Algoritmos de Aproximação

Segundo Semestre de 2012

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que
maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$
sujeito a $x_i x_i = 1$ para cada i em V.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$

sujeito a $x_i x_i = 1$ para cada i em V.

Se x é solução viável do programa, $x_i \in \{-1, 1\}$ para todo i.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$

sujeito a $\quad x_i x_i = 1 \quad$ para cada i em V.

Se x é solução viável do programa, $x_i \in \{-1, 1\}$ para todo i.

Tome $S = \{i \ : \ x_i = 1\}$.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$

sujeito a $x_i x_i = 1$ para cada i em V.

Se x é solução viável do programa, $x_i \in \{-1, 1\}$ para todo i.

Tome $S = \{i \ : \ x_i = 1\}$.

Se $ij \in \delta(S)$ então $1 - x_i x_j = 2$, senão $1 - x_i x_j = 0$, logo $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j) = w(\delta(S))$.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que
maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$
sujeito a $x_i x_i = 1$ para cada i em V.

Se x é solução viável do programa, $x_i \in \{-1, 1\}$ para todo i.

Tome $S = \{i \ : \ x_i = 1\}$.

Se $ij \in \delta(S)$ então $1 - x_i x_j = 2$, senão $1 - x_i x_j = 0$,
logo $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j) = w(\delta(S))$.

Se $\delta(S)$ é corte, seja $x_i = 1$ se $i \in S$,
e $x_i = -1$ caso contrário.

Programação Semidefinida

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação alternativa: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j)$

sujeito a $\quad x_i x_i = 1 \quad$ para cada i em V.

Se x é solução viável do programa, $x_i \in \{-1, 1\}$ para todo i.

Tome $S = \{i \ : \ x_i = 1\}$.

Se $ij \in \delta(S)$ então $1 - x_i x_j = 2$, senão $1 - x_i x_j = 0$,
logo $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j) = w(\delta(S))$.

Se $\delta(S)$ é corte, seja $x_i = 1$ se $i \in S$,
e $x_i = -1$ caso contrário. Tal x é viável de peso $w(\delta(S))$.

Relaxação vetorial

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação quadrática: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - x_i x_j)$

sujeito a $x_i x_i = 1$ para cada i em V.

($w_{ij} = 0$ se $ij \notin E$)

Relaxação vetorial

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação quadrática: encontrar vetor x que

maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - x_i x_j)$

sujeito a $x_i x_i = 1$ para cada i em V.

($w_{ij} = 0$ se $ij \notin E$)

Relaxação vetorial:

encontrar matrix real Y $n \times n$, onde $n = |V|$, que

maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - (YY^T)_{ij})$

sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Relaxação vetorial

MAX CUT: Dados $G = (V, E)$ e pesos $w_e \geq 0$ para cada e em E, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação quadrática: encontrar vetor x que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - x_i x_j)$
sujeito a $x_i x_i = 1$ para cada i em V.

($w_{ij} = 0$ se $ij \notin E$)

Relaxação vetorial:
encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

As incógnitas são os vetores: as linhas da matriz Y.

Relaxação vetorial

MAX CUT: Dados G completo e pesos $w_{ij} \geq 0$ para cada ij, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação quadrática: encontrar vetor x que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - x_i x_j)$
sujeito a $x_i x_i = 1$ para cada i em V.

Relaxação vetorial:
encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Relaxação vetorial

MAX CUT: Dados G completo e pesos $w_{ij} \geq 0$ para cada ij, encontrar corte $\delta(X)$ em G de peso máximo.

Formulação quadrática: encontrar vetor x que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - x_i x_j)$
sujeito a $x_i x_i = 1$ para cada i em V.

Relaxação vetorial:
encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

É relaxação: solução viável x da formulação quadrática corrresponde a solução viável do programa vetorial.
(Tome Y com uma única coluna não nula, dada pelo x.)

Relaxação vetorial

Encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Relaxação vetorial

Encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Algumas questões:

- Mas como resolver um programa vetorial?

Relaxação vetorial

Encontrar matrix real Y $n \times n$, onde $n = |V|$, que
maximize $\frac{1}{2}\sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Algumas questões:

- Mas como resolver um programa vetorial?
- Aliás, tal programa tem solução?

Relaxação vetorial

Encontrar matrix real Y $n \times n$, onde $n = |V|$, que maximize $\frac{1}{2} \sum_{ij} w_{ij}(1 - (YY^T)_{ij})$
sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Algumas questões:

- Mas como resolver um programa vetorial?
- Aliás, tal programa tem solução?
- E o que fazer com a solução depois?

Tal programa tem solução?

Dados G e w, encontrar matrix real Y que

maximize $\frac{1}{2}\sum_{ij} w_{ij}(1-(YY^T)_{ij})$

sujeito a $(YY^T)_{ii}=1$ para cada i em V.

Tal programa tem solução?

Dados G e w, encontrar matrix real Y que

maximize $\frac{1}{2}\sum_{ij} w_{ij}(1-(YY^T)_{ij})$

sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Seja $V = V(G)$ e $n = |V|$.

Conjunto das soluções viáveis é

subconjunto de $\mathbb{R}^{n^2}$ fechado e limitado.

Tal programa tem solução?

Dados G e w, encontrar matrix real Y que

maximize $\frac{1}{2}\sum_{ij} w_{ij}(1 - (YY^T)_{ij})$

sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Seja $V = V(G)$ e $n = |V|$.

Conjunto das soluções viáveis é subconjunto de $\mathbb{R}^{n^2}$ fechado e limitado.

Fechado pois é interseção de hiperplanos.
($(YY^T)_{ii} = 1$ para cada i)

Tal programa tem solução?

Dados G e w, encontrar matrix real Y que

maximize $\frac{1}{2}\sum_{ij} w_{ij}(1-(YY^T)_{ij})$

sujeito a $(YY^T)_{ii} = 1$ para cada i em V.

Seja $V = V(G)$ e $n = |V|$.

Conjunto das soluções viáveis é subconjunto de $\mathbb{R}^{n^2}$ fechado e limitado.

Fechado pois é interseção de hiperplanos.
($(YY^T)_{ii} = 1$ para cada i)

Limitado pois $||Y||^2 = \sum_i Y_{i*}Yi* = \sum_i (YY^T)_{ii} = n$.
(Y_{i*}: linha i de Y)

O que fazer com sua solução?

Algoritmo de Goemans e Williamson:

MAXCUT-GW (G, w)

1 seja $\hat{Y}$ uma solução ótima do programa vetorial

2 $s \leftarrow$ RANDESFERA$(V(G))$

3 $S \leftarrow \{i \in V(G) \; : \; s\hat{Y}_{i*} > 0\}$

4 **devolva** $\delta(S)$

O que fazer com sua solução?

Algoritmo de Goemans e Williamson:

MAXCUT-GW (G, w)

1 seja $\hat{Y}$ uma solução ótima do programa vetorial
2 $s \leftarrow$ RANDESFERA$(V(G))$
3 $S \leftarrow \{i \in V(G) \ : \ s\hat{Y}_{i*} > 0\}$
4 **devolva** $\delta(S)$

RANDESFERA: devolve um r indexado por $V(G)$ tq $||r|| = 1$
(ou seja, um ponto na esfera unitária)
com probabilidade uniforme.

O que fazer com sua solução?

Algoritmo de Goemans e Williamson:

MAXCUT-GW (G, w)

1 seja $\hat{Y}$ uma solução ótima do programa vetorial
2 $s \leftarrow$ RANDESFERA$(V(G))$
3 $S \leftarrow \{i \in V(G) \ : \ s\hat{Y}_{i*} > 0\}$
4 **devolva** $\delta(S)$

RANDESFERA: devolve um r indexado por $V(G)$ tq $||r|| = 1$ (ou seja, um ponto na esfera unitária) com probabilidade uniforme.

Cada linha da matriz Y é um ponto desta mesma esfera.

Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\, ij \in \delta(S) \,] \;\geq\; \frac{1}{\pi} \arccos\left((\hat{Y}\hat{Y}^T)_{ij} \right)$.

Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\, ij \in \delta(S) \,] \;\geq\; \frac{1}{\pi} \arccos\left((\hat{Y}\hat{Y}^T)_{ij} \right)$.

Prova: Usemos $\hat{y}$ e $\hat{z}$ para denotar $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$. Então

Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\,ij \in \delta(S)\,] \;\geq\; \frac{1}{\pi}\,\mathrm{arccos}\,((\hat{Y}\hat{Y}^T)_{ij})$.

Prova: Usemos $\hat{y}$ e $\hat{z}$ para denotar $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$. Então $\Pr[\,ij \in \delta(S)\,] \;=\; \Pr[\,s\hat{y} > 0$ e $s\hat{z} \leq 0\,] + \Pr[\,s\hat{y} \leq 0$ e $s\hat{z} > 0\,]$.

Termo vermelho: probabilidade de s estar na esfera unitária na intersecção dos semi-espaços $\{r : r\hat{y} > 0\}$ e $\{r : r\hat{z} \leq 0\}$.

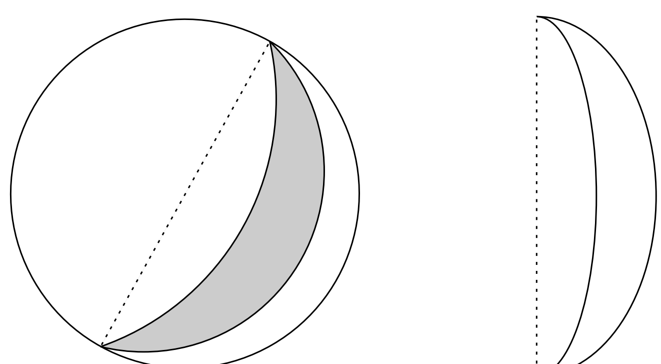
Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\,ij \in \delta(S)\,] \;\geq\; \frac{1}{\pi}\,\mathrm{arccos}\,((\hat{Y}\hat{Y}^T)_{ij})$.

Prova: Usemos $\hat{y}$ e $\hat{z}$ para denotar $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$. Então
$\Pr[\,ij \in \delta(S)\,] \;=\; \Pr[\,s\hat{y} > 0$ e $s\hat{z} \leq 0\,] + \Pr[\,s\hat{y} \leq 0$ e $s\hat{z} > 0\,]$.

Termo vermelho: probabilidade de s estar na esfera unitária na intersecção dos semi-espaços $\{r : r\hat{y} > 0\}$ e $\{r : r\hat{z} \leq 0\}$.

Tal região é a "fatia" da esfera de ângulo igual ao ângulo entre os vetores $\hat{y}$ e $\hat{z}$, que denotamos por Θ.



Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\,ij \in \delta(S)\,] \;=\; \frac{1}{\pi}\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right)$.

Prova: $\hat{y}$ e $\hat{z}$ denotam $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$ e
$\Pr[\,ij \in \delta(S)\,] \;=\; \Pr[\,s\hat{y} > 0$ e $s\hat{z} \leq 0\,] + \Pr[\,s\hat{y} \leq 0$ e $s\hat{z} > 0\,]$.

Termo vermelho: probabilidade de s estar na esfera unitária na intersecção dos semi-espaços $\{r : r\hat{y} > 0\}$ e $\{r : r\hat{z} \leq 0\}$.

Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\,ij \in \delta(S)\,] \;=\; \frac{1}{\pi}\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right)$.

Prova: $\hat{y}$ e $\hat{z}$ denotam $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$ e
$\Pr[\,ij \in \delta(S)\,] \;=\; \Pr[\,s\hat{y} > 0$ e $s\hat{z} \leq 0\,] + \Pr[\,s\hat{y} \leq 0$ e $s\hat{z} > 0\,]$.

Termo vermelho: probabilidade de s estar na esfera unitária na intersecção dos semi-espaços $\{r : r\hat{y} > 0\}$ e $\{r : r\hat{z} \leq 0\}$.

Pela distribuição uniforme, a probabilidade de tal região é proporcional à área da região, e portanto ao ângulo Θ:

$$\Pr[\,s\hat{y} > 0 \text{ e } s\hat{z} \leq 0\,] \;=\; \frac{\Theta}{2\pi}\,.$$

Análise

Lema: Para S e $\hat{Y}$ do algoritmo e quaisquer dois vértices i e j de G, $\Pr[\,ij \in \delta(S)\,] \;=\; \frac{1}{\pi}\,\text{arccos}\,((\hat{Y}\hat{Y}^T)_{ij})\;.$

Prova: $\hat{y}$ e $\hat{z}$ denotam $\hat{Y}_{i\star}$ e $\hat{Y}_{j\star}$ e
$\Pr[\,ij \in \delta(S)\,] \;=\; \Pr[\,s\hat{y} > 0 \text{ e } s\hat{z} \le 0\,] + \Pr[\,s\hat{y} \le 0 \text{ e } s\hat{z} > 0\,]\;.$

Termo vermelho: probabilidade de s estar na esfera unitária na intersecção dos semi-espaços $\{r : r\hat{y} > 0\}$ e $\{r : r\hat{z} \le 0\}$.

Pela distribuição uniforme, a probabilidade de tal região é proporcional à área da região, e portanto ao ângulo Θ:

$$\Pr[\,s\hat{y} > 0 \text{ e } s\hat{z} \le 0\,] \;=\; \frac{\Theta}{2\pi}\;.$$

Do mesmo modo, $\Pr[\,s\hat{y} \le 0 \text{ e } s\hat{z} > 0\,] = \Theta/2\pi$,
o que termina a prova, já que $\Theta = \text{arccos}(\hat{y}\hat{z})$.

O que falta para concluir?

Seja W o valor da solução produzida por MAXCUT-GW.

O que falta para concluir?

Seja W o valor da solução produzida por MAXCUT-GW.

Lembre-se que, por linearidade da esperança,

$$\mathrm{E}[W] \;=\; \sum_{ij} w_{ij}\,\Pr[\,ij \in \delta(S)\,].$$

O que falta para concluir?

Seja W o valor da solução produzida por MAXCUT-GW.

Lembre-se que, por linearidade da esperança,

$$\mathrm{E}[W] \;=\; \sum_{ij} w_{ij}\, \Pr[\, ij \in \delta(S)\,].$$

Pelo lema então

$$\mathrm{E}[W] \;=\; \frac{1}{\pi} \sum_{ij} w_{ij} \arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right) .$$

O que falta para concluir?

Seja W o valor da solução produzida por MAXCUT-GW.

Lembre-se que, por linearidade da esperança,

$$\mathrm{E}[W] \;=\; \sum_{ij} w_{ij}\, \Pr[\,ij \in \delta(S)\,].$$

Pelo lema então

$$\mathrm{E}[W] \;=\; \frac{1}{\pi} \sum_{ij} w_{ij} \arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right).$$

Vamos ver, na aula que vem, que

$$\mathrm{E}[W] \;\geq\; 0{,}878 \sum_{ij} w_{ij}(1-(\hat{Y}\hat{Y}^T)_{ij}) \;\geq\; 0{,}878\,\mathrm{OPT}.$$