Chaînes de Markov (et applications)

Raphael Lachieze-Rey*

3 décembre 2018

M1 Paris Descartes.

Table des matières

*lr.raphael@gmail.com

Rappels

0.1 Espérances et probas conditionnelles

On rappelle la formule de probabilités conditionnelles : $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$, pour $\mathbb{P}(B) \neq 0$. On note parfois $\mathbb{P}_B(A) = \mathbb{P}(A|B)$, rappelons que $\mathbb{P}_B(\cdot)$ est une mesure de probabilités à part entière.

Le double conditionnement se traite ainsi :

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap B|C)\mathbb{P}(C) = \mathbb{P}_C(A \cap B)\mathbb{P}(C) = \mathbb{P}_C(A|B)\mathbb{P}_C(B)\mathbb{P}(C) = \mathbb{P}(A|\underbrace{B, C}_{i.e. B\cap C})\mathbb{P}(B|C)\mathbb{P}(C).$$

Etant donné des variables aléatoires X et Y à valeurs réélles,

$$\mathbb{E}(X|Y) = \varphi(Y)$$

est une variable aléatoire qui est entièrement déterminée par Y. Par exemple, si X, Y sont des variables de Bernoulli de paramètre $1/2$ indépendantes,

$$\mathbb{E}((X+Y)^2|Y) = \mathbb{E}(X^2|Y) + 2\mathbb{E}(XY|Y) + \mathbb{E}(Y^2|Y) = \mathbb{E}X^2 + 2Y\mathbb{E}X + Y^2 = \frac{1}{2} + Y + Y^2 = \varphi(Y).$$

0.2 Familles sommables et théorème de Fubini

Etant donné un espace mesuré (Ω, μ) et une fonction $f : \Omega \to \mathbb{R}$, l'intégrale

$$\int_\Omega f(x)\mu(dx),$$

n'a de sens que si f est intégrable

$$\int_\Omega |f(x)|\mu(dx) < \infty$$

(exemple de $x \mapsto \frac{\sin(x)}{x}$ sur $\mathbb{R}$). Par contre, si $f \geqslant 0$, alors

$$\int_\Omega f(x)\mu(dx) \in \mathbb{R}_+ \cup \{+\infty\}$$

est défini sans ambiguité. De même, étant donné une série $(a_n; n \in \mathbb{N})$,

$$\sum_{n=0}^{\infty} a_n$$

n'a de sens que si $a_n \geqslant 0$ ou si $(a_n; n \geqslant 0)$ est sommable, c'est-à-dire

$$\sum_{n=0}^{\infty} |a_n| < \infty.$$

Pour ce qui est de l'interversion, le théorème de Fubini nous dit que pour une fonction bi-mesurable $f(x, y)$ sur un produit d'espaces mesurés $\Omega \times \Omega'$,

$$\int_\Omega \left(\int_{\Omega'} f(x,y)\mu(dx) \right) \mu'(dy) = \int_{\Omega'} \left(\int_\Omega f(x,y)\mu'(dy) \right) \mu(dx)$$

si $f(x, y) \geqslant 0$ ou si

$$\int_{\Omega \times \Omega'} |f(x,y)|\mu(dx)\mu'(dy) < \infty$$

ou si, de manière équivalente,

$$\int_\Omega \left(\int_{\Omega'} |f(x,y)|\mu(dx) \right) \mu(dy) < \infty.$$

Les fonctions positives peuvent être intégrées dans l'ordre qu'on veut. Si X_n sont des variables aléatoires sur l'espace probabilisé $(\Omega, \mathbb{P})$, $\Omega' = \mathbb{N}$ et μ' est la mesure de comptage, ça nous donne avec $f(\omega, n) = X_n(\omega)$,

$$\mathbb{E} \sum_{n \in \mathbb{N}} X_n = \sum_{n \in \mathbb{N}} \mathbb{E} X_n$$

sans besoin de justification si $f(\omega, n) \geqslant 0$, ou si

$$\mathbb{E} \sum_{n \in \mathbb{N}} |X_n| = \sum_{n \in \mathbb{N}} \mathbb{E}|X_n| < \infty.$$

0.3 Formule des probabilités totales

Soit U, V deux variables à valeurs dans un espace dénombrable E. Alors pour $x \in E$,

$$\mathbb{P}(U = x) = \sum_{y \in E} \mathbb{P}(U = x, V = y) = \sum_{y \in E} \mathbb{P}(U = x | V = y)\mathbb{P}(V = y).$$

1 Chaînes de Markov homogènes

$(\Omega, \mathbb{P})$ est un espace probabilisé.

1.1 Exemples et définitions

Idée : Une chaîne de Markov est une suite de variables aléatoires dans le temps ou **conditionnellement au présent, le futur ne dépend pas du passé**, ou autrement dit le futur ne dépend du passé que par le présent.

Formellement, soit E un espace fini ou dénombrable. Ce sera l'**espace d'états**.

Définition 1. *Soit $X = \{X_n; n \geq 0\}$ une suite de variables aléatoires à valeurs dans E. On dit que X est une chaîne de Markov si, pour tout $x_1, \ldots, x_{n+1} \in E$, on a*

$$\mathbb{P}(\underbrace{X_{n+1} = x_{n+1}}_{\text{Le futur}} \mid \underbrace{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n}_{\text{Le passé (et le présent)}}) = \mathbb{P}(\underbrace{X_{n+1} = x_{n+1}}_{\text{Le futur}} \mid \underbrace{X_n = x_n}_{\text{Le présent}})$$

Cette propriété des chaînes de Markov est aussi connue comme **propriété de Markov**.

Exercice 1.
Parmi les exemples suivants, lesquels correspondent a une chaîne de Markov ?
- Les records du monde du 100m
- La population mondiale
- La position d'une voiture (car le passé nous renseigne sur sa vitesse, et donc sur sa position future)
- Le nombre de personnes dans une file d'attente
- Un marcheur aléatoire qui ne revient jamais sur ses pas.
- Le couple (position, vitesse) d'une voiture de course
- une marche aléatoire

Exemple 1. Modélisation
- Séquence d'ADN : ACGGTAACTC... peut-être vue en première approximation comme une chaîne de Markov
- Evolution de population : Chaque jour, un individu nait ou un individu meurt
- Généalogie/Epidémiologie : Chaque jour, un individu donne naissance (ou contamine) un nombre aléatoire d'individu, ou meurt (guérit)
- Intelligence artificielle
- Simulation. Exemple : jeu de cartes mélangé. On part d'un jeu de cartes (fictif) dans l'ordre, et à chaque coup on applique l'interversion de 2 cartes tirées au hasard. La "loi" du jeu de cartes converge vers la loi d'un jeu de cartes mélangé selon une permutation uniforme

Les questions auxquelles on va tenter de répondre dans ce cours :
- Conaissant la loi de X_0, quelle est la loi de $X_n, n \in \mathbb{N}$? La loi de X_n converge-t-elle ?
- Partant d'un certain $x \in E$, et pour $y \in E$, quelle est la proba que la chaîne passe par y, i.e. qu'il existe un temps $T < \infty$, aléatoire, pour que $X_T = y$? Quel est l'espérance de T ?
- ...

Exercice 2. Soit $R_n, n \geq 0$ des variables indépendantes à valeurs dans $E = \mathbb{N}$. Montrer que $S_n = \sum_{i=1}^n R_i$ et $P_n = \prod_{i=1}^n R_i$ sont des chaînes de Markov.

Une chaîne de Markov peut être vue comme un **système dynamique**, ce qui veut dire que $X_{n+1} = f_n(X_n)$, ou f_n est une "transformation aléatoire" indépendante du passé. Dans l'exemple précédent, $f_n(X_n)$ est la somme (ou le produit) de X_n avec R_{n+1}.

Si la transformation aléatoire f_n ne dépend pas de n, i.e. $X_{n+1} = f(X_n)$ pour tout n pour une certaine transformation f, on dit que X est une **chaîne de Markov homogène**.

Cela veut dire que si à un certain instant $n \geq 0$ la chaîne se trouve à l'état x ($X_n = x$), alors la probabilité qu'elle se trouve à l'état y au temps $n + 1$ est la même que si l'on était au temps initial.

Définition 2. *Une chaîne de Markov est* **homogène** *si pour tout $n \geq 0$, x et y dans E*

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

Dans ce cas, on pose

$$Q(x, y) = \mathbb{P}(X_1 = y \,|\, X_0 = x),\ x, y \in E.$$

Q est la **matrice de transition** *de la chaîne X, on dit aussi* **noyau de transition** *quand E est infini.*

Dans ce cours, toutes les chaîne de Markov sont supposées homogènes, ce ne sera pas forcément explicitement écrit.

Remarque 1. C'est éventuellement une matrice infinie

Une chaîne de Markov homogène "saute" donc aléatoirement d'états en états, et la probabilité de chaque saut est donnée par la matrice Q.

Notation : Comme E est un espace dénombrable, une mesure μ sur E est défini par sa valeur sur les atomes :

$$\mu(\{x\}), x \in E.$$

Réciproquement, si $\mu(\{x\}) \geqslant 0$ pour $x \in E$, μ définit bien une mesure sur E via

$$\mu(A) = \sum_{x \in A} \mu(\{x\}), A \subset E.$$

On abrège parfois cette notation en $\mu(x) = \mu(\{x\})$. μ est alors une mesure de probabilité si

$$\sum_{x \in E} \mu(x) = 1.$$

On peut également noter cette mesure comme un vecteur si $E = \{x_1, x_2, ...\}$ est ordonné : $\mu = (\mu(x))_{x \in E} = (\mu_1, \mu_2, \mu_3, ...)$ est équivalent à $\mu(x_i) = \mu_i$.

Définition 3. *Soit Q la matrice de transition d'une chaîne de Markov homogène. Etant donné un état $x \in E$, la mesure*

$$Q_x(y) = Q(x, y)$$

est une distribution (i.e. une mesure de probabilité) sur E, appelée mesure de saut depuis x.

Démonstration:
En effet, pour $x \in E$,

$$\sum_{y \in E} Q_x(y) = \sum_{y \in E} Q(x, y) = \sum_{y \in E} \mathbb{P}(X_1 = y | X_0 = x) = \mathbb{E}(\sum_{y \in E} \mathbf{1}_{\{X_1 = y\}} | X_0 = x) = \mathbb{E}(1 | X_0 = x) = 1.$$

Exercice 3. Markov lui-même a analysé la succession de voyelles et de consonnes dans le livre *Eugene Onekin* de Pushkin sous l'angle des chaînes de Markov. Il est parti des données suivantes : *A. Markov, studied the sequence of 20,000 letters in A. S. Pushkin's poem "Eugeny Onegin" discovering that the stationary vowel probability is* $p = 0.432$*, that the probability of a vowel following a vowel is* $p_1 = 0.128$*, and that the probability of a vowel following a consonant is* $p_2 = 0.663$.

1. Quel est l'espace d'état ? La Matrice de transition ?
2. Mêmes questions avec "The Childhood of Bagrov, the Grandson" : *Markov also gave the results of his other tests ; he studied the sequence of 100,000 letters in S. T. Aksakov's novel. For that novel, the probabilities were p = 0.449, p1 = 0.552, and p2 = 0.365.*

Correction:
$E = \{V, C\}$

$$Q = \begin{pmatrix} p_1 & 1 - p_1 \\ p_2 & 1 - p_2 \end{pmatrix}$$

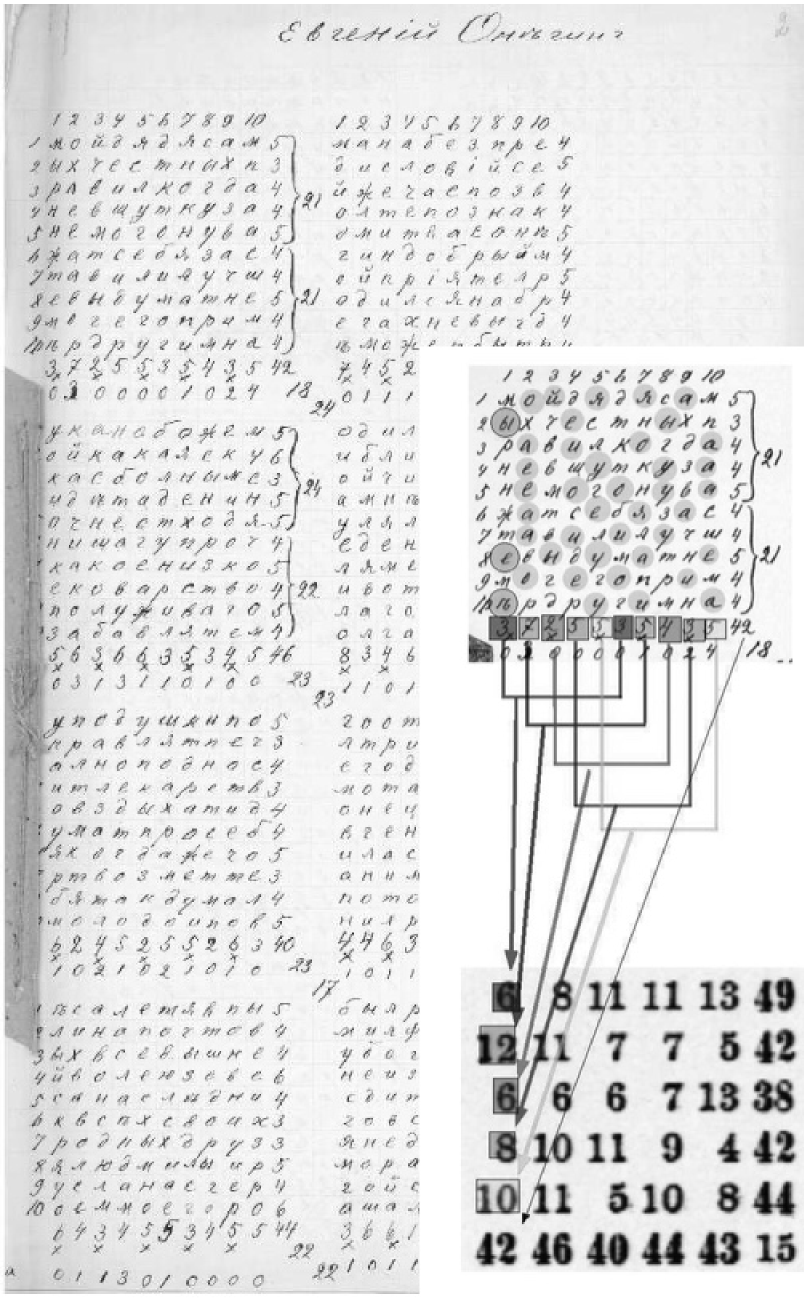


FIG. 2.1. *Left background: The first 800 letters of 20,000 total letters compiled by Markov and taken from the first one and a half chapters of Pushkin's poem "Eugeny Onegin." Markov omitted spaces and punctuation characters as he compiled the cyrillic letters from the poem. Right foreground: Markov's count of vowels in the first matrix of 40 total matrices of* 10×10 *letters. The last row of the* 6×6 *matrix of numbers can be used to show the fraction of vowels appearing in a sequence of 500 letters. Each column of the matrix gives more information. Specifically, it shows how the sums of counted vowels are composed by smaller units of counted vowels. Markov argued that if the vowels are counted in this way, then their number proved to be stochastically independent.*

Exemple 2. Une grenouille monte sur une échelle. Chaque minute, elle peut monter d'un barreau avec probabilité 1/2, ou descendre d'un barreau avec probabilité 1/2. L'échelle a 5 barreaux. Si la grenouille arrive tout en haut, elle saute immédiatement en bas de l'échelle et recommence.

On appelle X_n la position de la grenouille sur l'échelle. L'espace d'états est donc $E = \{0, 1, 2, \ldots, 5\}$. Si a un instant n la grenouille est au niveau $x \in \{1, 2, 3, 4\}$ de l'échelle, alors à l'instant $n+1$ elle sera

$$\begin{cases} \text{au barreau } x+1 \text{ avec probabilité } 1/2, \\ \text{au barreau } x-1 \text{ avec probabilité } 1/2, \end{cases}$$

ce qui se traduit par

$$\mathbb{P}(X_{n+1} = x+1 \,|\, X_n = x) = 1/2 \quad (= \mathbb{P}(X_1 = x+1 \,|\, X_0 = x)),$$
$$\mathbb{P}(X_{n+1} = x-1 \,|\, X_n = x) = 1/2 \quad (= \mathbb{P}(X_1 = x-1 \,|\, X_0 = x)$$

Comme les probabilités ne dépendent pas de n, il semble que l'on tienne le bon bout pour avoir une chaîne de Markov homogène. Si c'est le cas, on peut écrire une partie de la matrice de transition :

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

Si la grenouille se retrouve à l'état 5, alors elle peut soit passer à l'état 4, soit passer à l'état 0. Il faut donc remplacer la dernière ligne de la matrice par

$$(1/2, 0, 0, 1/2, 0),$$

(encore une fois cela ne dépend pas de l'instant n). Si la grenouille est à l'état 0, elle ne peut que passer à l'état 1. La première ligne de la matrice est donc

$$(0, 1, 0, 0, 0).$$

X_n est donc bien une chaîne de Markov homogène, avec matrice de transition Q.

Exercice 4. Introduisons un facteur de fatigue $f \in (0, 1)$, et imaginons qu'à chaque instant la grenouille reste à son état actuel avec probabilité f. X_n est toujours une chaîne de Markov ? Si oui, quelle est sa matrice de transition ?

Imaginons désormais que le facteur de fatigue $f = f_n$ dépend du temps. Que cela change-t-il ?

Si désormais le facteur de fatigue dépend de tout le chemin parcouru par la grenouilleque cela change-t-il ?

Exercice 5. Le nombre d'individus d'une population évolue de la manière suivante : A chaque instant, un individu nait avec la probabilité $p \in (0, 1)$, ou un individu meurt avec la probabilité $q = 1 - p$.

Ecrire le noyau de transition.

Ecrire la chaîne de Markov en termes des variables introduites à l'exercice 2.

Comment corriger la matrice de transition pour qu'il n'y ait pas un nombre négatif d'individus ?

Correction: X_n est le nombre d'individus au temps n, pour $n \in \mathbb{N}$. Au temps $n+1$, on a soit X_n+1 individus, avec probabilité p, ou X_n-1 individus, avec probabilité $q = 1-p$. Donc pour tout x, $Q(x, x+1) = \mathbb{P}(X_{n+1} = X_n+1 | X_n = x) = 1/2 = \mathbb{P}(X_1 = x+1 | X_0 = x)$, et de même $Q(x, x-1) = 1/2$.

Pour $y \notin \{x-1, x+1\}$, $Q(x, y) = 0$. On a donc une matrice de transition infinie qui ressemble à ça

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

On peut en fait écrire que $X_{n+1} = X_n + R_n$, où R_n vaut soit $+1$ soit -1. On appelle ce type de variable une variable de Rademacher :

$$\mathbb{P}(R_n = \pm 1) = 1/2.$$

Le problème est qu'on peut avoir un nombre négatif d'individus. On suppose que quand il y a 0 individus, la population s'éteint. Cela signifie que $Q(0,0) = 1$ et $Q(0,n) = 0$ pour $n \neq 0$. Il faut donc corriger la MDT en remplaçant la 1re ligne par $(1, 0, 0, \dots)$ (et les nombres négatifs ne font plus partie de l'espace d'état.) Ca nous donne

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

Définition 4. *On dit qu'une matrice Q (ou un noyau éventuellement infinie) est* **stochastique** *ssi tous ses coefficients sont ≥ 0 et si la somme de chaque ligne fait 1 : $\forall x \in E$,*

$$\sum_{y \in E} Q(x,y) = 1.$$

On dit aussi **matrice markovienne**.

Remarque 2. Les coefficients d'une matrice stochastique sont dans $[0,1]$, ils peuvent donc représenter une probabilité...

Proposition 1. *Si Q est la matrice de transition d'une chaîne de Markov, alors elle est stochastique.*

Démonstration: C'est une conséquence directe de la Définition 3.

On va voir que réciproquement, pour toute matrice stochastique Q sur un produit $E \times E$, il existe une chaîne de Markov sur E qui admet Q comme matrice de transition . Mais elle n'est pas unique, car il faut encore préciser la loi du premier état X_0.

Exercice 6. Soit $X = (X_n)_{n \geqslant 0}$ une chaîne de Markov . Soit $k \in \mathbb{N}$ et $X' = (X'_n)_{n \geqslant 0}$ la suite de variables aléatoires définie par

$$\begin{aligned} X'_0 &= X_k \\ X'_1 &= X_{k+1} \\ X'_2 &= X_{k+2} ... \end{aligned}$$

Montrer que X' est une chaîne de Markov de matrice de transition Q. On la note parfois $\tau_k X$, c'est-à-dire que τ_k est l'opérateur de translation, ou de *shift temporel*, agissant sur la chaîne de Markov X.

On appelle graphe d'une chaîne de Markov (ou d'une matrice de transition) le graphe dont les sommets sont les états possibles et étant donné $x, y \in E$, il y a une flèche de x vers y si $Q(x,y) > 0$.

1.2 Loi des marginales

Le comportement d'une chaîne de Markov X dépend entièrement de sa matrice de transition Q, et de la position initiale X_0.

On appelle **loi initiale** de X la loi de X_0, c'est une mesure définie par

$$\mu_0(x) = \mu_0(\{x\}) = \mathbb{P}(X_0 = x).$$

Connaissant μ_0 et Q, on peut calculer directement la loi de X_n, aussi appelée *loi marginale*. Pour cela, il est plus facile de parler de la notion de **chemin**. Un chemin est simplement la donnée d'un vecteur fini $(x_0, x_1, \dots, x_n) \in E^n$. La probabilité de suivre un chemin donné est facile à calculer :

Proposition 2. *Pour tout chemin* $(x_0, x_1, \dots, x_n)$ *dans* E*, on a*

$$\begin{aligned}\mathbb{P}(X_0 = x_0, X_1 = x_1, &X_2 = x_2, \dots, X_n = x_n)\\ &= \mu_0(x_0)Q(x_0, x_1)Q(x_1, x_2)\dots Q(x_{n-1}, x_n).\end{aligned}$$

Démonstration:

On a (en utilisant la propriété de Markov)

$$\begin{aligned}\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)&\\ &= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0)\mathbb{P}(X_{n-1} = x_{n-1}, \dots, X_0 = x_0)\\ &= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1})\mathbb{P}(X_{n-1} = x_{n-1}, \dots, X_0 = x_0)\\ &= Q(x_{n-1}, x_n)\mathbb{P}(X_{n-1} = x_{n-1}, \dots, X_0 = x_0)\\ &= Q(x_{n-1}, x_n)\mathbb{P}(X_{n-1} = x_{n-1} | X_{n-2} = x_{n-2}, \dots, X_0 = x_0)\mathbb{P}(X_{n-2} = x_{n-2}, \dots, X_0 = x_0)\\ &= Q(x_{n-1}, x_n)\mathbb{P}(X_{n-1} = x_{n-1} | X_{n-2} = x_{n-2})\mathbb{P}(X_{n-2} = x_{n-2}, \dots, X_0 = x_0)\\ &= Q(x_{n-1}, x_n)Q(x_{n-2}, n-1)\mathbb{P}(X_{n-2} = x_{n-2}, \dots, X_0 = x_0)\\ &= \dots(\text{récurrence})\\ &= Q(x_{n-1}, x_n)Q(x_{n-2}, x_{n-1})\dots Q(x_0, x_1)\underbrace{\mathbb{P}(X_0 = x_0)}_{\mu_0(x_0)}.\end{aligned}$$

Exercice 7. Soit $a \in]0, 1[$. On considère une chaîne de Markov dont la matrice de transition est la suivante :

$$\begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$$

Calculer la probabilité de l'événement A_3 de passer pour $n \leq 4$ trois fois par l'état 2

1. Avec $\mu_0 = (1, 0)$,
2. Avec $\mu_0 = (1/2, 1/2)$.

Correction:

Il faut dénombrer tous les chemins possibles qui contiennent trois fois l'état 2. En partant de 1 :

— $x^1 = (1, 1, 2, 2, 2)$
— $x^2 = (1, 2, 1, 2, 2)$
— $x^3 = (1, 2, 2, 1, 2)$
— $x^4 = (1, 2, 2, 2, 1)$

Pour le premier chemin, sa probabilité est, d'après la formule précédente,

$$\mathbb{P}(X = x^1) = Q(1, 1)Q(1, 2)Q(2, 2)Q(2, 2) = a(1-a)a^2 = a^3(1-a).$$

Avec d'autres calculs, on montre

$$\mathbb{P}(A_3) = a^3(1-a) + 2a(1-a)^3 + a^2(1-a)^2.$$

Dans un chemin $(x_0, \dots, x_n)$, si il existe i tel que $Q(x_i, x_{i+1}) = 0$, cela signifie qu'il est impossible de passer de l'état x_i à l'état x_{i+1}, et il est donc impossible de suivre le chemin $(x_0, \dots, x_n)$. On dit qu'un chemin $x_0, \dots, x_n$ est **possible**, et on le note

$$x_0 \to x_1 \to \cdots \to x_n,$$

si $\forall 1 \leqslant i \leqslant n, Q(x_{i-1}, x_i) > 0$.

Soit $x, y \in E$. On dit qu'il existe un chemin **possible** entre x et y si il existe un chemin possible de la forme $(x, x_1, \dots, x_{n-1}, y)$.

Pour une même chaîne X, on considère souvent plusieurs lois initiales différentes. Dans ce cas on précise la loi utilisée en notant

$$\mathbb{P} = \mathbb{P}_{\mu_0}$$

dans chaque calcul de probabilité, et l'espérance est alors notée $\mathbb{E}_{\mu_0}$. Si la loi est un "Dirac" : $\mu_0 = \delta_x$ pour un certain $x \in E$ (ce qui veut dire $X_0 = x$ p.s.), alors on note plus simplement $\mathbb{P}_{\delta_x} = \mathbb{P}_x, \mathbb{E}_{\delta_x} = \mathbb{E}_x$.

Exemple 3. Pour reprendre l'exemple de la grenouille,

$$\begin{aligned}
&\mathbb{P}_0(X_1 = 1) = 1,\\
&\mathbb{P}_0(X_1 = 3) = 0,\\
&\mathbb{P}_2(X_1 = 3) = 1/2,\\
&\mathbb{P}_0(X_3 = 3) = (1/2)^3 = 1/8,\\
&\mathbb{P}_0(X_3 = 4) = 0,\\
&\dots
\end{aligned}$$

On peut calculer directement la loi de X_n en utilisant le produit matriciel.

Notation 1. *Pour une mesure μ_0 et une matrice Q, on note la mesure*

$$(\mu_0 Q)(y) = \sum_{x \in E} \mu_0(x) Q(x, y).$$

Cela revient à multiplier (matriciellement) la mesure μ_0 vue comme un vecteur $\mu_0 = (\mu_0(x_1), \mu_0(x_2), \dots)$ par la matrice Q (la taille du vecteur et de la matrices peuvent être infinies).

Proposition 3. *Si μ_0 est la loi de X_0, considérée comme un vecteur, alors $(\mu_0 Q)$ est la loi de X_1.*

Démonstration: Soit $y \in E$.

$$\begin{aligned}
\mathbb{P}(X_1 = y) &= \sum_x \mathbb{P}(X_1 = y, X_0 = x)\\
&= \sum_x \mathbb{P}(X_1 = y | X_0 = x) \mathbb{P}(X_0 = x) = \sum_x \mu(x) Q(x, y) = (\mu Q)(y)
\end{aligned}$$

Proposition 4. *Pour tout n, la loi de X_n est μQ^n. En particulier, si $\mu = \delta_x$, l'élément à la ligne x et colonne y se retrouve via*

$$(\delta_x Q^n)(y) = Q^n(x, y) = \mathbb{P}_x(X_n = y).$$

Démonstration: D'après la proposition 3, la loi de X_1 est $\mu_1 = \mu Q$.

On peut alors "oublier" la variable X_0, et ne considérer que la chaîne qui part de X_1 (formellement, poser $X'_0 = X_1, X'_1 = X_2, etc...$). La matrice de transition est toujours Q, par contre la loi initiale n'est plus μ_0, c'est μ_1.

La loi de X_2 (i.e. X'_1) est donc, en réutilisant la proposition 3,

$$\mu_2 = (\mu_1 Q) = ((\mu_0 Q) * Q).$$

Comme le produit matriciel est associatif, $((\mu_0 Q) * Q) = \mu_0 * Q * Q = \mu_0 * (Q * Q) = \mu_0 * Q^2$.

La loi de X_2 est donc bien μQ^2, comme annoncé. En raisonnant par récurrence, on montre bien $\mu_3 = (\mu_0 * Q^2) * Q = \mu_0 Q^3$, ..., ... et $\mu_n = \mu_0 Q^n$.

Exercice 8. Une chaîne de Markov avec états $E = \{1, 2\}$ a la matrice de transition

$$Q = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$$

pour un nombre $a \in (0, 1)$.

1. Diagonaliser Q
2. Calculer Q^n pour $n \geqslant 1$.
3. Calculer la loi de X_n pour tout n, sachant que l'on part de l'état $X_0 = 1$. Donner par exemple $\mathbb{P}_1(X_n = 1)$ la probabilité de retour en 1 en n coups.

Correction:

1. On a $\mu_0 = \delta_1$. La loi de X_n est donnée par $\mu_0 Q^n$, il faut donc calculer la puissance n-ème de Q (c'est la difficulté de ce type d'exercice).

   Pour cela, le plus simple est de diagonaliser la matrice. Comme c'est une matrice stochastique,

   $$Q * 1 = 1 \quad ,$$

   ou $1 = (1, 1, ..., 1)^t$, et $*$ est le produit matriciel ($Q * 1(x) = \sum_y 1 Q(x, y) = 1$). Cela veut dire que 1 est vecteur propre associé à la valeur propre 1.

   Comme la matrice Q est symétrique, elle est diagonalisable. La première valeur propre est 1, et la trace est

   $$Tr(Q) = 2a.$$

   Il s'ensuit que la seconde valeur propre est $\lambda = 2a - 1$. Pour trouver le vecteur propre associé, résolvons le système $Q * (u, v) = \lambda(u; v)$:

   $$\begin{cases} au + (1-a)v = (2a-1)u \\ (1-a)u + av = (2a-1)v, \end{cases} \quad \begin{cases} (1-a)v = (a-1)u \\ (1-a)u = (a-1)v, \end{cases}$$

   i.e $u = -v$. Le vecteur propre $(1; -1)$ correspond donc à la valeur propre λ.
2. On a $Q = PDP^{-1}$ où

   $$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

   et

   $$P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

3. Calculer $\mathbb{P}_1(X_n = 1)$ revient à calculer $\mathbb{P}(X_n = 1)$ pour la chaîne de Markov de loi initiale $\mu_0 = (1, 0)$ et de matrice de transition Q. Donc la loi de X_n est $\mu_0 Q^n = (1, 0) P D^n P^{-1} = (1/2 + 1/2(2a-1)^n, 1/2 - 1/2(2a-1)^n)$.

   Après n coups, X_n a un tout petit peu plus de chances d'être en 1 qu'en 2, mais ce tout petit peu est en λ^n et s'atténue rapidement avec le temps. Plus a(=probabilité de rester à sa place) est grand, plus ça décroît lentement.

Exercice 9. (exercice 1.1.4 du Norris, p. 9) Une puce saute aléatoirement sur les sommets d'un triangle, sans préférence. Quelle est la probabilité qu'après n sauts la puce soit de retour à son point de départ ?

Remarque : De manière similaire à l'exercice précédent, on devrait trouver 1/3 plus un terme correctif qui décroît (vite) avec le temps.

Recommencer si cette fois la puce saute dans le sens des aiguilles d'une montre avec probabilité 2/3 et dans l'autre sens avec probabilité 1/3.

Remarque 3. TRES IMPORTANT ! !

$$Q^k(x, y) \neq Q(x, y)^k.$$

membre de gauche : multiplication matricielle.

membre de droite : multiplication de réels (beaucoup plus facile).

Proposition 5. *On a pour $n \geq 0, k \geq 0$*

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

Démonstration: Par la propriété de Markov,

$$\mathbb{P}(X_{n+k} = y | X_n = x) = \mathbb{P}(X_k = y | X_0 = x).$$

Supposons que $\mu_0 = \delta_x$, c'est-à-dire que la chaîne démarre toujours en x. On a

$$\mathbb{P}_x(X_k = y | X_0 = x) = \mathbb{P}_x(X_k = y) = (\mu_0 Q^k)(y) = (\delta_x Q^k)(y) = Q^k(x, y).$$

Dans le cas général (μ_0 quelconque), on a également $\mathbb{P}(X_k = y | X_0 = x) = \mathbb{P}_x(X_k = y) = Q^k(x, y)$.

Remarque 4. Une chaîne de Markov est une variable àléatoire à valeurs dans $\mathcal{S}_E := E^{\mathbb{N}}$ l'ensemble des suites à valeurs dans E. On munit $E^{\mathbb{N}}$ de la tribu $\mathcal{B}(\mathcal{S}_E)$ engendrée par les produits finis de parties de E, c'est-à-dire de la forme $\mathbf{A} := A_1 \times \cdots \times A_n \times E^{\mathbb{N}} \subset \mathcal{S}_E$. Cela signifie que les évènements de la forme

$$(X_1 \in A_1, \dots, X_n \in A_n)$$

sont mesurables (car il est inutile d'écrire $X_{n+1} \in E, X_{n+2} \in E$, ..., cette partie de l'évènement est automatiquement réalisée). La loi d'une chaîne de Markov est donc une mesure de probabilité sur l'espace mesurable $(\mathcal{S}_E, \mathcal{B}(\mathcal{S}_E))$.

Il est clair que si deux chaîne de Markov $X = (X_n)$ et $Y = (Y_n)$ ont la même loi initiale μ_0 et la même matrice de transition , alors elles ont la même loi (On peut montrer par récurrence que pour tout n, $(X_0, \dots, X_n)$ a la même loi que $(Y_0, \dots, Y_n)$, CQFD).

Exercice 10. Soit E dénombrable et $x \in E$. Soit $X = (X_n)_{n \geqslant 0}$ une chaîne de Markov sur E.

1. Soit $n, p \in \mathbb{N}$. Montrer que l'évènement $\mathbf{A}_n^p$=“X passe au moins n fois par l'état x avant le temps p” est mesurable.
2. Soit $n \in \mathbb{N}$. Montrer que l'évènement $\mathbf{A_n}$ =“X passe n fois au moins par x” est mesurable.
3. Montrer que l'évènements $\mathbf{A}$ =“X passe une infinité de fois par x” est mesurable.
4. On suppose $E \subset \mathbb{R}$. Soit $l \in \mathbb{R}$. Soit $\overline{X_n} = \frac{1}{n} \sum_{k=1}^{n} X_k$. Montrer que l'évènement “$\overline{X_n} \xrightarrow[n \to \infty]{} x$” est mesurable.

Correction:

1.
$$\mathbf{A}_n^p = \bigcup_{\{i_1,\dots,i_n\}\subset\{1,\dots,p\}} \underbrace{(X_{i_1} = x, \dots, X_{i_n} = x)}_{\text{mesurable}}$$
Donc c'est un union fini d'ensembles mesurables.
2. $\mathbf{A}_n = \cap_{p=1}^{\infty} \mathbf{A}_n^p$ intersection dénombrable d'évènements mesurables.
3.
$$\mathbf{A} = \cap_{n\geqslant 1} \mathbf{A}_n$$
intersection dénombrable d'évènements mesurables.

Théorème 1. *Soit μ_0 une distribution sur E, et Q une matrice de transition sur E. Alors il existe une chaîne de Markov $X = (X_n)_{n\geqslant 0}$ de loi initiale μ_0 et de matrice de transition Q.*

Démonstration:
Soit $n \geqslant 1$ et $X^{(n)} = (X_0, \dots, X_n)$ un vecteur aléatoire dont la loi est définie par, pour tout n-uplet $x = (x_0, \dots, x_n)$,

$$\mathbb{P}(X^{(n)} = x) = \mu_0(x) \prod_{i=1}^{n} Q(x_{i-1}, x_i).$$

Ces probabilités définissent bien une loi car les évènements $(X^{(n)} = x), x \in E^{n+1}$, engendrent la tribu discrète sur E^{n+1}, i.e. la tribu $\mathcal{P}(E^{n+1})$ de toutes les sous-parties de E.

Il faut ensuite invoquer le théorème de projection de Kolmogorov pour montrer qu'il existe une unique variable $X = (X_0, X_1, ...)$ sur $\mathcal{B}(\mathcal{S}_E)$ dont les projections sur $\mathcal{P}(E^{n+1})$ sont les lois des $X^{(n)}$. On a donc bien une unique variable $X \in E^{\mathbb{N}}$ telle que les $X^{(n)}$ aient la loi presecrite pour les projection de X, i.e.

$$\mathbb{P}(X_0 \in A_1, \dots, X_n \in A_n) = \mathbb{P}(X^{(n)} \in A_1 \times ... \times A_n).$$

On peut donc assimiler (la loi d') une chaîne de Markov sur E a la donnée d'un couple (μ_0, Q) où μ_0 est une probabilité et Q est une matrice stochastique.

On peut étendre la propriété de Markov au futur et au passé au sens large :

Théorème 2. *Soit $n \in \mathbb{N}^*$. Soit $F \in \sigma(X_{n+1}, X_{n+2}, ...)$ un évènement du "futur". Soit $P \in \sigma(X_n)$ un évènement du "présent". Soit $S \in \sigma(X_0, \dots, X_{n-1})$ un évènement du "passé". Alors*

$$\mathbb{P}(F|P \cap S) = \mathbb{P}(F|P).$$

Démonstration:
On suppose dans un premier temps que F est de la forme $(X_{n+1} = x, X_{n+2} = y)$, $P = (X_n = z)$ et $S = (x_{n-1} = t)$. Ainsi

$$\mathbb{P}(F|P \cap S) = \mathbb{P}(X_{n+2} = y, X_{n+1} = x | X_n = x_n, X_{n-1} = t) = \mathbb{P}(X_{n+2} = y)$$

1.3 Exercices

Exercice 11. Soit $R_n, n \geq 0$ des variables indépendantes à valeurs dans $E = \mathbb{N}$. Montrer que $S_n = \sum_{i=1}^n R_i$ et $P_n = \prod_{i=1}^n R_i$ sont des chaînes de Markov.

Exercice 12. Markov lui-même a analysé la succession de voyelles et de consonnes dans le livre *Eugene Onekin* de Pushkin sous l'angle des chaînes de Markov. Il est parti des données suivantes : *A. Markov, studied the sequence of 20,000 letters in A. S. Pushkin's poem "Eugeny Onegin" discovering that the stationary vowel probability is* $p = 0.432$*, that the probability of a vowel following a vowel is* $p_1 = 0.128$*, and that the probability of a vowel following a consonant is* $p_2 = 0.663$.

1. Quel est l'espace d'état ? La Matrice de transition ?
2. Mêmes questions avec "The Childhood of Bagrov, the Grandson" : *Markov also gave the results of his other tests ; he studied the sequence of 100,000 letters in S. T. Aksakov's novel. For that novel, the probabilities were p = 0.449, p1 = 0.552, and p2 = 0.365.*

Correction:
$E = \{V, C\}$

$$Q = \begin{pmatrix} p_1 & 1-p_1 \\ p_2 & 1-p_2 \end{pmatrix}$$

Exercice 13. Le nombre d'individus d'une population évolue de la manière suivante : A chaque instant, un individu nait avec la probabilité $p \in (0,1)$, ou un individu meurt avec la probabilité $q = 1-p$.

Ecrire le noyau de transition.

Ecrire la chaîne de Markov en termes des variables introduites à l'exercice 2.

Comment corriger la matrice de transition pour qu'il n'y ait pas un nombre négatif d'individus ?

Correction: X_n est le nombre d'individus au temps n, pour $n \in \mathbb{N}$. Au temps $n+1$, on a soit X_n+1 individus, avec probabilité p, ou $X_n - 1$ individus, avec probabilité $q = 1-p$. Donc pour tout x, $Q(x, x+1) = \mathbb{P}(X_{n+1} = X_n + 1 | X_n = x) = 1/2 = \mathbb{P}(X_1 = x+1 | X_0 = x)$, et de même $Q(x, x-1) = 1/2$.

Pour $y \notin \{x-1, x+1\}$, $Q(x,y) = 0$. On a donc une matrice de transition infinie qui ressemble à ça

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

On peut en fait écrire que $X_{n+1} = X_n + R_n$, où R_n vaut soit $+1$ soit -1. On appelle ce type de variable une variable de Rademacher :

$$\mathbb{P}(R_n = \pm 1) = 1/2.$$

Le problème est qu'on peut avoir un nombre négatif d'individus. On suppose que quand il y a 0 individus, la population s'éteint. Cela signifie que $Q(0,0) = 1$ et $Q(0,n) = 0$ pour $n \neq 0$. Il faut donc corriger la MDT en remplaçant la 1re ligne par $(1, 0, 0, \dots)$ (et les nombres négatifs ne font plus partie de l'espace d'état.) Ca nous donne

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

Exercice 14. Soit $k \in \mathbb{N}^*$. Soit $X = (X_n)$ une chaîne de Markov (homogène) de loi initiale μ_0 et de matrice de transition Q. Montrer que X' définie par $X'_n = X_{kn}$ est une chaîne de Markov homogène, et donner sa matrice de transition .

Exercice 15. Soit $X = (X_n)_{n\geqslant 0}$ une chaîne de Markov . Soit $k \in \mathbb{N}$ et $X' = (X'_n)_{n\geqslant 0}$ la suite de variables aléatoires définie par

$$\begin{aligned} X'_0 =& X_k \\ X'_1 =& X_{k+1} \\ X'_2 =& X_{k+2}... \end{aligned}$$

Montrer que X' est une chaîne de Markov de matrice de transition Q. On la note parfois $\tau_k X$, c'est-à-dire que τ_k est l'opérateur de translation, ou de *shift temporel*, agissant sur la chaîne de Markov X.

Exercice 16. Soit $a \in]0, 1[$. On considère une chaîne de Markov dont la matrice de transition est la suivante :

$$\begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$$

Calculer la probabilité de l'événement A_3 de passer pour $n \leq 4$ trois fois par l'état 2

1. Avec $\mu_0 = (1, 0)$,
2. Avec $\mu_0 = (1/2, 1/2)$.

Correction:
Il faut dénombrer tous les chemins possibles qui contiennent trois fois l'état 2. En partant de 1 :
— $x^1 = (1, 1, 2, 2, 2)$
— $x^2 = (1, 2, 1, 2, 2)$
— $x^3 = (1, 2, 2, 1, 2)$
— $x^4 = (1, 2, 2, 2, 1)$
Pour le premier chemin, sa probabilité est, d'après la formule précédente,

$$\mathbb{P}(X = x^1) = Q(1, 1)Q(1, 2)Q(2, 2)Q(2, 2) = a(1-a)a^2 = a^3(1-a).$$

Avec d'autres calculs, on montre

$$\mathbb{P}(A_3) = a^3(1-a) + 2a(1-a)^3 + a^2(1-a)^2.$$

Exercice 17. Une chaîne de Markov avec états $E = \{1,2\}$ a la matrice de transition

$$Q = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$$

pour un nombre $a \in (0,1)$.

1. Diagonaliser Q
2. Calculer Q^n pour $n \geqslant 1$.
3. Calculer la loi de X_n pour tout n, sachant que l'on part de l'état $X_0 = 1$. Donner par exemple $\mathbb{P}_1(X_n = 1)$ la probabilité de retour en 1 en n coups.

Correction:

1. On a $\mu_0 = \delta_1$. La loi de X_n est donnée par $\mu_0 Q^n$, il faut donc calculer la puissance n-ème de Q (c'est la difficulté de ce type d'exercice).
   Pour cela, le plus simple est de diagonaliser la matrice. Comme c'est une matrice stochastique,
   $$Q * 1 = 1 \quad ,$$
   ou $1 = (1,1,...,1)^t$, et $*$ est le produit matriciel ($Q * 1(x) = \sum_y 1Q(x,y) = 1$). Cela veut dire que 1 est vecteur propre associé à la valeur propre 1.
   Comme la matrice Q est symétrique, elle est diagonalisable. La première valeur propre est 1, et la trace est
   $$Tr(Q) = 2a.$$
   Il s'ensuit que la seconde valeur propre est $\lambda = 2a-1$. Pour trouver le vecteur propre associé, résolvons le système $Q*(u,v) = \lambda(u;v)$:
   $$\begin{cases} au + (1-a)v = (2a-1)u \\ (1-a)u + av = (2a-1)v, \end{cases} \quad \begin{cases} (1-a)v = (a-1)u \\ (1-a)u = (a-1)v, \end{cases}$$
   i.e $u = -v$. Le vecteur propre $(1;-1)$ correspond donc à la valeur propre λ.
2. On a $Q = PDP^{-1}$ où
   $$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
   et
   $$P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
3. Calculer $\mathbb{P}_1(X_n = 1)$ revient à calculer $\mathbb{P}(X_n = 1)$ pour la chaîne de Markov de loi initiale $\mu_0 = (1,0)$ et de matrice de transition Q. Donc la loi de X_n est $\mu_0 Q^n = (1,0)PD^nP^{-1} = (1/2 + 1/2(2a-1)^n, 1/2 - 1/2(2a-1)^n)$.
   Après n coups, X_n a un tout petit peu plus de chances d'être en 1 qu'en 2, mais ce tout petit peu est en λ^n et s'atténue rapidement avec le temps. Plus a(=probabilité de rester à sa place) est grand, plus ça décroît lentement.

Exercice 18. (exercice 1.1.4 du Norris, p. 9) Une puce saute aléatoirement sur les sommets d'un triangle, sans préférence. Quelle est la probabilité qu'après n sauts la puce soit de retour à son point de départ ?

Remarque : De manière similaire à l'exercice précédent, on devrait trouver 1/3 plus un terme correctif qui décroît (vite) avec le temps.

Recommencer si cette fois la puce saute dans le sens des aiguilles d'une montre avec probabilité 2/3 et dans l'autre sens avec probabilité 1/3.

Exercice 19. Soit E dénombrable et $x \in E$. Soit $X = (X_n)_{n \geqslant 0}$ une chaîne de Markov sur E.

1. Soit $n, p \in \mathbb{N}$. Montrer que l'évènement $\mathbf{A}_n^p$="X passe au moins n fois par l'état x avant le temps p" est mesurable.
2. Soit $n \in \mathbb{N}$. Montrer que l'évènement $\mathbf{A_n}$ ="X passe n fois au moins par x" est mesurable.
3. Montrer que l'évènements $\mathbf{A}$ ="X passe une infinité de fois par x" est mesurable.
4. On suppose $E \subset \mathbb{R}$. Soit $l \in \mathbb{R}$. Soit $\overline{X_n} = \frac{1}{n}\sum_{k=1}^n X_k$. Montrer que l'évènement "$\overline{X_n} \xrightarrow[n\to\infty]{} x$" est mesurable.

Correction:

1. $$\mathbf{A}_n^p = \bigcup_{\{i_1,\ldots,i_n\}\subset\{1,\ldots,p\}} \underbrace{(X_{i_1} = x, \ldots, X_{i_n} = x)}_{\text{mesurable}}$$

   Donc c'est un union fini d'ensembles mesurables.
2. $\mathbf{A}_n = \cap_{p=1}^{\infty} \mathbf{A}_n^p$ intersection dénombrable d'évènements mesurables.
3. $$\mathbf{A} = \cap_{n \geqslant 1} \mathbf{A}_n$$

   intersection dénombrable d'évènements mesurables.

Exercice 20. Un homme possède un stock de bois aléatoire. Il a au jour 0 un nombre X_0 de bûches, et chaque jour il consomme une bûche. On appelle X_n le nombre de bûches à la fin du n-ème jour. Si à un jour n il utilise sa dernière bûche, il renouvelle son stock le lendemain en allant ramasser des bûches, il en ramène un nombre aléatoire $X_n \geq 1$ tel que, pour $y \geq 1$,

$$\mathbb{P}(X_n = y) = f(y),$$

ou

$$f : \mathbb{N}^* \to \mathbb{R}^+$$

satisfait $\sum_{y \geq 1} f(y) = 1$. Le jour du ramassage, il consomme également immédiatement une bûche.

Identifier la chaîne de Markov et donner sa matrice de transition .

(Ce modèle s'appelle modèle de renouvellement, et peut s'appliquer à tout stock qui se renouvelle par un nombre aléatoire dont la loi ne varie pas.)

Exercice 21. On considère une variable aléatoire Y_0 telle que

$$\mathbb{P}(Y_0 = 2) = \frac{1}{3}$$
$$\mathbb{P}(Y_0 = -1) = \frac{2}{3}.$$

Soit $(Y_n)_{n \geqslant 1}$ une famille de variables aléatoires indépendantes et identiquement distribuées, de meme loi que Y_0. Soit $X_0 = 0, X_n = \sum_{k=1}^n Y_k$ pour $n \geqslant 1$.

1. Montrer que $(X_n)_{n \geqslant 0}$ est une chaîne de Markov.
2. Montrer que, pour $n \geqslant 0$, $\mathbb{P}_0(X_n = 0) = 0$ si n n'est pas un multiple de 3.

Correction: Comme $Y_n \equiv 2$ modulo 3 pour tout $n \geqslant 1$, $X_1 = X_0 + Y_0 \equiv 0 + 2 \equiv 2$ modulo 3, $X_2 = X_1 + Y_1 \equiv 2 + 2 \equiv 1$ modulo 3, et $X_3 \equiv 1 + 2 \equiv 0$ modulo 3. Donc X_1 et X_2 ne peuvent etre égaux à 0. On montre par récurrence sur $k \geqslant 1$ que pour tout $k \geqslant 1$, $X_{3k} \equiv 0$ modulo 3, $X_{3k+1} \equiv 2$ modulo 3, $X_{3k+2} \equiv 1$ modulo 3.

3. Soit $k \geqslant 1$, et $n = 3k$.

   (a) Soit un n-uplet $c \in \mathbb{R}^n$ constitué uniquement de $+2$ et de -1 (par exemple : $c = (2, 2, -1, 2)$ ou $(2, -1, -1, -1, 2)$...). Sous quelle condition la somme de tous les éléments de c est-elle égale à 0 ?

   Correction: Soit p le nombre de $+2$ (donc le nombre de -1 est $n - p$). Donc la somme de tous les éléments d est $2p - (n - p)$, cette somme est nulle ssi $3p = n \Leftrightarrow p = k$.

   (b) En déduire le nombre de chemins **possibles** qui vont de 0 à 0 en n coups.

   Correction: Le nombre de chemins qui vont de 0 à 0 (en repassant éventuellement par 0) correspond au nom de n-uplets constitués de $+2$ et -1 dont la somme vaut 0, c'est-à-dire au nombre de n-uplets contiennent k fois 2 et de $n - k$ fois -1. Choisir un tel n-uplet équivaut à choisir le sous-ensem de $\{1, \dots, n\}$ de k éléments qui indique les positions des "+2". Le nombre de ces sous ensembles $C_k^{3k} = \frac{(3k)!}{k!2k!}$.

   (c) Donner un équivalent quand k tend vers l'infini de

   $$\mathbb{P}_0(X_{3k} = 0).$$

   On rappelle la formule de Stirling : $k! \sim_{k \to \infty} k^k e^{-k} \sqrt{2\pi k}$.

   Correction: La probabilité d'un chemin qui va de 0 à 0 en $n = 3k$ coups est la probabilité que k variables Y_i tombent sur $+2$, et $n - k = 2k$ tombent sur -1, c'est donc $(1/3)^k (2/3)^{2k}$

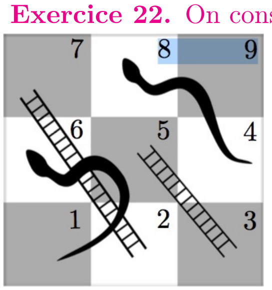
   $$\mathbb{P}(X_{3k} = 0) = (\#\{\text{ chemins qui vont de 0 à 0}\})(1/3)^k (2/3)^{2k}$$
   $$= \frac{2^{2k} C_k^{3k}}{3^{3k}} = \frac{2^{2k}(3k)!}{3^{3k} k!(2k)!} \sim \frac{2^{2k}(3k)^{3k} e^{-3k} \sqrt{6\pi k}}{3^{3k} k^k e^{-k} \sqrt{2\pi k} (2k)^{2k} e^{-2k} \sqrt{4\pi k}} \sim \frac{2^{2k} 3^{3k} k^{3k} \sqrt{6\pi}}{3^{3k} 2^{2k} k^{3k} \sqrt{2\pi \cdot 4\pi} \sqrt{k}} \sim \frac{\sqrt{3}}{\sqrt{2k}}$$

   Comme

   $$\sum_{r > 0} \mathbb{P}_0(X_r = 0) = 0 + \sum_{k \geqslant 1} \mathbb{P}_0(X_{3k} = 0) = \infty,$$

   la chaine est récurrente.

Exercice 22. On considère un jeu d'échelles et serpents qui se joue sur ce plateau :



Les règles sont les suivantes (à 1 joueur) :

— A chaque tour on lance un dé équilibré à 6 faces, et on avance du nombre de cases correspondantes, en s'arrêtant si on arrive sur la case 9.
— Si on arrive en bas d'une échelle, on monte instantanément en haut
— Si on arrive en haut d'un serpent, on descent instantanément en bas de sa queue.
— Dans les autres case de figure, on reste sur la case.

On considère la chaîne de Markov qui donne le position d'un joueur à la fin du tour.

1. Quels sont les états qui vont être effectivement visités par la chaîne de Markov ? Dessiner le graphe de la chaîne de Markov. *Correction:* En réalité, seuls les états 1,4,5,7,9 sont utiles

2. Donner la matrice de transition.

Correction:

$$\begin{pmatrix} 1 \\ 4 \\ 5 \\ 7 \\ 9 \end{pmatrix} \begin{pmatrix} 1/6 & 1/6 & 1/3 & 1/3 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/3 \\ 1/6 & 1/6 & 0 & 1/6 & 1/2 \\ 0 & 1/6 & 0 & 0 & 5/6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Exercice 23. On considère un championnat sportif à 4 équipes, qui se prolonge indéfiniment. A chaque journée, l'équipe 1 affronte une équipe tirée au hasard parmi les 3 autres, et les 2 équipes restantes s'affrontent. Une victoire rapporte 1 points, 1 défaite -1 points. Chaque équipe à 50% de chances de gagner un match. On note $E = \mathbb{Z}^4$. On assimile un état $x = (x_1, x_2, x_3, x_4) \in E$ à la situation du championnat où pour $i = 1, \dots, 4$, l'équipe numéro i a x_i points. On note $Q(x, y)$ la matrice de transition de la chaîne de Markov correspondante.

1. Soit $x, y \in E$. Montrer que si $Q(x, y) > 0$, alors $\sum_{i=1}^4 x_i = \sum_{i=1}^4 y_i$. A-t-on la réciproque ?

   Correction: On appelle δ_i la quantité de laquelle varie le score de x_i après une journée de championnat. Comme deux équipes gagnent et deux équipes perdent, on a 2 δ_i égaux à 1, et 2 égaux à -1. Donc $\sum_{i=1}^4 \delta_i = 0$. Si $Q(x, y) > 0$, alors il est possible de passer de x à y en une journée de championnat, donc

   $$\sum_{i=1}^4 y_i = \sum_{i=1}^4 (x_i + \delta_i) = \sum_{i=1}^4 x_i + \sum_{i=1}^4 \delta_i = \sum_{i=1}^4 x_i.$$

   Si $x = (0, 0, 0, 0)$ et $y = (1, 0, 0, 0)$ par exemple, $\sum_{i=1}^4 x_i = 0 \neq 1 = \sum_{i=1}^4 y$ donc x et y ne sont pas dans la même classe d'équivalence. Il y a donc plusieurs classes de récurrence et la chaîne de Markov n'est pas irréductible.

2. Soit $(x, y) \in E$. Montrer qu'il existe un chemin possible entre x et y ssi $\sum_{i=1}^4 x_i = \sum_{i=1}^4 y_i$.

3. On considère les score a_n et b_n des équipes 1 et 2 après n matchs, pour $n \in \mathbb{N}$, et $Y_n = (a_n, b_n)$, en supposant $a_0 = b_0 = 0$. Montrer que $(Y_n)_{n \geqslant 0}$ est une chaîne de Markov homogène dans $\mathbb{Z}^2$. Donner sa matrice de transition . Montrer que pour 2 couples $x = (a, b), y = (a', b')$ de $\mathbb{Z}^2$, il existe un chemin possible qui va de x à y.

   Correction: Soit $(a, b) \in \mathbb{Z}^2$. On a

   $$Q'((a, b), (a + 1, b + 1)) = Q'((a, b), (a - 1, b - 1)) = \frac{1}{4}.$$

   Comme la somme de toutes les possibilités fait 1, on a

   $$Q'((a, b), (a + 1, b - 1)) = Q'((a, b), (a - 1, b + 1)) = \frac{1}{2}(1 - \frac{1}{4}) = \frac{1}{4}.$$

   Les probabilités de transitions vers les autres états sont nulles. Y_n est la marche aléatoire symmétrique sur $\mathbb{Z}^2$, elle est donc irréductible récurrente (théorème de Polya).

4. On appelle $Z_n = (x_n, y_n, z_n)$ le score des trois premières équipes. matrice de transition ?

   Correction: On appelle A_n l'accroissement.

   $$\mathbb{P}(A_n = (1, 1, -1)) = \mathbb{P}(A_n = (-1, -1, 1)) = 1/4.$$

   On a donc

   $$\begin{aligned} \mathbb{P}(A_n = (1, -1, 1)) =& \mathbb{P}(A_n = (1, -1, -1)) = \mathbb{P}(A_n = (-1, 1, 1)) = \mathbb{P}(A_n = (-1, 1, -1)) \\ =& \frac{1}{4}(1 - 2 \cdot \frac{1}{4}) = \frac{1}{8}. \end{aligned}$$

5. On considère désormais qu'à la première journée, les matchs sont 1 contre 2, et 3 contre 4. A la seconde journée, 1 contre 3 et 2 contre 4, à la 3ème journée, 1 contre 4 et 2 contre 3, puis ça recommence indéfiniment. Qu'est-ce qui change sur les questions précédente ? (réponse en 3 lignes maximum)

   Correction: Dans cette organisation, on n'a plus affaire à une chaîne de Markov homogène, car la matrice de transition dépend de la classe de congruence de n modulo 3.

6. On considère désormais la suite qui contient le classement de la première équipe. Est-ce une chaîne de Markov ?

2 Temps d'absorption

Dans ce chapitre on se pose la question suivante : Etant donné une chaîne X et x dans l'espace d'états E, quel est le temps moyen (éventuellement infini) que met X à arriver au temps x.

Ce temps dépend évidemment de la loi initiale μ_0 : Si $\mu_0 = \delta_x$, le temps d'attente est en moyenne 0. Si la chaîne n'est pas trop compliquée, il est possible de mener des calculs explicites pour trouver ce temps moyen.

2.1 Temps d'arrêt

Avant d'expliquer comment faire, il faut comprendre un fait fondamental sur les chaînes de Markov.

Pour $x \in E$ on définit le temps aléatoire

$$T_x = \min\{n \geq 0 \,:\, X_n = x\},$$

premier moment où la chaîne atteint x. On étend cette définition à toute partie $A \subset E$:

$$T_A = \min\{n \geqslant 0 \,:\, X_n \in A\} = \min\{T_x \,:\, x \in A\}.$$

Définition 5. *Soit X une chaîne de Markov . Soit T une variable aléatoire à valeurs dans $\mathbb{N}$. T est un* **temps d'arrêt** *si pour tout n, l'évènement $(T = n)$ dépend uniquement du passé, c'est-a-dire si l'évènement $(T = n)$ est entièrement déterminé par les variables $X_0, \ldots, X_n$ (c'est-à-dire mesurable par rapport à $\sigma(X_0, \ldots, X_n)$).*

Exemple 4. Pour $x \in E$, le temps T_x est un temps d'arrêt : Pour $n \in \mathbb{N}$,

$$(T_x = n) = (X_0 \neq x, X_2 \neq x, \ldots, X_{n-1} \neq x, X_n = x) = \varphi(X_0, \ldots, X_n)$$

où φ est mesurable. C'est bien un évènement qui est entièrement déterminé si on connait les valeurs de $X_0, \ldots, X_n$.

Exercice 24.
1. $(T = n)$ peut être remplacé par $(T \leq n)$, ou $(T > n)$ dans la définition précédente.
2. Par contre on ne peut pas le remplacer par $(T < n)$. Montrez que $S_x = T_x - 1$ est un contre-exemple.

Correction: Supposons que $(T = n) \in \sigma(X_0, \ldots, X_n)$ pour tout $n \geqslant 1$. Alors

$$(T \leqslant n) = \cup_{0 \leqslant k \leqslant n}(T = k) \in \cup_{0 \leqslant k \leqslant n}\sigma(X_0, \ldots, X_k) \subset \sigma(X_0, \ldots, X_n).$$

Si réciproquement $(T \leqslant n) \in \sigma(X_0, \ldots, X_n)$ pour tout $n \geqslant 1$, alors $(T = n) = (T \leqslant n) \setminus (T \leqslant n-1) \in \sigma(X_0, \ldots, X_n)$ car $(T \leqslant n-1) \in \sigma(X_0, \ldots, X_{n-1}) \subset \sigma(X_0, \ldots, X_n)$. Donc les deux définitions sont équivalentes.

Comme $(T > n) = (T \leqslant n)^c$, $(T \leqslant n) \in \sigma(X_0, \ldots, X_n) \Leftrightarrow (T > n) \in \sigma(X_0, \ldots, X_n)$ donc ces deux définitions sont également équivalentes.

On a bien $(S_x < n) = (T_x \leqslant n) \in \sigma(X_0, \ldots, X_n)$, donc S_x satisfiat à la définition avec $(\ldots < n)$, mais par contre déterminer $(S_x \leqslant n)$ requiert de connaître X_{n+1}, donc $(S_x \leqslant n) \notin \sigma(X_0, \ldots, X_n)$, et S_x n'est pas un temps d'arrêt.

On peut donc donner une autre preuve que T_x est un temps d'arrêt :

$$(T_x > n) = \cap_{k=0}^{n}(X_k \neq x) \in \sigma(X_0, \ldots, X_n)$$

Exercice 25. Soit $R_k, k \geq 0$, des variables iid de Rademacher, c'est-à-dire

$$\begin{cases} \mathbb{P}(R_k = 1) = 1/2, \\ \mathbb{P}(R_k = -1) = 1/2, \end{cases}$$

et X_n leur somme. Soit T le premier temps ou la chaîne est passé deux fois par l'état 10 :

$$T = \min\{n \geqslant 0 : X_n = 10 \text{ et il existe un seul } k < n \text{ tel que } X_k = 10\}.$$

Montrer que T est un temps d'arrêt.

On peut alors élargir la **propriété de Markov** aux temps aléatoires, à condition que ceux-ci soient des temps d'arrêt .

Proposition 6 (propriété de Markov forte). *Soit $k \geq 1$, et T un temps d'arrêt . Pour $x, y \in E$,*

$$\mathbb{P}(X_{T+k} = y | X_T = x) = \mathbb{P}(X_k = y | X_0 = x) = Q^k(x, y).$$

Démonstration:

Prouvons-le pour $T = T_x$, pour un x quelconque de E, et $k = 1$.

$$\mathbb{P}(X_{T_x+1} = y | X_{T_x} = x) = \sum_n \mathbb{P}(X_{T_x+1} = y | X_{T_x} = x, T_x = n)\mathbb{P}(T_x = n). \tag{1}$$

Or $T_x = n$ est équivalent à "$X_0 \neq x, X_1 \neq x, \dots, X_{n-1} \neq x, X_n = x$". On a donc

$$\begin{aligned} &\mathbb{P}(X_{T_x+1} = y | X_{T_x} = x, T_x = n) \\ &\qquad = \mathbb{P}(X_{n+1} = y | X_n = x; X_0 \neq x, X_1 \neq x, \dots, X_{n-1} \neq x) \\ &\qquad = \mathbb{P}(X_{n+1} = y | X_n = x) = Q(x, y) \end{aligned}$$

car "le futur ne dépend que du présent" (**propriété de Markov**). En reportant dans (1),

$$\begin{aligned} \mathbb{P}(X_{T_x+1} = y | X_{T_x} = x) &= \sum_n \mathbb{P}(X_1 = y | X_0 = x)\mathbb{P}(T_x = n) \\ &= Q(x, y) \sum_n \mathbb{P}(T_x = n) = Q(x, y). \end{aligned}$$

En appliquant ce raisonnement itérativement k fois, on obtient le résultat.

Exercice 26. Traiter le cas où T est un temps d'arrêt quelconque.

Si T est un temps d'arrêt quelconque, il faut juste remarquer que l'on peut toujours écrire

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

car $T = n$ ne dépend que du passé (incarné par les variable $X_1, \dots, X_n$), et la propriété de Markov nous dit que connaître le passé est équivalent à connaître la valeur de X_n : $X_n = x$.

Exercice 27. 1. Soit $X_n, n \geqslant 1$ des variables aléatoires iid positives de même loi qu'une variable X. On sait que

$$S_n = \sum_{k=1}^{n} X_k$$

est une chaîne de Markov. Soit T un temps d'arrêt. Montrer que

$$\mathbb{E}(\sum_{k=1}^{T} X_k) = \mathbb{E}(T)\mathbb{E}(X).$$

(Indice : le terme de gauche est égal à $\mathbb{E} \sum_{k=1}^{\infty} X_k 1_{\{T > k-1\}}$)

2. Sans l'hypothèse des X_k positifs, montrez que le résultat est toujours vrai si $E(T) < \infty$ et les X_k sont bornées.
3. Appliquez ces résultats au modèle de population aléatoire avec paramètre p qui s'éteint cette fois si la population atteint -1. Qu'observe-t-on ?
4. On considère désormais une population $X = (X_n)_{n \geqslant 0}$ de même matrice de transition mais tel que le nombre initial d'individus est fixé à $X_0 = k \in \mathbb{N}^*$ (i.e. $\mu_0 = k$). Soit $0 \leqslant j < k$. Soit T_j le temps d'atteinte de $j \leqslant k$ pour X. Soit $X'_n = X_{T_{j+1}+n}$. Soit T'_j le temps d'atteinte de j pour la suite X'. Montrer que $\mathbb{E}_k(T_j - T_{j+1}) = \mathbb{E}(T'_j)$.
5. Quel est le temps moyen que met une population initiale de taille k pour s'éteindre ?

Correction:

1. $$\mathbb{E}(\sum_{k=1}^{T} X_k) = \mathbb{E}\sum_{k=1}^{\infty} X_k 1_{\{T>k-1\}} = \sum_{k=1}^{\infty} \mathbb{E}(X_k 1_{\{T>k-1\}}).$$
   On peut intervertir car toutes les quantiés sont positives. On a
   $$(T > k-1) = (T \leq k-1)^c \in \sigma(X_1, \ldots, X_{k-1}),$$
   donc cet événement est indépendant de X_k, et
   $$\mathbb{E}(\sum_{k=1}^{T} X_k) = \sum_{k=1}^{\infty} \mathbb{E}(X_k)\mathbb{E}(1_{\{T>k-1\}}) = \mathbb{E}(X_1)\mathbb{E}(T).$$
2. Le seul point problèmatique est de justifier l'interversion de $\mathbb{E}$ et $\sum_k$. Pour appliquer le théorème de Fubini, il faut montrer que
   $$\sum_k \mathbb{E}[|X_k \mathbf{1}_{\{T>k-1\}}|] < \infty.$$
   Comme les X_k sont bornées, c'est équivalent à
   $$\sum_k \mathbb{E}[\mathbf{1}_{\{T>k-1\}}] = \mathbb{E}\sum_k \mathbf{1}_{\{T\geqslant k\}} = \mathbb{E}T < \infty.$$
3. Prenons le modèle de population avec paramètre $p \in [0,1]$ et extinction en -1, i.e. $T = T_{-1}$ (car $S_0 = 0$).
   Pour $p = 1/2$, on aurait
   $$\mathbb{E}\sum_{k=1}^{T_{-1}} X_k = \mathbb{E}(X_1)\mathbb{E}(T_{-1}) = 0$$
   car $\mathbb{E}X_1 = 0$. Or, par définition de T_{-1}, $\sum_{k=1}^{T_{-1}} X_k = -1$, ce qui amène à $-1 = 0$. La contradiction vient du fait que l'on n'a pas vérifié $\mathbb{E}T_{-1} < \infty$ (car les X_k ne sont pas négatives). Les hypothèses du résultat précédent ne sont pas vérifiées, c'est-à-dire qu'on a montré $\mathbb{E}T_{-1} = \infty$, le temps moyen d'extinction est infini.
   Pour $p > 1/2$, l'extinction est encore plus longue à arriver, donc on a $\mathbb{E}T_{-1} = \infty$ aussi.
   Pour $p < 1/2$, si l'on peut montrer $\mathbb{E}T_{-1} < \infty$, on a
   $$-1 = \mathbb{E}(T_{-1})\mathbb{E}(X) = \mathbb{E}(T_{-1})(2p-1);$$
   et donc
   $$\mathbb{E}T_{-1} = \frac{1}{1-2p}.$$
   On récupère très facilement le temps d'extinction.

4. D'après la propriété de Markov forte, X' est une chaîne de Markov de même matrice de transition que X. Soit T'_j le premier temps d'atteinte de j par X'. Comme T_{j+1} est le premier temps d'atteinte de $j+1$ par X, et T_j ..., $T'_j = T_{j+1} - T_j$:

   $$\mathbb{E}_k(T_j - T_{j+1}) = \mathbb{E}'_{j+1}(T'_j).$$

   La chaîne de Markov $S'_n = X'_n - j$ a la même loi que S_n. Hors, T'_j est le premier temps d'atteinte en 0 pour S'_n. Donc $\mathbb{E}'_{j+1}(T'_j) = \mathbb{E}_0(T_{-1})$.
5. $\mathbb{E}_k(T_0) = \sum_{j=0}^{k-1} \mathbb{E}(T_j - T_{j+1}) = \sum_{j=0}^{k-1} \mathbb{E}_0(T_{-1}) = k/(1-2p)$ ou $+\infty$.

On pourra utiliser la formulation suivante de la **propriété de Markov**.

Proposition 7. *Pour tout temps d'arrêt T, la chaîne*

$$\tau_T X := X' = (X'_0 = X_T, X'_1 = X_{T+1}, \dots)$$

est une chaîne de Markov dont la matrice de transition est Q et la loi initiale est X_T. De plus, la loi de X' est indépendante de $(X_0, \dots, X_{T-1})$ conditionnellement à X_T.

2.2 Probabilités et temps d'absorptions

Avec le langage introduit dans la section précédente, on s'intéresse pour $A \subset E$ aux quantités

$$h^A = \mathbb{P}(T_A < \infty),$$
$$k^A = \mathbb{E}(T_A).$$

Remarquons que si $h_A \neq 1, k^A = \infty$, donc il faut calculer h^A en premier, et ensuite k^A si ça a du sens. Si l'on conditionne par l'état de départ $x \in E$, on a

$$h_x^A = \mathbb{P}_x(T_A < \infty) \quad \text{Probabilité d'arriver un jour en } A \text{ en partant de } x \text{ ,}$$
$$k_x^A = \mathbb{E}_x(T_A) \quad \text{Temps moyen pour y arriver.}$$

Si $x \in A$, h_x^A est trivialement 1, et T_x^A est trivialement 0. Si $A = \{y\}$ est constitué d'un unique point, on note $h_x^{\{y\}} = h_x^y, k_x^{\{y\}} = k_x^y$.

Théorème 3. *Si $x \notin A$, $(h_x^A)_{x \in E}$ et $(k_x^A)_{x \in E}$ sont solutions des systèmes (éventuellement infinis) d'équations :*

$$h_x^A = \sum_{y \in E} Q(x,y) h_y^A, \text{ pour tout } x \in E$$
$$k_x^A = 1 + \sum_{y \in E} Q(x,y) k_y^A, \text{ pour tout } x \in E.$$

Si le premier (resp. le second) système a plusieurs solutions, $(h_x^A)_x$ (resp.$(k_x^A)_x$) est la plus petite solution positive du système vérifiant $h_x^A = 1$ (reps. $k_x^A = 0$) pour $x \in A$: pour toute solution $(u_x)_{x \in E}, \forall x \in E, h_x^A \leqslant u_x$.

Remarque 5. Comme Q est une matrice stochastique, $h = (1,1,...,1)^t$ est toujours solution du système, mais ici on recherche la plus petite solution.

Traitons un exemple d'utilisation avant d'étudier la preuve :

Exemple 5. Une puce saute sur un triangle, avec une probabilité 2/3 dans le sens horaire, et 1/3 dans le sens anti-horaire.

— On numérote les sommets : 1,2,3. Matrice de transition :

— Intéressons-nous aux temps d'atteinte. Soit $A = \{x_1\}$ le sommet 1. On a pour $i = 2, 3$:

$$k_i^1 = 1 + \sum_{j=1}^{3} Q(i,j)k_j^1.$$

On a donc $k_1^A = 1$ et 2 équations :

$$k_2^1 = 1 + \sum_{j=1}^{3} Q(2,j)k_j^1 = 1 + \frac{1}{3}k_1^1 + 0 + \frac{2}{3}k_3^1 = 1 + \frac{2}{3}k_3^1.$$

$$k_3^1 = 1 + \sum_{j=1}^{3} Q(3,j)k_j^1 = 1 + \frac{2}{3}k_1^1 + \frac{1}{3}k_2^1 + 0 = 1 + \frac{1}{3}k_2^1.$$

Il faut résoudre ce système à 2 équations à 2 inconnues.

Démonstration: Les deux expressions sont obtenues en conditionnant par la valeur de X_1 dans le calcul de $\mathbb{P}(X_n \in A)$:

$$\begin{aligned} h_x^A &= \mathbb{P}(\exists n \geq 1, X_n \in A | X_0 = x) = \sum_{y \in E} \mathbb{P}(\exists n \geq 1, X_n \in A | X_1 = y, X_0 = x)\mathbb{P}(X_1 = y | X_0 = x) \\ &= \sum_{y \in E} \mathbb{P}(\exists n \geq 1, X_n \in A | X_1 = y)Q(x,y) \\ &= \sum_{y \in E} Q(x,y)h_y^A. \end{aligned}$$

De même pour k_x^A :

$$\begin{aligned} k_x^A &= \sum_n n\mathbb{P}_x(T_A = n | X_0 = x) = \sum_n n \sum_{y \in E} \mathbb{P}(T_A = n | X_1 = y)\mathbb{P}(X_1 = y | X_0 = x) \\ &= \sum_y Q(x,y)\mathbb{E}(T_A | X_1 = y) \end{aligned}$$

T_A est le temps que met la chaîne $(X_0 = x, X_1, X_2, \dots)$ pour arriver en A. $T'_A = T_A - 1$ est donc le temps mis par la chaîne $X' = (X'_0 = X_1 = y, X'_1 = X_2, X'_2 = X_3, \dots)$ pour arriver en A (car au lieu de partir au temps 0 on part au temps 1).

On a donc

$$\mathbb{E}(T_A | X_1 = y) = \mathbb{E}(T'_A + 1 | X_1 = y) = 1 + \mathbb{E}_y(T_{A'}).$$

T'_A est le temps moyen mis par la chaîne $X' = (X'_0 = X_1, X'_1 = X_2, \dots)$ pour arriver en A. Ce temps est exactement le même que le chaîne $(X_0, X_1, \dots)$ conditionnée par $X_0 = y$ pour arriver en y. Donc

$$\mathbb{E}_y(T'_A) = k_y^A.$$

En reportant, on a

$$k_x^A = \sum_y Q(x,y)(1 + k_y^A) = 1 + \sum_y Q(x,y)k_y^A$$

car Q est une matrice de transition .

On admet la minimalité dans ce cours.

2.3 Exercices

Exercices

Exercice 28. Soit $R_k, k \geq 0$, des variables iid de Rademacher, c'est-à-dire

$$\begin{cases} \mathbb{P}(R_k = 1) = 1/2, \\ \mathbb{P}(R_k = -1) = 1/2, \end{cases}$$

et X_n leur somme. Soit T le premier temps ou la chaîne est passé deux fois par l'état 10 :

$$T = \min\{n \geqslant 0 : X_n = 10 \text{ et il existe un seul } k < n \text{ tel que } X_k = 10\}.$$

Montrer que T est un temps d'arrêt.

Exercice 29. Soit une population X_n qui s'éteint si $X_n = 0$ et autrement croît (ou décroît) à chaque temps n par

$$\begin{cases} \mathbb{P}(X_{n+1} = X_n + 1) = p, \\ \mathbb{P}(X_{n+1} = X_n - 1) = q = 1 - p. \end{cases}$$

La matrice de transition est

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

et $Q(x, y) = 0$ partout ailleurs.

Le but de l'exercice est de trouver la probabilité d'atteindre 0 en partant de $n \in \mathbb{N}$ (extinction)

1. Trouver une relation entre $h^0_n, h^0_{n-1}, h^0_{n+1}$.
   Rappel sur les suites récurrentes d'ordre 2 : Si (U_n) vérifie $aU_{n+2} + bU_{n+1} + cU_n = 0$ pour tout $n \geq 0$, il faut considérer les racines (éventuellement complexes) x_1 et x_2 du polynôme $aX^2 + bX + c$. Si $x_1 \neq x_2$, U_n est de la forme
   $$U_n = \alpha x_1^n + \beta x_2^n,\ n \geq 0$$
   avec α et β a déterminer en calculant des valeurs particulières de la suite. Si $x_1 = x_2$ est une racine double,
   $$U_n = (\alpha + \beta n)x_1^n.$$
2. En déduire la forme générale de la suite $(h^0_n)_n$.
3. En utilisant le fait que $h^0_n \in [0, 1]$, montrer que si $p \leq 1/2$, $\beta = 0$. Qu'en déduisez-vous pour la population ?
4. Donner la probabilité d'extinction en partant de n si $p > 1/2$.

Exercice 30. Un joueur joue à pile-ou face. La pièce a une probabilité p de tomber sur pile, et $q = 1 - p$ de tomber sur face. Le joueur mise 1$ a chaque fois et parie toujours sur pile.

1. Quel est la probabilité de perdre tout son capital de départ ? (Faire le lien avec l'exo précédent).

2. Le joueur décide que si jamais il atteint 10$, il repart avec son pactole. Quelle est la probabilité qu'il y parvienne ?

Exercice 31. Un joueur a 2\$ et a besoin rapidement d'avoir 10\$. Il peut jouer son argent à pile ou face selon les règles suivantes : Si il gagne, il double sa mise, et si il perd, il perd sa mise. Il décide d'adopter une stratégie où il joue tout son argent à chaque fois si il a moins de 5\$, et sinon il ne joue que la quantité nécessaire pour arriver à 10\$.

1. Identifier la chaîne de Markov.
2. Donner le graphe de la chaîne de Markov.
3. Quelle est la probabilité pour le joueur d'atteindre son objectif ?
4. Quel temps mettra-t-il en moyenne pour que le jeu s'arrête (soit parce qu'il perd tout soit parce qu'il atteint 10\$) ?

Exercice 32. On considère une puce qui saute sur les 8 sommets d'un cube en 3 dimensions. A chaque saut, elle choisit arbitrairement entre tous les sommets adjacents. On appelle $X = (X_n)$ la chaîne de Markov correspondante.

1. Combien a-t-elle de choix à chaque saut ? Donner la matrice de transition .
2. On appelle arbitrairement O l'un des sommets. On appelle $C_i, i \geqslant 0$, l'ensemble des sommets qui sont atteignables en i sauts minimum. On impose $C_0 = O$.
   (a) Décrire les ensembles C_i.
   (b) Soit $I_n, n \in \mathbb{N}$ défini par $X_n \in C_{I_n}$. Montrer que $(I_n)_{n\in\mathbb{N}}$ est une chaîne de Markov et donner sa matrice de transition .
   (c) Calculer, pour la chaîne de Markov X, le temps moyen requis pour atteindre le sommet complètement opposé.

Exercice 33. Pour la population de l'exercice 29, on souhaite calculer la distribution du temps d'extinction T_0 en partant de 1 individu. On note plus généralement T_j le premier temps de passage à j.

La méthode la plus efficace consiste à considérer la fonction caractéristique

$$\varphi(s) = \mathbb{E}_1(s^{T_0}) = \sum_{n\geq 0} s^n \mathbb{P}_1(T_0 = n),\ 0 \leq s < 1.$$

1. Montrer que

$$\mathbb{E}_2(s^{T_0}) = \mathbb{E}_2(s^{T_1})^2 = \varphi(s)^2.$$

On pourra introduire

$$T_{2\to 1} = \min\{n : X_n = 1\}$$
$$T_{1\to 0} = \min\{n : X_n = 0\} - T_{2\to 1}$$

et montrer que ces temps sont indépendants.

2. Montrer que

$$\mathbb{E}_1(s^{T_0} | X_1 = 2) = \mathbb{E}_2(s^{1+T_0}).$$

3. En déduire que pour tout s $\varphi(s)$ vérifie la relation

$$ps\varphi(s)^2 - \varphi(s) + qs = 0.$$

On pourra conditionner par la valeur de X_1.

4. Montrer que

$$\mathbb{E}_1(T_0) = \lim_{s\uparrow 1} \varphi'(s).$$

En déduire la valeur du temps moyen d'extinction lorsqu'il est fini (on peut utiliser les résultats de l'exercice 29).

3 Classification des états

On s'intéresse à la "dynamique" des chaînes de Markov dans ce chapitre. On dit qu'un état $x \in E$ mène à un état $y \in E$ si

$$\mathbb{P}_x(\exists n \geq 0, X_n = y) > 0 \quad \Leftrightarrow Q^n(x,y) > 0$$

La probabilité de passer par y après être passée par x est non-nulle. Cela équivaut à dire qu'il y a des **chemins possibles entre** x **et** y, c'est-à-dire qu'il existe $x_0 = x, x_1 \in E \ldots, x_{n-1} \in E, x_n = y$ tels que

$$Q(x_0, x_1) > 0, \ldots, Q(x_{n-1}, x_n) > 0.$$

On note dans ce cas

$$x \rightsquigarrow y.$$

Si $x \rightsquigarrow y$ et $y \rightsquigarrow x$, on note

$$x \leftrightsquigarrow y$$

et on dit que x et y communiquent.

Exemple 6. Dans l'exemple d'une population, en général, pour tout $n \geq 1$, $n \rightsquigarrow 0$, mais $0 \rightsquigarrow n$ pour aucun n excepté 0. On a aussi $n \leftrightsquigarrow m$ pour tous $n, m > 0$.

On dit que $\{0\}$ est un état absorbant.

Cette propriété d'existence d'un chemin est en fait nécessaire : Si aucun chemin possible ne mène de x à y, alors on n'a aucune chance d'atteindre y.

D'une manière différente, on sait que la loi de X_n en partant de x est $Q^n(x,y)$ (proposition 4). Donc en notant $Q^n(x,y)$ les coefficients de la matrice de transition Q^n, $\boxed{x \rightsquigarrow y \text{ ssi } \exists n \geq 0 \text{ tel que } Q^n(x,y) > 0}$.

Théorème 4. *La relation* $\leftrightsquigarrow$ *est une relation d'équivalence et on peut partitionner* E *par l'ensemble des classes d'équivalences*

$$E = \cup_{x \in E} C_x$$

avec $C_x = C_y$ *si* $x \leftrightsquigarrow y$*, et* $C_x \cap C_y = \emptyset$ *sinon. Par convention,* $x \in C_x$ *pour tout* $x \in E$.

Démonstration: Prouvons que si deux états x et y ne communiquent pas, alors $C_x \cap C_y = \emptyset$. En effet, s'il existe $z \in C_x \cap C_y$, alors il existe un chemin de x à z car $x \rightsquigarrow z$ et il existe un chemin de z à y. En mettant ces chemins bout a bout, on obtient un chemin $x \rightsquigarrow z$. En raisonnant en partant de y, on trouve aussi un chemin $y \rightsquigarrow x$; on a donc $x \leftrightsquigarrow y$, contradiction.

Cela revient en fait à montrer la transitivité et la réflexivité de $\leftrightsquigarrow$ (nécessaires d'après la définition) pour montrer que c'est une relation d'équivalence.

Exercice 34. Un sondeur se déplace au hasard dans un immeuble, sonnant parfois plusieurs fois chez la même personne. Cet immeuble comporte trois étages, et il n'y a qu'un ascenseur, qui de surcroît ne peut que monter. Trouvez la chaîne de Markov, et déterminer ses classes d'équivalences. *Correction:* $\{1\}, \{2\}, \{3\}$

Exercice 35. Une chaîne de Markov a la matrice de transition

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

Quelles sont ses classes d'équivalence ?

On pourra dessiner un diagramme expliquant comment passer d'un état à l'autre. *Correction:* $\{1,2,3\}, \{4\}, \{5\}, \{6\}$.

3.1 Récurrence et transience

Définition 6. *On rappelle que T_x est le temps de 1er passage en x. Pour $r \geq 0$, on note*

$$T_x^{(r)} \text{ le temps de } r\text{-ème } \underline{\text{retour}} \text{ en } x \text{ (ou } (r+1)\text{-ème } \underline{\text{passage}}\text{)},$$

défini par récurrence par

$$T_x^{(0)} = T_x; \quad T_x^{(r+1)} = \inf\{n > T_x^{(r)} : X_n = x\}.$$

Définition 7. *Un état x est dit* **récurrent** *si la probabilité de retour est 1, c'est-à-dire si*

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

Si cette propriété n'est pas vérifiée, on dit que l'état est **transient***.*

Exercice 36. Soit $p \in]0,1[$. On considère la marché aléatoire sur $\mathbb{Z}$:

$$S_n = \sum_{k=1}^{n} X_k$$

où les X_k sont iid de loi

$$\mathbb{P}(X_k = 1) = p, \; \mathbb{P}(X_k = -1) = q := 1 - p.$$

1. 0 est-il récurrent ou transient ?
2. Quelles sont les classes d'équivalences ?

Correction: Si $p = 1/2$, d'après le modèle de population, on a $\mathbb{P}_1(T_0 < \infty) = 1$. De manière symétrique, on a $\mathbb{P}_{-1}(T_0 < \infty) = 1$ (il suffit de considérer la chaîne $S'_n = -S_n$). Donc

$$\begin{aligned}\mathbb{P}_0(T_0^{(1)} < \infty) =& \mathbb{P}_0(T_0^{(1)} < \infty | X_1 = 1)\mathbb{P}(X_1 = 1) + \mathbb{P}_0(T_0^{(1)} < \infty | X_1 = -1)\mathbb{P}(X_1 = -1) \\ =& \mathbb{P}_1(T_0 < \infty)\frac{1}{2} + \mathbb{P}_{-1}(T_0 < \infty)\frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1.\end{aligned}$$

Donc 0 est récurrent.

Si $p > 1/2$, $\mathbb{P}_1(T_0 < \infty) < 1$. Donc

$$\mathbb{P}_0(T_0^{(1)} < \infty) = p\mathbb{P}_1(T_0 < \infty) + q\mathbb{P}_{-1}(T_0 < \infty) < p + q = 1.$$

Donc 0 n'est pas récurrent. Idem si $p < 1/2$ (avec $S'_n = -S_n$).

Exercice 37. Dans l'exercice 34, quels sont les points récurrents ? Transients ?

Proposition 8. *Pour tous $x \in E, r \geq 0$, $T_x^{(r)}$ est un temps d'arrêt.*

Démonstration: On peut commencer par réfléchir à ce que ça veut dire et voir que c'est assez évident : "au temps n, je peux dire si oui ou non je suis passé r fois par x"...Plus formellement :

Il suffit de montrer que pour $n \geq 0$ l'événement $(T_x^{(r)} = n)$ est déterminé par ce qui se passe avant n :

$$\begin{aligned}T_x^{(r)} = n \Leftrightarrow& \exists \text{ une sous-partie } P \subseteq \{1, \ldots, n-1\} \text{ de cardinal } |P| = r \\ & \text{telle que } (X_i = x) \text{ pour } i \in P \text{ et } X_i \neq x \text{ pour } i \notin P \text{ et } X_n = x.\end{aligned}$$

Donc

$$(T_x^{(r)} = n) = \bigcup_{P\subseteq\{1,\dots,n-1\},|P|=r} (X_n = x;\, X_i = x, i \in P; X_i \neq x, i \notin P),$$

On pouvait le montrer par récurrence en remarquant

$$(T_x^{(r)} = n) = \cup_{0\le k<n}(T_x^{(r-1)} = k; X_m \neq x, k < m < n; X_n = x).$$

On appelle r-ème excursion

$$S_x^{(r)} = T_x^{(r+1)} - T_x^{(r)}$$

le temps passé loin de x entre le r-ème et le $r+1$-ème retour.

Proposition 9. *La loi de $S_x^{(r)}$ ne dépend pas de r : Pour tout $k \ge 1$*

$$\mathbb{P}(S_x^{(r)} = k) = \mathbb{P}(S_x^{(0)} = k) = \mathbb{P}_x(T_x = k).$$

De plus, $S_x^{(r)}$ est indépendante de $(X_k; k \le T_x^{(r)})$. Les $S_x^{(r)}, r \geqslant 0$, forment donc une suite de variables IID (indépendantes et identiquement distribuées).

Démonstration: En appliquant la **propriété de Markov forte** , la chaîne

$$X^{(r)} = (X_0^{(r)} = X_{T_x^{(r)}} = x, X_1^{(r)} = X_{T_x^{(r)}+1}, X_2^{(r)} = X_{T_x^{(r)}+2}, \dots)$$

est est une chaîne de Markov de matrice de transition Q qui démarre de $X_0^{(r)} = x$ (donc de loi initiale $\mu_0 = \delta_x$), et qui est indépendante de $(X_k)_{k\leqslant T_x^{(r)}}$. Donc $S_x^{(r)} \in \sigma(X')$ est indépendante des $S_x^{(k)}, k < r$ qui sont mesurables par rapport aux $X_k, k < T_x^{(r)}$. Il reste à montrer qu'ils ont la même distribution.

On peut considérer simultanément plusieurs évènements du futur :

$$\mathbb{P}(X_{T_x^{(r)}+1} = y_1, X_{T_x^{(r)}+2} = y_2, \dots, X_{T_x^{(r)}+k} = y_k | X_{T_x^{(r)}} = x) = \mathbb{P}_x(X_1 = y_1, \dots, X_k = y_k).$$

En faisant la somme sur tous les k-uplets $(y_1, \dots, y_k)$ tels que $y_k = x, y_i \neq x$ pour $i < k$, on a

$$\mathbb{P}(S_x^{(r)} = k) = \mathbb{P}(T_x = k),\ k \ge 0.$$

Une autre manière de voir les choses est de considérer le nombre de visites en un point x après 1 sachant $X_0 = x$:

$$V_x = \#\{n \ge 1 :\ X_n = x\} = \sum_{n=1}^{\infty} \mathbf{1}_{X_n = x}.$$

Exercice 38. Soit $p_r = \mathbb{P}_x(V_x > r), r \in \mathbb{N}$. Montrer que $p_r = p_1^r$.

Correction:

On peut en fait calculer la loi de V_x par récurrence. C'est évident pour $r = 0$. Supposons donc $p_r = p_1^r$. Remarquons que $(V_x > r)$ équivaut à $(T_x^{(r)} < \infty)$ (le r-ème retout, i.e. le $r+1$-ème passage, survient en un temps fini). Calculons p_r par récurrence :

$$\begin{aligned}
p_{r+1} = \mathbb{P}(T_x^{(r+1)} < \infty) &= \mathbb{P}(T_x^{(r)} < \infty \text{ et } T_x^{(r+1)} - T_x^{(r)} < \infty) \\
&= \mathbb{P}(T_x^{(r)} < \infty \text{ et } S_x^{(r)} < \infty) \\
&= \mathbb{P}(S_x^{(r)} < \infty | T_x^{(r)} < \infty)\mathbb{P}_x(T_x^{(r)} < \infty) \\
&= \mathbb{P}(T_x^{(1)} < \infty)\mathbb{P}(T_x^{(r)} < \infty) \\
&= p_1 p_r = p_1 p_1^r = p_1^{r+1}.
\end{aligned}$$

Proposition 10. *Soit X une chaîne de Markov irréductible. Les assertions suivantes sont équivalents :*

- *(i)$x \in E$ est récurrent*
- *(ii)si $X_0 = x$ p.s.,*
$$V_x = \infty \quad p.s.$$
- *(iii)*
$$\sum_{n \geq 0} Q^n(x, x) = \infty.$$

Démonstration:

Lemme 1. *Soit V une variable aléatoire à valeurs dans $\mathbb{N}$. Alors*
$$\mathbb{E}(V) = \sum_{r \in \mathbb{N}} \mathbb{P}(V > r).$$

Donc
$$\mathbb{E}(V_x) = \sum_r p_1^r.$$

On en déduit, comme $p_1 \in [0, 1]$,
$$x \text{ récurrent} \Leftrightarrow p_1 = 1 \Leftrightarrow \sum p_1^r = \infty \Leftrightarrow \mathbb{E}_x(V_x) = \infty$$

On a également $p_1 = 1 \Leftrightarrow \lim_{r \to \infty} p_1^r = 1 \Leftrightarrow \lim_{r \to \infty} \mathbb{P}(V > r) = 1 \Leftrightarrow \mathbb{P}_x(V_x = \infty) = 1$. Donc (i) $\Leftrightarrow$ (iii), et au passage $V_x = \infty$ p.s. est équivalent à $\mathbb{E}_x(V_x) = \infty$.

Remarquons que par le théorème de Fubini/Beppo-Levi
$$\mathbb{E}_x V_x = \mathbb{E}_x(\sum_{n \geq 1} \mathbf{1}_{X_n = x}) = \sum_{n \geq 1} \mathbb{E}_x(\mathbf{1}_{X_n = x}) = \sum_{n \geq 1} \mathbb{P}_x(X_n = x) = \sum_{n \geq 1} Q^n(x, x).$$

En particulier, cela implique que $V_x < \infty$ p.s. ssi $\mathbb{E}_x(V_x) = \infty$ ssi $\sum_n Q^n(x, x) < \infty$.

Exercice 39. Marche aléatoire sur $\mathbb{Z}$

Soit $X_n; n \geq 1$ des variables aléatoires iid de Rademacher (de loi
$$\begin{cases} \mathbb{P}(X_1 = 1) = 1/2, \\ \mathbb{P}(X_1 = -1) = 1/2. \end{cases}$$

On définit la Marche aléatoire sur $\mathbb{Z}$
$$\begin{cases} S_0 = 0; \\ S_n = S_{n-1} + X_n,\ n \geq 1. \end{cases}$$

Le but de l'exercice est de montrer directement que la marche aléatoire est récurrente ; i.e. tout état est récurrent. Nous utiliserons 2 méthodes.

1. Montrer que pour tout x, x est récurrent ssi 0 est récurrent.
2. On veut calculer explicitement la quantité $Q^n(0, 0) = \mathbb{P}_0(X_n = 0)$.
   (a) Montrer que $Q^n(0, 0) = 0$ si n est impair

(b) En utilisant le fait que tout chemin menant de x à x en n coups est constitué d'autant de "+1" que de "-1", montrer que le nombre de tels chemins est

$$\Gamma_{n,0\to 0} = \begin{cases} 0 \text{ si } n \text{ est impair,} \\ C^n_{n/2} = \begin{pmatrix} n \\ n/2 \end{pmatrix} \text{ si } n \text{ est pair.} \end{cases}$$

(c) En montrant que chaque chemin a une probabilité 2^{-n}, donner une expression explicite de $Q^n(x,x)$.

(d) En déduire le résultat. On pourra utiliser la formule de Stirling

$$n! \sim_{n\to\infty} n^n e^{-n}\sqrt{2\pi n}.$$

(e) (A rendre pour la semaine prochaine) On considère la même chaîne de Markov avec cette fois $\mathbb{P}(X_k = 1) = p, \mathbb{P}(X_k = -1) = q := 1 - p$, où $p \in]0,1[$, et $p \neq 1/2$. Montrer par la même méthode que 0 est transient.

Correction:
2-d)

$$Q^n(x,x) = 2^{-n}C_n^{n/2} = 2^{-n}\frac{n!}{(n/2)!^2} \sim \frac{2^{-n}n^n e^{-n}\sqrt{2\pi n}}{(n/2)^{2(n/2)}e^{-2n/2}2\pi n/2} \sim 2^{-n}\left(\frac{n}{n/2}\right)^n \frac{2}{\sqrt{2\pi n}} \sim \frac{\sqrt{2}}{\sqrt{\pi n}}$$

On a $\sum_n Q^n(x,x) = \infty$, donc la chaîne est récurrente en tout x.
3) Si $p \neq 1/2$, on a pour $n = 2k$ pair, avec la meme méthode ;

$$Q^n(0,0) = \begin{pmatrix} k \\ 2k \end{pmatrix} p^k(1-p)^k \sim \frac{C}{\sqrt{k}}(4p(1-p))^k.$$

Une étude de fonction montre que pour $p \neq 1/2$, $4p(1-p) < 1$, et donc la série $\sum_n Q^n(0,0)$ converge, donc la chaîne est transiente.

Théorème 5 (Polya,1921)**.** *La marche aléatoire sur $\mathbb{Z}^d$ est récurrente (en tous points) si $d \leq 2$ et transiente (en tous points) si $d \geq 3$.*

Proposition 11. *Au sein d'une même classe d'équivalence, les états sont soit tous récurrents, soit tous transients. On parle alors de classe récurrente ou de classe transiente.*

Démonstration:
Soit $x \leftrightsquigarrow y$ deux états d'une même classe, avec x transient (donc $\sum_r Q^r(x,x) < \infty$). Il existe alors n, m tels que

$$\mathbb{P}_x(X_m = y) = Q^m(x,y) > 0, \quad \mathbb{P}_y(X_n = x) = Q^n(y,x) > 0.$$

Pour $r \geq 0$, on a

$$\mathbb{P}_x(X_{n+m+r} = x) \geq \mathbb{P}_x(X_m = y; X_{m+r} = y; X_{m+r+n} = x),$$

qui correspond à la probabilité du chemin $(x \xrightarrow{m} y \xrightarrow{r} y \xrightarrow{n} x)$. En décomposant la chaîne selon les temps d'arrivée en x et de r-ème passage en y, on en déduit :

$$\mathbb{P}_x(X_{n+m+r} = x) \geq Q^m(x,y)Q^r(y,y)Q^n(y,x).$$

Donc

$$\sum_r Q^r(y,y) \leq \frac{1}{Q^m(x,y)Q^n(y,x)}\sum_r Q^r(x,x) < \infty.$$

Si x est transient, $\sum_r Q^r(x,x) < \infty$, donc idem pour y. Si y est récurrent, $\sum_r Q^r(y,y) = \infty$, donc idem pour x. En inversant les rôles de x et y on montre l'équivalence.

Remarque 6. Si E est fini, il y a toujours au moins une classe récurrente. (Il est impossible que tous les états n'aient été visités qu'un nombre fini de fois en un temps infini).

Il peut y avoir plusieurs classes récurrentes.

Définition 8. *On dit qu'une chaîne de Markov est* <u>*irréductible*</u> *si il n'y a qu'une seule classe. On parle alors de chaîne récurrente ou de chaîne transiente.*

Exemple 7. La marche aléatoire sur $\mathbb{Z}$ est irréductible. Attention : Il n'y a qu'une seule classe et pourtant cette classe peut être transiente ! (Lorsque $p \neq 1/2$).

3.2 Exercices

Exercices

Exercice 40. Une chaîne de Markov a la matrice de transition

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

Quelles sont ses classes d'équivalence ?

On pourra dessiner un diagramme expliquant comment passer d'un état à l'autre.

Exercice 41. Soit $p \in]0,1[$. On considère la marché aléatoire sur $\mathbb{Z}$:

$$S_n = \sum_{k=1}^{n} X_k$$

où les X_k sont iid de loi

$$\mathbb{P}(X_k = 1) = p,\ \mathbb{P}(X_k = -1) = q := 1 - p.$$

1. 0 est-il récurrent ou transient ?
2. Quelles sont les classes d'équivalences ?

Exercice 42. Soit $p_r = \mathbb{P}_x(V_x > r), r \in \mathbb{N}$. Montrer que $p_r = p_1^r$.

Exercice 43. Marche aléatoire sur $\mathbb{Z}$

Soit $X_n; n \geq 1$ des variables aléatoires iid de Rademacher (de loi

$$\begin{cases} \mathbb{P}(X_1 = 1) = 1/2, \\ \mathbb{P}(X_1 = -1) = 1/2. \end{cases}$$

On définit la Marche aléatoire sur $\mathbb{Z}$

$$\begin{cases} S_0 = 0; \\ S_n = S_{n-1} + X_n,\ n \geq 1. \end{cases}$$

Le but de l'exercice est de montrer directement que la marche aléatoire est récurrente ; i.e. tout état est récurrent. Nous utiliserons 2 méthodes.

1. Montrer que pour tout x, x est récurrent ssi 0 est récurrent.
2. On veut calculer explicitement la quantité $Q^n(0,0) = \mathbb{P}_0(X_n = 0)$.
   (a) Montrer que $Q^n(0,0) = 0$ si n est impair
   (b) En utilisant le fait que tout chemin menant de x à x en n coups est constitué d'autant de "+1" que de "−1", montrer que le nombre de tels chemins est

$$\Gamma_{n,0\to 0} = \begin{cases} 0 \text{ si } n \text{ est impair}, \\ C_{n/2}^{n} = \begin{pmatrix} n \\ n/2 \end{pmatrix} \text{ si } n \text{ est pair}. \end{cases}$$

(c) En montrant que chaque chemin a une probabilité 2^{-n}, donner une expression explicite de $Q^n(x,x)$.

(d) En déduire le résultat. On pourra utiliser la formule de Stirling

$$n! \sim_{n\to\infty} n^n e^{-n}\sqrt{2\pi n}.$$

(e) (A rendre pour la semaine prochaine) On considère la même chaîne de Markov avec cette fois $\mathbb{P}(X_k = 1) = p, \mathbb{P}(X_k = -1) = q := 1-p$, où $p \in]0,1[$, et $p \neq 1/2$. Montrer par la même méthode que 0 est transient.

Exercice 44 (Encore une autre manière de montrer que la marche aléatoire dissymétrique est transiente)**.** Soit $S_n = \sum_{k=0}^n X_k$ où les X_k sont des variables iid de Rademacher dissymétriques, c'est-à-dire $p = \mathbb{P}(X_k = 1) = p \neq 1/2$. Appliquer la loi des grands nombres à S_n et en déduire que la chaîne de Markov est transiente.

Exercice 45. Marche aléatoire sur $\mathbb{Z}^2$. On définit X_n des variables aléatoires iid telles que

$$\mathbb{P}(X_n = (-1,0)) = \mathbb{P}(X_n = (1,0)) = \mathbb{P}(X_n = (0,1)) = \mathbb{P}(X_n = (0,-1)) = 1/4,$$

et $S_0 = 0; S_{n+1} = S_n + X_n$ est la marche aléatoire sur $\mathbb{Z}^2$.

On va montrer que c'est une chaîne de Markov récurrente.

1. On appelle X_n^+ la projection de X_n sur la droite d'équation $y = x$, et X_n^- la projection sur la droite d'équation $y = -x$.
   (a) Montrer que
   $$X_n^+ = \begin{cases} \frac{1}{\sqrt{2}} \text{ avec probabilité } 1/2, \\ \frac{-1}{\sqrt{2}} \text{ avec probabilité } 1/2, \end{cases}$$
   et qu'il en est de même de X_n^-.
   (b) Montrer que X_n^+ et X_n^- sont indépendantes.
2. On appelle $S_n^+ = \sum_{k\leq n} X_n^+$ la projection de S_n sur la droite "$y = x''$ et $S_n^- = \sum_{k\leq n} X_n^-$.
   (a) Quelle est la probabilité $\mathbb{P}_0(S_n^- = S_n^+ = 0)$?
   (b) En déduire que la chaîne est récurrente.
3. Montrer que la marche aléatoire sur $\mathbb{Z}^d$ est transiente pour $d \geq 3$.

Exercice 46. Soit $X = (X_n)$ une chaîne de Markov transiente sur un espace d'états E dénombrable. Soit $x, y \in E$. Montrer que $\mathbb{P}_y(X_n = x) \to 0$.

Exercice 47. Soit X la chaîne de Markov de l'exercice 22 du chapitre 1. Donner les classes transientes et les classes récurrentes.

Exercice 48. Soit X la chaîne de Markov de l'exercice 21 du TD1.

1. Donner les différentes classes d'équivalence.
2. En utilisant les résultats de la question 3, dire si ces classes sont transientes ou récurrentes.

Exercice 49. Soit $X = (X_n), Y = (Y_n)$ et $Z = (Z_n)$ les chaîne de Markov de l'exercice 23.

1. Montrer que X a une infinité de classes d'équivalence.
2. Montrer que Z est irréductible.
3. Montrer que Y a une infinité de classes d'équivalence.

4 Distributions invariantes

Analyse asymptotique d'une chaîne de Markov :

1. Découper les états en classes.
2. On analyse chaque classe séparément :
   — Si la classe est transiente, la chaîne s'échappe "à l'infini", donc asymptotiquement elle n'est nulle part en particulier (ex : Marche aléatoire asymétrique sur $\mathbb{Z}$)...(ou alors elle est passée dans une classe récurrente et ne reviendra plus) : pour $x \in E$, quel que soit l'état de départ $y \in E$, $V_x < \infty$ p.s., $S_x < \infty$ p.s. et $\mathbb{P}_y(X_n = x) \xrightarrow[n\to\infty]{} 0$.
   — Si la chaîne est récurrente, elle visite les états les uns après les autres, en revenant sur ses pas (C'est notamment le cas lorsque la classe est finie). On se pose alors les questions suivantes :
   - Combien de temps passe-t-elle en moyenne dans chaque état ? C'est-à-dire, que vaut
   $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k = x\}} \text{ pour } x \text{ dans une classe récurrente ?}$$
   - Y'a-t-il convergence (pour chaque x, y de la classe) de $\mathbb{P}_y(X_n = x)$? Si oui, la limite dépend-elle de y ?

Pour simplifier, on considère qu'il n'y a qu'une seule classe (quand il y en a plusieurs, on peut utiliser les résultats de la partie 2 pour savoir dans quelle classe on finira et avec quelle probabilité)

Définition 9. *Soit μ une mesure sur E. On dit que μ est invariante pour la chaîne de Markov X de matrice de transition Q si $\mu Q = \mu$.*

μ est une mesure invariante ssi pour tout $x \in E$

$$\mu(x) = \sum_{y\in E} Q(y, x)\mu(y).$$

(Il suffit de regarder la coordonnée x de l'égalité $\mu = \mu Q$.)

C'est équivalent à dire que si X_0 a la loi μ, alors X_1 aussi (et donc par récurrence tous les autres temps aussi) :

Proposition 12. *Si μ_0 la distribution initiale est invariante, alors μ_0 est également la distribution de $X_1, X_2, \ldots,$ c'est-à-dire pour tout $x \in E$*

$$\mathbb{P}_{\mu_0}(X_1 = x) = \mathbb{P}_{\mu_0}(X_2 = x) = \cdots = \mu_0(x).$$

Démonstration: La preuve est évidente quand on se rappelle de la proposition 4 qui stipule que la loi de X_1 est $\mu_0 Q = \mu_0$, celle de X_2 est $\mu_0 Q^2 = (\mu_0 Q)Q = \mu_0 Q = \mu_0$, etc...

Rappel : Une distribution est une mesure dont la masse totale est 1 (c'est-à-dire une mesure pour laquelle la somme des masses de tous les termes vaut 1). On dit aussi une mesure de probabilité, ou juste une probabilité.

Donc une distribution invariante est une mesure invariante dont la masse vaut 1.

La mesure est dite invariante car si le premier état est tiré au hasard selon cette loi μ, alors le second point aura la même loi (mais dans ce cas il ne peut s'agir que de lois de probabilités) :

Exercice 50. On considère une puce qui saute sur les trois sommets d'un triangle. Quand elle est sur le sommet 1, elle saute automatiquement sur le sommet 2. Quand elle est en 2 ou 3, elle choisit un des 2 sommets avec probabilité 1/2.

1. Donner la matrice de transition .
2. Trouver la (les) distribution(s) invariante(s).

Correction: La distribution uniforme est invariante.

Remarque 7. Les équations (où les inconnues sont les $a_x, x \in E$)

$$a_x = \sum_{y \in E} Q(y,x) a_y, x \in E,$$

qui caractérisent les mesures invariantes, sont a ne pas confondre avec les équations

$$a_x = \sum_{y \in E} Q(x,y) a_y,$$

dont la plus petite solution donne les probabilités d'absorption.

L'idée cachée derrière une distribution invariante est la suivante (nous allons la concrétiser dans cette partie) : Le temps moyen passé par la chaîne en un état x est proportionnel à $\mu_0(x)$. De plus dans certains cas, on a la convergence $\mathbb{P}_y(X_n = x) \to \mu_0(x)$ indépendamment de y (ces résultats sont faux dans un cadre général mais on précisera par la suite).

Quelle est la conséquence de l'existence d'une mesure invariante ? Intuitivement, la fréquence de retour à un état x dépend du nombre de points y qui "pointent" vers x, c'est-à-dire tels que $Q(y,x)$ est "grand", pondérés par la probabilité de passer vers ces points y. En d'autres termes, le temps moyen passé en x sera proportionnel à

$$\sum_y \mu_0(y) Q(y,x) = \mu_0(x)$$

car μ_0 est invariante !

Si la mesure de départ n'est pas invariante, on espère qu'on va "converger" vers la distribution invariante. Ce sera l'objet du prochain chapitre.

Exercice 51. Donner toutes les mesures et les probabilités invariantes de la chaîne de Markov qui a pour matrice de transition

$$\begin{pmatrix} a & 1-a \\ 1-a & 0 \end{pmatrix}$$

en fonction de a.

Correction: Soit $\mu = (\alpha, \beta)$ une mesure invariante. Alors

$$(\alpha, \beta) = (\alpha, \beta) Q = (a\alpha + (1-a)\beta, (1-a)\alpha + a\beta).$$

Les solutions sont les vecteurs tels que $\alpha = \beta \geqslant 0$. La seule probabilité invariante est donc bien $(1/2, 1/2)$.

Remarquons aussi que si μ est une mesure invariante, et $t \geqslant 0$, alors $t\mu$ aussi est invariante :

$$Q(t\mu) = t(Q\mu) = t\mu.$$

On dit que μ et $t\mu$ sont égales à une constante près. En particulier, si la masse totale de μ, c'est-a-dire $\mu(E)$, est finie,

$$\pi(x) := \frac{\mu(x)}{\mu(E)}$$

est une probabilité invariante.

Modulo cette liberté, y'a-t-il beaucoup de mesures invariantes ?

Théorème 6 (Admis)**.** *Toute chaîne de Markov irréductible* <u>*récurrente*</u> *admet au plus une* <u>*mesure*</u> *invariante à une constante multiplicative près (et donc au plus une probabilité invariante).*

Exercice 52. 1. Donner une mesure invariante pour la marche aléatoire symétrique sur $\mathbb{Z}$ et montrer qu'il n'existe pas de distribution invariante.

2. Montrer que la marche aléatoire dissymétrique sur $\mathbb{Z}$ admet plus d'une mesure invariante (même à une constante près). Existe-t-il une distribution invariante ?

Correction:

1. Marche aléatoire sur $\mathbb{Z}$.
   Si μ était une distribution invariante sur $\mathbb{Z}$ elle vérifierait
   $$\mu(n) = \frac{1}{2}\mu(n-1) + \frac{1}{2}\mu(n+1),$$
   et donc $\mu(n) = (A + Bn), n \geq 1$ (Voir exo 29). Si $B \neq 0$, il existe n tel que $\mu(n) < 0$, ce qui ne colle pas avec la définition d'une mesure. Donc $B = 0$.
   Si on cherche une distribution invariante π, alors π satisfait la même équation, et donc $\pi(n) = A$. La contrainte $\sum_n \pi(n) = 1 < \infty$ impose $A = 0$, et $\pi = 0$; contradiction.
   Il n'y a donc pas de distribution invariante.
   En revanche, toute mesure constante $\mu(\{x\}) = \mu(\{0\}), x \in \mathbb{Z}$ est invariante.
2. En ce qui concerne la marche aléatoire dissymétrique sur $\mathbb{Z}$, une mesure invariante est une mesure qui vérifie
   $$\mu(k) = \mu(k-1)p + \mu(k+1)(1-p),$$
   et donc
   $$\mu(k) = \alpha + \beta\Big[\frac{p}{1-p}\Big]^k$$
   pour certains nombres $\alpha, \beta \geqslant 0$. Aucune de ces mesures n'est une distribution.

Aucun point n'est "oublié" par une mesure invariante :

Exercice 53. Soit μ une distribution invariante sur une chaîne irréductible. Alors $\mu(x) \neq 0$ pour tout $x \in E$.

Correction: Comme $\sum_{y \in E} \mu(y) = 1$, il existe $y \in E$ pour lequel $\mu(y) \neq 0$. Comme la chaîne est irréductible, $\exists n \geq 0$ tel que $Q^n(y, x) > 0$. Or
$$\mu(x) = \sum_{z \in E} \mu(z) Q^n(z, x) \geq \mu(y) Q^n(y, x) > 0.$$

Il existe toujours une mesure invariante pour une chaîne récurrente. On la construit comme ceci :

Proposition 13. *Soit X une chaîne de Markov IR (irréductible récurrente), et $x \in E$. On appelle μ_x la mesure définie par*
$$\mu_x(y) = \mathbb{E}_x\Big(\sum_{n=1}^{T_x^{(1)}} \mathbf{1}_{X_n = y}\Big), y \in E$$
le nombre moyen de visites en y entre 2 passages en x. Alors pour tout x, μ_x est une mesure invariante. C'est donc la seule à une constante près. Remarquons que $\mu_x(\{x\}) = 1$.

De plus, $0 < \mu_x(y) < \infty$ pour tout $x, y \in E$.

Démonstration: Vérifions que cette mesure est bien invariante.

Soit $y \in E$.

$$
\begin{aligned}
\mu_x(y) &= \mathbb{E}_x(\sum_{n\geq 1} \mathbf{1}_{X_n=y \text{ et } n\leq T_x^{(1)}}) \\
&= (\sum_{n\geq 1} \mathbb{P}_x(X_n = y \text{ et } n \leq T_x^{(1)})) \\
&= \sum_{n\geq 1}\sum_{z\in E} \mathbb{P}_x(X_n = y, X_{n-1} = z, T_x^{(1)} \geq n) \\
&= \sum_{n\geq 1}\sum_{z\in E} \mathbb{P}_x(X_n = y | X_{n-1} = z, T_x^{(1)} \geq n)\mathbb{P}_x(X_{n-1} = z, T_x^{(1)} \geq n) \text{ on conditionne par la valeur de } X_{n-1} \\
&= \sum_{n\geq 1}\sum_{z\in E} \mathbb{P}_x(X_n = y | X_{n-1} = z)\mathbb{P}_x(X_{n-1} = z, T_x^{(1)} \geq n) \\
&\quad \text{car } (T_x^{(1)} \geq n) = (T_x^{(1)} \leq n-1)^c \text{ est mesurable par rapport à } X_0, X_1, \ldots, X_{n-1} \\
&= \sum_{n\geq 1}\sum_{z\in E} Q(z,y)\mathbb{P}_x(X_{n-1} = z, T_x^{(1)} \geq n) \\
&= \sum_{z\in E} Q(z,y) \sum_{n\geq 1} \mathbb{P}_x(X_{n-1} = z, T_x^{(1)} \geq n) \\
&= \sum_{z\in E} Q(z,y) \sum_{n\geq 0} \mathbb{P}_x(X_n = z, T_x^{(1)} \geq n+1) \\
&= \sum_{z\in E} Q(z,y)\mathbb{E}_x(\sum_{n=0}^{T_x^{(1)}-1} \mathbf{1}_{X_n=z}) \\
&= \sum_{z\in E} Q(z,y)\mathbb{E}_x(\sum_{n=1}^{T_x^{(1)}} \mathbf{1}_{X_n=z}) \text{ car } X_0 = z \Leftrightarrow z = x \Leftrightarrow X_{T_x^{(1)}} = z \\
&= \sum_{z\in E} Q(z,y)\mu_x(z)
\end{aligned}
$$

Montrons que $0 < \mu_x(y) < \infty$. Comme x et y communiquent, il existe $n, m > 0$ tels que $Q^n(x,y) > 0, Q^m(x,y) > 0$. Comme la mesure est invariante, $\mu_x = \mu_x Q^n$:

$$\mu_x(y) = (\mu_x Q^n)(y) = \sum_z Q^n(z,y)\mu_x(z) \geqslant Q^n(x,y)\mu_x(x) > 0.$$

De manière similaire,

$$1 = \mu_x(x) = \sum_{z\in E} Q^m(z,x)\mu_x(z) \geqslant Q^m(y,x)\mu_x(y),$$

donc $\mu_x(y) \leqslant (Q^m(y,x))^{-1} < \infty$.

Comme μ_x est une mesure invariante, si sa masse est finie, alors

$$\pi(y) = \frac{\mu_x(y)}{\mu_x(E)}, y \in E.$$

est une distribution invariante (et ne dépend pas de x).

On a

$$\mu_x(E) = \sum_{y\in E} \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} 1_{X_n=y} = \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} \sum_y 1_{X_n=y} = \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} 1 = \mathbb{E}_x T_x^{(1)}.$$

Comme la chaîne est récurrente, on sait que $T_x^{(1)} < \infty$ p.s.. Par contre, rien n'indique que $\mathbb{E}T_x^{(1)} < \infty$. Si c'est le cas, on en déduit

$$\pi(y) = \frac{\mu_x(y)}{\mathbb{E}_x T_x^{(1)}}$$

Avec $x = y$, ça nous donne notamment une relation entre la valeur de la distribution invariante en x et le temps de retour moyen :

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

Théorème 7. *Soit X une chaîne de Markov irréductible récurrente. Alors on a les équivalences suivantes :*

(i)X admet une distribution invariante unique, définie par

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

(ii) Tout état x vérifie

$$\mathbb{E}_x(T_x^{(1)}) < \infty$$

(iii) Un état x le vérifie.

On peut énoncer le même théorème en supposant la chaîne irréductible tout court (et pas irréductible et récurrente), il est toujours vrai car pour une chaîne transiente, $\mathbb{P}(T_x^{(1)} = \infty) > 0$ et donc $\mathbb{E}_x(T_x^{(1)}) = \infty$, et si il y a une distribution invariante il y en a plusieurs (voir exo plus loin).

Si X vérifie la condition (i) (ou de manière équivalente (ii) et (iii)), on dit que X est **récurrente positive**, sinon on dit qu'elle est **récurrente nulle**.

Exemple 8. Comme la marche aléatoire symmétrique sur $\mathbb{Z}$ n'admet pas de distribution stationnaire, elle est récurrent nulle. C'est une autre manière de prouver que $\mathbb{E}T_0 = \infty$.

Le temps moyen de retour pour la marche aléatoire symmétrique sur $\mathbb{Z}^2$ est également infini, c'est donc également une chaîne de Markov irréductible récurrente nulle.

Démonstration: (ii)$\Rightarrow$ (iii) est évident.

$(iii) \Rightarrow (i)$ On a vu que $\mu_x(E) = \mathbb{E}_x(T_x^{(1)})$. Donc si $\mathbb{E}_x(T_x^{(1)}) < \infty$, $\frac{\mu_x}{\mathbb{E}_x(T_x^{(1)})}$ est une distribution invariante . Comme il existe au plus une distribution invariante (à constante près), tout autre distribution invariante π' vérifie $\pi' = C\pi$ pour une certaine constante C, mais la contrainte $\sum_{x\in E} \pi(x) = \sum_{x\in E} \pi'(x) = 1$ impose $C = 1$. Donc $\pi = 1$.

$(i) \Rightarrow (ii)$ Si π est une distribution invariante, $\mu_x = \lambda\pi$ pour un certain $\lambda < \infty$, et donc

$$\mathbb{E}_x(T_x^{(1)}) = \mu_x(E) = \lambda < \infty,$$

pour tout $x \in E$.

Même si c'est théoriquement intéressant, $\mathbb{E}_x T_x^{(1)}$ (ou $\pi_x(y)$) est dur à calculer en pratique : Il faut résoudre un systèmes de $|E|$ équations a $|E|$ inconnues.

Voici un outil plus pratique :

Proposition 14. *On dit qu'une distribution μ est réversible si*

$$\mu(x)Q(x, y) = \mu(y)Q(y, x), x, y \in E,$$

c'est-à-dire $\mathbb{P}_\mu(X_0 = x, X_1 = y) = \mathbb{P}_\mu(X_0 = y, X_1 = x)$. Toute distribution réversible est aussi invariante. Si une telle distribution existe la chaîne de Markov est dite **réversible**.

Démonstration:

$$\mathbb{P}_\mu(X_1 = x) = \sum_y \mathbb{P}_\mu(X_0 = y, X_1 = x) = \sum_y \mathbb{P}_\mu(X_0 = x, X_1 = y) = \mathbb{P}_\mu(X_0 = x).$$

c'est-à-dire $\mu Q = \mu$.

Remarque 8. Une chaîne de Markov est dite **symétrique** si sa matrice de transition Q vérifie $Q(x,y) = Q(y,x)$ pour tous $x, y \in E$. Dans ce cas, toute distribution invariante (s'il en existe) est réversible.

Exercice 54. L'urne d'Ehrenfest N particules sont dans une boite, séparées par un petit trou. A chaque instant une particule passe de la moitié gauche à la moitié droite, ou le contraire. On note X_n le nombre de particules à gauche au temps n. ($N - X_n$ est donc le nombre de particules à droite). La probabilité qu'une particule soit éjectée dépend de la "pression" présente :

$$\mathbb{P}(X_{n+1} = X_n - 1) = X_n/N \quad \text{(pression à gauche)},$$
$$\mathbb{P}(X_{n+1} = X_n + 1) = (N - X_n)/N \quad \text{(pression à droite)}.$$

1. Montrer que c'est un chaîne de Markov irréductible et récurrente.
2. Chercher une éventuelle distribution invariante dans l'urne d'Ehrenfest.

Correction:
Commençons par chercher une mesure réversible. Une telle mesure μ doit vérifier

$$\begin{aligned}\mu(k)Q(k,k+1) &= \mu(k+1)Q(k+1,k), 0 \leq k < N\\ \mu(k)(1-k/N) &= \mu(k+1)(k+1)/N\\ \mu(k+1) &= \frac{N-k}{k+1}\mu(k).\end{aligned}$$

On en déduit par récurrence

$$\mu(k) = \frac{(N-(k-1))(N-(k-2))\dots}{k(k-1)\dots}\mu(1) = \frac{(N-k+1)(N-k+2)\dots}{k(k-1)\dots}\mu(0) = C_N^k \mu(0).$$

Pour être en présence d'une distribution, il faut que la masse totale fasse 1, c'est-à-dire

$$\mu(0)\sum_{k=0}^N C_N^k = \mu(0)(1+1)^N = 1,$$

donc $\mu(0) = 2^{-N}$, donc $\mu = (2^{-N} C_N^k)_{k=0\dots N}$ est une distribution invariante (c'est la loi binomiale).

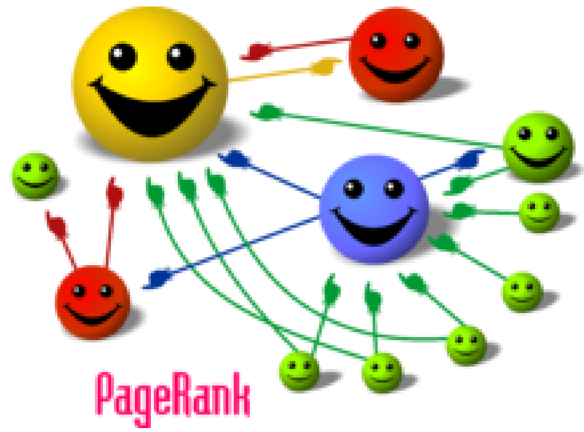
Théorème 8. *Soit X une chaîne de Markov irréductible sur un espace d'états fini. Alors X est IRP.*

Démonstration: Montrons tout d'abord qu'elle est récurrente. Comme E est fini, il existe p.s. un état x tel que $V_x = \infty$ p.s. Donc $\sum_{x \in E} \mathbb{P}(V_x = \infty) > 0$, et donc il existe un état récurrent x, donc tous les états sont récurrents.

Soit $x \in E, \mu_x$ la mesure définie à la proposition 13. Comme $\mu_x(y) < \infty$ pour $y \in E$, $\mu_x(E) = \sum_{y \in E} \mu_x(y) < \infty$, et donc

$$\pi = \frac{\mu_x}{\mu_x(E)}$$

définit une distribution invariante.



Exercice 55 (Moteur de recherche)**.**

On appelle N le nombre de pages web. On considère un internaute qui surfe sur le web en cliquant au hasard sur chaque page web. On note $x \to y$ si une page x pointe vers une page y. On appelle crédit accordé par une page x à une page y la valeur

$$c_{x \to y} = \frac{\mathbf{1}_{x \to y}}{\#\text{liens dans la page } x}$$

1. Quelle condition internet doit vérifier pour que la chaîne de Markov soit irréductible ?
2. On suppose que c'est le cas. Pourquoi y'a-t-il une unique distribution invariante π ? Quelles relations doit vérifier μ ?
3. On propose d'utiliser $\pi(x)$ comme mesure de l'importance d'une page dans les résultats d'un moteur de recherche. Qu'en pensez-vous ? Sans rien prouver, comment feriez-vous pour estimer μ ?

Correction:

1. Il faut que de chaque page x on puisse accéder à chaque page y en un nombre fini de clics.
2. On a une chaîne de Markov irréductible sur un ensemble fini E. Elle est donc irréductible récurrente positive. On a

$$\pi(y) = \sum_{y \in E} \frac{\mathbf{1}_{\{x \to y\}}}{\#\{\text{liens de } x\}} \pi(x).$$

3. L'importance d'une page $(\pi(y))$ est proportionelle à la somme des crédits $(c_{x \to y})$ que lui apportent les autres pages, pondérées par leurs "importances" $(\pi(x))$. Ca semble être une définition acceptable. On a aussi

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

donc on peut estimer $\pi(x)$ en estimant $\mathbb{E}_x(T_x^{(1)})$ en surfant aléatoirement sur le web, et en mesurant différentes réalisation indépendantes du temps pris pour retourner en x.

Exercice 56. On considère une particule qui saute de sommets en sommets sur un cube en trois dimensions. Elle ne peut sauter que sur un sommet adjacent, c'est-à-dire relié par une arête. Elle n'a pas de préférence de direction.

1. Pourquoi y'a-t-il une unique distribution invariante ? Quelle est-elle ?
2. Quel est le temps moyen de retour en un sommet donné ?

3. Soit x et y deux sommets du cubes. En moyenne, combien de temps la puce passe-t-elle en x entre deux passages en y ?
4. * Soit x et y deux sommets opposés du cube (c'est-à-dire pas sur la même face). Quel est le temps moyen pour aller de x à y ? (Indice : utiliser l'exercice 32.)

Correction: 1.chaîne de Markov I avec E fini : récurrente positive.
Un calcul direct montre que l'unique distribution invariante est $(1/8, 1/8, 1/8, 1/8, 1/8, 1/8)$.
2. Donc

$$1/8 = 1/\mathbb{E}_x(T_x)$$

et $\mathbb{E}_x T_x = 8$.
3. On sait que $\frac{\mu_y(x)}{\mathbb{E}_y(T_y)}$ est une (et donc la) distribution invariante π. Donc

$$1/8 = \pi(o) = \mu_y(x)\frac{1}{\mathbb{E}_y(T_y)} = \frac{\mu_y(x)}{8}.$$

Donc $\mu_y(x) = 1$.

Exercice 57 (Remplacement de machines)**.** On modélise le cycle de renouvellement d'une machine par une chaîne de Markov. Au temps 0 on affecte une machine à une certaine fonction. La machine a une probabilité $p_i \in (0, 1)$ de passer de la i-ème à la $i+1$-ème année, et si elle flanche, elle est remplacée par une machine neuve identique.

1. Ecrire le graphe de la chaîne de Markov. Montrer qu'elle est irréductible.

2. (a) Montrer qu'il existe une mesure invariante μ telle que $\mu(0) = 1$ ssi

$$v_N := \prod_{k=0}^{N} p_k \to 0 \text{ quand } N \to \infty.$$

   (b) En utilisant le rappel sur les produits infinis, vérifier que cette condition est vérifiée ssi $\sum_{k\geq 1}(1 - p_k) = \infty$.
   (c) En déduire que si $\sum_{k\geq 0}(1 - p_k) < \infty$, $\mathbb{E}(\text{temps de remplacement}) = \infty$.
3. On suppose qu'il n'y a pas de vieillissement : la probabilité de passer de l'année i à l'année $i + 1$ est la même pour tout i. On note $p \in (0, 1)$ cette probabilité. Quel est le temps moyen de remplacement ? Qu'en déduit-on pour une machine qui vieillit normalement ?

Correction:

1. Soit x, y deux états. Si $x \leqslant y$, $x \to x + 1 \to ... \to y$ est un chemin possible de x vers y, et $y \to 0 \to 1 \to ... \to x$ est un chemin possible de y vers x. Donc x et y communiquent. Donc la chaîne de Markov est irréductible.
2. On note

$$v_N = p_0 \dots p_N, N \geq 0, v_{-1} = 1 \text{ (produit vide)}.$$

a)Remarquons que $v_N \in (0, 1)$ est décroissante et converge donc vers $l \in [0, 1]$.
Une mesure invariante telle que $\mu(0) = 1$ vérifie

$$\mu(0) = 1 = \sum_{k\geq 0}(1 - p_k)\mu(k)$$

$$\mu(k) = p_{k-1}\mu(k - 1), k \geqslant 1, \text{ par récurrence } \mu(k) = p_0 \dots p_{k-1} = v_{k-1}, k \geq 1,$$

$$\mu(0) = 1 = v_{-1} \text{ marche aussi}$$

il est donc nécessaire et suffisant que

$$1 = \sum_{k\geqslant 0}(1-p_k)v_{k-1} = \sum_{k\geq 0} v_{k-1} - v_k = v_{-1} - \lim_N v_N = 1 - l.$$

La condition est donc bien $l = 0$.

AUTRE METHODE : On a

$$\mathbb{P}_0(T_0^{(1)} \geqslant N) = p_0 ... p_{N-1}$$

et par le TCD

$$\mathbb{P}_0(T_0^{(1)} = +\infty) = \lim_N \mathbb{P}_0(T_0^1 \geqslant N) = \lim_N v_N.$$

Donc la chaîne de Markov est récurrente ssi $v_N \to 0$. Si elle est récurrente, elle admet une mesure invariante ν telle que $\nu(x) \neq 0$ pour $x \in E$, et donc $\mu = \nu/\mu(0)$ vérifie la condition demandée. Pour la réciproque il faut faire le raisonnement précédent (l'existence d'une mesure invariante ne garantit pas que la chaîne de Markov soit récurrente).

b) voir ci-après.

c) Dans les notations du cours, le temps de remplacement est $T_0^{(1)}$. On sait que $\mathbb{E}(T_0^{(1)}) < \infty$ est équivalent au fait que la chaîne de Markov soit récurrente positive. Si la chaîne est transiente, il est clair que $\mathbb{E}(T_0^{(1)}) = \infty$. On suppose dans la suite qu'elle est récurrente.

Donc si $\sum_{k=0}^{\infty}(1-p_k) < \infty$, il n'existe pas de mesure invariante telle que $\mu(0) = 1$. Ca implique qu'il n'existe pas de distribution invariante π. Si c'était le cas, la mesure $\mu(x) = \pi(x)/\pi(1)$ vérifie $\mu(1) = 1$ et est invariante (on rappelle qu'une mesure invariante μ sur une chaîne de Markov récurrente vérifie $\mu(x) \neq 0$ pour tout $x \in E$).

En particulier, la chaîne de Markov ne peut pas être récurrente positive, et donc $\mathbb{E}(T_0^{(0)}) = \infty$.

3. Le produit partiel vaut $v_N = p^N$, il converge donc bien vers 0, et la mesure stationnaire telle que $\mu(0) = 1$ vérifie
$$\mu(k) = p^k.$$
La distribution stationnaire est donc
$$\pi(k) = \frac{p^k}{\sum_{k\geq 0} p^k} = p^k(1-p),$$
et le temps moyen de remplacement est $1/\pi(0) = 1/(1-p)$.

On peut penser que dans le cas d'un vieillissement "normal", la probabilité de flancher est plus grande chaque année, c'est-à-dire que p_k décroit avec k. Dans ce cas le temps de remplacement sera encore plus petit que dans le cas géométrique (vieillissement constant), il sera en particulier fini.

Rappel sur les produits infinis.

Théorème 9. *Pour toute suite de nombres $p_n \in (0, 1]$,*

$$\prod_{k=1}^{n} p_k \to l \in [0, 1]$$

avec

$$l = 0 \Leftrightarrow \sum_{k\geq 0}(1-p_k) = \infty.$$

Démonstration:

On pose

$$\Pi_n = \prod_{k=1}^{n} p_k,$$

on s'intéresse à la convergence de la suite Π_n. On a $\Pi_n > 0$ et

$$\log(\Pi_n) = \sum_{k=1}^{n} \log(p_k)$$

et Π_n converge ssi $\log(\Pi_n)$ converge. Il faut donc au moins que $\log(p_k) \to 0$ et donc que $p_k \to 1$.

Remarquons que si p_k ne tend pas vers 1 il existe $\eta < 1$ tel que $p_k \leq \eta$ pour une infinité de k et donc pour tout $q \geq 0$ on a $\Pi_n \leq \eta^q$ pour n suffisamment grand, d'ou $\Pi_n \to 0$.

Si par contre $p_k \to 1$ on pose $q_k = 1 - p_k \to 0$ et on a

$$\log(p_k) = \log(1 - q_k) \sim -q_k$$

et donc les stg $\log(p_k)$ et $-q_k$ ont même nature.

Donc $\Pi_n \to 0$ ssi $\log(\Pi_n) \to -\infty$ ssi $\sum q_k = \infty$.

Bilan

- Une chaîne de Markov irréductible récurrente admet exactement une mesure invariante à une constante près (ex : marche aléatoire symétrique sur $\mathbb{Z}$).
- Une chaîne de Markov irréductible récurrente admet une distribution invariance $\Leftrightarrow$ elle est “récurrente positive” $\Leftrightarrow$ la masse d'une mesure invariante non nulle est finie $\Leftrightarrow$ le temps moyen de retour en un point (n'importe lequel) est fini
- Une chaîne de Markov irréductible transiente peut admettre plusieurs mesures invariantes non-liées par une constante (ex : marche aléatoire dissymétrique sur $\mathbb{Z}$)
- Sur une chaîne de Markov non-irréductible, voir exercice ci-dessous : il existe au moins une mesure irréductible μ par classe récurrente (et $\mu(C) = 0$ pour toute classe transiente C)

4.1 Exercices

exercices

Exercice 58. On considère une puce qui saute sur les trois sommets d'un triangle. Quand elle est sur le sommet 1, elle saute automatiquement sur le sommet 2. Quand elle est en 2 ou 3, elle choisit un des 2 sommets avec probabilité 1/2.

1. Donner la matrice de transition .
2. Trouver la (les) distribution(s) invariante(s).

Correction: La distribution uniforme est invariante.

Exercice 59. chaînes non-irréductibles Soit X une chaîne de Markov à espace d'états E fini, avec des classes récurrentes $R_1, R_2, R_3, \ldots$ et des classes transientes $T_1, T_2, T_3, \ldots$.

1. Montrer que toute mesure invariante μ vérifie $\mu(T_i) = 0$ pour tout i.
2. Comme les classes récurrentes sont fermées (on ne peut pas passer d'une classe récurrente à une autre classe), on peut considérer la chaîne de Markov sur chaque classe séparément. On note μ_i une mesure invariante pour la classe R_i.

   Pour tout i, soit $\lambda_i \geq 0$ un nombre positif. Montrer que

   $$\mu(x) = \begin{cases} \lambda_i \mu_i(x) \text{ si } i \in R_i, \\ 0 \text{ si } x \in T_j \text{ pour un certain } j, \end{cases}$$

   est une mesure invariante.
3. Montrer que toute mesure invariante peut s'écrire de cette manière.
4. En déduire qu'il peut y avoir plusieurs distributions invariantes.

Correction:

1. Soit μ une mesure invariante. On a alors $\mu Q^n = \mu$ par récurrence. Soit $x \in T_i$ un élément d'une classe transiente. On a

$$\mu(x) = \sum_y Q^n(y, x)\mu(y).$$

Pour chaque y, $Q^n(y, x) \leq \mathbb{P}_y(X_n \in T_i) \to 0$ car T_i est une classe récurrente. Donc le membre de droite tend vers 0 et on a bien $\mu(x) = 0$.

2. On écrit la matrice par blocs, en notant Q_i la matrice de transition de la restriction de la chaîne de Markov à la classe R_i. On a par exemples avec 3 classes récurrentes et 2 classes transientes :

$$Q = \begin{pmatrix} Q_1 & 0 & 0 & 0 & 0 \\ 0 & Q_2 & 0 & 0 & 0 \\ 0 & 0 & Q_3 & 0 & 0 \\ ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? \end{pmatrix}$$

Comme la mesure considérée s'écrit

$$\mu = (\lambda_1\mu_1, \lambda_2\mu_2, \lambda_3\mu_3, 0, 0),$$

la multiplication μQ donne

$$\mu Q = (\lambda_1\mu_1 Q_1, \lambda_2\mu_2 Q_2, \lambda_3\mu_3 Q_3, 0, 0) = (\lambda_1\mu_1, \lambda_2\mu_2, \lambda_3\mu_3, 0, 0) = \mu,$$

donc la mesure est invariante.

3. Soit μ une mesure invariante. Alors $\mu(x) = 0$ pour $x \in T_i$ d'après la question 1, et la restriction μ_i de la mesure à chaque R_i est aussi invariante.

4. Supposons par exemple qu'il y ait deux classes récurrentes R_1, R_2, de distributions stationnaires π_1, π_2 (on rappelle que l'espace d'états étant fini, tout chaîne récurrente est récurrente positive et admet donc une distribution invariante). Alors pour $a, b \geq 0$ on pose

$$\pi(x) = \begin{cases} a\pi_1(x) \text{ si } x \in R_1, \\ b\pi_2(x) \text{ si } x \in R_2. \end{cases}$$

D'après 3., c'est une mesure invariante, et sa masse est

$$\pi(E) = \pi(R_1) + \pi(R_2) = a\pi_1(R_1) + b\pi_2(R_2) = a + b.$$

Donc si $a+b = 1$, π est une distribution invariante. Il y a donc une infinités de distributions invariantes.

Exercice 60. Donner toutes les mesures et les probabilités invariantes de la chaîne de Markov qui a pour matrice de transition

$$\begin{pmatrix} a & 1-a \\ 1-a & 0 \end{pmatrix}$$

en fonction de a.

Correction: Soit $\mu = (\alpha, \beta)$ une mesure invariante. Alors

$$(\alpha, \beta) = (\alpha, \beta)Q = (a\alpha + (1-a)\beta, (1-a)\alpha + a\beta).$$

Les solutions sont les vecteurs tels que $\alpha = \beta \geqslant 0$. La seule probabilité invariante est donc bien $(1/2, 1/2)$.

Exercice 61. 1. Donner une mesure invariante pour la marche aléatoire symétrique sur $\mathbb{Z}$ et montrer qu'il n'existe pas de distribution invariante.

2. Montrer que la marche aléatoire dissymétrique sur $\mathbb{Z}$ admet plus d'une mesure invariante (même à une constante près). Existe-t-il une distribution invariante ?

Correction:

1. Marche aléatoire sur $\mathbb{Z}$.
   Si μ était une distribution invariante sur $\mathbb{Z}$ elle vérifierait
   $$\mu(n) = \frac{1}{2}\mu(n-1) + \frac{1}{2}\mu(n+1),$$
   et donc $\mu(n) = (A + Bn), n \geq 1$ (Voir exo 29). Si $B \neq 0$, il existe n tel que $\mu(n) < 0$, ce qui ne colle pas avec la définition d'une mesure. Donc $B = 0$.
   Si on cherche une distribution invariante π, alors π satisfait la même équation, et donc $\pi(n) = A$. La contrainte $\sum_n \pi(n) = 1 < \infty$ impose $A = 0$, et $\pi = 0$; contradiction.
   Il n'y a donc pas de distribution invariante.
   En revanche, toute mesure constante $\mu(\{x\}) = \mu(\{0\}), x \in \mathbb{Z}$ est invariante.
2. En ce qui concerne la marche aléatoire dissymétrique sur $\mathbb{Z}$, une mesure invariante est une mesure qui vérifie
   $$\mu(k) = \mu(k-1)p + \mu(k+1)(1-p),$$
   et donc
   $$\mu(k) = \alpha + \beta\Big[\frac{p}{1-p}\Big]^k$$
   pour certains nombres $\alpha, \beta \geqslant 0$. Aucune de ces mesures n'est une distribution.

Exercice 62. Soit μ une distribution invariante sur une chaîne irréductible. Alors $\mu(x) \neq 0$ pour tout $x \in E$.

Correction: Comme $\sum_{y \in E} \mu(y) = 1$, il existe $y \in E$ pour lequel $\mu(y) \neq 0$. Comme la chaîne est irréductible, $\exists n \geq 0$ tel que $Q^n(y,x) > 0$. Or

$$\mu(x) = \sum_{z \in E} \mu(z) Q^n(z,x) \geq \mu(y) Q^n(y,x) > 0.$$

Exercice 63. L'urne d'Ehrenfest N particules sont dans une boite, séparées par un petit trou. A chaque instant une particule passe de la moitié gauche à la moitié droite, ou le contraire. On note X_n le nombre de particules à gauche au temps n. ($N - X_n$ est donc le nombre de particules à droite). La probabilité qu'une particule soit éjectée dépend de la "pression" présente :

$$\mathbb{P}(X_{n+1} = X_n - 1) = X_n/N \quad \text{(pression à gauche)},$$
$$\mathbb{P}(X_{n+1} = X_n + 1) = (N - X_n)/N \quad \text{(pression à droite)}.$$

1. Montrer que c'est un chaîne de Markov irréductible et récurrente.
2. Chercher une éventuelle distribution invariante dans l'urne d'Ehrenfest.

Correction:

Commençons par chercher une mesure réversible. Une telle mesure μ doit vérifier

$$\begin{aligned} \mu(k)Q(k,k+1) &= \mu(k+1)Q(k+1,k), 0 \leq k < N \\ \mu(k)(1 - k/N) &= \mu(k+1)(k+1)/N \\ \mu(k+1) &= \frac{N-k}{k+1}\mu(k). \end{aligned}$$

On en déduit par récurrence

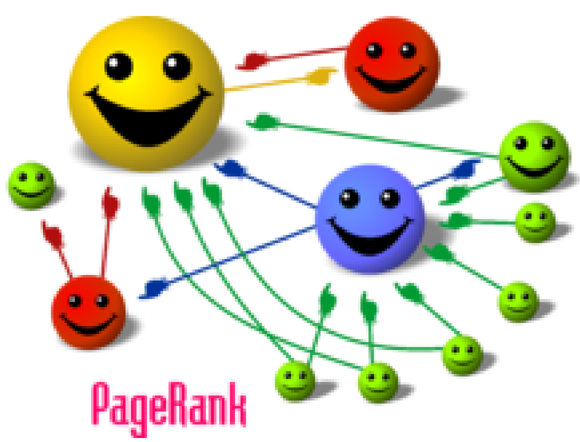
$$\mu(k) = \frac{(N-(k-1))(N-(k-2))\dots}{k(k-1)\dots}\mu(1) = \frac{(N-k+1)(N-k+2)\dots}{k(k-1)\dots}\mu(0) = C_N^k \mu(0).$$

Pour être en présence d'une distribution, il faut que la masse totale fasse 1, c'est-à-dire

$$\mu(0) \sum_{k=0}^N C_N^k = \mu(0)(1+1)^N = 1,$$

donc $\mu(0) = 2^{-N}$, donc $\mu = (2^{-N} C_N^k)_{k=0\dots N}$ est une distribution invariante (c'est la loi binomiale).

Exercice 64.



On appelle N le nombre de pages web. On considère un internaute qui surfe sur le web en cliquant au hasard sur chaque page web. On note $x \to y$ si une page x pointe vers une page y. On appelle crédit accordé par une page x à une page y la valeur

$$c_{x\to y} = \frac{\mathbf{1}_{x\to y}}{\#\text{liens dans la page } x}$$

1. Quelle condition internet doit vérifier pour que la chaîne de Markov soit irréductible ?
2. On suppose que c'est le cas. Pourquoi y'a-t-il une unique distribution invariante π ? Quelles relations doit vérifier μ ?
3. On propose d'utiliser $\pi(x)$ comme mesure de l'importance d'une page dans les résultats d'un moteur de recherche. Qu'en pensez-vous ? Sans rien prouver, comment feriez-vous pour estimer μ ?

Correction:

1. Il faut que de chaque page x on puisse accéder à chaque page y en un nombre fini de clics.
2. On a une chaîne de Markov irréductible sur un ensemble fini E. Elle est donc irréductible récurrente positive. On a
$$\pi(y) = \sum_{y\in E} \frac{\mathbf{1}_{\{x\to y\}}}{\#\{\text{liens de } x\}}\pi(x).$$
3. L'importance d'une page ($\pi(y)$) est proportionelle à la somme des crédits ($c_{x\to y}$) que lui apportent les autres pages, pondérées par leurs "importances" ($\pi(x)$). Ca semble être une définition acceptable. On a aussi
$$\pi(x) = \frac{1}{\mathbb{E}_x(T_x^{(1)})}$$
donc on peut estimer $\pi(x)$ en estimant $\mathbb{E}_x(T_x^{(1)})$ en surfant aléatoirement sur le web, et en mesurant différentes réalisation indépendantes du temps pris pour retourner en x.

Exercice 65. On considère une particule qui saute de sommets en sommets sur un cube en trois dimensions. Elle ne peut sauter que sur un sommet adjacent, c'est-à-dire relié par une arête. Elle n'a pas de préférence de direction.

1. Pourquoi y'a-t-il une unique distribution invariante ? Quelle est-elle ?
2. Quel est le temps moyen de retour en un sommet donné ?
3. Soit x et y deux sommets du cubes. En moyenne, combien de temps la puce passe-t-elle en x entre deux passages en y ?
4. * Soit x et y deux sommets opposés du cube (c'est-à-dire pas sur la même face). Quel est le temps moyen pour aller de x à y ? (Indice : utiliser l'exercice 32.)

Correction: 1.chaîne de Markov I avec E fini : récurrente positive.
Un calcul direct montre que l'unique distribution invariante est $(1/8, 1/8, 1/8, 1/8, 1/8, 1/8)$.
2. Donc
$$1/8 = 1/\mathbb{E}_x(T_x)$$
et $\mathbb{E}_x T_x = 8$.
3. On sait que $\frac{\mu_y(x)}{\mathbb{E}_y(T_y)}$ est une (et donc la) distribution invariante π. Donc
$$1/8 = \pi(o) = \mu_y(x)\frac{1}{\mathbb{E}_y(T_y)} = \frac{\mu_y(x)}{8}.$$
Donc $\mu_y(x) = 1$.

Exercice 66. On modélise le cycle de renouvellement d'une machine par une chaîne de Markov. Au temps 0 on affecte une machine à une certaine fonction. La machine a une probabilité $p_i \in (0,1)$ de passer de la i-ème à la $i+1$-ème année, et si elle flanche, elle est remplacée par une machine neuve identique.

1. Ecrire le graphe de la chaîne de Markov. Montrer qu'elle est irréductible.

2. (a) Montrer qu'il existe une mesure invariante μ telle que $\mu(0) = 1$ ssi

$$v_N := \prod_{k=0}^{N} p_k \to 0 \text{ quand } N \to \infty.$$

   (b) En utilisant le rappel sur les produits infinis, vérifier que cette condition est vérifiée ssi $\sum_{k\geq 1}(1-p_k) = \infty$.
   (c) En déduire que si $\sum_{k\geq 0}(1-p_k) < \infty$, $\mathbb{E}(\text{temps de remplacement}) = \infty$.
3. On suppose qu'il n'y a pas de vieillissement : la probabilité de passer de l'année i à l'année $i+1$ est la même pour tout i. On note $p \in (0,1)$ cette probabilité. Quel est le temps moyen de remplacement ? Qu'en déduit-on pour une machine qui vieillit normalement ?

Correction:

1. Soit x, y deux états. Si $x \leqslant y$, $x \to x+1 \to \ldots \to y$ est un chemin possible de x vers y, et $y \to 0 \to 1 \to \ldots \to x$ est un chemin possible de y vers x. Donc x et y communiquent. Donc la chaîne de Markov est irréductible.
2. On note

$$v_N = p_0 \ldots p_N, N \geq 0, v_{-1} = 1 \text{ (produit vide)}.$$

a)Remarquons que $v_N \in (0,1)$ est décroissante et converge donc vers $l \in [0,1]$.
Une mesure invariante telle que $\mu(0) = 1$ vérifie

$$\mu(0) = 1 = \sum_{k\geq 0}(1-p_k)\mu(k)$$
$$\mu(k) = p_{k-1}\mu(k-1), k \geqslant 1, \text{ par récurrence } \mu(k) = p_0 \ldots p_{k-1} = v_{k-1}, k \geq 1,$$
$$\mu(0) = 1 = v_{-1} \text{ marche aussi}$$

il est donc nécessaire et suffisant que

$$1 = \sum_{k\geqslant 0}(1-p_k)v_{k-1} = \sum_{k\geq 0} v_{k-1} - v_k = v_{-1} - \lim_N v_N = 1 - l.$$

La condition est donc bien $l = 0$.
AUTRE METHODE : On a

$$\mathbb{P}_0(T_0^{(1)} \geqslant N) = p_0 \ldots p_{N-1}$$

et par le TCD

$$\mathbb{P}_0(T_0^{(1)} = +\infty) = \lim_N \mathbb{P}_0(T_0^1 \geqslant N) = \lim_N v_N.$$

Donc la chaîne de Markov est récurrente ssi $v_N \to 0$. Si elle est récurrente, elle admet une mesure invariante ν telle que $\nu(x) \neq 0$ pour $x \in E$, et donc $\mu = \nu/\mu(0)$ vérifie la condition demandée. Pour la réciproque il faut faire le raisonnement précédent (l'existence d'une mesure invariante ne garantit pas que la chaîne de Markov soit récurrente).

b) voir ci-après.

c) Dans les notations du cours, le temps de remplacement est $T_0^{(1)}$. On sait que $\mathbb{E}(T_0^{(1)}) < \infty$ est équivalent au fait que la chaîne de Markov soit récurrente positive. Si la chaîne est transiente, il est clair que $\mathbb{E}(T_0^{(1)}) = \infty$. On suppose dans la suite qu'elle est récurrente.

Donc si $\sum_{k=0}^{\infty}(1-p_k) < \infty$, il n'existe pas de mesure invariante telle que $\mu(0) = 1$. Ca implique qu'il n'existe pas de distribution invariante π. Si c'était le cas, la mesure $\mu(x) = \pi(x)/\pi(1)$ vérifie $\mu(1) = 1$ et est invariante (on rappelle qu'une mesure invariante μ sur une chaîne de Markov récurrente vérifie $\mu(x) \neq 0$ pour tout $x \in E$).

En particulier, la chaîne de Markov ne peut pas être récurrente positive, et donc $\mathbb{E}(T_0^{(0)}) = \infty$.

3. Le produit partiel vaut $v_N = p^N$, il converge donc bien vers 0, et la mesure stationnaire telle que $\mu(0) = 1$ vérifie

$$\mu(k) = p^k.$$

La distribution stationnaire est donc

$$\pi(k) = \frac{p^k}{\sum_{k\geq 0} p^k} = p^k(1-p),$$

et le temps moyen de remplacement est $1/\pi(0) = 1/(1-p)$.

On peut penser que dans le cas d'un vieillissement "normal", la probabilité de flancher est plus grande chaque année, c'est-à-dire que p_k décroit avec k. Dans ce cas le temps de remplacement sera encore plus petit que dans le cas géométrique (vieillissement constant), il sera en particulier fini.

5 Convergence à l'équilibre

Rappel Soit $\mu_n, n \geqslant 0$, une suite de mesures sur E. On dit que la suite $\mu_n, n \geqslant 1$ converge vers une mesure μ, noté $\mu_n \xrightarrow[n\to\infty]{} \mu$ si

$$\forall x \in E, \mu_n(x) \to \mu(x).$$

Proposition 15. *On suppose E fini. On sait que pour $x \in E$, pour tout $n \geq 0$, $\mu_{x,n} = (Q^n(x,y))_{y\in E} = (\mathbb{P}_x(X_n = y))_{y\in E}$ est une mesure de probabilité.*

Si il existe $x \in E$ et une mesure de probabilité π telle que $\mu_{x,n} \to \pi$ alors π est une distribution invariante.

Démonstration: Soit $y \in E$. Alors on conditionne par la valeur de X_n

$$\begin{aligned}
\pi(y) = \lim_n \mathbb{P}_x(X_{n+1} = y) &= \lim_n \sum_{z\in E} \mathbb{P}_x(X_n = z, X_{n+1} = y) \\
&= \lim_n \sum_{z\in E} \mathbb{P}_x(X_n = z)\mathbb{P}_x(X_{n+1} = y|X_n = z) \\
&= \sum_{z\in E} \lim_n \mathbb{P}_x(X_n = z)Q(z,y) \\
&= \sum_z \pi(z)Q(z,y) = (\pi Q)(y).
\end{aligned}$$

C'est en fait vrai pour toute mesure initiale (pas uniquement δ_x).

Exemple 9. On considère la matrice de transition de l'exo 8

$$Q = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$$

On a vu à l'exo 8 que si $a = 1/2$,

$$\begin{aligned}
\mathbb{P}_1(X_n = 1) = 1/2 + \lambda^n/2 \to 1/2, \\
\mathbb{P}_1(X_n = 2) = 1/2 - \lambda^n/2 \to 1/2, \\
\mathbb{P}_2(X_n = 2) = 1/2 + \lambda^n/2 \to 1/2, \\
\mathbb{P}_2(X_n = 1) = 1/2 - \lambda^n/2 \to 1/2,
\end{aligned}$$

ou $\lambda = (2a-1) \in (-1,1)$. donc $\mu = (1/2)\delta_1 + (1/2)\delta_2$ est une mesure invariante. (On pouvait aussi faire le calcul directement...)

Peut-on dire que si π est une distribution invariante, alors la loi de X_n converge vers π ? La réponse est vraie sous certaines hypothèses, mais il faut tout de même exclure certaines situation désagréables...

5.1 Périodicité

Exemple 10 (Urne d'Ehrenfest)**.** L'unique distribution invariante est $\mu(k) = (2^{-N}C_N^k), 0 \leqslant k \leqslant N$.

Pourtant, étant donné un point k, on a $\mu_{k,2n+1}(k) = 0$ car la chaîne de Markov ne peut pas revenir en k en un nombre impair de pas. Donc $\mu_{k,n}(k)$ ne converge pas vers $\mu(k)$ quand $n \to \infty$.

Exemple 11. Soit la matrice de transition

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Le distribution $(1/2)(\delta_1 + \delta_2)$ est invariante, mais comme X_n oscille indéfiniment entre 1 (n pair) et 2 (n impair), sa loi ne peut pas converger...

On dit qu'une chaîne irréductible est périodique de période p si l'on peut décomposer l'espace d'états en une union disjointe

$$E = E_1 \cup \cdots \cup E_p$$

telle que $\mathbb{P}(X_1 \in E_{\overline{k+1}} | X_0 \in E_k) = 1$ (avec $\overline{p+1} = 1, \overline{k} = k$ autrement), et p est le plus grand entier tel que l'on puisse le faire (en effet, c'est toujours vrai pour $p = 1$). Autrement dit, la chaîne saute toujours de E_k vers $E_{\overline{k+1}}$:

$$E_1 \to E_2 \to \cdots \to E_p \to E_1 \to E_2 \ldots$$

Exemple 12. En appelant I les entiers impairs et P les entiers pairs, la marche aléatoire sur $\mathbb{Z}$ est 2-périodique avec la décomposition $\mathbb{Z} = I \cup P$. Il en est de même avec l'urne d'Ehrenfest.

Une chaine qui n'a pas de période est dite **apériodique**.

Théorème 10. *Soit X une chaîne de Markov irréductible. Alors les quatre propositions sont équivalentes*

(i) X a pour période p.
(ii) $\mathrm{pgcd}(\{n : Q^n(x,x) > 0\}) = p$ pour un $x \in E$
(iii) Pour n suffisamment grand, $Q^{pn}(x,x) > 0$ pour un $x \in E$.
(iv) (ii) et (iii) valent pour tout $x \in E$

Démonstration: Dans le cas $p = 1$. (le cas général est identique)

(iii) implique (ii) : Il suffit de voir que pgcd(n,n+1)=1

(ii) implique (i) : Par contraposée, si (i) n'est pas vérifiée $Q^n(x,x) > 0$ impliqe que n est un multiple de p, donc p divise le pgcd, et (ii) n'est pas vérifiée

(ii) implique (iii) : Soit i, j premiers entre eux tels que $Q^i(x,x) > 0, Q^j(x,x) > 0$. Donc il existe des chemins probables $x \xrightarrow{i} x$ $x \xrightarrow{j} x$. On veut montrer que pour n suffisament grand, $Q^n(x,x) > 0$. Pour ce faire, on va construire un "grand" chemin constitué de sous-chemins de taille i ou j. Par exemple pour des entiers $a, b \in \mathbb{N}$, pour $n = ai + bj$, on a le chemin probable de taille n

$$\underbrace{x \xrightarrow{i} x \xrightarrow{i} x \ldots \xrightarrow{i} x}_{a \text{ fois}} \underbrace{\xrightarrow{j} x \xrightarrow{j} x \ldots \xrightarrow{j} x}_{b \text{ fois}}.$$

Si on arrive à montrer pour n assez grand, la propriété $\mathcal{P}(n)$:

$$\exists a, b \in \mathbb{N} : n = ai + bj$$

, alors c'est gagné.

Exercice 67. Soit $i, j \in \mathbb{N}$ premiers entre eux. Alors il existe $n_0 \in \mathbb{N}$ tel que pour $n \geqslant n_0$, $\exists a, b \in \mathbb{N} : n = ai + bj$.

Correction:
Le point de départ est que i et j sont premiers entre eux, et il existe donc $n_0, m_0 \in \mathbb{N}$ tels que $n_0 i = 1 + m_0 j$ (raisonner dans le groupe $(\mathbb{Z}/j\mathbb{Z} \setminus \{0\}, \times)$), ou de manière équivalente il

existe $n_0', m_0' \in \mathcal{B}$ tels que $n_0'j = m_0'i + 1$. On va tenter de montrer $\mathcal{P}(n)$ par récurrence sur $n \geqslant m_0 j + m_0' i$. Si $n = ai + bj$, alors montrons que c'est vrai pour $n+1$: si $b \geqslant m_0$

$$n + 1 = n + n_0 i - m_0 j = (a + n_0)i + (b - m_0)j,$$

et donc on a bien $Q^{n+1}(x,x) > 0$. Si en revanche $b < m_0$, comme $n \geqslant m_0 j + m_0' i$ par hypothèse, nécessairement $a \geqslant m_0'$. Et donc

$$n + 1 = (a - m_0')i + (b + n_0')j.$$

On a bien montré par récurrence $\mathcal{P}(n)$: $\exists a', b' \in \mathbb{N} : n + 1 = a'i + b'j$. On a bien montré par récurrence que pour tout $n \geqslant m_0 j + m_0' i$, n s'écrit $ai + bj$. Donc $Q^n(x,x) > 0$ pour $n \geqslant n_0$.

(iii)$\Leftrightarrow$ (iv) Soit $y \in E$. Soit m, k tels que $Q^k(x,y) > 0, Q^m(y,x) > 0$. Alors pour n suffisament grand, $Q^{n+k+m}(y,y) > 0$. Et vice-versa.

(i) implique (ii) Supposons que le pgcd soit p. On choisit un élément x arbitraire, et on appelle E_1 la classe de tous les y tels qu'il existe un chemin probable de longueur kp de x vers y, où $k \in \mathbb{N}$. On appelle E_2 les $z \in E$ tel que $Q(y,z) > 0$ pour $y \in E_1$, et ainsi de suite. On définit ensuite

$$\tilde{X}_n = k \in \{1, \ldots, p\} \text{ ou } k \text{ est tel que } X_n \in E_k.$$

En pratique, pour montrer qu'une chaîne de Markov est apériodique, on cherche $x \in E$ et deux nombres “petits” m, k et premiers entre eux tels que $Q^m(x,x) > 0, Q^k(x,x) > 0$.

Exercice 68. Soit $q \in \mathbb{N}^*$ Soit $X = (X_n)$ la marche aléatoire symétrique sur $E = \{1, \ldots, q\}$ tel qu'on puisse passer de q à 1 et réciproquement. Est-ce périodique ?

Correction:

— La période n'est jamais de 3 ou plus car la trajectoire $1 \to 2 \to 1$ est possible

— Si q est pair. Passer de q à 1 ou de 1 à q veut dire passer d'un impair à un pair ou réciproquement. Donc on peut encore décomposer la chaîne de Markov en les pairs et les impairs, et elle est 2-périodique.

— Si q est impair, on va montrer que le (ii) du théorème est vérifié. Soit

$$P = \{n : Q^n(1,1) > 0\}.$$

En faisant un tour complet, on voit que $q \in P$. En faisant $1 \to 2 \to 1$ on voit que $2 \in A$. Comme le pgcd de 2 et q est 1, la chaîne est apériodique.

Le prochain théorème nous dit qu'une bonne chaîne de Markov (IRP apériodique) converge vers sa distribution invariante.

Théorème 11 (Convergence à l'équilibre, cas général)**.** *Soit X une chaîne de Markov IRPA. Soit π l'unique distribution invariante. Soit μ_0 la distribution initiale. Alors $X_n \to \pi$ en distribution (quelle que soit μ_0).*

Comme dans toute convergence, il est intéressant de connaître la vitesse, et pour cela il faut introduire une mesure entre deux distribution.

On pose pour μ, μ' des distributions

$$d_{TV}(\mu, \mu') = \frac{1}{2} \sum_{x \in E} |\mu(x) - \mu'(x)|.$$

Cette distance s'appelle la distance en “variation totale”.

Remarque 9.

$$d_{TV}(\mu, \mu') \leqslant \frac{1}{2} \left(\sum_{x \in E} \mu(x) + \sum_{x \in E} \mu'(x) \right) \leqslant \frac{1}{2}(1 + 1) \leqslant 1.$$

On vérifie bien que la topologie héritée de cette distance est plus fine que la topologie que l'on a définie pour la convergence entre mesures : Soit $\mu_n, n \in \mathbb{N}$ une suite de distributions telles que $d_{TV}(\mu_n, \mu) \to 0$. Alors $\mu_n(x) \to \mu(x)$ pour $x \in E$.

Exercice 69. Montrer que

$$d_{TV}(\mu, \mu') = \sup_{A \subset E} |\mu(A) - \mu'(A)|.$$

Condition 1 (Condition de Doeblin)**.** *Il existe $l \in \mathbb{N}^*, \beta > 0$, et une mesure de probabilité μ non-nulle telle que pour tout $x, y \in E$,*

$$Q^l(x, y) > \beta\mu(y).$$

En faisant la somme sur y, on voit que $\beta < 1$.

Théorème 12 (Convergence exponentielle à l'équilibre)**.** *Soit X une chaîne de Markov qui vérifie la condition de Doeblin. Supposons que π soit une distribution invariante pour X. Alors $X_n \to \pi$ en distribution (quelle que soit μ_0).*

De plus il existe $\alpha \in (0, 1)$ tel que

$$d_{TV}(\mu Q^n, \pi) \leqslant \alpha^n.$$

L'idée générale de ce résultat est que pour la distance en variation totale, une chaîne de Markov IRPA est contractante, c'est à dire qu'il existe $\alpha \in (0, 1)$ tel que pour deux distributions μ, μ',

$$d_{TV}(\mu Q, \mu' Q) \leqslant \alpha d_{TV}(\mu, \mu') \tag{2}$$

où Q est la matrice de transition . On rappelle que $\mu_0 Q$ est la loi de X_1 lorsque la loi de X_0 est μ_0. Ainsi, si π est la distribution invariante unique, on montre par récurrence que

$$d_{TV}(\text{loi de } X_n, \pi) = d_{TV}(\mu_0 Q^n, \pi) = d_{TV}(\mu_0 Q^n, \pi Q^n)$$
$$\leqslant \alpha d_{TV}(\mu_0 Q^{n-1}, \pi Q^{n-1}) \leqslant ... \leqslant \alpha^n d_{TV}(\mu_0, \pi) = \alpha^n \frac{1}{2} \sum_x (|\mu_0(x)| + |\pi(x)|) \leqslant \alpha^n,$$

ce qui prouve bien que la loi de X_n converge vers π, et de plus elle le fait à vitesse exponentielle

$$d_{TV}(\mu Q^n, \pi) \leqslant \alpha^n.$$

Proposition 16. *Soit X une chaîne de Markov irréductible apériodique sur un espace E fini. Alors X vérifie la condition de Doeblin.*

Démonstration:

Soit $x \in E$. On sait qu'il existe $n_x \in \mathbb{N}$ tel que $Q^{n_x}(x, x) > 0$ pour tout $n \geqslant n_x$. Pour tout $x, y \in E$, il existe $l_{x,y}$ tel que $Q^{l_{x,y}}(x, y) > 0$. Donc $Q^{n_x + l_{x,y}}(x, y) > 0$ pour tout $n \geqslant n_x$. On pose

$$l = \max_{x,y \in E} n_x + l_{x,y}.$$

On a bien $l = n + l_{x,y}$ pour un $n \geqslant n_x$, et donc $Q^l(x, y) > 0$. Soit désormais

$$0 < \mu(y) < \inf_{x \in E} Q^l(x, y) > 0.$$

et $\beta = M(E), \mu = M/\beta$. Alors

$$Q^l(x, y) \geqslant \beta\mu(y).$$

Démonstration: [Convergence quand la condition de Doeblin est vérifiée]

Commençons par le cas $l = 1$. On a

$$d_{TV}(\mu Q, \mu' Q) = \frac{1}{2} \sum_y \left| \sum_x [\mu(x) - \mu'(x)] Q(x, y) \right|$$

Astuce :

$$\sum_{x \in E} (\mu(x) - \mu'(x)) \beta m(y) = \beta m(y) \sum_{x \in E} \mu(x) - \beta m(y) \sum_{x \in E} \mu'(x) = \beta m(y)(1 - 1) = 0.$$

Donc

$$\begin{aligned} d_{TV}(\mu Q, \mu' Q) &= \frac{1}{2} \sum_y \left| \sum_x (\mu(x) - \mu'(x))(Q(x, y) - \beta m(y)) \right| \\ &\leqslant \frac{1}{2} \sum_x |\mu(x) - \mu'(x)| \sum_y (\underbrace{Q(x, y) - \beta m(y)}_{\geqslant 0 \text{ par Doeblin}}) \\ &= \frac{1}{2} \sum_x |\mu(x) - \mu'(x)|(1 - \beta). \end{aligned}$$

Avec $\alpha = 1 - \beta$, l'application est bien α-contractante.

Dans le cas général, on a besoin d'un lemme.

Exercice 70. Soit $X = (X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice de transition Q. Alors $X'_n := X_{kn}, n \in \mathbb{N}$, est une chaîne de Markov apériodique de matrice de transition $Q' := Q^p$.

Correction: Il suffit de montrer que Q' est stochastique, i.e.avec $v = (1, 1, \ldots, 1)^t$, $Q'v = v$. Comme $Qv = v, Q^p v = v$.

Soit donc X' la chaîne de Markov de loi initiale δ_x (pour n'importe quel $x \in E$) et définie par $\mathbb{P}(X'_{n+1} = y | X'_n = x) = Q^p(x, y)$. Elle est apériodique.

on a la division euclidienne : pour tout $n \in \mathbb{N}, \exists k \in \mathbb{N}, r \in \{0, ..., l-1\}$ tels que $n = kl + r$. Donc pour des distributions μ, μ'

$$d_{TV}(\mu Q^n, \mu' Q^n) = d_{TV}((\mu Q^r)(Q^l)^k, (\mu' Q^r)(Q^l)^k) \leqslant (1 - \beta)^k d_{TV}(\mu Q^r, \mu' Q^r) \leqslant (1 - \beta)^k.$$

Comme $k \geqslant n/l - 1$, $(1-\beta)^k \leqslant (1-\beta)^{n/l-1} \leqslant C[(1-\beta)^{1/l}]^n$ Donc on a le résultat avec $\alpha = (1-\beta)^{1/l}$.

C'est un peu cyclique, mais grâce à la convergence on peut se faire une idée des constantes impliquées dans le condition de Doeblin :

$$Q^n(x, y) \to \pi(y)$$

donc pour n suffisamment grand, et $\beta < 1$,

$$Q^n(x, y) \geqslant \beta \pi(y).$$

La distribution $\mu(y) = \pi(y)$, par exemple, est un bon candidat pour la condition de Doeblin (si on la connaît).

Démonstration: [Convergence pour une chaîne de Markov IRPA]

On va utiliser un couplage : Soit $Y = (Y_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice de transition Q et de distribution initiale π. Alors on définit

$$\mathcal{X}_n = (X_n, Y_n).$$

"Le futur ne dépend que du présent", donc $\mathcal{X}$ est une chaîne de Markov .

Elle est de plus irréductible car pour aller d'un état (x,x') à un état (y,y') on choisit $q > q'$ tels que $\mathbb{P}(x \overset{q}{\to} y) > 0$ et $\mathbb{P}(x' \overset{q'}{\to} y') > 0$. On sait de plus que pour tout $n \geq n_0$, $\mathbb{P}(x \overset{n}{\to} x) > 0$ et $\mathbb{P}(x' \overset{n}{\to} x') > 0$. On a alors pour la chaîne $\mathcal{X}$:

$$\mathbb{P}\begin{pmatrix} x & \overset{n_0}{\to} x & \overset{q}{\to} y \\ x' & \overset{n_0+q-q'}{\to} x' & \overset{q'}{\to} y' \end{pmatrix} > 0$$

c'est-à-dire

$$\mathbb{P}((x,x') \overset{n_0+q}{\to} (y,y')) > 0,$$

donc $\mathcal{X}$ est irréductible.

De plus la distribution $\Pi(x,y) = \pi(x)\pi(y)$ est invariante pour $\mathcal{X}$ car

$$\mathbb{P}_\Pi(\mathcal{X}_1 = (x,y)) = \mathbb{P}_\pi(X_1 = x)\mathbb{P}_\pi(Y_1 = y) = \pi(x)\pi(y)$$

. Le chaîne $\mathcal{X}$ est donc récurrente positive.

On choisit arbitrairement un état $x \in E$ et on pose

$$\mathcal{T} = \min\{n : \mathcal{X}_n = (x,x)\} = \min\{n : X_n = Y_n = x\}.$$

$\mathcal{T}$ est un temps d'arrêt car c'est le temps de retour en (x,x). Comme $\mathcal{X}$ est irréductible récurrente, $\mathcal{T} < \infty$ p.s..

On pose

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

On utilise une astuce pour montrer que $Z = (Z_n)_{n \geq 0}$ est une chaîne de Markov de même matrice de transition que X et Y et de loi initiale $(Z_0 = X_0 = x)$.

Comme $\mathcal{T}$ est un temps d'arrêt, d'après la Propriété de Markov forte,

$$(\mathcal{X}_{\mathcal{T}+n})_{n \geqslant 0}$$

est une chaîne de Markov de même matrice de transition que $\mathcal{X}$, de loi initiale (x,x), et indépendante de $(\mathcal{X}_n, n \leqslant \mathcal{T})$.

Soit $\mathcal{X}' = (Y,X)$ obtenue en échangeant les coordonnées de $\mathcal{X}$. Pour les mêmes raisons, $(\mathcal{X}'_{\mathcal{T}+n}, n \geqslant 0)$ a la même loi que $(\mathcal{X}'_n, n \geqslant 0)$, et est indépendante de $(\mathcal{X}'_n, n \leqslant \mathcal{T})$, et donc de $(\mathcal{X}_n, n \leqslant \mathcal{T})$. Remarquons que $(\mathcal{X}_n, n \leqslant \mathcal{T})$ et $(\mathcal{X}'_n, n \leqslant \mathcal{T})$.

Donc, $\mathcal{X}''$, qu'on construit en collant $(\mathcal{X}_n, n \leqslant \mathcal{T})$, et $(\mathcal{X}'_n, \mathcal{T}+n, n \geqslant 0)$, a la même loi que $\mathcal{X}$. En regardant la 1re coordonnée de cette égalité en loi, on en déduit que Z a la même loi que X.

Donc pour tout n, Z_n a la même loi que X_n (c'est-à-dire $\mathbb{P}_x(Z_n = y) = Q^n(x,y)$). Comme π est invariante et est la loi de Y_0, c'est aussi la loi de Y_n et on a

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

car $\mathcal{T} < \infty$ p.s..

Il suffit d'en savoir plus sur la distribution de $\mathcal{T}$ pour avoir une idée de la vitesse de décroissance.

Conclusion : Une "bonne" chaîne de Markov est une chaîne de Markov IRPA ; car elle admet automatiquement une distribution invariante, et y converge.

Exercice 71. Soit X une chaîne de Markov IRP. On suppose que $X_0 \sim \pi$ suit la loi invariante. On pose

$$\tau = \min\{n \geq 1 : X_n = X_0\}.$$

1.Montrer que $\mathbb{E}(\tau) = |E|$ est le nombre d'états possibles (en particulier $\mathbb{E}(\tau) = \infty$ si il y a une infinité d'états). Est-ce en contradiction avec le fait que pour tout $x \in E$, $\mathbb{E}_x(T_x^{(1)}) = \frac{1}{\pi(x)} < \infty$?

2. En déduire que si E est infini, "de nombreux états ont un grand temps de retour", c'est-à-dire que pour chaque $M > 0$, il y a une infinité de $x \in E$ tels que $\mathbb{E}_x(T_x) \geq M$. (En lien avec $\pi(x) = \mathbb{E}_x(T_x)^{-1}$, ce sont les états les "moins probables" qui ont les plus grands temps de retour). Connaissez-vous un exemple de chaîne de Markov qui rentre dans ce cadre ?

6 chaînes de Markov et simulation

Etant donné une probabilité π, le but de cette section est de proposer des méthodes algorithmiques pour simuler π, c'est-à-dire construire une variable aléatoire de loi π, ou proche de π.

Exemple 13 (Mélange d'un paquet de cartes)**.** On numérote les cartes d'un jeu de 52 cartes de 1 à 52. Un mélange du jeu de carte est donc la donnée des nombres de 1 à 52 dans le désordre, comme par exemple

$$(12, 1, 23, 9, 11, ...).$$

On assimile un jeu "mélangé", ou plus précisément une configuration du jeu, à la bijection $\sigma : \{1, \ldots, 52\} \to \{1, \ldots, 52\}$ qui donne l'ordre des cartes

$$(\sigma(1), \sigma(2), \sigma(3), \ldots, \sigma(52)).$$

L'exemple ci-dessus correspond donc à la bijection σ définie par $\sigma(1) = 12, \sigma(2) = 1, \sigma(3) = 23$, etc...

Les sites de jeu en ligne doivent fournir des jeux de cartes "parfaitement mélangés", c'est-à-dire tels que, à partir d'une configuration donné, on ne puisse absolument rien prédire sur le nouveau jeu mélangé. Mélanger un jeu revient à choisir une permutation σ tel qu'on ne puisse rien prédire sur le jeu $\sigma(1), \sigma(2), \ldots, \sigma(52)$.

La permutation doit donc être aléatoire, c'est-à-dire que c'est une variable aléatoire dans l'ensemble Σ_{52} de toutes les permutations sur un ensemble à 52 éléments. Pour qu'on ne puisse rien prédire sur σ, il faut idéalement que la distribution de σ soit uniforme, c'est-à-dire qu'étant donné une permutation $\sigma_0 \in \Sigma_{52}$,

$$\mathbb{P}(\sigma = \sigma_0) = \frac{1}{\#\Sigma_{52}}.$$

Problèmatique : $\#\Sigma_{52} = 52! \sim 8 \cdot 10^{67}$, il est impossible de tirer un point uniformément avec les ordinateurs actuels.

En appellant π la distribution uniforme sur Σ_{52}, c'est-à-dire $\pi(\sigma_0) = 1/52!$ pour tout $\sigma_0 \in \Sigma_{52}$, on cherche donc une manière de simuler π. C'est très dur à faire exactement, donc on se contentera parfois d'une simulation approximative (i.e. une convergence).

Idée On va chercher une chaîne de Markov $\sigma = (\sigma_n)_{n \geqslant 0}$ dans l'espace d'états $E = \Sigma_{52}$ dont la distribution stationnaire est la distribution uniforme π. On va tenter de s'arranger pour que de plus X soit IRPA. Ainsi, d'après la section précédente, la loi de X_n converge vers π quelle que soit la configuration initiale X_0 du paquet de cartes :

$$d_{TV}(X_n, \pi) \leqslant \alpha^n C$$

pour une certaine constante C. En quelques itérations, la loi de X_n est donc une approximation acceptable de π.

6.1 Algorithme de Metropolis

Etant donné une mesure de probabilité π sur un espace E, le but de l'algorithme de Metropolis est de construire une chaîne de Markov $X = (X_n)$ telle que la loi de X_n converge vers π,

$$X_n \xrightarrow{\mathcal{L}} \pi.$$

On dit que la chaîne de Markov $X = (X_n)$ simule approximativement la loi π.

On suppose sans perte de généralité que $\pi(x) > 0$ sur E (autrement il suffit d'ôter de E les points où π s'annule). Une manière pour approximer π de cette manière est de trouver une matrice stochastique $Q(x, y)$ telle que la chaîne de Markov correspondante soit IRPA et π est invariante pour Q. L'algorithme de Metropolis consiste en les étapes suivantes :

— Construire matrice de transition $P(x, y)$ quelconque telle que la chaîne de Markov correspondante qui vit dans le bon espace d'états E soit irréductible apériodique. Il faut de plus que P soit symétrique : $P(x, y) = P(y, x)$. Pour le bon fonctionnement de l'algorithme de simulation, il faut que la chaîne de Markov correspondante soit facile à simuler, c'est-à-dire que la loi $P(x, \cdot)$ doit être facile à calculer.

— Tirer X_0 suivant une loi quelconque μ (typiquement $\mu = \delta_x$ pour une certaine configuration $x \in E$)

— Pour $x \in E$, on appelle μ_x la distribution de probabilité correspondant à la ligne x de la matrice de transition :

$$\mu_x(y) = P(x, y).$$

Pour chaque n, tirer Y_{n+1} suivant la loi μ_{X_n}.

— Tirer U_n une variable de loi uniforme sur $[0, 1]$ indépendamment de (X_n) et (Y_n).

— Si $\pi(Y_{n+1})/\pi(X_n) \geqslant U_n$, poser $X_{n+1} = Y_{n+1}$

— Sinon, garder $X_{n+1} = X_n$.

En d'autre termes, on fait évoluer $X = (X_n)$ comme une chaîne de Markov normale de matrice de transition P, à la différence qu'à chaque itération on ne garde la nouvelle valeur X_{n+1} que si le nouveau ratio $\pi(X_{n+1})/\pi(X_n)$ est suffisamment élevé, autrement on laisse l'ancienne valeur $X_{n+1} = X_n$.

Exercice 72. Pourquoi (X_n) est une chaîne de Markov (homogène) ? Quelle est sa matrice de transition ? Montrer qu'elle est irréductible et réversible. Qu'en déduisez-vous sur la limite de X_n ? Par quel type plus général de condition peut-on remplacer

$$U_n \leq \frac{\pi(Y_{n+1})}{\pi(X_n)}?$$

Barker a proposé la condition

$$U_n \leq \frac{\pi(Y_{n+1})}{\pi(X_n) + \pi(Y_{n+1})}$$

Correction:

La matrice de transition est, pour $x \neq y$

$$\begin{aligned} Q(x, y) &= \mathbb{P}(X_1 = y | X_0 = x) = \mathbb{P}(U_n \leq \pi(Y_1)/\pi(X_0), Y_1 = y | X_0 = x) \\ &= \mathbb{P}(Y_1 = y | X_0 = x)\mathbb{P}(U_n \leq \pi(Y_1)/\pi(X_0) | Y_1 = y, X_0 = x) = P(x, y) \min(1, \frac{\pi(y)}{\pi(x)}), \end{aligned}$$

et pour $x \in E$

$$Q(x, x) = 1 - \sum_{y \neq x} Q(x, y).$$

La chaîne de Markov est IRPA :

Irréductible. P est irréductible, donc pour x, y il existe un chemin probable $x_0 = x \to x_1 \to \cdots \to x_q = y$ pour la matrice de transition P. Pour tout x, y tels que $P(x, y) > 0$, la matrice de transition Q vérifie aussi $Q(x, y) > 0$ car $\pi > 0$. Donc le chemin est aussi probable pour Q, et donc la chaîne de Markov est irréductible.

réversible Pour $x = y, \pi(x)Q(x, x) = \pi(x)Q(x, x)$, pour $x \neq y$,

$$\pi(x)Q(x, y) = P(x, y)\min(\pi(x), \pi(y)) = P(y, x)\min(\pi(y), \pi(x)) = \pi(y)Q(y, x).$$

Donc π est une distribution invariante. Comme Q est aussi apériodique (exo), on a la convergence en moyenne et en loi de X_n vers π. Toute fonction $\varphi(x, y) : [0, 1]^2 \to [0, 1]$ telle que $x\varphi(x, y) = y\varphi(y, x)$ conviendrait également dans $U_n \leq \varphi(\pi(Y_{n+1}), \pi(X_n))$, comme par exemple $\varphi(a, b) = \min(1, b/a)$ ou $\varphi(a, b) = \frac{a}{a+b}$ (Condition de Barker).

Exercice 73. Utiliser l'algorithme de Metropolis pour simuler approximativement une variable de Poisson de paramètre $\theta > 0$. Quelle est la matrice de transition correspondante si l'on utilise la règle de Barker ? Cette chaîne peut-elle vérifier la condition de Doeblin ?

Correction: La loi limite est

$$\pi(x) = e^{-\theta}\theta^x/x!, x \in \mathbb{N},$$

sur l'espace d'états $E = \mathbb{N}$. On choisit simplement comme loi de départ

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

et $P = 0$ ailleurs. :

Algorithme :

1. On part de $X_0 = x$ (=0 par exemple, ou partie entière de θ pour être plus proche de la loi d'arrivée).
2. Si X_n est défini, on pose $x = X_{n+1}$. On détermine X_{n+1} de la manière suivante :
   (a) On tire Y_{n+1} indépendamment du reste suivant la loi μ_{X_n} :

   $$\mathbb{P}(Y_{n+1} = y) = P(x, y) = \mu_x(y).$$

   (c'est bien une distribution car P est une matrice stochastique). Dans notre cas ça donne

   $$\begin{aligned} \mathbb{P}(Y_{n+1} = x + 1) =& \frac{1}{2} \\ \mathbb{P}(Y_{n+1} = x - 1) =& \frac{1}{2} \text{ si } x \neq 0 \\ \mathbb{P}(Y_{n+1} = 0) =& \frac{1}{2} \text{ si } x = 0. \end{aligned}$$

   (b) On tire indépendamment du reste $U_n \sim \mathcal{U}_{[0,1]}$

   (c) Le ratio de l'algorithme de Métropolis avec condition de Barker est, avec $y = Y_{n+1}$

   $$r := \frac{\pi(y)}{\pi(x) + \pi(y)} = \frac{\theta^y/y!}{\theta^x/x! + \theta^y/y!} = \begin{cases} \frac{\theta/y}{1+\theta/y} \text{ si } y = x + 1 \\ \frac{1}{1+\theta/x} \text{ si } x = y + 1 \\ 1 \text{ si } x = y = 0. \end{cases}$$

   (calculer ces quantités n'implique pas de calculer de grandes quantités comme $x!$ ou $y!$, c'est l'avantage de cetteméthode) Si $r \geqslant U_n$, on pose $X_{n+1} = Y_{n+1}$. Sinon on pose $X_{n+1} = X_n$.
3. Retour au point 2.

6.2 Exercices

Exercice 74. Montrer que le modèle de l'urne d'Ehrenfest ne vérifie pas la condition de Doeblin (ou plus généralement le montrer pour toute chaîne de Markov périodique).

Correction: On cherche l tel que $Q^l(x,y) \geqslant \beta m(y)$ pour tout x,y. Comme m a au moins une valeur non-nulle, il existe $y_0 \in E$ tel que $m(y_0) > 0$. Donc, pour tout $x \in E$,

$$Q^l(x, y_0) > 0.$$

On sait que $Q^l(y_0, y_0) = 0$ si l est impair, et $Q^l(x, y_0) = 0$ si x est un voisin immédiat de y_0 et l est pair. Donc il n'est pas possible de trouver l qui convienne. Il y aura le même problème pour toute chaîne périodique.

Exercice 75. Soit Q une matrice stochastique de taille N, et λ une valeur propre complexe de Q de module 1. Montrer que λ est une racine de l'unité. (Indication : Commencer par le cas ou la chaîne de Markov correspondante est irréductible et apériodique).

Correction: L'espace d'état est fini.

Supposons la chaîne irréductible apériodique. Alors elle est récurrente positive. On a donc

$$\mathbb{P}_i(X_n = j) \to \pi(j).$$

Soit $v = (v(1), \ldots, v(N))$ un vecteur propre associé à λ. On a $Q^n v = \lambda^n v$ d'une part et d'autre part

$$(Q^n v)_j = \sum_i v(i)\mathbb{P}_i(X_n = j) \to \sum_i v(i)\pi(j).$$

En particulier $\lambda^n v$ converge vers une valeur non nulle, et donc λ^n converge vers une valeur non nulle, donc $|\lambda| = 1$. Si λ n'était pas égal à 1 (comme $|\lambda| = 1$), λ^n tournerait indéfiniment autour du disque unité et λ^n ne pourrait pas converger. On en déduit $\lambda = 1$ (en particulier c'est une racine de l'unité...).

Si désormais on ne suppose plus la chaîne irréductible, on considère la restriction de la chaîne à une classe récurrente (il en existe au moins 1 car l'espace d'états est fini), et sa matrice de transition Q'. Comme $Qv = \lambda v$, en appelant v' la restriction de v aux indices dans la classe récurrente, on a $Q'v' = \lambda v'$ (dessin).

De plus, pour x dans cette classe récurrente, on a $Q^n(x,x) = (Q')^n(x,x) > 0$ pour n suffisamment grand, car la classe ne communique pas avec les autres classes. Donc la chaîne est apériodique. D'après ce qu'on a montré avant, on a donc $\lambda = 1$.

On ne suppose plus la chaîne apériodique, donc supposons qu'elle a une période p. Donc Q^p est apériodique. λ^p est une valeur propre de Q^p, donc d'après ce qu'on vient de montrer $\lambda^p = 1$, donc λ est une racine de l'unité.

Exercice 76 (Simulation d'un processus de répulsion)**.** On lance n particules chargées positivement dans $[0,1]^2$, que l'on approxime par

$$A_N = \{(k/N, j/N); 0 \leq k, j \leq N\} = \left(\frac{1}{N}\mathbb{Z}\right)^2 \cap [0,1]^2.$$

Une étude des propriétés électromagnétiques du système permet de montrer que la probabilité $\pi(x)$ d'une configuration $x = (x_1, \ldots, x_n) \in E = A_N^n$ est proportionnelle à

$$\mu(x) := \prod_{1 \leq i < j \leq n} \|x_i - x_j\|.$$

En d'autres termes,

$$\pi(x) = \frac{1}{Z}\mu(x)$$

où $Z = \mu(E)$ est une constante très difficile à déterminer.

Pour simplifier, on suppose que la grille est un **tore**, c'est-à-dire que l'on peut passer d'un côté au côté opposé.

Remarquons que chaque point $x_0 \in A_N$ a ainsi 4 voisins. On dit que $x, y \in E$ sont des **configurations voisines** si il existe $i_0 \in \{1, .., n\}$, noté $i_0(x, y)$, tel que x_{i_0} et y_{i_0} sont voisins dans la grille (torique), et $x_i = y_i$ pour $i \neq i_0$.

1. On définit la matrice de transition suivante : pour $x, y \in E$;

$$P(x,y) = \begin{cases} \frac{1}{4n} \text{ si } x \text{ et } y \text{ sont des configurations voisines,} \\ 0 \text{ sinon} \end{cases}$$

   Montrer que P est une matrice de transition . Soit $X = (X_k)_{k \geqslant 0}$ une chaîne de Markov ayant P comme matrice de transition . Décrire comment programmer le passage d'une configuration X_k à la configuration X_{k+1}. Montrer que X est irréductible et symétrique.
2. Montrer qu'elle n'est pas apériodique. On modifie la chaîne de Markov en introduisant $\varepsilon \in]0, 1[$, et en décrétant qu'à chaque transition, X_k a une probabilité ε de rester sur place. On appelle P' la nouvelle matrice de transition . Que vaut P' ? Pourquoi est-elle apériodique ?
3. Soit x, y deux configurations voisines. Calculer $r(x, y) := \mu(x)/\mu(y)$.
4. Proposer une procédure informatique pour simuler informatiquement approximativement ce processus, en se basant sur l'algorithme de Metropolis.

Correction:

1. Pour faire la transition entre un état $X_k = x = (x_1, \dots, x_n)$ vers le prochain état on procède d ela manière suivante. On tire $I \in \{1, \dots, n\}$ uniformément (index de la particule de x qui va se déplacer). Puis on tire Y uniformément parmi les 4 voisins de x_I. La nouvelle configuration est $X_{k+1} = y = (y_1, \dots, y_n)$ où

$$\begin{cases} y_I = & Y \\ y_i = & x_i \text{ pour } i \neq I. \end{cases}$$

   On vérifie que pour deux configurations $x, y \in E$, $\mathbb{P}(X_1 = y | X_0 = x) = 0$ si x, y ne sont pas des configurations voisines. Si elles le sont, on note $i_0 = i_0(x, y)$. On a alors x_{i_0} et y_{i_0} voisins dans la grille, et si l'on applique la procédure précédente avec les variables I et Y,

$$\mathbb{P}(X_1 = y | X_0 = x) = \mathbb{P}(I = i_0, Y = y_{i_0}) = \mathbb{P}(I = i_0)\mathbb{P}(Y = y_{i_0}) = \frac{1}{n}\frac{1}{4}.$$

   Dans tous les cas $\mathbb{P}(X_1 = y | X_0 = x) = P(x, y)$.
2. Montrons tout d'abord que $P^k(x, x) = 0$ pour tout nombre impair k et configuration $x = (x_1, \dots, x_n) \in E$. Soit $k \geqslant 1$. On suppose que $X_0 = x$. On note m_i le nombre de fois que la i-ème particule est tirée au hasard parmi les k transitions entre X_0 et X_k. Comme il y a eu en tout k transitions, $\sum_{i=1}^k m_i = k$. Entre $X_0 = x$ et X_k, la i-ème particule s'est donc déplacée m_i fois. Pour que $X_k = x$, il faut que chaque particule soit revenue à son état initial, et ça implique que m_i soit un nombre pair. En particulier, $k = \sum_{i=1}^n m_i$ est un nombre pair. Donc finalement, pour k impair,

$$Q^k(x, x) = \mathbb{P}_x(X_k = x) = 0.$$

Donc $A = \{k : Q^k(x,x) > 0\}$ ne contient que des nombres pairs. Cela implique $pgcd(A) \geqslant 2$. Comme $Q^2(x,x) > 0$, $2 \in A$, et donc $pgcd(A) = 2$. Donc la chaîne de Markov a pour période 2. On pose désormais

$$P'(x,y) = \begin{cases} \varepsilon \text{ si } x = y \\ (1-\varepsilon)P(x,y) \text{ si } x \neq y \end{cases}$$

Comme $P'(x,x) > 0$ A contient 1 et $pgcd(A) = 1$, elle est apériodique.

3. Soit $i_0 = i_0(x,y)$. En particulier $x_i = y_i$ pour $i \neq i_0$. On a

$$\frac{\mu(x)}{\mu(y)} = \frac{\prod_{i,j=1}^n \|x_i - x_j\|}{\prod_{i,j=1}^n \|y_i - y_j\|} = \frac{\prod_{i,j=1,i\neq i_0,j\neq j_0}^n \|x_i - x_j\|}{\prod_{i,j=1,i\neq i_0,j\neq i_0}^n \|x_i - x_j\|} \frac{\prod_{i=1,i\neq i_0}^n \|x_i - x_{i_0}\|}{\prod_{i=1,i\neq i_0}^n \|y_i - y_{i_0}\|} = \frac{\prod_{i=1,i\neq i_0}^n \|x_i - x_{i_0}\|}{\prod_{i=1,i\neq i_0}^n \|x_i - y_{i_0}\|}.$$

4. Algo de Metropolis avec la matrice de transition initiale $P'(x,y)$ pour un certain $\varepsilon > 0$. On calcule la ratio $\frac{\pi(Y_{n+1})}{\pi(X_n)} = r(Y_{n+1}, X_n)$ grâce à la question précédente.

6.3 Théorème ergodique

Le théorème 12 nous dit que si X est une chaîne IRPA la loi de X_n converge vers la distribution invariante π,

$$X_n \xrightarrow{\mathcal{L}} \pi.$$

En revanche, pour chaque trajectoire, X_n ne converge pas car il ne cesse de sauter d'un état à l'autre. On peut par contre dire que p.s. X_n converge en moyenne en un certain sens (rappelons que X_n n'est pas forcément un nombre...)

Le théorème de convergence en moyenne p.s. par excellence est la loi des grands nombres :

Théorème 13 (LGN)**.** *Soit $(V_n)_n$ des variables IID positives avec $\mathbb{E}|V_1| < \infty$. Alors p.s.*

$$\frac{1}{n}\sum_{k=1}^{n} V_k \to \mathbb{E}(V_1).$$

Si l'on n'a pas l'hypothèse $\mathbb{E}|V_1| < \infty$, le résultat est encore vrai si les V_i sont positifs ou nuls.

Théorème 14 (Théorème ergodique)**.** *Soit X une chaîne de Markov irréductible de distribution initiale une probabilité μ_0. Pour $n \geq 1$, on note $V_x(n)$ le nombre de visites en x avant le temps n*

$$V_x(n) = \sum_{k=1}^{n} \mathbf{1}_{X_k = x}.$$

Alors pour tout état $x \in E$ p.s.

$$\frac{1}{n}V_x(n) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k = x} \to \frac{1}{\mathbb{E}_x(T_x^{(1)})}$$

avec la convention $1/+\infty = 0$. Cette quantité est non-nulle uniquement si la chaîne de Markov est récurrente positive, auquel cas elle vaut $\pi(x)$ où π est l'unique distribution invariante.

Pour toute fonction $f : E \mapsto \mathbb{R}$ telle que

$$\sum_{x \in E} |f(x)| < \infty,$$

alors

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \to \overline{f} : \sum_{x \in E} f(x)\frac{1}{\mathbb{E}_x(T_x^{(1)})}.$$

En particulier, ce résultat est valide si f est bornée ou à support fini.

Remarque 10. Comme on a pas fait d'hypothèse de récurrence, cette dernière quantité peut être nulle.

Selon la légende, une des raisons pour lesquelles Markov a introduit son modèle est qu'il a eu un débat avec un autre mathématicien qui prétendait qu'il était impossible d'avoir une Loi des Grands Nombres avec des variables qui ne sont pas indépendantes.

Cela donne une manière d'estimer l'espérance. Ce type de méthode rentre dans les “méthodes de Monte-Carlo”.

Démonstration: Si la chaîne est transiente, le nombre de visites est p.s. fini et

$$\frac{V_x(n)}{n} \to 0 = \frac{1}{\mathbb{E}(T_x^{(1)})}$$

p.s..

Supposons la chaîne récurrente. Supposons $X_0 = x$, et posons $T_x^{(0)} = 0$. On rappelle que $S_x^r = T_x^{(r+1)} - T_x^{(r)}$ est la r-ème excursion entre 2 passages en x, pour $r \geq 0$. On a vu que les S_x^r sont des variables iid d'espérance

$$\mathbb{E}(S_x^1) = \mathbb{E}(T_x^{(1)}).$$

D'après la LGN classique,

$$\frac{U_t = S_x^{(1)} + S_x^{(2)} + \cdots + S_x^t}{t} \to \mathbb{E}_x(T_x^{(1)}) p.s. \tag{3}$$

quand $t \to \infty$. Comme il y a eu $V_x(n)$ visites au temps n, on a

$$U_{V_x(n)} \leq n \leq U_{V_x(n)+1}.$$

On a donc pour n suffisamment grand

$$\frac{U_{V_x(n)}}{V_x(n)} \leq \frac{n}{V_x(n)} \leq \frac{U_{V_x(n)+1}}{V_x(n)} = \frac{U_{V_x(n)+1}}{V_x(n)+1} \frac{V_x(n)+1}{V_x(n)},$$

en utilisant 3 et le fait que $V_x(n) \to \infty$ p.s. on a

$$\frac{n}{V_x(n)} \to \mathbb{E}_x(T_x^{(1)}) p.s.$$

On a pour la fonction f,

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) = \frac{1}{n} \sum_{x \in E} f(x) \sum_{k=1}^{n} \mathbf{1}_{\{X_k = x\}} = \frac{1}{n} \sum_{x \in E} f(x) V_x(n).$$

On pose $\pi(x) = \frac{1}{\mathbb{E}_x(T_x^{(1)})}$. Comme E est dénombrable

$$\mathbb{P}\left(\forall x \in E, \frac{V_x(n)}{n} \to \pi(x)\right) = 1,$$

On veut passer la limite en n a l'extérieur de la somme. Ecrivons ceci sous forme d'intégrale pour appliquer le théorème de Lebesgue : on veut montrer, avec μ la mesure de comptage ($\mu(x) = 1$) pour tout $x \in E$),

$$\int_E f_n(x) \mu(dx) \to \int_E f(x) \pi(x) \mu(dx)$$

où pour tout $x \in E$, $f_n(x) = f(x) \frac{V_x(n)}{n} \to f(x)\pi(x)$. On utilise la domination $|f_n(x)| \leqslant g(x) := |f(x)|$ qui vérifie bien

$$\int_E g(x) dx = \sum_{x \in E} |f(x)| dx < \infty.$$

Donc d'après le théorème de Lebesgue,

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

Exemple 14. Utilisons ce résultats pour classer les pages du web selon les critéres vus prévédemment. La convergence du théorème est une convergence en moyenne, et non pas une convergence p.s. Il ne suffit donc pas de lancer un seul robot-surfeur pour estimer $\pi(x)$, il faut en lancer N indépendants, qui constituent autant de chaînes de Markov $X^{(1)}, \ldots, X^{(N)}$, avec N grand, où chacune débute avec une distribution μ_0 sur E.

Pour estimer $\pi(x)$ pour une certaine page x, on estime $\mathbb{P}_{\mu_0}(X_n^{(1)} = x)$ avec la LGN :

$$\mathbb{P}_{\mu_0}(X_n^{(1)} = x) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_{\{X_n^{(k)} = k\}}$$

donc

$$\pi(x) = \lim_{n \to \infty} \mathbb{P}_{\mu_0}(X_n^{(1)} = x).$$

On a même une estimation de la vitesse de convergence :

$$\|\pi(x) - \mathbb{P}_{\mu_0}(X_n^{(1)} = x)\| \leqslant \alpha^n$$

pour un certain $\alpha \in]0,1[$.

6.4 Exercices

Exercice 77. On reprend l'exercice de renouvellement des machines, avec l'hypothèse que chaque jour, la machine a une probabilité $p \in]0,1[$ constante de casser avant d'atteindre le jour suivant (modèle sans mémoire). On suppose désormais que la casse d'une machine a un cout très élevé, noté $a > 0$. L'utilisateur décide alors d'adopter la politique suivante : Il fixe un âge limite L, et si la machine arrive à l'âge L, alors il la démonte et la remplace le lendemain par une nouvelle machine neuve, ce qui lui occasionne un cout $b > 0$. On suppose évidemment que $b < a$, sinon cette politique est inutile.

1. Décrire la chaîne de Markov correspondante et donner son graphe. Etudier les mesures et probabilités invariantes.
2. On suppose qu'au temps 0 l'utilisateur doit dépenser b pour acheter sa première machine. Montrer que le coût au temps n s'écrit
$$C_n = (b-a)V_L(n) + aV_0(n) - a,$$
ou $V_x(n)$ est le nombre de visites en x avant le temps n.
3. Calculer le cout limite moyen
$$m = \lim_n \frac{C_n}{n}.$$

Correction: 1) Avec les notations de l'exo 57, en appelant p_k les probabilités de vieillissement, les probabilités de vieillissement avec la politique de remplacement sont

$$\tilde{p}_k = \begin{cases} p_k \text{ si } k \leqslant L, \\ 0 \text{ si } k > L. \end{cases}$$

On en déduit que le produit infini converge et que la mesure invariante telle que $\mu(0) = 1$ est donnée par

$$\mu(k) = \begin{cases} v_k = \prod_{0 \leqslant i \leqslant k-1} p_k = p^k, 1 \leq k \leq L, \\ 0 \text{ si} k > L \end{cases}$$

elle a pour somme

$$s = \mu(\{0,\dots,L\}) = \sum_{k=0}^{L} p_0 \dots p_{k-1} = \sum_{k=0}^{L-1} p^k = \frac{1-p}{1-p^L}$$

La proba invariante est donc

$$\pi(k) = p^k \frac{1-p^L}{1-p}$$

.

2) Le cout correspond en fait à

$$C_n = a\#\{casses\} + b\#\{achats\}.$$

Le nombre d'achats est $V_0(n)$. Dans les cycles passés, il y a eu casse à chaque fois qu'il y a eu un passage en 0 sans être passé par L (sans compter le démarrage au temps 0) :

$$C_n = bV_L(n) + a(V_0(n) - 1 - V_L(n)).$$

3) Comme la chaîne est irréductible récurrente positive, le théorème ergodique nous donne

$$n^{-1}V_L(n) \to \pi(L)p.s.$$
$$n^{-1}V_0(n) \to \pi(0)p.s.,$$

donc

$$\frac{C_n}{n} \to \pi(L)b + \pi(0)(b-a),$$

4) Dans le cas d'un vieillissement constant, on a

$$\pi(0) = \frac{\mu(0)}{s} = \frac{1-p}{1-p^{L+1}}$$
$$\pi(L) = \frac{\mu(L)}{s} = \frac{p^L(1-p)}{1-p^{L+1}},$$

et donc le cout moyen est

$$\gamma(L) = \frac{C_n}{n} = \frac{1-p}{1-p^{L+1}}\left(bp^L + (b-a)\right)$$

Exercice 78. On jette un dé à 6 faces n fois, et on appelle X_n la somme de tous les résultats obtenus. On veut calculer

$$\lim_n \mathbb{P}(X_n \text{ est un multiple de } 13).$$

1. Montrer que X_n est une chaîne de Markov. Pourquoi ne peut-on pas lui appliquer les théorèmes du cours ?

   Correction:

   X_n n'est pas récurrente (elle passe au plus en chaque état de $\mathbb{Z}$).

2. On considère $\overline{X_n}^{13} \in \{0,1,2,\dots,12\}$ le reste de la division euclidienne de X_n par 13. Montrer que c'est une chaîne de Markov irréductible récurrente positive apériodique. En déduire le résultat.

   Correction:

   Soit deux états $x, y \in \{0,1,2,\dots,12\}$. Supposons par exemple $x < y$. Si on fait $y-x$ 1 de suite (ce qui a une probabilité non-nulle), alors on passe de x à y, donc $Q^n(x,y) > 0$ pour $n = y-x$. Si $x > y$, on peut suivre le chemin $x \to x+1 \to \dots 13 = 0 \to 1 \to \dots \to y$ avec une probabilité non-nulle. Donc la chaîne est irréductible. Comme l'espace d'états est fini, elle est aussi récurrente positive.

   Comme chaque point x à 6 antécédents $x-1,\dots,x-6$ et 6 destinations équiprobables $(x+1,\dots,x+6)$, on montre facilement que la distribution uniforme $\pi(x) = \frac{1}{13}$ est la distribution stationnaire.

   Pour montrer qu'elle est apériodique, on considère $x \in E$. Comme on peut faire six 1 de suite avec probabilité > 0, $Q^6(x,x) > 0$. Si on fait un 2 puis quatre 1, on retombe sur x, et on a donc $Q^5(x,x) > 0$. On a donc $Q^{n_0}(x,x) > 0$ et $Q^{n_0+1}(x,x) > 0$ avec $n_0 = 5$, et la chaîne est donc apériodique. En utilisant le théorème de convergence à l'équilibre, on en déduit

   $$\lim_n \mathbb{P}(X_n \text{ est un multiple de } 13) = \lim_n \mathbb{P}(\overline{X_n}^{13} = 0) \to \pi(0) = \frac{1}{13}.$$

3. On pose $Y_n = 5X_n$. Calculer

   $$\lim_n \mathbb{P}(Y_n \text{ est un multiple de } 13).$$

   Correction: La situation est un peu plus compliquée car on fait des sauts de 5 en 5 (modulo 13), on a le graphe suivant

   $0 \to 5 \to 10 \to 15 = 2 \to 7 \to 12 \to 17 = 4 \to 9 \to 14 = 1 \to 6 \to 11 \to 16 = 3 \to 8 \to 0 \to$ *etc...*

On peut encore aller de n'importe quel point à n'importe quel autre point en faisant suffisamment de 1 de suite, la chaîne est donc irréductible, et donc récurrente positive. La distribution stationnaire est encore $\pi(x) = 1/13$.

L'apériodicité se traite comme dans le cas précédent : En considérant deux chemins allant d'un point x à lui-meme, l'un ne comportant que des 1, et l'autre comportant un 2 et que des 1, on montre que la chaîne est apériodique. On a donc la même conclusion qu'à la question précédente.

4. Calculer
$$\lim_n \mathbb{P}(X_1 + Y_n \text{ est un multiple de } 10).$$

   Correction:

   Si on s'intéresse désormais à la chaîne $\overline{X_n}^{10}$, la situation n'est pas la même. On ne peut effectuer que des transitions du type
$$x \to x+5 \to x+10 = x \to x+5 \to x+10...$$
   avec x l'état initial entre 1 et 6. La chaîne de Markov est donc 2-périodique et on ne peut pas appliquer le théorème de convergence en loi.

   Le plus simple dans ce cas est de décomposer selon la vaeur de X_1 :
$$\mathbb{P}(X_1 + Y_n \equiv 0[10]) = \sum_{i=1}^{6} \mathbb{P}(X_1 = 6) = \lim_{n\to\infty} \sum_{i=1}^{6} \frac{1}{6} \mathbb{P}_i(x + Y_n \equiv 0[10]).$$
   On voit facilement que pour $i \neq 5$, $\mathbb{P}(x + Y_n \equiv 0[10]) = 0$. Donc la limite recherchée est $\frac{1}{6} \lim_{n\to\infty} \mathbb{P}_5(x + Y_n \equiv 0[10])$

   Donc $\overline{X_n} \in \{0, 5\}$ tout le temps. Si on est en 0 et qu'on fait $2, 4$ ou 6, on reste en 0, autrement on reste en 5. La chaîne a donc la matrice de transition
$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$
   et la distribution invariante $(1/2, 1/2)$. Cette chaîne de Markov est IRPA et donc la limite recherchée est $\frac{1}{6} \times \frac{1}{2} = \frac{1}{12}$.

5. On considère cette fois P_n le produit de tous les résultats obtenus après n lancers, avec la convention $P_0 = 1$. On considère la chaîne de Markov $\overline{P_n}^{13}$

   (a) Montrer que p.s. pour tout n, P_n n'est pas un multiple de 13.

   Correction:

   Montrons par récurrence que p.s. pour tout n, P_n n'est pas un multiple de 13. C'est vrai pour P_1 car P_1 est compris entre 1 et 6. Supposons que c'est vrai pour un n quelconque. Comme $P_{n+1} = P_n X_n$, si $P_{n+1} = 13k$ pour un certain k, alors par unicité de la décomposition en facteurs premiers $X_n = 13$, ce qui est impossible. Donc
$$\mathbb{P}(P_n \equiv 0[13]) = 0 \to 0.$$

   (b) Montrer que la chaîne de Markov $\overline{P_n}^{13}$ est irréductible.

   Correction: Soit $E = \mathbb{Z}/13\mathbb{Z}$.

   On peut observer que si l'on fait 2 à chaque coup on a
$$1 \to 2 \to 4 \to 8 \to 3 \to 6 \to 12 \to 11 \to 9 \to 5 \to 10 \to 7 \to 1.$$
   Donc tous les états communiquent et la chaîne de Markov est irréductible.

(c) Montrer que la mesure constante $\mu(x) = 1$ est invariante. *Correction:*
Tout élément x de E admet un élément y de E, noté x^{-1}, tel que $xy \equiv 1[13]$. On cherche μ tel que

$$\mu(x) = \sum_{z=1}^{12} \mu(z)\mathbb{P}(X_1 = x | X_0 = z)$$

Les états menant à x sont $x, 2^{-1}x, 3^{-1}x, 4^{-1}x, 5^{-1}x, 6^{-1}x$. Ces états sont tous différents car $zx \equiv z'x$ implique (après multiplication par x^{-1}) que $z = z'$. On a donc

$$\mu(x) = \sum_{u=1}^{6} \mu(xu^{-1})\frac{1}{6}.$$

On voit que $\mu(x) = 1$ (ou n'impore quelle autre constante) satisfait cette équation pour tout x.

(d) Calculer $\lim_{n\to\infty} \mathbb{P}(P_n \equiv 1[13])$.
Correction: La chaîne de Markov est irréductible sur un espace d'états fini, elle est donc IRP.
La chaîne de Markov admet $\pi = \mu/13$ comme distribution invariante, c'est donc la seule.
Il reste à montrer que la chaîne de Markov est apériodique, mais pour ça il suffit de trouver 2 chemins de 1 vers 1 qui ont des longueurs dont le pgcd est 1, ou qui ont des longueurs l et $l+1$ (pour un certain l - le pgcd sera forcément 1). Par exemple

$$1 \to 2 \to 4 \to 8 \to 3 \to 6 \to 12 \to 11 \to 9 \to 5 \to 10 \to 7 \to 1$$

et

$$1 \to 4 \to 8 \to 3 \to 6 \to 12 \to 11 \to 9 \to 5 \to 10 \to 7 \to 1.$$

On a donc la convergence en loi $\mathbb{P}_x(\overline{P_n}^{13} \equiv y) \to \pi(y) = 1/13$ pour n'importe quels x, y

(e) On étudie $Y_n = \overline{P_n}^4$ le reste de la division euclidienne de P_n par 4. Calculer

$$\lim_n \mathbb{P}(Y_n \text{ est un multiple de } 4).$$

Correction: Cette fois on considère la valeur de Y_n modulo 4. On a

$$\begin{aligned} 0 &\to 0 \\ 1 &\to 1, 2, 3, 0 \\ 2 &\to 2, 0 \\ 3 &\to 3, 2, 0. \end{aligned}$$

Donc $\{0\}$ est une classe récurrente, et comme tout autre état mène à 0 (en exo), les autres états sont transients. On en déduit que la chaîne passera par 0 et qu'elle y restera constamment, donc

$$\lim_n \mathbb{P}_n(X_n = 0[4]) = 1.$$

La différence avec 13 est la suivante : Si $X_n = 13$, cela signifie qu'on a pu multiplier les lancés successifs $L_1, L_2, \dots$ pour obtenir 13 :

$$13 = D_1 D_2 ... D_n.$$

Ceci dit, comme la seule décomposition de 13 en nombres premiers est

$$13 = 1 \times 13,$$

cela veut dire qu'un des lancés est égal à 13, ce qui est impossible.

(f) Que se passe-t-il si on remplace 4 par un autre nombre ?

Correction:

On peut en déduire de la même manière que tout nombre k pour lequel un nombre premier différent de $\{1,2,3,4,5,6\}$ apparait dans sa composition vérifiera

$$\mathbb{P}(X_n = 0[k]) \to 0.$$

Si par contre k peut s'écrire

$$k = p_1 p_2 ... p_q$$

avec $p_q \in \{0,1,2,\ldots,6\}$, ce n'est qu'une question de temps avant que la chaîne n'effectue la séquence

$$p_1, p_2, ..., p_q$$

ce qui implique que k sera inéluctablement un facteur de X_n pour un n suffisamment large, et donc

$$\lim \mathbb{P}(X_n = 0[k]) \to 1.$$

6. Dans toutes les situations précédentes, calculer le nombre moyen de visites à l'état désiré, c'est-à-dire

$$\lim_n \frac{\#\{n : X_n \text{ ou } Y_n \text{ est un multiple de } 13/10/4...\}}{n}.$$

Correction: Le théorème ergodique s'applique à toute chaîne irréductible. Pour les questions $1 \to$ 5.b), on a donc

$$... \to \pi(x)$$

qui vaut selon les questions $\frac{1}{13}, \frac{1}{10}, \frac{1}{2}, \frac{1}{12}$ ou 1.

Rappel sur l'anneau $\mathbb{F}_p := (\mathbb{Z}/p\mathbb{Z}, +, \times)$. Soit $p \geqslant 2$. On note $\mathbb{Z}/p\mathbb{Z} = \{0,1,2,\ldots,p-1\}$. L'addition et la multiplication se font modulo p. On dit que $x, y \in \mathbb{F}_p$ sont inversibles si $xy = 1$ dans $\mathbb{F}_p$ (i.e. $xy \equiv 1[p]$). Si p est premier, $(\mathbb{F}_p, +, \times)$ est un corps, et donc tous ses éléments sont inversibles sauf 0. De plus d'après le (petit) théorème de Fermat, pour $x \in \overline{F}_p \setminus \{0\}$, $x^{p-1} \equiv 1[p]$.

Exercice 79. Algorithme de mélange L'exercice a pour but d'évaluer l'efficacité d'algorithmes pour mélanger un paquet de 52 cartes. On considère les cartes d'un jeu, numérotées de 1 à 52. On assimile à une configuration du paquet un élément du groupe des permutations Σ_{52}. Un élément σ de Σ_{52} est une bijection de $\{1,2,\ldots,52\}$ vers lui-même. La configuration correspondante du paquet se retrouve en mettant les cartes dans l'ordre $\sigma(1)$ (au-dessus du paquet) $,\sigma(2),\ldots,\sigma(52)$. On admet le résultat suivant sur les éléments de Σ_{52} :Toute bijection peut s'écrire comme le produit de transpositions.

L'espace d'états est l'ensemble des configurations possibles du paquet de cartes, c'est-à-dire $E = \Sigma_{52}$. On appelle μ_0 la distribution uniforme sur E_{52}.

1. Préliminaire

(a) Soit σ une configuration. Que vaut $\mu_0(\sigma)$?

Correction: Comme $\#\Sigma_{52} = 52!$, $\mu_0(\sigma) = 1/52!$

(b) Soit i, j deux éléments distincts de $\{1,\ldots,52\}$. On note $\tau_{i,j}$ la transposition qui échange les éléments i et j de $\{1,\ldots,52\}$, i.e.

$$\tau_{i,j}(k) = \begin{cases} j \text{ si } k = i \\ i \text{ si } k = j \\ k \text{ sinon.} \end{cases}$$

Montrer que $\sigma \circ \tau$ est la permutation qui correspond à la configuration du paquet de cartes après qu'on ait échangé les cartes i et j.

Correction: Dans la configuration σ, les cartes sont dans l'ordre $\sigma(1), \ldots, \sigma(52)$. On a $\sigma \circ \tau(i) = \sigma(j)$ et $\sigma \circ \tau(j) = \sigma(i)$. Donc dans la configuration $\sigma \circ \tau$, les cartes sont dans l'ordre $\sigma(1), \ldots, \sigma(i-1), \sigma(j), \sigma(i+1), \ldots, \sigma(j-1), \sigma(i), \sigma(j+1), \ldots, \sigma(52)$ (en supposant $i < j$). C'est bien le paquet qu'on obtient en échangeant les i-èmes et j-èmes cartes.

2. **Mélange par transpositions**. A chaque étape, on choisit deux cartes au hasard uniformément dans le paquet et on les échange.

   (a) Formaliser la chaîne de Markov correspondante, et donner la matrice de transition. Est-elle symmétrique ?

   Correction:

   Etant donné une configurations, on peut former $52 \cdot 51/2 = 1326$ couples $\{i, j\}$ avec $i \neq j$. On peut donc lui appliquer 1326 transpositions différentes. Soit $\sigma, \sigma' \in \Sigma_{52}$.

   $$Q(\sigma, \sigma') = \begin{cases} \frac{1}{1326} \text{ si } \sigma' \text{ s'obtient à partir de } \sigma \text{ par une transposition} \\ 0 \text{ sinon} \end{cases}$$

   En particulier, elle est symmétrique, car on peut appliquer la même transposition pour passer de σ' à σ.

   (b) A-t-on une chaîne de Markov irréductible ? Récurrente ? Positive ?

   Correction: Soit $\sigma, \sigma' \in \Sigma_{52}$. Il existe des transpositions $\tau_1, \ldots, \tau_q$ et $\tau'_1, \ldots, \tau'_p$ telles que $\sigma = \tau_1 \circ \cdots \circ \tau_q$ et $\sigma' = \tau'_1 \circ \cdots \circ \tau'_q$. On peut donc passer de la configuration σ à la configuration σ' en appliquant les transpositions $\tau_q, \ldots, \tau_1, \tau'_1, \ldots, \tau'_q$ à au paquet. Comme l'espace d'états est fini, la chaîne de Markov est également récurrente positive.

   (c) A-t-on une chaîne de Markov apériodique ? (on pourra utiliser la notion de signature d'une permutation)

   Correction: Si $s(\sigma)$ est la signature de σ, pour toute transposition τ, $s(\sigma \circ \tau) = -s(\sigma)$. Donc je ne pourrai pas revenir à σ en appliquant un nombre impair de transpositions $\tau_1, \ldots, \tau_p$: en effet,

   $$s(\sigma \circ \tau_1 \circ \cdots \circ \tau_p) = (-1)s(\sigma \circ \tau_1 \circ \cdots \circ \tau_{p-1}) = \cdots = (-1)^p s(\sigma)$$

   , et on ne peut donc pas avoir la signature de la permutation initiale σ. Donc il n'y a aucun nombre impair dans l'ensemble $A = \{n : Q^n(\sigma, \sigma) > 0\}$. En revanche, $2 \in A$ car pour toute permutation $\tau, \sigma \circ \tau \circ \tau = \sigma$. Donc $pgcd(A) = 2$, et la chaine est 2-périodique.

   (d) Soit σ_0 une configuration aléatoire tirée selon la distribution μ_0, et σ_1 la configuration obtenue après avoir appliqué le mélange par transposition une fois. Soit $\sigma \in \Sigma_{52}$. Que vaut $\mathbb{P}(\sigma_1 = \sigma)$? Qu'en déduit-on pour μ_0 ?

   Correction: On appelle τ_1 la transpositino aléatoire appliquée à σ_0. $\mathbb{P}(\sigma_1 = \sigma) = \mathbb{P}(\sigma_0 \circ \tau_1 = \sigma) = \mathbb{P}(\sigma_0 = \sigma \circ \tau_1) = 1/52!$ car σ_0 est tirée uniformément. Donc σ_1 a également la distribution uniforme, et cette distribution est invariante pour notre chaîne de Markov .

   (e) Soit σ une configuration aléatoire tirée selon la distribution μ_0. On considère la variable aléatoire

   $$Y = \#\{ \text{ nombre de coeurs dans la première moitié du paquet.} \}$$

   Donner une procédure pour estimer $\mathbb{E}Y$.

   Correction: Comme la chaîne de Markov est irréductible, on peut appliquer le théorème ergodique. On peut écrire Y de la forme

   $$Y = \sum_{k=1}^{26} \mathbf{1}_{\{\text{la } k\text{-ème carte de } \sigma \text{ est un coeur}\}} = f(\sigma).$$

On appelle $(\sigma_n)_{n\geqslant 0}$ la chaîne de Markov décrite précédemment. On appelle Y_n le nombre de coeurs dans la première moitié du paquet σ_n. D'après le théorème ergodique

$$\frac{1}{n}\sum_{m=1}^{n} Y_m = \frac{1}{n}\sum_{m=1}^{n} f(\sigma_n) \to \bar{f}(y) = \sum_{\sigma\in E} \mu_0(\sigma) f(\sigma) = \mathbb{E}Y,$$

ce qui nous donne une manière d'estimer Y sur cette trajectoire de la chaîne de Markov .

3. **Mélange par coupe**. Pour effectuer le mélange, on tire une variable K binomiale de paramètres $n = 52, p = 1/2$. On prend les K cartes du dessus, et on les place telles quelles en dessous des cartes restantes.

   (a) Décrire la permutation $\tau_K \in \Sigma_{52}$ telle que, si σ' s'obtient à partir de σ via la coupe décrite ci-dessus, $\sigma = \sigma' \circ \tau_K$. (on suppose que la carte du dessus est numérotée 1).

   Correction: Soit $k \in \{0, 52\}$. On définit τ_k par

$$\tau_k(i) = \begin{cases} i + 52 - k \text{ si } i \leqslant k \\ i - k \text{ si } i > k. \end{cases}$$

   Donc si σ' est la configuration obtenue après avoir appliqué ce mélange, on a $\sigma'(\tau_k(i))$ est la carte de σ qui a été amenée à la position $\tau_k(i)$, c'est $\sigma(i)$.

   (b) Décrire la matrice de transition de la chaîne de Markov correspondante. Est-elle symmétrique ?

   Correction:

   Soit deux permutations distinctes σ, σ'. Si il existe k tel que $\sigma' = \sigma \circ \tau_k$ alors ce k est unique. On le note $k_{\sigma,\sigma'}$. Remarquons que l'opération inverse consiste à couper au niveau $52 - k$. Donc $k_{\sigma',\sigma} = 52 - k_{\sigma,\sigma'}$. On a

$$Q(\sigma, \sigma') = \begin{cases} \mathbb{P}(K = k_{\sigma,\sigma'}) \text{ si il existe un tel } k \\ 0 \text{ sinon,} \end{cases}$$

   en rappelant

$$\mathbb{P}(K = k) = \binom{k}{52} 2^{-52}.$$

   Donc $\mathbb{P}(K = k) = \mathbb{P}(K = 52 - k)$ et la chaîne de Markov est symmétrique.

   (c) La chaîne de Markov est-elle apériodique ?

   Correction: Soit $0 \leqslant k, k' \leqslant 52$ tels que $k + k' \leqslant 52$ (on peut prendre $k = k' = 1$). Alors faire la coupe au niveau k puis la faire au niveau k' revient à faire la coupe au niveau $k + k'$. En terme de chaîne de Markov , cela veut dire qu'on peut faire le même trajet en un temps 1 ou en un temps 2, et cela implique que la chaîne de Markov est apériodique.

   (d) μ_0 est-elle une mesure invariante ?

   Correction: Oui μ_0 est invariante. Même raisonnement que le mélange par transposition avec $\mathbb{P}(\sigma_1 = \sigma) = \mathbb{P}(\sigma_0 = \sigma \circ \tau_k^{-1}) = 1/52!$.

   (e) La chaîne est-elle irréductible ? Donner la décomposition en classes de l'ensemble d'états.

   Correction: Non car 2 cartes proches ne pourront être séparés (sauf en étant aux deux extrémités du paquet). Ainsi on a plusieurs classes de cardinal 52. Le nombre de classes est donc $52!/52 = 51!$ Ces classes sont toutes récurrentes

Exercice 80. Simulation d'un modèle de polymère

Soit $p \geqslant 3$. On appelle p-polymère un ensemble de p points distincts de $\mathbb{Z}^2$ $x_1, \ldots, x_p$, tels que pour $1 \leqslant i < j \leqslant p$, les segments $[x_i, x_{i+1}]$ et $[x_j, x_{\overline{j+1}}]$ ne se croisent pas, en notant $1 = \overline{p+1}$, et tel que 0 soit à l'intérieur du polygone délimité par $x_1, \ldots, x_p$. On appelle E l'ensemble de tous les polymères possibles, en notant que E est strictement inclus dans $(\mathbb{Z}^2)^p$. Pour $x = (x_1, \ldots, x_p) \in E$, on note $\mathrm{Per}(x)$ le périmètre du polymère, et $S(x)$ la surface à l'intérieur du polymère.

On note m la mesure $m(x) = S(x)/\mathrm{Per}(x)^4$.

Soit $x, y \in E$ deux configurations de polymère. On dit que x et y sont des configurations voisines si il existe $i_0 \in \{1, \ldots, p\}$ tels que pour $i \neq i_0, x_i = y_i$, et x_{i_0} et y_{i_0} sont distincts mais voisins dans la grille définie par $\mathbb{Z}^2$. On note $V(x)$ l'ensemble des configurations voisines de x.

1. Trouver, pour $p = 3$, 2 configurations voisines x, y telles que $\#V(x) \neq \#V(y)$.
2. Soit P définie par $P(x, x) = 0, x \in E$ et pour $x \neq y$,

$$P(x,y) = \begin{cases} \frac{1}{\#V(x)} \text{ si } x \text{ est une configuration voisine de } y \\ 0 sinon \end{cases}$$

3. On définit une nouvelle matrice de transition P' par

$$P'(x,x) = \begin{cases} 0 \text{ si } \#V(x) = 4n \\ \frac{1}{4n}(\#V(x) - 4n) \text{ sinon.} \end{cases}$$

Montrer que P' est apériodique et symétrique.

4. On veut une distribution de probabilités π proportionelle à m. Pourquoi imposer que 0 soit à l'intérieur d'un polymère de $\mathbb{Z}^2$?
5. On suppose $p = 3$. Que doit-on vérifier pour l'existence de π ?
6. En admettant cette condition vérifiée, comment simuler une chaîne de Markov dont la loi converge vers π ?

Correction:

1.
2.
3.
4. Etant donné $x \in E$, sans la contrainte que 0 soit à l'intérieur, on aurait une infinité de polymères qui s'obtiennent à partir de x via une translation, et donc $m(E)$ serait infini.
5. Il faut que la masse totale soit finie :

$$m(E) = \sum_{\substack{x_1, x_2, x_3 \in \mathbb{Z}^2 \\ (x_1, x_2, x_3) \in E}} \frac{S(x)}{\mathrm{Per}(x)^4} < \infty.$$

6. Similaire à l'exercice précédent

Exercice 81. On appelle graphe un ensemble de points E, et un ensemble d'arêtes A qui à chaque paire de points x, y associe $a(x, y)$ qui vaut 0 ou 1. Si $a(x, y) = 1$, on dit que x et y sont connectés, ou voisins, et on note $x \sim y$. On appelle degré de x et on note $d(x)$ le nombre de voisins de x.

On considère la chaîne de Markov X_n qui se déplace aléatoirement en sautant d'un point à un autre, sachant que :

1. D'un point x, on ne peut aller que sur un voisin de x,
2. Tous les voisins de x ont la même probabilité d'être choisis.

1. **Introduction.** Donner une expression de $Q(x, y)$, la matrice de transition.

2. **Irréductibilité.**

On dit que deux points x et y sont reliés dans le graphe si il existe une suite de points $x_0 = x, x_1, \dots, x_k = y$ tels que $x_i \sim x_{i+1}$. On dit que le graphe est connexe si tous les points sont reliés.

Donner un exemple de graphe qui n'est pas connexe. Donner une condition nécessaire et suffisante pour que la chaîne soit irréductible.

<u>On suppose dans la suite que le graphe est connexe.</u>

3. **récurrence**.

a) On suppose que E est fini. Montrer que la chaîne est irréductible récurrente positive.

b) Donner des exemples de graphes ou la chaîne est récurrente mais pas récurrente positive, et ou la chaîne de Markov n'est même pas récurrente. (On pourra utiliser les résultats sur la marche aléatoire dans $\mathbb{Z}^d$).

Calculer $d(x)$ pour la marche aléatoire symétrique sur $\mathbb{Z}^d$.

4. **mesures invariantes** On définit la mesure suivante sur le graphe :

$$\mu(x) = d(x).$$

a) Montrer que μ est une mesure réversible.

b) On suppose que E est fini. Donner une distribution invariante de la chaîne de Markov. En existe-t-il d'autres ?

c) On suppose le graphe fini et connexe. Donner l'espérance du temps de retour en un point x.

d) On suppose dans cette question que le graphe est fini, mais plus qu'il est connexe. Peut-il exister plusieurs mesures invariantes ? Donner la forme générale de toutes les mesures invariantes.

5. **Mesure d'occupation**

a) Donner un exemple de graphe non-apériodique.

b) On suppose le graphe apériodique. Soit x un point du graphe. Montrer que, quelle que soit la distribution initiale,

$$\mathbb{P}(X_n = x) \to \frac{d(x)}{2\#A}.$$

6. **Théorème ergodique.** Déterminer le temps moyen passé par la chaîne en un point x.

7. **Application aux échecs**.

a) On considère une tour que l'on déplace aléatoirement sur un échiquier (8x8 cases). Chacune des cases qui lui sont accessibles ont même probabilité à chaque coup. On rappelle qu'une tour ne peut faire qu'un mouvement horizontal ou vertical à chaque coup. Quel est le temps moyen de retour au point de départ ? (en fonction du point de départ ?). Quelle est la période ?

b) même question pour un cavalier (mouvements autorisés : 2 cases dans une direction puis 1 case dans l'autre direction).

c) même question pour un fou (mouvement uniquement diagonaux).

Correction:

1.

$$Q(x, y) = \frac{1_{x \text{voisin de } y}}{d(x)}.$$

2. La chaîne est irréductible si pour tous points x, y il existe des états $x_1 = x, \dots, x_q = y$ tels que chaque transition $Q(x_i, x_{i+1})$ soit > 0, c'est-à-dire si le graphe est connexe.

3.a) Toute chaîne de Markov irréductible finie est récurrente positive.

b)On a vu dans l'exo correspondant que la marche aléatoire sur $\mathbb{Z}^2$ était récurrente, mais pas récurrente positive. La marche aléatoire sur $\mathbb{Z}^d, d \geq 3$ n'est même pas récurrente. Pourtant ces chaîne de Markov ont bien une structure de graphe, les voisins d'un point x étant tous les points reliés à x dans la grille $\mathbb{Z}^d$ (on a d'ailleurs dans ce cas $d(x) = 2d$).

4. a) Soit x, y. On a

$$\mu(x)Q(x,y) = d(x)\frac{\mathbf{1}_{x\sim y}}{d(x)} = \mathbf{1}_{x\sim y} = \mu(y)Q(y,x),$$

donc la mesure est réversible.

b) Comme la chaîne est IRP il existe une unique distribution stationnaire π. Comme μ est réversible, elle est invariante, et comme elle a une masse finie et est invariante, on a

$$\pi(x) = \frac{\mu(x)}{\mu(E)}.$$

On a

$$\mu(E) = \sum_{x\in E}\mu(x) = \sum_{x\in E} d(x) = \sum_{x\in E, y\in E}\mathbf{1}_{x\sim y} = 2\#A,$$

le facteur 2 dans la somme précédente venant du fait que chaque couple (x, y) est compté 2 fois ((x, y) et (y, x)).

On a donc

$$\pi(x) = \frac{d(x)}{2\#A}.$$

d) Supposons que E soit l'union disjointe de sous-ensembles $E_1, E_2, \ldots$, chaque sous-ensemble E_i étant connexe pour la structure d'arêtes A. Sur chaque E_i, on note π_i l'unique distribution invariante correspondante, donnée d'après c) par

$$\pi_i(x) = \frac{d(x)}{\#A_i}, x \in E_i,$$

A_i étant le nombre d'arêtes impliquant des points de E_i.

Soit $\alpha_1, \alpha_2, \ldots$ des nombres positifs tels que

$$\sum_i \alpha_i = 1.$$

On pose

$$\pi(x) = \begin{cases}\alpha_i\pi_i(x) \text{ si } x \in E_i.\end{cases}$$

π est une mesure, et c'est une mesure de probabilités car

$$\mu(E) = \sum_i \mu(E_i) = \sum_i \alpha_i\pi_i(E_i) = \sum_i \alpha_i = 1.$$

Montrons que π est stationnaire.

$$\begin{aligned}\mathbb{P}_\pi(X_1 = x) &= \sum_i \sum_{x\in E_i}\mathbb{P}(X_0 = x)\mathbb{P}_x(X_1 = x)\\ &= \sum_i \alpha_i(\sum_{x\in E_i}\pi_i(x)\mathbb{P}_x(X_1 = x))\\ &= \sum_i \alpha_i\pi_i(x) = \pi(x)\end{aligned}$$

ou à l'avant-dernière ligne on a utilisé le fait que π_i est stationnaire pour la restriction de la chaîne à E_i.

On remarque en particulier qu'il y a une infinités de distributions invariantes.

Supposons pour la réciproque que π soit une distribution invariante. On appelle toujours π_i l'unique distribution invariante sur E_i.

On pose $\alpha_i = \pi(E_i) = \mathbb{P}(X_0 \in E_i)$ si X_0 a probabilité π. Si $\alpha_i \neq 0$ on considère la chaîne $X^{(i)}$ définie comme X conditionnée par le fait que $X_0 \in E_i$. C'est une chaîne de Markov à valeurs dans E_i. Sa distribution initiale est, pour $x \in E_i$,

$$\mathbb{P}(X_0^{(i)} = x) = \mathbb{P}(X_0 = x | X_0 \in E_i) = \frac{\mathbb{P}(X_0 = x)}{\mathbb{P}(X_0 \in E_i)} = \frac{\pi(x)}{\alpha_i}.$$

La distribution de $X_1^{(i)}$ est, pour $x \in E_i$,

$$\mathbb{P}(X_1^{(i)} = x) = \frac{\mathbb{P}_\pi(X_1 = x, X_0 \in E_i)}{\mathbb{P}(X_0 \in E_i)} == \frac{\mathbb{P}_\pi(X_1 = x)}{\mathbb{P}(X_0 \in E_i)} = \frac{\pi(x)}{\alpha_i}$$

car π est invariante pour X. La troisième inégalité provient du fait que si $X_1 \in E_i$, comme les E_i sont non-connectés, alors X_0 est aussi nécessairement dans E_i. On en déduit que $\frac{\pi}{\alpha_i}$ est une distribution stationnaire pour $X^{(i)}$, elle est donc égale à l'unique distribution stationnaire π_i sur E_i. On a donc

$$\pi(x) = \alpha_i \pi_i(x), x \in E_i$$

(le cas $\alpha_i = 0$ se traite aisément). On a de plus

$$1 = \sum_x \pi(x) = \sum_i \sum_{x \in E_i} \alpha_i \pi_i(x) = \sum_i \alpha_i.$$

On a donc montré que toute distribution invariante s'écrivait sous cette forme.

5.c) Il faut appliquer le théorème de convergence à l'équilibre, la chaîne étant IRPA.

6)La chaîne étant irréductible, on a

$$\frac{V_n(x)}{x} \to \pi(x) = \frac{d(x)}{2\#A}.$$

7)a) Chaque case de l'échiquier est connectée avec 7 autres cases horizontalement, et 7 autres cases verticalement. On a donc pour tout x $d(x) = 14$. La chaîne est irréductible (la tour peut aller n'importe ou en 2 coups), et donc IRP. La distribution invariante est

$$\pi(x) = \frac{d(x)}{\sum_{x \in E}} = \frac{14}{64 * 14} = \frac{1}{64},$$

c'est la distribution uniforme.

On a donc pour tout x

$$\mathbb{E}_x(T_x^{(1)}) = \frac{1}{\pi(x)} = 64.$$

Comme la tour peut aller n'importe ou en $n_0 = 2$ ou $n_0 + 1 = 3$ coups, la chaîne est apériodique (on peut donc appliquer le théorème de convergence à l'équilibre.)

b)Cette fois, les cases n'ont pas toutes le même nombre de voisins. Les cases situées à au moins 2 cases du bord, de type 1, au nombre de n_1=4x4=16, ont chacune $v_1 = 8$ voisins.

Les cases situés à une case d'un bord et au moins 2 cases d'un autre, de type 2, au nombre de n_2 =16, ont chacune $v_2 = 6$ voisins.

Les cases situées sur un bord, et à au moins 2 cases d'un autre bord, de type 3, au nombre de n_3 =16, ont chacune $v_3 = 4$ voisins.

Les cases situées en diagonale d'u coin, c'est-à-dire B2 et symétriques, au nombre de n_4 =4, de type 5, ont chacune $v_4 = 4$ voisins.

Les cases à une case du coin, c'est-à-dire A2, B1 et symétriques, de type 5, au nombre de n_5 =8, ont chacune $v_5 = 3$ voisins.

Enfin les coins, de type 6, au nombre de $n_6 = 4$, ont chacun $v_6 = 2$ voisins.
On a bien $\sum_i n_i$ =16+16+16+4+8+4=64 cases.
Pour calculer la distribution de retour, il faut connaître le nombre total d'arêtes,

$$\sum_x d(x) = \frac{1}{2}\sum_i n_i v_i = \frac{1}{2}(16(8+6+4)+4*4+8*3+4*2) = \frac{16}{2}(19+2) = 8*21 = 168.$$

Le temps de retour en partant d'une case x qui est de type i est donc

$$\mathbb{E}_x(T_x^{(1)}) = \frac{v_i}{168}.$$

Nous allons montrer que la période est de 2. Etant donné une case x sur la i-ème ligne et la j-ème colonne, on pose

$$n_x = i + j[2] \in \{0,1\}.$$

Si l'on bouge le cheval d'une case $x = (i,j)$ à une case $y = (i',j')$, on a

$$n_x - n_y = i - i' + j - j' = 2 + 1 = 1[2].$$

Donc si X_n est la position de la chaîne de Markov la suite n_{X_n} effectue ...$0 \to 1 \to 0 \to 1 \to 0 \to$ .... Il est donc impossible pour X_n de revenir sur ses pas en un nombre impair de coups (si on part à $n_{X_0} = 0$, pour n impair on a $n_{X_n} = 1$ donc $X_n \neq X_0$). La période est donc un multiple de 2.

On peut montrer avec un dessin qu'on peut revenir sur ses pas en $n_0 = 4$ ou $n_0 + 2 = 6$ coups. La période est donc au maximum de 2. On en déduit que la période est de 2.

c) Concernant le fou, il faut tout d'abord noter que la chaîne n'est pas irréductible (le fou reste soit sur les cases noires, soit sur les cases blanches). Il faut donc étudier la chaîne au sein des 2 classes irréductibles. Il faut compter tous les cas comme dans le cas du cavalier pour obtenir les temps de retour. Comme pour la tour, la chaîne est apériodique.

7 Chaînes de Markov en temps continu, processus de Poisson

Le but de ce chapitre est de généraliser à des processus continus la notion de processus de Markov. Le temps passe donc de $\{0, 1, 2, \ldots\}$ à un intervalle réél. On garde pour l'instant un espace d'états discret E. Pour simplifier on prend $E = \mathbb{Z}$.

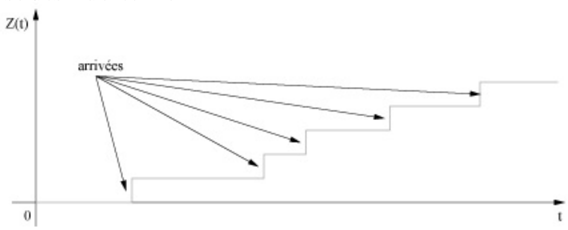
On appelle processus stochastique la donnée de variables X_t, pour $t \in I$ où I est un intervalle de $\mathbb{R}$. On prendra ici en général $I = \mathbb{R}_+$.

Enonçons la **propriété de Markov continue** pour un tel processus :

$$\forall n \geqslant 1, 0 < t_1 < t_2 < \cdots < t_n, x_1 \in \mathbb{N}, \ldots, x_n \in \mathbb{N},$$
$$\mathbb{P}(X_{t_n} = m_n | X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n-1}) = \mathbb{P}(X_{t_n} = m_n | X_{t_{n-1}} = x_{n-1})$$

7.1 Processus de sauts

On s'intéresse à un processus indexé par $I = \mathbb{R}_+$, qui ne peut évoluer que par accroissements de 1, ou "sauts" de taille 1.



On appelle $T_1, T_2, \ldots$ les variables aléatoires positives tels que les sauts interviennent aux instants $T_1, T_1 + T_2, T_1 + T_2 + T_3, \ldots$. Ainsi, le processus $(X_t)_{t \geqslant 0}$ correspondant peut être défini de la manière suivante :

$$X_t = \max\{k : T_1 + \cdots + T_k \leqslant t < T_1 + \cdots + T_{k+1}\}.$$

Cela revient à dire que X_t est égal à T_k sur tout intervalle $[T_1 + \cdots + T_k, T_1 + \cdots + T_{k+1}[, k \in \mathbb{N}$.

Si les T_k ne sont pas indépendants, cela signifie que posséder des informations sur, par exemple, T_1, me donne des informations sur, par exemple, T_2, et dans ce cas le passé nous renseigne sur le futur. On a donc peu de chances d'avoir un processus de Markov.

On suppose désormais que les T_i sont indépendantes. Pour avoir l'homogénéité en temps (i.e. le comportement à un temps t est le même qu'à tout autre temps $s \neq t$), on supposé également que les T_i sont identiquement distribués.

Exercice 82. On suppose que les T_i sont IID et $T_1 \sim \mathcal{U}_{[0,1]}$. Montrer que $(X_t)_{t \geqslant 0}$ n'est pas un processus de Markov.

Correction: Soit $t \in]0, 1[$. Comme $T_2 < 1$ presque sûrement,

$$\mathbb{P}(X_{t+1} = 1 | X_{t+1/2} = 1, X_t = 1) = 0$$

alors que

$$\mathbb{P}(X_{t+1} = 1 | X_{t+1/2}) \geqslant \mathbb{P}(T_1 \in [t, t + 1/2], T_1 + T_2 > t + 1) > 0.$$

Si $(X_t)_{t \geqslant 0}$ était un processus de Markov, ces deux quantités devraient être égales.

Le fait que les variables uniformes sont bornées implique une forme de mémoire : Si je sais qu'au bout d'un temps 1/2 le saut ne s'est pas réalisé, il ne reste plus qu'un temps 1/2 pour qu'il se réalise. La probabilité qu'il se réalise augment au fur et à mesure qu'on approche de 1.

Pour que ce processus soit de Markov, il faut que les temps d'attente T_i entre chaque saut soient sans mémoire.

On dit qu'une variable aléatoire T n'a pas de mémoire si elle vérifie la condition suivante. Pour $t, s \geqslant 0$,

$$\mathbb{P}(T \in [t, t+s])|T \geqslant t) = \mathbb{P}(T \in [0, s]).$$

Si les T_i sont toutes exponentielles avec le même paramètre $\lambda > 0$, on dit que le processus $(X_t)_{t \geqslant 0}$ est un processus de Poisson de paramètre λ.

Théorème 15. *Le processus de Poisson de paramètre $\lambda > 0$ est un processus de Markov.*

Exercice 83. Les seules variables aléatoires sans mémoire sont les variables exponentielles, c'est-a-dire avec fonction de distribution

$$\mathbb{P}(T \geqslant t) = \exp(-\lambda t), t \geqslant 0,$$

où $\lambda > 0$ est le paramètre de la loi.

Correction: Il est équivalent d'écrire la propriété de non-mémoire par, pour une variable $T \geqslant 0$.

$$\mathbb{P}(T \geqslant t+s|T \geqslant t) = \mathbb{P}(T \geqslant s)$$

Soit F la fonction de distribution d'une variable T sans mémoire.

$$\mathbb{P}(T \geqslant t+s|T \geqslant t) = \mathbb{P}(T \geqslant t)^{-1}\mathbb{P}(T \geqslant t+s) = F(t+s)F(t)^{-1} = F(s).$$

On obtient une équation fonctionnelle dont les seules solutions sont les fonctions exponentielles $t \mapsto \exp(-\lambda t)$.

Pour le montrer, on peut poser $L(t) = \ln(T)$ (à valeurs éventuellement $= -\infty$), et remarquer que $L(t+s) = L(t) + L(s)$ pour $t, s \geqslant 0$. On veut montrer que $L(t) = tL(1)$ pour $t \geqslant 0$, et que donc L est une fonction linéaire.

Comme $L(0)+L(0) = L(0+0) = L(0)$, on a $L(0) = 0$. On a ensuite $L(2) = L(1)+L(1) = 2L(1)$. De même, $L(3) = L(2) + L(1) = 3L(1)$, etc...Par récurrence, $L(k) = kL(1)$ pour $k \in \mathbb{N}$.En raisonnant de la même manière pour $n \geqslant 1$, $k \in \mathbb{N}$, $kL(1/n) = L(k/n)$. Pour $k = n$, on a $L(1) = L(n/n) = nL(1/n)$, d'où $L(1/n) = n^{-1}L(1)$. On a donc montré que pour ton nombre rationnel $x = k/n$, $L(x) = kL(1/n) = \frac{k}{n}L(1) = xL(1)$. Comme la fonction F est continue à droite, il en est de même de la fonction L. Donc pour $x \in \mathbb{R}$ irrationnel, on peut approximer x par des rationnels $x_q quil'approximentparladroite$, i.e. $x_q \to x$ pour$q \to \infty$ e t$x_q > x$, et $x_q L(1) = L(x_q) \to L(x)$, donc $L(x) = xL(1)$.

Donc on a $F(t) = \exp(tL(1)), t \geqslant 0$. Comme $F(t) \to 0$ quand $t \to \infty$, $L(1) < 0$, donc on a bien $F(t) = \exp(-\lambda t)$ pour un $\lambda > 0$.

Au lieu de matrice de transition, on parle pour le processus X_t de noyau de transition $Q_t(x, y), t > 0, x, y \in \mathbb{N}$,

$$Q_t(x, y) = \mathbb{P}(X_t = y|X_0 = x).$$

Ainsi, la plupart des évènements qui concernant le processus X_t sont entièrement déterminés par le noyau de transition.

Poisson vérifie-t-il Markov ? La réponse est oui, pour le montrer, commençons par le résultat suivant.

Proposition 17. *Pour $t > 0$, soit N_t une variable de Poisson de paramètre λt. Pour le processus de Poisson,*

$$Q_t(x, y) = \mathbb{P}(N_t = y - x),\ t > 0, x \leqslant y \in \mathbb{N}.$$

Exercice 84. :Preuve! On remarque que $\mathbb{P}(X_t = k) = \mathbb{P}(T_1 + \cdots + T_k \leqslant t,\ T_1 + \cdots + T_{k+1} > t)$

1. Ecrire la densité $f_{k+1}(t_1, \ldots, t_{k+1})$ de $(T_1, \ldots, T_{k+1})$
2. Calculer

$$\mathbb{E}\left[\mathbf{1}_{\{T_1+\cdots+T_k \leqslant t\}} \mathbb{P}(T_{k+1} > t - (T_1 + \cdots + T_k) | T_1, \ldots, T_k)\right]$$

et en déduire le résultat.

1

Correction:

1. Comme les T_i sont IID, la densité est

$$\prod_{i=1}^{k+1} \lambda e^{-\lambda t_i} = \lambda^{k+1} e^{-\lambda(t_1+\cdots+t_{k+1})}$$

2. La probabilité recherchée est donc

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\{T_1+\cdots+T_k \leqslant t\}} \exp(-\lambda(t - (T_1 + \cdots + T_k)))\right] \\
&= \int_{(\mathbb{R}_+)^k} \lambda^k \mathbf{1}_{\{t_1+\cdots+t_k \leqslant t\}} \exp(-\lambda(t_1 + \cdots + t_k)) \exp(-\lambda(t - (T_1 + \cdots + T_k))) dt_1 \ldots dt_k \\
&= \lambda^k \exp(-\lambda t) \int_{(\mathbb{R}_+)^k} \mathbf{1}_{\{t_1+\cdots+t_k \leqslant t\}} dt_1 \ldots dt_k \\
&= \lambda^k \exp(-\lambda t) t^k \underbrace{\int_{(\mathbb{R}_+)^k} \mathbf{1}_{\{t_1+\cdots+t_k \leqslant 1\}} dt_1 \ldots dt_k}_{\alpha_k}
\end{aligned}$$

Pour montrer $\alpha_k = 1/k!$, on peut utiliser une récurrence : $\alpha_1 = 1$, et

$$\begin{aligned}
\alpha_k &= \int_{(\mathbb{R}_+)^k} \mathbf{1}_{\{t_1+\cdots+t_{k-1} \leqslant 1-t_k\}} dt_1 \ldots dt_k \\
&= \int_0^1 \int_{(\mathbb{R}_+)^{k-1}} \mathbf{1}_{\{\frac{t_1+\cdots+t_{k-1}}{1-t_k} \leqslant 1\}} dt_1 \ldots dt_{k-1} dt_k \\
&= \int_0^1 (1-t_k)^{k-1} \int_{(\mathbb{R}_+)^{k-1}} \mathbf{1}_{\{t_1+\cdots+t_{k-1} \leqslant 1\}} dt_1 \ldots dt_{k-1} dt_k \\
&= \alpha_{k-1} \int_0^1 (1-t_k)^{k-1} dt_k = \frac{\alpha_{k-1}}{k},
\end{aligned}$$

d'où $\alpha_k = \frac{1}{k!}$.

Proposition 18. *On déduit de l'exercice précédent le résultat suivant : Le processus de Poisson est à accroissements stationnaires, c'est-à-dire*

$$\forall t_1, t_2, s > 0,\ X_{t_1+s} - X_{t_1} \overset{(d)}{=} X_{t_2+s} - X_{t_2}.$$

7.2 Générateur infinitésimal

Plutôt que la matrice de transition, avec un processus de Markov X_t en temps continu, on préfère travailler avec le *générateur infinitésimal* du processus, défini de la manière suivante : Soit f une fonction bornée et dérivable. On pose

$$Lf(x) = \lim_{\varepsilon \to 0} \frac{\mathbb{E}\left[f(X_{t+\varepsilon}) - f(X_t)\right]}{\varepsilon}$$

en supposant que la limite existe. Remarquons que si le processus est homogène, la limite ne dépend pas de t.

Si de plus le processus est à accroissements stationnaires, comme c'est le cas pour le processus de Poisson, cet opérateur ne dépend pas de t : En effet, $X_{t+\varepsilon} - X_\varepsilon \stackrel{(d)}{=} X_\varepsilon - X_0$ pour tous $t, \varepsilon > 0$.

Dans ce cas, l'opérateur L transforme une fonction en une autre fonction, qui dénote la manière dont $f(X_t)$ varie au voisinage de 0 si $X_0 = x$.

Dans le cas Poissonien, $X_\varepsilon - X_0$ est une variable de Poisson de paramètre $\lambda\varepsilon$. Pour ε petit, on a

$$\begin{aligned}
\mathbb{P}(X_\varepsilon = 0) &= \exp(-\varepsilon\lambda), \\
\mathbb{P}(X_\varepsilon = 1) &= \exp(-\varepsilon\lambda)\frac{\lambda\varepsilon}{1!} \\
\mathbb{P}(X_\varepsilon \geqslant 2) &= \sum_{k=2}^{\infty} (\lambda\varepsilon)^k \frac{e^{-\lambda\varepsilon}}{k!} \leqslant \sum_{k=0}^{\infty} (\lambda\varepsilon)^{k+2} \frac{1}{(k+2)!} \leqslant \varepsilon^2 \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{(k+2)!}}_{<\infty}.
\end{aligned}$$

Donc

$$\begin{aligned}
\mathbb{E}[f(X_\varepsilon) - f(X_0)|X_0 = x] &= (f(x+0) - f(x)) * \exp(-\lambda\varepsilon) + (f(x+1) - f(x))\lambda\varepsilon\exp(-\lambda\varepsilon) + 2\|f\|_\infty o(\varepsilon) \\
Lf(x) &= \lambda(f(x+1) - f(x))
\end{aligned}$$

- On dit que X_t est un *processus de sauts*, car il ne peut varier que par discontinuités.
- Dans ce cas, X_t est un processus de saut à taux constant λ : λ est l'intensité avec laquelle le processus *saute*, sous-entendu la répartition des sauts est Poissonnienne, c'est-à-dire que les sauts sont séparés par des variables exponentielles IID.

7.3 Processus de Poisson composé

Modifions le modèle précédent de la manière suivante. A chaque "saut" du processus, au lieu de sauter de 1, le processus varie d'une quantité aléatoire.

On introduit une loi de probabilité μ sur $\mathbb{R}$, qui représente la loi d'un saut typique.

Par exemple, si la probabilité de faire un saut négatif est de 1/2, cela nous donne $\mu(]-\infty, 0[) = \frac{1}{2}$, et $\mu(]0, \infty[) = \frac{1}{2}$. On suppose qu'il n'y a pas de saut de taille 0 ($\mu(\{0\}) = 0$). Pour s'en faire une représentation plus physique, on peut imaginer que chaque temps de saut correspond au temps d'arrivée dans une banque donnée, et que la taille du saut représente une quantité d'argent que le client va déposer à la banque (ou retirer si le saut est négatif).

Pour modéliser la quantité d'argent déposée à un instant $t \geqslant 0$, on introduit une suite de variables iid $Y_k, k \geqslant 1$, indépendants de μ_t, avec comme loi commune μ, et on suppose que le i-ème client a apporté une quantité d'argent $Y_k \in \mathbb{R}$.

En appelant X_t le processus de Poisson défini au chapitre précédent (avec paramètre d'intensité $\lambda > 0$), la quantité d'argent déposée à l'instant t est donc

$$Z_t = \sum_{k=1}^{\infty} Y_k \mathbf{1}_{\{X_t \geqslant k\}},$$

en gros, on comptabilise les sommes de tous les clients qui sont effectivement déjà passés à l'instant t.

En utilisant le fait que X_t est une variable de Poisson de paramètre λt, et que les Y_i sont IID de loi μ, on peut déterminer la loi de Z_t, via la fonction caractéristique.

Rappel sur les fonctions caractéristiques Pour toute variable aléatoire $Y \in \mathbb{R}$, on introduit sa fonction caractéristique (ou transformée de Fourier)

$$\psi_Y(\theta) = \mathbb{E}[\exp(i\theta Y)] = \int_{\mathbb{R}} e^{i\theta y}\mu(dy),$$

où μ est la loi de Y. Cette fonction possède de nombreuses propriétés, comme par exemple

— $\psi_Y(0) = 1$

— $|\psi_Y(\theta)| \leqslant 1, \theta \in \mathbb{R}$,

— ψ_Y caractérise la loi de Y, dans le sens où $\psi_Y = \psi_{Y'}$ implique $Y \overset{(d)}{=} Y'$.

Donc, déterminer la loi de Y revient à déterminer sa fonction caractéristique. De plus, si $\mathbb{E}[|Y|] < \infty$, on peut calculer l'espérance de Y avec sa fonction caractéristique

$$\mathbb{E}[Y] = -i\psi_Y'(0) = -i\frac{d}{d\theta}|_{\theta=0}\psi_Y(\theta).$$

Si $\mathbb{E}[Y^2] < \infty$, on peut calculer son moment d'ordre 2,

$$\mathbb{E}[Y^2] = -\psi_Y''(0) = -\frac{d^2}{d\theta^2}|_{\theta=0}\psi_Y(\theta).$$

Théorème 16. *Avec les notations précédentes, pour $t > 0$, la fonction caractéristique de Z_t est*

$$\psi_{Z_t}(\theta) = \exp\left(\lambda t \left(\psi_{Y_1}(\theta) - 1\right)\right), \theta \in \mathbb{R}.$$

Démonstration:

$$\begin{aligned}
\mathbb{E}[\exp(i\theta Z_t)] &= \sum_{n=0}^{\infty} \mathbb{P}(X_t = n)\mathbb{E}\left[\exp\left(i\theta\sum_{k=1}^{\infty}\mathbf{1}_{\{X_t \leqslant k\}}Y_k\right)|X_t = n\right] \\
&= e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}\mathbb{E}\left[\exp\left(i\theta\sum_{k=1}^{n}Y_k\right)\right] \\
&= e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}\mathbb{E}[\exp(i\theta Y_1)]^n \text{ grâce à l'indépendance entre les } Y_k \\
&= e^{-\lambda t}\exp(\lambda t\mathbb{E}[\exp(i\theta Y_1)])
\end{aligned}$$

Cela nous permet par exemple de déterminer les premiers moments de Z_t :

$$\begin{aligned}
\mathbb{E}Z_t &= -i\frac{d}{d_\theta}|_{\theta=0}\psi_{Z_t}(\theta) = -i\lambda t\psi_{Y_1}'(0)\exp\left(\lambda t(\underbrace{\psi_{Y_1}(0)}_{=1} - 1)\right) = \lambda t\mathbb{E}[Y_1] \\
\mathbb{E}Z_t^2 &= -\frac{d^2}{dt^2}|_{t=0}\psi_{Z_t}(\theta) = -\exp(\ldots)\left[\lambda t\psi_{Y_1}''(0) + (\lambda t\psi_{Y_1}'(0))^2\right] = (\lambda t\mathbb{E}Y_1^2 + \lambda^2t^2(\mathbb{E}Y_1)^2)
\end{aligned}$$

Du coup, la variance est

$$\mathrm{Var}(Z_t) = \mathbb{E}[Z_t^2] - (\mathbb{E}Z_t)^2 = \lambda t\mathbb{E}Y_1^2.$$

Exercice 85. Quel est le générateur infinitésimal du processus Z_t ?

Correction: Soit f une fonction bornée. Conditionnellement à $Z_0 = x$, pour $\varepsilon > 0$

$$\begin{aligned}
\mathbb{E}[f(Z_\varepsilon) - f(x)] =& \sum_{n=0}^{\infty}\mathbb{P}(X_\varepsilon = n)\mathbb{E}\left[f(x + \sum_{k=1}^{n}Y_k) - f(x)|X_\varepsilon = n\right] \\
=& 0 + \lambda\varepsilon\exp(-\lambda\varepsilon)\mathbb{E}[f(x + Y_1) - f(x)] + o(\varepsilon^2)
\end{aligned}$$

Donc

$$Lf(x) = \lambda \mathbb{E}[f(x + Y_1) - f(x)].$$

Le GI de Z_t est

$$Lf(x) = \lambda \mathbb{E}[f(x + Y_1) - f(x)].$$

Exercice 86. Supposons que la quantité d'argent déposée par chaque client soit le produit de 10 par une variable de Poisson de paramètre 2. Déterminer la fonction caractéristique, l'espérance, et la variance, de Z_t pour $t > 0$.

7.4 Exercices

8 TP 1 - Moteur de recherche

Le but de ce TP est de construire un moteur de recherche qui classe les pages du web selon une recherche par mots-clés.

8.1 Construction du "Web"

Dans ce TP, une page web est modélisée par un mot-clé, qui est réduit ici pour simplifier à une lettre entre `A` et `Z`, et par une liste de liens. Le nombre de pages est de `n`, et les liens sont représentées par des `1` dans une matrice `n`×`n`. La présence d'un `1` à la ligne `i` et la colonne `j` indique la présence d'un lien de la page `i` vers la page `j`. A noter que par convention, une page ne pointe jamais vers elle-même.

1. Créer un programme `W=randomWeb(n)` qui renvoie une liste de `n` "mots-clés". Chaque entrée de cette liste représente une page web.
2. Créér une matrice aléatoire `M=linkMatrix(n)` (variable globale) qui contient les informations de liens entre les pages web, en respectant les conditions suivantes :
   — Une page ne peut pointer vers elle-même
   — Chaque page contient au moins un lien

Correction:

```
function  M=randomWeb(n)

val="A":"Z";
probas=ones(1,26)/26;

Web=discrete_rnd (val,probas,1,n);

function  M=linkMatrix(n)
disp(1)
M=zeros(n,n);

for i=1:n
 do
  for j=1:n
  if j!=i
   M(i,j)=(rand()<0.5);
  end
end
  until (sum(M(i,:))!=0)
end
```

8.2 Calcul de PageRank via les valeurs propres

On rappelle que le crédit qu'une page web x appporte à une autre page web y est défini par

$$C(x,y) = \frac{1_{\{\text{x pointe vers y}\}}}{\#\{\text{liens présents sur la page x}\}}.$$

On suppose qu'il existe une quantité $I(x)$ appelée "importance" pour chaque page web x telle que

$$I(y) = \sum_x C(x,y)I(x).$$

1. Montrer que cela équivaut à : I^t est vecteur propre de C^t associé à la valeur 1. Pourquoi I existe bien ?
2. Calculer C à partir de `M` défini par `M=randomWeb(10)`. En déduire I à l'aide de l'instruction `eig(M')`.

Correction: C'est la définition du vecteur propre.

Comme C est une matrice stochastique, $C \cdot (1, 1, \ldots, 1)^t = (1, \ldots, 1)^t$, donc $(1, \ldots, 1)^t$ est vecteur propre associé à la valeur propre 1. Donc 1 est dans le spectre de C, donc il est également dans le spectre de C^t ($C - Id$ pas inversible équivaut à $C^t - Id$ pas inversible).

```
function I=importance(M)

%matrice de transition
%nombre reel de liens
numberOfLinks=sum(M);
D=diag(1./numberOfLinks);
C=D*M

%calcul de l'importance
[V, LAMBDA] = eig (C');
%I est le vecteur propre associé a la plus grande valeur propre, à savoir 1
I=V(:,1) ;
```

8.3 Estimation du PageRank avec un surfeur aléatoire

1. Soit `M=linkMatrix(n)`. Construire un programme `v=surfer(N,M)` qui renvoie un vecteur `v` de taille `(1,n)` contenant le nombre de visites du surfeur aléatoire sur chaque page après `N` clics sur le web dont les liens sont définis par la matrice `M`.
2. Comparer le vecteur `v` ainsi obtenu au vecteur `I` calculé précédemment. On divisera chacun de ces vecteurs par sa somme pour avoir la bonne normalisation.

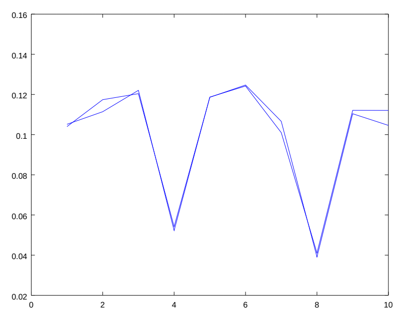
   Correction:

```
  function v=surfer(N,M)
 n=size(M,1)
      %v contient le nombre de visites pour chaque page
      v=zeros(1,n);
      % on commence à la page 1
      position=1;
      for k=1:N
        v(position)=v(position)+1;
           %on choisit la transition d?après la ligne de la page
          ligne=find(M(position,:)==1);
          l=length(ligne);
        position= ligne(unidrnd(l));

endfor
```

Ensuite on fait dans la fenêtre de commande :

```
n=10;
N=10000;
M=linkMatrix(n);
I=importance(M);
v=surfer(N,M);
I=I./sum(I);
v=v./sum(v);

plot(v)
hold on
plot(I)
```

On obtient la figure suivante :



On voit que `v` et `I` sont en effet très proches. L'algorithme semble donc donner une manière d'approximer la distribution invariante.

8.4 Interprétation

Soit $X = (X_n)$ la chaîne de Markov d'état initial 1 et de matrice de transition $C(x, y)$.

3. Comment tester empiriquement si la chaîne de Markov est irréductilbe ?

   Correction: Calculons `v=surfer(N,M)` pour une grande valeur de `N`, comme `N=10000`. On obtient

```
\texttt{v =

   1076    1137    1214     539    1161    1225    1048     387    1098    1115}
```

On voit que tous les états ont été visités un grand nombre de fois. Si la chaîne de Markov n'avait pas été irréductible, les états d'une des classes n'auraient plus été visités à partir d'un certain temps, et donc le nombre de visites en ces états aurait été faible.

4. On suppose qu'elle est irréductible. Pourquoi la chaîne de Markov correspondante a-t-elle une unique distribution invariante π ? Construire une variable `piEstim(M)` qui donne une estimation

de π. Pour observer la convergence, on pourra représenter sur le même graphe le vecteur `pi`, le vecteur `I` trouvé précédemment, la variation totale entre les deux :

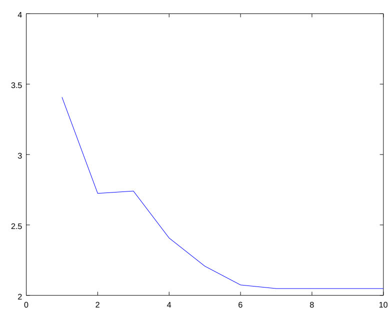
$$d_{VT}(\texttt{pi}, \texttt{I}) = \sum_{k=1}^{n} |\texttt{pi[k]} - \texttt{I[k]}|$$

et commenter.

Correction: Elle est irréductible sur un espace fini, donc IRP, donc elle admet une unique distribution invariante π. Remarquons que π est l'unique 1-vecteur propre de C^t, donc π est égal à `I./sum(I)` (modulo les approximations numériques du programme). On modifie le programme surfer pour mesurer au fur et à mesure la distance :

```
 function estim=piEstim(N,M)
 n=size(M,1)
      %v contient le nombre de visites pour chaque page
      v=zeros(1,n);

      %on calcule la "vraie" valeur de pi:
      I=importance(M);
      pi=I./sum(I);

      %on va observer l'evolution de la distance entre v et pi dans la liste dist
      dist=zeros(1,N);

      position=1;
      for k=1:N
        v(position)=v(position)+1;
           %on choisit la transition d?après la ligne de la page
          ligne=find(M(position,:)==1);
          l=length(ligne);
        position= ligne(unidrnd(l));


        estim=v./sum(v);
        dist(k)=sum(abs(estim-I'));

endfor
dist
plot(dist)
```

Pour $N = 10$, on observe la décroissance exponentielle :



8.5 Construction du moteur de recherche

Construire le programme `pageRank(a)` qui fonctionne de la manière suivante :

1. Tirage d'un web aléatoire de taille `n=1000000`.
2. Demande d'un mot clé à l'utilisateur (entre `'A'` et `'Z'`).
3. Estimer l'importance des pages (on fera attention à bien calibrer N).
4. Affichage des 10 pages contenant le mot clé les plus "importantes" dans l'ordre, avec leur score.

8.6 Raffinements

1. Supposons que l'on tente de renforcer la pertinence d'une page web en créant artificiellement 10 pages web qui pointent uniquement vers cette page. Est-ce que ça modifie la pertinence ? Et si les 10 nouvelles pages pointent également entre elles ?
2. Imaginer d'autres manières de tirer `linkMatrix` où certaines pages seraient plus influentes, et voir comment cela affecte la pertinence.

9 TP2 Simulation de polymères

Le but de ce TP est de simuler informatiquement un modèle simplifié de polymère. Ici, on choisit un nombre entier p, qui sera le nombre de particules de chaque polymère. Un polymère est une suite $x_1, \dots, x_p$ de points de $\mathbb{Z}^2$ qui satisfait aux contraintes suivantes :

— Pour $1 \leqslant i \neq j \leqslant p$, $[x_i, x_{\overline{i+1}}]$ et $[x_j, x_{\overline{j+1}}]$ ne se touchent pas, sauf éventuellement aux extrémités, où

$$\bar{i} = \begin{cases} i \text{ si } i < p \\ 1 \text{ si } i = p+1 \end{cases}$$

— 0 est dans "l'intérieur" du polymère

Le but est d'approcher vie l'algorithme de Metropolis une loi de distribution proportionnelle à la mesure

$$m(x) = \frac{\text{surface(x)}}{\text{Perimetre(x)}^4}.$$

9.1 Package `geometry`

On va utiliser le package `geometry`, qu'il faut télécharger sour la forme d'un fichier `.tar.gz`, puis installer via `pkg install geometry-xxx.tar.gz`, et charger via `load pkg geometry`.

On va utiliser les concepts suivants :

— `polygon` : matrice `p`×`2` (chaque ligne représente un point à 2 coordonnées)
— `drawPolygon(polygon)` : Représente le polygone `polygon` sur un graphique.
— `polygonArea(polygon)` : Calcule la surface de `polygon`. Il faut prendre la valeur absolue car la surface est "signée"
— `edgeLength([a b])` : Calcule la longueur du segment `[a,b]`, où `a`,`b` sont deux matrices 2×1.
— `randomPolygon(p)` : Tire aléatoirement un polygone à `p` sommets (fourni pour le TP)

Code de `randomPolygon` :

```
function  poly=randomPolygon(p)
tire au hasard un polygone aléatoire à p sommets

points=zeros(p,2);

val=-20:20;
probas=ones(1,41)/41;

for k=1:p
  points(k,1)=discrete_rnd(val,probas,1);
  points(k,2)=discrete_rnd(val,probas,1);
 end

 poly=closedPath(points);
```

9.2 Evolution libre du polymère

1. On va tout d'abord créér un polygone qui évolue librement, sans contraintes. Créer un programme `evolve(poly)` qui, étant donné une matrice `poly`, de taille `p`×`2`, renvoie une autre matrice `p`×`2`, où l'une des lignes du vecteur initial, correspondant à un point du polygone, s'est déplacé dans une des 4 directions possibles (ne pas tenir compte d'éventuels croisements ou superpositions pour l'instant).

```
function newPoly=evolve(poly)


p=size(poly,1);

val=1:p;
probas=ones(1,p)/p;
index=discrete_rnd(val,probas,1);



directions=["h","b","g","d"];
probas=ones(1,4)/4;
direction=discrete_rnd(directions,probas,1);

if direction=="h"
  increment=[0,1];
elseif direction=="b"
  increment=[0,-1];
elseif direction=="g"
  increment=[-1,0];
elseif direction=="d"
  increment=[1,0];
endif

newPoly=poly;
newPoly(index,:)=poly(index,:)+increment;
```

2. Créer un programme `polyEvolve(n,polygon)`, où `n` est un entier et `polygon` une liste de points (matrice p$\times$2) qui contient une boucle `for k=1:n`, où à chaque itération, `evolve` est appliqué à `polygon`. On observera l'évolution à l'aide de la commande `drawPolygon`.

```
function polyEvolve(n,polygon)


figure(1)

for k=1:n
  figure(1)
  clf;

  drawPolygon(polygone);
  drawnow;

  polygon=evolve(polygon);

end
```

3. Créer une fonction `m=score(poly)` qui renvoie la valeur

$$\texttt{m} = \frac{\texttt{polygonArea(poly)}}{\texttt{perimetre(poly)}^4}$$

La fonction `perimetre` n'existe pas, il faut la créer à l'aide de la fonction `edgeLength`.

```
function m=score(polygon);

surf=abs(polygonArea(polygon));

p=size(polygon,1);

polygonCyclic=[polygon;polygon(1,:)];

perimetre=0;
for k=1:p
  point1=polygonCyclic(k,:);
  point2=polygonCyclic(k+1,:);
  perimetre=perimetre+edgeLength([point1 point2]);
end

m=surf/(perimetre^4);
```

4. Modifier la fonction `evolve` pour qu'elle ne garde l'évolution que si le rapport des mesures nouveau/ancien excède la valeur d'une variable uniformément distribuée dans $[0, 1]$.

```
function newPoly=evolve(poly)

p=size(poly,1);

val=1:p;
probas=ones(1,p)/p;
index=discrete_rnd(val,probas,1);

directions=["h","b","g","d"];
probas=ones(1,4)/4;
direction=discrete_rnd(directions,probas,1);

if direction=="h"
  increment=[0,1];
elseif direction=="b"
  increment=[0,-1];
elseif direction=="g"
  increment=[-1,0];
elseif direction=="d"
  increment=[1,0];
endif

newPoly=poly;
newPoly(index,:)=poly(index,:)+increment;

mInit=score(poly);
```

```
m=score(newPoly);
u=unifrnd(0,1);
if (m>u*mInit)
'garde';
else
'jette';
  newPoly=poly;
end
```

5. Observer et interpréter l'évolution avec `p=4,n=1000`. On pourra ensuite changer la valeur de `p`.

9.3 Test de polymères

Dans cette dernière partie, on va s'assurer que les polymères respectent les contraintes prescrites (pas de croisement, etc...).

1. Créér un programme `isPolymer(polygon)`, qui renvoie `1` si `polygon` respecte les contraintes, et `0` sinon. On pourra utiliser les instructions suivantes
   — `polygonSelfIntersections (polygon)` : Donne les coordonnées des auto-intersections de `polygon`
   — `isPointInPolygon(a,polygon)` : Teste si le point `a` est à l'intérieur de `polygon`.

```
function answer=isPolymer(polygon)

answer=1;

intersections=polygonSelfIntersections (polygon);
if size(intersections,1)>0
  answer=0;
 end

 if isPointInPolygon([0 0],polygon)==0
  answer=0;
 endif
```

2. Modifier `polyEvolve` pour qu'une évolution ne soit acceptée que si le résultat respecte les contraintes. Observer les résultats comme à la fin de la partie précédente.

Programme qui étudie l'évolution conjointe de la surface, du périmètre, et du score :

```
function [surf,per,m,newPoly]=evolve(surfInit,perInit,mInit,poly)


p=size(poly,1);

val=1:p;
probas=ones(1,p)/p;
index=discrete_rnd(val,probas,1);



directions=["h","b","g","d"];
probas=ones(1,4)/4;
```

```
direction=discrete_rnd(directions,probas,1);

if direction=="h"
  increment=[0,1];
elseif direction=="b"
  increment=[0,-1];
elseif direction=="g"
  increment=[-1,0];
elseif direction=="d"
  increment=[1,0];
endif

newPoly=poly;
newPoly(index,:)=poly(index,:)+increment;

u=unifrnd(0,1);

[surf,per,m]=score(newPoly);
if (m>u*mInit && isPolymer(newPoly))
'garde';
else
'jette';
  surf=surfInit;
  per=perInit;
  m=mInit;
  newPoly=poly;
end

%%%%%%%%

function polyEvolve(n,p)

mValues=zeros(1,n);
sValues=zeros(1,n);
pValues=zeros(1,n);

polygone=randomPolygon(p);
[surf,per,m]=score(polygone);

figure(1)

for k=1:n
  figure(1)
  clf;

  subplot(4,1,1);
  plot(mValues(max(1,k-99):k-1))

  subplot(4,1,2);
  plot(sValues(max(1,k-99):k-1))

  subplot(4,1,3);
```

```
  plot(pValues(max(1,k-99):k-1))

  subplot(4,1,4)
  drawPolygon(polygone);


  drawnow;

  [surf,per,m,polygone]=evolve(surf,per,m,polygone);

  mValues(k)=m;
  sValues(k)=surf;
  pValues(k)=per;



end
mValues(n-1)
sValues(n-1)
pValues(n-1)
```

10 TP : Problème du voyageur de commerce

Soit $\mathcal{A}_N = \{1, ..., N\} \times \{1, .., N\}$ avec $N \in \mathbb{N}$. Soit $n \geqslant 1$ et $x = (x_1, \ldots, x_n)$ des points disjoints de $\mathcal{A}_N$. Le but de ce TP est de trouver rapidement un chemin court qui passe par tous les points x_i et revient à la position initiale. On prendra $N = 30, n = 15$ dans les simulations.

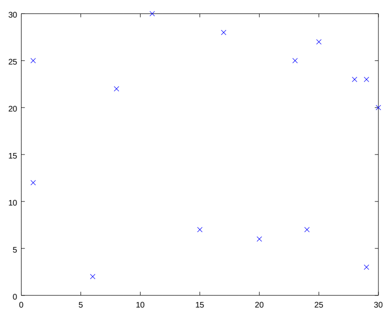
10.1 Préliminaires

1. Ecrire un programme `x=locations(n,N)` qui renvoie un vecteur `x` contenant `n` points disjoints de $\mathcal{A}_N$ tirés uniformément et les représente par des croix sur une figure. On pourra utiliser une matrice de taille $N \times N$ représentant toutes les locations possibles.

   Correction:

```
function x=locations(n,N)



M=zeros(N,N);

x=zeros(2,n);

for k=1:n
  test=1;
  while test==1
    i=unidrnd(N);
    j=unidrnd(N);
    test=M(i,j);
  end
  M(i,j)=1;
  x(1,k)=i;
  x(2,k)=j;
  plot(i,j,"x")
  hold on
end
```

   Exemple avec `x=locations(15)` :



2. Ecrire un programme `l=longueur(x)` qui calcule la longueur totale à parcouir pour aller de x_1 à x_2, puis de x_2 à x_3, ..., et revenir de x_n à x_1 (avec la norme 2 de $\mathbb{R}^2$).

   Correction:

```
function l=longueur(x)

n=length(x);
l=0;

for k=1:n
  a=x(:,k);
  if k<n
    b=x(:,k+1);
  elseif k==n
    b=x(:,1);
  end
  l=l+sqrt((b(1)-a(1))^2+(b(2)-a(2))^2);
end
```

   Dans la suite, on souhaite changer l'ordre de visite des x_i pour diminuer la longueur totale du trajet. Pour simplifier, on va chercher via l'algorithme de Metropolis un chemin aléatoire Y qui passe une fois par chaque x_i et tel que pour toute possibilité $y = (y_1, ..., y_n)$,

$$\mathbb{P}(Y = y) = \pi(y) := \frac{1}{Z} \exp(-\text{longueur}(y))$$

   pour une certaine constante Z qu'on n'a pas besoin de connaître.

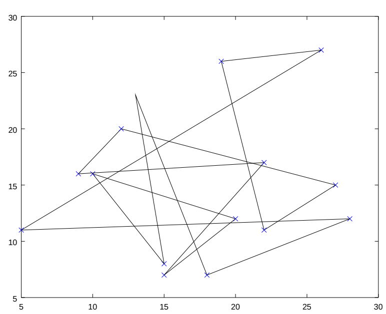
3. Ecrire un programme `draw(y)` qui trace avec des lignes l'itinéraire défini par un chemin $y \in (\mathcal{A}_N)^n$ à l'aide de l'instruction `line`.

   Correction:

```
function draw(y)
```

```
n=length(y);
hold on
for k=1:n-1
  a=y(1,k);
  b=y(2,k);
  c=y(1,k+1);
  d=y(2,k+1);
  plot(a,b,"x")
  line([a c],[b d])

  %pause
end
line([c y(1,1)],[d y(2,1)])
```

On voit que le chemin est loin d'être le plus court possible :



10.2 Algorithme d'optimisation

4. Ecrire un programme `y=echange(x,i,j)` qui échange les destinations i et j du chemin x.
   Correction:

   ```
   function y=echange(x,i,j)

   y=x;
   y(:,i)=x(:,j);
   y(:,j)=x(:,i);
   ```

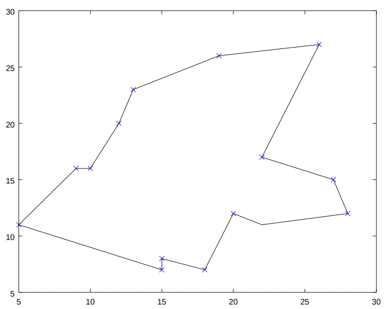
5. Formaliser la chaîne de Markov et le problème proposé. Etant donné $x \in (\mathcal{A}_N)^n$, quel est l'espace d'état ? Quelle est le noyau de transition $P(x, y)$ qui correspond à la question précédente ?
6. Etant donné deux chemins x, y, que vaut $\pi(x)/\pi(y)$? Ecrire un programme `r=ratio(x,y)` qui renvoie la valeur de $\pi(x)/\pi(y)$.

7. Ecrire un programme `y=Metropolis(x,k)` qui renvoie le chemin y obtenu après k itérations de l'algorithme de Metropolis sur les locations définies par le vecteur x. On dessinera le nouveau chemin à chaque itération.

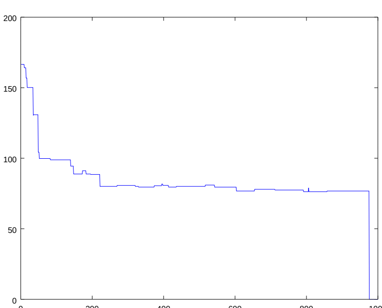
   Correction:

```
function [l,y]=Metropolis(x,k)


n=length(x)
l=zeros(1,k);
oldl=longueur(x);

y=x


figure(1)
clf
for i=1:k

  a=unidrnd(n);
  b=unidrnd(n);
  z=echange(y,a,b);
  U=rand();
  newl=longueur(z);
  if (-newl+oldl>log(U))

    y=z;
    oldl=newl;
    %pause
    figure(1)
    clf
    draw(y)
    drawnow
    figure(2)
    clf
    plot(l)
    drawnow
  end
  l(i)=longueur(y);


end
```

   Avec `Metropolis(x,1000)` on obtient le chemin suivant :



L'évolution des longueurs est la suivante :



10.3 Etude de la convergence

8. Modifier le programme précédent pour qu'il renvoie deux arguments, `l` et `y`, où `l` contient les longueurs du chemin à toutes les itérations. Tracer `l`.
9. Estimer visuellement le temps pris pour atteindre le minimum pour plusieurs valeurs de n différentes.
10. Dans cet algorithme, à chaque itération, les transitions vers un chemin plus court sont favorisées, mais on laisse tout de même la possibilité d'un chemin plus long, pour ne pas être piégé dans

un “minimum local”. Donnez un exemple simple de configuration permettant un minimum local qui n’est pas un mimimum global.

10.4 Estimation

11. Soit $X_1, \ldots, X_n$ n points uniformément répartis dans $[0, 1]$. Soit L_n la longueur du plus court chemin qui passe par tous les X_i et revient à sa position initiale. Donner une estimation de $\mathbb{E}(L_n)$ pour $1 \leqslant n \leqslant 15$.

11 Sujet d'examen

11.1 Juin 2011

Exercice 87. On considère une chaîne de Markov (X_n) dans $E = \{1, 2, 3\}$ avec matrice de transition

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

1. Quelles sont les classes de récurrence/transience ?
2. Calculer une mesure de probabilité invariante. Est-elle unique ? Est-elle réversible ?
3. Calculer pour tout $x \in E$ le temps moyen de retour à x, $\mathbb{E}(T_x^{(1)})$. Calculer la période de tout $x \in E$. Quelle est la limite des probabilités

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

quand $n \to \infty$.

4. Calculer le temps moyen de trajet entre 1 et 3

$$\mathbb{E}_1(T_3).$$

5. Répéter les questions 1-3 avec la matrice de transition

$$Q' = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Exercice 88 (Théorie du renouvellement)**.** Un homme possède un stock de bois aléatoire. Il a au jour 0 un nombre X_0 de bûches, et chaque jour il consomme une bûche. On appelle X_n le nombre de bûches à la fin du n-ème jour. Si à un jour n il utilise sa dernière bûche, il renouvelle son stock le lendemain en allant ramasser des bûches, il en ramène un nombre aléatoire $X_n \geq 1$ tel que, pour $y \geq 1$,

$$\mathbb{P}(X_n = y) = f(y),$$

ou

$$f : \mathbb{N}^* \to \mathbb{R}^+$$

satisfait $\sum_{y \geq 1} f(y) = 1$. Le jour du ramassage, il consomme également immédiatement une bûche.

(Ce modèle s'appelle modèle de renouvellement, et peut s'appliquer à tout stock qui se renouvelle par un nombre aléatoire dont la loi ne varie pas.)

1. Montrer que (X_n) est une chaîne de Markov avec pour matrice de transition

$$Q(x, y) = \begin{cases} 1 \text{ si } y = x - 1 \geq 0, \\ f(y+1) \text{ si } x = 0, \\ 0 \text{ autrement.} \end{cases}$$

2. Montrer que (X_n) est une chaîne de Markov irréductible.
3. Montrer que (X_n) est une chaîne de Markov récurrente.
4. On appelle

$$Z = \{y \in \mathbb{N} : f(y) \neq 0\}.$$

Donner une condition simple sur Z pour que la chaîne soit apériodique.

5. On définit la mesure sur $\mathbb{N}$

$$\lambda(x) = \sum_{y=x+1}^{\infty} f(y), x \in \mathbb{N}.$$

Montrer que λ est invariante.

6. On suppose que la condition de la question 4. est vérifiée. On pose

$$m = \sum_{n \geq 0} n f(n).$$

Montrer que si $m < \infty$, on peut définir à partir de λ une distribution invariante π. Existe-t-il dans ce cas d'autres distributions invariantes ?

7. On pose

$$u(n) = \mathbb{P}_0(X_n = 0).$$

Donner la limite de $u(n)$ quand $n \to \infty$. Que dire sur la limite de $u(n)$ lorsque $\mu = \infty$?

8. On considère désormais le problème de renouvellement de machines : Une machine est affectée à une fonction particulière, et elle est remplacée par une machine identique neuve lorsqu'elle casse. La durée de vie d'une machine suit la loi Y, d'espérance $\mu < \infty$. On s'intéresse à la probabilité $v(n)$ qu'une machine se cassé précisément à l'instant n. On note $f(n) = \mathbb{P}(Y = n)$.

a) On appelle Y_n l'âge de la machine en marche à l'instant $n \geq 0$. Montrer que Y_n est une chaîne de Markov et donner sa matrice de transition.

b) En définissant judicieusement $f(n)$, montrer que Y_n peut s'écrire de manière explicite en fonction de (X_n) définie au début de l'exercice.

c) En déduire la valeur de $\lim_n v(n)$.

Exercice 89 (Question de cours)**.** Rappeler ce qu'est une mesure réversible et montrer que toute mesure réversible est invariante.

Correction:

Correction

.

Exercice 1

Exercice 2 1. Si $X_n = 0$ à la fin du jour $n \geq 0$, alors pour passer à $X_n = y$ au jour $n+1$, il faut ramasser $y+1$ bûches (une bûche sera utilisée le jour même), ce qui a pour probabilité $f(y+1)$. On en déduit $Q(0, y) = f(y+1)$.

Les autres lignes de la matrice de transition ne posent pas de problème.

2. On appelle

$$Y = \max\{y : f(y+1) > 0\}.$$

Remarquons que la suite de l'exercice a peu d'intérêt si $Y < \infty$... Cela signifie qu'il est impossible de ramasser plus de $Y+1$ buches, autrement dit $X_n \leq Y$ pour tout $n \geq 0$, et la chaîne de Markov est à valeurs dans $E = \{0, 1, \ldots, Y\}$ ($E = \mathbb{N}$ si $Y = \infty$).

De tout état $y \in E$ on passe à l'état 0 avec probabilité 1 (en y jours). Pour $0 \leq y \leq Y$, on peut passer de l'état 0 à l'état Y avec une probabilité positive (si par exemple on passe par Y, on repassera par y à un moment).

Tous les états sont donc équivalents à 0, la chaîne est donc irréductible.

3. 0 est récurrent car on est sûr de passer une infinité de fois par 0. Comme la chaîne est irréductible, elle est récurrente.

4. Supposons qu'il existe $y \geq 1$ tel que y et $y+1$ soient dans E. Alors on peut revenir à 0 en $n_0 = y$ ou en $n_0 + 1 = y + 1$ coups, ce qui implique que la chaîne est apériodique.

5. On a, pour tout $z \in E$,

$$\begin{aligned}
(\lambda Q)(z) = \sum_x \lambda(x) Q(x, z) &= \lambda(0) Q(0, z) + \lambda(z+1) \\
&= 1.f(z+1) + \sum_{y \geq z+2} f(y) \\
&= \sum_{y \geq z+1} f(z) = \lambda(z).
\end{aligned}$$

Donc λ est invariante.

6. Calculons la masse de λ :

$$\begin{aligned}
\lambda(E) &= \sum_{x \geq 0} \lambda(x) \\
&= \sum_{x \geq 0} \sum_{y \geq x+1} f(y) \\
&= \sum_{x \geq 0, y \geq x+1} f(y) \\
&= \sum_{y \geq 1, 0 \leq x \leq y-1} f(y) \\
&= \sum_{y \geq 1} f(y) y = m.
\end{aligned}$$

Donc, si l'hypothèse est vérifiée, $m = \lambda(E) < \infty$, et

$$\pi(y) = \frac{1}{m} \lambda(y)$$

est une distribution invariante.

L'existence d'une distribution invariante garantit que la chaîne est récurrente positive, et il existe alors une unique distribution invariante.

7. La chaîne est irréductible récurrente positive apériodique avec pour distribution invariante π, d'après le théorème de convergence à l'équilibre, on a donc,

$$\mathbb{P}_0(X_n = 0) \to \pi(0) = \frac{1}{m}.$$

8. On met en parallèle les problèmes de remplacement de machines et de ramassage de buches de la manière suivante : On suppose qu'un certain monsieur A s'occupe du ramasse des bûches pendant que monsieur B s'occupe du remplacement des machines. Chaque fois que monsieur A arrive au bout de ses buches, la machine de monsieur B casse (mais monsieur A et B ne se connaissent pas).

Formellement, en appelant $Z_1, Z_2, ...$ des variables iid de loi

$$\mathbb{P}(Z_1 = k) = f(k), k \geq 1,$$

on appelle X_n le ramassage de buches ou le n-ème ramassage ramène Z_n buches et Y_n le renouvellement des machines ou Z_n est la durée de vie de la n-ème machine. On montre facilement que

$$X_n = 0 \text{ ssi } Y_n = 0,$$

et on en déduit d'après la question ? que

$$\mathbb{P}_0(Y_n = 0) \to \frac{1}{m}.$$

12 Sujet d'examen 2016

Exercice 1 Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov sur $E = \{1,2,3,4,5,6\}$ de matrice de transition
Correction:

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

1. Dessiner le graphe de la chaîne de Markov. Déterminer la ou les classes irréductibles de cette chaîne.
2. Quels sont les classes récurrentes ?
3. On suppose que $X_0 = 1$.
   (a) Quelle est la probabilité qu'on ne repasse jamais par 1 ?
   (b) Soit $n \geqslant 1$. Quelle est la probabilité de repasser pour la première fois en 1 à l'instant n ?

Correction:

1. Les classes irréductibles sont $\{1,2\}, \{3,4\}$, et $\{5,6\}$
2. Les classes récurrentes sont $\{1,2\}$ et $\{5,6\}$.
3. Comme la classe $\{1,2\}$ est récurrente, et que $X_0 = 1$, la chaîne ne ressortira jamais de cette classe avec probabilité 1, et tous les états de cette classe seront visités une infinité de fois. Donc la probabilité de ne jamais repasser par 1 est 0.
   Cette probabilité correspond à la probabilité du chemin $(1,2,2,\dots,2,1)$. On a donc
   $$\mathbb{P}_1((X_0,\dots,X_n) = (1,2,2,\dots,2,1)) = Q(1,2)Q(2,2)^{n-2}Q(2,1) = \frac{1}{2}\left(\frac{3}{4}\right)^{n-2}\frac{1}{4}$$

Exercice 2
On considère une variable aléatoire Y_0 telle que

$$\mathbb{P}(Y_0 = 2) = \frac{1}{3}$$
$$\mathbb{P}(Y_0 = -1) = \frac{2}{3}.$$

Soit $(Y_n)_{n\geqslant 1}$ une famille de variables aléatoires indépendantes et identiquement distribuées, de meme loi que Y_0. Soit $X_0 = 0, X_n = \sum_{k=1}^n Y_k$ pour $n \geqslant 1$. On admet que $(X_n)_{n\geqslant 0}$ est une chaîne de Markov.

1. Montrer que, pour $n \geqslant 0$, $\mathbb{P}_0(X_n = 0) = 0$ si n n'est pas un multiple de 3 (on pourra raisonner dans la classe de congruences modulo 3).

   Correction: Comme $Y_n \equiv 2$ modulo 3 pour tout $n \geqslant 1$, $X_1 = X_0+Y_0 \equiv 0+2 \equiv 2$ modulo 3, $X_2 = X_1+Y_1 \equiv 2+2 \equiv 1$ modulo 3, et $X_3 \equiv 1+2 \equiv 0$ modulo 3. Donc X_1 et X_2 ne peuvent etre égaux à 0. On montre par récurrence sur $k \geqslant 1$ que pour tout $k \geqslant 1$, $X_{3k} \equiv 0$ modulo 3, $X_{3k+1} \equiv 2$ modulo 3, $X_{3k+2} \equiv 1$ modulo 3.

2. Soit $k \geqslant 1$, et $n = 3k$.
   (a) Soit un n-uplet $c \in \mathbb{R}^n$ constitué uniquement de $+2$ et de -1 (par exemple : $c = (2,2,-1,2)$ ou $(2,-1,-1,-1,2)$...). Sous quelle condition la somme de tous les éléments de c est-elle égale à 0 ?

   Correction: Soit p le nombre de $+2$ (donc le nombre de -1 est $n-p$). Donc la somme de tous les éléments est $2p-(n-p)$, cette somme est nulle ssi $3p = n \Leftrightarrow p = k$.

(b) En déduire le nombre de chemins qui vont de 0 à 0 en n coups et ont une probabilité non-nulle.

Correction: Le nombre de chemins qui vont de 0 à 0 (en repassant éventuellement par 0) correspond au nom de n-uplets constitués de $+2$ et -1 dont la somme vaut 0, c'est-à-dire au nombre de n-uplets contiennent k fois 2 et de $n-k$ fois -1. Choisir un tel n-uplet équivaut à choisir le sous-ensem de $\{1,\dots,n\}$ de k éléments qui indique les positions des "+2". Le nombre de ces sous ensembles $C_k^{3k} = \frac{(3k)!}{k!2k!}$.

(c) Donner un équivalent quand k tend vers l'infini de

$$\mathbb{P}_0(X_{3k} = 0).$$

On rappelle la formule de Stirling : $k! \sim_{k\to\infty} k^k e^{-k}\sqrt{2\pi k}$.

Correction: La probabilité d'un chemin qui va de 0 à 0 en $n = 3k$ coups est la probabilité que k variable tombent sur $+2$, et $n-k = 2k$ tombent sur -1, c'est donc $(1/3)^k(2/3)^{2k}$

$$\mathbb{P}(X_{3k} = 0) = (\#\{\text{ chemins qui vont de 0 à 0}\})(1/3)^k(2/3)^{2k}$$

$$= \frac{2^{2k}C_k^{3k}}{3^{3k}} = \frac{2^{2k}(3k)!}{3^{3k}k!(2k)!} \sim \frac{2^{2k}(3k)^{3k}e^{-3k}\sqrt{6\pi k}}{3^{3k}k^k e^{-k}\sqrt{2\pi k}(2k)^{2k}e^{-2k}\sqrt{4\pi k}} \sim \frac{2^{2k}3^{3k}k^{3k}\sqrt{6\pi}}{3^{3k}2^{2k}k^{3k}\sqrt{2\pi\cdot 4\pi}\sqrt{k}} \sim$$

Comme

$$\sum_{r>0} \mathbb{P}_0(X_r = 0) = 0 + \sum_{k\geqslant 1} \mathbb{P}_0(X_{3k} = 0) = \infty,$$

la chaine est récurrente.

3. La chaîne de Markov est-elle récurrente ?

Exercice 3

On considère l'ensemble $G = \{1,2,3\} \times \{1,2,3\}$. On voit G comme une grille 3×3 et chaque élément de G, assimilé à une particule, est donc un couple de coordonnées : $(1,1), (3,1)$, etc...

On considère l'ensemble E de tous les couples de particules distinctes : $E = \{\{(1,1),(1,2)\},\{(1,1),(1,3)\},\dots\}$ On fera bien attention que les positions des deux particules sont distinctes, et la position n'a pas d'importance. Par exemple, $\{(1,3),(2,2)\} = \{(2,2),(1,3)\}$, et $\{(1,3),(1,3)\} \notin E$.

1. Montrer que E a 36 éléments.

Correction: Il y a 9 positions particules possibles, le nombre de couples de particules distinctes est donc $9*8 = 72$. Comme les particules ne sont pas ordonnés, on divise par 2, ce qui fait 36.

2. On considère la chaîne de Markov suivante à valeurs dans E. $X_0 = \{(1,1),(1,2)\}$. Etant donné $n \geqslant 1$, et $X_n \in E$, X_{n+1} est construite de la manière suivante :
— Soit u, v les deux particules de X_n. On choisit au hasard une particule de X_n, à pile-ou-face.
— On la change d'emplacement dans G, afin qu'elle ne tombe pas sur l'autre particule (il y a donc 7 possibilités pour le nouvel emplacement).
— X_{n+1} est constitué de la nouvelle particule, ainsi que de la particule de X_n qui n'a pas bougé.
Ainsi, on peut par exemple passer de $\{(1,1),(1,2)\}$ à $\{(1,1),(3,3)\}$, ou à $\{(2,1),(1,2)\}$, mais pas à $\{(2,1),(3,3)\}$, ou $\{(1,1),(1,1)\}$.

(a) Soit $u, v, u', v' \in G$, et $x = \{u,v\}, y = \{u',v'\} \in E$ deux couples de particules. Donner la valeur de $Q(x,y)$, où Q est la matrice de transition de cette chaîne de Markov, en fonction de u, v, u', v'. Q est-elle symétrique ?

Correction: Si u, u', v, v' sont 4 particules distinctes, $Q(x,y) = 0$ car on ne peut changer qu'une particule à la f
Si $u = u'$ et $v = v'$ (ou $u = v', v = u'$), $x = y$, $Q(x,y) = 0$ car la particule qui est bougée ne p
retourner à la meme place.
Si $u \neq u'$ et $v = v'$ (ou $u = v'$ et $u' \neq v$), il y a 1 chance sur 2 de choisir u comme particule à bou
dans x, puis 1 chance sur 7 de choisir u' comme nouvelle position de u. On a donc $Q(x,y) = \frac{1}{14}$.
On observe que $Q(x,y) = Q(y,x)$ dans tous les cas, la chaine est donc bien symétrique.

(b) Montrer que c'est une chaîne de Markov irréductible.

Correction: Soit deux couples $x = \{u, v\}, x' = \{u', v'\}$, avec $x \neq x'$. Si les 4 particules u, v, u', v' sont distinctes
a

$$Q(x, \{u, v'\}) = \frac{1}{14},\ Q(\{u, v'\}, \{u', v'\}) = \frac{1}{14},$$

ce qui fournit un chemin de probabilité non-nulle entre x et x'.
Si $u = u', v \neq v'$ (ou $u = v', u' \neq v$), on a $Q(x, x') = \frac{1}{14}$.
Dans tous les cas, il y a un chemin de probabilité $14^{-2} > 0$ entre x et x', quels que soient x et x'
chaine est donc irréductible.

(c) Montrer qu'elle est apériodique.

Correction: Soit $x = \{u, v\} \in E$, et $u', u'' \in G$ deux positions distinctes, et distinctes de u, v. On a

$$\mathbb{P}(x \to \{u', v\} \to x) = \frac{1}{14}\frac{1}{14} > 0,$$

et

$$\mathbb{P}(x \to \{u', v\} \to \{u'', v\}) = \frac{1}{14^3} > 0.$$

Les nombres 2 et 3 appartiennent donc à $\{n : Q^n(x,x) \neq 0\}$, le pgcd de cet ensemble est donc 1
donc la chaine est apériodique.

3. Montrer qu'il existe une unique distribution invariante, et la donner.
4. Quel est le temps moyen de retour à l'état initial ? Soit $x \in E$. La limite $\lim_{n\to\infty} \mathbb{P}_x(X_n = x)$ existe-t-elle ? Si oui, que vaut-elle ?

Correction: Comme l'espace d'états est fini et la chaine est irréductible, elle est nécessairement récurrente positive.
Par symétrie, on voit que la distribution uniforme $\pi(\{x\}) = \frac{1}{36}, x \in E$, est invariante. On peut aussi
le montrer en utilisant la symétrie de Q : Pour tout $x, y \in E$,

$$\pi(x)Q(x,y) = Q(y,x)\pi(x) = Q(y,x)\pi(y),$$

donc π est réversible, donc invariante.
Le temps de retour T_{x_0} en un état $x_0 \in E$ vérifie donc

$$\mathbb{E}T_{x_0} = \frac{1}{\pi(x_0)} = 36.$$

Comme la chaine est IRPA,

$$\lim_n \mathbb{P}_x(X_n = x) \to \pi(x) = \frac{1}{36}.$$

5. On veut simuler approximativement une variable aléatoire U à valeurs dans E, de probabilité,

pour $u, v \in G, u \neq v$,

$$\mathbb{P}(U = \{u, v\}) = \frac{1}{Z} \exp(-\|u - v\|),$$

où Z est une constante appropriée, et $\|u - v\|$ est la distance Euclidienne entre u et v dans $\mathbb{R}^2$. Que proposez-vous ?

Correction: On utilise l'algorithme de Metropolis avec la matrice Q.

13 Rattrapage 2016

Exercice 1

Un joueur joue à pile ou face avec une pièce qui a 3 fois plus de chances de tomber sur "pile" que sur "face". Le joueur joue toujours "face". Pour miser, il procède de la façon suivante.

— A chaque coup, il met en jeu 1 jeton.
— S'il gagne, il remporte 3 jetons en plus de sa mise.
— S'il perd, il perd le jeton qu'il avait mis en jeu.
— Dès qu'il a 5 jetons ou plus devant lui, il met 5 jetons dans sa poche.
— Il s'arrête de jouer quand il n'a plus de jetons devant lui.

Pour $n \in \mathbb{N}$, on appelle X_n le nombre de jetons que le joueur a en face de lui après le n-ème coup.

1. Pourquoi $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov ?
2. Dresser le graphe de la chaîne de Markov, en indiquant bien la probabilité de transition entre chaque état (ou, de manière équivalente, la matrice de transition).
3. Donner les classes transientes et récurrentes.
4. On suppose qu'il démarre avec un jeton devant lui, et 0 en poche.
   (a) Donner les équations qui permettent de déterminer combien de coups le joueurs jouera en moyenne avant d'arrêter. (*On ne demande pas de résoudre ces équations !*).

Correction: Le temps moyen t_x^0 pour aller de x en 0, pour $x \neq 0$ vérifie $t_x^0 = 1 + \sum_{y\in E, y\neq x} Q(x,y)t_y^0$. Ca n
donne les équations suivantes (avec $t_0^0 = 0$),

$$
\begin{aligned}
t_1^0 &= 1 + \frac{1}{4}t_4^0 \\
t_4^0 &= 1 + \frac{3}{4}t_3^0 + \frac{1}{4}t_2^0 \\
t_3^0 &= 1 + \frac{1}{4}t_1^0 + \frac{3}{4}t_2^0 \\
t_2^0 &= 1 + \frac{3}{4}t_3^0.
\end{aligned}
$$

   (b) Quelle est la probabilité qu'il reparte avec 0 jetons en poche ?

Correction: On ne peut pas dénombrer tous les chemins qui mènent de 1 à 0 sans que le joueur n'empoche 5 ca
y en a une infinité :

— $1 \to 0$
— $1 \to 4 \to 3 \to 2 \to 1 \to 0$
— $1 \to 4 \to 3 \to 2 \to 1 \to 4 \to 3 \to 2 \to 1 \to 0$
— ...

Il faut introduite l'état P, où le joueur empoche 5 euros, ce qui définit une nouvelle CDM, a
$\mathbb{P}(2 \to P) = \mathbb{P}(3 \to P) = \mathbb{P}(4 \to P) = 1/4$. En appelant ensuite p_x^0 la probabilité d'atteindre 0
partant de x, on a

$$\begin{aligned}
p_1^0 &= \mathbb{P}(1 \to 0) + \mathbb{P}(1 \to 4)p_4^0 = \frac{3}{4} + \frac{1}{4}p_4^0 \\
p_2^0 &= \mathbb{P}(2 \to 1)p_1^0 + \mathbb{P}(2 \to P)p_P^0 = \frac{3}{4}p_1^0 + \frac{1}{4} \cdot 0 = \frac{3}{4}p_1^0 \\
p_3^0 &= \frac{3}{4}p_2^0 \\
p_4^0 &= \frac{3}{4}p_3^0.
\end{aligned}$$

En substituant on obtient

$$\begin{aligned}
p_1^0 &= \frac{3}{4} + \frac{1}{4}(\frac{3}{4}(\frac{3}{4}(\frac{3}{4}p_1^0))) = \frac{3}{4} + \frac{27}{4^4}p_1^0 \\
\frac{4^4 - 27}{4^4}p_1^0 &= \frac{3}{4} \\
p_1^0 &= \frac{3 \cdot 4^3}{4^4 - 27} = \frac{3 \cdot 4^3}{229}
\end{aligned}$$

Exercice 2

Lors d'une réception mondaine, il y a N invités, où $N \in \mathbb{N}^*$. Le lieu de la réception est constitué d'un salon et une terrasse. On fait les hypothèses suivantes :

— Régulièrement, un invité passe soit du salon à la terasse, soit de la terasse au salon. On appelle X_n le nombre d'invités dans le salon après n passages entre les deux, pour $n \in \mathbb{N}$.
— Si il y a un nombre $k \in \mathbb{N}$ de convives au salon, la probabilité que la prochaine personne à passer la porte du salon à la terrasse est k/N. Donc, la probabilité que la prochaine personne à passer le fasse de la terrasse au salon est $1 - k/N$.
— Aucun invité ne part ou n'arrive à la réception.

1. Donner un schéma simplifié du graphe de la chaîne de Markov $(X_n)_{n \geqslant 0}$.
2. Cette chaîne de Markov est-elle irréductible ? Récurrente ? Récurrente positive ?
3. Cette chaîne de Markov est-elle apériodique ?
4. On suppose qu'il existe une mesure réversible μ. Donner les relations que doit vérifier μ.
5. Existe-t-il une mesure invariante ? Si oui donner son expression.
6. On suppose que tous les invités sont dans le salon au début de la réception. On appelle $T > 0$ le nombre de passages qu'il faudra avant que tous les invités ne se retrouvent dans le salon à nouveau. Quelle est l'espérance de T ?

Exercice 3

On appelle E l'ensemble des mots que l'on peut former avec les lettres de l'alphabet $A, B, C, D, \ldots$. Par exemple, les éléments "ARBRE,Z,PPPPP,..." sont des mots, même si certains n'ont aucune signification.

Un bébé joue à former des mots avec des cartes représentant les lettres de l'alphabet. A chaque coup :

— Il enlève la dernière carte avec une probabilité $1/3$
— Il rajoute une carte à la fin avec une probabilité $1/3$
— Il change la dernière carte avec une probabilité $1/3$. Dans ce cas, il remet cette dernière dans le paquet, en pioche une nouvelle, et la met à la fin du mot qu'il est en train de former.
— Quand il pioche une nouvelle carte, la lettre représentée sur cette carte a autant de chances d'être un A, qu'un B, un C, etc...Toutes les lettres sont équiprobables.

Soit $n \geqslant 1$. On appelle $X_n \in E$ le mot formé après n modifications.

1. Montrer que la chaîne (X_n) est irréductible.
2. Montrer que la chaîne (X_n) est récurrente.
3. Montrer que la chaîne (X_n) est apériodique.
4. Soit deux mots m_1 et m_2. On suppose que m_1 a une lettre de plus que m_2. Calculer
$$\mathbb{P}(X_1 = m_1 | X_0 = m_2) \text{ et } \mathbb{P}(X_1 = m_2 | X_0 = m_1).$$
5. On suppose que le bébé modifie son comportement de la manière suivante : La probabilité d'ajouter une lettre est désormais de $1/2$, et la probabilité de retirer la dernière lettre est désormais de $1/52$. Montrer que la nouvelle chaîne de Markov correspondante à ce comportement est symétrique.
6. Le but du bébé est d'écrire un mot qui ait l'air français. Il ne sait pas exactement quels mots sont français, mais il peut instinctivement affecter une probabilité $p \in [0, 1]$ que le mot qu'il est en train de former est français. Si ce mot ne lui convient pas, il peut revenir au mot qu'il avait avant de faire la modification.

   En utilisant l'algorithme de Métropolis, donner une manière que la loi de X_n converge vers une mesure μ qui soit proportionnelle a la mesure $m \mapsto p(m), m \in E$. On suppose que
$$\sum_{m \in E} p(m) < \infty.$$

14 Examen 2017

Université Paris Descartes
UFR de Mathématiques et Informatique
45, rue des Saints-Pères 75270 Paris cedex 06

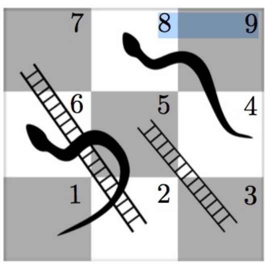
Chaînes de Markov
Examen

3 mai 2017

Nombre de pages de l'énoncé : 2 . Durée 2h

L'utilisation de tout appareil électronique, même à titre d'horloge, est interdit.

Exercice 90. On considère un jeu d'échelles et serpents qui se joue sur ce plateau :



Les règles sont les suivantes (à 1 joueur) :

— A chaque tour on lance un dé équilibré à 6 faces, et on avance du nombre de cases correspondantes, en s'arrêtant si on arrive sur la case 9.

— Si on arrive en bas d'une échelle, on monte instantanément en haut

— Si on arrive en haut d'un serpent, on descent instantanément en bas de sa queue.

— Dans les autres case de figure, on reste sur la case.

On considère la chaîne de Markov qui donne le position d'un joueur à la fin du tour.

1. Quels sont les états qui vont être effectivement visités par la chaîne de Markov ? Dessiner le graphe de la chaîne de Markov. *Correction:* En réalité, seuls les états 1,4,5,7,9 sont utiles

2. Donner la matrice de transition.

   Correction:

$$\begin{pmatrix} 1 \\ 4 \\ 5 \\ 7 \\ 9 \end{pmatrix} \begin{pmatrix} 1/6 & 1/6 & 1/3 & 1/3 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/3 \\ 1/6 & 1/6 & 0 & 1/6 & 1/2 \\ 0 & 1/6 & 0 & 0 & 5/6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

3. Donner les classes, et leurs natures (récurrentes, transientes). *Correction:* La classe $\{1, 4, 5, 7\}$ est transiente, la classe $\{9\}$ récurrente.

4. Donner un système linéaire d'équations qui permet de déterminer le temps moyen requis pour terminer le jeu (on ne demande pas de le résoudre) ?

Correction: On appelle k_i^9 le temps moyen pour atteindre 9 en partant de i. On a d'après le cours, avec $k_9^9 = 0$,

$$\begin{aligned}
k_1^9 =& 1 + \frac{1}{6}k_1^9 + \frac{1}{6}k_4^9 + \frac{1}{3}k_5^9 + \frac{1}{3}k_7^9\\
k_4^9 =& 1 + \frac{1}{6}k_1^9 + \frac{1}{6}k_4^9 + \frac{1}{6}k_4^9 + \frac{1}{6}k_7^9\\
k_5^9 =& 1 + \frac{1}{6}k_1^9 + \frac{1}{6}k_4^9 + \frac{1}{6}k_7^9\\
k_7^9 =& 1 + \frac{1}{6}k_4^9.
\end{aligned}$$

5. Donner un système linéaire pour déterminer la probabilité qu'un joueur ayant atteint la case 5 termine le jeu sans repasser par la case 1 (sans le résoudre) ? (On pourra considérer la case 1 comme un état absorbant).

   Correction: On considère désormais 1 comme absorbant (la probabilité de rester en 1 est de 1). On appelle h_i^9 la probabilité d'aller à 9 en un nombre fini d'étapes en partant de i. On a d'après le cours, avec $h_9^9 = 1, h_1^9 = 0$

$$\begin{aligned}
h_4^9 =& \frac{1}{6}h_4^9 + \frac{1}{6}h_4^9 + \frac{1}{6}h_7^9 + \frac{1}{3}\\
h_5^9 =& \frac{1}{6}h_4^9 + \frac{1}{6}h_7^9 + \frac{1}{2}\\
h_7^9 =& \frac{1}{6}h_4^9 + \frac{5}{6}.
\end{aligned}$$

Exercice 91. On considère un championnat sportif à 4 équipes, qui se prolonge indéfiniment. A chaque journée, l'équipe 1 affronte une équipe tirée au hasard parmi les 3 autres, et les 2 équipes restantes s'affrontent. Une victoire rapporte 1 points, 1 défaite -1 points. Chaque équipe à 50% de chances de gagner un match. On note $E = \mathbb{Z}^4$. On assimile un état $x = (x_1, x_2, x_3, x_4) \in E$ à la situation du championnat où pour $i = 1, \dots, 4$, l'équipe numéro i a x_i points. On note $Q(x, y)$ la matrice de transition de la chaîne de Markov correspondante.

1. Soit $x, y \in E$. Montrer que si $Q(x, y) > 0$, alors $\sum_{i=1}^4 x_i = \sum_{i=1}^4 y_i$. Déduisez-en que cette chaîne de Markov n'est pas irréductible.

   Correction: On appelle δ_i la quantité de laquelle varie le score de x_i après une journée de championnat. Comme deux équipes gagnent et deux équipes perdent, on a 2 δ_i égaux à 1, et 2 égaux à -1. Donc $\sum_{i=1}^4 \delta_i = 0$. Si $Q(x, y) > 0$, alors il est possible de passer de x à y en une journée de championnat, donc

$$\sum_{i=1}^4 y_i = \sum_{i=1}^4 (x_i + \delta_i) = \sum_{i=1}^4 x_i + \sum_{i=1}^4 \delta_i = \sum_{i=1}^4 x_i.$$

   Si $x = (0, 0, 0, 0)$ et $y = (1, 0, 0, 0)$ par exemple, $\sum_{i=1}^4 x_i = 0 \neq 1 = \sum_{i=1}^4 y$ donc x et y ne sont pas dans la même classe d'équivalence. Il y a donc plusieurs classes de récurrence et la chaîne de Markov n'est pas irréductible.

2. On considère la chaîne de Markov $(X_n)_{n \geqslant 0}$ de matrice de transition Q et telle que $X_0 = (0, 0, 0, 0)$. Donner sa période (la période est de 1 si la chaîne est apériodique).

   Correction: Si on regarde uniquement le score de la première équipe, il ne peut pas revenir à 0 en un nombre impair de matchs. Donc l'ensemble

$$A = \{n \geqslant 0 : Q^n(0, 0) > 0\}$$

ne contient que des nombres pairs. Comme il contient 2, son pgcd est 2, et la chaîne de Markov est donc 2-périodique.

3. On considère les score a_n et b_n des équipes 1 et 2 après n matchs, pour $n \in \mathbb{N}$, et $Y_n = (a_n, b_n)$, en supposant $a_0 = b_0 = 0$. On admet que $(Y_n)_{n \geqslant 0}$ est une chaîne de Markov homogène dans $\mathbb{Z}^2$. Donner sa matrice de transition Q'. $(Y_n)_{n \geqslant 0}$ est-elle irréductible sur $\mathbb{Z}^2$? Récurrente ?

   Correction: Soit $(a, b) \in \mathbb{Z}^2$. On a
$$Q'((a,b),(a+1,b+1)) = Q'((a,b),(a-1,b-1)) = \frac{1}{4}.$$
   Comme la somme de toutes les possibilités fait 1, on a
$$Q'((a,b),(a+1,b-1)) = Q'((a,b),(a-1,b+1)) = \frac{1}{2}(1-\frac{1}{4}) = \frac{1}{4}.$$
   Les probabilités de transitions vers les autres états sont nulles. Y_n est la marche aléatoire symétrique sur $\mathbb{Z}^2$, elle est donc irréductible récurrente (théorème de Polya).

4. Montrer que toute mesure uniforme sur $\mathbb{Z}$ est invariante pour la chaîne de Markov $(Y_n)_{n \geqslant 0}$. $(Y_n)_{n \geqslant 0}$ est-elle récurrente positive ?

   Correction:

   Soit μ qui associe $c \in \mathbb{R}$ à tout état (a, b) de $\mathbb{Z}^2$. Cette mesure est invariante car pour $(a, b) \in \mathbb{Z}^2$
$$\begin{aligned}\mathbb{P}_\mu(Y_1 = (a,b)) &= \sum_{(a',b') \in \mathbb{Z}^2} c\mathbb{P}(Y_1 = (a,b) | Y_0 = (a',b')) \\ &= c(\frac{Q((a-1,b-1)}{4}, (a,b)) + \frac{Q((a+1,b+1),(a,b))}{4} \\ &+ \frac{Q((a-1,b+1),(a,b))}{4} + \frac{Q((a+1,b-1),(a,b))}{4}) = c.\end{aligned}$$
   Donc μ est invariante. Comme la chaîne de Markov est récurrente, elle admet au plus une mesure invariante à une constante près. Si la chaîne était récurrente positive, elle admettrait une probabilité invariante π, qui serait proportionelle à μ, ce qui est impossible car la masse totale de π doit être de 1. Donc elle n'est pas récurrente positive.

5. On considère désormais qu'à la première journée, les matchs sont 1 contre 2, et 3 contre 4. A la seconde journée, 1 contre 3 et 2 contre 4, à la 3ème journée, 1 contre 4 et 2 contre 3, puis ça recommence indéfiniment. Qu'est-ce qui change sur les questions précédente ? (réponse en 5 lignes maximum)

   Correction: Dans cette organisation, on n'a plus affaire à une chaîne de Markov homogène, car la matrice de transition dépend de la classe de congruence de n modulo 3.

Exercice 92. Un biologiste cherche à modéliser un brin d'ADN, c'est-à-dire une séquence de n nucléotides, où $n \geqslant 1$ est fixé dans cet exercice, et une nucléotide est un élément de l'ensemble $N = \{A, C, G, T\}$. Un brin d'ADN est donc une séquence $(x_1, \ldots, x_n)$, où $x_i \in N$ pour $1 \leqslant i \leqslant n$.

Certains couples sont plus fréquents que d'autres dans une séquence d'ADN. Ainsi, il définit les probabilités $p_{A \to C}$ qu'une nucléotide C soit à droite d'une nucléotide A, , $p_{A \to G}, p_{A \to T}, \ldots$ dans $[0, 1]$ telles que
$$\sum_{a,b \in N : a \neq b} p_{a \to b} = 1.$$

Pour simplifier, on considère que la probabilité d'observer une séquence $x = (x_1, \ldots, x_n) \in N^n$ est simplement le produit des probabilités couple par couple :
$$\pi(x) = \frac{1}{Z} \prod_{i=1}^{n-1} p_{x_i \to x_{i+1}},$$

où Z est un nombre strictement positif qu'on ne demande pas de déterminer.

On définit l'état initial $x_0 = (A, A, A, \dots, A)$. On introduit la chaîne de Markov $(X_n)_{n\geqslant 0}$ à valeurs dans l'espace d'états $E = N^n$ d'état initial x_0, et telle que à chaque itération, une nucléotide est tirée au hasard dans $\{1, \dots, n\}$, et sa valeur est tirée au hasard dans N. On appelle Q la matrice de transition de cette chaîne de Markov .

1. Montrer que cette chaîne de Markov est irréductible et symmétrique.

   Correction: Irréducibilité : Il suffit de montrer que de x_0 on peut aller à n'importe quel état $x \in E$, et vice-versa. Partant de $x_0 = (A, A, \dots, A)$, il suffit de considérer le chemin qui au temps 1 change la 1-ère nucléotide en x_1, au temps 2 change le 2-ème nucléotide en x_2, etc...On a ainsi un chemin de probabilité strictement positive qui amène de x_0 à x. Le chemin inverse (qui change chaque nucléotide en A), ramène de x à x_0. Ainsi, x_0 communique avec tous les états, et la chaîne de Markov est irréductible.

   Symmétrique : Facile

2. Montrer qu'elle est apériodique. *Correction:* Il est possible à chaque transition de rester sur place : $(A, A, A, \dots) \to (A, A, A, \dots)$. Comme le pgcd d'un ensemble qui contient 1 est 1, la chaîne est apériodique.

3. On souhaite simuler une séquence d'ADN aléatoire de distribution π à l'aide de l'algorithme de Metropolis. Décrire l'algorithme à programmer. On décrira précisément comment tirer le nouveau brin d'ADN à chaque itération, et on donnera une expression simple du ratio d'acceptation en fonction des probabilités $p_{A\to C}, p_{A\to G}, \dots$.

   Correction: La matrice de transition Q correspondant à la chaîne de Markov décrite précédemment est symmétrique, récurrente, apériodique. On peut donc l'utiliser pour l'algorithme de Metropolis de la manière suivante :

   Pour chaque $n \geqslant 0$ (en commençant à $n = 0$) :

   (a) On part de l'état $X_n = (x_1, \dots, x_n)$.

   (b) On tire Y_{n+1} selon la loi $Q(X_n, \cdot)$, c'est-à-dire que Y_{n+1} est une variable aléatoire de loi
   $$\mathbb{P}(Y_{n+1} = x) = Q(X_n, x),$$
   la valeur de X_n étant déjà fixée. En pratique, on tire au hasard un nombre I uniformément dans $\{1, \dots, n\}$, et une nucléotide aléatoire N_n uniformément dans N. On pose ensuite $Y_{n+1} = (y_1, \dots, y_n)$ avec
   $$\begin{cases} y_i = x_i \text{ si } i \neq I \\ y_i = N_n \text{ sinon.} \end{cases}$$

   (c) On calcule le ratio $r_n = \frac{\pi(Y_{n+1})}{\pi(X_n)}$. On a
   $$r_n = \prod_{i=1}^{n-1} \frac{p_{y_i \to y_{i+1}}}{p_{x_{i+1}\to x_i}} = \frac{p_{x_{I-1}\to N_n}}{p_{x_{I-1}\to x_I}} \frac{p_{N_n \to x_{I+1}}}{p_{x_I \to x_{I+1}}}$$

   (d) On tire une variable aléatoire U_n uniformément dans $[0, 1]$.

   (e) Si $r_n \geqslant U_n$, on pose $X_{n+1} = Y_n$. Sinon on pose $X_{n+1} = X_n$.

   (f) On recommence à l'étape (a) avec $n + 1$.

4. Le biologiste s'intéresse à la probabilité que la séquence $AAGACTTTGACCACG$ apparaîsse dans un brin d'ADN aléatoire de distribution π, car elle correspond à une molécule importante. Comment estimer cette probabilité avec la chaîne $(X_n)_{n\geqslant 0}$ construite à la question précédente ?

Correction: La chaîne de Markov $(X_n)_{n\geqslant 0}$ est irréductible récurrente finie, donc récurrente positive, et admet comme distribution invariante π. On a donc d'après le théorème ergodique pour tout état $x \in E$,

$$\frac{1}{n} V_x(n) \to \pi(x),$$

où $V_x(n)$ est le nombre de visites en x de la chaîne (X_n) entre les temps 0 et n. En appelant A le sous-ensemble de E qui contient toutes les configurations d'ADN faisant apparaître la séquence d'intêret (une ou plusieurs fois), on a la convergence presque sûre

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_n \text{ contient la séquence}\}} = \frac{1}{n}\sum_{x\in A} V_x(n) \to \sum_{x\in A}\pi(x) = \pi(A)$$

qui est bien la quantité recherchée.

15 Examen 2018

Chaînes de Markov
Examen

Nombre de pages de l'énoncé : 2 . Durée 2h

L'utilisation de tout appareil électronique, même à titre d'horloge, est interdit.

Justifiez vos réponses ! Il sera tenu compte de la présentation.

Ecrivez clairement le numéro de l'exercice, et de la question traitée.

Question de cours

On considère un espace d'états dénombrable E, sur lequel on a une matrice stochastique symmétrique $P(x,y), x,y \in E$, telle qu'une chaîne de Markov X dont la matrice de transition est P soit apériodique et irréductible. Soit π une distribution de probabilité sur E. On souhaite simuler approximativement une variable aléatoire Z de loi μ. Indiquez les étapes de l'algorithme de Metropolis.

Exercice 1

Soit $E = \{1,2,3,4,5,6\}$. Soit $a,b,c,d,e,f,g,h,i,j,k,l,m,n$ des nombres de $]0,1[$. Soit $\alpha \in [0,1]$. On considère la matrice Q suivante :

$$\begin{pmatrix} a & b & 0 & 0 & 0 & 0 \\ c & d & e & 0 & 0 & 0 \\ 0 & 0 & f & g & 0 & h \\ 0 & 0 & i & j & 0 & 0 \\ \alpha & 0 & 0 & 0 & k & l \\ 0 & 0 & 0 & 0 & m & n \end{pmatrix}$$

1. Donner les relations que doivent satisfaire les paramètre $\alpha, a, b, c..., m, n$ pour que Q soit la matrice de transition d'une chaîne de Markov. Dans la suite on suppose ces conditions vérifiées, et on appelle $X = (X_n)_{n\geqslant 0}$ une chaîne de Markov de matrice de transition Q.

   Correction:

$$\begin{aligned} a+b &= 1 \\ c+d+e &= 1 \\ f+g+h &= 1 \\ i+j &= 1 \\ \alpha+k+l &= 1 \\ m+n &= 1 \end{aligned}$$

2. Donner le graphe de la chaîne de Markov dans le cas où $\alpha = 0$ et dans le cas $\alpha > 0$.
3. Donner la décomposition de E en classes récurrentes et transientes dans le cas où $\alpha = 0$ et dans le cas où $\alpha > 0$.

   Correction: Si $\alpha = 0$, $\{1,2\}$ et $\{3,4\}$ forment des classes transientes, et $\{5,6\}$ une classe récurrente. Si $\alpha > 0$, il n'y a qu'une seule classe (récurrente).
4. On suppose $\alpha > 0$. X est-elle apériodique ? Admet-elle une distribution invariante π ?

   Correction: Comme (au moins) un état renvoie sur lui-même, elle est apériodique (pgcd($\{1,...\}$)=1). Comme chaîne de Markov récurrente sur un espace fini elle est récurrente positive et donc admet une distribution invariante π.

5. On suppose $\alpha = c = d = e = k = l = f = g = h = 1/3$, $a = b = i = j = m = n = 1/2$. Donner la valeur de $\mathbb{E}_1(T_1^{(1)})$.

Exercice 2

Soit E, F deux ensembles dénombrables, et f une fonction de E dans F. Soit $X = (X_n)_{n\geqslant 0}$ une chaîne de Markov homogène sur E de noyau de transition $Q(x,y), x, y \in E$. On pose $Y_n = f(X_n)$ et $Y = (Y_n)_{n\geqslant 0}$.

1. On suppose que f est bijective. Montrer que Y est une chaîne de Markov.

   Correction: Soit $y_0, \ldots, y_{n+1} \in F$. On a

$$\begin{aligned}\mathbb{P}(Y_{n+1} = y_{n+1} | Y_i = y_i, 0 \leqslant i \leqslant n) =&\mathbb{P}(X_{n+1} = f^{-1}(y_{n+1}) | X_i = f^{-1}(y_i), 0 \leqslant i \leqslant n)\\ =&\mathbb{P}(X_{n+1} = f^{-1}(y_{n+1}) | X_n = f^{-1}(y_n)) = Q(f^{-1}(y_n), f^{-1}(y_{n+1})).\end{aligned}$$

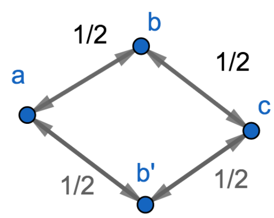
   donc Y est une chaîne de Markov .

2. On suppose que f est injective. Montrer que Y est une chaîne de Markov et donner son espace d'états et son noyau de transition.

   Correction: Il suffit de se restreindre à $F' = f(E)$ car f est bijective de E vers F'. D'après le calcul précédent, Y est une chaîne de Markov sur F' de noyau de transition

$$Q_Y(x,y) = Q(f^{-1}(x), f^{-1}(y)), x, y \in F'.$$

3. On suppose que $E = \{a, b, b', c\}, F = \{\alpha, \beta, \gamma\}$. On définit f par $f(b) = f(b') = \beta, f(a) = \alpha, f(c) = \gamma$. On suppose que X a le graphe suivant :



   Montrer que Y est une chaîne de Markov en justifiant bien et donner sa matrice de transition.

   Correction: Soit $y_0, \ldots, y_{n+1} \in F$. Si X_n est en 3 ou 4, X_{n+1} est forcément en 2 ou 1. Ca se traduit pour Y par

$$\mathbb{P}(Y_{n+1} = y_{n+1} | Y_n = y_n, ..., Y_0 = y_0) = \begin{cases} 1 \text{ si } y_{n+1} = 1 \text{ et } y_n \in \{2,3\} \\ 0 \text{ si } y_{n+1} \in \{2,3\} \text{ et } y_n \in \{2,3\}. \end{cases}$$

   Il reste donc à traiter le cas $y_n = 1, y_{n+1} \in \{2,3\}$.

$$\mathbb{P}(Y_{n+1} = 2 | Y_n = 1, Y_i = y_i) = \mathbb{P}(X_{n+1} = 3 | X_n = 2, X_i \in f^{-1}(y_i))p + \mathbb{P}(X_{n+1} = 3 | X_n = 1, X_i \in f^{-1}(y_i))(1-p)$$

   où

$$p = \mathbb{P}(X_n = 2 | Y_n = 1, Y_i = y_i).$$

   On a donc

$$\mathbb{P}(Y_{n+1} = 2 | Y_i = y_i, i \leqslant n) = \frac{1}{2}p + \frac{1}{2}(1-p) = \frac{1}{2}$$

indépendamment du passé $y_0, \dots, y_{n-1}$. Montre la même chose pour $Y_{n+1} = 3$ et on a donc bien une chaîne de Markov dont la matrice de transition est

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

4. On suppose que $E = \{a, b, c, d, e, b'\}$ et la matrice de transition est de la forme

$$Q = \begin{pmatrix} 0 & \cdot & 0 & 0 & 0 & \cdot \\ \cdot & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & \cdot & 0 & 0 \\ 0 & 0 & \cdot & 0 & \cdot & 0 \\ 0 & 0 & 0 & \cdot & 0 & \cdot \\ \cdot & 0 & 0 & 0 & \cdot & 0 \end{pmatrix}$$

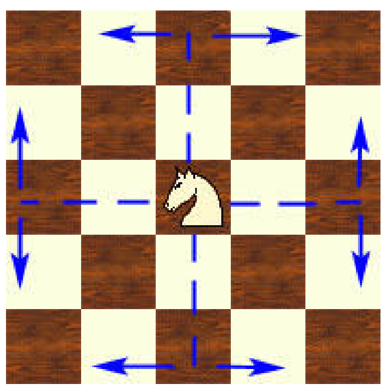
où les points représentent des nombres strictement positifs.

(a) Ecrire le graphe de X.

(b) On considère $F = \{\alpha, \beta, \gamma, \delta, \varepsilon\}$ et $f : E \to F$ définie par $f(a) = \alpha, f(b) = f(b') = \beta, f(c) = \gamma, f(d) = \delta, f(e) = \varepsilon$. Remplacez les points par des valeurs telles que Y ne soit pas une chaîne de Markov .

Exercice 3

Un échiquier comporte 8 cases sur 8 cases. On l'assimile à l'espace $E = \{1, 2, \dots, 8\} \times \{1, 2, ..., 8\}$. Pour se déplacer sur un échiquier, un cavalier peut faire deux pas dans une direction et un pas dans l'autre, comme indiqué sur la figure, si cette case se trouve bien sur l'échiquier.



1. Pour $x \in E$, on note $n(x)$ le nombre de cases où le cheval peut se déplacer depuis la case x, entre 1 et 8. Représentez l'échiquier sur votre copie et indiquez pour chaque case x le nombre $n(x)$.

Correction:

On considère la chaîne de Markov $X = (X_n)_{n \geqslant 0}$ sur E qui correspond à un déplacement aléatoire du cheval : la matrice de transition est

$$Q(x, y) = \frac{\mathbf{1}_{\{\text{le cavalier peut aller de } x \text{ à } y\}}}{n(x)}.$$

A chaque n, on décompose $X_n = (a_n, b_n)$ où $a_n \in \{1, 2, ..., 8\}$ et $b_n \in \{1, 2, .., 8\}$ représentent la ligne (en partant du haut) et la colonne (en partant de la droite) de X_n sur l'échiquier.

2. Montrer que pour tout $n \geqslant 0, x_n + y_n \equiv x_{n+1} + y_{n+1} + 1 \bmod 2$. Déduisez-en que la chaîne est périodique.

   Correction: On a soit

   — $x_{n+1} = x_n \pm 2$ et $y_{n+1} = y_n \pm 1$, auquel cas $x_{n+1} + y_{n+1} = x_n + y_n \pm 2 \pm 1 \equiv \pm 1 \equiv 1 \text{mod } 2$

   — Soit $x_{n+1} = x_n \pm 1$ et $y_{n+1} = y_n \pm 2$ auquel cas on a la même conclusion.

3. Montrer que pour deux cases voisines x et y dans E il existe un chemin probable qui va de x à y. Que peut-on en déduire pour les classes d'équivalence de X ?

   Correction: On peut trouver un chemin de longueur 3 qui va de x à y, et le chemin inverse va de y à x.

4. Montrer que $\mu(x) = n(x)$ est une mesure invariante.

   Correction: Pour tout $x \in E$,

$$\sum_{y \in E} Q(y,x)\mu(y) = \sum_{y \in E : x \sim y} \frac{1}{n(y)} n(y) = n(x)$$

5. On suppose que le cavalier commence à l'état $X_0 = x_0$, où $x_0 = (1,1)$. Donner l'espérance du temps de retour $\mathbb{E}_{x_0}(T_{x_0}^{(1)})$.

   Correction: On a une chaîne de Markov récurrente positive, donc une unique proba invariante donnée par

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

   Comme $n(x)$ est invariante, $\pi = n/n(E)$ et Donc

$$\mathbb{E}_{x_0}(T_{x_0}^{(1)}) = \frac{1}{\pi(x_0)} = \frac{n(E)}{n(x_0)} = \frac{n(E)}{4}.$$

   Il ne reste plus qu'à calculer

$$n(E) = \sum_{x \in E} n(x) = 4 \cdot 4 + ...$$