Step2_CausalStructures

March 11, 2020

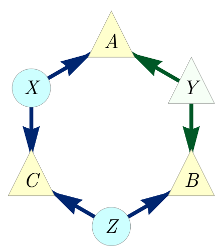
1 Application to causal structure discovery

See the paper about the technique. Read in particular Section I, section II can be skimmed, as well as III.A. The main idea can be understood by reading III.B; we'll use Example 1 in our description below; read the pen-and-paper proof in Example 1, we'll formulate it as a linear program.

Optional, read III.C, skimming through the technical definitions, pay attention to Example 4. The use of linear programming is detailed in Section IV, particularly IV.B. Skip the rest. Appendix A provides the data we need.

1.1 An example of causal structure

Consider the following causal structure:



scen15DAG.svg

where we observe the variables A, B and C that take the values $a, b, c = 0, 1$. We observe the following correlations:

$$P_{ABC}(a,b,c) = \begin{cases} 1/2, & \text{if } a = b = c, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

We try to understand whether this perfectly correlated distribution can arise in a causal structure where the variables A, B and C only depend on information that is shared with another party only.

1.1.1 A causal model

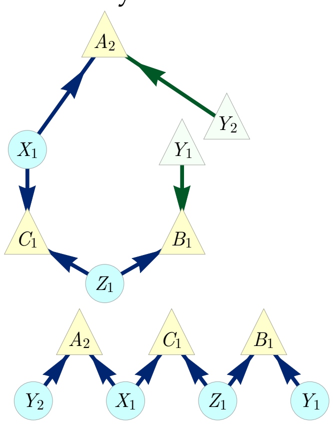
Thus, there are unobserved variables X, Y, Z, with distributions $P_X(x)$, $P_Y(y)$, $P_Z(z)$, such that the variable A is fully described by $P_{A|XY}(a|x,y)$, the variable B by $P_{B|YZ}(b|y,z)$ and the variable C by $P_{C|XZ}(c|x,z)$, and we have

$$P_{ABC}(a,b,c) = \sum_{xyz} P_X(x)P_Y(y)P_Z(z)P_{A|XY}(a|x,y)P_{B|YZ}(b|y,z)P_{C|XZ}(c|x,z). \tag{2}$$

Now, assume that the variables x, y, z are integers between 0 and N. Then, testing if our P_{ABC} has a model of the form (2) would be a polynomial feasibility problem (and a hard one!, already for $N \geq 3$). But we do not even know the type of the unobserved variables X, Y, Z (still, see https://arxiv.org/abs/1709.00707).

1.1.2 Test using the "inflation technique" which maps to LP

We will use another method, amenable to linear programming. We will make a (numerical) proof by contradication. Assume that P_{ABC} has a model of the form (2). Then, we can imagine a variation on that model, where we duplicate the variable Y, and wire the relations between the variables a bit differently.



This is called an *inflated* scenario. There, we obtain the slightly different correlations:

$$P_{A_2B_1C_1}(a_2,b_1,c_1) = \sum_{x_1y_1y_2z} P_{X_1}(x_1)P_{Y_1}(y_1)P_{Y_2}(y_2)P_{Z_1}(z_1)P_{A_2|X_1Y_2}(a_2|x_1,y_2)P_{B_1|Y_1Z_1}(b_1|y_1,z_1)P_{C_1|X_1Z_1}(c_1|x_1,z_1). \tag{3}$$

Note that, however, the marginal distribution of the inflated correlations

$$P_{A_2C_1} = \sum_{b_1} P_{A_2B_1C_1}(a_2,b_1,c_1) = \sum_{x_1y_2z} P_{X_1}(x_1)P_{Y_2}(y_2)P_{Z_1}(z_1)P_{A_2|X_1Y_2}(a_2|x_1,y_2)P_{C_1|X_1Z_1}(c_1|x_1,z_1) \tag{4}$$

has the same form as the marginal distribution of the original scenario

$$P_{AC}(a,c) = \sum_{b} P_{ABC}(a,b,c) = \sum_{xyz} P_X(x)P_Y(y)P_Z(z)P_{A|XY}(a|x,y)P_{C|XZ}(c|x,z). \tag{5}$$

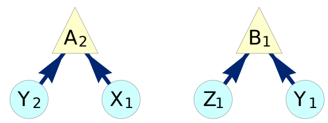
We thus have

$$P_{AC}(i,k) = P_{A_2C_1}(i,k), \qquad \forall i,k\,. \tag{6}$$

The same argument holds for

$$P_{BC}(j,k) = P_{B_1C_1}(j,k), \qquad \forall j,k\,. \tag{7}$$

Now, let us examine $P_{A_2B_1}(a_2,b_1) = \sum_{c_1} P_{A_2B_1C_1}(a_2,b_1,c_1)$. It corresponds, after removal of C_1, to the graph:



Marginal.svg

where the variables A_2 and B_1 are independent. We thus have:

$$P_{A_2B_1}(a_2,b_1) = P_{A_2}(a_2)P_{B_1}(b_1) \tag{8}$$

which we can now match with the original problem:

$$P_{A_2B_1}(i,j) = P_A(i)P_B(j), \qquad \forall i,j \;. \tag{9}$$

1.2 The linear program

Writing all these constraints together, we have ($\forall i,j,k$ is implicit):

$$\sum_j P_{A_2B_1C_1}(i,j,k) = \sum_j P_{ABC}(i,j,k), \tag{10}$$
$$\sum_i P_{A_2B_1C_1}(i,j,k) = \sum_i P_{ABC}(i,j,k), \tag{11}$$
$$\sum_k P_{A_2B_1C_1}(i,j,k) = P_A(i)P_B(j), \tag{12}$$
$$P_{A_2B_1C_1}(i,j,k) \geq 0 \tag{13}$$

Now, the inflated correlations $P_{A_2B_1C_1}$ may obey additional constraints, but we remark that the constraints listed above correspond to a linear program in the primal form: indeed, the right-hand side of the equations are constant values that depend only on the coefficients $P_{ABC}(i,j,k)$ which are known.

$$\begin{aligned} \text{minimize} \quad & 0 \\ \text{over} \quad & \vec{v} \in \mathbb{R}^n \\ & M\vec{v} = \vec{b} \\ & \vec{v} \geq 0 \end{aligned} \tag{14}$$

where the objective is trivial, the constraint right-hand side $\vec{b}$ is the only part of the problem that depends on P_{ABC}, and the matrix M only depends on the problem structure (matching the marginals).

Now, if this linear program is infeasible, it proves by contradiction that no model exists for the original problem (because original problem has model => inflation has a model).

Homework 1 Write the linear program using Convex.jl, and verify if the distribution P_{ABC} is compatible with the inflation (hint: it should not).

You can use the numerically better behaved variant that has the slack variable z:

$$
\begin{aligned}
\text{maximize} \quad & z \\
\text{over} \quad & z \in \mathbb{R}, \vec{v} \in \mathbb{R}^n \\
& M\vec{v} = \vec{b} \\
& \vec{v} \geq z
\end{aligned}
\tag{15}
$$

Homework 2 For which values of t the following distribution is compatible with the inflated model?

$$
P_{ABC}(a,b,c) = \begin{cases} 1/2t, & \text{if } a = b = c, \\ (1-t)/6, & \text{otherwise.} \end{cases}
\tag{16}
$$

Homework 3 Test the distribution:

$$
P_{ABC}(a,b,c) = \begin{cases} 1/3, & \text{if } a + b + c = 1, \\ 0, & \text{otherwise.} \end{cases}
\tag{17}
$$

This distribution should be compatible with the inflation above; nevertheless it is not compatible with the causal structure we test (see Example 2 of the paper). Why is the linear program feasible then?

Homework 4 Solve one of the following questions:

- Consider the dual problem of (15). How to interpret the dual variables and the dual objective? Read the part about infeasibility certificates in the Mosek Cookbook, section 2.3. Can you derive a causal compatibility inequality as in Example 4 of the paper)?
- Implement the Spiral inflation given in FIG. 3 of the paper, and verify that the distribution (17) is incompatible.

[ ]: