Weak Limits of Sinkhorn Optimal Transport

Dr. Alberto Gonzalez Sanz
(Université Toulouse III - Paul Sabatier)

May 03, 2023

One of the most classic and important statistical problems is hypotheses testing, with a non-parametric goodness-of-fit test „$(H0 : P = QvsH1 : P \neq Q)$“ being a standout among them. To perform this test, a notion of similarity between probability measures is necessary. Optimal transportation defines a distance between probability measures that takes into account weak and moment convergence as well as the geometry of the latent space. However, the curse of dimensionality affects optimal transportation, reducing its usefulness in this framework.

This presentation will delve into the topic of the asymptotic distribution of entropic regularized optimal transport –one of the most widely used variations of the optimal transport problem. It defines a measure of similarity between probability measures, the Sinkhorn divergence. The talk will analyze the effect of dimensionality on statistical complexity and demonstrate that the Sinkhorn divergence, unlike its non-regularized counterpart, is not impacted by dimensionality. In conclusion, the asymptotic distribution of the Sinkhorn divergence under both hypotheses will be studied, with a view to establishing asymptotically consistent tests that are based on optimal transport. This analysis will provide a deeper understanding of the behaviour of the Sinkhorn divergence and its usefulness in hypothesis testing.

The primary objective of this talk is to provide an educational overview of the techniques used to obtain the results in this specific problem. The aim is to present these techniques in a clear and systematic manner that makes them readily reproducible for other statistical and/or machine learning problems. The focus will be on imparting a thorough understanding of the methods employed, rather than simply presenting the results.