Clustering of Multivariate Data with Varying Dimensions

Bin Cheng

Columbia University
E-mail: bc2159@cumc.columbia.edu

Abstract: In many biomedical research, it is of interest to classify patients according to their medical information. Such information is not only of high but also varying dimensions. For example, the toxicity profile, microbial profile, or health record profile. We propose a method to cluster data of varying dimensions. Simulations demonstrate that the method has nice discriminating ability. A physical activity example is discussed as an illustration of the method.