Formulação de Programação Linear Inteira para o Problema de Particionamento em Conjuntos Convexos

Teobaldo L. Bulhões Júnior[a]

[a] *Instituto de Computação, Universidade Federal Fluminense, Niterói, RJ, Brazil*
tbulhoes@ic.uff.br

Resumo

Dado um grafo $G = (V, E)$, um conjunto $S \subseteq V$ é dito convexo se, para quaisquer dois vértices $u, v \in S$, todo vértice w que está contido em algum caminho mínimo entre u e v também pertence a S. Este trabalho trata do Problema de Particionamento em Conjuntos Convexos, o qual consiste em particionar o conjunto de vértices de um grafo G em um número mínimo de conjuntos convexos. A solução com apenas um conjunto convexo não é válida, do contrário o problema seria trivial. É apresentada uma formulação de programação linear inteira para o problema, bem como experimentos preliminares com grafos aleatórios.

1. Introdução

Seja $G = (V, E)$ um grafo simples. Um conjunto $S \subseteq V$ é dito convexo se, para quaisquer dois vértices $u, v \in S$, todo vértice w que está contido em algum caminho mínimo entre u e v também pertence a S. Este trabalho trata do Problema de Particionamento em Conjuntos Convexos (PPCC), o qual consiste em particionar o conjunto de vértices de um grafo G em um número mínimo de conjuntos convexos. A solução com apenas um conjunto convexo não é válida, do contrário o problema seria trivial.

Um grafo é dito p-convexo se ele pode ser particionado em exatamente p conjuntos convexos. Em [1], é mostrado que é NP-Completo decidir se um grafo simples G é p-convexo, para um $p \geq 2$ fixo. Além disso, dentre outros resultados teóricos, os autores mostraram que: (i) todo grafo cordal é

Email: `tbulhoes@ci.ufpb.br` (Teobaldo L. Bulhões Júnior)

June 29, 2016

p-convexo, para qualquer valor de p; (ii) é possível verificar em tempo linear se um cografo é p-convexo, para qualquer valor de p.

Este trabalho propõe uma formulação de programação linear inteira para o PPCC, a qual é apresentada na Seção 2. A Seção 3 apresenta alguns resultados obtidos em grafos aleatórios.

2. Formulação Matemática

Seja $G = (V, E)$ o grafo de entrada. Para cada par de vértices $i, j \in V$ com $i < j$, seja $P(i,j) = \{w \in V \setminus \{i,j\} \mid w \text{ pertence a algum caminho mínimo entre } i \text{ e } j\}$. Observe que os conjuntos $P(i,j)$ podem ser computados em tempo $O(|V|^3)$ através do algoritmo de Floyd-Warshall. Ainda para os vértices i e j, defina uma variável binária x_{ij}, a qual assume o valor 1 caso eles pertençam ao mesmo conjunto convexo da solução. Além disso, defina uma variável binária y_j para cada vértice $j \in V$, a qual assume o valor 1 caso o vértice j seja o representante do seu conjunto convexo na solução. O representante de um conjunto convexo C é o vértice de C que possui o menor índice. As variáveis y_j servem apenas para contar o número de conjuntos convexos da solução. Considere a seguinte formulação, denotada por $\mathcal{F}_1$.

$$(\mathcal{F}_1) \quad \min \sum_{j \in V} y_j \tag{1}$$

$$\begin{aligned}
\text{s.a.} \quad x_{wi} &\geq x_{ij}, & \forall i, j \in V, i < j, \forall w \in P(i,j) \quad & (2)\\
x_{wj} &\geq x_{ij}, & \forall i, j \in V, i < j, \forall w \in P(i,j) \quad & (3)\\
y_j &\leq 1 - x_{ij}, & \forall i, j \in V, i < j \quad & (4)\\
y_j &\geq 1 - \sum_{i<j} x_{ij}, & \forall j \in V, \quad & (5)\\
\sum_{j \in V} y_j &\geq 2, & & (6)\\
x_{ij} &\in \{0,1\}, & \forall i, j \in V, i < j, \quad & (7)\\
y_j &\in \{0,1\}, & \forall j \in V. \quad & (8)
\end{aligned}$$

A função objetivo (1) minimiza o número de conjuntos convexos da solução. As restrições (2) e (3) determinam que, se os vértices i e j estão em um mesmo conjunto convexo C da solução (i.e., $x_{ij} = 1$), então todo vértice w que pertence a algum caminho mínimo entre i e j também pertence a C — ou seja, $x_{iw} = x_{jw} = 1$. As restrições (4) e (5) determinam quando um vértice j é o representante de seu conjunto convexo. Se algum vértice $i \in V$ tal que $i < j$ pertence ao mesmo conjunto convexo que j, então j não é o representante de seu conjunto convexo (restrições (4)). Por outro lado, se nenhum vértice $i \in V$ tal que $i < j$ pertence ao mesmo conjunto convexo que j, então j é o representante de seu conjunto convexo (restrições (5)). A restrição (6) proíbe a solução trivial, que é a solução com apenas um conjunto convexo. Por fim, (7) e (8) definem os domínios das variáveis.

A formulação ($\mathcal{F}_1$) produz uma particionamento em conjuntos convexos do grafo G, sendo cada partição definida por uma componente conexa do grafo solução $\bar{G} = (\overline{V}, \overline{E})$, onde $\overline{V} = V$ e $\overline{E} = \{(i,j) \mid i, j \in V, i < j, x_{ij} = 1\}$. Porém, essa formulação não é correta, pois nela não há a correspondência entre representante e conjunto convexo. Isto é, em uma solução da formulação $\mathcal{F}_1$, é possível que dois vértices que pertencem a um mesmo conjunto convexo sejam representantes de conjunto. A Figura 1 mostra um exemplo de um solução com apenas um conjunto convexo, mas com duas variáveis y_j iguais a 1.

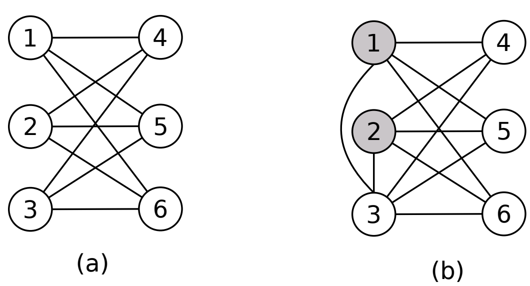

Figura 1: (a) Instância do problema. (b) Solução obtida por $\mathcal{F}_1$ com apenas um conjunto convexo, mas com dois representantes (destacados em cinza).

Devemos garantir então que cada componente conexa do grafo solução possua exatamente um representante. Uma maneira é transformar o grafo solução em uma união disjunta de cliques através da adição das desigualdades triangulares:

$$-x_{ij} + x_{ik} + x_{jk} \leq 1, \qquad \forall i,j,k \in V, i < j < k, \tag{9}$$

$$x_{ij} - x_{ik} + x_{jk} \leq 1, \qquad \forall i,j,k \in V, i < j < k, \tag{10}$$

$$x_{ij} + x_{ik} - x_{jk} \leq 1, \qquad \forall i,j,k \in V, i < j < k, \tag{11}$$

Observe que as desigualdades (9)-(11) garantem que a solução não contém P_3 como subgrafo induzido, ou seja, cada componente do grafo solução é uma clique. Claramente, uma clique possui apenas um representante. A Figura 2 exibe a solução obtida por $\mathcal{F}_1$ para a instância da Figura 1(a) após a adição das desigualdades triangulares.

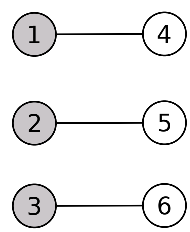

Figura 2: Nova solução obtida por $\mathcal{F}_1$ para a instância ilustrada na Figura 1(a) após a adição das desigualdades triangulares.

Outra forma de garantir que os representantes pertencem a componentes distintas do grafo solução é através da adição das desigualdades de árvore. Seja T uma árvore tal que $V(T) \subseteq V(G)$. Então:

$$\sum_{(i,j) \in E(T)} x_{ij} \leq |V(T)| - \sum_{j \in V(T)} y_j \tag{12}$$

Observe que, como qualquer aresta (i,j) pode existir na solução (independente dela existir no grafo de entrada ou não), as arestas da árvore T não estão necessariamente contidas no grafo de entrada.

3. Experimentos

Esta seção reporta alguns resultados obtidos com a formulação descrita na Seção 2. A forma adotada para garantir que os representantes estão desconectados no grafo solução foi através das desigualdades triangulares, que são mais fáceis de serem implementadas. A formulação foi implementada em C++, utilizando o CPLEX 12.4 como resolvedor de problemas de programação inteira.

3.1. Geração de instâncias

As instâncias geradas seguem o modelo aleatório de Gilbert [2]. Nesse modelo, obtemos um grafo $G(n, p)$ com n vértices no qual cada possível aresta existe com uma probabilidade p. Dessa forma, o parâmetro p pode ser visto como a densidade esperada para o grafo $G(n, p)$.

3.2. Impacto da densidade

Esta seção discute brevemente o impacto da densidade no tamanho mínimo de um particionamento em conjuntos convexos de um grafo G. Para isso, foram gerados 10 grafos $G(30, p)$, com $p \in \{0.1, 0.2, \ldots, 1\}$. Vale ressaltar que a pesquisa em questão é muito incipiente. O objetivo desta seção não é apresentar conclusões sólidas a respeito do impacto da densidade na dificuldade das instâncias, mas sim gerar uma discussão em sala sobre o assunto.

A Figura 3 apresenta os custos ótimos das 10 instâncias em questão. Se considerarmos a métrica de dificuldade como sendo "número mínimo de conjuntos convexos", podemos perceber que os grafos de densidade moderada (em torno de 0.5) são os mais difíceis. Também podemos perceber que a dificuldade da instância diminui a medida que a densidade se aproxima de 1.

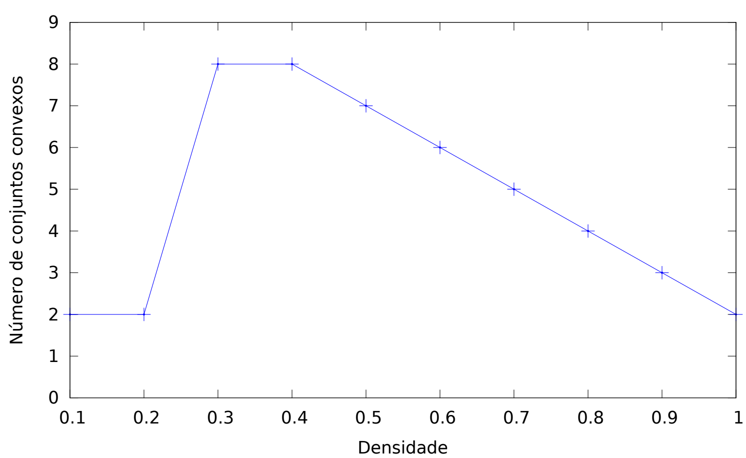

Figura 3: Custos ótimos para 10 instâncias com 30 vértices e densidades $d \in \{0.1, 0.2, \ldots, 1\}$

Referências

[1] D. Artigas, S. Dantas, M. Dourado, J. Szwarcfiter, Partitioning a graph into convex sets, Discrete Mathematics 311 (17) (2011) 1968 – 1977.

[2] E. N. Gilbert, Random graphs, in: Annals of Mathematical Statistics, vol. 3, 1959, pp. 1141–1144.