EI339 Artificial Intelligence Homework 3

Zhou Litao 518030910407 F1803016

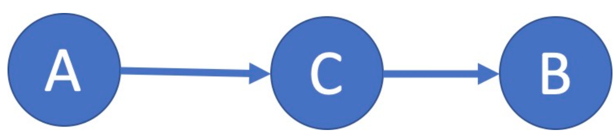
2020 年 12 月 4 日, Fall Semester

Exercise 1 (BAYES NETWORK) Evaluate the distributions p(a),p(b|c),and p(c|a) corresponding to the joint distribution given in Table 1. Hence show by direct evaluation that p(a,b,c) = p(a)p(c|a)p(b|c). Draw the corresponding directed graph (30)

a	b	c	$p(a,b,c)$
0	0	0	0.192
0	0	1	0.144
0	1	0	0.048
0	1	1	0.216
1	0	0	0.192
1	0	1	0.064
1	1	0	0.048
1	1	1	0.096

Table 1 The joint distribution over three binary variables.

Solution. The directed graph is shown below, and the distributions are listed on the next page.



□

a	p(a)
0	0.6
1	0.4

b	c	p(b|c)
0	0	0.8
1	0	0.2
0	1	0.4
1	1	0.6

a	c	p(c|a)
0	0	0.4
0	1	0.6
1	0	0.6
1	1	0.4

a	b	c	p(a)	p(c|a)	p(b|c)	p(a,b,c)
0	0	0	0.6	0.4	0.8	0.192
0	0	1	0.6	0.6	0.4	0.144
0	1	0	0.6	0.4	0.2	0.048
0	1	1	0.6	0.6	0.6	0.216
1	0	0	0.4	0.6	0.8	0.192
1	0	1	0.4	0.4	0.4	0.064
1	1	0	0.4	0.6	0.2	0.048
1	1	1	0.4	0.4	0.6	0.096

Exercise 2 (Markov Model) Use the technique of d-separation, to verify that the Markov model shown in Figure 1 having N nodes in total satisfies the conditional independence properties

$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})$$

for n = 2,.,N. Similarly, show that a model described by the graph in Figure 2 in which there are N nodes in total (30)

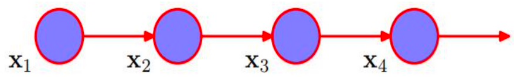


图 1: A first-order Markov chain of observations $\{x_n\}$ in which the distribution $p(x_n \mid x_{n-1})$ of a particular observation x_n is conditioned on the value of the previous observation x_{n-1}

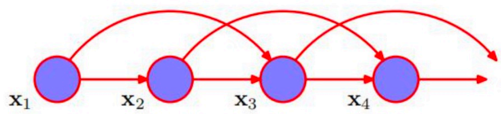


图 2: A second-order Markov chain, in which the conditional distribution of a particular observation x_n depends on the values of the two previous observations x_{n-1} and x_{n-2}.

Solution. For the first formula, we can check for every X_i,$i < N-1$, the only undirected path to X_N will necessarily pass through X_{N-1}. Thus if we set X_{N-1} to be evidence, there will be no active path for any X_i, $i < N-1$ to reach X_N. Thus $X_N \perp X_1 \ldots X_{N-2} | X_{N-1}$. i.e.

$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})$$

For second-order Markov Chain, we assert that $X_N \perp X_{N-3}, \ldots, X_1 | X_{N-2}, X_{N-1}$, i.e.

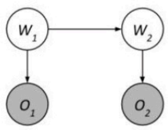
$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2})$$

□

To prove this, for every $X_i, i < N-2$, we check the undirected path in the graph between X_i and X_N. Note that it must pass through either X_{N-1} or X_{N-2}, forming a casual relation. Since X_{N-1} and X_{N-2} colored grey, all the paths between X_i and X_N are inactive. Thus, by d-separation, the assertion holds.

Exercise 3 (HMM) Consider the following Hidden Markov Model. O_1 and O_2 are supposed to be shaded. Suppose that we observe $O_1 = a$ and $O_2 = b$ Using the forward algorithm, compute the probability distribution $P(W_2 \mid O_1 = a, O_2 = b)$ one step at a time. (40)

1. Compute $\mathrm{P}(W_1, O_1 = a)$.
2. Using the previous calculation,compute $\mathrm{P}(\mathrm{W}_2, \mathrm{O}_1 = a)$
3. Using the previous calculation,compute $\mathrm{P}(\mathrm{W}_2, \mathrm{O}_1 = a, \mathrm{O}_2 = b)$
4. Finally, compute $P(W_2 \mid O_1 = a, O_2 = b)$.



W_1	$P(W_1)$
0	0.7
1	0.3

W_t	W_{t+1}	$P(W_{t+1} \mid W_t)$
0	0	0.9
0	1	0.1
1	0	0.5
1	1	0.5

W_t	O_t	$P(O_t \mid W_t)$
0	a	0.4
0	b	0.6
1	a	0.8
1	b	0.2

Solution.

1. $P(W_1 = 0, O_1 = a) = P(W_1 = 0) \times P(O_1 = a|W_1 = 0) = 0.4 \times 0.7 = 0.28$

   $P(W_1 = 1, O_1 = a) = P(W_1 = 1) \times P(O_1 = a|W_1 = 1) = 0.3 \times 0.8 = 0.24$

2. $$\begin{aligned}&P(W_2 = 1, O_1 = a)\\&= P(W_2 = 1|O_1 = a) \times P(O_1 = a)\\&= P(W_2 = 1|W_1 = 1) \times P(W_1 = 1, O_1 = a) + P(W_2 = 1|W_1 = 0) \times P(W_1 = 0, O_1 = a)\\&= 0.5 \times 0.24 + 0.1 \times 0.28 = 0.148\end{aligned} \tag{1}$$

   $$\begin{aligned}&P(W_2 = 0, O_1 = a)\\&= P(W_2 = 0|O_1 = a) \times P(O_1 = a)\\&= P(W_2 = 0|W_1 = 1) \times P(W_1 = 1, O_1 = a) + P(W_2 = 0|W_1 = 0) \times P(W_1 = 0, O_1 = a)\\&= 0.5 \times 0.24 + 0.9 \times 0.28 = 0.372\end{aligned} \tag{2}$$

3. $$\begin{aligned}&P(W_2 = 0, O_1 = a, O_2 = b)\\&= P(O_2 = b|W_2 = 0, O_1 = a) \times P(W_2 = 0, O_1 = a)\\&= P(O_2 = b|W_2 = 0) \times P(W_2 = 0, O_1 = a)\\&= 0.6 \times 0.372 = 0.2232\end{aligned} \tag{3}$$

   $$\begin{aligned}&P(W_2 = 1, O_1 = a, O_2 = b)\\&= P(O_2 = b|W_2 = 1, O_1 = a) \times P(W_2 = 1, O_1 = a)\\&= P(O_2 = b|W_2 = 1) \times P(W_2 = 1, O_1 = a)\\&= 0.2 \times 0.148 = 0.0296\end{aligned} \tag{4}$$

4.

$$P(W_2 = 0|O_1 = a, O_2 = b) = \frac{0.2232}{0.2232 + 0.0296} \approx 0.8829 \tag{5}$$

$$P(W_2 = 1|O_1 = a, O_2 = b) = \frac{0.0296}{0.2232 + 0.0296} \approx 0.1171 \tag{6}$$

□