Lecture 16: Semidefinite Programming and Duality Theorems

Rafael Oliveira

University of Waterloo
Cheriton School of Computer Science

rafael.oliveira.teaching@gmail.com

July 6, 2021

Overview

- Part I
  - Administrivia
  - Why Semidefinite Programming?
  - Convex Algebraic Geometry
- Part II
  - Duality Theory
- Conclusion
- Acknowledgements

Career Workshop this Thursday!

- open to *all CS students* who are curious about the variety of careers to which they can apply their computing skills.
- speakers who hold careers in medical imaging, digital forensics, fintech, gaming/VR, film and entertainment, and *graduate studies/research*!
- Registration is via Eventbrite
  `https://www.eventbrite.com/e/161507745013`

Mathematical Programming

Mathematical Programming deals with problems of the form

Mathematical Programming

Mathematical Programming deals with problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \geq 0 \\ & \quad \vdots \\ & g_m(x) \geq 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

Mathematical Programming

Mathematical Programming deals with problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \geq 0 \\ & \quad \vdots \\ & g_m(x) \geq 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

- Very general family of problems.

Mathematical Programming

Mathematical Programming deals with problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \geq 0 \\ & \vdots \\ & g_m(x) \geq 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

- Very general family of problems.
- Special case when all $f, g_1, \ldots, g_m$ are *linear*. *Linear Programming*

Mathematical Programming

Mathematical Programming deals with problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \geq 0 \\ & \vdots \\ & g_m(x) \geq 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

- Very general family of problems.
- Special case when all $f, g_1, \ldots, g_m$ are *linear*. *Linear Programming*
- More general case: *Semidefinite Programming*
  1. $A_1, \ldots, A_n, B \in \mathcal{S}^m$ are $m \times m$ symmetric matrices

Mathematical Programming

Mathematical Programming deals with problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \geq 0 \\ & \vdots \\ & g_m(x) \geq 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

- Very general family of problems.
- Special case when all $f, g_1, \ldots, g_m$ are *linear*. *Linear Programming*
- More general case: *Semidefinite Programming*
  1. $A_1, \ldots, A_n, B \in \mathcal{S}^m$ are $m \times m$ symmetric matrices
  2. Constraints:

     $$x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B$$

  3. Minimize linear function $c^T x$

What is a Semidefinite Program?

- $\mathcal{S}^m := \mathcal{S}^m(\mathbb{R})$ space of all $m \times m$ symmetric matrices (real entries)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad m=2$$

What is a Semidefinite Program?

- $\mathcal{S}^m := \mathcal{S}^m(\mathbb{R})$ space of all $m \times m$ symmetric matrices (real entries)
- $A \in \mathcal{S}^m$ is *positive semidefinite* (PSD), denoted $A \succeq 0$ if

What is a Semidefinite Program?

- $\mathcal{S}^m := \mathcal{S}^m(\mathbb{R})$ space of all $m \times m$ symmetric matrices (real entries)
- $A \in \mathcal{S}^m$ is *positive semidefinite* (PSD), denoted $A \succeq 0$ if
  1. all eigenvalues of A are *non-negative*
  2. $A = Y^T Y$ for some $Y \in \mathbb{R}^{d \times m}$ where $d \leq m$
  3. $z^T A z \geq 0$ for any $z \in \mathbb{R}^m$
  4. and more...

Spectral theorem ⇒ all symmetric matrices have real eigenvalues

$A \succeq 0 \iff \lambda_i(A) \geq 0 \quad \forall\ i \in [m]$

g_i: constraints in SDP

What is a Semidefinite Program?

- $\mathcal{S}^m := \mathcal{S}^m(\mathbb{R})$ space of all $m \times m$ symmetric matrices (real entries)
- $A \in \mathcal{S}^m$ is *positive semidefinite* (PSD), denoted $A \succeq 0$ if
  1. all eigenvalues of A are *non-negative*
  2. $A = Y^T Y$ for some $Y \in \mathbb{R}^{d \times m}$ where $d \leq m$
  3. $z^T A z \geq 0$ for any $z \in \mathbb{R}^m$
  4. and more...

Semidefinite Programming deals with problems of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^m \end{array}$$

linear function

Positive semidefinite constraint

What is a Semidefinite Program?

- $\mathcal{S}^m := \mathcal{S}^m(\mathbb{R})$ space of all $m \times m$ symmetric matrices (real entries)
- $A \in \mathcal{S}^m$ is *positive semidefinite* (PSD), denoted $A \succeq 0$ if
  1. all eigenvalues of A are *non-negative*
  2. $A = Y^T Y$ for some $Y \in \mathbb{R}^{d \times m}$ where $d \leq m$
  3. $z^T A z \geq 0$ for any $z \in \mathbb{R}^m$
  4. and more...

Semidefinite Programming deals with problems of the form

$$\begin{array}{rl} \text{minimize} & c^T x \\ \text{subject to} & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^n \end{array}$$

Where we use $C \succeq D$ to denote that $C - D \succeq 0$ (i.e., $C - D$ is PSD).

How does it generalize Linear Programming?

Linear Programming

$$\begin{aligned} \text{minimize} \quad & a^T x \\ \text{subject to} \quad & Cx \geq b \\ & x \in \mathbb{R}^n \end{aligned}$$

How does it generalize Linear Programming?

Linear Programming		**Semidefinite Programming**	
minimize	$a^T x$	minimize	$c^T x$
subject to	$Cx \geq b$	subject to	$x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B$
	$x \in \mathbb{R}^n$		$x \in \mathbb{R}^n$

Linear Programming $\subset$ Semidefinite Programming

idea: encode each linear constraint of LP into a diagonal entry of SDP constraint.

for each row $i \in [m]$ $\sum_{j=1}^{n} C_{ij} x_j \geq b_i$

How does it generalize Linear Programming?

Linear Programming		**Semidefinite Programming**	
minimize	$a^T x$	minimize	$c^T x$
subject to	$Cx \geq b$	subject to	$x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B$
	$x \in \mathbb{R}^n$		$x \in \mathbb{R}^n$

Set A_i's to be diagonal matrices, and $B = diag(b_1, \ldots, b_m)$

$\sum c_{ij} x_j \geq b_i$

$B = \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_m \end{pmatrix}$ symmetric

$A_i = \begin{pmatrix} c_{1i} & & & \\ & c_{2i} & & \\ & & \ddots & \\ & & & c_{mi} \end{pmatrix}$

(diagonal ∴ symmetric)

$x_1 A_1 + \cdots + x_n A_n \succeq B$ iff

$\begin{pmatrix} \sum c_{ij} x_j - b_1 & & \\ & \ddots & \\ & & \sum c_{mj} x_j - b_m \end{pmatrix} \succeq 0$

non-negative

Why should I care?

- Linear Programs appear everywhere in life: many problems of interest (resource allocation problems) can be modelled as linear program!

Why should I care?

- Linear Programs appear everywhere in life: many problems of interest (resource allocation problems) can be modelled as linear program!
- Semidefinite Programming is no different!

Why should I care?

- Linear Programs appear everywhere in life: many problems of interest (resource allocation problems) can be modelled as linear program!
- Semidefinite Programming is no different!
  - equilibrium analysis of dynamics and control (flight controls, robotics, etc.)
  - robust optimization
  - statistics and ML
  - continuous games
  - software verification
  - filter design
  - quantum computation and information
  - automated theorem proving
  - packing problems
  - many more

Why should I care?

- Linear Programs appear everywhere in life: many problems of interest (resource allocation problems) can be modelled as linear program!
- Semidefinite Programming is no different!
  - equilibrium analysis of dynamics and control (flight controls, robotics, etc.)
  - robust optimization
  - statistics and ML
  - continuous games
  - software verification
  - filter design
  - quantum computation and information
  - automated theorem proving
  - packing problems
  - many more
- See more here

`https://windowsontheory.org/2016/08/27/proofs-beliefs-and-algorithms-through-the-lens-of-sum-of-squares/`

Important Questions

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^n \end{aligned}$$

Important Questions

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\
& x \in \mathbb{R}^n
\end{aligned}
$$

1. When is a Semidefinite Program *feasible*?
   - Is there a solution to the constraints at all?

Important Questions

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^n \end{aligned}$$

1. When is a Semidefinite Program *feasible*?
   - Is there a solution to the constraints at all?
2. When is a Semidefinite Program *bounded*?
   - Is there a minimum? Is the minimum achievable? Or is the minimum $-\infty$?

Important Questions

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^n \end{aligned}$$

1. When is a Semidefinite Program *feasible*?
   - Is there a solution to the constraints at all?
2. When is a Semidefinite Program *bounded*?
   - Is there a minimum? Is the minimum achievable? Or is the minimum $-\infty$?
3. Can we characterize *optimality*?
   - How can we know that we found a minimum solution?
   - Do these solutions have nice description?
   - Do the solutions have *small bit complexity*?

Important Questions

$$\begin{array}{rl} \text{minimize} & c^T x \\ \text{subject to} & x_1 \cdot A_1 + \cdots + x_n \cdot A_n \succeq B \\ & x \in \mathbb{R}^n \end{array}$$

1. When is a Semidefinite Program *feasible*?
   - Is there a solution to the constraints at all?
2. When is a Semidefinite Program *bounded*?
   - Is there a minimum? Is the minimum achievable? Or is the minimum $-\infty$?
3. Can we characterize *optimality*?
   - How can we know that we found a minimum solution?
   - Do these solutions have nice description?
   - Do the solutions have *small bit complexity*?
4. How do we design *efficient algorithms* that find *optimal solutions* to Semidefinite Programs?

- Part I
  - Administrivia
  - Why Semidefinite Programming?
  - Convex Algebraic Geometry
- Part II
  - Duality Theory
- Conclusion
- Acknowledgements

Spectrahedra

To understand SDPs, we need to understand their *feasible regions*, which are called *spectrahedra* and described as *Linear Matrix Inequalities* (LMIs).

$$x_1 A_1 + \cdots + x_n A_n \succeq B$$

Spectrahedra

To understand SDPs, we need to understand their *feasible regions*, which are called *spectrahedra* and described as *Linear Matrix Inequalities* (LMIs).

Definition (Linear Matrix Inequalities)

A linear matrix inequality is an inequality of the form:

$$A_0 + \sum_{i=1}^{n} A_i x_i \succeq 0,$$

where $A_0, \ldots, A_n$ are *symmetric matrices*.

Spectrahedra

To understand SDPs, we need to understand their *feasible regions*, which are called *spectrahedra* and described as *Linear Matrix Inequalities* (LMIs).

Definition (Linear Matrix Inequalities)

A linear matrix inequality is an inequality of the form:

$$A_0 + \sum_{i=1}^{n} A_i x_i \succeq 0,$$

where $A_0, \ldots, A_n$ are *symmetric matrices*.

Definition (Spectrahedron)

A spectrahedron is a set defined by finitely many LMIs. In other words, it can be defined as:

$$S = \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} A_i x_i \succeq B, \quad A_i, B \in \mathcal{S}^m \right\}$$

Spectrahedra

To understand SDPs, we need to understand their *feasible regions*, which are called *spectrahedra* and described as *Linear Matrix Inequalities* (LMIs).

Suppose S defined by two LMIs $\begin{cases} \sum D_i x_i \succeq E \\ \sum F_i x_i \succeq G \end{cases}$

$$A_i = \begin{pmatrix} D_i & 0 \\ 0 & F_i \end{pmatrix} \qquad B = \begin{pmatrix} E & 0 \\ 0 & G \end{pmatrix}$$

$\rightarrow \left\{ \sum A_i x_i \succeq B \right.$

Definition (Spectrahedron)

A spectrahedron is a set defined by finitely many LMIs. In other words, it can be defined as:

$$S = \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} A_i x_i \succeq B, \quad A_i, B \in \mathcal{S}^m \right\}$$

Spectrahedra

To understand SDPs, we need to understand their *feasible regions*, which are called *spectrahedra* and described as *Linear Matrix Inequalities* (LMIs).

Definition (Spectrahedron)

A spectrahedron is a set defined by finitely many LMIs. In other words, it can be defined as:

$$S = \left\{ x \in \mathbb{R}^n \ \middle|\ \sum_{i=1}^{n} A_i x_i \succeq B, \quad A_i, B \in \mathcal{S}^m \right\}$$

Example of Spectrahedron

Polyhedron: $Cx \geq b$

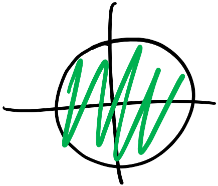
$$A_j = \begin{pmatrix} c_{1j} & & & \\ & c_{2j} & & \\ & & \ddots & \\ & & & c_{mj} \end{pmatrix} \qquad B = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_m \end{pmatrix}$$

Example of Spectrahedron

Circle: $\mathcal{C} = \{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1 \}$

$$\mathcal{C} = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix} 1+x & y \\ y & 1-x \end{pmatrix} \succeq 0 \right\}$$

$$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + x \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} + y \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \succeq 0$$



$1 + x \geq 0$

$1 - x \geq 0$

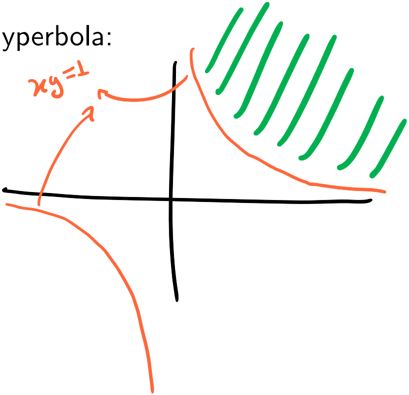
$|x| \le 1$

$$\det \begin{pmatrix} 1+x & y \\ y & 1-x \end{pmatrix} \geq 0$$

$$1 - x^2 - y^2 \geq 0 \iff x^2 + y^2 \le 1$$

Example of Spectrahedron

Hyperbola:



$$\mathcal{H} = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} x \geq 0,\ y \geq 0 \\ xy \geq 1 \end{array} \right\}$$

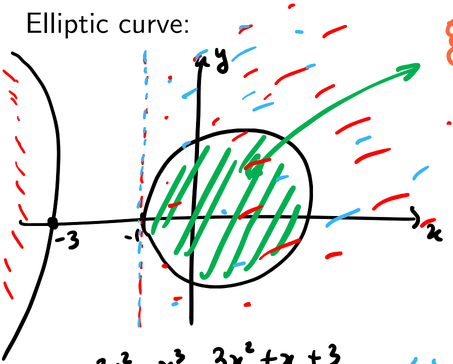
$$= \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix} \succeq 0 \right\}$$

$$x \geq 0,\ y \geq 0$$

$$\det \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix} \geq 0 \leftrightarrow xy - 1 \geq 0$$

Example of Spectrahedron

Elliptic curve:



$$0 = \underbrace{-2y^2 - x^3 - 3x^2 + x + 3}_{\det(A(x,y))}$$

$$\mathcal{E} = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; A(x,y) = \begin{bmatrix} x+1 & 0 & y \\ 0 & 2 & -x-1 \\ y & -x-1 & 2 \end{bmatrix} \succeq 0 \right\}$$

$A(x,y) \succeq 0$ iff $A(x,y)$ has only ≥ 0 eigenvalues

$\Leftrightarrow$ $\det(tI - A(x,y))$ has only ≥ 0 roots

$$\det(tI - A(x,y)) = t^3 - (x+5)t^2 + (-x^2 + 2x - y^2 + 7)t - \det(A(x,y))$$

≥ 0 inequalities determine the oval region.

Projected Spectrahedron

For both LPs and SDPs, it is enough to obtain a *linear projection of spectrahedron* (or polyhedron, if in LP).

Projected Spectrahedron

For both LPs and SDPs, it is enough to obtain a *linear projection of spectrahedron* (or polyhedron, if in LP).

Definition (Projected Spectrahedron)

A set $S \in \mathbb{R}^n$ is a *projected spectrahedron* if it has the form:

$$S = \left\{ x \in \mathbb{R}^n \;\middle|\; \exists y \in \mathbb{R}^t \text{ s.t. } \sum_{i=1}^{n} A_i x_i + \sum_{j=1}^{t} B_j y_j \succeq C, \quad A_i, B_j, C \in \mathcal{S}^m \right\}$$

S is projection of spectrahedron T

$T := \{ (x, y) \in \mathbb{R}^{n+t} \mid \sum A_i x_i + \sum B_j y_j \succeq C \}$

S has a semidefinite representation if S is a projection of a spectrahedron.

Projected Spectrahedron

For both LPs and SDPs, it is enough to obtain a *linear projection of spectrahedron* (or polyhedron, if in LP).

Definition (Projected Spectrahedron)

A set $S \in \mathbb{R}^n$ is a *projected spectrahedron* if it has the form:

$$S = \left\{ x \in \mathbb{R}^n \ \middle|\ \exists y \in \mathbb{R}^t \text{ s.t. } \sum_{i=1}^{n} A_i x_i + \sum_{j=1}^{t} B_j y_j \succeq C, \quad A_i, B_j, C \in \mathcal{S}^m \right\}$$

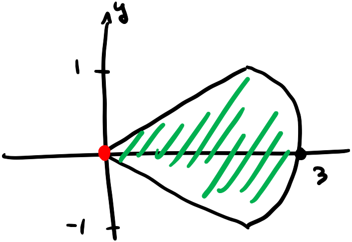
$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{s.t.} & z \in T \end{array} \iff \begin{array}{ll} \text{minimize} & \gamma^T x \\ \text{s.t.} & \begin{pmatrix} x \\ y \end{pmatrix} \in T \end{array} \iff x \in S$$

$c = \begin{pmatrix} \gamma \\ 0 \end{pmatrix}$ $\quad z = \begin{pmatrix} x \\ y \end{pmatrix}$

since S is a projection of T.

Example of Projected Spectrahedron

Projection quadratic cone intersected with halfspace:



$$S = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \exists\, z \in \mathbb{R} \text{ s.t. } \begin{pmatrix} z+y & 2z-x \\ 2z-x & z-y \end{pmatrix} \succeq 0,\ z \le 1 \right\}$$

Over $\mathbb{R}^3$, (x,y,z) would be given by $z \le 1$ (halfspace) intersect the cone $z^2 \ge y^2 + (2z-x)^2$

Convince yourself that the inequalities above $\Rightarrow x \ge 0$.

Remark: unlike in polyhedral case, projection of spectrahedron MAY NOT be a spectrahedron. (the above is an example)

Working with Symmetric Matrices

Definition (Frobenius Inner Product)

$A, B \in \mathcal{S}^m$, define the *Frobenius inner product* as

$$\langle A, B \rangle := \mathrm{tr}[AB] = \sum_{i,j} A_{ij} B_{ij}$$

- This is the "usual inner product" if you think of the matrices as vectors
- Thus, have the norm

$$\|A\|_F = \sqrt{\langle A, A \rangle} = \sqrt{\sum_{i,j} A_{ij}^2}$$

- With this norm, can talk about the *polar dual* to a given spectrahedron $S \subseteq \mathcal{S}^m$:

$$S^\circ = \{ Y \in \mathcal{S}^m \mid \langle Y, X \rangle \leq 1, \ \forall X \in S \}$$

Standard Primal Form

Just like in Linear Programming, we can represent SDPs in standard form:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Where now:

X symmetric matrix of variables

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} = x_{11}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_{12}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_{22}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Standard Primal Form

Just like in Linear Programming, we can represent SDPs in standard form:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Where now:

- the variables are encoded in a positive semidefinite matrix X,

Standard Primal Form

Just like in Linear Programming, we can represent SDPs in standard form:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

linear function over variables x_{ij}

Where now:

- the variables are encoded in a positive semidefinite matrix X,
- each constraint is given by an inner product $\langle A_i, X \rangle = b_i$

$tr(A_i X) = b_i$

Standard Primal Form

Just like in Linear Programming, we can represent SDPs in standard form:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Where now:

- the variables are encoded in a positive semidefinite matrix X,
- each constraint is given by an inner product $\langle A_i, X \rangle = b_i$
- Note the similarity with LP standard primal. Can obtain LP standard form by making X and A_i's to be diagonal

$x_{ij} = 0$ if $i \neq j$ $\langle Z_{ij}, X \rangle = 0$

Standard Primal Form

Just like in Linear Programming, we can represent SDPs in standard form:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Where now:

- the variables are encoded in a positive semidefinite matrix X,
- each constraint is given by an inner product $\langle A_i, X \rangle = b_i$
- Note the similarity with LP standard primal. Can obtain LP standard form by making X and A_i's to be diagonal
- How is that an LMI though?

Standard Primal Form as LMI

$$\begin{array}{ll} \text{minimize} & \langle C, X \rangle \\ \text{subject to} & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{array}$$

$$\langle A_i, X \rangle = b_i \qquad \sum_{k,l} (A_i)_{kl} \cdot x_{kl} = b_i$$

writing all equalities as diagonal matrices

$$\sum x_{kl} \underbrace{\begin{pmatrix} (A_1)_{kl} & & & \\ & (A_2)_{kl} & & \\ & & \ddots & \\ & & & (A_t)_{kl} \end{pmatrix}}_{D_{kl}} \succeq \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_t \end{pmatrix}$$

block diagonalize to get 1 LMI

$$\sum x_{kl} (-D_{kl}) \succeq \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_t \end{pmatrix}$$

$$X \succeq 0$$

Example

value (min) = $1 - \sqrt{2}$

Primal

$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

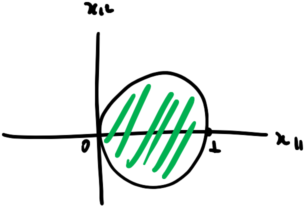
$b_1 = 1$

$C = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$

$$\begin{aligned} \text{minimize} \quad & 2x_{11} + 2x_{12} \\ \text{subject to} \quad & x_{11} + x_{22} = 1 \\ & \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \succeq 0 \end{aligned}$$

$X^* = \text{OPT} = \begin{bmatrix} \frac{2-\sqrt{2}}{4} & -\frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{2+\sqrt{2}}{4} \end{bmatrix}$

both OPT and value may not be rational!



Semidefinite Programming Duality

Consider our SDP:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Semidefinite Programming Duality

Consider our SDP:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

- If we look at what happens when we multiply i^{th} equality by a variable y_i:

$$\sum_{i=1}^{t} y_i \cdot \langle A_i, X \rangle = \sum_{i=1}^{t} y_i \cdot b_i \quad \Rightarrow \quad \left\langle \sum_{i=1}^{t} y_i A_i \ , \ X \right\rangle = y^T b$$

Semidefinite Programming Duality

Consider our SDP:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

- If we look at what happens when we multiply i^{th} equality by a variable y_i:

$$\sum_{i=1}^{t} y_i \cdot \langle A_i, X \rangle = \sum_{i=1}^{t} y_i \cdot b_i \quad \Rightarrow \quad \left\langle \sum_{i=1}^{t} y_i A_i \ , \ X \right\rangle = y^T b$$

- Thus, if $\sum_{i=1}^{t} y_i A_i \preceq C$, then we have:

$$y^T b = \left\langle \sum_{i=1}^{t} y_i A_i \ , \ X \right\rangle \leq \langle C, X \rangle$$

$\langle A, B \rangle \geq 0$ if $A, B \succeq 0$

Semidefinite Programming Duality

Consider our SDP:

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

- If we look at what happens when we multiply i^{th} equality by a variable y_i:

$$\sum_{i=1}^{t} y_i \cdot \langle A_i, X \rangle = \sum_{i=1}^{t} y_i \cdot b_i \quad \Rightarrow \quad \left\langle \sum_{i=1}^{t} y_i A_i \ , \ X \right\rangle = y^T b$$

- Thus, if $\sum_{i=1}^{t} y_i A_i \preceq C$, then we have:

$$y^T b = \left\langle \sum_{i=1}^{t} y_i A_i \ , \ X \right\rangle \leq \langle C, X \rangle$$

- $y^T b$ is a *lower bound* on the solution to our SDP!

Semidefinite Programming Duality

Consider the following SDPs:

Primal SDP		*Dual SDP*	
minimize	$\langle C, X \rangle$	maximize	$y^T b$
subject to	$\langle A_i, X \rangle = b_i$	subject to	$\sum_{i=1}^{t} y_i A_i \preceq C$
	$X \succeq 0$		

best lower bound

conditions we need for lower bound primal

Semidefinite Programming Duality

Consider the following SDPs:

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- From previous slide

$$\sum_{i=1}^{t} y_i A_i \preceq C \Rightarrow y^T b \text{ is a lower bound on value of Primal}$$

Semidefinite Programming Duality

Consider the following SDPs:

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- From previous slide

$$\sum_{i=1}^{t} y_i A_i \preceq C \Rightarrow y^T b \text{ is a lower bound on value of Primal}$$

- Thus, the optimal (maximum) value of *dual LP* lower bounds the optimal (minimum) value of the *Primal LP*!

Semidefinite Programming Duality

Consider the following SDPs:

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- From previous slide

$$\sum_{i=1}^{t} y_i A_i \preceq C \Rightarrow y^T b \text{ is a lower bound on value of Primal}$$

- Thus, the optimal (maximum) value of *dual LP* lower bounds the optimal (minimum) value of the *Primal LP*!

Theorem (Weak Duality)

Let X be a feasible solution of the primal SDP and y be a feasible solution of the dual SDP. Then

value of dual   value of primal

$$y^T b \leq \langle C, X \rangle.$$

Remarks on Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

Remarks on Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

Theorem (Complementary Slackness)

Let X be a feasible solution of the primal SDP and y be a feasible solution of the dual SDP. If (X, y) satisfy the complementary slackness condition

$$\left(C - \sum_{i=1}^{t} y_i A_i \right) X = 0$$

Then (X, y) are primal and dual optimum solutions of the SDP problem.

Remarks on Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

Theorem (Complementary Slackness)

Let X be a feasible solution of the primal SDP and y be a feasible solution of the dual SDP. If (X, y) satisfy the complementary slackness condition

$$\left(C - \sum_{i=1}^{t} y_i A_i \right) X = 0$$

Then (X, y) are primal and dual optimum solutions of the SDP problem.

Complementary slackness gives us *sufficient* conditions to check optimality of our solutions.

Strong Duality

Primal SDP

$$\begin{array}{ll} \text{minimize} & \langle C, X \rangle \\ \text{subject to} & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{array}$$

Dual SDP

$$\begin{array}{ll} \text{maximize} & y^T b \\ \text{subject to} & \sum_{i=1}^{t} y_i A_i \preceq C \end{array}$$

Strong Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- Strong duality in SDPs is a bit more complex than in LPs.

Strong Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- Strong duality in SDPs is a bit more complex than in LPs.
- Both primal and dual may be feasible, and yet strong duality may not hold!

Home work

Strong Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- Strong duality in SDPs is a bit more complex than in LPs.
- Both primal and dual may be feasible, and yet strong duality may not hold!
- But under mild conditions, strong duality holds!

Strong Duality

Primal SDP		*Dual SDP*	
minimize	$\langle C, X\rangle$	maximize	$y^T b$
subject to	$\langle A_i, X\rangle = b_i$	subject to	$\sum_{i=1}^{t} y_i A_i \preceq C$
	$X \succeq 0$		

- Strong duality in SDPs is a bit more complex than in LPs.
- Both primal and dual may be feasible, and yet strong duality may not hold!
- But under mild conditions, strong duality holds!
- Primal SDP is *strictly feasible* if there is feasible solution $X \succ 0$. strict inequality
- Dual SDP is *strictly feasible* if there is feasible $\sum_{i=1}^{t} y_i A_i \prec C$. strict inequality

Slater conditions

Strong Duality

Primal SDP

$$\begin{aligned} \text{minimize} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i \\ & X \succeq 0 \end{aligned}$$

Dual SDP

$$\begin{aligned} \text{maximize} \quad & y^T b \\ \text{subject to} \quad & \sum_{i=1}^{t} y_i A_i \preceq C \end{aligned}$$

- Strong duality in SDPs is a bit more complex than in LPs.
- Both primal and dual may be feasible, and yet strong duality may not hold!
- But under mild conditions, strong duality holds!
- Primal SDP is *strictly feasible* if there is feasible solution $X \succ 0$.
- Dual SDP is *strictly feasible* if there is feasible $\sum_{i=1}^{t} y_i A_i \prec C$.

Theorem (Strong Duality under Slater Conditions)

If primal SDP and dual SDP are both strictly feasible, then

$$\textit{max dual} \;=\; \textit{min of primal}.$$

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard

[1]pun intended

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard
- Special cases have very striking applications!
  - Linear Programming (previous lectures)
  - Today: *Semidefinite Programming*

[1] pun intended

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard
- Special cases have very striking applications!
  - Linear Programming (previous lectures)
  - Today: *Semidefinite Programming*
- Semidefinite Programming and Duality - fundamental concepts, lots of applications!
  - Applications in Combinatorial Optimization (Max-Cut in *next lecture*!)
  - Applications in Control Theory
  - many more!

[1] pun intended

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard
- Special cases have very striking applications!
  - Linear Programming (previous lectures)
  - Today: *Semidefinite Programming*
- Semidefinite Programming and Duality - fundamental concepts, lots of applications!
  - Applications in Combinatorial Optimization (Max-Cut in *next lecture*!)
  - Applications in Control Theory
  - many more!
- Check out connections to Sum of Squares and a **bold**[1] attempt to have one algorithm to solve all problems! (i.e., one algorithm to rule them all)

`https://windowsontheory.org/2016/08/27/proofs-beliefs-and-algorithms-through-the-lens-of-sum-of-squares/`

[1] pun intended

Acknowledgement

- Lecture based largely on:
    - [Blekherman, Parrilo, Thomas 2012, Chapter 2]

References I

Blekherman, Grigoriy and Parrilo, Pablo and Thomas, Rekha (2012)
Convex Algebraic Geometry