BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Probabilistic Graphical Model

Session # 7 : Undirected Graphical Model

Seetha Parameswaran
seetha.p@pilani.bits-pilani.ac.in

The instructor is gratefully acknowledging the authors who made their course materials freely available online.

Table of Contents

Active Trail in Markov Network

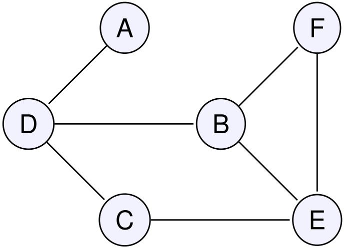
Definition

Let $\mathcal{H}$ be a Markov network structure and let $X_1 -, \ldots, -X_n$ be a path in $\mathcal{H}$.
Let $Z \subseteq X$ be a set of observed variables.
The path $X_1 -, \ldots, -X_n$ is **active** given Z if none of the X_i is in Z.

- Influence has to flow through unobserved variables along the trail.
- Once a variable is observed along the trail, the influence is blocked.

Markov Network - Example

- Find the active trails given B is observed.



$A - D - C - E - F$

Separation in Markov Network

Definition

A set of nodes Z separates X and Y in $\mathcal{H}$, a Markov network structure, if there is no active path between any node in X and Y given Z.

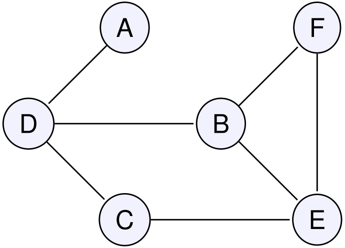
- Denote Separation as $sep_{\mathcal{H}}(X; Y|Z)$
- Global Independencies

$$\mathcal{I}_g(\mathcal{H}) = \{(X \perp Y|Z) : sep_H(X; Y|Z)\} \tag{1}$$

- The Independencies in $\mathcal{I}(\mathcal{H})$ are precisely those that are guaranteed to hold for every distribution P over $\mathcal{H}$.

Markov Network - Example

- Find the global Independencies.



$$\mathcal{I}_g(\mathcal{H}) = \{(A \perp B|D) : sep_H(A; B|D)\}$$
$$= \{(A \perp C|D) : sep_H(A; C|D)\}$$
$$= \{(A \perp E|D) : sep_H(A; E|D)\}$$
$$= \{(A \perp F|D) : sep_H(A; F|D)\}$$

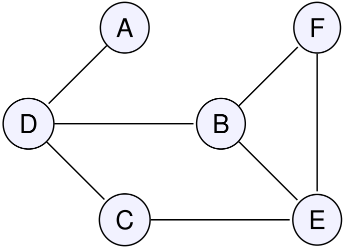
Factorization implies Independence

Theorem

Let P be a distribution over $\mathcal{X}$ and $\mathcal{H}$ a Markov Network structure over $\mathcal{X}$.
*If P is a Gibbs distribution that factorizes over $\mathcal{H}$ and $\mathcal{H}$ is an **I-map** for P.*

Markov Network - Example

- Identify a possible factorization for the Markov Network.



$$P_G = \phi_1(A, D)\phi_2(B, E, F)$$
$$\phi_3(D, B)\phi_4(D, C)\phi_5(C, E)$$

Independence implies Factorization

Theorem

Hammersley-Clifford theorem:

Let P be a ***positive*** *distribution over $\mathcal{X}$ and $\mathcal{H}$ a Markov Network structure over $\mathcal{X}$.*

If $\mathcal{H}$ is an ***I-map*** *for P, then P is a Gibbs distribution that factorizes over $\mathcal{H}$.*

Pairwise Independencies

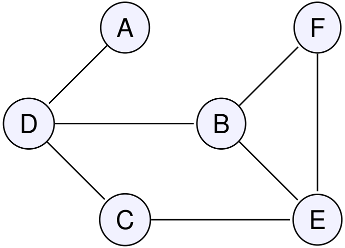
Definition

Let $\mathcal{H}$ be a Markov network structure. Pairwise Independencies associated with $\mathcal{H}$ is defined as

$$\mathcal{I}_p(\mathcal{H}) = \{(X \perp Y | \mathcal{X} - \{X, Y\}) : edge(X - Y) \notin \mathcal{H}\} \tag{2}$$

Markov Network - Example

- Find the pairwise independencies.



For node A

$$\begin{aligned}
\mathcal{I}_p(\mathcal{H}) &= (A \perp C | D, B, E, F : edge(A - C) \notin \mathcal{H}) \\
&= (A \perp B | D, C, E, F : edge(A - B) \notin \mathcal{H}) \\
&= (A \perp E | D, B, C, F : edge(A - E) \notin \mathcal{H}) \\
&= (A \perp F | D, B, E, C : edge(A - F) \notin \mathcal{H})
\end{aligned}$$

Markov Blanket

Definition

For a given graph $\mathcal{H}$, the Markov blanket of X in $\mathcal{H}$ is defined as neighbours of X in $\mathcal{H}$.

$$MB_{\mathcal{X}} = \{Pa(X), Ch(X), Pa(Ch(X))\} \tag{3}$$

- Markov blanket is the set of nodes containing parents, children, and children's parents.

Local Independencies

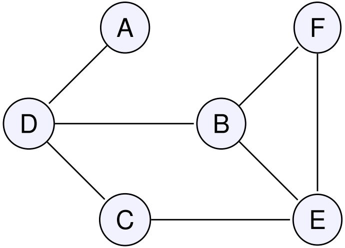
Definition

Let $\mathcal{H}$ be a Markov network structure. Local Independencies associated with $\mathcal{H}$ is defined as

$$\mathcal{I}_l(\mathcal{H}) = \{(X \perp \mathcal{X} - \{X + MB_{\mathcal{H}}(X)\} | MB_{\mathcal{H}}(X)) : X \in \mathcal{X}\} \tag{4}$$

Markov Network - Example

- Find the local independencies and Markov blanket.



For node A

$$MB_A = \{A, D, B, C\}$$
$$\mathcal{I}_l(\mathcal{H}) = (A \perp E, F | MB_A)$$

Independencies

Definition

For any Markov network $\mathcal{H}$ and any distribution P,

$$\text{if} \quad P \models \mathcal{I}_l(\mathcal{H}) \quad \text{then} \quad P \models \mathcal{I}_p(\mathcal{H}) \tag{5}$$
$$\text{if} \quad P \models \mathcal{I}_g(\mathcal{H}) \quad \text{then} \quad P \models \mathcal{I}_l(\mathcal{H}) \tag{6}$$

- $\mathcal{I}_p(\mathcal{H})$ is strictly weaker that $\mathcal{I}_l(\mathcal{H})$ which is strictly weaker that $\mathcal{I}_g(\mathcal{H})$
- For a positive distribution P,

$$P \models \mathcal{I}_p(\mathcal{H})$$
$$P \models \mathcal{I}_l(\mathcal{H})$$
$$P \models \mathcal{I}_g(\mathcal{H})$$

Constructing Graphs from Distributions

- A fully connected graph has no independence conditions and, hence, it can be an I-Map of any probability distribution.
- To encode the Independencies in a given distribution P using a graph structure, use minimal I-map.
- Two approaches for constructing a minimal I-map
  1. using the pairwise Markov Independencies.
  2. using the local Independencies.

Constructing Graphs from Distributions

Pairwise Markov independencies

- Let P be a positive distribution.
- Let $\mathcal{H}$ be defined by introducing an edge $\{X, Y\}$.
- If the edge $\{X, Y\}$ is not in $\mathcal{H}$, then X and Y must be independent given all other nodes in the graph.
- To guarantee that $\mathcal{H}$ is an I-map, add direct edges between all pairs of nodes X and Y such that

$$P \not\models (X \perp Y | \mathcal{X} - \{X, Y\})$$

- Then Markov network $\mathcal{H}$ is the unique minimal I-map for $\mathcal{H}$.
- To guarantee that $\mathcal{H}$ is an I-map, add direct edges between all pairs of nodes X and Y, such that they are dependent even on observing all the other variables in the network.

Constructing Graphs from Distributions

Local Independencies

- Let P be a positive distribution.
- For each variable X, define the neighbors of X to be a minimal set of nodes Y that render X independent of the rest of the nodes. i.e. Markov Blanket of X.
- A set U is a Markov blanket of X in a distribution P if $X \notin U$ and if U is a minimal set of nodes such that

$$(X \perp \mathcal{X} - \{X + U\} | U) \in \mathcal{I}(P)$$

- Then define a graph $\mathcal{H}$ by introducing an edge $\{X, Y\}$ for all X and all $Y \in MB_P(X)$
- Then Markov network $\mathcal{H}$ is the unique minimal I-map for $\mathcal{H}$.
- For each variable X, find the minimal set of nodes. Observing these makes the variable independent of all the variables. Then, add an edge between the variable and all the nodes in the set.

Table of Contents

Bayesian Network and Markov Network

Both

- Parametrize a probability distribution using a graphical model.
- Encode the Independencies among the random variable.

Convert Bayesian Network to Markov Network

Two perspectives

1. Parameterization perspective – represent the probability distribution of the Bayesian model using a fully parameterized Markov model.
2. Independencies perspective – represent the independence constraints encoded by the Bayesian model using the Markov model.

Convert Bayesian Network to Markov Network

Parameterization perspective

- Probability distribution P_B, $\mathcal{B}$ is a parameterized Bayesian network over a graph $\mathcal{G}$.
- The parameterization of the Bayesian network $\mathcal{B}$, can also be viewed as a parameterization of a Gibbs distribution.
- Each CPD $P(X_i | Pa_{X_i})$ is a factor with scope $\{X_i, Pa_{X_i}\}$.
- This set of factors defines a Gibbs distribution with the partition function equal to 1.

Convert Bayesian Network to Markov Network

Independencies perspective

1. Replace all the directed edges between the nodes with undirected edges.
2. Add additional undirected edges between nodes that are parents of the node.

- This new structure is called moral graph of Bayesian network.
- $\mathcal{M}[\mathcal{G}]$ is an I-Map for distribution P_B,
- Moral graph $\mathcal{M}[\mathcal{G}]$ of a Bayesian model G loses some information regarding the Independencies.
- $\mathcal{M}[\mathcal{G}]$ is a minimal I-Map for $\mathcal{G}$.

$$\mathcal{I}(\mathcal{M}[\mathcal{G}]) \subseteq \mathcal{I}(\mathcal{G})$$

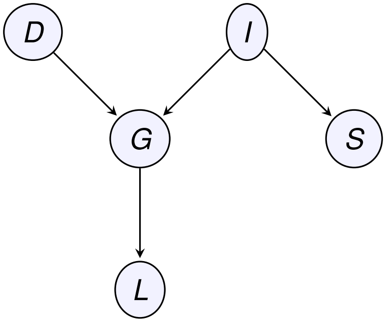
Moral Graph

Definition

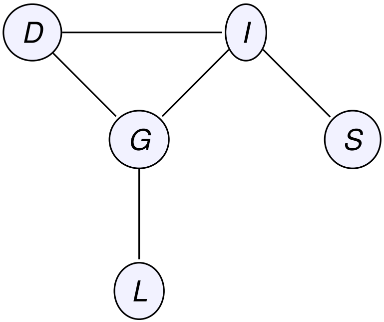
The moral graph $\mathcal{M}[\mathcal{G}]$ of a Bayesian network structure $\mathcal{G}$ over $\mathcal{X}$ is the undirected graph over $\mathcal{X}$ that contains an undirected edge between X and Y if:

- A) there is a directed edge between them (in either direction) or
- B) X and Y are both parents of the same node.

Convert Bayesian Network to Markov Network



Bayesian Network



Markov Network

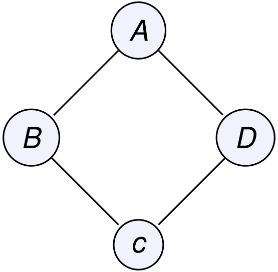
Convert Markov Network to Bayesian Network

Independencies perspective

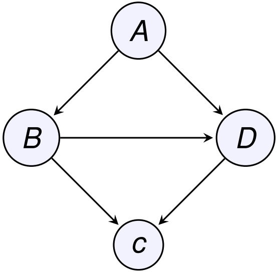
1. Replace all the undirected edges between the nodes with directed edges.
2. Partition all loops into triangles. Add edges to the network to make it chordal.

- Any Bayesian network I-map for the given Markov network must add triangulating edges into the graph, so that the resulting graph is chordal. This process is called triangulation.
- A triangulated or chordal graph is a graph in which each of its cycles of four or more vertices has a chord.
- By simply converting edges of a non-triangulated graph into directed ones, introduces immoralities. An immorality is a v-structure $X \rightarrow Z \leftarrow Y$, if there is no directed edge between X and Y.

Convert Markov Network to Bayesian Network



Markov Network



Bayesian Network

QUESTIONS

1. Given a Markov Network, find the appropriate factorization of joint distribution.
2. Given a Markov Network, identify the active trails.
3. Given a Markov Network, identify the I-maps.
4. Given a Markov Network, identify the d-separations.
5. Given a toy application, generate Markov Network and the factors associated with it.
6. Given factors, generate a Markov Network.
7. Given a joint distribution in the factorized form, generate a Markov Network.
8. Given a Markov Network, identify the conditional Independencies.

References

1. Probabilistic Graphical Models: Principles and Techniques by Daphne Koller and Nir Friedman. MIT Press. 2009
2. Artificial Intelligence: A Modern Approach (3rd Edition) by Stuart Russell, Peter Norvig
3. Mastering Probabilistic Graphical Models using Python by Ankur Ankan, Abhinash Panda. Packt Publishing 2015.
4. Learning in Graphical Models by Michael I. Jordan. MIT Press. 1999

Thank You !!!