Langage C et aléa, séance 2

École des Mines de Nancy, séminaire d'option Ingénierie Mathématique

Frédéric Sur
http://www.loria.fr/~sur/enseignement/coursCalea/

1 La méthode de Monte-Carlo

Origine du nom[1] :

> *Rolling dice, shuffling decks of cards, spinning roulette wheels, etc., are fascinating pastimes for just about everybody. These traditional uses of random numbers have suggested the name "Monte-Carlo method," a general term used to describe any algorithm that employs random numbers.*

1.1 Justification de la méthode

Le but est d'évaluer une espérance $\mathbf{E}(f(X))$ où f est une fonction mesurable sur $\mathbb{R}^d$ ($d \geq 1$) muni d'une probabilité $\rho(x)dx$, à valeurs dans $\mathbb{R}$.

Soit (X_i) une suite de variables aléatoires indépendantes identiquement distribuées selon ρ sur $\mathbb{R}^d$.

D'après la loi des grands nombres, l'estimateur $\displaystyle\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ converge presque sûrement vers

$$\mathbf{E}(f(X_1)) = \int_{\mathbb{R}^d} f(x)\rho(x)\;dx.$$

Si de plus $\mathbf{E}(f(X_1)^2) < +\infty$, par le théorème de la limite centrale :

$$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i) - \mathbf{E}(f(X_1))\right)$$

converge en loi vers la loi normale centrée réduite $\mathcal{N}(0,1)$, où

$$\sigma^2 = \int_{\mathbb{R}^d}\Big(f - \mathbf{E}(f(X_1))\Big)^2 \rho(x)\;dx \quad \left(= \int_{\mathbb{R}^d} f^2(x)\rho(x)\;dx - \mathbf{E}(f(X_1))^2\right).$$

On déduit un intervalle de confiance pour $\mathbf{E}(f(X_1))$, avec un niveau de confiance de 95%, à l'aide d'une table de la loi normale :

$$\mathbf{E}(f(X_1)) \in \left[\frac{1}{n}\sum_{i=1}^{n} f(X_i) - \frac{1.96\sigma}{\sqrt{n}},\; \frac{1}{n}\sum_{i=1}^{n} f(X_i) + \frac{1.96\sigma}{\sqrt{n}}\right].$$

En approchant σ^2 par la variance empirique (avec un nombre de tirages n assez grand), on obtient un encadrement de $\mathbf{E}(f(X_1))$ à un niveau de confiance de 95%. En toute rigueur, il faudrait utiliser une loi de Student à $n-1$ degrés de liberté pour l'estimation de l'intervalle de confiance précédent (car l'écart-type est estimé), mais celle-ci se confond avec la loi normale lorsque n est assez grand…

[1] cf *The art of computer programming*, D.E. Knuth, chapitre 3, tome 2. Il existe aussi dans d'autres contextes des méthodes de Las Vegas ou d'Atlantic City…

Quelques remarques :

– La taille de l'intervalle de confiance est en $O(1/\sqrt{n})$, donc la convergence est lente.
– Une stratégie pour améliorer la vitesse de convergence est de diminuer la variance σ^2, par exemple en choisissant la loi ρ de manière adéquate. (plus d'informations dans la suite de votre cursus...)
– Avantages de la méthode : la convergence ne dépend pas de la dimension d de l'espace sur lequel est définie la fonction f, ni de la régularité de f ; la méthode est parallélisable (les estimations des $f(X_i)$ sont indépendantes).

1.2 En pratique

On voit qu'il « suffit » de simuler les réalisations d'une variable aléatoire de loi ρ sur $\mathbb{R}^d$.

Dans la section suivante, la distribution ρ est portée par un ensemble compact A de $\mathbb{R}^d$ d'aire m_A, et $\rho(x) = \frac{1}{m_A} 1_A(x)$, c'est-à-dire les X_i sont distribués selon une loi uniforme sur A.

Dans ce cas, l'espérance estimée par $\frac{1}{n} \sum_{i=1}^{n} f(X_i)$ sera $\mathbf{E}(f(X_1)) = \frac{1}{m_A} \int_A f(x)\, dx$.

2 Exemples d'estimation d'aires

Les deux exercices suivants consistent à calculer des aires. Dans les deux cas il existe des méthodes plus rapides d'approximation : il s'agit d'exercices de programmation avant tout.

Exercice 1 (Approximation de π)

On cherche l'aire du disque unité. Il s'agit donc d'estimer :

$$I = \int_{\mathbb{R}^2} 1_{x^2+y^2 \leq 1}(x, y)\, dxdy.$$

Si $A = [-1, 1] \times [-1, 1]$, on a aussi :

$$I = \int_A 1_{x^2+y^2 \leq 1}(x, y)\, dxdy.$$

Écrivez un programme qui donne un encadrement de I (qui vaut bien sûr π) à un niveau de confiance de 95%. Vous écrirez tout d'abord une fonction qui simule une variable aléatoire de loi uniforme sur $[-1, 1]$.

Représentez à l'aide de Gnuplot le graphe de l'estimation de π en fonction du nombre d'itérations. Vous superposerez également les courbes représentants les extrémités de l'intervalle de confiance.

Exercice 2 (Approximation de l'aire de l'ensemble de Mandelbrot)

Soit c un nombre complexe, on définit la suite (z_n) par la relation de récurrence :

$$\begin{cases} z_0 = c \\ \forall\, n \in \mathbb{N},\ z_{n+1} = z_n^2 + c \end{cases}$$

L'ensemble de Mandelbrot $\mathcal{M}$ est l'ensemble des complexes c tels que la suite (z_n) associée est bornée. On peut montrer que $\mathcal{M}$ est mesurable et on cherche à estimer son aire par la méthode de Monte-Carlo.

On peut montrer facilement que $\mathcal{M}$ est inclus dans le disque $\mathcal{C}$ de centre 0 et de rayon 2. Donc l'intégrale à estimer est :

$$I = \int_{\mathcal{C}} 1_{\mathcal{M}}(x+iy)\ dxdy = \int_{\rho=0}^{2} \int_{\theta=0}^{2\pi} 1_{\mathcal{M}}(\rho e^{i\theta})\ \rho\ d\rho d\theta.$$

(pourquoi cette dernière égalité ?)

Écrivez une fonction qui retourne 1 si le point d'affixe $x + iy$ appartient à $\mathcal{M}$ et 0 sinon. On se limitera à vérifier que la suite des (z_n) reste dans le disque de rayon 2 pour $n < 1000$ par exemple.
La fonction sera du type : `int Mandel(double x, double y)`

Écrivez deux fonctions, l'une simulant une variable aléatoire uniforme sur $[0, 2]$ et l'autre simulant une variable aléatoire uniforme sur $[0, 2\pi]$.

En utilisant la fonction `Mandel` sur le point d'affixe $\rho\cos(\theta) + i\rho\sin(\theta)$, donnez une estimation de l'aire de l'ensemble de Mandelbrot à l'aide de la méthode de Monte-Carlo. (*réponse* : l'approximation à 10^{-2} de l'aire est 1.50)

Autre méthode : on peut aussi tirer les points uniformément sur le carré $[-2, 2] \times [-2, 2]$ (qui contient $\mathcal{C}$).

---

3 Pricing d'options européennes

Les *options* sont des produits financiers. Une option d'achat (*call*) donne le droit à son détenteur d'acheter à une date future un actif (par exemple une certaine quantité d'actions) à un certain prix (dit *prix d'exercice*, également fixé au moment de l'achat). Les options considérées dans cet exercice sont dites *européennes* : la date à laquelle le droit de vente peut s'exercer est aussi fixée au moment de l'achat.

Nous nous plaçons dans le cadre du modèle de Black et Scholes, où le marché se réduit à deux actifs, l'un sans risque (avec un taux d'intérêt r), et l'autre risqué sur lequel est basée l'option. Le cours S_t de l'actif risqué suit un *mouvement brownien géométrique*. S_t est alors décrit par la variable aléatoire :

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

où : σ est la *volatilité* du prix de l'action, et B est un mouvement brownien (par suite de quoi la loi de $B(t)$ est normale de moyenne nulle et de variance t). Le terme de gauche dans l'exponentielle décrit une tendance, le terme de droite les fluctuations aléatoires. Qualitativement, si la volatilité est très faible (et donc le « risque » aussi), S_t évolue comme l'actif non-risqué, alors que si la volatilité est élevée, le terme de gauche est négatif (tendance à la baisse), mais il est contrebalancé par la composante aléatoire qui permet d'« espérer » des gains significativement plus important que l'actif non-risqué.

Une explication intuitive du fait que l'évolution de S_t dépende de r est que les investisseurs ont dans ce modèle une stratégie favorisant le placement sans risque lorsque sa rentabilité est supérieure à celle de l'actif risqué, et inversement.

On suppose que l'option est achetée au temps $t = 0$ et qu'elle s'exerce au temps $t = T$, à un prix d'exercice K. Le gain est alors $\max(S_T - K, 0)$ (si le cours de l'actif est supérieur au prix d'exercice alors l'option est exercée, sinon on ne fait rien car on perdrait de l'argent). L'espérance de gain est donc : $\mathbf{E}\left(\max(S_T - K, 0)\right)$.

La question que l'on cherche à résoudre est : comment fixer le prix P de l'option à la date $t = 0$? Pour que celui-ci soit « juste » pour l'acheteur et le vendeur, il faut que le gain espéré à la date $t = T$ soit égal au gain que procurerait la somme d'argent P investie sur le placement sans risque, *i.e.* $Pe^{rT} = \mathbf{E}\left(\max(S_T - K, 0)\right)$.
D'où :

$$P = \mathbf{E}\left(\max\left(S_0 e^{-\frac{\sigma^2}{2}T+\sigma B(T)} - Ke^{-rT}\ , 0\right)\right).$$

Il existe une formule explicite[2] permettant de calculer P dans ce cas de figure (connaissant bien sûr S_0, σ, K, et r).

Par contre nous ne disposons pas d'une telle formule dans le cas suivant. Imaginons cette fois que l'actif sous-jacent est un *panier* de N actions, qui évoluent toutes de manière indépendante[3] selon un mouvement brownien géométrique. Le prix de l'option sera alors fixé selon :

$$P = \mathbf{E}\left(\max\left(\sum_{i=1}^{N} S_0^i e^{-\frac{\sigma_i^2}{2}T+\sigma_i B_i(T)} - Ke^{-rT}\ ,0\right)\right),$$

où S_0^i et σ_i sont respectivement la prix initial de l'action i et sa volatilité propre, et où les B_i sont des browniens indépendants.

Exercice 3 (simulation d'un mouvement brownien)

Voici une définition pour le *mouvement brownien*.

Le processus (i.e. la famille de variables aléatoires) $(B_t)_{t>0}$ est un mouvement brownien si :

- $B_0 = 0$ p.s.,
- les réalisations $t \mapsto B_t(\omega)$ sont continues (presque sûrement),
- les variables aléatoires $B_{t+h} - B_t$ sont indépendantes de loi gaussienne centrée de variance h.

On déduit de cette définition la construction suivante, qui permet de simuler les valeurs d'un mouvement brownien en des points $t_i = iT/L$ ($i \in \{0, \dots, L\}$:

- $B_0 = 0$
- $B_{t_i} = B_{t_{i-1}} + \sqrt{\frac{T}{L}}Y_i$, où les Y_i sont des variables aléatoires indépendantes identiquements distribuées selon la loi normale centrée réduite.

Écrivez une fonction (de T et L) générant les B_{t_i}.

Affichez diverses réalisations à l'aide de Gnuplot.

Simulez (et affichez) le cours d'un actif risqué suivant le modèle de Black et Scholes (écrivez une fonction).

Simulez le cours d'un panier d'actifs avec les valeurs numériques de l'exercice suivant.

Exercice 4 (pricing d'une option sur un panier d'actifs)

À l'aide de la méthode de Monte-Carlo[4], déterminez un intervalle de confiance à 95% du prix de vente de l'option sur panier dans le cas : $N = 5$, $S_0^i = \{10, 20, 20, 30, 50\}$, $\sigma_i = \{0.3, 0.3, 0.1, 0.5, 0.1\}$, $K = 100$, $r = 0.05$, $T = 10$.

Remarque : il n'est pas nécessaire d'utiliser les simulations de l'exercice précédent : $B_i(T) \sim \mathcal{N}(0, \sqrt{T})$.

---

[2] Le succès du modèle de Black et Scholes vient certainement de sa simplicité et du fait qu'il permet de faire des calculs facilement…Certains scientifiques (dont B. Mandelbrot) pensent que la crise financière de 2008 est due à la mauvaise utilisation de ce modèle et au fait qu'il conduise à minorer les risques.

[3] Cette hypothèse est-elle réaliste ?

[4] Ici, qui joue le rôle de f ?