Partie III: chaînes de Markov

Christophe Sabot, Université Lyon 1
Notes informelles de cours

May 4, 2020

Principe, heuristique

$(X_n)_{n\geqslant 0}$ suite de variables aléatoires, la propriété de Markov correspond à dire que X_{n+1} ne dépend du passé qu'à travers X_n.

$$\mathbb{P}(X_{n+1} \in A \mid X_0, \ldots, X_n) = \mathbb{P}(X_{n+1} \in A \mid X_n).$$

On supposera que la suite est à valeurs dans E fini ou dénombrable. Ce qui détermine la loi de la chaîne c'est donc:

$$\mathbb{P}(X_{n+1} = y \mid X_n = x) := Q_n(x, y).$$

On supposera que la chaîne est **homogène**, c'est-à-dire que $Q_n(x, y)$ ne dépend pas de n. On notera $Q := Q_n$.

Donc pour tout x, $Q(x, \cdot)$ sera une probabilité.

1- Définition, premières propriétés

Définition

Soit E fini ou dénombrable (muni de la tribu des parties $\mathcal{P}(E)$). (i) On appelle matrice stochastique une matrice $(Q(x,y))_{x,y\in E}$ telle que

$$0 \leqslant Q(x,y) \leqslant 1$$

$$\sum_{y\in E} Q(x,y) = 1, \quad \forall x \in E$$

Donc pour chaque $x \in E$, $Q(x,\cdot)$ est une probabilité.
(ii) Une suite de v.a. $(X_n)_{n\geqslant 0}$ à valeurs dans E est une chaîne de Markov de matrice de transition Q si

$$\mathbb{P}(X_{n+1} = y \mid X_0, \dots, X_n) = Q(X_n, y), \quad p.s.$$

Pour tout entier n et tout $y \in E$.

1- Définition, premières propriétés

Définition

Soit E fini ou dénombrable (muni de la tribu des parties $\mathcal{P}(E)$). (i) On appelle matrice stochastique une matrice $(Q(x,y))_{x,y\in E}$ telle que

$$0 \leqslant Q(x,y) \leqslant 1$$

$$\sum_{y\in E} Q(x,y) = 1, \quad \forall x \in E$$

Donc pour chaque $x \in E$, $Q(x,\cdot)$ est une probabilité.
(ii) Une suite de v.a. $(X_n)_{n\geqslant 0}$ à valeurs dans E est une chaîne de Markov de matrice de transition Q si

$$\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \dots, X_n = x_n) = Q(x_n, y),$$

Pour tout entier n et tout $x_0, \dots, x_n, y \in E$, dès que $\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) > 0$.

Proposition

$(X_n)_{n\geqslant 0}$ est une chaîne de Markov (noté CM) de matrice de transition Q si $\forall x_0, \ldots, x_n$ dans E:

$$\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_0 = x_0) Q(x_0, x_1) \cdots Q(x_{n-1}, x_n).$$

En particulier, si $\mathbb{P}(X_0 = x) > 0$, alors

$$\mathbb{P}(X_n = y \mid X_0 = x) = Q^n(x, y),$$

où Q^n est la puissance n-ième de la matrice sto Q:

$$Q^n(x, y) = \sum_{z_1, \ldots, z_{n-1}} Q(x, z_1) Q(z_1, z_2) \ldots Q(z_{n-1}, y).$$

Preuve: par récurrence, en notant que

$$\begin{aligned}
&\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \\
=&\mathbb{P}(X_n = x_n \mid X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) \mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) \\
=&Q(x_{n-1}, x_n) \mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1})
\end{aligned}$$

On appelle $\mu(x) := \mathbb{P}(X_0 = x)$ la **loi intiale** de la CM.
On notera aussi Q pour l'opérateur sur les fonctions

$$\begin{array}{rcl} Q : \mathbb{R}^E & \mapsto & \mathbb{R}^E \\ f & \to & Qf \end{array}$$

avec

$$Qf(x) = \sum_{y \in E} Q(x,y) f(y).$$

On a alors

$$\mathbb{E}(f(X_{n+1}) \mid X_0, \ldots, X_n) = Qf(X_n), \quad p.s.$$

$$\mathbb{E}(f(X_n) \mid X_0 = x) = Q^n f(x).$$

Facile à vérifier: exo.

Théorème

Pour toute matrice sto Q, pour toute probabilité μ sur E, $\exists$ un espace de proba $(\Omega, \mathcal{F}, \mathbb{P})$ et une suite de v.a. $(X_n)_{n\in\mathbb{N}}$ sur cet espace, tel que (X_n) soit une CM de matrice de transition Q et de loi initiale μ.

Preuve. Nous donnons une preuve constructive de ce résultat, qui est essentiellement basée sur la construction d'une v.a. de loi fixée sur E à partir d'une v.a. de loi uniforme sur E. Ce résultat peut aussi être conséquence du théorème d'extension de Kolmogorov.

Lemme

Soit ν une proba sur E. Il existe une fonction $f^\nu : [0,1[\mapsto E$ telle que si U est une variable aléatoire uniforme sur $[0,1]$, alors

$$f^\nu(U) \text{ a pour loi } \nu.$$

Preuve du lemme. E étant fini ou dénombrable on peut ordonner ses éléments:

$$Y = \{y_i\}_{i \in I}, \quad I = \{0, \ldots, |E|\} \text{ ou } I = \mathbb{N}.$$

On définit alors pour $u \in [0, 1[$

$$f^{\nu}(u) = y_i \text{ si } u \in \left[\sum_{j=0}^{i-1} \nu(y_j), \sum_{j=0}^{i} \nu(y_j)\right[$$

Rq: les intervalles décrits forment clairement une partition de $[0, 1[$ lorsque i parcours I, la fonction f^{ν} est donc bien définie sur $[0, 1[$.

Si U est uniforme sur $[0, 1]$ alors pour $i \in I$:

$$\mathbb{P}(f^{\nu}(U) = y_i) = \mathbb{P}(U \in \left[\sum_{j=0}^{i-1} \nu(y_j), \sum_{j=0}^{i} \nu(y_j)\right[)$$
$$= \nu(y_i).$$

Preuve du Théorème. On se donne une suite $(U_i)_{i\in\mathbb{N}}$ de v.a. i.i.d. de loi uniforme sur $[0,1]$, définies sur un espace de proba $(\Omega,\mathcal{F},\mathbb{P})$. (C'est possible, cf cours de L3).

On définit alors la suite $(X_n)_{n\in\mathbb{N}}$ par recurrence par

$$X_0 = f^{\mu}(U_0) \text{ et } X_{n+1} = f^{Q(X_n,\cdot)}(U_{n+1}).$$

Alors par le lemme précédent, X_0 a la loi μ et

$$\begin{aligned}\mathbb{P}(X_{n+1}=y \mid X_0,\ldots,X_n) &= \mathbb{P}(f^{Q(X_n,\cdot)}(U_{n+1})=y \mid X_n)\\ &= Q(X_n,y).\end{aligned}$$

Donc (X_n) est une CM de loi initiale μ et de mat. de transition Q.

Remarque: cette preuve donne aussi une méthode générale pour construire une CM à partir d'une suite de v.a. i.i.d., méthode qui peut être utile pour la simulation de CMs.

Examples classiques

Example 1: (X_n) suite de v.a. i.i.d. à valeurs dans E de loi ν.
C'est une CM de loi initiale ν et de mat. de trans. $Q(x, y) = \nu(y)$

Example 2: Marche aléatoire sur $\mathbb{Z}^d$ issue de $a \in \mathbb{Z}^d$:

$X_n = a + Y_1 + \ldots + Y_n$ avec $(Y_i)_{i \geqslant 1}$ i.i.d. à valeurs dans $\mathbb{Z}^d$ de loi ν

Alors $\mu = \delta_a$ et $Q(x, y) = \nu(y - x)$.

Example 2-bis: Marche aléatoire simple (MAS) sur $\mathbb{Z}^d$ issue de $a \in \mathbb{Z}^d$: idem avec $(Y_i)_{i \geqslant 1}$ i.i.d. et

$$\mathbb{P}(Y_k = \pm e_i) = \frac{1}{2d}, \quad \forall i = 1, \ldots, e_d.$$

où $\{e_1, \ldots, e_d\}$ base canonique de $\mathbb{Z}^d$.

Example 3: Processus de Galton-Watson. Rappel:

$$(\xi_k^n)_{n\in\mathbb{N},k\in\mathbb{N}^*} \text{ i.i.d. à valeurs dans } \mathbb{N} \text{ de loi } \mu$$

ξ_k^n est le nombre de descendants du k-ième individu vivant au temps n. Alors $(Z_n)_{n\geqslant 0}$ défini par $Z_0 = 1$ et

$$Z_{n+1} = \sum_{k=1}^{Z_n} \xi_k^n.$$

Alors (Z_n) est une CM à valeurs dans $\mathbb{N}$ de mat. de trans.

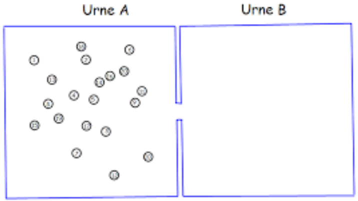
$$Q(x,y) = \mu^{*x}(y).$$

où μ^{*x} est la loi convoluée x fois de μ. (exo: facile par rec).

Rappel: *Si μ et ν sont deux lois sur $\mathbb{Z}$, $\mu * \nu$ est la loi*

$$\mu * \nu(z) = \sum_k \mu(k)\nu(z-k),$$

c'est la loi de $X+Y$ si X et Y sont indépendants de lois μ et ν resp.



Example 4: Urnes d'Ehrenfest.
N boules dans deux urnes A et B. A chaque temps entier une boule prise au hasard change d'urne. On note $(X_n)_{n\geqslant 0}$ le nombre de boules au temps n dans l'urne A.

A chaque temps, (X_n) change de ± 1, c'est une CM et

$$\begin{cases} Q(x, x+1) = \frac{N-x}{N} \\ Q(x, x-1) = \frac{x}{N} \end{cases}$$

2-Espace canonique. Propriété de Markov.

Soit E espace d'état fini ou dénombrable. On note

$$\Omega = E^{\mathbb{N}} = \{w = (w_0, \ldots, w_n, \ldots),\ \omega_k \in E\},$$

l'ensemble des suites (trajectoires) à valeurs dans E. On appelle n-cylindre un ensemble du type, pour $(x_0, \ldots, x_n) \in E^{n+1}$,

$$C_{x_0,\ldots,x_n} = \{w \in \Omega,\ \ w_0 = x_0, \ldots, w_n = x_n\}.$$

On définit X_n, $n \geqslant 0$, l'application $X_n : \Omega \mapsto E$ définie par

$$X_n(\omega) = \omega_n.$$

C'est l'application n-ième coordonnée. On note $\mathcal{F}_n$ la tribu engendrée par les n-cylindres, c'est donc aussi

$$\mathcal{F}_n = \sigma\{C_{x_0,\ldots,x_n},\ x_0, \ldots, x_n \in E\} = \sigma\{X_0, \ldots, X_n\} \overset{\sim}{=} \mathcal{P}(E)^{\otimes n+1},$$

la plus petite tribu qui rend mesurable $X_0, \ldots, X_n$, et

$$\mathcal{F} = \sigma\{\mathcal{F}_n,\ n \geqslant 0\} = \sigma\{X_k,\ k \geqslant 0\} = \mathcal{P}(E)^{\otimes \mathbb{N}},$$

qui est la tribu produit sur les trajectoires.

Proposition

Soit Q mat. sto. sur $E \times E$. Pour tout x, $\exists!$ probabilité $\mathbb{P}_x$ sur $(\Omega, \mathcal{F})$ telle que sous $\mathbb{P}_x$, $(X_n)_{n\in\mathbb{N}}$ est une CM de mat. de trans. Q et de loi intiale δ_x.
Si μ proba sur E, $\mathbb{P}_\mu := \sum_{x\in E} \mu(x)\mathbb{P}_x$ telle que $(X_n)_{n\in\mathbb{N}}$ est une CM de mat. de trans. Q et de loi initiale μ.
On notera $\mathbb{E}_x$ et $\mathbb{E}_\mu$ l'espérance associée aux probabilités $\mathbb{P}_x$ et $\mathbb{P}_\mu$.

Preuve. Existence. Par le Théorème 1, on peut construire un espace de proba $(\Omega', \mathcal{F}', \mathbb{P}')$ et une suite de v.a. (X'_n) à valeurs dans E telles que (X'_n) est une CM de mat de trans. Q et de loi initiale δ_x.
Alors

$$(X'_n) : \Omega' \mapsto \Omega$$

est $\mathcal{F}$-mesurable (car X'_n est $\mathcal{F}_n$-mesurable forcément). On définit $\mathbb{P}_x$ comme la loi de (X'_n) (càd la loi image de $\mathbb{P}'$ par (X'_n)). Alors,

$$\begin{aligned}\mathbb{P}_x(X_0 = x_0, \dots, X_n = x_n) &= \mathbb{P}'(X'_0 = x'_0, \dots, X'_n = x'_n) \\ &= \mathbb{1}_{x=x_0} Q(x_0, x_1) \dots Q(x_{n-1}, x_n).\end{aligned}$$

Donc, sous $\mathbb{P}_x$, la suite de v.a. canonique (X_n) est une CM partant de x et de mat. de trans. Q.

Preuve. Unicité. Soit $\tilde{\mathbb{P}}$ proba sur l'espace canonique $(\Omega, \mathcal{F})$ sous laquelle (X_n) est une CM de loi initiale δ_x et de mat. trans. Q. Alors, $\forall n \geqslant 0$, $\forall x_0, \ldots, x_n$ dans E:

$$\begin{aligned}\tilde{\mathbb{P}}(X_0 = x_0, \ldots, X_n = x_n) &= \mathbb{1}_{x=x_0} Q(x_0, x_1) \ldots Q(x_{n-1}, x_n), \\ &= \mathbb{P}_x(X_0 = x_0, \ldots, X_n = x_n)\end{aligned}$$

donc $X_0, \ldots, X_n$ ont même loi sous $\mathbb{P}_x$ et $\tilde{\mathbb{P}}$, donc $\mathbb{P}_x$ et $\tilde{\mathbb{P}}$ sont égales sur $\mathcal{F}_n$, $\forall n \geqslant 0$.
Comme $\mathcal{F} = \sigma\{\mathcal{F}_n,\ n \in \mathbb{N}\}$, $\tilde{\mathbb{P}}$ et $\mathbb{P}_x$ sont égales.

Remarques: Sur l'espace canonique $(\Omega, \mathcal{F})$, on peut donc définir une famille de probabilités $(\mathbb{P}_\mu)$ pour μ proba sur E, sous lesquelles (X_n) est une CM de loi initiale μ et mat. de trans. Q. Il sera pratique de travailler sur l'espace canonique et de disposer de cette famille de loi. C'est ce que nous ferons dans la suite.

Sur l'espace canonique on définit le décalage en temps $\theta : \Omega \mapsto \Omega$,

$$\theta((\omega_0, \ldots, \omega_n, \ldots)) = (\omega_1, \ldots, \omega_n, \ldots)$$

et le k-décalage, θ^k:

$$\theta^k((\omega_0, \ldots, \omega_n, \ldots)) = (\omega_k, \ldots, \omega_n, \ldots).$$

Théorème (Propriété de Markov)

Soit F et G deux fonctions mesurables sur $(\Omega, \mathcal{F})$. Alors si F est $\mathcal{F}_n$-mesurable:

$$\mathbb{E}_\mu \left(F \cdot G \circ \theta^n\right) = \mathbb{E}_\mu \left(F \cdot \mathbb{E}_{X_n}\left(G\right)\right)$$

autrement dit

$$\mathbb{E}_\mu \left(G \circ \theta^n \mid \mathcal{F}_n\right) = \mathbb{E}_{X_n}\left(G\right)$$

dès que chaque terme a un sens (intégrable ou positif).

Autrement dit, ça veut dire que pour f et g mesurables:

$$\begin{aligned} &\mathbb{E}_\mu \left(f(X_0, \ldots, X_n) \cdot g((X_{n+k})_{k\geqslant 0})\right) \\ =&\mathbb{E}_\mu \left(f(X_0, \ldots, X_n) \cdot \mathbb{E}_{X_n}\left(g((X_k)_{k\geqslant 0})\right)\right) \end{aligned}$$

Preuve.

On commence par le montrer pour des indicatrices de trajectoires:

$$F = \mathbb{1}_{X_0=x_0,\dots,X_n=x_n}, \quad G = \mathbb{1}_{X_0=y_0,\dots,X_k=y_k},$$

pour n entier donné dans l'énoncé et k entier. Alors,

$$\begin{aligned}
&\mathbb{E}_\mu\left(F \cdot G \circ \theta^n\right) \\
=&\mathbb{E}_\mu\left(\mathbb{1}_{X_0=x_0,\dots,X_n=x_n} \mathbb{1}_{X_n=y_0,\dots,X_{n+k}=y_k}\right) \\
=&\mathbb{E}_\mu\left(\mathbb{1}_{X_0=x_0,\dots,X_n=x_n} \mathbb{E}_\mu\left(\mathbb{1}_{X_n=y_0,\dots,X_{n+k}=y_k} \mid X_0 = x_0, \dots, X_n = x_n\right)\right) \\
=&\mathbb{E}_\mu\left(\mathbb{1}_{X_0=x_0,\dots,X_n=x_n} \mathbb{E}_\mu\left(\mathbb{1}_{X_n=y_0,\dots,X_{n+k}=y_k} \mid X_n = x_n\right)\right) \\
=&\mathbb{E}_\mu\left(\mathbb{1}_{X_0=x_0,\dots,X_n=x_n} \mathbb{E}_{x_n}\left(\mathbb{1}_{X_0=y_0,\dots,X_k=y_k}\right)\right) \\
=&\mathbb{E}_\mu\left(F \cdot \mathbb{E}_{X_n}\left(G\right)\right)
\end{aligned}$$

Donc la propriété est vraie pour tout $F = \mathbb{1}_A$ et $G = \mathbb{1}_B$ pour $A \in \mathcal{F}_n$ et $B \in \mathcal{F}_k$, $\forall k \geqslant 0$. Par classes monotones, la propriété est vraie pour tout B dans $\mathcal{F}$. Puis vraie pour tout F $\mathcal{F}_n$-mesurable et G $\mathcal{F}$-mesurable par approximation par des fonctions plateau.

Théorème (Propriété de Markov fort)

Soit T un $(\mathcal{F}_n)$-temps d'arrêt. Soit F et G deux fonctions mesurables sur $(\Omega, \mathcal{F})$. Alors si F est $\mathcal{F}_T$-mesurable:

$$\mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} F \cdot G \circ \theta^T\right) = \mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} F \cdot \mathbb{E}_{X_T}(G)\right)$$

autrement dit

$$\mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} G \circ \theta^T \mid \mathcal{F}_T\right) = \mathbb{1}_{T<\infty} \mathbb{E}_{X_T}(G)$$

dès que chaque terme a un sens (intégrable ou positif).

Autrement dit, ça veut dire que pour f et g mesurables:

$$\begin{aligned} &\mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} f(X_0, \dots, X_T) \cdot g((X_{T+k})_{k\geqslant 0})\right) \\ =&\mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} f(X_0, \dots, X_T) \cdot \mathbb{E}_{X_T}\left(g((X_k)_{k\geqslant 0})\right)\right) \end{aligned}$$

Preuve.

On commence par $F = \mathbb{1}_A$ avec $A \in \mathcal{F}_T$.

$$\begin{aligned}
\mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} F \cdot G \circ \theta^T\right) &= \sum_{n=0}^{\infty} \mathbb{E}_\mu\left(\mathbb{1}_{\{T=n\}\cap A} G \circ \theta^n\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_\mu\left(\mathbb{1}_{\{T=n\}\cap A} \mathbb{E}_{X_n}(G)\right) \\
&= \mathbb{E}_\mu\left(\mathbb{1}_{T<\infty} F \cdot \mathbb{E}_{X_n}(G)\right)
\end{aligned}$$

en appliquant la propriété de Markov (faible) dans la 2-ième égalité car $\{T = n\} \cap A \in \mathcal{F}_n$ car $A \in \mathcal{F}_T$.

On conclue pour tout F $\mathcal{F}$-mesurable par approximation par fonctions plateaux.

3- Récurrence, classification des états.

Soit Q mat. sto. sur E, on se placera par défaut sur l'espace canonique $(\Omega, \mathcal{F})$ muni de la famille de probabilités $(\mathbb{P}_x)_{x\in E}$ avec (X_n) les v.a. canoniques. Donc sous $\mathbb{P}_x$, (X_n) est une CM de mat. trans. Q partant de x.

On note $N(x)$ la v.a. à valeurs dans $\mathbb{N} \cup \{\infty\}$ qui donne le nombre de visites en $x \in E$

$$N(x) := \sum_{n=0}^{\infty} \mathbb{1}_{X_n = x}.$$

On appelle $G = (G(x,y))_{x,y\in E}$ la **fonction de Green** (ou **matrice potentielle**),

$$G(x,y) = \mathbb{E}_x\left(N(y)\right), \quad \forall x, y \in E.$$

On a donc

$$G(x,y) = \sum_{n=0}^{\infty} \mathbb{E}_x\left(\mathbb{1}_{X_n = y}\right) = \sum_{n=0}^{\infty} Q^n(x,y).$$

De façon générale $G(x,y)$ est à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$.

Soit $x \in E$, on note

$$\tau_x = \inf\{n \geqslant 0,\ X_n = x\}, \quad \tau_x^+ = \inf\{n > 0,\ X_n = x\}.$$

resp. le premier temps de visite de x ou le premier temps strictement positif de visite en x. Sous $\mathbb{P}_x$, càd si la CM part de x, τ_x^+ est le premier temps de retour en x.

Théorème

Soit $x \in E$, on a équivalence entre les 3 énoncés suivants

$$\big(\mathbb{P}_x(\tau_x^+ < \infty) = 1\big) \Leftrightarrow (N(x) = \infty,\ \mathbb{P}_x \ p.s.) \Leftrightarrow (G(x,x) = \infty)$$

ainsi que les 3 énoncés complémentaires suivants

$$\big(\mathbb{P}_x(\tau_x^+ < \infty) < 1\big) \Leftrightarrow (N(x) < \infty,\ \mathbb{P}_x \ p.s.) \Leftrightarrow (G(x,x) < \infty)$$

Dans le premier cas, on dit que x est un point **récurrent**, *dans le le deuxième cas on dit que x est un point* **transitoire**.

Lemme

Pour tout x, y dans E, tout $n \geqslant 1$,

$$\mathbb{P}_x(N(y) \geqslant n) = \mathbb{P}_x(\tau_y < \infty)\mathbb{P}_y(\tau_y^+ < \infty)^{n-1}.$$

Preuve du Théorème. On déduit du lemme appliqué à $x = y$ que

$$\mathbb{P}_x(N(x) \geqslant n) = \mathbb{P}_x(\tau_x^+ < \infty)^{n-1}.$$

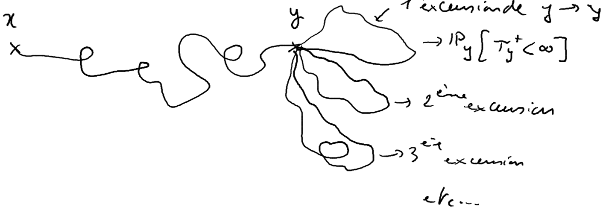
Donc si $\mathbb{P}_x(\tau_x^+ < \infty) = 1$ alors $N(x) = \infty$ $\mathbb{P}_x$ p.s. et donc $G(x, x) = \mathbb{E}_x(N(x)) = \infty$. D'autre part, par la formule classique

$$G(x, x) = \mathbb{E}_x(N(x)) = \sum_{n=1}^{\infty} \mathbb{P}_x(N(x) \geqslant n) = \sum_{n=1}^{\infty} \mathbb{P}_x(\tau_x^+ < \infty)^{n-1}.$$

Donc $G(x, x) = \infty$ implique $\mathbb{P}_x(\tau_x^+ < \infty) = 1$ (sinon la somme est géométrique finie).
On a donc montré l'équivalence des 3 premiers énoncés. La preuve est similaire pour les 3 autres (exo).

Preuve du lemme : <u>idée</u> .

$$\mathbb{P}_x\left[N(y) \geq n\right] = \mathbb{P}_x\left[T_y < +\infty\right] \mathbb{P}_y\left[T_y^+ < +\infty\right]^{n-1}$$



Preuve du Lemme.

On considère les temps de visites successifs de y.

$$\begin{cases} \tau_y^1 = \tau_y \\ \tau_y^{n+1} = \inf\{k > \tau_y^n,\ X_k = y\} = \tau_y^n + \tau_y^+ \circ \theta^{\tau_y^n} \end{cases}$$

avec la convention que $\inf \emptyset = \infty$. Donc si $\tau_y^n = \infty$, les suivants aussi. Avec cette notation, on a pour $n \geqslant 1$,

$$\begin{aligned} \mathbb{P}_x(N(y) \geqslant n) = \mathbb{P}_x(\tau_y^n < \infty) &= \mathbb{P}_x\left(\tau_y^{n-1} < \infty \text{ et } \tau_y^+ \circ \theta^{\tau_y^{n-1}} < \infty\right) \\ &= \mathbb{E}_x\left(\mathbb{1}_{\tau_y^{n-1}<\infty} \mathbb{1}_{\tau_y^+ \circ \theta^{\tau_y^{n-1}} < \infty}\right) \\ &= \mathbb{E}_x\left(\mathbb{1}_{\tau_y^{n-1}<\infty} \mathbb{E}_{X_{\tau_y^{n-1}}}\left(\mathbb{1}_{\tau_y^+<\infty}\right)\right) \\ &= \mathbb{P}_y\left(\tau_y^+ < \infty\right) \mathbb{P}_x\left(\tau_y^{n-1} < \infty\right) \\ &= \mathbb{P}_y\left(\tau_y^+ < \infty\right) \mathbb{P}_x\left(N(y) \geqslant n-1\right) \end{aligned}$$

par Markov fort et car $X_{\tau_y^{n-1}} = y$ (car temps de visite de y).

On conclue la preuve par récurrence car pour $n = 1$,

$$\mathbb{P}_x(N(y) \geqslant 1) = \mathbb{P}_x(\tau_y < \infty).$$

On peut déduire les corollaires suivants du lemme.

Corollaire

(i) Si $\mathbb{P}_x(\tau_x^+ < \infty) < 1$ alors, sous $\mathbb{P}_x$, $N(x)$ suit une loi géométrique de paramètre $p = 1 - \mathbb{P}_x(\tau_x^+ < \infty) = \mathbb{P}_x(\tau_x^+ = \infty)$.
(ii) Si, x, y sont dans E

$$G(x, y) = \mathbb{P}_x(\tau_y < \infty)\frac{1}{1 - \mathbb{P}_y(\tau_y^+ < \infty)}$$

$$G(x, x) = \frac{1}{1 - \mathbb{P}_x(\tau_x^+ < \infty)}$$

(avec la convention $\frac{1}{0} = +\infty$ et $0 \times \infty = 0$.)

Preuve. (i) Conséquence directe du lemme avec $x = y$.
(ii) $\mathbb{E}_x(N(y)) = \sum_{n \geqslant 1} \mathbb{P}_x(N(y) \geqslant n)$, puis c'est une somme géométrique.

Définition

(i) On dit que x conduit à y, on note $x \rightsquigarrow y$, si $\mathbb{P}_x(\tau_y < \infty) > 0$.
(ii) On dit que x communique à y, on note $x \sim y$ si $x \rightsquigarrow y$ et $y \rightsquigarrow x$.

On a facilement que

$$(x \rightsquigarrow y) \Leftrightarrow (\exists n \geqslant 0 \text{ t.q. } Q^n(x,y) > 0) \Leftrightarrow (G(x,y) > 0),$$

car $G(x,y) = \sum_{n\geqslant 0} Q^n(x,y)$ et $G(x,y) = \mathbb{P}_x(\tau_y < \infty)G(x,x)$.

On en déduit facilement que $\sim$ est une relation d'équivalence. Ses classes d'équivalence sont appelées **classes de communication**.

Lemme

(i) Si $x \rightsquigarrow y$ alors (x rec.)$\Rightarrow$(y rec. et $\mathbb{P}_y(\tau_x < \infty) = 1$).
(ii) Si $x \sim y$ alors (x rec.)$\Leftrightarrow$(y rec.) et dans ce cas $\mathbb{P}_y(\tau_x < \infty) = \mathbb{P}_x(\tau_y < \infty) = 1$.

Remarque. En particulier, on déduit de (i) que si x est rec. alors $(x \rightsquigarrow y) \Leftrightarrow (x \sim y)$.

Preuve. (i) On a par la propriété de Markov forte:

$$\begin{aligned}\mathbb{P}_x(N(x) < \infty) &\geqslant \mathbb{P}_x(\tau_y < \infty\ ,\ \tau_x \circ \theta^{\tau_y} = \infty)\\ &= \mathbb{E}_x(\mathbb{1}_{\tau_y < \infty} \mathbb{1}_{\tau_x < \infty} \circ \theta^{\tau_y})\\ &= \mathbb{P}_x(\tau_y < \infty)\mathbb{P}_y(\tau_x = \infty).\end{aligned}$$

Comme $\mathbb{P}_x(\tau_y < \infty) > 0$ (car $x \rightsquigarrow y$), si x est rec. alors $\mathbb{P}_x(N(x) < \infty) = 0$ implique que $\mathbb{P}_y(\tau_x = \infty) = 0$, donc

$$\mathbb{P}_y(\tau_x < \infty) = 1.$$

En particulier $y \rightsquigarrow x$. Donc $\exists$ des entiers n_1 et n_2 tels que $Q^{n_1}(x, y) > 0$ et $Q^{n_2}(y, x) > 0$. Donc

$$\begin{aligned}G(y, y) = \sum_{n=0}^{\infty} Q^n(y, y) &\geqslant Q^{n_2}(y, x) \left(\sum_{n=0}^{\infty} Q^n(x, x)\right) Q^{n_1}(x, y)\\ &= Q^{n_2}(y, x) G(x, x) Q^{n_1}(x, y)\end{aligned}$$

Donc $G(x, x) = \infty$ implique $G(y, y) = \infty$ donc y est rec.

(ii) découle de (i)

Par le lemme précédent, dans une classe de communication, soit tous les points sont récurrent, soit tous transitoires. On note $\mathcal{R}$ l'ensemble des points récurrents, alors $\mathcal{R}$ se partitionne en ses classes de communications

$$\mathcal{R} = \sqcup_{i \in I} \mathcal{R}_i,$$

les $\mathcal{R}_i$ sont appelées classes de récurrence.

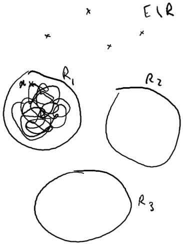
Théorème (Théorème de classification des états)

(i) Soit x rec. et i t.q. $x \in \mathcal{R}_i$, alors $\mathbb{P}_x$ p.s.
- $N(y) = \infty$, si $y \in \mathcal{R}_i$,
- $N(y) = 0$, si $y \in \mathcal{R}_i$.

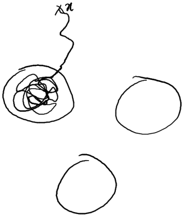
(ii) Soit $x \in E \setminus \mathcal{R}$, et $T = \inf\{n \geqslant 0,\ X_n \in \mathcal{R}\}$, alors $\mathbb{P}_x$ p.s.
- soit $T = \infty$ et $N(y) < \infty$ pour tout $y \in E$.
- soit $T < \infty$, alors $\exists\ j \in I$ tel que $X_T \in \mathcal{R}_j$ et, $\forall n \geqslant T$ $X_n \in \mathcal{R}_j$, $N(y) < \infty$ pour $y \in E \setminus \mathcal{R}_j$, $N(y) = \infty$ pour $y \in \mathcal{R}_j$.

Preuve. Admis, cf Le Gall p. 202.

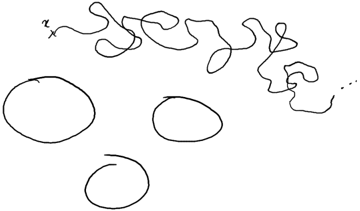
Cas a :



Cas b :



Cas C :



Définition

On dit que la chaîne est irréductible si $x \sim y$ *pour tous* $x, y \in E$ *(càd que* $\mathbb{P}_x(\tau_y < \infty) > 0$ *pour tous* $x, y \in E$*).*

Proposition

Si la CM est irréductible alors

- Soit tous les points sont récurrents, E *est une classe de récurrence et on a* $\forall x, y \in E$, $\mathbb{P}_x(N(y) = \infty) = 1$.

- Soit tous les points sont transitoires, et dans ce cas $\forall x, y \in E$ $\mathbb{P}_x(N(y) < \infty) = 1$.

Dans le premier cas on dit que la CM est récurrente irréductible, dans le deuxième transitoire (ou transiente) irréductible.

Preuve. Application directe du théorème de classification des états.

Cas des CM à espace d'états fini.

On note $\{\mathcal{C}_j\}_{j\in J}$ les classes de communications de E, donc

$$E = \sqcup_{j\in J}\mathcal{C}_j.$$

Proposition

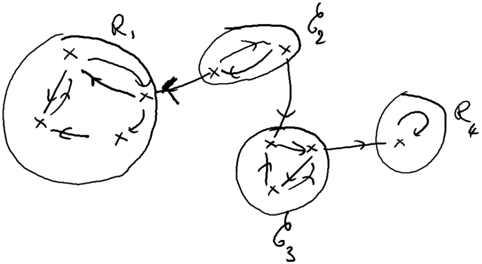
On suppose E **fini**.
(i) les classes de récurrences sont exactement les classes de communication $\mathcal{C}_j$ telles qu'il n'existe pas $x \in \mathcal{C}_j$ et $y \in E \setminus \mathcal{C}_j$ tels que $x \rightsquigarrow y$.
(ii) si la CM est irréductible alors elle est récurrente.

Preuve. (i) Si $\mathcal{C}_j$ est tel que ($x \in \mathcal{C}_j$ et $x \rightsquigarrow y$)$\Rightarrow$ ($y \in \mathcal{C}_j$), alors sous $\mathbb{P}_x$, $X_n \in \mathcal{C}_j$ $\forall n \in \mathbb{N}$. Comme $\mathcal{C}_j$ est fini, il existe au moins un point y dans $\mathcal{C}_j$ tel que $N(y) = \infty$. Donc y est récurrent, donc $\mathcal{C}_j$ est une classe de récurrence.
Réciproquement, si il existe $x \in \mathcal{C}_j$ et $y \in E \setminus \mathcal{C}_j$ tels que $x \rightsquigarrow y$, alors $y \not\rightsquigarrow x$ sinon y serait dans $\mathcal{C}_j$. Donc

$$\mathbb{P}_x(N(x) < \infty) \geqslant \mathbb{P}_x(\tau_y < \infty)\mathbb{P}_y(\tau_x = \infty) = \mathbb{P}_x(\tau_y < \infty) > 0.$$

Donc x est transitoire et la classe $\mathcal{C}_j$ est transitoire.

Exemple :



La marche aléatoire simple sur $\mathbb{Z}^d$: théorème de Polya

Théorème (Théorème de Polya)

La marche aléatoire simple (MAS) sur $\mathbb{Z}^d$ est irréductible et

$$\begin{cases} \text{récurrente, si } d = 1 \text{ ou } d = 2 \\ \text{transitoire, si } d \geqslant 3. \end{cases}$$

Preuve. Nous présentons une preuve qui permet de calculer la transformée de Fourier de la fonction de Green. Il y a de nombreuses autres preuves dont une par calcul direct (ou estimation) combinatoire de la probabilité que la MAS revienne à l'origine au temps n.

On note $(X_n)_{n\geqslant 0}$ la MAS partant de 0. Pour $\theta \in [-\pi, \pi[^d$ on note

$$\varphi_n(\theta) = \mathbb{E}\left(\exp(i\,\langle \theta, X_n \rangle\right),$$

la fonction caractéristique de X_n ($\langle \cdot, \cdot \rangle$ denote le produit scalaire dans $\mathbb{R}^d$.) Comme $X_n \in \mathbb{Z}^d$, φ_n serait 2π-périodique en chaque coordonnée si définie sur $\mathbb{R}^d$, on la restreint donc à $[-\pi, \pi[^d$.

On a $X_n = Y_1 + \cdots + Y_n$ où les Y_i sont les pas i.i.d. de la MAS (cf definition), on a donc

$$\varphi_n(\theta) = (\mathbb{E}\left(\exp(i\left\langle \theta, Y_1 \right\rangle)\right))^n = \left(\frac{1}{2d} \sum_{i=1}^{d} (e^{i\theta_i} + e^{-i\theta_i}) \right)^n$$
$$= \left(\frac{1}{d} \sum_{i=1}^{d} \cos(\theta_i) \right)^n$$

D'autre part, on a

$$\varphi_n(\theta) = \sum_{x \in \mathbb{Z}^d} e^{i\langle \theta, x \rangle} \mathbb{P}_0(X_n = x) = \sum_{x \in \mathbb{Z}^d} e^{i\langle \theta, x \rangle} Q^n(0, x).$$

Donc

$$\int_{[-\pi,\pi[^d} \varphi_n(\theta) d\theta_1 \cdots d\theta_d = \sum_{x \in \mathbb{Z}^d} Q^n(0, x) \left(\prod_{j=0}^{d} \int_{[-\pi,\pi[} e^{ix_j\theta} d\theta \right)$$
$$= (2\pi)^d Q^n(0, 0).$$

(En effet comme $\sum_{x \in \mathbb{Z}^d} Q^n(0, x) = 1$ on peut utiliser Fubini car la somme est dominée.)

Pour $0 < \rho < 1$, on note

$$G_\rho(0,0) = \sum_{n=0}^{\infty} \rho^n Q^n(0,0),$$

de sorte que $G(0,0) = \lim_{\rho\to 1, \rho<1} G_\rho(0,0)$. On a par le calcul précédent et par Fubini pour $\rho < 1$, en notant $d\theta = d\theta_1 \cdots d\theta_d$,

$$\begin{aligned} G_\rho(0,0) &= \sum_{n=0}^{\infty} \rho^n \frac{1}{(2\pi)^d} \int_{[-\pi,\pi[^d} \varphi_n(\theta) d\theta \\ &= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi[^d} \sum_{n=0}^{\infty} \rho^n \left(\frac{1}{d} \sum_{i=1}^{d} \cos(\theta_i) \right)^n d\theta \\ &= \frac{1}{(2\pi)^d} \int_{[-\pi,\pi[^d} \frac{1}{1 - \frac{\rho}{d} \sum_{i=1}^{d} \cos(\theta_i)} d\theta \end{aligned}$$

En effet, on a $\left| \rho^n \left(\frac{1}{d} \sum_{i=1}^{d} \cos(\theta_i) \right)^n \right| \leqslant \rho^n$, on a donc des séries absolument convergentes.

Ensuite on montre que

$$\lim_{\rho\to 1}\int_{[-\pi,\pi[^d}\frac{1}{1-\frac{\rho}{d}\sum_{i=1}^{d}\cos(\theta_i)}d\theta=\int_{[-\pi,\pi[^d}\frac{1}{1-\frac{1}{d}\sum_{i=1}^{d}\cos(\theta_i)}d\theta$$

(Notez que les intégrants sont toujours positifs car $|\frac{1}{d}\sum_{i=1}^{d}\cos(\theta_i)|\leqslant 1$, l'intégrale de droite a donc un sens et est à valeurs dans $\mathbb{R}_+\cup\{+\infty\}$.)

En effet, sur le domaine $\{\theta,\ \frac{1}{d}\sum_{i=1}^{d}\cos(\theta_i)>0\}$ on peut utiliser la convergence monotone et sur le domaine $\{\theta,\ \frac{1}{d}\sum_{i=1}^{d}\cos(\theta_i)\leqslant 0\}$, on peut utiliser la convergence dominée car sur ce domaine $\frac{1}{1-\frac{\rho}{d}\sum_{i=1}^{d}\cos(\theta_i)}\leqslant 1$.

On a donc démontré

$$G(0,0)=\frac{1}{(2\pi)^d}\int_{[-\pi,\pi[^d}\frac{1}{1-\frac{1}{d}\sum_{i=1}^{d}\cos(\theta_i)}d\theta.$$

A partir de cette expression, on peut voir si $G(0,0) < \infty$ ou $G(0,0) = \infty$. En effet, quand $\theta \to 0$

$$1 - \frac{1}{d}\sum_{i=1}^{d} \cos(\theta_i) \sim \frac{1}{2}\|\theta\|^2.$$

Comme la divergence de l'intégrale se fait au voisinage de 0, $G(0,0) < \infty$ ssi

$$\int_{B_d(0,1)} \frac{1}{\|\theta\|^2} d\theta < \infty.$$

En intégrant sur le rayon, càd sur $r := \|\theta\|$ on a

$$\begin{aligned}\int_{B_d(0,1)} \frac{1}{\|\theta\|^2} d\theta &= d|B_d(0,1)| \int_0^1 \frac{1}{r^2} r^{d-1} dr \\ &= d|B_d(0,1)| \int_0^1 r^{d-3} dr\end{aligned}$$

où $|B_d(0,1)|$ est le volume de la boule unité de centre 0 et de rayon 1. (cf par ex. "Garet, Kurtzman" propriété 4.48).

On voit donc que $G(0,0) < \infty$ ssi $d > 2$. CQFD.

Exemple du processus de Galton Watson

On suppose $p_0 > 0$. Dans ce cas, 0 est un point absorbant (c'est à dire que si on est en 0, on y reste avec proba 1) et $k \rightsquigarrow 0$ pour tout $k \in \mathbb{N}$. En effet,

$$\mathbb{P}(Z_{n+1} = 0 \mid Z_n = k) = p_0^k.$$

car il faut que chaque individu ait 0 descendant.

Si on applique le Théorème de classification des états, on voit que $\{0\}$ est la seule classe de récurrence, donc soit la population s'éteint (càd Z_n atteint 0), soit chaque $k \in \mathbb{N}$ est visité un nombre fini de fois, donc $Z_n \to \infty$. Donc,

$$\mathbb{P}(\{Z_n \text{ s'éteint}\} \cup \{Z_n \to \infty\}) = 1.$$

Remarque: En L3, vous avez vu que $\mathbb{P}(Z_n \to \infty) > 0$ ssi $m > 1$.

Marche aléatoire générale sur $\mathbb{Z}$

Soit $X_n = x + \sum_{k=1}^{n} \xi_k$, où $(\xi_k)_{k\geqslant 1}$ est une suite de v.a. i.i.d. à valeurs dans $\mathbb{Z}$, $x \in \mathbb{Z}$ la position initiale. On suppose les (ξ_k) intégrable et on note $m = \mathbb{E}(\xi_k)$.

Proposition

(i) Si $m \neq 0$ alors $X_n \cdot m \to +\infty$ et tous les points sont transitoires.
(ii) Si $m = 0$ alors tous les points sont récurrents. La marche est irréductible ssi tout élément de $\mathbb{Z}$ s'écrit comme somme d'éléments de $\mathrm{Supp}(loi(\xi))$.

Preuve Par la LGN, on a $\frac{X_n}{n} \to m$, p.s..
(i) Si $m \neq 0$, ça implique $X_n \cdot m \to +\infty$ p.s.

(ii) Supposons $m = 0$ et supposons 0 transitoire. On note $G(x,y)$ la fonction de Green. On a donc $G(0,0) < \infty$ et

$$G(0,x) \leqslant G(x,x) = G(0,0),$$

par invariance par translation. Donc pour $\epsilon > 0$, pour $N \in \mathbb{N}$

$$\sum G(0,x) \leqslant G(0,0)(2\epsilon N + 1) \tag{1}$$

Par la LGN on a $\frac{X_n}{n} \xrightarrow{\mathcal{P}} 0$ et pour tout $\epsilon > 0$, $\exists\ N_0 \in \mathbb{N}$ tel que

$$\mathbb{P}_0(\frac{X_n}{n} \leqslant \epsilon) \geqslant \frac{1}{2}, \text{ pour } n \geqslant N_0.$$

Donc, pour $n \geqslant N_0$,

$$\sum_{|x| \leqslant \epsilon n} Q^n(0,x) = \mathbb{P}_0(\frac{X_n}{n} \leqslant \epsilon) \geqslant \frac{1}{2}.$$

Finalement, pour $N \geqslant N_0$,

$$\sum_{|x| \leqslant \epsilon N} G(0,x) \geqslant \sum_{n=N_0}^{N} \sum_{|x| \leqslant \epsilon N} Q^n(0,x) \geqslant (N - N_0)\frac{1}{2}. \qquad (2)$$

On choisit maintenant ϵ tel que $2\epsilon G(0,0) < 1/4$ puis $N > 4G(0,0)$ tel que $(N - N_0) > \frac{1}{2}N$. On a donc $G(0,0)(2\epsilon N + 1) < N/4$, on arrive avec les inégalités (1) et (2) à

$$\frac{1}{4}N < \sum_{|x| \leqslant \epsilon N} G(0,x) < \frac{1}{4}N.$$

Donc à une absurdité. Donc 0 est récurrent, donc tous les x sont récurrents. Le reste est simple.

4-Mesures invariantes

Définition

Soit μ une mesure positive sur E telle que $\mu(\{x\}) < \infty$ pour tout $x \in E$ et $\mu \neq 0$. On dit que μ est une mesure invariante si

$$\mu(x) = \sum_{y \in E} \mu(y) Q(y, x), \quad \forall x \in E. \tag{3}$$

Si $\mu(E) = 1$ on dit que μ est une probabilité invariante. On notera souvent $\mu Q = \mu$ pour l'équation (3).

Remarque: Si μ est une probabilité invariante, alors si la loi initiale est μ (càd sous $\mathbb{P}_\mu$) alors X_n a pour loi μ. En effet, par récurrence, si on suppose que, sous $\mathbb{P}_\mu$, $X_n \sim \mu$ alors

$$\mathbb{P}_\mu(X_{n+1} = x) = \sum_{y \in E} \mathbb{P}_\mu(X_n = y) Q(y, x) = \sum_{y \in E} \mu(y) Q(y, x) = \mu(x).$$

Exemple 1. Sur $\mathbb{Z}^d$, la mesure uniforme $\mu(\{x\}) = 1$, $\forall x \in \mathbb{Z}^d$, est une mesure invariante pour toute marche aléatoire.

Remarque: Quand E est fini, l'équation $\mu Q = \mu$ correspond à μ est vecteur propre à coeff positifs de ${}^tQ$ avec vap 1. La matrice Q étant stochastique, on sait que $\mathbb{1}$ est vecteur propre de Q avec vap 1, donc 1 est aussi vap de ${}^tQ$. Le fait qu'on puisse trouver un vecteur propre à coefficients positifs est moins simple (voir plus loin) et aussi lié au théorème de Perron-Frobenius dans le cas E fini.

Définition-Proposition

Soit μ une mesure positive sur E telle que $\mu(\{x\}) < \infty$ pour tout $x \in E$ et $\mu \neq 0$. On dit que μ est une mesure réversible si

$$\mu(x)Q(x,y) = \mu(y)Q(y,x), \quad \forall x, y \in E.$$

Une mesure réversible est invariante.

Preuve: $\sum_y \mu(y)Q(y,x) = \mu(x)\sum_y Q(x,y) = \mu(x)$.

Remarque: La réciproque est fausse (exo: trouvez un contre-exemple) ! Néanmoins beaucoup de CM qui proviennent de la modélisation ou de la physique sont réversibles. Les CM réversibles ont des propriétés remarquables et sont plus simples à étudier (explication peut-être en fin du cours).

Exemple 2. La MAS sur $\mathbb{Z}^d$ admet la mesure uniforme comme mesure réversible. Par contre, cette mesure est invariante pour toute marche aléatoire mais pas forcément réversible. (Exo: trouvez une condition sur les pas de la marche aléatoire pour que cette mesure soit réversible.)

Exemple 3. (Urne d'Ehrenfest). Soit μ mesure sur $\{0, \ldots, N\}$ donnée par

$$\mu(j) = \binom{N}{j}, \quad j \in \{0, \ldots, N\}.$$

Alors μ est mesure réversible, car

$$\mu(j)Q(j, j+1) = \binom{N}{j} \frac{N-j}{N} = \binom{N}{j+1} \frac{j+1}{N} = \mu(j+1)Q(j+1, j).$$

Donc $\nu(j) = \frac{1}{2^n}\binom{N}{j}$ est une probabilité réversible (donc invariante).

Interprétation. Si on place chaque boule au hasard dans l'urne A ou B, le nombre de boules dans A a pour loi ν. Donc la probabilité invariante correspond à placer chaque boule au hasard.

Proposition (Existence)

Si x est un point récurrent de la CM, alors la mesure ν_x donnée par

$$\nu_x(y) = \mathbb{E}_x\left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=y}\right), \quad \forall y \in E,$$

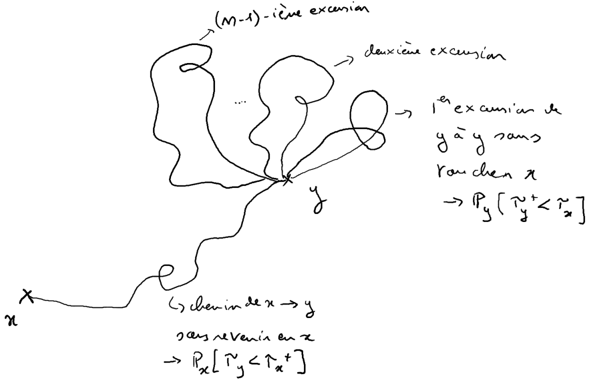
est une mesure invariante. Son support est la classe de récurrence de x.

Preuve. Si on note $N_x(y) = \sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=y}$ le nombre de visites en y avant le premier retour en x, par la propriété de Markov fort on a

$$\mathbb{P}_x(N_x(y) \geqslant n) = \mathbb{P}_x(\tau_y < \tau_x^+)\mathbb{P}_y(\tau_y^+ < \tau_x)^{n-1},$$

pour $y \neq x$ et $N_x(x) = 1$. (La preuve se fait de la même façon que la proposition sur $N(x)$: **exo**). Comme x est rec., $x \rightsquigarrow y$ ssi $x \sim y$. Donc $\mathbb{P}_x(\tau_y < \tau_x^+) > 0$ ssi $x \sim y$. De plus si $x \sim y$, alors $\mathbb{P}_y(\tau_y^+ < \tau_x) < 1$ donc $\mathbb{E}_x(N_x(y)) < \infty$. La mesure ν_x est donc finie et strict positive sur la classe de récurrence de x, nulle ailleurs.

Heuristique formule : $\mathbb{P}_x\left[N_x(y) \geq n\right] = \mathbb{P}_x\left[\tau_y < \tau_x^+\right] \mathbb{P}_y\left[\tau_y^+ < \tau_x\right]^{n-1}$



On prouve maintenant que ν_x est invariante. Pour $y \in E$,

$$\begin{aligned}
\sum_{z\in E} \nu_x(z)Q(z,y) &= \sum_{z\in E} \mathbb{E}_x \left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=z} \right) Q(z,y) \\
&= \sum_{z\in E} \mathbb{E}_x \left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=z} Q(z,y) \right) \\
&= \sum_{z\in E} \mathbb{E}_x \left(\sum_{k=0}^{+\infty} \mathbb{1}_{k<\tau_x^+} \mathbb{1}_{X_k=z} \mathbb{P}_z(X_1 = y) \right) \\
&= \sum_{z\in E} \sum_{k=0}^{+\infty} \mathbb{E}_x \left(\mathbb{1}_{k<\tau_x^+} \mathbb{1}_{X_k=z} \mathbb{P}_z(X_1 = y) \right)
\end{aligned}$$

Comme τ_x^+ est un temps d'arrêt on utilise que

$$\{k < \tau_x^+\} \cap \{X_k = z\} = \{\tau_x^+ \leqslant k\}^c \cap \{X_k = z\} \in \mathcal{F}_k.$$

et on applique alors Markov faible pour déduire

$$\mathbb{E}_x \left(\mathbb{1}_{k<\tau_x^+} \mathbb{1}_{X_k=z} \mathbb{P}_z(X_1 = y) \right) = \mathbb{E}_x \left(\mathbb{1}_{k<\tau_x^+} \mathbb{1}_{X_k=z} \mathbb{1}_{X_{k+1}=y} \right).$$

Donc,

$$\begin{aligned}
\sum_{z\in E}\nu_x(z)Q(z,y) &= \sum_{z\in E}\sum_{k=0}^{+\infty}\mathbb{E}_x\left(\mathbb{1}_{k<\tau_x^+}\mathbb{1}_{X_k=z}\mathbb{1}_{X_{k+1}=y}\right)\\
&= \sum_{k=0}^{+\infty}\mathbb{E}_x\left(\mathbb{1}_{k<\tau_x^+}\mathbb{1}_{X_{k+1}=y}\sum_{z\in E}\mathbb{1}_{X_k=z}\right)\\
&= \sum_{k=0}^{+\infty}\mathbb{E}_x\left(\mathbb{1}_{k<\tau_x^+}\mathbb{1}_{X_{k+1}=y}\right)\\
&= \mathbb{E}_x\left(\sum_{k=0}^{+\infty}\mathbb{1}_{k<\tau_x^+}\mathbb{1}_{X_{k+1}=y}\right)\\
&= \mathbb{E}_x\left(\sum_{k=1}^{+\infty}\mathbb{1}_{k<\tau_x^++1}\mathbb{1}_{X_k=y}\right)\\
&= \mathbb{E}_x\left(\sum_{k=1}^{\tau_x^+}\mathbb{1}_{X_k=y}\right)
\end{aligned}$$

On conclue de la façon suivante: si $y \neq x$,

$$\mathbb{E}_x\left(\sum_{k=1}^{\tau_x^+} \mathbb{1}_{X_k=y}\right) = \mathbb{E}_x\left(\sum_{k=1}^{\tau_x^+-1} \mathbb{1}_{X_k=y}\right) = \mathbb{E}_x\left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=y}\right) = \nu_x(y)$$

car sous $\mathbb{P}_x$, $X_0 = x$ et $X_{\tau_x^+} = x$. Si $y = x$ alors

$$\mathbb{E}_x\left(\sum_{k=1}^{\tau_x^+} \mathbb{1}_{X_k=x}\right) = 1 = \mathbb{E}_x\left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=x}\right) = \nu_x(x)$$

On a donc démontré que

$$\sum_{z\in E} \nu_x(z) Q(z,y) = \nu_x(y).$$

On a aussi remarqué la chose simple suivante.

Proposition

On a pour tout $x \in V$,

$$\nu_x(x) = 1.$$

Théorème

Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov irréductible et récurrente. Alors il existe une mesure invariante, elle est unique à constante multiplicative près (càd que deux mesures invariantes sont toujours proportionnelles). De plus elle est strictement positive sur tout E.

Preuve. *Existence.* C'est une conséquence du lemme précédent, en prenant ν_x pour un $x \in E$. La classe de récurrence de x est E car la chaîne est irréductible et récurrente, donc le support de ν_x est tout E.

Unicité. Soit μ une mesure invariante de la CM. On choisit $x \in E$, on veut montrer que μ est proportionnelle à ν_x donc que

$$\mu = \mu(x)\nu_x, \ \ (\text{càd } \mu(y) = \mu(x)\nu_x(y), \ \forall y \in E)$$

car $\nu_x(x) = 1$. Si on démontre cela, on saura que tout mesure invariante est proportionnelle à ν_x donc on aura unicité à constante multiplicative près.

On commence par montrer que $\mu \geqslant \mu(x)\nu_x$. Pour $p \in \mathbb{N}$, on note

$$\nu_x^p(y) = \mathbb{E}_x\left(\sum_{k=0}^{(\tau_x^+-1)\wedge p} \mathbb{1}_{X_k=y}\right), \quad \forall y \in E,$$

Pour $p = 0$ on a $\nu_x^0 = \delta_x$, donc $\mu \geqslant \mu(x)\nu_x^0$. On va montrer la propriété par récurrence. Si on suppose la propriété vraie pour p, en procédant comme dans la preuve du lemme, pour $y \neq x$,

$$\begin{aligned}\mu(y) &= \sum_{z\in E}\mu(z)Q(z,y)\\ &\geqslant \mu(x)\sum_{z\in E}\nu_x^p(z)Q(z,y)\\ &= \mu(x)\sum_{z\in E}\mathbb{E}_x\left(\sum_{k=0}^{(\tau_x^+-1)\wedge p}\mathbb{1}_{X_k=z}Q(z,y)\right)\\ &= \mu(x)\mathbb{E}_x\left(\sum_{k=1}^{\tau_x^+\wedge(p+1)}\mathbb{1}_{X_k=y}\right) = \mu(x)\nu_x^{p+1}(y)\end{aligned}$$

Pour $y = x$, on a toujours $\nu^p(x) = 1$ pour tout p, donc $\mu(x) = \mu(x)\nu^p(x)$.
On a donc par récurrence que pour tout p, $\mu \geqslant \mu(x)\nu_x^p$. D'où

$$\mu(y) \geqslant \mu(x) \lim_{p\to\infty} \nu_x^p(y) = \mu(x)\nu_x(y), \quad \forall y \in E.$$

et on a égalité en $y = x$.
On note

$$E_0 = \{y \in E,\ \mu(y) = \mu(x)\nu_x(y)\}, \quad E_1 = \{y \in E,\ \mu(y) > \mu(x)\nu_x(y)\}.$$

Donc $E_0 \neq \emptyset$ car $x \in E_0$. Supposons que $E_1 \neq \emptyset$. Comme la CM est irréductible on peut trouver $y' \in E_1$, $y'' \in E_0$ tels que $Q(y', y'') > 0$. On a

$$\begin{aligned}\mu(y'') &= \sum_{z\in E} \mu(z)Q(z, y'') \\ &> \sum_{z\in E} \mu(x)\nu_x(z)Q(z, y'') = \mu(x)\nu_x(y'') = \mu(y'').\end{aligned}$$

d'où contradiction, donc $E_1 = \emptyset$. (L'inégalité stricte précédente vient de ce que dans la somme l'inégalité est stricte pour $z = y'$ et large pour les autres termes.)

Proposition

Soit $(X_n)_{n\geqslant 0}$ une CM irréductible récurrente.

(i) ($\exists$ une proba invariante ν) $\Leftrightarrow$ ($\exists x \in E$ tel que $\mathbb{E}_x(\tau_x^+) < \infty$)
Dans ce cas toutes les mesures invariante sont de masse finie, $\mathbb{E}_y(\tau_y^+) < \infty$ pour tout $y \in E$. De plus, la probabilité invariante ν est donnée par

$$\nu(y) = \frac{1}{\mathbb{E}_y(\tau_y^+)}, \quad \forall y \in E.$$

(ii) ($\nexists$ de proba invariante) $\Leftrightarrow$ ($\exists x \in E$ tel que $\mathbb{E}_x(\tau_x^+) = \infty$)
Dans ce cas toutes les mesures invariante sont de masse infinie, et $\mathbb{E}_y(\tau_y^+) = \infty$ pour tout $y \in E$.

Dans le cas (i) on dit que la CM est irréductible **récurrente positive**, *dans le cas (ii) irréductible* **récurrente nulle**.

Preuve. (i) Pour $x \in E$ on a que

$$\nu_x(E) = \sum_{y \in E} \nu_x(y) = \mathbb{E}_x \left(\sum_{k=0}^{\tau_x^+ - 1} \sum_{y \in E} \mathbb{1}_{X_k = y} \right) = \mathbb{E}_x \left(\sum_{k=0}^{\tau_x^+ - 1} 1 \right) = \mathbb{E}_x(\tau_x^+)$$

Donc $\nu_x(E) < \infty$ ssi $\mathbb{E}_x(\tau_x^+) < \infty$. Comme toutes les mesures invariantes sont égales à constantes près, on a $\nu_x(E) < \infty$ ssi $\nu_y(E) < \infty$ pour tout $y \in E$, donc

$$\begin{aligned} (\exists x \in E, \, \mathbb{E}_x(\tau_x^+) < \infty) &\Leftrightarrow (\forall y \in E, \, \mathbb{E}_y(\tau_y^+) < \infty) \\ &\Leftrightarrow (\text{les mes. inv. ont une masse finie}) \end{aligned}$$

Dans ce cas la mesure invariante est donnée par

$$\nu = \frac{\nu_x}{\nu_x(E)}, \quad \forall x \in E.$$

Comme $\nu_x(x) = 1$, et $\nu_x(E) = \mathbb{E}_x(\tau_x^+)$, ça donne le résultat.
(ii) Similaire.

Remarque: Si E est fini et la CM est irréductible, elle est récurrente positive.

Proposition

Si la CM est irréductible et si il existe une **probabilité** *invariante, alors elle est récurrente (positive).*

Remarque: Faux en général si il existe seulement une mesure invariante: ex. la MAS sur $\mathbb{Z}^d$, $d \geqslant 3$.

Preuve. Soit ν probabilité invariante. Soit $x \in E$.

$$\sum_{y\in E} \nu(y) G(y,x) = \sum_{y\in E} \nu(y) \sum_{n=0}^{\infty} Q^n(y,x) = \sum_{n=0}^{\infty} \nu(x) = +\infty.$$

D'autre part,

$$\sum_{y\in E} \nu(y) G(y,x) \leqslant \sum_{y\in E} \nu(y) G(x,x) = \nu(E) G(x,x) = G(x,x).$$

Donc $G(x,x) = \infty$. La CM est récurrente et récurrente positive par définition.

Exemple. *Chaîne de naissance et de mort.* On se donne $(\omega_i)_{i\geqslant 1} \in]0,1[^{\mathbb{N}^*}$. On considère la CM sur $E = \mathbb{N}$ qui saute de ± 1 avec probabilité de transition

$$\begin{cases} Q(0,1) = 1, & \\ Q(i,i+1) = \omega_i, & \text{si } i \geqslant 1 \\ Q(i,i-1) = 1-\omega_i, & \text{si } i \geqslant 1 \end{cases}$$

On cherche $(\nu(i))_{i\in\mathbb{N}}$, mesure reversible, telle que $\nu(0) = 1$: on doit avoir

$$\nu(0) = \nu(1)(1-\omega_1), \;\; \nu(i)\omega_i = \nu(i+1)(1-\omega_{i+1}), \; i \geqslant 1.$$

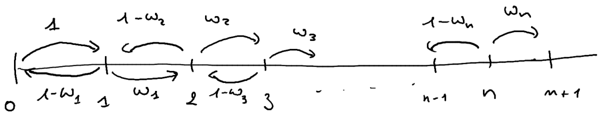
On pose $\rho_i = \frac{\omega_i}{1-\omega_i}$, pour $i \geqslant 1$. Par récurrence, ça donne pour $n \geqslant 1$,

$$\nu(n) = \rho_1 \cdots \rho_{n-1} \frac{1}{1-\omega_n} = \rho_1 \cdots \rho_{n-1}(1+\rho_n).$$

La CM est donc récurrente positive ssi $\sum_{n=1}^{\infty} \rho_1 \cdots \rho_n < \infty$.

Chaîne de naissance et mort.

$$(\omega_i)_{i\geq 1} \in \,]0,1[^{\mathbb{N}^*}$$



5- Le théorème ergodique

Théorème (Version générale)

Soit (X_n) *une CM* **irréductible et récurrente**. *Soit* μ *une mesure invariante de la CM. Soient* f, g, *deux fonctions définies sur* E. *On suppose de plus que* f *est* μ*-intégrable et* $g \geqslant 0$ *et* $\sum_E g(y)\mu(y) > 0$. *Alors, pour tout point initial* $x \in E$,

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} f(X_k)}{\sum_{k=0}^{n-1} g(X_k)} = \frac{\sum_{y\in E} f(y)\mu(y)}{\sum_{y\in E} g(y)\mu(y)}, \quad \mathbb{P}_x \text{ p.s.}$$

Théorème (version récurrente positive)

Soit (X_n) *une CM* **irréductible et récurrente positive**, *et* ν *la* **probabilité invariante**. *Soient* f *une fonction définie sur* E *et* ν*-intégrable. Alors, pour tout point initial* $x \in E$.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{y\in E} f(y)\nu(y), \quad \mathbb{P}_x \text{ p.s.}$$

Remarque: Le théorème ergodique s'occupe donc de la convergence p.s. pour des moyennes de Cezaro de fonctions prises le long de la trajectoire $f(X_n)$. Le premier théorème peut aussi s'écrire comme le rapport des moyennes de Cezaro pour f et g, en divisant par n chaque terme.

Preuve du 2ème théorème. Conséquence directe du premier en prenant $g = 1$.

Preuve du 1er théorème. On suppose d'abord $f \geqslant 0$. On se place toujours sous $\mathbb{P}_x$. On considère les temps de visites successifs de x:

$$\begin{cases} \tau_x^0 = \tau_x = 0 \\ \tau_x^{i+1} = \inf\{k > \tau_x^i,\ X_k = x\} = \tau_x^i + \tau_x^+ \circ \theta^{\tau_x^i}, \quad i \geqslant 1 \end{cases}$$

Pour $n \in \mathbb{N}$, on note $N_n(x)$ le nombre de retours en x avant le temps $n - 1$:

$$N_n(x) = |\{0 < k \leqslant n-1,\ X_k = x\}| = \sum_{k=1}^{n-1} \mathbb{1}_{X_k = x}.$$

On peut décomposer la somme suivant les excursions de la CM:

$$\sum_{k=0}^{n-1} f(X_k) = \left(\sum_{i=1}^{N_n(x)} Z_i(f) \right) + Z(f)$$

où

$$Z_i(f) = \sum_{k=\tau_x^{i-1}}^{\tau_x^i - 1} f(X_k), \quad Z(f) = \sum_{k=\tau_x^{N_n(x)}}^{n-1} f(X_k)$$

Lemme

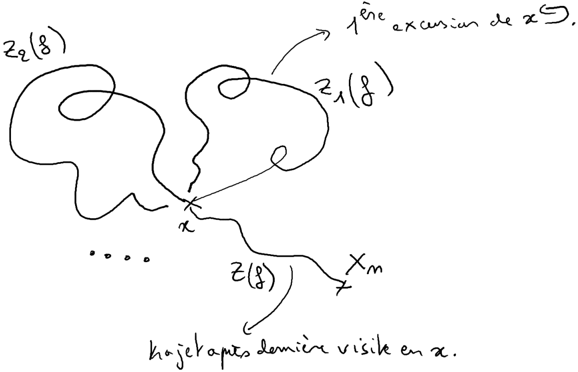
Sous $\mathbb{P}_x$, les $(Z_i(f))_{i \geqslant 1}$ sont i.i.d. et $\mathbb{E}_x(Z_i(f)) = \int_E f \, d\nu_x$.

Preuve Soient $\varphi_1, \ldots, \varphi_l$ des fonctions test, par Markov fort,

$$\begin{aligned}
\mathbb{E}_x\left(\varphi_1(Z_1) \cdots \varphi_l(Z_l)\right) &= \mathbb{E}_x\left(\varphi_1(Z_1)\varphi_2(Z_1 \circ \theta_{\tau_x^+}) \cdots \varphi_l(Z_{l-1} \circ \theta_{\tau_x^+}) \right) \\
&= \mathbb{E}_x\left(\varphi_1(Z_1)\mathbb{E}_x\left(\varphi_2(Z_1) \cdots \varphi_l(Z_{l-1})\right)\right) \\
&= \mathbb{E}_x\left(\varphi_1(Z_1)\right)\mathbb{E}_x\left(\varphi_2(Z_1) \cdots \varphi_l(Z_{l-1})\right)
\end{aligned}$$

On conclue par récurrence que

$$\mathbb{E}_x\left(\varphi_1(Z_1) \cdots \varphi_l(Z_l)\right) = \mathbb{E}_x\left(\varphi_1(Z_1)\right) \cdots \mathbb{E}_x\left(\varphi_l(Z_1)\right)$$


$Z_2(f)$
1ère excursion de x.
$Z_1(f)$
x
....
X_n
$Z(f)$
trajet après dernière visite en x.

Les $Z_i(f)$ sont donc i.i.d., et on a

$$\mathbb{E}_x(Z_1(f)) = \mathbb{E}_x\left(\sum_{k=0}^{\tau_x^+-1} f(X_k)\right) = \sum_{y\in E} f(y)\mathbb{E}_x\left(\sum_{k=0}^{\tau_x^+-1} \mathbb{1}_{X_k=y}\right)$$
$$= \sum_{y\in E} f(y)\nu_x(y).$$

Ceci conclue la preuve du lemme.

On remarque maintenant, comme $f \geqslant 0$, que

$$Z(f) \leqslant Z_{N_n(x)+1}(f)$$

On a donc

$$\frac{\sum_{i=1}^{N_n(x)} Z_i(f)}{N_n(x)} \leqslant \frac{\sum_{k=0}^{n-1} f(X(k))}{N_n(x)} \leqslant \frac{\sum_{i=1}^{N_n(x)+1} Z_i(f)}{N_n(x)}$$

Comme $\lim_{n\to\infty} N_n(x) = \infty$ p.s. car x est récurrent, comme les $Z_i(f)$ sont i.i.d., on a, p.s.

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{N_n(x)} Z_i(f)}{N_n(x)} = \lim_{n\to\infty} \frac{\sum_{i=1}^{N_n(x)+1} Z_i(f)}{N_n(x)} = \mathbb{E}_x(Z_1(f)) = \int f \, d\nu_x.$$

On a donc

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} f(X(k))}{N_n(x)} = \int f \, d\nu_x, \quad \text{p.s.}$$

On peut faire la même chose pour la somme des $g(X_k)$, et on obtient, si $\int f \, d\nu_x < \infty$,

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} f(X_k)}{\sum_{k=0}^{n-1} g(X_k)} = \frac{\int f \, d\nu_x}{\int g \, d\nu_x} = \frac{\int f \, d\mu}{\int g \, d\mu} \;\text{ p.s.}$$

La dernière égalité vient du fait que toutes les mesures invariantes sont proportionnelles.
Pour f générale on écrit $f = f^+ - f^-$.

Corollaire

Soit (X_n) une CM irréductible récurrente.

(i) Si la chaîne est récurrente positive, et ν la probabilité invariante:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{X_k=y} = \nu(y), \quad \mathbb{P}_x \ \ p.s. \ \ \forall x, y \in E.$$

(ii) Si la chaîne est récurrente nulle:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{X_k=y} = 0, \quad \mathbb{P}_x \ \ p.s. \ \ \forall x, y \in E.$$

Remarque. Tous les théorèmes précédents sont vrais sous $\mathbb{P}_{\mu_0}$ à la place de $\mathbb{P}_x$ par additivité.

Preuve. Pour (i) on applique le 2ème théorème à $f = \delta_y$.

(ii) On applique le 1er théorème à $f = \delta_y$ et $g = 1$. Comme la CM est récurrente nulle, $\int g d\mu = \mu(E) = +\infty$.

6- Convergence en loi vers la mesure invariante

Définition-Proposition

(i) Soit x *un point récurrent de la CM. Par définition, la période de* x*, notée* $d(x)$*, est le PGCD de l'ensemble des temps de retour possibles en* x*:*

$$d(x) := \mathrm{PGCD}(L_x), \quad \textit{où} \quad L_x = \{n \geqslant 0,\ Q^n(x,x) > 0\}.$$

De plus on a $d(x)\mathbb{Z} = L_x - L_x$.

(ii) Si la CM est irréductible et récurrente alors tous les points ont la même période, notée d*, appelée* **période de la CM**. *La CM est dite* **apériodique** *si la période est 1.*

(iii) Si la CM est irréductible récurrente et apériodique, alors pour tout $x, y \in E$*, il existe* $n_0 \in \mathbb{N}$ *(dépendant de* x *et* y*) tel que*

$$Q^n(x,y) > 0, \quad \forall n \geqslant n_0.$$

Preuve.

(i) L_x est stable par addition car si $Q^{n_1}(x,x) > 0$ et $Q^{n_2}(x,x) > 0$ alors $Q^{n_1+n_2}(x,x) > 0$. Donc $L_x - L_x$ est un sous groupe de $\mathbb{Z}$. Comme $L_x \subseteq d(x)\mathbb{Z}$ car $d(x) = \mathrm{PGCD}(L_x)$, on a que $L_x - L_x$ est le sous-groupe engendré par $d(x)$.

(ii) Montrons que $d(y)\mathbb{Z} \subseteq d(x)\mathbb{Z}$, $\forall x, y \in E$. Comme $x \sim y$, $\exists n_1, n_2$ tels que $Q^{n_1}(x,y) > 0$ et $Q^{n_2}(y,x) > 0$. Donc si $n \in L(y)$ alors $Q^{n_1+n+n_2}(x,x) > 0$ et $n_1 + n + n_2 \in L_x$.
Donc, si $n \in L_y$, $k \in L_y$,

$$n - k = (n_1 + n + n_2) - (n_1 + k + n_2) \in L_x - L_x.$$

Donc,

$$L_y - L_y \subseteq L_x - L_x, \quad \text{d'où} \quad d(y)\mathbb{Z} \subseteq d(x)\mathbb{Z}.$$

Par symétrie du role de x et y, on obtient $d(x) = d(y)$.

(iii) Soit $x \in E$. Si $d(x) = d = 1$, il existe n_1 et m_1 dans L_x tels que $n_1 - m_1 = 1$. Tout entier n s'écrit

$$n = m_1 k + j, \quad \text{avec } k \geqslant 0,\ 0 \leqslant j \leqslant m_1 - 1.$$

Pour $k \geqslant m_1$, on peut écrire

$$n = m_1 k + j = m_1 k + j(n_1 - m_1) = m_1(k - j) + n_1 j \in L_x,$$

car $k - j \geqslant 0$ et $L_x + L_x \subseteq L_x$. Donc pour $n \geqslant m_1^2$, on a

$$Q^n(x, x) > 0.$$

Comme $x \rightsquigarrow y$, il existe k_1 tel que $Q^{k_1}(x, y) > 0$, et donc

$$Q^{n+k_1}(x, y) > 0,$$

pour $n \geqslant m_1^2$. Ceci conclue la preuve avec $n_0 = m_1^2 + k_1$.

Définition-Proposition

(i) Soit μ et ν deux mesures de **probabilité** *sur E. On appelle distance en variation totale entre μ et ν*

$$\mathrm{d_{TV}}(\mu,\nu) = \frac{1}{2}\sum_{x\in E}|\mu(x)-\nu(x)|.$$

(ii) On a $0 \leqslant \mathrm{d_{TV}}(\mu,\nu) \leqslant 1$, $\mathrm{d_{TV}}(\mu,\nu)=0$ ssi $\mu=\nu$, et $\mathrm{d_{TV}}(\mu,\nu)=1$ ssi μ et ν sont à supports disjoints. De plus, $\mathrm{d_{TV}}$ est une distance sur l'espace des probabilités sur E.
(iii) Si (μ_n) est une suite de probabilités et ν une probabilité, alors

$$\left(\lim_{n\to\infty}\mathrm{d_{TV}}(\mu_n,\nu)=0\right) \Rightarrow \left(\mu_n \xrightarrow{\mathcal{L}} \nu\right).$$

Remarque. En fait (iii) est une équivalence dans le cas E fini ou dénombrable, mais nous n'en aurons pas besoin.
Preuve. (ii) facile, exo.
(iii) $|\mu_n(x)-\nu(x)| \leqslant 2\,\mathrm{d_{TV}}(\mu_n,\nu)$, pour tout x, d'où l'implication.

On adoptera la notation suivante: si μ est une probabilité sur E, on note $\mathcal{L}_\mu(X_n)$ la loi de X_n partant de la loi μ (i.e. sous $\mathbb{P}_\mu$). Pour simplifier on note, $\mathcal{L}_x := \mathcal{L}_{\delta_x}$ quand la CM part de x.

Théorème

Soit (X_n) une CM irréductible, récurrente positive et apériodique, et ν sa probabilité invariante. Alors, pour toute distribution initiale μ,

$$\lim_{n\to\infty} \mathrm{d}_{\mathrm{TV}}(\mathcal{L}_\mu(X_n), \nu) = 0.$$

En particulier, sous $\mathbb{P}_\mu$,

$$X_n \xrightarrow{\mathcal{L}} \nu.$$

Remarque. La dernière assertion est équivalente à dire que, pour tout $y \in E$,

$$\lim_{n\to\infty} \mathbb{P}_\mu(X_n = y) = \nu(y).$$

On se ramène d'abord au cas $\mu = \delta_y$. On considère une suite croissante E_k de sous-ensembles finis de E telle que

$$\cup E_k = E.$$

On a

$$\begin{aligned}
\mathrm{d}_{\mathrm{TV}}(\mathcal{L}_\mu(X_n), \nu) &= \frac{1}{2} \sum_y |\mathbb{P}_\mu(X_n = y) - \nu(y)| \\
&= \frac{1}{2} \sum_y \left| \sum_x \mu(x) \mathbb{P}_x(X_n = y) - \sum_x \mu(x) \nu(y) \right| \\
&= \sum_x \mu(x) \frac{1}{2} \sum_y |\mathbb{P}_x(X_n = y) - \nu(y)| \\
&\leqslant \mu(E_k^c) + \sum_{x \in E_k} \mu(x) \, \mathrm{d}_{\mathrm{TV}}(\mathcal{L}_x(X_n), \nu).
\end{aligned}$$

Pour tout $\epsilon > 0$, on peut trouver k tel que $\mu(E_k^c) \leqslant \epsilon$. Donc si $\mathrm{d}_{\mathrm{TV}}(\mathcal{L}_x(X_n), \nu) \to 0$ pour tout x, on obtient

$$\limsup_{n \to \infty} \mathrm{d}_{\mathrm{TV}}(\mathcal{L}_\mu(X_n), \nu) \leqslant \epsilon.$$

On considère (X'_n), resp. (X''_n), deux CM **indépendantes** de même matrice de transition Q et de lois initiales resp. δ_x et ν (définies sur un espace $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$). On considère alors

$$\hat{X}_n = (X'_n, X''_n) \in E \times E.$$

Alors $(\hat{X}_n)$ est une CM de matrice de transition de transition $\overline{Q}$ sur $E \times E$ définie par,

$$\overline{Q}((x', x''), (y', y'')) = Q(x', y')Q(x'', y''), \quad \forall x', x'', y', y'' \in E,$$

et de loi initiale $\delta_x \otimes \nu$. (Facile: exo).
D'autre part, on peut vérifier que $\nu \otimes \nu$ est mesure invariante de $\overline{Q}$.
En effet, pour tout $x', x'' \in E$,

$$\begin{aligned}
&\sum_{y', y''} \nu \otimes \nu((y', y''))\overline{Q}((y', y''), (x', x'')) \\
=&\sum_{y', y''} \nu(y')\nu(y'')Q(y', x')Q(y'', x'') \\
=&(\sum_{y'} \nu(y')Q(y', x'))(\sum_{y''} \nu(y'')Q(y'', x'')) \\
=&\nu(x')\nu(x'') = \nu \otimes \nu((x', x'')).
\end{aligned}$$

Comme Q est irr. rec. apériodique, pour $x', x'', y', y'' \in E$ il existe $n'_0,\ n''_0 \in \mathbb{N}$ tel que

$$Q^n(x', y') > 0, \;\; n \geqslant n'_0, \quad \text{et} \quad Q^n(x'', y'') > 0, \;\; n \geqslant n''_0,$$

donc

$$\overline{Q}^n((x', x''), (y', y'')) > 0, \quad \forall n \geqslant \max(n'_0, n''_0).$$

On en déduit que $\overline{Q}$ est irréductible. Comme d'autre part elle admet une probabilité invariante, $\nu \otimes \nu$, elle est récurrente. (Cf Proposition de la Section 4).

On considère

$$T := \inf\{n \geqslant 0,\ X'_n = X''_n\} = \inf\{n \geqslant 0,\ \hat{X}_n \in D\},$$

où

$$D = \{(x', x'') \in E^2,\ x' = x''\}.$$

T un temps d'arrêt pour la filtration canonique de $(\hat{X}_n)$. Comme $(\hat{X}_n)$ est irréductible rec.

$$T < \infty, \quad \text{p.s.}$$

$$
\begin{aligned}
\mathrm{d}_{\mathrm{TV}}(\mathcal{L}_x(X_n),\nu) &= \frac{1}{2}\sum_{y\in E}|\mathbb{P}_x(X_n=y)-\nu(y)| \\
&= \frac{1}{2}\sum_{y\in E}\left|\hat{\mathbb{P}}(X_n'=y)-\hat{\mathbb{P}}(X_n''=y)\right| \\
&= \frac{1}{2}\sum_{y\in E}\left|\hat{\mathbb{E}}(\mathbb{1}_{X_n'=y}-\mathbb{1}_{X_n''=y})\right| \\
&\leqslant \frac{1}{2}\sum_{y\in E}\left|\hat{\mathbb{E}}(\mathbb{1}_{T\leqslant n}(\mathbb{1}_{X_n'=y}-\mathbb{1}_{X_n''=y}))\right| \\
&\quad + \frac{1}{2}\sum_{y\in E}\left|\hat{\mathbb{E}}(\mathbb{1}_{T>n}(\mathbb{1}_{X_n'=y}-\mathbb{1}_{X_n''=y}))\right|
\end{aligned}
$$

Considérons le dernier terme,

$$
\begin{aligned}
\frac{1}{2}\sum_{y\in E}\left|\hat{\mathbb{E}}(\mathbb{1}_{T>n}(\mathbb{1}_{X_n'=y}-\mathbb{1}_{X_n''=y}))\right| &\leqslant \hat{\mathbb{E}}(\mathbb{1}_{T>n}\sum_{y\in E}\frac{1}{2}(\mathbb{1}_{X_n'=y}+\mathbb{1}_{X_n''=y})) \\
&= \hat{\mathbb{P}}(T>n).
\end{aligned}
$$

En utilisant Markov fort pour (X'_n) et (X''_n),

$$
\begin{aligned}
&\hat{\mathbb{E}}(\mathbb{1}_{T\leqslant n}(\mathbb{1}_{X'_n=y} - \mathbb{1}_{X''_n=y})) \\
=&\sum_{k=0}^{n}\sum_{z\in E}\hat{\mathbb{E}}(\mathbb{1}_{T=k}\mathbb{1}_{X'_k=X''_k=z}(\mathbb{1}_{X'_n=y} - \mathbb{1}_{X''_n=y})) \\
=&\sum_{k=0}^{n}\sum_{z\in E}\hat{\mathbb{E}}(\mathbb{1}_{T=k}\mathbb{1}_{X'_k=X''_k=z}(Q^{n-k}(z,y) - Q^{n-k}(z,y)) \\
=&0
\end{aligned}
$$

On a donc montré que

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{L}_x(X_n),\nu) \leqslant \hat{\mathbb{P}}(T > n).$$

Comme $(\hat{X}_n)$ est récurrent, $\lim_{n\to\infty}\hat{\mathbb{P}}(T > n) = 0$. CQFD.

Lorsque l'espace d'états E est fini, la vitesse de convergence est géométrique.

Théorème

Si la CM est irréductible, rec. et apériodique et E est fini, alors il existe $C > 0$ et $0 < \rho < 1$ tel que

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{L}_\mu(X_n), \nu) \leqslant C\rho^n, \quad \forall n \geqslant 0.$$

Preuve. Comme E fini, il existe $n_0 \in \mathbb{N}$ tel que,

$$Q^{n_0}(x, y) > 0, \ \ \forall x, y \in E.$$

On note

$$\Pi = (\Pi(x, y))_{x,y \in E} \quad \text{où} \quad \Pi(x, y) = \nu(y).$$

Π est une matrice stochastique, c'est la matrice de transition d'un tirage i.i.d. sous ν.

Comme $\inf_{x,y} Q^{n_0}(x,y) > 0$, on peut trouver $0 < \epsilon < 1$, tel que

$$Q^{n_0} \geqslant \epsilon \Pi$$

et donc

$$Q^{n_0} = \epsilon \Pi + (1-\epsilon) P,$$

où P est une matrice stochastique. (Ceci car Π est une matrice stochastique).

On note $(Y_k)_{k\geqslant 0} = (X_{n_0 k})_{k\geqslant 0}$ qui est une CM de matrice de transition $\overline{Q} := Q^{n_0}$.
On se donne une suite de v.a. i.i.d. $(\eta_k)_{k\geqslant 1}$ de loi $\mathcal{B}(\epsilon)$.
On définit la CM $(\tilde{Y}_k)_{k\geqslant 1}$: au temps k, conditionnement à $(Y_0, \dots, Y_k)$,

$$\begin{cases} \text{Si } \eta_{k+1} = 1, \text{ on tire } Y_{k+1} \text{ suivant la loi } \nu \\ \text{Si } \eta_{k+1} = 0, \text{ on tire } Y_{k+1} \text{ suivant la loi } P(Y_k, \cdot) \end{cases}$$

Alors (exo), $(\tilde{Y}_k)$ est une CM de matrice de trans. $\overline{Q}$, càd $(\tilde{Y}_k)$ a la même loi que (Y_k).

On note alors,
$$T = \inf\{k \geqslant 0,\ \eta_k = 1\}.$$
En procédant comme précédemment avec le temps T (exo), on peut montrer que, $\forall k \geqslant 0$

$$\mathrm{d}_{\mathsf{TV}}(\mathcal{L}_\mu(X_{n_0 k}), \nu) = \mathrm{d}_{\mathsf{TV}}(\mathcal{L}_\mu(Y_k), \nu) \leqslant \mathbb{P}(T > k) = (1-\epsilon)^k.$$

L'idée est que au temps T la loi de X_T est ν, comme ν est invariante, la loi de X reste ν après ce temps là...

On montre ensuite que $\mathrm{d}_{\mathsf{TV}}(\mathcal{L}_\mu(X_n), \nu)$ est décroissant (exo). On obtient alors la décroissance géométrique voulue.

Complément: fonction harmonique

Définition

On dit que $f : E \mapsto \mathbb{R}$ *est* Q*-harmonique en* $x \in V$ *si*

$$Qf(x) = 0$$

On rappelle que $Qf(x) = \sum_{y \in E} Q(x,y) f(y)$ *(et implicitement on suppose que* $\sum_y Q(x,y) f(y)$ *est sommable.)*

On dit que f *est harmonique sur* $F \subseteq E$ *si* $Qf_{|F} = 0$ *(et simplement harmonique si harmonique sur tout* E*).*

Proposition

Soit (X_n) *une CM de mat. de trans.* Q*.*
(i) Si f *est harmonique et* $f(X_n)$ *integrable alors* $f(X_n)$ *est une martingale.*
(ii) Soit f *est harmonique sur* $F \subseteq E$ *et* $T = \inf\{n \geqslant 0, \quad X_n \in F^c\}$*. Si* $f(X_{n \wedge T})$ *est intégrable alors* $f(X_{n \wedge T})$ *est une martingale.*

Preuve. (i) Par Markov,

$$\mathbb{E}_x(f(X_{n+1}) \mid \mathcal{F}_n) = \mathbb{E}_{X_n}(f(X_1)) = Qf(X_n) = f(X_n).$$

(ii) Par Markov,

$$\begin{aligned}
&\mathbb{E}_x(f(X_{T\wedge(n+1)}) \mid \mathcal{F}_n) \\
=&\mathbb{E}_x(\mathbb{1}_{T\leqslant n} f(X_{T\wedge(n+1)}) \mid \mathcal{F}_n) + \mathbb{E}_x(\mathbb{1}_{T>n} f(X_{T\wedge(n+1)}) \mid \mathcal{F}_n) \\
=&\mathbb{E}_x(\mathbb{1}_{T\leqslant n} X_T \mid \mathcal{F}_n) + \mathbb{E}_x(\mathbb{1}_{T>n} f(X_{n+1}) \mid \mathcal{F}_n) \\
=&\mathbb{1}_{T\leqslant n} X_T + \mathbb{1}_{T>n} \mathbb{E}_{X_n}(f(X_1)) \\
=&\mathbb{1}_{T\leqslant n} X_T + \mathbb{1}_{T>n} f(X_n) \\
=&f(X_{T\wedge(n)})
\end{aligned}$$

Pour aller plus loin: equation de Dirichlet, cf e.g. notes de Le Gall page 216, Thm 13.7.2

Pour aller plus loin

- Martingales rétrogrades et applications à la LGN, à Hewitt-Savage ([1] Le Gall section 12.6) ou au thm de de Finetti (Durrett [2], Theorem 4.7.9 page 253)
- Chaînes de Markov réversibles et réseaux électriques [3], ou [4] chapter 2.

[1] Le Gall, Intégration, Probabilités et Processus Aléatoires, notes de cours, https://www.imo.universite-paris-saclay.fr/~jflegall/IPPA2.pdf

[2] Durrett, Probability: Theory and Examples, v. 5, https://services.math.duke.edu/~rtd/PTE/PTE5_011119.pdf

[3] Doyle, Snell, Random walks and electric networks, https://math.dartmouth.edu/~doyle/docs/walks/walks.pdf

[4] Lyons, Peres, Probability on Trees and Networks, http://mypage.iu.edu/~rdlyons/prbtree/book.pdf