Intervention Target Estimation in the Presence of Latent Variables

Burak Varıcı Ali Tajer Karthikeyan Shanmugam Prasanna Sattigeri

Rensselaer Polytechnic Institute IBM Research AI

Rensselaer IBM

uai2022

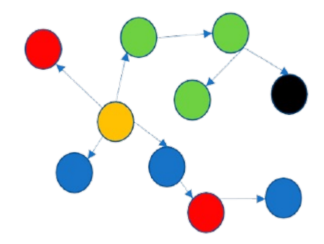
Motivation

- Directed Acyclic Graphs (DAG): encode cause-effect relationships
- **Causally insufficient systems**: unobserved confounders
- **Maximal Ancestral Graphs (MAG):** ancestral and confounding relationships
- **Interventions**: forced changes on target variables

Motivation and Applications

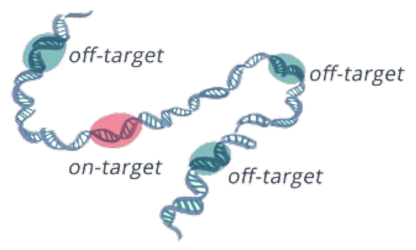
Fault localization:

- Causal model: set of loosely coupled services
- Vulnerable to unwanted changes, e.g., delays, attacks
- Interventions: root causes of the faulty operations



Biological applications:

- Causal model: gene regulatory networks
- Off-target genome sites can also be affected



Courtesy of SJTU-BioX

How it helps interventional structure learning?

Interventional Structure Learning:

- Causally insufficient models are common yet understudied
- Known targets → strong assumption
- Unknown targets → not scalable methods
- Combine with scalable observational algorithms!

observational algorithm + **infer interventional knowledge** = **full information in scalable fashion**

Model

- **Linear Structural Equation Model**: $X = [X_1, \dots, X_p]^\top$ and $\epsilon \sim (N, \Omega)$

$$X = B^\top X + \epsilon$$

- Precision matrix: $\Theta = (I - B)\Omega^{-1}(I - B)^\top$
- Models: $B^{(s)}, \epsilon^{(s)}, \Sigma^{(s)}, \Theta^{(s)}, \Omega^{(s)}$ for $s \in \{1, 2\}$
- **Soft Interventions:** Change in noise variations of targets.

$$\mathcal{I} \triangleq \{i : \sigma_i^{(1)} \neq \sigma_i^{(2)}\}$$

- Causally insufficient systems: $\mathcal{I}$ is not exactly identifiable!

Augmented graph and interventional MAG

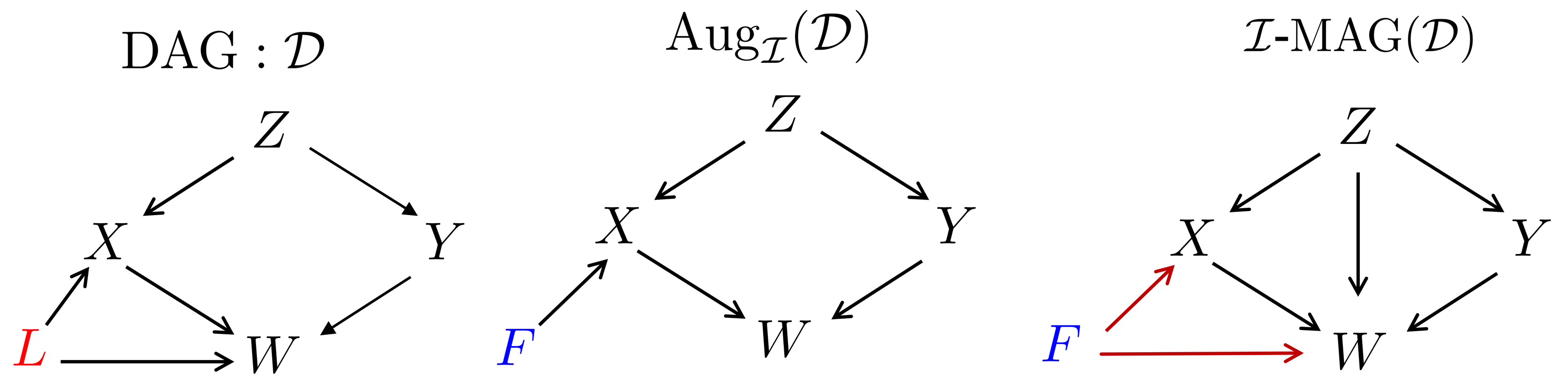
- $\mathrm{Aug}_{\mathcal{I}}$: Create a node ($F$) for each pair of settings, draw edges to targets (Jaber et al., NeurIPS'20)
- $\mathcal{I}$-MAG($\mathcal{D}$): MAG of $\mathrm{Aug}_{\mathcal{I}}$, with edges $\mathcal{E}_{\mathcal{I}}$
- $\mathrm{Aug}_{\mathcal{I}}(\mathcal{D})$ and $\mathcal{I}$-MAG($\mathcal{D}$) exactly represents the separation statements.
- Effective intervention targets:

$$\mathcal{K} \triangleq \{i : (F, i) \in \mathcal{E}_I\}$$

- "Parents-or-spouses" of effective interventions: $\mathrm{ps}(\mathcal{K})$

Objective

Estimate $\mathcal{K}$ and $\mathrm{ps}(\mathcal{K})$ from $\Sigma^{(1)}$ and $\Sigma^{(2)}$



$\mathcal{I} = \{\emptyset, \{X\}\}$, $(F, X), (F, W) \in \mathcal{E}_{\mathcal{I}}$

$\mathcal{K} = \{X, W\}$

$\mathrm{ps}(\mathcal{K}) = \{\{Z\}, \{X, Y, Z\}\}$

Results

- Marginal SEMs: $X_S \to (B_S, \epsilon_S)$
- Precision Difference Estimation (PDE) $\Delta_S = \Theta_S^{(1)} - \Theta_S^{(2)}$.
- Lasso formulation and solution through ADMM (Jiang et al. JMLR 2018).

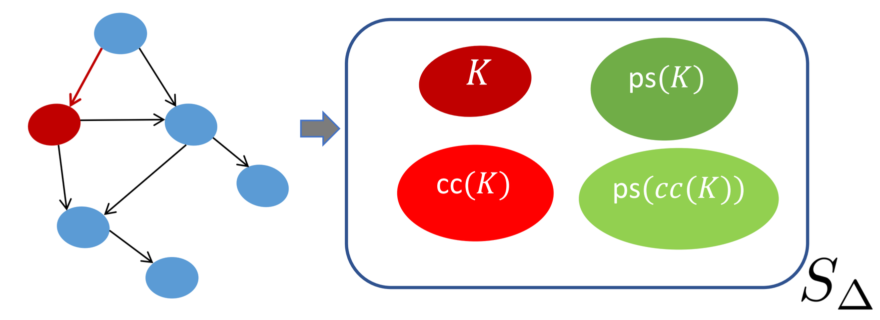
$$\hat{\Delta}_S = \min_{\Delta_S} \left\{ \frac{1}{2} \mathrm{Tr}(\Delta_S^\top \hat{\Sigma}^{(1)} \Delta_S \hat{\Sigma}^{(2)}) - \mathrm{Tr}(\Delta_S(\hat{\Sigma}^{(1)} - \hat{\Sigma}^{(2)})) + \lambda \|\Delta_S\|_1 \right\}$$

PreDITEr
- Intervened $K \in \mathcal{K}$: ϵ_S is never invariant
- Non-intervened J : ϵ_S can be made invariant
- **Goal** : search for an S to make $[\Delta_S]_{J,J} = 0$
- **Steps** : prune, then check subsets of the pruning set

Theorem 1: Consider an observed node $V \in \mathbf{V}$. Then,

$$V \in \mathcal{K} \iff \nexists S \subseteq \mathbf{V} \text{ such that } [\Delta_S]_{V,V} = 0$$

Pruning set



$$\Delta = \Theta_{\mathbf{V}}^{(1)} - \Theta_{\mathbf{V}}^{(2)}$$

$$S_\Delta = \{V \in \mathbf{V} : [\Delta]_{V,V} \neq 0\}$$

(contains $\mathcal{K}, \mathrm{ps}(\mathcal{K})$)

Results for restricting search to S_Δ

- Consider a node $V \in S_\Delta \setminus \mathcal{K}$.

$$S = S_\Delta \cap \mathrm{an}(V) \to [\Delta_S]_{V,V} = 0$$

- Consider $K \in \mathcal{K}$ and $J \in \mathbf{V} \setminus \mathcal{K}$. Then,

$$J \in \mathrm{ps}(K) \iff \nexists S \subseteq S_\Delta \text{ such that } [\Delta_S]_{K,J} = 0$$

Theorem 2 (Main result): Given the true covariances $\Sigma^{(1)}$ and $\sigma^{(2)}$, the algorithm **perfectly** estimates
- Effective interventions $\mathcal{K}$
- Parents-or-spouses of them $\mathrm{ps}(\mathcal{K})$

- Knowing $\mathcal{K}$ and $\mathrm{ps}(\mathcal{K})$: all the interventional knowledge
- Markov equivalence class of $\mathcal{I}$-MAG's: ψ-PAG

Theorem 3 (Secondary result): Given the PAG of a MAG $\mathcal{M}$, the algorithm **perfectly** recovers the ψ-PAG.

Experiments

- **Synthetic data:** Erdös-Renyi random DAGs with graph size p
- $|\mathbf{L}| = 5$ latents, $|\mathcal{K}| = 5$ targets
- Compare to FCI-JCI123 (Mooij et al. JMLR'2020) at 5000 samples

Method	PreDITEr	FCI-JCI123	PreDITEr	FCI-JCI123
Graph size	20	20	40	40
Precision	1.0	1.0	1.0	0.96
Recall	0.83	1.0	0.87	0.96
Runtime(s)	< 1	80.9	< 1	1301.9

- **Real data**: Protein signaling network (Sachs et al. 2005).
- PreDITEr recovers most of the skeleton correctly (details in paper).

Conclusions

- **Intervention target estimation** in linear models with latent nodes.
- A **consistent** and **scalable** algorithm using precision differences.
- Furthermore, refined the observational PAG to interventional PAG.

Corresponding author: Burak Varıcı (varicb@rpi.edu)