Méthodes de Monte Carlo en Finance

Notes de cours

Bruno Bouchard

Université Paris VI, LPMA, et CREST

bouchard@ccr.jussieu.fr

Cette version : Mai 2006[1]

[1]Première version: 2002

Table des matières

1 Généralités et nombres aléatoires **7**
1.1 Notions générales 7
1.2 Loi uniforme 9
1.3 Autres lois 10
1.3.1 Méthode d'inversion 10
1.3.2 Méthode du rejet 11
1.3.3 Méthode par transformation 14
1.3.4 Variables corrélées et techniques par conditionnement 17
1.4 Notion de discrépance 18
1.4.1 Généralités 18
1.4.2 Cas des suites uniformément distribuées 20
1.4.3 Suites à discrépance faible 21

2 Méthodes de discrétisation **23**
2.1 Modèle de Black-Scholes 23
2.2 Schéma de discrétisation 24
2.2.1 Discrétisation d'Euler 24
2.2.2 Schéma de Milshtein 29
2.3 Options Asiatiques 31
2.3.1 Schémas simples 31
2.3.2 Schéma avec développement dans le modèle de Black-Scholes 34
2.4 Options à barrière 35
2.4.1 Approche naive 35
2.4.2 Approche par les ponts de diffusion 36
2.5 Options Lookback 40
2.6 Options Américaines 41

3 Réduction de la variance **43**
3.1 Contrôle antithétique 43
3.2 Régularisation du payoff 45
3.3 Variable de contrôle 46
3.3.1 Motivation et exemples 46

3.3.2 Approche systématique . . . 49
3.3.3 Méthode adaptative . . . 50
3.4 Fonction d'importance . . . 50
3.4.1 Un exemple . . . 50
3.4.2 Approche systématique . . . 52
3.4.3 Fonction d'importance optimale et algorithme de Robbins Monro 52

4 Calcul des sensibilités 57
4.1 Approche par différences finies . . . 57
4.2 Grecques dans le modèle de Black et Scholes . . . 58
4.3 Processus tangent . . . 60
4.3.1 Notions de processus tangent . . . 60
4.3.2 Processus tangent et delta . . . 61
4.3.3 Processus tangent et véga (exercice) . . . 62
4.4 Calcul de Malliavin . . . 63
4.4.1 Introduction au calcul de Malliavin . . . 63
4.4.2 Calcul de Malliavin et sensibilités . . . 67

5 Calcul d'espérances conditionnelles 71
5.1 Espérances conditionnelles et densités . . . 71
5.2 Application à l'évaluation d'options américaines . . . 77

6 Compléments (exercices) 81
6.1 Simulation d'un CIR : schéma exact . . . 81
6.2 Simulation d'un CIR : schéma d'Euler implicite . . . 83
6.3 Discrétisation d'EDSR . . . 84
6.4 Algorithme de Robbins Monro . . . 87
6.5 Méthode de rejet avec recyclage . . . 88
6.6 Estimation d'espérances conditionnelles . . . 89
6.7 Variables antithétiques . . . 91
6.8 Suites uniformes randomisées . . . 93

Introduction

Dans les modèles sans friction ni contrainte, le prix d'une option européenne, sur un sous-jacent $X = (X_t)_{t\geq 0}$, de payoff actualisé $g(X)$, s'écrit comme l'espérance sous une probabilité risque neutre, unique si le marché est complet, de sa valeur liquidative : $p(0,x) = \mathbb{E}\left[g(X) \mid X_0 = x\right]$.

Il existe de multiples manières d'évaluer cette espérance : approche par arbres, par équations aux dérivées partielles (EDP), ... Ici, nous nous concentrons sur l'approche Monte-Carlo. Elle consiste à simuler un grand nombre de réalisations de $g(X)$ puis à en prendre la moyenne $\hat{p}(0,x)$, la loi des grands nombres assurant la convergence de $\hat{p}(0,x)$ vers $p(0,x)$.

Le Chapitre 1 est consacré à des généralités sur les méthodes de Monte-Carlo et aux principaux modes de génération de nombres "aléatoires".

Si g dépend de toute la trajectoire de X, ou si X est solution d'une équation différentielle stochastique (EDS) dont on ne connaît pas explicitement la solution, il est généralement impossible de simuler parfaitement $g(X)$. Il faut alors recourir à une approximation qui correspond à une discrétisation en temps. Ces techniques seront étudiées dans le Chapitre 2.

Même lorsqu'on sait simuler $g(X)$, il arrive que la variance de l'estimateur soit trop grande pour que celui-ci soit fiable. Le Chapitre 3 est consacré aux techniques de réduction de variance.

Connaître le prix auquel on peut vendre une option est important, mais il faut ensuite pouvoir se couvrir. La couverture étant définie à partir du delta, la dérivée de $p(0,x)$ par rapport à x, il faut pouvoir l'estimer. Le Chapitre 4 est consacré à ce sujet. Il sera l'occasion d'introduire les notions de processus tangent et de dérivée de Malliavin.

Enfin, lorsque l'option est de type américaine, les méthodes de Monte-Carlo standard conduisent à estimer un très grand nombre d'espérances conditionnelles. Nous verrons dans le dernier Chapitre comment la formule d'intégration par parties du calcul de Malliavin permet de les ré-écrire sous une forme exploitable numériquement.

Nous renvoyons à [38], [44] et [51] pour une présentation théorique des techniques d'évaluation et de gestion de portefeuille en finance.

Notations

On commence par introduire quelques notations. On indentifiera un élément x de $\mathbb{R}^d$ au vecteur colonne associé de coordonnées x^j, $j = 1, \ldots, d$. On notera $\mathbb{M}^d$, l'ensemble des matrices carrées de dimension d, de coordonnées a^{ij}. On notera I_d la matrice identité de $\mathbb{M}^d$ et e_d le vecteur unité de $\mathbb{R}^d$. La norme euclidienne sur $\mathbb{R}^d$ ou $\mathbb{M}^d$ sera notée $\|\cdot\|$. Pour un élément a de $\mathbb{M}^d$, on notera $a^{i\cdot}$ le vecteur ligne correspondant à sa i-ème ligne et $a^{\cdot j}$ le vecteur colonne correspondant à sa j-ème colonne. La transposée de $a \in \mathbb{M}^d$ sera notée a^*. Pour un ensemble de point $(x^i)_{i=1}^d$, on notera $\mathrm{Vect}[(x^i)_{i=1}^d]$ le vecteur colonne de componsantes x^i. Pour $x \in \mathbb{R}^d$, on notera $\mathrm{diag}\,[x]$ la matrice diagonale de $\mathbb{M}^d$ dont les éléments diagonaux sont les x^i. Pour une fonction $f : \mathbb{R}^d \mapsto \mathbb{R}^n$, on notera ∇f sa matrice Jacobienne, i.e. $(\nabla f)^{ij} = \partial f^i / \partial x^j$. Si $f : (\mathbb{R}^d)^k \mapsto \mathbb{R}^n$, on notera $\nabla_\ell f$, la matrice Jacobienne obtenue en dérivant par rapport à sa ℓ-ème composante vectorielle, i.e. $(\nabla_\ell f(x_1, \ldots, x_k))^{ij} = \partial f^i(x_1, \ldots, x_k)/\partial x_\ell^j$. On écrira parfois simplement $\nabla_{x_\ell} f$. On notera C_p^k (resp. C_b^k) l'ensemble des fonctions C^k à croissance polynomiale (resp. bornées) dont les dérivées sont également à croissance polynomiale (resp. bornées). Lorsque la fonction est seulement définie sur A, on notera $C^k(A)$, $C_p^k(A)$ et $C_b^k(A)$. $\mathcal{N}(m, \Sigma)$ désignera la loi normale de moyenne m et de matrice de variance-covariance Σ. On notera souvent par $C > 0$ une constante générique qui peut changer de valeur d'une ligne à l'autre.

Chapitre 1

Généralités et nombres aléatoires

1.1 Notions générales

Les méthodes de Monte-Carlo sont basées sur la loi des grands nombres : on simule un grand nombre de variables aléatoires (v.a.) indépendantes et de même loi, $(Y_n)_n$ de même loi que Y, on prend ensuite la moyenne des valeurs prises, $N^{-1}\sum_{n=1}^{N} Y_n$, et on obtient une approximation de $\mathbb{E}[Y]$. La convergence est assurée par le théorème classique suivant.

Théorème 1.1.1 *(Loi des grands nombres) Soit $(Y_n)_{n\geq 1}$ une suite de v.a. réelles intégrables i.i.d.*[1] *Alors,*

$$\frac{1}{N}\sum_{n=1}^{N} Y_n \quad \longrightarrow \quad \mathbb{E}[Y_1] \qquad \textit{p.s. et dans } L^1 \; .$$

Si en plus les (Y_n) sont de carré intégrable, on peut montrer que la convergence a lieu dans L^2 et on obtient un contrôle de l'erreur p.s. et de l'erreur L^2.

Théorème 1.1.2 *Soit $(Y_n)_{n\geq 1}$ une suite de v.a. réelles i.i.d. de carré intégrable. On pose*

$$\hat{m}_N \quad := \quad \frac{1}{N}\sum_{n=1}^{N} Y_n \; .$$

Alors ,

$$\mathbb{E}\left[|\hat{m}_N - \mathbb{E}[Y_1]|^2\right] \quad = \quad \frac{\mathrm{Var}(Y_1)}{N} \; ,$$

et

$$\begin{aligned} 1 \quad &= \quad \limsup_{N\to\infty} \sqrt{\frac{N}{\mathrm{Var}(Y_1)\, 2\ln\ln N}}\,(\hat{m}_N - \mathbb{E}[Y_1]) \\ &= \quad -\liminf_{N\to\infty} \sqrt{\frac{N}{\mathrm{Var}(Y_1)\, 2\ln\ln N}}\,(\hat{m}_N - \mathbb{E}[Y_1]) \qquad \textit{p.s.} \end{aligned}$$

[1] Identiquement Indépendemment distribuées, i.e. de même loi et indépendantes

Preuve. La première égalité s'obtient en développant le terme de gauche et en utilisant le fait que les $Y_n - \mathbb{E}[Y_1]$ sont indépendants et centrés. La seconde est connue sous le nom de Loi du Logarithme itéré. $\square$

Remarque 1.1.3 Il est facile de vérifier que, si $Y_1 \in L^2$,

$$\hat{\sigma}_N^2 := \frac{1}{N-1} \sum_{n=1}^{N} (Y_n - \hat{m}_N)^2$$

est un estimateur sans biais de $\mathrm{Var}(Y_1)$. D'après le Théorème précédent, $\hat{\sigma}_N^2/N$ est donc un estimateur sans biais de $\mathrm{Var}(\hat{m}_N)$.

Enfin, le théorème de la limite centrale permet de contruire des intervalles de confiance asymptotiques.

Théorème 1.1.4 *(Théorème de la limite centrale) Soit $(Y_n)_{n\geq 1}$ une suite de v.a. réelles i.i.d. de carré intégrable. On suppose que $\mathrm{Var}(Y_1) > 0$ et on pose*

$$\hat{m}_N := \frac{1}{N} \sum_{n=1}^{N} Y_n \quad \textit{et} \quad \hat{\sigma}_N := \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (Y_n - \hat{m}_N)^2} \; .$$

Alors,

$$\sqrt{N} \left(\frac{\hat{m}_N - \mathbb{E}[Y_1]}{\hat{\sigma}_N} \right) \mathbf{1}_{\hat{\sigma}_N > 0} \;\longrightarrow\; \mathcal{N}(0,1) \quad \textit{en loi} \; .$$

En particulier, pour tout réel $c > 0$,

$$\mathbb{P}\left[\sqrt{N} \left| \frac{\hat{m}_N - \mathbb{E}[Y_1]}{\hat{\sigma}_N} \right| < c \right] \;\longrightarrow\; 1 - \alpha_c$$

où $\alpha_c = \mathbb{P}[|X| > c]$ pour $X \sim \mathcal{N}(0,1)$.

On en déduit que, pour N suffisamment grand, la probabilité d'avoir $\mathbb{E}[Y_1] \in [\hat{m}_N \pm c\hat{\sigma}_N/\sqrt{N}]$ est proche de $1 - \alpha_c$.

Remarque 1.1.5 Une estimation seule ne veut rien dire. L'estimateur est une variable aléatoire qui a sa propre variance. Deux estimations différentes peuvent conduire à deux valeurs très différentes. Il est donc important de connaître la variance de l'estimateur ou, et c'est encore mieux, de fournir un intervalle de confiance. En particulier, lorsque l'on voudra donner un prix d'option calculé par simulation, il faudra toujours le mettre en perspective avec la variance de l'estimateur ou avec un intervalle de confiance, de manière à pouvoir juger de la précision du résultat obtenu.

1.2 Loi uniforme

En pratique un ordinateur ne sait engendrer que des suites de nombres déterministes : il est incapable de générer une suite "réellement" aléatoire. Par contre, il est possible de constuire des suites de nombres qui se comportent (statistiquement) comme des suites aléatoires.

Les suites les plus courantes produites par les ordinateurs sont calculées à partir d'un nombre fini d'entiers $\{0, \ldots, M-1\}$. En divisant par M, on obtient ainsi une suite sur $[0,1[$. Elles sont construites sur la base de récurrences de la forme

$$u_{n+1} \;=\; g(u_n) \quad n \in \mathbb{N}$$

où g est une fonction de $\{0, \ldots, M-1\}$ dans lui-même et u_0, appelé graine, est à initialiser dans $\{0, \ldots, M-1\}$. On pose alors

$$x_n \;=\; \frac{u_n}{M} \;\in\; [0,1[\quad n \in \mathbb{N}\,.$$

L'exemple le plus simple est celui de la congruence mixte :

$$g(u) \;=\; (Au + C) \mod M$$

où A et C sont des réels positifs à choisir (dans le cas $C = 0$, on parle de congruence multiplicative).

Remarque 1.2.1 1. Il est clair que de telles suites ne permettent pas de générer tout les réels entre 0 et 1.
2. Par construction, les suites ainsi contruites sont périodiques de période au plus égale à M. Il faut donc s'assurer que la période est suffisamment longue par rapport à l'utilisation que l'on compte en faire. On peut choisir A, C et M de manière à avoir une période maximale, voir [24].
3. Il est nécessaire d'initialiser la graine u_0. Pour obtenir des "tirages" différents, il faut bien évidemment partir d'une graine différente. On peut par exemple initialiser la graine en utilisant l'horloge de l'ordinateur. Ceci doit se faire en tout début de programme.
4. La plupart des générateurs pré-programmés dans les languages de type C sont de ce type. Il faut toujours vérifier que la période est suffisemment grande par rapport à l'utilisation que l'on veut en faire (RAND_MAX en C).

Nous ne rentrerons pas plus ici dans la description de ces générateurs (voir par exemple [24] pour plus de détails). Vous pouvez trouver de bons générateurs sur internet, en consultant par exemple la version en ligne du livre *Numerical Recipes*[2].

Remarque 1.2.2 1. Si $U \sim \mathcal{U}_{[0,1]}$ alors $(b-a)U + a \sim \mathcal{U}_{[a,b]}$. On passe ainsi d'un générateur de $\mathcal{U}_{[0,1]}$ à un générateur de $\mathcal{U}_{[a,b]}$.
2. Si $U_1, \ldots, U_d$ sont indépendantes de même loi $\mathcal{U}_{[0,1]}$ alors $(U_1, \ldots, U_d) \sim \mathcal{U}_{[0,1]^d}$.

[2] $http://www.library.cornell.edu/nr/nr\_index.cgi$

Voici un exemple de programme utilisant le générateur $rand()$ du C++.

```
#include<iostream.h> /* pour cout */
#include <stdlib.h> /* pour rand(),....*/
#include <time.h> /* pour time() */

void main( ) {

  int i ;

  srand((unsigned)time( NULL )) ; /*initialisation de la graine par l'horloge */

  /* Affiche 100 simulations d'une uniforme sur [0,1] */
  for( i = 0 ; i < 100 ;i ++ )
  cout << (float) rand()/(RAND_MAX ) << "\n" ; /* rand() renvoie un nombre
                                                   entier entre 0 et RAND_MAX */
  }
```

1.3 Autres lois

Cette section est consacrée à l'étude de différentes méthodes permettant de ramener la simulation d'une variable aléatoire quelconque à celle d'une variable aléatoire uniformément distribuée sur $[0,1]$. On suppose ici que l'on sait parfaitement simuler une suite de v.a. uniformément distribuées sur $[0,1]^d$ (ou sur $]0,1[^d$, voir Section 1.3.2).

1.3.1 Méthode d'inversion

La méthode d'inversion (de la fonction de répartition) est basée sur le résultat élémentaire suivant :

Proposition 1.3.1 *Soit Y une variable aléatoire réelle à valeurs dans $\mathbb{R}$ de fonction de répartition F_Y. On pose*

$$F_Y^{-1}(u) \ := \ \inf\{y \in \mathbb{R} \ : \ F_Y(y) \geq u\} \quad u \in [0,1] \, .$$

Si $U \sim \mathcal{U}_{dom(F_Y^{-1})}$ alors $F_Y^{-1}(U)$ et Y ont la même loi.

Preuve. Il suffit de remarquer que

$$\mathbb{P}\left[F_Y^{-1}(U) \leq y\right] \ = \ \mathbb{P}\left[U \leq F_Y(y)\right] \ = \ F_Y(y)$$

pour tout $y \in \mathbb{R}$. □

Remarque 1.3.2 Si Y a pour loi $\mathbb{P}[Y = y_k] = p_k$, $k \in \mathbb{N}$, alors, si $U \sim \mathcal{U}_{[0,1]}$

$$y_0 \mathbf{1}_{U \le p_0} + \sum_{i \ge 1} y_i \mathbf{1}_{\left\{\sum_{j=0}^{i-1} p_j < U \le \sum_{j=0}^{n} p_j\right\}}$$

a même loi que Y. Pour simuler Y, on simule donc $U \sim \mathcal{U}_{[0,1]}$ et on utilise la boucle

$p = p_0$; $j = 0$;
Tant que($p < U$) faire { $j = j + 1$; $p = p + p_j$; }
$Y = y_j$;

Evidemment cela peut être très coûteux si la loi de Y est très dispersée.

Remarque 1.3.3 Certains générateurs de nombres pseudo-aléatoire peuvent renvoyer la valeur 0. C'est par exemple le cas de la fonction $ran()$ en C. Si on n'y prend pas garde, on est alors amené à calculer $F_Y^{-1}(0)$ ce qui peut poser un problème à l'ordinateur et génèrer une erreur si cette quantité n'est pas définie (comme dans l'exemple ci-dessous). On verra dans la sous-section suivante comment gérer ce problème.

Exemple 1.3.4 *(Variable à support continu) Si Y suit une loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$, on a :*

$$F_Y(y) \quad = \quad 1 - e^{-\lambda y} \; .$$

Si $U \sim \mathcal{U}_{]0,1[}$, $-\lambda^{-1} \ln(1-U)$ et Y on donc la même loi. Comme $1-U$ suit également une loi $\mathcal{U}_{]0,1[}$, $-\lambda^{-1}\ln(U)$ suit également une $\mathcal{E}(\lambda)$. Pour simuler une loi $\mathcal{E}(\lambda)$, il suffit donc de simuler une variable $U \sim \mathcal{U}_{]0,1[}$ et de poser $Y = -\lambda^{-1}\ln(U)$. Pour simuler une suite $(Y_n)_{n\in\mathbb{N}}$ de variables indépendantes de loi $\mathcal{E}(\lambda)$, on simule une suite $(U_n)_{n\ge1}$ de variables indépendantes de loi $\mathcal{U}_{]0,1[}$ et on pose $Y_n = -\lambda^{-1}\ln(U_n)$, $n \ge 1$.

Exemple 1.3.5 *(Variable à support discret) Si Y suit une loi de Bernouilli de paramètre p, on tire $U \sim \mathcal{U}_{[0,1]}$ et on pose $Y = \mathbf{1}_{U \le p}$.*
Si Y suit une loi Binomiale de paramètre (n,p), c'est la somme de n variables de Bernouilli indépendantes de paramètre p. On pose donc

$$Y \quad = \quad \sum_{k=0}^{n} \mathbf{1}_{U_k \le p}$$

où les U_k sont i.i.d. de loi $\mathcal{U}_{[0,1]}$.

1.3.2 Méthode du rejet

a- Lois conditionnelles et lois uniformes sur un domaine

Proposition 1.3.6 *Soit $(Z_n)_{n\ge1}$ une suite de v.a. i.i.d. à valeurs dans $\mathbb{R}^d$ et soit $D \subset \mathbb{R}^d$ tel que $\mathbb{P}[Z_1 \in D] > 0$. On pose*

$$\begin{aligned}
\nu_1 &:= \inf\{k \ge 1 \;:\; Z_k \in D\} \\
\nu_{n+1} &:= \inf\{k > \nu_n \;:\; Z_k \in D\} \\
Y_n &:= Z_{\nu_n} \quad n \ge 1 \; .
\end{aligned}$$

Alors, $(Y_n)_{n\geq 1}$ est une suite de v.a. i.i.d. de loi ρ donnée par

$$\rho(A) \;=\; \mathbb{P}\left[Z_1 \in A \mid Z_1 \in D\right] \quad , \quad A \subset \mathbb{R}^d \; .$$

On vérifie facilement (voir preuve ci-dessous) que $\mathbb{E}\left[\nu_1\right] = \mathbb{E}\left[\nu_{n+1} - \nu_n\right] = 1/\alpha$ où $\alpha := \mathbb{P}\left[Z_1 \in D\right]$. Donc, plus α est petit et plus il y a de rejet, i.e. plus la simulation est coûteuse.

Preuve. Les (Z_n) étant indépendants et de même loi que Z_1, on a

$$\mathbb{P}\left[\nu_1 = k\right] \;=\; \mathbb{P}\left[Z_1 \notin D, \ldots, Z_{k-1} \notin D, Z_k \in D\right] \;=\; (1-\alpha)^{k-1}\alpha$$

où $\alpha := \mathbb{P}\left[Z_1 \in D\right]$. On calcule alors

$$\begin{aligned}
\mathbb{P}\left[Z_{\nu_1} \in A\right] &= \sum_{k\geq 1} \mathbb{P}\left[\nu_1 = k \;,\; Z_k \in A \cap D\right] \\
&= \sum_{k\geq 1} (1-\alpha)^{k-1} \mathbb{P}\left[Z_k \in A \cap D\right] \\
&= \mathbb{P}\left[Z_1 \in A \cap D\right] / \mathbb{P}\left[Z_1 \in D\right] \; .
\end{aligned}$$

On suppose maintenant que

$$\mathbb{P}\left[Z_{\nu_i} \in A_i, \; i = 1, \ldots, n-1\right] \;=\; \prod_{i=1}^{n-1} \mathbb{P}\left[Z_1 \in A_i \cap D\right] / \mathbb{P}\left[Z_1 \in D\right] \; .$$

Etant donné le résultat précédent, il suffit de montrer que cette propriété implique que

$$\mathbb{P}\left[Z_{\nu_i} \in A_i, \; i = 1, \ldots, n\right] \;=\; \prod_{i=1}^{n} \mathbb{P}\left[Z_1 \in A_i \cap D\right] / \mathbb{P}\left[Z_1 \in D\right] \; .$$

Or,

$$\begin{aligned}
&\mathbb{P}\left[Z_{\nu_i} \in A_i, \; i = 1, \ldots, n\right] \\
&= \sum_{j\geq 1, k\geq n-1} \mathbb{P}\left[Z_{\nu_1} \in A_1, \ldots, Z_{\nu_{n-1}} \in A_{n-1}, \nu_{n-1} = k, \nu_n = k+j, Z_{k+j} \in A_n\right] \\
&= \sum_{k\geq n-1} \mathbb{P}\left[Z_{\nu_1} \in A_1, \ldots, Z_{\nu_{n-1}} \in A_{n-1}, \nu_{n-1} = k\right] \sum_{j\geq 1} (1-\alpha)^{j-1} \mathbb{P}\left[Z_1 \in A_n \cap D\right] \\
&= \mathbb{P}\left[Z_{\nu_1} \in A_1, \ldots, Z_{\nu_{n-1}} \in A_{n-1}\right] \mathbb{P}\left[Z_1 \in A_n \cap D\right] / \mathbb{P}\left[Z_1 \in D\right] \; .
\end{aligned}$$

□

Corollaire 1.3.7 *Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes suivant une loi uniforme sur $\Delta = \prod_{i=1}^d [a_i, b_i] \subset \mathbb{R}^d$ et $D \subset \Delta$ tel que $\mathbb{P}\left[Z_1 \in D\right] > 0$. On pose*

$$\begin{aligned}
\nu_1 &:= \inf\{k \geq 1 \;:\; Z_k \in D\} \\
\nu_{n+1} &:= \inf\{k > \nu_n \;:\; Z_k \in D\} \\
Y_n &:= Z_{\nu_n} \quad n \geq 1 \; .
\end{aligned}$$

Alors, $(Y_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de même loi uniforme sur D.

Preuve. D'après la proposition précédente, $(Y_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de même loi ρ donnée par

$$\rho(A) \;=\; \mathbb{P}\left[Z_1 \in A \cap D\right]/\mathbb{P}\left[Z_1 \in D\right] = \frac{1}{\int_{\mathbb{R}^d} \mathbf{1}_D(z)dz} \int_{\mathbb{R}^d} \mathbf{1}_A(z)\mathbf{1}_D(z)dz\ .$$

Il en découle que Y_1 a pour densité $\mathbf{1}_D/(\int_{\mathbb{R}^d} \mathbf{1}_D(z)dz)$ sur $\mathbb{R}^d$. □

Exemple 1.3.8 *Lorsque l'on veut simuler une suite $(Y_n)_{n\geq 1}$ de v.a. indépendantes uniformséments distribuées sur $D = \{y \in \mathbb{R}^d \ :\ 0 < \|y\| < 1\}$, il suffit de simuler une suite de v.a indépendantes $(U_n)_{n\geq 1} \sim \mathcal{U}_{[-1,1]^d}$ et de poser*

$$\begin{aligned} \nu_1 &:= \inf\{k \geq 1 \ :\ 0 < \|U_k\| < 1\} \\ \nu_{n+1} &:= \inf\{k > \nu_n \ :\ 0 < \|U_k\| < 1\} \\ Y_n &:= U_{\nu_n} \quad n \geq 1\ . \end{aligned}$$

b- Lois à densité

La méthode de rejet est également utilisée lorsque l'on ne connaît pas l'inverse de la fonction de répartition de Y mais seulement sa densité f_Y. La proposition suivante est à l'origine d'une première méthode. Une seconde, plus simple à mettre en oeuvre, sera présentée dans la section suivante, voir Proposition 1.3.18.

Proposition 1.3.9 *Soit f une densité sur $\mathbb{R}^d$, $(Z_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de densité g sur $\mathbb{R}^d$, et soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes suivant une loi uniforme sur $[0,1]$, indépendantes de la suite $(Z_n)_{n\geq 1}$. On pose*

$$\begin{aligned} \nu_1 &:= \inf\{k \geq 1 \ :\ f(Z_k) > aU_k g(Z_k)\} \\ \nu_{n+1} &:= \inf\{k > \nu_n \ :\ f(Z_k) > aU_k g(Z_k)\} \\ Y_n &:= Z_{\nu_n} \quad n \geq 1\ , \end{aligned}$$

où a est un réel fixé vérifiant $f(z) \leq ag(z)$ pour tout $z \in \mathbb{R}^d$. Alors, la suite $(Y_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de même densité f.

Preuve. Utiliser la Proposition 1.3.6. □

Remarque 1.3.10 D'après l'exercice précédent, on a $\mathbb{P}\left[\nu_1 = k\right] = (1-\alpha)^{k-1}\alpha$ d'où $\mathbb{E}\left[\nu_1\right] = \alpha^{-1} = a$. Etant clair que $\nu_{n+1} - \nu_n$ a la même loi que ν_1 pour $n \geq 1$, on en déduit que $\mathbb{E}\left[\nu_{n+1} - \nu_n\right] = a$. Ceci montre que plus a est grand et plus le temps de simulation est grand (en moyenne). Il faut donc choisir a le plus petit possible.

Exemple 1.3.11 *Soit la densité définie par*

$$f(y) \;=\; \frac{2}{\pi}\sqrt{1-y^2}\mathbf{1}_{[-1,1]}(y)\;.$$

On peut alors prendre g définie par

$$g(z) \;=\; \frac{1}{2}\mathbf{1}_{[-1,1]}(z)$$

(loi uniforme sur $[-1,1]$) et $a = 4/\pi$.

1.3.3 Méthode par transformation

Cette méthode consiste à écrire une v.a. Y comme une fonction g d'une autre v.a. X facilement simulable. Elle repose sur la formule de changement de variables suivante :

Lemme 1.3.12 *(Formule de changement de variables) Soit ϕ un difféomorphisme d'un ouvert $D \subset \mathbb{R}^d$ sur un ouvert $\Delta \subset \mathbb{R}^d$, et soit g une fonction borélienne bornée de Δ dans $\mathbb{R}$, alors*

$$\int_\Delta g(v)dv \;=\; \int_D g(\phi(u))\left|\det(\nabla(\phi)(u))\right| du\;.$$

Preuve. Voir par exemple [50]. □

Comme Corollaire immédiat on obtient

Corollaire 1.3.13 *Soit X un vecteur aléatoire de densité f sur $\mathbb{R}^d$. On suppose que $X \in D$ p.s. où D est un ouvert de $\mathbb{R}^d$. Soit ψ un difféomorphisme de D sur un ouvert Δ. Alors $Y := \psi(X)$ a pour densité*

$$f(\psi^{-1}(\cdot))\left|\det(\nabla(\psi^{-1})(\cdot))\right|\mathbf{1}_\Delta(\cdot)\;.$$

Preuve. Soit h une fonction borélienne bornée de $\mathbb{R}^d$ dans $\mathbb{R}$. On calcule

$$\mathbb{E}\left[h(Y)\right] \;=\; \mathbb{E}\left[h(\psi(X))\right] \;=\; \int_D h(\psi(x))f(x)dx\;.$$

En applicant le Lemme 1.3.12 à $\phi := \psi^{-1}$ et $g := h\circ\psi$, on obtient :

$$\begin{aligned}\mathbb{E}\left[h(Y)\right] &= \int_\Delta h(\psi(\psi^{-1}(y)))f(\psi^{-1}(y))\left|\det(\nabla(\psi^{-1})(y))\right| dy\\ &= \int_\Delta h(y)f(\psi^{-1}(y))\left|\det(\nabla(\psi^{-1})(y))\right| dy\;.\end{aligned}$$

La fonction h étant quelconque, ceci conclut la preuve. □

a. Cas Gaussien

En utilisant cette formule et le fait que $\nabla(\psi^{-1})(\cdot) = \{\nabla(\psi)(\psi^{-1}(\cdot))\}^{-1}$, on obtient directement le résultat suivant sur lequel repose les deux principales méthodes de simulation de la loi normale : la méthodes de Box-Müller et l'algorithme polaire.

Proposition 1.3.14 *(Box-Müller) Soit* $(U,V) \sim \mathcal{U}_{]0,1[^2}$ *et*

$$X := \sqrt{-2\ln(U)}\cos(2\pi V) \quad , \quad Y := \sqrt{-2\ln(U)}\sin(2\pi V) \ .$$

Alors $(X,Y) \sim \mathcal{N}(0, I_2)$.

Cette formule permet donc de simuler deux v.a. gaussiennes centrées réduites indépendantes à partir de deux v.a. indépendantes uniformément distribuées sur $]0,1[$.

Remarque 1.3.15 Encore une fois, il faut faire attention aux valeurs retournées par le générateur de $\mathcal{U}_{]0,1[}$ utilisé. En général, les générateurs pré-programmés en C renvoient des valeurs dans $[0,1[$ et non $]0,1[$. Il faut donc contrôler qu'il ne renvoit pas 0 (sinon $\ln(0)$ provoque un bug). Ceci revient à utiliser la méthode du rejet, Exemple 1.3.8.

L'algorithme polaire est construit sur le même principe mais évite de faire appel aux fonctions trigonométriques qui peuvent être coûteuses en temps de calcul. Il nécessite par contre de mettre en oeuvre la méthode du rejet pour simuler une v.a. uniformément distribuée sur $D = \{(u,v) \in \mathbb{R}^2 \ : 0 < u^2 + v^2 < 1\}$, voir Exemple 1.3.8.

Proposition 1.3.16 *(Algorithme polaire) Soit* (U,V) *uniformément distribuée sur* $\{(u,v) \in \mathbb{R}^2 \ : 0 < u^2 + v^2 < 1\}$*, soit* $R^2 := U^2 + V^2$ *et*

$$X := U\sqrt{-2\ln(R^2)/R^2} \quad , \quad Y := V\sqrt{-2\ln(R^2)/R^2} \ .$$

Alors $(X,Y) \sim \mathcal{N}(0, I_2)$.

Preuve. On vérifie tout d'abord facilement que $(R^2, \theta/2\pi) \sim \mathcal{U}_{]0,1[^2}$ si (R,θ) sont les coordonnées polaires de $U + iV$ (utiliser le Corollaire 1.3.13). D'après la Proposition 1.3.14, le couple

$$\left(\sqrt{-2\ln(R^2)}\cos(\theta), \sqrt{-2\ln(R^2)}\sin(\theta)\right)$$

suit une gaussienne $\mathcal{N}(0, I_2)$. Or, $(\cos(\theta), \sin(\theta)) = (U/R, V/R)$ par construction. □

Il est facile à partir de ces deux méthodes de simuler une suite de vecteurs gaussiens indépendants de même loi $\mathcal{N}(0, I_d)$ puisque $Y = (Y^1, \ldots, Y^d)^* \sim \mathcal{N}(0, I_d)$ si et seulement si $(Y^i)_{i=1}^d$ est une suite de variables aléatoires indépendantes de même loi $\mathcal{N}(0,1)$.

Etant donné une matrice définie positive $\Gamma \in \mathbb{M}^d$ et un vecteur μ de $\mathbb{R}^d$, on simule également facilement un vecteur de loi $\mathcal{N}(\mu, \Gamma)$ en utilisant la procédure de factorisation de Cholesky.

Proposition 1.3.17 *Soit μ un vecteur de $\mathbb{R}^d$ et Γ une matrice carrée $d \times d$ définie positive.*
(1) Il existe $A \in \mathbb{M}^d$ telle que $AA^* = \Gamma$.
(2) Si $Y \sim \mathcal{N}(0, I_d)$, alors $X = \mu + AY \sim \mathcal{N}(\mu, \Gamma)$.

Pour simuler un vecteur gaussien de loi $\mathcal{N}(\mu, \Gamma)$ on procède donc ainsi. On commence par calculer la matrice A en utilisant l'algorithme de décomposition de Cholesky (voir par exemple [17]). On simule ensuite un vecteur $Y \sim \mathcal{N}(0, I_d)$ en utilisant la formule de Box-Müller ou l'algorithme polaire, puis, on calcule $\mu + AY$.

b. Loi de Poisson

Si $(T_k)_{k \geq 1}$ est une suite i.i.d. de loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$, alors

$$Y \quad := \quad \sum_{n=1}^{\infty} n \mathbf{1}_{\left\{\sum_{k=1}^{n} T_k \leq 1 \leq \sum_{k=1}^{n+1} T_k\right\}}$$

suit une loi de Poisson $\mathcal{P}(\lambda)$, i.e.

$$\mathbb{P}[Y = k] \quad = \quad \frac{\lambda^k}{k\,!} e^{-\lambda} \quad , \quad k \geq 0 \; .$$

On déduit de l'Exemple 1.3.4 que, si $(U_k)_{k \geq 1}$ est une suite i.i.d de loi $\mathcal{U}_{[0,1]}$, alors

$$Y \quad := \quad \sum_{n=1}^{\infty} n \mathbf{1}_{\left\{\prod_{k=1}^{n+1} U_k \leq e^{-1/\lambda} \leq \prod_{k=1}^{n} U_k\right\}}$$

admet comme loi $\mathcal{P}(\lambda)$.

c. Autres lois à densité : méthode du ratio de lois uniformes

On peut également utiliser cette approche pour écrire un algorithme proche de celui considéré dans la Section 1.3.2 pour les lois à densité mais plus simple à mettre en oeuvre dans la mesure où il ne nécessite pas de trouver une densité g et un réel a telle que $f \leq ag$, voir Proposition 1.3.9.

Proposition 1.3.18 *Soit f une densité sur $\mathbb{R}$ et deux réels $a_1 \in \mathbb{R}$ et $a_2 > 0$. On suppose que l'ensemble*

$$D_f \quad := \quad \left\{(u, v) \in \mathbb{R}^2 \;:\; 0 < u^2 \leq f\left(a_1 + a_2 \frac{v}{u}\right)\right\}$$

est borné. Alors, si (U, V) est uniformément distribuée sur D_f, $Z := a_1 + a_2 V/U$ a pour densité f.

Preuve. (U,V) ayant pour densité $\mathbf{1}_{D_f}/|D_f|$, on déduit du Corollaire 1.3.13 que (U^2,Z) a pour densité $f_{U^2,Z}(w,z) = \mathbf{1}_{0<w\leq f(z)}/(2a_2|D_f|)$. La densité de Z est donc $f/(2a_2|D_f|)$. Comme f est déjà une densité, on a forcément $2a_2|D_f| = 1$. □

On peux utiliser la méthode du rejet pour simuler le couple (U,V) uniformément distribué sur D, en partant d'une suite i.i.d. de couples uniformément distribués sur $[0,u^*]\times[v_*,v^*]\supset D$ où $u^* = \sup_{x\in\mathbb{R}} f(x)^{1/2}$, $v_* = \inf_{x\in\mathbb{R}}[(x-a_1)f(x)^{1/2}/a_2]$, $v^* = \sup_{x\in\mathbb{R}}[(x-a_1)f(x)^{1/2}/a_2]$. On peut remarquer que si $x^2f(x)$ et $f(x)$ sont bornés alors u^*, v_* et v^* sont bien finis et donc D_f est borné.

1.3.4 Variables corrélées et techniques par conditionnement

Variables corrélées

Si (X,Y) sont corrélées, on peut réécrire leur densité f sous la forme

$$f(x,y) \;=\; f_X(x)f(y \mid X=x)$$

où f_X est la loi de X et $f(\cdot \mid X=x)$ est la loi de Y sachant $X=x$. Pour simuler (X,Y), on commence donc par simuler X selon f_X puis on simule Y (indépendemment) selon la loi $f(\cdot \mid X=\bar{x})$ où $\bar{x}$ est la valeur prise par la simulation de X. Lorsque X et Y sont indépendantes, on a simplement $f(\cdot \mid X=\bar{x}) = f_Y$ et cela revient a simuler X et Y indépendemment, chacun selon sa loi marginale.

Techniques par conditionnement

1er cas. On suppose que la loi de Y s'écrit

$$f_Y(y) \;=\; \sum_{i\geq 0} p_i f_i(y)\ ,$$

où les f_i sont des densités et les p_i sont positifs (et donc ont une somme égale à 1 car f_Y est aussi une densité). On peut voir $f_Y(y)$ comme la densité marginale du couple (X,Y) où X a pour loi $\mathbb{P}[X=i]=p_i$ et Y a pour loi f_i conditionnellement à $\{X=i\}$. On peut donc procéder comme ci-dessus. Cela n'a évidemment d'intérêt que si l'on sait simuler les f_i.

Par exemple, si Y a pour densité

$$f_Y(y) \;=\; \alpha f_\sigma(y) + (1-\alpha)f_\gamma(y)$$

où $\alpha\in\,]0,1[$ et f_σ (resp. f_γ) est la densité de la loi $\mathcal{N}(0,\sigma^2)$ (resp. $\mathcal{N}(0,\gamma^2)$), on commence par tirer une loi uniforme U sur $[0,1]$. Si $U\leq\alpha$, on tire ensuite Y selon la loi $\mathcal{N}(0,\sigma^2)$. Si $U>\alpha$, on tire Y selon la loi $\mathcal{N}(0,\gamma^2)$. Ceci revient à poser $p_1=\alpha$, p_0

$= (1-\alpha)$, $f_1 = f_\sigma$ et $f_0 = f_\gamma$. X suit alors une loi de Bernouilli de paramètre α, i.e. $\mathbb{P}[X = 1] = \alpha$. On parle de mélange de gaussiennes.

2ème cas. On peut écrire la loi de Y sous la forme

$$f_Y(y) \;=\; \int g(y,x)dx \;,$$

où g est une fonction positive. Là encore, g est la densité d'un couple (X,Y) où Y a pour loi marginale f_Y. On peut donc commencer par simuler X selon sa loi marginale $f_X(x) = \int g(y,x)dy$ puis on simule Y selon $g(y,\bar{x})/f_X(\bar{x})$ où $\bar{x}$ est la valeur prise par la simulation de X.

1.4 Notion de discrépance

1.4.1 Généralités

Si $(U_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de loi uniforme sur $[0,1]$, $\mathcal{U}_{[0,1]}$, on a alors d'après la loi des grands nombres

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{1}_D\left((U_{kn}, U_{kn+1}, \ldots, U_{kn+k-1})\right) \;=\; \prod_{j=1}^{k}(b_j - a_j)$$

pour tout rectangles

$$D \;:=\;]a_1,b_1]\times]a_2,b_2]\times\cdots\times]a_k,b_k] \;\;,\;\; 0\leq a_j < b_j \leq 1 \;\; j = 1,\ldots,k$$

de $[0,1]^k$.

Il est donc naturel de chercher à générer des suites de nombres qui ont le même comportement. Ceci conduit à définir la notion de suite (réelle) uniforme.

Définition 1.4.1 *Une suite $(u_n)_{n\geq 1}$ dans $[0,1]^d$ est dite uniforme sur $[0,1]^d$ si pour tout $x \in [0,1]^d$ on a*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{d} \mathbf{1}_{u_n^i \leq x^i} \;=\; \prod_{i=1}^{d} x^i \;.$$

Il est important de ne pas confondre *suite de v.a. uniformément distribuées sur* $[0,1]^d$ et *suite uniforme sur* $[0,1]^d$. Une suite uniforme sur $[0,1]^d$ n'est pas forcément une suite de v.a. uniformément distribuées sur $[0,1]^d$. Par contre, il est clair, d'après la loi des grands nombres, qu'une suite de v.a. uniformément distribuées sur $[0,1]^d$ est presque surement uniforme.

Etant donnée une suite $u = (u_n)_{n\geq 1}$ uniforme sur $[0,1]^d$, on définit les discrépances

$$D_p^*(u,N) \;:=\; \|F(x) - F_N^u(x)\|_{L^p(\mathbb{R}^d,dx)} \quad p \in \mathbb{N}\cup\{+\infty\} \;,$$

où pour $x \in \mathbb{R}^d$ on a posé

$$F(x) \ := \ \prod_{i=1}^{d} x^i \quad \text{et} \quad F_N^u(x) = \frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{d}\mathbf{1}_{y^i \leq x^i} \ .$$

A $x \in \mathbb{R}^d$ fixé, le terme $F(x) - F_N^u(x)$ correspond à la différence entre le volume théorique du rectangle $\prod_{i=1}^d [0, x^i]$ et le volume "empiriquement" estimé par la suite u. Si u est uniforme sur $[0,1]^d$, alors cette différence doit tendre vers 0. La discrépance $D_p^*(u,N)$ mesure en quelque sorte la bonne répartition des points de la suite u dans l'espace $[0,1]^d$.

Théorème 1.4.2 *Il y a equivalence entre*
(i) u est uniforme sur $[0,1]^d$
(ii) $\lim_{N\to\infty} D_p^*(u,N) = 0$ pour tout $p \in \mathbb{N}\cup\{+\infty\}$
(iii) La mesure[3] $\mu_N^u := \frac{1}{N}\sum_{n=1}^N \delta_{u_n}$ converge étroitement[4] vers la mesure de Lebesgue sur $[0,1]^d$.

Preuve. Voir par exemple [13] □

L'intérêt de la notion de discrépance réside dans l'inégalité de Koksma-Hlawka. Avant d'énoncer ce résultat, on a besoin d'introduire la notion de fonction à variation finie. On note $\mathbf{1}_d$ le vecteur de $\mathbb{R}^d$ dont toutes les composantes valent 1.

Définition 1.4.3 *On dit que $f : [0,1]^d \mapsto \mathbb{R}$ est à variation finie si il existe une mesure (signée) ν sur $[0,1]^d$ (muni de la tribu borélienne associée) telle que $\nu(\{\mathbf{1}_d\}) = 0$ et*

$$f(x) \ = \ f(1) - \nu\left(\prod_{i=1}^{d}[x^i, 1]\right) \ = \ f(0) - \nu\left([0,1]^d - \prod_{i=1}^{d}[x^i,1]\right) \ .$$

On note alors $V(f) = |\nu|\left([0,1]^d\right)$ la variation totale de f.

Remarque 1.4.4 Pour $I \subset \{1,\ldots,d\}$ et $x \in \mathbb{R}^d$, on note x_I le vecteur de i-ème composante x^i si $i \in I$, 1 sinon. Si $f : [0,1]^d \mapsto \mathbb{R}$ vérifie[5]

$$\int_{[0,1]^{|I|}} \left| \frac{\partial f}{\partial x_{i_1} \ldots x_{i_{|I|}}} \right| dx_I \ < \ \infty$$

pour tout $I = \{i_1, \ldots, i_{|I|}\} \subset \{1,\ldots,d\}$, alors f est à variation finie.

[3] δ_x denote la mesure de Dirac en x

[4] Ceci signifie que $\int_{[0,1]^d} f d\mu_N^u = \frac{1}{N}\sum_{n=1}^N f(u_n)$ tend vers $\int_{[0,1]^d} f d\lambda_d$ pour toute fonction f continue bornée.

[5] les dérivées étant prises au sens des distributions.

Théorème 1.4.5 *(Inégalité de Koksma-Hlawka) Soit $f : [0,1]^d \mapsto \mathbb{R}$ à variation finie. Pour toute suite $u = (u_n)_{n\geq 1}$ dans $[0,1]^d$ on a*

$$\left| \int_{[0,1]^d} f(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(u_n) \right| \;\leq\; V(f) D^*_\infty(u,N) \;.$$

Pour calculer $\mathbb{E}[f(U)]$, $U \sim \mathcal{U}_{[0,1]^d}$, on peut donc utiliser une suite $(u_n)_{n\geq 1}$ uniforme sur $[0,1]^d$ et calculer $\frac{1}{N}\sum_{n=1}^{N} f(u_n)$. Plus la discrépance de la suite sera faible et meilleur sera l'approximation de $\mathbb{E}[f(U)]$. C'est une alternative aux méthodes de Monte-Carlo présentées ci-dessus. Lorsque la suite $(u_n)_{n\geq 1}$ est purement déterministe, comme c'est le cas pour les suites à discrépance faible présentées ci-dessous, on parle de Quasi-nombres aléatoires.

1.4.2 Cas des suites uniformément distribuées

Si $(U_n)_{n\geq 1}$ est une suite de v.a. i.i.d. uniformément distribuées sur $[0,1]^d$, elle est p.s. uniforme sur $[0,1]^d$. On peut donc calculer sa discrépance $D^*_p(U,N)$ qui est alors une quantité aléatoire.

D'après la loi du logarithme itéré, voir Theorem 1.1.2, on peut déjà calculer que

$$\begin{aligned}
\limsup_{N\to\infty} \sqrt{\frac{N}{2\ln\ln N}} D^*_\infty(U,N) \;&\geq\; \sup_{x\in[0,1]^d} \limsup_{N\to\infty} \sqrt{\frac{N}{2\ln\ln N}} \left(F^U_N(x) - F(x)\right) \\
&=\; \sup_{x\in[0,1]^d} \sqrt{F(x)(1-F(x))} \\
&=\; \frac{1}{2} \;.
\end{aligned}$$

On a en fait

Théorème 1.4.6 *Si $(U_n)_{n\geq 1}$ est une suite de v.a. i.i.d. uniformément distribuées sur $[0,1]^d$, alors*

$$\limsup_{N\to\infty} \sqrt{\frac{N}{2\ln\ln N}} D^*_\infty(U,N) \;=\; \frac{1}{2} \;.$$

Par ailleurs,

$$\mathbb{E}\left[D^*_\infty(U,N)\right] \;=\; \frac{c^d_\infty}{\sqrt{N}} \;,$$

où $c^d_\infty \to \infty$ lorsque $d \to \infty$.

Ce résultat montre en particulier que

$$D^*_\infty(U,N) \;\leq\; O\left(\sqrt{\frac{\ln\ln N}{N}}\right) \quad \text{p.s.}$$

1.4.3 Suites à discrépance faible

On peut contruire des suites dans $[0,1]^d$ dont la discrépance est plus faible que celle obtenue pour les suites de v.a. uniformément distribuées sur $[0,1]^d$. Il existe une vaste littérature sur ce sujet et nous n'allons présenté ici que les plus simples à mettre en oeuvre.

Il faut bien garder en mémoire que si ces suites ont une discrépances plus faible asymtotiquement lorsque N tend vers l'infini, celle-ci dépend généralement de la dimension de manière non-triviale et peut devenir très grande, à N fixé, lorsque la dimension augmente. En grande dimension, il est généralement préférable d'utiliser des suites i.i.d. uniformément distribuées sur $[0,1]^d$ pour lesquelles le comportement asymptotique (p.s.) ne dépend pas de la dimension (voir Théorème 1.4.6).

On pourra consulter [57] pour une étude comparative de différentes suites et pour plus de référence. Voir également [13].

a- Suite de Halton

Etant donné un nombre premier $p \geq 2$ et un entier n, il existe une unique décomposition (décomposition p-adique) de n sous la forme :

$$n \;=\; a_0^{n,p} + a_1^{n,p} p + \ldots + a_{k_{n,p}}^{n,p} p^{k_{n,p}}$$

où $k_{n,p}$ est un entier et les $a_k^{n,p}$ sont des entiers vérifiant $0 \leq a_k^{n,p} \leq p-1$ avec $a_{k_{n,p}}^{n,p} \neq 0$. Cette décomposition s'obtient facilement en utilisant la récurrence

$$\begin{aligned}
r_0^{n,p} &= n \\
a_k^{n,p} &= r_k^{n,p} \bmod p \ , \ \ r_k^{n,p} \ = \ (r_{k-1}^{n,p} - a_{k-1}^{n,p})/p \quad 1 \leq k \leq k_{n,p} \\
\text{avec } k_{n,p} &= \min\{k \geq 1 \ : \ r_{k+1}^{n,p} = 0\} \ .
\end{aligned}$$

On note ensuite

$$\Phi_p(n) \;:=\; \frac{a_0^{n,p}}{p} + \ldots + \frac{a_{k_{n,p}}^{n,p}}{p^{k_{n,p}+1}} \ .$$

On construit la suite de Halton $u = (u_n)_{n \geq 1}$ de la manière suivante. Soit $p_1, \ldots, p_d$ les d premiers nombres premiers. Pour chaque $n \geq 1$ et $1 \leq i \leq d$, on pose :

$$u_n^i \;=\; \Phi_{p_i}(n) \ .$$

Pour cette suite, on a :

$$D_\infty^*(u,N) \;\leq\; O\left(\frac{\ln(N)^d}{N}\right) \ ,$$

voir par exemple [57]. C'est évidemment mieux que ce que l'on obtient avec une suite de v.a. uniformément distribuées sur $[0,1]^d$ mais la majoration dépend de la dimension, ce qui n'est pas le cas pour les suites de v.a. uniformément distribuées sur $[0,1]^d$ (sauf en moyenne). En pratique, c'est beaucoup moins efficace en grande dimension.

Pour $d=1$, on parle de suite de Van der Corput.

b- Suite SQRT

Pour tout réel x, on pose $\mathrm{Ent}[x]$ sa partie entière. Soit $p_1, \ldots, p_d$ les d premiers nombres entiers. Pour tout $n \geq 1$ et $1 \leq i \leq d$, on pose $u_n^i = n\sqrt{p_i} - \mathrm{Ent}[n\sqrt{p_i}]$. On a alors

$$D_\infty^*(u, N) \quad = \quad O\left(\frac{1}{N^{1-\varepsilon}}\right) \qquad \text{pour tout } \varepsilon > 0 \,.$$

c- Suite de Faure

Soit $p_{*,d}$ le plus petit nombre premier impair plus grand que d. On définit l'application $\mathcal{Q}$ qui a un rationel x de la forme

$$x \quad = \quad \sum_{i \geq 0} \frac{x_i}{p_{*,d}^{i+1}}$$

associe

$$\mathcal{Q}(x) \quad = \quad \sum_{i \geq 0} \frac{\sum_{j \geq i} \binom{j}{i} x_j \bmod p_{*,d}}{p_{*,d}^{i+1}}$$

et on note $\mathcal{Q}_k$ l'application obtenue en composant k fois $\mathcal{Q}$ avec la convention que $\mathcal{Q}_0$ est l'identité. On définit alors la suite par

$$u_n^i \quad = \quad \mathcal{Q}_{i-1}(n-1) \qquad 1 \leq n \text{ et } 1 \leq i \leq d.$$

On a

$$D_\infty^*(u, N) \quad \leq \quad \frac{1}{N}\left[\frac{1}{d!}\left(\frac{p_{*,d}-1}{2}\right)^d \left(\frac{\ln(2N)}{\ln(p_{*,d})} + d + 1\right)^d\right]$$

(voir [65]).

Chapitre 2

Méthodes de discrétisation

La base des méthodes de Monte-Carlo en finance est la capacité à simuler l'évolution des processus de prix. On commence donc par s'intéresser aux méthodes de simulation d'une diffusion de la forme :

$$dX_t \;=\; b(X_t)dt + a(X_t)dW_t \quad , \quad X_0 \;=\; x \;\in\; \mathbb{R}^d \, . \tag{2.0.1}$$

Ici, X modélise l'évolution de d sous-jacents sur le marché (actions par exemples) et W est un mouvement brownien standard d-dimensionnel sur un espace de probabilité filtré complet $(\Omega, \mathbb{F}, \mathbb{P})$. On suppose que $\mathbb{F} = \{\mathcal{F}_t,\ t \in [0,T]\}$ est la filtration naturelle complétée de W.

En général, on supposera que a et b sont lipschitziennes, i.e. qu'il existe $K > 0$ tel que

$$\|a(x) - a(y)\| + \|b(x) - b(y)\| \;\leq\; K\,\|x - y\| \quad , \quad \text{pour tout } x,\ y \in \mathbb{R}^d \, . \tag{2.0.2}$$

Cette hypothèse garantit l'existence d'une solution forte à (2.0.1).

2.1 Modèle de Black-Scholes

Dans le modèle de Black-Scholes, la dynamique du prix sous la probabilité risque neutre s'écrit

$$dX_t \;=\; rX_t dt + \mathrm{diag}\,[X_t]\,\sigma dW_t \quad , \quad X_0 \;\in\; \mathbb{R}^d \, , \tag{2.1.1}$$

où $r \in \mathbb{R}_+$ est le taux sans risque et $\sigma \in \mathbb{M}^d$ est la matrice de volatilité (supposée inversible pour assurer la complétude du marché). Ce système se ré-écrit sous la forme :

$$dX^i_t \;=\; rX^i_t + X^i_t \sigma^{i\cdot} dW_t \quad , \quad X^i_0 \;\in\; \mathbb{R}\, , \;\; i = 1, \ldots, d \, , \tag{2.1.2}$$

dont la solution est

$$X^i_t \;=\; X^i_0 \exp\left\{\left(r - \frac{1}{2}\left\|\sigma^{i\cdot}\right\|^2\right) t + \sigma^{i\cdot} W_t\right\} \, , \;\; i = 1, \ldots, d \, . \tag{2.1.3}$$

Pour simuler X_t, il suffit donc de simuler la valeur de W en t. Les W^i_t, $i = 1, \ldots, d$, étant indépendants et de même loi $\mathcal{N}(0,t)$, il suffit donc de simuler d variables aléatoires indépendantes de loi $\mathcal{N}(0,t)$.

2.2 Schéma de discrétisation

2.2.1 Discrétisation d'Euler

2.2.1.1 Construction

La simulation du processus X dans le modèle de Black-Scholes est "exacte" dans la mesure où, à t donné, on peut simuler exactement dans la loi de X_t. Cela est possible grâce à la forme particulièrement simple de (2.1.2).

En général, la solution de (2.0.1) n'a pas une forme aussi simple et l'on est obligé de recourir à une approximation de X qui correspond à une discrétisation en temps de l'équation (2.0.1).

On se donne une grille de $[0, T]$

$$\pi^n \ := \ (t_0, \ldots, t_i, \ldots, t_n) \quad \text{avec} \quad t_i \ = \ iT/n \ .$$

On notera $\Delta^n t = T/n$ et $\Delta^n W_{i+1} = W_{t_{i+1}} - W_{t_i}$. L'idée de la discrétisation d'Euler est très simple. Pour tout $i = 1, \ldots, n$, on a :

$$\begin{aligned} X_{t_i} &= X_{t_{i-1}} + \int_{t_{i-1}}^{t_i} b(X_s)ds + \int_{t_{i-1}}^{t_i} a(X_s)dW_s \\ &\backsimeq X_{t_{i-1}} + b(X_{t_{i-1}})\Delta^n t + a(X_{t_{i-1}})\Delta^n W_i \end{aligned}$$

ce qui conduit à la construction du schéma

$$\bar{X}_0^n \ = \ X_0 \tag{2.2.1}$$

$$\bar{X}_{t_i}^n \ = \ \bar{X}_{t_{i-1}}^n + b(\bar{X}_{t_{i-1}}^n)\Delta^n t + a(\bar{X}_{t_{i-1}}^n)\Delta^n W_i \ , \quad i = 1, \ldots, n \ . \tag{2.2.2}$$

C'est l'équivalent stochastique du schéma d'Euler utilisé pour les équations différentielles ordinaires.

La simulation de $\bar{X}^n$ se ramène à la simulation des accroissements de W, où $\Delta^n W_i \sim \mathcal{N}(0, \Delta^n t I_d)$ pour tout $i = 1, \ldots, n$.

On introduit maintenant un schéma d'Euler "continu". Pour cela on définit la fonction

$$\phi_t^n \ := \ \max\{t_i, \ i = 0, \ldots, n, \ \ \text{tel que } t_i \leq t\} \quad \text{sur } [0, T] \ ,$$

et, à $t \in [0, T]$, on associe

$$\bar{X}_t^n \ = \ \bar{X}_{\phi_t^n}^n + b(\bar{X}_{\phi_t^n}^n)(t - \phi_t^n) + a(\bar{X}_{\phi_t^n}^n)\left(W_t - W_{\phi_t^n}\right) \ , \tag{2.2.3}$$

ce que l'on peut réécrire sous forme intégrale

$$\bar{X}_t^n \ = \ X_0 + \int_0^t b(\bar{X}_{\phi_s^n}^n)ds + \int_0^t a(\bar{X}_{\phi_s^n}^n)dW_s \ . \tag{2.2.4}$$

2.2.1.2 Convergence L^p

A n fixé, une simple récurrence montre que, sous l'hypothèse (2.0.2), $\bar{X}^n_{t_i} \in L^p$ pour tout $i \in \{1, \ldots, n\}$ et $p \geq 1$. On vérifie maintenant que ceci est vrai uniformément en n.

Lemme 2.2.1 *Sous l'hypothèse* (2.0.2), *pour tout* $p \geq 1$

$$\sup_{n\geq 1} \mathbb{E}\left[\sup_{t\in[0,T]} \left\|\bar{X}^n_t\right\|^p\right]^{1/p} \;<\; \infty\ .$$

Preuve. On peut toujours supposer que p est pair et plus grand que 4 en utilisant l'inégalité des normes $\|\cdot\|_{L^p} \leq \|\cdot\|_{L^{2p}}$. On commence par utiliser l'inégalité de Jensen appliquée à l'application convexe $x \mapsto \|x\|^p$ pour écrire que, pour tout $t \in [0, T]$,

$$\begin{aligned}
\sup_{s\leq t} \left\|\bar{X}^n_s\right\|^p &= 3^p \sup_{s\leq t} \left\| \frac{1}{3} \left(X_0 + \int_0^s b(\bar{X}^n_{\phi^n_r})dr + \int_0^s a(\bar{X}^n_{\phi^n_r})dW_r \right) \right\|^p \\
&\leq 3^{p-1} \left(\|X_0\|^p + \sup_{s\leq t} \left\| \int_0^s b(\bar{X}^n_{\phi^n_r})dr \right\|^p + \sup_{s\leq t} \left\| \int_0^s a(\bar{X}^n_{\phi^n_r})dW_r \right\|^p \right)\ .
\end{aligned}$$

En utilisant de nouveau l'inégalité de Jensen appliquée à l'application convexe $x \mapsto \|x\|^p$, on obtient

$$\left\| \int_0^s b(\bar{X}^n_{\phi^n_r})dr \right\|^p \;=\; s^p \left\| \int_0^s \frac{1}{s}\, b(\bar{X}^n_{\phi^n_r})dr \right\|^p \;\leq\; s^{p-1} \int_0^s \left\|b(\bar{X}^n_{\phi^n_r})\right\|^p dr\ .$$

En utilisant ensuite (2.0.2), on déduit que

$$\begin{aligned}
\left\| \int_0^s b(\bar{X}^n_{\phi^n_r})dr \right\|^p &\leq s^{p-1} \int_0^s \left\| \|b(X_0)\| + K \left\|\bar{X}^n_{\phi^n_r}\right\| + K\|X_0\| \right\|^p dr \\
&\leq C \left(1 + \int_0^s \left\|\bar{X}^n_{\phi^n_r}\right\|^p dr \right)\ .
\end{aligned}$$

Comme, à n fixé, $\bar{X}^n_{t_i} \in L^p$ pour tout i, on vérifie en utilisant (2.0.2) que $(\int_0^t a(\bar{X}^n_{\phi^n_r})dW_r)_{t\leq T}$ est une martingale continue de carré intégrable. En utilisant l'inégalité de Burkholder-Davis-Gundy (Lemme 2.2.2 ci-après), l'inégalité de Jensen et (2.0.2), on obtient alors que

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\leq t} \left\| \int_0^s a(\bar{X}^n_{\phi^n_r})dW_r \right\|^p \right] &\leq \mathbb{E}\left[\left(\int_0^t \left\|a(\bar{X}^n_{\phi^n_s})\right\|^2 ds \right)^{p/2} \right] \\
&\leq \mathbb{E}\left[t^{p/2-1} \int_0^t \left(\left\|a(\bar{X}^n_{\phi^n_s})\right\|^2 \right)^{p/2} ds \right] \\
&\leq C \left(1 + \int_0^t \mathbb{E}\left[\left\|\bar{X}^n_{\phi^n_s}\right\|^p \right] ds \right)\ .
\end{aligned}$$

On déduit des inégalités précédentes que

$$\mathbb{E}\left[\sup_{s\leq t} \left\|\bar{X}^n_s\right\|^p\right] \;\leq\; C\left(1 + \int_0^t \mathbb{E}\left[\left\|\bar{X}^n_{\phi^n_s}\right\|^p\right] ds\right) \;\leq\; C\left(1 + \int_0^t \mathbb{E}\left[\sup_{r\leq s} \left\|\bar{X}^n_r\right\|^p\right] ds\right)\ .$$

Il suffit maintenant d'appliquer le Lemme de Gronwall (Lemme 2.2.3 ci-après) à $s \longmapsto \mathbb{E}\left[\sup_{r\leq s}\left\|\bar{X}_r^n\right\|^p\right]$ pour conclure. $\square$

On donne maintenant les deux Lemmes techniques que nous avons utilisés (voir par exemple [37]).

Lemme 2.2.2 *(Inégalités de Burkholder-Davis-Gundy) Soit $(M_t)_{t\in[u,v]}$ une martingale continue sur $[u,v]$ de carré intégrable. Alors, pour tout $m>0$, il existe k_m et $K_m>0$, ne dépendant que de m, tels que*

$$k_m\mathbb{E}\left[(\langle M\rangle_v)^m\right] \leq \mathbb{E}\left[\sup_{u\leq t\leq v}\|M_t\|^{2m}\right] \leq K_m\mathbb{E}\left[(\langle M\rangle_v)^m\right] \ .$$

Lemme 2.2.3 *(Lemme de Gronwall) Soit g une fonction continue positive vérifiant*

$$g(t) \leq \alpha(t) + \beta\int_0^t g(s)ds \quad \forall\, t\in[0,T],$$

avec $\beta\geq 0$ et $\alpha : [0,T]\mapsto\mathbb{R}$ intégrable. Alors,

$$g(t) \leq \alpha(t) + \beta\int_0^t \alpha(s)e^{\beta(t-s)}ds \quad \forall\, t\in[0,T].$$

On étudie maintenant le comportement des incréments de $\bar{X}^n$ entre deux dates de discrétisation.

Lemme 2.2.4 *Sous* (2.0.2), *pour tout $p\geq 1$*

$$\max_{0\leq i\leq n-1}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1})}\left\|\bar{X}_t^n-\bar{X}_{\phi_t^n}^n\right\|^p\right]^{1/p} \quad\leq\quad C/\sqrt{n}\ .$$

Preuve. On pose $p\geq 4$ et pair sans perte de généralité. En argumentant comme dans le lemme précédent, on montre que, pour tout $i\in\{0,\ldots,n-1\}$,

$$\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1})}\left\|\bar{X}_t^n-\bar{X}_{\phi_t^n}^n\right\|^p\right] \quad\leq\quad C\mathbb{E}\left[n^{-p/2+1}\int_{t_i}^{t_{i+1}}\left\|b(\bar{X}_{\phi_s^n}^n)\right\|^p+\left\|a(\bar{X}_{\phi_s^n}^n)\right\|^p ds\right]$$

Ceci implique, d'après (2.0.2) et le Lemme 2.2.1 que

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1})}\left\|\bar{X}_t^n-\bar{X}_{\phi_t^n}^n\right\|^p\right] &\leq C\left(n^{-p/2+1}\int_{t_i}^{t_{i+1}}C\left(1+\mathbb{E}\left[\left\|\bar{X}_{\phi_s^n}^n\right\|^p\right]\right)ds\right)\\ &\leq Cn^{-p/2}\ ,\end{aligned}$$

ce qui conclut la preuve. $\square$

Remarque 2.2.5 Les mêmes arguments que ceux de la preuve précédente montrent que, sous (2.0.2), pour tout $p \geq 1$

$$\max_{0 \leq i \leq n-1} \mathbb{E}\left[\sup_{t \in [t_i, t_{i+1})} \left\|X_t - X_{\phi_t^n}\right\|^p\right]^{1/p} \quad \leq \quad C/\sqrt{n}\;.$$

On peut maintenant étudier la vitesse de convergence L^p du schéma d'Euler.

Théorème 2.2.6 *Sous l'hypothèse* (2.0.2), *pour tout* $p \geq 1$

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left\|\bar{X}_t^n - X_t\right\|^p\right]^{1/p} \quad \leq \quad C/\sqrt{n}\;.$$

Preuve. On pose $p \geq 4$ et pair sans perte de généralité. En argumentant comme dans le lemme précédent, on montre que

$$\mathbb{E}\left[\sup_{s \in [0,t]} \left\|\bar{X}_s^n - X_s\right\|^p\right] \quad \leq \quad C\mathbb{E}\left[\int_0^t \left\|\bar{X}_{\phi_s^n}^n - X_s\right\|^p ds\right]\;.$$

On utilise maintenant le Lemme 2.2.4

$$\begin{aligned}
\mathbb{E}\left[\sup_{s \in [0,t]} \left\|\bar{X}_s^n - X_s\right\|^p\right] \quad &\leq \quad C\mathbb{E}\left[\int_0^t \left\|\bar{X}_s^n - X_s\right\|^p + \left\|\bar{X}_{\phi_s^n}^n - \bar{X}_s^n\right\|^p ds\right] \\
&\leq \quad C\left(n^{-p/2} + \int_0^t \mathbb{E}\left[\sup_{r \leq s} \left\|\bar{X}_r^n - X_r\right\|^p\right] ds\right)
\end{aligned}$$

et on conclut en utilisant le Lemme de Gronwall. □

Remarque 2.2.7 On considère un payoff qui dépend de la trajectoire de X en un nombre fini de dates $(s_1, \ldots, s_k)$, $k \geq 1$. Si g est lipschitzienne, on obtient comme corollaire du Théorème 2.2.6

$$\mathbb{E}\left[\left\|g\left(\bar{X}_{\phi_{s_1}^n}^n, \ldots, \bar{X}_{\phi_{s_k}^n}^n\right) - g\left(X_{s_1}, \ldots, X_{s_k}\right)\right\|^p\right]^{1/p} \quad \leq \quad C/\sqrt{n}\;.$$

Exemple 2.2.8 *La remarque précédente montre que le schéma d'Euler permet d'approximer correctement une moyenne discrète*

$$A_{(s_i)_{i=1}^k} \;:=\; \frac{1}{k}\sum_{i=1}^k X_{s_i} \quad \textit{par} \quad \bar{A}^n_{(s_i)_{i=1}^k} \;:=\; \frac{1}{k}\sum_{i=1}^k \bar{X}^n_{\phi^n_{s_i}}\;.$$

pour évaluer les options sur moyenne du type $[A_{(s_i)_{i=1}^k} - K]^+$ *(call sur moyenne de strike* K*),* $[X_T - A_{(s_i)_{i=1}^k}]^+$ *(call avec strike flottant), et les puts correspondants. Le Théorème 2.2.6 nous assure une convergence* L^p *en* $1/\sqrt{n}$*.*
De la même manière, on peut approximer correctement le maximum discret

$$M_{(s_i)_{i=1}^k} \;:=\; \max\left\{X_{s_i},\; i \in \{1, \ldots, k\}\right\} \quad \textit{par} \quad \bar{M}^n_{(s_i)_{i=1}^k} \;:=\; \max\left\{\bar{X}^n_{\phi^n_{s_i}},\; i \in \{1, \ldots, k\}\right\}\;.$$

2.2.1.3 Convergence faible

En pratique, ce qui intéresse surtout les praticiens c'est la convergence du prix "discrétisé" vers le vrai prix, i.e. la convergence de $\mathbb{E}g(\bar{X}_T^n)-\mathbb{E}g(X_T)$ vers 0. C'est la convergence faible.

Théorème 2.2.9 *Si b, $a \in C_b^4$ et $g \in C_p^4$, alors*

$$\left\|\mathbb{E}\left[g(\bar{X}_T^n) - g(X_T)\right]\right\| \;\; \leq \;\; C/n \, .$$

Ce résultat est du à [62]. On va le démontrer dans un cadre simplifié. Pour cela, on va partiellement utiliser le Théorème de Feynman-Kac que l'on commence par rappeler, voir [37] et [43].

Théorème 2.2.10 *(i) Supposons que b et a sont lipschitziennes et qu'il existe une solution $u \in C_p^{1,2}$ à l'équation parabolique*

$$\begin{aligned} 0 &= \frac{\partial u}{\partial t} + \sum_{j=1}^{d} b^j \frac{\partial u}{\partial x^j} + \frac{1}{2}\sum_{i,j=1}^{d} (aa^*)^{ij} \frac{\partial^2 u}{\partial x^i \partial x^j} \, , \quad \textit{sur } [0,T] \times \mathbb{R}^d \\ g &= u(T,\cdot) \quad \textit{sur } \mathbb{R}^d \, , \end{aligned} \tag{2.2.5}$$

alors

$$u(t,x) = \mathbb{E}\left[g(X_T) \mid X_t = x\right] \quad \textit{et} \quad g(X_T) = \mathbb{E}\left[g(X_T)\right] + \int_0^T \nabla_x u(t,X_t) a(X_t) dW_t \, . \tag{2.2.6}$$

(ii) Si $a \in C_b^4$, $b \in C_b^2$ et $g \in C_p^4$, alors la réciproque est vraie et $u \in C_p^4$.

Eléments de preuve du Théorème 2.2.9. Pour simplifier, on suppose que b et a sont bornées, que $d = 1$ et que la solution u de (2.2.5) est C_b^∞. On a alors par Itô sur le schéma d'Euler continu

$$\begin{aligned} \mathbb{E}\left[g(\bar{X}_T^n) - g(X_T)\right] &= \mathbb{E}\left[u(T,\bar{X}_T^n)\right] - u(0,X_0) \\ &= \mathbb{E}\left[\int_0^T u_t(s,\bar{X}_s^n) + b(\bar{X}_{\phi_s^n}^n) u_x(s,\bar{X}_s^n) + \frac{1}{2} a^2(\bar{X}_{\phi_s^n}^n) u_{xx}(s,\bar{X}_s^n) ds\right] \, . \end{aligned}$$

Puisque u est solution de (2.2.5), on a également

$$\begin{aligned} \mathbb{E}\left[g(\bar{X}_T^n) - g(X_T)\right] &= \mathbb{E}\left[\int_0^T u_t(s,\bar{X}_s^n) - u_t(s,\bar{X}_{\phi_s^n}^n) + b(\bar{X}_{\phi_s^n}^n)\left(u_x(s,\bar{X}_s^n) - u_x(s,\bar{X}_{\phi_s^n}^n)\right) ds\right] \\ &\quad + \mathbb{E}\left[\int_0^T \frac{1}{2} a^2(\bar{X}_{\phi_s^n}^n)\left(u_{xx}(s,\bar{X}_s^n) - u_{xx}(s,\bar{X}_{\phi_s^n}^n)\right) ds\right] \, . \end{aligned}$$

On considère maintenant le premier terme. Comme $u \in C_b^\infty$, a et b sont bornées et $\mathbb{E}\left[\|\bar{X}_t^n\|\right]$ est uniformément borné en $t \in [0,T]$ et n (Lemme 2.2.1), on obtient, en utilisant le Lemme d'Itô, que pour tout $s \in [0,T]$

$$\begin{aligned}\left\|\mathbb{E}\left[u_t(s,\bar{X}_s^n) - u_t(s,\bar{X}_{\phi_s^n}^n)\right]\right\| &= \left\|\mathbb{E}\left[\int_{\phi_s^n}^s b(\bar{X}_{\phi_r^n}^n)u_{tx}(s,\bar{X}_r^n) + \frac{1}{2}a^2(\bar{X}_{\phi_r^n}^n)u_{txx}(s,\bar{X}_r^n)dr\right]\right\| \\ &\leq C/n \;.\end{aligned}$$

On obtient le même type de majoration pour les autres termes, ce qui implique que

$$\left\|\mathbb{E}\left[g(\bar{X}_T^n) - g(X_T)\right]\right\| \;\leq\; C/n \;,$$

et conclut la preuve. □

On obtient donc une vitesse de convergence faible en $1/n$. Evidemment, l'hypothèse $g \in C_p^4$ n'est pas très satisfaisante car généralement non vérifiée en finance. Le résultat précédent peut être amélioré de la manière suivante.

Théorème 2.2.11 *Si b, $a \in C_b^\infty$ et si l'une des deux conditions suivantes est vérifiée :*
(i) $g \in C_p^\infty$,
(ii) *g est à croissance polynomiale et il existe $\varepsilon > 0$ tel que,*

$$\forall\ \zeta,\, x \in \mathbb{R}^d,\ \ \zeta^*(aa^*)(x)\zeta \;\geq\; \varepsilon\,\|\zeta\| \;, \tag{2.2.7}$$

alors

$$\mathbb{E}g(\bar{X}_T^n) - \mathbb{E}g(X_T) \;=\; \frac{C_1}{n} + \ldots + \frac{C_k}{n^k} + O(n^{-k-1}) \;,\ \ k \in \mathbb{N} \;.$$

Remarque 2.2.12 Ce résultat est très important en pratique car il permet de ramener la vitesse de convergence en $1/n^2$. En effet, d'après le théorème précédent, on a

$$\mathbb{E}\left[2g(\bar{X}_T^{2n}) - g(\bar{X}_T^n)\right] - \mathbb{E}g(X_T) \;=\; \frac{2C_1}{2n} - \frac{C_1}{n} + \frac{2C_2}{4n^2} - \frac{C_2}{n^2} + O(1/n^3) \;=\; O(1/n^2) \;.$$

En utilisant $3n$ pas, on obtient une erreur en $O(1/n^2)$. Cette technique est connue sous le nom d'extrapolation de Romberg. On peut remarquer que l'hypothèse (ii) du théorème est souvent satisfaite en finance. Dans le modèle de Black-Scholes, on pourra par exemple considérer le ln de la diffusion pour se ramener au cas où $a(x)$ est constante et satisfait l'hypothèse d'uniforme ellipticité (2.2.7).

2.2.2 Schéma de Milshtein

Dans la présentation du schéma d'Euler, on a utilisé l'approximation

$$\int_{t_{i-1}}^{t_i} a(X_s)dW_s \;\sim\; a(X_{t_{i-1}})\left(W_{t_i} - W_{t_{i-1}}\right)$$

mais on aurait pû utiliser une approximation d'ordre supérieur. Pour fixer les idées, on se place en dimension 1 et on suppose que $b=0$. Alors, pour $s \in (t_{i-1}, t_i]$

$$
\begin{aligned}
a\left(X_s\right) &= a\left(X_{t_{i-1}}+\int_{t_{i-1}}^s a(X_t)dW_t\right) \\
&\sim a\left(X_{t_{i-1}}+a(X_{t_{i-1}})\left(W_s-W_{t_{i-1}}\right)\right) \\
&\sim a\left(X_{t_{i-1}}\right)+a'\left(X_{t_{i-1}}\right)a(X_{t_{i-1}})\left(W_s-W_{t_{i-1}}\right) \ .
\end{aligned}
$$

On utilise maintenant l'égalité

$$
\int_{t_{i-1}}^{t_i}\left(W_s-W_{t_{i-1}}\right)dW_s = \frac{1}{2}\left\{\left(W_{t_i}-W_{t_{i-1}}\right)^2-\Delta^n t\right\} \ ,
$$

ce qui conduit à l'approximation

$$
\int_{t_{i-1}}^{t_i} a(X_s)dW_s \sim a(X_{t_{i-1}})\left(W_{t_i}-W_{t_{i-1}}\right)+\frac{1}{2}a'\left(X_{t_{i-1}}\right)a(X_{t_{i-1}})\left\{\left(W_{t_i}-W_{t_{i-1}}\right)^2-\Delta^n t\right\} \ .
$$

Le terme de dérive b ayant une contribution inférieure dans l'erreur d'approximation par rapport au terme de diffusion, il n'est pas nécessaire de le corriger. Pour $d=1$, on obtient donc le schéma d'approximation suivant :

$$
\begin{aligned}
\tilde{X}_0^n &= X_0 \\
\tilde{X}_{t_i}^n &= \tilde{X}_{t_{i-1}}^n+b(\tilde{X}_{t_{i-1}}^n)\Delta^n t+a(\tilde{X}_{t_{i-1}}^n)\left(W_{t_i}-W_{t_{i-1}}\right) \\
&\quad +\frac{1}{2}a'\left(\tilde{X}_{t_{i-1}}^n\right)a(X_{t_{i-1}})\left\{\left(W_{t_i}-W_{t_{i-1}}\right)^2-\Delta^n t\right\} \ , \ i=1,\ldots,n \ .
\end{aligned}
\tag{2.2.8}
$$

Exemple 2.2.13 *Dans le modèle de Black-Scholes en dimension 1, la solution de* (2.1.1) *est approchée par*

$$
\begin{aligned}
\tilde{X}_0^n &= X_0 \\
\tilde{X}_{t_i}^n &= \tilde{X}_{t_{i-1}}^n\left\{1+\left(r-\sigma^2/2\right)\Delta^n t+\sigma\left(W_{t_i}-W_{t_{i-1}}\right)+\frac{1}{2}\sigma^2\left(W_{t_i}-W_{t_{i-1}}\right)^2\right\} \ .
\end{aligned}
$$

En dimension d quelconque, le même raisonnement conduit à consider l'approximation suivante :

$$
\begin{aligned}
\tilde{X}_0^n &= X_0 \\
\tilde{X}_{t_i}^n &= \tilde{X}_{t_{i-1}}^n+b(\tilde{X}_{t_{i-1}}^n)\Delta^n t+a(\tilde{X}_{t_{i-1}}^n)\left(W_{t_i}-W_{t_{i-1}}\right) \\
&\quad +\sum_{j,l=1}^d(\nabla a^{\cdot j}a^{\cdot l})(\bar{X}_{t_{i-1}}^n)\int_{t_{i-1}}^{t_i}(W_s^j-W_{t_{i-1}}^j)dW_s^l \ , \ \ i=1,\ldots,n \ .
\end{aligned}
\tag{2.2.9}
$$

Le théorème suivant justifie l'introduction de ce schéma dans le sens où il permet d'obtenir une vitesse de convergence L^p en $1/n$ au lieu de $1/\sqrt{n}$ pour le schéma d'Euler.

Théorème 2.2.14 *Si b, $a \in C_b^2$, alors pour tout $p \geq 1$*

$$\max_{0\leq i\leq n} \mathbb{E}\left[\left\|\tilde{X}^n_{t_i} - X_{t_i}\right\|^p\right]^{1/p} \;\leq\; C/n\ .$$

La mise en oeuvre numérique du schéma (2.2.9) suppose de savoir simuler correctement l'intégrale d'Itô $\int_{t_{i-1}}^{t_i}(W_s^j - W_{t_{i-1}}^j)dW_s^l$ ce qui est très difficile en pratique pour $d \neq 1$. En général, on n'utilise ce schéma pour $d \geq 2$ que lorsque l'hyposèse de commutativité

$$\nabla a^{\cdot j} a^{\cdot l} = \nabla a^{\cdot l} a^{\cdot j} \quad \forall\, j,\ l\ \in \{1, \ldots, d\} \tag{2.2.10}$$

est vérifiée. Dans ce cas, la formule d'intégration par parties du calcul d'Itô permet de ré-écrire (2.2.9) sous la forme :

$$\begin{aligned}
\tilde{X}^n_0 &= X_0 \\
\tilde{X}^n_{t_i} &= \tilde{X}^n_{t_{i-1}} + b(\tilde{X}^n_{t_{i-1}})\Delta^n t + a(\tilde{X}^n_{t_{i-1}})\left(W_{t_i} - W_{t_{i-1}}\right) - \frac{1}{2}\sum_{j=1}^{d}(\nabla a^{\cdot j}a^{\cdot j})(\tilde{X}^n_{t_{i-1}})\Delta^n t \\
&\quad + \frac{1}{2}\sum_{j,l=1}^{d}(\nabla a^{\cdot j}a^{\cdot l})(\tilde{X}^n_{t_{i-1}})(W^j_{t_i} - W^j_{t_{i-1}})(W^l_{t_i} - W^l_{t_{i-1}})\ ,\quad i = 1, \ldots, n\ .
\end{aligned}$$

Il suffit alors de simuler les accroissements de W.

Remarque 2.2.15 On peut définir des schémas d'ordre supérieur, voir par exemple [40] et [41]. Mais en général, ils sont très difficiles à mettre en oeuvre numériquement surtout en dimension $d > 1$. On peut se référer à [18] pour une discussion sur le sujet. Voir également [23] pour un exposé complet sur le schéma d'Euler et le schéma de Milshtein.

2.3 Options Asiatiques

On s'intéresse ici aux options sur moyenne $g(A_T, X_T)$ avec $A_T := \int_0^T X_s ds$.

2.3.1 Schémas simples

Une première solution pour approcher A_T consiste à introduire le schéma d'Euler de

$$A_t \;:=\; \int_0^t X_s ds$$

dont la dynamique est donnée par

$$A_0 \;=\; 0 \quad , \quad dA_t \;=\; X_t dt\ .$$

On approche donc A_T par $\bar{A}^n_T$ défini par l'équation de récurrence stochastique

$$\begin{aligned}
\bar{A}^n_0 &= X_0 \\
\bar{A}^n_{t_i} &= \bar{A}^n_{t_{i-1}} + \bar{X}^n_{t_{i-1}}\Delta^n t\ ,\quad i \in \{1, \ldots, n\}.
\end{aligned}$$

Cela revient à utiliser l'approximation

$$\int_0^T X_s ds \quad \sim \quad \frac{1}{n}\sum_{i=0}^{n-1} \bar{X}^n_{t_i} \, .$$

On déduit du Théorème 2.2.6 la vitesse de convergence L^p de ce schéma.

Corollaire 2.3.1 *Sous* (2.0.2)*, si g est lipschitzienne, alors, pour tout $p \geq 1$,*

$$\left\| g(A_T, X_T) - g(\bar{A}^n_T, \bar{X}^n_T) \right\|_{L^p} \quad \leq \quad C/\sqrt{n} \, .$$

Dans le modèle de Black-Scholes , on peut simuler parfaitement X en un nombre fini de dates. On obtient alors un résultat plus précis.

Proposition 2.3.2 *Si $b(x) = rx$, $a(x) = \mathrm{diag}\,[x]\,\sigma$, et si g est lipschitzienne, alors, pour tout $p \geq 1$,*

$$\left\| g(A_T, X_T) - g(\bar{A}^n_T, X_T) \right\|_{L^p} \quad \leq \quad C/n$$

où

$$\bar{A}^n_T \quad = \quad \frac{1}{n}\sum_{i=0}^{n-1} X_{t_i} \, .$$

Preuve. Il suffit de démontrer le résultat pour p pair. Pour $t \in [t_i, t_{i+1}]$, on a :

$$A_t - \bar{A}^n_t \quad = \quad A_{t_i} - \bar{A}^n_{t_i} + \int_{t_i}^t (X_s - X_{t_i})\, ds$$

où $(\bar{A}^n_t)_{t \leq T}$ est le schéma d'Euler continu de A. Par ailleurs, en utilisant la version stochastique du théorème de Fubini (voir par exemple [58])

$$\begin{aligned}
\int_{t_i}^t (X_s - X_{t_i})\, ds \quad &= \quad \int_{t_i}^t \left(\int_{t_i}^s rX_u du + \int_{t_i}^s \mathrm{diag}\,[X_u]\,\sigma dW_u \right) ds \\
&= \quad \int_{t_i}^t rX_u (\int_u^t ds) du + \int_{t_i}^t (\int_u^t ds) \mathrm{diag}\,[X_u]\,\sigma dW_u \\
&= \quad \int_{t_i}^t (t-u) rX_u du + \int_{t_i}^t (t-u) \mathrm{diag}\,[X_u]\,\sigma dW_u \, .
\end{aligned}$$

On utilise maintenant le résultat suivant dont la preuve sera donnée à la fin de la section.

Lemme 2.3.3 *Soit $z \in \mathbb{R}^d$ et*

$$Z_t \quad = \quad z + \int_0^t b_s ds + \int_0^t a_s dW_s$$

où a, b sont adaptés et satisfont

$$\int_0^T \mathbb{E}\left[\|b_s\|^k\right] + \mathbb{E}\left[\|a_s\|^k\right] ds < \infty$$

pour tout $k \geq 1$. Alors, pour tout $q \geq 1$, il existe $C > 0$ tel que

$$\mathbb{E}\left[\|Z_t\|^{2q}\right] \;\; \leq \;\; \|z\|^{2q} + C \int_0^t \mathbb{E}\left[\|Z_s\|^{2q} + \|b_s\|^{2q} + \|a_s\|^{2q}\right] ds \; .$$

En appliquant ce Lemme aux égalités précédentes et en utilisant le fait que $\|X_s\|_{L^p}$ est uniformément bornée sur $[0, T]$, on obtient que pour tout $t \in [t_i, t_{i+1}]$

$$\mathbb{E}\left[\left\|A_t - \bar{A}^n_t\right\|^p\right] \;\; \leq \;\; \mathbb{E}\left[\left\|A_{t_i} - \bar{A}^n_{t_i}\right\|^p\right] + C \int_{t_i}^t \left(\mathbb{E}\left[\left\|A_s - \bar{A}^n_s\right\|^p\right] + (t-s)^p\right) ds$$

ce qui implique en utilisant le Lemme de Gronwall que

$$\mathbb{E}\left[\left\|A_t - \bar{A}^n_t\right\|^p\right] \;\; \leq \;\; \left(\mathbb{E}\left[\left\|A_{t_i} - \bar{A}^n_{t_i}\right\|^p\right](1 + \frac{C}{n}) + C\frac{1}{n^{p+1}}\right) e^{C/n} \; ,$$

d'où, pour tout $i \in \{0, \ldots, n-1\}$,

$$\mathbb{E}\left[\left\|A_{t_{i+1}} - \bar{A}^n_{t_{i+1}}\right\|^p\right] \;\; \leq \;\; \left(\mathbb{E}\left[\left\|A_{t_i} - \bar{A}^n_{t_i}\right\|^p\right](1 + \frac{C}{n}) + C\frac{1}{n^{p+1}}\right) e^{C/n} \; .$$

Comme $A_0 - \bar{A}^n_0 = 0$, cette récurrence montre le résultat. □

L'approche précédente consiste à utiliser une méthode de rectangles pour discrétiser l'intégrale. Même si d'un point de vue théorique, on obtient la même convergence L^p que pour le schéma d'Euler, il est en général nettement préférable d'utiliser une méthode de trapèzes en considérant

$$\tilde{A}^n_T \;\; = \;\; \frac{1}{n} \sum_{i=0}^{n-1} \left(\bar{X}^n_{t_i} + \bar{X}^n_{t_{i+1}}\right) / 2 \; .$$

Exemple 2.3.4 *On évalue un call asiatique dans le modèle de Black-Scholes de paramètres $r = 0.1$, $\sigma = 0.2$, $T = 1$, $X_0 = K = 100$ avec $n = 50$. Le tableau ci-dessous donne les intervalles de confiance simulés pour les deux méthodes, le prix réel étant d'environ 7.04.*

Nombre de simulations	*Euler*	*Trapèzes*
10 000	[6.93 , 7.26]	[6.96 , 7.30]
20 000	[6.88 , 7.11]	[6.97 , 7.21]
50 000	[6.85 , 6.99]	[6.97 , 7.12]
100 000	[6.86 , 6.96]	[6.98 , 7.09]

La sous-estimation du schéma d'Euler est flagrante.

Preuve du Lemme 2.3.3. Pour simplifier, on suppose que $d = 1$. Par le Lemme d'Itô, on obtient sur $[0, T]$:

$$(Z_t)^{2q} = z^{2q} + \int_0^t 2q(Z_s)^{2q-1}b_s + q(2q-1)(Z_s)^{2q-2}a_s^2 ds + \int_0^t 2q(Z_s)^{2q-1}a_s dW_s \,.$$

En utilisant les hypothèses sur a et b, on vérifie facilement que

$$\int_0^T \mathbb{E}\left[\left\|a_s(Z_s)^{2q-1}\right\|^2\right] ds < \infty \,,$$

de sorte que

$$\mathbb{E}\left[(Z_t)^{2q}\right] = z^{2q} + \int_0^t 2q\mathbb{E}\left[(Z_s)^{2q-1}b_s\right] + q(2q-1)\mathbb{E}\left[(Z_s)^{2q-2}a_s^2\right] ds \quad , \quad t \in [0, T] \,.$$

Il suffit maintenant d'utiliser les inégalités

$$\begin{aligned} \left\|(Z_s)^{2q-1}b_s\right\| &\leq \|Z_s\|^{2q}\,\mathbf{1}_{\|Z_s\| \geq \|b_s\|} + \|b_s\|^{2q}\,\mathbf{1}_{\|Z_s\| < \|b_s\|} \leq \|Z_s\|^{2q} + \|b_s\|^{2q} \\ \left\|(Z_s)^{2q-2}a_s^2\right\| &\leq \|Z_s\|^{2q} + \|a_s\|^{2q} \end{aligned}$$

pour conclure. □

2.3.2 Schéma avec développement dans le modèle de Black-Scholes

Dans le modèle de Black-Scholes, on peut améliorer la vitesse de convergence en utilisant un développement de Taylor à l'ordre 1 du type, pour $d = 1$,

$$e^{(r-\sigma^2/2)(t-t_i)+\sigma(W_t - W_{t_i})} \sim 1 + (r(t - t_i) + \sigma(W_t - W_{t_i}))$$

pour obtenir

$$\begin{aligned} A_T &= \sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} e^{(r-\sigma^2/2)(t-t_i)+\sigma(W_t - W_{t_i})} dt \\ &\sim A_T^n := \sum_{i=0}^{n-1} X_{t_i} \left(\frac{T}{n} + \frac{rT^2}{2n^2} + \sigma \int_{t_i}^{t_{i+1}} (W_t - W_{t_i}) dt\right) \,. \end{aligned}$$

On procède alors de la manière suivante : on commence par simuler $(W_{t_i})_{i=1}^n$ puis on simule A_T^n conditionnellement à $(W_{t_i})_{i=1}^n$. Pour cela, on utilise le résultat suivant qui sera démontré dans un cadre plus général par la suite (Lemme 2.4.3) :

Corollaire 2.3.5 *Soit $u < v$. Conditionnellement à $(W_u = x,\ W_v = y)$, le processus $(W_t)_{u \leq t \leq v}$ est un processus gaussien vérifiant pour $u \leq s \leq t \leq v$*

$$\mathbb{E}\left[W_t \mid W_u = x,\ W_v = y\right] = \frac{v-t}{v-u}x + \frac{t-u}{v-u}y \,, \tag{2.3.1}$$

$$\mathrm{Cov}\left(W_t,\ W_s \mid W_u = x,\ W_v = y\right) = \frac{(v-t)(s-u)}{v-u} \,. \tag{2.3.2}$$

On en déduit que, conditionnellement à $(W_{t_i}, W_{t_{i+1}}) = (x, y)$, $\int_{t_i}^{t_{i+1}} W_t dt$ suit une loi normale avec pour moments

$$\mathbb{E}\left[\int_{t_i}^{t_{i+1}} W_t dt \mid W_{t_i} = x,\ W_{t_{i+1}} = y\right] \quad = \quad \int_{t_i}^{t_{i+1}} \mathbb{E}\left[W_t \mid W_{t_i} = x,\ W_{t_{i+1}} = y\right] dt$$

et

$$\begin{aligned} &\mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} W_t dt\right)^2 \mid W_{t_i} = x,\ W_{t_{i+1}} = y\right] \\ &= 2\int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \mathbb{E}\left[W_t W_u \mid W_{t_i} = x,\ W_{t_{i+1}} = y\right] du dt\ , \end{aligned}$$

la seconde égalité étant obtenue en utilisant le Lemme d'Itô. On en déduit une formule explicite en utilisant (2.3.1) et (2.3.2).

En utilisant ce schéma, on améliore la vitesse de convergence.

Proposition 2.3.6 *Si g est lipschitzienne, alors, pour tout $p \geq 1$,*

$$\|g(A_T, X_T) - g(A_T^n, X_T)\|_{L^p} \quad \leq \quad C/n^{3/2}\ .$$

Nous renvoyons à [46] et [64] pour les preuves de ces résultats ainsi que l'étude d'autres approximations.

2.4 Options à barrière

Dans cette section, on s'intéresse à l'évaluation d'espérance de la forme :

$$\mathbb{E}\left[\mathbf{1}_{\tau > T} g(X_T)\right] \quad \text{où} \quad \tau \ := \ \inf\{t \in [0, T] \ :\ X_t \notin D\} \tag{2.4.1}$$

est le temps de sortie de X d'un borélien D de $\mathbb{R}^d$, avec pour convention $\inf \emptyset = +\infty$.

2.4.1 Approche naïve

On commence par l'approche la plus simple qui consiste à approximer (2.4.1) par son équivalent discret

$$\mathbb{E}\left[\mathbf{1}_{\bar{\tau}^n > T} g(\bar{X}_T^n)\right] \quad \text{où} \quad \bar{\tau}^n \ := \ \inf\{t_i,\ i \in \{0, \ldots, n\} \ :\ \bar{X}_{t_i}^n \notin D\} \tag{2.4.2}$$

est l'équivalent discret de τ. Cette quantité est facilement simulable et on a le résultat de convergence suivant démontré dans [31].

Théorème 2.4.1 *Si D est borné de frontière ∂D de classe[1] C^3, b et $a \in C^3$ avec a strictement uniformément elliptique[2] sur D, alors, pour toute fonction mesurable g bornée qui s'annule sur un voisinage de ∂D, on a*

$$\mathbb{E}\left[\mathbf{1}_{\bar{\tau}^n > T} g(\bar{X}_T^n)\right] - \mathbb{E}\left[\mathbf{1}_{\tau > T} g(X_T)\right] \;=\; O\left(\frac{1}{\sqrt{n}}\right) .$$

Il s'agit d'un résultat assez négatif dans la mesure où l'on perd la vitesse de convergence faible obtenue pour les options vanilla.

2.4.2 Approche par les ponts de diffusion

Dans cette partie, nous présentons une autre approche qui permet d'améliorer la vitesse de convergence faible du Théorème 2.4.1. Cette fois-ci on approche (2.4.1) par

$$\mathbb{E}\left[\mathbf{1}_{\tau^n > T} g(\bar{X}_T^n)\right] \quad \text{où} \quad \tau^n \;:=\; \inf\{t \in [0,T] \;:\; \bar{X}_t^n \notin D\} \tag{2.4.3}$$

c'est-à-dire que l'on écrit le problème sur le schéma d'Euler continu. On a alors le résultat de convergence suivant démontré dans [32].

Théorème 2.4.2 *Si D est un demi-espace, b et $a \in C^5$ avec a strictement uniformément elliptique sur D, alors, pour toute fonction mesurable g bornée qui s'annule sur un voisinage de ∂D, on a*

$$\mathbb{E}\left[\mathbf{1}_{\tau^n > T} g(\bar{X}_T^n)\right] - \mathbb{E}\left[\mathbf{1}_{\tau > T} g(X_T)\right] \;=\; \frac{C_1}{n} + o\left(\frac{1}{n}\right) .$$

On retrouve ainsi une vitesse de convergence faible en $1/n$.

2.4.2.1 Sur les ponts de diffusions

Afin d'implémenter cette approximation, on aura besoin du résultat suivant sur la loi du schéma d'Euler continu conditioné.

Lemme 2.4.3 *On suppose que $\gamma(x) = (a(x)a^*(x))^{1/2}$ est inversible pour tout $x \in \mathbb{R}^d$. Alors, conditionnellement à $(\bar{X}_{t_i}^n = x_i\ ,\ \bar{X}_{t_{i+1}}^n = x_{i+1})$, le processus $(\bar{X}_t^n)_{t_i \le t \le t_{i+1}}$ a la loi de*

$$\left(x_i + \gamma(x_i)\tilde{W}_{t-t_i}\right)_{t_i \le t \le t_{i+1}} \quad \textit{conditionnellement à } \tilde{W}_{t_{i+1}-t_i} \;=\; \gamma(x_i)^{-1}\left(x_{i+1} - x_i\right)$$

où $\tilde{W}$ est un mouvement brownien. C'est un processus gaussien d'espérance $x_i \frac{t_{i+1}-t}{t_{i+1}-t_i} + x_{i+1}\frac{t-t_i}{t_{i+1}-t_i}$ et de matrice de variance-convariance $\frac{(s-t_i)(t_{i+1}-t)}{t_{i+1}-t_i}\gamma(x_i)^2$ pour tout $t_i \le s \le t \le t_{i+1}$.

---

[1] i.e. pour tout $y \in \partial D$, il existe un voisinage $V(y)$ de y et un difféomorphisme φ de $V(y) \mapsto B \subset \mathbb{R}^d$ tels que
(i) $\varphi(V(y) \cap D) \subset \mathbb{R}_+^d := \left\{x \in \mathbb{R}^d \;:\; x^1 \ge 0\right\}$
(ii) $\varphi(V(y) \cap \partial D) \subset \partial \mathbb{R}_+^d$
(iii) $\varphi \in C^3(V(y))$ et $\varphi^{-1} \in C^3(B)$.

[2] i.e. satisfait (2.2.7) pour tout $x \in D$.

Eléments de preuve. Conditionnellement à $\bar{X}^n_{t_i} = x_i$, $(\bar{X}^n_t)_{t_i \le t \le t_{i+1}}$ est un processus de diffusion homogène qui admet pour densité de transition

$$p_h(x,z) \;=\; \frac{1}{\sqrt{(2\pi h)^d \det[\gamma^2(x_i)]}} \exp\left\{-\frac{1}{2h}\left(z-x-b(x_i)h\right)^* \left(\gamma^{-2}(x_i)\right)\left(z-x-b(x_i)h\right)\right\} .$$

Pour $t_i \le s \le t \le t_{i+1}$

$$\mathbb{P}\left[\bar{X}^n_s \in dx \;,\; \bar{X}^n_t \in dy \;,\; \bar{X}^n_{t_{i+1}} \in dx_{i+1} \mid \bar{X}^n_{t_i} = x_i\right] \;=\; p_{s-t_i}(x_i,x)p_{t-s}(x,y)p_{t_{i+1}-t}(y,x_{i+1})dxdydx_{i+1}$$

$$\mathbb{P}\left[\bar{X}^n_s \in dx,\; \bar{X}^n_{t_{i+1}} \in dx_{i+1} \mid \bar{X}^n_{t_i} = x_i\right] \;=\; p_{s-t_i}(x_i,x)p_{t_{i+1}-s}(x,x_{i+1})dxdx_{i+1} .$$

En divisant le premier terme par le second, on obtient :

$$\mathbb{P}\left[\bar{X}^n_t \in dy \mid \bar{X}^n_s = x,\; \bar{X}^n_{t_i} = x_i,\; \bar{X}^n_{t_{i+1}} = x_{i+1}\right] \;=\; \frac{p_{t-s}(x,y)p_{t_{i+1}-t}(y,x_{i+1})}{p_{t_{i+1}-s}(x,x_{i+1})}\, dy .$$

Ceci montre que

$$p^{t_{i+1},x_{i+1}}_{t_i,x_i}(s,x,t,y) \;:=\; \frac{p_{t-s}(x,y)p_{t_{i+1}-t}(y,x_{i+1})}{p_{t_{i+1}-s}(x,x_{i+1})}$$

est la densité de transition du processus $(\bar{X}^n_t)_{t_i \le t \le t_{i+1}}$ conditionnellement à $(\bar{X}^n_{t_i} = x_i, \bar{X}^n_{t_{i+1}} = x_{i+1})$. Le reste en découle par des calculs directs. □

Remarque 2.4.4 En utilisant la propriété de Markov de $\bar{X}^n$, on vérifie facilement que les processus $(\bar{X}^n_t)_{t_i \le t \le t_{i+1}}$ pour i allant de 0 à $n-1$ sont indépendants conditionnellement à $\{\bar{X}^n_{t_0}, \ldots, \bar{X}^n_{t_n}\}$.

2.4.2.2 Implémentation

On peut maintenant décrire la méthode. On commence par simuler le schéma d'Euler $(\bar{X}^n_{t_i})^n_{i=1}$ et on écrit

$$\mathbb{E}\left[\mathbf{1}_{\tau^n > T} g(\bar{X}^n_T)\right] \;=\; \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\tau^n > T} \mid (\bar{X}^n_{t_i})^n_{i=1}\right] g(\bar{X}^n_T)\right]$$

ce qui signifie qu'il faut calculer $\mathbb{E}\left[\mathbf{1}_{\bar{\tau}^n > T} \mid (\bar{X}^n_{t_i})^n_{i=1}\right]$, une fois la trajectoire discrète $(\bar{X}^n_{t_i})^n_{i=1}$ simulée. Evidemment si un des $\bar{X}^n_{t_i}$ n'est pas dans D, il n'y a rien à faire et le payoff de l'option donne 0. On ne calcule donc cette probabilité que si les $\bar{X}^n_{t_i}$ simulés sont tous dans D.

Pour cela, on va utiliser les résultats de la section précédente. Tout d'abord, la Remarque 2.4.4 implique que

$$\mathbb{E}\left[\mathbf{1}_{\tau^n > T} \mid (\bar{X}^n_{t_i})^n_{i=1}\right] \;=\; \prod_{i=0}^{n-1} \mathbb{P}\left(\forall t \in [t_i, t_{i+1}],\; \bar{X}^n_t \in D \mid \bar{X}^n_{t_i},\; \bar{X}^n_{t_{i+1}}\right) .$$

Nous allons maintenant montrer que, dans le cas où D est un demi-espace, le calcul est explicite. On écrit D sous la forme

$$D \;=\; \{y \in \mathbb{R}^d \;:\; \zeta^*(y-\kappa) > 0\} \tag{2.4.4}$$

i.e. $\partial\bar{D}$ est l'hyperplan passant par $\kappa \in \mathbb{R}^d$ othogonal à $\zeta \in \mathbb{R}^d$.

Exemple 2.4.5 *Pour une barrière haute U en dimension 1, on a $D = (-\infty, U)$, ce qui donne $\kappa = U$ et $\zeta = -1$.*

On fixe $i \in \{0, \ldots, n-1\}$. D'après le Lemme 2.4.3, on a

$$\begin{aligned}
&\mathbb{P}\left(\exists t \in [t_i, t_{i+1}],\ \bar{X}^n_t \notin D \mid \bar{X}^n_{t_i} = x_i,\ \bar{X}^n_{t_{i+1}} = x_{i+1}\right)\\
&= \mathbb{P}\left(\exists t \in [t_i, t_{i+1}],\ \zeta^*\gamma(x_i)\tilde{W}_{t-t_i} \le \zeta^*(\kappa - x_i) \mid \tilde{W}_{t_{i+1}-t_i} \;=\; \gamma(x_i)^{-1}(x_{i+1}-x_i)\right)\ .
\end{aligned}$$

On choisit maintenant une matrice P orthogonale[3] telle que $P^{1\cdot} = \frac{1}{\|\gamma(x_i)\zeta\|}\zeta^*\gamma(x_i)$. Le processus $\hat{W}_{\cdot - t_i} = (P\tilde{W}_{t-t_i})_{t\in[t_i,t_{i+1}]}$ a la même loi que $\tilde{W}$. Par ailleurs

$$\zeta^*\gamma(x_i)\tilde{W}_{t-t_i} \;=\; \zeta^*\gamma(x_i)P^*\hat{W}_{t-t_i} \;=\; \|\gamma(x_i)\zeta\|\,\hat{W}^1_{t-t_i}$$

car le vecteur ligne $\zeta^*\gamma(x_i)P^*$ a seulement sa première composante non nulle, égale à $\|\gamma(x_i)\zeta\|$. En ré-injectant ce résultat dans les égalités précédentes, on obtient

$$\begin{aligned}
&\mathbb{P}\left(\exists t \in [t_i, t_{i+1}],\ \bar{X}^n_t \notin D \mid \bar{X}^n_{t_i} = x_i,\ \bar{X}^n_{t_{i+1}} = x_{i+1}\right)\\
&= \mathbb{P}\left(\exists t \in [t_i, t_{i+1}],\ \hat{W}^1_{t-t_i} \le \frac{\zeta^*(\kappa - x_i)}{\|\gamma(x_i)\zeta\|} \mid \hat{W}^1_{t_{i+1}-t_i} = \frac{\zeta^*(x_{i+1}-x_i)}{\|\gamma(x_i)\zeta\|}\right)\\
&= \mathbb{P}\left(\min_{t\in[t_i,t_{i+1}]} \hat{W}^1_{t-t_i} \le \frac{\zeta^*(\kappa - x_i)}{\|\gamma(x_i)\zeta\|} \mid \hat{W}^1_{t_{i+1}-t_i} = \frac{\zeta^*(x_{i+1}-x_i)}{\|\gamma(x_i)\zeta\|}\right).
\end{aligned}$$

Cette probabilité se calcule facilement en utilisant le principe de réflexion du brownien (voir Lemme 2.4.8)

$$\mathbb{P}\left(\min_{t\in[t_i,t_{i+1}]} \hat{W}^1_{t-t_i} \le a \mid \hat{W}^1_{t_{i+1}-t_i} = b\right) = e^{-2\frac{n}{T}a(a-b)}\ \ \forall\, a \le 0 \text{ et } b \ge a \tag{2.4.5}$$

que l'on applique à $b = \frac{\zeta^*(x_{i+1}-x_i)}{\|\gamma(x_i)\zeta\|}$ et $a = \frac{\zeta^*(\kappa - x_i)}{\|\gamma(x_i)\zeta\|}$. Puisque l'on ne fait ce calcul que si $x_i \in D$, on vérifie bien que $a \le 0$ et $b \ge a$. Finalement, on a montré que

Proposition 2.4.6 *Si $\gamma(x)^2$ est inversible pour tout $x \in \mathbb{R}^d$ et si D est donné par* (2.4.4)*, alors pour tout $x_i,\ x_{i+1} \in D$, on a*

$$\begin{aligned}
&\mathbb{P}\left(\forall\, t \in [t_i, t_{i+1}],\ \bar{X}^n_t \in D \mid \bar{X}^n_{t_i} = x_i,\ \bar{X}^n_{t_{i+1}} = x_{i+1}\right)\\
&= 1 - \exp\left(-2\frac{n}{T}\frac{(\zeta^*(\kappa - x_i))^*\,(\zeta^*(\kappa - x_{i+1}))}{\|\gamma(x_i)\zeta\|^2}\right)\ .
\end{aligned} \tag{2.4.6}$$

[3] i.e. $PP^* = P^*P = I_d$.

Lorsque D n'est pas un demi-espace, on n'a en général plus de forme explicite pour (2.4.6). On peut toutefois essayer d'approcher D par son hyperplan tangent en $\Pi_{\partial D}(x_i)$, le projeté de x_i sur la frontière de D. Si D est de classe C^5, on retrouve la vitesse en $1/n$ du Théorème 2.4.2, voir [32].

Exemple 2.4.7 *On évalue un call up-and-out*

$$\mathbb{E}\left[\mathbf{1}_{\max_{t\in[0,T]} X_t<U}[X_T-K]^+\right]$$

dans le modèle de Black-Scholes de paramètres $r=0$, $\sigma=0.15$, $T=1$, $X_0=100$, $K=90$, $U=130$. *Le tableau ci-dessous donne les intervalles de confiance simulés pour les deux méthodes (naïve et par pont), le prix réel étant d'environ* 9.21. *On effectue à chaque fois* 30.000 *simulations.*

Nombre de pas de temps	*Méthode Naïve*	*Méthode par pont*
10	[9.84 , 9.97]	[9.18 , 9.31]
50	[9.46 , 9.60]	[9.14 , 9.27]
100	[9.40 , 9.54]	[9.16 , 9.30]

La sur-estimation de la méthode naïve est flagrante : il est absolument nécessaire d'utiliser la méthode tenant compte de la probabilité de sortie de D entre deux dates de discrétisation.

On conclut cette section par le Lemme que l'on a utilisé pour obtenir (2.4.5).

Lemme 2.4.8 *Soit W un mouvement brownien unidimensionelle et $\{\mathcal{F}_t^W,\ t\geq 0\}$ sa filtration naturelle. Alors pour tout $a\leq 0$, $b\geq a$ et $h>0$*

$$\mathbb{P}\left(\min_{t\in[0,h]} W_t\leq a \mid W_h=b\right) \;=\; e^{-\frac{2}{h}a(a-b)}\ .$$

Preuve. Soit $\tau_a := \inf\{t\geq 0\ :\ W_t=a\}$, le temps d'atteinte de a par W. On a alors,

$$\mathbb{P}\left(\min_{t\in[0,h]} W_t\leq a\ ,\ W_h\geq b\right) \;=\; \mathbb{P}\left(\tau_a\leq h\ ,\ W_h\geq b\right)=\mathbb{P}\left(\tau_a\leq h\ ,\ W_h-W_{\tau_a}\geq b-a\right)\ .$$

Comme τ_a est $\mathcal{F}_{\tau_a}^W$−mesurable et que $W_h-W_{\tau_a}$ est indépendant de $\mathcal{F}_{\tau_a}^W$ par la propriété de Markov forte du mouvement brownien, on obtient

$$\begin{aligned}\mathbb{P}\left(\min_{t\in[0,h]} W_t\leq a\ ,\ W_h\geq b\right) &= \mathbb{P}\left(\tau_a\leq h\ ,\ W_h-W_{\tau_a}\leq a-b\right)\\ &= \mathbb{P}\left(\tau_a\leq h\ ,\ W_h\leq 2a-b\right)\end{aligned}$$

car $W_h-W_{\tau_a}$ et $W_{\tau_a}-W_h$ ont la même loi (propriété de symétrie). Comme $2a-b\leq a$, on a donc

$$\mathbb{P}\left(\min_{t\in[0,h]} W_t\leq a\ ,\ W_h\geq b\right) \;=\; \mathbb{P}\left(W_h\leq 2a-b\right)\ .$$

Finalement, comme

$$\mathbb{P}\left(\min_{t\in[0,h]} W_t \leq a \mid W_h = b\right) \quad = \quad \frac{\frac{\partial}{\partial b}\mathbb{P}\left(\min_{t\in[0,h]} W_t \leq a\ ,\ W_h \leq b\right)}{\frac{\partial}{\partial b}\mathbb{P}\left(W_h \leq b\right)}\ ,$$

un calcul direct donne le résultat. □

2.5 Options Lookback

On s'intéresse maintenant aux payoffs de la forme $g\left(X_T, \max_{t\in[0,T]} X_t\right)$ dans un modèle en dimension 1. Pour de telles options, on considère l'approximation

$$\mathbb{E}\left[g\left(X_T, \max_{t\in[0,T]} X_t\right)\right] \ \approx\ \mathbb{E}\left[g\left(\bar{X}^n_T, \max_{t\in[0,T]} \bar{X}^n_t\right)\right]\ .$$

Pour cela, il faut pouvoir simuler $\max_{t\in[0,T]} \bar{X}^n_t$, le maximum du schéma d'Euler continu. On commence par simuler les $(\bar{X}^n_{t_i})_{i=1}^n$ puis on simule

$$\max_{t\in[0,T]} \bar{X}^n_t \ =\ \max_{0\leq i\leq n-1} \max_{t\in[t_i,t_{i+1}]} \bar{X}^n_t \tag{2.5.1}$$

sachant $(\bar{X}^n_{t_i})_{i=1}^n$. D'après la Remarque 2.4.4,

$$\mathrm{Loi}\left(\max_{t\in[t_i,t_{i+1}]} \bar{X}^n_t \mid (\bar{X}^n_{t_i})_{i=1}^n\right) \ =\ \mathrm{Loi}\left(\max_{t\in[t_i,t_{i+1}]} \bar{X}^n_t \mid \bar{X}^n_{t_i},\ \bar{X}^n_{t_{i+1}}\right)\ ,$$

dont la fonction de répartition est donnée par la Proposition 2.4.6 avec $D = (-\infty, M)$ (i.e $\zeta = -1$ et $\kappa = M$)

$$\begin{aligned}&\mathbb{P}\left(\max_{t\in[t_i,t_{i+1}]} \bar{X}^n_t \leq M \mid \bar{X}^n_{t_i} = x_i,\ \bar{X}^n_{t_{i+1}} = x_{i+1}\right)\\ &= 1 - \exp\left(-2\frac{n}{T}\frac{(M-x_i)(M-x_{i+1})}{a^2(x_i)}\right) \ =:\ F_n(M; x_i, x_{i+1}) \quad \text{pour } M \geq \max\{x_i, x_{i+1}\}\ .\end{aligned}$$

Son inverse est

$$F_n^{-1}(U; x_i, x_{i+1}) \ =\ \frac{1}{2}\left(x_i + x_{i+1} + \sqrt{(x_{i+1}-x_i)^2 - 2\frac{a^2(x_i)T}{n}\ln(1-U)}\right)\ .$$

On en déduit que

$$\left(\max_{t\in[t_i,t_{i+1}]} \bar{X}^n_t\right)_{i=0}^{n-1} \ \text{sachant}\ (\bar{X}^n_{t_i} = x_i,\ 0 \leq i \leq n) \ \stackrel{\text{loi}}{=}\ \left(F_n^{-1}(U_i; x_i, x_{i+1})\right)_{i=0}^{n-1}$$

où $(U_i)_{i=0}^{n-1}$ est une suite de variables aléatoires indépendantes uniformes sur $[0,1]$. D'après (2.5.1) cela permet de simuler parfaitement $\max_{t\in[0,T]} \bar{X}^n_t$ conditionnellement à $(\bar{X}^n_{t_i},\ 0 \leq i \leq n)$. On peut procéder de la même manière pour $d \geq 1$, en considérant les maxima des différentes composantes de X.

2.6 Options Américaines

Dans un marché complet sans friction, le prix d'une option américaine de payoff g admet la représentation

$$p(0,x) \;=\; \sup_{\tau\in\mathcal{T}_{[0,T]}} \mathbb{E}\left[e^{-r\tau}g(X_\tau)\mid X_0=x\right]$$

où X est la solution de (2.0.1) avec $b=rX$, $r>0$ est le taux sans risque et $\mathcal{T}_{[0,T]}$ est l'ensemble des temps d'arrêts à valeurs dans $[0,T]$ (voir par exemple [44]). On discrétise cette équation naturellement en considérant l'ensemble des temps d'arrêts $\bar{\mathcal{T}}^n_{[0,T]}$ à valeurs dans $\{0=t_0,t_1,\ldots,t_n=T\}$ et en remplaçant X par son schéma d'Euler $\bar{X}^n$. On obtient alors l'approximation

$$\bar{p}^n(0,x) \;=\; \sup_{\tau\in\bar{\mathcal{T}}^n_{[0,T]}} \mathbb{E}\left[e^{-r\tau}g(\bar{X}^n_\tau)\mid \bar{X}^n_0=x\right] .$$

On peut alors vérifier que $\bar{p}^n$ satisfait l'équation de la programmation dynamique

$$\begin{aligned}\bar{p}^n(t_n,\bar{X}^n_{t_n}) &= g(\bar{X}^n_{t_n})\\ \bar{p}^n(t_i,\bar{X}^n_{t_i}) &= \max\left\{g(\bar{X}^n_{t_i})\ ,\ e^{-\frac{rT}{n}}\mathbb{E}\left[\bar{p}^n(t_{i+1},\bar{X}^n_{t_{i+1}})\mid \bar{X}^n_{t_i}\right]\right\}\ ,\ i\in\{0,\ldots,n-1\} .\end{aligned}$$

La preuve du résultat suivant peut être trouvée dans [5].

Théorème 2.6.1 *Si g et a sont lipschitziennes, pour tout $p\geq 1$,*

$$\sup_{i\in\{0,\ldots,n\}} \left\|\bar{p}^n(t_i,\bar{X}^n_{t_i})-p(t_i,X_{t_i})\right\|_{L^p} \;\leq\; C/\sqrt{n} .$$

On pourra se référer à [4], [5], [9], [11], [12] et [66] pour une approche similaire appliquée à la discrétisation d'équations forward-backward plus générales.

Chapitre 3

Réduction de la variance

On s'intéresse ici à l'évaluation d'option dont le payoff est de la forme $F = g(\bar{X}^n)$ où $\bar{X}^n$ est le schéma d'Euler de la solution X de (2.0.1). L'impact de la discrétisation est donné par les théorèmes du Chapitre 2. On rappelle que dans certains modèles comme celui de Black-Scholes, (2.1.1), on peut simuler de manière exacte X, il n'y a donc pas d'erreur de discrétisation. Dans ce cas on écrira indifféremment $\bar{X}^n$ ou X pour désigner le processus X.

Une autre source d'erreur est l'erreur de Monte-Carlo : même s'il est convergent, un estimateur reste une variable aléatoire avec sa propre variance. On a déjà vu dans le Chapitre 1 comment la mesurer. En pratique, celle-ci peut être très grande, rendant le résultat obtenu peu précis. Dans les sections suivantes, nous allons étudier différentes méthodes permettant de la réduire.

3.1 Contrôle antithétique

L'idée du contrôle antithétique est très simple. Nous la présentons dans le cadre du modèle de Black-Scholes en dimension 1, i.e. X est solution de (2.1.1). Elle est basée sur la propriété de symétrie du mouvement brownien : $W \overset{\text{loi}}{=} -W$. On en déduit que $X_T \overset{\text{loi}}{=} X_T^-$ où

$$X_T^- \;=\; X_0 \exp\big((r - \sigma^2/2)T - W_T\big) \;.$$

Ceci implique que

$$\mathbb{E}\left[\frac{g(X_T) + g(X_T^-)}{2}\right] \;=\; \mathbb{E}\left[g(X_T)\right] \;.$$

Si

$$\mathrm{Var}\left(\frac{g(X_T) + g(X_T^-)}{2}\right) \;<\; \frac{\mathrm{Var}\left(g(X_T)\right)}{2}$$

ce qui revient à dire que

$$\mathrm{Cov}\left(g(X_T)\;,\; g(X_T^-)\right) \;<\; 0\;,$$

on peut obtenir une meilleure précision en simulant deux fois moins d'accroissement du brownien. En effet, si N est le nombre de simulations par la méthode sans contrôle antithétique, alors sous la condition précédente

$$\frac{\mathrm{Var}\left(\frac{g(X_T)+g(X_T^-)}{2}\right)}{N/2} \ < \ \frac{\mathrm{Var}\left(g(X_T)\right)}{N} \ ,$$

ce qui signifie que l'estimateur obtenu en simulant $\frac{g(X_T)+g(X_T^-)}{2}$ avec $N/2$ trajectoires est plus précis que celui qui consiste à simuler $g(X_T)$ avec N trajectoires. On peut montrer qu'il y a un gain de variance dès que $W_T \mapsto g(X_T)$ est monotone.

Lemme 3.1.1 *Si g est monotone et $\sigma > 0$, alors*

$$\mathrm{Cov}\left(g(X_T)\ ,\ g(X_T^-)\right) \ \leq \ 0$$

avec inégalité stricte si g est strictement monotone sur un domaine de mesure non nulle.

Preuve. Soit f la densité de la loi normale centrée de variance T. On note $m = \mathbb{E}\left[g(X_T)\right]$ et $\tilde{g}(W_T) = g(X_T)$. Sans perte de généralité, on peut supposer que $\tilde{g}$ est croissante. Soit $c := \inf\{y \in \mathbb{R} \ : \ \tilde{g}(y) \geq m\}$. On a alors

$$\begin{aligned}\int \left(\tilde{g}(w) - m\right)\left(\tilde{g}(-w) - m\right) f(w)dw \ &= \ \int \left(\tilde{g}(w) - m\right)\left(\tilde{g}(-w) - \tilde{g}(-c)\right) f(w)dw \\ &\quad + \left(\tilde{g}(-c) - m\right) \int \left(\tilde{g}(w) - m\right) f(w)dw \ .\end{aligned}$$

Par définition de m, on a $\int \left(\tilde{g}(w) - m\right) f(w)dw = 0$. On utilise ensuite la monotonie de $\tilde{g}$ pour montrer que pour tout w

$$\left(\tilde{g}(w) - m\right)\left(\tilde{g}(-w) - \tilde{g}(-c)\right) \ \leq \ 0 \ ,$$

avec inégalité stricte sur un domaine de mesure non nulle g est strictement monotone sur un domaine de mesure non nulle. □

Exemple 3.1.2 *On applique cette méthode à l'évaluation d'un put européen dans le modèle de Black-Scholes. On prend $r = 0$, $\sigma = 0.3$, $T = 1$, $X_0 = 100$, et un strike $K = 105$. Le tableau suivant résume les écart-types et les intervalles de confiance estimés avec et sans contrôle antithétique. Le nombre de trajectoires correspond au N ci-dessus. Le prix exact est de 14.88.*

Nombre de trajectoires	*Avec*	*Sans*
20 000	0.0256 , [14.83 , 14.93]	0.1163 , [14.62 , 15.08]
200 000	0.0181 , [14.85 , 14.92]	0.03684 , [14.77 , 14.91]
1 000 000	0.0080 , [14.87 , 14.90]	0.0165 , [14.85 , 14.91]

Le gain est évident, surtout pour un petit nombre de trajectoires.

On refait maintenant la même étude pour le payoff (non monotone) $[K-X_T]^+ + [X_T - K]^+$ *:*

Nombre de trajectoires	*Avec*	*Sans*
20 000	0.1723 , [24.40 , 25.08]	0.1349 , [24.57 , 25.10]
200 000	0.0554 , [24.72 , 24.93]	0.0421 , [24.66 , 24.83]
1 000 000	0.0247 , [24.71 , 24.80]	0.0188 , [24.72 , 24.79]

3.2 Régularisation du payoff

On se place en dimension $d = 1$ et on considère $F = g(\bar{X}^n_{t_1}, \ldots, \bar{X}^n_{t_n})$. Si la fonction g est trop irrégulière, on risque d'avoir une mauvaise approximation de l'espérance. La proposition suivante montre comment on peut la régulariser.

Proposition 3.2.1 *On suppose que* $a \geq \varepsilon$ *sur* $\mathbb{R}$ *pour un* $\varepsilon > 0$. *Soient* g *et* G *deux fonctions à croissance polynomiale telles que* $g = \partial^n G/\partial x^1 \cdots \partial x^n$ *au sens des distributions, alors :*

$$\mathbb{E}[F] \;=\; \mathbb{E}\left[G\left(\bar{X}^n_{t_1}, \ldots, \bar{X}^n_{t_n}\right) \prod_{i=1}^{n} \frac{n\left(W_{t_i} - W_{t_{i-1}}\right)}{Ta(\bar{X}^n_{t_{i-1}})}\right] .$$

Preuve. On se restreint au cas où $F = g(\bar{X}^n_T)$ pour simplifier les arguments. Le cas général est obtenu de la même manière. Par construction du schéma d'Euler, on a

$$\mathbb{E}\left[g(\bar{X}^n_T)\right] \;=\; \mathbb{E}\left[\int_{\mathbb{R}} g\left(\bar{X}^n_{t_{n-1}} + b(\bar{X}^n_{t_{n-1}})\Delta^n t + a(\bar{X}^n_{t_{n-1}})\frac{\omega T^{1/2}}{n^{1/2}}\right) f(\omega)d\omega\right]$$

où f est la densité de la loi $\mathcal{N}(0,1)$. En intégrant par parties, on obtient alors

$$\mathbb{E}[F] \;=\; \mathbb{E}\left[\int_{\mathbb{R}} G\left(\bar{X}^n_{t_{n-1}} + b(\bar{X}^n_{t_{n-1}})\Delta^n t + a(\bar{X}^n_{t_{n-1}})\frac{\omega T^{1/2}}{n^{1/2}}\right) \frac{n^{1/2}\omega f(\omega)d\omega}{T^{1/2}a(\bar{X}^n_{t_{n-1}})}\right]$$

ce qui donne le résultat. $\square$

Ceci permet de régulariser le payoff g en passant à une primitive. Le prix à payer est l'apparition du poids aléatoire $\prod_{i=1}^n \frac{n\left(W_{t_i} - W_{t_{i-1}}\right)}{Ta(\bar{X}^n_{t_{i-1}})}$, dont la variance est très élevée si les t_i sont proches.

Exemple 3.2.2 *Soit* $g(x) = \mathbf{1}_{[a,b]}(x)$ *avec* $d = 1$. *On introduit la fonction* $G(x) = (x - a)\mathbf{1}_{[a,b]}(x) + (b-a)\mathbf{1}_{x>b}$. *Alors,* $G' = g$ *et on obtient :*

$$\mathbb{E}\left[g(\bar{X}^n_T)\right] \;=\; \mathbb{E}\left[G(\bar{X}^n_T)\frac{n\left(W_T - W_{t_{n-1}}\right)}{Ta(\bar{X}^n_{t_{n-1}})}\right] .$$

Dans le modèle de Black-Scholes , si $\tilde{X} = \ln(X)$ est le log du prix, on obtient

$$\mathbb{E}\left[g(\tilde{X}_T)\right] \quad = \quad \mathbb{E}\left[G(\tilde{X}_T)\frac{W_T}{\sigma T}\right] \; . \tag{3.2.1}$$

On teste cette méthode pour $\mathbb{E}\left[\mathbf{1}_{\{X_T \in [a,b]\}}\right]$ dans le modèle de Black-Scholes de paramètres $r = 0$, $\sigma = 0.25$, $T = 1$, $X_0 = 100$. On prend $a = 95$ et $b = 105$. On passe au ln pour utiliser la formule (3.2.1). Le tableau ci-dessous donne les écart-types et les intervalles de confiance simulés avec et sans régularisation, le prix réel étant d'environ 15.7.

Nombre de simulations	*Sans*	*Avec*
10 000	0.43 , [15.6 , 17.3]	0.18 , [15.3 , 16.0]
20 000	0.30 , [15.5 , 16.7]	0.12 , [15.4 , 15.9]
50 000	0.24 , [15.6 , 16.5]	0.10 , [15.5 , 15.9]

Exercice 3.2.3 *Soit φ telle que $\varphi = 1$ sur $[a,b]$. Montrer que si φ et sa dérivée (au sens des distributions) sont à croissance polynomiale, alors :*

$$\mathbb{E}\left[\mathbf{1}_{[a,b]}(\bar{X}^n_T)\right] \quad = \quad \mathbb{E}\left[G(\bar{X}^n_T)\left(\varphi(\bar{X}^n_T)\frac{n\left(W_T - W_{t_{n-1}}\right)}{Ta(\bar{X}^n_{t_{n-1}})} - \varphi'(\bar{X}^n_T)\frac{n^{1/2}}{T^{1/2}a(\bar{X}^n_{t_{n-1}})}\right)\right]$$

où $G(x) = (x-a)\mathbf{1}_{[a,b]}(x) + (b-a)\mathbf{1}_{x>b}$.

3.3 Variable de contrôle

3.3.1 Motivation et exemples

L'idée du *Control Variate* est d'introduire une variable Y d'espérance nulle telle que $\mathrm{Var}\,(F+Y) << \mathrm{Var}\,(F)$. Puisque Y est d'espérance nulle, on peut estimer $\mathbb{E}\,[F] = \mathbb{E}\,[F+Y]$ par

$$\hat{p}^N_Y \quad := \quad \frac{1}{N}\sum_{j=1}^{N}\left(F^{(j)} + Y^{(j)}\right) \; .$$

Evidemment cet estimateur a les mêmes propriétés de convergence que $\hat{p}^N$, mais est plus précis à N fixé.

Exemple 3.3.1 (Option Vanilla) *On considère le payoff $F = g(\bar{X}^n_T)$ où g est à valeurs dans $\mathbb{R}$ et $\bar{X}$ est solution de (2.2.1) avec $b = 0$, i.e. on considère par exemple les prix actualisés des actifs sous-jacents au taux sans risque. Dans ce cas, $\mathbb{E}[\bar{X}^n_T] = X_0$. La variable (vectorielle) $\bar{X}^n_T$ est donc un candidat naturel pour servir de variable de contrôle. On peut alors chercher à minimiser sur $\rho \in \mathbb{R}^d$*

$$\mathrm{Var}\left(g(\bar{X}^n_T) + \rho^*\bar{X}^n_T\right) \quad = \quad \mathrm{Var}\left(g(\bar{X}^n_T)\right) + 2\rho^*\mathrm{Cov}\left(g(\bar{X}^n_T), \bar{X}^n_T\right) + \rho^*\mathrm{Var}\left(\bar{X}^n_T\right)\rho \; ,$$

où $\mathrm{Cov}\left(g(\bar{X}_T^n), \bar{X}_T^n\right) = \left(\mathrm{Cov}\left(g(\bar{X}_T^n), \bar{X}_T^{n^j}\right)\right)_j$ *et* $\mathrm{Var}\left(\bar{X}_T^n\right) = \left(\mathrm{Cov}\left(\bar{X}_T^{n^i}, \bar{X}_T^{n^j}\right)\right)_{i,j}$. *L'optimum est atteint pour*

$$\hat{\rho} \quad := \quad -\mathrm{Var}\left(\bar{X}_T^n\right)^{-1} \mathrm{Cov}\left(g(\bar{X}_T^n), \bar{X}_T^n\right) \ .$$

En général, on ne connaît pas explicitement $\hat{\rho}$ mais on peut essayer de l'approcher numériquement par Monte-Carlo.

Dans l'exemple suivant on se place dans le modèle de Black-Scholes en dimension 1. On prend $r = 0$, $\sigma = 0.3$, $X_0 = 100$, $T = 1$, et on veut évaluer un call européen de strike $K = 80$. Comme le strike est faible par rapport au prix du sous-jacent, on peut s'attendre à un forte corrélation entre le payoff de l'option et X_T. On estime le $\hat{\rho} = -0.825$. Le tableau ci-dessous résume les écart-types et les intervalles de confiance estimés avec et sans variable de contrôle, le prix exact étant de 23.53*.*

Nombre de simulations	*Avec*	*Sans*
5 000	0.0938 , [23.29 , 23.66]	0.3887 , [23.42 , 24.95]
10 000	0.0664 , [23.39 , 23.65]	0.2739 , [23.50 , 24.57]
100 000	0.0208 , [23.51 , 23.60]	0.0851 , [23.47 , 23.80]
500 000	0.0093 , [23.52 , 23.55]	0.0379 , [23.48 , 23.63]

On observe que l'utilisation de la variable de contrôle améliore largement la précision de l'estimateur. Bien entendu, il s'agit d'un cas favorable puisque la corrélation est très forte.

On refait maintenant la même étude avec un strike $K = 150$. Dans ce cas, la corrélation n'est plus que de 0.148*, le gain de variance est beaucoup plus faible. On obtient les résultats suivants, le prix exact étant de* 1.49*,*

Nombre de simulations	*Avec*	*Sans*
5 000	0.06842 , [1.496 , 1.764]	0.0828 , [1.445 , 1.770]]
10 000	0.0296 , [1.466 , 1.582]	0.0360 , [1.4592 , 1.600]
100 000	0.0210 , [1.470 , 1.552]	0.02567 , [1.468 , 1.568]
500 000	0.0092 , [1.477 , 1.513]	0.0113 , [1.475 , 1.519]

Exemple 3.3.2 (Option asiatique) *On suppose que X est solution de l'équation de Black-Scholes* (2.1.1) *en dimension 1. Si σ est faible, on peut suivre l'approche de [39] qui consiste à utiliser l'approximation*

$$\frac{1}{T}\int_0^T X_t dt \quad \sim \quad \exp\left(\frac{1}{T}\int_0^T \ln(X_t)dt\right) \quad =: \ Z_T \ ,$$

où

$$\frac{1}{T}\int_0^T \ln(X_t)dt \quad = \quad (r - \sigma^2/2)\frac{T}{2} + \frac{\sigma}{T}\int_0^T W_t dt \ . \tag{3.3.1}$$

Comme

$$Loi\left(\frac{1}{T}\int_0^T \ln(X_t)dt\right) = \mathcal{N}\left(T(r/2-\sigma^2/4)\ ,\ \sigma^2 T/3\right)\ ,$$

il est souvent possible de calculer facilement $\mathbb{E}\left[g(Z_T)\right]$*. Par exemple, si* $g(x)=[x-K]^+$*, il suffit d'utiliser la formule de Black-Scholes avec les paramètres* $\tilde r = r/2-\sigma^2/12$ *et* $\tilde\sigma = \sigma/\sqrt{3}$*, en faisant attention à corriger le résultat du facteur d'actualisation. On peut alors se servir de* $g(Z_T)$ *comme variable de contrôle.*

On teste cette méthode pour évaluer un call asiatique de paramètres $r=0.1$*,* $\sigma=0.2$*,* $T=1$*,* $X_0=K=100$*. On utilise la discrétisation par trapèzes pour approcher* $\int_0^T X_t dt$ *avec* $n=50$*. La formule de Black-Scholes permet de calculer* $e^{-rT}\mathbb{E}\left[g(Z_T)\right]=6.77$ *et on prend pour variable de contrôle*

$$Y = e^{-rT}\left(\mathbb{E}\left[g(Z_T)\right]-\left[e^{(r-\sigma^2/2)\frac{T}{2}+\frac{\sigma}{2n}\sum_{i=0}^{n-1}\left(W_{t_i}+W_{t_{i+1}}\right)}-K\right]^+\right)\ .$$

Le tableau ci-dessous donne les intervalles de confiance simulés avec et sans cette variable de contrôle, le prix réel étant d'environ 7.04.

Nombre de simulations	*Avec*	*Sans*
10 000	[7.034 , 7.049]	[6.957 , 7.296]
20 000	[7.034 , 7.046]	[6.970 , 7.207]
50 000	[7.039 , 7.045]	[6.974 , 7.124]
100 000	[7.037 , 7.042]	[6.980 , 7.086]

Exemple 3.3.3 (Option américaine) *Dans la Section 2.6, on a proposé un schéma de discrétisation de l'équation backward réfléchie associée au problème d'évaluation de l'option américaine. On donnera dans la Section 5.2 une méthode pour estimer par Monte-Carlo les espérances conditionnelles qui apparaissent dans ce schéma. Comme le prix de l'option américaine est fortement corrélé à celui de l'option européenne, on peut l'utiliser comme variable de contrôle dans les modèles de type Black-Scholes pour lesquelles on connaît le prix exacte de l'européenne. Pour un put américain de strike* K*, le schéma devient*

$$\begin{aligned}
\bar p^n(t_n,\bar X^n_{t_n}) &= \left[K-\bar X^n_{t_n}\right]^+\\
\bar p^n(t_i,\bar X^n_{t_i}) &= \max\Big\{g(\bar X^n_{t_i})\ ,\ e^{-\frac{rT}{n}}\mathbb{E}\left[\bar p^n(t_{i+1},\bar X^n_{t_{i+1}})-\left[K-\bar X^n_{t_{i+1}}\right]^+\ \mid \bar X^n_{t_i}\right]\\
&\qquad +BS(\bar X^n_{t_i},t_{i+1})\Big\}\quad ,\quad i\in\{0,\ldots,n-1\}\ .
\end{aligned}$$

où $BS(\bar X^n_{t_i},t_{i+1})$ *est le prix de Black-Scholes du put de maturité* t_{i+1} *si le sous-jacent vaut* $\bar X^n_{t_i}$ *à la date* t_i*. On peut également utiliser le prix de l'option européenne comme variable de contrôle à la place du payoff. Ceci revient à remplacer* $[K-\bar X^n_{t_{i+1}}]^+$ *par* $BS(\bar X^n_{t_{i+1}},T)$ *et* $BS(\bar X^n_{t_i},t_{i+1})$ *par* $BS(\bar X^n_{t_i},T)$*.*

3.3.2 Approche systématique

Dans les exemples précédents, on choisit une classe de variables de contrôle que l'on suppose être "bonne" *a priori*. Le théorème de représentation des martingales (voir par exemple [37]) permet d'obtenir une variable de contrôle permettant d'éliminer en théorie (de réduire en pratique) la variance de l'estimateur.

Théorème 3.3.4 (Représentation des martingales) *Soit $F \in L^2$ une variable aléatoire $\mathcal{F}_T$ mesurable, alors il existe $\phi^F \in L^2([0,T] \times \Omega, dt \otimes dP)$ adapté tel que*

$$F \;=\; \mathbb{E}[F] + \int_0^T \phi_s^F dW_s \;.$$

Pour $F = g(\bar{X}_T^n)$, il est donc optimal d'utiliser $Y := -\int_0^T \phi_s^F dW_s$ comme variable de contrôle. Evidemment, ceci est complètement théorique puisqu'en général on ne connaît pas ϕ^F.

Si $F = g(X_T)$, on peut estimer ϕ^F en utilisant le Théorème 2.2.10 :
1/ On calcule numériquement une approximation $\tilde{u}$ de la solution de (2.2.5) par une méthode de différences finies. On obtient ainsi une approximation $\nabla_x \tilde{u}$ du gradient de u par rapport à x.
2/ On utilise l'approximation

$$\int_0^T \nabla_x u(s, X_s) a(X_s) dW_s \;\sim\; \sum_{i=1}^n \nabla_x \tilde{u}(t_{i-1}, \bar{X}_{t_{i-1}}^n) a(\bar{X}_{t_{i-1}}^n)(W_{t_i} - W_{t_{i-1}}) \;.$$

3/ On pose

$$Y = \sum_{i=1}^n \nabla_x \tilde{u}(t_{i-1}, \bar{X}_{t_{i-1}}^n) a(\bar{X}_{t_{i-1}}^n)(W_{t_i} - W_{t_{i-1}}) \;.$$

Comme $\mathbb{E}[Y] = 0$, on peut s'en servir comme variable de contrôle.

Remarque 3.3.5 Evidemment, si l'on pouvait calculer une approximation suffisamment fine de la solution de (2.2.5), on n'aurait pas besoin d'utiliser de méthode de Monte-Carlo, cf (2.2.6). Toutefois, on peut se restreindre à une résolution grossière de (2.2.5) qui soit suffisante pour être utilisée dans la méthode de réduction de variance. On peut également approcher le gradient par celui correspondant à un payoff/modèle proche pour lequel on a une formule explicite. Par exemple, dans un modèle à volatilité stochastique, on peut approcher le delta en t_i par le delta de Black-Scholes correspondant à la valeur de la volatilité en t_i.

Remarque 3.3.6 Cette approche par EDP s'étend aux options asiatiques, lookback ou à barrière. Le cas des options asiatiques est traité simplement en augmentant la taille du processus X, le résultat est similaire à celui du Théorème 2.2.10 pour le processus augmenté, voir également [45] Proposition 5.2.11. On pourra consulter [61] pour les options lookback et [32] pour les options à barrière.

3.3.3 Méthode adaptative

On se place dans le cadre du modèle de Balck-Scholes, i.e. $b(x) = rx$ et $a(x) = \mathrm{diag}\,[x]\,\sigma$, avec $d = 1$ et un payoff de la forme

$$G \;=\; \psi\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_\kappa} - W_{t_{\kappa-1}}\right)\;.$$

On cherche un processus h déterministe et constant, égal à h^i, sur les intervalles $[t_{i-1}, t_i]$. On note $Z = (W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_\kappa} - W_{t_{\kappa-1}})$ et, par abus de notation, $h = (h^1, \ldots, h^\kappa)$. Le problème est donc de minimiser sur $\mathbb{R}^\kappa$ la fonctionnelle

$$H(h) \;:=\; \mathbb{E}\left[(\psi(Z) - h^* Z)^2\right]\;.$$

En écrivant les conditions du premier ordre sur ce problème de minimisation strictement convexe et coercif, on déduit que la solution est donnée par $\hat{h}$ défini par

$$\hat{h}_i \;=\; \mathbb{E}[Z^i \psi(Z)]/(t_i - t_{i-1})\;.$$

On considère alors la suite $(\zeta_n)_n$ définie par

$$\zeta_n^i \;=\; \frac{1}{n}\sum_{k=1}^{n}(Z_n^i \psi(Z_n))/(t_i - t_{i-1})$$

où $(Z_n)_n$ est une suite i.i.d. de même loi que Z. Il est clair que $\zeta_n \to \hat{h}$ p.s. Si ψ est au plus à croissance exponentielle, on déduit alors des résultats généraux obtenus par [2] que

$$\begin{aligned}
&\hat{m}_N^c := \frac{1}{N}\sum_{n=1}^{N}\left(\psi(Z_{n+1}) - h_n^* Z_{n+1}\right) \longrightarrow \mathbb{E}[Z] \;\; p.s. \\
&\sqrt{N}\left(\hat{m}_N^c - \mathbb{E}[Z]\right) \longrightarrow N(0, v^2) \\
&\hat{\sigma}_N^2 := \frac{1}{N}\sum_{n=1}^{N}\left(\psi(Z_{n+1}) - h_n^* Z_{n+1}\right)^2 - (\hat{m}_N^c)^2 \longrightarrow v^2 \;\; p.s.
\end{aligned}$$

où $v^2 := \mathrm{Var}(\psi(Z) - \hat{h}^* Z)$.

3.4 Fonction d'importance

3.4.1 Un exemple

Supposons que l'on veuille évaluer $\mathbb{E}[X_T - K]^+$. Si K est proche de, ou inférieur à, X_0, on aura beaucoup de simulations $X_T^{(j)}$ supérieures à K. Mais si K est beaucoup plus grand que X_0, il y en aura peu et l'estimateur de Monte-Carlo risque d'avoir une variance très forte.

Une façon de remédier à ce problème est de forcer X_T "à aller au dessus de K" en lui ajoutant un drift positif. Ceci peut être fait sans changer la valeur de l'espérance grâce au théorème de Girsanov, voir par exemple Théorème 5.1 dans [37].

Théorème 3.4.1 (Girsanov) *Soit h un processus adpaté à valeurs dans $\mathbb{R}^d$ tel que*

$$\int_0^T \|h_s\|^2 ds \quad < \quad \infty \;\; \mathbb{P}-p.s.$$

Soit

$$H_T^h \quad := \quad \exp\left\{-\frac{1}{2}\int_0^T \|h_s\|^2 ds + \int_0^T h_s^* dW_s\right\} \; .$$

Si $\mathbb{E}[H_T^h] = 1$ alors $\mathbb{P}^h = H_T^h \cdot \mathbb{P}$ est équivalente à $\mathbb{P}$ et

$$W_.^h \quad := \quad W_. - \int_0^. h_s ds$$

est un $\mathbb{P}^h$-mouvement brownien.

Si h vérifie les hypothèses du théorème précédent, et si on note $\mathbb{E}^h$ l'espérance sous $\mathbb{P}^h$, on a :

$$\mathbb{E}\left[g(X_T)\right] \quad = \quad \mathbb{E}^h\left[(H_T^h)^{-1} g(X_T)\right] \; .$$

Dans le cas présenté en introduction, on peut choisir h positif de telle sorte que $\mathbb{P}^h[X_T \geq K]$ soit grande. On simule X_T sous $\mathbb{P}^h$ en utilisant l'équation

$$X_t \quad = \quad X_0 + \int_0^t \left(b(X_s) + a(X_s)h_s\right) ds + \int_0^t a(X_s) dW_s^h, \;\; 0 \leq t \leq T \; ,$$

où W^h est un $\mathbb{P}^h$-mouvement brownien. Il faut également simuler

$$H_T^h \quad := \quad \exp\left\{\frac{1}{2}\int_0^T \|h_s\|^2 ds + \int_0^T h_s^* dW_s^h\right\}$$

sous $\mathbb{P}^h$.

Exemple 3.4.2 *On se place dans le modèle de Black-Scholes en dimension 1. On prend $r = 0$, $\sigma = 0.25$, $X_0 = 100$, $T = 1$, et on veut évaluer un call européen de strike $K =$ 150. Le prix exact est d'environ 0.672. On utilise la méthode avec $h = 2$.*

Nombre de simulations	*Sans Importance Sampling*	*Avec h=2*
10 000	0.0484 , [0.584 , 0.774]	0.0046 , [0.668 , 0.686]
100 000	0.0149 , [0.658 , 0.716]	0.0015 , [0.669 , 0.675]

3.4.2 Approche systématique

Comme dans la Section 3.3, on peut trouver un h optimal. On suppose que $F \geq \varepsilon > 0$. Si le payoff est bornée inférieurement, ce qui est en général le cas dans les applications en finance, on peut toujours supposer qu'il est uniformément strictement positif en lui ajoutant une constante (ce qui ne fait qu'ajouter une constante à l'espérance que l'on veut calculer). On suppose également que $F \in L^2$. Dans ce cas, on peut définir

$$h_t = -\frac{1}{\mathbb{E}[F \mid \mathcal{F}_t]}\phi_t^F \ , \quad t \in [0,T] \ ,$$

avec ϕ^F défini comme dans le Théorème 3.3.4. On vérifie que $h \in L^2([0,T] \times \Omega, dt \otimes dP)$ et que $\mathbb{E}[H_T^h] = 1$. On est donc sous les conditions du Théorème 3.4.1. Soit

$$\zeta_t = \mathbb{E}[F \mid \mathcal{F}_t]/\mathbb{E}[F] \ , \quad t \in [0,T] \ .$$

On a d'après la définition de h et ϕ^F

$$\zeta_t = 1 + \frac{1}{\mathbb{E}[F]}\int_0^t (\phi_s^F)^* dW_s = 1 - \int_0^t \zeta_s h_s^* dW_s$$

d'où

$$\zeta_T = \exp\left\{-\frac{1}{2}\int_0^T \|h_s\|^2 \, ds - \int_0^T h_s^* dW_s\right\} = H_T^h \ .$$

On en déduit que $(H_T^h)^{-1}F = \mathbb{E}[F]$ de sorte que la variance de $(H_T^h)^{-1}F$ est nulle.

Comme dans la Section 3.3.2, on peut chercher à utiliser les techniques d'EDP pour approcher h et donc H_T^{-1}. Par exemple, pour les options vanilla, si la solution $u(t,x)$ de l'EDP associée (voir Théorème 2.2.10) est suffisamment régulière, on aura :

$$h_t = -\frac{1}{u(t,X_t)}\nabla_x u(t,X_t) a(X_t) \ .$$

On peut également l'approcher par celui d'un payoff/modèle proche de celui considéré, pour lequel la formule est explicite, voir Remarque 3.3.5. C'est par exemple l'approche suivie par [28].

3.4.3 Fonction d'importance optimale et algorithme de Robbins Monro

Ré-écriture du problème

On se place dans le cadre du modèle de Balck-Scholes, i.e. $b(x) = rx$ et $a(x) = \operatorname{diag}[x]\,\sigma$, avec $d = 1$ et un payoff de la forme

$$G = \psi\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_\kappa} - W_{t_{\kappa-1}}\right) \ .$$

On cherche un changement de mesure optimal avec h déterministe et constant, égal à h^i, sur les intervalles $[t_{i-1}, t_i]$. On note $Z = (W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_\kappa} - W_{t_{\kappa-1}})$ et, par abus de notation, $h = (h^1, \dots, h^\kappa)$. Le problème est donc de minimiser sur $\mathbb{R}^\kappa$ la fonctionnelle

$$H(h) \; := \; \mathbb{E}\left[\left(e^{-\frac{1}{2}\|h\|^2 - h^* Z}\psi(Z+h)\right)^2\right] \; .$$

En utilisant le théorème de Girsanov, on observe que

$$\begin{aligned} H(h) &= \mathbb{E}\left[e^{-\|h\|^2 - 2h^* Z}\psi(Z+h)^2\right] = \mathbb{E}\left[e^{-\frac{1}{2}\|h\|^2 + h^* Z}e^{\|h\|^2 - 2h^* Z}\psi(Z)^2\right] \\ &= \mathbb{E}\left[e^{\frac{1}{2}\|h\|^2 - h^* Z}\psi(Z)^2\right] \; . \end{aligned}$$

Si $\mathbb{E}[\psi(Z)^{2+\varepsilon}] < \infty$ pour un $\varepsilon > 0$, on vérifie facilement en utilisant l'inégalité de Holder que H est deux fois continument dérivable. En particulier

$$\nabla H(h) \; = \; \mathbb{E}\left[(h - Z)e^{\frac{1}{2}\|h\|^2 - h^* Z}\psi(Z)^2\right] \; .$$

Comme il s'agit d'un problème strictement convexe (calculer la Hessienne) et coercif dès que $\psi \geq 0$ et $\psi \neq 0$ (minorer le log en utilisant l'inégalité de Jensen), il existe un optimum $\hat{h}$ qui en outre est la solution de

$$\nabla H(\hat{h}) = \mathbb{E}\left[(\hat{h} - Z)e^{\frac{1}{2}\|\hat{h}\|^2 - \hat{h}^* Z}\psi(Z)^2\right] = 0 \; . \tag{3.4.1}$$

Algorithme de Robbins-Monro

D'après l'équation (3.4.1), on doit résoudre un problème du type

$$f(h) := \mathbb{E}\left[F(h, Z)\right] \; = \; 0$$

où F est une fonction de $\mathbb{R}^\kappa \times \mathbb{R}^d$ dans $\mathbb{R}^\kappa$ et Z une variable aléatoire sur $\mathbb{R}^d$. L'algorithme de Robbins-Monro consiste à simuler une suite $(Z_n)_{n \geq 1}$ de copies indépendantes de Z et à définir la suite $(\zeta_n)_{n \geq 1}$ de variables aléatoires par

$$\zeta_{n+1} = \zeta_n - \gamma_{n+1} F(\zeta_n, Z_{n+1}) \; ,$$

la condition initale ζ_0 étant donnée. Sous des hypothèses classiques on obtient le résultat de convergence suivant.

Théorème 3.4.3 *Soit $\mathcal{F}_n := \sigma(\zeta_k, Y_k \; ; \; k \leq n)$ où $Y_{n+1} := F(\zeta_n, Z_{n+1})$. On suppose qu'il existe $\hat{h}$ vérifiant $f(\hat{h}) = 0$ tel que $(h - \hat{h})^* f(h) > 0$ pour tout $h \in \mathbb{R}^\kappa \setminus \{\hat{h}\}$. On suppose également que la suite $(\gamma_n)_n$ vérifie*

$$\sum_{n \geq 1} \gamma_n = \infty \quad \text{et} \quad \sum_{n \geq 1} \gamma_n^2 < \infty \; . \tag{3.4.2}$$

S'il existe $C > 0$ tel que $\mathbb{E}[\|Y_{n+1}\|^2 \mid \mathcal{F}_n] \leq C(1 + \|\zeta_n\|^2)$ p.s. pour tout $n \geq 1$, alors $\zeta_n \to \hat{h}$ p.s.

Il est clair que les deux conditions sur $\hat{h}$ sont vérifiées dans notre problème 3.4.1. Par ailleurs, on peut toujours choisir $(\gamma_n)_n$ vérifiant (3.4.2). Par contre la dernière condition $\mathbb{E}[\|Y_{n+1}\|^2 \mid \mathcal{F}_n] \leq C(1+\|\zeta_n\|^2)$ p.s. n'a aucune raison d'être vérifiée à cause de l'exponentielle intervenant dans la formule.

Pour palier à ce problème, [3] a proposé une version tronquée de cet algorithme qui évite l'explosion de la suite ζ_n quand la condition d'intégrabilité n'est pas vérifiée. L'algorithme consiste à se donner $x^1 \neq x^2$ et $M > 0$ tels que

$$\max\{f(x^1), f(x^2)\} \;<\; \min\{M, \inf_{\|x\|\geq M} f(x)\}\ ,$$

puis une suite de réels $(u_n)_n$ strictment croissante et tendant vers l'infini, avec $u_0 > M$. On définit ensuite la suite $(\zeta_n)_n$ par

$$\zeta_{n+1} = \begin{cases} \zeta_n - \gamma_{n+1}Y_{n+1} & \text{si} \quad \|\zeta_n - \gamma_{n+1}Y_{n+1}\| \leq u_{\rho(n)}\ , \\ x_n & \text{sinon} \end{cases}$$

avec $x_n = x^1$ si $\rho(n)$ est pair et x^2 sinon, et $\rho(n) = \sum_{k=0}^{n-1} \mathbf{1}_{\|\zeta_k - \gamma_{k+1}Y_{k+1}\| > u_{\rho(k)}}$ avec $\rho(0) = 0$. Le résultat de convergence suivant est démontré dans [3]. Il discute également le choix de la suite $(u_n)_n$.

Théorème 3.4.4 *On suppose qu'il existe $\varepsilon > 0$ tel que $\mathbb{E}[\|\psi(Z)\|^{4+\varepsilon}] < \infty$ et on définit l'algorithme précédent avec F associée au problème* (3.4.1)*. Alors, on peut choisir $(u_n)_n$ et $(\gamma_n)_n$ telles que* (3.4.2) *soit vérifiée et*

$$\sum_{n\geq 1} \gamma_n^2 \mathbb{E}[\|Y_{n+1}\|^2 \mid \mathcal{F}_n] < \infty \ \ p.s.$$

Dans ce cas, $\zeta_n \to \hat{h}$ p.s. Par ailleurs, si

$$\hat{m}_N^r := \frac{1}{N}\sum_{n=1}^{N}\left(e^{-\frac{1}{2}\|\zeta_{n-1}\|^2 - (\zeta_{n-1})^* Z_n}\right)\psi(Z_n + \zeta_{n-1}) \quad et \quad m := \mathbb{E}[\psi(Z)]\ ,$$

alors

$$\hat{m}_N^r \to m \ \ p.s. \quad et \quad \sqrt{N}(\hat{m}_N^r - m) \to N(0, v^2)$$

où $v^2 = Var\Big(\Big(e^{-\frac{1}{2}\|\hat{h}\|^2 - \hat{h}^* Z}\Big)\,\psi(Z + \hat{h})\Big)$.

Preuve. Voir la section 6.4 pour une preuve de la convergence p.s. sous des conditions supplémentaires. □

Ce résultat est en fait un cas une application du cas général traité dans [2]. En particulier, le théorème de la limite centrale énoncé reste vrai si on remplace v^2 par l'estimation

$$\hat{\sigma}_N^2 := \frac{1}{N}\sum_{n=1}^{N}\left(e^{-\frac{1}{2}\|\zeta_{n-1}\|^2 - (X_{n-1})^* Z_n}\right)^2 \psi(Z_n + \zeta_{n-1})^2 - (\hat{m}_N^r)^2$$

dans le sens où $\sqrt{N(\hat{\sigma}_N^2)^{-1}}(\hat{m}_N^r - m) \to N(0,1)$. On a en fait $\hat{\sigma}_N^2 \to v^2$ p.s.

On conclut cette section par la preuve du Théorème 3.4.3.

Preuve du Théorème 3.4.3. Pour simplifier, on suppose que $\kappa = 1$ et que $f(x) = \mathbb{E}[F(x,Z)]$ est bornée uniformément par C. On commence par calculer

$$\begin{aligned}\mathbb{E}\left[|\zeta_{n+1} - \hat{h}|^2 \mid \mathcal{F}_n\right] &= |\zeta_n - \hat{h}|^2 - 2\gamma_{n+1}\mathbb{E}\left[(\zeta_n - \hat{h})Y_{n+1} \mid \mathcal{F}_n\right] \\ &+ \mathbb{E}\left[\gamma_{n+1}^2 Y_{n+1}^2 \mid \mathcal{F}_n\right] .\end{aligned}$$

Comme $\mathbb{E}\left[(\zeta_n - \hat{h})Y_{n+1} \mid \mathcal{F}_n\right] = (\zeta_n - \hat{h})f(\zeta_n) \geq 0$, on en déduit que

$$\mathbb{E}\left[|\zeta_{n+1} - \hat{h}|^2 \mid \mathcal{F}_n\right] \leq |\zeta_n - \hat{h}|^2 + \mathbb{E}\left[\gamma_{n+1}^2 Y_{n+1}^2 \mid \mathcal{F}_n\right] . \tag{3.4.3}$$

L'équation précédente implique également

$$0 \leq 2\mathbb{E}\left[\sum_{n\geq 0}\gamma_{n+1}(\zeta_n - \hat{h})Y_{n+1}\right] \leq \hat{h}^2 + \sum_{n\geq 0}\mathbb{E}\left[\gamma_{n+1}^2 Y_{n+1}^2\right] < \infty \tag{3.4.4}$$

puisque $|f| \leq C$ par hypothèse. En particulier la suite $(S_n)_n$ définie par

$$S_n := \sum_{k=0}^{n}\mathbb{E}\left[\gamma_{n+1}^2 Y_{n+1}^2 \mid \mathcal{F}_n\right]$$

est bornée p.s. par C et donc converge presque surement en tant que suite croissante. Comme $(Z_n)_n$ définie par

$$Z_n := |\zeta_n - \hat{h}|^2 + C - S_n$$

est une sur-martingale par (3.4.3) et qu'une surmartingale bornée inférieurement admet une limite p.s., on en déduit que $(|\zeta_n - \hat{h}|^2)_n$ admet p.s. une limite. On déduit alors de (3.4.4), de l'identité $\mathbb{E}\left[\sum_{n\geq 0}(\zeta_n - \hat{h})Y_{n+1}\right] = \mathbb{E}\left[\sum_{n\geq 0}(\zeta_n - \hat{h})f(\zeta_n)\right]$, de la condition de monotonie sur f et des hypothèses sur $(\gamma_n)_n$ que $\mathbb{P}[\lim_n |\zeta_n - \hat{h}| \geq \delta] = 0$ pour tout $\delta > 0$. Il s'en suit que $\zeta_n \to \hat{h}$ p.s. $\square$

Chapitre 4

Calcul des sensibilités

Connaître la sensibilité d'un portefeuille par rapport aux variations des sous-jacents est tout aussi important que de connaître sa valeur. En effet, ce sont ces sensibilités qui vont permettre de se couvrir. Par ailleurs, on a vu que la connaissance du delta permet de mettre en oeuvre des techniques de réduction de variance.

4.1 Approche par différences finies

On note X^x la solution de (2.0.1). Une première approche pour calculer le delta et le gamma de

$$u(0,x) \;=\; \mathbb{E}\left[g(X^x)\right] \;,$$

consiste à utiliser une méthode de différences finies, i.e.

$$\begin{aligned} \Delta^i(0,x) \;&:=\; \frac{\partial u}{\partial x^i}(0,x) \;\sim\; \frac{1}{2\varepsilon}\left(u(0,x+\varepsilon e_d^i) - u(0,x-\varepsilon e_d^i)\right) \;=:\; \tilde{\Delta}^i(0,x) \\ \Gamma^{ij}(0,x) \;&:=\; \frac{\partial^2 u}{\partial x^i \partial x^j}(0,x) \;\sim\; \frac{1}{2\varepsilon}\left(\tilde{\Delta}^i(0,x+\varepsilon e_d^j) - \tilde{\Delta}^i(0,x-\varepsilon e_d^j)\right) \;=:\; \tilde{\Gamma}^{ij}(0,x) \end{aligned}$$

où $e_d = \mathrm{Vect}[1]$.

L'algorithme consiste à estimer par Monte-Carlo la fonction u en partant des différentes conditions initiales puis à utiliser les approximations $\tilde{\Delta}$ et $\tilde{\Gamma}$.

Il y a essentiellement deux possibilités que nous exposons pour le calcul du delta d'une option vanilla en dimension $d = 1$ (on suppose évidemment que l'on peut simuler $g(X^x)$) :

1/ On utilise N simulations pour estimer une approximation $\hat{u}(0,x+\varepsilon)$ de $u(0,x+\varepsilon)$ et N (autres et indépendantes des premières) simulations pour estimer une approximation

$\hat{u}(0, x-\varepsilon)$ de $u(0, x-\varepsilon)$. Dans ce cas, on a :

$$\begin{aligned}
\operatorname{Var}\left(\frac{\hat{u}(0,x+\varepsilon)-\hat{u}(0,x-\varepsilon)}{2\varepsilon}\right) &= \frac{1}{4\varepsilon^2}\left(\operatorname{Var}\left(\hat{u}(0,x+\varepsilon)\right)+\operatorname{Var}\left(\hat{u}(0,x-\varepsilon)\right)\right) \\
&\sim \frac{1}{4\varepsilon^2}\left(\frac{\operatorname{Var}\left(g(X_T^x)\right)}{N}+\frac{\operatorname{Var}\left(g(X_T^x)\right)}{N}\right) \\
&= \frac{1}{2N\varepsilon^2}\operatorname{Var}\left(g(X_T^x)\right) .
\end{aligned}$$

2/ On simule N trajectoires du brownien et on construit $X^{x+\varepsilon}$ et $X^{x-\varepsilon}$ avec les mêmes trajectoires. Dans ce cas

$$\begin{aligned}
\operatorname{Var}\left(\frac{\hat{u}(0,x+\varepsilon)-\hat{u}(0,x-\varepsilon)}{2\varepsilon}\right) &= \frac{1}{N}\operatorname{Var}\left(\frac{g(X_T^{x+\varepsilon})-g(X_T^{x-\varepsilon})}{2\varepsilon}\right) \\
&\sim \frac{1}{N}\operatorname{Var}\left(g'(X_T^x)\right) .
\end{aligned}$$

Si ε est petit et g régulière, la seconde méthode sera en général préférable à la première.

Cette méthode est très simple à mettre en oeuvre mais le choix du ε n'est pas évident. Si ε est trop petit, la variance de l'estimateur peut être très grande, c'est le cas si le payoff est très irrégulier. Si ε est trop grand, l'approximation des dérivées est mauvaise, voir l'Exemple 4.2.4 ci-dessous. Les vitesses de convergence de ces méthodes ont été étudiées par [29], [30] et [47].
Typiquement, en utilisant l'approche 2/, on obtient une vitesse de convergence en loi en $\sqrt{N}$ si $N^{\frac{1}{4}}\varepsilon \to 0$ lorsque u est C^3. Si on utilise un schéma de type "forward", i.e. $(\hat{u}(0,x+\varepsilon)-\hat{u}(0,x))/\varepsilon$, on obtient une vitesse en loi en $\sqrt{N}$ si $N^{\frac{1}{2}}\varepsilon \to 0$ lorsque u est C^2.

On peut également consulter [22] pour des résultats sur les vitesses lorsque l'on remplace X par son schéma d'Euler.

L'objet des sections suivantes est de donner une interprétation probabiliste de ces sensibilités qui ne fasse par intervenir d'approximation.

4.2 Grecques dans le modèle de Black et Scholes

On note X^x la solution de (2.1.1) avec $d = 1$ et la condition initiale $X_0 = x$. On considère le payoff $g(X_T^x)$ et on cherche à donner une représentation probabiliste des dérivées de

$$u(0,x) = \mathbb{E}\left[g(X_T^x)\right] .$$

Proposition 4.2.1 *Supposons que g est à croissance polynomiale. Alors,*

$$\frac{\partial u}{\partial x}(0,x) = \frac{1}{x\sigma T}\mathbb{E}\left[g\left(X_T^x\right)W_T\right] .$$

Preuve. On suppose que g est C_b^1, le résultat général étant obtenu par densité. Tout d'abord, on remarque que

$$\frac{\partial X_T^x}{\partial x} = e^{\left(b-\sigma^2/2\right)T+\sigma W_T} = \frac{1}{x}X_T^x \;\; \mathbb{P}-\text{p.s.}$$

On en déduit que

$$\frac{\partial g}{\partial x}\left(X_T^x\right) = \frac{1}{x}g'\left(X_T^x\right)X_T^x \;\; \mathbb{P}-\text{p.s.}$$

Par ailleurs,

$$\frac{1}{\varepsilon}\left\|g\left(X_T^{x+\varepsilon}\right)-g\left(X_T^x\right)\right\| \leq \|g'\|_\infty \frac{\left\|X_T^{x+\varepsilon}-X_T^x\right\|}{\varepsilon} = \|g'\|_\infty \left\|e^{\left(b-\sigma^2/2\right)T+\sigma W_T}\right\| \in L^2 .$$

Par convergence dominée, on a donc :

$$\frac{\partial u}{\partial x}(0,x) = \frac{1}{x}\mathbb{E}\left[g'\left(X_T^x\right)X_T^x\right] .$$

Soit f la densitée de la loi normale centrée réduite. L'équation précédente se ré-écrit :

$$\begin{aligned}
\frac{\partial u}{\partial x}(0,x) &= \int_{\mathbb{R}} g'\left(xe^{\left(b-\sigma^2/2\right)T+\sigma\sqrt{T}w}\right)e^{\left(b-\sigma^2/2\right)T+\sigma\sqrt{T}w}f(w)dw \\
&= \int_{\mathbb{R}} \frac{1}{x\sigma\sqrt{T}}\frac{\partial g}{\partial w}\left(xe^{\left(b-\sigma^2/2\right)T+\sigma\sqrt{T}w}\right)f(w)dw \\
&= \int_{\mathbb{R}} \frac{1}{x\sigma\sqrt{T}}g\left(xe^{\left(b-\sigma^2/2\right)T+\sigma\sqrt{T}w}\right)wf(w)dw
\end{aligned}$$

en intégrant par parties, ce qui donne le résultat.

□

Exercice 4.2.2 *Montrer que si g est à croissance polynomiale :*

$$\frac{\partial^2 u}{\partial x^2}(0,x) = \frac{1}{x^2\sigma T}\mathbb{E}\left[g\left(X_T^x\right)\left(\frac{W_T^2}{\sigma T}-W_T-\frac{1}{\sigma}\right)\right] . \tag{4.2.1}$$

Donner une représentation similaire pour le véga : $\frac{\partial u}{\partial \sigma}(0,x)$.

Exercice 4.2.3 *Touver une formulation du delta et du gamma pour les options asiatiques.*

Exemple 4.2.4 *On utilise le résultat de* (4.2.1) *pour estimer*

$$\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\mathbf{1}_{\{X_T^x\in[a,b]\}}\right] .$$

On prend comme paramètres $x = 1$, $\sigma = 0.25$, $r = 0$, $a = 0.95$ *et* $b = 1.05$. *On donne les intervalles de confiance obtenus par l'approche par différences finies avec différentes valeurs de* ε *et par* (4.2.1). *On fait à chaque fois* 50.000 *simulations. La valeur exacte est d'environ* -2.53.

Méthode	*Intervalle estimé*
Par (4.2.1)	$[-2.61\ ,\ -2.51]$
Diff. finies $\varepsilon = 0.5$	$[-0.32\ ,\ -0.31]$
Diff. finies $\varepsilon = 0.1$	$[-2.47\ ,\ -2.06]$
Diff. finies $\varepsilon = 0.05$	$[-3.34\ ,\ -1.68]$
Diff. finies $\varepsilon = 0.005$	$[-20.31\ ,\ 23.91]$

On observe que le résultat est extrêmement biaisé pour $\varepsilon = 0.5$ *et* $\varepsilon = 0.1$. *Pour* $\varepsilon = 0.005$, *il est trop volatil. Même pour* $\varepsilon = 0.05$, *le résultat est bien moins précis que celui obtenu par* (4.2.1).

Remarque 4.2.5 Dans la preuve de la Proposition 4.2.1, on a utilisé deux notions de différentiation :
(i) On a dérivé X_T^x par rapport à sa condition initiale x. Le processus $Y_\cdot = \frac{\partial X_\cdot^x}{\partial x}$ s'appelle le processus tangent à X^x.
(ii) On a utilisé la dérivée de $g\left(xe^{\left(b-\sigma^2/2\right)T+\sigma\sqrt{T}w}\right)$ par rapport à w. A une normalisation près, cela revient à dériver $g\left(X_T^x\right)$ par rapport à W_T. On a ensuite utiliser une intégration par parties par rapport à w, c'est-à-dire "par rapport à" W.

Au vu de la remarque précédente, on a besoin de deux notions : celle de processus tangent, et celle de dérivation par rapport au mouvement Brownien. L'objet des sections suivantes est de définir ces deux notions. On les utilisera ensuite pour généraliser les résultats obtenus dans le modèle de Black et Scholes.

4.3 Processus tangent

4.3.1 Notions de processus tangent

On note X^x la solution de (2.0.1) avec pour condition initiale $X_0^x = x$.

Théorème 4.3.1 *Si* $a,\ b \in C_b^1$, *alors, pour tout* $t \in [0,T]$, *l'application* $x \longmapsto X_t^x$ *est p.s.* C^1. *Le processus (matriciel) gradient* $\nabla_x X^x = \{(\frac{\partial X^{x,i}}{\partial x^j})_{i,j}\}$ *est solution de l'EDS :*

$$\nabla_x X_t^x \quad = \quad I_d + \int_0^t \nabla b(X_s^x)\nabla_x X_s^x ds + \sum_{j=1}^d \int_0^t \nabla a^{\cdot j}(X_s^x)\nabla_x X_s^x dW_s^j\ . \qquad (4.3.1)$$

(voir Théorème 39 dans [58].)

Le processus $\nabla_x X^x$ est appelé processus tangent ou de dérivée première.

Exemple 4.3.2 *Pour $d = 1$, l'équation (4.3.1) s'écrit*

$$\nabla_x X_t^x \;=\; 1 + \int_0^t b'(X_s^x)\nabla_x X_t^x ds + \int_0^t a'(X_s^x)\nabla_x X_t^x dW_s$$

dont la solution est

$$\nabla_x X_t^x \;=\; e^{\int_0^t \left(b'(X_s^x) - |a'(X_s^x)|^2/2\right)ds + \int_0^t a'(X_s^x)dW_s} .$$

Exemple 4.3.3 *Si $b(x) = rx$ et $a(x) = \operatorname{diag}[x]\,\sigma$ comme dans le modèle de Black-Scholes, on obtient*

$$\nabla_x X_t^x \;=\; \operatorname{diag}\left[\left(e^{(r-\left\|\sigma^{j\cdot}\right\|^2/2)t+\sigma^{j\cdot}W_t}\right)_{j=1}^d\right] \;=\; \operatorname{diag}[x]^{-1} X_t^x \;=\; X_t^{e_d} .$$

Remarque 4.3.4 En imposant plus de régularité sur a et b, on peut définir des processus dérivée d'ordre supérieur. On considère alors le processus tangent du processus tangent, etc...

Remarque 4.3.5 On peut également définir le processus dérivée, $\nabla_x \bar{X}^{x,n}$, du schéma d'Euler $\bar{X}^{x,n}$ de X^x. Il est obtenu par récurrence

$$\begin{aligned} \nabla_x \bar{X}_0^{x,n} &= I_d \\ \nabla_x \bar{X}_{t_i}^{x,n} &= \nabla_x \bar{X}_{t_{i-1}}^{x,n} + \nabla b(\bar{X}_{t_{i-1}}^{x,n})\nabla_x \bar{X}_{t_{i-1}}^{x,n}\Delta^n t + \sum_{j=1}^d \nabla a^{\cdot j}(\bar{X}_{t_{i-1}}^{x,n})\nabla_x \bar{X}_{t_{i-1}}^{x,n}\Delta^n W_i^j \end{aligned} \tag{4.3.2}$$

L'équation (4.3.1) peut être vue comme limite de (4.3.2) quand le pas de temps tend vers 0, $\nabla_x \bar{X}^{x,n}$ coincinde avec le schéma d'Euler de $\nabla_x X^x$.

4.3.2 Processus tangent et delta

On peut maintenant reprendre les arguments de la Section 4.2 pour calculer ∇u. On suppose tout d'abord que g est C_b^1 de sorte que par convergence dominée :

$$\nabla u(0,x) \;=\; \mathbb{E}\left[\nabla g(X_T^x)\nabla_x X_T^x\right] .$$

Par passage par densité, on obtient le résultat pour toute fonction g dont les dérivées premières sont à croissance polynomiale. On peut également l'obtenir pour des payoffs dépendant de la trajectoire de X^x aux dates t_i.

Proposition 4.3.6 *Soit g une fonction C_p^1 de $\mathbb{R}^{dn}$ dans $\mathbb{R}$. Alors,*

$$\nabla \mathbb{E}\left[g(X_{t_1}^x, \ldots, X_{t_n}^x)\right] \;=\; \sum_{i=1}^n \mathbb{E}\left[\nabla_i g(X_{t_1}^x, \ldots, X_{t_n}^x)\nabla_x X_{t_i}^x\right] , \tag{4.3.3}$$

où $\nabla_i g$ denote le gradient de g par rapport à son i-ème argument (vectoriel).

Remarque 4.3.7 On obtient le même résultat si on remplace X^x par son schéma d'Euler

$$\nabla \mathbb{E}\left[g(\bar{X}^{x,n}_{t_1},\dots,\bar{X}^{x,n}_{t_n})\right] \;=\; \sum_{i=1}^{n} \mathbb{E}\left[\nabla_i g(\bar{X}^{x,n}_{t_1},\dots,\bar{X}^{x,n}_{t_n})\nabla_x \bar{X}^{x,n}_{t_i}\right] \;.$$

Au vu de la Remarque 4.3.5, cela revient en fait à discrétiser l'Equation (4.3.3).

Exemple 4.3.8 *On considère un call sur moyenne discrète dans le modèle de Black-Scholes en dimension* 1*. La maturité est* 1 *et la moyenne est calculée à intervalles réguliers de longueur* 1/24*, i.e. tous les* 15 *jours. On prend comme paramètre* $K = 100$*,* $X_0 = 100$*,* $\sigma = 0.35$ *et* $r = 0.1$*. On considère* $n = 24$ *pas de temps, et on estime le delta par la formule*

$$\mathbb{E}\left[\mathbf{1}_{A_1 \geq K} \sum_{i=1}^{24} \frac{1}{24} e^{\frac{i}{24}(r-\sigma^2/2)+\sigma W_{\frac{i}{24}}}\right] \;=\; \frac{1}{100}\mathbb{E}\left[A_1 \mathbf{1}_{A_1 \geq K}\right]$$

avec

$$A_1 \;:=\; \frac{1}{24}\sum_{i=1}^{24} X_0 e^{\frac{i}{24}(r-\sigma^2/2)+\sigma W_{\frac{i}{24}}} \;.$$

Les résultats suivants ont été obtenus en utilisant le contrôle antithétique.

Nombre de simulations	*Intervalle de confiance du delta*
5 000	$[0.598, 0.605]$
10 000	$[0.601, 0.606]$
50 000	$[0.601, 0.604]$

Exercice 4.3.9 *Donner une représentation du gamma sur le modèle de la Proposition 4.3.6.*

4.3.3 Processus tangent et véga (exercice)

On peut facilement étendre l'approche précédente pour estimer le ϑ, i.e. la sensibilité du prix de l'option par rapport à une pertubation sur la matrice de volatilité a. On se donne une fonction $\tilde{a}$ de $\mathbb{R}^d$ dans $\mathbb{M}^d$, $\varepsilon > 0$ et on note $X^{\varepsilon\tilde{a}}$ la solution de (2.0.1) avec la matrice de volatilité $a + \varepsilon\tilde{a}$ à la place de a. On cherche à estimer

$$\vartheta(\tilde{a}) \;:=\; \frac{\partial}{\partial\varepsilon}\mathbb{E}\left[g(X^{\varepsilon\tilde{a}}_T)\right]\Big|_{\varepsilon=0} \;.$$

On suppose que a, b et $\tilde{a} \in C^1_p$

1/ En considérant ε comme un processus déterministe, déduire du Théorème 4.3.1 que $Z^{\tilde{a}} := \frac{\partial}{\partial \varepsilon} X^{\varepsilon \tilde{a}}\big|_{\varepsilon=0}$ est bien défini et est solution de

$$Z_t^{\tilde{a}} \;=\; \int_0^t \nabla b(X_s) Z_s^{\tilde{a}} ds + \int_0^t \tilde{a}(X_s) dW_s + \sum_{j=1}^d \int_0^t \nabla a^{\cdot j}(X_s) Z_s^{\tilde{a}} dW_s^j \;\; , \;\; t \in [0,T] \, .$$

2/ En déduire une formulation de la dérivée directionnelle $\vartheta(\tilde{a})$ sur le modèle de la Proposition 4.3.6.

4.4 Calcul de Malliavin

Les deux derniers résultats donnent une représentation précieuse du gradient, puisqu'elle fournissent des estimateurs de Monte-Carlo très naturels. Evidemment, elles sont utiles à des fins de couverture mais peuvent également être employées dans le cadre des techniques de réduction de variance présentées dans le Chapitre 3. Toutefois, ces formulations imposent la dérivabilité du payoff (au moins dans un sens faible) et surtout la connaissance de cette dérivée. Ceci n'est pas si évident. Si l'on veut calculer le delta d'un book d'options, on ne connaît pas forcément dans le détail la forme exacte de tout les payoffs (ils sont généralement fournis par un ordinateur qui agit comme une boîte noire). Dans la Proposition 4.2.1, on a vu que, dans le cadre particulier du modèle de Black-Scholes, on pouvait obtenir une formulation ne faisant pas intervenir le gradient de g. Pour cela, on avait utilisé une idée d'intégration par parties par rapport à la densité gaussienne. Dans cette section, on va introduire une notion de calcul différentiel par rapport à la trajectoire du mouvement brownien, et une formule d'intégration par parties associée. Nous renvoyons à [54] et [55] pour une présentation complète du calcul de Malliavin.

4.4.1 Introduction au calcul de Malliavin

Définition 4.4.1 *On dira qu'une variable aléatoire F de L^2 est simple s'il existe un entier k_F, une suite $0 \leq s_1^F \ldots < s_{k_F}^F \leq T$ et une fonction ϕ^F de $(\mathbb{R}^d)^{k_F}$ dans $\mathbb{R}$, continue, C^1 par morceaux, tels que :*

$$F \;=\; \phi^F\left(W_{s_1^F}, \ldots, W_{s_{k_F}^F}\right) \quad \textit{avec} \quad \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t,T]}(s_j^F) \in L^2 \quad \textit{pour tout } t \in [0,T] \, .$$

On note $\mathcal{S}$ l'espace de telles fonctions.

Bien qu'il soit défini sur un espace beaucoup plus gros que $\mathcal{S}$, voir [54] et [56], on n'introduira ici le calcul de Malliavin que pour des fonctions simples. Il y a deux raisons à cela :

1/ En général, si on sait simuler parfaitement F, c'est que $F \in \mathcal{S}$.

2/ C'est beaucoup plus simple, et nous pourrons mener les preuves jusqu'au bout.

3/ Discrétiser des formules obtenues en travaillant sur X ou travailler directement sur le problème discrétisé associé à $\bar{X}^n$ revient généralement au même, voir Remarque 4.4.18. On ne perd donc rien en se plaçant tout de suite dans un cadre plus simple à gérer.

Définition 4.4.2 *Soit $F \in \mathcal{S}$, pour tout $t \in [0,T]$, on définit*

$$D_t F := \lim_{\varepsilon \to 0} \frac{\phi^F\left(W + \varepsilon \mathbf{1}_{[t,T]}\right) - \phi^F(W)}{\varepsilon} = \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t,T]}(s_j^F) \ \mathbb{P}\text{-}p.s.$$

Cette définition peut être comprise comme ceci : on choque légèrement la trajectoire du brownien W en la remplaçant par une trajectoire $W^{\varepsilon \mathbf{1}_{[t,T]}} = W + \varepsilon \mathbf{1}_{[t,T]}$, i.e. on "shifte" le brownien de ε après t. On regarde ensuite l'impact de ce choc sur F.

On appelle DF le processus dérivée de Malliavin.

Exemple 4.4.3 *On fixe s, $t \in [0,T]$.*
1/ s est indépendant de W et donc $D_t s = 0$.
2/ Si $F = W_s$ on a

$$D_t F = \lim_{\varepsilon \to 0} \frac{W_s + \varepsilon \mathbf{1}_{[t,T]}(s) - W_s}{\varepsilon} = \mathbf{1}_{[t,T]}(s) .$$

3/ Dans le modèle de Black-Scholes , on a simplement

$$D_t x e^{(r-\sigma^2/2)s + \sigma W_s} = \sigma x e^{(r-\sigma^2/2)s+\sigma W_s} \mathbf{1}_{[t,T]}(s) .$$

Remarque 4.4.4 $D_t F$ n'est en général pas adapté, voir l'exemple 3/ ci-dessus.

Les propriétés suivantes découlent immédiatement de la définition.

Propriété 4.4.5 *Soient F et $G \in \mathcal{S}$ et $\psi \in C_p^1$. Alors,*
(i) FG et $\psi(F) \in \mathcal{S}$,
(ii) $D\psi(F) = \psi'(F) DF$,
(iii) $D(FG) = (DF)G + F(DG)$.

Remarque 4.4.6 Si b et $a \in C_b^1$, alors $\bar{X}_{t_i}^{x,n}$ est une fonction déterministe de $(W_{t_j},\ 0 \leq j \leq i)$, et $\bar{X}_{t_i}^{x,n} \in \mathcal{S}$. Sa dérivée de Malliavin en t peut être calculée récursivement :

$$\begin{aligned}
D_t \bar{X}_{t_i}^{x,n} &= 0 \quad \text{si} \quad t_i < t \\
D_t \bar{X}_{t_i}^{x,n} &= a(\bar{X}_{t_{i-1}}^{x,n}) \quad \text{si} \quad t_{i-1} < t \leq t_i \qquad (4.4.1) \\
D_t \bar{X}_{t_i}^{x,n} &= D_t \bar{X}_{t_{i-1}}^{x,n} + \nabla b(\bar{X}_{t_{i-1}}^{x,n}) D_t \bar{X}_{t_{i-1}}^{x,n} \Delta^n t + \sum_{j=1}^{d} \nabla a^{\cdot j}(\bar{X}_{t_{i-1}}^{x,n}) D_t \bar{X}_{t_{i-1}}^{x,n} \Delta^n W_i^j \quad \text{si} \quad t < t_{i-1} .
\end{aligned}$$

Remarque 4.4.7 On peut observer que le gradient et la dérivée de Malliavin en t de $\bar{X}^n$ suivent la même équation, voir Remarque 4.3.5, à la condition en ϕ_t^n près

$$D_t \bar{X}^{x,n}_{\phi_t^n} \;=\; 0 \;\neq\; \nabla_x \bar{X}^{x,n}_{\phi_t^n} \,.$$

Si $a(x)$ est inversible pour tout $x \in \mathbb{R}^d$, un calcul direct montre que

$$\nabla_x \bar{X}^{x,n}_{t_i} \;=\; D_t \bar{X}^{x,n}_{t_i} a(\bar{X}^{x,n}_{\phi_t^n})^{-1} \nabla_x \bar{X}^{x,n}_{\phi_t^{n,+}} \quad \text{pour tout } t \leq t_i \leq T \,.$$

où $\phi^{n,+}(t) = \inf\{t_i,\ i \in \{0,\ldots,n\} \ :\ t \leq t_i\}$.

L'intérêt de cette notion réside dans le fait que la stratégie de couverture d'une option peut s'écrire en fonction de la dérivée de Malliavin du payoff.

Théorème 4.4.8 (Formule de Clark-Ocone) *Soit $F \in \mathcal{S}$, alors*

$$F \;=\; \mathbb{E}\left[F\right] + \int_0^T \mathbb{E}\left[D_t F \mid \mathcal{F}_t\right] dW_t \,.$$

Preuve. Pour simplifier, on se place dans le cas où $d = 1$. On suppose également que $\phi^F \in C_b^2$. Le cas général est obtenu par densité. A $s_{i-1}^F < t \leq s_i^F$ fixé, $\mathbb{E}\left[F \mid \mathcal{F}_t\right]$ est une fonction $\psi(t, W_t, (W_{s_j^F})_{j \leq i-1})$. On note ψ' sa dérivée par rapport au deuxième argument. On a alors pour $s_{i-1}^F < t \leq s_i^F$

$$\psi'(t,w,z) \;=\; \lim_{\varepsilon \to 0} \varepsilon^{-1} \mathbb{E}\left[\phi^F\left(z, \tilde{\omega}^i + w + \varepsilon, \ldots, \tilde{\omega}^n + w + \varepsilon\right) - \phi^F\left(z, \tilde{\omega}^i + w, \ldots, \tilde{\omega}^n + w\right)\right]$$

où $\tilde{\omega}$ est une variable aléatoire à valeurs dans $\mathbb{R}^{n-i+1}$ distribuée selon une loi normale $\mathcal{N}(0, \Sigma)$ avec

$$\Sigma^{lk} \;=\; \min\{s_l^F - t\ ,\ s_k^F - t\} \,.$$

On a donc par convergence dominée

$$\psi'(t,w,z) \;=\; \mathbb{E}\left[\sum_{j=i}^{k_F} \nabla_j \phi^F\left(z, \omega^i + w, \ldots, \omega^n + w\right)\right]$$

d'où pour $t \in (s_{i-1}^F, s_i^F]$

$$\psi'\left(t, W_t, (W_{s_j^F})_{j \leq i-1}\right) \;=\; \mathbb{E}\left[\sum_{j=i}^{k_F} \nabla_j \phi^F(W) \;\mid \mathcal{F}_t\right] \;=\; \mathbb{E}\left[D_t \phi^F(W) \;\mid \mathcal{F}_t\right] \,.$$

Comme $\phi^F \in C_b^2$, on peut vérifier par des arguments similaires que ψ est $C_b^{1,2}$ par rapport à ces deux premiers arguments. En appliquant le Lemme d'Itô à la martingale $\psi(t, W_t, (W_{s_j^F})_{j \leq i-1})$ sur $(s_{i-1}^F, s_i^F]$, on obtient alors

$$\begin{aligned}
\mathbb{E}\left[F \mid \mathcal{F}_{s_i^F}\right] \;&=\; \mathbb{E}\left[F \mid \mathcal{F}_{s_{i-1}^F}\right] + \int_{s_{i-1}^F}^{s_i^F} \psi'\left(t, W_t, (W_{s_j^F})_{j \leq i-1}\right) dW_t \\
&=\; \mathbb{E}\left[F \mid \mathcal{F}_{s_{i-1}^F}\right] + \int_{s_{i-1}^F}^{s_i^F} \mathbb{E}\left[D_t \phi^F(W) \;\mid \mathcal{F}_t\right] dW_t \,.
\end{aligned}$$

Cette relation étant vérifiée pour tout $i \in \{1, \ldots, k\}$, en sommant le système d'équations obtenu, on en déduit que

$$\begin{aligned} F \;=\; \mathbb{E}\left[F \mid \mathcal{F}_{s_k^F}\right] &= \mathbb{E}\left[F \mid \mathcal{F}_0\right] + \int_0^{s_k^F} \mathbb{E}\left[D_t \phi^F(W) \mid \mathcal{F}_t\right] dW_t \;. \\ &= \mathbb{E}\left[F\right] + \int_0^{T} \mathbb{E}\left[D_t \phi^F(W) \mid \mathcal{F}_t\right] dW_t \;. \end{aligned}$$

□

On peut maintenant énoncer le théorème d'intégration par parties du calcul de Malliavin.

Théorème 4.4.9 (Formule d'intégration par parties) *Soit $F \in \mathcal{S}$ et soit un processus adapté $h \in L^2([0,T] \times \Omega, dt \otimes dP)$, alors*

$$\mathbb{E}\left[F \int_0^T h_t^* dW_t\right] \;=\; \mathbb{E}\left[\int_0^T D_t F h_t dt\right] \;.$$

Preuve. On pose $X_t = \mathbb{E}\left[F \mid \mathcal{F}_t\right]$ et $H_t = \int_0^t h_s^* dW_s$. On remarque que $H_0 = 0$. D'après le Théorème 4.4.8, on a donc par le Lemme d'Itô :

$$F \int_0^T h_t^* dW_t \;=\; X_T H_T \;=\; \int_0^T \left(X_t h_t^* + H_t \mathbb{E}\left[D_t F \mid \mathcal{F}_t\right]\right) dW_t + \int_0^T \mathbb{E}\left[D_t F \mid \mathcal{F}_t\right] h_t dt \;.$$

On obtient donc

$$\mathbb{E}\left[F \int_0^T h_t^* dW_t\right] \;=\; \mathbb{E}\left[\int_0^T \mathbb{E}\left[D_t F \mid \mathcal{F}_t\right] h_t dt\right] \;=\; \mathbb{E}\left[\int_0^T D_t F h_t dt\right]$$

où l'on a utilisé le Lemme de Fubini et le fait que h est adapté. □

Remarque 4.4.10 Le Théorème 4.4.9 reste vrai, dans une certaine mesure, même si h n'est pas adapté. Dans ce cas, l'intégrale $\int_0^T h_t^* dW_t$ est définie en tant qu'intégrale de Skorohod, voir [54]. Lorsque $h = Fu$ où F est une variable aléatoire $\mathcal{F}_T$-mesurable et u est un processus adapté, on obtient, sous certaines hypothèses de régularité et d'intégrabilité, un lien entre l'intégrale d'Itô et de Skorohod :

$$\int_0^T F u_t^* dW_t \;=\; F \int_0^T u_t^* dW_t - \int_0^T D_t F u_t dt \;.$$

4.4.2 Calcul de Malliavin et sensibilités

4.4.2.1 Delta pour les variables simples

On commence par utiliser les résultats de la section précédente pour donner une représentation probabiliste du delta. Soit $F \in \mathcal{S}$. Alors, d'après le Théorème 4.4.8

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] \, dW_t \, .$$

On fixe maintenant t et i tels que $s_{i-1}^F < t \leq s_i^F$, on a alors

$$D_t F = \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t,T]}(s_j^F) = \sum_{j=i}^{k_F} \nabla_j \phi^F(W)$$

qui est indépendant de $t \in (s_{i-1}^F, s_i^F]$. On a donc

$$D_t F = \int_t^{s_i^F} \frac{D_s F}{s_i^F - t} ds$$

d'où, en utilisant le Théorème 4.4.9

$$\mathbb{E}[D_t F \mid \mathcal{F}_t] = \mathbb{E}\left[\int_t^{s_i^F} \frac{D_s F}{s_i^F - t} ds \mid \mathcal{F}_t\right] = \mathbb{E}\left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} \mid \mathcal{F}_t\right] \, .$$

On obtient finalement le résultat suivant

Proposition 4.4.11 *Soit $F \in \mathcal{S}$, alors*

$$F = \mathbb{E}[F] + \sum_{i=1}^{k_F} \int_{s_{i-1}^F}^{s_i^F} \mathbb{E}\left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} \mid \mathcal{F}_t\right] dW_t \, .$$

A $s_{i-1}^F < t \leq s_i^F$ fixé, la stratégie de couverture est donc donnée par la quantité $\mathbb{E}\left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} \mid \mathcal{F}_t\right]$, après renormalisation, qui peut être approchée par Monte-Carlo.

Exercice 4.4.12 *Ecrire la stratégie de couverture d'une option européenne dans un marché complet, en fonction du delta. Retrouver le résultat de la Proposition 4.2.1 en utilisant la Proposition 4.4.11.*

4.4.2.2 Delta et Gamma pour les fonctions du schéma d'Euler

On suppose maintenant que $F = g(\bar{X}_T^{x,n})$. On suppose également que $g \in C_p^1$, que a est inversible sur $\mathbb{R}^d$ et on pose

$$u(0, x) := \mathbb{E}\left[g(\bar{X}_T^{x,n})\right] \, .$$

D'après la Proposition 4.3.6, on a

$$\nabla u(0,x) \;=\; \mathbb{E}\left[\nabla g(\bar{X}_T^{x,n})\nabla_x \bar{X}_T^{x,n}\right] \;,$$

ce qui peut se réécrire, en utilisant la Remarque 4.4.7 et la Propriété 4.4.5

$$\begin{aligned}\nabla u(0,x) &= \frac{1}{T}\mathbb{E}\left[\int_0^T \nabla g(\bar{X}_T^{x,n})D_t\bar{X}_T^{x,n}a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_x\bar{X}_{\phi_t^{n,+}}^{x,n}dt\right]\\ &= \frac{1}{T}\mathbb{E}\left[\int_0^T D_t g(\bar{X}_T^{x,n})a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_x\bar{X}_{\phi_t^{n,+}}^{x,n}dt\right] \;.\end{aligned}$$

En utilisant le Théorème 4.4.9, on obtient

$$\nabla u(0,x) \;=\; \frac{1}{T}\mathbb{E}\left[g(X_T^{x,n})\left(\int_0^T\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_x\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right)^*\right] \;.$$

En passant par densité, on obtient

Proposition 4.4.13 *Si $a(x)$ est inversible pour tout $x\in\mathbb{R}^d$, a^{-1} et g sont à croissance polynomiale, a et b sont C_b^1, alors*

$$\nabla\mathbb{E}\left[g(\bar{X}_T^{x,n})\right] \;=\; \frac{1}{T}\mathbb{E}\left[g(X_T^{x,n})\left(\int_0^T\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_x\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right)^*\right] \;.$$

On retrouve en particulier le résultat de la Proposition 4.2.1 en utilisant l'Exemple 4.3.3.

Exercice 4.4.14 *Montrer que sous les conditions de la Proposition 4.4.13*

$$\nabla\mathbb{E}\left[g\left(X_{t_1}^{x,n},\dots,X_{t_n}^{x,n}\right)\right] \;=\; \mathbb{E}\left[g\left(X_{t_1}^{x,n},\dots,X_{t_n}^{x,n}\right)\left(\int_0^T h_t\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_x\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right)^*\right] \;,$$

où $h\in L^2([0,T],dt)$ vérifie

$$\int_0^{t_i} h_t dt \;=\; 1 \quad \textit{pour tout} \quad i\in\{1,\dots,n\} \;.$$

On s'intéresse maintenant au gamma de $\mathbb{E}\left[g(\bar{X}_T^{x,n})\right]$. Si on suppose g, a, a^{-1} et b suffisamment régulières, on obtient en argumentant comme dans la Proposition 4.3.6 et en utilisant la Proposition 4.4.13 que :

$$\begin{aligned}\frac{\partial^2 u}{\partial x^i\partial x^j}(0,x) &= \frac{1}{T}\frac{\partial}{\partial x^j}\mathbb{E}\left[g\left(X_T^{x,n}\right)\int_0^T\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_{x^i}\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right]\\ &= \frac{1}{T}\,\mathbb{E}\left[\nabla g(X_T^{x,n})\nabla_{x^j}\bar{X}_T^{x,n}\int_0^T\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_{x^i}\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right]\\ &\quad+\frac{1}{T}\,\mathbb{E}\left[g(X_T^{x,n})\int_0^T\left(\sum_{k=1}^d\nabla\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\right)^{.k}\nabla_{x^j}\bar{X}_{\phi_t^n}^{x,n}\left(\nabla_{x^i}\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^k\right)^* dW_t\right]\\ &\quad+\frac{1}{T}\,\mathbb{E}\left[g(X_T^{x,n})\int_0^T\left(a(\bar{X}_{\phi_t^n}^{x,n})^{-1}\nabla_{x^j}\nabla_{x^i}\bar{X}_{\phi_t^{n,+}}^{x,n}\right)^* dW_t\right] \;,\end{aligned}$$

où ∇_{x^ℓ} est la dérivée par rapport à x^ℓ, i.e. $\nabla_{x^\ell}\bar{X}^{x,n}_{\phi_t^{n,+}} = \left(\nabla_x \bar{X}^{x,n}_{\phi_t^{n,+}}\right)^{.\ell}$. On s'intéresse uniquement au premier terme qui fait intervenir ∇g. Pour une processus adapté h de carré intégrable, on note maintenant

$$\delta\left(h_t\right) \ := \ \int_0^T h_t^* dW_t \ .$$

En utilisant la Remarque 4.4.7 et la Propriété 4.4.5, on obtient

$$\begin{aligned}
A \ &:= \ \mathbb{E}\left[\nabla g(X_T^{x,n})\nabla_{x^j}\bar{X}_T^{x,n}\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\right] \\
&= \ \frac{1}{T}\,\mathbb{E}\left[\int_0^T \nabla g(\bar{X}_T^{x,n})D_t\bar{X}_T^{x,n}\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}dt\right] \\
&= \ \frac{1}{T}\,\mathbb{E}\left[\int_0^T D_t\left(g(\bar{X}_T^{x,n})\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\right)a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}dt\right] \\
&\quad -\frac{1}{T}\mathbb{E}\left[\int_0^T g(\bar{X}_T^{x,n})D_t\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}dt\right] \ .
\end{aligned}$$

D'après le Théorème 4.4.9, ceci implique que

$$\begin{aligned}
A \ &= \ \frac{1}{T}\,\mathbb{E}\left[g(\bar{X}_T^{x,n})\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\right] \\
&\quad -\frac{1}{T}\mathbb{E}\left[\int_0^T g(\bar{X}_T^{x,n})D_t\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}dt\right] \ .
\end{aligned}$$

En regroupant les termes et en passant par densité, on obtient

Proposition 4.4.15 *Si a est inversible sur $\mathbb{R}^d$, $a^{-1} \in C_p^1$, g est à croissance polynomiale, a et b sont C_b^2, alors*

$$\frac{\partial}{\partial x^i \partial x^j}\mathbb{E}\left[g(\bar{X}_T^{x,n})\right] \ = \ \mathbb{E}\left[g(X_T^{x,n})\Pi^{ij}\right] \ ,$$

où

$$\begin{aligned}
\Pi^{ij} \ &:= \ \frac{1}{T^2}\left(\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)\right) \\
&\quad -\frac{1}{T^2}\left(\int_0^T D_t\delta\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}dt\right) \\
&\quad +\frac{1}{T}\,\delta\left(\sum_{k=1}^d \nabla\left(a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\right)^{.k}\nabla_{x^j}\bar{X}^{x,n}_{\phi_t^{n,+}}\left(\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right)^k + a(\bar{X}^{x,n}_{\phi_t^n})^{-1}\nabla_{x^j}\nabla_{x^i}\bar{X}^{x,n}_{\phi_t^{n,+}}\right) \ .
\end{aligned}$$

Exercice 4.4.16 *Vérifier que dans le modèle de Black-Scholes $\nabla_{x^i}\nabla_{x^j}X^x = 0$ et retrouver ainsi la formule* (4.2.1).

Exercice 4.4.17 *Etendre la formule de la Proposition 4.4.15 à des payoffs de la forme*

$$g\left(\bar{X}^{x,n}_{t_1}, \ldots, \bar{X}^{x,n}_{t_n}\right)$$

sur le modèle de l'Exercice 4.4.14.

4.4.2.3 Véga pour les fonctions du schéma d'Euler (exercice)

On reprend l'exercice de la Section 4.3.3. On se donne une fonction $\tilde{a}$ de $\mathbb{R}^d$ dans $\mathbb{M}^d$, $\varepsilon > 0$ et on note $\bar{X}^{\varepsilon\tilde{a},n}$ la solution de 2.2.3 avec la matrice de volatilité $a + \varepsilon\tilde{a}$ à la place de a. On veut estimer

$$\vartheta(\tilde{a}) \quad := \quad \frac{\partial}{\partial\varepsilon}\mathbb{E}\left[g(\bar{X}_T^{\varepsilon\tilde{a},n})\right]\Big|_{\varepsilon=0} .$$

On suppose que a, b et $\tilde{a} \in C_p^1$.
1/ Donner l'équation satisfaite par $\bar{Z}^{\tilde{a},n} := \frac{\partial}{\partial\varepsilon}\bar{X}^{\varepsilon\tilde{a},n}\big|_{\varepsilon=0}$.
2/ Vérifier que pour tout $t \in (t_{n-1}, T]$

$$D_t\bar{Z}_T^{\tilde{a},n} \quad = \quad \sum_{j=1}^{d} \nabla a^{\cdot j}(\bar{X}_{t_{n-1}}^n)\bar{Z}_{t_{n-1}}^{\tilde{a},n} + \tilde{a}(\bar{X}_{t_{n-1}}^n) .$$

3/ En supposant que a^{-1} et g sont à croissance polynomiale, déduire de 2/, de la fomulation obtenue dans la Section 4.3.3 pour $\vartheta(\tilde{a})$, du Théorème 4.4.9 et de la Propriété 4.4.5 que

$$\vartheta(\tilde{a}) \quad = \quad \mathbb{E}\left[g(\bar{X}_T^n)a(\bar{X}_{t_{n-1}}^n)^{-1}\left(\frac{n}{T}\bar{Z}_T^{\tilde{a},n}\Delta^n W_n - \sum_{j=1}^{d}\nabla a^{\cdot j}(\bar{X}_{t_{n-1}}^n)\bar{Z}_{t_{n-1}}^{\tilde{a},n} - \tilde{a}(\bar{X}_{t_{n-1}}^n)\right)\right] .$$

4/ Etendre cette formulation aux payoffs de la forme $g\left(\bar{X}_{t_1}^n, \ldots, \bar{X}_{t_n}^n\right)$.
5/ On suppose maintenant que X^ε est solution de (2.1.1) pour $d = 1$ et $\sigma + \varepsilon$ à la place de σ. Calculer $\frac{\partial}{\partial\varepsilon}X^\varepsilon\big|_{\varepsilon=0}$. Faire le lien avec D_tX_T et montrer, en utilisant le Théorème 4.4.9 et la Propriété 4.4.5, que

$$\vartheta(1) \quad = \quad \mathbb{E}\left[g(X_T)\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right] .$$

Etendre cette formule au cas $d > 1$. (On peut également obtenir cette formule en considérant le "schéma d'Euler" de $\ln(X)$. Dans ce cas, il faut faire attention car σ intervient dans le coefficient de drift de $\ln(X)$ ce qui modifie la forme de "$\bar{Z}^{1,n}$" par rapport à celle obtenue dans 1/).

Remarque 4.4.18 Cette approche a été initiée par [26]. Les formules que nous démontrons ici correspondent à la discrétisation des formules obtenues en considérant un payoff dépendant de X au lieu de $\bar{X}^n$. L'étude de ces méthodes a été poursuivie dans [25], voir également [20]. Des extensions au cas des options barrière et lookback on été obtenues par [33]. On pourra également consulter [7] qui traite notamment le cas des options asiatiques et [35] qui développe des approches alternatives.

Chapitre 5

Calcul d'espérances conditionnelles

Dans la Section 2.6, on a montré que le prix d'une option américaine peut être approchée par un schéma discret de la forme

$$
\begin{aligned}
\bar{p}^n(t_n, \bar{X}^n_{t_n}) &= g(\bar{X}^n_{t_n}) \\
\bar{p}^n(t_i, \bar{X}^n_{t_i}) &= \max\left\{ g(\bar{X}^n_{t_i}) \ , \ e^{-rT/n}\mathbb{E}\left[\bar{p}^n(t_{i+1}, \bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right]\right\} \ , \ i \in \{0, \ldots, n-1\} \ .
\end{aligned}
$$

Le problème de ce schéma est qu'à chaque date de discrétisation, il nécessite le calcul d'une espérance conditionnelle. Il faut donc trouver un moyen de les estimer de manière efficace. La formule d'intégration par parties du calcul de Malliavin va nous permettre de construire un estimateur facilement implémentable.

5.1 Espérances conditionnelles et densités

Par souci de simplification, on va se limiter au cas $d = 1$, le cas général est obtenu de la même manière (voir [10]).

Théorème 5.1.1 *Soit $\bar{X}^n$ le schéma d'Euler de* (2.0.1). *Soit ψ à croissance polynomiale et φ, $\tilde{\varphi} \in C^1_p(\mathbb{R}_+)$ telles que $\varphi(0) = \tilde{\varphi}(0) = 1$. On suppose que a, $b \in C^1_b$ avec $a \geq \varepsilon > 0$. On a :*

$$
\mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i} = x\right] = \frac{\mathbb{E}\left[\mathbf{1}_{x \leq \bar{X}^n_{t_i}} \psi(\bar{X}^n_{t_{i+1}}) \delta_i\left(\varphi(\bar{X}^n_{t_i} - x)\right)\right]}{\mathbb{E}\left[\mathbf{1}_{x \leq \bar{X}^n_{t_i}} \delta_i\left(\tilde{\varphi}(\bar{X}^n_{t_i} - x)\right)\right]}
$$

où pour une fonction $f \in C^1_p(\mathbb{R}_+)$

$$
\begin{aligned}
\delta_i\left(f(\bar{X}^n_{t_i} - x)\right) &= -f'(\bar{X}^n_{t_i} - x) + \frac{n}{T} f(\bar{X}^n_{t_i} - x)\Big\{ a(\bar{X}^n_{t_{i-1}})^{-1}\Delta^n W_i \\
&\quad - a(\bar{X}^n_{t_i})^{-1}\left[(1 + \frac{T}{n} b'(\bar{X}^n_{t_i}))\Delta^n W_{i+1} + a'(\bar{X}^n_{t_i})(\Delta^n W^2_{i+1} - \frac{T}{n})\right]\Big\} \ .
\end{aligned}
$$

Remarque 5.1.2 On obtient en particulier une formulation de la densité de $\bar{X}^n_{t_i}$ sous la forme

$$\mathbb{E}\left[\mathbf{1}_{x\leq \bar{X}^n_{t_i}}\delta_i\left(\tilde{\varphi}(\bar{X}^n_{t_i}-x)\right)\right] \;=\; \mathbb{E}\left[\mathbf{1}_{x\leq \bar{X}^n_{t_i}}\left(\frac{n}{T}\,\tilde{\varphi}(\bar{X}^n_{t_i}-x)a(\bar{X}^n_{t_{i-1}})^{-1}\Delta^n W_i-\tilde{\varphi}'(\bar{X}^n_{t_i}-x)\right)\right] \;,$$

cette égalité étant obtenue en observant que

$$\mathbb{E}\left[(1+\frac{T}{n}b'(\bar{X}^n_{t_i}))\Delta^n W_{i+1}+a'(\bar{X}^n_{t_i})(\Delta^n W^2_{i+1}-\frac{T}{n})\mid \mathcal{F}_{t_i}\right] \;=\; 0\;.$$

Si on note

$$\tilde{\delta}_i(\tilde{\varphi}(\bar{X}^n_{t_i}-x)) \;:=\; \frac{n}{T}\,\tilde{\varphi}(\bar{X}^n_{t_i}-x)a(\bar{X}^n_{t_{i-1}})^{-1}\Delta^n W_i-\tilde{\varphi}'(\bar{X}^n_{t_i}-x)$$

on obtient donc

$$\mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}})\mid \bar{X}^n_{t_i}=x\right] \;=\; \frac{\mathbb{E}\left[\mathbf{1}_{x\leq \bar{X}^n_{t_i}}\psi(\bar{X}^n_{t_{i+1}})\delta_i\left(\varphi(\bar{X}^n_{t_i}-x)\right)\right]}{\mathbb{E}\left[\mathbf{1}_{x\leq \bar{X}^n_{t_i}}\tilde{\delta}_i\left(\tilde{\varphi}(\bar{X}^n_{t_i}-x)\right)\right]}\;.$$

Remarque 5.1.3 Cette formulation de la densité a été largement utilisée par [54] pour étudier la régularité des densités de variables aléatoires. Une étude sur les méthodes de réduction de variance de l'estimateur ainsi obtenu a été menée par [42] en dimension 1.

Remarque 5.1.4 Les fonctions φ et $\tilde{\varphi}$ agissent comme des fonctions de localisation : on sélectionne les trajectoires qui ne sont pas trop éloignées de x. Typiquement, elles atteignent leur maximum en 0 et décroissent ensuite rapidement. On peut choisir les fonctions φ et $\tilde{\varphi}$ de manière à minimiser la variance intégrée des estimateurs du numérateur et dénominateur. Il est montré dans [10] que les fonctions φ et $\tilde{\varphi}$ optimales sont de type exponentielle $\varphi(y)=e^{-\eta y}$, $\tilde{\varphi}(y)=e^{-\tilde{\eta} y}$, où

$$\begin{aligned}\eta^2 \;&=\; \mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}})^2\delta_i(1)^2\right]/\mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}})^2\right]\\ \tilde{\eta}^2 \;&=\; \mathbb{E}\left[\tilde{\delta}_i(1)^2\right]\;.\end{aligned}$$

On peut vérifier que η et $\tilde{\eta}$ ont une évolution avec n de l'ordre de $\sqrt{n}$.

Remarque 5.1.5 Si X est le log des prix dans le modèle de Black-Scholes, i.e. $b(x)=r-\sigma^2/2$ et $a(x)=\sigma$, on obtient

$$\delta_i\left(f(X_{t_i}-x)\right) \;=\; -f'(X_{t_i}-x)+\frac{n}{T}\,\frac{f(X_{t_i}-x)}{\sigma}\left\{\Delta^n W_i-\Delta^n W_{i+1}\right\}\;.$$

Preuve du Théorème 5.1.1 On commence par fixer $[a,b]$ tel que $x\in[a,b]$ et on écrit que

$$\mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}})\mid \bar{X}^n_{t_i}\in[a,b]\right] \;=\; \frac{\mathbb{E}\left[\mathbf{1}_{[a,b]}(\bar{X}^n_{t_i})\psi(\bar{X}^n_{t_{i+1}})\right]}{\mathbb{E}\left[\mathbf{1}_{[a,b]}(\bar{X}^n_{t_i})\right]}\;. \tag{5.1.1}$$

On se concentre sur le numérateur, le dénominateur étant obtenu de la même manière en remplaçant ψ par 1. On suppose que $\psi \in C_b^1$, le résultat général étant obtenu par passage à la limite. On pose

$$\Phi = \int_{-\infty}^{\bar{X}_{t_i}^n} \mathbf{1}_{[a,b]}(\xi)\psi(\bar{X}_{t_{i+1}}^n)\varphi(\bar{X}_{t_i}^n - \xi)d\xi .$$

Par la Propriété 4.4.5, la Remarque 4.3.5 et le fait que $\varphi(0) = 1$, on a, pour tout $t \in [0, T]$,

$$\begin{aligned} D_t\Phi &= \mathbf{1}_{[a,b]}(\bar{X}_{t_i}^n)\psi(\bar{X}_{t_{i+1}}^n)D_t\bar{X}_{t_i}^n \\ &\quad + \int_{-\infty}^{\bar{X}_{t_i}^n} \mathbf{1}_{[a,b]}(\xi)\left(\psi'(\bar{X}_{t_{i+1}}^n)D_t\bar{X}_{t_{i+1}}^n\varphi(\bar{X}_{t_i}^n - \xi) + \psi(\bar{X}_{t_{i+1}}^n)\varphi'(\bar{X}_{t_i}^n - \xi)D_t\bar{X}_{t_i}^n\right) d\xi .\end{aligned}$$

On pose

$$h_{i,t} := \frac{n}{T}\left(a(\bar{X}_{t_{i-1}}^n)^{-1}\mathbf{1}_{[t_{i-1},t_i)}(t) - a(\bar{X}_{t_i}^n)^{-1}(1 + \frac{T}{n}b'(\bar{X}_{t_i}^n) + a'(\bar{X}_{t_i}^n)\Delta^n W_{i+1})\mathbf{1}_{[t_i,t_{i+1})(t)}\right) .$$

On remarque que, par construction de h_i et (4.4.1),

$$\int_{t_{i-1}}^{t_{i+1}} D_t\bar{X}_{t_{i+1}}^n h_{i,t}dt = 0 \quad \text{et} \quad \int_{t_{i-1}}^{t_{i+1}} D_t\bar{X}_{t_i}^n h_{i,t}dt = 1 \tag{5.1.2}$$

de sorte que

$$\begin{aligned} \int_{t_{i-1}}^{t_{i+1}} D_t\Phi h_{i,t}dt &= \mathbf{1}_{[a,b]}(\bar{X}_{t_{i-1}}^n)\psi(\bar{X}_{t_i}^n) + \int_{-\infty}^{\bar{X}_{t_i}^n} d\xi\mathbf{1}_{[a,b]}(\xi)\psi(\bar{X}_{t_i}^n)\varphi'(\bar{X}_{t_{i-1}}^n - \xi) \\ &= \mathbf{1}_{[a,b]}(\bar{X}_{t_{i-1}}^n)\psi(\bar{X}_{t_i}^n) + \int_{[a,b]} \mathbf{1}_{\xi \le \bar{X}_{t_{i-1}}^n}\psi(\bar{X}_{t_i}^n)\varphi'(\bar{X}_{t_{i-1}}^n - \xi)d\xi \end{aligned} \tag{5.1.3}$$

On calcule maintenant

$$\begin{aligned} \int_{t_{i-1}}^{t_{i+1}} D_t\Phi h_{i,t}dt &= \frac{n}{T}\int_{t_{i-1}}^{t_i} D_t\Phi a(\bar{X}_{t_{i-1}}^n)^{-1}dt \\ &\quad -\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t\Phi a(\bar{X}_{t_i}^n)^{-1}\left(1 + \frac{T}{n}b'(\bar{X}_{t_i}^n) + a'(\bar{X}_{t_i}^n)\Delta^n W_{i+1}\right) dt .\end{aligned}$$

D'après le Théorème 4.4.9 conditionnellement à $\mathcal{F}_{t_{i-1}}$,

$$\mathbb{E}\left[\frac{n}{T}\int_{t_{i-1}}^{t_i} D_t\Phi\, a(\bar{X}_{t_{i-1}}^n)^{-1}dt\right] = \mathbb{E}\left[\frac{n}{T}\,\Phi\, a(\bar{X}_{t_{i-1}}^n)^{-1}\Delta^n W_i\right]$$

et

$$\begin{aligned}
&\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t\Phi a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})+a'(\bar{X}^n_{t_i})\Delta^n W_{i+1}\right)dt\right]\\
&=\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t\Phi a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})\right)dt\right]\\
&+\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t\Phi a(\bar{X}^n_{t_i})^{-1}a'(\bar{X}^n_{t_i})\Delta^n W_{i+1}dt\right]\\
&=\mathbb{E}\left[\frac{n}{T}\,\Phi a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})\right)\Delta^n W_{i+1}\right]+\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t(\Phi\Delta^n W_{i+1})a(\bar{X}^n_{t_i})^{-1}a'(\bar{X}^n_{t_i})dt\right]\\
&-\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} \Phi D_t(\Delta^n W_{i+1})a(\bar{X}^n_{t_i})^{-1}a'(\bar{X}^n_{t_i})dt\right]
\end{aligned}$$

comme $D_t\Delta^n W_{i+1}=1$ sur $(t_i,t_{i+1}]$, on a donc par le Théorème 4.4.9 conditionnellement à $\mathcal{F}_{t_i}$

$$\begin{aligned}
&\mathbb{E}\left[\frac{n}{T}\int_{t_i}^{t_{i+1}} D_t\Phi a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})+a'(\bar{X}^n_{t_i})\Delta^n W_{i+1}\right)dt\right]\\
&=\mathbb{E}\left[\frac{n}{T}\,\Phi a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})\right)\Delta^n W_{i+1}\right]+\mathbb{E}\left[\frac{n}{T}\,\Phi a(\bar{X}^n_{t_i})^{-1}a'(\bar{X}^n_{t_i})(\Delta^n W_{i+1})^2\right]\\
&-\mathbb{E}\left[\Phi a(\bar{X}^n_{t_i})^{-1}a'(\bar{X}^n_{t_i})\right]\ .
\end{aligned}$$

On a montré que

$$\begin{aligned}
\mathbb{E}\left[\int_{t_{i-1}}^{t_{i+1}} D_t\Phi h_{i,t}dt\right] \;=\;& \frac{n}{T}\,\mathbb{E}\left[\Phi\left(a(\bar{X}^n_{t_{i-1}})^{-1}\Delta^n W_i-a(\bar{X}^n_{t_i})^{-1}\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})\right)\Delta^n W_{i+1}\right)\right]\\
&-\mathbb{E}\left[\Phi a(\bar{X}^n_{t_i})^{-1}\left(\frac{n}{T}a'(\bar{X}^n_{t_i})(\Delta^n W_{i+1})^2-a'(\bar{X}^n_{t_i})\right)\right]\ .
\end{aligned}$$

En combinant cette équation avec (5.1.3), on obtient

$$\begin{aligned}
&\frac{n}{T}\,\mathbb{E}\Big\{\Phi\Big[a(\bar{X}^n_{t_{i-1}})^{-1}\Delta^n W_i\\
&-a(\bar{X}^n_{t_i})^{-1}\left(\left(1+\frac{T}{n}b'(\bar{X}^n_{t_i})\right)\Delta^n W_{i+1}+a'(\bar{X}^n_{t_i})(\Delta^n W_{i+1})^2-a'(\bar{X}^n_{t_i})\frac{T}{n}\right)\Big]\Big\}\\
&=\int_{[a,b]}\mathbb{E}\left[\mathbf{1}_{[a,b]}(\bar{X}^n_{t_{i-1}})\psi(\bar{X}^n_{t_i})+\int_{[a,b]}\mathbf{1}_{\xi\le\bar{X}^n_{t_{i-1}}}\psi(\bar{X}^n_{t_i})\varphi'(\bar{X}^n_{t_{i-1}}-\xi)\right]d\xi\ ,
\end{aligned}$$

ce qui implique, par définition de Φ,

$$\mathbb{E}\left[\mathbf{1}_{[a,b]}(\bar{X}^n_{t_i})\psi(\bar{X}^n_{t_{i+1}})\right] \;=\; \int_{[a,b]}\mathbb{E}\left[\mathbf{1}_{\xi\le\bar{X}^n_{t_i}}\psi(\bar{X}^n_{t_{i+1}})\delta_i\left(\varphi(\bar{X}^n_{t_{i-1}}-\xi)\right)\right]d\xi\ .$$

En utilisant (5.1.1), on obtient

$$\mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}})\mid\bar{X}^n_{t_i}\in[a,b]\right] \;=\; \frac{\int_{[a,b]}\mathbb{E}\left[\mathbf{1}_{\xi\le\bar{X}^n_{t_i}}\psi(\bar{X}^n_{t_{i+1}})\delta_i\left(\varphi(\bar{X}^n_{t_i}-\xi)\right)\right]d\xi}{\int_{[a,b]}\mathbb{E}\left[\mathbf{1}_{\xi\le\bar{X}^n_{t_i}}\delta_i\left(\varphi(\bar{X}^n_{t_i}-\xi)\right)\right]d\xi}\ ,$$

et comme la loi de $\bar{X}^n_{t_i}$ n'a pas d'atome, on en déduit le résultat. □

La formulation du Théorème 5.1.1 fournit un estimateur naturel pour l'espérance conditionnelle. On peut ainsi considérer N copies indépendantes $(\bar{X}^{n(j)})_{j=1}^N$ de $\bar{X}^n$ et définir un estimateur de l'espérance conditionnelle par

$$\tilde{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] \;=\; \frac{\sum_{j=1}^N \mathbf{1}_{\bar{X}^n_{t_i} \le X^{n(j)}_{t_i}} \psi(\bar{X}^{n(j)}_{t_{i+1}}) \delta_i^{(j)} \left(\varphi(\bar{X}^{n(j)}_{t_i} - \bar{X}^n_{t_i})\right)}{\mathbf{1}_{\bar{X}^n_{t_i} \le \bar{X}^{n(j)}_{t_i}} \tilde{\delta}_i^{(j)} \left(\varphi(\bar{X}^{n(j)}_{t_i} - \bar{X}^n_{t_i})\right)}$$

où $\delta_i^{(j)}$ et $\tilde{\delta}_i^{(j)}$ sont les opérateurs correspondant à la j-ème trajectoire simulée. En pratique ces estimateurs risquent d'être instables à cause de la division (ils n'ont d'ailleurs aucune raison d'être dans L^1 !). Il faut donc les corriger légèrement. Si ψ est à croissance polynomiale, il est assez facile de construire des polynômes $\underline{\rho}$ et $\bar{\rho}$ tels que

$$\underline{\rho}(\bar{X}^n_{t_i}) \;\le\; \mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] \;\le\; \bar{\rho}(\bar{X}^n_{t_i}) \; .$$

On définit alors l'estimateur de l'espérance conditionnelle par

$$\hat{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] \;:=\; \underline{\rho}(\bar{X}^n_{t_i}) \vee \tilde{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] \wedge \bar{\rho}(\bar{X}^n_{t_i}) \; .$$

Comme, d'après le Lemme 2.2.1, $\bar{X}^n_{t_i}$ est dans tous les L^p, $p \ge 1$, on obtient ainsi un estimateur avec de bonnes propriétés d'intégrabilité. On a le résultat de convergence suivant.

Théorème 5.1.6 *Sous les hypothèses du Théorème 5.1.1, si on prend φ et $\tilde{\varphi}$ de la forme $e^{-\eta_n |x|}$ avec $\eta_n \sim \sqrt{n}$ quand $n \to \infty$, alors, pour tout $p \ge 1$,*

$$\left\| \hat{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] - \mathbb{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right] \right\|_{L^p} \;\le\; C \frac{n^{\frac{1}{4p}}}{N^{\frac{1}{2p}}}$$

où C ne dépend pas de n et N.

La preuve de ce résultat n'est pas très difficile mais un peu longue. Nous renvoyons à [12] pour les détails et pour l'étude du cas général.

Remarque 5.1.7 D'après la Remarque 5.1.2, si on connaît explicitement la densité $f_{\bar{X}^n_{t_i}}$ de $\bar{X}^n_{t_i}$, on peut également utiliser comme estimateur

$$\check{E}\left[\psi(\bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i} = x\right] \;=\; \frac{\frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\bar{X}^n_{t_i} \le X^{n(j)}_{t_i}} \psi(\bar{X}^{n(j)}_{t_{i+1}}) \delta_i^{(j)} \left(\varphi(\bar{X}^{n(j)}_{t_i} - \bar{X}^n_{t_i})\right)}{f_{\bar{X}^n_{t_i}}(x)} \; .$$

Remarque 5.1.8 Lorsque $d \geq 2$, un résultat similaire à celui des Théorèmes 5.1.1 et 5.1.6 peut être obtenu mais l'écriture de l'estimateur est plus compliquée. Dans [10], vous trouverez une écriture en terme d'intégrales de Skorohod itérées. Le cas du modèle de Black-Scholes a été étudié par [48], la formulation est alors plus simple. On peut néanmoins obtenir une forme élégante de ce résultat dans le cas où $\bar{X}^n = X = W$. A un changement de variable près, il est souvent possible de se ramener à ce cas. On a alors sous des conditions analogues à celles du Théorème 5.1.1

$$\mathbb{E}\left[\psi(W_{t_{i+1}}) \mid W_{t_i} = x\right]$$
$$=$$
$$\frac{\mathbb{E}\Big[\psi(W_{t_{i+1}}) \prod_{\ell=1}^d \mathbf{1}_{x^\ell \leq W^\ell_{t_i}} e^{-\eta^\ell (W^\ell_{t_i} - x)} \big\{\frac{n}{T}\big(\Delta^n W^\ell_i - \Delta^n W^\ell_{i+1}\big) + \eta^\ell\big\}\Big]}{\mathbb{E}\Big[\prod_{\ell=1}^d \mathbf{1}_{x^\ell \leq W^\ell_{t_i}} e^{-\tilde{\eta}^\ell (W^\ell_{t_i} - x)} \big\{\frac{n}{T}\Delta^n W^\ell_i + \eta^\ell\big\}\Big]} .$$

La convergence de l'estimateur L^p est alors en $n^{\frac{d}{4p}}/N^{\frac{1}{2p}}$. Dans ce cas particulier, on peut simplifier la formulation de l'estimateur puisque l'on connaît explicitement la densité des W_{t_i}, voir Remarque 5.1.7.

Exemple 5.1.9 *On considère le processus suivant*

$$dX_t = \text{diag}\,[X_t]\,\sigma dW_t \;\;,\;\; X^1_0 = X^2_0 = X^3_0 = 1 \;.$$

avec

$$\sigma = \begin{bmatrix} 0.2 & 0 & 0 \\ 0.08 & 0.4 & 0 \\ 0.03 & -0.15 & 0.32 \end{bmatrix} .$$

On estime la densité du processus et l'espérance conditionnelle

$$r(x) \;=\; 100 \times \mathbb{E}\left[\left(\frac{X^1_2 + X^2_2}{2} - X^3_2\right)^+ \mid X_1 = x\right]$$

en différents points. On donne le résultat moyen obtenu ainsi que l'écart-type de l'estimateur avec ou sans fonction de localisation.

Densité $x^1 = 1.3$

$x^3 \backslash x^2$		0.7	1.0	1.3
0.7	*Valeur exacte*	0.29	0.56	0.48
	Avec localisation	0.30[0.03]	0.56[0.02]	0.48[0.01]
	Sans localisation	0.30[0.41]	0.57[0.43]	0.51[0.44]
1.0	*Valeur exacte*	0.59	0.70	0.43
	Avec localisation	0.59[0.02]	0.70[0.01]	0.45[0.27]
	Sans localisation	0.60[0.47]	0.72[0.48]	0.47[0.49]
1.3	*Valeur exacte*	0.44	0.38	0.18
	Avec localisation	0.44[0.01]	0.38[0.00]	0.18[0.00]
	Sans localisation	0.45[0.48]	0.40[0.49]	0.22[0.51]

Espérance conditionnelle $x^1 = 0.9$

$x^3 \backslash x^2$		0.9	1.0	1.1
0.9	*Valeur exacte*	17.26	20.56	24.05
	Avec localisation	17.28[1.12]	20.49[1.19]	24.06[1.17]
	Sans localisation	16.88[5.01]	20.62[7.14]	25.04[11.52]
1.0	*Valeur exacte*	13.70	16.59	19.61
	Avec localisation	13.72[0.82]	16.59[0.92]	19.73[1.00]
	Sans localisation	12.97[6.20]	16.08[10.41]	21.35[25.48]
1.1	*Valeur exacte*	10.88	13.39	16.11
	Avec localisation	10.94[0.85]	13.48[0.90]	16.32[1.05]
	Sans localisation	10.58[17.54]	13.81[28.01]	13.19[166.22]

On observe une très forte volatilité de l'estimateur en l'absence de fonction de localisation. Il est donc absolument nécessaire de l'utiliser. Même avec celle-ci, la variance reste forte. Cette approche doit donc être combinée avec d'autres techniques de réduction de variance dès que cela est possible.

5.2 Application à l'évaluation d'options américaines

On revient maintenant au problème d'évaluation d'options américaines présenté dans la Section 2.6. On considère le problème discrétisé

$$\begin{aligned}\bar{p}^n(t_n, \bar{X}^n_{t_n}) &= g(\bar{X}^n_{t_n}) \\ \bar{p}^n(t_i, \bar{X}^n_{t_i}) &= \max\left\{g(\bar{X}^n_{t_i}) \ , \ e^{-rT/n}\mathbb{E}\left[\bar{p}^n(t_{i+1}, \bar{X}^n_{t_{i+1}}) \mid \bar{X}^n_{t_i}\right]\right\} \quad , \ i \in \{0, \ldots, n-1\} \ .\end{aligned}$$

Si g est lipschitzienne, il est facile de construire une suite de polynômes $\bar{\rho}_i(x) = \alpha_i + \beta_i \|x\|^2$ tels que les (α_i, β_i) sont uniformément bornés en i et n et tels que

$$\begin{aligned}&\left\|\mathbb{E}\left[g(\bar{X}^n_{t_n}) \mid \bar{X}^n_{t_{n-1}}\right]\right\| \le \bar{\rho}_{n-1}(\bar{X}^n_{t_{n-1}}) \\ \text{et } \forall\, i \quad &\left\|\mathbb{E}\left[g(\bar{X}^n_{t_i}) \vee e^{-rT/n}\bar{\rho}_i(\bar{X}^n_{t_i}) \mid \bar{X}^n_{t_{i-1}}\right]\right\| \le \bar{\rho}_{i-1}(\bar{X}^n_{t_{i-1}}) \ .\end{aligned}$$

On peut donc utiliser l'estimateur du Théorème 5.1.6 pour calculer les espérances conditionnelles. L'argorithme utilisé est le suivant :

On considère nN copies $(\bar{X}^{n^{(1)}}, \ldots, \bar{X}^{n^{(nN)}})$ de $\bar{X}^n =: \bar{X}^{n^{(0)}}$, et on pose $\mathcal{N}_i := \{(i-1)N+1, \ldots, iN\}$.

Initialisation : Pour $j \in \{0\} \cup \mathcal{N}_n$, on pose : $\hat{p}^n(t_n, \bar{X}^{n^{(j)}}_{t_n}) = g\left(\bar{X}^{n^{(j)}}_{t_n}\right)$.

Récurrence rétrograde : Pour $i = n, \ldots, 1$, on pose, pour $j \in \{0\} \cup \mathcal{N}_{i-1}$:

$$\hat{p}^n(t_{i-1}, \bar{X}^{n^{(j)}}_{t_{i-1}}) = \max\left\{g(X^{n^{(j)}}_{t_{i-1}}) \ , \ e^{-rT/n}\hat{E}\left[\hat{p}^n(t_i, \bar{X}^{n^{(j)}}_{t_i}) \mid \bar{X}^{n^{(j)}}_{t_{i-1}}\right]\right\}$$

où

$$\hat{E}\left[\hat{p}^n(t_i,\bar{X}^{n(j)}_{t_i}) \mid \bar{X}^{n(j)}_{t_{i-1}}\right] = -\bar{\rho}_{i-1}(\bar{X}^{n(j)}_{t_{i-1}}) \vee \tilde{E}\left[\hat{p}^n(t_i,\bar{X}^{n(j)}_{t_i}) \mid \bar{X}^{n(j)}_{t_{i-1}}\right] \wedge \bar{\rho}_{i-1}(\bar{X}^{n(j)}_{t_{i-1}})$$

$$\tilde{E}\left[\hat{p}^n(t_i,\bar{X}^{n(j)}_{t_i}) \mid \bar{X}^{n(j)}_{t_{i-1}}\right] := \frac{\sum_{\ell\in\mathcal{N}_i} \mathbf{1}_{\bar{X}^{n(j)}_{t_{i-1}} \le \bar{X}^{n(\ell)}_{t_{i-1}}} \hat{p}^n(t_i,\bar{X}^{n(\ell)}_{t_i})\delta^{(\ell)}_{i-1}\left(\varphi(\bar{X}^{n(\ell)}_{t_{i-1}}-\bar{X}^{n(j)}_{t_{i-1}})\right)}{\sum_{\ell\in\mathcal{N}_i} \mathbf{1}_{\bar{X}^{n(j)}_{t_{i-1}} \le \bar{X}^{n(\ell)}_{t_{i-1}}} \tilde{\delta}^{(\ell)}_{i-1}\left(\tilde{\varphi}(\bar{X}^{n(\ell)}_{t_{i-1}}-\bar{X}^{n(j)}_{t_{i-1}})\right)}$$

pour $i \geq 1$ et pour $i = 1$

$$\hat{E}\left[\hat{p}^n(t_1,\bar{X}^{n(j)}_{t_1})\right] := \frac{1}{N}\sum_{\ell\in\mathcal{N}_1} \hat{p}^n(t_1,\bar{X}^{n(\ell)}_{t_1}) .$$

Comme on fait n approximations de l'espérance conditionnelle, l'erreur est de l'ordre de n fois celle donnée par le Théorème 5.1.6, i.e.

$$\max_{0\le i\le n} \left\|\hat{p}^n(t_i,\bar{X}^n_{t_i}) - \bar{p}^n(t_i,\bar{X}^n_{t_i})\right\|_{L^p} \le Cn\left(n^{1/4p}/N^{1/2p}\right) .$$

Remarque 5.2.1 On pourra bien évidemment, et c'est même fortement recommandé, coupler cet algorithme avec une méthode de réduction de variance de type control variate comme celle proposée dans l'Exemple 3.3.3.

Remarque 5.2.2 Tout ceci se généralise à une large classe d'équations forward-backward. La présentation générale de la méthode et les vitesses de convergence sont dues à [12]. Lorsque le processus $\bar{X}^n$ est de dimension d, la vitesse de convergence devient $n\left(n^{d/4p}/N^{1/2p}\right)$. Cette approche a été initiée par [16] et [49] mais en utilisant des outils différents pour le calcul des espérances conditionnelles, voir aussi [19], [34] et les références de la section 2.6.

Exemple 5.2.3 *On utilise cette approche pour évaluer un put américain sur moyenne dans le modèle de Black-Scholes en dimension 2. On prend* $X^1_0 = X^2_0 = K = 100$, $r = 0.05$, $T = 1$, $\sigma^{11} = 0.15$, $\sigma^{12} = 0$, $\sigma^{21} = -0.05$ *et* $\sigma^{22} = 0.10$. *On s'intéresse à deux types de moyenne, géométrique et arithmétique, i.e. aux deux payoffs :*

$$G^{geo} = \left[K - \sqrt{X^1_T X^2_T}\right]^+ \quad et \quad G^{arith} = \left[K - \frac{X^1_T + X^2_T}{2}\right]^+ .$$

On écrit le prix en fonction du brownien, $p(t_i, W_{t_i})$. *Le prix de l'option européenne géométrique est donné par la formule de Black-Scholes avec dividende valant* 6.25%, *une volatilité de* $7,07\%$ *et un sousjacent correspondant à la moyenne géométrique de* X^1 *et* X^2 *(à vérifier en exercice en considérant la loi de* $\sqrt{X^1_T X^2_T}$*). On appelle* $BS^{geo}(W_{t_i}, T)$ *le prix de l'option européenne sur moyenne géométrique de maturité* T *à la date* t_i *pour une valeur* W_{t_i} *du Brownien. On considère l'estimateur suivant :*

$$\mathbb{E}\left[p(t_{i+1},W_{t_{i+1}}) \mid W_{t_i}=x\right] = \frac{\mathbb{E}\left[\left(p(t_{i+1},W_{t_{i+1}}) - \rho^1 BS^{geo}(W_{t_{i+1}},T) - \rho^2 W_{t_{i+1}}\right) Q_{t_i}(x,W_{t_i})\right]}{f_{W_{t_i}}(x)}$$
$$+\rho^1 e^{r(t_{i+1}-t_i)} BS^{geo}(W_{t_i},T) + \rho^2 W_{t_i}$$

où

$$Q_{t_i}(x, W_{t_i}) \quad := \quad \left\{ H_x(W_{t_i}) \delta_i^1 \left(\varphi(W_{t_i}^1 - x^1) \right) \delta_i^2 \left(\varphi(W_{t_i}^2 - x^2) \right) - c \delta_i^1 (1) \, \delta_i^2 (1) \right\} ,$$

δ^1 et δ^2 correspondent aux δ du premier et du second Brownien, $f_{W_{t_i}}$ est la densité de W_{t_i}, voir Remarque 5.1.7, et $H_x(y) = 1$ si $x^1 \leq y^1$ et $x^2 \leq y^2$, 0 sinon. La fonction de localisation φ est de type exponentielle avec un paramètre estimé selon la formule de la Remarque 5.1.4. Le paramètre ρ est le paramètre de corrélation estimé, voir l'Exemple 3.3.1. En ce qui concerne le c, on utilise la propriété (à vérifier en exercice en utilisant la formule d'intégration par parties)

$$\mathbb{E} \left[\left(p(t_{i+1}, W_{t_{i+1}}) - \rho^1 BS^{geo}(W_{t_{i+1}}, T) - \rho^2 W_{t_{i+1}} \right) \delta_i^1 (1) \, \delta_i^2 (1) \right] \quad = \quad 0$$

et on choisit le c de manière à minimiser la variance de l'estimateur. Le prix de l'option américaine sur moyenne géométrique vaut environ 1.54. *On donne les prix moyens estimés et leur écart-type entre [ ].*

n	N	*Géométrique*
5	2 048	1.48 [0.007]
10	2 048	1.52 [0.013]
20	4 096	1.54 [0.014]

n	N	*Arithmétique*
5	2 048	1.38 [0.017]
10	2 048	1.44 [0.029]
20	2 048	1.50 [0.143]
5	8 192	1.37 [0.007]
10	8 192	1.43 [0.009]
20	8 192	1.46 [0.033]

Chapitre 6

Compléments (exercices)

6.1 Simulation d'un CIR : schéma exact

On considère le processus de Cox Ingersoll Ross (CIR) X^x défini sur $[0,T]$ par

$$X_t^x \;=\; x+\int_0^t (a-bX_s^x)ds+\sigma\int_0^t \sqrt{X_s^x}dW_s$$

où x, a, b, $\sigma>0$ et W est un mouvement brownien sur l'espace $(\Omega,\mathcal{F},\mathbb{P})$ muni de la filtration naturelle $\mathbb{F}=(\mathcal{F}_t)_{t\le T}$ engendrée par W complétée. On suppose que $a\ge\sigma^2/2$ ce qui implique que $X^x>0$ $\mathbb{P}-$p.s.

Par la suite on se donne $\pi:=\{0=t_0<t_1<\ldots<t_i<\ldots<t_n=T\}$ une partition de $[0,T]$.

1. Peut-on approcher X^x par un schéma d'Euler classique ? Justifiez.

2. Soit $\lambda\ge0$. On suppose qu'il existe une fonction $F\in C_b^{1,2}([0,T]\times\mathbb{R}_+;\mathbb{R}_+)$ vérifiant

$$\left\{\begin{array}{rcl} \frac{\partial}{\partial t}F(t,x) &=& (a-bx)\frac{\partial}{\partial x}F(t,x)+\frac12\sigma^2x\frac{\partial^2}{\partial x^2}F(t,x) \quad \text{sur} \;\; [0,T]\times\mathbb{R}_+ \\ F(0,x) &=& e^{-\lambda x} \quad \text{sur} \;\; \mathbb{R}_+\ , \end{array}\right. . \tag{6.1.1}$$

2.1. Montrer que $\sup_{k>0}\sup_{t\le T}\mathbb{E}\left[X_{t\wedge\tau_k}^x\right]<\infty$ où $\tau_k:=\inf\{t\ge0\ :\ X_t^x\ge k\}\wedge T$.

2.2. Soit $t\in(0,T]$, montrer que le processus M défini par $M_s:=F(t-s,X_s^x)$ est une martingale sur $[0,t]$.

2.3. En déduire que $F(t,x)=\mathbb{E}\left[e^{-\lambda X_t^x}\right]$.

3. On suppose maintenant que la solution de (6.1.1) est de la forme $F(t,x)=e^{-a\phi(t)-x\psi(t)}$. Quelles équations doivent vérifier ϕ et ψ ?

4. On admet que $F(t,x)=(2\lambda L(t)+1)^{-2a/\sigma^2}e^{-\frac{\lambda L(t)\zeta(t,x)}{2\lambda L(t)+1}}$ avec $L(t)=(\sigma^2/4b)(1-e^{-bt})$ et $\zeta(t,x)=4xbe^{-bt}/(\sigma^2(1-e^{-bt}))$. Quelle est la transformée de Laplace $\phi_{Y_t^x}$ de $Y_t^x:=X_t^x/L(t)$?

5. On suppose que $4a/\sigma^2=:k\in\mathbb{N}$.

5.1. Soit N une gaussienne de variance 1 et de moyenne m. Calculer la densité f_{N^2} de N^2 et sa transformée de Laplace ϕ_{N^2}.

5.2. En déduire une méthode de simulation des accroissements $(X^x_{t_{i+1}} - X^x_{t_i})_{i<n}$ lorsque $4a/\sigma^2$ est un entier.

6. On considère maintenant le cas où $4a/\sigma^2$ est un réel strictement positif quelconque.

6.1. Soient $\alpha > 0$, $\beta > 0$ et

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbf{1}_{x>0} ,$$

la densité de la loi Gamma $G(\alpha, \beta)$, où Γ est la fonction Gamma. Comment simuler une variable aléatoire (U, V) de loi uniforme sur $D := \{(u, v) \in (0, \infty)^2 \ : \ 0 \leq u \leq \sqrt{f_{\alpha,\beta}(v/u)}\}$ quand $\alpha > 1$? On commencera par montrer que D est contenu dans un rectangle Δ.

6.2. On suppose $\alpha > 1$. Soit (U, V) une variable aléatoire uniformément distribuée sur D. Quelle est la loi de V/U ?

6.3. Déduire des questions précédentes une méthode de simulation de copies i.i.d. de loi $G(\alpha, \beta)$ quand $\alpha > 1$. Donner le coût moyen d'un tirage en fonction du volume $|D|$ de D.

6.4. Que faire quand $\alpha = 1$?

6.5. On rappelle que, pour tout $\alpha > 0$ et $\beta > 0$, la transformée de Laplace de la loi $G(\alpha, \beta)$ est donnée par

$$\phi_{\alpha,\beta}(y) \ = \ (1 + y/\beta)^{-\alpha} \ \ y \geq 0 .$$

Soit $\nu > 0$, M une variables aléatoire de loi de Poisson de paramètre $p > 0$ et $(\chi_{\nu+i})_{i\geq 0}$ une suite de variables aléatoires indépendantes (et indépendantes de M) telle que $\chi_{\nu+i} \sim G((\nu+i)/2, 1/2)$ pour chaque $i \geq 0$. Calculer la transformée de Laplace $\phi_{\nu+M}$ de $\chi_{\nu+M} := \sum_{i\geq 0} \chi_{\nu+i} \mathbf{1}_{i=M}$.

6.6. Comment simuler des copies i.i.d. de $\chi_{\nu+M}$?

6.7. Déduire des questions précédentes un mode de simulation des accroissements $(X^x_{t_{i+1}} - X^x_{t_i})_{i<n}$.

Remarque : En pratique cette méthode est assez coûteuse numériquement. Lorsque le pas de temps est petit, on préfèrera discrétiser l'EDS associée à X^x en adaptant l'approche par schéma d'Euler (voir à ce sujet les travaux récents de [1]).

6.2 Simulation d'un CIR : schéma d'Euler implicite

On s'intéresse à l'équation différentielle stochastique

$$\begin{cases} dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}\, dW_t \\ X_0 = x_0 \end{cases} \tag{6.2.1}$$

où $a, b, \sigma, x > 0$ et $(W_t)_{t\geq 0}$ est un mouvement brownien réel.
Soit $T > 0$ et $N \in \mathbb{N}^*$. On pose $\Delta t = T/N$ et pour $0 \leq k \leq N$, on note $t_k = kT/N = k\Delta t$.
1. Quel résultat assure l'existence d'une unique solution à l'équation

$$\begin{cases} dY_t = -\frac{a}{2}Y_t dt + \frac{\sigma}{2}\, dB_t \\ Y_0^i = \sqrt{x_0} \end{cases}$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard. Que peut-on dire du processus $W_t = \int_0^t \left(1_{\{Y_s\geq 0\}} - 1_{\{Y_s<0\}}\right) dB_s$? Vérifier que $X_t = (Y_t)^2$ est solution de l'équation (6.2.1) pour $b = \frac{\sigma^2}{4a}$.
2. On se donne maintenant $(W_t^1, \ldots, W_t^d)_{t\geq 0}$ un mouvement brownien à valeurs $\mathbb{R}^d$ et pour $1 \leq i \leq d$ on note Y_t^i la solution de

$$\begin{cases} dY_t^i = -\frac{a}{2}Y_t^i dt + \frac{\sigma}{2}\, dW_t^i \\ Y_0^i = 1_{\{i=1\}}\sqrt{x_0}. \end{cases}$$

3. On pose $X_t = \sum_{i=1}^d (Y_t^i)^2$. Que peut-on dire du processus

$$W_t = \int_0^t 1_{\{X_s>0\}} \frac{1}{\sqrt{X_s}} \sum_{i=1}^d Y_s^i dW_s^i + \int_0^t 1_{\{X_s=0\}} dW_s^1?$$

En déduire que X_t est solution de l'équation (6.2.1) pour $b = \frac{d\sigma^2}{4a}$.

A partir de maintenant, on admet que l'équation (6.2.1) **possède une unique solution** $(X_t)_{t\geq 0}$ **telle que** $\mathbb{P}(\forall t \geq 0,\ X_t \geq 0) = 1$.

4. On pose $\alpha \overset{\text{déf}}{=} \frac{2ab}{\sigma^2} > 1$ et

$$\forall x > 0,\ s(x) = \int_1^x y^{-\alpha} e^{\alpha y/b} dy.$$

Pour $k, n \in \mathbb{N}^*$ tels que $1/n \leq x_0 \leq k$, on pose

$$\tau_n^k = \inf\{t \geq 0,\ X_t \notin]1/n, k[\}\ ,\ \tau^k = \inf\{t \geq 0,\ X_t \geq k\}\ .$$

On admet que $\mathbb{P}(\tau_n^k < +\infty) = 1$.
5. Vérifier que $\forall x > 0,\ s''(x) = -\alpha\frac{(b - x)}{bx}s'(x)$.

6. Appliquer Itô à $s(X)$. En déduire que $\mathbb{E}\left[s\left(X_{\tau_n^k}\right)\right] = s(x_0)$.
7. On suppose que $\alpha > 1$. Donner $\lim_{n\to+\infty} s(1/n)$.
En déduire $\lim_{n\to+\infty} \mathbb{P}\left(X_{\tau_n^k} = 1/n\right)$ puis que $\mathbb{P}(\forall t \leq \tau_k,\ X_t > 0) = 1$ en utilisant le lemme de Borel Cantelli (on utilisera une minoration adéquate de $s(1/n)$).
Conclure que $\mathbb{P}(\forall t \geq 0,\ X_t > 0) = 1$.

On se place désormais dans le cas $\alpha > 1$.

8. Calculer $< \sqrt{X}, W >_t$.
9. En remarquant que pour $0 \leq k \leq N-1$, $\sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)$ est égal à

$$\left(\sqrt{X_{t_{k+1}}} - \sqrt{X_{t_k}}\right)\left(W_{t_{k+1}} - W_{t_k}\right) + \sqrt{X_{t_k}}\left(W_{t_{k+1}} - W_{t_k}\right),$$

donner la limite en probabilité de $\sum_{k=0}^{N-1} \sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)$ lorsque $N \to +\infty$.
10. Conclure que

$$x_0 + \sum_{k=0}^{N-1}\left[\left(a(b - X_{t_{k+1}}) - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)\right]$$

converge vers X_T lorsque $N \to +\infty$.

Ce résultat suggère l'utilisation du schéma implicite suivant

$$\begin{cases} \tilde{X}_0 = x_0 \ \text{ et pour } 0 \leq k \leq N-1, \\ \tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + \left(a(b - \tilde{X}_{t_{k+1}}) - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\tilde{X}_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right) \end{cases} \tag{6.2.2}$$

pour l'équation (6.2.1). En raison du caractère implicite, il faut vérifier que l'équation de récurrence qui précède a bien une solution.

11. Vérifier que pour $x > 0$ et $w \in \mathbb{R}$, l'équation

$$(1 + a\Delta t)y^2 - wy - (ab(\alpha - 1) + x) = 0$$

portant sur la variable y admet une unique racine strictement positive $f(x, w)$.
12. En déduire l'existence d'une suite de variables strictement positives $(\tilde{X}_{t_k})_{0\leq k\leq N}$ vérifiant (6.2.2).

6.3 Discrétisation d'EDSR

On considère un marché financier formé d'un actif risqué S de dynamique

$$S_t = S_0 + \int_0^t \sigma(S_u)dW_u \ ,\ t \leq T\ ,$$

où σ est bornée, uniformément lipschitzienne et vérifie $\sigma \geq \varepsilon$ pour un $\varepsilon > 0$. On suppose que le taux sans risque est nul.

1. On note $\mathcal{H}^2$ l'ensemble des processus prévisibles ξ vérifiant

$$\|\xi\|_{\mathcal{H}^2} := \mathbb{E}\left[\int_0^T |\xi_t|^2 dt\right]^{\frac{1}{2}} \quad < \quad \infty \, .$$

On note ϕ le processus de $\mathcal{H}^2$ correspondant au nombre d'unités de S détenues dans le portefeuille.
1.1. Ecrire la dynamique de la richesse $X^{x,\phi}$ associée à une dotation initiale x et une stratégie de portefeuille ϕ sous la contrainte d'autofinancement.

1.2. On suppose à partir de maintenant que la gestion du portefeuille entraine un coût continu f, fonction uniformément lipschitzienne de constante de Lipschitz L, qui dépend de la valeur du portefeuille et de celle du sous-jacent S. La dynamique de la richesse s'écrit alors

$$X_t^{x,\phi} \quad = \quad x - \int_0^t f(X_u^{x,\phi}, S_u)du + \int_0^t \phi_u dS_u \, .$$

Montrer que pour tout $(x, \phi) \in \mathbb{R} \times \mathcal{H}^2$, $X^{x,\phi}$ appartient à l'ensemble $\mathcal{S}^2$ des processus prévisibles χ vérifiant

$$\|\chi\|_{\mathcal{S}^2} := \mathbb{E}\left[\sup_{t\leq T} |\chi_t|^2\right]^{\frac{1}{2}} \quad < \quad \infty \, .$$

2. On veut couvrir l'option $g(S_T)$, i.e. trouver $(x, \phi) \in \mathbb{R} \times \mathcal{H}^2$ tels que

$$X_T^{x,\phi} \quad = \quad g(S_T) \, . \tag{6.3.1}$$

2.1. Montrer que cela revient à trouver (Y_0, Z) où $(Y, Z) \in \mathcal{S}^2 \times \mathcal{H}^2$ est solution de

$$\begin{cases} Y_t &= Y_0 - \int_0^t f(Y_u, S_u)du + \int_0^t Z_u dW_u \;\; \forall \, t \leq T \\ Y_T &= g(S_T) \, . \end{cases} \tag{6.3.2}$$

On supposera par la suite qu'un tel couple (Y, Z) existe.

2.2. Montrer que pour tout $0 \leq s \leq t \leq T$

$$Y_s \quad = \quad \mathbb{E}\left[Y_t + \int_s^t f(Y_u, S_u)du \mid \mathcal{F}_s\right] \, . \tag{6.3.3}$$

3. Dans la suite de l'exercice, on étudie une méthode de discrétisation de l'équation (6.3.2). On fixe un entier $n > 0$ et on considère la grille $\pi := \{t_i = ih, \; i = 0, \ldots, n\}$ de $[0, T]$, où $h = T/n$. A partir de maintenant, on suppose que g est uniformément lipschitzienne, de constante de Lipschitz L.

3.1. Rappeler la définition du schéma d'Euler $\bar{S}$ de S.

3.2. On définit la suite $(\bar{Y}_{t_i})_{0\leq i\leq n}$ de manière rétrograde par

$$\left\{\begin{array}{rcl} \bar{Y}_{t_n} & = & g(\bar{S}_{t_n}) \\ \bar{Y}_{t_i} & = & \mathbb{E}\left[\bar{Y}_{t_{i+1}} + hf(\bar{Y}_{t_{i+1}}, \bar{S}_{t_{i+1}}) \mid \mathcal{F}_{t_i}\right] \quad , \quad i = n-1, \ldots, 0 \, . \end{array}\right. \tag{6.3.4}$$

En utilisant une récurrence, montrer que $\bar{Y}_{t_i} \in L^2$ pour tout $i \leq n$.

3.3. Montrer que pour tout a, $b \in \mathbb{R}$ et $\alpha > 0$: $(a+b)^2 \leq (1+\alpha)a^2 + (1+4/\alpha)b^2$.

3.4. Déduire de (6.3.3) et (6.3.4) que, pour tout $i \in \{0, \ldots, n-1\}$ et $\alpha > 0$,

$$\begin{array}{rcl} |\bar{Y}_{t_i} - Y_{t_i}|^2 & \leq & (1+\alpha+3L^2(1+4/\alpha)h^2)\mathbb{E}\left[|\bar{Y}_{t_{i+1}} - Y_{t_{i+1}}|^2 \mid \mathcal{F}_{t_i}\right] \\ & + & 3L^2(1+4/\alpha)\mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} |\bar{S}_{t_{i+1}} - S_s| + |Y_{t_{i+1}} - Y_s| ds\right)^2 \mid \mathcal{F}_{t_i}\right] \, . \end{array}$$

3.5. On suppose à partir de maintenant que $Z \in \mathcal{S}^2$. Montrer que, pour tout $i \in \{0, \ldots, n-1\}$,

$$\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]} |Y_{t_{i+1}} - Y_s|^2\right] \leq C_0(t_{i+1} - t_i) \, ,$$

où C_0 est une constante qui ne dépend ni de i ni de π (Rem : on peut en fait vérifier cette propriété même si $Z \notin \mathcal{S}^2$).

3.6. En utilisant l'inégalité de Jensen, déduire des questions précédentes que, pour tout $i \in \{0, \ldots, n-1\}$,

$$|\bar{Y}_{t_i} - Y_{t_i}|^2 \leq (1+\alpha+3L^2(1+4/\alpha)h^2)\mathbb{E}\left[|\bar{Y}_{t_{i+1}} - Y_{t_{i+1}}|^2 \mid \mathcal{F}_{t_i}\right] + C\,(1+1/\alpha)\,h^3 \, ,$$

où $C > 0$ est une constante indépendante de i et π.

3.7. En choisissant $\alpha > 0$ correctement, en déduire que

$$\max_{0\leq i\leq n} \mathbb{E}\left[|\bar{Y}_{t_i} - Y_{t_i}|^2\right] \leq C_1\left(\mathbb{E}\left[|\bar{Y}_{t_n} - Y_{t_n}|^2\right] + h\right) \leq C_2\,h \, ,$$

où C_1 et C_2 sont deux constantes indépendantes de h.

3.8. Que peut-on dire de $\mathrm{Err}_n := \max_{0\leq i\leq n} \mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]} |\bar{Y}_{t_{i+1}} - Y_s|^2\right]^{\frac{1}{2}}$? Justifier.

4.1. Montrer qu'il existe un processus $\tilde{Z} \in \mathcal{H}^2$ tel que

$$\bar{Y}_{t_i} - h\mathbb{E}\left[f(\bar{Y}_{t_{i+1}}, \bar{S}_{t_{i+1}}) \mid \mathcal{F}_{t_i}\right] = \bar{Y}_{t_{i+1}} - \int_{t_i}^{t_{i+1}} \tilde{Z}_s dW_s \, .$$

4.2. Montrer que si $f \equiv 0$, alors $\|Z - \tilde{Z}\|^2_{\mathcal{H}^2} \leq Ch$ où C ne dépend pas de h (Rem : on peut le montrer également quand $f \not\equiv 0$).

4.3. Dans le cas générale, i.e. $f \not\equiv 0$, calculez $\mathbb{E}\left[\bar{Y}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i}) \mid \mathcal{F}_{t_i}\right]$. Que proposez-vous pour estimer Z et ϕ?
4.4. Soit $\hat{Z}$ défini sur chaque intervalle $[t_i, t_{i+1})$ par $\hat{Z}_t = (t_{i+1} - t_i)^{-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z_s ds \mid \mathcal{F}_{t_i}\right]$. En admettant qu'il existe C_3 indépendant de h tel que $\|Z - \tilde{Z}\|^2_{\mathcal{H}^2} + \|Z - \hat{Z}\|^2_{\mathcal{H}^2} \leq C_3 h$, montrer que $\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2\right] \leq C_4 h$, où C_4 est indépendant de h.

Remarque : Pour simuler les trajectoires $(\bar{Y}_{t_i})_{i\geq 0}$, on simule $\bar{S}$ et on utilise des estimations de type Monte-Carlo pour approcher les espérances conditionnelles (voir le Chapitre 5).

6.4 Algorithme de Robbins Monro

Soit une fonction continue, bornée, strictement croissante $f : \mathbb{R} \mapsto \mathbb{R}$. On suppose qu'il existe $x^* \in \mathbb{R}$ tel que $f(x^*) = 0$.

Etant donnée une suite $(U_n)_{n\geq 1}$ de variables aléatoires indépendantes de même loi uniforme sur $[-1, 1]$, on considère la suite $(X_n, Y_n)_{n\geq 0}$ définie par $X_0 = 0$ et

$$\begin{aligned} Y_{n+1} &= f(X_n) + U_{n+1} \\ X_{n+1} &= X_n - \frac{1}{n+1} Y_{n+1} \ , \quad n \geq 0 \ . \end{aligned}$$

On note $\mathcal{F}_n = \sigma(X_i \ , \ 0 \leq i \leq n)$.
1. Ecrire $\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n]$ en fonction de X_n. Calculer $|X_{n+1} - x^*|^2 - |X_n - x^*|^2$ et en déduire, en utilisant les hypothèses faites sur f, que, pour tout $n \geq 0$,

$$\mathbb{E}\left[|X_{n+1} - x^*|^2 \mid \mathcal{F}_n\right] \ \leq \ |X_n - x^*|^2 + \frac{1}{(n+1)^2}\mathbb{E}\left[|Y_{n+1}|^2 \mid \mathcal{F}_n\right] \ .$$

2. Montrer qu'il existe une constante $C > 0$ telle que, pour tout $n \geq 1$,

$$S_n := \sum_{k=0}^{n-1} \frac{1}{(k+1)^2}\mathbb{E}\left[|Y_{k+1}|^2 \mid \mathcal{F}_k\right] \ \leq \ C \quad \mathbb{P} - p.s.$$

En déduire que $(S_n)_{n\geq 0}$ converge presque sûrement.
3. Déduire de 1. que le processus $(Z_n)_{n\geq 1}$ défini par

$$Z_n \ := \ |X_n - x^*|^2 + C - S_n$$

est une $(\mathcal{F}_n)_{n\geq 0}$-sur-martingale.
4. En utilisant le fait que pour une sur-martingale positive $(W_n)_{n\geq 1}$, il existe une variable aléatoire W_∞ telle que $\lim_{n\to\infty} W_n = W_\infty$ $\mathbb{P}$-p.s., déduire des questions précédentes que la suite $(|X_n - x^*|)_{n\geq 0}$ admet une limite presque sûre.
5. Montrer que pour tout $n \geq 1$

$$0 \leq \sum_{k=0}^{n} \frac{2}{k+1}\mathbb{E}\left[f(X_k)(X_k - x^*)\right] \leq |x^*|^2 + C \ .$$

6. Déduire de 5. que

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \frac{2}{k+1} f(X_k)(X_k - x^*)\right] \quad \leq \quad |x^*|^2 + C \; .$$

7. En utilisant les hypothèses faites sur f, montrer que, pour tout $\delta > 0$,

$$\inf_{|x-x^*|>\delta} (x - x^*) f(x) \quad > \quad 0 \; .$$

8. Déduire de 6. et 7. que, pour $\delta \in \mathbb{Q} \cap \,]0, \infty[$, $\mathbb{P}[A_\delta] = 0$ où

$$A_\delta := \left\{ \lim_{n\to\infty} |X_n - x^*| \geq \delta \right\} \; .$$

Que peut-on dire de $\lim_{n\to\infty} X_n$? Justifiez.
9. On reprend le modèle précédent mais maintenant on définit $(X_n)_{n\geq 0}$ par $X_0 = 0$ et

$$X_{n+1} \quad = \quad X_n - \gamma_n Y_{n+1} \quad , \quad n \geq 0 \; ,$$

où $(\gamma_n)_{n\geq 0}$ est une suite réelle positive. Quelles conditions doit-on imposer sur la suite $(\gamma_n)_{n\geq 0}$ pour que X_n converge vers x^* presque surement ? Justifiez brièvement.
10. Proposer une méthode basée sur des simulations permettant de résoudre

$$g(x) \quad = \quad \alpha \; ,$$

où g est une fonction continue, bornée, strictement croissante et $\alpha \in g(\mathbb{R})$.

6.5 Méthode de rejet avec recyclage

On veut estimer par Monte-Carlo $\mathbb{E}[g(X)]$ où X est une variable aléatoire de densité f_X strictement positive sur $\mathbb{R}$ et g une fonction réelle bornée. On considère un couple (U, Z) de densité

$$f_{(U,Z)}(u, z) = \mathbf{1}_{\{u\in[0,1]\}} f_Z(z)$$

où f_Z est une densité strictement positive sur $\mathbb{R}$. On suppose qu'il existe $a > 0$ tel que $f_X(z) < a f_Z(z)$ pour tout $z \in \mathbb{R}$.
On considère maintenant une suite $(U_k, Z_k)_{k\geq 1}$ de variables aléatoires indépendantes de même loi que (U, Z). On pose $\mu_0 = \nu_0 = 0$ et on définit de manière récursive

$$\begin{aligned} \nu_{i+1} \quad &:= \quad \inf\{k > \nu_i \; : \; f_X(Z_k) > a U_k f_Z(Z_k)\} \\ \mu_{i+1} \quad &:= \quad \inf\{k > \mu_i \; : \; f_X(Z_k) \leq a U_k f_Z(Z_k)\} \; . \end{aligned}$$

1- Donner (sans justification) la loi de $(Z_{\nu_1}, \ldots, Z_{\nu_n})$?
2- Décrire une méthode de rejet permettant d'estimer $\mathbb{E}[g(X)]$.

3- Calculer $\alpha := \mathbb{P}\left[(U_1, Z_1) \in D\right]$ où

$$D := \{(u,z) \in [0,1] \times \mathbb{R} \ : \ f_X(z) > au f_Z(z)\} \ .$$

4- En déduire $\mathbb{E}\left[\nu_1\right]$ en fonction de a (on commencera par calculer la loi de ν_1).
5- En admettant que

$$\mathbb{E}\left[\nu_{i+1} - \nu_i\right] = \mathbb{E}\left[\nu_1\right] \quad i \geq 1$$

calculer $\mathbb{E}\left[\nu_n\right]$. Quelle interprétation donner à ce nombre ?
6- Calculer la loi de ν_n, i.e. $\mathbb{P}[\nu_n = k]$ pour $k \geq 1$.
7- Montrer que pour toute suite $(A_i)_{i\geq 1}$ de Boréliens de $\mathbb{R}$

$$\mathbb{P}\left[Z_{\nu_i} \in A_i, \ i = 1, \ldots, n \mid \nu_n = n+p\right] = \prod_{i=1}^{n} \int_{A_i} f_X(z) dz$$

et que

$$\mathbb{P}\left[Z_{\mu_i} \in A_i, \ i = 1, \ldots, p \mid \nu_n = n+p\right] = (a-1)^{-p} \prod_{i=1}^{p} \int_{A_i} (af_Z(z) - f_X(z)) dz \ .$$

8- En déduire

$$\mathbb{E}\left[\frac{(a-1) f_X(Z_{\mu_i})}{(af_Z - f_X)(Z_{\mu_i})} g(Z_{\mu_i}) \mid \nu_n = n+p\right] \quad , \quad 1 \leq i \leq p \ ,$$

et

$$\mathbb{E}\left[\frac{1}{\nu_n} \sum_{i=1}^{n} g(Z_{\nu_i}) \mid \nu_n = n+p\right] \ .$$

9- En utilisant les résultats de la question 8., proposer une méthode d'estimation de $\mathbb{E}\left[g(X)\right]$ tenant compte de tous les éléments de la suite $(Z_i)_{i=1}^{\nu_n}$. Justifier.

6.6 Estimation d'espérances conditionnelles

Soit U une variable aléatoire uniforme sur $[0,1]$ et X une variable aléatoire de fonction de répartition Φ continue et strictement croissante sur $\mathbb{R}$.
1. Soit I un intervalle de $[0,1]$ et $(U_n)_{n\geq 1}$ une suite i.i.d.[1] de même loi que U. On définit la suite $(T_n)_{n\geq 1}$ par

$$T_{n+1} \quad := \quad \inf\{k > T_n \ : \ U_k \in I\} \ , \ \ n \geq 1 \ ,$$

avec comme convention $T_0 = 0$. Soient $m < M$ tels que $P[X \in [m, M]] > 0$. Déterminer I tel que $(\Phi^{-1}(U_{T_n}))_{n\geq 1}$ soit une suite i.i.d de loi égale à celle de X sachant $X \in \Delta := [m, M]$ (on admettra que la suite $(U_{T_n})_{n\geq 1}$ est i.i.d.).

[1] i.e. dont les éléments sont indépendants et de même loi.

2. On pose $\delta_n = T_n - T_{n-1}$ pour $n \geq 1$. Montrer que les δ_n sont i.i.d. de loi géométrique de paramètre $\alpha > 0$ à préciser.
3. Soit I défini comme en 1. Montrer que, pour toute fonction bornée f,

$$\hat{m}_N := \frac{\sum_{n=1}^N f(\Phi^{-1}(U_n))\mathbf{1}_I(U_n)}{\sum_{n=1}^N \mathbf{1}_I(U_n)}\mathbf{1}_{\delta_1 \leq N} \quad \to \quad \mathbb{E}[f(X) \mid X \in \Delta] \ \ \mathbb{P}-p.s.$$

(avec la convention $0/0 = 0$).
4. On veut estimer la vitesse de convergence L^2 de l'estimateur $\hat{m}_N$. On suppose que $|f| \leq f_m$ où f_m est une constante réelle. On note

$$A_N := \frac{1}{N}\sum_{n=1}^N f(\Phi^{-1}(U_n))\mathbf{1}_I(U_n)\ ,\ B_N := \frac{1}{N}\sum_{n=1}^N \mathbf{1}_I(U_n)$$
$$A := \mathbb{E}[f(X)\mathbf{1}_{X\in\Delta}]\ ,\ B := \mathbb{P}[X \in \Delta] \ \text{ et } \ r = \frac{A}{B}\ .$$

4.1. Montrer que

$$\mathbb{E}[|\hat{m}_N - r|^2]^{\frac{1}{2}} \quad \leq \quad \mathbb{E}\left[\left|\frac{A_N}{B_N} - \frac{A}{B}\right|^2 \mathbf{1}_{\delta_1 \leq N}\right]^{\frac{1}{2}} + r\mathbb{E}[\mathbf{1}_{\delta_1 > N}]^{\frac{1}{2}}\ .$$

4.2. Montrer que

$$|\hat{m}_N - r|\mathbf{1}_{\delta_1 \leq N} \quad = \quad \left|\frac{A_N - A}{B_N} + r\frac{B - B_N}{B_N}\right|\mathbf{1}_{\delta_1 \leq N}\ ,$$

et en déduire que

$$|\hat{m}_N - r|^2\mathbf{1}_{\delta_1 \leq N} \quad \leq \quad \frac{4}{B^2}\left|A_N - A + r(B - B_N)\right|^2 + 4f_m^2\mathbf{1}_{\tilde{\Omega}^c}\ ,$$

où $\tilde{\Omega} := \{|B_N - B| \leq 2^{-1}B\}$.
4.3. Déduire de la question précédente que

$$\mathbb{E}[|\hat{m}_N - r|^2\mathbf{1}_{\delta_1 \leq N}]^{\frac{1}{2}} \quad \leq \quad \frac{C}{\sqrt{N}}\ ,$$

où $C > 0$ est une constante indépendante de N.
4.4. Déduire que

$$\mathbb{E}[|\hat{m}_N - r|^2]^{\frac{1}{2}} \quad \leq \quad C\left(N^{-\frac{1}{2}} + (1-\alpha)^{N/2}\right)\ ,$$

où $C > 0$ est une constante indépendante de N. Conclure sur la vitesse de convergence L^2 de l'estimateur $\hat{m}_N$.
5. On se propose maintenant d'utiliser une autre méthode.
5.1. Calculer la loi de la suite $(Y_n)_{n\geq 1}$ définie par

$$Y_n := \Phi^{-1}\left(\Phi(m) + U_n(\Phi(M) - \Phi(m))\right)\ \ ,\ n \geq 1\ .$$

5.2. Construire à partir de $(Y_n)_{n\geq 1}$ un estimateur de $\mathbb{E}[f(X) \mid X \in \Delta]$. Quel est son biais ? Quelle est sa vitesse de convergence L^2 ? Comparer avec la méthode précédente.
6. On suppose maintenant que f est continue. Soit $(\xi_n)_{n\geq 1}$ une suite (déterministe) équi-répartie sur $[0,1]$, i.e. telle que

$$\frac{1}{N}\sum_{n=1}^{N} g(\xi_n) \quad \to \quad \int_{[0,1]} g(x)dx \quad \forall\, g \text{ continue bornée de } [0,1] \text{ dans } \mathbb{R}\ .$$

En utilisant l'approche des points 3. et 5. proposer deux modes d'estimation de $\mathbb{E}[f(X) \mid X \in \Delta]$ à partir de $(\xi_n)_{n\geq 1}$. Justifier la convergence des deux méthodes.

6.7 Variables antithétiques

On se donne $(G_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées suivant la loi gaussienne centrée réduite.
On dit qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est monotone (resp. croissante) en chacune de ses variables si pour tout $k \in \{1, \dots, n\}$, l'application qui à $(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n, x, y) \in \mathbb{R}^{n+1}$ associe

$$(y - x)\left(f(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_n) - f(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_n)\right)$$

est de signe constant (resp. positive) sur $\mathbb{R}^{n+1}$.

1. Quelle est la loi de $-G_i$?
2. Soit $f, g : \mathbb{R} \to \mathbb{R}$ croissantes et bornées.
   (a) Quel est le signe de $(f(G_1) - f(G_2))(g(-G_1) - g(-G_2))$?
   En déduire que $\mathrm{Cov}\left((, f)\,(G_1), g(-G_1)\right) \leq 0$.
   (b) Pour l'estimation de $\mathbb{E}\left[(\right]f(G_1))$, quelle variable faut-il préférer entre $\frac{1}{2I}\sum_{i=1}^{2I} f(G_i)$ et $\frac{1}{2I}\sum_{i=1}^{I}(f(G_i)+f(-G_i))$? (comparer à la fois la précision et l'effort de calcul).
3. Nous allons voir que la technique de réduction de variance précédente s'étend à la dimension $n \geq 2$ et même à la dimension infinie.
   On admet que pour ϕ et $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ croissantes en chacune de leurs variables et bornées, on a

   $$\mathrm{Cov}\left((, \phi)\,(G_1, \dots, G_{n-1}), \psi(-G_1, \dots, -G_{n-1})\right) \leq 0\ .$$

   On se donne f et $g : \mathbb{R}^n \to \mathbb{R}$ croissantes en chacune de leurs variables et bornées.
   (a) Pour $x \in \mathbb{R}$, quel est le signe de

   $$\mathbb{E}\left[(\right]f(G_1, \dots, G_{n-1}, x)g(-G_1, \dots, -G_{n-1}, -x)) - \Lambda(x)\Gamma(-x) \qquad (6.7.1)$$

   où $\Lambda(x) = \mathbb{E}\left[(\right]f(G_1, \dots, G_{n-1}, x))$ et $\Gamma(x) = \mathbb{E}\left[(\right]g(-G_1, \dots, -G_{n-1}, x))$?

(b) Que peut-on dire des fonctions $\Lambda(x)$ et $\Gamma(x)$? En déduire le signe de

$$\mathbb{E}\left[(\right]\Lambda(G_n)\Gamma(-G_n))-\mathbb{E}\left[(\right]f(G_1,\ldots,G_{n-1},G_n))\mathbb{E}\left[(\right]g(-G_1,\ldots,-G_{n-1},-G_n)).$$

(c) En intégrant l'inégalité (6.7.1) contre une densité bien choisie, vérifier que

$$\mathrm{Cov}\left((,f\right)(G_1,\ldots,G_n),g(-G_1,\ldots,-G_n))\leq 0.$$

(d) Conclure que pour tout $n\in\mathbb{N}^*$, pour toute fonction $f:\mathbb{R}^n\to\mathbb{R}$ monotone en chacune de ses variables et bornée,

$$\mathrm{Cov}\left((,f\right)(G_1,\ldots,G_n),f(-G_1,\ldots,-G_n))\leq 0.$$

(Remarquer que si $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)\in\{-1,1\}^n$, le vecteur $(\varepsilon_1G_1,\ldots,\varepsilon_nG_n)$ a même loi que $(G_1,\ldots,G_n)$.)
Qu'indique ce résultat pour l'estimation de $\mathbb{E}\left[(\right]f(G_1,\ldots,G_n))$ par la méthode de Monte-Carlo ?

4. On se place maintenant dans le modèle de Black-Scholes

$$S_t(W)=S_0e^{\sigma W_t+(r-\sigma^2/2)t}$$

où $S_0\in\mathbb{R}_+^*$ et $W=(W_t)_{t\geq 0}$ est un mouvement brownien standard. On fixe une maturité $T>0$ et $f:\mathbb{R}\to\mathbb{R}$ une fonction de payoff monotone, continue et bornée. Pour $N\in\mathbb{N}^*$, on pose $\Delta t=T/N$ et $t_k=kT/N=k\Delta t,\ k\in\{0,\ldots,N\}$.

(a) Quel est le signe de $\mathrm{Cov}\left((,f\right)(S_T(W)),f(S_T(-W)))$?

(b) On note $M_T^N(W)=\frac{1}{N}\sum_{k=0}^{N-1}S_{t_k}(W)$.

i. Donner, en la justifiant, la limite presque sûre de $M_T^N(W)$ lorsque N tend vers $+\infty$.

ii. Montrer que

$$M_T^N(W)=\frac{S_0}{N}\ \varphi_{N-1}\left(\frac{W_{t_1}}{\sqrt{\Delta t}},\frac{W_{t_2}-W_{t_1}}{\sqrt{\Delta t}},\ldots,\frac{W_{t_{N-1}}-W_{t_{N-2}}}{\sqrt{\Delta t}}\right),$$

où pour $n\in\mathbb{N}^*$,

$$\varphi_n:\ (z_1,\ldots,z_n)\in\mathbb{R}^n\to\sum_{k=0}^{n}e^{\sigma\sqrt{\Delta t}\sum_{j=1}^{k}z_j+k(r-\sigma^2/2)\Delta t}.$$

iii. En déduire le signe de $\mathrm{Cov}\left(f(M_T^N(W)),f(M_T^N(-W))\right)$.

iv. Conclure que $\mathrm{Cov}\left(f\left(\frac{1}{T}\int_0^T S_t(W)dt\right),f\left(\frac{1}{T}\int_0^T S_t(-W)dt\right)\right)\leq 0.$

(c) Montrer de même que $\mathrm{Cov}\left(f\left(\max_{t\in[0,T]}S_t(W)\right),f\left(\max_{t\in[0,T]}S_t(-W)\right)\right)\leq 0.$

6.8 Suites uniformes randomisées

Pour une fonction réelle f sur $[0,1]$, on note $\|f\|_p = (\int_0^1 |f(x)|^p dx)^{1/p}$ pour $p > 0$ et $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$.

On considère une suite $(\xi_n)_{n\geq 1}$ à valeurs dans $[0,1]$ et on lui associe la fonction

$$F_N(x) \quad := \quad \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{\xi_n \leq x} \quad , \quad x \in [0,1] \; .$$

1. On définit la discrépance L^p à l'origine de la suite par

$$D_N^{(p)} \quad := \quad \|F_N - F\|_p \;\; , \;\; p \in \mathbb{N}^* \cup \{\infty\}$$

où $F(x) := x$ sur $[0,1]$.
a. Montrer que si $x \in]0,1[$ vérifie $|F_N(x) - x| \geq b$, $b \in \mathbb{R}_+$, alors $D_N^{(1)} \geq b^2/2$.
b. Déduire de a. que $D_N^{(1)} \to 0$ implique $D_N^{(\infty)} \to 0$.
c. En déduire que s'il existe $p \in \mathbb{N}^* \cup \{\infty\}$ tel que $D_N^{(p)} \to 0$, alors $D_N^{(q)} \to 0$ pour tout $q \in \mathbb{N}^* \cup \{\infty\}$.
d. Rappeler la définition d'une suite uniforme sur $[0,1]$ et montrer qu'une suite est uniforme si et seulement si $D_N^{(p)} \to 0$ pour tout $p \in \mathbb{N}^* \cup \{\infty\}$.
e. En déduire que si $(\xi_n)_{n\geq 1}$ est une suite de v.a. indépendantes uniformément distribuées sur $[0,1]$ alors $\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{\xi_n \leq x}$ tend vers x uniformément sur $[0,1]$ p.s.
2. Soit f une fonction réelle à dérivée bornée sur $[0,1]$ et U une loi uniforme sur $[0,1]$. Montrer que pour $p \in \mathbb{N}^*$

$$\left| \mathbb{E}[f(U)] - \frac{1}{N}\sum_{n=1}^{N} f(\xi_n) \right| \quad \leq \quad \|f'\|_q \; D_N^{(p)}$$

où q vérifie $\frac{1}{p} + \frac{1}{q} = 1$ (on pensera a utiliser l'égalité $f(x) = f(0) + \int_0^x f'(z)dz$).
3. En utilisant la loi du log itéré[2], montrer que si $(\xi_n)_n$ est une suite de v.a. indépendantes uniformément distribuées sur $[0,1]$ alors

$$\limsup_{N\to\infty} \sqrt{\frac{N}{2\ln\ln N}}\; D_N^{(p)} \quad \leq \quad \|\gamma\|_p \;\; , \;\; p \in \mathbb{N}^*$$

où $\gamma(x) := \sqrt{x(1-x)}$ sur $[0,1]$.
4. a. Montrer que

$$(D_N^{(2)})^2 \quad = \quad \frac{1}{N^2}\sum_{i,j\leq N} (1 - \max(\xi_i, \xi_j)) - \frac{1}{N}\sum_{i=1}^{N}(1 - \xi_i^2) + \frac{1}{3} \; .$$

[2]Soit $(X_n)_n$ est une suite de v.a. indépendantes de même loi, d'espérance nulle et de variance égale à 1. On pose $S_N = (\sum_{n=1}^{N} X_n)/\sqrt{N}$. Alors $\limsup_{N\to\infty} S_N/\sqrt{2\ln\ln N} = 1$ p.s.

b. On suppose que $(\xi_n)_{n\geq 1}$ est une suite de v.a. indépendantes uniformément distribuées sur $[0,1]$. Montrer que $\mathbb{E}\left[(D_N^{(2)})^2\right] = 1/(6N)$. Que peut-on dire sur $\mathbb{E}[|\mathbb{E}[f(U)] - N^{-1}\sum_{n=1}^N f(\xi_n)|^2]$ si f et U sont comme dans la question 2. ?

5. On revient au cas où la suite $(\xi_n)_{n\geq 1}$ est déterministe. Etant donnée une v.a V uniformément distribuée sur $[0,1]$, on définit la suite $(\tilde{\xi}_n)_n$ par $\tilde{\xi}_n := \{\xi_n + V\}$ où $\{y\}$ désigne la partie fractionnaire de y. On note

$$\tilde{D}_N^{(\infty)} \;:=\; \|\tilde{F}_N - F\|_\infty \quad \text{où} \quad \tilde{F}_N(x) \;:=\; \frac{1}{N}\sum_{n=1}^N \mathbf{1}_{\tilde{\xi}_n \leq x} \,.$$

a. Montrer que pour tout n, $\tilde{\xi}_n$ est de loi uniforme sur $[0,1]$.

b. On fixe $N \geq 1$. En décomposant $\{1,\ldots,N\}$ en $J_1 := \{j \leq N \;:\; 0 \leq \xi_j + V - 1 \leq x\}$ et $J_2 := \{j \leq N \;:\; 0 \leq \xi_j + V \leq x\}$, montrer que

$$|\tilde{F}_N(x) - x| \;\leq\; \sup_{0\leq y_1 \leq y_2 \leq 1} \left| \frac{1}{N}\sum_{n=1}^N \mathbf{1}_{y_1 \leq \xi_n \leq y_2} - (y_2 - y_1) \right| \;,\; \forall\, x \in [0,1]$$

et en déduire que

$$\tilde{D}_N^{(\infty)} \;\leq\; 2D_N^{(\infty)} \,.$$

c. Montrer que $\tilde{\sigma}_N(x) := \mathrm{Var}\left[\tilde{F}_N(x)\right] \leq 4(D_N^{(\infty)})^2$ pour tout $x \in [0,1]$, et que, si f et U sont comme dans la question 2., alors

$$\mathbb{E}\left[|\mathbb{E}[f(U)] - \hat{m}_N|^2\right] \;\leq\; 4(\|f'\|_1 D_N^{(\infty)})^2 \,,$$

où $\hat{m}_N := \frac{1}{N}\sum_{n=1}^N f(\tilde{\xi}_n)$.

d. Calculer $\mathbb{E}[\hat{m}_N]$ et proposer une méthode pour estimer $\mathrm{Var}\,[\hat{m}_N]$.

Remerciements : Je tiens à remercier sincèrement Fabian Astic, Laure Elie et Nizar Touzi pour leurs conseils et leur aide lors de l'écriture de la première version de ces notes.

Bibliographie

[1] Alfonsi A., On the discretization schemes for the CIR (and Bessel squared) processes, *rapport CERMICS [2005-279]*, 2005.

[2] Arouna B., Adaptative Monte Carlo method, a variance reduction technique, *prépublication*.

[3] Arouna B., Robbins-Monro algorithms and variance reduction in finance, *The Journal of Computational Finance*, 7 (2), 2003.

[4] Bally V. et G. Pagès, A quantization algorithm for solving multi-dimensional optimal stopping problems, *Bernoulli*, 9(6), 1003-1049, 2003.

[5] Bally V. et G. Pagès, Error analysis of the quantization algorithm for obstacle problems, *Stochastic Processes and Their Applications*, 106(1), 1-40, 2003.

[6] Bally V. et D. Talay, The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function, *Probability Theory and Related Fields*, 104, 43-60, 1996.

[7] BenHamou E., *Application of Malliavin calculus and Wiener chaos to option pricing theory*, Phd thesis, The London School of Economics and Political Science, 2000.

[8] BenHamou E., An Application of Malliavin Calculus to Continuous Time Asian Options Greeks, preprint, 2000.

[9] Bouchard B. et J.-F. Chassagneux, Discrete time approximation for continuously and discretely reflected BSDE's, prépublication PMA, 2006.

[10] Bouchard B., I. Ekeland et N. Touzi, On the Malliavin approach to Monte Carlo approximation of conditional expectations, *Finance and Stochastics*, 8 (1), 45-71, 2004.

[11] Bouchard B. et R. Elie, Discrete time approximation of decoupled Forward-Backward SDE with jumps, prépublication PMA, 2006.

[12] Bouchard B. et N. Touzi, Discrete Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations, *Stochastic Processes and their Applications*, 111 (2), 175-206, 2004.

[13] Bouleau N. et D. Lépingle, *Numerical methods for stochastic processes*, Wiley series in probability and mathematical statistics, New-York, 1994.

[14] Boyle P., M. Broadie et P. Glasserman, Monte-Carlo Methods for security pricing, *Journal of Economic Dynamics and Control*, 21, 1267-1321, 1997.

[15] Broadie M. et J. Detemple, Recent advances in numerical methods for pricing derivative securities, in *Numerical Methods in Finance*, edt L.G.G. Rogers et D. Talay, Cambridge University Press, Cambridge, U.K, 43-66, 1998.

[16] Carriere E., Valuation of the Early-Exercise Price for Options using Simulations and Nonparametric Regression, *Insurance : mathematics and Economics*, 19, 19-30, 1996.

[17] Ciarlet P.G., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Collection Mathématiques Appliquées pour la Maîtrise, Masson.

[18] Clark J.M.C. et R.J. Cameron, The maximum rate of convergence of discrete approximation for stochastic differential equations, in *Lecture Notes in Control and Information Sciences*, *Stochastic Differential Systems*, 25, Edt B. Grigelionis, Springer Verlag, Berlin, 1980.

[19] Clément E., D. Lamberton, et P. Protter, An analysis of a least squares regression method for American option pricing, *Finance and Stochastics*, 6, 449-472, 2002.

[20] Cvitaniè J., J. Ma et J. Zhang, Efficient Computation of Hedging Portfolios for Options with Discontinuous Payoffs, to appear in *Mathematical Finance*, 2001.

[21] Detemple J., R. Garcia et M. Rindisbacher, Representation formulas for Malliavin derivatives of diffusion processes, preprint, 2002.

[22] Detemple J., R. Garcia et M. Rindisbacher, Asymptotic Properties of Monte Carlo Estimators of Derivatives, preprint, 2005.

[23] Faure O., *Simulation du mouvement brownien et des diffusions*, thèse de l'Ecole Nationale des Ponts et Chaussées, 1992.

[24] Fishman G. S., *Monte Carlo. Concepts, Algorithms, and Applications*, Springer Series in Operation Research, Springer, 1995.

[25] Fournié E. , J.-M. Lasry, J. Lebuchoux et P.-L. Lions, Applications of Malliavin calculus to Monte Carlo methods in finance II, *Finance and Stochastics*, 5, 201-236, 2001.

[26] Fournié E. , J.-M. Lasry, J. Lebuchoux, P.-L. Lions et N. Touzi, Applications of Malliavin calculus to Monte Carlo methods in finance, *Finance and Stochastics*, 3, 391-412, 1999.

[27] Fournié E. , J.-M. Lasry et P.-L. Lions, Some nonlinear methods for studying far-from-the-money contingent claims, in *Numerical Methods in Finance*, edt L.G.G. Rogers et D. Talay, Cambridge University Press, Cambridge, U.K, 115-145, 1998.

[28] Fournié E. , J. Lebuchoux et N. Touzi, Small Noise Expansion and Importance Sampling, Asymptotic Analysis, 14, 361-376, 1997.

[29] Glasserman P. et D.D. Yao, Some guidelines and guarantees for common random numbers, *Manag. Sci.*, 38, 884-908, 1992.

[30] Glynn P.W., Optimization of stochastic systems via simulation, in *Proceedings of the 1989 Winter simulation Conference*, San Diego, Society for Computer Simulation, 90-105, 1989.

[31] Gobet E., Weak approximation of killed diffusion using Euler schemes, *Stochastic Processes and their Applications*, 87, 167-197, 2000.

[32] Gobet E., Euler schemes and half-space approximation for the simulation of diffusion in a domain, *ESAIM Probability and Statistics*, 5, 261-297, 2001.

[33] Gobet E. et A. Kohatsu-Higa, Computation of Greeks for barrier and lookback options using Malliavin calculus, *Electronic Communications in Probability*, 8, 51-62, 2003.

[34] Gobet E., J.P. Lemor et X. Warin, Rate of convergence of empirical regression method for solving generalized BSDE. *preprint Ecole Polytechnique*, R.I.574, 2005.

[35] Gobet E. et R. Munos, Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to stochastic optimal, *SIAM Journal on Control and Optimization*, 43(5), 1676-1713, 2005.

[36] Jäckel P., *Monte-Carlo methods in finance*, Wiley Finance Series, Wiley, 2002.

[37] Karatzas I. et S.E. Shreve, *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, 113, Springer-Verlag, New York, 1998.

[38] Karatzas I. et S.E. Shreve, *Methods of mathematical finance*, Applications of Mathematics, Stochastic Modelling and Applied Probability, 39, Springer-Verlag, New York, 1999.

[39] Kemna A.G.Z. et A.C.F Vorst, A pricing method for options based on average asset values, *J. Banking Finan.*, 14 (1), 1990.

[40] Kloeden P.E. et E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics, 23, Springer-Verlag, New York, 1992.

[41] Kloeden P.E., E. Platen et H. Schurz, *Numerical solution of SDE through computer experiments*, Universititext, Springer-Verlag, New York, 1992.

[42] Kohatsu-Higa A. et R. Pettersson, Variance reduction methods for simulation of densities on Wiener space, *SIAM Journal on Numerical Analysis*, to appear, 2001.

[43] Kunita H., Stochastic differential equations and stochastic flows of diffeomorphisms, *Cours de l'école d'été de Saint-Flour, Lec. Notes in Math.*, 1097, Springer-Verlag, 1982.

[44] Lamberton D. et B. Lapeyre, *Une introduction au calcul stochastique appliqué à la finance*, Collection Mathématiques et Applications, Ellipse, 1991.

[45] Lapeyre B., E. Pardoux et R. Sentis, *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*, Mathématiques et Applications, 29, Springer, 1998.

[46] Lapeyre B., E. Temam, Competitive Monte Carlo methods for the pricing of Asian Options, *Journal of Computational Finance*, 5 (1), 2001.

[47] L'Ecuyer P. et G. Perron, On the convergence rates of IPA and FDC derivatives estimators, *Operation Research*, 42, 643-656, 1994.

[48] Lions P.-L. et H. Regnier, Calcul du prix et des sensibilités d'une option américaine par une méthode de Monte Carlo, preprint, 2001.

[49] Longstaff F.A. et R.S. Schwartz, Valuing American Options By Simulation : A simple Least-Square Approach, *Review of Financial Studies*, 14, 113-147, 2001.

[50] Malliavin P. et H. Airault, *Intégration, analyse de Fourier, Probabilités, analyse gaussienne*, Collection Maîtrise de Mathématiques pures, Masson.

[51] Musiela M. et M. Rutkowski, *Martingale methods in financial modelling*, Applications of Mathematics, Stochastic Modelling and Applied Probability, 36, Springer-Verlag, New York, 1997.

[52] Newton N. J., Variance reduction for simulated diffusions, *SIAM Journal on Applied Mathematics*, 54 (6), 1780-1805, 1994.

[53] Newton N. J., Continuous-time Monte-Carlo methods and variance reduction, in *Numerical Methods in Finance*, edt L.G.G. Rogers et D. Talay, Cambridge University Press, Cambridge, U.K, 22-42, 1998.

[54] Nualart D., *The Malliavin Calculus and Related Topics* Springer Verlag, Berlin, 1995.

[55] Nualart D., Analysis on Wiener space and anticipating stochastic calculus, Lect. Notes Math. 1690, Springer, Berlin, 1998.

[56] Øksendal B., *An Introduction to Malliavin Calculus with Applications to Economics*, Lecture Notes from a course given 1996 at the Norwegian School of Economics and Business Administration (NHH), NHH Preprint Series, 1996.

[57] Pagès G. et Y. J. Xiao, Sequences with low discrepancy and pseudo-random numbers : theoretical remarks and numerical tests, *Journal of Statistical Computation and Simulation*, 56, pp.163-183, 1997.

[58] Protter P., *Stochastic Integration and Differential Equations*, Applications of Mathematics, Stochastic Modelling and Applied Probability, 21, Springer-Verlag, New York, 1990.

[59] Revuz D. et M. Yor, *Continuous martingales and Brownian motion*, 2nd Ed. Grundlehren der Mathematischen Wissenschaften, 293, Springer, 1994.

[60] Rogers L.C.G., Monte Carlo valuation of American options, *Mathematical Finance*, 12, 271-286, 2002.

[61] Seumen-Tonou P., *Méthodes numériques probabilistes pour la résolution d'équations du transport et pour l'évaluation d'options exotiques*. Thèse de doctorat, Université Provence-Aix-Marseille I, 1997.

[62] Talay D., Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution, *Mathematical Modelling and Numerical Analysis*, 20, 141-179, 1986.

[63] Talay D. et L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, *Stochastic Analysis and Applications*, 8 (4), 483-509, 1990.

[64] Temam E., *Couverture approchée d'options exotiques - Pricing des options asiatiques*, thèse de doctorat de Paris VI, 2001.

[65] Xiao Y. J., *Contributions aux méthodes arithmétiques pour la simulation accélérée*, Thèse de l'Ecole Nationale des Ponts et Chaussées, Paris, 1990.

[66] Zhang J., *Some fine properties of backward stochastic differential equations*, PhD thesis, Purdue University, 2001.