
Chaînes de Markov : Exemples et Applications.

ECH-CHAMMAKHY MOHAMMED

EDUCATION FELLOW
MATHÉMATIQUES

Encadré par :
DECREUSEFOND LAURENT



*A tous les malheureux, où qu'ils se trouvent,
aux âmes accablées par la dépression, aux malades atteints du cancer (que Dieu les
guérissent), à ceux emprisonnés injustement, et notamment à nos frères palestiniens,
qui font face à une profonde injustice ; espérons que la guerre prend fin et que l'eau
de rivière retrouve son cours !*

Mohammed ECH-CHAMMAKHY

Table des matières

Introduction	3
1 Chaînes de Markov	4
1.1 Introduction	4
1.2 Définitions et propriétés	5
1.3 Graphe de transition	6
1.4 Classification des états	12
1.5 Communication des classes	14
1.6 théorème ergodique	15
2 TD	18
3 TP	20
3.1 File d'attente	20
3.1.1 Enoncé	20
3.1.2 Quelques simulations obtenues	22
3.2 Urnes et particules	22
3.2.1 Le modèle microscopique	23
3.2.2 L'urne d'Ehrenfest	24
3.2.3 Temps de retour	25
3.2.4 Aspect numérique	25
3.2.5 Code	25
3.2.6 Quelques simulations obtenues	27
4 Examen	28
5 Projet	30
5.1 Processus de Poisson : Gestion du Capital d'un Casino	30
5.2 Files d'Attente	30
5.3 Gestion de Stock avec les Chaînes de Markov	30
5.4 Résolution de Systèmes Linéaires par les Chaînes de Markov	30
5.5 Martingales	31
5.6 Mouvement Brownien	31
6 Planning et Evaluation	32

Introduction

Ce cours est destiné aux étudiants en deuxième année en cycle d'ingénieurs dans une école généraliste.

Il s'agit d'une introduction aux processus stochastiques qui a comme préalable la connaissance des notions de probabilités discrètes et les principaux résultats sur les variables aléatoires, notamment la notion d'indépendance et de conditionnement.

On se focalise principalement sur les chaînes de Markov à temps discret, et le point central est le théorème ergodique permettant de décrire le comportement asymptotique d'une chaîne de Markov à l'aide de la distribution stationnaire.

Les travaux pratiques sont issus de deux textes de modélisation de l'agrégation des mathématiques externe au Maroc.

La partie Projet a pour vocation d'encourager l'étudiant à choisir un thème en lien avec les chaînes de Markov et avec un domaine « d'ingénierie » de son intérêt.

Vous trouverez sur cette page web le support de cours ainsi que les TP : Ma page personnelle, des améliorations sont prévues dans l'avenir !

Prérequis

- Algèbre linéaire.
- Probabilités discrètes.
- Conditionnement.
- Indépendance.
- Programmation avec Python.

Objectifs

- Découvrir des applications des chaînes de Markov liées à l'ingénierie.
- Savoir démontrer quelques propriétés élémentaires sur les chaînes de Markov.
- Classer les états d'une chaîne de Markov sur des exemples concrets.
- Simuler numériquement une chaîne de Markov.
- Valider numériquement des résultats théoriques.

1 | Chaînes de Markov

1.1 Introduction

Un processus stochastique est tout simplement un modèle aléatoire dépendant du temps. Lorsque l'espace d'états est fini, ou même dénombrable, on est dans le cadre des chaînes de Markov à temps discret. Il existe également des chaînes de Markov à temps continu, mais cela ne fait pas l'objet de ce cours.

Considérons une puce se trouvant à l'origine à l'instant initial. A chaque unité de temps, on lance une pièce de monnaie. Si l'on obtient 'face' (respectivement 'pile') notre puce se déplace d'une unité vers la gauche (respectivement à droite). Si on note par X_n la variable aléatoire décrivant la position de notre puce au bout de n lancers, en supposant qu'on a répété l'expérience énormément de fois, de manière indépendante et dans les mêmes conditions, alors le processus stochastique $\{X_n, n = 0, 1, \dots\}$ est une chaîne de Markov, qu'on qualifie de marche aléatoire.

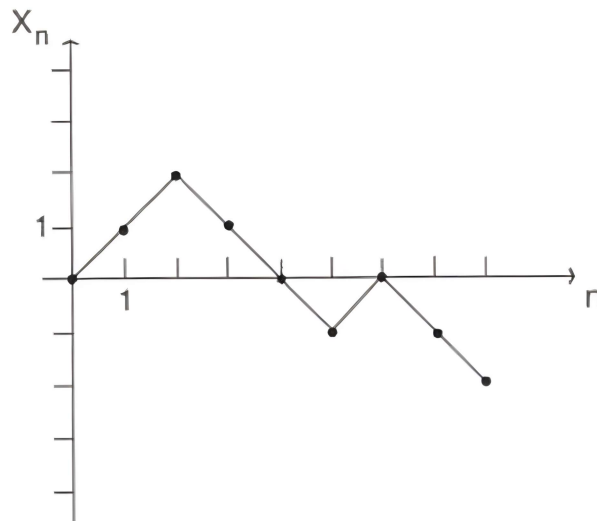


FIGURE 1.1 – Marche aléatoire d'une puce.

D'ailleurs, les chaînes de markov permettent également de modéliser des situations compliquées dans divers domaines, à l'instar de la biologie, la physique, l'économie et la recherche opérationnelle, et la liste est loin d'être exhaustive.

A ce propos, voici un exemple intéressant en sûreté de fonctionnement et gestion de maintenance.

Une machine peut être ou bien en état de fonctionnement, ou bien en état de réparation. On fait l'hypothèse que les temps de fonctionnement et de réparation sont des variables aléatoires indépendantes. De plus, les périodes de fonctionnement et de réparation se succèdent en alternance. Elles se terminent respectivement par une défaillance ou par une remise en service. Le nombre de changements d'état au cours du temps est une chaîne de Markov, et on peut exprimer quelques indicateurs de performance de la machine en fonction de temps caractéristiques issus de la chaîne de Markov.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé dans lequel sont définies toutes les variables aléatoires considérées dans la suite de ce cours. Considérons le processus $X = (X_n)_{n \geq 0}$ défini par une suite de variables aléatoires à valeurs dans un ensemble dénombrable E appelé espace d'état, et soit $\mu = (\mu(i))_{i \in E}$ une loi de probabilité sur E . Pour $i \in E$, la notation $X_n = i$ signifie que le processus X est dans l'état i à l'instant n .

1.2 Définitions et propriétés

Définition 1.2.1. On dit que le processus $X = (X_n)_{n \geq 0}$ est une chaîne de Markov à temps discret ($n \in \mathbb{N}$), à espace d'état E et de loi initiale μ si et seulement si

$$\forall n \geq 0, \forall i, i_0, \dots, i_{n-1}, j \in E,$$

$$\mathbb{P}(X_{n+1} = j / X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j / X_n = i)$$

$$\mathbb{P}(X_0 = i) = \mu(i)$$

La probabilité $\mathbb{P}(X_{n+1} = j / X_n = i)$ est appelée probabilité de transition, de la chaîne X , de l'état i à l'état j entre les instants n et $n + 1$.

Remarque 1.2.2. L'indice de la variable aléatoire représentant le temps, X_n traduit l'observation du processus à l'instant n . En particulier, X_0 est l'état initial du processus.

Remarque 1.2.3. La propriété markovienne se reformule autrement : L'état futur du processus ne dépend du passé qu'à travers le présent. De manière plus formelle ; la variable aléatoire X_{n+1} (l'état futur) ne dépend que de la variable aléatoire X_n (l'état présent) et non pas des variables aléatoires X_0, \dots, X_{n-1} (les états passés).

Exemple 1 : (Gestion de stock) Un magasin garde en inventaire au plus 3 unités d'un produit, alors que chaque jour la demande est de 0, 1, 2 et 3 avec probabilités $1/8$, $2/8$, $3/8$, et $2/8$

Il renouvelle l'inventaire à 3 unités pour le lendemain matin seulement si le nombre

d'unités du produit est inférieur strictement à 2 en fin de journée. Alors le nombre d'unités à la fin de la journée est une chaîne de Markov d'espace d'états 0,1,2,3.

Exemple 2 : (Sûreté de fonctionnement) Une papeterie comprend 3 petites machines servant lors de la première phase pour préparer la pâte, fonctionnant indépendamment l'une de l'autre. Chacune ayant une fiabilité de 0.92 au cours d'une journée, ce qui revient à dire que sa probabilité de tomber en panne dans cette journée est de 0.08. Il faut une nuit pour réparer une machine qui tombe en panne, mais une seule à la fois peut être réparée. Le nombre de machines en état de fonctionnement au début d'une journée est une chaîne de Markov d'espace d'état 0,1,2,3.

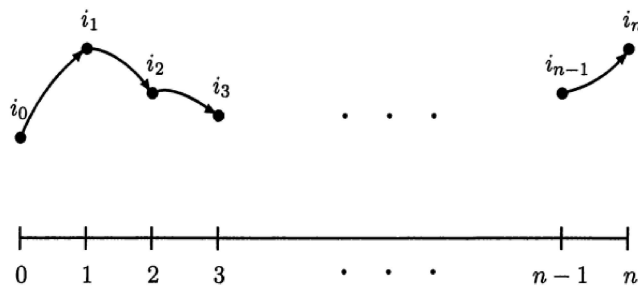


FIGURE 1.2 – Illustration de transitions de l'instant 0 à l'instant n .

Définition 1.2.4. La chaîne de Markov X est dite homogène dans le temps si et seulement si la probabilité $\mathbb{P}(X_{n+1} = j / X_n = i)$ est indépendante du temps n . Dans ce cas on note $K(i, j) = \mathbb{P}(X_{n+1} = j / X_n = i)$ (ou tout simplement $p_{i,j}$), et la matrice ainsi obtenue $K = (K(i, j))_{i,j \in E}$ est appelée matrice de transition de la chaîne de Markov homogène X .

Remarque 1.2.5. Une chaîne de Markov à temps discret et homogène X est caractérisée par le triplet (E, μ, K) , où E est son espace d'état, $\mu = (\mu(i))_{i \in E}$ est sa loi initiale, $\mu(i) = \mathbb{P}(X_0 = i)$, et $K = (K(i, j))_{i,j \in E}$ est sa matrice de transition.

Remarque 1.2.6. Pour tout état i fixé dans E , $\{K(i, j), j \in E\}$ est une loi de probabilité sur E . Les lignes de la matrice de transitions sont donc des lois de probabilité sur E .

Notations : On note $X \sim (E, \mu, K)$ pour désigner une chaîne de Markov $X = (X_n)_{n \geq 0}$ à temps discret, à espace d'état E , de loi initiale $\mu = (\mu(i))_{i \in E}$, et de matrice de transition $K = (K(i, j))_{i,j \in E}$.

1.3 Graphe de transition

Une chaîne de Markov est généralement représentée par son graphe de transition, graphe orienté et évalué dont les sommets sont les états de la chaîne de Markov (éléments de E) et les arcs sont les transitions possibles entre états; Au dessus de chaque arête on écrit la probabilité de transition correspondante.

Remarque 1.3.1. Une matrice possédant ces deux propriétés $p_{i,j} \geq 0$ et $\sum p_{i,j} = 1$ est dite matrice stochastique.

Proposition 1. Soit P une matrice stochastique. Alors P admet comme valeur propre 1 et un vecteur associé à cette valeur propre est $V = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

Démonstration. Calcul immédiat. □

Considérons l'exemple suivant :

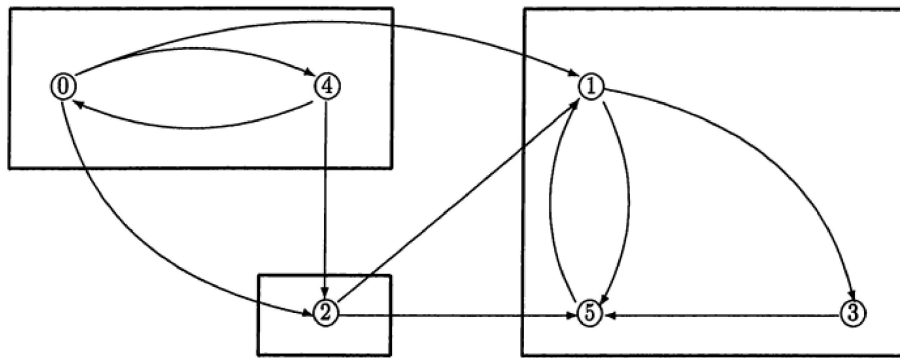


FIGURE 1.3 – Graphe de transition à 6 états.

dont la matrice de transition est :

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

Définition 1.3.2. On dit qu'une chaîne de Markov $X \sim (E, \mu, K)$ est irréductible si et seulement si

$$\forall i, j \in E, \quad \exists m \in \mathbb{N}^* \quad K^{(m)}(i, j) > 0$$

où $K^{(m)}(i, j)$ est l'entrée (i, j) de la matrice K^m . Cela représente la probabilité de passer de l'état i à l'état j en m transitions. Nous avons donc pour tout $i, j \in E$,

$$K^{(m)}(i, j) = \sum_{i_1, \dots, i_{m-1} \in E} K(i, i_1) K(i_1, i_2) \cdots K(i_{m-1}, j)$$

Proposition 2. Soit une chaîne de Markov homogène $X \sim (E, \mu_0, K)$.
Alors $\forall n \geq 0, \forall i_0, i_1, \dots, i_n \in E$

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mu_0(i_0) K(i_0, i_1) K(i_1, i_2) \cdots K(i_{n-1}, i_n) \quad (1.1)$$

Démonstration. Elle se fait par récurrence sur n . La propriété est vraie pour $n = 0$ par définition de μ_0 . Supposons la vraie au rang $n - 1$. Distinguons deux cas :

1. Si $P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) = 0$, il résulte de l'hypothèse de récurrence que $\mu_0(\{i_0\}) P_{i_0 i_1} \cdots P_{i_{n-2} i_{n-1}} = 0$, et donc

$$\mu_0(\{i_0\}) P_{i_0, i_1} \cdots P_{i_{n-1} i_n} = 0.$$

Or $P(X_0 = i_0, \dots, X_n = i_n) = 0$ et la propriété est vraie dans ce cas.

2. Si maintenant $P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) > 0$, il vient

$$\begin{aligned} & P(X_0 = i_0, \dots, X_{n-1} = i_n) \\ &= P(X_n = i_n \mid X_{n-1} = i_{n-1}, \dots, X_0 = i_0) P(X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\ &= P\left(X_n = i_n \mid \bigcap_{0 \leq j \leq n-1} \{X_j = i_j\}\right) P\left(\bigcap_{0 \leq j \leq n-1} \{X_j = i_j\}\right) \\ &= P(X_n = i_n \mid X_{n-1} = i_{n-1}) \mu_0(\{i_0\}) P_{i_0 i_1} \cdots P_{i_{n-2} i_{n-1}} \\ &= \mu_0(\{i_0\}) P_{i_0 i_1} \cdots P_{i_{n-1} i_n}, \\ &= \mu_0(i_0) K(i_0, i_1) K(i_1, i_2) \cdots K(i_{n-1}, i_n) \end{aligned}$$

ce qui démontre la proposition. □

On peut généraliser cette égalité :

$$\mathbb{P}(X_r = i_r, X_1 = i_1, \dots, X_n = i_n) = \mu(i_r) K(i_r, i_1) K(i_1, i_2) \cdots K(i_{n-1}, i_n) \quad (1.2)$$

théorème 1.3.3. Pour tout $n \in \mathbb{N}$, pour tout $m \geq 1$, pour tout $x_0, x_1, \dots, x_{n+m} \in \{1, \dots, N\}$, on a :

$$\begin{aligned} & \mathbb{P}(X_{n+m} = x_{n+m}, \dots, X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) \\ &= \mathbb{P}(X_m = x_{n+m}, \dots, X_1 = x_{n+1} \mid X_0 = x_n) \end{aligned}$$

Démonstration. D'après la définition des probabilités conditionnelles et la proposi-

tion précédente on a

$$\begin{aligned}
\mathbb{P}(X_{n+m} = x_{n+m}, \dots, X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) &= \\
&= \frac{\mathbb{P}(X_{n+m} = x_{n+m}, \dots, X_0 = x_0)}{\mathbb{P}(X_n = x_n, \dots, X_0 = x_0)} \\
&= \frac{\mu_0(x_0) K(x_0, x_1) \dots K(x_{n+m-1}, x_{n+m})}{\mu_0(x_0) K(x_0, x_1) \dots K(x_{n-1}, x_n)}, \text{ d'après la proposition précédente} \\
&= K(x_n, x_{n+1}) \dots K(x_{n+m-1}, x_{n+m}) \\
&= \frac{\mathbb{P}(X_m = x_{n+m}, \dots, X_0 = x_n)}{\mathbb{P}(X_0 = x_n)}, \text{ d'après la proposition précédente} \\
&= \mathbb{P}(X_m = x_{n+m}, \dots, X_1 = x_{n+1} \mid X_0 = x_n).
\end{aligned}$$

□

Proposition 3. Soit une chaîne de Markov homogène $X \sim (E, \mu, K)$. Alors

$$\forall m > 0, \quad \mathbb{P}(X_m = j) = \sum_{i \in E} \mu(i) K^{(m)}(i, j), \quad \forall j \in E$$

Démonstration. A faire en TD.

□

Proposition 4. Soit une chaîne de Markov homogène $X \sim (E, \mu, K)$. Alors $\forall n \geq 0, \forall m > 0, \forall i, j \in E$,

$$\mathbb{P}(X_{n+m} = j \mid X_n = i) = \mathbb{P}(X_m = j \mid X_0 = i) = K^{(m)}(i, j)$$

Démonstration. A faire en TD.

□

théorème 1.3.4. (Equation de **Chapman-Kolmogorov**) Soit X une chaîne de Markov d'espace d'états E , de mesure initiale μ_0 et de matrice de transition K . Alors on a pour tout $n \in \mathbb{N}$

$$K^{(n)} = K^n$$

où $K^{(n)}$ est la matrice de transition en n pas.

Démonstration. On procède par récurrence sur n . Pour $n = 0$ on a $K^{(0)} = K^0 = \mathbb{I}$ où \mathbb{I} est la matrice identité. Soit $n \in \mathbb{N}^*$.

$$\begin{aligned}
p_{ij}^{(n+1)} &= \mathbb{P}(X_{n+1} = j \mid X_0 = i) \\
&= \sum_{k \in E} \mathbb{P}(X_{n+1} = j, X_n = k \mid X_0 = i) \\
&= \sum_{k \in E} \mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) \cdot \mathbb{P}(X_n = k \mid X_0 = i)
\end{aligned}$$

D'une part X est une chaîne de Markov, alors

$$\mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) = \mathbb{P}(X_{n+1} = j \mid X_n = k)$$

d'autre part X est homogène, donc $\mathbb{P}(X_{n+1} = j \mid X_n = k) = p_{kj}$. D'où

$$p_{ij}^{(n+1)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}$$

Ainsi

$$K^{(n+1)} = K^{(n)} \times K$$

On en déduit le résultat par hypothèse de récurrence. \square

Proposition 5. Soit une chaîne de Markov homogène $X \sim (E, \mu, K)$. Nous avons alors les équations de Chapman-Kolmogorov suivantes :

$$\forall m \geq 2, \forall 0 < r < m, \forall i, j \in E, \quad K^{(m)}(i, j) = \sum_{k \in E} K^{(m-r)}(i, k) K^{(r)}(k, j)$$

Démonstration. A faire en TD. \square

Exercice : Dépenses énergétiques, d'après le livre "Exercices de probabilité" Cottrel

Supposons qu'on dispose de deux systèmes de chauffage, l'un de base l'autre d'appoint. On dira que l'on est dans l'état 1 si seul le chauffage de base fonctionne, dans l'état 2 si les deux systèmes fonctionnent. Si un jour on est dans l'état 1, on estime qu'on y reste le lendemain avec probabilité $1/2$; par contre si l'on est dans l'état 2, le lendemain la maison est chaude et l'on passe à l'état 1 avec probabilité $3/4$.

Soit X_n l'état du système au jour numéro n ; on admet que (X_n) est une chaîne de Markov.

1. Déterminer la matrice de transition et son graphe.
2. On pose $p_n = \mathbb{P}(X_n = 1)$. Déterminer une relation de récurrence entre p_n et p_{n+1} , puis exprimer p_n en fonction de p_0 . Que vaut $\lim_{n \rightarrow \infty} p_n$?
3. On suppose que l'on est dans l'état 1 un dimanche, trouver la probabilité d'être dans le même état le dimanche suivant.
4. Montrer que si un jour on se retrouve dans l'état 1 avec probabilité $\frac{3}{5}$ alors il en est de même les jours qui suivent.

Corrigé

1. La chaîne de Markov $(X_n)_{n \geq 0}$ est à valeurs dans $E = \{1, 2\}$. Sa matrice de transition $P = (p(x, y))$ est de taille 2×2 et d'après les hypothèses de l'énoncé, elle est donnée par :

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

2. D'après la formule des probabilités totales et conditionnelles,

$$\begin{aligned}
p_{n+1} &= \mathbb{P}(X_{n+1} = 1) = \mathbb{P}(X_{n+1} = 1, X_n = 1) + \mathbb{P}(X_{n+1} = 1, X_n = 2) \\
&= \mathbb{P}(X_{n+1} = 1 \mid X_n = 1)\mathbb{P}(X_n = 1) + \mathbb{P}(X_{n+1} = 1 \mid X_n = 2)\mathbb{P}(X_n = 2) \\
&= p(1, 1)p_n + p(2, 1)(1 - p_n), \quad \text{car } \mathbb{P}(X_n = 2) = 1 - \mathbb{P}(X_n = 1) \\
&= \frac{3}{4}p_n + \frac{1}{4}(1 - p_n) \\
&= \frac{3}{4} - \frac{1}{4}p_n.
\end{aligned}$$

On peut se ramener à une suite géométrique en soustrayant de l'équation ci-dessus la solution de l'équation $x = \frac{3}{4} - \frac{1}{4}x$, soit $x = \frac{3}{5}$.

$$\begin{aligned}
p_n - \frac{3}{5} &= -\frac{1}{4} \left(p_{n-1} - \frac{3}{5} \right) \\
&= \left(-\frac{1}{4} \right)^n \left(p_0 - \frac{3}{5} \right),
\end{aligned}$$

d'où

$$p_n = \frac{3}{5} + \left(-\frac{1}{4} \right)^n \left(p_0 - \frac{3}{5} \right).$$

et

$$\lim_{n \rightarrow \infty} p_n = \frac{3}{5}.$$

3. On cherche $\mathbb{P}(X_{n+7} = 1 \mid X_n = 1) = \mathbb{P}(X_7 = 1 \mid X_0 = 1)$. Supposer que l'on est dans l'état 1 le dimanche revient à supposer que $\mathbb{P}(X_0 = 1) = 1 = p_0$ (et donc $\mathbb{P}(X_0 = 2) = 0$). Donc,

$$\begin{aligned}
p_7 &= \mathbb{P}(X_7 = 1) = \mathbb{P}(X_7 = 1 \mid X_0 = 1)\mathbb{P}(X_0 = 1) + \mathbb{P}(X_7 = 1 \mid X_0 = 2)\mathbb{P}(X_0 = 2) \\
&= \mathbb{P}(X_7 = 1 \mid X_0 = 1).
\end{aligned}$$

On cherche donc p_7 . D'après la question précédente,

$$p_7 = \frac{3}{5} + \left(-\frac{1}{4} \right)^7 \left(p_0 - \frac{3}{5} \right) \approx \frac{3}{5}.$$

4. La relation de récurrence établie sur (p_n) nous dit que si $p_{n-1} = \frac{3}{5}$ alors $p_n = \frac{3}{5}$.

1.4 Classification des états

La classification des états permet une description de la chaîne de Markov, ainsi que son comportement à long terme.

Notons P_x la probabilité conditionnelle $P_{X_0=x}[\cdot]$ sachant l'événement $\{X_0 = x\}$ réalisé.

Définition 1.4.1. Soit $x \in E$ un état. Le temps d'atteinte de x , noté T_x , est le premier instant où x est visité après le départ. Par convention, le temps d'atteinte est infini si nous n'atteignons jamais x :

$$\forall \omega \in \Omega, T_x(\omega) = \begin{cases} \inf\{n > 0 \mid X_n(\omega) = x\} & \text{si un tel entier existe} \\ +\infty & \text{sinon.} \end{cases}$$

Si la chaîne part de l'état x lui-même, nous employons plutôt le terme de temps de retour.

Définition 1.4.2. Un état $x \in E$ est dit récurrent si

$$P_x(T_x < +\infty) = 1.$$

L'état x est dit transient ou transitoire sinon, c'est-à-dire quand

$$P_x(T_x < +\infty) < 1, \text{ autrement dit, } P_x(T_x = \infty) > 0.$$

Remarque 1.4.3. Autrement dit, Un état i est récurrent si

$$f_{ii} = P(\text{retour à } i \mid \text{départ de } i) = 1.$$

Sinon ($f_{ii} < 1$), alors l'état est transient.

Définition 1.4.4. Soit une chaîne de Markov et x un état récurrent. Alors x est dit récurrent positif si $\mathbb{E}_x(T_x) < \infty$, il est dit récurrent nul si $\mathbb{E}_x(T_x) = \infty$.

Remarque 1.4.5. Un état récurrent i est récurrent nul si

$$\mu_i = \mathbb{E}(\text{temps de premier retour à } i \mid \text{départ de } i) = \infty.$$

Sinon ($\mu_i < \infty$), alors l'état est récurrent positif. Ici, le temps est mesuré en nombre de transitions.

Définition 1.4.6. La période $d(i)$ d'un état i est définie par

$$d(i) = \text{pgcd}\{n \geq 1 : P(X_n = i \mid X_0 = i) > 0\} = \text{pgcd}\{n \geq 1 : K_{ii}^{(n)} > 0\}.$$

Par convention, $d(i) = 0$ si l'ensemble ci-dessus est vide.

Un état i est périodique si $d(i) > 1$ et apériodique sinon.

Définition 1.4.7. Un état récurrent positif apériodique est dit ergodique. et si $K_{ii} = 1$, l'état i est dit absorbant.

Exemple :

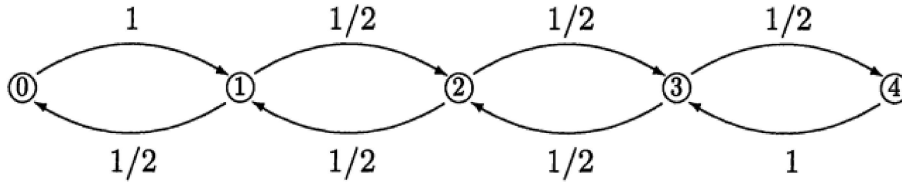


FIGURE 1.4 – Exemple pour la classification des états.

Dans l'exemple de la figure 1.4, l'état 2 est récurrent, car la probabilité de non retour à cet état est nulle.

De plus, $d(2) = 2$, car on retourne à 2 à partir de 2 avec une probabilité strictement positive seulement en $2k$ pas, pour $k \geq 1$.

théorème 1.4.8. Un état i est récurrent si et seulement si

$$\sum_{n \geq 1} K_{ii}^{(n)} = \infty.$$

Sinon, il est transient, et alors nécessairement $\lim_{n \rightarrow \infty} K_{ii}^{(n)} = 0$.

(b) Un état récurrent i est récurrent nul si et seulement si

$$\lim_{n \rightarrow \infty} K_{ii}^{(n)} = 0.$$

Sinon, il est récurrent positif.

Démonstration. Conséquence du théorème ergodique : à venir. □

1.5 Communication des classes

Définition 1.5.1. Soient deux états i et j . On dit que i conduit à j ou que j est accessible depuis i , noté $i \rightarrow j$, si :

$$\exists n \in \mathbb{N}, K_{i,j}^{(n)} = P(X_n = j \mid X_0 = i) > 0.$$

Cette relation signifie que, partant de i , nous avons une probabilité non nulle d'atteindre j après n transitions.

Nous dirons que i communique avec j , noté $i \leftrightarrow j$, si chacun des états i et j est accessible depuis l'autre, c'est-à-dire $i \rightarrow j$ et $j \rightarrow i$.

Proposition 6. La relation de communication \leftrightarrow est une relation d'équivalence.

Démonstration. — La relation est réflexive : $i \leftrightarrow i$, puisque $K_{ii}^{(0)} = 1 > 0$;
— Elle est symétrique : $j \leftrightarrow i$ si $i \leftrightarrow j$, par définition de la relation ;
— Elle est de plus transitive : $i \leftrightarrow k$ si $i \leftrightarrow j$ et $j \leftrightarrow k$, puisqu'alors $K_{ij}^{(m)} P_{jk}^{(n)} > 0$ pour certains $m, n \geq 0$, et donc par les équation de Chapman-Kolmogorov :

$$K_{ik}^{(r+m)} = \sum_{n \geq 0} K_{ij}^{(n)} K_{jk}^{(m)} \geq K_{ij}^{(n)} K_{jk}^{(m)} > 0.$$

□

Proposition 7. Soient deux états i et j qui communiquent, alors i et j sont soit tous les deux récurrents, soit tous les deux transients. De plus, Si i et j sont périodiques alors ils ont la même période.

Définition 1.5.2. — Les états E de la chaîne peuvent être partitionnés en classes d'équivalence appelées classes irréductibles. Si E est réduit à une seule classe, la chaîne de Markov est dite irréductible.

— La relation d'accessibilité s'étend aux classes : une classe d'équivalence C' est accessible depuis une classe C , noté $C \rightarrow C'$, si

$$\forall (x, x') \in C \times C', x \rightarrow x'.$$

— Une classe d'équivalence C est dite fermée si, pour tout x, y , tels que $(x \in C \text{ et } x \rightarrow y) \Rightarrow y \in C$.

Remarque 1.5.3. Il résulte de la proposition 7 que les éléments de la même classe sont de même nature, et pour ceci, elle sera qualifiée de récurrente, positive ou nulle, ou bien transiente en fonction de la classification de ses éléments. De plus, on peut montrer que toute classe récurrente est fermée, et que toute chaîne irréductible sur un ensemble d'états fini est récurrente positive.

Exemple

Reprenons l'exemple de la page 8 :

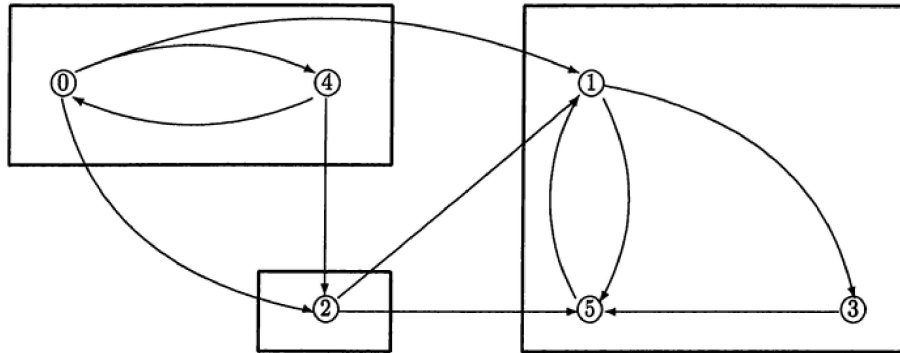


FIGURE 1.5 – Exemple de classes de communication

Au vu de ce qui précède, on constate que :

- $\{2\}$ est transiente.
- $\{0,4\}$ est transiente.
- $\{1,3,5\}$ est récurrente positive.

1.6 théorème ergodique

Dans cette partie, on aborde la notion de mesure invariante permettant de caractériser la chaîne de Markov à long terme. Le théorème central est le théorème ergodique. Cette partie est fortement inspirée de [1] et vous y trouvez toutes les démonstrations.

théorème 1.6.1. (Théorème ergodique)

(a) Dans tous les cas sauf peut-être dans celui où j est un état récurrent positif périodique, on a

$$\lim_{n \rightarrow \infty} K_{ij}^{(n)} = \frac{f_{ij}}{\mu_j}$$

où

$$f_{ij} = P(\text{visite de } j \mid \text{départ de } i)$$

et

$$\mu_j = \mathbb{E}(\text{temps de premier retour à } j \mid \text{départ de } j).$$

(b) Dans tous les cas, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} K_{ij}^{(k)} = \frac{f_{ij}}{\mu_j}.$$

Définition 1.6.2. (loi stationnaire ou invariante). Une suite finie ou infinie dénombrable $\pi = (\pi_j)$ est appelée une distribution stationnaire pour une chaîne irréductible, donc avec une seule classe d'états, si et seulement si les conditions suivantes sont satisfaites :

1. $\pi_j > 0$ pour tout j ;
2. $\sum_i \pi_i = 1$;
3. $\pi_j = \sum_i \pi_i K_{ij}$ pour tout j , c'est-à-dire $\pi = \pi K$ en notation matricielle.

théorème 1.6.3. Une chaîne irréductible a une distribution stationnaire (π_j) si et seulement si elle est récurrente positive. Dans ce cas, la distribution stationnaire est unique et donnée par $\pi_j = \frac{1}{\mu_j}$ pour tout état j , où $\mu_j = \mathbb{E}(\text{temps de premier retour à } j \mid \text{départ de } j)$.

Remarque 1.6.4. Toute chaîne de Markov irréductible sur un espace d'états fini est récurrente positive.

théorème 1.6.5. Soit (X_n) une chaîne de Markov irréductible et apériodique alors P^n converge vers la matrice dont tous les lignes sont constantes égales à μ , en particulier quel que soit la loi de X_0 , X_n converge en loi vers μ .

Démonstration. La preuve se base sur Perron-Frobenius en utilisant le fait que la matrice de transition admet 1 comme unique valeur propre de module 1. \square

Remarque 1.6.6. Remarquons que la condition de matrice stochastique irréductible n'est pas suffisante pour cette conclusion. Par exemple, si l'on considère la matrice stochastique irréductible

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

alors, $A^{2n} = I_2$ et $A^{2n+1} = A$. Par conséquent, la suite $(A^n)_n$ ne converge pas (mais les deux suites extraites $(A^{2n})_n$ et $(A^{2n+1})_n$ convergent). On remarque que dans ce cas,

$$\text{PGCD} \left\{ n \in \mathbb{N}, [A^n]_{i,i} > 0 \right\} = 2.$$

Exemple : Gestion de maintenance

Un appareil peut être soit en réparation (état 0), soit en état de fonctionnement bon, moyen ou mauvais (états 1, 2 ou 3, respectivement). Si l'appareil est dans l'état 1 (bon) ou l'état 2 (moyen) un jour donné, le jour suivant, il peut soit rester dans le même état, soit passer à un état de fonctionnement inférieur, c'est-à-dire 2 ou 3, respectivement. Lorsque l'appareil est en mauvais état de fonctionnement (état 3), il est envoyé en réparation le lendemain, puis remis en bon état (état 1) le jour suivant. Les différents états et les transitions possibles sont illustrés dans la figure 1.6

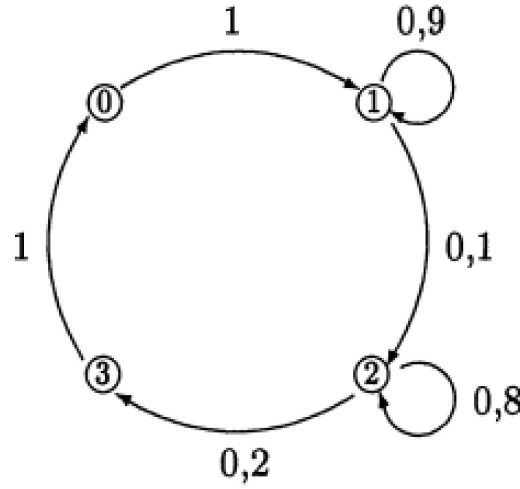


FIGURE 1.6 – Graphe de transition

Voici la matrice de transition pour l'état de l'appareil (0, 1, 2 ou 3) d'un jour au suivant :

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0,9 & 0,1 & 0 \\ 0 & 0 & 0,8 & 0,2 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Tous les états communiquent entre eux, et donc la chaîne est irréductible. Puisque le nombre d'états est fini, la chaîne est récurrente positive. De plus, $P_{11} = 0,9 > 0$, ce qui implique que la chaîne est apériodique, donc ergodique. Dans ce cas, on a

$$\lim_{n \rightarrow \infty} P^{(n)} = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 \end{pmatrix}$$

où $\{\pi_0, \pi_1, \pi_2, \pi_3\} = \pi$ satisfait $\pi = \pi P$, c'est-à-dire

$$\begin{cases} \pi_0 = \pi_3, \\ \pi_1 = \pi_0 + 0,9\pi_1, \\ \pi_2 = 0,1\pi_1 + 0,8\pi_2, \\ \pi_3 = 0,3\pi_2, \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1. \end{cases}$$

On a alors $\pi_0 + 10\pi_0 + 5\pi_0 + \pi_0 = 1$, d'où $\pi_0 = \frac{1}{17}$. C'est la proportion moyenne de jours à long terme que l'appareil passe en réparation. De plus, $\mu_0 = \frac{1}{\pi_0} = 17$ est le nombre moyen de jours qu'un appareil nouvellement réparé fonctionne avant de retourner en réparation.