Lecture 8

Last time
- Conic optimization
- Examples
- Duality

Today
- Subdifferential calculus
- Lagrange duality
- Fenchel biconjugation

Subdifferential Calculus

The beauty of differential calculus is that we can compute gradients by breaking up our functions into simpler functions. In turn, we can do something similar for the convex subdifferential.

Theorem (subdifferential calculus). For any $f: E \to \overline{\mathbb{R}}$ and $g: Y \to \overline{\mathbb{R}}$ and $A: E \to Y$ linear. Then,

$$\partial(f + g \circ A)(\bar{x}) \supseteq \partial f(\bar{x}) + A^* \partial g(A\bar{x}).$$

If further f and g are convex and $0 \in \operatorname{int}(\operatorname{dom} g - A \operatorname{dom} f)$. Then, equality holds.

$\dashv$

Proof: Suppose $w_f \in \partial f(\bar{x})$ and $w_g \in \partial g(A\bar{x})$, then $\forall x \in E$

$$f(\bar{x}) + g(A\bar{x}) + \langle w_f + A^* w_g , x - \bar{x}\rangle$$
$$= f(\bar{x}) + \langle w_f, x - \bar{x}\rangle + g(A\bar{x}) + \langle w_g, Ax - A\bar{x}\rangle$$
$$\leq f(x) + g(Ax).$$

Thus, $w_f + A^* w_g \in \partial(f + g \circ A)(\bar{x})$.

For the converse, suppose that $w \in \partial(f + g\circ A)(\bar{x})$ so $\bar{x}$ minimizes $f - \langle w, \cdot\rangle + g \circ A$. Then, the dual is also attained for some $\bar{y} \in Y$.

By our characterization of solutions

$$A^* \bar{y} \in \partial(f - \langle w, \cdot\rangle)(\bar{x}) = \partial f(\bar{x}) - w$$

(why?)

and

$$-\bar{y} \in \partial g(A\bar{x}).$$

Thus,

$$w \in \partial f(\bar{x}) - A^*\bar{y} \subseteq \partial f(\bar{x}) + A^* \partial g(A\bar{x}).$$

$\square$

Lagrange Duality

Let us now comeback to problems

with functional constraints:

$$p^* = \begin{cases} \inf & f(x) \\ \text{s.t.} & g_i(x) \le 0 \quad \forall i \in [m] \end{cases} \qquad g: E \to \mathbb{R}^m$$

Assume that $f, g_i : E \to \overline{\mathbb{R}}$ are convex and $\emptyset \neq \operatorname{dom} f \subseteq \bigcap_{i \in [m]} \operatorname{dom} g_i$.

We can rewrite

$$p^* = \inf_x \sup_{\substack{\lambda \ge 0 \\ \lambda \in \mathbb{R}^m}} \underbrace{\left\{ f(x) + \lambda^T g(x) \right\}}_{\substack{\text{Lagrangian} \\ L(x;\lambda).}}$$

The key idea to derive weak duality in this context is swapping the inf and the sup:

$$p^* \overset{\text{(why?)}}{\ge} \sup_{\substack{\lambda \ge 0 \\ \lambda \in \mathbb{R}^m}} \underbrace{\inf_x \left\{ f(x) + \lambda^T g(x) \right\}}_{\Phi(\lambda)} = d^*$$

Q: When do we have equality?

Once more, the key is to consider a value function $\nu : \mathbb{R}^m \to \overline{\mathbb{R}}$:

(B) $\nu(z) = \inf\{f(x) \mid g(x) \le z\}$.

Proposition: We have

$$\nu(0) = p^* \quad \text{and} \quad \nu^{**}(0) = d^*.$$

Proof: The first claim is immediate. To show the second one,

$$\nu^*(y) = \sup_{z} \{ y^T z - \nu(z) \}$$

$$= \sup_{z,x} \{ y^T z - f(x) \mid g(x) \le z \}$$

$$= \sup_{\substack{z,x \\ s \ge 0}} \{ y^T z - f(x) \mid g(x) + s = z \}$$

$$= \sup_{\substack{x \\ s \ge 0}} \{ y^T g(x) + y^T s - f(x) \}$$

$$= \begin{cases} -\Phi(-y) & \text{if } y \le 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Then,

$$d^* = \sup_{\lambda \in \mathbb{R}^m_+} \Phi(\lambda) = \sup_{\lambda \in \mathbb{R}^m_+} \langle \lambda, 0 \rangle - (-\Phi(\lambda))$$

$$= \sup_{\lambda \in \mathbb{R}^m} \langle \lambda, 0 \rangle - \nu^*(\lambda)$$
$$= \nu^{**}(0).$$

$\square$

Then, to understand duality we need to understand when is it that $\nu(0) = \nu^{**}(0)$.

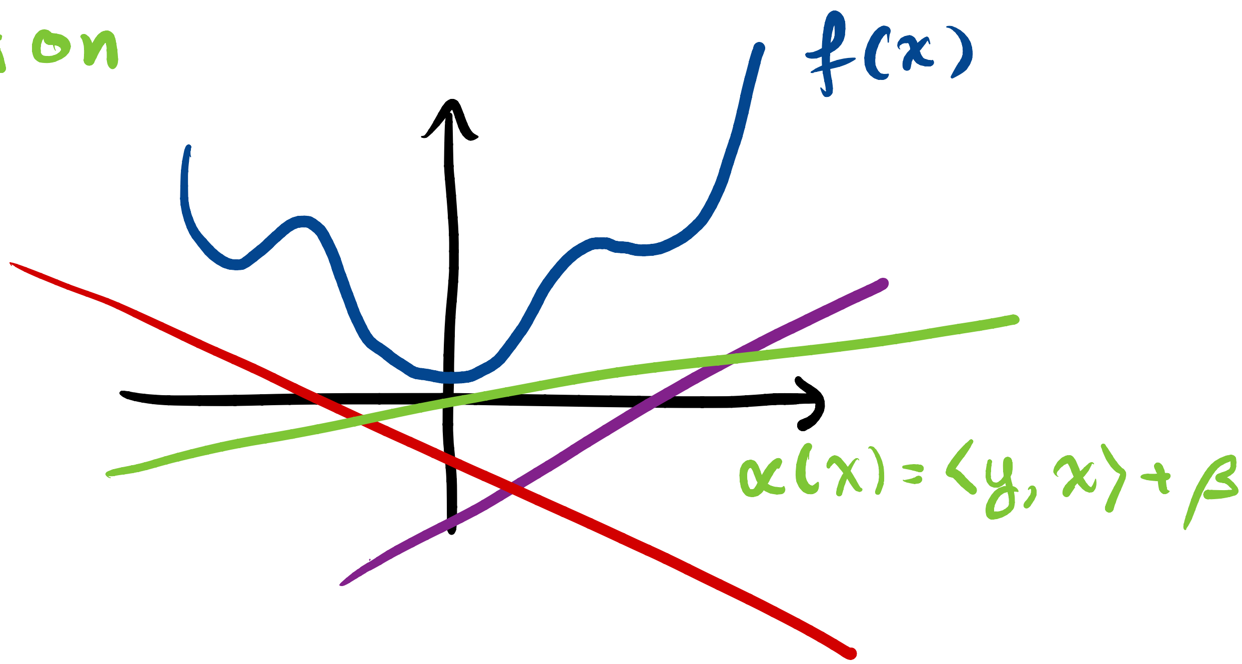
Fenchel biconjugation

In turn, biconjugates are suprema of affine minorants.

Proposition (✓): Given any $f : E \to \overline{\mathbb{R}}$ and $\bar{x} \in E$, we have

$$f^{**}(\bar{x}) = \sup \left\{ \alpha(\bar{x}) \;\middle|\; \alpha \le f \text{ is an affine function} \right\}$$

Intuition



Proof: The RHS is equal to

$$\sup_{y,\beta}\{\langle y,x\rangle-\beta \mid \langle y,x\rangle-\beta\le f(x)\ \ \forall x\in E\}$$

$$=\sup_{y,\beta}\{\langle y,x\rangle-\beta \mid \langle y,x\rangle-f(x)\le \beta\ \ \forall x\in E\}$$

$$=\sup_{y,\beta}\{\langle y,x\rangle-\beta \mid f^*(y)\le\beta\}$$

$$=\sup_{y}\{\langle y,x\rangle-f^*(y)\}$$

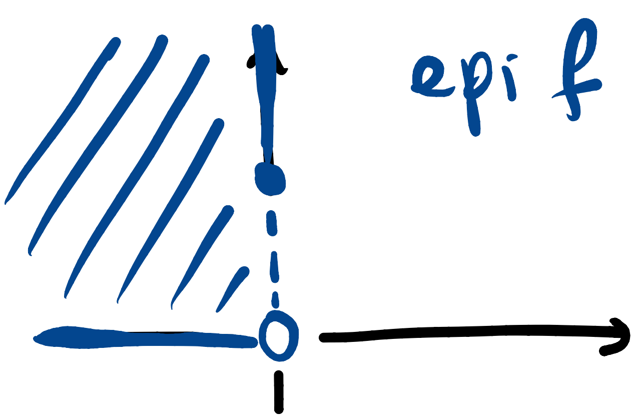
$$=f^{**}(x).$$

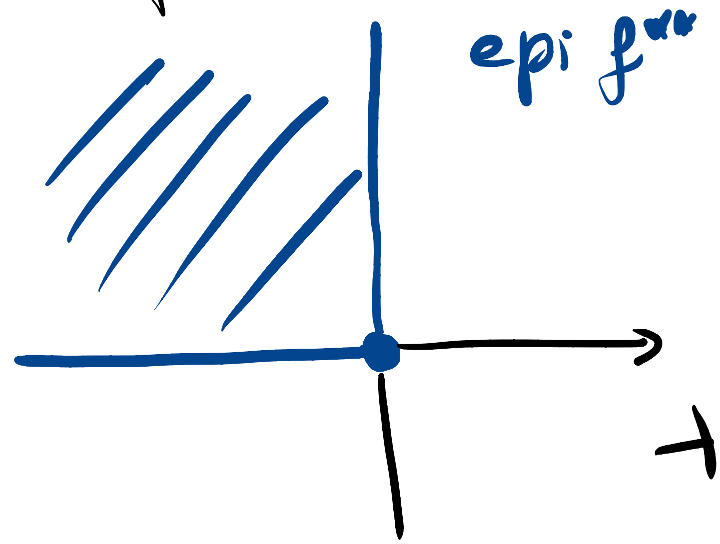
$\square$

Example: Consider

$$f(x)=\begin{cases}0 & \text{if } x<0,\\ 1 & \text{if } x=0,\\ +\infty & \text{otherwise.}\end{cases}$$

Then, $f^*(x)=\iota_{\mathbb{R}_-}(x)$ and f is not closed, while f^{**} is.





Recall that we already proved that f^* is convex and closed.

Theorem (Biconjugates)
Suppose $f: E \to \mathbb{R}$, then, $f = f^{**}$ if, and only if, f is closed and convex.

Proof: ($\Rightarrow$) This follows since f^{**} is closed and convex.
($\Leftarrow$) It suffices to show $f(0) = f^{**}(0)$. By Proposition ($\checkmark$) we want to prove

$$f(0) = \sup \{ \alpha(0) \mid \alpha \leq f \text{ affine} \}.$$

Let's consider two cases:
Case 1: Suppose $0 \in \operatorname{cl} \operatorname{dom} f$. Then, pick $r < f(0)$. Since $(0, r) \notin \operatorname{epi} f$, then $\exists\, \varepsilon > 0$ s.t. $\forall x \in \varepsilon B$ $f(x) > r$. Thus

$$r < \inf_x f(x) + \iota_{\varepsilon B}(x),$$

since $\varepsilon B \cap \operatorname{dom} f \neq \emptyset$, then by

Fenchel duality, there exists a $y \in E$ s.t.

$$r \le -f^*(y) - \iota^*_{\varepsilon B}(-y)$$

(why?) $$= -f^*(y) - \varepsilon \|y\|.$$

By the definition of Fenchel conjugate this is equivalent to

$$\underbrace{\langle y, x\rangle + r + \varepsilon\|y\|}_{\text{affine } \alpha(x).} \le f(x) \quad \forall x \in E$$

Note that $\alpha(0) \ge r$ and so we conclude $f^{**}(0) \ge r$. Since r was arbitrary $f^{**}(0) = f(0)$.

Case 2: $0 \notin \text{cl dom} f$. By our previous argument there is an affine minorant (just take any $x \in \text{dom} f$). Since $\text{dom} f$ is a closed convex set, there exist $z \in E \setminus \{0\}$ and β s.t.

$$\langle z, x\rangle \le \beta < 0 \quad \forall x \in \text{dom} f.$$

(Using Basic separation Thm from)

Then, for any $K \geq 0$

$$\underbrace{\alpha(x) + K(\langle z, x\rangle - \beta)}_{\text{affine}} \leq \alpha(x) \leq f(x) \quad \forall x \in E.$$

Moreover, $\alpha(0) - K\beta$ goes to $+\infty$ as $K \to +\infty$. Thus $f^{**}(0) = +\infty = f(0)$; completing the proof. $\square$

Theorem: Given $\bar{x} \in E$, then $f(\bar{x}) = f^{**}(\bar{x})$ if either

1. f is closed, convex, and proper (never $-\infty$).
2. We have $f(\bar{x})$ is finite and $\partial f(\bar{x}) \neq \emptyset$.

$+$

The proof is an easy exercise.

Back to Lagrangian duality.

Theorem: Suppose f, g_i are convex. Suppose there exists a Slater $\hat{x} \in E$, i.e.,

$$g_i(\hat{x}) < 0 \quad \forall i \in [m].$$

Then, $p^* = d^*$. Moreover if d is finite, there exists a dual optimal $\bar{\lambda}$.

Proof: One can show ν in $(\heartsuit)$ is convex (Exercise). Then, $0 \in \operatorname{int} \operatorname{dom} \nu$ due to $\hat{x}$. Therefore, $\exists \lambda \in \partial \nu(0)$. We conclude from the previous theorem, that

$$p^* = \nu(0) = \nu^{**}(0) = d^*.$$

Exercise: Show that λ is a solution to the dual.

$\square$