Data Science Seminar

Hosted by Department of Mathematical Sciences

- Date: Tuesday, March 26, 2020
- Time: 12:00pm – 1:00pm
- Room: WH-100E
- Speaker: Wangshu Tu (Binghamton University)
- Title: A family of mixture models for biclustering

Abstract

Biclustering allows for simultaneous clustering of the observations and variables. Martella et. al (2008) introduced biclustering in a model-based clustering framework by utilizing a structure similar to a mixture of factor analyzer structures such that observed variables are modelled using a latent variable that is assumed to be from a MVN(0, I). In Martella et. al (2008), clustering of variables was introduced by imposing constraints on the entries of the factor loading matrix to be 0 and 1. However, this approach restricts the non-zero off-diagonal entries of the covariance matrix to be 1, which is very restrictive. Here, we assume the latent variable to be from a MVN(0,T) where T is a diagonal matrix and hence, the non-zero off-diagonal entires of the covariance matrix are not restricted to be equal to 1. A family of models are developed by imposing constraints on the components of the covariance matrix. An alternating expectation conditional maximization(AECM) algorithm is used for parameter estimation. Proposed method will be illustrated using simulated and real datasets. The presentation will conclude with some on-going work and future research directions.