pcalg: Estimating and visualizing high-dimensional dependence structures using the PC-algorithm

Markus Kalisch

February 22, 2006

We consider the PC-algorithm (Spirtes, Glymour and Scheines (2000)) for estimating the skeleton of a potentially very high-dimensional acyclic directed graph (DAG) with corresponding Gaussian distribution. The PC-algorithm is computationally feasible for sparse problems with many nodes, i.e. variables, and it has the attractive property to automatically achieve high computational efficiency as a function of sparseness of the true underlying DAG.

The restriction of underlying Gaussian distribution can be relaxed by using a robust scale estimator with high precision and high breakdown point (Q_n estimator).

We provide theoretical consistency results on this algorithm and analyze its properties in simulations. Furthermore, we introduce the new R-package pcalg, which performs both the standard and the robust version of the discussed algorithm. This package was used to obtain all simulation results.