Projet X-MAP-311-2012-CHAFAI-1

Matrices doublement stochastiques

Proposé par Djalil Chafaï

Second semestre 2011-2012

On dispose de n informations secrètes de même longueur, qu'on souhaite découper et répartir dans n fichiers de même longueur. Il est commode de modéliser la longueur d'une information ou d'un fichier par un nombre réel dans $[0,1]$. Si $P_{i,j} \in [0,1]$ désigne la portion de l'information i mise dans le fichier j, on a donc les contraintes suivantes :

$$\forall i,j \in \{1,\dots,n\}, \quad P_{i,j} \in [0,1], \quad \sum_{k=1}^{n} P_{i,k} = 1, \quad \sum_{k=1}^{n} P_{k,j} = 1.$$

On dit que la matrice $(P_{i,j})_{1\leq i,j\leq n}$ est doublement stochastique car chacune de ses lignes et de ses colonnes est une loi de probabilité discrète sur $\{1,\dots,n\}$.

Q1 Proposer un algorithme permettant de générer des matrices doublement stochastiques aléatoires $n \times n$. Implémenter cet algorithme en langage Scilab ;

Q2 Considérons la version discrète du problème : $P_{i,j} \in \{0,1,\dots,L\}/L$ pour tous $i,j \in \{1,\dots,n\}$ et pour un entier $L > 0$ fixé ne dépendant pas de n. La matrice LP est une sorte de carré magique $L \times L$ (pas de contrainte diagonale). Proposer un algorithme permettant de générer tous ces carrés magiques $L \times L$. Évaluer la complexité de l'algorithme. Implémenter cet algorithme en langage Scilab.

Étude de l'ensemble des matrices doublement stochastiques

On note $\mathcal{B}_n$ l'ensemble des matrices doublement stochastiques $n \times n$, et $\mathcal{S}_n$ le groupe des permutations de $\{1,\dots,n\}$.

Q3 Montrer que $\mathcal{B}_n$ est un sous-ensemble convexe et compact de $[0,1]^{n^2}$, possédant $(n-1)^2$ degrés de liberté, et qu'il s'agit d'un polytope de $\mathbb{R}^{n^2}$ (appelé polytope de Birkhoff). Donner un sens à la notion de loi de probabilité uniforme sur $\mathcal{B}_n$;

Q4 Proposer un algorithme de simulation de la loi uniforme sur $\mathcal{B}_n$ basé sur la méthode du rejet, et étudier sa complexité ;

Q5 On considère l'application $\Phi_n : \mathcal{S}_n \to \mathcal{M}_n(\{0,1\})$ définie par $\Phi_n(\sigma) = (\mathbf{1}_{j=\sigma(i)})_{1\leq i,j\leq n}$. Montrer que $\Phi_n(\mathcal{S}_n) \subset \mathcal{B}_n$ et que Φ_n est un homomorphisme de groupes injectif de $\mathcal{S}_n$ dans $\mathrm{GL}_n(\{0,1\})$. L'ensemble $\mathcal{B}_n$ est-il un groupe ?

Q6 Un théorème de Birkhoff et von Neumann affirme que $\Phi_n(\mathcal{S}_n)$ constitue l'ensemble des points extrémaux de $\mathcal{B}_n$. Trouver une preuve courte de ce théorème ;

Q7 Peut-on utiliser le théorème de Birkhoff et von Neumann pour simuler la loi uniforme sur $\mathcal{B}_n$?

Un algorithme markovien

Cette section est consacrée à un algorithme de simulation approchée de la loi uniforme sur le polytope de Birkhoff $\mathcal{B}_n$. On construit la suite récurrente aléatoire $(P^{(k)})_{k\geq 0}$ dans $\mathcal{B}_n$ en posant $k := 0$ et $P^{(k)} := \frac{1}{n} J_n$ où J_n est la matrice $n \times n$ pleine de 1, puis on exécute l'algorithme suivant :

I. sélectionner aléatoirement sans remise deux numéros de lignes i_1, i_2 dans $\{1, \ldots, n\}$ et aléatoirement sans remise deux numéros de colonnes j_1, j_2 dans $\{1, \ldots, n\}$;

II. simuler un nombre réel p dans l'intervalle à déterminer

III. construire $P^{(k+1)}$ à partir de $P^{(k)}$ en modifiant seulement et aléatoirement les entrées (i_1, j_1),(i_1, j_2),(i_2, j_1),(i_2, j_2) (transfert de masse conservatif) ;

IV. poser $k := k + 1$ et revenir à la première étape.

Q8 Proposer plusieurs critères d'arrêt pour l'algorithme ;

Q9 Implémenter l'algorithme en langage Scilab ;

Q10 Trouver une fonction $\Psi_{r,u} : \mathcal{M}_n(\mathbb{R}) \mapsto \mathcal{M}_n(\mathbb{R})$ paramétrée par $r \in \{1, \ldots, n\}^4$ et $u \in [0, 1]$ telle que $P^{(k+1)} = \Psi_{R_k,U_k}(P^{(k)})$ où (R_k, U_k) est un couple de variables aléatoires dont on précisera la loi sachant $P^{(k)}$;

Q11 Montrer que pour tous $P, Q \in \mathcal{B}_n$, il existe un entier $m \geq 1$ et une suite $R^{(1)}, \ldots, R^{(m)}$ dans $\mathcal{B}_n$ telle que $R^{(1)} = P$, $R^{(m)} = Q$, et $R^{(k+1)} = \Psi_{r_k,u_k}(R^{(k)})$ pour tout $k \in \{1, \ldots, m - 1\}$ (pour des paramètres $(r_1, u_1), \ldots, (r_{m-1}, u_{m-1})$ bien choisis) ;

Cet algorithme, utilisé avec un critère d'arrêt convenable, fournit des matrices doublement stochastiques aléatoires dont la loi est assez proche de la loi uniforme sur $\mathcal{B}_n$. On peut l'utiliser pour explorer des propriétés de cette loi.

Q12 Soit P une matrice doublement stochastique aléatoire suivant la loi uniforme sur $\mathcal{B}_n$. Utiliser votre implémentation de l'algorithme pour tester la conjecture suivante : lorsque $n \to \infty$, les coefficients de nP deviennent indépendants et convergent en loi vers une loi exponentielle ;

Q13 Montrer que pour tout $P \in \mathcal{B}_n$, le spectre de P est inclus dans $\{z \in \mathbb{C} : |z| \leq 1\}$ et contient toujours la valeur 1 ;

Q14 Soit P une matrice doublement stochastique aléatoire suivante la loi uniforme sur $\mathcal{B}_n$, et soit $\mu_n = \frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda_\ell}$ la mesure de comptage des valeurs propres de $\sqrt{n}P$. Utiliser votre implémentation de l'algorithme pour tester la conjecture suivante : avec probabilité 1, lorsque $n \to \infty$, la mesure de comptage μ_n converge vers la loi uniforme sur un disque de $\mathbb{C}$.

Matrices doublement stochastiques et couplage

La distance de couplage entre deux lois de probabilités μ et ν sur $\mathbb{R}$ est définie par

$$W(\mu, \nu) = \sqrt{\inf_{\eta \in \mathcal{C}(\mu,\nu)} \mathbb{E}(|X - Y|^2)}$$

où $\mathcal{C}(\mu, \nu)$ désigne l'ensemble des lois de probabilités sur $\mathbb{R}^2$ dont les lois marginales sont μ et ν, et où (X, Y) est un couple de variables aléatoires sur $\mathbb{R}^2$ de loi η.

Q15 Montrer que $\mathcal{C}(\mu,\nu)$ est un ensemble non vide et convexe ;

Q16 On considère le cas purement atomique où $\mu = \frac{1}{n}\sum_{\ell=1}^n \delta_{a_i}$ et $\nu = \frac{1}{n}\sum_{\ell=1}^n \delta_{b_i}$. Montrer que pour toute loi de probabilité η sur $\mathbb{R}^2$, on a $\eta \in \mathcal{C}(\mu,\nu)$ si et seulement si $n(\eta(\{(a_i,b_j)\}))_{1\le i,j\le n} \in \mathcal{B}_n$, et que

$$nW(\mu,\nu)^2 = \min_{P\in\mathcal{B}_n} \sum_{1\le i,j\le n} P_{i,j}(a_i-b_j)^2;$$

Q17 En déduire, à l'aide du théorème de Birkhoff et von Neumann, que

$$nW(\mu,\nu)^2 = \min_{\sigma\in\mathcal{S}_n} \sum_{1\le i\le n} (a_i-b_{\sigma(i)})^2;$$

Q18 En déduire que si $a,b\in\mathbb{R}_+^n$ alors, en notant $a_{(1)}\le\cdots\le a_{(n)}$ et $b_{(1)}\le\cdots\le b_{(n)}$ le réordonement croissant des suites a,b, on a

$$W(\mu,\nu)^2 = \frac{1}{n}\sum_{i=1}^n (a_{(i)}-b_{(i)})^2.$$

Q19 Montrer que si $A,B\in\mathcal{M}_n(\mathbb{C})$ sont normales, c'est-à-dire que $AA^*=A^*A$ et $BB^*=B^*B$, alors, en notant les spectres $\lambda_1(A),\ldots,\lambda_n(A)$ et $\lambda_1(B),\ldots,\lambda_n(B)$,

$$nW^2\left(\frac{1}{n}\sum_{\ell=1}^n \delta_{|\lambda_\ell(A)|}, \frac{1}{n}\sum_{\ell=1}^n \delta_{|\lambda_\ell(B)|}\right) \le \mathrm{Trace}((A-B)(A-B)^*) = \sum_{1\le i,j\le n} |A_{i,j}-B_{i,j}|^2.$$

Indication : utiliser la décomposition en valeurs singulières (SVD) de A et de B, et le fait que si $U\in\mathcal{M}_n(\mathbb{C})$ est unitaire, alors $(|U_{i,j}|^2)_{1\le i,j\le n}\in\mathcal{B}_n$.

Ouverture culturelle

Les concepts et formules abordés précédemment sont reliés au problème de l'affectation optimale randomisé, un grand classique en optimisation combinatoire randomisée. On dispose de n tâches et de n machines. On note $c_{i,j}$ le coût d'exécution de la tâche i sur la machine j. On souhaite affecter une et une seule tâche par machine de sorte que le coût d'exécution total soit minimal. Si on code une répartition des tâches par une permutation de $\{1,\ldots,n\}$, appartenant au groupe symétrique $\mathcal{S}_n$, le coût total optimal C_n ainsi que l'ensemble des répartitions optimales $\mathcal{O}_n$ sont

$$C_n := \min_{\sigma\in\mathcal{S}_n}\sum_{i=1}^n c_{i,\sigma(i)} \quad \text{et} \quad \mathcal{O}_n = \arg\min_{\sigma\in\mathcal{S}_n}\sum_{i=1}^n c_{i,\sigma(i)}.$$

Lorsque les $(c_{i,j})_{1\le i,j\le n}$ sont aléatoires, indépendantes, et de loi exponentielle, alors

$$\mathbb{E}(C_n) = \sum_{\ell=1}^n \frac{1}{\ell^2}.$$

Cette formule, attribuée à Parisi, peut être établie de manière élémentaire en utilisant les propriétés remarquables de la loi exponentielle, comme l'a montré Wästlund. On a donc $\lim_{n\to\infty}\mathbb{E}(C_n)=\zeta(2)=\frac{\pi^2}{6}$. Concernant la variance, Wästlund a également obtenu $\mathrm{Var}(C_n)\sim\frac{4}{n}(\zeta(2)-\zeta(3))$, précisant un résultat de Talagrand. Ces comportements asymptotiques sont universels : ils sont valables au delà de la loi exponentielle.

Références

- Sourav Chatterjee, Persi Diaconis, Allan Sly, *Properties of Uniform Doubly Stochastic Matrices*, arXiv:1010.6136;
- Johan Wästlund, *An easy proof of the* $\zeta(2)$ *limit in the random assignment problem*, Electronic Communications in Probability, vol. 14 paper num. 26 (2009).