


Estimating network-mediated causal effects via spectral embeddings

Alex Hayes, Mark M. Fredrickson, Keith Levin
Presented at ACIC 2023

Causal inference for network data is an area of active interest in the social sciences. Unfortunately, the complicated dependence structure of network data presents an obstacle to many causal inference procedures. 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. To estimate network-mediated effects, we embed nodes into a low-dimensional space and fit two regression models: (1) an outcome model describing how nodal outcomes vary with treatment, controls, and position in latent space; and (2) a mediator model describing how latent positions vary with treatment and controls. We prove that the estimated coefficients are asymptotically normal about the true coefficients under a sub-gamma generalization of the random dot product graph, a widely-used latent space model. We show that these coefficients can be used in product-of-coefficients estimators for causal inference. Our method is easy to implement, scales to networks with millions of edges, and can be extended to accommodate a variety of structured data.




UNIVERSITY OF MICHIGAN

Estimating network-mediated causal effects via spectral embeddings

Alex Hayes¹ Mark M. Fredrickson² Keith Levin¹

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



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

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 $1, \dots, n$.

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	$C_i \in \mathbb{R}^p$
Friend group (latent)	$X_i \in \mathbb{R}^d$

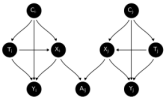


Figure 2: A directed acyclic graph (DAG) representing the causal pathways in a network with homophilous mediation, for node i and j .

We are interested in the causal effect of T_i on Y_i mediated by the latent position X_i . More precisely, we want to estimate the *natural direct effect* and the *natural indirect effect*

$$\Psi_{\text{nd}}(t, t') = \mathbb{E}[Y_i(t, X_i(t')) - Y_i(t', X_i(t'))]$$

$$\Psi_{\text{in}}(t, t') = \mathbb{E}[Y_i(t, X_i(t)) - Y_i(t, X_i(t'))]$$

Semi-parametric network model

Let $A \in \mathbb{R}^{n \times n}$ be a random matrix, such as the adjacency matrix of an undirected graph. Let $P = \mathbb{E}[A|X] = XX^T$ be the expectation of A conditional on $X \in \mathbb{R}^{n \times d}$, which has independent and identically distributed rows X_1, \dots, X_n . That is, P has $\text{rank}(P) = d$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0 = \lambda_{d+1} = \dots = \lambda_n$. Conditional on X , the upper-triangular elements of $A - P$ are independent (λ_{ij}, b_{ij}) -sub-gamma random variables.

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

The outcome regression functional is linear in T_i, C_i , and X_i , and the mediator regression functional is linear in T_i, C_i , and $T_i \cdot C_i$:

$$\mathbb{E}[Y_i | T_i, C_i, X_i] = \beta_0 + \frac{T_i}{n} \beta_1 + \frac{C_i}{n} \beta_2 + \frac{X_i}{n} \beta_3 \quad (\text{outcome model})$$

$$\mathbb{E}[X_i | T_i, C_i] = \frac{\theta_0}{n} + \frac{T_i}{n} \theta_1 + \frac{C_i}{n} \theta_2 + \frac{T_i \cdot C_i}{n} \theta_3 \quad (\text{mediator model})$$

Under these moment assumptions, and DAG of Figure 2, letting μ_i denote the mean of C_i , we have the following identification result:

$$\Psi_{\text{nd}}(t, t') = (t - t') \beta_1 \quad \text{and} \quad \Psi_{\text{in}}(t, t') = (t - t') \beta_1 \beta_3 + (t - t') \mu_i \theta_3 \beta_3$$

Estimation challenge: friend groups X unknown!

The adjacency spectral embedding (ASE) of A is well-known estimate of X under the network model described above, defined as

$$\hat{X} = \hat{D} \hat{S}^{1/2} \in \mathbb{R}^{n \times d}$$

where $\hat{D} \hat{S} \hat{D}^T$ is the rank- d truncated singular value decomposition of A . Under a suitably well-behaved model, if d is correctly specified, there is an orthogonal matrix Q such that

$$\max_{1 \leq i \leq n} \|\hat{X}_i - X_i Q\| = o_p(1).$$

Let $\hat{D} = [1 \ T \ C \ \hat{X}] \in \mathbb{R}^{n \times (2+p+d)}$ and $L = [1 \ T \ C \ T \cdot C] \in \mathbb{R}^{n \times (2+p+2)}$. We estimate β_0, β_1 and θ via ordinary least squares as follows

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = (\hat{D}^T \hat{D})^{-1} \hat{D}^T Y \quad \text{and} \quad \hat{\theta} = (L^T L)^{-1} L^T \hat{X}.$$

To estimate Ψ_{nd} and Ψ_{in} , let $\hat{\mu}_i$ be the sample mean of C_i , and combine regression coefficients from the network regression models

$$\hat{\Psi}_{\text{nd}} = \hat{\Psi}_{\text{nd}} = (t - t') \hat{\beta}_1 \quad \text{and} \quad \hat{\Psi}_{\text{in}} = (t - t') \hat{\beta}_1 \hat{\beta}_3 + (t - t') \cdot \hat{\mu}_i \cdot \hat{\theta}_3 \hat{\beta}_3$$

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

$$\sqrt{n} \frac{\hat{Q}_n^{1/2}}{\sqrt{\lambda_d}} \left(\text{vec}(\hat{\theta} Q_n^T) - \text{vec}(\theta) \right) \rightarrow \mathcal{N}(0, I_{p+d}), \quad \text{and}$$

$$\sqrt{n} \frac{\hat{Q}_n^{1/2}}{\sqrt{\lambda_d}} \left(\hat{\beta}_n - \beta_n \right) \rightarrow \mathcal{N}(0, I_d).$$

Further,

$$\sqrt{n} \hat{\sigma}_{\text{nd}}^2 \left(\hat{\Psi}_{\text{nd}} - \Psi_{\text{nd}} \right) \rightarrow \mathcal{N}(0, 1), \quad \text{and}$$

$$\sqrt{n} \hat{\sigma}_{\text{in}}^2 \left(\hat{\Psi}_{\text{in}} - \Psi_{\text{in}} \right) \rightarrow \mathcal{N}(0, 1),$$

where $\hat{\sigma}_{\text{nd}}^2$ and $\hat{\sigma}_{\text{in}}^2$ are derived via the Delta method.

Results applied to Glasgow data

- Estimated effects are adjusted for possible confounding by age and church attendance.
- Estimated effects vary with the chosen dimension d of the latent space.
- Over-specifying d is typically okay, but under-specifying d leads to a failure to capture social structure in X .
- Once we capture enough social structure in X , we see a significant indirect social effect that leads adolescent girls to smoke more.

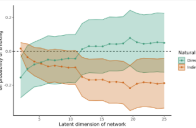


Figure 3: Estimated direct and indirect effects of sex on tobacco use in the Glasgow social network. Positive values indicate a greater propensity for adolescent boys to smoke, negative effects a greater propensity for adolescent girls to smoke.

References

Hayes, Alex, Mark M. Fredrickson, and Keith Levin. "Estimating Network-Mediated Causal Effects via Spectral Embeddings." arXiv, April 14, 2023. <http://arxiv.org/abs/2212.12041>.