Introdução à emparelhamento em grafos bipartidos

Marcio T. I. Oshiro

Universidade de São Paulo
Instituto de Matemática e Estatística
Departamento de Ciência da Computação
MaratonIME

29 de outubro de 2015

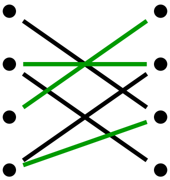
Notação

- $G = (A, B, E)$ — grafo bipartido
- $n = |A| + |B|$ — número de vértices
- $m = |E|$ — número de arestas
- $E \triangle E' = (E \cup E') \setminus (E \cap E')$ — diferença simétrica

em alguns casos, por conveniência, trataremos (sub)grafos por seus conjuntos de arestas

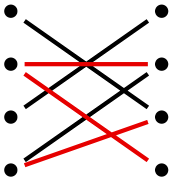
Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes



Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes

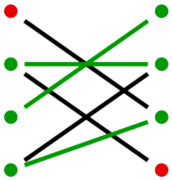


Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes

dado um emparelhamento M

- vértices são: emparelhados ou livres

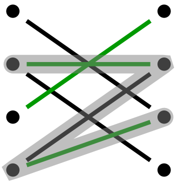


Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes

dado um emparelhamento M

- vértices são: emparelhados ou livres
- **caminho alternante**: alterna entre arestas de M e $E \setminus M$

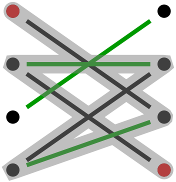


Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes

dado um emparelhamento M

- vértices são: emparelhados ou livres
- **caminho alternante**: alterna entre arestas de M e $E \setminus M$
- **caminho de aumento**: caminho alternante que começa e termina em vértices livres (sempre ímpar)

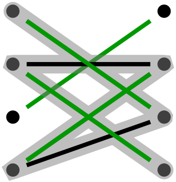


Emparelhamento

- **emparelhamento**: conjunto de arestas não adjacentes

dado um emparelhamento M

- vértices são: emparelhados ou livres
- **caminho alternante**: alterna entre arestas de M e $E \setminus M$
- **caminho de aumento**: caminho alternante que começa e termina em vértices livres (sempre ímpar)



Emparelhamento de cardinalidade máxima

- emparelhamento M é máximo se $|M|$ é máximo

Emparelhamento de cardinalidade máxima

- emparelhamento M é máximo se $|M|$ é máximo

Teorema (Berge)

Um emparelhamento M de um grafo G é máximo se e somente se não exite caminho de aumento em G.

Emparelhamento de cardinalidade máxima

- emparelhamento M é máximo se $|M|$ é máximo

Teorema (Berge)

Um emparelhamento M de um grafo G é máximo se e somente se não exite caminho de aumento em G.

Demonstração.

$(\Rightarrow)$ Trivial.
$(\Leftarrow)$ Seja M^* um emparelhamento máximo.

- $M^* \cup M$ consiste de circuitos alternantes e caminhos alternantes (pares).
- $|M| = |M^*|$

Logo, M também é máximo. $\square$

Emparelhamento de cardinalidade máxima

Algoritmo

empMax(G)

$M \leftarrow \emptyset$

enquanto existe caminho de aumento P faça

$M \leftarrow M \triangle P$

devolva M

Emparelhamento de cardinalidade máxima

Implementação

- M é armazenado em vetores (*arrays*) M_A e M_B

$$ab \in M, a \in A, b \in B \Rightarrow M_A[a] = b \text{ e } M_B[b] = a$$

- caminhos de aumento são encontrados por *dfs* a partir de vértices de A livres

temAumento(a) $\quad \triangleright$ G, M_A, M_B implícitos na chamada

    para cada vizinho $b \in B$ de a ainda não visitado faça
        se b é livre ou temAumento($M_B[b]$) $=$ `sim` então
            $M_A[a] \leftarrow b$
            $M_B[b] \leftarrow a$
            devolva `sim`
    devolva `não`

Emparelhamento de cardinalidade máxima

Implementação

- não precisa marcar que visitou vértices de A
- não precisa buscar em B também
- melhor fazer a busca em B se $|B| < |A|$
- tempo de pior caso: $\mathrm{O}(nm)$
- dá para resolver em $\mathrm{O}(\sqrt{n}m)$ (Hopcroft-Karp)

Emparelhamento de cardinalidade máxima

Aplicações

$C \subseteq A \cup B$ é uma **cobertura por vértices** se $\{a, b\} \cap C \neq \emptyset$, para toda aresta $ab \in E$.

Emparelhamento de cardinalidade máxima

Aplicações

$C \subseteq A \cup B$ é uma **cobertura por vértices** se $\{a, b\} \cap C \neq \emptyset$, para toda aresta $ab \in E$.

Teorema (Kőnig)

$$\left.\begin{array}{r} G \text{ bipartido} \\ M^* \text{ emparelhamento máximo} \\ C^* \text{ cobertura mínima} \end{array}\right\} \iff |M^*| = |C^*|$$

Demonstração.

Claramente $|M^*| \leq |C^*|$.
Seja C inicialmente vazio. Para cada $ab \in M^*$, com $a \in A$ e $b \in B$, se a tem vizinho livre, insere a em C. Caso contrário insere b em C. No final, C é uma cobertura de G.
Logo, $|M^*| = |C| \geq |C^*|$. $\square$

Emparelhamento máximo de peso máximo

- $G = (A, B, E)$ — grafo bipartido completo
- $w \colon E \to \mathbb{R}$ — peso das arestas
- $n = |A| = |B|$ — número de vértices
- $m = |E|$ — número de arestas

Note que sempre existe um emparelhamento perfeito ($|M| = n$).

Primal

$$
\begin{aligned}
\text{Maximizar} \quad & \sum_{ij \in E} w(ij) x_{ij} \\
\text{sujeito a} \quad & \sum_{j \in B} x_{ij} = 1, \forall i \in A \\
& \sum_{i \in A} x_{ij} = 1, \forall j \in B \\
& x_{ij} \geq 0, \forall ij \in E
\end{aligned}
$$

x é vetor de variáveis

$x_{ij} = 1 \Rightarrow ij \in M$

Emparelhamento máximo de peso máximo

Programação linear

Primal

$$\text{Maximizar} \quad \sum_{ij \in E} w(ij) x_{ij}$$

$$\text{sujeito a} \quad \sum_{j \in B} x_{ij} = 1, \ \forall i \in A$$

$$\sum_{i \in A} x_{ij} = 1, \ \forall j \in B$$

$$x_{ij} \geq 0, \ \forall ij \in E$$

Dual

$$\text{Minimizar} \quad \sum_{i \in A} y_i + \sum_{j \in B} z_j$$

$$\text{sujeito a} \quad y_i + z_j \geq w(ij), \ \forall ij \in E$$

x, y e z são vetores de variáveis

$x_{ij} = 1 \Rightarrow ij \in M$

Emparelhamento máximo de peso máximo

- y e z são chamadas de variáveis duais
- (y, z) é **viável** se $y_i + z_j \geq w(ij)$, $\forall i \in A$ e $\forall j \in B$

A teoria de programação linear garante que

$$\sum_{ij \in E} w(ij) x_{ij} \leq \sum_{i \in A} y_i + \sum_{j \in B} z_j,$$

para qualquer x viável no primal e (y, z) viável no dual. Mais ainda, vale a igualdade se e somente se x e (y, z) são soluções ótimas do primal e do dual, respectivamente.

Emparelhamento máximo de peso máximo

dada (y, z) viável no dual

- $G_{yz} = (A, B, E_{yz})$, onde $E_{yz} = \{ij \in E\colon y_i + z_j = w(ij)\}$

Teorema ("folgas complementares")

$$\left.\begin{array}{r} G \text{ bipartido} \\ (y,z) \text{ viável do dual} \\ M \text{ emp. perf. de } G_{yz} \end{array}\right\} \Rightarrow \begin{array}{c} M \text{ é emparelhamento perfeito de} \\ \text{peso máximo de G} \end{array}$$

Demonstração.

Seja M^* um emparelhamento perfeito de peso máximo de G.

$$\sum_{ij \in M^*} w(ij) \leq \sum_{ij \in M^*} y_i + z_j = \sum_{ij \in M} y_i + z_j = \sum_{ij \in M} w(ij). \qquad \square$$

Emparelhamento máximo de peso máximo

dada (y, z) viável no dual

- $G_{yz} = (A, B, E_{yz})$, onde $E_{yz} = \{ij \in E \colon y_i + z_j = w(ij)\}$

Teorema ("folgas complementares")

$$\left.\begin{array}{r} G \text{ bipartido} \\ (y,z) \text{ viável do dual} \\ M \text{ emp. perf. de } G_{yz} \end{array}\right\} \Rightarrow \begin{array}{c} M \text{ é emparelhamento perfeito de} \\ \text{peso máximo de G} \end{array}$$

Demonstração.

Seja M^* um emparelhamento perfeito de peso máximo de G.

$$\sum_{ij \in M^*} w(ij) \leq \sum_{ij \in M^*} y_i + z_j = \sum_{ij \in M} y_i + z_j = \sum_{ij \in M} w(ij). \qquad \square$$

Baseado no teorema anterior, o método húngaro

- começa com (y, z) viável e emparelhamento M de G_{yz}
- enquanto M não é perfeito, muda (y, z) adequadamente, adicionando mais arestas, e tenta aumentar M

Inicialmente

- $M = \emptyset$
- $y_i = \max\limits_{j \in B} w(ij)$, para todo $i \in A$
- $z_j = 0$, para todo $j \in B$

Emparelhamento máximo de peso máximo

Método húngaro O(n^4)

Algoritmo temAumento(i) deve considerar G_{yz}.
Se encontrar caminho de aumento, devolve ∞.
Caso contrário, devolve $\delta < \infty$.

temAumento(i) $\triangleright$ G, M_A, M_B implícitos na chamada
- $\delta \leftarrow \infty$
- para cada vizinho $j \in B$ de i ainda não visitado faça
  - se $y_i + z_j = w(ij)$ então
    - se j é livre ou temAumento$(M_B[j]) = \infty$ então
    - $M_A[i] \leftarrow j$, $M_B[j] \leftarrow i$
    - devolva ∞
  - senão
    - $\delta \leftarrow \min(\delta, y_i + z_j - w(ij))$
- devolva δ

Emparelhamento máximo de peso máximo

Método húngaro O(n^4)

hungaroN4(G)
 inicializa M_A, M_B e (y, z)
 enquanto M não é perfeito faça
  seja $i \in A$ livre
  $\delta \leftarrow$ temAumento(i)
  se $\delta < \infty$ então
   para cada $i \in A$ visitado faça
    $y_i \leftarrow y_i - \delta$
   para cada $j \in B$ visitado faça
    $z_j \leftarrow z_j + \delta$
 devolva M_A, M_B

Emparelhamento máximo de peso máximo

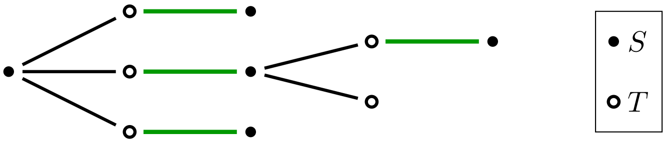
Método húngaro $O(n^3)$

É possível implementar com tempo $O(n^3)$.

Principais diferenças com relação à implementação em $O(n^4)$:

- atualização de (y, z)
- aproveitamento de buscas anteriores no caso de não encontrar caminho de aumento, mantendo vértices da árvore alternante ($S \subseteq A$ e $T \subseteq B$ são os conjuntos de vértice da árvore)

Árvore alternante:



Emparelhamento máximo de peso máximo

Método húngaro O(n^3): atualização das variáveis duais

S e T também corresponde aos vértices visitados.

Para todo $j \in B \setminus T$, $d_j = \min_{i \in S}(y_i + z_j - w(ij))$

atualiza-dual(y, z, d, S, T)

$\delta \leftarrow \min_{j \in B \setminus T} d_j$ $\quad \triangleright\ \delta = \min_{i \in S,\, j \in B \setminus T}(y_i + z_j - w(ij))$

para todo $i \in S$ faça $y_i \leftarrow y_i - \delta$

para todo $j \in B$ faça

se $j \in T$ então

$z_j \leftarrow z_j + \delta$

senão

$d_j \leftarrow d_j - \delta$

devolva y, z e d

Emparelhamento máximo de peso máximo

Método húngaro O(n^3)

húngaro(G, w)
- inicializa (y, z) e M
- enquanto M não é perfeito em G_{yz} faça
  - seja $i \in A$ livre
  - $S \leftarrow \{i\}$, $T \leftarrow \emptyset$
  - atualize d
  - enquanto não encontrar caminho de aumento faça
    - se $N_{yz}(S) = T$ então ▷ não dá para aumentar a árvore alternante
      - $y, z, d \leftarrow$atualiza-dual(y, z, d, S, T)
    - seja $j \in N_{yz}(S) \setminus T$
    - se j é livre então ▷ achou caminho de aumento
      - seja P o caminho de i a j
      - $M \leftarrow M \triangle P$
    - senão
      - j está emparelhado com algum $i' \in A \setminus S$
      - $S \leftarrow S \cup \{i'\}$, $T \leftarrow T\,\{j\}$
      - atualize d
- devolva M

Emparelhamento máximo de peso máximo

Método húngaro $O(n^3)$

- laço externo é executado no máximo n vezes
- atualização de d consome tempo $O(n)$
- laço interno termina em no máximo n iterações
- complexidade de tempo do algoritmo é $O(n^3)$

Emparelhamento máximo de peso máximo

- e se $|A| \neq |B|$?

- e se G não é completo?

- e se, em vez de maximizar o peso, quero minimizar o peso?

Emparelhamento máximo de peso máximo

- e se $|A| \neq |B|$?
  - insere novos vértices até que $|A| = |B|$

- e se G não é completo?

- e se, em vez de maximizar o peso, quero minimizar o peso?

Emparelhamento máximo de peso máximo

- e se $|A| \neq |B|$?
  - insere novos vértices até que $|A| = |B|$
- e se G não é completo?
  - insere as arestas que faltam com peso 0
- e se, em vez de maximizar o peso, quero minimizar o peso?

Emparelhamento máximo de peso máximo

- e se $|A| \neq |B|$?
  - insere novos vértices até que $|A| = |B|$
- e se G não é completo?
  - insere as arestas que faltam com peso 0
- e se, em vez de maximizar o peso, quero minimizar o peso?
  - troque os pesos para $w' = -w$

Emparelhamento máximo de peso máximo

- e se $|A| \neq |B|$?
  - insere novos vértices até que $|A| = |B|$
- e se G não é completo?
  - insere as arestas que faltam com peso 0
- e se, em vez de maximizar o peso, quero minimizar o peso?
  - troque os pesos para $w' = -w$

Uma referência mais formal:
`http://www.ime.usp.br/~mksilva/cursos/2015-2/mac0325-5781/aulas/10-06-o-metodo-hungaro/`