Adaptive Covers of the Mapper Graph using Information Theory

Nithin Chalapathi
University of Utah

UUCS-21-009

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

27 April 2021

Abstract

Topological data analysis (TDA) has recently grown in popularity for analyzing and visualizing high-dimensional data. One of the most popular tools from TDA is the mapper construction, known as the mapper graph in the 1-dimensional setting. An obstacle in the applications of the mapper graph is parameter selection, specically the choice of a cover. In this thesis, we explore strategies for computing an adaptive cover for the mapper graph using information theoretic measures, namly the Akaike information criterion and Bayesian information criterion. We develop a new strategy inspired by X-means, called multi-pass X-means, for adaptively splitting cover elements. We demonstrate that multi-pass X-means produces mapper graphs that approximate the topology of hand-tuned mapper graphs via experimental results on synthetic and real-world datasets. Our preliminary results show that the usage of information theoretic measures is a promising direction for parameter selection. Finally, we study a variant of the mapper graph, called the enhanced mapper graph, introduced by Brown et al. [3]. We provide an open-source library with both our adaptive cover strategies and one of the rst implementations of the enhanced mapper graph.