Convex Analyses and Conic 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 2.3-2.6, 3.1-3.3, 6.1-6.2

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

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

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

There is a construct proof of the theorem (page 26 of the text).

Basic and Basic Feasible Solution I

Now consider the feasible set $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ for given data $A \in R^{m \times n}$ and $\mathbf{b} \in R^m$. 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 A_B. If a basic solution $\mathbf{x}_B \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution, or BFS.

An equivalent statement of Carathéodory's theorem is:

Theorem 2 *If there is a feasible solution* $\mathbf{x}$ *to* $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$*, then there is a basic feasible solution to the system (page 26 of the text), and it is an extreme or corner point of the feasible set and vice versa.*

Corollary 1 *The set* $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ *is a polyhedral set.*

Basic and Basic Feasible Solution of the Inequality Form

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 to with respect to the basis A_B in polyhedron set $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$.

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

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 affine combination is also in H.

Hyperplanes dominate the entire theory of optimization. Let $\mathbf{a}$ be a nonzero n-dimensional (slope) vector, and let b be a real (intercept) 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, upper and lower 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 (page 510 of the text).

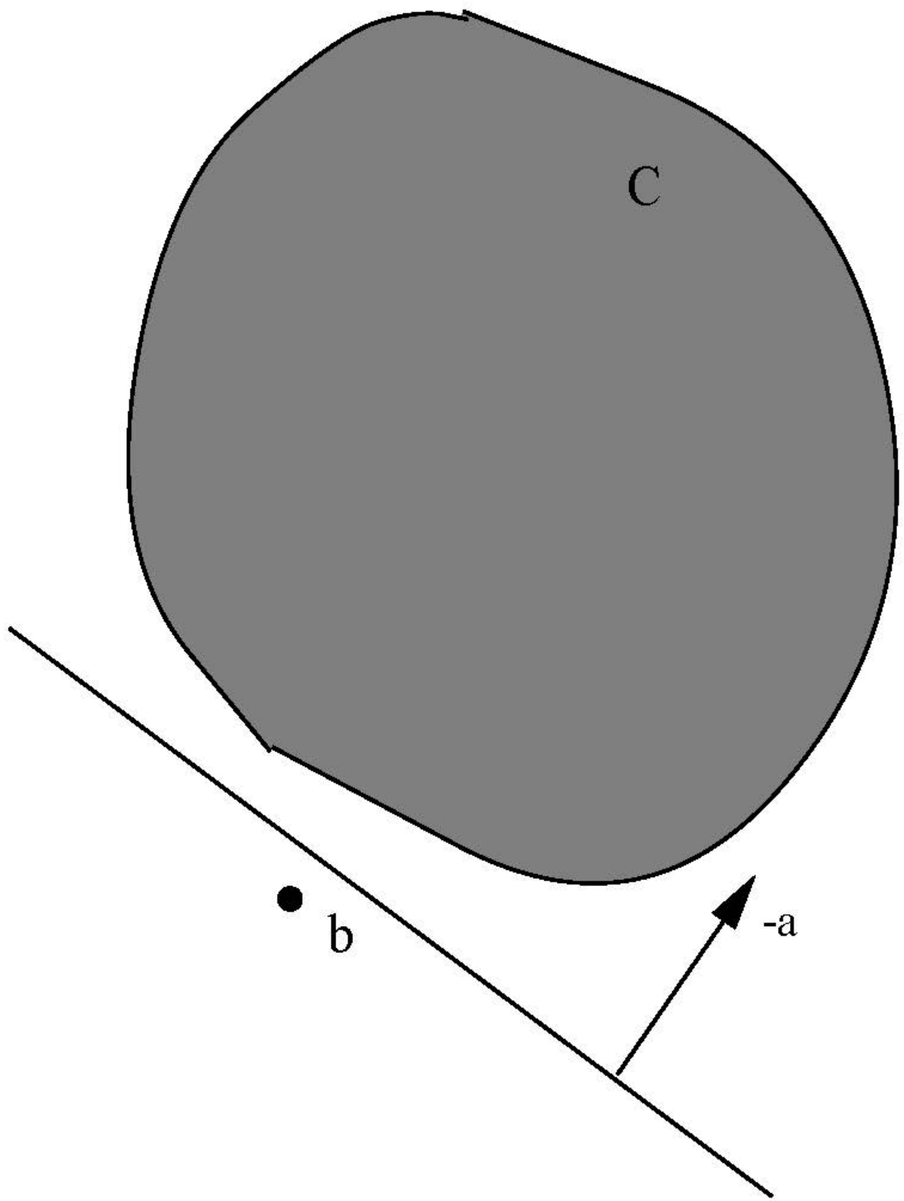
Theorem 3 *(Separating hyperplane theorem) Let C be a closed convex set in $\mathcal{R}^m$ and let $\mathbf{b}$ be a point exterior to C. Then there is a vector $\mathbf{y} \in \mathcal{R}^m$ such that*

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

Theorem 4 *(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*

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

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.


C
b
-a

Farkas' Lemma

The following results are Farkas' lemma and its variants.

Theorem 5 *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 its alternative system* $-A^T\mathbf{y} \geq \mathbf{0}$ *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 $\mathbf{a}_{.1}, ..., \mathbf{a}_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from $\text{cone}(\mathbf{a}_{.1}, ..., \mathbf{a}_{.n})$.

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,

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

from $\mathbf{x}\geq \mathbf{0}$ 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, or $\mathbf{b}\notin C:=\{A\mathbf{x}: \ \mathbf{x}\geq \mathbf{0}\}$. We now prove that its alternative system has a solution. We first prove

Lemma 1 $C=\{A\mathbf{x}: \ \mathbf{x}\geq \mathbf{0}\}$ *is a closed convex set.*

That is, any convergent sequence $\mathbf{b}^k\in C, \ k=1.2....$ has its limit point $\bar{\mathbf{b}}$ also in C. Let $\mathbf{b}^k=A\mathbf{x}^k, \ \mathbf{x}^k\geq \mathbf{0}$. Then by Carathéodory's theorem, we must have $\mathbf{b}^k=A_{B^k}\mathbf{x}_{B^k}, \ \mathbf{x}_{B^k}\geq \mathbf{0}$ where A_{B^k} is a basis of A. Therefore, $\mathbf{x}_{B^k}$, together with zero values for the nonbasic variables, is **bounded** for all k, so that it has sub-sequence, say indexed by $l=1,...$, where $\mathbf{x}^l=\mathbf{x}_{B^l}$ has a limit point $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}}\geq \mathbf{0}$. Consider this very sub-sequence $\mathbf{b}^l=A\mathbf{x}^l$ we must also have $\mathbf{b}^l\to\bar{\mathbf{b}}$. Then from

$$\|\bar{\mathbf{b}}-A\bar{\mathbf{x}}\|=\|\bar{\mathbf{b}}-\mathbf{b}^l+A\mathbf{x}^l-A\bar{\mathbf{x}}\|\leq\|\bar{\mathbf{b}}-\mathbf{b}^l\|+\|A\mathbf{x}^l-A\bar{\mathbf{x}}\|\leq\|\bar{\mathbf{b}}-\mathbf{b}^l\|+\|A\|\|\mathbf{x}^l-\bar{\mathbf{x}}\|$$

we must have $\bar{\mathbf{b}} = A\bar{\mathbf{x}}$, that is, $\bar{\mathbf{b}} \in C$; since otherwise the right-hand side of the above inequality is strictly greater than zero which is a contradiction.

Now 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

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

From $\mathbf{0} \in C$ 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 $\bar{x}_1 = \alpha > 0, \bar{x}_2 = ... = \bar{x}_n = 0$, 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 ∞ as $\alpha \to \infty$. 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 6 *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 Closed Convex Cones?

Sre the pair

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

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1(> 0)$$

alternative systems for a general closed convex cone K?

Here operator $\mathcal{A}\mathbf{x}$ and Adjoint-Operator $\mathcal{A}^T\mathbf{y}$ minimic matrix-vector production $A\mathbf{x}$ and its transpose operation $A^T\mathbf{y}$, where

$$\mathcal{A} = (\mathbf{a}_1; \mathbf{a}_2; ...; \mathbf{a}_m), \quad \mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}), \quad \text{and} \quad A^T\mathbf{y} = \sum_i y_i \mathbf{a}_i^T.$$

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m \text{ and } \mathcal{A}^T\mathbf{y} = \sum_i^m y_i \mathbf{a}_i.$$

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

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

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

When Farkas' Lemma Holds for General Cones?

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

If there is $\mathbf{y}$ such that $-\mathcal{A}^T\mathbf{y} \in \operatorname{int} K^*$, then $C := \{\mathcal{A}\mathbf{x} : \mathbf{x} \in K\}$ is a closed convex cone. Consequently,

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

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1(> 0)$$

are an alternative system pair.

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

Primal and Dual of Conic LP

Recall the pair of

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ (\mathcal{A}\mathbf{x} = \mathbf{b}), \ \mathbf{x} \in K;$$

and it dual problem

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

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

Cone K can be also a product of different cones, that is, $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; ...)$ where $\mathbf{x}_1 \in K_1$, $\mathbf{x}_2 \in K_2$,... and so on with linear constraints:

$$\mathcal{A}_1 \mathbf{x}_1 + \mathcal{A}_2 \mathbf{x}_2 + ... = \mathbf{b}.$$

LP, SOCP, and SDP Primal-Dual Examples

$$
\begin{array}{llll}
\text{min} & (2;\ 1;\ 1)^T\mathbf{x} & \text{max} & y \\
\text{s. t.} & \mathbf{e}^T\mathbf{x} = 1, & \text{s.t.} & \mathbf{e}\cdot y + \mathbf{s} = (2;\ 1;\ 1), \\
 & \mathbf{x} \geq \mathbf{0}. & & \mathbf{s} \geq \mathbf{0}.
\end{array}
$$

$$
\begin{array}{llll}
\text{min} & (2;\ 1;\ 1)^T\mathbf{x} & \text{max} & y \\
\text{s.t.} & \mathbf{e}^T\mathbf{x} = 1, & \text{s.t.} & \mathbf{e}\cdot y + \mathbf{s} = (2;\ 1;\ 1), \\
 & x_1 - \|\mathbf{x}_{-1}\| \geq 0. & & s_1 - \|\mathbf{s}_{-1}\| \geq 0.
\end{array}
$$

$$
\begin{array}{ll}
\text{min} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\
\text{s.t.} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\
& \mathbf{x} = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0},
\end{array}
$$

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

Rules to Construct the Dual of CLP

$$(CLP) \quad \text{minimize} \quad \sum_k \mathbf{c}_k \bullet \mathbf{x}_k \qquad (CLD) \quad \text{minimize} \quad \mathbf{b}^T\mathbf{y}$$

$$\text{subject to} \quad \sum_k \mathcal{A}_k \mathbf{x}_k = \mathbf{b}, \qquad \text{subject to} \quad \mathcal{A}_k^T \mathbf{y} + \mathbf{s}_k = \mathbf{c}_k,\ \forall k,$$

$$\mathbf{x}_k \in K_k,\ \forall k. \qquad \mathbf{s}_k \in K_k^*,\ \forall k.$$

obj. coef. vector	right-hand-side
right-hand-side	obj. coef. vector
$\mathcal{A}$	$\mathcal{A}^T$
Max model	Min model
$\mathbf{x}_k \in K$	kth block-constraint slack $\mathbf{s}_k \in K^*$
$\mathbf{x}_k$ "free"	kth block-constraint slack $\mathbf{s}_k = \mathbf{0}$
ith block-constraint slack $\mathbf{s}_i \in K$	$\mathbf{y}_i \in K^*$
ith block-constraint slack $\mathbf{s}_i = \mathbf{0}$	$\mathbf{y}_i$ "free"

The dual of the dual is primal!

CLP Duality Theorems

Theorem 7 *(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).*

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

Corollary 2 *Let* $\mathbf{x}^* \in \mathcal{F}_p$ *and* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$. *Then,* $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$ *implies that* $\mathbf{x}^*$ *is optimal for (CLP) and* $(\mathbf{y}^*, \mathbf{s}^*)$ *is optimal for (CLD).*

Is the reverse also true? That is, given $\mathbf{x}^*$ optimal for (CLP), then there is $(\mathbf{y}^*, \mathbf{s}^*)$ feasible for (CLD) and $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$?

This is called the Strong Duality Theorem.

"True" when $K = \mathcal{R}^n_+$, that is, the polyhedral cone case, but not true in general.

Proof of Strong Duality Theorem for LP

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.

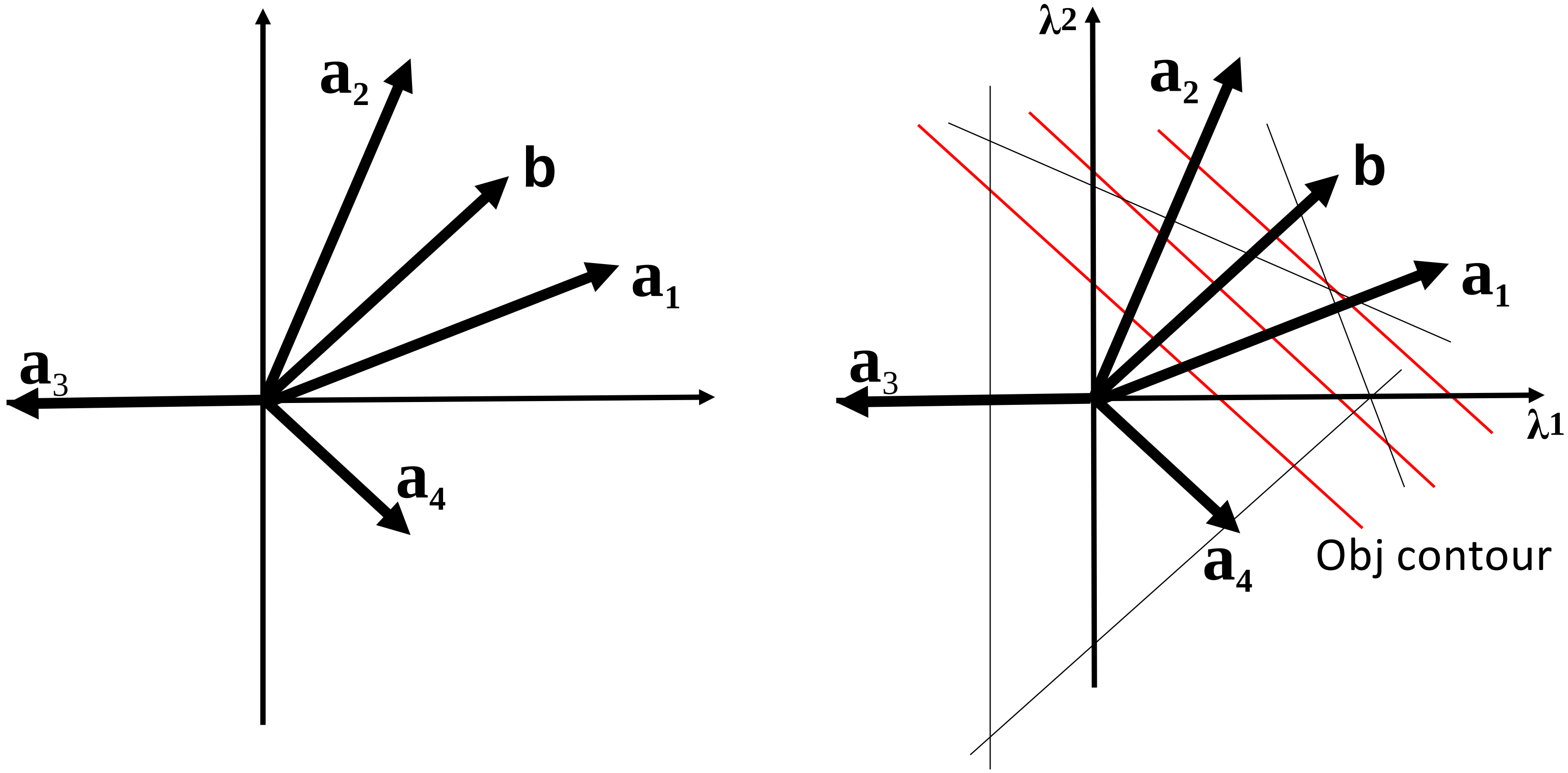
Geometric Interpretation

Consider

$$\begin{array}{lrrrrl} \text{minimize} & 18x_1 + & 12x_2 + & 2x_3 & + 6x_4 & \\ \text{subject to} & 3x_1 + & x_2 & - 2x_3 & + x_4 & = 2 \\ & x_1 + & 3x_2 & & - x_4 & = 2 \\ & \multicolumn{5}{l}{x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0,\ x_4 \geqslant 0.} \end{array}$$

and its dual

$$\begin{array}{lrrl} \text{maximize} & 2\lambda_1 + & 2\lambda_2 & \\ \text{subject to} & 3\lambda_1 + & \lambda_2 & \leqslant 18 \\ & \lambda_1 + & 3\lambda_2 & \leqslant 12 \\ & -2\lambda_1 & & \leqslant 2 \\ & \lambda_1 & - \lambda_2 & \leqslant 6. \end{array}$$



Each column $\mathbf{a}_j$ of the primal defines a constraint of the dual as a half-space whose boundary is orthogonal to that column vector and is located at a point determined by c_j. The dual objective is maximized at an extreme point of the dual feasible region. The active constraints at optimal solution correspond to an optimal basis of the primal. In the specific example, $\mathbf{b}$ is a conic combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. The weights in this combination are the x_i's in the optimal solution of the primal.

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."

The LP Primal and Dual Relations

Dual \ Primal	F-B	F-UB	IF
F-B	☺		
F-UB			☹
IF		☹	☹

$$
\begin{array}{llcl}
\min & -x_1 - x_2 & & \\
\text{s.t.} & x_1 - x_2 & = & 1 \\
 & -x_1 + x_2 & = & 1 \\
 & x_1, \quad x_2 & \geq & 0
\end{array}
$$

$$
\begin{array}{llcl}
\max & y_1 + y_2 & & \\
\text{s.t.} & y_1 - y_2 & \leq & -1 \\
 & -y_1 + y_2 & \leq & -1
\end{array}
$$

Optimality Conditions for LP

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

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.

Complementarity Gap

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.

If $\mathbf{x}^T\mathbf{s} = 0$, then we say $\mathbf{x}$ and $\mathbf{s}$ are complementary to each other.

Since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative, $\mathbf{x}^T\mathbf{s} = 0$ implies that $\mathbf{x}_. * \mathbf{s} = \mathbf{0}$ or $x_j s_j = 0$ for all $j = 1, \ldots, n$.

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

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* $\mathbf{x}^* \geq \mathbf{0}$ *and* $\mathbf{s}^* \geq \mathbf{0}$ *meaning*

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

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, solving which can be viewed as a binary classification problem for a given data $(A, \mathbf{b}, \mathbf{c})$.

$\{\mathbf{x} : \; A_{P^*}\mathbf{x}_{P^*} = \mathbf{b}, \; \mathbf{x}_{P^*} \geq \mathbf{0}, \; \mathbf{x}_{Z^*} = \mathbf{0}\}$ is called the primal optimal face, and
$\{y : \; \mathbf{c}_{Z^*} - A_{Z^*}^T\mathbf{y} \geq \mathbf{0}, \; \mathbf{c}_{P^*} - A_{P^*}^T\mathbf{y} = \mathbf{0}\}$ is called the dual optimal face.

An Example

Consider the primal problem:

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

The dual problem is

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

where $P^* = \{3\}$ and $Z^* = \{1, 2\}$.

Proof of the LP strict complementarity theorem

For any given index $1 \le j \le n$, consider

$$\begin{array}{rll} \bar{z}_j := & \text{minimize} & -x_j \\ & \text{subject to} & A\mathbf{x} = \mathbf{b}, \\ & & -\mathbf{c}^T\mathbf{x} \ge -z^*, \\ & & \mathbf{x} \ge \mathbf{0}; \end{array}$$

and its dual

$$\begin{array}{ll} \text{maximize} & \mathbf{b}^T\mathbf{y} - z^*\tau \\ \text{subject to} & A^T\mathbf{y} - \mathbf{c}\tau + \mathbf{s} = -\mathbf{e}_j, \\ & \mathbf{s} \ge \mathbf{0},\ \tau \ge 0. \end{array}$$

If $\bar{z}_j < 0$, then we have an optimal solution for (LP) such that $x_j^* > 0$. On the other hand, if $\bar{z}_j = 0$, from the LP strong duality theorem, we have a solution $(\mathbf{y}, \mathbf{s}, \tau)$ such that

$$\mathbf{b}^T\mathbf{y} - z^*\tau = 0,\ A^T\mathbf{y} - \mathbf{c}\tau + \mathbf{s} = -\mathbf{e}_j,\ (\mathbf{s}, \tau) \ge \mathbf{0}.$$

In this solution, if $\tau > 0$, then we have

$$\mathbf{b}^T(\mathbf{y}/\tau) - z^* = 0,\ A^T(\mathbf{y}/\tau) + (\mathbf{s} + \mathbf{e}_j)/\tau = \mathbf{c}$$

which is an optimal solution for the dual with slack $s_j^* > 0$ where $\mathbf{y}^* = \mathbf{y}/\tau, \mathbf{s}^* = (\mathbf{s} + \mathbf{e}_j)/\tau$. If $\tau = 0$, we have $\mathbf{b}^T\mathbf{y} = 0,\ A^T\mathbf{y} + \mathbf{s} + \mathbf{e}_j = 0,\ \mathbf{s}, \tau \geq \mathbf{0}$. Then for any optimal dual solution (y^*, s^*), $(y^* + y, s^* + s + \mathbf{e}_j)$ is also an optimal dual solution where the jth slack is strictly positive.

Thus, for each given $1 \leq j \leq n$, there is an optimal solution pair $(\mathbf{x}^j, \mathbf{s}^j)$ such that either $x_j^j > 0$ or $s_j^j > 0$. Let an optimal solution pair be

$$\mathbf{x}^* = \frac{1}{n}\sum_j \mathbf{x}^j \quad \text{and} \quad \mathbf{s}^* = \frac{1}{n}\sum_j \mathbf{s}^j.$$

Then it is a strict complementarity solution pair.

Let $(\mathbf{x}^1, \mathbf{s}^1)$ and $(\mathbf{x}^2, \mathbf{s}^2)$ be two strict complementarity solution pairs. Note that we still have

$$0 = (\mathbf{x}^1)^T\mathbf{s}^2 = (\mathbf{x}^2)^T\mathbf{s}^1$$

from the Strong Duality theorem. This indicates that they must have same strict complementarity partition, since, otherwise, we must have an j such that $x_j^1 > 0$ and $s_j^2 > 0$ or $(\mathbf{x}^1)^T\mathbf{s}^2 > 0$.

Strong Duality for General CLP?

The strong duality theorem may not hold for general convex cones:

$$\mathbf{c} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and

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

When Strong Duality Theorems Holds for CLP

Theorem 10 *The following statements hold for every pair of (CLP) and (CLD):*

i) *If (CLP) and (CLD) both are feasible, and furthermore one of them have an interior, then there is no duality gap between (CLP) and (CLD). However, one of the optimal solution may not be attainable.*

ii) *If (CLP) and (CLD) both are feasible and have interior, then, then both have attainable optimal solutions with no duality gap.*

iii) *If (CLP) or (CLD) is feasible and unbounded, then the other has no feasible solution.*

iv) *If (CLP) or (CLD) is infeasible, and furthermore the other is feasible and has an interior, then the other is unbounded.*

In case i), one of the optimal solution may not be attainable although no gap.

SDP Example with Zero-Duality Gap but not Attainable

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad b_1 = 2.$$

The primal has an interior, but the dual does not.

Proof of CLP Strong Duality Theorem under Nonempty Interior Cond.

i) Let $\mathcal{F}_p$ be feasible and have an interior, and let z^* be its infimum. Then we consider the alternative system pair

$$\mathcal{A}\mathbf{x} - \mathbf{b}\tau = \mathbf{0},\ \mathbf{c} \bullet \mathbf{x} - z^*\tau < 0,\ (\mathbf{x}, \tau) \in K \times R_+,$$

and

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ -\mathbf{b}^T\mathbf{y} + \kappa = -z^*,\ (\mathbf{s}, \kappa) \in K^* \times R_+.$$

But the former is infeasible, so that we have a solution for the latter. From the Weak Duality theorem, we must have $\kappa = 0$, that is, we have a solution $(\mathbf{y}, \mathbf{s})$ such that

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{b}^T\mathbf{y} = z^*,\ \mathbf{s} \in K^*.$$

ii) We only need to prove that there exist a solution $\mathbf{x} \in \mathcal{F}_p$ such that $\mathbf{c} \bullet \mathbf{x} = z^*$, that is, the infimum of (CLP) is attainable. But this is just the other side of the proof given that $\mathcal{F}_d$ is feasible and has an interior, and z^* is also the supremum of (CLD).

iii) The proof by contradiction follows the Weak Duality Theorem.

iv) Suppose $\mathcal{F}_d$ is empty and $\mathcal{F}_p$ is feasible and have an interior. Then, we have $\bar{\mathbf{x}} \in \mathrm{int}\, K$ and $\bar{\tau} > 0$ such that $\mathcal{A}\bar{\mathbf{x}} - \mathbf{b}\bar{\tau} = \mathbf{0},\ (\bar{\mathbf{x}}, \bar{\tau}) \in \mathrm{int}(K \times R_+)$. Then, for any z^*, we again consider the alternative system pair

$$\mathcal{A}\mathbf{x} - \mathbf{b}\tau = \mathbf{0},\ \mathbf{c} \bullet \mathbf{x} - z^*\tau < 0,\ (\mathbf{x}, \tau) \in K \times R_+,$$

and

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ -\mathbf{b}^T\mathbf{y} + s = -z^*,\ (\mathbf{s}, s) \in K^* \times R_+.$$

But the latter is infeasible, so that the formal has a feasible solution for any z^*. At such an solution, if $\tau > 0$, we have $\mathbf{c} \bullet (\mathbf{x}/\tau) < z^*$; if $\tau = 0$, we have $\hat{\mathbf{x}} + \alpha\mathbf{x}$, where $\hat{\mathbf{x}}$ is any feasible solution for (CLP), being feasible for (CLP) and its objective value goes to $-\infty$ as α goes to ∞.

The CLP Primal and Dual Relations

Dual \ Primal	F-B	F-UB	IF
F-B	☺		☹
F-UB			☹
IF	☹	☹	☹

$$
\begin{array}{llll}
\min & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \bullet X & & \\
\text{s.t.} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet X & = 0 \\
& \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X & = 2 \\
& X & \succeq \mathbf{0}
\end{array}
$$

$$
\begin{array}{llll}
\max & 2y_2 & & \\
\text{s.t.} & y_1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + S & = & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
& S & \succeq & \mathbf{0}
\end{array}
$$

The Dual is feasible and bounded, but Primal is infeasible.

Optimality and Complementarity Conditions for SDP

$$\begin{array}{rcl} C \bullet X - \mathbf{b}^T\mathbf{y} &=& 0 \\ \mathcal{A}X &=& \mathbf{b} \\ -\mathcal{A}^T\mathbf{y} - S &=& -C \\ X, S &\succeq& \mathbf{0}, \end{array}, \quad \text{or} \quad \begin{array}{rcl} XS &=& \mathbf{0} \\ \mathcal{A}X &=& \mathbf{b} \\ -\mathcal{A}^T y - S &=& -C \\ X, S &\succeq& \mathbf{0} \end{array}$$

Let X^* and S^* be optimal solutions with zero duality gap. Then

$$\text{rank}(X^*) + \text{rank}(S^*) \leq n.$$

Hint of the Proof: for any symmetric PSD matrix $P \in S^n$ with rank r, there is a factorization $P = V^TV$ where $V \in R^{r \times n}$ and columns of V are nonzero-vectors and orthogonal to each other.

When $\text{rank}(X^*) + \text{rank}(S^*) = n$, then they are strictly complemetary

SDP Strict Complementarity May Not exists

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \; K = \mathcal{S}^3_+.$$

The maximal solution rank of either the primal or dual is one.