EXERCISES 4

The assigment problem and discrete optimal transport. The classical *assignment problem* (also known as the *bivariate perfect matching problem* in graph theory) is the problem to, given an N times N matrix (c_{ij}) minimize the functional

$$\sigma \mapsto \sum_{i=1}^{N} c_{i\sigma(i)} \tag{0.1}$$

In economical terms we have N workers and N jobs to conduct and c_{ij} is the cost of assigning work j to a worker i. The problem is to minimize the total cost, if all the every workers are assigned different jobs, i.e. worker i is assigned the job j where $j = \sigma(i)$ for some permutation $\sigma \in S_N$.

The goal of this "guided exercise" is to show that this problem can be formulated in terms of optimal transport. Indeed, consider the general setting of a cost function $c(x, y)$ defined on $X \times Y$. Set $\mu := \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$ for a given configuration $(x_1, ..., x_N)$ of (say different) N points on X and similarly $\nu := \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i}$. Now, any given permutation $\sigma \in S_N$ determines a transport plan γ_σ from μ to ν defined by

$$\gamma_\sigma := \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i} \otimes \delta_{y_{\sigma(i)}}$$

[check this!] whose cost is given by

$$C(\frac{1}{N}\sum_{i=1}^{N} \delta_{x_i} \otimes \delta_{p_{\sigma(i)}}) = \frac{1}{N}\sum_{i=1}^{N} c(x_i, y_{\sigma(i)})$$

[check this!]. The non-trivial fact in this story is that the infimum over the cost of all transport plans between the discrete measures μ and ν is equal to the infimum over all *discrete* transport plans, i.e. those of the form γ_σ, i.e.

$$\inf_{\gamma} C(\gamma) = \inf_{\sigma \in S_N} C(\gamma_\sigma). \tag{0.2}$$

(the point is that we do not have to assume that X and Y are discrete!). To see how this comes about first show that any transport plan γ between μ and ν is supported on the discrete set obtained as the product of the support $\{x_1, ...x_N\}$ of μ and the support $\{y_1, ...y_N\}$ of ν, i.e.

$$\gamma := \sum_{i=1}^{N} A_{ij}\delta_{x_i} \otimes \delta_{y_i}$$

for some matrix (A_{ij}). Next, show that the transport condition corresponds to the matrix A beeing a doubly stochastic matrix, i.e.

$$\sum_i A_{ij} = \sum_j A_{ij} = 1$$

Denoting by B_N the convex set of all such matrices the optimal transport problem may thus be recasted as an optimization problem on B_N. Now, using Choquet's theorem, deduce that the optimum is realized at an extremal point of the set B_N. Finally, invoking the Birkhoff-von Neumann theorem saying that the extremal points of B_N are permutation matrices, conclude that 0.2 indeed holds (compare the warm-up exercises 3-5 on p. 14-15 in Villani's book).

The Legendre transform formalism for non-quadratic costs. For a given cost functional $c(x,y)$, say defined on all of $\mathbb{R}^{2n}$, one can define a generalization of the Legendre transform by

$$\phi^{*,c}(y) := \sup_{x\in}(-c(x,y) - \phi(x)),$$

so that the ordinary Legendre transform is obtained in the standard quadratic case, $c(x,y) = -\langle x, y\rangle$ [be aware that there are many different sign conventions in the litterature!]. A lsc function $\phi(x)$ is said to be $c-$*convex* if $P_c(\phi) = \phi$ where the operator P_c is the "double Legendre transform":

$$P_c\phi := (\phi^{*,c})^{*,c}$$

The goal of this rather open exercise is to go through the notes in the course concerning the quadratic case and verify that they generalize to the non-quadratic case using this generalized form of the Legendre transform [you don't have to worry to much about regularity issues etc].

As an application of this formalism check that Kantorovich duality in the case when $c(x,y) = |x-y|$ (i.e. the L^1-distance) gives the following classical formula for the Wasserstein 1-distance W_1 on $\mathbb{R}^n$

$$W_1(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)} \int\int |x-y| d\gamma(x,y) = \sup_{f\in Lip_1} \int f(\mu-\nu),$$

where the first equality is a definition and Lip_1 denotes the space of all functions on $\mathbb{R}^n$ which are Lipschitz continuous with constant 1, i.e.

$$|f(x) - f(y)| \leq |x-y|$$