Lecture 6

Last time

- Gordon's Theorem of alternatives
- Optimality conditions with functional constraints.

Today

- Fenchel conjugate
- Fenchel duality

Fenchel conjugate

We will introduce a transformation that plays a key role in duality.

Question: Given a function $f : E \to \overline{\mathbb{R}}$ and a point $z \in E$, how to find $\bar{x}$ s.t. $z \in \partial f(\bar{x})$?

Recall that $g \in \partial f(\bar{x})$ iff

$$f(\bar{x}) - \langle z, \bar{x} \rangle \leq f(x) - \langle z, x \rangle \quad \forall x \in E.$$

This is equivalent to

$$\bar{x} \in \operatorname{argmin} f(x) - \langle z, x \rangle$$

$$\Updownarrow$$

$$\bar{x} \in \operatorname{argmax} \langle z, x \rangle - f(x).$$

This is the core idea behing the following definition.

Def: The Fenchel conjugate of a function $f : E \to \overline{\mathbb{R}}$ is

$$f^*(z) = \sup_{x \in E} \langle z, x\rangle - f(x). \quad \dashv$$

Lemma: f^* is convex and closed.

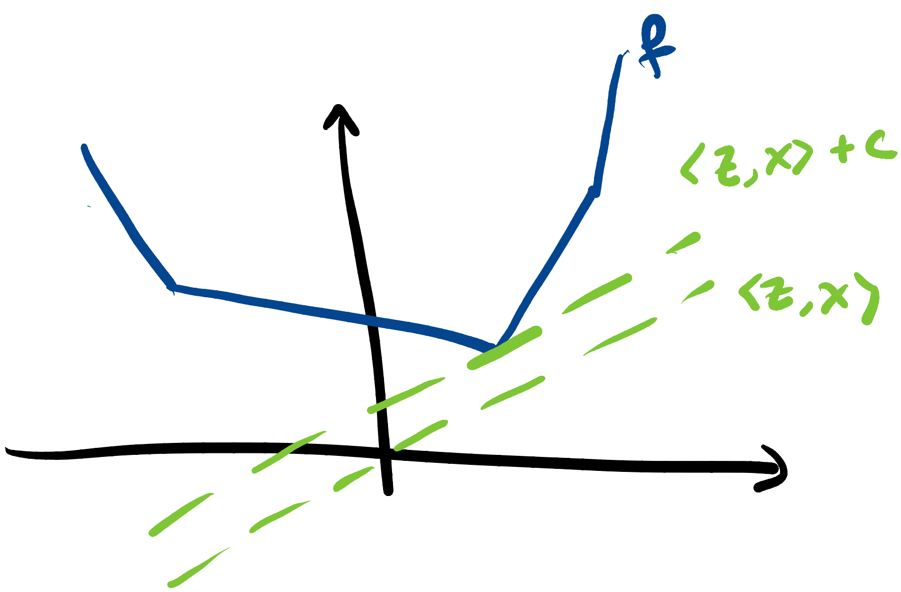
Proof: Note that

$$t \geq f^*(z) \iff t \geq \langle z, x\rangle - f(x) \quad \forall x \in E.$$

Then, epi f^* is an intersection of halfspaces. $\square$

Another interpretation

By our derivation, the sup in $f^*(z)$ is attained whenever $z \in \partial f(x)$. Thus, we could compute f^* geometrically by raising the graph of $x \mapsto \langle z, x\rangle + c$ until it is tangent to f.



$$f^*(z) = -c$$

Proposition: For any $f: E \to \overline{\mathbb{R}}$ we have that

$$f(x) + f^*(z) \geq \langle z, x \rangle$$

If $f(x)$ is finite, then equality holds iff $z \in \partial f(x)$. $\dashv$

Proof: Follows easily by our discussion above. $\square$

Examples

▷ **Linear functions** Suppose $f(x) = \langle c, x \rangle$

Then,

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x \rangle - \langle c, x \rangle \\ &= \sup_x \langle z - c, x \rangle \\ &= \begin{cases} 0 & \text{if } g = c \\ +\infty & \text{otherwise.} \end{cases} \\ &= \iota_{\{c\}}(z). \end{aligned}$$

▷ **Indicator of a point** $f(x) = \iota_{\{c\}}(x)$

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x \rangle - \iota_{\{c\}}(x) \\ &= \sup_{x \in \{c\}} \langle z, x \rangle \\ &= \langle c, z \rangle. \end{aligned}$$

We got $f = f^{**}$. We will see this is a recurrent pattern for convex closed functions.

▷ Indicator of $[-1,1]$ suppose $f(x) = \imath_{[-1,1]}(x)$

$$
\begin{aligned}
f^*(z) &= \sup_x zx - \imath_{[-1,1]}(x) \\
&= \sup_{x\in[-1,1]} zx \\
&= \text{sign}(z)\, z \\
&= |z|.
\end{aligned}
$$

▷ Absolute value $f(x) = |x|$

$$f^*(z) = \sup_x zx - |x|$$

Let's consider two cases if $|z| \le 1$

$$\Rightarrow \sup_x x\,(\underbrace{z - \text{sign}(x)}) \le 0$$ ← achieved with $x=0$

always have opposite signs

if $|z|>1$, then taking any $x = \lambda z$ for any $\lambda > 0$ yields

$$f^*(z) \ge \lambda z^2 - \lambda|z| > \lambda(z^2 - z) > 0$$

Taking $\lambda \uparrow \infty$ gives $f^*(z) = \infty$.

Therefore, $f^*(z) = \imath_{[-1,1]}(x)$.

Once more the two coincide.

▷ p norms Suppose that $f(x) = \frac{|x|^p}{p}$ ← $p \ge 1$.

Then

$$f^*(z) = \frac{|z|^q}{q} \quad \text{with} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

(Exercise)

Fenchel duality

Recall that the adjoint A^* of a linear map $A: E \to Y$ is defined via
↑ Euclidean.

$$\langle Ax, y \rangle = \langle x, A^* y \rangle \quad \forall x \in E, y \in Y.$$

Theorem (Fenchel Duality): For any functions $f: E \to \overline{\mathbb{R}}$ and $g: Y \to \overline{\mathbb{R}}$, and linear map $A: E \to Y$. Define the primal problem

$$p^* = \inf_{x \in E} f(x) + g(Ax)$$

and the dual problem

$$d^* = \sup_{z \in Y} -f^*(A^* z) - g^*(-z).$$

Then, without extra assumptions,

$$p^* \geq d^* \qquad \text{(Weak duality).}$$

If further f and g are never $-\infty$ and convex, and

$$0 \in \operatorname{int}(\operatorname{dom} g - A \operatorname{dom} f) := \operatorname{int}\{u - Av \mid u \in \operatorname{dom} g, v \in \operatorname{dom} f\}$$

(Constraint qualification)

then, we have

$$p^* = d^* \qquad \text{(Strong duality).}$$

In this case d^* is finite, and it is always attained. Moreover, $\bar{x}$ and $\bar{y}$ are optimal if, and only if,

$$A^*\bar{y} \in \partial f(\bar{x}) \quad \text{and} \quad -\bar{y} \in \partial g(A\bar{x})$$

(Complementary slackness)

Moreover, constraint qualification holds if either

1. There is $\hat{x} \in \operatorname{dom} f$ s.t. $A\hat{x} \in \operatorname{int} \operatorname{dom} g$.
2. There is $\hat{x} \in \operatorname{int} \operatorname{dom} f$ s.t. $A\hat{x} \in \operatorname{dom} g$, and A is surjective.

$\dashv$

Proof: Weak duality follows from Fenchel-Young since $\forall (x,y) \in E \times Y$

$$\begin{array}{rl} & f(x) + f^*(A^*y) \geq \langle x, A^*y \rangle \\ + & g(Ax) + g^*(-y) \geq \langle Ax, -y \rangle \\ \hline & f(x) + g(Ax) + f^*(A^*y) + g^*(y) \geq 0 \end{array}$$

Taking an inf yields $p^* \geq d^*$.

The key object to prove strong duality is the value function

$$v(z) := \inf_{x \in E} \{ f(x) + g(Ax + z) \} \quad \text{for } z \in Y.$$

So $p^* = \nu(0)$. Note that if $p = -\infty$, we are strong duality holds trivially by weak duality. Assume that $z \in \mathrm{dom}\, \nu$, thus the infimum is not $+\infty$. Therefore, $z \in \mathrm{dom}\, \nu$ iff $x \in \mathrm{dom}\, f(x)$ and $Ax + z \in \mathrm{dom}\, g$. Thus, $z \in \mathrm{dom}\, g - A\,\mathrm{dom}\, f$. Hence, constraint gualification is equivalent to $0 \in \mathrm{int}\,\mathrm{dom}\, \nu$.

Claim (Exercise): The function ν is convex. Then, since $0 \in \mathrm{int}\,\mathrm{dom}\, \nu$, we have $\nu(0)$ is finite and by HW1 P1c, ν is never $-\infty$ in int dom f. Thus, there exists $-y \in \partial \nu(0)$.

Moreover,

$$\nu(0) + \nu^*(-y) = \langle 0, -y \rangle = 0.$$

Therefore,

$$\begin{aligned}
d^* \le p^* &= \nu(0) \\
&= -\nu^*(-y) \\
&= -\sup_z -\langle y, z\rangle - \nu(z) \\
&= \inf_z \nu(z) + \langle y, z\rangle \\
&= \inf_z \inf_x \{ f(x) + g(Ax+z) \} + \langle y, z\rangle \\
&= \inf_x \underbrace{\inf_z \{ \langle y, Ax+z\rangle + g(Ax+z) \}}_{-g^*(-y)} + \langle y, -Ax\rangle + f(x)
\end{aligned}$$

$$= -g^*(-y) + \overbrace{\inf_x \langle A^*y, x\rangle + f(x)}^{-f^*(A^*y)}$$

$$= -g^*(-y) - f^*(A^*y)$$

$$\leq d^*$$

Thus, by weak duality $p^* = -g^*(-y) - f^*(A^*y) = d^*$. Complementary slackness ensures that the inequalities in (:-)) become equalities.

Now, let's prove that constraint qualification holds given (1) or (2). If (1) holds Then, $0 \in \operatorname{dom} g - A \operatorname{dom} f$ and since $A\hat{x} \in \operatorname{int} \operatorname{dom} g$ $\exists r > 0$ s.t. $0 \in A\hat{x} + rB - A\hat{x} \subseteq \operatorname{dom} g - A \operatorname{dom} f$, thus $0 \in \operatorname{int}(\operatorname{dom} g - A \operatorname{dom} f)$.

If (2) holds, then we can use the following.

Theorem (Open Mapping) Any surjective mapping $A: E \to Y$ maps open sets to open sets. ⊣

We have that $\exists r > 0$ s.t. $\hat{x} + rB \subseteq \operatorname{dom} f$. Therefore $A\hat{x} + r\hat{A}B$ is an open map contained in $A \operatorname{dom} f$. A similar argument as before, yields that constraint qualification holds. □

Proof of the Open Mapping Theorem:

A is one-to-one

We start by assuming $\dim Y = \dim E$.
Notice it suffices to show $0 \in \operatorname{int} AB$.
Seeking contradiction assume $0 \notin \operatorname{int} AB$,
$\Rightarrow \exists (x_n) \subseteq B^c$ s.t. $(Ax_n)_n \subseteq (AB)^c$ with
$Ax_n \to 0$. Therefore, $\frac{Ax_n}{\|x_n\|} \to 0$.
WLOG assume $\frac{x_n}{\|x_n\|} \to x^*$. This is a contradiction since $Ax^* \neq 0$.
If $\dim E > \dim Y$, we reduce to the previous case. Again assume $0 \notin \operatorname{int} AB$. From Linear Algebra we know

$$A(\ker A)^\perp = \text{range } A \quad \text{and} \quad A(B \cap (\ker A)^\perp) = AB.$$

We can run the same argument setting $E' = (\ker A)^\perp$.

$\square$