A Bayesian Approach to Constraint Based Causal Inference

Tom Claassen TOMC@CS.RU.NL
Radboud University, Nijmegen

Tom Heskes TOMH@CS.RU.NL
Radboud University, Nijmegen

Keywords: causal discovery, structure learning, graphical models

We target the problem of accuracy and robustness in causal inference from finite data sets. Our idea is to combine the inherent robustness of Bayesian approaches to causal structure discovery, such as GES, with the theoretical strength and clarity of constraint-based methods such as IC and PC/FCI. We obtain probability estimates on the input statements in a constraint-based procedure, which are then processed in decreasing order of reliability.

Interactions between real-world variables are often modeled in the form of a *causal DAG* $\mathcal{G}_C$. A directed path from A to B in $\mathcal{G}_C$ indicates a **causal relation** $A \Rightarrow B$ in the system. The *causal Markov* and *faithfulness* assumptions link the structure of the graph $\mathcal{G}_C$ to observed probabilistic in/dependencies, which forms the basis behind existing causal discovery procedures.

Method

We break up this inference process into a series of modular steps on basic **logical causal statements** of the form $L : (Z \Rightarrow X) \vee (Z \Rightarrow Y)$, and $L : Z \not\Rightarrow X$. Subsequent statements follow from deduction on the causal properties *transitivity* and *acyclicity*.

We obtain **probability estimates** on logical causal statements by summing the normalized posteriors of all structures $\mathcal{G}$ that entail L through d-separation:

$$p(L|\mathbf{D}) \propto \sum_{\mathcal{G} \in (L)} p(\mathbf{D}|\mathcal{G})p(\mathcal{G}).$$

Structures over different (small) *subsets* of variables $\mathbf{X} \subset \mathbf{V}$ can already suffice to derive a specific L. This is used in an efficient search strategy over increasing subsets of nodes, where it suffices to keep track of only the *maximum* probabilities obtained so far.

For the likelihood estimates $p(\mathbf{D}|\mathcal{G})$ on possible DAG structures we employ the well-known *Bayesian Dirichlet* (BD) metric.

We still need to account for the fact that the minimal DAG over subset $\mathbf{X} \subset \mathbf{V}$ may be **unfaithful** (uDAG) to the underlying structure. This leads to a modified inference rule, where d-separation remains valid, but the identifiable dependencies are restricted. From this we build a mapping from (possibly unfaithful) uDAGs $\mathcal{G}$ to *valid* logical causal statements $\mathcal{L}$.

Implementation and results

Tests show that a basic implementation of the resulting Bayesian Constraint-based Causal Discovery (BCCD) algorithm already outperforms established procedures such as FCI and Conservative PC. It can also indicate which causal decisions in the output have high reliability and which do not.

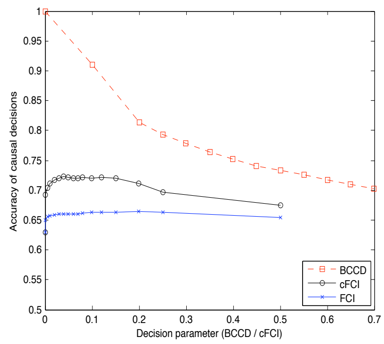

Figure 1. Tunable accuracy of causal decisions in BCCD

The approach is easily adapted into a powerful new independence test that actually *increases* in power for larger conditioning sets. Future extension include scoring MAGs, and allowing for continuous/mixed data.

Acknowledgments

This research was supported by NWO Vici grant nr.639.023.604.