

Introduction to Optimal Transport

Matthew Thorpe

Office: F2.08

Email: m.thorpe@maths.cam.ac.uk

Lent Term 2018

1. Overview: Optimal transport crosses many branches of mathematics such as partial differential equations, probability, fluid mechanics and functional analysis. Applications of Optimal Transport are increasing as numerical developments have made computations ever more efficient. We now see applications of optimal transport in (i) image retrieval, registration and morphing, (ii) color and texture analysis, (iii) image denoising and restoration, (iv) morphometry, (v) super resolution, and (vi) machine learning. In this course I aim to give an overview of the theory of optimal transport. Whilst we will cover some of the numerical methods I will largely skip applications.

2. Scheduling: We meet for an hour, twice a week on Mondays and Fridays at 10am in MR5 for the first six weeks in the Lent term.

3. Prerequisites: Little prior knowledge will be needed with just some basic understanding of measure theory and functional analysis.

4. Texts: I recommend Villani's *Topics in Optimal Transportation*. I will mostly use this book, for more recent topics I will use the relevant research papers. Other good references are Villani's 'other' book *Optimal Transport Old and New*, Santambrogio's book *Optimal Transport for Applied Mathematicians*, and Ambrosio, Gigli and Savaré's book *Gradient Flows in Metric Spaces and in the Space of Probability Measures*.

5. Content of Course: I aim to cover the following topics.

Kantorovich Duality. Kantorovich duality forms the basis for many theoretical results regarding optimal transport, for example the equivalence of Monge and Kantorovich's formulation.

Existence and Characterisations of Optimal Transport Maps. We prove existence of optimal transport plans, and their characterisation as the subgradient of a convex function.

Connections to Fluid Mechanics: Benamou and Brenier's Formulation. In their seminal work Benamou and Brenier wrote the Wasserstein distance (an example of an optimal transport distance) as the minimum kinetic energy of an evolving (in time) fluid satisfying the continuity equation and with the endpoints fixed (at $t = 0$ and $t = 1$). This leads to one numerical approach for computing the Wasserstein distance and leads to the understanding of the Riemannian structure in the Wasserstein metric space.

Wasserstein spaces, Geodesics, and Riemannian Structure. Via the Benamou and Brenier formulation we can characterise tangent spaces and the Riemannian structure of the Wasserstein metric space.

Gradient Flows for the Fokker-Planck Equation. Jordan, Kinderlehrer and Otto showed how the Wasserstein distance arises naturally in the gradient flow approach for computing solutions to the Fokker-Planck equation.

Numerical Methods: Cuturi's Entropy Regularised Approach. Arguably the biggest development (at least in recent years) in the computation of optimal transport distances was due to Cuturi's entropy regularised approach. The idea is to use entropy to regularise the distance, then some simple rearrangements reveal this is a Kullback-Liebler divergence. Standard methods, e.g. Sinkhorn's algorithm, can then be used to find minimizers of the entropy regularised distance.