Iterative Estimation Approach for ERGMs with Nodal Random Effects

Sevag Kevork[1], Goeran Kauermann[1]

[1] *Department of Statistics, LMU Munich, Germany*

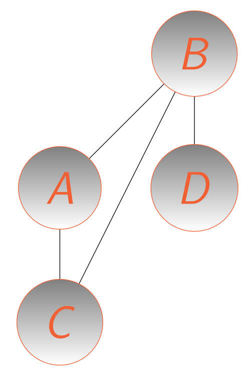
Sevag.Kevork@stat.uni-muenchen.de

Research Question

Setting

The presence of unobserved heterogeneity in Exponential Random Graph Models (ERGM) is an obvious concern. We extend the well-known Exponential Random Graph Model (ERGM) by including random effects to account for unobserved heterogeneity in the network. This leads to an ERGM with random structure on the coefficients. Estimation is carried out by combining approximate penalized pseudo-likelihood estimation for the random effects with maximum likelihood estimation for the remaining parameters in the model. This allows to fit nodal heterogeneity effects even for large scale networks.

Definition of a Network



$$= \begin{pmatrix} 0&1&1&0\\ 1&0&1&1\\ 1&1&0&0\\ 0&1&0&0 \end{pmatrix}$$

$$Y_{ij} = \begin{cases} 1, & \text{if vertex i and vertex j are connected} \\ 0, & \text{otherwise or if i = j} \end{cases}$$

where $i, j = 1, ..., N$.

$\Rightarrow$ Undirected network, this implies $Y_{ij} = Y_{ji}$.

$Y = (Y_{ij})_{i,j=1,...,N}$ denotes the network adjacency matrix.

Exponential Random Graph Model

$$P(Y = y|\theta) = \frac{exp\left(\theta^T s(y)\right)}{\kappa(\theta)},$$

Where

- where Y is a random network on n nodes.
- θ is a vector of parameters.
- $s(y)$ a known vector of so-called network statistics on y.
- $\kappa(\theta)$ is a normalizing constant, makes all the probabilities sum to 1.

Conditional "log-odds" of an edge

- $y_{ij} = 0$ or 1, depending whether there is an edge.
- y_{-ij} denotes the status of all pairs in y other than (i, j).
- y_{ij}^+ denotes the same network as y but with $y_{ij} = 1$.
- y_{ij}^- denotes the same network as y but with $y_{ij} = 0$.

$$\frac{P(Y_{ij} = 1|Y_{-ij}, \theta)}{P(Y_{ij} = 0|Y_{-ij}, \theta)} = \exp\left[\frac{\theta^T s(y_{ij}^+)}{\theta^T s(y_{ij}^-)}\right]$$

$$\Leftrightarrow \text{logit}\left[P(Y_{ij} = 1\,|Y_{-ij}, \theta)\right] = \theta^T \underbrace{\left[s(y_{ij}^+) - s(y_{ij}^-)\right]}_{=:\, s_{ij}(y)},$$

- so $\theta^T\left[s(y_{ij}^+) - s(y_{ij}^-)\right]$ is the conditional log-odds of an edge (i, j).
- where $s_{ij}(y)$ denotes the vector of change statistics.

1 Modeling process

In our approach we postulate

$$\text{logit}\left[P(Y_{ij} = 1\,|Y_{-ij}, \theta)\right] = \theta^T s_{ij}(y) + u_i + u_j,$$

with $u_i \sim N(0, \sigma_u^2)$, for $i = 1, ..., N$.
This leads to the entire model

$$P(Y = y|\theta, u) = \frac{exp\left(\theta^T s(y) + u^T t(y)\right)}{\kappa(\theta, u)},$$

where $t(y) = \left(\sum_{j\neq 1} y_{1j}, \sum_{j\neq 2} y_{2j}, \ldots, \sum_{j\neq N} y_{Nj}\right)$.

Iterative Estimation Algorithm

1. Estimate $\boldsymbol{u}$ with pseudo-likelihood estimation approach (i.e. GLMM):
   i. $\log(y_{ij}|y, \boldsymbol{\theta}, \boldsymbol{u}) = \theta^{(0)} + u_i^{(1)} + u_j^{(1)}$
   ii. extract the vector of the random effects $\boldsymbol{u}^{(1)}$
2. Estimate $\boldsymbol{\theta}$ with stepping algorithm approach and set $\boldsymbol{u}^{(1)}$ as offset parameter:
   i. $\log\left(P(y_{ij} = 1|y_{-ij}, \boldsymbol{\theta}, \boldsymbol{u})\right) = \boldsymbol{\theta}^{(1)} s_{ij}(y) + \underbrace{\boldsymbol{u}^{(1)} t(y)}_{\text{offset}}$
   ii. extract $\boldsymbol{\theta}^{(1)} s_{ij}(y)$
3. Update step 1 and estimate $\boldsymbol{u}^{(2)}$ taking $\boldsymbol{\theta}^{(1)} s_{ij}(y)$ as offset parameter:
   i. $\log(y_{ij}|y, \boldsymbol{\theta}, \boldsymbol{u}) = \underbrace{\boldsymbol{\theta}^{(1)} s_{ij}(y)}_{\text{offset}} + u_i^{(2)} + u_j^{(2)}$
   ii. extract the vector of the random effects $\boldsymbol{u}^{(2)}$
4. Start again with step 2 until $\boldsymbol{u}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$. converge

2 Network statistics in our model

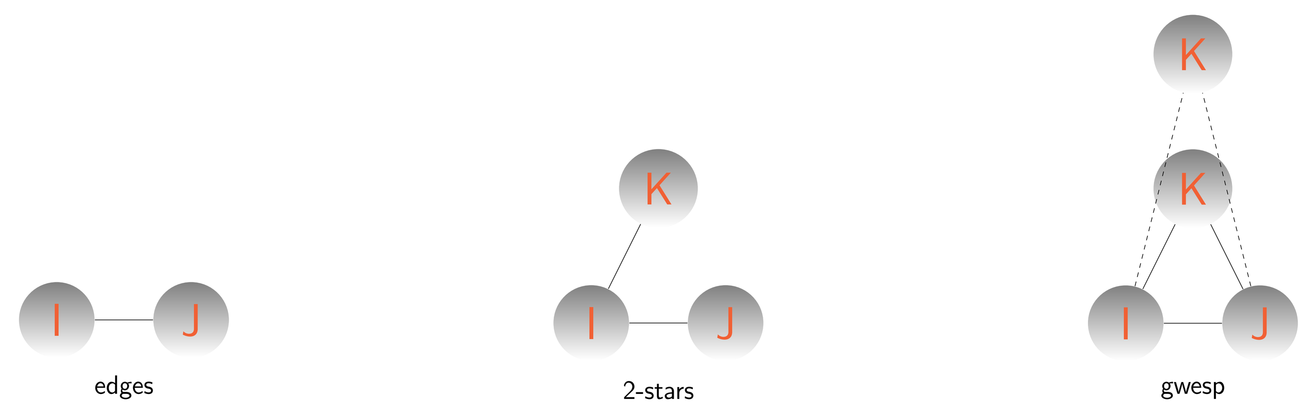


Figure 1: *Left: The network statistics included in the model which have to be estimated with our iterative estimation algorithm – edges: $s(y) = \sum_{i<j} y_{ij}$, 2-stars: $s(y) = \sum_{i<j<k} y_{ij} y_{ik}$ and gwesp: $s(y, \alpha) = \sum_d g_d(\alpha) \sum_{i<j} y_{ij} 1_{\{\sum_k y_{ik} y_{jk} = d\}}$.*

3 Data example and results

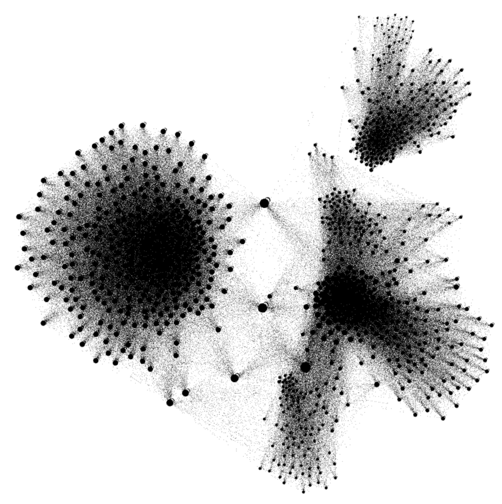


Figure 2: *Facebook network data example of 1000 nodes, computation time 11h 40min using five 2.2Ghz cores in parallel for 100 iterations.*

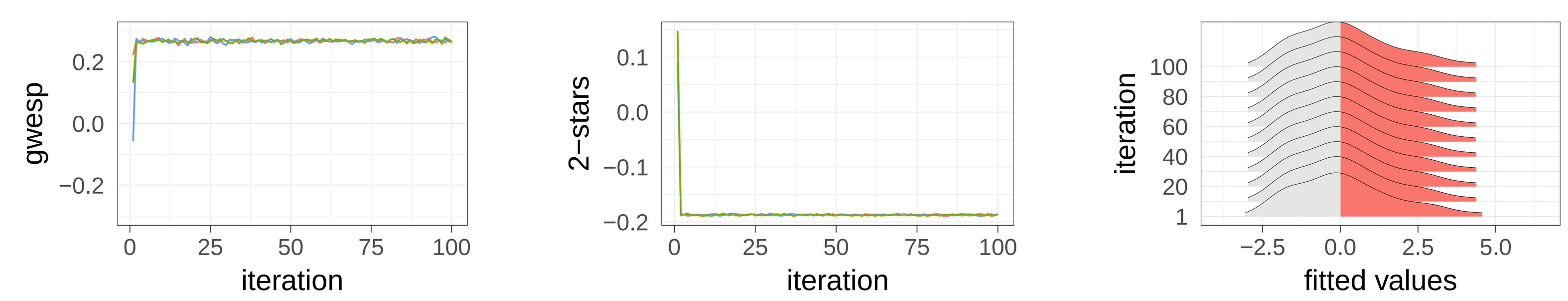


Figure 3: *From left to right: Trend of the gwesp network statistics through 100 iterations, the blue and the green lines have different starting values $\theta^{(0)}$ for convergence reasons. The trend of the 2-stars network statistics through 100 iterations. The third plot shows the density of the random effects which are estimated in each iteration step.*

4 Model Evaluation & Outlook

Likelihood of the extended ERGM:

$$l(\theta, \sigma^2) = log \int \frac{\exp\{\theta^T s(y) + u^T t(y)\}}{\kappa(\theta, u)} \cdot \prod_{i=1}^{n} \phi\left(\frac{u_i}{\sigma_u^2}\right) du.$$

$\Rightarrow$ Laplace Approximation:

$$l(\theta, \sigma^2) \approx log\left\{\frac{\exp\{\theta^T s(y) + \hat{u}^T t(y)\}}{\kappa(\theta, \hat{u})} \cdot \prod_{i=1}^{n} \phi\left(\frac{\hat{u}_i}{\sigma_u^2}\right) \left|J_u\left(\theta,\, \hat{u};\, \sigma_u^2\right)\right|^{-\frac{1}{2}}\right\}.$$

where,

$$J_u(\theta, u; \sigma_u^2) = \frac{1}{N}\sum_{j=1}^{N}\left[t(y^{*(j)}) - \mathbb{E}(t(y)^*)\right] \cdot \left[t(y^{*(j)}) - \mathbb{E}(t(y)^*)\right]^t, \hat{\kappa}(\hat{\theta}, \hat{u}) = \frac{1}{N}\sum_{j=1}^{N} exp(\hat{\theta}^T s(y^{*(j)}) + \hat{u}^T t(y^{*(j)})).$$

Akaike Information Criterium for the extended ERGM

$$AIC = -2\, l(\hat{\theta}, \sigma^2) + 2p$$

Outlook

1. Simulation based study for the model evalution.
2. Speeding up the implementation.

References

Handcock, Mark S. and Hunter, David R. and Butts, Carter T. and Goodreau, Steven M. and Krivitsky, Pavel N. and Morris, Martina (2008). ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *Journal of Statistical Software*.

Kolaczyk, Eric D. (2009). Statistical Analysis of Network Data: Methods and Models. *Springer Series In Statistics*.

Douglas Bates and Martin Mächler and Ben Bolker and Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. DOI: 10.18637/jss.v067.i01.