



Department of Decision Sciences

Statistics Seminars

Faster Wasserstein distance estimation with the Sinkhorn divergence

Lénaïc Chizat

École polytechnique fédérale de Lausanne (EPFL)

Thursday, 16th September 2021

12:30 pm Room 3-E4-SR03 Via Roentgen 1 Milano

Abstract

The squared Wasserstein distance is a natural quantity to compare probability distributions in a non-parametric setting. This quantity is usually estimated with the plug-in estimator, defined via a discrete optimal transport problem. It can be solved to ε -accuracy by adding an entropic regularization of order ε and using for instance Sinkhorn's algorithm. In this work, we propose instead to estimate it with the Sinkhorn divergence, which is also built on entropic regularization but includes debiasing terms. We show that, for smooth densities, this estimator has a comparable sample complexity but allows higher regularization levels, of order $\varepsilon^{1/2}$, which leads to improved computational complexity bounds and a speedup in practice. We also propose and analyze an estimator based on Richardson extrapolation of the Sinkhorn divergence which enjoys improved statistical and computational efficiency guarantees, under a condition on the regularity of the approximation error, which is in particular satisfied for Gaussian densities. This is joint work with Pierre Roussillon, Flavien Léger, François-Xavier Vialard and Gabriel Peyré