Cause-Effect Models

Congyuan Duan

January 10, 2022

Contents

Contents

Definition

Definition 3.1 (Structural causal models) *An SCM $\mathfrak{C}$ with graph $C \to E$ consists of two* ***assignments***

$$C := N_C, \tag{3.1}$$

$$E := f_E(C, N_E), \tag{3.2}$$

where $N_E \perp\!\!\!\perp N_C$, that is, N_E is independent of N_C.

In this model, we call the random variables C the **cause** and E the **effect**. Furthermore, we call C a **direct cause of** E, and we refer to $C \to E$ as a **causal graph**.

Graphical Representation

Formally, a structural causal model consists of two sets of variables U and V, and a set of functions f that assigns each variable in V a value based on the values of the other variables in the model. A variable X is a direct cause of a variable Y if X appears in the function that assigns Y's value.

- The variables in U are called **exogenous variables**, meaning that they are external to the model; we choose not to explain how they are caused.
- The variables in V are **endogenous variables**. Every endogenous variable in a model is a descendant of at least one exogenous variable.

Exogenous variables cannot be descendants of any other variables.

Example: College Acceptance

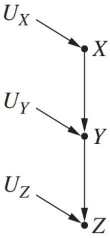
SCM 2.2.1 (School Funding, SAT Scores, and College Acceptance)

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

$$f_X : X = U_X$$

$$f_Y : Y = \frac{x}{3} + U_Y$$

$$f_Z : Z = \frac{y}{16} + U_Z$$



Contents

Cause-effect interventions

Example 3.2 (Cause-effect interventions) Suppose that the distribution $P_{C,E}$ is entailed by an SCM $\mathfrak{C}$

$$\begin{aligned} C &:= N_C \\ E &:= 4 \cdot C + N_E, \end{aligned} \tag{3.3}$$

with $N_C, N_E \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, and graph $C \to E$. Then,

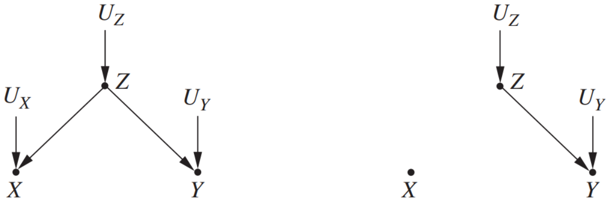
$$\begin{aligned} P_E^{\mathfrak{C}} = \mathcal{N}(0,17) \neq \mathcal{N}(8,1) = P_E^{\mathfrak{C};do(C:=2)} &= P_{E|C=2}^{\mathfrak{C}} \\ \neq \mathcal{N}(12,1) = P_E^{\mathfrak{C};do(C:=3)} &= P_{E|C=3}^{\mathfrak{C}}. \end{aligned}$$

Intervening on C changes the distribution of E. But on the other hand,

$$P_C^{\mathfrak{C};do(E:=2)} = \mathcal{N}(0,1) = P_C^{\mathfrak{C}} = P_C^{\mathfrak{C};do(E:=314159265)} \left(\neq P_{C|E=2}^{\mathfrak{C}} \right). \tag{3.4}$$

Graphical Representation

When we intervene on a variable, we remove all edges directed into it.



Interventions and Conditioning

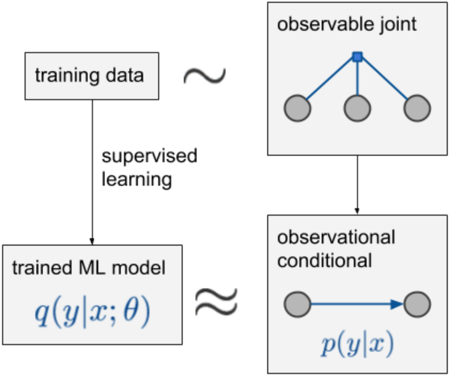
$p(y|do(x))$ and $p(y|x)$ are not generally the same.

- Conditioning $p(y|x)$: What is the distribution of Y given that I **observe** variable X takes value x.
- Interventions $p(y|do(x))$: What is the distribution of Y if I were to **set** the value of X to x.

In summary, y and x are statistically dependent and therefore seeing x allows us to predict the value of y, but y is not caused by x so setting the value of x won't effect the distribution of y.

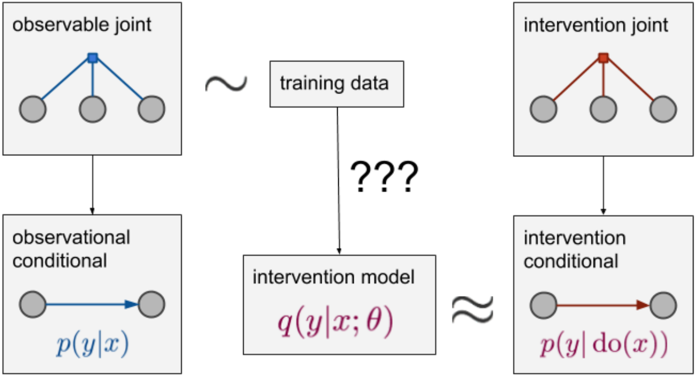
Interventions and Conditioning

We are interested in $p(y|x)$. In the simple supervised learning case, we can build a model $q(y|x;\theta)$ from the training data to approximate $p(y|x)$.



Interventions and Conditioning

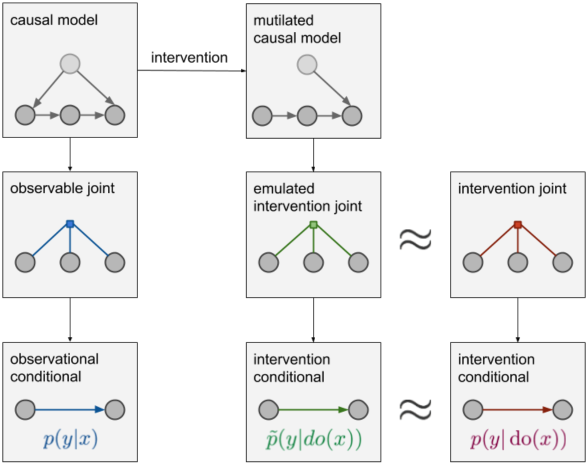
What if we're actually interested in $p(y|do(x))$ rather than $p(y|x)$? The intervention joint is a joint distribution over the same domain as p but it's a different distribution.



We want to estimate the red conditional $p(y|do(x))$ from the blue joint.

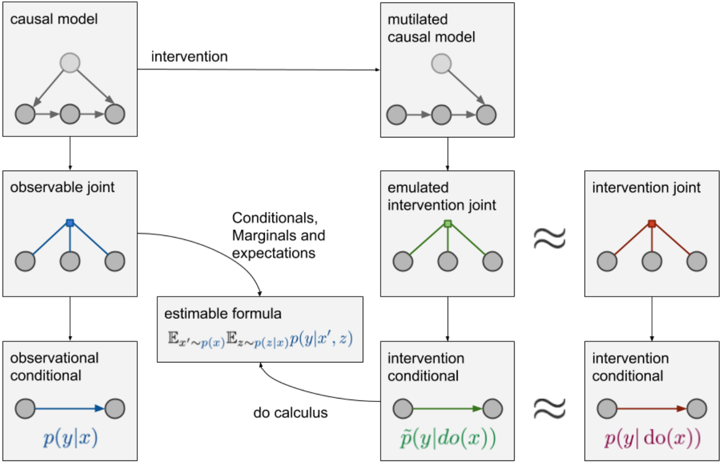
Interventions and Conditioning

If we know about the causal structure, we can establish a connection between the blue and the red joints. Then we can emulate the effect of intervention by deleting all edges that lead into nodes in a *do* operator.



Interventions and Conditioning

Do-calculus allows us to massage the green conditional distribution until we can express it in terms of various marginals, conditionals and expectations under the blue distribution.



Contents

Example: Eye disease

Example 3.4 (Eye disease) There exists a rather effective treatment for an eye disease. For 99% of all patients, the treatment works and the patient gets cured ($B = 0$); if untreated, these patients turn blind within a day ($B = 1$). For the remaining 1%, the treatment has the opposite effect and they turn blind ($B = 1$) within a day. If untreated, they regain normal vision ($B = 0$).

Which category a patient belongs to is controlled by a rare condition ($N_B = 1$) that is unknown to the doctor, whose decision whether to administer the treatment ($T = 1$) is thus independent of N_B. We write it as a noise variable N_T.

Assume the underlying SCM

$$\mathfrak{C}: \begin{array}{rcl} T & := & N_T \\ B & := & T \cdot N_B + (1-T)\cdot(1-N_B) \end{array} \tag{3.5}$$

with Bernoulli distributed $N_B \sim \mathrm{Ber}(0.01)$; note that the corresponding causal graph is $T \to B$.

A specific patient with poor eyesight comes to the hospital and goes blind ($B = 1$) after the doctor administers the treatment ($T = 1$). **"What would have happened had the doctor administered treatment T = 0?"**

Steps in Computing Counterfactuals

- Every assignment $U = u$ corresponds to a single member or "unit" in a population, or to a "situation" in nature.
- Since we have gleaned knowledge about the previously unknown noise variables for the given individual, we update the noise distributions.

(i) Abduction: Use evidence $E = e$ to determine the value of U.
(ii) Action: Modify the model, M, by removing the structural equations for the variables in X and replacing them with the appropriate functions $X = x$, to obtain the modified model, M_x.
(iii) Prediction: Use the modified model, M_x, and the value of U to compute the value of Y, the consequence of the counterfactual.

Remark: Nondeterministic Counterfactuals.

Interventions and Counterfactuals

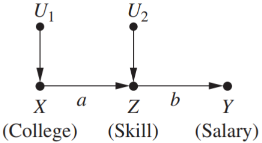
- Interventions: If I take aspirin, will my headache be cured?(population, new distribution)
- Counterfactuals: Was it the aspirin that stopped my headache?(individual)

Interventions and Counterfactuals

Consider the model

$$X = U_1 \quad Z = aX + U_2 \quad Y = bZ$$

Let $X = 1$ stand for having a college education, $U_2 = 1$ for having professional experience, Z for the level of skill needed for a given job, and Y for salary. Suppose our aim is to compute $E[Y_{X=1}|Z = 1]$, which stands for the expected salary of individuals with skill level $Z = 1$, had they received a college education.



Interventions and Counterfactuals

$X = u_1 \quad Z = aX + u_2 \quad Y = bZ$								
u_1	u_2	$X(u)$	$Z(u)$	$Y(u)$	$Y_0(u)$	$Y_1(u)$	$Z_0(u)$	$Z_1(u)$
0	0	0	0	0	0	ab	0	a
0	1	0	1	b	b	$(a+1)b$	1	$a+1$
1	0	1	a	ab	0	ab	0	a
1	1	1	$a+1$	$(a+1)b$	b	$(a+1)b$	1	$a+1$

Using this table, we can verify that

$$E[Y_1|Z=1] = (a+1)b$$

$$E[Y_0|Z=1] = b$$

$$E[Y|do(X=1), Z=1] = b$$

$$E[Y|do(X=0), Z=1] = b$$

and

$$E[Y_1 - Y_0|Z=1] = ab \neq 0$$

Contents

Reference

https://www.inference.vc/untitled/

Pearl J, Glymour M, Jewell N P. Causal inference in statistics: A primer[M]. John Wiley & Sons, 2016.