

## 6.825 Homework 3: Solutions

### 1 Easy EM

You are given the network structure shown in Figure 1 and the data in the following table, with actual observed values for A, B, and C, and expected counts for D.

A	B	C	$\Pr(D A, B, C)$
1	1	1	.6
0	0	0	.3
0	0	1	.7
0	1	1	.9
1	1	1	.6
0	0	0	.3
1	0	1	.5
1	1	0	.4
0	1	1	.9
0	0	1	.7
0	0	1	.7
0	1	0	.1
0	1	1	.9
0	1	1	.9
1	0	0	.2
1	0	1	.5
1	1	1	.6

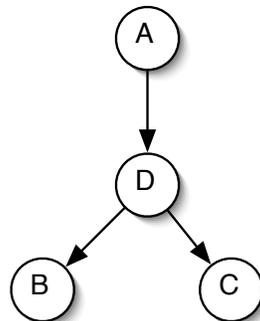


Figure 1: Network Structure

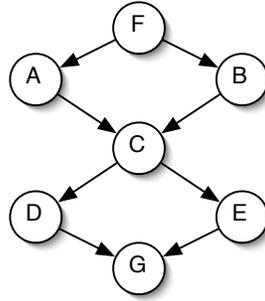


Figure 2: Network Structure

1. What are the maximum likelihood estimates for  $\Pr(A)$ ,  $\Pr(D|A)$ ,  $\Pr(B|D)$ , and  $\Pr(C|D)$ ?

**Solution:**

$$\Pr(A) = \frac{7}{17} = 0.412$$

$$\Pr(D|A) = \frac{E\#[D \wedge A]}{E\#[A]} = \frac{3.4}{7} = 0.486$$

$$\Pr(B|D) = \frac{E\#[B \wedge D]}{E\#[D]} = \frac{5.9}{9.8} = 0.602$$

$$\Pr(C|D) = \frac{E\#[C \wedge D]}{E\#[D]} = \frac{8.5}{9.8} = 0.867$$

2. What's another way to represent this data in a table with only 8 entries?

**Solution:**

Use a table that tallies the occurrences of each data sample.

A	B	C	$\Pr(D a, b, c, d)$	Number of Occurrences
0	0	0	.3	2
0	0	1	.7	3
0	1	0	.1	1
0	1	1	.9	4
1	0	0	.2	1
1	0	1	.5	2
1	1	0	.4	1
1	1	1	.6	3

## 2 Harder EM

Consider the Bayesian network structure shown in Figure 2.

1. Write down an expression for  $\Pr(a, b, c, d, e, f, g)$  using only conditional probabilities that would be stored in that network structure.

**Solution:**

$$\Pr(a, b, c, d, e, f, g) = \Pr(f) \Pr(a|f) \Pr(b|f) \Pr(c|a, b) \Pr(d|c) \Pr(e|c) \Pr(g|e, d)$$

2. We'd like to use variable elimination to compute  $\Pr(d|b)$ . Which variables are irrelevant?

**Solution:**

Variables E and G are irrelevant. If we try to sum them out, we get factors that are identical to 1.

3. Using elimination order A, B, C, D, E, F, G to compute  $\Pr(d|b)$ , what factors get created?

**Solution:**

We've already established that we can ignore E and G. And since we're naming specific values d and b in our query, we really only need to sum over A, C, and F. Summing over F gives us factor

$$f_1(a, b) = \sum_F \Pr(f) \Pr(a|f) \Pr(b|f) .$$

Summing over C gives us factor

$$f_2(a, b, d) = \sum_C \Pr(c|a, b) \Pr(d|c) .$$

Summing over A gives us

$$\Pr(b, d) = \sum_A f_1(a, b) f_2(a, b, d) .$$

To get the conditional probability we want, we first need to compute

$$\Pr(b) = \sum_D \Pr(b, d) ;$$

then  $\Pr(d|b) = \Pr(b, d) / \Pr(b)$ .

4. Assume variables F and C are hidden. You have a data set consisting of vectors of values of the other variables:  $\langle a, b, d, e, g \rangle$ . You start with an initial set of parameters,  $\theta_0$ , that contains CPTs for the whole network. You decide to estimate  $\Pr(f|a, b, d, e, g)$  and  $\Pr(c|a, b, d, e, g)$  separately, for each row of the table, as shown.

A	B	D	E	G	$\Pr(F a, b, d, e, g)$	$\Pr(C a, b, d, e, g)$
0	1	0	1	1		
0	0	0	1	1		
0	0	0	0	1		
0	1	0	1	1		
0	0	0	0	1		
...						

Is this justified? Why or why not?

**Solution:**

I should have said, above, "you decide to calculate" rather than "you decide to estimate" above, since we calculate these values directly from the data in each row and the model.

The answer is, in this case, yes, but not always. In particular, because F and C are independent given A and B, we can write

$$\Pr(F, C|a, b, d, e, g) = \Pr(F|a, b) \Pr(C|a, b, d, e) ,$$

and so it's okay to fill in values for F and C independently given the other variables. (Note that we can save time by figuring out  $\Pr(F|a, b)$  for every value of  $a$  and  $b$ , and using those values to fill in the whole column.)

If, instead, we had A and B as the hidden variables, we would have to estimate their values jointly, since they are not independent given values of all the other variables. Thus, we'd have to make a table that looked something like:

C	D	E	F	G	A	B	$\Pr(a, b c, d, e, f, g)$
0	1	0	1	1	0	0	.4
					0	1	.3
					1	0	.2
					1	1	.1
0	0	0	1	1	0	0	.05
					0	1	.1
					1	0	.05
					1	1	.8
...							

This table contains the conditional joint of A and B, for each actually observed data point. You could also make it more compact by collapsing all duplicate data points and counting them, as we did above.

- Now you think of another shortcut: You decide to estimate  $\Pr(G|d, e)$  directly from the data and leave G out of the EM computations all together. Is this justified?

**Solution:**

Yes.

$$\Pr(G|a, b, c, d, e, f) = \Pr(G|d, e).$$

EM starts with some values in the CPT, and then it loops for a while, alternately

- estimating the counts for the missing data from the CPT, and then
- re-estimating the CPT given the new estimated counts for the missing data.

But, G's CPT entry does not depend on the missing data, F and G; it only depends on D and E. So, it's ok to calculate its CPT parameters once, and then leave them out of the EM loop.

### 3 Markov Chain

Consider a 3-state Markov chain, where  $R(s_1) = 1$ ,  $R(s_2) = 3$ , and  $R(s_3) = -1$ . The transition probabilities are given by the table below, where the entry in row  $i$ , column  $j$ , is the probability of making a transition from  $s_i$  to  $s_j$ .

	s <sub>1</sub>	s <sub>2</sub>	s <sub>3</sub>
s <sub>1</sub>	0.8	0.1	0.1
s <sub>2</sub>	0.2	0.1	0.7
s <sub>3</sub>	0.9	0.1	0.0

1. Compute the values of each state, assuming a discount factor of  $\gamma = 0.9$ .
2. Compute the values with a discount factor of  $\gamma = 0.1$ .

**Solution:**

Writing out the full equation for the values of each state we get:

$$\begin{aligned} v_1 &= 1 + 0.8\gamma v_1 + 0.1\gamma v_2 + 0.1\gamma v_3 \\ v_2 &= 3 + 0.2\gamma v_1 + 0.1\gamma v_2 + 0.7\gamma v_3 \\ v_3 &= -1 + 0.9\gamma v_1 + 0.1\gamma v_2 \end{aligned}$$

Or, in matrix form:

$$\begin{pmatrix} 0.8\gamma - 1 & 0.1\gamma & 0.1\gamma \\ 0.2\gamma & 0.1\gamma - 1 & 0.7\gamma \\ 0.9\gamma & 0.1\gamma & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$$

With  $\gamma = 0.9$  and  $0.1$ , respectively, we find:

$$\begin{aligned} v_1 &= 9.26, & 1.11 \\ v_2 &= 10.27, & 2.99 \\ v_3 &= 7.42, & -0.87 \end{aligned}$$

## 4 Markov Decision Process

Now, consider a Markov decision process with the same states and rewards as in the previous problem. It has two actions. The first action is described by the transition matrix given above. The second action is described by the following transition matrix:

	s <sub>1</sub>	s <sub>2</sub>	s <sub>3</sub>
s <sub>1</sub>	0.1	0.1	0.8
s <sub>2</sub>	0.7	0.1	0.2
s <sub>3</sub>	0.9	0.0	0.1

Starting with a value function that is 0 for all states, perform value iteration for 10 iterations for  $\gamma = 0.9$ . Do it again for  $\gamma = 0.1$ . Can you see it converging faster in one case? If so, why? Can you see what the optimal policy is? How could you compute the values of the states under the optimal policy?

**Solution:**

The solution can be implemented in a few lines of matlab:

```

g = 0.9; %or 0.1
r = [1; 3; -1];
v = [0; 0; 0];
a1 = [0.8, 0.1, 0.1; 0.2, 0.1, 0.7; 0.9, 0.1, 0];
a2 = [0.1, 0.1, 0.8; 0.7, 0.1, 0.2; 0.9, 0.0, 0.1];
for i = 1:10,
    v = r + g*max(a1*v, a2*v);
end,

```

With  $\gamma = 0.9$  and  $0.1$ , respectively, we find after ten iterations:

$$\begin{aligned}
v_1 &= 6.51, & 1.11 \\
v_2 &= 8.35, & 3.09 \\
v_3 &= 4.68, & -0.87
\end{aligned}$$

We see that when  $\gamma = 0.1$ , the values for each state are greater than or equal to those found in problem 3. We should expect this, since value iteration finds a global optimum; the policy of always choosing the first action would give exactly the discounted rewards found in problem 3, and the optimal policy will perform better than or equal to that.

So why isn't this true when  $\gamma = 0.9$ ? The algorithm hasn't converged yet. Running for 1000 iterations, we find the values converge to:

$$\begin{aligned}
v_1 &= 10.0 \\
v_2 &= 11.83 \\
v_3 &= 8.16
\end{aligned}$$

These values *are* better than those found in problem 3 for  $\gamma = 0.9$ . Intuitively, the algorithm converges quickly with smaller values of  $\gamma$  because we are putting less of an emphasis on future rewards and so don't have to consider as many time steps to get a good estimate. Concretely, the value of future rewards decays to zero much more rapidly when  $\gamma = 0.1$ .

To compute the optimal policy, it is only necessary to see which action gives the highest expected value at each state once value iteration has converged. For this problem, the optimal actions are 1, 2, 1, for states  $s_1$ ,  $s_2$ , and  $s_3$  respectively. So, to compute the optimal values exactly we can treat this problem as a Markov Chain with transition probabilities:

	$s_1$	$s_2$	$s_3$
$s_1$	0.8	0.1	0.1
$s_2$	0.7	0.1	0.2
$s_3$	0.9	0.1	0.0

Solving this system as in the previous problem will produce the same values as value iteration.

## 5 At the Races, again

You go to the horse races, hoping to win some money. Your options are to:

- Bet \$2 on Moon. You think that he'll win with probability 0.7. If he wins, you'll get back \$4. If he loses, you'll get back nothing.

- Bet \$2 on Jeb. You think that he'll win with probability 0.2. If he wins, you'll get back \$22. If he loses, you'll get back nothing.
- Bet \$1 on Moon and \$1 on Jeb. With probability 0.7, Moon will win and you'll get back \$2. With probability 0.2, Jeb will win, and you'll get back \$11. If they both lose, you'll get back nothing.

1. What's the expected value of each bet? What action would a risk-neutral bettor choose?

**Solution:**

Expected Values:

- Bet \$2 on Moon:  $.7 \cdot U(2) + .3 \cdot U(-2) = 0.8$
- Bet \$2 on Jeb:  $.2 \cdot U(20) + .8 \cdot U(-2) = 2.4$
- Bet \$1 on Moon and \$1 on Jeb:  $.7 \cdot U(0) + .2 \cdot U(9) + .1 \cdot (-2) = 1.6$

So a risk-neutral bettor would choose the second option, putting all their money on Jeb.

2. Now, consider a risk-averse bettor, whose utility function for money is

$$U(x) = \sqrt{x + 20} .$$

What are the expected utilities of each bet? What action would that person choose?

**Solution:**

Expected utilities:

- Bet \$2 on Moon: 4.56
- Bet \$2 on Jeb: 4.66
- Bet \$1 on Moon and \$1 on Jeb: 4.63

This person would choose the second option, as well, but it's very close to the other options.

If we changed the utility function to be  $U(x) = \sqrt{x + 3}$ , the values for the three bets are now 1.87, 1.76, and 2.01. So now, we'd prefer option 3.

3. How do the outcomes of the previous parts relate to investment strategy?

**Solution:**

Investment strategy should differ depending on a person's utility function.

It's common to reduce risk by simultaneously betting on a variety of different options. This is called "hedging one's bets," and it's why, when we're really risk averse (the last utility function simulates having only \$3 in our pockets), we hedge.