'Unit-Tests' for Good and Bad Controls: Task 1

Saksham Jain

13 Nov, Autumn 2022

Contents

1 Outline

This report is organized as follows: in Section 2 we highlight the main objectives for this first task. In Section 3 we highlight the main ideas and results of the background reading for this task. Finally, in Section 4, we propose a unit test for M-bias and conduct an experimental analysis.

2 Goals

- Develop background on formulating low-dimensional problems as 'unit-tests'

- Develop background on good and bad controls
- Propose a 'unit test' based on a selected example from Cinelli et al. [2021]
- Show experimental results for this unit test

3 Background

3.1 Linear Unit-Tests for Invariance Discovery

Main Idea: Formulate low-dimensional linear problems [Aubin et al., 2021] for algorithms that supposedly learn only invariant correlations.

For all problems, $X^e = [X^e_{\text{inv,i}}, X^e_{\text{spu,i}}]$ and the dimension $d = d_{\text{inv}} + d_{\text{spu}}$.

As the relevant inspiration for our proposed unit test (described later in Section 4), we reiterate only the linear least squares problem from [Aubin et al., 2021] in its simple, default setting below:-

3.1.1 Linear least squares

$\forall\ e \in \varepsilon$ and $i = 1, ..., n_e$,

$$Y^e_i = \frac{2}{d} 1^\top_{d_{\text{inv}}} \tilde{Y^e_i}\ , \qquad \begin{aligned} X^e_{\text{inv,i}} &\sim \mathcal{N}_{d_{\text{inv}}}(0, (\sigma^e)^2) \\ \tilde{Y^e_i} &\sim \mathcal{N}_{d_{\text{inv}}}(W_{yx} X^e_{\text{inv,i}}, (\sigma^e)^2)) \\ X^e_{\text{spu,i}} &\sim \mathcal{N}_{d_{\text{spu}}}(W_{xy} \tilde{Y^e_i}, 1) \end{aligned}$$

where $W_{yx} \in \mathbb{R}^{d_{\text{inv}} \times d_{\text{inv}}}, W_{xy} \in \mathbb{R}^{d_{\text{spu}} \times d_{\text{inv}}}$ are i.i.d. Gaussian. Therefore, the features contain both causes ($X^e_{\text{inv,i}}$) and effects ($X^e_{\text{spu,i}}$) of the outcome variable, and only the causes remain invariant across the environments e.

3.2 A Crash Course on Good and Bad Controls

Main Idea: Block all spurious paths between X and Y, and do not perturb the causal paths.

3.2.1 Terminology

1. *Mediator*: A variable that lies in the causal path between X and Y., e.g. X→Z→Y. Controlling for Z blocks the flow of association.

2. *Common Cause*: A variable that affects both X and Y, e.g. X←Z→Y. Controlling for Z blocks the flow of association.

3. *Collider*: A variable that is a common effect of both of X and Y. e.g. X→Z←Y. Controlling for Z opens the flow of association.

4. *Back-door path*: For a given variable, it is a confounding path that begins with an arrow pointing into it.

5. *d-separation*: Z d-separates X and Y, if controlling for it blocks all paths form X to Y. It implies X⊥Y | Z

6. *Average Causal Effect (ACE)*: It is the expected increase in Y with a unit increase in X due to some intervention.

3.2.2 Examples

1. **Good Controls**: Conditioning on these blocks the back-door paths between X and Y, and thus allows us to get an unbiased estimate for the ACE.
   The classic case is when Z is a common cause of X and Y, in which case Z clearly d-separates X and Y.
   This holds true when Z is a mediator between an unobserved common cause (of X and Y) and either X or Y.
   This insight can further be extended to the case where Z is a common cause of X and a mediator M between X and Y.
   And thus also holds true when Z is a mediator between an unobserved common cause (of X and M) and either X or M.

2. **Bad Controls**: Conditioning on these biases the ACE estimate by inducing non-causal associations, e.g. by opening colliding paths.
   M-bias: When Z is correlated with both X and Y and is treated as a "pre-treatment" variable. Note that it becomes undecidable (without further assumptions) if Z has a direct effect on Y.
   Bias amplification: When X and Y have an unobserved confounder U, but Z is a "pre-treatment" variable and only a cause of X.
   Overcontrol bias: When Z is a mediator between X and Y, or is the descendant of a mediator between X and Y.

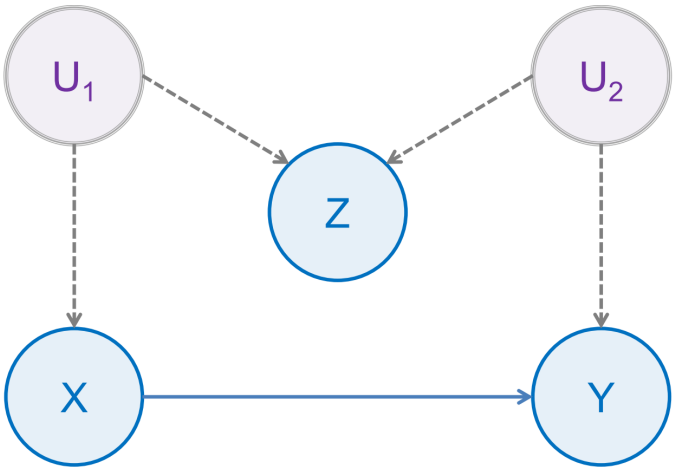

Figure 1: A causal DAG depicting M-Bias. Note that U_1 and U_2 are unobserved.

Selection bias: When Z is a collider of X and Y, or is a collider of X and an unobserved parent of Y.
Case-control bias: When Z is a collider of Y only. This is a special case of selection bias. Note that adjusting for Z is valid for testing if the effect of X on Y is zero.

3. **Neutral Controls (possibly helpful)**: Conditioning on these does not block or open any back-door paths between X and Y, but may improve precision of the ACE estimate or help in the case of selection bias.
   One case is when Z is a direct parent of Y only.
   Another case is when Z is a direct cause of a mediator M between X and Y.
   Also, adjusting for Z may be helpful if in the case of selection bias, Z is actually a mediator between X and the collider of X and an unobserved parent of Y.

4. **Neutral Controls (not helpful)**: Conditioning on these does not block or open any back-door paths between X and Y, but may be bad for precision of the estimate.
   When Z is a direct parent of X, controlling for it will reduce the variation of X.

4 Proposed Unit Test

We propose a unit test for M-bias. For the purpose of this paper, the 'algorithm' our unit test is meant to test refers to any process that outputs a set of variables to include for regression.

We formulate a linear least square regression problem where the features contain a variable that is correlated with both X and Y through unobserved causes of X and Y, which,

if included in the regression, will bias the ACE estimate. The simplest setting (shown in Figure 1 [1]) for constructing the dataset is as follows:-

$$\begin{aligned} Y &= \beta_0 + \beta_1 * X + \beta_2 U_2 + \epsilon \qquad & \beta_1 &\sim \mathcal{N}(\mu_{\beta_1}, s^2_{\beta_1}) \\ X &= \beta'_0 + \beta'_1 U_1 \qquad , & U_1 &\sim \mathcal{N}(\mu_1, s^2_1) \\ Z &= \beta''_1 U_1 + \beta''_2 U_2 & U_2 &\sim \mathcal{N}(\mu_2, s^2_2) \end{aligned}$$

where X is the treatment, Z is the bad control, and U_1, U_2 are unobserved. Note that while the coefficients $\beta_0, \beta'_0, \beta'_1, \beta''_1, \beta''_2$ are scalar, β_1 has mean μ_{β_1} (which is the ACE) and therefore, the quantity we want to recover using regression. The noise term ϵ accounts for the uncertainty in β_1.

Thus, if Z is included in the regression, it will bias the ACE estimate. Please see the companion Jupyter Notebook for an experimental example of this unit test.

Of course, in the real world, we usually have more variables than just X and Z. So we can extend the above setting to a slightly more general case, where we have more variables we can control for. However, since we want to test only for the presence of a variable that induces M-bias, we will add only neutral controls (that directly affect Y) to the linear model (see section 3.2 Examples). We construct the problem as follows:-

$$\begin{aligned} Y &= \beta_0 + \beta_1 * X + \beta_2 U_2 + \mathbf{W}\theta + \epsilon + \delta \qquad & \beta_1 &\sim \mathcal{N}(\mu_{\beta_1}, s^2_{\beta_1}),\ \delta \sim \mathcal{N}(0, \sigma^2) \\ X &= \beta'_0 + \beta'_1 U_1 \qquad , & U_1 &\sim \mathcal{N}(\mu_1, s^2_1) \\ Z &= \beta''_1 U_1 + \beta''_2 U_2 & U_2 &\sim \mathcal{N}(\mu_2, s^2_2) \end{aligned}$$

where $\mathbf{W}$ contains variables $W_1, ..., W_d$ drawn i.i.d. from $\mathcal{N}(0, 1)$, and $\theta \in \mathbb{R}^d$ are their coefficients which we can set to, for example, $\theta_j = \{10/j\sqrt{n}\}$, n being the sample size. The additional noise term δ allows us to choose σ^2 such that we can set R^2 for $W_1, ..., W_d$ and Y to a pre-specified value. Here we set $R^2 = 0.8$.

Now, if we control for variables from $\mathbf{W}$ that are selected by the algorithm, it may help with the precision of the ACE. However, if Z is selected, the ACE estimate will be biased. Please see the companion Jupyter Notebook for an experimental example of this generalized unit test.

[1] Figure from Cárdenas Hurtado [2016]

References

Carlos Cinelli, Andrew Forney, and Judea Pearl. A crash course in good and bad controls. *Sociological Methods & Research*, page 00491241221099552, 2021.

Benjamin Aubin, Agnieszka Słowik, Martin Arjovsky, Leon Bottou, and David Lopez-Paz. Linear unit-tests for invariance discovery. *arXiv preprint arXiv:2102.10867*, 2021.

Camilo Alberto Cárdenas Hurtado. Causal inference in the presence of causally connected units: a semi-parametric hierarchical structural equation model approach. *Departamento de Estadística*, 2016.