Chapitre 4

Chaînes de Markov Homogènes

4.1 Définitions, Caractéristiques et Théorème de Chapman-Kolmogorov

4.1.1 Définitions

Définition 4.1. *Une* chaîne de Markov homogène $(X_n)_{n\in\mathbb{N}}$ *est un processus stochastique à* temps discret *et à* espace d'état E dénombrable *qui est* markovien, *i.e.*

$$\forall n \in \mathbb{N}, \quad \forall i_0, \ldots, i_n, \quad \mathbb{P}(X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}),$$

*et vérifie la propriété d'*homogénéité *:* $\forall n, m \in \mathbb{N}$, $m < n$, $\forall i, j \in E$,

$$\mathbb{P}(X_n = j | X_m = i) = \mathbb{P}(X_{n+h} = j | X_{m+h} = i) = \mathbb{P}(X_{n-m} = j | X_0 = i),\ h > 0.$$

Définition 4.2. *Un vecteur* $\lambda = (\lambda_i)_{i\in E}$ *avec* $0 \leq \lambda_i < +\infty$ *est une* mesure *sur* E.

Cette mesure est finie *si* $\sum_{i\in E} \lambda_i < +\infty$.

C'est une probabilité *si* $\sum_{i\in E} \lambda_i = 1$.

Définition 4.3. *Une matrice* $P = (p_{ij})_{i,j\in E}$ *est dite* stochastique *si* $p_{ij} \geq 0$ *pour tous* $i, j \in E$ *et* $\sum_{j\in E} p_{ij} = 1$ *pour tout* $i \in E$.

4.1.2 Caractérisation et premières propriétés

Caractérisation 4.1. *Une chaîne de Markov homogène est* entièrement caractérisée *par*

- *la loi de* X_0 *notée* λ *(c'est la donnée des* $\lambda_i = \mathbb{P}(X_0 = i)$ *pour tout* $i \in E$*) ;*
- *les probabilités de passage de l'état* i *à l'état* j *pour tous* $i, j \in E$,

$$p_{ij} = \mathbb{P}(X_n = j | X_{n-1} = i) = \mathbb{P}(X_1 = j | X_0 = i),$$

par homogénéité (réel dans $[0, 1]$ *qui ne dépend que* (i, j)*).*

La matrice de passage P *est* carrée, *à termes* positifs *et* stochastique. *En effet,* $0 \leq p_{ij} \leq 1$ *car c'est une probabilité et pour tout* $i \in E$,

$$\sum_{j \in E} p_{ij} = \sum_{j \in E} \mathbb{P}(X_1 = j | X_0 = i) = \mathbb{P}(X_1 \in E | X_0 = i) = 1 \quad \textit{par union disjointe.}$$

De plus, P *admet* 1 *comme* valeur propre *de vecteur propre associé* $(1 \cdots 1)^t$ *car* $P(1 \cdots 1)^t = (1 \cdots 1)^t$.

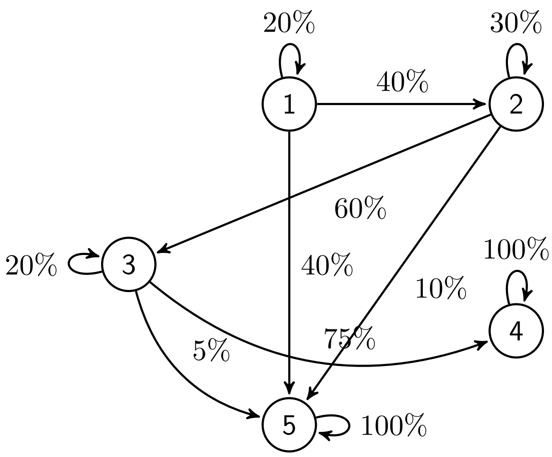
Notation. On notera alors $CM(\lambda, P)$ une chaîne de Markov de loi initiale λ et de matrice de passage P.

Exemple 4.1. Un étudiant en licence a 3 possibilités pour l'année suivante
- soit de passer dans l'année supérieure,
- soit de redoubler,
- soit d'abandonner.

On note $E = \{L_1, L_2, L_3, M, A\}$ l'espace d'états et sa matrice de transition est

$$P = \begin{pmatrix} 0,2 & 0,4 & 0 & 0 & 0,4 \\ 0 & 0,3 & 0,6 & 0 & 0,1 \\ 0 & 0 & 0,2 & 0,75 & 0,05 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Diagramme.



Théorème 4.1. *Soient* λ *une mesure de probabilité sur* E *et* P *une matrice stochastique. Une suite de variables aléatoires* $(X_n)_{0 \leq n \leq N}$ *est une* $CM(\lambda, P)$ *si et seulement si, pour toute suite d'états* $i_0, i_1, \ldots, i_N \in E$,

$$\mathbb{P}(X_0 = i_0, \ldots, X_N = i_N) = \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{N-1}, i_N}.$$

Démonstration. (partielle). **Rappel Formule de Bayes.** $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B)$.

Si $(X_n)_{0\leq n\leq N}$ est une $CM(\lambda, P)$, alors

$$\begin{aligned}
\mathbb{P}(X_0 = i_0, \dots, X_N = i_N) &= \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1|X_0 = i_0)\mathbb{P}(X_2 = i_2|X_0 = i_0, X_1 = i_1)\cdots \\
&\quad \cdots \mathbb{P}(X_N = i_N|X_0 = i_0, \dots, X_{N-1} = i_{N-1}) \\
&= \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1|X_0 = i_0)\mathbb{P}(X_2 = i_2|X_1 = i_1)\cdots \\
&\quad \cdots \mathbb{P}(X_N = i_N|X_{N-1} = i_{N-1}) \\
&= \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{N-1}, i_N}.
\end{aligned}$$

D'autre part, en prenant la somme dans (4.1) sur $i_N, \dots, i_{k+1}$ et en se servant du fait que $\sum_{j\in E} p_{ij} = 1$, on obtient pour tout $k \in \{0, \dots, N-1\}$,

$$\mathbb{P}(X_0 = i_0, \dots, X_k = i_k) = \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{k-1}, i_k}.$$

En particulier, $\mathbb{P}(X_0 = i_0) = \lambda_{i_0}$ et pour $k \in \{0, \dots, N-1\}$,

$$\mathbb{P}(X_{k+1} = i_{k+1}|X_0 = i_0, \dots, X_k = i_k) = \frac{\mathbb{P}(X_{k+1} = i_{k+1}, X_0 = i_0, \dots, X_k = i_k)}{\mathbb{P}(X_0 = i_0, \dots, X_k = i_k)} = p_{i_k, i_{k+1}}.$$

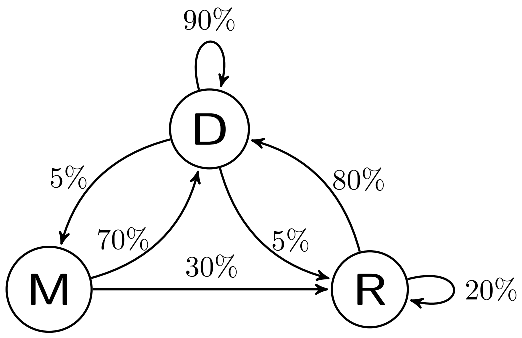
□

Exemple 4.2. Doudou le hamster (voir Wikipedia). 3 états : dort, mange ou exercice dans sa roue. Il change d'activités toutes les minutes de la façon suivante

- quand il dort, 9/10 de ne pas se réveiller la minute suivante,
- quand il se réveille, 1/2 de manger, 1/2 de jouer dans sa roue,
- le repas dure 1 minute après il fait autre chose,
- après manger, 3/10 de jouer dans sa roue et 7/10 de dormir,
- courir est fatigant, donc 8/10 qu'il dorme après, sinon il continue en oubliant qu'il est fatigué.

Doudou suit une chaîne de Markov puisque sa position à la minute n ne dépend que de celle à la minute $n-1$ (le reste des positions passées n'intervient pas). On note $E = \{D, M, R\}$ son espace d'états.

Diagramme de transition.



Matrice de transition.

$$P = \begin{pmatrix} 9/10 & 1/20 & 1/20 \\ 7/10 & 0 & 3/10 \\ 8/10 & 0 & 2/10 \end{pmatrix}.$$

4.1.3 Théorème de Chapman-Kolmogorov

Théorème 4.2. *Soit $(X_n)_{n\geq 0}$ une $CM(\lambda, P)$ à valeurs dans E dénombrable. Pour tout triplet $m, n, p \in \mathbb{N}$, $m < n < p$, pour tout couple d'états $(i,j) \in E^2$,*

$$\mathbb{P}(X_p = j | X_m = i) = \sum_{k\in E} \mathbb{P}(X_p = j | X_n = k)\mathbb{P}(X_n = k | X_m = i).$$

Démonstration. Par la formule de Bayes, on a

$$\mathbb{P}(X_p = j | X_m = i) = \frac{\mathbb{P}(X_p = j, X_m = i)}{\mathbb{P}(X_m = i)}. \tag{4.1}$$

On décompose l'événement $\{X_p = j\} \cap \{X_m = i\} = \bigcup_{k\in E}\{X_p = j\} \cap \{X_m = i\} \cap \{X_n = k\}$ (union disjointe). D'où

$$\begin{aligned}
\mathbb{P}(X_p = j, X_m = i) &= \mathbb{P}\left(\bigcup_{k\in E}\{X_p = j\} \cap \{X_m = i\} \cap \{X_n = k\}\right) \\
&= \sum_{k\in E} \mathbb{P}(X_p = j, X_m = i, X_n = k) \text{ union disjointe} \\
&= \sum_{k\in E} \mathbb{P}(X_p = j | X_m = i, X_n = k)\mathbb{P}(X_m = i, X_n = k) \\
&= \sum_{k\in E} \mathbb{P}(X_p = j | X_m = i, X_n = k)\mathbb{P}(X_n = k | X_m = i)\mathbb{P}(X_m = i).
\end{aligned}$$

D'où le résultat en réutilisant (4.1). □

Théorème 4.3. *Soit $(X_n)_{n\geq 0}$ une $CM(\lambda, P)$. Alors*

$$\forall m, n \geq 0, \quad \forall i, j \in E, \quad \mathbb{P}(X_n = j) = (\lambda P^n)_j, \quad \mathbb{P}(X_{n+m} = j | X_m = i) = (P^n)_{ij} := p_{ij}^{(n)}.$$

Démonstration. Par le Théorème 4.1, pour tous $n \geq 0$ et $j \in E$,

$$\begin{aligned}
\mathbb{P}(X_n = j) &= \sum_{i_0\in E} \cdots \sum_{i_{n-1}\in E} \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = j) \\
&= \sum_{i_0\in E} \cdots \sum_{i_{n-1}\in E} \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1},j} = (\lambda P^n)_j.
\end{aligned}$$

Montrons la seconde propriétés par récurrence. Pour $n = 1$, on a

$$\mathbb{P}(X_{m+1} = j | X_m = i) = p_{ij} = p_{ij}^{(1)}.$$

La propriété est donc vraie au rang $n = 1$. Supposons la vraie au rang n et montrons la au rang $n + 1$. On a alors, en utilisant le Théorème 4.2,

$$\begin{aligned}
\mathbb{P}(X_{m+n+1} = j | X_m = i) &= \sum_{k\in E} \mathbb{P}(X_{m+n+1} = j | X_{m+n} = k)\mathbb{P}(X_{m+n} = k | X_m = i) \\
&= \sum_{k\in E} p_{kj} p_{ik}^{(n)} \text{ par hypothèse de récurrence} \\
&= p_{ij}^{(n+1)}.
\end{aligned}$$

La propriété est donc vraie pour tout rang n. □

Remarque. On aurait pu commencer avec $n = 0$ car par convention $p_{ii}^{(0)} = 1$ et $p_{ij}^{(0)} = 0$ pour $i \neq j$. Les $p_{ij}^{(n)}$ sont les éléments de la puissance $n^{ème}$ de la matrice de passage ou encore de la matrice de passage en n étapes.

Exercice 4.1. Un joueur fréquente 3 casinos numérotés 1, 2 et 3. Chaque jour il choisit l'un des deux casinos où il n'est pas allé la veille, avec probabilité $\frac{1}{2}$. Le premier jour, ($n = 0$), il choisit l'un des trois casinos $1, 2, 3$ avec probabilités $1/4, 1/4, 1/2$ respectivement.

On note X_n la v.a. égale au numéro du casino fréquenté le jour n par le joueur.

1. *Montrer que* $(X_n)_{n\in\mathbb{N}}$ *est une chaîne de Markov et calculer sa matrice de transition* P.
   La loi donnant le numéro du casino visité au $n+1$-ème jour, étant donnée la liste de tous les casinos visités jusqu'au n-ème jour ne dépend que de celui visité au n-ème jour. $(X_n)_{n\geq 0}$ est donc une chaîne de Markov homogène (les probabilités de transitions ne dépendent pas de n) sur un espace à 3 états. Le coefficient $p_{i,j}$ de la matrice de transition est égal à la probabilité d'aller de i à j :
   $$p_{i,j} = P(X_{n+1} = j | X_n = i).$$
   Dans notre cas, la matrice de transition vaut donc
   $$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

2. *Donner* $\mathbb{P}(X_1 = 1 \mid X_0 = 3)$, $\mathbb{P}(X_3 = 3 \mid X_1 = 2)$, $\mathbb{P}(X_8 = 3 \mid X_5 = 2)$.
   On a $P(X_1 = 1|X_0 = 3) = p_{3,1} = 1/2$, $P(X_3 = 3|X_1 = 2) = p_{2,3}^{(2)}$ et $P(X_8 = 3|X_5 = 2) = p_{2,3}^{(3)}$. Nous devons calculer P^2 et P^3. On obtient alors
   $$P^2 = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \quad \text{et} \quad P^3 = \begin{pmatrix} 1/4 & 3/8 & 3/8 \\ 3/8 & 1/4 & 3/8 \\ 3/8 & 3/8 & 1/4 \end{pmatrix}.$$
   Ainsi $P(X_3 = 3|X_1 = 2) = 1/4$ et $P(X_8 = 3|X_5 = 2) = 3/8$.

3. *Donner* $\mathbb{P}(X_1 = 1, X_4 = 3, X_5 = 2, X_7 = 1)$.
   $$\begin{aligned} \mathbb{P}(X_1 = 1, X_4 = 3, X_5 = 2, X_7 = 1) &= P(X_7 = 1|X_5 = 2)P(X_5 = 2|X_4 = 3) \\ &\quad \times \mathbb{P}(X_4 = 3|X_1 = 1)P(X_1 = 1) \\ &= p_{2,1}^{(2)}\, p_{3,2}\, p_{1,3}^{(3)}\, (\lambda P)_1. \end{aligned}$$
   On a que la loi initiale est le vecteur ligne $\lambda = (1/4, 1/4, 1/2)$. Il ne reste donc qu'à calculer $\lambda P = (3/8, 3/8, 1/4)$. Donc
   $$P(X_1 = 1, X_4 = 3, X_5 = 2, X_7 = 1) = \frac{1}{4} \times \frac{1}{2} \times \frac{3}{8} \times \frac{3}{8} = \frac{9}{512}.$$

Exercice 4.2. Une souris effectue une suite de déplacements aléatoires, indépendants les uns des autres entre trois pièces numérotées 1, 2 et 3. La règle des déplacements est alors la suivante

- Lorsque la souris est dans la pièce 1, elle y reste avec la probabilité $\frac{1}{3}$ ou bien passe dans l'une des deux autres pièces suivant la même probabilité $\frac{1}{3}$.
- Lorsque la souris est dans la pièce 2, elle y reste avec la probabilité $\frac{1}{2}$ ou passe dans la pièce 3 avec la probabilité $\frac{1}{2}$.
- Lorsque la souris est dans la pièce 3, elle y reste.

On note X_0 le numéro de la pièce initialement occupée par la souris, (X_0 peut être aléatoire), X_n, $n \geq 1$, le numéro de la pièce occupée par la souris après son n-ième déplacement.

1. *Justifier que X_n est une chaîne de Markov à valeurs dans $\{1,2,3\}$ et donner sa matrice de transition P.*
   La loi de la position de la souris à l'instant $n+1$, étant donné toute l'histoire, ne dépend que de sa position à l'instant n. La suite des pièces visitées (X_n) par la souris est donc une chaîne de Markov à trois états, dont la matrice de transition P est
   $$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. *Calculer $\mathbb{P}(X_{n+3} = 1 \mid X_n = 2)$.*
   On a $\mathbb{P}(X_{n+3} = 1 \mid X_n = 2) = p^{(3)}_{2,1}$. Calculons donc les puissances de P
   $$P^2 = \begin{pmatrix} 1/9 & 5/18 & 11/18 \\ 0 & 1/4 & 3/4 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^3 = \begin{pmatrix} 1/27 & 19/108 & 85/108 \\ 0 & 1/8 & 7/8 \\ 0 & 0 & 1 \end{pmatrix}.$$
   Donc $\mathbb{P}(X_{n+3} = 1 \mid X_n = 2) = 0$.

3. *Calculer $\mathbb{P}(X_0 = 1, X_2 = 3, X_5 = 2, X_7 = 1)$ si X_0 prend des valeurs $1, 2, 3$ avec probabilités $1/6, 1/2, 1/3$ respectivement.*
   La loi initiale est le vecteur ligne $\lambda = (1/6, 1/2, 1/3)$. On a donc
   $$\begin{aligned} \mathbb{P}(X_0 = 1, X_2 = 3, X_5 = 2, X_7 = 1) &= P(X_7 = 1 | X_5 = 2) P(X_5 = 2 | X_2 = 3) \\ &\quad \times \mathbb{P}(X_2 = 3 | X_0 = 1) P(X_0 = 1) \\ &= p^{(2)}_{2,1}\, p^{(3)}_{3,2}\, p^{(2)}_{1,3}\, \lambda_1 \\ &= 0 \times 0 \times \frac{11}{18} \times \frac{1}{6} = 0. \end{aligned}$$

4. *Diagonaliser P. Calculer la matrice P^n. Donner $\mathbb{P}(X_{k+n} = 3 \mid X_k = 1)$.*
   La matrice P est triangulaire, donc ses valeurs propres sont sur la diagonale : 1, $\frac{1}{2}$, $\frac{1}{3}$. Comme elles sont toutes distinctes, la matrice P est diagonalisable. On voit directement que le premier vecteur de base $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ est vecteur propre pour la valeur propre $\frac{1}{3}$. D'autre part, comme la matrice P est stochastique, les coefficients sur chaque ligne somment à 1, donc le vecteur contenant que des 1 est vecteur propre pour la valeur propre 1. Après

calcul, on trouve que $\begin{pmatrix}2\\1\\0\end{pmatrix}$ est vecteur propre associé à la valeur propre $\frac{1}{2}$. Donc, on a $P = RDR^{-1}$, avec

$$R = \begin{pmatrix}1 & 2 & 1\\0 & 1 & 1\\0 & 0 & 1\end{pmatrix}, \qquad D = \begin{pmatrix}1/3 & 0 & 0\\0 & 1/2 & 0\\0 & 0 & 1\end{pmatrix}.$$

Alors

$$D^n = \begin{pmatrix}1/3^n & 0 & 0\\0 & 1/2^n & 0\\0 & 0 & 1\end{pmatrix}.$$

L'inverse de R peut être facilement calculée par le pivot de Gauss

$$R^{-1} = \begin{pmatrix}1 & -2 & 1\\0 & 1 & -1\\0 & 0 & 1\end{pmatrix}.$$

On a l'égalité suivante pour tout $n \geq 0$

$$P^n = RD^nR^{-1} = \begin{pmatrix}3^{-n} & -2.3^{-n} + 2^{1-n} & 3^{-n} - 2^{1-n} + 1\\0 & 2^{-n} & -2^{-n} + 1\\0 & 0 & 1\end{pmatrix}.$$

Donc $P(X_{k+n} = 3 \mid X_k = 1) = p_{1,3}^{(n)} = 3^{-n} - 2^{1-n} + 1$.

4.2 Communication entre états, Périodicité d'un état

4.2.1 Communication entre états

Définition 4.4. *Soit $(X_n)_{n\geq 0}$ une $CM(\lambda, P)$ à valeurs dans E. Soient $i, j \in E$. On dit que i mène à j et on le note $i \to j$ si $\mathbb{P}(\exists n \geq 0 : X_n = j | X_0 = i) > 0$ ou encore $\exists\, n \geq 0$ tel que $p_{ij}^{(n)} > 0$, avec $p_{ij}^{(0)} = \delta_{ij}$. On dit que i et j communiquent (et on le note $i \leftrightarrow j$) si i mène à j et j mène à i.*

Proposition 4.1. *La relation de communication entre états $\leftrightarrow$ est une relation d'équivalence dans l'espace d'états E dénombrable et elle partitionne E en classes disjointes.*

Démonstration. Il est clair que

1. i communique avec i (en 0 pas) car $p_{ii}^{(0)} = 1$ (réflexive) ;
2. si i communique avec j, alors, j communique avec i par définition (symétrique) ;
3. si i communique avec j qui communique avec k, alors i communique avec k (transitive).

Donc $\leftrightarrow$ est une relation d'équivalence.

L'espace d'états E se divise en classes disjointes d'états qui communiquent $E = C_1 \cup C_2 \cup C_3 \cup \cdots$ où pour tout $r = 1, 2, \ldots$ et tout $i, j \in C_r$, i communique avec j. $\square$

Définition 4.5. *Une chaîne de Markov homogène dont tous les états communiquent est dite irréductible (dans le cas contraire, elle est dite réductible).*

Définition 4.6. *Un état i est absorbant s'il ne mène à aucun autre état, i.e. $p_{i,i} = 1$ et sa classe est réduite à lui-même.*

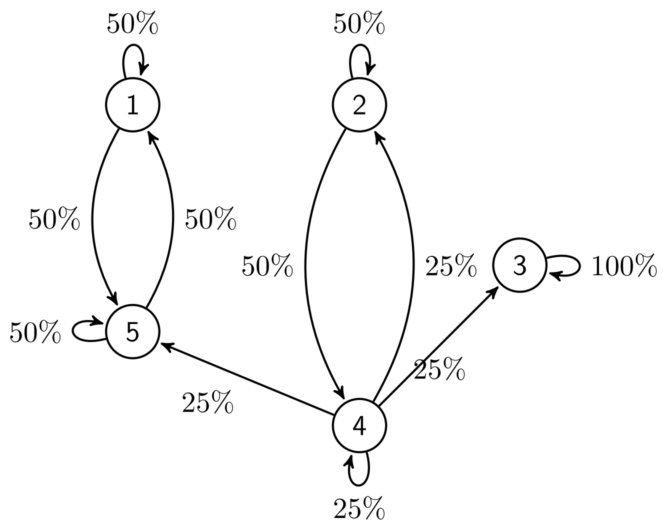
Définition 4.7. *Une classe d'états C qui communiquent est fermée si chaque état j tel qu'il existe un état $i \in C$ menant à j appartient à C.*

Théorème 4.4. *Toute chaîne de Markov homogène sur un espace d'états E fini a au moins une classe fermée d'états qui communiquent.*

Exemple 4.3. 1.

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

Diagramme.

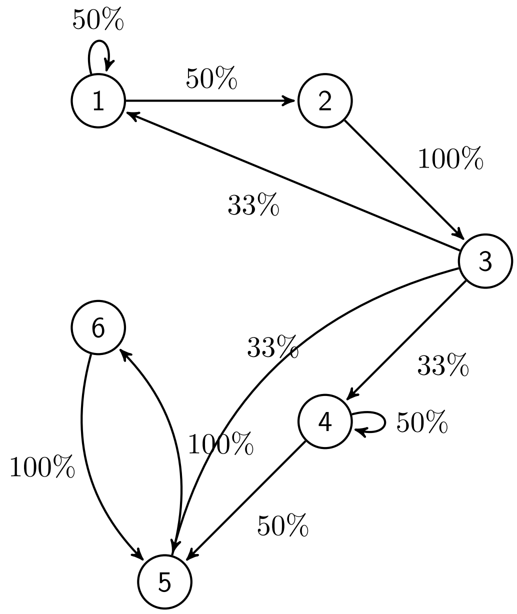


Classes d'états : $\{1, 5\}$ fermée, $\{2, 4\}$ non fermée et $\{3\}$ fermée.

2.

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

Diagramme.



Classes d'états : $\{1, 2, 3\}$ non fermée, $\{5, 6\}$ fermée et $\{4\}$ non fermée.

4.2.2 Périodicité d'un état

Définition 4.8. *Soit $i \in E$. On définit sa périodicité par*

$$d_i = PGCD\{n \geq 1 : p_{ii}^{(n)} > 0\}$$

avec la convention $d_i = +\infty$ s'il n'y a pas d'indice n tel que $p_{ii}^{(n)} > 0$ (on ne revient pas en i).
Si $d_i = 1$, l'état est dit apériodique.

Remarque. Si i est absorbant, alors $d_i = 1$ (la réciproque est fausse).

Proposition 4.2. *Deux états d'une même classe ont même périodicité, i.e. si $i, j \in E$ tels que $i, j \in C$, alors $d_i = d_j$.*

Définition 4.9. *Une chaîne irréductible a tous ces états de même périodicité d et la périodicité commune d est appelée périodicité de P ou de la chaîne.*
Si $d = 1$, la matrice de transition et la chaîne sont dites apériodiques.

4.3 État transitoire ou récurrent, Nombre moyen de passages, Probabilités d'absorption

4.3.1 Récurrence, transience et nombre moyen de passages

Soit $(X_n)_{n\geq 0}$ une $CM(\lambda, P)$ sur E dénombrable. On définit le temps de retour en i par $T^i = \min\{n > 0 : X_n = i\}$ avec la convention usuelle $\inf \emptyset = +\infty$ ($T_i = +\infty$ si $\forall n > 0$, $X_n \neq i$). On notera la probabilité conditionnelle $\mathbb{P}_i$ sur $\mathcal{F}$ définie par : $\forall A \in \mathcal{F}$, $\mathbb{P}_i(A) = \mathbb{P}(A|X_0 = i)$. De même, $\mathbb{E}_i[\cdot]$ désignera l'espérance par rapport à cette mesure.

Définition 4.10. *1. Un état i est dit récurrent si*

$$\mathbb{P}_i(\omega\ : X_n(\omega) = i \text{ pour une infinité de } n\ ) = 1 \quad i.e. \quad \mathbb{P}_i(T_i < +\infty) = 1.$$

2. Un état i est transient si

$$\mathbb{P}_i(\omega\ : X_n(\omega) = i \text{ pour une infinité de } n\ ) = 0 \quad i.e. \quad \mathbb{P}_i(T_i < +\infty) < 1.$$

Remarques.

1. Un état absorbant est récurrent. En effet, sur l'événement $\{X_n = i\}$, on a $T_i = 1$ puisque $p_{ii} = 1$ donc $\mathbb{P}_i(T_i < +\infty) = 1$.
2. L'état i est récurrent si $\mathbb{P}_i(T_i = +\infty) = 0$.
3. L'état i est transient si $\mathbb{P}_i(T_i = +\infty) > 0$.
4. Si $\mathbb{P}_i(T_i = +\infty) = 1$, alors $d_i = +\infty$.

Définition 4.11. *Soit N_i le nombre total de visites de l'état i*

$$N_i = \sum_{n\geq 0} \mathbb{1}_{\{X_n=i\}}.$$

Alors $\mathbb{E}_i[N_i] = \sum_{n\geq 0} p_{ii}^{(n)}$. De même,

$$\mathbb{E}_i[N_j] = \sum_{n\geq 0} \mathbb{P}_i(X_n = j) = \sum_{n\geq 0} p_{ij}^{(n)}.$$

Théorème 4.5. *1. L'état i est récurrent $\Leftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} = +\infty \Leftrightarrow \mathbb{E}_i[N_i] = +\infty$.*

2. L'état i est transient $\Leftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} < +\infty \Leftrightarrow \mathbb{E}_i[N_i] = \frac{1}{1-f_{ii}}$, où $f_{ii} = \mathbb{P}_i(T_i < +\infty)$.

Théorème 4.6. *1. Soit C une classe d'états qui communiquent. Alors tous les états de cette classe sont récurrents ou transients.*

2. Chaque classe récurrente est fermée.

3. Chaque classe fermée finie est récurrente.

Remarque pour les exercices. Si l'espace d'états E est fini, chaque classe fermée est récurrente et chaque classe non-fermée est transiente.

Finalement, le dernier théorème sur ce sujet donne un fait important pour les exercices : dans une classe récurrente **tout état est visité avec probabilité 1** et par conséquent (comme la classe est récurrente) **tout état est visité une infinité de fois avec probabilité 1**.

Théorème 4.7. *Soit une chaîne de Markov homogène irréductible. Alors, pour tout $j \in E$, on a $\mathbb{P}(T^j < \infty) = 1$.*

Proposition 4.3. *Soit $i, j \in E$. Alors*

$$\mathbb{P}_i(T_j < +\infty) = p_{ij} + \sum_{k\neq j} p_{ik}\mathbb{P}_k(T_j < +\infty).$$

Proposition 4.4. *Soit $i, j \in E$ tels que $i \neq j$. Alors*

1. *soit $\mathbb{P}_i(T_j < +\infty) = 0$ et j est non accessible depuis i, donc $\mathbb{E}_i[N_j] = 0$;*
2. *soit $\mathbb{P}_i(T_j < +\infty) > 0$ et i mène à j, donc $\mathbb{E}_i[N_j] = \mathbb{P}_i(T_j < +\infty)\mathbb{E}_j[N_j]$.*

Exercice 4.3. Identifier les classes des matrices de transition suivantes et trouver les classes fermées.

$$P_1 = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ 1/5 & 0 & 2/5 & 0 & 0 & 0 & 2/5 \\ 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 1/4 \\ 1/5 & 1/5 & 0 & 1/5 & 0 & 1/5 & 1/5 \\ 1/5 & 1/5 & 0 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$P_2 = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 & 0 & 0 \\ 2/7 & 1/7 & 0 & 1/7 & 0 & 1/7 & 2/7 \\ 0 & 0 & 0 & 0 & 1/5 & 0 & 4/5 \\ 0 & 0 & 2/7 & 1/7 & 4/7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/5 & 0 & 3/5 \end{pmatrix},$$

$$P_3 = \begin{pmatrix} 1/5 & 0 & 0 & 0 & 4/5 \\ 0 & 1/3 & 2/3 & 0 & 0 \\ 0 & 1/4 & 3/4 & 0 & 0 \\ 1/8 & 2/8 & 1/8 & 3/8 & 1/8 \\ 2/5 & 0 & 0 & 0 & 3/5 \end{pmatrix}.$$

Pour P_1, $\{1,2\}$, $\{7\}$ sont les classes fermées, $\{3\}$, $\{4\}$ et $\{5,6\}$ sont les classes NON fermées.
Pour P_2, $\{1,2,3\}$ et $\{5,7\}$ sont fermées, $\{4,6\}$ est NON fermée.
Pour P_3, $\{1,5\}$, $\{2,3\}$ sont fermées, $\{4\}$ est NON fermée.

4.3.2 Probablité et temps moyen d'absorption

Voici le cheminement de l'étude d'un chaîne de Markov homogène finie que nous avons vu jusqu'à présent (pour les exercices) :

- On trouve d'abord les classes : les classes fermées sont alors récurrentes.
- Si la chaîne de Markov commence dans une classe récurrente, elle va tourner **toujours** dans cette classe et **chaque** état de cette classe sera visité une infinité de fois.
- Si la chaîne de Markov commence dans un des états transients, elle va y revenir seulement un nombre **fini** de fois. Comme le nombre d'états transients pour des chaînes de Markov finies est **fini**, toute classe transiente sera **abandonnée** pour toujours avec probabilité 1 ! ! ! Donc avec probabilité 1, la chaîne de Markov atterrira dans une des classes récurrentes. Avec quelle probabilité ? Dans quelle classe ?

4.3.2.a Probabilité d'absorption

Pour cela il faut savoir calculer les quantités

$$h_i^C = \mathbb{P}(\exists n \geq 0 : X_n \in C \mid X_0 = i), \quad C \subset E.$$

Dans ce calcul, le Théorème 4.7 joue un grand rôle. En effet,

- Soit i récurrent.

$\triangleright$ Si $i \in C$ (et donc C récurrente), alors

$$h_i^C = 1 \quad \text{par le Théorème 4.7}$$

(si la chaîne par de la classe récurrente de i, elle y reste toute sa vie).

$\triangleright$ Si $i \notin C$ et C est récurrente, alors $h_i^C = 0$ (si la chaîne part d'une autre classe récurrente que C, il lui est impossible d'en sortir et donc d'aller dans une autre classe récurrente).

- Si i est transient, on écrit alors un système

$$\forall i \text{ transient}, \quad h_i^C = \sum_{j \in E} h_j^C p_{i,j}.$$

On remplace h_j^C par 1 ou 0 pour tous les états j récurrents dans ce système. Il reste un système où le nombre d'inconnues est égal au nombre d'équations et au nombre d'états transients. On résout le système et on trouve h_i^C pour tout état i transient.

4.3.2.b Temps moyen d'absorption

Finalement, on s'intéresse à k_i le temps moyen d'absorption dans une des classes récurrentes à partir de l'état transient i. On écrit alors

$$k_i = \sum_{j \in E} p_{i,j}(1 + k_j) = 1 + \sum_{j \in E} p_{i,j} k_j$$

(on part de l'état i, en une étape on est à l'état j avec probabilité $p_{i,j}$ puis on regarde le temps d'absorption de j à l'une des classes récurrentes, mais il faut une étape de plus qu'en partant de i).

Ici $k_j = 0$ pour tout j récurrent. On aboutit à un système où le nombre d'inconnues est égal au nombre d'équations et au nombre d'états transients. On le résout et on trouve les k_i pour tout état i transient.

Exercice 4.4. 1. *Que peut-on dire du comportement des chaînes de Markov des exercices 4.1, 4.2 lorsque $n \to \infty$?*
   ○ Dans le cas de l'Exercice 4.1, la chaîne de Markov est irréductible finie, donc récurrente. Chaque état sera visité une infinité de fois *p.s.*.
   ○ Dans le cas de l'Exercice 4.2, La chaîne de Markov va s'absorber dans l'état 3 *p.s.*.

2. *Dans l'Exercice 4.3,*

(a) *Pour la matrice P_1, calculer h_2^1 et h_3^1. En déduire h_3^2 et h_3^7. Donner k_3 le temps moyen d'absorption dans une classe récurrente au départ de l'état 3.*
Pour la matrice P_1, $h_2^1 = 1$ car 1 et 2 sont dans la même classe récurrente. On a aussi

$$h_3^1 = 1/5h_1^1 + 2/5h_3^1 + 2/5h_7^1 = 1/5 + 2/5h_3^1 + 0 \Leftrightarrow h_3^1 = 1/3.$$

Comme $\{1, 2\}$ est une classe récurrente on a $h_3^2 = h_3^1$. On a enfin $h_3^7 = 1 - 1/3 = 2/3$.
Pour le temps moyen d'absorption dans une classe récurrente au départ de l'état 3, on a

$$k_3 = 1 + 1/5k_1 + 2/5k_3 + 2/5k_7 = 1 + 0 + 2/5k_3 + 0 \Leftrightarrow k_3 = 5/3.$$

(b) *Pour la matrice P_2, donner h_4^A et h_6^A quelque soit $A \subset \{1, 2, 3\}$. Que peut-on dire de h_4^B et h_6^B pour $B \subset \{5, 7\}$? Donner k_4 le temps moyen d'absorption dans une classe récurrente au départ de l'état 4.*
Pour la matrice P_2,

$$\begin{aligned} h_4^A &= 2/7h_1^A + 1/7h_2^A + 0h_3^A + 1/7h_4A + 0h_5^A + 1/7h_6^A + 2/7h_7^A \\ h_6^A &= 0h_1^A + 0h_2^A + 2/7h_3^A + 0h_4^A 1/5h_5^A + 0h_6^A + 4/5h_7^A. \end{aligned}$$

Ici $h_1^A = h_2^A = h_3^A = 1$ car $\{1, 2, 3\}$ est une classe récurrente. De même, $h_5^A = h_7^A = 0$ car $\{5, 7\}$ est récurrente fermée. On résout ce système et on obtient $h_4^A = 23/41$, $h_6^A = 15/41$. Donc $h_4^B = 1 - 23/41$, $h_6^B = 1 - 15/41$.
Pour le temps moyen d'absorption dans une classe récurrente au départ de l'état 4, on doit résoudre le système suivant

$$\begin{aligned} k_4 &= 1 + 2/7k_1 + 1/7k_2 + 0k_3 + 1/7k_4 + 0k_5 + 1/7k_6 + 2/7k_7 \\ k_6 &= 1 + 0k_1 + 0k_2 + 2/7k_3 + 1/7k_4 + 4/7k_5 + 0k_6 + 0k_7. \end{aligned}$$

Ici $k_1 = k_2 = k_3 = k_5 = k_7 = 0$. On retrouve $k_4 = 56/41$ et $k_6 = 49/41$.

(c) *Pour la matrice P_3, donner h_4^1 et h_4^2. Calculer k_4.*
Pour la matrice P_3,

$$h_4^1 = 1/8h_1^1 + 2/8h_2^1 + 1/8h_3^1 + 3/8h_4^1 + 1/8h_5^1.$$

Ici $h_1^1 = h_5^1 = 1$ car $\{1, 5\}$ sont dans la même classe récurrente. De même, $h_2^1 = h_3^1 = 0$ car $\{2, 3\}$ est une autre classe récurrente. Donc $h_4^1 = 2/5$. Par conséquent $h_4^2 = 1 - h_4^1 = 3/5$.
Pour le temps moyen d'absorption dans une classe récurrente au départ de l'état 4, on a

$$k_4 = 1 + (1/8 + 2/8 + 1/8)0 + 3/8k_4 + (1/8)0,$$

d'où la valeur de $k_4 = 8/5$.

4.4 Mesure invariante et comportement asymptotique d'une chaîne de Markov homogène finie

4.4.1 Définition de la mesure invariante

Définition 4.12. *Une mesure π sur l'ensemble E est invariante pour la matrice P si et seulement si $\pi = \pi P$.*

Théorème 4.8. *Soit* $(X_n)_{n\geq 0}$ *une* $CM(\pi, P)$ *de mesure invariante* π. *Alors*

$$\forall n \geq 0, \quad \forall i \in E, \quad \mathbb{P}(X_n = i) = \pi_i.$$

Démonstration. $\mathbb{P}(X_n = i) = (\pi P)_i = \pi_i$. □

On voit que si la mesure initiale d'une CM est invariante, alors sa loi à tout instant $n \geq 0$ est la même. La mesure invariante π est juste un vecteur propre de gauche de la matrice P associée à la valeur propre 1.

Proposition 4.5. *Une chaîne de Markov homogène finie admet au moins une mesure invariante (ou distribution stationnaire).*

4.4.2 Calcul de la mesure invariante

Proposition 4.6. *Soit* $(X_n)_{n\geq 0}$ *une* $CM(\lambda, P)$ *de mesure invariante* π. *Alors*

$$\forall i \text{ transient}, \quad \pi_i = 0.$$

Ainsi le long système $\pi = \pi P$, quand on pose $\pi_i = 0$ pour tout i transient, se décompose en "petits" systèmes pour les choses fermées récurrentes. Pour chaque "petit" système, π est unique à une constante près et la mesure invariante est en général la combinaison de ces "solutions" avec des constantes.

Théorème 4.9. *Soit P irréductible. Supposons qu'il existe une mesure invariante* π *finie. Alors, si* $\sum_{i\in E} \pi_i = 1$, $\pi_i = \frac{1}{\mathbb{E}_i[T_i]}$.

Remarque. *Attention!!!* On peut avoir une infinité de mesures invariantes!

Exemple 4.4.
1. Si $P = I$, alors toute distribution est stationnaire.
2. $P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. Les classes sont $\{1\}$ non fermée donc transiente (de période $+\infty$), $\{2\}$ (de période 1) et $\{3, 4\}$ (de période 2) fermées donc récurrentes. On résout le système

$$\pi = \pi P \Leftrightarrow \begin{cases} \pi_1 = 0 \\ \pi_2 = 1/2\pi_1 + \pi_2 \\ \pi_3 = 1/2\pi_1 + \pi_4 \\ \pi_4 = \pi_3 \end{cases} \Leftrightarrow \begin{cases} \pi_1 = 0 \\ \pi_3 = \pi_4 \end{cases}$$

donc $\mathcal{S} = \{(0\,1 - 2k\,k\,k); k \in [0, 1/2]\}$.

Exercice 4.5. *Donner toutes les lois de probabilité stationnaires pour les chaîne de Markov suivantes*

1. *la chaîne de Markov de l'exercice 4.1. Supposons de plus que le loi initiale de cette chaîne de Markov est* $(1/3, 1/3, 1/3)$. *Donner sans calcul* $\mathbb{P}(X_{345} = 2)$.

On résout $\pi = \pi P$, où π est un vecteur ligne ! Comme P est irréductible, π est unique à une constante près. On a le système suivant

$$\begin{cases} \pi_1 = \pi_2(1/2) + \pi_3(1/2) \\ \pi_2 = (1/2)\pi_1 + (1/2)\pi_3 \\ \mu_3 = (1/2)\mu_1 + (1/2)\mu_2 \end{cases} \Longleftrightarrow \pi_1 = \pi_2 = \pi_3.$$

Donc $\pi = c(1,1,1)$, $c \in \mathbb{R}$. Pour avoir une loi de probabilité, on prend $c = 1/3$.

La loi initiale proposée est la loi stationnaire. Donc sous cette loi, $\mathbb{P}(X_n = 2) = \pi_2 = 1/3$ pour tout $n \geq$.

2. *la chaîne de Markov de l'exercice 4.2.*
   Les états 1 et 2 sont transients, alors $\pi_1 = \pi_2 = 0$. On a $\pi_3 = c$ avec $c \in \mathbb{R}$, on prend $c = 1$ pour avoir une loi de probabilité.

3. *la chaîne de Markov de matrice P_1 de l'exercice 4.3. Supposons de plus que la loi initiale pour cette chaîne de Markov est* $(1/3, 1/2, 0, 0, 0, 0, 1/6)$. *Donner (sans calcul)* $\mathbb{P}(X_{1000} = 1)$ *et* $\mathbb{P}(X_{2000} = 7)$.
   Il faut résoudre $\pi = \pi P$. Comme c'est un système assez grand, essayons d'être astucieux. Les états 3,4,5,6 sont transients. Alors $\pi_3 = \pi_4 = \pi_5 = \pi_6 = 0$. Il reste deux systèmes d'équations. L'un correspond à la classe $\{1, 2\}$:

$$\begin{cases} \pi_1 = (1/2)\pi_1 + (1/3)\pi_2 \\ \pi_2 = (1/2)\pi_1 + (2/3)\pi_2 \end{cases} \Longleftrightarrow (\pi_1, \pi_2) = c_1(1/3, 1/2).$$

   L'autre système est de la classe $\{7\}$: $\pi_7 = \pi_7$. Donc $\pi = ((1/3)c_1, (1/2)c_1, 0, 0, 0, 0, c_2)$. Les constantes $c_1 \geq 0$ et $c_2 \geq 0$ sont liées par le fait que $c_1(1/3 + 1/2) + c_2 = 1$ pour avoir une loi de probabilité.

   La loi initiale proposée est une loi de probabilité stationnaire avec $c_1 = 1, c_2 = 1/6$. Donc $\mathbb{P}(X_n = i) = \pi_i$ pour tout $n \geq 0$, donc $\mathbb{P}(X_n = 1) = 1/3$, $\mathbb{P}(X_n = 7) = 1/6$ pour tout $n \geq 0$.

4. *la chaîne de Markov de matrice P_2 de l'exercice 4.3.*
   Posons tout de suite $\pi_4 = \pi_6 = 0$ et résolvons deux systèmes : l'un pour $\{1, 2, 3\}$

$$\begin{cases} \pi_1 = 1/2\pi_1 + 1/4\pi_3 \\ \pi_2 = 1/2\pi_1 + 2/3\pi_3 \\ \pi_3 = 1/3\pi_2 + 3/4\pi_3 \end{cases},$$

   l'autre pour $\{5, 7\}$

$$\begin{cases} \pi_5 = 1/5\pi_5 + 2/5\pi_7 \\ \pi_7 = 4/5\pi_5 + 3/5\pi_7 \end{cases},$$

   On a $\pi = (c_1, 3/2c_1, 2c_1, 0, c_2, 0, 2c_2)$ avec $c_1(1 + 3/2 + 2) + c_2(1 + 2) = 1$, $c_1, c_2 \geq 0$.

5. *la chaîne de Markov de matrice P_3 de l'exercice 4.3.*
   $\mu = (c_1, c_2, 8/3c_2, 0, 2c_1)$ avec $c_1, c_2 \geq 0$, $c_1 + c_2 + 8/3c_2 + 2c_1 = 1$.

4.4.3 Convergence vers la mesure invariante

Théorème 4.10. *Toute CM homogène finie irréductible apériodique admet une mesure invariante vers laquelle la chaîne converge en loi quelque soit la distribution initiale, i.e.*

$$\forall \lambda_0, \quad (\mathbb{P}(X_n = j))_{j \in E} \underset{n \to +\infty}{\longrightarrow} \pi,$$

où $\forall i \in E$, $\pi_i > 0$ *et* $\pi_i = \mathbb{E}_i[T_i]^{-1}$ *avec* $T_i = \inf\{n \geq 0 : X_n = i\}$.

Remarque. *Attention !!!* Une chaîne de Markov homogène finie admettant une unique distribution stationnaire ne converge pas nécessairement en loi vers cette mesure.

Exemple 4.5. $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $P^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Chaîne irréductible de période 2. Sa distribution stationnaire est $(1/2 1/2)$ mais la loi de X_n est

- pour $n = 2p$ pair, X_1 et X_0 ont même loi : $P^{2p} = I_2$ et $\lambda_{2p} = \lambda_0$;
- pour $n = 2p+1$ impair, $\lambda_{2p+1} = \lambda_0 P = \lambda_1 \neq \lambda_0$ sauf si $\lambda_0 = (1/2 1/2)$.

Si $\lambda_0 \neq \pi$, on a divergence de la suite $(\lambda_n)_n$ quand $n \to \infty$.

Cas particulier. $P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$, avec $0 < a, b < 1$. C'est une chaîne irréductible apériodique. Par une méthode purement algébrique, on peut calculer P^n et en déduire sa limite quand $n \to \infty$.

Calculons les valeurs propres de P

$$\begin{aligned} \det(P - \lambda I) &= \begin{vmatrix} 1-a-\lambda & a \\ b & 1-b-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & a \\ 1-\lambda & 1-b-\lambda \end{vmatrix} \\ &= (1-\lambda) \begin{vmatrix} 1 & a \\ 1 & 1-b-\lambda \end{vmatrix} = (1-\lambda)(1-a-b-\lambda). \end{aligned}$$

Donc les valeurs propres de P sont donc 1 et $1-a-b$. Le vecteur propre $(1,1)^t$ est associé à la valeur propre 1 et $(-a, b)$ est associé à la valeur propre $1-a-b$. On a donc

$$R = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix}, \quad R^{-1} = \frac{1}{a+b} \begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix},$$

et

$$P^n = RD^nR^{-1} = \begin{pmatrix} \dfrac{a(1-a-b)^n + b}{a+b} & \dfrac{a(1-a-b)^n + a}{a+b} \\[2ex] \dfrac{b - b(1-a-b)^n}{a+b} & \dfrac{b(1-a-b)^n + a}{a+b} \end{pmatrix} \underset{n \to \infty}{\longrightarrow} \begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\[2ex] \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{pmatrix}.$$