# Lecture 5: Gaussian graphical models. + Interventions. + iDAGs

$G =$ DAG (directed acyclic graph)
$= ([p], E)$

For us the nodes are random variables $X_i$, $i \in [p]$.

$$f_{A \cup B | C}(x_A, x_B | x_c) = f_{A|C}(x_A | x_c) \cdot f_{B|C}(x_A | x_c)$$
$$\Updownarrow$$

- local Markov properties associated to $G$:
  set of all conditional indep statements $\underbrace{X_v}_{A} \perp\!\!\!\perp \underbrace{X_{nd(v) \setminus pa(v)}}_{B} \mid \underbrace{X_{pa(v)}}_{C}$

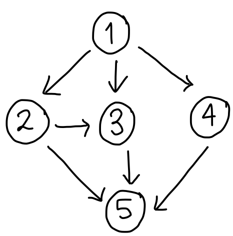
- Recursive factorization: $f(x) = \prod_{j \in V} f_j(x_j | x_{pa(j)})$

<u>Def</u>: The parametric directed graphical model associated to the DAG $G$ consist of all probability densities that factorize as

$$f(x) = \prod_{j \in V} f_j(x_j | x_{pa(j)})$$

<u>Theorem</u>: A prob. distrib. satisfies the recursive factorization prop. associated to the directed acyclic graph $G \Leftrightarrow$ it satisfies the directed local Markov prop. assoc. to $G$.

<u>Example</u>:




$v = 2$ $\quad 2 \perp\!\!\!\perp 4 \mid 1$

$pa(2) = \{1\}$

$nd(2) = 4$

$v = 3$ $\quad 3 \perp\!\!\!\perp 4 \mid 1, 2$

$$f(x) = f(x_1) f(x_2 | x_1) f(x_3 | x_1, x_2) f(x_4 | x_1) f(x_5 | x_2, x_3)$$

- Separation for undirected $G$.

<u>Def</u>:
- A pair of vertices $a, b \in V$ is separated by a set of vertices $C \subseteq V \setminus \{a, b\}$ if every path from $a$ to $b$ contains a vertex in $C$.
- If $A, B, C$ are disjoint subsets of $V$, we say that $C$ separates $A$ and $B$ if $a$ and $b$ are separated by $C$ for all $a \in A$, $b \in B$.

For what comes we need

* directed paths: $\bullet \rightarrow \bullet \rightarrow \bullet \cdots \bullet \rightarrow \bullet$
* undirected paths: $\bullet \longrightarrow \bullet \longrightarrow \underbrace{\bullet}_{\text{Collider}} \longleftarrow \bullet \longleftarrow \underbrace{\bullet}_{\text{non-collider.}}$

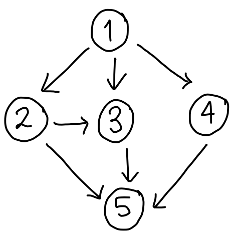
**Def:** Two nodes $v$ and $w$ in a DAG are d-connected given a set $C \subseteq V \setminus \{v, w\}$ if there is an undirected path $\pi$ from $v$ to $w$ such that

(1) all colliders on $\pi$ are in $C \cup an(C)$ and

(2) no noncollider on $\pi$ is in $C$

* If $A, B, C \subseteq V$ are pairwise disjoint with $A, B \neq \emptyset$, then $C$ d-separates $A$ and $B$ if no pair of nodes $a \in A$ and $b \in B$ are d-connected given $C$.

**Example:**



* 2 is d-connected to 4 given 5, $2 \rightarrow 5 \leftarrow 4$
* 2 is not d-connected to 4 given 1

$2 \longleftarrow 1 \longrightarrow 4$

$2 \longrightarrow 5 \longleftarrow 4$

$C = \{1\} \quad an(C) = \{\}$

→ The global Markov property captures all the conditional indep. properties that are implied by the graph.

**Def:** The directed global Markov property associated to a DAG $G$ consists of all conditional indep. statements

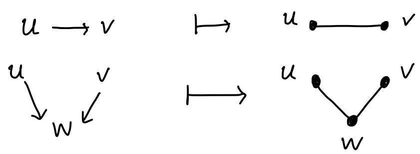
$X_A \perp\!\!\!\perp X_B \mid X_C$, for all disjoint sets $A, B, C$ such that $C$ d-separates $A$ and $B$ in $C$.

* Let $\mathcal{M}_{local}(G) = \{$ distrib. that satisfy local Markov prop$\}$.

$\mathcal{M}_{global}(G) = \{$ " ——————— " global " ——— " $\}$.

**Thm:** If the random vector $X$ has a joint prob. distrib. $P$ that obeys the directed local Markov property for the directed acyclic graph $G$, then $P$ obeys the directed global Markov property for the graph $G$.

Simpler criteria for d-separation:
For a DAG $G$, let $G^m$ denote its moralization:



**Prop:** Let $G$ be a DAG. Then $C$ d-separates $A$ and $B$ in $G$ iff $C$ separates $A$ and $B$ in the moralization

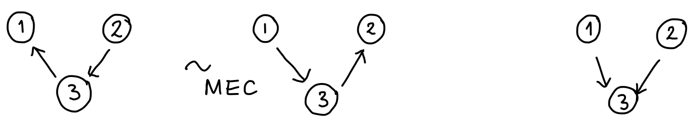
$$\Big(\underbrace{G_{an(A \cup B \cup C)}}_{\text{Induced subgraphs on these nodes}}\Big)^m$$

→ Up to here we have characterized distributions that are Markov to a DAG. What about distributions for different DAGs?

- Let $\mathcal{M}(G)$ be the set of strictly positive densities that are Markov w.r.t. $G$.

- Two DAGs $G_1, G_2$ are Markov equivalent if $\mathcal{M}(G_1) = \mathcal{M}(G_2)$
  MEC := Markov Equivalence Class.

Ex:



$1 \perp\!\!\!\perp 2 \mid 3$ $\qquad$ $1 \perp\!\!\!\perp 2 \mid 3$

- Graphical criteria for Markov Equiv. (Verma & Pearl 91)
  Two DAGs $G_1$ and $G_2$ belong to the same MEC $\Longleftrightarrow$ they have the same skeleta (i.e. underlying und. graph) an v-structures $i \to j \leftarrow k$

| $p$ | 4 | 7 | ? |
|---|---|---|---|
| #MEC | 185 | 312510571 | ? |
| $\frac{\#\text{MEC}}{\#\text{DAGS}}$ | 0.34 | 0.27 | $\approx 0.25$ CONJECTURE as $p \to \infty$ |

Conjecture $\Rightarrow$ a MEC has $\approx 4$ DAGs $\Rightarrow$ few experiments are needed to identify the DAG.

Interventions and their DAGs.

$\to$ Modify the distributions, soft vs. perfect interventions

Let $I \subseteq [p]$ be an intervention target. i.e. nodes to be intervened.

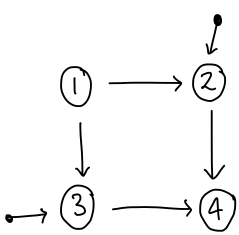
<u>Def:</u> Under a general intervention target $I \subset [p]$ the interventional distribution $f^{(I)}$ can be factorized as

$$f^{(I)}(x) = \prod_{i \in I} f^{(I)}(x_i \mid x_{pa_G(i)}) \prod_{j \notin I} f^{(\emptyset)}(x_j \mid x_{pa_G(j)})$$

where $f^{(I)}$ and $f^{(\emptyset)}$ are the interv. and obsv. distrib. over $X$ respect.

Note $f^{(I)}(x_j \mid x_{pa_G(j)}) = f^{\emptyset}(x_j \mid x_{pa(j)})$ $\forall j \notin I$ i.e. conditional distributions of non-targeted variables are invariant to the interv.

<u>Ex:</u>




$I = \{2, 3\}$

$f(x_1)$, $f(x_4 \mid x_3, x_2)$ are invariant

Let $\{f^{(I)}\}_{I \in \mathcal{I}}$ denote a collection of distr. over $X$ indexed by $\mathcal{I}$.

<u>Def 3.3:</u> For a DAG $G$ and interventional target set $\mathcal{I}$, define

$$\mathcal{M}_{\mathcal{I}}(G) := \Big\{ \{f^{(I)}\}_{I \in \mathcal{I}} \;\Big|\; \forall I \in \mathcal{I}: \; f^{(I)} \in \mathcal{M}(G) \text{ and } f^{(I)}(X_j \mid X_{pa_G(j)}) = f^{\emptyset}(X_j \mid X_{pa_G(j)}) \;\; \forall j \in I \Big\}$$

$\mathcal{M}_{\mathcal{I}}(G)$ is the set of interventional distributions that can be generated from a DAG $G$ by intervening according to $\mathcal{I}$.

<u>Def:</u> ($\mathcal{I}$-Markov equivalence). Two DAGs $G_1$ and $G_2$ for which $\mathcal{M}_{\mathcal{I}}(G_1) = \mathcal{M}_{\mathcal{I}}(G_2)$ belong to the same $\mathcal{I}$-Markov equivalence class ($\mathcal{I}$-MEC).

(1) $\mathcal{I}$-DAGs
(2) Generalize Verma & Pearl.

<u>Def:</u> Let $G$ be a DAG and $\mathcal{I}$ intervention targets.

$G^{\mathcal{I}}$ is the graph $G$ augmented with $\mathcal{I}$-vertices $\{\zeta_I\}_{I \in \mathcal{I}, I \neq \emptyset}$ and $\mathcal{I}$-edges $\{\zeta_I \to i\}_{i \in I \in \mathcal{I}, I \neq \emptyset}$

<u>Example:</u>

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[r] \arrow[dr] & 3 \arrow[d] \\ & & 4 \end{tikzcd} \qquad \mathcal{I} = \{ \underset{\{2,3\}}{\underset{\shortparallel}{I_1}}, \underset{\{4\}}{\underset{\shortparallel}{I_2}} \}.$$

<u>Def:</u> ($\mathcal{I}$-Markov Property). Let $\mathcal{I}$ be intervention targets with $\emptyset \in \mathcal{I}$ and suppose $\{f^{(I)}\}_{I \in \mathcal{I}}$ is a set of (strict. positive) probab. distrib. over $X_1, \dots, X_p$ indexed by $I \in \mathcal{I}$. $\{f^{(I)}\}_{I \in \mathcal{I}}$ satisfies the $\mathcal{I}$-Markov property w.r.t. the $\mathcal{I}$-DAG $G^{\mathcal{I}}$ iff

(1) $f^{I} \in \mathcal{M}(G) \;\; \forall I \in \mathcal{I}$

(2) $f^{(I)}(X_A \mid X_C) = f^{\emptyset}(X_A \mid X_C)$ for any $I \in \mathcal{I}$ and any disjoint $A, C \subset [p]$ s.t. $C \cup \zeta_{\mathcal{I} \setminus I}$ d-separates $A$ and $\zeta_I$ in $G^{\mathcal{I}}$.

$\zeta_{\mathcal{I} \setminus I} = \{ \zeta_J : J \in \mathcal{I}, J \neq I \}$

**Prop:** Suppose $\emptyset \in \mathcal{I}$. Then $\{f^{(I)}\}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(G) \iff \{f^{(I)}\}_{I \in \mathcal{I}}$ satisfies the $\mathcal{I}$-Markov prop. w.r.t $\mathcal{G}^{\mathcal{I}}$.

**Thm:** Suppose $\emptyset \in \mathcal{I}$. Two DAGs $G_1$ and $G_2$ belong to the same $\mathcal{I}$-MEC $\iff$ their $\mathcal{I}$-DAGS $G_1^{\mathcal{I}}$, $G_2^{\mathcal{I}}$ have the same skeleta and v-structures.

**Examples:**

$$1 \to 2 \to 3$$

$$\begin{tikzcd} 1 \arrow[r] \arrow[d] & 2 \arrow[d] \\ 4 \arrow[r] & 3 \end{tikzcd}$$