Week 9

Theorem: Let $1 \le p < +\infty$, $x_0 \in X$. Let $\{\mu_n\}_n \subseteq \mathcal{P}_p(X)$, $\mu \in \mathcal{P}_p(X)$.

Then, it is equivalent:

① $\mu_n \longrightarrow \mu$ narrowly and $\int_X d^p(x,x_0)\, d\mu_n \longrightarrow \int_X d^p(x,x_0)\, d\mu$

② $W_p(\mu_n,\mu) \longrightarrow 0$

Remark: Let $\{\mu_n\}_n \subseteq \mathcal{P}(X)$, $\mu \in \mathcal{P}(X)$. Then: $\mu_n \xrightarrow{*} \mu$ weakly* $\Leftrightarrow$ $\mu_n \longrightarrow \mu$ narrowly.

Proof: we showed ② $\Rightarrow$ ① in general, ① $\Rightarrow$ ② for X compact.

① $\Rightarrow$ ② in general. Fix $\delta > 0$, define $M_n := \int (1 + d^p(x,x_0))\, d\mu_n$

$\downarrow$ by assumption

$M := \int (1 + d^p(x,x_0))\, d\mu$

Step 1: $$\underbrace{\frac{1 + d(x,x_0)^p}{M_n}\mu_n}_{\in\, \mathcal{P}(X)} \longrightarrow \frac{1 + d(x,x_0)^p}{M}\mu \text{ narrowly } (*).$$

Indeed, testing with $\varphi \in C_c(X)$, we observe $(1 + d(x_0,x)^p)\varphi(x) \in C_c(X)$.

Use this test in $\mu_n \longrightarrow \mu$: $\frac{1}{M_n}\int (1 + d^p(x,x_0))\varphi\, d\mu_n \longrightarrow \frac{1}{M}\int (1 + d^p(x,x_0))\varphi\, d\mu$.

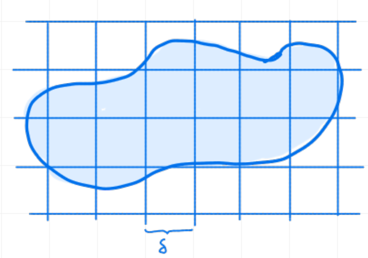
So $(*)$ is true with weak-* convergence. By Remark, also for narrow convergence.

Step 2: By Prokhorov's theorem applied to $(*)$, $\exists K \subseteq X$ compact such that:

$$\int_{X \setminus K} (1 + d^p(x,x_0))\, d\left(\frac{\mu_n}{M_n} + \frac{\mu}{M}\right) \le \delta \quad \forall n \Rightarrow \int_{X \setminus K} (1 + d^p(x,x_0))\, d(\mu_n + \mu) \le 3M\delta \quad \forall n \text{ large}$$

$\hookrightarrow M/M_n \le 2$

Step 3: Let Q_i be a grid of cubes of size δ covering K.



By weak convergence, $\mu_n(Q) \to \mu(Q)$ $(+)$ for "many" cubes Q (as soon as $\mu(\partial Q) = 0$).

Since X locally compact, we can find a finite family of $(\varphi_i)_{i \in I} \subseteq C_c(X)$, $\varphi_i \ge 0$, s.t. $\sum_{i \in I} \varphi_i \le 1$ in X, $\sum_{i \in I} \varphi_i \equiv 1$ in K, $\operatorname{diam}(\operatorname{supp}(\varphi_i)) \le \delta$ for all $i \in I$.

Set $\Lambda_{n,i} := \int_X \varphi_i \, d\mu_n$, $\Lambda_i := \int_X \varphi_i \, d\mu$, $\lambda_{n,i} := \min\{\Lambda_{n,i}, \Lambda_i\}$

and define the measures: $\alpha_{n,i} := \frac{\lambda_{n,i}}{\Lambda_{n,i}} \varphi_i \, d\mu_n$, $\beta_{n,i} := \frac{\lambda_{n,i}}{\Lambda_i} \varphi_i \, d\mu$,

$$\alpha_n := \mu_n - \sum_i \alpha_{n,i}, \quad \beta_n := \mu - \sum_i \beta_{n,i}.$$

Note $\alpha_{n,i}(X) = \beta_{n,i}(X) = \lambda_{n,i}$, $\alpha_n(X) = \beta_n(X) = 1 - \sum_i \lambda_{n,i}$.

Take $\gamma_n := \sum_i \frac{\alpha_{n,i} \otimes \beta_{n,i}}{\lambda_{n,i}} + \frac{\alpha_n \otimes \beta_n}{1 - \sum_i \lambda_{n,i}}$.

probability measures

We can check:

- $\gamma_n \in \Gamma(\mu_n, \mu)$. $\pi_1 \# \gamma_n = \sum_i \pi_1 \# \frac{\alpha_{n,i} \otimes \beta_{n,i}}{\lambda_{n,i}} + \pi_1 \# \frac{\alpha_n \otimes \beta_n}{1 - \sum_i \lambda_{n,i}}$
$$= \sum_i \alpha_{n,i} + \alpha_n = \mu_n.$$

- $W_p^p(\mu_n, \mu) \leq \int d(x,y)^p \, d\gamma_n(x,y) =$

integrates $\lambda_{n,i}$

integrates 1

$\leq 2^p d(x,x_0)^p + 2^p d(y,x_0)^p$

$$= \sum_i \int d(x,y)^p \frac{d\alpha_{n,i}(x) \, d\beta_{n,i}(y)}{\lambda_{n,i}} + \int d(x,y)^p \frac{d\alpha_n(x) \, d\beta_n(y)}{1 - \sum_i \lambda_{n,i}}$$

$$\leq \delta^p \overbrace{\sum_{i \in I} \lambda_{n,i}}^{\leq 1} + 2^p \int_{K \cup K^c} d(x,x_0)^p \, d\alpha_n(x) + 2^p \int_{K \cup K^c} d(y,x_0)^p \, d\beta_n(y)$$

$$\leq \delta^p + 2^p \sup_K d(x,x_0)^p \underbrace{\alpha_n(K)}_{\downarrow\, n \to \infty \atop 0} + \underbrace{\int_{X \setminus K} d(x,x_0)^p \overbrace{d\alpha_n(x)}^{\leq \mu_n}}_{\leq 3M\delta} + \ldots \left(\text{Same with } \beta\right) \ldots$$

$$\alpha_n(K) = \mu_n(K) - \sum_i \int_K \underbrace{\frac{\lambda_{n,i}}{\Lambda_{n,i}}}_{\to 1 \text{ as } n \to \infty} \varphi_i \, d\mu_n \longrightarrow \mu_n(K) - \int_K \underbrace{\sum \varphi_i}_{= 1} \, d\mu_n = 0$$

$\mu_n \xrightarrow{*} \mu$

$\square$

Same for $\beta_n(x)$.

Geodesics in the Wasserstein space

Def: A curve $\eta(t): [0,1] \to X$ is a (constant speed) geodesic

if $d(\eta(t), \eta(s)) = |t-s|\, d(\eta(0), \eta(1)) \quad \forall\, 0 \le s \le t \le 1.$ $\quad (\times)$

Remark (motivation): Let M be an (embedded) manifold. A constant speed geodesic between x and $y \in M$ would be

$$\underset{\substack{\eta(0)=x \\ \eta(1)=y}}{\operatorname{argmin}} \left(\int_0^1 |\eta'|^2 \, dt \right)^{1/2} \quad (1) \qquad (*)$$

or equivalently: $$\underset{\substack{\eta(0)=x \\ \eta(1)=y}}{\operatorname{argmin}} \int_0^1 |\eta'| \, dt \quad (2)$$

← this function is invariant by reparametrization of the curve

(reparametrized to be) constant speed minimizer

↪ namely, $|\eta'| \equiv$ constant

I consider the geodesic distance on the manifold

Recall $\int |\eta'| \le \left(\int |\eta'|^2 \right)^{1/2} \Longrightarrow (1) \geq (2)$. Equality $\Longleftrightarrow |\eta'| \equiv$ constant.

Since any min. in (2) can be reparametrized to be constant speed, we get $(1) = (2)$.

Lemma: $\eta(t)$ geodesic $(*)$ $\Longrightarrow$ $(\times)$ holds.

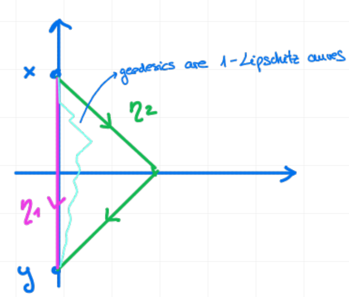
- $d(\eta(0), \eta(1)) \le d(\eta(0), \eta(s)) + d(\eta(s), \eta(t)) + d(\eta(t), \eta(1))$

  $\le (s + (t-s) + (1-t))\, d(\eta(0), \eta(1)) = d(\eta(0), \eta(1))$.

  $|\eta'| \equiv d(\eta(0), \eta(1))$, $\; d(\eta(0), \eta(s)) \le \int_0^s |\eta'|$

  → in all the inequalities we actually get equalities.

Example: $(\mathbb{R}^2, \|\cdot\|_{\ell^\infty})$, $\; x = (0,1),\; y = (0,-1)$. $\qquad \|(x,y)\|_{\ell^\infty} = \sup\{|x|, |y|\}$



$\eta_1(t) = (0, 1-2t)$ is a geodesic from x to y

$$\eta_2(t) = \begin{cases} (2t, 1-2t) & t \in (0, 1/2) \\ (2(1-t), 1-2t) & t \in (1/2, 1) \end{cases}$$ is a geodesic.

Theorem: Let $X = \mathbb{R}^d$, $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. Let γ optimal coupling for W_p.

Then, a Wasserstein geodesic between μ and ν is $\boxed{\mu_t = \pi_t \# \gamma,}$

where $\pi_t(x,y) = (1-t)x + ty$.

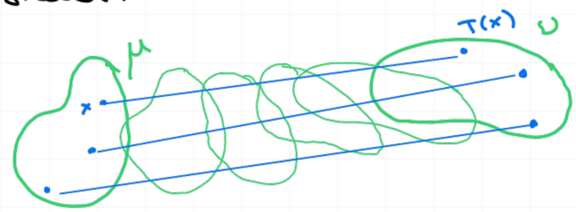
<u>Remark</u> : in the particular case $\gamma = (Id, T)_\# \mu$, it is induced by a map:

$$\mu_t = \pi_t \# (Id, T)_\# \mu = (\pi_t \circ (Id, T))_\# \mu = T_t \# \mu ,$$

$$\pi_t \circ (Id, T)(x) = \pi_t(x, T(x)) = (1-t)x + tT(x) =: T_t(x) .$$

Morally, to find the W_p geodesic we should:

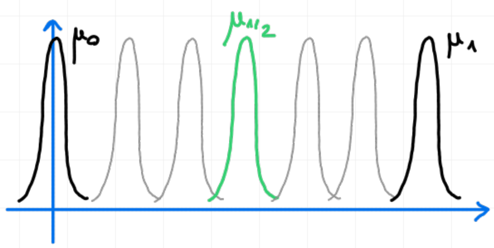
- take optimal map from μ to ν.
- see initial measure as collection of point-like masses.
- let each point x move along the geodesic to its final destination $T(x)$ (at constant speed).



<u>Examples</u>: • Geodesic from δ_x to δ_y ?

• Geodesic from δ_{x_0} to μ ?

•



Instead what would be the mid-point of the geodesic with resp. to $\|\cdot\|_{L^1}$?

<u>Proof of theorem</u>: $s<t$. Define $\gamma_{st} = (\pi_s, \pi_t)_\# \gamma \in \Gamma(\mu_s, \mu_t)$:

$$\pi_1 \# \gamma_{st} = \pi_s \# \gamma = \mu_s .$$

$$W_p(\mu_s, \mu_t) \leq \left(\int_{X\times X} |z - z'|^p \, d\gamma_{st}(z,z') \right)^{1/p} = \left(\int_{X\times X} |\overbrace{\pi_s(x,y) - \pi_t(x,y)}^{(1-s)x + sy - (1-t)x - ty = (s-t)(x-y)}|^p \, d\gamma(x,y) \right)^{1/p}$$

$$= |s-t| \left(\int_{X\times X} |x-y|^p \, d\gamma(x,y) \right)^{1/p} = |s-t| \, W_p(\mu_0, \mu_1) \qquad (*)$$

$$W_p(\mu_0, \mu_1) \leq W_p(\mu_0, \mu_s) + W_p(\mu_s, \mu_t) + W_p(\mu_t, \mu_1) \overset{(*)}{\leq} (s + (t-s) + 1 - t) \, W_p(\mu_0, \mu_1) .$$

$\longrightarrow$ All the inequalities are equalities.

<u>Def</u>: F is <u>geodesically convex</u> if $\forall x, y$, $\exists \eta : [0,1] \to X$ constant speed geodesic from x to y s.t. $F \circ \eta : [0,1] \to \mathbb{R} \cup \{+\infty\}$ is convex, i.e.,

$$F(\eta(t)) \leq (1-t) F(x) + t F(y) .$$