Midterm Review

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

http://www.stanford.edu/~yyye
Symmetric Matrices and Quadratic Functions

• $\mathcal{M}^n$

• Positive Definite (PD): $Q \succ 0$ iff $x^T Q x > 0$, for all $x \neq 0$

• Positive SemiDefinite (PSD): $Q \succeq 0$ iff $x^T Q x \geq 0$, for all $x$

• $f(x) = x^T Q x$, then $\nabla f(x) = Q x$ and $\nabla^2 f(x) + Q$. 
Convex Cones

- A set $C$ is a cone if $x \in C$ implies $\alpha x \in C$ for all $\alpha > 0$
- The intersection of cones is a cone
- A convex cone is a cone and a convex set.
- Dual cone:

  \[ C^* := \{ y : x \cdot y \geq 0 \ \text{for all} \ x \in C \} \]
Calculus of Dual Cones

- Let \((x_1; x_2) \in C_1 \oplus C_2\). Then the dual cone is \(C_1^* \oplus C_2^*\).
- Let \(x \in C_1 \cup C_2\). Then, the dual cone is \(C_1^* \cap C_2^*\).
- Let \(x \in C_1 \cap C_2\). Then, the dual cone is the convex hall of \(x \in C_1^* \cup C_2^*\).
- In general, the dual cone can be constructed as

\[
\{ s : 0 \leq \inf_{x} s \cdot x, \ s.t. \ x \in C, \|x\| \leq 1 \}.
\]
The most important theorem about the convex set is the following separating theorem.

**Theorem 1** (Separating hyperplane theorem) Let \( C \) be a convex set in \( \mathbb{R}^m \) and let \( b \) be a point exterior to the closure of \( C \). Then there is a vector \( y \in \mathbb{R}^m \) such that

\[
b \cdot y > \sup_{x \in C} x \cdot y.
\]
Farkas’ Lemma and Alternative Systems for Linear Systems

Given matrix $A \in \mathbb{R}^{m \times n}$ vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$, exactly one of the following system pairs is feasible: Alternative systems: \{x : Ax = b\} or \{y : A^T y = 0, b^T y \neq 0\}.

\{x : Ax = b, x \geq 0\} or \{y : A^T y \leq 0, b^T y = 1\}

\{x : Ax = 0, c^T x = -1, x \geq 0\} or \{y : A^T y \leq c\}
Duality Theory

\[(CLP) \quad \text{minimize} \quad c \cdot x \]
\[\text{subject to} \quad a_i \cdot x = b_i, \; i = 1, 2, \ldots, m, \; x \in C.\]

\[(CLD) \quad \text{maximize} \quad b^T y \]
\[\text{subject to} \quad \sum_i^m y_i a_i + s = c, \; s \in C^*, \]
where \(y \in R^m\), \(s\) is called the dual slack vector/matrix, and \(C^*\) is the dual cone of \(C\).

\[x \cdot s \geq 0\]
for any feasible \(x\) of (CLP) and \(s\) of (CLD).

Linear Programming (LP): \(c, a_i, x \in R^n\) and \(C = R^n_+\)

Semidefinite Programming (SDP): \(c, a_i, x \in M^n\) and \(C = M^n_+\)
Theorem 2 (Weak duality theorem) Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then,

\[ c \cdot x \geq b^T y \quad \text{where} \quad x \in \mathcal{F}_p, \ (y, s) \in \mathcal{F}_d. \]

\[ c \cdot x - b^T y = c \cdot x - (Ax)^T y = x \cdot (c - A^T y) = x \cdot s \geq 0. \]

This theorem shows that a feasible solution to either problem yields a bound on the value of the other problem. We call $c \cdot x - b^T y$ the duality gap.
**Strong Duality Theorem for CLP**

**Theorem 3** *(Strong duality theorem)* Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty and at least one of them has an interior and it has an bounded optimizer, say (CLP). Then, $x^\ast$ is optimal for (CLP) and $(y^\ast, s^\ast)$ is optimal for (CLD) if and only if

$$c \cdot x^\ast = b^T y^\ast.$$ 

If both $\mathcal{F}_p$ and $\mathcal{F}_d$ have interior, then there are bounded $x^\ast$ feasible for (CLP) and $(y^\ast, s^\ast)$ feasible for (CLD) such that

$$c \cdot x^\ast = b^T y^\ast.$$ 

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad b = 2.$$
Complementarity Gap for LP

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

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

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

\[
\begin{align*}
\mathbf{x} \ast \mathbf{s} &= 0 \\
A\mathbf{x} &= \mathbf{b} \\
-A^T \mathbf{y} - \mathbf{s} &= -\mathbf{c}.
\end{align*}
\]

This system has total $2n + m$ unknowns and $2n + m$ equations including $n$ nonlinear equations.
Let both the primal and dual LP be feasible. Then

- **Strong duality** always holds;
- There is always a **strictly complementarity** solution;
- There is always an optimal primal basic feasible solution where at most \( m \) entries are nonzero; and the is always an optimal dual basic feasible solution where at least \( m \) constraints are active.
### Rules to construct the dual

<table>
<thead>
<tr>
<th></th>
<th>obj. coef. vector</th>
<th>right-hand-side</th>
<th>obj. coef. vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>right-hand-side</td>
<td>$A$</td>
<td>$A^T$</td>
<td></td>
</tr>
<tr>
<td>Max model</td>
<td>$x_j \in C$</td>
<td>$-x_j \in C$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_j \in S$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$i$th constraint $\leq C$</td>
<td></td>
<td>$i$th constraint $\geq C$</td>
</tr>
<tr>
<td></td>
<td>$i$th constraint $\geq C$</td>
<td></td>
<td>$i$th constraint $= S$</td>
</tr>
<tr>
<td>Min model</td>
<td>$j$th constraint $\geq C^*$</td>
<td></td>
<td>$j$th constraint $\leq C^*$</td>
</tr>
<tr>
<td></td>
<td>$j$th constraint $= S$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$y_i \in C^*$</td>
<td>$-y_i \in C^*$</td>
<td>$y_i \in S$</td>
</tr>
</tbody>
</table>

The dual of the dual is the primal!
Optimality Conditions: When it is an optimizer

The question: How does one recognize a local optimal solution to a nonlinear programming problem? We first consider the linearly constrained optimization:

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{(LCOP)} & \\
\text{subject to} & \quad x \in F,
\end{align*}
\]

where \( F \) is a convex polyhedron, for example,

\[
F = \{x : Ax = b, \; x \geq 0\}.
\]

A local optimal solution or local minimizer:

\[
\bar{x} \in F \quad \text{and} \quad f(\bar{x}) \leq f(x) \quad \forall x \in F \cap N(\bar{x})
\]

where \( N(\bar{x}) \) is a small ball centred at \( \bar{x} \).
Descent Direction

Let \( f \) be a differentiable function on \( R^n \). If point \( \bar{x} \in R^n \) and there exists a vector \( d \) such that

\[
\nabla f(\bar{x})d < 0,
\]

then there exists a scalar \( \bar{\tau} > 0 \) such that

\[
f(\bar{x} + \tau d) < f(\bar{x}) \text{ for all } \tau \in (0, \bar{\tau}).
\]

The vector \( d \) (above) is called a descent direction at \( \bar{x} \). If \( \nabla f(\bar{x}) \neq 0 \), then \( \nabla f(\bar{x}) \) is the direction of steepest ascent and \( -\nabla f(\bar{x}) \) is the direction of steepest descent at \( \bar{x} \).

Denote by \( D_{\bar{x}}^d \) the set of descent directions at \( \bar{x} \), that is,

\[
D_{\bar{x}}^d = \{ d \in R^n : \nabla f(\bar{x})d < 0 \}.
\]
At feasible point $\bar{x}$, a feasible direction is

$$\mathcal{D}_{\bar{x}}^f := \{ d \in \mathbb{R}^n : d \neq 0, \bar{x} + \lambda d \in F \text{ for all small } \lambda > 0 \}.$$

Examples:

$$F = \mathbb{R}^n \Rightarrow \mathcal{D}_{\bar{x}}^f = \mathbb{R}^n.$$

$$F = \{ x : Ax = b \} \Rightarrow \mathcal{D}_{\bar{x}}^f = \{ d : Ad = 0 \}.$$

$$F = \{ x : Ax \geq b \} \Rightarrow \mathcal{D}_{\bar{x}}^f = \{ d : A_i d \geq 0, \forall i \in \mathcal{A}(\bar{x}) \},$$

where the active or binding constraint set $\mathcal{A}(\bar{x}) := \{ i : A_i \bar{x} = b_i \}$. 
A General Answer to Optimality Conditions

What are the necessary conditions in order to have $\bar{x}$ as a local optimizer?

A general answer: the intersection of the descent and feasible direction sets at $\bar{x}$, $D^d_{\bar{x}}$ and $D^f_{\bar{x}}$, must be empty. Or

$$(\text{LCLP})$$

minimize $\nabla f(\bar{x})d$

subject to $d \in D^f_{\bar{x}}$

is bounded from below.
Sample Problem 1: Construction of the Dual Cone

Let $Q$ be positive definite and the convex cone $C(Q)$ be

$$C(Q) = \{(t; x) : \sqrt{x^T Q x} \leq t\}.$$

We like to find the dual cone of $C(Q)$:

$$C(Q)^* = \{(t'; x') : t't + (x')^T x \geq 0, \forall (t; x) \in C(Q)\}.$$

For given $(t'; x')$, consider the optimization problem

minimize \[ t't + (x')^T x \]
subject to \[ \sqrt{x^T Q x} \leq t. \]

Let the minimal value function $z(t'; x') \geq 0$. Then $(t'; x')$ is in the dual cone; and the converse is also true.
Construction of the Dual Cone II

It is sufficient to consider $t = 1$:

\[
\begin{align*}
\text{minimize} & \quad t' + (x')^T x \\
\text{subject to} & \quad x^T Q x \leq 1.
\end{align*}
\]

This is an ellipsoid constrained problem and the unique minimizer is

\[
x = -\frac{1}{\sqrt{(x')^T Q^{-1} x'}} \cdot Q^{-1} x'.
\]

Thus

\[
z(t'; x) = t' - \sqrt{(x')^T Q^{-1} x'} \geq 0
\]

is the dual cone.
Sample Problem 2: Combinatorial Auction

\[
\begin{align*}
\text{max} \quad & \pi^T x - \alpha \\
\text{s.t.} \quad & Ax + s - e \cdot \alpha = 0, \\
& x \leq q, \\
& x, s \geq 0. \\
\end{align*}
\]

\[
\begin{align*}
\text{min} \quad & q^T y \\
\text{s.t.} \quad & A^T p + y \geq \pi, \\
& e^T p = 1, \\
& (p, y) \geq 0.
\end{align*}
\]
Combinatorial Auction: convex programming model

\[
\begin{align*}
\text{min} & \quad f(s) \\
\text{s.t.} & \quad Ax - s \leq 0, \\
& \quad \pi^T x - \max\{s\} \geq \mu, \\
& \quad x \leq q, \\
& \quad x \geq 0,
\end{align*}
\]

where \(f(s)\) can be viewed as organizer’s risk function on outstanding shares.