MASKING VARIABLES IN MIXTURE MODELING

Emilie Rausch

Dr. Douglas Steinley, Dissertation Supervisor

ABSTRACT

Finite normal mixture modeling is a popular technique for clustering individuals into distinct subpopulations. A characteristic of mixture modeling that researchers may be unaware of is its tendency to arrive at locally, rather than globally, optimal solutions, and these local solutions can lead to an incorrect partition of the individuals into groups. This paper examines the behavior of mixture modeling under different conditions, specifically focusing on when data contain noise and when the mixtures' means are various distances (effect sizes) from one another. Monte Carlo simulations were conducted and it was found that the ability of the mixture model to obtain correct partitions and stable parameter estimates degenerates when noise is added. In addition, the results suggest that as the distance between clusters increases, recovery only reaches a moderate level at almost three times a large effect size.