Master MFA première année 2011-2012

Module D4MA1C20, Compléments magistère
Devoir n^{0}2
A rendre le 4 mai 2012

1. La tour rase t'elle les murs ?

On s'intéresse à la marche au hasard stationnaire $\Xi = (X_n)_{n\in\mathbb{Z}}$ de la tour sur un jeu d'échecs 8×8. On se demande quelle proportion du temps la marche passe sur le pourtour du jeu.

1– Soit $\Upsilon = (Y_n)_{n\in\mathbb{Z}}$ une chaîne de Markov irréductible et stationnaire, à valeurs dans un ensemble fini E. Soit $f : E \to \mathbb{R}$ une fonction. Montrer que la suite de variables aléatoires $\frac{1}{n}\sum_{k=1}^{n} f(Y_k)$ converge presque sûrement. Quelle et sa limite ?

2– Montrer que la marche au hasard Ξ est irréductible et apériodique. Quelle est la mesure stationnaire ?

3– Montrer que la proportion du temps, entre les temps 1 et n, que la marche passe sur le pourtour converge presque sûrement. Quelle est sa limite ?

4– Qu'en est il si on remplace la tour par le roi ?

Solution :

1– Soit μ la mesure sur $E^{\mathbb{Z}}$ invariante par le décalage T associée à la chaîne. On peut supposer que la variable aléatoire Y_k est la fonction $E^{\mathbb{Z}} \to \mathbb{R}$, $(x_n)_{n\in\mathbb{Z}} \mapsto x_k$. Soit $g : E^{\mathbb{Z}} \to \mathbb{R}$ la fonction définie par $g((x_n)_{n\in\mathbb{Z}}) = f(x_0)$, i.e. $g = f(Y_0)$. Par définition, $g \circ T^k = f(Y_k)$. Alors g est mesurable pour la tribu produit. En effet, g ne prend qu'un nombre fini de valeurs, et les ensembles $g^{-1}(t) = \bigcup_{f(x)=t}\{x_0 = x\}$ sont des réunions finies de cylindres. Elle est bornée donc intégrable. Par hypothèse, Υ est irréductible, donc μ est ergodique pour T. D'après le Théorème Ergodique de Birkhoff, la suite de fonctions

$$g_n = \frac{1}{n}\sum_{k=1}^{n} g \circ T^k$$

converge presque partout vers $\mathbb{E}(g)$. Par conséquent, $\frac{1}{n}\sum_{k=1}^{n} f(Y_n)$ converge presque sûrement vers $\mathbb{E}(f(Y_0))$.

Autre façon de trouver la valeur de la limite (pour Adrien) : il s'agit de $\int g\,d\mu$. Or, pour une fonction qui ne dépend que de x_0, $\int g\,d\mu = \sum_{x\in E} g(x)p(x)$ où p est la distribution stationnaire. En effet, lorsque g est une fonction caractéristique, i.e. $g((x_n)_{n\in\mathbb{Z}}) = 1 \Leftrightarrow x_0 = y$, alors $\int g\,d\mu = \mu\{x_0 = y\} = p(y)$, la formule est vraie. En général, g est une combinaison linéaire de telles fonctions, donc la formule est vraie en général.

2– La tour peut aller de toute position vers toute autre position en 2 pas, donc sa marche est irréductible. Elle peut revenir à son point de départ en 3 pas, donc elle est apériodique. La distribution stationnaire est unique. Le graphe de la marche possède 64 sommets et $64\times14/2=448$ arêtes, toutes ayant le même poids. La matrice des probabilités de transitions est symétrique, et donc bi-stochastique. La distribution stationnaire est uniforme.

3– On applique le résultat de la question 1. à la fonction caractéristique f des cases du pourtour (il y en a 28). La proportion du temps passé sur le pourtour vaut $\frac{1}{n}\sum_{k=1}^{n} f(X_n)$, elle converge presque sûrement vers la moyenne de f pour la distribution stationnaire, qui vaut $28/64 = 7/16$. La tour passe donc un peu moins de la moitié de son temps sur le pourtour de l'échiquier.

4– La marche du roi est aussi irréductible et apériodique, elle se déroule sur un graphe dont 4 sommets ont 3 arêtes, 24 ont 5 arêtes et 36 ont 8 arêtes. La distribution stationnaire donne un poids de $3/420$ aux angles, $5/420$ aux aux cases du pourtour, et $8/420$ aux cases intérieures (marche aléatoire sur un graphe, exemple 58 du cours, traité en Exercice 1 de la feuille 3). L'espérance, relativement à cette distribution, de la fonction f est $4 \times 3/420 + 24 \times 5/420 = 132/420 = 0.31...$. On conclut que le roi passe bien moins de temps sur le pourtour du jeu que la tour.

2. Capacité du canal radio-shadok

Radio-shadok émet au moyen d'un poste alimenté par la génératrice d'un vélo. La qualité du signal est assez mauvaise. Quand on l'écoute, on prend souvent un mot pour un autre, suivant la matrice de probabilités de transition suivante.

	ga	bu	zo	meu
ga	3/4	1/8	1/16	1/16
bu	1/8	3/4	1/16	1/16
zo	1/8	1/8	1/2	1/4
meu	1/8	1/8	1/4	1/2

On se demande comment calculer la capacité de ce canal.

1– Soit P la matrice des probabilités de transition d'un canal. Montrer que le calcul de la capacité $\kappa(P)$ revient à chercher la distribution de probabilité (matrice ligne p) qui maximise $H(pP) - ph(P)$, où $h(P)$ est le vecteur colonne dont les composantes sont les entropies des lignes de P.

2– On suppose P carrée et inversible. Montrer que $\kappa(P) = \sup_q I$ où $I(q) = H(q) - qe$, pour un vecteur colonne e que l'on précisera.

3– Montrer que I est concave, qu'elle atteint son maximum en une distribution q_0 qui charge tous les points. En utilisant le théorème des extrema liés, donner une expression pour q_0 qui s'apparente à la distribution de Maxwell-Boltzmann.

4– En déduire une formule exacte pour $\kappa(P)$. Application numérique : calculer la capacité de radio-shadok (une valeur approchée suffit).

Solution :

1– Soit X une variable aléatoire de loi p. Si on met X en entrée du canal, la variable aléatoire Y lue en sortie a pour loi q donnée par

$$q(y) = \sum_{x \in E_X} p_X(x) P_{xy},$$

soit $q = pP$. Pour chaque $x \in E_X$, la variable conditionnée $Y|X = x$ a pour loi

$$p_{Y|X=x}(y) = \mathbb{P}(Y = y | X = x) = P_{xy},$$

donc $H(Y|X=x) = h(P)(x)$ est l'entropie de la x-ème ligne, et

$$H(Y|X) = \sum_{x \in X} p_X(x)h(P)(x) = ph(P).$$

Il vient

$$I(X;Y) = H(Y) - H(Y|X) = H(pP) - ph(P).$$

C'est cette quantité qu'il faut maximiser sur l'ensemble $\mathcal{P}(E_X)$ des distributions de probabilité sur E_X.

2– Si P est inversible, $q = pP$ détermine $p = qP^{-1}$. Maximiser $H(pP) - ph(P)$ sur $\mathcal{P}(E_X)$ équivaut à maximiser $H(q) - qP^{-1}h(P)$ sur $\mathcal{P}(E_Y)$. On pose $e = P^{-1}h(P)$ et $\mathcal{E}(q) = qe$. Il vient

$$\kappa(P) = \sup_{q \in \mathcal{P}(E_Y)} H(q) - \mathcal{E}(q).$$

3– I est la différence d'une entropie (qui est concave) et d'une forme linéaire, c'est donc une fonction strictement concave. Elle est continue sur le compact $\mathcal{P}(E_Y)$ donc elle atteint son maximum. Par stricte concavité, le maximum est atteint en un point q_0 qui est unique et appartient à l'intérieur de $\mathcal{P}(E_Y)$. Au voisinage de ce point, la fonction I est différentiable.

L'ensemble $\mathcal{P}(E_Y)$ est contenu dans l'espace vectoriel des fonctions réelles sur E_Y. C'est un polyèdre convexe de l'hyperplan d'équation $\phi = 0$, où $\phi(q) = \sum_{y \in E_Y} q(y)$. D'après le Théorème des extrema liés, en q_0, les formes linéaires $dI = \sum_{y \in E_Y}(-\log_2(q_0(y)) - 1 - e(y))dq(y)$ et $d\phi = \sum_{y \in E_Y} dq(y)$ sont proportionnelles. Il existe un réel λ tel que pour tout $y \in E_Y$,

$$-\log_2(q_0(y)) - 1 - e(y) = \lambda.$$

Cela donne $q_0(y) = 2^{-\lambda-1-e(y)}$, et λ est déterminé par la condition

$$\sum_{y \in E_Y} 2^{-\lambda-1-e(y)} = 1.$$

Il vient $\lambda = -1 + \log_2(\sum_{y \in E_Y} 2^{-e(y)})$ et

$$q_0(y) = \frac{2^{-e(y)}}{\sum_{y \in E_Y} 2^{-e(y)}}.$$

4– On calcule

$$\begin{aligned} I(q_0) &= H(q_0) - q_0 e \\ &= \sum_{y \in E_Y} (\lambda + 1 + e(y))q_0(y) - q_0(y)e(y) \\ &= \sum_{y \in E_Y} (\lambda+1)q_0(y) \\ &= \lambda + 1 = \log_2(\sum_{y \in E_Y} 2^{-e(y)}). \end{aligned}$$

Application numérique :

$$P^{-1} = \frac{1}{10}\begin{pmatrix} 14 & -2 & -1 & -1 \\ -2 & -14 & -1 & -1 \\ -2 & -2 & 27 & 13 \\ -2 & -2 & -13 & 27 \end{pmatrix}, \quad h(P) = \begin{pmatrix} (19 - 6\log_2 3)/8 \\ (19 - 6\log_2 3)/8 \\ 7/4 \\ 7/4 \end{pmatrix}, \quad e = \begin{pmatrix} 1.0735 \\ 1.0735 \\ 1.9755 \\ 1.9755 \end{pmatrix}$$

puis

$$\sum_{y \in E_Y} 2^{-e(y)} = 1.4589, \quad \kappa = \log_2(1.4589) = 0.5449.$$