Estimation of Causal Peer Influence effects

Toulis, Kao

Proceedings of the 30th International Conference on Machine Learning (2013)

Link

Introduction

- Interference among units:
  - When a treatment on a unit has an effect on the response of another.
- Also called spillover effects in economics.
- In the context of humans and their peers, called *peer influence effects*.
- How to do causal inference in the presence of interference?

Related work

- Most works (at the time) treat interference as a nuisance.
  - Especially, in the estimation of average treatment effect.
- In this work, they estimate the peer influence effect itself.
- Hudgens & Halloran (2008): estimating causal effect of vaccination under interference by comparing *reference groups*.
- Peer effects have been researched in the absence of a causal model (Shah & Zaman, 2011).

Why estimate peer influence effects?

- Take the case study in Bakshy et al., 2012:
  - Facebook users were shown ads with and without their friends' product affiliation.
  - How much causal effect does peer endorsement have?
- Risks of ignoring peer influence effects (Sobel, 2006):
  - Social program for financing poor households to move to better neighborhoods.
  - Ignoring interference leads to erroneous conclusions about effectiveness of the program.

The potential outcomes model

Unit	"Ideal" world		Real world	
	$Z_i = 0$	$Z_i = 1$	$Z_i = 0$	$Z_i = 1$
1	$Y_1(0)$	$Y_1(1)$	$Y_1(0)$	?
2	$Y_2(0)$	$Y_2(1)$	?	$Y_2(1)$

Table 1. Causal inference as a missing data problem under the Rubin model

If we observed all possible outcomes,

$$\mu = (1/2) \cdot [Y_1(1) + Y_2(1) - Y_1(0) - Y_2(0)]$$

In the real world, we only observe one outcome for a unit.
We may randomize the treatment assignment and estimate the causal effect by $Y_2(1) - Y_1(0)$.

The SUTVA assumption

- Stable Unit Treatment Value assumption:
  - The outcome of individual i is a function of only its treatment Z_i. $Y_i(\boldsymbol{Z}) = Y_i(Z_i)$.
- However, in the presence of interference, this assumption is violated.
- This work relaxes the assumption to:
  - The outcome of i is a function of its own treatment Z_i as well as those of its neighbors.

$$Y_i(\boldsymbol{Z}) = Y_i(Z_i, \boldsymbol{Z}_{N_i})$$

A few definitions

- If $Z_i = 1$, then i is said to have *primary* effects.
- If i has at least one neighbor being treated, then i is said to have *peer influence* effects.
- A unit is *k-exposed* if exactly k of its neighbors are being treated. $\boldsymbol{Z}$ in this case is a *k-level assignment*.
- Let D_i be the set of all assignments $\boldsymbol{Z}_{N_i}$ such that i is k-exposed.
- A unit is *non-exposed* if $Z_i = 0$ and $\boldsymbol{Z}_{N_i} = 0$.
- Let V_k be the set of all nodes with at least k neighbors.
- $\boldsymbol{Z}(N_i; k)$ be the set of assignments on N_i where exactly k neighbors of i get treated. Size: $C_k^{n_i}$

Causal estimands

The primary treatment effect is defined

$$\xi \equiv \frac{1}{N}\sum_i Y_i(1, \mathbf{Z}_{\mathcal{N}_i} = \mathbf{0}) - Y_i(\mathbf{0})$$

The k-level peer influence effect is defined as

$$\delta_k \equiv \frac{1}{|V_k|}\sum_{i \in V_k}\left[\binom{n_i}{k}^{-1}\sum_{\mathbf{z}\in\mathbf{Z}(\mathcal{N}_i;k)} Y_i(0,\mathbf{z}) - Y_i(\mathbf{0})\right]$$

How to estimate these?

- The primary effect ξ is estimated in the usual way: a randomized experiment.
- However, estimating δ_k is more involved.
  - Any randomization of the treatment vector must $Z_i = 0$ for all $i \in V_k$.
  - The randomization must happen only within their neighborhoods.

A simple sequential design

Algorithm 1 Estimation of δ_k: Simple Sequential Randomization $\texttt{SSR}(G, \mathbf{Z})$

Input: G network, $\mathbf{Z}$ current treatment vector

Output: $\mathbf{Z}$ treatment vector (in-place)

while $i \leftarrow \texttt{sample}\{i : i \in V_k \;\&\; \mathbf{s}(\mathbf{Z}_{\mathcal{N}_i}) \leq k\}$ **do**

  $T_i = \{j \in \mathcal{N}_i : Z_j \neq \text{NA}\}$

  $\mathbf{W} \leftarrow \texttt{sample}\{\mathbf{W} : \mathbf{W} \in \mathcal{D}_i \;\&\; \mathbf{W}_{T_i} = \mathbf{Z}_{T_i}\}$

  $\mathbf{Z}_{\mathcal{N}_i} \leftarrow \texttt{sample}\{\mathbf{W}, \mathbf{0}\}$

  $Z_i \leftarrow 0$

  $V_k \leftarrow V_k \setminus \{i\}$

end while

A few takeaways

- The paper has a lot of more results.
  - A model-based approach where the response function Y_i is a linear function of units and their neighbors.
  - How to model network uncertainty?
  - Performance analysis and comparisons.