Invariant Ancestry Search

Phillip B. Mogensen

Nikolaj Thams

Jonas Peters

Copenhagen Causality Lab (CoCaLa),
University of Copenhagen

Setup

- Response variable Y, predictor variables $X_1, \ldots, X_d$, observed in different environments E (e.g., observational/interventional data).
- Existing idea: Invariant Causal Prediction (Peters et al., 2016)
  - A causal model is invariant across environments (formally: $Y \perp\!\!\!\perp E \mid X_S$).
  - Output intersection of all invariant models.
- S_{ICP} is a subset of PA_Y.
- Problems:
  - S_{ICP} is often empty.
  - S_{ICP} is not necessarily invariant.
- Goal: Improve on above shortcomings of ICP.
- Key insight: Minimal invariance.

Example

$$\begin{tikzcd}
 & X_1 \arrow[dr] & Y \arrow[dr] & \\
E \arrow[ur] \arrow[r] & X_2 \arrow[r] \arrow[rr, bend right] & X_3 \arrow[u] & X_4
\end{tikzcd}$$

Inv. sets: $\{1,2\}$, $\{3\}$, $\{1,2,3\}$, $\ldots$ $\Rightarrow S_{\mathrm{ICP}} = \emptyset$

$$\begin{tikzcd}
 & X_1 \arrow[r] & Y \arrow[dr] & \\
E \arrow[ur] \arrow[r] & X_2 \arrow[r] \arrow[rr, bend right] & X_3 \arrow[u] & X_4
\end{tikzcd}$$

Inv. sets: $\{1,2\}$, $\{1,3\}$, $\{1,2,3\}$, $\ldots$ $\Rightarrow S_{\mathrm{ICP}} = \{1\}$

Invariant Ancestry Search

- S min. inv. $:\Leftrightarrow$ S inv. and no $S' \subsetneq S$ inv.
- Min. inv. sets are ancestors of Y (Tian et al., 1998)
- IAS: Output union of min. inv. sets.

ICP ($S_{\text{ICP}} := \bigcap_{S:S \text{ inv.}} S$)

- $S_{\text{ICP}} \subseteq \text{PA}_Y$
- $\mathbb{P}(\hat{S}_{\text{ICP}} \subseteq \text{PA}_Y) \geq 1 - \alpha$.
- Can be empty.
- Not necessarily invariant.

IAS ($S_{\text{IAS}} := \bigcup_{S:S \text{ min. inv.}} S$)

- $S_{\text{IAS}} \subseteq \text{AN}_Y$
- $\lim_{n\to\infty} \mathbb{P}(\hat{S}_{\text{IAS}} \subseteq \text{AN}_Y) \geq 1 - \alpha$.
- Non-empty.
- Invariant.
- $S_{\text{ICP}} \subseteq S_{\text{IAS}}$ (Equal iff ICP invariant.)
- Retains guarantees if restricted to inv. sets up to size $m < d$.

Example

$$\begin{tikzcd} & X_1 \arrow[dr] & Y \arrow[dr] & \\ E \arrow[ur] \arrow[r] & X_2 \arrow[r] \arrow[rr, bend right] & X_3 \arrow[u] & X_4 \end{tikzcd}$$

Inv. sets: $\{1,2\}$, $\{3\}$, $\{1,2,3\}$, $\ldots$ $\Rightarrow S_{\text{ICP}} = \emptyset$

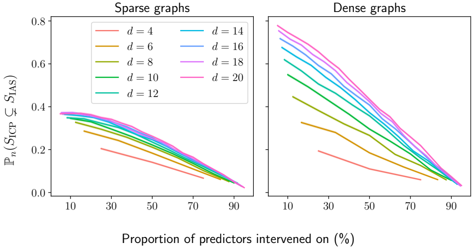
Min. inv. sets: $\{1,2\}$, $\{3\}$ $\Rightarrow S_{\text{IAS}} = \{1,2,3\}$

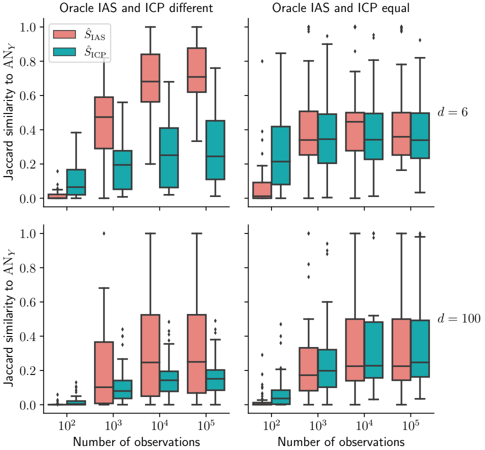
$$\begin{tikzcd} & X_1 \arrow[r] & Y \arrow[dr] & \\ E \arrow[ur] \arrow[r] & X_2 \arrow[r] \arrow[rr, bend right] & X_3 \arrow[u] & X_4 \end{tikzcd}$$

Inv. sets: $\{1,2\}$, $\{1,3\}$, $\{1,2,3\}$, $\ldots$ $\Rightarrow S_{\text{ICP}} = \{1\}$

Min. inv. sets: $\{1,2\}$, $\{1,3\}$ $\Rightarrow S_{\text{IAS}} = \{1,2,3\}$

Numerical experiments





- IAS often strictly larger than ICP.
- IAS finds large sets of ancestors in finite samples

Summary

- IAS finds non-empty, invariant sets of predictors that are ancestors of Y.
- Numerical experiments: $S_{\text{ICP}} \subseteq S_{\text{IAS}}$ often strict; IAS finds subsets of AN_Y in finite samples with high probability.

For more of our work, visit our website:

`https://cocala.github.io/`

Thanks for listening!

Phillip B. Mogensen
<pbm@math.ku.dk>

Nikolaj Thams
<thams@math.ku.dk>

Jonas Peters
<jonas.peters@math.ku.dk>

References I

Peters, Jonas, Peter Bühlmann, and Nicolai Meinshausen. "Causal inference by using invariant prediction: identification and confidence intervals." Journal of the Royal Statistical Society. Series B (Statistical Methodology) (2016): 947-1012.

Tian, Jin, Azaria Paz, and Judea Pearl. Finding minimal d-separators. Computer Science Department, University of California, 1998.

CoCaLab Causality Copenhagen