Semidefinite Programming

Fin Bauer and Stefan Heidekrüger

21. Mai 2015

Outline

What Is Semidefinite Programming?

$$
\begin{aligned}
\min_{X} \quad & \langle C, X \rangle := Tr(CX) = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} \\
\text{s.t.} \quad & \langle A_i, X \rangle \; = b_i, \quad i = 1, \ldots, m \\
& \qquad X \; \succeq 0
\end{aligned}
$$

C, A_i, $X \in \mathbb{R}^{nxn}$ symmetric
$X \succeq 0 \triangleq$ X is positive semidefinite (p.s.d.)

An Easy Example

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

An Easy Example

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

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

Problem Session

Problem 1

Are Linear Programs (LP) and Convex Quadratically Constrained Quadratic Programs (CQCQP) Semidefinite Programs (SDP)?

Problem 2

Model *MAXCUT* as an IQP.

Problem 3

Repitition of $\alpha(G), \omega(G), \chi(G)$

Problem 4

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

Problem 1:
Have I Ever Seen Semidefinite Programming Before?

- Linear Programming

Linear Program

$$\begin{aligned} \min_x \quad & b_0^T x + c_0 \\ \text{s.t.} \quad & b_i^T x + c_i \leq 0, \quad \text{i in 1,...,n} \\ & x \geq 0 \end{aligned}$$

Hint

Diagonal Matrix

- Convex Quadratically Constrained Quadratic Programming

CQCQP

$$\begin{aligned} \min \quad & x^T A_0 x + b_0^T x + c_0 \\ \text{s.t.} \quad & x^T A_i x + b_i^T x + c_i \leq 0, \quad \text{i in 1,...,n} \end{aligned}$$

Hint

$$\text{Given } A_i = M_i^T M_i$$

$$\text{then } x^T A_i x + b_i^T x + c_i \leq 0$$

$$\begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - b_i^T x \end{pmatrix} \succeq 0$$

Optimization Hierarchy

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

What's the Dual?

Primal Problem in Standard Form

$$\mathcal{P} = \inf_{X} \{ Tr(CX); Tr(A_i X) = b_i \ (i = 1, ..., m), \ X \in \mathcal{S}_n^+ \}$$

Dual Problem in Standard Form

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

Weak Duality

Duality Gap

Let $X \in \mathcal{P}$ and $(y, S) \in \mathcal{D}$. The quantity

$$\langle C, X \rangle - b^T y$$

is called the duality gap of $\mathcal{P} and \mathcal{D}$ at (X, y, S).

Weak Duality

Let $X \in \mathcal{P}$ and $(y, S) \in \mathcal{D}$. One has

$$\langle C, X \rangle - b^T y = \langle S, X \rangle \geq 0$$

Example with Duality Gap

Primal Problem

$$\min \; -x_2 \quad \text{s.t.} \quad \begin{pmatrix} x_2 - a & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{pmatrix} \preceq 0$$

Dual Problem

$$\max \; -aw_{11} \quad \text{s.t.} \quad \Omega \succeq 0, \; w_{22} = 0, w_{11} + 2w_{23} = 1$$

Strong Duality

Strict Feasibility

There exists $X \in \mathcal{P}$ and $S \in \mathcal{D}$ such that $X \prec 0$ and $S \prec 0$.

Strong Duality

Let $\mathcal{P}$ and $\mathcal{D}$ be strictly feasible. Then the duality gap is zero and the optimal sets of both the primal and the dual solution are nonempty.

When is the Solution Optimal?

Optimality Conditions

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

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

$$XS = 0$$

How to solve it?

Interior Point Algorithm

$$\min_{X} \langle C, X \rangle - \mu \log \det(X); \ \langle A_i, X \rangle = b_i \quad (i = 1; \ldots, m)$$

The Lovász number

Definition (SDP-variant)

Let $G = (V, E)$ be a graph. Then the Lovász-number of its complement $\bar{G}$ is defined by

$$
\begin{aligned}
\vartheta(\bar{G}) := \max_X \; & \langle ee^T, X \rangle = e^T X e \\
s.t. \; & x_{ij} = 0, \quad \forall i \neq j : (i,j) \notin E \\
& Tr(X) = 1 \\
& X \text{ sym. pos. sem. def.}
\end{aligned}
$$

The sandwich theorem

Theorem

$$\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$$

The sandwich theorem

Theorem

$$\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$$

Example: Pentagon

$$\omega(G) = ? \qquad \vartheta(\bar{G}) = \sqrt{5} \qquad \chi(G) = ?$$

The sandwich theorem

Theorem

$$\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$$

Example: Pentagon

$$\omega(G) = ? \qquad \vartheta(\bar{G}) = \sqrt{5} \qquad \chi(G) = ?$$

Proof idea

SQP-relaxations of MAX CLIQUE (easy) and MIN COLORING (not so easy), respectively.

So what?

- Both the clique and the coloring problems are NP-complete, but ϑ can be found in polynomial time (using ellipsoid methods)
- This gives an n-approximation to ω (and thus α) and χ
- It can be shown that no $(n - \varepsilon)$-approximation is possible in polynomial time, unless $NP = ZPP$

Problem 2:
What is MAX CUT?

Group presentation time.

Problem 2:
What is MAX CUT?

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 (P)
- Show: (P) equivalent to SDP (SP)
- Strong duality holds for (SP) (ommitted)
- Solve SQP's dual in polynomial time

QP Relaxation of (MC)

(P)

$$W_P^* := \max \frac{1}{2} \sum_{i<j} w_{ij}(1 - v_i^T v_j)$$

$$s.t.\ v_i \in \mathbb{S}^n \qquad \forall i \in V$$

A randomized algorithm using (P)

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

A randomized algorithm using (P)

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

For cut W obtained this way:

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

A bound given by $E[W]$

Theorem

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

with

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

A bound given by $E[W]$

Theorem

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

with

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

Corollary

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

SDP formulation of (P)

(SD)

$$W_P^* := \max \frac{1}{2} \sum_{i<j} w_{ij}(1 - y_{ij})$$

$$s.t. \; y_{ii} = 1 \qquad \forall i \in V$$

$$Y \text{ sym. pos. sem. def.}$$

SDP formulation of (P)

(SD)

$$W_P^* := \max \frac{1}{2} \sum_{i<j} w_{ij}(1 - y_{ij})$$

$$s.t. \; y_{ii} = 1 \qquad \forall i \in V$$

$$Y \text{ sym. pos. sem. def.}$$

How is this an SDP? Rewrite the objective!

$$= \frac{1}{4} \sum_{i \in V} \sum_{j \in V} w_{ij}(1 - y_{ij})$$

$$= \frac{1}{2} W_{tot} - \frac{1}{4} \langle W, Y \rangle$$

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

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

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

- 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)

(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}

'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 $E[W] < .8786 W^*_{MC}$

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?

$$
\begin{aligned}
\max \quad & \rho \\
\text{s.t.} \quad & (a_i - c)^T E (a_i - c) \leq 1 \ \forall i \\
& (b_j - c)^T E (b_j - c) \geq \rho^2 \ \forall j \\
& E \in \mathbb{S}_+^n
\end{aligned}
$$

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

$$\begin{aligned} \max \quad & \rho \\ \text{s.t.} \quad & (a_i - c)^T E (a_i - c) \leq 1 \ \forall i \\ & (b_j - c)^T E (b_j - c) \geq \rho^2 \ \forall j \\ & E \in \mathbb{S}_+^n \end{aligned}$$

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

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