Bayesian Networks 4
I-Equivalence, Distributions to Graphs

I-Equivalence

Very different BN structures can actually encode the same set of conditional independence assertions eg. the three structures below encode $(X \perp Y \mid Z)$:

\[ X \rightarrow Z \rightarrow Y \]
\[ X \rightarrow Z \leftarrow Y \]

I-Equivalence

I-equivalence of two graphs implies:

- Any distribution $P$ that can be factorized over one of these graphs can be factorized over other
- There is no intrinsic property of $P$ that would allow us to associate it with one graph rather than an equivalent one

I-Equivalence

• Suppose we know that $X$ and $Y$ are correlated in the distribution $P(X, Y)$
• We don’t know if the correct structure is:

\[ X \rightarrow Y \]

Or

\[ X \leftarrow Y \]

This has big implications for inferring causality! We'll cover this later in the course if we have time.

I-Equivalence

Two graph structures $\mathcal{K}_1$ and $\mathcal{K}_2$ over $\mathcal{X}$ are I-equivalent if $I(\mathcal{K}_1) = I(\mathcal{K}_2)$.
I-Equivalence

The skeleton of a Bayesian network graph $G$ over $\mathcal{X}$ is an undirected graph over $\mathcal{X}$ that contains an edge $\{X, Y\}$ for every edge $(X, Y)$ in $G$.

These two BNs have the same skeleton

I-Equivalence

• If two networks have a common skeleton, then the set of trails between two variables is the same in both networks.
• But...having the same trails is not enough for I-equivalence eg.

I-Equivalence

Theorem 3.7: Let $G_1$ and $G_2$ be two graphs over $\mathcal{X}$. If $G_1$ and $G_2$ have the same skeleton and the same set of v-structures then they are I-equivalent.

But there are graphs that are I-equivalent but do not have the same set of v-structures.
• eg. two complete (fully-connected) graphs have the same skeleton but not the same v-structures.

Can we provide a stronger condition that corresponds to I-Equivalence?
I-Equivalences

A v-structure $X \rightarrow Z \leftarrow Y$ is an *immorality* if there is no direct edge between $X$ and $Y$. If there is such an edge, it is called a *covering edge* for the v-structure.

Let $G_1$ and $G_2$ be two graphs over $\mathcal{X}$. Then $G_1$ and $G_2$ have the same skeleton and the same set of immoralities if and only if they are I-equivalent.

Distributions to Graphs

Given a distribution $P$, to what extent can we construct a graph $\mathcal{G}$ whose independencies reflect those of $P$?

One solution: Take any graph that is an I-map for $P$. But a complete graph is an I-map for any distribution $P$.
Distributions to Graphs

How do we obtain a minimal I-map given a set of independencies $I$? See Algorithm 3.2
• First assume you are given predetermined variable ordering $(X_1, \ldots, X_n)$.
• For each $X_i$
  – Pick some minimal subset $U$ of $(X_1, \ldots, X_{i-1})$ to be $X_i$’s parents in $G$.
  – To pick this subset, we require that $U$ satisfy $(X_i \perp \{X_1, \ldots, X_{i-1}\} - U | U)$ and that no node can be removed from $U$ without violating this property.

Distributions to Graphs

If $G$ is a minimal I-map for distribution $P$, can we “read off” all of the independencies in $P$ directly from $G$?
• This is false! (See next slide for counterexample)

Distributions to Graphs

Suppose we have a distribution $P$ whose independencies are reflected in the graph $G$ below that is. $I(P) = I(G)$

We want to construct a minimal I-map for $P$ using three different orderings:
1. D, I, S, G, L
2. L, S, G, I, D
3. L, D, S, I, G

Distributions to Graphs

Ordering: D, I, S, G, L
Ordering: L, S, G, I, D
Ordering: L, D, S, I, G

These are all minimal I-maps but the last two fail to capture some independence relationships. The minimal I-map criterion doesn’t quite cut it. We want a graph $G$ that precisely captures the independencies in a given distribution $P$.2

Distributions to Graphs

Ordering: D, I, S, G, L
Distributions to graphs

We say that a graph $\mathcal{K}$ is a perfect map (P-map) for a set of independencies $I$ if we have that $I(\mathcal{K}) = I$. We say that $\mathcal{K}$ is a perfect map for $P$ if $I(\mathcal{K}) = I(P)$.

Distributions to Graphs

Does every distribution have a perfect map?
No...2 common counterexamples
1. Regularity in the parameterization of the distribution (eg. XOR relationships) that cannot be captured in the graph structure
2. Independence assumptions imposed by the structure of BNs is not appropriate

Counterexample of type 1:

$P(x, y, z) = \begin{cases} 1/12 & x \oplus y \oplus z = \text{false} \\ 1/6 & x \oplus y \oplus z = \text{true} \end{cases}$

$X \perp Y \in I(P)$ but $X \perp Z \mid Y \notin I(P)$ and $Y \perp Z \mid X \notin I(P)$

One possible minimal I-map:

But it is not a perfect map since $(X \perp Z) \in I(P)$ which cannot be "read off" the graph by d-separation

Counterexample of type 2:

Suppose we have $(A \perp C \mid \{B, D\}) \in I(P)$ and $(B \perp D \mid \{A, C\}) \in I(P)$

Can we draw a P-map with just these independencies? No

Can't express these independencies with a BN. You need an undirected graphical model
Distributions to Graphs

- Assuming that $P$ has a perfect map, we can in fact find a perfect map for $P$
- More about this in the structure learning part of the course...