Abstract

Gaussian graphical models explore dependent relationships between random variables, through estimation of the corresponding inverse covariance (precision) matrices. We develop an estimator for such models appropriate for heterogeneous data; specifically, data obtained from different categories that share some common structure, but also exhibit differences. We propose a method which jointly estimates several graphical models corresponding to the different categories present in the data. The method aims to preserve the common structure, while allowing for differences between the categories. This is achieved through a hierarchical penalty that targets the removal of common zeroes in the precision matrices across categories. We establish the asymptotic consistency and persistency of the proposed estimator in the high-dimensional case, and illustrate its superior performance on a number of simulated networks. An application to learning semantic connections between terms from webpages collected from computer science departments is also included. Some extensions to Markov networks (suitable for binary/categorical variables) are also discussed.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.