The Annals of Statistics
2012, Vol. 40, No. 4, 1973–1977
DOI: 10.1214/12-AOS980
Main article DOI: 10.1214/11-AOS949

DISCUSSION: LATENT VARIABLE GRAPHICAL MODEL SELECTION VIA CONVEX OPTIMIZATION

BY STEFFEN LAURITZEN AND NICOLAI MEINSHAUSEN

University of Oxford

We want to congratulate the authors for a thought-provoking and very interesting paper. Sparse modeling of the concentration matrix has enjoyed popularity in recent years. It has been framed as a computationally convenient convex ℓ_1-constrained estimation problem in Yuan and Lin (2007) and can be applied readily to higher-dimensional problems. The authors argue—we think correctly—that the sparsity of the concentration matrix is for many applications more plausible after the effects of a few latent variables have been removed. The most attractive point about their method is surely that it is formulated as a convex optimization problem. Latent variable fitting and sparse graphical modeling of the conditional distribution of the observed variables can then be obtained through a single fitting procedure.

Practical aspects. The method deserves wide adoption, but this will only be realistic if software is made available, for example, as an R-package. Not many users will go to the trouble of implementing the method on their own, so we will strongly urge the authors to do so.

An imputation method. In the absence of readily available software, it is worth thinking whether the proposed fitting procedure can be approximated by methods involving known and well-tested computational techniques. The concentration matrix of observed and hidden variables is

$$K = \begin{pmatrix} K_O & K_{OH} \\ K_{HO} & K_H \end{pmatrix},$$

where we have deviated from the notation in the paper by omitting the asterisk. The proposed estimator $\hat{S}_n = \hat{K}_O$ of K_O was defined as

$$(\hat{K}_O, \hat{L}_n) = \operatorname{argmin}_{S,L} -\ell(S - L; \Sigma_O^n) + \lambda_n\big(\gamma \|S\|_1 + \operatorname{tr}(L)\big) \tag{1}$$

$$\text{such that } S - L \succ 0, L \succ 0, \tag{2}$$

where Σ_O^n is the empirical covariance matrix of the observed variables.

An alternative would be to replace the nuclear-norm penalization with a fixed constraint κ on the rank of the hidden variables, replacing problem (1) with

$$\begin{aligned} (\hat{K}_O, \hat{L}_n) &= \operatorname{argmin}_{S,L} -\ell(S - L; \Sigma_O^n) + \lambda_n \|S\|_1 \\ &\text{such that } S - L \succ 0 \text{ and } L \succ 0 \text{ and } \operatorname{rank}(L) \le \kappa. \end{aligned} \tag{3}$$

Received February 2012.