Questions existentielles
La méthode probabiliste

Exercice 1 Comptage

On note K_n le graphe complet à n sommets. Une ***clique*** est un sous-graphe complet. Soit C le nombre de cliques à k sommets, on les numérote de 1 à C.

▷ 1. On colorie chaque arête de K_n indépendemment, avec probabilité $1/2$ pour chacune des deux couleurs. Quelle est la probabilité de l'événement A_i « la i^e^ clique est monochromatique. » ?

▷ 2. On suppose dans le reste de l'exercice que $\frac{C_n^k}{2^{C_k^2-1}} < 1$ et $k \geqslant 3$. Montrer qu'alors $\mathbb{P}\left(\bigcup_{i=1}^{C} A_i\right) < 1$.

▷ 3. En déduire qu'il existe un 2-coloriage de K_n tel qu'il n'y ait pas de clique monochromatique de taille k.

Solution

▷ 1. On commence par choisir l'une des deux couleurs. Puis chaque arête a une chance sur dexu d'être de la bonne couleur. $\mathbb{P}(A_i) = 2 \times 2^{-C_k^2} = 2^{-C_k^2+1}$.

▷ 2. $\mathbb{P}\left(\bigcup_{i=1}^{C} A_i\right) \leqslant \sum_1^C \mathbb{P}(A_i) = \frac{C_n^k}{2^{C_k^2-1}}$

▷ 3. $\cup_{i=1}^{C} A_i$ est l'événement « Il existe une clique monochromatique de taille k ». On vient de montrer que le contraire a une probabilité non nulle. Lors d'un tirage on a une probabilité non nulle de trouver un coloriage avec la proriété voulue, donc un tel coloriage existe.

Exercice 2 Espérance

Une ***coupe*** d'un graphe est une partition des sommets en deux ensembles A et B. Le ***poids*** d'une coupe est le nombre d'arêtes ayant une extrémité dans A et l'autre dans B.

▷ 1. *(lemme)* Soit $X : \Omega \to E$ une variable aléatoire d'espérance μ, avec E dénombrable. Montrer que $\mathbb{P}(X \geqslant \mu) > 0$ et $\mathbb{P}(X \leqslant \mu) > 0$.

▷ 2. On affecte indépendemment chaque sommet à A ou B avec probabilité $\frac{1}{2}$. Soit C le poids de la coupe obtenue. Calculer $\mathbb{E}(C)$.

▷ 3. En déduire qu'il existe une coupe de poids au moins $\frac{m}{2}$ où m est le nombre d'arêtes.

Solution

▷ 1. Supposons par l'absurde $\mathbb{P}(X \geqslant \mu) = 0$.

$$\begin{aligned}\mathbb{E}(X) &= \textstyle\sum_{a\in E} a\mathbb{P}(X = a)\\ &= \textstyle\sum_{a\in E, a<\mu} a\mathbb{P}(X = a)\\ &< \textstyle\sum_{a\in E, a<\mu} \mu\mathbb{P}(X = a) < \mu\end{aligned}$$

Contradiction. L'autre cas est symétrique.

▷ 2. Soit m le nombre d'arêtes. Soit E_i la variable aléatoire qui vaut 1 si la i^{e}arête est dans la coupe, 0 sinon. $\mathbb{E}(C) = \mathbb{E}\left(\sum_{i=1}^m E_i\right) = \sum_{i=1}^m \mathbb{E}(E_i) = \frac{m}{2}$.

▷ 3. D'après la question 1, $\mathbb{P}\left(C \geqslant \frac{m}{2}\right) > 0$ donc il existe une coupe de poids au moins $\frac{m}{2}$.

Exercice 3 Dérandomisation par espérance conditionnelle

On garde les notations de l'exercice précédent. Soit p la probabilité d'obtenir une coupe de poids au moins $\frac{m}{2}$ par le tirage précédent.

▷ 1. Montrer que $\frac{1}{p} = O(m)$ (i.e. minorer p).

▷ 2. En déduire un algorithme pour trouver une telle coupe dont l'espérance du temps de calcul est $O(m^2)$.

▷ 3. On cherche maintenant un algorithme déterministe. On note les sommets $s_1, \ldots, s_n$. Soit X_i l'ensemble auquel est affecté s_i (X_i est donc une variable aléatoire qui vaut A ou B). Montrer que $\mathbb{E}(C \mid X_1) = \mathbb{E}(C)$.

▷ 4. Supposons $X_1, \ldots, X_k$ fixés. Montrer que l'on peut choisir $x \in \{A, B\}$ de façon à ce que

$$\mathbb{E}(C \mid X_1, \ldots, X_k, X_{k+1} = x) \geqslant \mathbb{E}(C \mid X_1, \ldots, X_k).$$

▷ 5. En déduire un algorithme déterministe qui fixe les valeurs de $X_1, \ldots, X_n$, tel que $\mathbb{E}(C \mid X_1, \ldots, X_n) \geqslant \mathbb{E}(C)$.

▷ 6. Exprimer $\mathbb{E}(C \mid X_1, \ldots, X_k, X_{k+1} = A) - \mathbb{E}(C \mid X_1, \ldots, X_k, X_{k+1} = B)$ en fonction des voisins de v_{k+1} et donner une description simple l'algorithme précédent.

Solution

▷ 1. On node $w(c)$ le poids de la coupe c. On note que pour tout c, $w(c) \leqslant m$.

$$\begin{aligned}\tfrac{m}{2} = \mathbb{E}(C) = \sum_c \frac{w(c)}{2^n} &= \sum_{w(c)<\frac{m}{2}} \frac{w(c)}{2^n} + \sum_{w(c)\geqslant\frac{m}{2}} \frac{w(c)}{2^n} \\ &\leqslant \sum_{w(c)<\frac{m}{2}} \frac{\frac{m}{2}-\frac{1}{2}}{2^n} + \sum_{w(c)\geqslant\frac{m}{2}} \frac{m}{2^n} \\ &\leqslant (1-p)(\frac{m}{2}-\frac{1}{2}) + pm\end{aligned}$$

Donc
$$p \geqslant \tfrac{1}{1+m}$$

▷ 2. On peut vérifier le poids d'une coupe en un temps m. Si on teste des coupes aléatoires jusqu'à trouver une coupe convenable, on teste en moyenne $1/p = O(m)$ coupes. L'espérance du temps de calcul est donc $O(m^2)$.

▷ 3. $\mathbb{E}(C \,|\, X_1)$ est une fonction de X_1, qui vaut $\mathbb{E}(C \,|\, X_1 = A)$ si $X_1 = A$ et idem pour B sinon. Par symétrie, $\mathbb{E}(C \,|\, X_1 = A) = \mathbb{E}(C \,|\, X_1 = B)$.
Or $\mathbb{E}(C) = \mathbb{P}(X_1 = A)\,\mathbb{E}(C \,|\, X_1 = A) + \mathbb{P}(X_1 = B)\,\mathbb{E}(C \,|\, X_1 = B)$ et $\mathbb{P}(X_1 = A) = \frac{1}{2}$.
Donc $\mathbb{E}(C \,|\, X_1)$ est une fonction constante, égale à $\mathbb{E}(C)$.

▷ 4.
$$\mathbb{E}(C \,|\, X_1, \ldots, X_k) = \mathbb{E}(C \,|\, X_1, \ldots, X_k, X_{k+1} = A) + \mathbb{E}(C \,|\, X_1, \ldots, X_k, X_{k+1} = B)$$
donc l'un des deux termes de droite est au moins égal à celui de gauche.

▷ 5. Pour k de 1 à n, on fixe X_k de façon à ce que
$$\mathbb{E}(C \,|\, X_1, \ldots, X_{k-1}, X_k) \geqslant \mathbb{E}(C \,|\, X_1, \ldots, X_{k-1}).$$
L'espérance commence donc à $\mathbb{E}(C)$ et augmente jusqu'à ce que l'on ait fixé tous les X_k.
C'est un algorithme glouton.

▷ 6. On note à nouveau E_i l'ensemle auquel est affecté la i^{e}arête. On veut exprimer $\sum_i \mathbb{E}(E_i \,|\, X_1, \ldots, X_k, X_{k+1} = A) - \sum_i \mathbb{E}(E_i \,|\, X_1, \ldots, X_k, X_{k+1} = B)$.
Si la i^{e}arête n'a pas v_{k+1} pour extrémité, $\mathbb{E}(E_i)$ ne change pas quand on affecte v_{k+1} à A ou B. Sinon, v_{k+1} est l'une des extrémité de la i^{e}arête, on distingue trois cas pour l'autre extrémité :
– Elle est déjà affectée à A. Alors $E_i = 0$ si on effecte v_{k+1} à A, 1 sinon.
– Elle est déjà affectée à B. Alors $E_i = 1$ si on effecte v_{k+1} à A, 0 sinon.
– Elle n'est pas encore affectée. Alors $\mathbb{E}(E_i)$ reste égal à $\frac{1}{2}$.
Ainsi, si v_{k+1} a plus de voisins dans B, affecter v_{k+1} à A fait augmenter l'espérance du poids de la coupe.
L'algorithme précédent consiste donc, pour chaque sommet, à l'effecter dans l'ensemble qui maximise la contribution de ce sommet au poids de la coupe (selon les sommets déjà affectés).

Exercice 4 Sample and modify

Soit G un graphe à n sommets et m arêtes, et $d := 2m/n$ le degré moyen de G. Un ***stable*** est un sous-ensemble des sommets ne contenant pas deux sommets voisins. On considère l'algorithme probabiliste suivant pour trouver un stable :
– Supprimer chaque sommet (et ses arêtes adjacentes) indépendemment avec probabilité $1 - \frac{1}{d}$.
– Supprimer chaque arête restante et supprimer l'une de ses extrémités.

▷ 1. Montrer que l'on obtient bien un stable.

▷ 2. Calculer l'espérance du nombre X de sommets à la fin de la première étape.

▷ 3. Calculer l'espérance du nombre Y d'arêtes à la fin de la première étape.

▷ 4. En déduire qu'il existe un stable de taille au moins $\frac{n^2}{4m}$.
Remarque. Touver un stable de taille maximale est NP-complet.

Solution

▷ 1. La deuxième étape garantit que l'on obtient bien un stable : deux sommets ne peuvent pas être voisins : s'ils le sont, on en supprime un des deux à cette étape.

▷ 2. Chaque sommet survit avec probabilité $\frac{1}{d}$ donc $\mathbb{E}(X) = \frac{n}{d} = \frac{n^2}{2m}$.

▷ 3. Une arête survit si et seulement si se deux extrémités survivent, donc elle survit avec probabilité $\frac{1}{d^2}$.
$\mathbb{E}(Y) = \frac{m}{d^2} = \frac{n^2}{4m}$.

▷ 4. L'espérance de la taille du stable est $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y) = \frac{n^2}{4m}$. Il existe donc un stable de taille au moins $\frac{n^2}{4m}$.