Statistical analysis of network data lecture 8

Sofia Olhede

EPFL

November 13, 2024

ERGMs

- Exponential random graph models are specifying a family of probability distributions on graphs.
- Depending on the application, we may model simple, loopy, multiple-edged, weighted or directed graphs.
- Let $\mathscr{G}_n$ be the set of all graphs on n vertices. Consider the following model

$$\Pr\{\mathcal{G} = G\} = \exp\left\{\sum_{i=1}^{k} \theta_i T_i(G) - c(\theta)\right\},$$

where θ_i, $1 \leq i \leq k$ are real valued parameters, and T_i are real valued statistic defined on $\mathscr{G}_n$; $c(\theta)$ is a normalizing constant.

ERGMs II

- ERGMs can be used to model relationships in sociology.
- Sometimes T_j is chosen to be d,
- Sometimes we look at the total number of edges,
- the number of triangles, or other sub-graph counts,
- the number of connected components in the graph.

ERGMs III

- One possible example is

$$\Pr\{\mathcal{G} = G\} = \exp\Big\{n^2\big(\beta_1 \hat{t}(K_2, G) + \beta_2 \hat{t}(K_3, G)\big) - c(\beta)\Big\},$$

with $\hat{t}(F, G) = X_F(G)/\big(n_{|v(F)|/\,\mathrm{aut}(F)}\big)$, and

$$\exp(c(\beta)) = \sum_{G \in \mathscr{G}_n} \exp\Big\{n^2\big(\beta_1 \hat{t}(K_2, G) + \beta_2 \hat{t}(K_3, G)\big)\Big\}.$$

- This is a special case of a **Gibbs distribution**. β_1 is known as the *particle parameter* and β_2 is known as the *energy parameter*.
- The exponential model is said to be *attractive* if β_2 is positive and *repulsive* if β_2 is negative.

ERGMs IV

- When β_2 is positive then Chatterjee and others have shown a typical graph drawn from the standard edge-triangle exponential random graph model always looks like an Erdos-Rényi random graph or a mixture of Erdos-Rényi random graphs.
- By raising the triangle density to an exponent $\gamma > 0$, Lubetzky and Zhao proposed a natural generalization:

$$\Pr\{\mathcal{G} = G\} = \exp\left\{n^2\left(\beta_1 \hat{t}(K_2, G) + \beta_2 \hat{t}^{\gamma}(K_3, G)\right) - c'(\beta)\right\},$$

  which enabled the model to exhibit a nontrivial structure even when β_2 is positive.
- When $\gamma \geq 2/3$ then this generalized model still features the Erdos-Rényi behavior; but for $\gamma < 2/3$, there exist regions of values of (β_1, β_2) for which a typical graph drawn from the model has symmetry breaking.

ERGMs V

- When Chatterjee introduced the framework to study the asymptotics, a lot was added to our understanding of ERGMs.
- The degeneracy of the ERGM had already been observed by practitioners, see for example the paper by David Hunter in JASA (2012). Hunter studied the behavior of friendship networks in US schools, and observed interesting features with network growth.

Latent Space I

- The latent space model (lecture 1) assumes a type of **conditional independence**.
- We assume (see e.g. Hoff (2002)) that conditionally on unobserved latent variables z_i (unobserved positions in a latent space) and x_{ij} we have
$$P(A \mid \mathbf{z}, \mathbf{X}, \boldsymbol{\theta}) = \prod_{i<j} \Pr\{a_{ij} \mid z_i, z_j, x_{ij}, \theta\}.$$
- A convenient way to parameterize $\Pr\{a_{ij} \mid z_i, z_j, x_{ij}, \theta\}$ is to use logistic regression.
- For a simple Bernoulli, this corresponds to parameterizing y_i in terms of its success probability p_i in terms of covariate x_i to form predictor $\eta_i = \eta(x_i)$:
$$\Pr\{Y_i = 1\} = \mathbb{E}(Y_i) = \mu_i = p_i, \quad g(\mu_i) = \eta_i.$$
Here $g(\mu) = \text{logit}(p_i) = \log(p_i/(1-p_i))$, and so can take any value in $\mathbb{R}$.

Latent Space II

- We take as $\eta_{ij} = \gamma + \beta^T x_{ij} - |z_i - z_j|$.
- For two different nodes, j and k, equidistant in the latent space, their log-odds ratio is determined by the covariate and so the difference is $\beta^T(x_{ij} - x_{ik})$.
- Hoff writes $d_{ij} = |z_i - z_j|$ and considers the network to be **represented** by d_{ij} if

$$d_{ij} > 1 \quad \Leftrightarrow a_{ij} = 0, \quad d_{ij} < 1 \quad \Leftrightarrow a_{ij} = 1.$$

- This is in some sense equivalent to the **geometric random graph**. The network is **representable** if such distances can be formed, e.g. the z_i in some underlying space with some dimensionality can be constructed.
- We can show by removing the covariate that the model becomes equivalent to Chung-Lu's modelling framework.

Latent Space III

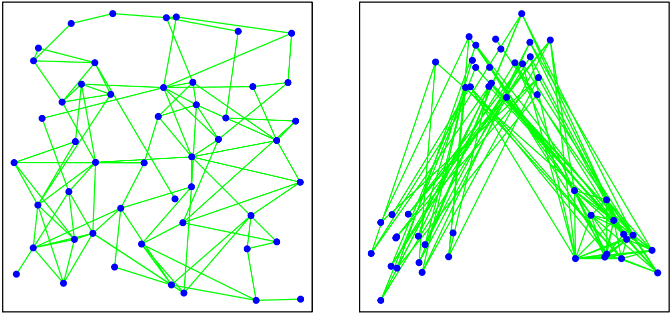
- Hoff has generalized this. He replaces logit by probit model:

$$\Pr\{A_{ij}=1\}=\mathbb{E}(A_{ij})=\mu_{ij}=p_{ij}, \quad g(\mu_{ij})=\eta_{ij},$$

with $g(\mu)=\Phi^{-1}(\mu)$ with $\eta_{ij}=\gamma+\beta^T x_{ij}+\alpha(z_i,z_j)$.

1. **Latent class/stochastic block model:**
   $\alpha(z_i,z_j)=\theta_{z_i,z_j}, \quad z_i\in\{1,\ldots,K\},\Theta=(\theta_{a,b})$. This encapsulates **stochastic equivalence**, a type of pattern often seen in network data in which the nodes can be divided into groups such that members of the same group have similar patterns of relationships.
2. **Latent distance model:**
   $\alpha(z_i,z_j)=-|z_i-z_j|, \quad z_i\in\mathbb{R}^K$. This model encapsulates **homophily**, so relationships between nodes with similar characteristics are stronger than those having different characteristics. Homophily explains **transitivity** ("a friend of a friend is a friend"), **balance** ("the enemy of my friend is an enemy") and cohesive subgroups of nodes.

Latent Space IV



Picture from Hoff 2007.

Latent Space V

3. **Latent eigenmodel:**
   $\alpha(z_i, z_j) = {z_i}^T \wedge z_j, \quad z_i \in \mathbb{R}^K$ and Λ a $K \times K$ matrix. An interpretation of the latent eigenmodel is that each node i has a vector of unobserved characteristics as given by z_i. Depending on the eigenvalues of Λ similar values of z_j will lead to contributing positively or negatively to the relationship between i and j.

- The eigenmodel with K-length vectors generalizes the latent class model.
- The eigenmodel with $K + 1$-length vectors almost generalizes the latent distance model with K dimensional latents.