# Theory of Polyhedron and Duality

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

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

LY, Appendix B, Chapters 2.3-2.4, 4.1-4.2

## Basic and Basic Feasible Solution I

Consider the polyhedron set $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ where $A$ is a $m \times n$ matrix with $n \geq m$ and full row rank, 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 $\mathbf{x}_B$. By setting the variables, $\mathbf{x}_N$, of $\mathbf{x}$ corresponding to the remaining columns of $A$ equal to zero, we obtain a solution $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$. (Here, the index set $N$ represents the indices of the remaining columns of $A$)

Then $\mathbf{x}$ is said to be a basic solution with respect to the basic variable set $B$. The variables of $\mathbf{x}_B$ are called basic variables, those of $\mathbf{x}_N$ are called nonbasic variables, and $A_B$ is called a basis.

If a basic solution satisfies $\mathbf{x}_B \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution (BFS). BFS is an extreme or corner point of the polyhedron.

# Basic and Basic Feasible Solution II

Consider the polyhedron set $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$ where $A$ is a $m \times n$ matrix with $n \ge m$ and full row rank, select $m$ linearly independent columns, denoted by the variable index set $B$, from $A$. Solve

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

for the $m$-dimension vector $\mathbf{y}$.

Then $\mathbf{y}$ is called a basic solution with respect to the basis $A_B$ in polyhedron set $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$.

If a basic solution satisfies $A_N^T\mathbf{y} \le \mathbf{c}_N$, then $\mathbf{y}$ is called a basic feasible solution (BFS) of $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$, where the index set $N$ represents the indices of the remaining columns of $A$. BFS is an extreme or corner point of the polyhedron.

# Separating Hyperplane Theorem

The most important theorem about convex sets is the following Separating Hyperplane Theorem (Figure 1).

**Theorem 1** *(Separating Hyperplane Theorem) Let $C \subset \mathcal{E}$, where $\mathcal{E}$ is either $\mathcal{R}^n$ or $\mathcal{M}^n$, be a closed convex set and let $\mathbf{b}$ be a point exterior to $C$. Then there is a vector $\mathbf{a} \in \mathcal{E}$ such that*

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

*where $\mathbf{a}$ is the norm direction of the hyperplane.*

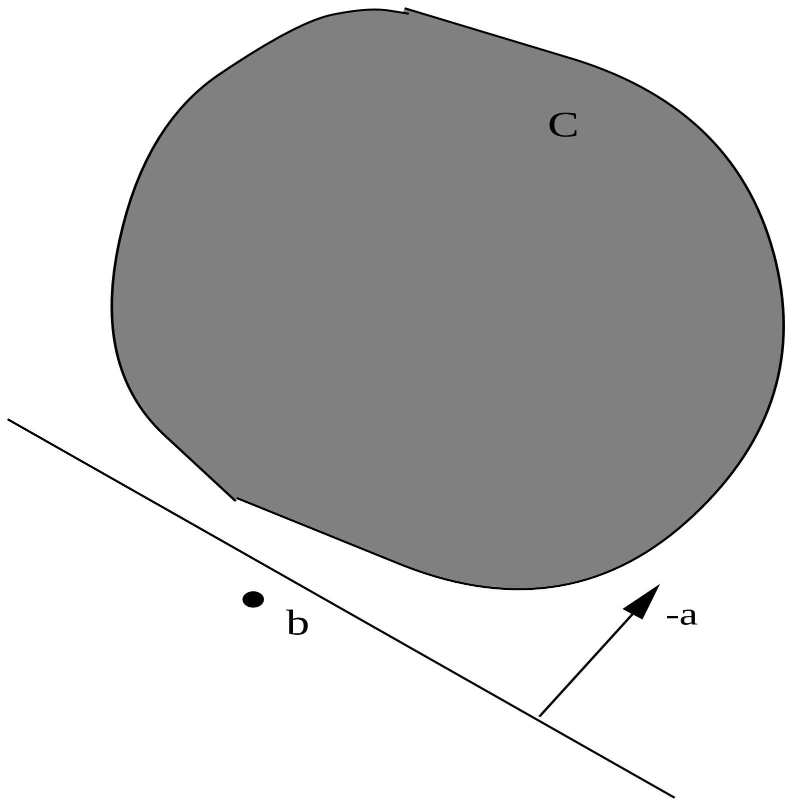


Figure 1: Illustration of the separating hyperplane theorem; an exterior point $\mathbf{b}$ is separated by a hyperplane from a convex set $C$.

## Examples

Let $C$ be a unit circle centered at point $(1;1)$. That is, $C = \{\mathbf{x} \in \mathcal{R}^2 : \ (x_1 - 1)^2 + (x_2 - 1)^2 \le 1\}$. 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

**Theorem 2** *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 the system $\{\mathbf{y} : \; A^T\mathbf{y} \leq \mathbf{0}, \; \mathbf{b}^T\mathbf{y} > 0\}$ has no feasible solution.*

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

Example

Let $A = (1, 1)$ and $b = -1$. Then $y = -1$ is an infeasibility certificate for $\{\mathbf{x} : \; A\mathbf{x} = b, \; \mathbf{x} \geq \mathbf{0}\}$.

## Alternative Systems

Farkas' lemma is also called the Alternative Theorem. That is, exactly one of the two systems

$$\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$$

and

$$\{\mathbf{y} : \; A^T\mathbf{y} \leq \mathbf{0}, \; \mathbf{b}^T\mathbf{y} > 0\}$$

is feasible.

## Geometric interpretation

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

$$\mathbf{b} \notin \{A\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}.$$

## 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,

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

since $\mathbf{x} \geq \mathbf{0}$ 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 (?), 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

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

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

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

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

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

# Farkas' Lemma Variant

**Theorem 3** *Let $A \in \mathcal{R}^{m \times n}$ and $\mathbf{c} \in \mathcal{R}^n$. Then the system $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$ has a solution $\mathbf{y}$ if and only if $A\mathbf{x} = \mathbf{0}$, $\mathbf{x} \ge \mathbf{0}$, $\mathbf{c}^T\mathbf{x} < 0$ has no feasible solution $\mathbf{x}$.*

Again, a vector $\mathbf{x} \ge \mathbf{0}$, with $A\mathbf{x} = \mathbf{0}$ and $\mathbf{c}^T\mathbf{x} < 0$, is called a infeasibility certificate for the system $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$.

Example

Let $A = (1; -1)$ and $\mathbf{c} = (1; -2)$. Then, $\mathbf{x} = (1; 1)$ is an infeasibility certificate for $\{y : \ A^T y \le \mathbf{c}\}$.

## Linear Programming and Its Dual

Consider the linear program (LP) in standard form, called the primal problem

$$
\begin{array}{lll}
(LP) & \text{minimize} & \mathbf{c}^T\mathbf{x} \\
 & \text{subject to} & A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge \mathbf{0},
\end{array}
$$

where $\mathbf{x} \in \mathcal{R}^n$.

The dual problem (LD) can be written as:

$$
\begin{array}{lll}
(LD) & \text{maximize} & \mathbf{b}^T\mathbf{y} \\
 & \text{subject to} & A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \ge \mathbf{0},
\end{array}
$$

where $\mathbf{y} \in \mathcal{R}^m$ and $\mathbf{s} \in \mathcal{R}^n$. The components of $\mathbf{s}$ are called dual slacks.

# Constructing the Dual

| | |
|---|---|
| Obj Coef Vector | RHS |
| RHS | Obj Coef Vector |
| $A$ | $A^T$ |
| Max Model | Min Model |
| $x_j \geq 0$ | $j$-th constraint $\geq$ |
| $x_j \leq 0$ | $j$-th constraint $\leq$ |
| $x_j$: free | $j$-th constraint $=$ |
| $i$-th constraint $\leq$ | $y_i \geq 0$ |
| $i$-th constraint $\geq$ | $y_i \leq 0$ |
| $i$-th constraint $=$ | $y_i$: free |

$$
Primal: \quad \begin{array}{llll} \text{maximize} & x_1 & +2x_2 & \\ \text{subject to} & x_1 & & \leq 1 \\ & & x_2 & \leq 1 \\ & x_1 & +x_2 & \leq 1.5 \\ & x_1, & x_2 & \geq 0. \end{array}
$$

$$
Dual: \quad \begin{array}{lllll} \text{minimize} & y_1 & +y_2 & +1.5y_3 & \\ \text{subject to} & y_1 & & +y_3 & \geq 1 \\ & & y_2 & +y_3 & \geq 2 \\ & y_1, & y_2, & y_3 & \geq 0. \end{array}
$$

# LP Duality Theories

**Theorem 4** *(Weak Duality Theorem) Let two feasible regions $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then,*

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y} \quad \textit{where} \quad \mathbf{x} \in \mathcal{F}_p,\ (\mathbf{y}, \mathbf{s}) \in \mathcal{F}_d.$$

The proof is simple:

$$\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} = \mathbf{c}^T\mathbf{x} - (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{x}^T\mathbf{s} \geq 0.$$

This theorem shows that a feasible solution to either problem yields a bound on the optimal value of the other problem. We call $(\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y})$ the duality gap. From this, we have the following important results.

**Theorem 5** *(Strong Duality Theorem) Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then $\mathbf{x}^*$ is optimal for LP if and only if the following conditions hold:*

**i)** $\mathbf{x}^* \in \mathcal{F}_p$

**ii)** *There exists* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$

**iii)** $\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$

Given $\mathcal{F}_p$ and $\mathcal{F}_d$ being non-empty, we would like to show that there exist $\mathbf{x}^* \in \mathcal{F}_p$ and $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$ such that $\mathbf{c}^T\mathbf{x}^* \leq \mathbf{b}^T\mathbf{y}^*$, or to show that

$$A\mathbf{x} = \mathbf{b},\ A^T\mathbf{y} \leq \mathbf{c},\ \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} \leq 0,\ \mathbf{x} \geq \mathbf{0}$$

is feasible.

# Proof of Strong Duality Theorem

Suppose not. Then from Farkas' lemma, we must have an infeasibility certificate $(\mathbf{x}', \tau, \mathbf{y}')$ such that

$$A\mathbf{x}' - \mathbf{b}\tau = \mathbf{0},\; A^T\mathbf{y}' - \mathbf{c}\tau \le \mathbf{0},\; (\mathbf{x}'; \tau) \ge \mathbf{0}$$

and

$$\mathbf{b}^T\mathbf{y}' - \mathbf{c}^T\mathbf{x}' = 1$$

If $\tau > 0$, then we have

$$0 \ge (-\mathbf{y}')^T(A\mathbf{x}' - \mathbf{b}\tau) + \mathbf{x}'^T(A^T\mathbf{y}' - \mathbf{c}\tau) = \tau(\mathbf{b}^T\mathbf{y}' - \mathbf{c}^T\mathbf{x}') = \tau$$

which gives a contradiction.

If $\tau = 0$, then the weak duality theorem also leads to a contradiction.

**Theorem 6** *(LP Duality Theorem) If LP and LD both have feasible solutions then both problems have optimal solutions and the optimal objective values of the objective functions are equal.*

*If one of LP or LD has no feasible solution, then the other is either unbounded or has no feasible solution. If one of LP or LD is unbounded then the other has no feasible solution.*

Above theorems show that if a pair of feasible solutions can be found to the primal and dual problems with equal objective values, then both of them are optimal. The converse is also true; there is no "gap".

## Optimality Conditions

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

which is a system of linear inequalities and equations. Now it is easy to verify whether or not a pair $(\mathbf{x},\mathbf{y},\mathbf{s})$ is optimal.

For feasible vectors $\mathbf{x}$ and $(\mathbf{y}, \mathbf{s})$, $\mathbf{x}^T\mathbf{s} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$ is called the complementarity gap.

Since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative, $\mathbf{x}^T\mathbf{s} = 0$ implies that $x_j s_j = 0$ for all $j = 1, \ldots, n$, and thus we can say $\mathbf{x}$ and $\mathbf{s}$ are complementary to each other.

$$
\begin{aligned}
X\mathbf{s} &= \mathbf{0} \\
A\mathbf{x} &= \mathbf{b} \\
-A^T\mathbf{y} - \mathbf{s} &= -\mathbf{c},
\end{aligned}
$$

where $X$ is the diagonal matrix of vector $\mathbf{x}$.

This system has total $(2n + m)$ unknowns and $(2n + m)$ equations including $n$ nonlinear equations.

**Theorem 7** *(Strict Complementarity Theorem) If both LP and LD have feasible solutions then both problems have a pair of strictly complementary solutions $x^* \geq 0$ and $s^* \geq 0$ meaning*

$$X^* s^* = 0 \quad \textit{and} \quad x^* + s^* > 0.$$

*Moreover, the supports*

$$P^* = \{j : \ x_j^* > 0\} \quad \textit{and} \quad Z^* = \{j : \ s_j^* > 0\}$$

*are invariant for all pairs of strictly complementary solutions.*

Given LP or LD, the pair of $P^*$ and $Z^*$ is called the (strict) complementarity partition. $\{x : \ A_{P^*} x_{P^*} = b, \ x_{P^*} \geq 0, \ x_{Z^*} = 0\}$ is called the primal optimal face and $\{y : \ c_{Z^*} - A_{Z^*}^T y \geq 0, \ c_{P^*} - A_{P^*}^T y = 0\}$ is called the dual optimal face.

## An Example

Consider a primal problem:

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

Then the dual problem is given by

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

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

## Proof Sketch for the Strict Complementarity Theorem

Let $z^*$ be the optimal objective value of LP and LD in the standard form. For any index $j$, consider the following auxiliary problem:

$$\begin{array}{lll} LP(j) & \text{minimize} & -x_j \\ & \text{subject to} & A\mathbf{x} = \mathbf{b},\ \mathbf{c}^T\mathbf{x} \le z^*,\ \mathbf{x} \ge \mathbf{0}. \end{array}$$

Clearly, any feasible solution of $LP(j)$ is an optimal solution of LP. If $LP(j)$ has a feasible solution with strictly negative objective value, we denote the solution by $\bar{\mathbf{x}}^j$ (that is, $\bar{\mathbf{x}}^j$ is an optimal solution for LP with $\bar{x}^j_j > 0$). Otherwise, the minimal value of $LP(j)$ must be zero.

Now consider the dual of $LP(j)$:

$$\begin{array}{lll} LD(j) & \text{maximize} & \mathbf{b}^T\mathbf{y} - z^*\tau \\ & \text{subject to} & A^T\mathbf{y} - \mathbf{c}\tau \le -\mathbf{e}_j,\ \tau \ge 0, \end{array}$$

where $\mathbf{e}_j$ is the $j$-th unit vector (whose $j$-th component is 1 and all other components are 0). Any optimal solution, $(\bar{\mathbf{y}}, \bar{\tau})$ for $LD(j)$ must have zero objective value:

$$\mathbf{b}^T\bar{\mathbf{y}} - z^*\bar{\tau} = 0.$$

Either $\bar{\tau} = 0$ (which case gives a homogeneous dual solution), or $\bar{\tau} > 0$ (which case gives an optimal dual solution by scaling), one can proceed to construct an optimal solution $(\bar{\mathbf{y}}^j, \bar{\mathbf{s}}^j)$ for LD with $\bar{s}_j^j > 0$.

Take the average of $\bar{\mathbf{x}}^j$ and $(\bar{\mathbf{y}}^j, \bar{\mathbf{s}}^j)$, respectively. Then, this pair will be a strictly complementary solution pair for LP and LD.