
VINAY JOSHY, University of Guelph

Sparse Group Lasso for Variable Selection in Finite Gaussian Mixture Regression Models

When analyzing heterogeneous data with latent subpopulations, finite mixture regression (FMR) models are effective as they allow for variations in regression coefficients across mixture components. Variable selection is important at two levels: (1) at the group level by removing completely irrelevant covariates across all subpopulations and (2) at the individual level by removing irrelevant covariates within each subpopulation. However, existing variable selection methods for FMR focus solely on individual-level selection using the least absolute shrinkage and selection operator (LASSO), providing no mechanism for group-level variable elimination. We introduce a sparse group LASSO regularization method that performs efficient variable selection at both levels simultaneously in finite Gaussian mixture regression models. We developed a novel optimization procedure through a Majorization-Minimization algorithm. Simulation studies demonstrate our proposed method's effectiveness in retaining relevant covariates and accuracy in clustering observations into subpopulations. We apply our proposed method to a Chiroptera (bat) dataset exhibiting morphological heterogeneity for identifying relevant environmental and morphological covariates that influence bat forearm development.