

# Modélisation et résolutions numérique et symbolique de problèmes via les logiciels Maple et MATLAB (MODEL)

## Cours n°10 : Méthode de Monte Carlo

Stef Graillat

Université Pierre et Marie Curie (Paris 6)



## Résumé du cours précédent

### Transformée de Fourier discrète :

- Algorithme FFT : version récursive et itérative
- Complexité en  $\mathcal{O}(n \log n)$
- Applications à la multiplication de polynômes, au traitement d'image et au traitement du signal

- ① Rappels élémentaires de probabilité et de statistique
- ② Présentation des méthodes de Monte Carlo
- ③ Calcul d'intégrale via les méthodes de Monte Carlo

# Plan du cours

- ① Statistique de base : nombre aléatoire et génération
- ② Méthode de Monte Carlo et calcul d'intégrales
- ③ Méthode de Monte Carlo et optimisation
- ④ Méthode de Monte Carlo et comptage

- Scientific Computing with Case Studies, Dianne P. O'Leary, SIAM, 2009
- Explorations in Monte Carlo Methods, R.W. Shonkwiler et F. Mendivil, Springer, 2009
- Monte Carlo Methods, M.H. Kalos et P.A. Whitlock, 2nd édition, Wiley-VCH, 2008
- Monte Carlo Strategies in Scientific Computing, Jun S. Liu, Springer, 2001

## Qu'est-ce qu'une méthode de Monte Carlo

### Définition 1

*On appelle **méthode de Monte-Carlo** toute méthode visant à calculer une valeur numérique en utilisant des procédés aléatoires.*

L'aléatoire se fait par des **tirages aléatoires**

Le nom de ces méthodes, qui fait allusion aux jeux de hasard pratiqués à Monte-Carlo, a été inventé en 1947 par Nicholas Metropolis

### Exemples :

- Jet d'un dé à 6 faces. Après 120 jets, on a une bonne estimation de la probabilité qu'apparaisse un chiffre.
- une « boîte noire » prend un nombre dans  $[0, 1]$  et renvoie un nombre dans  $[0, 1]$ . Si on envoie  $m$  entrées et que l'on étudie les  $m$  sorties, on peut utiliser la moyenne des sorties pour avoir une approximation de la moyenne du processus défini par la boîte noire

Le développement des méthodes de Monte-Carlo s'est effectué lors de la Seconde Guerre mondiale pour les recherches sur la fabrication de la bombe atomique dans le cadre du « Projet Manhattan ».

Utilisation de ces méthodes probabilistes pour résoudre des équations aux dérivées partielles dans le cadre du « Monte-Carlo N-Particle » [MCNP].

## Quelques applications actuelles :

- Calcul d'intégrales multiples
- Filtrage (filtre particulaire) : application en traitement du signal, imagerie, apprentissage, etc.
- Finance : calcul du prix de produits dérivés (option)
- Bio-informatique : détermination d'un arbre phylogénétique à partir de données génomiques
- Raytracing

## Deux principes de base

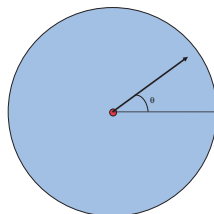
- Il y a une différence importante entre
  - les **méthodes de Monte Carlo** qui utilisent des données aléatoires
  - les **pseudo-méthodes de Monte Carlo** qui utilisent des données pseudo-aléatoires

On peut dire qu'en pratique, tous les méthodes sont des pseudo-méthodes puisque les nombres aléatoires générés sur un ordinateur ne sont pas vraiment aléatoires !

- Les méthodes de Monte Carlo sont des méthodes de dernier recours. Elles sont en général **coûteuses**. On ne les applique généralement que pour des problèmes difficiles à résoudre par des méthodes **déterministes**

## Exemple de génération de nombres aléatoires

- Mettre des boules numérotées de 1 à  $n$  dans une urne et tirer une boule (en remettant la boule dans l'urne ensuite). On obtient un tirage aléatoire dans  $\{1, 2, \dots, n\}$  uniformément distribué
- Tourner la roue! loi uniforme sur  $[0, 2\pi]$



- Si, en moyenne, un élément radioactif émet des particules  $\alpha$  toutes les  $\mu$  secondes alors le temps entre deux émissions successives suit une loi exponentielle
- La loi normale est un bon modèle dans de nombreuses situations
  - caractéristique physique des plantes et des animaux (hauteurs, poids, etc)
  - loi de distribution de vitesses des molécules en équilibre thermodynamique (distribution de Maxwell-Boltzmann)

## Propriétés d'un échantillon

### Définition 2 (Moyenne)

La *moyenne* d'un échantillon  $\{x_i\}$ ,  $i = 1, \dots, n$  est définie par

$$\mu_n = \frac{1}{n} \sum_{i=1}^n x_i$$

### Définition 3 (Variance)

La *variance* d'un échantillon  $\{x_i\}$ ,  $i = 1, \dots, n$  est définie par

$$\sigma_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_n)^2$$

Soit  $X$  une variable aléatoire réelle.

### Définition 4 (densité)

On dit que  $X$  admet une densité de probabilité  $f : \Omega \rightarrow \mathbb{R}^+$  avec  $\Omega \subset \mathbb{R}$  si

$$\int_{\Omega} f(x) dx = 1$$

et

$$P(X \in B) = \int_B f(x) dx \quad \text{pour } B \subset \Omega$$

Soit  $X$  une variable aléatoire réelle de densité de probabilité  $f$

### Définition 5 (Moyenne et variance)

La moyenne (on dit aussi l'espérance)  $\mu$  et la variance  $\sigma^2$  de  $X$  sont définis par

$$\begin{aligned}\mu &= \int_{\Omega} xf(x) dx \\ \sigma^2 &= \int_{\Omega} (x - \mu)^2 f(x) dx\end{aligned}$$

On note aussi  $E[X]$  pour  $\mu$  et  $\text{Var}(X)$  pour  $\sigma^2$

## Exemple de fonctions de distribution

- Loi uniforme sur l'intervalle  $[0, m]$  :

$$\begin{aligned}f(x) &= \frac{1}{m} \\ \mu &= \int_0^m \frac{x}{m} dx = \frac{m}{2} \\ \sigma^2 &= \int_0^m \frac{1}{m} \left(x - \frac{m}{2}\right)^2 dx = \frac{m^2}{12}\end{aligned}$$

- Loi exponentielle de paramètre  $\mu$  sur l'intervalle  $[0, +\infty[$  :

$$\begin{aligned}f(x) &= \frac{1}{\mu} e^{-x/\mu} \\ \mu &= \mu \\ \sigma^2 &= \mu^2\end{aligned}$$

## Exemple de fonctions de distribution (suite)

- Loi normale de paramètres  $\mu$  et  $\sigma$  sur  $]-\infty, +\infty[$  :

$$\begin{aligned}f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \\ \mu &= \mu \\ \sigma^2 &= \sigma^2\end{aligned}$$

## Théorème centrale limite

Si l'on tire  $n$  variables aléatoires  $X_i$  indépendantes suivant la même loi et définies sur le même espace de probabilité, et que l'on note  $\bar{x}_n$  la moyenne des tirages ( $\bar{x}_n = \frac{1}{n}(X_1 + \dots + X_n)$ ) alors la variable aléatoire  $\bar{x}_n$  approche une **loi normale**.

Si l'on pose de plus

$$y_n = \sqrt{n} \frac{\bar{x}_n - \mu}{\sigma},$$

alors la variable aléatoire  $y_n$  approche la **loi normale de moyenne 0 et de variance 1** pour  $n$  suffisamment grand.

## Génération de nombres aléatoires

- On pourrait générer des nombres aléatoires via les méthodes décrites précédemment
- En pratique, il est difficile de générer des nombres aléatoires sur un ordinateur
- Les nombres aléatoires générés sur un ordinateur le sont de façon déterministe : il s'agit de nombres pseudo-aléatoires (mais ils semblent aléatoires)
- En fait, les nombres pseudo-aléatoires font un cycle : si on regarde une suite suffisamment longue de nombres, on verra qu'elle est périodique

- Il est peu coûteux de générer des nombres pseudo-aléatoires uniformément distribués
- La simulation des autres lois de probabilité se fait généralement à partir de plusieurs nombres pseudo-aléatoires uniformément distribués
- En Matlab, il y a la commande `rand` pour générer des nombres pseudo-aléatoires uniformément distribués et `randn` pour générer des nombres pseudo-aléatoires suivant la loi normale

On sait maintenant qu'il existe des outils pour générer des nombres pseudo-aléatoires

À partir de maintenant on abandonnera le préfixe pseudo pour parler de nombres aléatoires

On va passer de ce qu'est un nombre aléatoire à comment on peut les utiliser pour résoudre des problèmes

## Méthode de Monte Carlo pour le calcul d'intégrale multiple

**Problème** : on veut estimer la quantité

$$I = \int_0^1 \cdots \int_0^1 f(x_1, \dots, x_{10}) p(x_1, \dots, x_{10}) dx_1 \cdots dx_{10} = \int_{\Omega} f(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}$$

Notation :

- $\mathbf{x} = [x_1, \dots, x_{10}]$
- $\Omega = [0, 1] \times \cdots \times [0, 1]$
- en général  $p(\mathbf{x})$  est une constante égale à 1 divisé par le volume de  $\Omega$  mais  $p$  peut être plus général.  
Il faut que  $p$  soit positif et que

$$\int_{\Omega} p(\mathbf{x}) d\mathbf{x} = 1$$

→ comment calculer une approximation de  $I$  ?

## Méthode n°1 : interpolation

**Idée** : interpoler  $f(\mathbf{x})p(\mathbf{x})$  par un polynôme et intégrer ensuite ce polynôme.

Par exemple, un polynôme de degré au plus 2 en chaque variable aura des monômes de la forme

$$x_1^{[]} x_2^{[]} x_3^{[]} x_4^{[]} x_5^{[]} x_6^{[]} x_7^{[]} x_8^{[]} x_9^{[]} x_{10}^{[]}$$

où chaque boîte  $[]$  contient 0,1 ou 2. Il y a donc  $3^{10} = 59049$  coefficients et donc 59049 évaluations à effectuer.

Il faut approximer la fonction par un polynôme sur une petite « boîte » si l'on veut avoir une bonne approximation.

Si l'on divise chaque intervalle  $[0, 1]$  en 5 segments, cela fait  $5^{10}$  boîtes avec 59049 évaluations à effectuer sur chaque boîte!!!!

Cette méthode est **très coûteuse** !

## Méthode n°2 : séparation des intégrales

Il se peut que l'on puisse séparer la fonction  $f(\mathbf{x})p(\mathbf{x})$  en

$$f(\mathbf{x})p(\mathbf{x}) \approx f_1(x_1)f_2(x_2) \cdots f_{10}(x_{10})$$

Dans ce cas, on peut approximer l'intégrale par

$$I \approx \int_0^1 f_1(x_1) dx_1 \cdots \int_0^1 f_{10}(x_{10}) dx_{10}$$

Méthode efficace mais il est en général rare que l'on puisse séparer  $f(\mathbf{x})p(\mathbf{x})$

Idée :

- Générer  $n$  points  $\{z^{(i)}\}$  aléatoirement distribués suivant la loi de densité  $p$

$$\mu_n = \frac{1}{n} \sum_{i=1}^n f(z^{(i)})$$

est une approximation de la valeur moyenne de  $f$  dans la région  $\Omega$  et donc

$$I \approx \mu_n \int_{\Omega} p(\mathbf{x}) dx_1 \cdots dx_{10} = \mu_n$$

# Méthode de Monte Carlo

Les méthodes de Monte Carlo reposent sur la **loi forte des grands nombres** :

Soit une suite  $(X_n)_{n \in \mathbb{N}}$  de variables aléatoires indépendantes qui suivent la même loi de probabilité, alors

$$\frac{X_1 + \cdots + X_n}{n} \rightarrow E[X_1] \text{ quand } n \rightarrow +\infty$$

Si  $X$  est une variable aléatoire de loi de densité  $p$  et si  $f : \mathbb{R} \rightarrow \mathbb{R}$  alors

$$E[f(X)] = \int_{\Omega} f(x)p(x)dx$$

D'ou si  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires indépendantes qui suivent la même loi de probabilité  $p$ , alors

$$\frac{f(X_1) + \cdots + f(X_n)}{n} \rightarrow E[f(X_1)] = \int_{\Omega} f(x)p(x)dx \text{ quand } n \rightarrow +\infty$$

- L'espérance de l'estimateur  $\mu_n$  est égale à la vraie valeur de l'intégrale. On parle d'estimateur sans-biais.
- Pour  $n$  suffisamment grand (via le Théorème centrale limite), la loi de l'estimateur approche une loi normale de variance  $\sigma^2/n$  où

$$\sigma^2 = \int_{\Omega} (f(\mathbf{x}) - \mu_n)^2 p(\mathbf{x}) d\mathbf{x}$$

- En effet, on a  $\mu_n - I \approx \frac{\sigma}{\sqrt{n}}Z$  où  $Z$  suit une loi normale centrée réduite (d'espérance 0 et de variance 1)
- On note que la variance est une constante indépendante de la dimension de l'intégrale

## Réduction de la variance : échantillonnage préférentiel

Supposons que l'on veuille estimer

$$I = \int_{\Omega} f(\mathbf{x}) d\mathbf{x}$$

où  $\Omega$  est une région de  $\mathbb{R}^{10}$  de volume 1

**Méthode 1** : on utilise une méthode de Monte Carlo en prenant une **distribution uniforme** sur  $\Omega$

**Méthode 2** : on choisit une fonction  $p(\mathbf{x}) > 0$  pour tout  $\mathbf{x} \in \Omega$  telle que

$$\int_{\Omega} p(\mathbf{x}) d\mathbf{x} = 1$$

On a alors

$$I = \int_{\Omega} \frac{f(\mathbf{x})}{p(\mathbf{x})} p(\mathbf{x}) d\mathbf{x}$$

Quand la méthode 2 est-elle meilleure que la méthode 1 ?

On rappelle que la variance de l'estimateur est de l'ordre de

$$\sigma^2 = \int_{\Omega} \left( \frac{f(x)}{p(x)} - I \right)^2 p(x) dx$$

avec

$$I = \int_{\Omega} f(x) dx$$

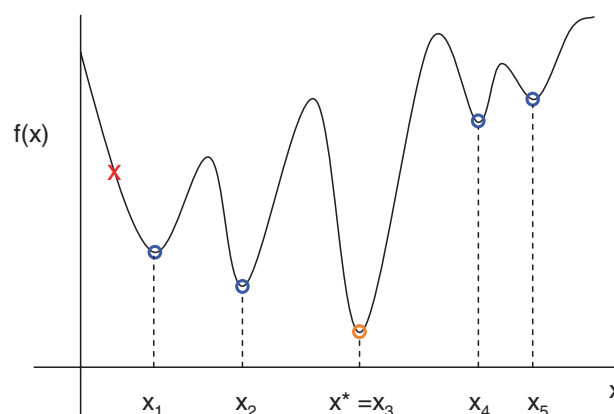
Si on choisit  $p$  de façon à ce que  $f(x)/p(x)$  soit constante, alors  $\sigma$  devient proche de 0 !

On essaye donc d'obtenir une fonction de densité  $p$  proche de :

$$f/I.$$

## Méthode de Monte Carlo pour l'optimisation continue

- Les méthodes de Monte Carlo fournissent un bon moyen pour générer des points initiaux dans les méthodes d'optimisation de fonctions non-convexes
- Un point initial peut entraîner l'algorithme à converger vers un minimum local plutôt que vers un minimum global
- Si l'on prend un ensemble aléatoire de points sur le domaine de la fonction qu'on essaie de minimiser alors on augmente la probabilité de trouver un minimum global



Une **option financière** est un produit dérivé qui donne, à l'acheteur, le droit :

- d'acheter (option d'achat, appelée aussi *call*),
- ou de vendre (option de vente, appelée aussi *put*),

une quantité donnée d'un actif financier (action, obligation, indice boursier, devise, matière première, etc.), appelé **actif sous-jacent**

- à un prix (en général) précisé à l'avance (**prix d'exercice** ou *strike* en anglais),
- à une date d'échéance donnée (option dite **européenne**),
- ou durant toute la période jusqu'à échéance (option dite **américaine**).

Ce droit lui-même se négocie, sur un marché d'options spécialisé (géré par une bourse, ou au gré à gré), contre un certain prix, appelé **prime**.

## Introduction aux produits dérivés en finance (suite)

Un **produit dérivé** est un instrument financier :

- dont la valeur fluctue en fonction de l'évolution du taux ou du prix d'un produit appelé sous-jacent ;
- qui ne requiert aucun placement net initial ou peu significatif ;
- dont le règlement s'effectue à une date future.

Il s'agit d'un contrat entre deux parties, un acheteur et un vendeur, qui fixe des flux financiers futurs fondés sur ceux d'un actif sous-jacent, réel ou théorique, généralement financier.

Les options décrites précédemment sont appelées des **options vanilles**

Il existe des options plus complexes dites **options exotiques**

On peut par exemple citer les **options asiatiques** :

- la valeur à l'échéance d'une option asiatique découle du prix moyen du sous-jacent dans un intervalle de temps déterminé. Elle est inférieure à celle d'une option vanille car la valeur moyenne d'un sous-jacent est moins volatile que sa valeur à un instant donné.

## Utilisation des options

Les options peuvent être utilisées :

- en **couverture de risque** de baisse ou hausse du prix du sous-jacent (par exemple un producteur de pétrole peut choisir d'acheter des puts afin de se prémunir d'une baisse trop importante des cours),
- pour **spéculer** à la baisse ou à la hausse du sous-jacent (c'est en ce sens qu'elles sont distribuées comme rémunération sous le nom de stock options),
- pour **spéculer** sur la volatilité.

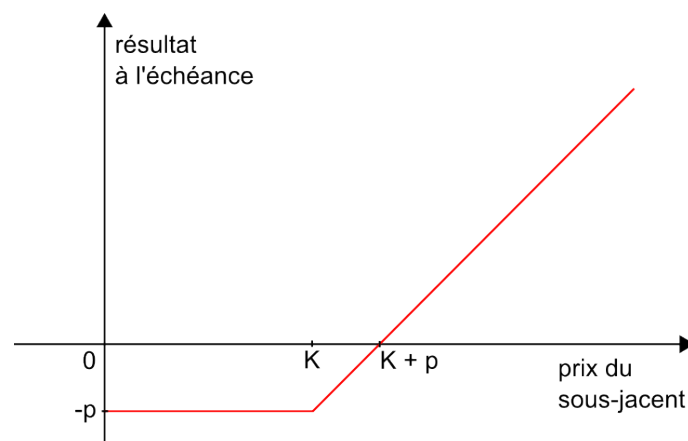
En l'absence d'une couverture spécifique et dans le cas le plus défavorable, l'acheteur d'une option aura une perte limitée à la prime qu'il aura payée. Son gain maximum théorique est en revanche illimité (ou limité au prix d'exercice diminué de la prime pour un put dont le sous-jacent ne peut avoir un prix négatif).

Symétriquement, le vendeur d'une option voit son gain maximum limité à la prime qu'il reçoit. Sa perte peut être illimitée ou limitée (vendeur d'un put dont le prix du sous-jacent ne peut être négatif). Il s'agit d'une stratégie spéculative très risquée.

Si l'option n'a pas été exercée à la date d'échéance, elle est dite abandonnée

## Prix d'une option d'achat (call)

- $K$  : le prix d'exercice de l'option
- $S$  : le prix du sous-jacent
- $p$  : la prime de l'option
- $R$  : le résultat à l'échéance



## Prix d'une option d'achat (call)

On note  $(X)_+ = \max(0, X)$

- Notons  $(S_t)$  le prix (aléatoire) de l'action (sous-jacent) considérée à l'instant  $t$ .
- Si  $S_T > K$ , alors le détenteur du call exerce son droit, et achète l'action au prix  $K$  et la revend tout de suite au prix  $S_T$ ; son gain est alors de  $R = (S_T - K) - p = (S_T - K)_+ - p$
- Si, par contre,  $S_T < K$ , alors le détenteur n'exerce pas son droit (il n'achète pas l'action), et son gain est alors de  $R = -p = (S_T - K)_+ - p$ .

Dans les deux cas, le gain est en moyenne de  $E[(S_T - K)_+] - p$  où  $E[\cdot]$  dénote l'espérance de la variable aléatoire.

## Prix d'une option d'achat (call)

- Pour que le problème soit équitable, il faut que l'espérance du gain soit nul donc que  $p = E[(S_T - K)_+]$ .
- Il reste à modéliser le prix des actions. On utilise le modèle suivant : entre les instants  $t_i$  et  $t_{i+1}$  le prix de l'action varie de

$$S_{t_{i+1}} - S_{t_i} = Y_i S_{t_i}$$

où les  $Y_i$  sont des variables iid de loi normale centrée  $\mathcal{N}(0, 1)$ .

- Le théorème central limite permet de déduire que  $S_T \approx e^Z$  avec  $Z$  de loi  $\mathcal{N}(0, 1)$ . La formule précédente a été introduite par Black et Scholes ce qui leur a valu le prix Nobel d'économie.



Fischer Black (1938 - 1995)



Myron Scholes (1941 - )

## Conclusion

- Les méthodes de Monte Carlo sont des méthodes de dernier recours
- Elles nécessitent la génération de nombres aléatoires ayant de bonnes propriétés
- Elles sont utilisées dans de nombreux domaines