Mach Learn
DOI 10.1007/s10994-013-5423-y

Least-squares independence regression for non-linear causal inference under non-Gaussian noise

Makoto Yamada · Masashi Sugiyama · Jun Sese

Received: 28 March 2011 / Accepted: 9 November 2013

Abstract The discovery of non-linear causal relationship under additive non-Gaussian noise models has attracted considerable attention recently because of their high flexibility. In this paper, we propose a novel causal inference algorithm called *least-squares independence regression* (LSIR). LSIR learns the additive noise model through the minimization of an estimator of the *squared-loss mutual information* between inputs and residuals. A notable advantage of LSIR is that tuning parameters such as the kernel width and the regularization parameter can be naturally optimized by cross-validation, allowing us to avoid overfitting in a data-dependent fashion. Through experiments with real-world datasets, we show that LSIR compares favorably with a state-of-the-art causal inference method.

Keywords Causal inference · Non-linear · Non-Gaussian · Squared-loss mutual information · Least-squares independence regression

1 Introduction

Learning *causality* from data is one of the important challenges in the artificial intelligence, statistics, and machine learning communities (Pearl 2000). A traditional method of learning causal relationship from observational data is based on the linear-dependence Gaussian-noise model (Geiger and Heckerman 1994). However, the linear-Gaussian assumption is too

Editor: Kristian Kersting.

M. Yamada (✉)
701 1st Ave., Sunnyvale, CA 94089, USA
e-mail: makotoy@yahoo-inc.com

M. Sugiyama · J. Sese
2-12-1, O-okayama, Meguro-ku, Tokyo 152-8552, Japan

M. Sugiyama
e-mail: sugi@cs.titech.ac.jp

J. Sese
e-mail: sesejun@cs.titech.ac.jp

Published online: 28 November 2013