

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement et de la Recherche Scientifique
Université Abderahamane Mira de Béjaia

Faculté des Sciences Exactes
Département de Mathématiques



Mémoire de fin d'étude

En vue de l'obtention du diplôme de Master en Mathématiques
Option : Probabilités statistique et applications.

Thème

**Les Méthodes d'approximation
stochastiques**

Présenté par :

M^{elle} SLIMANI Fadia

Soutenu devant le Jury composé de :

<i>M^r</i>	Y. BOUMZAID	Président	M.C.B	U.A/ Mira, Bejaia
<i>M^r</i>	F. Maouche	Rapporteur	M.C.B	U.A/ Mira, Bejaia
<i>M^r</i>	M. Bouraine	Examineur	M.A.A	U.A/ Mira, Bejaia
<i>M^r</i>	W. Saci	Examineur	Docteur	U.A/ Mira, Bejaia

Année universitaire : 2019/2020

Remerciement

Je tiens d'abord à remercier Dieu de m'avoir donné le courage, la patience et la force
pour accomplir ce travail.

J'exprime ma profonde reconnaissance à Monsieur Maouche

Mon promoteur, qui a dirigé mon mémoire

Pour ses conseils et commentaires précieux

J'exprime mes vifs remerciements à Monsieur Y.Boumzaid, Monsieur W. Saci et

Monsieur M. Bouraine pour avoir accepté de faire partie du jury de ce mémoire.

Enfin un grand merci à mes parents, qui à travers leur patience et leur compréhension,
m'ont permis d'achever à bien ce modeste travail.

Et aussi à mon fiancé, mes sœurs et mes amies.

Dédicace

Je dédie ce modeste travail qui n'aurait pu aboutir et voir la lumière qu'avec l'aide de dieu le tout puissant à mes très chers parents qui ont toujours été à mes côtés

A mes chères sœurs, Kamilia, Radia, Sinaya, Souad, Lilia

A mes chers frères, Fahem et Imad

A mon fiancé Adel et sa famille et à mes amies

Table des matières

Introduction générale	5
1 Les approximations stochastiques	7
1.1 Introduction	7
1.2 Une approche déterministe	8
1.3 L'algorithme de Robbins-Monro	9
1.3.1 Exemple :	10
1.3.2 Hypothèses	11
1.3.3 Etude de la convergence de l'algorithme Robbins-Monro	12
1.3.4 La convergence presque sûre de l'algorithme Robbins-Monro	14
1.3.5 La convergence en loi de l'algorithme	17
1.4 L'algorithme de Kiefer-Wolfowitz	23
1.5 L'algorithme du gradient stochastique	24
2 Exemples d'application des procédures stochastiques	26
2.1 Introduction	26
2.2 Application aux problèmes inverses	26

2.2.1	Définition de problème inverses mal posé	26
2.2.2	Position du problème	27
2.2.3	Etude de convergence	28
2.2.4	Vitesse de convergence	31
2.3	Application au problème de calibration	34
2.3.1	Calibration d'un modèle multidimensionnel	34
3	Application numérique	40
	Conclusion générale	47
	Bibliographie	49

Introduction générale

Au cours des dernières années, les méthodes d'approximation stochastiques se sont beaucoup développées tant sur le plan de l'analyse mathématique que vers diverses applications : automatique, images, statistique.

Ce mémoire a pour but de présenter les méthodes d'approximation stochastique qui font partie des techniques modernes de résolution numérique de nombreux problèmes pratiques et sont à la base de diverses applications industrielles avancées : traitement de signal [3,4,5,14], control adaptif [11], estimation récursive [6].

Ces méthodes sont relativement anciennes, les initiateurs en sont H. Robbins et S. Monro [19] d'une part, J. Kiefer et J. Wolfowitz [10] d'autre part. Quant à la convergence et à la vitesse de convergence de ces méthodes d'approximation B. T. Polyak [13,18] a donné quelques conditions.

Notre travail est organisé comme suit :

- Chapitre 1 : Dans ce chapitre nous avons fait une présentation générale des méthodes d'approximation stochastique, en particulier l'algorithme de Robbins- Monro et nous avons présenté les hypothèses nécessaires pour atteindre la convergence de ces algorithmes.
- Chapitre 2 : Dans ce chapitre, nous avons donné deux exemples d'applications pour

comprendre comment l'algorithme de Robbins -Monro est appliqué dans divers domaines :

- **Dans premier exemple** on a considéré un problème linéaire décrit par une équation à opérateur où le second membre est mesuré avec des erreurs aléatoires dans un espace de Hilbert. On a estimé la solution en utilisant une procédure itérative stochastique de type Robbins-Monro puis on a montré la convergence presque complète (p.c) de celle-ci tout en précisant la vitesse de convergence.
- **Le deuxième** exemple montre comment utiliser un algorithme stochastique pour résoudre un problème de calibration.
- Chapitre 3 : Représente notre application numérique et concerne l'application de l'algorithme de Robbins-Monro à la recherche de quantiles d'une fonction de répartition inconnue.

Le mémoire s'achève par une conclusion générale.

Chapitre 1

Les approximations stochastiques

1.1 Introduction

Les méthodes d'approximation stochastique sont une famille d'algorithmes stochastiques itératifs qui permettent d'approcher les zéros ou les extremas de fonctions ne pouvant être calculées directement, mais seulement estimées à partir d'observations bruitées. Plus précisément, soit ϕ une fonction inconnue définie sur \mathbb{R}^d pour $d \geq 1$, peuvent s'écrire, pour tout $x \in \mathbb{R}^d$, sous la forme

$$\phi(x) = \mathbb{E}[\Phi(x, \varepsilon)],$$

où l'on sait évaluer $\Phi(x, \varepsilon)$ qui dépend d'un vecteur aléatoire ε . Alors, les méthodes d'approximation stochastique sont des algorithmes de la forme

$$X_{n+1} = X_n + \gamma_n \Phi(X_n, \varepsilon_{n+1}), \tag{1.1}$$

où (ε_n) est une suite de vecteurs aléatoires et (γ_n) est une suite de pas déterministes que l'on explicitera dans la suite.

Les algorithmes stochastiques permettent en particulier de résoudre deux types de problèmes qui sont : d'une part, la recherche de points x^* pour lesquels une fonction inconnue ϕ de \mathbb{R}^d dans \mathbb{R}^d atteint un niveau donné α , C'est -à-dire la recherche de

$$\{x \in \mathbb{R}^d : \phi(x) = \alpha\},$$

et, d'autre part, la recherche de maximum d'une fonction inconnue ϕ à valeurs dans \mathbb{R} ,

$$\arg \max_{x \in \mathbb{R}^d} \phi(x).$$

L'algorithme stochastique le plus célèbre pour résoudre le premier problème est l'algorithme de Robbins-Monro, tandis que le second se résout grâce à l'algorithme de Kiefer-Wolfowitz.

Ce chapitre est consacré à la présentation de ces deux algorithmes stochastiques. Mais avant de détailler la présentation de ces deux algorithmes, nous présentons un algorithme déterministe dont les algorithmes stochastiques sont des versions perturbées.

1.2 Une approche déterministe

Pour ϕ une fonction réelle et $\alpha \in \mathbb{R}$, on souhaite approcher le point x^* satisfaisant $\phi(x^*) = \alpha$. La proposition qui suit est une version simplifiée de la Proposition 1.2.3 du livre de Duflo [6].

Proposition 1.2.1. *soit ϕ une fonction continue satisfaisant*

- *Il existe $x^* \in \mathbb{R}$ tel que $\phi(x^*) = \alpha$.*
- *Pour tout $x \neq x^*$, $(x - x^*)(\phi(x) - \alpha) < 0$.*
- *Il existe $C > 0$ telle que, pour tout $x \in \mathbb{R}$, $|\phi(x)| \leq C(1 + |x|)$.*

De plus, soit (γ_n) une suite positive déterministe décroissante vers 0 et qui vérifie

$$\sum_{n=0}^{+\infty} \gamma_n = +\infty.$$

Alors, l'algorithme déterministe défini récursivement par $x_0 \in \mathbb{R}$ et pour tout $n \geq 0$,

$$x_{n+1} = x_n + \gamma_n(\phi(x_n) - \alpha) \tag{1.2}$$

converge vers x^* .

Démonstration. voir [6].

Il est difficile d'utiliser des algorithmes déterministes du type (1.2) pour la recherche du point x^* . Pour cela, on a alors recours à des algorithmes stochastiques dont celui de Robbins-Monro est le plus célèbre d'entre eux.

1.3 L'algorithme de Robbins-Monro

Dans cette section, pour $\alpha \in \mathbb{R}^d$ et pour une fonction ϕ inconnue de \mathbb{R}^d dans \mathbb{R}^d , on s'intéresse au problème de chercher x^* satisfaisant

$$\phi(x^*) = \alpha. \tag{1.3}$$

Commençons par une heuristique de cet algorithme. On suppose que $d = 1$ et que la fonction ϕ est décroissante de \mathbb{R} dans \mathbb{R} .

Soit X_0 arbitrairement choisi. A l'instant $n \geq 1$, on détermine X_n en fonction des observations antérieures, en supposant qu'on observe $\phi(X_n)$ à travers une variable aléatoire Y_{n+1} qui vérifie $\mathbb{E}[Y_{n+1} | \mathfrak{F}_n] = \phi(X_n)$.

- Si $Y_{n+1} \geq \alpha$, alors $\phi(X_n) \geq \alpha$ donc $X_n \leq x^*$.

Pour approcher x^* il faudra ajouter une quantité positive à X_n .

Au contraire,

- Si $Y_{n+1} \leq \alpha$, il sera nécessaire de soustraire une quantité positive à X_n . Ainsi, l'algorithme de Robbins-Monro est de la forme, pour tout $n \geq 0$

$$X_{n+1} = X_n + \gamma_n(Y_{n+1} - \alpha). \quad (1.4)$$

Dans toute la suite, la fonction ϕ est une fonction qui va de \mathbb{R}^d dans \mathbb{R}^d .

1.3.1 Exemple :

Dosage : Procédure de Robbins-Monro[6]

Une dose x d'un produit chimique crée un effet aléatoire mesuré par $f(x, \varepsilon)$, tel que ε est une variable aléatoire et f une fonction inconnue de \mathbb{R}^2 dans \mathbb{R} . L'effet moyen est supposé croissant, c'est à dire $f(x, \varepsilon)$ est une variable aléatoire de moyenne $\Phi(x) = \mathbb{E}[f(x, \varepsilon)]$, tel que Φ est croissante et inconnue.

On cherche à déterminer la dose x^* qui crée un effet moyen de niveau donné α , c'est -à-dire à résoudre l'équation $\Phi(x^*) = \alpha$.

La procédure utilisée est la suivante, on choisit un dosage X_0 arbitrairement et on l'administre à un sujet qui réagit avec l'effet $u_1 = f(X_0, \varepsilon_1)$, puis on continue par récurrence ; à l'instant n , on choisit, à partir des observations antérieures, une dose X_n administrée à un sujet indépendant de ceux qui furent déjà traités : l'effet est $u_{n+1} = f(X_n, \varepsilon_{n+1})$. Les sujets traités sont de même nature et indépendants entre eux, ce qui traduit par le fait que (ε_n) est une suite de variables aléatoires indépendantes de même loi que ε . Ainsi, puisque

X_n ne dépend que des observations passées, on a

$$E[u_{n+1}|u_1, \dots, u_n] = \Phi(X_n).$$

Si u_{n+1} dépasse α ; on réduit la dose ; Si u_{n+1} est inférieur à α ; on l'augmente. L'algorithme de Robbins-Monro de choix des X_n est de la forme suivante :

$$X_{n+1} = X_n - a_n(u_{n+1} - \alpha), \text{ avec } a_n \geq 0.$$

1.3.2 Hypothèses

Nous allons étudier l'algorithme défini par (1.4). On note \mathfrak{F}_n la σ -algèbre des événements s'étant produits avant l'instant n , $\mathfrak{F}_n = \sigma(X_0, Y_0, \dots, X_n, Y_n)$.

Afin de prouver la convergence de cet algorithme, on fait des hypothèses sur la fonction ϕ , sur la suite d'observations Y_n et sur la suite de pas γ_n .

Les hypothèses sont les suivantes :

(H_1) ϕ est continue.

(H_2) $\forall x \neq x^*, (x - x^*)(\phi(x) - \alpha) < 0$

(H_3) $\forall n \geq 0, \mathbb{E}[Y_{n+1}|\mathfrak{F}_n] = \phi(X_n)$ p.s

(H_4) il existe $C > 0$ telle que $\mathbb{E}[Y_{n+1}^2|\mathfrak{F}_n] \leq C(1 + X_n^2)$ p.s.

(H_5) γ_n décroît vers 0,

$$\sum_{n=0}^{+\infty} \gamma_n = +\infty, \quad \sum_{n=0}^{+\infty} \gamma_n^2 < +\infty.$$

- L'hypothèse (H_2) est vérifiée si ϕ est décroissante,

- les hypothèses (H_3) et (H_4) sont essentielles, l'hypothèse (H_3) assurant qu'on observe une

valeur "pas très loin" de la vraie fonction ϕ , et l'hypothèse (H_4) qui impose que le moment conditionnel de Y_n^2 soit à croissance polynômiale de degré au plus 2.

Un exemple de suite Y_n est la suite $Y_{n+1} = \phi(X_n) + \varepsilon_{n+1}$ où ε_{n+1} est une suite de variables i.i.d. centrées indépendantes de (X_n) ayant un moment d'ordre 2 et la fonction ϕ^2 est à croissance polynômiale de degré au plus 2.

Enfin, l'hypothèse (H_5) est une hypothèse classique sur les pas d'algorithme stochastique.

1.3.3 Etude de la convergence de l'algorithme Robbins-Monro

Rappelle sur la Convergence d'une suite de variables aléatoires

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires réelles, et X une variable aléatoire réelle.

Définition 1.1. Convergence en probabilité [16]

La suite $(X_n)_{n \geq 0}$ *converge vers X en probabilité*, ce que l'on note aussi $X_n \xrightarrow[n]{(p)} X$, si

$$\text{pour tout } \delta > 0, \quad P(|X_n - X| > \delta) \xrightarrow[n]{} 0.$$

Définition 1.2. Convergence presque sûre [16]

La suite $(X_n)_{n \geq 0}$ *converge vers X presque sûrement*, noté $X_n \xrightarrow[n]{(p.s)} X$, si

$$P(X_n \xrightarrow[n]{} X) = 1.$$

Autrement dit $(X_n)_{n \geq 0}$ converge vers X p.s s'il existe un événement $\Omega' \subset \Omega$ tel que $P(\Omega') = 1$ et, pour toute réalisation $\omega \in \Omega'$, la suite réelle $X_n(\omega)$ converge vers $X(\omega)$.

Proposition 1.3.1. [16]

- (a) Si $X_n \xrightarrow[n]{(p.s)} X$, alors $X_n \xrightarrow[n]{(p)} X$
- (b) $X_n \xrightarrow[n]{p.s} X$, si, et seulement si, pour tout $\delta > 0$, $P(\limsup_{n \rightarrow +\infty} \{|X_n - X| > \delta\}) = 0$.

En particulier, si

$$\sum_{n \geq 0} P(|X_n - X| > \delta) < \infty$$

pour tout $\delta > 0$, alors $X_n \xrightarrow[n]{(p)} X$.

- (c) Si $X_n \xrightarrow[n]{(p)} X$, alors il existe une sous suite $(X_{\varphi(k)})_k$ qui converge vers X p.s.

Définition 1.3. Convergence en loi [16]

La suite $(X_n)_{n \geq 0}$ **converge vers X en loi**, noté $X_n \xrightarrow[n]{(loi)} X$, si pour toute fonction continue bornée

$$\varphi : \mathbb{R} \longrightarrow \mathbb{R}, E[\varphi(X_n)] \xrightarrow[n]{} E[\varphi(X)].$$

Une remarque, essentielle est que la variable aléatoire X pourrait être remplacée par n'importe quelle variable aléatoire ayant la même loi, pour cette raison, on peut aussi dire que X_n **converge en loi vers la loi μ** , noté $X_n \xrightarrow[n]{(loi)} \mu$ si $X_n \xrightarrow[n]{(loi)} X$ ou X suit la loi de μ .

De plus, on note que cette convergence ne dépend que de la loi de X_n pour chaque n , et non pas de la loi jointe de $(X_n)_{n \geq 0}$: il s'agit en fait de la convergence de la loi de X_n vers la loi de X . En cela, cette convergence est relativement différente des précédentes.

La convergence en loi est plus faible que toutes les autres, en effet :

Proposition 1.3.2. [16]

(a) Si $X_n \xrightarrow[n]{(p)} X$, alors $X_n \xrightarrow[n]{(loi)} X$.

(b) Dans le cas où X est constante égale à $c \in \mathbb{R}$, $X_n \xrightarrow[n]{(p)} c$ équivaut à $X_n \xrightarrow[n]{(loi)} c$

Le schéma suivant résume les implications entre les modes de convergence

$$\boxed{X_n \xrightarrow[n]{(p.s)} X} \longrightarrow \boxed{X_n \xrightarrow[n]{(p)} X} \longrightarrow \boxed{X_n \xrightarrow[n]{(loi)} X}$$

1.3.4 La convergence presque sûre de l’algorithme Robbins-Monro

Nous commençons par énoncer un lemme central pour la démonstration de la convergence presque sûre de l’algorithme Robbins-Monro.

Lemme 1.3.3. (Robbins-Siegmund [6])

Soient (V_n) , (A_n) , (B_n) et (C_n) quatre suites positives adaptées à \mathfrak{F}_n . On suppose que V_0 est intégrable et pour tout $n \geq 0$.

$$\mathbb{E}[V_{n+1} | \mathfrak{F}_n] \leq V_n(1 + C_n) + A_n - B_n \quad p.s$$

Alors, si

$$\sum_{n=0}^{+\infty} C_n < +\infty, \quad \sum_{n=0}^{+\infty} A_n < +\infty$$

(V_n) converge p.s vers une variable aléatoire V_∞ finie p.s et

$$\sum_{n=0}^{+\infty} B_n < +\infty \quad p.s$$

On donne aussi un lemme très utile dans les preuves, lemme dû à Toeplitz, qui est une généralisation du lemme de Césaro.

Lemme 1.3.4. (Toeplitz [6])

Soit (C_n) une suite de \mathbb{R}_+^* qui vérifie

$$\sum_{n=0}^{+\infty} C_n < +\infty$$

Soit de plus une suite (x_n) de \mathbb{R} qui converge vers $x_\infty \in \mathbb{R}$. Alors, on a

$$\lim_{n \rightarrow +\infty} \frac{\sum_{k=1}^n C_k x_k}{\sum_{k=1}^n C_k} = x_\infty$$

Le théorème suivant assure la convergence presque sûre de X_n vers x^* .

Théorème 1.3.5. (Robbins-Monro [6])

sous les hypothèses (H_{1-5}) , la suite (X_n) définie par (1.4) converge presque sûrement vers x^* .

Démonstration. Pour tout $n \geq 0$, soit $V_n = (X_n - x^*)^2$. Nous allons montrer que V_n converge presque sûrement vers 0. On a

$$\begin{aligned} V_{n+1} &= (X_{n+1} - x^*)^2 \\ &= (X_n - x^* + \gamma_n(Y_{n+1} - \alpha))^2, \\ &= V_n + \gamma_n^2(Y_{n+1} - \alpha)^2 + 2\gamma_n(Y_{n+1} - \alpha)(X_n - x^*) \end{aligned}$$

Ainsi,

$$\mathbb{E}[V_{n+1} | \mathfrak{F}_n] = V_n + \gamma_n^2 \mathbb{E}[(Y_{n+1} - \alpha)^2 | \mathfrak{F}_n] + 2\gamma_n \mathbb{E}[(Y_{n+1} - \alpha)(X_n - x^*) | \mathfrak{F}_n]$$

$$= V_n + \gamma_n^2 \mathbb{E}[(Y_{n+1} - \alpha)^2 | \mathfrak{F}_n] + 2\gamma_n(X_n - x^*) \mathbb{E}[(Y_{n+1} - \alpha) | \mathfrak{F}_n] \quad (1.5)$$

De plus, l'hypothèse (H_3) implique que $\mathbb{E}[(Y_{n+1} - \alpha) | \mathfrak{F}_n] = \phi(X_n) - \alpha$ p.s. Et on a $\mathbb{E}[(Y_{n+1} - \alpha)^2 | \mathfrak{F}_n] \leq 2\mathbb{E}[Y_{n+1}^2 | \mathfrak{F}_n] + 2\alpha^2$ p.s

Donc avec l'hypothèse (H_4) , il existe une constante $C' > 0$ telle que $\mathbb{E}[(Y_{n+1} - \alpha)^2 | \mathfrak{F}_n] \leq C'(1 + X_n^2)$ p.s. De plus, on peut trouver une constante C'' telle que $C'(1 + X_n^2) \leq C''(1 + V_n)$ p.s. Finalement, de (1.5), on déduit que

$$\mathbb{E}[V_{n+1} | \mathfrak{F}_n] \leq V_n(1 + C'' \gamma_n^2) + C'' \gamma_n^2 + 2\gamma_n(X_n - x^*)(\phi(X_n) - \alpha) \quad p.s \quad (1.6)$$

On applique le lemme 1.3.3 aux suites $C_n = A_n = C'' \gamma_n^2$ et $B_n = -2\gamma_n(X_n - x^*)(\phi(X_n) - \alpha)$. V_n, A_n, B_n et C_n sont bien positives (B_n est positive grâce à l'hypothèse (H_2)). De plus, l'hypothèse (H_5) assure que $\sum_{n=0}^{+\infty} C_n < +\infty$ et $\sum_{n=0}^{+\infty} A_n < +\infty$. Donc grâce au lemme 1.3.3, on déduit que V_n converge p.s. vers une variable aléatoire V_∞ finie p.s. et que

$$\sum_{n=0}^{+\infty} -2\gamma_n(X_n - x^*)(\phi(X_n) - \alpha) < +\infty \quad p.s \quad (1.7)$$

Il reste donc à montrer que $V_\infty = 0$ p.s. Par l'absurde, supposons que $V_\infty \neq 0$ p.s.

Alors, on peut trouver $0 < a < b < 1/2$ tels que, pour n assez grand,

l'évènement $\{a < |X_n - x^*| < b\}$ n'est pas négligeable. Cependant, sur cet anneau, on peut aussi trouver une constante $c > 0$ telle que $(x^* - X_n)(\phi(X_n) - \alpha) > c$ qui par (1.7),

implique

$$\sum_{n=1}^{\infty} \gamma_n < +\infty,$$

ce qui contredit (H_5) . Par conséquent $V_\infty = 0$ p.s.

1.3.5 La convergence en loi de l'algorithme

Afin d'établir un théorème central limite pour la suite (X_n) définie par (1.4), on a besoin de rajouter les hypothèses suivantes aux hypothèses (H_{1--5}) :

(H₆) ϕ est de classe $C^2(\mathbb{R})$.

(H₇) La fonction g définie par $g^2(X_n) = \mathbb{E}[Y_{n+1}^2 | \mathfrak{F}_n]$ est continue au voisinage de x^* .

(H₈) Il existe $p > 2$ tel que $\sup \mathbb{E}[|Y_{n+1} - \phi(X_n)|^p | \mathfrak{F}_n] < +\infty$ p.s

On prend $\gamma_n = 1/n$.

Théorème 1.3.6. *Sous les hypothèses (H_{1--8}) et si on note $q = -\phi'(x^*)$ et $\sigma^2 = g^2(x^*) - \alpha^2$, alors*

1. Si $q > 1/2$

$$\sqrt{n}(X_n - x^*) \xrightarrow{\text{loi}} \mathcal{N}\left(0, \frac{\sigma^2}{2q - 1}\right)$$

2. Si $q = 1/2$,

$$\sqrt{\frac{n}{\ln(n)}}(X_n - x^*) \xrightarrow{\text{loi}} \mathcal{N}(0, \sigma^2).$$

3. Si $0 < q < 1/2$,

$$n^q(X_n - x^*) \xrightarrow{\text{loi}} Z,$$

où Z est une variable aléatoire finie p.s

Démonstration. On fait la démonstration pour $q > 1/2$. On trouvera une démonstration pour $q = 1/2$ et $q < 1/2$ dans [6]. Plus précisément si $q > 1/2$, pour tout $n \geq 0$, nous déduisons de (1.4) que

$$X_{n+1} - x^* = X_n - x^* + \gamma_n(Y_{n+1} - \alpha) \tag{1.8}$$

De plus,

$$\begin{aligned} Y_{n+1} &= (Y_{n+1} - \phi(X_n)) + \phi(X_n) \\ &= (Y_{n+1} - \phi(X_n)) + \alpha + (\phi(X_n) - \alpha) \end{aligned} \quad (1.9)$$

Et un développement de Taylor de ϕ permet d'écrire

$$\begin{aligned} \phi(X_n) - \alpha &= \phi(X_n) - \phi(x^*) \\ &= (X_n - x^*)\phi'(x^*) + (\phi(X_n) - \phi(x^*) - (X_n - x^*)\phi'(x^*)). \end{aligned}$$

En notant $Z_{n+1} = Y_{n+1} - \phi(X_n)$ et $\delta_n = \phi(X_n) - \phi(x^*) - (X_n - x^*)\phi'(x^*)$, on déduit de (1.9) que

$$Y_{n+1} = Z_{n+1} + \alpha + (X_n - x^*)\phi'(x^*) + \delta_n, \quad (1.10)$$

et donc avec la notation $q = -\phi'(x^*)$, (1.8) devient

$$\begin{aligned} X_{n+1} - x^* &= X_n - x^* + \gamma_n Z_{n+1} + \gamma_n (X_n - x^*)\phi'(x^*) + \gamma_n \delta_n \\ &= (1 - q\gamma_n)(X_n - x^*) + \gamma_n Z_{n+1} + \gamma_n \delta_n \end{aligned} \quad (1.11)$$

En notant $\alpha_n = 1 - q\gamma_n$, une récurrence immédiate mène, pour tout $n \geq 2$, à

$$X_{n+1} - x^* = \beta_{n-1}(X_1 - x^*) + \beta_{n-1}M_n + \beta_{n-1}R_{n-1} \quad (1.12)$$

où

$$\beta_{n-1} = \prod_{k=1}^{n-1} \alpha_k, \quad (1.13)$$

$$M_n = \sum_{k=1}^{n-1} \frac{\gamma_k}{\beta_k} Z_{k+1}, \quad (1.14)$$

Et

$$R_n = \sum_{k=1}^n \frac{\gamma_k}{\beta_k} \delta_k. \quad (1.15)$$

De plus, avec $\gamma_k = 1/k$, il est facile de voir que

$$\beta_n \sim n^{-q} \quad (1.16)$$

La suite (M_n) est une martingale dont le crochet est donné par

$$\begin{aligned} \langle M \rangle_n &= \sum_{k=1}^{n-1} \left(\frac{\gamma_k}{\beta_k} \right)^2 \mathbb{E}[Z_{k+1}^2 | \mathfrak{F}_k], \\ &= \sum_{k=1}^{n-1} \left(\frac{\gamma_k}{\beta_k} \right)^2 (g^2(X_k) - \phi(X_k)^2) \end{aligned} \quad (1.17)$$

De plus, comme X_k converge p.s. vers x^* , $g^2(X_k) - \phi(X_k)^2$ converge p.s. vers $g^2(x^*) - \alpha^2 = \sigma^2$ et par le lemme 1.3.4 de Toeplitz, il s'ensuit de (1.17) que si $2q - 1 > 0$

$$\langle M \rangle_n \sim \sigma^2 \sum_{k=1}^{n-1} \left(\frac{\gamma_k}{\beta_k} \right)^2 \sim \frac{\sigma^2}{2q-1} n^{2q-1}. \quad (1.18)$$

Donc,

$$\frac{M_n}{n^{q-1/2}} \xrightarrow{loi} \mathcal{N} \left(0, \frac{\sigma^2}{2q-1} \right)$$

Ce qui implique par (1.16) que

$$\sqrt{n} \beta_{n-1} M_n \xrightarrow{loi} \mathcal{N} \left(0, \frac{\sigma^2}{2q-1} \right) \quad (1.19)$$

Ensuite, comme $2q - 1 > 0$, on déduit immédiatement de (1.16) que

$$\sqrt{n} \beta_{n-1} (X_1 - x^*) = O(n^{1/2-q}) \quad (1.20)$$

qui converge p.s. vers 0 quand n tend vers l'infini, il reste à voir que

$$\sqrt{n}\beta_{n-1}R_{n-1}$$

converge p.s. vers 0 quand n tend vers l'infini. En effet, pour tout $h > 0$, il existe $\varepsilon > 0$ tel

que si $|x - x^*| < \varepsilon$,

$$|\phi(x) - \phi(x^*) - (x - x^*)\phi'(x^*)| \leq h(x - x^*)^2$$

donc

$$-h(x - x^*)^2 \leq \phi(x) - \phi(x^*) - (x - x^*)\phi'(x^*) \leq h(x - x^*)^2,$$

en particulier,

$$(x - x^*)(\phi(x) - \phi(x^*)) - (x - x^*)^2\phi'(x^*) \leq h\varepsilon(x - x^*)^2. \quad (1.21)$$

De plus, rappelons que si $V_n = (X_n - x^*)^2$ alors on l'équation (1.6) donnée par

$$\mathbb{E}[V_{n+1}|\mathfrak{F}_n] \leq V_n(1 + C''\gamma_n^2) + C''\gamma_n^2 + 2\gamma_n(X_n - x^*)(\phi(X_n) - \alpha) \quad p.s$$

où $C'' > 0$, De plus, pour $|X_n - x^*| \leq \varepsilon$ alors d'après (1.21)

$$(X_n - x^*)(\phi(X_n) - \alpha) \leq (X_n - x^*)^2(\phi'(x^*) + h\varepsilon) = V_n(-q + h\varepsilon),$$

ce qui mène à

$$\mathbb{E}[V_{n+1}|\mathfrak{F}_n] \leq V_n(1 + C''\gamma_n^2 - 2\gamma_n(q - h\varepsilon)) + C''\gamma_n^2 \quad p.s. \quad (1.22)$$

Posons alors pour $N \geq 1$.

$$T_N = \inf\{n \geq N : |X_n - x^*| > \varepsilon\}.$$

Comme $\{T_N > n\} = \bigcap_{k=N}^n \{|X_n - x^*| \leq \varepsilon\} \in \mathfrak{F}_n$, on déduit de (1.22)

$$\mathbb{E}[V_{n+1}I_{\{T_N > n\}} | \mathfrak{F}_n] \leq V_n I_{\{T_N > n\}} (1 + C'' \gamma_n^2 - 2\gamma_n(q - h\varepsilon)) + C'' \gamma_n^2 \text{ p.s.} \quad (1.23)$$

Et $\{T_N > n + 1\} \subset \{T_N > n\}$ donc

$$\mathbb{E}[V_{n+1}I_{\{T_N > n+1\}} | \mathfrak{F}_n] \leq V_n I_{\{T_N > n\}} (1 + C'' \gamma_n^2 - 2\gamma_n(q - h\varepsilon)) + C'' \gamma_n^2 \text{ p.s.} \quad (1.24)$$

En passant à l'espérance, on déduit donc que si

$$v_n = \mathbb{E}[V_n I_{\{T_N > n\}}]$$

Alors pour $n \geq N$

$$v_{n+1} \leq v_n (1 + C'' \gamma_n^2 - 2\gamma_n(q - h\varepsilon)) + C'' \gamma_n^2 \quad (1.25)$$

$$\leq v_n (1 - 2\gamma_n p_{h,n}) + C'' \gamma_n^2 \text{ p.s.} \quad (1.26)$$

où on a noté

$$p_{h,n} = \left(-\frac{C''}{2}\gamma_n + (q - h\varepsilon)\right).$$

On obtient par une récurrence immédiate que pour $\gamma_k = 1/k$,

$$v_n \leq v_N \omega_{n-1} + \omega_{n-1} \sum_{k=N}^{n-1} \frac{1}{k^2 \omega_k} \text{ p.s.} \quad (1.27)$$

Où

$$\omega_n = \prod_{k=N}^n \left(1 - 2\frac{p_{h,k}}{k}\right) \quad (1.28)$$

Et il est facile de voir que $n^{2p_{h,N}} \omega_n$ converge vers une constante et

$$\sum_{k=N}^{n-1} \frac{1}{k^2 \omega_k} = O(n^{-1+2p_{h,N}}).$$

Donc

$$v_n = O(n^{-2p_{h,N}} + n^{-1}) = O(n^{-1})$$

Si $p_{h,N} \geq 1/2$. Or par hypothèse, $q > 1/2$, on choisit donc h et N de telle sorte que

$$q - 1/2 \geq h\varepsilon + \frac{C''}{N} > 0.$$

Ainsi, la suite

$$\sum_{k=1}^n \frac{\gamma_k}{\beta_k} \mathbb{E}[V_k I_{\{T_N > k\}}]$$

a le même comportement asymptotique que

$$\sum_{k=1}^n \frac{\gamma_k^2}{\beta_k} = O\left(\frac{1}{n\beta_n}\right).$$

Ainsi,

$$\sqrt{n}\beta_n \sum_{k=1}^n \frac{\gamma_k^2}{\beta_k} = O\left(\frac{1}{\sqrt{n}}\right), \tag{1.29}$$

donc la suite

$$\sqrt{n}\beta_n \sum_{k=1}^n \frac{\gamma_k}{\beta_k} \mathbb{E}[V_k I_{\{T_N > k\}}]$$

tend vers 0 quand n tend vers l'infini. Rappelons que R_n est donné par (1.15)

$$R_n = \sum_{k=1}^n \frac{\gamma_k}{\beta_k} \delta_k.$$

Et comme X_k tend p.s. vers x^* , il suffit de travailler sur la suite

$$\hat{R}_n = \sum_{k=1}^n \frac{\gamma_k}{\beta_k} V_k. \tag{1.30}$$

Et comme $\{T_N = +\infty\} \subset \{T_N > k\}$ pour tout k alors

$$\hat{R}_n I_{\{T_N = +\infty\}} = \sum_{k=1}^n \frac{\gamma_k}{\beta_k} V_k I_{\{T_N = +\infty\}}$$

$$\leq \sum_{k=1}^n \frac{\gamma_k}{\beta_k} V_k I_{\{T_N > k\}}$$

Ainsi d'après (1.29),

$$\hat{R}_n I_{\{T_N = +\infty\}} = \mathbf{op}\left(\frac{1}{\sqrt{n}\beta_n}\right). \tag{1.31}$$

Or la convergence p.s. de X_k vers x^* entraine que l'ensemble

$$\bigcup_{N \geq 1} \{T_N = +\infty\}$$

Est un ensemble de probabilité 1 et donc que

$$\hat{R}_n = \mathbf{op}\left(\frac{1}{\sqrt{n}\beta_n}\right). \tag{1.32}$$

Ainsi, $\sqrt{n}\beta_n R_n$ converge en probabilité vers 0. On déduit alors de (1.12), (1.19), (1.20) que

$$\sqrt{n}(X_n - x^*) \xrightarrow{loi} \mathcal{N}\left(0, \frac{\sigma^2}{2q-1}\right)$$

1.4 L'algorithme de Kiefer-Wolfowitz

On s'intéresse maintenant au deuxième problème évoqué au début du chapitre qui est la recherche de x^* , pour une fonction inconnue ϕ , défini par

$$x^* = \arg \max_{x \in \mathbb{R}^d} \phi(x). \tag{1.33}$$

Comme dans le cas de l'algorithme de Robbins-Monro, commençons par une heuristique de l'algorithme, soit ϕ une fonction qui va de \mathbb{R} dans \mathbb{R} . Si ϕ est dérivable, alors par définition de x^* , on a $\phi'(x^*) = 0$. On sait alors que si on dispose d'observation Y_{n+1} telle que $\mathbb{E}[Y_{n+1} | \mathfrak{F}_n] = \phi'(X_n)$, l'algorithme de Robbins-Monro défini par $X_0 \in \mathbb{R}$ et pour tout $n \geq 0$, $X_{n+1} = X_n + \gamma_n Y_{n+1}$ va converger p.s vers x^* sous de bonnes hypothèses sur

ϕ . Néanmoins, il se peut que nous n'ayons que des observations Y_{n+1} nous permettant d'approcher ϕ et non ϕ' . On a alors recours à l'algorithme de Kiefer-Wolfowitz.

L'idée est que même si on ne peut pas approcher directement ϕ' , on peut quand même approcher la dérivée par différentiation de ϕ qui sera une bonne approximation de ϕ' .

En particulier, la suite d'observations Y_{n+1} définie par

$$Y_{n+1} = \frac{Y_{n+1}^+ - Y_{n+1}^-}{2c_n}$$

où les suites (Y_{n+1}^+) et (Y_{n+1}^-) sont telles que

$$\mathbb{E}[Y_{n+1}^+ | \mathfrak{F}_n] = \phi(X_n + c_n) \quad p.s$$

$$\mathbb{E}[Y_{n+1}^- | \mathfrak{F}_n] = \phi(X_n - c_n) \quad p.s$$

Avec (c_n) une suite déterministe à préciser, et une bonne approximation de ϕ' . L'algorithme de Kiefer-Wolfowitz est donc de la forme

$$X_{n+1} = X_n + \frac{\gamma_n}{2c_n}(Y_{n+1}^+ - Y_{n+1}^-)$$

1.5 L'algorithme du gradient stochastique

Soit (Ω, A, \mathbb{P}) un espace de probabilité et ε une variable aléatoire définie sur Ω à valeurs dans W . On note μ la loi de probabilité résultant du transport de la loi \mathbb{P} par ε .

Soit H un espace de Hilbert muni du produit scalaire $\langle \cdot, \cdot \rangle$ et de la norme $\|\cdot\|$, une partie convexe fermée non vide U de H et une fonction j définie sur $H \times W$ supposée intégrable pour tout $x \in U$ et à valeurs dans $\bar{\mathbb{R}}$.

On note J l'espérance de la fonction j :

$$J(x) = \mathbb{E}[j(x, \varepsilon)] = \int_{\Omega} j(x, \varepsilon) d\mathbb{P} = \int_W j(x, \varepsilon) d\mu. \quad (1.34)$$

On s'intéresse à la problématique suivante :

$$\min_{x \in U} J(x). \quad (1.35)$$

Pour obtenir la solution du problème (1.35) et calculer le gradient $J'(x)$ de J en tout point x , on peut utiliser un algorithme de type gradient sous les hypothèses classiques de convexité et de différentiabilité.

Cette approche peut s'avérer coûteuse en temps de calcul, car chaque évaluation du gradient passe par le calcul d'une espérance.

La méthode du gradient stochastique consiste à mettre en œuvre un algorithme au cours duquel la variable à optimiser x évolue en fonction du gradient de j évalué pour une valeur particulière de la variable aléatoire ε , et non en fonction du gradient de l'espérance de j . C'est là une idée nouvelle par rapport aux méthodes d'optimisation classiques, car on cherche simultanément à faire progresser la méthode de gradient et à effectuer un calcul d'espérance.

L'algorithme que l'on propose alors est le suivant.

$$X_{n+1} = \text{proj}_U(X_n - a_n j'_x(X_n, \varepsilon_{k+1})), \quad (1.36)$$

où, (ε_k) est une suite de variable aléatoire indépendantes et de même loi μ , et a_n est une suite des variables réelles.

Remarque

Les algorithmes de Kiefer-Wolfowitz et du gradient stochastique ainsi que l'étude de leur convergence sont traités dans la bibliographie suivante [20]

Chapitre 2

Exemples d'application des procédures stochastiques

2.1 Introduction

Les algorithmes stochastiques ont été largement utilisés dans de nombreuses applications.

Notre but dans ce chapitre est d'appliquer la procédure d'approximation de type Robbins-Monro pour approximer la solution de l'équation à opérateur suivante $Ax = u$.

Puis, nous présentons un exemple sur le marché financier et on donne la modélisation mathématique du problème et expliquer comment l'algorithme de Robbins-Monro peut aider à résoudre de tels problèmes. Ensuite, nous traitons le cas d'une option panier.

2.2 Application aux problèmes inverses

2.2.1 Définition de problème inverses mal posé

En 1923, Hadamard a introduit la notion de problème bien posé. il s'agit d'un problème dont :

- La solution existe
- la solution est unique
- la solution dépend continument des données.

Un problème qui n'est pas bien posé au sens de la définition ci-dessus est dit mal posé.

2.2.2 Position du problème

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et H un espace de Hilbert séparable. On considère l'équation à opérateur suivante

$$Ax = u \tag{2.1}$$

Où A est un opérateur compact linéaire de H . On pose

$$u_k = u_{ex} + \xi_k$$

où u_{ex} est la valeur exacte et inconnue du second membre.

$(\xi_k)_{k \in \mathbb{N}^*}$ est une suite de variables aléatoire fonctionnelles indépendantes à valeur dans H tel que :

$$E\|\xi_k\|^2 \leq L < +\infty \tag{2.2}$$

Et

$$E\|\xi_k\|^m \leq \frac{m!}{2} E\|\xi_k\|^2 l^{m-2}, \quad m \geq 2 \tag{2.3}$$

E : l'espérance mathématique

l, L sont des constantes positives.

2.2.3 Etude de convergence

On utilise la méthode d'approximation de type Robbins-Monro suivante

$$X_{k+1} = X_k - \frac{1}{k}[AX_k - u_k], \quad k \in \mathbb{N}. \quad (2.4)$$

On note u_{ex} la mesure directe exacte et X_{ex} la solution de notre problème inverse tel que $u_{ex} = AX_{ex}$, et comme $u_k = u_{ex} + \xi_k$ Alors, la procédure (2.4) devient comme suite

$$X_{k+1} = X_k - \frac{1}{k}[AX_k - u_{ex} - \xi_k] \quad (2.5)$$

La procédure (2.5) peut être réécrite comme suit

$$X_{k+1} - X_{ex} = X_k - X_{ex} - \frac{1}{k}[AX_k - u_{ex}] + \frac{1}{k}\xi_k \quad (2.6)$$

$$X_{k+1} - X_{ex} = X_k - \frac{1}{k}AX_k + \frac{1}{k}u_{ex} + \frac{1}{k}\xi_k - X_{ex} \quad (2.7)$$

Puisque $u_{ex} = AX_{ex}$, on aura

$$X_{k+1} - X_{ex} = \left(I - \frac{1}{k}A\right) X_k + \left[\frac{1}{k}AX_{ex} - X_{ex}\right] + \frac{1}{k}\xi_k. \quad (2.8)$$

Alors, la relation suivante est obtenue

$$X_{k+1} - X_{ex} = \left(I - \frac{1}{k}A\right) (X_k - X_{ex}) + \frac{1}{k}\xi_k. \quad (2.9)$$

Posons $a_k = \frac{1}{k}$

Pour résoudre l'équation à opérateur linéaire (2.1) on va énoncer un théorème général de la convergence qui donne une borne exponentielle de la solution itérative obtenue par la procédure (2.5).

Théorème 2.2.1 (.16). *Soit $A \in \mathfrak{L}(H)$, X_i, ξ_i des éléments prenant leurs valeurs dans H , $a_i = \frac{1}{i}, i \in \mathbb{N}^*$. La solution du problème (2.1) vérifie l'inégalité exponentielle suivante*

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp \left(- \frac{\varepsilon^2}{16 \sum_{i=1}^k \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 E \|\xi\|^2 a_i^2} \right). \quad (2.10)$$

Démonstration. On a la relation suivante

$$\begin{aligned} \mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) &= \mathbb{P} \left(\left\| \prod_{i=1}^k (I - a_k A)(X_1 - X_{ex}) + \sum_{i=1}^k \prod_{j=i+1}^k (I - a_j A) a_i \xi_i \right\| > \varepsilon \right) \\ &\leq \mathbb{P} \left(\left\| \sum_{i=1}^k \prod_{j=i+1}^k (I - a_j A) a_i \xi_i \right\| > \varepsilon - \left\| \prod_{i=1}^k (I - a_k A)(X_1 - X_{ex}) \right\| \right) \end{aligned}$$

Et pour k assez grand, on a

$$\left\| \prod_{i=1}^k (I - a_k A)(X_1 - X_{ex}) \right\| < \frac{\varepsilon}{2}.$$

Donc, On obtient

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq \mathbb{P} \left(\left\| \sum_{i=1}^k \prod_{j=i+1}^k (I - a_j A) a_i \xi_i \right\| > \frac{\varepsilon}{2} \right) \quad (2.11)$$

Posons

$$\eta_i = \prod_{j=i+1}^k (I - a_j A) a_i \xi_i.$$

Grâce aux les inégalités suivantes

$$\mathbb{P} \left(\left\| \sum_{i=1}^k \eta_i \right\| > \frac{\varepsilon}{2} \right) \leq \exp^{-\frac{t\varepsilon}{2}} E \left(\exp \left(t \left\| \sum_{i=1}^k \eta_i \right\| \right) \right) \leq 2 \exp^{-\frac{t\varepsilon}{2}} E \cosh t \left\| \sum_{i=1}^k \eta_i \right\| \quad (2.12)$$

$$\begin{aligned} E \cosh t \left\| \sum_{i=1}^k \eta_i \right\| &\leq \prod_{i=1}^k E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \\ \mathbb{P} \left(\left\| \sum_{i=1}^k \eta_i \right\| > \frac{\varepsilon}{2} \right) &\leq 2 \exp^{-\frac{t\varepsilon}{2}} \prod_{i=1}^k E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \end{aligned} \quad (2.13)$$

Lorsqu'on développe la fonction exponentielle au voisinage de zéro, on obtiendra

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) = 1 + \sum_{m=2}^{\infty} \frac{t^m E\|\eta_i\|^m}{m!}. \quad (2.14)$$

En remplaçant la valeur de η_i , on trouve

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \leq 1 + \sum_{m=2}^{\infty} \frac{t^m \left\| \prod_{j=i+1}^k (I - a_j A) a_i \right\|^m E\|\xi_i\|^m}{m!}$$

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \leq 1 + \frac{t^2 \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2}{2} \sum_{m=2}^{\infty} t^{m-2} \left\| \prod_{j=i+1}^k (I - a_j A) a_i \right\|^{m-2} l^{m-2}$$

On pose

$$t \leq \frac{1}{2 \left\| \prod_{j=i+1}^k (I - a_j A) a_i \right\| l}$$

Donc

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \leq 1 + \frac{t^2 \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2}{2} \sum_{m=2}^{\infty} \frac{1}{2^{m-2}}$$

Ce qui implique que

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \leq 1 + t^2 \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2$$

Et comme $1 + x \leq \exp(x)$, on a

$$E(\exp(t\|\eta_i\|) - t\|\eta_i\|) \leq \exp \left(t^2 \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2 \right).$$

D'ou

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp^{-\frac{t\varepsilon}{2}} \prod_{j=i+1}^k \exp \left(t^2 \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2 \right).$$

qui est équivalente à la relation suivante

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp \left(-\frac{t\varepsilon}{2} + t^2 \sum_{i=1}^k \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E\|\xi_i\|^2 \right). \quad (2.15)$$

Le second membre de l'équation (2.15) atteint son minimale lorsque

$$t^* = \frac{\varepsilon}{4 \sum_{i=1}^k \left\| \prod_{j=i+1}^k (I - a_j A) \right\|^2 a_i^2 E \|\xi_i\|^2}.$$

Si on remplace t^* dans (2.15) on trouve l'équation (2.10)

2.2.4 Vitesse de convergence

Corollaire 2.2.2 (16). *Sous les conditions du théorème 2.2.1. La relation suivante est vérifiée*

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp\left(-\frac{(k+1)^{2c} \varepsilon^2}{16D}\right) \quad (2.16)$$

Avec, D, c des constantes, $1/2 < c \leq 1$.

Démonstration. En utilisant le fait que : $\ln(1+x) \leq x$, pour $x > -1$, il est facile de prouver :

$$\left\| \prod_{j=i+1}^k (I - a_j A) \right\| \leq \prod_{j=i+1}^k \left(1 - \frac{c}{j}\right) \leq \left(\frac{i+1}{k+1}\right)^c. \quad (2.17)$$

Pour plus de détail, voir [17].

et comme

$$\ln\left(1 - \frac{c}{j}\right) \leq -\frac{c}{j}$$

Alors,

$$\prod_{j=i+1}^k \left(\left(1 - \frac{c}{j}\right)\right) \leq \left(\frac{i+1}{k+1}\right)^c$$

Donc, la relation (2.10) est équivalente à

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{16 \left(\sum_{i=1}^k \left(\frac{i+1}{k+1}\right)^{2c} E \|\xi_i\|^2 a_i^2\right)}\right)$$

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp \left(-\frac{\varepsilon^2}{16 \sum_{i=1}^k \left(\frac{i+1}{k+1}\right)^{2c} \frac{a^2}{i^2} E\|\xi_i\|^2} \right)$$

Puisque $E\|\xi_i\|^2 \leq L$, on trouve

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp \left(-\frac{\varepsilon^2}{16La^2 \sum_{i=1}^k \frac{i+1^{2c}}{i} \frac{1}{k+1}^{2c}} \right)$$

Avec $C = La^2$ et si on pose : $w_i = \frac{1}{i^2}$, $b_i = (i+1)^{2c}$, alors,

$\frac{1}{b_k} \sum_{i=1}^k b_i w_i$ tend vers 0 lorsque k tend vers l'infinité. i.e.

$$\lim_{k \rightarrow +\infty} \frac{C}{(K+1)^{2c}} \sum_{i=1}^k \frac{i+1^{2c}}{i^2} = 0$$

Donc, $\exists N \in \mathbb{N}$ tel que

$$\frac{C}{(K+1)^{2c}} \sum_{i=N+1}^{\infty} \frac{i+1^{2c}}{i^2} \leq \varepsilon.$$

D'ou

$$\frac{C}{(K+1)^{2c}} \sum_{i=1}^{\infty} \frac{i+1^{2c}}{i^2} \leq \frac{C}{(K+1)^{2c}} \sum_{i=1}^N \frac{i+1^{2c}}{i^2} + \varepsilon = \frac{CC'}{(K+1)^{2c}} + \varepsilon$$

Ou $C' = \sum_{i=1}^N \frac{(i+1)^{2c}}{i^2}$ Finalement, on a

$$\mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq 2 \exp \left(-\frac{\varepsilon^2}{\frac{16CC'}{(k+1)^{2c}}} \right) = 2 \exp \left(-\frac{(k+1)^{2c} \varepsilon^2}{16D} \right), \quad D = CC', \quad 1/2 < c \leq 1$$

Corollaire 2.2.3. 1. La procédure (2.5) converge presque complètement (p.co) vers la solution de l'équation (2.1) i.e.

$$\sum_{k=1}^{+\infty} \mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) \leq +\infty \tag{2.18}$$

2. On a

$$\|X_{k+1} - X_{ex}\| = O(k^{-2c}), \quad 1/2 < c \leq 1 \tag{2.19}$$

Démonstration. [16]

1. Posons

$$v_k = 2 \exp\left(-\frac{(k+1)^{2c}\varepsilon^2}{16D}\right) \leq 2 \exp(-(k+1)^{2c}\varepsilon^2)$$

Si on applique la règle de Cauchy sur la série à termes positive v_k , on trouve que

$\sum_{k=1}^{+\infty} v_k$ est convergente. Ce qui implique que

$$\forall \varepsilon > 0, \sum_{k=1}^{+\infty} \mathbb{P}(\|X_{k+1} - X_{ex}\| > \varepsilon) = 2 \sum_{k=1}^{+\infty} \exp(-(k+1)^{2c}\varepsilon^2) < +\infty \quad (2.20)$$

Ce qui assure la convergence presque complète.

2. Pour obtenir (2.19) il suffit de choisir $A = \varepsilon k^{2c}$ dans l'inégalité précédente pour avoir

$$\sum_{k=1}^{+\infty} \mathbb{P}(\|X_{k+1} - X_{ex}\| > Ak^{-2c}) < +\infty$$

Corollaire 2.2.4. *Sous les hypothèses du théorème (2.2.1) et pour un niveau donné γ , on peut trouver un entier naturel k_γ , tel que la solution X_{ex} appartienne à une boule fermée du centre X_{k_γ} et du rayon ε avec une probabilité supérieure ou égale à $1 - \gamma$,*

Démonstration. On a

$$\lim_{k \rightarrow +\infty} 2 \exp(-(k+1)^{2c}\varepsilon^2) = 0.$$

Ce qui implique l'existence d'un entier naturel k_γ , tel que

$$k \geq k_\gamma \implies 2 \exp(-(k+1)^{2c}\varepsilon^2) \leq \gamma.$$

Donc

$$\sum_{k=1}^{+\infty} \mathbb{P}(\|X_{k_\gamma+1} - X_{ex}\| > \varepsilon) \geq 1 - \gamma.$$

2.3 Application au problème de calibration

Dans cette partie, nous fournirons un exemple supplémentaire sur l'application de l'algorithme de Robbins-Monro dans le domaine des marchés financiers et prouverons que ces algorithmes sont appliqués dans divers domaines. Pour plus de détails voir [8].

2.3.1 Calibration d'un modèle multidimensionnel

L'une des principales préoccupations des banques est de trouver le meilleur modèle d'ajustement possible marché donné.

L'idée est de déterminer l'ensemble des paramètres qui correspondent au meilleur prix du marché pour les options liquides. Nous sommes particulièrement intéressés à trouver le paramètre de corrélation dans les modèles de marché multidimensionnels.

Dans cette section, nous présentons la modélisation mathématique du problème et expliquons comment les algorithmes stochastiques et en particulier l'algorithme de Robbins-Monro peut aider à résoudre de tels problèmes. Ensuite, nous traitons le cas d'une option panier.

La modélisation mathématique du problème

Nous considérons un marché financier modélisé par un espace de probabilité $(\Omega, \mathfrak{F}, \mathbb{P})$. Nous supposons que \mathbb{P} est déjà la mesure neutre au risque et que $\mathfrak{F} = (\mathfrak{F}_t)_{t \geq 0}$ est la filtration générée par d mouvements browniens standards unidimensionnels $B^i = \{B_t^i, t \geq 0\}$, Sous ces conditions,

les prix \bar{S}_t^i des actifs sont donnés par l'équation suivante

$$d\bar{S}_t^i = \bar{S}_t^i(rdt + \sigma_i dB_t^i), \quad \bar{S}_0^i = x_i,$$

Où $d\langle B^i, B^j \rangle_t = \rho_{ij} dt$ si $i \neq j$, 1 sinon, r est le taux d'intérêt et σ_i le volatilité de l'actif i .

Soit $\Gamma = (\rho_{ij})_{1 \leq i, j \leq d}$ est la matrice de covariance des mouvements browniens et doit donc être positive.

Nous supposons que certains prix des options de panier sont indiqués par $(C_j)_{j \in J}$. Dans ce cas, nous devons calibrer le paramètre de corrélation entre les différents actifs.

Notre objectif est alors de trouver les corrélations qui ont conduit à ces prix. Nous considérons certains gains $(\phi_j(\bar{S}_T))_{j \in J}$, et leurs prix d'option correspondants $(C_j)_{j \in J}$.

L'idée est de minimiser, sur l'ensemble des matrices admissibles Γ , un critère bien choisi qui permet de formuler le problème comme un problème des moindres carrés en considérant la différence relative entre les prix du marché et ceux calculés

$$\sum_{j \in J} \frac{1}{C_j^2} (\mathbb{E}[e^{-rT} \phi_j(\bar{S}_T)] - C_j)^2. \tag{2.21}$$

Supposons que les coefficients de corrélation $\rho_{ij} = \rho \quad \forall i \neq j$. Bien que cette hypothèse est assez restrictive en ce qui concerne la structure de covariance, les praticiens supposent souvent une telle structure pour garantir que le problème est bien défini. Comme rappelé ci-dessus, la matrice Γ doit être définitivement positive. Tout d'abord, nous devons trouver l'ensemble des valeurs admissibles pour ρ . Les valeurs propres de Γ sont $1 - \rho$ avec la multiplicité 1 et $(d - 1)\rho + 1$ avec multiplicité $d - 1$. Pour garantir que Γ est une matrice de covariance bien définie il est nécessaire et suffisant de choisir ρ dans l'intervalle ouvert $U =]\frac{-1}{d-1}, 1[$. Pour des raisons de simulation, nous considérons des mouvements browniens

indépendants. Donc, nous introduisons $\mathcal{W} = \{W_t, t \geq 0\}$ un mouvement brownien standard de dimension d . Supposons \mathfrak{L} désigner la factorisation de Cholesky. On sait que le vecteur $(B_T^i)_{i=1\dots d}$ est de même distribution que W_T . Si l'on considère le processus $(S_t)_{t \geq 0}$ tel que $dS_t^i = S_t^i(rdt + \sigma_i \mathfrak{L}_i dW_t)$ où \mathfrak{L}_i désigne la i -ème ligne de la matrice \mathfrak{L} . Alors S_T et \bar{S}_T sont même distribution.

Les critères peuvent être réécrits avec S_T .

$$\sum_{j \in J} \frac{1}{C_j^2} (\mathbb{E}[e^{-rT} \phi_j(S_T)] - C_j)^2. \tag{2.22}$$

Nous essayons de trouver le zéro de la dérivée des critères ci-dessus (2.22) par rapport à ρ . Si nous admettons que nous pouvons échanger la différenciation et l'esperance (voir Paragraphe 3.1.5 [8] pour une preuve de celui-ci), nous cherchons la racine de la fonction f

$$\begin{aligned} f(\rho) &= \frac{1}{C^2} \mathbb{E} \left[e^{-rT} \frac{d\phi(S_T)}{d\rho} \right] \mathbb{E} [e^{-rT} \phi(S_T) - C], \\ f(\rho) &= \frac{1}{C^2} \mathbb{E}[e^{-rT} \nabla \phi(S_T) \cdot \nabla_\rho S_T] \mathbb{E}[e^{-rT} \phi(S_T) - C], \\ f(\rho) &= \frac{1}{C^2} \mathbb{E}[e^{-rT} (e^{-rT} \phi(\bar{S}_T) - C) \nabla \phi(\bar{S}_T) \cdot \nabla_\rho S_T], \end{aligned}$$

Où \bar{S}_t est une copie indépendante de S_t

La fonction f est définie par une esperance mathématique

$$f(\rho) = \mathbb{E}[F(\rho, G^1, G^2)]$$

où G^1 et G^2 sont deux variables aléatoires indépendantes de loi $\mathcal{N}(0, I_{\mathbb{R}^d})$.

Minimisation des critères

Nous allons maintenant expliquer comment minimiser les critères (2.22) à l'aide d'un algorithme stochastique.

Cas d'une option panier

Cas d'une option panier à partir de maintenant, nous considérons les options sur un panier d'actifs avec gains ϕ des éléments de type suivant

$$\phi(S_T) = \left(\sum_{i=1}^d \lambda_i S_T^i - K \right)_+$$

Formulation des critères

On peut réécrire le gain en fonction du paramètre de corrélation et d'une variable standard normale

$$\psi(G, \rho) = \left(\sum_{i=1}^d \lambda_i x_i \exp((\tau - \sigma_i/2)T + \sigma_i \sqrt{T} \mathfrak{L}_i G) - K \right)_+$$

Les critères de la fonction f peuvent alors être réécrits

$$\begin{aligned} f(\rho) &= \frac{1}{C^2} \mathbb{E} \left(e^{-\tau T} (e^{-\tau T} \phi(\bar{S}_T) - C) \left(\sum_{i=2}^d \lambda_i \sigma_i S_T^i d\mathfrak{L}_i W_T \mathbf{1}_{\phi(S_T) > 0} \right) \right) \\ f(\rho) &= \frac{1}{C^2} \mathbb{E} \left((e^{-\tau T} \psi(G^1, \rho) - C) \left(\sum_{i=2}^d \lambda_i \sigma_i S_0^i e^{-\sigma_i^2 T/2 + \sigma_i \mathfrak{L}_i(\rho) \sqrt{T} G^{(2)}} d\mathfrak{L}_i(\rho) \sqrt{T} G^{(2)} \mathbf{1}_{\psi(G^{(2)}, \rho) > 0} \right) \right), \\ f(\rho) &= \frac{1}{C^2} \mathbb{E}(F(\rho, G^{(1)}, G^{(2)})) \end{aligned}$$

Où $G^{(1)}$ et $G^{(2)}$ sont deux variables normales standard indépendantes dans \mathbb{R}^d .

Nous considérons $(G_n^{(1)})_{n \geq 0}$ et $(G_n^{(2)})_{n \geq 0}$ deux séquences de variables normales standard indépendantes dans \mathbb{R}^d et nous introduisons la séquence ρ_n définie par récurrence pour tout ρ_0 arbitraire dans U .

$$\rho_{n+1} = \prod_{U_\varepsilon} (\rho_n - \gamma_{n+1} F(\rho_n, G_{n+1}^{(1)}, G_{n+1}^{(2)})) \quad (2.23)$$

Où U_ε est un ensemble compact strictement inclus dans U tel que $d(U_\varepsilon, \partial U) \geq \varepsilon$

Exemple

Considérons une option sur un panier de trois actifs de gain $(S_T^1 + S_T^2 + S_T^3 - K)$ avec les caractéristiques suivantes :

$$S_0^1 = 30, \quad S_0^2 = 10, \quad S_0^3 = 15,$$

$$\sigma_1 = 0.3, \quad \sigma_2 = 0.4, \quad \sigma_3 = 0.25$$

$$\tau = 0.035, \quad T = 3, \quad K = 90$$

Et avec une corrélation entre les différents actifs égale à 0,3.

La figure 2.1 montre l'estimation du paramètre de corrélation à l'aide de la méthode décrite ci-dessus avec $\gamma = 5$.

Nous obtenons une corrélation approximative de 0,303 tandis que la valeur réelle est de 0,3, ce qui donne une estimation assez précise.

Les résultats semblent bons, même si l'influence du choix du paramètre de gain sur la précision de l'approximation n'est pas à négligée. Le choix du paramètre γ est définitivement un problème brûlant.

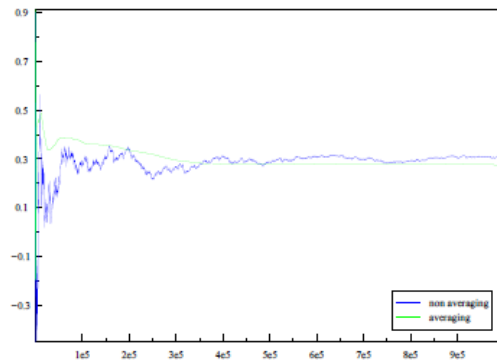


Figure 2.1 : estimation de la corrélation.

Chapitre 3

Application numérique

La procédure développée dans cette partie est l'application de l'algorithme de Robbins-Monro sur la recherche de quantiles d'une fonction de répartition inconnue.

Définition 3.1. Soit Z une variable aléatoire réelle de fonction de répartition F inconnue. Pour $0 < \alpha < 1$, on définit le quantile d'ordre α de Z comme le plus petit réel z_α satisfaisant

$$F(z_\alpha) = \mathbb{P}(Z \leq z_\alpha) = \alpha.$$

c'est à dire $\min\{z_\alpha : F(z_\alpha) = \alpha\}$.

Exemple

Pour $\alpha = 1/2$, $F(z_\alpha) = \mathbb{P}(Z \leq z_\alpha) = 1/2$, c'est la recherche de la médiane.

Pour $\alpha = 1/4$, $F(z_\alpha) = \mathbb{P}(Z \leq z_\alpha) = 1/4$, c'est la recherche de premier quartile Q_1 .

Pour $\alpha = 3/4$, $F(z_\alpha) = \mathbb{P}(Z \leq z_\alpha) = 3/4$, c'est la recherche de troisième quartile Q_3 .

Et comme F est croissante, continue à droite avec limite à gauche.

Afin de garantir l'unicité de z_α , on suppose que F est continue et strictement croissante.

Ainsi F vérifie (H_1) .et pour tout $z \in \mathbb{R}$ avec $z \neq z_\alpha$,

$$(z - z_\alpha)(F(z) - \alpha) > 0$$

De plus, on suppose qu'on dispose d'une suite de variables aléatoire (Z_n) indépendantes et de même loi que Z .

Alors, pour estimer z_α , on définit l'algorithme de Robbins-Monro par $X_0 \in \mathbb{R}$ et pour tout $n \geq 0$,

$$X_{n+1} = X_n - \gamma_n(Y_{n+1} - \alpha)$$

Où

$$Y_{n+1} = I_{\{Z_{n+1} \leq X_n\}}.$$

Et comme la suite (Y_n) vérifié facilement les hypothèses suivantes :

1.

$$\forall n \geq 0, \mathbb{E}[Y_{n+1} \mid \mathfrak{F}_n] = \phi(X_n) \quad p.s$$

2. il existe $C > 0$ telle que $\mathbb{E}[Y_{n+1}^2 \mid \mathfrak{F}_n] \leq C(1 + X_n^2) \quad p.s.$

Nous prenons l'exemple d'une variable aléatoire Z de loi exponentielle de paramètre $\lambda > 0$ et nous allons notamment voir l'influence de la croissance de F et du choix du pas γ_n sur la convergence numérique de l'algorithme.

La fonction de répartition de Z est définie, pour tout $x \in \mathbb{R}$ par :

$$F_\lambda(x) = (1 - \exp(-\lambda x))I_{x>0}.$$

Sur la figure 2.2 on a tracé les fonctions de répartition pour les paramètres

$$\lambda_1 = 0.5 \quad \lambda_2 = 1 \quad \lambda_3 = 3 \quad \lambda_4 = 5$$

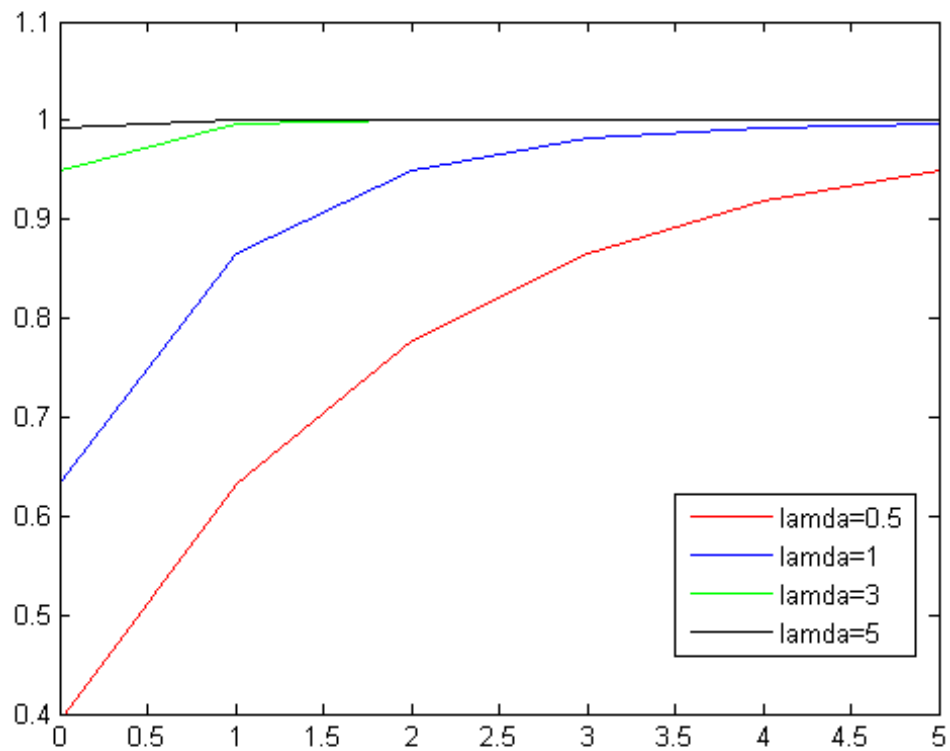


Figure 2.2 : Fonction de répartition de la loi exponentielle.

Pour $0 < \alpha < 1$, on a :

$$F(z_\alpha) = \alpha$$

$$1 - \exp^{-\lambda z_\alpha} = \alpha$$

Donc

$$z_\alpha = -\frac{\log(1 - \alpha)}{\lambda}.$$

_ Influence de la croissance de F_λ :

On a lancé l’algorithme de Robbins-Monro avec un pas $\gamma_n = 1/n$ pour différentes valeurs de n et différentes valeurs du niveau α . Les résultats obtenus pour la valeur absolue de la différence entre X_n et z_α sont donnés dans le tableau suivant.

$\lambda_1=0.5$	n=100	n=500	n=1000
$\alpha = 0.25$	0.1062	0.0321	0.0062
$\alpha = 0.5$	0.0226	0.0466	0.0875
$\alpha = 0.9$	1.8062	1.7186	1.6099

$\lambda_2=1$	n=100	n=500	n=1000
$\alpha = 0.25$	0.0999	0.0017	0.0081
$\alpha = 0.5$	0.2209	0.1253	0.0502
$\alpha = 0.9$	0.1533	0.4902	0.5039

$\lambda_3=3$	n=100	n=500	n=1000
$\alpha = 0.25$	0.0240	0.0042	0.0066
$\alpha = 0.5$	0.0597	0.0148	0.0021
$\alpha = 0.9$	0.0023	0.0078	0.0161

$\lambda_4=5$	n=100	n=500	n=1000
$\alpha = 0.25$	0.0020	0.0022	0.0021
$\alpha = 0.5$	0.0625	0.0188	0.0106
$\alpha = 0.9$	0.1001	0.2396	0.0903

Interprétation des résultats :

D’après l’étude comparative que nous avons réalisée nous remarquons que :

- Plus le paramètre λ est grand, plus l’estimation de z_α est meilleure.
- Plus le niveau α est grand, moins l’algorithme de Robbins-Monro est performant.

C’est à dire l’estimation de z_α par l’algorithme de Robbins-Monro est meilleur quand la croissance de la fonction F_λ est forte et quand la valeur à estimer se trouve à un endroit où F_λ est fortement croissante.

- l’algorithme de Robbins-Monro est sensible par rapport au point de départ X_0 . On remarque que l’algorithme donne de meilleurs resultats lorsque X_0 est proche de la solution exacte, ce qui represente un inconvenient car la solution exacte n’est pas

forcement connue.

– Influence du choix de γ_n :

Jusqu'à présent, nous avons considéré un pas $\gamma_n = 1/n$ ou $\gamma_n = 1/n^\alpha$ avec $1/2 < \alpha < 1$ qui sont les choix naturels pour vérifier l'hypothèse suivante :

$$\sum_{n=0}^{+\infty} \gamma_n = +\infty \quad \text{et} \quad \sum_{n=0}^{+\infty} \gamma_n^2 < +\infty$$

Comme nous allons le voir dans le tableau suivant, le choix du pas a une véritable influence sur la convergence numérique de X_n vers z_α .

Aussi, nous avons choisi trois types de pas vérifiant l'hypothèse précédant qui sont :

$$\gamma_n = 1/n, \quad \gamma_n = 1/n^{3/4}, \quad \gamma_n = 1/n^{1.1/2}$$

Et nous nous sommes placés dans le cas $\lambda_1 = 0.5$ et $\alpha = 0.9$.

On a alors répertorié dans le tableaux ci-dessous les résultats obtenus pour la valeur absolue de la différence entre X_n et z_α pour chaque choix de pas

	n=100	n=500	n=1000
$\gamma_n = 1/n$	1.8062	1.7186	1.6099
$\gamma_n = 1/n^{3/4}$	1.8029	0.8301	0.4088
$\gamma_n = 1/n^{1.1/2}$	0.9127	0.2558	0.0893

Donc le tableau précédent montre que le meilleur pas pour ces simulations est $\gamma_n = 1/n^{1.1/2}$ et que les différences de convergence entre les algorithmes dépendent fortement du pas choisi.

On donnera l'algorithme implémenté sous Matlab du problème définie plus haut.

Algorithme :

```
clear all;
n = input('donner la valeur de n');
alpha = input('donner la valeur de alpha');
lamda = input('donner la valeur de lamda');
X(1) = 1;
zal = -(log(1 - alpha)/lamda);
for j = 1 : n + 1
Z = -(log(rand(j, 1))/lamda);
end
for j = 1 : n - 1
if (Z(j + 1) <= X(j))
T(j + 1) = 1;
else
T(j + 1) = 0;
end
X(j + 1) = X(j) - (1/j) * (T(j + 1) - alpha);
end
v = abs(zal - X(n));
disp(zal);
disp(v);
```

Conclusion générale

Nous nous sommes intéressés dans ce mémoire aux méthodes d'approximations stochastiques, particulièrement l'algorithme de Robbins Monro. Ces méthodes font partie de méthodes modernes de résolutions de problèmes issus des sciences de l'ingénieur.

Dans le premier chapitre, nous avons abordé des généralités sur l'algorithme de Robbins Monro qui fut introduit en 1952 et donné les conditions générales pour la convergence de cette procédure . Nous avons abordé par la suite, l'algorithme de Kiefer Wolfowitz qui est une variante de Robins Monro et qui permet de chercher un maximum d'une fonction inconnue. Et à la fin du chapitre, nous avons donné l'algorithme du gradient stochastique.

Dans le deuxième chapitre, nous avons donné deux exemples d'application de l'algorithme de Robbins Monro, le premier exemple concerne son application aux problèmes inverses. L'autre, sur le problème de calibration dans les marchés financiers qui fait l'estimation du paramètre de corrélation.

Le chapitre 3, constitue notre application numérique où nous avons réalisé et programmé la procédure de Robins Monro pour la recherche des quantiles d'une fonction de répartition inconnue. Nous avons conclu que l'algorithme de Robbins Monro est un algorithme très efficace et qui converge très rapidement vers la solution. Néanmoins, nous avons remarqué aussi qu'il est sensible par rapport au point initial, car celui-ci ne doit pas être éloigné de

la solution exacte.

Comme perspective, on s'est mis au défi d'étudier d'autres algorithmes qui évitent la sensibilité vis-a-vis du point initial, par exemple la méthode LASSO.

Bibliographie

- [1] Benveniste, A., Métivier, M. and Priouret, P. *Adaptive algorithm and stochastic approximations*. Springer-Verlag.1990
- [2] Cavalier, L. *Problèmes inverses en statistique.*, Mémoire présenté pour obtenir l'habilitation à diriger des recherches,2003.
- [3] Ding, F. *Hierarchical estimation algorithms for multivariable systems using measurement information*.Information Sciences, 2014
- [4] Ding, F. *State filtering and parameter estimation for state space systems with scarce measurements*. Signal Processing, 2014
- [5] Ding, F. *Decomposition based fast least squares algorithm for output error systems*.Signal Processing. 2013
- [6] Duflo, M. *Méthodes récursives aléatoires*. Masson. 1990
- [7] Hadamard, j. *Lectures on Cauchy's Probleme in linear Parial Differential Equation*,Yale University Press, 1923
- [8] Jérôme Lelong. *Etude asymptotique des algorithmes stochastiques et calcul du prix*

- des options Parasiennes*, Ecole nationale des pont et chaussees, 2007
- [9] K. Aki and P. Richards. *Quantitative Seismology : Theory and Methods*. Freeman, 1980.
- [10] Kiefer, J., Wolfowitz, J. *Stochastic estimation of the maximum of a regression function*. The Annals of Mathematical Statistics, 1952
- [11] Kushner, H. J. *Numerical methods for controls for non linear stochastic systems with delays and jumps : Applications to admission control*. *Stochastics : An International Journal of Probability and Stochastic Processes*, 2011
- [12] Kern M. *Problèmes inverses*. Syllabus du cours à l'Ecole supérieure d'ingénieurs Léonard de Vinci.
- [13] Levitin, E. S., Polyak, B.T. *Convergence of minimizing sequences in conditional extremum problems*. Doklady Akademii Nauk SSSR, 1966
- [14] Li, P., Lilja, D. J. *Using stochastic computing to implement digital image processing algorithms*. In *Computer Design*, 2011
- [15] Mokkadem, A. and Pelletier, M. *Convergence rate and averaging of nonlinear twotime-scale stochastic approximation algorithms*. Ann. Appl. Probab., 16(3) :1671–1702, 2006. ISSN 1050-5164.
- [16] Maouche Fouad, *les approches stochastiques pour la résolution de problème mal posés*, Université de Bijaia, 2015
- [17] Philippe Frayesse, *Estimation récursive dans certains modèles de déformation*, Université Science et Technologies-Bordeaux, 2013
- [18] Polyak, B. T. *The conjugate gradient method in extremal problems*. Computational Mathematics and Mathematical Physics, 1969

-
- [19] Robbins, H., Monro, S *A stochastic approximation method*, the annals of mathematical statistics, 1951
- [20] Dahmani, A., Rahmani, S. *On the Rate of convergence in the central limit theorem for martingale difference sequences of the Kiefer wolfowitz algorithm*. Afr Diaspora J. Math.7.2008, no 1.9-19.