CS181 Midterm 2 Practice Questions

1. **Convergence of K-Means**

Consider Lloyd's algorithm for finding a K-Means clustering of N data, i.e., minimizing the "distortion measure" objective function

$$J(\{\boldsymbol{r}_n\}_{n=1}^N, \{\boldsymbol{\mu}_k\}_{k=1}^K) = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\boldsymbol{x}_n - \boldsymbol{\mu}_k||_2^2.$$

Show that as a consequence of there being a finite number of possible assignments for the set of responsibilities $r_{n,k}$, and that for each such assignment there is a unique optimum for the means $\{\boldsymbol{\mu}_k\}_{k=1}^K$, the K-Means algorithm must converge after a finite number of iterations.

2. **K-Means++**

One way to initialize Lloyd's algorithm for K-Means is to randomly select some of the data to be the first cluster centers. The easiest version of this would pick uniformly from among the data. K-Means++ biases this distribution so that it is not uniform. Explain in words how the distribution is non-uniform and why it should lead to better initializations.

3. **Standardizing Input Data**

Standardizing data helps ensure that distances makes sense and that the different properties of the items are balanced. Give an example of a kind of data for which standardization might be necessary to get good results from K-Means clustering.

4. **K-Medoids**

K-Medoids clustering is similar to K-Means, except that it requires the cluster centers to be data examples. Describe a situation in which this is desirable or necessary.

5. **Value Iteration**

Say an MDP has state space S and reward R where all rewards are positive. If you run value iteration, what is the largest k for which $V_k(s)$ is zero?

6. **Infinite Horizon**

You are on a linear space and can move only right or left. Each position has reward r_i and $\gamma \approx 1$. Describe your optimal policy given any state i (don't forget about ties).

7. **Value Iteration vs Expectimax Search**

What is the running time of Expectimax Search and Value Iteration as a function of the horizon, T, the number of actions, M, the the number of transitions for any state and action, L, and the number of steps, N? Why don't we always choose the algorithm with better asymptotic running time?

8. **Setting Rewards to Compute Shortest Path**

Suppose we have a grid-world where each state is represented as a point in $\{(x, y)|0 \leq x \leq 4, 0 \leq y \leq 4\}$. Suppose you have a robot that starts in the lower left corner and is given a goal point that is needs to reach. At each state, the robot can move one point to the left, right, up, or down, and each action has a 90% chance of success (otherwise you stay at the current point).

- What is the size of the state space in the resulting MDP?
- Suppose we want to minimize the length of path to the goal, but we have no preference for which path the robot should take (and all points are reachable). What rewards could we assign to each state so that we recover a shortest path to the goal as our optimal policy?

9. **Bounds**

Show that as a consequence of the constraint $0 \leq p(x_n|\mu_k) \leq 1$ for the discrete variable x_n, the incomplete-data log likelihood function for a mixture of Bernoulli distributions is bounded above, and hence that there are no singularities for which the likelihood goes to infinity.

10. **Gaussian**

Consider a Gaussian mixture model in which the marginal distribution $p(\boldsymbol{z})$ for the latent variable is given by

$$p(\boldsymbol{z}) = \prod_{k=1}^{K} \pi_k^{z_k} ,$$

and the conditional distribution $p(\boldsymbol{x} \mid \boldsymbol{z})$ for the observed variable is given by

$$p(\boldsymbol{x} \mid \boldsymbol{z}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k} .$$

Show that the marginal distribution $p(\boldsymbol{x})$, obtained by summing $p(\boldsymbol{z})p(\boldsymbol{x} \mid \boldsymbol{z})$ over all possible values of $\boldsymbol{z}$, is a Gaussian mixture of the form

$$p(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) .$$

11. **EM and Mixture Models**

Show that if we maximize

$$\mathbb{E}_z[\ln p(\boldsymbol{X}, \boldsymbol{Z} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma})] = \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \{\ln \pi_k + \ln \mathcal{N}(\boldsymbol{x}_n \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\}$$

with respect to $\boldsymbol{\mu}_k$ while keeping the responsibilities $\gamma(z_{nk})$ fixed, we obtain the closed form solution given by

$$\boldsymbol{\mu}_k = \frac{\sum_{n=1}^{N} \gamma(z_{nk}) \boldsymbol{x}_n}{\sum_{n=1}^{N} \gamma(z_{nk})}.$$

12. **Belief State Methods**

In belief state methods for POMDPs, the agent has to maintain a probability distribution over the current state of the world.

$$b(s_t) = P(s_t|o_1,,o_t,a_1,,a_{t-1})$$

Express $b(s_t)$ as a function of $b(s_{t-1})$.

13. **PCA on Circular Data (inspired by Barber, 15.2)**

Consider the following generative process for generating $2N$ data points along the circumference of a unit circle. (Recall that a unit circle satisfies $x^2 + y^2 = 1$.)

- Draw a point x_n along the x-axis $\sim$ Unif$(-1, 1)$
- Given x_n, set $y_{n_1} = \sqrt{1 - x^2}$ and $y_{n_2} = -\sqrt{1 - x^2}$, and add both $\langle x_n, y_{n_1} \rangle$ and $\langle x_n, y_{n_2} \rangle$ to the dataset.

Suppose we generate N points x_n in this way, giving us a total of $2N$ datapoints.

(a) If N is very large, what do you expect $\bar{\boldsymbol{x}}$, the empirical mean of the data, to (approximately) be?

(b) Using the approximation of $\bar{\boldsymbol{x}}$ from part (a), write out the sample covariance matrix $\boldsymbol{S}$ of the data in terms of the x_n.

(c) What (numerical) values will the entries of $\boldsymbol{S}$ approach as N gets large? (Hint: You may want to use a rearrangement of the identity $Var[x] = E[x^2] - (E[x])^2$, and the fact that $x_n \sim$ Unif$(-1, 1)$.)

(d) If we wanted to use PCA to reduce the datapoints to a single dimension, given your answer to part (c), what will the first principal component (i.e., eigenvector) and its corresponding eigenvalue approximately be?

(e) Is this surprising?

14. **High Dimensional Data (Bishop 12.1.4)**

Suppose we have a design matrix $\boldsymbol{X} \in \mathbb{R}^{N \times D}$ which has been centered, so the sample covariance matrix is $\boldsymbol{S} = \frac{1}{N}\boldsymbol{X}^\mathsf{T}\boldsymbol{X}$. Also, let $\boldsymbol{u}_d$, where $d = 1..D$, be the eigenvectors of $\boldsymbol{S}$.

(a) Show that the D vectors defined by $\boldsymbol{v}_d = \boldsymbol{X}\boldsymbol{u}_d$ are eigenvectors of $\frac{1}{N}\boldsymbol{X}\boldsymbol{X}^\mathsf{T}$, and that they have the same eigenvalues as their corresponding $\boldsymbol{u}_d$.

(b) Assuming we can recover the $\boldsymbol{u}_d$ from the $\boldsymbol{v}_d$ with reasonable time and memory, explain why calculating the $\boldsymbol{v}_d$ first might be useful if $N < D$.

(c) Show that the $\hat{\boldsymbol{u}}_d = \boldsymbol{X}^\mathsf{T}\boldsymbol{v}_d$ is, like $\boldsymbol{u}_d$, an eigenvector of $\boldsymbol{S}$.

15. **Reconstruction Error in Matrix Form (inspired by Barber, 15.2.1)**

Suppose we have a set of centered data vectors $\boldsymbol{x}_n \in \mathbb{R}^D$, and we would like to project them onto a subspace of dimensionality K using a basis of K (not necessarily orthonormal) vectors $\boldsymbol{w}_k$. As in the notes, we define $\hat{\boldsymbol{x}}_n \equiv \sum_{k=1}^{K} \alpha_k^{(n)} \boldsymbol{w}_k$, and so the reconstruction error is

$$\begin{aligned} J(\{\boldsymbol{u}_k\}_{k=1}^K) &= \sum_{n=1}^{N} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)^\mathsf{T} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n) \\ &= \sum_{n=1}^{N} (\boldsymbol{x}_n - \sum_{k=1}^{K} \alpha_k^{(n)} \boldsymbol{w}_k)^\mathsf{T} (\boldsymbol{x}_n - \sum_{k=1}^{K} \alpha_k^{(n)} \boldsymbol{w}_k). \end{aligned}$$

Now, let $\boldsymbol{X} \in \mathbb{R}^{N \times D}$ be the design matrix of data vectors, $\boldsymbol{W} \in \mathbb{R}^{D \times K}$ be a matrix of horizontally stacked $\boldsymbol{w}_k$ column vectors, and $\boldsymbol{Z} \in \mathbb{R}^{K \times N}$ be a matrix of horizontally stacked α column vectors. That is, $\boldsymbol{Z} = \begin{bmatrix} \alpha_1^{(1)} & \cdots & \alpha_1^{(N)} \\ \vdots & \ddots & \vdots \\ \alpha_K^{(1)} & \cdots & \alpha_K^{(N)} \end{bmatrix}$.

(a) Write the reconstruction error J only in terms of the matrices $\boldsymbol{X}$,$\boldsymbol{W}$, $\boldsymbol{Z}$, and matrix operations.

(b) Argue from the form of your answer to (a) that there are no unique $\boldsymbol{W}, \boldsymbol{Z}$ that minimize J.

16. **Q learning**

Suppose you are standing on a linear board: you can take action L or R (walk left or right), or sleep. If you walk, you have probability p_a that you actually walk to the next square, where $a \in L, R$. Otherwise, your cat distracted you and you are still on the same square. Staying a square gives you reward r_i. Your learning rate is $\alpha = .5$ and $\gamma = .5$.
You are on square 1, you choose $a = Left$ and receive $r = 4$. What is your updated value of $Q(1, L)$?

17. **Q learning**

In the next step from the previous problem, you are on square 0, you choose $a = R$, receive $r = 3$ and end up in $s = 1$. What is your updated value of $Q(0, R)$?

18. **Model Based Reinforcement Learning**

In model-based RL, how might we compute the Maximum-Likelihood estimate of the expected reward from taking action a in state s?

19. **ϵ-Greedy**

Why do we generally use an ϵ-Greedy algorithm when choosing the current action during the Q-Learning algorithm? Describe how you would change the value of ϵ over time to encourage early exploration/learning and good decision making after the state space was explored.

20. **Convergence of Q-Learning**

Suppose we have a deterministic world where each state-action pair (s, a) is visited infinitely often. Consider an interval during which every such pair is visited. Suppose that the largest error in the approximation of $\hat{Q}_n$ after n iterations is e_n, i.e.

$$e_n = \max_{s,a} |\hat{Q}_n(s, a) - Q(s, a)|$$

Show that e_{n+1} is bounded above by γe_n, where γ is the usual discount factor.