In this set of exercises, an *optimal transport map* should be understood as an optimal transport plan that is also a map (so, it must be optimal among all possible transport plans). The linear cost is $c(x, y) = |x - y|$, whereas the quadratic cost is $c(x, y) = \frac{1}{2}|x - y|^2$.

Exercise 1.1 (Translations are optimal)**.** Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be the translation map $T(x) \coloneqq x + x_0$. 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 quadratic cost.

Exercise 1.2 (Homotheties are optimal)**.** Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be the homothety $T(x) \coloneqq \lambda x$ 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.

Exercise 1.3. Let $\mu \coloneqq \frac{1}{\pi}\chi_{B(0,1)}\mathcal{L}^2$ be the uniform probability measure on $B(0, 1) \subset \mathbb{R}^2$ and let $p_0 \coloneqq (1, 0)$, $p_1 \coloneqq (2, 0)$ be two fixed points in $\mathbb{R}^2$. Describe the optimal transport map between μ and $\frac{1}{2}(\delta_{p_0} + \delta_{p_1})$ in the following two cases:

(a) when the cost is the quadratic cost;

(b) when the cost is the linear cost.

Exercise 1.4 (Counterexamples)**.** For any of the following statements, find two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with compact support such that the statement holds (you can choose also the dimension $d \in \mathbb{N}$). Each of the statements should be treated independently.

(a) There is more than one[1] optimal transport map from μ to ν with respect to the linear cost.

(b) There is more than one optimal transport map from μ to ν with respect to the quadratic cost.

(c) There is not an optimal transport plan between μ and ν with respect to the cost $c(x, y) = \lfloor |x - y| \rfloor$ (the floor function[2] of the distance).

(d) There is an optimal transport map from μ to ν with respect to the linear cost, but there is none with respect to the quadratic cost.

(e) There is an optimal transport map from μ to ν with respect to the quadratic cost, but there is none with respect to the linear cost.

Hint: To solve (c), show that the infimum of the Kantorovich problem for $\mu = \chi_{[0,1]}\mathcal{L}^1$, $\nu = \chi_{[1,2]}\mathcal{L}^1$ is 0 but any transport plan has *positive* cost.

To solve (e), it might be useful to first solve Exercise 1.3.

[1]Uniqueness should be understood in the μ-almost everywhere sense.
[2]Given $t \geq 0$, $\lfloor t \rfloor$ is the largest integer n such that $n \leq t$.

Exercise 1.5 (Birkhoff-von Neumann theorem)**.** An $n \times n$ matrix $A \in \mathscr{M}(n, \mathbb{R})$ with nonnegative entries is:

- *a doubly-stochastic matrix* if $\sum_{i=1}^n A_{ij} = 1$ for any $j = 1, \dots, n$ and $\sum_{j=1}^n A_{ij} = 1$ for any $i = 1, \dots, n$.
- *a permutation matrix* if there is a permutation $\sigma \in S^n$ such that $A_{i\sigma(i)} = 1$ and $A_{ij} = 0$ if $j \neq \sigma(i)$.

Prove that any doubly-stochastic matrix can be written as a finite convex combination of permutation matrices.
Hint: Use Hall's marriage theorem to show that for any doubly stochastic matrix A, there is a permutation matrix P and a number $\varepsilon > 0$ such that $A_{ij} \geq \varepsilon P_{ij}$ for any $1 \leq i, j \leq n$.

Exercise 1.6 (Discrete optimal transport)**.** Given two families $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$ of points in $\mathbb{R}^d$, let $\mu := \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$ and $\nu := \frac{1}{n} \sum_{i=1}^n \delta_{y_i}$. Prove that, for *any* choice of a finite cost $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, there exists an optimal transport map from μ to ν.
Hint: Use Exercise 1.5.

Exercise 1.7 (Prokhorov's theorem)**.** Let X be a complete and separable metric space, a family of probability measures $\mathcal{F} \subset \mathcal{P}(X)$ is sequentially relatively compact with respect to the narrow topology if $\mathcal{F}$ is equi-tight, i.e. for every $\varepsilon > 0$ there exists $K \subset X$ compact such that $\mu(X \setminus K) < \varepsilon$ for all $\mu \in \mathcal{F}$.
Hint:

1. Prove the result in the case of X compact using Banach-Alaoglu-Bourbaki theorem.
2. Deduce the result in its full generality taking an exhaustion of X by compact sets.

Remark. Actually Prokhorov's theorem states the "if and only if", namely $\mathcal{F}$ is sequentially relatively compact with respect to the narrow topology *if and only if* $\mathcal{F}$ is equi-tight. However, for our aims, we are interested only in the aforementioned implication.