ABSTRACT

Graphical Gaussian models have proven to be useful tools for exploring network structures based on multivariate data. Applications to studies of gene expression have generated substantial interest in these models. Recent progress includes the development of fitting methodology involving penalization of the likelihood function and Bayesian approaches putting explicit priors on graphs. In this paper we advocate the use of multivariate $t$-distributions for more robust inference of graphs.

We demonstrate that penalized likelihood inference combined with an application of the EM algorithm provides a computationally efficient approach to model selection in the $t$-distribution case. We consider two versions of multivariate $t$ distributions, the Classical multivariate $t$ distribution and the more flexible Alternative $t$ distribution which requires the use of approximation techniques. For this distribution, we describe a Markov chain Monte Carlo EM algorithm based on a Gibbs sampler as well as a simple variational approximation that makes the resulting method feasible in large problems.

We also show how Bayesian approaches based on Gaussian distributions can be extended to $t$ distributions. The main challenge here is to develop extensions that are computationally efficient. We introduce a third $t$ distribution model involving Dirichlet process priors that maintains much of the flexibility of the Alternative model, while reducing computational costs.