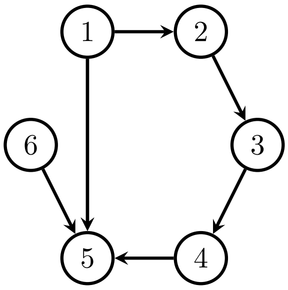
Graphical Models: Worksheet 3 MT 2023

A: Warm Up

A1. Directed Acyclic Graphs



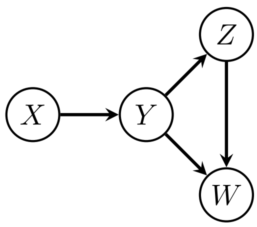
(a) List all the conditional independences implied by applying the local Markov property to the DAG $\mathcal{G}$ shown above.

(b) Find the moral graph. Is it decomposable?

(c) Find all the sets of vertices $C \subseteq \{2, 3, 5, 6\}$ such that $X_1 \perp\!\!\!\perp X_4 \mid X_C$ according to the global Markov property.

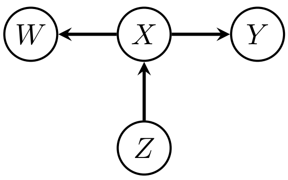
A2. Markov Equivalence

List all the graphs (either undirected or directed acyclic) that are Markov equivalent to the one shown.



A3. Junction Trees

Consider the graphical model shown.



(a) Draw a junction tree suitable for performing probability inference on a distribution that is Markov with respect to this graph.

(b) Suppose that the distribution is given by

$$p(z) = \begin{array}{cc} z=0 & 1 \\ \hline 0.4 & 0.6 \end{array} \qquad p(x \mid z) = \begin{array}{c|cc} z & x=0 & 1 \\ \hline 0 & 0.9 & 0.1 \\ 1 & 0.4 & 0.6 \end{array}$$

$$p(w \mid x) = \begin{array}{c|cc} x & w=0 & 1 \\ \hline 0 & 0.1 & 0.9 \\ 1 & 0.2 & 0.8 \end{array} \qquad p(y \mid x) = \begin{array}{c|cc} x & y=0 & 1 \\ \hline 0 & 0.7 & 0.3 \\ 1 & 0.4 & 0.6 \end{array}$$

Give an initialization of potentials in your junction tree consistent with this joint distribution. *[Hint: you shouldn't need to do any calculations.]*

(c) Using the junction tree algorithm, calculate the consistent potentials for this junction tree.

(d) Use the junction tree to compute $p(w \mid y = 1)$.

B: Core Questions

B1. Markov Blanket

Let $\mathcal{G}$ be a DAG. The *Markov blanket* of a vertex v is

$$\mathrm{mb}_{\mathcal{G}}(v) \equiv \mathrm{ch}_{\mathcal{G}}(v) \cup \mathrm{pa}_{\mathcal{G}}(\{v\} \cup \mathrm{ch}_{\mathcal{G}}(v)) \setminus \{v\}.$$

(i.e., the parents of v, children of v, and the other parents of children of v, but not v itself).

(a) Show that, in the moral graph $\mathcal{G}^m$, the boundary of v is precisely $\mathrm{bd}_{\mathcal{G}^m}(v) = \mathrm{mb}_{\mathcal{G}}(v)$.

(b) Deduce that if p is Markov with respect to $\mathcal{G}$ then

$$X_v \perp\!\!\!\perp X_{V \setminus (\mathrm{mb}(v) \cup \{v\})} \mid X_{\mathrm{mb}(v)} \; [p] \qquad\qquad v \in V. \qquad (1)$$

(c) Suppose p satisfies (1). Does this imply that p is Markov with respect to $\mathcal{G}$? Justify your answer.

B2. Structural Equation Models

Let $\mathcal{G}$ be a DAG and let X_V be a multivariate normal vector with zero mean and positive definite covariance matrix Σ.

(a) Let v be a vertex that has no children in $\mathcal{G}$, and denote $W = V \setminus \{v\}$. Show that

$$X_v \mid X_W = x_W \sim N(b_{vW} x_W, \; \Sigma_{vv \cdot W}).$$

where $\Sigma_{vv \cdot W} = \Sigma_{vv} - \Sigma_{vW}(\Sigma_{WW})^{-1}\Sigma_{Wv}$ is the Schur complement (see Worksheet 0, question 5) and $b_{vW} = (b_{vw})_{w \in W}$ is a vector which you should find.

(b) Show that Σ is Markov with respect to $\mathcal{G}$ if and only if both (i) Σ_{WW} is Markov with respect to $\mathcal{G}_W$ and (ii) $b_{vw} = 0$ for each $w \notin \mathrm{pa}_{\mathcal{G}}(v)$.

(c) Deduce that Σ is Markov with respect to $\mathcal{G}$ if and only if we can write

$$X_v = \sum_{w \in \mathrm{pa}_{\mathcal{G}}(v)} b_{vw} X_w + \varepsilon_v,$$

for all $v \in V$, where $\varepsilon_V = (\varepsilon_v)_{v \in V}$ is a Gaussian random vector with independent components. (Here an empty sum is zero by convention.)

(d) By writing the previous result in matrix form, show that

$$\Sigma = (I - B)^{-1} D (I - B)^{-T},$$

where I is the identity matrix, D is diagonal, and B is a lower triangular matrix with (i, j)th entry b_{ij}.

(e) Let $K = \Sigma^{-1}$ be the concentration matrix for X_V. Show that if $i \neq j$ then

$$k_{ij} = \sum_{\ell \in C_{ij}} d_{\ell\ell}^{-1} b_{\ell i} b_{\ell j} - d_{jj}^{-1} b_{ji} - d_{ii}^{-1} b_{ij}$$

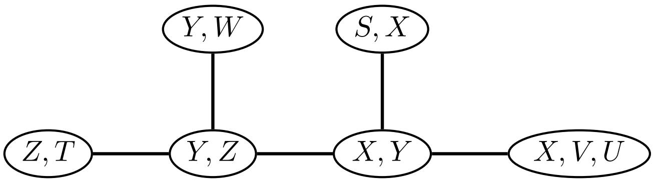
where $C_{ij} = \mathrm{ch}_{\mathcal{G}}(i) \cap \mathrm{ch}_{\mathcal{G}}(j)$. Deduce a graphical condition (i.e. a condition on $\mathcal{G}$) that will ensure $X_i \perp\!\!\!\perp X_j \mid X_{V \setminus \{i,j\}}$.

B3. Evidence Propagation

Let $\mathcal{T}$ be a junction tree with cliques $C_1, \ldots, C_k$, and suppose that all potentials are consistent.

(a) Let $e \in C_i$ and $f \in C_j$. Explain why the calculation of $p(x_f \mid \{X_e = y_e\})$ only requires messages to be passed from ψ_{C_i} along the (unique) path in $\mathcal{T}$ to ψ_{C_j}.

Suppose we have random variables S, T, U, V, W, X, Y, Z all taking values in $\{0, 1\}$, arranged in the junction tree below. Initially, the potentials are all consistent.



(b) How would you calculate $p(z = 0 \mid s = 1)$ in the most efficient way possible using the tree?

(c) How many additions and multiplications do you need to perform in order to calculate $p(z = 0 \mid s = 1)$ using (i) the method above; (ii) from the joint distribution directly?

(d) How would you calculate $p(t = 0 \mid s = 1, y = 1)$?

C: Optional

C1. Junction Tree Efficiency

Let $X_1, \ldots, X_k$ be binary random variables arranged in a junction tree with maximum clique size c and diameter d (the diameter is the length of the longest path in the tree). What is the maximum complexity (in terms of the number of additions, multiplications, divisions) required to calculate $p(x_i \,|\, x_j = 0)$? What about $p(x_i \,|\, x_j = 0, x_k = 0)$?

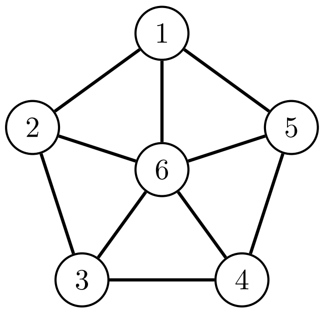
C2. Triangulation

The *Tarjan elimination algorithm* is a method for taking an undirected graph $\mathcal{G}$ and an ordering of the vertices of $\mathcal{G}$, and returning a triangulated graph $\mathcal{G}' \supseteq \mathcal{G}$.

1. Pick the largest element v of V under the ordering;
2. Join together all neighbours of v, and remove v from $\mathcal{G}$;
3. Repeat 1–2 until all vertices have been eliminated;
4. Construct a new graph which contains all the additional edges.

The ordering of the vertices used above is called an *elimination order*. An elimination order is said to be *perfect* if $\mathcal{G}' = \mathcal{G}$.

(a) Apply the algorithm to the graph below using the elimination orderings (i) $1, 2, 3, 4, 5, 6$; (ii) $6, 1, 2, 3, 4, 5$. (NB: the *last* vertex is eliminated first.) What are the resulting cliques?



(b) Show that the graph returned by the Tarjan Elimination algorithm is triangulated.

(c) Show that there exists a perfect elimination order if and only if $\mathcal{G}$ is triangulated.