Homework for Chapter 21: A Gallery of Rogues

1. When doing difference-in-differences, we worry about whether the treated and control groups seem to be following the same trend before treatment. We don’t need to worry about this so much with synthetic control.
   1. Why not?
   2. (Bonus, not covered in the chapter) If the prior trends didn’t match in a synthetic control estimation, what conclusion might we come to?
2. You hear that, in 1980, a school district in India changed the way they teach math considerably, using a new curriculum that nobody else was using. You want to know the effect of this curriculum on whether students entered more math-oriented jobs as adults. Which of the following features of this analysis would make you concerned about using synthetic control to answer this question?
   1. You are looking at an outcome that occurs long after treatment.
   2. Your data set consists of students from the school years 1979 and 1981.
   3. You have a large list of other Indian school districts that didn’t change curriculum but aren’t sure which would make the best control groups
   4. Only one district changed its curriculum
3. Matrix completion uses information from other observations in the same time period, or the same group, to predict missing “untreated” outcomes.
   1. Is this a back-door-closing or front-door-isolating procedure?
   2. Intuitively, what back doors does this process allow us to close (by back-door closing) or ignore (by front-door isolating), whichever one applies given your answer to a?
4. In this problem you will follow the SGS causal discovery algorithm as described in the book and figure out the correct causal diagram. There are four relevant variables in your diagram: A, B, C, and D. You have theoretical reason to believe that C -> B should be on the graph (so you can start with that), but aren’t sure about the rest. Given the following findings from your data, draw the proper diagram.

* Without any controls, A and B have a nonzero relationship
* Controlling for C, or D, or C and D, A and B have a nonzero relationship (A -> B or A <- B)
* Without any controls, B and D have a nonzero relationship
* Controlling for A, B and D have a zero relationship (B ! D)
* Without any controls, C and A have a nonzero relationship
* Controlling for B, C and A have a zero relationship (no C -> A or C <- A, and also not both A -> B and C -> B. So since C -> B, must be B -> A)
* Controlling for B and D, C and A have a nonzero relationship (A -> D, C -> D)

1. Describe an example of a policy intervention that would be made much more effective by an estimate of how heterogeneous treatment effects vary at the individual level, relative to just having an aggregate measure of an average treatment effect like ATE, ATT, or LATE.