Algoritmos de Aproximação

Segundo Semestre de 2012

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.

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 $x_i x_i = 1 \quad$ 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.

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.

Então $\frac{1}{2} \sum_{ij \in E} w_{ij}(1 - (\hat{Y}\hat{Y}^T)_{ij}) \;\geq\; \mathrm{OPT}(G, w)$.

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)

seja $\hat{Y}$ uma solução ótima do programa vetorial
$r \leftarrow$ RANDESFERA$(V(G))$
$S \leftarrow \{i \in V(G) \; : \; r\,\hat{Y}_{i*} > 0\}$
devolva $\delta(S)$

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)

seja $\hat{Y}$ uma solução ótima do programa vetorial
$r \leftarrow$ RANDESFERA$(V(G))$
$S \leftarrow \{i \in V(G) \;:\; r\,\hat{Y}_{i*} > 0\}$
devolva $\delta(S)$

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

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)

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

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

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

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.

Teorema: O algoritmo é uma $0{,}878$-aproximação.

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, $\Pr[\,ij \in \delta(S)\,] \;\geq\; \frac{1}{\pi}\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right)$.

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G,w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G,w) \,.$$

Análise

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

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G, w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G, w)\,.$$

Prova: Pelo lema,

$$\begin{aligned}\mathrm{E}[W] &= \textstyle\sum_{ij\in E} w_{ij}\Pr[\,ij \in \delta(S)\,] \\ &\geq \textstyle\sum_{ij\in E} w_{ij}\,\dfrac{1}{\pi}\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right)\ldots\end{aligned}$$

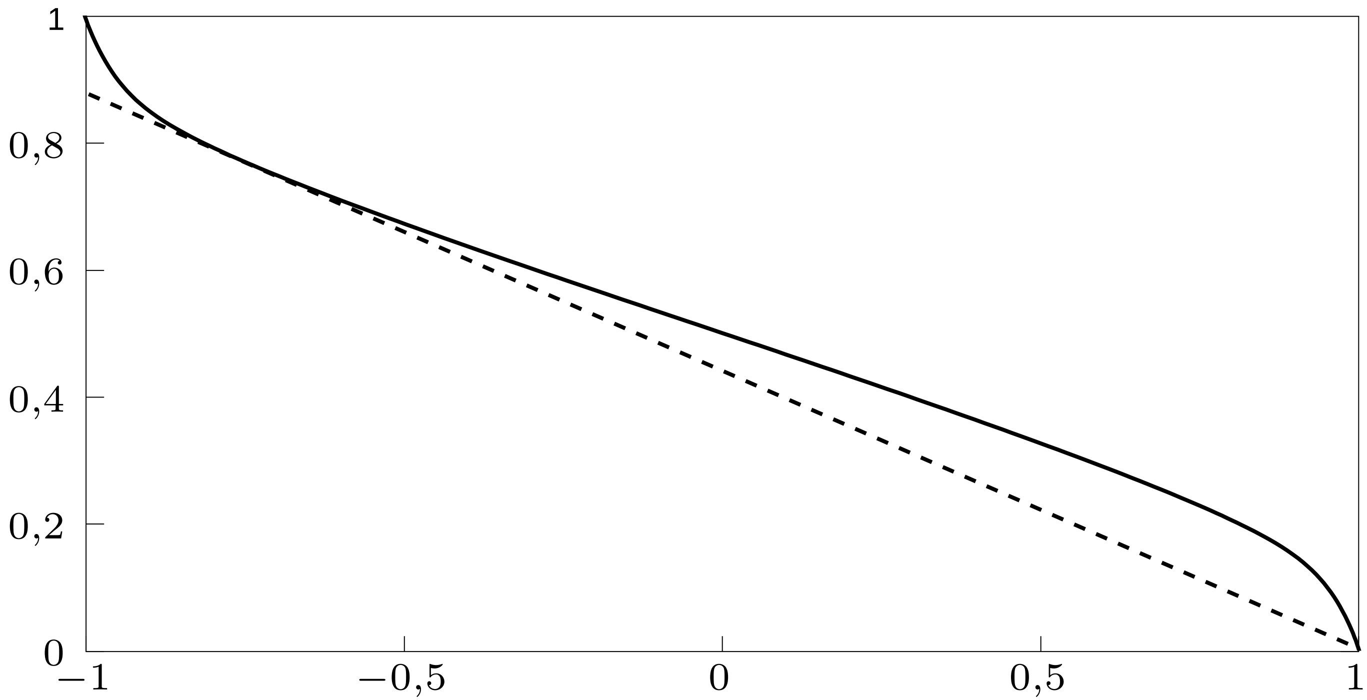
e antes de continuar...

Análise

Cálculos elementares mostram que
$\frac{1}{\pi}\arccos(x) \;\geq\; 0{,}878\,\frac{1}{2}(1-x)$, para todo x no intervalo $[-1,1]$.

Análise

Cálculos elementares mostram que
$\frac{1}{\pi}\arccos(x) \;\geq\; 0{,}878\,\frac{1}{2}(1-x)$, para todo x no intervalo $[-1,1]$.



Análise

Cálculos elementares mostram que
$\frac{1}{\pi}\arccos(x) \;\geq\; 0{,}878\,\frac{1}{2}(1-x)$, para todo x no intervalo $[-1,1]$.

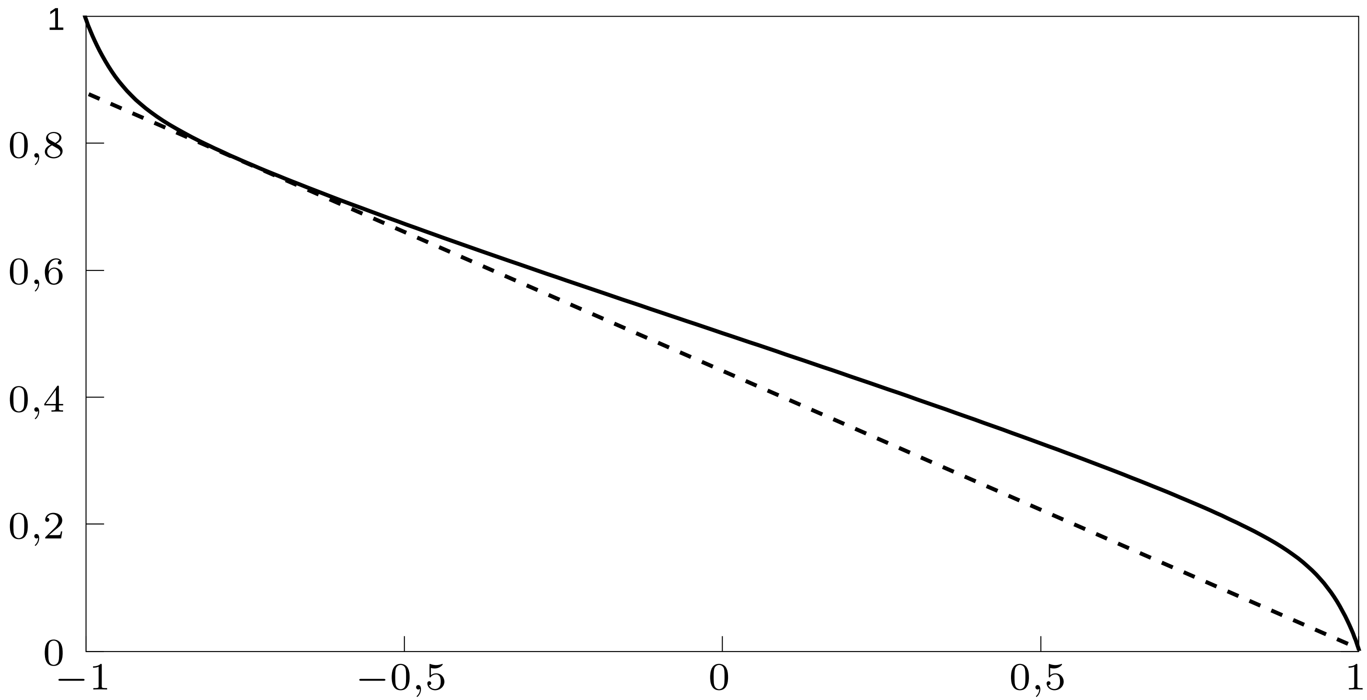


Gráfico de $f(x) := \frac{1}{\pi}\arccos(x)$ e $g(x) := 0{,}878\,\frac{1}{2}(1-x)$.

Verifica-se que $f(x) \geq g(x)$ no intervalo $[-1,1]$.

Finalizando a análise

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G, w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G, w)\,.$$

Finalizando a análise

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G, w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G, w)\,.$$

Prova: Pelo lema,

$$\begin{aligned}
\mathrm{E}[W] &= \textstyle\sum_{ij\in E} w_{ij}\,\Pr[\,ij \in \delta(S)\,] \\
&\geq \textstyle\sum_{ij\in E} w_{ij}\,\dfrac{1}{\pi}\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right)
\end{aligned}$$

Finalizando a análise

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G, w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G, w)\,.$$

Prova: Pelo lema e pela observação prévia,

$$\begin{aligned}
\mathrm{E}[W] &= \textstyle\sum_{ij\in E} w_{ij}\,\Pr[\,ij \in \delta(S)\,] \\
&\geq \textstyle\sum_{ij\in E} w_{ij}\,\dfrac{1}{\pi}\arccos\big((\hat{Y}\hat{Y}^T)_{ij}\big) \\
&\geq 0{,}878\,\dfrac{1}{2}\textstyle\sum_{ij\in E} w_{ij}(1 - (\hat{Y}\hat{Y}^T)_{ij})
\end{aligned}$$

Finalizando a análise

Teorema: Se W é o peso do corte produzido pelo algoritmo MAXCUT-GW (G, w) então

$$\mathrm{E}[W] \;\geq\; 0{,}878\,\mathrm{OPT}(G, w)\;.$$

Prova: Pelo lema e pela observação prévia,

$$\begin{aligned}
\mathrm{E}[W] &= \sum\nolimits_{ij \in E} w_{ij}\,\mathrm{Pr}[\,ij \in \delta(S)\,] \\
&\geq \sum\nolimits_{ij \in E} w_{ij}\,\frac{1}{\pi}\,\arccos\left((\hat{Y}\hat{Y}^T)_{ij}\right) \\
&\geq 0{,}878\,\frac{1}{2}\sum\nolimits_{ij \in E} w_{ij}(1 - (\hat{Y}\hat{Y}^T)_{ij}) \\
&\geq 0{,}878\,\mathrm{OPT}(G, w)\;,
\end{aligned}$$

onde a última desigualdade segue da relaxação.

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)
1 seja $\hat{Y}$ uma solução ótima do programa vetorial
2 $r \leftarrow$ RANDESFERA$(V(G))$
3 $S \leftarrow \{i \in V(G) \ : \ r\hat{Y}_{i*} > 0\}$
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.

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)

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

- Como se executa a linha 1?

Algoritmo de Goemans e Williamson

MAXCUT-GW (G, w)

1 seja $\hat{Y}$ uma solução ótima do programa vetorial
2 $r \leftarrow$ RANDESFERA$(V(G))$
3 $S \leftarrow \{i \in V(G) \;:\; r\,\hat{Y}_{i*} > 0\}$
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.

- Como se executa a linha 1?
- Como se implementa RANDESFERA?

Distribuição uniforme na esfera

Sejam $x_1, \ldots, x_n$ são variáveis aleatórias independentes, cada uma com distribuição normal.

Distribuição uniforme na esfera

Sejam $x_1, \ldots, x_n$ são variáveis aleatórias independentes, cada uma com distribuição normal.

Então o vetor s com componentes

$$s_i = \frac{x_i}{\sqrt{x_1^2 + \cdots + x_n^2}}$$

tem distribuição uniforme na esfera unitária.

Distribuição uniforme na esfera

Sejam $x_1, \ldots, x_n$ são variáveis aleatórias independentes, cada uma com distribuição normal.

Então o vetor s com componentes

$$s_i \;=\; \frac{x_i}{\sqrt{x_1^2 + \cdots + x_n^2}}$$

tem distribuição uniforme na esfera unitária.

Se u e v são variáveis aleatórias independentes, com distribuição uniforme em $[0,1)$, então

$$x \;=\; \cos(2\pi\, v)\,\sqrt{-2\,\ln u}$$

é uma variável aleatória com distribuição normal.

Matrizes positiva semidefinidas

Matriz real X quadrada é positiva semidefinida se, para todo vetor y, vale que $yXy \geq 0$.

Notação: $X \succeq 0$.

Matrizes positiva semidefinidas

Matriz real X quadrada é positiva semidefinida se, para todo vetor y, vale que $yXy \geq 0$.

Notação: $X \succeq 0$.

Lema: Para toda matriz quadrada simétrica X, as seguintes afirmações são equivalentes:

1. X é positiva semidefinida;
2. Todos os autovalores de X são não negativos;
3. Existe uma matriz quadrada Y tal que $X = YY^T$;
4. $X = \sum_{i=1}^{n} \lambda_i w_i w_i^T$ onde $\lambda_i \geq 0$ e $w_i^T w_i = 1$ para todo i, e $w_i^T w_j = 0$ para todo $j \neq i$.

Programas semidefinidos

Dados a, b e c, encontrar matrix real X $n \times n$, que

maximize $\sum_{ij} c_{ij} x_{ij}$

sujeito a $\sum_{i,j} a_{ijk}\, x_{ij} = b_k$ para cada k,

$x_{ij} = x_{ji}$ para todo par i, j

$X \succeq 0$.

Programas semidefinidos

Dados a, b e c, encontrar matrix real X $n \times n$, que

maximize $\sum_{ij} c_{ij} x_{ij}$

sujeito a $\sum_{i,j} a_{ijk}\, x_{ij} = b_k$ para cada k,

$x_{ij} = x_{ji}$ para todo par i, j

$X \succeq 0$.

Sob certas condições técnicas, resolve-se tal programa dentro de um erro aditivo de $\epsilon > 0$, em tempo polinomial no tamanho da entrada (soma dos tamanhos de a, b e c) e $\lg \frac{1}{\epsilon}$.

Programas semidefinidos

Dados a, b e c, encontrar matrix real X $n \times n$, que

maximize $\sum_{ij} c_{ij} x_{ij}$

sujeito a $\sum_{i,j} a_{ijk}\, x_{ij} = b_k$ para cada k,

$x_{ij} = x_{ji}$ para todo par i, j

$X \succeq 0$.

Sob certas condições técnicas, resolve-se tal programa dentro de um erro aditivo de $\epsilon > 0$, em tempo polinomial no tamanho da entrada (soma dos tamanhos de a, b e c) e $\lg \frac{1}{\epsilon}$.

Condições técnicas:
A região viável não é vazia e está contida em uma bola de tamanho polinomial (na entrada).

Matrizes positiva semidefinidas

Matriz real quadrada X é positiva semidefinida se existe matriz quadrada Y tal que $X = YY^T$.

Tal X é simétrica.

Matrizes positiva semidefinidas

Matriz real quadrada X é positiva semidefinida se existe matriz quadrada Y tal que $X = YY^T$.

Tal X é simétrica.

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Matrizes positiva semidefinidas

Matriz real quadrada X é positiva semidefinida se existe matriz quadrada Y tal que $X = YY^T$.

Tal X é simétrica.

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Primeiro, se $X = Y\,Y^T$, então

$$yXy \;=\; y(YY^T)y \;=\; (yY)(yY)^T \;\geq\; 0,$$

ou seja, no máximo uma das alternativa vale.

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Autovalores $\lambda_1, \ldots, \lambda_n$ de X.
(Matriz real simétrica tem todos os autovalores reais.)

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Autovalores $\lambda_1, \ldots, \lambda_n$ de X.
(Matriz real simétrica tem todos os autovalores reais.)

Λ: matriz diagonal com $\lambda_1, \ldots, \lambda_n$ na diagonal.
U: matriz ortogonal cuja coluna i é autovetor de λ_i.

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Autovalores $\lambda_1, \ldots, \lambda_n$ de X.
(Matriz real simétrica tem todos os autovalores reais.)

Λ: matriz diagonal com $\lambda_1, \ldots, \lambda_n$ na diagonal.
U: matriz ortogonal cuja coluna i é autovetor de λ_i.
Verifique que $X = U\Lambda\,U^T$.

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Autovalores $\lambda_1, \ldots, \lambda_n$ de X.
(Matriz real simétrica tem todos os autovalores reais.)

Λ: matriz diagonal com $\lambda_1, \ldots, \lambda_n$ na diagonal.
U: matriz ortogonal cuja coluna i é autovetor de λ_i.
Verifique que $X = U\Lambda\,U^T$.

Se X é positiva semidefinida, $\lambda_i \geq 0$ para todo i, e
$X \;=\; U\sqrt{\Lambda}\sqrt{\Lambda}\,U^T \;=\; (U\sqrt{\Lambda})(U\sqrt{\Lambda})^T$, ou seja, $Y = U\sqrt{\Lambda}$.

Continuação da prova

Lema: Para toda matriz simétrica X,
vale uma e apenas uma das seguintes alternativas:
existe uma matriz quadrada Y tal que $X = Y\,Y^T$
ou existe um vetor y tal que $yXy < 0$.

Prova: Autovalores $\lambda_1, \ldots, \lambda_n$ de X.
(Matriz real simétrica tem todos os autovalores reais.)

Λ: matriz diagonal com $\lambda_1, \ldots, \lambda_n$ na diagonal.
U: matriz ortogonal cuja coluna i é autovetor de λ_i.
Verifique que $X = U\Lambda\,U^T$.

Se X é positiva semidefinida, $\lambda_i \geq 0$ para todo i, e
$X \;=\; U\sqrt{\Lambda}\sqrt{\Lambda}\,U^T \;=\; (U\sqrt{\Lambda})(U\sqrt{\Lambda})^T$, ou seja, $Y = U\sqrt{\Lambda}$.

Se X não é positiva semidefinida, $\lambda_i < 0$ para algum i.
Logo $Xu_i = \lambda_i\,u_i$ e $u_iXu_i = \lambda_i\,u_i\,u_i \;=\; \lambda_i \;<\; 0$.

Matrizes positiva semidefinidas

Programa 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.

Matrizes positiva semidefinidas

Programa 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.

Programa semidefinido correspondente:

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

maximize $\sum_{ij} w_{ij}(1 - x_{ij})$

sujeito a $x_{ii} = 1$ para cada i em V,

$x_{ij} = x_{ji}$ para todo par i, j em V,

$X \succeq 0$.

Matrizes positiva semidefinidas

Programa 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.

Programa semidefinido correspondente:

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

maximize $\sum_{ij} w_{ij}(1 - x_{ij})$

sujeito a $x_{ii} = 1$ para cada i em V,

$x_{ij} = x_{ji}$ para todo par i, j em V,

$X \succeq 0$.

As condições técnicas estão satisfeitas.

O lema anterior dá um algoritmo de separação.

Correlational clustering

Dados $G = (V, E)$ e pesos $w_{ij}^{+} \geq 0$ e $w_{ij}^{-} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Correlational clustering

Dados $G = (V, E)$ e pesos $w_{ij}^{+} \geq 0$ e $w_{ij}^{-} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Dado um clustering $\mathcal{C}$,
$\delta(\mathcal{C})$: arestas com extremos em conjuntos distintos de $\mathcal{C}$
$E(\mathcal{C})$: arestas com extremos num mesmo conjunto de $\mathcal{C}$.

Correlational clustering

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Dado um clustering $\mathcal{C}$,
$\delta(\mathcal{C})$: arestas com extremos em conjuntos distintos de $\mathcal{C}$
$E(\mathcal{C})$: arestas com extremos num mesmo conjunto de $\mathcal{C}$.

Problema: Dados G, w^+ e w^-, encontrar
um clustering $\mathcal{C}$ que maximize $w^+(E(\mathcal{C})) + w^-(\delta(\mathcal{C}))$.

Correlational clustering

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Dado um clustering $\mathcal{C}$,
$\delta(\mathcal{C})$: arestas com extremos em conjuntos distintos de $\mathcal{C}$
$E(\mathcal{C})$: arestas com extremos num mesmo conjunto de $\mathcal{C}$.

Problema: Dados G, w^+ e w^-, encontrar um clustering $\mathcal{C}$ que maximize $w^+(E(\mathcal{C})) + w^-(\delta(\mathcal{C}))$.

$\frac{1}{2}$-aproximação:
devolva o melhor entre $\mathcal{C}_1 = \{V\}$ e $\mathcal{C}_2 = \{\{v\} : v \in V\}$.

Correlational clustering

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Dado um clustering $\mathcal{C}$,
$\delta(\mathcal{C})$: arestas com extremos em conjuntos distintos de $\mathcal{C}$
$E(\mathcal{C})$: arestas com extremos num mesmo conjunto de $\mathcal{C}$.

Problema: Dados G, w^+ e w^-, encontrar um clustering $\mathcal{C}$ que maximize $w^+(E(\mathcal{C})) + w^-(\delta(\mathcal{C}))$.

$\frac{1}{2}$-aproximação:
devolva o melhor entre $\mathcal{C}_1 = \{V\}$ e $\mathcal{C}_2 = \{\{v\} : v \in V\}$.

$$\mathrm{OPT} \;\leq\; \sum_{ij \in E} (w^+_{ij} + w^-_{ij}) \;=\; w^+(E(\mathcal{C}_1)) + w^-(\delta(\mathcal{C}_2))$$

Uma formulação

Dados $G = (V, E)$ e pesos $w_{ij}^{+} \geq 0$ e $w_{ij}^{-} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Vamos mostrar uma $\frac{3}{4}$-aproximação para o problema.

Uma formulação

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Vamos mostrar uma $\frac{3}{4}$-aproximação para o problema.

e_k: vetor com k-ésima coordenada 1, zero nas demais.

Uma formulação

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Vamos mostrar uma $\frac{3}{4}$-aproximação para o problema.

e_k: vetor com k-ésima coordenada 1, zero nas demais.

Ideia: $x_k = e_k$ sse v_k está no cluster k.

Uma formulação

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Vamos mostrar uma $\frac{3}{4}$-aproximação para o problema.

e_k: vetor com k-ésima coordenada 1, zero nas demais.

Ideia: $x_k = e_k$ sse v_k está no cluster k.

Formulação alternativa:

Dados G, w^+ e w^-, encontrar $x_1, \ldots, x_n$, onde $n = |V_G|$, que
maximizem $\sum_{ij} \left(w^+_{ij}(x_i\, x_j) + w^-_{ij}(1 - x_i\, x_j)\right)$
sujeito a $x_i \in \{e_1, \ldots, e_n\}$ para $i = 1, \ldots, n$.

Relaxação vetorial

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Formulação alternativa:

Dados G, w^+ e w^-, encontrar $x_1, \ldots, x_n$, onde $n = |V_G|$, que

maximizem $\sum_{ij} \left(w^+_{ij}(x_i\, x_j) + w^-_{ij}(1 - x_i\, x_j)\right)$

sujeito a $x_i \in \{e_1, \ldots, e_n\}$ para $i = 1, \ldots, n$.

Relaxação vetorial

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Formulação alternativa:

Dados G, w^+ e w^-, encontrar $x_1, \ldots, x_n$, onde $n = |V_G|$, que

maximizem $\sum_{ij} \left(w^+_{ij}(x_i\, x_j) + w^-_{ij}(1 - x_i\, x_j) \right)$

sujeito a $x_i \in \{e_1, \ldots, e_n\}$ para $i = 1, \ldots, n$.

Relaxação vetorial:

Dados G, w^+ e w^-, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w^+_{ij}(v_i\, v_j) + w^-_{ij}(1 - v_i\, v_j) \right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i, j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

Relaxação vetorial

Dados $G = (V, E)$ e pesos $w^+_{ij} \geq 0$ e $w^-_{ij} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Relaxação vetorial:

Dados G, w^+ e w^-, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w^+_{ij}(v_i\, v_j) + w^-_{ij}(1 - v_i\, v_j) \right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i,\, j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

Relaxação vetorial

Dados $G = (V, E)$ e pesos $w_{ij}^{+} \geq 0$ e $w_{ij}^{-} \geq 0$ para cada ij em E, um clustering é uma partição de V.

Relaxação vetorial:

Dados G, w^{+} e w^{-}, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w_{ij}^{+}(v_i\, v_j) + w_{ij}^{-}(1 - v_i\, v_j) \right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i,\ j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

Como é uma relaxação do problema, vale que $z_{CC}^{*} \geq \mathrm{OPT}$.

Relaxação vetorial

Dados $G = (V, E)$ e pesos $w_{ij}^+ \geq 0$ e $w_{ij}^- \geq 0$ para cada ij em E, um clustering é uma partição de V.

Relaxação vetorial:

Dados G, w^+ e w^-, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w_{ij}^+ (v_i\, v_j) + w_{ij}^- (1 - v_i\, v_j) \right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i,\, j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

Como é uma relaxação do problema, vale que $z_{CC}^* \geq \mathrm{OPT}$.

Pode ser resolvida (dentro de um erro aditivo) em tempo polinomial.

Algoritmo

Dados G, w^+ e w^-, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w^+_{ij}(v_i\, v_j) + w^-_{ij}(1 - v_i\, v_j)\right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i,\ j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

Algoritmo

Dados G, w^+ e w^-, encontrar vetores $v_1, \ldots, v_n$ que

maximizem $\sum_{ij} \left(w^+_{ij}(v_i\, v_j) + w^-_{ij}(1 - v_i\, v_j)\right)$

sujeito a $v_i\, v_i = 1$ para $i = 1, \ldots, n$

$v_i\, v_j \geq 0$ para $i,\, j = 1, \ldots, n$

$v_i \in \mathbb{R}^n$ para $i = 1, \ldots, n$.

MAXCC (G, w^+, w^-)

1 seja v uma solução ótima do programa vetorial

2 $r_1 \leftarrow$ RANDESFERA$(V(G))$ $r_2 \leftarrow$ RANDESFERA$(V(G))$

3 $R_1 \leftarrow \{i \in V(G) \ : \ r_1\, v_i \geq 0$ e $r_2 v_i \geq 0\}$

4 $R_2 \leftarrow \{i \in V(G) \ : \ r_1\, v_i \geq 0$ e $r_2 v_i < 0\}$

5 $R_3 \leftarrow \{i \in V(G) \ : \ r_1\, v_i < 0$ e $r_2 v_i \geq 0\}$

6 $R_4 \leftarrow \{i \in V(G) \ : \ r_1\, v_i < 0$ e $r_2 v_i < 0\}$

7 **devolva** $\{R_1, R_2, R_3, R_4\}$

Análise

Lema: Para x em $[0,1]$,

$$\frac{(1-\frac{1}{\pi}\arccos(x))^2}{x} \;\geq\; 0{,}75 \quad \text{e} \quad \frac{1-(1-\frac{1}{\pi}\arccos(x))^2}{1-x} \;\geq\; 0{,}75.$$

Análise

Lema: Para x em $[0,1]$,

$$\frac{(1-\frac{1}{\pi}\arccos(x))^2}{x} \;\geq\; 0{,}75 \quad \text{e} \quad \frac{1-(1-\frac{1}{\pi}\arccos(x))^2}{1-x} \;\geq\; 0{,}75.$$

Teorema: Se W é o peso do clustering $\mathcal{C}$ produzido pelo algoritmo MAXCC (G, w^+, w^-) então

$$\mathrm{E}[W] \;\geq\; 0{,}75\,\mathrm{OPT}(G, w^+, w^-)\,.$$

Análise

Lema: Para x em $[0,1]$,

$$\frac{(1-\frac{1}{\pi}\arccos(x))^2}{x} \;\geq\; 0{,}75 \;\text{ e }\; \frac{1-(1-\frac{1}{\pi}\arccos(x))^2}{1-x} \;\geq\; 0{,}75.$$

Teorema: Se W é o peso do clustering $\mathcal{C}$ produzido pelo algoritmo MAXCC (G, w^+, w^-) então

$$\mathrm{E}[W] \;\geq\; 0{,}75\,\mathrm{OPT}(G, w^+, w^-)\,.$$

Prova: Seja X_{ij} a v.a. que vale 1 se i e j estão no mesmo cluster em $\mathcal{C}$, 0 caso contrário. Então

Análise

Lema: Para x em $[0,1]$,

$$\frac{(1-\frac{1}{\pi}\arccos(x))^2}{x} \;\geq\; 0{,}75 \;\text{ e }\; \frac{1-(1-\frac{1}{\pi}\arccos(x))^2}{1-x} \;\geq\; 0{,}75.$$

Teorema: Se W é o peso do clustering $\mathcal{C}$ produzido pelo algoritmo MAXCC (G, w^+, w^-) então

$$\mathrm{E}[W] \;\geq\; 0{,}75\,\mathrm{OPT}(G, w^+, w^-)\,.$$

Prova: Seja X_{ij} a v.a. que vale 1 se i e j estão no mesmo cluster em $\mathcal{C}$, 0 caso contrário. Então

$$\mathrm{E}[W] \;=\; \textstyle\sum_{ij\in E}\,(w^+_{ij}\,\mathrm{E}[X_{ij}] + w^-_{ij}\,(1-\mathrm{E}[X_{ij}])).$$

Vamos calcular $\mathrm{E}[X_{ij}]$.

Análise

A probabilidade um hiperplano aleatório deixar v_i e v_j do mesmo lado é $1 - \frac{1}{\pi} \arccos(v_i\, v_j)$.

Análise

A probabilidade um hiperplano aleatório deixar v_i e v_j do mesmo lado é $1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\, v_j)$.

A probabilidade de v_i e v_j estarem no mesmo lado dos hiperplanos dados por r_1 e de r_2 é $(1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\, v_j))^2$, ou seja, $\mathrm{E}[X_{ij}] = (1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\, v_j))^2$.

Análise

A probabilidade um hiperplano aleatório deixar v_i e v_j do mesmo lado é $1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\,v_j)$.

A probabilidade de v_i e v_j estarem no mesmo lado dos hiperplanos dados por r_1 e de r_2 é $(1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\,v_j))^2$, ou seja, $\mathrm{E}[X_{ij}] = (1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\,v_j))^2$.

Segue que

$$\begin{aligned}
\mathrm{E}[W] &= \textstyle\sum_{ij\in E} (w^+_{ij}\,\mathrm{E}[X_{ij}] + w^-_{ij}\,(1 - \mathrm{E}[X_{ij}])) \\
&= \textstyle\sum_{ij\in E} \Big[w^+_{ij}\,(1 - \frac{1}{\pi}\,\mathrm{arccos}(v_i\,v_j))^2 \\
&\qquad\qquad + w^-_{ij}\,(1 - (1 - \dfrac{1}{\pi}\,\mathrm{arccos}(v_i\,v_j))^2)\Big] \\
&\geq 0{,}75 \textstyle\sum_{ij\in E} \big(w^+_{ij}\,(v_i\,v_j) + w^-_{ij}\,(1 - v_i\,v_j)\big) \\
&\geq 0{,}75\,\mathrm{OPT}.
\end{aligned}$$