

FEUILLE D'EXERCICES # 12

Exercice 1 *La chaîne serpent*

Soit $(X_n)_{n \geq 0}$ une CdM sur E de transition Q . Pour $\ell \geq 1$, on définit $Y_n := (X_n, X_{n+1}, \dots, X_{n+\ell})$.

1. La suite $(Y_n)_{n \geq 0}$ prend ses valeurs dans $F = E^{\ell+1}$. Montrez que c'est une chaîne de Markov et précisez sa matrice de transition.
2. Montrez que si $(X_n)_{n \geq 0}$ est irréductible, il en est de même pour $(Y_n)_{n \geq 0}$ si on restreint l'espace d'état à $\tilde{F} = \{(i_0, \dots, i_\ell) \in E^{\ell+1}, Q(i_0, i_1)Q(i_1, i_2) \dots Q(i_{\ell-1}, i_\ell) > 0\}$.
3. Montrez que si $(X_n)_{n \geq 0}$ a une distribution stationnaire π , alors $(Y_n)_{n \geq 0}$ a aussi une distribution stationnaire.

Exercice 2 *Estimation des probabilités de transition*

Soit $(X_n)_{n \geq 0}$ une CdM sur E , irréductible récurrente positive, de distribution stationnaire π .

1. Quelle est la loi invariante de la chaîne $((X_n, X_{n+1}, \dots, X_{n+\ell}))_{n \geq 0}$?
2. En déduire que si $g : E^{\ell+1} \rightarrow \mathbb{R}_+$ est telle que $\mathbb{E}_\pi[g(X_0, \dots, X_\ell)] < +\infty$, alors pour toute loi initiale μ , \mathbb{P}_μ -presque sûrement, lorsque n tend vers l'infini, on a :

$$\frac{1}{n} \sum_{k=0}^{n-1} g(X_k, X_{k+1}, \dots, X_{k+\ell}) \longrightarrow \mathbb{E}_\pi[g(X_0, \dots, X_\ell)].$$

On suppose que Q et π sont inconnues. On souhaite estimer les probabilités de transition $Q(i, j)$ à l'aide de la seule observation d'une trajectoire $(X_n(\omega))_{n \geq 0}$.

3. Soit $(i, j) \in E^2$, déterminer les limites presque sûre lorsque n tend vers l'infini de

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{\{X_k=i\}} \quad \text{et} \quad \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{\{X_k=i, X_{k+1}=j\}}.$$

4. En déduire un estimateur consistant de la probabilité de transition $Q(i, j)$.

Exercice 3 *Une version simplifiée de l'algorithme Pagerank de Google*¹

Voici comment fonctionne une version simplifiée de l'algorithme de classement utilisé par Google. Soit G le graphe (fini) dont les sommets s_1, \dots, s_N sont les pages du réseau internet et dont les arêtes sont les hyperliens entre les pages web. On ajoute un sommet s_0 au graphe G , ainsi qu'une arête entre s_0 et tous les sommets s_i , $i \in \{1, \dots, N\}$. On note \widehat{G} le graphe obtenu, A sa matrice d'adjacence i.e. $A_{ij} = 1$ s'il existe une arête entre s_i et s_j , $A_{ij} = 0$ sinon, et $N(i)$ le nombre de liens sortant de s_i et pointant vers s_j avec $j \neq i$. Étant donné un paramètre $p \in]0, 1[$, on définit une matrice Q sur \widehat{G} de la façon suivante : $\forall 1 \leq i \leq N$, $Q(s_0, s_i) = 1/N$ et $Q(s_i, s_i) = 0$, et

$$Q(s_i, s_0) = \begin{cases} 1 & \text{si } N(i) = 1 \\ 1-p & \text{si } N(i) \neq 1 \end{cases} \quad \forall 1 \leq i \leq N,$$

$$Q(s_i, s_j) = \begin{cases} 0 & \text{si } A_{ij} = 0 \\ p/N(i) & \text{si } A_{ij} = 1 \end{cases} \quad \forall 1 \leq i, j \leq N, i \neq j.$$

¹Un moteur de recherche fonctionne de la façon suivante : des "robots" scannent les pages web à la recherche de mots clefs ; l'information recueillie permet d'associer à chaque page web un répertoire de mots clefs puis inversement chaque mot clef se voit associé à un répertoire inverse de sites web. Lorsque qu'un utilisateur lance une recherche, le répertoire inverse produit une liste de sites web. Cette liste est ensuite ordonnée par un algorithme de classement (ranking). La liste ordonnée est enfin proposée à l'utilisateur.

1. Montrer que la matrice A est apériodique et que Q est une matrice stochastique apériodique.
2. Interpréter l'évolution de la chaîne de Markov associée à Q .
3. Montrer que la chaîne admet une unique probabilité invariante π . L'algorithme PageRank utilisé par Google interprète $0 \leq \pi(i) \leq 1$ comme la popularité de la page i , et ordonne les pages web par popularité.
4. Vous semble-t-il raisonnable de calculer explicitement la probabilité invariante? Comment faire alors?

Exercice 4 *La méthode MCMC et l'algorithme de Metropolis*

Soit E un espace d'états fini et π une probabilité sur E . Étant donnée une fonction f sur E , on souhaite calculer la quantité

$$\mathbb{E}_\pi[f] = \int_E f(x) d\pi(x) = \sum_{x \in E} f(x) \pi(x).$$

Il peut arriver (et c'est souvent le cas) que l'espace d'états E soit "très gros" et que les réels $\pi(x)$ ne soient pas explicitement calculables ou soient trop petits². Pour calculer $\mathbb{E}_\pi[f]$, on peut alors se dire qu'il faudrait mettre en place une méthode de Monte-Carlo. Mais alors, il faut disposer d'un algorithme générant des variables aléatoires distribuées selon la loi π sans disposer vraiment de la dite loi! La méthode de Monte Carlo via les chaînes de Markov (MCMC) consiste à construire une chaîne de Markov $(X_n)_{n \geq 0}$ sur E dont les probabilités de transition sont très simples et qui admet la loi π comme mesure invariante. Si l'on parvient à construire une telle chaîne, le théorème ergodique assure que, lorsque n tend vers l'infini :

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \longrightarrow \mathbb{E}_\pi[f].$$

Considérons une matrice de transition \tilde{Q} sur E telle que, $\tilde{Q}(x, y) > 0 \Rightarrow \tilde{Q}(y, x) > 0$ pour tout $x, y \in E$. Pour $x \neq y$, on pose alors

$$R(x, y) = \begin{cases} 1 \wedge \left(\frac{\pi(y)\tilde{Q}(y,x)}{\pi(x)\tilde{Q}(x,y)} \right) & \text{si } \tilde{Q}(x, y) \neq 0 \\ 0 & \text{sinon,} \end{cases}$$

et l'on définit une matrice de transition Q via les formules :

$$Q(x, y) = \tilde{Q}(x, y)R(x, y) \text{ pour } x \neq y, \quad Q(x, x) = 1 - \sum_{y \neq x} Q(x, y).$$

1. Montrer que si $\pi(x) > 0$ pour tout $x \in E$, alors la matrice Q est réversible par rapport à π .
2. Montrer que si \tilde{Q} est irréductible, alors Q est irréductible et apériodique.
3. Interpréter et commenter l'algorithme suivant :

étape 0 : - initialiser X_0

étape $n + 1$: $\begin{cases} - \text{choisir } y \text{ avec la loi } \tilde{Q}(X_n, \cdot) \\ - \text{tirer un nombre } U \text{ uniformément au hasard dans } [0, 1] \\ - \text{si } U < R(X_n, y) \text{ alors } X_{n+1} = y, \text{ sinon } X_{n+1} = X_n. \end{cases}$

²Par exemple, on peut considérer le modèle de physique statistique suivant $\Lambda = \{0, 1\}^n$ et $E = \{0, 1\}^\Lambda$. Le cardinal de E est alors $2^{(n+1)^2}$. Étant donnée une fonction H (comme Hamiltonien) sur E , on définit une mesure π de la façon suivante : $\forall x \in E, \pi(x) := \frac{1}{Z} e^{-H(x)}$, où $Z = \sum_{y \in E} e^{-H(y)}$. Même si H est connue, le calcul de $\pi(x)$ demande la connaissance de Z , ce qui n'est pas toujours raisonnable...