Programação semidefinida

João Gouveia

· U C ·

FCTUC FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE DE COIMBRA

Jornadas de Matemática - 10 de fevereiro 2018 - Universidade do Porto

Capítulo 1

O que é?

Programação linear - Exemplo

Um criador de porcos pretende determinar as quantidades de cada tipo de ração que devem ser dadas diariamente a cada animal por forma a conseguir uma certa qualidade nutritiva a um custo mínimo.

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $\quad y \rightarrow$ Quantidade de Ração B.

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $\quad$ $y \rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & \end{array}$$

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $y \rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \text{min} & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \end{array}$$

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $y \rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \\ & 50x + 10y \geq 150 \end{array}$$

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \to$ Quantidade de Ração A $y \to$ Quantidade de Ração B.

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \\ & 50x + 10y \geq 150 \\ & 30x + 30y \leq 600 \end{array}$$

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $\quad$ $y \rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \\ & 50x + 10y \geq 150 \\ & 30x + 30y \leq 600 \\ & x, y \geq 0 \end{array}$$

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x\rightarrow$ Quantidade de Ração A $\quad y\rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \text{min} & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \\ & 50x + 10y \geq 150 \\ & 30x + 30y \leq 600 \\ & x, y \geq 0 \end{array}$$

$$\begin{array}{ll} \text{min} & 10x + 5y \\ \text{s. a} & \begin{bmatrix} -200 \\ -150 \\ 600 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 20 & 50 \\ 50 & 10 \\ -30 & -30 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \end{array}$$

Forma canónica

Programação Linear - Exemplo (cont.)

	Ração A	Ração B	Quantidade Requerida
Proteína	20	50	≥ 200
Vitamina	50	10	≥ 150
Hid. Carbono	30	30	≤ 600
Custo	10	5	

$x \rightarrow$ Quantidade de Ração A $\quad y \rightarrow$ Quantidade de Ração B.

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & 20x + 50y \geq 200 \\ & 50x + 10y \geq 150 \\ & 30x + 30y \leq 600 \\ & x, y \geq 0 \end{array}$$

$$\begin{array}{ll} \min & 10x + 5y \\ \text{s. a} & \begin{bmatrix} 20 & 50 & -1 & 0 & 0 \\ 50 & 10 & 0 & -1 & 0 \\ -30 & -30 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 200 \\ 150 \\ -600 \end{bmatrix} \\ & x, y\ z_1, z_2, z_3 \geq 0 \end{array}$$

Forma normal

Programação linear - Definição

Programa Linear - Forma normal

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array}$$

Programa Linear - Forma canónica

$$\begin{array}{ll} \min & \langle c, y\rangle \\ \text{s.t.} & b + Ay \geq 0 \end{array}$$

Programação linear - Definição

Programa Linear - Forma normal

$$\begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & Ax = b \\ & x \in \mathbb{R}^n_+. \end{array}$$

Programa Linear - Forma canónica

$$\begin{array}{ll} \min & \langle c, y \rangle \\ \text{s.t.} & b + Ay \geq 0 \end{array} \qquad \begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & b + Ay \in \mathbb{R}^n_+. \end{array}$$

Programação linear - Definição

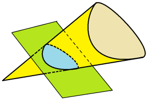
Programa Linear - Forma normal

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \mathbb{R}^n_+. \end{array}$$

Programa Linear - Forma canónica

$$\begin{array}{ll} \min & \langle c, y\rangle \\ \text{s.t.} & b + Ay \geq 0 \end{array} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & b + Ay \in \mathbb{R}^n_+. \end{array}$$

Em ambos os casos estamos a otimizar sobre fatias afins do ortante não-negativo.



Podemos jogar esse jogo com outros cones.

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t Ax \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^{n} v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^{n} v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.
- $A = B^t B$ para alguma matriz B.

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^{n} v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.
- $A = B^t B$ para alguma matriz B.

Exemplo:

$$A = \begin{bmatrix} 9 & -8 & -2 \\ -8 & 26 & 4 \\ -2 & 4 & 6 \end{bmatrix} \succeq 0$$

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^{n} v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.
- $A = B^t B$ para alguma matriz B.

Exemplo:

$$A = \begin{bmatrix} 9 & -8 & -2 \\ -8 & 26 & 4 \\ -2 & 4 & 6 \end{bmatrix} \succeq 0 \longrightarrow \mathrm{Spec}(A) = \{5, 6, 30\}$$

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^n v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.
- $A = B^t B$ para alguma matriz B.

Exemplo:

$$A = \begin{bmatrix} 9 & -8 & -2 \\ -8 & 26 & 4 \\ -2 & 4 & 6 \end{bmatrix} \succeq 0 \longrightarrow \text{Spec}(A) = \{5, 6, 30\} \longrightarrow A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & -5 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \\ 2 & -5 & -1 \end{bmatrix}$$

O meu cone favorito

Uma matriz A simétrica $n \times n$ diz-se semidefinida positiva, e escrevemos $A \succeq 0$ se

- $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$.
- Os valores próprios de A forem todos não negativos.
- $A = \sum_{i=1}^n v_i v_i^t$ para alguns $v_i \in \mathbb{R}^n$.
- $A = B^t B$ para alguma matriz B.

Exemplo:

$$A = \begin{bmatrix} 9 & -8 & -2 \\ -8 & 26 & 4 \\ -2 & 4 & 6 \end{bmatrix} \succeq 0 \longrightarrow \mathrm{Spec}(A) = \{5, 6, 30\} \longrightarrow A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & -5 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \\ 2 & -5 & -1 \end{bmatrix}$$

O conjunto das matrizes semidefinidas positivas, S_+^n, é fechado para adição e multiplicação por escalares positivos: é um *cone convexo*.

Programação semidefinida - Definição

Programa semidefinido - Forma normal

$$\begin{array}{rl} \min_{X \in S^n} & \langle C, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle = b_i, \ \ i = 1, \cdots, m \\ & X \succeq 0. \end{array}$$

Programa semidefinido - Forma canónica

$$\begin{array}{rl} \min_{y \in \mathbb{R}^m} & \langle c, y \rangle \\ \text{s.t.} & A_0 + \sum y_i A_i \succeq 0 \end{array}$$

Programação semidefinida - Definição

Programa semidefinido - Forma normal

$$\begin{array}{rl} \min_{X \in \mathcal{S}^n} & \langle C, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle = b_i, \;\; i = 1, \cdots, m \\ & X \succeq 0. \end{array}$$

Programa semidefinido - Forma canónica

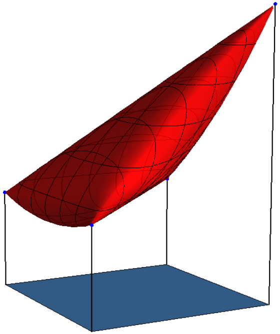
$$\begin{array}{rl} \min_{y \in \mathbb{R}^m} & \langle c, y \rangle \\ \text{s.t.} & A_0 + \sum y_i A_i \succeq 0 \end{array}$$

Mais uma vez, em ambos os casos estamos a otimizar sobre fatias afins do cone das matrizes semidefinidas positivas.

Programação semidefinida - Exemplos

$$\begin{array}{ll} \min & x - y + z \\ \text{s.t.} & \begin{bmatrix} 1 & x & y \\ x & x & z \\ y & z & y \end{bmatrix} \succeq 0 \end{array}$$

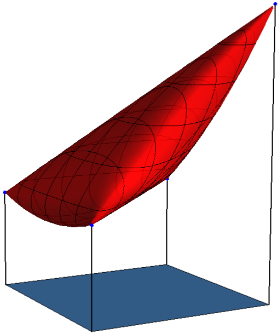
$$\begin{array}{ll} \min & x - y + z \\ \text{s.t.} & \begin{bmatrix} 1 & x & y \\ x & x & z \\ y & z & y \end{bmatrix} \succeq 0 \end{array}$$



Programação semidefinida - Exemplos

$$\begin{array}{ll} \min & x - y + z \\ \text{s.t.} & \begin{bmatrix} 1 & x & y \\ x & x & z \\ y & z & y \end{bmatrix} \succeq 0 \end{array}$$

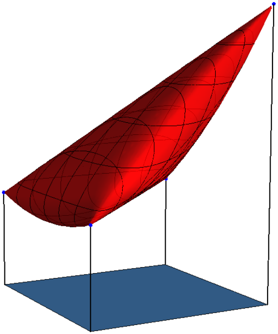
$$\begin{array}{ll} \min & x + y + z \\ \text{s.t.} & \begin{bmatrix} 1+x & y-z & x+y \\ y-z & 1-x & z \\ x+y & z & z \end{bmatrix} \succeq 0 \end{array}$$

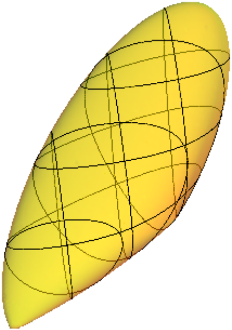


Programação semidefinida - Exemplos

$$\begin{array}{ll} \min & x - y + z \\ \text{s.t.} & \begin{bmatrix} 1 & x & y \\ x & x & z \\ y & z & y \end{bmatrix} \succeq 0 \end{array}$$

$$\begin{array}{ll} \min & x + y + z \\ \text{s.t.} & \begin{bmatrix} 1+x & y-z & x+y \\ y-z & 1-x & z \\ x+y & z & z \end{bmatrix} \succeq 0 \end{array}$$





Capítulo 2

Para que serve?

1. Problema do corte máximo

Problema do corte máximo

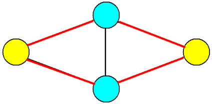
Definition

Dado um grafo $G = (V, E)$ e uma partição V_1, V_2 de V, o conjunto C de arestas entre V_1 e V_2 diz-se um **corte**.

Problema do corte máximo

Definition

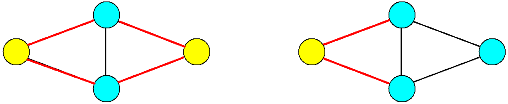
Dado um grafo $G = (V, E)$ e uma partição V_1, V_2 de V, o conjunto C de arestas entre V_1 e V_2 diz-se um **corte.**



Problema do corte máximo

Definition

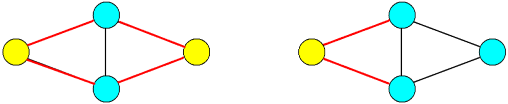
Dado um grafo $G = (V, E)$ e uma partição V_1, V_2 de V, o conjunto C de arestas entre V_1 e V_2 diz-se um **corte**.



Problema do corte máximo

Definition

Dado um grafo $G = (V, E)$ e uma partição V_1, V_2 de V, o conjunto C de arestas entre V_1 e V_2 diz-se um **corte**.



O problema

Queremos encontrar o corte C com o maior número possível de arestas.

Como encontrar um corte

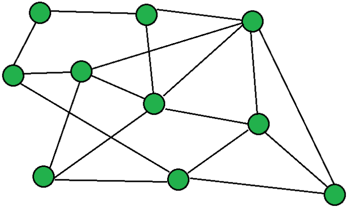
Ideia Simples

Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.

Como encontrar um corte

Ideia Simples

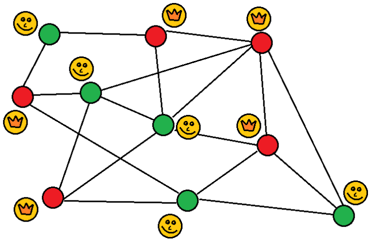
Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.



Como encontrar um corte

Ideia Simples

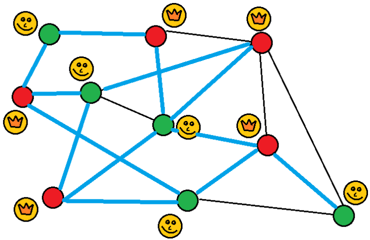
Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.



Como encontrar um corte

Ideia Simples

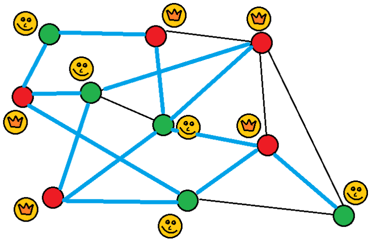
Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.



Como encontrar um corte

Ideia Simples

Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.

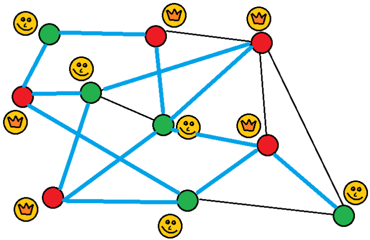


A probabilidade de uma aresta estar no corte é de $1/2$, logo o número médio de arestas obtido é metade das do grafo, e portanto pelo menos metade do corte ótimo.

Como encontrar um corte

Ideia Simples

Lançamos a moeda ao ar para cada vértice e dividimo-los em cara e coroa.



A probabilidade de uma aresta estar no corte é de $1/2$, logo o número médio de arestas obtido é metade das do grafo, e portanto pelo menos metade do corte ótimo.

Podemos fazer melhor?

Problema reformulado

Seja $x_i = 1$ se $i \in V_1$ e -1 se não. Então $x_i x_j = -1$ se $\{i, j\}$ estiver no corte e 1 se não.

Problema reformulado

Seja $x_i = 1$ se $i \in V_1$ e -1 se não. Então $x_i x_j = -1$ se $\{i, j\}$ estiver no corte e 1 se não.

MaxCut

$$\max_{x \in \{-1,1\}^n} \sum_{i,j} \frac{1}{2}(1 - x_i x_j).$$

Problema reformulado

Seja $x_i = 1$ se $i \in V_1$ e -1 se não. Então $x_i x_j = -1$ se $\{i, j\}$ estiver no corte e 1 se não.

MaxCut

$$\max_{x \in \{-1,1\}^n} \sum_{i,j} \frac{1}{2}(1 - x_i x_j).$$

Note-se que $X = xx^t = [x_i x_j]_{i,j=1,\dots,n}$, para $x \in \{-1, 1\}^n$ se e só se

$$X \succeq 0, \ X_{ii} = 1 \text{ e car}(X) = 1.$$

Problema reformulado

Seja $x_i = 1$ se $i \in V_1$ e -1 se não. Então $x_i x_j = -1$ se $\{i, j\}$ estiver no corte e 1 se não.

MaxCut

$$\max_{x \in \{-1,1\}^n} \sum_{i,j} \frac{1}{2}(1 - x_i x_j).$$

Note-se que $X = xx^t = [x_i x_j]_{i,j=1,\ldots,n}$, para $x \in \{-1, 1\}^n$ se e só se

$$X \succeq 0,\ X_{ii} = 1 \text{ e } \mathrm{car}(X) = 1.$$

MaxCut

$$\max_{X} \sum_{i,j} \frac{1}{2}(1 - X_{ij}),$$

com $X \succeq 0$, $\mathrm{diag}(X) = 1$, $\mathrm{car}(X) = 1$.

Problema reformulado

Seja $x_i = 1$ se $i \in V_1$ e -1 se não. Então $x_i x_j = -1$ se $\{i, j\}$ estiver no corte e 1 se não.

MaxCut

$$\max_{x \in \{-1,1\}^n} \sum_{i,j} \frac{1}{2}(1 - x_i x_j).$$

Note-se que $X = xx^t = [x_i x_j]_{i,j=1,\ldots,n}$, para $x \in \{-1, 1\}^n$ se e só se

$$X \succeq 0, \; X_{ii} = 1 \text{ e } \mathrm{car}(X) = 1.$$

Esquecendo a última condição obtemos uma relaxação.

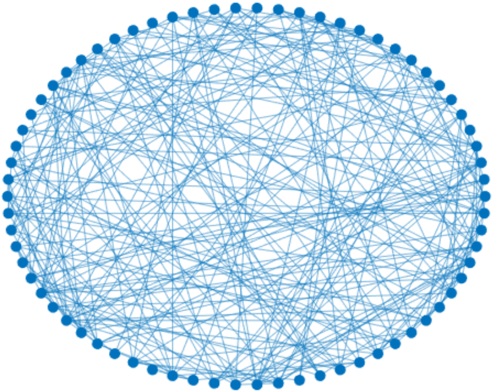
MaxCut relaxado

$$\max_{X} \sum_{i,j} \frac{1}{2}(1 - X_{ij}),$$

com $X \succeq 0$, $\mathrm{diag}(X) = 1$

Exemplo

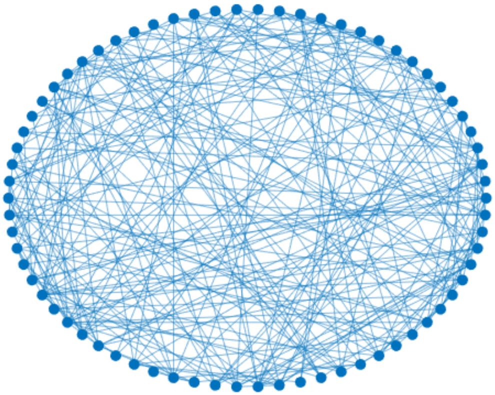
Grafo aleatório com 70 vértices e 250 arestas.



Qual o maior corte?

Exemplo

Grafo aleatório com 70 vértices e 250 arestas.

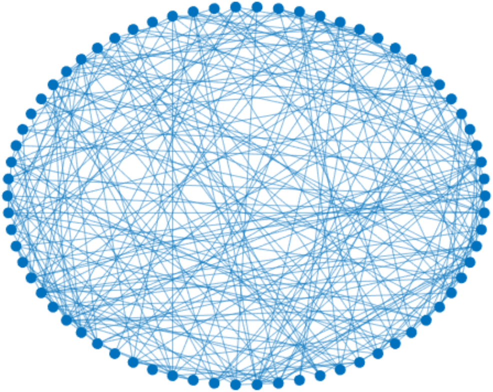


Qual o maior corte?

No máximo 195, porque o ótimo da relaxação é 195.56

Exemplo

Grafo aleatório com 70 vértices e 250 arestas.



Qual o maior corte?

No máximo 195, porque o ótimo da relaxação é 195.56 **Mas quão boa é a aproximação?**

Goemans-Williamson

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Goemans-Williamson

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Como o encontrar?

Goemans-Williamson

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Como o encontrar?

1. Escrever $X = CC^T$. Linhas de $C \leftrightarrow$ vetores em $\mathbb{R}^n$ de norma 1.

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Como o encontrar?

1. Escrever $X = CC^T$. Linhas de $C \leftrightarrow$ vetores em $\mathbb{R}^n$ de norma 1.
2. A cada vértice v de G corresponde um ponto p_v na esfera S^{n-1}.

Goemans-Williamson

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Como o encontrar?

1. Escrever $X = CC^T$. Linhas de $C \leftrightarrow$ vetores em $\mathbb{R}^n$ de norma 1.
2. A cada vértice v de G corresponde um ponto p_v na esfera S^{n-1}.
3. Geramos um hiperplano aleatório e particionamos os vértices conforme o lado em que calham.

Theorem (Goemans-Williamson 95)

O valor real do MAXCUT é pelo menos 0.87856 vezes o valor da relaxação.

Como o encontrar?

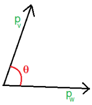
1. Escrever $X = CC^T$. Linhas de $C \leftrightarrow$ vetores em $\mathbb{R}^n$ de norma 1.
2. A cada vértice v de G corresponde um ponto p_v na esfera S^{n-1}.
3. Geramos um hiperplano aleatório e particionamos os vértices conforme o lado em que calham.

Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?

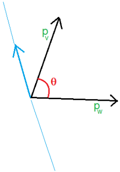
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



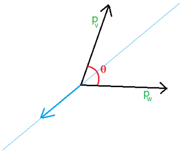
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



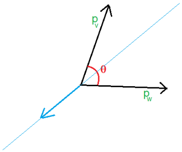
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?

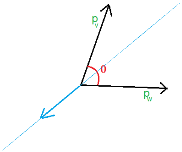


Probabilidade da aresta v, w estar no corte é

$$\frac{\theta}{\pi}$$

Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?

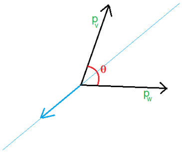


Probabilidade da aresta v, w estar no corte é

$$\frac{\theta}{\pi} = \frac{\arccos\langle u, w\rangle}{\pi}$$

Porque funciona?

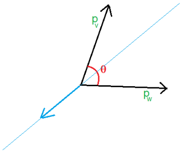
Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



Probabilidade da aresta v, w estar no corte é

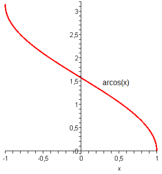
$$\frac{\theta}{\pi} = \frac{\operatorname{arccos}\langle u, w\rangle}{\pi} = \frac{\operatorname{arccos}(X_{ij})}{\pi}$$

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



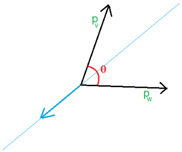
Probabilidade da aresta v, w estar no corte é

$$\frac{\theta}{\pi} = \frac{\arccos\langle u, w\rangle}{\pi} = \frac{\arccos(X_{ij})}{\pi}$$



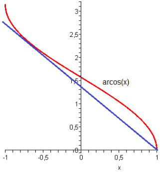
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



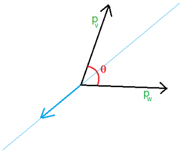
Probabilidade da aresta v, w estar no corte é

$$\frac{\theta}{\pi} = \frac{\arccos\langle u, w\rangle}{\pi} = \frac{\arccos(X_{ij})}{\pi}$$



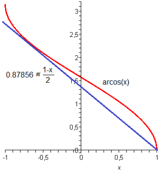
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



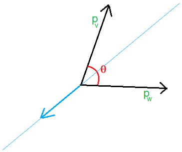
Probabilidade da aresta v, w estar no corte é

$$\frac{\theta}{\pi} = \frac{\arccos\langle u, w\rangle}{\pi} = \frac{\arccos(X_{ij})}{\pi} \geq 0.87856\frac{1 - X_{ij}}{2}.$$



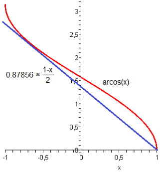
Porque funciona?

Dados dois vetores p_v e p_w qual a probabilidade de (v, w) estar no corte?



Probabilidade da aresta v, w estar no corte é

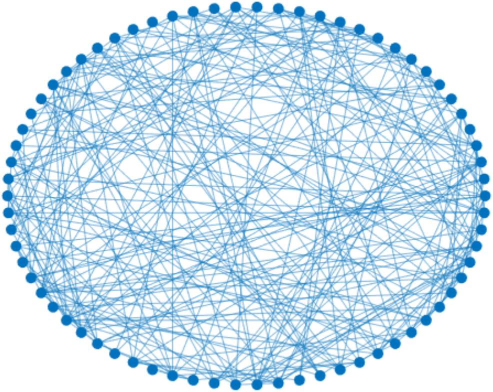
$$\frac{\theta}{\pi} = \frac{\arccos\langle u, w\rangle}{\pi} = \frac{\arccos(X_{ij})}{\pi} \geq 0.87856 \frac{1 - X_{ij}}{2}.$$



O número esperado de arestas é 0.87856 o valor da relaxação.

Exemplo revisitado

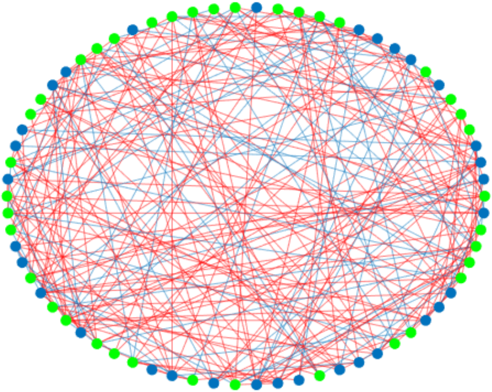
Grafo aleatório com 70 vértices e 250 arestas.



Ótimo da relaxação é 195.56.

Exemplo revisitado

Grafo aleatório com 70 vértices e 250 arestas.



Ótimo da relaxação é 195.56.

Corte com $176 > 171.8 = 0.87856 \times 195.56$ arestas

Capítulo 3

Para que serve?

2. Otimização polinomial

Positividade

Dado um polinómio **real**, como verificar se é sempre não negativo?

Positividade

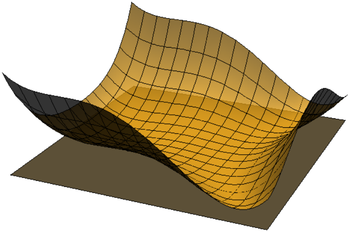
Dado um polinómio **real**, como verificar se é sempre não negativo?

Será $x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ não negativo?

Positividade

Dado um polinómio **real**, como verificar se é sempre não negativo?

Será $x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ não negativo?



Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2$$

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2 = \sum_i (\hat{h}_i^t X_d)^2$$

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2 = \sum_i (\hat{h}_i^t X_d)^2 = \sum_i X_d^t h_i \hat{h}_i^t X_d$$

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2 = \sum_i (\hat{h}_i^t X_d)^2 = \sum_i X_d^t h_i \hat{h}_i^t X_d = X_d^t \left(\sum_i h_i \hat{h}_i^t \right) X_d$$

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2 = \sum_i (\hat{h}_i^t X_d)^2 = \sum_i X_d^t h_i \hat{h}_i^t X_d = X_d^t \left(\sum_i h_i \hat{h}_i^t \right) X_d = X_d^t H X_d$$

com $H \succeq 0$.

Somas de quadrados

Para certificar não negatividade neste caso basta escrever

$$x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4 = (x^2 - y^2 + y)^2 + (xy - x)^2.$$

- Um polinómio soma de quadrados (sdq) é sempre não negativo.
- Mas como verificamos se é sdq?

Seja grau$(p) = 2d$ e X_d o vetor dos monómios de grau $\leq d$.

$$p = \sum_i h_i(x)^2 = \sum_i (\hat{h}_i^t X_d)^2 = \sum_i X_d^t h_i \hat{h}_i^t X_d = X_d^t \left(\sum_i h_i \hat{h}_i^t \right) X_d = X_d^t H X_d$$

com $H \succeq 0$. Assim p é sdq se e só se $p(x) = X_d^t H X_d$ para algum $H \succeq 0$.

Exemplo

$p(x) = x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ tem grau 4 logo $d = 2$. Temos:

$$X_2 = [1, x, y, x^2, xy, y^2]^t.$$

Exemplo

$p(x) = x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ tem grau 4 logo $d = 2$. Temos:

$$X_2 = [1, x, y, x^2, xy, y^2]^t.$$

$p(x)$ é soma de quadrados sse existir $H \succeq 0$ tal que

$$p(x) = \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} & h_{16} \\ h_{12} & h_{22} & h_{23} & h_{24} & h_{25} & h_{26} \\ h_{13} & h_{23} & h_{33} & h_{34} & h_{35} & h_{36} \\ h_{14} & h_{24} & h_{34} & h_{44} & h_{45} & h_{46} \\ h_{15} & h_{25} & h_{35} & h_{45} & h_{55} & h_{56} \\ h_{16} & h_{26} & h_{36} & h_{46} & h_{56} & h_{66} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}.$$

Exemplo

$p(x) = x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ tem grau 4 logo $d = 2$. Temos:

$$X_2 = [1, x, y, x^2, xy, y^2]^t.$$

$p(x)$ é soma de quadrados sse existir $H \succeq 0$ tal que

$$p(x) = \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} & h_{16} \\ h_{12} & h_{22} & h_{23} & h_{24} & h_{25} & h_{26} \\ h_{13} & h_{23} & h_{33} & h_{34} & h_{35} & h_{36} \\ h_{14} & h_{24} & h_{34} & h_{44} & h_{45} & h_{46} \\ h_{15} & h_{25} & h_{35} & h_{45} & h_{55} & h_{56} \\ h_{16} & h_{26} & h_{36} & h_{46} & h_{56} & h_{66} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}.$$

$$p(x) = h_{11} + 2h_{12}x + 2h_{13}y + (2h_{14} + h_{22})x^2 + (2h_{15} + 2h_{23})xy + \cdots$$

Expandindo são equações lineares nas entradas de H logo temos um programa semidefinido que podemos resolver.

Exemplo

$p(x) = x^2 + x^4 + y^2 - x^2y^2 - 2y^3 + y^4$ tem grau 4 logo $d = 2$. Temos:

$$X_2 = [1, x, y, x^2, xy, y^2]^t.$$

$p(x)$ é soma de quadrados sse existir $H \succeq 0$ tal que

$$p(x) = \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}.$$

$$p(x) = h_{11} + 2h_{12}x + 2h_{13}y + (2h_{14} + h_{22})x^2 + (2h_{15} + 2h_{23})xy + \cdots$$

Expandindo são equações lineares nas entradas de H logo temos um programa semidefinido que podemos resolver.

Hilbert

Theorem (Hilbert 1888)

Para polinómios de grau $2d$ *em* n *variáveis não negatividade e sdq são equivalentes se e só se uma das seguintes condições se verificar:*

Hilbert

Theorem (Hilbert 1888)

Para polinómios de grau $2d$ em n variáveis não negatividade e sdq são equivalentes se e só se uma das seguintes condições se verificar:

- $n = 1$;

Theorem (Hilbert 1888)

Para polinómios de grau $2d$ em n variáveis não negatividade e sdq são equivalentes se e só se uma das seguintes condições se verificar:

- $n = 1$;
- $d = 1$;

Theorem (Hilbert 1888)

Para polinómios de grau $2d$ em n variáveis não negatividade e sdq são equivalentes se e só se uma das seguintes condições se verificar:

- $n = 1$;
- $d = 1$;
- $n = 2$ and $d = 2$.

Hilbert

Theorem (Hilbert 1888)

Para polinómios de grau $2d$ em n variáveis não negatividade e sdq são equivalentes se e só se uma das seguintes condições se verificar:

- $n = 1$;
- $d = 1$;
- $n = 2$ *and* $d = 2$.

Theorem (Artin 1927 - *17.º Problema de Hilbert*)

Um polinómio é não negativo se e só se for soma de quadrados de funções racionais.

Exemplo de Motzkin

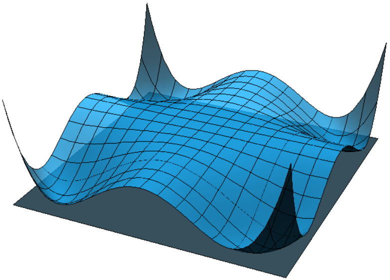
Em 1967 Motzkin deu o primeiro exemplo explícito de polinómio não negativo não sdq.

$$M(x, y) = x^4y^2 + y^4x^2 + 1 - 3x^2y^2.$$

Exemplo de Motzkin

Em 1967 Motzkin deu o primeiro exemplo explícito de polinómio não negativo não sdq.

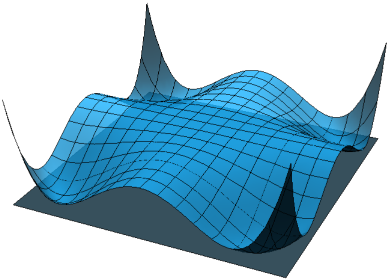
$$M(x,y) = x^4y^2 + y^4x^2 + 1 - 3x^2y^2.$$



Exemplo de Motzkin

Em 1967 Motzkin deu o primeiro exemplo explícito de polinómio não negativo não sdq.

$$M(x, y) = x^4y^2 + y^4x^2 + 1 - 3x^2y^2.$$



$$M(x, y) = (x^2 + y^2 + 1)\left(\frac{x^3y + xy^3 - 2xy}{x^2 + y^2}\right)^2 + \left(\frac{x^2 - y^2}{x^2 + y^2}\right)^2.$$

Otimização

Como usar esta ideia na minimização de polinómios?

Otimização

Como usar esta ideia na minimização de polinómios?

Otimização polinomial global

$p_{\min} = \min p(x)$ sobre todo $x \in \mathbb{R}^n$.

Como usar esta ideia na minimização de polinómios?

Otimização polinomial global

$p_{\min} = \min p(x)$ sobre todo $x \in \mathbb{R}^n$.

Transformamos num problema de não negatividade.

Otimização polinomial global 2.0

$p_{\min} = \max \lambda$ tal que $p(x) - \lambda \geq 0$ para todo $x \in \mathbb{R}^n$.

Otimização

Como usar esta ideia na minimização de polinómios?

Otimização polinomial global

$p_{\min} = \min p(x)$ sobre todo $x \in \mathbb{R}^n$.

Transformamos num problema de não negatividade.

Otimização polinomial global 2.0

$p_{\min} = \max \lambda$ tal que $p(x) - \lambda \geq 0$ para todo $x \in \mathbb{R}^n$.

Agora substituimos a não negatividade.

Otimização polinomial aproximada

$p_{\mathrm{sdq}} = \max \lambda$ tal que $p(x) - \lambda$ é sdq.

Otimização

Como usar esta ideia na minimização de polinómios?

Otimização polinomial global

$p_{\min} = \min p(x)$ sobre todo $x \in \mathbb{R}^n$.

Transformamos num problema de não negatividade.

Otimização polinomial global 2.0

$p_{\min} = \max \lambda$ tal que $p(x) - \lambda \geq 0$ para todo $x \in \mathbb{R}^n$.

Agora substituimos a não negatividade.

Otimização polinomial aproximada

$p_{\text{sdq}} = \max \lambda$ tal que $p(x) - \lambda$ é sdq.

Este é um problema semidefinido e dá-nos um minorante para $p_{\min}$.

Otimização 2

E se quisermos realmente o mínimo e não um minorante?

Otimização 2

E se quisermos realmente o mínimo e não um minorante?

Otimização polinomial aproximada

$p_{\text{sdq}} = \max \lambda$ tal que $p(x) - \lambda$ é sdq.

Otimização 2

E se quisermos realmente o mínimo e não um minorante?

Otimização polinomial aproximada

$p^k_{\text{sdq}} = \max \lambda$ tal que $(x_1^2 + x_2^2 + \cdots x_n^2)^k p(x) - \lambda$ é sdq.

Otimização 2

E se quisermos realmente o mínimo e não um minorante?

Otimização polinomial aproximada

$p^k_{\text{sdq}} = \max \lambda$ tal que $(x_1^2 + x_2^2 + \cdots x_n^2)^k p(x) - \lambda$ é sdq.

Usando geometria algébrica real (Positivstellensatz) podemos provar que

$$\lim_{k \to \infty} p^k_{\text{sdq}} = p_{\min}$$

Otimização 2

E se quisermos realmente o mínimo e não um minorante?

Otimização polinomial aproximada

$p^k_{\text{sdq}} = \max \lambda$ tal que $(x_1^2 + x_2^2 + \cdots x_n^2)^k p(x) - \lambda$ é sdq.

Usando geometria algébrica real (Positivstellensatz) podemos provar que

$$\lim_{k \to \infty} p^k_{\text{sdq}} = p_{\text{min}}$$

Não é difícil adaptar os métodos para trabalhar com desigualdades e equações polinomiais. Lasserre (2001) e Parrilo (2000) popularizaram esta ideia.

Obrigado pela vossa atenção

jgouveia@mat.uc.pt