Structures et Algorithmes Aléatoires

F. Baccelli, P. Brémaud

2012

Les exercices sont entièrement **indépendants**.
Les notes de cours sont autorisées. Celles de TD ne le sont pas.
Le barème est indicatif.

1 Exercice 1 – Matrice Aléatoire [6 points]

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes prenant les deux valeurs -1 et $+1$ équiprobablement. On pose $X = X_1 + \cdots + X_n$.
1. (a) Montrer que pour tout $t > 0$ on a la borne

$$E\left[e^{tX_1}\right] \leq e^{\frac{t^2}{2}} .$$

(b) Montrer que pour tout $a > 0$

$$P(X \geq a) \leq e^{-\frac{a^2}{2n}} .$$

(c) Montrer que pour tout $a > 0$

$$P(|X| \geq a) \leq 2e^{-\frac{a^2}{2n}} .$$

2. Soit $A = \{a_{ij}\}_{1 \leq i \leq n, 1 \leq j \leq m}$ une matrice $m \times n$ dont les éléments sont à valeurs dans $\{0, 1\}$. Soit $b = \{b_j\}_{1 \leq j \leq m}$ un vecteur colonne dont les éléments sont à valeurs dans $\{-1, +1\}$ et soit $c = \{c_i\}_{1 \leq i \leq n}$ le vecteur colonne défini par $c = Ab$. On pose

$$||c||_\infty = ||Ab||_\infty = \max_{1 \leq i \leq n} |c_i| .$$

On choisit b aléatoire: $b_1, \ldots, b_n$ sont des variables aléatoires indépendantes prenant les deux valeurs dans -1 et $+1$ équiprobablement. Montrer que

$$P\left(||c||_\infty \geq \sqrt{4m \log n}\right) \leq \frac{2}{n} .$$

(pour cela on obtiendra une inégalité du même type pour chacun des c_i en considérant les deux cas $k \leq \sqrt{4m \log n}$ et $k > \sqrt{4m \log n}$, où k est le nombre de $+1$ dans la i-ème ligne de A.)

2 Exercice 2 – Pair-à-Pair [2 points]

Dans un système pair-à-pair, s'il y a n pairs $n > 0$, chacun d'entre eux reçoit un débit constant de chacun des $n-1$ autres. Le débit total est donc proportionnel à $n(n-1)$. On considère donc la chaîne de Markov homogène à temps discret sur $\mathcal{E} = \mathbb{N}^*$ de matrice de transition

$$p_{i,i+1} = \frac{\lambda}{\lambda + i(i-1)\mu}, i \geq 1 \quad p_{i,i-1} = \frac{i(i-1)\mu}{\lambda + i(i-1)\mu}, i \geq 2,$$

où λ représente l'intensité d'arrivée des pairs et $i(i-1)\mu$ l'intensité de leurs départs quand il y en a i.

Sous quelles conditions cette chaîne de Markov est-elle récurrente positive? Donner sa distribution stationnaire sous forme d'une série.

3 Exercice 3 – Récurrence Géométrique [7 points]

Soit $\{X_n\}$, $n \in \mathbb{N}$, une chaîne de Markov homogène à temps discret sur $\mathcal{E} = \mathbb{N}$, de matrice de transition $\mathbb{P} = (p_{ij})$ irréductible. On suppose que cette chaîne de Markov vérifie les propriétés suivantes : il existe un entier K positif, un réel $0 < \varepsilon < 1$ et une fonction $h : \mathbb{N} \to \mathbb{R}_+$, avec $h(i) \geq 1$ pour tout i, tels que

$$(\alpha) \qquad \sum_{j \in \mathbb{N}} p_{ij} h(j) \leq h(i)(1-\varepsilon) \quad \forall i \geq K,$$

$$(\beta) \qquad \sum_{j \in \mathbb{N}} p_{ij} h(j) < \infty \quad \forall i < K.$$

1. Montrer que $\{X_n\}$ est récurrente positive.

2. Montrer qu'il existe une fonction $g : \mathbb{N} \to \mathbb{R}_+$, avec $g(i) \geq 1$ pour tout i, telle que

$$(\gamma) \qquad \sum_{j \geq K} p_{ij} g(j) \leq g(i)(1-\varepsilon) \quad \forall i \in \mathbb{N}.$$

3. Soit T le temps de retour dans l'ensemble $\{0, 1, \ldots, K-1\}$:

$$T = \inf\{n \geq 1,\ X_n < K\}.$$

Calculer $P_i(T > m) = P(T > m | X(0) = i)$, $m, i \in \mathbb{N}$, en fonction des puissances de la matrice ${}_K\mathbb{P} = ({}_K p_{\ell j})$, $\ell, j \in \mathbb{N}$ définie par

$$\begin{aligned} {}_K p_{\ell j} &= p_{\ell j} \quad \text{si } j \geq K; \\ & 0 \quad \text{sinon.} \end{aligned}$$

4. Etablir la propriété de récurrence géométrique suivante : pour tout $m \in \mathbb{N}$, $m \geq 1$ et tout $i \in \mathbb{N}$,

$$P_i(T > m) \leq (1-\varepsilon)^m g(i),$$

où g est la fonction construite dans la question 2.

5. Soit $\{X_n\}$ la chaîne de Markov sur $\mathbb{N}$ matrice de transition

$$p_{i,i+1} = \frac{\lambda}{\lambda + \mu(i \wedge K)}, i \geq 0 \quad p_{i,i-1} = \frac{\mu(i \wedge K)}{\lambda + \mu(i \wedge K)}, i \geq 1.$$

Montrer que cette chaîne de Markov est irréductible. Donner sa condition de récurrence positive. Montrer que sous cette condition, le temps de retour T de la chaîne de Markov dans l'ensemble $\{0, 1, \ldots, K-1\}$ est tel qu'il existe des réels positifs x et ε, avec $\varepsilon < 1$, tels que

$$P_i(T > m) \leq (1+x)^i (1-\varepsilon)^m, \quad \forall i \geq K, m \geq 1.$$

4 Exercice 4 – Auto-configuration d'un Réseau Wifi [5 points]

On se donne un graphe d'interférence dont les nœuds $\{1, \ldots, n\}$ sont n points d'accès wifi. Deux points d'accès p et $q \in \{1, \ldots, n\}$ sont reliés par une arête dans ce graphe s'ils sont voisins et peuvent interférer. On suppose ce graphe connecté.

Il y a K (11 dans le Wifi) fréquences possibles et chaque point d'accès peut choisir l'une d'entre elles. On notera $f(p) \in \{1, \ldots, K\}$ la fréquence du point d'accès p. Si p et q choisissent la même fréquence, chacun reçoit de l'autre une interférence $g(p,q) = g(q,p) > 0$.

L'interférence vue par p sur la féquence k vaut donc

$$I(p,k) = \sum_{q \in \mathcal{N}_p} 1_{f(q)=k},$$

où $\mathcal{N}_p$ est l'ensemble des voisins de p.

On cherche un algorithme produisant un état $f(1), \ldots, f(n)$ où l'interférence totale

$$\mathcal{E}(f(1), \ldots, f(n)) = \sum_{p=1}^{n} I(p, f(p))$$

est petite.

Dans ce contexte un algorithme est dit distribué si à chaque étape de l'algorithme, chaque point d'accès p base ses décisions sur la connaissance de $I(p,k)$, $k = 1, \ldots, K$ seulement.

Soit $T > 0$, un paramètre connu de tous. Donner (en pseudo-code) un algorithme aléatoire distribué permettant de conduire le réseau dans une configuration où l'état $f(1), \ldots, f(n)$ est de probabilité

$$\pi(f(1), \ldots, f(n)) = \frac{1}{Z_T} e^{-\frac{1}{T}\mathcal{E}(f(1), \ldots, f(n))},$$

où Z_T est une constante de normalisation. Etudier la limite de cette probabilité lorsque T tend vers 0.

5 Solution Exercice 1

1. (a)

$$\begin{aligned}
E\left[e^{tX_1}\right] &= \frac{1}{2}(e^{+t}+e^{-t}) \\
&= 1+\frac{t^2}{2!}+\frac{t^4}{4!}+\frac{t^6}{6!}+\cdots \\
&\le 1+\frac{t^2}{2\times 1!}+\frac{t^4}{2\times 2\times 2!}++\frac{t^6}{2\times 2\times 2\times 3!}+\cdots = e^{\frac{t^2}{2}}
\end{aligned}$$

(b) Pour $t>0$,

$$\begin{aligned}
P(X\ge a) &= P\left(e^{tX}\ge e^{ta}\right) \\
&\le \frac{E\left[e^{tX}\right]}{e^{ta}}.
\end{aligned}$$

Mais (indépendance des X_i)

$$E\left[e^{tX}\right]=E\left[e^{t(X_1+\cdots+X_n)}\right]=E\left[e^{tX_1}\right]^n\le e^{\frac{nt^2}{2}}$$

et donc

$$\begin{aligned}
P(X\ge a) &\le e^{-(at-n\frac{t^2}{2})} \\
&\le e^{-\min_{t>0}(at-n\frac{t^2}{2})}=e^{-\frac{a^2}{2n}}.
\end{aligned}$$

(c) Par symétrie ($-X$ a la même distribution de probabilité que X),

$$P(X\le -a)=P(-X\ge a)=P(X\ge a)\le e^{-\frac{a^2}{2n}}.$$

Comme $\{|X|\ge a\}=\{X\ge a\}\cup\{X\le -a\}$, par sigma-additivité,

$$P(|X|\ge a)=P(X\ge a)+P(-X\ge a)\le 2e^{-\frac{a^2}{2n}}.$$

2.

Si le nombre k d'éléments égaux à $+1$ dans la i-ème ligne de A est $\le \sqrt{4m\ln n}$, alors $|c_i|\le\sqrt{4m\ln n}$, et si $k>\sqrt{4m\ln n}$, c_i est la somme de k variables IID prenant les valeurs -1 et $+1$ équiprobablement. Donc, comme $k\le m$

$$P\left(|c_i|\ge\sqrt{4m\ln n}\right)\le 2e^{-\frac{4m\ln n}{2k}}\le 2e^{-2\ln n}=\frac{2}{n^2}$$

et finalement

$$\begin{aligned}
P\left(||c||_\infty\ge\sqrt{4m\ln n}\right) &= P\left(\cup_{i=1}^n\{|c_i|\ge\sqrt{4m\ln n}\}\right) \\
&\le \sum_{i=1}^n P\left(|c_i|\ge\sqrt{4m\ln n}\right)\le\frac{2}{n}.
\end{aligned}$$

6 Solution Exercice 3

1. Comme $h \geq 1$, (α) implique

$$(\alpha') \qquad \sum_{j \in \mathbb{N}} p_{ij} h(j) \leq h(i) - \varepsilon \quad \forall i \geq K.$$

   La fonction h peut donc être prise comme fonction test dans le critère de Foster avec $F = \{0, 1, \ldots, K-1\}$ (les notations sont celles du cours).

2. Utilisant (α),

$$\sum_{j \geq K} p_{ij} h(j) \leq \sum_{j \in \mathbb{N}} p_{ij} h(j) \leq h(i)(1-\varepsilon), \quad \forall i \geq K,$$

   et on peut donc prendre $g(i) = h(i) \geq 1$ pour tout $i \geq K$. On cherche ensuite des réels $g(i)$ tels que $g(i) \geq 1$ et

$$g(i)(1-\varepsilon) \geq \sum_{j \geq K} p_{ij} g(j) = \sum_{j \geq K} p_{ij} h(j), \quad \forall i = 0, \ldots, K-1.$$

   On peut donc prendre

$$g(i) = \max\left(1, \frac{1}{1-\varepsilon} \sum_{j \geq K} p_{ij} h(j)\right), \quad i = 0, \ldots, K-1.$$

   Le fait que $g(i)$ ainsi défini est fini pour tout $i = 0, \ldots, K-1$ découle de l'hypothèse (β).

3. Montrons par récurrence sur m que pour tout $m \geq 1$ et tout $j \geq K$, et tout i,

$$P_i(T > m, X_m = j) = ({}_K\mathbb{P}^m)_{ij}\,.$$

   Cette propriété est trivialement vraie pour $m = 1$. Supposons la vraie pour $m \geq 1$. Alors pour tout $j \geq K$,

$$\begin{aligned}
P_i(T > m+1, X_{m+1} = j) &= \sum_{k \geq K} P_i(T > m+1, X_m = k, X_{m+1} = j) \\
&= \sum_{k \geq K} P_i(T > m, X_m = k, X_{m+1} = j) \\
&= \sum_{k \geq K} P_i(X_{m+1} = j | T > m, X_m = k) P_i(T > m, X_m = k) \\
&= \sum_{k \geq K} P_i(X_{m+1} = j | X_m = k)({}_K\mathbb{P}^m)_{ik} \\
&= \sum_{k \geq K} ({}_K\mathbb{P})_{kj}({}_K\mathbb{P}^m)_{ik} \\
&= \sum_{k \in \mathbb{N}} ({}_K\mathbb{P})_{kj}({}_K\mathbb{P}^m)_{ik} \\
&= ({}_K\mathbb{P}^{m+1})_{ij},
\end{aligned}$$

où on a utilisé la propriété de Markov puis l'hypothèse de récurrence. On en déduit:

$$P_i(T > m) = ({}_K\mathbb{P}^m\mathbf{1})_i\,,$$

où on note $\mathbf{1}$ le vecteur dont toutes les coordonnées sont égales à 1.

4. La preuve est par récurrence sur m. Pour $m = 1$, pour tout i,

$$\begin{aligned} P_i(T > 1) &= ({}_K\mathbb{P}\mathbf{1})_i = \sum_{j \geq K} ({}_K\mathbb{P})_{ij} \\ &= \sum_{j \geq K} p_{ij} \leq \sum_{j \geq K} p_{ij} g(j) \\ &\leq g(i)(1-\varepsilon). \end{aligned}$$

De même, si la propriété est vraie pour $m \geq 1$, alors

$$\begin{aligned} P_i(T > m+1) &= (({}_K\mathbb{P})({}_K\mathbb{P})^m\mathbf{1})_i \\ &= \sum_{j \geq K} p_{ij} P_j(T > m) \\ &\leq \sum_{j \geq K} p_{ij} g(j)(1-\varepsilon)^m \\ &\leq g(i)(1-\varepsilon)^{m+1}. \end{aligned}$$

5. On a

$$p_{i,i+1} = \frac{\lambda}{\lambda + \mu(i \wedge K)}, i \geq 0 \quad p_{i,i-1} = \frac{\mu(i \wedge K)}{\lambda + \mu(i \wedge K)}, i \geq 1.$$

La chaîne de Markov est ergodique ssi $\rho = \frac{\lambda}{K\mu} < 1$ d'après un théorème du cours. Si on pose $h(i) = (1+x)^i$, avec $x > 0$, la condition (β) est trivialement vérifiée et la fonction h est bien bornée inférieurement par 1; (α) est vérifiée ssi

$$(1+x)\frac{\lambda}{\lambda + \mu K} + \frac{1}{1+x}\frac{\mu K}{\lambda + \mu K} \leq \beta,$$

avec $\beta = 1 - \varepsilon$. Ceci est équivalent à

$$(1+x)\lambda + \frac{\mu K}{1+x} \leq (\lambda + \mu K)\beta,$$

ou encore à

$$x\lambda - \frac{\mu K x}{1+x} \leq (\lambda + \mu K)(\beta - 1),$$

ou encore à

$$x\left(\frac{\mu K}{1+x} - \lambda\right) \geq \varepsilon(\lambda + \mu K).$$

Comme $\rho < 1$, pour tout $x > 0$ suffisamment petit, $\frac{\mu K}{1+x} - \lambda > 0$. Un tel x étant fixé, on choisit alors

$$\varepsilon = \frac{x}{\lambda + \mu K}\left(\frac{\mu K}{1+x} - \lambda\right).$$

On peut donc trouver x et ε tels que (α) et (β) sont vérifiés. Comme $g(i) = h(i)$ pour $i \geq K$, on en déduit immédiatement la propriété demandée par application du résultat de la question 4.