UNIVERSITY OF CAMBRIDGE

MAT3, MAMA

MATHEMATICAL TRIPOS Part III

Monday, 10 June, 2019 1:30 pm to 4:30 pm

PAPER 348

INTRODUCTION TO OPTIMAL TRANSPORT

*Your mark will be determined by answers to your best **FOUR** questions.*

*There are **FIVE** questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

(a) Let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ where X, Y are Polish spaces. (i) Define the *Monge* and *Kantorovich optimal transport problems* between μ and ν. (ii) Give an example of two measures μ and ν for which there does not exist a transport map between μ and ν.
[5 marks]

(b) Show that the infimum in Monge's optimal transport problem is always greater or equal than the infimum in Kantorovich's optimal transport problem. **[8 marks]**

(c) If $\mu \in \mathcal{P}(\mathbb{R})$ and $\nu \in \mathcal{P}(\mathbb{R})$ have cumulative distribution functions F and G respectively, where G is invertible, then write down (without proof) the optimal transport map for the Monge optimal transport problem with cost $c(x, y) = d(x-y)$ where $d : \mathbb{R} \to \mathbb{R}$ is convex and continuous. **[2 marks]**

(d) Using the same assumptions as in part (c), show that any non-decreasing map $T : \mathbb{R} \to \mathbb{R}$ satisfying $T_{\#}\mu = \nu$ is a solution of Monge's optimal transport problem.
[*Hint: Show that the only non-decreasing map T with $T_{\#}\mu = \nu$ is the map given in part (c).*] **[10 marks]**

2

(a) Let $X, Y \subset \mathbb{R}^d$ and $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ have densities f and g with respect to the Lebesgue measure respectively. Assume $T : X \to Y$ is continuously differentiable and bijective. Show that $\nu = T_{\#}\mu$ if and only if $f(x) = g(T(x))|\det(\nabla T(x))|$ for almost every $x \in X$. **[6 marks]**

(b) (i) Show that $\varphi \in C^2(\mathbb{R}^d)$ is convex if it satisfies $\gamma^\top D^2\varphi(x)\gamma \geqslant 0$ for any $\gamma \in \mathbb{R}^d$ (i.e. the Hessian matrix $D^2\varphi(x)$ is positive semi-definite).
(ii) Now consider $\zeta \in C_c^\infty(\mathbb{R}^d)$, $\varepsilon > 0$ with $\varepsilon \max_{x\in\mathbb{R}^d} \|D^2\zeta\| \leqslant \lambda$ and $\varphi \in C^2(\mathbb{R}^d)$ with $\gamma^\top D^2\varphi(x)\gamma \geqslant \lambda$ for all $x \in \mathbb{R}^d$ (i.e. φ is λ-convex). Show that if $T = \nabla\varphi$ then $T_\varepsilon := T + \varepsilon\nabla\zeta$ can be written in the form $T_\varepsilon = \nabla\varphi_\varepsilon$ where $\varphi_\varepsilon \in C^2(\mathbb{R}^d)$ is convex. **[5 marks]**

(c) Let μ be the uniform measure on $[0, 1]$ and ν be the uniform measure on $[1, 2]$. Assume $c(x, y) = h(|x - y|)$ where $h : \mathbb{R}_+ \to \mathbb{R}$ is convex. Prove that $T^\dagger(x) = x + 1$ is an optimal transport map in the Monge sense between μ and ν. **[7 marks]**

(d) Let $c(x, y) = \sqrt{|x - y|}$, μ be the uniform measure on $[0, 1]$ and ν be the uniform measure on $[1, 2]$. Show that $T^\dagger(x) = x + 1$ is the worst transport map for the Monge optimal transport problem between μ and ν with cost c. **[7 marks]**

3

(a) Write down the dual form for the Kantorovich optimal transport problem and write down the assumptions sufficient for duality to hold (i.e. state the *Kantorovich Duality Theorem*). **[4 marks]**

(b) Show that (without referring to weak duality results from convex programming), under the conditions stated in you answer to part (a), the maximum to the dual problem is always less than, or equal to, the minimum to the Kantorovich optimal transport problem. **[7 marks]**

(c) State the *Kantorovich-Rubinstein Theorem.* **[4 marks]**

(d) Let $T^\dagger : \mathbb{R}^d \to \mathbb{R}^d$ be defined by $T^\dagger(x) = \lambda x$ for some fixed $\lambda > 0$ and assume $X \subset \mathbb{R}^d$ is compact. For any probability measure $\mu \in \mathcal{P}(X)$ with finite second moment, show that $T^\dagger$ is an optimal map from μ to $T^\dagger_{\#}\mu$ with respect to the cost $c(x,y) = |x-y|^2$. [*Hint: consider the potential* $\varphi(x) = (1-\lambda)|x|^2$.] **[10 marks]**

4

(a) State the *Knott-Smith Optimality Criterion.* **[4 marks]**

(b) State *Brenier's Theorem.* **[4 marks]**

(c) Let $\varepsilon > 0$ and $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have finite second moments. Let $\pi_\varepsilon \in \Pi(\mu, \nu)$ and $\varphi_\varepsilon \in L^1(\mu)$ be a proper lower semi-continuous convex function such that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} (\varphi_\varepsilon(x) + \varphi_\varepsilon^*(y) - x \cdot y) \, \mathrm{d}\pi_\varepsilon(x, y) \leqslant \varepsilon.$$

Show that $\mathbb{K}(\pi_\varepsilon) \leqslant \inf_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) + \varepsilon$ where $\mathbb{K}$ is the Kantorovich cost with cost function $c(x, y) = \frac{1}{2}|x - y|^2$. **[8 marks]**

(d) Let $c(x, y) = |x - y|^2$ and define μ and ν to be the probability measures on $\mathbb{R}^2$ with densities f and g where

$$f(x) = \frac{1}{\pi} \chi_{B(0,1)}(x), \qquad g(y) = \frac{2|y|^2}{\pi} \chi_{B(0,1)}(y).$$

Show that $T^\dagger : B(0, 1) \to B(0, 1)$ defined by $T^\dagger(x) = \frac{x}{\sqrt{|x|}}$ is an optimal transport map for the Monge problem with cost c. **[9 marks]**

5

(a) Define the *p-Wasserstein distance* d_{W^p} on the space of probability measures on X with bounded p^{th} moment (i.e. $\mathcal{P}_p(X)$) where $X \subseteq \mathbb{R}^d$. **[3 marks]**

(b) Let $X \subset \mathbb{R}^d$ be bounded, $\mu_n, \mu \in \mathcal{P}(X)$, and $p, q \in [1, +\infty)$. Show that $\mu_n \to \mu$ in the metric d_{W^p} if and only if $\mu_n \to \mu$ in the metric d_{W^q}. **[5 marks]**

(c) Assume $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have finite p^{th} moments and there exists a transport map $T^\dagger : \mathbb{R}^d \to \mathbb{R}^d$ that minimises the Monge optimal transport problem with cost $c(x, y) = |x - y|^p$. Define $P_t = (1 - t)\mathrm{Id} + tT^\dagger$ and $\mu_t = [P_t]_{\#}\mu$. Assume that P_t is invertible for all $t \in [0, 1]$ and that the Kantorovich optimal transport cost is equal to the Monge optimal transport cost. Show that

$$d_{W^p}(\mu_t, \mu_s) = |t - s| d_{W^p}(\mu, \nu)$$

where d_{W^p} is the p-Wasserstein distance. **[10 marks]**

(d) Show that, if $X = Y = \mathbb{R}^d$ and $c(x, y) = \mathbb{I}_{x \neq y}$, then $\inf_{\pi \in \Pi(\mu, \nu)} \mathbb{K}(\pi) = \frac{1}{2} \|\mu - \nu\|_{\mathrm{TV}}$ where $\|\mu\|_{\mathrm{TV}} = 2 \sup_{A \subseteq \mathbb{R}^d} |\mu(A)|$. **[7 marks]**

END OF PAPER