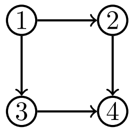
Probabilistic Graphical Models: Problem Set 5 (Solutions)

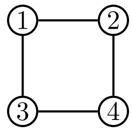
Svante Linusson, Liam Solus
KTH Royal Institute of Technology

28 September 2021

1. What is the treewidth of each of the following graphs?



$\mathcal{G}$



$\mathcal{H}$

Solution:

Recall that the width of an induced graph $\mathcal{G}_{\Phi,\prec}$ is the size of any largest maximal clique in $\mathcal{G}_{\Phi,\prec}$ minus 1. If $\mathcal{G} = ([m], E)$ is a DAG, its set of factors is $\Phi = \{P(x_i \mid x_{\mathrm{pa}_{\mathcal{G}}(i)})\}_{i=1}^m$. The induced-width of $\mathcal{G}$ with respect to the elimination ordering $\prec$, where $\prec$ is an ordering (permutation) of $X_1, ..., X_m$, is the width of $\mathcal{G}_{\Phi,\prec}$, and we denote it by $\omega_{\Phi,\prec}$. The tree-width of the DAG $\mathcal{G}$ is then $\omega_{\mathcal{G}}$, which is the minimum of the $\omega_{\Phi,\prec}$ over all permutations of $X_1, \ldots, X_m$. For the given DAG $\mathcal{G}$, the set of factors we begin with is

$$\Phi = \{P(x_1), P(x_2 \mid x_1), P(x_3 \mid x_1), P(x_4 \mid x_2, x_3)\}.$$

Recall that the graph $\mathcal{G}_{\Phi,\prec}$ has an edge $i - j$ whenever x_i and x_j appear together in the initial set of factors Φ or if x_i and x_j appear together in a factor ψ produced in the variable elimination process. The former condition implies that $\mathcal{G}_{\Phi,\prec}$ always contains the edges $1-2, 1-3, 2-3, 2-4$, and $3-4$. Denote this subgraph of $\mathcal{G}_{\Phi,\prec}$ by

$$\mathcal{H}_\Phi = ([4], \{1-2, 1-3, 2-3, 2-4, 3-4\}).$$

$\mathcal{H}_\Phi$ has two maximal cliques with vertex sets $\{1, 2, 3\}$ and $\{2, 3, 4\}$. Since we would then add in edges based on the factors ψ produced by variable elimination with respect to $\prec$, we know that the size of the largest maximal cliques in $\mathcal{G}_{\Phi,\prec}$ is always at least 3. So a lower bound on the tree-width of $\mathcal{G}$ is two.

We claim that the $\omega_{\mathcal{G}} = 2$. To prove this, we need only find a permutation of X_1, X_2, X_3, X_4 such that variable elimination with respect to this ordering doesn't add in any new edges. Consider the ordering $\prec = (X_1, X_2, X_3, X_4)$. To eliminate variable X_1, we compute the factors

$$\psi_1(x_1, x_2, x_3) = P(x_1)P(x_2 \mid x_1)P(x_3 \mid x_1),$$
$$\tau_1(x_2, x_3) = \sum_{x_1} \psi_1(x_1, x_2, x_3).$$

We then get the updated set of factors $\Phi_1 = \{\tau_1(x_2, x_3), P(x_4 \mid x_2, x_3)\}$. Since x_1, x_2, x_3 appear together in $\psi_1(x_1, x_2, x_3)$, we need to add edges to $\mathcal{H}_\Phi$ for every i and j for which x_i and x_j appear together in ψ_1. These would be the edges $1-2, 1-3, 2-3$, all of which are already in $\mathcal{H}_\Phi$. Hence, we do not need to add any edges! We then eliminate X_2 by creating the factors

$$\psi_2(x_2, x_3, x_4) = \tau_1(x_2, x_3)P(x_4 \mid x_2, x_3),$$
$$\tau_2(x_3, x_4) = \sum_{x_2} \psi_2(x_2, x_3, x_4).$$

We then get the updated set of factors $\Phi_2 = \{\tau_2(x_3, x_4)\}$. Similar to the previous case, since $2-3, 2-4, 3-4$ are all already edges in $\mathcal{H}_\Phi$, then we do not need to add any edges. The remaining two steps of the variable elimination algorithm also do not add in any new edges, so we conclude that $\mathcal{G}_{\Phi,\prec} = \mathcal{G}_\Phi$, which has induced width two. Hence, we can conclude that $\omega_\mathcal{G} = 2$.

For the case when $\mathcal{G}$ is the given undirected graph, our inital set of factors is

$$\Phi = \{\phi_{\{1,2\}}(x_1, x_2), \phi_{\{1,3\}}(x_1, x_3), \phi_{\{2,3\}}(x_2, x_3), \phi_{\{2,4\}}(x_2, x_4)\},$$

This is because the maximal cliques in $\mathcal{G}$ are its edges, and, assuming $P(X_1, X_2, X_3, X_4)$ factors according to $\mathcal{G}$, there are nonnegative functions $\phi_{\{1,2\}}(x_1, x_2), \phi_{\{1,3\}}(x_1, x_3), \phi_{\{2,3\}}(x_2, x_3)$, and $\phi_{\{2,4\}}(x_2, x_4)$, such that

$$P(X_1, X_2, X_3, X_4) = \phi_{\{1,2\}}(x_1, x_2)\phi_{\{1,3\}}(x_1, x_3)\phi_{\{2,3\}}(x_2, x_3)\phi_{\{2,4\}}(x_2, x_4).$$

In this case $\mathcal{H}_\Phi = \mathcal{G}$. Now consider variable elimination with respect to the ordering $\prec = (X_1, X_2, X_3, X_4)$. In the first step we would produce the factors

$$\psi_1(x_1, x_2, x_3) = \phi_{\{1,2\}}(x_1, x_2)\phi_{\{1,3\}}(x_1, x_3),$$
$$\tau_1(x_2, x_3) = \sum_{x_1} \psi_1(x_1, x_2, x_3),$$

and we get the updated set of factors $\Phi_1 = \{\tau_1(x_2, x_3), \phi_{\{2,3\}}(x_2, x_3), \phi_{\{2,4\}}(x_2, x_4)\}$. In this case, we see that x_2 and x_3 appear together in ψ_1 but $2-3$ is not already an edge in $\mathcal{H}_\Phi$. So we need to add this edge in to produce $\mathcal{G}_{\Phi,\prec}$. It can then be checked that continuing variable elmination with respect to this ordering adds in no other edges, so

$$\mathcal{G}_{\Phi,\prec} = ([4], \{1-2, 1-3, 2-3, 2-4, 3-4\}),$$

which has induced-with two. By the symmetry of $\mathcal{G}$, we also see that no matter which variable we eliminate first we will always add in a new edge in the first step of variable elimination. Hence, a lower bound on the treewidth of $\mathcal{G}$ is two. Since we have found an ordering $\prec$ that gives $\mathcal{G}_{\Phi,\prec}$ with induced-width two, we conclude that $\omega_\mathcal{G} = 2$.

---

2. Let $\mathcal{G} = ([m], E)$ be a DAG and let $\Phi = \{P(X_i \mid X_{\mathrm{pa}_\mathcal{G}(i)})\}_{i=1}^m$ be its set of factors. What is the relationship between $\mathcal{G}$ and $\mathcal{H}_\Phi$?

Solution:

Recall that $\mathcal{H}_\Phi$ is the undirected graph that contains the edge $i - j$ whenever x_i and x_j appear together in a factor in Φ. For our set of factors, x_i and x_j appear together in a factor in Φ if and only if $i \in \mathrm{pa}_\mathcal{G}(j)$, $j \in \mathrm{pa}_\mathcal{G}(i)$ or $i, j \in \mathrm{pa}_\mathcal{G}(k)$ for some $k \in [m] \setminus \{i, j\}$. Hence, $\mathcal{H}_\Phi$ is the moral graph of $\mathcal{G}$.

---

3. Show that every induced graph $\mathcal{G}_{\Phi,\prec}$ is chordal.

4. Let $\mathcal{G}$ be a connected, undirected graph. Show that $\mathcal{G}$ is a tree if and only if its treewidth is $\omega_\mathcal{G} = 1$.

Solution:

Suppose that $\mathcal{G}$ is an undirected tree. Then the set of maximal cliques in $\mathcal{G}$ are its edges. Hence, if $P(X_1, \dots, X_m)$ factors according to $\mathcal{G}$ then there exist nonnegative functions $\phi_{\{i,j\}}(x_i, x_j)$ for $i - j \in E$ such that

$$P(X_1, \dots, X_m) = \prod_{i-j \in E} \phi_{\{i,j\}}(x_i, x_j).$$

To prove that $\omega_\mathcal{G} = 1$, we need to show that there exists an ordering (permutation) $\prec$ of $X_1, \dots, X_m$ such that any intermediate factor produced by variable elimination does not produce any ψ which have x_i, x_j in its scope where $i - j \notin E$. This can be done inductively.

Suppose first that $G = ([2], \{1 - 2\})$. It is immediate that $\mathcal{G}$ has treewidth 1. Suppose now that the claim is true for all trees with $m-1$. Since $G = ([m], E)$ a tree, it has at least one leaf (i.e. a node with exactly one neighbor). Call such a leaf i^*, and call its unique neighbor j^*. Now pick an ordering $\prec$ that begins with the variable X_{i^*}. Then $\psi_1(x_{i^*}, x_{j^*}) = \phi_{\{i^*,j^*\}}(x_{i^*}, x_{j^*})$, and $\tau_1(x_{j^*}) = \sum_{x\ i^*} \psi_1(x_{i^*}, x_{j^*})$. It follows that the marginal distribution $\sum_{x_{i^*}} P(x_1, \ldots, x_m)$ factors according to $\mathcal{H} = ([m] \setminus \{i^*\}, E \setminus \{i^* - j^*\})$. Hence, by the inductive hypothesis $\mathcal{G}$ has treewidth 1.

Conversely, suppose that $\mathcal{G}$ has treewidth 1, but, for the sake of contradiction, $\mathcal{G}$ contains a cycle. Without loss of generality, suppose the cycle is the path $\langle 1, 2, \ldots, k, 1\rangle$. Let $\prec$ be an ordering of $X_1, \ldots, X_m$ and suppose that X_t is the first variable eliminated with respect to this ordering with $t \in [k]$. The ψ factor resulting from this elimination will be of the form

$$\psi = \gamma \cdot \phi_{\{t-1,t\}}(x_{t-1}, x_t)\phi_{\{t,t+1\}}(x_t, x_{t+1}),$$

and hence, at least three variables x_{t-1}, x_t, and x_{t+1} appear together in ψ (γ is the product of the remaining factors that contain x_t). It follows that there is at least a clique of size three in $\mathcal{G}_{\Phi,\prec}$ for any choice of $\prec$. Hence, the treewidth of $\mathcal{G}$ is at least two, which is a contradiction.

5. (Medical diagnosis again.) Suppose in our medical diagnosis example we assume that the distribution $\mathbf{X} = [Con, M, F, H, CC, C19, S]^T$ is Markov to the DAG presented in the lecture. We assume $\mathbf{X}$ abides by the following hierarchy:

$$\begin{array}{rcc}
S \sim (\{0,1,2,3\}) & C19 \sim (\theta_9) & \\
H|S = 0 \sim (\theta_1) & Con|H=0, CC=0, C19=0 \sim (\theta_{10}) & M|CC=0, C19=0 \sim (\theta_{18}) \\
H|S = 1 \sim (\theta_2) & Con|H=0, CC=0, C19=1 \sim (\theta_{11}) & M|CC=0, C19=1 \sim (\theta_{19}) \\
H|S = 2 \sim (\theta_3) & Con|H=0, CC=1, C19=0 \sim (\theta_{12}) & M|CC=1, C19=0 \sim (\theta_{20}) \\
H|S = 3 \sim (\theta_4) & Con|H=1, CC=0, C19=0 \sim (\theta_{13}) & M|CC=1, C19=1 \sim (\theta_{21}) \\
CC|S = 0 \sim (\theta_5) & Con|H=0, CC=1, C19=1 \sim (\theta_{14}) & F|CC=0, C19=0 \sim (\theta_{22}) \\
CC|S = 1 \sim (\theta_6) & Con|H=1, CC=0, C19=1 \sim (\theta_{15}) & F|CC=0, C19=1 \sim (\theta_{23}) \\
CC|S = 2 \sim (\theta_7) & Con|H=1, CC=1, C19=0 \sim (\theta_{16}) & F|CC=1, C19=0 \sim (\theta_{24}) \\
CC|S = 3 \sim (\theta_8) & Con|H=1, CC=1, C19=1 \sim (\theta_{17}) & F|CC=1, C19=1 \sim (\theta_{25})
\end{array}$$

We assume that, based on historical data, we have fit the model parameters as

$$\begin{aligned}
[\theta_1, \ldots, \theta_{25}]^T = [0.03, & 0.70, 0.48, 0.65, 0.73, \\
& 0.16, 0.22, 0.96, 0.81, 0.57, \\
& 0.56, 0.36, 0.01, 0.42, 0.32, \\
& 0.64, 0.12, 0.53, 0.57, 0.52, \\
& 0.11, 0.53, 0.59, 0.54, 0.06]^T.
\end{aligned}$$

Let $\mathbf{Y} = [S, Con, M, F]^T$. We observe for a new patient with the data $\mathbf{y} = [0, 1, 0, 1]^T$. Use variable elimination to compute the posterior expectations

$$[H|\mathbf{Y} = \mathbf{y}], \qquad [CC|\mathbf{Y} = \mathbf{y}], \qquad \text{and} \qquad [C19|\mathbf{Y} = \mathbf{y}].$$

Solution:

The conditional distributions specified above give us the complete conditional probability table for the joint distribution $\mathbf{X} = [Con, M, F, H, CC, C19, S]^T$ under the assumption that the distribution is markov to the DAG drawn in the lecture notes example (see slides). Hence, we can apply variable elimination, first to compute the unnormalized probability mass function

$$f_{Con,M,F,H,CC,C19,S}(1, 0, 1, h, cc, c19, 0).$$

From here, we then eliminate further to compute the unnormalized probability mass functions:

$$f_{Con,M,F,H,S}(1, 0, 1, h, 0), \quad f_{Con,M,F,CC,S}(1, 0, 1, cc, 0), \quad f_{Con,M,F,C19,S}(1, 0, 1, c19, 0).$$

We apply a second round of VE to compute the marginal probability

$$f_{Con,M,F,S}(1,0,1,0).$$

Dividing each of the first three results with the value $f_{Con,M,F,S}(1,0,1,0)$ yields the conditional distributions

$$H|\mathbf{Y}=\mathbf{y}, \qquad CC|\mathbf{Y}=\mathbf{y}, \qquad \text{and} \qquad C19|\mathbf{Y}=\mathbf{y}.$$

Since the variables are binary with outcomes $\{0,1\}$, the desired conditional expectations are the values of these distributions at $H=1$, $CC=1$ and $C19=1$, respectively.