ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

G. PERENNOU

Convergence presque sûre et en moyenne des processus d'approximation stochastique, (1ère Partie) processus de Dvoretzky

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 58, série *Mathématiques*, n° 12 (1976), p. 132-165

<http://www.numdam.org/item?id=ASCFM_1976__58_12_132_0>

CONVERGENCE PRESQUE SURE ET EN MOYENNE DES PROCESSUS D'APPROXIMATION STOCHASTIQUE (1ère Partie) PROCESSUS DE DVORETZKY

G. PERENNOU, Université de Toulouse

Introduction et Résumé

Soient, pour tout entier n, T_n une application d'un ensemble H dans lui-même (ou de H^{n+1} dans H), U_n une variable aléatoire à valeurs dans H (v.a. H.). Soit encore X_o une v.a.H. et $X_{n+1} = T_n(X_n) + U_n$ pour $n = 0,1,2,\ldots$.

Si l'on suppose que U_n est une erreur (ou une perturbation) on conçoit l'intérêt du problème qui consiste à rechercher les conditions sous lesquelles, pour un mode de convergence donné, le comportement limite de $(X_n)_{\mathbb{N}}$ n'est pas affecté par $(U_n)_{\mathbb{N}}$.
Dans le cas où $H=\mathbb{R}$ c'est sous cette forme que Dvoretzky généralisa les processus d'approximations stochastiques introduits par Robbins et Monro.
L'étude de ces processus, souvent appelés processus de Dvoretzky s'est ensuite développée grâce à la contribution, notamment, de Wolfowitz dans le cas où $H = \mathbb{R}$, de Derman et Saks dans le cas où $H = \mathbb{R}^n$ de Schmetterer et Venter dans le cas où H est un espace de Hilbert séparable. En ce qui concerne les processus de Robbins et Monro proprement dits, nous y reviendrons dans la deuxième partie.

Le présent travail est à rapprocher de celui de Venter par rapport auquel il apporte des extensions (corollaire 1.1.) et des résultats nouveaux (corollaire 1.2.). Une démonstration d'un énoncé de Venter est également donnée (Théorème 2).

L'originalité dans la présentation du travail réside dans le fait que les résultats de convergence apparaissent comme conséquence de propriétés de stabilité de classes de suites d'applications de $H^{n+1} \times \Omega$ dans H.

I-1- Généralités - notations - L'ensemble $\mathcal{E}_{\Sigma}(\alpha,\mathfrak{X})$.

I-1-1- Notations

Nous nous donnerons une fois pour toute un espace de Hilbert H séparable, un espace de probabilité $(\Omega,\mathcal{Q},P)$ et une famille croissante de sous-tribus de $\mathcal{Q}$ noté $\mathcal{C} = (\mathcal{Q}_n)_{\mathbb{N}}$. $\mathcal{B}_H$ désignera d'une manière générale la tribu borélienne de H ($\mathcal{B}_{\mathbb{R}}$ dans le cas $H = \mathbb{R}$).

Nous utiliserons sans les repréciser les notations suivantes :

$<.,.>$: produit scalaire dans H (où $\mathbb{R}^n$),
$||.||$: norme associée au produit scalaire précédent,
$(.)_+$: $= \max\{o,.\}$ (dans $\mathbb{R}$)
ℓ: ensemble des suites numériques de séries convergentes,
ℓp : ensemble des suites numériques $(u_n)_{\mathbb{N}}$ telle que $\Sigma|u_n|^p<\infty$,

$\ell_\pi = \{(u_n)_{\mathbb{N}} \mid \forall n \geq o,\ \Pi_{i \leq n}(1+u_i) = B_n \in]o,\infty[, B_n \to B \in]o,\infty[\}$,

$(\ell^c)^+$: ensemble des suites numériques positives de séries divergentes,

$(\ell^c)^+_o$: sous-ensemble du précédent où de plus la suite converge vers zéro,

$\mathcal{S}_A(B)$: ensemble des suites d'applications $(T_n)_{\mathbb{N}}$ où $T_n : B^{n+1} \times \Omega \to A$.

Nous aurons à utiliser l'espérance mathématique E de variables aléatoires à valeurs dans un espace de Hilbert séparable H ainsi que l'espérance mathématique conditionnelle $E^{\mathcal{G}_n}$ relativement à $\mathcal{G}_n$. Précisons à cet égard que ceci est possible pour H espace de Hilbert séparable et X variable aléatoire à valeurs dans H quand $E\|X\| < \infty$ ($\|X\|$ est dans ce cas une variable aléatoire réelle) ; on a alors (voir par exemple [9] et [10]) :

a) $\| EX\| \leq E\|X\|$,

b) $E < h,X > = < h,EX >$ pour tout $h \in H$,

c) $\|E^{\mathcal{G}_n}X\| \leq E^{\mathcal{G}_n} \|X\|$ p.s.,

d) $E^{\mathcal{G}_n} < h,X > = < h, E^{\mathcal{G}_n}X >$ pour tout h variable aléatoire à valeur dans H et $\mathcal{G}_n$ mesurable.

L'abréviation suivante : v.a.C. (en particulier v.a.$\mathbb{R}$. ou v.a.H.), désignera une variable aléatoire à valeurs dans C.

Enfin si $X = (X_n)_{\mathbb{N}}$ est une suite d'applications de Ω dans H et $(T_n)_{\mathbb{N}}$ une suite de $\mathcal{S}_A(H)$ nous désignerons par $T_n(X)$ l'application composée $T_n(X_o(.),\ldots,X_n(.),.)$ de Ω dans A , ceci pour tout $n \geq o$.

I-1.2. QUELQUES DEFINITIONS ET PROPRIETES GENERALES

Nous définirons d'abord une classe de $\mathcal{S}_H(H)$ qui joue dans la suite un rôle fondamental. Son intérêt apparaîtra clairement au paragraphe I.3. .

DEFINITION 1.1.

Soient $\Sigma = (\sigma_n)_{\mathbb{N}}$ une suite d'applications de H dans lui-même, $\mathfrak{X}$ une famille de suites d'applications de Ω dans H, α une application de $\mathfrak{X} \times \Omega$ dans $\mathbb{R}_+$ (pour X fixé on notera $\alpha(X)$ cette application) et $\mathcal{C} = (\mathfrak{a}_n)_{\mathbb{N}}$ une suite croissante de sous-tribus de $\mathfrak{a}$.

Nous désignerons par $\mathcal{C}_\Sigma(\alpha, \mathfrak{X})$ (resp. $\mathcal{C}_\Sigma(\alpha, \mathfrak{X}, \mathcal{C})$) le sous-ensemble de suites $(T_n)_{\mathbb{N}}$ de $\mathcal{S}_H(H)$ vérifiant pour tout $X = (X_n)_{\mathbb{N}} \in \mathfrak{X}$:

(1.0) Pour tout $n \in \mathbb{N}$ $T_n(X)$ est une v.a.H (resp. v.a.H $\mathfrak{a}_n$- mesurable); de plus il existe trois suites $(\alpha_n)_{\mathbb{N}}$, $(\beta_n)_{\mathbb{N}}$ et $(\gamma_n)_{\mathbb{N}}$ de $\mathcal{S}_{\mathbb{R}}(H)$ telles que :

$\exists \Omega_0(X) = \Omega$ p.s. , $\forall \omega \in \Omega_0(X)$, $\exists N(X)(\omega)$ fini .

$\forall n \geq N(X)(\omega)$,

(1.1) $\|T_n(X) \leq \max\{\alpha_n(X), (1+\beta_n(X)) \| \sigma_n X_n\| - \gamma_n(X)\}$

(1.2) $(\alpha_n(X))_{\mathbb{N}}$ converge vers $\alpha(X)$,

(1.3) $(\beta_n(X))_{n \geq N(X)} \in \ell_\pi$,

(1.4) $(\gamma_n(X))_{n \geq N(X)} \in (\ell^c)_0^+$ à chaque fois que $(\|\sigma_n X_n(\omega)\|)_{\mathbb{N}}$ est bornée.

Notons immédiatement que l'on peut manifestement choisir $(\alpha_n(X))_{\mathbb{N}}$ décroissante vers $\alpha(X)$ et $\beta_n(X)$ tel $1+\beta_n(X) > 0$ pour tout n (sur $\Omega_0(X)$).

Les propositions 1.1. et 1.2. suivantes permettent encore d'assouplir la définition précédente.

Proposition 1.1.

Dans la définition 1.1. il est équivalent d'imposer que $(\gamma_n)_{n \geq N(X)}$ soit dans $(\ell^c)^+$, les autres conditions demeurant inchangées.

Ceci résulte facilement du

Lemme 1.1.

Toute suite de $(\ell^c)^+$ est minorée par une suite de $(\ell^c)_0^+$.

Ce résultat est bien connu. Nous le laissons sans démonstration et de même pour le lemme 1.2. qui sera utile pour la démonstration de la proposition 1.2.

Lemme 1.2.

Soit $(u_n)_{\mathbb{N}}$ une suite numérique positive telle que $\sum u_n = \infty$. Il est toujours possible de trouver une suite $(\Theta_n)_{\mathbb{N}}$ positive qui décroit vers zéro et telle que $\sum u_n \Theta_n = \infty$.

Proposition 1.2.

Dans la définition 1.1. la condition (1.1.) est équivalente à la suivante (l'argument X est omis) (1.1. bis) $\|T_n\|^p \leq \max\{\alpha_n^p, (1+\beta_n)^p \|\sigma_n X_n\|^p - \gamma_n\}$ pourvu que $0 < p < \infty$, les autres conditions demeurant inchangées.

Démonstration

Nous montrerons que (1.1) $\Longrightarrow$(1.1 bis). L'implication inverse se démontre de même, essentiellement en remplaçant p par 1/p.
L'argument X sera omis dans la démonstration. Compte-tenu du lemme 1.2 ci-dessus, nous supposerons, sans perte de généralité, que la suite $(\alpha_n)_{\mathbb{N}}$ de la définition 1.1. vérifie (avec $(\gamma_n)_{\mathbb{N}}$) : $\alpha_n > o$ pour tout $n \geq N$,

$$\sum \alpha_n^{p-1} \gamma_n = \infty \text{ sur l'ensemble où } \sum \gamma_n = \infty ,$$

et que $(1+\beta_n) \geq o$ pour tout n.
Deux cas sont à envisager.

1er Cas : $p \leq 1$

Il résulte de (1.1.) que sur Ω_o et pour $n \geq N$:
$\|T_n\|^p \leq \max\{\alpha_n^p, (1+\beta_n)^p \|\sigma_n X_n\|^p - \gamma'_n\}$ (1.1. bis),
avec $\gamma'_n = p(1+\beta_n)^{p-1} \|\sigma_n X_n\|^{p-1} \gamma_n$.
En effet si $(1+\beta_n) \cdot \|\sigma_n X_n\| \geq \gamma_n + \alpha_n$, (1.1.) entraîne que $\|T_n\| \leq (1+\beta_n) \|\sigma_n X_n\| - \gamma_n$.
De l'inégalité générale $(X-\gamma)^p \leq X^p - p.\, X^{p-1}\gamma$ valable pour $p \leq 1$, $X \geq o$ et $X \geq \gamma$ résulte que : $\|T_n\|^p \leq (1+\beta_n)^p \|\sigma_n X_n\|^p - \gamma'_n$, ce qui implique (1.1. bis).
Si au contraire $(1+\beta_n)\|\sigma_n X_n\| < \alpha_n + \gamma_n$, (1.1.) se réduit à $\|T_n\| \leq \alpha_n$ ce qui implique aussi (1.1. bis).
La suite $(\alpha_n)_{\mathbb{N}}$ vérifie par hypothèse (1.2) et de même $(\beta_n)_{\mathbb{N}}$ vérifie (1.3) .
Quant à la suite $(\gamma'_n)_{\mathbb{N}}$ elle vérifie (1.4) vu que si $\|\sigma_n X_n\| \leq r < \infty$ on a $\gamma'_n \geq p\gamma_n r^{p-1}(1+\beta_n)^{p-1}$; sur Ω_o et pour $n \geq N$ il vient alors que : $\gamma'_n \geq o$ et $\sum \gamma'_n = \infty$.

2ème Cas : $p>1$

De (1.1.) résulte encore que sur Ω_o et pour $n \geq N$:
$\|T_n\|^p \leq \max\{\alpha_n^p, (1+\beta_n)^p \|\sigma_n X_n\|^p - \gamma'_n\}$ où $\gamma'_n = \gamma_n (\gamma_n + \alpha_n)^{p-1}$
En effet si $(1+\beta_n)\|\sigma_n X_n\| \geq \alpha_n + \gamma_n$, alors (1.1.) implique

$$\|T_n\|^p \leq (1+\beta_n)^p \|\sigma_n X_n\|^p (1 - \gamma_n (1+\beta_n)-1 \|\sigma_n X_n\|^{-1})^p$$

$$\leq (1+\beta_n)^p \|\sigma_n X_n\|^p (1 - \gamma_n (1+\beta_n)^{-1} \|\sigma_n X_n\|^{-1})$$

(car la dernière parenthèse est un réel entre 0 et 1 et on a $X^p \leq X$ si $X \in (0,1)$ et $p \geq 1$),
soit : $\|T_n\|^p \leq (1+\beta_n)^p \|\sigma_n X_n\|^p - \gamma'_n$, ce qui entraîne (1.1. bis) .

Si au contraire $(1+\beta_n)\|\sigma_n X_n\| < \alpha_n + \gamma_n$, (1.1.) se réduit à $\|T_n\| \leq \alpha_n$, inégalité qui implique (1.1. bis).
Par ailleurs (1.2.) et (1.3.) sont vérifiées comme au premier cas. Compte-tenu du choix de $(\alpha_n)_{\mathbb{N}}$ il est clair que $(\gamma'_n)_{\mathbb{N}}$ vérifie (1.4.) .

I-2. PROPRIETES DE STABILITE DE $\mathcal{C}_{\Sigma}(\alpha,\mathfrak{X})$

Nous établirons les propriétés de convergence des processus d'approximation stochastique en nous appuyant sur les propriétés de stabilité étudiées dans ce paragraphe.

Proposition 1.3.

Soient $\mathfrak{X}$ une famille de suites d'applications de Ω dans H et pour tout X de $\mathfrak{X}$ $(a_n(X))_{\mathbb{N}}$ une suite d'applications de Ω dans $\mathbb{R}_+$ qui convergent vers $\alpha(X)$ (application de Ω dans $\mathbb{R}_+$) en dehors d'un ensemble négligeable.

Si pour tout $k \in \mathbb{N}$ la suite $(T_n)_{\mathbb{N}}$ est dans $\mathcal{C}_{\Sigma}(a_k,\mathfrak{X})$ alors on a aussi $(T_n)_{\mathbb{N}}$ dans $\mathcal{C}_{\Sigma}(\alpha,\mathfrak{X})$.

Cette proposition permet de simplifier les démonstrations des propositions 1.4 et 1.5 .

Cependant en affaiblissant légèrement leurs énoncés, on conserverait des démonstrations simples et des résultats suffisants pour la convergence des processus d'approximation stochastique.

Nous nous bornerons donc, afin d'alléger, à donner un schéma de démonstration.

1°) Il est possible de se restreindre au cas où $(a_n(X))$ est décroissante vers $\alpha(X)$;

2°) Soient $(\alpha_n(k))$, $(\beta_n(k))$, $(\gamma_n(k))$ les trois suites de la définition 1.1 relative à l'assertion $(T_n)_{\mathbb{N}} \in \mathcal{C}_{\Sigma}(a_k,\mathfrak{X})$. Il est possible de choisir ces suites de manière que $\alpha_n(k) \geq \alpha_n(k+1)$;

3°) En dehors d'un ensemble négligeable il est possible de construire trois suites $(\alpha_n^*(X))$, $(\beta_n^*(X))$ et $(\gamma_n^*(X))$ comme suit : supposons ces suites construites jusqu'au rang $N_{k-1} \geq 0 (k \geq 1)$ et posons :

$$N_k^{(1)} = \min \{n : \sum_{\ell=N_{k-1}}^{n} \gamma_\ell(k) \geq 1/k\} \text{ (= 0 si } (\sigma_n X_n)_{\mathbb{N}} \text{ n'est pas bornée).}$$

$$N_k^{(2)} = \min \{n : \prod_{\ell=r}^{s} (1+\beta_\ell(k)) \in \left[1-1/k^2,\ 1+1/k^2\right]$$

$$N_k = \max \{N_k^{(1)}, N_k^{(2)}\}$$

$$\left.\begin{array}{l} \alpha_m^* = \alpha_m(k) \\ \beta_m^* = \beta_m(k) \\ \gamma_m^* = \gamma_m(k) \end{array}\right\} \quad \text{pour } N_{k-1} < m \leq N_k$$

4°) $\|T_n\| \leq \max\{\alpha_n^*, (1+\beta_n^*)\sigma_n X_n - \gamma_n^*\}$ et les conditions (1.0), (1.1), (1.2) , (1.3) et (1.4) sont vérifiées.

Nous établirons maintenant un lemme utile pour les deux propositions suivantes.

Lemme 1.3.

Soient $(T_n)_{\mathbb{N}}$ une suite de $\mathcal{C}_\Sigma(\alpha,\mathfrak{X})$, $(A_n)_{\mathbb{N}}$, $(B_n)_{\mathbb{N}}$ et $(C_n)_{\mathbb{N}}$ trois suites de $\mathcal{S}_{\mathbb{R}}(H)$ vérifiant pour tout $X \in \mathfrak{X}$, tout $\omega \in \Omega_1(X) = \Omega$ p.s., tout $n \geq N_1(X)$ ($N_1(X)$ étant une application de $\Omega_1(X)$ dans $\mathbb{N}$) :

(i) $A_n(X) \geq 0$, $B_n(X) \geq 0$ et $\sum A_n(X) < \infty$, $\sum B_n(X) < \infty$

(ii) $(C_n(X))_{n\geq N_1(X)} \in \ell_\pi$.

Soit $(T'_n)_{\mathbb{N}}$ une suite de $\mathcal{S}_H(H)$ vérifiant pour tout X de $\mathfrak{X}$, sur $\Omega_1(X)$, pour tout $n \geq N_1(X)$ et pour un réel $p \in \,]0, \infty[$:

(iii) $\|T'_n(X)\|^p \leq \|T_n(X)\|^p(1+C_n(X)) + A_n(X) + B_n(X)\|\sigma_n X_n\|^p$.

Alors $(T'_n)_{\mathbb{N}}$ est une suite de $\mathcal{C}_\Sigma(\alpha,\mathfrak{X})$ dès que $T'_n(X)$ est une v.a.H pour tout $n \in \mathbb{N}$ et tout $X \in \mathfrak{X}$.

Démonstration

Il résulte de la proposition 1.3. que l'on peut se restreindre au cas où $\inf_\Omega \alpha(X) = a(X) > 0$. Sinon il suffirait de faire la démonstration pour $\alpha_k(X) = \max\{1/k, \alpha(X)\}$ et de faire tendre k vers l'infini.
Sur X pour tout $\omega \in \Omega_0(X) \cap \Omega_1(X)$ et tout $n \geq \max\{N(X), N_1(X)\} = N_2(X)$ (Ω_0 et N déterminés par la définition 1.1. appliquée à $(T_n)_{\mathbb{N}}$) il est possible d'écrire (l'argument X étant omis) :

$$\|T'_n\|^p \leq \max\{\alpha_n^p, (1+\beta_n)^p\|\sigma_n X_n\|^p - \gamma_n\}(1+C_n) + A_n + B_n\|\sigma_n X_n\|^p .$$

Posons :

1°) $\alpha_n^{*p} = \alpha_n^p(1+C_n) + A_n + B_n(\alpha_n^p + \gamma_n)(1+\beta_n)^{-p}$, $(\alpha_n \geq 0)$:

on voit immédiatement que $(\alpha_n^*)_{\mathbb{N}}$ vérifie (1.2) définition 1.1. ;

2°) $(1+\beta_n^*)^p = (1+\beta_n)^p(1+C_n)\left(1+\frac{A_n}{(1+C_n)(\alpha_n^p+\gamma_n)} + \frac{B_n}{(1+C_n)(1+\beta_n)^p}\right)$, $(\beta_n^* \geq 0)$, où, puisque :

$\sum_{n\geq N_2} \frac{A_n}{(1+C_n)(\alpha_n+\beta_n)} < \infty$, $\sum_{n\geq N_2} \frac{B_n}{(1+C_n)(1+\beta_n)^p} < \infty$ le deuxième membre se présente comme le produit de trois termes, chacun extrait d'un produit infini convergent ; il s'en suit que $(\beta_n^*)_{n\geq N_2} \in \ell_\pi$ sur $\Omega_0 \cap \Omega_1$ et donc (1.3) définition 1.1. est vérifiée ;

3°) $\gamma_n^* = \gamma_n(1+C_n)$: la suite $(\gamma_n^*)_{n\geq N_2}$ vérifie (1.4) définition 1.1..

Par ailleurs une vérification simple montre que sur $\Omega_0 \cap \Omega_1$ et pour $n \geq N_2$ on a

$\|T'_n\|^2 \leq \max\{\alpha_n^{*p}, (1+\beta_n^*)^p\|\sigma_n X_n\|^p - \gamma_n^*\}$.

Si $T'_n(X)$ est, pour tout $n \in \mathbb{N}$ et tout $X \in \mathfrak{X}$, une v.a.H. les conditions de la définition 1.1. et de la proposition 1.2 sont vérifiées, ce qui établit le lemme.

Proposition 1.4.

Soient $(T_n)_{\mathbb{N}}$ une suite de $\mathscr{C}_{\Sigma}(\alpha,\mathfrak{X},\mathscr{C})$ et $(U_n)_{\mathbb{N}}$ une suite de $\mathscr{J}_H(H)$ qui pour tout $X \in \mathfrak{X}$ est une suite de v.a.H. (c'est-à-dire $U_n(X)$ est une v.a.H pour tout n). Supposons que les hypothèses suivantes soient vérifiées pour tout $X \in \mathfrak{X}$:

(1.5) $\sum E\|U_n(X)\|^2 < \infty$,

(1.6) $\sum \|E^{\mathcal{A}_n} U_n(X)\| < \infty$ p.s.,

alors $(T_n+U_n)_{\mathbb{N}}$ est une suite de $\mathscr{C}_{\Sigma}(\alpha,\mathfrak{X})$.

Démonstration

Moyennant la proposition 1.3 il est possible de supposer qu'il existe un réel $a(X)>0$ tel que $\alpha(X) \geqslant a(X)$ pour tout $X \in \mathfrak{X}$ (sinon, il suffirait de faire la démonstration en remplaçant $\alpha(X)$ par $a_k(X) = \max\{\frac{1}{k},\alpha(X)\}$) puis de faire tendre k vers l'infini.

D'une manière générale on peut écrire pour tout X de $\mathfrak{X}$ (l'argument X étant maintenant omis) :

$$\|T_n+U_n\|^2 = \|T_n\|^2 + 2<T_n,U_n> + \|U_n\|^2$$

$$\leqslant \max\{a^2,\|T_n\|^2\}(1+\beta_n') + \|U_n\|^2$$

où l'on a posé : $\beta_n' = 2<T_n, U_n> \max\{a^2,\|T_n\|^2\}^{-1}$ pour tout $n \in \mathbb{N}$.

Soit $T_n^* = \max\{\frac{a}{\|T_n\|}, 1\}.\ T_n$ (en convenant que $\|T_n^*\| = a$ quand $T_n = 0$).

On remarque d'abord que T_n^* est une v.a.H.

Soient $(\alpha_n)_{\mathbb{N}}$, $(\beta_n)_{\mathbb{N}}$ et $(\gamma_n)_{\mathbb{N}}$ les trois suites de la définition 1.1. relative à $(T_n)_{\mathbb{N}}$. Il est toujours possible de choisir la première décroissante vers α.

Par suite on a (pour $\omega \in \Omega_0(X)$, $n \geqslant N(X)$) :

1°) $\|T_n^*\| = \max\{a, \|T_n\|\} \leqslant \max\{\alpha_n, (1+\beta_n)\|\sigma_n X_n\| - \gamma_n\}$ en remarquant que $a \leqslant \alpha_n$ pour tout n), d'où l'on voit que $(T_n^*)_{\mathbb{N}} \in \mathscr{C}_{\Sigma}(\alpha,\mathfrak{X})$,

2°) $\|T_n+U_n\|^2 \leqslant \|T_n^*\|^2(1+\beta_n') + \|U_n\|^2$, où $\sum\|U_n\|^2 < \infty$ p.s. d'après (1.5)

D'après le lemme 1.3 le seul point à démontrer est que $(\beta_n')_{n \geqslant N_1} \in \ell_\pi$ p.s. pour $N_1(X)$ fini p.s.

A cette fin remarquons que :

$$|E^{\mathcal{A}_n}\beta_n'| \leqslant 2|<T_n, E^{\mathcal{A}_n}U_n>|\max\{a,\|T_n\|\}^{-2} \text{ p.s.}$$

$$\leqslant 2\, a^{-1}\|E^{\mathcal{A}_n} U_n\| \text{ p.s. } (1.8)$$

(en se rappelant que T_n est $\mathcal{A}_n$ mesurable) ;

Semblablement :

$$E|\beta_n'|^2 \leqslant 4.a^{-2}\, E\,\|U_n\|^2 \,(1.9).$$

De (1.5), (1.6), (1.8) et (1.9) résulte que :

$\sum |E^{\mathcal{A}_n} \beta'_n| < \infty$ p.s. , $\sum E|\beta'_n|^2 < \infty$, puis en utilisant un résultat classique de suites centrées :

$\sum \beta'_n$ converge p.s. et $\sum \beta'^2_n < \infty$ p.s..

Il suffit maintenant de choisir $N_1 = \min\{n : \beta'_n < -1/2\}$ pour voir que $(\beta_n)_{n \geq N_1} \in \ell_\pi$ p.s. avec N_1 p.s. fini.

Définition 1.2.

Nous désignerons par $\mathcal{D}_\Sigma(\alpha,\mathfrak{X})$ (resp. $\mathcal{D}_\Sigma(\alpha,\mathfrak{X},\mathcal{C})$) le sous-ensemble de $\mathcal{C}_\Sigma(\alpha,\mathfrak{X})$ resp. $\mathcal{C}_\Sigma(\alpha,\mathfrak{X},\mathcal{C})$) constitué des suites $(T_n)_{\mathbb{N}}$ assujetties à vérifier de plus pour tout $X \in \mathfrak{X}$

(1.11) $\exists M(X) \in \mathbb{N}$, $\exists r(X) \in \mathbb{R}_+$, $\exists k(X) > 0$, $\forall_n \geq M(X)$:

$$\|\sigma_n X_n\| \geq r(X) \Rightarrow \|T_n(X)\| \geq k(X) \cdot \|\sigma_n X_n\| \quad \text{p.s. .}$$

L'intérêt de la classe $\mathcal{D}_\Sigma$ ainsi définie tient à la propriété de stabilité ci-après, où le "bruit" est de la forme $\lambda_n U_n$ et peut croître avec $\|\sigma_n X_n\|$ au sens précis donné en (1.12).

Proposition 1.5.

Soient $(U_n)_{\mathbb{N}}$ et $(\lambda_n)_{\mathbb{N}}$ deux suites de $\mathcal{S}_H(H)$ et $\mathcal{S}_{\mathbb{R}}(H)$ respectivement vérifiant pour toute suite $X = (X_n)_{\mathbb{N}}$ d'un ensemble $\mathfrak{X}$ et pour une suite croissante $\mathcal{C}$ de sous-tribus de $\mathcal{A}$:

(1.12) (i) pour tout $n \in \mathbb{N}$, $U_n(X)$ est une v.a.H, $\lambda_n(X)$ est $\mathcal{A}_n$- mesurable et il existe $c_n(X)$ et $d_n(X) \in \mathbb{R}_+$ tel que

$$|\lambda_n(X)| \leq c_n(X) + d_n(X) \|\sigma_n X_n\| \quad \text{p.s.,}$$

$$\text{(ii)} \sum c_n(X) \|E^{\mathcal{A}_n} U_n(X)\| < \infty \quad \text{p.s.,}$$

$$\sum d_n(X) \|E^{\mathcal{A}_n} U_n(X)\| < \infty \quad \text{p.s.,}$$

$$\text{(iii)} \sum c_n^2(X)\, E\|U_n(X)\|^2 < \infty , \sum d_n^2(X)\, E\|U_n(X)\|^2 < \infty .$$

La suite $(T_n + \lambda_n U_n)_{\mathbb{N}}$ est alors dans $\mathcal{C}_\Sigma(\alpha,\mathfrak{X})$.

Démonstration

La démonstration est semblable à celle de la proposition précédente. Nous omettrons encore l'argument X. On peut encore supposer l'existence d'un réel $a>0$ tel que $\alpha \geq a$. D'une manière générale on peut écrire :

$\|T_n + \lambda_n U_n\|^2 = \max\{a^2, \|T_n\|^2\}(1+\beta'_n) + |\lambda_n|^2 \|U_n\|^2$, où l'on a posé $\beta'_n = 2\lambda_n \langle T_n, U_n\rangle \max\{a^2,\|T_n\|^2\}^{-1}$

Il suffit encore ici de prouver que $(\beta'_n)_{n \geq N_1} \in \ell_\pi$ p.s. pour $N_1(X)$ fini p.s. .

$|\overset{\mathcal{G}_n}{E}\beta_n| \leq 2|\lambda_n|.|<T_n, \overset{\mathcal{G}_n}{E}U_n>|.\max\{a^2,\|T_n\|^2\}^{-1}$ p.s.
(en se rappelant que T_n et λ_n sont $\mathcal{G}_n$ - mesurables)

$\leq 2(c_n + d_n\|\sigma_n X_n\|)\max\{a,\|T_n\|\}^{-1}.\ \|\overset{\mathcal{G}_n}{E}U_n\|$ p.s.

(d'après (1.12)(i)et puisque $\|T_n\| \leq \max\{a,\|T_n\|\}$)

$\leq 2(c_n\ a^{-1} + d_n\|\sigma_n X_n\|.\|T_n\|^{-1}).\ \|\overset{\mathcal{G}_n}{E}U_n\|$ p.s.

$\leq 2(c_n\ a^{-1} + d_n\ \max\{ra^{-1},k^{-1}\}).\|\overset{\mathcal{G}_n}{E}U_n\|$ p.s.

(en utilisant (1.11)).

Semblablement on a :

$E|\beta_n|^2 \leq 4(c_n\ a^{-1} + d_n\ \max\{r\ a^{-1},k^{-1}\})^2\ .\ E\|U_n\|^2$

De [(ii) et (iii)](1.12) résulte alors que $\sum|\overset{\mathcal{G}_n}{E}\beta'_n| < \infty$ p.s.
et $\sum E|\beta'_n|^2 < \infty$. La démonstration s'achève comme pour la proposition 1.4. .

Remarque. Avec les notations de la proposition 1.4. on voit que les hypothèses clés sont : (1.5) et $(\beta'_n)_{n\geq N_1} \in \ell_\pi$ jointes aux conditions de la définition 1.1. . Les propositions 1.4 et 1.5 sont alors deux versions particulières d'un résultat plus général basé sur ces hypothèses.

Dans la suite nous n'utiliserons que les propositions 1.4. et 1.5. .

I-3- CONVERGENCE PRESQUE SURE DES PROCESSUS D'APPROXIMATION STOCHASTIQUE

Soit $(T_n)_{\mathbb{N}}$ un élément de $\mathcal{B}_H(H)$ et notons $X=(X_n)_{\mathbb{N}} = \mathcal{A}(X_o,(T_n))$ la suite obtenue comme suit :

$$\begin{cases} X_o \text{ v.a.H,} \\ X_{n+1} = T_n(X_o(.),\dots,X_n(.),.) \quad n = 0,1,2,\dots \end{cases}$$

Plus généralement, soit $\Sigma = (\sigma_n)_{\mathbb{N}}$ une suite d'applications de H dans H et notons $X = (X_n)_{\mathbb{N}} \in \mathcal{A}_\Sigma(X_o,(T_n))$ toute suite définie comme suit :

$$\begin{cases} \sigma_o X_o \text{ v.a. H,} \\ \sigma_{n+1} X_{n+1} = T_n(X_o(.),\dots,X_n(.),.) \quad n=0,1,2,\dots \end{cases}$$

Nous appellerons d'une manière générale une telle suite processus d'approximation stochastique.

Théorème 1.1.

<u>Soient</u> X_o <u>une</u> v.a.H, $(T_n)_{\mathbb{N}} \in \mathcal{S}_H(H)$, $\Sigma = (\sigma_n)_{\mathbb{N}}$ <u>une suite d'applications de</u> H <u>dans</u> H <u>et</u> $(X_n)_{\mathbb{N}} \in \mathcal{A}_\Sigma(X_o,(T_n))$.
<u>Si</u> $(T_n)_{\mathbb{N}} \in \mathcal{C}_\Sigma(\alpha,\{(X_n)_{\mathbb{N}}\})$ <u>alors</u> $\lim_n \sup \|\sigma_n X_n\| \leq \alpha$ p.s. .
<u>En particulier si</u> α <u>peut être pris arbitrairement petit</u> : $\lim_n \|\sigma_n X_n\| = 0$ p.s.

Démonstration

L'hypothèse sur $(T_n)_{\mathbb{N}}$ permet d'écrire sur un ensemble $\Omega_o = \Omega$ p.s., pour $n \geq N(.,X)$ fini :

$$\|\sigma_{n+1} X_{n+1}\|^2 \leq \max\{\alpha_n^2 , (1+\beta_n)^2 \|\sigma_n X_n\|^2 - \gamma_n\} \tag{1.1 bis}$$

où $(\alpha_n)_{\mathbb{N}}$, $(\beta_n)_{\mathbb{N}}$, $(\gamma_n)_{\mathbb{N}}$ vérifie (1.2), (1.3) et 1.4).

Posons $B_n = \prod_{i=N(.,X)}^{n} (1+\beta_i)^2$ sur Ω_o et pour $n \geq N(.,X)$. On a $B_n > 0$ et $\lim_n B_n = B > 0$.

En itérant (1.1 bis) il vient :

$$\begin{aligned} \|\sigma_{n+1}X_{n+1}\|^2 &\leq \max\{\alpha_n^2 ,(1+\beta_n)^2 \max\{\alpha_{n-1}^2 ,(1+\beta_{n-1})\|\sigma_{n-1} X_{n-1}\| - \gamma_{n-1}\} - \gamma_n\} \\ &\leq \max\{\alpha_n^2,(1+\beta_n)^2\alpha_{n-1}^2,(1+\beta_n)^2(1+\beta_{n-1})^2\|\sigma_{n-1}X_{n-1}\|^2 - \gamma_{n-1}(1+\beta_n)^2 \gamma_n\} \\ &\leq \max\{\alpha_n^2,\dots,\alpha_k^2 B_n B_k^{-1},\dots,\alpha_N^2 B_n B_N^{-1}, B_N\|\sigma_N X_N\|^2 - \sum_{i=N}^{n} B_n B_i^{-1}\gamma_i\} \qquad (1.15) \end{aligned}$$

(en écrivant N pour N(.,X))

$$\leq \max\{\sup_{n\geq N} B_n . \sup_{n\geq N} \alpha^2_n . B_n^{-1} , B_N\|\sigma_N X_N\|^2\} \tag{1.16}$$

D'après les hypothèses (1.2), (1.3) et ce qui a été dit de $(B_n)_{\mathbb{N}}$ on voit que $(\|\sigma_n X_n\|)_{n\geq N}$ est une suite bornée sur Ω_o. On a alors $\sum \gamma_n = \infty$ d'après (1.4) et par suite $\sum B_n^{-1}\gamma_n = \infty$. En se reportant à (1.15) on peut alors écrire pour $N \leq n_o \leq n$:

$$\|\sigma_{n+1}X_{n+1}\|^2 \leq \max\{\sup_{k\geq n_o} B_k \cdot \sup_{k\geq n_o} \alpha_k^2 B_k^{-1},\ B_{n_o}\|\sigma_{n_o}X_{n_o}\|^2 - B_n \sum_{k=n_o}^{n} B_k^{-1}\gamma_k\} .$$

En faisant tendre n vers $+\infty$ il en résulte, compte-tenu de $\inf_{n\geq n_o} B_n > 0$ et de $\sum_{k=n_o}^{\infty} B_k^{-1}\gamma_k = \infty$, que $\lim_n \sup \|\sigma_{n+1} X_{n+1}\|^2 \leq \sup_{k\geq n_o} B_k \cdot \sup_{k\geq n_o} \alpha_k^2 B_k^{-1}$ (sur Ω_o) ;

il suffit de faire tendre n_o vers l'infini, de remarquer que $(B_b)_{\mathbb{N}}$ converge, pour voir que $\lim \sup_n \|\sigma_n X_n\|^2 \leq \alpha^2$ (sur Ω_o). La seconde partie de la proposition est évidente.

<u>Remarque</u> : le théorème et sa démonstration sont à rapprocher des lemmes 8 et 9 de Derman et Sacks et du lemme 10 de Venter donnés dans Wasan pages 185, 186 et 187 [16].

Les deux corollaires suivants constituent à proprement parlé les résultats de convergence au sens presque sûr des processus d'approximation stochastique.

<u>Corollaire 1.1.</u>(légère extension d'un théorème de Venter)

<u>Soient</u> X_o <u>une</u> v.a.H, $\Sigma = (\sigma_n)_{\mathbb{N}}$ <u>une suite d'applications de</u> H <u>dans</u> H, $(T_n)_{\mathbb{N}}$ <u>et</u> $(U_n)_{\mathbb{N}}$ <u>deux suites de</u> $\mathcal{S}_H(H)$, $\mathcal{C}$ <u>une suite croissante de sous-tribus de</u> $\mathcal{A}$.
<u>Posons</u> $X = (X_n)_{\mathbb{N}} \in \mathcal{A}_\Sigma(X_o, (T_n + U_n))$ <u>et supposons vérifiées les hypothèses suivantes</u> :

(i) $(T_n)_{\mathbb{N}} \in \mathcal{C}_\Sigma(\alpha, \{X\}, \mathcal{C})$,
(ii) $(U_n(X))_{\mathbb{N}}$ <u>est une suite de</u> v.a.H <u>qui vérifie</u> (1.5)<u>et</u>(1.6),
<u>alors on a</u> : $\lim_n \sup \|\sigma_n X_n\| \leq \alpha$ <u>p.s.</u> ,
<u>et en particulier, lorsque</u> α <u>est arbitrairement petit</u> :

$$\lim \|\sigma_n X_n\| = 0 \text{ p.s.} .$$

<u>Corollaire 1.2.</u>

<u>Soient</u> X_o, Σ, $(T_n)_{\mathbb{N}}$, $(U_n)_{\mathbb{N}}$ <u>et</u> $\mathcal{C}$ <u>comme au corollaire</u> 1.1. <u>et de plus</u> $(\lambda_n)_{\mathbb{N}} \in \mathcal{S}_R(H)$. <u>Posons</u> $X = (X_n) \in \mathcal{A}_\Sigma(X_o, (T_n + \lambda_n U_n))$.
<u>Supposons vérifiées les hypothèses suivantes</u>

(i) $(T_n)_{\mathbb{N}} \in \mathcal{D}_\Sigma(\alpha, \{X\}, \mathcal{C})$,
(ii) $(\lambda_n(X))_{\mathbb{N}}$ vérifie (1.12) (i),
(iii) $(U_n(X))_{\mathbb{N}}$ vérifie (1.12).
<u>Alors on a</u> : $\lim_n \sup \|\sigma_n X_n\| \leq \alpha$ p.s.,
<u>en particulier lorsque</u> α <u>est arbitrairement petit</u> :

$$\lim \|\sigma_n X_n\| = 0 \text{ p.s.} .$$

Le premier corollaire découle directement de la proposition 1.4 et du théorème 1.1. et le second de la proposition 1.5. et du théorème 1.1. .

Commentaire

L'étude de la convergence vers un point x^* de H se fait en posant $\sigma X_n = X_n - x^* = \sigma_n X_n$. Cependant un changement de variable $X'_n = X_n - x^*$ permettrait aussi de se ramener au cas où σ_n est l'application identique pour tout n. Plus généralement lorsque, pour tout n, σ_n est inversible, le changement de variable $X'_n = \sigma_n X_n$ ramène au cas où σ_n est l'application identique.

Ce n'est donc pas ce type de problème que nous avons en vue en introduisant Σ; dans la deuxième partie nous aurons à étudier des cas où σ_n n'est pas inversible.

1-4- CONVERGENCE EN MOYENNE DES PROCESSUS D'APPROXIMATION STOCHASTIQUE

Les résultats de convergence en moyenne sont moins nombreux et supposent des hypothèses plus fortes que ceux de convergence p.s. (1).

Des résultats généraux ont été obtenus par Schmetterer dans le cas où H est un espace de Hilbert séparable et $E^{\mathcal{G}_n} U_n = 0$ p.s. (pour tout n) et par Venter qui donne un énoncé avec la condition plus large : $\sum (E\|E^{\mathcal{G}_n} U_n\|^2)^{1/2} < \infty$.

Lemme 1.4.

Soit $(Y_n)_{\mathbb{N}}$ une suite de v.a.$\mathbb{R}_+$ vérifiant :

(1.16) $Y_n \to 0$ en probabilité,

(1.17) $\exists a_o \in \mathbb{R}_+, \exists n_o \in \mathbb{N}, \forall a > a_o, \forall n \geq n_o$:

$$EY_{n_o}^2 < \infty \text{ et } E(Y_{n+1} - a)_+^2 \leq (1+b_n)\ E(Y_n - a)_+^2 + r_n ,$$

où $b_n = b_n(a)$ et $r_n = r_n(a)$ sont dans $\mathbb{R}_+$,

(1.18) $\forall a > a_o$: $(b_n)_{\mathbb{N}} \in \ell_1$ et $(r_n)_{\mathbb{N}} \in \ell_1$

Pour tout $p \in]0,2[$ la suite $(Y_n)_{\mathbb{N}}$ converge alors vers zéro en moyenne d'ordre p (il suffit du reste pour cela que (1.17) ait lieu pour un $a > a_o$).

Si de plus b_n et r_n ne dépendent pas de a, alors (Y_n) converge en moyenne quadratique vers zéro

(1) Il s'agit évidemment toujours des processus de Dvoretzky.

Démonstration

L'inégalité (1.17) montre d'abord que $E\,Y_n^2 < \infty$ dès que $n > n_o$. En l'itérant il vient aussi pour $a > a_o$:

$$E(Y_n - a)_+^2 \leqslant \left[\prod_{i=n_o}^{n-1} (1+b_i)\right] . E(Y_{n_o} - a)_+^2 + \sum_{i=n_o}^{n-1} (r_i . \prod_{j=i+1}^{n-1} (1+b_j))$$

en convenant que le produit Π sur un ensemble vide vaut 1.

Posons $K(a) = K = \prod_{i \geqslant 0} (1+b_i(a)) < \infty$, $\sigma_m(a) = \sigma_m = \sum_{i \geqslant m} r_i < \infty$.

Il vient donc pour tout $n \geqslant n_o$:

$$E(Y_n - a)_+^2 \leqslant K(E(Y_{n_o} - a)_+^2 + \sigma_{n_o}) < \infty \qquad (1.19)$$

ce qui prouve déjà que $((Y_n - a)_+^p)_{\mathbb{N}}$ est équiintégrable pour $p \in]0,2[$. Ceci, joint à l'hypothèse (1.16) de l'énoncé, entraîne que $E(Y_n - a)_+^p \to 0$ (en utilisant le résultat classique selon lequel, pour une suite équiintégrable la convergence en probabilité équivaut à la convergence en moyenne - voir par exemple [8].).

Supposons maintenant que l'on ait K et σ_n indépendants de a.

Fixons a (a>0). Pour tout $b \geqslant \max\{a_o, a\}$ l'inégalité $E(Y_n - a)_+^2 \leqslant E(Y_n - b)_+^2 + 2(b-a)$ $E(Y_n - b)_+ + (b-a)^2\, P\{Y_n > a\}$ jointe à (1.19) entraîne que pour M entier compris entre n_o et n :

$$E(Y_n - a)_+^2 \leqslant K(E(Y_M - b)_+^2 + \sigma_M) + 2(b-a)\, E(Y_n - b)_+ + (b-a)^2\, P\{Y_n > a\} \qquad (1.20) .$$

Pour tout $\varepsilon > 0$ il est possible de fixer M assez grand pour que $K\sigma_M \leqslant \varepsilon/2$ (puisque $\sum r_i < \infty$), puis b assez grand pour que $K\, E(Y_M - b)_+^2 \leqslant \varepsilon/2$ (car $E\, Y_M^2 < \infty$).

La première partie de la démonstration prouve que $E(Y_n - b)_+ \to 0$ ($n \to \infty$) et l'hypothèse (1.16) entraîne que $P\{Y_n > a\} \to 0$ ($n \to \infty$). Il s'en suit que $\lim_n E(Y_n - a)_+^2 \leqslant \varepsilon$ (d'après 1.20)). Finalement en faisant tendre ε vers zéro on obtient : $\lim_n E(Y_n - a)_+^2 = 0$ pour tout $a > 0$ et donc : $\lim_n E\, Y_n^2 = 0$.

Théorème 2

Soient X_o une v.a.H, $\Sigma = (\sigma_n)_{\mathbb{N}}$ une suite d'applications de H dans H, $(T_n)_{\mathbb{N}}$ et $(U_n)_{\mathbb{N}}$ deux suites de $\mathcal{J}_H(H)$, $\mathcal{C}$ une suite croissante de sous-tribus de $\mathcal{Q}$. Posons $X = (X_n)_{\mathbb{N}} \in \mathcal{A}_{\Sigma}(X_o, (T_n + U_n))$ et supposons vérifiées les hypothèses suivantes :

(1.21) $(T_n)_{\mathbb{N}} \in \mathcal{E}_{\Sigma}(\alpha, \{X\}, \mathcal{C})$ avec α réel positif ;

(1.22) Dans la définition 1.1. relative à $(T_n)_{\mathbb{N}}$ l'inégalité (1.1) est vérifiée à partir d'un rang n_o (indépendant de ω) avec : α_n = cte p.s., β_n = cte p.s., $\beta_n \geqslant 0$, $(\beta_n)_{n \geqslant n_o} \in \ell_1$ p.s. ;

(1.23) $\sum E\|U_n(X)\|^2 < \infty$, $\sum (E\| E^{\mathcal{C}_n} U_n(X)\|^2)^{1/2} < \infty$;

(1.24) $E\|\sigma_{n_o} X_{n_o}\| < \infty$;

<u>On a alors</u> :

$$\limsup_n E\|\sigma_n X_n\|^2 \leqslant \alpha^2.$$

<u>En particulier quand</u> α <u>peut être choisi arbitrairement petit</u> $(\sigma_n X_n)_{\mathbb{N}}$ <u>converge fortement en moyenne quadratique vers zéro</u>.

<u>Démonstration</u>

Nous allons supposer que $\beta_n = 0$ p.s. pour $n > n_o$. En effet si ce n'était pas le cas on s'y ramènerait en posant sur un ensemble $\Omega_o = \Omega$ p.s. :

$$\sigma'_n X_n = \sigma_n X_n \Big(\prod_{n_o < i < n} (1+\beta_i)\Big)^{-1} \qquad \text{pour } n > n_o .$$

Soit alors un réel a, plus grand que α(p.s.). Du corollaire I.1. résulte que :

$$(\|\sigma_n X_n\| - a)_+ \to 0 \text{ p.s.} \quad \text{quand } n \to \infty ,$$

Soit $a > \alpha$. Nous allons montrer que $E(\|\sigma_n X_n\| - a)_+^2 \to 0$ quand $n \to \infty$, ce qui achèvera la démonstration du théorème en vertu de l'inégalité

$E(\|\sigma_n X_n\| - a)_+^2 + 2a.E(\|\sigma_n X_n\| - a)_+ + a^2 \geqslant E\|\sigma_n X_n\|^2$ où l'on fait $n \to \infty$ puis $a \downarrow \alpha$.

On a successivement :

$$\begin{aligned}
(\|\sigma_{n+1} X_{n+1}\| - a)_+ &\leqslant (\|\sigma_{n+1} X_{n+1}\| - \frac{a}{\max\{a, \|T_n\|\}}\|T_n\|)_+ \\
&\leqslant \|\sigma_{n+1} X_{n+1} - \frac{a}{\max\{a, \|T_n\|\}} . T_n\| \\
&\leqslant \|T_n + U_n - \frac{a}{\max\{a, \|T_n\|\}} T_n\| = \\
&= \|(\|T_n\| - a)_+ (\max\{a, \|T_n\|\})^{-1} . T_n + U_n\| ,
\end{aligned}$$

puis en élevant au carré et en développant :

$$(\|\sigma_{n+1} X_{n+1}\| - a)_+^2 \leqslant (\|T_n\| - a)_+^2 + 2\beta'_n (\|T_n\| - a)_+ + \|U_n\|^2 \quad (1.25)$$

où on a posé $\beta'_n = (\max\{a, \|T_n\|\})^{-1} . \langle T_n, U_n \rangle$.

Par ailleurs si n est assez grand pour que $a \geqslant \alpha_n$:

$$\begin{aligned}
(\|T_n\| - a)_+ &\leqslant (\|\sigma_n X_n\| - a - \gamma_n)_+ \text{ p.s.} \\
&\leqslant (\|\sigma_n X_n\| - a)_+
\end{aligned} \qquad (1.26)$$

(en utilisant (1.1.) avec $\beta_n = 0$).

En reportant (1.26) dans (1.25) puis en prenant l'espérance mathématique il vient (en se rappelant que T_n est un v.a.H. $\mathcal{A}_n$ mesurable) :

$$\begin{aligned}
E(\|\sigma_{n+1} X_{n+1}\| - a)_+^2 &\leqslant E(\|\sigma_n X_n\| - a)_+^2 + 2|E\,(E^{\mathcal{A}_n} \beta'_n .(\|T_n\| - a)_+)| + E\|U_n\|^2 \\
&\leqslant E(\|\sigma_n X_n\| - a)_+^2 + 2.\big(E(E^{\mathcal{A}_n}\beta'_n)^2 . E(\|\sigma_n X_n\| - a)_+^2\big)^{1/2} + E\|U_n\|^2
\end{aligned}$$

(en utilisant l'inégalité de Schwartz et (1.26))

$$\leq E(\|\sigma_n X_n\| -a)_+^2 \; (1+(E(\overset{\mathcal{G}_n}{E} \beta_n'^2))^{1/2}) + E\|U_n\|^2 + (E|\overset{\mathcal{G}_n}{E} \beta_n'|^2)^{1/2} \quad (1.27)$$

(où l'on a utilisé l'inégalité générale $x^{1/2} \leq \frac{1+x}{2}$ valable dès que $x \geq 0$).

Comme on le vérifie immédiatement on a : $|\overset{\mathcal{G}_n}{E} \beta_n'| \leq \|\overset{\mathcal{G}_n}{E} U_n\|$ p.s. .

Dans (1.27) on reconnait l'inégalité du lemme 1.4. , Plus précisément pour $Y_n = (\|\sigma_n X_n\| - \alpha)_+ (n>o)$

(1.16) est vérifié (Y_n est p.s. finie ; la convergence en probabilité résulte alors de la convergence p.s.)

(1.17) est vérifié en vertu du (1.24) et (1.27)

(1.18) est vérifié d'après (1.23) pour des coefficients ne dépendant pas de a.

Il s'en suit bien que Y_n et a forciori $(\|\sigma_n X_n\| -a)_+$ tend vers zéro en moyenne quadratique.

R E F E R E N C E S

(1) - BLUM J.R. (1954) Approximation methods wich converge with probability one. *Ann. Math. Statis.* 25 382-386

(2) - BLUM J.R. (1954) . Multidimensional stochastic approximation methods. *Ann. Math. Statist.* 25 737-744

(3) - CHUNG, K.L. (1954). On a stochastic approximation method. *Ann. Math. Statist.* 25.463-483.

(4) - DERMAN, C. and SACKS,J. (1959). On Dvoretzky's stochastic approximation theorem *Ann. Math. Statist.* 30. 601-605.

(5) - DRIML, M. and HANS, O (1959). Conditional expectations for genreralized random variables. Transactions of the Second Prague Conference on Information Theory, Statistical Decision Functions and Random Processes. 123-143.

(6) - DVORETZKY, A. (1956) On stochastic approximation. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 1 39-55

(7) - LEOVE, M. (1960). Probability Theory (2nd Ed.) *Van Nostrand, Princeton.*

(8) - MEYER, P.A. (1966) Probabilités et Potentiel. *Hermann.*

(9) - MOURIER, E. (1952-53). Eléments aléatoires dans un espace de Banach. *Ann. Inst. H. Poincaré* 13 161-244.

(10)- NEVEU, J. (1972). Martingales à temps discret. *Masson et Cie - Paris -.*

(11)- HENNEQUIN P.L. et A. TORTRAT (1965). Théorie des probabilités et quelques applications. - *Masson et Cie - Paris -*

(12)- SCHMETTERER, L. (1958). Sur l'itération stochastique. *Le calcul des Probabilités et ses Applications.* 87 55-63.

(13)- SCHMETTERER, L. Cours A.E.A.-D.E.A. . Math. Appliquées. Clermont-Ferrand 2ème édition 1972.

(14)- VENTER, J.H. (1966). On Dvoretzky stochastic approximation theorems. *Ann. Math. Statist.* 37 181-190.

(15)- WASAN, M.T. (1969). Stochastic approximation (*Cambridge University Press*).

(16)- WOLFOWITZ, J. (1956). On stochastic approximation methods. *Ann. Math. Statist.* 27 1151-1156.

2ème PARTIE - PROCESSUS GENERALISES DE ROBBINS-MONRO

2.0. INTRODUCTION ET RESUME

Soient (Ω, G, P) un espace de probabilité, H un espace de Hilbert séparable et l'équation $M(x) = 0$ $(M : H \to H)$ ou, ce qui revient au même quand M est monotone et continue au voisinage d'une solution x^* :

$(\mathcal{P}1)$ Trouver $x^* \in H$ tel que $\langle M(x), x - x^* \rangle \geq 0$ pour tout x de H.

Nous envisagerons plus généralement le problème $(\mathcal{P}2)$ suivant :

$(\mathcal{P}2)$ Trouver $x^* \in C$ tel que $\langle M(x), x - x^* \rangle \geq 0$ pour tout $x \in C$, où C est un (ensemble) convexe fermé de H.

Les méthodes d'approximations stochastiques s'introduisent pour la résolution de ces problèmes lorsque $M(x)$ n'est pas explicitement connue, mais est l'espérance d'un variable aléatoire $g(x,.)$ (ceci pour tout $x \in H$) sur laquelle il est possible de procéder à des observations. Nous supposerons que $E\|g(x,.)\| < \infty$ pour tout $x \in H$ afin que $M(x) = E\, g(x,.)$ ait effectivement un sens.

La situation que nous venons de décrire se rencontre fréquemment lorsque l'on veut optimiser un critère en moyenne ou lorsque la grandeur $M(x)$ est altérée par une erreur ou un bruit aléatoire.

Les processus de Robbins-Monro [7] permettent de construire des systèmes d'itérations stochastiques qui convergent p.s. ou en moyenne quadratique vers les solutions des problèmes $(\mathcal{P}1)$ et $(\mathcal{P}2)$. Cependant il sera nécessaire de les étudier sous des formes généralisées pour obtenir des résultats assez forts et généraux pour s'adapter aux divers cas concrets.

La forme générale des processus de Robbins-Monro est la suivante,

$$(2.0)\quad \begin{cases} X_0 \text{ v.a.H. arbitraire,} \\ \forall\omega\in\Omega,\ \forall n\in\mathbb{N} : X_{n+1}(\omega) = X_n(\omega) - a_n \cdot Y_n(X_0(\omega),\dots, X_n(\omega),\omega) \\ (a_n)_{\mathbb{N}} \text{ suite de réels positifs.} \end{cases}$$

L'objet de cet exposé est de donner un théorème général et diverses méthodes pour la résolution de ($\mathcal{P}$1) et de ($\mathcal{P}$2) qui apparaissent comme des corollaires du théorème.

Ce théorème général nous permet d'éliminer l'hypothèse d'unicité de solution de ($\mathcal{P}$1) ou de ($\mathcal{P}$2), contrairement à ce qui est classiquement fait.

Il s'avère d'un usage commode pour la méthode de fonctions de pénalité ainsi que pour l'étude des processus de Robbins-Monro projetés utilisés lors de la résolution de ($\mathcal{P}$2).

2.1. UN THEOREME FONDAMENTAL

Soient, pour tout $n \geq 0$, Y_n une application de $H^{n+1} \times \Omega$ dans H, σ_n une application de H dans H et X_n une variable aléatoire à valeurs dans H (v.a.H.). Nous adopterons les notations :

$$\sigma_n(X_n(.)) = \sigma_n X_n \ , \ Y_n(X_0(.), X_1(.), \dots, X_n(.), .) = Y_n(X)$$

et introduirons les processus de Robbins-Monro sous la forme généralisée suivante :

$$(2.1) \quad \begin{cases} X_0 \text{ v.a.H.} \ , \ \Sigma = (\sigma_n)_{\mathbb{N}} \ , \ \sigma_n \text{ est mesurable pour tout } n, \\ \sigma_{n+1} X_{n+1} = \sigma_n X_n - a_n \cdot Y_n(X) \ , \text{ où } a_n \in \mathbb{R}_+ \text{ et} \\ Y_n(X) \text{ v.a.H.} \ , \ n = 0,1,2,\dots \ . \end{cases}$$

Il est clair que si les applications σ_n sont inversibles, un changement de variable $X'_n = \sigma_n X_n$ permettrait de se ramener à la forme (2.0). Mais dans la suite σ_n sera toujours mesurable et souvent non inversible. Ainsi, C_0 étant un ensemble convexe fermé de H, π_{C_0} le projecteur sur C_0, nous poserons $\sigma_n X_n = \sigma X_n = X_n - \pi_{C_0} X_n$; nous définirons une application $\sigma (= \sigma_n)$ continue (et donc mesurable), non inversible en général.

Voici un théorème général sur les processus (2.1) et qui contient en les renforçant ou en les généralisant beaucoup de résultats classiques.

Théorème 2.1.

Soit $(X_n)_{\mathbb{N}}$ une suite de v.a.H. définie par (2.1), $(\mathcal{A}_n)_{\mathbb{N}}$ une suite croissante de sous-tribus de $\mathcal{A}$ à laquelle est adaptée la suite $(\sigma_n X_n)_{\mathbb{N}}$ de v.a.H. Considérons les hypothèses suivantes :

(2.2) $\forall n \in \mathbb{N}$, $\exists c_n$ et $d_n \in \mathbb{R}_+$: $E^{\mathcal{A}_n} \| Y_n(X) \|^2 \leq c_n + d_n \cdot \| \sigma_n X_n \|^2$ p.s., (2.3) $\exists\, \alpha > 0,\ \forall r \in [\alpha, \infty[,\ \exists\, (e_n(\alpha, r) > 0,\ c'_n$ <u>et</u> $d'_n \in \mathbb{R}_+)$:

$$\|\sigma_n X_n\| \in [\alpha, r] \Longrightarrow \langle \sigma_n X_n , M_n(X)\rangle \geq e_n(\alpha, r) - c'_n - d'_n . \|\sigma_n X_n\|^2 \ \underline{\text{p.s.}} ,$$

<u>où on a posé</u> $M_n(X) = E^{\mathcal{A}_n} Y_n(X)$ p.s. ,

(2.4) $\Sigma c_n . a_n^2 < \infty,\ \Sigma d_n . a_n^2 < \infty,\ \Sigma c'_n . a_n < \infty,$

$\Sigma d'_n . a_n < \infty,\ \Sigma e_n(\alpha, r) . a_n = \infty\ (\forall r \in [\alpha, \infty[),$

(2.2 bis) $\quad E\,\|\sigma_0 X_0\|^2 < \infty$

et $\sum a_n^2\, E\,\|Y_n(X) - M_n(X)\|^2 < \infty$.

<u>Sous les hypothèses</u> (2.2), (2.3) <u>et</u> (2.4) <u>on a :</u>

$$\limsup_n \|\sigma_n X_n\| \leq \alpha \ \underline{\text{p.s.}} ;$$

<u>Si, de plus,</u> (2.2 bis) <u>est vérifiée alors :</u>

$$\limsup_n E\,\|\sigma_n X_n\|^2 \leq \alpha^2 .$$

<u>Démonstration</u>

Posons $T_n(X) = \sigma_n X_n - a_n . M_n(X)$ et $Z_n(X) = a_n(M_n(X) - Y_n(X))$, de sorte que l'on peut écrire le processus (2.1) sous la forme de Dvoretzky généralisée : $\sigma_{n+1} X_{n+1} = T_n(X) + Z_n(X)$.

<u>1er Point</u> : $(T_n)_{\mathbb{N}} \in \mathscr{C}_{\Sigma}(\alpha, \{X\}, (\mathcal{A}_n)_{\mathbb{N}})$ (voir 1$^{\text{ère}}$ partie : Déf. 1.1).

Tout d'abord, nous remarquerons que $T_n(X)$ est $\mathcal{A}_n$-mesurable en tant que somme de deux v.a. $\mathcal{A}_n$-mesurables : $\sigma_n X_n$ et $- a_n . M_n(X)$.

Par ailleurs on a, dès que $\|\sigma_n X_n\| \in [\alpha, r]$:

$$\| T_n(X) \|^2 = \| \sigma_n X_n \|^2 - 2 \cdot a_n \cdot \langle \sigma_n X_n , M_n(X) \rangle + a_n^2 \| M_n(X) \|^2$$

$$\leq (1 + a_n^2 \cdot d_n + a_n d'_n) \cdot \| \sigma_n X_n \|^2 + a_n^2 \cdot c_n + a_n \cdot c'_n - e_n(\alpha, r) \quad (2.5)$$

Posons alors : $\alpha_n^* = \alpha + a_n \sqrt{c_n + d_n \cdot \alpha^2}$, $\gamma_n^* = e_n(\alpha, \max\{\alpha, \| \sigma_n X_n \| \}) \cdot a_n$.

$(1 + \beta_n^*)^2 = 1 + a_n^2 d_n + a_n d'_n + (a_n^2 c_n + a_n c'_n) \cdot \alpha^{-2}$ avec $\beta_n^* \geq 0$.

* Si $\| \sigma_n X_n \| \leq \alpha$ on a : $\| T_n \| \leq \| \sigma_n X_n \| + a_n \cdot \| M_n X \| \leq \alpha_n^*$, puisque de (2.2) résulte immédiatement que $\| M_n(X) \|^2 \leq c_n + d_n \cdot \| \sigma_n X_n \|^2$.

* Si, au contraire, $\| \sigma_n X_n \| > \alpha$ il résulte de (2.5) et de $\alpha^{-2} \| \sigma_n X_n \|^2 \geq 1$ que $\| T_n(X) \| \leq (1 + \beta_n^*)^2 \cdot \| \sigma_n X_n \|^2 - \gamma_n^*$.

Finalement la suite vérifie l'inégalité :

$$\| T_n(X) \|^2 \leq \max \{\alpha_n^{*2} , (1 + \beta_n^*)^2 \| \sigma_n X_n \|^2 - \gamma_n^*\} ,$$

dans laquelle

* la suite $(\alpha_n^*)_{\mathbb{N}}$ converge vers α (car $a_n^2 (c_n + d_n \alpha^2) \to 0$ d'après (2.4))

* la suite $(\beta_n^*)_{\mathbb{N}}$ vérifie que $\prod_{n=n_0}^{p} (1 + \beta_n^*) = s_p > 0$ (pour n_0 assez grand et $p \geq n_0$) et s_n converge vers un réel strictement positif ; Ceci résulte facilement de (2.4) ; On a donc $(\beta_n^*)_{\mathbb{N}} \in \ell_\pi$ conformément aux notations de la première partie.

* la suite $(\gamma_n^*)_{\mathbb{N}}$ est positive et vérifie d'après (2.4) que $\Sigma \gamma_n^* = \infty$ pour toute suite $(\| \sigma_n X_n \|)_{\mathbb{N}}$ bornée.

Les conditions de la définition de $\underset{\Sigma}{\mathscr{C}} (\alpha, \{X\}, (\mathcal{A}_n)_{\mathbb{N}})$ sont donc vérifiées.

<u>2ème Point</u> : $(T_n)_{\mathbb{N}} \in \mathscr{D}_\Sigma (\alpha, \{X\}, (\mathcal{A}_n)_{\mathbb{N}})$.

Pour établir ce deuxième point, il suffit, conformément à la définition

1.2 de $\mathcal{L}_\Sigma$ (première partie), de montrer que :

$\exists$ k et $\rho > 0$, $\exists\, M \in \mathbb{N}$, $\forall n \geq M$: $\|\sigma_n X_n\| \geq \rho \Rightarrow \|T_n(X)\| \geq k \cdot \|\sigma_n X_n\|$.

A cette fin, remarquons que $\|T_n\|^2 \geq \|\sigma_n X_n\|^2 - 2 \cdot a_n \cdot \|M_n(X)\| \cdot \|\sigma_n X_n\|$

$$\geq \|\sigma_n X_n\|^2 \left(1 - 2 a_n \frac{\|M_n(X)\|}{\|\sigma_n X_n\|}\right) \geq \|\sigma_n X_n\|^2 \left(1 - 2 a_n \sqrt{\frac{c_n}{\rho^2} + d_n}\right),$$

si $\|\sigma_n X_n\| \geq \rho > 0$. On peut alors déterminer N assez grand pour que

$n \geq N \Rightarrow 2 a_n \sqrt{\frac{c_n}{\rho^2} + d_n} \leq \frac{1}{2}$ (vu que $a_n^2 (c_n . \rho^{-2} + d_n) \to 0$ d'après (2.4)).

Ainsi $n \geq N$, $\|\sigma_n X_n\| \geq \rho \Rightarrow \|T_n\|^2 \geq \frac{1}{2} \|\sigma_n X_n\|^2$.

<u>3ème point</u> : propriétés de $(Z_n)_{\mathbb{N}}$.

Posons $\lambda_n^2(X) = c_n + d_n \|\sigma_n X_n\|^2$ et $Z_n(X) = \lambda_n(X) \cdot U_n(X)$.

On a ainsi : $E^{\mathcal{A}_n} U_n = 0$ p.s. et $E \|U_n\|^2 = E\, \lambda_n^{-2}(X)\, E^{\mathcal{A}_n} \|Z_n(X)\|^2$

$\leq a_n^2\, E\, (\lambda_n^{-2}(X) \cdot E^{\mathcal{A}_n} \|Y_n(X)\|^2) \leq a_n^2$ p.s. d'après (2.2).

Par suite $\Sigma\, c_n \cdot E \|U_n(X)\|^2 < \infty$ et $\Sigma\, d_n \cdot E \|U_n(X)\|^2 < \infty$.

Si de plus (2.2 bis) est vérifiée alors : $E^{\mathcal{A}_n} Z_n(X) = 0$ p.s. (déjà vérifiée précédemment) et $\Sigma\, E\, (\|Z_n(X)\|^2) < \infty$.

<u>4ème point</u> : Conclusion de la démonstration.

Sous les hypothèses (2.2), (2.3) et (2.4) on applique le corollaire 1.2 (première partie) et si de plus (1.3 bis) alors on applique le corollaire 1.1 et le théorème 1.2 (première partie) ce qui démontre le théorème 2.1 ci-dessus.

<u>Remarque</u> : Bien entendu si α peut être choisi arbitrairement petit, les deux limites supérieures de l'énoncé se transforment en limites simples :

* $\|\sigma_n X_n\| \to 0$ p.s.,
* $E \|\sigma_n X_n\|^2 \to 0$.

2.2. APPLICATION A LA RESOLUTION DE ($\mathcal{P}$2) PAR LE PROCESSUS DE ROBBINS-MONRO PROJETE

Soit pour tout x de H, g(x,.) une v.a.H. de norme intégrable. Posons $M(x) = E(g(x,.))$.

Nous allons construire un algorithme pour la résolution du problème ($\mathcal{P}$2) suivant :

$$(\mathcal{P}2) \quad \begin{cases} \text{Trouver } x^* \in C, \text{ C convexe fermé de H,} \\ \text{Tel que : } \langle x - x^*, M(x)\rangle \geq 0 \text{ pour tout x de C.} \end{cases}$$

Soit C_0 l'ensemble convexe fermé des solutions de ($\mathcal{P}$2). Nous allons plus précisément construire un algorithme permettant d'approcher C_0 (toujours supposé non vide).

Nous désignerons dans la suite par π_C (resp. π_{C_0}) le projecteur sur C (resp. sur C_0) et considèrerons les hypothèses :

(2.6) $\quad \forall\, x \in C\ ,\ E\ \|g(x,.)\|^2 \leq c + d\ .\ \|x - \pi_{C_0} x\|^2,$

où c et d sont deux réels positifs,

(2.6 bis) $\quad \forall\, x \in C,\ E\ \|g(x,.) - M(x)\|^2 \leq K < \infty\ ,$

(2.7) $\quad \exists\ \alpha > 0\ ,\ \forall\ r \in [\alpha,\infty[,\ \exists\ e(\alpha,r) > 0\ ,\ \forall\ x \in C :$

$$\|x - \pi_{C_0} x\| \in [\alpha,r] \Rightarrow \langle x - \pi_{C_0} x\ ,\ M(x)\rangle \geq e(\alpha,r)\ ,$$

ainsi que la suite de v.a.C :

$$(2.8) \quad \begin{cases} X_0 \text{ v.a.C arbitraire} \\ X_{n+1} = \pi_C\,(X_n - a_n\, Y_n(X)) \qquad n = 0,1,2,\ldots \\ \text{où : } a_n \geq 0\ ,\ \Sigma\, a_n = \infty\ ,\ \Sigma\, a_n^2 < \infty, \\ Y_n(X) \text{ v.a.H qui pour } X_n = x \text{ a la même distribution que } g(x,.). \end{cases}$$

Remarquons que <u>si</u> M <u>est fortement monotone</u>, (c'est-à-dire : $<Mx - My , x - y> \geq k \, . \, \|x - y\|^2$, $\forall (x,y) \in C^2$, avec $k > 0$) l'hypothèse (2.7) est manifestement vérifiée pour tout $\alpha > 0$.

<u>Corollaire 1</u>

<u>Sous les hypothèses (2.6) et (2.7) le processus défini en (2.8) vérifie que :</u>

$$\limsup_n \|X_n - \pi_{C_0} X_n\| \leq \alpha \quad \text{p.s.}$$

<u>et si de plus (2.6 bis) est vérifié alors</u> :

$$\limsup_n E \, \|X_n - \pi_{C_0} X_n\|^2 \leq \alpha^2.$$

<u>Démonstration</u>

Nous poserons pour tout n : $\mathcal{A}_n$ tribu engendré par $(X_i)_{i \leq n}$ et

$\sigma_n X_n = X_n - \pi_{C_0} X_n = \sigma X_n$ (σ est continue et donc mesurable)

$a_n Y'_n(X) = X_n - \pi_C [X_n - a_n Y_n(X)] + \pi_{C_0} X_{n+1} - \pi_{C_0} X_n$

de telle manière que :

$\sigma_{n+1} X_{n+1} = \sigma_n X_n - a_n \, . \, Y'_n(X)$.

Montrons alors que le théorème 2.1 s'applique.

<u>Vérification de (2.2)</u>

$a_n^2 E^{\mathcal{A}_n} \|Y'_n(X)\|^2 = E^{\mathcal{A}_n} \|X_n - X_{n+1} + \pi_{C_0} X_{n+1} - \pi_{C_0} X_n\|^2$ p.s.

$\leq 4 \, E^{\mathcal{A}_n} \|X_{n+1} - X_n\|^2$ p.s. (en vertu de l'inégalité générale $\|x + y\|^2 \leq 2 \, \|x\|^2 + 2 \, \|y\|^2$ et de la propriété de contraction des projecteurs : $\|\pi_{C_0} x - \pi_{C_0} y\| \leq \|x - y\|$)

$\leq 4 \, . \, a_n^2 E^{\mathcal{A}_n} \|Y_n(X)\|^2$ p.s. (en appliquant encore la propriété de contraction de π_{C_0} : $\|\pi_{C_0}(x + y) - x\| \leq \|y\|$ si $x \in C_0$)

$$\leq 4 \,.\, a_n^2 \,(c + d \,.\, ||\sigma X_n||^2) \quad \text{p.s.}$$

(compte tenu de (2.6) et de la définition de $Y_n(X)$).

Finalement en posant $c_n = 4c$, $d_n = 4c$ nous obtenons bien l'inégalité (2.2) du théorème.

Vérification de (2.3)

Nous nous appuierons ici sur le lemme suivant :

Lemme 2.1. - Soit C un convexe fermé de H.

$\forall\, (x,y) \in H^2$, $\forall\, (u,v) \in C^2$ ont lieu les inégalités suivantes :

(i) $<x - \pi_C\, x, \pi_C\, x - \pi_C\, y> \leq ||x-y||^2$,

(ii) $<x - \pi_C\, x, u - v> \geq - ||x - u||^2$.

Démonstration du lemme

Posons $\mu(x,y) = <x - \pi_C\, x - \pi_C\, y>$. De la définition du projecteur résulte que $\mu(x,y) \geq 0$. Il vient alors :

$$0 \leq \mu(x,y) \leq \mu(x,y) + \mu(y,x) = <x-y \,,\, \pi_C x - \pi_C y> - ||\pi_C x - \pi_C y||^2$$

ce qui implique (i).:

$$<x - \pi_C x \,,\, u - v> = <x - \pi_C x \,,\, u - \pi_C x> + <x - \pi_C x \,,\, \pi_C x - v>$$

$$= - \mu(x,u) + \mu(x,v) \geq - ||x - u||^2 .$$

Revenons à la vérification de (2.3) et posons : $E^{\mathcal{a}_n}\, Y_n'(X) = M_n'(X)$

$$a_n \,.\, <M_n'(X) \,,\, \sigma X_n> = E^{\mathcal{a}_n} <a_n\, Y_n(X) \,,\, \sigma X_n>$$

$$+ E^{\mathcal{a}_n} <X_{n+1}' - \pi_C\, X_{n+1}' \,,\, \sigma X_n> + E^{\mathcal{a}_n} <\pi_{C_0}\, X_{n+1} - \pi_{C_0}\, X_n \,,\, \sigma X_n> \quad \text{p.s.}$$

(où $X_{n+1}' = X_n - a_n\, Y_n(X)$).

$$\text{p.s.} \geq a_n \cdot \langle M_n(X), \sigma X_n \rangle - E^{\mathcal{A}_n} \| X'_{n+1} - X_n \|^2 - E^{\mathcal{A}_n} \| X_{n+1} - X_n \|^2$$

(où on a utilisé (ii) relatif à C et (i) relatif à C_0) ,

$$\text{p.s.} \geq a_n \cdot \langle M_n(X), \sigma X_n \rangle - 2\, a_n^2 \cdot E^{\mathcal{A}_n} \| Y_n(X) \|^2$$

$$\text{p.s.} \geq a_n \cdot \langle M_n(X), \sigma X_n \rangle - 2\, a_n^2 \cdot (c + d \, \| \sigma X_n \|^2)$$

Si $\| \sigma X_n \| \in [\alpha, r]$ alors de (2.7) résulte que :

$$\langle M'_n(X), \sigma X_n \rangle \geq e(\alpha, r) - c'_n - d'_n \cdot \| \sigma X_n \|^2 \text{ p.s.}$$

où $c'_n = 2\, a_n c$ et $d'_n = 2\, a_n d$.

<u>Vérification de (2.2 bis)</u>

$$a_n^2\, E^{\mathcal{A}_n} \| Y'_n(X) - M'_n(X) \|^2 \leq 2.E^{\mathcal{A}_n} \| \pi_C\, X'_{n+1} - E^{\mathcal{A}_n} \pi_C\, X'_{n+1} \|^2 +$$

$$+ 2.E^{\mathcal{A}_n} \| \pi_{C_0}\, X_{n+1} - E^{\mathcal{A}_n} \pi_{C_0}\, X_{n+1} \|^2 \text{ p.s.}$$

$$\leq 2.E^{\mathcal{A}_n} \| X'_{n+1} - E^{\mathcal{A}_n} X'_{n+1} \|^2 + 2\, E^{\mathcal{A}_n} \| X_{n+1} - E^{\mathcal{A}_n} X_{n+1} \|^2 \text{ p.s.}$$

$$\leq 4.E^{\mathcal{A}_n} \| X'_{n+1} - E^{\mathcal{A}_n} X'_{n+1} \|^2 \text{ p.s.}$$ (en vertu de l'inégalité générale, valable pour toute v.a.H Z, toute sous tribu $\mathcal{B}$ de $\mathcal{A}$ et tout convexe fermé C de H :

$$E^{\mathcal{B}} \| \pi_C\, Z - E^{\mathcal{B}} \pi_C\, Z \|^2 \ (\leq E^{\mathcal{B}} \| \pi_C\, Z - \pi_C\, E^{\mathcal{B}} Z \|^2) \leq E^{\mathcal{B}} \| Z - E^{\mathcal{B}} Z \|^2 \text{ p.s.})$$

$$\leq 4\, a_n^2 \cdot E^{\mathcal{A}_n} \| Y_n(X) - M_n(X) \|^2 \text{ p.s.}$$

Si (2.6 bis) est vérifié, nous avons donc :

$$\Sigma\, a_n^2 \cdot E^{\mathcal{A}_n} \| Y_n(X) - M_n(X) \|^2 \leq 4\, K \cdot \Sigma\, a_n^2 \text{ p.s.} < \infty \text{ p.s.} .$$

<u>Vérification de (2.4)</u>

Nous avons successivement :

$\Sigma\, a_n^2\, c_n = 4\, c \,.\, \Sigma\, a_n^2 < \infty \,,\ \Sigma\, a_n^2\, d_n = 4\, d \,.\, \Sigma\, a_n^2 < \infty \,,$

$\Sigma\, a_n\, c'_n = 2\, c \,.\, \Sigma\, a_n^2 < \infty \,,\ \Sigma\, a_n\, d'_n = 2\, d \,.\, \Sigma\, a_n^2 < \infty \,,$

$\Sigma\, e_n(\alpha,r)\, a_n = e(\alpha,r)\, \Sigma a_n = \infty \,.$

Les conditions d'application du théorème 2.1 sont donc réunies. Le corollaire 1 en découle immédiatement.

Commentaires

Nous avons en fait rassemblé dans le corollaire 1 les deux cas (fréquemment distingués) $C = H$ et $C \neq H$. Lorsque $C = H$ la formulation se simplifie puisque π_C est alors l'opérateur identique.

Nous avons également présenté dans ce corollaire le cas général où C_0 est un convexe non vide, sans introduire l'hypothèse d'unicité de solution. L'intérêt du point de vue choisi pour présenter les résultats apparaît ici clairement : les généralisations sont absorbées sans difficulté particulière.

2.3. APPLICATION A LA RESOLUTION DE ($\mathcal{P}$2) PAR LA METHODE DES FONCTIONS DE PENALITE

Soit pour tout x de H et tout $n \in \mathbb{N}$ $g_n(x,.)$ une v.a.H de norme intégrable. Posons $M_n(x) = E\,(g_n\,(x,.))$.

Désignons par Γ_n l'ensemble des solutions du problème ($\mathcal{P}$1) relatif à M_n : $\Gamma_n = \{x^*: < M_n\, x,\, x-x^*> \geq 0 \quad \forall x \in H\}$.

Définissons aussi, pour les ensembles convexes fermés de H, une distance δ définie par :

$$\delta\,(A,B) = \sup\,\{\,||\pi_A\, x - \pi_B\, x\,|| : x \in H\,\}\,.$$

La méthode des fonctions de pénalité consiste à introduire une famille P_n d'application de H dans H telle que, posant $M_n = M + P_n$, $g_n = g + P_n$ on ait : $\Gamma_n \to C_0$ au sens de δ (ou d'une autre distance).

L'intérêt des processus de Robbins Monro relativement à ce type de méthode tient au corollaire 2 ci-dessous :

Introduisons les hypothèses :

(2.9) $\quad \forall x \in H : E\,||g_n\,(x,.)\,||^2 \leq c_n^* + d_n^*\,||x - \pi_{\Gamma_n}\,||^2\,.$

avec $\quad c_n^*$ et $d_n^* \in \mathbb{R}_+$ et $n = 0,1,2,\ldots,$

(2.10) $\quad \exists\, \alpha > 0,\ \forall r \in [\alpha, \infty[,\ \exists\, e_n\,(\alpha, r) > 0,$

$$||x - \pi_{\Gamma_n}\, x\,|| \in [\alpha, r] \Rightarrow < x - \pi_{\Gamma_n}\, x,\, M_n(x) > \geq e_n\,(\alpha,r)$$

$(n = 0,1,2,\ldots)$

(2.11) $\quad$ Pour tout n $\Gamma_n \neq \phi$ et $\Sigma\, \delta\,(\Gamma_n, \Gamma_{n+1}) < \infty$

ainsi que la suite $(X_n)_{\mathbb{N}}$ suivante :

$$\begin{cases} X_0 \text{ v.a.H arbitraire} \\ X_{n+1} = X_n - a_n \,.\, Y_n(X) \end{cases}$$

où : $Y_n(X)$ a la même distribution que $g_n(x,.)$ quand $X_n = x$ et :

$$\Sigma\, a_n^2\, c_n^* < \infty, \quad \Sigma\, a_n^2\, d_n^* < \infty\,, \quad \Sigma\, e_n\,(\alpha,\, r)\, a_n = \infty \qquad (\forall r \geq \alpha)\ ;$$

Introduisons encore l'hypothèse :

(2.9 bis) $\quad E\, \| g_n(x,.) - M_n(x)\, \|^2 \leq K_n$ et $\Sigma\, a_n^2\, K_n < \infty$.

Corollaire 2

Sous les hypothèses (2.9), (2.10), (2.11) et (2.12) on a :

$$\lim_n \sup \| X_n - \pi_{\Gamma_n}\, X_n \| \leq \alpha \quad \text{p.s.}$$

et si de plus (2.9 bis) est vérifiée alors :

$$\lim_n \sup E\, \| X_n - \pi_{\Gamma_n}\, X_n\, \|^2 \leq \alpha^2.$$

Démonstration

Nous poserons $\sigma_n\, X_n = X_n - \pi_{\Gamma_n}\, X_n$, $\mathcal{A}_n$ la tribu engendrée par $(X_i)_{i \leq n}$, $a_n\, Y'_n\,(X) = a_n\, Y_n\,(X) + \pi_{\Gamma_{n+1}}\, X_{\Gamma_{n+1}} - \pi_{\Gamma_n} X_n$, $M'_n(X) = E^{\mathcal{A}_n} Y'_n(X)$ $(n = 0,1,2,\dots)$.

Nous avons encore : σ_n application continue, donc mesurable, et

$$\sigma_{n+1}\, X_{n+1} = \sigma_n\, X_n - a_n\, Y'_n\,(X) \qquad (n = 0,1,2,\dots)\ .$$

Le corollaire 2 est une conséquence simple du théorème 2.1.

Vérification de (2.2)

$$a_n^2 .\, E^{\mathcal{A}_n}\, \| Y'_n\,(X)\, \|^2 \leq 2\, a_n^2 .\, E^{\mathcal{A}_n}\, \| Y_n\,(X)\, \|^2 + 2.E^{\mathcal{A}_n}\, \| \pi_{\Gamma_{n+1}}\, X_{n+1} - \pi_{\Gamma_n}\, X_n \|^2 \text{ p.s.}$$

$$\text{p.s.} \leq 2\, a_n^2 \,.\, E^{\mathcal{A}_n}\, \| Y_n(X) \|^2 + 4.E^{\mathcal{A}_n} \| \pi_{\Gamma_{n+1}}\, X_{n+1} - \pi_{\Gamma_n}\, X_{n+1} \|^2$$

$$+\ 4.E^{\mathcal{A}_n}\, \| \pi_{\Gamma_n}\, X_{n+1} - \pi_{\Gamma_n}\, X_n\, \|^2$$

$$\text{p.s.} \leq 2\, a_n^2\, E^{\mathcal{A}_n} \| Y_n\,(X) \|^2 + 4.\delta^2(\Gamma_n,\, \Gamma_{n+1}) + 4\, E^{\mathcal{A}_n}\, \| X_{n+1} - X_n\, \|^2$$

(en se rappelant la définition de δ et la propriété de contraction de π_C)

$$\text{p.s.} \leq 6\, a_n^2 .\ E^{\mathcal{A}_n} \| Y_n\ (X) \|^2 + 4.\ \delta^2(\Gamma_n\ ,\ \Gamma_{n+1}).$$

Finalement, il vient en utilisant (2.9) :

$$E^{\mathcal{A}_n} \| Y'_n\ (X) \|^2 \leq c_n + d_n \|\sigma_n\ X_n\|^2 \text{ p.s.} \quad ,$$

où $\qquad c_n = (6\, c_n^* + 4\, a_n^{-2}\ \delta^2(\Gamma_n\ ,\ \Gamma_{n+1}))$ et $d_r = 6\, d_n^*$.

Vérification de (2.3)

$$< \sigma_n\ X_n,\ M'_n\ (X) > = < \sigma_n\ X_n,\ M_n(X) > + a_n^{-1} < \sigma_n X_n, E^{\mathcal{A}_n} \pi_{\Gamma_{n+1}} X_{n+1} - \pi_{\Gamma_n} X_n > \text{ p.s.}$$

$$\text{p.s.} \geq < \sigma_n\ X_n,\ M_n(X) > + a_n^{-1}\ E^{\mathcal{A}_n} < \sigma_n\ X_n,\ \pi_{\Gamma_{n+1}}\ X_{n+1} - \pi_{\Gamma_n}\ X_{n+1} >$$

$$+ a_n^{-1}\ E^{\mathcal{A}_n} < \sigma_n\ X_n,\ \pi_{\Gamma_n}\ X_{n+1} - \pi_{\Gamma_n}\ X_n >$$

$$\text{p.s.} \geq < \sigma_n\ X_n,\ M_n\ (X) > - a_n^{-1} \|\sigma_n\ X_n\| . \delta\ (\Gamma_{n+1}\ ,\ \Gamma_n)$$

$$- a_n^{-1}\ E^{\mathcal{A}_n} \| X_{n+1} - X_n \|^2$$

(en utilisant l'inégalité de Schwartz, la définition de δ et l'inégalité (i) du lemme 2.1)

$$\text{p.s.} \geq < \sigma_n\ X_n,\ M_n(X) > - a_n^{-1}\ \delta\ (\Gamma_n,\ \Gamma_{n+1}) . \|\sigma_n\ X_n\|$$

$$- a_n^{-1}\ E^{\mathcal{A}_n} \| a_n\ Y_n(X) \|^2 .$$

Ainsi, dès que $\|\sigma_n\ X_n\| \in [\alpha,\ r]$ on a :

$$< \sigma_n\ X_n,\ M'_n\ (X) > \geq e_n\ (\alpha,\ r) - c'_n - d'_n \|\sigma_n\ X_n\|^2 \qquad \text{p.s.}$$

où $c'_n = a_n\ c_n^*$ et $d'_n = a_n\ d_n^* + a_n^{-1}\ \delta(\Gamma_n,\ \Gamma_{n+1})\ \ \alpha^{-1}$,

compte tenu de (2.9) et de : $\|\sigma_n\ X_n\| \leq \alpha^{-1} .\ \|\sigma_n\ X_n\|^2$.

Vérification de (2.2 bis) quand (2.9 bis) est vérifiée

$$E^{\mathcal{A}_n}\| Y'_n(X) - M'_n(X)\|^2 = E^{\mathcal{A}_n}\| (Y_n(X) - M_n(X)) + a_n^{-1}.\,(\pi_{\Gamma_{n+1}} X_{n+1} - E_n^{\mathcal{A}_n} \pi_{\Gamma_{n+1}} X_{n+1})\|^2$$

$$\text{p.s.} \leq 2.E^{\mathcal{A}_n}\| Y_n(X) - M_n(X)\|^2 + 2\, a_n^{-2}.\, E^{\mathcal{A}_n}\| \pi_{\Gamma_{n+1}} X_{n+1} - E^{\mathcal{A}_n} \pi_{\Gamma_{n+1}} X_{n+1}\|^2$$

$$\text{p.s.} \leq 2\, K_n + a_n^{-2}\, E^{\mathcal{A}_n}\| X_{n+1} - E^{\mathcal{A}_n} X_{n+1}\|^2$$

(voir corollaire 1, vérification 2.2 bis)

$$\text{p.s.} \quad \leq 2\, K_n + 2\, E^{\mathcal{A}_n}\| Y_n(X) - M_n(X)\|^2$$

$$\text{p.s.} \quad \leq 4\, K_n\,,$$

(d'après 2.9 bis)

Finalement nous avons, en utilisant (2.9 bis) :

$$\Sigma\, a_n^2\, E^{\mathcal{A}_n}\| Y'_n(X) - M'_n(X)\|^2 \leq \Sigma\, a_n^2\, K_n < \infty\,;$$

(2.2 bis) est donc bien vérifiée.

Vérification de (2.4)

$\Sigma\, a_n^2\, c_n = 6\, \Sigma\, a_n^2\, c_n^* + 4\, \Sigma\, \delta^2(\Gamma_n, \Gamma_{n+1}) < \infty$ d'après (2.11) et (2.12),

$\Sigma\, a_n^2\, d_n = 6\, \Sigma\, a_n^2\, d_n^* < \infty$ d'après (2.12),

$\Sigma\, a_n\, c'_n = \Sigma\, a_n^2\, c_n^* < \infty$ d'après (2.12),

$\Sigma\, a_n\, d'_n = \Sigma\, a_n^2\, d_n^* + \alpha^{-1}\, \Sigma\, \delta\, (\Gamma_n, \Gamma_{n+1}) < \infty$ d'après (2.11) et (2.12),

$\Sigma\, a_n\, e_n\, (\alpha, r) = \infty$ d'après (2.12).

Les conditions sont donc réunies pour appliquer le théorème 2.1..

Les méthodes de pénalité laissent souvent un ensemble de contraintes de signes sur certaines variables.

Si par exemple on veut minimiser J(x) sous des contraintes $C(x)\geq o$ et $D(x) = o$ on se ramène au problème approché suivant :

$$\begin{cases} \min J(x)+\frac{1}{\varepsilon}(|C(x)-t|^2+(D(x)|^2), \\ \text{avec } t\geq o. \end{cases}$$

Modifions donc la définition de Γ_n comme suit :

$\Gamma_n = \{x^*: \langle M_n x, x-x^* \rangle \geq o, \forall x \in C\}$

pour C convexe fermé de H et considérons le processus projeté sur C suivant :

$$\begin{cases} X_o \text{ v.a.C arbitraire,} \\ X_{n+1} = \Pi_C\left(X_n - a_n Y_n(X)\right) . \end{cases}$$

Corollaire 3

L'énoncé du corollaire 2 demeure alors valable avec les modifications précédentes.

Démonstration

Nous poserons : $\mathcal{G}_n$ tribu engendrée par $(X_i)_{i\leq n}$ et $\sigma_n X_n = X_n - \pi_{\Gamma_n} X_n$,

$a_n Y'_n(X) = X_n - \pi_C\left(X_n - a_n Y_n(X)\right) + \pi_{\Gamma_{n+1}} X_{n+1} - \pi_{\Gamma_n} X_n$

de manière que :

$\sigma_{n+1} X_{n+1} = \sigma_n X_n - a_n Y'_n(X)$.

Vérification de (2.2)

$$a_n^2 \overset{\mathcal{G}_n}{E} \| Y'_n(X)\|^2 \leq 2 \overset{\mathcal{G}_n}{E} \| X_{n+1} - X_n\|^2 + 2 \overset{\mathcal{G}_n}{E} \| \pi_{\Gamma_{n+1}} X_{n+1} - \pi_{\Gamma_n} X_n\|^2 \quad \text{p.s.}$$

$$\text{p.s.} \leq 6 \overset{\mathcal{G}_n}{E} \| X_{n+1} - X_n\|^2 + 4\, \delta^2(\Gamma_{n+1}, \Gamma_n)$$

$$\text{p.s.} \leq 6\, a_n^2 \overset{\mathcal{G}_n}{E} \| Y_n(X)\|^2 + 4\, \delta^2(\Gamma_{n+1}, \Gamma_n)$$

$$\text{p.s.} = c_n + d_n \|\sigma_n X_n\|^2 \; (c_n \text{ et } d_n \text{ comme au corollaire 2})$$

(en utilisant des éléments de la démonstration du corollaire 1 et en combinant avec celle du corollaire 2)

Vérification de (2.3) (en combinant les démonstrations des corollaires 1 et 2)

Posons encore $M'_n = \overset{\mathcal{G}_n}{E} Y'_n$.

$$a_n \langle \sigma_n X_n, M'_n(X) \rangle = \overset{\mathcal{G}_n}{E} \langle a_n Y_n(X), \sigma_n X_n \rangle + \overset{\mathcal{G}_n}{E} \langle X'_{n+1} - \pi_C X'_{n+1}, \sigma_n X_n \rangle$$

$$+ \overset{\mathcal{G}_n}{E} \langle \pi_{\Gamma_{n+1}} X_{n+1} - \pi_{\Gamma_n}, \sigma_n X_n \rangle \quad \text{p.s.}$$

(avec $X'_{n+1} = X_n - a_n Y_n(X)$ et bien sûr $X_{n+1} = \pi_C X'_{n+1}$)

$$\text{p.s.} \geq a_n \langle M_n(X), \sigma_n X_n \rangle - \overset{\mathcal{G}_n}{E} \| X'_{n+1} - X_n\|^2 - \overset{\mathcal{G}_n}{E} \| X_{n+1} - X_n\|^2$$

$$- \delta(\Gamma_n, \Gamma_{n+1}) \|\sigma_n X_n\|$$

$$\geq a_n \langle M_n(X), \sigma_n X_n \rangle - 2\, a_n^2 \overset{\mathcal{G}_n}{E} Y_n - \delta(\Gamma_n, \Gamma_{n+1}) \|\sigma_n X_n\|$$

En remplaçant $E^{\mathcal{G}_n} Y_n$ par $2 E^{\mathcal{G}_n} Y_n$ la démonstration se poursuit comme au corollaire 2.

Vérification de (2.2 bis) quand (2.9 bis) est vérifiée

$$a_n^2 E^{\mathcal{G}_n} \| Y'_n(X) - M'_n(X) \| \leq 2 E^{\mathcal{G}_n} \| \pi_C X'_{n+1} - E^{\mathcal{G}_n} \pi_C X'_{n+1} \|^2 +$$
$$2 E^{\mathcal{G}_n} \| \pi_{\Gamma_{n+1}} X_{n+1} - E^{\mathcal{G}_n} \pi_{\Gamma_{n+1}} X_{n+1} \|^2 \text{ p.s.}$$

$$\text{p.s.} \leq 4 E^{\mathcal{G}_n} \| X'_{n+1} - E^{\mathcal{G}_n} X'_{n+1} \|^2 \text{ p.s.}$$ (voir démonstration du corollaire 1).

$$\text{p.s.} = 4 a_n^2 E^{\mathcal{G}_n} \| Y_n(X) = M_n(X) \|^2 \leq 4 a_n^2 K_n$$

Vérification (2.4)

Elle se fait comme aux corollaires 1 et 2.

R E F E R E N C E S

(1) - BERTRAN. J.P. (1973) - Optimisation stochastique dans un espace de Hilbert. Méthode de série divergente. *Colloque d'analyse numérique - La Colle sur Loup -*

(2) - BLUM. J.P. (1954) - Approximation methods witch converge with probability one. *Ann. Math. Stat. 25 - 382-386-*

(3) - BLUM. J.P. (1954) - Multidimentionnal stochastic approximation methods. *Ann. Math. Stat. 25 - 462-483-*

(4) - DAUBEZE. P. (1974) - Contribution à l'étude des masques pour la lecture automatique des caractères imprimés. *Doctorat de Spécialité - Université Paul Sabatier - TOULOUSE - FRANCE - n° 1662 -*

(5) - HIRIART-URRUTY. J.B. (1974) - Algorithmes stochastiques de résolution d'équations et d'inéquations variationnelles. *Document interne - Université de CLERMONT-FERRAND - Département de Mathématiques Appliquées (B.P. 45 - 63170 - AUBIERE - FRANCE).*

(6) - MACCHI. C. (1971) - Thèse de Doctorat d'Etat ès-Sciences Mathématiques - Université de Paris VI - n° CNRS AO-7150-

(7) - ROBBINS. H. , MONRO. S. (1951) - A stochastic approximation method. *Ann. Math. Stat. 22 - 400-407 -*

(8) - SCHMETTERER. L. (1958) - Sur l'itération stochastique. *Ann. Math. Stat. 29 - 373-405-*

(9) - SCHMETTERER. L.- Cours A.E.A. - D.E.A. Mathématiques Appliquées - Université de Clermont-Ferrand (B.P. 45 - 63170 AUBIFRE - FRANCE).

(10)-VENTER. J.H. (1966) - An extension of the Robbins-Monro procedure. *Ann. Math. Stat. 37 - 181-190-*

(11)-WASAN. M.T. (1969) - Stochastic approximation. *Cambridge University Press.*

-:-:-:-:-:-:-:-:-