Recursive Causal Discovery (Abstract Reprint)

Ehsan Mokhtarian[1], Sepehr Elahi[1], Sina Akbari[1], Negar Kiyavash[2]

[1]School of Computer and Communication Sciences, EPFL, 1015 Lausanne, Switzerland
[2]College of Management of Technology, EPFL, 1015 Lausanne, Switzerland

Abstract Reprint. This is an abstract reprint of the journal article by Mokhtarian, Elahi, Akbari, and Kiyavash (2025).

Abstract

Causal discovery from observational data, i.e., learning the causal graph from a finite set of samples from the joint distribution of the variables, is often the first step toward the identification and estimation of causal effects, a key requirement in numerous scientific domains. Causal discovery is hampered by two main challenges: limited data results in errors in statistical testing and the computational complexity of the learning task is daunting. This paper builds upon and extends four of our prior publications (Mokhtarian et al., 2021; Akbari et al., 2021; Mokhtarian et al., 2022, 2023a). These works introduced the concept of removable variables, which are the only variables that can be removed recursively for the purpose of causal discovery. Presence and identification of removable variables allow recursive approaches for causal discovery, a promising solution that helps to address the aforementioned challenges by reducing the problem size successively. This reduction not only minimizes conditioning sets in each conditional independence (CI) test, leading to fewer errors but also significantly decreases the number of required CI tests. The worst-case performances of these methods nearly match the lower bound. In this paper, we present a unified framework for the proposed algorithms, refined with additional details and enhancements for a coherent presentation. A comprehensive literature review is also included, comparing the computational complexity of our methods with existing approaches, showcasing their state-of-the-art efficiency. Another contribution of this paper is the release of RCD, a Python package that efficiently implements these algorithms. This package is designed for practitioners and researchers interested in applying these methods in practical scenarios. The package is available at github.com/ban-epfl/rcd, with comprehensive documentation provided at rcdpackage.com.

References

Mokhtarian, E.; Elahi, S.; Akbari, S.; and Kiyavash, N. 2025. Recursive Causal Discovery. *Journal of Machine Learning Research*, 26: 1–65.