The Annals of Statistics
2012, Vol. 40, No. 4, 1984–1988
DOI: 10.1214/12-AOS984
Main article DOI: 10.1214/11-AOS949

DISCUSSION: LATENT VARIABLE GRAPHICAL MODEL SELECTION VIA CONVEX OPTIMIZATION

BY CHRISTOPHE GIRAUD AND ALEXANDRE TSYBAKOV

Ecole Polytechnique and CREST-ENSAE

Recently there has been an increasing interest in the problem of estimating a high-dimensional matrix K that can be decomposed in a sum of a sparse matrix S^* (i.e., a matrix having only a small number of nonzero entries) and a low rank matrix L^*. This is motivated by applications in computer vision, video segmentation, computational biology, semantic indexing, etc. The main contribution and novelty of the Chandrasekaran, Parrilo and Willsky paper (CPW in what follows) is to propose and study a method of inference about such decomposable matrices for a particular setting where K is the precision (concentration) matrix of a partially observed sparse Gaussian graphical model (GGM). In this case, K is the inverse of the covariance matrix of a Gaussian vector X_O extracted from a larger Gaussian vector (X_O, X_H) with sparse inverse covariance matrix. Then it is easy to see that K can be represented as a sum of a sparse precision matrix S^* corresponding to the observed variables X_O and a matrix L^* with rank at most h, where h is the dimension of the latent variables X_H. If h is small, which is a typical situation in practice, then L^* has low rank. The GGM with latent variables is of major interest for applications in biology or in social networks where one often does not observe all the variables relevant for depicting sparsely the conditional dependencies. Note that formally this is just one possible motivation and mathematically the problem is dealt with in more generality, namely, postulating that the precision matrix satisfies

$$K = S^* + L^* \tag{1}$$

with sparse S^* and low-rank L^*, both symmetric matrices. A small amendment to that inherited from the latent variables motivation is that L^* is assumed negative definite (in our notation, L^* corresponds to $-L^*$ in the paper). We believe that this is not crucial and all the results remain valid without this assumption.

CPW propose to estimate the pair (S^*, L^*) from a n-sample of X_O by the pair $(\widehat{S}, \widehat{L})$ obtained by minimizing the negative log-likelihood with mixed ℓ^1 and nuclear norm penalties; cf. (1.2) of the paper. The key issue in this context is identifiability. Under what conditions can we identify S^* and L^* separately? CPW provide geometric conditions of identifiability based on transversality of tangent spaces to the varieties of sparse and low-rank matrices. They show that, under these conditions, with probability close to 1, it is possible to recover the support of S^*, the rank

Received February 2012.