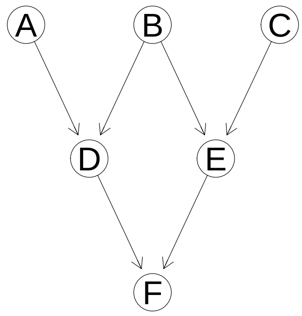
1. Consider the causal Bayesian network with variables A, B, C, D, E, F determined by A, B, C being mutually independent and binary with values $\{-1, 1\}$ and $P(A = 1) = P(B = 1) = P(C = 1) = 1/2$ and $D = AB$, $E = BC$, $F = DE$.

   (a) Draw the graph of the associated Bayesian network



   (b) Find $P(C = 1 \mid A = -1)$.
   C and A are marginally independent and hence

$$P(C = 1 \mid A = -1) = 1/2.$$

   (c) Find $P(E = 1 \mid F = 1, B = -1)$.
   This can be found by probability propagation, but can also be calculated semidirectly:

$$P(A, C, D, E \mid F = 1, B = -1) \propto$$
$$P(A)P(D \mid A, B = -1)P(C)P(E \mid C, B = -1)P(F = 1 \mid D, E).$$

   Adding over A and C yields

$$P(D, E \mid F = 1, B = -1) \propto$$
$$P(F = 1 \mid D, E).$$

   Finally adding over D yields this probability to be 1/2.
   The result can also be obtained directly by realising that, conditional on B, D and E are independent and uniformly distributed.

   (d) Find the intervention probability $P(E = 1 \mid F \leftarrow 1, B \leftarrow -1)$.
   It makes no change to E that F is set, so the setting of F can be ignored. Since B is a founder, $P(\cdot \mid B \leftarrow -1) = P(\cdot \mid B = -1)$ and hence

$$P(E = 1 \mid F \leftarrow 1, B \leftarrow -1) = P(E = 1 \mid B = -1) = 1/2.$$

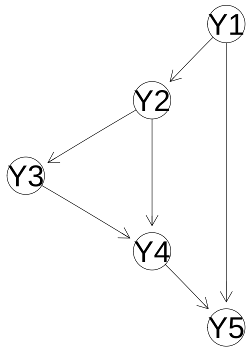
Clearly, this example is very special because of the massive symmetry in the specification of the network.

2. Let X_1, X_2, X_3, X_4, X_5 be independent with $X_i \sim \mathcal{N}(0, 1)$. Define recursively the structural equation system

$$Y_1 \leftarrow X_1, \quad Y_2 \leftarrow X_2 + Y_1, \quad Y_3 \leftarrow X_3 + Y_2, \quad Y_4 \leftarrow X_4 + Y_2 + Y_3, \quad Y_5 \leftarrow X_5 + Y_1 + Y_4$$

and assume intervention in the system is made by replacement, so the associated Bayesian network is causal

(a) Draw the causal DAG associated with this system



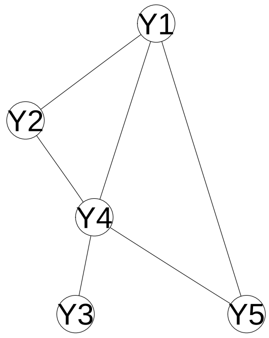
(b) Find the concentration matrix $K = \Sigma^{-1}$ of Y.
Using the recursive definition we find the joint density to be

$$f(y) \propto e^{-\{y_1^2+(y_2-y_1)^2+(y_3-y_2)^2+(y_4-y_3-y_2)^2+(y_5-y_4-y_1)^2\}/2}.$$

Expanding the squares and identifying coefficients we find the concentration matrix to be

$$K_Y = \begin{pmatrix} 3 & -1 & 0 & 1 & -1 \\ -1 & 3 & 0 & -1 & 0 \\ 0 & 0 & 2 & -1 & 0 \\ 1 & -1 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

(c) Construct the dependence graph of Y;
From the concentration matrix we find:



Note this is smaller than the moral graph, which also has a link between Y_2 and Y_3.

(d) Find the conditional distribution of Y_5 given $Y_3 = 0, Y_1 = 0$.

The concentration matrix of Y_2, Y_4, Y_5 conditional on Y_1 and Y_3 is

$$K_{245|13} = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

The variance of the conditional distribution of Y_5 is thus the lower corner of the inverse of this matrix, ie

$$V(Y_5 \mid Y_1, Y_3) = \frac{\det \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}}{\det K_{245|13}} = 5/2.$$

thus, $Y_5 \mid Y_1 = 0, Y_3 = 0 \sim \mathcal{N}(0, 5/2)$.

(e) Find the intervention distribution of Y_5 given $Y_3 \leftarrow 0, Y_1 \leftarrow 0$.

By intervention with replacement, this corresponds to the system

$$Y_1 \leftarrow 0, \quad Y_2 \leftarrow X_2, \quad Y_3 \leftarrow 0, \quad Y_4 \leftarrow X_4 + Y_2, \quad Y_5 \leftarrow X_5 + Y_4$$

Hence $V(Y_5) = 3$ in this system and thus

$$Y_5 \mid Y_3 \leftarrow 0, Y_1 \leftarrow 0 \sim \mathcal{N}(0, 3).$$

3. Let $\mathcal{A} = \mathcal{C}$ be the cliques of a chordal graph $\mathcal{G} = (V, E)$. For $A \subseteq V$ let $H(A)$ denote the entropy of X_A.

Show that

$$H(V) = \sum_{C \in \mathcal{C}} H(C) - \sum_{S \in \mathcal{S}} \nu(S) H(S)$$

where $\mathcal{S}$ are the minimal complete separators of $\mathcal{G}$ and $\nu(S)$ the number of times that the set S appears as an intersection between neighbouring cliques in a junction tree for $\mathcal{A}$.

For a junction tree with two cliques, this was proved in Problem 2, sheet 4.

Assume now the result is true for any junction tree with at most n cliques and let $\mathcal{T}$ be a junction tree with $n+1$ cliques. Let L be a *leaf* of $\mathcal{T}$, and let $\mathcal{C}' = \mathcal{C} \setminus \{L\}$ denote the remaining cliques. Similarly, let $\mathcal{T}'$ be the junction tree for $\mathcal{C}'$, obtained by removing the leaf L, and let $V' = \cup_{C \in \mathcal{C}'} C$.

The inductive hypothesis yields

$$H(V') = \sum_{C \in \mathcal{C}'} H(C) - \sum_{S \in \mathcal{S}'} \nu'(S) H(S). \tag{1}$$

Let now $S^* = L \cap V'$ be the separator in $\mathcal{T}$ associated with the leaf L. This S^* separates L from V' since $\mathcal{T}$ is a junction tree. Thus, using the result for $n = 2$ we get

$$H(V) = H(V') + H(L) - H(S^*). \tag{2}$$

Combining (1) with (2) using that

$$\nu(S) = \begin{cases} \nu'(S) & \text{if } S \neq S^* \\ \nu'(S^*) + 1 & \text{if } S = S^* \end{cases}$$

the result follows.

Note that this induction proof is a model for almost all proofs involving junction trees.

4. Consider the following directed acyclic graphs, and in each case, list all DAGs in their Markov equivalence class and verify in every single case whether they are Markov equivalent to an undirected graph.

   (a) $1 \to 2, 3 \to 2, 2 \to 4, 4 \to 5, 2 \to 5$;
   Arrow between 4 and 5 can be reversed. As 1 and 3 are unmarried parents, the DAG is not Markov equivalent to an undirected graph.

   (b) $1 \to 2, 2 \to 3, 2 \to 4, 4 \to 5, 6 \to 5$;
   The following DAGs are all equivalent to this:
   $2 \to 1, 2 \to 3, 2 \to 4, 4 \to 5, 6 \to 5$;
   $2 \to 1, 3 \to 2, 2 \to 4, 4 \to 5, 6 \to 5$;
   $2 \to 1, 2 \to 3, 4 \to 2, 4 \to 5, 6 \to 5$;
   None of them are equivalent to an undirected graph as 4 and 6 are unmarried parents.

(c) $1 \to 2, 2 \to 3, 2 \to 4, 4 \to 5, 5 \to 6$;

This is a tree, and is therefore Markov equivalent to the undirected version of the tree and also to any directed version obtained by choosing any of the six vertices as root and directing arrows away from the root.