# GRAPHICAL MODELS

Piotr GRACZYK

## 6. DECOMPOSABLE GRAPHS
## ( triangulated graphs, chordal graphs)

A part of this chapter is based on lectures of Prof. S. Lauritzen at CIMPA Summer School Hammamet 2011, with his kind permission.

Only graphical models governed by **DECOMPOSABLE GRAPHS** have good statistical properties:

• one can compute easily MLE estimators $\widehat{K}$ and $\widehat{\Sigma}$ of the precision and covariance matrices

• statistical tests can be performed

• Bayesian statistics is possible and performant

That's why we shall learn some theory of DECOMPOSABLE GRAPHS

Consider an undirected graph $\mathcal{G} = (V, E)$ with vertices $V$ and edges $E$.

If $W \subset V$, the **induced graph** is $\mathcal{G}_W = (W, E_W)$ where $\{i, j\} \in E_W$ if and only if $\{i, j\} \in E$ and $i, j \in W$. The edges of the **induced graph** $\mathcal{G}_W$ are all the edges of $\mathcal{G}$ connecting vertices from $W$.

A **path** of length $n$ from $\alpha \in V$ to $\beta \in V$ is a sequence

$$\alpha_0 = \alpha, \alpha_1, \ldots, \alpha_n = \beta$$

of vertices distinct for $i = 0, \ldots, n-1$ such that $\{\alpha_i, \alpha_{i+1}\} \in E$ for each $i = 1, \ldots, n$.

A subset $S \subset V$ is an $(\alpha, \beta)$-**separator** if every path from $\alpha$ to $\beta$ intersects $S$.
$S$ separates $A \subset V$ from $B \subset V$ if $S$ is an $(\alpha, \beta)$-**separator** for every $\alpha \in A$ and $\beta \in B$.
A separator of $A$ and $B$ is **minimal** if no proper subset $T \subsetneq S$ separates $A$ and $B$.

A graph is **complete** if all vertices are joined by an edge. A subset $W$ is complete if its induced graph $\mathcal{G}_W$ is complete.

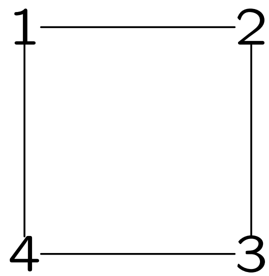
A **clique** of $\mathcal{G}$ is a **maximal complete** subset of $V$.

A **cycle** of length $n$ is a **path** of length $n$ from $\alpha$ to $\alpha$. The shortest cycles are triangles=cycles of length 3.

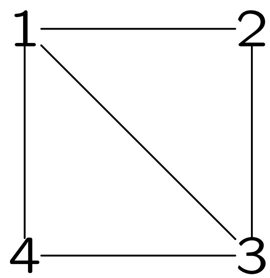
A **tree** is a connected graph without cycles. It has a unique path between any two vertices.

A graph is **triangulated(chordal)** if **every cycle of length** $n \geq 4$ has a **chord**, that is two non-consecutive vertices that are connected by an edge(chord).

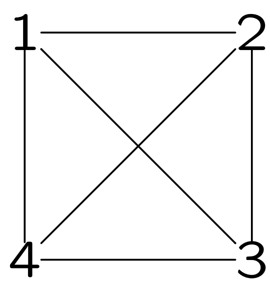
**Examples.**

The graph 

is the smallest non-chordal graph.

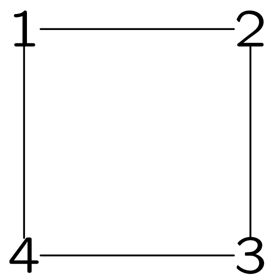
The graph 

is chordal and non-complete.

The graph 

is complete $\Rightarrow$ chordal.

In the graph

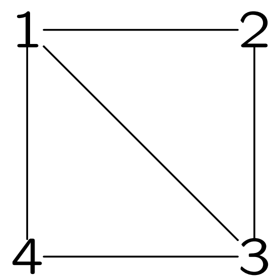



the set $S = \{1, 3\}$ is a $(2, 4)-$separator.
The separator $S$ is minimal. $S$ is not complete.

the set $S' = \{2, 4\}$ is a $(1, 3)-$separator.
The separator $S'$ is minimal. $S'$ is not complete.

There are no other separators.
No separator is complete.

The cliques are $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$ and $\{1, 4\}$

In the graph 


the set $S = \{1, 3\}$ is a $(2, 4)$−separator. $S$ is minimal and complete. There are no other separators.
( the set $S' = \{2, 4\}$ is NOT a $(1, 3)$−separator)

Every minimal separator is complete.

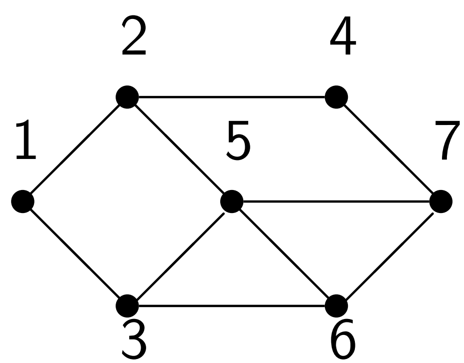
The cliques are $\{1, 2, 3\}$ and $\{1, 3, 4\}$.

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

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

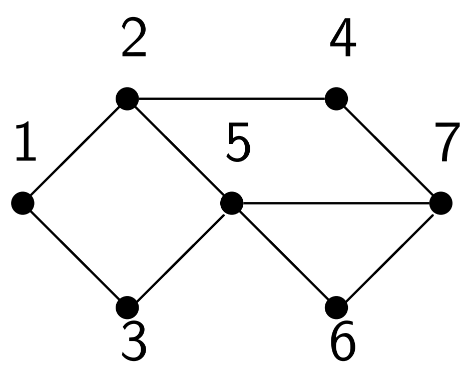
The decomposition is *proper* if $A \neq \emptyset$ and $B \neq \emptyset$.
The *components* of $\mathcal{G}$ are the induced subgraphs $\mathcal{G}_{A \cup S}$ and $\mathcal{G}_{B \cup S}$.
A graph is *prime* if no proper decomposition exists.

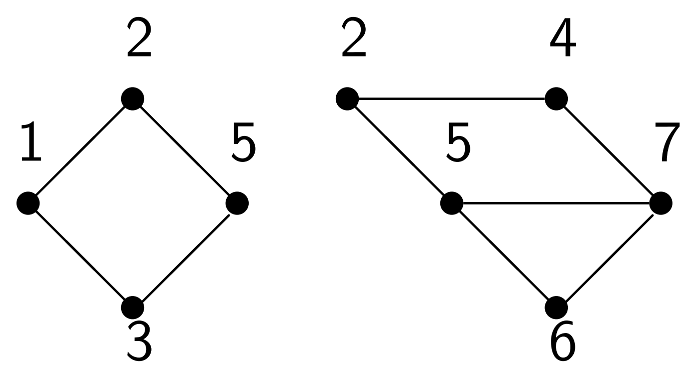
# Examples




The graph to the left is prime

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



Suppose $P$ satisfies (F) w.r.t. $\mathcal{G}$ and $(A, B, S)$ is a decomposition. Then

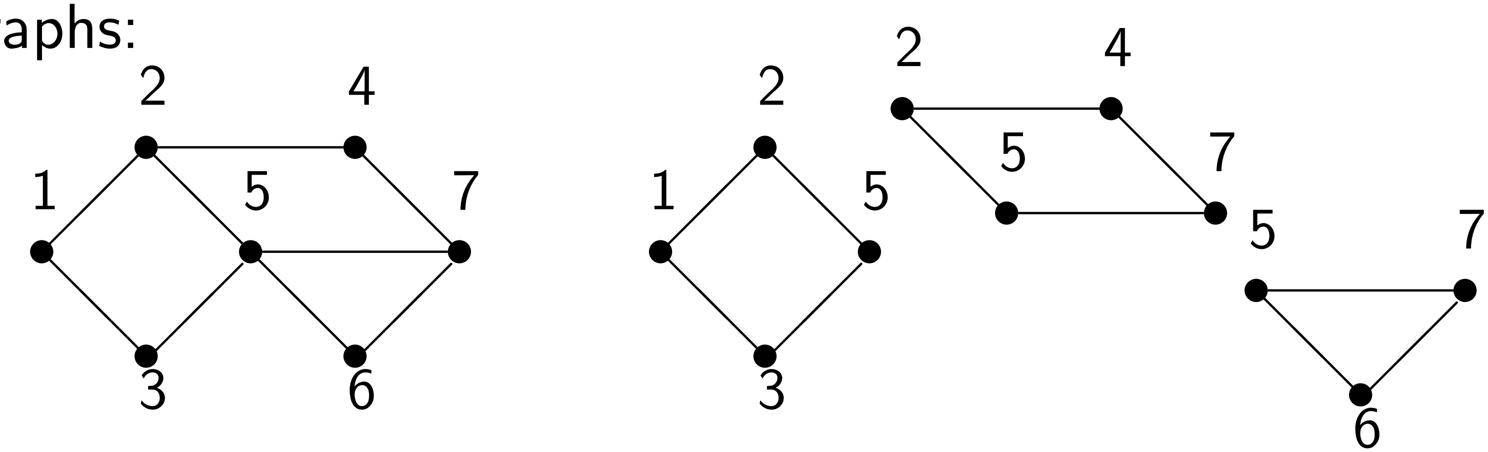
(i) *$P_{A\cup S}$ and $P_{B\cup S}$ satisfy (F) w.r.t. $\mathcal{G}_{A\cup S}$ and $\mathcal{G}_{B\cup S}$ respectively;*

(ii) $f(x)f_S(x_S) = f_{A\cup S}(x_{A\cup S})f_{B\cup S}(x_{B\cup S})$.

The converse also holds in the sense that *if (i) and (ii) hold, and $(A, B, S)$ is a decomposition of $\mathcal{G}$, then $P$ factorizes w.r.t. $\mathcal{G}$.*

# Decomposability

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.

# Perfect numbering

A numbering $V = \{1, \ldots, |V|\}$ of the vertices of an undirected graph is *perfect* if

$$\forall j = 2, \ldots, |V| : \mathrm{bd}(j) \cap \{1, \ldots, j-1\} \text{ is complete in } \mathcal{G}.$$

A set $S$ is an *$(\alpha, \beta)$-separator* if $\alpha \perp_{\mathcal{G}} \beta \mid S$,

# Characterizing chordal graphs

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 $(\alpha, \beta)$-separator are complete.*

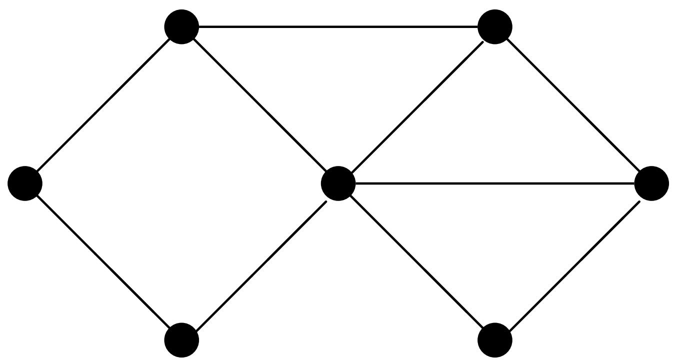
*Trees are chordal graphs* and thus decomposable.

Here is a (greedy) algorithm for checking chordality:

1. Look for a vertex $v^*$ with $\mathrm{bd}(v^*)$ complete. *If no such vertex exists, the graph is not chordal.*
2. Form the subgraph $\mathcal{G}_{V \setminus v^*}$ and let $v^* = |V|$;
3. Repeat the process under 1;
4. *If the algorithm continues until only one vertex is left, the graph is chordal and the numbering is perfect.*

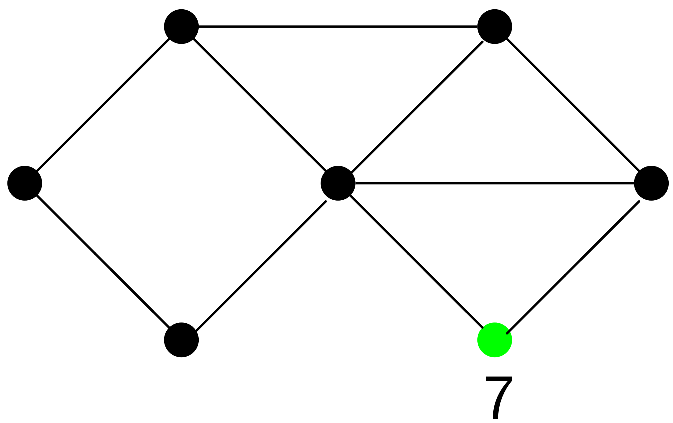
The complexity of this algorithm is $O(|V|^2)$.

# Greedy algorithm



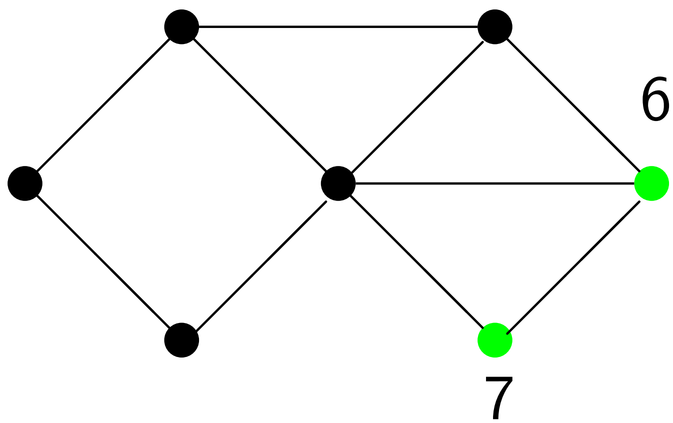
Is this graph chordal?

# Greedy algorithm



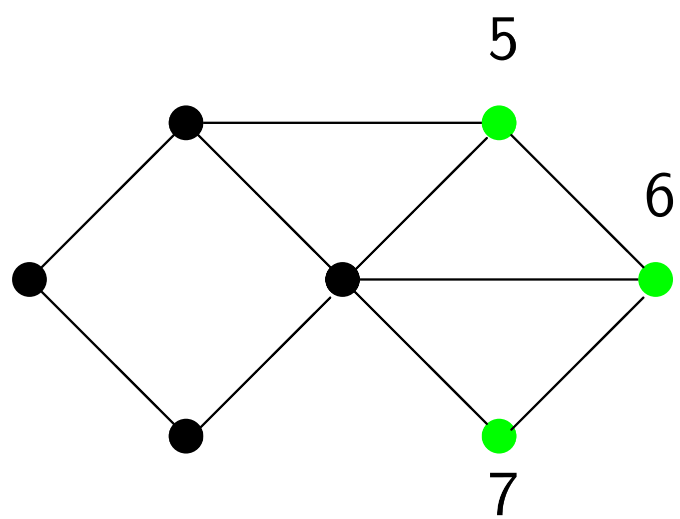

Is this graph chordal?

# Greedy algorithm



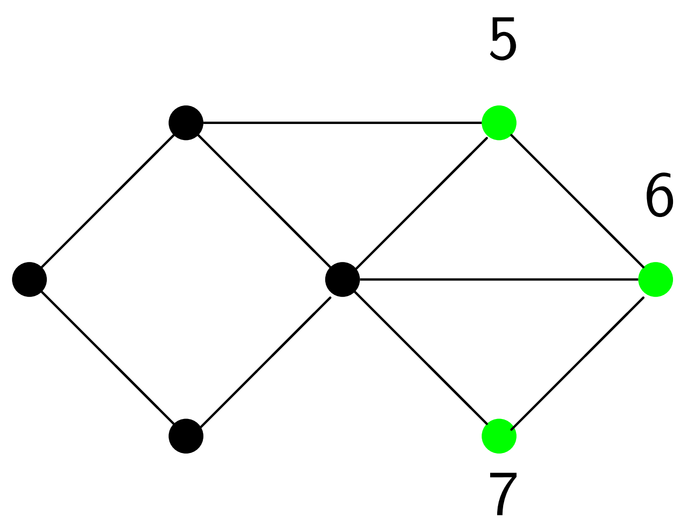

Is this graph chordal?

# Greedy algorithm



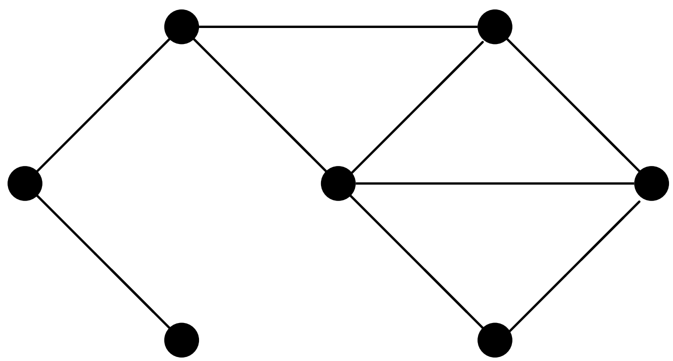

Is this graph chordal?

# Greedy algorithm



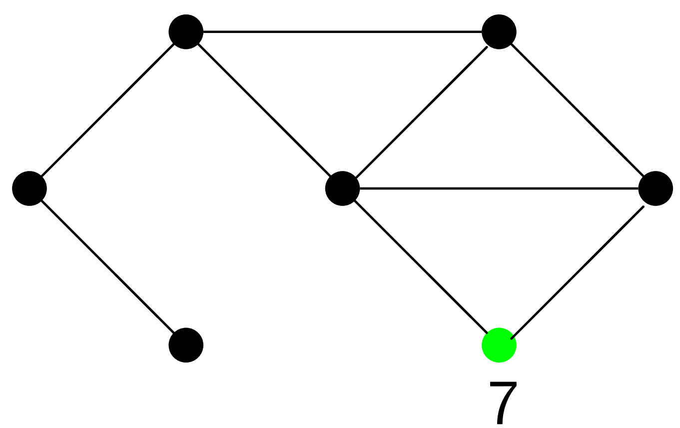

This graph is *not* chordal, as there is no candidate for number 4.

# Greedy algorithm



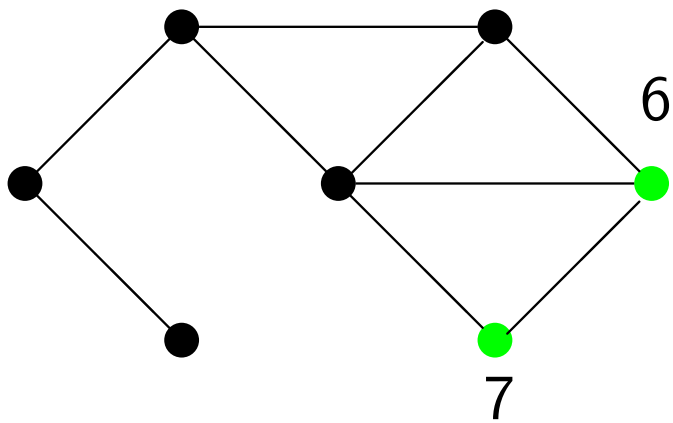
Is this graph chordal?

# Greedy algorithm



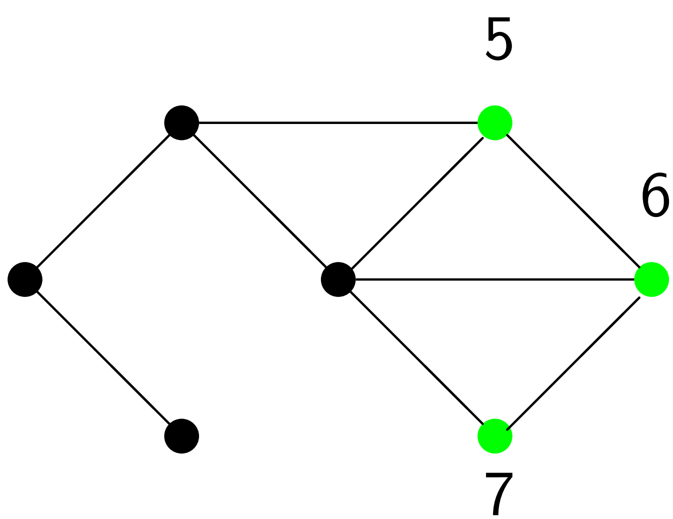

Is this graph chordal?

# Greedy algorithm



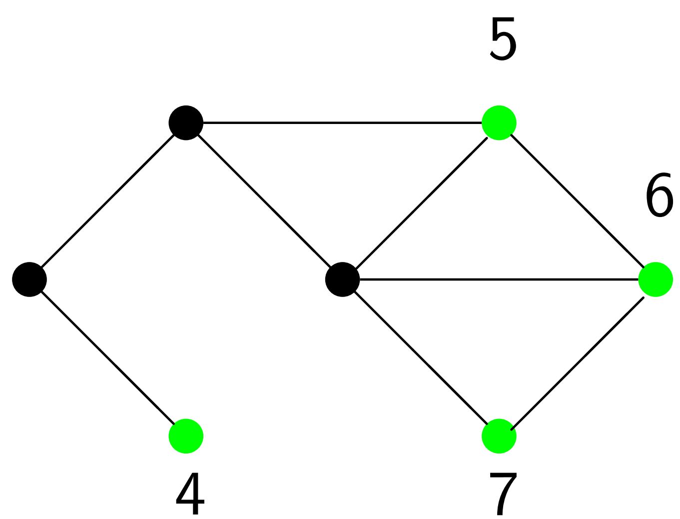

Is this graph chordal?

# Greedy algorithm



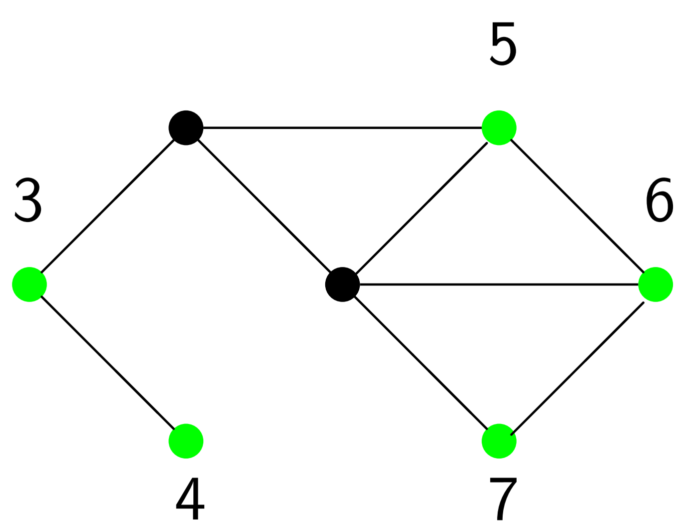

Is this graph chordal?

# Greedy algorithm



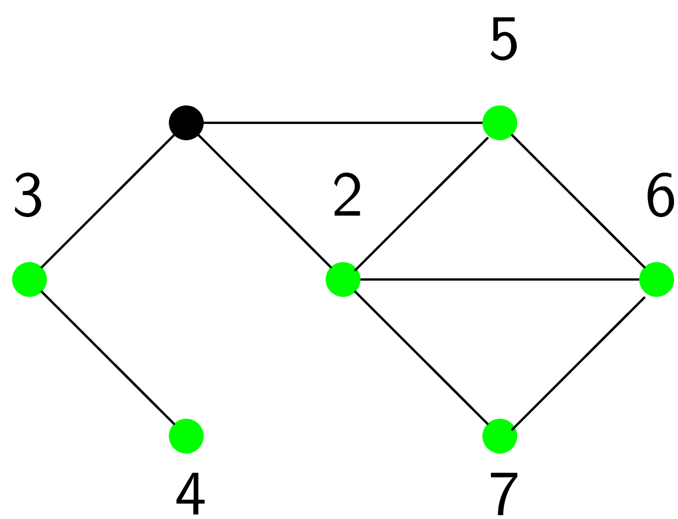

Is this graph chordal?

# Greedy algorithm



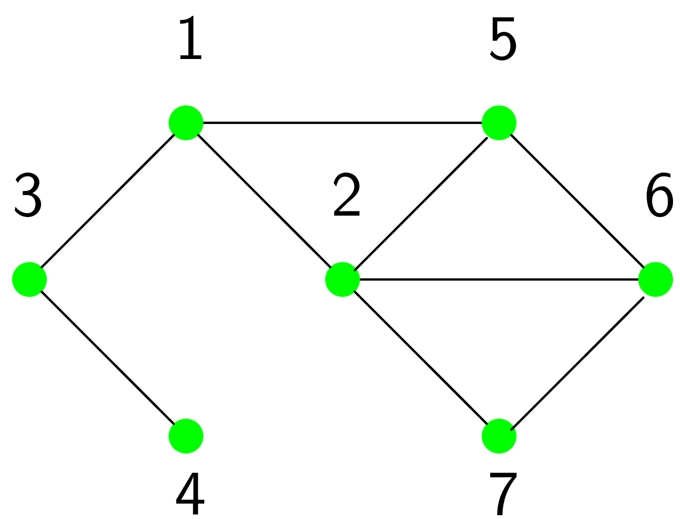

Is this graph chordal?

# Greedy algorithm




Is this graph chordal?

# Greedy algorithm

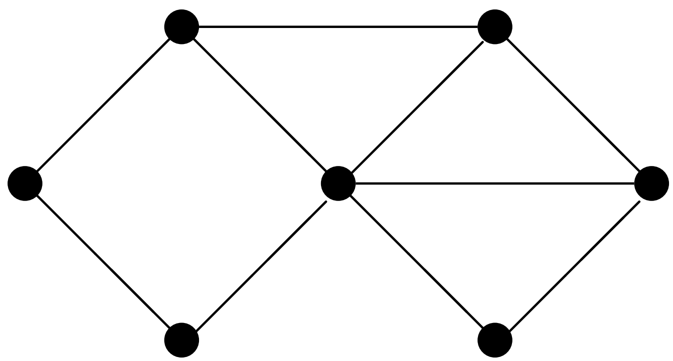



*This graph is chordal!*

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

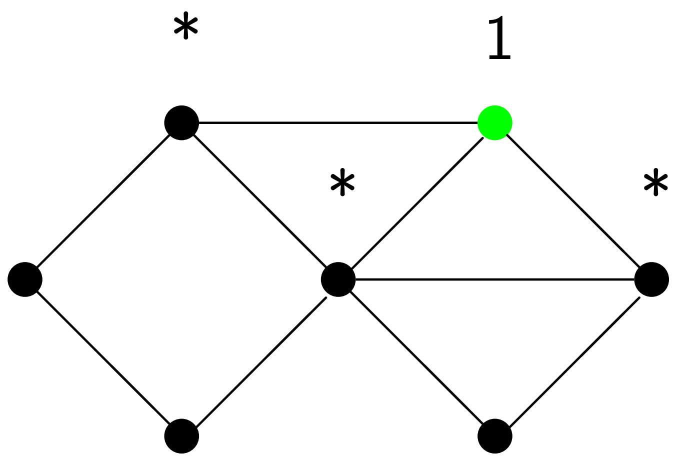
1. Choose $v_0 \in V$ arbitrary and let $v_0 = 1$;
2. When vertices $\{1, 2, \ldots, j\}$ have been identified, choose $v = j + 1$ among $V \setminus \{1, 2, \ldots, j\}$ with highest cardinality of its numbered neighbours;
3. *If* $\mathrm{bd}(j + 1) \cap \{1, 2, \ldots, j\}$ *is not complete,* $\mathcal{G}$ is not chordal;
4. Repeat from 2;
5. *If the algorithm continues until no vertex is left, the graph is chordal and the numbering is perfect.*

# Maximum Cardinality Search



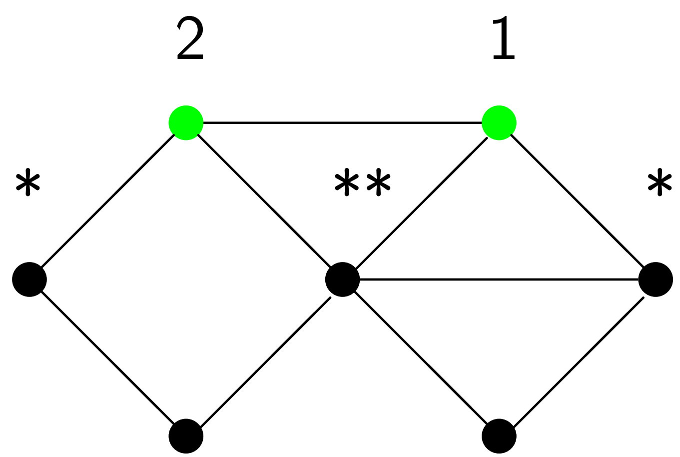
Is this graph chordal?

# Maximum Cardinality Search



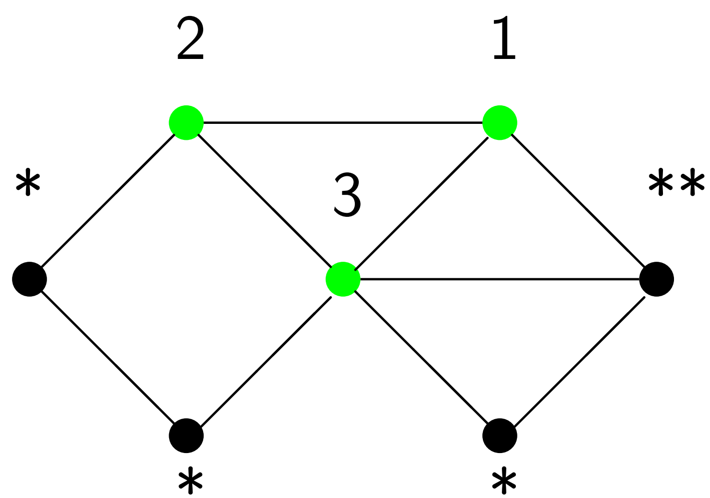

Is this graph chordal?

# Maximum Cardinality Search



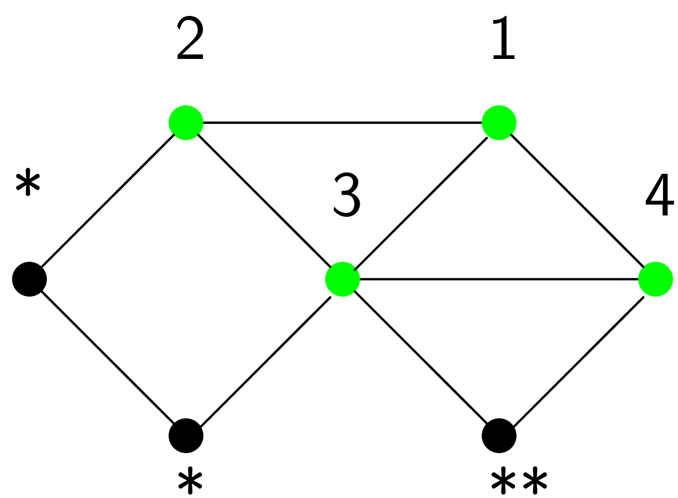

Is this graph chordal?

# Maximum Cardinality Search



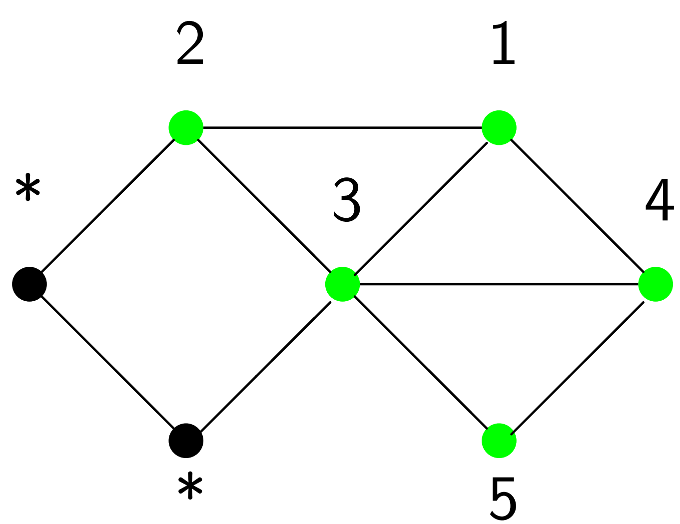

Is this graph chordal?

# Maximum Cardinality Search



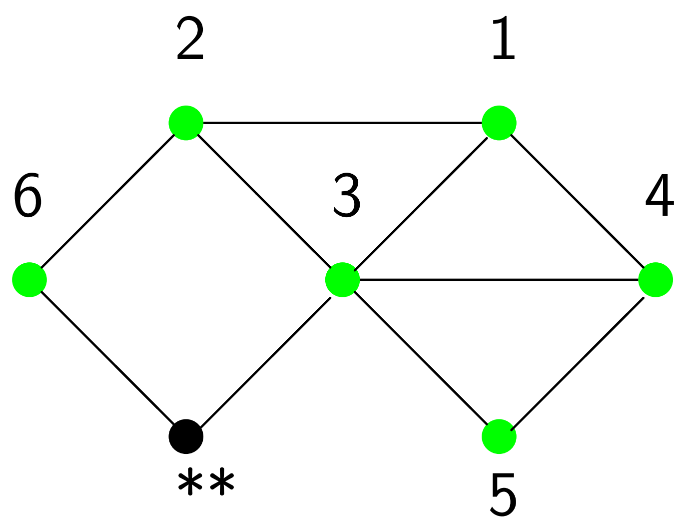

Is this graph chordal?

# Maximum Cardinality Search



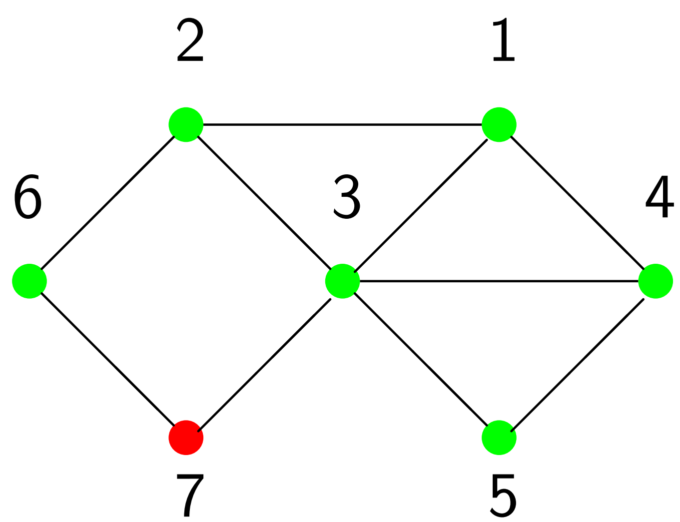

Is this graph chordal?

# Maximum Cardinality Search



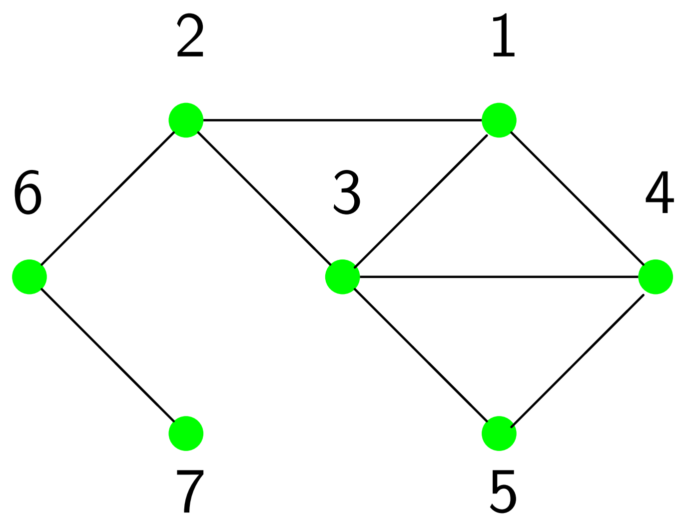

Is this graph chordal?

# Maximum Cardinality Search



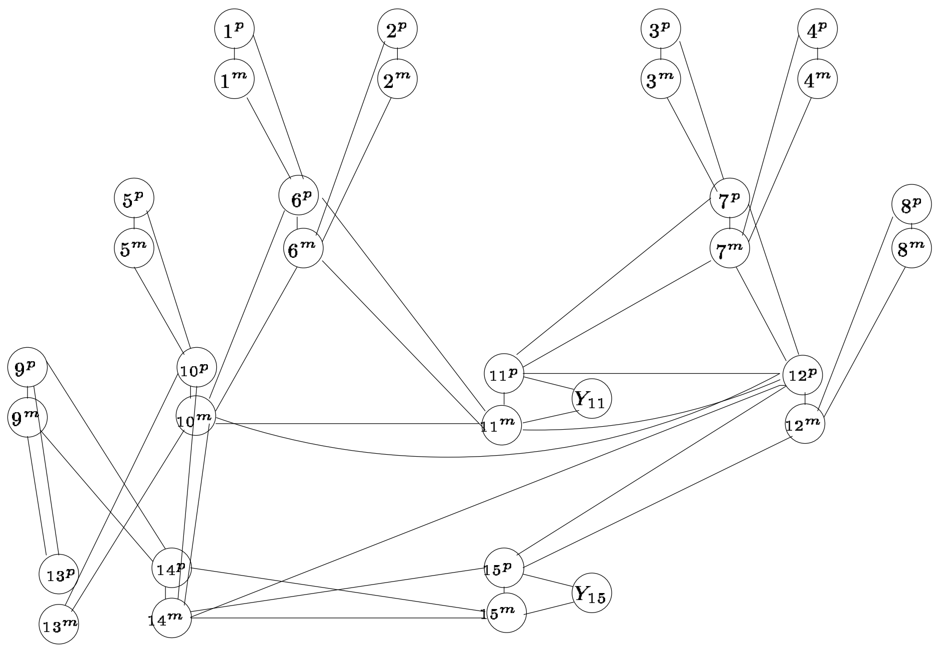

*The graph is not chordal!* because 7 does not have a complete boundary.

# Maximum Cardinality Search




MCS numbering for the chordal graph. Algorithm runs essentially as before.

# A chordal graph



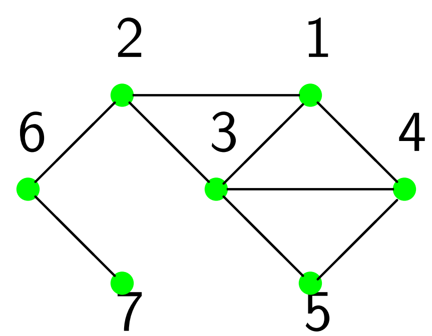

This graph is chordal, but it might not be that easy to see. . . Maximum Cardinality Search is handy!

# 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 $\lambda$ a *ladder vertex* if $\lambda = |V|$ or if $\pi_{\lambda+1} < \pi_\lambda + 1$. Let $\Lambda$ be the set of ladder vertices.




$\pi_\lambda$: 0,1,2,2,2,1,1.

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

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 $(\alpha, \beta)$-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 $C_\lambda$ are the cliques: Since the MCS-numbering is perfect, $C_\lambda, \lambda > \lambda_{\min}$ all satisfy

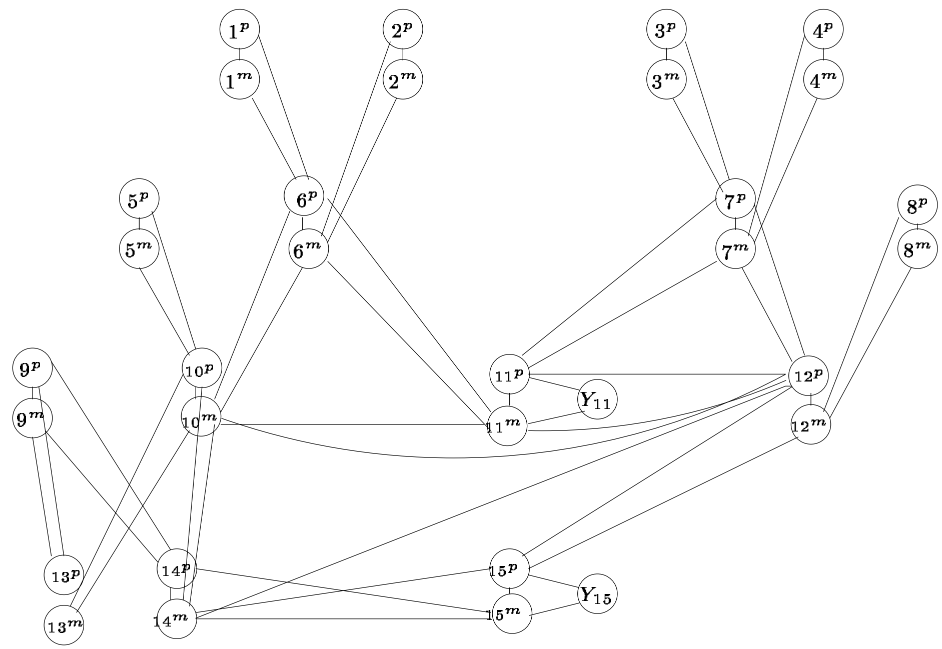
$$C_\lambda \cap (\cup_{\lambda' < \lambda} C_{\lambda'}) = C_\lambda \cap C_{\lambda^*} = S_\lambda$$

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

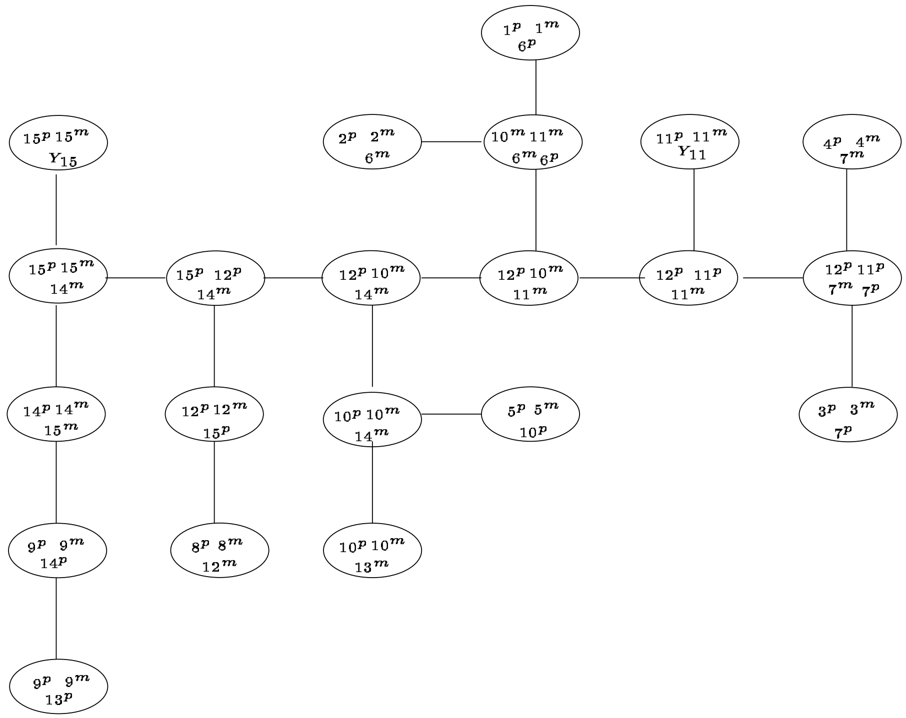
A junction tree is now easily constructed by attaching $C_\lambda$ to any $C_{\lambda^*}$ satisfying the above. Although $\lambda^*$ may not be uniquely determined, $S_\lambda$ is.

Indeed, the sets $S_\lambda$ 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 prime components 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.

If the graph $\mathcal{G}$ is chordal, we say that the graphical model is *decomposable*.

In this case, *the IPS-algorithm converges in a finite number of steps*.

We also have the familiar *factorization of densities*

$$f(x \mid \Sigma) = \frac{\prod_{C \in \mathcal{C}} f(x_C \mid \Sigma_C)}{\prod_{S \in \mathcal{S}} f(x_S \mid \Sigma_S)^{\nu(S)}} \qquad (1)$$

where $\nu(S)$ is the number of times $S$ appear as intersection between neighbouring cliques of a junction tree for $\mathcal{C}$.

# Relations for trace and determinant

Using the factorization (1) we can for example match the expressions for the trace and determinant of $\Sigma$

$$\operatorname{tr}(KW) = \sum_{C\in\mathcal{C}} \operatorname{tr}(K_C W_C) - \sum_{S\in\mathcal{S}} \nu(S)\operatorname{tr}(K_S W_S)$$

and further

$$\det\Sigma \quad = \quad \{\det(K)\}^{-1} = \frac{\prod_{C\in\mathcal{C}}\det\{\Sigma_C\}}{\prod_{S\in\mathcal{S}}\{\det(\Sigma_S)\}^{\nu(S)}}$$

These are some of many relations that can be derived using the decomposition property of chordal graphs.

The same factorization clearly holds for the maximum likelihood estimates:

$$f(x \mid \hat{\Sigma}) = \frac{\prod_{C \in \mathcal{C}} f(x_C \mid \hat{\Sigma}_C)}{\prod_{S \in \mathcal{S}} f(x_S \mid \hat{\Sigma}_S)^{\nu(S)}} \quad (2)$$

Moreover, it follows from the general likelihood equations that

$$\hat{\Sigma}_A = W_A/n \text{ whenever } A \text{ is complete.}$$

Exploiting this, we can obtain an explicit formula for the maximum likelihood estimate in the case of a chordal graph.

For a $|d| \times |e|$ matrix $A = \{a_{\gamma\mu}\}_{\gamma \in d, \mu \in e}$ we let $[A]^V$ denote the matrix obtained from $A$ by filling up with zero entries to obtain full dimension $|V| \times |V|$, i.e.

$$\left([A]^V\right)_{\gamma\mu} = \begin{cases} a_{\gamma\mu} & \text{if } \gamma \in d, \mu \in e \\ 0 & \text{otherwise.} \end{cases}$$

*The maximum likelihood estimates exists if and only if $n \geq C$ for all $C \in \mathcal{C}$. Then the following simple formula holds for the maximum likelihood estimate of $K$:*

$$\hat{K} = n \left\{ \sum_{C \in \mathcal{C}} \left[ (w_C)^{-1} \right]^V - \sum_{S \in \mathcal{S}} \nu(S) \left[ (w_S)^{-1} \right]^V \right\}.$$

## "Clique-separator formula" for $\widehat{K}$.

Suppose that the graph $\mathcal{G}$ is decomposable. Let $Cliq$ be the set of all cliques of $\mathcal{G}$ and $Sep$ the set of all minimal separators of $\mathcal{G}$.
Suppose that $n \geq |C|$ (the number of elements of $C$) for each clique $C$.
If the mean $\xi$ of the model is known and $\tilde{\Sigma}$ is the sample covariance matrix then

$$\widehat{K} = \sum_{C \in Cliq} [\tilde{\Sigma}_C^{-1}]^V - \sum_{S \in Sep} \nu(S) [\tilde{\Sigma}_S^{-1}]^V$$

If the mean is unknown, then $\hat{\xi} = \bar{X}$ and one uses the "Clique-separator formula" for $\widehat{K}$ with the corrected sample covariance matrix $\frac{n}{n-1}\tilde{\Sigma}$.

**Back to Example "Simpson paradox"** $\mathcal{G}:\ 1 —— 3 —— 2$

Suppose that $\xi = 0$ and the sample covariance matrix equals

$\tilde{\Sigma} = \begin{pmatrix} 1 & 0.5 & 1 \\ 0.5 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$. The graph $\mathcal{G}$ governs the model.

We computed "by hand" $\hat{\Sigma} = \begin{pmatrix} 1 & \frac{2}{3} & 1 \\ \frac{2}{3} & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$

Let us find $\hat{K}$ and $\hat{\Sigma}$ by "Clique-separator formula".

The cliques of $\mathcal{G}$ are $C_1 = \{1, 3\}$ and $C_2 = \{2, 3\}$.
The minimal separator is $S = \{3\}$.

$\tilde{\Sigma} = \begin{pmatrix} 1 & 0.5 & 1 \\ 0.5 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$. We only use $\pi_{\mathcal{G}}(\tilde{\Sigma}) = \begin{pmatrix} 1 & & 1 \\ & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$

Apply the "Clique-separator formula" for $\hat{K}$:
$\hat{K} = [\tilde{\Sigma}_{1,3}^{-1}]^V + [\tilde{\Sigma}_{2,3}^{-1}]^V - [\tilde{\Sigma}_{3}^{-1}]^V$.

$$\tilde{\Sigma}_{1,3}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}; \quad [\tilde{\Sigma}_{1,3}^{-1}]^V = \frac{1}{2}\begin{pmatrix} 3 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\tilde{\Sigma}_{2,3}^{-1} = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix}; \quad [\tilde{\Sigma}_{2,3}^{-1}]^V = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

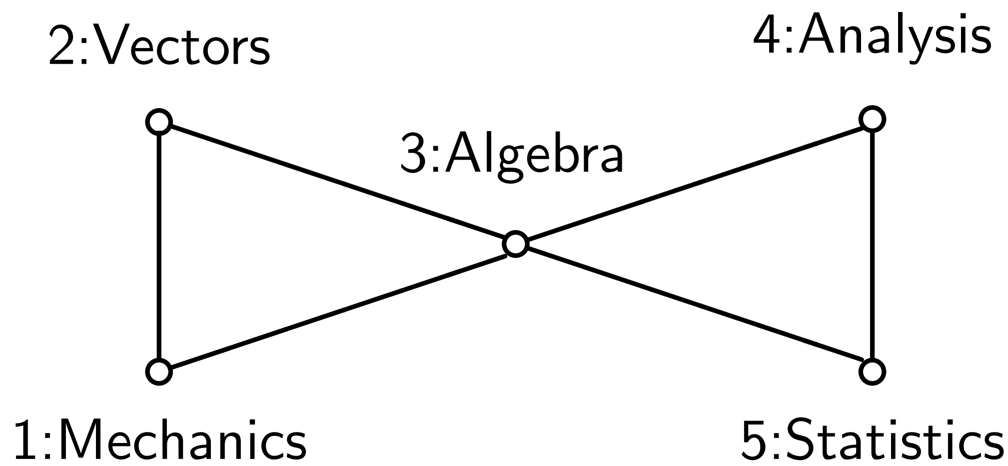
$$[\tilde{\Sigma}_{3}^{-1}]^V = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

$$\hat{K} = \begin{pmatrix} \frac{3}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{3}{2} & -1 \\ -\frac{1}{2} & -1 & \frac{7}{6} \end{pmatrix}; \quad \hat{\Sigma} = \hat{K}^{-1} = \begin{pmatrix} 1 & \frac{2}{3} & 1 \\ \frac{2}{3} & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

**Exercise.** Suppose that $\mathcal{G}$ : 1—2—3, the mean $\xi = 0$ and $\tilde{\Sigma} = \begin{pmatrix} 1 & 1 & 0.9 \\ 1 & 2 & 2 \\ 0.9 & 2 & 3 \end{pmatrix}$.

Compute by the clique-separator formula the MLEs $\hat{K}$ and $\hat{\Sigma}$.

# Mathematics marks




This graph is chordal with cliques $\{1, 2, 3\}$, $\{3, 4, 5\}$ with separator $S = \{3\}$ having $\nu(\{3\}) = 1$.

Since one degree of freedom is lost by subtracting the average, we get in this example

$$\hat{K} = 87 \begin{pmatrix} w^{11}_{[123]} & w^{12}_{[123]} & w^{13}_{[123]} & 0 & 0 \\ w^{21}_{[123]} & w^{22}_{[123]} & w^{23}_{[123]} & 0 & 0 \\ w^{31}_{[123]} & w^{32}_{[123]} & w^{33}_{[123]} + w^{33}_{[345]} - 1/w_{33} & w^{34}_{[345]} & w^{35}_{[345]} \\ 0 & 0 & w^{43}_{[345]} & w^{44}_{[345]} & w^{45}_{[345]} \\ 0 & 0 & w^{53}_{[345]} & w^{54}_{[345]} & w^{55}_{[345]} \end{pmatrix}$$

where $w^{ij}_{[123]}$ is the $ij$th element of the inverse of

$$W_{[123]} = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix}$$

and so on.