Causal reasoning with cyclic graphs

Simon Ferreira simon.ferreira@sorbonne-universite.fr

L'Institut Pierre Louis d'Epidémiologie et de Santé Publique, Inserm, Sorbonne Université

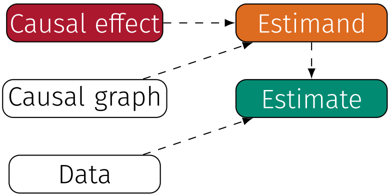
Nov, 2025

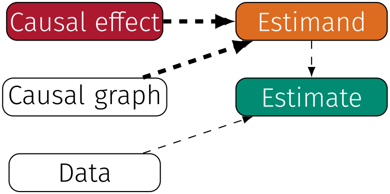
SUMMARY

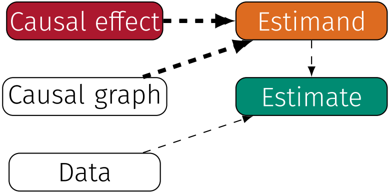
1

The SCM framework

Causal inference and identifiability




Causal effect
Estimand
Causal graph
Estimate
Data



The causal effect is said to be identifiable if it is uniquely computable from $P(\mathbb{V})$.

Total effect a.k.a Average Treatment Effect (ATE)

Total effect

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x}'))$$

$$\sim P(\mathbb{y} \mid do(\mathbb{x}))$$

Where $do(\cdot)$ is an operator representing an intervention.

Pearl, "Causal diagrams for empirical research". Biometrika, 1995

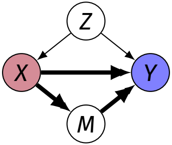
Total effect a.k.a Average Treatment Effect (ATE)

Total effect

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x}'))$$

$$\sim P(\mathbb{y} \mid do(\mathbb{x}))$$

Where $do(\cdot)$ is an operator representing an intervention.



$P(\mathbb{y} \mid do(\mathbb{x}))$ is identifiable if it is uniquely computable from a **positive observationnal** distribution $P(\mathbb{V})$.

Pearl, "Causal diagrams for empirical research". Biometrika, 1995

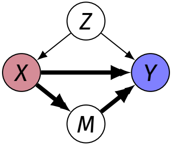
Total effect a.k.a Average Treatment Effect (ATE)

Total effect

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{X} = \mathbb{x}'))$$

$$\sim P(\mathbb{y} \mid do(\mathbb{x}))$$

Where $do(\cdot)$ is an operator representing an intervention.



$P(\mathbb{y} \mid do(\mathbb{x}))$ is identifiable if it is uniquely computable from a **positive observationnal** distribution $P(\mathbb{V})$.

The **do-calculus** is a **sound and complete** for identification.

Pearl, "Causal diagrams for empirical research". Biometrika, 1995

Structural Causal Model

$$
\begin{aligned}
\forall x,\ & \xi_x \leftarrow \mathcal{D}_x \\
A &:= f_A(\xi_a, \xi_{ab}) \\
G &:= f_G(\xi_g) \\
H &:= f_H(G, \xi_h) \\
I &:= f_I(G, \xi_i) \\
B &:= f_B(A, H, \xi_b, \xi_{ab}) \\
C &:= f_C(A, B, I, \xi_c) \\
F &:= f_F(C, G, \xi_f) \\
D &:= f_D(C, F, \xi_d) \\
E &:= f_E(B, D, G, \xi_e)
\end{aligned}
$$

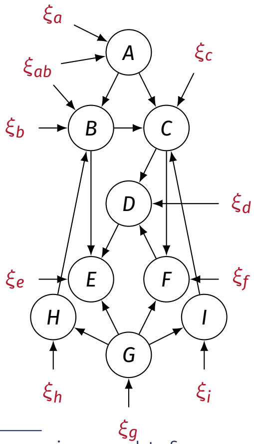
Pearl, Causality: Models, Reasoning, and Inference. Cambridge University Press, 2009

Causal graphs

Structural Causal Model

Directed Acyclic Graph (DAG)

$$\forall x,\ \xi_x \leftarrow \mathcal{D}_x$$
$$A := f_A(\xi_a, \xi_{ab})$$
$$G := f_G(\xi_g)$$
$$H := f_H(G, \xi_h)$$
$$I := f_I(G, \xi_i)$$
$$B := f_B(A, H, \xi_b, \xi_{ab})$$
$$C := f_C(A, B, I, \xi_c)$$
$$F := f_F(C, G, \xi_f)$$
$$D := f_D(C, F, \xi_d)$$
$$E := f_E(B, D, G, \xi_e)$$



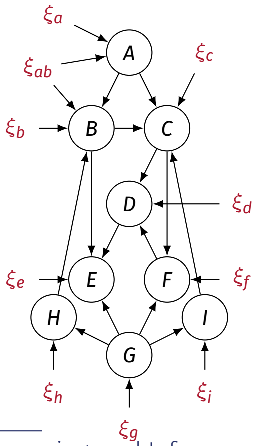
Pearl, Causality: Models, Reasoning, and Inference. Cambridge University Press, 2009

Causal graphs

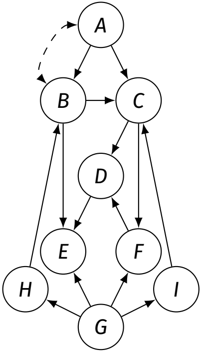
Structural Causal Model

$$\forall x,\ \xi_x \leftarrow \mathcal{D}_x$$
$$A := f_A(\xi_a, \xi_{ab})$$
$$G := f_G(\xi_g)$$
$$H := f_H(G, \xi_h)$$
$$I := f_I(G, \xi_i)$$
$$B := f_B(A, H, \xi_b, \xi_{ab})$$
$$C := f_C(A, B, I, \xi_c)$$
$$F := f_F(C, G, \xi_f)$$
$$D := f_D(C, F, \xi_d)$$
$$E := f_E(B, D, G, \xi_e)$$

Directed Acyclic Graph (DAG)



Acyclic Directed Mixed Graph (ADMG)



Pearl, Causality: Models, Reasoning, and Inference. Cambridge University Press, 2009

Causal graphs

Structural Causal Model

$\forall x,\ \xi_x \leftarrow \mathcal{D}_x$

$A := f_A(\xi_a, \xi_{ab})$

$G := f_G(\xi_g)$

$H := f_H(G, \xi_h)$

$I := f_I(G, \xi_i)$

$B := b$

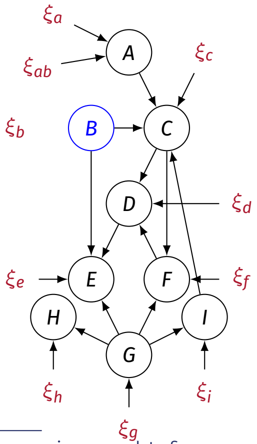
$C := f_C(A, B, I, \xi_c)$

$F := f_F(C, G, \xi_f)$

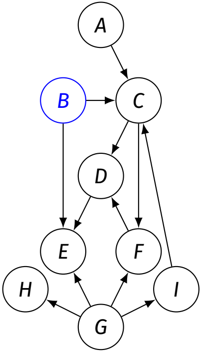
$D := f_D(C, F, \xi_d)$

$E := f_E(B, D, G, \xi_e)$

Directed Acyclic Graph (DAG)



Acyclic Directed Mixed Graph (ADMG)



Pearl, Causality: Models, Reasoning, and Inference. Cambridge University Press, 2009

D-SEPARATION FOR ADMGS

A path is said to be blocked by a set of vertices $\mathbb{Z} \subset \mathbb{V}$ if:

- it contains a chain $\langle A {*\!\!\rightarrow} B \rightarrow C \rangle$ or $\langle A \leftarrow B {\leftarrow\!\!*} C \rangle$ or a fork $\langle A \leftarrow B \rightarrow C \rangle$ and $B \in \mathbb{Z}$; or
- it contains a collider $\langle A {*\!\!\rightarrow} B {\leftarrow\!\!*} C \rangle$ such that no descendant of B is in $\mathbb{Z}$.

Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc. 1988

D-SEPARATION FOR ADMGS

A path is said to be blocked by a set of vertices $\mathbb{Z} \subset \mathbb{V}$ if:

- it contains a chain $\langle A {*}\!\!\rightarrow B \rightarrow C \rangle$ or $\langle A \leftarrow B \leftarrow\!\!{*} C \rangle$ or a fork $\langle A \leftarrow B \rightarrow C \rangle$ and $B \in \mathbb{Z}$; or
- it contains a collider $\langle A {*}\!\!\rightarrow B \leftarrow\!\!{*} C \rangle$ such that no descendant of B is in $\mathbb{Z}$.

$\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between $\mathbb{X}$ and $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $(\mathbb{X} \perp\!\!\!\perp_d \mathbb{Y} \mid \mathbb{Z})_{\mathcal{G}}$.

Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc. 1988

D-SEPARATION FOR ADMGs

A path is said to be blocked by a set of vertices $\mathbb{Z} \subset \mathbb{V}$ if:

- it contains a chain $\langle A {*\!\!\rightarrow} B \rightarrow C \rangle$ or $\langle A \leftarrow B {\leftarrow\!\!*} C \rangle$ or a fork $\langle A \leftarrow B \rightarrow C \rangle$ and $B \in \mathbb{Z}$; or
- it contains a collider $\langle A {*\!\!\rightarrow} B {\leftarrow\!\!*} C \rangle$ such that no descendant of B is in $\mathbb{Z}$.

$\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between $\mathbb{X}$ and $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $(\mathbb{X} \perp\!\!\!\perp_d \mathbb{Y} \mid \mathbb{Z})_{\mathcal{G}}$.

Theorem

$(\mathbb{X} \perp\!\!\!\perp_d \mathbb{Y} \mid \mathbb{Z})_{\mathcal{G}} \Rightarrow \mathbb{X} \perp\!\!\!\perp_{\mathrm{Pr}} \mathbb{Y} \mid \mathbb{Z}$

Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc. 1988

DO-CALCULUS FOR ADMGS

The do-calculus consists of three rules:

Rule 1 $P(y|do(x), z, w) = P(y|do(x), w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(y|do(x, z), w) = P(y|do(x), z, w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(y|do(x, z), w) = P(y|do(x), w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\overline{\mathbb{Z}(\mathbb{W})}}}$

Pearl, "Causal diagrams for empirical research". Biometrika, 1995

DO-CALCULUS FOR ADMGs

The do-calculus consists of three rules:

Rule 1 $P(y|do(x), z, w) = P(y|do(x), w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(y|do(x, z), w) = P(y|do(x), z, w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(y|do(x, z), w) = P(y|do(x), w)$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}\mathbb{Z}(\mathbb{W})}}}$

Theorem

$P(y \mid do(x))$ is identifiable if and only if there exists a finite sequence of transformations, each conforming to either one of the Rules 1-3 or some standard probability manipulations, that reduces $P(y \mid do(x))$ into a do-free formula.

Pearl, "Causal diagrams for empirical research". Biometrika, 1995

The do-calculus consists of three rules:

Rule 1 $P(\mathbb{y}|do(\mathbb{x}),\mathbb{z},\mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}),\mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X},\mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(\mathbb{y}|do(\mathbb{x},\mathbb{z}),\mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}),\mathbb{z},\mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X},\mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(\mathbb{y}|do(\mathbb{x},\mathbb{z}),\mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}),\mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_d \mathbb{Z} \mid \mathbb{X},\mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}\mathbb{Z}(\mathbb{W})}}}$

Theorem

$P(\mathbb{y} \mid do(\mathbb{x}))$ is identifiable if and only if there exists a finite sequence of transformations, each conforming to either one of the Rules 1-3 or some standard probability manipulations, that reduces $P(\mathbb{y} \mid do(\mathbb{x}))$ into a do-free formula.

The do-calculus is sound and complete!

Pearl, “Causal diagrams for empirical research”. Biometrika, 1995

Back-Door Criterion

Definition

A set of variables $\mathbb{Z}$ satisfies the back-door criterion relative to (X, Y) if:

- $\mathbb{Z} \cap \boldsymbol{De}(X, \mathcal{G}) = \emptyset$, and
- $\mathbb{Z}$ blocks every back-door path (*i.e.*, $\langle X \leftarrow \cdots Y \rangle$).

Back-Door Criterion

Definition

A set of variables $\mathbb{Z}$ satisfies the back-door criterion relative to (X, Y) if:

- $\mathbb{Z} \cap \boldsymbol{De}(X, \mathcal{G}) = \emptyset$, and
- $\mathbb{Z}$ blocks every back-door path (*i.e.*, $\langle X \leftarrow \cdots Y \rangle$).

Theorem

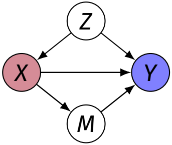
If $\mathbb{Z}$ satisfies the back-door criterion relative to (X, Y) then:

$$\Pr(y \mid do(x)) = \sum_{\mathbb{z}} \Pr(y \mid x, \mathbb{z}) \Pr(\mathbb{z})$$

Definition

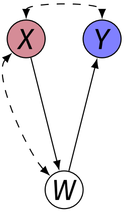
A set of variables $\mathbb{Z}$ satisfies the back-door criterion relative to (X, Y) if:

- $\mathbb{Z} \cap \boldsymbol{De}(X, \mathcal{G}) = \emptyset$, and
- $\mathbb{Z}$ blocks every back-door path (*i.e.*, $\langle X \leftarrow \cdots Y \rangle$).



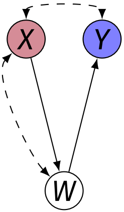
$$
\begin{aligned}
P(y \mid do(x)) \\
&= \sum_z P(y \mid do(x), z) P(z \mid do(x)) \\
&= \sum_z P(y \mid x, z) P(z \mid do(x)) \quad \text{(Rule 2)} \\
&= \sum_z P(y \mid x, z) P(z) \quad \text{(Rule 3)}
\end{aligned}
$$

The do-calculus can time be consuming



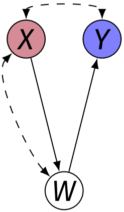
The do-calculus can time be consuming

$P(\mathbb{y} \mid do(\mathbb{x}))$

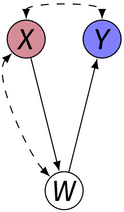


THE DO-CALCULUS CAN TIME BE CONSUMING

$$P(\mathbb{y} \mid do(\mathbb{x})) = \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$



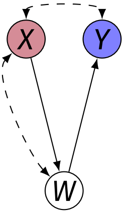
The do-calculus can time be consuming



$$P(\mathbb{y} \mid do(\mathbb{x})) = \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$

$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid do(\mathbb{x})) \quad \text{(Rule 2)}$$

The do-calculus can time be consuming

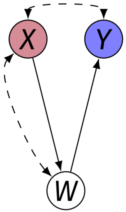


$$P(\mathbb{y} \mid do(\mathbb{x})) = \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$

$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid do(\mathbb{x})) \quad \text{(Rule 2)}$$

$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{w})) P(\mathbb{w} \mid do(\mathbb{x})) \quad \text{(Rule 3)}$$

The do-calculus can time be consuming



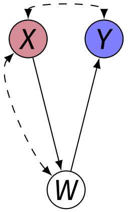
$$P(y \mid do(x)) = \sum_{w} P(y \mid do(x), w) P(w \mid do(x))$$

$$= \sum_{w} P(y \mid do(x), do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

$$= \sum_{w} P(y \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 3)}$$

$$= \sum_{w,x'} P(y \mid do(w), x') P(x' \mid do(w)) P(w \mid do(x))$$

The do-calculus can time be consuming



$$P(y \mid do(x)) = \sum_{w} P(y \mid do(x), w) P(w \mid do(x))$$

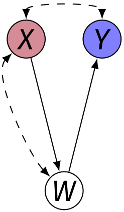
$$= \sum_{w} P(y \mid do(x), do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

$$= \sum_{w} P(y \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 3)}$$

$$= \sum_{w,x'} P(y \mid do(w), x') P(x' \mid do(w)) P(w \mid do(x))$$

$$= \sum_{w,x'} P(y \mid w, x') P(x' \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

THE DO-CALCULUS CAN TIME BE CONSUMING



$$P(y \mid do(x)) = \sum_{w} P(y \mid do(x), w) P(w \mid do(x))$$

$$= \sum_{w} P(y \mid do(x), do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

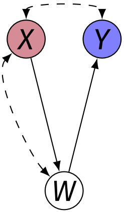
$$= \sum_{w} P(y \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 3)}$$

$$= \sum_{w,x'} P(y \mid do(w), x') P(x' \mid do(w)) P(w \mid do(x))$$

$$= \sum_{w,x'} P(y \mid w, x') P(x' \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

$$= \sum_{w,x'} P(y \mid w, x') P(x') P(w \mid do(x)) \quad \text{(Rule 3)}$$

The do-calculus can time be consuming



$$P(y \mid do(x)) = \sum_w P(y \mid do(x), w) P(w \mid do(x))$$

$$= \sum_w P(y \mid do(x), do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

$$= \sum_w P(y \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 3)}$$

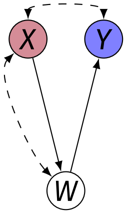
$$= \sum_{w,x'} P(y \mid do(w), x') P(x' \mid do(w)) P(w \mid do(x))$$

$$= \sum_{w,x'} P(y \mid w, x') P(x' \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 2)}$$

$$= \sum_{w,x'} P(y \mid w, x') P(x') P(w \mid do(x)) \quad \text{(Rule 3)}$$

$$= \dots$$

The do-calculus can time be consuming



$$
\begin{aligned}
P(y \mid do(x)) &= \sum_{w} P(y \mid do(x), w) P(w \mid do(x)) \\
&= \sum_{w} P(y \mid do(x), do(w)) P(w \mid do(x)) \quad \text{(Rule 2)} \\
&= \sum_{w} P(y \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 3)} \\
&= \sum_{w,x'} P(y \mid do(w), x') P(x' \mid do(w)) P(w \mid do(x)) \\
&= \sum_{w,x'} P(y \mid w, x') P(x' \mid do(w)) P(w \mid do(x)) \quad \text{(Rule 2)} \\
&= \sum_{w,x'} P(y \mid w, x') P(x') P(w \mid do(x)) \quad \text{(Rule 3)} \\
&= \ldots
\end{aligned}
$$

In this specific case, the total effect is not identifiable

$\Longrightarrow$ We can never find a do-free formula!

Hedge criterion for ADMGs

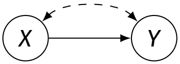
A hedge for $P(\mathbb{y} \mid do(\mathbb{x}))$ is a subgraph with specific graphical constraints related to $\mathbb{X}$ and $\mathbb{Y}$.

Shpitser and Pearl, "Complete identification methods for the causal hierarchy". JMLR, 2008

Hedge criterion for ADMGs

A hedge for $P(\mathbb{y} \mid do(\mathbb{x}))$ is a subgraph with specific graphical constraints related to $\mathbb{X}$ and $\mathbb{Y}$.

- no bidirected dashed edges $\implies$ no hedges
- This graph is a hedge:

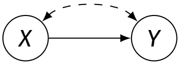


Shpitser and Pearl, "Complete identification methods for the causal hierarchy". JMLR, 2008

Hedge criterion for ADMGs

A hedge for $P(\mathbb{y} \mid do(\mathbb{x}))$ is a subgraph with specific graphical constraints related to $\mathbb{X}$ and $\mathbb{Y}$.

- no bidirected dashed edges $\Longrightarrow$ no hedges
- This graph is a hedge:



Theorem

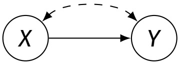
$P(\mathbb{y} \mid do(\mathbb{x}))$ is not identifiable in $\mathcal{G}$ <u>if and only if</u> there is a hedge for the ordered pair $(\mathbb{X}, \mathbb{Y})$ in $\mathcal{G}$.

Shpitser and Pearl, "Complete identification methods for the causal hierarchy". JMLR, 2008

Hedge criterion for ADMGs

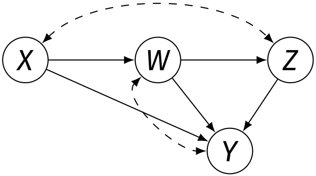
A hedge for $P(\mathbb{y} \mid do(\mathbb{x}))$ is a subgraph with specific graphical constraints related to $\mathbb{X}$ and $\mathbb{Y}$.

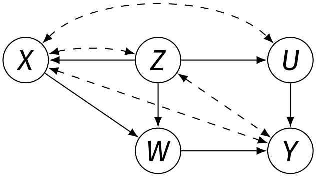
- no bidirected dashed edges $\Longrightarrow$ no hedges
- This graph is a hedge:



Theorem

$P(\mathbb{y} \mid do(\mathbb{x}))$ is not identifiable in $\mathcal{G}$ if and only if there is a hedge for the ordered pair $(\mathbb{X}, \mathbb{Y})$ in $\mathcal{G}$.





Shpitser and Pearl, "Complete identification methods for the causal hierarchy". JMLR, 2008

Hedge criterion for ADMGs: definition

C-component: Set of vertices which are all connected via bidirected arrows.

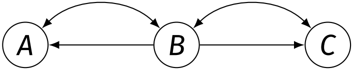


Figure: C-component

Forest: Acyclic graph in which every edge has at most one children.
We call roots the nodes with no children.

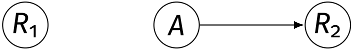


Figure: Forest

Hedge for $(\mathbb{X}, \mathbb{Y})$: Pair of $\mathbb{R}$-rooted C-forests $(\mathcal{F}, \mathcal{F}')$ in the graph such that:

- $\mathbb{R} \subseteq An(\mathbb{Y}, \mathcal{G} \backslash \{V\})$
- $\mathcal{F} \subseteq \mathcal{F}'$
- $\mathcal{F} \cap \mathbb{X} = \emptyset$
- $\mathcal{F}' \cap \mathbb{X} \neq \emptyset$

2

Cluster DMGs over ADMGs

- Definitions
- Summary causal graphs
- Causal tools

Inserm iPLesp SANTÉ SORBONNE UNIVERSITÉ

2

Cluster DMGs over ADMGs

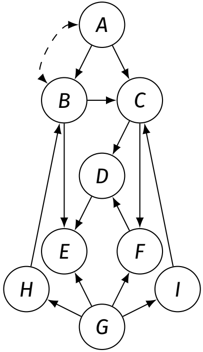
Definitions

Summary causal graphs

Causal tools

Cluster DMGs over ADMGs

ADMG

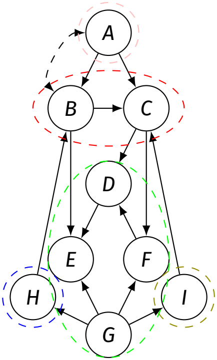


C-DMG over ADMGs

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

Cluster DMGs over ADMGs

ADMG

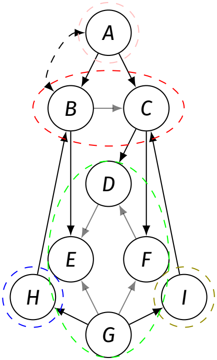


C-DMG over ADMGs

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

Cluster DMGs over ADMGs

ADMG

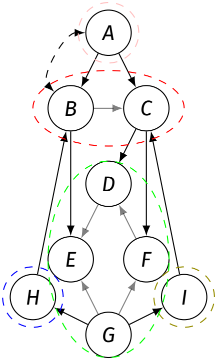


C-DMG over ADMGs

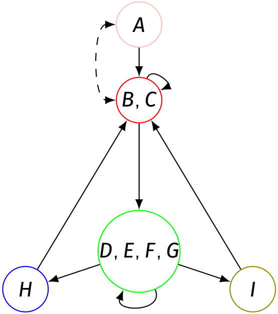
Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

Cluster DMGs over ADMGs

ADMG

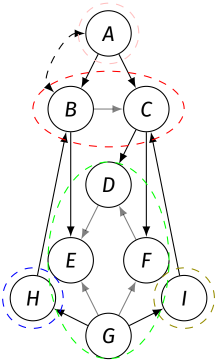


C-DMG over ADMGs

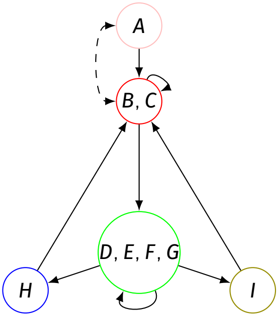


Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

ADMG



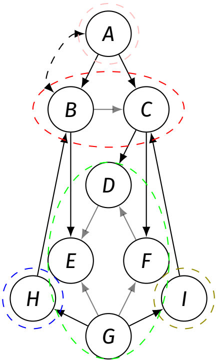
C-DMG over ADMGs



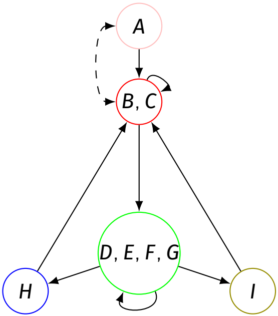
Cycles !

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

ADMG



C-DMG over ADMGs



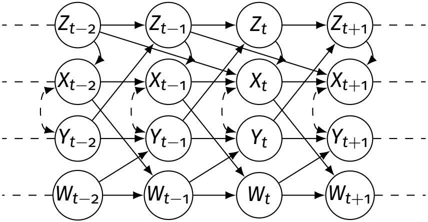
Assumption: For every cycle (e.g., $(B, C) \rightarrow (D, E, F, G) \rightarrow (H) \rightarrow (B, C)$) no 2 adjacent clusters are of size 1.

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs.

2

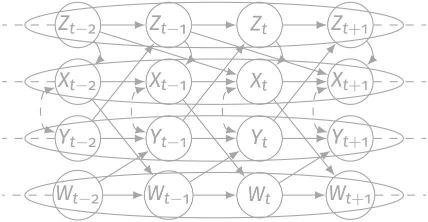
Cluster DMGs over ADMGs

ADMGs with time and summary causal graphs

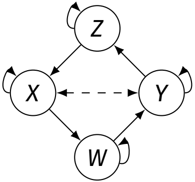


Assaad, Devijver, and Gaussier, "Survey and Evaluation of Causal Discovery Methods for Time Series". JAIR, 2022

ADMGs with time and summary causal graphs

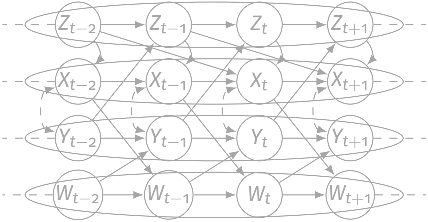


Summary causal graph (SCG)

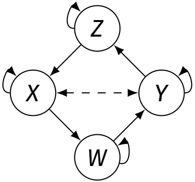


Assaad, Devijver, and Gaussier, "Survey and Evaluation of Causal Discovery Methods for Time Series". JAIR, 2022

ADMGs WITH TIME AND SUMMARY CAUSAL GRAPHS



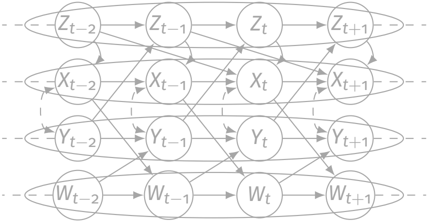
Summary causal graph (SCG)



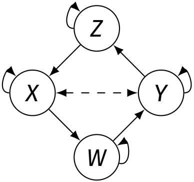
- SCGs can contain cycles

Assaad, Devijver, and Gaussier, "Survey and Evaluation of Causal Discovery Methods for Time Series". JAIR, 2022

ADMGs with time and summary causal graphs



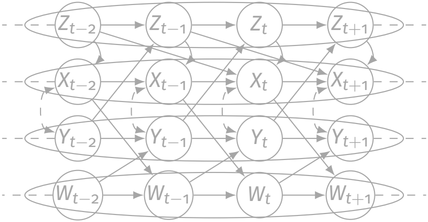
Summary causal graph (SCG)



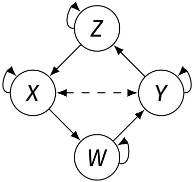
- SCGs can contain cycles
- Each vertex does not map to one single random variable

Assaad, Devijver, and Gaussier, “Survey and Evaluation of Causal Discovery Methods for Time Series”. JAIR, 2022

ADMGs with time and summary causal graphs



Summary causal graph (SCG)



- SCGs can contain cycles
- Each vertex does not map to one single random variable
- There might exists many ADMGs compatible with one SCG

Assaad, Devijver, and Gaussier, “Survey and Evaluation of Causal Discovery Methods for Time Series”. JAIR, 2022

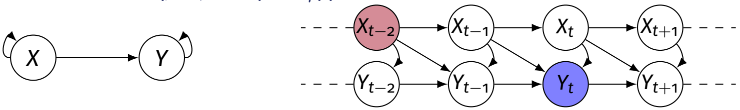
Two types of total effects in C-DMGs

A micro total effect is a total effect from a single variable in a cluster (e.g., $X_{t-\gamma}$) to another single variable (e.g., Y_t). For example: $\Pr(y_t \mid do(x_{t-\gamma}))$.

Ferreira and Assaad, "Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding".

Two types of total effects in C-DMGs

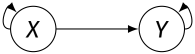
A micro total effect is a total effect from a single variable in a cluster (e.g., $X_{t-\gamma}$) to another single variable (e.g., Y_t). For example: $\Pr(y_t \mid do(x_{t-\gamma}))$.

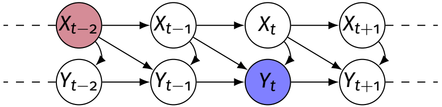


Ferreira and Assaad, "Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding".

Two types of total effects in C-DMGs

A micro total effect is a total effect from a single variable in a cluster (e.g., $X_{t-\gamma}$) to another single variable (e.g., Y_t). For example: $\Pr(y_t \mid do(x_{t-\gamma}))$.



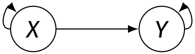


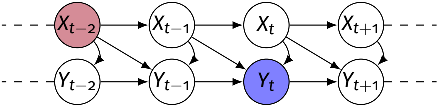
A macro total effect is a total effect from a whole cluster (e.g., $\{X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t\}$) to a whole other cluster (e.g., $\{Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t\}$). For example: $\Pr(\{y_{t_0}, y_{t_0+1}, \cdots, y_{t-1}, y_t\} \mid do(\{x_{t_0}, x_{t_0+1}, \cdots, x_{t-1}, x_t\}))$.

Ferreira and Assaad, "Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding".

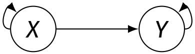
Two types of total effects in C-DMGs

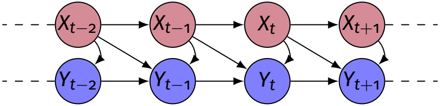
A micro total effect is a total effect from a single variable in a cluster (e.g., $X_{t-\gamma}$) to another single variable (e.g., Y_t). For example: $\Pr(y_t \mid do(x_{t-\gamma}))$.





A macro total effect is a total effect from a whole cluster (e.g., $\{X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t\}$) to a whole other cluster (e.g., $\{Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t\}$). For example: $\Pr(\{y_{t_0}, y_{t_0+1}, \cdots, y_{t-1}, y_t\} \mid do(\{x_{t_0}, x_{t_0+1}, \cdots, x_{t-1}, x_t\}))$.

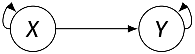


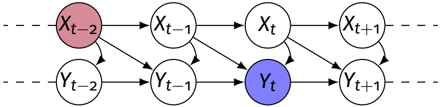


Ferreira and Assaad, "Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding".

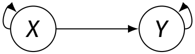
Two types of total effects in C-DMGs

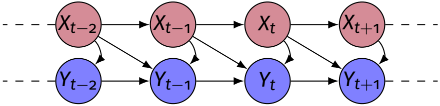
A micro total effect is a total effect from a single variable in a cluster (e.g., $X_{t-\gamma}$) to another single variable (e.g., Y_t). For example: $\Pr(y_t \mid do(x_{t-\gamma}))$.





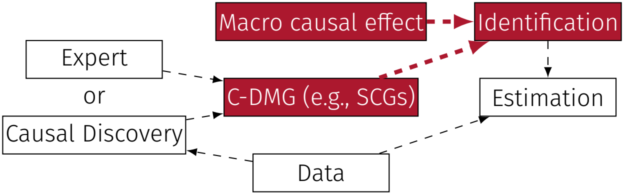
A macro total effect is a total effect from a whole cluster (e.g., $\{X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t\}$) to a whole other cluster (e.g., $\{Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t\}$). For example: $\Pr(\{y_{t_0}, y_{t_0+1}, \cdots, y_{t-1}, y_t\} \mid do(\{x_{t_0}, x_{t_0+1}, \cdots, x_{t-1}, x_t\}))$.





Ferreira and Assaad, "Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding".

Causal Reasoning



Inserm

iPLesp

SANTÉ SORBONNE UNIVERSITÉ

2

Cluster DMGs over ADMGs

d-separation for C-DMGs over ADMGs

Is d-separation applicable for C-DMGs over ADMGs?

d-separation for C-DMGs over ADMGs

Is d-separation applicable for C-DMGs over ADMGs?
Yes!

d-separation for C-DMGs over ADMGs

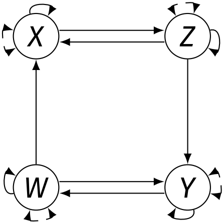
Is d-separation applicable for C-DMGs over ADMGs?
Yes!

Theorem

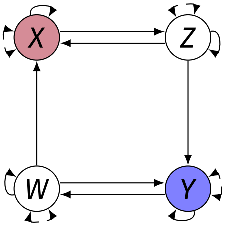
d-separation is valid in C-DMGs over ADMGs.

- *If a d-separation holds in a given C-DMG, then it holds in every compatible ADMG.*
- *If a d-separation does not hold in a given C-DMG, then there exists a compatible ADMG in which it does not hold.*

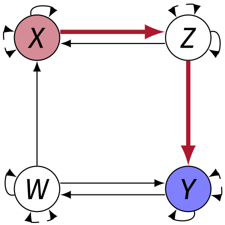
Demo: d-separation for C-DMGs over ADMGs



DEMO: D-SEPARATION FOR C-DMGS OVER ADMGS

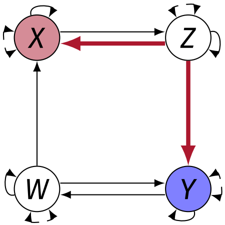


$$X \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Y \mid Z, W$$



Is the path $\langle X \rightarrow Z \rightarrow Y \rangle$ blocked by $\{Z, W\}$?

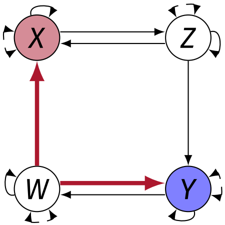
$$X \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Y \mid Z, W$$



Is the path $\langle X \leftarrow Z \rightarrow Y \rangle$ blocked by $\{Z, W\}$?

$$X \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Y \mid Z, W$$

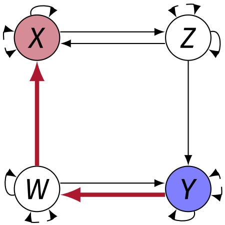
Demo: d-separation for C-DMGs over ADMGs



Is the path $\langle X \leftarrow W \rightarrow Y \rangle$ blocked by $\{Z, W\}$?

$$X \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Y \mid Z, W$$

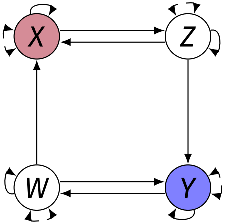
Demo: d-separation for C-DMGs over ADMGs



Is the path $\langle X \leftarrow W \leftarrow Y \rangle$ blocked by $\{Z, W\}$?

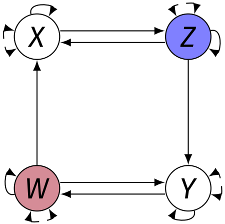
$$X \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Y \mid Z, W$$

Demo: d-separation for C-DMGs over ADMGs



$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

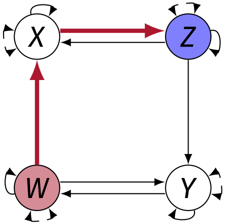
Demo: d-separation for C-DMGs over ADMGs



$$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$$

$$W \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Z \mid \mho \setminus \{X\}$$

Demo: d-separation for C-DMGs over ADMGs

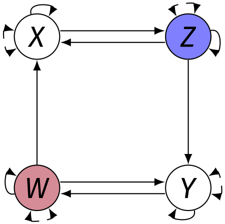


Is the path $\langle W \to X \to Z \rangle$ blocked by $\mho \setminus \{X\}$?

$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

$W \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Z \mid \mho \setminus \{X\}$

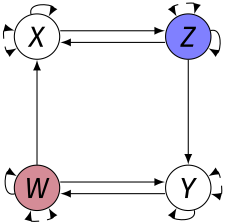
Demo: d-separation for C-DMGs over ADMGs



$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

$W \not\!\perp\!\!\!\perp_{\mathcal{G}} Z \mid \Omega \setminus \{X\}$

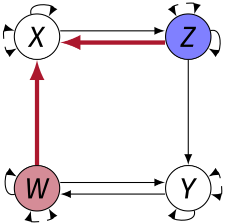
Demo: d-separation for C-DMGs over ADMGs



$$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$$

$$W \not\perp\!\!\!\perp_{\mathcal{G}} Z \mid \mho \setminus \{X\}$$

$$W \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Z \mid \mho \cup \{X\}$$



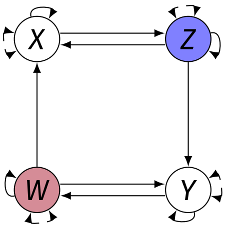
Is the path $\langle W \rightarrow X \leftarrow Z \rangle$ blocked by $\mho \cup \{X\}$?

$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

$W \not\perp\!\!\!\perp_{\mathcal{G}} Z \mid \mho \setminus \{X\}$

$W \overset{?}{\perp\!\!\!\perp}_{\mathcal{G}} Z \mid \mho \cup \{X\}$

Demo: d-separation for C-DMGs over ADMGs

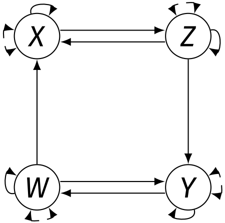


$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

$W \not\!\perp\!\!\!\perp_{\mathcal{G}} Z \mid \mho \setminus \{X\}$

$W \not\!\perp\!\!\!\perp_{\mathcal{G}} Z \mid \mho \cup \{X\}$

Demo: d-separation for C-DMGs over ADMGs



$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

$W \not\perp\!\!\!\perp_{\mathcal{G}} Z \mid \Omega \setminus \{X\}$

$W \not\perp\!\!\!\perp_{\mathcal{G}} Z \mid \Omega \cup \{X\}$

$X \perp\!\!\!\perp_{\mathcal{G}} Y \mid Z, W$

do-calculus for C-DMGs over ADMGs

Is do-calculus applicable for C-DMGs over ADMGs?

do-calculus for C-DMGs over ADMGs

Is do-calculus applicable for C-DMGs over ADMGs?

Yes!

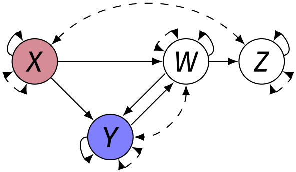
Is do-calculus applicable for C-DMGs over ADMGs?
Yes!

Theorem

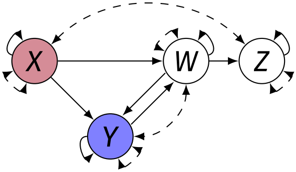
The Rules 1-3 of the do-calculus are valid in C-DMGs over DMGs.

- *If a sequence of rules apply in a given C-DMG, then it applies in every compatible ADMG.*
- *If a sequence of rules of the do-calculus does not apply in a given C-DMG, then there exists a compatible ADMG in which it does not apply.*

Demo: do-calculus for C-DMGs over ADMGs



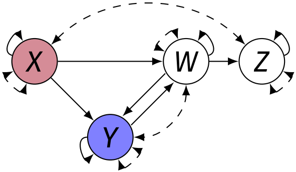
Demo: do-calculus for C-DMGs over ADMGs

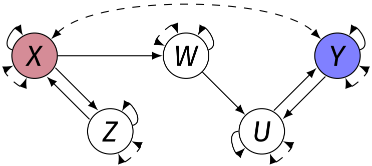


$P(\mathbb{y} \mid do(\mathbb{x}))$

$= P(\mathbb{y} \mid \mathbb{x})$ (Rule 2)

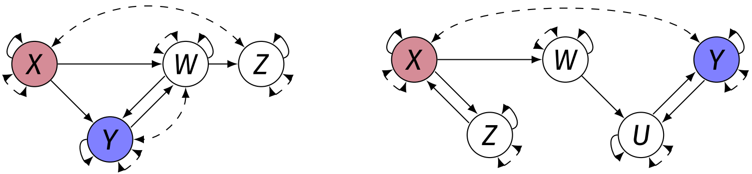
Demo: do-calculus for C-DMGs over ADMGs





$P(\mathbb{y} \mid do(\mathbb{x}))$

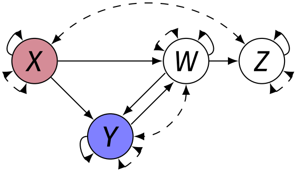
$= P(\mathbb{y} \mid \mathbb{x})$ (Rule 2)

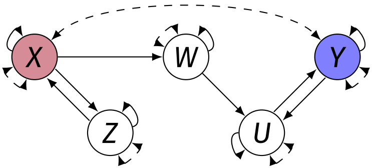


$P(\mathbb{y} \mid do(\mathbb{x}))$

$P(\mathbb{y} \mid do(\mathbb{x}))$

$= P(\mathbb{y} \mid \mathbb{x})$ (Rule 2)

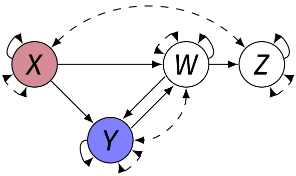


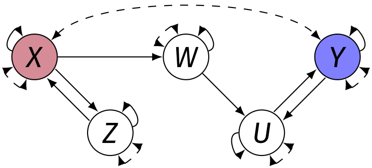


$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$

$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

Demo: do-calculus for C-DMGs over ADMGs

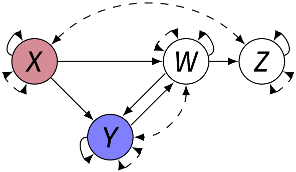


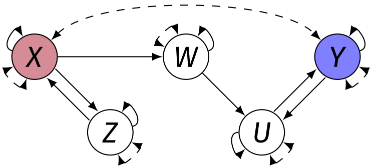


$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2) x2}$$

$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

DEMO: DO-CALCULUS FOR C-DMGS OVER ADMGS

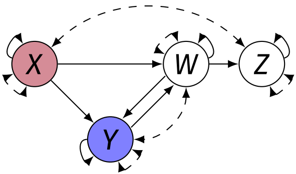




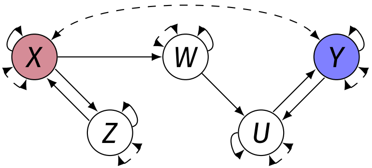
$$P(\mathbb{y} \mid do(\mathbb{x})) = P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

$$\begin{aligned} P(\mathbb{y} \mid do(\mathbb{x})) &= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x})) \\ &= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2) x2} \\ &= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 3)} \end{aligned}$$

Demo: do-calculus for C-DMGs over ADMGs

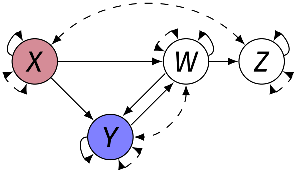


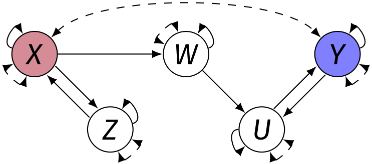
$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$



$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2) x2}$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 3)}$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid do(\mathbb{w}), \mathbb{x}') P(\mathbb{x}' \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x})$$

Demo: do-calculus for C-DMGs over ADMGs

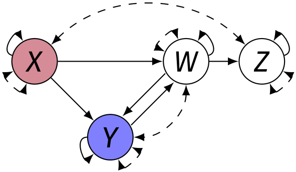


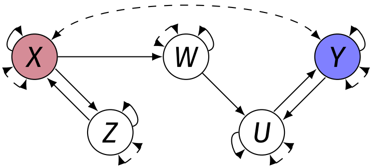


$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2) x2}$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 3)}$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid do(\mathbb{w}), \mathbb{x}') P(\mathbb{x}' \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x})$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid \mathbb{w}, \mathbb{x}') P(\mathbb{x}' \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

Demo: do-calculus for C-DMGs over ADMGs





$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), \mathbb{w}) P(\mathbb{w} \mid do(\mathbb{x}))$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{x}), do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2) x2}$$
$$= \sum_{\mathbb{w}} P(\mathbb{y} \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 3)}$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid do(\mathbb{w}), \mathbb{x}') P(\mathbb{x}' \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x})$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid \mathbb{w}, \mathbb{x}') P(\mathbb{x}' \mid do(\mathbb{w})) P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 2)}$$
$$= \sum_{\mathbb{w}, \mathbb{x}'} P(\mathbb{y} \mid \mathbb{w}, \mathbb{x}') P(\mathbb{x}') P(\mathbb{w} \mid \mathbb{x}) \quad \text{(Rule 3)}$$

$$P(\mathbb{y} \mid do(\mathbb{x}))$$
$$= P(\mathbb{y} \mid \mathbb{x}) \quad \text{(Rule 2)}$$

Completeness of d-separation and do-calculus for C-DMGs over ADMGs

Are the d-separation and do-calculus complete for C-DMGs over ADMGs?

Completeness of d-separation and do-calculus for C-DMGs over ADMGs

Are the d-separation and do-calculus complete for C-DMGs over ADMGs? No!

Completeness of d-separation and do-calculus for C-DMGs over ADMGs

Are the d-separation and do-calculus complete for C-DMGs over ADMGs? No!

Assumption: For every cycle, no two adjacent clusters are of size 1.

Completeness of d-separation and do-calculus for C-DMGs over ADMGs

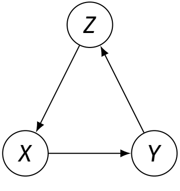
Are the d-separation and do-calculus complete for C-DMGs over ADMGs? No!

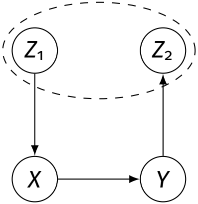
Assumption: For every cycle, no two adjacent clusters are of size 1. Then, Yes!

Completeness of d-separation and do-calculus for C-DMGs over ADMGs

Are the d-separation and do-calculus complete for C-DMGs over ADMGs? No!

Assumption: For every cycle, no two adjacent clusters are of size 1. Then, Yes!





Hedge for C-DMGs over ADMGs

Is the hedge criterion applicable for C-DMGs over ADMGs?

Hedge for C-DMGs over ADMGs

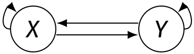
Is the hedge criterion applicable for C-DMGs over ADMGs?

No!

Hedge for C-DMGs over ADMGs

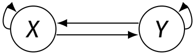
Is the hedge criterion applicable for C-DMGs over ADMGs?

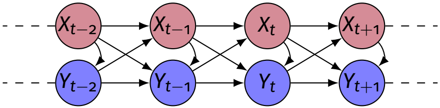
No!



Hedge for C-DMGs over ADMGs

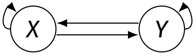
Is the hedge criterion applicable for C-DMGs over ADMGs?
No!

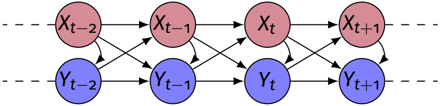


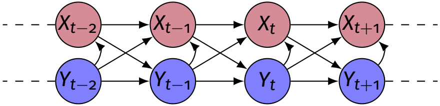


Is the hedge criterion applicable for C-DMGs over ADMGs?

No!



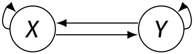




SC-Hedges for C-DMGs over ADMGs

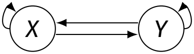
Consider an C-DMG $\mathcal{G}^s$. The SC-projection $\mathcal{H}^s$ of $\mathcal{G}^s$ is the graph that includes all vertices and edges from $\mathcal{G}^s$, plus a bidirected dashed edge between each pair $X, Y \in \mathbb{S}$ such that $X \in Scc(Y, \mathcal{G}) = De(Y, \mathcal{G}^s) \cap An(Y, \mathcal{G}^s)$ and $X \neq Y$.

Consider an C-DMG $\mathcal{G}^s$. The SC-projection $\mathcal{H}^s$ of $\mathcal{G}^s$ is the graph that includes all vertices and edges from $\mathcal{G}^s$, plus a bidirected dashed edge between each pair $X, Y \in \mathbb{S}$ such that $X \in Scc(Y, \mathcal{G}) = De(Y, \mathcal{G}^s) \cap An(Y, \mathcal{G}^s)$ and $X \neq Y$.

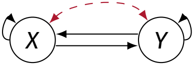


C-DMG

Consider an C-DMG $\mathcal{G}^s$. The SC-projection $\mathcal{H}^s$ of $\mathcal{G}^s$ is the graph that includes all vertices and edges from $\mathcal{G}^s$, plus a bidirected dashed edge between each pair $X, Y \in \mathbb{S}$ such that $X \in Scc(Y, \mathcal{G}) = De(Y, \mathcal{G}^s) \cap An(Y, \mathcal{G}^s)$ and $X \neq Y$.



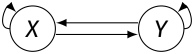
C-DMG



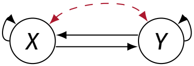
SC-projection

SC-Hedges for C-DMGs over ADMGs

Consider an C-DMG $\mathcal{G}^s$. The SC-projection $\mathcal{H}^s$ of $\mathcal{G}^s$ is the graph that includes all vertices and edges from $\mathcal{G}^s$, plus a bidirected dashed edge between each pair $X, Y \in \mathbb{S}$ such that $X \in Scc(Y, \mathcal{G}) = De(Y, \mathcal{G}^s) \cap An(Y, \mathcal{G}^s)$ and $X \neq Y$.



C-DMG



SC-projection

Theorem

Consider that $\mathcal{H}^s$ is the SC-projection of the C-DMG $\mathcal{G}^s$.
$P(\mathbb{y} \mid do(\mathbb{x}))$ *is not identifiable in $\mathcal{G}^s$ if there exists a hedge for the ordered pair $(\mathbb{X}, \mathbb{Y})$ in $\mathcal{H}^s$.*

What about micro-queries?

Reminder:

- A macro causal effect is a causal effect from a whole cluster $\mathbb{X}$ to a another whole cluster $\mathbb{Y}$, e.g., $\Pr(Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t \mid do(X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t))$.
- A micro causal effect is a total effect from a single variable X to another single variable Y, e.g., $\Pr(y_t \mid do(x_{t-\gamma}))$.

What about micro-queries?

Reminder:

- A macro causal effect is a causal effect from a whole cluster $\mathbb{X}$ to a another whole cluster $\mathbb{Y}$, e.g., $\Pr(Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t \mid do(X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t))$.
- A micro causal effect is a total effect from a single variable X to another single variable Y, e.g., $\Pr(y_t \mid do(x_{t-\gamma}))$.

Can we identify micro causal effects in C-DMGs over ADMGs?

What about micro-queries?

Reminder:

- A macro causal effect is a causal effect from a whole cluster $\mathbb{X}$ to a another whole cluster $\mathbb{Y}$, e.g., $\Pr(Y_{t_0}, Y_{t_0+1}, \cdots, Y_{t-1}, Y_t \mid do(X_{t_0}, X_{t_0+1}, \cdots, X_{t-1}, X_t))$.
- A micro causal effect is a total effect from a single variable X to another single variable Y, e.g., $\Pr(y_t \mid do(x_{t-\gamma}))$.

Can we identify micro causal effects in C-DMGs over ADMGs?

In specific cases and/or with prior knowledge, such as time, then yes!

What about micro-queries?

Can we identify micro causal effects in C-DMGs over ADMGs?

In specific cases and/or with prior knowledge, such as time, then yes!

- Assaad et al., "Identifiability of total effects from abstractions of time series causal graphs".
- Assaad, "Towards identifiability of micro total effects in summary causal graphs with latent confounding: extension of the front-door criterion".
- Ferreira and Assaad, "Identifiability of Direct Effects from Summary Causal Graphs".
- Ferreira and Assaad, "Average Controlled and Average Natural Micro Direct Effects in Summary Causal Graphs".

3

The Input/Output SCM framework

- Definitions
- Causal tools in DMGs

3

The Input/Output SCM framework

Definitions

Causal tools in DMGs

i/o SCM

$$
\begin{aligned}
\forall x,\ \xi_x &\leftarrow \mathcal{D}_x \\
A &:= f_A(\xi_a, \xi_{ab}) \\
B &:= f_B(A, D, E, \xi_b, \xi_{ab}) \\
C &:= f_C(B, \xi_c) \\
D &:= f_D(C, \xi_d) \\
E &:= f_E(C, \xi_e) \\
(B, C, D) &:= f_{(B,C,D)}(A, E, \xi_b, \xi_{ab}, \xi_c, \xi_d) \\
(B, C, E) &:= f_{(B,C,E)}(A, D, \xi_b, \xi_{ab}, \xi_c, \xi_e) \\
(B, C, D, E) &:= f_{(B,C,D)}(A, \xi_b, \xi_{ab}, \xi_c, \xi_d, \xi_e)
\end{aligned}
$$

Assumption: Compatibility of the generating processes in cycles.

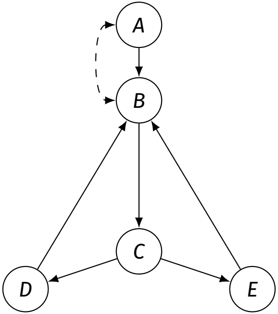
e.g., $(b, c, d) = f_{(B,C,D)}(a, e, \xi_b, \xi_{ab}, \xi_c, \xi_d) \implies b = f_B(a, d, e, \xi_b, \xi_{ab})$

Input/Output SCMs

i/o SCM

~~Acyclic~~ Directed Mixed Graph (DMG)

$$
\begin{aligned}
\forall x,\ \xi_x &\leftarrow \mathcal{D}_x \\
A &:= f_A(\xi_a, \xi_{ab}) \\
B &:= f_B(A, D, E, \xi_b, \xi_{ab}) \\
C &:= f_C(B, \xi_c) \\
D &:= f_D(C, \xi_d) \\
E &:= f_E(C, \xi_e) \\
(B, C, D) &:= f_{(B,C,D)}(A, E, \xi_b, \xi_{ab}, \xi_c, \xi_d) \\
(B, C, E) &:= f_{(B,C,E)}(A, D, \xi_b, \xi_{ab}, \xi_c, \xi_e) \\
(B, C, D, E) &:= f_{(B,C,D)}(A, \xi_b, \xi_{ab}, \xi_c, \xi_d, \xi_e)
\end{aligned}
$$



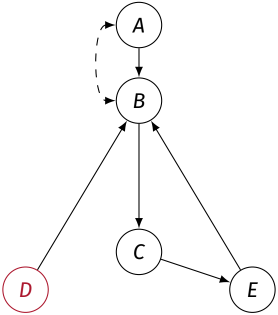
Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias"

Input/Output SCMs

i/o SCM

$$\forall x,\ \xi_x \leftarrow \mathcal{D}_x$$
$$A := f_A(\xi_a, \xi_{ab})$$
$$B := f_B(A, D, E, \xi_b, \xi_{ab})$$
$$C := f_C(B, \xi_c)$$
$$D := d$$
$$E := f_E(C, \xi_e)$$
$$\sout{(B, C, D) := f_{(B,C,D)}(A, E, \xi_b, \xi_{ab}, \xi_c, \xi_d)}$$
$$(B, C, E) := f_{(B,C,E)}(A, D, \xi_b, \xi_{ab}, \xi_c, \xi_e)$$
$$\sout{(B, C, D, E) := f_{(B,C,D)}(A, \xi_b, \xi_{ab}, \xi_c, \xi_d, \xi_e)}$$

Directed Mixed Graph (DMG)



Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias"

3

The Input/Output SCM framework

Definitions

Causal tools in DMGs

SIGMA-SEPARATION

Strongly connected component: $\boldsymbol{Scc}(V, \mathcal{G}) := \boldsymbol{An}(V, \mathcal{G}) \cap \boldsymbol{De}(V, \mathcal{G})$

A walk $\tilde{\pi} = \langle V_1, \cdots, V_n \rangle$ is said to be σ-blocked by a set of vertices $\mathbb{Z} \subseteq \mathbb{V}$ if:

1. $\exists 1 < i < n$ such that $\langle V_{i-1} {*\!\!\rightarrow} V_i {\leftarrow\!\!*} V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \notin \mathbb{Z}$, or
2. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i {\leftarrow\!\!*} V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}(V_{i-1}, \mathcal{G})$, or
3. $\exists 1 < i < n$ such that $\langle V_{i-1} {*\!\!\rightarrow} V_i \rightarrow V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}(V_{i+1}, \mathcal{G})$, or
4. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i \rightarrow V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \left(\boldsymbol{Scc}(V_{i-1}, \mathcal{G}) \cap \boldsymbol{Scc}(V_{i+1}, \mathcal{G}) \right)$.

Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias". PMLR, 2020.

SIGMA-SEPARATION

Strongly connected component: $\boldsymbol{Scc}(V, \mathcal{G}) := \boldsymbol{An}(V, \mathcal{G}) \cap \boldsymbol{De}(V, \mathcal{G})$

A walk $\tilde{\pi} = \langle V_1, \cdots, V_n \rangle$ is said to be σ-blocked by a set of vertices $\mathbb{Z} \subseteq \mathbb{V}$ if:

1. $\exists 1 < i < n$ such that $\langle V_{i-1} \ast\!\!\to V_i \leftarrow\!\!\ast V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \notin \mathbb{Z}$, or
2. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i \leftarrow\!\!\ast V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}(V_{i-1}, \mathcal{G})$, or
3. $\exists 1 < i < n$ such that $\langle V_{i-1} \ast\!\!\to V_i \to V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}(V_{i+1}, \mathcal{G})$, or
4. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i \to V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash (\boldsymbol{Scc}(V_{i-1}, \mathcal{G}) \cap \boldsymbol{Scc}(V_{i+1}, \mathcal{G}))$.

$\mathbb{X}$ and $\mathbb{Y}$ are σ-separated by $\mathbb{Z}$ if every walk between $\mathbb{X}$ and $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $(\mathbb{X} \perp\!\!\!\perp_\sigma \mathbb{Y} \mid \mathbb{Z})_\mathcal{G}$.

Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias". PMLR, 2020.

SIGMA-SEPARATION

Strongly connected component: $\boldsymbol{Scc}\left(V, \mathcal{G}\right) := \boldsymbol{An}\left(V, \mathcal{G}\right) \cap \boldsymbol{De}\left(V, \mathcal{G}\right)$

A walk $\tilde{\pi} = \langle V_1, \cdots, V_n \rangle$ is said to be σ-blocked by a set of vertices $\mathbb{Z} \subseteq \mathbb{V}$ if:

1. $\exists 1 < i < n$ such that $\langle V_{i-1} {*\!\!\rightarrow} V_i {\leftarrow\!\!*} V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \notin \mathbb{Z}$, or
2. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i {\leftarrow\!\!*} V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}\left(V_{i-1}, \mathcal{G}\right)$, or
3. $\exists 1 < i < n$ such that $\langle V_{i-1} {*\!\!\rightarrow} V_i \rightarrow V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \boldsymbol{Scc}\left(V_{i+1}, \mathcal{G}\right)$, or
4. $\exists 1 < i < n$ such that $\langle V_{i-1} \leftarrow V_i \rightarrow V_{i+1} \rangle \subseteq \tilde{\pi}$ and $V_i \in \mathbb{Z} \backslash \left(\boldsymbol{Scc}\left(V_{i-1}, \mathcal{G}\right) \cap \boldsymbol{Scc}\left(V_{i+1}, \mathcal{G}\right)\right)$.

$\mathbb{X}$ and $\mathbb{Y}$ are σ-separated by $\mathbb{Z}$ if every walk between $\mathbb{X}$ and $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\left(\mathbb{X} \perp\!\!\!\perp_{\sigma} \mathbb{Y} \mid \mathbb{Z}\right)_{\mathcal{G}}$.

Theorem

$\left(\mathbb{X} \perp\!\!\!\perp_{\sigma} \mathbb{Y} \mid \mathbb{Z}\right)_{\mathcal{G}} \Rightarrow \mathbb{X} \perp\!\!\!\perp_{\Pr} \mathbb{Y} \mid \mathbb{Z}$

Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias". PMLR, 2020.

do-calculus for DMGs

The σ-based do-calculus consists of three rules:

Rule 1 $P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\overline{\mathbb{Z}(\mathbb{W})}}}$

Forré and Mooij, “Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias”. PMLR, 2020.

The σ-based do-calculus consists of three rules:

Rule 1 $P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\overline{\mathbb{Z}(\mathbb{W})}}}$

Theorem

$P(\mathbb{y} \mid do(\mathbb{x}))$ is identifiable if ~~and only if~~ there exists a finite sequence of transformations, each conforming to either one of the Rules 1-3 or some standard probability manipulations, that reduces $P(\mathbb{y} \mid do(\mathbb{x}))$ into a do-free formula.

Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias". PMLR, 2020.

DO-CALCULUS FOR DMGs

The σ-based do-calculus consists of three rules:

Rule 1 $P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}}}$

Rule 2 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{z}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\underline{\mathbb{Z}}}}$

Rule 3 $P(\mathbb{y}|do(\mathbb{x}, \mathbb{z}), \mathbb{w}) = P(\mathbb{y}|do(\mathbb{x}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp_\sigma \mathbb{Z} \mid \mathbb{X}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{X}}\overline{\mathbb{Z}(\mathbb{W})}}}$

Theorem

$P(\mathbb{y} \mid do(\mathbb{x}))$ is identifiable if ~~and only if~~ there exists a finite sequence of transformations, each conforming to either one of the Rules 1-3 or some standard probability manipulations, that reduces $P(\mathbb{y} \mid do(\mathbb{x}))$ into a do-free formula.

The do-calculus is sound!

Forré and Mooij, "Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias". PMLR, 2020.

4

Cluster DMGs over DMGs

- Definitions
- Causal tools

4

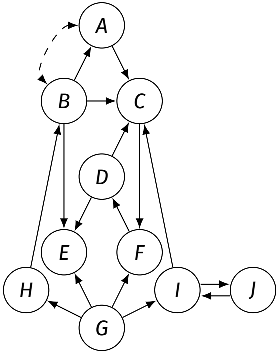
Cluster DMGs over DMGs

Definitions

Causal tools

Cluster DMGs over DMGs

DMG

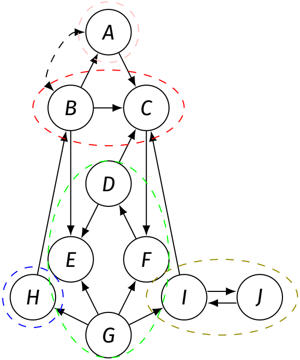


C-DMG over DMGs

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs over DMGs.

Cluster DMGs over DMGs

DMG

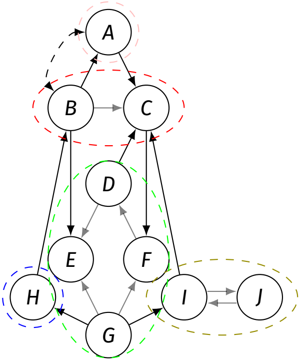


C-DMG over DMGs

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs over DMGs.

Cluster DMGs over DMGs

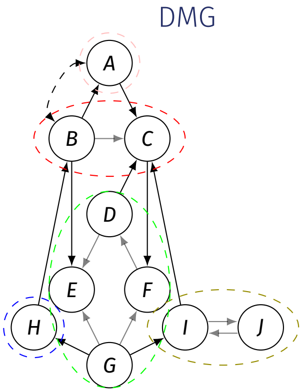
DMG

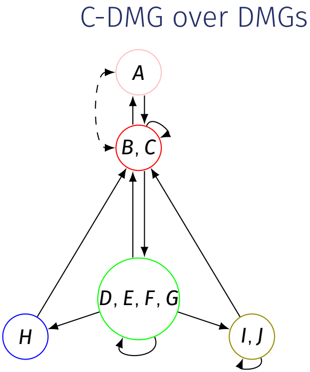


C-DMG over DMGs

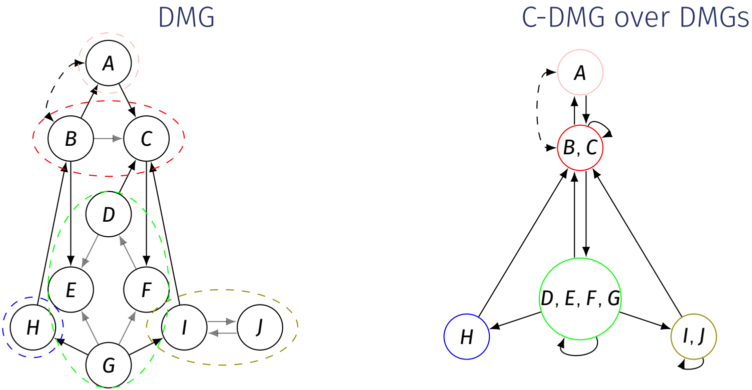
Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs over DMGs.

Cluster DMGs over DMGs





Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs over DMGs.



When clustering, cycles can:

- "disappear" (e.g., $I \to J \to I$) and/or
- "remain" (e.g., $C \to F \to D \to C$) and/or
- "appear" (e.g., $A \to (B, C) \to A$).

Ferreira and Assaad, Identifying Macro Causal Effects in C-DMGs over DMGs.

4

Cluster DMGs over DMGs

Definitions

Causal tools

σ-separation for C-DMGs over DMGs

Is σ-separation applicable for C-DMGs over DMGs?

σ-separation for C-DMGs over DMGs

Is σ-separation applicable for C-DMGs over DMGs?
Yes!

σ-separation for C-DMGs over DMGs

Is σ-separation applicable for C-DMGs over DMGs?
Yes!

Theorem

σ-separation is valid in C-DMGs over DMGs.

- *If a σ-separation holds in a given C-DMG, then it holds in every compatible DMG.*
- *If a σ-separation does not hold in a given C-DMG, then there exists a compatible DMG in which it does not hold.*

σ-based do-calculus for C-DMGs over DMGs

Is σ-based do-calculus applicable for C-DMGs over DMGs?

σ-based do-calculus for C-DMGs over DMGs

Is σ-based do-calculus applicable for C-DMGs over DMGs?

Yes!

σ-based do-calculus for C-DMGs over DMGs

Is σ-based do-calculus applicable for C-DMGs over DMGs?
Yes!

Theorem

The Rules 1-3 of the do-calculus are valid in C-DMGs over DMGs.

- *If a sequence of rules apply in a given C-DMG, then it applies in every compatible DMG.*
- *If a sequence of rules of the do-calculus does not apply in a given C-DMG, then there exists a compatible DMG in which it does not apply.*

5

Conclusion

Key Message

Use causal graphs !

- Even if the causal graph is not fully specified,
- Even if the causal graph is cyclic.

Now, there exist tools which accomodate these cases.

THANK YOU FOR YOUR ATTENTION!

ANY QUESTIONS?

Inserm

iPLesp

SANTÉ SORBONNE UNIVERSITÉ

References I

Assaad, Charles K. "Towards identifiability of micro total effects in summary causal graphs with latent confounding: extension of the front-door criterion". In: Transactions on Machine Learning Research (2025). ISSN: 2835-8856. URL: https://openreview.net/forum?id=5f7YlSKG1l.

Assaad, Charles K., Emilie Devijver, and Eric Gaussier. "Survey and Evaluation of Causal Discovery Methods for Time Series". In: 73 (May 2022). ISSN: 1076-9757. DOI: 10.1613/jair.1.13428. URL: https://doi.org/10.1613/jair.1.13428.

References II

- Assaad, Charles K. et al. "Identifiability of total effects from abstractions of time series causal graphs". In: Proceedings of the Fortieth Conference on Uncertainty in Artificial Ed. by Negar Kiyavash and Joris M. Mooij. Vol. 244. Proceedings of Machine Learning Research. PMLR, July 2024, pp. 173–185.
- Ferreira, Simon and Charles K. Assaad. "Average Controlled and Average Natural Micro Direct Effects in Summary Causal Graphs". In: (2025). Submitted.
- — ."Identifiability of Direct Effects from Summary Causal Graphs". In: Proceedings of the AAAI Conference on Artificial Intelligence 38.18 (Mar. 2024), pp. 20387–20394. DOI: 10.1609/aaai.v38i18.30021.

References III

Ferreira, Simon and Charles K. Assaad. Identifying Macro Causal Effects in C-DMGs. 2025. arXiv: 2504.01551 [cs.AI].

— .“Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding”. In: (2025). Causal Inference for Time Series Data Workshop at the 40th Conference on Uncertainty in Artificial Intelligence.

Forré, Patrick and Joris M. Mooij. “Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias”. In: Proceedings of The 35th Uncertainty in Artificial Intelligence Confe Ed. by Ryan P. Adams and Vibhav Gogate. Vol. 115. Proceedings of Machine Learning Research. PMLR, 2020, pp. 71–80.

References IV

- Pearl, Judea. "Causal diagrams for empirical research". In: Biometrika 82.4 (Dec. 1995), pp. 669–688. ISSN: 0006-3444. DOI: 10.1093/biomet/82.4.669.
- — .Causality: Models, Reasoning, and Inference. Cambridge University Press, 2009.
- — .Probabilistic Reasoning in Intelligent Systems: Networks of Pla San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988. ISBN: 0-934613-73-7.
- Shpitser, Ilya and Judea Pearl. "Complete identification methods for the causal hierarchy". In: Journal of Machine Learning Research 9 (2008). URL: https://www.jmlr.org/papers/volume9/shpitser08a/shpitser08a.pdf.