Adjacency Faithfulness and Conservative Causal Inference

Joseph Ramsey Peter Spirtes
Department of Philosophy
Carnegie Mellon University
Pittsburgh, PA 15213
{jdramsey, ps7z}@andrew.cmu.edu

Jiji Zhang
Division of Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125
jiji@hss.caltech.edu

Abstract

Most causal discovery algorithms in the literature exploit an assumption usually referred to as the Causal Faithfulness or Stability Condition. In this paper, we highlight two components of the condition used in constraint-based algorithms, which we call "Adjacency-Faithfulness" and "Orientation-Faithfulness." We point out that assuming Adjacency-Faithfulness is true, it is possible to test the validity of Orientation-Faithfulness. Motivated by this observation, we explore the consequence of making only the Adjacency-Faithfulness assumption. We show that the familiar PC algorithm has to be modified to be correct under the weaker, Adjacency-Faithfulness assumption. The modified algorithm, called Conservative PC (CPC), checks whether Orientation-Faithfulness holds in the orientation phase, and if not, avoids drawing certain causal conclusions the PC algorithm would draw. However, if the stronger, standard causal Faithfulness condition actually obtains, the CPC algorithm outputs the same pattern as the PC algorithm does in the large sample limit.

We also present a simulation study showing that the CPC algorithm runs almost as fast as the PC algorithm, and outputs significantly fewer false causal arrowheads than the PC algorithm does on realistic sample sizes.

1 MOTIVATION: FAITHFULNESS DECOMPOSED

Directed acyclic graphs (DAGs) can be interpreted both probabilistically and causally. Under the causal interpretation, a DAG represents a causal structure such that A is a direct cause of B just in case there is a directed edge from A to B in . Under the probabilistic interpretation, a DAG , also referred to as a Bayesian network, represents a probability distribution P that satisfies the Markov Property: each variable in is independent of its non-descendants conditional on its parents. The Causal Markov Condition

i. every non-collider on p is not a member of **C**;

ii. every collider on p is an ancestor of some member of **C**.

Two sets of variables **A** and **B** are said to be **d-separated** by **C** if there is no active path between any member of **A** and any member of **B** relative to **C**.

A well-known important result is that for any three disjoint sets of variables **A**, **B** and **C** in a DAG , **A** and **B** are entailed (by the Markov condition) to be independent conditional on **C** if and only if they are d-separated by **C** in . So the causal Faithfulness condition can be rephrased as saying that for every three disjoint sets of variables **A**, **B** and **C**, if **A** and **B** are not d-separated by **C** in the causal DAG, then **A** and **B** are not independent conditional on **C**.

Two simple facts about d-separation are particularly relevant to our purpose (see e.g. Neapolitan 2004, pp. 89 for proofs):

Proposition 1. Two variables are adjacent in a DAG if and only if they are not d-separated by any subset of other variables in the DAG.

Call a triple of variables $\langle X, Y, Z \rangle$ in a DAG an unshielded triple if X and Z are both adjacent to Y but are not adjacent to each other.

Proposition 2. In a DAG, any unshielded triple $\langle X, Y, Z \rangle$ is a collider if and only if all sets that d-separate X from Z do not contain Y; it is a non-collider if and only if all sets that d-separate X from Z contain Y.

Below we focus on two implications of the Causal Faithfulness Condition, easily derivable given Propositions 1 and 2. We call them Adjacency-Faithfulness and Orientation-Faithfulness, respectively.

Implication 1 (Adjacency-Faithfulness ar Imp60621660tateand be 29

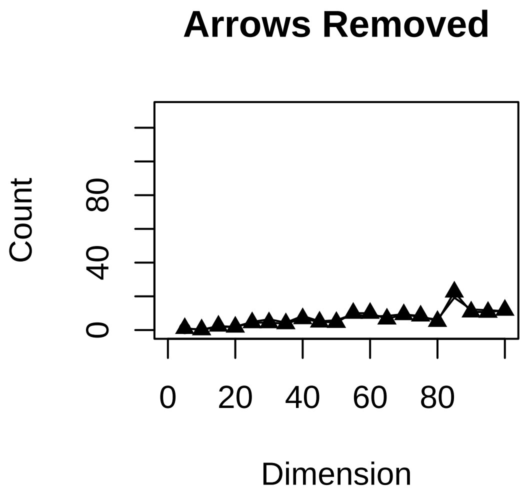
the causal Markov and Adjacency-Faithfulness conditions are both satisfied, but Orientation-Faithfulness is not true of the triple $\langle A, B, C \rangle$. Now, given the correct conditional independence oracle, the PCTj [illegible]

X2 X3

X1

X

ment can be made. Instead we wish to show that the CPC algorithm in practice performs better than the PC algorithm, regardless of whether Orientation-Faithfulness holds or not. That is, even when the data are generated from a


Arrows Removed
Count
0
40
80
0
20
40
60
80
Dimension