18.095: Statistical and Computational Optimal Transport

Lecturer: Philippe Rigollet — Jan. 15, 2020

---

1. EXERCISES

Problem 1 Let X, Y be two random variables and let $Z \sim \mathcal{N}(0,1)$ be a standard Gaussian random variable that is independent of both X and Y.

1. Show that
$$\big|\mathbb{E}[X] - \mathbb{E}[Y]\big| \le W_1(X,Y) \le \mathbb{E}|X| + \mathbb{E}|Y|$$

2. Show that
$$W_1(X+Z, Y+Z) \le W_1(X,Y)$$
Find an example where this inequality is an equality and one where it is a strict inequality.

3. Assume now that Y is a point mass at $y \in \mathbb{R}$. Show that
$$\mathsf{var}(X) \le W_2^2(X,Y)$$
where var denotes the variance.

4. Assume now that $X \in [0,1]$ and $Y \in [0,1]$ have pdfs f and g respectively. Show that
$$W_2(X,Y) \le \frac{1}{2}\int_0^1 |f(x) - g(x)| \mathrm{d}x$$

Problem 2 Find two distributions P and Q such that the optimal transport map between P and Q is not unique

Problem 3 Assume that $X, Y \in \mathbb{R}^2$ are such that
$$X \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right), \qquad Y \sim \mathcal{N}_2\left(\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}\right)$$

For each of the following maps, say whether it is (or not) an optimal transport map between X and Y and why.

1. $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ given by
$$T_1(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \cdot x$$

2. $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ given by
$$T_2(x) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \cdot x$$

3. $T_3 : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T_3(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

4. $T_4 : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T_4(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & -1 \\ 1 & -1 \end{pmatrix} \cdot x$$

Problem 4 *All logs are natural.* Given $c \in \mathbb{R}^d$ and $\beta > 0$, consider the optimization problem

$$\min_{x \in \Delta_d} \left\{ c^\top x + \beta \sum_{j=1}^{d} x_j \log x_j \right\},$$

where Δ_d

$$\Delta_d = \left\{ x = (x_1, \ldots, x_d)^\top \in [0,1]^d \ : \ x_j \geq 0, \sum_{j=1}^{d} x_j = 1 \right\}$$

1. Compute the solution $x^\star(\beta)$ of the above constrained optimization problem.

2. Show that

$$\min_{1 \leq j \leq d} c_j \leq c^\top x^\star(\beta) \leq \min_{1 \leq j \leq d} c_j + \beta \log d$$

3. Can you suggest how to pick a (random) index J such that c_J is close to $\min_{1 \leq j \leq d} c_j$?

4. Let now x, y be any vectors in $[0, \infty)^d$ and define the Kullback-Leibler (KL) divergence between them by

$$\mathrm{KL}(x \| y) := \sum_{j=1}^{d} x_j \log \left(\frac{x_j}{y_j} \right)$$

with the convention that $\log(1/0) = +\infty$ and $0 \log 0 = 0$

(a) Show that if $x, y \in \Delta_d$ then $\mathrm{KL}(x \| y) \geq 0$ with equality iff $x = y$ `[Hint: Jensen's inequality]`

(b) For $y \in [0, \infty)$, compute the KL projection

$$\bar{y} = \operatorname*{argmin}_{x \in \Delta_d} \mathrm{KL}(x \| y)$$

(c) Find an example where it is different from the Euclidean projection

$$\hat{y} = \operatorname*{argmin}_{x \in \Delta_d} \|y - x\|$$

where $\|z\|^2 = \sum_j z_j^2$ is the squared Euclidean norm.