Dynamical Optimal Transport and Multimarginal Optimal transport (along with flow matching)

Dynamical optimal transport tries to find not just a map to transpot one measure to other. but a path on which the transportation could be made.

Benamou – Brenier Theorem:

Suppose $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. then

$$W_2^2(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int \|x-y\|^2 \, d\pi(x,y)$$

$$= \min \left\{ \int_0^1 \int_{\mathbb{R}^d} \|\underbrace{v_t}_{\text{Velocity field.}}(x)\|^2 \, d\mu_t(x)\, dt \;:\; \frac{d}{dt}\mu_t + \operatorname{div}(v_t \mu_t) = 0 \text{ in } (0,1) \times \mathbb{R}^d \right\}.$$

where minimization is among all μ_t in the space of probability measures such that $\mu_0 = \mu$ and $\mu_1 = \nu$.

We show a rough sketch of the proof.

Proof (Outlines):

First we show

$$W_2^2(\mu, \nu) \geq \int_0^1 \int_{\mathbb{R}^d} \|v_t(x)\|^2 \, d\mu_t(x)$$

for some (v_t, μ_t) s.t.

$$\frac{d\mu_t}{dt} + \operatorname{div}(v_t \mu_t) = 0$$

$$\mu_0 = \mu \text{ and } \mu_1 = \nu$$

consider $\pi^{MK} \in \Pi(\mu, \nu)$ which solves Monge Kantorovich for $c(x,y) = \|x-y\|^2$ between μ and ν.

Define $T_t(x,y) = (1-t)x + ty$ for $(x,y) \in \text{support}(\pi^{MK})$.

Note $T_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$.

Define $\mu_t = T_t \# \pi^{MK}$. then $\mu_0 = \mu$ and $\mu_1 = \nu$.

Consider a signed measure

$$T_t \# \left((y-x)\pi^{MK}\right)$$

this measure is obtained by multiplying $(y-x)$ to the density of π^{MK}.

is absolutely continuous w.r.t. μ_t with density v_t. (this is how v_t is defined)

Note $v_t(\cdot) : \mathbb{R}^d \to \mathbb{R}$.

$$\therefore \quad \int_{\mathbb{R}^d} \|v_t(x)\|^2 \, d\mu_t(x)$$

$$= \int \|y-x\|^2 \, d\pi^{MK}(x,y)$$

$$\Rightarrow \int_0^1 \int_{\mathbb{R}^d} \|v_t(x)\|^2 \, d\mu_t(x) \, dt$$

$$= \int \|y-x\|^2 \, d\pi^{MK}(x,y).$$

Have to now show (μ_t, v_t) satisfies the continuity equation.

Roughly,

$$\int \phi(z) \frac{d\mu_t(z)}{dt} \, dz = \frac{d}{dt} \int \phi(z) \, d\mu_t(z)$$

$$= \frac{d}{dt} \int \phi\big((1-t)x + ty\big) \, d\pi^{MK}(x,y)$$

$$= \int \langle \nabla \phi((1-t)x+ty), y-x \rangle \, d\pi^{MK}(x,y)$$

$$= \int \langle \nabla \phi(z), v_t(z) \rangle \, d\mu_t(z)$$

$$= \int \phi(z) \, \mathrm{div}\big(v_t(z) \, d\mu_t(z)\big)$$

Now we show.

$$W_2^2(\mu, \nu) \leq \int_0^1 \int_{\mathbb{R}^d} \|v_t(x)\|^2 \, d\mu_t(x) \, dt$$

where (μ_t, v_t) satisfies

$$\frac{d}{dt}\mu_t + \mathrm{div}(v_t \mu_t) = 0.$$

Suppose (μ_t, v_t) is a pair satisfying

$$\frac{d}{dt}\mu_t + \mathrm{div}\,(v_t \mu_t) = 0 \quad \text{and} \quad \begin{array}{l}\mu_0 = \mu. \\ \mu_1 = \nu\end{array}$$

v_t is the velocity field.

Define a flow X_t as

$$X_t = X_0 + \int_0^t v_s(X_s)\, ds$$

$$dX_t = v_s(X_t)\, dt.$$

Density of X_t at time is given by

$$\frac{d}{dt}\mu_t + \mathrm{div}\,(\mu_t v_t) = 0.$$

Since $\mu_0 = \mu$ and $\mu_1 = \nu$,
(X_0, X_1) is a coupling of μ, ν

$$W_2^2(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int \|x-y\|^2\, d\pi$$

$$\leq \mathbb{E}_{\{X_t : [0,1]\}} \|X_0 - X_1\|^2$$

$$= \mathbb{E}_{\{X_t : [0,1]\}} \left\| \int_0^1 v_t(X_t)\, dt \right\|^2$$

$$\leq \int_0^1 \mathbb{E}_{\{X_t : \in [0,1]\}} \|v_t(X_t)\|^2\, dt$$

$$= \int_0^1 \int_{\mathbb{R}^d} \|v_t(x)\|^2 d\mu_t(x)\, dt$$

Hence the proof is complete.

Dynamic Entropic Optimal Transport.

There is a Benamou-Brenier type description of dynamic entropic optimal transport.

For instance,

$$\inf_{\pi \in \Pi(\mu,\nu)} \int \|x-y\|^2 \, d\pi(x,y) + \varepsilon \, KL(\pi \,\|\, \mu \otimes \nu)$$

$$= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} \left(\|v_t(x)\|^2 + \frac{\varepsilon}{4} \|\nabla \log(\mu_t)(x)\|^2 \right) d\mu_t(x) \, dt : \begin{array}{l} \frac{d}{dt}\mu_t + \mathrm{div}(v_t \mu_t) = 0 \\ \text{and} \\ \mu_0 = \mu, \ \mu_1 = \nu \end{array} \right\}$$

We will not prove this relation here.

Instead we describe how the solution (μ_t, v_t) of the above variational problem looks like.

Let π^ε be the solution of EOT.

Let $\gamma_t(x,y)$ be a Brownian bridge with end point (x,y).

$\Rightarrow \gamma_t : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$

Let $\pi_t^\varepsilon = \mathrm{Law}(\gamma_t \# \pi^\varepsilon)$, similarly $v_t^\varepsilon(x)$ defined by $\mathrm{Law}((y-x)\gamma_t \# \pi^\varepsilon) = v_t^\varepsilon \pi_t^\varepsilon$