Penalized Likelihood Method for Estimation of Directed Acyclic Graphs

Ali Shojaie

Joint work with George Michailidis

Department of Statistics, University of Michigan

May 16, 2009

Outline

The setting

- Graphical models
- Penalized likelihoods
- Some asymptotics

Outline

The setting

- Graphical models
- Penalized likelihoods
- Some asymptotics

The difference ...

Outline

The setting

- Graphical models
- Penalized likelihoods
- Some asymptotics

The difference ...

- Directed graphs (DAGs) vs undirected graphs
- Network structure
- Small n, large p asymptotics

Directed vs. Undirected Graphs

- Undirected Graphs represent conditional (in)dependence structure
- For Gaussian random variables, equivalent to finding zeros of the precision matrix
- Many recent algorithms developed

Directed vs. Undirected Graphs

- Undirected Graphs represent conditional (in)dependence structure
- For Gaussian random variables, equivalent to finding zeros of the precision matrix
- Many recent algorithms developed
- Directed Graphs often represent causal relationship: Causality vs Correlation!
- Zeros of the precision matrix do not correspond to conditional independence
- NP-hard, few "efficient" algorithms available

Penalized Likelihood Estimation of DAGs

- When data has (known) natural ordering, (i.e. direction known) estimation of DAGs reduces to estimation of network structure
- Write the likelihood as a function of adjacency matrix of the graph
- Develop efficient algorithms to estimate the structure of the network

Lasso vs. Adaptive Lasso

- when $p \gg n$
- **Lasso** not consistent for variable selection unless incoherence (neighborhood stability) exists
- **Adaptive Lasso** with initial weights derived from lasso, consistent for variable selection, even without incoherence

Thank You!