TP : Introduction aux chaînes de Markov

Soit $\mathfrak{X}$ un ensemble fini, dont on numérote les éléments de 1 à $N = \operatorname{card} \mathfrak{X}$: $\mathfrak{X} = \{x_1, x_2, \dots, x_N\}$. Une *matrice stochastique* (ou *matrice de transition*) sur $\mathfrak{X}$ est une matrice carrée de taille $N \times N$ à coefficients positifs, dont on note $P(x_i, x_j)$ l'élément sur la i-ième ligne et la j-ième colonne, et telle que :

$$\forall i \in \{1, 2, \dots, N\}, \ \sum_{j=1}^{N} P(x_i, x_j) = 1.$$

Une *chaîne de Markov* de matrice P est une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ avec chaque $X_n \in \mathfrak{X}$, telle que pour tout $n \geq 0$ et tout $(n+1)$-uplet $(a_0, a_1, \dots, a_n) \in \mathfrak{X}^{n+1}$, on a :

$$\mathbb{P}[X_0 = a_0, X_1 = a_1, \dots, X_n = a_n] = \mathbb{P}[X_0 = a_0]\, P(a_0, a_1)\, P(a_1, a_2) \cdots P(a_{n-1}, a_n).$$

Partie 1. Programmes généraux.

Question 1. On encode une probabilité $\pi = (\pi(x_1), \dots, \pi(x_N))$ sur $\mathfrak{X}$ par un *array* `numpy`, et une matrice stochastique par un *array* bidimensionnel : par exemple,

```
pi = np.array([0.2,0.5,0.3])
P = np.array([[1/3,1/2,1/6],[1/4,0,3/4],[0,1/2,1/2]])
```

définissent respectivement un vecteur de probabilité et une matrice stochastique sur un ensemble $\mathfrak{X}$ de taille 3. Écrire deux programmes `verifier_proba(pi)` et `verifier_matrice_sto(P)` qui vérifient qu'un vecteur est une probabilité, et qu'une matrice est stochastique. Pour tester l'égalité de nombres réels avec décimales, on pourra utiliser `np.isclose` (pour éviter les imprécisions dues à l'écriture décimale).

Question 2. On suppose données sur un ensemble fini $\mathfrak{X}$ une probabilité π_0 et une matrice stochastique P. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice P, et de *loi initiale* π_0 :

$$\mathbb{P}[X_0 = a] = \pi_0(a) \quad \forall a \in \mathfrak{X}.$$

Pour $n \geq 1$, exprimer en fonction de π_0 et de la matrice stochastique P le vecteur de probabilité $\pi_n = (\mathbb{P}[X_n = x_1], \dots, \mathbb{P}[X_n = x_N])$. Écrire un programme `loi_marginale(P, pi0, n)` qui calcule ce vecteur. Pour effectuer des produits matriciels, on pourra utiliser la commande `np.matmul`.

Question 3. Si $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice P, on a pour tout $n \geq 0$ et tous $a, b \in \mathfrak{X}$:

$$\mathbb{P}[X_{n+1} = b \,|\, X_n = a] = P(a, b).$$

Supposons donnés $X_0, \dots, X_n$ et une variable aléatoire U de loi uniforme sur $[0, 1]$, indépendante de $X_0, \dots, X_n$. Pour construire X_{n+1} sachant la valeur $X_n = a$, on peut donc prendre l'unique $k \in \{1, 2, \dots, N\}$ tel que

$$\sum_{j=1}^{k-1} P(a, x_j) < U \leq \sum_{j=1}^{k} P(a, x_j).$$

Écrire un programme `transition(P, a, u)` qui renvoie l'élément k, k étant caractérisé par les inégalités ci-dessus.

Question 4. En utilisant un vecteur de variables aléatoires $(U_0, U_1, \dots, U_n)$ uniformes sur $[0, 1]$ et indépendantes, déduire de la question précédente un programme `chaine_markov(P, pi0, n)` qui prend en argument une matrice stochastique de taille $N \times N$ et une loi initiale π_0, et qui renvoie une simulation de $X_0, X_1, \dots, X_n$ chaîne de Markov correspondante. Dans ce programme, $\mathfrak{X}$ sera représenté par les entiers de $\{1, 2, \dots, N\}$.

Partie 2. Un premier exemple.

Question 5. On prend le vecteur π et la matrice P de la question 1. comme loi initiale et matrice de transition d'une chaîne de Markov sur un espace à 3 états. Calculer la loi marginale π_{10}, puis comparer avec un diagramme en bâtons cette loi théorique avec la loi empirique de la variable X_{10} (on pourra prendre la loi empirique avec 1000 essais).

Question 6. Toujours avec la même loi initiale et la même matrice de transition, comparer les vecteurs π_{10} et π_{100}. Que peut-on conjecturer pour

$$\lim_{n\to\infty} \mathbb{P}[X_n = a]?$$

Question 7. En utilisant la méthode `np.linalg.eig` et la transposition `P.T` de matrices, trouver les valeurs propres $\lambda_1, \lambda_2, \lambda_3$ de P, et des vecteurs propres à gauche ν_1, ν_2, ν_3 tels que $\nu_i P = \lambda_i \nu_i$. En déduire une preuve de la conjecture de la question précédente.

Les résultats des questions 6. et 7. se généralisent à de nombreuses chaînes de Markov : sous certaines hypothèses naturelles, il existe un vecteur de probabilité π tel que, quelque soit la loi initiale π_0, on a

$$\lim_{n\to\infty} \pi_n = \pi.$$

On dit alors que π est la *loi stationnaire* ou *invariante* de la chaîne de Markov ; c'est la limite des lois marginales.

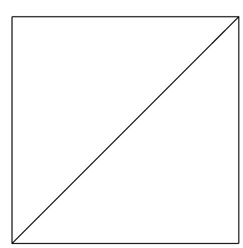
Question 8. Écrire un programme dépendant de n et qui trace les trois fonctions

$$f_i : k \in \{1, \dots, n\} \mapsto \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{X_j = i}$$

avec $i \in \{1, 2, 3\}$. Illustrer la convergence des fonctions f_i vers $\pi(\{i\})$, $i \in \{1, 2, 3\}$.

Ainsi, la loi invariante est aussi la limite des fréquences empiriques des états. Les deux dernières parties illustrent ces résultats pour deux autres modèles de chaînes de Markov.

Partie 3. Marche aléatoire sur un graphe.
On considère la marche aléatoire sur le graphe suivant :



À chaque étape, la nouvelle position X_{n+1} est choisie uniformément au hasard parmi tous les sommets voisins de X_n.

Question 9. Déterminer la matrice de transition P de cette chaîne de Markov, et écrire un programme simule jusqu'au temps n une chaîne de Markov de matrice de transition P, et telle que X_0 soit le coin supérieur gauche.

Question 10. Calculer la mesure invariante de cette chaîne, c'est dire la probabilité π sur $\{1, 2, 3, 4\}$ telle que $\pi P = \pi$).

Question 11. Écrire un programme dépendant de n qui trace les 4 fonctions

$$f_i : k \in \{1, \dots, n\} \to \frac{1}{k} \sum_{j=1}^{k} 1_{X_j=i}$$

où $i \in \{1, 2, 3, 4\}$ est un des sommets du graphe. Écrire aussi un programme dépendant de n qui représente les mesures empiriques $\mu_k(\{i\})$ pour $k \in \{1, \dots, n\}$. Que dire de la convergence de ces fonctions ?

Partie 4. Le modèle des urnes d'Ehrenfest. On fixe un entier $N \geq 1$, et on considère une boîte contenant N particules numérotées et tombant dans deux compartiments A et B. À chaque étape :

- on tire au hasard l'une des particules numérotées de 1 à N, chaque particule étant équiprobable.
- avec probabilité $\frac{1}{2}$, on fait passer la particule tirée au hasard de son compartiment vers l'autre compartiment (donc, de A vers B si elle était en A et de B vers A si elle était en B).
- avec probabilité $\frac{1}{2}$, on laisse la boîte invariante.

On note X_n le nombre de particules dans le compartiment A au temps n ; alors, il y a $N - X_n$ particules dans le compartiment B au temps n. On admet que si les choix décrits ci-dessus sont fait indépendamment à chaque étape, alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur $\mathfrak{X} = \{0, 1, \dots, N\}$ (attention, par rapport aux questions précédentes, $\mathfrak{X}$ est maintenant de cardinal $N + 1$).

Question 12. Calculer la matrice de transition de la chaîne définie ci-dessus. On remarquera que $P(k, l) = 0$ si $l \notin \{k - 1, k, k + 1\}$, de sorte qu'il suffit de calculer les valeurs $P(k, k - 1)$, $P(k, k)$ et $P(k, k + 1)$.

Question 13. Trouver l'unique vecteur de probabilité π qui vérifie pour tous $x, y \in \{0, 1, \dots, N\}$ l'*équation de réversibilité*

$$\pi(x)\, P(x, y) = \pi(y)\, P(y, x).$$

Notant $\alpha = \pi(0)$, on pourra essayer de calculer en fonction de α les valeurs $\pi(1)$, $\pi(2)$, etc. pour déterminer π à un coefficient multiplicatif près, puis utiliser la condition $\sum_{x=0}^{N} \pi(x) = 1$ pour trouver la valeur de α.

Question 14. Montrer que le vecteur π de la question précédente est invariant par P : $\pi P = \pi$.

Question 15. Avec $N = 10$, comparer l'histogramme de la loi π avec le diagramme en bâtons de la loi empirique de X_{100} (avec disons 1000 simulations indépendantes de cette variable aléatoire, et en partant de $X_0 = 0$). Commenter ces figures.