Stat 375 Inference in Graphical Models

Homework 6

Due - 5/21/2012

Please return this homework in class or in the bin in Packard second floor, or by e-mail to Yiqun Liu (liuyiqun1124@gmail.com).

In this homework we will consider structural learning of Gaussian graphical models. Namely we will consider the gaussian measure over $x \in \mathbb{R}^n$

$$\mu(\mathrm{d}x) = \frac{1}{Z(\Theta)} \exp\Big\{ - \frac{1}{2}\langle x, \Theta\, x\rangle \Big\}\, \mathrm{d}x\,, \tag{1}$$

with an inverse covariance matrix Θ that is sparse. The objective is to reconstruct the graph structure (equivalently, the support of Θ) from data. We will follow the approach of

N. Meinshausen and P. Bühlmann, *High-dimensional graphs and variable selection with the Lasso*, Ann. Statist. 34, 3 (2006), 1436-1462.

The graph structure to be reconstructed will be $G = (V, E)$ a 10×10 two-dimensional grid. The 'true' values of the parameters is given by

$$\Theta_{ij} = \begin{cases} -1 & \text{if } (i,j) \in E, \\ \deg(i) + 4 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We denote by p the number of node of V: $p = |V|$.

(1) Check that the above Gaussian measure is well defined, i.e. that $\Theta \in \mathbb{R}^{p\times p}$ is positive definite.

(2) Write a code that generates samples $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ i.i.d. with distribution μ.
[Hint: Diagonalize $\Theta = UDU^T$ where U is orthogonal and D is diagonal. Then generate $z \sim \mathsf{N}(0, I_{p\times p})$ and let $x \equiv UD^{-1/2}z$.]

(3) Consider vertex $i \in V$, and let $\mathbf{x}_i = (x_i^{(1)}, \dots, x_i^{(n)})^T \in \mathbb{R}^n$ be the (column) vector of realizations of variable x_i in the n i.i.d. samples, generated as above. Further, let $\mathbf{X}_{\setminus i} \in \mathbb{R}^{n\times(p-1)}$ be the matrix with columns $(\mathbf{x}_j, j \in [p]\setminus i)$. We estimate the neighborhood of i by solving the ℓ_1-penalized least squares problem

$$\widehat{\beta}(i) = \arg \min_{\beta\in\mathbb{R}^{p-1}} \Big\{ \frac{1}{2n}\big\|\mathbf{x}_i - \mathbf{X}_{\setminus i}\beta\big\|_2^2 + \lambda_n \|\beta\|_1 \Big\}\,, \tag{2}$$

and estimating the neighborhood of i by letting

$$(i,j) \in \widehat{E} \text{ if and only if } \widehat{\beta}(i)_j \neq 0\,. \tag{3}$$

Here $\widehat{E}$ is the estimated edge set. If the procedure is successful, then this definition is consistent i.e. $\widehat{\beta}(i)_j \neq 0$ if and only if $\widehat{\beta}(j)_i \neq 0$.

Write a program that, given data $x^{(1)}, \dots, x^{(n)} \in \mathbb{R}^p$, outputs $\widehat{E}$ according to the above procedure.

[Hint 1: Packages for solving (2) are available in R (e.g. lars, glmnet), Matlab (e.g. cvx), Python (e.g. cvxopt), etc. If you prefer to write an algorithm for solving (2) from scratch, ant want to discuss it further, you are welcome to office hours.]
[Hint 2: For reasons of numerical accuracy, it is often preferable to replace the rule $\widehat{\beta}(i)_j \neq 0$ in Eq. (3) with something like $|\widehat{\beta}(i)_j| \geq 0.001$.]

(4) Carry out numerical experiments with your program. Namely, attempt reconstruction of the graph G from data, using $n \in \{30, 60, 120, 240\}$. In each case, select a value of λ_n that yields a reasonably good reconstruction (if such a value can be found), and report the size of the symmetric difference between E and $\widehat{E}$, i.e.

$$\Delta \equiv |E \cup \widehat{E}| - |E \cap \widehat{E}| \,. \tag{4}$$