Lecture 5. Theorems of Alternatives and Self-Dual Embedding
A system of linear equations may not have a solution.

It is well known that either

\[ Ax = c \]

has a solution, or

\[ A^T y = 0, \quad c^T y = -1 \]

has a solution (but not both nor neither).

Such type of statement is called *theorem of alternatives*.

There are in fact quite a few such theorems of alternatives, in the context of linear algebra, as we will discuss shortly.
A very general form of \textit{theorem of alternatives} is Hilbert’s Nullstellensatz.

Let $P_i(x)$ be a polynomial of $n$ variables $x_1, \ldots, x_n$ in the complex field $(x \in \mathbb{C}^n)$, $i = 1, \ldots, m$.

\textbf{Theorem 1} \textit{(Nullstellensatz)} Either

$$P_1(x) = 0, \ P_2(x) = 0, \ldots, P_m(x) = 0$$

has a solution $x \in \mathbb{C}^n$, or there exist polynomials $Q_1(x), \ldots, Q_m(x)$, such that

$$P_1(x)Q_1(x) + P_2(x)Q_2(x) + \cdots + P_m(x)Q_m(x) = -1$$

has a solution.

For more details, visit Terence Tao’s blog at:

For the system of linear inequalities, the most famous theorem of alternative is the following Farkas’ lemma (shown by Farkas around the end of the 19th century).

**Theorem 2 (Farkas)** Either

\[ Ax = b, \ x \geq 0 \]

has a solution, or

\[ A^T y \geq 0, \ b^T y < 0 \]

has a solution.

In fact, some more refined forms of the theorems of alternatives exist, in the style of the Farkas lemma. Some examples are shown in the next few slides.
Let $x$ and $y$ be two vectors. Denote $x \geq y$ to be ‘$x \geq y$ and $x \neq y$’; similarly for ‘$x \not\preceq y$’.

**Theorem 3 (Gordan)** Either

$$Ax > 0$$

has a solution, or

$$A^Ty = 0, \ y \geq 0$$

has a solution.

**Theorem 4 (Stiemke)** Either

$$Ax \geq 0$$

has a solution, or

$$A^Ty = 0, \ y > 0$$

has a solution.
**Theorem 5** *(Gale)* Either

\[ Ax \not\leq b \]

has a solution, or

\[ A^T y = 0, \quad b^T y = -1, \quad y \geq 0 \]

has a solution.

**Theorem 6** *(Tucker)* Suppose \( A \neq 0 \). Either

\[ Ax \not\geq 0, \quad Bx \geq 0, \quad Cx = 0 \]

has a solution, or

\[ A^T u + B^T v + C^T w = 0, \quad u > 0, \quad v \geq 0 \]

has a solution.
Theorem 7 (Motzkin) Suppose $A \neq 0$. Either

$$Ax > 0, \ Bx \geq 0, \ Cx = 0$$

has a solution, or

$$A^T u + B^T v + C^T w = 0, \ u \geq 0, \ v \geq 0$$

has a solution.
The most important theorem of alternatives is still the Farkas lemma. The Farkas lemma can be proven either by the simplex method for linear programming, or it can be proven by the separation theorem. Interestingly, the Farkas lemma can also be proven in an elementary manner, due to a recent brilliant expository article by Terence Tao.

Let us slightly rephrase the Farkas lemma:

Let \( P_i(x) = \sum_{j=1}^{n} p_{ij} x_j - r_i, \, i = 1, 2, \ldots, m. \) If the system \( P_i(x) \geq 0, \, i = 1, 2, \ldots, m, \) does not have a solution, then there are \( y_i \geq 0 \) such that \( \sum_{i=1}^{m} y_i P_i(x) = -1. \)

We prove the result by induction on \( n. \) When \( n = 1, \) we can scale \( P_i(x) \geq 0 \) possibly into three parts: \( x_1 - a_i \geq 0, \, i \in I_+; \, -x_1 + b_j \geq 0, \, j \in I_-; \, c_k \geq 0, \, k \in I_0. \)

If there is \( k \in I_0 \) with \( c_k < 0 \) then we may let \( y_k = -1/c_k \) and \( y_\ell = 0 \) for all \( \ell \neq k. \) Otherwise there is \( i \in I_+ \) and \( j \in I_- \) with \( a_i > b_j. \) We let \( y_i = y_j = 1/(a_i - b_j) \) and 0 elsewhere, yielding \( \sum_{i=1}^{m} y_i P_i(x) = -1. \)
Now suppose that the result is proven for the dimension (of \( x \)) no more than \( n \). Consider the case when we have \( n + 1 \) variables:
\[
\bar{x} = (x; x_{n+1}), \text{ where } x \in \mathbb{R}^n.
\]
Depending on the coefficients of \( x_{n+1} \) we scale the original inequalities to get the following equivalent system of inequalities:
\[
\begin{align*}
P_i(\bar{x}) &= x_{n+1} - Q_i(x) \geq 0, \quad i \in I_+; \\
P_j(\bar{x}) &= -x_{n+1} + Q_j(x) \geq 0, \quad j \in I_-; \\
P_k(\bar{x}) &= Q_k(x) \geq 0, \quad k \in I_0.
\end{align*}
\]
If
\[
\begin{align*}
-Q_i(x) + Q_j(x) &\geq 0, \quad i \in I_+ \text{ and } j \in I_-; \\
Q_k(x) &\geq 0, \quad k \in I_0
\end{align*}
\]
has a solution then the original system would have a solution too.
By the condition of the lemma, therefore, it cannot have a solution. Now using the induction hypothesis, there are $y_{ij}, y_k \geq 0$ such that

$$\sum_{i \in I_+: j \in I_-} y_{ij}(-Q_i(x) + Q_j(x)) + \sum_{k \in I_0} y_k Q_k(x) = -1.$$ 

Hence,

$$\sum_{i \in I_+} \left( \sum_{j \in I_-} y_{ij} \right) (x_{n+1} - Q_i(x)) + \sum_{j \in I_-} \left( \sum_{i \in I_+} y_{ij} \right) (-x_{n+1} + Q_j(x)) + \sum_{k \in I_0} y_k Q_k(x) = -1.$$

The Farkas lemma is thus proven by this simple induction argument.
In fact, the strict separation theorem of convex polytopes follows from the Farkas lemma.

The separation theorem that we are talking about here is:

If polytopes $P_1$ and $P_2$ do not intersect then there is an affine linear function $f(x) = \langle y, x \rangle + y_0$ such that

$$f(x) \geq 1 \text{ for } x \in P_1, \text{ and } f(x) \leq -1 \text{ for } x \in P_2.$$ 

Let the extreme points of $P_1$ be $\{p_1, ..., p_s\}$, and the extreme points of $P_2$ be $\{q_1, ..., q_t\}$.

Then “$f(x) \geq 1 \text{ for } x \in P_1, \text{ and } f(x) \leq -1 \text{ for } x \in P_2$” $\iff$

$$\langle y, p_i \rangle + y_0 \geq 1, \text{ for } i = 1, ..., s; \text{ and } \langle y, q_j \rangle + y_0 \leq -1, \text{ for } j = 1, ..., t.$$
Let us now use a contradiction argument. Suppose the above does not have a solution in $y$ and $y_0$. Then the Farkas lemma asserts that there are $u_i \geq 0$ ($i = 1, ..., s$) and $v_j \geq 0$ ($j = 1, ..., t$), such that

$$\sum_{i=1}^{s} u_i \left( \langle y, p_i \rangle + y_0 - 1 \right) + \sum_{j=1}^{t} v_j \left( -\langle y, q_j \rangle - y_0 - 1 \right) = -1.$$ 

This implies that

$$\sum_{i=1}^{s} u_i p_i - \sum_{j=1}^{t} v_j q_j = 0, \quad \sum_{i=1}^{s} u_i - \sum_{j=1}^{t} v_j = 0, \quad \text{and} \quad \sum_{i=1}^{s} u_i + \sum_{j=1}^{t} v_j = 1.$$ 

Therefore $\sum_{i=1}^{s} u_i = \sum_{j=1}^{t} v_j = 1/2$. Moreover,

$$P_1 \ni 2 \sum_{i=1}^{s} u_i p_i = 2 \sum_{j=1}^{t} v_j q_j \in P_2,$$

which contradicts to the assumption that $P_1 \cap P_2 = \emptyset$.

The same arguments will work if ‘polytopes’ are replaced by ‘polydra’.

Shuzhong Zhang
Let us return to conic optimization models.

To study the geometric structure of the problem, introduce the linear subspace

\[ \mathcal{L} := \{x \mid Ax = 0\}, \]  
and its orthogonal complement \( \mathcal{L}^\perp = \{A^T y \mid \exists y \in \mathbb{R}^m\} \).

Let \( a = A^T (AA^T)^{-1} b \), and so \( a + \mathcal{L} = \{x \mid Ax = b, \ x \geq 0\} \). In this notation, the generic conic optimization problem can be written as

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad x \in (a + \mathcal{L}) \cap \mathcal{K}.
\end{align*}
\]  

(1)
Geometrically, the conic optimization is nothing but linear optimization over the intersection of an affine linear subspace $a + \mathcal{L}$ and a given closed convex cone $\mathcal{K}$. Such problems can be either feasible or infeasible; having unbounded optimal value or finite optimal value; attainable optimal solution exists or not. We shall see first that the situation for a generic conic optimization gets more complicated than for linear programming.
Let us consider the case of infeasibility.

It is well known that for linear programming the state of being ‘infeasibility’ is rather stable, in the sense that if the data \((a, c)\) corresponds to an infeasible problem instance, then a small perturbation on \((a, c)\) will remain the infeasibility status.

Why?

Consider the following equivalence due to the Farkas lemma:

\[
\{x \in \mathbb{R}^n \mid Ax = b, \ x \in \mathbb{R}^n_+\} = \emptyset \iff \{y \in \mathbb{R}^m \mid A^T y \in \mathbb{R}^n_+, \ b^T y < 0\} \neq \emptyset.
\]
For conic optimization, however, this no longer holds. Consider Example 1

\[
\begin{align*}
\text{minimize} & \quad x_{22} \\
\text{subject to} & \quad x_{11} = b_1, x_{12} = b_2 \\
& \begin{bmatrix} x_{11}, & x_{12}, \\ x_{21}, & x_{22} \end{bmatrix} \succeq 0.
\end{align*}
\]

For \((b_1, b_2) = (0, 1)\) the problem is infeasible. However, for \((b_1, b_2) = (\epsilon, 1)\) where \(\epsilon\) is any positive number, the problem is always feasible.
Definition 1  We call the conic optimization problem:

weakly infeasible  $\iff (a + \mathcal{L}) \cap \mathcal{K} = \emptyset$ but $\text{dist}(a + \mathcal{L}, \mathcal{K}) = 0$;

strongly infeasible  $\iff \text{dist}(a + \mathcal{L}, \mathcal{K}) > 0$;

weakly feasible  $\iff (a + \mathcal{L}) \cap \mathcal{K} \neq \emptyset$ but $(a + \mathcal{L}) \cap \text{int} \mathcal{K} = \emptyset$;

strongly feasible  $\iff (a + \mathcal{L}) \cap \text{int} \mathcal{K} \neq \emptyset$. 

Shuzhong Zhang
Theorem 8 (Alternatives of strong infeasibility).

\{x \in \mathbb{R}^n \mid Ax = b, x \in \mathcal{K}\} \text{ is not strongly infeasible } \iff \{y \in \mathbb{R}^m \mid A^T y \in \mathcal{K}^*, b^T y < 0\} = \emptyset.

Proof. Note that for \(s \in \mathcal{L}^\perp\), i.e. \(s = A^T y\) for some \(y \in \mathbb{R}^m\), we have

\[a^T s = a^T A^T y = b^T y.\]

Observe the following chain of equivalent statements:

\[\{y \in \mathbb{R}^m \mid A^T y \in \mathcal{K}^*, b^T y < 0\} = \emptyset \iff \text{“for all } s \in \mathcal{L}^\perp \cap \mathcal{K}^* \text{ it must follow } a^T s \geq 0\} \iff a \in (\mathcal{L}^\perp \cap \mathcal{K}^*)^* = \text{cl} (\mathcal{K} + \mathcal{L}) \iff \exists x_i \in \mathcal{K}, \bar{x}_i \in \mathcal{L} \text{ such that } \|a - (x_i + \bar{x}_i)\| = \|(a - \bar{x}_i) - x_i\| \to 0 \iff (a + \mathcal{L}) \cap \mathcal{K} \text{ is not strongly infeasible. The theorem is proven.} \]
Definition 2 We call

- $d_P$ an primal improving direction if $d_P \in \mathcal{L} \cap \mathcal{K}$, and $c^T d_P < 0$;
- $d_P^i$ an primal improving direction sequence if
  $$d_P^i \in \mathcal{K}, \text{dist}(d_P^i, \mathcal{L}) \to 0$$
  and $\limsup_i c^T d_P^i < 0$;
- $d_D$ a dual improve direction if $d_D \in \mathcal{L}^\perp \cap \mathcal{K}^*$ and $a^T d_D < 0$;
- $d_D^i$ a dual improve direction sequence if
  $$d_D^i \in \mathcal{K}^*, \text{dist}(d_D^i, \mathcal{L}^\perp) \to 0$$
  and $\limsup_i a^T d_D^i < 0$. 
If a problem is not strongly infeasible, then it can be either feasible or weakly infeasible. Next question is: can we further characterize the feasibility of a conic problem? Naturally if it is possible, then by exclusion we will also have characterized the weak feasibility.

**Theorem 9** *(Alternatives of primal feasibility).*

\[
\{ x \in \mathbb{R}^n \mid Ax = b, x \in \mathcal{K}\} \text{ is infeasible if and only if there is a dual improving direction sequence.}
\]
Proof. Let

$$\tilde{L} = \text{span}(\mathcal{H}(\{x \mid Ax = b\})) = \left\{ \begin{pmatrix} x_0 \\ x \end{pmatrix} \mid -x_0b + Ax = 0 \right\}$$

and $\tilde{\mathcal{K}} = \mathbb{R}_+ \times \mathcal{K}$. We have

$$\tilde{L}^\perp = \left\{ \begin{pmatrix} s_0 \\ s \end{pmatrix} \mid \exists y \in \mathbb{R}^m, s_0 = -b^T y, s = A^T y \right\} \quad \text{and} \quad \tilde{\mathcal{K}}^* = \mathbb{R}_+ \times \mathcal{K}^*.$$

Let $\bar{c} = \begin{pmatrix} -1 \\ 0_n \end{pmatrix}$. 
Theorem 8 states that the primal problem is strongly infeasible is equivalent to the existence of a dual improve direction. To state it in the dual way, Theorem 8 can also be written as

\[(\bar{c} + \bar{L}^\perp) \cap \bar{K}^* \text{ is strongly infeasible } \iff \exists \bar{x} \in \bar{L} \cap \bar{K}, \bar{c}^T \bar{x} < 0. \quad (2)\]

It is easy to see that \(\{x \mid Ax = b, x \in \mathcal{K}\}\) is feasible \(\iff\)
\[\exists \bar{x} \in \bar{L} \cap \bar{K}, \bar{c}^T \bar{x} < 0 \iff \text{(by (2)) } (\bar{c} + \bar{L}^\perp) \cap \bar{K}^* \text{ is strongly infeasible.}\]

Shuzhong Zhang
Therefore, \( \{ x \mid Ax = b, \ x \in \mathcal{K} \} \) is infeasible \( \iff \) \( (\bar{c} + \bar{L}^\perp) \cap \bar{\mathcal{K}}^* \) is not strongly infeasible \( \iff \exists s^i_0 \geq 0, s^i \in \mathcal{K}^* \) such that

\[
\text{dist} \left( \begin{pmatrix} s^i_0 \\ s^i \end{pmatrix}, \bar{c} + \bar{L}^\perp \right) \to 0 \iff s^i \in \mathcal{K}^*, \ \text{dist} (s^i, L^\perp) \to 0 \text{ and } \limsup_i a^T s^i < 0 \iff \exists \text{ a dual improve direction sequence.}
\]

\[\square\]
As a consequence, by exclusion, this implies the following result:

**Corollary 1** \( \{ x \in \mathbb{R}^n \mid Ax = b, \ x \in K \} \) is weakly infeasible \( \iff \) there is a dual improving direction sequence but no dual improving direction.

So far we have been concerned with characterizing infeasible conic programs. How about feasible ones? Recall that we have two types of feasibilities: the strong feasibility (also known as to satisfy the Slater condition), and the weak feasibility, which means that it is feasible but the Slater condition is not satisfied. It is helpful to derive an equivalent condition for the Slater condition, using the dual information.
Lemma 1 \((a + \mathcal{L}) \cap \interior \mathcal{K} \neq \emptyset\) if and only if for any \(d \neq 0\) and \(d \in \mathcal{L}^\perp \cap \mathcal{K}^*\) it must follow \(a^T d > 0\).

Proof. Note that \(x \in \interior \mathcal{K} \iff \forall 0 \neq y \in \mathcal{K}^*\) it follows \(x^T y > 0\).

‘\(\Longrightarrow\)’: Let \(x \in (a + \mathcal{L}) \cap \interior \mathcal{K}\). In particular, write \(x = a + z\) with \(z \in \mathcal{L}\). Take any \(d \neq 0\) and \(d \in \mathcal{L}^\perp \cap \mathcal{K}^*\):

\[
0 < x^T d = (a + z)^T d = a^T d.
\]

‘\(\Longleftarrow\)’: If \((a + \mathcal{L}) \cap \interior \mathcal{K} = \emptyset\), then \(a + \mathcal{L}\) and \(\interior \mathcal{K}\) can be separated by a hyperplane. That is, there is \(s \neq 0\) such that

\[
s^T x \geq 0 \text{ for all } x \in \mathcal{K} \text{ and } s^T (a + y) \leq 0 \text{ for all } y \in \mathcal{L}.
\]

The last inequality means that \(s^T a \leq 0\) and \(s^T y = 0\) for all \(y \in \mathcal{L}\). Thus, we have found a non-zero vector \(s \in \mathcal{L}^\perp \cap \mathcal{K}^*\) with \(a^T s \leq 0\), which contradicts the assumption. \(\square\)
**Theorem 10** If the primal conic program satisfies the Slater condition, and its dual is feasible, then the dual has a non-empty and compact optimal solution set.

**Proof.** Clearly, the objective value of the dual problem is bounded below by the weak duality theorem. The only possibility for the dual problem not to possess an optimal solution is that the ‘sup’ is non-attainable; i.e. there is an infinite sequence

\[ s^i = c - A^T y^i \in (a + \mathcal{L}^\perp) \cap \mathcal{K}^* \]

such that

\[ a^T s^i = a^T c - b^T y^i \to a^T c - \text{optimal value of the dual problem}, \]

and \( \|s^i\| \to \infty \). Without loss of generality, suppose that \( s^i/\|s^i\| \) converges to \( d \). Then we have

\[ d \in \mathcal{L}^\perp \cap \mathcal{K}^* \text{ and } a^T d = 0. \]

This contradicts with the primal Slater condition, in light of Lemma 1. Hence the optimal solutions must be bounded. \( \square \)
Therefore, if a pair of primal-dual conic programs satisfy the Slater condition, then attainable optimal solutions exist for both problems, according to Theorem 10. The only remaining issue in this regard is: how about the strong duality theorem? In other words, do these optimal solutions satisfy the complementarity slackness conditions? The answer is yes.

**Theorem 11** If the primal conic program and its dual conic program both satisfy the Slater condition, then the optimal solution sets for both problems are non-empty and compact. Moreover, the optimal solutions are complementary to each other with zero duality gap.
Proof. The first part of the theorem is simply Theorem 10, stated both for the primal and for the dual parts separately.

We now focus on the second part. Let \( x^* \) and \( s^* \) be optimal solutions for the primal and the dual conic optimization problems respectively. Write the primal problem in the parametric form:

\[
(P(a)) \quad \text{minimize} \quad c^T x \\
\text{subject to} \quad x \in a + \mathcal{L} \\
x \in \mathcal{K}.
\]

Let the optimal value of \( (P(a)) \), as a function of the vector \( a \), be \( v(a) \). Clearly \( v(a) \) is a convex function in its domain, \( \mathcal{D} \), whenever \( (P(a)) \) is feasible. Obviously, \( \mathcal{D} \supseteq \mathcal{K} \). Also we see that \( v(0) = 0 \) due to Theorem 8.
By the Slater condition we know that $a \in \text{int } \mathcal{D}$. Let $g$ be a subgradient for $v$ at $a$; that is

$$v(z) \geq v(a) + g^T(z - a)$$

(3)

for all $z \in \mathcal{D}$.

Observe that for any $d \in \mathcal{L}$ we have $a + \mathcal{L} = a + d + \mathcal{L}$. Thus, $v(a + d) = v(a)$ for any $d \in \mathcal{L}$. This implies $g \in \mathcal{L}^\perp$. Take any $u \in \mathcal{K}$. Since $x^* + u$ is a feasible solution for $(P(a + u))$, it follows that

$$v(a + u) \leq c^T(x^* + u) = v(a) + c^Tu.$$ 

Together with the subgradient inequality (3) this yields

$$(c - g)^Tu \geq 0$$

for any $u \in \mathcal{K}$; hence $c - g \in \mathcal{K}^*$. Thus, $c - g$ is a dual feasible solution, and so

$$a^Ts^* \leq a^T(c - g).$$
Taking \( z = 0 \), the subgradient inequality gives us

\[
0 = v(0) \geq v(a) + g^T(0 - a) = c^T x^* - g^T a. \tag{4}
\]

Combining with (4) we have

\[
a^T c - c^T x^* - a^T s^* \geq 0.
\]

Hence,

\[
0 = (x^* - a)^T (s^* - c) = (x^*)^T s^* + a^T c - c^T x^* - a^T s^* \geq (x^*)^T s^* \geq 0.
\]

Thus, \((x^*)^T s^* = 0\), and the complementarity slackness condition is satisfied. \(\Box\)
So far we have discussed the relationship between a duality pair of conic optimization problems

\[(P) \quad \text{minimize} \quad c^T x \]
\[\text{subject to} \quad Ax = b \]
\[x \in \mathcal{K}, \]

and

\[(D) \quad \text{maximize} \quad b^T y \]
\[\text{subject to} \quad A^Ty + s = c \]
\[s \in \mathcal{K}^*.\]
In case $\mathcal{K} = \mathcal{K}^* = \mathbb{R}^n_+$ — the linear programming case — the following facts are well known:

- If $(P)$ and $(D)$ both are feasible, then both problems are solvable, with attainable optimal solutions;

- It is possible that $(P)$ is feasible and unbounded in its objective. In that case, $(D)$ is infeasible.

- Symmetrically, if $(D)$ is feasible and unbounded, then $(P)$ is infeasible.

- It is possible that both $(P)$ and $(D)$ are infeasible.
This can be summarized by the following table, where the row is the status of the primal problem and the column represents the status of the dual problem:

<table>
<thead>
<tr>
<th></th>
<th>Feasible</th>
<th>Unbounded</th>
<th>Infeasible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasible</td>
<td>((P), (D)) solvable</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Unbounded</td>
<td>×</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Infeasible</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>
For general conic programs, assuming \((D)\) is feasible, then the relationship with \((P)\) is as follows:

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>s. f.</td>
<td>(\exists) bd opt sol set (Theorem 10)</td>
</tr>
<tr>
<td>w. f.</td>
<td>no improving direction sequence (Theorem 9)</td>
</tr>
<tr>
<td>w. inf.</td>
<td>(\exists) improving dir seq but no improving dir (Corollary 1)</td>
</tr>
<tr>
<td>strongly infeasible</td>
<td>(\exists) improving direction (Theorem 8)</td>
</tr>
</tbody>
</table>
Example 2

\[ \text{minimize} \quad x_{12} + x_{21} \]
\[ \text{subject to} \quad x_{12} + x_{21} + x_{33} = 1, \quad x_{22} = 0 \]
\[
\begin{bmatrix}
  x_{11}, & x_{12}, & x_{13} \\
  x_{21}, & x_{22}, & x_{23} \\
  x_{31}, & x_{32}, & x_{33}
\end{bmatrix} \succeq 0,
\]

with the dual

\[ \text{maximize} \quad y_1 \]
\[ \text{subject to} \quad \begin{bmatrix}
  0, & 1 - y_1, & 0 \\
  1 - y_1, & -y_2, & 0 \\
  0, & 0, & -y_1
\end{bmatrix} \succeq 0. \]

The primal problem is weakly feasible, while the dual is weakly infeasible.

Shuzhong Zhang
Example 3

\[ \begin{array}{c}
\text{minimize} \\
\text{subject to }
\end{array} \]
\[ x_{12} + x_{21} \]
\[ x_{12} + x_{21} + x_{23} + x_{32} = 1 \]
\[ x_{22} = 0 \]
\[ \begin{bmatrix}
  x_{11}, & x_{12}, & x_{13} \\
  x_{21}, & x_{22}, & x_{23} \\
  x_{31}, & x_{32}, & x_{33}
\end{bmatrix} \succeq 0, \]

with the dual

\[ \begin{array}{c}
\text{maximize} \\
\text{subject to }
\end{array} \]
\[ y_1 \]
\[ \begin{bmatrix}
  0, & 1 - y_1, & 0 \\
  1 - y_1, & -y_2, & -y_1 \\
  0, & -y_1, & 0
\end{bmatrix} \succeq 0. \]

Both the primal and the dual are weakly infeasible.
Example 4

\[\begin{array}{c}
\text{minimize} & x_{12} + x_{21} - x_{33} \\
\text{subject to} & x_{12} + x_{21} + x_{32} = 1, x_{22} = 0 \\
& \begin{bmatrix}
  x_{11}, & x_{12}, & x_{13} \\
  x_{21}, & x_{22}, & x_{23} \\
  x_{31}, & x_{32}, & x_{33}
\end{bmatrix} \succeq 0,
\end{array}\]

with the dual

\[\begin{array}{c}
\text{maximize} & y_1 \\
\text{subject to} & \begin{bmatrix}
  0, & 1 - y_1, & 0 \\
  1 - y_1, & -y_2, & -y_1 \\
  0, & -y_1, & -1
\end{bmatrix} \succeq 0.
\end{array}\]

In this case, the primal is weakly infeasible, while the dual is strongly infeasible.
Example 5

\[
\begin{align*}
\text{minimize} \quad & x_{12} + x_{21} - x_{33} \\
\text{subject to} \quad & x_{12} + x_{21} + x_{23} + x_{32} = 1, \ x_{22} = -1 \\
& \begin{bmatrix}
  x_{11}, & x_{12}, & x_{13} \\
  x_{21}, & x_{22}, & x_{23} \\
  x_{31}, & x_{32}, & x_{33}
\end{bmatrix} \succeq 0,
\end{align*}
\]

with the dual

\[
\begin{align*}
\text{maximize} \quad & y_1 - y_2 \\
\text{subject to} \quad & \begin{bmatrix}
  0, & 1-y_1, & 0 \\
  1-y_1, & -y_2, & -y_1 \\
  0, & -y_1, & -1
\end{bmatrix} \succeq 0.
\end{align*}
\]

In this case, both the primal and the dual are strongly infeasible.
Consider the standard conic optimization problems \((P)\) and \((D)\).
Take any \(x^0 \in \text{int } K\), \(s^0 \in \text{int } K^*\), and \(y^0 \in \mathbb{R}^m\). Moreover, define
\[
\begin{align*}
  r_p &= b - Ax^0, \\
  r_d &= s^0 - c + A^T y^0, \\
  r_g &= 1 + c^T x^0 - b^T y^0.
\end{align*}
\]
The following model is self-dual:
\[
\begin{align*}
  (SD) \quad \min \quad & \beta \theta \\
  \text{s.t.} \quad & Ax - b \tau + r_p \theta = 0 \\
  & -A^T y + c \tau + r_d \theta - s = 0 \\
  & b^T y - c^T x + r_g \theta - \kappa = 0 \\
  & -r_p^T y - r_d^T x - r_g \tau = -\beta \\
  & x \in K, \quad \tau \geq 0, \quad s \in K^*, \quad \kappa \geq 0,
\end{align*}
\]
where \(\beta = 1 + (x^0)^T s^0 > 1\), and the decision variables are 
\((y, x, \tau, \theta, s, \kappa)\).
The above self-dual model admits an interior feasible solution

\[(y, x, \tau, \theta, s, \kappa) = (y^0, x^0, 1, 1, s^0, 1)\].

**Proposition 1** The problem \((SD)\) has a maximally complementary optimal solution, denoted by \((y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*)\), such that \(\theta^* = 0\) and \((x^*)^T s^* + \tau^* \kappa^* = 0\). Moreover, if \(\tau^* > 0\), then \(x^*/\tau^*\) is an optimal solution for the primal problem, and \((y^*/\tau^*, s^*/\tau^*)\) is an optimal solution for the dual problem. If \(\kappa^* > 0\) then either \(c^T x^* < 0\) or \(b^T y^* > 0\); in the former case the dual problem is strongly infeasible, and in the latter case the primal problem is strongly infeasible.
If strict complementarity fails, then one can only conclude that the original problem does not have complementary optimal solutions. In that case, one may wish to identify whether the problems are weakly feasible or weakly infeasible. However, the weak infeasibility is in a slippery terrace; it cannot be certified by a finite proof. The type of certificate that one can expect is in the spirit of “if the problem is ever feasible, then any of its feasible solution must have a very large norm”.
For linear programming, the situation is particularly interesting.

Remember we have shown before that if a primal-dual LP pair is strongly feasible, then the primal-dual central path solutions converge to a *strictly complementary* optimal solutions. That is,

\[
\begin{align*}
    Ax(\mu) &= b \\
    A^T y(\mu) + s(\mu) &= c \\
    x_i(\mu)s_i(\mu) &= \mu, \ i = 1, 2, \ldots, n
\end{align*}
\]

and \(x^* := \lim_{\mu \downarrow 0} x(\mu)\) and \(s^* := \lim_{\mu \downarrow 0} s(\mu)\) are strictly complementary.
When we apply this to \((SD)\), we have

\[
\begin{cases}
Ax - b\tau = 0 \\
-A^T y + c\tau - s = 0 \\
b^T y - c^T x - \kappa = 0
\end{cases}
\]

with

\[
x \geq 0, \ s \geq 0, \ \tau \geq 0, \ \kappa \geq 0 \text{ and } x + s > 0, \ \tau + \kappa > 0.
\]
This constructively proves a well known result of Goldman and Tucker:

**Theorem 12 (Goldman and Tucker, 1956).** Suppose that $M$ is skew-symmetric. Then the following system always has a solution:

\[
\begin{align*}
Mx - s &= 0 \\
x &\geq 0, \ s \geq 0, \ x + s > 0.
\end{align*}
\]

In fact, the theorem of Goldman and Tucker can also be used to prove the previous result: they are equivalent.
Key References:


