The Wasserstein Barycenter Problem: Formulation, Computation and Applications

Marco Cuturi
Kyoto University, Japan
mcuturi@i.kyoto-u.ac.jp

How can we define the average of a set of probability measures? This question is important because (1) averaging ranks among the most elementary operations used in statistics to summarize data (2) probability measures are ubiquitous in machine learning, where they are used to represent datasets, generative models or complex objects (an image as a bag-of-features, a text as a bag-of-words).

I will present in this talk a possible answer to this question grounded on the optimal transport (a.k.a. Wasserstein/ earth mover's) geometry. The problem I will describe, known as the Wasserstein barycenter problem, tries to find, given a set of probability measures of interest, the probability measure that minimizes the sum of all its Wasserstein-distances to those probability measures. After providing a few self-contained reminders on optimal transport in the first part of the talk, I will illustrate using toy data that Wasserstein barycenters have several intuitive and appealing properties. I will then show that in its original form theWasserstein barycenter problem is intractable, but that it can be solved approximately, very efficiently, and to arbitrary precision in practice by regularizing it with an entropic term. I will provide details of very recent algorithmic advances in this nascent field followed by an application to the visualization of datasets of brain activation maps.