

Causal inference with interference

Dawei Geng and Yunran Chen

December 6, 2018

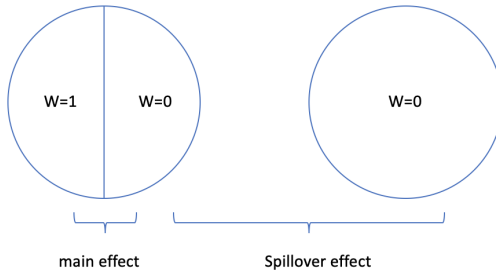
Outline

- 1 Introduction
 - Background
 - Goals
- 2 Problem formulation
 - Notations
 - Potential outcomes
 - Assignment mechanism
 - Causal estimands
- 3 GPS-based estimator
 - Definition of GPS
 - Properties of GPS
 - Estimating procedure
- 4 Simulation

Definition of interference

Interference

Potential outcomes of a unit depend on its treatment and on the treatments of other local units.



Problems and Goals

Problems:

- Potential outcomes (SUTVA \rightarrow SUTNVA)
- Assignment mechanisms (Unconfoundness \rightarrow extension)
- Estimand (Biased naive estimand \rightarrow extension)
- Estimation method (Propensity Score \rightarrow Generalized PS)

Goals:

- Known network + unknown assignment mechanism
- Estimate: main effect and spillover effect

Notations

- Treatment $W_i \in \{0, 1\}$
- Outcome $Y_i \in \mathcal{Y}$
- Covariate $\mathbf{X}_i \in \mathcal{X}$;

Notations

- Treatment $W_i \in \{0, 1\}$
- Outcome $Y_i \in \mathcal{Y}$
- Covariate $\mathbf{X}_i \in \mathcal{X}$; individual $\mathbf{X}_i^{ind} \in \mathcal{X}^{ind}; \text{neighborhood}$
 $\mathbf{X}_i^{neigh} \in \mathcal{X}^{neigh}$
- Undirected network $G = (\mathcal{N}, \mathcal{E})$
- For node i consider partitions $(i, \mathcal{N}_i, \mathcal{N}_{-i})$:
 $(W_i, \mathbf{W}_{\mathcal{N}_i}, \mathbf{W}_{\mathcal{N}_{-i}}), (Y_i, \mathbf{Y}_{\mathcal{N}_i}, \mathbf{Y}_{\mathcal{N}_{-i}})$

SUTNVA and potential outcomes

Assumption 1 No Multiple Versions of Treatment(Consistency)

$$Y_i = Y_i(\mathbf{W})$$

Assumption 2 Neighborhood Interference

Given a function $g_i : \{0, 1\}^{N_i} \rightarrow \mathcal{G}_i$, $\forall i \in \mathcal{N}$, $\forall \mathbf{W}_{\mathcal{N}_{-i}}, \mathbf{W}'_{\mathcal{N}_{-i}}$ and $\forall \mathbf{W}_{\mathcal{N}_i}, \mathbf{W}'_{\mathcal{N}_i}, g_i(\mathbf{W}_{\mathcal{N}_i}) = g_i(\mathbf{W}'_{\mathcal{N}_i})$, the following equality holds:

$$Y_i(W_i, \mathbf{W}_{\mathcal{N}_i}, \mathbf{W}_{\mathcal{N}_{-i}}) = Y_i(W_i, \mathbf{W}'_{\mathcal{N}_i}, \mathbf{W}'_{\mathcal{N}_{-i}})$$

Potential outcomes

$$Y_i(w, g) = Y_i(W_i = w, G_i = g)$$

for a subset of nodes $V_g = \{i : g \in \mathcal{G}_i\}$ where $G_i = g_i(\mathbf{W}_{\mathcal{N}_i})$

Assignment mechanism

Assignment mechanism

$$P(\mathbf{W}, \mathbf{G} | \mathbf{X}, \{\mathbf{Y}(w, g), w = 0, 1; g \in \mathcal{G}\}) = \begin{cases} P(\mathbf{W} | \mathbf{X}, \{\mathbf{Y}(w, g), w = 0, 1; g \in \mathcal{G}\}) & \text{if } \mathbf{G} = g(\mathbf{W}) \\ 0 & \text{otherwise} \end{cases}$$

Assumption 3 Unconfoundness of Individual and Neighborhood Treatment

$$Y_i(w, g) \perp\!\!\!\perp W_i, G_i | \mathbf{X}_i \quad \forall w \in \{0, 1\}, g \in \mathcal{G}_i, \forall i$$

Causal estimands

Main effect $\tau(g)$ and overall main effect τ

$$\tau(g) = E[Y_i(W_i = 1, G_i = g) - Y_i(W_i = 0, G_i = g) | i \in V_g]$$

$$\tau = \sum_{g \in \mathcal{G}} \tau(g) P(G_i = g)$$

Spillover effect $\delta(g; w)$ and overall spillover effect $\Delta(w)$

$$\delta(g; w) = E[Y_i(W_i = w, G_i = g) - Y_i(W_i = w, G_i = 0) | i \in V_g]$$

$$\Delta(w) = \sum_{g \in \mathcal{G}} \delta(g; w) P(G_i = g)$$

Total effect TE

$$\sum_{g \in \mathcal{G}} E[Y_i(W_i = 1, G_i = g) - Y_i(W_i = 0, G_i = 0) | i \in V_g] P(G_i = g)$$

$$= \tau + \Delta(0)$$

Generalized propensity score

$$\begin{aligned}\psi(w; g; x) &= P(W_i = w, G_i = g | \mathbf{X}_i = \mathbf{x}) \\ &= P(G_i = g | W_i = w, \mathbf{X}_i^g = \mathbf{x}^g) P(W_i = w | \mathbf{X}_i^w = \mathbf{x}^w) \\ &= \lambda(g; w; x^g) \phi(w; x^w)\end{aligned}$$

- Joint propensity score: $\psi(w; g; x)$
- Neighborhood propensity score: $\lambda(g; w; x^g)$
- Individual propensity score: $\phi(w; x^w)$

Assumption 1-3

Assumption 1 No Multiple Versions of Treatment(Consistency)

$$Y_i = Y_i(\mathbf{W})$$

Assumption 2 Neighborhood Interference

Given a function $g_i : \{0, 1\}^{N_i} \rightarrow \mathcal{G}_i$, $\forall i \in \mathcal{N}$, $\forall \mathbf{W}_{\mathcal{N}_{-i}}, \mathbf{W}'_{\mathcal{N}_{-i}}$ and $\forall \mathbf{W}_{\mathcal{N}_i}, \mathbf{W}'_{\mathcal{N}_i}, g_i(\mathbf{W}_{\mathcal{N}_i}) = g_i(\mathbf{W}'_{\mathcal{N}_i})$, the following equality holds:

$$Y_i(W_i, \mathbf{W}_{\mathcal{N}_i}, \mathbf{W}_{\mathcal{N}_{-i}}) = Y_i(W_i, \mathbf{W}'_{\mathcal{N}_i}, \mathbf{W}'_{\mathcal{N}_{-i}})$$

Assumption 3 Unconfoundedness of Individual and Neighborhood Treatment

$$Y_i(w, g) \perp\!\!\!\perp W_i, G_i | \mathbf{X}_i \quad \forall w \in \{0, 1\}, g \in \mathcal{G}_i, \forall i$$

Properties of GPS

Properties

- **Balancing:**

$$P(W_i, G_i | \mathbf{X}_i, \psi(w; g; \mathbf{X}_i)) = P(W_i, G_i | \psi(w; g; \mathbf{X}_i))$$

- **Conditional Unconfoundedness(joint)**

$$Y_i(w, g) \perp\!\!\!\perp W_i, G_i | \psi(w; g; \mathbf{X}_i) \quad \forall w \in \{0, 1\}, g \in \mathcal{G}_i$$

- **Conditional Unconfoundedness**

$$Y_i(w, g) \perp\!\!\!\perp W_i, G_i | \lambda(g; w; \mathbf{X}_i^g), \phi(1; \mathbf{X}_i^w) \quad \forall w \in \{0, 1\}, g \in \mathcal{G}_i$$

GPS estimator

$$E[E[Y_i | W_i = w, G_i = g, \psi(w; g; \mathbf{X}_i)] | W_i = w, G_i = g]$$

$$E[E[Y_i | W_i = w, G_i = g, \lambda(g; w; \mathbf{X}_i^g), \phi(1; \mathbf{X}_i^w)] | W_i = w, G_i = g]$$

Estimating procedure

Subclassification and GPS

- ① Subclassification on $\phi(1; \mathbf{X}_i^w)$
 - ① Estimate $\phi(1; \mathbf{X}_i^w)$: logistic regression $W_i \sim \mathbf{X}_i^w$
 - ② Predict $\phi(1; \mathbf{X}_i^w)$
 - ③ Subclassify J subclasses B_j by $\phi(1; \mathbf{X}_i^w)$ where $\mathbf{X}_i^w \perp\!\!\!\perp W_i | i \in B_j$
- ② Within B_j , estimate $\mu_j(w, g) = E[Y_i(w, g) | i \in B_j^g]$ where $B_j^g = V_g \cap B_j$
 - ① Estimate $\lambda(g; w; \mathbf{X}_i^g)$: $G_i \sim W_i + \mathbf{X}_i^g$
 - ② Estimate outcome model $Y_i \sim \lambda(g; w; \mathbf{X}_i^g) + W_i + G_i$
 - ③ Predict $Y_i(w, g)$
 - ④ Estimate $\hat{\mu}_j(w, g) = \sum_{i \in B_j^g} \hat{Y}_i(w, g) / |B_j^g|$
- ③ Estimate $\hat{\mu}(w, g) = \sum_{j=1}^J \hat{\mu}_j(w, g) \pi_j^g$ where $\pi_j^g = |B_j^g| / v_g$

Notice: $\mu(w, g) = E[Y_i(w, g) | i \in V_g], \forall w \in \{0, 1\}, g \in \mathcal{G}$

Simulation

A friendship network from Facebook is obtained from Stanford Network Analysis Project(SNAP). It consists of 10 egocentric networks of over 4000 nodes and 80000 edges.

① Individual covariates $X_i^{ind} = (Age_i, Gender_i)$

② Neighborhood covariates

$$X_i^{neighbor} = \left(\frac{\sum_{j \in V_i} Age_j}{N_i}, \frac{\sum_{j \in V_i} Gender_j}{N_i}, N_i \right)$$

Simulate 2 scenarios on this real network

① $Z \perp\!\!\!\perp G | X^{ind}$

② $Z \perp\!\!\!\perp G | X^{ind}, X^{neighbor}$

Data Generation Mechanism

Assignment Generation Mechanism

- ① $\text{logit}(P(Z_i = 1)) = -1.5 + 1.2\text{gender}_i + 0.3\text{age}_i$
- ② $\text{logit}(P(Z_i = 1)) = -4.5 + 0.4\text{gender}_i + 0.1\text{age}_i + 2.5\text{friends.gender}_i + 0.3\text{friends.age}_i$

Outcome Generation Mechanism

$$Y_i(z, g) \sim N(\mu(z, g, X_i^{\text{ind}}, X_i^{\text{neighbor}}), 1)$$

$$\mu(z, g, X_i^{\text{ind}}, X_i^{\text{neighbor}}) = 15 - 7I(\phi(1, X^{\text{ind}}) \geq 0.85) - 15z$$

$$+ 3zI(\phi(1, X^{\text{ind}}) \geq 0.85) + \delta g$$

Results

Comparison Procedure

- ① A simple difference in mean outcome of treated and untreated.
- ② A regression estimator on individual treatment Z_i & covariates
- ③ Subclassification on individual propensity score & covariates

Scenarios	δ	1	2	3	GPS
$Z \perp\!\!\!\perp G X^{ind}$	low	-2.3	-0.58	-0.37	-0.02
	med	-2.46	-0.60	-0.41	-0.05
	high	-2.52	-0.53	-0.38	-0.01
$Z \perp\!\!\!\perp G X^{ind}, X^{neighbor}$	low	-2.47	-1.80	-1.90	-0.7
	med	-2.78	-2.06	-2.18	-0.69
	high	-2.99	-2.23	-2.37	-0.69

Table 1: Bias of Estimators