15-381 Artificial Intelligence: Representation and Problem Solving
Homework 3 - Solutions

1 [15 pts] Maximum Likelihood

Our favourite robot, R2D2, has been sent on a mission to the outer world to locate the latest weapon of the dark lord. On its way, it gets partly damaged by an enemy ship damaging its encryption mechanism.

R2D2 wants to send out the range of the enemy weapon. Since he is worried that the message might be intercepted, he instead sends a series of randomly generated values $y_1, y_2, y_3, \dots y_m$ which are related to ϕ, the actual range of the weapon in the following manner: $P(Y = y|\phi) = \phi e^{-\phi y}$($y$'s are always positive).

Unfortunately, the Jedi knights, powerful though they might be, are somewhat lacking in their knowledge of math and ask the good students of 15-381 to help them out in this task. Can you help them defeat the dark lord?

1. Write down the likelihood(probability) of the data $y_1 \dots y_m$ assuming they were independently generated from the distribution.

2. Determine the value of ϕ that maximizes the likelihood of the data, the is the MLE estimate. Recall from the discussion in the class that the MLE estimate of a parameter is the value of the parameter that maximizizes the likelihood of the data. One way to obtain the maximum value of a function, is to solve for its derivative being zero, estimate the value of the function at that point. This is guaranteed only to give you a stationary point(it could be a minimum point for example). You'll need to then show that it is a maxima and not a minima.

   Warning: The Jedi can read your mind, but the TAs are somewhat lacking in that aspect(in other words, show us your working :))

Solution

1. Likelihood L $= \prod_{i=1}^{m} \phi e^{-\phi y_i} = \phi^m e^{-\phi \sum_i y_i}$

2. $\frac{dL}{d\phi} = m\phi^{m-1} e^{-\phi \sum_i y_i} - \phi^m (\sum_i y_i) e^{-\phi \sum_i y_i} = 0 \Rightarrow \phi = 0$, or, $\phi = m/(\sum_i y_i)$

   To see that this is a maximum, it is enough to note that $\frac{dL}{d\phi} > 0$ when $\phi < m/(\sum_i y_i)$ and $\frac{dL}{d\phi} < 0$ when $\phi > m/(\sum_i y_i)$. Thus, L is a decreasing function to the right of the solution and an increasing function to the left implying that the stationary point is a point of maxima. An alternate but equivalent way of showing this is to compute the second derivative and show that it is negative at this point.

2 [15 pts] HMM miscellany

1. Derive the update formula that was shown in class for $\beta_t(i)$.

$$\beta_t(i) = P(O_{t+1}, \dots O_T | q_t = s_i) = \sum_j a_{j,i} b_j(O_{t+1}) \beta_{t+1}(j)$$

   Use $\beta_T(i)$ =1 for all i to initiate the function (T is the total number of values observed.)

Hint: Sum over all possible states in time t+1 and then use Bayes rule.

2. Show how to derive the equation for $S_t(i)$ (the probability of being in state i at time t) which was used in class as part of the EM algorithm.

$$S_t(i) \triangleq P(q_t = s_i | O_1, \ldots O_T) = \frac{\alpha_t(i)\beta_t(i)}{\sum_j \alpha_t(j)\beta_t(j)}$$

Hint: Use chain rule as shown on slide. Then use chain rule again for the numerator. Note that the denominator is a sum of the numerator for all possible states.

Solution

1.

$$\beta_t(i) = P(O_{t+1}, \ldots O_T | q_t = s_i) = \sum_j P(q_{t+1} = s_j, O_{t+1}, \ldots O_T | q_t = s_i)$$

(Using $P(A) = \sum_b P(A, b)$)

$$= \sum_j P(O_{t+1}, \ldots O_T | q_{t+1} = s_j, q_t = s_i) P(q_{t+1} = s_j | q_t = s_i)$$

(Using chain rule of the form $P(A, B|C) = P(B|A, C)P(B|C)$)

$$= \sum_j P(O_{t+1}, \ldots O_T | q_{t+1} = s_j) P(q_{t+1} = s_j | q_t = s_i)$$

(Using markov property: $O_{t+1}, \ldots O_T \perp q_t | q_{t+1}$)

$$= \sum_j P(O_{t+2}, \ldots O_T | O_{t+1}, q_{t+1} = s_j) P(O_{t+1} | q_{t+1} = s_j) P(q_{t+1} = s_j | q_t = s_i)$$

(Using chain rule of the form $P(A, B|C) = P(B|A, C)P(B|C)$ with $B = O_{t+1}$)

$$= \sum_j P(O_{t+2}, \ldots O_T | q_{t+1} = s_j) P(O_{t+1} | q_{t+1} = s_j) P(q_{t+1} = s_j | q_t = s_i)$$

(Using markov property: $O_{t+2}, \ldots O_T \perp O_{t+1} | q_{t+1}$)

$$= \sum_j a_{j,i} b_j(O_{t+1}) \beta_{t+1}(j)$$

(Using definitions of $a_{j,i}, b_j(O_{t+1})\beta_{t+1}(j)$)

2.

$$S_t(i) \triangleq P(q_t = s_i | O_1, \ldots O_T) = P(q_t = s_i, O_1, \ldots O_T)/P(O_1, \ldots O_T)$$

(Using definition of conditional distribution: $P(A|B) = P(A, B)/P(B)$)

$$= P(q_t = s_i, O_1, \ldots O_T)/(\sum_i P(q_t = s_i, O_1, \ldots O_T))$$

(Using $P(A) = \sum_b P(A, b)$ with $B = q_t$)

$$Numerator = P(q_t = s_i, O_1, \ldots O_T) = P(q_t = s_i, O_1, O_2, \ldots O_t, O_{t+1} \ldots O_T)$$

$$= P(O_{t+1} \dots O_T | q_t = s_i, O_1, O_2, \dots O_t) P(q_t = s_i, O_1, O_2, \dots O_t)$$

(Using chain rule)

$$= P(O_{t+1} \dots O_T | q_t = s_i) P(q_t = s_i, O_1, O_2, \dots O_t)$$

(Using markov property: $O_{t+1}, \dots O_T \perp O_1 \dots O_t | q_t$)

$$= \alpha_t(i)\beta_t(i)$$

Therefore, $S_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_j \alpha_t(j)\beta_t(j)}$

3 [15 pts] Value Iteration

Consider an MDP with 6 states: Bathroom, Kitchen, Bedroom, Dining Room, Under-Attack, and Dead. This represents a domain of a robotic rat foraging for food in a house with four rooms (Bathroom, Kitchen, Bedroom, and Dining Room). In states Bathroom, Kitchen, Bedroom, and Dining Room, there are three available actions: stay in place (S), move horizontally (H), and move vertically (V). While foraging, the rat may be attacked by a robotic cat that also inhabits the house, which causes the rat to enter the Under-Attack state. From the Under-Attack state there is only one action, Die. In state Dead there is only one action, Stay Dead. The rewards and transition probabilities are as follows:

T(s,a,s')

	s'					
s, a	Bathroom	Kitchen	Bedroom	Dining Room	Under-Attack	Dead
Bathroom, H	0	0.6	0.4	0	0	0
Bathroom, V	0	0.4	0.6	0	0	0
Bathroom, S	0.75	0	0	0	0.25	0
Kitchen, H	0.6	0	0	0.4	0	0
Kitchen, V	0.4	0	0	0.6	0	0
Kitchen, S	0	0.75	0	0	0.25	0
Bedroom, H	0.4	0	0	0.6	0	0
Bedroom, V	0.6	0	0	0.4	0	0
Bedroom, S	0	0	0.75	0	0.25	0
Dining Room, V	0	0.6	0.4	0	0	0
Dining Room, H	0	0.4	0.6	0	0	0
Dining Room, S	0	0	0	0.75	0.25	0
Under-Attack, Die	0	0	0	0	0	1.0
Dead, Stay Dead	0	0	0	0	0	1.0

s	$R(s)$
Bathroom	+4
Kitchen	+10
Bedroom	0
Dining Room	+2
Under-Attack	-50
Dead	0

1. What is the total number of possible policies?
2. Trace value iteration by hand on this problem. Initialize the value of each state with 0. As robotic rats are somewhat shortsighted, use $\gamma = 0.5$. Give the values for all states after each iteration. You may stop after 6 iterations.

3. What is the optimal policy given your values for the states?

Solution

1. There are four states each with 3 possible actions, and two states with only one possible action. By Markov assumption, the choice in each state is independent of the other states. The total number of valid policies is therefore 3^4.

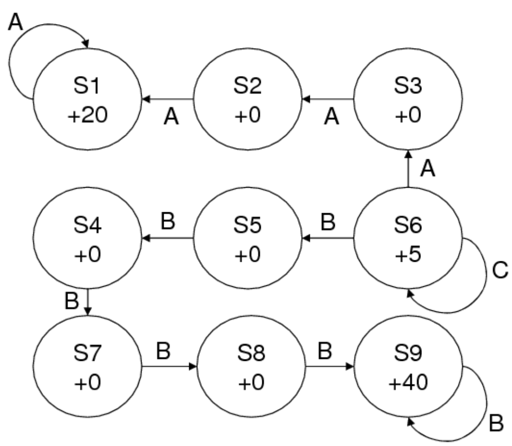
2. First six iterations of value-iteration

Iter	Bathroom	Kitchen	Bedroom	Dining Room	Under-attack	Dead
0	0	0	0	0	0	0
1	4	10	0	2	-50	0
2	7	11.6	1.6	5	-50	0
3	7.8	13.1	3.1	5.8	-50	0
4	8.55	13.5	3.5	6.55	-50	0
5	8.75	13.875	3.875	6.75	-50	0
6	8.9375	13.975	3.975	6.9375	-50	0

3. Optimal policy $\pi*$: $\pi*$ (Bathroom)= H, $\pi*$ (Kitchen)=H, $\pi*$ (Bedroom)=V,$\pi*$ (Dining Room)=V,$\pi*$ (Under Attach)= Die,$\pi*$ (Dead) = Stay Dead.

4 [15 pts] Discount Rate

Consider the following MDP.



Actions A, B, and C can be taken from state S6. All transitions are deterministic, and the immediate rewards for the states are given in the nodes. Let π^* denote the optimal policy.

1. If $\pi^*(S6) = A$, would you expect γ to be small, moderate, or large? Why?

2. If $\pi^*(S6) = B$, would you expect γ to be small moderate, or large? Why?

3. If $\pi^*(S6) = C$, would you expect γ to be small, moderate, or large? Why?

4. Find the specific ranges of γ (if they exist) so that $\pi^*(S6)$ will be A, B, and C. (For example, one incorrect solution is $\pi^*(S6) = A$ when $\gamma \in [0, 0.9]$, $\pi^*(S6) = B$ when $\gamma \in [0.9, 0.99]$, and $\pi^*(S6) = C$ when $\gamma \in [0.99, 1.0]$; another incorrect solution is $\pi^*(S6) = A$ for all γ.) You may find it helpful to recall the formula for a geometric series:
$\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r}$ if $r \in [0, 1)$

Solution

1. Moderate. If γ is small, the optimal policy would be C, since the gains of staying in S_6 would be greater than the longterm gains of going to S_1 or S_9. If γ is large, the optimal policy would choose B since that would lead to S_9 since the length of the path would be offset by the higher gains of reaching S_9.

2. Large.

3. Small.

4. Utility of choosing A is $5 + 20\gamma^3 + 20\gamma^4 + \ldots = 5 + 20\gamma^3/(1-\gamma)$
   Utility of choosing B is $5 + 5\gamma + 5\gamma^2 + \ldots = 5/(1-\gamma)$
   Utility of choosing C is $5 + 40\gamma^5 + 5\gamma^6 + \ldots = 5 + 40\gamma^5/(1-\gamma)$

   Solving for the points of intersection and comparing functions in each interval, we get $\gamma \in [0, 1/2], \pi*(S_6) = C, \gamma \in [1/2, 1/\sqrt{2}], \pi*(S_6) = A, \gamma \in [1/\sqrt{2}, 1], \pi*(S_6) = B$

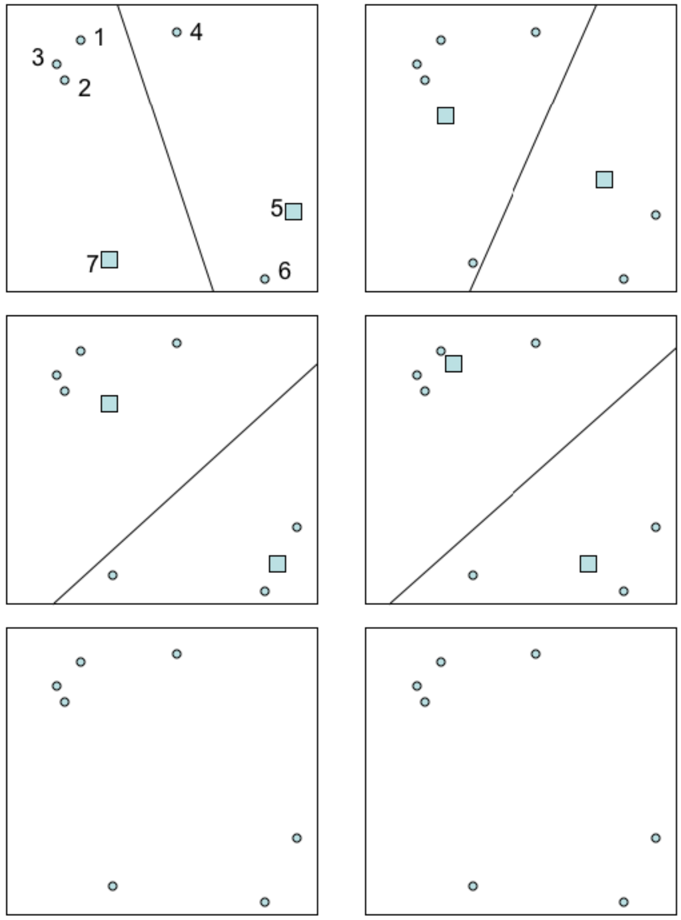
5 [10 pts] Clustering

1. Perform K-means clustering on the dataset given below. Circles are data points and there are two initial cluster centers, at data points 5 and 7. Draw the cluster centers(as squares) and the decision boundaries that define each cluster(we've already done this for the initial state, as shown in the first figure). If no points belong to a particular cluster, assume its center does not change. Use as many pictures as you need for convergence.

2. Give one advantage of hierarchical clustering over K-means clustering, and one advantage of K-means clustering over hierarchical clustering.

Solution

1. Possible solution



Full points were awarded even if the intermediate steps were slightly different(but plausbile). The final state should be as specified, if the method was correct.

2. Possible answers include: K-means is faster than hierarchical clustering, while hierarchical clustering doesn't require to pre-specify the number of clusters.

6 [30 pts] Forward-Backward in HMMS

Cecilia and Slackernerney, two graduate students in Markovford, have gotten into an argument over the weather. Cecilia thinks summer has ended and that we are into Autumn(Fall), while Slackenerney (who has been spending all his time in a room without windows) thinks that this is summer. Since they weren't able to resolve it themselves, they have decided to ask the fine students at CMU to help them decide.

A bit about Markovford: there are three seasons in Markovford – Summer(S), Autumn(A) and Winter(W). Given the season the previous day, the season on a day is conditionally independent of the season on all previous days. The weather is either Hot(H), Rainy(R) or Freezing(F). Given the season on any given day, the weather that day is independent of all other variables.

More formally, if we let q_i denote the season on the i^{th} day (taking values S,A,W) and O_i denote the observed weather pattern(one of H,R,F); the model $q_1 \dots q_N, O_1 \dots O_N$ is an HMM as covered in class.

1. Does the weather in Markovford reflect normal weather patterns? Why, or why not?

2. Implement forward-backward in MATLAB, for this HMM. In order to fully specify the HMM, you will need to fill the values for the CPT. These are in the tranprob and emitprob matrices of 'model' structure saved in input.mat (use load(input.mat) to get the structure; model.tranprob and model.emitprob are the respective probability distributions). readme.txt contains the numerical mapping of states. You will also need the initial state probabilities; please assume them to be uniform, ie $P(q_1 = A) = P(q_1 = S) = P(q_1 = W) = 1/3$

3. Cecilia has made 20 observation of the weather over the last 20 days(ie $O_1 \dots O_N$). This data is in the weather array of the model structure(model.weather).
   (a) Apply inference, to compute the probability of Autumn for each of these days (ie $P(q_i = A|O_1 \dots O_N)\forall i$)and write down the result you get. You must include this output both in your written submission and in a file called ps3_6a.txt and include it in your tarball.
   (b) Cecilia made the last observation today. Who do you think was correct – Cecilia or Slackernerney?
   (c) What is the probability that it is currently winter($P(q_N = W|O_1 \dots O_N)$) in Markovford?

Submit your code and a readme describing briefly what each file does, along with the output listed above. Your code musn't modify input.mat so I can test it with other inputs(by replacing this input.mat with a different file)

For those who don't plan to use matlab, the same data is stored in input.txt in human readable format. readme.txt has an explanation of input.txt. Again, your code should take in input.txt as a command line argument.

A note of caution on the programming: The transition probabilities in the data are defined to be from state i to state j and not j to i. So you'll need to transpose the matrix if you want to use the equation in question 2.1 to compute β_t.