On subdifferential calculus – highlights 4

October 2010

Appendix B: Fenchel conjugation

Definition B.1: Let $f : \mathbb{R}^n \to (-\infty, +\infty]$. The *Fenchel conjugate of* f is $f^* : \mathbb{R}^n \to [-\infty, +\infty]$, given by

$$f^*(\xi) := \sup_{x \in \mathbb{R}^n} [\xi^t x - f(x)].$$

Define the *Fenchel biconjugate* of f by repetition:

$$f^{**}(x) := \sup_{\xi \in \mathbb{R}^n} [\xi^t x - f^*(\xi)],$$

so f^{**} is the conjugate of f^*.
Simple general properties (see Proposition B.1):

- f^* and f^{**} are convex, even when f isn't.
- $f \geq g \Rightarrow f^* \leq g^*$.
- $\exists_\xi \; f^*(\xi) = -\infty \Leftrightarrow f \equiv +\infty$ on all of $\mathbb{R}^n$.
- $\forall_{x_0, \xi \in \mathbb{R}^n}$

$$f^*(\xi) \geq \xi^t x_0 - f(x_0) \text{ (Young's inequality).}$$

- $f \geq f^{**}$.
- $\forall_{x_0, \xi \in \mathbb{R}^n}$

$$f^*(\xi) = \xi^t x_0 - f(x_0) \Leftrightarrow \xi \in \partial f(x_0).$$

Example B.2: a. Let $f : \mathbb{R} \to (-\infty, +\infty]$ be given by

$$f(x) := \begin{cases} x \log x & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ +\infty & \text{if } x < 0. \end{cases}$$

For fixed ξ calculate $f^*(\xi)$:
Step 1. Can restrict maximization to dom f:

$$f^*(\xi) = \sup_{x \geq 0} [\xi x - x \log x].$$

Here $0 \log 0 := 0$ captures $f(0) = 0$.
Step 2: So now we maximize over $\mathbb{R}_+$!
Step 2a: Search for *interior* maximum. Note: f is convex, so $\psi(x) := \xi x - x \log x$ is concave. Hence for $x_0 > 0$:

$$x_0 \text{ gives } \textit{interior} \text{ maximum} \Leftrightarrow \psi'(x_0) = 0$$

by Prop. 2.6. So obtain $\xi - \log x_0 - 1 = 0$, i.e., $x_0 = \exp(\xi - 1)$ (note: $x_0 > 0$, so is indeed interior!). Get $\psi(x_0) = \exp(\xi - 1)$.
Step 2b: Search for maximum on boundary. Only point in boundary is $x = 0$, with $\psi(0) = 0$.

$0 < \exp(\xi - 1)$, so combining steps 2a-b gives $f^*(\xi) = \exp(\xi - 1)$.

Next, fix x and calculate $f^{**}(x)$:

$$f^{**}(x) = \sup_{\xi \in \mathbb{R}} \phi(\xi) := \xi x - \exp(\xi - 1)].$$

Here ϕ is concave and differentiable on $\mathbb{R}$. So for $\xi_0 \in \mathbb{R}$

$$\xi_0 \text{ gives maximum} \Leftrightarrow g'(\xi_0) = 0$$

by Prop. 2.6. So

$$\xi_0 \text{ gives maximum} \Leftrightarrow \xi_0 = \log x + 1$$

Note: $\log x$ makes only sense for $x > 0$. So distinguish

Case 1: $x > 0$. Then $f^{**}(x) = \phi(\log x + 1) = x \log x$.

Case 2: $x < 0$. Then $f^{**}(x) = +\infty$ (let $\xi \to -\infty$).

Case 3: $x = 0$. Now

$$f^{**}(0) = \sup_{\xi \in \mathbb{R}} -\exp(\xi - 1) = 0$$

by $\lim_{\xi \to -\infty} -\exp(\xi - 1) = 0$.

Combining cases 1,2,3 gives $f^{**} = f$.

b. Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be given by

$$f(x) := \begin{cases} -\sum_{i=1}^n \log(x_i) & \text{if } x \in \mathbb{R}^n_{++}, \\ +\infty & \text{otherwise.} \end{cases}$$

Can exclude dom f from sup:

$$f^*(\xi) = \sup_{x \in \mathbb{R}^n_{++}} \psi(x) := \sum_i \xi_i x_i + \sum_i \log(x_i).$$

Here $-\psi$ is convex and differentiable. So for any $x \in \mathbb{R}^n_{++}$:

$$x \text{ gives maximum in sup } \Leftrightarrow \nabla\psi(x) = 0,$$

by Prop. 2.6. Gives $\xi_i + x_i^{-1} = 0$ for each i. So

$$x \text{ gives maximum in sup } \Leftrightarrow \forall_i \xi_i = -x_i^{-1}.$$

By requirement $x \in \mathbb{R}^n_{++}$ distinguish:

Case 1: $\forall_i \xi_i < 0$: then $x_i := -\xi_i^{-1} > 0$ for all i. So

$$f^*(\xi) = -n + \sum_i \log(1/-\xi_i) = -n + f(-\xi).$$

Case 2: $\exists_j\ \xi_j > 0$: then above sup not attained and actually $f^*(\xi) = +\infty$.

Combining cases 1-2 gives

$$f^*(\xi) = \begin{cases} -n - f(-\xi) & \text{if } \xi \in \mathbb{R}^n_{--} \\ +\infty & \text{otherwise} \end{cases}$$

Next, fix x and calculate $f^{**}(x)$: by $f^*(\xi)$ above obtain

$$f^{**}(x) = \sup_{\xi \in \mathbb{R}^n_{--}} [\xi^t x + n - f(-\xi)].$$

Trick: change of variable $\zeta := -\xi$ gives

$$f^{**}(x) = \sup_{\zeta \in \mathbb{R}^n_{++}} [-\zeta^t x + n - f(\zeta)] \overset{!}{=} n + f^*(-x)$$

by above expression for $f^*(\xi)$. If $x \in \mathbb{R}^n_{--}$ this implies $f^{**}(x) = n - n + f(x) = f(x)$ and if $x \notin \mathbb{R}^n_{--}$ one gets $f^{**}(x) = n + \infty = +\infty = f(x)$. Hence, $f^{**} = f$.

Found twice $f^{**} = f$! *What is the explanation?*

Definition: Let $f : \mathbb{R}^n \to (-\infty, +\infty]$. Then f is *lower semicontinuous* (l.s.c.) at $x_0 \in \mathbb{R}^n$ if

$$\forall_{r \in \mathbb{R}, r < f(x_0)} \exists_{\delta > 0} \forall_{x, |x - x_0| < \delta} \, f(x) > r.$$

Also: f is *l.s.c.* if $\forall_{x_0}$ f l.s.c. at x_0. Further: f is *upper semicontinuous* (u.s.c.) at $x_0 \in \mathbb{R}^n$ if $-f$ is l.s.c. at x_0.

Elementary facts:

1. For $f : \mathbb{R}^n \to (-\infty, +\infty]$ and $x_0 \in \text{int dom } f$:

$$f \text{ continuous at } x_0 \Leftrightarrow f \text{ l.s.c. and u.s.c. at } x_0.$$

2. If $\{f_\kappa : \kappa\}$ is collection of functions $f_\kappa : \mathbb{R}^n \to (-\infty, +\infty]$, such that

$$\forall_\kappa f_\kappa \text{ is l.s.c. at } x_0 \in \mathbb{R}^n,$$

then $f : \mathbb{R}^n \to (-\infty, +\infty]$, defined by $f(x) := \sup_\kappa f_\kappa(x)$, is l.s.c. at x_0.

3. Fact 2 implies that

$$\bar{f}(x) := \sup_q \{q(x) : q : \mathbb{R}^n \to \mathbb{R}, q \text{ l.s.c.}\}$$

defines a l.s.c. function $\bar{f} : \mathbb{R}^n \to (-\infty, +\infty]$. Name: *l.s.c. (lower) hull* of f (it is the "largest l.s.c. function $\leq f$").

4. For $f : \mathbb{R}^n \to (-\infty, +\infty]$:

$$f \text{ is l.s.c. } \Leftrightarrow \text{epi } f \text{ is closed set.}$$

Theorem B.5 (Fenchel-Moreau): Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be convex. Then

$$f(x_0) = f^{**}(x_0) \Leftrightarrow f \text{ is l.s.c. at } x_0.$$

Proof of $\Rightarrow$ is not difficult, because f^{**} is l.s.c. by elementary fact 2 above. $\Leftarrow$ uses the separation Theorem A.4, applied to epi $\bar{f}$, which is convex and also closed by above facts 3, 4.

Example B.2 (continued): Concrete calculations gave $f^{**} = f$. Explanation: f's in Example B.2a-b are convex and l.s.c.

Definition: Let $K \subset \mathbb{R}^n$ be a cone at 0 (i.e., $\forall_{\alpha>0, x\in K} \alpha x \in K$). The (negative) *polar* of K is

$$K^* := \{\xi \in \mathbb{R}^n : \forall_{x\in K} \xi^t x \leq 0\}.$$

Example: Let $S \subset \mathbb{R}^n$ and $x_0 \in S$. Then $K := \cup_{\alpha>0} \alpha(S - x_0)$ is cone at 0. Here

$$K^* = N_S(x_0) = \{\xi : \forall_{x\in S} \xi^t (x - x_0) \leq 0\}.$$

Corollary B.6 (bipolar theorem for cones): Let K be a closed convex cone in $\mathbb{R}^n$. Then $K = K^{**} := (K^*)^*$.

Proof. Set $f := \chi_K$. Then f is l.s.c. and convex. So $f^{**} = f$ by F-M theorem. Now check that

$$f^*(\xi) = \sup_{x \in K} \xi^t x = \chi_{K^*}(\xi)$$

for all ξ. Consequence:

$$f^{**}(x) = \sup_{\xi \in K^*} \xi^t x$$

and $f^{**} = \chi_{K^{**}}$ follows. So $\chi_{K^{**}} = \chi_K$. Conclusion: $K^{**} = K$. QED

Corollary: Let L be a linear subspace of $\mathbb{R}^n$. Then $L = L^{\perp\perp} := (L^\perp)^\perp$.

Farkas' Lemma: Let A be $p \times n$-matrix, $c \in \mathbb{R}^n$. Then precisely one of the following is true:

$$(1)\ \exists_{x \in \mathbb{R}^n} Ax \leq 0 \text{ and } c^t x > 0,$$

$$(2)\ \exists_{y \in \mathbb{R}^p_+} A^t y = c.$$

Proof. Hint: (2) $\Rightarrow$ not (1) is easy. Next, not (1) means:

$$\forall_{x \in \mathbb{R}^n, Ax \leq 0} c^t x \leq 0.$$

Thus, $\forall_{x \in K^*} c^t x \leq 0$, i.e., $c \in K^{**}$. Here $K := A^t(\mathbb{R}^p_+)$ is the closed convex cone generated by all nonnegative linear combinations of colums of A^t (= rows of A, viewed as colums). By F-M Theorem, $K^{**} = K$, so we obtain $c \in K^{**} = K = A^t(\mathbb{R}^p_+)$. QED