Lecture 7

I will attempt to have a sychronous zoom option starting next week.

Recall:

We seek to prove existence of solutions to:

Kantorovich's Optimal Transport Problem

Given $\mu, \nu \in \mathcal{P}(X)$, solve

$$\min_{\gamma : \gamma \in \Pi(\mu,\nu)} \underbrace{\int_{X \times X} d(x^1, x^2)^p \, d\gamma(x^1, x^2)}_{\mathbb{K}_p(\gamma) :=} , \quad p \geq 1$$

Lemma: Given Polish spaces (X, d_X), (Y, d_Y), and $\{\mu_n\}_{n=1}^{+\infty} \subseteq \mathcal{P}(X)$ narrowly converging to μ, then for any continuous function $t : X \to Y$, $t\#\mu_n$ narrowly converges to $t\#\mu$.

Prop: Given a Polish space (X, d) and $\mu, \nu \in \mathcal{P}(X)$, $\Pi(\mu,\nu)$ is compact in the narrow topology.

Lemma: Suppose $g: X \to \mathbb{R} \cup \{+\infty\}$ is lsc and bounded below. Then $\exists \{g_k\}_{k=1}^{+\infty} \subseteq C_b(X)$ s.t. $\lim_{k \to +\infty} g_k(x) \nearrow g(x) \ \forall x \in X$.

Thm (Portmanteau): For any $g: X \to \mathbb{R} \cup \{+\infty\}$ lsc and bounded below, the functional $\mu \mapsto \int g \, d\mu$ is lsc wrt narrow conv in $P(X)$, that is

$$\mu_n \to \mu \text{ narrowly} \Rightarrow \varliminf \int g \, d\mu_n \geq \int g \, d\mu.$$

Combining these results, by the Direct Method of the Calculus of Variations, we obtain...

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

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

In particular, taking $c(x^1,x^2) = d(x^1,x^2)^p$, we obtain that a solution to Kantorovich's problem exists.

Thus, for any $\mu,\nu \in \mathcal{P}(X)$, there exists an optimal transport plan γ_*

Why was narrow topology the "right" topology?

Recall:

(X,d) separable locally compact metric space

Banach space	$(C_0(X), \Vert\cdot\Vert_\infty)$	$(C_b(X), \Vert\cdot\Vert_\infty)$
Dual space	$(\mathcal{M}^s(X), \Vert\cdot\Vert_{TV})$	[big space, containing $(\mathcal{M}(X), \Vert\cdot\Vert_{TV})$]
Weak-* topology	wide topology	narrow topology

$\Gamma(\mu,\nu)$ is narrowly cpt $\Rightarrow$ widely cpt

But the objective function may not be lsc in wide topology.

Let $c(x^1,x^2) = d(x^1,x^2) - 1$.

Consider $\gamma_n := \delta_{(n,n)} \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$.

Exercise: $\gamma_n \to \underset{\gamma}{\overset{\shortparallel}{0}}$ in the wide topology

However, $\underline{\lim} \int c(x^1,x^2)\, d\gamma_n(x^1,x^2)$

$$\begin{aligned} &= \underline{\lim} \int (d(x^1,x^2) - 1)\, d\gamma_n(x^1,x^2) \\ &= -1 \\ &< 0 \\ &= \int c(x^1,x^2)\, d\gamma(x^1,x^2). \end{aligned}$$

Moral: Wide convergence can allow mass to escape to infinity.

However...

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

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

Then, $\mu_n \to \mu$ narrowly.

In particular,
wide convergence + mass conservation $\Leftrightarrow$ narrow conv.

Pf: Fix $f \in C_b(X)$.

increasing sequence of compact sets

Since $\{\mu\}$ is tight, $\forall k \in \mathbb{N}$, $\exists K_k \subset\subset X$ s.t. $\mu(X \setminus K_k) \leq \frac{1}{k}$. By Tietze extension theorem, $\exists \eta_k \in C(X)$ s.t. $\eta_k \equiv 1$ on K_k, $\operatorname{supp} \eta_k \subset\subset X$, $0 \leq \eta_k \leq 1$ and $\eta_k \nearrow 1$ pointwise.

Then, $\forall k \in \mathbb{N}$,

$$\begin{aligned}\liminf_{n\to\infty} \int f + \|f\|_\infty \, d\mu_n &\geq \liminf_{n\to\infty} \int (f + \|f\|_\infty)\eta_k \, d\mu_n \\ &= \int (f + \|f\|_\infty)\eta_k \, d\mu\end{aligned}$$

$$\begin{aligned}\limsup_{n\to\infty} \int f - \|f\|_\infty \, d\mu_n &\leq \limsup_{n\to\infty} \int (f - \|f\|_\infty)\eta_k \, d\mu_n \\ &= \int (f - \|f\|_\infty)\eta_k \, d\mu\end{aligned}$$

Finally, by conservation of mass,

$$\int f\,d\mu = \|f\|_\infty + \int f - \|f\|_\infty\,d\mu$$

$$\overset{MCT}{=} \|f\|_\infty + \lim_{k\to\infty} \int (f - \|f\|_\infty)\eta_k\,d\mu$$

$$\geq \|f\|_\infty + \limsup_{n\to\infty} \int f - \|f\|_\infty\,d\mu_n$$

$$= \limsup_{n\to\infty} \int f\,d\mu_n$$

$$\geq \liminf_{n\to\infty} \int f\,d\mu_n$$

$$= -\|f\|_\infty + \liminf_{n\to\infty} \int f + \|f\|_\infty\,d\mu_n$$

$$\overset{MCT}{\geq} -\|f\|_\infty + \lim_{k\to\infty} \int (f + \|f\|_\infty)\eta_k\,d\mu$$

$$= -\|f\|_\infty + \int f + \|f\|_\infty\,d\mu$$

$$= \int f\,d\mu$$

Thus, equality must hold throughout. Since f was arbitrary, $\mu_n \to \mu$ narrowly. $\square$

Rmk: The distinction between narrow and wide convergence is especially important when $\mathcal{X}$ is not locally compact, e.g. $\mathcal{X} = C([0,1];\mathbb{R})$ or $\mathcal{X} = \mathcal{P}(\mathbb{R}^d)$.

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 $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a $\underbrace{\text{normed vector space}}_{\text{nvs}}$.

y

||

Let $(\mathcal{X}^*, \|\cdot\|_{\mathcal{X}^*})$ be its dual space, that is, the set of all bounded linear functionals on $\mathcal{X}$ with

$$\|y\|_{\mathcal{X}^*} = \sup_{\substack{x \in \mathcal{X} \\ \|x\| \le 1}} y(x)$$

Given $x \in \mathcal{X}$, $y \in \mathcal{X}^*$, let $\langle y, x \rangle =: y(x)$

Exercise:

- For any collection of convex functions $(f_i)_{i \in I}$ on $\mathcal{X}$, $\sup_{i \in I} f_i(x)$ is convex.
- For any collection of lsc functions $(f_i)_{i \in I}$ on $\mathcal{X}$, $\sup_{i \in I} f_i(x)$ is lsc.

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

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

Ex: Suppose $\mathcal{X} = \mathbb{R}$ and $f(x) = e^x$.
Then
$$f^*(y) = \sup_{x\in\mathbb{R}} \{yx - e^x\} = -\inf_{x\in\mathbb{R}} \{e^x - yx\}.$$

If $y<0$, $f^*(y) = +\infty$.
If $y=0$, $f^*(y) = 0$.

If $y>0$... then $x \mapsto e^x - yx$ is a convex, differentiable fn, so a critical point is a global minimizer

$$e^{x_*} - y = 0 \iff x_* = \log(y) \iff -\{e^{x_*} - yx_*\} = -y + y\log y$$

Thus,
$$f^*(y) = \begin{cases} +\infty & \text{if } y<0, \\ 0 & \text{if } y=0, \\ y\log y - y & \text{if } y>0. \end{cases}$$

Exercise: If $f(x) = \frac{1}{p}|x|^p$, find $f^*(y)$.

An immediate consequence of the defn is...

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

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

Another immediate consequence of the defn and the above exercises is...

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

In a similar way, we may define

<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

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

Note that, for all f proper, $x \in X$,

$$f^*(y) + f(x) \geq \langle y, x \rangle \quad \forall y \in X^* \text{ (Young)}$$
$$\Updownarrow$$
$$f(x) \geq \langle y, x \rangle - f^*(y) \quad \forall y \in X^*$$
$$\Updownarrow$$
$$f(x) \geq f^{**}(x)$$

Note that, since f^{**} is always convex and lower semicontinuous, a necessary condition for $f = f^{**}$ is that f is convex and lsc.
In fact, this is sufficient!

Thm (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 $x_0 \iff f(x_0) = f^{**}(x_0)$.

Pf: by Hahn-Banach :) $\square$

Exercise:
- If x_0 is a minimizer of f, what is $f^*(0)$?
- If x_0 is a local minimizer of f, that is, there exists a neighborhood $\mathcal{U}$ of x_0 s.t. $f(x_0) \leq f(y)$ $\forall y \in \mathcal{U}$, and f is convex, prove that x_0 is