Bridging Centrality and Extremity: Refining Empirical Data Depth Using Extreme Value Theory

Regina Liu*
Rutgers University, USA rliu@stat.rutgers.edu
John Einmahl
Tilburg University, The Netherlands J.H.J.Einmahl@uvt.nl
Jun Li
University of California, Riverside, USA jun.li@ucr.edu

A data depth is a measure of centrality of a point with respect to a given distribution or data set. It provides a natural center-outward ordering of multivariate data points, and yields a new systematic nonparametric multivariate analysis. The depth approaches to multivariate analysis have had many successes so far. In particular, the approaches derived from geometric depths (e.g. halfspace depth and simplicial depth) are especially useful since they reflect more accurately the true probabilistic geometry underlying the data. However, the empirical geometric depths are defined to be zero for the regions outside the convex hull that spans the data set. This property has long restricted much the utility of depth approach in applications where the extreme outlying probability mass may be the focal point, such as in problems of classification or quality control charts with small false alarm rates. To overcome this shortcoming, we propose to apply extreme value theory to refine the empirical estimates of half-space depth for points which are outside the data hull. This provides an important linkage between data depth, which is useful for inference on centrality, and extreme value theory, which is useful for inference on extremity. The refined estimate of the half-space depth can thus extend depth utilities beyond its data hull to the entire sample space, and broaden greatly the applicability of data depth. We show that the proposed refined depth is a consistent estimator and significantly improves upon the original empirical estimator. The improvement is shown to have significant impact on the applications of data depth such as in multivariate classification and construction of control charts.

Key Words: Data depth, extreme value theory, half-space depth, simplicial depth