Conditionally Independent Dyadic Models for Complex Networks

Hierarchical Networks Group

Department of Statistics, Carnegie Mellon University

April 18, 2015

Grant Support Provided By: IES R305D120004, NSF MMS-1129271
A completely specified network of $n$ nodes has $\binom{n}{2}$ dyads, each with an outcome of interest:

<table>
<thead>
<tr>
<th>Node 1</th>
<th>Node 2</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>76</td>
<td>0</td>
</tr>
</tbody>
</table>
Conditionally Independent Dyadic Models

Premise: We model the underlying strength of each tie (dyad) as a Gaussian random variable, such as

\[ Z_{ij} = \mu + X_{ij} \beta + U_{ij} + \varepsilon_{ij}. \]

Realization: The outcome \( Y_{ij} \) is a function of this. Gaussian is the identity. Binary:

\[ Y_{ij} = \mathbb{I}(Z_{ij} > 0) \]

Ordinal data with \( K > 2 \) categories:

\[ Y_{ij} = \mathbb{I}(Z_{ij} > 0) + \sum_{k=1}^{K-2} \mathbb{I}(Z_{ij} > C_k) \]

Other partitions are possible (and so are multivariate versions). Interesting parts: What comprises \( U_{ij} \)?
To directly specify edge covariates, we can use the standard specification:

\[ U_{ij} = X_{ij}/\beta \]

We have a few varieties available:

- **EdgeCOV**: directly specified for each edge.
- **SenderCOV**: each row is a node, each column is a variable on the node; we process this into an edge-covariate matrix given the sender.
- **ReceiverCOV**: each row is a node, each column is a variable on the node; we process this into an edge-covariate matrix given the receiver.
- **SendRecCOV**: combines the two above.
- **IdenticalCOV**: each row is a node, each column is a variable. We make a new edge-covariate matrix that identifies if two node have the same value.
BETA(): “Beta” Model for Sender-Receiver Strength

\[ U_{ij} = \beta_i + \beta_j \]
LSM(): Latent Space Model (Hoff, Raftery and Handcock 2002)

\[ U_{ij} = -|d_i - d_j| \]
$U_{ij} = d_i' d_j$
$U_{ij} = S_i BR_j$
MMSBM(): Mixed Membership Stochastic Block Model (Airoldi et al 2008)

\[ U_{ij} = S_{ij} BR_{ji} \]

\[ S_{ij} \sim Mult(\pi_i) \]
National Longitudinal Study on Adolescent Health ("Add Health"): Extensive studies of social network data in hundreds of schools.

Public versions: fake generated networks consistent with schools in the study. "Mesa" in the Southwest; "Magnolia" in the South.

Originally in the R package "ergm"; reprocessed as edge and node data frames for CIDnetworks.
Examples in Education
Example: let’s try a block model for the fit.

\[ Z_{ij} = \mu + SBM(4) + \varepsilon_{ij} \]
Examples in Education

Mesa High Network
Example: let’s try a latent space model for the fit.

\[ Z_{ij} = \mu + LSM(2) + \epsilon_{ij} \]
Examples in Education

Mean Latent Space Positions from Gibbs Sampler