36-780: Social Network Modeling

QAP, ERGM, Graphical Models, and Networks

Brian Junker
132E Baker Hall
brian@stat.cmu.edu

Outline

- A few words about QAP
  - Simulation: Permutation tests
- ERGMs
  - An explicit computation
  - ERGMs for valued ties
  - Goodness of fit
    - Simulation: parametric bootstrap
- Graphical models vs social network models
- D-separation, identifiability
- HW03: Read Shalizi & Thomas for Thursday!

QAP Regression

- Let Y = $[Y_{ij}]$ be an $n \times n$ response matrix
  - An adjacency matrix, or
  - A matrix of edge values
- Let $X^{(1)}, X^{(2)}, \ldots, X^{(k)}$ be k $n \times n$ matrices of predictors.
- We can easily imagine

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 X_{ij}^{(1)} + \cdots + \beta_k X_{ij}^{(k)} + \epsilon_{ij} \quad \text{, or} \\ \text{logit}\, P[Y_{ij} = 1] &= \beta_0 + \beta_1 X_{ij}^{(1)} + \cdots + \beta_k X_{ij}^{(k)} + \epsilon_{ij} \end{aligned}$$

Etc.

How can we do inference?

- Typically, the β's are estimates with OLS for

$$Y_{ij} = \beta_0 + \beta_1 X_{ij}^{(1)} + \cdots + \beta_k X_{ij}^{(k)} + \epsilon_{ij}$$

and estimated with the usual IRWLS algorithm for

$$\text{logit}\, P[Y_{ij} = 1] = \beta_0 + \beta_1 X_{ij}^{(1)} + \cdots + \beta_k X_{ij}^{(k)} + \epsilon_{ij}$$

- These methods assume Y_{ij}'s are independent – probably not!
  - If you talk to Mauricio and I talk to Mauricio, then chances are you and I talk to each other too!
  - Another *pseudo-likelihood method* (like old ERGM fitting*)
  - Need a way of estimating standard errors, p-values, …

*Strauss & Ikeda (1990)

QAP inference: random permutation null distribution

- We need a "null" distribution for the β's under H_0*: No relationship between the X's and Y*
- If we randomly permute the order of the graph nodes in Y but keep the weights between the permuted dyads the same, this should be H_0.
- Algorithm: repeat (100, or 1000, or *m*, times):
  - Randomly permute Y
  - Re-run the regression
  - Write down the estimated β's

QAP comments

- QAP = *Quadratic Assignment Procedure*
  - Krackhardt, David. 1987. "QAP Partialling as a Test of Spuriousness." *Social Networks* 9: 171-186.
  - Krackhardt, David. 1988. "Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data." *Social Networks* 10: 359-381.
- It's not always clear (to me!) that H_0*: "no relationship at all"* is the right null to simulate from
  - There appear to be several possible null distributions to choose from in "netlm" and "netlogit" (from the "statnet" package(s) [ergm/network/sna…])
- David Krackhardt will give guest lecture later in mini

ERGM for binary ties

- Recall that the general form of an ERGM is

$$P(Y = y) = \frac{\exp\{\vec{\theta}^{\mathsf{T}}\vec{g}(y)\}}{\kappa(\vec{\theta},\mathcal{Y})}$$

  - $\vec{\theta}$ is a vector of parameters
  - $\vec{g}(y)$ a vector of "network statistics"
  - $\kappa(\vec{\theta},\mathcal{Y})$ is a normalizing constant, i.e.

$$\kappa(\vec{\theta},\mathcal{Y}) = \sum_{y\in\mathcal{Y}} \exp\{\vec{\theta}^{\mathsf{T}}\vec{g}(y)\}$$

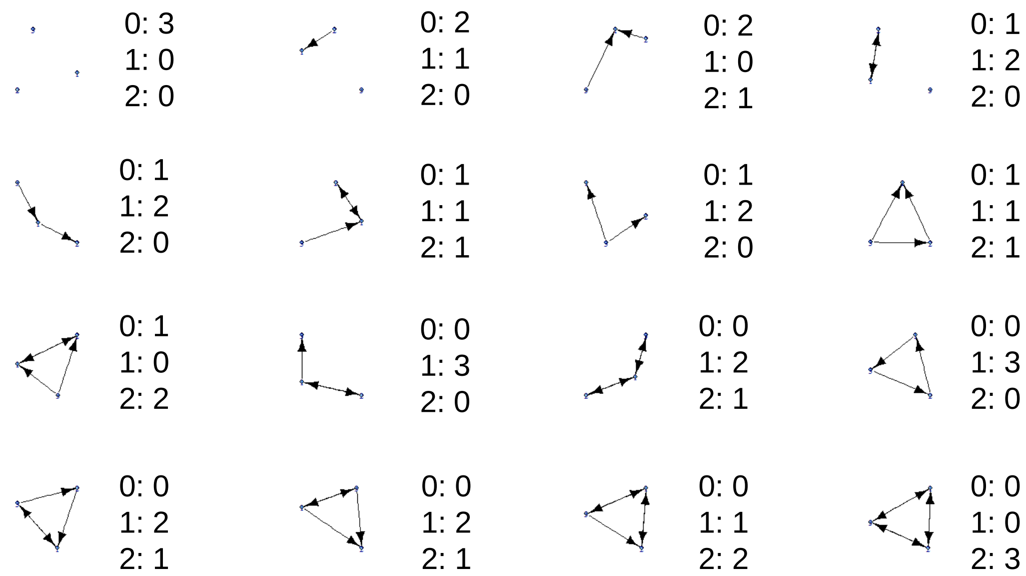
where $\mathcal{Y}$ is the set of all graphs on the vertices of *y*.

ERGMs: an actual model

- Suppose we take the model `Y ~ indegree` from the `ergm()` function in R.
  - *"This term adds one network statistic to the model for each element in d; the ith such statistic equals the number of nodes in the network of in-degree d[i]"*
- Let's see what this looks like on a network with three nodes…

d_0(y): # of nodes with in-degree = 0
d_1(y): # of nodes with in-degree = 1
d_2(y): # of nodes with in-degree = 2

In-degree statistics for 3 nodes…




```
for (i in 0:15) plot(graph.isocreate(3,i),edge.col="black")
```


…so the ERGM model will look like:

- Numerator:

$$\exp\{\theta_0 d_0(y) + \theta_1 d_1(y) + \theta_2 d_2(y)\}$$

- Denominator:

$$\begin{aligned}
\kappa(\vec{\theta}, \mathcal{Y}) \;=\; & \\
& \exp\{\theta_0 3 + \theta_1 0 + \theta_2 0\} + \exp\{\theta_0 2 + \theta_1 1 + \theta_2 0\} + \exp\{\theta_0 2 + \theta_1 0 + \theta_2 1\} + \\
& \exp\{\theta_0 1 + \theta_1 2 + \theta_2 0\} + \exp\{\theta_0 1 + \theta_1 2 + \theta_2 0\} + \exp\{\theta_0 1 + \theta_1 1 + \theta_2 1\} + \\
& \exp\{\theta_0 1 + \theta_1 2 + \theta_2 0\} + \exp\{\theta_0 1 + \theta_1 1 + \theta_2 1\} + \exp\{\theta_0 1 + \theta_1 0 + \theta_2 2\} + \\
& \exp\{\theta_0 3 + \theta_1 3 + \theta_2 0\} + \exp\{\theta_0 0 + \theta_1 2 + \theta_2 1\} + \exp\{\theta_0 0 + \theta_1 3 + \theta_2 0\} + \\
& \exp\{\theta_0 0 + \theta_1 1 + \theta_2 2\} + \exp\{\theta_0 3 + \theta_1 2 + \theta_2 1\} + \exp\{\theta_0 3 + \theta_1 1 + \theta_2 2\} + \\
& \exp\{\theta_0 0 + \theta_1 0 + \theta_2 3\}
\end{aligned}$$

Computing κ…

- Clearly for many nodes or many network statistics, this becomes unwieldy…
- As sketched in Hunter et al (2008), Snijders (2002) applied a trick of Geyer & Thompson (1992):

$$\begin{aligned}
\ell(\vec{\theta}) - \ell(\vec{\theta}_0) &= (\vec{\theta} - \vec{\theta}_0)^{\mathsf{T}} \vec{g}(y) - \log \left[\frac{\kappa(\vec{\theta}, \mathcal{Y})}{\kappa(\vec{\theta}_0, \mathcal{Y})} \right] \\
&= (\vec{\theta} - \vec{\theta}_0)^{\mathsf{T}} \vec{g}(y) - \log E_{Y|\vec{\theta}_0} \left[\exp\{(\vec{\theta} - \vec{\theta}_0)^{\mathsf{T}} \vec{g}(Y)\} \right] \\
&\approx (\vec{\theta} - \vec{\theta}_0)^{\mathsf{T}} \vec{g}(y) - \log \left[\frac{1}{M} \sum_{m=1}^{M} \exp\{(\vec{\theta} - \vec{\theta}_0)^{\mathsf{T}} \vec{g}(Y_m)\} \right]
\end{aligned}$$

where $Y_m \sim p(Y=y|\theta_0)$ from MCMC…

ERGM for valued ties

- What if y = $[y_{ij}]$ is a matrix of edge values (amount of money paid from i to j, or number of advice-seeking overtures, etc.)?
- In principle,

$$P(Y = y) = \frac{\exp\{\vec{\theta}^{\mathsf{T}} \vec{g}(y)\}}{\kappa(\vec{\theta}, \mathcal{Y})}$$

is just an exponential family model, so should generalize to arbirary y, with

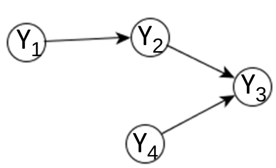
$$\kappa(\vec{\theta}, \mathcal{Y}) = \int_{\mathcal{Y}} \exp\{\vec{\theta}^{\mathsf{T}} \vec{g}(y)\} d\mu(y)$$

ERGM for valued ties…

- In practice, it is difficult to specify and compute
  - Again, mostly $\kappa(\vec{\theta}, \mathcal{Y})$ is the culprit
- Krivitsky (2012) has worked out some general theory
- Krivitsky also developed an R package for the specific case of counted ties, `ergm.count`

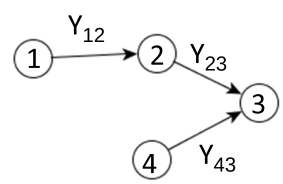
Graphical Models vs Social Network Models

- Graphical Model:



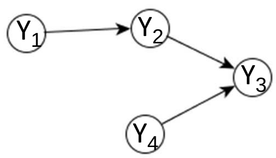
- *Nodes are random* variables
- *Edges are fixed* entities expressing conditional independence (non-dyadic!)

- Social Network Model:



- *Nodes are fixed* entities
- *Edges are random* variables expressing some dyadic relationship

A fast introduction/review for graphical models



- Motivated from path models/SEMs:

$$
\begin{aligned}
Y_2 &= \beta_0^{12} + \beta_1^{12} Y_1 + \epsilon_{12} \\
Y_3 &= \beta_0^{243} + \beta_1^{243} Y_2 + \beta_2^{243} Y_4 + \epsilon_{243}
\end{aligned}
$$

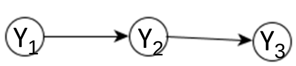
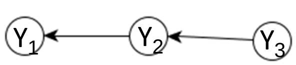
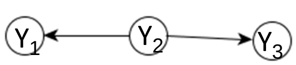
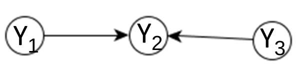
- But for any kind of Y's
  - Can write joint model from implied conditionals (like Bayes nets)

$$
\begin{aligned}
p(y_1, y_2, y_3, y_4) &= p(y_3|y_2, y_4) \times \\
&\quad p(y_2|y_1)\ p(y_1)\ p(y_4)
\end{aligned}
$$

  - Can "read off" conditional independencies
    - $Y_1 \amalg (Y_3, Y_4) | Y_2$
  - More generally, have *graphical Markov property*

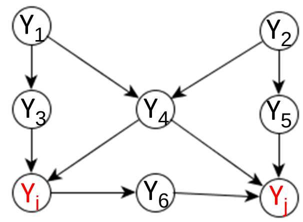
Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669-688.

Conditional independence and d-separation in graphical models

	Unconditional:	Conditional on Y_2:	
	$Y_1 \not\amalg Y_3$	$Y_1 \amalg Y_3 \| Y_2$	chain
	$Y_1 \not\amalg Y_3$	$Y_1 \amalg Y_3 \| Y_2$	chain
	$Y_1 \not\amalg Y_3$	$Y_1 \amalg Y_3 \| Y_2$	common cause
	$Y_1 \amalg Y_3$	$Y_1 \not\amalg Y_3 \| Y_2$	collider, common effect

Let A, B, C be disjoint sets of nodes; and let p be an undirected path from A to B. *C blocks p*, if (i) there exists a collider y on p with neither y nor its descendants in C; or (ii) there is a non-collider on p in C. *C d-separates A from B* if it blocks every path between them. This occurs if and only if, $A \amalg B \mid C$.

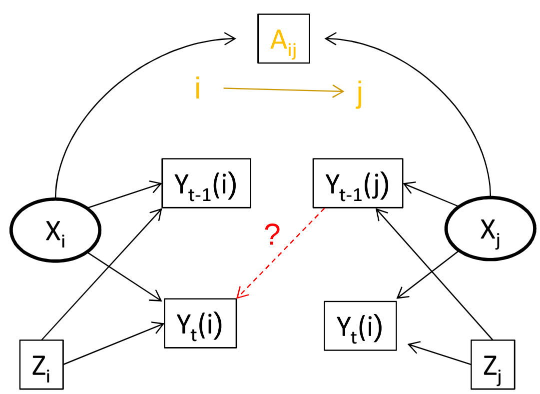
Backdoor paths and estimating the effect of Y_i on Y_j



- Let (Y_i, Y_j) be an ordered pair. A set of nodes C satisfies the *backdoor criterion* for Y_i and Y_j if
  - No node of C is a descendant of Y_i; and
  - C blocks every path with an arrow into Y_i
- Conditioning on C makes the effect of Y_i on Y_j identifiable.
- C={Y_3, Y_4} or C={Y_4, Y_5} satisfy the backdoor criterion
- C = {Y_4} does not, since Y_4 does not block path Y_i, Y_3, Y_1, Y_4, Y_2, Y_5, Y_j
- A backdoor path is a confounding path

Shalizi & Thomas's main result

- Consider just one edge A_{ij} in a social network
- Consider measured outcomes on the nodes at two different times
  - $Y_{t-1}(i)$, $Y_{t-1}(j)$
  - $Y_t(i)$, $Y_t(j)$
- When can we estimate effect of $Y_{t-1}(j)$ on $Y_t(i)$?



- X_i, X_i are unobserved (latent) features of i and j
- Z_i and Z_i are observed covariates on i and j

Summary

- A few words about QAP
  - Simulation: Permutation tests
- ERGMs
  - An explicit computation
  - ERGMs for valued ties
  - Goodness of fit
    - Simulation: parametric bootstrap
- Graphical models vs social network models
- D-separation, identifiability
- HW03: Read Shalizi & Thomas for Thursday!