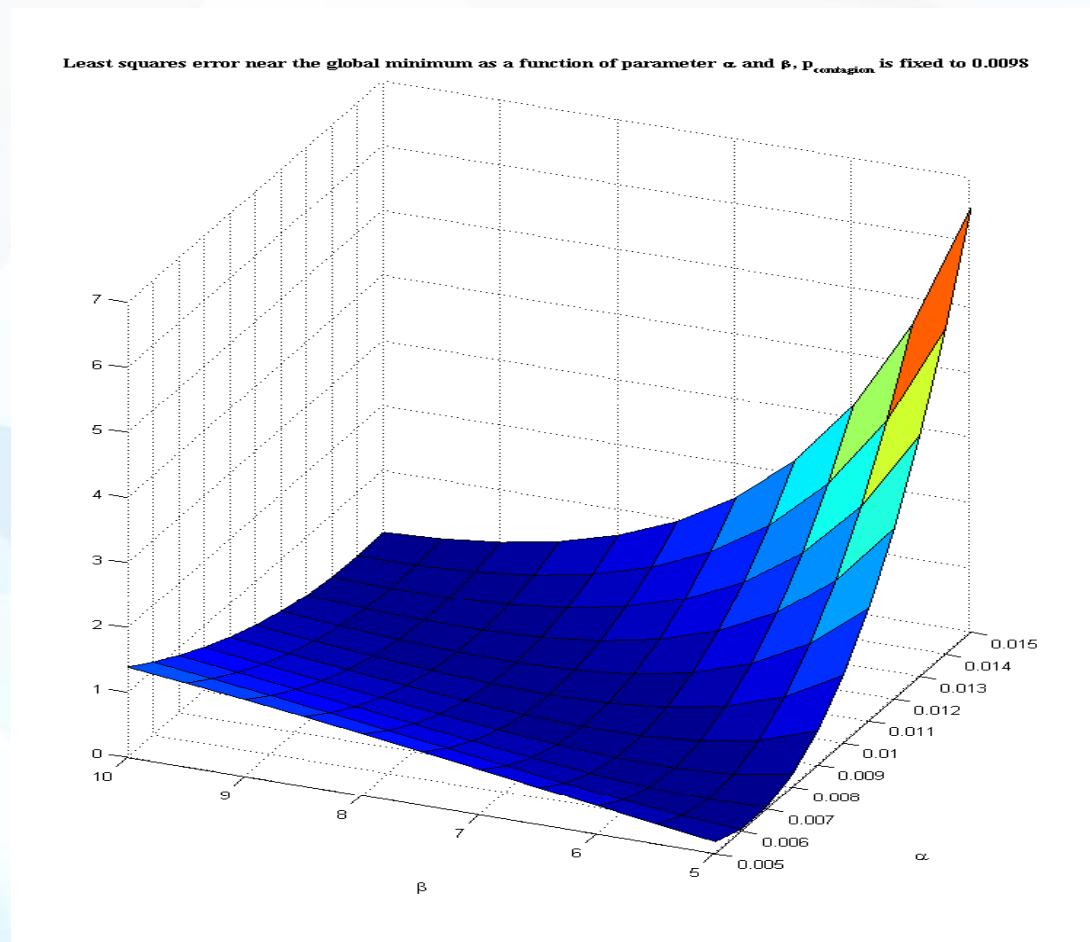


Techniques numériques en finance et actuariat

Armand Bernou - armand.bernou@univ-lyon1.fr



Année 2023-2024



Chapitre 3: Méthodes de Monte-Carlo

Principe de base et rappels

Méthodes de réduction de la variance

Importance Sampling

Variables antithétiques

Variables de contrôle

Discrétisation et simulation de processus aléatoires

Simulation du Brownien géométrique

Problématique

On veut toujours estimer $\mathbb{E}[f(X_T)]$. Mais, pour les équations qui nous intéressent, les EDS n'ont souvent pas de solution explicite (typiquement le cas des modèles de taux d'intérêt).

On a aussi des cas où la solution est explicite, mais trop difficile à simuler directement.

Monte-Carlo: on estime $\mathbb{E}[f(X_T)]$ par $\frac{1}{n} \sum_{i=1}^n f(\tilde{X}_t^i)$ où \tilde{X}_t^i est un schéma.

On a ainsi deux erreurs: une erreur de discrétisation (du schéma) et une erreur statistique.

Problématique II

En général, on cherche donc une méthode pour simuler

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t,$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien, en dimension $d \geq 1$.

Ces dynamiques sont importantes

- ▶ en finance (pricing)
- ▶ en assurance: simulation de courbes des taux, évaluation...

On peut réécrire l'EDS

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$$

Problématique III

© Théo Jalabert



On peut réécrire l'EDS

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$$

On cherche donc à trouver un algorithme/schéma pour résoudre cette équation, et en fait pour approcher

$$\int_0^t b(s, X_s) ds \quad \text{et} \quad \int_0^t \sigma(s, X_s) dW_s.$$

Rappels sur la méthode I

Soit X_1, \dots, X_N une suite de variables aléatoires (v.a.) indépendante et identiquement distribuées (i.i.d.) de même loi que X (avec $\text{Var}(X) = \sigma^2$).

On définit la moyenne arithmétique de cette suite :

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i.$$

L'estimateur de Monte-Carlo de $\mathbb{E}[X]$ est donné par

$$\mathbb{E}[X] \approx \bar{X}_N,$$

et sa variance par:

$$\text{Var}(\bar{X}_N) \approx \frac{1}{N^2} \sum_{i=1}^N \text{Var}(X_i) = \frac{\sigma^2}{N}.$$

Rappels II

- On suppose que $\mathbb{E}[|X|] < \infty$, alors (Loi Forte des Grands Nombres) :

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \rightarrow \mathbb{E}[X], \text{ p.s.}$$



- On suppose que $\mathbb{E}[|X|] < \infty$, alors (Loi Forte des Grands Nombres) :

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \rightarrow \mathbb{E}[X], \text{ p.s.}$$

- Si de plus, $\sigma^2 = \text{Var}(X) < \infty$, alors (Théorème de la Limite Centrale)

$$\sqrt{N} \frac{\bar{X}_N - \mathbb{E}[X]}{\sigma} \rightarrow \mathcal{N}(0, 1)$$

avec une convergence en loi.

- $\bar{\sigma}_N^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X}_N)^2$: estimateur sans biais de σ^2

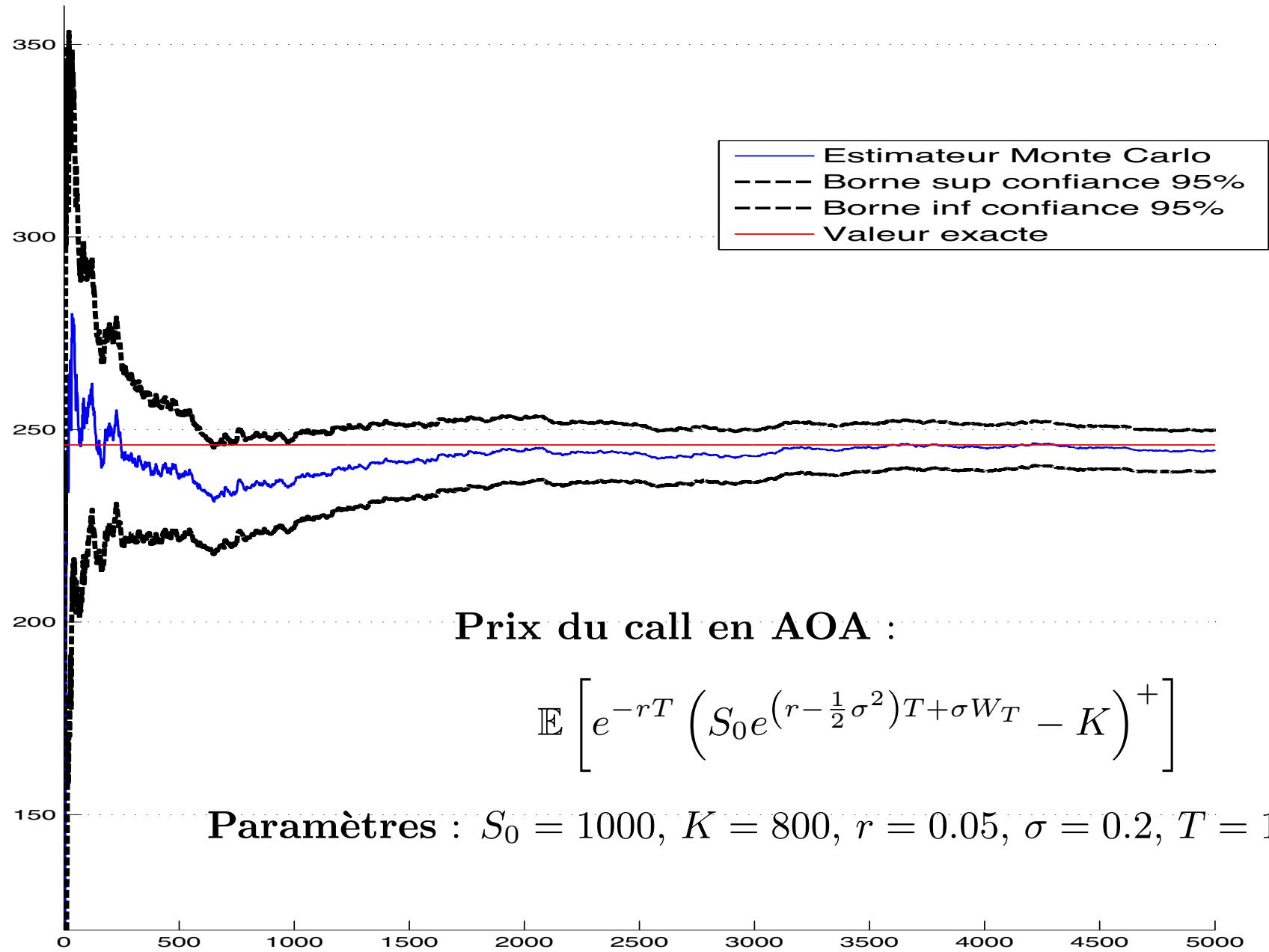
Règle approximative : à 95%,

$$\mathbb{E}[X] \in \left[\bar{X}_N - 1.96 \frac{\bar{\sigma}_N}{\sqrt{N}}, \bar{X}_N + 1.96 \frac{\bar{\sigma}_N}{\sqrt{N}} \right]$$

- Vitesse de convergence lente en \sqrt{N} mais indépendante de la dimension.

Exemple avec intervalle de confiance

© Théo Jalabert





Chapitre 3: Méthodes de Monte-Carlo

Principe de base et rappels

Méthodes de réduction de la variance

Importance Sampling

Variables antithétiques

Variables de contrôle

Discrétisation et simulation de processus aléatoires

Simulation du Brownien géométrique

Problématique

On a vu précédemment que le TCL fournit, pour l'estimation de $\mathbb{E}[f(X)]$, une estimation de l'erreur de Monte-Carlo en \sqrt{N} , où N est le nombre de v.a. utilisés dans la méthode.

C'est une vitesse de convergence relativement lente: pour améliorer la précision d'une décimale, il faut multiplier par 100 le nombre de variables utilisées !

Pour ces raisons, on a souvent recours à des méthodes de *réduction de la variance*. Elle consiste à remplacer l'estimateur

$$\hat{E}_n = \frac{1}{n} \sum_{i=1}^n f(X_i)$$

utilisé classiquement pour estimer $\mathbb{E}[f(X)]$ par des estimateurs modifiés.

On va voir essentiellement trois méthodes:

- ▶ Importance Sampling;
- ▶ Variables antithétiques;
- ▶ Variables de contrôle.

Importance Sampling - Principe général

© Théo Jalabert



Soit $X \sim p$ et $Y \sim q$ avec p, q deux densités sur \mathbb{R} . On cherche à estimer $\mathbb{E}[f(X)]$.

Si q ne s'annule pas sur le support de pf , alors

$$\mathbb{E}[f(X)] = \mathbb{E}\left[\frac{f(Y)p(Y)}{q(Y)}\right].$$

Le nouvel estimateur est donc

$$\hat{E}_n = \frac{1}{n} \sum_{i=1}^n \frac{f(Y_i)p(Y_i)}{q(Y_i)},$$

avec $(Y_i)_{1 \leq i \leq n}$ suite i.i.d. de densité q .

Ainsi, si l'on sait simuler Y , et que l'on a un petit peu d'information sur fp , on peut:

- ▶ Estimer $\mathbb{E}[f(X)]$ en simulant selon la densité q ;
- ▶ Réduire la variance en choisissant adroitemment la densité q (critère : $|f|p/q$ quasi constant et de variance finie).

Importance Sampling - Exemple

© Théo Jalabert



On cherche à estimer $\mathbb{P}(X > 5)$ pour $X \sim \mathcal{N}(0, 1)$. L'événement est très rare, donc l'estimation de Monte-Carlo fonctionne très mal.

On introduit la densité

$$q(x) = e^{-x+5} \mathbf{1}_{x>5}$$

qui correspond à une loi exponentielle de paramètre 1 translatée de 5.

On a alors

$$\mathbb{P}(X > 5) = \int_5^\infty \phi(x) dx = \int_5^\infty \phi(x) \frac{q(x)}{q(x)} = \mathbb{E}\left[\frac{\phi(Y)}{q(Y)} \mathbf{1}_{Y>5}\right].$$

En prenant 10^6 simulations, l'estimateur “naïf” renvoie 0, l'estimateur d'importance sampling renvoie 3 chiffres significatifs.



On va utiliser des symétries pour réduire le nombre de simulations nécessaires. Par exemple:

- ▶ Si $U \sim \mathcal{U}([a, b])$, $b + a - U \sim \mathcal{U}([a, b])$.
- ▶ Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $2\mu - X \sim \mathcal{N}(\mu, \sigma^2)$.

Plus généralement, si A est une transformation qui laisse la loi de X invariante, i.e. si $A(X) \stackrel{\mathcal{L}}{=} X$, on utilisera l'estimateur

$$\hat{E}_n = \frac{1}{n} \sum_{i=1}^n \frac{f(X_i) + f(A(X_i))}{2}.$$

Exemple: Modèle de Black-Scholes

On a

$$S_t = S_0 e^{(r - \frac{\sigma^2}{2})t + \sigma W_t},$$

avec r le taux d'intérêt sans risque, σ la volatilité. Pour le calcul du call, on veut obtenir $\delta = \mathbb{E}[e^{-rT}(S_T - K)_+]$. Comme $W_T \sim \mathcal{N}(0, T)$, on a, pour $X \sim \mathcal{N}(0, 1)$,

$$\delta = \mathbb{E}[h(X)], \quad \text{avec } h(x) = e^{-rT} \left(e^{(r - \frac{\sigma^2}{2})t + \sigma \sqrt{T}x} - K \right)_+.$$

On peut montrer que comme h est croissante et, comme $W \rightarrow -W$ est décroissante, la variance de l'estimateur antithétique:

$$\hat{E}_n = \frac{1}{n} \sum_{i=1}^n \frac{h(X_i) + h(-X_i)}{2},$$

est réduite par rapport à l'estimateur naïf.

Variables de contrôle - Principe

Soit $h_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ avec $\mathbb{E}[h_0(X)] = \alpha$ donné et $\text{Var}[h_0(X)] < \infty$.

L'estimateur proposé est

$$\hat{E}_n = \frac{1}{n} \sum_{i=1}^n \left[h(X_i) - b[h_0(X_i) - \alpha] \right],$$

pour $b \in \mathbb{R}$ à choisir.

On a une réduction de variance si

$$b^2 \text{Var}[h_0(X)] - 2b \text{Cov}[h(X), h_0(X)] < 0.$$

Il existe une version multidimensionnelle (avec plusieurs variables de contrôle).



Chapitre 3: Méthodes de Monte-Carlo

Principe de base et rappels

Méthodes de réduction de la variance

Importance Sampling

Variables antithétiques

Variables de contrôle

Discrétisation et simulation de processus aléatoires

Simulation du Brownien géométrique

Génération de nombres aléatoires

© Théo Jalabert



- ▶ Nécessite un générateur de nombres pseudo-aléatoires : Random(), seed à mettre à jour (ou pas) après chaque jeu de simulations.
- ▶ Voir également méthodes quasi Monte-Carlo, basées sur des suites à discrépance faible (recouvrement de $[0, 1]^d$ le plus uniformément possible). Ex : suites de Halton, Faure, Sobol

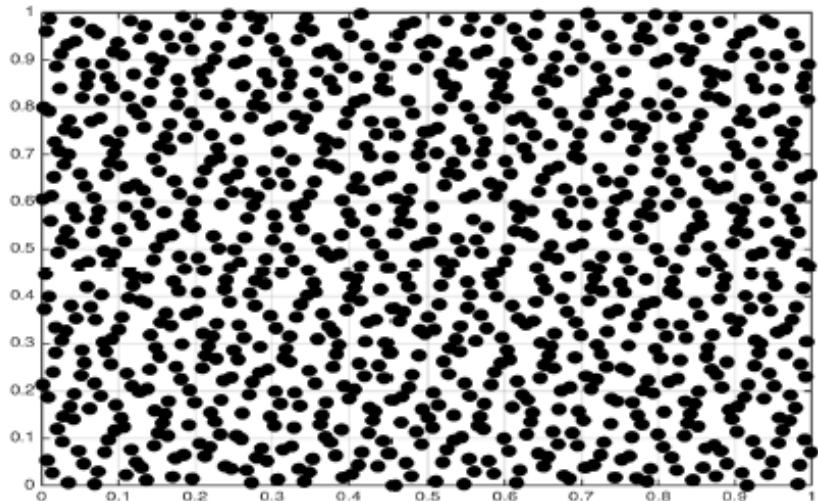
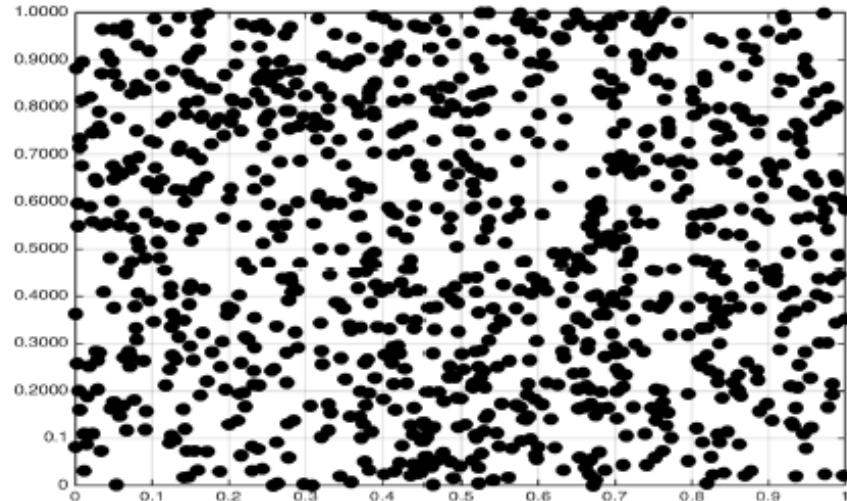


Figure: Simulation pseudo-aléatoire (gauche), suite de Halton (droite)

Ces dernières méthodes ne sont pas applicables en grande dimension/si la fonction f n'est pas suffisamment régulière.



- De manière générale, **pas de méthode universelle**. Efficacité dépend des caractéristiques de la variable à simuler.
- Méthode d'inversion;
- Méthode de rejet;
- Méthode de Walker (Voir Poly ISFA 1A - Alexis Bienvenüe);
- Simulation de variables Gaussiennes (Algorithme de Box-Muller, Algorithme de Marsiglia).

Première étape cruciale vers la simulation du Brownien.

Soit $X \sim \mathcal{N}(m, K)$ avec $K \in \mathbb{M}_{d \times d}^{\text{Sym}}$ (matrice de taille $d \times d$ symétrique).

Proposition (décomposition de Cholesky)

Il existe une matrice L telle que $K = LL^\top$. L'algorithme de Cholesky construit L triangulaire inférieure :

$$K_{i,j} = \sum_{k=1}^d L_{i,k} L_{k,j}^\top = \sum_{1 \leq k \leq i \wedge j} L_{i,k} L_{j,k}$$

Algorithme de Cholesky

© Théo Jalabert



Algorithme (décomposition de Cholesky)

Colonne 1 de L :

- ▶ $L_{1,1} = \sqrt{K_{1,1}},$
- ▶ $L_{i,1} = \frac{K_{1,i}}{L_{1,1}}, i \geq 1$

Colonne j de L :

- ▶ $L_{j,j} = \sqrt{K_{j,j} - \sum_{k=1}^{j-1} L_{j,k}^2},$
- ▶ $L_{i,j} = \frac{K_{i,j} - \sum_{k=1}^{j-1} L_{i,k} L_{j,k}}{L_{j,j}}, i > j$

Complexité en $\mathcal{O}(d^3)$.

Exercice : Écrire la décomposition de Cholesky en dimension 2.

Corollaire: Soit X_1, \dots, X_d des variables de loi $\mathcal{N}(0, 1)$ indépendantes et $\tilde{X} = (X_1, \dots, X_d)^\top$ avec m un vecteur de \mathbb{R}^d et K une matrice de covariance, alors:

$$X = m + L\tilde{X} \sim \mathcal{N}(m, K)$$

où L est une racine de K , i.e., $K = LL^\top$.

Le mouvement Brownien

On appelle mouvement Brownien uni-dimensionnel standard défini sur $[0, T]$ un processus $W = (W_t)_{t \in [0, T]}$ avec

1. $W_0 = 0$,
2. Presque sûrement, $t \mapsto W_t$ est continue;
3. Pour tout $k \geq 1$ et $0 \leq t_0 < t_1 < \dots < t_k \leq T$, les accroissements $\{W_{t_1} - W_{t_0}, \dots, W_{t_k} - W_{t_{k-1}}\}$ sont indépendants;
4. Pour tout $0 \leq s < t \leq T$,

$$W_t - W_s \sim \mathcal{N}(0, t - s).$$

Simulation d'un Mouvement Brownien

En dimension 1:

Simulation d'un mouvement Brownien unidimensionnel standard

Soit $n \geq 1$ et $0 = t_0 < t_1 < \dots < t_n = T$. On pose

$$W_{t_0} = 0,$$

$$W_{t_{k+1}} = W_{t_k} + \sqrt{t_{k+1} - t_k} X_k, \quad k = 1, \dots, n$$

avec X_1, \dots, X_n i.i.d. de loi $\mathcal{N}(0, 1)$.

Simulation d'un Mouvement Brownien

En dimension 1:

Simulation d'un mouvement Brownien unidimensionnel standard

Soit $n \geq 1$ et $0 = t_0 < t_1 < \dots < t_n = T$. On pose

$$W_{t_0} = 0,$$

$$W_{t_{k+1}} = W_{t_k} + \sqrt{t_{k+1} - t_k} X_k, \quad k = 1, \dots, n$$

avec X_1, \dots, X_n i.i.d. de loi $\mathcal{N}(0, 1)$.

En dimension d , pour simuler un mouvement Brownien multivarié de matrice de corrélation K , i.e. tel que $\text{Cov}(W_t^i, W_t^j) = tK_{i,j}$, $1 \leq i, j \leq d$:

Simulation d'un mouvement Brownien multivarié de matrice de corrélation K

Soit L tel que $K = LL^\top$. On simule \tilde{W} un mouvement Brownien standard (avec d composantes indépendantes) à l'aide de l'algorithme précédent (appliqué d fois). Puis, pour $t = t_0, \dots, t_n$:

$$W_t = L\tilde{W}_t.$$

Raffiner les trajectoires sans nouvelles simulations ?

© Théo Jalabert



Soit $0 = t_0 < t_1 < \dots < t_n = T$ une suite de temps croissante.

- On suppose que l'on a simulé un mouvement brownien 1d standard à ces dates, i.e., W_{t_0}, \dots, W_{t_n} .
- Comment simuler W_t , pour $t_k < t < t_{k+1}$ sachant W_{t_k} et $W_{t_{k+1}}$?

Pont Brownien :

La loi conditionnelle de W_t (pour $t_k < t < t_{k+1}$), sachant $W_{t_k} = x$ et $W_{t_{k+1}} = y$ est

$$\mathcal{N}\left(x \frac{t_{k+1}-t}{t_{k+1}-t_k} + y \frac{t-t_k}{t_{k+1}-t_k}, \frac{(t-t_k)(t_{k+1}-t)}{t_{k+1}-t_k}\right)$$

Lorsque $t = \frac{t_k+t_{k+1}}{2}$, cela devient

$$\mathcal{N}\left(\frac{x+y}{2}, \frac{t_{k+1}-t_k}{4}\right)$$



Chapitre 3: Méthodes de Monte-Carlo

Principe de base et rappels

Méthodes de réduction de la variance

Importance Sampling

Variables antithétiques

Variables de contrôle

Discrétisation et simulation de processus aléatoires

Simulation du Brownien géométrique

Simuler un mouvement Brownien géométrique

© Théo Jalabert



On considère d sous-jacents et p Browniens indépendants dans la modélisation du prix, ainsi que le **cas de coefficients constants**. On a donc, pour $i \in \{1, \dots, d\}$

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_i dt + \sum_{j=1}^p \sigma_{i,j} dW_{j,t}.$$

Simulation à une date précise t

1. Simuler un mouvement Brownien standard de composantes indépendantes $W_t = (W_{1,t}, \dots, W_{p,t})$ à la date t .
2. Pour $i = 1, \dots, d$, calculer

$$S_{i,t} = S_{i,0} e^{(\mu_i - \frac{1}{2} \sum_{j=1}^p \sigma_{i,j}^2) t + \sum_{j=1}^p \sigma_{i,j} W_{j,t}}$$

Simulation d'une trajectoire en $t_0 < \dots < t_n$

1. Simuler les accroissements indépendants du Brownien $W_{t_{k+1}} - W_{t_k}$, $k = 0, \dots, n - 1$
2. Pour $i = 1, \dots, d$, calculer

$$S_{i,t_{k+1}} = S_{i,t_k} e^{(\mu_i - \frac{1}{2} \sum_{j=1}^p \sigma_{i,j}^2)(t_{k+1} - t_k) + \sum_{j=1}^p \sigma_{i,j} (W_{j,t_{k+1}} - W_{j,t_k})}$$

Simulation (exacte) d'un mouvement Brownien géométrique II

© Théo Jalabert



Cas de coefficients dépendants du temps : on a maintenant

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_i(t)dt + \sum_{j=1}^p \sigma_{i,j}(t)dW_{j,t}, \quad i = 1, \dots, d.$$

Simulation d'une trajectoire en $t_0 < \dots < t_n$

$$S_{i,t_{k+1}} = S_{i,t_k} e^{\int_{t_k}^{t_{k+1}} (\mu_i(t) - \frac{1}{2} \sum_{j=1}^p \sigma_{i,j}^2(t)) dt + \sum_{j=1}^p \int_{t_k}^{t_{k+1}} \sigma_{i,j}(t) dW_{j,t}}, \quad i = 1, \dots, d$$

Comment simuler cette dernière somme dans l'exponentielle ?

Sur chaque intervalle $[t_k, t_{k+1}]$, $k = 0, \dots, n - 1$, simuler un vecteur gaussien centré de matrice de covariance $K = (K_{l,m})$, $1 \leq l, m \leq d$ donnée par:

$$K_{l,m} = \sum_{j=1}^p \int_{t_k}^{t_{k+1}} \sigma_{l,j}(t) \sigma_{m,j}(t) dt = \int_{t_k}^{t_{k+1}} \left(\sigma(t) \sigma(t)^\top \right)_{l,m} dt$$

où $\sigma(t) = (\sigma(t)_{i,j})_{1 \leq i \leq d, 1 \leq j \leq p}$ est la matrice de volatilité instantanée associée au BG.



Supposons que l'actif risqué est modélisé par l'EDS suivante :

$$dS_t = S_t r dt + S_t \sigma(t, S_t) dW_t$$

Problème : Même si $\sigma(t, x)$ est une fonction bornée, le coefficient de volatilité $x\sigma(t, x)$ n'est pas borné. Pour des questions de stabilité et de qualité de résultats numériques, on va à nouveau travailler sur des "log-prix", i.e., $X_t = \ln(S_t)$:

$$X_0 = x, \quad dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

- Dans le cas multivarié, le logarithme est à considérer composante par composante.
- Dans l'exemple précédent $b(t, x) = r - \frac{1}{2}\sigma^2(t, x)$ (cf Chapitre 2).



Simulation d'une EDS vectorielle :

$$X_0 = x, \quad dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t.$$

Soit la grille temporelle $(t_k)_{k=0, \dots, n}$ tel que $t_k = k\Delta_n$ où $\Delta_n = \frac{T}{n}$. L'idée du schéma d'Euler est issue de l'approximation suivante :

$$\begin{aligned} X_{t_{k+1}} &= X_{t_k} + \int_{t_k}^{t_{k+1}} b(s, X_s)ds + \int_{t_k}^{t_{k+1}} \sigma(s, X_s)dW_s \\ &\simeq X_{t_k} + b(t_k, X_{t_k})\Delta_n + \sigma(t_k, X_{t_k})(W_{t_{k+1}} - W_{t_k}) \end{aligned}$$

Schéma de discrétisation d'Euler

© Théo Jalabert



Pour simuler une trajectoire du processus X , on construit l'approximation $(X_{t_0}^n, \dots, X_{t_n}^n)$ suivante :

$$\begin{cases} X_0^n = x \\ X_{t_{k+1}}^n = X_{t_k}^n + b(t_k, X_{t_k}^n) \Delta_n + \sigma(t_k, X_{t_k}^n) (W_{t_{k+1}} - W_{t_k}), \quad k = 0, \dots, n-1 \end{cases}$$

- Nécessite la simulation des accroissements d'un brownien;
- Procédure exacte dans le cas de coefficients b et σ constant;
- Facilité d'implémentation, rapidité de simulation.

Idée d'accélération antithétique de la méthode de Monte-Carlo avec ce schéma ?

Une alternative: schéma de Milstein

Soit t dans l'intervalle $[t_k, t_{k+1}]$,

$$X_t = X_{t_k} + \int_{t_k}^t b(s, X_s) ds + \int_{t_k}^t \sigma(s, X_s) dW_s$$

On a

$$\mathbb{E} \left[\left(\int_{t_k}^t b(s, X_s) ds \right)^2 \right] \leq (t - t_k) \mathbb{E} \left(\int_{t_k}^t b^2(s, X_s) ds \right) \sim C(t - t_k)^2$$

$$\text{et } \mathbb{E} \left[\left(\int_{t_k}^t \sigma(s, X_s) dW_s \right)^2 \right] = \mathbb{E} \left(\int_{t_k}^t \sigma^2(s, X_s) ds \right) \sim C(t - t_k)$$

→ Terme d'intégrale stochastique dominant lorsque le pas de temps est petit.

Cas de la dimension 1 :

$$\begin{aligned} X_0^n &= x, \quad X_{t_{k+1}}^n = X_{t_k}^n + b(t_k, X_{t_k}^n) \Delta_n + \sigma(t_k, X_{t_k}^n) (W_{t_{k+1}} - W_{t_k}) \\ &\quad + \frac{1}{2} \sigma(t_k, X_{t_k}^n) \partial_x \sigma(t_k, X_{t_k}^n) \left[(W_{t_{k+1}} - W_{t_k})^2 - \Delta_n \right] \end{aligned}$$

- Coïncide avec le schéma d'Euler si $\sigma(t, x) = \sigma(t)$;
- Meilleure vitesse de convergence (s'attaque directement au terme d'erreur le plus fort);
- En dimension $d > 1$, extension possible mais en toute généralité, nécessite de simuler $\int_{t_k}^{t_{k+1}} W_{i,s} dW_{j,s}$ avec $i \neq j$.

Erreur faible : $\mathbb{E}[f(X_T)] - \mathbb{E}[f(X_T^n)] = \varepsilon(f, n)$

Vitesse de convergence

b et σ sont fonctions Lipschitziennes de t et de X_t .

Hypothèses	Euler $\varepsilon(f, n)$	Milstein $\varepsilon(f, n)$
H1 : f est régulière ($f \in C^4$)	$\frac{C_1}{n} + \frac{C_2}{n^2} + \dots$	$\frac{\hat{C}_1}{n} + \frac{\hat{C}_2}{n^2} + \dots$
H2 : $\sigma\sigma^\top(0, X_0)$ inversible	idem	idem

H2 est le cadre intéressant pour nous

- f sans régularité supposée, par exemple $f(x) = (x - K)^+$, $f(x) = \mathbf{1}_{x \geq K}$
- hypothèse de non-dégénérescence sur σ : ellipticité. Cela implique que la loi de X_T a une densité régulière (théorème d'Hörmander).

A propos du schéma d'Euler

Cette méthode de simulation propose un bon arbitrage entre simplicité de la stratégie et vitesse de convergence.

L'extrapolation de Romberg permet d'améliorer la vitesse de convergence:

Alors que le schéma repose sur l'égalité:

$$\mathbb{E}[f(X_T^n)] - \mathbb{E}[f(X_T)] = O\left(\frac{1}{n}\right)$$

On peut plutôt utiliser

$$2\mathbb{E}[f(X_T^{2n})] - \mathbb{E}[f(X_T^n)] = \mathbb{E}[f(X_T)] + O\left(\frac{1}{n^2}\right).$$