Lecture 8

Recall:

<u>Def</u>: Given $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, $x \in \mathrm{dom}(f)$, the Frechet 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 $x \notin \mathrm{dom}(f)$, $\partial f(x) = \emptyset$.

$$y \in \partial f(x) \iff \liminf_{x' \to x} \frac{f(x') - f(x) - \langle y, x'-x \rangle}{\|x'-x\|} \geq 0$$

$$\iff \liminf_{z \to 0} \frac{f(x+z) - f(x) - \langle y, z \rangle}{\|z\|} \geq 0$$

... compare to....

<u>Def</u>: Given $f:\mathcal{X}\to\mathbb{R}\cup\{+\infty\}$ proper, $x\in\operatorname{dom}(f)$, the <u>Gâteaux subdifferential</u> of f at x is the set valued operator

$$\partial^G f(x) = \left\{ y\in\mathcal{X}^* : \liminf_{h\to 0} \frac{f(x+\overbrace{hx'}^{z})-f(x)}{h} \geq \langle y,x'\rangle \ \overset{\text{forall}}{\underset{}{\forall}}\, x'\in\mathcal{X} \right\}$$

<u>Rmk</u> $\partial f(x) \subseteq \partial^G f(x)$

To see this, suppose $y\in\partial f(x)$. Then, for any $x'\in\mathcal{X}\setminus\{0\}$, $z := hx' \xrightarrow{h\to 0} 0$, so

$$\liminf_{z\to 0} \frac{f(x+\overset{hx'}{z})-f(x)-\overbrace{\langle y,z\rangle}^{h\langle y,x'\rangle}}{\underbrace{\|z\|}_{h\|x'\|}} \geq 0$$

$$\Downarrow$$

$$\liminf_{h\to 0} \left(\frac{f(x+hx')-f(x)}{h} - \langle y,x'\rangle\right)\frac{1}{\|x'\|} \geq 0$$

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

$$\lim_{z \to 0} \frac{f(x+z) - f(x) - \langle y, z \rangle}{\|z\|} = 0.$$

We denote y by $Df(x)$.

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.

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

We denote y by $D^G f(x)$. Rmk: $D^G f(x) \in \mathcal{X}^*$

Rmk: Frechet diff $\Rightarrow$ Gâteaux.

Thm: If f is Gâteaux differentiable at $x \in \operatorname{dom}(f)$, then $\partial^G f(x) = \{D^G f(x)\}$. If f is Frechet differentiable at $x \in \operatorname{dom}(f)$, then $\partial f(x) = \{Df(x)\}$

Pf: Exercise 18

Rmk: If f is proper, lsc, and convex, it is Frechet differentiable iff it is Gateaux differentiable.

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

Exercise 13 shows that convexity is necessary.

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

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

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

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

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

<u>Prop</u>: If f is lsc at x_0,
$x_0 \in \partial f^*(y_0) \Rightarrow y \in \partial f(x_0)$

<u>Thm</u>: If f is lsc everywhere,
$y \in \partial f(x) \Leftrightarrow x \in \partial f^*(y)$.

<u>Def</u>: 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 $\Rightarrow \chi_C$ is lsc
C convex $\Rightarrow \chi_C$ 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: $D_0 := \sup_{v \in \mathcal{U}^*} g(v)$, $g(v) = -F^*(0, v)$

we always have $P_0 \geq D_0$.

Thus, we will seek conditions on F that ensure $P_0 = D_0$, 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 $P_0 < +\infty$. ← "primal problem is feasible"

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

Then,

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

(ii) $P_0 = D_0$ and v^* is a maximizer of dual problem $\iff v^* \in \partial P(0)$.

Pf:

Step 1: Show $P(u)$ is proper, convex.

Since $P(0) = \inf_{x \in \mathcal{X}} F(x, 0) = P_0 < +\infty$,

$\mathcal{P}$ is proper.

$$\mathcal{P}(\overbrace{(1-\alpha)u_0+\alpha u_1}^{u_\alpha})$$

$$= \inf_{x\in\mathcal{X}} F(x, u_\alpha)$$

$$\le F((1-\alpha)x_0+\alpha x_1, u_\alpha) \quad \text{for any } (1-\alpha)x_0+\alpha x_1 \in \mathcal{X}$$

$$= F((1-\alpha)x_0+\alpha x_1, (1-\alpha)u_0+\alpha u_1) \quad F \text{ convex}$$

$$\le (1-\alpha)F(x_0, u_0) + \alpha F(x_1, u_1)$$

Taking inf over $x_0 \in \mathcal{X}$, $x_1 \in \mathcal{X}$

$$\mathcal{P}(u_\alpha) \le (1-\alpha)\mathcal{P}(u_0) + \alpha \mathcal{P}(u_1)$$

<u>Step 2</u>: Prove part (i).

By defn,

$$\mathcal{P}^*(v) = \sup_{u\in\mathcal{U}} \langle v, u\rangle - \mathcal{P}(u)$$

$$= \sup_{u\in\mathcal{U}} \sup_{x\in\mathcal{X}} \langle 0, x\rangle + \langle v, u\rangle - F(x, u)$$

$= F^*(0, v)$
$= -g(v)$

Also, $P^{**}(u) = \sup_{v \in \mathcal{U}^*} \langle v, u \rangle + g(v)$

$\Rightarrow P^{**}(0) = \sup_{v \in \mathcal{U}^*} g(v) = D_0$.

Since P is convex, $0 \in \operatorname{dom}(P)$, Fenchel-Moreau ensures
$P_0 = D_0 \Leftrightarrow P(0) = P^{**}(0) \stackrel{?}{\Leftrightarrow} P$ is lsc at 0

<u>Step 3</u>: We now show (ii).
Assume $\boxed{v_* \in \partial P(0)}$. Since P is convex, $0 \in \operatorname{dom}(P)$, for any sequence $u_n \to 0$, we have

$$\liminf_{n\to\infty} \Phi(u_n) \geq \liminf_{n\to\infty} \Phi(0) + \langle v_*, u_n - 0 \rangle$$
$$= \Phi(0).$$

Thus Φ is lsc at zero, so by previous part, $P_0 = D_0$

Furthermore, $v_* \in \partial\Phi(0)$ implies equality holds in Young's inequality,

$$\Phi(0) + \Phi^*(v_*) = \langle v_*, 0 \rangle = 0,$$

so $\sup_{v \in U^*} g(v) = D_0 = P_0 = \Phi(0) = -\Phi^*(v_*) = g(v_*)$.

It remains to prove the converse. Suppose $P_0 = D_0$ and v_* is a maximizer of dual problem.

By (i), φ is lsc at 0. By Fenchel - Moreau,

$$
\begin{aligned}
\varphi(0) = \varphi^{**}(0) &= \sup_{v \in \mathcal{U}^*} g(v) \\
&= g(v_*) \\
&= -\varphi^*(v_*) \\
&= \langle 0, v_* \rangle - \varphi^*(v_*)
\end{aligned}
$$

Thus, equality holds in Young's inequality, so $v_* \in \partial \varphi(0)$. $\square$

Kantorovich Duality

(X, d_X), (Y, d_Y) Polish spaces
$\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$
$c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ lsc, bdd below

$$\min_{\gamma : \gamma \in \Gamma(\mu,\nu)} \overbrace{\int_{X \times Y} c(x,y) \, d\gamma(x,y)}^{\mathbb{K}(\gamma)} \qquad \text{(KP)}$$

This is a convex optimization problem.
To find its dual...

① Rewrite as unconstrained optimization problem.
② Identify "perturbation" function $F(x,u)$ so that (KP) $= D_0$.

We will do this via introducing a Lagrange multiplier.

Recall: Lagrange multipliers in Calculus...

Given $A \in M_{m\times n}(\mathbb{R})$, $b \in \mathbb{R}^m$

$$\inf_{Ax=b} f(x) = \inf_{x\in\mathbb{R}^n} \underbrace{f(x) + \chi_{\{x : Ax=b\}}(x)}$$

$$= \inf_{x\in\mathbb{R}^n} \underbrace{\sup_{\lambda\in\mathbb{R}^m} f(x) + \langle \lambda, Ax-b\rangle}$$

Relation to Primal/Dual Problem:

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

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

We can also write dual problem as a saddle point problem:

$$g(v) = -F^*(0,v) = -\sup_{(x,u)\in \mathcal{X}\times\mathcal{U}} \langle 0,x\rangle + \langle v,u\rangle - F(x,u)$$

$$= \inf_{(x,u)\in \mathcal{X}\times\mathcal{U}} F(x,u) - \langle v,u\rangle$$

$$\Rightarrow D_0 = \sup_{v\in\mathcal{U}^*} \inf_{(x,u)\in\mathcal{X}\times\mathcal{U}} F(x,u) - \langle v,u\rangle$$

Moral: Introducing a Lagrange multiplier to remove constraint can shed light on how to choose perturbation function $F(x,u)$.

|

$=$

How to do this for (KP)?

Recall: $\Gamma(\mu,\nu) = \{\gamma \in \mathcal{P}(X \times Y) : \pi_1 \# \gamma = \mu, \pi_2 \# \gamma = \nu\}$

Want to introduce Lagrange multiplier to enforce this constraint.

Fact: If X is locally cpt, $\mu \in \mathcal{M}(X)$, then $C_c(X)$ are dense in $L^1(\mu)$. (See Folland Thm 7.8, 7.9.)

Lemma: Given $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, $\gamma \in \mathcal{M}(X \times Y)$,

$$\chi_{\Gamma(\mu,\nu)}(\gamma) = \sup_{\varphi \in C_b(X),\, \psi \in C_b(Y)} \langle \mu - \pi^1 \# \gamma, \varphi \rangle + \langle \nu - \pi^2 \# \gamma, \psi \rangle$$

Therefore, we may rewrite (KP) as the following saddle point problem:

$$D_0 = \sup_{v \in \mathcal{U}^*} \inf_{(x,u) \in \mathcal{X} \times \mathcal{U}} F(x,u) - \langle v, u \rangle$$

What is the corresponding primal problem?

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

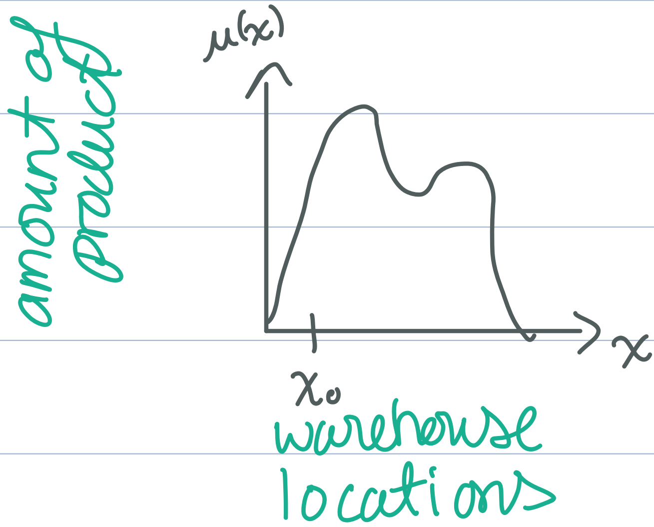
Optimal Transport
MAD LIBS!

Name of retailer:

Type of product:

Worthy cause:

The Shipper's Problem (Caffarelli)



- It costs ____________ $c(x,y)$ dollars to move one ______________ from x to y.

- You want to make extra \$\$\$ to support ____

- You charge ______________ $\varphi(x)$ dollars to pick up one ______________ from location x and $\psi(y)$ dollars to deliver to y.

Obviously, if ______________ will let you ship, the following must be true:

$$\varphi(x) + \psi(y) \leq c(x, y).$$

Now: prove $\mathcal{P}_0 = D_0$ for (KP)

Suppose X, Y cpt Polish spaces.

Thm: For all $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$,

$$\underbrace{\inf_{\gamma \in \Gamma(\mu,\nu)} \mathbb{K}(\gamma)}_{= -D_0} = \underbrace{\sup_{\substack{(\varphi,\psi) \in C(X)\times C(Y) \\ \varphi \oplus \psi \le c}} \int \varphi \, d\mu + \int \psi \, d\nu}_{= -P_0}$$