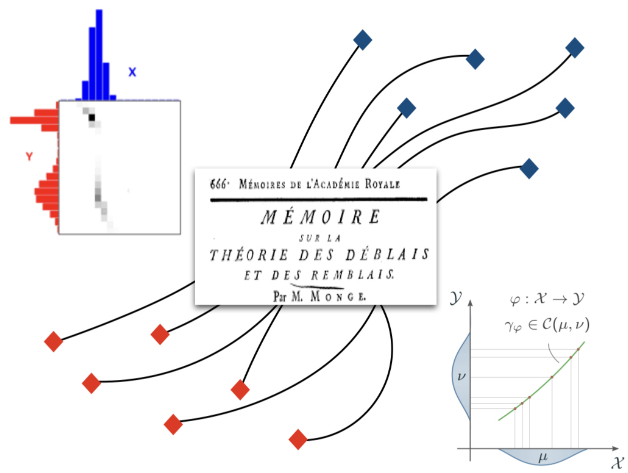
Seminar on
Statistical Aspects of Optimal Transport

winter term 2021/22

Key information

Time:	05 Nov 2021 – 11 Feb 2022, on Fridays, 10.15-12.00
Format:	Room 5.101 (IMS), if Covid-rules permit face-to-face teaching
Possible Modules:	B.Mat.3444: Seminar on mathematical statistics
	M.Mat.4844: Seminar on mathematical statistics
	B.Mat.3447: Seminar on statistical foundations of data science
	M.Mat.4847: Seminar on statistical foundations of data science
Instructors:	Dr. Tobias Kley and Prof. Dr. Axel Munk
Intended Audience:	**Advanced Bachelor and beginning to intermediate Master students**
Language:	English

Prerequisites



This seminar covers an advanced topic alongside the lecture series *Statistical foundations of data science.*

Good knowledge in probability theory is required.

Participants must have successfully attended

- Measure & probability theory (B.Mat.1400), and
- Stochastics (B.Mat.2410),

or provide proof of equivalent qualification.

A background in stochastic processes, optimization and graph theory and having completed the following modules is considered helpful:

- Introduction to mathematical statistics (B.Mat.3144) or
  Introduction to statistical foundations of data science (B.Mat.3147) or
  Mathematical statistics (M.Mat.3140),
- Advances in mathematical statistics (B.Mat.3344) or
  Advances in statistical foundations of data science (B.Mat.3347),
- Seminar im Zyklus 'Mathematische Statistik' (B.Mat.3444) or
  Seminar im Zyklus 'Statistische Grundlagen der Data Science' (B.Mat.3447) or
  Seminar on mathematical statistics (M.Mat.4844) or
  Seminar on statistical foundations of data science (M.Mat.4847).

In particular, this includes the following:

- Statistical foundations of data science, parts I & II (Vst.-Nr. 503026),
- Seminar of Empirical Processes (Vst.-Nr. 503239).

Description

Heuristically, optimal transport can be thought of as the problem of transporting goods to some target at a minimum possible cost. In a more general formulation this can be casted as the problem to find a coupling between two probability measures with given marginals μ and ν such that the expected cost to "transform" μ into ν is minimised. This smallest expected cost to achieve the transport defines a distance on the set of distributions on an underlying metric measure space which takes into account the geometric properties of this space. This yields the so called Wasserstein space and provides the foundation for a remarkably rich theory. In fact, optimal transport is a highly active area in mathematics (Rachev and Rüschendorf, 1998, Santambrogio, 2015, Villani, 2009), and more recently in computer science, machine learning (Peyré and Cuturi, 2019), and statistics (Panaretos and Zemel, 2020). In addition, recently it has been recognised as a novel and highly valuable tool for many tasks in data science (see, e. g., Tameling et al., 2021).

Alternatively, the optimal transport problem can be described using concepts from graph theory. More precisely, if the ground space is finite or countable, then each transport map between two sets of locations can be viewed as a matching in a weighted bipartite graph. Then, the optimal transport solution minimizes the sum of the edge weights. In this sense the solution to the optimal transport problem defines a functional on point sets. One path to understanding the stochastic behavior of optimal transport therefore is an analysis of such functionals.

In this seminar we investigate the subject from a general point of view and cover an approach to describing the stochastic behavior of a wide range of such functionals that also include other interesting minimization problems such as the Traveling salesman problem, minimum spanning trees, and other combinatorial optimisation tasks. After covering basic properties of these functionals its asymptotics for randomly drawn points will be discussed. This provides important insight into the statistical precision of optimal transport estimates from data which, for example, can be used to control the behavior of stochastic algorithms for optimization.

Application

To provide participants with the material to be presented at an early stage, we ask you to preregister for this seminar. To this end, please email Tobias Kley (tobias.kley@uni-goettingen.de) and indicate your interest to give a seminar talk. Please include information about relevant courses you have completed in your email. The deadline for preregistration is **8 October 2021**.

A preparatory virtual meeting, during which topics will be assigned to participating students, is scheduled for **15 October 2021 (10:15-12am)**. Notably, the seminar is limited to 13 participants. Should preregistrations exceed 13, then participants will be chosen based on the information provided in their preregistration email.

Recommended literature

Main Reference

Topics for presentations will be assigned along the lines of

- Yukich, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems.* Springer (Lecture Notes in Mathematics), Berlin.

References for further reading

- Bertsekas, D. P. (1991). *Linear Network Optimization: Algorithms and Codes*, MIT Press. Chapters 1–4 are vailable online: `http://web.mit.edu/dimitrib/www/net.html`.
- Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to Linear Optimization*, Athena Scientfic.
- Herrmann, M. (2020). *Graphentheorie*. Lecture Notes.
  Available online: `https://www.tu-braunschweig.de/ipde/personal/herrmann/skripte`
  (English translation will be made available on StudIP.)
- Panaretos, V. M., and Zemel, Y. (2020). An Invitation to Statistics in Wasserstein Space. Springer. Available online: `https://link.springer.com/book/10.1007/978-3-030-38438-8`.
- Peyré, G., and Cuturi, M. (2019). Computational Optimal Transport. Foundations and Trends in Machine Learning. Available online: `https://optimaltransport.github.io/`.
- Rachev, S. T., and Rüschendorf, L. (1998). *Mass Transportation Problems: Volume I: Theory*. Springer Science & Business Media.
- Rachev, S. T., and Rüschendorf, L. (1998). *Mass Transportation Problems: Volume II: Applications*. Springer Science & Business Media.
- Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*, Springer.
- Tameling, C., Stoldt, S., Stepahn, T., Naas, J., Jakobs, S., Munk, A. (2021). Colocalization for super-resolution microscopy via optimal transport. Nature Computational Science, 1, 199–211.
- Tarjan, R. E. (1983). *Data Structures and Network Algorithms*, SIAM.
- Villani, C. (2009). *Optimal Transport: Old and New*. Springer, Berlin.