Graphical Models

October 5, 2020

We are interrested in the patterns of dependence in a p-dimensional Normal distribution. The data x are centered so we assume:

$$x \sim N(0, \Sigma)$$

with variance matrix Σ and precision matrix $\Omega = \Sigma^{-1}$. Elements of Ω and Σ are ω_{ij} and σ_{ij}. Write x_i for the ith univariate element of x and, for any subset of indices $K \in \{1, ..., p\}$ write $x_K = \{x_i : i \in K\}$. Important subsets are $[-i] = \{1, ..., p\} \setminus \{i\}$ and $k : m = \{k, k+1, ..., m\}$. The regression coefficients for any x_j on x_i with $i \neq j$ are ω_{ij}/ω_{ii}

1 sparse gaussian graphical models

$ne(i) = \{j : \omega_{ij} \neq 0\}$ is the neighborhood of i.
The focus is on sparse models so $|ne(i)| \ll p$
The graph $\Upsilon = (V, E)$, that is the vertices V and the edges E.

2 compositional networks

Now suppose we fit p univariate regressions using a subset $cpv(i) \subseteq \{(i+1) : p\}$

$$x_i = \Sigma_{j \in cpv(i)} \gamma_{ij} x_j + \epsilon_i$$

The join distribution becomes $x = \Gamma x + \epsilon$; $\epsilon \sim N_p(0, \Psi)$ and $\Gamma_{ij} = \gamma_{ij}$ Create graph by connecting each element of $cpv(i)$ to i. This gives a directed acyclic graph (DAG). Now turn all arrows into undirected edges and marry all parents. An egde between i and j corresponds to a non-zero element Ω_{ij}.

3 Analysis and selection of regressions

Variable selection can be achieved through variable selection priors and forward/backward selection. This entails fitting all models in a typical forward/backward selection fashion and than comparing posterior probabilities

4 Banerjee et al 2006

Estimate a dense Σ with ML and an ℓ_1-norm to encourage sparsity in Ω.

5 Referneces

Dobra, A., Hans, C., Jones, B., Nevins, J. R., Yao, G., & West, M. (2004). Sparse graphical models for exploring gene expression data. Journal of Multivariate Analysis, 90, 196–212. https://doi.org/10.1016/j.jmva.2004.02.009