(Rough) Introduction to properties of $\mathbb{W}_p$

—

May 25, 2020

Abstract

These notes are an abridged version of Chapter 5 of "Optimal Transport for Applied Mathematicians", by Filippo Santambrogio, for the Summer 2020 reading group on Topics in Optimal Transport. This is meant to quickly outline some of the fundamental definitions and notions that are necessary for reading papers in the general area.

Distance and triangle inequality

Let $\Omega \subseteq \mathbb{R}^d$ and define the set of probability measures over Ω be denoted by $\mathcal{P}(\Omega)$. We restrict our understanding to the set of probability measures with finite p moments, denoted by the set

$$\mathcal{P}_p(\Omega) = \left\{ \mu \in \mathcal{P}(\Omega) \;\middle|\; \int |x|^p \, d\mu < +\infty \right\}.$$

Note that if Ω is bounded, then $\mathcal{P}(\Omega) = \mathcal{P}_p(\Omega)$.

For two measures $\mu, \nu \in \mathcal{P}_p(\Omega)$, we define

$$W_p(\mu, \nu) := \min_{\gamma \in \Pi(\mu,\nu)} \left\{ \int_{\Omega \times \Omega} |x - y|^p \, d\gamma \right\}^{1/p}.$$

Recall that $\Pi(\mu, \nu)$ is the set of joint probability measures with marginals μ and ν. Since the measures have p finite moments, this gives us finite-ness of $W_p(\mu, \nu)$.

Proposition 1. *The quantity W_p is a distance over $\mathcal{P}_p(\Omega)$ and satisfies the triangle inequality.*

With this, we can define a new space, and hence we can work with new notions of topologies. The following definition can be generalized to any complete separable metric (Polish) space.

Definition 2. For each $p \in [1, +\infty)$ we define the Wasserstein space of order p, $\mathbb{W}_p(\Omega)$, as the space $\mathcal{P}_p(\Omega)$ endowed with the distance W_p.

Topology induced by W_p

In this section, we formalize notions of convergence in $\mathbb{W}_p$. First, we recall the notion of *"weak convergence"*, which denotes convergence in the duality with C_b (space of bounded continuous functions).

Definition 3. **(“Weak convergence”)** For a sequence $(\mu_n) \subseteq \mathcal{P}_p(\Omega)$, we say that $\mu_n \rightharpoonup \mu$ if and only if for every $\varphi \in C_b(\Omega)$, we have

$$\int \varphi \, d\mu_n \to \int \varphi \, d\mu.$$

Note that if Ω is compact, then the spaces $C_b(\Omega), C_0(\Omega), C(\Omega)$ and $C_c(\Omega)$ coincide. For the remainder of this section, we tie together statements about weak convergence and convergence in W_p.

Theorem 4. *If $\Omega \subseteq \mathbb{R}^d$ is compact and $p \in [1, +\infty)$, in the space $\mathbb{W}_p(\Omega)$, we have $\mu_n \rightharpoonup \mu$ if and only if $W_p(\mu_n, \mu) \to 0$.*

Theorem 5. *In the space $\mathbb{W}_p(\mathbb{R}^d)$, we have $W_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \rightharpoonup \mu$ and $\int |x|^p \, d\mu_n \to \int |x|^p \, d\mu$.*

Curves in $\mathbb{W}_p$

A *curve* is a continuous function $\omega : [0, 1] \to X$, where (X, d) is a metric space. If it exists, we define the *metric derivative* of ω at time t, denoted by $|\omega'|(t)$ through

$$|\omega'|(t) := \lim_{h \to 0} \frac{d(\omega(t+h), \omega(t))}{|h|}.$$

Theorem 6. *Suppose $\omega \in C([0,1], X)$ is Lipschitz continuous. Then the metric derivative $|\omega'|(t)$ exists for almost every $t \in [0,1]$. Morever, for $t < s$, we have*

$$d(\omega(t), \omega(s)) \leq \int_t^s |\omega'|(\tau) \, d\tau.$$

A more general notion of (continuous) curves are *absolutely continuous* curves, denoted by $\mathrm{AC}(X)$. A curve $\omega \in C([0,1], X)$ is said to be absolutely continuous if, for every $t_0 < t_1$, there exists a $g \in L^1([0,1])$ such that

$$d(\omega(t_0), \omega(t_1)) \leq \int_{t_0}^{t_1} g(s) \, ds.$$

Connection with the continuity equation

We henceforth assume Ω is compact. We denote a vector field (parameterized by t) as $\boldsymbol{v}_t$, and we assume they lie in L^p. The goal here is to link curves in $\mathrm{AC}(\mathbb{W}_p(\Omega))$ and the *continuity equation*, defined as

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0. \tag{1}$$

One can view the divergence term as excess of mass which is injected into the motion at every point. Some intuition on the continuity equation: suppose a family of particles is initialized with respect to some density μ_0, with each particle following the ODE

$$\begin{cases} y_x'(t) = \boldsymbol{v}_t(y_x(t)) \\ y_x(0) = x, \end{cases}$$

with $y_x(t)$ being the position of the particle x at time t. Defining $Y_t(x) = y_x(t)$ and $\mu_t := (Y_t)_{\#}(\mu_0)$. Then the pair $(\mu_t, \boldsymbol{v}_t)$ satisfy (1) *in the weak sense*: for all $\varphi \in C_c^1((0,1), \overline{\Omega})$, we have

$$\int_0^1 \int_\Omega (\partial_t \varphi) \; d\mu_t dt + \int_0^1 \int_\Omega \nabla\varphi \cdot \boldsymbol{v}_t \; d\mu_t dt = 0.$$

We are now ready to state the main result of this section

Theorem 7. *Let $(\mu_t)_{t\in[0,1]}$ be an absolutely continuous curve in $\mathbb{W}_p(\Omega)$ with $p > 1$. Then for a.e. $t \in [0,1]$ there exists a vector field $\boldsymbol{v}_t \in L^p(\mu_t; \mathbb{R}^d)$ such that*

- *$(\mu_t, \boldsymbol{v}_t)$ satify (1) in the weak sense,*
- *for a.e. t, we have $\|\boldsymbol{v}_t\|_{L^p((\mu_t)} \leq |\mu'|(t)$, the metric derivative at time t of the curve $t \mapsto \mu_t$ with respect to W_p.*

Conversely, if $(\mu_t)_{t\in[0,1]}$ is a family of measures in $\mathcal{P}_p(\Omega)$ and for each t we have a vector field $\boldsymbol{v}_t \in L^p(\mu_t; \mathbb{R}^d)$ with $\int_0^1 \|\boldsymbol{v}_t\|_{L^p(\mu_t)} < +\infty$ solving the continuity equation, then $(\mu_t)_t$ is absolutely continuous in $\mathbb{W}_p(\Omega)$ and, for a.e. t, we have $|\mu'|(t) \leq \|\boldsymbol{v}_t\|_{L^p(\mu_t)}$.

Constant-speed geodesics in $\mathbb{W}_p$

This section is mostly definitions. Let ω be some general curve in a metric space (X, d). The *length* of a curve is formally defined as

$$\text{Length}(\omega) := \sup \left\{ \sum_{k=0}^{n-1} d(\omega(t_k), \omega(t_{k+1})) \;\middle|\; n \geq 1, 0 = t_0 < t_1 < \cdots < t_n = 1 \right\}.$$

Proposition 8. *For any curve $\omega \in AC(X)$, we have*

$$\textit{Length}(\omega) = \int_0^1 |\omega'|(t) \; dt.$$

A curve $\omega : [0,1] \to X$ is a *geodesic* between x_0 and $x_1 \in X$ if it minimizes the length among all curves such that $\omega(0) = x_0$ and $\omega(1) = x_1$. A space (X, d) is a *length space* if

$$d(x, y) = \inf\{\text{Length}(\omega) \mid \omega \in \text{AC}(X), \omega(0) = x, \omega(1) = y\}.$$

A space (X, d) is a *geodesic space* if there exists a geodesic that attains the infimum above. In a length space, a curve is a *constant-speed geodesic* between $\omega(0)$ and $\omega(1) \in X$ if it satisfies

$$d(\omega(t), \omega(s)) = |t - s| d(\omega(0), \omega(1)), \quad \forall t, s \in [0,1].$$

Equivalently, ω is a constant-speed geodesic if (and only if) $\omega \in \text{AC}(X)$ and $|\omega'|(t) = d(\omega(0), \omega(1))$ (constant speed modulus) a.e.

Theorem 9. *Suppose Ω is convex. Let $\mu, \nu \in \mathcal{P}_p(\Omega)$ and let $\gamma^* \in \Pi(\mu, \nu)$ be an optimal transport plan for $W_p(\mu, \nu)$. Define $\pi_t : \Omega \times \Omega \to \Omega$ through $\pi_t(x, y) = (1-t)x + ty$. Then the curve $\mu_t := (\pi_t)_{\#}\gamma$ is a constant speed geodesic in $\mathbb{W}_p$ connecting $\mu_0 = \mu$ to $\mu_1 = \nu$. Thus, $\mathbb{W}_p(\Omega)$ is a geodesic space.*

Remark 10. If $\gamma = \gamma_T$ (i.e. the plan is given by an optimal transport map), then the curve in the aforementioned theorem is given by $((1-t)Id + tT)_{\#}\mu$.

Proposition 11. *Suppose γ is induced by an optimal transport map T; then then particles initialized at x move on a straight line with constant speed $T(x) - x$. Define $T_t := (1-t)Id + tT$ (it is invertible), the velocity field $\boldsymbol{v}_t(y) = (T - I)(T_t^{-1}(y))$ and the geodesic $\mu_t = (\pi_t)_{\#}\gamma$. Then $(\mu_t, \boldsymbol{v}_t)$ satisfy the continuity equation with $\|\boldsymbol{v}_t\|_{L^p(\mu_t)} = |\mu'|(t) = W_p(\mu, \nu)$.*