Serie 10

Optimal transport, Fall semester

EPFL, Mathematics section, Dr. Xavier Fernández-Real

Definition 1 (Middle point). Given two probability measures $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$, let $\mathcal{C}(\mu_0, \mu_1)$ be the infimum of the Kantorovich problem with respect to the quadratic cost

$$\mathcal{C}(\mu_0, \mu_1) := \inf_{\gamma \in \Gamma(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|x-y|^2}{2} \, d\gamma(x, y).$$

Let $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$ be two probability measures with compact support. A probability measure $\mu_{\frac{1}{2}}$ is a middle point of μ_0 and μ_1 if $\mathcal{C}(\mu_0, \mu_{\frac{1}{2}}) = \mathcal{C}(\mu_1, \mu_{\frac{1}{2}}) = \frac{1}{4}\mathcal{C}(\mu_0, \mu_1)$.

Exercise 10.1. For any $\rho \in \mathcal{P}(\mathbb{R}^d)$, it holds

$$\mathcal{C}(\mu_0, \rho) + \mathcal{C}(\rho, \mu_1) \geq \frac{1}{2}\mathcal{C}(\mu_0, \mu_1).$$

Moreover, if equality holds, then there is an optimal plan $\gamma \in \Gamma_{opt}(\mu_0, \mu_1)$ such that $(\frac{x+z}{2})_\# \gamma = \rho$ (here, x, z denote the first and second coordinate of $\mathbb{R}^d \times \mathbb{R}^d$).

Hint: See the proof of Theorem 3.1.5. Try to prove the inequality without directly using the triangular inequality for the Wasserstein distance.

Solution: Let $\gamma_0 \in \Gamma(\mu_0, \rho)$ and $\gamma_1 \in \Gamma(\rho, \mu_1)$ be two optimal plans from μ_0 to ρ and from ρ to μ_1, respectively. The gluing lemma (see the proof of Theorem 3.1.5) ensures the existence of $\tilde{\gamma} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ such that (here, x, y, z denote the coordinates of $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$)

$$(x, y)_\# \tilde{\gamma} = \gamma_0 \quad \text{and} \quad (y, z)_\# \tilde{\gamma} = \gamma_1.$$

Let $\gamma := (x, z)_\# \tilde{\gamma}$. It follows directly from the properties of $\tilde{\gamma}$ that γ is an admissible plan from μ_0 to μ_1. Therefore it holds

$$\begin{aligned}
\mathcal{C}(\mu_0, \mu_1) &\leq \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - z|^2 \, d\gamma(x, z) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} |x - z|^2 \, d\tilde{\gamma}(x, y, z) \\
&\leq \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \left(2|x - y|^2 + 2|y - z|^2\right) d\tilde{\gamma}(x, y, z) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, d\gamma_0(x, y) + \int_{\mathbb{R}^d \times \mathbb{R}^d} |y - z|^2 \, d\gamma_1(y, z) = 2\left(\mathcal{C}(\mu_0, \rho) + \mathcal{C}(\rho, \mu_1)\right).
\end{aligned}$$

If equality holds, then all the inequalities we have applied must be equalities. Hence (consider the first inequality of the chain) γ has to be an optimal plan and (consider the second inequality) $|x - z|^2 = 2|x - y|^2 + 2|y - z|^2$ has to be true $\tilde{\gamma}$-almost everywhere. The latter identity implies that $\tilde{\gamma}$-almost everywhere it holds $y = \frac{x+z}{2}$, thus

$$\rho = y_\# \tilde{\gamma} = (\tfrac{x+z}{2})_\# \tilde{\gamma} = (\tfrac{x+z}{2})_\# \gamma,$$

as desired.

Exercise 10.2. Thanks to the previous Exercise, show that a measure $\mu_{\frac{1}{2}}$ is a middle-point if and only if

$$\mu_{\frac{1}{2}} \text{ is a middle-point} \qquad \Leftrightarrow \qquad \mathcal{C}(\mu_0,\mu_{\frac{1}{2}}) \le \frac{1}{4}\mathcal{C}(\mu_0,\mu_1) \quad \text{and} \quad \mathcal{C}(\mu_1,\mu_{\frac{1}{2}}) \le \frac{1}{4}\mathcal{C}(\mu_0,\mu_1). \qquad (1)$$

Solution: Indeed these two inequalities, together with $\mathcal{C}(\mu_0,\mu_{\frac{1}{2}}) + \mathcal{C}(\mu_{\frac{1}{2}},\mu_1) \ge \frac{1}{2}\mathcal{C}(\mu_0,\mu_1)$ (this inequality follows from the triangle inequality for W_2, since $\mathcal{C} = \frac{1}{2}W_2^2$), imply that $\mathcal{C}(\mu_0,\mu_{\frac{1}{2}}) = \mathcal{C}(\mu_1,\mu_{\frac{1}{2}}) = \frac{1}{4}\mathcal{C}(\mu_0,\mu_1)$.

Exercise 10.3. Let $\mu_0,\mu_1 \in \mathcal{P}(\mathbb{R}^d)$ be two probability measures with compact support. A probability measure $\mu_{\frac{1}{2}}$ is a middle point of μ_0 and μ_1 if $\mathcal{C}(\mu_0,\mu_{\frac{1}{2}}) = \mathcal{C}(\mu_1,\mu_{\frac{1}{2}}) = \frac{1}{4}\mathcal{C}(\mu_0,\mu_1)$.

(i) If $\mu_0 = \delta_{p_0}$ and $\mu_1 = \delta_{p_1}$, show that the middle point is unique and $\mu_{\frac{1}{2}} = \delta_{\frac{p_1+p_2}{2}}$.

(ii) Prove that there is always at least one middle point.

(iii) Find two probability measures μ_0,μ_1 such that they have more than one middle point.

(iv) Show that if the optimal transport plan between μ_0 and μ_1 is unique, then there is a unique middle point.

(v) Prove that if $\mu_0,\mu_1 \ll \mathcal{L}^d$, then the middle point is unique and $\mu_{\frac{1}{2}} \ll \mathcal{L}^d$.

Solution:

Let us consider an optimal plan $\gamma \in \Gamma(\mu_0,\mu_1)$. We claim that $\mu_{\frac{1}{2}} := (\frac{x+z}{2})_\#\gamma$ is a middle-point. Indeed, since $(x, \frac{x+z}{2})_\#\gamma$ is an admissible plan from μ_0 to $\mu_{\frac{1}{2}}$, it holds

$$\mathcal{C}(\mu_0,\mu_{\frac{1}{2}}) \le \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d} \left| x - \frac{x+z}{2}\right|^2 d\gamma(x,z) = \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{4}|x-z|^2\, d\gamma(x,z) = \frac{1}{4}\mathcal{C}(\mu_0,\mu_1).$$

The same reasoning yields also $\mathcal{C}(\mu_1,\mu_{\frac{1}{2}}) \le \frac{1}{4}\mathcal{C}(\mu_0,\mu_1)$ and therefore, thanks to (1), $\mu_{\frac{1}{2}}$ is a middle-point.

Hence, given an optimal plan γ we can produce a middle point via the formula $(\frac{x+z}{2})_\#\gamma$. Vice versa, thanks to the lemma above, given a middle point $\mu_{\frac{1}{2}}$ there is an optimal plan γ such that $(\frac{x+z}{2})_\#\gamma = \mu_{\frac{1}{2}}$. As an observation, one may be tempted to deduce from these observations that the map between optimal plans and middle points is an isomorphism. In order to show it, one should check that if $\gamma,\gamma' \in \Gamma(\mu_0,\mu_1)$ are optimal and such that $(\frac{x+z}{2})_\#\gamma = (\frac{x+z}{2})_\#\gamma'$, then $\gamma = \gamma'$. Such a statement is true but not straightforward, and we shall not prove it here.

Now we are ready to tackle the statements of the exercise.

(i) Since there is a unique optimal plan (that is $\delta_{p_0} \times \delta_{p_1}$) there can be only one middle point and it must be $(\frac{x+z}{2})_\#(\delta_{p_0} \times \delta_{p_1}) = \delta_{\frac{p_0+p_1}{2}}$.

(ii) The existence of a middle point follows directly from the existence of an optimal plan.

(iii) Consider the two probability measures constructed in the solution of Exercise 1.4(b). Since every plan induces a middle point as explained above, one can check that the two mentioned probability measures are a good example.

(iv) Let $\gamma \in \Gamma(\mu_0, \mu_1)$ be the unique optimal coupling. Then, thanks to the observations above, $\mu_{\frac{1}{2}} = (\frac{x+z}{2})_\#\gamma$ has to be the unique middle point.

(v) Theorem 2.5.9 asserts that there is a unique optimal map between μ_0 and μ_1, that we denote $T : \mathbb{R}^d \to \mathbb{R}^d$. Thus, our observations imply that $\mu_{\frac{1}{2}} := (\frac{x+T(x)}{2})_\#\mu_0$ is the unique middle point. Also, again by Theorem 2.5.9, it holds $T = \nabla\varphi$ where $\varphi : \mathbb{R}^d \to \mathbb{R}$ is a convex function. Hence

$$\frac{x + T(x)}{2} = \frac{1}{2}\nabla\left(\frac{|x|^2}{2} + \varphi\right).$$

We now note that, if $X := \frac{\mathrm{id}+T}{2}$, then X^{-1} is 2-Lipschitz. Indeed, observe that for any $x, x' \in \mathbb{R}^d$,

$$|X(x) - X(x')||x - x'| \geq \langle X(x) - X(x'), x - x'\rangle = \frac{1}{2}|x - x'|^2 + \frac{1}{2}\langle\nabla\varphi(x) - \nabla\varphi(x'), x - x'\rangle.$$

Now, observe that the last term is nonnegative, since φ is convex, so that we have

$$|X(x) - X(x')| \geq \frac{1}{2}|x - x'|,$$

from where we deduce X^{-1} is 2-Lipschitz.

In particular, for any Borel set $E \subset \mathbb{R}^d$ we have

$$|X^{-1}(E)| \leq \int_E |\det(\nabla X^{-1})|(y)\, dy \leq 2^d|E|,$$

hence $\left(\frac{x+T(x)}{2}\right)_\# dx \ll dx$, and we conclude that

$$\mu_{\frac{1}{2}} = \left(\frac{\mathrm{id} + T}{2}\right)_\# \mu_0 \ll \left(\frac{\mathrm{id} + T}{2}\right)_\# dx \ll dx.$$