Mathematical Optimization Model/Theory Review I

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

Chapters 1, 2.1-2, 6.1-2, 7.2, 11.3, 11.6

Mathematical Optimization

The class of mathematical optimization/programming problems considered in this course can all be expressed in the form

$$\text{(P)} \qquad \text{minimize} \qquad f(\mathbf{x})$$

$$\text{subject to} \qquad \mathbf{x} \in \mathcal{X}$$

NGO

where $\mathcal{X}$ usually specified by constraints:

$$c_i(\mathbf{x}) \quad (\leq, =, \geq) \; 0 \quad i = 1, ..., m$$

If the constraint functions are linear/affine type, then $\mathcal{X}$ is a convex polyhedral set/region and the problem becomes classical LP.

Structured/Deciplined Convex Optimization: Conic Linear Programming (CLP)

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$

$\sum_i c_i x_i$ | $\sum_i \sum_j c_{ij} x_{ij}$

$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K,$$

$$(\ \mathcal{A}\mathbf{x} = \mathbf{b}\ ),$$

$= \begin{pmatrix} A_i \cdot X \\ \vdots \end{pmatrix}$ $i = 1 \dots m$

where K is a closed and pointed convex cone.

$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

Linear Programming (LP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = \mathcal{R}^n_+$

Second-Order Cone Programming (SOCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = SOC = \{\mathbf{x} : \ x_1 \geq \|\mathbf{x}_{-1}\|_2\}$.

$x_1 \geq \left\| \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \right\|_2$

Semidefinite Programming (SDP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{S}^n$ and $K = \mathcal{S}^n_+$

p-Order Cone Programming (POCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = POC = \{\mathbf{x} : \ x_1 \geq \|\mathbf{x}_{-1}\|_p\}$.

Here, $\mathbf{x}_{-1}$ is the vector $(x_2; ...; x_n) \in R^{n-1}$.

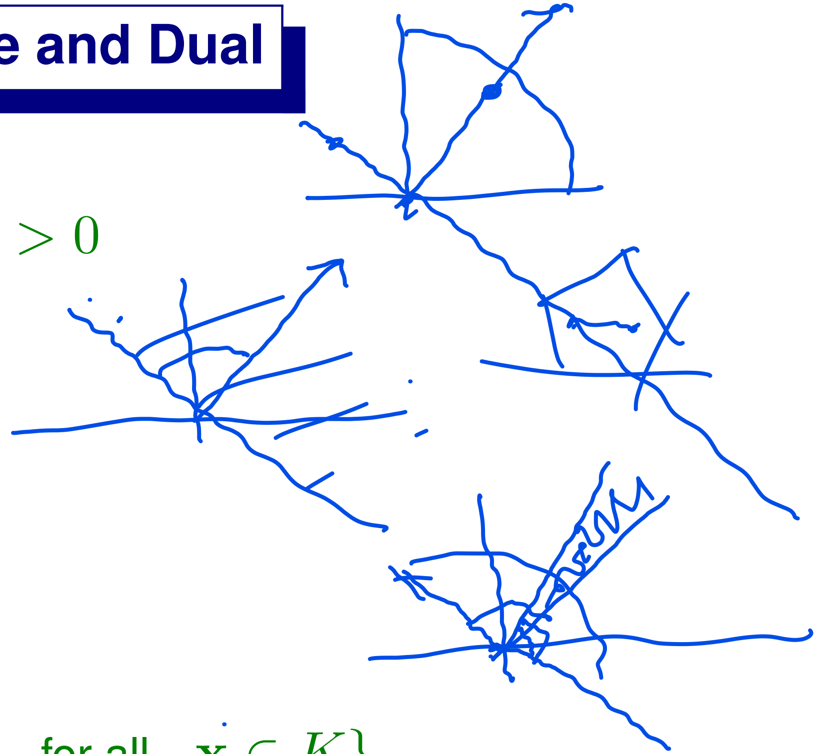
Cone K can be also a product of different cones, that is, $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; ...)$ where $\mathbf{x}_1 \in K_1$, $\mathbf{x}_2 \in K_2$,... and so on with linear constraints:

$$\mathcal{A}_1\mathbf{x}_1 + \mathcal{A}_2\mathbf{x}_2 + ... = \mathbf{b}.$$

Cone, Convex Cone and Dual

- A set K is a cone if $\mathbf{x} \in K$ implies $\alpha\mathbf{x} \in K$ for all $\alpha > 0$
- The intersection of cones is a cone
- A convex cone is a cone and also a convex set
- A pointed cone is a cone that does not contain a line
- Dual of Cone:

$$K^* := \{\mathbf{y} : \mathbf{x} \bullet \mathbf{y} \geq 0 \quad \text{for all} \quad \mathbf{x} \in K\}.$$



Theorem 1 *The dual is always a closed convex cone, and the dual of the dual is the closure of convex hall of K.*

Cone Examples

- Example 1: The n-dimensional non-negative orthant, $\mathcal{R}_+^n = \{\mathbf{x} \in \mathcal{R}^n : \ \mathbf{x} \geq \mathbf{0}\}$, is a convex cone. Its dual is itself.
- Example 2: The set of all PSD matrices in $\mathcal{S}^n$, $\mathcal{S}_+^n$, is a convex cone, called the PSD matrix cone. Its dual is itself.
- Example 3: The set $\{(t; \mathbf{x}) \in \mathcal{R}^{n+1} : \ t \geq \|\mathbf{x}\|_p\}$ for a $p \geq 1$ is a convex cone in $\mathcal{R}^{n+1}$, called the p-order cone. Its dual is the q-order cone with $\frac{1}{p} + \frac{1}{q} = 1$.
- The dual of the second-order cone ($p = 2$) is itself.

LP, SOCP, and SDP Examples

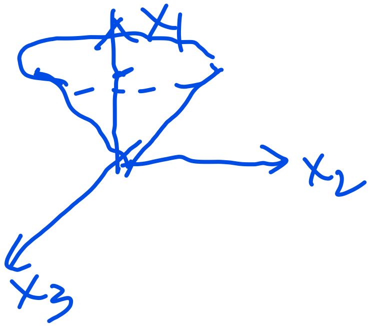
$1+0 :=$

$$(LP) \quad \text{minimize} \quad 2x_1 + x_2 + x_3$$
$$\text{subject to} \quad x_1 + x_2 + x_3 = 1,$$
$$(x_1; x_2; x_3) \geq \mathbf{0}.$$

$x_1 + (x_1 + x_2 + x_3)$

$1 + x_1$

$\sqrt{2}$



$$(SOCP) \quad \text{minimize} \quad 2x_1 + x_2 + x_3$$
$$\text{subject to} \quad x_1 + x_2 + x_3 = 1,$$
$$x_1 - \sqrt{x_2^2 + x_3^2} \geq 0.$$

$\sqrt{2}-1$

$1 + x_1$

$\boxed{x_1 = \sqrt{2} - 1}$

$x_2 + x_3 = 2 - \sqrt{2}$

$x_2 = x_3 = 1 - \frac{\sqrt{2}}{2}$

$2\left(1 - \frac{\sqrt{2}}{2}\right)^2$

$= (\sqrt{2}-1)^2$

$1+0 :=$

$$(SDP) \quad \text{minimize} \quad 2x_1 + x_2 + x_3$$
$$\text{subject to} \quad x_1 + x_2 + x_3 = 1,$$
$$\begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}.$$

$1 + x_1$

(SDP) can be rewriten as

$$
\begin{array}{ll}
\text{minimize} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\
\text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\
& \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0},
\end{array}
$$

that is

$$
\mathbf{c} = \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_1 = \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}.
$$

Dual of Conic LP

m

$$(CLD) \quad \text{maximize} \quad \mathbf{b}^T\mathbf{y} = \sum_{i=1}^{m} b_i y_i$$

$$\text{subject to} \quad \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \; (\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c}), \; \boxed{\mathbf{s} \in K^*,}$$

where $\mathbf{y} \in \mathcal{R}^m$, $\mathbf{s}$ is called the dual slack vector/matrix, and K^* is the dual cone of K.

Here, operator $\mathcal{A}\mathbf{x}$ and Adjoint-Operator $\mathcal{A}^T\mathbf{y}$ minimic matrix-vector production $A\mathbf{x}$ and its transpose operation $A^T\mathbf{y}$, where

$$\mathcal{A} = (\mathbf{a}_1; \mathbf{a}_2; ...; \mathbf{a}_m), \quad \mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}), \quad \text{and} \quad \mathcal{A}^T\mathbf{y} = \sum_i y_i \mathbf{a}_i^T.$$

LP, SOCP, and SDP Examples Again

$c = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$

$a_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = e$

$$
D: \quad \begin{array}{ll} \text{min} & 2x_1 + x_2 + x_3 \\ \text{s. t.} & x_1 + x_2 + x_3 = 1, \; (y) \\ & (x_1; x_2; x_3) \geq \mathbf{0}. \end{array}
\qquad
D = \begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e} \cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & (s_1; s_2; s_3) \geq \mathbf{0}. \end{array}
$$

$2 - y \geq 0$

$1 - y \geq 0$

$1 - y \geq 0$

$\sqrt{2}$

$x_1 = \sqrt{2} - 1$

$$
\begin{array}{ll} \text{min} & 2x_1 + x_2 + x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1, \\ & x_1 - \sqrt{x_2^2 + x_3^2} \geq 0. \end{array}
\qquad
\sqrt{2} \quad \boxed{\begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e} \cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & s_1 - \sqrt{s_2^2 + s_3^2} \geq 0. \end{array}}
$$

$2 - y \geq \sqrt{2(1-y)^2}$

$= y \ldots$

$\geq \sqrt{2}(y-1)$

$y = \sqrt{2}$

For the SOCP case: $2 - y \geq \sqrt{2(1-y)^2}$. Since $y = 1$ is feasible for the dual, $y^* \geq 1$ so that the dual constraint becomes $2 - y \geq \sqrt{2}(y-1)$ or $y \leq \sqrt{2}$. Thus, $y^* = \sqrt{2}$, and there is no duality gap.

$$1 = \quad \begin{array}{ll} \text{minimize} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\ & \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}, \end{array}$$

$$1 = \quad \begin{array}{ll} \text{maximize} & y \\ \text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} y + \mathbf{s} = \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix}, \\ & \mathbf{s} = \begin{pmatrix} s_1 & s_2 \\ s_2 & s_3 \end{pmatrix} \succeq \mathbf{0}. \end{array}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} 2-y & 0.5(1-y) \\ 0.5(1-y) & 1-y \end{pmatrix} \succeq 0$$

$$0 \qquad 0$$

CLP Duality Theorems

Theorem 2 *(Weak duality theorem)* $\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T\mathbf{y} = \mathbf{x} \bullet \mathbf{s} \geq 0$ *for any feasible* $\mathbf{x}$ *of (CLP) and* $(\mathbf{y}, \mathbf{s})$ *of (CLD).*

$\in K$ $\in K^*$

The weak duality theorem shows that a feasible solution to either problem yields a bound on the value of the other problem. We call $\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T\mathbf{y}$ the duality gap.

Corollary 1 *Let* $\mathbf{x}^* \in \mathcal{F}_p$ *and* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$. *Then,* $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$ *implies that* $\mathbf{x}^*$ *is optimal for (CLP) and* $(\mathbf{y}^*, \mathbf{s}^*)$ *is optimal for (CLD).*

Is the reverse also true? That is, given $\mathbf{x}^*$ optimal for (CLP), then there is $(\mathbf{y}^*, \mathbf{s}^*)$ feasible for (CLD) and $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$?

This is called the Strong Duality Theorem.

"True" when $K = \mathcal{R}^n_+$, but not true in general.

Classical LP and LD Cases

Theorem 3 *The following statements hold for every pair of (LP) and (LD) :*

i) *If (LP) and (LD) are both feasible, then both problems have optimal solutions and the optimal objective values of the objective functions are equal, that is, optimal solutions for both (LP) and (LD) exist and there is no duality gap.*

ii) *If (LP) or (LD) is feasible and bounded, then the other is feasible and bounded.*

iii) *If (LP) or (LD) is feasible and unbounded, then the other has no feasible solution.*

iv) *If (LP) or (LD) is infeasible, then the other is either unbounded or has no feasible solution.*

A case that neither (LP) nor (LD) is feasible: $\mathbf{c} = (-1;\ 0), \quad A = (0,\ -1), \quad b = 1.$

$K = \mathbb{R}^n_+$

a_i, c

The LP Primal and Dual Relation

F IF LP ←

Dual \ Primal	F-B	F-UB	IF
F-B	☺	x	x
F-UB	x	x	☹
IF	x	☹	☹

$$
\begin{array}{llllll}
\min & -x_1 & - & x_2 & & \\
\text{s.t.} & x_1 & - & x_2 & = 1 \\
 & -x_1 & + & x_2 & = 1 \\
 & x_1, & & x_2 & \geq 0
\end{array}
$$

$$
\begin{array}{llllll}
\max & y_1 & + & y_2 & & \\
\text{s.t.} & y_1 & - & y_2 & \leq -1 \\
 & -y_1 & + & y_2 & \leq -1
\end{array}
$$

Figure 1: Both primal and dual are infeasible

LP Farkas Lemma and Duality

The Farkas lemma concerns the system the system $\boxed{\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}}$ and its alternative $\{\mathbf{y} : \ -A^T\mathbf{y} \geq \mathbf{0}, \mathbf{b}^T\mathbf{y} > 0\}$ for given data $(A, \mathbf{b})$. This pair can be represented as a primal-dual LP pair

$$0 = \quad \begin{array}{ll} \text{min} & \underbrace{\mathbf{0}^T\mathbf{x}}_{} = 0 \\ \text{s. t.} & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}; \end{array} \qquad \begin{array}{ll} \text{max} & \mathbf{b}^T\mathbf{y} \to +\infty \\ \text{s.t.} & A^T\mathbf{y} + s \leq \mathbf{0}. \end{array}$$

$\to \infty$

$y = 0 \quad b^T y = \alpha$

$\alpha \quad A^T y \leq 0$

$s \geq 0 \quad (\alpha y)$

If the primal is infeasible, then the dual must be feasible and unbounded since it is always feasible.

Geometric Interpretation: Let $A = [\mathbf{a}_1, \ \mathbf{a}_2, ..., \mathbf{a}_n]$, then if $\mathbf{b} \notin \{A\mathbf{x} : \ \mathbf{x} \geq \mathbf{0}\}$, then there must be a vector $\mathbf{y}$ where the angle between $\mathbf{y}$ and $\mathbf{b}$ is strictly acute, and the angle with $\mathbf{a}_j$ is either right or obtuse for all i.

$K = R^n_+$

LP Optimality Conditions and Solution Support

$$\left\{ (\mathbf{x},\mathbf{y},\mathbf{s}) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+) : \begin{array}{rcl} \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} & = & \mathbf{0} \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c} \end{array} \right\}; \quad \text{or}$$

$c^T x - b^T y = x \cdot s$

$0 = \sum_{j=1}^{n} x_j s_j$

complementarity $x_j \cdot s_j = 0$ $\forall j$

$$\begin{array}{rcl} \mathbf{x}.\cdot\mathbf{s} & = & \mathbf{0} \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c}. \end{array}$$

Let $\mathbf{x}^*$ and $\mathbf{s}^*$ be optimal solutions with zero duality gap. Then

$$|\text{supp}(\mathbf{x}^*)| + |\text{supp}(\mathbf{s}^*)| \leq n.$$

$K = R^n_+$

There are $\mathbf{x}^*$ and $\mathbf{s}^*$ such that the sum of support sizes of $\mathbf{x}^*$ and $\mathbf{s}^*$ equal n: called a strict complemetarity pair.

Sometimes we look $\mathbf{x}^*$ such that the support size of $\mathbf{x}^*$ is minimal.

If there is $\mathbf{s}^*$ such that $|\text{supp}(\mathbf{s}^*)| \geq n - d$, then the support size for $\mathbf{x}^*$ is at most d.

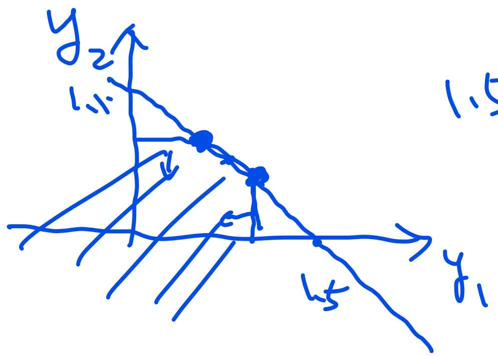
An Example

Consider the primal and dual problem pair:

$$x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad obj = 2$$

$$1.5 \quad \left\{ \begin{array}{llllll} \text{minimize} & x_1 & +x_2 & +1.5 \cdot x_3 & & \\ \text{subject to} & x_1 & & + \quad x_3 & = 1 \\ & & x_2 & + \quad x_3 & = 1 \\ & x_1, & x_2, & \quad x_3 & \geq 0; \end{array} \right.$$

$$x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad obj = 1.5$$



$$1.5 \quad \begin{array}{lllll} \text{maximize} & y_1 & +y_2 & & \\ \text{subject to} & y_1 & & +s_1 & = 1 \\ & & y_2 & +s_2 & = 1 \\ & y_1 & +y_2 & +s_3 & = 1.5 \\ & & & \mathbf{s} & \geq 0; \end{array}$$

$$\left\{ \begin{array}{l} 1 - y_1 \geq 0 \\ 1 - y_2 \geq 0 \\ 1.5 - y_1 - y_2 \geq 0 \end{array} \right.$$

$y_1 + y_2 = 1.5$

$y_1 = .75 = y_2$

$$s = \begin{pmatrix} 2 \\ -.25 \\ 0 \end{pmatrix} \quad \begin{array}{l} y_1 = 1 \\ y_2 = .5 \end{array}$$

where $P^* = \{3\}$ and $Z^* = \{1, 2\}$

Face and Extreme Point

Let P be a polyhedron in $\mathcal{R}^n$, F is a face of P if and only if there is a vector $\mathbf{c}$ for which F is the set of points attaining max $\{\mathbf{c}^T\mathbf{x} : \ \mathbf{x} \in P\}$ provided this maximum is finite.

A polyhedron has only finite many faces; each face is a nonempty polyhedron.

A vector $\mathbf{x} \in P$ is an extreme point or a vertex of P if $\mathbf{x}$ is not a convex combination of more than one distinct points.

Basic Feasible Solution

In the LP standard form, select m linearly independent columns, denoted by the index set B, from A.

$$A_B\mathbf{x}_B = \mathbf{b}$$

for the m-vector $\mathbf{x}_B$. By setting the variables, $\mathbf{x}_N$, of $\mathbf{x}$ corresponding to the remaining columns of A equal to zeros, we obtain a solution $\mathbf{x}$ such that

$$A\mathbf{x} = \mathbf{b}.$$

Then, $\mathbf{x}$ is said to be a (primal) basic solution to (LP) with respect to the basis A_B. The components of $\mathbf{x}_B$ are called basic variables.

If a basic solution $\mathbf{x} \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution.

If one or more components in $\mathbf{x}_B$ has value zero, the basic feasible solution $\mathbf{x}$ is said to be (primal) degenerate.

Dual Basic Feasible Solution

A dual vector $\mathbf{y}$ satisfying

$$A_B^T \mathbf{y} = \mathbf{c}_B$$

is said to be the corresponding dual basic solution with respect to the basis A_B.

If the dual basic solution is also feasible, that is

$$\mathbf{s} = \mathbf{c} - A^T \mathbf{y} \geq \mathbf{0},$$

then primal feasible $\mathbf{x}$ is called an optimal basic solution and A_B an optimal basis.

If one or more components in $\mathbf{s}_N$ has value zero, the basic feasible solution $\mathbf{y}$ is said to be (dual) degenerate.

Theorem 4 *(Carathéodory's theorem) Given (LP) and (LD) where A has full row rank m,*

i) *if there is a feasible solution, there is a basic feasible solution;*

ii) *if there is an optimal solution, there is an optimal basic solution.*

When Strong Duality Theorems Holds for CLP

$K = \mathbb{R}^n_+$

$Ax = b \quad A\tilde{x} = b$

$x \geq 0, \quad \tilde{x} > 0$

Theorem 5 *The following statements hold for every pair of (CLP) and (CLD):*

i) *If (CLP) and (CLD) both are feasible, and furthermore one of them have an interior, then there is no duality gap between (CLP) and (CLD). However, one of the optimal solution may not be attainable.*

ii) *If (CLP) and (CLD) both are feasible and have interior, then, then both have attainable optimal solutions with no duality gap.*

iii) *If (CLP) or (CLD) is feasible and unbounded, then the other has no feasible solution.*

iv) *If (CLP) or (CLD) is infeasible, and furthermore the other is feasible and has an interior, then the other is unbounded.*

In case i), one of the optimal solution may not be attainable although no gap.

Duality Theorem for CLP

Theorem 6 *(CLP duality theorem) If one of (CLP) or (CLD) is unbounded then the other has no feasible solution.*

If (CLP) and (CLD) are both feasible, then both have bounded optimal objective values and the optimal objective values may have a duality gap.

If one of (CLP) or (CLD) has a strictly or interior feasible solution and it has an optimal solution, then the other is feasible and has an optimal solution with the same optimal value.

The CLP Primal and Dual Relation

Dual \ Primal	F-B	F-UB	IF
F-B	☺		☹
F-UB			☹
IF	☹	☹	☹

CLP

$$
\begin{array}{lll}
\min & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \bullet X & \\
\text{s.t.} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet X & = 0 \\
 & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X & = 2 \\
 & X & \succeq \mathbf{0}
\end{array}
$$

$$
\begin{array}{lll}
\max & 2y_2 & \\
\text{s.t.} & y_1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + S & = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
 & S & \succeq \mathbf{0}
\end{array}
$$

The Dual is feasible and bounded, but Primal is infeasible.

Rules to Construct the Dual in General

$$(CLP) \quad \text{minimize} \quad \sum_k \mathbf{c}_k \bullet \mathbf{x}_k$$
$$\text{subject to} \quad \sum_k \mathcal{A}_k \mathbf{x}_k = \mathbf{b},$$
$$\mathbf{x}_k \in K_k, \ \forall k.$$

$$(CLD) \quad \text{minimize} \quad \mathbf{b}^T \mathbf{y}$$
$$\text{subject to} \quad \mathcal{A}_k^T \mathbf{y} + \mathbf{s}_k = \mathbf{c}_k, \ \forall k,$$
$$\mathbf{s}_k \in K_k^*, \ \forall k.$$

obj. coef. vector	right-hand-side
right-hand-side	obj. coef. vector
$\mathcal{A}$	$\mathcal{A}^T$
Max model	Min model
$\mathbf{x}_k \in K$	kth block-constraint slack $\mathbf{s}_k \in K^*$
$\boxed{\mathbf{x}_k \text{ "free"}}$	kth block-constraint slack $\mathbf{s}_k = \mathbf{0}$
ith block-constraint slack $\mathbf{s}_i \in K$	$\mathbf{y}_i \in K^*$
ith block-constraint slack $\mathbf{s}_i = \mathbf{0}$	$\mathbf{y}_i$ "free"

The dual of the dual is primal!

Optimality and Complementarity Conditions for SDP

| supp(x) |

rank(X)

$$\begin{aligned} \mathbf{c} \bullet X - \mathbf{b}^T\mathbf{y} &= 0 \\ \mathcal{A}X &= \mathbf{b} \\ -\mathcal{A}^T\mathbf{y} - S &= -\mathbf{c} \\ X, S &\succeq \mathbf{0} \end{aligned} \qquad (1)$$

, Rk(X) =

$$\begin{aligned} XS &= \mathbf{0} \\ \mathcal{A}X &= \mathbf{b} \\ -\mathcal{A}^T\mathbf{y} - S &= -\mathbf{c} \\ X, S &\succeq \mathbf{0} \end{aligned} \qquad (2)$$

The Rank Theorem of SDP

$$
\begin{array}{lll}
(SDP) & \text{min} & C \bullet X \\
 & \text{subject to} & A_i \bullet X = b_i, i = 1, 2, ..., m,\ X \succeq 0,
\end{array}
$$

where $C,\ A_i \in \mathcal{S}^n$.

Or simply for the SDP Feasibility problem:

$$
\text{Solve} \quad A_i \bullet X = b_i, i = 1, 2, ..., m,\ X \succeq 0,
$$

Solution Rank for SDP

$$\begin{array}{rcl} C \bullet X - \mathbf{b}^T \mathbf{y} & = & 0 \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T \mathbf{y} - S & = & -C \\ X, S & \succeq & \mathbf{0}, \end{array} , \quad \text{or} \quad \begin{array}{rcl} XS & = & \mathbf{0} \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T y - S & = & -C \\ X, S & \succeq & \mathbf{0} \end{array}$$

Let X^* and S^* be optimal solutions with zero duality gap. Then

$$\boxed{\text{rank}(X^*)} + \text{rank}(S^*) \leq n.$$

Hint of the Proof: for any symmetric PSD matrix $P \in S^n$ with rank r, there is a factorization $P = V^T V$ where $V \in R^{r \times n}$ and columns of V are nonzero-vectors and orthogonal to each other.

There are X^* and S^* such that the ranks of X^* and S^* are maximal, respectively.

There are X^* and S^* such that the ranks of X^* and S^* are minimal, respectively.

If there is S^* such that $\text{rank}(S^*) \geq n - d$, then the maximal rank of X^* is at most d.

A Bound on Support/Rank

Theorem 7 *(Carathéodory's theorem)*

- *If there is a minimizer for (LP), then there is a minimizer of (LP) whose support size r satisfying $r \le m$.*
- *If there is a minimizer for (SDP), then there is a minimizer of (SDP) whose rank r satisfying $\frac{r(r+1)}{2} \le m$. Moreover, such a solution can be find in polynomial time.*

How Sharp is the Rank Bound? The rank bound is sharp: consider $n = 4$ and the SDP problem:

$$\begin{aligned} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet X &= 1, \ \forall i < j = 1, 2, 3, 4, \\ X &\succeq 0, \end{aligned}$$

Applications: Finding the extreme eigenvalue of a symmetric matrix and the singular value of any matrix are convex optimization!

Low-Rank SDP Solution

For simplicity, consider the SDP feasibility problem

$$A_i \bullet X = b_i \qquad i = 1, \ldots, m, \quad X \succeq \mathbf{0}$$

where $A_1, \ldots, A_m$ are positive semidefinite matrices and scalars $(b_1, \ldots, b_m) \geq \mathbf{0}$.

$$x_1 + x_2 + x_3 = 1,$$
$$\begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}.$$

We try to find an approximate $\hat{X} \succeq 0$ of rank at most d:

$$\beta(m, n, d) \cdot b_i \leq A_i \bullet \hat{X} \leq \alpha(m, n, d) \cdot b_i \qquad \forall\, i = 1, \ldots, m.$$

Here, $\alpha \geq 1$ and $\beta \in (0, 1]$ are called the distortion factors. Clearly, the closer are both to 1, the better.

Approximate Low-Rank SDP Theorem

Theorem 8 *Let* $r = \max\{\textit{rank}(A_i)\}$ *and* $\bar{X} = VV^T$ *be a feasible solution. Then, for any* $d \geq 1$*, the randomly generated*

$$\hat{X} = V[\sum_{i=1}^{d} \xi_i \xi_i^T]V^T, \quad \xi_i \in N(\mathbf{0}, \frac{1}{d}I)$$

$$\alpha(m,n,d) = \begin{cases} 1 + \dfrac{12\ln(4mr)}{d} & \textit{for } 1 \leq d \leq 12\ln(4mr) \\ 1 + \sqrt{\dfrac{12\ln(4mr)}{d}} & \textit{for } d > 12\ln(4mr) \end{cases}$$

and

$$\beta(m,n,d) = \begin{cases} \dfrac{1}{e(2m)^{2/d}} & \textit{for } 1 \leq d \leq 4\ln(2m) \\ \max\left\{\dfrac{1}{e(2m)^{2/d}}, 1 - \sqrt{\dfrac{4\ln(2m)}{d}}\right\} & \textit{for } d > 4\ln(2m) \end{cases}$$