



Contribution ID: 152

Type: not specified

Alexander Cloninger - Learning on and near Low-Dimensional Subsets of the Wasserstein Manifold

Tuesday, October 4, 2022 5:00 PM (1 hour)

Detecting differences and building classifiers between distributions $\{\mu_i\}_{i=1}^N$, given only finite samples, are important tasks in a number of scientific fields. Optimal transport (OT) has evolved as the most natural concept to measure the distance between distributions, and has gained significant importance in machine learning in recent years. There are some drawbacks to OT: computing OT can be slow, and because OT is a distance metric, it only yields a pairwise distance matrix between distributions rather than embedding those distributions into a vector space. If we make no assumptions on the family of distributions, these drawbacks are difficult to overcome. However, in the case that the measures are generated by push-forwards by elementary transformations, forming a low-dimensional submanifold of the Wasserstein manifold, we can deal with both of these issues on a theoretical and a computational level. In this talk, we'll show how to embed the space of distributions into a Hilbert space via linearized optimal transport (LOT), and how linear techniques can be used to classify different families of distributions generated by elementary transformations and perturbations. The proposed framework significantly reduces both the computational effort and the required training data in supervised settings. Similarly, we'll demonstrate the ability to learn a near isometric embedding of the low-dimensional submanifold. Finally, we'll provide non-asymptotic bounds on the error induced in both the supervised and unsupervised algorithms from finitely sampling the target distributions and projecting the LOT Hilbert space into a finite dimensional subspace. We demonstrate the algorithms in pattern recognition tasks in imaging and provide some medical applications.