Workshop Talk on Approximate Inference of Influences by Bounds on Confounding

Philipp Geiger, Dominik Janzing and Bernhard Schölkopf

Max Planck Institute for Intelligent Systems, Tübingen, Germany

July 27, 2014
Outline

1. Motivating example
2. General problem description
3. Outline of our approach
4. Part 1 of our approach: Formalizing dependence versus effect
5. Part 2 of our approach: Prototypical application scenarios
6. Conclusion
Motivating example

Imagine you observe

Whenever a patient takes a certain drug, he/she recovers the same day.

Intuitive conclusion

Drug is effective because in this scenario no confounder can be that strong (deterministic).

Later we will see ...

... how this intuition can be formalized.
General problem description

Given

- The joint distribution $P(X, Y)$, induced by a causal model with potential hidden confounder $U$ and causal DAG as in Fig. 1.
- Some form of additional knowledge which implies “bounds on confounding”.

Goal

Estimate post-interventional distribution $P(Y | \text{do } X=x)$ [Pearl, 2000] or related quantities such as the effect of treatment on the treated (ETT) [Pearl, 2000] or the causal strength from $X$ to $Y$, denoted by $\mathcal{C}_{X \rightarrow Y}$ [Janzing et al., 2013].
Outline of our approach

Part 1: Formalizing dependence versus effect

Our approach consists of two parts.

In **part 1**, we propose various **formalizations** of the following intuitive notions:

- **Observed dependence**: the statistical dependence between $X$ and $Y$ (i.e. based on $P(X, Y)$).
- **Back-door dependence**: the “spurious association” [Pearl, 2000] between $X$ and $Y$ due to $U$.
- **Causal effect**: what happens to $Y$ upon intervening on $X$. 
Outline of our approach

Part 1: Formalizing dependence versus effect (ctd.)

For all formalizations one can derive inequalities of the form:

\[
\begin{bmatrix}
\text{back-door dependence} \\
\end{bmatrix} \geq d\left(\begin{bmatrix}
\text{observed dependence} \\
\end{bmatrix}, \begin{bmatrix}
\text{causal effect} \\
\end{bmatrix}\right).
\]

\(d(\cdot, \cdot)\) either stands for “\(\cdot - \cdot\)” (the usual difference) or “\(D[\cdot \| \cdot]\)” (the KL divergence).

In the former case we can also write:

\[
\begin{bmatrix}
\text{causal effect} \\
\end{bmatrix} \geq \begin{bmatrix}
\text{observed dependence} \\
\end{bmatrix} - \begin{bmatrix}
\text{back-door dependence} \\
\end{bmatrix}.
\]

NB: For some formalizations we can even prove equalities, instead of inequalities.
Outline of our approach

Part 2: Prototypical application scenarios

For the inequalities from part 1 to be applicable, we need bounds on the back-door dependence.

Therefore in part 2 we discuss several prototypical scenarios which imply such bounds, and apply part 1.

Remark: how to find such scenarios?

- **Main difficulty:** Find conditions which are *weak enough* to be met in *realistic scenarios*, but *strong enough* to imply bounds on back-door dependence.

- **General intuition:** strength of back-door dependence = deviation from randomized experiment.
Part 1: Formalizing dependence versus effect

Estimating the causal effect \( p(Y|\text{do } X=x) \)

Here we formalize the intuitive notions by:

- observed dependence: \( p(Y|X=x) \),
- back-door dependence: \( \min \{ \mathcal{C}_{U \rightarrow X}, \mathcal{C}_{U \rightarrow Y} \} \), \( I(X:U) \),
- causal effect: \( p(Y|\text{do } X=x) \).

Theorem

\[
D[ p(Y|X) \parallel p(Y|\text{do } X) ] \leq \min \{ \mathcal{C}_{U \rightarrow X}, \mathcal{C}_{U \rightarrow Y} \} \leq I(X:U).
\]
Part 1: Formalizing dependence versus effect

Estimating the causal strength $C_{X \rightarrow Y}$

Here we formalize the intuitive notions by:

- observed dependence: $I(X : Y)$,
- back-door dependence: $C_{U \rightarrow X}$,
- causal effect: $C_{X \rightarrow Y}$.

Theorem

$$C_{U \rightarrow X} \geq I(X : Y) - C_{X \rightarrow Y}.$$
Part 1: Formalizing dependence versus effect

Estimating the differential ETT from $X$ to $Y$

Here we formalize the intuitive notions by:

- **observed dependence:** $d_x \mathbb{E}[Y|X=x],$
- **back-door dependence:**
  $\partial_2 \mathbb{E}[Y_{do\;X=x}|X=x]$  
  ($\partial_i$ is the partial derivative),
- **causal effect:** $\partial_1 \mathbb{E}[Y_{do\;X=x}|X=x],$

**NB:** this is a differential version of the effect of treatment on the treated (ETT),

$$\mathbb{E}[Y_{do\;X=x'|X=x'}] - \mathbb{E}[Y_{do\;X=x}|X=x'].$$

**Theorem**

For all $x$ we have

$$\partial_2 \mathbb{E}[Y_{do\;X=x}|X=x] = d_x \mathbb{E}[Y|X=x] - \partial_1 \mathbb{E}[Y_{do\;X=x}|X=x].$$
Formalizing the drug efficacy scenario

- Let $X$ be weekday of intake, $Y$ be weekday of recovery.
- Recall the observation: $p(y|x) = \delta_{yx}$ (Kronecker delta).
- Recall the intuition: $U$ cannot fully determine $X$.
  $\Rightarrow I(U : X) < H(X)$ (bound on back-door dependence).

Results from part 1 imply

$\exists x, x'$ such that $p(Y|\text{do } X=x) \neq p(Y|\text{do } X=x')$.
$\Rightarrow$ Drug is effective.
Part 2: Prototypical application scenarios

Estimating the effect of advertisement letters

Scenario: mail order company, efficacy of ad. letters

▷ $X$: advertisement letter was sent out to a person.
▷ $Y$: ordering behavior of this person afterwards.
▷ $W$: letter was sent out based on guidelines $U$ or randomly. We do not observe $W$, but know $p(W)$.

$X := \begin{cases} f(U, N_X), & \text{if } W = 1, \\ f'(N_X), & \text{if } W = 0. \end{cases}$

Results from part 1 imply

$$\mathcal{C}_{X \rightarrow Y} \geq I(X:Y) - H(X)p(W=1).$$
Part 2: Prototypical application scenarios

**Variant of the regression discontinuity design (RDD)**

**NB:** RDD [Thistlewaite et al., 1960] uses discontinuity in $Z \rightarrow X$.

**Scenario: “differential version” of the RDD**

- Observe $Z, X, Y$, know causal structure.
- $X := f(Z)$ differentiable, invertible.
  
  $g := f^{-1}$.

**Results from part 1 imply**

\[
\partial_2 \mathbb{E}[Y_{do\,X=x}|Z=g(x)] \quad g'(x) = d_x \mathbb{E}[Y|X=x] - \partial_1 \mathbb{E}[Y_{do\,X=x}|X=x].
\]

**Therefore**

If $|f'(x_0)|$ big ($|g'(f(x_0))|$ small) and (*) bounded,

then the above equation helps to estimate the causal effect.
Conclusion

- Presented various formalizations of the intuitive notions of observed dependence, back-door dependence and causal effect.
- Always: back-door dependence bounds deviation between observed dependence and causal effect.
- Discussed prototypical scenarios which allow to bound the back-door dependence and thus to apply the above bound.
- **Open problem:** finding real applications.
Paper