MÉTHODE DE MONTE-CARLO
EXERCICES DU 3 MARS 2015

par Rémi Peyre

EXERCICE 1 — Entrainement aux variables antithétiques

Le sujet de cet exercice est à propos d'une marine militaire qui voudrait estimer l'efficacité pratique de son sonar pour détecter des navires ennemis tentant de traverser le bras de mer qui les sépare le pays ennemi du sien. (Nous nous placerons dans la suite du point de vue des ressortissants du pays au service duquel est cette marine, et dirons donc « notre » pays pour ce dernier). On fera les hypothèses suivantes [*] *:*

- *La mer est un rectangle de* 100 km *de large sur* 200 km *de long, disposé horizontalement (quand on place le nord en haut) ; notre pays en occupe la côte sud, et l'ennemi la côte nord. Les coordonnées des points sont données dans un repère orthonormal dont l'axe des abscisses pointe vers l'est et l'axe des ordonnées vers le nord, dont l'origine est située à l'extrêmité ouest de la côte sud.*
- *Lorsqu'un navire ennemi tente la traversée, afin qu'on ne puisse pas anticiper sa trajectoire, il choisit son point de départ uniformément sur la côte nord, son point d'arrivée uniformément sur la côte sud, et navigue en ligne droite entre les deux.*
- *Notre armée dispose d'un unique sonar, qui détecte un navire dès lors que ce dernier passe à moins de* 40 km *de lui.*
- *Afin que l'ennemi ne puisse pas anticiper la position de notre sonar, celle-ci est changée chaque jour, en la choisissant à chaque fois de façon uniforme sur l'étendue de la mer.*

1. Écrire une fonction `detection` qui prend en arguments le point de départ d'un navire, son point d'arrivée, et la position du sonar (ces trois positions étant données sous la forme du 2-vecteur[-ligne] de leurs coordonnées), et qui renvoie `true` si le segment joignant le point de départ au point d'arrivée passse à moins de `RAYON` du sonar. La variable `RAYON` sera déclarée dans le code comme une *variable globale* définie dans une fonction-mère [†].

Indication : On pourra utiliser le paramétrage naturel du segment reliant les points de départ et d'arrivée par l'intervalle $[0, 1]$. Si on considère le *carré* de la distance au sonar, celui-ci évolue alors comme un polynome de degré 2, dont il est facile d'étudier les variations pour déterminer s'il passe au-dessous du carré de `RAYON`.

Indication : On prendra bien garde à un piège : il s'agit de savoir si le *segment* parcouru par le navire passe à portée de sonar, et non pas la *droite* qui supporte ce segment... !

2. Estimer par la méthode de Monte-Carlo la probabilité que le sonar détecte un navire tentant une traversée.

On définit dans la suite de l'exercice une fonction position_ conjuguee, qui va de l'étendue de la mer dans elle-même, de la façon suivante :

---

[*]. Précisons que ce type d'usage de la méthode de Monte-Carlo dans des situations militaires, où on est obligé de *simuler* la stratégie supposée de l'ennemi, est une application tout ce qu'il y a de plus véritable, quoique l'exemple présenté ici soit simplifié à outrance.

[†]. Ce n'est donc pas un *argument* de la fonction `interception`... !

– *Si le point x est situé dans la moitié* ouest *de la mer*[‡], *position_conjuguee*(x) *sera le translaté de x de* 100 *km vers l'*est. *(Cette valeur de* 100 km *correspondant à la moitié de la longueur de la mer).*
– *Si le point x est situé dans la moitié* est *de la mer, position_conjuguee*(x) *sera le translaté de x de* 100 *km vers l'*ouest.

3. Écrire une fonction `conjugaison` qui prend en argument un point x dans la mer (donné sous la forme du 2-vecteur de ses coordonnées) et qui renvoie *position_conjuguee*(x). Les largeur et longueur de la mer seront déclarées (s'il y a lieu) sous la forme de variables globales définies dans une fonction-mère, plutôt que par leurs valeurs numériques.

4. Justifier (sommairement) que la mesure-image de la probabilité uniforme sur l'étendue de la mer par *position_conjuguee* est à nouveau la probabilité uniforme sur l'étendue de la mer. En déduire qui si on garde les mêmes positions de départ et d'arrivée pour le navire ennemi, et qu'on remplace la position de notre sonar par sa position conjuguée, on se retrouve exactement dans la même situation probabiliste qu'au départ, et que par conséquent la probabilité d'interception du navire ennemi reste exactement la même.

5. Expliquer intuitivement pourquoi on peut légitimement s'attendre à ce que, si on compare l'évènement de détection du navire ennemi quand le sonar est en position x avec l'élèvenement de détection quand le sonar est en *position_conjuguee*(x), ces deux évènements seront vraisemblablement négativement corrélés : autrement dit, lorsqu'on a détecté le navire avec le sonar en x, la probabilité de le détecter si on avait placé le sonar en *position_conjuguee*(x) est moindre que lorsqu'on n'a pas détecté le navire avec le sonar en x.

6. Implémenter la technique de variables antithétiques suggérée par les considérations ci-dessus.
Indication : On procèdera donc par couples de simulations ; dans chacun de ces couples, la trajectoire du navire ennemi sera la même, mais la position du sonar dans la première simulation sera remplacée par sa position conjuguée dans la seconde.

7. Chronométer les méthodes de Monte-Carlo pour comparer leurs efficacité avec et sans technique de variables antithétiques, et conclure quant à l'utilité de cette technique pour notre problème[§].

EXERCICE 2 — Entrainement aux variables de contrôle

On cherche dans cet exercice à déterminer avec précision la valeur de $\mathbb{E}(\log|G|)$ [¶] *lorsque G suit une loi normale standard, en utilisant une méthode de Monte-Carlo.*

1. Vérifier empiriquement, par une estimation de covariance, que $\log|G|$ et $\frac{1}{2}G^2$ sont positivement corrélées.

2. **(✿)** Que vaut $\mathbb{E}(\frac{1}{2}G^2)$? (exactement s'entend).

[‡]. La définition de *position_conjuguee* est ambigüe lorsqu'on est pile au milieu de la mer, mais c'est sans importance ici.

[§]. Bien sûr, il n'y a aucun intérêt à déterminer la probabilité de détection du navire ennemi avec une précision extrême ! En revanche, une situation réaliste serait qu'on aurait à estimer *beaucoup* de probabilités différentes pour diverses situations similaires à celle-ci (correspondant par exemple à des stratégies différentes pour la trajectoire du navire ennemi et la disposition du sonar), auquel cas l'ensemble des calculs serait long (même pour une précision modérée), et l'utilisation d'une technique de réduction de la variance alors bienvenue.

[¶]. « log » désigne ici le logarithme népérien.

3. Estimer $\mathbb{E}(\log|G| - \frac{1}{2}G^2)$ par la méthode de Monte-Carlo ; en déduire un estimateur de $\mathbb{E}(\log|G|)$ (en précisant la marge d'erreur). Comparer comment se comporte la précision de cet estimateur par rapport à l'estimateur naïf de $\mathbb{E}(\log|G|)$ (à nombre de simulations égal), et vérifier que l'efficacité *à nombre de simulations fixé* est améliorée comme on pouvait s'y attendre. Chronométrer les deux calculs et en déduire le gain en efficacité *à temps de calcul fixé.*

4. En s'inspirant du programme `mathfi_ctrl1` du polycopié, utiliser la version paramétrée de la technique de la variable de contrôle pour déterminer $\mathbb{E}(|G|)$, la variable de contrôle utilisée étant G^2. Nous savons qu'une telle méthode est essentiellement équivalente à passer par l'évaluation naïve de $\mathbb{E}(|G| - aG^2)$ pour une certaine valeur a ; quelle est la valeur a utilisée par votre programme en l'occurrence ? Si cette valeur était égale à $1/2$, l'efficacité par simulation de notre technique paramétrée serait évidemment la même qu'avec la technique non paramétrée ; vu que a est ici différent de $1/2$, vérifier que comme prévu l'efficacité (à nombre de simulations fixé) a été strictement améliorée par rapport à la version non paramétrée.

5. Montrer que $\mathbb{E}(|G|) = \sqrt{2/\pi}$.
Indication : Utiliser la densité de la loi normale standard pour écrire cette espérance comme une intégrale sur $\mathbb{R}$, et calculer cette intégrale par les moyens analytiques habituels — en utilisant que $|x|$ vaut $-x$ lorsque $x \leqslant 0$ et x lorsque $x > 0$.

6. En s'inspirant du programme `mathfi_ctrl3` du polycopié, déterminer $\mathbb{E}(|G|)$ en utilisant *à la fois* les variables de contrôle G^2 et $|G|$. Nous savons qu'une telle méthode est essentiellement équivalente à passer par l'évaluation naïve de $\mathbb{E}(|G| - a'G^2 - b'|G|)$ pour certaines valeurs a' et b' ; quelles sont les valeurs a' et b' utilisées par votre programme en l'occurrence ? La valeur a' est-elle égale (ou sensiblement égale) au a de la question précédente ; pouvait-on le prévoir ? Si on avait $(a', b') = (a, 0)$, l'efficacité par simulation de notre technique à deux variables de contrôle serait évidemment la même qu'avec seulement la variable de contrôle G^2 ; vu que ce n'est pas le cas ici, vérifier que comme prévu l'efficacité (à nombre de simulations donnée) a été strictement améliorée par rapport à la version à une seule variable de contrôle.

7. **($\bigstar$)** On pourrait aussi utiliser G comme variable de contrôle, dans la mesure où on sait que $\mathbb{E}(G) = 0$. Pourquoi cela ne serait-il pas pertinent ici ? Qu'observerait-on si on utilisait quand même G comme variable de contrôle ? (en plus de G^2 et $|G|$, s'entend).

EXERCICE 3 — Entrainement à la stratification "à postériori"

Une compagnie d'assurances couvre 10 000 *clients contre l'incendie de leurs maisons. Au cours de l'année, chacune de ces maisons a, indépendamment, une (mal)chance sur* 1 000 *de bruler. Néanmoins, toutes les maisons assurées n'ont pas la même valeur : ainsi, la maison du client numéro i est assurée pour une valeur de* $(10\ 000\ /\ i)^{0{,}7} \times 100\ 000$ *€. Au début de l'année, grâce aux primes versées par les clients, l'assureur a provisionné* 5 000 000 *€. L'assureur souhaite évaluer son risque que ces primes ne suffisent pas à couvrir l'ensemble des sinistres qu'il aura à indemniser dans l'année.*

1. Écrire une fonction `simulation` qui simule la survenue des sinistres chez chacun des clients au cours de l'année : cette fonction ne prendra aucun argument et renverra le montant total des sinistres survenus. À l'aide de cette fonction, estimer le risque encouru par l'assureur par une méthode de Monte-Carlo naïve.

2. Modifier la fonction `simulation` en une fonction `simulation_` qui, outre le montant total des sinistres survenus, renvoie aussi le montant du pire sinistre de l'année.

L'assureur distingue trois catégories de sinistres : les petits *sinistres, qui concernent un montant de moins de* 1 000 000 €*, les* gros *sinistres, qui concernent un montant compris entre* 1 000 000 € *et* 2 500 000 €*, et les* très gros *sinistres, qui concernent un montant supérieur à* 2 500 000 €*.*

3. Démontrer que la probabilité que le pire sinistre qui survient dans l'année soit un petit, resp. un gros, resp. un très gros sinistre, vaut exactement :

$$(1-1\,000^{-1})^{372}, \qquad \text{resp.} \quad (1-1\,000^{-1})^{100} - (1-1\,000^{-1})^{372}, \qquad \text{resp.} \quad 1-(1-1\,000^{-1})^{100}.$$

4. À l'aide du calcul précédent et de la fonction `simulation_`, modifier la méthode de Monte-Carlo de la question 1 pour y implémenter une technique de stratification, les strates de cette technique correspondant au type du pire sinistre survenu dans l'année.

EXERCICE 4 — Entrainement à l'allocation interstrates optimale

Dans cet exercice, on considère un produit financier structuré comme un panier d'actifs *: il s'agit de plusieurs actifs financiers complètement différents (par exemple un kilo d'uranium, une action de* Microsoft*, et une obligation d'état italienne), mais qu'on est obligé d'acheter ensemble. Nous supposons ici que notre panier est composé de trois actifs notés respectivement* 1, 2 *et* 3 *; le juste prix du panier étant égal à la somme des justes prix de chacun des actifs le composant. Nous supposons en outre que le juste prix de chacun des actifs* 1, 2 *et* 3 *peut être déterminé en fonction d'une espérance :*

- *Le juste prix (en €) de l'actif* 1 *est* $4\,\mathbb{E}\big(\sin^2(U_1+U_2+U_3)\big)$*, où* U_1, U_2, U_3 *sont des variables aléatoires de même loi* Uniforme$([-1,1])$ *;*
- *Le juste prix de l'actif* 2 *est* $0{,}1\,\mathbb{E}\big(\max(e^{G_1}, e^{G_2})\big)$*, où* G_1, G_2 *sont des variables aléatoires indépendantes de même loi* $\mathcal{N}(3,2)$ *;*
- *Le juste prix de l'actif* 3 *est* $\mathbb{E}(P^{0,5})$*, où* P *est une variable aléatoire suivant une loi de Poisson de paramètre* 8*.*

1. Écrire une méthode de Monte-Carlo pour évaluer les juste prix de chacun des trois actifs du panier.
Indication : Pour simuler une loi de Poisson, il peut être utile de se rappeler (par exemple) la propriété suivante : si $E_1, E_2, \ldots$ suivent des lois exponentielles indépendantes de même taux λ (càd. que $\mathbb{P}(E. \geqslant x) = e^{-\lambda x}$), alors $\sup\{n \in \mathbb{N} \mid \sum_{i=1}^n E_i \leqslant 1\}$ suit une loi de Poisson de paramètre λ.

2. Déterminer les efficacités des calculs du juste prix de chacun de ces actifs [|||]. (Dans cet exercice, le cout d'un calcul sera assimilé à sa durée).
Indication : Attention, il y a un piège ! En effet, la variance de l'estimateur *du juste prix* de l'actif i n'est pas la même chose que la variance de l'estimateur *de l'espérance* qui intervient dans la formule de ce juste prix...

[|||]. Notez que ces efficacités dont des grandeurs *dimensionnées*... Quelle est leur homogénéité en l'occurrence ?

3. Grâce aux résultats de la question précédente, calculer comment allouer intelligemment le temps de calcul entre la détermination des justes prix de chacun des trois actifs pour déterminer le juste prix du panier avec une efficacité maximale ; puis en déduire ce que cela signifie en terme du *nombre de simulations* consacrées aux différents calculs.

4. Comparer l'efficacité de la détermination du juste prix du panier lorsqu'un utilise l'allocation optimale du tempsde calcul, par rapport à celle où on utilise l'allocation "naïve" consistant à faire le même nombre de simulations pour chaque actif.