HEART: Statistics and Data Science With Networks

Joshua Agterberg

Johns Hopkins University

Fall 2021

Outline

Outline

Multilayer Graphs

- Observe L graphs ($L =$ number of layers)

Multilayer Graphs

- Observe L graphs ($L =$ number of layers)
- Assume that graphs share probability structure

Multilayer Graphs

- Observe L graphs ($L =$ number of layers)
- Assume that graphs share probability structure
- Usually assume graphs have the same set of nodes

Multilayer Graphs

- Observe L graphs ($L =$ number of layers)
- Assume that graphs share probability structure
- Usually assume graphs have the same set of nodes
- Easy example: friendships evolving each year

Multilayer Stochastic Blockmodels

- Typically assume community structure is the same across graphs

Multilayer Stochastic Blockmodels

- Typically assume community structure is the same across graphs
- May allow the probabilities to change (i.e have a layer-specific matrix $B^{(l)}$

Multilayer Stochastic Blockmodels

- Typically assume community structure is the same across graphs
- May allow the probabilities to change (i.e have a layer-specific matrix $B^{(l)}$
- How to perform spectral clustering?

Multilayer Stochastic Blockmodel

Could just use the average adjacency matrix:

$$\bar{A} := \frac{1}{L} \sum_{l=1}^{L} A^{(l)}.$$

Multilayer Stochastic Blockmodel

Could just use the average adjacency matrix:

$$\bar{A} := \frac{1}{L} \sum_{l=1}^{L} A^{(l)}.$$

but there's a problem! Example: $L = 2$,

$$B^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad B^{(2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Multilayer Stochastic Blockmodel

Could just use the average adjacency matrix:

$$\bar{A} := \frac{1}{L}\sum_{l=1}^{L} A^{(l)}.$$

but there's a problem! Example: $L = 2$,

$$B^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad B^{(2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then averaging results in the matrix

$$\bar{B} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

which is just Erdos-Renyi! So $\bar{A}$ loses information.

Multilayer Stochastic Blockmodel

Could just use the average adjacency matrix:

$$\bar{A} := \frac{1}{L}\sum_{l=1}^{L} A^{(l)}.$$

but there's a problem! Example: $L = 2$,

$$B^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad B^{(2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then averaging results in the matrix

$$\bar{B} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

which is just Erdos-Renyi! So $\bar{A}$ loses information.
Ways around this: Paper 1, Paper 2, among others.

Multilayer Random Graphs beyond Stochastic Blockmodels

- COSIE Model
- Multilayer RDPG
- Mixture Multilayer SBM
- Dimple
- Multilayer Networks with Vertex Community Flips
- More General Models

Computing Graph Embeddings from Multiple Graphs

- Omnibus Embedding (L is not too big)
  - Paper 1
  - Paper 2
  - Paper 3

Computing Graph Embeddings from Multiple Graphs

- Omnibus Embedding (L is not too big)
  - Paper 1
  - Paper 2
  - Paper 3
- What if L is big?
  - Tensor Methods
  - Bias-Adjusted Spectral Embedding
  - MASE (parallelizable)

Practical Considerations

- Typically want to obtain a single vertex representation for all graphs
- If L and n are both big, may not want to do complicated techniques for computing reasons
- Sometimes L is small, so omnibus embedding is not too bad
- Sometimes graphs are sparse, so need to design algorithms that allow for very sparse graphs