COMPUTING BUMP-FUNCTION LANDSCAPE IN TOPOLOGICAL DATA ANALYSIS

C.W. Sahabandu

Department of Mathematics, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka
chathuriws@sci.pdn.ac.lk

Topological Data Analysis (TDA) is a modern field that emerged from various work in computational geometry and applied topology. Persistent homology is a tool that is used extensively in TDA to analyse the shape of data. To find the persistent homology, data must be first represented as a simplicial complex, then the geometric properties of the data are computed using a tool called homology that appears in algebraic topology. The process of constructing a simplicial complex using raw data uses a parameter, the radius r that is used to draw a ball around each point. Changing this parameter, one can obtain the persistence homology of the dataset, and a signature that tells about the persisting topological features of the dataset. Persistent homology can be summarized in a variety of ways; the most prominent ones being the barcode, persistence diagrams and persistent landscape. Although barcodes contain all the information, it is difficult to do statistical analysis using them. Persistence landscape, introduced by Peter Bubenik, solves this problem by mapping the summary into an L^p space. The function used in the construction of persistence landscape is a piecewise linear function. In this study, a new topological summary for data called the bump landscape is proposed. A bump function is a smooth function which has a compact support. The advantages of using a bump function over a piecewise linear function is that most properties of the function are preserved under the sum, product, average and the convolution. The topological summary given by the bump landscape retains all the properties of persistence landscape. The smoothness of the bump function may give additional geometric features of data. In addition, a code was written to generate the bump landscape for synthetic data using ripser, scikit-TDA and java script.

Keywords: Barcodes, Bump functions, Persistence homology, Persistence landscape, Topological data analysis