EVERY CORRELATION EVERYWHERE ALL AT ONCE: MODELING RECIPROCITY IN MULTI-LAYER DIRECTED SOCIAL NETWORKS USING OBSERVATIONAL AND SELF-REPORT DATA

RICHARD MCELREATH

April 13, 2023

1. Goal

Suppose a community of N individuals who may practice M different directed behaviors (layers) towards one another. Reciprocity could exist within a dyad $[i,j]$ both within each behavior (i and j are balanced for a given layer) and between different behaviors (i and j are balanced between layers 1 and 2). Our goal is to infer these different reciprocities at the dyad level.

We also anticipate different kinds of data about each layer, both observations of directed behavior and self-report. These data have different reliabilities, and self-report is prone to disagreement within dyads. Therefore we need to model error in self-report on top of the underlying network.

2. Strategy

We have two problems to solve. First, how to add layers to an SRM. Second, how to use both observations and self-report while respecting the different reliabilities of each.

2.1. **Layers.** We develop a layered social relations model (SRM) for any arbitrary number of directed behaviors. In a typical SRM, there is a single directed behavior. A single correlation parameter is sufficient to model balance in behavior between individuals within dyads:

$$\begin{pmatrix} d_{ij} \\ d_{ji} \end{pmatrix} \sim \text{MVNormal} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$

where d_{ij} and d_{ji} are dyad-specific varying effects for the dyad comprising individuals i and j.

DEPARTMENT OF HUMAN BEHAVIOR, ECOLOGY AND CULTURE, MAX PLANCK INSTITUTE FOR EVOLUTIONARY ANTHROPOLOGY, LEIPZIG, GERMANY

E-mail address: richard_mcelreath@eva.mpg.de.

21   To extend this to more than one behavior (layer), we need a larger covariance structure. Consider two behaviors. Now we require four varying effects for each dyad, each expressing the offset for a behavior k and a direction ij
24 or ji. It also implies four correlation parameters:

   (1) ρ_{W12}: The correlation within an individual between layers. Do individuals who perform behavior 1 also tend to perform behavior 2
27   within the same dyads?
   (2) ρ_{B1}: The correlation between individuals in layer 1. This is analogous to the ρ parameter in the single-layer SRM.
30 (3) ρ_{B2}: The correlation between individuals in layer 2.
   (4) ρ_{B12}: The correlation between individuals between layers 1 and 2. This allows for balance across layers. For example, it is possible that
33   individual i directs behavior 1 to individual j, but individual j directs only behavior 2 to individual i.

The resulting correlation matrix is highly structured. For ease of notation, let
36 w_{kl} indicate the within-person correlation between layers k and l. Let r_k be the between-person correlation in layer k. And let r_{kl} be the between-person correlation between layers k and l. Then for the 2-layer case we have:

$$\mathbf{R} = \begin{bmatrix} 1 & w_{12} & r_1 & r_{12} \\ & 1 & r_{12} & r_2 \\ & & 1 & w_{12} \\ & & & 1 \end{bmatrix}$$

39 This defines a vector of dyad varying effects $\mathbf{d} = \{d_{121}, d_{122}, d_{211}, d_{212}\}$, where the the indexes are: actor, recipient, layer.
   For M layers, we require:

42 • $\frac{M!}{2!(M-2)!}$ within-person w_{kl} parameters
   • $\frac{M!}{2!(M-2)!}$ between-person, within-layer r_k parameters
   • $\frac{M!}{2!(M-2)!}$ between-person, between-layer r_{kl} parameters

45 So for M layers, we require a total of $\frac{3M!}{2!(M-2)!}$ correlation parameters. For $M = 3$, we require 9. For $M = 5$, we require 30. The resulting correlation matrix has dimension $2M$. Just as an example, here is $\mathbf{R}$ for $M = 3$:

$$\mathbf{R} = \begin{bmatrix} 1 & w_{12} & w_{13} & r_1 & r_{12} & r_{13} \\ & 1 & w_{23} & r_{12} & r_2 & r_{23} \\ & & 1 & r_{13} & r_{23} & r_3 \\ & & & 1 & w_{12} & w_{13} \\ & & & & 1 & w_{23} \\ & & & & & 1 \end{bmatrix}$$

In principle, each layer could have a unique standard deviation, implying another M scale parameters used in constructing the covariance matrix for the $2M$ varying effects.

The matrix $\mathbf{R}$ is symmetric both in the usual way for a correlation matrix, the upper and lower triangles are reflections of one another, but also in the other direction, along the other diagonal from bottom-left to upper-right. There is a sub-matrix in the upper-left for the w correlations and another in the lower-right that are identical, because the individuals i and j in the dyad are exchangeable. And the sub-matrix in the upper-right, for the r parameters, could be reflected along the bizarro-diagonal because the layers are exchangeable. This makes it easy to generate the matrix for any M, by construction first the two separate sub-matrices. Just for fun, here's $M = 9$:

$$\begin{bmatrix}
1 & w_{12} & w_{13} & w_{14} & w_{15} & w_{16} & w_{17} & w_{18} & w_{19} & r_1 & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} & r_{17} & r_{18} & r_{19} \\
 & 1 & w_{23} & w_{24} & w_{25} & w_{26} & w_{27} & w_{28} & w_{29} & r_{12} & r_2 & r_{23} & r_{24} & r_{25} & r_{26} & r_{27} & r_{28} & r_{29} \\
 & & 1 & w_{34} & w_{35} & w_{36} & w_{37} & w_{38} & w_{39} & r_{13} & r_{23} & r_3 & r_{34} & r_{35} & r_{36} & r_{37} & r_{38} & r_{39} \\
 & & & 1 & w_{45} & w_{46} & w_{47} & w_{48} & w_{49} & r_{14} & r_{24} & r_{34} & r_4 & r_{45} & r_{46} & r_{47} & r_{48} & r_{49} \\
 & & & & 1 & w_{56} & w_{57} & w_{58} & w_{59} & r_{15} & r_{25} & r_{35} & r_{45} & r_5 & r_{56} & r_{57} & r_{58} & r_{59} \\
 & & & & & 1 & w_{67} & w_{68} & w_{69} & r_{16} & r_{26} & r_{36} & r_{46} & r_{56} & r_6 & r_{67} & r_{68} & r_{69} \\
 & & & & & & 1 & w_{78} & w_{79} & r_{17} & r_{27} & r_{37} & r_{47} & r_{57} & r_{67} & r_7 & r_{78} & r_{79} \\
 & & & & & & & 1 & w_{89} & r_{18} & r_{28} & r_{38} & r_{48} & r_{58} & r_{68} & r_{78} & r_8 & r_{89} \\
 & & & & & & & & 1 & r_{19} & r_{29} & r_{39} & r_{49} & r_{59} & r_{69} & r_{79} & r_{89} & r_9 \\
 & & & & & & & & & 1 & w_{12} & w_{13} & w_{14} & w_{15} & w_{16} & w_{17} & w_{18} & w_{19} \\
 & & & & & & & & & & 1 & w_{23} & w_{24} & w_{25} & w_{26} & w_{27} & w_{28} & w_{29} \\
 & & & & & & & & & & & 1 & w_{34} & w_{35} & w_{36} & w_{37} & w_{38} & w_{39} \\
 & & & & & & & & & & & & 1 & w_{45} & w_{46} & w_{47} & w_{48} & w_{49} \\
 & & & & & & & & & & & & & 1 & w_{56} & w_{57} & w_{58} & w_{59} \\
 & & & & & & & & & & & & & & 1 & w_{67} & w_{68} & w_{69} \\
 & & & & & & & & & & & & & & & 1 & w_{78} & w_{79} \\
 & & & & & & & & & & & & & & & & 1 & w_{89} \\
 & & & & & & & & & & & & & & & & & 1
\end{bmatrix}$$

Once $\mathbf{R}$ is defined, the remainder of the model is a standard SRM. For example, for Poisson observations of behavior in each layer:

$$Y_{ijk} \sim \text{Poisson}(\lambda_{ijk})$$
$$\log \lambda_{ijk} = \alpha_k + G_{ik} + R_{jk} + d_{ijk}$$
$$\mathbf{d} \sim \text{MVNormal}(0, \mathbf{SRS})$$

where $\mathbf{S}$ is an appropriate vector of scale parameters for the layers and G_{ik} and R_{jk} are standard SRM generalized giving and receiving offsets in layer k.

The size of the matrix $\mathbf{R}$ presents challenges for defining coherent prior correlation matrices, since the w and r parameters cannot in principle be

independent and still construct valid correlation matrices. In practice, independent priors can be assigned to each w and r, but prior predictive simulations are necessary to understand what such a prior actually implies in the sample space. As M grows, priors may need to be increasingly narrow (and centered on zero) so that the joint prior covers valid correlation matrices.

An approach that seems to work well is to define independent LKJ correlation matrix priors for the w sub-matrix, the r sub-matrix without its within-layer diagonal (top-left to bottom-right), and the r within-layer diagonal. This does not ensure positive semidefinite matrices, once the sub-matrices are composed into the full matrix. But it is good enough to allow the sampler to get started and it seems to function well on synthetic data.

2.2. General behavior.

The terms G_{ik} and R_{jk} model general behavior that is not dependent upon network ties (the d parameters). In a single-layer SRM, these varying effects are simply individual-level varying effects with a two-dimensional Gaussian prior. With $M > 1$, this prior is naturally of higher dimension. We require two parameters for each layer, so $2M$ in total for each individual in the sample, and if every correlation is of interest, a covariance matrix of dimension $2M$ as well. Or each layer can be assigned an independent 2D Gaussian prior, if correlations across layers are not of interest or value in partial pooling.

2.3. Self-report data.

In the previous sections, we've outlined an estimator for the rate λ_{ijk} of observed directed behavior k from individual i to j. Now we anticipate having in addition to direct observation of k self-reports from both i and j. Self-report data present the additional challenge of being unreliable. People may under-report behavior or over-report behavior in the network. And these biases could be directional, with for example under-reporting of in-degree and over-reporting of out-degree.

When there is observational data to ground-truth self-report with, for at least some individuals, then we can use this information to improve the value of self-reported behavior. Also, when self-report is double-sampled so that both members of a dyad report on the same directed edge, then it is sometimes possible to estimate the reliability this way as well. We aim to do both in the same model.

Let Y_{ijk} be observed behavior from i to j in layer k. Let $X_{ijk,i}$ be i's self-report of behavior to j in layer k. Let $X_{ijk,j}$ be similarly j's report of i's behavior to j in layer k. We consider the case where X is binary, but the approach could be extended to count data.

The most general model for X is to adapt an item-response framework. In principle X is an item response that reflects Y through unobserved aspects of the reporting individual. Let $\ell_{ijk} = \log \lambda_{ijk}$ be the expected log rate of

behavior from i to j in layer k, as defined in the previous sections. Then the log-odds that i reports out-directed behavior, i.e. $X_{ijk,i} = 1$, is:

$$\operatorname{logit} p_{ijk,i} = \delta_{ik}(\ell_{ijk} - \beta_{ik}) + \gamma_{ik}$$

The self-report process is governed by three latent parameters.

(1) β_{ik}: Individual i's threshold for reporting behavior. Larger values of β require larger values of ℓ for a report.

(2) δ_k: Individual i's discrimination. This is a positive real value that determines how quickly the probability of a report increases as ℓ increases. When $\delta = 0$, or is just very small, then reports are essentially random with respect to ℓ, the real rate of behavior.

(3) γ_{ik}: Individual i's tendency to over- or under-report behavior across all dyads. If $\delta = 0$, then γ_{ik} provides a base rate of reporting. If it is omitted, then the "guessing" rate is always one-half.

We also need to consider reports of received behavior. So these parameters need an additional subscript for direction, 1 for out and 2 for in. This gives us:

$$\operatorname{logit} p_{ijk,i} = \delta_{ik1}(\ell_{ijk} - \beta_{ik1}) + \gamma_{ik1}$$
$$\operatorname{logit} p_{jik,i} = \delta_{ik2}(\ell_{jik} - \beta_{ik2}) + \gamma_{ik2}$$

Obviously we cannot estimate all of these. So either informative priors or simpler formulations are necessary. However it is important to outline the general case and be transparent about which decisions are made to make the model tractable. For example, δ could vary only by layer, not by individual. But this isn't as bad as it would be in an ordinary IRT, because the variable ℓ_{ijk} is a sum of parameters that are informed by observed behavior as well.