Generalized Exponential Random Graph Models: Inference for Weighted Graphs

James D. Wilson
University of North Carolina at Chapel Hill
June 18th, 2015

Political Networks, 2015
1. Setting the Stage: ERGMs for Unweighted Graphs
2. Generalized Exponential Random Graphs (GERGMs)
   - Model
   - Inference via MCMC
   - Applications and Degeneracy Case Study
3. Tutorial on `gergm.R` Software
4. Using `gergm.R` in person
Networks / Graphs

- Means to visualize, model, and analyze interactions of a complex system
- Treat an “actor” of the system as a vertex and place edges between actors that interact
Setting:

- Network $x$ composed of $n$ nodes and $M$ edges
- Assume binary edges between nodes $i, j$: $x_{i,j} \in \{0, 1\}$
- $X$: family of all graphs with $n$ nodes and binary edges
- $\mathbb{P}_X$: probability measure on $X$

Aim: Identify (and then estimate) the relational covariates that capture network structure through $\mathbb{P}_X$
Exponential Random Graph Models (ERGMs)

The probability (likelihood) of observing network $x$:

$$P_X(x, \theta) = \frac{\exp(\theta' h(x))}{\sum_{z \in X} \exp(\theta' h(z))}, \quad x \in \{0, 1\}^M$$

- $h : \{0, 1\}^M \rightarrow \mathbb{R}^p$: Network covariates
- $\theta \in \mathbb{R}^p$: unknown parameters (we have to estimate these!)
ERGM: Covariate Specification

- **Endogeneous**: network-level
- **Exogeneous**: relational data-driven covariates
ERGMs: Food for Thought

- Model selection
- Goodness of fit diagnostics
- Computational tools for estimation
  - Markov Chain Monte-Carlo estimation
  - Maximum pseudo-likelihood estimation
- Likelihood degeneracy: how can we avoid trivial models?
Weighted Graphs

Setting:

- Network $y$ composed of $n$ nodes and $M$ edges
- Edges are continuous valued between nodes $i, j$: $y_{i,j} \in (-\infty, \infty)$
- $Y$: family of all graphs with $n$ nodes
- $\mathbb{P}_Y$: probability measure on $Y$

Examples:

- Correlation networks
- Financial Lending Networks
- Voting Networks
- Brain Networks
Inference for weighted networks

GERGM:
Generative exponential model for $y$ that incorporates

- network covariates
- population edge-level covariates

References:

Generalized Exponential Random Graph Model

Probability measure on family of weighted graphs w/ $n$ nodes, $M$ edges

**Model:** Two Steps

1) Model joint structure of $Y$ on restricted network $X \in [0, 1]^M$:

$$f_X(x, \theta) = \frac{\exp (\theta' h(x))}{\int_{[0,1]^m} \exp (\theta' h(z)) \, dz}, \quad x \in [0, 1]^M$$

- $h : [0, 1]^M \to \mathbb{R}^p$: Network covariates (endogeneous or exogenous)
- $\theta \in \mathbb{R}^p$: unknown structural parameters
2) Transform onto space of continuous weights:

\[ f_Y(y, \theta, \Lambda) = \frac{\exp (\theta' h(T(y, \Lambda)))}{\int_{[0,1]^m} \exp (\theta' h(z)) \, dz} \prod_{ij} t_{ij}(y, \Lambda), \quad y \in \mathbb{R}^M \]

- \( T : \mathbb{R}^M \to [0, 1]^m \): parametric transformation function
  - Monotonically increasing
  - Most obvious choice: cumulative distribution functions
- \( \Lambda \in \mathbb{R}^q \): unknown transformation parameters
Features of the GERGM

- Flexible model for any type of weighted network

- Like ERGM, requires computationally efficient algorithms for estimation like Markov Chain Monte-Carlo or Maximum pseudo-likelihood

- Reduces to classical regression when $h(x) = 0$. 
## Model Specification: Weighted Network Statistics

<table>
<thead>
<tr>
<th>Network Statistic</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reciprocity</td>
<td>$\theta_R$</td>
<td>$\left( \sum_{i&lt;j} x_{ij} x_{ji} \right)^{\alpha_R}$</td>
</tr>
<tr>
<td>Cyclic Triads</td>
<td>$\theta_{CT}$</td>
<td>$\left( \sum_{i&lt;j&lt;k} (x_{ij} x_{jk} x_{ki} + x_{ik} x_{kj} x_{ji}) \right)^{\alpha_{CT}}$</td>
</tr>
</tbody>
</table>
| In-Two-Stars      | $\theta_{ITS}$ | $\left( \sum_{i} \sum_{j<k 
eq i} x_{ji} x_{ki} \right)^{\alpha_{ITS}}$ |
| Out-Two-Stars     | $\theta_{OTS}$ | $\left( \sum_{i} \sum_{j<k 
eq i} x_{ij} x_{ik} \right)^{\alpha_{OTS}}$ |
| Edge Density      | $\theta_E$ | $\left( \sum_{i 
eq j} x_{ij} \right)^{\alpha_E}$ |
| Transitive Triads | $\theta_{TT}$ | $\left( \sum_{i<j<k} (x_{ij} x_{jk} x_{ik} + x_{ij} x_{kj} x_{ji} + x_{ij} x_{kj} x_{ki}) \right)^{\alpha_{TT}} + \sum_{i<j<k} (x_{ji} x_{jk} x_{ki} + x_{ji} x_{jk} x_{ik} + x_{ji} x_{kj} x_{ki})^{\alpha_{TT}}$ |
Likelihood of GERGM is intractable; relies on MCMC

**Gibbs** *(Desmarais, et al., 2012)*

**Major Issue**: Restricts model specification by requiring first order network statistics

\[
\frac{\partial^2 h(x)}{\partial x_{ij}^2} = 0, \quad i, j \in [n]
\]

**Metropolis-Hastings** *(Wilson, et al. 2015)*: Removes above restriction on GERGM
**Framework**: Acceptance/Rejection algorithm for weighted edges w/ multivariate truncated normal proposal distribution.

**Proposal**: 
\[
q_\sigma(w|x) = \frac{\sigma^{-1} \phi\left(\frac{w-x}{\sigma}\right)}{\Phi\left(\frac{1-x}{\sigma}\right) - \Phi\left(\frac{-x}{\sigma}\right)}, \quad 0 \leq w \leq 1
\]

**Advantages:**
- Flexible model specification
  - Interaction effects
  - Exponential weighting of covariates
- New available models can avoid likelihood degeneracy
For $i, j \in [n]$, generate proposal edge $\tilde{x}_{i,j}^{(t)} \sim q_\sigma(\cdot|x_{i,j}^{(t)})$ independently across edges.

Set

$$x^{(t+1)} = \begin{cases} 
\tilde{x}^{(t)} = (\tilde{x}_{i,j}^{(t)})_{i,j \in [n]} & \text{w.p. } \rho(x^{(t)}, \tilde{x}^{(t)}) \\
x^{(t)} & \text{w.p. } 1 - \rho(x^{(t)}, \tilde{x}^{(t)})
\end{cases}$$

where

$$\rho(x, y) = \min \left( \frac{f_X(y|\theta)}{f_X(x|\theta)} \prod_{i=1}^{m} \frac{q_\sigma(x_i|y_i)}{q_\sigma(y_i|x_i)}, 1 \right)$$

$$= \min \left( \exp \left( \theta' (h(y) - h(x)) \right) \prod_{i=1}^{m} \frac{q_\sigma(x_i|y_i)}{q_\sigma(y_i|x_i)}, 1 \right)$$

(1)
Case Study: In-Two-Stars Model

Model:

\[ f_X(x, \theta, \alpha) = \frac{\exp(\theta_E h_E(x) + \theta_{ITS} h_{ITS}(x))}{C(\theta_E, \theta_{ITS})}, \quad x \in [0, 1]^m \]

- Edge density: \( h_E(x) = \sum_{i \neq j} x_{ij}/m \)

- In-Two-Stars: \( h_{ITS}(x, \alpha) = \left( \sum_i \sum_{j<k \neq i} x_{ji} x_{ki} \right)^\alpha \)

When \( \alpha = 1 \), this is closely related to the well-known degenerate triangle model and the Ising model.
Case Study: In-Two-Stars Model

Degeneracy Issues:

In $\alpha = 1$ model, simulated networks tend to be nearly empty or nearly complete leading to *likelihood degeneracy*:

![Graphs showing density functions](image-url)
**Upshot:** Exponential weighting of network statistics alleviates degeneracy!
Lending volume (in millions of US $) between 18 countries
Collected by Bank of International Settlements (BIS)
Originally published in Oatley et al. (2013)
Heavy Tailed Lending volumes

- Log normalized volume for edge weights
Dynamic Summary of Covariates

- Highly transitive, not reciprocal
- Temporal trends suggest a change in lending structure
Figure: Goodness of fit evaluation of estimated model for 1980 (reciprocity and transitivity – highlighted in yellow) on the international lending network generated using 10K simulated networks from the fitted model (with $\alpha = 0.1$). Results indicate positive significant reciprocity effect.
Desmarais et al. 2012

- Describes the inter-state migration in U.S. from 2006 to 2007.
- Fit a GERGM with 5 network statistics, 11 demographic covariates
Covariate Estimates

Model
- Gibbs
- M-H
- MPLM

Variable
- Transitive Triads
- Reciprocity
- Out-Two-Stars
- In-Two-Stars
- Cyclic Triads
- Unemployment Sender
- Unemployment Receiver
- Population Sender
- Population Receiver
- Jan. Temp. Sender
- Jan. Temp. Receiver
- Intercept
- Income Sender
- Income Receiver
- Distance
- Dispersion

Coefficient
-0.2
-0.1
0.0
0.1
-60
-30
0
30
## Inference on the Migration Data

<table>
<thead>
<tr>
<th>Statistic</th>
<th>↑↓ in Migration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitivity</td>
<td>↑</td>
</tr>
<tr>
<td>Out-two-stars</td>
<td>↑</td>
</tr>
<tr>
<td>In-two-stars</td>
<td>↓</td>
</tr>
<tr>
<td>Cyclic triads</td>
<td>↓</td>
</tr>
<tr>
<td>Unemployment Send.</td>
<td>↑</td>
</tr>
<tr>
<td>Population Rec.</td>
<td>↓</td>
</tr>
<tr>
<td>January Temp. Send.</td>
<td>↓</td>
</tr>
</tbody>
</table>

Unemployment, population, and average January temperature each influence U.S. migration!
M-H, Gibbs, and MPLE have good performance
We’ll now go through two examples of how to use the GERGM in practice.

Go to http://jameswd.web.unc.edu under "Presentations" to get started.
Future Work

- Model Selection for random graph models
- Mixing Time Analysis
- Temporal Extensions
- New applications (always looking for exciting data)
Thank you

Collaborators: Skyler Cranmer, Bruce Desmarais, Matthew Denny, and Shankar Bhamidi.

Funding: National Science Foundation
- Grants: DMS-1105581, DMS-1310002, SES-1357622, SES-1357606, SES-1461493, and CISE-1320219

Contact:
- Email: jdwilson1212@gmail.com
- Place: University of San Francisco, Department of Mathematics