Optimal Transport: Lecture #9

- Recall primal KP: for measures μ, ν with finite second moment, $\Pi(\mu,\nu) = \{ \gamma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \mid \gamma(A \times \mathbb{R}^d) = \mu(A),\ \gamma(\mathbb{R}^d \times B) = \nu(B) \}$ and (KP) solves

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi], \text{ where } I[\pi] := \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2}\|x-y\|^2 \, d\pi(x,y).$$

- This has dual formulation $\sup_{(\varphi,\psi) \in \Phi_c} J[\varphi,\psi]$, where $\overset{(\varphi,\psi)}{\Phi_c} \Leftrightarrow \varphi(x) + \psi(y) \leq \frac{1}{2}\|x+y\|^2$

$$J[\varphi,\psi] = \int_{\mathbb{R}^d} \varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d} \psi(y)\,d\nu(y).$$

- We saw in L7 that using special properties of quadratic cost and the assumption of finite second moment, the problem can be written as

$$\inf_{\pi \in \Pi[\mu,\nu]} I[\pi] = \sup_{(\varphi,\psi) \in \Phi_c} J[\varphi,\psi] = \inf_{(\varphi,\psi) \in \tilde{\Phi}} \tilde{J}[\varphi,\psi],$$

where $(\varphi,\psi) \in \tilde{\Phi} \Leftrightarrow x \cdot y \leq \varphi(x) + \psi(y)$. The double convexification trick established that we can focus on the subset of $\tilde{\Phi}$ of the form $(\varphi^{**}, \varphi^{*})$ for some φ.

automatically l.s.c. and convex

Theorem (Existence of Optimal Pair of Convex Conjugate Functions for $\tilde{J}$): Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have finite second moment. Let $\tilde{\Phi}$ be as above: $(\varphi,\psi) \in \tilde{\Phi} \Leftrightarrow \varphi(x) + \psi(y) \geq x \cdot y$. Then $\exists (\varphi, \varphi^*)$ s.t. each is l.s.c., convex, and proper, s.t. $\inf_{(\varphi,\psi) \in \tilde{\Phi}} \tilde{J}[\varphi,\psi] =$

$J[\varphi, \varphi^*]$.

- The basic idea is to double convexify, pass to a limit, and convexify again. That is, suppose $\{(\varphi_k, \psi_k)\}_{k=1}^\infty$ is s.t. $\lim_{k\to\infty} J[\varphi_k, \psi_k] = \inf_{(\varphi,\psi)} J[\varphi,\psi]$. We want to show that after passing to a subsequence, and using double convexification to assume (φ_k, ψ_k) are convex conjugates,

$$\left.\begin{aligned} \lim_{k\to\infty} \varphi_k &=: \varphi \\ \lim_{k\to\infty} \psi_k &=: \psi \end{aligned}\right\} (\star)$$

exist, are in $L^1(d\mu)/L^1(d\nu)$, $(\varphi,\psi) \in \hat{\Phi}$, and that $J[\varphi,\psi] \leq \lim_{k\to\infty} J[\varphi_k,\psi_k]$.
Then applying double convexification would give the result.

- To make this precise, we need to ensure $(\star)$ occurs in the correct L^1 sense, which is subtle due to the fact that even if $\{(\varphi_k, \psi_k)\}_{k=1}^\infty$ is an optimizing sequence, so is $\{(\varphi_k - C, \psi_k + C)\}_{k=1}^\infty$ for any constant C, which in general precludes L^1 convergence.

<u>Lemma (Double Convexification)</u>: Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be supported on sets X, Y satisfying $M_2 := \int_X \frac{\|x\|^2}{2}\, d\mu(x) + \int_Y \frac{\|y\|^2}{2}\, d\nu(y) < +\infty$. Let, for any measurable $\varphi, \psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$,

$$\varphi^*(y) = \sup_{x\in X} \{x \cdot y - \varphi(x)\}$$
$$\psi^*(x) = \sup_{y\in Y} \{x \cdot y - \psi(y)\}$$

Let $\tilde{\Phi} = \{(\varphi,\psi) \mid x\cdot y \leq \varphi(x) + \psi(y) \ \forall_{a.e.} x, y\}$. Let $\{(\varphi_k, \psi_k)\}_{k=1}^\infty$ be a minimizing sequence for J on $\tilde{\Phi}$. Then:

(i) One can modify $\{(\varphi_k, \psi_k)\}_{k=1}^{\infty}$ on measure 0 sets s.t.

$$x \cdot y \leq \varphi_k(x) + \psi_k(y)$$

holds for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, and without changing the value of $J[\varphi_k, \psi_k]$.

(ii) $\exists \{a_k\}_{k=1}^{\infty}$ s.t. $\{(\bar{\varphi}_k, \bar{\psi}_k)\}_{k=1}^{\infty} := \{(\varphi_k^{**} - a_k, \varphi_k^{*} + a_k)\}_{k=1}^{\infty}$ is still a minimizing sequence for J on $\tilde{\Phi}$, and s.t.

$$\text{(Lower)} \quad \forall x \in X, \ \forall y \in Y, \quad \bar{\varphi}_k(x) \geq \frac{-\|x\|^2}{2}$$
$$\bar{\psi}_k(y) \geq \frac{-\|y\|^2}{2}$$

$$\text{(Upper)} \quad \liminf_{k \to \infty} \left\{ \inf_{x \in X} \bar{\varphi}_k(x) + \frac{\|x\|^2}{2} \right\} \leq \inf_{\tilde{\Phi}} J + M_2$$

$$\liminf_{k \to \infty} \left\{ \inf_{y \in Y} \bar{\psi}_k(y) + \frac{\|y\|^2}{2} \right\} \leq \inf_{\tilde{\Phi}} J + M_2$$

(iii) With $X = Y = \mathbb{R}^d$, $*$ is the usual LT and $\inf_{\tilde{\Phi}} J = \inf_{\varphi \in L^1(d\mu)} J(\varphi^{**}, \varphi^{*})$.

Proof: Since $(\varphi_k, \psi_k) \in \tilde{\Phi}$, $x \cdot y \leq \varphi_k(x) + \psi_k(y)$ a.e.

Let N_k^x, N_k^y be s.t. $\mu[N_k^x] = \nu[N_k^y] = 0$ and $x \cdot y \leq \varphi_k(x) + \psi_k(y)$ holds on $(N_k^x)^c \times (N_k^y)^c$. Setting

$$\tilde{\varphi}_k(x) = \begin{cases} \varphi_k(x), & x \notin N_k^x \\ +\infty, & x \in N_k^y \end{cases}$$

$$\tilde{\psi}_k(y) = \begin{cases} \psi_k(y), & y \notin N_k^y \\ +\infty, & y \in N_k^y \end{cases}$$

is a measure 0 change that makes $x \cdot y \le \tilde{\varphi}_k(x) + \hat{\psi}_k(y)$ hold a.e. and $J[\varphi_k, \psi_k] = J[\tilde{\varphi}_k, \hat{\psi}_k]$. This gives (i).

To see (ii), let us construct the requisite $\{a_k\}_{k=1}^{\infty}$ as follows. By (i), we know that for an arbitrary minimizing sequence $\{(\varphi_k, \psi_k)\}_{k=1}^{\infty}$, we may WLOG assume $x \cdot y \le \varphi_k(x) + \psi_k(y)$ holds for all (x, y). Then

$$\varphi_k(x) \ge x \cdot y - \psi_k(y)$$

Since ψ_k is proper, it is not $\psi_k \equiv +\infty$ and thus we may assume there exists $y_0 \in Y$, $b_0 \in \mathbb{R}$ s.t. $\varphi_k(x) \ge x \cdot y_0 + b_0$ holds for all x.

$$\Rightarrow \varphi_k^*(y_0) = \sup_{x \in X} \{x \cdot y_0 - \varphi_k(x)\}$$

$$\le -b_0$$

$\Rightarrow \varphi_k^* \not\equiv +\infty$. By an analogous argument, φ_k^* is bounded below by an affine function. Set

$$a_k := \inf_{y \in Y} \left\{ \varphi_k^*(y) + \frac{\|y\|^2}{2} \right\} < \infty$$

Then consider $\{(\bar{\varphi}_k, \bar{\psi}_k)\}_{k=1}^{\infty} = \{(\varphi_k^{**} - a_k, \varphi_k^* + a_k)\}_{k=1}^{\infty}$. Clearly $\inf_{y \in Y} \left\{ \bar{\psi}_k(y) + \frac{\|y\|^2}{2} \right\} = 0$, and moreover

$$\bar{\varphi}_k^*(x) + \frac{\|x\|^2}{2}$$

$$= [\bar{\psi}_k]^*(x) + \frac{\|x\|^2}{2}$$

$$= \sup_{y \in Y} \left\{ x \cdot y - \bar{\psi}_k(y) + \frac{\|x\|^2}{2} \right\}$$

$$\geq \sup_{y \in Y} \left\{ -\frac{\|y\|^2}{2} - \bar{\psi}_k(y) \right\}$$

$$= 0.$$

Thus, $\bar{\varphi}_k(x) \geq -\frac{\|x\|^2}{2}$, $\bar{\psi}_k(y) \geq \frac{-\|y\|^2}{2}$, as desired.

Note that desirable properties follow by double convexification:

$$J[\bar{\varphi}_k, \bar{\psi}_k] = J[\varphi_k^{**}, \varphi_k^{*}] \leq J[\varphi_k, \psi_k] < \infty$$

$\Rightarrow \{\bar{\varphi}_k\}_{k=1}^{\infty} \subset L^1(d\mu)$, $\{\bar{\psi}_k\}_{k=1}^{\infty} \subset L^1(d\nu)$ and $\{(\bar{\varphi}_k, \bar{\psi}_k)\}_{k=1}^{\infty}$ remains a minimizing sequence.

Direct calculation yields that $\{(\bar{\varphi}_k, \bar{\psi}_k)\}_{k=1}^{\infty}$ satisfy the desired upper bounds, yielding (ii). The final statement (iii) is immediate from (ii) upon noting

$$(\bar{\varphi}_k, \bar{\psi}_k) = (\varphi_k^{**} - a_k, \varphi_k^{*} + a_k) = ([\varphi_k - a_k]^{**}, [\varphi_k - a_k]^{*})$$