)\mu_{\ell}.$$

Donc, en prenant $n = n_{\ell}$, on a $Y_{n_{\ell}+1} \rightarrow 0$. Comme $\text{card } I_{\ell}$ est majoré, on obtient ainsi $Y_n \rightarrow 0$.

3.2 - Comportement asymptotique du moment d'ordre deux dans le cas linéaire

Théorème 4 :

Soit $a_n = \frac{1}{n^{\alpha}}$, avec $\alpha > 0$. On suppose les hypothèses suivantes vérifiées :

- (1) Pour tout n , G_n est symétrique, de type positif.
- (2) Il existe un réel λ positif, un entier r , un réel $0 < \beta \leq 1$, un réel $\gamma > 1 - \beta$ tels que :
 - (i) pour $\beta = 1$, $\gamma r^{\alpha} < 2\lambda$;
 - (ii) pour tout ℓ , $\lambda \left(\sum_{i \in I_{\ell}} G_i \right) \geq \lambda \ell^{\alpha-\beta}$, uniformément en $\omega \in \Omega$,
 où I_{ℓ} est l'intervalle d'entiers $[(\ell - 1)r + 1, \ell r]$.
- (3) $\| G_n \|^2 = o(n^{2\alpha-\beta})$, uniformément en ω .
- (4) $\| \theta_n - \theta_{n+1} \| = O(n^{-\beta-\gamma})$.
- (5) $E [\| (G_n - I)W_n \|^2]^{\frac{1}{2}} = O(n^{\alpha-\beta-\gamma})$.

Alors, $\overline{\lim} n^{\gamma} E [\| X_n + W_n - \theta_n \|^2] < \infty$.

Ce théorème est en fait un corollaire, en prenant $\mu_{\ell} = \frac{1}{\ell^{\beta}}$, de la proposition suivante que nous allons démontrer.

On note $R_n = (G_n - I)W_n + \frac{1}{a_n}(\theta_n - \theta_{n+1})$, $\xi_{\ell}^2 = E \left[\left(\sum_{i \in I_{\ell}} a_i \| R_i \| \right)^2 \right]$.

Proposition 2 :

On suppose les hypothèses suivantes vérifiées :

(1) Il existe un réel λ positif, un entier r , une suite croissante d'entiers (n_ℓ) vérifiant $n_{\ell+1} \leq n_\ell + r$, une suite (μ_ℓ) dans \mathbb{R}^+ tels que :

$$(i) \mu_\ell = o(1); \sum_1^\infty \mu_\ell = +\infty;$$

$$(ii) \text{ pour tout } \ell, \lambda \left(\min_{i \in I_\ell} a_i (G_i + G'_i) \right) \geq \lambda \mu_\ell, \text{ uniformément en } \omega \in \Omega,$$

où I_ℓ est l'intervalle d'entiers $[n_\ell, n_{\ell+1}[$.

$$(2) \sum_{i \in I_\ell} a_i^2 \|G_i\|^2 = o(1) \mu_\ell, \text{ uniformément en } \omega.$$

$$(3) \sum_1^\infty g_\ell < \infty.$$

(4) Il existe une suite (α_ℓ) dans \mathbb{R}^+ vérifiant $\sum_1^\infty \alpha_\ell < \infty$ et un réel $\gamma < \lambda$

tels que, pour tout ℓ :

$$\frac{\mu_{\ell+1}}{g_{\ell+1}} \leq \frac{\mu_\ell}{g_\ell} (1 + \gamma \mu_\ell + \alpha_\ell + o(\mu_\ell)).$$

Soit, pour tout n , $\ell(n)$ tel que $n \in I_{\ell(n)}$.

Alors, $\overline{\lim} \frac{\mu_{\ell(n)}}{g_{\ell(n)}} E [\|X_n + W_n - \theta_n\|^2] < \infty$.

Démonstration de la proposition 2 :

D'après (3.1.5) dans la démonstration du théorème 3, on a, en notant

$$Y_n = X_n + W_n - \theta_n, H_n = a_n G_n, Q_\ell = \prod_{i \in I_\ell} (I - H_i), Q_{i,\ell} = \prod_{j=i+1}^{n_{\ell+1}-1} (I - H_j) :$$

$$Y_{n_{\ell+1}} = Q_\ell Y_{n_\ell} + \sum_{i \in I_\ell} a_i Q_{i,\ell} R_i$$

$$\|Y_{n_{\ell+1}}\|^2 \leq \|Q_\ell\|^2 \|Y_{n_\ell}\|^2 + \left\| \sum_{i \in I_\ell} a_i Q_{i,\ell} R_i \right\|^2 + 2 \langle Q_\ell Y_{n_\ell}, \sum_{i \in I_\ell} a_i Q_{i,\ell} R_i \rangle.$$

Sous (1), (2), d'après le lemme 4, on a, uniformément en ω :

$$\|Q_\ell\|^2 \leq 1 - (\lambda + o(1))\mu_\ell.$$

Comme $\gamma < \lambda$, il existe $\nu > \gamma$ tel que, à partir d'un certain rang :

$$\|Q_\ell\|^2 \leq 1 - \nu\mu_\ell.$$

Sous (2), il existe $M > 0$ tel que, uniformément en ω :

$$\forall \ell, \forall i \in I_\ell, \|Q_{i,\ell}\| \leq M.$$

Alors, en notant $e_\ell^2 = E[\|Y_{n_\ell}\|^2]$, on a :

$$e_{\ell+1}^2 \leq (1 - \nu\mu_\ell)e_\ell^2 + M^2g_\ell^2 + 2M^2e_\ell g_\ell.$$

En écrivant $e_\ell \leq 1 + e_\ell^2$, comme, sous (3), $g_\ell \rightarrow 0$, il existe $M_1 > 0$ tel que :

$$e_{\ell+1}^2 \leq (1 - \nu\mu_\ell + 2M^2g_\ell)e_\ell^2 + M_1g_\ell.$$

Les hypothèses $\mu_\ell = o(1)$, $\sum_1^\infty \mu_\ell = \infty$, $\sum_1^\infty g_\ell < \infty$ impliquent, d'après le

lemme 5, que $e_\ell \rightarrow 0$. On en déduit qu'il existe $C > 0$ tel que :

$$e_{\ell+1}^2 \leq (1 - \nu\mu_\ell)e_\ell^2 + Cg_\ell.$$

Sous (1), (4), l'application du lemme 2 conduit à :

$$\overline{\lim} \frac{\mu_\ell}{g_\ell} e_\ell^2 < \infty.$$

Sous (2), il existe $M_2 > 0$ tel que, pour $n \in I_\ell$:

$$\frac{\mu_\ell}{g_\ell} E[\|Y_{n+1}\|^2] \leq M_2 \frac{\mu_\ell}{g_\ell} E[\|Y_n\|^2] + M_2\mu_\ell g_\ell.$$

Comme $\mu_\ell g_\ell \rightarrow 0$, on a, en prenant $n = n_\ell$:

$$\overline{\lim} \frac{\mu_\ell}{g_\ell} E[\|Y_{n_\ell+1}\|^2] < \infty.$$

Comme $\text{card } I_\ell$ est majoré, on obtient ainsi, en notant, pour tout n , $\ell(n)$ l'entier tel que $n \in I_{\ell(n)}$:

$$\overline{\lim} \frac{\mu_{\ell(n)}}{g_{\ell(n)}} E[\|Y_n\|^2] < \infty.$$

3.3 - Normalité asymptotique

On note, pour tout n , F_n la tribu engendrée par $X_1, G_1, U_1, \dots, G_{n-1}, U_{n-1}$, pour tout n , $R_n = (G_n - I) W_n + \frac{1}{a_n} (\theta_n - \theta_{n+1})$, pour tout $\ell, n_\ell = (\ell - 1)r + 1$,

$$Y_\ell = \sum_{i \in I_\ell} (R_i - E[R_i / F_{n_{\ell-1}}]), \text{ où } I_\ell \text{ est l'intervalle d'entiers } [n_\ell, n_{\ell+1}[.$$

Théorème 5 :

Soit $a_n = \frac{1}{n^\alpha}$, avec $0 < \alpha \leq 1$. On suppose les hypothèses suivantes vérifiées :

- (1) (i) Pour tout ℓ , pour $i \in I_\ell$, G_i est $F_{n_{\ell-1}}$ -mesurable.
 (ii) Il existe un entier r , une matrice carrée d'ordre k symétrique définie positive Γ tels que :
- (a) $2 \lambda_{\min}(\Gamma) > r^\alpha \beta_+$, avec $\beta_+ = 1$ si $\alpha = 1$, $\beta_+ = 0$ si $\alpha \neq 1$;
- (b) $\sum_{i \in I_\ell} G_i \xrightarrow{P.S.} \Gamma$.
- (iii) $n^{-\alpha} \|G_n\|^2 \xrightarrow{P.S.} 0$.

- (2) (i) Il existe une variable aléatoire R dans \mathbb{R}^+ telle que, pour tout n ,

$$\|R_n\| \leq R.$$

(ii) $\lambda^{\frac{\alpha}{2}} \sum_{i \in I_\ell} E[R_i / F_{n_{\ell-1}}] \xrightarrow{P.S.} 0$.

- (iii) Il existe une matrice carrée d'ordre k , Σ , un réel $C > 0$ tels que

$$C > \|E[Y_\ell Y_\ell' / F_{n_{\ell-1}}] - \Sigma\| \xrightarrow{P.S.} 0.$$

On note Λ la réduite diagonale de Γ , P la matrice orthogonale de passage telles que $P' \Gamma P = \Lambda$, M la matrice d'élément général (i,j) , $M^{(ij)} = (P' \Sigma P)^{(ij)} (\Lambda^{(ii)} + \Lambda^{(jj)} - r^\alpha \beta_+)^{-1}$.

Alors, $n^{\frac{\alpha}{2}} (X_n + W_n - \theta_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, P M P')$.

Ce théorème est un corollaire d'un théorème de Fabian [7] ; sa démonstration est donnée dans l'appendice.

APPENDICE

1 - Démonstration du lemme 5

On note $c_n = \lambda \alpha_n - \beta_n$, $\Delta_n = \gamma_n + \frac{\Delta}{\lambda} \beta_n$, $\prod_{jn} = \prod_{k=j+1}^n (1 - c_k)$ pour $j < n$,

$\prod_{jn} = 1$ pour $j \geq n$.

Sous (2), $c_n \rightarrow 0$; on suppose que, pour tout n , $1 - c_n > 0$. Sous (3), on obtient :

$$b_{n+1} \leq \prod_{0n} b_1 + \sum_{j=1}^n \prod_{jn} \Delta_j + \frac{\Delta}{\lambda} \sum_{j=1}^n \prod_{jn} c_j.$$

Sous (2), $\prod_{0n} \rightarrow 0$; $\sum_{j=1}^n \prod_{jn} c_j \rightarrow 1$. Donc :

$$\overline{\lim} b_{n+1} \leq \frac{\Delta}{\lambda} + \overline{\lim} \sum_{j=1}^n \prod_{jn} \Delta_j.$$

Soit $\epsilon > 0$; soit n_0 tel que, pour tout $n > n_0$, on ait $\sum_1^n c_i > \text{Log } \frac{1}{\epsilon}$; soit, pour

$n > n_0$, $K(n) = \max \left\{ k : \sum_{i=k+1}^n c_i > \text{Log } \frac{1}{\epsilon} \right\}$. On a : $0 \leq K(n) < n$, $K(n) \leq K(n + 1)$,

$K(n) \rightarrow \infty$.

On considère la décomposition :

$$\sum_{j=1}^n \prod_{jn} \Delta_j = \sum_{j=1}^{K(n)} \prod_{K(n)n} \prod_{jK(n)} \Delta_j + \sum_{j=K(n)+1}^n \prod_{jn} \Delta_j.$$

On a : $\prod_{K(n)n} \leq \exp\left(-\sum_{j=K(n)+1}^n c_j\right) < \varepsilon$; pour tout $\ell, p, \prod_{\ell p} \leq \prod_1^{\infty} (1 + \beta_j) < \infty$;
à partir d'un certain rang, $\sum_{j=K(n)+1}^{\infty} \Delta_j < \varepsilon$. On en déduit que $\overline{\lim} \sum_{j=1}^n \prod_{j_n} \Delta_j = 0$, donc
 $\overline{\lim} b_{n+1} \leq \frac{\Delta}{\lambda}$.

2 - Démonstration du lemme 2

On note $b_n^1 = \frac{a_n}{g_n} b_n$. Sous (1), (2), (3), (4), on a :

$$b_{n+1}^1 \leq (1 - (\lambda - \gamma)a_n + \gamma_n + o(a_n))b_n^1 + \mu a_n(1 + \gamma a_n + \gamma_n + o(a_n)).$$

Soit $0 < \varepsilon < 1$. Sous (2) et (4), il existe n_0 tel que, pour $n > n_0$:

$$b_{n+1}^1 \leq (1 - (\lambda - \gamma)(1 - \varepsilon)a_n + \gamma_n)b_n^1 + \mu(1 + \varepsilon)a_n.$$

On peut appliquer le lemme 5 : $\overline{\lim} b_{n+1}^1 \leq \frac{\mu}{\lambda - \gamma} \frac{1 + \varepsilon}{1 - \varepsilon}$; comme ε est quel-

conque, $\overline{\lim} b_{n+1}^1 \leq \frac{\mu}{\lambda - \gamma}$.

3 - Démonstration du théorème 5

Théorème (Fabian [7]) :

Soit, pour tout n , U_n, V_n, T_n des variables aléatoires dans \mathbb{R}^k , Γ_n, Φ_n des variables aléatoires matricielles carrées d'ordre k , définies sur un espace probabilisé (Ω, \mathcal{C}, P) ; soit un réel $0 < \alpha \leq 1$, un réel $\beta \geq 0$ tels que :

$$U_{n+1} = (I - n^{-\alpha}\Gamma_n)U_n + n^{-(\alpha+\beta)/2}\Phi_n V_n + n^{-\alpha-\beta/2}T_n.$$

Soit (F_n) une suite non décroissante de sous-tribus de \mathcal{C} .

On suppose les hypothèses suivantes vérifiées :

(1) (i) Pour tout n , Γ_n est F_n -mesurable.

(ii) Il existe une matrice carrée d'ordre k symétrique définie positive Γ

telle que :

- (a) $2 \lambda_{\min}(\Gamma) > \beta_+$ avec $\beta_+ = \beta$ si $\alpha = 1$, $\beta_+ = 0$ si $\alpha \neq 1$;
- (b) $\Gamma_n \xrightarrow{\text{p.s.}} \Gamma$.

(2) (i) Pour tout n , Φ_{n-1} est F_n -mesurable.

(ii) Il existe une matrice carrée d'ordre k , Φ , telle que $\Phi_n \xrightarrow{\text{p.s.}} \Phi$.

(3) (i) Pour tout n , V_{n-1} est F_n -mesurable.

(ii) Pour tout n , $E[V_n / F_n] = 0$ p.s.

(iii) Il existe une matrice carrée d'ordre k , Σ , un réel $C > 0$ tels que :

$$C > \|E[V_n V_n' / F_n] - \Sigma\| \xrightarrow{\text{p.s.}} 0.$$

(iv) Pour tout $r > 0$, $\lim_{j \rightarrow \infty} \sigma_{j,r}^2 = 0$, avec $\sigma_{j,r}^2 = E[I_{\{\|V_j\|^2 \geq r\}} \|V_j\|^2]$,

ou, pour $\alpha = 1$, $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \sigma_{j,r}^2 = 0$.

(4) Il existe $T \in \mathbb{R}^k$ tel que $T_n \xrightarrow{\text{p.s.}} T$ ou $E[\|T_n - T\|] \rightarrow 0$.

On note Λ la réduite diagonale de Γ , P la matrice orthogonale de passage telles que $P' \Gamma P = \Lambda$, M la matrice d'élément général (i, j) , $M^{(ij)} = (P' \Phi \Sigma \Phi' P)^{(ij)} (\Lambda^{(ii)} + \Lambda^{(jj)} - \beta_+)^{-1}$.

Alors, $n^{\frac{\beta}{2}} U_n \xrightarrow{\mathcal{L}} \mathcal{N}((\Gamma - (\beta_+ / 2)I)^{-1} T, PMP')$.

Démonstration du théorème 5 :

D'après (3.1.5) dans la démonstration du théorème 3, on a, en notant

$$H_n = n^{-\alpha} G_n, Q_\ell = \prod_{i \in I_\ell} (I - H_i), Q_{i,\ell} = \prod_{j=i+1}^{\ell r} (I - H_j), S_m = \sum_{\ell_1 > \ell_2 > \dots > \ell_m} H_{\ell_1} H_{\ell_2} \dots H_{\ell_m}$$

pour $\ell_j \in I_\ell$, $S_{im} = \sum_{\substack{\ell_1 > \ell_2 > \dots > \ell_m \\ \in [i+1, \ell r]}} H_{\ell_1} H_{\ell_2} \dots H_{\ell_m}, Y_n = X_n + W_n - \theta_n, E_\ell = Y_{n,\ell} :$

$$E_{\ell+1} = Q_\ell E_\ell + \sum_{i \in I_\ell} i^{-\alpha} Q_{i,\ell} R_i$$

$$E_{\ell+1} = Q_{\ell}E_{\ell} + (\ell r)^{-\alpha}Y_{\ell} + \ell^{-\frac{3\alpha}{2}}T_{\ell}$$

avec

$$T_{\ell} = \ell^{-\frac{3\alpha}{2}} \left(\sum_{i \in I_{\ell}} i^{-\alpha} Q_{i\ell} R_i - (\ell r)^{-\alpha} Y_{\ell} \right).$$

Sous (1i), Q_{ℓ} est $F_{n_{\ell}-1}$ -mesurable.

$$\begin{aligned} Q_{\ell} &= I - \sum_{i \in I_{\ell}} H_i + \sum_{m=2}^r (-1)^m S_m \\ &= I - (\ell r)^{-\alpha} \sum_{i \in I_{\ell}} G_i - \sum_{i \in I_{\ell}} (1 - i^{\alpha} (\ell r)^{-\alpha}) H_i + \sum_{m=2}^r (-1)^m S_m. \end{aligned}$$

Pour tout ℓ , pour $i \in I_{\ell}$, il existe $\beta > 0$ tel que $1 - i^{\alpha} (\ell r)^{-\alpha} \leq \beta (\ell r)^{-\alpha}$.

Sous (1iii), $i^{\frac{\alpha}{2}} H_i \xrightarrow{\text{p.s.}} 0$; d'après (3.1.3) dans la démonstration du lemme 4, on a à partir d'un certain rang :

$$\sum_{m=2}^r \|S_m\| \leq 2 \left(\sum_{i \in I_{\ell}} \|H_i\| \right)^2.$$

On en déduit que :

$$\ell^{\alpha} \left(\sum_{i \in I_{\ell}} (1 - i^{\alpha} (\ell r)^{-\alpha}) H_i - \sum_{m=2}^r (-1)^m S_m \right) \xrightarrow{\text{p.s.}} 0.$$

Sous (1ii), $\ell^{\alpha} (\ell r)^{-\alpha} \sum_{i \in I_{\ell}} G_i \xrightarrow{\text{p.s.}} r^{-\alpha} \Gamma$. Donc, on peut écrire :

$$Q_{\ell} = I - \ell^{-\alpha} \Gamma_{\ell}, \text{ avec } \Gamma_{\ell} \xrightarrow{\text{p.s.}} r^{-\alpha} \Gamma.$$

On va appliquer le théorème de Fabian, en prenant $\Phi_{\ell} = I$, $V_{\ell} = r^{-\alpha} Y_{\ell}$, $\beta = \alpha$.

Comme, sous (2i), $\|Y_\ell\|$ est majorée par une variable aléatoire, l'hypothèse concernant $\sigma_{j,r}^2$ est vérifiée.

Montrons que $T_\ell \xrightarrow{P.S.} 0$.

$$\text{On a } Q_{i,\ell} = I - \sum_{j=i+1}^{\ell r} H_j + \sum_{m=2}^{\ell r-i} (-1)^m S_{im}$$

$$T_\ell = \ell^{\frac{3\alpha}{2}} \left((\ell r)^{-\alpha} \sum_{i \in I_\ell} E [R_i / F_{n_{\ell r-1}}] + \sum_{i \in I_\ell} i^{-\alpha} (1 - i^{-\alpha} (\ell r)^{-\alpha}) I \right. \\ \left. - \sum_{j=i+1}^{\ell r} H_j + \sum_{m=2}^{\ell r-i} (-1)^m S_{im} R_i \right).$$

$$\text{Sous (2) : } \ell^{\frac{3\alpha}{2}} (\ell r)^{-\alpha} \sum_{i \in I_\ell} E [R_i / F_{n_{\ell r-1}}] \xrightarrow{P.S.} 0$$

$$\ell^{\frac{3\alpha}{2}} \sum_{i \in I_\ell} i^{-\alpha} (1 - i^{-\alpha} (\ell r)^{-\alpha}) R_i \xrightarrow{P.S.} 0.$$

Sous (1iii) et (2), en majorant $\sum_{m=2}^{\ell r-i} \|S_{im}\|$ par $2 \left(\sum_{i \in I_\ell} \|H_i\| \right)^2$, on déduit que

$$\ell^{\frac{3\alpha}{2}} \sum_{i \in I_\ell} i^{-\alpha} \left(- \sum_{j=i+1}^{\ell r} H_j + \sum_{m=2}^{\ell r-i} (-1)^m S_{im} \right) R_i \xrightarrow{P.S.} 0.$$

Donc, $T_\ell \xrightarrow{P.S.} 0$.

On applique alors le théorème de Fabian :

$$\ell^{\frac{\alpha}{2}} E_\ell \xrightarrow{\mathfrak{L}} \mathcal{N}(0, PM_1 P')$$

avec $M_1^{(ij)} = r^{-\alpha} (P' \Sigma P)^{(ij)} (\Lambda^{(ii)} + \Lambda^{(jj)} - r^\alpha \beta_+)^{-1}$.

Donc : $n^{\frac{\alpha}{2}} Y_{n,\ell} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \text{PMP}')$ avec $M = r^\alpha M_1$.

Or : $n^{\frac{\alpha}{2}} Y_{n+1} - n^{\frac{\alpha}{2}} Y_n = -n^{\frac{\alpha}{2}} G_n Y_n + n^{\frac{\alpha}{2}} R_n$. Sous (liii) et (2i), comme $\text{card } I_\ell$ est majoré, on obtient :

$$n^{\frac{\alpha}{2}} Y_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \text{PMP}').$$

BIBLIOGRAPHIE

- [1] ALBERT A.E., GARDNER L.A., Stochastic approximation and nonlinear regression. Research Monograph 42, M.I.T. Press, Cambridge, Massachusetts, 1967.
- [2] ASSILI O., Régression séquentielle par approximation stochastique. Thèse de Doctorat de troisième cycle, Université de Nancy I, 1985.
- [3] BENNAR A., Approximation stochastique : convergence dans le cas de plusieurs solutions et étude de modèles de corrélation. Thèse de Doctorat de troisième cycle, Université de Nancy I, 1985.
- [4] BENVENISTE A., METIVIER M., PRIOURET P., Algorithmes adaptatifs et approximations stochastiques, théorie et applications. Masson, 1987.
- [5] CHUNG K.L., On a stochastic approximation method. A.M.S., vol. 25, pp. 463-483, 1954.
- [6] EL YASSA M., Rapidité de convergence d'un processus général d'approximation stochastique sous contraintes convexes. Thèse de Doctorat de troisième cycle, Université de Nancy I, 1985.
- [7] FABIAN V., On asymptotic normality in stochastic approximation. A.M.S., vol. 39, pp. 1327-1332, 1968.
- [8] GLADYSHEV E.G., On stochastic approximation. Theory of Probability and its Applications, vol. 10, pp. 275-278, 1965.
- [9] I.E.E.E., Transactions on Information Theory, vol. 30, n° 2, 1984.
- [10] LJUNG L., Strong convergence of a stochastic approximation algorithm. A.S., vol. 6, n° 3, pp. 680-696, 1978.
- [11] LJUNG L., Analysis of stochastic gradient algorithms for linear regression problems. I.E.E.E.-I.T., vol. 30, n° 2, pp. 151-160, 1984.
- [12] LJUNG L., SODERSTROM T., Theory and practice of recursive identification. M.I.T. Press, Cambridge, Massachusetts, 1983.

- [13] **METIVIER M., PRIOURET P.**, Applications of a Kushner and Clark lemma to general classes of stochastic algorithms. I.E.E.E.-I.T., vol. 30, n° 2, pp. 140-151, 1984.
- [14] **METIVIER M., PRIOURET P.**, Théorèmes de convergence presque sûre pour une classe d'algorithmes stochastiques à pas décroissant. Prob. Th. Rel. Fields, vol. 74, pp. 403-428, 1987.
- [15] **MONNEZ J.M.**, Convergence presque sûre de processus d'approximation stochastique dynamique. Statistique et Analyse des Données, vol. 14, n° 2, pp. 55-79, 1989.
- [16] **ROBBINS H., MONRO S.**, A stochastic approximation method. A.M.S., vol. 22, pp. 400-407, 1951.
- [17] **ROBBINS H., SIEGMUND D.**, A convergence theorem for nonnegative almost supermartingales and some applications. Optimizing methods in Statistics, edited by J.S. Rustagi, Academic Press, New York, pp. 233-257, 1971.