M1 MAPI3 - Simulations stochastiques 2024-2025
Xavier Bressaud *xavier.bressaud@math.univ-toulouse.fr*
Manon Costa *manon.costa@math.univ-toulouse.fr*
Bastien Mallein *bastien.mallein@math.univ-toulouse.fr*

TD 8-9 & TP 5 - Méthode de Monte Carlo par chaînes de Markov - I

Exercice 1 - Algorithme de Metropolis-Hastings

Soit E un espace d'état fini et μ une mesure de probabilité sur E qui charge tous les états. Rappelons que l'algorithme de Metropolis-Hastings pour un noyau auxiliaire de transition Q (tel que $Q(x, y) = 0$ si et seulement si $Q(y, x) = 0$) s'écrit:

Etape 0 :
`Initialiser` X_0 ;
Etape n+1 :
`Choisir` y `selon la loi` $Q(X_n, \cdot)$;
`Poser` $X_{n+1} = y$ `avec proba` $\min\left(1, \frac{\mu(y)Q(y,X_n)}{\mu(X_n)Q(X_n,y)}\right)$, `sinon poser` $X_{n+1} = X_n$

1) Montrer que si la probabilité d'acceptation-rejet $\min\left(1, \frac{\mu(y)Q(y,x)}{\mu(x)Q(x,y)}\right)$ est remplacée par

$$\frac{\mu(y)Q(y, x)}{\mu(y)Q(y, x) + \mu(x)Q(x, y)},$$

la chaîne de Markov définie par l'algorithme de Metropolis-Hastings aura encore μ pour mesure invariante.

2) On remplace plus généralement la probabilité d'acceptation-rejet $\min\left(1, \frac{\mu(y)Q(y,x)}{\mu(x)Q(x,y)}\right)$ par

$$\alpha\left(\frac{\mu(y)Q(y, x)}{\mu(x)Q(x, y)}\right),$$

avec $\alpha : \mathbb{R}_+ \to]0, 1]$. Donner une condition suffisante sur la fonction α pour que la mesure μ soit la mesure invariante de la chaîne de Markov définie par l'algorithme de Metropolis-Hastings.

Exercice 2 - Méthode de Métropolis-Hastings sur l'espace $\{1, 2, 3\}$

On considère dans cet exercice l'espace d'états $\{1, 2, 3\}$ de dimension 3 et Q le noyau markovien donné par

$$\begin{pmatrix} 0 & a_1 & a_2 \\ c_1 & b_1 & a_3 \\ c_2 & c_3 & b_2 \end{pmatrix}.$$

La matrice Q est symétrique dès lors que $(a_1, a_2, a_3) = (c_1, c_2, c_3)$.

1. Quelles sont les conditions sur les paramètres pour avoir une matrice de transition ?
2. **TP -** Ecrire un programme qui permet de simuler une chaine de markov de matrice de transition Q (prendre une des matrices ci-dessous).
3. **TP -** On veut dans la suite simuler la loi de probabilité $\mu = (1/6, 1/2, 1/3)$. Utiliser l'algorithme de Metropolis Hasting pour simuler une chaine de Markov de loi invariante μ.
4. **TP -**Vérifier numériquement si (X_n) est la chaine de markov précédente, alors

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i=x} \xrightarrow[n\to\infty]{p.s.} \mu(x)$$

5. **TP -**Pour comparer les vitesses de convergences pour différences matrices, nous allons calculer la distance en variation totale entre la mesure empirique du vecteur $X_1, \dots X_n$ et mu.

$$d_{VT}(X, \mu) = \sum_{x=1}^{3} |\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i = x} - \mu(x)|.$$

On tracera l'évolution de cette quantité pour différentes valeurs de n (entre 0 et 300) et les matrices ci-dessous

$$Q_1 = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.6 & 0.3 & 0.1 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 0 & 0.3 & 0.7 \\ 0.3 & 0.4 & 0.3 \\ 0.7 & 0.3 & 0.0 \end{pmatrix}, \quad Q_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}.$$

Exercice 3 - Metropolis-Hastings vs méthode du rejet pour une densité continue

Le but est de générer des échantillons selon la loi sur $\mathbb{R}$ de densité

$$\mu_\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}(1 + \sin(\alpha x)),$$

pour $\alpha > 0$.

1. **TP -** Tracer la densité pour plusieurs valeurs de $\alpha > 0$.
2. Ecrire un pseudo code pour simuler un échantillon de variables i.i.d. de loi de densité μ_α par l'algorithme de Métropolis Hasting avec le noyau $Q(x, .) = \mathcal{N}(x, 1)$.
3. **TP -**On choisira maintenant $\alpha = 2$. Coder l'algorithme précédent pour simuler un échantillon de taille $n = 1000$. Mesurer le temps de calcul nécessaire avec la fonction $time.clock()$
4. **TP -**Simuler un échantillon de taille identique en utilisant la méthode du rejet avec comme fonction auxilliaire la densité gaussienne. Mesurer le temps de calcul et comparer.
5. **TP -**Calculer les approchées de l'espérance en utilisant les échantillons ci-dessus et comparer.

Exercice 4 - Metropolis-Hastings aux échecs

Une pièce d'échec se déplace aléatoirement sur un échiquier de 8×8 cases. En partant d'une position fixée (disons en bas à gauche) et en ne faisant que des déplacements autorisés, on veut modéliser une distribution uniforme μ sur l'échiquier.

1) On considère une tour (tous les mouvements horizontaux et verticaux sont autorisés). Montrer que choisir de façon équiprobable un déplacement autorisé à chaque étape conduit à une chaîne de Markov de mesure invariante μ. La loi de la position de la tour converge donc vers μ.
2) On considère une dame (tous les mouvements horizontaux, verticaux et en diagonal sont autorisés). Montrer que choisir de façon équiprobable un déplacement autorisé à chaque étape conduit à une chaîne de Markov qui n'a pas μ pour mesure invariante. La loi de la position de la dame ne converge donc pas vers μ.
3) Décrire une chaîne de Markov pour la dame utilisant les déplacements autorisés et le stationnement (laisser la dame sur place) de mesure invariante μ. On pourra par exemple utiliser l'algorithme de Metropolis-Hastings.

Exercice 5 - Recuit simulé et refroidissement trop rapide

Soit $E = \{1, 2, 3\}$ et $V : E \to \mathbb{R}$ donnée par $V(1) = 0$, $V(2) = 1$ et $V(3) = -1$ (la fonction V possède un minimum local en 1 et un minimum global en 3).

1) Ecrire l'algorithme de recuit simulé pour la matrice de transition

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}$$

et la suite de température $T_n = \frac{1}{c \log(n+1)}$, $n \geq 1$, $c > 0$.

2) Ecrire la matrice de transition P_n correspondante à la température T_n.

3) **TP-** Simuler numériquement l'algorithme du recuit simuler et tracer plusieurs trajectoires pour différentes valeurs de $c > 0$.

4) On suppose maintenant que la chaîne de Markov $(X_n)_{n\geq 0}$ évoluant selon cet algorithme part de $X_0 = 1$. Montrer que, pour tout $n \geq 2$, on a

$$\mathbb{P}[X_n = 3] \leq \sum_{k=0}^{n-2} \mathbb{P}[X_k = 1, X_{k+1} = 2].$$

5) Montrer que $\mathbb{P}[X_k = 1, X_{k+1} = 2] \leq \mathbb{P}[X_{k+1} = 2 | X_k = 1]$ et calculer cette valeur en fonction de c.

6) En déduire que pour c suffisamment grand, $\mathbb{P}[X_n = 3]$ ne tend pas vers 1 lorsque n tend vers l'infini (il n'y a pas convergence en probabilité de X_n vers le minimum global 3).

7) **TP -** Vérifier numériquement ces résultats et essayer de calibrer valeur de c permettant la convergence.

Exercice 6 - Recuit simulé et problème du voyageur de commerce

On considère le problème du voyageur de commerce suivant. Un commerçant doit visiter N clients dans N villes différentes puis revenir à son point de départ (en ne visitant qu'une seule fois chaque ville). On cherche à minimiser la distance totale parcourue.

On représente les villes par N points $v_1, \ldots, v_N$ dans $[0,1] \times [0,1] \subset \mathbb{R}^2$, un trajet possible par une suite $(v_{i(1)}, \ldots, v_{i(N)})$ de points et la distance à minimiser est

$$d\Big((v_{i(1)}, \ldots, v_{i(N)})\Big) = \sum_{k=1}^{N} |v_{i(k)} - v_{i(k+1)}|,$$

où $v_{i(N+1)}$ est $v_{i(1)}$.

1) Justifier que $N!$ est le nombre de trajets possibles.

2) **TP -** Simuler la position de $N = 10$ villes au hasard dans $[0,1] \times [0,1]$.

3) **TP -** Ecrire l'algorithme de recuit simulé pour le noyau d'exploration Q qui consiste à échanger 2 villes au hasard dans un trajet $(v_{i_1}, \ldots, v_{i_N})$.

3) **TP -** Représenter graphiquement les étapes de l'algorithme en reliant les villes des parcours.

4) **TP -** Représenter graphiquement la distance parcouru par le trajet à l'étape n pour différentes simulations de l'algorithme ayant les mêmes villes initiales. A partir de quelles valeurs de n pensez vous que l'algorithme a convergé ?

5) **TP -** Comparer la situation pour différentes suites de temps $T_{n+1} = \lambda T_n$ avec $\lambda \in \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1\}$.