Undirected Graphical Models 2: Independencies

Independencies (Bayesian Networks)

Use d-separation to read off independencies in a Bayesian network

Takes a bit of effort!
Independencies (Markov networks)

Use separation to determine independencies (really easy!)

\[ X \perp Y \mid Z \]

\[ X \quad Z \quad Y \]

Independencies

Formally: let $\mathcal{H}$ be a Markov network structure, and let $X_1 - \ldots - X_k$ be a path in $\mathcal{H}$. Let $Z \subseteq \mathcal{X}$ be a set of observed variables. The path $X_1 - \ldots - X_k$ is active given $Z$ if none of the $X_i$'s in $i=1, \ldots, k$, is in $Z$.

\[ X_1 \quad X_2 \quad X_3 \]

Path not active if $X_2$ is in $Z$ and it separates $X_1$ and $X_3$
Independencies

A set of nodes $Z$ separates $X$ and $Y$ in $\mathcal{H}$, denoted $\text{sep}_{\mathcal{H}}(X; Y | Z)$, if there is no active path between any node $X \in X$ and $Y \in Y$ given $Z$.

We define the global independencies associated with $\mathcal{H}$ to be:

$$I(\mathcal{H}) = \{(X \perp Y | Z) : \text{sep}_{\mathcal{H}}(X; Y | Z)\}$$

Separation is monotonic in $Z$ ie. If $\text{sep}_{\mathcal{H}}(X; Y | Z)$ then $\text{sep}_{\mathcal{H}}(X; Y | Z')$ for any $Z' \supset Z$.

Example:

$$(X_1 \perp X_4 | X_2)$$

Can’t encode non-monotonic independence relations with separation in a Markov network (more on this later)
Independencies

Properties we want separation to have:
1) **Soundness**: i.e. Separation in Graph $\mathcal{H}$
   $\iff$ Independence in distribution $P$
2) **Completeness**: i.e. Separation in Graph $\mathcal{H}$ finds all independences in distribution $P$

Do these properties hold?

**Soundness**

**Soundness**: Separation in Graph $\mathcal{H} \iff$ Independence in distribution $P$

- $\Rightarrow$ direction: true. See Theorem 4.1
- $\Leftarrow$ direction: true*

*true only for positive distributions (i.e. probability of all events $> 0$)

**Hammersley-Clifford Theorem**: Let $P$ be a positive distribution over $\mathcal{X}$, and $\mathcal{H}$ a Markov network graph over $\mathcal{X}$. If $\mathcal{H}$ is an I-map for $P$, then $P$ is a Gibbs distribution that factorizes over $\mathcal{H}$. 
Independencies

Properties we want separation to have:

1) **Soundness**: i.e. Separation in Graph $\mathcal{H}$
   ⇔ Independence in distribution $P^*$

2) **Completeness**: i.e. Separation in Graph $\mathcal{H}$
   finds all independences in distribution $P$

Completeness

- **Strong version (not true)**: every pair of nodes $X$ and $Y$ that are not separated in $\mathcal{H}$ are dependent in every distribution which factorizes over $\mathcal{H}$
- **Weaker version needed**: If $X$ and $Y$ are not separated given $Z$ in $\mathcal{H}$, then $X$ and $Y$ are dependent given $Z$ in some distribution $P$ that factorizes over $\mathcal{H}$. 
Independencies

Properties we want separation to have:

1) **Soundness**: i.e. Separation in Graph $\mathcal{H}$ $\iff$ Independence in distribution $P^*$

   *See fine print

2) **Completeness**: i.e. Separation in Graph $\mathcal{H}$ finds all independences in distribution $P^*$

   *See fine print

Independencies

We had two definitions of independencies in Bayesian networks:

1. **Global independencies**
   - D-separation

2. **Local independencies**:
   - $(X_i \perp \text{NonDescendants}(X_i) \mid \text{Parents}(X_i))$
Independencies

We can do the same thing with Markov Networks:

1. Global independencies: Separation
   \[ I(\mathcal{H}) \]

2. “Local” independencies:
   a) Pairwise independencies \[ I_p(\mathcal{H}) \]
   b) Local independencies (Markov Blanket)
      \[ I_l(\mathcal{H}) \]

Pairwise Independencies

Intuitively: when two variables are not directly connected, we can make them conditionally independent through other mediating variables

Let \( \mathcal{H} \) be a Markov network. We define the pairwise independencies associated with \( \mathcal{H} \) to be:

\[
I_p(\mathcal{H}) = \{(X \perp Y | X - \{X,Y\}); X - Y \notin \mathcal{H}\}
\]
Local Independencies

Markov Blanket
• Intuitively: block all influences on a node by conditioning on its immediate neighbors

Grey nodes are the Markov blanket

Local Independencies

Markov Blanket
• Formally: for a given graph $\mathcal{H}$, we define the Markov blanket of $X$ in $\mathcal{H}$, denoted $MB_{\mathcal{H}}(X)$, to be the neighbors of $X$ in $\mathcal{H}$. We define the local independencies associated with $\mathcal{H}$ to be:

$$I_\mathcal{L}(\mathcal{H}) = \{(X \perp \mathcal{X} - \{X\} \mid MB_{\mathcal{H}}(X)) : X \in \mathcal{X}\}.$$
Independencies

• For general distributions: $I_p(H)$ weaker than $I_i(H)$ which is weaker than $I(H)$
• For positive distributions: All three are equivalent