[M. Gubinelli - Processus discrets - M1 MMD 2009/2010 - 20091224 - v.7]

II Chaînes de Markov

1 Définitions et premières propriétés

On s'intéresse ici à des processus discrets $\{X_n\}_{n \geqslant 0}$ avec une simple propriété de dépendance qui néanmoins admettent une grande variété des comportements et de possibilités de modélisation. Ils sont nommées en honneur du mathématicien russe A.Markov qui n'a introduit l'étude.

Définition 1. *Soit $\{X_n\}_{n \geqslant 0}$ un processus discret défini sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans l'ensemble dénombrable M. On dit que $\{X_n\}_{n \geqslant 0}$ est une chaîne de Markov (sur M) ssi $\forall n \geq 0$ et $\forall x_0, \ldots, x_{n+1} \in M$,*

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \tag{1}$$

Autrement dit, la loi conditionnelle de X_{n+1} sachant X_0, ..., X_n (le passé) ne dépend que de X_n (le présent). On appelle M *l'espace d'états* de la chaîne $\{X_n\}_{n \geqslant 0}$, X_0 l'*état initiale* de la chaîne et la loi de X_0 *loi initiale*.

Définition 2. *On dit que la chaîne est homogène ssi la loi conditionnelle de X_{n+1} sachant X_n ne dépends pas de n, i.e.*

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \mathbb{P}(X_1 = y | X_0 = x) = P(x, y) \qquad \forall n \geqslant 0, x, y \in M$$

La fonction P: $M \times M \to [0, 1]$ est la matrice (ou probabilité) de transition de la chaîne $\{X_n\}_{n \geqslant 0}$.

En général une *matrice de transition* sur M est une application P: $M \times M \to [0, 1]$ telle que pour tout $x \in M$

$$\sum_{y \in M} P(x, y) = 1. \tag{2}$$

(vérifier que la matrice de transition d'une chaîne de Markov satisfait cette équation).

Dans la suite on va considérer seulement des chaînes de Markov homogènes (sauf indication explicite du contraire).

Exemple 3. MARCHE ALÉATOIRE SUR $\mathbb{Z}$. Un joueur lance successivement et de manière indépendante une pièce de monnaie, éventuellement biaisée. Chaque fois qu'il obtient un Pile il reçoit un euro, chaque fois qu'il obtient un Face, il perd un euro. Soit $k_0 \in \mathbb{Z}$ sa fortune initiale. On note S_n sa fortune à l'étape n. Nous avons $S_0 = k_0$ et $S_{n+1} = S_n + X_{n+1}$, où X_{n+1} est une variable aléatoire représentant le gain (ou la perte, selon si positif ou négatif) à l'étape $n + 1$: $X_{n+1} = 1$ avec probabilité $p \in]0, 1[$, et -1 sinon. Les variables aléatoires X_1,...,X_k,... sont i.i.d. par hypothèse. La suite $(S_n)_{n \geq 0}$ est donc une chaîne de Markov homogène de matrice de transition

$$P(k, m) = \begin{cases} p & \text{si } m = k + 1; \\ 1 - p & \text{si } m = k - 1; \\ 0 & \text{autrement .} \end{cases} \qquad \forall k, m \in \mathbb{Z}$$

Exemple 4. MODÈLE DE WRIGHT-FISCHER. Ce modèle décrit l'évolution d'un ensemble de N chromosomes. On suppose qu'il y a 2 types de chromosomes, A et B, et on note X_n le nombre de chromosomes de type A présents à la génération n (il y en a donc $N - X_n$ de type B). Le modèle évolue de la façon suivante : chaque chromosome de la génération $n+1$ choisit au hasard et uniformément un chromosome parent dans la génération n, ceci indépendamment des autres chromosomes. Le chromosome fils a alors le même type que son chromosome parent. Si $X_n = i$, chaque chromosome de la génération $n + 1$ sera donc de type A avec probabilité i/N. On en déduit que la suite $X_{n\geq 0}$ est une chaîne de Markov homogène à valeurs dans $\{0, 1, ..., N\}$, de probabilité de transition

$$P(i,j) = C_N^j \left(\frac{i}{N}\right)^j \left(\frac{N-i}{N}\right)^{N-j} \qquad \forall i,j \in \{0,1,...,N\}.$$

Exemple 5. PANNES ALÉATOIRES. Soit $\{U_n\}_{n\geqslant 0}$ une suite iid à valeurs dans $\{1, 2, ..., +\infty\}$. La v.a. U_k s'interprète comme durée de vie d'une quelque machine (la k-eme) qui est remplace par un autre (la $k+1$-eme) de que elle défaille. Au temps initial 0 la machine 1 est mise en service et elle dure jusq'au temps U_1, subitement remplacée par la machine 2 que dure pour un intervalle de temps U_2 et donc jusq'au temps $U_1 + U_2$ et ainsi de suite. On note X_n le temps de service de la machine en utilisation au temps n. Le processus $\{X_n\}_{n\geqslant 0}$ est un processus à valeurs dans $\mathbb{N}$. Si $R_k = \sum_{m=1}^k U_m$ (avec la convention que $R_0 = 0$) est l'instant de remplacement de la k-eme machine ($k \geqslant 1$) on a la relation suivante:

$$X_n = n - R_k \qquad \text{pour} \quad R_k \leqslant n < R_{k+1}.$$

Le processus $\{X_n\}_{n\geqslant 0}$ est tel que $X_{R_k} = 0$ pour tout $k \geqslant 0$ et il croit linéairement dans les intervalles $[R_k, R_{k+1} - 1]$. Il est une chaîne de Markov homogène sur $\mathbb{N}$ de matrice de transition

$$P(i,j) = \begin{cases} \frac{\mathbb{P}(U_1 > i+1)}{\mathbb{P}(U_1 > i)} & \text{si } j = i+1 \ ; \\ 1 - P(i, i+1) & \text{si } j = 0 \ ; \\ 0 & \text{autrement} \, . \end{cases} \qquad \forall i,j \in \mathbb{N}$$

Remarque 6. Etant données une matrice de transition P et une loi de probabilité μ, on peut toujours construire une chaîne de Markov de matrice de transition P issue d'une variable aléatoire initiale X_0 distribuée suivant une quelconque probabilité μ sur M.

Notations. Pour tout $x \in M$ on note $\mathbb{P}_x$ la probabilité conditionnelle sachant que $X_0 = x$ (i.e. $\mathbb{P}_x(A) = \mathbb{P}(A \mid X_0 = x)$ pour tout événement $A \in \mathcal{F}$); et $\mathbb{E}_x$ l'espérance correspondante.

1.1 Systèmes dynamiques aléatoires

Soit $(\Theta, \mathcal{B}, m)$ un espace de probabilité et $\theta_1, \theta_2, ...$ une suite infinie de variables aléatoires i.i.d., à valeurs dans Θ, de loi m:

$$\mathbb{P}(\theta_i \in A) = m(A), \qquad A \in \mathcal{B}.$$

Soit $f: \Theta \times M \longrightarrow M$ une application mesurable ($\{\theta: f(\theta, x) = y\} \in \mathcal{B}, \forall x, y \in M$).

On considère X_0 une variable aléatoire indépendante de la suite $\{\theta_j\}_{j\in\mathbb{N}}$, et récursivement on pose

$$X_{n+1} = f(\theta_{n+1}, X_n), \qquad n \geq 0. \tag{3}$$

Définition 7. *(f, m) s'appelle un système dynamique aléatoire.*

Exercice 1. Montrer que la suite $(X_n, n \geq 0)$ définie par l'eq. (3) est une chaîne de Markov homogène.

Exemple 8. (Marche aléatoire sur $\mathbb{Z}$) Soit X_0 une v.a. à valeurs dans $\mathbb{Z}$ et $\{Z_n\}_{n\geqslant 1}$ une suite iid indépendant of X_0, à valeurs dans l'ensemble $\{-1, +1\}$ et tel que $\mathbb{P}(Z_1 = +1) = p$. Le processus discret défini par

$$X_n = X_{n-1} + Z_n, \qquad n \geqslant 1$$

est une chaîne de Markov homogène de même loi que la marche aléatoire sur $\mathbb{Z}$ définie dans l'exemple 3.

1.2 Equation de Chapman-Kolmogorov

Soit $\{X_n\}_{n\geq 0}$ une chaîne de Markov homogène de matrice de transition P. Pour toute fonction bornée $f\colon M\to\mathbb{R}$ on pose

$$Pf(x)=\sum_{y\in M} P(x,y)f(y)=\mathbb{E}_x[f(X_1)]. \tag{4}$$

Soit $\Pi(M)=\left\{\mu\colon M\to[0,1]\ ,\ \sum_{x\in M}\mu(x)=1\right\}$, l'ensemble des mesures de probabilité sur M. Pour toute $\mu\in\Pi(M)$ on pose

$$\mu P(x)=\sum_{y\in M}\mu(y)P(y,x). \tag{5}$$

On remarque que $\mu P\in\Pi(M)$. On définit aussi P^2, P^n, etc, par récurrence en utilisant la règle usuelle de multiplication des matrices :

$$P^{n+1}(x,y)=\sum_{z\in M}P(x,z)P^n(z,y). \tag{6}$$

Théorème 9. *Soit $\{X_n\}_{n\geq 0}$ une chaîne de Markov de matrice de transition P. On note $\mu_0\in\Pi(M)$ la loi de X_0. Alors*

1. *La loi de $(X_0,\ldots,X_n)$ est donnée par*

$$\mathbb{P}(X_0=x_0,\ldots,X_n=x_n)=\mu_0(x_0)P(x_0,x_1)\ldots P(x_{n-1},x_n), \tag{7}$$

   $\forall$ les états x_0, ..., $x_n\in M$, et $\forall n\in\mathbb{N}$. Réciproquement, tout processus $(X_n)_{n\geq 0}$ vérifiant l'équation (7) est une chaîne de Markov de matrice de transition P et de loi initiale μ_0.

2. *La loi de X_n est $\mu_n=\mu_0P^n$. Elle est donc entièrement caractérisée par μ_0 et P.*

3. *Pour toute fonction bornée $f\colon M\longrightarrow\mathbb{R}$, tout $x\in M$,*

$$\mathbb{E}_x[f(X_n)]=P^nf(x).$$

Démonstration. Exercice. □

Remarque 10. La suite $(Y_n)_{n\geq 0}$ définie par $Y_n=X_{k+n}$, k étant fixé, est aussi une chaîne de Markov de matrice de transition P.

Proposition 11. Propriété de Markov simple. *Soient $(X_n)_{n\geq 0}$ une chaîne de Markov et $\mathcal{F}_n=\sigma(X_1,\ldots,X_n)$ la tribu engendrée par $X_1,\ldots,X_n$ (i.e. la tribu représentant "le passé jusqu'à l'instant n"). Alors la propriété de Markov peut s'écrire*

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n)=\mathbb{E}(f(X_{n+1})|X_n)=Pf(X_n), \tag{8}$$

pour toute fonction bornée $f\colon M\to\mathbb{R}$.

1.3 Classification des états

Définition 12. *On dit que x* communique *avec y (et l'on note $x\to y$) ssi une des propriétés equivalentes suivantes est verifié:*

a) *il existe $n\geqslant 1$ et un $n+1$-plet d'états $(x_1=x,\ldots,x_{n+1}=y)$ tels que $P(x_i,x_{i+1})>0$ pour tout $1\leqslant i\leqslant n$.*

b) *il existe $n\geqslant 1$ tel que $P^n(x,y)>0$.*

c) *$\mathbb{P}_x(\exists k\geqslant 1\colon X_k=y)>0$.*

Exercice 2. Verifier que les propriétes a),b),c) sont equivalentes. Sugg.:

$$P(x, x_2)\cdots P(x_n, y) \leqslant P^n(x, y)$$

et

$$P^n(x, y) \leqslant \mathbb{P}_x(\exists k \geqslant 1: X_k = y) \leqslant \sum_{k \geqslant 1} P^k(x, y).$$

Si $x \rightarrow y$ et $y \rightarrow x$ ou si $x = y$ alors on dit que x et y *communiquent* et l'on note $x \leftrightarrow y$. La relation $\leftrightarrow$ est transitive, symétrique et reflexive. Elle est donc une relation d'équivalence et définit des classes $\{C_k \subseteq M\}_k$ d'éléments qui communiquent entre eux (classes de communication) et qui forment une partition de M. On dit que un ensemble $A \subseteq M$ est fermé si $x \in A, x \rightarrow y \Rightarrow y \in A$. Un état x est dit absorbant ssi $\{x\}$ est une classe fermé. Si M est formé d'une seule classe de communication (c-à-d si tout les états communiquent entre eux) on dit que la chaîne X (ou la matrice de transition P) est *irréductible*.

Exemple 13. La marche aléatoire sur $\mathbb{Z}$ et le modèle de l'urne d'Ehrenfest sont irréductibles. Par contre, le modèle de Wright-Fisher n'est pas irréductible: les états 0 et N ne communiquent qu'avec eux-même. La matrice (13) (plus en bas) n'est pas non plus irréductible.

1.4 Recurrence

Soit $N_x = \text{card}\{n \geqslant 0: X_n = x\}$ les nombres des visites à l'état x :

$$N_x = \sum_{n \geqslant 0} 1_{X_n = x}.$$

Définition 14. *Un état $x \in M$ est appelé* récurrent *si $\mathbb{P}_x(N_x = +\infty) = 1$ et* transient *si $\mathbb{P}_x(N_x = +\infty) = 0$.*

On revient toujours à un état recurrent, mais presque surement on visite un état transient seulement un nombre fini des fois. On va montrer que les état sont soit recurrent, soit transients.

Pour tout $x \in M$, on considère

$$T_x = \inf\{k > 0 : X_k = x\} \in \mathbb{N} \cup \{+\infty\} \tag{9}$$

le premier instant (strictement positif) de passage en x, avec la convention $\inf \emptyset = +\infty$. Puis, de manière récursive, on introduit

$$T_x^1 = T_x, \qquad T_x^{n+1} = \inf\{k > T_x^n : X_k = x\} \text{ pour } n \geqslant 1,$$

le $n+1$-ème instant de passage en x. Pour $n \geqslant 1$ si $T_x^{n-1} < +\infty$ soit $\tau_x^n = T_x^n - T_x^{n-1}$ (avec $T_x^0 = 0$).

Proposition 15. (REGENERATION) *Soit $x \in M$ et $n \geqslant 1$. Conditionellement à l'evenement $\{T_x^n < +\infty\}$ la loi de τ_x^{n+1} est independande de $(T_x^1, \ldots, T_x^n)$ et*

$$\mathbb{P}(\tau_x^{n+1} = k | T_x^n < +\infty) = \mathbb{P}_x(T_x = k), \qquad k \in \mathbb{N} \cup \{+\infty\}.$$

Démonstration. Il suffit de calculer la loi jointe de l'$(n+1)$-plet $(T_x^1, \ldots, T_x^n, \tau_x^{n+1})$: pour tout $1 \leqslant t_1 < t_2 < \cdots < t_n < +\infty$, en exploitant la propriété de Markov à l'instant t_n et l'homogénéité on obtient

$$\mathbb{P}(T_x^1 = t_1, \ldots, T_x^n = t_n, \tau_x^{n+1} = k) = \mathbb{P}(T_x^1 = t_1, \ldots, T_x^n = t_n, A_{t_n+1, t_n+k})$$

$$= \mathbb{P}(T_x^1 = t_1, \ldots, T_x^n = t_n)\mathbb{P}(A_{t_n+1, t_n+k} | X_{t_n} = x) = \mathbb{P}(T_x^1 = t_1, \ldots, T_x^n = t_n)\mathbb{P}_x(A_{1,k})$$

$$= \mathbb{P}(T_x^1 = t_1, \ldots, T_x^n = t_n)\mathbb{P}_x(T_x = k)$$

où $A_{t,s} = \{X_i \neq x \text{ pour } t \leqslant i < s \text{ et } X_s = x\}$ si $s < +\infty$ et $A_{t,+\infty} = \{X_i \neq x, \forall i \geqslant t\}$. On remarque que cette identité n'est pas vraie si $t_n = +\infty$. En revanche on a que

$$\mathbb{P}(T_x^1 = t_1, ..., T_x^n = t_n, \tau_x^{n+1} = k \,|T_x^n < +\infty) = \mathbb{P}(T_x^1 = t_1, ..., T_x^n = t_n |T_x^n < +\infty)\mathbb{P}_x(T_x = k)$$

pour tout $1 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant +\infty$ car $\mathbb{P}(T_x^1 = t_1, ..., T_x^n = t_n | T_x^n < +\infty) = 0$ si $t_n = +\infty$, ce qui donne la thése. □

Lemme 16. *Pour $n \geqslant 0$ on a que $\mathbb{P}_x(N_x \geqslant n) = f_x^n$ avec $f_x = \mathbb{P}_x(T_x < +\infty)$.*

Démonstration. On a que $\mathbb{P}_x(N_x \geqslant 0) = 1$ et que

$$\begin{aligned}
\mathbb{P}_x(N_x \geqslant n) &= \mathbb{P}_x(T_x^n < +\infty) \\
&= \mathbb{P}_x(T_x^{n-1} < +\infty \quad \text{et} \quad \tau_x^n < +\infty) \\
&= \mathbb{P}_x(T_x^{n-1} < +\infty)\mathbb{P}_x(T_x < +\infty) \qquad \text{(par la Prop. 15)} \\
&= \mathbb{P}_x(N_x \geqslant n-1)\, f_x
\end{aligned}$$

pour tout $n \geqslant 1$. Par recurrence on a la thése. □

Remarque 17. Pour tout v.a. X à valeurs dans $\mathbb{N}$ on a que

$$\mathbb{E}[X] = \mathbb{E}[\sum_{k \geqslant 1} 1_{k \leqslant X}] = \sum_{k \geqslant 1} \mathbb{P}(X \geqslant k) \tag{10}$$

Théorème 18. *On a la dichotomie suivante:*

i. $\mathbb{P}_x(T_x < \infty) = 1 \Rightarrow$ *l'état x est recurrent et* $\sum_{n \geqslant 0} P^n(x, x) = +\infty$;

ii. $\mathbb{P}_x(T_x < \infty) < 1 \Rightarrow$ *l'état x est transient et* $\sum_{n \geqslant 0} P^n(x, x) < +\infty$.

En particulier tout état x est soit transient, soit recurrent.

Démonstration. Si $f_x = \mathbb{P}_x(T_x < \infty) = 1$ alors par le Lemme 16 on a que

$$\mathbb{P}_x(N_x = +\infty) = \lim_{n \to \infty} \mathbb{P}_x(N_x > n) = \lim_{n \to \infty} f_x^n = 1$$

et donc x est un état recurrent et

$$\infty = \mathbb{E}_x[N_x] = \mathbb{E}_x[\sum_{n \geqslant 0} 1_{X_n = x}] = \sum_{n \geqslant 0} P^n(x, x).$$

En revanche, si $f_x < 1$ alors, par l'eq. (10) et par le Lemme 16,

$$\sum_{n \geqslant 0} P^n(x, x) = \mathbb{E}_x[N_x] = \sum_{n \geqslant 1} \mathbb{P}_x(N_x \geqslant n) = \sum_{n \geqslant 1} f_x^n = \frac{f_x}{1 - f_x} < +\infty$$

ce qu'implique aussi que $\mathbb{P}_x(N_x = +\infty) = 0$ et donc que x est un état transient. □

Théorème 19. *Si $x \leftrightarrow y$ alors il sont tout les deux du même type (soit transients, soit récurrents). Donc la récurrence ou la transience sont des propriétés des classes de communication.*

Démonstration. Si x, y communiquent, alors il existent N, M tels que $P^N(x, y) > 0$ et $P^M(y, x) > 0$. Une simple majoration donne

$$P^{2N+n+2M}(x, x) \geqslant P^N(x, y)P^{N+n+M}(y, y)P^M(y, x) \geqslant [P^N(x, y)P^M(y, x)]^2 P^n(x, x)$$

pour tout $n \geqslant 1$. Soit $\alpha = P^N(x, y)P^M(y, x) > 0$, alors on peut minorer

$$\sum_{k \geqslant 0} P^k(x, x) \geqslant \sum_{k \geqslant 2N+2M} P^k(x, x) \geqslant \alpha \sum_{k \geqslant N+M} P^k(y, y) \geqslant \alpha^2 \sum_{k \geqslant 0} P^k(x, x)$$

et donc les états x et y sont soit tout les deux transients, soit récurrents. □

Remarque 20. On dit alors d'une chaîne irréductible qu'elle est transiente, récurrente (car tout les états sont du même type).

Proposition 21. *Un ensemble $A \subseteq M$ fermé et fini contient au moins un état recurrent. Une chaîne finie et irreductible est donc recurrent.*

Démonstration. Soit $|A| < +\infty$ et supposons par absurd que pour tout $z \in A$, $\mathbb{P}_z(N_z = +\infty) = 0$. On fixe $x \in A$, pour tout $z \in A$ l'eq. (11) donne que

$$\mathbb{P}_x(N_z \geqslant r) = \mathbb{P}_x(T_y < +\infty)\mathbb{P}_z(N_z \geqslant r).$$

En prenant la limite pour $r \to +\infty$ on obtient que

$$\mathbb{P}_x(N_z = +\infty) = \mathbb{P}_x(T_z < +\infty)\mathbb{P}_z(N_z = +\infty) = 0$$

pour tout $z \in A$ e par consequent

$$1 = \mathbb{P}_x(\cap_{z \in A} \{N_z < +\infty\}) = \mathbb{P}_x(\sum_{z \in A} N_z < +\infty) = \mathbb{P}_x(\sum_{n \geqslant 0} 1_{X_n \in A} < +\infty)$$

car $\sum_{z \in A} N_z = \sum_{z \in A} \sum_{n \geqslant 0} 1_{X_n = z} = \sum_{n \geqslant 0} 1_{X_n \in A}$ est le temps passé dans A par la chaîne. L'ensemble A etant fermé on a que $\mathbb{P}_x(X_n \in A) = 1$ pour tout $n \geqslant 0$ et donc aussi que le temps passé dans A est infini (si on demarre de $x \in A$):

$$1 = \mathbb{P}_x(\forall n \geqslant 0 \colon X_n \in A) \leqslant \mathbb{P}_x(\sum_{n \geqslant 0} 1_{X_n \in A} = +\infty).$$

On obtient ainsi une contradiction. □

Exemple 22. Lorsque $\mathbb{P}_x(T_x = 1) = 1$ l'état x est *absorbant*. Par exemple les états 0 et N du modèle de Wright-Fisher sont absorbants; les autres états étant transients.

Exercice 3. Montrer que la loi de N_y sous $\mathbb{P}_x$ est

$$\mathbb{P}_x(N_y = r) = \begin{cases} f_{xy} f_{yy}^{r-1}(1 - f_{yy}) & \text{si } r \geqslant 1 \\ 1 - f_{xy} & \text{si } r = 0 \end{cases} \tag{11}$$

où $f_{xy} = \mathbb{P}_x(T_y < +\infty)$ est la probabilité de repasser par y en démarrant de x.

Solution. On pose

$$A_{n,m} = \{X_i \neq y \text{ pour } n \leqslant i < m \text{ et } X_m = y\} \quad B_n = \{X_i \neq y \text{ pour } i \geqslant n\}$$

alors

$$f_{xy} = \sum_{k \geqslant 1} \mathbb{P}_x(T_y = k) = \sum_{k \geqslant 1} \mathbb{P}_x(A_{1,k})$$

$$= \sum_{k \geqslant 1} \sum_{x_1, \ldots, x_{k-1} \neq y} P(x, x_1) P(x_1, x_2) \cdots P(x_{k-1}, y)$$

et $\mathbb{P}_x(B_1) = 1 - f_{xy}$. Supposons que $r \geqslant 1$, il est facile de voir que

$$\mathbb{P}_x(N_y = r) = \sum_{1 \leqslant n_1 < n_2 < \cdots < n_r} \mathbb{P}_x(A_{1,n_1}, A_{n_1+1,n_2}, \ldots, A_{n_{r-1}+1,n_r}, B_{n_r+1})$$

Donc, si on pose $k_i = n_{i+1} - n_i \geqslant 1$ et on utilise la propriété de Markov et l'homogénéité on trouve

$$\mathbb{P}_x(N_y = r) = \sum_{k_1, \ldots, k_r \geqslant 1} \mathbb{P}_x(A_{1,k_1}) \mathbb{P}_y(A_{1,k_2}) \cdots \mathbb{P}_y(A_{1,k_r}) \mathbb{P}_y(B_1) = f_{xy}(f_{yy})^{r-1}(1 - f_{xx}).$$

Quand $r = 0$ on a $\mathbb{P}_x(N_y = 0) = \mathbb{P}_x(B_1) = 1 - f_{xy}$.

1.5 Probabilités stationnaires

Définition 23. *Une probabilité* $\pi \in \Pi(M)$ *est dite* stationnaire *(ou* invariante*) pour la matrice de transition* P*, si* $\pi = \pi P$*, i.e.*

$$\pi(x) = \sum_{y \in M} \pi(y) P(y, x), \qquad \forall\, x \in M. \tag{12}$$

Par récurrence, on a $\pi = \pi P^n$ pour tout $n \geq 1$. Par conséquent, si l'état initial de la chaîne X_0 a pour loi π, alors X_n a la même loi π que X_0, $\forall n \geq 1$.

Exemple 24. Soient $a, b \in [0, 1]$ et

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

alors

$$\pi = \left(\frac{b}{a+b}, \frac{a}{a+b} \right).$$

Remarque 25.

1. Il peut y avoir plusieurs probabilités stationnaires. Par exemple, la matrice de transition

$$\begin{pmatrix} 1-a & a & 0 & 0 \\ b & 1-b & 0 & 0 \\ 0 & 0 & 1-a' & a' \\ 0 & 0 & b' & 1-b' \end{pmatrix}, \tag{13}$$

avec $a, b, a', b' \in [0, 1]$, admet comme probabilités stationnaires tout quadruplé de la forme

$$\left(\frac{\alpha b}{a+b}, \frac{\alpha a}{a+b}, \frac{(1-\alpha) b'}{a'+b'}, \frac{(1-\alpha) a'}{a'+b'} \right), \alpha \in [0, 1].$$

2. Lorsque M est infini, il se peut aussi qu'il n'y ait pas de probabilité stationnaire. Par exemple, dans le cas de la marche aléatoire sur $\mathbb{Z}$, l'équation (12) devient, pour tout $x \in \mathbb{Z}$,

$$\pi(x) = \pi(x-1) p + \pi(x+1)(1-p)$$

et on vérifie facilement (exercice) qu'il n'y a pas de probabilité satisfaisant cette équation. Par contre il existe des *mesures* (c'est à dire des mesure positives non-finie) satisfaisant cette équation, par exemple la mesure de comptage $\pi(x) = 1, \forall x \in \mathbb{Z}$.

Proposition 26. *Si* M *est fini, alors l'ensemble* $\mathcal{I}(P)$ *des probabilités stationnaires pour une matrice de transition* P *est un sous-ensemble non-vide, compact et convexe de* $\Pi(M)$.

Démonstration. $\Pi(M)$ est un sous-ensemble convexe, fermé et borné de $\mathbb{R}^m$, où m est le cardinal (fini) de M (exercice). En particulier $\Pi(M)$ est compact.

Soit $\mu \in \Pi(M)$ une probabilité quelconque. On considère la combinaison convexe

$$\hat{\mu}_n = \frac{1}{n} \sum_{k=0}^{n-1} \mu P^k.$$

Alors

$$\hat{\mu}_n P - \hat{\mu}_n = \frac{1}{n} \sum_{k=0}^{n-1} \left(\mu P^{k+1} - \mu P^k \right) = \frac{1}{n} (\mu P^n - \mu) \longrightarrow_{n \to \infty} 0.$$

Puisque $\Pi(M)$ est compact, il existe une sous-suite $\hat{\mu}_{n_k}$ de $\hat{\mu}_n$ convergente. Soit π sa limite. Alors

$$\pi P = \lim_k \hat{\mu}_{n_k} P = \lim_k \hat{\mu}_{n_k} = \pi,$$

donc π est stationnaire.

La compacité et la convexité de $\mathcal{I}(P)$ sont laissées en exercice. □

D'une manière générale, étant donnée une matrice de transition P, une mesure μ satisfaisant $\mu = \mu P$ est dite *invariante* pour P.

1.6 Probabilités réversibles

Définition 27. *Une probabilité $\pi \in \Pi(M)$ est dite* réversible *par rapport à P si pour tous $x, y \in M$*

$$\pi(x)P(x, y) = \pi(y)P(y, x). \tag{14}$$

Proposition 28. *Si π est réversible, alors elle est stationnaire.*

Démonstration. Exercice. □

Exemple 29. Si P est symétrique, i.e. si $P(x, y) = P(y, x)$ pour tout couple $(x, y) \in M \times M$, et si M est fini de cardinal $|M|$, alors la la probabilité uniforme $\left(\frac{1}{|M|}, \ldots, \frac{1}{|M|}\right)$ est réversible.

Exercice 4. L'URNE D'EHRENFEST. N molécules de gaz sont réparties dans un récipient divisé en deux enceintes séparées par une paroi poreuse. A chaque étape une particule choisie uniformément au hasard change d'enceinte. On note X_n le nombre de particules dans la première enceinte à l'étape n. Montrer que $(X_n)_{n \geq 0}$ est une chaîne de Markov à valeurs dans $\{1, \ldots, N\}$ de matrice de transition

$$P(i, j) = \begin{cases} \dfrac{N - i}{N} & \text{si } j = i + 1 \text{et } 0 \leq i \leq N - 1 \\ \dfrac{i}{N} & \text{si } j = i - 1 \text{et } 1 \leq i \leq N \\ 0 & \text{sinon} \end{cases}$$

et que π est une probabilité réversible ssi $\pi(i) = 2^{-N} C_N^i$.

Proposition 30. *Soit π une probabilité réversible pour P et X_0 une variable aléatoire de loi π. On fixe $n \in \mathbb{N}$. Alors la suite $\{X_j^* = X_{n-j}\}_{0 \leq j \leq n}$ est une chaîne de Markov de matrice de transition P, et X_0^* a pour loi π.*

C'est-à-dire qu'à l'équilibre, la loi de la suite $X_0, \ldots, X_n$ est invariante par retournement de temps.

Démonstration. Exercice. □

Une mesure μ (non nécessairement tel que $\mu(M) = 1$) sur M est dite réversible par rapport à P si

$$\mu(x)P(x, y) = \mu(y)P(y, x). \tag{15}$$

Proposition 31. *(*CONDITION DE CYCLE DE KOLMOGOROV*) Soit P une matrice de transition irreductible sur M. Les assertions suivantes sont équivalentes :*

a) Pour toute suite $x_0, x_1, \ldots, x_n$ dans M avec $x_n = x_0$ et telle que $\prod_{1 \leq i \leq n} P(x_i, x_{i-1}) > 0$, on a

$$\prod_{i=1}^{n} \frac{P(x_{i-1}, x_i)}{P(x_i, x_{i-1})} = 1.$$

b) Il existe une mesure μ réversible par rapport à P.

Démonstration. On voit immédiatement que b) implique a). En fait, par irreducibilité, le support de la mesure μ est tout M. Soient $x_0, ..., x_{n-1} \in M$ et $x_n = x_0$ tels que $\prod_{1 \leq i \leq n} P(x_i, x_{i-1}) > 0$. Pour la reversibilité de μ

$$\prod_{i=1}^{n} \frac{P(x_{i-1}, x_i)}{P(x_i, x_{i-1})} = \frac{\mu(x_0)}{\mu(x_n)} \prod_{i=1}^{n} \frac{P(x_{i-1}, x_i)}{P(x_i, x_{i-1})} = \prod_{i=1}^{n} \frac{\mu(x_{i-1}) P(x_{i-1}, x_i)}{\mu(x_i) P(x_i, x_{i-1})} = 1.$$

Supposons maintenant que a) est vérifiée. Soit x_0 un point quelconque de M. Pour tout $x \in M$ soit $x = x_0, x_1, ..., x_n = x$ un chemin tel que $\prod_{1 \leq i \leq n} P(x_i, x_{i-1}) > 0$ (il existe par irreducibilité de P). On pose alors $\mu(x_0) = 1$ et

$$\mu(x) = \prod_{i=1}^{n} \frac{P(x_{i-1}, x_i)}{P(x_i, x_{i-1})}.$$

La condition a) garantit que cette définition ne dépend pas du chemin choisi, et il s'ensuit facilement que μ est réversible. □

Par conséquent, quand M est fini la condition a) implique l'existence d'une probabilité réversible.

1.7 Chaînes irréductibles

Proposition 32. *Soit P une matrice de transition irréductible et supposons qu'il existe une mesure stationnaire π. Alors*

1. *$\pi(x) > 0$ pour tout $x \in M$,*
2. *$Pf = f$ implique que $f = $ constante,*
3. *tout mesure stationnaire est multiple de π et si π est une probabilité stationnaire, alors elle est la seule probabilité stationaire pour P.*

Démonstration. Soit $x \in M$ tel que $\pi(x) > 0$. Pour tout $y \in M$, il existe un entier k tel que $P^k(x, y) > 0$. Par ailleurs, par stationnarité, $\pi = \pi P^k$. On a donc

$$\pi(y) = \sum_{z \in M} \pi(z) P^k(z, y) \geq \pi(x) P^k(x, y) > 0,$$

ce qui prouve 1). Pour démontrer 2), on considère

$$\begin{aligned} &\sum_{x, y \in M} \pi(x) P(x, y) (f(x) - f(y))^2 \\ &= 2 \sum_{x \in M} \pi(x) f(x)^2 - 2 \sum_{x, y \in M} \pi(x) P(x, y) f(x) f(y) \\ &= 2 \sum_{x \in M} \pi(x) f(x) \left(f(x) - \sum_{y \in M} P(x, y) f(y) \right) = 0. \end{aligned}$$

Donc $\forall x, y \in M$

$$\pi(x) P(x, y) (f(x) - f(y))^2 = 0.$$

Vu (1), ceci implique que $\forall x, y \in M$

$$P(x, y) (f(x) - f(y))^2 = 0$$

i.e. $f(x) = f(y)$ si $P(x, y) > 0$. Par ailleurs, puisque la chaîne est irréductible, $\forall x, y \in M$, il existe un chemin $x_0 = x, x_1, ..., x_k = y$ tel que $P(x_i, x_{i-1}) > 0\ \forall\ 1 \leq i \leq k$, et donc $f(x) = f(x_1) = ... = f(y)$. Enfin, pour montrer l'unicité 3) (à moins d'un facteur constante), considérons une mesure ν (donc $\nu = \nu P$). On pose

$$Q(x, y) = P(y, x) \frac{\pi(y)}{\pi(x)} \qquad \text{et} \qquad f(x) = \frac{\nu(x)}{\pi(x)}.$$

Alors Q est une matrice de transition irréductible (exercice) et $Qf = f$. Donc f est constante et on en déduit que $\nu(x) = c\pi(x)$ pour tout $x \in M$. Si ν et π sont des mesures de probabilité alors $c = 1$ et on a l'unicité. □

1.8 Excursions

Pour tous $x, y \in M$, on compte le nombre de passages en y avant de toucher x pour la première fois :

$$N_x^y = \sum_{n=0}^{T_x - 1} 1_{[X_n = y]}. \tag{16}$$

Remarque 33.

1. Si $X_0 = x$, alors $N_x^x = 1$.
2. Si $X_0 \neq x$, alors $N_x^x = 0$.
3. $\sum_{y \in M} N_x^y = T_x$.

On introduit ensuite, pour tout $x, y \in M$

$$\mu_x(y) = \mathbb{E}_x[N_x^y] = \mathbb{E}_x[\sum_{n \geqslant 0} 1_{T_x > n, X_n = y}] \tag{17}$$

On voit que pour tout $x \in M$, μ_x définit une mesure positive sur M (qui n'est pas une probabilité en général), et que

$$\mu_x(M) = \sum_{y \in M} \mu_x(y) = \mathbb{E}_x(T_x) \in [0, \infty]. \tag{18}$$

Proposition 34. *Pour tout $x \in M$ recurrent, μ_x est une mesure stationnaire, i.e.*

$$\mu_x(y) = \sum_{z \in M} \mu_x(z) P(z, y).$$

Démonstration. Il faut d'abord remarquer que

$$\mu_x(y) = \mathbb{E}_x[\sum_{n \geqslant 0} 1_{T_x > n, X_n = y}] = \mathbb{E}_x[\sum_{n \geqslant 1} 1_{T_x \geqslant n, X_n = y}]$$

car $\mathbb{P}_x(T_x < \infty, X_0 = X_{T_x} = x) = 1$. On note aussi que $\{T_x \geqslant n\} \in \sigma(X_0, ..., X_{n-1}) = \mathcal{F}_{n-1}$ ce qui nous permet de montrer que

$$\begin{aligned}
\mathbb{E}_x[1_{T_x \geqslant n} 1_{X_n = y} 1_{X_{n-1} = z}] &= \mathbb{E}_x[\mathbb{E}[1_{T_x \geqslant n} 1_{X_n = y} 1_{X_{n-1} = z} | \mathcal{F}_{n-1}]] \\
&= \mathbb{E}_x[\mathbb{E}[1_{X_n = y} | \mathcal{F}_{n-1}] 1_{T_x \geqslant n} 1_{X_{n-1} = z}] \qquad (\{T_x \geqslant n\} \in \mathcal{F}_{n-1}) \\
&= \mathbb{E}_x[\mathbb{E}[1_{X_n = y} | X_{n-1}] 1_{T_x \geqslant n} 1_{X_{n-1} = z}] \qquad \text{(Markov)} \\
&= \mathbb{E}_x[P(X_{n-1}, y) 1_{T_x \geqslant n} 1_{X_{n-1} = z}] \qquad \text{(Def. de } P) \\
&= P(z, y) \, \mathbb{E}_x[1_{T_x \geqslant n} 1_{X_{n-1} = z}]
\end{aligned}$$

Or

$$\begin{aligned}
\mu_x(y) &= \sum_{n \geqslant 1} \mathbb{E}_x[1_{T_x \geqslant n, X_n = y} \sum_{z \in M} 1_{X_{n-1} = z}] \qquad (\text{car } \sum_{z \in M} 1_{X_{n-1} = z} = 1) \\
&= \sum_{x \in M} \sum_{n \geqslant 1} \mathbb{E}_x[1_{T_x \geqslant n} 1_{X_n = y} 1_{X_{n-1} = z}] \\
&= \sum_{x \in M} \sum_{n \geqslant 1} P(z, y) \, \mathbb{E}_x[1_{T_x \geqslant n} 1_{X_{n-1} = z}] \\
&= \sum_{x \in M} P(z, y) \underbrace{\sum_{k \geqslant 0} \mathbb{E}_x[1_{T_x > k} 1_{X_k = z}]}_{= \mu_x(z) \text{ (par l'eq. (17))}} \qquad (\text{où on pose } k = n - 1) \\
&= \sum_{z \in M} P(x, z) \, \mu_x(z)
\end{aligned}$$

□

Corollaire 35. *Si P est irréductible et recurrent, alors $\mu_x(y) < \infty$ $\forall x, y \in M$.*

Démonstration. P est irréductible, donc pour tous x, $y \in M$, il existe un entier k tel que $P^k(y,x) > 0$. Alors,

$$1 = \mu_x(x) = \sum_{z \in M} \mu_x(z) P^k(z,x) \geq \mu_x(y) P^k(y,x)$$

donc

$$\mu_x(y) \leq \left(P^k(y,x)\right)^{-1} < \infty. \qquad \square$$

Définition 36. *On dit que l'état $x \in M$ est* recurrent positif *si $\mathbb{E}_x[T_x] < +\infty$. Un état recurrent tel que $\mathbb{E}_x[T_x] = +\infty$ est appelé* recurrent nul.

Remarquons qu'il résulte de la formule (18), qu'un état $x \in M$ recurrent est récurrent positif si et seulement si μ_x est une mesure finie ($\mu_x(M) < +\infty$). Par conséquent,

Corollaire 37. *Si $|M| < \infty$ et P est irréductible, la chaîne est récurrente positive, i.e. $\mathbb{E}_x(T_x) < \infty$ pour tout $x \in M$.*

Soit maintenant x un état récurrent positif. On peut définir la probabilité sur M

$$\pi_x(y) = \frac{\mu_x(y)}{\mu_x(M)} = \frac{\mu_x(y)}{\mathbb{E}_x(T_x)}, \qquad \forall y \in M. \tag{19}$$

D'après la Proposition (34), π_x est une probabilité stationnaire. Par ailleurs, si P est irréductible, on sait, d'après la Proposition 32, qu'il existe une seule probabilité stationnaire et que tout les états sont recurrents (car au moins l'état x est recurrent). Cela signifie que on peut definit une mesure invariante μ_y pour tout $y \in M$ et par irreducibilité que $\mu_y(z) = C_{x,y}\, \mu_x(z)$ (car les mesures invariantes d'une chaîne irreductible sont toutes proportionelles). Pour la recurrence positive de x on a que $\mathbb{E}_y[T_y] = \sum_{z \in M} \mu_y(z) = C_{x,y} \sum_{z \in M} \mu_x(z) = C_{x,y} \mathbb{E}_x[T_x] < +\infty$ et donc que $\mathbb{E}_y[T_y] < +\infty$ et que tout état $y \in M$ est recurrent positif. On peut definir $\pi_y(z) = \mu_y(z) / \mathbb{E}_y[T_y]$ et par irreductibilité on obtient que $\pi_x = \pi_y$ pour tout $y \in M$. Dans ce cas, $\pi_x(z)$ ne dépend pas de x et on peut ecrire $\pi(z) = \pi_x(z)$ pour tout $x \in M$ et donc $\pi(x) = \pi_x(x) = \mu_x(x) / \mathbb{E}_x[T_x] = 1 / \mathbb{E}_x[T_x]$. D'où le résultat suivant.

Proposition 38. *Si $\{X_n\}_{n \geq 0}$ est une chaîne irréductible avec au moins un état récurrent positif, alors tous les états sont récurrent positifs et*

$$\pi(x) = \frac{1}{\mathbb{E}_x(T_x)} > 0 \tag{20}$$

est l'unique probabilité stationnaire. De plus,

$$\mathbb{E}_x(N_x^y) = \frac{\pi(y)}{\pi(x)}. \tag{21}$$

Exemple 39. On peut montrer que dans le cas de la marche aléatoire sur $\mathbb{Z}$, la chaîne est
(a) transiente si $p \neq 1/2$
(b) récurrente nulle si $p = 1/2$.

Exemple 40. Par le Corollaire (37) dans le modèle de l'urne d'Ehrenfest, la chaîne est récurrente positive (espace d'états fini et chaîne irreductible).