OSE SEMINAR 2011

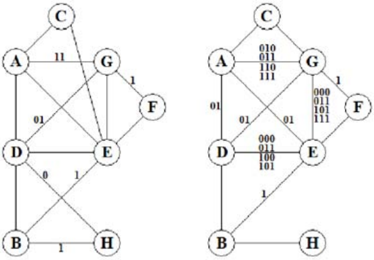
Labeled Graphical Models

Henrik Nyman

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
ÅBO AKADEMI UNIVERSITY

ÅBO, DECEMBER 8 2011

Introduction

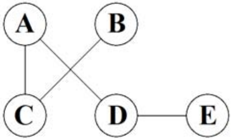


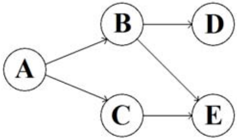
- Collaboration with Jukka Corander and Johan Pensar.
- Graphical representation of conditional and context specific independencies in probability distributions.

Graphical Models

Graphical models are used to give a graphical representation of the dependency structure in a probability distribution.

- Markov Networks - Undirected graphs.
- Bayesian Networks - Directed acyclic graphs.





- Can represent different dependency structures.
- Not all probability distributions can be represented by Bayesian Networks or Markov Networks.

Labeled Graphical Models

- Consists of an undirected graph whose edges might contain labels.
- Markov networks can represent conditional independencies.

$$A \perp\!\!\!\perp B \mid C$$

- Labeled graphical models can in addition to conditional independencies also represent context specific independencies.

$$A \perp\!\!\!\perp B \mid C = 1$$

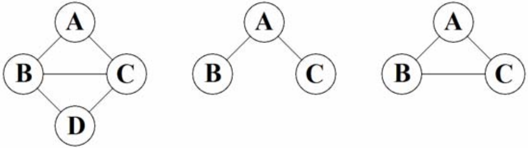
Conditional Independencies in Markov Networks

In Markov networks two variables A and B are conditionally independent given a set of variables S if

1. There is no edge between A and B.
2. There is no path between A and B, $(A, X_1, \ldots, X_n, B)$, such that

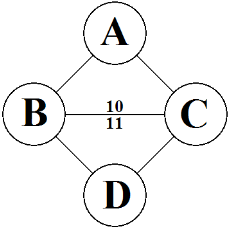
$$(\cup_{i=1}^{n} X_i) \cap S = \emptyset$$

It is then said that S separates A and B.



Context Specific Independencies in Labeled Graphical Models

- LGMs represent conditional independencies in the same way as Markov networks.
- Labels on the edges represent context specific independencies.



- The length of the label on the edge (A, B) is the number of nodes adjacent to both A and B.
- Ordering in the labels according to the assigned topological order.

Classic Example

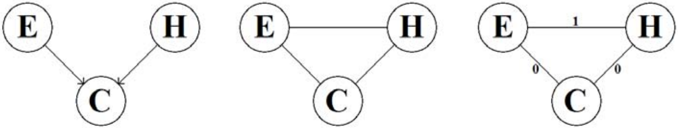
E - Electricity.
H - Hardware functional.
C - Computer functional.
How is the dependency structure represented in this case by

- Bayesian networks?
- Markov networks?
- Labeled graphical models?

Classic Example

E - Electricity.
H - Hardware functional.
C - Computer functional.
How is the dependency structure represented in this case by

- Bayesian networks?
- Markov networks?
- Labeled graphical models?



Objectives

Given a dataset we want to be able to:

- Grade and compare LGMs.
- Find the optimal LGM describing the data.

Cliques and Separators in Undirected Graphs

Definition

A clique in an undirected graph is a maximal set of nodes with edges between each pair of nodes.

Definition

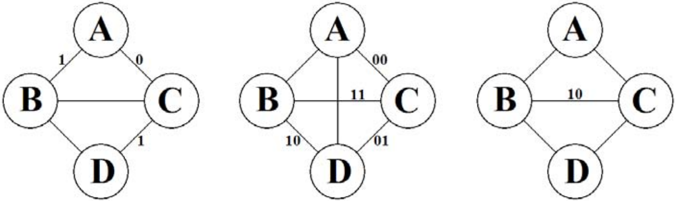
A separator in an undirected graph is the set of nodes that a pair of cliques have in common.

Analytic LGM

Definition

Let $C(G_L)$ and $S(G_L)$ denote the cliques and separators of the labeled graphical model G_L. We define G_L as analytic iff

1. No labels are placed on edges in any separator.
2. For every clique all labels are on edges with one node in common.



Marginal Likelihood

Given an analytic LGM G_L it is possible to calculate the marginal likelihood of a dataset $\mathbf{X}$ according to

Marginal Likelihood

Given an analytic LGM G_L it is possible to calculate the marginal likelihood of a dataset $\mathbf{X}$ according to

$$P(\mathbf{X}|G_L) = \frac{\prod_{c \in C(G)} \prod_{j=1}^{d_c} \prod_{l=1}^{q_j} \frac{\Gamma(\sum_{i=1}^{k_j} a_{jil})}{\Gamma(n(\pi_j^l) + \sum_{i=1}^{k_j} a_{jil})} \prod_{i=1}^{k_j} \frac{\Gamma(n(x_j^i|\pi_j^l) + a_{jil})}{\Gamma(a_{jil})}}{\prod_{s \in S(G)} \frac{\Gamma(a_s)}{\Gamma(t + a_s)} \prod_{i=1}^{k} \frac{\Gamma(n_s^{(i)} + a_{s,i})}{\Gamma(a_{s,i})}}$$

Score Function

- Given a non-analytic LGM it is not possible to calculate the marginal likelihood of a dataset.
- For non-analytic LGMs we need to use a score function.

$$S(G_L|\mathrm{X}) = \frac{\prod_{i=1}^{k}(\theta_i^*)^{n_i}}{n^{0.5\cdot\dim(\Theta^*|G_L)}}$$

Projections

- θ^* is the MLE in the parameter space that satisfies all the conditional independencies induced by G_L.
- θ^* is attained by projecting the original MLE into the parameter space induced by G_L.
- This projection is done by sequentially projecting according to each one of the context specific independence statements in G_L.

Search Method

- Complete search not viable.
- Implement a Markov Chain Monte Carlo (MCMC) method.
- Non-reversible Metropolis-Hastings algorithm.

Size of Model Space

Nodes	UG	LGM
1	1	1
2	2	2
3	8	71
4	64	16777531
5	1024	$> 10^{24}$
6	32768	$> 10^{72}$
7	2097152	$> 10^{202}$

Size of Model Space

Nodes	UG	LGM	Unique LGM
1	1	1	1
2	2	2	2
3	8	71	<32
4	64	16777531	<4096
5	1024	$>10^{24}$	<134217728
6	32768	$>10^{72}$	$<3\cdot10^{17}$
7	2097152	$>10^{202}$	$<3\cdot10^{36}$

- Number of atoms in observable universe $\approx 10^{80}$.

References

J. Corander.
Labelled graphical models.
Scand J Stat, 30:493–508, 2003.

J. Corander, M. Ekdahl, and T. Koski.
Parallell interacting mcmc for learning of topologies of graphical models.
Data Min Knowl Disc, 17:431–456, 2008.

T. Koski and J. M. Noble.
Bayesian Networks : An Introduction.
Wiley, 2009.

H. Nyman and J. Pensar.
Calculating the marginal likelihood of a dataset given an lgm using inverse projections.
Article in the making, 2011.

J. Whittaker.
Graphical Models in Applied Multivariate Statistics.
Wiley, New York, 1990.

The end of the presentation

Thank you for listening!

The end of the presentation

Thank you for listening!

Questions?