Chapitre 2

Chaînes de Markov

Introduction

Une chaîne de Markov est une suite aléatoire $\{X_n; n = 0, 1, 2, \ldots\}$, définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs dans un ensemble E qui peut être arbitraire, mais qui sera ici soit fini soit dénombrable, et qui jouit de la propriété de Markov. Sans donner tout de suite une définition formelle, indiquons qu'une suite de Markov a la propriété que connaissant X_n, on peut oublier le passé pour prédire l'avenir. Une façon de fabriquer une telle suite est de se donner une suite de v.a.$\{Y_n,\ n \geq 1\}$ qui sont mutuellement indépendantes, et globalement indépendantes de X_0, à valeurs dans F, et une application $f : \mathbb{N} \times E \times F \to E$, t.q. pour $n \geq 1$,

$$X_n = f(n, X_{n-1}, Y_n).$$

D'une certaine façon, c'est le modèle le plus simple d'une suite de v. a. non indépendantes.

Les deux chapitres qui suivent donneront un aperçu des nombreuses applications des chaînes de Markov. Notons que nous allons présenter la théorie des chaînes de Markov *homogènes* (dans la formule de récurrence ci–dessus, cela revient à prendre f indépendante de n, et les Y_n de même loi), même si dans plusieurs applications on a besoin de chaînes non homogènes. Même dans ces cas–là, les propriétés asymptotiques des chaînes homogènes sont cruciales.

2.1 Définition et propriétés élémentaires.

Nous allons définir et étudier les chaînes de Markov $\{X_n; n \in \mathbb{N}\}$ à valeurs dans un espace (fini ou) dénombrable E. On notera $x, y, \ldots$ les points de E. Pour alléger la rédaction, on conviendra que pour chaque condition faisant intervenir une probabilité conditionnelle $\mathbb{P}(A|B)$, celle–ci n'est supposée vérifiée que dans le cas où $\mathbb{P}(B) > 0$.

Définition 2.1.1. *Le processus stochastique $\{X_n; n \in \mathbb{N}\}$ à valeurs dans E est appelé une chaîne de Markov si pour tout $n \in \mathbb{N}$, la loi conditionnelle de X_{n+1} sachant $X_0, X_1, \ldots, X_n$ est égale à la loi conditionnelle sachant X_n, i.e $\forall x_0, \ldots, x_n,\ y \in E$,*

$$\mathbb{P}(X_{n+1} = y | X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_{n+1} = y | X_n = x_n).$$

Un critère simple qui permet dans la plupart des cas de vérifier qu'un processus est une chaîne de Markov est donné par le :

Lemme 2.1.2. *Soit E et F deux ensembles dénombrables, et f une application de $\mathbb{N} \times E \times F$ dans E. Soit $X_0, Y_1, Y_2, \ldots$ des v.a. mutuellement indépendantes, X_0 à valeurs dans E et les Y_n à valeurs dans F, et $\{X_n,\ n \geq 1\}$ le processus à valeurs dans E défini par*

$$X_{n+1} = f(n, X_n, Y_{n+1}),\ n \in \mathbb{N}.$$

Alors $\{X_n,\ n \in \mathbb{N}\}$ est une chaîne de Markov.

Preuve

$$\begin{aligned}
\mathbb{P}(X_{n+1} = y | X_0 = x_0, \ldots, X_n = x_n)& \\
&= \frac{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n, X_{n+1} = y)}{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)} \\
&= \sum_{\{z; f(n, x_n, z) = y\}} \frac{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n, Y_{n+1} = z)}{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)} \\
&= \sum_{\{z; f(n, x_n, z) = y\}} \mathbb{P}(Y_{n+1} = z) \\
&= \frac{\mathbb{P}(X_n = x_n, X_{n+1} = y)}{\mathbb{P}(X_n = x_n)}
\end{aligned}$$

□

Une chaîne de Markov est l'analogue d'un système déterministe défini par une relation de récurrence du type :

$$x_{n+1} = f(n, x_n),$$

par opposition aux systèmes "avec mémoire"du type :

$$x_{n+1} = f(n, x_n, x_{n-1}, \dots, x_1, x_0).$$

Ici la fonction $f(n, \cdot)$ est remplacée par la "matrice de transition" :

$$P_{xy} = \mathbb{P}(X_{n+1} = y | X_n = x).$$

Dans toute la suite, cette matrice $P = (P_{xy};\ x, y \in E)$ sera indépendante de l'instant n. On dit que la chaîne de Markov est **homogène**.

La matrice P est dite **markovienne** (ou **stochastique**), au sens où elle vérifie la propriété que $\forall x \in E$, le vecteur ligne $(P_{xy};\ y \in E)$ est une mesure de probabilité sur E, c'est à dire :

$$P_{xy} \geq 0,\ \forall y \in E;\ \sum_{y \in E} P_{xy} = 1.$$

Comme on va le voir, la loi d'une chaîne de Markov est entièrement déterminée par la donnée d'une "loi initiale" $(\mu_x;\ x \in E)$, qui est la loi de X_0, et de la matrice de transition de la chaîne.

Définition 2.1.3. *Soit μ une probabilité sur E, et P une matrice markovienne. Un processus stochastique $\{X_n;\ n \in \mathbb{N}\}$ à valeurs dans E, et défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, est une chaîne de Markov (μ, P) (i.e. de loi initiale μ et de matrice de transition P) si :*

(i) $\mathbb{P}(X_0 = x) = \mu_x,\ \forall x \in E.$

(ii) $\mathbb{P}(X_{n+1} = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x) = P_{xy},$
$\forall x_0, \dots, x_{n-1}, x, y.$

Proposition 2.1.4. *Une CNS pour qu'un processus $\{X_n,\ n \in \mathbb{N}\}$ à valeurs dans E soit une chaîne de Markov (μ, P) est que $\forall n \in \mathbb{N}$, la loi du v.a. $(X_0, X_1, \dots, X_n)$ soit donnée par*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n-1} x_n}.$$

Preuve

CN. Si $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) > 0$,

$$\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_n = x_n | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) \\ \times \cdots \times \mathbb{P}(X_1 = x_1 | X_0 = x_0)\mathbb{P}(X_0 = x_0)$$

et l'égalité ci–dessus résulte de la définition. Sinon, les deux membres de l'égalité de l'énoncé sont nuls (considérer le plus petit indice k tel que $\mathbb{P}(X_0 = x_0, \ldots, X_k = x_k) = 0$).

CS. Le (i) de la définition résulte de l'égalité de l'énoncé. Le (ii) s'en déduit aisément. Démontrons en fait plus que (ii).

$$\mathbb{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+p} = x_{n+p} | X_0 = x_0, \ldots, X_n = x_n) \\ = \frac{\mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n+p-1} x_{n+p}}}{\mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n-1} x_n}}$$

(ii) s'en déduit en choisissant $p = 1$. □

On a en fait montré le :

Corollaire 2.1.5. *Si $\{X_n;\ n \in \mathbb{N}\}$ est une chaîne de Markov (μ, P), alors pour tous $n, p, x_0, \ldots, x_{n+p}$*

$$\mathbb{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+p} = x_{n+p} | X_0 = x_0, \ldots, X_n = x_n) \\ = P_{x_n x_{n+1}} \times \cdots \times P_{x_{n+p-1} x_{n+p}}.$$

Une probabilité μ sur E est considérée comme un vecteur ligne, une application $g : E \to \mathbb{R}$ est considérée comme un vecteur colonne, ce qui justifie les notations

$$(\mu P)_y = \sum_{x \in E} \mu_x P_{xy}$$

$$(Pg)_x = \sum_{y \in E} P_{xy} g_y,$$

et l'intégrale d'une fonction g par rapport à une mesure μ s'écrit (si la somme converge absolument)

$$\mu g = \sum_{x \in E} \mu_x g_x$$

Proposition 2.1.6. *Soit $\{X_n, n \in \mathbb{N}\}$ une chaîne de Markov (μ, P). Alors* *(i)* $\mathbb{P}(X_n = y|X_0 = x) = \mathbb{P}(X_{n+p} = y|X_p = x) = (P^n)_{xy}$

(ii) $\mathbb{P}(X_n = y) = (\mu P^n)_y$

(iii) $\mathbb{E}[g(X_n)|X_0 = x] = (P^n g)_x$

Preuve

(i)

$$\begin{aligned}\mathbb{P}(X_n = y|X_0 = x) &= \sum_{x_1,\dots,x_{n-1}} \mathbb{P}(X_n = y, X_{n-1} = x_{n-1}, .., X_1 = x_1|X_0 = x)\\ &= \sum_{x_1,\dots,x_{n-1}} \frac{\mu_x P_{xx_1} \times \cdots \times P_{x_{n-1}y}}{\mu_x}\\ &= \sum_{x_1,\dots,x_{n-1}} P_{xx_1} \times \cdots \times P_{x_{n-1}y}\\ &= (P^n)_{xy}\end{aligned}$$

(ii) On remarque que

$$\begin{aligned}\mathbb{P}(X_n = y) &= \sum_{x\in E} \mathbb{P}(X_n = y, X_0 = x)\\ &= \sum_{x\in E} \mathbb{P}(X_n = y|X_0 = x)\mu_x,\end{aligned}$$

et on utilise (i).

(iii) On utilise à nouveau (i) à partir de :

$$\mathbb{E}[g(X_n)|X_0 = x] = \sum_{y\in E} g_y \mathbb{P}(X_n = y|X_0 = x)$$

2.2 Exemples

2.2.1 Marche aléatoire sur $E = \mathbf{Z}^d$

Soit $\{Y_n; n \in \mathbb{N}^*\}$ une suite de v.a. i.i.d. à valeurs dans $\mathbf{Z}^d$, de loi commune λ, et X_0 une v.a. à valeurs dans $\mathbf{Z}^d$, indépendante des Y_n. Alors la suite $\{X_n,\ n \geq 0\}$ définie par

$$X_{n+1} = X_n + Y_{n+1},\ \in \mathbb{N}$$

est une chaîne de Markov (μ, P), avec μ=loi de X_0, et $P_{xy} = \lambda_{y-x}$. Le cas le plus classique des marches aléatoires dans $\mathbf{Z}^d$ est le cas des marches aléatoires symétriques partant de 0, i.e.

$$\mu = \delta_0$$
$$\lambda_{\pm e_1} = \frac{1}{2d},$$

où $(e_1, \ldots, e_d)$ est une base orthonormée de $\mathbb{R}^d$.

2.2.2 Processus de Galton–Watson

C'est un processus de branchement $\{Z_n; n \in \mathbb{N}\}$ où Z_n désigne le nombre d'individus de sexe mâle de la n–ième génération portant un nom donné, ces individus étant tous issus d'un même ancêtre, l'unique mâle formant la génération 0 ($Z_0 = 1$ p.s.). On suppose que chacun des mâles de la n–ième génération engendre ξ_i^n enfants mâles ($1 \leq i \leq Z_n$) de telle sorte que

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^n.$$

Notre hypothèse fondamentale est que les v.a. $\{\xi_i^n, i = 1, 2, \ldots, n = 0, 1, 2, \ldots\}$ sont i.i.d., de telle sorte que en particulier Z_n et $\{\xi_1^n, \ldots, \xi_p^n, \ldots\}$ sont indépendants.

Les $\{Z_n,\ n \in \mathbb{N}\}$ forment une chaîne de Markov (μ, P), à valeurs dans $\mathbb{N}$, avec $\mu = \delta_1$,

$$P_{xy} = (p^{*x})_y,$$

où p^{*x}, la x–ème puissance de convolution de la loi p sur $\mathbb{N}$ des ξ_n^k, i.e. la loi de la somme de x v.a. i.i.d. de loi commune p.

2.2.3 File d'attente en temps discret

On considère une file d'attente qui se forme à un guichet. X_n désigne le nombre de clients dans la file en attente ou en train de se faire servir à l'instant n. Entre les instants n et $n+1$ arrivent Y_{n+1} clients, et si $X_n > 0$ partent Z_{n+1} clients. On suppose que X_0, Y_1, Z_1, Y_2, $Z_2 \ldots$ sont indépendantes, vérifiant $0 < \mathbb{P}(Y_n = 0) < 1$, et les Z_n vérifiant $\mathbb{P}(Z_n = 1) = p = 1 - \mathbb{P}(Z_n = 0)$. C'est à dire que

$$X_{n+1} = X_n + Y_{n+1} - \mathbf{1}_{\{X_n > 0\}} Z_{n+1}.$$

2.3 Propriété de Markov forte

Soit $\{X_n; n \in \mathbb{N}\}$ une chaîne de Markov à valeurs dans E, définie sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Etant donné μ une probabilité sur E, on utilisera la notation $\mathbb{P}_\mu$ pour désigner n'importe quelle probabilité sur $(\Omega, \mathcal{F})$ telle que sous $\mathbb{P}_\mu$ la suite $\{X_n,\ n \geq 0\}$ est une chaîne de Markov de loi initiale μ, i. e. la loi de X_0 est la probabilité μ, soit

$$\mathbb{P}_\mu(X_0 = x) = \mu_x,\ x \in E.$$

Dans le cas $\mu = \delta_x$, on notera $\mathbb{P}_x$ pour $\mathbb{P}_{\delta_x}$. $\mathbb{P}_x$ s'interprète comme la probabilité conditionnelle, sachant que $X_0 = x$. Pour tout $n \geq 0$, on notera $\mathcal{F}_n$ la tribu des événements "déterminés par $X_0, X_1, \ldots, X_n$", i.e.

$$\mathcal{F}_n = \left\{\{\omega; (X_0(\omega), \ldots, X_n(\omega)) \in B_n\},\ B_n \in \mathcal{P}(E^{n+1})\right\},$$

où l'on a utilisé une notation qui reviendra de temps en temps dans ce livre, à savoir $\mathcal{P}(F)$ désigne l'ensemble des parties (ou sous–ensembles) de l'ensemble F.

Théorème 2.3.1. *Soit $\{X_n; n \geq 0\}$ une chaîne de Markov (μ, P). Alors pour tout $n \in \mathbb{N}$, $x \in E$ conditionnellement en $\{X_n = x\}$, $\{X_{n+p}; p \geq 0\}$ est une chaîne de Markov (δ_x, P) indépendante de $(X_0, \ldots, X_n)$, autrement dit pour tout $A \in \mathcal{F}_n$ et pour tout $m > 0$, $x_1, \ldots, x_m \in E$,*

$$\begin{aligned}&\mathbb{P}\left(A \cap \{X_{n+1} = x_1, \ldots, X_{n+m} = x_m\} | X_n = x\right)\\ &\quad = \mathbb{P}(A | X_n = x)\mathbb{P}_x(X_1 = x_1, \ldots, X = x_m)\end{aligned}$$

Preuve

Il suffit de faire la preuve dans le cas $A = \{X_0 = y_0, X_1 = y_1, \ldots, X_n = y_n\}$ (A est une réunion au plus dénombrable d'événements disjoints de cette forme, et le cas général s'en déduira donc par la σ–additivité de $\mathbb{P}$). Il suffit de considérer le cas $y_n = x$, car dans le cas contraire les deux membres de l'égalité sont nuls. Le membre de gauche de l'égalité de l'énoncé vaut

$$\frac{\mathbb{P}(X_0 = y_0, \ldots, X_n = x, X_{n+1} = x_1, \ldots, X_{n+m} = x_m)}{\mathbb{P}(X_n = x)}$$

ce qui vaut, en appliquant deux fois la Proposition 2.1.4,

$$\frac{\mathbb{P}(A)}{\mathbb{P}(X_n = x)} \times P_{xx_1} \times P_{x_1x_2} \times \cdots \times P_{x_{m-1}x_m},$$

soit

$$\mathbb{P}(A|X_n = x)\mathbb{P}_x(X_1 = x_1, \dots, X_m = x_m).$$

□

Le résultat précédent dit en particulier que le passé et le futur sont conditionnellement indépendants, sachant la position de la chaîne à l'instant présent n.

Nous allons maintenant étendre la propriété de Markov en remplaçant l'instant fixe n par un instant aléatoire qui vérifie une propriété particulière.

Définition 2.3.2. *Une v.a.* T *à valeurs dans* $\mathbb{N} \cup \{+\infty\}$ *est appelée un* **temps d'arrêt** *si* $\forall n \in \mathbb{N}$,

$$\{T = n\} \in \mathcal{F}_n.$$

Autrement dit, l'observation de $X_0, X_1, \dots, X_n$, la chaîne jusqu'à l'instant n, permet de décider si oui ou non T vaut n.

Exemple 2.3.3. *i)* $\forall i \in n$, *l'instant* S_i *du premier passage à l'état* x *:*

$$S_x = \begin{cases} \inf\{n \geq 0; X_n = x\} & \textit{si un tel } n \textit{ existe}, \\ +\infty, & \textit{sinon}; \end{cases}$$

est un temps d'arrêt, ainsi que l'instant du "premier retour " à l'état x *:*

$$T_x = \begin{cases} \inf\{n \geq 1; X_n = x\} & \textit{si un tel } n \textit{ existe}, \\ +\infty, & \textit{sinon}. \end{cases}$$

Remarque 2.3.4. *Avec la convention que l'inf de l'ensemble vide vaut* $+\infty$, *il est inutile de définir ces temps sur deux lignes :*

$$T_x = \inf\{n \geq 1;\ X_n = x\}.$$

T_x *est bien un temps d'arrêt, puisque*

$$\{T_x = n\} = \{X_1 \neq x\} \cap \dots \cap \{X_{n-1} \neq x\} \cap \{X_n = x\}.$$

ii) $\forall A \subset E$, *le temps d'atteinte de* A

$$T_A = \inf\{n \geq 1; X_n \in A\}$$

est un temps d'arrêt.
iii) Par contre, l'instant de dernier passage en A

$$L_A = \sup\{n \geq 1; X_n \in A\}$$

n'est pas un temps d'arrêt.

On notera $\mathcal{F}_T$ la tribu des événements "déterminés par $X_0, X_1, \dots, X_T$", définie comme la tribu des événements $B \in \mathcal{F}$ tels que $\forall n \in \mathbb{N}$,

$$B \cap \{T = n\} \in \mathcal{F}_n.$$

Théorème 2.3.5. *(Propriété de Markov forte) Soit* $\{X_n : n \geq 0\}$ *une chaîne de Markov* (μ, P)*, et* T *un temps d'arrêt. Conditionnellement en* $\{T < \infty\} \cap \{X_T = x\}$, $\{X_{T+n}; n \geq 0\}$ *est une chaîne de Markov* (δ_x, P) *indépendante de* $\mathcal{F}_T$*. Autrement dit, pour tout* $A \in \mathcal{F}_T$*, et pour tout* $m > 0$, $x_1, \dots, x_m \in E$,

$$\begin{aligned}
&\mathbb{P}(A \cap \{X_{T+1} = x_1, \dots, X_{T+m} = x_m\} | X_T = x, T < \infty) \\
&= \mathbb{P}(A | X_T = x,\ T < \infty) \times \mathbb{P}_x(X_1 = x_1, \dots, X_m = x_m)
\end{aligned}$$

Preuve Il suffit de montrer que $\forall n \in \mathbb{N}$,

$$\begin{aligned}
&\mathbb{P}(A \cap \{T = n\} \cap \{X_{T+1} = x_1, \dots, X_{T+m} = x_m\} | X_T = x) \\
&= \mathbb{P}(A \cap \{T = n\} | X_T = x) \mathbb{P}_x(X_1 = x_1, \dots, X_m = x_m),
\end{aligned}$$

ce qui résulte du Théorème 2.3.1, puis de sommer sur n.

2.4 Etats récurrents et transitoires

On définit comme ci–dessus

$$T_x = \inf\{n \geq 1; X_n = x\}, \text{ et on pose la}$$

Définition 2.4.1. *L'état* $x \in E$ *est dit* **récurrent** *si* $\mathbb{P}_x(T_x < \infty) = 1$*, et* **transitoire** *dans le cas contraire (i.e. si* $\mathbb{P}_x(T_x < \infty) < 1$*).*

On définit le nombre de retours à l'état x :

$$N_x = \sum_{n \geq 1} \mathbf{1}_{\{X_n = x\}}$$

Proposition 2.4.2. *a) Si x est récurrent,*

$$\mathbb{P}_x(N_x = +\infty) = 1.$$

b) Si x est transitoire,

$$\mathbb{P}_x(N_x = k) = (1 - \Pi_x)\Pi_x^k, \ k \geq 0,$$

avec $\Pi_x = \mathbb{P}_x(T_x < \infty)$ (en particulier $N_x < \infty$, $\mathbb{P}_x$ p.s.)

Preuve Posons

$$\begin{aligned} T_x^2 &= \inf\{n > T_x, X_n = x\} \\ &= T_x + \inf\{n > 0, X_{T_x+n} = x\}. \end{aligned}$$

On vérifie (exercice) que T_x^2 est un temps d'arrêt.

$$\begin{aligned} \mathbb{P}_x(T_x^2 < \infty) &= \mathbb{P}_x(T_x^2 < \infty | T_x < \infty)\mathbb{P}_x(T_x < \infty) \\ &= \sum_{n=1}^{\infty} \mathbb{P}_x(T_x^2 = T_x + n | T_x < \infty)\mathbb{P}_x(T_x < \infty). \end{aligned}$$

Mais, en utilisant le Théorème 2.3.5 on obtient

$$\begin{aligned} \mathbb{P}_x(T_x^2 =& T_x + n | T_x < \infty) \\ &= \mathbb{P}_x(X_{T_x+1} \neq x, \ldots, X_{T_x+n-1} \neq x, X_{T_x+n} = x | T_x < \infty) \\ &= \mathbb{P}_x(X_1 \neq x, \ldots, X_{n-1} \neq x, X_n = x) \\ &= \mathbb{P}_x(T_x = n) \end{aligned}$$

Finalement

$$\begin{aligned} \mathbb{P}_x(T_x^2 < \infty) = (\mathbb{P}_x(T_x < \infty))^2, \ \text{soit} \\ \mathbb{P}_x(N_x \geq 2) = (\mathbb{P}_x(T_x < \infty))^2, \end{aligned}$$

et on montre en itérant le même argument que

$$\mathbb{P}_x(N_x \geq k) = (\mathbb{P}_x(T_x < \infty))^k, \ \ k \in \mathbb{N}.$$

Les deux parties de la Proposition se déduisent aisément de cette relation.

Corollaire 2.4.3. *L'état x est récurrent ssi*

$$\sum_{n=0}^{\infty} (P^n)_{xx} = +\infty$$

PREUVE

$$\begin{aligned}\mathbb{E}_x(N_x) &= \sum_{n\geq 1} \mathbb{P}_x(X_n = x) \\ &= \sum_{n\geq 1} (P^n)_{xx}\end{aligned}$$

D'après la proposition, cette quantité est infinie lorsque x est récurrent. Mais si x est transitoire,

$$\begin{aligned}\mathbb{E}_x(N_x) &= \sum_{k=1}^{\infty} k(1-\Pi_x)\Pi_x^k \\ &= \prod_x (1-\Pi_x)^{-1} < \infty\end{aligned}$$

Définition 2.4.4. *On dit que l'état y est* **accessible** *à partir de x (noté $x \to y$) s'il existe $n \geq 0$ tel que $\mathbb{P}_x(X_n = y) > 0$. On dit que les états x et y* **communiquent** *(noté $\leftrightarrow$) si $x \to y$ et $y \to x$.*

La relation $x \leftrightarrow y$ est une relation d'équivalence (exercice), et on peut donc partitionner E en classes d'équivalence modulo la relation $\leftrightarrow$.

Notons que $x \to y \Leftrightarrow \exists\, n \geq 0$ t.q. $(P^n)_{xy} > 0$, puisque $\mathbb{P}_x(X_n = y) = (P^n)_{xy}$ (Proposition 2.1.6(i)).

Théorème 2.4.5. *Soit $C \subset E$ une classe d'équivalence pour la relation $\leftrightarrow$. Alors tous les états de C sont soit récurrents soit transitoires.*

PREUVE Soit $x, y \in C$. Il suffit de montrer que x transitoire $\Rightarrow$ y transitoire (car alors par l'absurde y récurrent $\Rightarrow$ x récurrent). Comme $x \leftrightarrow y$, d'après la dernière remarque $\exists\, n, m > 0$ t.q. $(P^n)_{xy} > 0$ et $(P^m)_{yx} > 0$. Mais $\forall r \geq 0$,

$$(P^{n+r+m})_{xx} \geq (P^n)_{xy}(P^r)_{yy}(P^m)_{yx})$$

et

$$\sum_{r=0}^{\infty} (P^r)_{yy} \leq \frac{1}{(P^n)_{xy}(P^m)_{yx}} \sum_{n=0}^{\infty} (P^{n+r+m})_{xx} < \infty.$$

Définition 2.4.6. *Une chaîne de Markov* (μ, P) *est dite* **irréductible** *si* E *est constitué d'une unique classe d'équivalence. Elle est dite* **récurrente irréductible** *si elle est irréductible et si tous les états sont récurrents.*

Proposition 2.4.7. *Toute chaîne de Markov irréductible sur un espace* **fini** E *est récurrente irréductible.*

PREUVE Par l'absurde, puisque E est fini, il y a au moins un état qui est visité une infinité de fois avec une probabilité non nulle, donc p.s. d'après la dichotomie de la Proposition 2.4.2, et cet état (donc tous les états) est (sont) récurrent(s).

2.5 Le cas récurrent irréductible

Dans cette section, on suppose la chaîne irréductible récurrente. Nous allons commencer par étudier les **excursions** successives de la chaîne entre deux retours successifs à l'état x :

$$\mathcal{E}_k = (X_{T_x^k}, X_{T_x^k+1}, \ldots, X_{T_x^{k+1}}),\ k \geq 0.$$

Ces excursions sont des suites aléatoires de longueur aléatoire finie ≥ 2, formées d'états de E distincts de x, sauf le premier et le dernier qui sont égaux à x. Notons U l'ensemble des suites

$$u = (x, x_1, \ldots, x_n, x),$$

avec $n \geq 0$, $x_\ell \neq x$, $1 \leq \ell \leq n$. U est dénombrable, et constitue l'ensemble des valeurs possible des excursions $\mathcal{E}_0, \mathcal{E}_1, \ldots$. Ces v.a. prennent donc leurs valeurs dans un ensemble dénombrable, et leur loi de probabilité est caractérisée par les quantités

$$\mathbb{P}(\mathcal{E}_k = u),\ \ u \in U.$$

Proposition 2.5.1. *Sous* $\mathbb{P}_x$ *la suite* $(\mathcal{E}_0, \mathcal{E}_1, \ldots)$ *des excursions est i.i.d., c'est à dire qu'il existe une probabilité* $\{p_u,\ u \in U\}$ *telle que pour tout* $k > 0$, $u_0, \ldots, u_k \in U$,

$$\mathbb{P}_x(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \ldots, \mathcal{E}_k = u_k) = \prod_{\ell=0}^{k} p_{u_\ell}.$$

PREUVE C'est une conséquence de la propriété de Markov forte. En effet, $\{\mathcal{E}_0 = u_0\} \in \mathcal{F}_{T_x}$, et l'événement

$$\{\mathcal{E}_1 = u_1, \dots, \mathcal{E}_k = u_k\}$$

est de la forme

$$\{X_{T_x+1} = x_1, \dots, X_{T_x+p} = x_p\},$$

pour $p > 0$, $x_1, \dots, x_p \in E$. Donc

$$\begin{aligned}
&\mathbb{P}_x(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \dots, \mathcal{E}_k = u_k) \\
&= \mathbb{P}_x(\{\mathcal{E}_0 = u_0\} \cap \{X_{T_1+1} = x_1, \dots, X_{T_1+p} = x\} | T_x < \infty) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(X_1 = x_1, \dots, X_p = x) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(\mathcal{E}_0 = u_1, \dots, \mathcal{E}_{k-1} = u_k) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(\mathcal{E}_0 = u_1) \times \ldots \times \mathbb{P}_x(\mathcal{E}_0 = u_k) \\
&= p_{u_0} p_{u_1} \times \cdots \times p_{u_k},
\end{aligned}$$

si $\{p_u, u \in U\}$ est la loi de $\mathcal{E}_0$ sous $\mathbb{P}_x$. □

On appelle mesure sur E un "vecteur ligne" $\{\gamma_x; x \in E\}$ tel que $0 \le \gamma_x < \infty$, $\forall x$. Dans le cas où la mesure est finie, $\sum_{x \in E} \gamma_x < \infty$, on peut la normaliser pour en faire une probabilité sur E, $\left(\frac{\gamma_x}{\sum_z \gamma_z},\ x \in E\right)$. Une mesure γ est dite invariante (par la matrice de transition P) si

$$\gamma P = \gamma, \text{ i.e.}$$

$$\sum_{y \in E} \gamma_y P_{yx} = \gamma_x,\ x \in E.$$

Une mesure γ sera dite strictement positive si $\gamma_x > 0,\ \forall x \in E$.

Une mesure de probabilité γ est invariante ssi la chaîne (γ, P) a la propriété que γ est la loi de X_n, $\forall n \in \mathbb{N}$, donc $\forall n$, $\{X_{n+m}; m \in \mathbb{N}\}$ est une chaîne (γ, P).

Remarque 2.5.2. *Soit π une probabilité invariante, c'est à dire une probabilité qui vérifie $\pi P = \pi$, ou encore $\forall x \in E$,*

$$\sum_{y \neq x} \pi_y P_{yx} = \pi_x (1 - P_{xx}),$$

soit

$$\mathbb{P}(X_n \neq x, X_{n+1} = x) = \mathbb{P}(X_n = x, X_{n+1} \neq x),$$

ce qui signifie qu'à l'équilibre, le nombre moyen de départs de l'état x entre les instants n et $n+1$ est égal au nombre moyen d'arrivées à l'état x entre n et $n+1$. On voit que l'équation qui caractérise la probabilité invariante est très intuitive.

Théorème 2.5.3. *Soit $\{X_n; n \in \mathbb{N}\}$ une chaîne de matrice de transition P récurrente irréductible. Alors il existe une mesure strictement positive invariante γ, unique à une constante multiplicative près.*

PREUVE **Existence** Posons γ_y^x=nombre moyen de visites à l'état y lors de l'excursion $\mathcal{E}_0$ partant de x, soit

$$\begin{aligned}
\gamma_y^x &= \mathbb{E}_x \sum_{n=1}^{T_x} \mathbf{1}_{\{X_n=y\}} \\
&= \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = y, n \leq T_x) \\
&= \sum_{z\in E} \sum_{n=1}^{\infty} \mathbb{P}_x(\{X_{n-1} = z, n-1 < T_x\} \cap \{X_n = y\}) \\
&= \sum_{z\in E} \left(\sum_{n=2}^{\infty} \mathbb{P}_x(X_{n-1} = z, n-1 \leq T_x) \right) P_{zy} \\
&= (\gamma^x P)_y.
\end{aligned}$$

Notons que l'on a utilisé la récurrence à l'avant dernière égalité. On va maintenant utiliser l'irréductibilité de la chaîne. $\exists\ n, m$ tels que $(P^n)_{xy} > 0$, $(P^m)_{yx} > 0$. Donc, puisque $\gamma_x^x = 1$,

$$\begin{aligned}
0 < (P^n)_{xy} = \gamma_x^x (P^n)_{xy} \leq (\gamma^x P^n)_y = \gamma_y^x \\
\gamma_y^x (P^m)_{yx} \leq (\gamma^x P^m)_x = \gamma_x^x = 1.
\end{aligned}$$

Donc γ^x est bien une mesure strictement positive, qui vérifie $\gamma_x^x = 1$.

Unicité Soit λ invariante telle que $\lambda_x = 1$. On va montrer que $\lambda \geq \gamma^x$, puis que $\lambda = \gamma^x$. Notons que cette partie de la preuve du théorème n'utilise que l'irréductibilité.

$$\begin{aligned}
\lambda_y &= P_{xy} + \sum_{z_1 \neq x} \lambda_{z_1} P_{z_1 y} \\
&= P_{xy} + \sum_{z_1 \neq x} P_{x z_1} P_{z_1 y} + \sum_{z_1, z_2 \neq x} \lambda_{z_2} P_{z_2 z_1} P_{z_1 y} \\
&\geq \sum_{n=0}^{\infty} \sum_{z_1, \ldots, z_n \neq x} P_{x z_n} P_{z_n z_{n-1}} \times \cdots \times P_{z_1 y} \\
&= \sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = y, T_x \geq n+1) \\
&= \gamma_y^x.
\end{aligned}$$

Donc $\mu = \lambda - \gamma^x$ est aussi une mesure invariante, et $\mu_x = 0$. Soit $y \in E$, et n tel que $(P^n)_{yx} > 0$. Alors

$$0 = \mu_x = \sum_{z \in E} \mu_z (P^n)_{zx} \geq \mu_y (P^n)_{yx}.$$

Donc $\mu_y = 0$, et ceci $\forall y \in E$. □

On a vu qu'un état x est récurrent si

$$\mathbb{P}_x(T_x < \infty) = 1.$$

Posons $m_x = \mathbb{E}_x(T_x)$.

Si cette quantité est finie, l'état x est dit **récurrent positif**, et sinon **récurrent nul**.

Théorème 2.5.4. *On se place à nouveau dans le cas irréductible. Un état x est récurrent positif ssi tous les états sont récurrent positifs, ssi il existe une probabilité invariante, et dans ce cas elle est donnée par $\pi = (\pi_x = \frac{1}{m_x},\ x \in E)$.*

Preuve Notons que

$$m_x = \sum_{y \in E} \gamma_y^x$$

Donc si x est récurrent positif, la probabilité $\pi = (\pi_y = \dfrac{\gamma_y^x}{m_x},\ y \in E)$ est une probabilité invariante.

Réciproquement, si π est une probabilité invariante, par l'irréductibilité (cf. la fin de la preuve de l'existence dans le Théorème 2.5.3), π est strictement positive, donc si x est un état arbitraire, $\lambda = \left(\lambda_y = \dfrac{\pi_y}{\pi_x},\ y \in E\right)$ est une mesure invariante qui vérifie $\lambda_x = 1$. Par l'irréductibilité et la preuve de l'unicité dans le Théorème 2.5.3,

$$m_x = \sum_{y \in E} \gamma_y^x = \sum_{y \in E} \frac{\pi_y}{\pi_x} = \frac{1}{\pi_x} < \infty.$$

Donc l'état x, comme tous les états, est récurrent positif. □

La dichotomie suivante résulte des deux théorèmes précédents : dans le cas *récurrent irréductible*, la chaîne est *récurrente positive* s'il existe une *probabilité invariante*, *récurrente nulle* si *toute mesure invariante* est *de masse infinie* ($\sum_i \pi_i = +\infty$). En particulier, si $|E| < \infty$, il n'existe pas d'état récurrent nul, tout état récurrent est récurrent positif.

Corollaire 2.5.5. *Soit $\{X_n\}$ une chaîne de Markov irréductible, récurrente positive. Pour tout $x \in E$, on pose $T_x = \inf\{n > 0,\ X_n = x\}$. Alors pour tout $y \in E$,*

$$\mathbb{E}_y(T_x) < \infty.$$

PREUVE Remarquons que

$$T_x \geq T_x \mathbf{1}_{\{T_y < T_x\}},$$

d'où en prenant l'espérance sous $\mathbb{P}_x$,

$$m_x \geq \mathbb{E}_x(T_x | T_y < T_x) \mathbb{P}_x(T_y < T_x).$$

Mais il résulte de propriété de Markov forte que $\mathbb{E}_x(T_x|T_y < T_x) > \mathbb{E}_y(T_x)$, et de l'irréductibilité que $\mathbb{P}_x(T_y < T_x) > 0$. Le résultat est établi. □

Remarque 2.5.6. ***Cas non irréductible****. Limitons nous pour simplifier au cas $|E| < \infty$. Il existe au moins une classe récurrente (nécessairement récurrente positive), donc il existe au moins une probabilité invariante. Toute* *probabilité invariante ne charge que les états récurrents. S'il y a une seule classe récurrente, alors la chaîne possède une et une seule probabilité invariante. Sinon, à chaque classe récurrente on associe une unique probabilité invariante dont le support est cette classe, et toutes les mesures invariantes s'obtiennent comme combinaison linéaire convexe des précédentes. Donc dès qu'il existe au moins deux classes récurrentes, il y a une infinité non dénombrable de probabilités invariantes.*

Revenons au cas irréductible. On peut maintenant établir le théorème ergodique, qui constitue une généralisation de la loi des grands nombres.

Théorème 2.5.7. *Supposons la chaîne irréductible et récurrente positive. Désignons par* $\pi = (\pi_x,\, x \in E)$ *son unique probabilité invariante. Si* $f : E \to \mathbb{R}$ *est bornée, alors* $\mathbb{P}$ *p. s., quand* $n \to \infty$,

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \to \sum_{x\in E} \pi_x f(x).$$

Preuve Par hypothèse, il existe c tel que $|f(x)| \le c$, $\forall x \in E$.

Définissons
$$N_x(n) = \sum_{1\le k\le n} \mathbf{1}_{\{X_k = x\}},$$
le nombre de retours à l'état x avant l'instant n. On veut étudier les limites quand $n \to \infty$ de
$$\frac{N_x(n)}{n}.$$
Désignons par $S_x^0, S_x^1, \ldots, S_x^k, \ldots$ les longueurs des excursions $\mathcal{E}_0, \mathcal{E}_1, \ldots, \mathcal{E}_k, \ldots$ partant de x. On a

$$S_x^0 + \cdots + S_x^{N_x(n)-1} \le n < S_x^0 + \cdots + S_x^{N_x(n)}.$$

Donc
$$\frac{S_x^0 + \cdots + S_x^{N_x(n)-1}}{N_x(n)} \le \frac{n}{N_x(n)} \le \frac{S_x^0 + \cdots + S_x^{N_x(n)}}{N_x(n)}$$
Mais par le caractère i.i.d. des $\mathcal{E}_k$ (donc aussi des S_x^k),

$$\frac{S_x^0 + \cdots + S_x^{N_x(n)}}{N_x(n)} \to \mathbb{E}_x(T_x) = m_x \ \ \mathbb{P}_x \text{ p.s.}$$

Comme $N_x(n) \to +\infty \ \ \mathbb{P}_x$ p.s.,

$$\frac{n}{N_x(n)} \to m_x \ \ \mathbb{P}_x \text{ p.s.}$$

$$\frac{N_x(n)}{n} \to \frac{1}{m_x} \ \ \mathbb{P}_x \text{ p.s.}$$

Cette convergence a également lieu $\mathbb{P}_\mu$ p.s., $\forall \mu$ loi initiale, puisque les limites de $\frac{N_x(n)}{n}$ sont les mêmes pour la chaîne $\{X_n; n \geq 0\}$ et pour la chaîne $\{X_{T_x+n}; n \geq 0\}$.

Soit maintenant $F \subset E$. On note $\bar{f} = \sum_{x \in E} \pi_x f(x)$.

$$\begin{aligned}
\left| \frac{1}{n} \sum_{k=1}^{n} f(X_k) - \bar{f} \right| &= \left| \sum_{x \in E} \left(\frac{N_x(n)}{n} - \pi_x \right) f(x) \right| \\
&\leq c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + c \sum_{x \notin F} \left(\frac{N_x(n)}{n} + \pi_x \right) \\
&= c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + c \sum_{x \in F} \left(\pi_x - \frac{N_x(n)}{n} \right) + 2c \sum_{x \notin F} \pi_x \\
&\leq 2c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + 2c \sum_{x \notin F} \pi_x
\end{aligned}$$

On choisit F fini tel que $\sum_{x \notin F} \pi_x \leq \frac{\varepsilon}{4c}$, puis $N(\omega)$ tel que $\forall n \geq N(\omega)$,

$$\sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| \leq \frac{\varepsilon}{4c},$$

et le résultat est établi dans le cas f bornée. $\square$

On va maintenant établir un théorème de la limite centrale, en nous limitant - pour simplifier - au cas E fini.

Théorème 2.5.8. *Soit $\{X_n; n \in \mathbb{N}\}$ une chaîne de Markov à valeurs dans E fini, de matrice de transition P irréductible. Soit π l'unique probabilité invariante de la chaîne, et $f : E \to \mathbb{R}$ telle que*

$$\pi f = \sum_{x \in E} \pi_x f_x = 0.$$

Alors il existe $\sigma_f \geq 0$ tel que, quand $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_1^n f(X_k) \text{ converge en loi vers } \sigma_f Z,$$

où $Z \simeq N(0,1)$.

PREUVE

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_1^n f(X_k) &= \sum_{x \in E} \sqrt{n} f_x \left(\frac{N_x(n)}{n} - \pi_x \right) \\
&= \sum_{x \in E} \pi_x f_x \sqrt{\frac{N_x(n)}{n}} \left(\frac{\sqrt{N_x(n)}}{\pi_x} - \frac{n}{\sqrt{N_x(n)}} \right)
\end{aligned}$$

Par le même raisonnement que dans la preuve du théorème ergodique, cette dernière quantité se comporte quand $n \to \infty$ comme

$$\sum_{x \in E} \pi_x f_x \sqrt{\frac{N_x(n)}{n}} \times \frac{\bar{S}_x^0 + \cdots + \bar{S}_x^{N_x(n)}}{\sqrt{N_x(n)}},$$

avec $\bar{S}_x^k = S_x^k - \frac{1}{\pi_x}$.

Par le théorème de la limite centrale pour les v.a. i.i.d., le vecteur dont la x–ième coordonnée est

$$\frac{\bar{S}_x^0 + \cdots + \bar{S}_x^{N_x(n)}}{\sqrt{N_x(n)}}$$

converge en loi quand $n \to \infty$ vers un vecteur aléatoire gaussien centré. En outre, on a les convergences p.s.

$$\frac{N_x(n)}{n} \to \pi_x, \ \ x \in E.$$

On en déduit la convergence en loi de la somme ci–dessus vers une v.a. gaussienne centrée. $\square$

Le calcul de la variance σ_f^2 ne se déduit pas aisément de la preuve ci–dessus. On va donner - sans démonstration - une formule pour σ_f^2. Au moins sous les hypothèses du Théorème 2.6.7 ci–dessous, la quantité suivante est bien définie :

$$(Qf)_x = \sum_{n=0}^{\infty} \mathbb{E}_x[f(X_n)], \ \ x \in E.$$

Notons que

$$(I - P)Qf = f.$$

On a

$$\begin{aligned}\sigma_f^2 &= \sum_{x\in E} \pi_x (Qf)_x^2 - \sum_x \pi_x (PQf)_x^2 \\ &= 2\sum_x \pi_x (Qf)_x f_x - \sum_x \pi_x f_x^2.\end{aligned}$$

2.6 Le cas apériodique

On vient de montrer en particulier que dans le cas irréductible récurrent positif,

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k=y\}} \to \pi_y \ \text{ p.s. },$$

quand $n \to \infty$. En prenant l'espérance sous $\mathbb{P}_x$, on obtient que

$$\frac{1}{n}\sum_{k=1}^{n}(P^k)_{xy} \to \pi_y, \ \ \forall x, y \in E.$$

On voit que les moyennes de Césaro des $(P^k)_{xy}$ convergent. On peut se poser la question : est–ce que

$$(P^n)_{xy} \to \pi_y, \ \ \forall x, y \in E \quad ?$$

Il est facile de voir que ce n'est pas toujours le cas sous l'hypothèse irréductible et récurrent positif.

Considérons une marche aléatoire sur $E = \mathbf{Z}/N$, N entier pair (i.e. on identifie 0 et N)

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

avec les Y_n i.i.d. à valeurs dans $\{-1, 1\}$, autrement dit

$$X_n = (X_0 + Y_1 + \cdots + Y_n) \bmod N.$$

Cette chaîne est irréductible, et récurrente positive car E est fini. Mais $(P^{2k+1})_{xx} = 0$, pour tout $x \in E$. Dans le cas particulier $N = 2$, on a $P^{2k} = I$ et $P^{2k+1} = P$.

Pour espérer avoir la convergence souhaitée, il faut faire une hypothèse supplémentaire

Définition 2.6.1. *Un état $x \in E$ est dit* **apériodique** *si $\exists N$ tel que*

$$(P^n)_{xx} > 0, \qquad \text{pour tout } n \geq N.$$

Lemme 2.6.2. *Si P est irréductible et s'il existe un état apériodique x, alors $\forall y, z \in E$, $\exists M$ tel que $(P^n)_{yz} > 0$, $\forall n \geq M$. En particulier, tous les états sont apériodiques.*

Preuve D'après l'irréductibilité $\exists r, s \in N$ tels que $(P^r)_{yx} > 0$, $(P^s)_{xz} > 0$. En outre

$$\left(P^{r+n+s}\right)_{yz} \geq (P^r)_{yx}(P^n)_{xx}(P^s)_{xz} > 0,$$

dès que $n \geq N$. Donc on a la propriété désirée avec $M = N + r + s$.

Remarque 2.6.3. *Plaçons nous dans le cas irréductible, récurrent positif. Soit π la probabilité invariante. $\pi_y > 0$, $\forall y \in E$. Donc $(P^n)_{xy} > 0$ à partir d'un certain rang N est une CN pour que $(P^n)_{xy} \to \pi_y$. On va voir maintenant que c'est une CS.*

Théorème 2.6.4. *Supposons P irréductible, récurrent positif et apériodique. Soit π l'unique probabilité invariante. Alors si $\{X_n; n \in \mathbb{N}\}$ est une chaîne de Markov (μ, P), $\forall y \in E$,*

$$\mathbb{P}(X_n = y) \to \pi_y, \;\; n \to \infty,$$

soit

$$(\mu P^n)_y \to \pi_y,$$

pour toute loi initiale μ. En particulier, $\forall x, y \in E$,

$$(P^n)_{xy} \to \pi_y.$$

Preuve On va utiliser un argument de couplage. Soit $\{Y_n, n \in \mathbb{N}\}$ une chaîne de Markov (π, P), indépendante de $\{X_n; n \in \mathbb{N}\}$, et $x \in E$ arbitraire. On pose

$$T = \inf\{n \geq 0;\;\; X_n = Y_n = x\}$$

Etape 1 Montrons que $\mathbb{P}(T < \infty) = 1$.

$\{W_n = (X_n, Y_n); n \in \mathbb{N}\}$ est une chaîne de Markov à valeurs dans $E \times E$, de loi initiale λ (avec $\lambda_{(x,u)} = \mu_x \pi_u$), et de matrice de transition $\tilde{P}_{(x,u)(y,v)} = P_{xy} P_{uv}$. Comme P est apériodique, $\forall x, u, y, v$, pour n assez grand

$$(\tilde{P}^n)_{(x,u)(y,v)} = (P^n)_{xy} (P^n)_{uv} > 0.$$

Donc $\tilde{P}$ est irréductible. En outre, $\tilde{P}$ possède la probabilité invariante

$$\tilde{\pi}_{(x,u)} = \pi_x \pi_u.$$

Donc, d'après le Théorème 2.5.4, $\tilde{P}$ est récurrent positif. T est le premier temps de passage de la chaîne $\{W_n\}$ au point (x, x) ; il est donc fini p.s.

Etape 2 On pose

$$Z_n = \begin{cases} X_n, & n \leq T; \\ Y_n, & n > T. \end{cases}$$

On va maintenant montrer que $\{Z_n; n \in \mathbb{N}\}$ est une chaîne de Markov (μ, P).

Notons que par la propriété de Markov forte, les deux processus $\{X_{T+n}; n \geq 0\}$ et $\{Y_{T+n}); n \geq 0\}$ sont des chaînes de Markov (δ_x, P), indépendantes de $(X_0, \ldots, X_T)$. Par conséquent, $\{Z_n, n \in \mathbb{N}\}$ est, comme $\{X_n\}$, une chaîne de Markov (μ, P).

Etape 3 On va conclure. On a trois identités

$$\begin{aligned}
\mathbb{P}(Z_n = y) &= \mathbb{P}(X_n = y) \\
\mathbb{P}(Y_n = y) &= \pi_y \\
\mathbb{P}(Z_n = y) &= \mathbb{P}(X_n = y, n \leq T) + \mathbb{P}(Y_n = y, n > T)
\end{aligned}$$

Donc

$$\begin{aligned}
|\mathbb{P}(X_n = y) - \pi_y| &= |\mathbb{P}(Z_n = y) - \mathbb{P}(Y_n = y)| \\
&\leq \mathbb{P}(n < T) \\
&\to 0,
\end{aligned}$$

quand $n \to \infty$. □

On va maintenant préciser la vitesse de convergence dans le théorème précédent, sous une hypothèse supplémentaire, dite condition de Doeblin :
$\exists\ n_0 \in \mathbb{N}$, $\beta > 0$ et une probabilité ν sur E tels que

$$(D) \quad (P^{n_0})_{xy} \geq \beta\nu_y, \ \forall x, y \in E.$$

Remarque 2.6.5. *La condition (D) est équivalente à la condition*

$$\exists x \in E, \ n_0 \geq 1 \quad \textit{tels que} \ \inf_{y \in E} (P^{n_0})_{yx} > 0.$$

Elle entraîne que cet état x est apériodique. Mais elle n'entraîne pas l'irréductibilité (il est facile de construire un contre–exemple). On verra à l'exercice 2.9.5 que cette condition entraîne l'existence d'une unique classe récurrente, donc d'une unique probabilité invariante.

Lemme 2.6.6. *Si P est irréductible et apériodique, et E fini, alors la condition (D) est satisfaite.*

PREUVE Choisissons $x \in E$. $\forall y \in E$, $\exists n_y$ tel que $n \geq n_y \Rightarrow (P^n)_{yx} > 0$. Posons $\bar{n} = \sup_{y \in E} n_y$, $\alpha = \inf_y (P^{\bar{n}})_{yx}$. Alors $\alpha > 0$, et $\forall y \in E$,

$$(P^{\bar{n}})_{yx} \geq \alpha.$$

Donc la condition (D) est satisfaite avec $n_0 = \bar{n}$, $\beta = \alpha$, $\nu = \delta_x$. □

Par ailleurs, la condition de Doeblin est très rarement satisfaite dans le cas card$E = +\infty$, car alors typiquement $\forall n \in \mathbb{N}$, $y \in E$,

$$\inf_{x \in E} (P^n)_{xy} = 0.$$

Théorème 2.6.7. *Supposons que P est irréductible et satisfait la condition (D). Alors P est apériodique, récurrente positive, et si π désigne sa probabilité invariante,*

$$\sum_{y \in E} |(P^n)_{xj} - \pi_y| \leq 2\,(1-\beta)^{[n/n_0]}, \ \forall x \in E, \ n \in \mathbb{N},$$

où $[n/n_0]$ désigne la partie entière de n/n_0.

Nous allons tout d'abord introduire un outil qui nous sera utile dans la preuve du théorème.

Définition 2.6.8. *On appelle couplage de deux probabilités p et q sur E tout couple aléatoire (X, Y) de deux v.a. à valeurs dans E, telles que p soit la loi de X et q celle de Y.*

Lemme 2.6.9. *Soit p et q deux probabilités sur E. On a l'égalité*

$$||p-q||_1 = 2 \inf_{(X,Y) \text{ couplage de } p,q} \mathbb{P}(X \neq Y).$$

PREUVE : Tout d'abord, si (X, Y) est un couplage de p et q,

$$\begin{aligned}\mathbb{P}(X = Y) &= \sum_{x\in E} \mathbb{P}(X = Y = x) \\ &\leq \sum_{x\in E} p_x \wedge q_x,\end{aligned}$$

d'où

$$\begin{aligned}\mathbb{P}(X \neq Y) &\geq 1 - \sum_{x\in E} p_x \wedge q_x \\ &= \sum_{x\in E} (p_x - q_x)^+,\end{aligned}$$

et

$$\begin{aligned}\|p-q\|_1 &= \sum_{x\in E} |p_x - q_x| \\ &\leq 2\mathbb{P}(X \neq Y).\end{aligned}$$

D'un autre coté, posons $\alpha = \sum_{x\in E} p_x \wedge q_x$. Si ξ, U, V et W sont des v. a. indépendantes, avec $\mathbb{P}(\xi = 1) = 1 - \mathbb{P}(\xi = 0) = \alpha$, U de loi r t.q. $r_x = \alpha^{-1} p_x \wedge q_x$, V de loi $\bar{p}$ t.q. $\bar{p}_x = (1-\alpha)^{-1}(p_x - q_x)^+$, W de loi $\bar{q}$ t.q. $\bar{q}_x = (1-\alpha)^{-1}(q_x - p_x)^+$, alors

$$\begin{aligned}X &= \xi U + (1-\xi)V, \\ Y &= \xi U + (1-\xi)W\end{aligned}$$

réalise un couplage (X, Y) de p et q tel que $2\mathbb{P}(X \neq Y) = \|p-q\|_1$. □

PREUVE DU THÉORÈME 2.6.7 : La chaîne étant irréductible, la condition (D) entraîne clairement son caractère apériodique.

Etape 1 On va tout d'abord montrer que pour toutes probabilités μ et ν sur E,

$$\|\mu P^n - \nu P^n\|_1 \leq 2(1-\beta)^{[n/n_0]}. \tag{2.1}$$

Pour cela, d'après le Lemme 2.6.9, il suffit de construire un couplage (X_n, Y_n) des probabilités μP^n et νP^n, tel que

$$\mathbb{P}(X_n \neq Y_n) \leq (1-\beta)^{[n/n_0]}.$$

Supposons que $n = kn_0 + m$, avec $m < n_0$. Etant donné (X_0, Y_0) de loi $\mu \times \nu$ sur $E \times E$, pour $\ell = 0, 1, \dots, k-1$, on définit $(X_{(\ell+1)n_0}, Y_{(\ell+1)n_0})$ en fonction de $(X_{\ell n_0}, Y_{\ell n_0})$ comme suit. On se donne une suite $\{\xi_\ell, U_\ell, V_\ell,\ \ell \geq 0\}$ de v.a.indépendantes, les ξ_ℓ de Bernoulli t.q. $\mathbb{P}(\xi_\ell = 1) = \beta = 1 - \mathbb{P}(\xi_\ell = 0)$, les U_ℓ de loi $\bar{m} = \beta^{-1} m$ et les V_ℓ de loi uniforme sur $[0,1]$. On note enfin

$$Q_{xy} = (1-\beta)^{-1}((P^{n_0})_{xy} - m_y),$$

et $f : E \times [0,1] \to E$ telle que si V suit la loi uniforme sur $[0,1]$, $f(x, V)$ suit la loi $Q_{x\cdot}$, $x \in E$. On pose alors

$$\begin{aligned} X_{(\ell+1)n_0} &= \xi_\ell U_\ell + (1-\xi_\ell) f(X_{\ell n_0}, V_\ell) \\ Y_{(\ell+1)n_0} &= \xi_\ell U_\ell + (1-\xi_\ell) f(Y_{\ell n_0}, V_\ell). \end{aligned}$$

Notons que l'on a bien construit un couplage $(X_{\ell n_0}, Y_{\ell n_0})$ de $\mu P^{\ell n_0}$ et $\nu P^{\ell n_0}$, pour $\ell = 0, \dots, k$, tel que

$$\mathbb{P}(X_{\ell n_0} \neq Y_{\ell n_0}) \leq \mathbb{P}(\cap_{m=0}^{\ell} \xi_m = 0) = (1-\beta)^\ell.$$

Il reste à construire un couplage (X_n, Y_n) de μP^n et νP^n, tel que $\{X_n \neq Y_n\} \subset \{X_{kn_0} \neq Y_{kn_0}\}$, ce qui est facile.

Etape 2 Montrons maintenant que pour toute probabilité μ, $\{\mu P^n,\ n \geq 0\}$ est une suite de Cauchy dans l'espace de Banach $\ell^1(E)$. Si $\nu = \mu P^m$, d'après (2.1),

$$\|\mu P^{n+m} - \mu P^n\|_1 = \|\nu P^n - \mu P^n\|_1 \leq 2c^{n-n_0},$$

où $c = (1-\beta)^{1/n_0}$, d'où le résultat.

Etape 3 Il résulte de la 2ème étape que la suite de probabilités $\{\mu P^n,\ n \geq 0\}$ converge dans $\ell^1(E)$, vers une probabilité π sur E. Mais

$$\pi P = \lim_{n\to\infty} \mu P^{n+1} = \pi,$$

donc π est invariante, et la chaîne est récurrente positive. En conséquence, d'après (2.1), pour toute probabilité μ sur E,

$$\|\mu P^n - \pi\|_1 \leq 2(1-\beta)^{[n/n_0]},$$

ce qui établit la vitesse de convergence annoncée, et l'apériodicité.

2.7 Chaîne de Markov réversible

On se place dans le cas irréductible, récurrent positif. La formulation de la propriété de Markov "passé et futur sont conditionnellement indépendants sachant le présent" nous indique que si $\{X_n; n \in \mathbb{N}\}$ est une chaîne de Markov, alors $\forall N$, $\{\hat{X}_n^N = X_{N-n}; 0 \leq n \leq N\}$ est aussi une chaîne de Markov. En général, la chaîne retournée n'est pas homogène, sauf si $\{X_n\}$ est initialisée avec sa probabilité invariante π.

Proposition 2.7.1. *Soit $\{X_n; n \in \mathbb{N}\}$ une chaîne de Markov (π, P), avec π probabilité invariante et P irréductible. Alors la chaîne retournée $\{\hat{X}_n^N; 0 \leq n \leq N\}$ est une chaîne de Markov $(\pi, \hat{P})$, avec*

$$\pi_y \hat{P}_{yx} = \pi_x P_{xy}, \ \forall\, x, y \in E$$

Preuve

$$\begin{aligned}
&\mathbb{P}(\hat{X}_{p+1} = x | \hat{X}_p = y) \\
&= \mathbb{P}(X_n = x | X_{n+1} = y) \\
&= \mathbb{P}(X_{n+1} = y | X_n = x) \times \frac{\mathbb{P}(X_n = x)}{\mathbb{P}(X_{n+1} = y)}.
\end{aligned}$$

□

On dit que la chaîne $\{X_n; n \in \mathbb{N}\}$ est **réversible** si $\hat{P} = P$, ce qui a lieu si et seulement si la relation suivante dite d'"équilibre ponctuel"est satisfaite :

$$\pi_x P_{xy} = \pi_y P_{yx}, \ \forall\, x, y \in E,$$

avec π probabilité invariante. Il est facile de vérifier que si une probabilité π satisfait cette relation, alors elle est invariante par P. La réciproque n'a aucune raison d'être vraie.

Remarque 2.7.2. *Si π est la probabilité invariante d'une chaîne irréductible (et donc aussi récurrente positive), la chaîne n'est pas nécessairement réversible par rapport à π. Supposons que $cardE \geq 3$. Alors il peut exister $x \neq y$ t.q. $P_{xy} = 0 \neq P_{yx}$. D'où $\pi_x P_{xy} = 0 \neq \pi_y P_{yx}$. Les transitions de y à x de la chaîne initiale correspondent aux transitions de x à y de la chaîne retournée, donc $P_{yx} \neq 0 \Rightarrow \hat{P}_{xy} \neq 0$, d'où $\hat{P} \neq P$.*

Remarque 2.7.3. *Étant donné une matrice de transition d'une chaîne de Markov récurrente irréductible P, un problème classique est de calculer sa probabilité invariante.*

Un autre problème, qui apparaîtra au chapitre suivant, est de déterminer une matrice de transition P irréductible, dont la chaîne associée admette comme probabilité invariante une mesure π donnée.

Le second problème est facile à résoudre. En fait il existe un grand nombre de solutions. Le plus simple est de chercher P telle que la chaîne correspondante soit réversible par rapport à la mesure donnée π. Autrement dit, il suffit de trouver une matrice markovienne irréductible P telle que la quantité $\pi_x P_{xy}$ soit symétrique en x, y.

Pour résoudre le premier problème, on peut là aussi chercher à résoudre l'équation

$$\pi_x P_{xy} = \pi_y P_{yx}, \quad \forall x,\, y \in E,$$

qui contrairement à l'équation $\pi P = \pi$, n'impose pas de sommation en i. Mais cette équation n'a de solution que si la chaîne est réversible par rapport à son unique probabilité invariante, ce qui n'est pas forcément le cas.

Supposons maintenant que l'on se donne le couple (P, π), et que l'on veuille vérifier si oui ou non π est la probabilité invariante associée à la chaîne de matrice de transition irréductible P. Si la quantité $\pi_x P_{xy}$ est symétrique en x, y, la réponse est oui, et on a même mieux, la réversibilité. Si non, il faut vérifier si oui ou non $\pi P = \pi$. Une façon de faire ce calcul est donnée par la Proposition suivante.

Proposition 2.7.4. *Soit P une matrice de transition irréductible, et π une probabilité sur E. On pose pour tout x, $y \in E$,*

$$\hat{P}_{xy} = \begin{cases} \frac{\pi_y}{\pi_x} P_{yx}, & si\ x \neq y, \\ P_{xx}, & si\ x = y. \end{cases}$$

Alors π est la probabilité invariante de la chaîne de matrice de transition P

et $\hat{P}$ est la matrice de la chaîne retournée ssi pour tout $x \in E$,

$$\sum_{y \in E} \hat{P}_{xy} = 1.$$

2.8 Statistique des chaînes de Markov

Le but de cette section est d'introduire des notions élémentaires sur la statistique des chaînes de Markov.

On a vu que pour tout $n > 0$, la loi du vecteur aléatoire $(X_0, X_1, \dots, X_n)$ ne dépend que de la loi initiale μ et de la matrice de transition P.

La première question que l'on peut se poser en statistique des chaînes de Markov est : peut–on estimer le couple (μ, P) au vu de l'observation de la suite $(X_0, X_1, \dots, X_n)$, d'une telle façon que l'erreur d'estimation soit "petite" quand n est "grand". Comme on va le voir ci–dessous, on peut estimer P. Par contre, on ne peut raisonnablement estimer la loi initiale μ que si celle–ci coïncide avec la probabilité invariante, et que l'*on est dans le cas irréductible et récurrent positif, ce que nous supposerons dans toute la suite de cette section*.

Commençons par l'estimation de la probabilité invariante μ.

Pour tout $x \in E$,

$$\hat{\mu}_x^n = \frac{1}{n+1} \sum_{\ell=0}^{n} \mathbf{1}_{\{X_\ell = x\}}$$

est un estimateur consistant de μ_x, puisqu'une conséquence immédiate du théorème ergodique est que

Proposition 2.8.1. *Pour tout $x \in E$, $\hat{\mu}_x^n \to \mu_x$ p.s., quand $n \to \infty$.*

Passons maintenant à l'estimation des P_{xy}, $x, y \in E$. On choisit l'estimateur

$$\hat{P}_{xy}^n = \frac{\displaystyle\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x, X_{\ell+1} = y\}}}{\displaystyle\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x\}}}$$

On a la

Proposition 2.8.2. *Pour tout $x, y \in E$, $\hat{P}^n_{xy} \to P_{xy}$ p.s. quand $n \to \infty$.*

Preuve On a bien sûr

$$\hat{P}^n_{xy} = \left(\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x\}}\right)^{-1} \frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x, X_{\ell+1}=y\}}.$$

On sait que

$$\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x\}} \to \mu_x.$$

Pour $n \geq 0$, posons $\widetilde{X}_n = (X_n, X_{n+1})$. Il n'est pas très difficile de vérifier que $\{\widetilde{X}_n,\ n \geq 0\}$ est une chaîne de Markov irréductible et récurrente positive à valeurs dans $\widetilde{E} = \{(x,y) \in E \times E, P_{xy} > 0\}$, de matrice de transition $\widetilde{P}_{(x,y)(u,v)} = \delta_{yu} P_{uv}$, et de probabilité invariante $\widetilde{\mu}_{(x,y)} = \mu_x P_{xy}$. Le théorème ergodique appliqué à la chaîne $\{\widetilde{X}_n\}$ entraîne que

$$\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x, X_{\ell+1}=y\}} \to \mu_x P_{xy}$$

2.9 Exercices

Exercice 2.9.1. *Montrer que la chaîne (X_n, $n \in \mathbb{N}$) à trois états 0, 1, 2, de matrice de transition*

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ p & 1-p-q & q \\ 0 & 0 & 1 \end{pmatrix} \ (p, q > 0, p+q \leq 1)$$

et d'état initial 1 ($P(X_0 = 1) = 1$) change d'état pour la première fois à un instant aléatoire $T \geq 1$ de loi géométrique. Montrer ensuite que X_T est une v.a. indépendante de T, de loi $(p/(p+q), 0, q/(p+q))$, et enfin que $X_t = X_T$ si $t \geq T$.

Exercice 2.9.2. *Soit $\{X_n; n \in \mathbb{N}\}$ une chaîne de Markov de matrice de transition*

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

Déterminer les classes d'équivalence, les états transitoires et récurrents, et les mesures invariantes.

Exercice 2.9.3. *On considère une chaîne de Markov* $\{X_n;\, n \in \mathbb{N}\}$ *à valeurs dans l'espace fini* $E = \{1, 2, 3, 4, 5, 6\}$*, de matrice de transition* P*, dont les termes hors–diagonaux sont donnés par*

$$P = \begin{pmatrix} \cdot & 2/3 & 1/3 & 0 & 0 & 0 \\ 1/4 & \cdot & 0 & 0 & 1/5 & 2/5 \\ 0 & 0 & \cdot & 1/2 & 0 & 0 \\ 0 & 0 & 2/3 & \cdot & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdot & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & \cdot \end{pmatrix}$$

1 *Déterminer les termes diagonaux de la matrice de transition* P.

2 *Montrer que* E *est constitué de trois classes d'équivalence que l'on précisera, l'une* $\mathcal{T}$ *étant transitoire, les deux autres* $\mathcal{R}_1$, $\mathcal{R}_2$ *récurrentes.*

3 *Déterminer une probabilité invariante admettant* $\mathcal{R}_1$ *comme support, et une probabilité invariante admettant* $\mathcal{R}_2$ *comme support. En déduire toutes les probabilités invariantes.*

Exercice 2.9.4. *Soit* P *une matrice markovienne sur un ensemble fini* E *de cardinal* d.

a *Montrer que la probabilité* π *est invariante ssi*

$$\pi(I - P + A) = a,$$

où A *est la matrice* $d \times d$ *dont tous les éléments sont égaux à 1, et* a *le vecteur de* $\mathbb{R}^d$ *dont tous les éléments sont égaux à 1.*

b *Montrer que* P *est irréductible ssi* $I - P + A$ *est inversible.*

c *Déduire des questions précédentes une méthode de calcul de* π.

Exercice 2.9.5. *Soit* P *une matrice markovienne sur un nensemble fini ou dénombrable* E*, qui satisfait la condition* (D) *de la section 2.6.*

1. *Onsuppose d'abord la condition* (D) *satisfaite avec* $n_0 = 1$*. Montrer qu'il existe un état récurrent, qui est p. s. visité par la chaîne une infinité de fois, quelque soit le point de départ. En déduire que la chaîne possède une unique classe récurrente. (Indication : on pourra montrer*

qu'il existe $x \in E$, $\beta > 0$ tels que la chaîne puisse être simulée comme suit : à chaque instant n, on va à l'état x avec la probabilité β, et on fait un autre mouvement (qui lui dépend de l'état où est la chaîne) avec probabilité $1-\beta$).

2. *Montrer que le résultat est encore vrai dans le cas général de l'hypothèse (D). (Indication : considérer la chaîne échantillonnée $\{X_{kn_0},\ k = 0, 1, \ldots\}$).*

Exercice 2.9.6. *Montrer que si x est récurrent, alors $\sum_{n\geq 0}(P^n)_{xy} = +\infty$ ssi $x \leftrightarrow y$, $= 0$ ssi $x \not\rightarrow y$.*

Exercice 2.9.7. Marche aléatoire sur Z *On pose*

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

où les X_n prennent leurs valeurs dans $\mathbf{Z}$, les Y_n dans $\{-1, 1\}$, $X_0, Y_1, .., Y_n, ..$ étant une suite indépendante, et pour tout n,

$$\mathbb{P}(Y_n = 1) = p = 1 - \mathbb{P}(Y_n = -1), \quad 0 < p < 1.$$

a *Montrer que la chaîne $\{X_n\}$ est irréductible.*

b *Montrer que dans le cas $p \neq 1/2$, la chaîne est transitoire (utiliser la loi des grands nombres).*

c *On considère le cas $p = 1/2$. Montrer que la chaîne est récurrente (on évaluera $\sum_{n\geq 1}(P^n)_{00}$ en utilisant la formule de Stirling $n! \simeq \sqrt{2\pi n}(\frac{n}{e})^n$). Montrer que la chaîne est récurrente nulle (on cherchera une mesure invariante). Préciser les valeurs de*

$$\limsup_{n\to\infty} X_n \text{ et } \liminf_{n\to\infty} X_n.$$

Exercice 2.9.8. Marche aléatoire sur $\mathbf{Z}^d$ *On pose*

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

où les X_n prennent leurs valeurs dans $\mathbf{Z}^d$, les Y_n étant i.i.d., globalement indépendants de X_0, de loi donnée par

$$\mathbb{P}(Y_n = \pm e_i) = (2d)^{-1}, \quad 1 \leq i \leq d,$$

$\{e_1, \ldots, e_d\}$ désignant la base canonique de $\mathbf{Z}^d$.

Montrer que $\{X_n\}$ est une chaîne de Markov irréductible à valeurs dans $\mathbf{Z}^d$, récurente nulle si $d = 1, 2$, transitoire si $d \geq 3$.

Exercice 2.9.9. *On reprend la marche aléatoire à valeurs dans* $\mathbf{Z}$ *dans le cas symmétrique (cas* $p = 1/2$*), et on va établir la récurrence par un argument totalement différent de celui de l'exercice 2.9.7. On suppose pour simplifier que* $X_0 = x \in \mathbf{Z}$*.*

Pour tous $a, b \in \mathbf{Z}$*, avec* $a < x < b$*, on pose*

$$T_{a,b} = \inf\{n \geq 0;\ X_n \notin]a, b[\},$$
$$T_a = \inf\{n \geq 0;\ X_n = a\},$$
$$T_b = \inf\{n \geq 0;\ X_n = b\},$$

et on remarque que

$$X_{n\wedge T_{a,b}} = x + \sum_{k=1}^{n} Y_k \mathbf{1}_{\{T_{a,b}>k-1\}}.$$

a *Montrer que les v.a.* Y_k *et* $\mathbf{1}_{\{T_{a,b}>k-1\}}$ *sont indépendantes. En déduire que*

$$\mathbb{E}X_{n\wedge T_{a,b}} = x.$$

b *Montrer que* $|X_{n\wedge T_{a,b}}| \leq \sup(|a|, |b|)$*,* $T_{a,b} < \infty$ *p.s., et*

$$\mathbb{E}X_{T_{a,b}} = x.$$

c *Etablir les identités*

$$\mathbb{P}(X_{T_{a,b}} = a) = \frac{b-x}{b-a}, \quad \mathbb{P}(X_{T_{a,b}} = b) = \frac{x-a}{b-a}.$$

c *Montrer que* $\mathbb{P}(T_a < T_n) \to 1$*, quand* $n \to \infty$*.*

d *Montrer que* $T_a < \infty$ *p.s., et que de même* $T_b < \infty$ *p.s. En déduire que la chaîne est récurrente.*

Exercice 2.9.10. Marche aléatoire réfléchie en 0 *Les* $\{Y_n\}$ *étant définis comme à l'exercice 2.9.7, on définit la chaîne de Markov* $\{X_n\}$ *à valeurs dans* $\mathbb{N}$ *par la formule de récurrence*

$$X_{n+1} = \mathbf{1}_{\{X_n>0\}} Y_{n+1} + \mathbf{1}_{\{X_n=0\}}.$$

On suppose bien sûr que $X_0 \in \mathbb{N}$*. On notera ci–dessous* $\{X'_n\}$ *la marche aléatoire de l'exercice 2.9.7, avec le même* X_0 *et les mêmes* $\{Y_n\}$ *que ceux utilisés dans la définition de la chaîne* $\{X_n\}$*. On utilisera dans cet exercice les résultats de l'exercice 2.9.7.*

– **a** *Montrer que la chaîne ainsi définie est irréductible à valeurs dans* $\mathbb{N}$. *Préciser sa matrice de transition.*
– **b** *Montrer que p.s.,* $X_n \geq X'_n$, $\forall n$. *En déduire que* $\{X_n\}$ *est transitoire dans le cas* $p > 1/2$.
– **c** *On pose* $T = \inf\{n \geq 0,\ X_n = 0\}$. *Montrer que* $X_n = X'_n$ *si* $T \geq n$. *En déduire que la chaîne est récurrente dans le cas* $p \leq 1/2$ *(on pourra par exemple montrer que l'état 1 est récurrent).*
– **d** *Montrer que la chaîne est récurrente nulle dans le cas* $p = 1/2$, *récurrente positive dans le cas* $p < 1/2$ *(on montrera que dans le premier (resp. le second) cas, la mesure* $(1/2, 1, 1, 1, \ldots)$ *(resp. la mesure* μ *définie par*

$$\mu_0 = \frac{1-2p}{2(1-p)}, \quad \mu_x = \frac{1-2p}{2}\frac{p^{x-1}}{(1-p)^{x+1}},\ x \geq 1$$

est une mesure invariante).

Exercice 2.9.11. Chaîne de naissance et de mort *Soit* $\{X_n\}$ *une chaîne de Markov à valeurs dans* $E = \mathbb{N}$, *de matrice de transition* P *définie par*

$$P_{x,x-1} = q_x, \quad P_{x,x} = r_x, \quad P_{x,x+1} = p_x,$$

avec pour tout $x \in \mathbb{N}$, $p_x + r_x + q_x = 1$, $q_0 = 0$, $q_x > 0$ *si* $x > 0$, *et* $p_x > 0$ *pour tout* $x \in \mathbb{N}$.

Pour $x \in \mathbb{N}$, *on pose* $\tau_x = \inf\{n \geq 0,\ X_n = x\}$. *Etant donnés trois états* a, x *et* b *tels que* $a \leq x \leq b$, *on pose* $u(x) = \mathbb{P}_x(\tau_a < \tau_b)$. *On définit* $\{\gamma_x,\ x \in \mathbb{N}\}$ *par* $\gamma_0 = 1$ *et pour* $x > 0$, $\gamma_x = q_1 \times \cdots \times q_x / p_1 \times \cdots \times p_x$.

1. *Montrer que la chaîne est irréductible à valeurs dans* $\mathbb{N}$.
2. *Pour* $a < x < b$, *établir une relation entre* $u(x) - u(x+1)$ *et* $u(x-1) - u(x)$. *Calculer* $u(a) - u(b)$ *en fonction des* γ_x, *et en déduire que pour* $a < x < b$,

$$u(x) = \sum_{y=x}^{y=b-1} \gamma_y / \sum_{y=a}^{y=b-1} \gamma_y.$$

   Traiter le cas particulier où $p_x = q_x$ *pour tout* $x > 0$.
3. *Calculer* $\mathbb{P}_1(\tau_0 = \infty)$ *et montrer que la chaîne est récurrente ssi* $\sum_0^\infty \gamma_y = +\infty$.

4. *Déterminer les mesures sur E invariantes par P, et en déduire que la chaîne est récurrente positive ssi*

$$\sum_{x=1}^{\infty} \frac{p_0 p_1 \times \cdots \times p_{x-1}}{q_1 q_2 \times \cdots \times q_x} < \infty.$$

5. *Montrer que dans le cas récurrent positif la chaîne est réversible. (Indication : on remarque tout d'abord que pour $x > 0$ la relation $\pi_x = (\pi P)_x$ s'écrit*

$$\pi_x P_{x,x-1} + \pi_x P_{x,x+1} = \pi_{x-1} P_{x-1,x} + \pi_{x+1} P_{x+1,x}.$$

On considère ensuite le cas $x = 0$, puis on montre par récurrence que

$$\pi_x P_{x,x+1} = \pi_{x+1} P_{x+1,x}, \quad \forall x \geq 0).$$

Exercice 2.9.12. File d'attente *On considère une file d'attente en temps discret qui se forme à un guichet, suivant le phénomène suivant : à chaque instant $n \in \mathbb{N}$, il arrive un client avec la probabilté p, $(0 < p < 1)$ et pas de client avec la probabilité $1 - p$. Lorsqu'il y a au moins un client en attente, à chaque instant un client est servi et quitte le système avec la probabilité q, $0 < q < 1$, et personne ne quitte le système avec la probabilité $1 - q$ (un client qui arrive à l'instant n repart au plus tôt à l'instant $n + 1$). Tous les tirages ci–dessus sont indépendants entre eux. On note X_n le nombre de clients présents dans la file à l'instant n.*

1. *Montrer que $\{X_n,\, n \in \mathbb{N}\}$ est une chaîne de Markov irréductible à valeurs dans $\mathbb{N}$. Préciser sa matrice de transition P_{xy}, $x, y \in \mathbb{N}$.*
2. *Donner une CNS sur p et q pour que la chaîne $\{X_n\}$ possède une probabilité invariante. On supposera dans la suite que cette condition est satisfaite et on notera $\{\pi_x,\, x \in \mathbb{N}\}$ l'unique probabilité invariante que l'on déterminera.*
3. *Calculer $\mathbb{E}_\pi(X_n)$.*
4. *On précise maintenant que les clients sont servis dans l'ordre de leur arrivée. On désigne par T le temps de séjour d'un client quelconque. Le système étant initialisé avec sa probabilité invariante, quelle est l'espérance de T ?*

Exercice 2.9.13. File d'attente *On considère une file d'attente qui se forme à un guichet. X_n désigne le nombre de clients dans la file en attente ou* *en train de se faire servir à l'instant* n. *Entre les instants* n *et* $n+1$ *arrivent* Y_{n+1} *clients, et si* $X_n > 0$ *partent* Z_{n+1} *clients. On suppose que* X_0, Y_1, Z_1, Y_2, $Z_2 \ldots$ *sont indépendantes, avec les* Y_n *i.i.d. vérifiant* $0 < \mathbb{P}(Y_n = 0) < 1$, *et les* Z_n *vérifiant* $\mathbb{P}(Z_n = 1) = p = 1 - \mathbb{P}(Z_n = 0)$.

1. *Montrer que* $(X_n; n \in \mathbb{N})$ *est une chaîne de Markov dont on précisera la probabilité de transition.*
2. *On note* φ *la fonction génératrice des* Y_n, ρ *celle des* Z_n, Ψ_n *celle des* X_n. *Calculer* Ψ_{n+1} *en fonction de* Ψ_n, φ *et* ρ.
3. *Montrer qu'il existe une unique probabilité invariante dont on calculera la fonction génératrice ssi* $\mathbb{E}(Y_1) < p$.

Exercice 2.9.14. Aloha discret *Le but de cet exercice est d'étudier le protocole de communication suivant : des usagers se présentent aux instants* $\{1, 2, \ldots, n, \ldots\}$ *pour transmettre un message via un canal, qui ne peut transmettre qu'un message à la fois. Lorsque deux ou plus usagers se présentent en même temps, aucun message ne passe, chaque usager en est averti, et il tente de se représenter plus tard. On cherche une politique de retransmission "distribuée", i.e. telle que chaque usager puisse décider quand se représenter, sans connaître les intentions des autres usagers. Le protocole "Aloha discret" stipule que chaque usager dont le message a été bloqué tente sa chance à l'instant suivant avec la probabilité* p. *Si son tirage aléatoire lui indique de ne pas se présenter, il tente sa chance à l'instant suivant avec la probabilité* p, *et ainsi de suite jusqu'à ce qu'un tirage lui indique de tenter sa chance. On appelle* Y_n *le nombre de messages "frais" (i.e. qui se présentent pour la première fois) arrivant au canal de transmission à l'instant* n. *On suppose que la suite* $\{Y_n\}$ *est i.i.d., avec* $\mathbb{P}(Y_n = i) = a_i$, $i \in \mathbb{N}$, *et* $\mathbb{E}(Y_n) > 0$. *Soit* X_n *le nombre de messages retardés en attente d'être transmis à l'instant* n.

1. *Montrer que* $\{X_n\}$ *est une chaîne de Markov dont on précisera la matrice de transition.*
2. *Montrer que* $\{X_n\}$ *est irréductible, mais n'est pas récurrente positive.*

Exercice 2.9.15. Programmation *On reprend la file d'attente de l'exercice 2.9.12.*

1. *Simuler et tracer une trajectoire de* $\{X_n,\ n \geq 0\}$ *de* $n = 1$ *à* $n = 1000$, *pour* $p = 1/2$ *et successivement* $q = 3/5, 7/13, 15/29, 1/2$.
2. *Puisque* $\{X_n\}$ *est irréductible, récurrente positive et apériodique,* $(P^n)_{yx} \to \pi_x$. *Tracer au choix l'histogramme (ou la fonction de répartition) empirique de* $(P^n)_{y\cdot}$, *pour* $n = 100, 500, 1000$, *pour une taille*

d'échantillon de 10^4*. On représentera l'histogramme (resp. la fonction de répartition) de* π *sur le même graphique. On traitera les cas* $p = 1/2$, $q = 3/5, 7/13$.

3. *Comparer graphiquement les quantités*

$$n^{-1}\sum_{k=1}^{n}\mathbf{1}_{\{X_k=x\}}, \quad x \in \mathbb{N}$$

et l'histogramme de π*, pour* $n = 10^3, 10^4, 10^5$*. On traitera les cas* $p = 1/2$*,* $q = 3/5, 7/13$*. Pour chaque valeur de* q*, on pourra choisir l'intervalle utile des valeurs de* x *au vu des tracés de la question précédente.*

2.10 Problèmes corrigés

2.10.1 Enoncés

Problème 2.10.1. *On considère une chaîne de Markov en temps discret* $\{X_n;\, n \in \mathbb{N}\}$ *à valeurs dans l'espace fini* $E = \{1, 2, 3, 4, 5, 6\}$*, de matrice de transition* P*, dont les termes hors–diagonaux sont donnés par*

$$P = \begin{pmatrix} \cdot & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & \cdot & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 7/8 & 0 \\ 1/4 & 1/4 & 0 & \cdot & 1/4 & 1/4 \\ 0 & 0 & 3/4 & 0 & \cdot & 0 \\ 0 & 1/5 & 0 & 1/5 & 1/5 & \cdot \end{pmatrix}$$

1. *Déterminer les termes diagonaux de la matrice de transition* P.
2. *Déterminer les classes d'équivalence de la chaîne.*
3. *Montrer que les états 4 et 6 sont transitioires, et que l'ensemble des autres états se décompose en deux classes récurrentes que l'on précisera. Dans la suite, on notera* $\mathcal{T} = \{4, 6\}$*,* $\mathcal{C}$ *la classe récurrente contenant 1, et* $\mathcal{C}'$ *l'autre classe récurrente. Pour tout* $x, y \in E$*, on définit* $\rho_x := \mathbb{P}_x(T < \infty)$*, où* $T := \inf\{n \geq 0;\, X_n \in \mathcal{C}\}$.
4. *Montrer que*

$$\rho_x \begin{cases} 1, & si\ x \in \mathcal{C}; \\ 0, & si\ x \in \mathcal{C}'; \end{cases}$$