On Covariate Adjustment in Missing Not at Random Models

Jacob M. Chen†, Rohit Bhattacharya†
{*jmc8@*, *rb17@*}williams.edu
† Dept. of Computer Science, Williams College

Motivation

- Missing data is ubiquitous across the sciences.
- Recent work [4, 1] introduces m-adjustment and fixing as ways to identify causal effects under missing data.
- We extend these techniques and show how they can be applied to a broad class of missing data models.
- We show that existing theory on optimal adjustment sets can be applied to certain classes of missing data models.

Directed Acyclic Graphs and M-Graphs

- Missing data directed acyclic graphs (m-DAGs) encode substantive assumptions on the target distribution & missingness mechanism.
- For each partially observed variable Z_i, there is a missingness indicator R_i and observed proxy Z_i^* such that

$$Z_i^* = Z_i \text{ if } R_{Z_i} = 1 \text{ and } Z_i^* = ? \text{ if } R_{Z_i} = 0.$$

- The full data distribution (target distribution + missingness mechanism) factorizes according to the m-DAG $\mathcal{G}$ as,

$$p(V) = \prod_{V_i \in V} p(V_i \mid \text{pa}_{\mathcal{G}}(V_i)).$$

SWIG Phrasing of The M-Adjustment Criterion [4, 2]

- Z is a valid m-adjustment set wrt A and Y if and only if
  - Z does not contain any potential outcomes in the SWIG.
  - $Y(a) \perp\!\!\!\perp A \mid Z, R_W$ holds in the SWIG.
  - $Y(a) \perp\!\!\!\perp R_W$ holds in the SWIG.

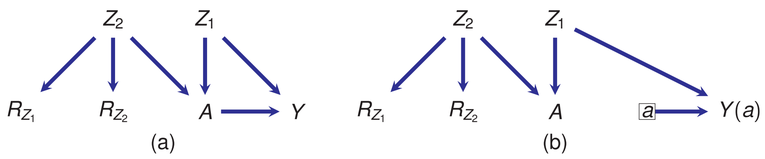

Figure: (a) is the original graph and (b) is the SWIG.

Fixability And Fixing In M-graphs

- Let M be the set of partially observed parents of R_Z, R_M be the set of missingness indicators for M, and S be the set of variables that have been selected on.
- R_Z is *fixable* if

$$R_Z \perp\!\!\!\perp \{R_M, S\} \setminus \text{pa}_{\mathcal{G}}(R_Z) \mid \text{pa}_{\mathcal{G}}(R_Z).$$

- **Probabilistic Operation:**

$$p(V \setminus R_Z \mid\mid R_Z = 1) = p(V)/p(R_Z \mid \text{pa}_{\mathcal{G}}(R_Z)) \mid_{R_Z=1}$$

- **Graphical Operation:** Delete incoming edges to R_Z and make Z a fully observed variable because R_Z is fixed to value 1.

Figure: Fixing in an m-graph: (a) is pre-fixing and (b) is post-fixing

Using M-Adjustment and Fixing

- (a) is an example where m-adjustment is sufficient, but in (b) and (c) different fixing techniques along with m-adjustment are required to identify the causal effect.

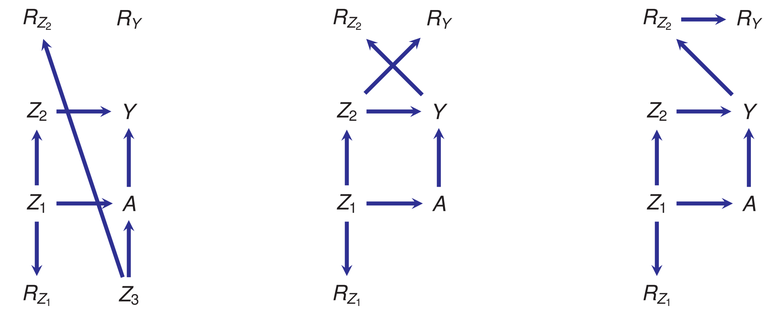

Figure: (a), (b), (c) are all MNAR models. The causal effect in (b) is obtained by fixing R_{Z_2} and R_Y in parallel while the causal effect in (c) is obtained by fixing R_Y then R_{Z_2} in sequence.

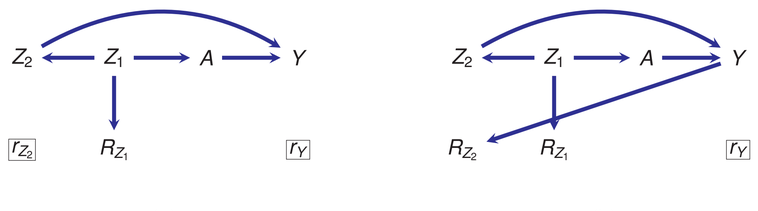

Figure: (a) is the graph we arrive at after fixing R_{Z_2} and R_Y in parallel. (b) is the graph we arrive at after fixing R_Y. We then fix R_{Z_2} to recover the causal effect.

Algorithms for Recovering Causal Effects under Missing Data

- We propose a simple algorithm for m-adjustment after fixing in parallel.
  - For each indicator R_Z, we check if R_Z is fixable in $\mathcal{G}$.
  - We then search for a valid m-adjustment set in the graph obtained after fixing all R_Z that satisfied fixability.
- We also propose an algorithm that greedily chooses missingness indicators in the m-DAG $\mathcal{G}$ sorted in reverse topological order and tries to fix them in sequence until there is a valid m-adjustment set.

Algorithm 2 Using sequential fixing to recover causal effect.

$R_A \leftarrow$ the set of all missingness indicators in the m-graph G
$R_S \leftarrow$ an empty set
while there is no valid m-adjustment set Z in G and $R_A \neq \emptyset$ **do**
  $fixedSomethingInSequence \leftarrow$ False
  for each R_a in R_A **do**
    $R_a \leftarrow$ the first missingness indicator in R_A sorted in reverse topological order
    $M \leftarrow$ the parents of R_a
    $R_M \leftarrow$ the missingness indicators for each partially observed variable in M
    $R_{M_2} \leftarrow \{R_M \cap M\}$
    $R_M \leftarrow R_M \setminus \{R_M \cap M\}$
    if ($R_M = \emptyset$ or $R_a \perp\!\!\!\perp R_M \mid M$) and ($R_a \perp\!\!\!\perp R_S \mid M$) **then**
      fix R_a in G
      $R_A \leftarrow R_A \setminus R_a$
      $fixedSomethingInSequence \leftarrow$ True
      $R_s \leftarrow$ the
      $R_S \leftarrow R_S \cup R_M$
      $R_S \leftarrow R_S \cup R_{M_2}$
      $R_A \leftarrow R_A \setminus R_S$
      break
  if $fixedSomethingInSequence$ is False **then**
    break ▷ nothing was fixable, so we break out of the while loop

Figure: Algorithm for Fixing in Sequence

Optimal Adjustment Sets

- Under fully observed data, the efficient adjustment set is the set of variables O such that $O_i \to M_i$ exists for any M_i that lies on a causal path between A and Y [3].

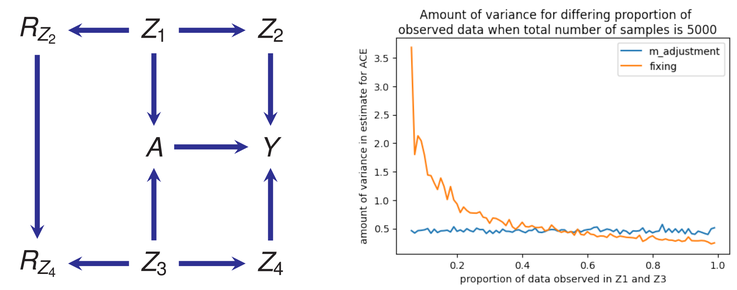

Figure: With no missingness, $\{Z_2, Z_4\}$ is the optimal adjustment set. With missingness, our experiments show that $\{Z_1, Z_3\}$ produce lower variance estimates if either (i) the amount of missingness in the optimal set is significant or (ii) the models required to fix R_{Z_2} and R_{Z_4} are complicated.

Optimal Adjustment Sets Under Missing Data

- When you have to fix every missingness indicator to recover the causal effect, the optimal adjustment set under missing data is the same as the optimal adjustment set under fully observed data.
- We show that this is true for the block-parallel model and the block-sequential model.

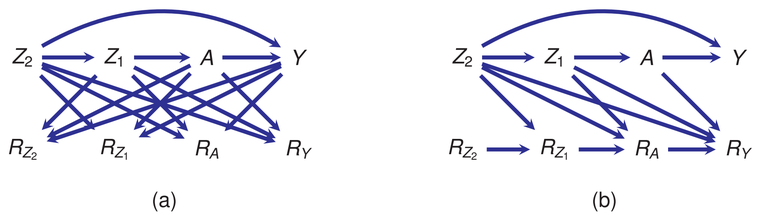

Figure: (a) is a block-parallel model and (b) is a block-sequential model. These classes of models can be generalized to any number of covariates.

Future Work

- How to formulate a complete algorithm for identifying causal effects in MNAR m-graphs via m-adjustment and fixing?
- Are there more classes of missingness models where we know the optimal adjustment set under missingness?
- Is there a general criterion for finding optimal adjustment sets for m-graphs?

References

[1] Rohit Bhattacharya, Razieh Nabi, Ilya Shpitser, and James M Robins. Identification in missing data models represented by directed acyclic graphs. In *Uncertainty in Artificial Intelligence*, pages 1149–1158. PMLR, 2020.

[2] Thomas S Richardson and James M Robins. Single world intervention graphs: a primer. In *Second UAI workshop on causal structure learning, Bellevue, Washington*. Citeseer, 2013.

[3] Andrea Rotnitzky and Ezequiel Smucler. Efficient adjustment sets for population average causal treatment effect estimation in graphical models. *Journal of Machine Learning Research*, 21:188–1, 2020.

[4] Mojdeh Saadati and Jin Tian. Adjustment criteria for recovering causal effects from missing data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 561–577. Springer, 2019.

[5] Yan Zhou, Roderick JA Little, and John D Kalbfleisch. Block-conditional missing at random models for missing data. *Statistical Science*, 25(4):517–532, 2010.