Processus Stochastiques Rémy Mahfouf
ENS Paris, 2021-2022 Bureau V2
remy.mahfouf@ens.fr

TD 11 : Temps d'atteinte et fonctions harmoniques

Vendredi 17 Décembre

Exercice 1 (Temps d'atteinte depuis la mesure invariante)
On se donne une chaîne de Markov X irreductible récurrente positive de matrice de transition Q. On pose $T_{x,0} := \inf\{n \geq 0 : X_n = x\}$, $T_x := \inf\{n \geq 1 : X_n = x\}$ et de manière plus générale on définit $T_{x,m} := \inf\{n \geq m : X_n = x\}$. On note μ la probabilité invariante associée à la chaine.

1. Montrez que $\mathbb{E}_x[T_{x,m}] < \infty$ puis que pour $y \in E$ on a

$$\mu(y) = \frac{\mathbb{E}_x[\sum_{n=0}^{T_{x,m}-1} \mathbb{1}_{X_n=y}]}{\mathbb{E}_x[T_{x,m}]}.$$

2. Montrez que l'on a

$$\sum_{k=0}^{m-1} Q^k(x,x) = \mu(x)\mathbb{E}_x[T_{x,m}] = \mu(x)(m + \mathbb{E}_{\mu_m}[T_{x,0}]),$$

où μ_m est la loi définie par $\mu_m(y) := Q^m(x,y)$. En déduire que

$$\sum_{k=0}^{m-1} \left(Q^k(x,x) - \mu(x)\right) = \mu(x)\mathbb{E}_{\mu_m}[T_{x,0}].$$

3. On suppose de plus que $\sup_y E_y[T_x] < \infty$. Montrez que l'on a

$$\mathbb{E}_\mu[T_{x,0}] = \mathbb{E}_x[T_x] \sum_{k=0}^{\infty} \left(Q^k(x,x) - \mu(x)\right).$$

On a ainsi exprimé le temps de premier retour en x en partant depuis la mesure stationnaire.

4. Considerons la chaîne de Markov des mots de taille 2 sur l'aphabet $\{A, B\}$ avec A qui apparait à droite avec probabilité p. Ecrire la matrice de probabilité associée et montrez que Q^2 est une matrice dont les lignes sont toutes égales à la mesure invariante. En déduire que $\mathbb{E}_\mu[T_{AA}] = \frac{1+p}{p^2} - 2$, $\mathbb{E}_\mu[T_{AB}] = \mathbb{E}_\mu[T_{BA}] = \frac{1}{p(1-p)} - 2$. Que dire sur l'esperance du temps d'écriture de ABRACADABRA par un singe tapant au hasard sur un clavier de 26 lettres ?

<u>*Solution de l'exercice 1*</u>

1. On montre tout d'abord que $\mathbb{E}[T_{x,m}] < \infty$. Le cas $m = 1$ provient du fait que la chaîne est récurrente positive. Pour le cas $m = 2$, on remarque que $\mathbb{E}_x[T_{x,2}] = \mathbb{E}_x[T_{x,2}\mathbb{1}_{T_{x,1}=1}] + \mathbb{E}_x[T_{x,2}\mathbb{1}_{T_{x,1}>1}]$. On applique au premier terme la propriété de Markov au temps 1, ce qui produit une contribution $\mathbb{P}_x(T_{x,1} = 1)(\mathbb{E}_x[T_{x,1}] + 1)$. Pour le second terme, on remarque que sur l'événement $\{T_{x,1} > 1\}$, on a $T_{x,1} = T_{x,2}$ presque sûrement. Cela permet de majorer la contribution du second terme par $\mathbb{E}[T_{x,1}] < \infty$. On réapplique la même stratégie pour montrer par récurrence le résultat. Ensuite, par unicité de la mesure invariante (à constante multiplicative près), il suffit de vérifier que l'application

$y \mapsto \mathbb{E}_x\big[\sum_{n=0}^{T_{x,m}-1} \mathbb{1}_{X_n=y}\big]$ est invariante ainsi que la normalisation. On procède comme pour le temps d'arret $T_{x,1}$ étudié en cours. On a alors

$$\mu(y) = \sum_{z\in E} \mathbb{E}_x\Big[\sum_{k=1}^{T_{x,m}} \mathbb{1}_{X_{k-1}=z,X_k=y}\Big] = \sum_{z\in E} \mathbb{E}_x\Big[\sum_{k=1}^{\infty} \mathbb{1}_{k\le T_{x,m},X_{k-1}=z}\Big] Q(z,y) = \sum_{z\in E} \mu(z)Q(z,y).$$

La normalisation est obtenue en sommant sur toutes les valeurs de y.

2. La première égalité est simple en écrivant $Q^k(x,x) = \mathbb{E}_x[\mathbb{1}_{X_k=x}]$. Pour la seconde on conditionne par la valeur au temps m et on applique la propriété de Markov simple au temps m.
3. Comme $\sup_y E_y[T_x] < \infty$, on voit directement que $\mathbb{E}_\mu[T_x] < \infty$. On fait tendre m vers l'infini, dans l'équation obtenue dans la réponse précédente. Reste à utiliser la convergence de μ_m vers μ. Comme E est dénombrable et que l'on travaille avec des mesures de probabilité, on voit directement que $\lim_{m\infty} \sum_{y\in E} |\mu_m(y) - \mu(y)| = 0$ ce qui permet de conclure (modulo le fait que la chaîne soit apériodique).
   Si la chaîne n'est pas apériodique, on peut la rendre légérement paraisseuse en un point, utiliser le résultat ci-dessus puis passer à la limite en envoyant le paramètre de paresse vers 0.
4. Sur la base (AA, AB, BA, BB) on a

$$Q = \begin{pmatrix} p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \\ p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \end{pmatrix}$$

   La mesure invariante est donnée par le vecteur $(p^2, p(1-p), p(1-p), (1-p)^2)$. On remarque directement que les lignes de Q^2 sont la mesure invariante. Pour voir cela on remarque que la matrice Q^2 est de rang 1 puisque 0 est valeur propre de multiplicité au moins 2, 1 est valeur prendre de multiplicité 1 et la somme des valeurs propres est nulle. On en déduit que 0 est encore valeur propre une troisième fois compté avec avec multiplicité. La matrice Q^2 est donc de rang 1, ce qui donne le résultat. Dans le cas de l'alphabet à 26 lettres, on remarque que la mesure invariante est la mesure uniforme (pour cela on peut utiliser le théorème ergodique et l'indépendance des tirages sucessifs). Le même argument reste vrai pour la chaîne mentionée ci-dessus. Il faut nécessairement 11 itérations de la chaine pour tomber sur la mesure invariante (on peut alors passer de tout mot de 11 lettres à un autre avec une probabilité 26^{-11}). Si $x = \text{ABRACADABRA}$, les entiers $k \le 10$ tel que $Q^k(x,x) > 0$ sont $0, 7, 10$, ce qui donne le résultat $\mathbb{E}_{\text{Unif}}[T_{\text{ABRACADABRA}}] = 26^{11} + 26^4 + 26 - 11$.

Exercice 2 (Fonctions harmoniques et martingales)
On rappelle qu'une fonction f positive définie sur un espace d'états dénombrable E (localement fini) est harmonique si pour tout $x \in E$ on a $f(x) = Qf(x)$.

1. Montrez que f est harmonique ssi pour tout $x \in E$, le processus $(f(X_n))_{n\ge 0}$ est une martingale sous $\mathbb{P}_x$ pour la filtration naturelle. Montrez que pour tout $F \subseteq E$ et $G = E\backslash F$, $(f(X_{n\wedge T_G}))_{n\ge 0}$ est encore une martingale.
2. On définit le bord de F (noté ∂F) l'ensemble des points de F qui ont un voisins hors de F. Si aucun point ne vérifie cela le bord est l'infini. Montrez le principe du maximum i.e. si h est harmonique sur F fini, on a $\sup_{x\in F} h(x) = \sup_{x\in\partial F} h(x)$.
3. On suppose F non vide et T_G presque surement fini et on se donne une fonction $g : G \to \mathbb{R}$ bornée. Montrez que $h(x) := \mathbb{E}_x[g(T_G)]$ est l'unique fonction bornée qui est harmonique sur F et coincide avec g sur G. On dit que h est solution du problème de Dirichelet avec condition aux bords g.
4. Soit $H_{G,y}(\cdot, y)$ la solution du problème de Dirichelet avec $g = \mathbb{1}_y(\cdot)$ (on suppose à nouveau que T_G presque surement fini). Montrez que toute fonction harmonique s'écrit

$$h = \sum_{y\in G} h(y) H_{G,y}(\cdot, y).$$

5. Donnez l'exemple d'un graphe connexe où il existe un fonction harmonique bornée non constante.

Solution de l'exercice 2

1. Si f est harmonique on a $\mathbb{E}_x[f(X_{n+1})|\mathcal{F}_n] = Qf(X_n) = f(X_n)$. En conséquence $\mathbb{E}_x[f(X_n)] = f(x)$ donc $f(X_n) \in L^1$. La réciproque est immédiate puisque $f(x) = \mathbb{E}_x[f(X_0)] = \mathbb{E}_x[f(X_1)] = Qf(x)$. Pour la seconde partie de la question il suffit d'appliquer la théorie usuelle des martingales.
2. Comme le graphe est localement fini et chaque point est la moyenne de ses voisins, aucun maximum strict ne peut se trouver à l'interieur de F, ce qui permet de conclure. Le résultat analogue est aussi vrai pour le minimum.
3. La fonction proposée est solution du problème demandé ($T_G = 0$ présque surement si on démarre la marche depuis un point de G). Pour l'unicité on remarque que la différence de 2 solution est encore harmonique nulle au bord donc identiquement nulle d'après le principe du maximum.
4. La fonction $h - \sum_{y\in G} h(y)H_{G,y}(\cdot, y)$ est harmonique nulle au bord donc identiquement nulle.
5. On peut considérer la marche aléatoire simple sur un arbre binaire infini, et on définit la fonction harmonique $h(x) := \mathbb{P}_x$(la marche s'enfonce dans la partie gauche de l'arbre). h est positive, bornée par 1 et clairement constante puisque si on démarre très loin dans la partie gauche de l'arbre, on a une probabilité très faible de remonter et finir dans la partie droite.

Exercice 3 (Les parapluies)
Michel possède k parapluies qu'il garde soit chez lui, soit dans son bureau. Le matin, avant d'aller au travail, il regarde par la fenêtre le temps qu'il fait :
— s'il fait beau, il ne prend pas de parapluie ;
— s'il pleut et qu'il a un parapluie chez lui, il prend le parapluie et va au travail ;
— s'il pleut et qu'il n'y a plus de parapluie, il décide de rester chez lui.
Le soir, il fait la même chose (s'il pleut et qu'il n'a pas de parapluies, il décide de rentrer quand même). On suppose que chaque demi-journée, il pleut avec probabilité p et il fait beau avec probabilité $1-p$ avec $0 < p < 1$, et que de plus, les météos des demi-journées sont indépendantes. Michel souhaite ne rater en moyenne qu'une journée de travail par mois. Sachant qu'il pleut 111 jours par an à Paris, combien de parapluies doit-il posséder pour cela ?

<u>*Solution de l'exercice 3*</u>
Notons X_n le nombre de parapluies que Michel a chez lui le soir du n-ième jour. Chaque jour :
— avec proba $p(1-p)$, il pleut le matin mais pas l'après-midi, auquel cas X_n diminue de 1 (sauf si $X_n = 0$),
— avec proba $p(1-p)$, il pleut l'après-midi mais pas le matin, auquel cas X_n augmente de 1 (sauf si $X_n = k$),
— avec proba $p^2 + (1-p)^2$, il pleut toute la journée ou pas du tout, auquel cas X_n ne change pas (y compris si $X_n = 0$ ou k).
On en déduit que X est une chaîne de Markov sur $\{0, 1, \ldots, k\}$ de matrice de transition Q avec

$$Q(i,j) = \begin{cases} p(1-p) & \text{si } |i-j| = 1, \\ p^2 + (1-p)^2 & \text{si } i = j \notin \{0,k\}, \\ 1 - p(1-p) & \text{si } i = j \in \{0,k\}, \\ 0 & \text{si } |i-j| \geq 2. \end{cases}$$

Cette chaîne de Markov est irréductible car $Q^{|i-j|}(i,j) > 0$ pour tous i, j et apériodique car $Q(0,0) > 0$. Elle admet donc une unique mesure de probabilité stationnaire et converge vers cette mesure stationnaire. On cherche une mesure μ réversible pour Q, i.e. telle que pour tous i et j,

$$\mu(i)Q(i,j) = \mu(j)Q(j,i). \tag{1}$$

La condition 1 est trivialement vérifiée pour $i = j$ (par symétrie) et pour $|i-j| \geq 2$ (les deux membres sont nuls), donc il suffit d'avoir

$$\mu(i)Q(i, i+1) = \mu(i+1)Q(i+1, i)$$

pour tout $0 \leq i \leq k-1$, soit $p(1-p)\mu(i) = p(1-p)\mu(i+1)$, soit $\mu(i+1) = \mu(i)$. Ainsi, la mesure uniforme μ sur $\{0, 1, \ldots, k\}$ est réversible, donc stationnaire, pour Q, donc

$$\mathbb{P}(X_n = 0) \xrightarrow[n\to+\infty]{} \mu(0) = \frac{1}{k+1},$$

et

$$\mathbb{P}\left(\text{Michel reste chez lui le jour } n\right) = \mathbb{P}\left(X_n = 0 \text{ et il pleut le } n\text{-ième matin}\right) = p\,\mathbb{P}(X_n = 0) \xrightarrow[n\to+\infty]{} \frac{p}{k+1}.$$

Or, il pleut en moyenne 111 jours par an à Paris. On a donc $1-(1-p)^2 = \frac{111}{365}$, soit $p \approx 0,1658$. Pour ne rater qu'un jour de travail par mois en moyenne, il faut donc $\frac{p}{k+1} \leq \frac{1}{20}$ (4 semaines de 5 jours), donc Michel doit avoir au moins $20p-1 \approx 2,3$ parapluies, donc au moins 3.