Optimisation

Algorithmes d'approximation par relaxation de programme linéaire

Louis-Claude Canon
louis-claude.canon@univ-fcomte.fr

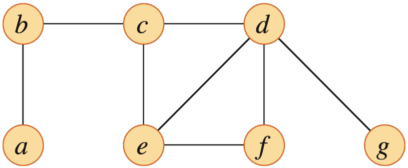
Master 2 Informatique – Semestre 9

Plan

Plan

Problème de la couverture de sommets de poids minimal

- Un graphe non-orienté $G = (V, E)$ avec une pondération sur les arêtes $w : E \to \mathbb{N}$.
- Le poids d'une couverture de sommets C est $w(C) = \sum_{v \in C} w(v)$.
- Le *problème de la couverture de sommets de poids minimal* consiste à trouver la couverture de sommets de poids minimal.
- Ce problème généralise le problème sans pondération. Quelle pondération permet d'obtenir ce problème plus simple ?



Formulation en programme entier

- La décision se porte sur chaque sommet que l'on conserve ou non.
- On définit donc une variable binaire par sommet : si elle vaut 1, on conserve le sommet associé dans la couverture, sinon elle vaut 0 et on ne le conserve pas.
- On obtient alors un *programme entier* (ou programme linéaire en nombre entier) dont la résolution est NP-difficile.

$$
\begin{array}{lll}
\text{minimiser} & \displaystyle\sum_{v \in V} w(v)x(v) & \\
\text{sous les contraintes} & x(u) + x(v) \geq 1 & \text{pour chaque } (u, v) \in E \\
& x(v) \in \{0, 1\} & \text{pour chaque } v \in V
\end{array}
$$

Relaxation d'un programme entier

- Pour permettre la résolution en temps polynomial, on transforme chaque variable binaire en une variable continue contrainte dans l'intervalle $[0; 1]$.
- Cette formulation est appelé le *programme linéaire relaxé* (ou sa relaxation) car les contraintes sont plus relâchées.
- Sa valeur de l'objectif optimale est donc au moins aussi bonne (c'est une **borne inférieure** de la valeur de l'objectif optimale).

$$
\begin{array}{ll}
\text{minimiser} & \displaystyle\sum_{v \in V} w(v)x(v) \\
\text{sous les contraintes} & x(u) + x(v) \geq 1 \text{ pour chaque } (u, v) \in E \\
& x(v) \leq 1 \text{ pour chaque } v \in V \\
& x(v) \geq 0 \text{ pour chaque } v \in V
\end{array}
$$

Question

Quelle contrainte est redondante dans cette formulation relaxée ?

$$
\begin{aligned}
\text{minimiser} \quad & \sum_{v \in V} w(v)x(v) \\
\text{sous les contraintes} \quad & x(u) + x(v) \geq 1 \text{ pour chaque } (u, v) \in E \\
& x(v) \leq 1 \text{ pour chaque } v \in V \\
& x(v) \geq 0 \text{ pour chaque } v \in V
\end{aligned}
$$

1. Aucune.
2. La première.
3. La seconde.
4. La troisième.

Question

Quelle contrainte est redondante dans cette formulation relaxée ?

$$\begin{aligned} \text{minimiser} \quad & \sum_{v \in V} w(v)x(v) \\ \text{sous les contraintes} \quad & x(u) + x(v) \geq 1 \text{ pour chaque } (u, v) \in E \\ & x(v) \leq 1 \text{ pour chaque } v \in V \\ & x(v) \geq 0 \text{ pour chaque } v \in V \end{aligned}$$

1. Aucune.
2. La première.
3. La seconde. ✓
4. La troisième.

Question

Comment prouver que la seconde contrainte est redondante ?

$$
\begin{array}{lr}
\text{minimiser} & \displaystyle\sum_{v \in V} w(v)x(v) \\
\text{sous les contraintes} & x(u) + x(v) \geq 1 \text{ pour chaque } (u,v) \in E \\
& x(v) \leq 1 \text{ pour chaque } v \in V \\
& x(v) \geq 0 \text{ pour chaque } v \in V
\end{array}
$$

1. Par le théorème de la dualité en programmation linéaire.
2. Par récurrence sur les sommets.
3. Par un argument d'échange.
4. Par l'absurde.

Question

Comment prouver que la seconde contrainte est redondante ?

$$\begin{aligned}
\text{minimiser} \quad & \sum_{v \in V} w(v)x(v) \\
\text{sous les contraintes} \quad & x(u) + x(v) \geq 1 \text{ pour chaque } (u, v) \in E \\
& x(v) \leq 1 \text{ pour chaque } v \in V \\
& x(v) \geq 0 \text{ pour chaque } v \in V
\end{aligned}$$

1. Par le théorème de la dualité en programmation linéaire.
2. Par récurrence sur les sommets. ✓
3. Par un argument d'échange.
4. Par l'absurde. ✓(une solution optimale avec $x(v) > 1$ peut être améliorée)

Couverture-Sommet-Pondéré-Approchée

On va arrondir les variables de la solution du programme linéaire pour construire une couverture :

Couverture-Sommet-Pondéré-Approchée(G, w)

$C \leftarrow \emptyset$
calculer $\bar{x}$, solution optimale du programme linéaire relaxé
pour tout $v \in V$ **faire**
  si $\bar{x}(v) \geq 1/2$ **alors**
    $C \leftarrow C \cup \{v\}$
retourner C

Résultat formel

Théorème

Couverture-Sommet-Pondéré-Approchée *est un algorithme d'approximation 2 à temps polynomial pour le problème de la couverture de sommets de poids minimal.*

Preuve de la complexité en temps

Démonstration.

- Le nombre de variables est un polynôme de la taille de l'entrée ($|V|$).
- Idem pour le nombre de contraintes ($|V| + |E|$).
- La résolution du programme linéaire peut donc se faire en temps polynomial.
- La procédure d'arrondissement consiste en $|V|$ itérations.
- Couverture-Sommet-Pondéré-Approchée est bien un algorithme à temps polynomial.

$\square$

Preuve de la validité

Démonstration.

- Pour chaque arête $(u, v) \in E$, $x(u) + x(v) \geq 1$.
- $x(u)$ ou $x(v)$ vaut au moins $1/2$ et est sélectionné dans la couverture.
- Chaque arête sera donc couverte.

$\square$

Preuve du facteur d'approximation

Démonstration.

- Soit C^* la couverture optimale.
- Soit z^* la valeur de l'objectif optimale du programme linéaire relaxé : $z^* \leq w(C^*)$ (c'est une **borne inférieure** de la valeur de l'objectif optimale).
- On va prouver que $w(C) \leq 2z^*$ (et on aura donc $w(C) \leq 2w(C^*)$) :

$$z^* = \sum_{v \in V} w(v)\bar{x}(v) \geq \sum_{v \in C} w(v)\bar{x}(v)$$

- Comme chaque sommet v est conservé si $\bar{x}(v) \geq 1/2$:

$$z^* \geq \frac{1}{2} \sum_{v \in C} w(v) = \frac{1}{2} w(C)$$

$\square$

Optimalité du facteur d'approximation

S'agit-il d'un *facteur d'approximation optimal* (ou fin) ? Peut-on trouver un pire cas qui l'atteigne et si oui lequel ?

Optimalité du facteur d'approximation

S'agit-il d'un *facteur d'approximation optimal* (ou fin) ? Peut-on trouver un pire cas qui l'atteigne et si oui lequel ?
Deux sommets de poids unitaire et une arête. On pourrait avoir $\bar{x}(v_1) = \bar{x}(v_2) = 1/2$. On aurait alors une couverture de poids 2.

Plan

Problème des triangles d'un graphe

- Un graphe non-orienté $G = (V, E)$.
- Trois sommets u, v, w forment un triangle si les arêtes (u, v), (v, w) et (w, u) sont dans E.
- On souhaite déterminer un sous-ensemble $S \subseteq V$ de sommets, tel quel le graphe $G \setminus S$ ne contient aucun triangle.
- Le problème consiste à trouver un sous-ensemble S de taille minimale.

Solutions gloutonnes

Quel principe glouton pourrait être mis en place ?

Solutions gloutonnes

Quel principe glouton pourrait être mis en place ?

TRIANGLE-GRAPHE-GLOUTON1(G)

$S \leftarrow \emptyset$
$V' \leftarrow G.V$
tant que il reste un triangle (u, v, w) dans V' **faire**
  $V' \leftarrow V' \setminus \{u\}$
  $S \leftarrow S \cup \{u\}$
retourner $\mathcal{S}$

TRIANGLE-GRAPHE-GLOUTON2(G)

$S \leftarrow \emptyset$
$V' \leftarrow G.V$
tant que il reste un triangle (u, v, w) dans V' **faire**
  $V' \leftarrow V' \setminus \{u, v, w\}$
  $S \leftarrow S \cup \{u, v, w\}$
retourner $\mathcal{S}$

Pire cas pour les solutions gloutonnes

Quelle stratégie est la meilleure ?

Pire cas pour les solutions gloutonnes

Quelle stratégie est la meilleure ?

- Bien que Triangle-Graphe-Glouton1 semble plus parcimonieux, son pire cas est $(|V|-1)/2$: $(|V|-1)/2$ triangles qui partagent tous un sommet central.
- Triangle-Graphe-Glouton2 atteint le facteur d'approximation 3 avec un graphe composé seulement d'un triangle.

Formulation en programme entier : variables de décision

Quelles sont les variables de décision ?

Formulation en programme entier : variables de décision

Quelles sont les variables de décision ?

- $x_v \in \{0, 1\}$ pour tout $v \in V$.
- Si $x_v = 1$, le sommet v fait parti de la solution S.
- Sinon, le sommet v n'en fait pas parti.

Formulation en programme entier : fonction objectif

Quelle est la fonction objectif ?

Formulation en programme entier : fonction objectif

Quelle est la fonction objectif ?
On souhaite minimiser le nombre de sommets retournés :

$$\sum_{v \in V} x_v$$

Formulation en programme entier : contraintes

Quelles sont les contraintes ?

Formulation en programme entier : contraintes

Quelles sont les contraintes ?
On souhaite sélectionner au moins un sommet pour chaque triangle : $x_u + x_v + x_w \geq 1$ si $(u, v) \in E$, $(v, w) \in E$ et $(w, u) \in E$.

Formulation en programme entier : final

$$
\begin{array}{lll}
\text{minimiser} & \displaystyle\sum_{v \in V} x(v) & \\
\text{sous les contraintes} & x(u) + x(v) + x(w) \geq 1 & \text{si } (u, v) \in E \text{ et } (v, w) \in E \text{ et } (w, u) \in E \\
 & x(v) \in \{0, 1\} & \text{pour chaque } v \in V
\end{array}
$$

Relaxation en programme linéaire

$$
\begin{aligned}
\text{minimiser} \quad & \sum_{v \in V} x(v) \\
\text{sous les contraintes} \quad & x(u) + x(v) + x(w) \geq 1 \text{ si } (u, v) \in E \text{ et } (v, w) \in E \text{ et } (w, u) \in E \\
& x(v) \leq 1 \text{ pour chaque } v \in V \\
& x(v) \geq 0 \text{ pour chaque } v \in V
\end{aligned}
$$

Triangle-Graphe-Approchée

Comment arrondir les variables de la solution du programme linéaire pour construire une solution valide ?

Triangle-Graphe-Approchée

Comment arrondir les variables de la solution du programme linéaire pour construire une solution valide ?

- On veut au moins sélectionner un sommet pour chaque contrainte $x(u) + x(v) + x(w) \geq 1$.
- Quelle valeur seuil doit franchir $x(u)$ pour qu'on le conserve ?

Triangle-Graphe-Approchée

Comment arrondir les variables de la solution du programme linéaire pour construire une solution valide ?

- On veut au moins sélectionner un sommet pour chaque contrainte $x(u) + x(v) + x(w) \geq 1$.
- Quelle valeur seuil doit franchir $x(u)$ pour qu'on le conserve ?

Triangle-Graphe-Approchée(G)

$C \leftarrow \emptyset$
calculer $\bar{x}$, solution optimale du programme linéaire relaxé
pour tout $v \in V$ **faire**
  si $\bar{x}(v) \geq 1/3$ **alors**
    $C \leftarrow C \cup \{v\}$
retourner C

Complexité en temps

Quel est le nombre de variables et de contraintes ?

Complexité en temps

Quel est le nombre de variables et de contraintes ?

- Il y a $|V|$ variables et $O(|V| + |E|^3)$ contraintes.
- La résolution du programme linéaire peut donc se faire en temps polynomial.
- La procédure d'arrondissement consiste en $|V|$ itérations.

Validité

- Pour chaque triangle $(u, v) \in E$, $x(u) + x(v) + x(w) \geq 1$.
- $x(u)$, $x(v)$ ou $x(w)$ vaut au moins $1/3$ et est sélectionné dans la solution.
- Chaque triangle perdra donc un sommet.

Facteur d'approximation

- Soit S^* la solution optimale.
- Soit z^* la valeur de l'objectif optimale du programme linéaire relaxé : $z^* \leq w(S^*)$ (c'est une **borne inférieure** de la valeur de l'objectif optimale).
- On va prouver que $w(S) \leq 3z^*$ (et on aura donc $w(S) \leq 3w(S^*)$) :

$$z^* = \sum_{v \in V} w(v)\bar{x}(v) \geq \sum_{v \in S} w(v)\bar{x}(v)$$

- Comme chaque sommet v est conservé si $\bar{x}(v) \geq 1/3$:

$$z^* \geq \frac{1}{3} \sum_{v \in S} w(v) = \frac{1}{3} w(S)$$

Optimalité du facteur d'approximation

Le facteur d'approximation 3 est-il optimal ?

Optimalité du facteur d'approximation

Le facteur d'approximation 3 est-il optimal ?
Un graphe avec seulement un triangle. On pourrait avoir $\bar{x}(v_1) = \bar{x}(v_2) = \bar{x}(v_3) = 1/3$. On aurait une solution avec 3 sommets.

Résultat formel

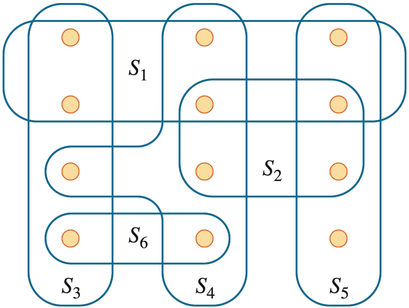
Théorème

TRIANGLE-GRAPHE-APPROCHÉE *est un algorithme d'approximation 3 à temps polynomial pour le problème des triangles d'un graphe.*

Plan

Problème de la couverture d'ensemble

- Un ensemble X et une famille $\mathcal{F}$ de sous-ensembles de X : $\forall S \in \mathcal{F}, S \subseteq X$.
- Chaque élément de X appartient à un sous-ensemble de $\mathcal{F}$: $X = \bigcup_{S \in \mathcal{F}} S$.
- Il faut déterminer quelle sous-famille $\mathcal{C} \subseteq \mathcal{F}$ permet de couvrir tous les éléments de X : $X = \bigcup_{S \in \mathcal{C}} S$.
- On cherche à minimiser le nombre des sous-ensembles sélectionnés $|\mathcal{C}|$.



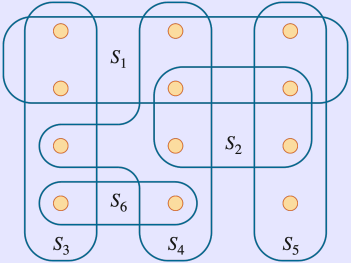
Fréquence de l'élément le plus fréquent

- On note f la fréquence de l'élément le plus fréquent :

$$f = \max_{x \in X} |\{S \in \mathcal{F} : x \in S\}|$$

Question

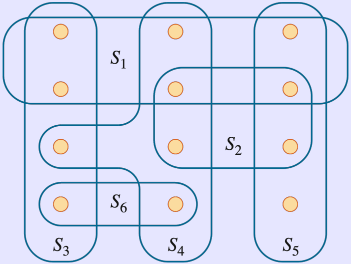
Que vaut f dans cet exemple ?



- 1
- 2
- 3
- 4

Question

Que vaut f dans cet exemple ?



- 1
- 2
- 3 ✓(tous dans $S_1 \cap S_2$)
- 4

Question

Le problème de la couverture d'ensemble généralise le problème de la couverture de sommets : une instance de ce dernier problème peut être convertie en instance pour le premier. Que vaut f, la fréquence de l'élément le plus fréquent, pour les instances issues de cette conversion ?

- 1
- 2
- 3
- 4

Question

Le problème de la couverture d'ensemble généralise le problème de la couverture de sommets : une instance de ce dernier problème peut être convertie en instance pour le premier. Que vaut f, la fréquence de l'élément le plus fréquent, pour les instances issues de cette conversion ?

- 1
- 2 ✓(arêtes $\rightarrow X$, sommets $\rightarrow \mathcal{F}$)
- 3
- 4

Formulation en programme entier

- La décision se porte sur chaque sous-ensemble que l'on conserve ou non. On définit donc x_S pour tout $S \in \mathcal{F}$.
- L'objectif est de minimiser le nombre de sous-ensembles choisis.
- On introduit une contrainte par élément de X : au moins un sous-ensemble doit le couvrir.

$$
\begin{array}{lll}
\text{minimiser} & \displaystyle\sum_{S \in \mathcal{F}} x_S & \\
\text{sous les contraintes} & \displaystyle\sum_{S \in \mathcal{F} : x \in S} x_S \geq 1 & \text{pour chaque } x \in X \\
& x(v) \in \{0, 1\} & \text{pour chaque } v \in V
\end{array}
$$

Couverture-Ensemble-Approchée

Comment arrondir les variables de la solution du programme linéaire relaxé pour construire une solution valide ?

COUVERTURE-ENSEMBLE-APPROCHÉE

Comment arrondir les variables de la solution du programme linéaire relaxé pour construire une solution valide ?

COUVERTURE-ENSEMBLE-APPROCHÉE$(X, \mathcal{F})$

$C \leftarrow \emptyset$
calculer $\bar{x}$, solution optimale du programme linéaire relaxé
pour tout $S \in \mathcal{F}$ **faire**
  si $\bar{x}_S \geq 1/f$ **alors**
    $C \leftarrow C \cup \{S\}$
retourner C

Résultat formel

Théorème

COUVERTURE-ENSEMBLE-APPROCHÉE *est un algorithme d’approximation f à temps polynomial pour le problème de la couverture d’ensemble.*

Preuve de la complexité en temps

Démonstration.

- Il y a $|\mathcal{F}|$ variables et $|X|$ contraintes.
- La résolution du programme linéaire peut donc se faire en temps polynomial.
- La procédure d'arrondissement consiste en $|X|$ itérations.

$\square$

Preuve de la validité et du facteur d'approximation

Démonstration.

- Pour chaque élément $x \in X$, il y a au moins un sous-ensemble choisi grâce aux contraintes et au choix de l'arrondi.
- L'arrondissement augmente la valeur de l'objectif d'un facteur f au maximum.
- Il s'agit donc bien d'un algorithme d'approximation f.

□

Optimalité du facteur d'approximation

Le facteur d'approximation f est-il optimal ?

Optimalité du facteur d'approximation

Le facteur d'approximation f est-il optimal ?
n éléments et f sous-ensembles C qui contiennent tous les éléments.

Plan

Résumé

- On peut concevoir des algorithmes d'approximation en temps polynomial en passant par la modélisation en programmation entier puis la résolution de programmation linéaire.
- Cette modélisation nécessite de définir les variables de décision, l'objectif et les contraintes.
- L'approche s'appuie sur une relaxation du programme entier en programme linéaire, la résolution de ce dernier, puis l'arrondissement des valeurs obtenues.
- On doit ensuite en analyser : sa complexité en temps, sa validité et son facteur d'approximation (borne puis optimalité).

Prochaines échéances

- Début du tournoi le 19/11.
- Premier rendu intermédiaire du tournoi le 2/12.
- Second rendu intermédiaire du tournoi le 9/12.
- Épreuve de TP le 17/12.
- Rendu final pour le tournoi le 8/1.
- Épreuve sur table et restitution du tournoi le 9/1.