Hierarchical Latent Space Models for Social Network Analysis

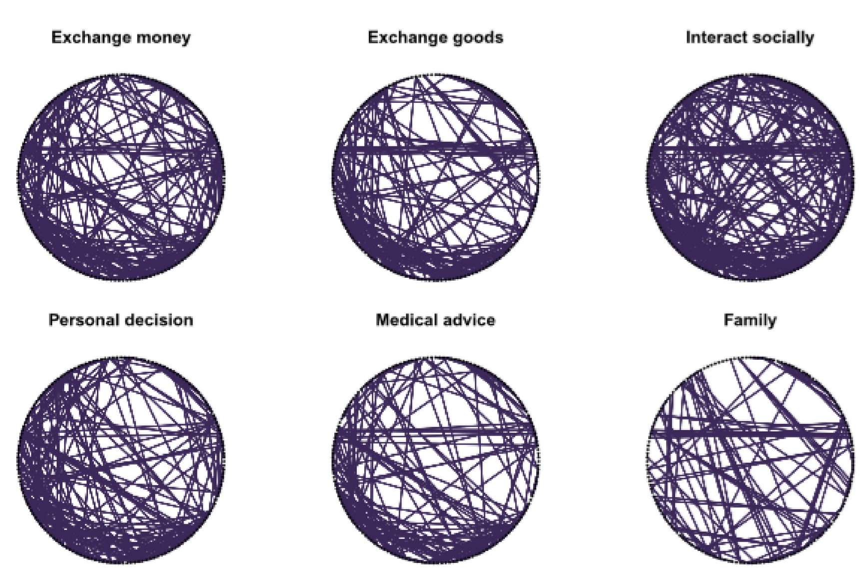
Alex Loewi[1,4], Francisco Ralston[2], Octavio Mesner[2,3,4]

[1]Public Policy, [2]Engineering & Public Policy, [3]Statistics & Data Science, [4]Machine Learning

Carnegie Mellon University

Introduction

Multigraphs: "Parallel" network models sharing identical nodes but different edges connecting nodes, used to model data where actors (nodes) maintain varying types of relationships.
Social Network Analysis (SNA), where nodes represent individual people and edges represent differing relationships, such as friendship, professional colleague, or facebook connected, for example.
Problem: Social science and network analysis rely on intuitive and interpretable models. Interpreting a graph with many nodes and differing, but related edges, is challenging.



Solution: Using a LASSO penalty, we aim to collapse sufficiently similar graphs subject to a tuning parameter.

Employee Relationships Data

Recently conducted study at a charter school in West Baltimore
Collecteda five-layer valued network between the approximately eighty employees of the school:

- Frequency of interaction
- Discussing Academics
- Discussing Behavior
- Social interaction
- Professionally Helpful Relationship

While empirically these layers are found to have very high correlations, modeling them in a way that demonstrates those correlations is challenging.

Latent Space Projection

- Map nodes, $i \mapsto z_i \in \mathbb{R}^n$ where proximity, $\|z_i - z_j\|_2 = d_{ij} < 1$, indicates nodes are connected (and not connected otherwise)
- Makes reciprocity ($j \to i \Rightarrow i \to j$) and transitivity ($i \to j, j \to k \Rightarrow i \to k$) probable
- Given $\sigma_{ij} = \mathbb{P}(Y_{ij} = 1 | z_i, z_j) = \text{logit}^{-1}\left(\alpha + \|z_i - z_j\|_2^2\right)$ where $Y_{ij} = 1$ indicates i, j are connected in the data, find z
- Likelihood:

$$\prod_{i<j} \sigma_{ij}^{y_{ij}} (1 - \sigma_{ij})^{1-y_{ij}}$$

Hierarchical Models

Goal: Collapse similar networks
We include a group-wise LASSO penalty on the log-likelihood:

$$\sum_k \sum_{i<j} [y_{ijk} \ln \sigma_{ijk} + (1 - y_{ijk}) \ln(1 - \sigma_{ijk})] + \lambda \sum_i \sum_k \|\epsilon_{ik}\|_1$$

where $z_{ik} = b_i + \epsilon_{ik}$ for initialization points, b_i

Optimization Approaches:

1. Proximal Gradient Descent
2. Coordinate-Wise Optimization (pending)
3. Hamiltonian Monte Carlo methods

Takeaway

Convex methods have valuable contributions for fitting complex Latent Space Models, and their extensions

Model's Advantages

Our model has the additional inclusion of a hierarchy on the latent variables:

$$z_{ik} = b_i + \epsilon_{ik}$$

each node has a separate latent position to explain its ties within a single layer of the multigraph, but these multiple positions are all tied together by a single "base" position for that node b_i, and a layer-specific perturbation, ϵ_{ik}.
This allows two major advantages. One is graph-wise regularization with a group lasso penalty, allowing entire redundant layers to be removed for meaningful dimensionality reduction. Second, regularization allows a manual tuning between layouts that are expressive, and those that are comparable, solving the difficult problem of visual comparability.

Results

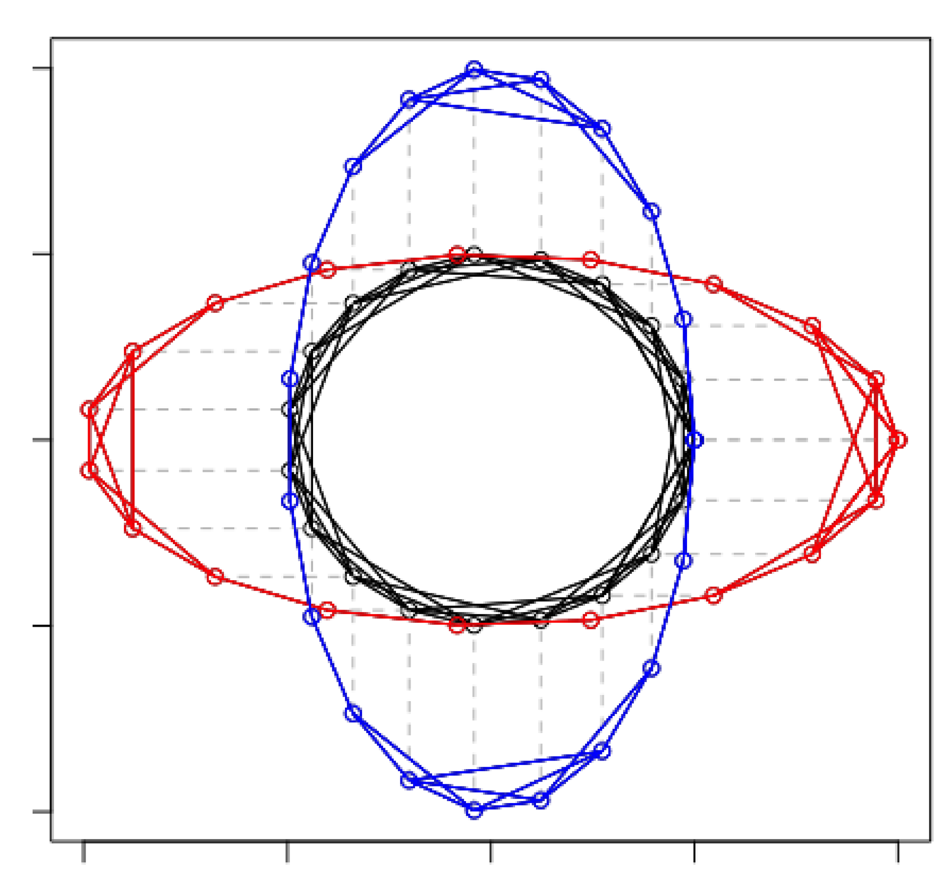


Figure 1: Figure caption

Challenges

- Negative log-likelihood is not convex in the positions
  Fix: We started from various random initializations
- Distances between positions will yield same likelihood under translations and rotation
  Fix: Fit each layer of the hierarchy separately then use linear transformations to align each to initialize the model

Conclusion

We found that proximal gradient methods are an effective tool for refining an HLSM fit, given a good starting point.

References

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. *Latent space approaches to social network analysis.* Journal of the american Statistical association 97.460 (2002): 1090-1098.
Salter-Townshend, Michael, and Tyler H. McCormick. *Latent space models for multiview network data.* Technical Report 622, Department of Statistics, University of Washington, 2013.
Salter-Townshend. M, *Personal communication*, 2017
Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. *Stan: A probabilistic programming language.* Journal of Statistical Software 76(1). DOI 10.18637/jss.v076.i01, 2017
Loewi, A., *Parent-Teacher Relationships and Student Outcomes*, 2017

Acknowledgements

The authors would like to thank the Network Analysis group in the Department of Statistics and Data Science for their invitation to present, and the helpful feedback they contributed.