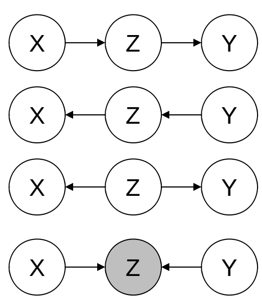
Undirected Graphical Models 2: Independencies

1

Independencies (Bayesian Networks)

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



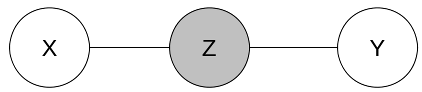
Takes a bit of effort!

2

Independencies (Markov networks)

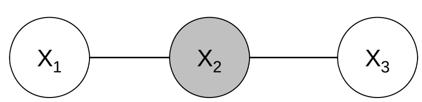
Use separation to determine independencies (really easy!)

$X \perp Y \mid Z$



Independencies

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



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

Independencies

A set of nodes ***Z*** separates ***X*** and ***Y*** in $\mathcal{H}$, denoted 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:

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

5

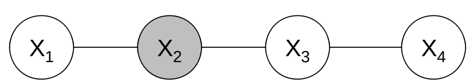
Independencies

Separation is monotonic in ***Z*** ie.

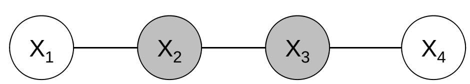
If *sep*$_{\mathcal{H}}$(***X***; ***Y*** | ***Z***) then *sep*$_{\mathcal{H}}$(***X***; ***Y*** | ***Z'***) for any ***Z'*** $\supset$ ***Z***.

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?

Soundness

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

- $\Longrightarrow$ direction: true. See Theorem 4.1
- $\Longleftarrow$ 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 $\mathcal{H}$ $\Leftrightarrow$ 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 $\boldsymbol{Z}$ in $\mathcal{H}$, then X and Y are dependent given $\boldsymbol{Z}$ in some distribution P that factorizes over $\mathcal{H}$.

Independencies

Properties we want separation to have:

1) Soundness: i.e. Separation in Graph $\mathcal{H}$ $\Leftrightarrow$ 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

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 $[\mathcal{I}(\mathcal{H})]$
2. "Local" independencies:
   a) Pairwise independencies $[\mathcal{I}_p(\mathcal{H})]$
   b) Local independencies (Markov Blanket) $[\mathcal{I}_\ell(\mathcal{H})]$

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:

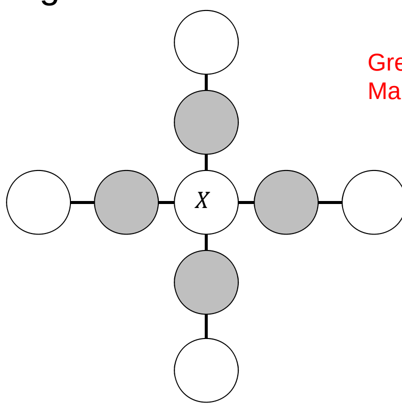
$\mathcal{I}_p(\mathcal{H}) = \{(X \perp Y \mid \mathcal{X} - \{X,Y\}) : X—Y \notin \mathcal{H}\}$

14

Local Independencies

Markov Blanket

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



Grey nodes are the Markov blanket

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:

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

16

Independencies

- For general distributions: $\mathcal{I}_p(\mathcal{H})$ weaker than $\mathcal{I}_l(\mathcal{H})$ which is weaker than $\mathcal{I}(\mathcal{H})$
- For positive distributions: All three are equivalent

17