

Mock Exam for the LTCC Course *Introduction to Optimal Transport*

March 2026

Exercise 1 (25 points).

Consider the Wasserstein distance W_2 between two probability measures μ and ν on $\mathbb{R}^n$, defined by

$$W_2(\mu, \nu) := \inf \mathbb{E}(\|X - Y\|_2^2)^{1/2},$$

where the infimum is taken over all random vectors (X, Y) of $\mathbb{R}^n \times \mathbb{R}^n$ with $X \sim \mu$ and $Y \sim \nu$. Show for the case of $W_2(\mathcal{N}(m_1, \Sigma_1), \mathcal{N}(m_2, \Sigma_2))$, that

$$W_2^2(\mathcal{N}(m_1, \Sigma_1), \mathcal{N}(m_2, \Sigma_2)) = \|m_1 - m_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2}).$$

This formula goes back to **Givens and Shortt**, **Knott and Smith**, **Olkin and Pukelsheim**, and **Dowson and Landau**. Note in particular that we have

$$\text{Tr}((\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2}) = \text{Tr}((\Sigma_2^{1/2} \Sigma_1 \Sigma_2^{1/2})^{1/2}).$$

Solution:

One way to prove this claim is to follow the arguments of *Knott and Smith*, and find an optimal transportation map. To construct a good candidate, note that

$$\mathcal{N}(m_2, \Sigma_2)$$

is the image law of

$$\mathcal{N}(m_1, \Sigma_1)$$

with the linear map

$$x \mapsto m_2 + A(x - m_1),$$

where

$$A = \Sigma_1^{-1/2} (\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2} \Sigma_1^{-1/2} = A^\top.$$

To check that this maps $\mathcal{N}(m_1, \Sigma_1)$ to $\mathcal{N}(m_2, \Sigma_2)$, say in the case $m_1 = m_2 = 0$ for simplicity, one may define the random column vectors $X \sim \mathcal{N}(m_1, \Sigma_1)$ and $Y = AX$, and write

$$\begin{aligned} \mathbb{E}(YY^\top) &= A \mathbb{E}(XX^\top) A^\top \\ &= \Sigma_1^{-1/2} (\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2} (\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2} \Sigma_1^{-1/2} \\ &= \Sigma_2. \end{aligned}$$

To check that the map is optimal, one may use

$$\begin{aligned}\mathbb{E}(\|X - Y\|_2^2) &= \mathbb{E}(\|X\|_2^2) + \mathbb{E}(\|Y\|_2^2) - 2\mathbb{E}(\langle X, Y \rangle) \\ &= \text{Tr}(\Sigma_1) + \text{Tr}(\Sigma_2) - 2\mathbb{E}(\langle X, AX \rangle) \\ &= \text{Tr}(\Sigma_1) + \text{Tr}(\Sigma_2) - 2\text{Tr}(\Sigma_1 A),\end{aligned}$$

and observe that by the cyclic property of the trace,

$$\text{Tr}(\Sigma_1 A) = \text{Tr}\left((\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2})^{1/2}\right).$$

Exercise 2

(a) [This exercise is taken from [[1], Examples 2.14, 2.15].]

Let $X, Y \subset \mathbb{R}$ be bounded and $c(x, y) = h(y - x)$ where $h : X \rightarrow Y$ is a linear function. Show that every admissible transport map is optimal, i.e., show that if $T : X \rightarrow Y$, $T\#f = g$, then

$$M(T, f, g) = M^{OT}(f, g).$$

(b) [This exercise is taken from [[1], Example 2.16].]

Let $X = [0, 2]$, $Y = [1, 3]$, $f = \frac{1}{2}\chi_{[0,2]}$, $g = \frac{1}{2}\chi_{[1,3]}$, $c(x, y) = h(y - x)$ with $h(s) = |s|$. Let $T_1(x) = x + 1$ and

$$T_2(x) = \begin{cases} x + 2 & \text{if } x \in [0, 1], \\ x & \text{if } x \in (1, 2]. \end{cases}$$

Observe that f and g have mass in common in the interval $[1, 2]$. The map T_2 leaves the common mass fixed and only transports mass from $[0, 1]$ to $[2, 3]$.

(i) Show that T_1 and T_2 are both optimal transport maps:

$$M(T_1, f, g) = M(T_2, f, g) = M^{OT}(f, g).$$

(ii) Derive an optimal Kantorovich potential pair for the dual problem.

Solution:

(a) We have

$$\begin{aligned}M(T, f, g) &= \int_X c(x, T(x))f(x) dx = \int_X h(T(x) - x)f(x) dx \\ &= \int_X h(T(x))f(x) dx - \int_X h(x)f(x) dx \quad (\text{since } h \text{ is linear}) \\ &= \int_Y h(y)g(y) dy - \int_X h(x)f(x) dx\end{aligned}$$

by the push forward constraint. Therefore $M(T, f, g)$ is independent of T and every admissible transport map is optimal.

(b) (i) We have

$$M(T_1, f, g) = \int_0^2 c(x, T_1(x)) f(x) dx = \int_0^2 |T_1(x) - x| \frac{1}{2} dx = \int_0^2 1 \cdot \frac{1}{2} dx = 1$$

and

$$M(T_2, f, g) = \int_0^2 c(x, T_2(x)) f(x) dx = \int_0^2 |T_2(x) - x| \frac{1}{2} dx = \int_0^1 2 \cdot \frac{1}{2} dx = 1.$$

These maps are optimal since, for any admissible map T such that $T\#f = g$,

$$\begin{aligned} \int_0^2 |T(x) - x| f(x) dx &\geq \left| \int_0^2 (T(x) - x) f(x) dx \right| = \left| \int_0^2 T(x) f(x) dx - \int_0^2 x f(x) dx \right| \\ &= \left| \int_1^3 y g(y) dy - \int_0^2 x f(x) dx \right| \quad (\text{by (3.1) with } \varphi(y) = y) \\ &= \left| \frac{1}{2} \int_1^3 y dy - \frac{1}{2} \int_0^2 x dx \right| = 1. \end{aligned}$$

(ii) One possible choice is $u(x) = -x, v(y) = y$.

Exercise 3 [This is Exercise 4 from [1]]

Let $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be defined by $T(x) = -x$. Characterize the probability measures μ in $\mathbb{R}^d$ such that T is an optimal transport map between μ and $T\#\mu$ for the quadratic cost.

Solution:

To find this characterization, we need to use c-cyclical monotonicity on two points x_1 and x_2 (i.e. monotonicity).

Exercise 4 [This is Exercise 19 from [1]]

Consider the cost $c_\epsilon(x, y) = \sqrt{\epsilon^2 + |x - y|^2}$, and let γ_ϵ be optimal for the cost c_ϵ in $\Gamma(\mu, \nu)$. Take now $\epsilon \rightarrow 0$. Prove that, up to subsequences, we have $\gamma_\epsilon \rightarrow \gamma$, where γ is optimal for the cost $|x - y|$.

Solution:

Consider the Taylor expansion of the cost in ϵ , which allows to guess the lower-order term.

References

- [1] Santambrogio, Filippo, *Optimal Transport for Applied Mathematicians*, Calculus of Variations, PDEs and Modeling, 1-356, 2015.
- [2] Bourne, David *A Brief Introduction to Optimal Transport Theory* LMS Lecture Notes, 2018.