Lecture 5

Recall:

Kantorovich's Optimal Transport Problem

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

$c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ lower semicts and bdd below

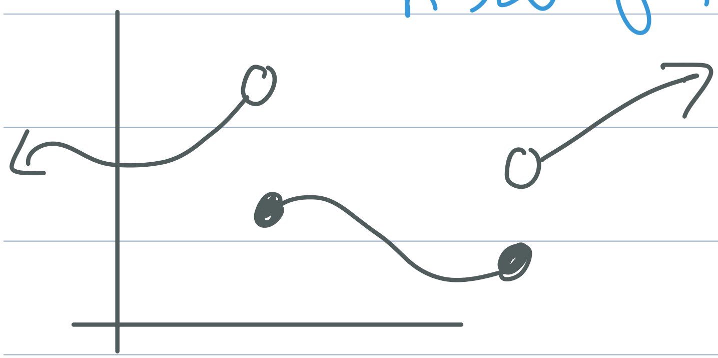
$$\min_{\gamma : \gamma \in \Gamma(\mu,\nu)} \overbrace{\int_{X \times Y} \underbrace{c(x,y)}_{\substack{\text{cost of moving mass} \\ \text{from location } x \text{ to} \\ \text{location } y}} d\gamma(x,y)}^{\mathbb{K}_c(\gamma)}$$

If γ_* attains the minimum, we will call it an <u>optimal transport plan</u>.

Recall... given a metric space (Z, d_Z),

Def: $f : Z \to \mathbb{R} \cup \{+\infty\}$ is <u>lower semicontinuous</u> if $z_n \xrightarrow{d} z$ implies

$$\liminf_{n\to\infty} f(z_n) \geq f(z)$$



Reasons Kantorovich Problem Better:

① the constraint set is always nonempty
② the constraint set is convex (Ex 8)
③ the objective fn is linear, hence convex (Ex 9)

Remark: Since Kantorovich's problem is the minimization of a linear objective function, subject to affine linear equality and inequality constraints, it is a linear program... hence is a convex optimization problem.

Ex (finite dim'l linear program)

$$\min_x \quad b_2^t x$$

$$A_0 x = b_0$$
$$A_1 x \geq b_1$$

4. Kantorovich's problem has a dual
5. We can easily prove existence of solns to Kantorovich prob via...

The Direct Method of the Calculus of Variations

Setup: $C \subseteq X$, (X, d) metric space

Goal: prove that $\min_{x \in C} F(x)$ exists

$F : X \to \mathbb{R} \cup \{+\infty\}$

Step 0: Show $\inf_{x \in C} F(x) < +\infty$.

that is, the problem is "feasible."

Step 1: Take a minimizing sequence.

That is, choose $\{x_n\}_{n=1}^{\infty} \subseteq C$ s.t.

$$\lim_{n \to \infty} F(x_n) = \inf_{x \in C} F(x).$$

Step 2: Prove that C is compact,

so $\exists$ subsequence $\{x_{n_k}\}_{k=1}^{\infty} \subseteq C$

so that $x_{n_k} \xrightarrow{k \to \infty} x_* \in C$

Step 3: Prove that F is lsc, so

$$\inf_{x \in C} F(x) = \lim_{n \to \infty} F(x_n) = \lim_{k \to \infty} F(x_{n_k})$$

$$= \ldots = \liminf_{k \to \infty} F(x_{n_k}) \geq F(x_*)$$

Thus x_* is a minimizer of F on C.

Key challenge: choosing the right metric d on C.

- weak enough to get compactness
- strong enough to get lsc

What is the right topology for Kantorovich's problem?

First: consider compactness of constraint set.

Thm (Prokhorov)

Given a Polish space (Z, d_Z) and $\mathcal{K} \subseteq \mathcal{P}(Z)$,

- $\mathcal{K}$ is relatively compact in narrow topology

$\Updownarrow$

- $\mathcal{K}$ is tight, $\forall \varepsilon > 0, \exists K_\varepsilon \subset\subset Z$ s.t. $\mu(Z \setminus K_\varepsilon) \leq \varepsilon \;\forall \mu \in \mathcal{K}$

An immediate corollary is...

Cor: If (Z, d_Z) is a Polish space, then for any $\sigma \in \mathcal{P}(Z)$, $\{\sigma\}$ is tight.

Other key step...

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

Pf: See exercise 5.

Prop: Given Polish spaces (X, d_X) and (Y, d_Y), $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, then $\Gamma(\mu, \nu)$ is compact in the narrow topology.

Next: want to show

$\mathbb{K}_p(\gamma)$ is lsc in narrow topology.

Lemma: Given metric space (Z, d_Z), suppose $g: Z \to \mathbb{R} \cup \{+\infty\}$ is lsc and bdd below. Then $\exists$ $\{g_k\}_{k=1}^{\infty} \subseteq C_b(Z)$ s.t.

$$g_k(z) \xrightarrow{k\to\infty} g(z), \quad \forall z \in Z,$$

and $\inf g_k \geq \min\{\inf g, 0\}$ $\forall k \in \mathbb{N}$.

Pf: Exercise 9, via Moreau-Yosida regularization.

Thm (Portmanteau): For any $g: Z \to \mathbb{R} \cup \{+\infty\}$ lsc, bdd below, the functional

$$\underset{\in P(Z)}{\mu} \mapsto \underset{\in \mathbb{R} \cup \{+\infty\}}{\int_Z g \, d\mu}$$

is lsc wrt narrow topology.

Pf: The result is trivially true if $g \equiv +\infty$, so assume WLOG $\inf g < +\infty$.

By lemma, $\exists \{g_k\}_{k=1}^{\infty} \subseteq C_b(Z)$ s.t. $g_k \nearrow g$ pointwise.

Fix $\{\mu_n\}_{n=1}^{\infty} \subseteq \mathcal{P}(Z)$ converging narrowly to $\mu \in \mathcal{P}(Z)$.

For any $k \in \mathbb{N}$,

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

$$= \int g_k \, d\mu.$$

Thus,

$$\liminf_{n\to\infty} \int g \, d\mu_n \geq \liminf_{k\to\infty} \int g_k \, d\mu$$

$\int d\mu = 1$

$c := \min\{\inf g, 0\}$

$$= \liminf_{k\to\infty} \int g_k - c + c \, d\mu$$

$$= \liminf_{k\to\infty} \int g_k - c \, d\mu + c$$

$$\overset{\text{Fatou}}{\geq} \int \liminf_{k\to\infty} g_k - c \, d\mu + c$$

$$= \int g - c \, d\mu + c$$

$$= \int g \, d\mu \qquad \square$$

As an immediate corollary,

Cor: Given metric spaces (X, d_X) and (Y, d_Y), for any function $c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is lsc and bdd below, $\mathbb{K}_c(\gamma)$ is lsc in the narrow topology.

Combining these compactness and lsc results, by the direct method of the calculus of variations...

<u>Thm</u>: Given Polish spaces (X, d_X) and (Y, d_Y), for any function $c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is lsc and bdd below and for any $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, there exists $\gamma_* \in \Gamma(\mu, \nu)$ satisfying

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

A solution of Kantorovich's problem exists!

Why was the narrow topology the "right" topology?

Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a normed vector space
n.v.s.

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

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

Given $x \in X, y \in X^*$ let

$$\langle y, x \rangle := y(x).$$

Topologies on n.v.s.

$\boxed{\text{on } X}$

- norm, $\|x_n - x\|_X \to 0$
- weak, $\underbrace{\langle y, x - x_n \rangle}_{y(x) - y(x_n)} \to 0 \ \forall y \in X^*$

$\boxed{\text{on } X^*}$

- norm, $\|y_n - y\|_{X^*} \to 0$
- weak, $\langle z, y_n - y \rangle \to 0 \ \forall z \in (X^*)^*$
- weak-*, $\langle y_n - y, x \rangle \to 0 \ \forall x \in X$

$\mathcal{X}$ is a locally compact metric space

<u>Topologies on $\mathcal{M}_s(\mathcal{X})$</u>

NOT METRIZABLE

- strong, $\mu_n(A) \to \mu(A) \quad \forall A \in \mathcal{B}(\mathcal{X})$
- total variation norm

$$\|\mu - \nu\|_{TV} = \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|$$

$$= \sup_{\substack{f \in C_0(\mathcal{X}) \\ \|f\|_\infty \le 1}} \int f \, d\mu - \int f \, d\nu$$

...where...

$$C_0(\mathcal{X}) = \{ f : \mathcal{X} \to \mathbb{R} : \forall \varepsilon > 0, K_\varepsilon \subset\subset \mathcal{X} \text{ s.t. } |f(x)| \le \varepsilon \ \forall x \in K_\varepsilon^c \}$$

Banach space	$(C_0(X), \|\cdot\|_\infty)$	$(C_b(X), \|\cdot\|_\infty)$
Dual space	$(\mathcal{M}_s(X), \|\cdot\|_{TV})$	$\ddot\frown$
Weak-* topology	wide topology $\Leftarrow$	narrow topology

$$\mu_n \xrightarrow{\text{wide}} \mu \iff \int f \, d\mu_n \to \int f \, d\mu \quad \forall f \in C_0(X)$$

Which topology for Kantorovich's problem?

Exercise 10: Give an example of $\mu \in P(\mathbb{R})$ for which $\Gamma(\mu, \mu)$ is not compact wrt $\|\cdot\|_{TV}$

Similarly, wide topology is too weak to ensure lsc of $\mathbb{K}_c(\gamma)$.

For example, $X = Y = \mathbb{R}$, $c(x,y) = |x-y| - 1$.
Consider $\gamma_n = \delta_{(n,n)} \in P(\mathbb{R} \times \mathbb{R})$

Fact: $\gamma_n \to 0$ in wide topology

However,

$$\liminf_{n\to\infty} \mathbb{K}_c(\gamma_n)$$

$$= \liminf_{n\to\infty} \int |x-y| - 1 \, d\gamma_n(x,y)$$

$$= \liminf_{n\to\infty} |n-n| - 1$$

$$= -1$$

But $\mathbb{K}_c(0) = 0$.

Moral: Wide convergence can allow mass to "escape to $+\infty$".
OTOH, given a narrowly convergent sequence of probability measures, the limit <u>must</u> be a prob measure.

However...

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

$$\mu_n \to \mu \text{ narrowly}$$

if and only if

(*) $\left(\lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu \quad \forall f \in C_c(X)\right)$

In particular,

wide convergence + conservation of mass & positivity $\iff$ narrow conv

(will discuss C_c^∞ next time)

Pf: Suppose (*) holds.

Since $\{\mu\}$ is tight, $\forall k \in \mathbb{N}$, $\exists$ an increasing sequence of compact sets K_k s.t. $\mu(X \setminus K_k) \leq \frac{1}{k}$.

...pick up here next time...

Crash course in convex analysis and optimization