

DM : Chaînes de Markov (Comportement asymptotique)

Gabriel PEREIRA DE CARVALHO

14 janvier 2024

Exercice 1 : Deux variantes de l'algorithme de Metropolis-Hastings

Algorithme de Metropolis-Hastings

1. Exprimer la matrice de transition de la chaîne de Markov $(X_n)_{n \geq 0}$ en fonction de π et Q . En déduire que la chaîne de Markov est irréductible.

Soit $R : (x, y) \in E^2 \mapsto \min \left\{ 1, \frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)} \right\}$ la probabilité d'acceptation.

Alors, $\forall y \neq X_n$ on a $X_{n+1} = y$ avec probabilité $Q(X_n, y)R(X_n, y)$ car les tirages de \tilde{X}_{n+1} et U_{n+1} sont indépendantes. Donc, soit P la matrice de transition, on a

$$\begin{cases} P(x, y) &= Q(x, y)R(x, y) & \forall y \neq x \\ P(x, x) &= 1 - \sum_{y \neq x} P(x, y) \end{cases} \quad (1)$$

Alors, on remarque que Q irréductible sur $E \implies P$ irréductible sur E .

2. Montrer que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

Premièrement, on remarque que $\frac{R(x, y)}{R(y, x)} = \frac{\min \left\{ 1, \frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)} \right\}}{\min \left\{ 1, \frac{1}{\frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)}} \right\}} = \frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)}.$

Montrons que

$$\forall x, y \in E, \quad \pi(x)P(x, y) = \pi(y)P(y, x)$$

Pour $x = y$ c'est vrai. Alors, supposons $x \neq y$.

On calcule

$$\frac{\pi(x)P(x, y)}{\pi(y)P(y, x)} = \frac{\pi(x)Q(x, y)R(x, y)}{\pi(y)Q(y, x)R(y, x)} \quad (2)$$

$$= \frac{\pi(x)Q(x, y)}{\pi(y)Q(y, x)} \frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)} \quad (3)$$

$$= 1. \quad (4)$$

On en conclue que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

3. En déduire un résultat de convergence des moyennes trajectorielles.

Soit $F : E \rightarrow \mathbb{R}$ tel que $\mathbb{E}_\pi(|F|) < \infty$. Soit $X_0 \in E$ la donnée initiale de la trajectoire.

On remarque que la chaîne de Markov $(X_n)_{n \geq 0}$ avec matrice de transition P est irréductible avec mesure invariante unique π . Donc, par le **Théorème ergodique**, les moyennes trajectorielles convergent presque sûrement,

$$\frac{1}{n} \sum_{i=0}^{n-1} F(X_i) \xrightarrow{n \rightarrow +\infty} \mathbb{E}_\pi(F)$$

Première variante

1. Exprimer la matrice de transition de la chaîne de Markov $(X_n)_{n \geq 0}$ en fonction de π et Q_1, Q_2 . En déduire que la chaîne de Markov est irréductible.

Soit $R : (x, y, z) \in E^3 \mapsto \min \left\{ 1, \frac{\pi(z)Q_1(z, y)Q_2(y, x)}{\pi(x)Q_1(x, y)Q_2(y, z)} \right\}$ la probabilité d'acceptation.

Alors, $\forall z \neq X_n$ on a $X_{n+1} = z$ avec probabilité $\sum_{y \in E} Q_1(x, y)Q_2(y, z)R(x, y, z)$.

Donc, soit P la matrice de transition, on a

$$\begin{cases} P(x, z) &= \sum_{y \in E} Q_1(x, y)Q_2(y, z)R(x, y, z) & \forall z \neq x \\ P(x, x) &= 1 - \sum_{z \neq x} P(x, z) \end{cases} \quad (5)$$

Alors, on remarque que Q_1, Q_2 irréductible sur $E \implies P$ irréductible sur E .

2. Montrer que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

Premièrement, on remarque que $\frac{R(x, y, z)}{R(y, x, z)} = \frac{\min \left\{ 1, \frac{\pi(z)Q_1(z, y)Q_2(y, x)}{\pi(x)Q_1(x, y)Q_2(y, z)} \right\}}{\min \left\{ 1, \frac{1}{\frac{\pi(z)Q_1(z, y)Q_2(y, x)}{\pi(x)Q_1(x, y)Q_2(y, z)}} \right\}} = \frac{\pi(z)Q_1(z, y)Q_2(y, x)}{\pi(x)Q_1(x, y)Q_2(y, z)}.$

Montrons que

$$\forall x, z \in E, \quad \pi(x)P(x, z) = \pi(z)P(z, x)$$

Pour $x = z$ c'est vrai. Alors, supposons $x \neq z$.

On calcule

$$\frac{\pi(x)P(x, z)}{\pi(z)P(z, x)} = \frac{\pi(x) \sum_{y \in E} Q_1(x, y)Q_2(y, z)R(x, y, z)}{\pi(z) \sum_{y \in E} Q_1(z, y)Q_2(y, x)R(z, y, x)} \quad (6)$$

$$= \frac{\sum_{y \in E} \pi(x)Q_1(x, y)Q_2(y, z)R(z, y, x) \frac{\pi(z)Q_1(z, y)Q_2(y, x)}{\pi(x)Q_1(x, y)Q_2(y, z)}}{\sum_{y \in E} \pi(z)Q_1(z, y)Q_2(y, x)R(z, y, x)} \quad (7)$$

$$= \frac{\sum_{y \in E} \pi(z)Q_1(z, y)Q_2(y, x)R(z, y, x)}{\sum_{y \in E} \pi(z)Q_1(z, y)Q_2(y, x)R(z, y, x)} \quad (8)$$

$$= 1. \quad (9)$$

On en conclue que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

3. En déduire un résultat de convergence des moyennes trajectorielles.

Soit $F : E \rightarrow \mathbb{R}$ tel que $\mathbb{E}_\pi(|F|) < \infty$. Soit $X_0 \in E$ la donnée initiale de la trajectoire.

On remarque que la chaîne de Markov $(X_n)_{n \geq 0}$ avec matrice de transition P est irréductible avec mesure invariante unique π . Donc, par le **Théorème ergodique**, les moyennes trajectorielles convergent presque sûrement,

$$\frac{1}{n} \sum_{i=0}^{n-1} F(X_i) \xrightarrow{n \rightarrow +\infty} \mathbb{E}_\pi(F)$$

Deuxième variante

1. Exprimer la matrice de transition de la chaîne de Markov $(X_n)_{n \geq 0}$ en fonction de π et Q_0, Q_1 . En déduire que la chaîne de Markov est irréductible.

Soit $\forall c \in \{0, 1\}$, $R_c : (x, y) \in E^2 \mapsto \min \left\{ 1, \frac{\pi(y)Q_{1-c}(y, x)}{\pi(x)Q_c(x, y)} \right\}$ la probabilité d'acceptation pour $C_{n+1} = c$.

Alors, $\forall y \neq X_n$ on a $X_{n+1} = y$ avec probabilité $\frac{1}{2}Q_0(X_n, y)R_0(X_n, y) + \frac{1}{2}Q_1(X_n, y)R_1(X_n, y)$.

Donc, soit P la matrice de transition, on a

$$\begin{cases} P(x, y) &= \frac{1}{2}Q_0(x, y)R_0(x, y) + \frac{1}{2}Q_1(x, y)R_1(x, y) & \forall y \neq x \\ P(x, x) &= 1 - \sum_{y \neq x} P(x, y) \end{cases} \quad (10)$$

Alors, on remarque que Q_0, Q_1 irréductible sur $E \implies P$ irréductible sur E .

2. Montrer que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

Premièrement, $\forall c \in \{0, 1\}$, on remarque que

$$\frac{R_c(x, y)}{R_{1-c}(y, x)} = \frac{\min \left\{ 1, \frac{\pi(y)Q_{1-c}(y, x)}{\pi(x)Q_c(x, y)} \right\}}{\min \left\{ 1, \frac{\pi(x)Q_c(x, y)}{\pi(y)Q_{1-c}(y, x)} \right\}} \quad (11)$$

$$= \min \left\{ 1, \frac{\pi(y)Q_{1-c}(y, x)}{\pi(x)Q_c(x, y)} \right\} \max \left\{ 1, \frac{\pi(y)Q_{1-c}(y, x)}{\pi(x)Q_c(x, y)} \right\} \quad (12)$$

$$= \frac{\pi(y)Q_{1-c}(y, x)}{\pi(x)Q_c(x, y)} \quad (13)$$

Montrons que

$$\forall x, y \in E, \quad \pi(x)P(x, y) = \pi(y)P(y, x)$$

Pour $x = y$ c'est vrai. Alors, supposons $x \neq y$.

On calcule

$$\frac{\pi(x)P(x,y)}{\pi(y)P(y,x)} = \frac{\pi(x) \left(\frac{1}{2}Q_0(x,y)R_0(x,y) + \frac{1}{2}Q_1(x,y)R_1(x,y) \right)}{\pi(y) \left(\frac{1}{2}Q_0(y,x)R_0(y,x) + \frac{1}{2}Q_1(y,x)R_1(y,x) \right)} \quad (14)$$

$$= \frac{\pi(x)Q_0(x,y)R_0(x,y) + \pi(x)Q_1(x,y)R_1(x,y)}{\pi(y)Q_0(y,x)R_0(y,x) + \pi(y)Q_1(y,x)R_1(y,x)} \quad (15)$$

$$= \frac{\pi(x)Q_0(x,y)R_1(y,x) \frac{\pi(y)Q_1(y,x)}{\pi(x)Q_0(x,y)} + \pi(x)Q_1(x,y)R_0(y,x) \frac{\pi(y)Q_0(y,x)}{\pi(x)Q_1(x,y)}}{\pi(y)Q_0(y,x)R_0(y,x) + \pi(y)Q_1(y,x)R_1(y,x)} \quad (16)$$

$$= \frac{\pi(y)Q_1(y,x)R_1(y,x) + \pi(y)Q_0(y,x)R_0(y,x)}{\pi(y)Q_0(y,x)R_0(y,x) + \pi(y)Q_1(y,x)R_1(y,x)} \quad (17)$$

$$= 1. \quad (18)$$

On en conclue que $(X_n)_{n \geq 0}$ est réversible par rapport à π .

3. En déduire un résultat de convergence des moyennes trajectorielles.

Soit $F : E \rightarrow \mathbb{R}$ tel que $\mathbb{E}_\pi(|F|) < \infty$. Soit $X_0 \in E$ la donnée initiale de la trajectoire.

On remarque que la chaîne de Markov $(X_n)_{n \geq 0}$ avec matrice de transition P est irréductible avec mesure invariante unique π . Donc, par le **Théorème ergodique**, les moyennes trajectorielles convergent presque sûrement,

$$\frac{1}{n} \sum_{i=0}^{n-1} F(X_i) \xrightarrow{n \rightarrow +\infty} \mathbb{E}_\pi(F)$$

Exercice 2 : Couplage et critère de Doeblin

On vérifie que pour chaque x_n , l'expression donné définit une probabilité.

Pour $n_0 = 1$, on a $\forall z \in E, P(x, z) \geq \alpha m(z)$.

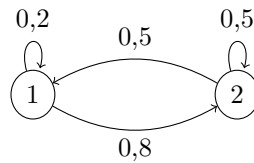
D'où $\forall z \in E, \frac{1}{1-\alpha} (P(x, z) - \alpha m(z)) \geq 0$.

Alors, on calcule la somme des probabilités

$$\sum_{z \in E} \frac{1}{1-\alpha} (P(x, z) - \alpha m(z)) = \frac{1}{1-\alpha} \sum_{z \in E} P(x, z) - \frac{\alpha}{1-\alpha} \sum_{z \in E} m(z) = 1.$$

0. Trouver un exemple de chaîne de Markov apériodique, irréductible et récurrente positive qui ne vérifie pas (1).

Considérons la chaîne de deux états avec matrice de transition $P = \begin{pmatrix} 0,2 & 0,8 \\ 0,5 & 0,5 \end{pmatrix}$.



On remarque que la chaîne est irréductible par connexité du graphe de transitions.

Par le **Théorème 4.6** du cours, on sait que cette chaîne est récurrente positive \iff elle admet l'unique mesure invariante. On calcule

$$\begin{cases} 0,2\pi(1) + 0,5\pi(2) = \pi(1) \\ 0,8\pi(1) + 0,5\pi(2) = \pi(2) \\ \pi(1) + \pi(2) = 1 \end{cases} \iff \begin{cases} \pi(1) = \frac{5}{13} \\ \pi(2) = \frac{8}{13} \end{cases} \quad (19)$$

donc, la chaîne est récurrente positive. Et la chaîne est apériodique car

$$\forall x, y \in E, \quad \forall n \in \mathbb{N}, \quad P^n(x, y) > 0.$$

1. Vérifier que P est la matrice de transition de la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$.

Calculons la probabilité de la transition $X_n \rightarrow z \in E$.

Fixons $X_n = x$. On a

$$\begin{aligned} \forall z \in E, \quad \mathbb{P}(X_{n+1} = z) &= \mathbb{P}(X_{n+1} = z | B_{n+1} = 0) \mathbb{P}(B_{n+1} = 0) + \mathbb{P}(X_{n+1} = z | B_{n+1} = 1) \mathbb{P}(B_{n+1} = 1) \\ &= \frac{1}{1-\alpha} (P(x, z) - \alpha m(z)) (1-\alpha) + m(z) \alpha \\ &= P(x, z) \end{aligned}$$

On en conclue que P est la matrice de transition de la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$.

2. Soit $\tau = \inf\{n \geq 1 : B_n = 1\}$.

Montrer que $\mathbb{P}_x(X_\tau = y) = m(y)$ puis que

$$\mathbb{P}_x(X_n = y, \tau \leq n) = \sum_{k=1}^n (1-\alpha)^{k-1} \alpha (m P^{n-k})(y).$$

On remarque que X_τ est tirée au sort dans E selon m , donc X_τ ne dépend pas de $X_0, \dots, X_{\tau-1}$ et $\mathbb{P}_x(X_\tau = y) = m(y)$.

Alors, calculons $\mathbb{P}_x(X_n = y, \tau \leq n)$. On a

$$\mathbb{P}_x(X_n = y, \tau \leq n) = \sum_{k=1}^n \mathbb{P}_x(X_n = y | \tau = k) \mathbb{P}_x(\tau = k) \quad (20)$$

$$= \sum_{k=1}^n \mathbb{P}_x(X_n = y | \tau = k) (1 - \alpha)^{k-1} \alpha \quad (21)$$

$$= \sum_{k=1}^n \left(\sum_{z \in E} \mathbb{P}_x(X_\tau = z, X_n = y | \tau = k) \right) (1 - \alpha)^{k-1} \alpha \quad (22)$$

$$= \sum_{k=1}^n \left(\sum_{z \in E} m(z) \mathbb{P}_z(X_{n-k} = y) \right) (1 - \alpha)^{k-1} \alpha \quad (23)$$

$$= \sum_{k=1}^n \left(\sum_{z \in E} m(z) P^{n-k}(z, y) \right) (1 - \alpha)^{k-1} \alpha \quad (24)$$

$$= \sum_{k=1}^n (m P^{n-k})(y) (1 - \alpha)^{k-1} \alpha. \quad (25)$$

3. Soient μ et ν deux probabilités sur E . Dédurre de la question précédente que

$$\mu P^n(y) - \nu P^n(y) = \sum_{x \in E} \mathbb{P}_x(X_n = y, \tau > n) (\mu(x) - \nu(x))$$

puis que

$$\|\mu P^n - \nu P^n\|_{VT} = \frac{1}{2} \sum_{y \in E} |\mu P^n(y) - \nu P^n(y)| \leq (1 - \alpha)^n$$

Soit $y \in E$, d'abord on calcule $\mu P^n(y) - \nu P^n(y)$. On a

$$\mu P^n(y) - \nu P^n(y) = \sum_{x \in E} (\mu(x) - \nu(x)) P^n(x, y) \quad (26)$$

$$= \sum_{x \in E} (\mu(x) - \nu(x)) \mathbb{P}_x(X_n = y) \quad (27)$$

$$= \sum_{x \in E} (\mu(x) - \nu(x)) (\mathbb{P}_x(X_n = y, \tau \leq n) + \mathbb{P}_x(X_n = y, \tau > n)) \quad (28)$$

Mais par la question précédente, on sait que $\mathbb{P}_x(X_n = y, \tau \leq n)$ ne dépend pas de x et comme $\sum_{x \in E} \mu(x) = \sum_{x \in E} \nu(x) = 1$, les termes avec $\tau \leq n$ s'annulent. On en conclue,

$$\mu P^n(y) - \nu P^n(y) = \sum_{x \in E} (\mu(x) - \nu(x)) \mathbb{P}_x(X_n = y, \tau > n).$$

Alors, on remarque que $\tau > n \iff B_1 = \dots = B_n = 0$. D'où,

$$\mathbb{P}(\tau > n) = (1 - \alpha)^n.$$

Donc, on a

$$\|\mu P^n - \nu P^n\|_{VT} = \frac{1}{2} \sum_{y \in E} |\mu P^n(y) - \nu P^n(y)| \quad (29)$$

$$= \frac{1}{2} \sum_{y \in E} \left| \sum_{x \in E} (\mu(x) - \nu(x)) \mathbb{P}_x(X_n = y, \tau > n) \right| \quad (30)$$

$$\leq \frac{1}{2} \sum_{y \in E} \left(\sum_{x \in E} |\mu(x) - \nu(x)| \mathbb{P}_x(X_n = y, \tau > n) \right) \quad (31)$$

$$\leq \frac{1}{2} \sum_{y \in E} \left(\sum_{x \in E} \mathbb{P}_x(X_n = y, \tau > n) \right) \quad (32)$$

$$\leq \frac{1}{2} \sum_{x \in E} \left(\sum_{y \in E} \mathbb{P}_x(X_n = y, \tau > n) \right) \quad (33)$$

$$= \frac{1}{2} \sum_{x \in E} \mathbb{P}_x(\tau > n) \quad (34)$$

$$= \frac{1}{2} \mathbb{P}(\tau > n) \quad (35)$$

$$= \frac{1}{2} (1 - \alpha)^n \leq (1 - \alpha)^n \quad (36)$$

4. Montrer que $y \mapsto P^n(x, y)$ converge vers une mesure π et que π ne dépend pas de x .
Montrer que π est une mesure de probabilité invariante pour le processus $\{X_n\}_{n \geq 0}$.

Montrons d'abord qu'il existe une mesure de probabilité invariante π .

Posons $X_0 = x \in E$, on veut calculer $\mathbb{E}_x(T_x^+)$.

Premièrement, on remarque que $\forall x, y \in E$ on a $P(x, y) \leq 1 - P(x, x) \leq 1 - \alpha m(x)$. Alors, soit $k \in \mathbb{N}$, on a

$$\mathbb{P}_x(T_x^+ > k) = \sum_{y \neq x} P(x, y) \mathbb{P}_y(T_x^+ > k - 1) \quad (37)$$

$$\leq (1 - \alpha m(x)) \mathbb{P}_x(T_x^+ > k - 1) \quad (38)$$

Par récurrence, on conclue $\mathbb{P}_x(T_x^+ > k) \leq (1 - \alpha m(x))^k$. Alors, on a

$$\mathbb{E}_x(T_x^+) = \sum_{k \geq 0} \mathbb{P}_x(T_x^+ > k) \leq \sum_{k \geq 0} (1 - \alpha m(x))^k < +\infty. \quad (39)$$

Donc, par le **Théorème 4.6** du cours, on conclue qu'il existe une mesure invariante π donnée par

$$\pi(x) = \frac{1}{\mathbb{E}_x(T_x^+)}$$

Finalement, montrons que $y \mapsto P^n(x, y)$ converge vers π .

Par l'exercice 3, on a

$$\|P^n(x, \cdot) - \pi P^n\|_{VT} = \|P^n(x, \cdot) - \pi\|_{VT} \leq (1 - \alpha)^n \xrightarrow{n \rightarrow +\infty} 0.$$

On en conclue que $y \mapsto P^n(x, y)$ converge vers π .

5. Montrer maintenant que lon peut toujours se ramener au cas $n_0 = 1$ dans 1.

Supposons $n_0 > 1$. Soit $(X_n)_{n \geq 0}$ une chaîne de Markov avec matrice de transition P et $X_0 = x \in E$.

Par division euclidienne,

$$\forall n \in \mathbb{N}, \quad \exists k, \ell \in \mathbb{N} \quad \text{tels que } n = kn_0 + \ell \quad \text{avec } 0 \leq \ell < n_0.$$

Soit $(\tilde{X}_n)_{n \geq 0}$ une chaîne de Markov avec matrice de transition P^{n_0} avec \tilde{X}_0 distribuée sous $P^\ell(x, \cdot)$.

On remarque que la chaîne $(\tilde{X}_n)_{n \geq 0}$ vérifie le **critère de Doeblin** avec $n_0 = 1$. On a

$$\mathbb{P}_x(X_n = y) = \sum_{z \in E} P^\ell(x, z) \mathbb{P}_z(X_{n-\ell} = y) \tag{40}$$

$$= \sum_{z \in E} P^\ell(x, z) \mathbb{P}_z(\tilde{X}_k = y) \tag{41}$$

donc si $y \mapsto P^{kn_0}(z, y)$ converge vers une mesure invariante, alors $y \mapsto P^n(x, y)$ converge aussi.