Foundation of Computing and Mathematical Science

— Optimization —

October 2009

Masakazu Kojima
Mituhiro Fukuda
Alex Fukunaga

This course will cover some recent topics in the theory and applications of optimizations. This year, topics will include semidefinite programming, gradient based methods, and constraint programming.

Class Schedule for Part I

Part I: Semidefinite Programming — Masakazu Kojima

10/02 — 1
10/09 — 2
10/16 — 3
10/23 — No class, Koudai-sai
10/30 — 4

Class Schedule for Part II and III

Part II: Gradient Based Methods — Mituhiro Fukuda

11/06 — 1
11/13 — 2
11/20 — 3
11/27 — 4
12/04 — 5

Part III: Constraint Programming— Alex Fukunaga

12/11 — 1
12/16 — 2
01/08 — 3
01/15 — No class, the national unified entrance examination
01/22 — 4
01/29 — 5

Examination on 02/05 or 02/12

Introduction to Semidefinite Programming

Masakazu Kojima, Tokyo Institute of Technology

October, 2009

http://www.is.titech.ac.jp/∼kojima/articles/IntroductionToSDP.pdf

Abstract

- The main purpose of this lecture is an introduction of semidefinite programs for graduate students and researchers who are not familiar to this subject and/or who want to look over SDPs quickly.
- Assuming the basics of linear programs and linear algebra, the lecture places the main emphasis on the basic theory of SDPs.
- Some examples and applications of SDPs are also presented to show the significance of SDPs in the field of optimization.

Contents

Contents

Chapter 1: Basic theory.

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form SDP
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Contents

Contents

Contents

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

SDP is an extension of LP to the space of symmetric matrices.

LP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7,\ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0,\ X_{12} \geq 0,\ X_{22} \geq 0.$

SDP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7,\ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0,\ X_{12} \geq 0,\ X_{22} \geq 0,$

$\begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} \succeq \boldsymbol{O}$ (positive semidefinite).

- Both LP and SDP have linear objective functions in real variables $X_{11},\ X_{12},\ X_{22}$.
- Both LP and SDP have linear equality and inequality constraints in real variables $X_{11},\ X_{12},\ X_{22}$.

SDP is an extension of LP to the space of symmetric matrices.

LP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7, \ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0, \ X_{12} \geq 0, \ X_{22} \geq 0.$

SDP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7, \ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0, \ X_{12} \geq 0, \ X_{22} \geq 0,$

$\begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} \succeq \boldsymbol{O}$ (positive semidefinite).

- SDP has a psd constraint in $\begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix}$, or $X_{11} \geq 0, \ X_{22} \geq 0, \ X_{11}X_{22} - X_{12}^2 \geq 0,$ which requires $X_{11}, \ X_{12}, \ X_{22}$ 'dependent nonlinearly', while $X_{11} \geq 0, \ X_{12} \geq 0, \ X_{22} \geq 0$ in LP are linear and separable.

SDP is an extension of LP to the space of symmetric matrices.

LP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7,\ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0,\ X_{12} \geq 0,\ X_{22} \geq 0.$

SDP: minimize $-X_{11} - 2X_{12} - 5X_{22}$

subject to $2X_{11} + 3X_{12} + X_{22} = 7,\ X_{11} + X_{12} \geq 1,$

$X_{11} \geq 0,\ X_{12} \geq 0,\ X_{22} \geq 0,$

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} \succeq \boldsymbol{O}$$ (positive semidefinite).

- The feasible region of LP and the feasible region of SDP are convex sets, but the former is polyhedral while the latter is non-polyhedral.

Exercise 1.

Draw a picture of the set $\left\{(X_{11}, X_{12}, X_{22}) : \begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} \succeq \boldsymbol{O}\right\}$.

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Lots of Applications to Various Problems

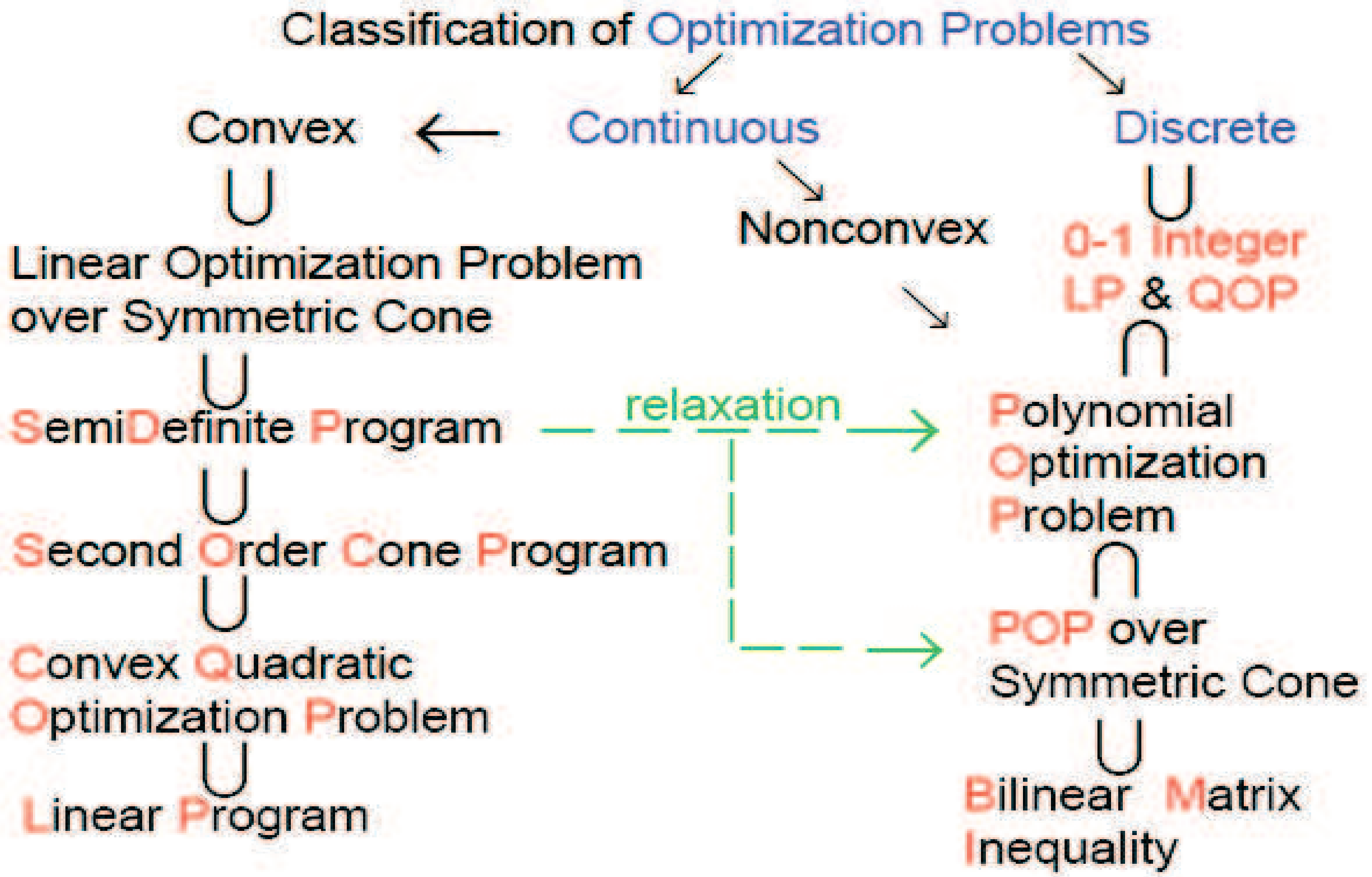
- Systems and control theory — Linear Matrix Inequality [6]
- SDP relaxations of combinatorial and nonconvex problems
  - Max cut and max clique problems [14]
  - 0-1 integer linear programs [24]
  - Polynomial optimization problems [22, 35]
- Robust optimization [4]
- Quantum chemistry [51]
- Moment problems (applied probability) [5, 23]
- . . .

Survey articles — Todd [39] ,Vandenberghe-Boyd [45]
Handbook of SDP — Wolkowicz-Saigal-Vandenberghe [46]
Web pages — Helmberg[15], Wolkowicz [47]

Theory

- Self-concordant theory [33]
- Euclidean Jordan algebra [10, 36]
- Polynomial-time primal-dual interior-point methods [1, 17, 20, 27, 34]

SDP serves as a core convex optimization problem


Classification of Optimization Problems
Convex
Continuous
Discrete
Nonconvex
Linear Optimization Problem over Symmetric Cone
0-1 Integer LP & QOP
SemiDefinite Program
relaxation
Polynomial Optimization Problem
Second Order Cone Program
POP over Symmetric Cone
Convex Quadratic Optimization Problem
Linear Program
Bilinear Matrix Inequality

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

(LP)	minimize	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$
	subject to	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (1 \leq p \leq m), \ \mathbb{R}^n \ni \boldsymbol{x} \geq \mathbf{0}$.

Here	$\mathbb{R}$	:	the set (linear space) of real numbers,
	$\mathbb{R}^n$	:	the linear space of n dim. vectors,
	$\boldsymbol{a}_p \in \mathbb{R}^n$	:	data, n dim. vector $(1 \leq p \leq m)$,
	$b_p \in \mathbb{R}$	:	data, real number $(1 \leq p \leq m)$,
	$\boldsymbol{x} \in \mathbb{R}^n$	:	variable, n dim. vector,
	$\boldsymbol{a}_p \cdot \boldsymbol{x}$	$=$	$\sum_{i=1}^n [\boldsymbol{a}_p]_i \boldsymbol{x}_i$ (the inner product of $\boldsymbol{a}_p$ and $\boldsymbol{x}$).

(LP)	minimize	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$
	subject to	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (1 \leq p \leq m), \ \mathbb{R}^n \ni \boldsymbol{x} \geq \mathbf{0}$.

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\mathbb{S}^n$: the linear space of $n \times n$ symmetric matrices,

$\boldsymbol{A}_p \in \mathbb{S}^n$: data, $n \times n$ symmetric matrix $(0 \leq p \leq m)$,

$b_p \in \mathbb{R}$: data, real number $(1 \leq p \leq m)$,

$\boldsymbol{X} \in \mathbb{S}^n$: $n \times n$ variable, symmetric matrix;

$$\boldsymbol{X} = (X_{ij}) = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{pmatrix} \in \mathbb{S}^n,$$

$X_{ij} = X_{ji} \in \mathbb{R} \ (1 \leq i \leq j \leq n)$, (LP) minimize $\boldsymbol{a}_0 \cdot \boldsymbol{x}$

subject to $\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (1 \leq p \leq m), \ \mathbb{R}^n \ni \boldsymbol{x} \geq \mathbf{0}$.

(SDP) minimize $\boldsymbol{A}_0 \bullet \boldsymbol{X}$

subject to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\boldsymbol{X} \in \mathbb{S}^n_+ \quad \Leftrightarrow \quad \boldsymbol{X} \in \mathbb{S}^n$ is positive semidefinite,

$\boldsymbol{X} \succeq \boldsymbol{O} \quad \Leftrightarrow \quad \boldsymbol{X} \in \mathbb{S}^n_+$ for some n,

$\boldsymbol{A}_p \bullet \boldsymbol{X} \quad = \quad \sum_{i=1}^n \sum_{j=1}^n [\boldsymbol{A}_p]_{ij} \boldsymbol{X}_{ij}$

(the inner product of $\boldsymbol{A}_p$ and $\boldsymbol{X}$).

(LP) minimize $\boldsymbol{a}_0 \cdot \boldsymbol{x}$
subject to $\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (1 \leq p \leq m), \ \mathbb{R}^n \ni \boldsymbol{x} \geq \mathbf{0}$.

(SDP) minimize $\boldsymbol{A}_0 \bullet \boldsymbol{X}$
subject to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$$\Uparrow \quad \left\{ \begin{array}{l} m = 2, \ n = 2, \ b_1 = 7, \ b_2 = 9, \\ \boldsymbol{X} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \ \boldsymbol{A}_0 = \begin{pmatrix} -1 & -1 \\ -1 & -5 \end{pmatrix}, \\ \boldsymbol{A}_1 = \begin{pmatrix} 2 & 1.5 \\ 1.5 & 1 \end{pmatrix}, \ \boldsymbol{A}_2 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 3 \end{pmatrix}. \end{array} \right.$$

minimize $-X_{11} - 2X_{12} - 5X_{22}$
subject to $2X_{11} + 3X_{12} + X_{22} = 7, \ 2X_{11} + X_{12} + 3X_{22} = 9,$

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix} \succeq \boldsymbol{O}.$$

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$: semidefinite constraint.

- Definition: $\boldsymbol{X} \succeq \boldsymbol{O}$ if
  $\boldsymbol{u}^T \boldsymbol{X} \boldsymbol{u} = \sum_{i=1}^n \sum_{j=1}^n X_{ij} u_i u_j \geq 0$ for $\forall \boldsymbol{u} \in \mathbb{R}^n$.
- Definition: $\boldsymbol{X} \succ \boldsymbol{O}$ if
  $\boldsymbol{u}^T \boldsymbol{X} \boldsymbol{u} = \sum_{i=1}^n \sum_{j=1}^n X_{ij} u_i u_j > 0$ for $\forall \boldsymbol{u} \neq \mathbf{0}$.
- $\boldsymbol{X} \in \mathbb{S}^n \Rightarrow$ all n e.values are real.
- $\boldsymbol{X} \succeq \boldsymbol{O}$ ($\succ \boldsymbol{O}$) $\Leftrightarrow$ all n e.values ≥ 0 (> 0).
- $\boldsymbol{X} \succeq \boldsymbol{O}$ ($\succ \boldsymbol{O}$) $\Leftrightarrow$ all principal minors ≥ 0 (> 0).
- $\boldsymbol{X} \succeq \boldsymbol{O}$ ($\succ \boldsymbol{O}$) $\Rightarrow$ all diagonal X_{ii}'s ≥ 0 (> 0).
- $\boldsymbol{X} \succeq \boldsymbol{O}$ and $X_{ii} = 0 \Rightarrow X_{ij} = 0 \ (\forall j)$.

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$: semidefinite constraint.

- $\boldsymbol{X} \succeq \boldsymbol{O}$ ($\succ \boldsymbol{O}$) $\Leftrightarrow \exists n \times n$ (nonsingular) $\boldsymbol{B}$; $\boldsymbol{X} = \boldsymbol{B}\boldsymbol{B}^T$ (factorization).
- $\boldsymbol{X} \succeq \boldsymbol{O} \Leftrightarrow \exists n \times n$ lower triang. $\boldsymbol{L}$; $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{L}^T$ (Cholesky factorization).
- $\boldsymbol{X} \succeq \boldsymbol{O} \Leftrightarrow \exists n \times n$ orthogonal P and $\exists n \times n$ diagonal $\boldsymbol{D}$;

  $$\boldsymbol{X} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T \text{ (orthogonal decompostion).}$$

  Here each diagonal element $\lambda_i = D_{ii}$ of D is an eigenvalue of $\boldsymbol{X}$ and each ith column p_i of P an eigenvector corresponding to λ_i.

- $\boldsymbol{X} \succeq \boldsymbol{O} \Leftrightarrow \exists \boldsymbol{C} \in \mathbb{S}^n_+$; $\boldsymbol{X} = \boldsymbol{C}^2 \Leftarrow$ Take $\boldsymbol{C} = \boldsymbol{P}(\boldsymbol{D})^{1/2}\boldsymbol{P}^T$;

  $$\boldsymbol{C}^2 = \left(\boldsymbol{P}(\boldsymbol{D})^{1/2}\boldsymbol{P}^T\right)\left(\boldsymbol{P}(\boldsymbol{D})^{1/2}\boldsymbol{P}^T\right) = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T = \boldsymbol{X}.$$

  We will write $\boldsymbol{X} = \left(\sqrt{X}\right)^2$.

$$\text{(SDP)} \quad \text{minimize} \quad \boldsymbol{A}_0 \bullet \boldsymbol{X}$$
$$\text{subject to} \quad \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}.$$

$\mathbb{S}^n$: a linear space with dimension $n(n+1)/2$.

- $\boldsymbol{X} + \boldsymbol{Y} \in \mathbb{S}^n$ for $\forall \boldsymbol{X} \in \mathbb{S}^n$ and $\forall \boldsymbol{Y} \in \mathbb{S}^n$.
- $\alpha \boldsymbol{X} \in \mathbb{S}^n$ for $\forall \alpha \in \mathbb{R}$ and $\forall \boldsymbol{X} \in \mathbb{S}^n$.
- linear independence.
- a basis consisting of $n(n+1)/2$.

Example. $n = 2$. Note that $X_{12} = X_{21}$.

$$2 \begin{pmatrix} 1.1 & -0.5 \\ -0.5 & 2.4 \end{pmatrix} + 0.5 \begin{pmatrix} 2.4 & 0.6 \\ 0.6 & 1.2 \end{pmatrix} = \begin{pmatrix} 3.4 & 0.7 \\ 0.7 & 5.4 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \ : \ \text{a basis of } \mathbb{S}^2.$$

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}.$

$\mathbb{S}^n$: a linear space with dimension $n(n+1)/2$.

- $\boldsymbol{X} + \boldsymbol{Y} \in \mathbb{S}^n$ for $\forall \boldsymbol{X} \in \mathbb{S}^n$ and $\forall \boldsymbol{Y} \in \mathbb{S}^n$.
- $\alpha \boldsymbol{X} \in \mathbb{S}^n$ for $\forall \alpha \in \mathbb{R}$ and $\forall \boldsymbol{X} \in \mathbb{S}^n$.
- linear independence.
- a basis consisting of $n(n+1)/2$.
- For every $\boldsymbol{A}, \ \boldsymbol{X} \in \mathbb{S}^n$, the inner product $\boldsymbol{A} \bullet \boldsymbol{X}$ is defined;

$$\begin{aligned} \boldsymbol{A} \bullet \boldsymbol{X} &= \sum_{i=1}^n \left(\sum_{j=1}^n A_{ij} X_{ij} \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n A_{ij} X_{ji} \right) = \text{trace } \boldsymbol{A}\boldsymbol{X}. \end{aligned}$$

(i,i)th element of $\boldsymbol{A}\boldsymbol{X}$

- $\boldsymbol{u}^T \boldsymbol{X} \boldsymbol{u} = \text{trace } \boldsymbol{u}^T \boldsymbol{X} \boldsymbol{u} = \text{trace } \boldsymbol{X} \boldsymbol{u} \boldsymbol{u}^T = \boldsymbol{X} \bullet \boldsymbol{u} \boldsymbol{u}^T$

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$ and the inner product $\boldsymbol{X} \bullet \boldsymbol{Y}$.

- $\mathbb{S}^n_+ \subseteq (\mathbb{S}^n_+)^* \equiv \{\boldsymbol{Y} \in \mathbb{S}^n : \boldsymbol{Y} \bullet \boldsymbol{X} \geq 0$ for $\forall \boldsymbol{X} \in \mathbb{S}^n_+\}$.

- $\mathbb{S}^n_+ \supseteq (\mathbb{S}^n_+)^*$. Hence $\mathbb{S}^n_+ = (\mathbb{S}^n_+)^*$ (self-dual).

.

Exercise 2. Prove $\mathbb{S}^n_+ = (\mathbb{S}^n_+)^*$.

(SDP)	minimize	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$
	subject to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \le p \le m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$.

$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}$ and the inner product $\boldsymbol{X} \bullet \boldsymbol{Y}$.

- $\boldsymbol{X}, \ \boldsymbol{Y} \succeq \boldsymbol{O}$ and $\boldsymbol{X} \bullet \boldsymbol{Y} = 0 \Rightarrow \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{O}$.

  (Proof) Apply the eigenvalue decomposition $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T$. Then

  $$0 = \boldsymbol{X} \bullet \boldsymbol{Y} = \text{trace}\ \boldsymbol{X}\boldsymbol{Y} = \text{trace}\ \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T\boldsymbol{Y}$$
  $$= \text{trace}\ \boldsymbol{D}\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P} = \sum_{i=1}^n D_{ii} \left(\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\right)_{ii}, \ D_{ii} \ge 0, \ \left(\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\right)_{ii} \ge 0$$
  $$\Rightarrow \forall i, \begin{cases} D_{ii} = 0 \text{ or} \\ \left(\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\right)_{ii} = 0; \text{the } i\text{th row of } \left(\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\right) = \mathbf{0}. \end{cases}$$

  Therefore $\boldsymbol{D}\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P} = \boldsymbol{O}$, which implies $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\boldsymbol{P}^T = \boldsymbol{O}$.

(SDP) minimize $\boldsymbol{A}_0 \bullet \boldsymbol{X}$
subject to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}.$

Common properties on

$$\mathbb{R}^n_+ \equiv \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \geq \mathbf{0}\}, \ \mathbb{S}^n_+ \equiv \{\boldsymbol{X} \in \mathbb{S}^n : \boldsymbol{X} \succeq \boldsymbol{O}\}.$$

- $\mathbb{R}^n_+$ is a cone; $\alpha\boldsymbol{X} \in \mathbb{R}^n_+$ if $\alpha \geq 0$, $\boldsymbol{x} \in \mathbb{R}^n_+$.
- $\mathbb{R}^n_+$ is convex;
  $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{x} \in \mathbb{R}^n_+$ if $0 \leq \lambda \leq 1$, $\boldsymbol{x}, \ \boldsymbol{y} \in \mathbb{R}^n_+$.
- self-dual;
  $(\mathbb{R}^n_+)^* \equiv \left\{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{y} \bullet \boldsymbol{x} \geq 0 \text{ for } \forall \boldsymbol{x} \in \mathbb{R}^n_+\right\} = \mathbb{R}^n_+.$
- $\boldsymbol{x}, \ \boldsymbol{y} \in \mathbb{R}^n_+$ and $\boldsymbol{x} \cdot \boldsymbol{y} = 0 \Longrightarrow x_i y_i = 0 \ (1 \leq i \leq n)$.

- $\mathbb{S}^n_+$ is a cone; $\alpha\boldsymbol{X} \in \mathbb{S}^n_+$ if $\alpha \geq 0$ and $\boldsymbol{X} \in \mathbb{S}^n_+$.
- $\mathbb{S}^n_+$ is convex;
  $\lambda\boldsymbol{X} + (1-\lambda)\boldsymbol{Y} \in \mathbb{S}^n_+$ if $0 \leq \lambda \leq 1$ and $\boldsymbol{X}, \ \boldsymbol{Y} \in \mathbb{S}^n_+$.
- self-dual;
  $(\mathbb{S}^n_+)^* \equiv \left\{\boldsymbol{Y} \in \mathbb{S}^n : \boldsymbol{Y} \bullet \boldsymbol{X} \geq 0 \text{ for } \forall \boldsymbol{X} \in \mathbb{S}^n_+\right\} = \mathbb{S}^n_+.$
- $\boldsymbol{X}, \ \boldsymbol{Y} \in \mathbb{S}^n_+$ and $\boldsymbol{X} \bullet \boldsymbol{Y} = 0 \Longrightarrow \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{O}$.

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Equality standard form (SDP):

min. $\boldsymbol{A}_0 \bullet \boldsymbol{X}$

sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}.$

$$\Uparrow \quad n = \sum_{q=1}^{t} n^q, \ \boldsymbol{A}_p \equiv \begin{pmatrix} \boldsymbol{A}_p^1 & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_p^2 & \cdots & \boldsymbol{O} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{A}_p^t \end{pmatrix}.$$

Equality standard form with multiple matrix variables (SDP)' :

min. $\sum_{q=1}^{t} \boldsymbol{A}_0^q \bullet \boldsymbol{X}^q$

sub.to $\sum_{q=1}^{t} \boldsymbol{A}_p^q \bullet \boldsymbol{X}^q = b_p \ (1 \leq p \leq m),$

$\mathbb{S}^{n^q} \ni \boldsymbol{X}^q \succeq \boldsymbol{O} \ (1 \leq q \leq t).$

- If $n^q = 1 \ (1 \leq q \leq t)$, (SDP)' is equivalent to the equality standard form of LP, where $\boldsymbol{A}_p^q \in \mathbb{R}$ and $X^q \in \mathbb{R}$.
- Can we transform the above (SDP)' (or the equality standard form of LP) into Equality standard form (SDP)?

Exercise 3. Prove (SDP)' is equivalent to (SDP). Hint: Construct an optimal solution of (SDP) from any optimal solution of (SDP)', and vice versa.

10/10/2008

Why do we need a standard from SDP?

(a) A unified SDP model for theory and method of SDPs.

(b) SDP software packages are available only for some standard forms.

Equality standard form (SDP):

$$\begin{array}{ll} \text{min.} & \boldsymbol{A}_0 \bullet \boldsymbol{X} \\ \text{sub.to} & \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}. \end{array}$$

$$\Uparrow \quad ?$$

An SDP from systems and control theory (SDP)':

$$\begin{array}{ll} \text{min} & \lambda \\ \text{sub. to} & \begin{pmatrix} \boldsymbol{X}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{X} + \boldsymbol{C}^T\boldsymbol{C} & \boldsymbol{X}\boldsymbol{B} + \boldsymbol{C}^T\boldsymbol{D} \\ \boldsymbol{B}^T\boldsymbol{X} + \boldsymbol{D}^T\boldsymbol{C} & \boldsymbol{D}^T\boldsymbol{D} - \boldsymbol{I} \end{pmatrix} \preceq \lambda \boldsymbol{I}, \\ & \boldsymbol{X} \succeq -\lambda \boldsymbol{I}. \end{array}$$

Here $\boldsymbol{X} \in \mathbb{S}^n$ and $\lambda \in \mathbb{R}$ are variables, and $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ are given data matrices.

- Can we transform the (SDP)' into Equality standard form (SDP)?
- "Yes" in theory, but not practical at all.
- Transform (SDP)' into an LMI standard form (with equality constraints), which corresponds to the dual of an equality standard form with free variables.

A general SDP:

min.	a linear function in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$ $(1 \le q \le t)$
sub.to	linear equalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear (matrix) inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	$x_1, \ldots, x_k \in \mathbb{R}$ (free real variables),
	$\boldsymbol{X}^q \succeq \boldsymbol{O}$ $(1 \le q \le t)$ (psd matrix variables).

Here a nonnegative x_i is regarded as a 1×1 psd matrix var., and a matrix variable $\boldsymbol{U} \in \mathbb{R}^{k \times m}$ a set of free variables U_{ij}s.

Any real-valued linear function in $\boldsymbol{X} \in \mathbb{S}^n$ can be written as $\boldsymbol{A} \bullet \boldsymbol{X} = \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{A}_{ij} \boldsymbol{X}_{ij}$ for $\exists \boldsymbol{A} \in \mathbb{S}^n$.

- We can transform ‘any SDP’ into Equality standard form. But such a transformation is neither trivial nor practical in many cases.
- It is easier to reduce an SDP to ‘an LMI standard form with equality constraints’ than to Equality standard form.

A general SDP:

min.	a linear function in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$ $(1 \leq q \leq t)$
sub.to	linear equalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear (matrix) inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	$x_1, \ldots, x_k \in \mathbb{R}$ (free real variables),
	$\boldsymbol{X}^q \succeq \boldsymbol{O}$ $(1 \leq q \leq t)$ (psd matrix variables).

Reduction to 'an LMI standard form with equality constraints'.

Represent each symmetric variable $\boldsymbol{X}^q \in \mathbb{S}^{n^q}$ as a linear combination of a basis $\boldsymbol{E}^q_{ij}$ $(1 \leq i \leq j \leq n^q)$ such that

$$\boldsymbol{X}^q = \sum_{1 \leq i \leq j \leq n^q} \boldsymbol{E}^q_{ij} y^q_{ij},$$

where y^q_{ij} denotes a free real variable and $\boldsymbol{E}^q_{ij}$ an $n^q \times n^q$ matrix with 1 at the (i,j)th and (j,i)th elements and 0 elsewhere. Then substitute it into the general SDP.

A general SDP:

min.	a linear function in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$ $(1 \leq q \leq t)$
sub.to	linear equalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	linear (matrix) inequalities in $x_1, \ldots, x_k$ and $\boldsymbol{X}^q$,
	$x_1, \ldots, x_k \in \mathbb{R}$ (free real variables),
	$\boldsymbol{X}^q \succeq \boldsymbol{O}$ $(1 \leq q \leq t)$ (psd matrix variables).

'An LMI standard form with equality constraints':

min	a linear function in $y_1, \ldots, y_\ell$
sub.to	linear equalities in $y_1, \ldots, y_\ell$,
	linear (matrix) inequalities in $y_1, \ldots, y_\ell$,
	$y_1, \ldots, y_\ell \in \mathbb{R}$ (free real variables).

- Take the dual $\Rightarrow$ an eq. standard form with free variables.
- We can apply existing software; CSDP, PENON, SDPA, SDPT3 and SeDuMi to this primal-dual pair.

Exercise 4. Transform the SDP

$$\min \quad w + \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \bullet \boldsymbol{X} \quad \text{sub.to} \quad \begin{pmatrix} \boldsymbol{X} & \begin{matrix} 2 \\ 1 \end{matrix} \\ \begin{matrix} 2 & 1 \end{matrix} & w \end{pmatrix} \succeq \boldsymbol{O}.$$

to an LMI standard form SDP

$$\begin{aligned}
\min \quad & w + 2y_1 + 2y_2 + 3y_3 \\
\text{sub.to} \quad & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} y_1 + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} y_2 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} y_3 \\
& + \begin{pmatrix} \boldsymbol{O} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 1 \end{pmatrix} w + \begin{pmatrix} \boldsymbol{O} & \begin{matrix} 2 \\ 1 \end{matrix} \\ \begin{matrix} 2 & 1 \end{matrix} & 0 \end{pmatrix} \succeq \boldsymbol{O}.
\end{aligned}$$

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Eigenvalues of a symmetric matrix $\boldsymbol{A}$

$$\begin{aligned} \text{the max. eigenvalue} &= \min\{\lambda : \lambda \boldsymbol{I} \succeq \boldsymbol{A}\} \\ &= \min\{\lambda : \lambda \boldsymbol{I} - \boldsymbol{A} \succeq \boldsymbol{O}\}. \\ \text{the min. eigenvalue} &= \max\{\lambda : \boldsymbol{A} - \lambda \boldsymbol{I} \succeq \boldsymbol{O}\}. \end{aligned}$$

- We can formulate many engineering problems involving eigenvalues of symmetric matrices via SDPs.
- A Linear Matrix inequality (LMI) $\boldsymbol{A}(\cdot) \succeq \boldsymbol{O}$, where $\boldsymbol{A}(\cdot)$ is a linear mapping in matrix and/or vector variables can be formulated in

  $$\text{maximize } \lambda \text{ subject to } \boldsymbol{A}(\cdot) - \lambda \boldsymbol{I} \succeq \boldsymbol{O}.$$

  For example,

  $$\boldsymbol{A}(\boldsymbol{X}) = \begin{pmatrix} \boldsymbol{X}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{X} + \boldsymbol{C}^T\boldsymbol{C} & \boldsymbol{X}\boldsymbol{B} + \boldsymbol{C}^T\boldsymbol{D} \\ \boldsymbol{B}^T\boldsymbol{X} + \boldsymbol{D}^T\boldsymbol{C} & \boldsymbol{D}^T\boldsymbol{D} - \boldsymbol{I} \end{pmatrix} \succeq \boldsymbol{O}.$$

For LMIs, see

[6] S. Boyd, L. El Ghaui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

The Schur complement. Let

$$\boldsymbol{A} \in \mathbb{S}^k, \text{ positive definite}, \quad \boldsymbol{X} \in \mathbb{R}^{k\times\ell}, \ \boldsymbol{Y} \in \mathbb{S}^\ell.$$

Then

$$\boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{X} \succeq \boldsymbol{O} \quad \Leftrightarrow \quad \begin{pmatrix} \boldsymbol{A} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{O}.$$

quadratic in $\boldsymbol{X}$ linear in $\boldsymbol{X}$

Proof:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{X} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{A}^{-1}\boldsymbol{X} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{A} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{A}^{-1}\boldsymbol{X} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}$$

Hence

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{O} \Leftrightarrow \begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}.$$

$\Updownarrow$ $\boldsymbol{A}$ is positive definite.

$$\boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{X} \succeq \boldsymbol{O}.$$

The Schur complement. Let

$$\boldsymbol{A} \in \mathbb{S}^k, \text{ positive definite}, \quad \boldsymbol{X} \in \mathbb{R}^{k\times\ell}, \ \boldsymbol{Y} \in \mathbb{S}^\ell.$$

Then

$$\boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{X} \succeq \boldsymbol{O} \iff \begin{pmatrix} \boldsymbol{A} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{O}.$$

quadratic in $\boldsymbol{X}$ linear in $\boldsymbol{X}$

- When $\boldsymbol{A} = \boldsymbol{I}$, $\boldsymbol{Y} - \boldsymbol{X}^T\boldsymbol{X} \succeq \boldsymbol{O} \iff \begin{pmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{O}.$
- When $\boldsymbol{A} = \boldsymbol{I}$, $\boldsymbol{X} = \boldsymbol{x} \in \mathbb{R}^k$ and $\boldsymbol{Y} = \boldsymbol{y} \in \mathbb{R}$,

$$\boldsymbol{y} - \boldsymbol{x}^T\boldsymbol{x} \geq 0 \iff \begin{pmatrix} \boldsymbol{I} & \boldsymbol{x} \\ \boldsymbol{x}^T & \boldsymbol{y} \end{pmatrix} \succeq \boldsymbol{O}.$$

- When $\boldsymbol{A} = \boldsymbol{I}\boldsymbol{y}$, $\boldsymbol{X} = \boldsymbol{x} \in \mathbb{R}^k$ and $\boldsymbol{Y} = \boldsymbol{y} \in \mathbb{R}$,

$$\boldsymbol{y} - \sqrt{\boldsymbol{x}^T\boldsymbol{x}} \geq 0 \iff \boldsymbol{y}^2 - \boldsymbol{x}^T\boldsymbol{x} \geq 0, \ \boldsymbol{y} \geq 0 \iff \begin{pmatrix} \boldsymbol{I}\boldsymbol{y} & \boldsymbol{x} \\ \boldsymbol{x}^T & \boldsymbol{y} \end{pmatrix} \succeq \boldsymbol{O}.$$

(SOCP constraint) $(y - \boldsymbol{x}^T\boldsymbol{x}/y \geq 0$ if $y > 0)$

A quasi-convex optimization problem

$$\min \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}} \text{ sub.to } \boldsymbol{Ax} \geq \boldsymbol{b}.$$

Here $\boldsymbol{L} \in \mathbb{R}^{k \times n}$, $\boldsymbol{c} \in \mathbb{R}^k$, $\boldsymbol{d} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{\ell \times n}$, $\boldsymbol{b} \in \mathbb{R}^\ell$, and $\boldsymbol{d}^T\boldsymbol{x} > 0$ for $\forall$ feasible $\boldsymbol{x} \in \mathbb{R}^n$.

$$\Updownarrow$$

$$\min \zeta \text{ sub.to } \zeta \geq \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}}, \ \boldsymbol{Ax} \geq \boldsymbol{b}.$$

$$\Updownarrow \ \zeta - \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}} \geq 0 \Leftrightarrow \begin{pmatrix} (\boldsymbol{d}^T\boldsymbol{x})\boldsymbol{I} & \boldsymbol{Lx}-\boldsymbol{c} \\ (\boldsymbol{Lx}-\boldsymbol{c})^T & \zeta \end{pmatrix} \succeq \boldsymbol{O}.$$

SDP: $\min \zeta$ sub.to $\begin{pmatrix} \boldsymbol{d}^T\boldsymbol{x}\boldsymbol{I} & \boldsymbol{Lx}-\boldsymbol{c} \\ (\boldsymbol{Lx}-\boldsymbol{c})^T & \zeta \end{pmatrix} \succeq \boldsymbol{O}, \ \boldsymbol{Ax} \geq \boldsymbol{b}.$

$$\Downarrow$$

SOCP

Chapter 1: Basic theory

1. LP versus SDP
2. Why is SDP interesting and important?
3. The equality standard form
4. Some basic properties on positive semidefinite matrices and their inner product
5. General SDPs
6. Some examples
7. Duality

Lagrangian function, Lagrangian Dual, 10/10/2008

A general nonlinear program (P):

min	$f(\boldsymbol{x})$
sub.to	$g_j(\boldsymbol{x}) = 0 \ (j = 1, \ldots, \ell), \ h_k(\boldsymbol{x}) \geq 0 \ (k = 1, \ldots, m).$

Here $f, \ g_j, \ h_k : \mathbb{R}^n \rightarrow \mathbb{R}$. Let ζ^* be the optimal value of (P).

Lagrangian function $L : \mathbb{R}^n \times \mathbb{R}^\ell \times \mathbb{R}^m_+ \rightarrow \mathbb{R}$

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \ = \ f(\boldsymbol{x}) - \sum_{j=1}^{\ell} y_j g_j(\boldsymbol{x}) - \sum_{k=1}^{m} z_k h_k(\boldsymbol{x}).$$

Let $(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^\ell \times \mathbb{R}^m_+$ be fixed. Then, for $\forall$ feas. sol. $\bar{\boldsymbol{x}}$ of (P)

$$\text{min}_{\boldsymbol{x} \in \mathbb{R}^n} L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \ \leq \ L(\bar{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z}) \leq f(\bar{\boldsymbol{x}});$$

hence $\text{min}_{\boldsymbol{x} \in \mathbb{R}^n} L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \leq \zeta^* \Rightarrow$ Lagrangian relaxation.

Lagrangian dual

$$\hat{\zeta} \ = \ \max_{(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^\ell \times \mathbb{R}^m} \text{min}_{\boldsymbol{x} \in \mathbb{R}^n} L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \leq L(\bar{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z}) \leq \zeta^*$$

$\hat{\zeta} < \zeta^*$ can occur in general. Convexity+Assumption $\Rightarrow \hat{\zeta} = \zeta^*$.

Equality standard form SDP:

$$\min \quad \boldsymbol{A}_0 \bullet \boldsymbol{X} \quad \text{sub.to} \quad \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \boldsymbol{X} \succeq \boldsymbol{O}.$$

Lagrangian function, Lagrangian dual

$$L(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) = \boldsymbol{A}_0 \bullet \boldsymbol{X} - \sum_{p=1}^{m} (\boldsymbol{A}_p \bullet \boldsymbol{X} - b_p)\, y_p - \boldsymbol{S} \bullet \boldsymbol{X}$$

$$\text{for } \forall \boldsymbol{X} \in \boldsymbol{S}^n, \ \forall \boldsymbol{y} = (y_1, \ldots, y_m)^T \in \mathbb{R}^m, \ \forall \boldsymbol{S} \succeq \boldsymbol{O}.$$

$$\max_{\boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O}} \min_{\boldsymbol{X} \in \mathbb{S}^n} L(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S})$$

$$= \max_{\boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O}} \{ L(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) : \nabla_X L(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) = \boldsymbol{O} \}$$

$$= \max_{\boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O}} \left\{ L(\boldsymbol{X}, y, \boldsymbol{S}) : \boldsymbol{A}_0 - \sum_{p=1}^{m} \boldsymbol{A}_p y_p - \boldsymbol{S} = \boldsymbol{O} \right\}$$

$$= \max_{\boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O}} \left\{ L(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) : \boldsymbol{S} = \boldsymbol{A}_0 - \sum_{p=1}^{m} \boldsymbol{A}_p y_p \right\}$$

$$= \max_{\boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O}} \left\{ \sum_{p=1}^{m} b_p y_p : \boldsymbol{S} = \boldsymbol{A}_0 - \sum_{p=1}^{m} \boldsymbol{A}_p y_p \right\}$$

(since $L(\boldsymbol{X}, y, \boldsymbol{A}_0 - \sum_{p=1}^{m} \boldsymbol{A}_p y_p) = \sum_{p=1}^{m} b_p y_p$)

$$= \max \left\{ \sum_{p=1}^{m} b_p y_p : \boldsymbol{S} = \boldsymbol{A}_0 - \sum_{p=1}^{m} \boldsymbol{A}_p y_p, \ \boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{S} \succeq \boldsymbol{O} \right\}$$

(Dual of Equality standard form).

A primal-dual pair of LPs

(P)	min	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$	s.t.	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p), \ \boldsymbol{x} \geq \mathbf{0}$.
(D)	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{a}_p y_p + s = \boldsymbol{a}_0, \ \mathbb{R}^n \ni s \geq \mathbf{0}$.

Weak duality

LP : $\boldsymbol{x} \cdot s = \boldsymbol{a}_0 \cdot \boldsymbol{x} - \sum_{j=1}^m b_p y_p \geq 0$ for $\forall$ feasible $\boldsymbol{x}, \ \boldsymbol{y}, \ s$.

SDP : $\boldsymbol{X} \bullet S = \boldsymbol{A}_0 \bullet \boldsymbol{X} - \sum_{j=1}^m b_p y_p \geq 0$ for $\forall$ feasible $\boldsymbol{X}, \ \boldsymbol{y}, \ \boldsymbol{S}$.

Exercise 5: Prove the weak duality

$$\boldsymbol{X} \bullet S = \boldsymbol{A}_0 \bullet \boldsymbol{X} - \sum_{j=1}^m b_p y_p \geq 0 \text{ for } \forall \text{ feasible } \boldsymbol{X}, \ \boldsymbol{y}, \ \boldsymbol{S}.$$

A primal-dual pair of SDPs

(P)	min.	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	sub.to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succeq \boldsymbol{O}$.
(D)	max.	$\sum_{p=1}^m b_p y_p$	sub.to	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succeq \boldsymbol{O}$.

A primal-dual pair of LPs

(P)	min	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$	s.t.	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p), \ \boldsymbol{x} \geq \mathbf{0}$.
(D)	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{a}_p y_p + s = \boldsymbol{a}_0, \ \mathbb{R}^n \ni s \geq \mathbf{0}$.

Strong duality: If $\exists$ feasible $(\boldsymbol{x}, \ \boldsymbol{y}, \ s)$ $(\boldsymbol{x} \geq 0, \ \boldsymbol{y} \geq \mathbf{0})$ then

$$\text{LP} \ : \ \bar{\boldsymbol{x}} \cdot \bar{s} = \boldsymbol{a}_0 \cdot \boldsymbol{x} - \sum_{j=1}^m b_p \bar{y}_p = 0 \text{ at } \forall \text{ optimal } (\bar{\boldsymbol{x}}, \ \bar{\boldsymbol{y}}, \ \bar{s}).$$

If $\exists$ interior feasible $(\boldsymbol{X}, \ \boldsymbol{y}, \ \boldsymbol{S})$ $(\boldsymbol{X} \succ \boldsymbol{O}, \ \boldsymbol{S} \succ \boldsymbol{O})$ then

$$\text{SDP} \ : \ \bar{\boldsymbol{X}} \bullet \bar{\boldsymbol{S}} = \boldsymbol{A}_0 \bullet \bar{\boldsymbol{X}} - \sum_{j=1}^m b_p \bar{y}_p = 0 \text{ at } \forall \text{ optimal } (\bar{\boldsymbol{X}}, \ \bar{\boldsymbol{y}}, \ \bar{\boldsymbol{S}}).$$

- For the strong duality, "$\exists$ int. feasible $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S})$ $(\boldsymbol{X} \succ \boldsymbol{O}, \boldsymbol{S} \succ \boldsymbol{O})$" is necessary! $\Rightarrow$ an example, next

A primal-dual pair of SDPs

(P)	min.	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	sub.to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succeq \boldsymbol{O}$.
(D)	max.	$\sum_{p=1}^m b_p y_p$	sub.to	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succeq \boldsymbol{O}$.

Example [45]: "$\exists$ interior feasible $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S})$ ($\boldsymbol{X} \succ \boldsymbol{O}, \boldsymbol{S} \succ \boldsymbol{O}$)" is necessary!

$$
\begin{array}{ll}
\text{(P) min} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet \boldsymbol{X} \\
\text{sub.to} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bullet \boldsymbol{X} = 0, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \bullet \boldsymbol{X} = 2, \boldsymbol{X} \succeq \boldsymbol{O}.
\end{array}
$$

or

$$
\text{(P) min} \quad X_{33} \quad \text{sub.to} \quad X_{11} = 0, X_{12} + X_{21} + 2X_{33} = 2, \boldsymbol{X} \succeq \boldsymbol{O}.
$$

Exercise 6. Show that the objective value $X_{33} = 1$ if $\boldsymbol{X}$ is feasible.

$$\text{(D) max} \quad 2y_2$$
$$\text{sub.to} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} y_1 + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} y_2 \preceq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

or

$$\text{(D) min} \quad 2y_2 \quad \text{sub.to} \quad \begin{pmatrix} -y_1 & -y_2 & 0 \\ -y_2 & 0 & 0 \\ 0 & 0 & 1-2y_2 \end{pmatrix} \succeq \boldsymbol{O}.$$

Exercise 7. Show that the objective value $2y_2 = 0$ if (y_1, y_2) is feasible.

A primal-dual pair of SDPs

$$\text{(P)} \quad \text{min.} \quad \boldsymbol{A}_0 \bullet \boldsymbol{X} \quad \text{sub.to} \quad \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succeq \boldsymbol{O}.$$

$$\text{(D)} \quad \text{max.} \quad \sum_{p=1}^m b_p y_p \quad \text{sub.to} \quad \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succeq \boldsymbol{O}.$$

The KKT optimality condition

$$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \le p \le m), \ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0,$$

$$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}, \ \mathbb{S}^n \ni \boldsymbol{S} \succeq \boldsymbol{O}, \ \boldsymbol{X}\boldsymbol{S} = \boldsymbol{O} \text{ (complementarity)}.$$

$\boldsymbol{O} = \boldsymbol{X}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{X} \Rightarrow \boldsymbol{X}$ and $\boldsymbol{S}$ are commutative; hence

$\Downarrow$ $\exists$ orthogonal $\boldsymbol{P} \in \mathbb{R}^{n \times n}$; $\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P} = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$,

$\boldsymbol{P}^T\boldsymbol{S}\boldsymbol{P} = \text{diag}\,(\nu_1, \ldots, \nu_n)$

$$\boldsymbol{O} = \boldsymbol{X}\boldsymbol{S} = P^T\boldsymbol{X}PP^T\boldsymbol{S}P = \text{diag}\,(\lambda_1, \ldots, \lambda_n)\text{diag}\,(\nu_1, \ldots, \nu_n),$$

$$P^T(\boldsymbol{X} + \boldsymbol{S})P = \text{diag}\,(\lambda_1, \ldots, \lambda_n) + \text{diag}\,(\nu_1, \ldots, \nu_n).$$

$$\Downarrow$$

$$\lambda_i \ge 0, \ \nu_i \ge 0, \ \lambda_i\nu_i = 0 \ (1 \le i \le n) \text{ (complementarity)},$$

$$X + S \succ O \Leftrightarrow \lambda_i + \nu_i > 0 \ (1 \le i \le n) \text{ (strict comp.)}.$$

LP: $x_i \ge 0, s_i \ge 0, x_i s_i = 0 \ (\forall i)$ (comp.), $x_i + s_i > 0 \ (\forall i)$

An equality standard form

(P) min. $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \boldsymbol{X} \succeq \boldsymbol{O}$.

An equality standard form with free variables

(P) min. $\boldsymbol{A}_0 \bullet \boldsymbol{X} + \boldsymbol{d}_0^T \boldsymbol{z}$

sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} + \boldsymbol{d}_p^T \boldsymbol{z} = b_p \ (1 \leq p \leq m),$

$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}, \ \boldsymbol{z} \in \mathbb{R}^\ell$ (a free vector variable).

Here $\boldsymbol{d}_p \in \mathbb{R}^\ell \ (0 \leq p \leq m)$.

$\Updownarrow$ dual

An LMI standard form with equality constraints

(D) max. $\sum_{p=1}^{m} b_p y_p$

sub.to $\sum_{p=1}^{m} \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \mathbb{S}^n \ni \boldsymbol{S} \succeq \boldsymbol{O}, \ \sum_{p=1}^{m} \boldsymbol{d}_p y_p = \boldsymbol{d}_0.$

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

Some existing numerical methods for SDPs

- IPMs (Interior-point methods)
  - Primal-dual scaling, CSDP(Borchers[7]), SDPA(Fujisawa-K-Nakata-Yamashita[49]), SDPT3(Toh-Todd-Tutuncu[42]), SeDuMi(Sturm[37])
  - Dual scaling, DSDP(Benson-Ye-Zhang[3])
- Nonlinear programming approaches
  - Spectral bundle method(Helmberg-Rendl[17])
  - Gradient-based log-barrier method(Burer-Monteiro[9])
  - PENON(Kocvara [19]) — Augmented Lagrangian
  - Saddle point mirror-prox algorithm (Lu-Nemirovski-Monteiro[26])

- Medium scale (e.g. $n,\ m \leq 5000$) and high accuracy.
- Large scale (e.g., $n, m \geq 10,000$) and low accuracy.

• Parallel implementation:

SDPA $\Rightarrow$ SDPARA(Y-F-K[49]), SDPARA-C(N-Y-F-K[31])

DSDP $\Rightarrow$ PDSDP(Benson[2]), CSDP $\Rightarrow$ Borchers-Young[8]

Spectral bundle method $\Rightarrow$ Nayakkankuppam[32]

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

3 approaches to primal-dual interior-point methods for SDPs

(1) Self-concordance.
Yu. E. Nesterov and A. Nemirovski ’94 [33]
Yu. E. Nesterov and M. J. Todd ’98 [34]

(2) Linear optimization problems over symmetric cones (Jordan algebra)
L. Faybusovich ’97 [10]
S. Schmieta and F. Alizadeh ’01 [36]
$\Longrightarrow$ Appendix.

(3) Extensions of primal-dual interior-point algorithms for LPs to SDPs [1, 17, 20, 27]
$\Longrightarrow$ Our approach in this lecture.

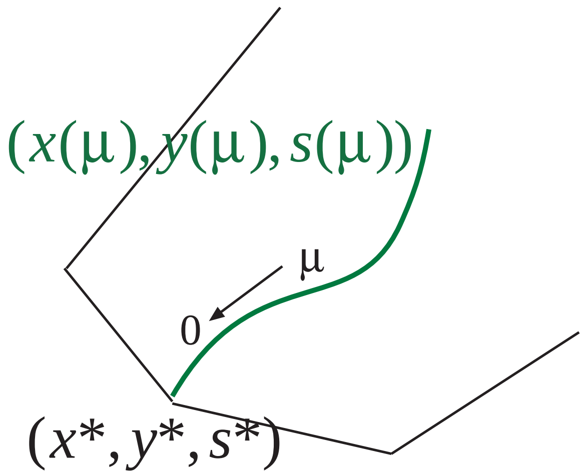
Chapter 2: Primal-dual interior-point methods

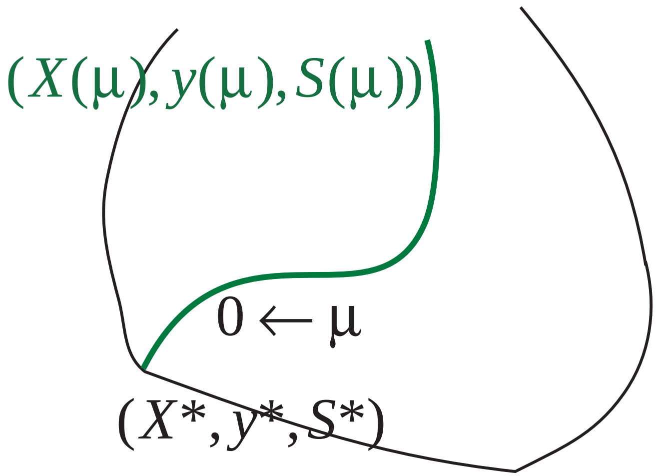
1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

LP:	P	min	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$	s.t.	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p = 1), \ \boldsymbol{x} \in \mathbb{R}^n_+$
	D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0, \ s \in \mathbb{R}^n_+$

SDP:	P	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$
	D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$

- Basic idea of the primal-dual interior-point method:
  Trace the central trajectory $\longrightarrow$ an opt. sol. in the p-d space.





- How do we define the central trajectory?
- How do we numerically trace the central trajectory?

LP:	P	min	$\boldsymbol{a}_0 \cdot \boldsymbol{x}$	s.t.	$\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p = 1), \ \boldsymbol{x} \in \mathbb{R}^n_+$
	D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0, \ s \in \mathbb{R}^n_+$

SDP:	P	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$
	D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$

- A log barrier to be away from the boundary $-\sum_{i=1}^m \log x_i$.

$\boldsymbol{x} \in$ the boundary of $\mathbb{R}^n_+$ $\Leftrightarrow x_i = 0 \ (i = 1, \ldots, n)$.
$\boldsymbol{x} \in$ the interior of $\mathbb{R}^n_+ \equiv \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \geq \mathbf{0}\} \Leftrightarrow x_i > 0 \ (i = 1, \ldots, n)$.

- A log barrier to be away from the boundary $-\log\det \boldsymbol{X}$.

$\boldsymbol{X} \in$ the interior of $\mathcal{S}^n_+ \equiv \{\boldsymbol{X} \in \mathbb{S}^n : \boldsymbol{X} \succeq \boldsymbol{O}\} \Leftrightarrow$ det $\boldsymbol{X} > 0$.
$\boldsymbol{X} \in$ the boundary of $\mathcal{S}^n_+$ $\Leftrightarrow$ det $\boldsymbol{X}$ = 0.

LP: P min $\boldsymbol{a}_0 \cdot \boldsymbol{x}$ s.t. $\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p = 1),\ \boldsymbol{x} \in \mathbb{R}^n_+$

D max $\sum_{p=1}^m b_p y_p$ s.t. $\sum_{p=1}^m \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0,\ s \in \mathbb{R}^n_+$

SDP: P min $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ s.t. $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p),\ \boldsymbol{X} \in \mathcal{S}^n_+$

D max $\sum_{p=1}^m b_p y_p$ s.t. $\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0,\ \boldsymbol{S} \in \mathcal{S}^n_+$

A primal-dual pair of LPs with logarithmic barrier terms, $\mu > 0$

P(μ) min $\boldsymbol{a}_0 \cdot \boldsymbol{x} - \mu \sum_{i=1}^m \log x_i$ s.t. $\boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p),\ \boldsymbol{x} > \mathbf{0}$

D(μ) max $\sum_{p=1}^m b_p y_p + \mu \sum_{i=1}^m \log s_i$ s.t. $\sum_{p=1}^m \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0,\ \boldsymbol{s} > \mathbf{0}$

A primal-dual pair of SDPs with logarithmic barrier terms, $\mu > 0$

P(μ) min $\boldsymbol{A}_0 \bullet \boldsymbol{X} - \mu \log \det \boldsymbol{X}$ s.t. $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p),\ \boldsymbol{X} \succ \boldsymbol{O}$

D(μ) max $\sum_{p=1}^m b_p y_p + \mu \log \det \boldsymbol{S}$

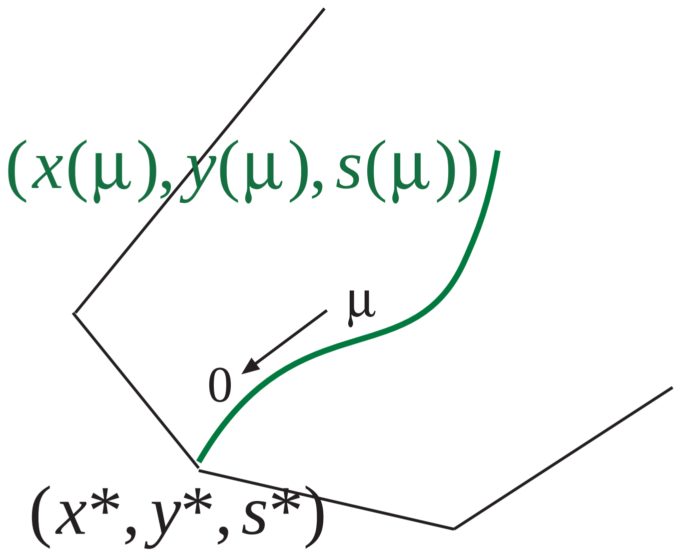
s.t. $\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0,\ \boldsymbol{S} \succ \boldsymbol{O}$

A primal-dual pair of LPs with logarithmic barrier terms, $\mu > 0$

$$\text{P}(\mu) \quad \min \quad \boldsymbol{a}_0 \cdot \boldsymbol{x} - \mu \sum_{i=1}^{m} \log x_i \ \text{s.t.}\ \boldsymbol{a}_p \cdot \boldsymbol{x} = b_p \ (\forall p),\ \boldsymbol{x} > \mathbf{0}$$

$$\text{D}(\mu) \quad \max \quad \sum_{p=1}^{m} b_p y_p + \mu \sum_{i=1}^{m} \log s_i \ \text{s.t.}\ \sum_{p=1}^{m} \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0,\ \boldsymbol{s} > \mathbf{0}$$

- For every $\mu > 0$, (P(μ),D(μ)) has a unique opt.sol. $(\boldsymbol{x}(\mu), \boldsymbol{y}(\mu), \boldsymbol{s}(\mu))$, which converges an opt. sol. of (P,D).



- $C = \{(\boldsymbol{x}(\mu), \boldsymbol{y}(\mu), \boldsymbol{s}(\mu)) : \mu > 0\}$: the central trajectory.

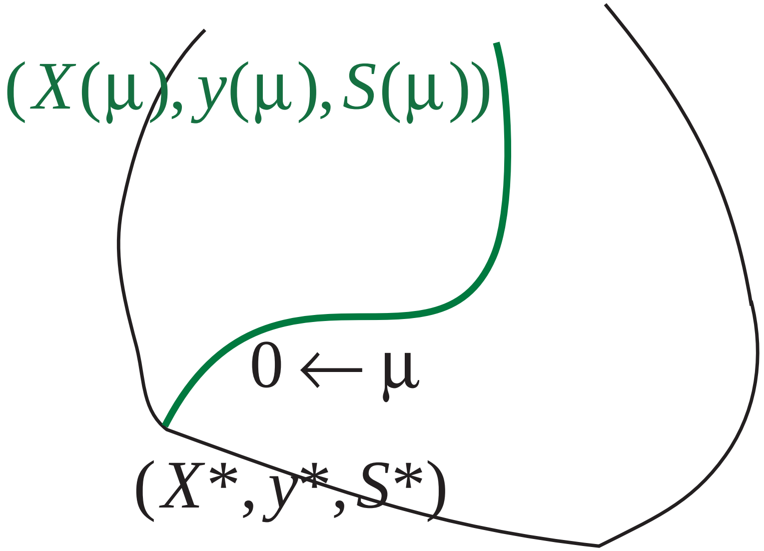
A primal-dual pair of SDPs with logarithmic barrier terms, $\mu > 0$

P(μ) min $\boldsymbol{A}_0 \bullet \boldsymbol{X} - \mu \log\det \boldsymbol{X}$ s.t. $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succ \boldsymbol{O}$

D(μ) max $\sum_{p=1}^m b_p y_p + \mu \log\det \boldsymbol{S}$

s.t. $\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succ \boldsymbol{O}$

- For every $\mu > 0$, (P(μ),D(μ)) has a unique opt.sol. $(\boldsymbol{X}(\mu), \boldsymbol{y}(\mu), \boldsymbol{S}(\mu))$, which converges an opt. sol. of (P,D).



- $C = \{(\boldsymbol{X}(\mu), \boldsymbol{y}(\mu), \boldsymbol{S}(\mu)) : \mu > 0\}$: the central trajectory.

A primal-dual pair of SDPs

(P)	min.	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	sub.to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succeq \boldsymbol{O}.$
(D)	max.	$\sum_{p=1}^m b_p y_p$	sub.to	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succeq \boldsymbol{O}.$

The KKT optimality condition

$$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (1 \leq p \leq m), \ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0,$$

$$\mathbb{S}^n \ni \boldsymbol{X} \succeq \boldsymbol{O}, \ \mathbb{S}^n \ni \boldsymbol{S} \succeq \boldsymbol{O}, \ \boldsymbol{X}\boldsymbol{S} = \boldsymbol{O}$$ (complementarity).

A primal-dual pair of SDPs with logarithmic barrier terms, $\mu > 0$

P(μ) min $\boldsymbol{A}_0 \bullet \boldsymbol{X} - \mu \log\det \boldsymbol{X}$ s.t. $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succ \boldsymbol{O}$

D(μ) max $\sum_{p=1}^{m} b_p y_p + \mu \log\det \boldsymbol{S}$

s.t. $\sum_{p=1}^{m} \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succ \boldsymbol{O}$

- For every $\mu > 0$, (P(μ),D(μ)) has a unique opt.sol. $(\boldsymbol{X}(\mu), \boldsymbol{y}(\mu), \boldsymbol{S}(\mu))$, which converges an opt. sol. of (P,D).
- For every $\mu > 0$, the objective function of P(μ) is convex in $\boldsymbol{X}$.
- For every $\mu > 0$, the objective function of D(μ) is concave in $(\boldsymbol{y}, \boldsymbol{S})$.
- For every $\mu > 0$, $(\boldsymbol{X}(\mu), \boldsymbol{y}(\mu), \boldsymbol{S}(\mu))$ is characterized as the Karush-Kuhn-Tucker optimality condition

$$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \sum_{p=1}^{m} \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0,$$
$$\boldsymbol{X} \succ 0, \ \boldsymbol{S} \succ 0, \ X\boldsymbol{S} = \mu \boldsymbol{I}.$$

Some properties of the central trajectory C

Suppose that

$(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) \in C$ or $\left\{ \begin{array}{l} \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \\ \boldsymbol{X} \succ 0, \ \boldsymbol{S} \succ 0, \ X\boldsymbol{S} = \mu \boldsymbol{I} \text{ for } \exists \mu > 0. \end{array} \right.$

- $\boldsymbol{X}$ and $\boldsymbol{S}$ are commutative; hence
  $\exists$ orthogonal $\boldsymbol{P} \in \mathbb{R}^{n \times n}$; $\boldsymbol{P}^T \boldsymbol{X} \boldsymbol{P} = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$,
  $\boldsymbol{P}^T \boldsymbol{S} \boldsymbol{P} = \text{diag}\,(\nu_1, \ldots, \nu_n)$.
- $\mu \boldsymbol{I} = P^T \mu \boldsymbol{I} \boldsymbol{P} = \boldsymbol{P}^T \boldsymbol{X} \boldsymbol{P} \boldsymbol{P}^T \boldsymbol{S} \boldsymbol{P} =$
  diag $(\lambda_1, \ldots, \lambda_n)$diag $(\nu_1, \ldots, \nu_n)$;
  $\lambda_i > 0, \ \nu_i > 0, \ \lambda_i \nu_i = \mu \ (\forall i)$
- As $\mu \to 0$, $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) \to$ an opt. sol. and $\lambda_i \nu_i \to 0 \ (\forall i)$.

In LP case: If $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ is on the central trajectory then

- $\boldsymbol{x}_i > 0, \ s_i > 0, \ x_i s_i = \mu \ (\forall i)$.
- As $\mu \to 0$, $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \to$ an opt. sol. and $x_i s_i \to 0 \ (\forall i)$.

Hence $\lambda_i \leftrightarrow x_i$ and $\nu_i \leftrightarrow s_i$.

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

$$(\boldsymbol{X},\boldsymbol{y},\boldsymbol{S}) = (\boldsymbol{X}(\mu),\boldsymbol{y}(\mu),\boldsymbol{S}(\mu)) \in C \ \Leftrightarrow$$

$$\begin{cases} \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succ 0, \\ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succ 0, \ \boldsymbol{X}\boldsymbol{S} = \mu \boldsymbol{I}. \end{cases}$$

(a continuation toward an opt.sol. as $\mu \to 0$)

One iteration: Suppose the current iterate $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)$ satisfies

$$\left(\boldsymbol{A}_p \bullet \boldsymbol{X}^k = b_p \ (\forall p), \ \sum_{p=1}^m \boldsymbol{A}_p y_p^k + \boldsymbol{S}^k = \boldsymbol{A}_0, \right) \ \boldsymbol{X}^k \succ 0, \ \boldsymbol{S}^k \succ 0.$$

(in theory but not in practice)

Choose $\beta \in [0,1]$ and let $\hat{\mu} = \beta \boldsymbol{X}^k \bullet \boldsymbol{S}^k / n$, which determines the target point $(\boldsymbol{X}(\hat{\mu}), \boldsymbol{y}(\hat{\mu}), \boldsymbol{S}(\hat{\mu}))$ on C satisfying

$$(\sharp) \ \ \boldsymbol{X}\boldsymbol{S} = \hat{\mu}\boldsymbol{I}, \ \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0.$$

Compute a search ("Newton") direction $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$.
Choose step lengths $\alpha_p > 0$ and $\alpha_d > 0$ ($\alpha_p = \alpha_d$ in theory);

$$\boldsymbol{X}^{k+1} = \boldsymbol{X}^k + \alpha_p d\boldsymbol{X} \succ 0, \boldsymbol{y}^{k+1} = \boldsymbol{y}^k + \alpha_d d\boldsymbol{y},$$

$$\boldsymbol{S}^{k+1} = \boldsymbol{S}^k + \alpha_d d\boldsymbol{S} \succ 0.$$

$$(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) = (\boldsymbol{X}(\mu), \boldsymbol{y}(\mu), \boldsymbol{S}(\mu)) \in C \iff$$

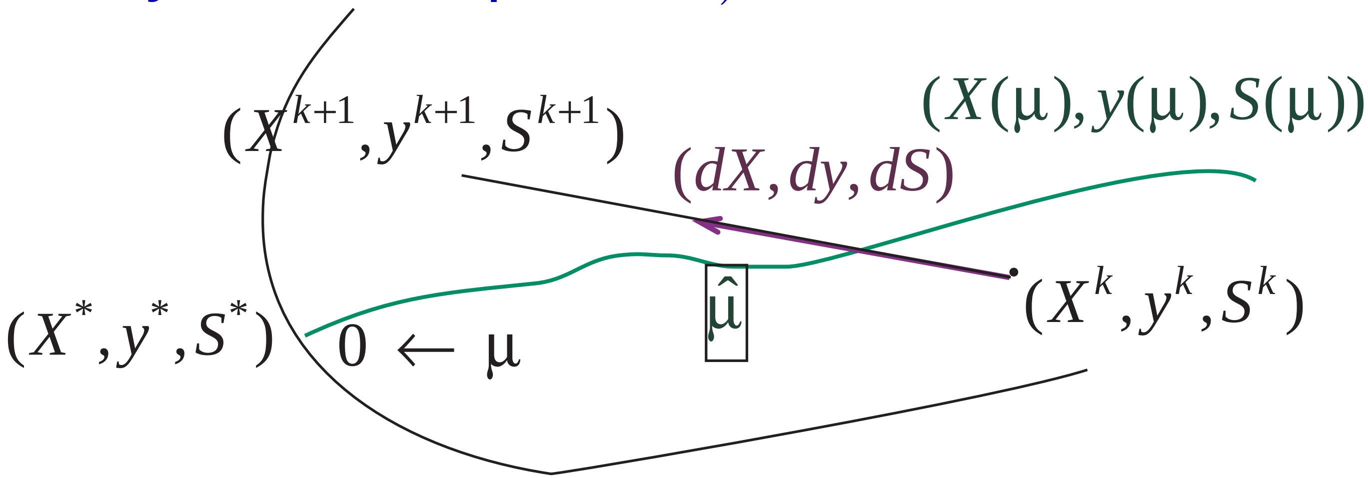
$$\begin{cases} \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succ 0, \\ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succ 0, \ \boldsymbol{X}\boldsymbol{S} = \mu \boldsymbol{I}. \end{cases}$$

(a continuation toward an opt.sol. as $\mu \to 0$)

One iteration: Suppose the current iterate $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)$ satisfies

$$(\boldsymbol{A}_p \bullet \boldsymbol{X}^k = b_p \ (\forall p), \ \sum_{p=1}^m \boldsymbol{A}_p y_p^k + \boldsymbol{S}^k = \boldsymbol{A}_0,) \ \boldsymbol{X}^k \succ 0, \ \boldsymbol{S}^k \succ 0.$$

(in theory but not in practice)



$$(\boldsymbol{X},\boldsymbol{y},\boldsymbol{S}) = (\boldsymbol{X}(\mu),\boldsymbol{y}(\mu),\boldsymbol{S}(\mu)) \in C \iff$$

$$\begin{cases} \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \succ 0, \\ \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \succ 0, \ \boldsymbol{X}\boldsymbol{S} = \mu \boldsymbol{I}. \end{cases}$$

(a continuation toward an opt.sol. as $\mu \to 0$)

- Computation of a search ("Newton") direction $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$

Choose $\beta \in [0,1]$ and let $\hat{\mu} = \beta \boldsymbol{X}^k \bullet S^k / n$, which determines the target point $(\boldsymbol{X}(\hat{\mu}), y(\hat{\mu}), \boldsymbol{S}(\hat{\mu}))$ on C satisfying

$$(\sharp) \ XS = \hat{\mu}\boldsymbol{I}, \ \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \sum_{p=1}^m \boldsymbol{A}_p y_p + S = \boldsymbol{A}_0.$$

Substitue $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k) + (d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$ into ($\sharp$). Solve

$$(\boldsymbol{X}^k + d\boldsymbol{X})(\boldsymbol{S}^k + d\boldsymbol{S}) = \hat{\mu}\boldsymbol{I} \Rightarrow \boldsymbol{X}^k \boldsymbol{S}^k + d\boldsymbol{X}\boldsymbol{S}^k + X^k d\boldsymbol{S} = \hat{\mu}\boldsymbol{I},$$

linearize

$$\boldsymbol{A}_p \bullet (\boldsymbol{X}^k + d\boldsymbol{X}) = b_p \ (\forall p),$$

$$\sum_{p=1}^m \boldsymbol{A}_p(\boldsymbol{y}_p^k + d\boldsymbol{y}_p) + (\boldsymbol{S}^k + d\boldsymbol{S}) = \boldsymbol{A}_0$$

in $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$.

- No sol. $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$; $d\boldsymbol{X} \in \mathbb{S}^n$, $d\boldsymbol{S} \in \mathbb{S}^n$

- No solution $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$; $d\boldsymbol{X} \in \mathbb{S}^n$, $d\boldsymbol{S} \in \mathbb{S}^n$.
- the total $\#$ of equations $>$ the total $\#$ of variables. In fact:

	# of equations	# of variables
$\boldsymbol{X}^k\boldsymbol{S}^k + d\boldsymbol{X}\boldsymbol{S}^k + \boldsymbol{X}^k d\boldsymbol{S} = \hat{\mu}\boldsymbol{I}$	$\rightarrow n^2$	$d\boldsymbol{X} \in \mathbb{S}^n \rightarrow \frac{n(n+1)}{2}$
$\boldsymbol{A}_p \bullet (\boldsymbol{X}^k + d\boldsymbol{X}) = b_p \; (\forall p)$	$\rightarrow m$	$d\boldsymbol{y} \rightarrow m$
$\sum_{p=1}^{m} \boldsymbol{A}_p(y_p^k + dy_p) + (\boldsymbol{S}^k + d\boldsymbol{S}) = \boldsymbol{A}_0$	$\rightarrow \frac{n(n+1)}{2}$	$d\boldsymbol{S} \rightarrow \frac{n(n+1)}{2}$
Total:	$3n^2/2 + n/2 + m$	$n^2 + n + m$

$\Rightarrow$ We need modification.

- More than 20 search directions and several family of search directions (Todd [38])
- Most popular ones are:
  - HKM direction [17, 20, 27]
  - NT direction [34]

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

Typical primal-dual interior-point methods (in theory)

(a) Path-following
(b) Mizuno-Todd-Ye Predictor-corrector
(c) Potential reduction — not stated, see [20, 43]
(d) Homogeneous self-dual embedding — not stated, see [26, 18]

Main issues studied:

- Polynomial-time convergence for (a), (b), (c) and (d)
- Local convergence for (b)

Some additional issues to be taken account:

(e) Search directions — NT [34], HKM [17, 20, 27]
(f) Feasible starting points or infeasible starting points

Primal-dual interior-point methods in practice — later

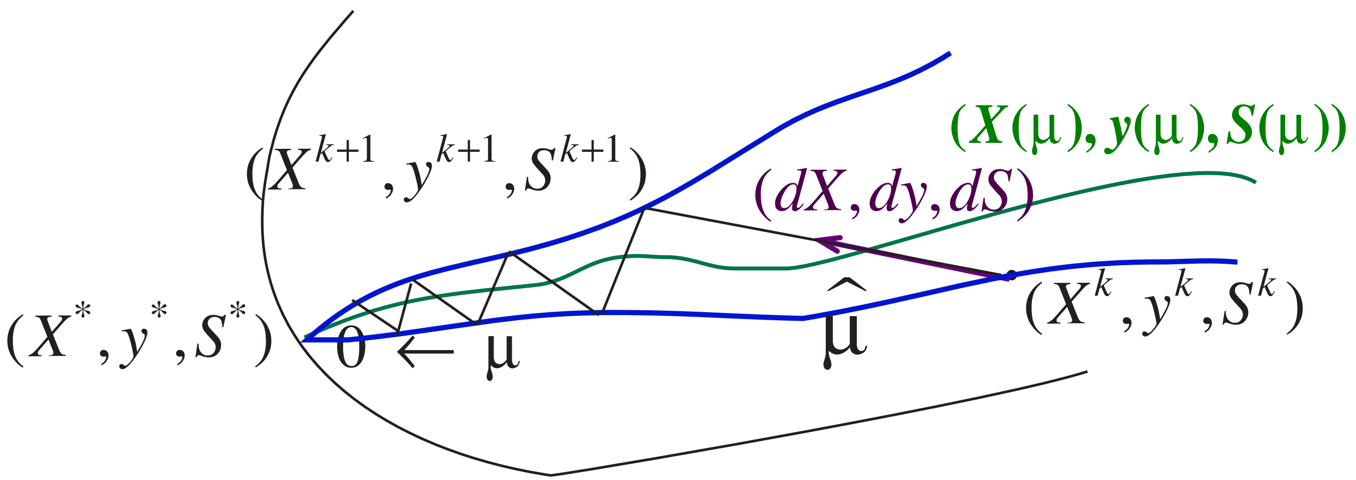
(a) Path-following primal-dual interior-point methods

- $\{(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)\}$ in a neighborhood N of the cent. trajectory.
- A step length

$$\bar{\alpha} = \max\{\alpha : (\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k) + \alpha(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S}) \in N\};$$

$$(\boldsymbol{X}^{k+1}, \boldsymbol{y}^{k+1}, \boldsymbol{S}^{k+1}) = (\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k) + \bar{\alpha}(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S}).$$

- Neighborhood?
- $\hat{\mu}$ to choose $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$ towards $(\boldsymbol{X}(\hat{\mu}), \boldsymbol{y}(\hat{\mu}), \boldsymbol{S}(\hat{\mu}))$?



(a) Path-following primal-dual interior-point methods

- $\{(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)\}$ in a neighborhood N of the cent. trajectory.
- A step length

  $\bar{\alpha} = \max\{\alpha : (\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k) + \alpha(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S}) \in N\};$

  $$(\boldsymbol{X}^{k+1}, \boldsymbol{y}^{k+1}, \boldsymbol{S}^{k+1}) = (\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k) + \bar{\alpha}(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S}).$$

- Neighborhood?
- $\hat{\mu}$ to choose $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$ towards $(\boldsymbol{X}(\hat{\mu}), \boldsymbol{y}(\hat{\mu}), \boldsymbol{S}(\hat{\mu}))$?

Best complexity: Short-step algorithms [20, 27, 34]

$$N_2(\tau) = \left\{ (\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) : \begin{array}{l} \text{feasible, } \left\| \sqrt{\boldsymbol{X}}\boldsymbol{S}\sqrt{\boldsymbol{X}} - \mu \boldsymbol{I} \right\|_F \leq \tau\mu, \\ \text{where } \mu = \boldsymbol{X} \bullet \boldsymbol{S}/n. \end{array} \right\},$$

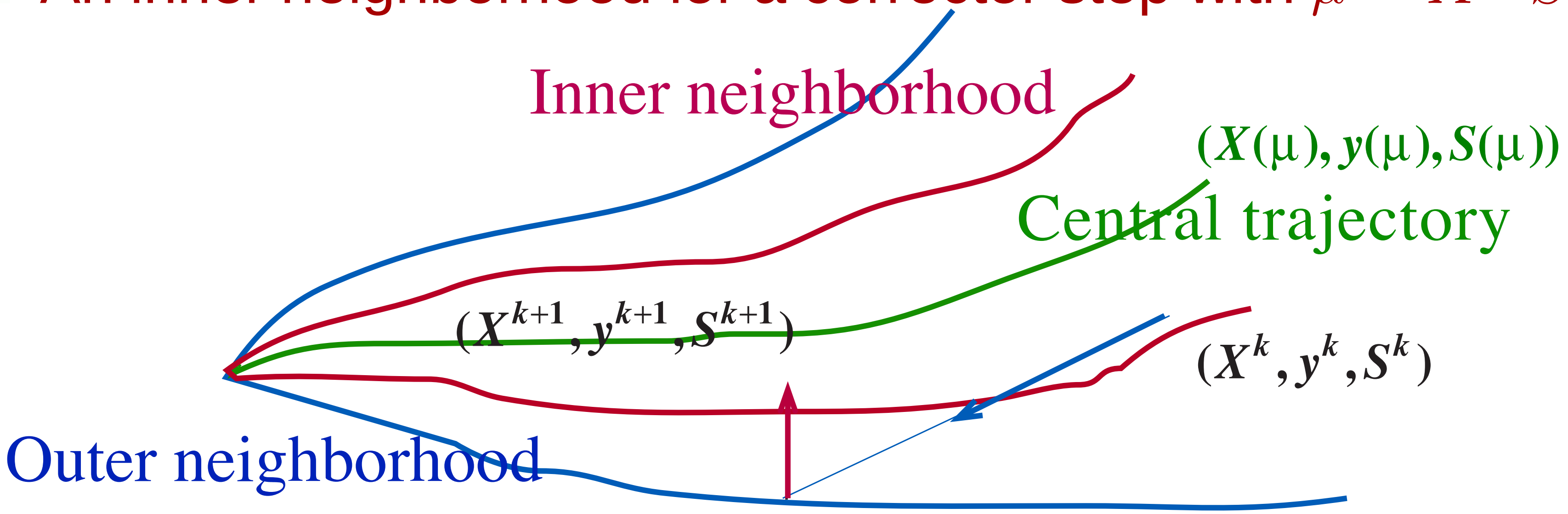
$$\hat{\mu} = (1 - \gamma/\sqrt{n})\boldsymbol{X}^k \bullet \boldsymbol{S}^k/n.$$

Here $\tau, \gamma \in (0,1)$, e.g., $\tau = \gamma = 0.5$.

$\Rightarrow O\left(\sqrt{n}\log\left(\frac{\boldsymbol{X}^0 \bullet \boldsymbol{S}^0}{\epsilon}\right)\right)$ iterations to attain $\boldsymbol{X}^k \bullet \boldsymbol{S}^k \leq \epsilon$

(b) Mizuno-Todd-Ye Predictor-corrector

- An outer neighborhood for a predictor step with $\hat{\mu} = 0$
- An inner neighborhood for a corrector step with $\hat{\mu} = X \bullet S/n$



- Not only polynomial-time but also superlinear linear convergence under a certain assumption; the strict complementarity and nondegeneracy of the unique optimal solution $(\boldsymbol{X}^*, \boldsymbol{y}^*, \boldsymbol{S}^*)$: $X^* + S^* \succ O$ and Inner and outer neighborhoods which enforce $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)$ to converge $(\boldsymbol{X}^*, \boldsymbol{y}^*, \boldsymbol{S}^*)$ tangentially to the central trajectory [21].

(P)	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+.$
(D)	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+.$

Primal-dual interior-point methods in practice

(i) An infeasible $(\boldsymbol{X}^0, \boldsymbol{y}^0, \boldsymbol{S}^0)$ such that $\boldsymbol{X}^0, \ \boldsymbol{S}^0 \succ \boldsymbol{O}$.

(ii) $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{S}^k)$; $\boldsymbol{X}^k, \boldsymbol{S}^k \succ \boldsymbol{O}$, $\left|\boldsymbol{A}_p \bullet \boldsymbol{X}^k - b_p\right| \to 0$,
$\left\|\boldsymbol{A}_p y_p^k + \boldsymbol{S}^k - \boldsymbol{A}_0\right\| \to 0$ and $X^k \bullet \boldsymbol{S}^k \to 0$.

(iii) Step size control: Let $\gamma \in (0,1)$, e.g., $\gamma = 0.80 \sim 0.98$,

$$\begin{aligned}
\alpha_p &= \gamma \times \max\{\alpha \in [0,1] : \boldsymbol{X}^k + \alpha d\boldsymbol{X} \succeq \boldsymbol{O}\}, \\
\alpha_d &= \gamma \times \max\{\alpha \in [0,1] : \boldsymbol{S}^k + \alpha d\boldsymbol{S} \succeq \boldsymbol{O}\}, \\
\boldsymbol{X}^{k+1} &= \boldsymbol{X}^k + \alpha_p d\boldsymbol{X}, \\
(\boldsymbol{y}^{k+1}, \boldsymbol{S}^{k+1}) &= (\boldsymbol{y}^k, \boldsymbol{S}^k) + \alpha_d(d\boldsymbol{y}, d\boldsymbol{S})
\end{aligned}$$

(Conservative compared to the LP case).

(P)	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+.$
(D)	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+.$

Primal-dual interior-point methods in practice

(iv) Mehrotra's predictor-corrector

- A reasonable choice of the target point $(\boldsymbol{X}(\beta\mu^k), \boldsymbol{y}(\beta\mu^k), \boldsymbol{S}(\beta\mu^k))$ on the central trajectory, where $\mu^k = \boldsymbol{X}^k \bullet \boldsymbol{S}^k / n$.
- A 2nd-order correction in computing the search direction $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{S})$.

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

P	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$.
D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$.

Important features in practice — large-scale SDPs arise often.

- $n \times n$ matrix variables $\boldsymbol{X}, \ \boldsymbol{S} \in \mathcal{S}^n$, each of which involves $n(n+1)/2$ real variables; for example, $n = 2000 \Rightarrow$ $n(n+1)/2 \approx 2$ million.
- m linear equality constraints in $\mathcal{P}$ or m $\boldsymbol{A}_p$'s $\in \mathcal{S}^n$.

Data matrices $\boldsymbol{A}_p \ (p = 1, \ldots, m)$ are sparse!

21 benchmark problems with $n \geq 500$ from SDPLIB [53]

the ratio of nonzero elements in $\boldsymbol{A}_p$s	10^{-2}	$\sim$	10^{-4}	$\sim$	10^{-6}	$\sim$	10^{-8}
# of problems		7		11		3	

$\Downarrow$

$\Diamond$ Exploit sparsity and structured sparsity

$\Diamond$ Enormous computational power$\Rightarrow$ parallel computation

P	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	s.t.	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+.$
D	max	$\sum_{p=1}^m b_p y_p$	s.t.	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+.$

Structured sparsity

The aggregate sparsity pattern $\widehat{\boldsymbol{A}}$: a symbolic $n \times n$ matrix:

$$\widehat{\boldsymbol{A}}_{ij} = \begin{cases} \star & \text{if the } (i,j)\text{th element of } \boldsymbol{A}_p \text{ is nonzero for } \exists p \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\star$ denotes a nonzero number.

Example: $m = 1$

$$\boldsymbol{A}_0 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \ \boldsymbol{A}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow \widehat{\boldsymbol{A}} = \begin{pmatrix} \star & \star & 0 \\ \star & \star & \star \\ 0 & \star & \star \end{pmatrix}.$$

Next — three types of structured sparsity

The aggregate sparsity pattern $\widehat{\boldsymbol{A}}$: a symbolic $n \times n$ matrix:

$$\widehat{\boldsymbol{A}}_{ij} = \begin{cases} \star & \text{if the } (i,j)\text{th element of } \boldsymbol{A}_p \text{ is nonzero for } \exists p \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\star$ denotes a nonzero number.

Structured sparsity-1 : $\widehat{\boldsymbol{A}}$ is block-diagonal.

Then $\boldsymbol{X}, \ \boldsymbol{S}$ have the same diagonal block structure as $\widehat{\boldsymbol{A}}$.

$$\widehat{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{B}_3 \end{pmatrix}, \ \boldsymbol{B}_i \text{ : symmetric.}$$

Example: CH_3N : an SDP from quantum chemistry, Fukuda et al. [13], Zhao et al. [51].

$m = 20,709, \ n = 12,802,$ "the number of blocks in $\widehat{\boldsymbol{A}}$" $= 22,$
the largest bl.size $= 3,211 \times 3,211,$
the average bl.size $= 583 \times 583.$

The aggregate sparsity pattern $\widehat{\boldsymbol{A}}$: a symbolic $n \times n$ matrix:

$$\widehat{\boldsymbol{A}}_{ij} = \begin{cases} \star & \text{if the } (i,j)\text{th element of } \boldsymbol{A}_p \text{ is nonzero for } \exists p \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\star$ denotes a nonzero number.

Structured sparsity-2 : $\widehat{\boldsymbol{A}}$ has a sparse Cholesky factorization.

"a small bandwidth" "a small bandwidth + bordered"

$$\widehat{\boldsymbol{A}} = \begin{pmatrix} \star & \star & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \star & \star & \star & \boldsymbol{O} & \boldsymbol{O} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \star & \star & \star \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \star & \star \end{pmatrix}, \quad \widehat{\boldsymbol{A}} = \begin{pmatrix} \star & \star & \boldsymbol{O} & \boldsymbol{O} & \star \\ \star & \star & \star & \boldsymbol{O} & \star \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \star & \star & \star \\ \star & \star & \cdots & \star & \star \end{pmatrix}$$

- $\boldsymbol{S}$: the same sparsity pattern as $\widehat{\boldsymbol{A}}$. • $\boldsymbol{X}$: fully dense.
- $\boldsymbol{X}^{-1}$: the same sparsity pattern as $\widehat{\boldsymbol{A}} \Rightarrow$ Use $\boldsymbol{X}^{-1}$ instead $\boldsymbol{X}$ (the positive definite matrix completion used in SDPARA-C)

The aggregate sparsity pattern $\widehat{\boldsymbol{A}}$: a symbolic $n \times n$ matrix:

$$\widehat{\boldsymbol{A}}_{ij} = \begin{cases} \star & \text{if the } (i,j)\text{th element of } \boldsymbol{A}_p \text{ is nonzero for } \exists p \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\star$ denotes a nonzero number.

Structured sparsity-3 : bl-diagonal $\widehat{\boldsymbol{A}}$ + blockwise orthogonal,

for most pairs (p,q) $1 \leq p < q \leq m$,

$\boldsymbol{A}_p$ and $\boldsymbol{A}_q$ do not share nonzero blocks; $\boldsymbol{A}_p \bullet \boldsymbol{A}_q = 0$

$\Rightarrow$ efficient computation of search directions

$$\boldsymbol{A}_1 = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}, \quad \boldsymbol{A}_2 = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_{22} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{23} \end{pmatrix}$$

- An engineering application, Ben-Tal et al. [52].
- A sparse SDP relaxation of poly. opt., Waki et al. [44].

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

Optimization Technology Center
http://www.ece.northwestern.edu/OTC/

$\Downarrow$

NEOS Solvers
http://www-neos.mcs.anl.gov/neos/solvers/index.html

$\Downarrow$

• Semidefinite Programming

software	lang.	method
csdp	c	p-d ipm
pensdp	matalb	augmented Lagrangian
sdpa	c++	p-d ipm
sdpt3	matlab	p-d ipm
sedumi	matlab	p-d ipm , self-dual embedding
. . .	. . .	. . .

- Binary and/or source codes are available.
- SDPA sparse format for all packages, matlab interface.
- Online solver — submit your SDP problem through Internet.

Some remarks on software packages.

- SDPs are more difficult to solve than LPs.
  - Degeneracy, no interior points in primal or dual SDPs.
  - Large scale problems.
- More accuracy requires more cpu time.
- Some package can solve SDPs faster with low accuracy.
- Sparse structure of SDPs.
- Some SDPs can be solved faster and/or more accurately by one package, but other SDPs by some other else.

Try some software packages that fit your problem.

SDPA Online Solver

http://sdpara.r.dendai.ac.jp/portal/

- SDPA on a single cpu.
- SDPARA on Opteron 32 cpu (1.4GHz), Xeon 16cpu (2.8GHz).
- SDPARA-C on Opteron 32 cpu (1.4GHz), Xeon 16cpu (2.8GHz).
- Submit your problem and choose one of the packages.

Chapter 2: Primal-dual interior-point methods

1. Existing numerical methods for SDPs
2. Three different approaches to primal-dual interior-point methods for SDPs
3. The central trajectory
4. Search directions
5. Various primal-dual interior-point methods
6. Exploiting sparsity
7. Software packages
8. Numerical results

Sparse SDP relaxation [44]: min g.Rosenbrock function

$f(\boldsymbol{x}) = \sum_{i=2}^{n} \left(100(x_i - x_{i-1}^2)^2 + (1 - x_i)^2\right), \; x_1 \geq 0.$

- SeDuMi.
- a single cpu, 2.4GHz Xeon.

		cpu in sec.	
n	ϵ_{obj}	Sparse	Dense
10	2.5e-08	0.2	10.6
15	6.5e-08	0.2	756.6
400	2.5e-06	3.7	—
800	5.5e-06	6.8	—

$$\epsilon_{\text{obj}} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}}.$$

When $n = 800$, SDP relaxation problem:

- $\boldsymbol{A}_p : 4794 \times 4794 \; (p = 1, 2, \ldots, 7, 988)$
- Each $\boldsymbol{A}_p$: 799 diagonal blocks with 6×6 matrices
- Structured sparsity, bl-diagonal + bl-wise orthogonal

$$\mathcal{P}: \quad \min \quad \boldsymbol{A}_0 \bullet \boldsymbol{X} \quad \text{sub.to} \quad \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$$
$$\mathcal{D}: \quad \max \quad \sum_{p=1}^m b_p y_p \quad \text{sub.to} \quad \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$$

From quantum chemistry, Fukuda et al. [13], Zhao et al. [51].

problem	m	n	#blocks	the sizes of largest blocks
O	7230	5990	22	[1450, 1450, 450, ...]
HF	15018	10146	22	[2520, 2520, 792, ...]
CH_3N	20709	12802	22	[3211, 3211, 1014, ...]

Parallel computation: cpu time in second

# of processors	16	64	128	256
O	14250.6	4453.3	3281.1	2951.6
HF	*	*	26797.1	20780.7
CH_3N	*	*	57034.8	45488.9

$\mathcal{P}$:	min	$\boldsymbol{A}_0 \bullet \boldsymbol{X}$	sub.to	$\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$
$\mathcal{D}$:	max	$\sum_{p=1}^m b_p y_p$	sub.to	$\sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$

Large-size SDPs by SDPARA-C [31] (64 CPUs)

3 types of test Problems:

(a) SDP relaxations of max. cut problems on lattice graphs with size 10×1000, 10×2000 and 10×4000.

(b) SDP relaxations of max. clique problems on lattice graphs with size 10×500, 10×1000 and 10×2000.

(c) Norm minimization problems

$$\text{min.} \ \left\| \boldsymbol{F}_0 - \sum_{i=1}^{10} \boldsymbol{F}_i y_i \right\| \ \text{sub.to} \ \ y_i \in \mathbb{R} \ (i = 1, 2, \ldots, 10)$$

where $\boldsymbol{F}_i : 10 \times 9990, \ 10 \times 19990$ or 10×39990 and $\|\boldsymbol{G}\| =$ the square root of the max. eigenvalue of $\boldsymbol{G}^T\boldsymbol{G}$.

In all cases, the aggregate sparsity pattern consists of one block and is very sparse.

$$\mathcal{P}: \quad \min \quad \boldsymbol{A}_0 \bullet \boldsymbol{X} \quad \text{sub.to} \quad \boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (\forall p), \ \boldsymbol{X} \in \mathcal{S}^n_+$$
$$\mathcal{D}: \quad \max \quad \sum_{p=1}^m b_p y_p \quad \text{sub.to} \quad \sum_{p=1}^m \boldsymbol{A}_p y_p + \boldsymbol{S} = \boldsymbol{A}_0, \ \boldsymbol{S} \in \mathcal{S}^n_+$$

Large-size SDPs by SDPARA-C (64 CPUs)

	Problem	n	m	time (s)	memory (MB)
	Cut(10×1000)	10000	10000	274.3	126
(a)	Cut(10×2000)	20000	20000	1328.2	276
	Cut(10×4000)	40000	40000	7462.0	720
	Clique(10×500)	5000	9491	639.5	119
(b)	Clique(10×1000)	10000	18991	3033.2	259
	Clique(10×2000)	20000	37991	15329.0	669
	Norm(10×9990)	10000	11	409.5	164
(c)	Norm(10×19990)	20000	11	1800.9	304
	Norm(10×39990)	40000	11	7706.0	583

Some exercises, 10/10/2008

Exercise 8. Describe a pair of the equality standard form SDP and its dual.

Exercise 9. Explain the basic idea of the primal-dual interior-point for SDPs briefly.

Exercise 10. Give the definition and the role of the log barrier used in the primal-dual interior-point for SDPs.

Exercise 11. Give the definition of the central trajectory for a pair of primal and dual SDPs.

Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

Let $\boldsymbol{F}_p$ be an $k \times \ell$ matrix $(0 \leq p \leq m)$. Approximate the matrix $\boldsymbol{F}_0$ as a linear combination of $\boldsymbol{F}_p$ $(1 \leq p \leq m)$;

minimize $\{\|\boldsymbol{F}(\boldsymbol{x})\| : \boldsymbol{x} \in \mathbb{R}^m\}$,

where $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 - \sum_{p=1}^{m} \boldsymbol{F}_p x_p$ for $\forall \boldsymbol{x} = (x_1, \ldots, x_m)^T$.

- Which norm?

$$\|\boldsymbol{A}\|_\infty = \max\{|A_{ij}| : 1 \leq i \leq k,\ 1 \leq j \leq \ell\}$$ (the ∞ norm)

$$\|\boldsymbol{A}\|_F = \left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} A_{ij}^2\right)^{1/2}$$ (the Frobenius norm)

$$\|\boldsymbol{A}\|_{L_2} = \max_{\|\boldsymbol{u}\|_2=1} \|\boldsymbol{A}\boldsymbol{u}\| = \big(\text{the maximum eigenvalue of } \boldsymbol{A}^T\boldsymbol{A}\big)^{1/2}$$

(the L_2 operator norm).

Let $\boldsymbol{F}_p$ be an $k \times \ell$ matrix $(0 \leq p \leq m)$. Approximate the matrix $\boldsymbol{F}_0$ as a linear combination of $\boldsymbol{F}_p$ $(1 \leq p \leq m)$;

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\| : \boldsymbol{x} \in \mathbb{R}^m\},$$

where $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 - \sum_{p=1}^m \boldsymbol{F}_p x_p$ for $\forall \boldsymbol{x} = (x_1, \ldots, x_m)^T$.

$$\|\boldsymbol{A}\|_\infty = \max\{|A_{ij}| : 1 \leq i \leq k,\ 1 \leq j \leq \ell\} \text{ (the } \infty \text{ norm)}$$

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\|_\infty : \boldsymbol{x} \in \mathbb{R}^m\}$$

$$\Downarrow$$

$$\text{minimize } \max\{|F_{ij}(\boldsymbol{x})| : 1 \leq i \leq k,\ 1 \leq j \leq \ell\}$$

$$\Downarrow$$

$$\text{minimize } \zeta \text{ sub.to } -\zeta \leq F_{ij}(\boldsymbol{x}) \leq \zeta \ (1 \leq i \leq k,\ 1 \leq j \leq \ell)$$

LP (Linear Programming)

Let $\boldsymbol{F}_p$ be an $k \times \ell$ matrix $(0 \leq p \leq m)$. Approximate the matrix $\boldsymbol{F}_0$ as a linear combination of $\boldsymbol{F}_p$ $(1 \leq p \leq m)$;

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\| : \boldsymbol{x} \in \mathbb{R}^m\},$$

where $\boldsymbol{F}(\boldsymbol{x}) \;=\; \boldsymbol{F}_0 - \sum_{p=1}^m \boldsymbol{F}_p x_p$ for $\forall \boldsymbol{x} = (x_1, \ldots, x_m)^T$.

$$\|\boldsymbol{A}\|_F \;=\; \left(\sum_{i=1}^{k} \sum_{j=1}^{\ell} A_{ij}^2 \right)^{1/2} \quad \text{(the Frobenius norm)}$$

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\|_F : \boldsymbol{x} \in \mathbb{R}^m\}$$

$$\Downarrow$$

$$\text{minimize}_{\boldsymbol{x} \in \mathbb{R}^m} \; \|\boldsymbol{F}(\boldsymbol{x})\|_F^2 \equiv \sum_{i=1}^{k} \sum_{j=1}^{\ell} F_{ij}(\boldsymbol{x})^2$$

the least square problem

convex QP (quadratic Programming)

Let $\boldsymbol{F}_p$ be an $k \times \ell$ matrix $(0 \leq p \leq m)$. Approximate the matrix $\boldsymbol{F}_0$ as a linear combination of $\boldsymbol{F}_p$ $(1 \leq p \leq m)$;

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\| : \boldsymbol{x} \in \mathbb{R}^m\},$$

where $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 - \sum_{p=1}^m \boldsymbol{F}_p x_p$ for $\forall \boldsymbol{x} = (x_1, \ldots, x_m)^T$.

$$\|\boldsymbol{A}\|_{L_2} = \max_{\|\boldsymbol{u}\|_2=1} \|\boldsymbol{A}\boldsymbol{u}\| = \left(\text{the maximum eigenvalue of } \boldsymbol{A}^T\boldsymbol{A}\right)^{1/2}$$

(the L_2 operator norm)

$$\text{minimize } \left\{\|\boldsymbol{F}(\boldsymbol{x})\|_{L_2} : \boldsymbol{x} \in \mathbb{R}^m\right\}$$

$$\Downarrow$$

minimize "the maximum eigenvalue of $\boldsymbol{F}(\boldsymbol{x})^T\boldsymbol{F}(\boldsymbol{x})$"

$$\Downarrow$$

minimize λ subject to $\lambda\boldsymbol{I} - \boldsymbol{F}(\boldsymbol{x})^T\boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{O}$

$\Downarrow$ the Schur complement

minimize λ subject to $\begin{pmatrix} \boldsymbol{I} & \boldsymbol{F}(\boldsymbol{x}) \\ \boldsymbol{F}(\boldsymbol{x})^T & \lambda\boldsymbol{I} \end{pmatrix} \succeq \boldsymbol{O}$ (SDP)

Exercise 12: Identify each $k \times \ell$ matrix $\boldsymbol{F}_p$ with a $k\ell$ column vector placing ℓ k-dimensional columns of $\boldsymbol{F}_p$ into one column. Then formulate the problem

$$\text{minimize } \{\|\boldsymbol{F}(\boldsymbol{x})\|_F : \boldsymbol{x} \in \mathbb{R}^m\}$$

as a standard least square problem, and derive the normal equation.

Exercise 13: Prove the identity

$$\|\boldsymbol{A}\boldsymbol{u}\| = \left(\text{the maximum eigenvalue of } \boldsymbol{A}^T\boldsymbol{A}\right)^{1/2}$$

Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

An SDP relaxation of QOP

$$\min \boldsymbol{x}^T \boldsymbol{Q}_0 \boldsymbol{x} + 2\boldsymbol{q}_0^T \boldsymbol{x} \text{ s.t. } \boldsymbol{x}^T \boldsymbol{Q}_p \boldsymbol{x} + 2\boldsymbol{q}_p^T \boldsymbol{x} + \gamma_p \leq 0 \ (1 \leq p \leq m)$$

$$\Updownarrow \ \widehat{\boldsymbol{Q}}_p = \begin{pmatrix} \gamma_p & \boldsymbol{q}_p^T \\ \boldsymbol{q}_p & \boldsymbol{Q}_p \end{pmatrix}, \ \gamma_0 = 0.$$

$$\min \widehat{\boldsymbol{Q}}_0 \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \text{ s.t. } \widehat{\boldsymbol{Q}}_p \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \ (1 \leq p \leq m)$$

An SDP relaxation of QOP

$$\min \widehat{Q}_0 \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \text{ s.t. } \widehat{Q}_p \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \; (1 \leq p \leq m)$$

relaxation $\Downarrow$

$$\boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \succeq \boldsymbol{O} \text{ (lifting)}$$

SDP-P: $\min \widehat{Q}_0 \bullet \boldsymbol{Y}$ s.t. $\widehat{Q}_p \bullet \boldsymbol{Y} \leq 0 \; (\forall p), \; Y_{11} = 1, \; \boldsymbol{Y} \succeq \boldsymbol{O}$.

- If $\boldsymbol{x}$ is a feas.sol. of QOP then $\boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix}$ is a feas.sol. of SDP-P with the same obj. val. $\Rightarrow$ relaxation.
- If $\overline{\boldsymbol{Y}}$ is an opt. sol. of SDP-P and $\overline{\boldsymbol{Y}} = \begin{pmatrix} 1 & \bar{\boldsymbol{x}}^T \\ \bar{\boldsymbol{x}} & \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T \end{pmatrix}$ for some $\bar{\boldsymbol{x}}$ (*i.e.*, rank$(\overline{\boldsymbol{Y}}) = 1$) then $\bar{\boldsymbol{x}}$ is an opt. sol. of QOP.
  Exercise 14. Prove this statement.
- If $\boldsymbol{Q}_p \; (0 \leq p \leq m)$ are p. semidefinite, then QOP = SDP-P.

An SDP relaxation of QOP

$$\min \widehat{Q}_0 \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \text{ s.t. } \widehat{Q}_p \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \ (1 \leq p \leq m)$$

relaxation $\Downarrow$

$$\boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \succeq \boldsymbol{O} \text{ (lifting)}$$

SDP-P: min $\widehat{Q}_0 \bullet \boldsymbol{Y}$ s.t. $\widehat{Q}_p \bullet \boldsymbol{Y} \leq 0 \ (\forall p), \ Y_{11} = 1, \ \boldsymbol{Y} \succeq \boldsymbol{O}.$

- Let $\overline{\boldsymbol{Y}} = \begin{pmatrix} 1 & \bar{\boldsymbol{x}}^T \\ \bar{\boldsymbol{x}} & \overline{\boldsymbol{X}} \end{pmatrix}$ be an opt. sol. of SDP-P, and ξ a random variable from the multivariate normal distribution $N(\bar{\boldsymbol{x}}, \overline{\boldsymbol{X}} - \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T)$. Then

$$E\left(\widehat{Q}_p \bullet \begin{pmatrix} 1 & \xi^T \\ \xi & \xi\xi^T \end{pmatrix}\right) = \widehat{Q}_p \bullet \begin{pmatrix} 1 & \bar{\boldsymbol{x}}^T \\ \bar{\boldsymbol{x}} & \overline{\boldsymbol{X}} \end{pmatrix} = \widehat{Q}_p \bullet \overline{\boldsymbol{Y}}$$

$$= \begin{cases} \text{the opt. value of SDP-P} & \text{if } p = 0, \\ \leq 0 & \text{if } 1 \leq p \leq m. \end{cases}$$

Exercise 15. Prove the identity

$$E(\xi\xi^T) = E(\overline{\boldsymbol{X}})$$

under the assumption that ξ is a random variable from the multivariate normal distribution $N(\bar{\boldsymbol{x}}, \overline{\boldsymbol{X}} - \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T)$ (recall the definition of $N(\bar{\boldsymbol{x}}, \overline{\boldsymbol{X}} - \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T)$).

The Lagrangian dual of QOP

min $\boldsymbol{x}^T \boldsymbol{Q}_0 \boldsymbol{x} + 2\boldsymbol{q}_0^T \boldsymbol{x}$ s.t. $\boldsymbol{x}^T \boldsymbol{Q}_p \boldsymbol{x} + 2\boldsymbol{q}_p^T \boldsymbol{x} + \gamma_p \leq 0 \ (1 \leq p \leq m)$

Lagrangian function

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{x}^T \boldsymbol{Q}_p \boldsymbol{x} + 2\boldsymbol{q}_p^T \boldsymbol{x} + \sum_{i=1}^m \lambda_i \left(\boldsymbol{x}^T \boldsymbol{Q}_p \boldsymbol{x} + 2\boldsymbol{q}_p^T \boldsymbol{x} + \gamma_p \right),$$
$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m_+.$$

Lagrangian dual: max $\min\{L(\boldsymbol{x}, \boldsymbol{\lambda}) : \boldsymbol{x} \in \mathbb{R}^n\}$ s.t. $\boldsymbol{\lambda} \in \mathbb{R}^m_+$

$$\Updownarrow$$

L. dual: max. ζ s.t. $L(\boldsymbol{x}, \boldsymbol{\lambda}) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in \mathbb{R}^n)$, $\boldsymbol{\lambda} \in \mathbb{R}^m_+$

$$\Updownarrow$$

L. dual: max. ζ s.t.

$$\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}^T \begin{pmatrix} \sum_{i=1}^m \lambda_i \gamma_i - \zeta & \boldsymbol{q}_0^T + \sum_{i=1}^m \lambda_i \boldsymbol{q}_i^T \\ \boldsymbol{q}_0 + \sum_{i=1}^m \lambda_i \boldsymbol{q}_i & \boldsymbol{Q}_0 + \sum_{i=0}^m \lambda_i \boldsymbol{Q}_i \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \geq 0$$

$(\forall \boldsymbol{x} \in \mathbb{R}^n)$ ($\boldsymbol{x}$ is not a variable.), $\boldsymbol{\lambda} \in \mathbb{R}^m_+$

$\Updownarrow$

SDP-D: max ζ s.t.

$$\begin{pmatrix} \sum_{i=1}^m \lambda_i \gamma_i - \zeta & \boldsymbol{q}_0^T + \sum_{i=1}^m \lambda_i \boldsymbol{q}_i^T \\ \boldsymbol{q}_0 + \sum_{i=1}^m \lambda_i \boldsymbol{q}_i & \boldsymbol{Q}_0 + \sum_{i=0}^m \lambda_i \boldsymbol{Q}_i \end{pmatrix} \succeq \boldsymbol{O}, \ \boldsymbol{\lambda} \in \mathbb{R}^m_+.$$

- SDP-D above is the dual SDP of SDP-P.

Exercise 16. Prove the equivalence $\Updownarrow$ above.

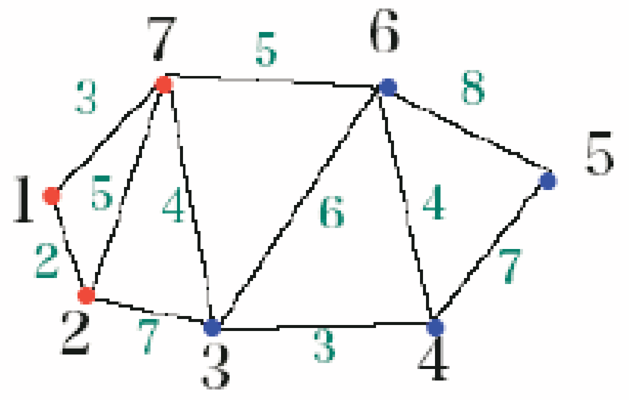
Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

[14] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM*, 42 (1995) 1115-1145.

The max-cut problem: Let $G = (N, E)$ be an undirected graph, and w_{ij} be weights of an edge $\{i, j\} \in E$.
For $\forall K \subset N$, let $\delta(K)$ denote $\{\{i, j\} : i \in K,\ j \notin K\}$ (the cut determined by K) and $w(\delta(K)) = \sum_{\{i,j\}\in\delta(K)} w_{ij}$.

Max-cut problem: max $w(\delta(K))$ s.t. $K \subset N$.



$N = \{1, 2, 3, 4, 5, 6, 7\}$, $w_{12} = w_{21} = 2, \ldots$

•$K = \{1, 2, 7\} \Rightarrow \delta(K) = \{\{2, 3\}, \{3, 7\}, \{6, 7\}\}$

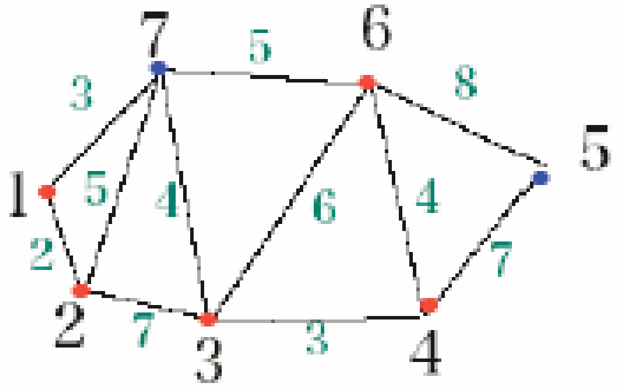
$$w(\delta(K)) = 7 + 4 + 5 = 16$$

$K = \{1, 2, 3, 4, 6\} \Rightarrow \delta(K) = \{\{1, 7\}, \{2, 7\}, \{3, 7\}, \{4, 5\}, \{5, 6\}, \{6, 7\}\}$

$$w(\delta(K)) = 3 + 5 + 4 + 7 + 8 + 5 = 32$$

The max-cut problem: Let $G = (N, E)$ be an undirected graph, and w_{ij} be weights of an edge $\{i, j\} \in E$.
For $\forall K \subset N$, let $\delta(K)$ denote $\{\{i, j\} : i \in K,\ j \notin K\}$ (the cut determined by K) and $w(\delta(K)) = \sum_{\{i,j\}\in\delta(K)} w_{ij}$.

Max-cut problem: max $w(\delta(K))$ s.t. $K \subset N$.



$N = \{1, 2, 3, 4, 5, 6, 7\}$, $w_{12} = w_{21} = 2, \ldots$

$K = \{1, 2, 7\} \Rightarrow \delta(K) = \{\{2, 3\}, \{3, 7\}, \{6, 7\}\}$

$w(\delta(K)) = 7 + 4 + 5 = 16$

$^{\bullet}K = \{1, 2, 3, 4, 6\} \Rightarrow \delta(K) = \{\{1, 7\}, \{2, 7\}, \{3, 7\}, \{4, 5\}, \{5, 6\}, \{6, 7\}\}$

$w(\delta(K)) = 3 + 5 + 4 + 7 + 8 + 5 = 32$

The max-cut problem: Let $G = (N, E)$ be an undirected graph, and w_{ij} be weights of an edge $\{i, j\} \in E$.
For $\forall K \subset N$, let $\delta(K)$ denote $\{\{i, j\} : i \in K,\ j \notin K\}$ (the cut determined by K) and $w(\delta(K)) = \sum_{\{i,j\}\in\delta(K)} w_{ij}$.

Max-cut problem: max $w(\delta(K))$ s.t. $K \subset N$.

Let $w_{ij} = 0$ if $\{i, j\} \notin E$, and let $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$; $x_i = \begin{cases} 1 & \text{if } i \in K, \\ -1 & \text{otherwise.} \end{cases}$ Then $w(\delta(K)) = \dfrac{1}{2}\sum_{i<j} w_{ij}(1 - x_i x_j) = \dfrac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(1 - x_i x_j) = \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}$, where $c_{ij} = -w_{ij}/4\ (i \neq j)$ and $c_{ii} = \sum_{j=1}^{n} w_{ij}$.

Exercise 17. Verify the identity $\dfrac{1}{4}\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(1 - x_i x_j) = \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}$.

$$\Downarrow$$

Max-cut prob. $\Leftrightarrow$ $c^* = \max \boldsymbol{C} \bullet \boldsymbol{x}^T\boldsymbol{x}$ s.t. $x_i^2 = 1 \ (i \in N)$

$\Longrightarrow$ relaxation

SDP: $\hat{c} = \max \boldsymbol{C} \bullet \boldsymbol{X}$
s.t. $X_{ii} = 1 \ (i \in N), \ \boldsymbol{X} \succeq \boldsymbol{O}$

- $c^* \leq \hat{c}$ Exercise 18. Show this inequality.
- How do we construct a cut from an opt.sol. $\widehat{\boldsymbol{X}}$ of SDP?

Step 1. Factorize $\widehat{\boldsymbol{X}}$ s.t. $\widehat{\boldsymbol{X}} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)^T (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$.
Step 2. Choose a vector ξ randomly from the unit sphere $\{\eta \in \mathbb{R}^n : \|\eta\| = 1\}$; hence ξ is a random variable vector.
Step 3. Let

$$\boldsymbol{x}_i(\xi) = \begin{cases} 1 & \text{if } \boldsymbol{v}_i^T\xi > 0, \\ -1 & \text{otherwise} \end{cases} \quad \text{or} \quad K(\xi) = \{i \in N : \boldsymbol{v}_i^T\xi > 0\}$$

$\Downarrow$

$$\frac{E(w(\delta(K(\xi))))}{\text{the value } c^* \text{ of max-cut}} \geq 0.878$$

Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

[22] J. B. Lasserre, Global optimization with polynomials and the problems of moments, *SIAM Journal on Optimization*, 11 (2001) 796–817.

[35] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems', *Mathematical Programming*, 96 (2003) 293-320.

$f(\boldsymbol{x})$: an SOS (Sum of Squares) polynomial

$$\Updownarrow$$

$$\exists \text{ polynomials } g_1(\boldsymbol{x}), \ldots, g_k(\boldsymbol{x});\ f(\boldsymbol{x}) = \sum_{i=1}^{k} g_i(\boldsymbol{x})^2.$$

$\mathcal{N}$: the set of nonnegative polynomials in $\boldsymbol{x} \in \mathbb{R}^n$.
SOS$_*$: the set of SOS. Obviously, **SOS**$_* \subset \mathcal{N}$.
SOS$_{2r} = \{f \in$ **SOS**$_*$: deg $f \leq 2r\}$: SOSs w. degree at most $2r$.

$n = 2$. $f(x_1, x_2) = (x_1^2 - 2x_2 + 1)^2 + (3x_1x_2 + x_2 - 4)^2 \in$ **SOS**$_4$.
$n = 2$. $f(x_1, x_2) = (x_1x_2 - 1)^2 + x_1^2 \in$ **SOS**$_4$.

- In theory, **SOS**$_*$ (SOS) $\subset \mathcal{N}$. **SOS**$_* \neq \mathcal{N}$ in general.
- If $n = 1$, **SOS**$_* = \mathcal{N}$. $\{f \in \mathcal{N}$: deg $f \leq 2\} \equiv$ **SOS**$_2$.
- In practice, $f(\boldsymbol{x}) \in \mathcal{N} \backslash$**SOS**$_*$ is rare.
- So we replace $\mathcal{N}$ by **SOS**$_* \Longrightarrow$ SOS Relaxations.

Exercise 19. Show $f(x_1, x_2) \equiv (x_1x_2 - 1)^2 + x_1^2 > 0$ for every $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ and $\inf_{\boldsymbol{x} \in \mathbb{R}^2} f(x_1, x_2) = 0$.

Representation of

$$SOS_{2r} \equiv \left\{ \sum_{j=1}^{k} g_j(\boldsymbol{x})^2 : k \geq 1,\ g_j(\boldsymbol{x}) \text{is a poly. of deg } \leq r \right\} \subset SOS_*.$$

$\forall$ poly. $g(\boldsymbol{x})$ of deg $\leq r$, $\exists \boldsymbol{a} \in \mathbb{R}^{d(r)}$; $g(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{u}_r(\boldsymbol{x})$, where

$$\boldsymbol{u}_r(\boldsymbol{x}) = (1, x_1, \ldots, x_n, x_1^2, x_1x_2, x_1x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r)^T$$

(a column vector of a basis for polynomials of degree $\leq r$),

$$d(r) = \begin{pmatrix} n+r \\ r \end{pmatrix}$$: the dimension of $\boldsymbol{u}_r(\boldsymbol{x})$.

$$\Downarrow$$

$$\begin{aligned} SOS_{2r} &= \left\{ \sum_{j=1}^{k} \left(\boldsymbol{a}_j^T \boldsymbol{u}_r(\boldsymbol{x}) \right)^2 : k \geq 1,\ \boldsymbol{a}_j \in \mathbb{R}^{d(r)} \right\} \\ &= \left\{ \boldsymbol{u}_r(\boldsymbol{x})^T \left(\sum_{j=1}^{k} \boldsymbol{a}_j \boldsymbol{a}_j^T \right) \boldsymbol{u}_r(\boldsymbol{x}) : k \geq 1,\ \boldsymbol{a}_j \in \mathbb{R}^{d(r)} \right\} \\ &= \left\{ \boldsymbol{u}_r(\boldsymbol{x})^T \boldsymbol{V} \boldsymbol{u}_r(\boldsymbol{x}) : \boldsymbol{V} \succeq \boldsymbol{O} \right\}. \end{aligned}$$

Example. $n = 1$, SOS polynomials of degree ≤ 3 in $x \in \mathbb{R}$.

$$SOS_6 \equiv \left\{ \sum_{i=1}^{k} g_i(\boldsymbol{x})^2 \;:\; k \geq 1,\; g_i(\boldsymbol{x}) \text{ is a poly. of degree } \leq 3 \right\}$$

$$= \left\{ \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \end{pmatrix}^T \boldsymbol{V} \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \end{pmatrix} \;:\; \boldsymbol{V} \text{ is } 4 \times 4 \text{ psd matrix} \right\}$$

Example. $n = 2$, SOS polynomials of degree ≤ 2 in x=(x_1, x_2).

$$SOS_4 \equiv \left\{ \sum_{i=1}^{k} g_i(\boldsymbol{x})^2 \; : \; k \geq 1, \; g_i(\boldsymbol{x}) \text{ is a poly. of degree } \leq 2 \right\}$$

$$= \left\{ \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}^T \boldsymbol{V} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix} \; : \; \boldsymbol{V} \text{ is a } 6 \times 6 \text{ psd matrix} \right\}$$

$f(\boldsymbol{x}) = -x_1 + 2x_2 + 3x_1^2 - 5x_1^2x_2^2 + 7x_2^4$ $\zeta = 3.1$: fixed

$f(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x}) \Longrightarrow$ LMI

$f(\boldsymbol{x}) - \zeta \in \mathit{SOS}_4$ (SOS of poly. of degree ≤ 2)

$\Updownarrow$

$\exists \boldsymbol{V} \in \mathbb{S}^6; \qquad f(\boldsymbol{x}) - \zeta =$

Sum of Squares

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} V_{11} & V_{12} & V_{13} & V_{14} & V_{15} & V_{16} \\ V_{12} & V_{22} & V_{23} & V_{24} & V_{25} & V_{26} \\ V_{13} & V_{23} & V_{33} & V_{34} & V_{35} & V_{36} \\ V_{14} & V_{24} & V_{34} & V_{44} & V_{45} & V_{46} \\ V_{15} & V_{25} & V_{35} & V_{45} & V_{55} & V_{56} \\ V_{16} & V_{26} & V_{36} & V_{46} & V_{56} & V_{66} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}$$

$(\forall (x_1, x_2)^T \in \mathbb{R}^n), \qquad \boldsymbol{V} \succeq \boldsymbol{O}$

$\Updownarrow$ Compare the coef. of $1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2, \boldsymbol{x}_2^2$ on both sides of $=$

LMI: $\exists \boldsymbol{V} \in \mathbb{S}^6$?;

$$-\zeta = V_{11},\ -1 = 2V_{12},\ 2 = 2V_{13},\ 3 = 2V_{14} + V_{22},$$
$$-5 = 2V_{46} + V_{55},\ 7 = V_{66},\ \text{all others}\ \ 0 = \cdots\ ,$$
$$\boldsymbol{V} \succeq \boldsymbol{O}$$

In general, each equality constraint is a linear equation in ζ and $\boldsymbol{V}$.

min $f(\boldsymbol{x}) = -x_1 + 2x_2 + 3x_1^2 - 5x_1^2x_2^2 + 7x_2^4$ ζ : variable

max ζ; $f(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x}) \Longrightarrow$ SDP

max ζ s.t. $f(\boldsymbol{x}) - \zeta \in \mathcal{SOS}_4$ (SOS of polynomials of degree ≤ 2)

$\Updownarrow$

max ζ s.t. $f(\boldsymbol{x}) - \zeta =$

Sum of Squares

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} V_{11} & V_{12} & V_{13} & V_{14} & V_{15} & V_{16} \\ V_{12} & V_{22} & V_{23} & V_{24} & V_{25} & V_{26} \\ V_{13} & V_{23} & V_{33} & V_{34} & V_{35} & V_{36} \\ V_{14} & V_{24} & V_{34} & V_{44} & V_{45} & V_{46} \\ V_{15} & V_{25} & V_{35} & V_{45} & V_{55} & V_{56} \\ V_{16} & V_{26} & V_{36} & V_{46} & V_{56} & V_{66} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}$$

$(\forall (x_1, x_2)^T \in \mathbb{R}^n)$, $\boldsymbol{V} \succeq \boldsymbol{O}$

$\Updownarrow$ Compare the coef. of $1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2, \boldsymbol{x}_2^2$ on both side of $=$

SDP (Semidefinite Program)

$$\max \quad \zeta \text{ s.t.} \quad -\zeta = V_{11},\ -1 = 2V_{12},\ 2 = 2V_{13},\ 3 = 2V_{14} + V_{22},$$
$$-5 = 2V_{46} + V_{55},\ 7 = V_{66},\ \text{all others } \ 0 = \cdots \ ,$$
$$\boldsymbol{V} \succeq \boldsymbol{O}$$

In general, each equality constraint is a linear equation in ζ and $\boldsymbol{V}$.

Chapter 3: Some applications

1. Matrix approximation problems
2. A nonconvex quadratic optimization problem
3. The max-cut problem
4. Sum of squares of polynomials
5. Sensor network localization problems

Sensor network localization problem: Let $s = 2$ or 3.

$$\boldsymbol{x}^p \in \mathbb{R}^s \quad : \quad \text{unknown location of sensors } (1 \leq p \leq m),$$

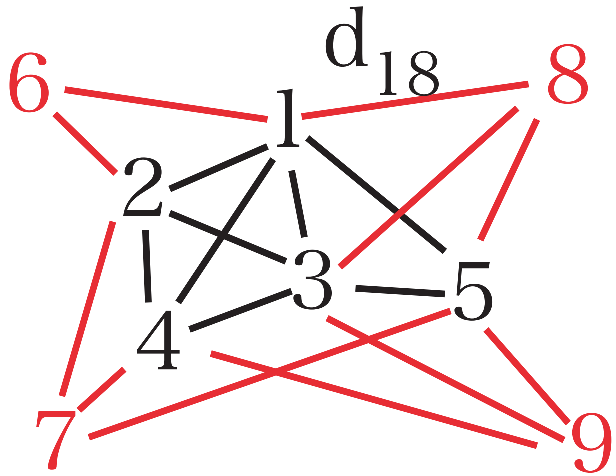
$$\boldsymbol{x}^r = \boldsymbol{a}^r \in \mathbb{R}^s \quad : \quad \text{known location of anchors } (m+1 \leq r \leq n),$$

$$d_{pq} \quad = \quad \|\boldsymbol{x}^p - \boldsymbol{x}^q\| + \epsilon_{pq} \text{ — given for } (p,q) \in \mathcal{N},$$

$$\mathcal{N} \quad = \quad \{(p,q) : \|\boldsymbol{x}^p - \boldsymbol{x}^q\| \leq \rho = \text{a given radio range}\}$$

Here ϵ_{pq} denotes a noise.

$m = 5, \ n = 9.$
$1, \ldots, 5$: sensors
$6, 7, 8, 9$: anchors



Anchors' positions are known.
A distance is given for $\forall$ edge.
Compute locations of sensors.

$\Rightarrow$ Some nonconvex QOPs

- SDP relaxation — FSDP by Biswas-Ye '06, ESDP by Wang et al '07, ... for $s = 2$.
- SOCP relaxation — Tseng '07 for $s = 2$.
- ...

Numerical results on 3 methods (a), (b) and (c) applied to a sensor network localization problem with

$m =$ the number of sensors dist. randomly in $[0, 1]^2$,
4 anchors located at the corner of $[0, 1]^2$,
$\rho =$ radio distance $= 0.1,$ no noise.

(a) FSDP
(b) FSDP + exploiting sparsity, as strong as (a)
(c) ESDP — a further relaxation of FSDP, weaker than (a);

m	SeDuMi cpu time in second (a)	(b)	(c)
500	389.1	35.0	62.5
1000	3345.2	60.4	200.3
2000		111.1	1403.9
4000		182.1	11559.8

A sensor network localization problem with

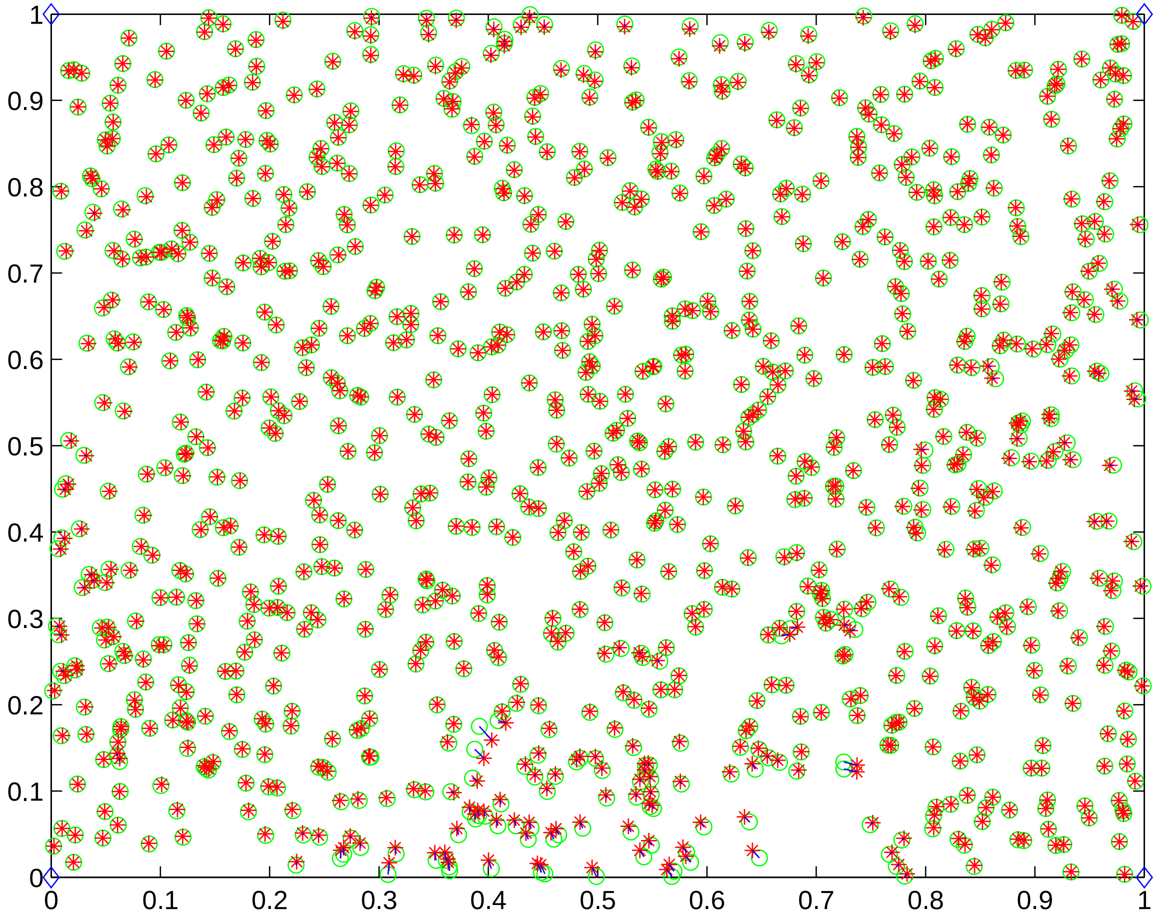
1000 sensors dist. randomly in $[0,1]^2$,

4 anchors located at the corner of $[0,1]^2$,

ρ = radio distance = 0.1,

no noise

(b) FSDP+ exploiting sparsity



anchor : ◇

true : ○

computed : $*$

deviation : —

A sensor network localization problem with

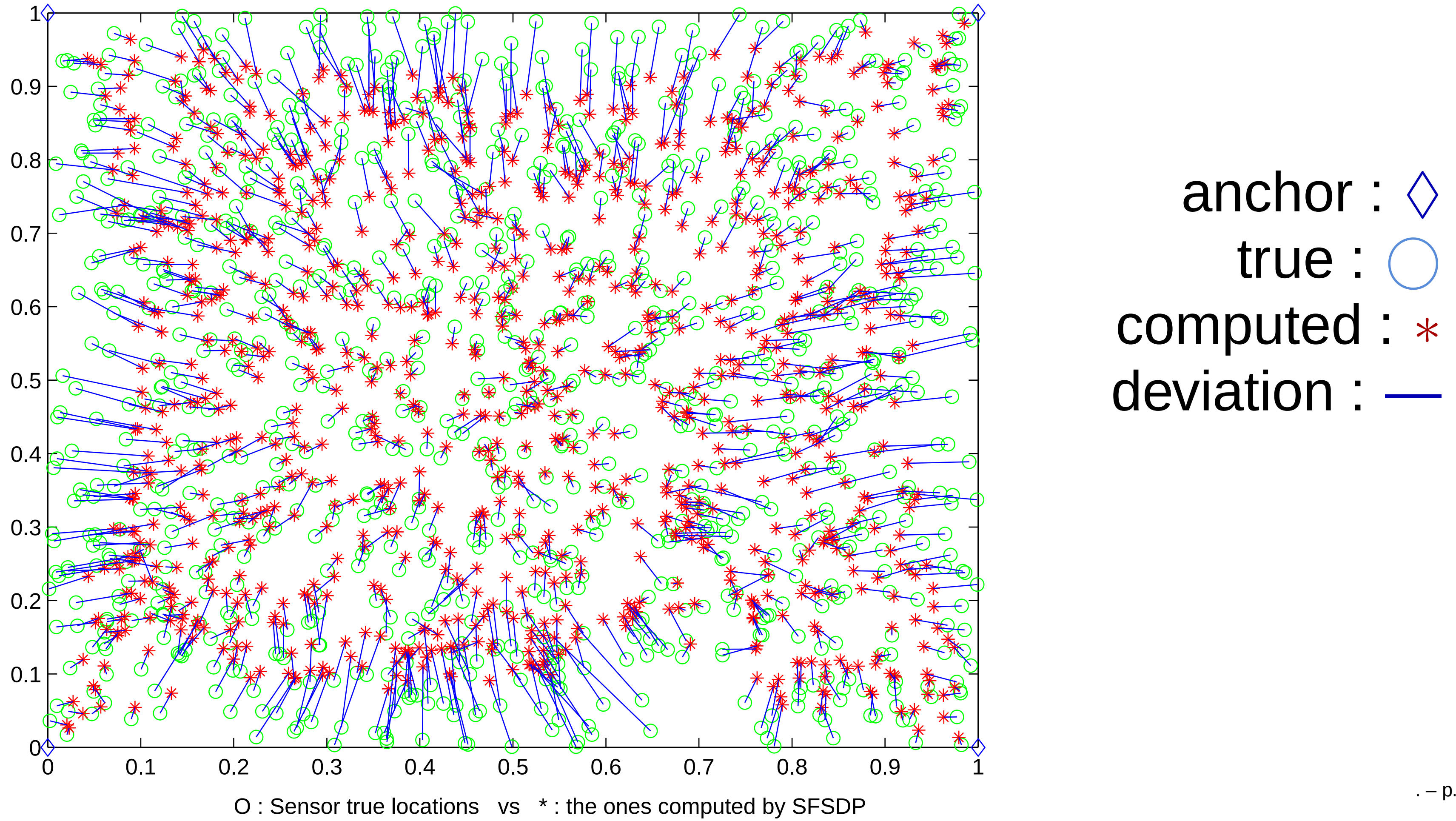
1000 sensors dist. randomly in $[0,1]^2$,

4 anchors located at the corner of $[0,1]^2$,

ρ = radio distance = 0.1,

noise = the true disttance $\times(1+\xi)$, $\xi \sim$ N(0,0.1).

(b) FSDP+ exploiting sparsity



A sensor network localization problem with

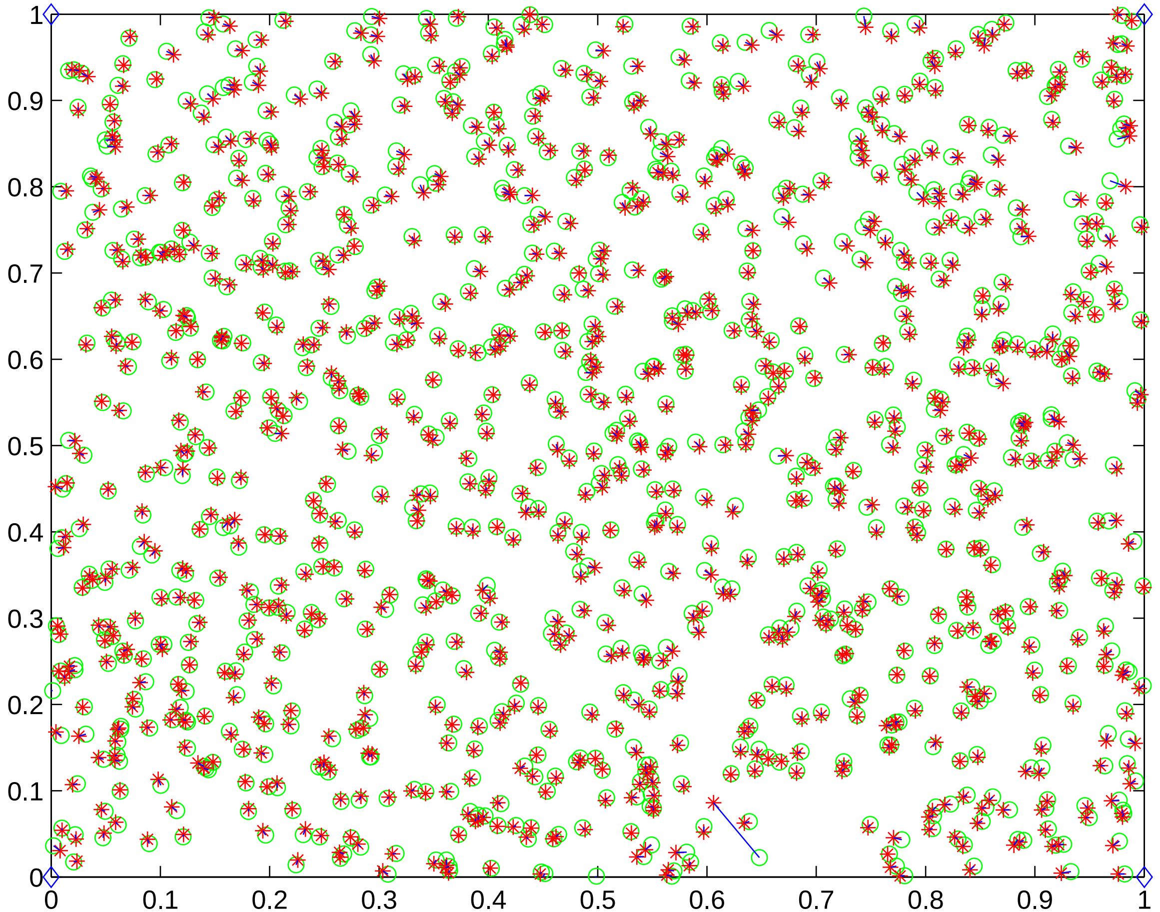
1000 sensors dist. randomly in $[0,1]^2$,

4 anchors located at the corner of $[0,1]^2$,

ρ = radio distance = 0.1,

noise = the true disttance $\times(1+\xi)$, $\xi \sim$ N(0,0.1).

(b) FSDP+ exploiting sparsity +Gradient method



anchor : ◇

true : ○

computed : $*$

deviation : —

O : Sensor true locations vs * : the ones computed by SFSDP + the refinepositions funct. by K. Toh

A sensor network localization problem with

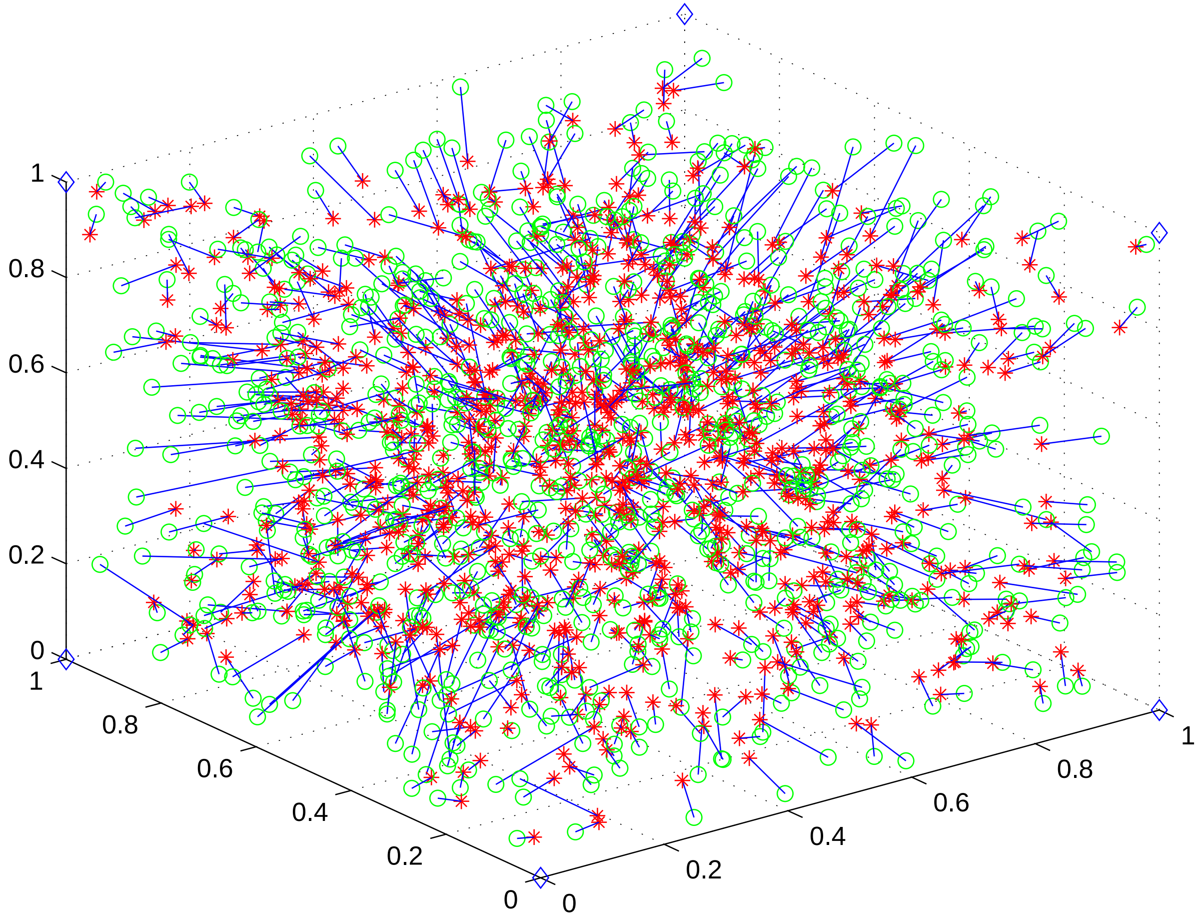
1000 sensors dist. randomly in $[0,1]^3$,

8 anchors located at the corner of $[0,1]^3$,

ρ = radio distance = 0.3,

noise = the true disttance $\times(1+\xi)$, $\xi \sim$ N(0,0.1).

(b) FSDP+ exploiting sparsity



anchor : ◇
true : ○
computed : $*$
deviation : —

A sensor network localization problem with

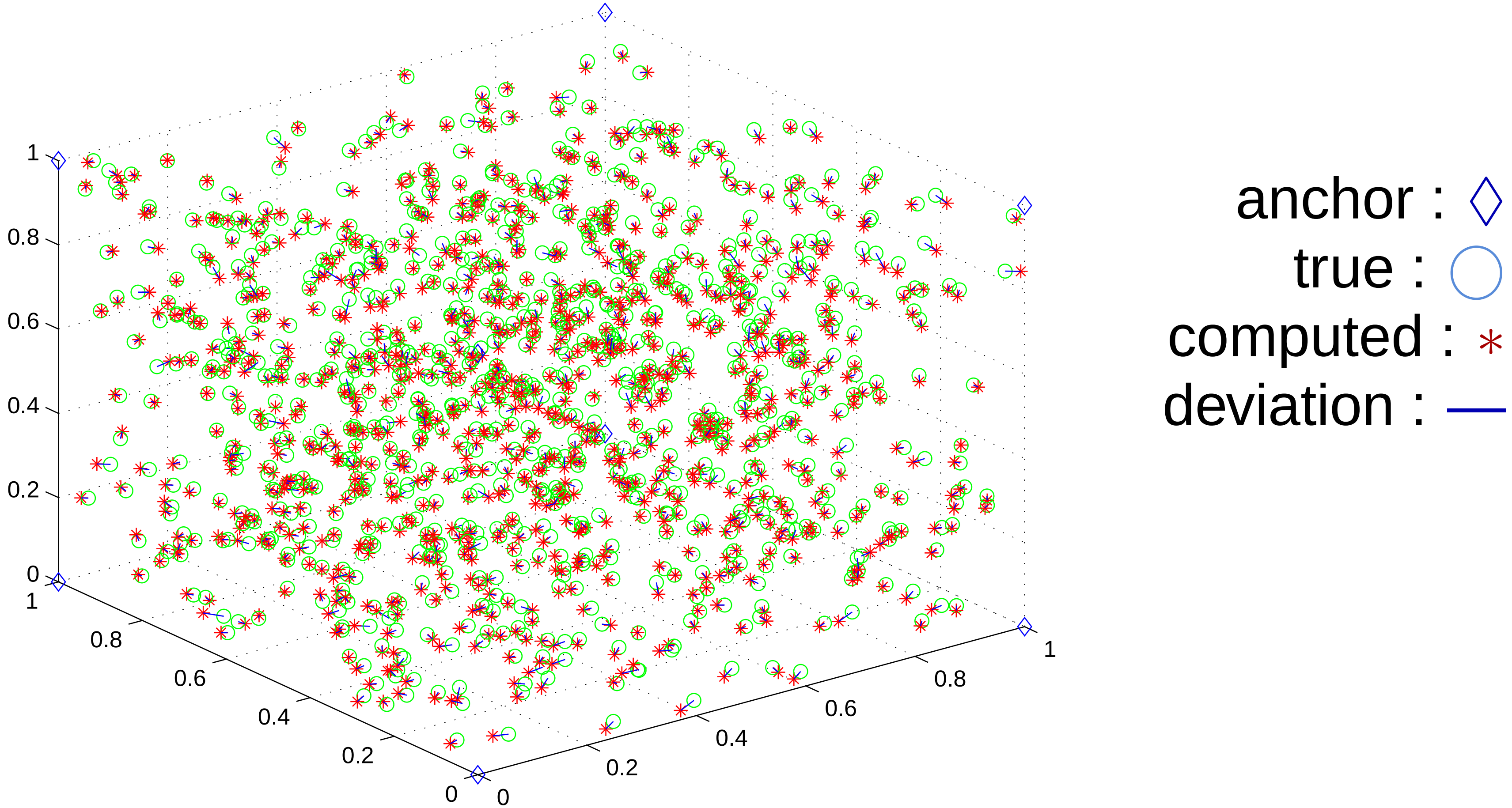
1000 sensors dist. randomly in $[0,1]^3$,

8 anchors located at the corner of $[0,1]^3$,

ρ = radio distance = 0.3,

noise = the true disttance$\times(1+\xi)$, $\xi \sim$ N(0,0.1).

(b) FSDP+ exploiting sparsity +Gradient method



Exercise 20. Describe the max-cut problem and formulate it as a 0-1 quadratic programming problem.

Exercise 21. Explain why the problem of minimizing $f(\boldsymbol{x})$ over $\boldsymbol{x} \in \mathbb{R}^n$ is equivalent to the problem

maximize ζ subject to $f(\boldsymbol{x}) - \zeta \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Exercise 22. Give the definition of an SOS polynomial of degree 2 in $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, and represent the set of such SOS polynomials in terms of a positive semidefinite matrix variable $\boldsymbol{V}$.

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP (Second Order Cone Program)
5. Some applications of SOCPs

[10] L. Faybusovich, Linear systems in Jordan algebra and primal-dual interior-point algorithms, *Journal of Computational and Applied Mathematics*, 86 (1997) 149-75.

[36] S. Schmieta and F. Alizadeh, Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones, *Mathematics of Operations Research*, 26 (2001) 543-564.

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP (Second Order Cone Program)
5. Some applications of SOCPs

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \le p \le m), \ \boldsymbol{x} \in K$

V : a finite dimensional vector space with an inner product $\langle \cdot, \cdot \rangle$,

$K \subset V$: closed, convex cone, nonempty-interior, $\nexists$ line $\subset \boldsymbol{K}$,

$\boldsymbol{a}_p \in V$: data $(0 \le p \le m)$, $b_p \in \mathbb{R}$: data,

$\boldsymbol{x} \in V$: variable,

$K^* = \{\boldsymbol{s} \in \boldsymbol{V} : \langle \boldsymbol{s}, \boldsymbol{x} \rangle \ge 0$ for $\forall \boldsymbol{x} \in K\}$ (the dual of K).

Examples

$\boldsymbol{V} = \mathbb{R}^n, \ K = \mathbb{R}^n_+ \ \Rightarrow$ LP; $K^* = K$

$\boldsymbol{V} = \mathbb{S}^n, \ K = \mathbb{S}^n_+ \ \Rightarrow$ SDP; $K^* = K$

$\boldsymbol{V} = \mathbb{R}^{1+n}, \ K = \{\boldsymbol{x} = (x_0, x_1) : x_0 \in \mathbb{R}, x_1 \in \mathbb{R}^n, x_0 \ge \|x_1\|\}$,

$\Rightarrow$ SOCP, Second Order Cone Program; $K^* = K$ — later.

$\boldsymbol{V} = \mathbb{S}^n, \ K = \left\{\boldsymbol{X} \in \mathbb{S}^n : \boldsymbol{u}^T \boldsymbol{X} \boldsymbol{u} \ge 0 \text{ for } \forall \boldsymbol{u} \in \mathbb{R}^n_+\right\} \supset \mathbb{S}^n_+$

$\Rightarrow$ Copositive Program; $K^* =$ cone$\{\boldsymbol{y}\boldsymbol{y}^T : \boldsymbol{y} \in \mathbb{R}^n_+\} \subset \mathbb{S}^n_+$

$\rightarrow$ graph partitioning, control theory.

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

V : a finite dimensional vector space with an inner product $\langle \cdot, \cdot \rangle$,

$K \subset V$: closed, convex cone, nonempty-interior, $\not\exists$ line $\subset \boldsymbol{K}$,

$\boldsymbol{a}_p \in V$: data $(0 \leq p \leq m), \ b_p \in \mathbb{R}$: data,

$\boldsymbol{x} \in V$: variable,

$K^* = \{ \boldsymbol{s} \in \boldsymbol{V} : \langle \boldsymbol{s}, \boldsymbol{x} \rangle \geq 0$ for $\forall \boldsymbol{x} \in K \}$ (the dual of K).

Convex program : min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$ sub.to $\boldsymbol{x} \in F$,

where F : compact, convex, nonempty interior.

$\Downarrow$

$$K \equiv \{ \lambda(1, \boldsymbol{x}) : \lambda \geq 0, \ \boldsymbol{x} \in F \}$$

$$\Downarrow$$

$\boldsymbol{x} \in F$ iff $(1, \boldsymbol{x}) \in K$

min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$ sub.to $\lambda = 1, \ (\lambda, \boldsymbol{x}) \in K$.

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

- V : a finite dimensional vector space with an inner product $\langle \cdot, \cdot \rangle$,
- $K \subset V$: closed, convex cone, nonempty-interior, $\nexists$ line $\subset \boldsymbol{K}$,
- $\boldsymbol{a}_p \in V$: data $(0 \leq p \leq m)$, $b_p \in \mathbb{R}$: data,
- $\boldsymbol{x} \in V$: variable,
- $K^* = \{ \boldsymbol{s} \in \boldsymbol{V} : \langle \boldsymbol{s}, \boldsymbol{x} \rangle \geq 0$ for $\forall \boldsymbol{x} \in K \}$ (the dual of K).

(D) (the dual of (P)): max $\sum_{p=1}^m b_p y_p$
sub.to $\sum_{p=1}^m \boldsymbol{a}_p y_p + \boldsymbol{s} = \boldsymbol{a}_0, \ \boldsymbol{s} \in K^*$.

- $0 \leq \langle \boldsymbol{x}, \boldsymbol{s} \rangle = \langle \boldsymbol{a}_0, \boldsymbol{x} \rangle - \sum_{p=1}^m b_p y_p$ for $\forall$ feasible $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$. (weak duality).
- If $\exists$ an int. feas. sol. $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ ($\boldsymbol{x} \in$ int $K, \boldsymbol{s} \in$ int K^*), then $0 = \langle \boldsymbol{a}_0, \boldsymbol{x}^* \rangle - \sum_{p=1}^m b_p y_p^*$ for $\forall$ opt. $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$. (strong duality)

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$ s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Self-concordance theory (Nesterov-Nemirovski [33]) $\Rightarrow$

- Polynomial-time interior-point methods for general LOP over cones. Construct a self-concordant barrier function in the interior of the feasible region — theoretically powerful but difficult in practice.
- Polynomial-time primal-dual interior-point methods for LOPs over symmetric cones — theoretically and practically powerful.
- LOPs over symmetric cones unifies LPs, SDPs and SOCPs.

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP (Second Order Cone Program)
5. Some applications of SOCPs

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Definition. $K \subset V$ is a symmetric cone if

- $K^* = K$ (self-dual).
- For $\forall\, \boldsymbol{x}, \ \boldsymbol{y}$ of int K, there is a linear transformation $T : V \to V$ such that $T(K) = K$ and $T(\boldsymbol{x}) = \boldsymbol{y}$ (homogeneous).

Symmetric cones are classified into the following cones

(a) the second order cone

$$\mathbb{Q}(n) \equiv \quad \{\boldsymbol{x} = (x_0, \boldsymbol{x}_1) : x_0 \in \mathbb{R}, \boldsymbol{x}_1 \in \mathbb{R}^n, x_0 \geq \|\boldsymbol{x}_1\|\},$$

where $\|\boldsymbol{x}_1\| = \sqrt{\boldsymbol{x}_1^T \boldsymbol{x}_1}$.

(b) the set $\mathbb{S}^n_+$ of $n \times n$ real, symmetric positive semidefinite matrices ($\supset$ the set of nonnegative numbers when $n = 1$).

(c) the set of $n \times n$ Hermitian psd mat. w. complex entries.

(d) the set of $n \times n$ Hermitian psd mat. w. quarternion entries.

(e) the set of 3×3 Hermitian psd mat. w. octonion entries.

(f) any cone $K_1 \times K_2$ where K_1 and K_2 are themselves symmetric cones.

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Definition. $K \subset V$ is a symmetric cone if

- $K^* = K$ (self-dual).
- For $\forall \, \boldsymbol{x}, \ \boldsymbol{y}$ of int K, there is a linear transformation $T : V \to V$ such that $T(K) = K$ and $T(\boldsymbol{x}) = \boldsymbol{y}$ (homogeneous).

Symmetric cones are classified into the following cones

(a) the second order cone

$$\mathbb{Q}(n) \equiv \quad \{\boldsymbol{x} = (x_0, \boldsymbol{x}_1) : x_0 \in \mathbb{R}, \boldsymbol{x}_1 \in \mathbb{R}^n, x_0 \geq \|\boldsymbol{x}_1\|\},$$

where $\|\boldsymbol{x}_1\| = \sqrt{\boldsymbol{x}_1^T \boldsymbol{x}_1}$.

(b) the set $\mathbb{S}^n_+$ of $n \times n$ real, symmetric positive semidefinite matrices ($\supset$ the set of nonnegative numbers when $n = 1$).

(f) any cone $K_1 \times K_2$ where K_1 and K_2 are themselves symmetric cones.

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP (Second Order Cone Program)
5. Some applications of SOCPs

(P) (LOP over a cone K): $\min \langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Theorem. A cone $K \subset V$ is symmetric iff it is the cone of squares of some Euclidean Jordan algebra in V (Jordan algebra characterization of symmetric cones); $K = \{\boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in V\}$.

Definition. $(V, \circ)$ is a Euclidean Jordan algebra if $(\boldsymbol{x}, \boldsymbol{y}) \in V \times V \rightarrow \boldsymbol{x} \circ \boldsymbol{y} \in V$ is a bilinear map satisfying

- $\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{x}$
- $\boldsymbol{x} \circ (\boldsymbol{x}^2 \circ \boldsymbol{y}) = \boldsymbol{x}^2 \circ (\boldsymbol{x} \circ \boldsymbol{y})$ where $\boldsymbol{x}^2 = \boldsymbol{x} \circ \boldsymbol{x}$
- $\langle \boldsymbol{x} \circ \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \circ \boldsymbol{z} \rangle$ for $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$

(a) the second order cone

$$\mathbb{Q}(n) \equiv \left\{ \boldsymbol{x} = (x_0, \boldsymbol{x}_1) \in \mathbb{R}^{1+n} : x_0 \geq \|\boldsymbol{x}_1\| \right\}:$$

$$\boldsymbol{x} \circ \boldsymbol{y} \equiv \left(x_0 y_0 + \boldsymbol{x}_1^T \boldsymbol{y}_1, x_0 \boldsymbol{y}_1 + y_0 \boldsymbol{x}_1 \right)$$

$$\Rightarrow \mathbb{Q}(n) = \left\{ \boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^{1+n} \right\}.$$

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Theorem. A cone $K \subset V$ is symmetric iff it is the cone of squares of some Euclidean Jordan algebra in V (Jordan algebra characterization of symmetric cones); $K = \{\boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in V\}$.

Definition. $(V, \circ)$ is a Euclidean Jordan algebra if $(\boldsymbol{x}, \boldsymbol{y}) \in V \times V \to \boldsymbol{x} \circ \boldsymbol{y} \in V$ is a bilinear map satisfying

- $\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{x}$
- $\boldsymbol{x} \circ (\boldsymbol{x}^2 \circ \boldsymbol{y}) = \boldsymbol{x}^2 \circ (\boldsymbol{x} \circ \boldsymbol{y})$ where $\boldsymbol{x}^2 = \boldsymbol{x} \circ \boldsymbol{x}$
- $\langle \boldsymbol{x} \circ \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \circ \boldsymbol{z} \rangle$ for $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$

(b) the set $\mathbb{S}^n_+$ of $n \times n$ real, symmetric positive semidefinite matrices

$$\boldsymbol{X} \circ \boldsymbol{Y} \equiv \frac{1}{2}(\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X}) \Rightarrow \mathbb{S}^n_+ = \{\boldsymbol{X} \circ \boldsymbol{X} = \boldsymbol{X}^2 : \boldsymbol{X} \in \mathbb{S}^n\}.$$

(P) (LOP over a cone K): min $\langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Theorem. A cone $K \subset V$ is symmetric iff it is the cone of squares of some Euclidean Jordan algebra in V (Jordan algebra characterization of symmetric cones); $K = \{\boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in V\}$.

Definition. $(V, \circ)$ is a Euclidean Jordan algebra if $(\boldsymbol{x}, \boldsymbol{y}) \in V \times V \to \boldsymbol{x} \circ \boldsymbol{y} \in V$ is a bilinear map satisfying

- $\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{x}$
- $\boldsymbol{x} \circ (\boldsymbol{x}^2 \circ \boldsymbol{y}) = \boldsymbol{x}^2 \circ (\boldsymbol{x} \circ \boldsymbol{y})$ where $\boldsymbol{x}^2 = \boldsymbol{x} \circ \boldsymbol{x}$
- $\langle \boldsymbol{x} \circ \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \circ \boldsymbol{z} \rangle$ for $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$

(b)' the nonnegative orthant $\mathbb{R}^n_+ = \prod_{i=1}^n \mathbb{S}^1_+$:

$$\boldsymbol{x} \circ \boldsymbol{y} = (x_1 y_1, \ldots, x_n y_n)$$

$$\Rightarrow \mathbb{R}^n_+ = \{x \circ x = (x_1^2, \ldots, x_n^2) : x \in \mathbb{R}^n\}.$$

(P) (LOP over a cone K): $\quad \min \langle \boldsymbol{a}_0, \boldsymbol{x} \rangle$
s.t. $\langle \boldsymbol{a}_p, \boldsymbol{x} \rangle = b_p \ (1 \leq p \leq m), \ \boldsymbol{x} \in K$

Theorem. A cone $K \subset V$ is symmetric iff it is the cone of squares of some Euclidean Jordan algebra in V (Jordan algebra characterization of symmetric cones); $K = \{\boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in V\}$.

Definition. $(V, \circ)$ is a Euclidean Jordan algebra if $(\boldsymbol{x}, \boldsymbol{y}) \in V \times V \to \boldsymbol{x} \circ \boldsymbol{y} \in V$ is a bilinear map satisfying

- $\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{x}$
- $\boldsymbol{x} \circ (\boldsymbol{x}^2 \circ \boldsymbol{y}) = \boldsymbol{x}^2 \circ (\boldsymbol{x} \circ \boldsymbol{y})$ where $\boldsymbol{x}^2 = \boldsymbol{x} \circ \boldsymbol{x}$
- $\langle \boldsymbol{x} \circ \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \circ \boldsymbol{z} \rangle$ for $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$

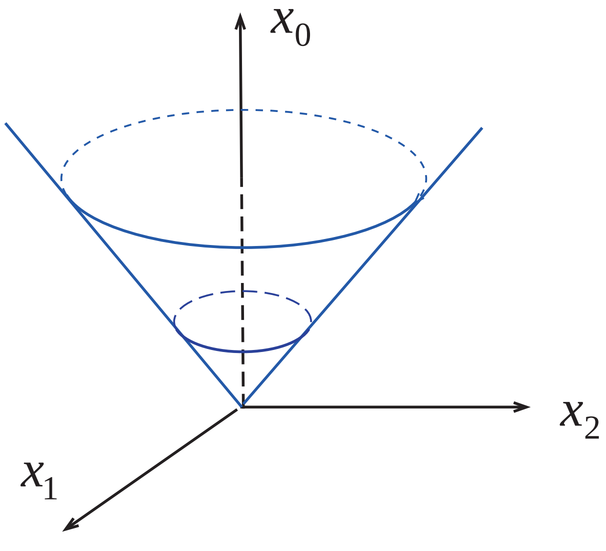
- We can extend and unify pdipm developed for LPs and SDPs using Euclidean Jordan algebra.
- We can define the central trajectory as the sol. of $\boldsymbol{x} \circ \boldsymbol{s} = \mu \boldsymbol{e}$, $\mu > 0$, where $\boldsymbol{e}$ denotes the identity element ; $\boldsymbol{e} \circ \boldsymbol{x} = \boldsymbol{x} \circ \boldsymbol{e}$ for $\forall \boldsymbol{x} \in V$.
- We can define $\det \boldsymbol{x}$ and the logarithmic barrier function $-\log \det \boldsymbol{x}$.

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP
5. Some applications of SOCPs

Second order cone with dimension $1+n$:

$$\mathbb{Q}(n) = \{x = (x_0, \boldsymbol{x}_1) \in \mathbb{R}^{1+n} : x_0 \geq \|\boldsymbol{x}_1\|\}$$



The second order cone ($n = 2$):

$$\mathbb{Q}(2) =$$
$$\left\{(x_0, x_1, x_2) \in \mathbb{R}^{1+2} : x_0 \geq \sqrt{x_1^2 + x_2^2}\right\}$$

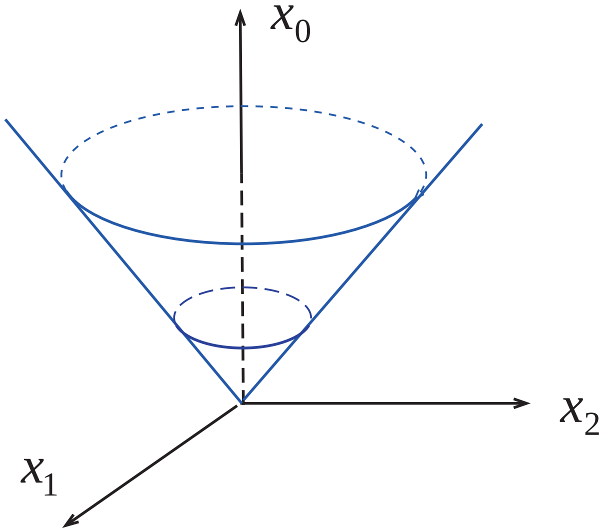
A primal-dual pair of SOCPs:

(P) min $\sum_{p=1}^{k} \boldsymbol{c}_p^T \boldsymbol{x}_p$ s.t. $\sum_{p=1}^{k} \boldsymbol{A}_p \boldsymbol{x}_p = b, \quad \boldsymbol{x}_p \in \mathbb{Q}(n_p)\ (\forall p).$

(D) max $\boldsymbol{b}^T \boldsymbol{y}$ s.t. $\boldsymbol{A}_p^T \boldsymbol{y} + \boldsymbol{s}_p = \boldsymbol{c}_p, \quad \boldsymbol{s}_p \in \mathbb{Q}(n_p)\ (\forall p).$

$\boldsymbol{A}_p$: a data matrix, $\boldsymbol{b}$: a data vector, $\boldsymbol{c}_p$: a data vector,
$\boldsymbol{x}_p$: a variable vector, $\boldsymbol{s}_p$: a variable vector,
$\boldsymbol{y}$: a variable vector.

- the boundary of $\mathbb{Q}(n_p)$
  $= \{\boldsymbol{x}_p = (x_{p0}, \boldsymbol{x}_{p1}) \in \mathbb{R}^{1+n_p}, x_{p0} = \|\boldsymbol{x}_{p1}\|\}$.
- the interior of $\mathbb{Q}(n_p)$
  $= \{\boldsymbol{x}_p = (x_{p0}, x_{p1}) \in \mathbb{R}^{1+n_p}, x_{p0} > \|\boldsymbol{x}_{p1}\|\}$;



The second order cone ($n = 2$):

$$\mathbb{Q}(2) = \left\{(x_0, x_1, x_2) \in \mathbb{R}^{1+2} : x_0 \geq \sqrt{x_1^2 + x_2^2}\right\}$$

A primal-dual pair of SOCPs:

(P) min $\sum_{p=1}^{k} \boldsymbol{c}_p^T \boldsymbol{x}_p$ s.t. $\sum_{p=1}^{k} \boldsymbol{A}_p \boldsymbol{x}_p = b, \quad \boldsymbol{x}_p \in \mathbb{Q}(n_p) \ (\forall p).$

(D) max $\boldsymbol{b}^T \boldsymbol{y}$ s.t. $\boldsymbol{A}_p^T \boldsymbol{y} + \boldsymbol{s}_p = \boldsymbol{c}_p, \quad \boldsymbol{s}_p \in \mathbb{Q}(n_p) \ (\forall p).$

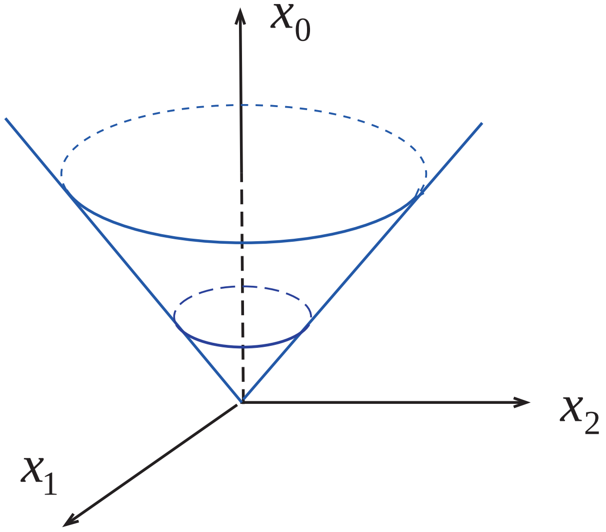
$\boldsymbol{A}_p$: a data matrix, $\boldsymbol{b}$: a data vector, $\boldsymbol{c}_p$: a data vector,
$\boldsymbol{x}_p$: a variable vector, $\boldsymbol{s}_p$: a variable vector,
$\boldsymbol{y}$: a variable vector.

- If $\boldsymbol{x}_p \in \mathbb{Q}(n_p)$ and $\boldsymbol{s}_p \in \mathbb{Q}(n_p)$ then

$$0 = \boldsymbol{x}_p^T \boldsymbol{s}_p \ \Leftrightarrow \ 0 = \boldsymbol{x}_p \circ \boldsymbol{s}_p \equiv (\boldsymbol{x}_p^T \boldsymbol{s}_p, x_{p0} \boldsymbol{s}_{p1} + s_{p0} \boldsymbol{x}_{p1}).$$

- the logarithmic barrier function:

$$\sum_{p=1}^{k} \boldsymbol{c}_p^T \boldsymbol{x}_p - \mu \sum_{p=1}^{k} \log(x_{p0} - \|\boldsymbol{x}_{p1}\|).$$



The second order cone ($n = 2$):

$\mathbb{Q}(2) =$
$\left\{ (x_0, x_1, x_2) \in \mathbb{R}^{1+2} : x_0 \geq \sqrt{x_1^2 + x_2^2} \right\}$

A primal-dual pair of SOCPs:

(P) min $\sum_{p=1}^{k} \boldsymbol{c}_p^T \boldsymbol{x}_p$ s.t. $\sum_{p=1}^{k} \boldsymbol{A}_p \boldsymbol{x}_p = b, \quad \boldsymbol{x}_p \in \mathbb{Q}(n_p) \ (\forall p).$

(D) max $\boldsymbol{b}^T \boldsymbol{y}$ s.t. $\boldsymbol{A}_p^T \boldsymbol{y} + \boldsymbol{s}_p = \boldsymbol{c}_p, \quad \boldsymbol{s}_p \in \mathbb{Q}(n_p) \ (\forall p).$

$\boldsymbol{A}_p$: a data matrix, $\boldsymbol{b}$: a data vector, $\boldsymbol{c}_p$: a data vector,
$\boldsymbol{x}_p$: a variable vector, $\boldsymbol{s}_p$: a variable vector,
$\boldsymbol{y}$: a variable vector.

Weak duality: $0 \leq \sum_{p=1}^{k} \boldsymbol{c}_p \boldsymbol{x}_p - \boldsymbol{b}^T \boldsymbol{y}$, $\forall$ feas. $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$.

Strong duality: If $\exists$ int. feas. $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$, then

$$0 = \sum_{p=1}^{k} \boldsymbol{c}_p \bar{\boldsymbol{x}}_p - \boldsymbol{b}^T \bar{\boldsymbol{y}}, \ \forall \text{ opt. } (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{s}}).$$

KKT optimality condition:

$$\sum_{p=1}^{k} \boldsymbol{A}_p \boldsymbol{x}_p = \boldsymbol{b}, \ \boldsymbol{x}_p \in \mathbb{Q}(n_p) \ (1 \leq p \leq k),$$
$$\boldsymbol{A}_p^T \boldsymbol{y} + \boldsymbol{s}_p = \boldsymbol{c}_p, \ \boldsymbol{s}_p \in \mathbb{Q}(n_p) \ (1 \leq p \leq k),$$
$$\boldsymbol{x}_p^T \boldsymbol{s}_p = 0 \text{ (or } \boldsymbol{x}_p \circ \boldsymbol{s}_p = \mathbf{0}\text{)} \ (1 \leq p \leq k).$$

Appendix.
Linear Optimization Problems over Symmetric Cones

1. Linear optimization problems over cones
2. Symmetric cones
3. Euclidean Jordan algebra
4. The equality standard form SOCP (Second Order Cone Program)
5. Some applications of SOCPs

Basic fact

Let $\boldsymbol{w} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$.

$$\begin{pmatrix} \boldsymbol{I}\alpha & \boldsymbol{w} \\ \boldsymbol{w}^T & \beta \end{pmatrix} \succeq \boldsymbol{O} \text{ (an SDP inequality).}$$

$\Updownarrow$ a special case of the Schur complement

$$\boldsymbol{w}^T\boldsymbol{w} \leq \alpha\beta,\ \alpha \geq 0 \text{ and } \beta \geq 0.$$

$$\Updownarrow$$

$$\left\| \begin{pmatrix} \alpha - \beta \\ 2\boldsymbol{w} \end{pmatrix} \right\| \leq \alpha + \beta.$$

$$\Updownarrow$$

$$(x_0, x_1, \ldots, x_{1+n}) \in \mathbb{Q}(1+n),$$
$$x_0 = \alpha + \beta,\ x_1 = \alpha - \beta, x_2 = 2w_1, \ldots, x_{2+n} = 2w_n$$

(an SOCP inequality).

A quasi-convex optimization problem.

$$\text{minimize } \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}} \text{ subject to } \boldsymbol{Ax} \geq \boldsymbol{b}.$$

Here we assume $\boldsymbol{d}^T\boldsymbol{x} > 0$ for $\forall$ feasible $\boldsymbol{x} \in \mathbb{R}^n$.

$$\Updownarrow$$

$$\text{minimize } \zeta \text{ subject to } \zeta \geq \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}},\ \boldsymbol{Ax} \geq \boldsymbol{b}.$$

$$\Updownarrow\ \zeta - \frac{(\boldsymbol{Lx}-\boldsymbol{c})^T(\boldsymbol{Lx}-\boldsymbol{c})}{\boldsymbol{d}^T\boldsymbol{x}} \geq 0 \Leftrightarrow \begin{pmatrix} (\boldsymbol{d}^T\boldsymbol{x})\boldsymbol{I} & \boldsymbol{Lx}-\boldsymbol{c} \\ (\boldsymbol{Lx}-\boldsymbol{c})^T & \zeta \end{pmatrix} \succeq \boldsymbol{O}.$$

$$\text{SDP: min } \zeta \text{ s.t. } \begin{pmatrix} \boldsymbol{d}^T\boldsymbol{x}\boldsymbol{I} & \boldsymbol{Lx}-\boldsymbol{c} \\ (\boldsymbol{Lx}-\boldsymbol{c})^T & \zeta\boldsymbol{x} \end{pmatrix} \succeq \boldsymbol{O},\ \boldsymbol{Ax} \geq \boldsymbol{b}.$$

$\Updownarrow$ Basic fact with $\boldsymbol{w} = \boldsymbol{Lx} - \boldsymbol{c}$, $\alpha = d^Tx$ and $\beta = \zeta$

$$\text{SOCP: min } \zeta \text{ s.t. } \left\| \begin{pmatrix} \boldsymbol{d}^T\boldsymbol{x} - \zeta \\ 2(\boldsymbol{Lx}-\boldsymbol{c}) \end{pmatrix} \right\| \leq \boldsymbol{d}^T\boldsymbol{x} + \zeta,\ \boldsymbol{Ax} \geq \boldsymbol{b}.$$

A convex quadratic optimization problem.

$$\min \quad \boldsymbol{x}^T \boldsymbol{Q}_0 \boldsymbol{x} + \boldsymbol{c}_0^T \boldsymbol{x} \qquad \text{s.t. } \boldsymbol{x}^T \boldsymbol{Q}_i \boldsymbol{x} + \boldsymbol{c}_i^T \boldsymbol{x} + \gamma_i \leq 0 \ (\forall i).$$

Here $\boldsymbol{Q}_i \in \mathbb{S}^n_+ \ (0 \leq i \leq m)$: data matrices, $\boldsymbol{c}_i \in \mathbb{R}^n \ (1 \leq i \leq m)$: data vectors and $\gamma_i \in \mathbb{R} \ (1 \leq i \leq m)$.

$\Downarrow$ factorize $\boldsymbol{Q}_i$ such that $\boldsymbol{Q}_i = \boldsymbol{L}_i^T \boldsymbol{L}_i,$ a new variable ζ

$$\min \quad \zeta \quad \text{s.t.} \quad \begin{array}{l} (\boldsymbol{L}_0 \boldsymbol{x})^T (\boldsymbol{L}_0 \boldsymbol{x}) + \boldsymbol{c}_0^T \boldsymbol{x} \leq \zeta, \\ (\boldsymbol{L}_i \boldsymbol{x})^T (\boldsymbol{L}_i \boldsymbol{x}) + \boldsymbol{c}_i^T \boldsymbol{x} + \gamma_i \leq 0 \ (1 \leq i \leq m). \end{array}$$

$\Downarrow$

$$\min \quad \zeta \quad \text{s.t.} \quad \begin{array}{l} (\boldsymbol{L}_0 \boldsymbol{x})^T (\boldsymbol{L}_0 \boldsymbol{x}) \leq (\zeta - \boldsymbol{c}_0^T \boldsymbol{x}), \\ (\boldsymbol{L}_i \boldsymbol{x})^T (\boldsymbol{L}_i \boldsymbol{x}) \leq (-\gamma_i - \boldsymbol{c}_i^T \boldsymbol{x}) \ (1 \leq i \leq m). \end{array}$$

Basic fact:

$\Downarrow$ $\boldsymbol{w}^T \boldsymbol{w} \leq \alpha\beta, \ \alpha \geq 0$ and $\beta \geq 0 \Leftrightarrow \left\| \begin{pmatrix} \alpha - \beta \\ 2\boldsymbol{w} \end{pmatrix} \right\| \leq \alpha + \beta.$

with $\boldsymbol{w} = \boldsymbol{L}_i \boldsymbol{x}, \ \alpha = (t - \boldsymbol{c}_0^T \boldsymbol{x})$ or $(-\gamma_i - \boldsymbol{c}_i^T \boldsymbol{x}), \ \beta = 1$

$$\min \quad \zeta \quad \text{s.t.} \quad \left\| \begin{pmatrix} \zeta - \boldsymbol{c}_0^T \boldsymbol{x} - 1 \\ 2\boldsymbol{L}_0 \boldsymbol{x} \end{pmatrix} \right\| \le \zeta - \boldsymbol{c}_0^T \boldsymbol{x} + 1$$

$$\left\| \begin{pmatrix} -\gamma_i - \boldsymbol{c}_i^T \boldsymbol{x} - 1 \\ 2\boldsymbol{L}_i \boldsymbol{x} \end{pmatrix} \right\| \le -\gamma_i - \boldsymbol{c}_i^T \boldsymbol{x} + 1$$

$$(1 \le i \le m).$$

Minimization of the sum of Euclidean norms.

$$\min \quad \sum_{i=1}^{m} c_i \|\boldsymbol{u}_i\| \quad \text{s.t. } \boldsymbol{u}_i = \boldsymbol{A}_i^T \boldsymbol{y} + \boldsymbol{b}_i \; 1 \leq i \leq m).$$

Here $\boldsymbol{A}_i \ (1 \leq i \leq m)$: data matrices, $\boldsymbol{b}_i \ (1 \leq i \leq m)$: data vectors and $0 < c_i \in \mathbb{R} \ (1 \leq i \leq m)$: data.

$\Downarrow$ new variables $t_i \ (1 \leq i \leq m)$

$$\min \quad \sum_{i=1}^{m} c_i t_i \quad \text{s.t. } \boldsymbol{u}_i = \boldsymbol{A}_i^T \boldsymbol{y} + \boldsymbol{b}_i, \ \|\boldsymbol{u}_i\| \leq t_i \ (1 \leq i \leq m).$$

A robust LP: min $\boldsymbol{c}^T\boldsymbol{x}$ subject to $\boldsymbol{x} \in F$, where

$$F = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i^T\boldsymbol{x} - b_i \geq 0 \text{ for } \forall \boldsymbol{a}_i \in U_i \ (1 \leq i \leq m)\}.$$

Here $U_i \equiv \{\tilde{\boldsymbol{a}}_i + \boldsymbol{P}_i\boldsymbol{u}_i : \|\boldsymbol{u}_i\| \leq 1\}$ denotes an ellipsoidal uncertain set for some $\tilde{\boldsymbol{a}}_i$ and some nonsingular matrix $\boldsymbol{P}_i$.

$$\boldsymbol{a}_i^T\boldsymbol{x} - b_i \geq 0 \text{ for } \forall \boldsymbol{a}_i \in U_i$$

$$\Updownarrow$$

$$\tilde{\boldsymbol{a}}_i{}^T\boldsymbol{x} + (\boldsymbol{P}_i\boldsymbol{u}_i)^T\boldsymbol{x} - b_i \geq 0 \text{ for } \forall \boldsymbol{u}_i \text{ with } \|\boldsymbol{u}_i\| \leq 1.$$

$$\Updownarrow$$

$$\min\{\tilde{\boldsymbol{a}}_i{}^T\boldsymbol{x} + \boldsymbol{u}_i{}^T(\boldsymbol{P}_i{}^T\boldsymbol{x}) - b_i : \|\boldsymbol{u}_i\| \leq 1\} \geq 0.$$

$$\Updownarrow \ \min. \text{ attains at } \boldsymbol{u}_i = -(\boldsymbol{P}_i^T\boldsymbol{x})/\|\boldsymbol{P}_i^T\boldsymbol{x}\|$$

$$\tilde{\boldsymbol{a}}_i^T\boldsymbol{x} - \|\boldsymbol{P}_i^T\boldsymbol{x}\| - b_i \geq 0 \ \text{ or } \|\boldsymbol{P}_i^T\boldsymbol{x}\| \leq \tilde{\boldsymbol{a}}_i^T\boldsymbol{x} - b_i.$$

Hence the robust LP $\Rightarrow$

SOCP $\min \ \boldsymbol{c}^T\boldsymbol{x}$ subject to $\|\boldsymbol{P}_i^T\boldsymbol{x}\| \leq \tilde{\boldsymbol{a}}_i^T\boldsymbol{x} - b_i \ (1 \leq i \leq m)$.

References

References

[1] F. Alizadeh, J.-P.A. Haeberly and M.L. Overton, Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results, *SIAM J. on Optim,*, 8 (1998) 746-768.

[2] S. J. Benson, Parallel Computing on Semidefinite Programs, Preprint ANL/MCS-P939-0302, http://www.msc.anl.gov/$\sim$benson/dsdp/pdsdp.ps (2002).

[3] S. J. Benson, Y. Ye and X. Zhang, Solving large-scale sparse semidefinite programs for combinatorial optimization, *SIAM Journal on Optimization*, 10 (2000) 443-461.

[4] A. Ben-Tal and A. Nemirovski, Robust convex optimization, *Mathematics of Operations Research*, 23 (1998) 769-805.

[5] D. Bertsimas and J. Sethuraman, Moment problems and semidefinite programming, in *Semidefinite programming*, H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., 469–509, 2000.

[6] S. Boyd, L. El Ghaui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

[7] B. Borchers, CSDP 2.3 user's guide, *Optimization Methods and Software*, 11 & 12 (1999) 597–611. Available at http://www.nmt.edu/˜borchers/csdp.html.

[8] B. Borchers and J. Young, Implementation of a primal-dual method for SDP on a parallel architecture, Dept. of Mathematics, New Mexico Tech, Socorro, NM 87801, September 2005.

[9] S. Burer and R.D.C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Mathematical Programming, 95 (2003) 329-357.

[10] L. Faybusovich, Linear systems in Jordan algebra and primal-dual interior-point algorithms, *Journal of Computational and Applied Mathematics*, 86 (1997) 149-75.

[11] K. Fujisawa, M. Kojima and K. Nakata, Exploiting sparsity in primal-dual interior-point methods for semidefinite programming, *Mathematical Programming*, 79 (1997) 235–253.

[12] M. Fukuda, M. Kojima, K. Murota and K. Nakata, Exploiting sparsity in semidefinite programming via matrix completion I: General framework, *SIAM Journal on Optimization*, 11 (2000) 647–674.

[13] M. Fukuda, B.J. Braams, M. Nakata, M.L. Overton, J.K. Percus,M. Yamashita and Z. Zhao, Large-scale semidefinite programs in electronic structure calculation, Research Report B-413, Department of Mathematical and Computing Sciences, Tokyo Inst. of Tech., Meguro, Tokyo 152-8552, Feb. 2005.

[14] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, tJournal of the ACM, 42 (1995) 1115-1145.

[15] C. Helmberg, Semidefinite programming home page, http://www-user.tu-chemnitz.de/∼helmberg/semidef.html

[16] C. Helmberg and F. Rendl, A spectral bundle method for semidefinite programming, *SIAM Journal on Optimization*, 10 (2000) 673-696.

[17] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, An interior-point method for semidefinite programming, *SIAM Journal on Optimization*, 6 (1996) 342–361.

[18] E. de. Klerk, T. Terlaky and K. Roos, Self-dual embeddings, in H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications*, *Kluwer Academic Publishers*, Massachusetts (2000).

[19] M. Kockvara, http://www2.am.uni-erlangen.de/∼kocvara/pennon/

[20] M. Kojima, S. Shindoh and S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices, *SIAM Journal on Optimization*, 7 (1997) 86–125.

[21] M. Kojima, M. Shida and S. Shindoh, Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs, *Mathematical Programming*, 80 (1998) 129-161.

[22] J. B. Lasserre, Global optimization with polynomials and the problems of moments, *SIAM Journal on Optimization*, 11 (2001) 796–817.

[23] J. B. Lasserre and T. Prieto-Rumeau, SDP vs. LP Relaxations for the moment approach in some performance evaluation problems, *Stochastic Models*, 20 (2004) 439-456.

[24] L. Lovasz and A. Schrijver, Cones of matrices and set-functions and 0-1 optimization, *SIAM Journal on Optimization* 1 (1991) 166?190.

[25] Z. Lu, A. Nemirovski and R.D.C. Monteiro. Large-Scale Semidefinite Programming via Saddle Point Mirror-Prox Algorithm, Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA, November, 2004.

[26] Z.-Q. Luo, J. F. Sturm and S. Zhang, Conic convex programming and self-dual embedding, *Optimization Methods and Software*, 14 (2000) 169-218.

[27] R. D. C. Monteiro, Primal-dual path-following algorithms for semidefinite programming, *SIAM Journal on Optimization*, 7 (1997) 663–678.

[28] R. D. C. Monteiro and Y. Zhang, A unified analysis for a class of a new family of primal-dual interior-point algorithms for semidefinite programming, *Mathematical Programming*, 81 (1998) 281-299.

[29] R.D.C. Monteiro, First- and second-order methods for semidefinite programming, *Mathematical Programming*, 97 (2003) 209-244.

[30] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima and K. Murota, Exploiting sparsity in semidefinite programming via matrix completion II: Implementation and numerical results, *Mathematical Programming*, 95 (2003) 303–327.

[31] K. Nakata, M. Yamashita, K. Fujisawa and M. Kojima, A parallel primal-dual interior-point method for semidefinite programs using positive definite matrix completion, Research Report B-398, Dept. of Mathematical and Computing Sciences, Tokyo Inst. Tech., Meguro, Tokyo 152-8552 (2003).

[32] M. V. Nayakkankuppam. Solving Large-Scale Semidefinite Programs in Parallel, Technical Report, Department of Mathematics and Statistics, University of Maryland, Baltimore County, March 2005.

[33] Yu. E. Nesterov and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Publication, Philadelphia, 1994.

[34] Yu. E. Nesterov and M. J. Todd, Primal-dual interior-point methods for self-scaled cones, *SIAM Journal on Optimization*, 8 (1998) 324-364.

[35] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems', *Mathematical Programming*, 96 (2003) 293-320.

[36] S. Schmieta and F. Alizadeh, Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones, *Mathematics of Operations Research*, 26 (2001) 543-564.

[37] J. F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optimization Methods and Software*, 11 & 12 (1999) 625–653.

[38] M. J. Todd, A study of search directions in primal-dual interior-point methods for semidefinite programming, *Optimization Methods and Software*, 11 & 12 (199) 1-46.

[39] M. J. Todd, Semidefinite optimization, *Acta Numerical* 10 (2001) 515-560.

[40] K. C. Toh, Solving large scale semidefinite programs via an iterative solver on the augmented systems, *SIAM Journal on Optimization*, 14 (2004) 670–698.

[41] K.C. Toh and M. Kojima, Solving some large scale semidefinite programs via the conjugate residual method, *SIAM Journal on Optimization*, 12 (2002) 669–691.

[42] K. C. Toh, M. J. Todd and R. H. Tütüncü, SDPT3 — a MATLAB software package for semidefinite programming, version 1.3, *Optimization Methods and Software*, 11 & 12 (1999) 545–581. Available at http://www.math.nus.edu.sg/˜mattohkc.

[43] L. Tuncel, Potential reduction and primal-dual methods, in H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications*, *Kluwer Academic Publishers*, Massachusetts (2000).

[44] H. Waki, S. Kim, M. Kojima and M. Muramatsu, Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity, Research Report B-411, Dept. of Mathematical and Computing Sciences, Tokyo Inst. Tech., Meguro, Tokyo 152-8552 (2004).

[45] L. Vandenberghe and S. Boyd, Semidefinite Programming, *SIAM Review* 38 (1996) 49–95.

[46] H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications*, *Kluwer Academic Publishers*, Massachusetts (2000).

[47] H. Wolkowicz, **http://liinwww.ira.uka.de/bibliography/Math/psd.html**

[48] M. Yamashita, K. Fujisawa and M. Kojima, Implementation and Evaluation of SDPA6.0 (SemiDefinite Programming Algorithm 6.0), *Optimization Methods and Software*, 18 (2003) 491–505.

[49] M. Yamashita, K. Fujisawa and M. Kojima, SDPARA: SemiDefinte Programming Algorithm Parallel version, *Parallel Computing*, 29 (2003) 1053–1067.

[50] Y. Ye and K. Anstreicher, On quadratic an d$O(nL)$ convergence of a predictor-corrector algorithm for LCP, *Mathematical Programming*, 62 (1993) 537-551.

[51] Z. Zhao, B.J. Braams, M. Fukuda, M.L. Overton and J.K. Percus, The reduced density matrix method for electronic structure calculations and the role of three-index representability, *The Journal of Chemical Physics*, 120 (2004) 2095–2104.

[52] A. Ben-Tal, M. Kocava, A. Nemirovski, and J. Zowe, Free material optimization via semidefinite programming: the multiload case with contact conditions, *SIAM J. Optim.*, 9 (1999) 813-832.

[53] B. Borchers, SDPLIB (a collection of semidefinite programming test problems), http://infohost.nmt.edu/∼borchers/sdplib.html.