Lagrangian duality and general duality – highlights 5

October 2010

Recall convex programming problem (new name primal problem):

\[(P) \inf_{x \in Z} f(x),\]

with 

\[Z := \{x \in S : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, Ax-b = 0\}.\]

Assume \(\inf(P) < +\infty \) (\(\Leftrightarrow Z \neq \emptyset\) and \(f \not\equiv +\infty\) on \(Z\)).

Define Lagrangian dual of \((P)\) by

\[(D) \sup_{(u,v) \in \mathbb{R}_+^m \times \mathbb{R}^p} \theta(u, v),\]

where

\[\theta(u, v) := \inf_{x \in S} [f(x) + \sum_{i=1}^m u_i g_i(x) + v^t(Ax - b)].\]

Note: \(\theta : \mathbb{R}_+^m \times \mathbb{R}^p \to [-\infty, +\infty) (\ast)\) is concave.
Example 1.1: Standard LP problem:

\[(P) \quad \inf_{x \in S} \{c^t x : Ax - b = 0\},\]

with \(S := \mathbb{R}_+^n\). Then above definition gives

\[\theta(u, v) := \inf_{x \geq 0} [c^t x + v^t (Ax - b)].\]

So

\[\theta(u, v) = \begin{cases} 
- b^t v & \text{if } c + A^t v \geq 0 \\
-\infty & \text{otherwise}
\end{cases}\]

Write \(w := -v\). Then \((D)\) becomes

\[(D) \quad \sup_{w \in \mathbb{R}^p} \{b^t w : A^t w \leq c\},\]

which is familiar form of dual in LP.

Example 1.2: Linear regression problem:

\[(P) \quad \inf_{x \in \mathbb{R}^n} \{|x|^2 : Ax - b = 0\},\]

Only equality constraints, so use for \((D)\)

\[\theta(v) := \inf_{x \in \mathbb{R}^n} [|x|^2 + v^t (Ax - b)].\]

To calculate, solve \(\nabla (|x|^2 + v^t (Ax - b)) = 0\), then \(\theta(v) = -v^t AA^t v/4 - b^t v\). Hence

\[(D) \quad \sup_{v \in \mathbb{R}^p} [-v^t AA^t v/4 - b^t v].\]

\((D)\) is solved via \(\nabla \theta(v) = 0\). Hence, \(\bar{v} := -2(AA^t)^{-1} b\)

is optimal (recall \(\theta\) is concave and recall \(A\) has rank \(p\), so \(AA^t\) is invertible).
Direct consequence of KKT theorem in [OSC]:

**Theorem 1.3 (Lagrangian duality):** (i) For all $x \in Z$ and $(u, v) \in \mathbb{R}_+^m \times \mathbb{R}^p$

$$\theta(u, v) \leq f(x) \text{ (weak duality).}$$

In particular, if for some $\bar{x} \in Z$ and $(\bar{u}, \bar{v}) \in \mathbb{R}_+^m \times \mathbb{R}^p$

$$\theta(\bar{u}, \bar{v}) = f(\bar{x})$$

then $\bar{x}$ is optimal for $(P)$, $(\bar{u}, \bar{v})$ is optimal for $(D)$ and complementary slackness holds for $\bar{x}$.

(ii) Conversely, if $\bar{x}$ is an optimal solution of $(P)$ and if both the regularity condition and Slater’s constraint qualification hold, then there exists $(\bar{u}, \bar{v}) \in \mathbb{R}_+^m \times \mathbb{R}^p$ such that

$$\theta(\bar{u}, \bar{v}) = f(\bar{x}) \text{ (strong duality).}$$
Duality by perturbations

Associate to
\[ (P) \quad \inf_{x \in S} \{ f(x) : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, Ax-b = 0 \}, \]
the perturbed optimization problems \((P_{y,z})\), \(y \in \mathbb{R}^m, z \in \mathbb{R}^p\), given by
\[ \inf_{x \in S} \{ f(x) : g_1(x) \leq y_1, \ldots, g_m(x) \leq y_m, Ax-b = z \} \]
(recall \(\inf \emptyset := +\infty\)). Then \((P_{0,0})\) coincides with \((P)\).

Define \(\nu : \mathbb{R}^m \times \mathbb{R}^p \to [-\infty, +\infty]\) by \(\nu(y, z) := \inf(P_{y,z})\). Observe: \(\nu\) is convex and \(\nu(0, 0) = \inf(P)\).

By definition of Fenchel conjugate
\[ \nu^*(-u, -v) := \sup_{y,z} \{ -u^t y - v^t z - \nu(y, z) \}. \]

By \(\nu(y, z) := \inf(P_{y,z})\) obtain
\[ \nu^*(-u, -v) = \sup_{x \in S, y} \{ -u^t y - v^t (Ax-b) - f(x) : g_j(x) \leq y_j \forall j \} \]

Hence,
\[ \nu^*(-u, -v) = \begin{cases} -\theta(u, v) & \text{if } u \geq 0 \\ +\infty & \text{otherwise} \end{cases} \]

So can rewrite Lagrangian dual of previous \((P)\) as
\[ (D) \quad \sup_{(u,v) \in \mathbb{R}^m \times \mathbb{R}^p} -\nu^*(-u, -v). \]
(note use of full space $\mathbb{R}^m \times \mathbb{R}^p$). Observe:

$$\sup(D) = \sup_{(u,v)}[0^t(-u) + 0^t(-v) - \nu^*(-u, -v)] =: \nu^{**}(0, 0)$$

by definition of biconjugate. Hence

$$\nu^{**}(0, 0) \leq \nu(0, 0) \iff \sup(D) \leq \inf(P) \iff \text{weak duality}$$

Also, if $\bar{x}$ optimal for $(P)$ then $\forall (\bar{u}, \bar{v}) \in \mathbb{R}^m \times \mathbb{R}^p$

$$(-\bar{u}, -\bar{v}) \in \partial \nu(0, 0) \iff \nu(0, 0) = -\nu^*(-\bar{u}, -\bar{v})$$

shows

$$f(\bar{x}) = \theta(\bar{u}, \bar{v}) \iff (-\bar{u}, -\bar{v}) \in \partial \nu(0, 0) \iff \text{strong duality}.$$
Let (general) primal problem be

$$\min_{x \in \mathbb{R}^n} \phi_0(x),$$

for given $\phi_0 : \mathbb{R}^n \to (-\infty, +\infty]$.

A perturbation scheme for $(\mathbb{P})$ consists of $l \in \mathbb{N}$ and $\phi : \mathbb{R}^n \times \mathbb{R}^l \to (-\infty, +\infty]$ such that

$$\forall x \phi(x, 0) = \phi_0(x).$$

Call

$$\min_{x \in \mathbb{R}^n} \phi(x, u)$$

for $u \in \mathbb{R}^l$ the $u$- perturbation of $(\mathbb{P})$.

**Example** Consider previous convex programming problem $(P)$. Define for $u := (y, z) \in \mathbb{R}^m \times \mathbb{R}^p$

$$\phi(x, u) := \begin{cases} f(x) & \text{if } x \in S, g(x) \leq y \text{ and } Ax - b = z \\ +\infty & \text{otherwise} \end{cases}$$

For this perturbation scheme $(\mathbb{P})$ is equivalent to $(P)$:

$$\phi(x, 0) = \begin{cases} f(x) & \text{if } x \in Z \\ +\infty & \text{if } x \in \mathbb{R}^n \setminus Z \end{cases}$$
Define **perturbation function** \( h : \mathbb{R}^l \to [-\infty, +\infty] \) by
\[
h(u) := \inf(\mathbb{P}_u) = \inf_{x \in \mathbb{R}^n} \phi(x, u).
\]

Define **dual problem** corresponding to above perturbation scheme:
\[
(\mathbb{D}) \sup_{v \in \mathbb{R}^l} -h^*(-v).
\]

Observe: dual objective function is \(-h^*(-v)\) and \(\sup(\mathbb{D}) = h^{**}(0)\).

**Theorem 3.1 (general duality-stability):**
(i) For all \( x \in \mathbb{R}^n \) and \( q \in \mathbb{R}^k \)
\[
-h^*(-q) \leq f(x) \quad \text{(weak duality)}.
\]
Consequently,
\[
\inf(\mathbb{P}) = h(0) \geq h^{**}(0) = \sup(\mathbb{D}).
\]
(ii) \( h \) is l.s.c. at 0 \(\Leftrightarrow\) \(\inf(\mathbb{P}) = \sup(\mathbb{D})\).
(iii) If \( h \) is continuous at 0 then
\[
\inf(\mathbb{P}) = \max(\mathbb{D}).
\]
Here set of optimal dual solutions is *nonempty* and equal to \(-\partial h(0)\).
Terminology:
$h$ l.s.c. at 0 is called weak stability
$h$ continuous at 0 is called (strong) stability.
Specialization: **Fenchel duality.**

Let $A$ be $m \times n$-matrix. Consider

$$(P_F) \quad \inf_{x \in \mathbb{R}^n} [f(x) + g(Ax)],$$

where $f : \mathbb{R}^n \to (-\infty, +\infty]$ and $g : \mathbb{R}^m \to (-\infty, +\infty]$ are convex functions. Suppose that $\inf(P_F) \in \mathbb{R}$.

Define associated **Fenchel dual problem:**

$$(D_F) \quad \sup_{q \in \mathbb{R}^m} [-f^*(A^tq) - g^*(-q)].$$

**Thm 3.2 (Fenchel’s duality theorem)**

(i) For all $x \in \mathbb{R}^n$ and $q \in \mathbb{R}^m$

$$-f^*(A^tq) - g^*(-q) \leq f(x) + g(Ax).$$

(ii) If $0 \in \text{int} (\text{dom } g - A(\text{dom } f))$, then

$$\inf(P_F) = \max(D_F).$$

Moreover, then $\bar{x} \in \mathbb{R}^n$ is optimal for $(P_F)$ and $\bar{q} \in \mathbb{R}^k$ is optimal for $(D_F)$ if and only if

$$A^t\bar{q} \in \partial f(\bar{x}) \text{ and } -\bar{q} \in \partial g(A\bar{x}).$$
**Examples** Let $b \in \mathbb{R}^m$ and let $K \subset \mathbb{R}^m$ be convex cone.

(a) $g := \chi\{b\}$ gives $Ax = b$.

(b) $g := \chi_{b+K}$ gives conical constraint $Ax \in b + K$.

(c) $f(x) := \mu^{-1}c^tx - \sum_{i=1}^n \log(x_i) + \chi_{\mathbb{R}^n^+}(x)$ and $g := \chi\{b\}$. Here $c \in \mathbb{R}^n$ and $\mu > 0$ is a scaling (penalty) parameter. Then $(P_F)$ is

$$\inf_{x \in \mathbb{R}^n_+, Ax = b} \mu^{-1}c^tx - \sum_{i=1}^n \log(x_i).$$

(*logarithmic barrier function*).

(d) $f(x) = \sum_{i=1}^n f_i(x_i)$ (additive separability) and $g = \chi\{b + K\}$, then $(P_F)$ is

$$\inf_{x, Ax \in b + K} \sum_i f_i(x_i).$$

Here each $f_i$ is convex function on $\mathbb{R}$.