Estimating network-mediated causal effects via spectral embeddings

Alex Hayes

2024-06-17 at NetSci 2014

Joint work with Mark Fredrickson and Keith Levin

Does mindfulness reduce psychological distress?

four preregistered mediators (Five Facet Mindfulness Questionnaire mindful action subscale, NIH Toolbox Loneliness, Drexel Defusion Scale, and Meaning in Life Questionnaire presence subscale).

1. I feel alone and apart from others
2. I feel left out
3. I feel that I am no longer close to anyone
4. I feel alone
5. I feel lonely

Never	Rarely	Sometimes	Usually	Always
1 ○	2 ○	3 ○	4 ○	5 ○

four preregistered mediators (Five Facet Mindfulness Questionnaire mindful action subscale, NIH Toolbox Loneliness, Drexel Defusion Scale, and Meaning in Life Questionnaire presence subscale).

1. I understand my life's meaning.
2. I am looking for something that makes my life feel meaningful.
3. I am always looking to find my life's purpose.
4. My life has a clear sense of purpose.
5. I have a good sense of what makes my life meaningful.
6. I have discovered a satisfying life purpose.
7. I am always searching for something that makes my life feel significant.
8. I am seeking a purpose or mission for my life
9. My life has no clear purpose.
10. I am searching for meaning in my life.

1 Absolutely untrue 2 Mostly untrue 3 Somewhat untrue 4 Can't say true or false 5 Somewhat true 6 Mostly true 7 Absolutely true

four preregistered mediators (Five Facet Mindfulness Questionnaire mindful action subscale, NIH Toolbox Loneliness, Drexel Defusion Scale, and Meaning in Life Questionnaire presence subscale).

1. When I do things, my mind wanders off and I'm easily distracted.
2. I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted.
3. I am easily distracted.
4. I find it difficult to stay focused on what's happening in the present.
5. It seems I am 'running on automatic' without much awareness of what I'm doing.
6. I rush through activities without being really attentive to them.
7. I do jobs or tasks automatically without being aware of what I'm doing.
8. I find myself doing things without paying attention.

1 Never or very rarely true 2 Rarely true 3 Sometimes true 4 Often true 5 Very often or always true

four preregistered mediators (Five Facet Mindfulness Questionnaire mindful action subscale, NIH Toolbox Loneliness, Drexel Defusion Scale, and Meaning in Life Questionnaire presence subscale).

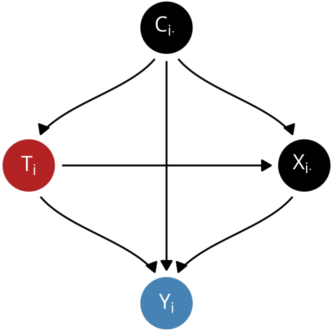
1. Feelings of anger. You become angry when someone takes your place in a long line. To what extent would you normally be able to defuse from feelings of anger?
2. Cravings for food. You see your favorite food and have the urge to eat it. To what extent would you normally be able to defuse from cravings for food?
3. Physical pain. Imagine that you bang your knee on a table leg. To what extent would you normally be able to defuse from physical pain?
4. Anxious thoughts. Things have not been going well at school or your job, and work just keeps piling up. To what extent would you normally be able to defuse from anxious thoughts like ”I’ll never get this done.”?
5. Thoughts of self. Imagine you are having a thought such as ”no one likes me.” To what extent would you normally be able to defuse from negative thoughts about yourself?

Causal mediation

Treatment	T_i	$\in \{0,1\}$
Outcome	Y_i	$\in \mathbb{R}$
Mediators	$X_{i\cdot}$	$\in \mathbb{R}^{1\times d}$
Confounders	$C_{i\cdot}$	$\in \mathbb{R}^{1\times p}$

Decompose effect of T_i on Y_i:

1. Effect operating along $T_i \to Y_i$ path (direct)
2. Effect operating along $T_i \to X_{i\cdot} \to Y_i$ path (indirect)



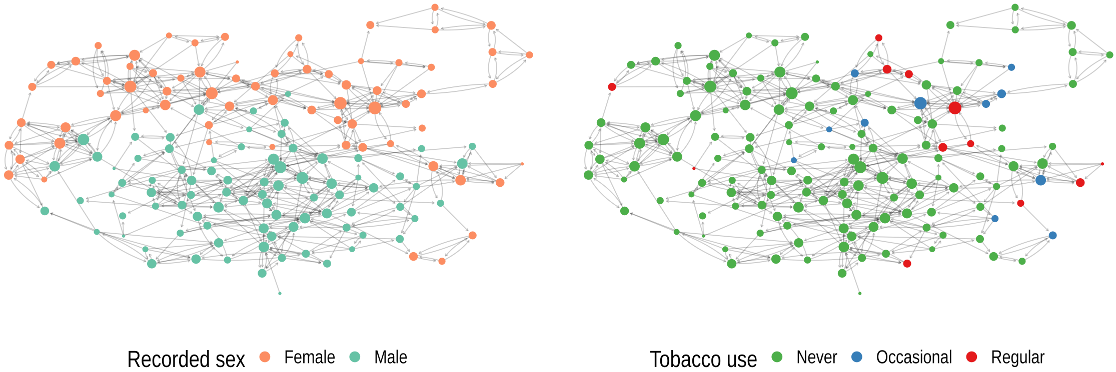
$$\Psi_{\text{ate}} = \Psi_{\text{nde}} + \Psi_{\text{nie}}$$

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

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

Social networks as mediators

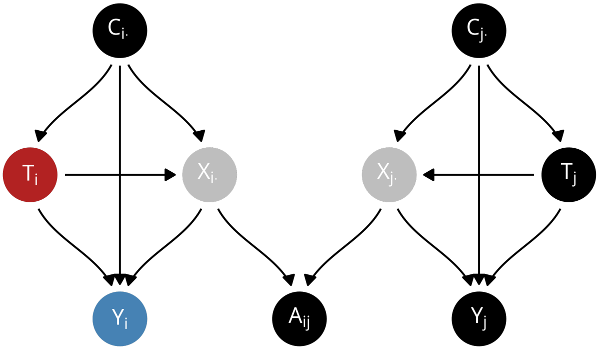
Mediation in an adolescent social network



Teenage Friends and Lifestyle Study (wave 1), Glasgow, 1996

We propose a model where social groups in networks mediate causal effects

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	$C_{i\cdot}$	$\in \mathbb{R}^{1 \times p}$
Friend group (latent)	$X_{i\cdot}$	$\in \mathbb{R}^{1 \times d}$



Structural causal model for network mediation in a network with two nodes i and j

The challenge: the friend groups $X_{i\cdot}$ are not observed!

Solution: estimate friend groups $X_{i\cdot}$ using community models!

We use a semi-parametric estimator that accommodates:

- Stochastic blockmodels
- Degree-corrected stochastic blockmodels
- Mixed-membership stochastic blockmodels
- Overlapping stochastic blockmodels
- Random dot product graphs
- Etc

Intuition: stochastic blockmodels

d communities or "blocks"

$X_{i\cdot} \in \{0,1\}^d$ one-hot indicator of node i's block

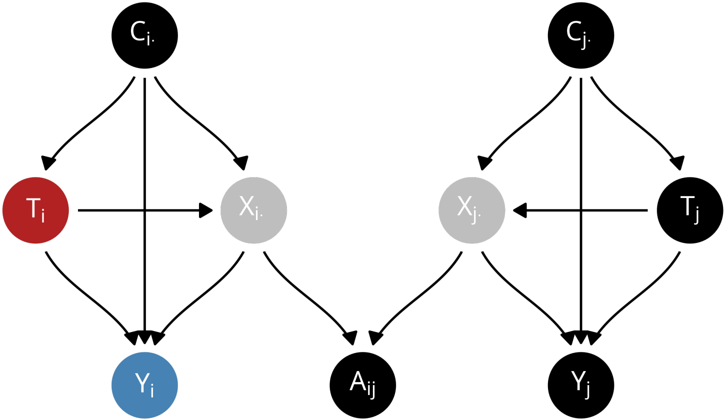
X is latent (i.e., unobserved)

$B \in [0,1]^{d \times d}$ inter-block edge probabilities

Friendships depend on group memberships and B

$\mathbb{P}(A_{ij} = 1 \mid X) = X_{i\cdot} B X_{j\cdot}^T$

Returning to the structural causal model for a moment



Don't know X but can estimate it!

Definition (ASE)

Given a network A, the d-dimensional adjacency spectral embedding of A is

$$\widehat{X} = \widehat{U}\widehat{S}^{1/2}$$

where $\widehat{U}\widehat{S}\widehat{U}^T$ is the rank-d truncated singular value decomposition of A.

Lemma

Under a suitable network model, there is a $d \times d$ orthogonal matrix Q such that

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

Note that we must correctly specify d

Semi-parametric regression models for mediation

If the previous DAG is correct, under the additional assumption that:

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

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

Then:

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

Estimation: plug $\widehat{X}$ into least squares

Let $\widehat{D} = \begin{bmatrix} 1 & T & C & \widehat{X} \end{bmatrix} \in \mathbb{R}^{n\times(2+p+d)}$ and $L = \begin{bmatrix} 1 & T & C & T\cdot C \end{bmatrix} \in \mathbb{R}^{n\times(2p+2)}$.

$$\begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_\mathsf{t} \\ \widehat{\beta}_\mathsf{c} \\ \widehat{\beta}_\mathsf{x} \end{bmatrix} = \left(\widehat{D}^T\widehat{D}\right)^{-1}\widehat{D}^T Y \quad \text{and} \quad \widehat{\Theta} = \left(L^T L\right)^{-1} L^T\widehat{X}.$$

$$\widehat{\Psi}_{\mathrm{nde}} = (t - t^*)\,\widehat{\beta}_\mathsf{t} \qquad \text{and}$$

$$\widehat{\Psi}_{\mathrm{nie}} = (t - t^*)\,\widehat{\theta}_\mathsf{t}\,\widehat{\beta}_\mathsf{x} + (t - t^*)\cdot\widehat{\mu}_c\cdot\widehat{\Theta}_{\mathsf{tc}}\,\widehat{\beta}_\mathsf{x}.$$

Main result

Theorem (Regression coefficients are asymptotically normal)

Under a suitably well-behaved network model and some moment conditions on regression errors, there is an unknown orthogonal matrix Q such that

$$\sqrt{n}\,\widehat{\Sigma}_{\beta}^{-1/2}\begin{pmatrix}\widehat{\beta}_w-\beta_w\\ Q\,\widehat{\beta}_x-\beta_x\end{pmatrix}\to\mathcal{N}(0,I_d),\ \textit{and}$$

$$\sqrt{n}\,\widehat{\Sigma}_{\mathrm{vec}(\Theta)}^{-1/2}\left(\mathrm{vec}\left(\widehat{\Theta}\,Q^T\right)-\mathrm{vec}(\Theta)\right)\to\mathcal{N}\left(0,I_{pd}\right).$$

where $\widehat{\Sigma}_{\mathrm{vec}(\Theta)}^{-1/2}$ *and* $\widehat{\Sigma}_{\beta}^{-1/2}$ *are the typical heteroscedasticity robust covariance estimators, with* $\widehat{X}$ *plugged in for X.*

Corollary

Theorem (Causal estimators are asymptotically normal)

Under the same statistical assumptions as before, plus mediating homophily,

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

where $\widehat{\sigma}^2_{\mathrm{nde}}$ and $\widehat{\sigma}^2_{\mathrm{nie}}$ are rather unfriendly variance estimators derived via the delta method and the previous theorem.

Application to Glasgow data

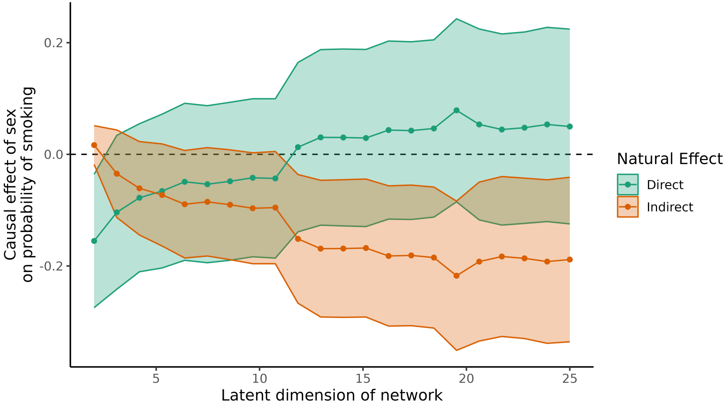


Figure 1: Estimated direct and indirect effects of sex on tobacco usage in the Glasgow social network, adjusted for age and church attendance. Positive values indicate a greater propensity for adolescent boys to smoke, negative effects a greater propensity for adolescent girls to smoke.

Thank you! Questions?

Read the manuscript at *https://arxiv.org/abs/2212.12041*

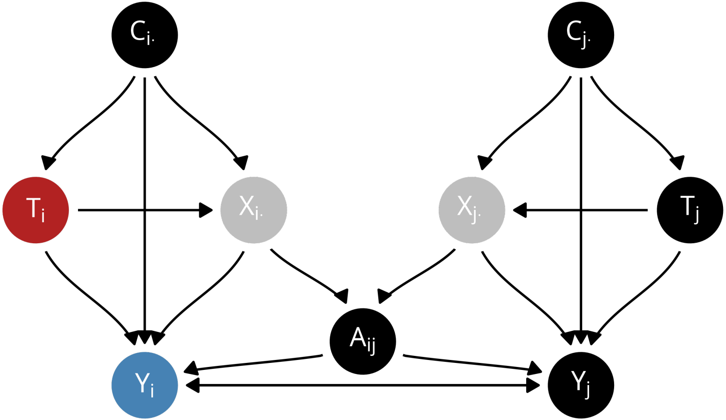
R package netmediate

Stay in touch

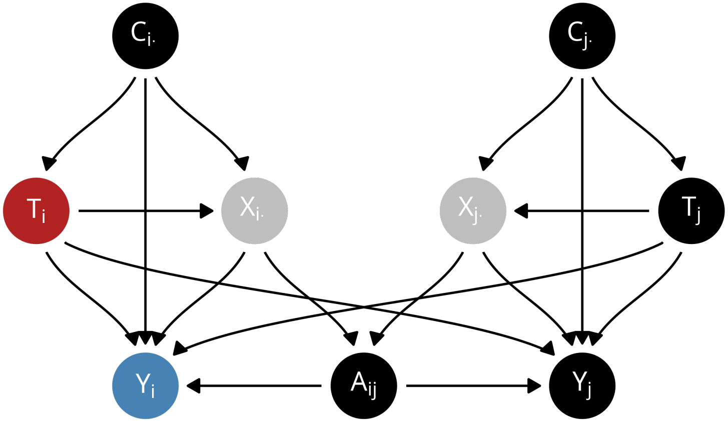
- @alexpghayes
- alex.hayes@wisc.edu
- *https://www.alexpghayes.com*
- *https://github.com/alexpghayes*

I'm looking for a post-doc starting Fall 2024, say hi if this work interests you!

Disambiguation: contagion ($Y_j \to Y_i$) is not allowed



Disambiguation: interference ($T_j \rightarrow Y_i$) is not allowed



More on interference and contagion

Interference and contagion effects are allowed so long as they happen in the latent space. Suppose

$$\mathbb{E}[Y_i \mid W_{i\cdot}, X_{i\cdot}] = W_{i\cdot}\beta_{\mathsf{w}} + X_{i\cdot}\beta_{\mathsf{x}}' + \delta_{\mathsf{y}} \sum_j X_{i\cdot}^T X_{j\cdot} Y_j$$

This latent space contagion model is a special parametric case of the regression outcome model (take $\beta_{\mathsf{x}} = \beta_{\mathsf{x}}' + X^T Y \delta_{\mathsf{y}}$).

Semi-parametric network model

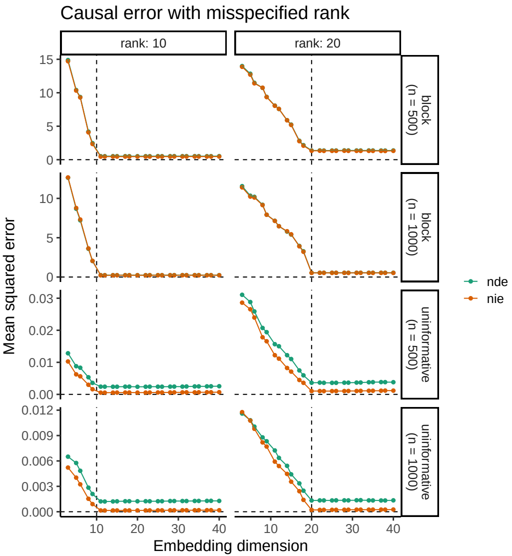
Definition

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 $X_{1\cdot}, \dots, X_{n\cdot}$. That is, P has $\text{rank}(P) = d$ 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.

Remark

$P = XX^T = (XQ)(XQ)^T$ for any $d \times d$ orthogonal matrix Q, the latent positions X are only identifiable up to an orthogonal transformation.

Choosing $\widehat{d}$: overestimating the embedding dimension is fine



Identifying assumptions

Definition

The random variables $(Y_i, Y_i(t,x), X_{i\cdot}, X_{i\cdot}(t), C_{i\cdot}, T_i)$ are independent over $i \in [n]$ and obey the following three properties.

1. Consistency:

$$\text{if } T_i = t\text{, then } X_{i\cdot}(t) = X_{i\cdot} \text{ with probability 1, and}$$
$$\text{if } T_i = t \text{ and } X_{i\cdot} = x\text{, then } Y_i(t,x) = Y_i \text{ with probability 1}$$

2. Sequential ignorability:

$$\{Y_i(t^*,x), X_{i\cdot}(t)\} \perp\!\!\!\perp T_i \mid C_{i\cdot} \quad \text{and} \quad \{Y_i(t^*,x)\} \perp\!\!\!\perp X_{i\cdot} \mid T_i = t, C_{i\cdot}$$

3. Positivity:

$$\mathbb{P}(x \mid T_i, C_{i\cdot}) > 0 \text{ for each } x \in \operatorname{supp}(X_{i\cdot})$$
$$\mathbb{P}(t \mid C_{i\cdot}) > 0 \text{ for each } t \in \operatorname{supp}(T_i)$$

Interventions allowed

Provided that controls $C_{i\cdot}$ are sufficiently informative about group membership $X_{i\cdot}$, treatment T_i is allowed to cause:

- Changes in popularity within a group
- Movement to a new friend group
- Becoming a member of a new friend group while remaining in current friend group
- Friendships becoming more or less likely between distinct friend groups
- Combinations of the above

See Appendix of manuscript for details.