Lecture 11

Recall:

Thm: Suppose $(X, d_X), (Y, d_Y)$ are cpt Polish spaces and $c \in C(X \times Y)$, $c \geq 0$. Then $\forall$ $\mu \in \mathbb{P}(X)$, $\nu \in \mathbb{P}(Y)$,

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

Furthermore, the maximum is attained.

Recall:

$$F((\varphi,\psi),u) = -\int \varphi \, d\mu - \int \psi \, d\nu + \chi_C(u)$$

$$C := \{u \in C_b(X \times Y) : u(x,y) + \varphi(x) + \psi(y) \leq c(x,y)\}$$

"Double convexification trick"

<u>Prop</u>: Suppose (X, d_X), (Y, d_Y) are cpt and $c: X \times Y \to [0, +\infty)$ cts.
Given

$\{(\varphi_i, \psi_i)\}_{i \in I} \subseteq C(X) \times C(Y)$, ~~$\varphi_i \geq 0$~~

$\{u_i\}_{i \in I} \subseteq C(X \times Y)$ unif bdd, e-cts

$\sup_{i \in \mathbb{N}} F((\varphi_i, \psi_i), u_i) < +\infty$

Then, $\{\tilde{\varphi}_i\}_{i \in I}$, $\{\tilde{\psi}\}_{i \in I}$ are unif bdd and e-cts and

$$F((\tilde{\varphi}_i, \tilde{\psi}_i), u_i) \leq F((\varphi_i, \psi_i), u_i).$$

<u>Pf</u>: Choose $C_0 \in \mathbb{R}$ s.t., $\forall i \in I$,

$$C_0 \geq F((\varphi_i, \psi_i), u_i) = -\int \varphi_i \, d\mu - \int \psi_i \, d\nu$$

$$\geq -\max_{x \in X} \varphi_i(x) - \max_{y \in Y} \psi_i(y)$$

$\exists x_i \in X$, $y_i \in Y$ s.t.

$$= -\varphi_i(x_i) - \psi_i(y_i)$$

That is $\varphi_i(x_i) - s_i + \psi_i(y_i) + s_i \geq -c_0$, for any $s_i \in \mathbb{R}$.

When $s_i \equiv 0$, we see $\varphi_i(x_i) + \psi_i(y_i) \geq -c_0$, so for each i either $\varphi_i(x_i) \geq -c_0/2$ (*)

or $\psi_i(y_i) \geq -c_0/2$. (**)

If (*) holds, let $s_i = \varphi_i(x_i) + c_0/2$.

Then, $\varphi_i(x_i) - s_i = -c_0/2$

$$\psi_i(y_i) + s_i = \psi_i(y_i) + \varphi_i(x_i) + c_0/2 \geq -\frac{c_0}{2}$$

Likewise, if (**), let $s_i = -\psi_i(y_i) + c_0/2$.

This shows we may always choose

$$s_i \text{ so that } \underbrace{\varphi_i(x_i) - s_i}_{\bar{\varphi}_i(x_i)} \geq -c_0/2$$

$$\underbrace{\psi_i(y_i) + s_i}_{\bar{\psi}_i(y_i)} \geq -c_0/2.$$

$$\text{Let } \left.\begin{aligned} \bar{\varphi}_i(x) &:= \varphi_i(x) - s_i \\ \bar{\psi}_i(y) &:= \psi_i(x) + s_i \end{aligned}\right\} \Rightarrow \begin{aligned} &F((\bar{\varphi}_i, \bar{\psi}_i), u) \\ &= F((\varphi_i, \psi_i), u) \end{aligned}$$

Define $\tilde{\varphi}_i(x) = \inf_{y \in Y} c(x,y) - u_i(x,y) - \overline{\psi}_i(y)$

Then, $\forall\, x, y$, we have

$$c(x,y) - u_i(x,y) - \overline{\psi}_i(y) \geq \overline{\varphi}_i(x) \quad (*)$$

so...

- $\tilde{\varphi}_i(x) \geq \overline{\varphi}_i(x) \quad \forall\, x \in X$
- $F((\tilde{\varphi}_i, \overline{\psi}_i), u_i) \leq F((\overline{\varphi}_i, \overline{\psi}_i), u_i)$
- $\{\tilde{\varphi}_i\}_{i \in I}$ are equicts (by lemma)
- to see $\inf_{i \in I, x \in X} \tilde{\varphi}_i(x) > -\infty$,

By $(*)$, $\forall\, y \in Y$

$$\overline{\psi}_i(y) \leq \underbrace{c(x_i, y) - u_i(x_i, y) - \overline{\varphi}_i(x_i)}_{\leq c_0/2}$$

so $\sup_{i \in I, y \in Y} \overline{\psi}_i(y) < +\infty$

The result follows by defn of $\tilde{\varphi}$.

- to see $\sup_{i\in I, x\in X} \tilde{\varphi}_i(x) < +\infty$,

by $\cancel{(*)}$ with $y = y_i$, $c_0/2$

$$c(x,y_i) - u_i(x,y_i) - \underbrace{\overline{\psi}_i(y_i)}_{\leq} \geq \overline{\varphi}_i(x)$$

Define $\tilde{\psi}_i(y) = \inf_{x\in X} c(x,y) - u_i(x,y) - \tilde{\varphi}_i(x)$

We always have, $\forall\, x, y$,

$$c(x,y) - u_i(x,y) - \tilde{\varphi}_i(x) \geq \overline{\psi}_i(y)$$

So...

- $\tilde{\psi}_i(y) \geq \overline{\psi}_i(y)$
- $F((\tilde{\varphi}_i, \tilde{\psi}_i), u_i) \leq F((\tilde{\varphi}_i, \overline{\psi}_i), u_i)$
- Again, $\{\tilde{\psi}_i\}_{i\in I}$ e-cts
- Since $\{\tilde{\varphi}_i\}_{i\in I}$ are unif bdd, so are $\{\tilde{\psi}_i\}_{i\in I}$. $\square$

Recall: Given a compact metric space X, $\mathcal{F} \subseteq C(X)$, Arzelá-Ascoli ensures

$\mathcal{F}$ unif bdd, equicts $\Leftrightarrow$ $\overline{\mathcal{F}}$ compact.

Now, use this fact to prove $P_0 = D_0$ for Kantorovich problem

Pf:

$$\mathcal{P}(u) = \inf_{(\varphi,\psi) \in C(X) \times C(Y)} F((\varphi,\psi), u)$$

Step 3: $\mathcal{P}$ is lsc at 0

$\boxed{\text{Step 3}}$ Suppose $u_n \to 0$.
uniformly on $X \times Y$

We must show

$$\liminf_{n \to \infty} \mathcal{P}(u_n) \geq \mathcal{P}(0).$$

WLOG, suppose $+\infty > \liminf_{n\to\infty} \mathcal{P}(u_n)$.
Let u_{n_k} denote a subsequence s.t.

$$\lim_{k\to\infty} \mathcal{P}(u_{n_k}) = \liminf_{n\to\infty} \mathcal{P}(u_n)$$

Thus $\sup_{k\in\mathbb{N}} \mathcal{P}(u_{n_k}) < +\infty$.

Likewise,

$$\begin{aligned} F((\varphi,\psi),u_{n_k}) &\geq -\int \varphi(x)+\psi(y)\, d\mu\otimes\nu(x,y) \\ &\geq \int u_{n_k}(x,y) - c(x,y)\, d\mu\otimes\nu \\ &\geq \int u_{n_k}\, d\mu\otimes\nu - \max_{x\in X,\, y\in Y} c(x,y) \end{aligned}$$

Since $u_{n_k} \to 0$ unif, RHS unif bdd below

Thus $\inf_{k\in\mathbb{N}} \mathcal{P}(u_{n_k}) > -\infty$.

so by defn of infimum, $\forall k \in \mathbb{N}$, $\exists (\varphi_{n_k}, \psi_{n_k}) \in C(\mathcal{X}) \times C(\mathcal{Y})$ s.t.

$$P(u_{n_k}) \geq F((\varphi_{n_k}, \psi_{n_k}), u_{n_k}) - \frac{1}{k}.$$

bdd above, unif in $k \in \mathbb{N}$

Since $u_{n_k} \to 0$, $\{u_{n_k}\}_{k \in \mathbb{N}}$ unif bdd, e-cts.

By Double Convexification Lemma, $\exists$ $\{\tilde{\varphi}_{n_k}\}$, $\{\tilde{\psi}_{n_k}\}$ are unif bdd and e-cts and

$$F((\tilde{\varphi}_{n_k}, \tilde{\psi}_{n_k}), u_{n_k}) \leq F((\varphi_{n_k}, \psi_{n_k}), u_{n_k}).$$

Thus,

$$+\infty > P(u_{n_k}) \geq F((\tilde{\varphi}_{n_k}, \tilde{\psi}_{n_k}), u_{n_k}) - \frac{1}{k}.$$

By Arzelá-Ascoli, $\exists$ further subsequences $\tilde{\varphi}_{n_k}$, $\tilde{\psi}_{n_k}$ s.t.

$\tilde{\varphi}_{n_k} \to \varphi_* \in C(X)$, $\tilde{\psi}_{n_k} \to \psi_* \in C(Y)$.

Since by definition,

$$\begin{aligned} u_{n_k}(x,y) &+ \tilde{\varphi}_{n_k}(x) + \tilde{\psi}_{n_k}(y) \le c(x,y) \\ \downarrow k\to\infty \\ 0 &+ \varphi_*(x) + \psi_*(y) \le c(x,y) \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{k\to\infty} P(u_{n_k}) &\ge \liminf_{k\to\infty} F((\tilde{\varphi}_{n_k}, \tilde{\psi}_{n_k}), u_{n_k}) \\ &= \liminf_{n\to\infty} -\int \tilde{\varphi}_{n_k}\, d\mu - \int \tilde{\psi}_{n_k}\, d\nu \\ &\overset{DCT}{=} -\int \varphi_*\, d\mu - \psi_*\, d\nu \\ &= F((\varphi_*, \psi_*), 0) \\ &\ge \inf_{(\varphi,\psi)} F((\varphi,\psi), 0) \\ &= P(0). \end{aligned}$$

Thus, Φ is lsc at zero, so $P_0 = D_0$.
(Rest of proof same as before)

Alert: common terminology abuse

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

"Primal" "Dual"

From Kantorovich back to Monge

Thm (Knott-Smith Optimality)
Fix $X \subseteq \mathbb{R}^d$ cpt, $c(x,y) = |x-y|^2$,
$\mu, \nu \in P(X)$.

(i) $\exists\, f_* \in L^1(\mu)$ proper, lsc, convex s.t.

$\bigstar$

$\boxed{\text{i.a}}$ $\sup_{\substack{\varphi,\psi \in C_b(\mathcal{X}) \\ \varphi(x)+\psi(y) \le c(x,y)}} \int \varphi \, d\mu + \int \psi \, d\nu = \int |x|^2 - 2f_*(x)\, d\mu + \int |y|^2 - 2f_*^*(y)\, d\nu$

$\boxed{\text{i.b}}$ For <u>any</u> OT plan γ_*, $y \in \partial f_*(x)$ γ_*-a.e.

(ii) Conversely, if $\gamma \in \Gamma(\mu,\nu)$ and $f \in L^1(\mu)$ proper, lsc, convex for which $y \in \partial f(x)$ γ-a.e., then...

$\boxed{\text{ii.a}}$ γ is optimal

$\boxed{\text{ii.b}}$ f satisfies $\bigstar$

Rmk: Suprisingly, $\boxed{\text{i.b}}$ does not ensure OT plans are unique

Exercise 22: Consider $(\mathcal{X}, d) = (\mathbb{R}^2, |\cdot|)$

$\mu = \frac{1}{2}(\delta_{(-1,-1)} + \delta_{(1,1)})$, $\nu = \frac{1}{2}(\delta_{(-1,1)} + \delta_{(1,-1)})$

Show that $f_*(\underline{x}) = f_*(x_1, x_2) = |x_1 - x_2|$ satisfies (i) above and find two distinct OT plans satisfying (ii).

Exercise 23: Double convexification for quadratic cost on $\mathbb{R}^d$.

Given $(\varphi,\psi) \in C(\mathbb{R}^d) \times C(\mathbb{R}^d)$ s.t. $F((\varphi,\psi),0) < +\infty$, define

$$\tilde{\varphi}(y) = \inf_{x \in \mathbb{R}^d} |x-y|^2 - \psi(y),$$

$$\tilde{\psi}(x) = \inf_{y \in \mathbb{R}^d} |x-y|^2 - \tilde{\varphi}(x).$$

Then,

(i) $f(x) := \frac{1}{2}(|x|^2 - \tilde{\varphi}(x)) \in L^1(\mu)$ and is proper, lsc, convex

(ii) $f^*(y) = \frac{1}{2}(|y|^2 - \tilde{\psi}(y))$

(iii) $F((\tilde{\varphi},\tilde{\psi}),0) \leq F((\varphi,\psi),0)$.

Pf:

First, we prove (i).

By Kantorovich duality, $\exists$ $\varphi_0, \psi_0 \in C(X)$ s.t. $\varphi_0(x) + \psi_0(y) \leq |x-y|^2$

$$F((\varphi_0, \psi_0), 0) = -\int \varphi_0 d\mu - \int \psi_0 d\nu = P_0$$

By Double convexification exercise, $\exists$ $f \in L^1(\mu)$ proper, lsc, convex where

$$F((\tilde{\varphi}, \tilde{\psi}), 0) \leq F((\varphi_0, \psi_0), 0).$$

for $f(x) := \frac{1}{2}(|x|^2 - \tilde{\varphi}(x))$

$$f^*(y) = \frac{1}{2}(|y|^2 - \tilde{\psi}(y))$$

Thus, if γ_0 is an OT plan,

$$
\begin{aligned}
P_0 &= F((\varphi_0,\psi_0),0)\\
&\geq F((\tilde{\varphi},\tilde{\psi}),0)\\
&= F((|x|^2-2f(x),\ |y|^2-2f^*(y)),0)\\
&= -\int |x|^2-2f(x)\,d\mu - \int |y|^2-2f^*(y)\,d\nu\\
&= -\int |x|^2-2f(x)+|y|^2-2f^*(y)\,d\gamma_0(x,y)\\
&= \int -|x|^2+2f(x)-|y|^2+2f^*(y)\,d\gamma_0(x,y)\\
&\overset{\text{Young}}{\geq} \int -|x|^2-|y|^2+2\langle x,y\rangle\,d\gamma_0(x,y)\\
&= -\int |x-y|^2\,d\gamma_0(x,y)\\
&= D_0\\
&= P_0
\end{aligned}
$$

Young: $\langle x,y\rangle \leq f(x)+f^*(y)$

Thus, equality holds throughout.

 implies $\boxed{\text{i.a}}$.

Subtracting  from , we see

$$\int 2f(x)+2f^*(y)-2\langle x,y\rangle \, d\gamma_0(x,y)=0$$

Since Young guarantees integrand always ≥ 0. Hence,

$$f(x)+f^*(y)-\langle x,y\rangle = 0 \quad \gamma_0\text{-a.e.}$$

$$\Downarrow$$

$$y\in\partial f(x) \quad \gamma_0\text{-a.e.}$$

This completes the proof of (i).