50.007 Machine Learning HW 5

Victoria Yong 1004455

Q1.

x_1 and x_6 are independent.

x_1 and x_6 are independent given x_5

x_1 and x_6 have induced dependence given x_{10}

Q2.

No. Effective parameters = 1(x_1) + (2*1)(x_2) + (2*1)(x_3) + (2*1)(x_4) + (2*1)(x_5) + 1(x_6) + (2*1)(x_7) + (2*1)(x_8) + (2*2*2*1)(x_9) + (2*1)(x_{10}) + (2*1)(x_{11}) = 26

if node X_3, X_8 and X_9 can take 5 different values: {1, 2, 3, 4, 5}, and all other nodes can only take 4 different values: {1, 2, 3, 4},

No. effective parameters = 3(x_1) + (4*3)(x_2) + (4*4)(x_3) + (5*3)(x_4) + (4*3)(x_5) + 3(x_6) + (4*3)(x_7) + (4*4)(x_8) + (4*4*5*4)(x_9) + (5*3)(x_{10}) + (4*3)(x_{11}) = 436

Q3.

a.

$$
\begin{aligned}
P(x_3 = 1) &= P(x_1)P(x_2|x_1)P(x_3 = 1|x_2) \\
&= P(x_1 = 1)P(x_2 = 1|x_1 = 1)p(x_3 = 1|x_2 = 1) \\
&\quad + P(x_1 = 1)P(x_2 = 2|x_1 = 1)p(x_3 = 1|x_2 = 2) \\
&\quad + P(x_1 = 2)P(x_2 = 1|x_1 = 1)p(x_3 = 1|x_2 = 1) \\
&\quad + P(x_1 = 2)P(x_2 = 2|x_1 = 1)p(x_3 = 1|x_2 = 2) \\
&= 0.5(0.2)(0.3) + 0.5(0.8)(0.3) + 0.5(0.3)(0.3) + 0.5(0.7)(0.3) \\
&= 0.3 \\
P(x_4 = 2) &= P(x_4 = 2|x_3)P(x_3) \\
&= P(x_4 = 2|x_3 = 1)P(x_3 = 1) + P(x_4 = 2|x_3 = 2)P(x_3 = 2) \\
&= 0.3(0.9) + 0.7(0.5) \\
&= 0.62 \\
P(x_3 = 1, x_4 = 2) &= P(x_4 = 2|x_3 = 1)P(x_3 = 1) \\
&= 0.3(0.9) \\
&= 0.27 \\
P(x_3 = 1|x_4 = 2) &= \frac{P(x_3 = 1, x_4 = 2)}{P(x_4 = 2)} \\
&= \frac{0.27}{0.62} \\
&= 0.435
\end{aligned}
$$

b.

$$P(x_1 = 1, x_2 = 1\ ) = P(x_2 = 1|x_1 = 1)P(x_1 = 1) = 0.5(0.2) = 0.1$$

$$Since\ (x_1, x_2)\ are\ independent\ from\ x_{11},$$
$$P(x_1 = 1, x_2 = 1, x_5 = 2, x_{11} = 2) = P(x_1 = 1, x_2 = 1, x_5 = 2)P(x_{11} = 2)$$

$$\begin{aligned} P(x_1 = 1, x_2 = 1, x_5 = 2) &= P(x_5 = 2|x_1 = 1, x_2 = 1)P(x_1 = 1, x_2 = 1) \\ &= 0.1P(x_5 = 2|x_4)P(x_4|x_3)p(x_3|x_2 = 1) \\ &= 0.1[0.3(0.1)(0.5) + 0.7(0.5)^2 + 0.3(0.9)(0.4) + 0.7(0.5)(0.4)] \\ &= 0.0438 \end{aligned}$$

$$\begin{aligned} P(x_{11} = 2) &= P(x_{10})P(x_{11} = 2|x_{10}) \\ &= P(x_{10} = 1)P(x_{11} = 2|x_{10} = 1) + P(x_{10} = 2)P(x_{11} = 2|x_{10} = 2) \\ &= 0.8(0.3) + 0.2(0.2) \\ &= 0.28 \end{aligned}$$

$$P(x_1 = 1, x_2 = 1, x_{11} = 2) = P(x_1 = 1, x_2 = 1\ )P(x_{11} = 2) = 0.1(0.28) = 0.028$$

$$Hence,$$
$$\begin{aligned} P(x_1 = 1, x_2 = 1, x_5 = 2, x_{11} = 2) &= P(x_1 = 1, x_2 = 1, x_5 = 2)P(x_{11} = 2) \\ &= 0.0438(0.28) \\ &= 0.012264 \end{aligned}$$

$$P(x_5 = 2|x_2 = 1, x_{11} = 2, x_1 = 1) = \frac{P(x_1 = 1, x_2 = 1, x_5 = 2, x_{11} = 2)}{P(x_1 = 1, x_2 = 1, x_{11} = 2)} = \frac{0.012264}{0.028} = 0.438$$

Q4.

a. We can use the maximum likelihood estimation method for nodes X_7 and X_9.

$$\theta_{x_7}(1) = \frac{count(x_5 = 1, x_7 = 1)}{count(x_5 = 1)} = \frac{1}{4}$$
$$\theta_{x_7}(2) = \frac{count(x_5 = 1, x_7 = 2)}{count(x_5 = 1)} = \frac{3}{4}$$
$$\theta_{x_7}(1) = \frac{count(x_5 = 2, x_7 = 1)}{count(x_5 = 2)} = \frac{6}{8}$$
$$\theta_{x_7}(2) = \frac{count(x_5 = 2, x_7 = 1)}{count(x_5 = 2)} = \frac{2}{8}$$

$$\theta_{x_9}(1) = \frac{count(x_6 = 1, x_7 = 1, x_8 = 1, x_9 = 1)}{count(x_6 = 1, x_7 = 1, x_8 = 1)} = \frac{1}{3}$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 1, x_7 = 1, x_8 = 1, x_9 = 2)}{count(x_6 = 1, x_7 = 1, x_8 = 1)} = \frac{2}{3}$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 1, x_7 = 1, x_8 = 2, x_9 = 1)}{count(x_6 = 1, x_7 = 1, x_8 = 2)} = 1$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 1, x_7 = 1, x_8 = 2, x_9 = 2)}{count(x_6 = 1, x_7 = 1, x_8 = 2)} = 0$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 1, x_7 = 2, x_8 = 1, x_9 = 1)}{count(x_6 = 1, x_7 = 2, x_8 = 1)} = 1$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 1, x_7 = 2, x_8 = 1, x_9 = 2)}{count(x_6 = 1, x_7 = 2, x_8 = 1)} = 0$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 2, x_7 = 1, x_8 = 1, x_9 = 1)}{count(x_6 = 2, x_7 = 1, x_8 = 1)} = 1$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 2, x_7 = 1, x_8 = 1, x_9 = 2)}{count(x_6 = 2, x_7 = 1, x_8 = 1)} = 0$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 2, x_7 = 2, x_8 = 1, x_9 = 1)}{count(x_6 = 2, x_7 = 2, x_8 = 1)} = 1$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 2, x_7 = 2, x_8 = 1, x_9 = 2)}{count(x_6 = 2, x_7 = 2, x_8 = 1)} = 0$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 2, x_7 = 1, x_8 = 2, x_9 = 1)}{count(x_6 = 2, x_7 = 1, x_8 = 2)} = 0$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 2, x_7 = 1, x_8 = 2, x_9 = 2)}{count(x_6 = 2, x_7 = 1, x_8 = 2)} = 1$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 1, x_7 = 2, x_8 = 2, x_9 = 1)}{count(x_6 = 1, x_7 = 2, x_8 = 2)} = \frac{1}{2}$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 1, x_7 = 2, x_8 = 2, x_9 = 2)}{count(x_6 = 1, x_7 = 2, x_8 = 2)} = \frac{1}{2}$$
$$\theta_{x_9}(1) = \frac{count(x_6 = 2, x_7 = 2, x_8 = 2, x_9 = 1)}{count(x_6 = 2, x_7 = 2, x_8 = 2)} = 0$$
$$\theta_{x_9}(2) = \frac{count(x_6 = 2, x_7 = 2, x_8 = 2, x_9 = 2)}{count(x_6 = 2, x_7 = 2, x_8 = 2)} = 1$$

From the MLE, the node tables are

	$x_7 = 1$	$x_7 = 2$
$x_5 = 1$	¼	3/4
$x_5 = 2$	6/8	2/8

x_6	x_7	x_8	x_9	
			1	2
1	1	1	1/3	2/3
1	1	2	1	0
1	2	1	1	0
2	1	1	1	0
2	2	1	1	0
2	1	2	0	1
1	2	2	½	½
2	2	2	0	1