University of Chicago Computational and Applied Math Student Seminar Fall 2022

LINSCAN — A Linearity-Based Clustering Algorithm

Andrew Dennehy (University of Chicago, CAM)

Monday, October 24, 12-1 pm at Jones 303

Abstract

DBSCAN and OPTICS are well-understood, stable clustering algorithms for nonparametrically analyzing the geometry of datasets in R^n. In this talk, we discuss the generalization of these algorithms to notions of distance besides the 2-norm; in particular, we discuss how stability of the algorithm depends on only the symmetry of the distance function and thus we can consider clustering using semi-metrics. We apply this to the goal of identifying approximately linear clusters while ignoring isotropic clusters by embedding points in probability space and clustering using a semi-metric on this space derived from a linearization of the KL Divergence; this has applications in geophysics, in particular for identifying slip faults from measurements of seismic activity.