Causal Inference via Kernel Deviance Measures

Jovana Mitrovic, Dino Sejdinovic, Yee Whye Teh

JM, YWT @ University of Oxford and DeepMind
DS @ University of Oxford

Poster #9 Today 10:45 AM – 12:45 PM @ Room 210

Motivation

Many scientific questions are fundamentally causal in nature, e.g.

- genetic drivers of diseases
- motives behind customer's purchasing behaviour

For answering these questions, we need experimental data!

Unfortunately, often only observational data is available due to ethical, financial and technical reasons.

How do we infer causal relationships from observational data?

Causal Discovery from Observational Data

Prior work:

- Test for conditional independence
  $\rightarrow$ No definitive answer and not robust
- Assume particular functional relationship between variables and use asymmetry between cause and effect
  $\rightarrow$ Restrictive due to fixed functional dependence

Causal Discovery from Observational Data

Prior work:

- Test for conditional independence
  $\rightarrow$ No definitive answer and not robust
- Assume particular functional relationship between variables and use asymmetry between cause and effect
  $\rightarrow$ Restrictive due to fixed functional dependence

Our contribution:
Kernel Conditional Deviance for Causal Inference (KCDC)

- general-purpose, fully non-parametric
- provides definitive answer
- does not impose *a priori* any assumptions on the functional relationship between variables

KCDC TL;DR

- Novel interpretation of asymmetry between cause and effect

Minimal description length independence

If $X \rightarrow Y$, the minimal description length of the mechanism mapping X to Y is independent of the value of X, but the minimal description length of the mechanism mapping Y to X is dependent on the value of Y.

KCDC TL;DR

- Novel interpretation of asymmetry between cause and effect

Minimal description length independence

If $X \to Y$, the minimal description length of the mechanism mapping X to Y is independent of the value of X, but the minimal description length of the mechanism mapping Y to X is dependent on the value of Y.

- Flexible and robust asymmetry measure using kernel embeddings of conditional distributions
  - KCDC measure – measuring structural variability in the RKHS

$$S_{X \to Y} = \frac{1}{n} \sum_{i=1}^{n} \left(\left\| \mu_{Y|X=x_i} \right\|_{\mathcal{H}_\mathcal{Y}} - \frac{1}{n} \sum_{j=1}^{n} \left\| \mu_{Y|X=x_j} \right\|_{\mathcal{H}_\mathcal{Y}} \right)^2$$

KCDC TL;DR

- Causal discovery framework with three causal decision rules and confidence measure
  - Direct comparison of KCDC measures
  - Majority voting from an ensemble of direct comparisons
  - Using KCDC measures as features within a classifier

<u>Confidence measure:</u> $\mathcal{T}^{\mathcal{KCDC}} = \dfrac{|S_{X\to Y} - S_{Y\to X}|}{\min(S_{X\to Y}, S_{Y\to X})}$

KCDC TL;DR

- Causal discovery framework with three causal decision rules and confidence measure
  - Direct comparison of KCDC measures
  - Majority voting from an ensemble of direct comparisons
  - Using KCDC measures as features within a classifier

Confidence measure: $$\mathcal{T}^{\mathcal{KCDC}} = \frac{|S_{X \to Y} - S_{Y \to X}|}{\min(S_{X \to Y}, S_{Y \to X})}$$

- Extensive experimental evaluation against competing methods on simulated data and state-of-the-art on benchmark dataset Tübingen Cause-Effect Pairs

Poster #9 Today 10:45 AM – 12:45 PM @ Room 210