Causality 2022, Assignment 1

Jonas Peters, TA: Sorawit (James) Saengkyongam

handed out: 27 April 2022 at 15:00
hand in until: 4 May 2022 at 09:00

Please specify which 2 exercises should count for passing the assignment sheet

Exercise 1 (d-separation)**.** Prove that one can d-separate any two nodes in a DAG $\mathcal{G}$ that are not directly connected by an edge.

(Hint: you may start with the following: "Let A and B be nodes of $\mathcal{G}$ that are not directly connected by an edge. We wish to find a set $S \subseteq V$ such that $A \perp\!\!\!\perp_{\mathcal{G}} B \,|\, S$, where $\perp\!\!\!\perp_{\mathcal{G}}$ denotes d-separation with respect to $\mathcal{G}$. Since $\mathcal{G}$ is acyclic, at least one of the following two statements holds true i) $A \in \mathbf{ND}_B$ or ii) $B \in \mathbf{ND}_A$. Without loss of generality, ...")

Exercise 2 (d-separation)**.** Find all DAGs over nodes $V = \{A, B, C, D\}$ that satisfies the following (and only that) d-separation statement:

$$A \;\; d\text{-sep. } C,$$

which is short-hand for A d-sep. $C \,|\, \emptyset$. That is, we are looking for all DAGs over V, such that U d-sep. $W \,|\, Z$ if and only if $\{U, W\} = \{A, C\}$ and $Z = \emptyset$. Prove that your answer is correct.

Exercise 3 (Observations and Interventions)**.** Consider the following SCM $\mathfrak{C}$.

$$\begin{aligned} X &:= N_X \\ Y &:= -X + N_Y \\ Z &:= X + 2Y + N_Z \end{aligned}$$

with $N_X, N_Y, N_Z \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$.

a) Draw the graph corresponding to the SCM.
b) Compute the joint (observational) distribution $P^{\mathfrak{C}}_{X,Y,Z}$ entailed by the SCM.
c) We now consider an intervention, where the variable Y is intervened on. It is set to a random variable $M \sim \mathcal{N}(0, 2)$ that is independent of N_X, N_Y, N_Z. Compute the joint intervention distribution $P^{\mathfrak{C};\, do(Y:=M)}_{X,Y,Z}$. Which of the three marginal distributions are different compared to b)?

Exercise 4 (prediction)**.** a) Consider a random vector $\mathbf{X} = (X_1, \ldots, X_d)^t \in \mathbb{R}^d$ and a target random variable $Y \in \mathbb{R}$. Assume that all these random variables have mean zero and finite variance. Prove that

$$\beta^0 := \mathrm{cov}(\mathbf{X})^{-1}\mathrm{cov}(\mathbf{X}, Y)$$

solves the optimization problem

$$\beta^0 = \underset{\beta}{\mathrm{argmin}}\, \mathbb{E}(Y - \beta^t \mathbf{X})^2.$$

It is thus the "population" version of the least squares regression problem. Here, the $d \times d$-matrix $\mathrm{cov}(\mathbf{X})$ is assumed to be invertible, and the jth entry of $\mathrm{cov}(\mathbf{X}, Y) \in \mathbb{R}^d$ is $\mathrm{cov}(X_j, Y)$.

b) Consider the example from Exercise 3 (the observational distribution) and compute β^0 for $\mathbf{X} = (X, Z)$. One can show (you do not need to do that) that even though Z is an effect and not a cause of Y, it can be a better predictor than X (e.g., in terms of reduction of the mean squared error).