TANYA SCHMAH, University of Ottawa Comparing the UMAP and CkNN graph constructions

Given data sampled from an arbitrary density on an unknown manifold, we consider two methods for constructing a discrete representation of the manifold. The Continuous k-Nearest Neighbours (CkNN) method (Berry and Sauer 2019) uses a "self-tuning" variable-bandwidth kernel (Zelnick-Manor and Perona 2005) to produce an unweighted graph, which the authors then use to compute homology. Independently, McInnes et al. (2018) have introduced the UMAP dimensionality reduction method, which includes a weighted graph construction with some similarities to the self-tuning kernel method used in CkNN. We compare these two graph constructions, and variations of them, and present computational experiments in topological data analysis and dimensionality reduction.