

III Théorèmes limites pour les chaînes de Markov

1 Le théorème ergodique

Corollaire 1. Soit $N_x = \sum_{n=0}^{\infty} 1_{\{X_n = x\}}$. Alors si x est transient

$$P_x(N_x = k) = (1-a)^k a, \quad k \geq 1, \quad a = P_x(T_x < \infty) < 1 \quad (1)$$

et si x est récurrent on a $P_x(N_x = \infty) = 1$.

Démonstration.

$$P_x(N_x = k) = P_x(T_x^{i+1} < \infty, i = 0, \dots, k-1; T_x^{k+1} = \infty) = (1-a)^k a$$

Théorème 2. Soit P une matrice irréductible récurrente positive et sa probabilité stationnaire. Alors pour tout $x, y \in M$

$$P_x \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_{\{X_k = y\}} = \pi(y) \quad P_x \text{ p.s.} \quad (2)$$

Démonstration. Par la récurrence on a que $T_y^k < \infty$ pour tout $k \geq 1$, donc les v.a. $\frac{k}{T_y^k} = \frac{T_y^{k+1} - T_y^k}{T_y^k}$ sont bien définies pour tout $k \geq 1$ et par le théorème de récurrence on a que la suite $(\frac{k}{T_y^k})_{k \geq 1}$ est iid et tel que $P_x(\frac{k}{T_y^k} = m) = P_y(T_y = m)$. D'après la loi des grands nombres

$$\frac{1}{k} \sum_{i=1}^{k-1} \frac{i}{T_y^i} E_y(T_y) = \frac{1}{\pi(y)} \quad P_x \text{ p.s.}$$

et $P_x(T_y < \infty) = 1$ donc

$$\frac{1}{k} T_y^k = \frac{1}{k} \sum_{i=1}^{k-1} (T_y^i + T_y - T_y^i) E_y(T_y^1) = \pi(y) \quad P_x \text{ p.s.}$$

Par ailleurs, si pour tout $n \geq 1$ on pose $N_y^n = \sum_{k=0}^{n-1} 1_{\{X_k = y\}}$ alors on a que

$$T_y^{N_y^n} \leq n \leq T_y^{N_y^n + 1}$$

et donc

$$\frac{N_y^n}{T_y^{N_y^n + 1}} \leq \frac{1}{n} \sum_{k=0}^{n-1} 1_{\{X_k = y\}} \leq \frac{N_y^n}{T_y^{N_y^n}}.$$

D'après le résultat.

Remarque 3. En modifiant légèrement cette preuve, on obtient le résultat plus général suivant.

Corollaire 4. Soit P une matrice irréductible récurrente positive et sa probabilité stationnaire. Soit f une fonction dans $L^1(\pi)$, i.e. $\sum_{x \in M} |f(x)| \pi(x) < \infty$. Alors

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{x \in M} f(x) \pi(x) \quad P_x \text{ p.s.}$$

2 Convergence vers l'équilibre

Un corollaire du théorème ergodique est que lorsque P est une matrice de transition irréductible et apériodique positive de probabilité invariante

$$\frac{1}{n} \sum_{j=0}^{n-1} P^j \xrightarrow{\text{loi}} \pi,$$

pour toute probabilité μ_0 sur (M) . Ceci n'implique pas en général que $P^n \xrightarrow{\text{loi}} \pi$:

Exemple 5. La matrice

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

a pour loi stationnaire $\pi = (\frac{1}{2}, \frac{1}{2})$ et comme puissances

$$P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Donc $P^n \xrightarrow{\text{loi}} \pi$ si n est pair et $P^n \xrightarrow{\text{loi}} \pi$ si n est impair et on voit que P^n ne converge pas vers π .

On cherche maintenant des conditions sur P pour que $P^n \xrightarrow{\text{loi}} \pi$. On cherche également à estimer la vitesse de cette convergence.

Définition 6. On définit une distance d sur (M) par

$$d(\mu_0) = \sup_{C \subset M} |\mu_0(C) - \pi(C)| = \frac{1}{2} \sup_{x \in M} |\mu_0(x) - \pi(x)|.$$

On a $d(\mu_0) \leq 1$ et $d(\mu_0) = 0 \iff \mu_0 = \pi$.

Démonstration. Il est clair que $d(\mu_0) \leq 1$ (car si $\mu_0(A) > \pi(A)$ alors $|\mu_0(A) - \pi(A)| \leq \pi(A) \leq 1$ pour des probabilités et que $d(\mu_0) = 0 \iff \mu_0 = \pi$). Pour montrer l'équivalence des deux expressions de $d(\mu_0)$ on observe que

$$d(\mu_0) = \sup_{C \subset M} |\mu_0(C) - \pi(C)| = \sup_{C \subset M} \left| \sum_{x \in C} \mu_0(x) - \sum_{x \in C} \pi(x) \right| = \sup_{x \in A} |\mu_0(x) - \pi(x)|$$

et si on pose $A = \{x \in M : \mu_0(x) > \pi(x)\}$ on obtient l'inégalité opposée

$$\begin{aligned} \frac{1}{2} \sup_{x \in M} |\mu_0(x) - \pi(x)| &= \frac{1}{2} \left(\sum_{x \in A} [\mu_0(x) - \pi(x)] + \sum_{x \in A^c} [\pi(x) - \mu_0(x)] \right) \\ &= \frac{1}{2} (\mu_0(A) - \pi(A) + (\pi(A^c) - \mu_0(A^c))) = d(\mu_0). \end{aligned}$$

Dans la suite on notera $P^n(x, A) = \sum_{y \in A} P^n(x, y) = P_x(X_n \in A)$ et on utilisera le lemme suivante

Lemme 7. Soit $f: M \rightarrow \mathbb{R}$ bornée alors

$$\inf_{c \in \mathbb{R}} \sup_{x \in M} |f(x) - c| = \frac{1}{2} \sup_{x, y \in M} |f(x) - f(y)|$$

Démonstration. D'une part, pour tout $c \in \mathbb{R}$

$$\sup_{x, y \in M} |f(x) - f(y)| = \sup_{x, y \in M} (|f(x) - c| + |f(y) - c|) = 2 \sup_{x \in M} |f(x) - c|$$

et donc

$$\inf_{c \in \mathbb{R}} \sup_{x \in M} |f(x) - c| = \frac{1}{2} \sup_{x, y \in M} |f(x) - f(y)|.$$

D'autre part on a que

$$\sup_{x,y \in M} |f(x) - f(y)| = \sup_{x \in M} f(x) - \inf_{y \in M} f(y),$$

soit alors $\hat{c} = (\sup_{x \in M} f(x) + \inf_{y \in M} f(y))/2$:

$$\hat{c} \leq \frac{1}{2} \sup_{x,y \in M} |f(x) - f(y)| = \inf_{x \in M} f(x) - f(x) \leq \sup_{x \in M} f(x) = \hat{c} + \frac{1}{2} \sup_{x,y \in M} |f(x) - f(y)|$$

ce qui donne

$$\inf_{x \in M} \sup_{y \in M} |f(x) - \hat{c}| \leq |f(x) - \hat{c}| \leq \frac{1}{2} \sup_{x,y \in M} |f(x) - f(y)|.$$

Proposition 8. Soit P une matrice de transition sur M . On pose

$$\alpha_n = \sup_{x,y \in M} d(P^n(x), P^n(y)) \quad (3)$$

Alors,

1. $\alpha_{n+m} \leq \alpha_n + \alpha_m$
2. Soit $\alpha_n = 1$ pour tout $n \geq 1$, soit $C < 1$ et $\alpha < 1$ tels que $\alpha_n \leq C \alpha^{n-1}$ pour $n \geq 1$.
3. Dans le cas $\alpha_n \leq C \alpha^{n-1}$ avec $\alpha < 1$, il existe une unique probabilité stationnaire et de plus,

$$|P^n(x, y) - \pi(y)| \leq C \alpha^n, \quad n \geq 1, x, y \in M. \quad (4)$$

Donc, lorsqu'il existe un indice n_0 tel que $\alpha_{n_0} < 1$, on a (4), ce qui implique que $\pi^n \rightarrow \pi$, π est la probabilité stationnaire, et que ces convergences ont lieu à des vitesses décroissant exponentiellement vite.

Démonstration.

$$\begin{aligned} P^{n+m}(x, A) - P^{n+m}(y, A) &= \sum_{z \in M} P^m(z, A) [P^n(x, z) - P^n(y, z)] \\ &= \inf_{z \in M} [P^m(z, A) - P^m(z, A)] [P^n(x, z) - P^n(y, z)] \\ &= \inf_{z \in M} |P^m(z, A) - P^m(z, A)| |P^n(x, z) - P^n(y, z)| \\ &= \inf_{z \in M} |P^m(z, A) - P^m(z, A)| |P^n(x, z) - P^n(y, z)| \\ &= \inf_{z \in M} |P^m(z, A) - P^m(z, A)| \inf_{z \in M} \sup_{z \in M} |P^m(z, A) - P^m(z, A)| \\ &= \frac{1}{2} \inf_{z \in M} |P^n(x, z) - P^n(y, z)| \sup_{z \in M} |P^m(z, A) - P^m(z, A)| \\ &= \frac{1}{2} \inf_{z \in M} |P^n(x, z) - P^n(y, z)| \sup_{z \in M} |P^m(z, A) - P^m(z, A)| \end{aligned}$$

et (1) s'ensuit.

Pour montrer (2), supposons qu'il existe n_0 tel que $\alpha_{n_0} < 1$. Soit alors $m < n_0$ tel que $n = \frac{n}{n_0} n_0 + m$. Alors, par (1), on a

$$\alpha_n \leq \frac{n}{n_0} \alpha_{n_0} + \alpha_m \leq \left(\frac{n}{n_0} \right)^{\frac{n}{n_0}} \alpha_{n_0} + \left(\frac{n}{n_0} \right)^m \alpha_m \leq \left(\frac{n}{n_0} \right)^{\frac{n}{n_0}} \alpha_{n_0} + \left(\frac{n}{n_0} \right)^m \alpha_m.$$

Donc, si on pose $C = \left(\frac{n}{n_0} \right)^{\frac{n}{n_0}} \alpha_{n_0}$ et $\alpha = \frac{n}{n_0} \alpha_m$, on obtient bien $\alpha_n \leq C \alpha^n$.

Pour le dernier point, on observe que

$$|P^{n+m}(x, y) - P^{n+m}(x, y)| = \left| \sum_{z \in M} P^m(x, z) (P^n(z, y) - P^n(x, y)) \right| \leq n C \alpha^n,$$

donc $P^n(x, y)$ est une suite de Cauchy. Soit $\pi(y)$ sa limite. On voit facilement qu'elle ne dépend pas de x , puisque

$$|P^n(x, y) - P^n(x, y)| \leq n C \alpha^n \rightarrow 0.$$

On note la limite de $P^n(x, y)$. On voit que est stationnaire, donc que

$$\begin{aligned} |P^n(x, y) - \pi(y)| &= |P^n(x, y) - P^n(\pi(y))| \\ &= \sum_z (P^n(x, y) - P^n(\pi(y)))^2 \\ &\leq \sum_z P^n(x, z) + P^n(\pi(y)) = 2. \end{aligned}$$

2.1 Chaînes fortement irréductibles

Définition 9. Une matrice de transition P est dite fortement irréductible s'il existe $n_0 \in \mathbb{N}$ tel que pour tous $x, y \in M$ on ait $P^{n_0}(x, y) > 0$.

Proposition 10. Si $|M| < \infty$ et si P est fortement irréductible, alors il existe n_0 tel que $n_0 < \infty$.

Démonstration. Par irréductibilité forte, il existe n_0 tel que $P^{n_0}(x, y) > 0 \quad x, y \in M$. Soit $A = \{z \in M : P^{n_0}(x, z) = P^{n_0}(y, z)\}$ on a que

$$\begin{aligned} n_0(x, y) &= \frac{1}{2} \sum_z |P^{n_0}(x, z) - P^{n_0}(y, z)|^2 \\ &= \frac{1}{2} \sum_{z \in A} (P^{n_0}(x, z) - P^{n_0}(y, z))^2 + \frac{1}{2} \sum_{z \notin A} (P^{n_0}(y, z) - P^{n_0}(x, z))^2 \\ &\leq \frac{1}{2} \sum_{z \in A} P^{n_0}(y, z) + \frac{1}{2} \sum_{z \notin A} P^{n_0}(x, z) < 1 \end{aligned}$$

et puisque M est fini, on en déduit que $n_0 = \sup_{x, y} n_0(x, y) < \infty$.

Et donc, dans ce cas, d'après la formule (4), P^n converge exponentiellement vite vers la unique probabilité stationnaire.

Exemple 11.

1. La matrice

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

est pas fortement irréductible.

2. La matrice

$$\begin{pmatrix} p & 1-p \\ 0 & 1 \end{pmatrix}$$

est pas irréductible.

3. La matrice

$$\begin{pmatrix} p & 1-p \\ 1 & 0 \end{pmatrix}$$

est fortement irréductible si $0 < p < 1$.

2.2 Aperiodicité

Définition 12. Soit $x \in M$ et $R(x) = \{n \in \mathbb{N} : P^n(x, x) > 0\}$. La période de x est le plus grand commun diviseur de $R(x)$.

Proposition 13. Supposons que P est irréductible. Alors tous les points de M ont la même période.

Démonstration. Soient $x, y \in M$ et $n_1 = n(x, y)$, $n_2 = n(y, x)$ tels que $P^{n_1}(x, y) > 0$, $P^{n_2}(y, x) > 0$. Alors

$$P^{n_1+n_2}(x, x) = \sum_{z \in M} P^{n_1}(x, z) P^{n_2}(z, x) \geq P^{n_1}(x, y) P^{n_2}(y, x) > 0$$

donc $n_1 + n_2 \in R(x)$. Si $r \in R(y)$, on a

$$P^{n_1+r+n_2}(x, x) \geq P^{n_1}(x, y) P^r(y, y) P^{n_2}(y, x) > 0$$

donc $n_1 + r + n_2 \in R(x)$. Par définition, la période $p(x)$ est un diviseur de $n_1 + n_2$ et de $n_1 + r + n_2$, donc $p(x)$ divise r , i.e. $p(x) \mid p(y)$. Si on reprend le même argument en échangeant les rôles de x et y , on obtient $p(y) \mid p(x)$. Donc $p(x) = p(y)$, ceci $\forall x, y \in M$.

Donc si P est irréductible, on peut parler de période de P .

Exemple 14.

1. La période de

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

est 2.

2. La période de la marche aléatoire sur \mathbb{Z} est également 2.

Lorsque P est irréductible de période 1, P est dite apériodique.

Proposition 15. Supposons que P est irréductible. S'il existe $x \in M$ tel que $P(x, x) > 0$, alors P est apériodique.

Démonstration. S'il existe $x \in M$ tel que $P(x, x) > 1$ alors $p(x) = 1$. Mais P est irréductible, donc tous les points ont période 1.

Lemme 16. Soit $A \subset \mathbb{N}$ un ensemble stable par addition et de p.g.c.d. égal à 1. Alors $A = \{N, N+1, \dots\}$ pour un certain $N \in \mathbb{N}$.

Démonstration. L'ensemble $B = \{x \in \mathbb{Z} : x \in A \text{ ou } x = 0\}$ est un sous-groupe de \mathbb{Z} et donc de la forme $d\mathbb{Z}$ où d est le plus petit élément non nul de B . De plus $A \subset B$ et par tout élément de A divisible par d on conclut que $d=1$ et donc qu'il existe $a, b \in A \setminus \{0\}$ tels que $a = b + 1$. Or $a \in A$ nécessairement. Si $b = 0$ alors $1 \in A$ et la preuve est terminée. Sinon $N = b^2$ convient puisque si $n \in \mathbb{N}$ on peut écrire $n = b^2 + bq + r$ avec $0 \leq r < b$ et donc $n = b(b + q) + r$ et $r \in A$.

Proposition 17. Une chaîne de Markov irréductible est fortement irréductible si et seulement si elle est apériodique.

Démonstration. Il est clair que une chaîne fortement irréductible est apériodique car pour tout $x \in M$ $R(x) = \{N, N+1, \dots\}$ pour un certain N . Considérons donc une chaîne apériodique et irréductible: l'ensemble $R(x)$ est stable par addition et son p.g.c.d. est 1. Par le lemme précédent il existe $N(x)$ tel que $P^n(x, x) > 0$ pour tout $n \geq N(x)$. Par irréductibilité pour tout $x, y, z \in M$, ils existent n_1 et n_2 tels que $P^{n_1}(x, z) P^{n_2}(z, y) > 0$ et donc $P^{n_1+n_2+N(x)}(x, y) \geq P^{n_1}(x, z) P^{n_2}(z, z) P^{N(x)}(z, y) > 0$ ce qui donne que $P^n(x, y) > 0$ pour tout $n \geq N(x, y) = n_1 + n_2 + N(z)$. Si on pose $N = \max_{x, y} N(x, y) < +\infty$ alors pour tout $x, y \in M$ et $n \geq N$ on a que $P^n(x, y) > 0$ ce qui signifie que la chaîne est fortement irréductible.