MANIFOLD LEARNING FROM THE USER'S PERSPECTIVE

MARINA MEILA

Principal Components Analysis/Karhunen-Loewe expansion (PCA) is a time honored method to visualize, understand or simply reduce data size. Why, then the need for other, non-linear dimension reduction methods? And why are there so many of them? This tutorial will explain how to navigate the landscape of modern dimension reduction algorithms like Isomap, Diffusion Maps, t-SNE and UMAP. The focus will be on how to safely interpret these algorithms' results in practice. This will be illustrated with data from MD simulations.

Thanks to Stefan Chmiela and Chris Fu, and to the Tkatchenko and Pfaendtner labs and to my students Wenyu Bo, Yu-chia Chen, Samson Koelle, Alex Kokot, James McQueen, Dominique-Perrault-Joncas, Weicheng Wu, Hanyu Zhang