Lecture 7

Recall:

So we solved Kantorovich's problem...

... how does this help us solve Monge's problem?

via the Kantorovich dual problem.

Crash course in convex analysis and optimization

Let $\mathcal{X}$ be a n.v.s.

Exercise 11: Given a collection of functions $f_\alpha : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $\alpha \in \mathcal{A}$,

- if f_α are convex, then $\sup_\alpha f_\alpha$ is convex
- if f_α are lsc, then $\sup_\alpha f_\alpha$ is lsc

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

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

Prop (Young's Inequality): Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper,

$$f^*(y) + f(x) \geq \langle y, x \rangle \quad \forall x \in X, y \in X^*$$

Lemma: For $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, f^* is convex and lsc.

Def: $\mathrm{dom}(f) = \{x \in X : f(x) < +\infty\}$.

Def: Given $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ s.t. f and f^* are proper, its biconjugate $f^{**}: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is

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

Rmk:

- $f^{**}(x) + f^*(y) \geq \langle y, x \rangle \quad \forall x \in \mathcal{X}, y \in \mathcal{X}^*$
- $f(x) \geq f^{**}(x) \quad \forall x \in \mathcal{X}$
- f^{**} always convex, lsc

Thm: (Fenchel-Moreau)
Given $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper,
(i) f is convex and lsc

$$\Updownarrow$$

f^* is proper and $f = f^{**}$

(ii) If f is convex and $f(x_0) < +\infty$,

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 $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, $x \in \mathrm{dom}(f)$, the subdifferential of f at x is the set valued operator

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

If $x \notin \mathrm{dom}(f)$, $\partial f(x) = \emptyset$.

Def: Given $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, $x \in \mathrm{dom}(f)$, f is Gâteaux differentiable at x if $\exists\, y \in \mathcal{X}^*$ s.t.

$$\lim_{h \downarrow 0} \frac{f(x + h\tilde{x}) - f(x)}{h} = \langle y, \tilde{x} \rangle, \quad \forall \tilde{x} \in \mathcal{X}.$$

We denote y by $Df(x)$.

Next time: example Gâteaux vs Frechet

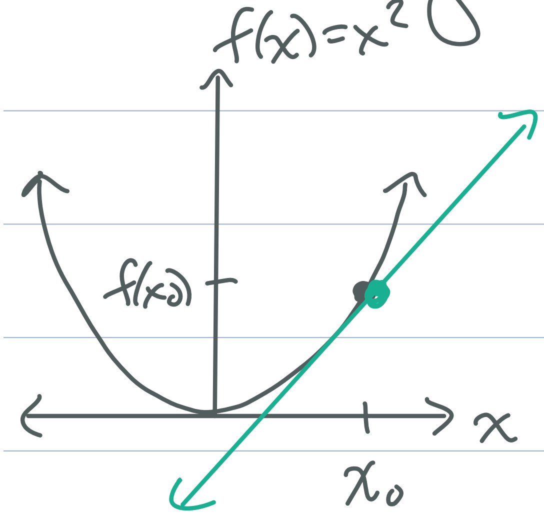
Thm: If f is Gâteaux differentiable at $x \in \mathrm{dom}(f)$, then $\partial f(x) = \{Df(x)\}$.

Rmk: If f is convex and cts $x \in \mathrm{dom}(f)$, then f is Gâteaux diff at x iff $\partial f(x) = \{Df(x)\}$.

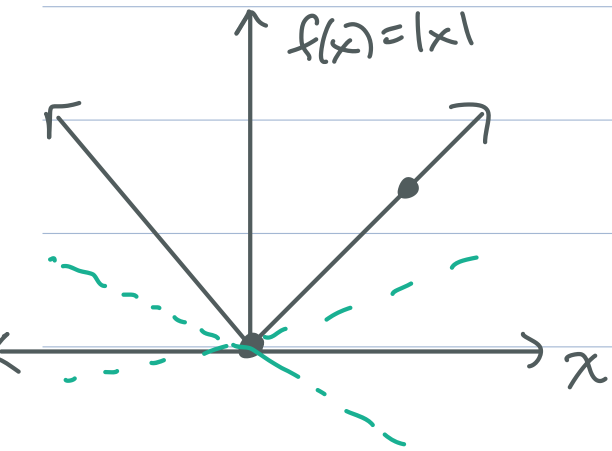
Prop: If $f : X \to \mathbb{R} \cup \{+\infty\}$ is proper and convex and $x \in \mathrm{dom}(f)$,

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

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



$x' \mapsto \underline{f(x_0)} + \langle y, \underline{x' - x_0} \rangle$ is the line passing through $(x_0, f(x_0))$ with slope y



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

A fact we will use in the proof:
If $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, convex, then any local minimizer of f is a global minimizer.
(Exercise 14)

Pf: By defn, $S \subseteq \partial f(x)$.
Now, show opposite containment.
Fix $y \in \partial f(x)$.

Fix $\varepsilon > 0$. Define
$$\psi(x') := f(x') - f(x) - \langle y, x' - x \rangle + \varepsilon \|x' - x\|$$

Since $y \in \partial f(x)$,
$$\lim_{\delta \to 0} \inf_{\{x' : 0 < \|x' - x\| < \delta\}} \frac{\psi(x')}{\|x' - x\|} \geq \varepsilon$$

In particular, $\exists\ \delta_0 > 0$ s.t.
$$\inf_{\{x' : 0 < \|x' - x\| < \delta_0\}} \frac{\psi(x')}{\|x' - x\|} \geq \frac{\varepsilon}{2}$$

Thus, for all $x' \in \{x' : 0 < \|x' - x\| < \delta_0\}$, $\psi(x') > 0$. Also, by def, $\psi(x) = 0$.

Thus x is a local min of ψ. Since ψ is convex, x is a global min of ψ, i.e., for all $x' \in X$

$$0 = \psi(x) \leq \psi(x') = f(x') - f(x) - \langle y, x' - x \rangle + \varepsilon \|x' - x\|.$$

Sending $\varepsilon \to 0$ gives

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

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

Subdifferentials provide insight into convex conjugation via following results...

Suppose $f: X \to \mathbb{R} \cup \{+\infty\}$ is proper and convex and $x_0 \in \operatorname{dom}(f)$.

<u>Prop</u>: $y \in \partial f(x_0) \iff f(x_0) + f^*(y_0) = \langle y_0, x_0 \rangle$

$$\implies x_0 \in \partial f^*(y_0)$$

<u>Prop</u>: If f is lsc at x_0,

$x_0 \in \partial f^*(y_0) \implies y \in \partial f(x_0)$

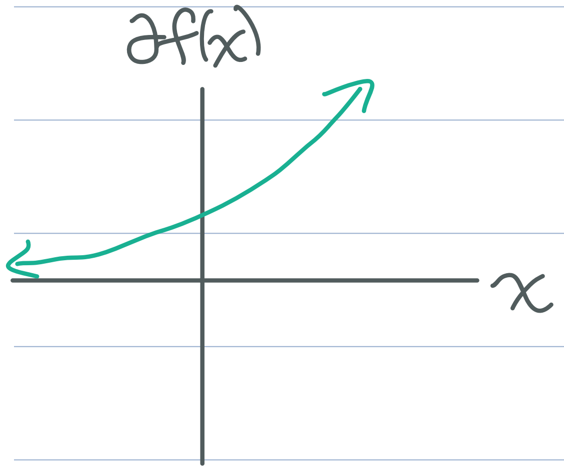
<u>Thm</u>: If f is lsc everywhere,

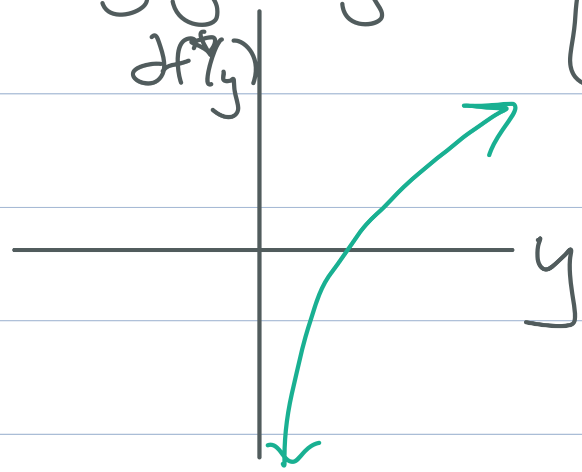
$$y \in \partial f(x) \iff x \in \partial f^*(y).$$

Example: Intuitive understanding of convex conjugates

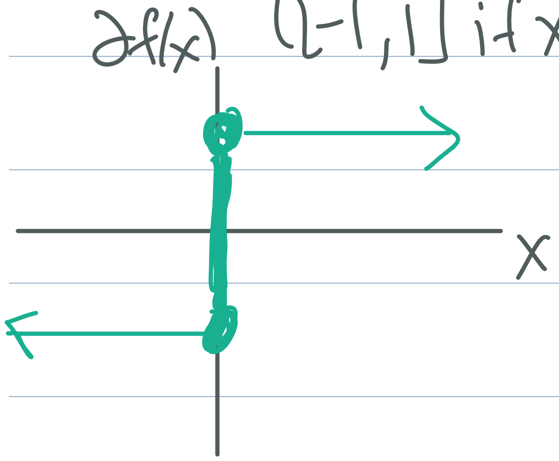
$(X, \|\cdot\|) = (\mathbb{R}, |\cdot|)$

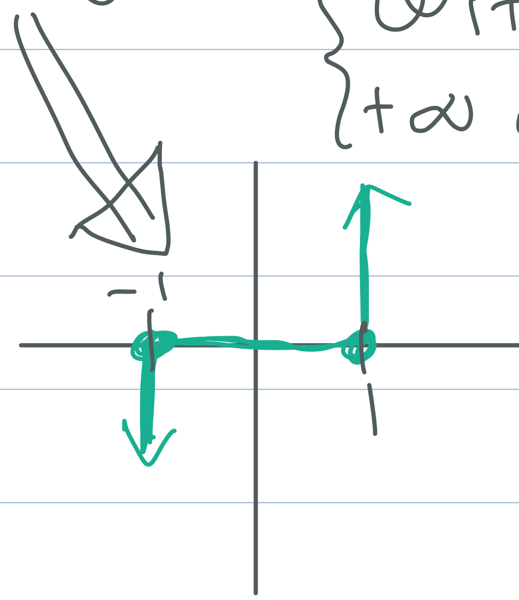
$f(x)$	$\partial f(x)$	$\partial f^*(y)$	$f^*(y)$
e^x	$\{e^x\}$	$\begin{cases} \emptyset & \text{if } y \leq 0 \\ \log(y) & \text{if } y > 0 \end{cases}$	$\begin{cases} +\infty & \text{if } y < 0 \\ 0 & \text{if } y = 0 \\ y\log y - y & \text{if } y > 0 \end{cases}$





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





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

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

Fact: C closed $\Rightarrow$ χ_C is lsc

C convex $\Rightarrow$ 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.

perturbations

Def: Given nvs $\mathcal{X}$ and $\mathcal{U}$ and a proper and convex function $F: \mathcal{X} \times \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$,

primal problem: $P_0 := \inf_{x \in \mathcal{X}} f(x)$, $f(x) = F(x, 0)$

dual problem: $D_0 := \sup_{v \in \mathcal{U}^*} g(v)$, $g(v) = -F^*(0, v)$

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

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

By the way we have defined our primal and dual problems, Young's inequality ensures...

$$F(x,u) + F^*(y,v) \geq \langle y,x \rangle + \langle v,u \rangle$$

In particular,
$f(x) - g(v) \geq 0 \quad \forall\, x \in \mathcal{X},\ v \in \mathcal{U}^*$

Thus, we always have $\mathcal{P}_0 \geq D_0$.

Thus, we will seek conditions on F that ensure $\mathcal{P}_0 = D_0$, i.e. "there is no duality gap"

Thm (Equivalence of Primal + Dual)
Given $F: \mathcal{X} \times \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$ proper, convex, suppose $\mathcal{P}_0 < +\infty$. ← "primal problem is feasible"

Define the inf-projection $\mathcal{P}(u) = \inf_{x \in \mathcal{X}} F(x, u)$.

Then,

(i) $P_0 = D_0 \iff P$ is lsc at $u=0$.

(ii) $P_0 = D_0$ and a maximizer of dual problem exists $\iff \partial P(0) \neq \emptyset$.