Minimax Localization of Structural Information in Large Noisy Matrices

Poster: W055

M. Kolar

S. Balakrishnan  A. Rinaldo  A. Singh
**Goal:** De-noise and re-order rows/columns of the matrix to infer biclusters that are activated.

**Observation model**

\[ A = \beta uv^T + R \]

- \( u \) – \( k_1 \) sparse unit vector
- \( v \) – \( k_2 \) sparse unit vector
- \( u, v \propto \{-1, 0, 1\} \)

\( R \sim \text{i.i.d. zero-mean subgaussian}(\sigma^2) \) perturbation
Identifying biclusters

**Information Theoretic minimax limit:** If

\[
\frac{\beta}{\sigma} \sim \sqrt{\frac{k_1 k_2 \log(n_1 n_2)}{\min(k_1, k_2)}}
\]

then, for any biclustering procedure, the probability of failure remains bounded away from zero by a constant.

**Note:**

Optimal performance achieved by scanning over all submatrices of size \(k_1 \times k_2\).
Computationally efficient procedures

Elementwise thresholding

Sparse Singular Value Decomposition

Row/Column Averaging (large clusters only $k \sim n^{1/2+\alpha}$)

$$\frac{\beta}{\sigma} \sim \sqrt{k_1 k_2 \log(n_1 n_2)}$$

$$\frac{\beta}{\sigma} \sim \frac{\sqrt{k_1 k_2 \log(n_1 n_2)}}{\min(n_1^\alpha, n_2^\alpha)}$$

Note:

These procedures do not achieve information theoretic lower bound.