

CS 181 Midterm 2

Chao Cheng

1. a. $p(A)p(C)p(B|A,C)p(D|C)p(E|B,D)$

$$1 + 1 + 4 + 2 + 4 = \boxed{12}$$

b. i. false (paths $C \rightarrow B \rightarrow E$ and $C \rightarrow D \rightarrow E$ are unblocked)

ii. false (path $A \rightarrow B \rightarrow E$ is unblocked)

iii True ($D \rightarrow E \rightarrow B$ is blocked since E is not observed)
 (and $D \rightarrow C \rightarrow B$ is blocked because C is observed)

c. 1 - missing a \sum_a in the beginning

2 - $p(D=d|C=c)$ should be after \sum_d not before

$$\begin{aligned} d. p(E) &= \sum_b \sum_c \sum_d p(c=c) p(d=d|c=c) p(e|b=b, d=d) \sum_a p(a=a) p(b=b|a=a, c=c) \\ &= \sum_b \sum_c \sum_d p(c=c) p(d=d|c=c) p(e|b=b, d=d) \psi_1(B, C) \end{aligned}$$

e. There are 4 numbers in the $\psi_1(B, C)$ table corresponding to each combination of B and C , where each number is the result of a sum-product calculation involving 2 products and 1 sum

f. It would add more as the factored expression terms would get extra conditions and therefore require more parameters

2. a. model one: $p(s_1)p(s_2)p(s_3)p(x_1|s_1)p(x_2|s_2)p(x_3|s_3)$

model two: $p(s_1)p(s_2|s_1)p(s_3|s_2)p(x_1|s_1)p(x_2|s_2)p(x_3|s_3)$

b. The HMM has the property that state s_{t+1} only depends on state s_t , which is more effective for tracking the location of a resident since the room they are in currently affects the probability of which room they will go to next (eg based on distance between rooms)

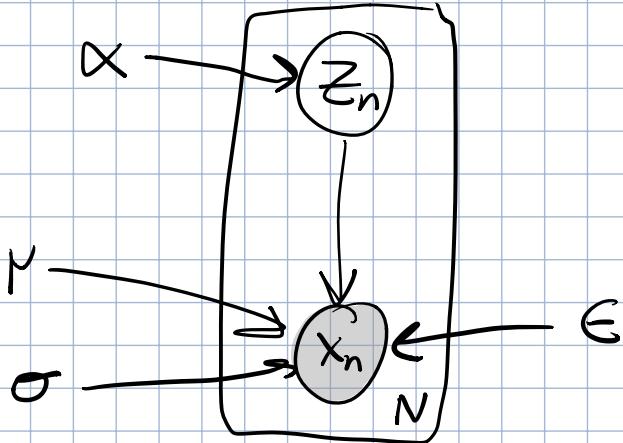
c. i. True (s_t and $x_{v \neq t}$ are not connected because no path)
(exists between them so they are independent)

ii False (s_t and $x_{v \neq t}$ are connected by a path so they are not conditionally independent)

d. Predicting the sequence of states is not a reliable way of getting to s_t so $p(s_{t+1}|s_t^*) \neq p(s_{t+1}|s_t)$

3. a. (i) and (iii) are consistent because the K-means loss graph shows that there are 3 clusters so after reducing dimensionality to 2 dimensions there will be 2 clusters in the projection
- b. (ii) and (iii) are consistent because the reconstruction loss graph shows that there is virtually no loss using 2 components and some loss using 1 component which indicates that the variability that could not be explained by using only 1 component is contained within a single axis/dimension which is what projections (ii) and (iii) show as they only vary along x axis whereas the other two projections have variation in both x and y.
- c. (iii) is consistent with both the K-means loss graph and the PCA reconstruction loss graphs so it is the projection that is most consistent with the data.
- d. not particularly effective; PCA tries to explain the variance in the data using the principal components (ideally the # of components is less than D) so if the variance of the data is low, using PCA will not be as effective in explaining the variance using a few principal components
- e. We would prefer K-means if the dimension of the data is high due to the curse of dimensionality

4. a.



b.

$$\begin{aligned} p(x_n, z_n; \alpha, \mu, \sigma, \epsilon) &= p(x_n | z_n, \mu, \sigma, \epsilon) p(z_n | \alpha) \\ &= N(x; \mu, \sigma^2)^{z_n} N(x; 0, \epsilon^2)^{1-z_n} \alpha \end{aligned}$$

c. $\ln(p(x_n, z_n; \alpha, \mu, \sigma, \epsilon))$

$$= \alpha \left[z_n \ln(N(x; \mu, \sigma^2)) + (1 - z_n) \ln(N(x; 0, \epsilon^2)) \right]$$

d. $q_n = p(z_n=1 | x_n; \alpha, \mu, \sigma, \epsilon)$

$$\begin{aligned} &= p(x_n | z_n=1; \mu, \sigma, \epsilon) p(z_n=1 | \alpha) \\ &= N(x_n; \mu, \sigma^2) \alpha \end{aligned}$$

e.

$E(\ln(p(x_n, z_n; \alpha, \mu, \sigma, \epsilon)))$

$$= \alpha E_{x_n \sim z_n} [\ln(x)] + (1 - \alpha) E_{x_n \sim z_n} [\ln(0)]$$

$$= \alpha \ln N(x; \mu, \sigma^2) + (1 - \alpha) \ln N(x; 0, \epsilon^2)$$

f. The MLE is closed form in EM so we can calculate the expectation easily and thus train the model efficiently

s.a. $V^\pi(s_0) = 0$

$$V^\pi(s_1) = -1 + 0.99 \times 0 + 0.99^2 \times 10 = 8.801$$

$$V^\pi(s_2) = 0 + 0.99 \times 10 = 9.9$$

$$V^\pi(s_3) = 10$$

b. The policy improvement step of policy iteration will update $\pi'(s_0) \leftarrow \text{right}$ since moving right has a higher value than moving left

c. $V^\pi(s_0) = 0$

$$V^\pi(s_1) = -1 + 0.01 \times 0 + 0.01^2 \times 10 = -0.999$$

$$V^\pi(s_2) = 0 + 0.01 \times 10 = 0.1$$

$$V^\pi(s_3) = 10$$

d. for s_0 :

$$a = \text{left} \rightarrow r(s_0, \text{left}) + \gamma V^\pi(s_0) = 0 + 0.01 \times 0 = 0$$

$$a = \text{right} \rightarrow r(s_0, \text{right}) + \gamma V^\pi(s_1) = 0 + 0.01 \times -0.999 = -0.00999$$

$0 > -0.00999$ so choose $\pi'(s_0) = \text{left}$

for s_1 :

$$a = \text{left} \rightarrow r(s_1, \text{left}) + \gamma V^\pi(s_0) = -1 + 0.01 \times 0 = 0$$

$$a = \text{right} \rightarrow r(s_1, \text{right}) + \gamma V^\pi(s_2) = -1 + 0.01 \times 0.1 = 0.001$$

$0.001 > 0$ so choose $\pi'(s_1) = \text{right}$,

so π is optimal.

e. No, because $\gamma V^\pi(s_2) > \gamma V^\pi(s_0)$ still, so the optimal policy is still $\pi(s_1) = \text{right}$

6a. We are not explicitly given a model (ie no transition function or reward function) and instead try to infer the optimal policy from samples of the world

b. False

c. We want to infer the optimal policy, which is given by $\pi^*(s) = \arg\max_a Q^*(s, a)$

therefore we keep track of state and action together instead of keeping them separate as in model based

d.

$$\text{action(state 1)} = \begin{cases} \text{random(left,right)} & \text{with probability } \epsilon \\ \text{left w/p=0.6} \\ \text{right w/p=0.4} & \text{otherwise} \end{cases}$$

e. update $Q(s_1, \text{left})$ with $s^1 = \text{state 2}$ and action $a^1 = \text{right}$

f. SARSA_λ is better than Q-learning when we want to get a firm grasp on the optimal policy by continuously exploring, however SARSA may get trapped by local minima whereas Q-learning (off-policy) is better for flexibility and does not get trapped as easily.

(or on-policy)