Maximal Ancestral Graphs

Part 1: Fundamentals

Sebastián Bejos, Eduardo Morales and Enrique Sucar.

Mixed Graphs

Definition

A (directed) **mixed graph** is a graph that may contain two kinds of edges: directed edges ($\to$) and bi-directed edges ($\leftrightarrow$).

- Between any two vertices there is at most one edge.
- The two ends of an edge we call **marks**.
- There are two kinds of marks: **arrowhead** ($>$) and **tail** ($-$).
- We say an edge is **into** (or **out of**) a vertex if the mark of the edge at the vertex is an arrowhead (or tail).

Mixed Graphs

We use the following terminology to describe relations between variables a mixed graph $\mathcal{G}$:

$$\text{If } \left\{ \begin{array}{c} X \leftrightarrow Y \\ X \rightarrow Y \\ X \leftarrow Y \end{array} \right\} \text{in } \mathcal{M} \text{ then } X \text{ is a } \left\{ \begin{array}{l} \textbf{spouse} \\ \textbf{parent} \\ \textbf{child} \end{array} \right\} \text{ of } Y \text{ and } \left\{ \begin{array}{c} X \in \boldsymbol{sp}(Y) \\ X \in \boldsymbol{pa}(Y) \\ X \in \boldsymbol{ch}(Y) \end{array} \right\}$$

Definition

A vertex X is said to be an **ancestor** of a vertex Y, $X \in \boldsymbol{an}(Y)$, if either there is a directed path $X \rightarrow \cdots \rightarrow Y$ from X to Y, or $X = Y$.

Ancestral Graphs

Definition

A mixed (directed) graph is an **ancestral graph** if:

(a) there are no directed cycles;

(b) whenever there is an edge $X \leftrightarrow Y$, then there is no directed path from X to Y, or from Y to X.

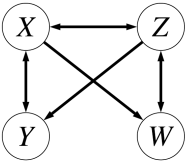


Figure 1: (a) An ancestral graph.

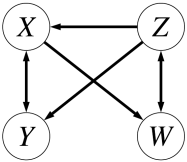


Figure 2: (b) No ancestral graph.

Collider Paths

Definition

(a) In an ancestral graph, a nonendpoint vertex X on a path is said to be a **collider** if two arrowheads meet at X (i.e., $\rightarrow X \leftarrow$, $\leftrightarrow X \leftrightarrow$, $\leftrightarrow X \leftarrow$. $\rightarrow X \leftrightarrow$).

(b) All other nonendpoint vertices on a path are **noncolliders** (i.e., $\rightarrow X \rightarrow$, $\leftarrow X \leftarrow$, $\leftarrow X \rightarrow$, $\leftrightarrow X \rightarrow$, $\leftarrow X \leftrightarrow$)

(c) A path along which every nonendpoint is a collider is called a **collider path**.

m-Connecting Paths

Definition

In an ancestral graph, a path π between vertices X and Y is **active** or ***m*-connecting** relative to a (possibly empty) set of vertices $\boldsymbol{Z}$, with $X, Y \notin \boldsymbol{Z}$ if

(i) every noncollider on π is not a member of $\boldsymbol{Z}$;

(ii) every collider on π is an ancestor of some member of $\boldsymbol{Z}$.

(iii) otherwise, $\boldsymbol{Z}$ **blocks** π.

Example: For the ancestral graph $A \rightarrow B \leftrightarrow C \leftarrow D$.

- The path $\pi_1 = (A, B, C, D)$ is active relative to $\boldsymbol{Z} = \{B, C\}$.
- The path π_1 is not *m*-connecting relative to $\boldsymbol{Z} = \emptyset$, $\boldsymbol{Z} = \{B\}$ or $\boldsymbol{Z} = \{C\}$, i.e., $\boldsymbol{Z} = \emptyset$, $\boldsymbol{Z} = \{B\}$ and $\boldsymbol{Z} = \{C\}$ blocks π_1.

m-Separation

Definition

- X and Y are said to be m-separated by $\boldsymbol{Z}$ if there is no active path between X and Y relative to $\boldsymbol{Z}$, i.e., if $\boldsymbol{Z}$ blocks all paths between X and Y.
- Two disjoint sets of variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ are m-separated by $\boldsymbol{Z}$ if every variable in $\boldsymbol{X}$ is m-separated from every variable in $\boldsymbol{Y}$ by $\boldsymbol{Z}$.

Example: For the ancestral graph $A \rightarrow B \leftrightarrow C \leftarrow D$.

- $(\{A\} \not\perp\!\!\!\perp_m \{D\} \mid \{B, C\})$ since $\pi_1 = (A, B, C, D)$ is active relative to $\boldsymbol{Z} = \{B, C\}$.
- $(\{A\} \perp\!\!\!\perp_m \{D\})$, $(\{A\} \perp\!\!\!\perp_m \{D\} \mid \{B\})$ and $(\{A\} \perp\!\!\!\perp_m \{D\} \mid \{C\})$, since there is no active path relative to $\boldsymbol{Z} = \emptyset$, $\boldsymbol{Z} = \{B\}$ and $\boldsymbol{Z} = \{C\}$, respectively.

Formal Independence Models

Definition

An independence model over a finite set $\boldsymbol{V}$ is a set $\boldsymbol{I}$ of ternary relations $\langle \boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{Z} \rangle$ where $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are disjoint subsets of $\boldsymbol{V}$, while $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not empty.

- The interpretation $\langle \boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{Z} \rangle$ is that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent given $\boldsymbol{Z}$.
- The independence model associated with an ancestral graph, $\boldsymbol{I}_m(\mathcal{G})$, is defined via m-separation as follows:

$$\boldsymbol{I}_m(\mathcal{G}) = \{\langle \boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{Z} \rangle | \boldsymbol{X} \text{ is } m\text{-separated from } \boldsymbol{Y} \text{ given in } \boldsymbol{Z}\}$$

Maximal Ancestral Graphs

Definition

An ancestral graph $\mathcal{G}$ is said to be **maximal** if, for every pair of nonadjacent vertices (X, Y), there exists a set $\boldsymbol{Z}$ $(X, Y \notin \boldsymbol{Z})$ such that X and Y are m-separated conditional on $\boldsymbol{Z}$, i.e., $\langle\{X\}, \{Y\} \mid \boldsymbol{Z}\rangle \in \boldsymbol{I}_m(\mathcal{G})$.

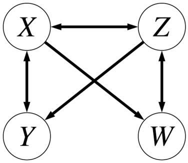


Figure 3: (a) A not maximal ancestral graph.

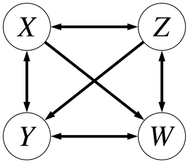
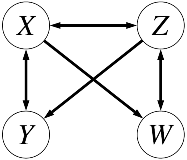


Figure 4: (b) A maximal ancestral graph.

Maximal Ancestral Graphs



(Y, W) is the only pair of nonadjacent vertices.

- For $\boldsymbol{Z} = \{X, Z\}$, $\pi_1 = (Y, X, Z, W)$ its a path that m-connects Y and W.
- For $\boldsymbol{Z} = \{X\}$, π_1 it is not an active path since $Z \notin \boldsymbol{an}(X)$, but $\pi_2 = (Y, Z, W)$ is active as there are no colliders in π_2, Z is noncollider for π_2 and $Z \notin \boldsymbol{Z}$.
- For $\boldsymbol{Z} = \{Z\}$, $\pi_3 = (Y, X, W)$ is an active path as there are no colliders in π_3, X is noncollider for π_3 and $X \notin \boldsymbol{Z}$.
- For $\boldsymbol{Z} = \emptyset$, π_1, π_2 and π_3 are active paths.

Maximal Ancestral Graphs

- **Maximal ancestral graphs (MAGs)** are maximal in the sense that no additional edge may be added to the graph without changing the independence model.

Proposition

If $\mathcal{G} = (\boldsymbol{V}, \boldsymbol{E})$ *is a maximal ancestral graph, and* $\mathcal{G}$ *is a subgraph of* $\mathcal{G}^* = (\boldsymbol{V}, \boldsymbol{E}^*)$*, then* $\boldsymbol{I}_m(\mathcal{G}) = \boldsymbol{I}_m(\mathcal{G}^*)$ *implies* $\mathcal{G} = \mathcal{G}^*$

Maximal Ancestral Graphs

- The following theorem gives the converse.

Theorem

If $\mathcal{G}$ is an ancestral graph then there exists a unique maximal ancestral graph $\mathcal{M}$ formed by adding $\leftrightarrow$ edges to $\mathcal{G}$ such that $\boldsymbol{I}_m(\mathcal{G}) = \boldsymbol{I}_m(\mathcal{M})$.

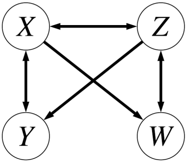


Figure 5: (a) An ancestral graph $\mathcal{G}$.

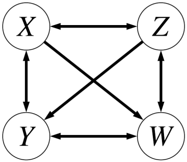


Figure 6: (b) The maximal ancestral graph $\mathcal{M}$ from $\mathcal{G}$.

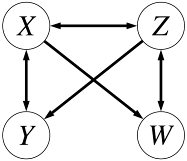
Inducing Paths

Definition

An inducing path π relative to a set $\boldsymbol{L}$, between vertices X and Y in an ancestral graph $\mathcal{G}$, is a path on which every nonendpoint vertex not in $\boldsymbol{L}$ is both a collider on π and an ancestor of at least one of the endpoints, X and Y.

- Any single-edge path is trivially an inducing path relative to any set of vertices.
- To simplify terminology, we will henceforth refer to inducing paths relative to the empty set simply as inducing paths

Inducing Paths



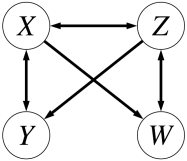
- The path (Y, Z, W) is an inducing path relative to $\{Z\}$, but not an inducing path relative to the empty set (because Z is not a collider)
- The path (Y, X, Z, W) is an inducing path relative to the empty set, because both X and Z are colliders on the path, X is an ancestor of W, and Z is an ancestor of Y.

Alternative definition to MAGs

Definition

A mixed graph is called a **maximal ancestral graph (MAG)** if

- i the graph does not contain any directed or almost directed cycles (**ancestral**); and
- ii there is no inducing path between any two non-adjacent vertices (**maximal**).



- The graph is not maximal because the path (Y, X, Z, W) is an inducing path between the non-adjacent vertices Y and W.

DAGs to MAGs

- Given any DAG $\mathcal{D}$ over $\boldsymbol{V} = \boldsymbol{O} \cup \boldsymbol{L}$ there is a MAG $\mathcal{M}$ over $\boldsymbol{O}$ alone, such that for any disjoint sets $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{O}$, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are *d*-separated by $\boldsymbol{Z}$ in $\mathcal{D}$ if and only if they are *m*-separated by $\boldsymbol{Z}$ in the MAG $\mathcal{M}$.
- The following construction gives us such a MAG:

DAGs to MAGs

Input: A DAG $\mathcal{D}$ over $\boldsymbol{O} \cup \boldsymbol{L}$
Output: A MAG $\mathcal{M}$ over $\boldsymbol{O}$

i For each pair of variables $X, Y \in \boldsymbol{O}$, X and Y are adjacent in $\mathcal{M}$ if and only if there is an inducing path between them relative to $\boldsymbol{L}$ in $\mathcal{D}$.

ii For each pair of adjacent variables X, Y in $\mathcal{M}$,
   (a) orient the edge as $X \to Y$ in $\mathcal{M}$ if X is an ancestor of Y in $\mathcal{D}$;
   (b) orient it as $X \leftarrow Y$ in $\mathcal{M}$ if Y is an ancestor of X in $\mathcal{D}$;
   (c) orient it as $X \leftrightarrow Y$ in $\mathcal{M}$ otherwise.

DAGs to MAGs

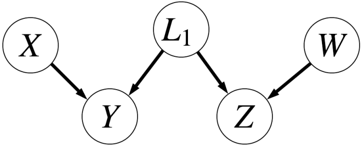


Figure 7: (a) A DAG $\mathcal{D}$ over $\boldsymbol{O} \cup \boldsymbol{L}$ with $\boldsymbol{L} = \{L_1\}$

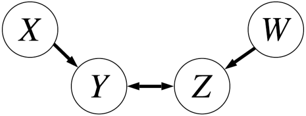


Figure 8: (b) The MAG $\mathcal{M}$ over $\boldsymbol{O}$ from the DAG $\mathcal{D}$

DAGs to MAGs

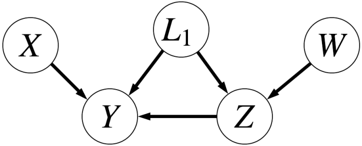


Figure 9: (a) A DAG $\mathcal{D}$ over $\boldsymbol{O} \cup \boldsymbol{L}$ with $\boldsymbol{L} = \{L_1\}$

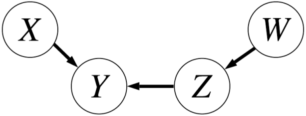


Figure 10: (b) The MAG $\mathcal{M}$ over $\boldsymbol{O}$ from the DAG $\mathcal{D}$

DAGs to MAGs

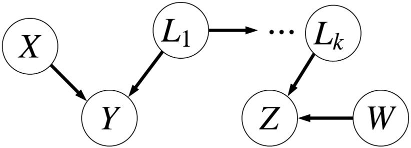


Figure 11: (a) A DAG $\mathcal{D}$ over $\boldsymbol{O} \cup \boldsymbol{L}$ with $\boldsymbol{L} = \{L_1, \ldots, L_k\}$

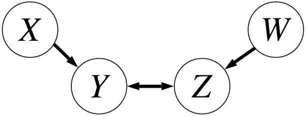


Figure 12: (b) The MAG $\mathcal{M}$ over $\boldsymbol{O}$ from the DAG $\mathcal{D}$

Meanning of the edges of a MAG

Directed edges as $X \rightarrow Y$ means:

- i X is an ancestor of Y.
- ii Y is not an ancestor of X.
- iii This **does not rule out** possible latent confounding between X and Y.

Bi-directed edges as $X \leftrightarrow Y$ means:

- i X is not an ancestor of Y.
- ii Y is not an ancestor of X.
- iii X and Y are **confounded**.

Canonical DAGs

Definition

If $\mathcal{G}$ is an ancestral graph with vertex set $\boldsymbol{V}$, then we define the **canonical DAG**, $\mathcal{D}(\mathcal{G})$ associated with $\mathcal{G}$ as follows:

- i Let $\boldsymbol{L}_{\mathcal{D}(\mathcal{G})} = \{\lambda_{XY} \mid X \leftrightarrow Y \text{ in } \mathcal{G}\}$
- ii DAG $\mathcal{D}(\mathcal{G})$ has vertex set $\boldsymbol{V} \cup \boldsymbol{L}_{\mathcal{D}(\mathcal{G})}$ and edge set defined as:

$$\text{If } \left\{ \begin{array}{c} X \rightarrow Y \\ X \leftrightarrow Y \end{array} \right\} \text{in } \mathcal{G} \text{ then } \left\{ \begin{array}{l} X \rightarrow Y \\ X \leftarrow \lambda_{XY} \rightarrow Y \end{array} \right\} \text{in } \mathcal{D}(\mathcal{G}).$$

Markov equivalence

- Several MAGs can also encode the same conditional independencies via *m*-separation.
- Such MAGs form a **Markov equivalence class** which can be described uniquely by a **partial ancestral graph (PAG)**.
- A PAG $\mathcal{P}$ has the same adjacencies as any MAG in the Markov equivalence class described by $\mathcal{P}$.
- We denote all MAGs in the Markov equivalence class described by a PAG $\mathcal{G}$ by $[\mathcal{G}]$.

Partial Ancestral Graphs

Definition

Let $[\mathcal{M}]$ be the Markov equivalence class of an arbitrary MAG $\mathcal{M}$. The **partial ancestral graph** (PAG) for $[\mathcal{M}]$, $\mathcal{P}_{[\mathcal{M}]}$, is a partial mixed graph such that

i $\mathcal{P}_{[\mathcal{M}]}$ has the same adjacencies as $\mathcal{M}$ (and any member of $[\mathcal{M}]$) does;

ii A mark of arrowhead is in $\mathcal{P}_{[\mathcal{M}]}$ if and only if it is shared by all MAGs in $[\mathcal{M}]$; and

iii A mark of tail is in $\mathcal{P}_{[\mathcal{M}]}$ if and only if it is shared by all MAGs in $[\mathcal{M}]$.

Causal Bayesian networks

Definition

A **Bayesian network** for a set of variables $\boldsymbol{V} = \{X_1, \dots Xp\}$ is a pair $(\mathcal{G}, f)$, where $\mathcal{G}$ is a DAG, and f is a joint density for $\boldsymbol{V}$ that factorizes as $\prod_{i=1}^{p} f(X_i \mid \boldsymbol{pa}(X_i))$.

Definition

We call a **DAG causal** if every edge $X_i \rightarrow X_j$ in $\mathcal{G}$ represents a direct causal effect of X_i on X_j.

Definition

A Bayesian network $(\mathcal{G}, f)$ is a **causal Bayesian network** if $\mathcal{G}$ is a causal DAG.

Consistent densities

- A density f is consistent with a causal DAG $\mathcal{D}$ if the pair $(\mathcal{D}, f)$ forms a causal Bayesian network.
- A density f is **consistent with a causal MAG** $\mathcal{M}$ if there exists a causal Bayesian network $(\mathcal{D}^*, f^*)$ such that $\mathcal{M}$ represents $\mathcal{D}^*$ and f is the observed marginal of f^*.
- A density f is **consistent with a causal PAG** $\mathcal{G}$ if it is consistent with a causal MAG in $[\mathcal{G}]$.