Lecture 5

Recall:

Kantorovich's Optimal Transport Problem

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

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

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

Reasons this is a better behaved problem:

① $\forall \mu,\nu \in \mathcal{P}(\mathcal{X})$, the constraint set is nonempty

② The constraint set is convex.

③ The objective function is convex.

④ Kantorovich's problem has a dual problem.

will discuss soon

⑤ We can prove a minimizer exists via the Direct Method of the Calculus of Variations.

$\downarrow$

Need a topology in which the constraint set is compact and the objective function is lsc.

Def: A sequence $\mu_n \in \mathcal{M}(X)$ is narrowly (metrizable) convergent to $\mu \in \mathcal{M}(X)$ if

$$\lim_{n\to\infty} \int_X \varphi \, d\mu_n = \int_X \varphi \, d\mu, \quad \forall \varphi \in C_b(X).$$

Def: $\mathcal{K} \subseteq \mathcal{P}(X)$ is tight if, $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset\subset X$ s.t. $\mu(X \setminus K_\varepsilon) \leq \varepsilon \quad \forall \mu \in \mathcal{K}$.

Suppose (X,d) is a Polish space. (complete, separable)

Thm (Prokhorov): Given a Polish space (X,d) and $\mathcal{K} \subseteq \mathcal{P}(X)$,

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

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

$\Updownarrow$

- $\mathcal{K}$ is tight.

<u>Cor</u>: If (X,d) is a Polish space, then for any $\mu \in \mathcal{P}(X)$, $\{\mu\}$ is tight.

We now have everything we need to prove compactness of the constraint set.

<u>Prop</u>: Given a Polish space (X,d) and $\mu, \nu \in \mathcal{P}(X)$, $\Gamma(\mu,\nu)$ is relatively compact in the narrow topology.

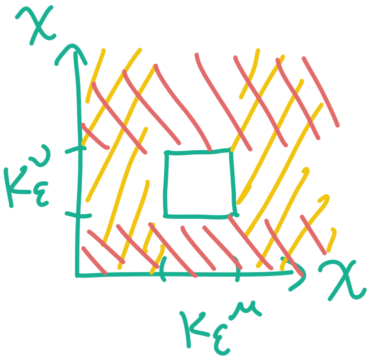
<u>Pf</u>: By corollary, $\forall \varepsilon > 0$, $\exists$ $K_\varepsilon^\mu, K_\varepsilon^\nu \subset\subset X$ s.t.

$$\mu(X \setminus K_\varepsilon^\mu) + \nu(X \setminus K_\varepsilon^\nu) \leq \varepsilon.$$

Define $K_\varepsilon := K_\varepsilon^\mu \times K_\varepsilon^\nu \subset\subset X \times X$.

Then, for all $\gamma \in \Gamma(\mu,\nu)$

$$\gamma((\mathcal{X}\times\mathcal{X})\setminus K_\varepsilon) \le \gamma(\underbrace{(\mathcal{X}\setminus K_\varepsilon^{\mu})\times\mathcal{X}}) + \gamma(\mathcal{X}\times\underbrace{(\mathcal{X}\setminus K_\varepsilon^{\nu})})$$



$$= \mu(\mathcal{X}\setminus K_\varepsilon^{\mu}) + \nu(\mathcal{X}\setminus K_\varepsilon^{\nu})$$

$$\le \varepsilon.$$

Thus, $\Gamma(\mu,\nu)$ is tight; hence by Prokhorov's theorem, it is relatively compact in the narrow topology.

A functional analysis perspective on the proof of Prokhorov's theorem

$(\mathcal{X}, d)$ locally compact metric space

Banachspace	$(C_0(\mathcal{X}), \lVert\cdot\rVert_\infty)$	$(C_b(\mathcal{X}), \lVert\cdot\rVert_\infty)$
Riesz Representation Theorem	$\downarrow$	
Dualspace	$\underline{(\mathcal{M}^s(\mathcal{X}), \lVert\cdot\rVert_{TV})}$	[big space, containing $(\mathcal{M}(\mathcal{X}), \lVert\cdot\rVert_{TV})$]
Weak-* topolgy	wide topology	narrow topology

Recall: $\mu \in \mathcal{M}^s(\mathcal{X}) \iff \mu = \mu^+ - \mu^-$ for $\mu^+, \mu^- \in \mathcal{M}(\mathcal{X})$

$\lVert\mu\rVert_{TV} = \mu^+(\mathcal{X}) + \mu^-(\mathcal{X})$

Remark: If (X,d) is compact, then all above notions coincide, i.e. $C_0(X) = C_b(X) = C(X)$.

Remark: $\mathcal{P}(X)$ is a convex subset of the unit ball in both dual spaces...

$(1-\alpha)\mu_0 + \alpha\mu_1$

$\{\mu : \|\mu\|_{TV} \leq 1\}$

but is not closed.

Exercise: Prove that $\mathcal{P}(X)$ is not closed in either the wide or narrow topologies.

Recall:

Thm (Banach, Alaoglu, Bourbaki): For any Banach space E, the closed unit ball

$$B_{E^*} = \{f \in E^* : \|f\| \leq 1\}$$

is compact in the weak-* topology.

Proof of Prokhorov:

We will show tight $\Rightarrow$ relatively compact in narrow topology.

If $\mathcal{K} \subseteq \mathcal{P}(X)$ is tight, then there exists a increasing sequence $K_m \subset\subset X$ s.t. $\mu(X \setminus K_m) \leq \frac{1}{m}, \forall \mu \in \mathcal{K}$.

Consider the restriction of μ to K_m: $\mu|_{K_m} \in \mathcal{M}(K_m)$, defined by

$$\mu|_{K_m}(B) = \mu(B \cap K_m).$$

Since $\|\mu|_{K_m}\|_{TV} = \mu(K_m) \leq 1$, we have $\{\mu|_{K_m} : \mu \in \mathcal{K}\}$ is a subset of the closed unit ball in $(C(K_m), \|\cdot\|_\infty)^*$.

Thus, B-A-B Theorem ensures that:

(i) $\exists$ a sequence $\{\mu_i^1\}_{i=1}^{+\infty}$ s.t. $\mu_i^1|_{K_1} \xrightarrow{i \to +\infty} \nu^1 \in \mathcal{M}(K_1)$

(ii) for each $n \geq 2$, choose a subsequence $\{\mu_i^n\}_{i=1}^{+\infty}$ of $\{\mu_i^{n-1}\}_{i=1}^{+\infty}$ s.t. $\mu_i^n|_{K_n} \xrightarrow{i \to +\infty} \nu^n \in \mathcal{M}(K_n)$.

By construction, for $m \leq n$, we have $\mu_i^n|_{K_m} \xrightarrow{i \to +\infty} \nu^m \in \mathcal{M}(K_m)$.

Consequently, for any $f \in C_b(X)$, $f \geq 0$,

$$\int_{K_n} f \, d\nu_n = \lim_{i\to+\infty} \int_{K_n} f \, d\mu_i^n\big|_{K_n} \geq \lim_{i\to+\infty} \int_{K_m} f \, d\mu_i^n\big|_{K_m}$$

$$= \int_{K_m} f \, d\nu^m$$

This shows:

(i) For any $f \in C_b(X)$, $f \geq 0$, thinking of ν^n and ν^m as $\mathcal{M}(X)$, we have $\int f \, d\nu^n \geq \int f \, d\nu^m$ for all $m \leq n$.

Since all bounded, monotonic sequences converge, $\int f \, d\nu^n$ converges for all $f \in C_b(X), f \geq 0$

(ii) For any closed set F, the function $f_\varepsilon(x) := \left(1 - \frac{d(x,F)}{\varepsilon}\right)^+ \in C_b(X)$ and $1_F(x) \leq f_\varepsilon(x) \leq 1_{F_\varepsilon}(x)$, for $F_\varepsilon = \{x : d(x,F) < \varepsilon\}$, so $f_\varepsilon \to 1_F$ pointwise and the Dominated Convergence Theorem ensures

$$\nu^n(F) \geq \lim_{\varepsilon\to 0} \int f_\varepsilon \, d\nu^n \geq \lim_{\varepsilon\to 0} \int f_\varepsilon \, d\nu^m = \nu^m(F).$$

(iii) Define $\nu \in \mathcal{M}(X)$ by $\nu(F) = \sup_n \nu^n(F)$, which exists by (ii).

(iv) Thus, for all $f \in C_b(X)$, $t \geq 0$

$$\nu(\overbrace{\{x : f(x) > t\}}^{\mathcal{U}\text{ open}}) = \nu(\overset{\text{closed}}{X}) - \nu(\overset{\text{closed}}{\mathcal{U}^c})$$
$$= \lim_{n\to\infty} \nu^n(X) - \nu^n(\mathcal{U}^c)$$
$$= \lim_{n\to\infty} \nu^n(\{x : f(x) > t\})$$

Hence, for all $f \in C_b(X)$, $f \geq 0$

$$\int f \, d\nu = \int_0^{+\infty} \nu(\{x : f(x) > t\}) dt$$
$$= \int_0^{\|f\|_\infty} \nu(\{x : f(x) > t\}) dt$$

Dominated convergence theorem

$$= \lim_{n\to\infty} \int_0^{\|f\|_\infty} \nu^n(\{x : f(x) > t\}) dt$$
$$= \lim_{n\to\infty} \int f \, d\nu^n$$

(ii) By taking positive and negative parts of f, we conclude that $\int f \, d\nu^n \to \int f \, d\nu$ for all $f \in C_b(X)$

Finally, via a diagonal argument, we obtain $\exists$ i_n s.t. $\mu^n_{i_n} \xrightarrow{n\to\infty} \nu$ narrowly.

Thus, $\mathcal{K}$ is relatively narrowly cpt. $\square$

Remaining key ingredient to apply Direct Method of Calculus of Variations:

need to show $\mathbb{K}_p(\gamma)$ is lsc in narrow topology

To determine what happens when $\mu_n \to \mu$ narrowly and φ is merely lsc...

$$\liminf_{n\to\infty} \int \varphi \, d\mu_n = ??$$

... we will show that any lower semicontinuous function may be approximated from below by cts fns.

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. $\lim_{k\to+\infty} g_k(x) \nearrow g(x) \ \forall x \in X$.

Def: A function $g : X \to \mathbb{R} \cup \{+\infty\}$ is proper if $\exists \, x$ s.t. $g(x) < +\infty$.

Def: Given $g: X \to \mathbb{R} \cup \{+\infty\}$, the Moreau-Yosida regularization is given by

$$g_k(x) := \inf_{y \in X} g(y) + k d(x,y), \quad k \geq 0.$$

Pf of Lemma:

Trivially true if $g \equiv +\infty$, so we may assume g is proper. Let $g_k(x)$ be the Moreau-Yosida regularization of $g(x)$.

(1) $g_k(x)$ is continuous

Suppose $x_n \xrightarrow{d} x$.

for fixed $y \in X$

$$\limsup_{n \to +\infty} g_k(x_n) \leq \limsup_{n \to \infty} g(y) + k d(x_n, y)$$

$= g(\text{c}$