Practical session 1: optimal transport on the real line

Lénaïc Chizat

March 10, 2021

The aim of this practical session is to familiarize ourselves with the various objects in optimal transport theory in the most simple setting: the real line $\mathbb{R}$ with the cost $c(x,y) = |y-x|^2$. This will be the occasion to observe in a concrete setting the behaviors that we have exhibited so far (stability, existence of transport maps, geodesics). You may use the language of your choice (Python, Julia, Matlab,...) and return a short report as a jupyter notebook or pdf file (with concise and justified answers, full proofs are not required).

Optimal transport via sorting. In this first part, we consider empirical distributions of the form $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\nu_n = \frac{1}{n}\sum_{i=1}^n \delta_{y_i}$ for $n \in \mathbb{N}^*$ and points $x_i, y_i \in \mathbb{R}$.

1. Write a function that takes as input the n-vectors $(x_i)_{i=1}^n$ and $(y_j)_{j=1}^n$ and returns the permutation of indices that characterize the optimal transport plan.
2. Let $\mu = \frac{3}{4}\mathcal{L}|_{[0,2/3]} + \frac{3}{2}\mathcal{L}|_{[2/3,1]}$ and $\nu = \frac{3}{4}\mathcal{L}|_{[1/3,1]} + \frac{3}{2}\mathcal{L}|_{[0,1/3]}$ where $\mathcal{L}$ is the Lebesgue measure. What is the optimal transport map between μ and ν?
3. Draw n independent samples from μ and from ν and represent the optimal transport plan between the empirical distributions $\hat{\mu}_n$ and $\hat{\nu}_n$ on $[0,1]^2$ (plot is via a point cloud). Plot for $n = 20$ and $n = 50$.
4. Compute the 2-Wasserstein distance between μ_n and ν_n and plot it as a function of n, for $n = 1$ to $n = 500$. What it is the (almost sure) limit of $W_2(\mu_n, \nu_n)$?
5. (optional) Take a random sample with large n (say, $n = 1000$) and represent the W_2 geodesic between μ_n and ν_n at times $t \in \{0, 1/4, 1/2, 3/4, 1\}$ (you may plot it using a histogram). What is the exact expression for the geodesic between μ and ν at $t = 1/2$?

Optimal transport via quantile functions. We now consider probability distributions that are given by their discretized density on a uniform grid. More specifically, we consider a fixed uniform grid $(x_i)_{i=1}^n$ on $[0,1]$ with $x_i = i/n - 1/(2n)$ (say, with $n = 200$) and measures of the form $\mu_a = \sum_{i=1}^n a_i \delta_{x_i}$ where $a \in \mathbb{R}_+^n$ satisfies $\sum a_i = 1$.

6. Write a function that takes as input two n-vectors a and b in $\mathbb{R}_+^2$ such that $\sum a_i = \sum b_i = 1$ and returns the discrete optimal transport plan represented by a matrix $P \in \mathbb{R}_+^{n\times n}$ (i.e. such that the optimal transport plan between μ_a and μ_b is $\sum_{i,j} P_{i,j}\delta_{(x_i,x_j)}$).
7. Let μ_a and μ_b be the discretization of truncated Gaussian distributions of (mean,variance) $(0.2, 0.1^2)$ and $(0.6, 0.2^2)$ respectively (do not forget to normalize the weight vectors after discretization). Represent the optimal transport plan in greyscale on $[0,1]^2$.
8. (optional) Because of the discretization, we see that the optimal transport plan is not deterministic (while it should be in the continuous world). A workaround is to define the *barycentric projection map*
$$T(x_i) = \frac{\sum_{j=1}^n x_j P_{i,j}}{\sum_{j=1}^n P_{i,j}}$$
which is well-defined whenever $a_i = \sum_{j=1}^n P_{i,j} > 0$. Using this map, plot the (approximate) geodesic between μ_a and μ_b at times $t \in \{0, 1/4, 1/2, 3/4, 1\}$.
9. (optional) What is the 2-Wasserstein distance between two Gaussians $\mathcal{N}(m_1, \sigma_1)$ and $\mathcal{N}(m_2, \sigma_2)$? You can make a conjecture using numerical experiments and prove it by constructing a monotone transport map.