Exponential Random Graph Models with Big Networks: Maximum Pseudolikelihood Estimation and the Parametric Bootstrap

Christian Schmid
In collaboration with Bruce Desmarais

The Pennsylvania State University

July 31st, 2017

This work was supported in part by NSF grants SES-1558661, SES-1637089, SES-1619644, and CISE-1320219.

Overview

- The Exponential Random Graph Model (ERGM) has an intractable normalizing constant.
- State-of-the-art approach: Approximization by Monte Carlo maximum likelihood (MCMLE).
- For large networks the maximum pseudolikelihood (MPLE) is an option.
- We introduce a resampling method - the parametric bootstrap - that combines advantages of MCMLE and MPLE.

The Exponential Random Graph Model

Idea: Take the adjacency matrix of an observed network A with N nodes as a manifestation of a matrix-like random variable Y.

Definition:

$$\mathrm{P}(Y = A|\theta) = \frac{\exp(\theta^T \cdot \Gamma(A))}{c(\theta)}$$

where

- $\theta \in \mathbb{R}^q$, is a vector of parameters
- $\Gamma : \mathcal{A}(N) \to \mathbb{R}^q \ , \ A \mapsto (\Gamma_1(A), \ldots, \Gamma_q(A))^T$, is a vector of network statistics
- $c(\theta) := \sum_{A^* \in \mathcal{A}(N)} \exp(\theta^T \cdot \Gamma(A^*))$, is a normalization constant

Parameter Estimation Method 1: MCMLE

Idea: Fix an auxiliary parameter vector $\theta_0 \in \mathbb{R}^q$. Then, we can show

$$\frac{c(\theta)}{c(\theta_0)} = \mathbb{E}_{\theta_0}\left[exp((\theta - \theta_0)^T \Gamma(A))\right]$$

Approximate with MCMC sample of networks $A_1, \dots, A_L$.

Then, by the law of big numbers, we get

$$\frac{c(\theta)}{c(\theta_0)} \approx \frac{1}{L}\sum_{i=1}^{L}\left[exp((\theta - \theta_0)^T \Gamma(A_i))\right]$$

Efficiency of MPLE/MCMLE

- MCMLE is approximately exact.
- MPLE approaches the MLE as the size of the network increases.
- MCMLE requires a large sample size to perform well.
- The required sample size increases as network size increases.

Simulation Study

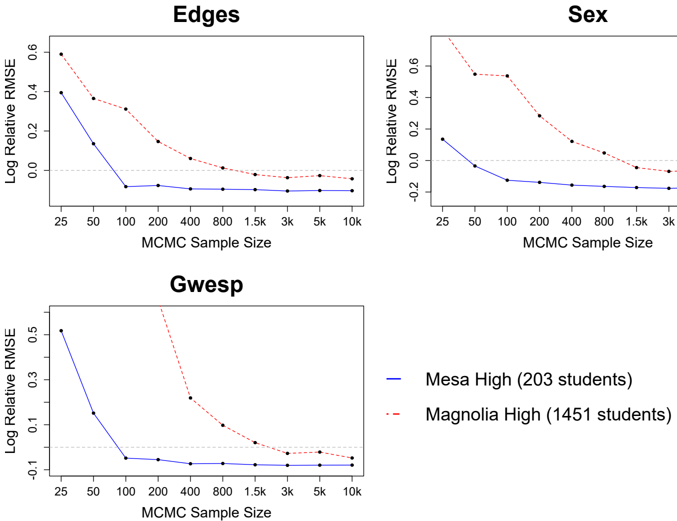
Data:
We use two friendship networks, Faux Mesa High (203 students) and Faux Magnolia High (1451 students), and use the same parametrization (edges, sex and gwesp)

Treatment:

- Take MCMLE of both networks as the 'true' coefficient θ and simulated 500 networks.
- Estimate the coefficients of these 500 networks using MPLE and MCMLE.
- For every single simulated network the MCMLE calculation is being repeated several times for 25 to 10.000 simulated networks used in the likelihood approximation.

Simulation Study Results

This plot visualizes the log of the ratio of the root MSE of the MCMLE to the MPLE.



Parameter Estimation Method 2: Bootstrapped MPLE

In contrast to the MPLE, the MCMLE does not underestimate the standard error.

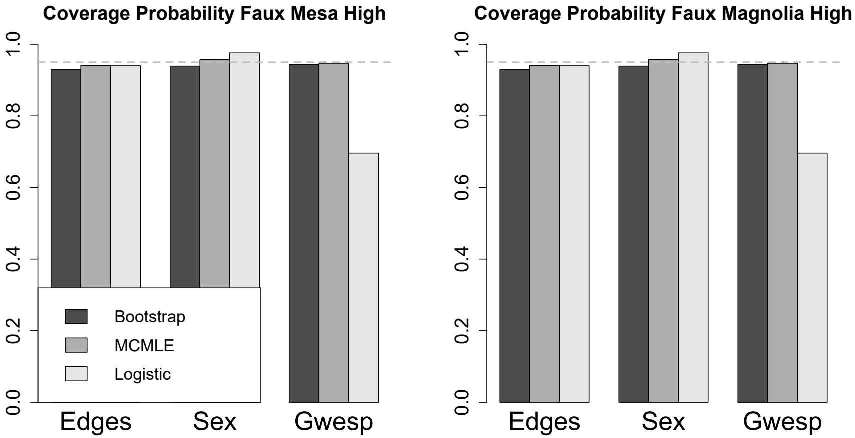
We introduce the **bootstrapped MPLE** in order to obtain more reliable confidence intervals:

1. Estimate the MPLE
2. Simulate a reasonable number of networks (e.g. 500)
3. Estimate the MPLE for each simulated network
4. Take the 2.5th and 97.5th percentile to obtain a 95% bootstrap confidence interval

Bootstrapped MPLE vs. MCMLE: Coverage Probability

Treatment: MCMLE as 'true' coefficients and simulate 1000 networks.

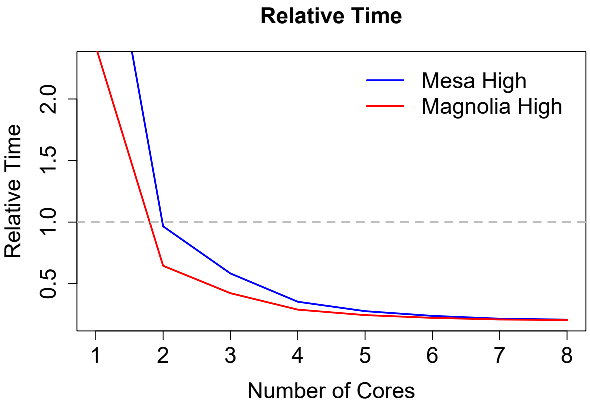
Calculate CI for each simulated network using bootstrapped MPLE, MCMLE and MPLE and report percentage the CI contains the 'true' value



Bootstrapped MPLE vs. MCMLE: Computation Time

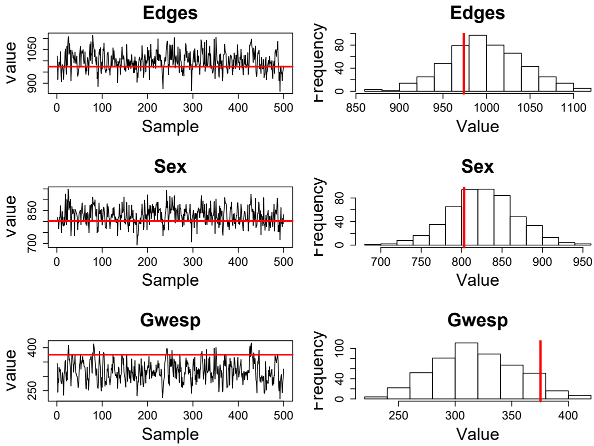
- The bootstrapped MPLE is embarrassingly parallel.
- By using multiple cores the computation time can be further reduced.

This plot visualizes the time ratio of the bootstrapped MPLE to the MCMLE.



Assessing Degeneracy

Degeneracy occurs if the stochastic process generated by the MCMC-algorithm does not hold through the model's defined distribution as stationary distribution.



Conclusion

	Pros	Cons
MCMLE	• approximately exact • can assess degeneracy	• Computationally expensive • θ_0 has to be picked close to θ
MPLE	• simple and fast • consistent estimator	• underestimates st. error
b. MPLE	• simple and fast • can assess degeneracy • consistent estimator • reasonable CIs	• slower than MPLE

References

B.A. Desmarais and S.J. Cranmer
Statistical mechanics of networks: Estimation and uncertainty
Physica A, 391(4):1865–1876, 2011.

M. van Duijn, K. Gile and M. Handcock
A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential random graph models
Social Networks, 31(1):52–62, 2009.

G. Robins, P. Pattison, Y. Kalish and D. Lusher
An introduction to exponential random graph (p) models for social networks*
Social Networks, 29(2):173–191, 2007.

D. Strauss and M. Ikeda
Pseudolikelihood estimation for social networks
JASA, 85:204–212, 1990.

Simulating Networks with MCMC

Choose a matrix $A^{(0)} \in \mathcal{A}(N)$ to start with and proceed as follows:

1. Randomly choose a dyad $A_{ij} := A[i,j]$ where $i \neq j$ from $A^{(k)}$
2. Compute the value

$$\pi := \frac{\mathrm{P}(Y_{ij} \neq A_{ij}^{(k)} | Y_{ij}^c = A_{ij}^c, \theta_0)}{\mathrm{P}(Y_{ij} = A_{ij}^{(k)} | Y_{ij}^c = A_{ij}^c, \theta_0)}$$

   where $Y_{ij}^c = A_{ij}^c$ is short for $Y_{pq} = A_{pq}$ for all $(p,q) \neq (i,j)$.

3. Fix $\delta := \min\{1, \pi\}$ and draw a random number Z from $\mathrm{Bin}(1, \delta)$. If
   - $Z = 0$, let $A^{(k+1)} := A^{(k)}$
   - $Z = 1$, set $A_{ij}^{(k+1)} = 1 - A_{ij}^{(k)}$
4. Start at step 1 with $A^{(k+1)}$.