Learning Hidden Causal Factors from Psychometrics Data Using Distributional Information

Roberto Legaspi
KDDI Research, Inc., Fujimino, Japan

Xinshuai Dong
Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

Donghuo Zeng
KDDI Research, Inc., Tokyo, Japan

Yuewen Sun
MBZUAI, Abu Dhabi, United Arab Emirates

Kazushi Ikeda
KDDI Research, Inc., Fujimino, Japan

Peter Spirtes
Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

Kun Zhang
Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

Abstract

Understanding latent variables and their causal mechanisms is central to psychological theory, yet most latent variable models in psychology have largely remained correlational. This work attempts to address three pivotal issues: identifying useful information from observational data that reveal latent causal factors, developing algorithms to leverage this distributional information, ensuring the identifiability of the recovered latent factors and their causal structure. We introduce a generalizable framework for discovering hidden causal structures from observed distributions in psychometric data. Applied to survey datasets on personality traits, teacher burnout, and multitasking behavior, our method uncovers hidden causal factors and their intricate interactions. Additionally, our findings offer an alternative perspective on psychometric scoring, grounded in the strength of the learned causal relations. These insights contribute to behavioral modeling and measurement and await further confirmatory studies to validate their implications for psychological science.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*