MAD M1 Actuariat/ES

Chapitre II: Chaine de Markov

Pierre-Olivier Goffard

Université de Lyon 1
ISFA
pierre-olivier.goffard@univ-lyon1.fr

ISFA
November 5, 2018

I. Premières définitions

Soit $(X_n)_{n\in\mathbb{N}}$ un processus stochastique, définit sur un espace probabilisé $(\Omega,\mathbb{A},\mathbb{P})$, à valeurs dans un espace d'état fini ou dénombrable E.

Exemple 1 (A propos de l'espace d'état)

E peut-être

- $\{a,b,c,d\}$
- $\mathbb{N}$ ou $\mathbb{Z}$

Definition 1 (Chaine de Markov)

Le processus $(X_n)_{n\in\mathbb{N}}$ est une chaine de Markov si

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

pour tout $n\in\mathbb{N}$ et tout $(x_0,x_1,\dots,x_{n+1})\in E^{n+2}$. La valeur futur du processus ne dépend que de l'état à l'instant présent.

Un évènement $(X_0 = x_0, \dots, X_n = x_n)$ est une trajectoire du processus $(X_n)_{n\in\mathbb{N}}$. La vraisemblance d'une telle trajectoire est donnée par

$$\begin{aligned}
\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) &= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0)\mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1}) \\
&= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1})\mathbb{P}(X_n = x_{n-1} | X_{n-2} = x_{n-2}) \\
&\times \mathbb{P}(X_0 = x_0, \dots, X_{n-2} = x_{n-2}) \\
&= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1})\mathbb{P}(X_{n-1} = x_{n-1} | X_{n-2} = x_{n-2}) \\
&\dots \mathbb{P}(X_1 = x_1 | X_0 = x_0)\mathbb{P}(X_0 = x_0) \\
&= Q_n(x_{n-1}, x_n) Q_{n-1}(x_{n-2}, x_{n-1}) \dots Q_1(x_0, x_1)\mu(x_0),
\end{aligned}$$

où

- $\mu(x) := \mathbb{P}(X_0 = x)$ est une loi de probabilité sur E appelée loi initiale de la chaine de Markov $(X_n)_{n\in\mathbb{N}}$.
- $Q_n(x,y) = \mathbb{P}(X_n = y | X_{n-1} = x)$ forme une suite de loi de probabilité sur E telle que

$$\begin{cases} Q_n(x,y) \geq 0, & (x,y) \in \mathbb{E}^2 \\ \sum_{y\in E} Q_n(x,y) = 1, & x \in \mathbb{E} \end{cases}, \text{ pour tout } n \geq 1,$$

  appelée probabilité de transition de $(X_n)_{n\in\mathbb{N}}$

Lorsque les probabilités de transtion dépendent de l'instant considéré alors on parle de chaine de Markov non-hommogène. Dans ce cours, on traitera majoritairement de chaine de Markov homogène.

Definition 2 (Matrice sto et chaine de Markov homogène)

1. Une matrice $(Q(x,y))_{(x,y)\in E^2}$ est dite stochastique si

$$\begin{cases} Q(x,y) \geq 0, & (x,y) \in \mathbb{E}^2 \\ \sum_{y\in E} Q(x,y) = 1, & x \in \mathbb{E} \end{cases}, \text{ pour tout } n \geq 1.$$

2. Une chaine de Markov $(X_n)_{n\in\mathbb{N}}$ est homogène si

$$\mathbb{P}(X_n = y | X_{n-1} = x) = Q(x,y), \text{ pour tout } (n,x,y) \in \mathbb{N} \times E^2.$$

   Les probabilités de transitions sont dites stationnaires, et la matrice $(Q(x,y))_{(x,y)\in E^2}$ est appelée matrice de transtion.

Exemple 2 (Espace d'état fini)

Soit $E = \{1, \ldots, K\}$ l'espace d'état (fini) d'une chaine de Markov homogène $(X_n)_{n\in\mathbb{N}}$. Elle est entièrement déterminée par la donnée de

- sa loi initiale $\mu(.)$
- sa matrice de transition

$$\mathbf{Q} = \begin{pmatrix} Q(1,1) & \ldots & Q(1,K) \\ \vdots & \ddots & \vdots \\ Q(K,1) & \ldots & Q(K,K) \end{pmatrix}$$

On associe souvent à la matrice de transitionn d'une chaine de Markov (homogène, sur un espace d'état fini) un graphe étiqueté et orienté.

Exemple 3 ((La chaine météo))

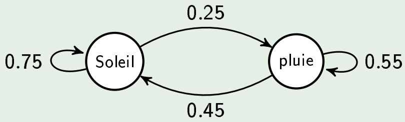
Soit $(X_n)_{n\in\mathbb{N}}$ un processus à valeur dans l'espace d'état {soleil, pluie} indiquant le temps qu'il fait heure par heure. On suppose que

- S'il fait beau maintenant alors on est sur qu'il fait beau l'heure d'après à 75%
- S'il pleut maintenant alors on est sur qu'il pleut l'heure d'après à 45%

La matrice de transition est donnée par

$$\mathbf{Q} = \begin{pmatrix} 0.75 & 0.25 \\ 0.55 & 0.45 \end{pmatrix}$$

et le graphe associé est



Exemple 4 (Pierre-O's life back in California)

Soit une chaine de Markov homogène $(X_n)_{n\in\mathbb{N}}$ indiquant la position de Pierre-O à un instant donnée. Il peut se trouver initialement dans un de ces trois endroits de manière équiprobable

- chez lui (Home)
- au travail (Uni)
- dans l'eau (Surf)

On suppose que d'un instant à l'autre il change d'endroit. De chez lui, il a une chance sur trois d'aller surfer. Du travail, il peut aller surfer ou retourner à la maison de manière équiprobable. Quand il est dans l'eau alors il retourne assurément chez lui.

1. Donner la loi initiale
2. Donner la matrice de transition
3. Donner le graphe

Exemple 5 (La marche aléatoire sur $\mathbb{Z}$)

Soit $(\xi_n)_{n\in\mathbb{N}}$ une suite de variable aléatoire i.i.d. distribuées comm ξ de loi de probabilité

$$\mathbb{P}(\xi = 1) = p, \text{ et } \mathbb{P}(\xi = -1) = 1 - p.$$

Le processus $(X_n)_{n\in\mathbb{N}}$ définit par

$$X_n = x_0 + \sum_{i=1}^{n} \xi_i, \text{ pour } n \geq 1, \text{ avec } x_0 \in \mathbb{Z},$$

est une chaine de Markov homogène de matrice de transition

$$Q(x,y) = \begin{cases} p & \text{si } y = x+1 \\ 1-p & \text{si } y = x-1 \\ 0 & \text{sinon} \end{cases}$$

Proposition 1 (Protocole générateur de chaine de Markov)

Soit $(U_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires i.i.d.,

- *sur un espace mesurable G,*
- *distribuées comme U,*
- *indépendantes de X_0,*

et une application mesurable $F : E \times G \mapsto E$. Le processus $(X_n)_{n\in\mathbb{N}}$ définit par

$$X_{n+1} = F(X_n, U_{n+1})$$

est une chaine de Markov homogène.

preuve:
Soit $n \in \mathbb{N}$ et $x_0, \ldots, x_{n+1} \in E$. Posons $A_n = \{X_0 = x_0, \ldots, X_n = x_n\}$, alors on a

$$\begin{aligned}
\mathbb{P}(X_{n+1} = x_{n+1} | A_n) &= \mathbb{P}[F(X_n, U_{n+1})) = x_{n+1} | A_n] \\
&= \mathbb{P}[F(x_n, U_{n+1})) = x_{n+1} | A_n] \\
&= \mathbb{P}[F(x_n, U_{n+1})) = x_{n+1}] \\
&= \mathbb{P}[F(x_n, U)) = x_{n+1}].
\end{aligned}$$

Le processus $(X_n)_{n\in\mathbb{N}}$ est une chaine de Markov homogène de matrice de transition

$$Q(x,y) = \mathbb{P}(F(x,U) = y), \ \forall (x,y) \in E^2.$$

$\square$

Exemple 6 (Inventaire)

Entre deux instants, une usine fabrique $q \in \mathbb{N}^*$ pièces. Les clients achètes D_{n+1} entre les instants n et $n+1$. On suppose que $(D_n)_{n\in\mathbb{N}}$ forme une suite de variables aléatoires discrètes i.i.d. On note

$$X_{n+1} = max(X_n + q - D_{n+1}, 0) = F(X_n, D_{n+1}), \ n \in \mathbb{N}.$$

le nombre de pièces en stock. $(X_n)_{n\in\mathbb{N}}$ est une chaine de Markov homogène de transition

$$\begin{cases} Q(x,0) = \mathbb{P}(D_1 \geq x+q), \\ Q(x,y) = \mathbb{P}(D_1 = x+q-y), \quad 0 < y < x+q. \end{cases}$$

Lorsque l'espace d'état E est fini, la loi initiale est donnée par un vecteur μ de dimansion Card(E). On désigne par $\mathbf{Q}^n$ la matrice de transition $\mathbf{Q}$ élevée à la puissance n. On note

$$\mu\mathbf{P}^n(y) = \sum_{x\in E} \mu(x)\mathbf{P}^n(x,y)$$

Proposition 2 (Petits calculs)

Soit $(X_n)_{n\in\mathbb{N}}$ *une chaine de Markov homogène, d'espace d'états* E*, de loi initiale* μ *et de matrice de transition* $\mathbf{Q}$*. On a les formules suivantes*

1. *Pour tout* $(x_0,\dots,x_n)\in E^{n+1}$,
$$\mathbb{P}(X_0=x_0,\dots,X_n=x_n)=\mu(x_0)\mathbf{Q}(x_0,x_1)\dots\mathbf{Q}(x_{n-1},x_n)$$
2. *Pour tout* $(x,y)\in E^2$ *et* $p\in\mathbb{N}$,
$$\mathbb{P}(X_{n+p}=y|X_p=x)=\mathbf{Q}^n(x,y)$$
3. $$\mathbb{P}(X_n=y)=\mu\mathbf{Q}^n(y)$$

preuve:

1. $$\begin{array}{rcl}\mathbb{P}(X_0=x_0,\dots,X_n=x_n) & = & \mathbb{P}(X_n=x_n|X_0=x_0,\dots,X_{n-1}=x_{n-1}) \\ & \times & \mathbb{P}(X_0=x_0,\dots,X_{n-1}=x_{n-1}) \\ & = & \mathbb{P}(X_n=x_n|X_{n-1}=x_{n-1})\mathbb{P}(X_{n-1}=x_{n-1}|X_{n-2}=x_{n-2}) \\ & \dots & \mathbb{P}(X_1=x_1|X_0=x_0)\mathbb{P}(X_0=x_0) \\ & = & \mathbf{Q}(x_{n-1},x_n)\mathbf{Q}(x_{n-2},x_{n-1})\dots\mathbf{Q}(x_0,x_1)\mu(x_0)\end{array}$$

on retrouve la formule annoncée après ré-arrangements des termes dans le produit.

2.

$$
\begin{aligned}
\mathbb{P}(X_{n+p} = y | X_p = x) &= \sum_{x_{p+1},\ldots,x_{n+p-1} \in E} \mathbb{P}\Big(X_{n+p} = y, X_{n+p-1} = x_{n+p-1}, \ldots, X_{p+1} = x_{p+1} \\
&\mid \quad X_p = x) \\
&= \sum_{x_{p+1},\ldots,x_{n+p-1} \in E} \mathbf{Q}(x, x_{p+1}) \ldots \mathbf{Q}(x_{n+p-1}, y) \\
&= \mathbf{Q}^n(x,y)
\end{aligned}
$$

3. $\mathbb{P}(X_n = y) = \sum_{x \in E} \mathbb{P}(X_n = y | X_0 = x) \mathbb{P}(|X_0 = x) = \mu \mathbf{Q}^n(y)$.

□

II. Simulation et calculs matriciels en Python

Soit $(X_n)_{n \in \mathbb{N}}$ une chaine de Markov homogène sur E, de matrice de transition $\mathbf{Q}$.

Definition 3 (Etat absorbant)

Un état $x \in E$ est dit absorbant si $Q(x,x) = 1$.

Definition 4 (Loi Catégorielle)

La loi catégorielle est la loi de probabilité d'une variable aléatoire discrete X à valeurs dans $\{1, \ldots, K\}$ valeurs distintes telles que

$$\mathbb{P}(X=1)=p_1, \ldots, \mathbb{P}(X=K)=p_K,$$

avec

$$p_i \geq 0, \text{ pour } i=1, \ldots, K \text{ et } \sum_{i=1}^{K} p_i = 1.$$

On note $X \sim \mathrm{Cat}(p_1, \ldots, p_K)$

L'idée sous-jacente est celle d'un sac rempli de balles pouvant être de K couleurs différentes. Une réalisation de loi catégorielle peut-être vu comme un vecteur $(X_1, \ldots, X_K)$ contenant des 0 et un unique 1 associé à la couleur tirée. La répétition de l'expérience (tirage aléatoire avec remise dans le sac) conduit à la définition de la loi multinomiale.

Definition 5 (Loi Multinomiale)

La loi multinomiale est la loi de probabilité d'une vecteur aléatoire discret $X = (X_1, \ldots, X_K)$ à valeurs dans $\{0, 1, \ldots, n\}^k$ telles que

$$\mathbb{P}(X_1 = x_1, \ldots, X_K = x_K) = \frac{n!}{x_1! \ldots x_K!} p_1^{x_1} \ldots p_K^{x_K},$$

où $\sum_{i=1}^{K} x_i = n$ et

$$p_i \geq 0, \text{ pour } i = 1, \ldots, K \text{ et } \sum_{i=1}^{K} p_i = 1.$$

On note $X \sim \text{Multinom}(n, p_1, \ldots, p_k)$.

Exemple 7 (Chaine de Markov à 5 états)

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène sur $E=\{1,2,3,4,5\}$, de matrice de transition

$$\mathbf{Q}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/8 & 1/4 & 0 & 1/8 \\ 1/4 & 1/2 & 1/8 & 1/8 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

et de loi initial $\mu=(0,1/3,1/3,1/3,0)$.

1. Cette chaine de Markov admet-elle des états absorbant? Si oui, lesquels?
2. Donner la loi de probabilité de X_4, c'est à dire $\mathbb{P}(X_4=x)$, $\forall x\in E$
3. On note par A l'ensemble des états absorbants et

$$\tau_A=\inf\{n\in\mathbb{N}\ ;\ X_n\in A\}$$

   le temps d'absorbtion de la chaine. Calculer $\mathbb{E}(\tau_A)$ par simulation.
4. Quel est l'état d'absorbtion le plus probable?

1. Etats accessibles, communiquants, classes ouvertes/fermées, chaine irréductible

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène d'espace d'état E et de matrice de transition $\mathbf{Q}$.

Definition 6 ($x \rightsquigarrow y$)

L'état $y \in E$ est accessible depuis l'état $x \in E$ s'il existe $n \in \mathbb{N}$ tel que

$$Q^n(x,y) > 0.$$

Proposition 3 (La relation d'accessibilité)

Soient $x,y,z \in E$,

- $x \rightsquigarrow x$,
- *Si $x \rightsquigarrow y$ et $y \rightsquigarrow z$ alors $x \rightsquigarrow z$.*

Sur le graph, cela signifie qu'il existe un chemin de x à y.

Definition 7 (Etats communiquants)

Soient $x,y \in E$. Si $x \rightsquigarrow y$ et $y \rightsquigarrow x$ alors x et y communiquent. On note $x \leftrightarrow y$.

Definition 8 (Classes d'équivalence)

Une classe d'équivalence est un sous-ensemble $\mathscr{C} \subset E$ formé d'états communiquant. Formellement

$$\mathscr{C} = \{x, y \in E \ ; \ x \leftrightarrow y\}$$

Remarque 1

1. On parle de classe d'équivalence car la relation de communication est symérique, réflexive, et transitive.
2. l'ensemble des classes d'équivalence forme une partition de l'espace d'état. Ce sont les classes irréductibles de la chaine de Markov.

Definition 9 (Irréductibilité)

La chaine de Markov $(X_n)_{n\in\mathbb{N}}$ est irréductible si pour tout $(x,y) \in E^2$, il existe $n \in \mathbb{N}$ tel que

$$Q^n(x,y) = \mathbb{P}(X_n = y | X_0 = x)$$

En remarquant que

$$Q^n(x,y) \sum_{x_1,\dots,x_{n-1}} Q(x,x_1)\dots Q(x_{n-1},y),$$

on constate l'existence d'une trajectoire de probabilité d'occurence non nulle avec l'existence de $x_i, i = 0,\dots,n-1$ tel que $Q(x_i, x_i + 1) > 0$ (en convenant que $x_0 = x$ et $x_n = y$).

Remarque 2

Dans une chaine irréductible tout les états communiquent. L'espace d'état est une classe d'équivalence.

Definition 10 (Classe d'équivalence ouverte/fermé)

Une classe (d'équivalence) $\mathscr{C}$ est

- ouverte si pour tout $x \in \mathscr{C}$ il existe $y \in E/\mathscr{C}$ et $n \in \mathbb{N}$ tels que $Q^n(x,y) > 0$,
- fermée sinon.

Exemple 8

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène d'espace d'état $\{1,2,3,4\}$ et de matrice de transition

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Combien de classe d'équivalence?
2. $(X_n)_{n\in\mathbb{N}}$ est-elle une chaine irréductible?
3. Distinguer les classes ouvertes et fermées.

Remarque 3

On peut définir une sous-chaine de Markov irréductible à partir d'une classe fermée.

Definition 11 (La périodicité d'un état)

La période d'un état $x \in E$ est défini par

$$\begin{aligned} d(x) &= \text{pgcd}\{n \in \mathbb{N}^* \ ; \ Q^n(x,x)\} \\ &= \text{pgcd}\{\text{Longueur des trajectoires menant de } x \text{ à } x\} \end{aligned}$$

Si $d(x) = 1$ alors x est un état apériodique.

Remarque 4

1. Si $Q(x,x) > 0$ alors x est apériodique.
2. Si $x \leftrightarrow y$ alors $d(x) = d(y)$
3. Soit $\mathscr{C}$ une classe d'équivalence, alors la période d'un état $x \in \mathscr{C}$ est la période de la classe d'équivalence $\mathscr{C}$.

Exemple 9

Soit $\{X_n\}_{n\in\mathbb{N}}$ une chaine de Markov homogène $\{1,2,3,4,5\}$ de matrice de transition

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

1. Chaine irréductible? Combien de classes d'équivalence? Ouvertes ou fermées?
2. Donner la période de chaque état?

2. Etats récurrents, états transients

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène d'espace d'état E. On notera la probabilité et l'espérance conditionnelle de X_n sachant $X_0 = x$ de la façon suivante

$$\mathbb{P}(X_n = y | X_n = x) = \mathbb{P}_x(X_n = y),\ y \in E,\ \text{et}\ \mathbb{E}(X_n | X_n = x) = \mathbb{E}_x(X_n).$$

Definition 12 (T_x, S_x, et N_x)

Soit $x \in E$, on note

1. $T_x = \inf\{n \in \mathbb{N}\ ;\ X_n = x\}$ est le temps d'atteinte de l'état x,
2. $S_x = \inf\{n \geq 1\ ;\ X_n = x\}$ est le temps de retour à l'état x (en supposant que $X_0 = x$),
3. $N_x = \sum_{n=1}^{+\infty} \mathbb{I}_{\{X_n = x\}}$ le nombre de passagge par l'état 0 (sans compter le point de départ).

Definition 13 (Etat récurrent/transient)

1. Un état $x \in E$ est récurrent si
$$\mathbb{P}_x(S_x < \infty) = 1$$
2. Un état $x \in E$ est transient
$$\mathbb{P}_x(S_x = \infty) > 0$$

Un état est transient lorsqu'il n'est pas récurrent.

Si $x \in E$ est un état récurrent, alors on est sur que la chaine repassera par x. Si x est transitoire alors il existe un évènement, de probabilité non nulle, pour lequel la chaine ne repasse pas par x.

Theoreme 1 (Charactérisation de la récurrence et de la transience)

Les assertions suivantes sont équivalentes

1. $x \in E$ *est récurrent* $\Leftrightarrow \mathbb{P}_x(N_x = \infty) = 1$ *et* $\sum_{n=1}^{+\infty} Q^n(x,x) = \infty$
2. $x \in E$ *est transient* $\Leftrightarrow \mathbb{P}_x(N_x < \infty) = 1$ *et* $\sum_{n=1}^{+\infty} Q^n(x,x) < \infty$. *De plus, on a*

$$\mathbb{P}_x(N_x = k) = \mathbb{P}(S_x = \infty)\left[1 - \mathbb{P}(S_x = \infty)\right]^{k-1}, k = 1, 2\ldots, \text{ (Distribution géométrique)}.$$

preuve:

On commence par remarquer que pour $x, y \in E$,

$$\mathbb{E}_y(N_x) = \mathbb{E}_y\left(\sum_{n=1}^{+\infty} \mathbb{I}_{\{X_n = x\}}\right) = \sum_{n=1}^{+\infty} \mathbb{E}_y(\mathbb{I}_{\{X_n = x\}}) = \sum_{n=1}^{+\infty} \mathbb{P}_y(X_n = x) = \sum_{n=1}^{+\infty} Q^n(y,x).$$

Posons $F_n = \{X_n = x,\ X_m \neq x\ ;\ m \geq n+1\}$ et $G = \{X_n \neq x\ ;\ n \geq 1\}$, on a

$$\{N_x < \infty\} = G \cup \bigcup_{n=1}^{+\infty} F_n.$$

Comme les évènements ci-dessus sont disjoints alors

$$\mathbb{P}_x(N_x < \infty) = \mathbb{P}_x(G) + \sum_{n=1}^{+\infty} \mathbb{P}_x(F_n). \tag{1}$$

Or

$$\begin{aligned}
\mathbb{P}_x(F_n) &= \mathbb{P}_x(X_n = x,\ X_m \neq x\ ;\ m \geq n+1) \\
&= \mathbb{P}(X_n = x,\ X_m \neq x\ ;\ m \geq n+1 | X_0 = x) \\
&= \mathbb{P}(X_m \neq x\ ;\ m \geq n+1 | X_n = x, X_0 = x)\mathbb{P}(X_n = x | X_0 = x) \\
&= \mathbb{P}(X_m \neq x\ ;\ m \geq n+1 | X_n = x)\mathbb{P}_x(X_n = x) \text{ (propriété de Markov)} \\
&= \mathbb{P}_x(X_n \neq x\ ;\ n \geq 1)\mathbb{P}_x(X_n = x) \text{ (Homogénéité)} \\
&= \mathbb{P}_x(G) Q^n(x,x)
\end{aligned}$$

En substituant dans l'équation (1), il vient

$$\mathbb{P}_x(N_x < \infty) = \mathbb{P}_x(G)\left[1 + \sum_{n=1}^{+\infty} Q^n(x,x)\right] = \mathbb{P}_x(G)\left[1 + \mathbb{E}_x(N_x)\right]$$

- Si x est récurrent alors $\mathbb{P}_x(G) = 0$, par suite $\mathbb{P}_x(N_x < \infty) = 0$, $\mathbb{P}_x(N_x = \infty)$. Cela implique $\mathbb{E}_x(N_x) = \sum_{n=1}^{+\infty} Q^n(x,x) = \infty$.
- Si x est transient alors $\mathbb{P}_x(G) > 0$ et

$$\mathbb{E}_x(N_x) = \frac{\mathbb{P}_x(N_x < \infty)}{\mathbb{P}_x(G) > 0} - 1 < \infty$$

ce qui implique que $\sum_{n=1}^{+\infty} Q^n(x,x) < \infty$ et $\mathbb{P}_x(N_x < \infty) = 1$

Supposons que x soit transient. Pour montrer que N_x suit une loi géométrique, nous avons besoin du lemme suivant

Lemme 1

$$\mathbb{P}_x(N_x = 0) = \mathbb{P}(S_x = \infty) \ \textit{et}\ \mathbb{P}_x(N_x = k | S_x < \infty) = \mathbb{P}_x(N_x = k-1),\ \textit{pour}\ k \geq 1.$$

preuve:

L'égalité $\mathbb{P}_x(N_x = 0) = \mathbb{P}(S_x = \infty)$ est évidente au vu de la définition de S_x. Nous avons

$$\begin{aligned}
\mathbb{P}_x(N_x = k | S_x < \infty) &= \mathbb{P}_x(N_x = k | S_x < \infty, X_0 = x) \\
&= \mathbb{P}(\sum_{k=1}^{+\infty} = k | S_x < \infty, X_{S_x} = x, X_0 = x) \\
&= \mathbb{P}(1 + \sum_{k=S_x+1}^{+\infty} = k | S_x < \infty, X_{S_x} = x) \\
&= \mathbb{P}(1 + \sum_{k=S_x+1}^{+\infty} = k | S_x < \infty, X_{S_x} = x) \\
&= \mathbb{P}(N_x = k-1)
\end{aligned}$$

$\square$

Par application du lemme, il vient

$$\begin{aligned}
\mathbb{P}_x(N_x = k) &= \mathbb{P}_x(N_x = k | S_x < \infty)\mathbb{P}_x(S_x < \infty) \\
&= \mathbb{P}_x(N_x = k-1)\mathbb{P}_x(S_x < \infty) \\
&= \mathbb{P}_x(N_x = k-1 | S_x < \infty)\mathbb{P}_x(S_x < \infty)^2 \\
&= \mathbb{P}_x(N_x = k-2)\mathbb{P}_x(S_x < \infty)^2 \\
&= \ldots \\
&= \mathbb{P}_x(N_x = 0)\mathbb{P}_x(S_x < \infty)^k. \\
&= \mathbb{P}_x(S_x = \infty)\mathbb{P}_x(S_x < \infty)^k.
\end{aligned}$$

Remarque 5

Une chaine de Markov irréductible, à espaces d'états fini, est récurrente au sens où tout les états sont récurrents.

Les états d'une classe ouverte sont transients, ceux d'une classe fermée sont récurrents.

Definition 14 (Filtration)

Soit $(\Omega, \mathscr{F}, \mathbb{P})$ un espace probabilisé. Une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ est une suite de sous-tribus de $\mathscr{F}$ telle que $\mathscr{F}_n \subset \mathscr{F}_{n+1}$

Exemple 10

La filtration $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ est la filtration naturelle de $(X_n)_{n\in\mathbb{N}}$. Ce sont les tribus engendré par toutes les trajectoires passées de $(X_n)_{n\in\mathbb{N}}$.

Definition 15 (Temps d'arrêt)

Un temps d'arrêt τ est une variable aléatoire telle que $\{\tau \leq n\} \subset \mathscr{F}_n$ pour tout $n\mathbb{N}$

Exemple 11

T_x et S_x sont des temps d'arrêts.

Theoreme 2 (Propriété de Markov forte)

Soit $(X_n)_{n\in N}$ une chaine de Markov homogène et τ un $\mathscr{F}_n$-temps d'arrêt alors, conditionellement à l'évènement $\{\tau < \infty\}$, le processus $(\widehat{X}_n = X_{\tau+n})_{n\in\mathbb{N}}$ est une chaine de Markov homogène de loi initiale $\widehat{\mu} \overset{\mathscr{D}}{=} X_\tau$ et de même probabilités de transition que $(X_n)_{n\in\mathbb{N}}$

preuve:

Soit $p \in \mathbb{N}^*$ et $(x, x_1, \dots, x_p) \in E^{p+1}$. On a

$$
\begin{aligned}
& \mathbb{P}(X_{\tau+1} = x_1, \dots, X_{\tau+p} = x_p | X_\tau = x, \tau < \infty) \\
= & \sum_{k=0}^{\infty} \mathbb{P}(X_{\tau+1} = x_1, \dots, X_{\tau+p} = x_p, \tau = k | X_\tau = x, \tau < \infty) \\
= & \sum_{k=0}^{\infty} \mathbb{P}(X_{k+1} = x_1, \dots, X_{k+p} = x_p, \tau = k | X_\tau = x, \tau < \infty) \\
= & \sum_{k=0}^{\infty} \mathbb{P}(X_{k+1} = x_1, \dots, X_{k+p} = x_p | \tau = k, X_\tau = x, \tau < \infty) \mathbb{P}(\tau = k | X_\tau = x, \tau < \infty) \\
= & \sum_{k=0}^{\infty} \mathbb{P}(X_{k+1} = x_1, \dots, X_{k+p} = x_p | \tau = k, X_\tau = x, \tau < \infty) \mathbb{P}(\tau = k | X_\tau = x, \tau < \infty) \\
= & \sum_{k=0}^{\infty} \mathbb{P}(X_{k+1} = x_1, \dots, X_{k+p} = x_p | \tau = k, X_k = x, \tau < \infty) \mathbb{P}(\tau = k | X_\tau = x, \tau < \infty) \\
= & \mathbb{P}_x(X_{k+1} = x_1, \dots, X_{k+p} = x_p) \sum_{k=0}^{\infty} \mathbb{P}(\tau = k | X_\tau = x, \tau < \infty) \\
= & \mathbb{P}_x(X_{k+1} = x_1, \dots, X_{k+p} = x_p)
\end{aligned}
$$

$\square$

Proposition 4

Soit $\tau_A = \inf\{n \geq 0 \ ; \ X_n \in A\}$, avec $A \subset E$ tel que $\tau_A < \infty$ presque-sûrement. Soit $x \in E$, alors

1. $$\mathbb{P}_x(\tau_A = n) = \begin{cases} 0, & x \in A, \\ \sum_{y \in E} Q(x,y)\mathbb{P}_y(\tau_A = n-1), & x \notin A. \end{cases}$$

2. $$\mathbb{E}_x(\tau_A) = \begin{cases} 0, & x \in A, \\ 1 + \sum_{y \in E} Q(x,y)\mathbb{E}_y(\tau_A), & x \notin A. \end{cases}$$

Exemple 12

On lance une pièce jusqu'à obtenir face deux fois de suite. On note $(X_n)_{n \in \mathbb{N}}$ le nombre de face consécutifs.

1. Donner la matrice de transition et le graph.
2. Donner le nombre moyen de lancer nécessaires à l'obtention de deux faces consécutifs.

Exemple 13

Soit $\{X_t \ ; \ t \in \mathbb{N}\}$ une chaine de Markov homogène d'espace d'état $\{-1,0,2,4\}$ et de matrice de transition,

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/10 & 3/10 & 5/10 & 1/10 \\ 2/10 & 1/10 & 6/10 & 1/10 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Soit $A = \{-1,4\}$, donner $\mathbb{E}_x(\tau_A)$ pour $x \in E$
2. Define F='Absorption dans l'état 4' and G="L'état 2 est visité juste avant l'absorption" Calculer $\mathbb{P}_x(E)$ et $\mathbb{P}_x(F)$ pour $x \in \{-1,0,2,4\}$.

Soit une chaine de Markov $(X_n)_{n\in\mathbb{N}}$ homogène et irréductibe sur un espace d'état E.

Proposition 5 (Temps moyen passé en x)

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbb{I}_{X_k=x}=\frac{1}{\mathbb{E}_x(S_x)}\ \ p.s.$$

On distingue deux cas

- *Soit $E_x(S_x)<\infty,\ \forall x\in E$ alors la chaine est récurrente positive*
- *Soit $E_x(S_x)=\infty,\ \forall x\in E$ alors la chaine est transiente ou récurrente nulle*

preuve

Posons $m(n)=\sum_{k=1}^{n}\mathbb{I}_{X_k=x}$. Si la chaine est transiente alors $\mathbb{P}(S_x=\infty)>0,\ \forall x\in E$ et donc $\mathbb{E}_x(S_x)=\infty$. De plus on remarque que $\forall\omega\in\Omega\ m(n)[\omega]$ (on comprend Ω comme l'ensemble des trajectoires possibles du processus) devient constant pour n suffisamment grand (comme x est transient alors au bout d'un moment la chaine ne visite plus x), cela implique que $m(n)/n \underset{n\to\infty}{\to} 0$ p.s.

Supposons la chaine récurrente, on introduit la suite de variables aléatoires

$$S_k=\inf\{n>S_k-1\ ;\ X_n=x\},\ \ k\geq 1$$

qui représente les temps de k^{eme} retour en x. On convient que $S_0=0$ et $S_1=S_x$. Soit $\delta_k^S=S_k-S_{k-1},\ k\geq 1$ les temps séparant deux passages par x consécutifs.

Lemme 2

$(\Delta_n^S)_{n\geq 1}$ est une i.i.d. de variables aléatoires distribuées comme $S_x|X_0 = x$

preuve: On va se contenter de montrer le résultat pour les deux premiers délais. On a pour $p, q \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}_x(S_1 = p, \Delta_2^S = q) &= \mathbb{P}(\Delta_2^S = q | S_1 = p, X_0 = x)\mathbb{P}_x(S_1 = p) \\
&= \mathbb{P}(X_{p+q} = x, \bigcap_{k=p+1}^{p+q-1}\{X_k \neq x\} | X_p = x, \bigcap_{k=1}^{p-1}\{X_k \neq x\}, X_0 = x)\mathbb{P}_x(S_1 = p) \\
&= \mathbb{P}(X_{p+q} = x, \bigcap_{k=p+1}^{p+q-1}\{X_k \neq x\} | X_p = x)\mathbb{P}_x(S_1 = p) \\
&= \mathbb{P}(X_q = x, \bigcap_{k=1}^{q-1}\{X_k \neq x\} | X_0 = x)\mathbb{P}_x(S_1 = p) \\
&= \mathbb{P}_x(S_1 = q)\mathbb{P}_x(S_1 = p) = \mathbb{P}_x(S_x = q)\mathbb{P}_x(S_x = p)
\end{aligned}$$

$\square$

En vertu de le loi forte des grands nombres, il vient

$$\frac{S_n}{n} = \frac{\sum_{k=1}^{n} \Delta_k^S}{n} \to \mathbb{E}_x(S_x) \text{ p.s.}$$

Comme x est récurrent alors $m(n) \to \infty$ et

$$\frac{S_m(n)}{m(n)} \to \mathbb{E}(S_x)$$

Soit $n \in \mathbb{N}$ tel que $m(n)[\omega] = l$, on a $S_l[\omega] \leq n \leq S_{l+1}[\omega]$ (On retourne l fois en x avant n puis la $l+1^{me}$ visite a lieu après n). Par suite,

$$\frac{S_{m(n)}}{m(n)} \leq \frac{n}{m(n)} \leq \frac{S_{m(n)+1}}{m(n)} = \frac{S_{m(n)+1}}{m(n)+1} \frac{m(n)+1}{m(n)}$$

Théorème des gendarmes $\Rightarrow$ $\frac{n}{m(n)} \to \mathbb{E}_x(S_x)$.
Vérifions que si $\mathbb{E}_x(S_x) = \infty$ alors $\mathbb{E}_y(S_y) = \infty$, $\forall y \in E$.
Notons que

$$\lim_{n \to +\infty} \mathbb{E}_y \left(\frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{X_k = y} \right) = \lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} Q^k(y, y)$$

d'une part et

$$\lim_{n \to +\infty} \mathbb{E}_y \left(\frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{X_k = y} \right) = \frac{1}{\mathbb{E}_y(S_y)}$$

en vertu du théorème de convergence dominée et le résultat de convergence précédent d'autre part. L'irréductibilité de la chaine entraine l'existence de $r, s \geq 1$ tels que $Q^r(x,y) > 0$ et $Q^s(y,x) > 0$. On remarque que $Q^{r+k+s}(x,x) \geq Q^r(x,y)Q^k(y,y)Q^s(y,x)$ puis

$$\frac{1}{n}\sum_{k=1}^{n} Q^k(y,y) \leq \frac{1}{nQ^r(x,y)Q^s(y,x)}\sum_{k=1}^{n} Q^{r+k+s}(x,x) \to 0.$$

On en déduit que $\mathbb{E}_y(S_y) = \infty$. On peut montrer de manière équivalente que $E_x(S_x) < \infty \Rightarrow E_y(S_y) < \infty,\ y \in E$

$\square$

Remarque 6

$$\frac{1}{n}\sum_{k=1}^{n} Q^k(x,x) \to \begin{cases} +\infty, & \text{récurrence nulle,} \\ 1/\mathbb{E}_x(S_x), & \text{récurrence positive.} \end{cases}$$

2. Loi invariante et stabilisation de la chaine de Markov

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov irréductible sur un espace d'état E, de matrice de transition $\mathbf{Q}$.

Definition 16 (Loi invariante)

Soit $\mathbf{Q}$ la matrice de transition sur E. Une probabilité π est invariante si $\pi\mathbf{Q} = \pi$.

Remarque 7

1. La notion de probabilité invariante de probabilité invariante est lié à la notion de stationarité (comportement de la chaine de Markov à horizon de temps long). En effet, si la loi initiale vérifie $\mu = \pi$ alors $\mu\mathbf{Q}^n = \mu$. La loi de X_n es indépendante de n.
2. Une mesure invariante π sur E n'est pas nécessairement unique, en effet $\lambda\pi$, pour $\lambda \in \mathbb{R}$ est aussi invariante. Si $0 < \sum_{x\in E}\pi(x) < \infty$, on normalise pour obtenir une mesure de probabilité avec
$$\frac{\pi(x)}{\sum_{x\in E}\pi(x) < \infty}$$

Definition 17 (Chaine de Markov réversible)

$(X_{n\in\mathbb{N}})$ est dite réversible si

$$(X_1,\ldots,X_T) \stackrel{\mathscr{D}}{=} (X_T,\ldots,X_1),\ \forall T \in \mathbb{N}.$$

Soit λ une mesure de probabilité sur E

Definition 18 (Chaine de Markov réversible)

λ réversible si

$$\lambda(x)Q(x,y) = \lambda(y)Q(y,x),\ \forall x,y \in E.$$

Lemme 3

$$\lambda \textit{ est réversible } \Rightarrow \lambda \textit{ est invariante }.$$

Theoreme 3

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov irréductible.

(i) *Transiente ou récurrente nulle $\Rightarrow$ Pas de probabilité invariante*

(ii) *Récurrente positive $\Rightarrow$ Probabilité invariante unique telle que $\pi(x) = 1/\mathbb{E}_x(S_x),\ \forall x \in E$*

preuve: D'après la Proposition 5 et la convergence dominée, on a

$$\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=1}^{n}\mathbb{1}_{X_k=x}\right)=\frac{1}{n}\sum_{k=1}^{n}\mathbb{P}_\mu(X_k=x)=\frac{1}{n}\sum_{k=1}^{n}\mu Q^k(x)\to\frac{1}{\mathbb{E}_x(S_x)}:=\pi(x)$$

pour toute loi initiale μ et en particulier $\forall x,y\in E$

$$\frac{1}{n}\sum_{k=1}^{n}\mu Q^k(y,x)\to\pi(x).$$

(i) Supposons la chaine transiente ou récurrent nulle. Supposons l'existence d'une probabilité variable μ alors on a $\mathbb{P}_\mu(X_k=x)=\mu(x)$, $\forall x\in E$ et d'autre part

$$\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=1}^{n}\mathbb{1}_{X_k=x}\right)=\frac{1}{n}\sum_{k=1}^{n}\mathbb{P}_\mu(X_k=x)\to 0.$$

On en déduit que nécéssairement $\mu(x)=0$, $\forall x\in E$, ce qui est impossible pour une probabilité.

(ii) Supposons que la chaine soit récurrente positive. Par les mêmes arguments que précédement, on a que s'il existe μ probabilité invariante alors $\mu = \pi$ ce qui implique son unicité. Nous devons vérifié que π est bien une probabilité invariante, pour l'instant nous savons simplement que $\pi(x)$, $x \in E$. On a, par le lemme de Fatou, l'inégalité

$$\begin{aligned}
\sum_{x \in E} \pi(x) &= \sum_{x \in E} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu Q^k(x) \\
&\leq \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{x \in E} \mu Q^k(x) \\
&= 1
\end{aligned}$$

et par les mêmes arguments

$$\begin{aligned}
\sum_{x \in E} \pi(x) Q(x,y) &= \sum_{x \in E} Q(x,y) \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu Q^k(x) \\
&\leq \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu Q^{k+1}(y) \\
&= \mu(y)
\end{aligned}$$

Or nous savons que

$$\begin{aligned}
\sum_{y \in E} \sum_{x \in E} \pi(x) Q(x,y) &= \sum_{x \in E} \pi(x) \sum_{y \in E} Q(x,y) \\
&= \sum_{x \in E} \pi(x)
\end{aligned}$$

ce qui permet de conclure que $\pi(x)Q(x,y) = \pi(y)$ car $\sum_{y\in E}\pi(y) < \infty$. π est bien une mesure invariante, est-elle une probabilité?

$$\begin{aligned}
\pi(x) &= \frac{1}{n}\sum_{k=1}^{n}\pi Q^n(x) \\
&= \sum_{y\in E}\pi(y)\frac{1}{n}\sum Q^n(y,x) \\
&\underset{n\to\infty}{\longrightarrow} \pi(x)\sum_{y\in E}\pi(y) \text{ (Convergence dominée)}
\end{aligned}$$

ce qui implique que $\sum_{y\in E}\pi(y) = 1$

Remarque 8

Nous venons de montrer que pour une chaine irréductible, récurrente positive, la suite $(\pi Q^n(x))$ converge en moyenne de Césarro vers $\pi(x)$ avec

$$\frac{\mu Q(x)+\mu Q^2(x)+\ldots+\mu Q^n(x)}{n} \underset{n\to\infty}{\to} \pi(x).$$

Mais que dire de la convergence de $\mu Q^n(x)=\mathbb{P}_\mu(X_n=x)$ au sens usuel? La chaine de Markov $(X_n)_{n\in\mathbb{N}}$ converge-t-elle en loi vers une variable aléatoire X_∞? Cette convergence est établie sous réserve d'apériodicité.

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov irréductible, récurrente positive.

Proposition 6

Si $(X_n)_{n\in\mathbb{N}}$ est apériodique alors il existe $N\in\mathbb{N}$ tel que pour tout $n\geq N$ on a

$$Q^n(x,y)>0 \text{ pour tout } x,y\in E.$$

preuve:
Soit $I(x)=\{n\in\mathbb{N}^* \ ; \ Q^n(x,x)>0\}$. On note que si $p,q\in I(x)$ alors $p+q\in I(x)$.

Lemme 4

Soit un ensemble $A\subset\mathbb{N}^*$, stable par addition de pgcd égal à 1 alors A contient tout les entiers plus grand que N.

preuve:
Soit $A' = A \cup \{0\}$, alors $A' - A'$ est un sous groupe de $\mathbb{Z}$ qui est donc de la forme $d\mathbb{Z}$, avec d le plus petit élément non nul de A'. Comme A' contient A alors d divise tout les élements de A donc $d = 1$. On en déduit qu'il existe $a, b \in A'$ tels que $a + 1 = b$. Nécécessairement, $a \in A$,

- si $b = 0$ alors $a = 1$ et A continent tout les entiers naturels
- Si $b \in A$ alors $N = b^2$ convient car pour $n \geq N$, on a pour $0 \leq r < b$,

$$n = b^2 + bq + r = b(q + 1 - r) + (b + 1)r \in A$$

□

$I(x) \subset \mathbb{N}$ est stable par addition, il existe $n_1 \in \mathbb{N}$ tel que $I(x)$ contient tout les entiers supérieurs à n_1 et $Q^n(x,x) > 0$ pour tout $n \geq n_1$. Par irréductibilité, on a aussi pour tout $z, y \in E$ l'existence de $n_2, n_3 \in \mathbb{N}$ tels que $Q^{n_2}(z,x) > 0$ et $Q^{n_3}(x,y) > 0$. On en déduit que

$$Q^{n_2+n+n_3}(z,y) \geq Q^{n_2}(z,x)Q^n(x,x)Q^{n_3}(x,y) > 0$$

On choisit alors $N = n_1 + n_2 + n_3$.

Definition 19 (Chaine de Markov couplée)

Soit $(X'_n)_{n\in\mathbb{N}}$ une copie indépendante de $(X_n)_{n\in\mathbb{N}}$, on appelle $((X_n,X'_n))_{n\in\mathbb{N}}$ la chaine couplée.

- (X_n,X'_n) est une chaine de Markov sur $E\times E$ de matrice de transition $\overline{Q}[(x,x'),(y,y')] = Q(x,y)Q(x',y')$
- Soit μ et ν les lois initiales respectives de $(X_n)_{n\in\mathbb{N}}$ et $(X'_n)_{n\in\mathbb{N}}$ alors la loi initiale de $((X_n,X'_n))_{n\in\mathbb{N}}$ est la mesure produit $\mu\otimes\nu$

Proposition 7

Si $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov irréductible, récurrente positive et apériodique alors la chaine couplée $((X_n,X'_n))_{n\in\mathbb{N}}$ est irréductible et récurrente positive.

preuve:
On vérifie que, il existe $N\in\mathbb{N}$ telle que $\forall n\geq N$

$$\overline{Q}^n[(x,x'),(y,y')] = Q^n(x,y)Q^n(x',y') > 0$$

d'après la Proposition 6. $((X_n,X'_n))_{n\in\mathbb{N}}$ est donc irréductible. On note ensuite que

$$\frac{1}{n}\sum_{k=1}^{n}\overline{Q}^n[(x,x'),(y,y')] = \frac{1}{n}\sum_{k=1}^{n}Q^n(x,y)Q^n(x',y') \leq \frac{1}{n}\sum_{k=1}^{n}Q^n(x,y) < \infty$$

et donc que $((X_n,X'_n))_{n\in\mathbb{N}}$ est récurrente positive.

Theoreme 4

Si $(X_n)_{n\in\mathbb{N}}$ est apériodique alors la chaine se stabilise au sens où, pour toutes lois initiales μ,ν sur E, on a

$$\sup_{y\in E}|\mu Q^n(y)-\nu Q^n(y)| \underset{n\longrightarrow\infty}{\rightarrow} 0.$$

et en particulier

$$\mu Q^n(y) \underset{n\longrightarrow\infty}{\rightarrow} \pi(y).$$

preuve:

Soit $((X_n,X'_n))_{n\in\mathbb{N}}$ la chaine couplée de $(X_n)_{n\in\mathbb{N}}$ de loi initiale $\mu\otimes\nu$ et $\tau_x=\inf\{n\in\mathbb{N}^*\ ;\ (X_n,X'_n)=(x,x)\}$. On a

$$\begin{aligned}
\mu Q^n(y) &= \mathbb{P}_{\mu\otimes\nu}(X_n=y)\\
&= \mathbb{P}_{\mu\otimes\nu}(X_n=y,\tau_x>n)+\sum_{k=1}^{n}\mathbb{P}_{\mu\otimes\nu}(X_n=y,\tau_x=k)\\
&= \mathbb{P}_{\mu\otimes\nu}(X_n=y,\tau_x>n)+\sum_{k=1}^{n}\mathbb{P}_{\mu\otimes\nu}(X_n=y,X_k=x,\tau_x=k)\\
&= \mathbb{P}_{\mu\otimes\nu}(X_n=y,\tau_x>n)+\sum_{k=1}^{n}\mathbb{P}_{\mu\otimes\nu}(X_n=y|X_k=x)\mathbb{P}_{\mu\otimes\nu}(\tau_x=k)\\
&= \mathbb{P}_{\mu\otimes\nu}(X_n=y,\tau_x>n)+\sum_{k=1}^{n}\mathbb{P}_{x}(X_{n-k}=y)\mathbb{P}_{\mu\otimes\nu}(\tau_x=k)
\end{aligned}$$

d'une part et

$$\nu Q^n(y) = \mathbb{P}_{\mu\otimes\nu}(X'_n = y, \tau_x > n) + \sum_{k=1}^{n} \mathbb{P}_x(X'_n = y)\mathbb{P}_{\mu\otimes\nu}(\tau_x = k)$$

d'autre part. On en déduit que

$$|\mu Q^n(y) - \nu Q^n(y)| = [\mathbb{P}_{\mu\otimes\nu}(X_n = y|\tau_x > n) - \mathbb{P}_{\mu\otimes\nu}(X'_n = y|\tau_x > n)]\mathbb{P}_{\mu\otimes\nu}(\tau_x > n) \leq K\mathbb{P}_{\mu\otimes\nu}(\tau_x > n)$$

avec K une constante. On constate que

$$\mu Q^n(y) \underset{n\longrightarrow\infty}{\rightarrow} \pi(y).$$

en choisissant $\nu = \pi$.

Proposition 8 (Loi stationnaire et espace d'état fini)

Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène dont l'espace d'état est fini.

- *Il existe au moins une loi stationaire*

Si la chaine $(X_n)_{n\in\mathbb{N}}$ est irréductible alors cette loi est unique.

Idée de la preuve:
On montre que 1 est une valeur propre et que l'on peut trouver un vecteur propre vérifiant les conditions pour être une loi de probabilité. Si la chaine est irréductible, les états sont récurrents positifs et la loi invariante est donc unique.

Exemple 14 ($\pi = (\pi_1 \ \pi_2 \ \pi_3)$)

Soit $\{X_t \ ; \ t \in \mathbb{N}\}$ une chaine de Markov homogène sur un espace d'état $E = \{1,2,3\}$ et de matrice de transition

$$\mathbf{Q} = \begin{pmatrix} 3/5 & 2/5 & 0 \\ 1/5 & 3/5 & 1/5 \\ 0 & 3/5 & 2/5 \end{pmatrix}.$$

1. Justifiez l'existence et l'unicité d'une loi invariante
2. Pour chaque $x \in E$, donnez π_x en résolvant $\begin{cases} \pi.\mathbf{Q} = \pi, \\ \sum_{x \in E} \pi_x = 1. \end{cases}$
3. pour chaque $x \in S$, donnez $\mathbb{E}_x(S_x)$.

Solution:

1. La chaine de Markov est irréductible avec un espace d'état fini
   $\hookrightarrow$ Il existe une unique distribution stationnaire
2. La loi de probabilité invariante est obtenue en résolvant le système
   $\begin{cases} \pi.\mathbf{Q} = \pi, \\ \sum_{x \in E} \pi_x = 1, \end{cases}$ et $\pi = (3/11 \ 6/11 \ 2/11)$.
3. On en déduit que $\mathbb{E}_1(S_1) = 11/3$, $\mathbb{E}_2(S_2) = 11/6$ and $\mathbb{E}_3(S_3) = 11/2$.

$x \in S$

x est transient

- $\mathbb{P}_x(S_x = \infty) > \infty$
- $\mathbb{P}_x(N_x = \infty) = 0$
- $\sum_{t=0}^{+\infty} Q(x,x)^t < +\infty$
- $\lim_{t \to +\infty} Q(x,x)^t = \pi_x = 0$

x est recurrent

- $\mathbb{P}_x(S_x < \infty) = 1$
- $\mathbb{P}_x(N_x = \infty) = 1$
- $\sum_{t=0}^{+\infty} Q(x,x)^t = +\infty$

x est positive recurrent

- $\mathbb{E}_x(S_x) < \infty$

x est recurrent nul

- $\mathbb{E}_x(S_x) = \infty$
- $\lim_{t \to +\infty} Q(x,x)^t = \pi_x = 0$

x est ergodique

- $d(x) = 1$
- $\lim_{t \to +\infty} Q(x,x)^t = \pi_x > 0$

Que peut-on dire d'une chaine de Markov qui n'est pas irréductible?
Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène sur un espace d'état fini. Ce dernier se divise en $k = k_1 + k_2$ classes d'équivalence avec $O_1, \dots, O_{k_1}$ les classes ouvertes et $F_1, \dots, F_{k_2}$ les classes fermées.

- $x \in \bigcup_{i=1}^{k_1} O_i$ alors x est transient et $\pi(x) = 0$
- Pour chaque classe fermée F_i, $i = 1, \dots, k_2$,
  - On définit une sous matrice de transition $\mathbf{Q}_{F_i}$, de loi stationaire unique π_{F_i}
  - On définit le vecteur $\Pi_{F_i} = \begin{pmatrix} \mathbf{0} & \pi_{F_i} & \mathbf{0} \end{pmatrix}$ de dimension Card(E).

Les lois stationnaires de $(X_n)_{n\in\mathbb{N}}$ admettent la forme suivante

$$\pi = \sum_{i=1}^{k_2} \alpha_i \Pi_{F_i}, \text{ avec } \begin{cases} 0 \le \alpha_i \le 1 \\ \sum_{i=1}^{k_2} \alpha_i = 1 \end{cases}$$

S'il n'y a qu'une seule classe fermée ($k_2 = 1$) alors la loi stationnaire est unique et $\alpha_1 = 1$.

Exemple 15

Soit $\{X_t \ ; \ t \in \mathbb{N}\}$ une chaine de Markov sur un espace d'état $E = \{1,2,3,4,5\}$ and de matrice de transition

$$\mathbf{Q} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 0 & 1/4 \\ 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

.

1. $\{X_t \ ; \ t \in \mathbb{N}\}$ est-elle irreductible, combien de classe de communication, ouvertes/fermées?
2. Donnez la forme générale des lois invariantes.

Solution:

1. La chaine n'est pas irréductible
   - Une classe ouverte $O_1 = \{1,2\}$, $\pi_1 = \pi_2 = 0$
   - Deux classes fermées $C_1 = \{3,4\}$ and $C_2 = \{5\}$
     - de lois invariantes $\pi_1 = (1/3 \ 2/3)$ and $\pi_2 = (1)$

2. Soient

$$\Pi_{C_1} = (0\ 0\ 1/3\ 2/3\ 0),$$

et

$$\Pi_{C_2} = (0\ 0\ 0\ 0\ 1).$$

Les lois invariantes de $\{X_t\ ;\ t \in \mathbb{N}\}$ sont de la forme

$$\pi = \alpha \Pi_{C_1} + (1-\alpha)\Pi_{C_2}, \text{ pour tout } \alpha \in [0,1].$$

3. Théorèmes ergodiques

Le théorème ergodique pour les chaines de Markov est l'analogue de la loi des grands nombres pour les variables aléatoires iid. Soit $(X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène irréductible et récurrente positive sur un espace d'état E et de loi invariante π. Le résultat suivant est démontré dans le cas où l'espace d'état est fini

Theoreme 5 (Théorème ergodique)

Soit $f : E \mapsto \mathbb{R}$ telle que $\int |f| d\pi = \sum_{x\in E} f(x)\pi(x) < \infty$, on a

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \int f(X) d\pi$$

preuve:

Supposons l'espace d'état fini, alors

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) = \sum_{x\in E} \frac{1}{n}\sum_{k=1}^{n} f(x)\mathbb{1}_{X_k=x} \underset{n\to+\infty}{\longrightarrow} \sum_{x\in E} f(x)\pi(x)$$

Remarque 9 (Algorithme MCMC)

Les méthodes de simulation de Monte Carlo par chaine de Markov consiste à estimer l'intégrale $\int f(X) d\pi$ par $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$, où $X_1, \dots, X_n$ sont les points d'une trajectoire d'une chaine de Markov dont la loi invariante est π.

Corollaire 1

Soit ν une mesure de probabilité sur E^{m+1} définie par

$$\nu(x_1,\ldots,x_{m+1}),\ x_1,\ldots,x_{m+1} \in E^{m+1}.$$

Si $g : E^{m+1} \mapsto \mathbb{R}$ est une fonction telle que $\int |g| d\nu < \infty$ alors

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g(X_k, X_{k+1}, \ldots, X_{k+m}) = \int g d\nu \ \text{p.s.}.$$

Idée de la preuve:
$\forall m \geq 1$, on montre que

$$Z_k = (X_k, \ldots, X_{k+m}),\ k \in \mathbb{N}$$

est une chaine de Markov irréductible sur

$$F = \{(x_1, \ldots, x_{m+1}) \in E^{m+1}\ ;\ Q(x_i, x_{i+1}),\ 1 \leq i \leq m\}$$

On montre que ν est la seule loi invariante puis le résultat suit par application du théorème ergodique.

Remarque 10

Une démonstration du théorème ergodique consiste à remarquer que les trajectoires entre deux passages par x (appelées excursions) de la forme

$$(X_{S_n}, X_{S_n+1}, \ldots, X_{S_{n+1}-1})$$

sont les réalisations d'une variable aléatoire i.i.d. puis à appliquer la loi des grands nombre.

V. La marche aléatoire sur $\mathbb{Z}$

La marche aléatoire sur $\mathbb{Z}$ est une chaine de Markov (X_n) dont l'espace d'état est $\mathbb{Z}$ définie par

$$X_n = X_{n-1} + \xi_n, \; n \geq 1.$$

où $\xi_1, \xi_2, \ldots,$ sont des variables aléatoires i.i.d. distribuées comme ξ avec

$$\mathbb{P}(\xi = 1) = p \text{ et } \mathbb{P}(\xi = -1) = 1 - p.$$

Theoreme 6

La marche aléatoire sur $\mathbb{Z}$ est irréductible et

- *Récurrente si $p = 1/2$.*
- *Transiente sinon.*

preuve:

Pour montrer ce résultat, on étudie la distribution du temps S_0 de retour à 0. On a

$$\mathbb{P}_0(S_0 < \infty) = \sum_{n=1}^{+\infty} \mathbb{P}(S_0 = n).$$

On remarque que les trajectoires allant de 0 à 0 sont nécessairement de longueur paire et

$$\mathbb{P}(S_0 = 2n+1) = 0, \text{ pour } n = 0, 1, \ldots.$$

et

$$\mathbb{P}_0(S_0 < \infty) = \sum_{n=1}^{+\infty} \mathbb{P}(S_0 = 2n).$$

On a exactement $\binom{2n}{n}$ trajectoires possibles, celle qui nous intéresse (pour lesquels $S_0 = 2n$) sont celles qui ne repassent pas par 0 entre l'instant 0 et $2n$ (On parle d'excursions). Leur nombre est donné par

$$2 \times C_{n-1} = 2 \times \frac{1}{n}\binom{2n-2}{n-1} \tag{2}$$

Definition 20 (Nombre de Catalan)

Les nombres de Catalan sont définis par

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \text{ pour } n \geq 0,$$

et vérifie

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k} \text{ pour } n \geq 1 \tag{3}$$

Exemple 16 (Mots de Dyck)

C_n correspond aux nombre de mots de $2n$ lettres comprenant respectivement n A et n B, tels que lu de gauche à droite le nombre de A demeure supérieur ou égal au nombre de B. La relation de récurrence (3) s'explique par le fait qu'un mot de Dyck contenant plus de deux lettres est obtenu par la concaténation de deux mots de Dyck.

Dans le problème considéré, on s'intéresse aux nombres de mots tels que le nombre de A (interprétés comme des $+1$) soit strictement supérieur au nombre de B (interprétés comme des -1). Alors notre mot commence nécessairement par un A et fini sur un B. La portion entre ce A et ce B est un mot de Dyck contenant $2n-2$ lettres. On a C_{n-1} possibilités. Le facteur 2 dans (2) s'explique par la symétrie du problème puisque l'on peut considérer les trajectoires dans lesquels les -1 dominent les $+1$. La probabilité d'une trajectoire quelconque de longueur $2n$ contenant n $+1$ et n -1 est donnée par $p^n(1-p)^n$, on en déduit que

$$\begin{aligned}\mathbb{P}(S_0 < \infty) &= \sum_{k=1}^{+\infty} \mathbb{P}(S_0 = k) = \sum_{k=1}^{+\infty} 2C_{n-1}[p(1-p)]^n \\ &= 2p(1-p) \sum_{k=0}^{+\infty} C_n[p(1-p)]^n = 2p(1-p)C[p(1-p)], \qquad (4)\end{aligned}$$

où $C(x) = \sum_{n=0}^{+\infty} C_n x^n$. Or, on a

$$\begin{aligned}C(x) &= 1 + \sum_{n=1}^{+\infty} C_n x^n = 1 + x \sum_{n=0}^{+\infty} C_{n+1} x^n \\ &= 1 + x \sum_{n=0}^{+\infty} \sum_{k=0}^{n} C_k C_{n-k} x^n = 1 + x \sum_{k=0}^{+\infty} \sum_{n=k}^{n} C_k C_{n-k} x^n \\ &= 1 + x \sum_{k=0}^{+\infty} C_k \sum_{n=0}^{n} C_n x^{n+k} = 1 + xC(x)^2\end{aligned}$$

Par suite, $C(x) = \frac{1-\sqrt{1-4x}}{2x}$. En substituant dans (4), on obtient

$$\mathbb{P}(S_0 < \infty) = 1 - |1-2p|$$

On en déduit que si $p \neq 1/2$ alors $\mathbb{P}(S_0 < \infty) < 1$ et la chaine est transiente sinon $\mathbb{P}(S_0 < \infty) = 1$ et la chaine est récurrente.

Remarque 11 (Divergence lorsque $p \neq 1/2$)

Dans le cas d'une chaine de Markov $(X_n)_{n\in\mathbb{N}}$ sur un espace ordonné et dénombrable (comme $\mathbb{N}$ ou $\mathbb{Z}$), si la chaine est transiente alors elle diverge vers ∞. Par exemple dans le cas de la chaine aléatoire sur $\mathbb{Z}$, on a par la loi des grands nombre

$$\frac{X_n}{n} = \frac{X_0}{n} + \frac{1}{n}\sum_{k=1}^{n} \xi_k \underset{n\to+\infty}{\longrightarrow} 2p-1.$$

On en déduit que

$$X_n \to \begin{cases} -\infty, & \text{si } p < 1/2, \\ 0, & \text{si } p = 1/2, \\ +\infty, & \text{si } p > 1/2. \end{cases}$$

Exemple 17 (Le problème de la ruine du parieur)

Un joueur entre dans un casino avec x\$ en poche, il paye 1\$ pour participer,

- Il gagne et remporte 2\$ avec une probabilité p
- Il perd avec une probabilité $q = 1 - p$

sa richesse après chaque partie est modélisée par un processus $(X_n)_{n\in\mathbb{N}}$. On suppose qu'il rentre chez lui si sa richesse devient nulle ou atteint un niveau $a \geq x$. On note $\phi(x,a)$ la probabilité qu'il rentre à la maison ruiné.

Proposition 9

La probabilité de ruine est donnée par

$$\phi(x,a) = \frac{(q/p)^x - (q/p)^a}{1 - (q/p)^a}.$$

preuve:
On note que

$$\phi(a,a) = 0 \text{ et } \phi(0,a) = 1$$

Soit $0 < x < a$, lors de la première partie,

- Il perd avec probabilité q et repart avec un niveau de richesse $x - 1$,
- Il gagne avec probabilité p et repart avec un niveau de richesse $x + 1$.

On en déduit que

$$\phi(x,a) = p\phi(x+1,a) + q\phi(x-1,a) \tag{5}$$

De plus, comme $p+q=1$ alors

$$\phi(x,a) = p\phi(x,a) + q\phi(x,a). \tag{6}$$

L'opération (5)-(6) donne

$$\phi(x+1,a) - \phi(x,a) = \frac{q}{p}\left[\phi(x,a) - \phi(x-1,a)\right]. \tag{7}$$

Soit

$$u_k = \phi(k+1,a) - \phi(k,a), \; k = 0,\dots,a-1$$

en remplaçant dans (7) cela donne

$$u_k = \left(\frac{q}{p}\right) u_{k-1} = \dots = \left(\frac{q}{p}\right)^k u_0. \tag{8}$$

En sommant (8) pour k allant de 1 à $a-1$, on obtient

$$-\phi(1,a) = [1-\phi(1,a)]\frac{q/p-(q/p)^a}{1-q/p} \Leftrightarrow \phi(1,a) = \frac{q/p-(q/p)^a}{1-(q/p)^a} \tag{9}$$

En sommant (8) pour k allant de 1 à $x-1$, on obtient

$$\phi(x,a) - \phi(1,a) = [1-\phi(1,a)]\frac{q/p-(q/p)^x}{1-q/p} \tag{10}$$

L'insertion de (9) dans (10) donne

$$\phi(x,a) = \frac{(q/p)^x - (q/p)^a}{1-(q/p)^a}$$

$\square$

Exemple 18 (Le problème de la double dépense dans les transactions validées par *blockchain*)

Marie achète un bien à Julien en l'échange de 10 BTCs.

- Julien attends que la transactions intègre un bloc, voir que plusieurs blocs soient créés avant d'expédier le bien.
- Une fois le bien reçu, Marie émet un transaction transférant les mêmes BTCs vers un porte-monnaie lui appartenant.
- Des mineurs malhonnêtes travaillent sur une chaine concurrente à la chaine de bloc principale
  - $\hookrightarrow$ les deux chaines sont identiques à la transaction frauduleuse prêt.
- Si la chaine malhonnête rattrape la chaine principale (en termes de nombre de bloc) alors la transaction de Marie à Julien est remplacée par la transaction de Marie à elle-même.

On modélise par $(X_n)_{n\in\mathbb{N}}$ la différence entre les nombres de blocs dans la chaine honnête et malhonnêtes à l'instant n, on suppose qu'à chaque instant un bloc est créé, il rejoint

- La chaine honnête avec probabilité p,
- La chaine malhonnête avec probabilité $q = 1 - p$.

On suppose que la chaine honnêtes a x blocs d'avance. La probabilité de succès de la double dépense est donnée par

$$\psi(x) = \mathbb{P}(X_n = 0, \text{ pour un certain } n \in \mathbb{N} | X_0 = x) = \lim_{a\to+\infty} \phi(x,a) = \left(\frac{q}{p}\right)^x.$$

Pour plus de détails, on pourra lire le *white paper* de Satoshi Nakamoto [4]

Mes notes se basent sur les documents [5, 2, 1, 3].

Maryann Hohn.
PSTAT160A: Applied Stochastic Processes - Lecture notes.
2017.

Nabil Kazi-Tani.
Modèles aléatoires discrets - Cours scannés ISFA.
2017.

J. Lacroix and P. Priouret.
Probabilités approfondies- Polycopié de cours.
Université Pierre et Marie Curie, 2005-2006.

Satoshi Nakamoto.
Bitcoin: A peer-to-peer electronic cash system.
2008.
`https://bitcoin.org/bitcoin.pdf`.

Lionel Truquet.
Statistique des processus 3A - Note de cours.
`http://www.ensai.fr/files/_media/documents/Enseignants%20chercheurs%20-%20doctorants/ltruquet%20-%20documents/polystatdesprocessus2.pdf`.