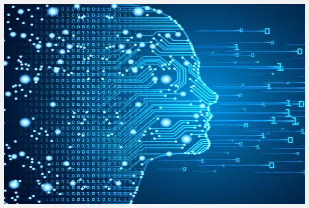


Principles of Causal Inference

Vasant G. Honavar

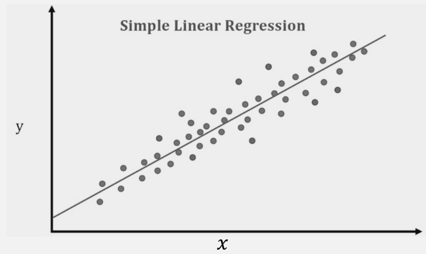
Dorothy Foehr Huck and J. Lloyd Huck Chair in Biomedical Data Sciences and Artificial Intelligence
Professor of Data Sciences, Informatics, Computer Science and Engineering, Bioinformatics & Genomics, Public Health Sciences and Neuroscience
Center for Artificial Intelligence Foundations and Scientific Applications
Institute for Computational and Data Sciences
Huck Institutes of the Life Sciences
Clinical and Translational Sciences Institute
Northeast Big Data Hub
Pennsylvania State University
vhonavar@psu.edu
http://faculty.ist.psu.edu/vhonavar
http://ailab.ist.psu.edu

Linear Structural Causal Models

- Linear Regression
- Introduction to Linear Structural Causal Models
- When regression can and cannot be used to find causal effects.
- Modern algorithmic approaches to identification in linear SCM

Regression

- Predict the value of Y based on X
- Supervised machine learning is often just regression on steroids
- How do we fit a regression line?
  - Given a data set of X, Y pairs, we fit them to
    $$y = mx + b$$
  - The least square regression is the line that minimizes sum squared error
    $$\sum_i (y_i - b - mx_i)^2$$
  - m denotes the slope and b the intercept along the Y axis



Regression Coefficient

- R_{YX} is slope of regression line of Y *on* X
- $m = R_{YX} = \sigma_{XY}/{\sigma_X}^2$
- Slope gives correlation
  - Positive slope → positive correlation
  - Negative slope → negative correlation
  - Zero slope → X and Y are independent or non-linearly correlated

Variance of X, i.e., ${\sigma_X}^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Covariance $\sigma_{XY} \triangleq \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$

Correlation coefficient $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

Multiple Regression

- $y = r_0 + r_1 \cdot x + r_2 \cdot z$
- How do we visualize?: a plane
- What happens if we fix X at some value?
  - $r_1 \cdot x$ becomes a constant
- r_2 is now the slope of slice along X–axis
- What happens if we fix Z at some value?
  - $r_2 \cdot z$ becomes a constant
- r_1 is now the slope of slice along Z–axis

Partial Regression Coefficient

- Regression coefficient of Y on X? R_{YX}
- Regression coefficient of Y on X when holding Z constant?
  - Partial regression coefficient $R_{YX\,.\,Z}$
  - What are the partial regression coefficients in

  $y = r_0 + r_1 \cdot x + r_2 \cdot z$?

  r_1 and r_2

Interpreting regression coefficients

Example: If $y = 1 + 2x_1 + 3x_2$

- Do not interpret the coefficients unless they are statistically significant.
- It *is NOT* accurate to say "For each change of 1 unit in x_1, y changes 2 units".
- What *is* correct to say is "If x_2 is fixed, then for each change of 1 unit in x_1, y changes 2 units."

Orthogonality principle of linear regression

- Suppose for simplicity, $\mathbb{E}[X]$= $\mathbb{E}[Y] = 0$ (intercept of regression is 0)
- Then the constant that minimizes the mean squared error of $\mathbb{E}[(Y - AX)^2]$ is such that $Y - AX$ is orthogonal to X
- That is, $\mathbb{E}\big[(Y - AX)X \quad \big] = 0$
- The corresponding minimum mean squared error is given by $\mathbb{E}\big[(Y - AX)Y \quad \big]$

Proof: See Papoulis, A. Probability, random variables, and stochastic processes, Chapter 7

Orthogonality principle of linear regression

- In general, if $Y = r_0 + r_1 \cdot X_1 + r_1 \cdot X_2 + \cdots + r_k X_k + \epsilon$
- Then regardless of the distribution of $Y, X_1, X_2 \cdots X_k$, the optimal least square solution is obtained when each of the regressors $X_1, X_2 \cdots X_k$ is uncorrelated with ϵ
  - $Cov(\epsilon, X_i) = 0 \; \forall i = 1, \cdots, k$

Orthogonality principle of linear regression

- Suppose we want to compute the expectation of $X = D_1$ given $Y = D_1 + D_2$
- Writing $X = \alpha + \beta Y + \epsilon$ we have:
  - $\mathbb{E}[X] = \alpha + \beta\mathbb{E}[Y] + 0$
- Multiplying both sides of the equation for X by X and taking expectation, we have
  - $\mathbb{E}[X^2] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[XY] + \mathbb{E}[X\epsilon]$
- Now, orthogonality principle dictates that $\mathbb{E}[X\epsilon] = 0$
- Solving the above two equations for α and β we get:
  - $\alpha = \mathbb{E}[X] - \mathbb{E}[Y]\frac{\sigma_{XY}}{\sigma_Y^2}$
  - $\beta = \frac{\sigma_{XY}}{\sigma_Y^2}$

Orthogonality principle of linear regression

- Suppose we use regression to get the best estimate of Z given $X = x$ and $Y = y$

But now, to obtain three equations for α, β_Y and β_X, we also multiply both sides by Y and X and take expectations. Imposing the orthogonality conditions $E[\epsilon Y] = E[\epsilon X] = 0$ and solving the resulting equations gives

$$\beta_Y = R_{ZY \cdot X} = \frac{\sigma_X^2 \sigma_{ZY} - \sigma_{ZX}\sigma_{XY}}{\sigma_Y^2 \sigma_X^2 - \sigma_{YX}^2}$$

$$\beta_X = R_{ZX \cdot Y} = \frac{\sigma_Y^2 \sigma_{ZX} - \sigma_{ZY}\sigma_{YX}}{\sigma_Y^2 \sigma_X^2 - \sigma_{YX}^2}$$

Linear Structural Causal Models

Linear SCM are defined as a system of linear equations representing ground-truth:

$$Y := \sum_i \lambda_{x_i y} X_i + \mathcal{E}_y$$

1. All correlations between $\mathcal{E}$ are explicitly specified.
2. X_i are the direct causes of Y, and $\lambda_{x_i y}$ is the change in Y per X_i.
3. WLOG assume normalized data ($\mathbb{E}[X] = 0$ and $\mathbb{E}[XX] = 1$) to simplify math
4. Assume $\mathcal{E}_y \sim \mathcal{N}$, meaning that the distribution is fully specified by covariance matrix Σ (σ_{ij}).

Causal Inference In Linear Systems

- What is the effect of salt intake on blood pressure after adjusting for confounders; or the total effect of an after-school study program on test scores;
- What is the direct effect or the unmediated by other variables, of the program on test scores.
- What is the effect of enrollment in an optional work training program on future earnings, when enrollment and earnings are confounded by a common cause (e.g., motivation).
- Continuous variables
  - We need to model with continuous variables. These traditionally been formulated as linear equation models .
  - We will assume linear functions and Normal distributions of errors.

Causal Inference In Linear Systems

Plan

- Efficient representation
- Substitutability of expectations for probabilities
- Linearity of expectations
- Invariance of regression coefficients

- Multivariate Gaussian can be expressed in terms of expectation and covariance on pairs of variables
- Conditional probability can be expressed in terms of conditional expectation

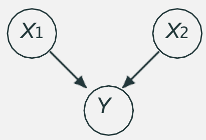
Non-Parametric to Linear

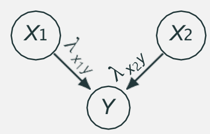
The only substantive change we are making is that the function f becomes linear:

$$V_i \leftarrow f_i(pa_i, U_i) \quad \Rightarrow \quad V_i \leftarrow \sum_{j | V_j \in pa_i} \lambda_{ji} V_j + \mathcal{E}_i$$

1. λ_{ji} is called the "Structural Coefficient".
2. Instead of using U_i, we rename it to $\mathcal{E}_i$ by convention.
3. If we know all λ_{ji}, we can find the causal effect of V_j on V_i.

Example: linear structural causal model





$$X_1 = f_{X_1}(U_{X_1})$$
$$X_2 = f_{X_2}(U_{X_2})$$
$$Y = f_Y(X_1, X_2, U_Y)$$

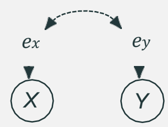
$$X_1 = \varepsilon_{x_1}$$
$$X_2 = \varepsilon_{x_2}$$
$$Y = \lambda_{x_1 y} X_1 + \lambda_{x_2 y} X_2 + \varepsilon_y$$

We can draw the structural coefficients directly on the graph, which then fully specifies the model.

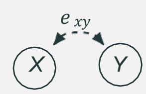
Example: linear structural causal model

The covariance between e_i and e_j is represented by e_{ij}, and is used as the value of a bidirected edge:

Latent Confounding



⇒

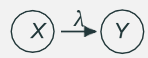


$$e_{xy} \equiv \mathbb{E}[e_x e_y]$$

- e_{xy} is unobserved, since it is covariance of latent variables. It is mathematically useful, however, so we draw it on the graph just like structural coefficients.

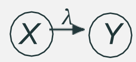
This is different from graph of non-parametric SCM, where a bidirected edge represents an explicit latent variable.

Linear SCM: Interventions



$$\mathbb{E}[Y|do(X = x)] = ?$$

Linear SCM: Interventions

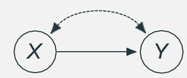


$$\mathbb{E}[Y|do(X = x)] = \mathbb{E}[\lambda x + e_y]$$
$$= \lambda x + \mathbb{E}[e_y]$$
$$= \lambda x$$

Note that *x* is a value of *X*

Linear SCM

- **Graph:** We are assuming that you have a hypothesized causal graph structure. In other words, you think you know what causes what, and which variables have an unknown common cause.
- **Observational Data:** You have a set of datapoints with measurements of all of the observable variables.
- **Goal: Find Structural Coefficients** You do **NOT** have knowledge of the underlying structural coefficients. These represent the actual causal effects that we want to find.

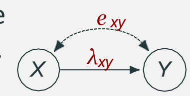


(x_1, y_1)

(x_2, y_2)

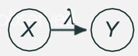
...

(x_n, y_n)



Linear SCM: Interventions

Remember that we assumed $e \sim N$, meaning that the distribution is fully specified by covariance matrix Σ (σ_{xy}).

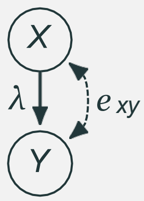


$$
\begin{aligned}
\sigma_{xy} &= \mathbb{E}[XY] \\
&= \mathbb{E}[X(\lambda X + e_y)] \\
&= \mathbb{E}[\lambda XX + X e_y] \\
&= \lambda \mathbb{E}[XX] + \mathbb{E}[X e_y] \\
&= \lambda 1 + 0 \\
&= \lambda
\end{aligned}
$$

Remember, we normailize
The mean to 0 and variance to 1

Connecting Observed with Unobserved

Solve for σ_{xy} in terms of the structural coefficients λ and e_{xy}.

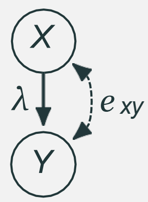


$\sigma_{xy} = ?$

Connecting Observed with Unobserved

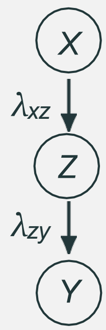
Solve for σ_{xy} in terms of the structural coefficients λ and e_{xy}

$$\sigma_{xy} = \mathbb{E}[XY]$$



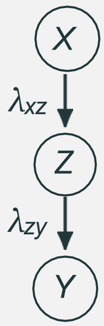
$$\begin{aligned}
\sigma_{xy} &= \mathbb{E}[XY] \\
&= \mathbb{E}[X(\lambda X + e_y)] \\
&= \mathbb{E}[\lambda XX + X e_y] \\
&= \lambda \mathbb{E}[XX] + \mathbb{E}[X e_y] \\
&= \lambda 1 + \mathbb{E}[X e_y] \\
&= \lambda 1 + \mathbb{E}[e_x e_y] \\
&= \lambda + e_{xy}
\end{aligned}$$

A Curious Property



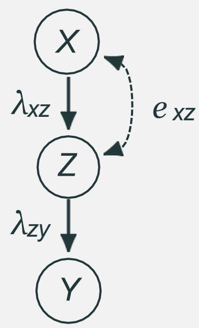
$\sigma_{xy} = ?$

A Curious Property of Linear Causal Models



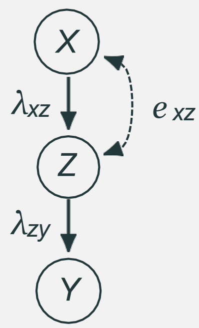
$$
\begin{aligned}
\sigma_{xy} &= \mathbb{E}[XY] \\
&= \mathbb{E}[X(\lambda_{zy} Z + e_y)] \\
&= \mathbb{E}[\lambda_{zy} XZ + X e_y] \\
&= \lambda_{zy}\,\mathbb{E}[XZ] + \mathbb{E}[X e_y] \\
&= \lambda_{zy}\,\mathbb{E}[XZ] \\
&= \lambda_{zy}\,\mathbb{E}[X(\lambda_{xz} X + e_z)] \\
&= \lambda_{zy}\lambda_{xz}\,\mathbb{E}[XX] + \lambda_{zy}\,\mathbb{E}[X e_z] \\
&= \lambda_{zy}\lambda_{xz}
\end{aligned}
$$

A Curious Property of Linear Causal Models



$\sigma_{xy} = ?$

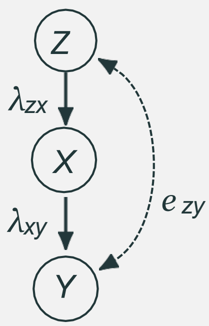
A Curious Property of Linear Causal Models



$$\begin{aligned}\sigma_{xy} &= \mathbb{E}[XY] \\ &= \mathbb{E}[X(\lambda_{zy} Z + e_y)] \\ &= \mathbb{E}[\lambda_{zy} XZ + X e_y] \\ &= \lambda_{zy}\, \mathbb{E}[XZ] + \mathbb{E}[X e_y] \\ &= \lambda_{zy}\, \mathbb{E}[XZ] \\ &= \lambda_{zy}\, \mathbb{E}[X(\lambda_{xz} X + e_z)] \\ &= \lambda_{zy} \lambda_{xz}\, \mathbb{E}[XX] + \lambda_{zy}\, \mathbb{E}[X e_z] \\ &= \lambda_{zy} \lambda_{xz} + \lambda_{zy}\, e_{xz}\end{aligned}$$

Paths and Covariances

There seems to be a relationship between covariances and paths in the graph.

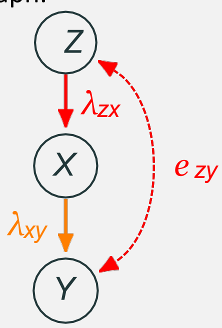


$$\sigma_{xy} = \mathbb{E}[XY] = \mathbb{E}[X(\lambda_{xy}X + e_y)]$$
$$= \lambda_{xy} \ \mathbb{E}[XX] + \mathbb{E}[Xe_y]$$
$$= \lambda_{xy} + \mathbb{E}[(\lambda_{zx}Z + e_x)e_y]$$
$$= \lambda_{xy} + \lambda_{zx} \ \mathbb{E}[e_z e_y] + \mathbb{E}[e_x e_y]$$
$$= \lambda_{xy} + \lambda_{zx} e_{zy}$$

e_x and e_y are uncorrelated
$\mathbb{E}[e_z e_y] = e_{zy}$ by definition

Paths and Covariances

There seems to be a relationship between covariances and paths in the graph.



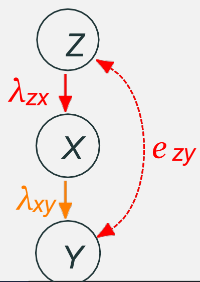
$$\sigma_{xy} = \lambda xy + \lambda_{zx}\, e_{zy}$$

The resulting terms correspond to paths between X and Y in the causal graph

Treks and Wright's Rule

The covariance between variables X and Y is the sum of the contributions of the paths between them in the causal graph, i.e. any non-self-intersecting path without colliding arrowheads ($\rightarrow\leftarrow$)

- $x \leftarrow \cdots \leftrightarrow \cdots \rightarrow y$
- $x \leftarrow \cdots \leftarrow w \rightarrow \cdots \rightarrow y$
- $x \leftarrow \cdots \leftarrow y$
- $x \rightarrow \cdots \rightarrow y$

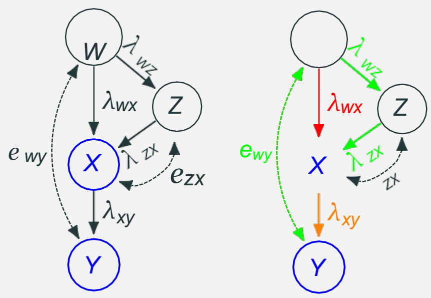


$$\sigma_{xy} = Assoc(X \rightarrow Z) + Assoc(X \leftarrow Z \leftarrow Y)$$

$$\sigma_{xy} = \lambda_{xy} + \lambda_{zx} e_{zy}$$

Reading Covariances off the Graph

The covariance between variables X and Y is the sum of open paths between them in the causal graph, so paths with no colliding arrowheads (→←)



$$\sigma_{xy} = \lambda_{xy} + \lambda_{wx}\, e_{wy} + \lambda_{zx}\lambda_{wz} e_{wy}$$

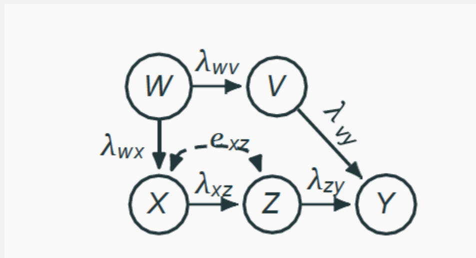
Wright's Rules

σ_{xy} = Sum of products of path coefficients along all open paths between X and Y

Wright's Rules (1921)

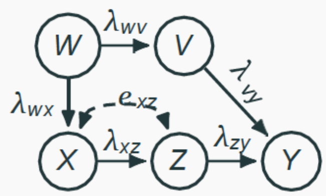
- σ_{xy} is 0 only when X and Y are d-separated.
- If there is an edge $X \overset{\alpha}{\rightarrow} Y$ in the model, then $\sigma_{xy} = \alpha +$ contributions of other paths between X and Y.
  - $\sigma_{xy} = \alpha$ if X and Y are d-separated in G_α (G with edge α removed)
- Wright's rules are defined for acyclic models (DAG)

Exercise



$\sigma_{xy} = ?$

Exercise

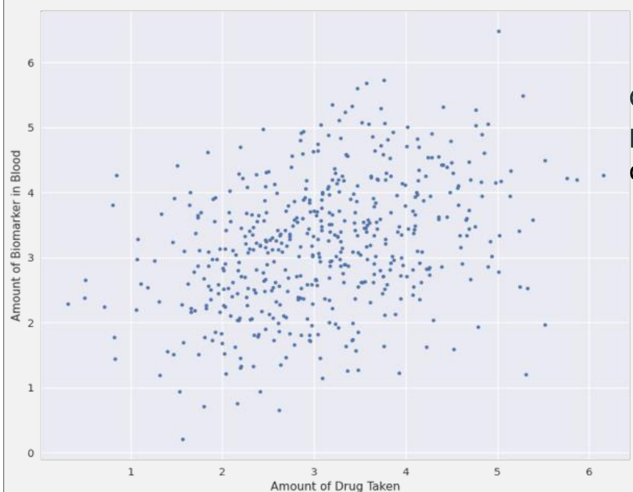


$$\sigma_{xy} = (\lambda_{xz} + e_{xz})\lambda_{zy} + \lambda_{wx}\lambda_{wv}\lambda_{vy}$$

Linear Regression

- Suppose you want to determine if a new drug is helpful for curing a disease

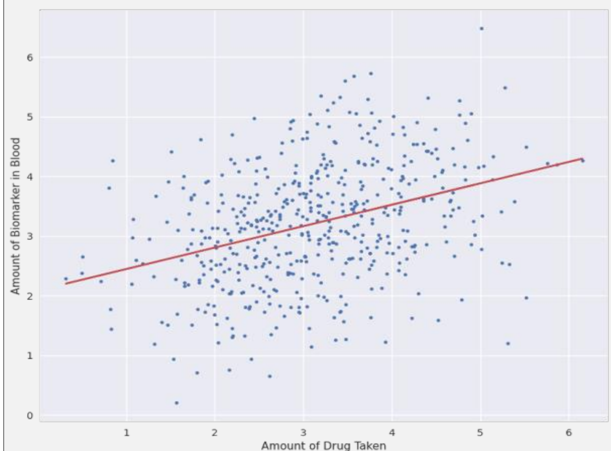
Gather data



Gather a dataset of patients who took the drug including:

1. Drug dosage
2. Blood biomarker (antibodies)

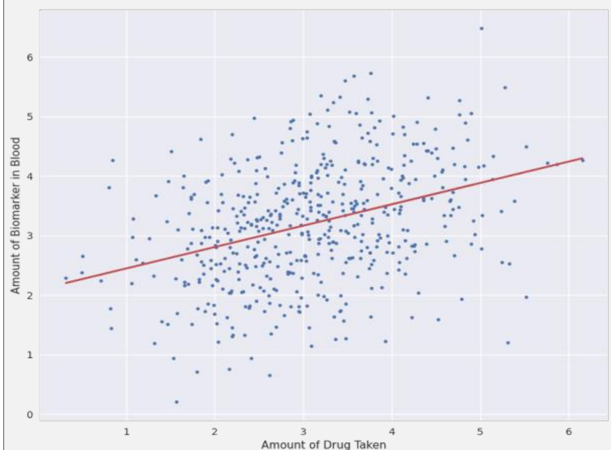
Perform regression



Perform a regression $Y = \beta X + e$ on the data, with X being drug dosage, and Y biomarker measured giving

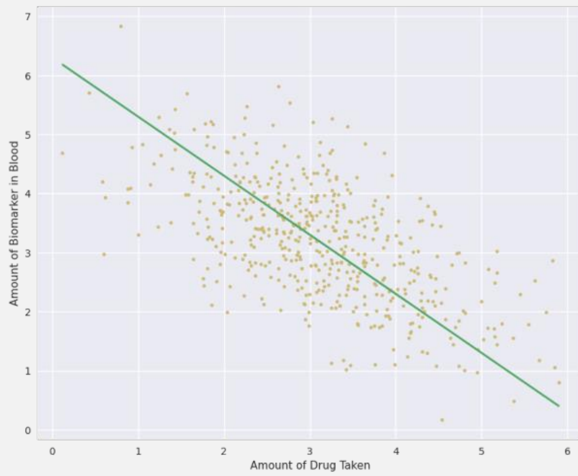
$$\beta = 0.375$$

Perform regression



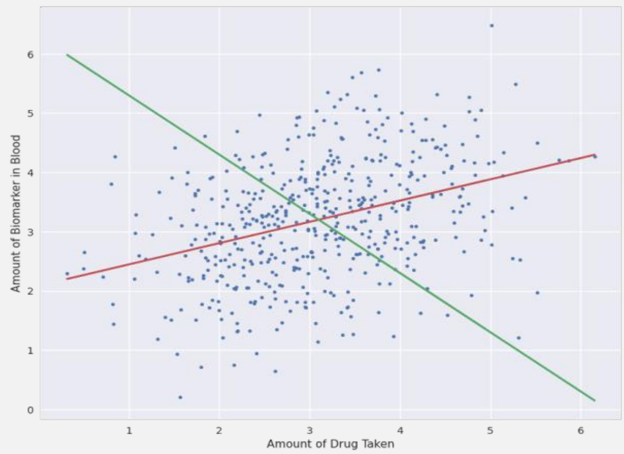
- Perform a regression $Y = \beta X + e$ on the data, with X being drug dosage, and Y biomarker measured giving $\beta = 0.375$
- Drug seems helpful, so you recommend it

What happens when the drug is given to everyone?



- When the drug is given to everyone in the population, you find a clear negative association between drug dosage and blood antibodies, with slope −1.
- This drug actually seems to hurt people!

Why did regression mislead us here?

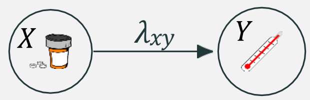


- Why was this negative effect (green line) not apparent from regression on the original dataset?
- Association $\neq$ causation!
- Can we get causation from the original dataset?

Why did regression mislead us here?

Key Assumption: Lack of Confounding

The following world model is implicitly assumed when attributing causal meaning to the regression coefficient:

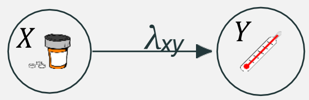


$$X := e_x$$
$$Y := \lambda_{xy} X + e_y$$

e_x, e_y independent

Why did regression mislead us here?

The following world model is implicitly assumed when attributing causal meaning to the regression coefficient:



$X := e_x$

$Y := \lambda_{xy} X + e_y$

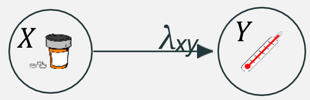
e_x, e_y independent

Regression $Y = \beta X + e$ gives correct $\beta = \lambda_{xy}$

The key assumption is lack of confounding!

Why did regression mislead us here?

The following world model (lack of confounding) is implicitly assumed when attributing causal meaning to the regression coefficient:



$$X := e_x$$
$$Y := \lambda_{xy} X + e_y$$

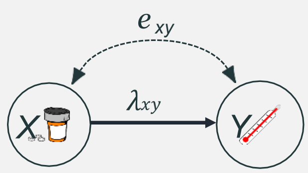
e_x, e_y independent

Covariance gives the same answer:

$$\sigma_{xy} = \mathbb{E}[XY] = \mathbb{E}[X(\lambda_{xy} X + e_y)] = \lambda_{xy}\,\mathbb{E}[XX] + \mathbb{E}[X e_y] = \lambda_{xy}$$

The True Scenario

If one is unable to ascertain the assumption of no confounding between X and Y, this is the corresponding graphical model



$$X := e_x$$
$$Y := \lambda_{xy} X + e_y$$

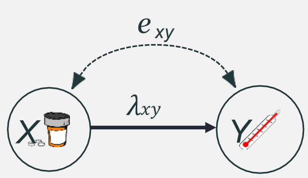
e_x, e_y correlated

May be

- The drug is expensive so mostly rich people are getting it.
- Rich people also tend to get better care overall and hence have a better chance of recovery
- But data about financial status not gathered

The True Scenario

If one is unable to ascertain the assumption of no confounding between X and Y, this is the corresponding graphical model



$X := e_x$
$Y := \lambda_{xy} X + e_y$

e_x, e_y correlated

- Regression $Y = \beta X + e$ gives a biased answer

  $\sigma_{xy} = \lambda_{xy} \mathbb{E}[XX] + \mathbb{E}[e_x e_y]$

  $\sigma_{xy} = \lambda_{xy} + e_{xy}$

- In this case, the causal effect of the drug X on blood antibodies Y is provably unidentifiable from observational data
- What can you do? Run an RCT!

What does Regression Compute?

$$Y = \beta X + e$$

- β is the regression coefficient
- What does β represent?
- Just the covariance between X and Y!

What does Regression Compute?

We want to minimize:

$$\mathbb{E}[(Y - \beta X)^2] = \mathbb{E}[YY - 2\beta XY + \beta^2 XX]$$
$$= \mathbb{E}[YY] - 2\beta\,\mathbb{E}[XY] + \beta^2\,\mathbb{E}[XX]$$
$$= 1 + \beta^2 - 2\beta\,\mathbb{E}[XY]$$
$$= 1 + \beta^2 - 2\beta\sigma_{xy}$$

Solving $\frac{\partial}{\partial\beta}(1 + \beta^2 - 2\beta\sigma_{xy}) = (2\beta - 2\sigma_{xy}) = 0$

We get: $\beta = \sigma_{xy}$

The regression coefficient is just the covariance between X and Y!

What does Regression Compute?

- The regression equation $Y = \beta X + e$ assumes $e \coprod X$
- The solution of the regression equation is: $\beta = \sigma_{xy}$.
- We will call this value r_{yx} (solved value of linear regression of Y on X)
- Knowledge of r_{yx} supports no causal claims.

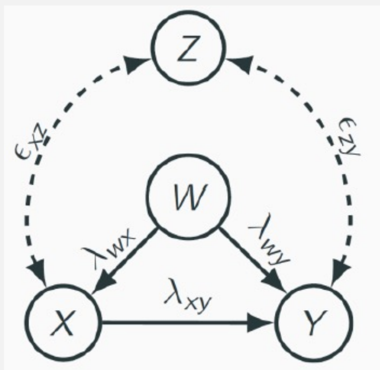
- In contrast, the structural causal model

- Corresponds to the structural equation $Y = \lambda X + e_y$
- which implies $\mathbb{E}[Y|do(X)] = \lambda X$

- The structural model makes causal claims, that is, claims about the interventional distribution which can be tested, and can be falsified.
- The SCM and regression equation look similar but have different interpretations.

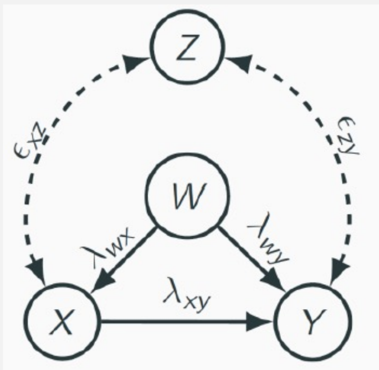
Be careful with regression

- Remember: alpha, beta denote regression coefficients and lambdas denote causal SCM parameters



Be careful with regression

- Remember: alpha, beta denote regression coefficients and lambdas denote causal SCM parameters



$$Y = \beta X + e$$

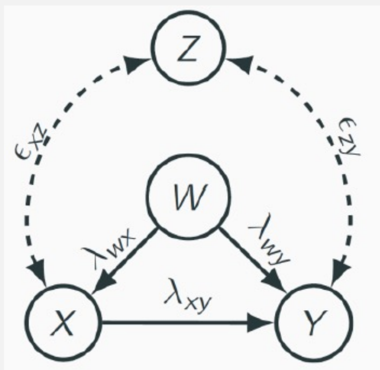
$$\beta = \sigma_{xy}$$

Show that

$$\sigma_{xy} = \lambda_{xy} + \lambda_{wx}\,\lambda_{wy}$$

Be careful with regression

- Remember: alpha, beta denote regression coefficients and lambdas denote causal SCM parameters



$\beta X + \alpha W + e$

$\beta = \sigma_{xy}$

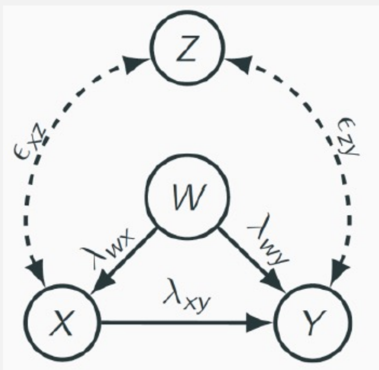
Show that

$\sigma_{xy} = \lambda_{xy}$

Note that we controlled for w

Be careful with regression

- Remember: alpha, beta denote regression coefficients and lambdas denote causal SCM parameters



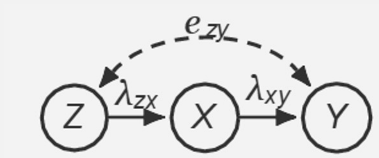
$$Y = \beta X + \alpha W + \gamma Z + e$$

$$\beta = \sigma_{xy}$$

$$\sigma_{xy} = \lambda_{xy} - \left(\frac{\epsilon_{xz}\epsilon_{zy}}{1 - \lambda_{wx}^2 - \epsilon_{xz}^2} \right)$$

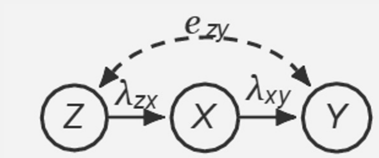
Correct use of regression

Single-Door Criterion



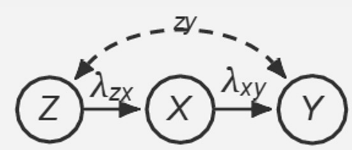
- Suppose we want to find λ_{xy}
- We have $r_{yx} = \sigma_{xy} = \lambda_{xy} + \lambda_{zx}\ e_{zy}$

Correct use of regression



- Suppose we want to find λ_{xy}
- We have $r_{yx} = \sigma_{xy} = \lambda_{xy} + \lambda_{zx}\ e_{zy}$

Regression



- What if we find the least squares regression parameters of the regression model given by

$$Y = \alpha X + \beta Z + e$$

- We find:

$$\alpha = \lambda xy \text{ and } \beta = ezy$$

Single door criterion for identification of (single) direct effects

Theorem: Single-Door Identification of (Single) Direct Causal Effects[1]

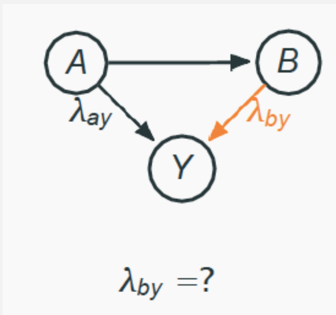
Let G be any path diagram in which λ is the path coefficient associated with the link $X \rightarrow Y$, and let G_λ denote the diagram obtained by removing $X \rightarrow Y$ from G. The coefficient λ is identifiable if there exists a set Z such that

- Z contains no descendants of Y, and
- Z d-separates X from Y in G_λ

If Z satisfies these conditions (single door criterion), $\lambda = r_{yxz}$ where r_{yxz} denotes the regression coefficient of X in the regression Y on X and Z.

Single door criterion for identification of (single) direct effects

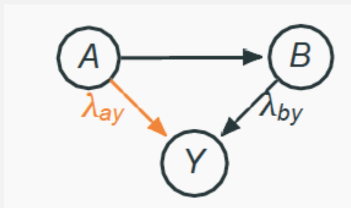
Example



$$\lambda_{by} = r_{yba}$$

Single door criterion for identification of direct causal effects

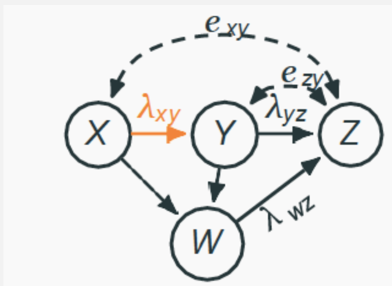
Example



$$\lambda_{ay} = r_{yab}$$

Single door criterion for identification of direct causal effects

Exercise

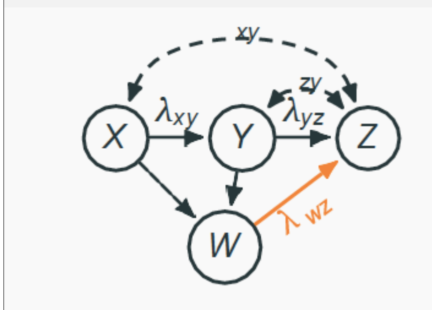


$\lambda_{xy} = r_{yx}$

- Why?
- ϕ (empty set) d–separates X from Y in $G_{\lambda_{xy}}$

Single door criterion for identification of direct causal effects

Exercise



$\lambda_{wz} = r_{zwyx}$

- Why?
- $\{X, Y\}$ *d*−separates W from Z in $G_{\lambda_{wz}}$

When can we use multiple regression to simultaneously solve for multiple path coefficients?

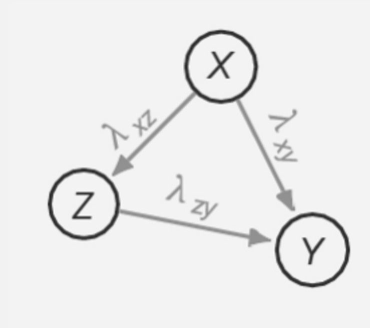
Theorem Back-Door (Identification of Total Effects) For any two variables X and Y in a causal diagram G, the total effect of X on Y is identifiable if there exists a set of measurements Z such that

1. No member of Z is a descendant of X, and
2. Z d-separates X from Y in the subgraph $G_{\underline{X}}$

Moreover, if Z satisfies the above criteria (backdoor criteria), the total effect of X on Y is given by $r_{yx \cdot z}$

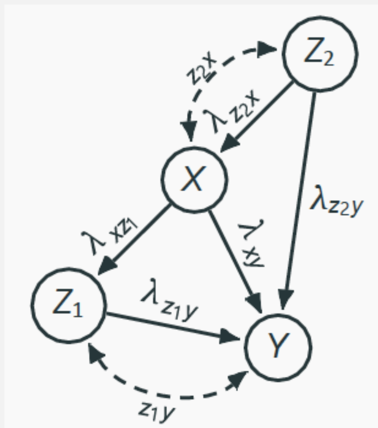
Recall that $G_{\underline{X}}$ is obtained by deleting from G, all outgoing edges from X .

Why exclude descendants of treatment from the conditioning set?



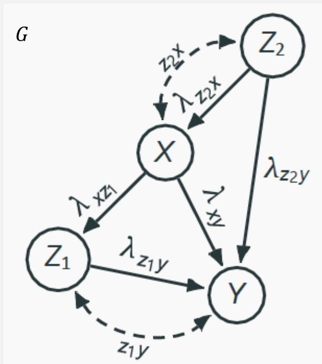
- Conditioning on the descendants of X, e.g., Z can block causal paths from X from Y through Z thus eliminating part of the total effect of X on Y

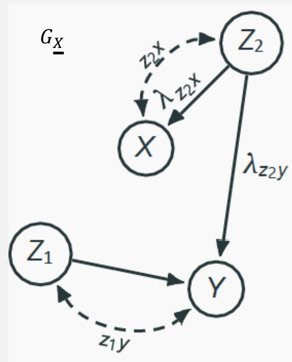
Example: total causal effect



- What is the total causal effect of X on Y?
- $\lambda_{XY} + \lambda_{XZ_1}\, \lambda_{Z_1 Y}$
- Can we find it using the back-door criterion?

Example: total causal effect of X on Y

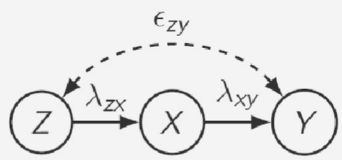




- Back-door conditioning set: $\{Z_2\}$
- $P(Y|do(X)) = r_{XYZ_2}$

Algorithms for Identifying Causal Effects in Linear Structural Causal Models

Equations for Causal Effect Identification in Linear Causal Models

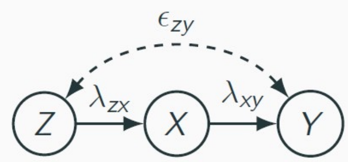


$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} = \begin{bmatrix} 1 & \lambda_{xy} + \lambda_{zx}\epsilon_{zy} & \lambda_{zx} \\ \lambda_{xy} + \lambda_{zx}\epsilon_{zy} & 1 & \lambda_{zx}\lambda_{xy} + \epsilon_{zy} \\ \lambda_{zx} & \lambda_{zx}\lambda_{xy} + \epsilon_{zy} & 1 \end{bmatrix}$$

- Note that the sigmas can be expressed in terms of lambdas using techniques previously introduced (path analysis)

PennState Institute for Computational and Data Sciences | Center for Artificial Intelligence Foundations and Scientific Applications | Artificial Intelligence Research Laboratory | CTSI Clinical and Translational Science Institute

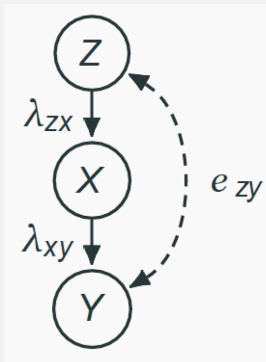
Equations for Causal Effect Identification in Linear Causal Models



$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} = \begin{bmatrix} 1 & \lambda_{xy} + \lambda_{zx}\epsilon_{zy} & \lambda_{zx} \\ \lambda_{xy} + \lambda_{zx}\epsilon_{zy} & 1 & \lambda_{zx}\lambda_{xy} + \epsilon_{zy} \\ \lambda_{zx} & \lambda_{zx}\lambda_{xy} + \epsilon_{zy} & 1 \end{bmatrix}$$

- Covariance matrix Σ is symmetric
- Only the entries in the lower or upper triangle need to be considered

Causal Effect Identification in Linear Causal Models

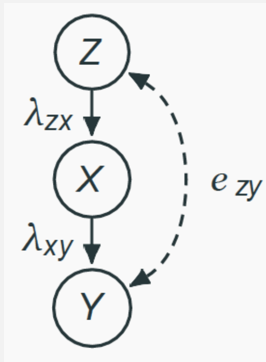


- Given a SCM and an observational dataset, is it possible to uniquely determine λ_{xy}?
- Can λ_{xy} be solved in terms of Σ?

$$\sigma_{xz} = \lambda_{zx}$$
$$\sigma_{xy} = \lambda_{xy} + \lambda_{zx} e_{zy}$$
$$\sigma_{zy} = \lambda_{zx}\lambda_{xy} + e_{zy}$$

- Σ can be estimated from the observational data (and hence known)
- The Λ need to be solved for

Causal Effect Identification in Linear Causal Models

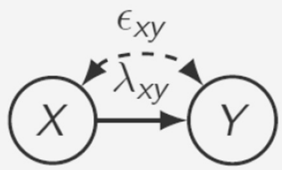


- Can Λ be solved in terms of Σ?

$$\sigma_{xz} = \lambda_{zx}$$
$$\sigma_{xy} = \lambda_{xy} + \lambda_{zx}\, e_{zy}$$
$$\sigma_{zy} = \lambda_{zx}\lambda_{xy} + e_{zy}$$

- λ_{zx} can be solved from the first equation
- Substituting λ_{zx} into the remaining 2 equations, we get 2 equations in 2 unknowns
- Hence, we can solve for Λ from Σ
- The given linear causal model can be identified from observational data

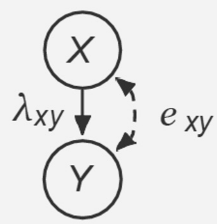
Causal Effect Identification in Linear Causal Models



$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} = \begin{bmatrix} 1 & \lambda_{xy} + \epsilon_{xy} \\ \lambda_{xy} + \epsilon_{xy} & 1 \end{bmatrix}$$

- Can Λ e solved in terms of Σ?

Causal Effect Identification in Linear Causal Models



- Can Λ e solved in terms of Σ?

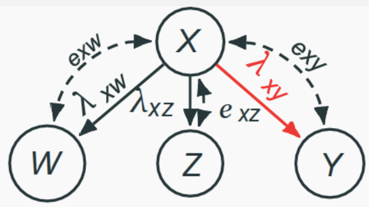
$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} = \begin{bmatrix} 1 & \lambda_{xy} + \epsilon_{xy} \\ \lambda_{xy} + \epsilon_{xy} & 1 \end{bmatrix}$$

- We have one equation in 2 unknowns

$$\sigma_{xy} = \lambda_{xy} + e_{xy}$$

- There is no unique solution for λ_{xy} or e_{xy}

Causal Effect Identification in Linear Causal Models



$\sigma_{xw} = \lambda_{xw} + e_{xw}$ $\quad$ $\sigma_{wz} = \lambda_{xw}\lambda_{xz} + \lambda_{xz}\, e_{xw} + \lambda_{xw}\, e_{xz}$

$\sigma_{xz} = \lambda_{xz} + e_{xz}$ $\quad$ $\sigma_{wy} = \lambda_{xw}\lambda_{xy} + \lambda_{xw}\, e_{xy} + \lambda_{xy}\, e_{xw}$

$\sigma_{xy} = \lambda_{xy} + e_{xy}$ $\quad$ $\sigma_{zy} = \lambda_{xz}\lambda_{xy} + \lambda_{xz}\, e_{xy} + \lambda_{xy}\, e_{xz}$

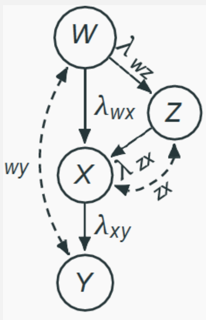
- Can we identify λ_{XY}?
- Yes, by solving the system of equations

Causal Effect Identification in Linear Causal Models

- $P(Y|do(X))$ Identifiable: Unique value of λ_{XY} consistent with observational data
- $P(Y|do(X))$ NOT identifiable: Infinite set of possible solutions for λ_{XY} consistent with observational data
- $P(Y|do(X))$ finite identifiable: if there is only a finite number of solutions for λ_{XY} that are consistent with observational data

Identification in Linear SCM

- Given an SCM and an observational dataset, is it possible to uniquely determine λ_{xy}?
- Can λ_{xy} be solved in terms of Σ?



$$\sigma_{wz} = \lambda_{wz}$$

$$\sigma_{wx} = \lambda_{wx} + \lambda_{wz}(\lambda_{zx} + e_{zx})$$

$$\sigma_{zx} = \lambda_{zx} + e_{zx} + \lambda_{wz}\lambda_{wx}$$

$$\sigma_{wy} = (\lambda_{wx} + \lambda_{wz}(\lambda_{zx} + zx))\lambda_{xy} + e_{wy}$$

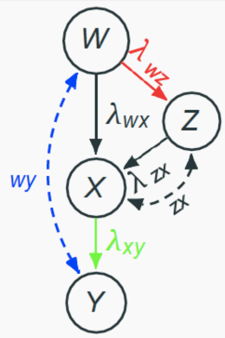
$$\sigma_{zy} = (\lambda_{zx} + e_{zx} + \lambda_{wz}\lambda_{wx})\lambda_{xy} + \lambda_{wz}$$

$$e_{wy}\ \sigma_{xy} = (\lambda_{wz}\lambda_{zx} + \lambda_{wx})\ e_{wy} + \lambda_{xy}$$

- Computer algebra approach doubly exponential in the number of variables

Identification Algorithm for Linear SCM

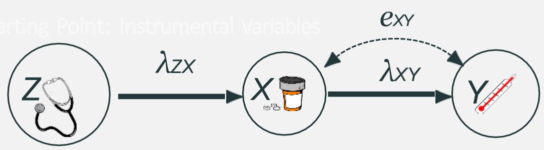
- Computer algebra approach doubly exponential in the number of variables
- Can we do better? Yes, by exploiting graphical criteria to identify solvable subsystems and/or useful substitutions



$$\sigma_{wz} = \lambda_{wz}$$
$$\sigma_{wx} = \lambda_{wx} + \lambda_{wz}(\lambda_{zx} + e_{zx})$$
$$\sigma_{zx} = \lambda_{zx} + e_{zx} + \lambda_{wz}\lambda_{wx}$$
$$\sigma_{wy} = (\lambda_{wx} + \lambda_{wz}(\lambda_{zx} + {}_{zx}))\lambda_{xy} + e_{wy}$$
$$\sigma_{zy} = (\lambda_{zx} + e_{zx} + \lambda_{wz}\lambda_{wx})\lambda_{xy} + \lambda_{wz} e_{wy}$$
$$\sigma_{xy} = (\lambda_{wz}\lambda_{zx} + \lambda_{wx}) e_{wy} + \lambda_{xy}$$

Instrumental variables

Suppose that an independent doctor's recommendation was added to the original drug/biomarker dataset



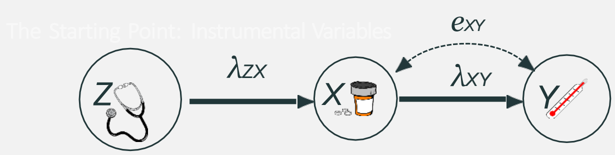
$$Z := e_Z$$
$$X := \lambda_{zx} Z + e_x$$
$$Y := \lambda_{xy} X + e_y$$

$$\sigma_{zx} = \lambda_{zx}$$
$$\sigma_{zy} = \lambda_{zx}\lambda_{xy}$$
$$\lambda_{xy} = \frac{\sigma_{zy}}{\sigma_{zx}}$$

- Is λ_{xy} identifiable non-parametrically?

Instrumental variables

Suppose that an independent doctor's recommendation was added to the original drug/biomarker dataset



A variable Z is an Instrumental variable for λ_{xy} from X to Y if

- Z is d–separated from Y in the subgraph $G_{\lambda_{xy}}$,
- Z is not d–separated from X in $G_{\lambda_{xy}}$

Conditional Instrumental variables

Conditional IV: A variable Z qualifies as a conditional IV given a set W for structural coefficient λ_{xy} from X to Y if

- W contains only non-descendants of Y

The Starting Point: Instrumental Variables

- W d-separates Z from Y in the subgraph $G_{\lambda_{xy}}$,
- W does not d-separate Z from X in $G_{\lambda_{xy}}$

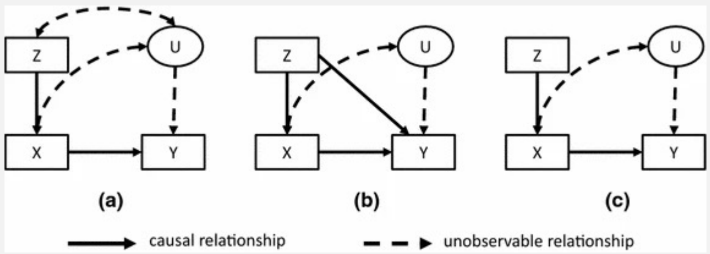
Ancestral IV: A variable Z qualifies as an ancestral IV given a set W for structural coefficient λ_{xy} from X to Y if

- W contains only ancestors of Z or Y or both that are non-descendants of Y
- W d-separates Z from Y in the subgraph $G_{\lambda_{xy}}$,
- W does not d-separate Z from X in $G_{\lambda_{xy}}$

Van der Zander et al., IJCAI 2015

Theorem: Conditional IV exists if and only if ancestral IV exists

Examples of valid and invalid instrumental variables



U – unobserved variable

Examples of situations where Z is an invalid instrument (**a**, **b**) and a valid instrument (**c**) for $X \rightarrow Y$

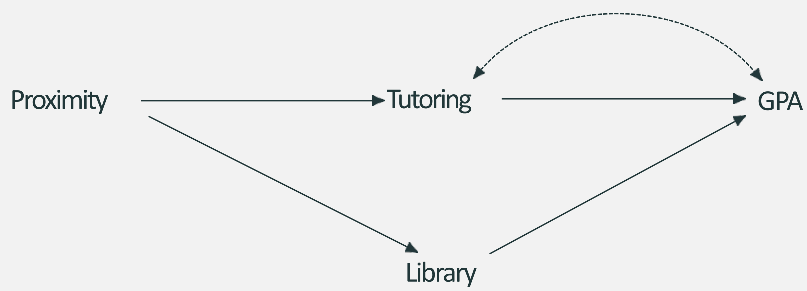
Example: Effect of tutoring program on GPA

- **Goal:** Estimate effect of tutoring program on GPA
- The relationship between attending the tutoring program and GPA may be confounded: students attending the program may care more about their grades or may be struggling with their work.
- If students are assigned dormitories at random, the proximity of the dorm to the tutors is a natural candidate instrumental variable

Example: Effect of tutoring program on GPA

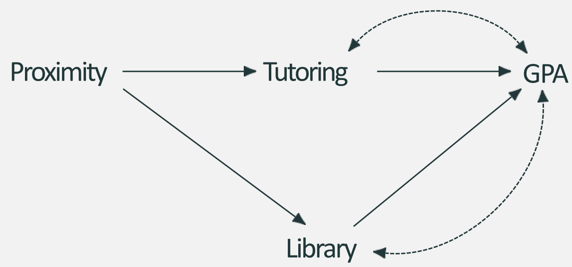
IV in Practice

What if the tutoring program is located in the college library? In that case, Proximity may also cause students to spend more time at the library, which in turn improves their GPA



Example: Effect of tutoring program on GPA

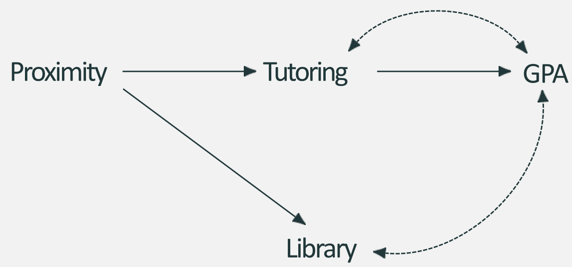
Now, suppose the student's studiousness affects time spent in the library as well as GPA.



Example: Effect of tutoring program on GPA

IV in Practice

Finally, suppose that Library Hours does not actually affect GPA because students who do not study in the library simply study elsewhere



Summary on direct and total effects in SEM

- Regression is essential for identification and causal effect computation.
- To estimate a causal effect we need to do a particular regression by specifying:
  - What variables should be included
  - Which coefficient we are interested in.
- As long as we have a Markovian system and errors are uncorrelated and every variable is measured, all structural parameters can be identified.
- But when some variables are unmeasured or errors are correlated we need to use additional tricks

Linear to non-linear causal interactions

- In nonlinear systems, on the other hand, the direct effect is defined through expressions such as $E[Y \mid do(x, z)] - E[Y \mid do(x', z)$ where $Z = z$ represents a specific stratum of all parents of Y (besides X).
- Even when the identification conditions are satisfied, and we are able to reduce the $do()$ operators (by adjustments) to ordinary conditional expectations, the result will still depend on the specific values of x, x', and z.
- The indirect effect cannot be given a definition in terms as do-expressions, since we cannot disable the capacity of Y to respond to X by holding variables constant.
- Nor can the indirect effect be defined as the difference between the total and direct effects, since differences do not faithfully reflect operations in non-linear systems.
- However, we can use do-calculus or equivalent graphical criteria for identification