Lecture 8

Recall:

Thm: Given a Polish space (X,d), for any function $c: X\times X \to \mathbb{R}\cup\{+\infty\}$ that is lsc and bounded below and $\mu,\nu \in P(X)$, there exists $\gamma_* \in \Gamma(\mu,\nu)$ s.t.

$$\gamma_* = \min_{\gamma:\gamma\in\Gamma(\mu,\nu)} \int_{X\times X} c(x^1,x^2)\, d\gamma(x^1,x^2).$$

Prop: Given a locally compact Polish space X and $\{\mu_n\}_{n=1}^{\infty}$ and μ in $P(X)$ satisfying

$$\lim_{n\to\infty} \int f\, d\mu_n = \int f\, d\mu \quad \text{for all } f\in C_c(X),$$

then $\mu_n \to \mu$ narrowly.

Crash course in convex analysis and optimization

Def: Given $f: X\to \mathbb{R}\cup\{+\infty\}$, its domain is
$D(f) = \{x\in X : f(x) < +\infty\}$.

<u>Def</u>: Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, its <u>conjugate</u> $f^*: X^* \to \mathbb{R} \cup \{+\infty\}$ is

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

<u>Prop</u> (Young's Inequality): $\forall\, x \in X,\ y \in X^*$

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

<u>Def</u>: Given a nvs X and $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, its <u>biconjugate</u> $f^{**}: X \to \mathbb{R} \cup \{+\infty\}$ is

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

<u>Lemma</u>: For any $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, f^* and f^{**} are convex and lsc.

<u>Thm</u> (Fenchel-Moreau): Given a nvs X and $f: X \to \mathbb{R} \cup \{+\infty\}$ proper,

(i) f convex and lsc $\iff f = f^{**}$

(ii) If f is convex and $f(x_0) < +\infty$, then f is lsc at $x_0 \iff f(x_0) = f^{**}(x_0)$.

Next: subdifferential of f

- provides another interpretation of f^*
- this is exactly the notion of "regularity" we'll need in our study of primal/dual optimization problems
- "right" generalization of gradient for gradient flows

Def: Given nvs X, $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, $x \in D(f)$, the subdifferential of f at x is the set valued operator

$$\partial f(x) = \{ y \in X^* : \underline{f(x') \geq f(x) + \langle y, x'-x \rangle + o(\|x'-x\|)}, \ \underline{\text{as } x' \to x} \}$$

If $x \notin D(f)$, $\partial f(x) = \emptyset$.

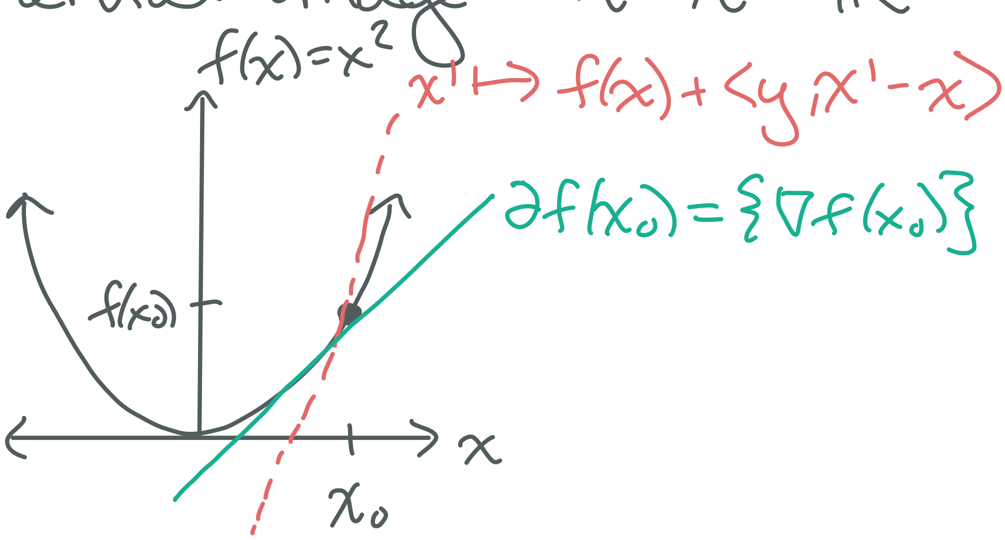
$$\underbrace{\liminf_{x'\to x}}_{\lim\limits_{\delta\to 0}\ \inf\limits_{\{x' : 0<\|x'-x\|<\delta\}}} \frac{f(x')-f(x)-\langle y, x'-x\rangle}{\|x'-x\|} \geq 0$$

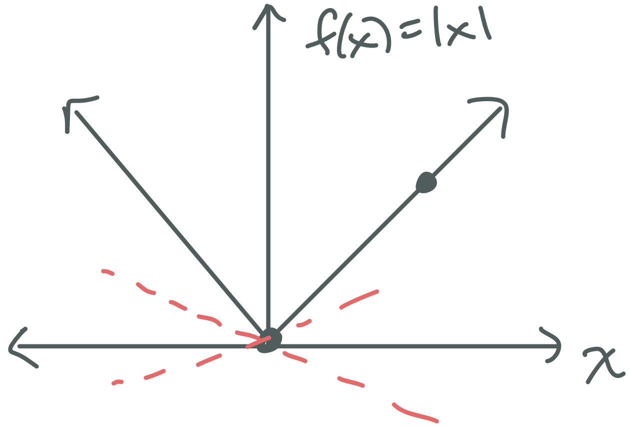
Thm: If f is defferentiable at $x \in D(f)$, then $\partial f(x) = \{\nabla f(x)\}$.

Prop: If $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is convex and $x \in D(f)$, then

$$\partial f(x) = \underbrace{\{y \in \mathcal{X}^* : f(x') \geq f(x) + \langle y, x'-x\rangle \ \forall x' \in \mathcal{X}\}}_{S}$$

Mental image: $\mathcal{X} = \mathcal{X}^* = \mathbb{R}$





$\partial f(0) = [-1, 1]$

Pf: If $y \in S$, then $y \in \partial f(x)$.

Now suppose $y \in \partial f(x)$. Want to show $y \in S$.

For any $\varepsilon > 0$, define

$$\Psi(x') := f(x') - f(x) - \langle y, x' - x \rangle + \varepsilon \|x' - x\|.$$

Since $y \in \partial f(x)$,

$$\lim_{\delta \to 0} \inf_{\{x' : 0 < \|x' - x\| < \delta\}} \frac{\Psi(x')}{\|x' - x\|} \geq \varepsilon.$$

Thus, $\exists\ \delta > 0$ s.t. for x' s.t. $0 < \|x' - x\| < \delta$, we have $\Psi(x') \geq \frac{\varepsilon}{2}$. Since $\Psi(x) = 0$, x must be a local minimizer of Ψ.

Since Ψ is a convex function and x is a local minimizer, in fact, x is a global minimizer. That is,

$$\underset{\displaystyle f(x')-f(x)-\langle y, x'-x\rangle+\varepsilon\|x'-x\|}{\underset{\shortparallel}{\Psi(x')}} \geq \Psi(x)=0 \qquad \forall x'\in X$$

Since $\varepsilon>0$ was arbitrary,

$$f(x')-f(x)-\langle y, x'-x\rangle \geq 0 \qquad \forall x'\in X.$$

Thus, $y\in S$. $\square$

Subdifferentials provide insight into convex conjugates via the following results...

Prop: Given $f : X \to \mathbb{R} \cup \{+\infty\}$ proper and convex,
$y_0 \in \partial f(x_0) \iff f(x_0) + f^*(y_0) = \langle y_0, x_0 \rangle$

"equality holds in Young's inequality"

Pf:

$$y_0 \in \partial f(x_0)$$
$$\Updownarrow$$
$$f(x') \geq f(x_0) + \langle y_0, x' - x_0 \rangle \quad \forall \; x' \in X$$
$$\Updownarrow$$
$$\langle y_0, x_0 \rangle - f(x_0) \geq \langle y_0, x' \rangle - f(x') \quad \forall \, x' \in X$$
$$\Updownarrow$$
$$\langle y_0, x_0 \rangle - f(x_0) \geq f^*(y_0)$$
$$\Updownarrow$$
$$\langle y_0, x_0 \rangle = f(x_0) + f^*(y_0)$$

Prop: Suppose f is convex and it is lsc at $x_0 \in D(f)$. Then $x_0 \in \partial f^*(y_0) \Rightarrow y_0 \in \partial f(x_0)$

Pf: By Fenchel-Moreau, $f(x_0) = f^{**}(x_0)$. Furthermore, by previous prop,

$$\langle y_0, x_0 \rangle = f^*(y_0) + f^{**}(x_0) = f^*(y_0) + f(x_0).$$

Thus, by previous prop again, $y \in \partial f(x_0)$. $\square$

Thm: Suppose f is proper, convex, and lsc. Then $y \in \partial f(x) \Leftrightarrow x \in \partial f^*(y)$.

Pf: By previous prop, we have $\Leftarrow$.

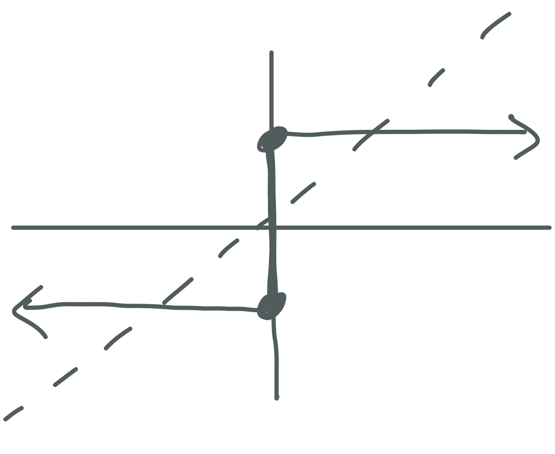
By Fenchel-Moreau, $f = f^{**}$. So if $y \in \partial f(x) \Rightarrow y \in \partial f^{**}(x) \underset{\text{prev. prop}}{\Rightarrow} x \in \partial f^*(y)$. $\square$

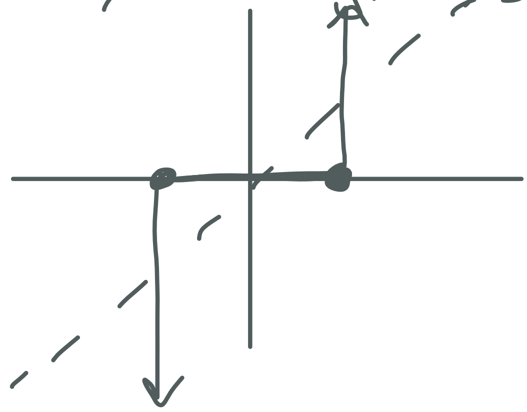
Example: Intuitive understanding of convex conjugates

$(X, \|\cdot\|) = (\mathbb{R}, |\cdot|)$

$f(x)$	$\partial f(x)$	$\partial f^*(y)$	$f^*(y)$
e^x	$\{e^x\}$	$\begin{cases} \emptyset & \text{if } y \le 0 \\ \{\log(y)\} & \text{if } y > 0 \end{cases}$	$\begin{cases} y\log y - y & \text{if } y > 0 \\ 0 & \text{if } y = 0 \\ +\infty & \text{if } y < 0 \end{cases}$

$f(x)$	$\partial f(x)$	$\partial f^*(y)$	$f^*(y)$
$\lvert x \rvert$	$\begin{cases} \operatorname{sgn}(x) & x \neq 0 \\ [-1,1] & x = 0 \end{cases}$	$\begin{cases} 0 & \text{if } y \in (-1,1) \\ (-\infty, 0] & \text{if } y = -1 \\ [0, +\infty) & \text{if } y = 1 \\ \emptyset & \text{if } y \notin [-1,1] \end{cases}$	$\chi_{[-1,1]}(y) = \begin{cases} 0 & \text{if } y \in [-1,1] \\ +\infty & \text{otherwise} \end{cases}$





Def: For $C \subseteq X$, its characteristic fn is

$$\chi_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$$

Exercise:

- If C is closed, χ_C is lsc.

- If C is convex, χ_C is convex.

Primal / Dual Optimization Problems

Goal of convex optimization: given f convex, C convex, solve

$$\inf_{x \in C} f(x) = \inf_{x \in X} f(x) + \chi_C(x)$$

Q's:

① What is the value of the infimum?
② Is inf attained?
③ Unique minimizer?
④ Characterize minimizer?

Key trick: observe the behavior of this optimization problem under **perturbations**.

Def: Given normed vector spaces X and U and a convex function $F : X \times U \to \mathbb{R} \cup \{+\infty\}$.

primal problem: $\mathcal{P}_0 := \inf_{x \in \mathcal{X}} f(x)$, $f(x) = F(x,0)$

dual problem: $\mathcal{D}_0 := \sup_{v \in \mathcal{U}^*} g(v)$, $g(v) = -F^*(0,v)$.

The function $F(x,u)$ encodes the perturbations of $f(x)$ that we consider.

We seek a "simple" $F(x,u)$ so that either $\mathcal{P}_0$ or $\mathcal{D}_0$ coincide with our problem.

Remark: $\mathcal{X} \times \mathcal{U}$ is a nvs with dual $\mathcal{X}^* \times \mathcal{U}^*$ and duality pairing

$$\langle (y,v),(x,u) \rangle = \langle y,x \rangle + \langle v,u \rangle$$