Midterm Presentation:

Risk Properties in Bandable Precision Matrix Estimation

Addison Hu
Statistics 490
01 March 2017

Outline

1. Refresher on Graphical Models & Multivariate Gaussian
2. Pairwise Inference for Entrywise Recovery of Σ^{-1}
3. Risk Bounds for Entrywise Recovery in $\|\cdot\|_\infty$
4. Next Steps

Refresher

Graphical Models

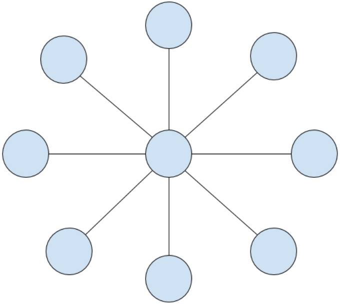
- Graphical models provide a framework within which to consider dependence structure within a group of variables.
- In doing so, we may relax the i.i.d. assumption and still perform inference feasibly.
- Examples:
  - Facebook users graph
  - Gene interaction networks

Markov Random Fields

- Consider a graph $G = (V, E)$, and a corresponding set of random variables $\{X_i\}_{i=1}^{|V|}$, where the random variables are indexed by $u \in V$.
- **Pairwise Markov property:** $X_u \perp\!\!\!\perp X_v | X_{V \setminus \{u,v\}}$ for any two non-adjacency nodes u, v.
- **Local Markov property:** $X_u \perp\!\!\!\perp X_{V \setminus \mathrm{cl}(u)} | X_{\mathrm{nb}(u)}$ for any node v.
- **Global Markov property:** $X_A \perp\!\!\!\perp X_B | X_S$ for disjoint $A, B \subset V$, and a separating subset S.
- Inference is easy when the edges are known; but is more interesting when they are unknown.

Example: Hub and Spoke Model

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$



Multivariate Gaussian

Suppose $X \sim \mathcal{N}(\mu, \Sigma)$. Its density function is given by:

$$p(\mathbf{x}) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu) \right\}$$

- Closure properties:
  - Sum of independent Gaussian random variables is Gaussian.
  - Marginal of a joint Gaussian distribution is Gaussian.
  - Condition of a joint Gaussian distribution is Gaussian.
- The sparsity pattern of Σ^{-1} concides with the adjacency matrix of the associated MRF.

Multivariate Gaussian, cont.

- Closure under marginalization: Suppose $A \subset V$. Then

$$\Sigma_A = \left(\Sigma_{ij}\right)_{i \in A, j \in A}$$

- Closure under conditioning: Suppose $A, B \subset V$, $A \cup B = V$, $A \cap B = \varnothing$. Then:

$$\left(\Omega_A\right)^{-1} = \Sigma_{A|B}$$
$$\left(\Sigma_A\right)^{-1} = \Omega_{A|B}$$

Precision Matrix Estimation

Maximum Likelihood Estimation

Assume $\mu = 0$. Then the maximum likelihood estimation problem is:

$$\begin{aligned} \underset{\Sigma}{\text{maximize}} \quad & -\log\det|\Sigma| - \langle \hat{\Sigma}, \Sigma^{-1} \rangle \\ \text{subject to} \quad & \Sigma \succeq 0 \end{aligned}$$

- Maximum Likelihood Estimate given by $\hat{\Sigma} = \frac{1}{n}\mathbf{X}^\top\mathbf{X}$.
- Idea: $\hat{\Omega} = \hat{\Sigma}^{-1}$.
- Issues:
  - Invertibility & Conditioning
  - Noise & Sparsity

Graphical Lasso

Asymptotic Normal Thresholding (ANT)

Risk Bounds in $\|\cdot\|_\infty$

Risk Upper Bound

Theorem. *Lorem ipsum.*

Oracle Inequalities

Coupling Argument

Risk Lower Bound

Le Cam's Two-Point Argument

Next Steps