# Chaînes de Markov 1

L'essentiel de ces notes provient de [BEK04]. On pourra aussi consulter [BL98] ou [Nor97] qui est très facile à lire (bien qu'en anglais) et contient de nombreux exemples.

Dans tout ce qui suit $E$ est un espace de cardinal **fini**, très souvent identifié à $\{1, \dots, d\}$.

## 1 Matrices de transition et chaînes de Markov

**Définition 1.** Une matrice de transition (ou noyau de transition ou matrice markovienne) sur l'ensemble $E$ est une application $P : E \times E \to [0, 1]$ telle que

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

Une chaîne de Markov de matrices de transition $(P_n)_{n \geq 0}$ et de loi initiale $\mu$ est une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $E$ telle que $\mathcal{L}(X_0) = \mu$ et

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \mu(\{x_0\}) \prod_{i=0}^{n-1} P_i(x_i, x_{i+1}). \tag{1}$$

pour tout $n \in \mathbb{N}$ et tout $(n + 1)$-uplet $(x_0, \dots, x_n)$ de $E^{n+1}$.

*Remarque* 2. On peut définir alternativement une chaîne de Markov sur $E$ de la façon suivante. Une suite $(X_n)_{n \geq 0}$ est une chaîne de Markov sur $E$ s'il existe une suite $(f_n)_{n \geq 0}$ de fonctions de $E \times [0, 1]$ dans $E$ et une suite de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$ telles que

$$\forall n \geq 0, \quad X_{n+1} = f_n(X_n, U_n).$$

**Définition 3.** La chaîne est dite homogène si la suite de matrices markoviennes $(P_n)_{n \geq 0}$ ou la suite de fonctions $(f_n)_{n \geq 0}$ est constante.

**Exemple 4** (Marche aléatoire sur les sommets du cube)**.** L'espace d'états est $E = \{0, 1\}^d$ les sommets du cube unité dans $\mathbb{R}^d$ et la matrice de transition est définie ainsi : pour $x, y \in E$,

$$P(x, y) = \begin{cases} 1/(2d) & \text{si } |x - y| = 1, \\ 0 & \text{sinon.} \end{cases}$$

**Exemple 5** (Urne d'Ehrenfest)**.** L'espace d'états est $\{0, 1, \dots, d\}$ et la matrice de transition est définie par

$$P(x, y) = \begin{cases} x/d & \text{si } y = x - 1, \\ (d - x)/d & \text{si } y = x + 1, \\ 0 & \text{sinon.} \end{cases}$$

**Exemple 6** (Modèle de Wright-Fisher)**.** L'espace d'états est $\{0, 1, \ldots, N\}$ et

$$P(x, y) = \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y}.$$

En d'autres termes, $\mathcal{L}(X_{n+1}|X_n = x)$ est la loi binomiale $\mathcal{B}(N, x/N)$.

Dans la suite, nous considérerons, sauf mention contraire, des chaînes de Markov homogènes.

## 2 Équation de Chapman-Kolmogorov

Il faut voir une matrice markovienne $P$ comme un opérateur agissant sur les fonctions (identifiées à des vecteurs colonnes) et les mesures identifiées à des vecteurs lignes. L'opérateur $P$ associe à une fonction $h$ (mesurable bornée) de $E$ dans $\mathbb{R}$ la fonction $Ph$ (mesurable bornée) de $E$ dans $\mathbb{R}$ définie par

$$\forall x \in E, \quad Ph(x) = \sum_{y \in E} P(x, y)h(y)$$

et à une mesure positive (respectivement une probabilité) $\mu$ sur $E$ la mesure positive (respectivement la probabilité) $\mu P$ sur $E$ définie par

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

**Théorème 7.** *Soit $X = (X_n)_{n \geq 0}$ une chaîne de Markov homogène sur $E$ de matrice de transition $P$ et de loi initiale $\mu_0$. Notons $\mu_n$ la loi de $X_n$. Alors*

— *la suite $(\mu_n)_{n \geq 0}$ vérifie la relation de récurrence, appelée équation de Chapman-Kolmogorov, suivante :*

$$\mu_{n+1} = \mu_n P = \mu_0 P^{n+1},$$

— *pour tous $x, y \in E$, $\mathbb{P}(X_n = y|X_0 = x) = P^n(x, y)$,*

— *pour toute fonction $h$ de $E$ dans $\mathbb{R}$,*

$$\mathbb{E}(h(X_n)|X_0 = x) = P^n h(x).$$

— *pour toute fonction $h$ de $E$ dans $\mathbb{R}$ et toute probabilité $\mu_0$ sur $E$, si $\mathcal{L}(X_0) = \mu_0$ alors*

$$\mathbb{E}(h(X_n)) = \mu_0(P^n h) = (\mu_0 P^n)h = \mu_0 P^n h.$$

## 3 Propriété de Markov

La propriété de Markov exprime le fait que « conditionnellement au présent, passé et futur sont indépendants ». Introduisons un peu de formalisme. Soit $X$ une chaîne de Markov sur $E$. On peut voir $X$ comme une variable aléatoire à valeurs dans l'espace des trajectoires $\mathbb{T} = E^{\mathbb{N}}$. Pour cela, on munit $\mathbb{T}$ de la tribu engendrée les tribus cylindriques $(\mathcal{B}_n(\mathbb{T}))_{n \geq 0}$ définies par $\mathcal{B}_n(\mathbb{T}) = \sigma(X_0, \ldots, X_n)$.

Si $X_0$ est de loi $\mu$, on note $\mathbb{P}_\mu$ la loi de $X$ et si $\mu = \delta_x$, on note $\mathbb{P}_{\delta_x} = \mathbb{P}_x$.

**Proposition 8.** *La loi d'une chaîne de Markov homogène sur $E$ de matrice de transition $P$ et de loi initiale $\mu$ est l'unique mesure de probabilité $\mathbb{P}_\mu$ sur $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ caractérisée par (1).*

On introduit à présent le processus décalé de $n$ : $X_{n+} = (X_{n+,k})_{k\in\mathbb{N}}$ défini par

$$X_{n+,k} = X_{n+k}, \quad \text{pour } k \geq 0.$$

**Théorème 9** (Propriété de Markov)**.** *Soit $X$ une chaîne de Markov sur $E$ de loi initiale $\mu$. Pour tout $A \in \mathcal{B}(\mathbb{T})$ et $k \in \mathbb{N}$,*

$$\mathbb{P}(X_{k+} \in A | X_k = x_k, \dots, X_0 = x_0) = \mathbb{P}_{x_k}(A).$$

*Remarque* 10. On peut montrer que, pour une chaîne de Markov à espace d'états fini, cette propriété est encore vraie si l'on remplace $k$ par un temps d'arrêt. On parle alors de propriété de Markov forte. Mais ceci nous entraine trop loin pour aujourd'hui...

## 4 Récurrence et transience

**Définition 11.** Soit $X$ une chaîne de Markov sur $E$ de noyau $P$ et $x \in E$. On définit la variable aléatoire $T_x = \inf\{n > 1,\ X_n = x\}$ à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$. L'état $x$ est dit

— transient pour $P$ si $\mathbb{P}_x(T_x < +\infty) < 1$,

— récurrent pour $P$ si $\mathbb{P}_x(T_x < +\infty) = 1$.

*Remarque* 12. On dit que $x \in E$ est absorbant si $P(x,x) = 1$. Un état absorbant est récurrent.

**Proposition 13.** *Si $x$ est récurrent alors la suite $(X_n)_{n\geq 0}$ revient une infinité de fois en $x$ avec probabilité 1.*

*Si $x$ est transient, le nombre (aléatoire) de retours en $x$ de la chaîne issue de $x$ est presque sûrement fini de loi géométrique de paramètre $\mathbb{P}_x(T_x = +\infty)$.*

On peut décomposer une chaîne de Markov en regroupant les états transients et les classes de récurrence que forment les classes d'équivalence de la relation $x \sim y$ si $x$ mène à $y$ et $y$ mène à $x$ (on dit que $x$ mène à $y$ s'il existe $k \in \mathbb{N}$ tel que $P^k(x,y) > 0$).

## 5 Probabilités invariantes

**Définition 14.** Une probabilité $\pi$ sur $E$ est dite invariante (ou stationnaire) pour $P$ si c'est un point fixe de l'équation de Chapman-Kolmogorov :

$$\pi = \pi P.$$

**Définition 15.** Soit $\pi$ une probabilité sur $E$. La matrice de transition $P$ est dite réversible par rapport à $\pi$ si

$$\forall x, y \in E, \quad \pi(x)P(x,y) = \pi(y)P(y,x).$$

**Proposition 16.** *Si $P$ est réversible par rapport à $\pi$ alors $\pi$ est une probabilité invariante pour le noyau $P$.*

**Exemple 17.** La marche aléatoire sur le cube admet une unique mesure invariante (la loi uniforme sur $E$) et la chaîne est réversible par rapport à cette probabilité. Toutes les mesures $p\delta_0 + (1-p)\delta_N$, pour $p \in [0,1]$ sont des mesures invariantes de la chaîne de Wright-Fisher.

### 5.1 Existence

Soit $\mathcal{P}(E)$ l'ensemble des mesures de probabilité sur $E$ fini de cardinal $d$. Cet ensemble est le simplexe unité de $\mathbb{R}^d$ :

$$\mathcal{P}(E) = \left\{ \mu \in \mathbb{R}^d \ : \ \forall x \in E,\ \mu(x) \geq 0,\ \sum_{x\in E} \mu(x) = 1 \right\}.$$

**Théorème 18.** *L'ensemble des probabilités invariantes pour $P$ est un sous-ensemble non vide compact et convexe de $\mathcal{P}(E)$.*

**Corollaire 19.** *Il existe une mesure de probabilité invariante pour $P$.*

### 5.2 Irréductibilité et unicité

Un graphe (orienté) sur un ensemble fini $E$ est un couple $G = (E, \mathcal{A})$ caractérisé par un ensemble d'arêtes $\mathcal{A} \subset E \times E$. Les points de $E$ sont appelés sommets et les points $y$ de $E$ tels que $(x, y) \in \mathcal{A}$ sont les voisins de $x$. On associe à $P$ la structure de graphe orienté $G = (E, \mathcal{A})$ définie par

$$\mathcal{A} = \{(x, y),\ P(x, y) > 0\}.$$

**Définition 20.** Le noyau $P$ est irréductible si, pour tout couple $(x, y) \in E \times E$, il existe $k \in \mathbb{N}$ et $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y \in E$ tels que $P(x_i, x_{i+1}) > 0$ pour $i = 0, \ldots, k-1$ ou encore s'il existe $k \in \mathbb{N}$ tel que $P^k(x, y) > 0$ (on dit que $x$ mène à $y$).

*Remarque* 21. L'irréductibilité de $P$ est équivalente à la connexité du graphe orienté associé.

**Proposition 22.** *Si $P$ est irréductible, alors*
- *les seules fonctions $P$-invariantes ($Pf = f$) sont les fonctions constantes,*
- *$P$ admet une unique probabilité invariante $\pi$. De plus, $\pi(x) > 0$ pour tout $x \in E$.*

Si la chaîne n'est pas irréductible, on peut décomposer son graphe associé en la réunion d'une partie transiente et de composantes connexes de sorte que sur chacune de ces composantes connexes, la chaîne soit irréductible. Toute mesure invariante pour $P$ s'écrit alors comme combinaison convexe des mesures invariantes des chaînes restreintes à chaque composante récurrente.

## 6 Théorème ergodique

**Théorème 23** (Théorème ergodique)**.** *Supposons $P$ irréductible de mesure invariante $\pi$. Alors, pour toute mesure initiale $\mu$, et tout $x \in E$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k = x\}} \xrightarrow[n\to\infty]{\mathbb{P}_\mu-p.s.} \pi(x),$$

*et, si pour tout $x \in E$, $T_x = \inf\{k \geq 1,\ X_k = x\}$ désigne le temps de retour en $x$, alors*

$$\pi(x) = \frac{1}{\mathbb{E}_x(T_x)}.$$

*Remarque* 24. Sous les mêmes hypothèses, pour toute fonction $h$ de $E$ dans $\mathbb{R}$,

$$\frac{1}{n}\sum_{k=0}^{n-1} h(X_k) \xrightarrow[n\to\infty]{\mathbb{P}_\mu-p.s.} \int h\,d\pi.$$

## 7 Convergence en loi

### 7.1 Irréductibilité forte

**Définition 25.** Le noyau $P$ est dit fortement irréductible s'il existe un entier $k$ tel que

$$\forall x, y \in E, \quad P^k(x,y) > 0.$$

*Remarque* 26. Si $P$ est fortement irréductible, non seulement tout état mène à tout autre mais il existe une longueur de chemin commune à tous les couples départ-arrivée.

**Théorème 27.** *Soit $P$ un noyau fortement irréductible de probabilité invariante $\pi$. Alors, pour toute probabilité $\mu$ sur $E$,*

$$\mu P^n \xrightarrow[n\to\infty]{} \pi,$$

*ou encore, pour toute loi initiale, $(X_n)_{n\geq 0}$ converge en loi vers $\mu$.*

### 7.2 Apériodicité

**Définition 28.** Soit $x \in E$ et $R(x) = \{n \in \mathbb{N}^*,\ P^n(x,x) > 0\}$. On appelle période de $x$, et on note $p(x)$, le plus grand commun diviseur de $R(x)$.

**Proposition 29.** *Si $P$ est irréductible, alors tous les points ont même période.*

**Définition 30.** Si la période de $P$ est égale à 1, on dit que $P$ est apériodique.

**Proposition 31.** *Le noyau $P$ est fortement irréductible si et seulement si $P$ est irréductible et apériodique.*

## 8 Exercices

**Exercice 32.** Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de loi initiale $\lambda$ et de matrice de transition $P$. Soit $k \in \mathbb{N}^*$. Montrer que $(X_{kn})_{n\geq 0}$ est une chaîne de Markov dont on précisera les paramètres.

**Exercice 33.** Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov sur $\{1,2,3\}$ de matrice de transition

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ p & 1-p & 0 \end{pmatrix}$$

Calculer $\mathbb{P}(X_n = 1 | X_0 = 1)$ dans les cas suivants : $p = 1/16$, $p = 1/6$ et $p = 1/12$.

**Exercice 34.** Soit Rec (resp. Tra) l'ensemble des états récurrents (resp. transients) d'une chaîne de Markov finie et homogène. Soit $T$ le temps d'atteinte dans l'ensemble Rec. Montrer que pour tout état $i \in$ Tra, on a

$$\mathbb{E}_i(T) = 1 + \sum_{j\in\mathrm{Tra}} p_{ij}\mathbb{E}_j(T).$$

**Exercice 35.** Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov sur $\{1,2,3,4\}$ de matrice de transition

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

— Classer les états de la chaîne.
— Quelle est la probabilité que la chaîne issue de 2 soit absorbée en 4 ?
— Quel est le temps moyen d'absorption de la chaîne issue de 2 ?
— Quelles sont les mesures invariantes de la chaîne ?

**Exercice 36.** On considère la matrice de transition suivante :

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

Représenter le graphe de la chaîne. Quelles sont les classes de la chaîne, les mesures invariantes ? Peut-on parler de chaîne périodique, apériodique ?

**Exercice 37.** Quelle est la mesure invariante de la chaîne de transition

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

**Exercice 38.** Soit $p \in ]0,1[$, $q = 1-p$ et $(X_n)_{n\geq 0}$ la chaîne de Markov sur $0,1,2,\dots,M$ de matrice de transition définie par

$$p_{00} = q, \quad p_{01} = p, \quad \forall i = 1,\dots,M-1, \quad p_{i,i-1} = q, \quad p_{i,i+1} = p, \quad p_{M,M-1} = q, \quad p_{M,M} = p.$$

Montrer que la chaîne admet une mesure réversible que l'on explicitera.

**Exercice 39.** Sur l'ensemble fini $E = \mathbb{Z}/m\mathbb{Z}$, on considère la chaîne de Markov $(X_n)_{n\geq 0}$ de transition $P$ définie par

$$P_{i,i+k} = P_{i,i-k} = 1/2, \quad P_{i,j} = 0 \text{ sinon pour } k \in \{1,\dots,m-1\}.$$

Donner, dans tous les cas, ses classes de récurrence et la mesure invariante de ces classes. Lorsque la chaîne est récurrente irréductible, déterminer quand elle est apériodique.

**Exercice 40.** On considère la chaîne d'Ehrenfest décrite dans l'exemple 5.
1. Grâce à la section 2, calculer l'espérance de $X_n$ sachant que $X_0 = i$.

2. Montrer que la loi binomiale $\mathcal{B}(d, 1/2)$ (que l'on notera $\pi$) est invariante (penser à la réversibilité...). Est-elle unique ?
3. On suppose que la chaîne est à l'état stationnaire. Comparer $\pi(d/2)$ et $\pi(0)$. Commenter...
4. Lorsque $d$ est grand, estimer que la probabilité d'observer la chaîne dans un état n'appartenant pas à l'intervalle $[(d-5\sqrt{d})/2, (d+5\sqrt{d})/2]$.
5. Déterminer la période de la chaîne $X$. Soit $F = \{0, 1, \ldots, d\} \cap (2\mathbb{N})$ et $z \in F$. Montrer que la suite $Y^z$ définie par $Y_n^z = X_{2n}^z$ est une chaîne de Markov sur $F$ dont on exprimera la mesure invariante en fonction de celle de $X$. Si $X_0$ est à valeurs dans $E$ quel est le comportement en temps en long de la loi de $X_n$ ? Les résultats de cette question admettent-ils une généralisation ?
6. Montrer que la chaîne d'Ehrenfest est l'image de la marche aléatoire sur le cube par une fonction que l'on précisera. Soit $E$ et $F$ deux ensembles finis, $f : E \to F$ et $X$ une chaîne de Markov sur $E$. La suite $Y$ définie par $Y_n = f(X_n)$ est-elle, en général, une chaîne de Markov ?

**Exercice 41.** On considère un cavalier fou sur un échiquier : à chaque coup, il se déplace de manière équiprobable sur toutes les cases de l'échiquier qu'il a le droit d'atteindre. On note $X_n$ sa position après $n$ déplacement. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov irréductible récurrente. Quelle est sa mesure invariante ?

**Exercice 42.** Soit $(V, \mathcal{E})$ un graphe connexe non orienté d'ensemble de sommets fini $V$ et d'ensemble d'arêtes $\mathcal{E} \in V \times V$. On associe à chaque arête $(i, j)$ un poids $w_{ij} = w_{ji} > 0$ et l'on pose $w_i = \sum_j w_{ij}$. Déterminer la mesure invariante de la chaîne de Markov sur $V$ de matrice de transition $P$ définie par $P_{ij} = w_{ij}/w_i$.

**Exercice 43.** Soit $(X_n)_{n\geq 0}$ une chaîne de Markov sur un espace d'états fini de matrice de transition $P$ avec $P_{ij} > 0$ pour tous $i, j$. On pose $\tau_j = \inf\{n \geq 1 \ ; \ X_n = j\}$. Montrer qu'il existe $\rho \in ]0, 1[$ tel que pour tout $n \geq 1$, tous $i, j$,

$$\mathbb{P}_i(\tau_j > n) \leq \rho^n.$$

**Exercice 44.** Soit $(X_n)_{n\geq 0}$ une chaîne de Markov sur un espace d'états fini de matrice de transition $P$ avec $P_{ij} > 0$ pour tous $i, j$ et de loi initiale $\mu$. On suppose cette chaîne irréductible et apériodique et on note $\pi$ sa mesure invariante. On considère $Y$ une chaîne de Markov de même matrice de transition et de loi initiale $\pi$ et on la suppose indépendante de $X$. On note $T = \inf\{n \geq 0, \ X_n = Y_n\}$ et on définit enfin $Z$ de la façon suivante :

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

1. Remarquer que $(X, Y)$ est une chaîne de Markov sur $E \times E$. Que dire des coefficients de la matrice de transition de la chaîne produit ?
2. Établir que $T$ est un temps d'arrêt pour la filtration du processus $(X, Y)$.
3. Déduire de la propriété de Markov forte, après l'avoir révisée ([BL98, p. 212]) que $Z$ a même loi que $X$.
4. Montrer qu'il existe $\rho \in ]0, 1[$ tel que pour tout $n \geq 1$, tous $i, j$,

$$\mathbb{P}_i(T > n) \leq \rho^n.$$

5. Déduire de ce qui précède que la convergence de la loi de $X_n$ vers $\pi$ est sous-géométrique.

## Références

[BEK04] M. Benaïm et N. El Karoui – *Promenade aléatoire*, Les éditions de l'école polytechnique, 2004.

[BL98] P. Barbe et M. Ledoux – *Probabilités*, De la licence à l'agrégation, Belin, 1998.

[Nor97] J. Norris – *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, 1997.