Lecture 25, Monday, 5/23/2022

Review: The Kantorovich duality — the strong duality

<u>Theorem</u> (Kantorovich duality) Let $X, Y$ be Polish, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, $c: X\times Y \to [0,\infty]$ lower semi-continuous.

Then $\inf_{\mathcal{A}(\mu,\nu)} E_K = \sup_{\Phi_c(\mu,\nu)\cap C_b} J = \sup_{\Phi_c(\mu,\nu)} J$.

Recall $\mathcal{A}(\mu,\nu) = \{\gamma \in \mathcal{P}(X\times Y): \pi^X_{\#}\gamma = \mu,\ \pi^Y_{\#}\gamma = \nu\}$.

$\Phi_c(\mu,\nu) = \{(\varphi,\psi)\in L^1(\mu)\times L^1(\nu): \varphi(x)+\psi(y) \le c(x,y),$
$\mu\text{-a.e. } x\in X,\ \nu\text{-a.e. } y \in Y\}$.

$$E_K[\gamma] = \int_{X\times Y} c(x,y)\,d\gamma(x,y).$$

$$J[\varphi,\psi] = \int_X \varphi\,d\mu + \int_Y \psi\,d\nu.$$

<u>Weak duality</u>. $\forall \gamma \in \mathcal{A}(\mu,\nu)$ $\forall (\varphi,\psi)\in \Phi_c(\mu,\nu)$:

$J[\varphi,\psi] \le E_K[\gamma]$, and $\sup_{\Phi_c(\mu,\nu)\cap C_b} J \le \sup_{\Phi_c(\mu,\nu)} J \le \inf_{\mathcal{A}(\mu,\nu)} E_K$.

<u>Today</u>: <u>The Kantorovich-Rubinstein Theorem — $W_1$ theory</u>.

Set up. Let $X$ be a Polish space. Denote

$\mathcal{P}_1(X) = \{\mu \in \mathcal{P}(X): \int_X d(x,x_0)\,d\mu(x) < \infty$ for some $x_0 \in X\}$.

Note: $\exists x_0\in X$ s.t. $\int_X d(x,x_0)\,d\mu(x)<\infty \iff \forall x_0 \in X$: $\int_X d(x,x_0)\,d\mu(x) < \infty$. Denote for any $\mu,\nu\in\mathcal{P}_1(X)$,

$$W_1(\mu,\nu) = \inf_{\gamma\in\mathcal{A}(\mu,\nu)} \int_{X\times X} d(x,y)\,d\gamma(x,y).$$

Recall: $\mathcal{M}(X)$ is the normed vector space of signed finite measures on $X$ with the norm $\|\nu\| = |\nu|(X)$. $\mathcal{M}_+(X)$ is the subset of $\mathcal{M}(X)$ of all nonnegative finite measures.

Let $\check{\mathcal{M}}_1(X) = \text{span}\, \mathcal{P}_1(X)$ in $\check{\mathcal{M}}(X)$. Define for any $\sigma \in \check{\mathcal{M}}_1(X)$

$$\|\sigma\|_{KR} = \sup \left\{ \int_X \varphi \, d\sigma : \varphi \in Lip(X) \cap L^1(\sigma),\ Lip(\varphi) \le 1 \right\},$$

where $Lip(\varphi) = \sup_{x, y \in X, x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)}$. Then, $(\check{\mathcal{M}}_1(X), \|\cdot\|_{KR})$ is a normed vector space.

<u>Theorem (Kantorovich-Rubinstein)</u> Let $(X, d)$ be a Polish space. Then $W_1(\mu, \nu) = \|\mu - \nu\|_{KR}$ $\forall \mu, \nu \in \mathcal{P}_1(X)$.

<u>Lemma</u> Let $X$ and $Y$ be Polish spaces, $\mu \in \mathcal{P}(X)$, and $\nu \in \mathcal{P}(Y)$. Let $c : X \times Y \to [0, \infty]$ be lower semicontinuous. Suppose $c_n : X \times Y \to [0, \infty]$ $(n = 1, 2, \dots)$ are nondecreasing, bounded and continuous, and $c_n \nearrow c$ on $X \times Y$. Then

$$\inf_{\gamma \in \mathcal{A}(\mu,\nu)} \int_{X \times Y} c_n \, d\gamma \to \inf_{\gamma \in \mathcal{A}(\mu,\nu)} \int_{X \times Y} c \, d\gamma \quad \text{as } n \to \infty.$$

<u>Proof</u> Denote $E_n[\gamma] = \int_{X \times Y} c_n \, d\gamma$ and $E[\gamma] = \int_{X \times Y} c \, d\gamma$.

Let $\gamma_n \in \mathcal{A}(\mu,\nu)$ be such that $E_n[\gamma_n] = \inf_{\mathcal{A}(\mu,\nu)} E_n$. Since $c_n \uparrow c$, $E_n[\gamma] \le E[\gamma]$ $\forall \gamma \in \mathcal{A}(\mu,\nu)$. Hence, $E_n[\gamma_n]$ is non-decreasing and is bounded by $\inf_{\mathcal{A}(\mu,\nu)} E$. In particular, $\lim_{n \to \infty} E_n[\gamma_n] \le \inf_{\mathcal{A}(\mu,\nu)} E$. Since $\mathcal{A}(\mu,\nu)$ is narrowly compact, up to a subseq., not relabeled, $\gamma_n \to \hat{\gamma} \in \mathcal{A}(\mu,\nu)$ narrowly. Since $c_n \uparrow c$, each $c_n \in C_b(X \times Y)$, we have for $n \ge m$, $E_n[\gamma_n] \ge E_m[\gamma_n]$, and thus

$$\lim_{n \to \infty} E_n[\gamma_n] \ge \lim_{n \to \infty} E_m[\gamma_n] = E_m[\hat{\gamma}].$$

By the monotone convergence theorem, $E_m[\hat{\gamma}] \to E[\hat{\gamma}]$. Hence, $\lim_{n \to \infty} E_n[\gamma_n] \ge E[\hat{\gamma}] \ge \inf_{\mathcal{A}(\mu,\nu)} E \ge \lim_{n \to \infty} E_n[\gamma_n]$. <u>QED</u>

**Definition** (c-concave functions) Let $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ be measurable spaces, $c : X \times Y \to [0, \infty]$ measurable, and $\varphi : X \to [-\infty, \infty)$, $\psi : Y \to [-\infty, \infty)$ measurable. Define

$$\varphi^c(y) = \inf_{x \in X} [c(x,y) - \varphi(x)] \quad \forall y \in Y.$$

$$\psi^c(x) = \inf_{y \in Y} [c(x,y) - \psi(y)] \quad \forall x \in X.$$

Call $\varphi^c, \psi^c$ c-conjugate functions corresponding to $\varphi$ and $\psi$, respectively. Call $(\varphi^{cc}, \varphi^c)$ a c-conjugate pair (if $\varphi^c \not\equiv +\infty$).

**Remark** We avoid $\infty - \infty$ in the definition.

**Definition** (c-Concavity) A function $\varphi : X \to [-\infty, \infty)$ is c-concave if it is the infimum of a family of c-affine functions $c(\cdot, y) + \alpha$. A function $\psi : Y \to [-\infty, \infty)$ is c-concave if it is the infimum of a family of c-affine functions $c(x, \cdot) + \beta$.

**Proposition** Let $X$ and $Y$ be Polish spaces, $c : X \times Y \to [0, \infty]$ lower semicontinuous, $\varphi : X \to [-\infty, \infty)$ and $\psi : Y \to [-\infty, \infty)$ Borel measurable, and $c \not\equiv +\infty$, $\varphi \not\equiv -\infty$, and $\psi \not\equiv -\infty$.

(1) Both $\varphi^c$ and $\psi^c$ are measurable.

(2) $\varphi^{cc} \geq \varphi$ $\varphi^{cc} = \varphi \Leftrightarrow \varphi$ is $c$-concave.

A similar result holds true for $\psi$.

<u>Proof</u> (1) Since $c \geq 0$ and $c$ is lower semi-continuous, we can approximate $c$ by $c_k$:

$$c_k(x,y) = \inf_{(x',y') \in X \times Y} \left[ c(x',y') \wedge k + k\, d_X(x',x) + k\, d_Y(y',y) \right].$$

Each $c_k \in BL(X \times Y)$, $0 \leq c_1 \leq \cdots \leq c_k \leq c \wedge k \leq c$, and $c = \sup_k c_k$. In particular, $c_k \uparrow c$. Define

$$\psi_k(y) = \inf_{x \in X} \left[ c_k(x,y) - \varphi(x) \right], \quad k = 1, 2, \cdots$$

Then $\psi_k \in \mathrm{Lip}(Y)$. In fact, $\forall y, y' \in Y$. $\forall \varepsilon > 0$. $\exists x$ s.t. $\psi_k(y) \leq c_k(x,y) - \varphi(x) + \varepsilon$. But $\psi_k(y') \geq c_k(x,y') - \varphi(x)$. So,

$$\psi_k(y) - \psi_k(y') \leq c_k(x,y) - c_k(x,y') + \varepsilon \leq \mathrm{Lip}(c)\, d_Y(y,y') + \varepsilon.$$

Let $\varepsilon \to 0$. Switch $y$ and $y'$. We get $\psi \in \mathrm{Lip}(Y)$.

We have $\psi_k \to \varphi^c$ on $Y$. Let $y \in Y$. $\forall \varepsilon > 0$, $\exists x \in X$ s.t. $\varphi^c(y) \geq c(x,y) - \varphi(x) - \varepsilon$. But, $\psi_k(y) \leq c_k(x,y) - \varphi(x)$. Thus, $\psi_k(y) - \varphi^c(y) \leq c_k(x,y) - c(x,y) + \varepsilon$.

Hence, $\limsup_{k \to \infty} \left( \psi_k(y) - \varphi^c(y) \right) \leq \varepsilon$, since $c_k \to c$.

A similar argument shows that

$$\liminf_{k \to \infty} \left( \psi_k(y) - \varphi^c(y) \right) \geq -\varepsilon.$$

But $\varepsilon > 0$ is arbitrary. So, $\psi_k(y) \to \varphi^c(y)$.

Since each $\psi_k$ is continuous, $\varphi^c$ is measurable.

Similarly, $\psi^c$ is measurable.

(2) Note that $\varphi^c : Y \to [-\infty, \infty)$ and $\psi^c : X \to [-\infty, \infty)$. So, $\varphi^{cc}$, $\psi^{cc}$ are well-defined. Let $x \in X$ and $\varepsilon > 0$. Then $\exists y \in Y$ s.t. $\varphi^{cc}(x) \geq c(x,y) - \varphi^c(y) - \varepsilon$. But $\varphi^c(y) \leq c(x,y) - \varphi(x)$. Hence, $\varphi^{cc}(x) \geq \varphi(x) - \varepsilon$. Hence $\varphi^{cc} \geq \varphi$.

Clearly, if $\varphi \leq \tilde{\varphi}$ then $\varphi^c \geq \tilde{\varphi}^c$. If $\varphi$ is c-concave, then $\varphi = \tilde{\psi}^c$ for some $\tilde{\psi}$. Hence, $\varphi^c = \tilde{\psi}^{cc} \geq \tilde{\psi}$, and $\varphi^{cc} \leq \tilde{\psi}^c = \varphi$. Hence, $\varphi^{cc} = \varphi$. On the other hand, if $\varphi^{cc} = \varphi$, $\varphi$ is the infimum of $c(\cdot, y) - \varphi^c(y)$ over $y$. Hence, $\varphi$ is c-concave. QED

Lemma Let $X, Y$ be Polish spaces, $\mu \in P(X)$, $\nu \in P(Y)$, and $c : X \times Y \to [0, \infty]$ bounded and continuous. Then

$$\sup_{(\varphi,\psi) \in \Phi_c(\mu,\nu)} J[\varphi, \psi] = \sup_{\varphi \in L^1(\mu)} J[\varphi^{cc}, \varphi^c] = \sup_{\psi \in L^1(\nu)} J[\psi^c, \psi^{cc}], \quad (*)$$

where $J[\varphi, \psi] = \int_X \varphi \, d\mu + \int_Y \psi \, d\nu$.

Remark ⊙ The sup. over a larger space $\Phi_c(\mu,\nu)$ is the same as that over a smaller space.

Proof of Lemma Clearly l.h.s $\geq$ r.h.s. in $(*)$, as $(\varphi^{cc}, \varphi^c) \in \Phi_c(\mu,\nu)$ if $\varphi \in L^1(\mu)$. $\forall \varphi, \psi \in \Phi_c(\mu,\nu) \cap C_b$:

$$\varphi(x) + \psi(y) \leq c(x,y) \quad \forall (x,y) \in X \times Y.$$

$\Rightarrow c(x,y) - \psi(y) \geq \varphi(x) \quad \forall (x,y)$. So, $\psi^c \geq \varphi$ on $X$,

$$J[\varphi, \psi] = \int_X \varphi \, d\mu + \int_Y \psi \, d\nu \leq \int_X \psi^c \, d\mu + \int_Y \psi \, d\nu = J[\psi^c, \psi].$$

Similarly, $J[\psi^c, \psi] \leq J[\psi^c, \psi^{cc}]$. QED

Proof of Thm For $n \in \mathbb{N}$, $d_n = d \, (1 + n^{-1} d)$ is a metric on $X$, $d_n \in C_b(X)$. Also, $d_n \uparrow d$ on $X$. Therefore by the above Lemma, we need only to consider that $d$ is bounded. Let us then assume so: $d$ is bounded. In this case, all Lipschitz functions are bounded, and hence are integrable against probability measures. By the Kantorovich duality, we then need only to show that

$$\sup_{(\varphi, \psi) \in \Phi_d(\mu, \nu)} J[\varphi, \psi] = \sup \left\{ \int_X \varphi \, d(\mu - \nu) : \varphi \in \mathrm{Lip}_b(X), \ \mathrm{Lip}(\varphi) \le 1 \right\}.$$

By the second Lemma, the l.h.s. is $\sup_{\varphi \in L^1(\mu)} J[\varphi^{dd}, \varphi^d]$, where

$$\varphi^d(y) = \inf_{x \in X} [d(x, y) - \varphi(x)],$$

$$\varphi^{dd}(x) = \inf_{y \in Y} [d(x, y) - \varphi^d(y)].$$

Now, $\varphi^d$ is 1-Lipschitz (cf. Proof of first lemma). So

$$-\varphi^d(x) \le \underbrace{\inf_{y \in Y} [d(x,y) - \varphi^d(y)]}_{= \varphi^{dd}(x)} \underset{\uparrow}{\le} -\varphi^d(x)$$

choosing $y = x$ in inf

Hence $\varphi^{dd} = -\varphi^d$. So,

$$\sup_{(\varphi, \psi) \in \Phi_c(\mu, \nu)} J[\varphi, \psi] = \sup_{\varphi \in L^1(\mu)} J[\varphi^{dd}, \varphi^d]$$

$$= \sup_{\varphi \in L^1(\mu)} J[-\varphi^d, \varphi^d] \le \sup_{\tilde{\varphi} \in L^1(\nu) \cap \mathrm{Lip}(Y), \ \|\tilde{\varphi}\|_{\mathrm{Lip}} \le 1} J[\tilde{\varphi}, -\tilde{\varphi}]$$

$$\le \sup_{\Phi_c(\mu, \nu)} J[\varphi, \psi].$$

QED