# Semidefinite Programming

#### Semidefinite Programming

Fin Bauer and Stefan Heidekroger

21. Mai 2015

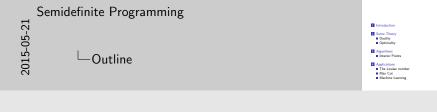
## Semidefinite Programming

Fin Bauer and Stefan Heidekrüger

21. Mai 2015

### Outline

- 1 Introduction
- 2 Some Theory
  - Duality
  - Optimality
- 3 Algorithms
  - Interior Points
- 4 Applications
  - The Lovász number
  - Max Cut
  - Max CutMachine Learning



# What Is Semidefinite Programming?

$$\min_{X} \langle C, X \rangle := Tr(CX) = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij}$$
s.t.  $\langle A_i, X \rangle = b_i, \quad i = 1, \dots, m$ 
 $X \succeq 0$ 

C,  $A_i$ ,  $X \in \mathbb{R}^{n \times n}$  symmetric  $X \succeq 0 \stackrel{\circ}{=} X$  is positive semidefinite (p.s.d.)



$$C = \begin{pmatrix} 1 & 2 \\ 2 & 9 \end{pmatrix}, \ A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \ A_2 = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}, \ b_1 = 11, \ b_2 = 19$$

Semidefinite Programming

—An Easy Example

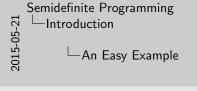
-Introduction

 $C = \begin{pmatrix} 1 & 2 \\ 2 & 9 \end{pmatrix}$ ,  $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ ,  $A_2 = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$ ,  $b_1 = 11$ ,  $b_2 = 19$ 

# An Easy Example

$$C = \begin{pmatrix} 1 & 2 \\ 2 & 9 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}, b_1 = 11, b_2 = 19$$

min 
$$x_{11} + 2x_{21} + 2x_{12} + 9x_{22}$$
  
s.t.  $x_{11} + x_{22} = 11$   
 $2x_{21} + 2x_{12} + 3x_{22} = 19$   
 $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \succeq 0$ 





 $C = \begin{pmatrix} 1 & 2 \\ 2 & 9 \end{pmatrix}$ ,  $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ ,  $A_2 = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$ ,  $b_1 = 11$ ,  $b_2 = 19$ 

# Problem Session

Problem 1

Are Linear Programs (LP) and Convex Quadratically

Constrained Quadratic Programs (CQCQP) Semidefinite Programs (SDP)?

Problem 4

Problem 2

max cut

Problem 3

lovasz

Is a certain optimization problem a SDP and what does it model?

Semidefinite Programming -Introduction └─Problem Session

# Problem 1: Have I Ever Seen Semidefinite Programming Before?

Linear Programming

### Linear Program

 $\min_{x} b_0^T x + c_0$ s.t.  $b_i^T x + c_i \le 0$ , i in 1,...,n  $x \ge 0$ 

Hint

 ${\sf Diagonal\ Matrix}$ 

Convex Quadratically Constrained Quadratic Programming

### CQCQP

min  $x^{T}A_{0}x + b_{0}^{T}x + c_{0}$ s.t.  $x^{T}A_{i}x + b_{i}^{T}x + c_{i} \le 0$ , i in 1,...,n Hint

Given  $A_i = M_i^T M_i$ then  $x^T A_i x + b_i^T x + c_i \le 0$  $\begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - b_i^T x \end{pmatrix} \succeq 0$  ☐ Introduction
☐ Problem 1:
☐ Have I Ever Seen Semidefinite Programming
☐ Introduction
☐ Problem 1:
☐ Have I Ever Seen Semidefinite Programming

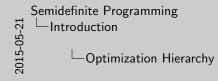
Semidefinite Programming

D-f--2



## Optimization Hierarchy

LP < CQCQP < SDP < Convex Programming



 $\mathsf{LP} < \mathsf{CQCQP} < \mathsf{SDP} < \mathsf{Convex Programming}$ 

# What's the Dual?

### Primal Problem in Standard Form

$$\inf_{X} \{ Tr(CX); Tr(A_{i}X) = b_{i} \ (i = 1, ..., m), \ X \in \mathcal{S}_{n}^{+} \}$$

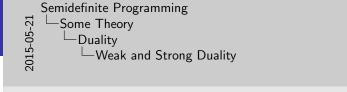
### Dual Problem in Standard Form

$$\sup_{y,S} \{ b^T y; \sum_{i=1}^m y_i A_i + S = C, \ S \in \mathcal{S}_n^+, y \in \mathbb{R}^m \}$$

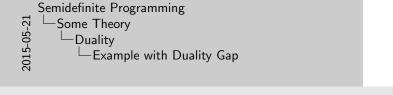
Semidefinite Programming
Some Theory
Duality
What's the Dual?



# Weak and Strong Duality



# Example with Duality Gap



## When is the Solution Optimal?

### Optimality Conditions

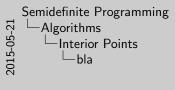
$$Tr(A_iX) = b_i, \quad X \succeq 0, \quad i = 1, ..., m$$

$$\sum_{i=1}^m y_i A_i + S = C, \quad S \succeq 0$$

$$XS = 0$$

Semidefinite Programming
Some Theory
Optimality
When is the Solution Optimal?





### The Lovász number

### Definition (SDP-variant)

Let G = (V, E) be a graph.

Semidefinite Programming

Applications

The Lovász number

The Lovász number

Definition (SDP-variant) Let G = (V, E) be a graph.

# Problem 2: What is MAX CUT?

Group presentation time.

What is MAX CUT?  $\max \frac{1}{2} \sum_{i < j} w_{ij} (1 - y_i y_j), \qquad s.t. y_i \in \{1, -1\}$ 

Semidefinite Programming

Problem 2:

-Applications

└─Max Cut

$$2 \sum_{i < j} q(x^{i})$$

Group presentation time.

# Problem 2: What is MAX CUT?

└─Applications └─Max Cut └─Problem 2: What is MAX CUT?

Semidefinite Programming

 $\max rac{1}{2} \sum_{i < j} w_{ij} (1-y_i y_j), \qquad s.t.y_i \in \{1,-1\}$ 

# IQP Model

Problem (MC) should be on the board now.

## An SDP-approximation algorithm (Goermans-Williamson)

#### Outline

- Relax (MC) into a QP (P)
- Find approximation bound of QP
- Show: equivalent SQP (SQ) to (P)
- Strong duality holds for (SQ) (ommitted)
- Solve SQP's dual in polynomial time



Outline

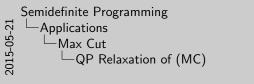
Relax (MC) into a QP (P)

- Find approximation bound of
- Strong duality holds for (SQ) (ommit

- 1995
- Approx: 0.87856
- eq. algorithm already existed, but bound wasn't known.

# QP Relaxation of (MC)

$$W_P^* := \max rac{1}{2} \sum_{i < j} w_{ij} (1 - v_i^{\mathsf{T}} v_j)$$
  $s.t. \ v_i \in \mathbb{S}^n \qquad orall i \in V$ 

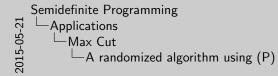




relaxation gives upper bound

## A randomized algorithm using (P)

- 1. Solve (P) to get vectors  $v_i$
- 2. Sample  $r \sim \mathit{UNIFORM}(\mathbb{S}^n)$
- 3. Set  $S := \{i | v_i^T r \ge 0\}$



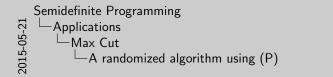
```
    Solve (P) to get vectors v;
    Sample r ~ UNIFORM(S")
    Set S := {i|v<sub>i</sub><sup>T</sup> r ≥ 0}
```

### A randomized algorithm using (P)

- 1. Solve (P) to get vectors  $v_i$
- 2. Sample  $r \sim UNIFORM(\mathbb{S}^n)$
- 3. Set  $S := \{i | v_i^T r \ge 0\}$

### For cut W obtained this way:

$$E[W] = \sum_{i < j} w_{ij} \frac{\operatorname{arccos}(v_i^T v_j)}{\pi}$$





 $E[W] = \sum_{i < j} w_{ij} \frac{\arccos(v_i^T v_j)}{\pi}$ 

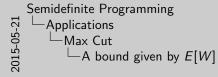
# A bound given by E[W]

### Theorem

$$E[W] \ge \alpha \frac{1}{2} \sum_{i < j} w_{ij} (1 - v_i^T v_j)$$

with

$$\alpha := \min_{0 \le \Theta \le \pi} \frac{2}{\pi} \frac{\Theta}{1 - \cos \Theta} > .87856\dots$$





# A bound given by E[W]

### Theorem

$$E[W] \ge \alpha \frac{1}{2} \sum_{i < j} w_{ij} (1 - v_i^T v_j)$$

with

$$\alpha := \min_{0 < \Theta < \pi} \frac{2}{\pi} \frac{\Theta}{1 - \cos \Theta} > .87856\dots$$

### Corollary

$$E[W] \ge \alpha W_P^* \ge \alpha W_{MC}^*$$

Semidefinite Programming

Applications

Max Cut

A bound given by E[W]



# SDP formulation of (P)

# (SD)

$$W_P^* := \max \frac{1}{2} \sum_{i < j} w_{ij} (1 - y_{ij})$$
 $s.t. \ y_{ii} = 1 \quad \forall i \in V$ 

Y sym. pos. sem. def.



# SDP formulation of (P)

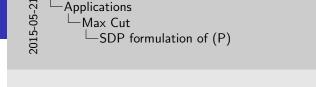
(SD)

$$egin{aligned} W_P^* := \max rac{1}{2} \sum_{i < j} w_{ij} (1 - y_{ij}) \ & ext{s.t. } y_{ii} = 1 \qquad orall i \in V \end{aligned}$$

 $Y_{ii} = 1$   $\forall i \in V$ Y sym. pos. sem. def.

How is this an SDP? Rewrite the objecti

How is this an SDP? Rewrite the objective! 
$$= \frac{1}{4} \sum_{i \in V} \sum_{j \in V} w_{ij} (1 - y_{ij})$$
 
$$= \frac{1}{2} W_{tot} - \frac{1}{4} \langle W, Y \rangle$$



Semidefinite Programming



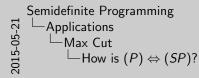
 $W_P^\star := \max \, \frac{1}{2} \sum_i w_{ij} (1-y_{ij})$ 

How is  $(P) \Leftrightarrow (SP)$ ?

■ Recall that a symmetric matrix  $A \in \mathbb{R}^n$  is positive semidefinite iff for some  $m \le n$ 

$$\exists B \in \mathbb{R}^{m \times n} : A = B^T B$$

- Given pos. semidef. A, such a B can be found in  $\mathcal{O}(n^3)$  using incomplete Cholesky decomposition.
- Interpret Y in (SP) as the Gram-Matrix of vectors  $v_i$  in (P)



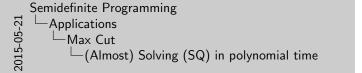
Recall that a symmetric matrix  $A \in \mathbb{R}^n$  is positive semidefinite iff for some  $m \leq n$ 

 $\exists B \in \mathbb{R}^{m \times n} : A = B^T E$ 

B Given pos. semidef. A, such a B can be found in O(n²) usin incomplete Cholesky decomposition.
B Interpret Y in (SP) as the Gram-Matrix of vectors v<sub>i</sub> in (P)

# (Almost) Solving (SQ) in polynomial time

- For this particular Problem, strong duality holds.
- Using the dual, a cut with weight at least  $W_{SQ}^* \varepsilon$  can be found in  $\mathcal{O}(\sqrt{n}(\log W_{tot} + \log \frac{1}{\varepsilon})$  iterations using an interior point algorithm. Each iteration can be implemented in  $\mathcal{O}(n^3)$
- This cut is a 0.878 approximation to  $W_{MC}^*$



For this particular Problem, strong duality holds.
 Using the dual, a cut with weight at least W<sup>\*</sup><sub>SQ</sub> − ε can be found in C(√R(log W<sub>cot</sub> + log ½) iterations using an interior

■ This cut is a 0.878 approximation to W\*...

Algo: Alizadeh 1995

## 'Quality' of the approximation

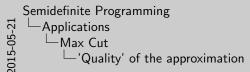
### Can $\alpha > 0.87856$ be improved?

No! The relaxation is tight.

- For  $C_5 : E[W] \approx .884 W_{MC}^*$
- For Peterson graph  $\approx$  .8787
- Examples are known such that  $E[W] < .8786W_{MC}^*$

### How does the algorithm do in practice?

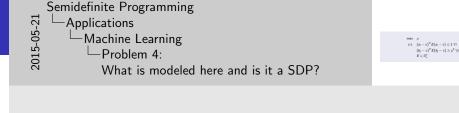
- Usually within 4% of  $W_M^*C$
- 'Almost always' within 9%





# Problem 4: What is modeled here and is it a SDP?

$$egin{array}{ll} \mathsf{max} & 
ho \ \mathsf{s.t.} & (a_i-c)^T E(a_i-c) \leq 1 \ orall i \ & (b_j-c)^T E(b_j-c) \geq 
ho^2 \ orall j \ & E \in \mathbb{S}^n_+ \end{array}$$



 $(b_j - c)^T E(b_j - c) \ge \rho^2 \forall j$   $E \in \mathbb{S}^n_+$ 

# Classification - Using SDP to tell two things apart (1)

### Ellipsoid

$$\mathcal{E} = \{ x \in \mathbb{R}^n; (x - c)^T E(x - c) <= 1, E \text{ is p.s.d.} \}$$

### First idea for SDP

$$egin{array}{ll} \mathsf{max} & 
ho \ & \mathsf{s.t.} & (\mathsf{a}_i-c)^\mathsf{T} \mathsf{E} (\mathsf{a}_i-c) \leq 1 \ orall i \ & (b_j-c)^\mathsf{T} \mathsf{E} (b_j-c) \geq 
ho^2 \ orall j \ & \mathsf{E} \in \mathbb{S}_+^n \end{array}$$

# Classification - Using SDP to tell two things apart (2)

max 
$$\rho$$
  
s.t.  $(1, a_i)^T \bar{E}(1, a_i) \leq 1 \ \forall i$   
 $(1, b_j)^T \bar{E}(1, b_j) \geq \rho^2 \ \forall j$   
 $E \in \mathbb{S}^{n+1}_+$ 

Semidefinite Programming

Applications

Machine Learning

Classification - Using SDP to tell two things
apart (2)

s.t.  $(1, a_i)^T \tilde{E}(1, a_i) \le 1 \ \forall i$