Chaînes de Markov :

théorèmes ergodiques

rappel : on a introduit précédemment la notion de mesure invariante d'une chaîne de Markov

irréductible récurrente $\Longrightarrow$ une unique mesure invariante à un coefficient près.

masse infinie
récurrence nulle
temps de retour d'espérance ∞

masse finie
récurrence positive
temps de retour d'espérance finie

1. Période d'une chaîne de Markov.

Definition Soit $(\mathcal{X}, P)$ chaîne de Markov, $x \in \mathcal{X}$.
La période $h(x)$ de l'état x est définie par :

$$h(x) = \mathrm{pgcd}\,\{n \geq 0 \mid P^n(x,x) > 0\}.$$

C'est donc le pgcd de l'ensemble des temps de retour possibles.

<u>remarques</u> : • L'ensemble $R(x) \mid h(x) = \mathrm{pgcd}\, R(x)$ est en général infini ; c'est une partie de $\mathbb{N}$ stable par addition.

$$n, m \in R(x) \Rightarrow n+m \in R(x).$$

• On a l'inclusion $R(x) \subset h(x)\mathbb{N}$.

En effet, $n \in R(x) \Rightarrow h(x)$ divise $n \Rightarrow n \in h(x)\mathbb{N}$.

• En fait, la différence $h(x)\mathbb{N} \setminus R(x)$ est finie

$\Longleftrightarrow$ $R(x)$ contient tous les multiples de $h(x)$ assez grands (caractérise $h(x)$).

Considérons une relation de Bezout :

$$a_1 n_1 + \cdots + a_r n_r - a_{r+1} n_{r+1} - \cdots - a_{r+s} n_{r+s} = h(x)$$

avec les $n_i \in R(x)$, les $a_i \in \mathbb{N}$.

Comme $R(x)$ est stable par addition :

$a = a_1 n_1 + \cdots + a_r n_r$, $b = a_{r+1} n_{r+1} + \cdots + a_{r+s} n_{r+s}$

$a - b = h(x)$ avec $a, b \in R(x)$.

Alors, $R(x)$ contient tous les multiples de $h(x)$ à partir de ab :

$$ab \in R(x)$$

$$ab + h = ab + a - b = b(a-1) + a \in R(x)$$

$$ab + 2h = ab + 2a - 2b = b(a-2) + 2a \in R(x)$$

$$\vdots \qquad\qquad \vdots$$

Si $k = qa + r$ (division euclidienne, $r < a$)

$$ab + kh = (qh)a + ab + r(a-b)$$
$$= \underbrace{b(a-r)}_{\in R(x)} + a\underbrace{(qh+r)}_{\in R(x)} \in R(x).$$

$\square$

<u>exemple</u> : si $R(x) = 3\mathbb{N} + 5\mathbb{N}$

$$= \{0, 3, 5, 6, 8, 9, 10, 11, \dots\}$$

alors R contient tous les multiples de $1 = \text{pgcd}(3,5)$ assez grand.

<u>Proposition</u> Si la chaîne de Markov est irréductible, alors $h(x)$ ne dépend pas de l'état x.

$h(x) = h(P) =$ période de la chaîne.

Preuve : Soit $x \neq y$ deux états.

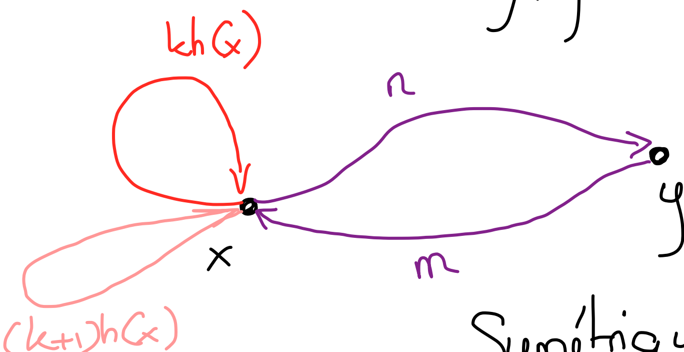
$\exists n, m \mid P^n(x,y) > 0$

$P^m(y,x) > 0.$

Soit $k \mid kh(x)$ et $(k+1)h(x) \in R(x)$ (existe car $kh(x) \in R(x)$ pour k assez grand)

$P^{kh(x)}(x,x) > 0, \quad P^{(k+1)h(x)} \in R(x)$

$\Longrightarrow P^{n+m+kh(x)}(y,y) > 0$ et $P^{n+m+(k+1)h(x)}(y,y) > 0$



Alors, $h(y)$ divise $\mathrm{pgcd}(n+m+kh(x), n+m+(k+1)h(x)) = h(x)$

Symétriquement, $h(x) \mid h(y) \Rightarrow h(x) = h(y)$ $\square$.

Une chaîne irréductible est dite apériodique si $h(P) = 1$.

$\iff \forall x, y, \quad P^n(x,y) > 0$ pour $n \geq n_{xy}$ assez grand.

Une condition suffisante simple pour cela est :

$\exists x \mid P(x,x) > 0$

$\Rightarrow 1 \in R(x) \Rightarrow h(x) = \text{pgcd}(R(x)) = 1$.

2. Convergence vers la loi stationnaire.

$(\mathcal{X}, P)$ chaîne de Markov irréductible ; on regarde $(X_n)_{n \in \mathbb{N}}$ sous la loi $\mathbb{P}_{\pi_0}$ avec π_0 loi initiale arbitraire.

$$\pi_n(x) = \mathbb{P}_{\pi_0}[X_n = x] = (\pi_0 P^n)(x).$$

Théorème 1) Si la chaîne est $\left|\begin{array}{l}\text{récurrente positive de loi stationnaire } \pi, \\ \text{apériodique}\end{array}\right.$

alors $\pi_n(x) \underset{n \to +\infty}{\longrightarrow} \pi(x) \quad \forall x \in \mathcal{X}$.

2) Si la chaîne est récurrente nulle ou transiente, alors $\pi_n(x) \underset{n \to +\infty}{\longrightarrow} 0$.

Remarque L'apériodicité est nécessaire :

si $\mathcal{X} = \{1, 2\}$ et $G_P = $
$$\begin{tikzcd} 1 \arrow[r, bend left, "1"] & 2 \arrow[l, bend left, "1"] \end{tikzcd}$$
, sous $\mathbb{P}_{\delta_1}$,

$$X_n = \begin{cases} 1 \text{ p.s. si } n \text{ pair}, \\ 2 \text{ p.s. si } n \text{ impair}. \end{cases}$$

donc $\pi_n \not\to$ loi invariante $\dfrac{\delta_1 + \delta_2}{2}$.

Ici on a une chaîne de période 2.

On va démontrer le point 1), qui est le plus important.

Introduisons la distance en variation totale entre deux mesures de probabilité π et μ sur $\mathcal{X}$.

$$d_{VT}(\pi, \mu) = \sup_{A \subset \mathcal{X}} |\pi(A) - \mu(A)|$$

On veut montrer que $d_{VT}(\pi_n, \pi) \xrightarrow[n \to +\infty]{} 0$.

Argument de couplage : on considère deux chaînes de Markov indépendantes $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ de matrice P, de lois initiales π_0 et μ_0.

La paire $(X_n, Y_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $\mathcal{X}^2$.

Sa matrice est $P^{\otimes 2}((x,y),(\bar{x},\bar{y})) = P(x,\bar{x})\, P(y,\bar{y})$.

Sa loi initiale est $\pi_0(x)\, \mu_0(y)$.

<u>Lemme</u> : Si P correspond à une chaîne irréductible récurrente positive et apériodique, il en va de même pour la matrice $P^{\otimes 2}$ de la chaîne $(X_n, Y_n)_{n\in\mathbb{N}}$.

<u>Preuve</u> : Pour une chaîne irréductible apériodique, on sait que

$$\left.\begin{array}{l} \forall x, \bar{x}, \quad P^n(x,\bar{x}) > 0 \\ \forall y, \bar{y}, \quad P^n(y,\bar{y}) > 0 \end{array}\right\} \text{pour } n \text{ assez grand.}$$

Alors, $\forall (x,y), (\bar{x},\bar{y})$, $(P^{\otimes 2})^n((x,y),(\bar{x},\bar{y})) > 0$ pour n assez grand.

Ceci prouve le caractère irréductible apériodique de $P^{\otimes 2}$.

Si π est la loi stationnaire de P, alors

$$\pi^{\otimes 2}(x,y) = \pi(x)\,\pi(y) \text{ est } \begin{cases} \text{de masse finie} \\ \text{invariante par } P^{\otimes 2} \end{cases}.$$

L'existence d'une mesure invariante de masse finie implique la récurrence positive. $\square$.

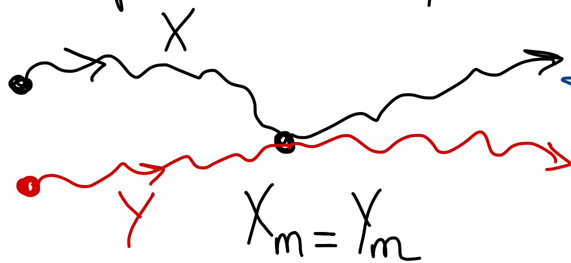
On pose
$$T = \inf\{n \in \mathbb{N} \mid X_n = Y_n\}$$
$$= \inf\{n \in \mathbb{N} \mid (X_n, Y_n) \in \text{diagonale de } \mathcal{X}^2\}$$
$$< +\infty \quad \text{p.s.} \quad \text{sous } \mathbb{P}_{\pi_0 \otimes \mu_0}.$$

<u>Lemme 2</u> : $d_{TV}(\pi_0 P^n, \mu_0 P^n) \leq \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n]$.

<u>Preuve</u> : notons que, pour tout état $x \in \mathcal{X}$:
tous temps $n \geq m$

$$\mathbb{P}_{\pi_0 \otimes \mu_0}\left[T = m \text{ et } X_n = x\right] = \mathbb{P}_{\pi_0 \otimes \mu_0}\left[T = m \text{ et } Y_n = x\right]$$

preuve intuitive : après coalescence, X et Y jouent un rôle symétrique



les deux trajectoires X et Y sont échangeables après le temps m.

preuve rigoureuse : les deux probabilités s'écrivent :

$$\sum_{x_0 \neq y_0, \dots, x_{m-1} \neq y_{m-1}, z_m} \pi_0(x_0) P(x_0, x_1) \dots P(x_{m-1}, z_m) \mu_0(y_0) P(y_0, y_1) \dots P(y_{m-1}, z_m) P^{n-m}(z_m, x).$$

Alors, $\forall x \in \mathcal{X}$:

$$\pi_n(x) = \mathbb{P}_{\pi_0}[X_n = x]$$
$$= \sum_{m=0}^{n} \mathbb{P}_{\pi_0 \otimes \mu_0}[T = m \text{ et } X_n = x] + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n \text{ et } X_n = x]$$
$$= \sum_{m=0}^{n} \mathbb{P}_{\pi_0 \otimes \mu_0}[T = m \text{ et } Y_n = x] + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n \text{ et } X_n = x]$$
$$= \mathbb{P}_{\pi_0 \otimes \mu_0}[T \leq n \text{ et } Y_n = x] + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n \text{ et } X_n = x]$$

puis, $\forall A \subset \mathcal{X}$:

$$\pi_n(A) = \mathbb{P}_{\pi_0 \otimes \mu_0}[T \leq n \text{ et } Y_n \in A] + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n \text{ et } X_n \in A]$$
$$\leq \mu_n(A) + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n].$$

Symétriquement, $\mu_n(A) \leq \pi_n(A) + \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n]$,

d'où : $\sup_{A \subset \mathcal{X}} |\mu_n(A) - \pi_n(A)| \leq \mathbb{P}_{\pi_0 \otimes \mu_0}[T > n]$ $\square$.

Fin de la preuve (cas récurrent positif apériodique) :

Prenons $\mu_0 = \pi$ unique loi invariante.

$$\mu_n = \pi \quad \forall n \in \mathbb{N}$$

et $d_{VT}(\pi_n, \pi) \leq \mathbb{P}_{\pi_0 \otimes \pi}[T > n] \xrightarrow[n \to +\infty]{} 0$

car T fini p.s. (on a une chaîne récurrente positive, qui visite tout le monde, donc la diagonale). $\square$.

3. Fréquences asymptotiques des visites

Les mesures stationnaires des CM sont également impliquées dans l'asymptotique des fréquences des visites.

Théorème : Soit $(\mathcal{X}, P)$ une chaîne irréductible.
On considère $(X_n)_{n\in\mathbb{N}}$ sous $\mathbb{P}_{\pi_0}$, π_0 arbitraire.

1) Si la chaîne est récurrente positive de loi stationnaire π, alors

$$\frac{1}{n} V_{x,n} = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}(X_k = x) \xrightarrow[p.s.]{} \pi(x) \quad \forall x \in \mathcal{X}.$$

2) Si la chaîne est récurrente nulle ou transiente, alors

$$\frac{1}{n} V_{x,n} \xrightarrow[p.s.]{} 0.$$

Preuve : le cas transient est très facile ($(V_{x,n})_{n\in\mathbb{N}}$ reste bornée).

Dans le cas récurrent, supposons pour simplifier $\pi_0 = \delta_x$.

Alors, $V_{x,n} = \sum_{k=1}^{\infty} \mathbb{1}_{\tau_x^{(k)} \leq n} = \max\{k : \tau_x^{(k)} \leq n\}$

$$\boxed{t \leq \frac{V_{x,n}}{n} \leq u} \quad (*) \iff \tau_x^{\lceil tn \rceil} \leq n \leq \tau_x^{\lfloor un \rfloor}$$

$$\iff \frac{\tau_x^{\lceil tn \rceil}}{n} \leq 1 \leq \frac{\tau_x^{\lfloor un \rfloor}}{n}.$$

$$\tau_x^{(k)} = \sum_{i=1}^{k} \underbrace{\tau_x^{(i)} - \tau_x^{(i-1)}}_{\text{somme de variables iiD}} \longrightarrow \text{loi des grands nombres}$$

$$\forall t, \ \frac{\tau_x^{\lfloor tn \rfloor}}{n} \longrightarrow t \cdot \mathbb{E}_x[\tau_x^+] = \begin{cases} t/\pi(x) & \text{récurrent positif} \\ +\infty & \text{récurrent nul} \end{cases}$$

Cas positif : asymptotiquement, (*) devient vrai avec proba 1 pour tout $t < \pi(x)$ et tout $u > \pi(x)$

$$\Longrightarrow \frac{V_{x,n}}{n} \xrightarrow[p.s.]{} \pi(x).$$

Cas nul : asymptotiquement, $\frac{V_{x,n}}{n} \leq u$ devient vrai pour tout $u > 0$

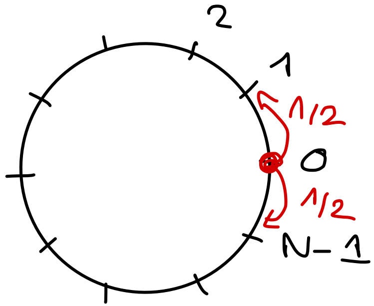
$$\Longrightarrow \frac{V_{x,n}}{n} \xrightarrow[p.s.]{} 0.$$

$\square$

remarque : théorème "ergodique" : moyenne temporelle = moyenne spatiale.

exemples

1. marche aléatoire sur le cercle



$$\mathcal{X} = \mathbb{Z}/N\mathbb{Z} \doteq [\![0, N-1]\!]$$

La chaîne est irréductible finie, donc récurrente positive.

apériodicité ? oui si N impair

mesure invariante : la loi uniforme $\pi(x) = \frac{1}{N} \quad \forall x \in \mathcal{X}$.

- convergence vers la loi stationnaire si N impair :

$$\mathbb{P}_0[X_n = k] \xrightarrow[n \to +\infty]{} \frac{1}{N} \qquad \forall k \in \mathbb{Z}/N\mathbb{Z}.$$

- théorème ergodique :

$$\forall k, \quad \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i = k) \xrightarrow[p.s]{} \frac{1}{N}.$$

Ceci implique, pour toute fonction $f : \mathbb{Z}/_{N\mathbb{Z}} \longrightarrow \mathbb{C}$:

$$\underbrace{\frac{1}{n} \sum_{i=1}^{n} f(X_i)}_{\text{moyenne temporelle}} = \sum_{k=1}^{N} f(k) \frac{V_{k,n}}{n} \xrightarrow[p.s.]{} \underbrace{\frac{1}{N} \sum_{k=1}^{N} f(k)}_{\text{moyenne spatiale}}.$$

2. marche aléatoire sur $\mathbb{Z}$

si $p \neq \frac{1}{2}$, on a exhibé deux mesures invariantes non proportionnelles

$\Longrightarrow$ la marche est irréductible transiente.

si $p = \frac{1}{2}$, on peut montrer la récurrence par le critère numérique

$$\left(P^{2n}(0,0) = O\left(\frac{1}{\sqrt{n}}\right), \sum_{n=1}^{\infty} P^{2n}(0,0) = +\infty \right)$$

Comme la mesure de comptage $\pi(x) = 1 \ \forall x \in \mathbb{Z}$ est invariante, on est dans le cas récurrent nul

$\Longrightarrow$ • $\forall x, \ \pi_n(x) \xrightarrow[n \to +\infty]{} 0$.

• les temps de retour sont finis p.s, mais d'espérance infinie.

• la fréquence asymptotique des visites d'un état tend vers 0.