Serie 6
Optimal transport, Fall semester
EPFL, Mathematics section, Dr. Xavier Fernández-Real

Exercise 6.1 (Translations are optimal). Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be the translation map $T(x) := x + x_0$, where $x_0 \in \mathbb{R}^d$. For any probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, show that T is an optimal transport map from μ to $T_\#\mu$ with respect to the quadratic cost.

Solution: We provide two solutions.

(i) We just need to show that T is the gradient of a convex function (see Corollary 2.5.10). Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be the convex function $\varphi(x) := \frac{1}{2}|x + x_0|^2$. Since $T = \nabla\varphi$, the optimality of T follows.

(ii) By the general case of Kantorovich duality (Theorem 2.6.6) we know that for any ϕ and ψ continuous and bounded such that $\phi(x) + \psi(y) \le c(x, y)$, we have

$$\int_X \phi d\mu + \int_Y \psi d\nu \le \min_{\gamma\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x, y)d\gamma(x, y)$$

where X and Y need not be compact and $\nu = T_\#\mu$.

Consider $\phi(x) = -\langle x, x_0\rangle$ and $\psi(y) = \langle y - x_0, x_0\rangle + \frac{1}{2}||x_0||^2$. Then

$$\begin{aligned}\phi(x) + \psi(y) &= -\langle x, x_0\rangle + \langle y - x_0, x_0\rangle + \frac{1}{2}\|x_0\|^2 = \langle y - x, x_0\rangle - \frac{1}{2}\|x_0\|^2 \\ &\le \|y - x\|\|x_0\| - \frac{1}{2}\|x_0\|^2 \le \frac{1}{2}\|y - x\|^2 = c(x, y).\end{aligned}$$

Hence, recalling $\nu = T_\#\mu$ with $T(x) = x + x_0$

$$\begin{aligned}\min_{\gamma\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x, y)d\gamma(x, y) &\ge \int_{\mathbb{R}^n} -\langle x, x_0\rangle d\mu(x) + \int_{\mathbb{R}^n} \left(\langle y - x_0, x_0\rangle + \frac{1}{2}\|x_0\|^2\right) d\nu(y) \\ &= -\int_{\mathbb{R}^n} \langle x, x_0\rangle d\mu(x) + \int_{\mathbb{R}^n} \left(\langle x, x_0\rangle + \frac{1}{2}\|x_0\|^2\right) d\mu(x) \\ &= \int_{\mathbb{R}^n} \frac{1}{2}\|x_0\|^2 d\mu(x) = \int_{\mathbb{R}^n} c(x, T(x))d\mu(x)\end{aligned}$$

which means that T is optimal.

Exercise 6.2 (Homoteties are optimal). Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be the homotety map $T(x) := \lambda x$, where $\lambda > 0$. For any compactly supported probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, show that T is an optimal transport map from μ to $T_\#\mu$ with respect to the quadratic cost.

Solution: We provide two solutions.

(i) It is sufficient to show that the homothety T is the gradient of a convex function (see Corollary 2.5.10). Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be the convex function $\varphi(x) := \frac{\lambda}{2}|x|^2$. Since $T = \nabla\varphi$, the optimality of T follows.

(ii) We consider the potential $\phi(x) = \frac{1}{2}(1-\lambda)\|x\|^2$ together with its c-transform ϕ^c. Since μ is compactly supported, we can work on a compact domain $\Omega = \max(1,\lambda)\text{spt}(\mu) \subset \mathbb{R}^n$, so that ϕ is bounded in Ω. We compute

$$\phi^c(y) = \inf_{x\in\mathbb{R}^n} c(x,y) - \phi(x) = \inf_{x\in\mathbb{R}^n} \frac{1}{2}\|x-y\|^2 - \frac{1}{2}(1-\lambda)\|x\|^2 \geq \inf_{x\in\mathbb{R}^n} \frac{1}{2}(\|x\|-\|y\|)^2 - \frac{1}{2}(1-\lambda)\|x\|^2$$

The previous is a quadratic expression and the infimum is achieved when $\|x\| = \frac{1}{\lambda}\|y\|$. In particular, we find

$$\phi^c(y) = -\frac{(1-\lambda)}{2\lambda}\|y\|^2.$$

The definition of c-transform implies $\phi(x) + \phi^c(y) \leq c(x,y)$. Thus, if we denote $\nu = T_{\#}\mu$ with $T(x) = \lambda x$,

$$\begin{aligned}\min_{\gamma\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x,y)d\gamma(x,y) &\geq \int_\Omega \phi(x)d\mu(x) + \int_\Omega \phi^c(x)d\nu(x) \\ &= \int_\Omega \frac{1}{2}(1-\lambda)\|x\|^2 d\mu(x) - \int_\Omega \frac{\lambda(1-\lambda)}{2}\|x\|^2 d\mu(x) \\ &= \int_\Omega \frac{(1-\lambda)^2}{2}\|x\|^2 d\mu(x) = \int_{\mathbb{R}^n} c(x,T(x))d\mu(x),\end{aligned}$$

which means that T is optimal.

Exercise 6.3. Let $S : \mathbb{R}^d \to \mathbb{R}^d$ be the function $S(x) := -x$. Characterize the probability measures $\mu \in \mathcal{P}(\mathbb{R}^d)$ with compact support such that S is an optimal transport map between μ and $S_{\#}\mu$ with respect to the quadratic cost.

Solution: Assume that S is optimal from μ to $S_{\#}\mu$, and let $\gamma := (\text{id} \times S)_{\#}\mu$ be the associated coupling. Since γ is optimal, Corollary 2.5.7 implies the existence of a cyclically monotone set $A \subset \mathbb{R}^d \times \mathbb{R}^d$ such that γ is supported on A. Since γ is also supported on graph(S), we can assume without loss of generality that $A \subset \text{graph}(S)$.

Take two points $x, y \in \mathbb{R}^d$ such that $(x, S(x)), (y, S(y)) \in A$. By cyclical monotonicity, it holds

$$\frac{1}{2}|x - S(x)|^2 + \frac{1}{2}|y - S(y)|^2 \leq \frac{1}{2}|x - S(y)|^2 + \frac{1}{2}|y - S(x)|^2.$$

Developing the squares and rearranging terms, this is equivalent $|x-y|^2 \leq 0$, thus $x = y$. Hence, this implies that A contains only one point $(x_0, S(x_0))$, and therefore $\mu = \delta_{x_0}$.

On the other hand, if μ is δ_{x_0} for some $x_0 \in \mathbb{R}^d$, then S is optimal from μ to $S_{\#}\mu = \delta_{S(x_0)}$ (since there is only one transport map).

Exercise 6.4. Let B_1 be the unit ball in $\mathbb{R}^2$. Write it as $B_1^+ \cup B_1^-$, where $B_1^+ = B_1 \cap \{(x,y) : x \geq 0\}$ and $B_1^- = B_1 \cap \{(x,y) : x < 0\}$. Find a convex function $\varphi : B_1 \to \mathbb{R}$ such that $\nabla\varphi$ is the optimal transport between $\mathbb{1}_{B_1}\mathscr{L}^2$ and $\big(\mathbb{1}_{(-1,0)+B_1^-} + \mathbb{1}_{(1,0)+B_1^+}\big)\mathscr{L}^2$ with the quadratic cost.

- *Observation:* Recall that $\mathbb{1}_A$ denotes the indicator function of a set $A \subset \mathbb{R}^2$.
- (💣) *Optional question (**not required for the hand-in serie**)*: For every small $\epsilon > 0$, define the strip

$$S_\epsilon := [-1,1] \times [-\epsilon, \epsilon] \subset \mathbb{R}^2.$$

Let $\mu = \mathbb{1}_{B_1}\mathscr{L}^2$, and consider the measures

$$\nu_\epsilon := \frac{1}{1+4\epsilon}\big(\mathbb{1}_{(-1,0)+B_1^-} + \mathbb{1}_{(1,0)+B_1^+}\big)\mathscr{L}^2 + \frac{1}{1+4\epsilon}\mathbb{1}_{S_\epsilon}\mathscr{L}^2.$$

Let $T_\epsilon : \mathbb{R}^2 \to \mathbb{R}^2$ be the optimal transport map from μ to ν_ϵ with respect to the quadratic cost. Prove that for ϵ sufficiently small, T_ϵ is discontinuous in the interior of B_1.

Solution:

We may conjecture that the optimal transport moves each point horizontally by +1 or -1 depending on which side of B_1 they are. Namely, we claim that the optimal transport map $T = \nabla\varphi$ is given by $\nabla\varphi(x,y) = (\operatorname{sign} x + x, y)$. Observe that this map is transporting the right measures, since

$$\begin{aligned}(\nabla\varphi(x,y))_\# \mathbb{1}_{B_1}\mathscr{L}^2 &= ((x,y) \to (1+x,y))_\# \mathbb{1}_{B_1^+}\mathscr{L}^2 + ((x,y) \to (-1+x,y))_\# \mathbb{1}_{B_1^-}\mathscr{L}^2 \\ &= \big(\mathbb{1}_{(-1,0)+B_1^-} + \mathbb{1}_{(1,0)+B_1^+}\big)\mathscr{L}^2.\end{aligned}$$

From here, we observe that the corresponding φ then could be given by (integrating in x and y the expression of $\nabla\varphi$)

$$\varphi(x,y) = |x| + \frac{1}{2}(x^2 + y^2)$$

which is a convex function.

By Theorem 2.5.9 and Corollary 2.5.10, every gradient of a convex function is the unique optimal map between its own marginals in the setting of this problem, so φ is the unique optimal transport map that we are looking for.