

## **Feasibility, Optimality and Conic Duality I**

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

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

## Carathéodory's theorem

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

**Theorem 1** Given matrix  $A \in R^{m \times n}$  where  $n > m$ , let convex polyhedral cone  $C = \{A\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}$ . For any  $\mathbf{b} \in C$ ,

$$\mathbf{b} = \sum_{i=1}^d \mathbf{a}_{j_i} x_{j_i}, \quad x_{j_i} \geq 0, \forall i$$

for some **linearly independent** vectors  $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_d}$  chosen from  $\mathbf{a}_1, \dots, \mathbf{a}_n$ .

## Basic and Basic Feasible Solution I

Now consider the polyhedron set  $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ . Select  $m$  linearly independent columns, denoted by the variable index set  $B$ , from  $A$ . Solve  $A_B \mathbf{x}_B = \mathbf{b}$  for the  $m$ -dimension vector  $\mathbf{x}_B$ , and set the remaining variables,  $\mathbf{x}_N$ , to zero. Then, we obtain a solution  $\mathbf{x}$  such that  $A\mathbf{x} = \mathbf{b}$ , that is called a basic solution to with respect to the basis  $B$ . If a basic solution  $\mathbf{x}_B \geq \mathbf{0}$ , then  $\mathbf{x}$  is called a basic feasible solution, or BFS.

BFS is an extreme or corner point of the polyhedron.

Carathéodory's theorem implies if there is a feasible solution to system  $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ , then there is a basic feasible solution to the system.

## Hyper Planes

The most important type of convex set is **hyperplane**, also called *linear variety* or *affine set*. if for any two points are in  $H$  then their **linear or affine combination** is also in  $H$ .

Hyperplanes dominate the entire theory of optimization. Let  $\mathbf{a}$  be a nonzero  $n$ -dimensional vector, and let  $b$  be a real number. The set

$$H = \{\mathbf{x} \in \mathcal{R}^n : \mathbf{a} \bullet \mathbf{x} = b\}$$

is a hyperplane in  $\mathcal{R}^n$ . Relating to hyperplane, positive and negative closed **half spaces** are given by

$$H_+ = \{\mathbf{x} : \mathbf{a} \bullet \mathbf{x} \geq b\}$$

$$H_- = \{\mathbf{x} : \mathbf{a} \bullet \mathbf{x} \leq b\}.$$

## Separating and supporting hyperplane theorem

The most important theorem about the convex set is the following **separating hyperplane theorem** (Figure 1).

**Theorem 2** (Separating hyperplane theorem) Let  $C$  be a closed convex set in  $\mathcal{R}^m$  and let  $\mathbf{b}$  be a point exterior to  $C$ . Then there is a vector  $\mathbf{y} \in \mathcal{R}^m$  such that

$$\mathbf{b} \bullet \mathbf{y} > \sup_{x \in C} \mathbf{x} \bullet \mathbf{y}.$$

**Theorem 3** (Supporting hyperplane theorem) Let  $C$  be a closed convex set and let  $\mathbf{b}$  be a point on the boundary of  $C$ . Then there is a vector  $\mathbf{y} \in \mathcal{R}^m$  such that

$$\mathbf{b} \bullet \mathbf{y} = \sup_{x \in C} \mathbf{x} \bullet \mathbf{y}.$$

## Separating Hyperplane Theorem

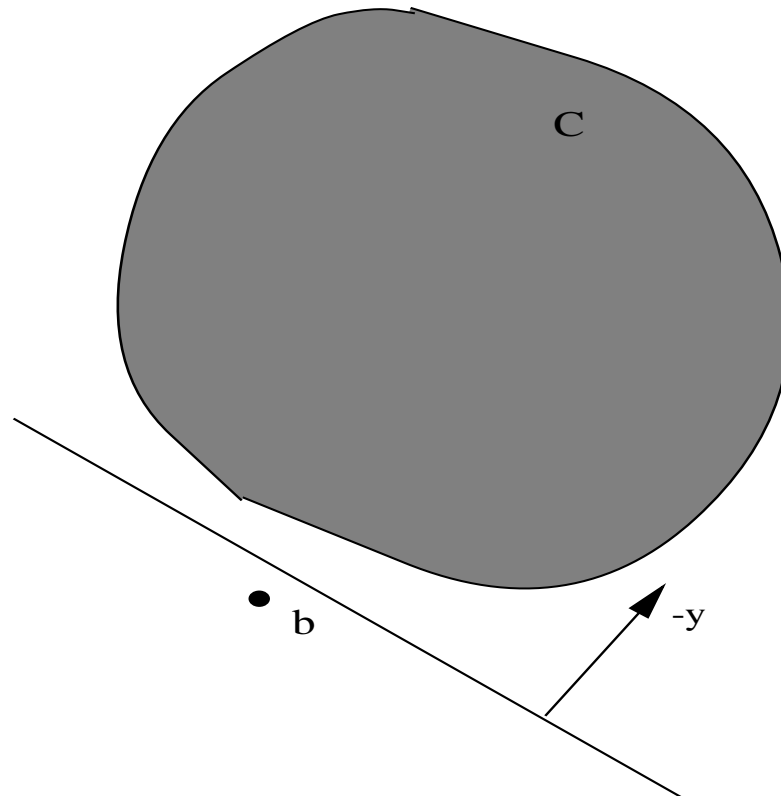


Figure 1: Illustration of the separating hyperplane theorem; an exterior point  $b$  is separated by a hyperplane from a convex set  $C$ .

### Example

Let  $C$  be a unit circle centered at point  $(1; 1)$ . That is,  
 $C = \{x \in \mathcal{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1\}$ . If  $\mathbf{b} = (2; 0)$ ,  $\mathbf{y} = (1; -1)$  is a **separating hyperplane** vector. If  $\mathbf{b} = (0; -1)$ ,  $\mathbf{y} = (0; -1)$  is a **separating hyperplane** vector. It is worth noting that these separating hyperplanes are not unique.

## Farkas' Lemma

The following results are Farkas' lemma and its variants.

**Theorem 4** Let  $A \in \mathcal{R}^{m \times n}$  and  $\mathbf{b} \in \mathcal{R}^m$ . Then, the system  $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$  has a feasible solution  $\mathbf{x}$  if and only if that  $-A^T \mathbf{y} \geq \mathbf{0}$  and  $\mathbf{b}^T \mathbf{y} > 0$  has no feasible solution  $\mathbf{y}$ .

Geometrically, Farkas' lemma means that if a vector  $\mathbf{b} \in \mathcal{R}^m$  does not belong to the convex cone generated by  $\mathbf{a}_{.1}, \dots, \mathbf{a}_{.n}$ , then there is a hyperplane separating  $\mathbf{b}$  from  $\text{cone}(\mathbf{a}_{.1}, \dots, \mathbf{a}_{.n})$ .

**Example** Let  $A = (1, 1)$  and  $b = -1$ . Then, there is  $y = -1$  such that  $-A^T y \geq 0$  and  $by > 0$ .



**Proof**

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

$$0 < \mathbf{b}^T \mathbf{y} = (A\bar{\mathbf{x}})^T \mathbf{y} = \bar{\mathbf{x}}^T (A^T \mathbf{y}) \leq 0$$

since  $\bar{\mathbf{x}} \geq \mathbf{0}$  and  $A^T \mathbf{y} \leq \mathbf{0}$ .

Now let  $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$  have no feasible solution, that is,  $\mathbf{b} \notin C := \{A\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}$ . Since  $C$  is a **closed** convex set from Carathéodory's theorem (?), by the separating hyperplane theorem, there is  $\mathbf{y}$  such that

$$\mathbf{y} \bullet \mathbf{b} > \sup_{\mathbf{c} \in C} \mathbf{y} \bullet \mathbf{c}$$

or

$$\mathbf{y} \bullet \mathbf{b} > \sup_{\mathbf{x} \geq \mathbf{0}} \mathbf{y} \bullet (A\mathbf{x}) = \sup_{\mathbf{x} \geq \mathbf{0}} A^T \mathbf{y} \bullet \mathbf{x}. \quad (1)$$

Since  $\mathbf{0} \in C$  we have  $\mathbf{y} \bullet \mathbf{b} > 0$ .

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

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

and it tends to  $\infty$  as  $\alpha \rightarrow \infty$ . This is a contradiction because  $\sup_{\mathbf{x} \geq \mathbf{0}} A^T \mathbf{y} \bullet \mathbf{x}$  is bounded from above by (1).

## Farkas' Lemma Variant

**Theorem 5** Let  $A \in \mathcal{R}^{m \times n}$  and  $\mathbf{c} \in \mathcal{R}^n$ . Then, the system  $\{\mathbf{y} : \mathbf{c} - A^T \mathbf{y} \geq \mathbf{0}\}$  has a solution  $\mathbf{y}$  if and only if that  $A\mathbf{x} = \mathbf{0}$ ,  $\mathbf{x} \geq \mathbf{0}$ , and  $\mathbf{c}^T \mathbf{x} < 0$  has no feasible solution  $\mathbf{x}$ .

**Example** Let  $A = (1; -1)$  and  $\mathbf{c} = (1; -2)$ . Then, there is  $\mathbf{x} = (1; 1) \geq \mathbf{0}$  such that  $A\mathbf{x} = 0$  and  $\mathbf{c}^T \mathbf{x} < 0$ .

## Alternative System Pair I

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}.$$

$$-A^T \mathbf{y} \geq \mathbf{0}, \quad \mathbf{b}^T \mathbf{y} = 1 (> 0)$$

A vector  $\mathbf{y}$ , with  $A^T \mathbf{y} \leq \mathbf{0}$  and  $\mathbf{b}^T \mathbf{y} = 1$ , is called an **infeasibility certificate** for the system  $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ .

## Alternative System Pair II

$$A\mathbf{x} = \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1 (< 0).$$

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

A vector  $\mathbf{x}$ , with  $A\mathbf{x} = \mathbf{0}$ ,  $\mathbf{x} \geq \mathbf{0}$  and  $\mathbf{c}^T \mathbf{x} = -1$ , is called an **infeasibility certificate** for the system  $\{\mathbf{y} : \mathbf{c} - A^T \mathbf{y} \geq \mathbf{0}\}$ .

## Farkas' Lemma for General Cones?

Given  $\mathbf{a}_i, i = 1, \dots, m$ , and  $\mathbf{b} \in \mathcal{R}^m$ . An analog lemma would be: the system  $\{\mathbf{x} : \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, \dots, m, \mathbf{x} \in K\}$  has a feasible solution  $\mathbf{x}$  if and only if that  $-\sum_i^m y_i \mathbf{a}_i \in K^*$  and  $\mathbf{b}^T \mathbf{y} > 0$  has no feasible solution  $\mathbf{y}$ ?

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; \dots; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m$$

and

$$\mathcal{A}^T \mathbf{y} = \sum_i^m y_i \mathbf{a}_i.$$

Is the following a **alternative system pair**:

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T \mathbf{y} \in K^*, \quad \mathbf{b}^T \mathbf{y} = 1?$$

## An SDP Cone Example when “Alternative System” failed

$$K = \mathcal{S}_+^2.$$

$$\mathbf{a}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Problem:  $C := \{\mathcal{A}\mathbf{x} : \mathbf{x} \in K\}$  is not **closed** even when  $K$  is a closed and pointed convex cone.

## When Farkas' Lemma Holds for General Cones?

Let  $K$  be a closed and pointed convex cone in the rest of the course.

If there is  $\mathbf{y}$  such that  $-\mathcal{A}^T \mathbf{y} \in \text{int } K^*$ , then

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T \mathbf{y} \in K^*, \quad \mathbf{b}^T \mathbf{y} = 1$$

are an alternative system pair.

And if there is  $\mathbf{x}$  such that  $\mathcal{A}^T \mathbf{x} = \mathbf{0}$ ,  $\mathbf{x} \in \text{int } K$ , then

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in K, \quad \mathbf{c} \bullet \mathbf{x} = -1 (< 0)$$

and

$$\mathbf{c} - \mathcal{A}^T \mathbf{y} \in K^*$$

are an alternative system pair.



## Conic LP

$$\begin{aligned} (CLP) \quad & \text{minimize} \quad \mathbf{c} \bullet \mathbf{x} \\ & \text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, \dots, m, \mathbf{x} \in K, \end{aligned}$$

where  $K$  is a closed and pointed convex cone.

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  $K = SOC = \{\mathbf{x} : x_1 \geq \|\mathbf{x}_{2:n}\|_2\}$ .

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  $K = POC = \{\mathbf{x} : x_1 \geq \|\mathbf{x}_{2:n}\|_p\}$ .

**LP, SOCP, and SDP Examples Again**

$$\begin{array}{ll}\text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & (x_1; x_2; x_3) \geq \mathbf{0}.\end{array}$$

$$\begin{array}{ll}\text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & x_1 - \sqrt{x_2^2 + x_3^2} \geq 0.\end{array}$$

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

where

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

## Dual of Conic LP

The **dual problem** to

$$\begin{aligned} (CLP) \quad & \text{minimize} \quad \mathbf{c} \bullet \mathbf{x} \\ & \text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, \dots, m, \mathbf{x} \in K. \end{aligned}$$

is

$$\begin{aligned} (CLD) \quad & \text{maximize} \quad \mathbf{b}^T \mathbf{y} \\ & \text{subject to} \quad \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \mathbf{s} \in K^*, \end{aligned}$$

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

**Theorem 6** (*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).

## LP, SOCP, and SDP Example Duals

$$\begin{array}{ll}\text{maximize} & y \\ \text{subject to} & \mathbf{e} \cdot \mathbf{y} + \mathbf{s} = (2; 1; 1), \\ & (s_1; s_2; s_3) \geq \mathbf{0}.\end{array}$$

$$\begin{array}{ll}\text{maximize} & y \\ \text{subject to} & \mathbf{e} \cdot \mathbf{y} + \mathbf{s} = (2; 1; 1), \\ & s_1 - \sqrt{s_2^2 + s_3^2} \geq 0.\end{array}$$

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

The weak duality theorem implies that the dual can serve as an **optimality certificate**.