The tiling by the minimal separators of a junction tree and applications to graphical models

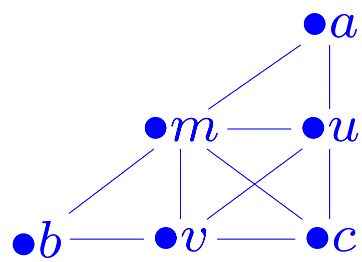
G. Letac et H. Massam

Summer school of St Flour, July 2006.

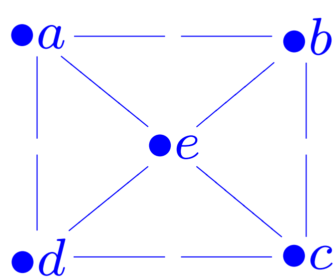
Gaussian graphical models. Let $G = (V, \mathcal{E})$ a finite undirected graph with $V = \{1, \ldots, n\}$. Consider a normal centered rv $X = (X_1, \ldots, X_n)$ with invertible covariance Σ such that $K = \Sigma^{-1} = (k_{ij})$ satisfies $k_{ij} = 0$ for $i \neq j$ and $\{i, j\} \notin \mathcal{E}$. This is equivalent to impose that X_i and X_j are conditionally independent knowing the remainder of $(X_1, \ldots, X_n)$.

Problems : estimation of Σ by maximum likelihood or Bayesian techniques.

Decomposable graphs The graph $G = (V, \mathcal{E})$ is *decomposable* if it is connected and if it does not contain any *induced* cycle of length ≥ 4. For instance in



$mucv$ is a cycle but it is not an induced cycle. In



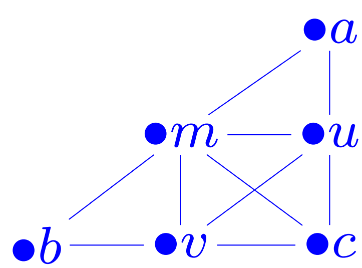
$abcd$ is an induced cycle (and this second graph is not decomposable). This has been considered for the first time by Chvatal en 1958.

What decomposable graphs are done for ? For us here : because the search of the maximum likelihood of Σ becomes a linear problem if G is decomposable and leads to equations of degree ≥ 5 if G is not decomposable. There are numerous applications of decomposable graphs in other parts of mathematics .

What one needs to know about decomposable graphs

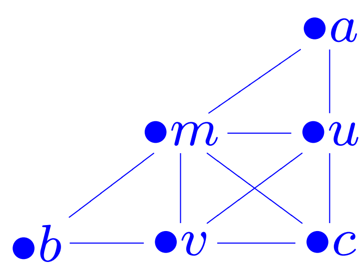
1. Cliques and junction trees.
2. Minimal separators
3. Perfect orderings of cliques
4. The two definitions of the multiplicity of a minimal separator.

Cliques and junction trees The cliques of a graph are its maximal complete subsets. A junction tree has the set of cliques as set of vertices and is such that if the clique C'' is on the unique path from C to C' then $C'' \supset C \cap C'$. For instance $\bullet 1 - \bullet 2 - \bullet 3$ is a junction tree for the decomposable graph



where the three cliques are $1 = (amu)$, $2 = (muvc)$ and $3 = (bmv)$. A connected graph is decomposable if and only if a junction tree exists (a neat proof of this is given by Blair and Peyton in 1991)

Minimal separators If a and b are not neighbors $S \subset V$ is a separator of a and b if any path from a to b hits S



For instance $muvc$ is a separator of a and b. If nothing can be taken out, S is a *minimal* separator of a and b. Finally S is *minimal* separator by itself if there exist non adjacent a and b such that S is a minimal separator of a and b. There are not so many of them, strictly less that the number of cliques anyway. They are mu and mv in the example. A connected graph is decomposable if and only if all the minimal separators are complete (Dirac 1961).

Perfect orderings of the cliques Let $\mathcal{C}$ be the family of the k cliques of the connected graph (not necessarily decomposable). Consider a bijection $P : \{1,\dots,k\} \to \mathcal{C}$ and

$$S_P(j) = [P(1) \cup P(2) \cup \ldots \cup P(j-1)] \cap P(j)$$

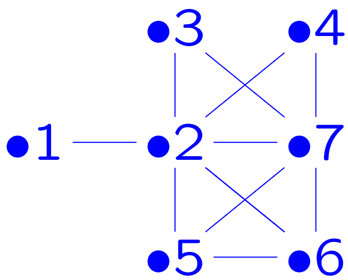
for $j \geq 2$. Then P is said to be *perfect* if for all $j \geq 2$ there exists $i_j < j$ such that

$$S_P(j) \subset P(i_j).$$

This is a deep notion : a connected graph is decomposable if and only if a perfect ordering of the cliques exists. If G is decomposable and if P is perfect then $S_P(j)$ is a minimal separator. Given a minimal separator S, the number $\nu_P(S)$ of $j \geq 2$ such that $S_P(j) = S$ is called the *multiplicity* de S. Lauritzen (1996) observes that $\nu_P(S) \geq 1$ and that $\nu_P(S))$ *does not depend on* P (we give a proof below). Thus by definition if $\mathcal{S}$ denotes the set of the minimal separators of a decomposable graph having k cliques then

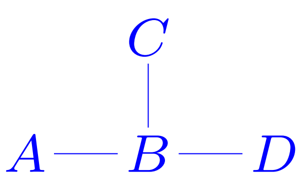
$$\sum_{S \in \mathcal{S}} \nu_P(S) = k - 1$$

Example :

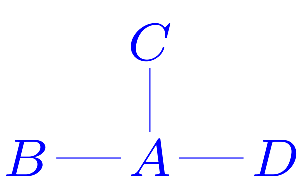


There are 4 cliques $A = \{1, 2\}$, $B = \{2, 3, 7\}$, $C = \{2, 4, 7\}$, $D = \{2, 5, 6, 7\}$ and two minimal separators $U = \{2\}$, $V = \{2, 7\}$. The ordering $ABCD$ is perfect with $S_2 = U$ et $S_3 = S_4 = V$. Therefore V has multiplicity 2 and U has multiplicity 1.

Remark :



is a junction tree, and

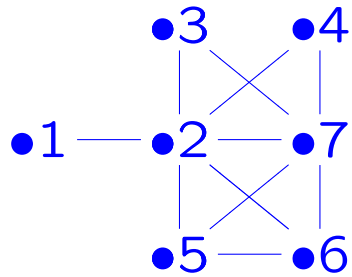


is not.

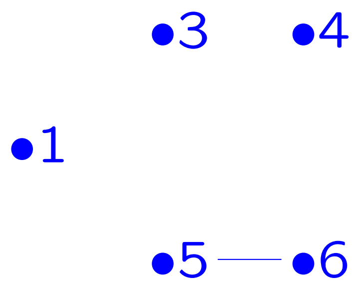
Topological multiplicity of a minimal separator Let S be a minimal separator of a decomposable graph $(V, \mathcal{E})$. Let $\{V_1, \ldots, V_p\}$ be the connected components of $V \setminus S$ (of course $p \geq 2$). Let q be the number of $j = 1, \ldots, p$ such that S is *NOT* a clique of $S \cup V_j$. The number $\nu(S) = q - 1$ is called the topological multiplicity of S.

(The notion is introduced by Lauritzen, Speed and Vivayan in 1979). Question : one observes that in all cases the two definitions of multiplicity coincide. Why ? Answer later on.

Example : If I remove the minimal separator $V = \{2, 7\}$ to its graph



four connected components are obtained :



If I add V to each of them, thus for component 1 I obtain the graph

whose $V = \{2, 7\}$ is a clique. This is not the case for the three other connected components 3, 4 et 56. Therefore $q = 3$ here and the topological multiplicity of V is 2.

Tiling of a junction tree by the minimal separators If $(H, \mathcal{E}(H))$ is a tree (undirected) with vertex set H and edge set $\mathcal{E}(H)$ a *tiling* of H is a family $\mathcal{T}$ of subtrees

$$\mathcal{T} = \{T_1, \ldots, T_p\}$$

of H such that if $\mathcal{E}(T_i)$ is the edge set of T_i then

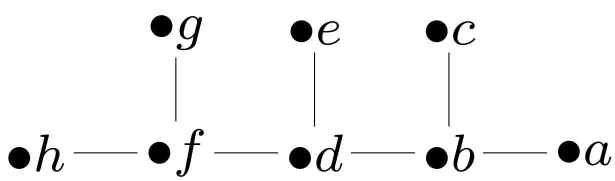
$$\{\mathcal{E}(T_1), \ldots, \mathcal{E}(T_q)\}$$

is a partition of $\mathcal{E}(H)$. This implies

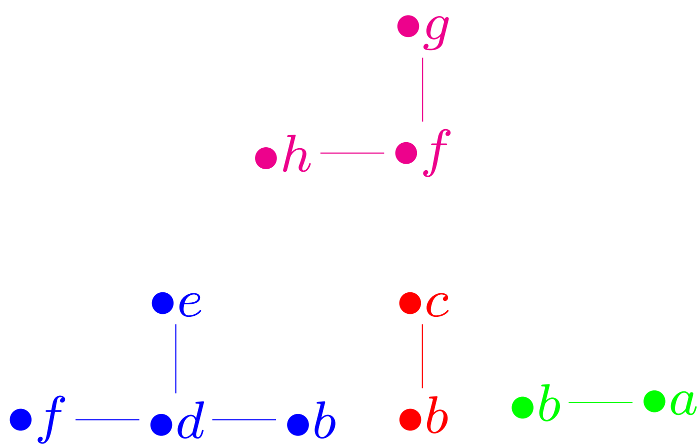
$$T_1 \cup \ldots \cup T_q = H$$

although $(T_1, \ldots, T_p)$ is not a partition of the set H.

Example



the tiles of the tiling can be chosen as



Theorem 1.

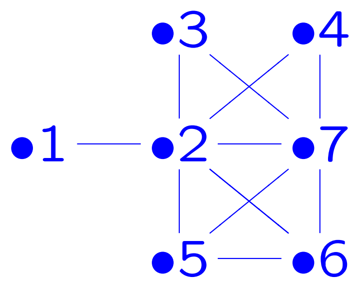
Let $G = (V, \mathcal{E})$ be a decomposable graph and let $(\mathcal{C}, \mathcal{E}(\mathcal{C}))$ be a junction tree of G. Let $\mathcal{S}$ be the family of minimal separators of G. There exists a unique tiling $\mathcal{T}$ of the tree $(\mathcal{C}, \mathcal{E}(\mathcal{C}))$ by subtrees and a bijection $S \mapsto T_S$ from $\mathcal{S}$ towards $\mathcal{T}$ with the following property : for all $S \in \mathcal{S}$ the edges of T_S are the edges $\{C, C'\}$ such that $S = C \cap C'$.

Under these circumstances the number of edges of T_S is the topological multiplicity of S. Furthermore if C and C' are two distinct cliques consider the unique path $(C = C_0, C_1, \ldots, C_q = C')$ from C to C'. Let $S_i \in \mathcal{S}$ such that $\{C_{i-1}, C_i\}$ is in T_{S_i}. Then

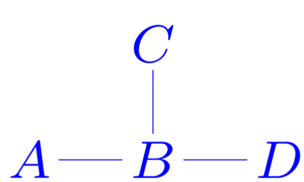
$$C \cap C' = \cap_{i=1}^{q} S_i.$$

In particular $C \cap C' = S$ if C and C' are in T_S.

Consider again the example :



There are 4 cliques $A = \{1, 2\}$, $B = \{2, 3, 7\}$, $C = \{2, 4, 7\}$, $D = \{2, 5, 6, 7\}$ and two minimal separators $U = \{2\}$, $V = \{2, 7\}$. The ordering $ABCD$ of the cliques is perfect with $S_2 = U$ et $S_3 = S_4 = V$. Thus V has multiplicity 2 and U has multiplicity 1. Consider the junction tree



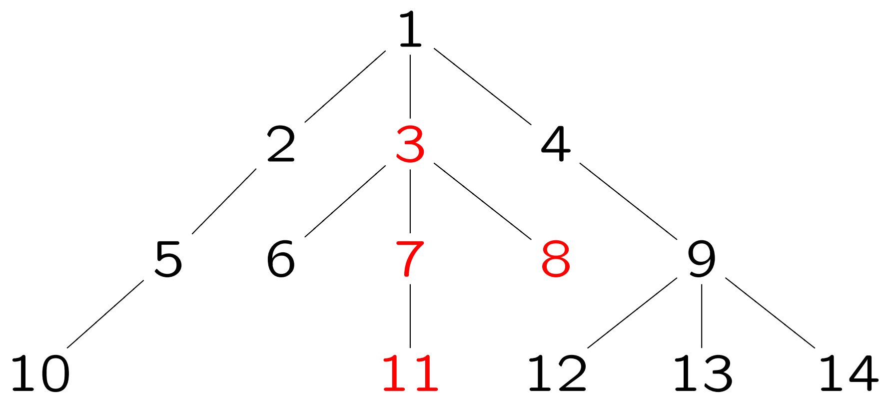
Then $T_U = AB$ et $T_V = BCD$.

Junction trees and perfect orderings of cliques. Recall that saying that is P is a perfect ordering of the set $\mathcal{C}$ of the k cliques of a decomposable graph is to say that there exists $i_j < j$ such that $S_P(j) \subset P(i_j)$. There exist in general several possible i_j's. Actually we fix one such i_j for each j and we create the graph having $\mathcal{C}$ as vertex set with having the $k-1$ edges $\{P(i_j), P(j)\}$. A beautiful result of Beeri, Fagin, Maier and Yannakakis (1983) claims that this graph is a junction tree and conversely that any junction tree can be constructed from a perfect ordering and from a choice of the $j \mapsto i_j$. Let us say that a junction tree is *adapted to* the perfect ordering P if there exists a choice $j \mapsto i_j$ giving the tree.

Tiling by minimal separators and perfect ordeings of the cliques. Let P be a perfect ordering of the set $\mathcal{C}$ of the k cliques of a decomposable graph and let S be in the set $\mathcal{S}$ of the minimal separators. Consider the set of cliques $J(P,S) =$

$$\{C \in \mathcal{C}\ ;\ \exists j \geq 2 \text{ tel que } P(j) = C \text{ et } S_P(j) = S\}.$$

Its importance in terms of Gaussian graphical models will be explained later on. Just remark that $\nu_P(S) = |J(P,S)|$. Consider now a junction tree adapted to P and let $\mathcal{T}$ be the tiling of this tree by the minimal separators. We transform this undirected tree into a rooted tree by taking $P(1)$ as a root. This transforms $\mathcal{C}$ into a partially ordered set : $C \preceq C'$ if the unique path from $P(1)$ to C' passes through C.



Now for all $S \in \mathcal{S}$ the subtree T_S has a *minimal point* $M(S)$ for this partial order. Here is now a useful result ruling out the old contest between multiplicities (recall that the number of vertices of a tree is the number of edges plus one) :

Theorem 2.

$J(P, S) = T_S \setminus \{M(S)\}$. In particular $\nu_P(S)$ is the topological multiplicity $|T_S| - 1$ of S.

Actually $J(P,S)$ depends on S and on $S_P(2)$ only

Theorem 3. Let P and P' two perfect orderings such that $P(1) \cap P(2) = P'(1) \cap P'(2)$, that is to say $S_P(2) = S_{P'}(2)$ (denoted S_2). Then $J(P,S) = J(P',S)$ if $S \neq S_2$ and

$$J(S_2, P) \cup \{P(1)\} = J(S_2, P') \cup \{P'(1)\}.$$

What is good for? For Gaussian graphical models Let $G = (V, \mathcal{E})$ be a decomposable graph with $V = \{1, \ldots, n\}$. Let $\mathcal{S}_n$ be the symmetric matrices of order n, let $\mathcal{P}_n \subset \mathcal{S}_n$ be the positive definite ones, let $ZS_G \subset \mathcal{S}_n$ be the subspace of matrices (z_{ij}) such that $z_{ij} = 0$ if $i \neq j$ and $\{i, j\} \notin \mathcal{E}$). Finally let $P_G = ZS_G \cap \mathcal{P}_n$ be the positive definite matrices with zeros prescribed by G. The model is therefore

$$\{N(0, \Sigma) \; ; \; \Sigma^{-1} \in P_G\}.$$

Denote by π the natural projection of $\mathcal{S}_n$ on ZS_G and denote $Q_G = \pi(P_G^{-1})$. This set Q_G is a convex cone with numerous properties : it carries the useful part of Σ^{-1} and of S^{-1} when the unknown covariance is Σ and the empirical covariance is S. The cone Q_G is the dual of the cone P_G. Finally, Q_G is characterized by the fact that the restriction x_C of $x \in Q_G$ to any clique C is positive definite.

The Wishart distributions for Gaussian graphical models are indexed by the minimal separators

Let us fix $\alpha : \mathcal{C} \rightarrow \mathbb{R}$ and $\beta : \mathcal{S} \rightarrow \mathbb{R}$ and let us introduce the function $x \mapsto H(\alpha, \beta; x)$ on Q_G by

$$H(\alpha, \beta; x) = \frac{\prod_{C \in \mathcal{C}} \det(x_C)^{\alpha(C)}}{\prod_{S \in \mathcal{S}} \det(x_S)^{\nu(S)\beta(S)}}.$$

Define the measure on Q_G by

$$\mu_G(dx) = H(-\frac{1}{2}(|C|+1), -\frac{1}{2}(|S|+1; x)\mathbf{1}_{Q_G}(x)dx.$$

An important result is that if P is a perfect ordering and if for all $S \in \mathcal{S}$ different from $S_P(2)$ one has

$$\sum_{C \in J(P,S)} (\alpha(C) - \beta(S)) = 0$$

then by a long calculation one sees that there exists a number $\Gamma(\alpha, \beta)$ with the following eigenvalue property : for all $y \in P_G$

$$\int_{Q_G} e^{-\operatorname{tr} xy} H(\alpha, \beta; x) \mu_G(dx) = \Gamma(\alpha, \beta) H(\alpha, \beta; \pi(y^{-1})).$$

A conclusion of the present lecture is the above fact is not linked to a particular perfect ordering P but only to the minimal separator S_2!