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

Convex polyhedral cones

$$C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$$

for some matrix A.

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

Theorem 1 *(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 $\mathbf{b}_1$,...,$\mathbf{b}_m$:*

$$C = \textit{cone}(\mathbf{b}_1, ..., \mathbf{b}_m) := \left\{ \sum_{i=1}^{m} \mathbf{b}_i y_i : \; y_i \geq 0, \; i = 1, ..., m \right\}.$$

Carathéodory's theorem

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

Theorem 2 *Let convex polyhedral cone* $C = \text{cone}(\mathbf{b}_1, ..., \mathbf{b}_m)$ *and* $x \in C$*. Then,* $x \in \text{cone}(\mathbf{b}_{i_1}, ..., \mathbf{b}_{i_d})$ *for some linearly independent vectors* $\mathbf{b}_{i_1}$*,...,*$\mathbf{b}_{i_d}$ *chosen from* $\mathbf{b}_1$*,...,*$\mathbf{b}_m$*.*

Example

$$C = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} y_1 + \begin{pmatrix} 2 \\ 1 \end{pmatrix} y_2 : \ y_1, y_2 \geq 0 \right\}$$

or

$$C = \left\{ \mathbf{x} \in \mathcal{R}^2 : \ \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \mathbf{x} \leq \mathbf{0} \right\}.$$

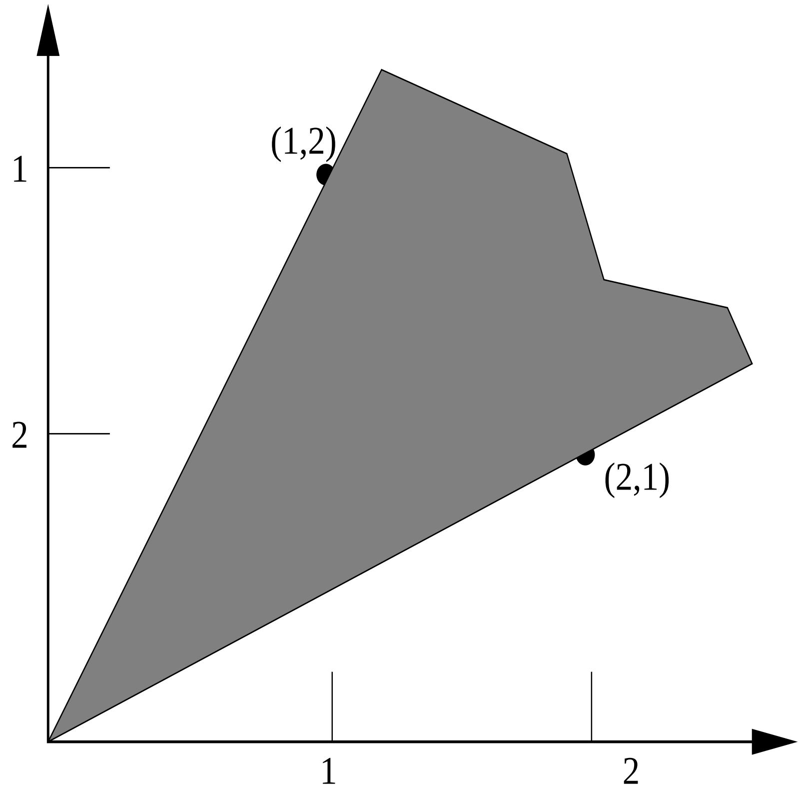

Figure 1: Representations of a polyhedral cone.

Point Cone

Pointed cone: a cone contains no straight line.

$$\{\mathbf{x} : \ A\mathbf{x} = \mathbf{0}\}?$$

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

$$\{\mathbf{x} : \ A\mathbf{x} \leq \mathbf{0}\}?$$

Convex polyhedra

A set is said to be a convex polyhedron if it can be written as

$$P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}.$$

A bounded polyhedron is called polytope.

Note that the feasible set $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \ge \mathbf{0}\}$ is polyhedral since it can be written as

$$\{\mathbf{x} : A\mathbf{x} \le \mathbf{b}, \quad , -A\mathbf{x} \le -\mathbf{b}, \quad \mathbf{x} \ge \mathbf{0}\}.$$

Also, if $\mathbf{b} = \mathbf{0}$, then P becomes a convex polyhedral cone.

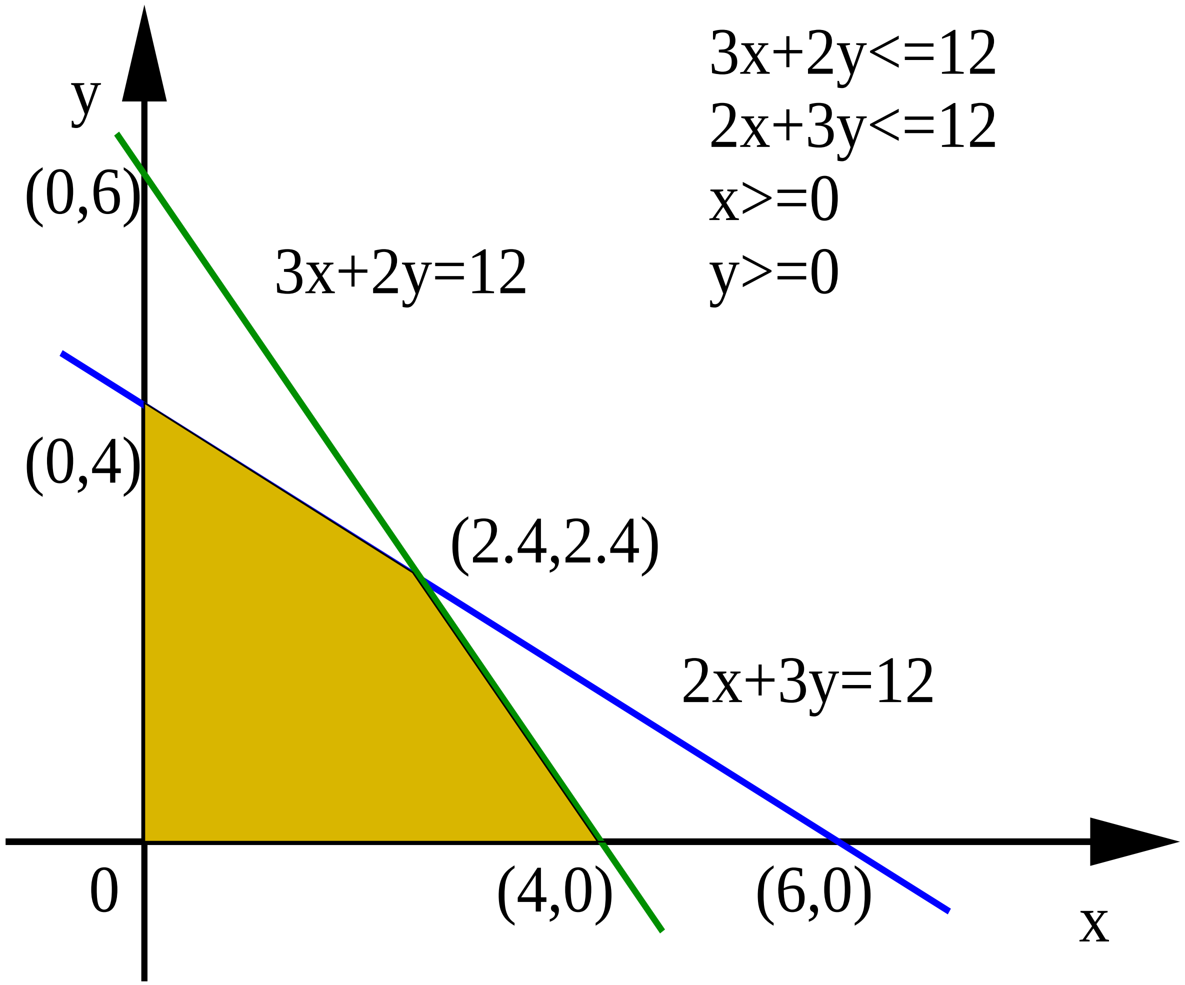

Figure 2: Polytope

Convex Combination

A convex hall, $CHull$, of a finite number of points $\mathbf{b}_1$,...,$\mathbf{b}_m$ is

$$CHull = \text{cone}(\mathbf{b}_1, ..., \mathbf{b}_m) :=$$

$$\left\{ \sum_{i=1}^{m} \mathbf{b}_i y_i : \sum_{i=1}^{m} y_i = 1,\ y_i \geq 0,\ i = 1, ..., m \right\}.$$

A point $\mathbf{x} \in CHull$ is said a convex combination of $\mathbf{b}_1$,...,$\mathbf{b}_m$.

This is the slice of the convex polyhedral cone generated from $\mathbf{b}_1$,...,$\mathbf{b}_m$ with a hyperplane generated from $\{\mathbf{y} : \sum_{i=1}^{m} y_i = 1\}$.

Separating hyperplane theorem

The most important theorem about the convex set is the following separating hyperplane theorem (Figure 3).

Theorem 3 *(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.

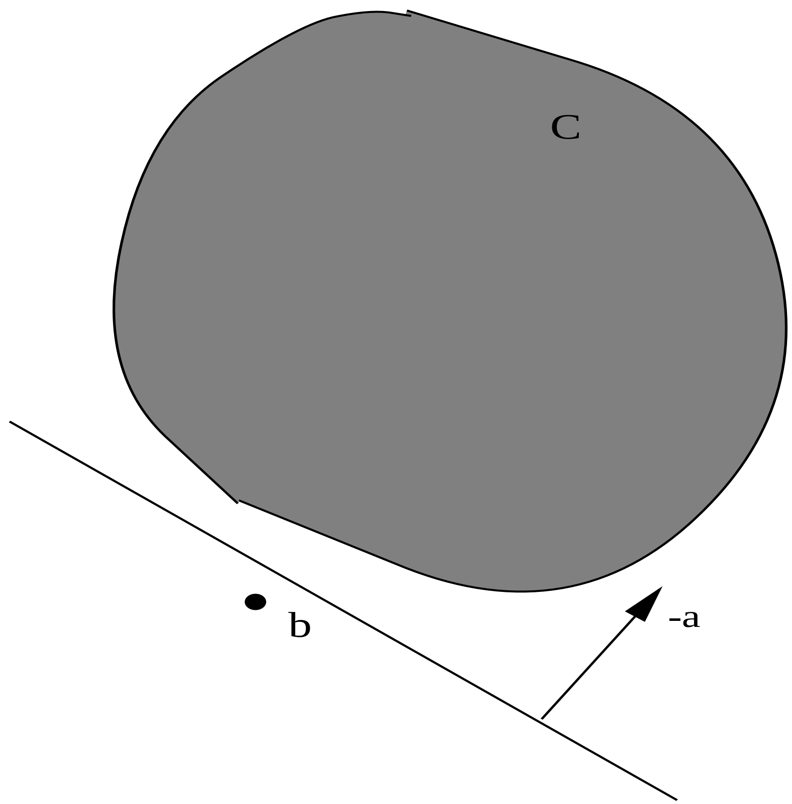

Figure 3: Illustration of the separating hyperplane theorem; an exterior point $\mathbf{b}$ is separated by a hyperplane from a convex set C.

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 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* $\{\mathbf{y} : \; A^T\mathbf{y} \leq \mathbf{0}, \; \mathbf{b}^T\mathbf{y} > 0$ *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 a 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, $\mathbf{b}\notin C:=\{A\mathbf{x}:\ \mathbf{x}\geq\mathbf{0}\}$. Then, 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

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

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 ∞ 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 5 *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 that $A\mathbf{x} = \mathbf{0}$, $\mathbf{x} \ge \mathbf{0}$, $\mathbf{c}^T\mathbf{x} < 0$ 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 $\{y : \; A^Ty \le \mathbf{c}\}$.

Dual of Linear Programming

Consider the linear program in standard form, called the primal problem,

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

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

The dual problem can be written as:

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

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

Rules to construct the dual

obj. coef. vector	right-hand-side
right-hand-side	obj. coef. vector
A	A^T
Max model	Min model
$x_j \geq 0$	jth constraint $\geq$
$x_j \leq 0$	jth constraint $\leq$
x_j free	jth constraint $=$
ith constraint $\leq$	$y_i \geq 0$
ith constraint $\geq$	$y_i \leq 0$
ith constraint $=$	y_i free

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

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

LP Duality Theories

Theorem 6 *(Weak duality theorem) Let feasible regions $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then,*

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

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

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

From this we have important results in the following.

Theorem 7 *(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 is* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$;

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

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

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

is feasible.

Proof of Strong Duality Theorem

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

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

and

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

If $\tau > 0$, then we have

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

which is a contradiction.

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

Theorem 8 *(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.

The above theorems show that if a pair of feasible solutions can be found to the primal and dual problems with equal objective values, then these are both 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 $\mathbf{x}$ 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 $j = 1, \ldots, n$, where we say $\mathbf{x}$ and $\mathbf{s}$ are complementary to each other.

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

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

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

Moreover, the supports

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

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 the primal problem:

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

The dual problem is

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

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