Lecture #16: Optimal Transport

Theorem: Suppose that $\Omega \subseteq \mathbb{R}^d$ is either bounded or equal to all of $\mathbb{R}^d$. Let $V_t : \Omega \to \mathbb{R}^d$ be Lipschitz in x uniformly in t, and uniformly bounded. Let $Y_t : \Omega \to \mathbb{R}^d$ be its induced flow. Suppose that for every $x \in \Omega$ and every $t \in [0,T]$, we have $Y_t(x) \in \Omega$. Then, for every probability measure $\tilde{\rho}_0 \in \mathcal{P}(\Omega)$, the measures induced by the flow

$$\rho_t := (Y_t)_{\#} \tilde{\rho}_0$$

Solve the continuity equation

$$(CE) \quad \begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t V_t) = 0 \\ \rho_t = \tilde{\rho}_0 \quad \text{at } t = 0 \quad \text{(i.e. initial distribution is } \tilde{\rho}_0\text{)} \end{cases}$$

Moreover, if ρ_t is a solution to the above (CE) with ρ_t a.c. w.r.t. Lebesgue measure on $\mathbb{R}^d$ for all $t \in [0,T]$, then $\rho_t = (Y_t)_{\#} \rho_0$. In particular, (CE) admits a unique solution.

Proof: Fix a test function $\phi \in C^1(\mathbb{R}^d)$ such that ϕ and $\nabla \phi$ are bounded. We want to show that (CE) is weakly satisfied when ρ_t is obtained by such a flow. We want to show

$$\frac{d}{dt} \int_\Omega \phi \, d\rho_t = \int_\Omega \nabla \phi \cdot V_t \, d\rho_t .$$

Well,

$$\frac{d}{dt} \int_\Omega \phi \, d\rho_t \overset{\boxed{\text{flow condition}}}{=} \frac{d}{dt} \int_\Omega \phi(y_x(t)) \, d\rho_0(x)$$

$$= \int_{\Omega} \nabla\phi(y_x(t)) \cdot y_x'(t)\, d\rho_0(x)$$

definition of ODE inducing the flow

$$= \int_{\Omega} \nabla\phi(y_x(t)) \cdot V_t(y_x(t))\, d\rho_0(x)$$

flow condition

$$= \int_{\Omega} \nabla\phi(y) \cdot V_t(y)\, d\rho_t(y)$$

This shows the kinds of purported solutions given by "flowing" ρ_0 are in fact weak solutions.

To show the converse, that every weak solution is in fact of this form, we proceed as follows. Suppose that $\{(\rho_t, V_t)\}$ is a weak and therefore distributional solution. Then for all bounded Lipschitz test function $\phi \in C^1_c((0,T)\times\Omega)$ satisfies

$$\int_0^T \int_{\Omega} \left[\partial_t \phi + V_t \cdot \nabla\phi\right] d\rho_t(x)\, dt = 0$$

By a smoothing argument, we can drop the C^1 assumption and just suppose ϕ is Lipschitz and compactly supported, as long as ρ_t is a.c. with respect to Lebesgue measure.

Now, let $\psi \in C_1^c(\mathbb{R}^d)$ and let $\phi(t,x) := \psi((Y_t)^{-1}x)$. If we can prove that

$$t \mapsto \int \phi(t,x)\, d\rho_t(x)$$

is in fact constant in t, then this will show $\rho_t = (Y_t)_\# \rho_0$, as desired. What properties does ϕ have? Well, Y_t is bi-Lipschitz for each t uniformly, so $\phi(t,x)$ is Lipschitz. It is compactly supported in x. Indeed, let

$$M := \sup_{t,x} |V_t(x)| < \infty$$

Then if $\mathrm{dist}(x, \underbrace{\mathrm{supp}(\psi)}_{\text{compact}}) > tM$, we have that $\phi(t,x)=0$.

Unfortunately, we do not have compact support in time, so we cannot say ϕ is Lipschitz and compactly support and therefore gives 0 in the (CE). We get around this with a cut-off function as follows. Let $\chi \in C_c^1((0,T))$. Then

constant w.r.t. x

$$\partial_t(\chi\phi) + v_t \cdot \nabla(\chi\phi) = \chi'(t)\phi(t,x) + \chi(t)\underbrace{\left(\partial_t \phi(t,x) + v_t \nabla\phi(t,x)\right)}_{=0 \text{ by definition of } \phi}$$

(exercise, or see Santambrogio Box 6.3/ transport equation literature)

Thus, we have that

$$\int_0^T\int_\Omega \partial_t(\chi\phi) + v_t\cdot\nabla(\chi\phi)\, d\rho_t(x)\, dt = \int_0^T\int_\Omega \chi'(t)\,\phi(t,x)\, d\rho_t(x)\, dt$$

$$\parallel$$

$$\int_0^T dt \int_\Omega \Big[\underbrace{\partial_t(\chi\phi) + v_t\cdot\nabla(\chi\phi)\, d\rho_t(x)}_{=0 \text{ by (CE)}}\Big]$$

$\Rightarrow$ for arbitrary χ, $\displaystyle\int_0^T\int_\Omega \chi'(t)\,\phi(t,x)\, d\rho_t(x)\,dt = 0.$

Since this holds for arbitrary χ, we must have that

$$t \mapsto \int_\Omega \phi(t,x)\, d\rho_t(x)$$

is constant, as desired. ∎

· Let us return to the setting of Wasserstein space. We can connect (CE) to curves in Wasserstein space as follows:

Theorem: Let $\{\mu_t\}_{t\in[0,1]}$ be an a.c. curve in $W_p(\Omega)$ for $p>1$, $\Omega\subset\mathbb{R}^d$ compact. Then for a.e. $t\in[0,1]$, there exists a vector field $v_t\in L^p(\mu_t,\mathbb{R}^d)$ such that:

(i) The continuity equation $\partial_t\mu_t+\nabla\cdot(v_t\mu_t)=0$ weakly

(ii) For a.e. t, we have $\|v_t\|_{L^p(\mu_t)}\le|\mu'|(t)$, where $|\mu'|(t)$ is the metric derivative of $t\mapsto\mu_t$ evaluated at time t.

Conversely, if $\{\mu_t\}_{t\in[0,1]}\subset\mathcal{P}_p(\Omega)$ and for each t we have a vector field $v_t\in L^p(\mu_t;\mathbb{R}^d)$ with $\int_0^1\|v_t\|_{L^p(\mu_t)}\,dt<\infty$ solving the continuity equation $\partial_t\mu_t+\nabla\cdot(v_t\mu_t)=0$, then $t\mapsto\mu_t$ is an a.c. curve in $W_p(\Omega)$ and for a.e. t, $|\mu'|(t)\le\|v_t\|_{L^p(\mu_t)}$.

· We will not prove this result, but it is done in Chapter 5.3 in OTAM.

· We can use this result to give the so-called *dynamical formulation* of OT, pioneered by Benamou and Brenier.

· To do so, we need to establish some properties of Wasserstein geodesics.

Theorem (McCann Interpolations are Wasserstein Geodesics): Suppose Ω is convex. Let $\mu,\nu\in\mathcal{P}_p(\Omega)$ and $\gamma\in\Pi(\mu,\nu)$ an OT plan realizing the Kantorovich problem for cost $c(x,y) = \|x-y\|^p$, $p \geq 1$. Define ⑤

$$\Pi_t : \Omega \times \Omega \to \Omega$$
$$(x,y) \mapsto [1-t]x + ty$$

Then the curve $\mu_t := [\Pi_t]_\# \gamma$ is a constant speed geodesic in W_p that starts at $\mu_0 = \mu$ and ends at $\mu_1 = \nu$. Moreover, if μ is a.c. or if γ is concentrated on the graph of the OT map $T: \Omega \to \Omega$ realizing the solution to the Monge problem, then $\Pi_t = ([1-t]\cdot \mathrm{Id} + t\cdot T)_\# \mu$.

Proof: Note that it is enough to show $W_p(\mu_t, \mu_s) \leq W_p(\mu,\nu)\cdot|t-s|$ for all t,s. (A) If we get this, then for $t<s$

$$W_p(\mu,\nu) \leq W_p(\mu,\mu_t) + W_p(\mu_t,\mu_s) + W_p(\mu_s,\nu)$$
$$\leq W_p(\mu,\nu)\cdot[t + s-t + 1-s]$$
$$= W_p(\mu,\nu)$$

$\Rightarrow$ equality holds, i.e. $W_p(\mu_t,\mu_s) = W_p(\mu,\nu)\cdot|t-s|$. So, let's show (A). Let $\gamma_t^s := (\Pi_t, \Pi_s)_\# \gamma \in \Pi(\mu_t,\mu_s)$. Then,

$$W_p(\mu_t,\mu_s) \leq \left[\int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^p \, d\gamma_t^s(x,y)\right]^{\frac{1}{p}}$$

$$= \left[\int_{\mathbb{R}^d\times\mathbb{R}^d} \|\Pi_t(x,y) - \Pi_s(x,y)\|^p \, d\gamma\right]^{\frac{1}{p}}$$

$$= \left[\int_{\mathbb{R}^d\times\mathbb{R}^d} \|(t-s)(x-y)\|^p \, d\gamma\right]^{\frac{1}{p}} = |t-s|\cdot W_p(\mu,\nu). \blacksquare$$