Chen-Yun Lin
(Duke University)

Data, Curse of Dimensionality, and Spectral Geometry

High dimensional data is increasingly available in many fields. However, the analysis of such data suffers the so-called curse of dimensionality. One powerful approach to nonlinear dimensionality reduction is the diffusion type maps. Its continuous counterpart is the embedding of a manifold using the eigenfunctions of the Laplace-Beltrami operator. Accordingly, one may ask, how many eigenfunctions are required in order to embed a given manifold.

In this talk, I will give some background regarding the dimensionality reduction problem, spectral geometry, and show theoretical results for a generalized diffusion map. Specifically, I will show a closed Riemannian manifold can be embedded into a finite dimensional Euclidean space by maps constructed based on the connection Laplacian at a certain time. This time and the embedding dimension can be bounded in terms of the dimension and the geometric bounds of the manifold. Furthermore, the map based on heat kernels can be made arbitrarily close to an isometry. In addition, I will give a "real world" example pertaining to paleontology, that demonstrates how heat kernels and diffusion maps can be used to quantify the similarity of shapes. The empirical results suggest that this framework is better than the metric commonly used in biological morphometrics.