Practice Questions for CS 181, Midterm 2 (Spring 2021)

Finale Doshi-Velez and David C. Parkes
Harvard College

April 18, 2021

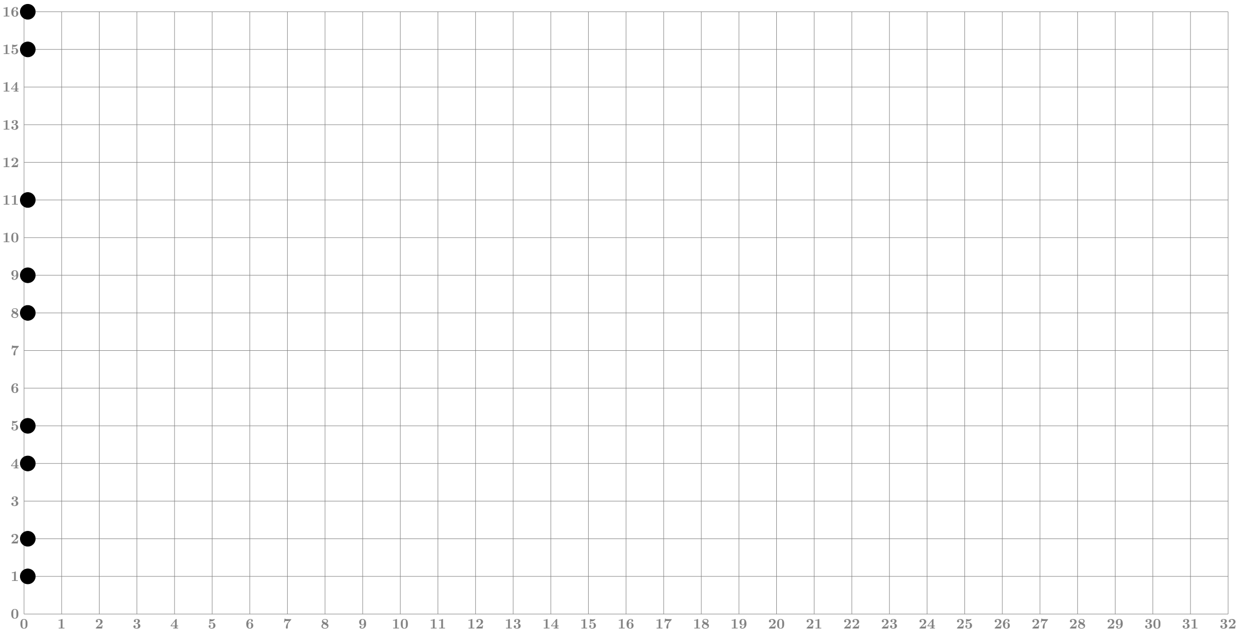
These practice questions are illustrative of the kinds of understanding that you should expect to be tested on the midterm. If anything they are slightly more difficult than the questions on the exam. You can expect around 6 questions on the midterm and you will have 100 minutes. This means that a typical question should take 15 minutes. But some will be shorter, some longer and say 10 mins vs 20 min questions. We've provide rough guidance here ("short"), ("typical") and ("long").

1. **Hierarchical Agglomerative Clustering** [Short]

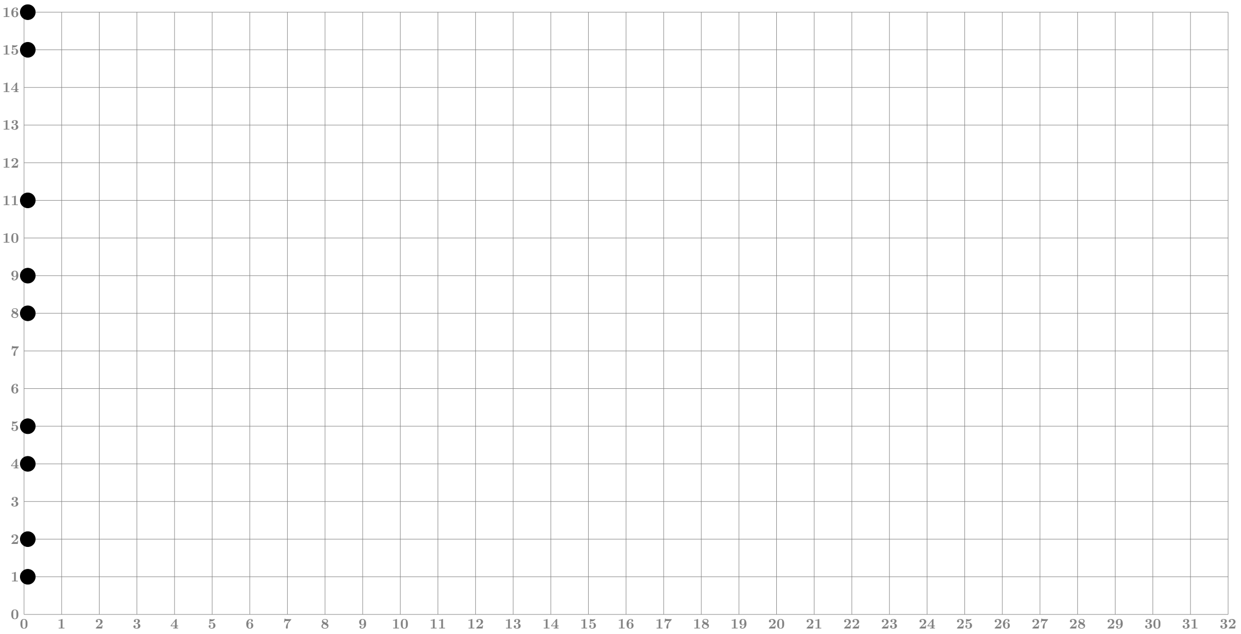
   Consider nine points $x_1, \ldots, x_9$ shown below, where the y-axis provides their values. We define $d(x, x') = |x - x'|$, and consider two different cluster distances.

   Draw the dendrogram for the data. Join together clusters one per step (on the horizontal-axis), breaking ties towards joining lower x values first. In the top figure, use the min-linkage distance and in the bottom figure use the max-linkage distance.

   (a) Min Linkage:

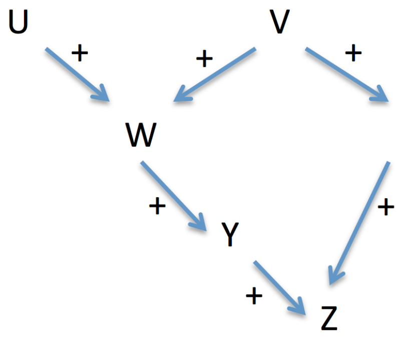


   (b) Max Linkage:



2. **Bayesian networks** [Typical]

Consider the following Bayesian network, where the variables are all Boolean.



The '+' annotations indicate the direction of the local effect; e.g., the '+' from U to W means that for each value v of V,

$$p(W = true \,|\, U = true, V = v) > p(W = true \,|\, U = false, V = v).$$

For each of the following questions, select one of the following, and also state which (if any) undirected paths are blocked (in the sense of d-separation):

$=$ if the two probabilities are necessarily equal;
$<$ if the first probability is necessarily smaller;
$>$ if the first probability is necessarily larger;
? if none of these cases hold.

(a) $p(V = true \mid Y = false)$ $\qquad$ $p(V = true \mid Y = true)$

(b) $p(V = true \mid Z = false)$ $\qquad$ $p(V = true \mid Z = true)$

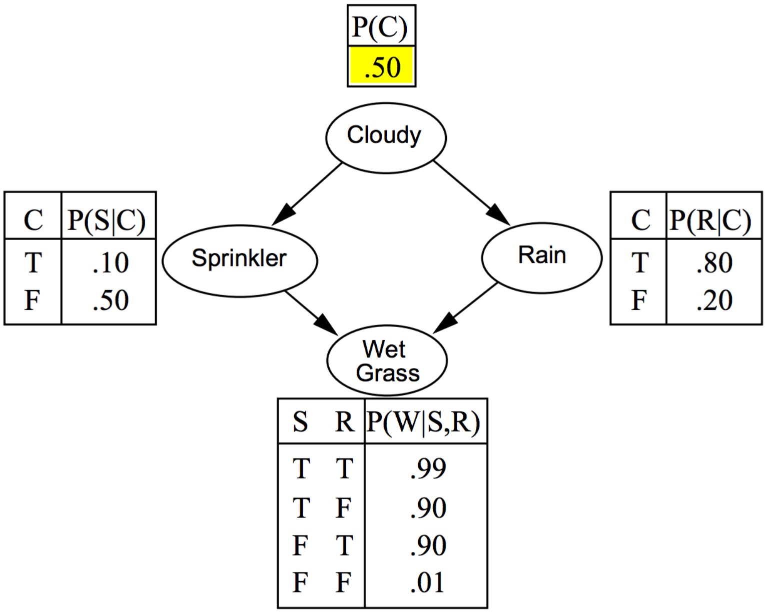
(c) $p(U = true \mid W = true, Y = false)$ $\qquad$ $p(U = true \mid W = true, Y = true)$

(d) $p(Y = true \mid Z = true, X = false)$ $\qquad$ $p(Y = true \mid Z = true, X = true)$

(e) $p(U = true \mid Y = true, Z = false)$ $\qquad$ $p(U = true \mid Y = true, Z = true)$

3. **Bayesian networks** [Long]

Consider this example of a Bayesian network with binary variables. It models a garden lawn and whether or not the grass is wet.



P(C)
.50

C	P(S|C)
T	.10
F	.50

C	P(R|C)
T	.80
F	.20

S	R	P(W|S,R)
T	T	.99
T	F	.90
F	T	.90
F	F	.01

(a) Construct an alternative Bayesian network that models the same distribution for variable ordering, S, C, R, W. That is, add S, then C with any required edge, then R with any required edges, then W with any required edges. **Don't specify conditional probability tables.** [Hint: Use the given Bayesian network to determine which conditional Independence properties hold amongst preceding variables, and only include needed edges.]

(b) Is this new Bayesian network a correct model of the distribution? Which network do you consider to be preferable, if any?

(c) Going back to the original network, what is the probability that it is not cloudy, rains, sprinkler doesn't run, and grass is wet?

(d) In the original network: write down the first two steps of variable elimination for $p(W)$, eliminating C and then S. Perform the numerical calculations!

4. **Markov Decision Process (modeling).** [Long] [Harder than an MDP modeling question you'd expect on Spring 2021 midterm]

   You are asked to develop a Markov Decision Process (MDP) to be used for the control of a single elevator. To model:

   - There are three floors
   - There are three buttons inside the car
   - There is a single call button outside on each floor
   - The door of the elevator opens and closes.

   The "agent" here is the elevator itself, and the aim of the system is to get passengers to their appropriate floors.

   (a) Describe in words the states, actions, reward function, and transition model for a suitable MDP model. Make sure that the reward function is clear.

   (b) Explain your model as you introduce it. From your explanation the reader should understand the idea for why an optimal policy should lead to an efficient system.

   NOTE: **There is no single correct answer here.**

5. **Alternate Reward Function for MDPs** [Short]

We have been assuming that the reward function for an MDP has the form $r(s, a)$. Also recall that we have written value iteration for infinite-horizon problems as:

$$V'(s) \leftarrow \max_a \left[r(s, a) + \gamma \sum_{s'} p(s' \mid s, a) V(s') \right] \tag{1}$$

Now, imagine that we have a reward function that depends on both the current state *and* the next state, i.e., $r(s, a, s')$.

(a) Explain why this kind of reward function can be useful from a modeling perspective

(b) Write an expression for the value iteration step that incorporates this alternative type of reward.

(c) Explain formally why this approach is neither more general nor less general than an MDP model that insists on just using $r(s, a)$.

6. **Planning in MDPs** [typical]

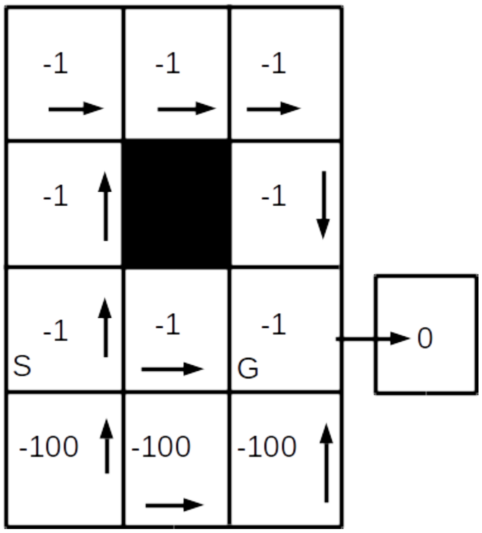
Consider a gridworld with the layout below. From each square, the agent may move into an adjoining square (up, down, left, right) or stay in place. If a policy specifies a move into a square which does not exist (i.e. down from one of the squares in the bottom row), the agent stays in place. Actions are deterministic, that is, they always have their intended effect. We use an infinite horizon with discount $\gamma = 1$. [This keeps the math simple in this example]

The robot starts in the state marked with an S. Upon reaching the state marked G the agent transitions into an absorbing state where it stays forever. **The rewards associated with a state are the reward for taking any action from that state**.

Recall the policy improvement step in policy iteration (where V^π is the value function of the current policy):

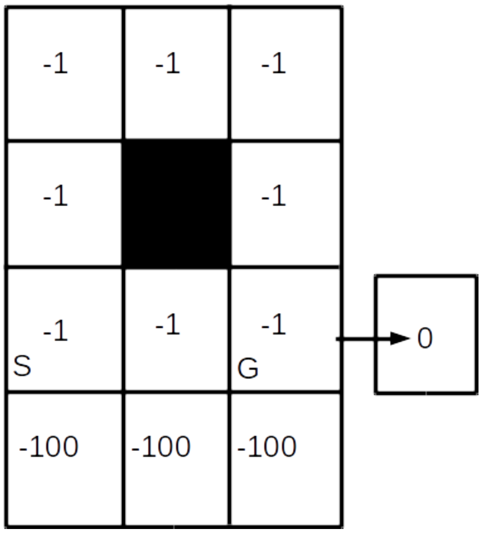
$$\pi'(s) \leftarrow \arg\max_{a \in A} \left[r(s,a) + \gamma \sum_{s' \in S} p(s' \mid s,a) V^\pi(s') \right], \quad \forall s$$

(a) Suppose that we follow the policy given by the arrows. What is the MDP value of each state under this policy? [You can figure this out by inspection of the policy and the environment]



(b) Can this policy be improved? To check this, (1) use policy improvement and draw the adjusted policy and (2) compute the new value function in each state.

(c) Is the new policy optimal? [Hint: you should be able to argue yes/no directly, without doing another round of policy iteration]



7. **Reinforcement learning** [typical]

The update rule for **SARSA** reinforcement learning is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[(r + \gamma Q(s',a')) - Q(s,a)]. \tag{2}$$

(a) What are the different quantities, how are they generated (e.g., which by the agent, which from the environment), and what is the idea of the update?

(b) What is meant by 'on-policy' and 'off-policy' reinforcement learning, and is SARSA an on-policy or off-policy method?

(c) What does it mean to **exploit** in the context of reinforcement learning?

(d) Consider this simple MDP world, where the reward is 100 for any action taken in state f and 0 in all other states and actions are deterministic (thus 'up' always moves 'up').

d	e	f
a	b	c

Assume the Q-values are initialized to 0, and the agent is initially in state c. What are the updates made by SARSA following each action (for $\alpha = 0.9$ and $\gamma = 0.9$).

Assume that no update is possible until the values of s, a, r, s', a' are all well-defined.

i. up (to state f)

ii. left (to state e)

iii. right (to state f)

iv. down (to state c)

8. **K-Means**[Typical]

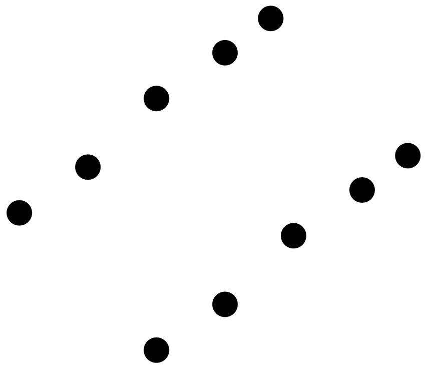
In K-Means, we are given a set of points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ and a fixed number of clusters K. Our aim is to find cluster centers $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$ that represent the data.

(a) Define the K-Means loss function.

(b) What two steps does Lloyd's algorithm repeat in order to find a good clustering?

(c) What is the asympotic run-time of each step of Lloyd's algorithm, as a function of the number of examples N and the number of clusters K?

(d) Given data that falls on two parallel diagonal lines as shown below, can Lloyd's algorithm with $K = 2$ find two clusters, such that each line is in one of the clusters?



9. **Hidden Markov Models** [Long] [You would need a calculator for this one!]

Consider a weather domain, with observations $\mathbf{x}_t \in \{D, R\}$ (dry, rain) and hidden state $\mathbf{s}_t \in \{C, S\}$ (cloud, sun). Assume the following parameters:

- initial prob: $p(\mathbf{s}_1 = C) = 0.7$
- transition

		Next State	
$p(\mathbf{s}_{t+1} \mid \mathbf{s}_t)$		C	S
State	C	0.8	0.2
	S	0.1	0.9

- output

		Output	
$p(\mathbf{x}_t \mid \mathbf{s}_t)$		D	R
State	C	0.25	.75
	S	0.6	0.4

(a) For a general HMM, if the total number of timesteps is T and $t < T$ is a timestep in the middle of the sequence, why is $p(\mathbf{s}_t \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \neq p(\mathbf{s}_t \mid \mathbf{x}_1, \ldots, \mathbf{x}_t)$? (An informal answer is fine.)

(b) (Forward-backward algorithm). Now suppose we observe $\mathbf{x}_1 = R$, $\mathbf{x}_2 = R$.
We can calculate:

$$\alpha_1(\mathbf{s}_1) = \begin{cases} 0.525 & \text{, if } \mathbf{s}_1 = C \\ 0.12 & \text{, if } \mathbf{s}_1 = S \end{cases}$$

Use

$$\alpha_2(\mathbf{s}_2) = p(\mathbf{x}_2 \mid \mathbf{s}_2) \sum_{\mathbf{s}_1} p(\mathbf{s}_2 \mid \mathbf{s}_1) \alpha_1(\mathbf{s}_1)$$

to compute the α_2-values.

(c) We have $\beta_2(\mathbf{s}_2) = 1$. In addition, we can calculate:

$$\beta_1(\mathbf{s}_1) = \begin{cases} 0.68 & \text{, if } \mathbf{s}_1 = C \\ 0.435 & \text{, if } \mathbf{s}_1 = S \end{cases}$$

Use these quantities, and

$$p(\mathbf{s}_t \,|\, \mathbf{x}_1, \ldots, \mathbf{x}_t) \propto \alpha_t(\mathbf{s}_t)\beta_t(\mathbf{s}_t)$$

to infer the values of $p(\mathbf{s}_1 \,|\, \mathbf{x}_1, \mathbf{x}_2)$ and $p(\mathbf{s}_2 \,|\, \mathbf{x}_1, \mathbf{x}_2)$.

(d) Use $p(\mathbf{x}_1, \mathbf{x}_2) = \sum_{\mathbf{s}_t} \alpha_t(\mathbf{s}_t)\beta_t(\mathbf{s}_t)$ to calculate the likelihood of the data.

10. **Mean of a Mixture Model** [Short]

For some class conditional distribution p_{class}, the details of which don't matter for this question, we are given a mixture model of the form

$$p(\mathbf{x}; \{\boldsymbol{\pi}_k\}_{k=1}^{K}, \boldsymbol{\theta}) = \sum_{k=1}^{K} \theta_k \, p_{\text{class}}(\mathbf{x} \,|\, \mathbf{z} = C_k; \boldsymbol{\pi}_k) \tag{3}$$

where example $\mathbf{x} \in \mathbb{R}^D$.

(a) Draw a graphical model with plates to show the form of this mixture distribution for a single example $\mathbf{x}$. [Note: if you're unfamiliar with the idea of "plate notation" for graphical models, take a quick look at p.363-365 in Bishop's book `https://bit.ly/3eajKx5`]

(b) Suppose that the mean of the class-conditional distribution for component k is given by $\boldsymbol{\mu}_k$. Show that the mean of the overall mixture model is given by

$$\mathbb{E}[\mathbf{x}] = \sum_{k=1}^{K} \theta_k \boldsymbol{\mu}_k.$$

11. **Expectation Maximization** [Typical]

We have a collection of binary images $\mathbf{x}_1, \ldots, \mathbf{x}_N$, each of which is 5×5. We treat each image $\mathbf{x}_n$ as a 25-dimensional binary vector where the dth pixel is $x_{n,d}$. We model an image as coming from a mixture distribution, with a product-of-Bernoulli distribution for each component k:

$$p(\mathbf{x}_n; \boldsymbol{\mu}_k) = \prod_{d=1}^{25} \mu_{k,d}^{x_{n,d}} (1 - \mu_{k,d})^{1 - x_{n,d}} \qquad \mathbf{x}_n \in \{0,1\}^{25} \qquad \boldsymbol{\mu}_k \in \{0,1\}^{25}\,.$$

Each of the K components has parameters $\boldsymbol{\mu}_k$, where each dimension $\mu_{k,d}$ specifies the probability that pixel d is black in an example from this component. The mixture weights are $\{\theta_k\}_{k=1}^K$ and known. You will use EM to estimate the $\{\boldsymbol{\mu}_k\}$ parameters.

(a) Write down the probability of generating a single image $\mathbf{x}$, i.e.,

$$p(\mathbf{x}; \{\boldsymbol{\mu}_k\}_{k=1}^K, \boldsymbol{\theta})$$

.

(b) What are the "latent variables" in this model? Draw the plate diagram for this model, writing it for N examples, and indicating what is known and unknown. [Hint: see the previous question for a pointer to plate diagrams]

(c) In the E-step, you find the probability with which example $\mathbf{x}_n$ belongs to each component fixing the parameters $\{\boldsymbol{\mu}_k\}_{k=1}^K$; i.e., $q_{n,k} = p(\mathbf{z}_n = C_k \,|\, \mathbf{x}_n; \{\boldsymbol{\mu}_k\}, \boldsymbol{\theta})$ for each k. Derive the expression for $q_{n,k}$.

(d) In the M-step, you update the parameters $\{\boldsymbol{\mu}_k\}_{k=1}^K$. Write down the expression for this, making use of the $\mathbf{q}$ values. [No need to derive the answer. As a hint, for the supervised case with class $\mathbf{z}_n$ of each image (one-hot coded), the MLE for the parameters of class k would be

$$\mu_{k,d} = \frac{\sum_{n=1}^N z_{n,k} x_{n,d}}{\sum_{n=1}^N z_{n,k}}$$

(intuitively, the percentage of times pixel d was black for the data in class k).]

12. **PCA** [Typical]

Consider a mean-centered data set of four points $\mathbf{x}_1 = (1, 0)$, $\mathbf{x}_2 = (-1, 0)$, $\mathbf{x}_3 = (0, -2)$, $\mathbf{x}_4 = (0, 2)$.

(a) Compute the empirical covariance matrix $\mathbf{S}$ for this dataset.

(b) Draw a rough sketch of the distribution $\mathcal{N}(0, \mathbf{S})$ formed with this covariance matrix.

(c) If we were to run PCA on this data, what algebraic properties of the empirical covariance matrix correspond to first and second principal components? What are the first and second principal components in this example? [Hint: you should be able to easily recognize them without computation]

(d) Graph the four points after running PCA and projecting down to a single dimension. What is lost in this transformation?