The Annals of Statistics
2012, Vol. 40, No. 4, 1997–2004
DOI: 10.1214/12-AOS1001
Main article DOI: 10.1214/11-AOS949

DISCUSSION: LATENT VARIABLE GRAPHICAL MODEL SELECTION VIA CONVEX OPTIMIZATION

BY EMMANUEL J. CANDÉS AND MAHDI SOLTANOLKOTABI

Stanford University

We wish to congratulate the authors for their innovative contribution, which is bound to inspire much further research. We find latent variable model selection to be a fantastic application of matrix decomposition methods, namely, the superposition of low-rank and sparse elements. Clearly, the methodology introduced in this paper is of potential interest across many disciplines. In the following, we will first discuss this paper in more detail and then reflect on the versatility of the low-rank + sparse decomposition.

Latent variable model selection. The proposed scheme is an extension of the *graphical lasso* of Yuan and Lin [15] (see also [1, 6]), which is a popular approach for learning the structure in an undirected Gaussian graphical model. In this setup, we assume we have independent samples $X \sim \mathcal{N}(0, \Sigma)$ with a covariance matrix Σ exhibiting a sparse dependence structure but otherwise unknown; that is to say, most pairs of variables are conditionally independent given all the others. Formally, the concentration matrix Σ^{-1} is assumed to be sparse. A natural fitting procedure is then to regularize the likelihood by adding a term proportional to the ℓ_1 norm of the estimated inverse covariance matrix S:

$$\text{minimize } -\ell(S, \Sigma_0^n) + \lambda \|S\|_1 \tag{1}$$

under the constraint $\mathbf{S} \succeq 0$, where Σ_0^n is the empirical covariance matrix and $\|S\|_1 = \sum_{ij} |S_{ij}|$. (Variants are possible depending upon whether or not one would want to penalize the diagonal elements.) This problem is convex.

When some variables are unobserved—the observed and hidden variables are still jointly Gaussian—the model above may not be appropriate because the hidden variables can have a confounding effect. An example is this: we observe stock prices of companies and would like to infer conditional (in)dependence. Suppose, however, that all these companies rely on a commodity, a source of energy, for instance, which is not observed. Then the stock prices might appear dependent even though they may not be once we condition on the price of this commodity. In fact, the marginal inverse covariance of the observed variables decomposes into two terms. The first is the concentration matrix of the observed variables in the full model conditioned on the latent variables. The second term is the effect of

Received April 2012.