Path Analysis Handout

Daniel Arnon and Adam Glynn

August 5, 2019

Note: this handout is somewhat informal and is meant to efficiently convey the key points of path analysis without taking the time for rigor.

1 Potential outcomes and structural linear models

The potential outcomes framework can be translated into the structural linear models framework in the following manner. Formally, potential outcomes are defined as:

$$\text{Potential outcome} = \begin{cases} Y_i(1), & if D_i = 1 \\ Y_i(0), & if D_i = 0 \end{cases}$$

Using the potential outcomes framework we can talk about both quantities: $Y_i(1)$ which is the potential outcome under treatment and $Y_i(0)$ which is the potential outcome under control. But since we only observe one of these two outcomes, the observed outcome Y_i can be rewritten as:

$$\begin{aligned} Y_i &= Y_i(1)D_i + Y_i(0)(1 - D_i) \\ &= Y_i(0) + (Y_i(1) - Y_i(0))D_i \\ &= \beta_{0i} + \beta_{1i}D_i \end{aligned}$$

where $\beta_{0i} = Y_i(0)$ is the unit specific intercept and $\beta_{1i} = (Y_i(1) - Y_i(0))$ is the unit specific slope. In order to convert to the familiar linear model we start by adding and subtracting $E[\beta_{0i}]$ from the equation to get:

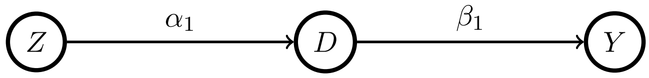
$$\begin{aligned} Y_i &= E[\beta_{0i}] + \beta_{1i}D_i + \beta_{0i} - E[\beta_{0i}] \\ &= E[\beta_{0i}] + \beta_{1i}D_i + \epsilon_i \\ &= \beta_0 + \beta_1 D_i + \epsilon_i \end{aligned}$$

Where to go from line 1 to line 2, we simply define $\epsilon_i = \beta_{0i} - E[\beta_{0i}]$. Note that at this stage, we have done nothing to the potential outcomes model other than re-parameterize the potential outcomes under treatment and control in terms of the average control outcome, deviations from this average, and treatment effects. The move from line 2 to line 3 however relies on the additional strong assumption that $\beta_{1i} = \beta_i \forall i$ (i.e., constant effects holds). The following section on path analysis relies on results analogous to this.

2 Path Analysis

2.1 DAG with Single Path

Consider the following DAG (Directed Acyclic Graph), where by convention, we do not include error terms on the graph unless they are correlated, point into more than one variable, or are pointed into themselves.



From the previous section, we know that we can transform potential outcomes each of these causal effects into linear models. The effect of $Z \rightarrow D$ can be parameterized as:

$$(1) D_i(z) = \alpha_0 + \alpha_1 z + \nu_i$$

The effect of $D \rightarrow Y$ can be parameterized as:

$$(2) Y_i(d) = \beta_0 + \beta_1 d + \epsilon_i$$

Note that if Z and/or D are not binary, then we have added a linearity assumption. With this model, we can combine these two equations by plugging (1) into (2) :

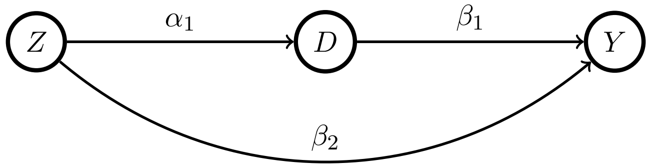
$$Y_i(D_i(z)) = Y_i(z) = \beta_0 + \beta_1(\alpha_0 + \alpha_1 z + \nu_i) + \epsilon_i$$

$$Y_i(z) = \underbrace{\beta_0 + \beta_1\alpha_0}_{\text{The intercept}} + \underbrace{(\beta_1\alpha_1)}_{\text{The Effect}} z + \underbrace{\beta_1\nu_i + \epsilon_i}_{\text{Error}}$$

Therefore, in the constant effects linear model, the equation for $Y_i(z)$ can be can be written in terms of the equations for $D_i(z)$ and $Y_i(d)$. **The total effect of Z on Y, is the *product* of the the path coefficients**. It is important to note the lack of i subscripts on these coefficients.

2.2 DAG with *Two* Paths

Consider the following DAG (Directed Acyclic Graph):



Now we can again restate the DAG in terms of the linear equations:
The effect of $Z \rightarrow D$ (which has not changed):

$$(1) D_i(z) = \alpha_0 + \alpha_1 z + \nu_i$$

But Y is now a function of both D and Z, and therefore the total effect on Y_i is:

$$(2) Y_i(d, z) = \beta_0 + \beta_1 d + \beta_2 z + \epsilon_i$$

Again, we can combine these two equations by plugging (1) into (2) :

$$Y_i(d_i(z), z) = Y_i(z) = \beta_0 + \beta_1(\alpha_0 + \alpha_1 Z + \nu_i) + \beta_2 Z + \epsilon_i$$

$$Y_i(z) = \underbrace{\beta_0 + \beta_1\alpha_0}_{Intercept} + \underbrace{(\beta_1\alpha_1 + \beta_2)}_{\text{Effect}} z + \underbrace{\epsilon_i + \beta_1\nu_i}_{\text{Error}}$$

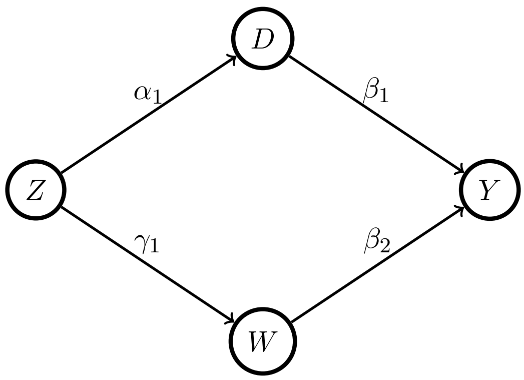
This demonstrates the second property of path analysis:

Theorem (1)**.** *The total effect of Z on Y is the* ***product*** *of the coefficients on each* ***directed*** *path, and the* ***sum*** *of all separate directed pathways.*

To explicate this carefully notice that:
The first pathway of $Z \to Y$ is the product of the coefficients along this path: $\alpha_1\beta_1$
The second pathway of $Z \to Y$ is single coefficient along this path: β_2
So by this proposition, the total effect is the *sum* of these two paths: $\beta_1\alpha_1 + \beta_2$

2.3 DAG with *two* Paths (but a little more complicated)

Just to make sure this is clear, lets look at this final DAG:

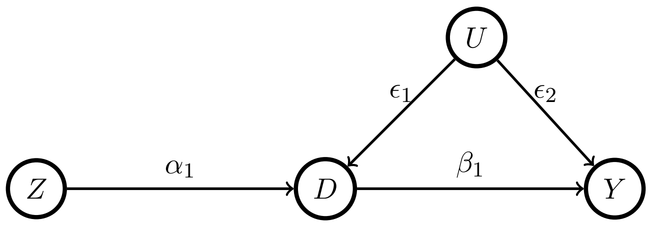


What is the total effect of Z on Y in terms of the path coefficients?

3 Back Door Criterion with Path Analysis

3.1 Using the Back Door Crtierion

Consider the DAG that was introduced in 2.1, but now with an additional variable U:



We know that the average effect of Z on Y can be identified by a regression of Y on Z without conditioning because the empty set satisfies the BDC. Furthermore, with the addition assumptions of linearity and constant effects, the average effect of Z on Y is constant and can be written as $\alpha_1 \cdot \beta_1$. Similarly, the average effect of Z on D can be identified by a regression of D on Z without conditioning, and this average effect can be written as α_1. Taking the ratio of these regression estimates therefore provides an estimate of β_1. This is the path analysis justification for the Wald instrumental variables estimator.

3.2 Further Implications of BDC and Path Analysis

Consider the following TSCS path analysis model. Note that this model is consistent with the LDV model of Equation 6 in Beck and Katz 2011. Using path analysis, we can confirm that the effect of X_1 on Y_3 is $\beta \cdot \phi^2$. Note also that because the regression of Y_3 on X_3 and Y_2 satisfies the BDC with respect to the treatment set $\{X_3, Y_2\}$, we can identify β and ϕ from this regression, and following Beck and Katz 2011, use the $\beta \cdot \phi^2$ formula to identify the effect of X_1 on Y_3. However, the BDC also shows us that we could identify the effect of X_1 on Y_3 by simply regressing Y_3 on X_1.