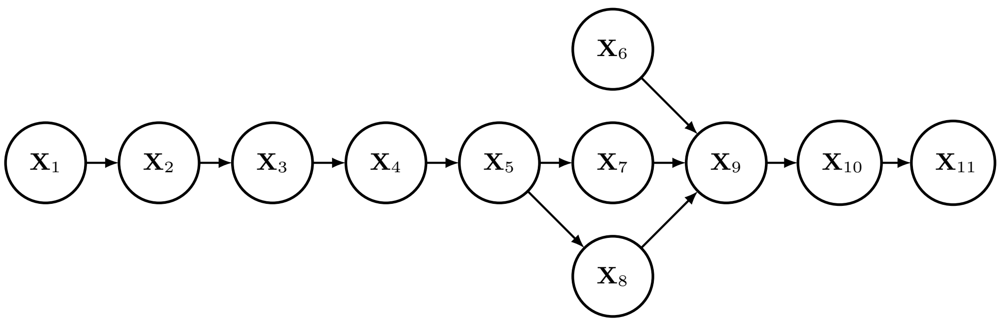
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

01.112 Machine Learning, Fall 2019
Homework 5

Due 13 Dec 2019, 5pm

This homework will be graded by Zihan Chen

In this homework, we would like to look at the Bayesian Networks. You are given a Bayesian network as below. All nodes can take 2 different values: $\{1, 2\}$.



Question 1. Without knowing the actual value of any node, are node $\mathbf{X}_2$ and $\mathbf{X}_6$ independent of each other? What if we know the value of node $\mathbf{X}_7$ and $\mathbf{X}_{11}$? *(5 points)*

Answer Without knowing the actual value of any node, node X_2 and X_6 are independent of each other. This is because there does not exist any path from X_2 to X_6 that is open. Based on the Bayes' ball algorithm, X_2 and X_6 are independent of each other.

If we know the value of node X_7 and X_{11}, then the two variables X_2 and X_6 become dependent. This is because there exist a path connecting X_2 and X_6 that is open: $X_2 - X_3 - X_4 - X_5 - X_8 - X_9 - X_{10} - X_{11} - X_{10} - X_9 - X_6$ or $X_2 - X_3 - X_4 - X_5 - X_7 - X_5 - X_8 - X_9 - X_{10} - X_{11} - X_{10} - X_9 - X_6$. Based on the Bayes' ball algorithm, X_2 and X_6 are dependent.

Question 2. What is the number of *free* parameters needed to for this Bayesian network? What would be the number of *free* parameters for the same network if node $\mathbf{X}_3$ and $\mathbf{X}_9$ can take 3 different values: $\{1, 2, 3\}$, and all other nodes can only take 5 different values: $\{1, 2, 3, 4, 5\}$? *(5 points)*

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

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

$1(X_1) + 2 \times 1(X_2) + 2 \times 1(X_3) + 2 \times 1(X_4) + 2 \times 1(X_5) + 1(X_6) + 2 \times 1(X_7) + 2 \times 1(X_8) + 2 \times 2 \times 2 \times 1(X_9) + 2 \times 1(X_{10}) + 2 \times 1(X_{11}) = 26$

If node X_3 and X_9 can take 3 different values: $\{1, 2, 3\}$, and all other nodes can only take 5 different values: $\{1, 2, 3, 4, 5\}$, the number of free parameters is:

$4(X_1) + 5 \times 4(X_2) + 5 \times 2(X_3) + 3 \times 4(X_4) + 5 \times 4(X_5) + 4(X_6) + 5 \times 4(X_7) + 5 \times 4(X_8) + 5 \times 5 \times 5 \times 2(X_9) + 3 \times 4(X_{10}) + 5 \times 4(X_{11}) = 392$

Question 3. If we have the following probability tables (next page) for the nodes. Compute the following probabilities. Clearly write down all the necessary steps.

	$\mathbf{X}_1$ = 1	$\mathbf{X}_1$ = 2
	0.5	0.5

$\mathbf{X}_1$	$\mathbf{X}_2$ = 1	$\mathbf{X}_2$ = 2
1	0.2	0.8
2	0.3	0.7

$\mathbf{X}_2$	$\mathbf{X}_3$ = 1	$\mathbf{X}_3$ = 2
1	0.3	0.7
2	0.3	0.7

$\mathbf{X}_3$	$\mathbf{X}_4$ = 1	$\mathbf{X}_4$ = 2
1	0.1	0.9
2	0.5	0.5

$\mathbf{X}_4$	$\mathbf{X}_5$ = 1	$\mathbf{X}_5$ = 2
1	0.5	0.5
2	0.6	0.4

	$\mathbf{X}_6$ = 1	$\mathbf{X}_6$ = 2
	0.6	0.4

$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$ = 1	$\mathbf{X}_9$ = 2
1	1	1	0.8	0.2
1	1	2	0.1	0.9
1	2	1	0.9	0.1
1	2	2	0.7	0.3
2	1	1	0.3	0.7
2	1	2	0.2	0.8
2	2	1	0.2	0.8
2	2	2	0.9	0.1

$\mathbf{X}_5$	$\mathbf{X}_7$ = 1	$\mathbf{X}_7$ = 2
1	0.2	0.8
2	0.3	0.7

$\mathbf{X}_5$	$\mathbf{X}_8$ = 1	$\mathbf{X}_8$ = 2
1	0.8	0.2
2	0.7	0.3

$\mathbf{X}_9$	$\mathbf{X}_{10}$ = 1	$\mathbf{X}_{10}$ = 2
1	0.8	0.2
2	0.8	0.2

$\mathbf{X}_{10}$	$\mathbf{X}_{11}$ = 1	$\mathbf{X}_{11}$ = 2
1	0.7	0.3
2	0.8	0.2

(a) Calculate the following conditional probability:

$$P(\mathbf{X}_3 = 1 | \mathbf{X}_4 = 1)$$

(6 points)

Answer One standard approach is to start by computing the following marginal probability:

$$P(X_3, X_4) = \sum_{X_1, X_2} P(X_1) P(X_2|X_1) P(X_3|X_2) P(X_4|X_3)$$

Compute $P(X_3 = 1, X_4 = 1)$ and $P(X_3 = 2, X_4 = 1)$ respectively, and then compute $P(X_4 = 1) = P(X_3 = 1, X_4 = 1) + P(X_3 = 2, X_4 = 1)$. The conditional probability $P(X_3 = 1 | X_4 = 1) = P(X_3 = 1, X_4 = 1)/P(X_4 = 1)$.

$$\begin{aligned}
P(X_3=1,X_4=1) &= P(X_1=1)P(X_2=1|X_1=1)P(X_3=1|X_2=1)P(X_4=1|X_3=1)\\
&+P(X_1=1)P(X_2=2|X_1=1)P(X_3=1|X_2=2)P(X_4=1|X_3=1)\\
&+P(X_1=2)P(X_2=1|X_1=2)P(X_3=1|X_2=1)P(X_4=1|X_3=1)\\
&+P(X_1=2)P(X_2=2|X_1=2)P(X_3=1|X_2=2)P(X_4=1|X_3=1)\\
&= 0.5\times 0.2\times 0.3\times 0.1+0.5\times 0.9\times 0.3\times 0.1+0.5\times 0.3\times 0.3\times 0.1\\
&+0.5\times 0.7\times 0.3\times 0.1 = 0.003+0.0135+0.0045+0.0105 = 0.0315
\end{aligned}$$

$$\begin{aligned}
P(X_3=2,X_4=1) &= P(X_1=1)P(X_2=1|X_1=1)P(X_3=2|X_2=1)P(X_4=1|X_3=2)\\
&+P(X_1=1)P(X_2=2|X_1=1)P(X_3=2|X_2=2)P(X_4=1|X_3=2)\\
&+P(X_1=2)P(X_2=1|X_1=2)P(X_3=2|X_2=1)P(X_4=1|X_3=2)\\
&+P(X_1=2)P(X_2=2|X_1=2)P(X_3=2|X_2=2)P(X_4=1|X_3=2)\\
&= 0.5\times 0.2\times 0.7\times 0.5+0.5\times 0.9\times 0.7\times 0.5+0.5\times 0.3\times 0.7\times 0.5\\
&+0.5\times 0.7\times 0.7\times 0.5 = 0.035+0.1575+0.0525+0.1225 = 0.3675
\end{aligned}$$

$$P(X_3=1|X_4=1) = \frac{P(X_3=1,X_4=1)}{P(X_4=1)} = \frac{0.0315}{0.0315+0.3675} \approx 0.079$$

(b) Calculate the following conditional probability:

$$P(\mathbf{X}_5=2|\mathbf{X}_3=2,\mathbf{X}_{11}=2,\mathbf{X}_1=2)$$

(9 points)

(*Hint: find a short answer. The values in some of the probability tables may reveal some useful information.*)

Answer We can have the following two observations from the tables:

- Distribution of X_3 doesn't change no matter what values X_2 takes.
- Distribution of X_{10} doesn't change no matter what values X_9 takes.

Thus, we have X_2 and X_3 are independent, X_9 and X_{10} are independent. There is no path connecting X_1 to X_5 and X_5 to X_{11}.

$$\begin{aligned}
P(X_5|X_3,X_{11},X_1) &= P(X_5|X_3)\\
&= \frac{P(X_3,X_5)}{P(X_3)}\\
&= \frac{\sum_{X_4} P(X_3)P(X_4|X_3)P(X_5|X_4)}{P(X_3)}\\
&= \sum_{X_4} P(X_4|X_3)P(X_5|X_4)
\end{aligned}$$

$$
\begin{aligned}
P(X_5 = 2 | X_3 = 2, X_{11} = 2, X_1 = 2) &= P(X_4 = 1 | X_3 = 2) P(X_5 = 2 | X_4 = 1) \\
&+ P(X_4 = 2 | X_3 = 2) P(X_5 = 2 | X_4 = 2) \\
&= 0.5 \times 0.5 + 0.5 \times 0.4 = 0.45
\end{aligned}
$$