Marches aléatoires, chaînes de Markov

Exercice 1 Bruit qui court

Les notes de cours seront-elles autorisées à l'examen ? La réponse est transmise à travers n intermédiaires. On suppose que chaque étudiant transmet le message de façon correcte avec une probabilité p telle que $0 < p < 1$ ou le déforme en son contraire avec probabilité $1 - p$. Les intermédiaires sont indépendants.

▷ 1. Modéliser cette situation par une chaîne de Markov a deux états et calculer la probabilité que l'information transmise par le n-ième intermédiaire soit conforme à l'information initiale.

▷ 2. Que se passe-t-il quand $n \to \infty$?

Solution

▷ 1. On modélise ce système par une chaîne de Markov ou le premier état signifie « La dernière personne ayant reçu l'information a reçu une information non déformée », l'autre état signifiant qu'elle a l'information inverse. Chaque transmission d'un intermédiaire au suivant correspond à une transition dans la chaîne de Markov.
La matrice correspondant à cette chaîne de Markov est :

$$\underbrace{\begin{pmatrix} p & (1-p) \\ (1-p) & p \end{pmatrix}}_{P} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{Q} \begin{pmatrix} 1 & 0 \\ 0 & 2p-1 \end{pmatrix} \underbrace{\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}}_{Q^{-1}}$$

avec $QQ^{-1} = Id$.

$$P^n = Q \begin{pmatrix} 1 & 0 \\ 0 & 2p-1 \end{pmatrix}^n Q^{-1}$$

La probabilité que l'information soit conforme à l'information initiale est donc

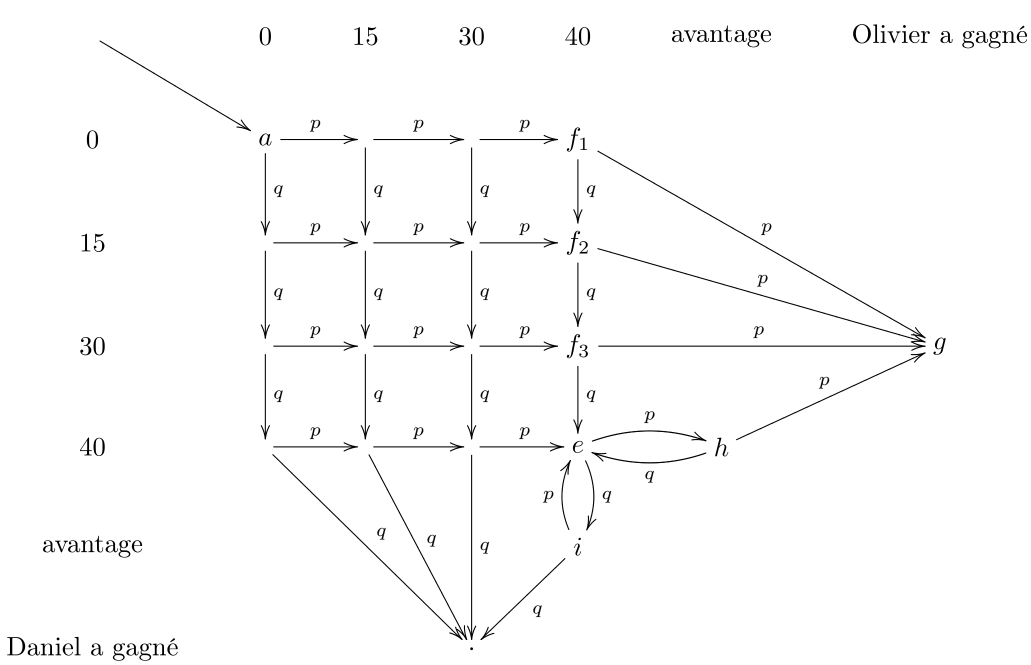
$$(1\ 0)P^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}(1 + (2p-1)^n)$$

▷ 2. Quand n tend vers ∞, la probabilité tends vers $\frac{1}{2}$.

Exercice 2 Un jeu de tennis

On considère un jeu classique du tennis (pas un jeu décisif) de Daniel contre Olivier. C'est à Olivier de servir. Il gagne chaque point avec probabilité p. Modéliser par une chaîne de Markov et donner la probabilité pour Olivier de gagner le jeu.

SolutionSoit $q := 1 - p$. Voici là chaîne de Markov de ce jeu de tennis :

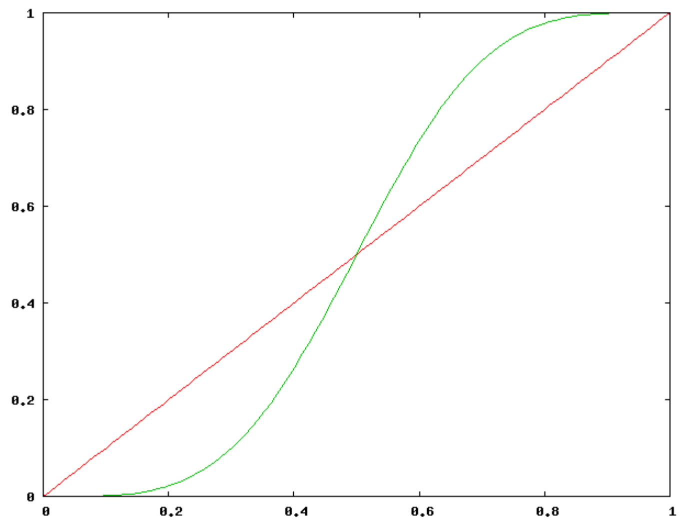


Correction par Jérémy Planul, relue par J.-B. Rouquier

La probabilité d'arriver en g est la somme sur tous les chemins menant à g de la probabilité de chaque chemin. On compte ces chemin selon le dernier état rencontré :
- s'il s'agit de f_1 la probabilité du chemin est p^4 ;
- pour f_2, il y a C_4^1 chemins possibles, chacun de probabilité $p^4 q$;
- pour f_3, il y a de même C_5^2 chemins, chacun de probabilité $p^4 q^2$;
- s'il s'agit de h, on décompose le chemin en trois étapes :
  - la portion $a...e$, il y a C_6^3 chemins, chacun de probabilité $p^3 q^3$,
  - un certain nombre de boucles ehe ou eie, la probabilité de faire n boucles est $2^n (pq)^n$,
  - la portion ehg, de probabilité p^2.

En sommant toutes les possibilité, la probabilité d'arriver en g est $p^4 + 4p^4 q + 10p^4 q^2 + 20p^3 q^3 \times \underbrace{\sum_{i=0}^{\infty} (2pq)^n}_{\frac{1}{1-2pq}} \times p^2$

Remarque. Voici en vert la probablité qu'Olivier gagne en fonction de p :



On a donc augmenté le contraste : le système de comptage des points avantage le plus fort joueur.

Exercice 3 D'un état à l'autre

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov à valeurs dans $\mathcal{S}$ de matrice de transition P. Pour tout état y, on note $T_y := \min\{n \mid X_n = y\}$ (avec la convention $\min \varnothing = +\infty$). Pour tout événement A on note $\mathbb{P}_x(A) := \mathbb{P}(A \mid X_0 = x)$.

Soient x et y deux états distincts, on pose $\rho_{xy} := \mathbb{P}_x(T_y < +\infty)$.

Les questions de cet exercices sont simples, le but est de manier les définitions en faisant une démonstration particulièrement détaillée.

▷ 1. Montrer que $\rho_{xy} > 0 \iff \exists n \geqslant 1, P^n(x,y) > 0$.

▷ 2. On suppose désormais $\rho_{xy} > 0$. Soit $n_0 := \min\{n \geqslant 1 \mid P^n(x,y) > 0\}$.
Soient $x_1, \ldots, x_{n_0-1}$ des états tels que $P(x,x_1)P(x_1,x_2)...P(x_{n_0-1},y) > 0$.
Montrer que les états $x =: x_0, x_1, \ldots, x_{n_0-1}, x_{n_0} := y$ sont distincts.

▷ 3. Supposons $\mathcal{S}$ fini de cardinal d. Montrer que $n_0 \leqslant d-1$. On suppose $\rho_{xy} > 0$, montrer que $\mathbb{P}_x(T_y \leqslant d-1) > 0$

Solution

▷ 1. Supposons $\rho_{xy} > 0$.

$$\begin{aligned}\rho_{xy} &= \mathbb{P}_x(T_y < \infty)\\ &= \mathbb{P}_x\left(\textstyle\bigcup_n T_y = n\right)\\ &= \textstyle\sum_n \mathbb{P}_x(T_y = n) \qquad \text{car les événements } T_y = n \text{ sont disjoints deux à deux.}\end{aligned}$$

Il existe donc n tel que $\mathbb{P}_x(T_y = n) > 0$. Or $P^n(x,y) = \mathbb{P}_x(X_n = y) \geqslant \mathbb{P}_x(T_y = n)$.

Réciproquement, supposons qu'il existe n tel que $P^n(x,y) > 0$.
Alors $\rho_{xy} = \mathbb{P}_x(T_y < \infty) \geqslant \mathbb{P}_x(T_y \leqslant n) \geqslant \mathbb{P}_x(X_n = y) > 0$.

Intuitivement, $\rho_{xy} > 0$ si et seulement si il existe un chemin de x à y.

▷ 2. Intuitivement, n_0 est la taille du plus court chemin de x à y, il ne peut donc avoir de boucle.
S'il existait $i, j \in [\![0,\, n_0]\!]$ tels que $x_i = x_j$ et $i < j$, on aurait alors $P(x, x_1)P(x_1, x_2)...P(x_i, x_j)...P(x_{n_0-1}, y) > 0$ et donc $P^{n_0-(j-i)}(x, y) > 0$, contradiction.

▷ 3. $x, x_1, \ldots, x_{n_0-1}, y$ sont n_0+1 états distincts donc $n_0+1 \leqslant d$. Donc $\mathbb{P}_x\,(T_y \leqslant d-1) \geqslant \mathbb{P}_x\,(T_y \leqslant n_0) \geqslant \mathbb{P}_x\,(T_y = n_0) > 0$.

Exercice 4 Walk the (real) line

Soit $\{X_k\}$ des variables aléatoires discrètes indépendantes et identiquement distribuées. On définit une marche aléatoire par $S_n := \sum_{k=1}^{n} X_k$. On s'intéresse à une marche aléatoire dans $\mathbb{R}$. La fonction de distribution des X_k est alors :

$$f_X(x) = \begin{cases} 1/2 & \text{si } x = \pm 1 \\ 0 & \text{sinon} \end{cases}$$

On s'intéresse à la probabilité d'un retour à l'origine (en un temps fini).

▷ 1. S'il y a eu un retour à l'origine au temps m, que peut-on dire de m ? Montrer qu'un retour à l'origine au temps $2m$ arrive avec une probabilité $u_{2m} = C_{2m}^{m} 2^{-2m}$.

▷ 2. On définit de même la probabilité f_{2m} qu'un premier retour à l'origine se fasse au temps $2m$. Montrer que les probabilités $\{f_{2k}\}$ et $\{u_{2k}\}$ vérifient la relation $u_{2n} = f_0 u_{2n} + f_2 u_{2n-2} + \cdots + f_{2n} u_0$ (on pose $u_0 := 1$ et $f_0 := 0$).

▷ 3. On définit les fonctions génératrices :

$$U(x) = \sum_{m=0}^{\infty} u_{2m} x^m \quad \text{et} \quad F(x) = \sum_{m=0}^{\infty} f_{2m} x^m$$

Déduire de la question précédente une relation simple entre $U(x)$ et $F(x)$.

▷ 4. On admet que $U(x) = \frac{1}{\sqrt{1-x}}$. En déduire que $F(x) = 1 - (1-x)^{1/2}$.

▷ 5. Montrer que $f_{2m} = \frac{C_{2m}^{m}}{(2m-1)2^{2m}}$. Indication : considérer F'.

▷ 6. Définissons w_n la probabilité qu'un retour à l'origine se fasse au plus tard au temps n. Notre but est de savoir si l'on va revenir en un temps fini, c'est-à-dire déterminer $w_* = \lim_{n\to\infty} w_n$. Montrer que $w_* = F(1)$. Conclure.

Correction : voir le poly ACM page 471.