Conic Duality

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

http://www.stanford.edu/~yyye
Vectors and Norms

- **Real numbers:** $\mathbb{R}$, $\mathbb{R}_+$, $\text{int} \mathbb{R}_+$
- **$n$-dimensional Euclidean space:** $\mathbb{R}^n$, $\mathbb{R}_+^n$, $\text{int} \mathbb{R}_+^n$
- **Component-wise:** $\mathbf{x} \geq \mathbf{y}$ means $x_j \geq y_j$ for $j = 1, 2, \ldots, n$
- **$\mathbf{0}$:** vector of all zeros; and **$\mathbf{e}$:** vector of all ones
- **Inner-product** of two vectors:
  \[ \mathbf{x} \cdot \mathbf{y} := \mathbf{x}^T \mathbf{y} = \sum_{j=1}^{n} x_j y_j \]
- **Euclidean norm:** $\| \mathbf{x} \|_2 = \sqrt{\mathbf{x}^T \mathbf{x}}$,
  - **Infinity-norm:** $\| \mathbf{x} \|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$,
  - **$p$-norm:** $\| \mathbf{x} \|_p = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p}$
• The **dual** of the $p$ norm, denoted by $\| \cdot \|^*$, is the $q$ norm, where $\frac{1}{p} + \frac{1}{q} = 1$

• Column vector:

$$\mathbf{x} = (x_1; x_2; \ldots; x_n)$$

Row vector:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

• Transpose operation: $A^T$

• A set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is said to be **linearly dependent** if there are scalars $\lambda_1, \ldots, \lambda_m$, not all zero, such that the **linear combination**

$$\sum_{i=1}^{m} \lambda_i \mathbf{a}_i = \mathbf{0}$$

• A **linearly independent** set of vectors that span $\mathbb{R}^n$ is a **basis**.
**Hyper plane and Half-spaces**

\[ H = \{ x : ax = \sum_{j=1}^{n} a_j x_j = b \} \]

\[ H^+ = \{ x : ax = \sum_{j=1}^{n} a_j x_j \leq b \} \]

\[ H^- = \{ x : ax = \sum_{j=1}^{n} a_j x_j \geq b \} \]
Figure 1: Plane and Half-Spaces

\[ 3x + 5y > 15 \]
\[ 3x + 5y = 15 \]
\[ 3x + 5y < 15 \]
Matrices and Norms

- **Matrix:** \( \mathbb{R}^{m \times n} \), \( i \)th row: \( a_i \), \( j \)th column: \( a_{.j} \), \( ij \)th element: \( a_{ij} \)

- \( A_I \) denotes the submatrix of \( A \) whose rows belong to index set \( I \), \( A_J \) denotes the submatrix whose columns belong to index set \( J \), \( A_{IJ} \) denotes the submatrix whose rows belong to index set \( I \) and columns belong to index set \( J \).

- **Determinant:** \( \det(A) \), **Trace:** \( \text{tr}(A) \)

- The operator norm of \( \|A\| \),

\[
\|A\|^2 := \max_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|^2}{\|x\|^2}
\]

- All-zero matrix: \( 0 \), and identity matrix: \( I \)

- Symmetric matrix: \( Q = Q^T \)
• Positive Definite: $Q \succ 0$ iff $x^T Q x > 0$, for all $x \neq 0$

• Positive Semidefinite: $Q \succeq 0$ iff $x^T Q x \geq 0$, for all $x$

• Null space and Row space of matrix $A$:

**Theorem 1**  The null space and row space of a matrix are perpendicular to each other, that is,

$$x^T s = 0, \quad \forall \ Ax = 0 \text{ and } s = A^T y.$$
Symmetric Matrix Space

- $n$-dimensional symmetric matrix space: $S^n$

- Inner Product:
  \[ X \bullet Y = \text{tr} X^T Y = \sum_{i,j} X_{i,j} Y_{i,j} \]

- Frobenius norm:
  \[ \|X\|_f = \sqrt{\text{tr} X^T X} \]

- Positive semidefinite matrix set: $S^n_+$, Positive definite matrix set: int $S^n_+$
• Decomposition of Symmetric Positive Semidefinite Matrices:

\[ X = \sum_{i=1}^{r} x_i x_i^T \]

where \( r \) is the rank of \( X \), and \( x_i^T x_j = 0 \) for \( i \neq j \).

• Let \( X \) be a positive semidefinite matrix of rank \( r \), \( A \) be any given symmetric matrix. Then, there is a decomposition of \( X \)

\[ X = \sum_{j=1}^{r} x_j x_j^T, \]

such that for all \( j \),

\[ x_j^T A x_j = A \bullet (x_j x_j^T) = A \bullet X/r. \]
**Affine and Convex Combination**

$S \subset \mathbb{R}^n$ is affine if

$$[x, y \in S \text{ and } \alpha \in \mathbb{R}] \implies \alpha x + (1 - \alpha)y \in S.$$

When $x$ and $y$ are two distinct points in $\mathbb{R}^n$ and $\alpha$ runs over $\mathbb{R}$,

$$\{z : z = \alpha x + (1 - \alpha)y\}$$

is the line set determined by $x$ and $y$.

When $0 \leq \alpha \leq 1$, it is called the convex combination of $x$ and $y$ and it is the line segment between $x$ and $y$. 
Convex Sets

- Set notations: \( x \in \Omega, \ y \not\in \Omega \ S \cup T, \ S \cap T \)

- \( \Omega \) is said to be a **convex set** if for every \( x^1, x^2 \in \Omega \) and every real number \( \alpha \in [0, 1] \), the point \( \alpha x^1 + (1 - \alpha)x^2 \in \Omega \).

- **Intersection** of convex sets is convex; the **convex hull** of a set \( \Omega \) is the intersection of all convex sets containing \( \Omega \).

- A point in a set is called an **extreme point** of the set if it cannot be represented as the convex combination of two distinct points of the set.

- A set is a **polyhedral** set if it has finitely many extreme points.
Let $C_1$ and $C_2$ be convex sets in a same space. Then,

- $C_1 \cap C_2$ is convex.
- $C_1 + C_2$ is convex, where
  \[ C_1 + C_2 = \{b_1 + b_2 : b_1 \in C_1 \text{ and } b_2 \in C_2\}. \]
- $C_1 \oplus C_2$ is convex, where
  \[ C_1 \oplus C_2 = \{(b_1; b_2) : b_1 \in C_1 \text{ and } b_2 \in C_2\}. \]
Cones

- A set $K$ is a cone if $x \in K$ implies $\alpha x \in K$ for all $\alpha > 0$

- A convex cone is cone and it’s also a convex-set.

- Dual cone:
  \[ K^* := \{ y : y \cdot x \geq 0 \text{ for all } x \in K \} \]
  $-K^*$ is also called the polar of $K$.

- The dual of a cone is always a closed convex cone.
Cone Examples

• Example 2.1: The \( n \)-dimensional non-negative orthant,
  \( \mathcal{R}_+^n = \{ x \in \mathcal{R}^n : x \geq 0 \} \), is a convex cone; and it’s self dual.

• Example 2.2: The set of all positive semi-definite symmetric matrices in \( S_+^n \),
  \( S_+^n \), is a convex cone, called the positive semi-definite matrix cone; and it’s self dual.

• Example 2.3: The set \( \{ x \in \mathcal{R}^n : x_1 \geq \| x_{-1} \| \} \), \( \mathcal{N}_2^n \), is a convex cone in \( \mathcal{R}^n \) called the second-order (norm) cone; and it’s self dual.

• Example 2.4: The set \( \{ x \in \mathcal{R}^n : x_1 \geq \| x_{-1} \|_p \} \), \( \mathcal{N}_p^n \), is a convex cone in \( \mathcal{R}^n \) for \( p \geq 1 \) called the \( p \)-order (norm) cone; and its dual is the \( q \)-order cone where \( \frac{1}{p} + \frac{1}{q} = 1 \).
Cone and Dual Facts

Let $K_1$ and $K_2$ be both closed convex cones. Then

i) $(K_1^*)^* = K_1$.

ii) $K_1 \subset K_2 \implies K_2^* \subset K_1^*$.

iii) $(K_1 \oplus K_2)^* = K_1^* \oplus K_2^*$.

iv) $(K_1 + K_2)^* = K_1^* \cap K_2^*$.

v) $(K_1 \cap K_2)^* = K_1^* + K_2^*$. 
Convex Polyhedral Cones I

- A cone $K$ is (convex) **polyhedral** if its intersection with a hyperplane is a polyhedral set.

- A convex cone $K$ is **polyhedral** if and only if $K$ can be represented by

  $K = \{ x : Ax \leq 0 \}$ or $\{ x : x = Ay, \ y \geq 0 \}$

  for some matrix $A$. In the latter case, $K$ is generated by the columns of $A$.

- The nonnegative orthant is a polyhedral cone but the second-order cone is not polyhedral.
Figure 2: Polyhedral and non-polyhedral cones.
Convex Polyhedral Cones II

It has been proved that for cones the concepts of “polyhedral” and “finitely generated” are equivalent according to the following theorem.

**Theorem 2** (*Minkowski and Weyl*) A convex cone $C$ is polyhedral if and only if it is finitely generated, that is, the cone is generated by a finite number of vectors $b_1, ..., b_m$:

$$C = \text{cone}(b_1, ..., b_m) := \left\{ \sum_{i=1}^{m} b_i y_i : y_i \geq 0, \ i = 1, ..., m \right\}.$$
Carathéodory’s theorem

The following theorem states that a polyhedral cone can be generated by a set of basic directional vectors.

**Theorem 3** Let convex polyhedral cone $C = \text{cone}(b_1, \ldots, b_m)$ and $x \in C$. Then, $x \in \text{cone}(b_{i_1}, \ldots, b_{i_d})$ for some linearly independent vectors $b_{i_1}, \ldots, b_{i_d}$ chosen from $b_1, \ldots, b_m$.

Some times we even have:

$$\begin{cases} x \in \mathbb{R}_+^2 : \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} x \leq 0 \end{cases} = \begin{cases} \begin{pmatrix} 1 \\ 2 \end{pmatrix} y_1 + \begin{pmatrix} 2 \\ 1 \end{pmatrix} y_2 : y_1, y_2 \geq 0 \end{cases}.$$
Figure 3: Representations of a polyhedral cone.
The most important theorem about the convex set is the following **separating hyperplane** theorem (Figure 4).

**Theorem 4**  (*Separating hyperplane theorem*) Let $C \subset \mathcal{E}$, where $\mathcal{E}$ is either $\mathbb{R}^n$ or $S^n$, be a closed convex set and let $b$ be a point exterior to $C$. Then there is a vector $a \in \mathcal{E}$ such that

$$a \cdot b > \sup_{x \in C} a \cdot x$$

where $a$ is the norm direction of the hyperplane.
Figure 4: Illustration of the separating hyperplane theorem; an exterior point $b$ is separated by a hyperplane from a convex set $C$. 
Examples

Let $C$ be a unit circle centered at point $(1; 1)$. That is,

$$C = \{ \mathbf{x} \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \}.$$ 

If $\mathbf{b} = (2; 0)$, $\mathbf{a} = (-1; 1)$ is a separating hyperplane vector.

If $\mathbf{b} = (0; -1)$, $\mathbf{a} = (0; 1)$ is a separating hyperplane vector. It is worth noting that these separating hyperplanes are not unique.
Farkas’ Lemma for Polyhedral Cone

**Theorem 5** Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then, the system
\[
\{ x : Ax = b, \ x \in \mathbb{R}^n_+ \}
\]
has a feasible solution $x$ if and only if that
\[
\{ y : -A^T y \in \mathbb{R}^n_+, \ b^T y > 0, \ (b^T y = 1) \}
\]
has no feasible solution.

A vector $y$, with $A^T y \leq 0$ and $b^T y > 0$, is called an infeasibility certificate for the system $\{ x : Ax = b, \ x \geq 0 \}$.

**Example:** Let $A = (1, 1)$ and $b = -1$. Then, $y = -1$ is an infeasibility certificate for $\{ x : Ax = b, \ x \geq 0 \}$. 
Farkas’ lemma is also called the alternative theorem, that is, exactly one of the two systems:

\[
\{ x : A x = b, \ x \geq 0 \}
\]

and

\[
\{ y : A^T y \leq 0, \ b^T y > 0, \ (b^T y = 1) \},
\]

is feasible.

Geometrically, Farkas’ lemma means that if a vector \( b \in \mathcal{R}^m \) does not belong to the cone generated by \( a_1, \ldots, a_n \), then there is a hyperplane separating \( b \) from \( \text{cone}(a_1, \ldots, a_n) \), that is,

\[
b \notin C := \{ A x : x \geq 0 \},
\]

which is a closed convex set(?).
Proof

Let \( \{x : Ax = b, \ x \geq 0\} \) have a feasible solution, say \( \bar{x} \). Then, \( \{y : A^T y \leq 0, \ b^T y > 0\} \) is infeasible, since otherwise,

\[
0 < b^T y = (Ax)^T y = x^T (A^T y) \leq 0
\]

since \( x \geq 0 \) and \( A^T y \leq 0 \).

Now let \( \{x : Ax = b, \ x \geq 0\} \) have no feasible solution, that is, \( b \not\in C := \{Ax : x \geq 0\} \). Then, by the separating hyperplane theorem, there is \( y \) such that

\[
y \cdot b > \sup_{c \in C} y \cdot c
\]

or

\[
y \cdot b > \sup_{x \geq 0} y \cdot (Ax) = \sup_{x \geq 0} A^T y \cdot x. \quad (1)
\]

Since \( 0 \in C \) we have \( y \cdot b > 0 \).
Furthermore, $A^T y \leq 0$. Since otherwise, say $(A^T y)_1 > 0$, one can have a vector $\bar{x} \geq 0$ such that $\bar{x}_1 = \alpha > 0$, $\bar{x}_2 = \ldots = \bar{x}_n = 0$, from which

$$
\sup_{x \geq 0} A^T y \cdot x \geq A^T y \cdot \bar{x} = (A^T y)_1 \cdot \alpha
$$

and it tends to $\infty$ as $\alpha \to \infty$. This is a contradiction because $\sup_{x \geq 0} A^T y \cdot x$ is bounded from above by (1).
Farkas’ Lemma variant

**Theorem 6** Let \( A \in \mathbb{R}^{m \times n} \) and \( c \in \mathbb{R}^n \). Then, the system \( \{ y : A^T y \leq c \} \) has a solution \( y \) if and only if that \( Ax = 0, x \geq 0, c^T x < 0 \) has no feasible solution \( x \).

Again, a vector \( x \geq 0 \), with \( Ax = 0 \) and \( c^T x < 0 \), is called a **infeasibility certificate** for the system \( \{ y : A^T y \leq c \} \).

**Example:** Let \( A = (1; -1) \) and \( c = (1; -2) \). Then, \( x = (1; 1) \) is an infeasibility certificate for \( \{ y : A^T y \leq c \} \).
Alternative Systems for General Cone?

\[ \{ x : Ax = b, \ x \in C \} \]

and

\[ \{ y : -A^T y \in C^*, \ b^T y > 0 \} ? \]

Counterexample:

\[ A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \]

and

\[ b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \ C = S^2_+ . \]
**Farkas Lemma for General Convex Cone**

**Theorem 7** Consider system \( \{ x : Ax = b, \ x \in K \} \) for a (closed) convex cone \( K \). Suppose that there exists vector \( \bar{y} \) such that \( -A^T \bar{y} \in \text{int} \ K^* \). Then,

- Set \( C := \{ Ax : x \in K \} \) is a closed convex set.
- The system \( \{ x : Ax = b, \ x \in K \} \) has a feasible solution \( x \) if and only if that \( \{ y : -A^T y \in K^*, \ b^T y > 0, \ (b^T y = 1) \} \) has no feasible solution.

**Corollary 1** Consider system \( \{(y, s) : A^T y + s = c, \ s \in K \} \) for a (closed) convex cone \( K \). Suppose that there exists vector \( \bar{x} \) such that \( A\bar{x} = 0, \ \bar{x} \in \text{int} \ K^* \). Then,

- Set \( C := \{ A^T y + s : s \in K \} \) is a closed convex set.
- The system \( \{ x : Ax = 0, \ c \cdot x = -1, \ x \in K^* \} \) has a feasible solution \( x \) if and only if that \( \{(y, s) : A^T y + s = c, \ s \in K \} \) has no feasible solution.