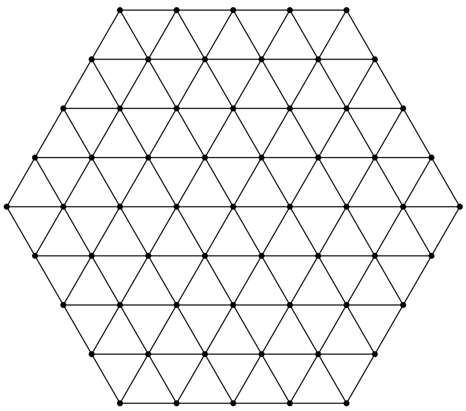
Probabilités de retour d'un marche aléatoire sur un graphe

mots-clés : marche aléatoire sur un graphe, martingales, temps d'arrêt.

Dans ce texte, on fixe un *graphe* fini $G = (V, E)$, $V = \{v_1, v_2, \ldots, v_m\}$ étant un ensemble fini de *sommets* et E un ensemble d'*arêtes*, c'est-à-dire une partie de $\{\{x, y\} \mid x, y \in V,\ x \neq y\}$. Les arêtes sont donc des paires de sommets distincts, et elles sont non orientées. Pour fixer les idées, on pourra garder en tête le graphe suivant, qui est une partie du réseau triangulaire :

$G = (V, E) =$ 

On supposera dans tout le texte que le graphe G est connexe et a au moins deux sommets ; en particulier, chaque sommet a au moins un voisin. Le *degré* d'un sommet $v \in V$ est son nombre de voisins :

$$\deg v = |\{w \in V \mid \{v, w\} \in E\}|.$$

Le degré du graphe est défini par $\deg G = \max\{\deg v \mid v \in V\}$. La *marche aléatoire* sur G est la chaîne de Markov d'espace d'états V, et de matrice de transition :

$$P(v, w) = \begin{cases} \frac{1}{\deg v} & \text{si } v \text{ et } w \text{ sont voisins,} \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, une transition de cette chaîne de Markov à partir d'un sommet $v \in V$ est réalisée en choisissant un sommet uniformément parmi les voisins de v, et en sautant vers ce voisin. Pour éviter des problèmes de périodicité, on étudiera plutôt la *marche aléatoire paresseuse*, qui est définie comme suit : à chaque instant de transition, soit l'on reste sur le même sommet v avec probabilité $\frac{1}{2}$, soit l'on saute vers l'un des voisins w de v avec probabilité $\frac{1}{2}$, ce voisin w étant alors choisi uniformément. La marche aléatoire paresseuse est donc la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice de transition $Q = \frac{I+P}{2}$, I étant la matrice identité. L'objectif de ce texte est d'étudier certaines propriétés de la marche paresseuse $(X_n)_{n\in\mathbb{N}}$. On s'intéressera en particulier aux *probabilités de retour* $Q^n(v, v) = \mathbb{P}[X_n = v \mid X_0 = v]$.

1 Mesure invariante et chaîne des ensembles évoluants

On montre facilement :

Théorème 1. *La marche aléatoire paresseuse est une chaîne de Markov irréductible apériodique sur V. Sa mesure de probabilité invariante π est donnée par :*

$$\pi(v) = \frac{\deg v}{2|E|},$$

où $|E|$ est le nombre total d'arêtes du graphe. Pour tous sommets $v, w \in V$, on a $\pi(w) = \lim_{n\to\infty} Q^n(v, w)$.

On cherche alors à estimer la différence $|Q^n(v, v) - \pi(v)|$ et à avoir une borne raisonnable sur cette quantité. Des informations importantes peuvent être obtenues en introduisant la *chaîne des ensembles évoluants*, qui est définie comme suit. On note $\mathfrak{P}(V)$ l'ensemble des parties de V, et si $S \in \mathfrak{P}(V)$ et $x \in [0, 1]$, on définit une nouvelle partie

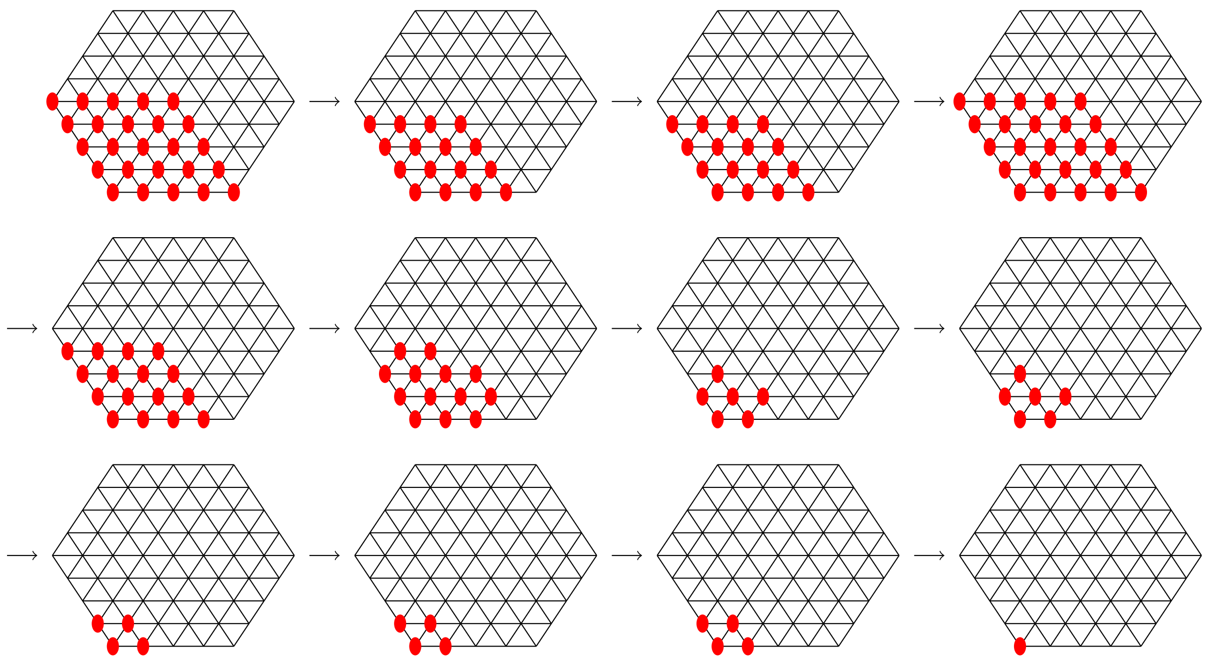
$$f(S, x) = \left\{ w \in V \mid \sum_{v \in S} \pi(v)\, Q(v, w) \geq x\, \pi(w) \right\}.$$

Définition 2. *Soit $(U_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes uniformes sur $[0, 1]$. Si $S_0 \in \mathfrak{P}(V)$, la chaîne des ensembles évoluants issue de S_0 est le processus $(S_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathfrak{P}(V)$ défini par la relation de récurrence*

$$S_{n+1} = f(S_n, U_n).$$

C'est une chaîne de Markov à valeurs dans $\mathfrak{P}(V)$, qui admet pour états absorbants $\emptyset$ et V.

Le graphique suivant montre une trajectoire de la chaîne des ensembles évoluants lorsque G est la partie du réseau triangulaire précédemment dessinée.



Théorème 3. *Soit $(S_n)_{n\in\mathbb{N}}$ la chaîne des ensembles évoluants issue d'une partie S_0. La suite des variables aléatoires $(\pi(S_n))_{n\in\mathbb{N}}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ engendrée par les ensembles aléatoires S_n.*

Démonstration. Il suffit de montrer que pour toute partie $S \subset V$, on a $\mathbb{E}[\pi(f(S,U))] = \pi(S)$. Pour tout $w \in V$, la probabilité pour que $w \in f(S,U)$ est la probabilité pour que

$$\frac{\sum_{v \in S} \pi(v)\, Q(v,w)}{\pi(w)} \geq U. \tag{1}$$

Notons que le terme de gauche dans cette inégalité est un réel entre 0 et 1. Par conséquent, la probabilité que l'on cherche est exactement le terme de gauche. On en déduit que :

$$\mathbb{E}[\pi(f(S,U))] = \sum_{w \in V} \mathbb{P}[w \in f(S,U)]\, \pi(w) = \sum_{\substack{v \in S \\ w \in V}} \pi(v)\, Q(v,w) = \sum_{v \in S} \pi(v) = \pi(S). \qquad \square$$

Corollaire 4. *Pour toute partie $S_0 \in \mathfrak{P}(V)$, la chaîne des ensembles évoluants $(S_n)_{n \in \mathbb{N}}$ issue de S_0 est avec probabilité 1 constante au bout d'un certain rang. Notons $S_\infty = \lim_{n \to \infty} S_n$ la limite presque sûre. La loi de la variable aléatoire S_∞ est*

$$\mathbb{P}[S_\infty = \emptyset] = 1 - \pi(S_0) \qquad ; \qquad \mathbb{P}[S_\infty = V] = \pi(S_0).$$

2 Temps d'atteinte des ensembles limites et probabilités de retour

Dans ce qui suit, $S_0 = \{v\}$ est un singleton fixé dans V, et on note comme précédemment $(S_n)_{n \in \mathbb{N}}$ la chaîne des ensembles évoluants issue de ce singleton. On définit

$$\tau_v = \inf\{n \in \mathbb{N} \mid S_n \in \{\emptyset, V\}\};$$

d'après ce qui précède, τ_v est un temps d'arrêt vis-à-vis de la filtration $\mathcal{F}_n$ et est fini presque sûrement. Le lien entre ce temps d'arrêt et la vitesse de convergence de $Q^n(v,v)$ vers $\pi(v)$ est donné par :

Proposition 5. *Pour tout $n \in \mathbb{N}$,*

$$|Q^n(v,v) - \pi(v)| \leq \mathbb{P}[\tau_v > n].$$

Démonstration. Notons que, pour tous sommets v et w, on a

$$Q^n(v,w) = \frac{\pi(w)}{\pi(v)}\, \mathbb{P}_{\{v\}}[w \in S_n],$$

où $(S_n)_{n \in \mathbb{N}}$ est la chaîne des ensembles évoluants issue de $\{v\}$, ce que l'on indique en utilisant la notation $\mathbb{P}_{\{v\}}$. En effet, cette identité est claire pour $n = 0$; supposons-la établie au rang $n \geq 0$, et calculons $\frac{\pi(w)}{\pi(v)}\, \mathbb{P}_{\{v\}}[w \in S_{n+1}]$. Conditionnellement à S_n, la probabilité pour qu'un élément w appartiennent à S_{n+1} a été calculée dans la preuve du théorème 3 :

$$\mathbb{P}[w \in S_{n+1} \,|\, S_n] = \frac{\sum_{z \in S_n} \pi(z)\, Q(z,w)}{\pi(w)}.$$

Alors,

$$\begin{aligned}
\frac{\pi(w)}{\pi(v)}\, \mathbb{P}_{\{v\}}[w \in S_{n+1}] &= \frac{\pi(w)}{\pi(v)}\, \mathbb{E}_{\{v\}}[\mathbb{P}[w \in S_{n+1} \,|\, S_n]] = \mathbb{E}_{\{v\}}\left[\sum_{z \in S_n} \frac{\pi(z)\, Q(z,w)}{\pi(v)}\right] \\
&= \sum_{z \in V} \mathbb{P}_{\{v\}}[z \in S_n]\, \frac{\pi(z)\, Q(z,w)}{\pi(v)} = \sum_{z \in V} Q^n(v,z)\, Q(z,w) = Q^{n+1}(v,w)
\end{aligned}$$

en utilisant l'hypothèse de récurrence. On en déduit que $Q^n(v,v) = \mathbb{P}_{\{v\}}[v \in S_n]$, et par ailleurs, $\pi(v) = \mathbb{P}_{\{v\}}[v \in S_\infty] = \mathbb{P}_{\{v\}}[v \in S_{\tau_v}]$. Par conséquent,

$$|Q^n(v,v) - \pi(v)| \leq |\mathbb{P}_{\{v\}}[v \in S_n] - \mathbb{P}_{\{v\}}[v \in S_{\tau_v}]| \leq \mathbb{P}[S_n \neq S_{\tau_v}] = \mathbb{P}[\tau_v > n]. \qquad \square$$

Lemme 6. *Si $S \notin \{\emptyset, V\}$, alors*

$$b(S) = \mathrm{var}_{(S_0=S)}(\pi(S_1)) = \mathbb{E}_S[(\pi(S_1))^2] - (\mathbb{E}_S[\pi(S_1)])^2 \geq \frac{1}{4|E|^2}.$$

Démonstration. Étant données deux variables aléatoires X et Y avec X de carré intégrable, on a toujours $\mathrm{var}(\mathbb{E}[X|Y]) \leq \mathrm{var}(X)$. Par conséquent,

$$b(S) = \mathrm{var}_S(\pi(S_1)) \geq \mathrm{var}_S\left(\mathbb{E}[\pi(S_1)\,|\,1_{U_0 \leq \frac{1}{2}}]\right).$$

— Si $w \in S$, alors $\frac{\sum_{v\in S}\pi(v)\,Q(v,w)}{\pi(w)} \geq Q(w,w) = \frac{1}{2}$, donc

$$\mathbb{P}\left[\sum_{v\in S}\pi(v)\,Q(v,w) \geq U_0\,\pi(w)\,|\,U_0 \leq \frac{1}{2}\right] = 1;$$
$$\mathbb{P}\left[\sum_{v\in S}\pi(v)\,Q(v,w) \geq U_0\,\pi(w)\,|\,U_0 \geq \frac{1}{2}\right] = 2\,\frac{\sum_{v\in S}\pi(v)\,Q(v,w)}{\pi(w)} - 1.$$

— Si $w \notin S$, alors $\frac{\sum_{v\in S}\pi(v)\,Q(v,w)}{\pi(w)} \leq \frac{1}{2}\,\frac{(\pi P)(w)}{\pi(w)} = \frac{1}{2}$, donc

$$\mathbb{P}\left[\sum_{v\in S}\pi(v)\,Q(v,w) \geq U_0\,\pi(w)\,|\,U_0 \leq \frac{1}{2}\right] = 2\,\frac{\sum_{v\in S}\pi(v)\,Q(v,w)}{\pi(w)};$$
$$\mathbb{P}\left[\sum_{v\in S}\pi(v)\,Q(v,w) \geq U_0\,\pi(w)\,|\,U_0 \geq \frac{1}{2}\right] = 0.$$

On en déduit la loi de l'espérance conditionnelle :

$$\mathbb{E}_S\left[\pi(S_1)\,|\,U_0 \leq \frac{1}{2}\right] = \pi(S) + 2\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w); \tag{2}$$
$$\mathbb{E}_S\left[\pi(S_1)\,|\,U_0 \geq \frac{1}{2}\right] = \pi(S) - 2\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w). \tag{3}$$

La variance de l'espérance conditionnelle est donc égale à $4\left(\sum_{v\in S,\,w\notin S}\pi(v)\,Q(v,w)\right)^2$. Or,

$$\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w) = \sum_{\substack{v\in S\\ w\notin S}}\frac{\deg v}{2|E|}\,\frac{1_{v\sim w}}{2\,\deg v} = \frac{1}{4|E|}\sum_{\substack{v\in S\\ w\notin S}}1_{v\sim w} \geq \frac{1}{4|E|};$$

en effet, puisque G est connexe, il existe au moins une arête $\{v,w\}$ avec $v \in S$ et $w \notin S$. $\square$

Théorème 7. *Pour tout $n \in \mathbb{N}$ et tout sommet $v \in V$ du graphe,*

$$|Q^n(v,v) - \pi(v)| \leq \frac{2\,(\deg G + 1)^{\frac{3}{2}}}{\sqrt{n}}.$$

Démonstration. On note $M_n = \pi(S_n)$, qui est une martingale positive bornée ; et $\sigma^2 = \frac{1}{4|E|^2}$. Si $h > 0$, introduisons les temps d'arrêt $T_{v,h} = \inf\{n \in \mathbb{N} \mid M_n = 0 \text{ ou } M_n \geq h\}$ et $\tau'_v = \inf\{n \in \mathbb{N} \mid M_n = 0\}$, $(S_n)_{n\in\mathbb{N}}$ étant issue de $\{v\}$. Comme $\tau_v \leq \tau'_v$, On a

$$\{\tau'_v > n\} \subset \{T_{v,h} > n\} \cup \{M_{T_{v,h}} \geq h\}$$
$$\mathbb{P}[\tau_v > n] = \mathbb{P}[\tau'_v > n \text{ et } \tau_v > n] \leq \mathbb{P}[\tau_v > n \text{ et } T_{v,h} > n] + \mathbb{P}[M_{T_{v,h}} \geq h].$$

Comme $(M_n)_{n\in\mathbb{N}}$ est une martingale bornée, l'inégalité de Markov et le théorème d'arrêt impliquent

$$\mathbb{P}[M_{T_{v,h}} \geq h] \leq \frac{\mathbb{E}[M_{T_{v,h}}]}{h} = \frac{M_0}{h} = \frac{\pi(v)}{h}.$$

Si $\tau_v > n$, alors par le lemme précédent $b(S_n) \geq \sigma^2$, donc

$$\mathbb{E}[(M_{n+1})^2\, 1_{\tau_v>n} | \mathcal{F}_n] = 1_{\tau_v>n}(b(S_n) + (M_n)^2) \geq 1_{\tau_v>n}(\sigma^2 + (M_n)^2).$$

Par conséquent, sur l'événement $\{\tau_v > n\}$, $G_n = (M_n)^2 - hM_n - n\sigma^2$ vérifie $\mathbb{E}[G_{n+1} \,|\, \mathcal{F}_n] \geq G_n$. Ceci revient à dire que $F_n = G_{n\wedge\tau_v}$ est une sous-martingale. Il en va de même pour $E_n = G_{n\wedge\tau_v\wedge T_{v,h}}$, et on peut alors écrire :

$$-h\pi(v) \leq (\pi(v))^2 - h\pi(v) = E_0 \leq \mathbb{E}[E_n] = \mathbb{E}[M_{n\wedge\tau_v\wedge T_{v,h}}(M_{n\wedge\tau_v\wedge T_{v,h}} - h)] - \mathbb{E}[(n\wedge\tau_v\wedge T_{v,h})\sigma^2].$$

Notons maintenant que pour toute partie $S \subset V$ et tout $x \in (0,1]$, on a l'inclusion

$$f(S,x) \subset S \sqcup \{w \in V \,|\, w \text{ a un voisin dans } S\}, \tag{4}$$

ce qui implique l'inégalité déterministe $\pi(f(S,x)) \leq (\deg G + 1)\,\pi(S)$, et donc presque sûrement $M_{n+1} \leq (\deg G + 1)\, M_n$. Il suit que $M_{n\wedge\tau_v\wedge T_{v,h}} \leq (\deg G + 1)h$ pour tout n et tout h, et donc

$$\mathbb{E}[M_{n\wedge\tau_v\wedge T_{v,h}}(M_{n\wedge\tau_v\wedge T_{v,h}} - h)] \leq (\deg G)\, h\, \mathbb{E}[M_{n\wedge\tau_v\wedge T_{v,h}}] = (\deg G)\, h\pi(v);$$
$$\mathbb{E}[n\wedge\tau_v\wedge T_{v,h}] \leq \frac{(\deg G + 1)\, h\pi(v)}{\sigma^2}.$$

Par convergence monotone, la dernière inégalité vaut aussi pour $\mathbb{E}[\tau_v \wedge T_{v,h}]$, et par l'inégalité de Markov, on obtient

$$\mathbb{P}[\tau_v > n \text{ et } T_{v,h} > n] \leq \frac{(\deg G + 1)\, h\pi(v)}{n\sigma^2}.$$

On choisit $h = \sqrt{\frac{n\sigma^2}{(\deg G+1)}}$ pour obtenir la borne

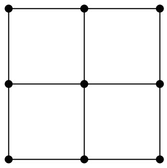
$$\mathbb{P}[\tau_v > n] \leq 4|E|\pi(v)\sqrt{\frac{\deg G + 1}{n}} \leq \frac{2\,(\deg G + 1)^{\frac{3}{2}}}{\sqrt{n}},$$

ce qui conclut avec la Proposition 5. □

Questions

1. Démontrer le théorème 1.

2. On considère comme graphe la grille G de taille 3×3 :



   Écrire un programme `MarcheAleatoire(n,v)` qui simule les n premiers pas de la marche aléatoire paresseuse $(X_n)_{n\in\mathbb{N}}$ sur ce graphe, avec $X_0 = v$. Donner la mesure invariante π pour ce graphe, et proposer une simulation qui montre la convergence en loi de X_n vers π.

3. Si S est une partie de V, démontrer la propriété (4).

4. Quels sont les états récurrents de la chaîne des ensembles évoluants ? Quels sont ses états transients ?

5. Avec le même graphe que précédemment, écrire un programme `EnsemblesEvoluants(n,S)` qui simule les n premiers pas de la chaîne des ensembles évoluants issue de $S_0 = S$. Vérifier expérimentalement le Corollaire 4.

6. Expliquer pourquoi le terme de gauche de l'inégalité (1) est bien toujours un réel entre 0 et 1.

7. Démontrer le corollaire 4.

8. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités et X une variable aléatoire réelle de carré intégrable. Si $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$, montrer que $\mathbb{E}[X \mid \mathcal{G}]$ est la projection orthogonale de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sur le sous-espace fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$. En déduire que
$$\mathrm{var}(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathrm{var}(X),$$
   ce qui est utilisé dans la preuve du lemme 6.

9. En utilisant les identités qui les précèdent, établir les deux égalités (2) et (3) relatives à l'espérance conditionnelle de $\pi(S_1)$.

10. Expliquer pourquoi la propriété (4) implique $M_{n+1} \leq (\deg G + 1) M_n$ presque sûrement.

11. Écrire un programme qui évalue la distribution du temps d'arrêt τ_v, G étant de nouveau la grille 3×3 et v l'un des sommets de ce graphe.

12. Si le graphe G est fixé, l'inégalité proposée par le théorème 7 vous paraît-elle optimale en n ? Quel est l'intérêt de cette inégalité ?