Introduction to Optimal Transport

Matthew Thorpe

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.

Pre-requisites

Familiarity with measure theory and functional analysis will be necessary. Prior exposure to convex analysis will be helpful but not essential.

Literature

1. C. Villani, *Topics in Optimal Transportation.* AMS, 2003.
2. C. Villani, *Optimal Transport: Old and New.* Springer, 2008. Also available at

   `http://cedricvillani.org/wp-content/uploads/2012/08/preprint-1.pdf`

3. F. Santambrogio, *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling.* Birkhäuser, 2015. Also available at

   `https://www.math.u-psud.fr/~filippo/OTAM-cvgmt.pdf`

4. L. Ambrosio, N. Gigli and G. Saveré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Springer Science & Business Media, 2008. Also available at

   `http://www2.stat.duke.edu/~sayan/ambrosio.pdf`

5. G. Peyré and M. Cuturi, *Computational Optimal Transport*, arXiv:1803.00567, 2018. Available at

   `https://arxiv.org/pdf/1803.00567.pdf`

6. S. Daneri and G. Saver'e, *Lecture Notes on Gradient Flows*, arXiv:1009.3737, 2010. Available at

   `https://arxiv.org/pdf/1009.3737.pdf`

7. I will also produce course notes which will be updated throughout term.

Additional support

Four examples sheets will be provided and four associated examples classes will be given (three in Lent term and one in Easter term). The final example class will also be used as a revision class.

Course Content

We will 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.

Wasserstein spaces, Geodesics, and Riemannian Structure. We define the Wasserstein distance and explore its topology, including proving existence of geodesics.

Gradient Flows for the Fokker-Planck Equation. We show how the Wasserstein distance arises naturally in the gradient flow approach for computing solutions to the Fokker-Planck equation.

Numerical Methods. We look at entropy regularised optimal transport which leads to a numerically efficient algorithm for the computation of optimal transport. Indeed, one can write entropy regularised optimal transport as a Kullback-Liebler divergence and use iterative optimisation methods such as Sinkhorns algorithm. Then we consider the flow minimisation approach which uses gradient descent on the Monge optimal transport problem.

Special Cases. Throughout the course we will consider the following special cases in detail: optimal transport on $\mathbb{R}$, discrete optimal transport, and semi-discrete optimal transport.