Undirected Network-Based Unsupervised Clustering

Hongmei Zhang
Professor, Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN 38152, USA

Xichen Mou
Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN 38152, USA

Abstract:

Development of certain diseases is driven by the concerted work of genetic or epigenetic factors. In unsupervised clustering, it is thus desired to cluster objects based on joint features of such factors. In this piece of work, under the Bayesian framework, we propose an approach to cluster subjects from the lenses of undirected networks, in particular, Gaussian graphs. Pseudo nodes are introduced to assist the selection of tuning parameters in graph construction. Furthermore, since the clustering focuses on the underlying network among the factors (nodes), within cluster variations among subjects with respect to the means of the nodes are allowed and modeled by the Dirichlet process. Simulations and real data applications are implemented to demonstrate the proposed method.