OT: Lecture #21

- So far, we have discussed convexity in the linear sense (i.e. the convex combinations of μ_0, μ_1 are the mixtures of the form $[1-t]\cdot\mu_0 + t\cdot\mu_1$ for $t \in [0,1]$). This does not respect the geometric structure of Wasserstein space and motivates the following.

Defn: Let X be a geodesic metric space and $F: X \to \mathbb{R} \cup \{+\infty\}$ a functional on X. We say F is geodesically convex if $\forall x_0, x_1 \in X$ there exists a constant speed geodesic $\omega: [0,1] \to X$ s.t. (i) $\omega(0) = x_0$, $\omega(1) = x_1$

(ii) $t \mapsto F(\omega(t))$ is convex.

Remark: if geodesics in X are line segments (e.g., in $\mathbb{R}^d$ or any normed vector space), this reduces to the usual notion of convex functional.

- In our setting of W_p equipped with the Wasserstein-p metric, we know that geodesics are given by certain pushforwards related to the optimal coupling (or map in certain cases), namely if $\mu_0, \mu_1 \in W_p$ have OT plan π, then a constant speed geodesic is given by

$$\mu_t = [P_t]_\# \pi, \qquad P_t(x,y) = (1-t)x + ty$$

- Let's look at conditions that guarantee certain functionals we've become familiar with are geodesically convex, which we call displacement convex in Wasserstein space, after McCann.

Proposition (Displacement Convexity of Potential Energies): Suppose $V: \mathbb{R}^d \to \mathbb{R}$ and let $\mathcal{V}: \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ be $\mathcal{V}(\mu) := \int V \, d\mu$. Then $\mathcal{V}$ is displacement convex $\Leftrightarrow$ V is convex.

Proof: $\Rightarrow$ Consider geodesics of the form $\mu_t = \delta_{[1-t]x+ty}$. Then notice $\mathcal{V}(\mu_t) = V([1-t]x+ty)$, so geodesic convexity of $\mathcal{V}$ gives convexity of V.

$\Leftarrow$: Suppose V is convex and let μ_t be a constant speed geodesic connecting μ_0 and μ_1. Then

$$\mathcal{V}(\mu_t) = \int_{\mathbb{R}^d} V(z)\, d\mu_t(z)$$

$$= \int_{\mathbb{R}^d} V_{(z)}\, d[P_t]_\# \pi(z)$$

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} V([1-t]x+ty)\, d\pi(x,y)$$

$$\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} \left([1-t]V(x) + tV(y)\right) d\pi(x,y)$$

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} [1-t]V(x)\, d\pi(x,y) + \int_{\mathbb{R}^d \times \mathbb{R}^d} tV(y)\, d\pi(x,y)$$

$$= [1-t]\int_{\mathbb{R}^d} V(x)\, d\mu_0(x) + t\int_{\mathbb{R}^d} V(y)\, d\mu_1(y)$$

$$= [1-t]\mathcal{V}(\mu_0) + t\mathcal{V}(\mu_1).$$

Proposition (Displacement Convexity of Interaction Energies): Suppose $W: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and let $\mathcal{W}: \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ be given by $\mathcal{W}(\mu) := \iint_{\mathbb{R}^d\, \mathbb{R}^d} W(x,y)\, d\mu(x)\, d\mu(y)$. Then

$\mathcal{W}$ is displacement convex if W is convex.

__Proof__: Let $\{\mu_t\}_{t\in[0,1]}$ be a constant speed geodesic between μ_0 and μ_1. Then

$$\mathcal{W}(\mu_t) = \int_{\mathbb{R}^d\times\mathbb{R}^d} W(x,y)\, d\mu_t(x)\, d\mu_t(y)$$

$$= \int_{\mathbb{R}^d\times\mathbb{R}^d} W(x,y)\, d[P_t]_\# \pi(x)\, d[P_t]_\# \pi(y)$$

$$= \int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d} W([1-t]x+ty, [1-t]\tilde{x}+t\tilde{y})\, d\pi(x,y)\, d\pi(\tilde{x},\tilde{y})$$

$$= \int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d} W\big([1-t](x,\tilde{x}) + t(y,\tilde{y})\big)\, d\pi(x,y)\, d\pi(\tilde{x},\tilde{y})$$

$$\leq \int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d} \big([1-t] W(x,\tilde{x}) + t W(y,\tilde{y})\big)\, d\pi(x,y)\, d\pi(\tilde{x},\tilde{y})$$

$$= (1-t)\int_{\mathbb{R}^d\times\mathbb{R}^d} W(x,\tilde{x})\, d\mu_0(x)\, d\mu_0(\tilde{x}) + t\int_{\mathbb{R}^d\times\mathbb{R}^d} W(y,\tilde{y})\, d\mu_1(y)\, d\mu_1(\tilde{y})$$

$$= (1-t)\,\mathcal{W}(\mu_0) + t\,\mathcal{W}(\mu_1). \quad \blacksquare$$

__Remark__: In fact, $\mathcal{W}$ convex iff W is convex; see Prop 7.25 in Santambrogio.

- Perhaps counterintuitively, $\mu \mapsto W_2^2(\mu,\nu)$ is not ~~[scribbled]~~ displacement convex:

__ex__: Let $\nu = \frac{1}{2}\delta_{(1,0)} + \frac{1}{2}\delta_{(-1,0)}$

$$\mu_t = \frac{1}{2}\delta_{(t,a)} + \frac{1}{2}\delta_{(-t,-a)}$$

If $a > 1$, the curve μ_t is the geodesic between $\frac{1}{2}\delta_{(-1,a)} + \frac{1}{2}\delta_{(1,-a)}$ and $\frac{1}{2}\delta_{(1,a)} + \frac{1}{2}\delta_{(-1,-a)}$ (exercise). But, $W_2^2(\mu_t,\nu) = a^2 + \min\{(1-t)^2, (1+t)^2\}$. (exercise), which is not convex in t.

- This motivates ~~[scribbled]~~ a new notion of geodesic, such that $\nu \mapsto W_2^2(\mu,\nu)$ is in fact "convex."

Def^n: Fix $\rho \in \mathcal{P}(\Omega)$. Let $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$. We call the following curve the generalized geodesic between μ_0 and μ_1 with base ρ:

$$\mu_t := \left[(1-t)T_{\rho\to\mu_0} + t\, T_{\rho\to\mu_1}\right]_\# \rho,$$

where $T_{\rho\to\mu_0}, T_{\rho\to\mu_1}$ are the respective OT maps from ρ to μ_i.

- So, we are fixing a base measure, taking convex combinations in the linear space $L^2(\Omega,\rho)$, then pushing forward. In fact, this idea is useful in a lot of situations and sometimes goes under the name "linearized OT."

- Note that for any ρ and with $\{\mu_t\}_{t\in[0,1]}$ as above, we have convexity along the curve:

$$W_2^2(\mu_t, \rho) = \min_{T_\# \rho = \mu_t} \int_{\mathbb{R}^d} \| T(x) - x \|^2 \, d\rho(x) \qquad (T_i := T_{\rho \to \mu_i})$$

$T = (1-t)T_0 + tT_1$

$$\leq \int_{\mathbb{R}^d} \| [1-t] T_0(x) + t T_1(x) - x \|^2 \, d\rho(x)$$

$$\leq (1-t) \int_{\mathbb{R}^d} \| T_0(x) - x \|^2 \, d\rho(x) + t \int_{\mathbb{R}^d} \| T_1(x) - x \|^2 \, d\rho(x)$$

$$= (1-t) W_2^2(\mu_0, \rho) + t \, W_2^2(\mu_1, \rho), \quad \text{as needed.}$$

- The key point is the T above is a transport map, but in general not the OT map.