Latent Space Approaches to Social Network Analysis

by Peter D. Hoff, Adrian E. Raftery, and Mark S. Handcock

Soubhik Barari, Jelena Bauer, Shusei Eshima, and Sun Young Park

GOV 2003: Topics in Quantitative Methods
Harvard University

April 8, 2019

Motivation

Stochastic Block Model

Motivation

Only illustrative!

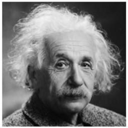
Stochastic Block Model

Motivation

Only illustrative!

Stochastic Block Model

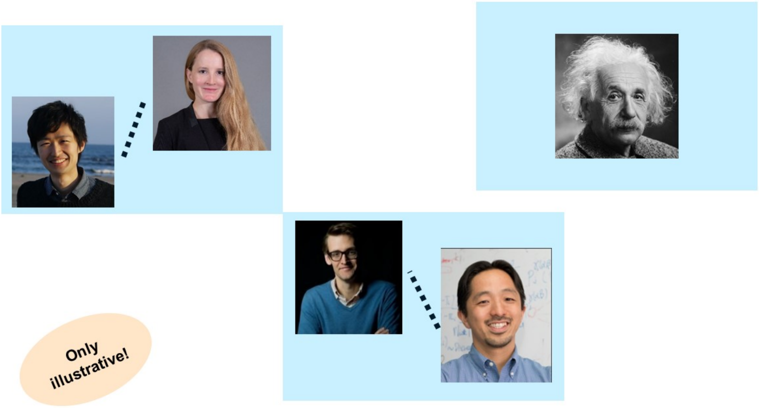
Motivation



Only illustrative!

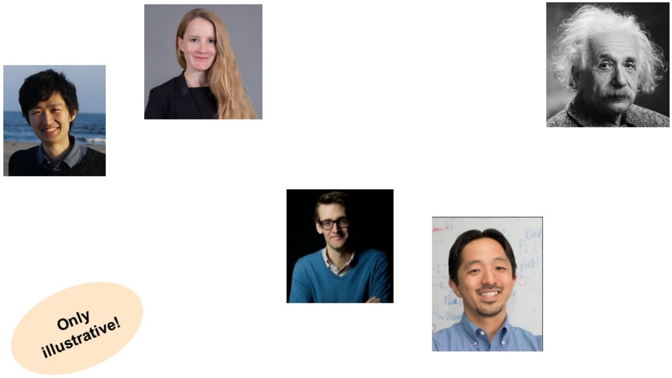
Stochastic Block Model

Motivation



Stochastic Block Model

Motivation



Latent Space Network Model

Motivation

Latent Space Network Model

Motivation

Latent Space Network Model

Motivation

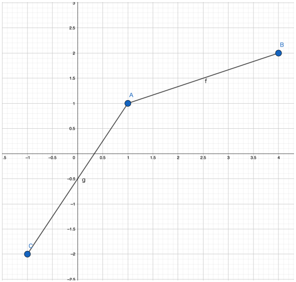
Latent Space Network Model

Distance vs Projection Model

- Goal of the model
  - We want to map actors on a latent space
  - If two actors are close in the space, they are more likely to have a tie

  $\Longrightarrow$ Distance model
- Incorporating how active actors are (asymmetric relation)

  $\Longrightarrow$ Projection model



Set Up for a Latent Space Network Model

Social network data set up

- n: number of actors
- $i, j = 1, ..., n$: actors
- y_{ij}: relational ties
- Y: $n \times n$ matrix representing these ties
- Direction of the path
    - Undirected: $i \sim j$, $y_{ij} = y_{ji}$
    - Directed: $i \rightarrow j$

Distance Models

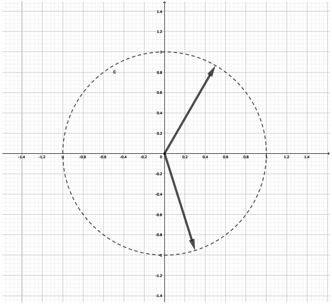
- i and j has a relationship $Y_{ij} \in \{0, \cdots, K-1\}$, typically $K = 2$
- We model $\pi_{ij} = p(Y_{ij} = 1 | z, x, \alpha, \beta)$
  - z: position in the latent space
  - x: covariates
  - α, β: parameters
- Same as the Logistic regression

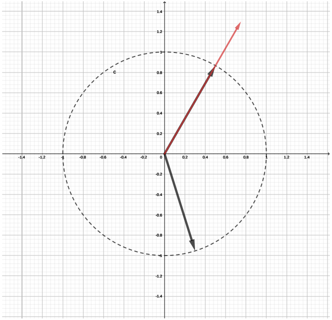
$$\begin{aligned} \text{logit}(\pi_{ij}) &= \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) \\ &= \alpha + \beta^\top x_{ij} - \underbrace{|z_i - z_j|}_{\text{distance in a latent space}} \end{aligned}$$

- All else equal, i and j are more likely to have a relationship if they are close in a latent space.

Projection Model

- We want to model asymmetric relationship
  - $|z_i - z_j|$ assumes symmetry
  - i and j have a tie but i is socially active
  - j is only a one of i's many ties
- Active level parameter $a_i > 0$
- A low-dimensional Euclidean space, e.g., $\vec{z} \in \mathbb{R}^2$
- Redefine: $\vec{z}_i = a_i \vec{e}_i$: $\color{red}{\vec{z}_i}$ $= 1.5\vec{e}_i$



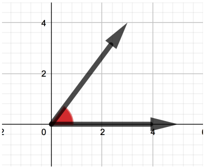


Projection Model

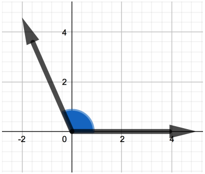
- Replace the distance term: $\operatorname{logit}(\pi_{ij}) = \alpha + \beta^\top x_{ij} + \frac{\vec{z}_i^\top \vec{z}_j}{|\vec{z}_j|}$

$$\frac{\vec{z}_i^\top \vec{z}_j}{|\vec{z}_j|} = \frac{|a_i \vec{e}_i||a_j \vec{e}_j| \cos\theta}{|a_j \vec{e}_j|} = a_i \cos\theta, \quad -a_i \le a_i \cos\theta \le a_i$$

- This is the model of homophily
- Given a_i, the smaller the angle between two vectors are, there more likely they have a tie



Prone to having a tie,
$\vec{e}_i^\top \vec{e}_j = \cos\theta > 0$



Averse to having a tie,
$\vec{e}_i^\top \vec{e}_j = \cos\theta < 0$

Great news: likelihood-based estimation methods are feasible...

Likelihood function:

$$\prod_{i\neq j} p(y_{ij}|z_i, z_j, x_{ij}, \alpha, \beta) = \prod_{i\neq j} (\pi_{ij})^{\mathrm{I}(y_{ij}=1)} (1-\pi_{ij})^{\mathrm{I}(y_{ij}=0)}, \ \ \pi_{ij} = p(Y_{ij} = 1|\cdot)$$

We define

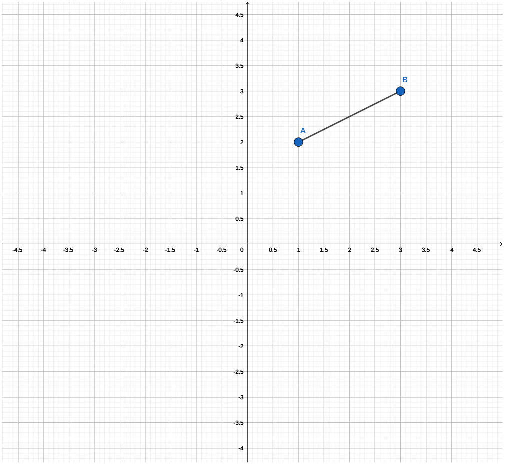
$$\eta_{ij} = \log \frac{\pi_{ij}}{1-\pi_{ij}} \ \Leftrightarrow \ \pi_{ij} = \frac{e^{\eta_{ij}}}{1+e^{\eta_{ij}}}$$

By taking log,

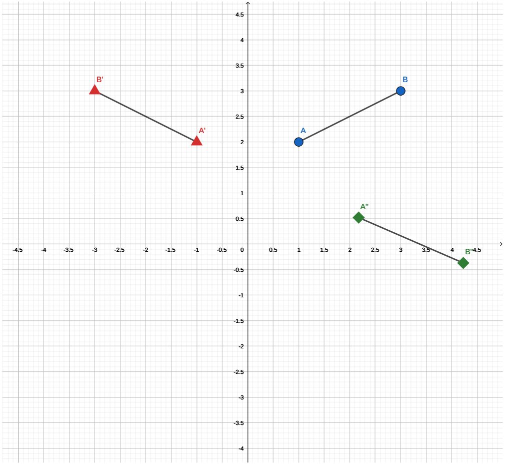
$$\begin{aligned} &\sum_{i\neq j} y_{ij}(\eta_{ij} - \log(1+e^{\eta_{ij}})) + (1-y_{ij})(-\log(1+e^{\eta_{ij}})) \\ &= \sum_{i\neq j} y_{ij}\eta_{ij} - \log(1+e^{\eta_{ij}}) \end{aligned}$$

Great news!

... but there is an identification problem



... but there is an identification problem



Distances between a set of points in Euclidean space are invariant under rotation, reflection, and translation

How to solve this

Idea:

- Define a class of positions equivalent to latent position Z under rotation, reflection, and translation ("configuration")
- Make inference on configurations via inference of particular elements of configurations that are comparable across configurations

Estimation

Procedure for sampling from posterior distribution:

0. Use prior information on α, β, and Z
1. Identify an MLE $\hat{Z}$ of Z, centered at origin, by direct maximization of likelihood
2. Construct Markov chain over model parameters:

   Start $Z_0 = \hat{Z}$

   a. Sample a proposal Z from a symmetric proposal distribution
   b. Accept that proposal for Z_{k+1} with some probability; otherwise, set $Z_{k+1} = Z_k$
   c. Store $\tilde{Z}_{k+1} = \arg\min_{TZ_{k+1}} tr(\hat{Z} - TZ_{k+1})' tr(\hat{Z} - TZ_{k+1})$
3. Update α and β with a Metropolis-Hastings algorithm

Applications

Actors	Tie (Y_{ij})	Distribution	Model
Monks	Affinity	$Y_{ij} \sim \text{Bern}(\pi_{ij})$	$\pi_{ij} = \text{logit}^{-1}(\alpha - \lVert Z_i - Z_j \rVert)$
Women	Contacts	$Y_{ij} \sim \text{Bern}(\pi_{ij})$	$\pi_{ij} = \text{logit}^{-1}(\alpha + \frac{Z_i Z_j}{\lVert Z_j \rVert})$
Senators	Bills co-sponsored	$Y_{ij} \sim \text{Pois}(\lambda_{ij})$	$\lambda_{ij} = \log^{-1}(\alpha - \lVert Z_i - Z_j \rVert)$

R package:
`latentnet`

Code repo:
`http://github.com/soubhikbarari/Gov2003-LatentNet`

Application 1: Monk Network (Sampson, 1986)

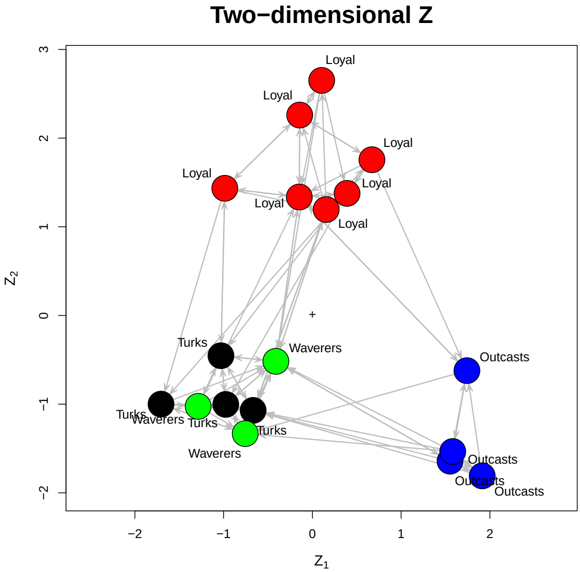
Political Preference and Social Relations between Novice Monks in New England

Actors	Tie (Y_{ij})	Distribution	Model
Monks	Affinity	$Y_{ij} \sim \text{Bern}(\pi_{ij})$	$\pi_{ij} = \text{logit}^{-1}(\underbrace{\alpha - \lVert Z_i - Z_j \rVert}_{\text{distance model}})$

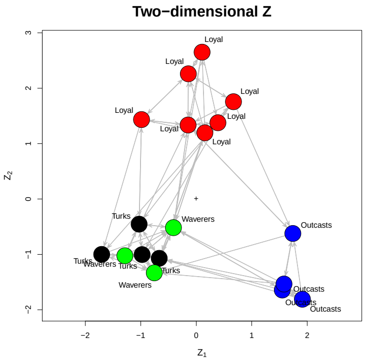
Code:

```
ergmm(samplike ~ euclidean(d = 2))
```

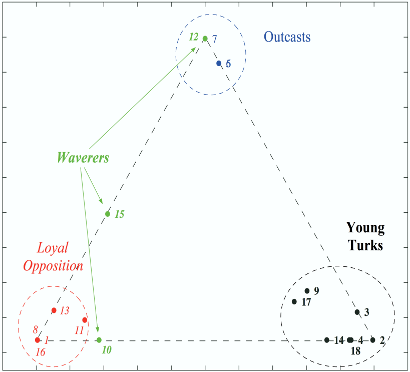
Application 1: Monk Network (Sampson, 1986)



Application 1: Monk Network (Sampson, 1986)



latent social position
(LSNM)



posterior mixed membership in a simplex
(MMSBM)

Application 1: Monk Network (Sampson, 1986)

Original distance model without clusters:

$$\begin{aligned} Y_{ij} &\sim \text{Bern}(\pi_{ij}) \\ \pi_{ij} &= \text{logit}^{-1}(\underbrace{\alpha - \|Z_i - Z_j\|}_{\text{distance model}}) \end{aligned}$$

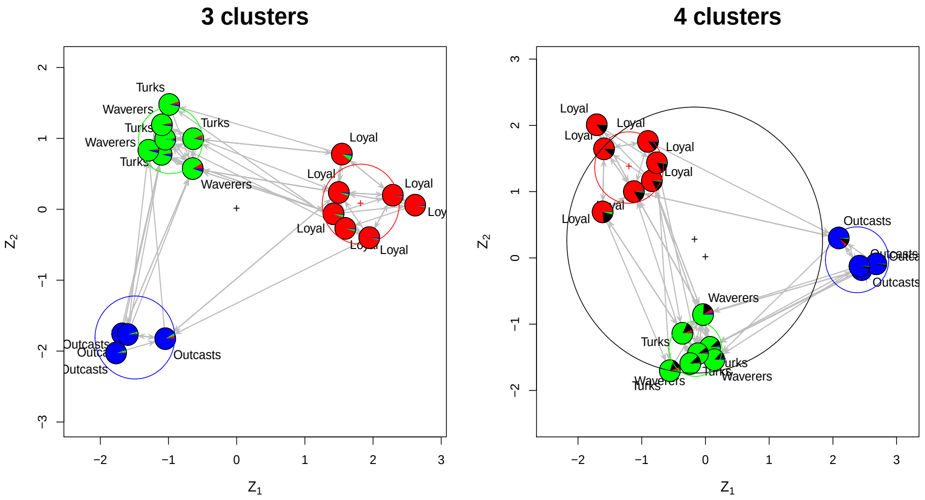
Extension of distance model to allow clustering:

$$\begin{aligned} Z_i &\sim \sum_{g=1}^{G} \lambda_g \text{MVN}_d(\mu_g, \sigma_g^2 I_d) \quad i = 1, \ldots, n, \\ \mu_g &\sim \text{MVN}_d(0, \omega^2 I_d) \quad g = 1, \ldots, G, \\ \sigma_g^2 &\sim \sigma_0^2 \text{Inv}\chi_\alpha^2 \quad g = 1, \ldots, G, \\ (\lambda_1, \ldots, \lambda_G) &\sim \text{Dirichlet}(\nu_1, \ldots, \nu_G) \end{aligned}$$

Code:

```
ergmm(samplike ~ euclidean(d = 2, G=3))
```

Application 1: Monk Network (Sampson, 1986)



Application 2: Women (Davis, Gardner, Gardner, 1941)

Social Stratification among Women in Mississippi

Actors	Tie (Y_{ij})	Distribution	Model
Women	Contacts	$Y_{ij} \sim \text{Bern}(\pi_{ij})$	$\pi_{ij} = \text{logit}^{-1}\Big(\underbrace{\alpha + Z_i Z_j}_{\text{projection model}}\Big)$

Code:

```
ergmm(davis ~ bilinear(d = 2, G = 2))
```

Application 2: Women (Davis, Gardner, Gardner, 1941)

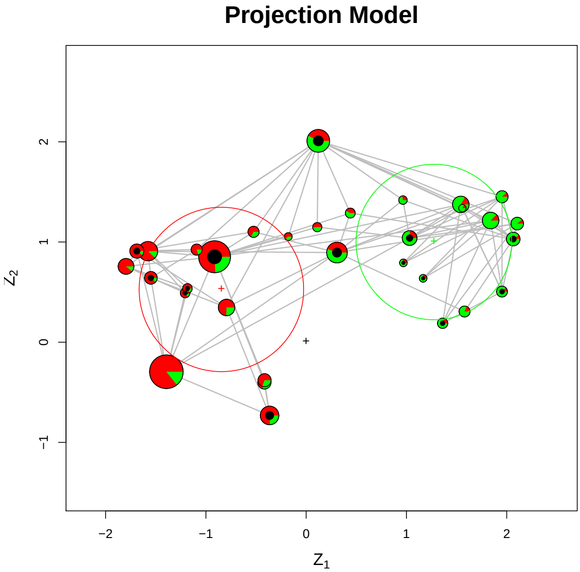
Social Stratification among Women in Mississippi

Actors	Tie (Y_{ij})	Distribution	Model
Women	Contacts	$Y_{ij} \sim \text{Bern}(\pi_{ij})$	$\pi_{ij} = \text{logit}^{-1}\Big(\underbrace{\alpha + Z_i Z_j}_{\text{projection model}} + \underbrace{\xi_i + \xi_j}_{\text{sociality effect}}\Big)$

Code:

```
ergmm(davis ~ bilinear(d = 2, G = 2) + rsociality)
```

Application 2: Women (Davis, Gardner, Gardner, 1941)



Application 3: Legislator Co-sponsorship (Fowler 2005)[1]

Hidden Social Network between U.S. Senators $\rightsquigarrow$ **Bill Co-sponsorship**. Can we discover the former using the latter?

Actors	Tie (Y_{ij})	Distribution	Model
Senators	Bills co-sponsored	$Y_{ij} \sim \text{Pois}(\lambda_{ij})$	$\lambda_{ij} = \log^{-1}(\alpha - \|Z_i - Z_j\|)$

Code:

```
cospons.latent <- ergmm(cospons.net ~ euclidean(d=2), family="Poisson")
```

[1] Fowler, James H. 'Legislative Cosponsorship Networks in the US House and Senate.' *Social Networks* 28.4 (2006): 454-465.

Application 3: Legislator Co-sponsorship (Fowler 2005)

Diagnostic tools

Application 3: Legislator Co-sponsorship (Fowler 2005)

Diagnostic tools

- MCMC Convergence
  - Trace plot, autocorrelation plot, posterior density
  - `mcmc.diagnostics(cospons.latent)`

Application 3: Legislator Co-sponsorship (Fowler 2005)

Diagnostic tools

- MCMC Convergence
  - Trace plot, autocorrelation plot, posterior density
  - `mcmc.diagnostics(cospons.latent)`
- Posterior predictive checks
  - Do posterior-simulated networks resemble observed Y?
  - `plot(gof(cospons.latent))`

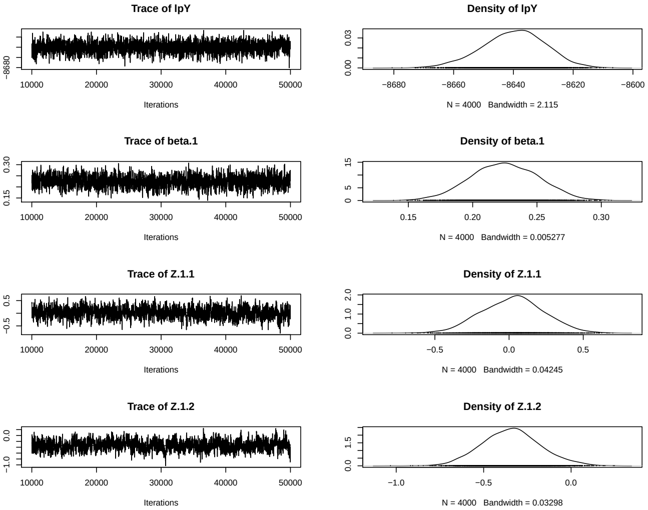
Application 3: Legislator Co-sponsorship (Fowler 2005)

Diagnostic tools

- MCMC Convergence
  - Trace plot, autocorrelation plot, posterior density
  - `mcmc.diagnostics(cospons.latent)`
- Posterior predictive checks
  - Do posterior-simulated networks resemble observed Y?
  - `plot(gof(cospons.latent))`
- Bayesian Information Criteria
  - $k\log(n) + 2\log\left(\frac{1}{L(\hat{Z}|Y)}\right)$ for n data points and k parameters
  - Model complexity additionally penalized by inverse likelihood
  - Can use for model comparison (pick lowest)
  - `summary(cospons.latent)`

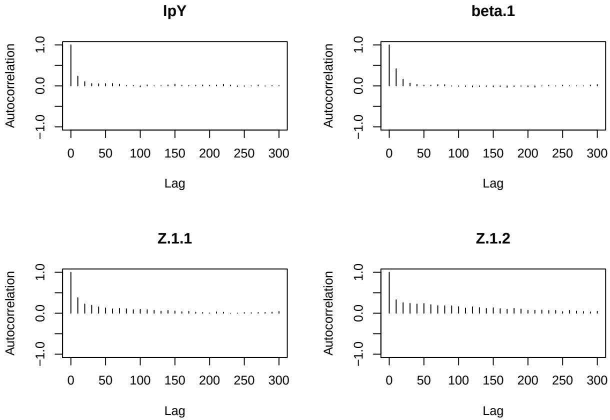
Application 3: Legislator Co-sponsorship (Fowler 2005)

MCMC Convergence – trace plot



Application 3: Legislator Co-sponsorship (Fowler 2005)

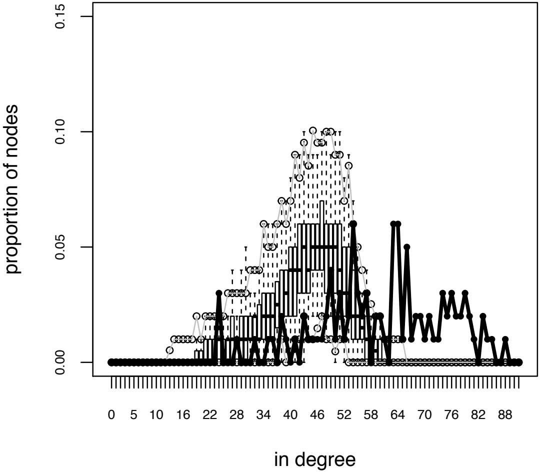
MCMC Convergence – autocorrelation plot



Application 3: Legislator Co-sponsorship (Fowler 2005)

Posterior predictive checks

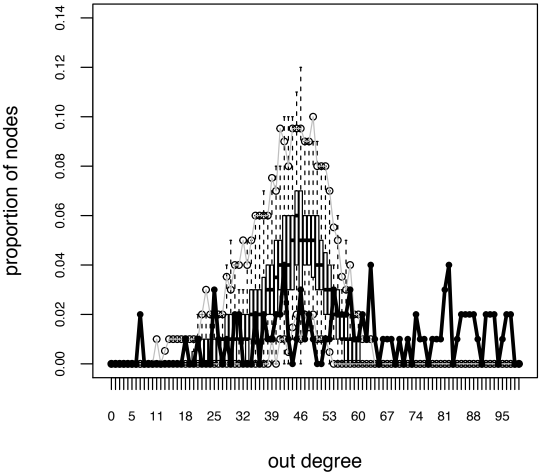
network summary statistic: in-degree count



Application 3: Legislator Co-sponsorship (Fowler 2005)

Posterior predictive checks

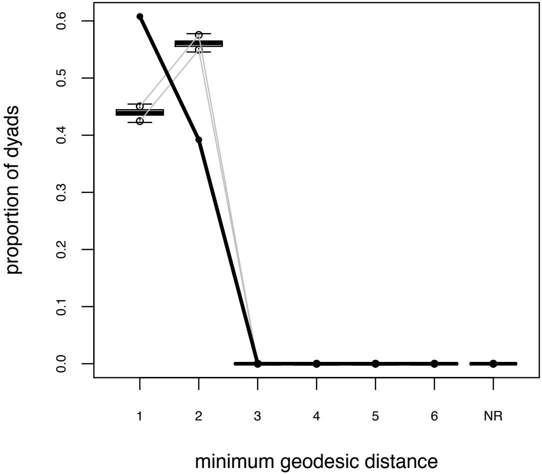
network summary statistic: out-degree count



Application 3: Legislator Co-sponsorship (Fowler 2005)

Posterior predictive checks

network summary statistic: geodesic distance (min dist. b/t any node pair)



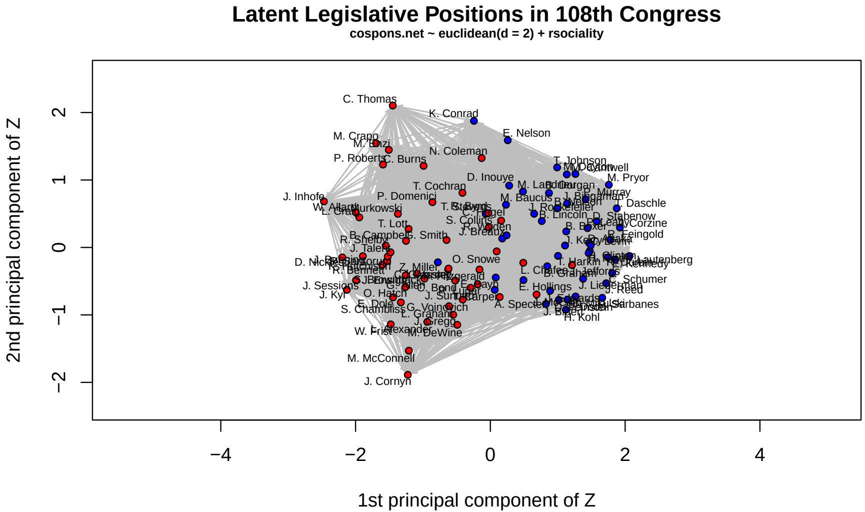
Application 3: Legislator Co-sponsorship (Fowler 2005)

Model comparison

Model	Overall BIC	Log lik.	Number of parameters
`cospons.net ~ euclidean(d=2)`	17325	-8640	$(102 \times 2) - 3 + 1$
`cospons.net ~ euclidean(d=3)`	11507	-5356	$(102 \times 3) - 5 + 1$
`cospons.net ~ euclidean(d=2) + rsociality`	11021	-5227	$(102 \times 3) - 3 + 1$
`cospons.net ~ euclidean(d=2, G=2)`	11486	-5531	$(102 \times 2) - 3 + 3$

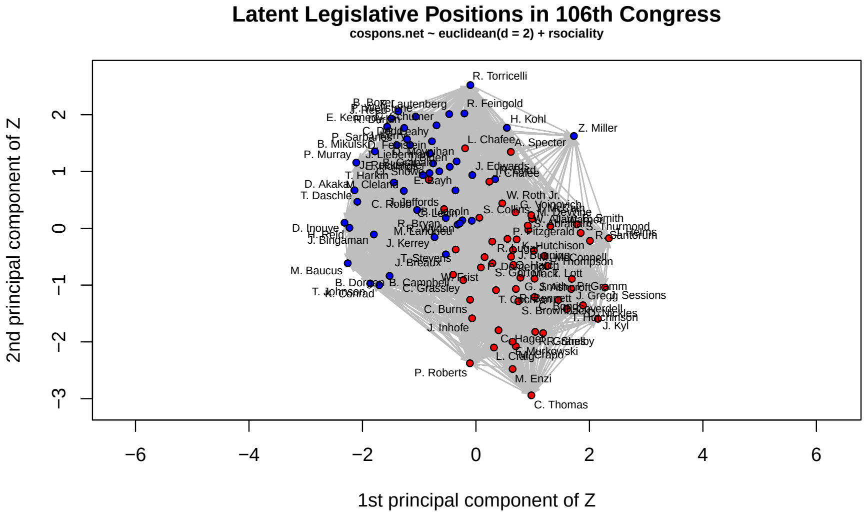
Application 3: Legislator Co-sponsorship (Fowler 2005)

2003-2005:



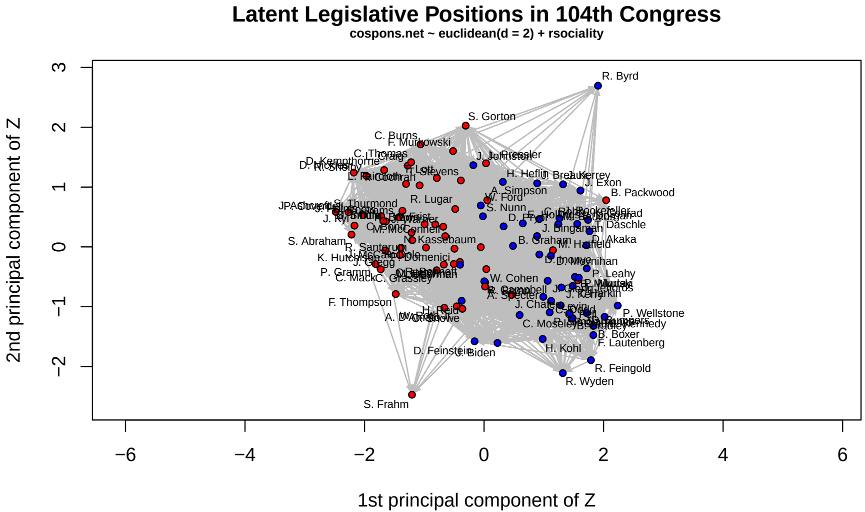
Application 3: Legislator Co-sponsorship (Fowler 2005)

1999-2000:



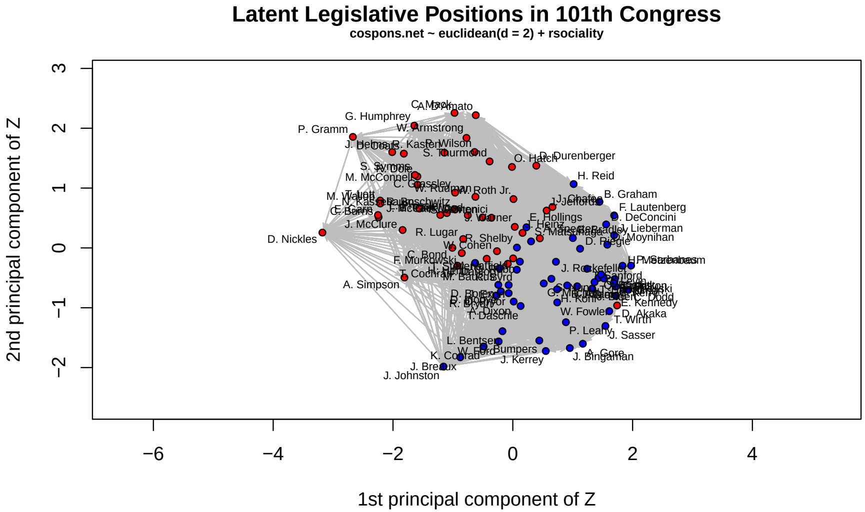
Application 3: Legislator Co-sponsorship (Fowler 2005)

1995-1997:



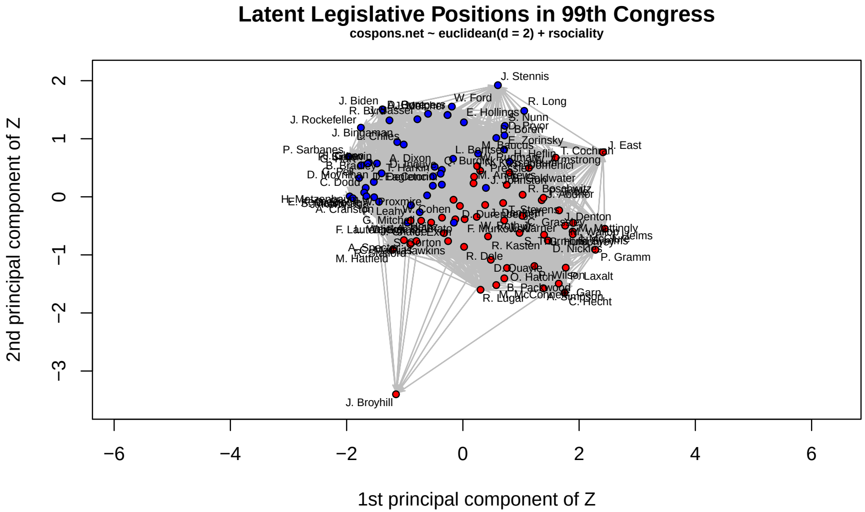
Application 3: Legislator Co-sponsorship (Fowler 2005)

1989-1991:



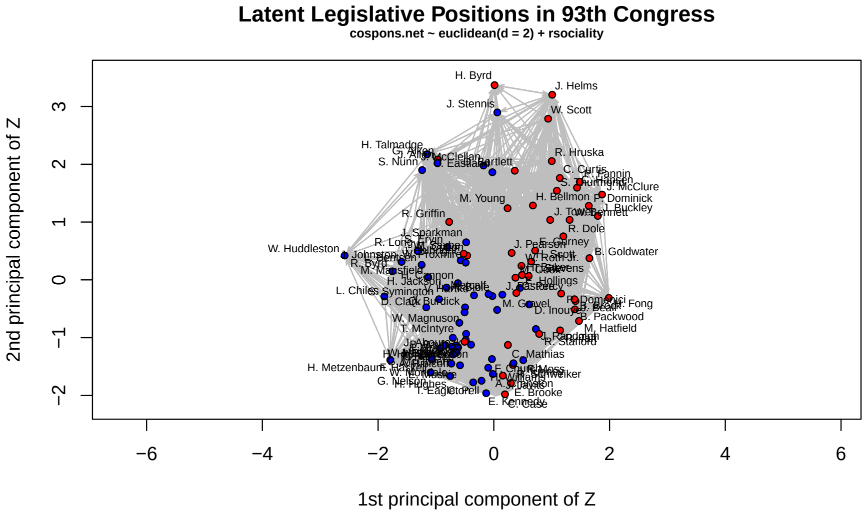
Application 3: Legislator Co-sponsorship (Fowler 2005)

1985-1987:



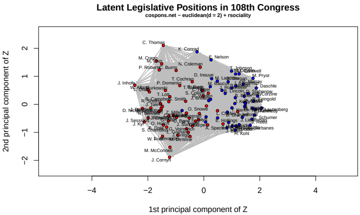
Application 3: Legislator Co-sponsorship (Fowler 2005)

1973-1975:

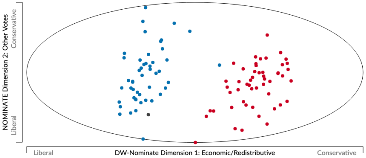


Application 3: Legislator Co-sponsorship (Fowler 2005)

Future extension: Connect **co-sponsorship** to **ideology**.



latent social position
(LSNM)



latent ideological position
(DW-nominate)

Concluding thoughts

- LSNM articulates an explicitly spatial model for network formation.
- Bayesian (e.g. posterior mean) and Frequentist methods (e.g. MLE) for estimating parameters.
- LSNM imposes fewer assumptions than MMSBM (conditionally independent ties).
  - freedom to specify the tie distribution, link function and priors.
- Choosing a larger $d \rightsquigarrow$ better fit (to a certain extent) but adds more parameters + exponentially increases run-time.
- Captures properties like transitivity[2] and reciprocity[3] in observed network.

[2] If $i \to j$ and $j \to k$ then probably $i \to k$.
[3] If $i \to j$ then probably ji.

Fun Fact

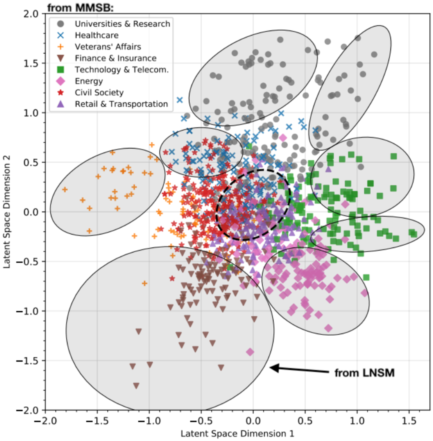
Even fictional researchers are using `ergmm` these days (from *The Wire*):

We hope you (real researchers) will too! ☺

Appendix

Comparing random graph models for validation

Scaling politicians and interest groups based on observed lobbying and sponsorship (Kim and Kunisky 2017)[4]:



[4] Kim, In Song, and Dmitriy Kunisky. 'Mapping Political Communities: A Statistical Analysis of Lobbying Networks in Legislative Politics'. Working Paper. `http://web.mit.edu/insong/www/pdf/network.pdf`, 2017.

LSNM vs. MMSBM

Supposing ties follow a Bernoulli distribution:

	Canonical LSNM	**Canonical MMSBM**
tie	$Y_{ij} \sim \text{Bern}(\text{logit}^{-1}(\alpha - \lVert Z_i - Z_j \rVert))$	$Y_{ij} \sim \text{Bern}(Z_{i\to j}^T \mathbf{B} W_{j\to i})$
structure	$Z_i \sim \underbrace{\text{MVN}_d([0],[100])}_{\text{diffuse prior}}$	$\text{Z}_{i\to j} \sim \text{Mult}(1,\pi_i)$ $\text{W}_{j\to i} \sim \text{Mult}(1,\pi_j)$ $\mathbf{B}_{sr} \sim \text{Beta}(\beta_{sr},\gamma_{sr})$
estimation	Metropolis-Hastings algorithm	Collapsed Gibbs sampling Variational Inference
process	position	community

Application: Tribes (Read, 1954)

Political Relationships between Tribes in New Guinea

Actors	Tie (Y_{ij})	Distribution	Model
Tribes	alliance neutral hostile	$Y_{ij} \sim \text{Mult}(1, \pi_{ij})$	?

Code:

```
ergmm(tribes ~ euclidean(d = 2, G = 3),
                         response = "sign.012",
                         family = "binomial.logit",
                         fam.par = list(trials = 2))
```

Application: Tribes (Read, 1954)