SUTD 2021 50.007 Homework 5

James Raphael Tiovalen   1004555

Bayesian Networks

Question 1

Without knowing the actual value of any node, nodes $\mathbf{X}_1$ and $\mathbf{X}_6$ are **independent of each other**. This is because there does not exist any path from $\mathbf{X}_1$ to $\mathbf{X}_6$ that is open. Hence, based on the Bayes' ball algorithm, $\mathbf{X}_1$ and $\mathbf{X}_6$ are independent of each other.

If we know the value of nodes $\mathbf{X}_5$ and $\mathbf{X}_{10}$, nodes $\mathbf{X}_1$ and $\mathbf{X}_6$ would also still be **independent of each other**. This is because while the $\mathbf{X}_9 \to \mathbf{X}_{10} \to \mathbf{X}_9$ path is now open, there are no open paths at the $\mathbf{X}_5$ juncture from both $\mathbf{X}_7$ and $\mathbf{X}_8$ (i.e., both the paths from $\mathbf{X}_4$ to $\mathbf{X}_8$ and from $\mathbf{X}_4$ to $\mathbf{X}_7$ are closed). Thus, based on the Bayes' ball algorithm, $\mathbf{X}_1$ and $\mathbf{X}_6$ are still independent of each other.

Question 2

The number of parameters in a Bayesian network correspond to the number of entries in the probability table of each node in the network. Assuming that the number of values that node k can take is r_k, then for node i with parents pa_i, the number of rows is $\prod_{j \in pa_i} r_j$. The number of columns would be r_i. However, since the values in the last column can actually be uniquely determined from the other columns since the values of each row would sum up to 1, then for node i, there are actually $(r_i - 1) \prod_{j \in pa_i} r_j$ free/effective/independent parameters involved.

Therefore, in the initial Bayesian network, the number of effective parameters is:

$$
\begin{aligned}
&1(\mathbf{X}_1) + 2 \times 1(\mathbf{X}_2) + 2 \times 1(\mathbf{X}_3) + 2 \times 1(\mathbf{X}_4) + 2 \times 1(\mathbf{X}_5) + 1(\mathbf{X}_6) \\
&+ 2 \times 1(\mathbf{X}_7) + 2 \times 1(\mathbf{X}_8) + 2 \times 2 \times 2 \times 1(\mathbf{X}_9) + 2 \times 1(\mathbf{X}_{10}) + 2 \times 1(\mathbf{X}_{11}) \\
&= 26
\end{aligned}
$$

If nodes $\mathbf{X}_3$, $\mathbf{X}_8$ and $\mathbf{X}_9$ can take 5 different values, while all other nodes can take 4 different values, then the number of effective parameters would be:

$$
\begin{aligned}
&3(\mathbf{X}_1) + 4 \times 3(\mathbf{X}_2) + 4 \times 4(\mathbf{X}_3) + 5 \times 3(\mathbf{X}_4) + 4 \times 3(\mathbf{X}_5) + 3(\mathbf{X}_6) \\
&+ 4 \times 3(\mathbf{X}_7) + 4 \times 4(\mathbf{X}_8) + 4 \times 4 \times 5 \times 4(\mathbf{X}_9) + 5 \times 3(\mathbf{X}_{10}) + 4 \times 3(\mathbf{X}_{11}) \\
&= 436
\end{aligned}
$$

Question 3

(a) Variables can be independent either because of the topology of the network (structural independence) or because of their conditional probability table entries (numerical independence). Variables that are structurally independent are guaranteed to also be numerically independent, and thus, it is sufficient to simply check only for numerical independence. Because of this fact,

we can see from the probability tables that $\mathbf{X}_3$ and $\mathbf{X}_2$ are independent (and hence, $\mathbf{X}_3$ and $\mathbf{X}_1$ are also independent). Thus, to calculate the conditional probability, we can simply calculate:

$$
\begin{aligned}
P(\mathbf{X}_3 = 1 | \mathbf{X}_4 = 2) &= \frac{P(\mathbf{X}_3 = 1, \mathbf{X}_4 = 2)}{P(\mathbf{X}_4 = 2)} \\
&= \frac{P(\mathbf{X}_3 = 1)P(\mathbf{X}_4 = 2 | \mathbf{X}_3 = 1)}{P(\mathbf{X}_3 = 1)P(\mathbf{X}_4 = 2 | \mathbf{X}_3 = 1) + P(\mathbf{X}_3 = 2)P(\mathbf{X}_4 = 2 | \mathbf{X}_3 = 2)} \\
&= \frac{0.3 \times 0.9}{(0.3 \times 0.9) + (0.7 \times 0.5)} \\
&= \frac{27}{62} \approx 0.435
\end{aligned}
$$

(b) Following a similar argument as for part **(a)**, we can see from the probability tables that $\mathbf{X}_{10}$ and $\mathbf{X}_9$ are independent. Thus, $\mathbf{X}_{11}$ and $\mathbf{X}_9$ are independent, and following that, $\mathbf{X}_{11}$ and $\mathbf{X}_5$ are also independent. Since from part **(a)** we know that $\mathbf{X}_3$ and $\mathbf{X}_2$ are independent and that $\mathbf{X}_3$ and $\mathbf{X}_1$ are also independent, we can thus conclude that to calculate the conditional probability, we simply have to calculate:

$$
\begin{aligned}
P(\mathbf{X}_5 = 2 | \mathbf{X}_2 = 1, \mathbf{X}_{11} = 2, \mathbf{X}_1 = 1) &= P(\mathbf{X}_5 = 2) \\
&= P(\mathbf{X}_3 = 1)P(\mathbf{X}_4 = 1 | \mathbf{X}_3 = 1)P(\mathbf{X}_5 = 2 | \mathbf{X}_3 = 1, \mathbf{X}_4 = 1) \\
&+ P(\mathbf{X}_3 = 1)P(\mathbf{X}_4 = 2 | \mathbf{X}_3 = 1)P(\mathbf{X}_5 = 2 | \mathbf{X}_3 = 1, \mathbf{X}_4 = 2) \\
&+ P(\mathbf{X}_3 = 2)P(\mathbf{X}_4 = 1 | \mathbf{X}_3 = 2)P(\mathbf{X}_5 = 2 | \mathbf{X}_3 = 2, \mathbf{X}_4 = 1) \\
&+ P(\mathbf{X}_3 = 2)P(\mathbf{X}_4 = 2 | \mathbf{X}_3 = 2)P(\mathbf{X}_5 = 2 | \mathbf{X}_3 = 2, \mathbf{X}_4 = 2) \\
&= (0.3 \times 0.1 \times 0.5) + (0.3 \times 0.9 \times 0.4) \\
&+ (0.7 \times 0.5 \times 0.5) + (0.7 \times 0.5 \times 0.4) \\
&= 0.438
\end{aligned}
$$

Question 4

We can use the maximum likelihood estimation to find the optimal model parameters.

$$\theta_{\mathbf{X}_7}(1,1) = \frac{\text{Count}(\mathbf{X}_5 = 1, \mathbf{X}_7 = 1)}{\text{Count}(\mathbf{X}_5 = 1)} = 1/4$$

$$\theta_{\mathbf{X}_7}(1,2) = \frac{\text{Count}(\mathbf{X}_5 = 1, \mathbf{X}_7 = 2)}{\text{Count}(\mathbf{X}_5 = 1)} = 3/4$$

$$\theta_{\mathbf{X}_7}(2,1) = \frac{\text{Count}(\mathbf{X}_5 = 2, \mathbf{X}_7 = 1)}{\text{Count}(\mathbf{X}_5 = 2)} = 6/8 = 3/4$$

$$\theta_{\mathbf{X}_7}(2,2) = \frac{\text{Count}(\mathbf{X}_5 = 2, \mathbf{X}_7 = 2)}{\text{Count}(\mathbf{X}_5 = 2)} = 2/8 = 1/4$$

$$\theta_{\mathbf{X}_9}(1,1,1,1) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1)} = 1/3$$

$$\theta_{\mathbf{X}_9}(1,1,1,2) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1)} = 2/3$$

$$\theta_{\mathbf{X}_9}(1,1,2,1) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2)} = 1$$

$$\theta_{\mathbf{X}_9}(1,1,2,2) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2)} = 0$$

$$\theta_{\mathbf{X}_9}(1,2,1,1) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1)} = 1$$

$$\theta_{\mathbf{X}_9}(1,2,1,2) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1)} = 0$$

$$\theta_{\mathbf{X}_9}(1,2,2,1) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2)} = 1/2$$

$$\theta_{\mathbf{X}_9}(1,2,2,2) = \frac{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 1, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2)} = 1/2$$

$$\theta_{\mathbf{X}_9}(2,1,1,1) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1)} = 1$$

$$\theta_{\mathbf{X}_9}(2,1,1,2) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 1)} = 0$$

$$\theta_{\mathbf{X}_9}(2,1,2,1) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2)} = 0$$

$$\theta_{\mathbf{X}_9}(2,1,2,2) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 1, \mathbf{X}_8 = 2)} = 1$$

$$\theta_{\mathbf{X}_9}(2,2,1,1) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1)} = 1$$

$$\theta_{\mathbf{X}_9}(2,2,1,2) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 1)} = 0$$

$$\theta_{\mathbf{X}_9}(2,2,2,1) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2, \mathbf{X}_9 = 1)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2)} = 0$$

$$\theta_{\mathbf{X}_9}(2,2,2,2) = \frac{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2, \mathbf{X}_9 = 2)}{\text{Count}(\mathbf{X}_6 = 2, \mathbf{X}_7 = 2, \mathbf{X}_8 = 2)} = 1$$

The resulting probability tables for $\mathbf{X}_7$ and $\mathbf{X}_9$ are:

$\mathbf{X}_5$	$\mathbf{X}_7$	
	1	2
1	1/4	3/4
2	3/4	1/4

$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$	
			1	2
1	1	1	1/3	2/3
1	1	2	1	0
1	2	1	1	0
1	2	2	1/2	1/2
2	1	1	1	0
2	1	2	0	1
2	2	1	1	0
2	2	2	0	1