Estimation of probability measures in high dimensions, with optimal transport and fast algorithms

We introduce a novel class of algorithms for the estimation of probability measures in high-dimensional spaces, given a finite number of samples. We are particularly interested in the case when the probability measure is concentrated near a low-dimensional set. These algorithms are based on geometric multiscale decompositions of probability measures, and we prove that with high probability, given a sufficiently large but finite number of samples, the algorithm returns a probability measure which is close, in Wasserstein-Kantorovich distance, to the target probability measure. We discuss applications to modeling high-dimensional noisy data sets, and anomaly detection in time-varying data. (Received October 02, 2013)