Lecture 10

Recall:

<u>Proposition</u>: Suppose $\mathcal{X}$ is a Polish space and $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is lsc and bdd below. Then $\forall \mu \in \mathcal{M}^s(\mathcal{X})$

$$\sup\left\{ \int g \, d\mu : g \in C_b(\mathcal{X}), g \leq f \right\} = \begin{cases} \int f \, d\mu & \text{if } \mu \in \mathcal{M}(\mathcal{X}) \\ +\infty & \text{if } \mu \in \mathcal{M}^s(\mathcal{X}) \setminus \mathcal{M}(\mathcal{X}) \end{cases}$$

Pf: Exercise 20

formerly $\mathcal{X}$ formerly x

How should we choose $Z, \mathcal{U}, F(z,u)$ so that KP coincides with...

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

$$D_0 = \sup_{v \in \mathcal{U}^*} \inf_{(z,u) \in Z \times \mathcal{U}} F(z,u) - \langle v, u \rangle \quad ?$$

$$F^*(y,v) = \sup_{(z,u) \in Z \times \mathcal{U}} \langle y, z \rangle + \langle v, u \rangle - F(z,u)$$

In the case of KP...

$$-KP = -\inf_{\gamma \in \mathcal{M}^s(X \times Y)} \mathbb{K}_c(\gamma) + \chi_{\Gamma(\mu,\nu)}(\gamma) + \chi_{\mathcal{M}(X \times Y)}(\gamma)$$

$$= -\inf_{\gamma \in \mathcal{M}^s(X \times Y)} \sup_{(\varphi,\psi) \in C_b(X) \times C_b(Y)} \mathbb{K}_c(\gamma) + \langle \mu - \pi^1{}_\# \gamma, \varphi \rangle + \langle \nu - \pi^2{}_\# \gamma, \psi \rangle + \chi_{\mathcal{M}(X \times Y)}(\gamma)$$

$$= \sup_{\gamma \in \mathcal{M}^s(X \times Y)} \inf_{(\varphi,\psi) \in C_b(X) \times C_b(Y)} -\mathbb{K}_c(\gamma) + \langle \pi^1{}_\# \gamma - \mu, \varphi \rangle + \langle \pi^2{}_\# \gamma - \nu, \psi \rangle - \chi_{\mathcal{M}(X \times Y)}(\gamma)$$

Gathering the γ's...

$$-\mathbb{K}_c(\gamma) + \langle \pi^1_{\#}\gamma - \mu, \varphi\rangle + \langle \pi^2_{\#}\gamma - \nu, \psi\rangle - \chi_{\mathcal{M}(X\times Y)}(\gamma)$$

$$= -\int \varphi \, d\mu - \int \psi \, d\nu - \underbrace{\int c(x,y) - \varphi(\pi^1(x,y)) - \psi(\pi^2(x,y))}_{\text{lsc and bdd below}} \, d\gamma$$
$$- \chi_{\mathcal{M}(X\times Y)}(\gamma)$$

$$= -\int \varphi \, d\mu - \int \psi \, d\nu - \sup_{\substack{u \in C_b(X\times Y), \\ u(x,y) \le c(x,y) - \varphi(x) - \psi(y)}} \int u(x,y) \, d\gamma$$

$$= \inf_{u \in C_b(X\times Y)} \underbrace{-\int \varphi \, d\mu - \int \psi \, d\nu + \chi_C(u)}_{F((\varphi,\psi),u)} - \int u(x,y) \, d\gamma$$

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

$\boxed{\text{THUS}}$

$-KP$

$$= \sup_{\gamma \in \mathcal{M}^s(X \times Y)} \inf_{(\varphi,\psi) \in C_b(X) \times C_b(Y)} \inf_{u \in C_b(X \times Y)} F((\varphi,\psi),u) - \langle u, \gamma \rangle,$$

For simplicity, assume (X, d_X) and (Y, d_Y) are compact metric spaces.

Fact: $(C(X))^* = \mathcal{M}^s(X)$

What is the corresponding primal problem?

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

$\downarrow$

primal problem: $\mathcal{P}_0 := \inf_{x \in X} f(x)$, $f(x) = F(x,0)$

By defn of F,

$$\mathcal{P}_0 = -\sup_{\substack{(\varphi,\psi)\in C_b(X)\times C_b(Y) \\ \varphi(x)+\psi(y)\le c(x,y)\ \forall x\in X,\, y\in Y}} \int \varphi\, d\mu + \int \psi\, d\nu.$$

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

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

$$\underbrace{\inf_{\gamma\in\Gamma(\mu,\nu)} \mathbb{K}(\gamma)}_{=-D_0} = \underbrace{\sup_{\substack{(\varphi,\psi)\in C(X)\times C(Y) \\ \varphi\oplus\psi\le c}} \int \varphi\, d\mu + \int \psi\, d\nu}_{=-\mathcal{P}_0}$$

Furthermore, the maximum is attained.

<u>Lemma</u>: Let (X, d_X) be a metric space and $\mathcal{A}$ a set. Suppose $\mathcal{F}$ is a collection of functions $f: X \times \mathcal{A} \to \mathbb{R}$. Suppose $\{f(\cdot, \alpha) : f \in \mathcal{F}, \alpha \in \mathcal{A}\}$ is e-cts. Then $\{\inf_{\alpha \in \mathcal{A}} f(\cdot, \alpha) : f \in \mathcal{F}\}$ is e-cts.

Pf: Exercise 19

"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

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

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

$$F((\varphi_i, \psi_i), u_i) < +\infty \quad \forall i \in I.$$

define

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

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

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

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

Pf:

Since $F((\varphi_i, \psi_i), u_i) < +\infty$,

$$u_i(x,y) + \varphi_i(x) + \psi_i(y) \leq c(x,y) \quad \forall x, y$$

$$\underset{\vee\!\mid}{} \quad u_i(x,y) + \psi_i(y)$$

Thus, $\sup_{i \in I, y \in Y} \psi_i(y) < +\infty$.

By definition of $\tilde{\varphi}_i$,

- $\varphi_i(x) \leq \tilde{\varphi}_i(x)$, $u_i(x,y) + \tilde{\varphi}_i(x) + \psi_i(y) \leq c(x,y)$
- By lemma $\{\tilde{\varphi}_i\}_{i \in I}$ c-cts
-

... will finish next time...

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

$$F((\varphi,\psi),u) = -\int \varphi \, d\mu - \int \psi \, d\nu + \chi_{\{u(x,y)+\varphi(x)+\psi(y) \leq c(x,y)\}}((\varphi,\psi),u)$$

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

By theorem on equivalence

of primal and dual problems, it suffices to show...

Step 1: F is convex
Step 2: $\mathcal{P}(0) < +\infty$
Step 3: $\mathcal{P}$ is lsc at 0

... to conclude $\mathcal{P}_0 = D_0$

Step 4: Prove the primal problem has a sol'n.

$\boxed{\text{Step 1}}$ This can be seen from the defn of F.

$\boxed{\text{Step 2}}$ $\mathcal{P}(0) \leq F((0,0),0) = 0 < +\infty.$

$\boxed{\text{Step 3}}$ Suppose $u_n \to 0$.

uniformly on $\mathcal{X} \times \mathcal{Y}$

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

This ensures $\mathcal{P}(u_{n_k}) < +\infty \quad \forall k \in \mathbb{N}$,
Suppose $\mathcal{P}(u_{n_k}) > -\infty$.

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.

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

Note that, $\forall\, C_n \in \mathbb{R}$, defining
$$\tilde{\varphi}_n := \varphi_n + C_n, \quad \tilde{\psi}_n := \psi_n - C_n,$$
we have
$$F((\varphi_n, \psi_n), u_n) = F((\tilde{\varphi}_n, \tilde{\psi}_n), u_n)$$

Thus, we may assume $\varphi_{n_k} \geq 0$.

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

By Arzelà-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,

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

Thus, $\mathcal{P}$ is lsc at zero, so $\mathcal{P}_0 = D_0$.

Step 4: It remains to show $\exists$ optimizer for primal problem.

Consider $u_n \equiv 0$ in previous argument. Taking φ_*, ψ_* as above,

$$\begin{aligned}
\mathcal{P}(0) &= \lim_{k\to\infty} \mathcal{P}(u_{n_k}) \\
&= F((\varphi_*, \psi_*), 0) \qquad = f((\varphi_*, \psi_*)) \\
&\geq \inf_{(\varphi,\psi)} F((\varphi,\psi), 0) = \inf_{(\varphi,\psi)} f((\varphi,\psi)) \\
&= \mathcal{P}(0).
\end{aligned}$$

Thus, equality holds throughout, and (φ_*, ψ_*) is optimizer for primal.

Exercise 21: Extend the previous theorem equating $P_0 = D_0$ to the case where c is lsc and bdd below.

What about X, Y noncompact?
Key difficulty: no more Arzela-Ascoli.

- For X, Y Polish spaces, one can still prove $P_0 = D_0$.
- In general, need to enlarge the space $C_b(X) \times C_b(X)$ to get existence of optimizers (φ^*, ψ^*).

Alert: common terminology abuse

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

"Primal" "Dual"

From Kantorovich back to Monge

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

Remark: $\tilde{\varphi}_i, \tilde{\psi}_i$ are the Moreau-Yosida regularizations of $-\psi_i, -\varphi_i$ with respect to the <u>square</u> distance