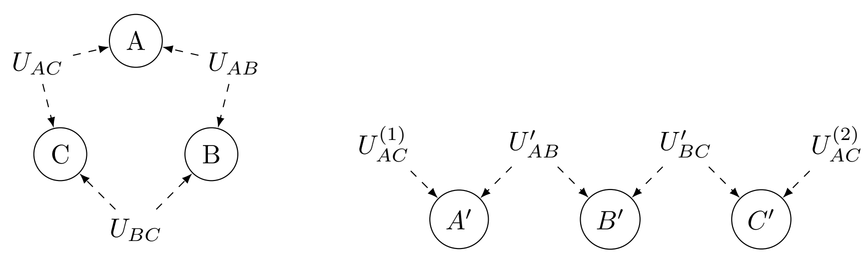
Meta-Review for "The Inflation Technique for Causal Inference with Latent Variables" (DGJCI.2017.0020)

January 1, 2018

The paper derives constraints for compatibility of discrete distributions with arbitrary causal diagrams, even if the graphs do not have d-separation constraints. The constraints derived are related to Instrumental Variable constraints as derived in Pearl [1995]. This is achieved by generating modified graphs (called "inflations") which have a subset of marginals compatible with the original graph. These modified graphs sometimes enable using d-separation statements that do not exist in the original graph to determine independence of variables. In turn, these independencies can sometimes be used to show a given distribution's marginals are incompatible with the modified graph. Finally, since the modified graph is generated such that marginals match with the original graph, the authors' lemma 4 claims that the given distribution is also incompatible with the original graph. The remaining of this note is organized as follows: a short overview of the method is provided in Sec. 1, challenges due to the notation is presented in Sec. 2, and my general observations and recommendation are provided in Sec. 3.

1 Overview

As an example, the authors used the "triangle scenario" to demonstrate the ideas behind their method (notation changed here to use conventions from Pearl):



From left to right, we have the triangle scenario, and a candidate "inflation" of the triangle scenario (to be described in a moment). One thing to notice is that there are no independence constraints possible to write down for the triangle scenario, which means that this structure seems not to be testable from data.

However, the problem may have some possible solution using the "inflation" construct. To witness, we construct a candidate "inflation" graph as follows: Take the original triangle scenario and split U_{AC} into two variables which are distributed identically to U_{AC}, where one of the variables only has an arrow to A, and the other to C. This is shown in the inflation above[1].

In this situation, it's not difficult to see that 2 marginal distributions can be matched, which follows since they are generated by identical ancestral sets of random variables:

$$P(AB) = P(A'B') \qquad\qquad P(BC) = P(B'C')$$

[1]Please note that the authors have a procedure for generating these inflations, as well as rules they must satisfy, which I have omitted to simplify the core idea.

This is valid for any distribution compatible with the original graph, meaning that if the original graph is compatible with a given distribution, there is a corresponding inflation compatible with these marginals.

The key insight is to use the converse. If it can be shown that there exists no distribution over $A'B'C'$ which matches the marginals from a candidate distribution (i.e., the inflated graph is incompatible), then the original graph should be incompatible as well.

To clarify, suppose the following candidate joint probability distribution:

$$P(ABC) = \begin{cases} \frac{1}{2} & \text{if } A = B = C \\ 0 & \text{otherwise} \end{cases}$$

Its the marginals are:

$$P(AB) = P(BC) = P(A'B') = P(B'C') = \begin{cases} \frac{1}{2} & \text{if both variables } = \\ 0 & \text{otherwise} \end{cases}$$

Looking at the marginals in the inflated graph, the only joint that satisfies this is one where all three variables are =, and all other cases have probability 0. However, C' is independent of A' in this case, so there is a situation where $C' \neq A'$, which shows that the triangle scenario is incompatible with the given graph.

The authors go on to generalize this idea to create a procedure for deriving inequalities that must be satisfied for discrete distributions to be compatible with the given graph.

Continuing the triangle scenario example shown above, assuming that the variables have domains $\{-1, 1\}$, they derive a set of inequalities on pg. 21. To do this, they generate an inequality that is always satisfied on 3 variables with such domains, and then exploit that A' and C' are independent in the inflated graph:

$$\mathbf{E}[AB] + \mathbf{E}[BC] \leq 1 + \mathbf{E}[AC] \qquad \Rightarrow \qquad \mathbf{E}[AB] + \mathbf{E}[BC] \leq 1 + \mathbf{E}[A]\mathbf{E}[C]$$

The authors also show the inflation method's relation to the Bell scenario in Physics, rewriting existing inequalities in the language of inflations of a causal graph, and generating new ones exploiting their method. Finally, the authors spend some time describing how their method could possibly be adapted to Generalized Probabilistic Theories, in particular to "Quantum Causal Models".

2 Notation & Readability

While the content of the paper is interesting and relevant, there are some critical issues with the paper's form. In particular, the notation and definitions differ from established literature in the field of Causality. It's important to note that this note in no way diminishes the significance of the results, and, indeed, there is a direct correspondence between established definitions/notation and the notation used in this work. However, since JCI is a journal targeted at researchers in many empirical fields, and the language of causality unifies and allows for their conversation, I believe that the effort of mapping between notations and definitions rests on the authors' shoulders rather than readers.

What follows is a list of the most critical issues in this regard, as well as suggestions of how to fix them.

i. Causal Model, Structure, and Parameters

The first, and arguably most critical example of this is section II (pg 5-6). Definitions of "causal model", "causal structure", and "causal parameters" are given without reference to previous literature. These are technical terms, and usually refer to specific definitions. It is initially confusing for a reader familiar with the literature and the previous definitions to see them redefined with different notation and variable names.

For reference, I have included the definition of "causal model" from Pearl [2003].

Definition 1. Structural Causal Model, Pearl [2003], pg 203

A **causal model** is a triple:

$$M = < U, V, F >$$

where:

1. U is a set of background variables (also called exogenous), that are determined by factors outside the model;

2. V is a set $\{V_1.V_2,...,V_n\}$ of variables, called endogenous, that are determined by the variables in the model - that is, variables in $U \cup V$; and
3. F is a set of functions $\{f_1, f_2, ..., f_n\}$ such that each f_i is a mapping from (the respective domains of) $U \cup PA_i$ to V_i, where $U_i \subseteq U$ and $PA_i \subseteq V \setminus V_i$ and the entire set F forms a mapping from U to V. In other words, each f_i in $v_i = f_i(pa_i, u_i)$ for $i = 1, ..., n$ assigns a value to v_i that depends on (the values of) a select set of variables in $V \cup U$, and the entire set F has a unique solution $V(u)$.

Each causal model M can be associated with a directed graph $G(M)$, in which each node corresponds to a variable and the directed edges point from members of PA_i and U_i towards V_i. We call such a graph the "causal diagram" associated with M.

In particular, notice that observed variables belong to the set V, and latent variables belong to the set U. This made the paper's usage of U and V on pg 5-8 as arbitrary subsets of the nodes initially confusing and distracting.

While there are multiple variants of the definition, most definitions which include latent variables fit closely with the above. I suggest modifying the definitions to be directly compatible with existing usage.

ii. Physics audience

As noted by the reviewers, the paper is very geared towards the physics, and even more specifically, the quantum community. There are expression very difficult to parse for the non-physicists, for instance, "Space-like separation" (page 11), which is not defined in any place in the paper.

More broadly, since the Bell scenario seems to be one of the big motivations for this work, a couple more sentences explaining it to an audience unfamiliar with the details of quantum theory would go a long way towards showing its use in Physics, and further clarify the importance of this work. Still, the applicability to Physics should be clearly separated from the real contribution to causal inference, which is immediately more broad. In particular, perhaps the authors can create one small section just for the physics readers, and remove all the other references from the paper.

Another example, the references to expectations and correlations follow standard physics notation. I recommend the authors changing $< AB >$ and $< A >$ to the notations more commonly used in causality, computer science, and statistics: $\mathbf{E}[AB]$ and $\mathbf{E}[A]$. This can be seen starting from eq. 30 on pg 13.

My intent is not to be exhaustive here since the paper was written with some difference audience in mind. In practice, I suggest the authors to shorten the references to Physics and try to, more broadly, consider JCI's audience.

iii. Main results

Lemma 4 is one of the most critical results of the paper. While it is likely true given the discussion above it in the text, a formal proof is desirable. You may be able to use the text above as a basis to prove the lemma. Also, perhaps you can call one of the main results of the paper by "theorem".

Another important contribution of this work is the generation os the inflation graph, which should be made more explicit. Give a more explicit construction and provide a better discussion of its computational complexity (e.g., linear, exponential).

Regarding the scope of the contribution, please clarify whether the result holds as it is currently stated for both discrete and continuous domains. Furthermore, the triangle example is useful for conveying the idea of the method, but we know that new constraints appear in the distribution whenever deterministic relations are present (see Pearl & Geiger D-separation around 1990's). The triangle example as initially provided seems to suggest that new equality constraints are being constructed, but this doesn't seem to be true. In reality, I believe the main contribution is of new inequality constraints and, as I understand, unrelated to the deterministic relations in this contrived example. Or, does the proposed method require that non-positivity holds (i.e., entries in the joint distribution with zero mass)? If not, please provide an intuitive example where determinism is not exploited and "confounding" the main result.

3 Overall Comments

The paper provides a novel, interesting, and possibly very impactful result to both causality and physics. While the paper provided a method for generating constraints (apparently in exponential time) for a given inflation, I did not see mention of how to choose these inflations in the first place, other than intuition. I am curious as to the applicability of this method to more complicated graphs, both due to the difficulty of generating inequalities, and to the lack of guidance towards which inflations might be useful for testing compatibility with given distributions. However, section 4D suggests that there exist heuristics which could be built upon to alleviate these issues. The many examples in the paper were a positive (see note above regarding determinism), which made it possible to comprehend the method even with unfamiliar notation.

The notation is clearly based on usage common in Physics. While this is not bad in itself, JCI is a journal of Causality, and unfortunately, the differing notations and definitions in our two fields make reading the paper more involved for someone outside Physics than I think a paper in JCI should be. My main concern is that almost no reader will have the energy to pass the barrier of such a different notation.

After addressing the reviewers questions and suggestions for improvement, I would be glad to recommend the paper for acceptance. In particular, I expect the authors to make the paper readable for the JCI's audience, including the suggestion provided in the section "Notations & Readability."

References

Judea Pearl. On the testability of causal models with latent and instrumental variables. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 435–443. Morgan Kaufmann Publishers Inc., 1995.

Judea Pearl. Causality: models, reasoning and inference. 2003.