STA6800 - Statistical Analysis of Network

Additive and Multiplicative Effects Network Models

Ick Hoon Jin

Yonsei University, Department of Statistics and Data Science

Introduction

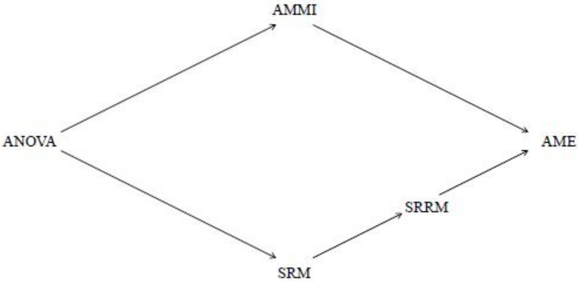


FIG 1. *A graph describing the relationships between models. An arrow is drawn from one model to another if the former can be viewed as a submodel of the latter.*

Addictive Effect Model

- Goal : Consider dependency.
- Addictive effect model (IID model)

$$y_{i,j} = \mu + a_i + b_j + \epsilon_{i,j}.$$

- a_i : sender effect (row means of the sociomatrix),
- b_j : receiver effect (column means of the sociomatrix).

Social Relations Model

- Goal : sender-receiver correlations(dyadic correlations)
- Social relations medel(SRM)

$$y_{i,j} = \mu + a_i + b_j + \epsilon_{i,j},$$

where $a_i's, b_j's, \epsilon_{i,j}'s$ are mean zero random variables for which

$$Var\left[\begin{pmatrix} a_i \\ b_j \end{pmatrix}\right] = \Sigma = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}, \quad Var\left[\begin{pmatrix} \epsilon_{i,j} \\ \epsilon_{j,i} \end{pmatrix}\right] = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

with effects otherwise being independent.

Social Relations Model Example

- n = 30 countries
- $y_{i,j}$: the 1990 export volume from country i to country j, in log billions of dollars for $i = 1, ..., n$, $j = 1, ..., n$
- $\hat{a}_i$: the i-th row mean minus the grand mean $\hat{\mu}$
- $\hat{b}_j$: the j-th row mean minus the grand mean $\hat{\mu}$

Social Relations Model Example

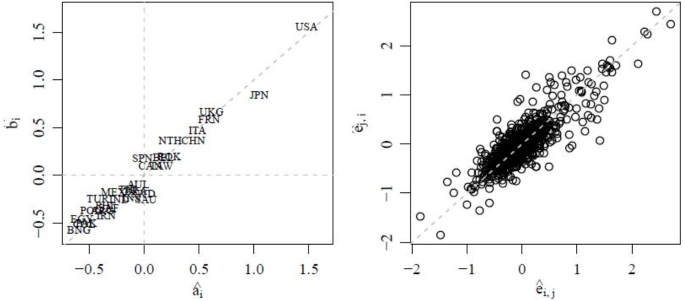


FIG 2. *Left panel: Scatterplot of country-level export effects versus import effects. Right panel: Scatterplot of dyadic residuals.*

Social Relations Covariance Model

- Social relations covariance model

$$Var[y_{i;j}] = \sigma_a^2 + \sigma_b^2 + \sigma^2,$$

where

- $Cov[y_{i,j}, y_{i,k}] = \sigma_a^2$ (within-row covariance)
- $Cov[y_{i,j}, y_{k,j}] = \sigma_b^2$ (within-column covariance)
- $Cov[y_{i,j}, y_{j,k}] = \sigma_{ab}$ (row-column covariance)
- $Cov[y_{i,j}, y_{j,i}] = 2\sigma_{ab} + \rho\sigma^2$ (row-column covariance plus reciprocity).

with all other covariances between elements of $\boldsymbol{Y}$ being zero.

Social Relations Regression Model

- Goal: Quantify the association between a particular dyadic variable and some other dyadic or nodal variables
- Social Relations Regression Model: Combines a linear regression model with the covariance structure of the SRM as follows:

$$y_{i,j} = \boldsymbol{\beta}^T \boldsymbol{x}_{i,j} + \mu + a_i + b_j + \epsilon_{i,j},$$

  where $\boldsymbol{x}_{i,j}$ is $p-$dimensional vector of regressors and $\boldsymbol{\beta}$ is a vector of regression coefficient to be estimated.
- Limitation: Unable to represent higher-order network patterns (lack of fit)

Social Relations Regression Model Example

- country-specific GDP and polity
- the geographic distance between pairs of county capitals.
- quantify the relationship between trade and polity after controlling for the effects of GDP and geographic distance.

$$y_{i,j} = \beta_0 + \beta_{r,1} polity_i + \beta_{r,2} gdp_i + \beta_{c,1} polity_j + \beta_{c,2} gdp_j + \beta_d distance_{i,j} + \epsilon_{i,j},$$

where $\epsilon_{i,j}^{\prime}s$ are assumed to be i.i.d. mean-zero error terms.

Social Relations Regression Model Example

	IID			SRRM			AME		
regressor	$\hat{\beta}$	se($\hat{\beta}$)	t-ratio	$\hat{\beta}$	se($\hat{\beta}$)	t-ratio	$\hat{\beta}$	se($\hat{\beta}$)	t-ratio
exporter polity	0.015	0.004	4.166	0.015	0.016	0.939	0.013	0.016	0.786
importer polity	0.022	0.004	6.070	0.022	0.016	1.420	0.018	0.015	1.173
exporter GDP	0.411	0.021	19.623	0.401	0.097	4.117	0.340	0.103	3.306
importer GDP	0.398	0.020	19.504	0.391	0.093	4.189	0.331	0.101	3.266
distance	-0.057	0.004	-13.360	-0.064	0.006	-11.578	-0.041	0.004	-10.724

TABLE 1

Parameter estimates and standard errors from the trade data using a normal linear regression model with i.i.d. errors, a SRRM, and an AME model.

Multiplicative Effects Models

- Goal: Capture higher-order network patterns
- Additive and Multiplicative Effects model (AME)

$$y_{i,j} = \boldsymbol{\beta}^T \mathbf{x}_{i,j} + \mathbf{u}_i^T \mathbf{v}_j + a_i + b_j + \epsilon_{i,j},$$

$$(\epsilon_{i,j}, \epsilon_{j,i}) : i < j \sim i.i.d. N_2 \left(\mathbf{0}, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

Additive and Multiplicative Effects Model Example

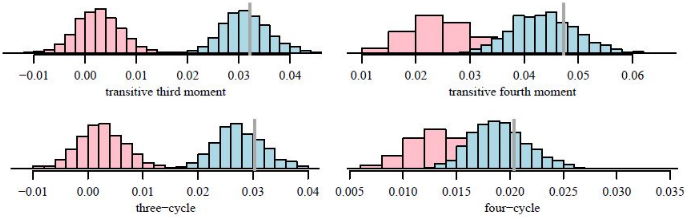


FIG 3. *Posterior predictive distributions of third- and fourth-order goodness-of-fit statistics. The pink histograms correspond to the SRRM fit, the blue to the AME fit. The observed values of the statistics are given by vertical gray lines.*

- pink histograms: SRRM fit
- blue histograms: AME fit
- vertical gray lines: The observed values of the statistics

Additive and Multiplicative Effects Model Example

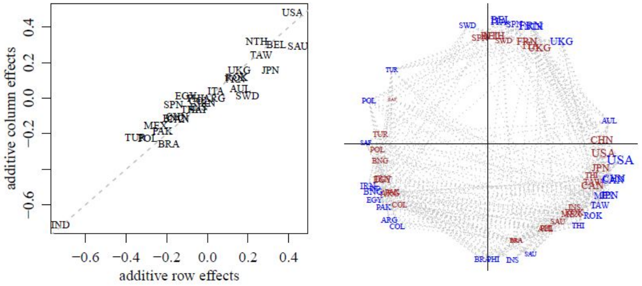


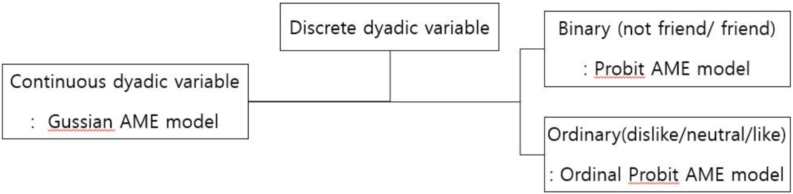
FIG 4. *Estimates of node-specific effects. The left panel gives additive row effects versus additive column effects. The plot on the right gives estimates of* $\mathbf{u}_i$ *in red and* $\mathbf{v}_i$ *in blue for each country* $i = 1, \ldots, n$. *The country names indicate the direction of these vectors, and the size of the plotting text indicates their magnitude. A dashed line is drawn between an export-import pair if their trade flow is larger than expected based on the other terms in the model.*

Multiplicative Effects Models

- Goal: prevent overfitting and provide summaries of certain network dependencies.
- Random effects AME model

$$(\boldsymbol{u}_1, \boldsymbol{v}_1), ..., (\boldsymbol{u}_n, \boldsymbol{v}_n) \sim i.i.d. \quad N_{2r}(\mathbf{0}, \Phi),$$
$$(a_1, b_1), ..., (a_n, b_n) \sim i.i.d. \quad N_2(\mathbf{0}, \Sigma).$$

Transformation Models for Non-Gaussian



ERGMs and Latent Variables Models

- ERGMs
  - Evaluating specific global network patterns of interest
  - Simply by including an appropriate sufficient statistic in the model
- Latent variable models
  - Description of local, micor-level patterns of relationships among specific nodes
- AME
  - Local patterns via estimating the node-specific effects $(a_i, b_i, \boldsymbol{u}_i, \boldsymbol{v}_i)$
  - Global patterns via estimating the parameters $\beta, \Sigma, \Psi, \sigma^2, \rho$

Comparisons to ERGMs vs SRM

$$P(Y) \sim \exp\left(\mu \sum_{i,j} y_{i,j} + \sum_{i}(a_i \sum_{j} y_{i,j} + b_i \sum_{j} y_{j,i}) + \rho \sum_{i,j} y_{i,j} y_{j,i}\right)$$

p1 model

- $\sum_{i,j} y_{i,j}$ = sufficient statistics, the total number of ties .... (1)
- $\sum_{i,j} y_{i,j} y_{j,i}$ = the number of reciprocated ties .... (2)
- $\sum_j y_{i,j}, \sum_j y_{j,i}$ = in- and out- degrees .... (3)

SRM

- Overall mean of the relations (μ).... (1)
- Dyadic correlation (ρ).... (2)
- Heterogeneity in row and column means (a_i, b_i).... (3)

- p2 model extends the p1 model by including regressors (as does the SRRM)
- Treats the node-level parameters a_i and b_i as potentially correlated random effects (as do the SRM and SRRM)

Limitation in P1, P2 Models (ERGMs)

- P1 model is **unable to describe more complex forms of** dependency such as transitivity or clustering
- P2 model or SRRM can represent some degree of higher-order dependency, still **exhibit lack-of-fit and so more complex models are desired**
- ERGMs approach to describe higher-order dependencies is to **include additional sufficient statistics**
  - However, it can **lead to model degeneracy**
  - How to solve?

1. Constraining the parameter space
2. Finding alternative summary statistics

AME Approach to Represent Complex Patterns

- AME represents complex patterns using low-rank matrix $\boldsymbol{UV}^T$
- $n \times n$ matrix Y can be approximated to an arbitrary degree of precision by a product UV^T; two $n \times r$ matrices U and V
- An AME model provides a model-based low-dimensional representation of the observed network

Limitation of Multiplicative Effects Approach

- Not all higher-order moments can be represented by the random effects model for the multiplicative effects (Gaussian random effect model)
- For example, when dimension r=1,

$$E[\gamma_{i,j}\gamma_{j,k}\gamma_{k,l}\gamma_{l,i}] = tr(\Psi_{uv}^4) = \sigma_{uv}^4,$$
$$E[\gamma_{i,j}\gamma_{j,k}\gamma_{k,i}] = tr(\Psi_{uv}^3) = \sigma_{uv}^3$$

  - These moments are not separately estimable because **both completely determined by the single parameter** σ_{uv}
- To separately estimate such moments requires the higher dimension : which is **very tricky**.

Pros of AME Models

- Multiplicative effects matrix UV^T provides **a reduced-rank representation of the sociomatrix** Y
- An estimate of Ψ provides a summary of the across node heterogeneity of the u_i's and v_i's which induces network dependency
- Multiplicative effects describe much broader range of patterns than the simple random effect model

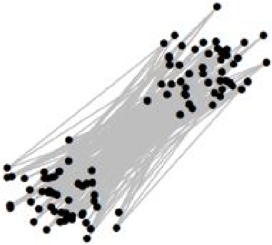
Cons of AME models

- An estimate of Ψ is an incomplete summary (it only describe the covariance of across node heterogeneity effects)
- Only provides limited summary of the potential network dependencies
- Higher-dependencies is somewhat opaque
- If primary goal is to evaluate specific types of higher-order network dependencies, ERMGs is more straightforward

Latent Variable Models

	Block Model	Latent distance model
Assumption	Each node belongs to an unobserved latent class or "block" "stochastic equivalence"	Each node has some unobserved location in a latent Euclidean "social space"
Relationships between nodes	Two nodes are determined (statistically) by their block memberships	The strength of a relation between two nodes is decreasing in the distance between them in latent space
How to define Membership	Members of the same group with the same distribution of relationships to other nodes	Closeness between two nodes
		Useful when there exists subgroups of nodes with strong within-group relations

Comparison between Latent Variable Models



Stochastic Block Model

Embedding of the nodes in Euclidean Space

Comparison between Latent Variable Models

- Stochastic Block Model
  - Within-group density of ties is lower than the between-group density.
  - In a latent distance model, two nodes are stochastically equivalent if they are in the same location in the social space.
- Latent Distance Model
  - The probability of a tie between two nodes is decreasing in the distance between them.
  - Blockmodel would require a large number of blocks.

Limitation of Latent Variable Models

- Real networks exhibit combinations of stochastic equivalence and transitivity in varying amounts
- Providing incomplete description of the heterogeneity across nodes
- How to solve?
  - Latent variable models based on multiplicative effects
  - Represent both types of network patterns
  - Generalization of both models

Simple Case Example

- Example, the simple case of an undirected dyadic variable; sociomatrix is symmetric
- $y_{i,j} \sim m_{i,j} + \alpha(u_i, u_j)$
    - α is function of node-specific latent variables $u_1, ...u_n$.
    - $m_{i,j}$ consists of any other terms in the model (e.g, a regression term or additive effects)
    - $y \sim x$ means that the distribution of y is stochastically increasing in x.
- Let's compare with the following three specification of the function α

Comparison with Multiplicative Effects Model

- Stochastic blockmodel: $\alpha(u_i, u_j) = u_i^T \Theta u_j$, where $u_i \in \mathbb{R}^r$ is a standard vector indicating block membership, Θ is $r \times r$ symmetric
- Latent distance model: $\alpha(u_i, u_j) = -|u_i - u_j|$, where $u_i \in \mathbb{R}^r$.
- Multiplicative effects model : $\alpha(u_i, u_j) = u_i^T \Lambda u_j$, where $u_i \in \mathbb{R}^r$ and Λ is $r \times r$ diagonal matrix.
- Similary for asymmetric AME models : u_i's and v_i's range over r-dimensional space, the multiplicative term UV^T ranges over the space of all $n \times n$ rank r matrices α

Gibbs Sampling for the AME

$$Y = M(X, \beta) + UV^T + a1^T + 1b^T + E$$

- a, b, E follows the social relations covariance model with parameters Σ, σ^2, ρ
- Let $(u_i, v_i) \sim N_{2r}(0, \Psi)$ independently across models ; $\Psi^{-1} \sim Wishart(\Psi_0^{-1}/\kappa_0, \kappa_0)$ as a prior

Gibbs Sampling for the AME

$$Y = M(X, \beta) + UV^T + a1^T + 1b^T + E$$

- Gibbs sampling
  1. Simulate β, a, b given $Y - UV^T, \Sigma, \sigma^2, \rho$
  2. Simulate σ^2 given $Y - UV^T, \beta, a, b, \rho$
  3. Simulate ρ given $Y - UV^T, \beta, a, b, \sigma^2$
  4. Simulate Σ given a, b
  5. Simulate missing values of $Y - UV^T$ given $\beta, a, b, \sigma^2, \rho$ and observed values of $Y - UV^T$

Summary

- The AME framework is a modular approach for network data analysis based on three statistical models

The AME framework	$\approx$	Applied statistician
1. The social relations covariance model		1. Linear random effects model
2. Low-rank matrix representations via multiplicative factors		2. Model-based singular value decomposition
3. Gaussian transformation models		3. Basis of many binary and ordinal regression models

Future Works

- Generalizing the framework to multiple sociomatrices on one or more nodesets
  - Multiple populations / multiple time points / multiple dyadic variables / or combinations of these
- Computational Improvement
  - Other integral approximation methods
  - Pseudo likelihood and method of moments estimation for blockmodels
  - Case-control likelihood approximation for the latent distance model