Junction Trees and Chordal Graphs

Steffen Lauritzen, University of Oxford

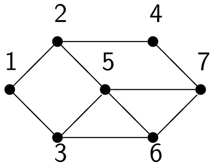
Graphical Models, Lecture 6, Michaelmas Term 2009

October 30, 2009

Consider an *undirected* graph $\mathcal{G} = (V, E)$. A partitioning of V into a triple (A, B, S) of subsets of V forms a *decomposition* of $\mathcal{G}$ if

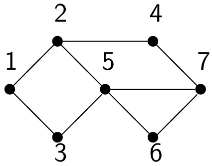
$$A \perp_{\mathcal{G}} B \mid S \text{ and } S \text{ is complete.}$$

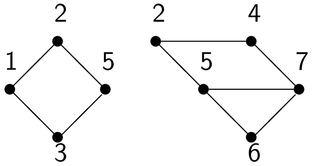
The decomposition is *proper* if $A \neq \emptyset$ and $B \neq \emptyset$.
The *components* of $\mathcal{G}$ are the induced subgraphs $\mathcal{G}_{A \cup S}$ and $\mathcal{G}_{B \cup S}$.
A graph is *prime* if no proper decomposition exists.



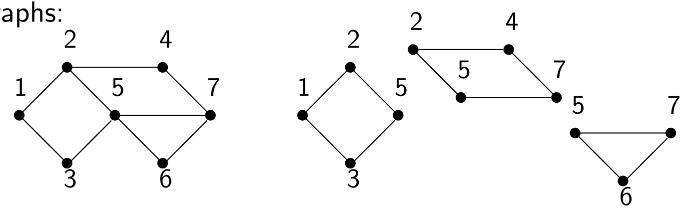
The graph to the left is prime

Decomposition with $A = \{1,3\}$, $B = \{4,6,7\}$ and $S = \{2,5\}$





Any graph can be recursively decomposed into its maximal prime subgraphs:



A graph is *decomposable* (or rather fully decomposable) if it is complete or admits a proper decomposition into *decomposable* subgraphs.

Definition is recursive. Alternatively this means that *all maximal prime subgraphs are cliques.*

Recursive decomposition of a decomposable graph into cliques yields the formula:

$$f(x)\prod_{S\in\mathcal{S}} f_S(x_S)^{\nu(S)} = \prod_{C\in\mathcal{C}} f_C(x_C).$$

Here $\mathcal{S}$ is the set of *minimal complete separators* occurring in the decomposition process and $\nu(S)$ the number of times such a separator appears in this process.

As we have a particularly simple factorization of the density, we have a similar factorization of the maximum likelihood estimate for a decomposable log-linear model.

The MLE for p under the log-linear model with generating class $\mathcal{A} = \mathcal{C}(\mathcal{G})$ for a chordal graph $\mathcal{G}$ is

$$\hat{p}(x) = \frac{\prod_{C \in \mathcal{C}} n(x_C)}{n \prod_{S \in \mathcal{S}} n(x_S)^{\nu(S)}}$$

where $\nu(S)$ is the number of times S appears as a separator in the total decomposition of its dependence graph.

The following are equivalent for any undirected graph $\mathcal{G}$.

(i) *$\mathcal{G}$ is chordal;*

(ii) *$\mathcal{G}$ is decomposable;*

(iii) *All maximal prime subgraphs of $\mathcal{G}$ are cliques;*

(iv) *$\mathcal{G}$ admits a perfect numbering;*

(v) *Every minimal (α, β)-separator are complete.*

Trees are chordal graphs and thus decomposable.

This simple algorithm has complexity $O(|V| + |E|)$:

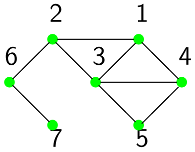
1. Choose $v_0 \in V$ arbitrary and let $v_0 = 1$;
2. When vertices $\{1, 2, \ldots, j\}$ have been identified, choose $v = j + 1$ among $V \setminus \{1, 2, \ldots, j\}$ with highest cardinality of its numbered neighbours;
3. *If* $\mathrm{bd}(j + 1) \cap \{1, 2, \ldots, j\}$ *is not complete,* $\mathcal{G}$ is not chordal;
4. Repeat from 2;
5. *If the algorithm continues until only one vertex is left, the graph is chordal and the numbering is perfect.*

Finding the cliques of a chordal graph

From an MCS numbering $V = \{1, \ldots, |V|\}$, let

$$B_\lambda = \mathrm{bd}(\lambda) \cap \{1, \ldots, \lambda - 1\}$$

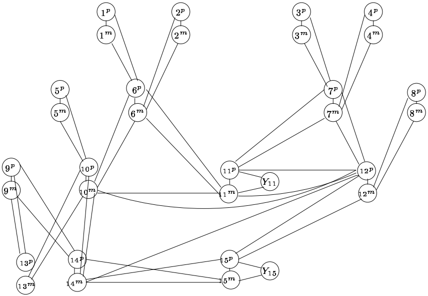
and $\pi_\lambda = |B_\lambda|$. Call λ a *ladder vertex* if $\lambda = |V|$ or if $\pi_{\lambda+1} < \pi_\lambda + 1$. Let Λ be the set of ladder vertices.



π_λ: 0,1,2,2,2,1,1.

The cliques are $C_\lambda = \{\lambda\} \cup B_\lambda, \lambda \in \Lambda$.

A chordal graph



This graph is chordal, but it might not be that easy to see... Maximum Cardinality Search is handy!

Let $\mathcal{A}$ be a collection of finite subsets of a set V. A *junction tree* $\mathcal{T}$ of sets in $\mathcal{A}$ is an undirected tree with $\mathcal{A}$ as a vertex set, satisfying the *junction tree property:*

> *If $A, B \in \mathcal{A}$ and C is on the unique path in $\mathcal{T}$ between A and B it holds that $A \cap B \subset C$.*

If the sets in an arbitrary $\mathcal{A}$ are pairwise incomparable, *they can be arranged in a junction tree if and only if $\mathcal{A} = \mathcal{C}$ where $\mathcal{C}$ are the cliques of a chordal graph*

The following are equivalent for any undirected graph $\mathcal{G}$.

(i) *$\mathcal{G}$ is chordal;*
(ii) *$\mathcal{G}$ is decomposable;*
(iii) *All prime components of $\mathcal{G}$ are cliques;*
(iv) *$\mathcal{G}$ admits a perfect numbering;*
(v) *Every minimal (α, β)-separator are complete.*
(vi) *The cliques of $\mathcal{G}$ can be arranged in a junction tree.*

The junction tree can be *constructed directly from the MCS ordering $C_\lambda, \lambda \in \Lambda$*, where C_λ are the cliques: Since the MCS-numbering is perfect, $C_\lambda, \lambda > \lambda_{\min}$ all satisfy

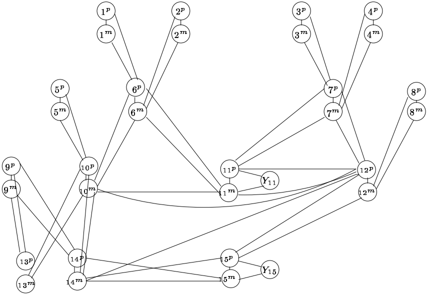
$$C_\lambda \cap (\cup_{\lambda' < \lambda} C_{\lambda'}) = C_\lambda \cap C_{\lambda^*} = S_\lambda$$

for some $\lambda^* < \lambda$.

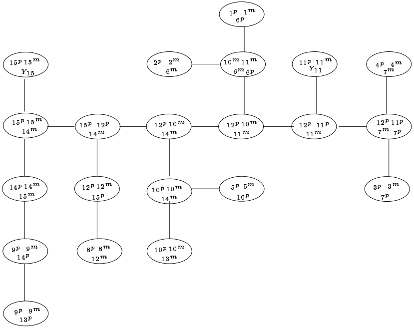
A junction tree is now easily constructed by attaching C_λ to any C_{λ^*} satisfying the above. Although λ^* may not be uniquely determined, S_λ is.

Indeed, the sets S_λ are the minimal complete separators and *the numbers $\nu(S)$ are $\nu(S) = |\{\lambda \in \Lambda : S_\lambda = S\}|$.*

A chordal graph



Junction tree



Cliques of graph arranged into a tree with $C_1 \cap C_2 \subseteq D$ for all cliques D on path between C_1 and C_2.

In general, the *prime components* of any undirected graph can be arranged in a junction tree in a similar way.

Then *every pair of neighbours (C, D) in the junction tree represents a decomposition of $\mathcal{G}$* into $\mathcal{G}_{\tilde{C}}$ and $\mathcal{G}_{\tilde{D}}$, where $\tilde{C}$ is the set of vertices in cliques connected to C but separated from D in the junction tree, and similarly with $\tilde{D}$.

The corresponding algorithm is based on a slightly more sophisticated algorithm known as *Lexicographic Search* (LEX) which runs in $O(|V|^2)$ time.

The MLE for p under a conformal log-linear model with generating class $\mathcal{A} = \mathcal{C}(\mathcal{G})$ for a chordal graph $\mathcal{G}$ is

$$\hat{p}(x) = \frac{\prod_{Q \in \mathcal{Q}} \hat{p}_Q(x_Q)}{\prod_{S \in \mathcal{S}} \{n(x_S)/n\}^{\nu(S)}}$$

where $\hat{p}_Q(x_Q)$ is the estimate of the marginal distrbution based on data from Q only and $\nu(S)$ is the number of times S appears as a separator in the decomposition of its dependence graph into prime components.

When the prime components are cliques it further holds that $\hat{p}_C(x_C) = n(x_C)/n$.

Local computation algorithms have been developed with a variety of purposes. For example:

- Kalman filter and smoother
- Solving sparse linear equations;
- Decoding digital signals;
- Estimation in hidden Markov models;
- Peeling in pedigrees;
- Belief function evaluation;
- Probability propagation.

Also dynamic programming, linear programming, optimizing decisions, calculating Nash equilibria in cooperative games, and many others. *List is far from exhaustive!*

All algorithms are using, explicitly or implicitly, a *graph decomposition* and *a junction tree* or similar to make the computations.