Estimating network-mediated causal effects via spectral embeddings

Alex Hayes [1] Mark M. Fredrickson [2] Keith Levin [1]

[1]Department of Statistics, University of Wisconsin-Madison [2]Department of Statistics, University of Michigan

Abstract

We consider the task of mediation analysis for network data, and present a model in which mediation occurs in a latent embedding space. Under this model, node-level interventions have causal effects on nodal outcomes, and these effects can be partitioned into a direct effect independent of the network, and an indirect effect induced by homophily.

Motivating example: smoking in adolescent social networks

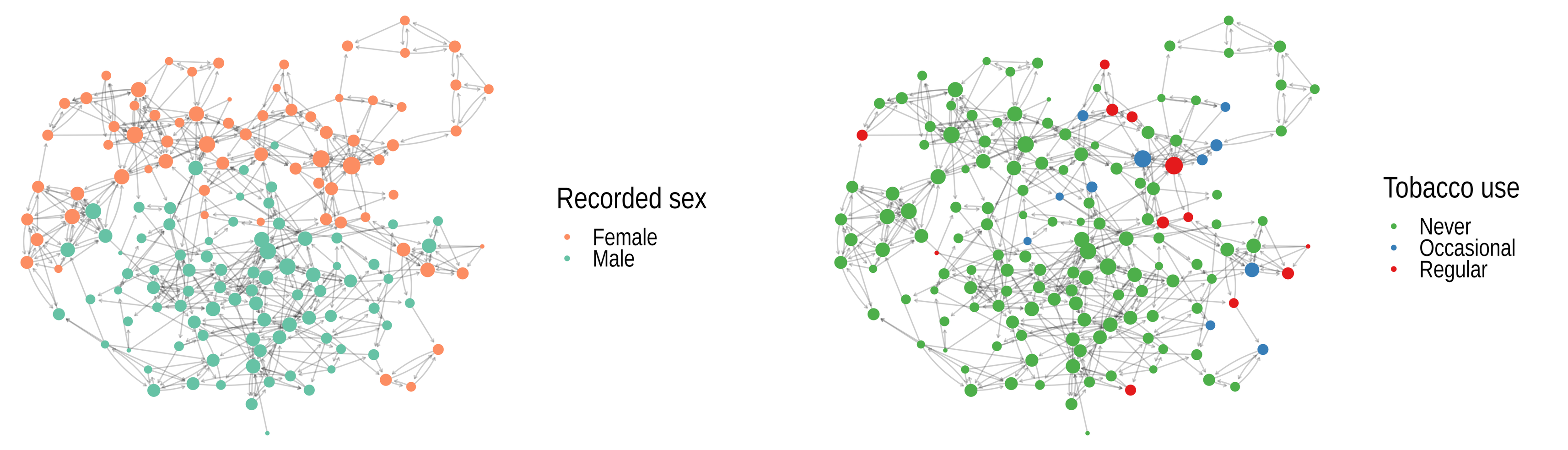

Figure 1. Directed friendships in a secondary school in Glasgow, reported in the Teenage Friends and Lifestyle Study (wave 1). Each node represents one student.

Notation & inferential targets

We assume we have a (symmetric) network with nodes $1, \ldots, n$.

Network adjacency matrix	A	$\in \mathbb{R}^{n \times n}$
Edge $i \sim j$	A_{ij}	$\in \mathbb{R}$
Treatment	T_i	$\in \{0, 1\}$
Outcome	Y_i	$\in \mathbb{R}$
Confounders	$C_{i \cdot}$	$\in \mathbb{R}^p$
Friend group (latent)	$X_{i \cdot}$	$\in \mathbb{R}^d$

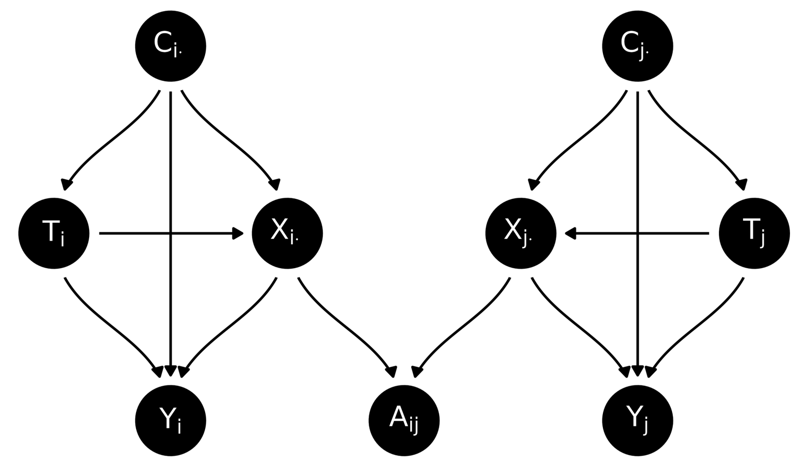

Figure 2. A directed acyclic graph (DAG) representing the causal pathways in a network with homophilous mediation, for node a network with two nodes called i and j.

We are interested in the causal effect of T_i on Y_i as mediated by the latent position $X_{i \cdot}$. More precisely, we want to estimate the *natural direct effect* and the *natural indirect effect*

$$\Psi_{\text{nde}}(t, t^*) = \mathbb{E}[Y_i(t, X_{i \cdot}(t^*)) - Y_i(t^*, X_{i \cdot}(t^*))]$$
$$\Psi_{\text{nie}}(t, t^*) = \mathbb{E}[Y_i(t, X_{i \cdot}(t)) - Y_i(t, X_{i \cdot}(t^*))]$$

Semi-parametric network model

Let $A \in \mathbb{R}^{n \times n}$ be a random symmetric matrix, such as the adjacency matrix of an undirected graph. Let $P = \mathbb{E}[A \mid X] = XX^T$ be the expectation of A conditional on $X \in \mathbb{R}^{n \times d}$, which has independent and identically distributed rows $X_{1 \cdot}, \ldots, X_{n \cdot}$. That is, P has $\operatorname{rank}(P) = d$ and is positive semi-definite with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d > 0 = \lambda_{d+1} = \cdots = \lambda_n$. Conditional on X, the upper-triangular elements of $A - P$ are independent (ν_n, b_n)-sub-gamma random variables.

Examples: (degree-corrected) stochastic blockmodels, mixed-membership blockmodels, overlapping blockmodels, (weighted, noisily-observed) random dot product graphs, LDA, factor models, etc

The outcome regression functional is linear in $T_i, C_{i \cdot}$, and $X_{i \cdot}$ and the mediator regression functional is linear in $T_i, C_{i \cdot}$, and $T_i \cdot C_{i \cdot}$:

$$\underbrace{\mathbb{E}[Y_i \mid T_i, C_{i \cdot}, X_{i \cdot}]}_{\mathbb{R}} = \underbrace{\beta_0}_{\mathbb{R}} + \underbrace{T_i}_{\{0,1\}} \underbrace{\beta_{\text{t}}}_{\mathbb{R}} + \underbrace{C_{i \cdot}}_{\mathbb{R}^{1 \times p}} \underbrace{\beta_{\text{c}}}_{\mathbb{R}^p} + \underbrace{X_{i \cdot}}_{\mathbb{R}^{1 \times d}} \underbrace{\beta_{\text{x}}}_{\mathbb{R}^d}, \qquad \text{(outcome model)}$$

$$\underbrace{\mathbb{E}[X_{i \cdot} \mid T_i, C_{i \cdot}]}_{\mathbb{R}^{1 \times d}} = \underbrace{\theta_0}_{\mathbb{R}^{1 \times d}} + \underbrace{T_i}_{\{0,1\}} \underbrace{\theta_{\text{t}}}_{\mathbb{R}^{1 \times d}} + \underbrace{C_{i \cdot}}_{\mathbb{R}^{1 \times p}} \underbrace{\Theta_{\text{c}}}_{\mathbb{R}^{p \times d}} + \underbrace{T_i}_{\{0,1\}} \underbrace{C_{i \cdot}}_{\mathbb{R}^{1 \times p}} \underbrace{\Theta_{\text{tc}}}_{\mathbb{R}^{p \times d}}. \qquad \text{(mediator model)}$$

Under these moment assumptions, and DAG of Figure 2, letting μ_c denote the mean of $C_{i \cdot}$, we have the following identification result:

$$\Psi_{\text{nde}}(t, t^*) = (t - t^*) \, \beta_{\text{t}}, \quad \text{and}$$
$$\Psi_{\text{nie}}(t, t^*) = (t - t^*) \, \theta_{\text{t}} \, \beta_{\text{x}} + (t - t^*) \, \mu_c \, \Theta_{\text{tc}} \, \beta_{\text{x}}.$$

Estimation challenge: friend groups X unknown!

The *adjacency spectral embedding* (ASE) of A is well-known to be a good estimate of X under a broad, semi-parametric class of network models. Given a network with adjacency matrix A, the d-dimensional ASE is defined as

$$\widehat{X} = \widehat{U} \widehat{S}^{1/2} \in \mathbb{R}^{n \times d},$$

where $\widehat{U} \widehat{S} \widehat{U}^T$ is the rank-d truncated singular value decomposition of A. Under a suitably well-behaved network model, if d is correctly specified or consistently estimated, there is some $d \times d$ orthogonal matrix Q such that

$$\max_{i \in [n]} \left\| \widehat{X}_{i \cdot} - X_{i \cdot} Q \right\| = o_p(1).$$

Estimation: plug in $\widehat{X}$ for X

Let $\widehat{D} = \begin{bmatrix} 1 & T & C & \widehat{X} \end{bmatrix} \in \mathbb{R}^{n \times (2+p+d)}$ and $L = \begin{bmatrix} 1 & T & C & T \cdot C \end{bmatrix} \in \mathbb{R}^{n \times (2p+2)}$. We estimate β_{w} and β_{x} via ordinary least squares as follows

$$\begin{bmatrix} \widehat{\beta}_{\text{w}} \\ \widehat{\beta}_{\text{x}} \end{bmatrix} = \left(\widehat{D}^T \widehat{D} \right)^{-1} \widehat{D}^T Y.$$

Similarly, we estimate Θ via ordinary least squares as

$$\widehat{\Theta} = \left(L^T L \right)^{-1} L^T \widehat{X}.$$

To estimate Ψ_{nde} and Ψ_{nie}, we combine regression coefficients from the network regression models

$$\widehat{\Psi}_{\text{cde}} = \widehat{\Psi}_{\text{nde}} = (t - t^*) \, \widehat{\beta}_{\text{t}} \qquad \text{and}$$
$$\widehat{\Psi}_{\text{nie}} = (t - t^*) \, \widehat{\theta}_{\text{t}} \, \widehat{\beta}_{\text{x}} + (t - t^*) \cdot \widehat{\mu}_c \cdot \widehat{\Theta}_{\text{tc}} \, \widehat{\beta}_{\text{x}},$$

where $\widehat{\mu}_c$ is the sample mean of $C_{i \cdot}$.

Theory

Under a suitable network model and moment bounds on the regression errors, there exists a sequence of orthogonal matrices $\{Q_n\}_{n=1}^{\infty}$ such that

$$\sqrt{n} \, \widehat{\Sigma}_{\text{vec}(\Theta)}^{-1/2} \left(\operatorname{vec}\left(\widehat{\Theta} \, Q_n^T \right) - \operatorname{vec}(\Theta) \right) \to \mathcal{N}(0, I_{pd}), \text{ and}$$
$$\sqrt{n} \, \widehat{\Sigma}_{\beta}^{-1/2} \begin{pmatrix} \widehat{\beta}_{\text{w}} - \beta_{\text{w}} \\ Q_n \, \widehat{\beta}_{\text{x}} - \beta_{\text{x}} \end{pmatrix} \to \mathcal{N}(0, I_d).$$

Further,

$$\sqrt{n \, \widehat{\sigma}^2_{\text{nde}}} \left(\widehat{\Psi}_{\text{nde}} - \Psi_{\text{nde}} \right) \to \mathcal{N}(0, 1), \text{ and}$$
$$\sqrt{n \, \widehat{\sigma}^2_{\text{nie}}} \left(\widehat{\Psi}_{\text{nie}} - \Psi_{\text{nie}} \right) \to \mathcal{N}(0, 1),$$

where $\widehat{\sigma}^2_{\text{nde}}$ and $\widehat{\sigma}^2_{\text{nie}}$ are derived via the Delta method.

Results applied to Glasgow data