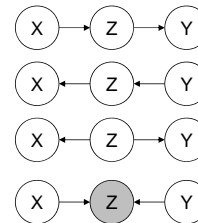


Undirected Graphical Models 2: Independencies

1

Independencies (Bayesian Networks)

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



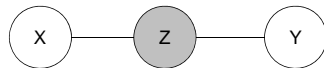
Takes a bit of effort!

2

Independencies (Markov networks)

Use **separation** to determine independencies
(really easy!)

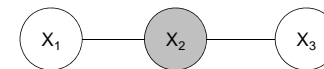
$$X \perp Y \mid Z$$



3

Independencies

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



Path not active if X_2 is in \mathbf{Z} and it separates X_1 and X_3

4

Independencies

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

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

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

5

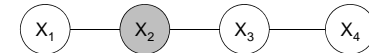
Independencies

Separation is **monotonic** in \mathbf{Z} i.e.

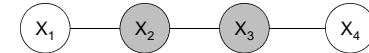
If $\text{sep}_{\mathcal{H}}(\mathbf{X}; \mathbf{Y} \mid \mathbf{Z})$ then $\text{sep}_{\mathcal{H}}(\mathbf{X}; \mathbf{Y} \mid \mathbf{Z}')$ for any $\mathbf{Z}' \supset \mathbf{Z}$.

Example:

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



$$(X_1 \perp X_4 \mid \{X_2, X_3\})$$



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

6

Independencies

Properties we want separation to have:

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

Do these properties hold?

7

Soundness

Soundness: Separation in Graph $\mathcal{H} \Leftrightarrow$ 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} .

8

Independencies

Properties we want separation to have:

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

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

9

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} .

10

Independencies

Properties we want separation to have:

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

*See fine print

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

*See fine print

11

Independencies

We had two definitions of independencies in Bayesian networks:

1. **Global independencies**

D-separation

2. **Local independencies**:

$(X_i \perp \text{NonDescendants}(X_i) \mid \text{Parents}(X_i))$

12

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) $[I_l(\mathcal{H})]$

13

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:

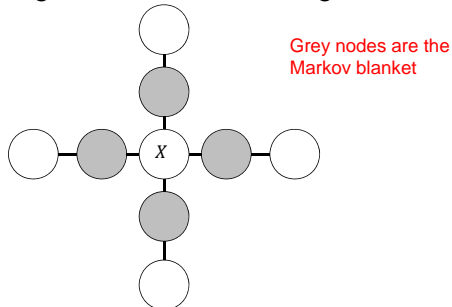
$$I_p(\mathcal{H}) = \{(X \perp Y \mid \mathcal{X} - \{X, Y\}) : X - Y \notin \mathcal{H}\}$$

14

Local Independencies

Markov Blanket

- Intuitively: block all influences on a node by conditioning on its immediate neighbors



15

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:

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

16

Independencies

- For general distributions: $I_\rho(\mathcal{H})$ weaker than $I_l(\mathcal{H})$ which is weaker than $I(\mathcal{H})$
- For **positive** distributions: All three are equivalent

17