Chaînes de Markov :

mesures invariantes et récurrence positive

Étant donnée une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ sur un espace d'états $\mathcal{X}$, on sait maintenant répondre à la question "combien de fois un état x est-il visité ?" (récurrence / transience).

La théorie des mesures invariantes va permettre en particulier de répondre à la question plus précise "à quelle fréquence moyenne un état x est-il visité ?", dans le cas d'une chaîne irréductible récurrente.

On fixe dans tout ce qui suit une matrice stochastique P irréductible sur un espace d'états $\mathcal{X}$.

1. Mesures invariantes : définition et exemples.

Définition : Une mesure invariante pour P est une mesure positive π sur $\mathcal{X}$, de masse $\sum_{x \in \mathcal{X}} \pi(x) = \pi(\mathcal{X})$ finie ou infinie, telle que $\pi P = \pi$.

On parle aussi de mesure stationnaire. Si π est de masse finie, alors en considérant $\frac{\pi}{\pi(\mathcal{X})}$, on peut se ramener au cas où π est de masse 1. On parle alors de probabilité invariante ou stationnaire.

$$\pi P = \pi \; : \quad \forall x \in \mathcal{X}, \quad \pi(x) = \sum_{t \in \mathcal{X}} \pi(t) P(t, x).$$

Proposition Si P est irréductible, alors une mesure invariante π non nulle est > 0 sur tout l'espace $\mathcal{X}$.

__Preuve__ : Soit x tel que $\pi(x) > 0$. Si $y \in \mathcal{X}$, $\exists n \geq 1 \mid$ $P^n(x,y) > 0$. Par récurrence sur n, on montre facilement que $\pi = \pi P^n$.

Alors, $\pi(y) = \sum_{t \in \mathcal{X}} \pi(t) P^n(t,y) \geq \pi(x) P^n(x,y) > 0$.

Ceci est valable pour tout $y \in \mathcal{X}$ par irréducibilité de P. $\square$.

__exemples__

1) marche aléatoire sur $\mathbb{Z}$: $P(k,k+1) = p$, $P(k,k-1) = 1-p$

Une mesure invariante est la mesure de comptage $\pi(k) = 1 \; \forall k \in \mathbb{Z}$.

En effet,
$$\begin{aligned} 1 = \pi(k) &= p + (1-p) \\ &= P(k-1,k) + P(k+1,k) \\ &= \pi(k-1) P(k-1,k) + \pi(k+1) P(k+1,k) \\ &= (\pi P)(k). \end{aligned}$$

Bien sûr, tout multiple $\lambda\pi$ avec $\lambda > 0$ est encore une mesure invariante.

Si $p \neq \frac{1}{2}$, une autre mesure invariante est :

$$\pi(k) = \left(\frac{p}{1-p}\right)^k \quad \forall k \in \mathbb{Z}.$$

En effet, on a de nouveau

$$\begin{aligned}
\pi(k) = \left(\frac{p}{1-p}\right)^k &= (1-p)\left(\frac{p}{1-p}\right)^k + p\left(\frac{p}{1-p}\right)^k \\
&= p\left(\frac{p}{1-p}\right)^{k-1} + (1-p)\left(\frac{p}{1-p}\right)^{k+1} \\
&= \pi(k-1)\,P(k-1,k) + \pi(k+1)\,P(k+1,k) \\
&= (\pi P)(k).
\end{aligned}$$

2) marche aléatoire sur un graphe fini.

Soit $G = (V, E)$ un graphe fini avec $\deg x \geq 1 \quad \forall x \in V$

$P = P_G$ la matrice de transition de la marche aléatoire sur ce graphe.

$$P(x, y) = \begin{cases} \dfrac{1}{\deg x} & \text{si } x \sim y, \\ 0 & \text{si } x \text{ et } y \text{ ne sont pas voisins.} \end{cases}$$

Une mesure invariante de masse finie est $\pi(x) = \deg x$.

$$(\pi P)(x) = \sum_{y \sim x} \pi(y) P(y, x) = \sum_{y \sim x} \deg y \frac{1}{\deg y}$$

$$= \sum_{y \sim x} 1 = \deg x = \pi(x)$$

Une probabilité invariante est donc

$$\pi_{\text{renormalisée}}(x) = \frac{\deg x}{\sum_{y \in V} \deg y} = \frac{\deg x}{2|E|}.$$

2. Existence et unicité pour une chaîne récurrente irréductible.

Théorème Soit $(\mathcal{X}, P)$ espace et matrice d'une chaîne de Markov récurrente irréductible. À un facteur multiplicatif positif près, il existe une unique mesure invariante π sur $\mathcal{X}$.

Preuve (existence) On fixe $x \in \mathcal{X}$. Comme on raisonne à un coefficient multiplicatif près, on pourra pour l'existence et l'unicité supposer que la mesure invariante donne à x une masse 1.

On définit :

$$\mu_x(y) = \mathbb{E}[\text{nbre de visites de } y \text{ lors de la première excursion } \mathcal{E}_x^{(1)}]$$

$$= \mathbb{E}\left[\sum_{n=0}^{\tau_x^+ - 1} \mathbb{1}_{(X_n = y)}\right].$$

Point non trivial : $\mu_x(\mathcal{X}) = \sum_{y \in \mathcal{X}} \mu_x(y)$

$$= \mathbb{E}\left[\sum_{n=0}^{\tau_x^+ - 1} \sum_{y \in \mathcal{X}} \mathbb{1}_{(X_n = y)}\right]$$

$$= \mathbb{E}\left[\sum_{n=0}^{\tau_x^+ - 1} 1\right] = \mathbb{E}[\tau_x^+]$$

peut valoir $+\infty$, ou être fini

Mais : $\forall y \in \mathcal{X}$, $\mu_x(y) < +\infty$.

$\#$ visites de y le long de $\mathcal{E}_x^{(1)}((X_n)_{n\in\mathbb{N}})$

$$= \mathbb{1}_{(\tau_x^+ < \tau_y)} \times 0 + \mathbb{1}_{(\tau_y < \tau_x^+)} \times (\# \text{ visites de } y \text{ de } (X_{n+\tau_y})_{n\in\mathbb{N}} \text{ avant d'atteindre } x)$$

Donc $\mu_x(y) = \mathbb{P}_x[\tau_y < \tau_x^+] \, \mathbb{E}_y\Big[\underbrace{\sum_{n=0}^{\tau_x - 1} \mathbb{1}_{X_n = y}}_{} \Big]$

(propriété de Markov forte)

variable géométrique de paramètre $\mathbb{P}_y[\tau_x < \tau_y^+]$.

donc d'espérance $\dfrac{1}{\mathbb{P}_y[\tau_x < \tau_y^+]}$.

$$= \frac{\mathbb{P}_x[\tau_y < \tau_x^+]}{\mathbb{P}_y[\tau_x < \tau_y^+]} \in (0, +\infty).$$

Montrons maintenant que μ_x est invariante. Remarquons qu'on a

$$\sum_{n=0}^{\tau_x^+ - 1} \mathbb{1}(X_n = y) = \sum_{n=1}^{\tau_x^+} \mathbb{1}(X_n = y).$$ (lemme cyclique)

Alors, $$(\mu_x P)(y) = \sum_{t \in \mathcal{X}} \mathbb{E}_x\left[\sum_{n=0}^{\tau_x^+ - 1} \mathbb{1}(X_n = t)\right] P(t, y)$$

$$= \sum_{n=0}^{\infty} \sum_{t \in \mathcal{X}} \mathbb{E}_x\left[\mathbb{1}(X_n = t)\, \mathbb{1}(\tau_x^+ > n)\right] P(t, y).$$

$$= \sum_{n=0}^{\infty} \sum_{t \in \mathcal{X}} \mathbb{P}_x\left[X_n = t,\ \tau_x^+ \geq n+1\right] P(t, y).$$

On remarque que :

$$\mathbb{P}_x[X_n = t, \tau_x^+ \geq n+1]\, P(t,y)$$

$$= \sum_{x_1, \dots x_{n-1} \neq x} P(x, x_1) P(x_1, x_2) \dots P(x_{n-1}, t)\, P(t,y)$$

$$= \mathbb{P}_x[X_n = t, X_{n+1} = y, \tau_x^+ \geq n+1].$$

$$\text{Donc } (\mu_x P)(y) = \sum_{n=0}^{+\infty} \sum_{\cancel{t \in \mathcal{X}}} \mathbb{P}_x[X_n = t, X_{n+1} = y, \tau_x^+ \geq n+1]$$

$$= \sum_{n=0}^{+\infty} \mathbb{P}_x[X_{n+1} = y, \tau_x^+ \geq n+1]$$

$$= \sum_{m=1}^{+\infty} \mathbb{E}_x[\mathbb{1}_{(X_m = y)} \mathbb{1}_{(\tau_x^+ \geq m)}]$$

$$= \mathbb{E}_x\Big[\sum_{m=1}^{\tau_x^+} \mathbb{1}_{(X_m = y)}\Big] = \mu_x(y) \qquad \square.$$

Preuve (unicité) : on utilise un argument assez classique en théorie des chaînes de Markov : μ_x a une propriété d'extrémalité.

Soit π une autre mesure invariante ; on peut supposer $\pi(x) = 1$.

Pour $y \neq x$, on a $\mu_x(y) = \sum_{n=1}^{\infty} \underbrace{\mathbb{P}_x[X_n = y, \tau_x^+ > n]}_{u_n}$

Par invariance de π :

$$\pi(y) = \sum_{x_1} \pi(x_1) P(x_1, y)$$

$$= \underbrace{P(x, y)}_{\mathbb{P}_x[X_1 = y, \tau_x^+ > 1]} + \sum_{x_1 \neq x} \pi(x_1) P(x_1, y)$$

$$= u_1 + \sum_{x_1 \neq x} \pi(x_1) P(x_1, y)$$

$$= u_1 + \sum_{x_1, x_2 \neq x} \pi(x_2)\, P(x_2, x_1)\, P(x_1, y)$$

$$+ \sum_{x_1 \neq x} P(x, x_1)\, P(x_1, y) \rightarrow \mathbb{P}_x[X_2 = y, \tau_x^+ > 2]$$

$$= u_1 + u_2 + \sum_{x_1, x_2 \neq x} \pi(x_2)\, P(x_2, x_1)\, P(x_1, y)$$

$$= \dots$$

$\rightarrow$ on peut extraire de $\pi(y)$ les termes u_n de $\mu_x(y)$

$\Rightarrow \pi(y) \geq \mu_x(y)$ si $y \neq x$, et $\pi(x) = \mu_x(x) = 1$.

$\Rightarrow \mu_x$ est minimale.

Mais alors, $\pi - \mu_x \geq 0$ et est une mesure invariante qui s'annule en x $\Longrightarrow$ elle s'annule partout $\Rightarrow \pi = \mu_x$.

$\square$

3. Chaînes récurrentes positives

Pour une chaîne irréductible récurrente, comme toutes les mesures invariantes sont proportionnelles, on a deux possibilités :

1. toutes les mesures invariantes sont de masse finie

$\Longleftrightarrow$ il existe une probabilité invariante (renormaliser une mesure invariante par sa masse)

$$\Longleftrightarrow \quad \forall x \in \mathcal{X}, \ \mu_x(\mathcal{X}) = \mathbb{E}_x[\tau_x^+] < +\infty$$

On parle alors de chaîne **récurrente positive**.

<u>OU</u>

2. toutes les mesures invariantes sont de masse infinie.

$$\Longleftrightarrow \forall x \in \mathcal{X}, \ \mu_x(\mathcal{X}) = \mathbb{E}_x[\tau_x^+] = +\infty.$$

On parle alors de chaîne récurrente nulle.

transiente	$\tau_x^+ = +\infty$ avec proba > 0
récurrente nulle	$\tau_x^+ < +\infty$, mais $\mathbb{E}_x[\tau_x^+] = +\infty$
récurrente positive	$\tau_x^+ < +\infty$ et $\mathbb{E}_x[\tau_x^+] < +\infty$

<u>Remarques</u> :

- Dans le cas récurrent positif, l'unique probabilité invariante est $\overline{\pi} = \dfrac{\mu_x(\cdot)}{\mu_x(\mathcal{X})}$ (pour n'importe quel x)

En évaluant en x : $\pi(x) = \dfrac{1}{\mathbb{E}_x[\tau_x^+]}$.

Ceci permet de calculer les espérances des temps d'excursions.

- Une chaîne irréductible finie est forcément récurrente positive.

<u>Théorème</u> Considérons réciproquement une chaîne $(\mathcal{X}, \mathbb{P})$ admettant une mesure invariante π de masse finie.

Alors, la chaîne est récurrente positive.

<u>Preuve</u> : On se donne $x \in \mathcal{X}$ et on peut supposer $\pi(x) = 1$.

Alors, l'inégalité $\pi \geq \mu_x$ n'utilisait pas la récurrence.

Donc, $+\infty > \pi(\mathcal{X}) \geq \mu_x(\mathcal{X}) = \mathbb{E}_x[\tau_x^+]$

et x est récurrent (positif) $\square$.

4. Spoiler : théorèmes ergodiques

Soit $(\mathcal{X}, P)$ chaîne irréductible récurrente positive.

π = probabilité invariante.

On suppose que $P(x,x) > 0$ pour au moins un état $x \in \mathcal{X}$.

1) $\pi_n(y) = \mathbb{P}_{\pi_0}(X_n = y) \underset{n \to +\infty}{\longrightarrow} \pi(y)$ $\forall y$
$\forall$ mesure initiale π_0.

2) $\frac{1}{n} \sum_{t=1}^{n} \mathbb{1}(X_t = y) \xrightarrow[n \to +\infty]{p.s.} \pi(y)$ $\forall y$
$\forall$ mesure initiale π_0.