Lecture #11: Optimal Transport

- We are now in a position to prove Knott-Smith optimality and Brenier's Theorem. Let's recall this omnibus result from Lecture #6.

Theorem (OT for Quadratic Cost): Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have finite second moments, that is,

$$\int_{\mathbb{R}^d} \|x\|_{\mathbb{R}^d}^2 \, d\mu(x), \quad \int_{\mathbb{R}^d} \|y\|^2 \, d\nu(y) < \infty.$$

(i) [Knott-Smith] $\pi \in \Pi(\mu, \nu)$ is optimal for (KP)

$$\Updownarrow$$

$\exists\ \varphi$ convex, l.s.c. s.t. $\operatorname{supp}(\pi) \subset \operatorname{Graph}(\partial \varphi)$.

For such a φ, the pair (φ, φ^*) is a minimizer to

$$\inf_{(\varphi, \psi)} \left\{ \int_{\mathbb{R}^d} \varphi(x) \, d\mu(x) + \int_{\mathbb{R}^d} \psi(y) \, d\nu(y) \;\middle|\; x \cdot y \leq \varphi(x) + \psi(y) \right\}.$$

(ii) [Brenier] If μ does not give mass to small sets, then $\exists!\ \pi$, optimal for (KP), of the form

$$d\pi(x, y) = d\mu(x) \cdot \delta_{\{y = \nabla\varphi(x)\}}(y),$$

where $\nabla \varphi$ is (unique up to μ-a.e. equivalence) the gradient of a convex function that pushes μ onto ν: $[\nabla \varphi]_{\#} \mu = \nu$, and $\operatorname{supp}(\nu) = \overline{\nabla \varphi(\operatorname{supp}(\mu))}$.

(iii) [Brenier for Monge] In the setting of (ii), $\nabla\varphi$ is the unique solution to the Monge Problem:

$$\int_{\mathbb{R}^d} \|\nabla\varphi(x)-x\|^2\, d\mu(x) = \inf_{\substack{T \text{ s.t.} \\ T_\#\mu=\nu}} \int_{\mathbb{R}^d} \|T(x)-x\|^2\, d\mu(x)$$

(iv) If ν also does not give mass to small sets, then

$$\nabla\varphi^* \circ \nabla\varphi(x) = x \quad \mu\text{-a.e.},$$
$$\nabla\varphi \circ \nabla\varphi^*(y) = y \quad \nu\text{-a.e.},$$

where $\nabla\varphi^*$ is the a.e. gradient of a convex function s.t.

$$\int_{\mathbb{R}^d} \|\nabla\varphi^*(y)-y\|^2\, d\nu(y) = \inf_{\substack{T \text{ s.t.} \\ T_\#\nu=\mu}} \int_{\mathbb{R}^d} \|T(y)-y\|^2\, d\nu(y).$$

Proof: We know that, by our result from Lecture 3 on Kantorovich duality, that there exists an optimal transport plan π^* s.t.

$$\pi^* = \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2\, d\pi(x,y).$$

As usual, we can boil our problem down to things in terms of $x\cdot y$, not $\|x-y\|^2$.

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2\, d\pi(x,y) \longleftrightarrow \inf_{\pi\in\Pi(\mu,\nu)} \overset{\text{sum of 2nd moments}}{M_2} - \sup_{\pi\in\Pi(\mu,\nu)} \int x\cdot y\, d\pi(x,y)$$

$$\sup_{\substack{\{\varphi(x)+\psi(y)\le\|x-y\|^2\} \\ \varphi,\psi\in L^1}} \left\{\int_{\mathbb{R}^d}\varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d}\psi(y)\,d\nu(y)\right\} \longleftrightarrow M_2 - \inf_{\underbrace{\{x\cdot y\le\varphi(x)+\psi(y)\}}_{\tilde{\Phi}}} \overbrace{\left\{\int_{\mathbb{R}^d}\varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d}\psi(y)\,d\nu(y)\right\}}^{J}$$

Now, note that if $\pi \in \Pi(\mu,\nu)$ and $T_{\#}\mu = \nu$, then

$$d\pi\text{-a.e. } y = T(x) \iff \pi = (Id \times T)_{\#}\mu \quad \circledast$$

Now, by our Theorem from the last lecture, there exist convex conjugate functions (φ, φ^*) optimal in the sense that $J[\varphi,\varphi^*] = \inf_{\tilde{\Phi}} J$. Then

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} x\cdot y \, d\bar{\pi}(x,y) = \int_{\mathbb{R}^d} \varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d} \varphi^*(y)\,d\nu(y)$$

$$= \iint_{\mathbb{R}^d\times\mathbb{R}^d} [\varphi(x) + \varphi^*(y)]\, d\bar{\pi}(x,y)$$

$$\iff \int_{\mathbb{R}^d\times\mathbb{R}^d} [\varphi(x) + \varphi^*(y) - x\cdot y]\, d\bar{\pi}(x,y) = 0$$

By properties of convex conjugates, we know $\boxed{\varphi(x) + \varphi^*(y) - x\cdot y \geq 0}$

$$\iff [\varphi(x) + \varphi^*(y) - x\cdot y] = 0 \quad \bar{\pi}\text{-a.e.}$$

$\iff$ for $d\bar{\pi}$-a.e. (x,y), $y \in \partial\varphi(x)$.

This gives "$\Rightarrow$" for Knott-Smith. Suppose now that $y \in \partial\varphi(x)$ for $d\bar{\pi}$-a.e. (x,y). Following the calculations above, we have

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (x\cdot y)\, d\pi(x,y) = \int_{\mathbb{R}^d} \varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d} \varphi^*(y)\,d\nu(y)$$

This implies that $\bar{\pi}$ is optimal by Kantorovich duality, establishing (i).

Suppose now μ does not give mass to small sets. Let φ be as above (convex, l.s.c., optimal with φ^* for (KP)). Then $\mathrm{Dom}(\varphi)$ is convex

→ really, this gives its boundary has measure 0

$$\Rightarrow \mu[\mathrm{Int}(\mathrm{Dom}(\varphi))] = 1.$$

Similarly, $\varphi \in L^1(d\mu) \Rightarrow \mu[\mathrm{Dom}(\varphi)] = 1$. On $\mathrm{Int}(\mathrm{Dom}(\varphi))$, the set of points at which φ is not differentiable is a small set (φ is convex). So, we have that $d\mu$-a.e. point of $\mathrm{Dom}(\varphi)$ is a point of differentiability

$$\Rightarrow \partial\varphi(x) = \{\nabla\varphi(x)\} \quad d\mu\text{-a.e.}$$

$$\Rightarrow y = \nabla\varphi(x) \quad \text{for } d\pi\text{-a.e. } (x,y),$$

which establishes, together with (★), the existence of a map needed in Brenier. It thus suffices to establish uniqueness.

Let $\varphi, \bar{\varphi}$ both be convex and s.t. $(\nabla\varphi)_\# \mu = (\nabla\bar{\varphi})_\# \mu = \nu$. We aim to show these are equal μ-a.e. By (i), $(\bar{\varphi}, \bar{\varphi}^*)$ is an optimal pair of dual variables, same with (φ, φ^*)

$$\Rightarrow \int_{\mathbb{R}^d} \varphi(x)\,d\mu(x) + \int_{\mathbb{R}^d} \varphi^*(y)\,d\nu(y) = \int_{\mathbb{R}^d} \bar{\varphi}(x)\,d\mu(x) + \int_{\mathbb{R}^d} \bar{\varphi}^*(y)\,d\nu(y) \quad (★★)$$

Suppose $\hat{\pi}$ is optimal for φ, i.e. $(\mathrm{Id} \times \nabla\varphi)_\# \mu = \hat{\pi}$. Then

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} [\bar{\varphi}(x) + \bar{\varphi}^*(y)]\,d\hat{\pi}(x,y)$$

margin constraints + (★★)

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} [\varphi(x) + \varphi^*(y)]\,d\hat{\pi}(x,y)$$

$$= \int_{\mathbb{R}^d\times\mathbb{R}^d} (x\cdot y)\, d\bar{\pi}(x,y)$$

$\bar{\pi} = (Id \times \nabla\varphi)_{\#}\mu$

$\Longrightarrow$

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \left[\bar{\varphi}(x) + \bar{\varphi}^*(\nabla\varphi(x))\right] d\mu(x) = \int_{\mathbb{R}^d} x\cdot\nabla\varphi(x)\, d\mu(x)$$

$$\Rightarrow \int_{\mathbb{R}^d\times\mathbb{R}^d} \left[\overbrace{\bar{\varphi}(x)+\bar{\varphi}^*(\nabla\varphi(x)) - x\cdot\nabla\varphi(x)}^{\boxed{\geq 0}\ \forall x}\right] d\mu(x) = 0$$

$$\Rightarrow \bar{\varphi}(x)+\bar{\varphi}^*(\nabla\varphi(x)) - x\cdot\nabla\varphi(x) = 0 \quad \mu \text{ a.e.}$$

$$\Rightarrow \nabla\varphi(x) \in \partial\bar{\varphi}(x) \text{ a.e.}$$

$\boxed{\Rightarrow}$ $\nabla\varphi(x) = \nabla\bar{\varphi}(x)$ a.e., which gives the needed uniqueness.

The set inclusions follow by technical arguments; (iii) is immediate; (iv) follows by similar analysis to above.