Serie 9

Optimal transport, Fall semester

EPFL, Mathematics section, Dr. Xavier Fernández-Real

Exercise 9.1 (Wasserstein and L^p-distances are not comparable)**.** Let $p \in [1, \infty)$. Give an example of two sequences of compactly supported nonnegative functions $f_n, g_n \in L^p(\mathbb{R}^d)$ with $\int f_n = \int g_n = 1$ for which, calling $\mu_n = f_n \mathscr{L}^d, \nu_n = g_n \mathscr{L}^d$ we have (give an example for each of the two scenarios):

i) $W_p(\mu_n, \nu_n) \to 0$ and $\|f_n - g_n\|_{L^p} \geq \epsilon > 0$.

ii) $W_p(\mu_n, \nu_n) \geq \epsilon > 0$ and $\|f_n - g_n\|_{L^p} \to 0$.

Solution: We work in dimension $d = 1$ for simplicity. Let us treat the two cases i) and ii) separately.

i) For every $n \geq 1$ define

$$f_n = 2\sum_{j=0}^{n-1} \mathbb{1}_{[\frac{2j}{2n}, \frac{2j+1}{2n})}, \qquad g_n = 2\sum_{j=0}^{n-1} \mathbb{1}_{[\frac{2j+1}{2n}, \frac{2j+2}{2n})}.$$

Then clearly $\int_{\mathbb{R}} f_n = \int_{\mathbb{R}} g_n = 1$ and $\|f_n - g_n\|_{L^p} = 2$. However, if $T : \mathbb{R} \to \mathbb{R}$ is a map sending monotonically $[\frac{2j}{2n}, \frac{2j+1}{2n})$ to $[\frac{2j+1}{2n}, \frac{2j+2}{2n})$ for every $j \in \{0, n-1\}$, then:

$$W_p(f_n \mathscr{L}^1, g_n \mathscr{L}^1)^p \leq \int_{\mathbb{R}} f_n |T(x) - x|^p dx \leq \left(\frac{1}{2n}\right)^p \overset{n\to\infty}{\longrightarrow} 0.$$

ii) For every $n \geq 1$ define

$$f_n = \frac{n+1}{2n} \mathbb{1}_{[-n-1,-n]} + \frac{n-1}{2n} \mathbb{1}_{[n,n+1]}, \qquad \frac{n-1}{2n} \mathbb{1}_{[-n-1,-n]} + \frac{n+1}{2n} \mathbb{1}_{[n,n+1]}.$$

Notice that $\int_{\mathbb{R}} f_n = \int_{\mathbb{R}} g_n = 1$ and

$$\int_{\mathbb{R}} |f_n - g_n|^p dx = \frac{2}{n^p} \overset{n\to\infty}{\longrightarrow} 0.$$

Take any transport map T from $f_n \mathscr{L}^1$ to $g_n \mathscr{L}^1$. Define $Z := \{x \in [-n-1, -n] \cup [n, n+1] : |T(x) - x| \geq 2n\}$. By definition of f_n and g_n we see immediately that $\mathscr{L}^1(Z) \geq 2/n$. Hence, for $n \geq 2$:

$$\int_{\mathbb{R}} |T(x) - x|^p f_n dx \geq \frac{n-1}{2n}(2n)^p \mathscr{L}^1(Z) \geq \frac{(2n)^{p-1} 2(n-1)}{n} \geq 1.$$

Therefore, we have $W_p(f_n \mathscr{L}^1, g_n \mathscr{L}^1) \geq 1$.

Exercise 9.2 (Convergence of p-Wasserstein distance as $p \downarrow 1$)**.** Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be a pair of probability measures. Show that if μ and ν are supported on a compact set, then

$$\lim_{p \downarrow 1} W_p(\mu, \nu) = W_1(\mu, \nu)\,.$$

Show a counterexample to the previous statement if we drop the assumption that μ and ν are supported on a compact set.

Hint: For the counterexample, set $\mu = \delta_0$ and find a measure ν such that $W_p(\mu, \nu) = \infty$ if $p > 1$ and $W_1(\mu, \nu)$ is finite.

Solution: If μ, ν are compactly supported, then without loss of generality we can take Ω to be compact. Observe that, on the one hand, using Hölder's inequality we have that for any $\gamma \in \Gamma(\mu, \nu)$, $q \geq p \geq 1$,

$$\left(\int |x-y|^p d\gamma(x,y)\right)^{\frac{1}{p}} \leq \left(\int |x-y|^q d\gamma(x,y)\right)^{\frac{1}{q}}$$

from where we deduce

$$W_p(\mu, \nu) \leq W_q(\mu, \nu).$$

On the other hand, since $|x-y|^p \leq |x-y| \text{diam}(\Omega)^{p-1}$, we have

$$W_p^p(\mu, \nu) \leq \text{diam}(\Omega)^{p-1} W_1(\mu, \nu).$$

By the first inequality, $W_p(\mu, \nu)$ is bounded and monotone increasing in $p \geq 1$. In particular, the limit as $p \downarrow 1$ exists, and by the second inequality it is equal to $W_1(\mu, \nu)$.

For the second part of the question suppose that $\mu = \delta_0$ and $\nu = \sum_{n=1}^{\infty} C 2^{-n} \frac{1}{n^2} \delta_{2^n}$. We take C such that $\sum_{n=1}^{\infty} C 2^{-n} \frac{1}{n^2} = 1$. As a result, for every $1 < p = 1 + \epsilon$

$$W_p(\mu, \nu) = \sum_{n=1}^{\infty} C 2^{n\epsilon} \frac{1}{n^2} = \infty$$

but

$$W_1(\mu, \nu) = \sum_{n=1}^{\infty} C \frac{1}{n^2} < \infty.$$

Exercise 9.3 (Convergence of p-Wasserstein distance as $p \uparrow \infty$)**.** Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be two compactly supported probability measures. The ∞-Wasserstein distance between μ and ν is defined as

$$W_\infty(\mu, \nu) := \inf \left\{ \|x - y\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d, \gamma)} : \gamma \in \Gamma(\mu, \nu) \right\}.$$

i) Prove that $W_p(\mu, \nu) \uparrow W_\infty(\mu, \nu)$ as $p \uparrow \infty$. Deduce that W_∞ defines a distance on $\mathcal{P}_\infty(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \ \mu \text{ has compact support}\}$.

ii) Give an example of $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^d)$ compactly supported in a common compact set for which

$$\begin{cases} W_p(\mu_n, \mu) \to 0 & \text{for every } p \in [1, \infty), \\ W_\infty(\mu_n, \mu) \geq \epsilon > 0 & \text{for every } n. \end{cases}$$

Solution:

i) As already noticed in the solution of the previous exercise, p-Wasserstein distances are monotonically nondecreasing as $p \uparrow \infty$. Therefore, given $\mu, \nu \in \mathcal{P}_\infty(\mathbb{R}^d)$, it only remains to show that $\liminf_{j\to\infty} W_{p_j}(\mu,\nu) \geq W_\infty(\mu,\nu)$ for some sequence $p_j \to \infty$. For this, take optimal plans $\gamma_j \in \Gamma(\mu,\nu)$ for the $|x-y|^{p_j}$-cost. By weak compactness of $\Gamma(\mu,\nu)$, there exist some $\gamma \in \Gamma(\mu,\nu)$ such that, up to extracting a subsequence, $\gamma_j \rightharpoonup \gamma$ narrowly. Call

$$\ell := \|x-y\|_{L^\infty(\mathbb{R}^d\times\mathbb{R}^d,\gamma)} = \sup\{|x-y| : (x,y) \in \text{supp}\gamma\}.$$

Notice that $\ell \geq W_\infty(\mu,\nu)$.

Since μ and ν are compactly supported, $\ell < \infty$, and, being $\text{supp}\gamma$ compact, there exists $(x_0, y_0) \in \text{supp}\gamma$ such that $|x-y| = \ell$.

Fix $\epsilon > 0$ and consider the open set

$$A := \{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d : |(x,y) - (x_0,y_0)| < \epsilon\}.$$

Owing to the fact that (x_0, y_0) is in the support of γ and the weak convergence of γ_j to γ, we get that

$$\liminf_{j\to\infty} \gamma_j(A) \geq \gamma(A) =: \delta > 0.$$

In particular,

$$\begin{aligned} \liminf_{j\to\infty} W_{p_j}(\mu,\nu) &\geq \liminf_{j\to\infty} \left(\int_A |y-x|^{p_j} d\gamma_j \right)^{1/p_j} \\ &\geq \liminf_{j\to\infty} \gamma_j(A)^{1/p_j}(\ell - \epsilon) = (\ell - \epsilon). \end{aligned}$$

The conclusion comes from the arbitrariness of ϵ.

ii) It is enough to take two distinct points $x, y \in \mathbb{R}^d$ and then choose $\mu = \delta_x$ and, for every $n \geq 1$,

$$\mu_n = \frac{n-1}{n}\delta_x + \frac{1}{n}\delta_y.$$

Exercise 9.4 (💣). For every $p \in [1, \infty)$, show that $(\mathcal{P}_p(\mathbb{R}^d), W_p)$ is a Polish space (separable and complete).

Hints: You can solve the exercise via the following steps:

i) For the separability, approximate each $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ with finite sums of dirac deltas in rational points and with rational coefficients.

ii) To prove completeness, take a Cauchy sequence $\{\mu_n\}_{n\geq 1} \subset \mathcal{P}_p(\mathbb{R}^d)$ and argue as follows:

- For every $k \geq 1$ take an optimal $\gamma_k \in \Gamma(\mu_k, \mu_{k+1})$. Use the disintegration Theorem to build a sequence of measures $\pi_n \in \mathcal{P}((\mathbb{R}^d)^n)$ with the following properties:

$$\begin{cases} p^{1,\dots,n-1}_{\#}\pi_n = \pi_{n-1} & \text{for every } n \geq 2, \\ p^{k,k+1}_{\#}\pi_n = \gamma_k & \text{for every } 1 \leq k < n. \end{cases}$$

  Here $p^{i,\dots,j}$ denotes the projection on the variables from i to j.

- Use Kolmogorov's extension Theorem to find $\pi_\infty \in \mathcal{P}((\mathbb{R}^d)^{\mathbb{N}})$ such that

$$p^{1,\dots,n}_{\#}\pi_\infty = \pi_n \quad \text{for every } n \geq 1.$$

- Observe that the L^p-space

$$\mathcal{X} := L^p((\mathbb{R}^d)^{\mathbb{N}}, \pi_\infty)$$

  is complete. Assuming without loss of generality that $\sum_n W_p(\mu_n, \mu_{n+1}) < \infty$, prove that the coordinate functions $p^n : (\mathbb{R}^d)^{\mathbb{N}} \to \mathbb{R}$ form a Cauchy sequence in $\mathcal{X}$, and deduce that $p^n \to \bar{p}$ in $\mathcal{X}$.

- Conclude that $\mu_n \to \bar{\mu}$ in $(\mathcal{P}_p(\mathbb{R}^d), W_p)$, where $\bar{\mu} := \bar{p}_{\#}\pi_\infty$.