Introduction to Semidefinite Programming: Motivation and Duality

Henry Wolkowicz*

June 16, 2001

*Department of Combinatorics & Optimization, University of Waterloo

Outline

- Basic Properties and Notation
- Examples
- Duality

1 Basic Properties and Notation

Basic linear **Semidefinite Programming** looks just like Linear Programming

$$
\textbf{(PSDP)} \qquad
\begin{array}{rccc}
p^* = & \max & \operatorname{trace} CX & (\langle C, X\rangle) \\
 & \text{s.t.} & \mathcal{A}X = b & \text{(linear)} \\
 & & X \succeq 0, \quad (X \in \mathcal{P}) & \text{(nonneg)}
\end{array}
$$

$$
\mathcal{A}(\alpha X_1 + \beta X_2) = \alpha \mathcal{A}(X_1) + \beta \mathcal{A}(X_2), \quad \forall X_i, \text{ and scalars } \alpha, \beta
$$

$\mathcal{S}^n$:= space of $n \times n$ real symmetric matrices, $(A = A^T)$
(Hermitian?)

$\preceq$ denotes the Löwner partial order
$A \preceq B$ if $B - A \succeq 0$ (positive semidefinite)

For $A \in \mathcal{S}^n$:
A is positive semidefinite (positive definite), (denoted $A \succeq 0$ $(A \succ 0)$,)
if $x^T A x \geq 0 (> 0), \ \forall x \neq 0$.

For $M, n \times n$, $Mx = \lambda x$ (eigenvector-eigenvalue equation, simplest linear operation is multiplication by scalar), equivalently, $\det(M - \lambda I) = 0$.

TFAE:

1. $A \succeq 0 \quad (A \succ 0)$
2. the eigenvalues $\lambda(A) \geq 0 \quad (\lambda(A) > 0)$
3. $A = R^T R$, for some R (Cholesky factorization if R is upper triangular - Gram matrix representation of n columns of R, $a_{ij} = u_i^T u_j$, R nonsingular if $A \succ 0$)
4. $A = \sum_i \alpha_i v_i v_i^T$ for some nonnegative α_i (sum of rank ones, $A \succ 0$??)
5. all principal minors (determinants of principal submatrices) are nonnegative (positive for $A \succ 0$, leading princ. minors??) (principal submatrix: delete some rows and corresp. cols.; history-Sylvester?)
6. $S = (S^{\frac{1}{2}})^2$, for some $S^{\frac{1}{2}} \succeq 0 \quad (\succ 0)$

For a general square matrix $M, n \times n$, define $\operatorname{trace} M = \sum_i M_{ii}$ Then

$$\operatorname{trace} M = \sum_i \lambda_i \quad \text{(trace is sum of eigenvalues)}$$

Theorem 1.0.0.1 *(Interlacing/Majorization Eigenvalues, Fan, 1954) Let $H, \bar{H}$ be an $n \times n$ Hermitian matrices of the form*

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \qquad \bar{H} = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix},$$

with H_{11} size $l \times l$. Then

$$(\lambda(H_{11}), \lambda(H_{22})) = \lambda(\bar{H}) \prec_M \lambda(H), \quad \text{on } \Re^n,$$

where $\prec_M$ denotes majorization, i.e. $\lambda \prec_M \mu$ <u>if</u> $\sum_{i=1}^k \lambda_{[i]} \leq \sum_{i=1}^k \mu_{[i]}$, with equality for $k = n$. ([.] denotes nonincreasing order). In particular,
$\lambda_j(H) \geq \lambda_j(H_{11}), \quad \lambda_{n-j+1}(H) \leq \lambda_{l-j+1}(H_{11}).$ ■

1.0.1 Further Properties and Definitions

- For matrices P, Q compatible for multiplication,

$$\operatorname{trace} PQ = \operatorname{trace} QP.$$

- Every symmetric matrix S has an orthogonal decomposition $S = Q\Lambda Q^T$, where Λ is diagonal with the eigenvalues of S on the diagonal, and Q has orthonormal columns consisting of eigenvectors of S.

- $X \succeq 0 \Rightarrow P^T X P \succeq 0 \quad$ (congruence)

Further properties of $A \succeq 0$:

• $A = \sum_{i=1}^{r} \alpha_i v_i v_i^T$ for minimum number of nonnegative α_i characterizes rank A;

• diag $(A) \geq 0$; and $a_{ii} = 0$ implies i-th and i-th col are 0;

• trace $A \geq 0$ and trace $A = 0$ iff $A = 0$;

• $\mathcal{P} \subset \mathcal{S}^n$ is a closed convex cone, i.e. $\mathcal{P} \subset \mathcal{P} + \mathcal{P}; \alpha\mathcal{P} \subset \mathcal{P}, \forall \alpha \geq 0$;

• $A, B \succeq 0$ implies trace $AB \geq 0$ with equality iff $AB = 0$ (proof?).

• geometry of $\mathcal{P}$ in $\mathcal{S}^n$ equipped with trace inner product: all vectors within 45 degrees of identity matrix;

• $\mathcal{P}$ is a *self-polar* cone (proof??), $\mathcal{P} = \mathcal{P}^*$, where

$$C^* = \{\phi \in \Re^n : \langle \phi, x \rangle \geq 0, \forall x \in C\}.$$

Lemma 1.0.1.1 (Schur Complement) *Let* $A = \begin{pmatrix} B & C^T \\ C & D \end{pmatrix}$ *be a symmetric matrix with the block* $B \succ 0$*. Then*

$$A \succ 0 \ (\succeq 0) \ \textit{iff} \ D - CB^{-1}C^T \succ 0 \ (\succeq 0).$$

Theorem 1.0.1.1 *(Perron-Frobenius) If an* $n \times n$ *matrix has nonnegative entries then it has a nonegative real eigenvalue* λ *which has maximum absolute value among all eigenvalues. This eigenvalue* λ *has a nonnegative real eigenvector. If, in addition, the matrix has no block-triangular decomposition (is irreducible - does not contain a* $k \times (n-k)$ *block of 0-s disjoint from the diagonal), then* λ *has multiplicity 1 and the corresponding eigenvector is positive.* ∎

The linear operator $\mathcal{A}$ and its adjoint:

$$\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m, \qquad \mathcal{A}^* : \mathbb{R}^m \to \mathcal{S}^n$$

$$\text{for given } A_i \in \mathcal{S}^n, i = 1, \ldots m :$$

$$(\mathcal{A}X)_i = \operatorname{trace}(A_i X), \quad \mathcal{A}^* y = \textstyle\sum_{i=1}^m y_i A_i$$

where

$$\langle \mathcal{A}X, y \rangle = \langle X, \mathcal{A}^* y \rangle, \quad \forall X, y$$

$\mathcal{P}$ - cone of positive semidefinite matrices

replaces

$\mathbb{R}^n_+$ - nonnegative orthant

$\mathrm{trace}\, CX = \langle C, X \rangle = \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij}$

SDP is equivalent to:

$$
\begin{array}{lrll}
 & p^* = & \max & \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \\
(\mathbf{PSDP}) & & \text{s.t.} & \langle A_i, X \rangle = b_i, \qquad i = 1, \dots m \\
 & & & X \succeq 0, \quad (X \in \mathcal{P})
\end{array}
$$

1.0.2 Duality (like LP)

payoff function, player Y to player X (Lagrangian)

$$L(X,y) := \operatorname{trace}(CX) + y^T(b - \mathcal{A}X)$$

Optimal (worst case) strategy for player X:

$$p^* = \max_{X \succeq 0} \min_y L(X,y)$$

Using the *hidden constraint* $b - \mathcal{A}X = 0$, recovers primal problem (PSDP).

$$
\begin{aligned}
L(X,y) &= \operatorname{trace}(CX) + y^T(b - \mathcal{A}X) \\
&= b^T y + \operatorname{trace}\,(C - \mathcal{A}^* y)\,X
\end{aligned}
$$

adjoint operator, $\quad \mathcal{A}^* y = \sum_i y_i A_i$

$$
\langle \mathcal{A}^* y, X \rangle = \langle y, \mathcal{A}X \rangle\,, \quad \forall X, y
$$

$$
p^* = \max_{X \succeq 0} \min_y L(X,y) \leq d^* := \min_y \max_{X \succeq 0} L(X,y)
$$

For dual, use *hidden constraint* $g(y) := C - \mathcal{A}^* y \preceq 0$

$$p^* = \max_{X \succeq 0} \min_y L(X,y) \le d^* := \min_y \max_{X \succeq 0} L(X,y)$$

dual obtained from optimal strategy of competing player Y;
use *hidden constraint* $g(y) = C - \mathcal{A}^* y \preceq 0$

$$(\textbf{DSDP}) \qquad \begin{array}{rcc} d^* = & \min & b^T y \\ & \text{s.t.} & \mathcal{A}^* y \succeq C \end{array}$$

for the primal

$$(\textbf{PSDP}) \qquad \begin{array}{rcc} p^* = & \max & \operatorname{trace} CX \\ & \text{s.t.} & \mathcal{A} X = b \\ & & X \succeq 0 \end{array}$$

1.0.3 Weak Duality - Optimality

Proposition 1.0.3.1 *(Weak Duality) If X feas. in (PSDP), y feas. in (DSDP), $Z = C - \mathcal{A}^* y \succeq 0$ is slack variable, then*

$$\textit{duality gap} \; := b^T y - \operatorname{trace} CX = \operatorname{trace} XZ \geq 0.$$

Proof. (Direct - using
$\operatorname{trace} ZX = \operatorname{trace} X^{1/2} X^{1/2} Z = \operatorname{trace} X^{1/2} Z X^{1/2} \geq 0$)

$$\begin{aligned} \operatorname{trace} CX - b^T y &= \operatorname{trace} (\mathcal{A}^* y - Z) X - b^T y \\ &= \operatorname{trace} y^T \mathcal{A} X - \operatorname{trace} ZX - b^T y \\ &= \operatorname{trace} y^T (\mathcal{A} X - b) - \operatorname{trace} ZX = -\operatorname{trace} ZX, \end{aligned}$$

∎

Characterization of optimality
(in the case of a zero duality gap at optimality and attainment)
for the dual pair X, y (slack $Z \succeq 0$)

$$\mathcal{A}^* y - Z = C \qquad \text{dual feasibility}$$

$$\mathcal{A} X = b \qquad \text{primal feasibility}$$

$$ZX = 0 \qquad \text{complementary slackness (equiv. } \operatorname{trace} ZX = 0\text{)}$$

$$ZX = \mu I \qquad \text{perturbed}$$

Forms the basis for:

interior point methods
(primal simplex method, dual simplex method)

1.0.4 Preliminary Examples

Example 1.0.4.1 Minimizing the Maximum Eigenvalue
Arises in e.g. stability of differential equations

- *The mathematical problem:*
  - ◇ *given $A(x)$ depending linearly on vector x*
  - ◇ *Find x to minimize the maximum eigenvalue of $A(x)$*
- *SDP Model:*
  - ◇ *the largest eigenvalue* $\lambda_{\max}(A(x)) \le \alpha$ <u>*iff*</u> $\lambda_{\max}(A(x) - \alpha I) \le 0$ <u>*iff*</u> $A(x) - \alpha I \preceq 0.$
  - ◇ *The DSDP (in dual form) is:*

$$\max -\alpha \qquad s.t. \quad A(x) - \alpha I \preceq 0.$$

HOW DOES SDP arise from quadratic approximations?

$$\text{Let} \quad q_i(y) = \frac{1}{2} y^T Q_i y + y^T b_i + c_i, \ y \in \mathbb{R}^n$$

$$(\mathbf{QQP}) \qquad \begin{cases} q^* = & \min & q_0(y) \\ & \text{s.t.} & q_i(y) = 0 \\ & & i = 1, \dots m \end{cases}$$

• Lagrangian (quadratic, linear, constant in y terms):

$$L(y,x) = \frac{1}{2} y^T (Q_0 - \sum_{i=1}^m x_i Q_i) y + y^T (b_0 - \sum_{i=1}^m x_i b_i) + (c_0 - \sum_{i=1}^m x_i c_i)$$

• Primal-Dual pair: $q^* = \min_y \max_x L(y,x) \geq d^* = \max_x \min_y L(y,x)$

• homogenize (add y_0): $\quad y_0 y^T (b_0 - \sum_{i=1}^m x_i b_i), \quad y_0^2 = 1.$

$$
\begin{array}{rcll}
d^* & = & \max_x \min_y & L(y,x) \\
 & = & \max_x \min\limits_{y_0^2=1} & \frac{1}{2} y^T (Q_0 - \sum_{i=1}^m x_i Q_i) y \quad (+t y_0^2) \\
 & & & + y_0 y^T (b_0 - \sum_{i=1}^m x_i b_i) \\
 & & & + (c_0 - \sum_{i=1}^m x_i c_i) \quad (-t)
\end{array}
$$

hidden semidefinite constraint yields SDP constraint

$$
(\mathcal{A} : \mathbb{R}^{m+1} \to \mathcal{S}_{n+1}) \qquad B - \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} \succeq 0.
$$

$$
B = \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix}, \quad \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{bmatrix} -t & \sum_{i=1}^m x_i b_i^T \\ \sum_{i=1}^m x_i b_i & \sum_{i=1}^m x_i Q_i \end{bmatrix}
$$

The dual program is equivalent to the SDP (with $c_0 = 0$)

$$
\textbf{(D)} \qquad
\begin{array}{rll}
d^* = & \sup & -\sum_{i=1}^m x_i c_i - t \\
& \text{s.t.} & \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} \preceq B \\
& & x \in \mathbb{R}^m, t \in \mathbb{R}
\end{array}
$$

As in linear programming, the dual is obtained from the optimal strategy of the competing player:

$$
\textbf{(DD)} \qquad
\begin{array}{rll}
d^* = & \inf & \operatorname{trace} BU \\
& \text{s.t.} & \mathcal{A}^* U = \begin{pmatrix} -1 \\ -c \end{pmatrix} \\
& & U \succeq 0.
\end{array}
$$

Example 1.0.4.2 Pseudoconvex (Nonlinear) Optimization Problem

$$
(\mathbf{PCP}) \qquad \begin{aligned} d^* = \quad & \min \quad \frac{(c^T x)^2}{d^T x} \\ & s.t. \quad Ax + b \geq 0, \end{aligned}
$$

(where $Ax + b \geq 0 \Rightarrow d^T x > 0$) is equivalent to

$$
\begin{aligned} d^* = \quad & \min \quad t \\ & s.t. \quad \begin{bmatrix} \mathrm{Diag}\,(Ax + b) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succeq 0 \end{aligned}
$$

- ***Alternate Name - Linear Matrix Inequality***

$$g(y) = C - \mathcal{A}^* y \preceq 0$$

adjoint operator $\qquad \mathcal{A}^* y = \sum_i y_i A_i, \quad A_i = A_i^T$

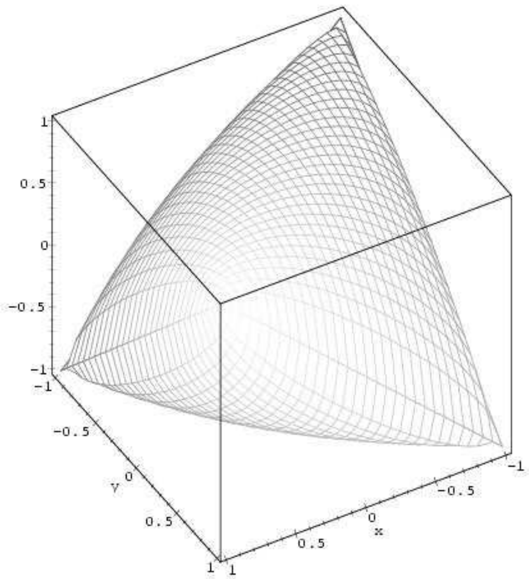
- $g(y)$ is a **convex constraint**, i.e.

$$g(\lambda x + (1-\lambda) y) \preceq \lambda g(x) + (1-\lambda) g(y), \quad \forall 0 \le \lambda \le 1, \forall x, y \in \mathbb{R}^m$$

equivalently, for every $P \succeq 0$, the real valued function $f_P(x) := \operatorname{trace} P g(x)$ is a convex function, i.e.

$$f_P(\lambda x + (1-\lambda) y) \le \lambda f_P(x) + (1-\lambda) f_P(y), \quad \forall 0 \le \lambda \le 1, \forall x, y \in \mathbb{R}^m$$

- *feasible set* $\{x \in \mathbb{R}^m : g(x) \preceq 0\}$ is convex set
- *optimum point* is on the boundary ($g(x)$ is singular)
- *boundary of feasible set* is *not* smooth (piecewise smooth, algebraic surface)



1.1 Introduction to Convex Cones and the Löwner Partial Order

Definition 1.1.0.1 *Let* $\alpha \in \mathbb{R}$ *and* $S, T \subset \mathbb{R}^n$. *Then*
$\alpha S = \{y : y = \alpha s, \text{ for some } s \in S\}$ *and*
$S + T = \{y : y = s + t, \text{ for some } s \in S, t \in T\}$

Definition 1.1.0.2 $\mathcal{K} \subset \mathbb{R}^n$ *is a* <u>*cone*</u> *if* $\alpha\mathcal{K} \subset K, \quad \forall \alpha > 0.$

Definition 1.1.0.3 *the cone* $\mathcal{K}$ *is a* <u>*convex cone*</u> *if* $\mathcal{K} + \mathcal{K} \subset \mathcal{K}$.

Definition 1.1.0.4 *A cone* $\mathcal{K}$ *is a* <u>*pointed cone*</u> *if* $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$.

Definition 1.1.0.5 *A cone* $\mathcal{K} \subset \mathbb{R}^n$ *is a* <u>*proper cone*</u> *if it is closed, pointed, and convex and has nonempty interior.*

Examples of Cones

Example 1.1.0.3 *open half line:* $\{x \in \mathbb{R} : x > 0\}$.

Example 1.1.0.4 *closed half line:* $\{x \in \mathbb{R} : x \geq 0\}$.

Example 1.1.0.5 *psd matrices,* $\mathcal{P}$*:* $\{X \in \mathcal{S}^n : X \succeq 0\}$.

Example 1.1.0.6 *Lorentz cone*
(ice-cream cone, second-order cone):
$L^m = \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m : x_m \geq \sqrt{x_1^2 + \ldots + x_m^2}$.

And, an important closed convex cone is:

Definition 1.1.0.6 *Polar (dual, conjugate) of set $\mathcal{S}$:*
$\mathcal{S}^+ := \{z : \langle x, z \rangle \geq 0, \forall x \in \mathcal{S}\} \quad (= S^*)$.

Example 1.1.0.7 *nonnegative orthant, psd cone $\mathcal{P}$, and Lorentz cone L^m, are all self-polar, i.e. $\mathcal{K} = \mathcal{K}^+$.*

Definition 1.1.0.7 *Direct sum of two cones*
$\mathcal{K} \oplus \mathcal{L} := \{(k\ l) : k \in \mathcal{K}, l \in \mathcal{L}\}$.

Example 1.1.0.8 *Let* L *denote the half line in* $\mathbb{R}$, *then* $\underbrace{L \oplus L \cdots \oplus L}_{n}$ *is the* n*-dimension nonnegative orthant.*

Suppose $\mathcal{K}$, $\mathcal{K}_1$, $\mathcal{K}_2$ are proper cones, then:

- $\mathcal{K}^+$ is proper.
- $(\mathcal{K}^+)^+ = \mathcal{K}$.
- $(\mathcal{K}_1 \cap \mathcal{K}_2)^+ = \overline{\mathcal{K}_1^+ + \mathcal{K}_2^+}$.
- $(\mathcal{K}_1 + \mathcal{K}_2)^+ = \mathcal{K}_1^+ \cap \mathcal{K}_2^+$.
- $\mathcal{K}_1 \oplus \mathcal{K}_2$ is proper.
- $(\mathcal{K}_1 \oplus \mathcal{K}_2)^+ = \mathcal{K}_1^+ \oplus \mathcal{K}_2^+$.

Useful Lemma, e.g. to Prove Farkas' Lemma

Lemma 1.1.0.1

$$\mathcal{K} \textit{ is a closed convex cone } \iff \mathcal{K} = \mathcal{K}^{++}.$$

Partial Orders:

Definition 1.1.0.8 $x \succeq_{\mathcal{K}} y$ *(respectively* $x \succ_{\mathcal{K}} y$*) if* $x - y \in \mathcal{K}$ *(respectively* $x - y \in \operatorname{int} \mathcal{K}$*).*

Remark 1.1.0.1 *If* $\mathcal{K}$ *is a pointed, convex cone, then "*$\succeq_{\mathcal{K}}$*" is a* *(linear) partial order* *(reflexive, transitive, and antisymmetric):*

- $0 \in \mathcal{K} \Rightarrow x \succeq_{\mathcal{K}} x$ *(reflexive);*
- $\mathcal{K}$ *is convex* $\Rightarrow$ *if* $x \succeq_{\mathcal{K}} y$ *and* $y \succeq_{\mathcal{K}} z$*, then* $x \succeq_{\mathcal{K}} z$ *(transitive);*
- $\mathcal{K}$ *is pointed* $\Rightarrow$ *if* $x \succeq_{\mathcal{K}} y$ *and* $y \succeq_{\mathcal{K}} x$*, then* $x = y$ *(antisymmetric);*
- $\mathcal{K}$ *is convex cone* $\Rightarrow$
  if $a, b \geq 0, u \succeq_{\mathcal{K}} x$ *and* $v \succeq_{\mathcal{K}} y$*, then* $au + bv \succeq_{\mathcal{K}} ax + by$ *(linear - homogeneous, additive);*

1.2 Convex Cone Program

Let: $\mathcal{K}, \mathcal{L}$ be convex cones;

f real valued convex function, g is $\mathcal{K}$-convex, i.e.

$g(\alpha u + (1-\alpha)v) \preceq_{\mathcal{K}} \alpha g(u) + (1-\alpha)g(v), \forall 0 \leq \alpha \leq 1, \forall u, v$

$$
\begin{array}{lccc}
 & \mu^* = \min & f(x) \\
(CP) & \text{s.t.} & g(x) \preceq_{\mathcal{K}} 0 \\
 & & x \succeq_{\mathcal{L}} 0
\end{array}
$$

with Lagrangian dual (weak duality)

$$
(DCP) \qquad \mu^* \geq \nu^* = \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \quad f(x) + \langle y, g(x) \rangle
$$

Example 1.2.0.9 *If $f(x) = \langle c, x \rangle$, $g(x) = b - \mathcal{A}x$, then we have a linear cone programming problem, (LCP). The dual*

$$
(DLCP) \qquad
\begin{aligned}
\mu^* \geq \nu^* &= \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \quad \langle c, x \rangle + \langle y, b - \mathcal{A}x \rangle \\
&= \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \quad \langle c - \mathcal{A}^* y, x \rangle + \langle y, b \rangle \\
&= \max_{\substack{y \in \mathcal{K}^+ \\ c - \mathcal{A}^* y \in \mathcal{L}^+}} \min_{x \in \mathcal{L}} \quad \langle c - \mathcal{A}^* y, x \rangle + \langle y, b \rangle
\end{aligned}
$$

reduces to the elegant (LP or SDP type) form

$$
(DLCP) \qquad
\begin{array}{rl}
\mu^* \geq \nu^* = \max & \langle y, b \rangle \\
s.t. & \mathcal{A}^* y \preceq_{\mathcal{L}^+} c \\
 & y \succeq_{\mathcal{K}^+} 0
\end{array}
$$

$$
(LCP) \qquad \begin{array}{rl} \mu^* = \min & f(x) \\ \text{s.t.} & \mathcal{A}x \succeq_{\mathcal{K}} b \\ & x \succeq_{\mathcal{L}} 0 \end{array}
$$

$$
(DLCP) \qquad \begin{array}{rl} \mu^* \geq \nu^* = \max & \langle y, b \rangle \\ \text{s.t.} & \mathcal{A}^* y \preceq_{\mathcal{L}^+} c \\ & y \succeq_{\mathcal{K}^+} 0 \end{array}
$$

Theorem 1.2.0.1 *Suppose that Slater's CQ holds for (LCP), i.e.* $\exists \hat{x}$ *such that* $\mathcal{A}\hat{x} \succ_{\mathcal{K}} b, \quad \hat{x} \succ_{\mathcal{L}} 0$. *Then* strong duality *holds for (LCP),(DLCP), i.e. there exists* y^* *feasible for (DLCP) such that*

$$
\mu^* = \nu^* = \langle y^*, b \rangle .
$$

Example

If the primal is

$$
\begin{array}{rll}
p^* = & \sup & x_2 \\
(\mathbf{P}) & \text{s. t.} & \begin{bmatrix} x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{array}
$$

$$
S^f = cone \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}
$$

Then the dual is

$$
(\mathbf{D}) \qquad \begin{array}{rcc} d^* = & \inf & U_{11} \\ & \text{s. t.} & U_{22} = 0 \\ & & U_{11} + 2U_{23} = 1 \\ & & U \succeq 0. \end{array}
$$

Then $p^* = 0 < d^* = 1$.

(But regularized dual (below) has $U \succeq_{(S^f)^+} 0$, i.e. only constraint is $U_{11} \geq 0$. So new dual optimal value is 0. (Attained.))

1.3 Optimality Conditions without constraint qualifications

1.3.1 Facial Structure

Definition 1.3.1.1 *$\mathcal{F}$ is a* face *of a convex cone $\mathcal{K}$ (denoted $\mathcal{F} \lhd \mathcal{K}$) if $\mathcal{F} \subset \mathcal{K}$ and*

$$\frac{1}{2}(x_1 + x_2) \in \mathcal{F} \Rightarrow x_1, x_2 \in \mathcal{F}, \qquad \forall x_1, x_2 \in \mathcal{K}$$

Definition 1.3.1.2 *A face $\mathcal{F} \lhd \mathcal{K}$ is called* exposed *if*

$$\mathcal{F} = \mathcal{K} \cap \phi^{\perp}, \qquad \text{for some } \phi \in \mathcal{K}^{+}$$

Definition 1.3.1.3 *A face $\mathcal{F} \lhd \mathcal{P}$ is called* projectionally exposed *if*

$$\mathcal{F} = Q\mathcal{P}Q^{t} \qquad \text{for some matrix } Q$$

1.3.2 Weak Duality

The weak duality: $c^T x \geq b^T y$ holds in Cone-LP:

$$c^T x - b^T y = c^T x - (Ax)^T y = x^T(c - A^T y) = x^T s \geq 0$$

1.3.3 Extended Farkas' Lemma

Theorem 1.3.3.1 ***(Seperation Theorem)*** *If $K \subseteq \mathbb{R}^n$ is closed and convex, $b \in \mathbb{R}^n$, $b \notin K$, then $\exists a \in \mathbb{R}^n$, such that* $\forall x \in K : a^T b < 0$ *and* $a^T x \geq 0$.

Lemma 1.3.3.1 ***(Farkas' Lemma)*** *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed, convex cone, $A \in \mathbb{R}^{m \times n}, A(\mathcal{K})$ closed, and $b \in \mathbb{R}^m$. Then:*

$$\{\exists x \succeq_{\mathcal{K}} 0 : Ax = b\} \iff \left\{A^T y \succeq_{\mathcal{K}^+} 0 \Rightarrow b^T y \geq 0\right\}.$$

Proof. (Necessity:) Suppose $\exists x \in \mathcal{K} : Ax = b$. And suppose $A^T y \succeq_{\mathcal{K}^+} 0$. Then $b^T y = (Ax)^T y = x^T (A^T y) \geq 0$, since $x \succeq_K 0$.

(Sufficiency:) Suppose that the RHS of (1.3.3.1) holds but the LHS fails, i.e. $b \notin A(\mathcal{K})$. Then separate, i.e. $\exists y \in \mathbb{R}_m$ such that $b^T y < 0$, and $\forall x \in \mathcal{K}, (Ax)^T y \geq 0$. Therefore, $A^T y \in K^+$ which implies $b^T y \geq 0$, contradiction. $\blacksquare$

Cone programming: S a closed convex cone induces a linear partial order ([11])

$$\min\{f(x) : g(x) \preceq_S 0,\ x \in \Omega\}$$

feasible set A

---

rewrite using minimal face $g(A) \subset -S^f$

$$\min\{f(x) : g(x) \preceq_{S^f} 0,\ x \in \Omega^f\}$$

where $\Omega^f = \Omega \cap g^{-1}(S^f - S^f) = \Omega \cap g^{-1}(S^f - S)$

---

optimality conditions for some $\Lambda \in (S^f)^+$:

$$f(x) + \Lambda g(x) \geq \mu^*,\ \forall x \in \Omega^f$$

Special case of linear cone programming, $\mathcal{K}, \mathcal{L}$ closed convex cones ([85]):

primal:

$$\mu^* = \min\{cx : \mathcal{A}x \succeq_{\mathcal{K}} 0,\ x \succeq_{\mathcal{L}} 0\}$$

minimal cones $\mathcal{K}^f, \mathcal{L}^f$

dual:

$$\mu^* = \nu^* = \max\{by : \mathcal{A}^* y \preceq_{(\mathcal{L}^f)^+} b,\ y \succeq_{(\mathcal{K}^f)^+} 0\}$$