Practice Questions for CS 181, Midterm 2 (Spring 2017)

April 16, 2017

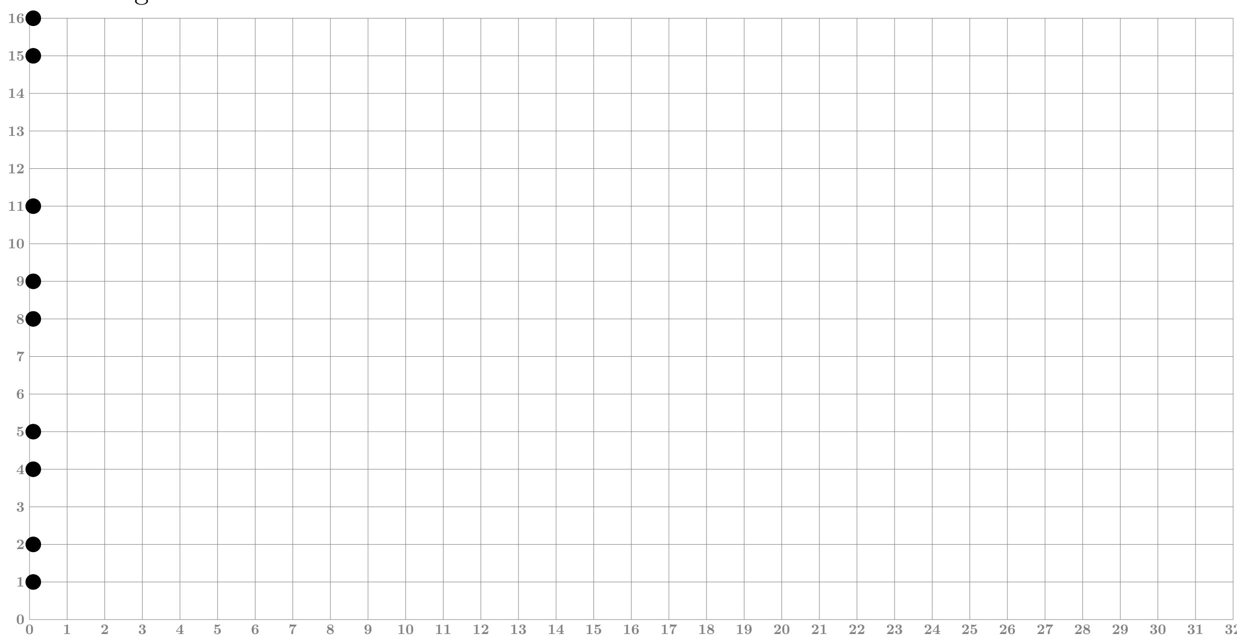
These practice questions are illustrative of the kinds of understanding that you should expect to be tested on the midterm. If anything they are slightly more difficult than the questions on the exam.

1. **Hierarchical Agglomerative Clustering**

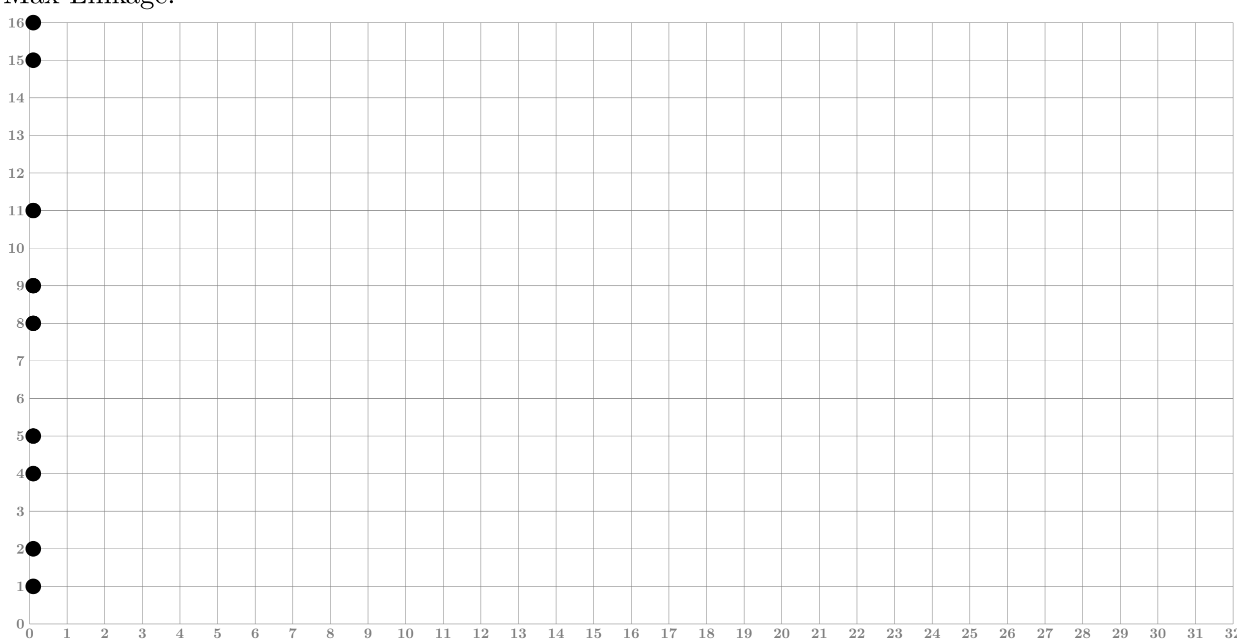
   Consider nine points $x_1, \ldots, x_9$ in $\mathbb{R}$ shown below, where the y-axis provides their values. We define $d(x, x') = |x - x'|$, and consider two different cluster distances.

   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:

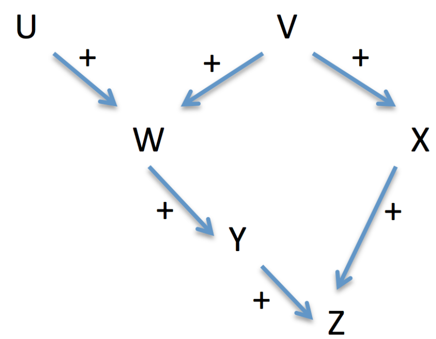


   (b) Max Linkage:



2. **Bayesian networks**

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) paths are blocked:

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

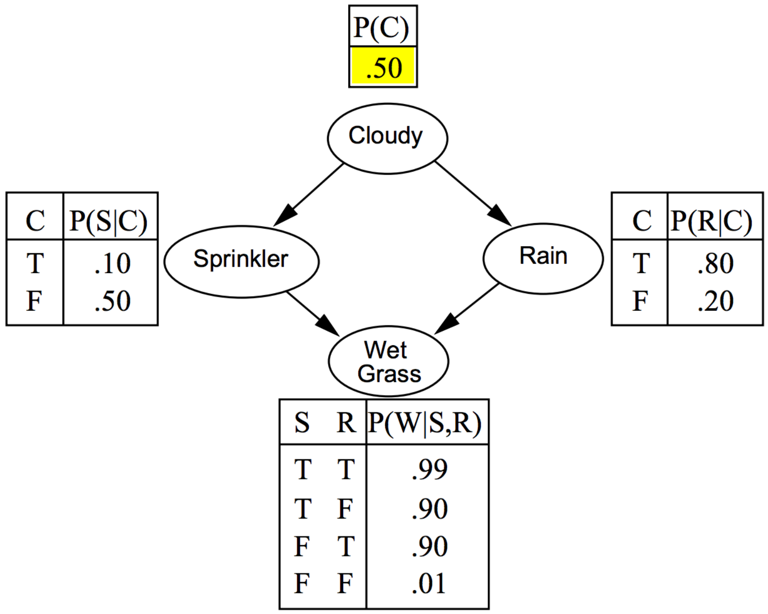
(c) $p(U = true \mid W = true, Y = false)$ $\qquad$ $p(U = true \mid W = true, Y = true)$

(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**

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



(a) Consider an alternate variable ordering, S, C, R, W. Use the given Bayesian network to determine which conditional indpendence properties (if any) hold amongst preceding variables, and construct the corresponding Bayesian network for this ordering. (Don't worry about specifying the conditional probability tables.)

(b) Is this new Bayesian network a correct model of the distribution? Which network is preferable?

(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!

(e) Explain how Gibbs sampling works for the inference problem $p(C \mid W = true)$.

4. **Markov Decision Process (modeling).**

   You are asked to develop a Markov Decision Process (MDP) to be used for the control of a single elevator. There are three floors, three buttons inside the car, and 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.

   Describe in words the states, actions, reward function, and transition model for a suitable MDP model. There is no single correct answer here, but a proper response will specify the size of each set and make the reward function clear. An optimal policy over this MDP should lead to an intuitively efficient system.

5. **Alternate Reward Function for MDPs**

We have been assuming that the reward function has the form $r(s, a)$, i.e., it only depends on the current state and action. In the discounted infinite-horizon case, we use this when computing the value function via:

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

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

(b) **Argue that this approach is strictly more general, by showing that any standard MDP (with $r(s, a)$) can be written in this form.**

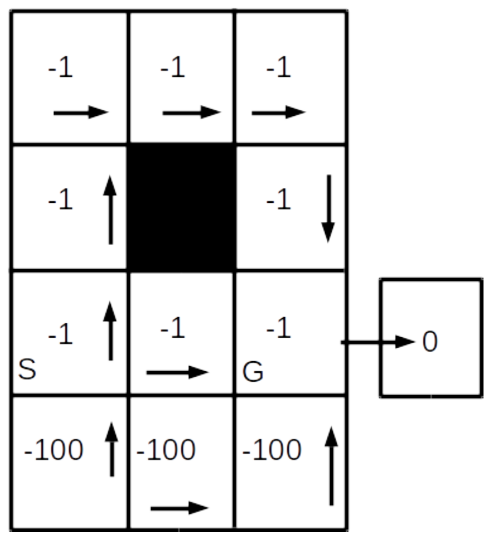
6. **Planning in MDPs**

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. Actions are deterministic, that is, they always have their intended effect. We use an infinite horizon with discount γ. We always start in 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 in that state. Assume a discount factor $\gamma = 1$.

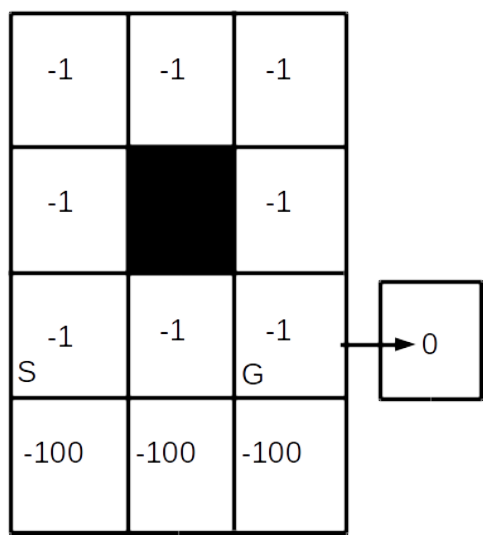
Recall the improvement step in policy iteration:

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

(a) Suppose that we follow the policy given by the arrows. What is the MDP value of each state under this policy? You may draw the values directly on the figure.



(b) Can this policy be improved? Draw the adjusted policy based on one round of policy iteration and compute the new value function in each state. Is the new policy optimal?



7. **Reinforcement learning**

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

$$w_{s,a} \leftarrow Q(s, a; \mathbf{w}) - \eta(Q(s, a; \mathbf{w}) - [r + \gamma Q(s', a'; \mathbf{w})])$$

(a) **Describe** the aim of the temporal difference update. What property will hold at convergence?

(b) **How** do we select state s, action a, reward r, state s' and action a' in SARSA learning with a policy π?

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

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

(e) Consider this simple MDP world, where the reward is 100 for any action 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 updates made by SARSA following each action (for $\eta = 0.9, \gamma = 0.9$). 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**

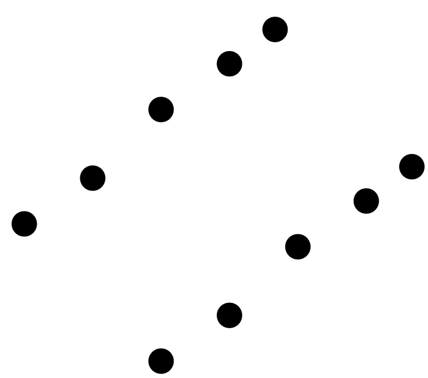
In K-Means, we are given a set of points $\mathbf{x}_1, , \mathbf{x}_n$ and a fixed number of clusters K. Our aim is to find cluster centers $\mathbf{z}_1, \ldots, \mathbf{z}_K$ that represent the data.

(a) Formally 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 find two clusters, such that each line is in one of the clusters?



9. **Hidden Markov Models**

Consider a weather domain, with outputs $\mathbf{x}_t \in \{D, R\}$ (no rain, rain) and state $\mathbf{s}_t \in \{C, S\}$ (cloud, sun). Assume the following parameters:

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

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

- output ($\boldsymbol{\pi}$)

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

(a) For a general HMM, if the total number of timesteps is n and $t < n$ 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 have a single sequence of observed data with $\mathbf{x}_1 = R$, $\mathbf{x}_2 = R$ in the weather domain.
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

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

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

(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**

We are given a mixture model of the form,

$$p(\mathbf{x} \,|\, \{\boldsymbol{\theta}_k\}_{k=1}^{c}) = \sum_{k=1}^{c} \pi_k \, p(\mathbf{x} \,|\, \boldsymbol{\theta}_k)$$

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

(a) **Draw** a graphical model with plates to show the form of the mixture distribution.

(b) The mean of the c'th component distribution $p(\mathbf{x} \,|\, \boldsymbol{\theta}_k)$ is given by $\boldsymbol{\mu}_k$. Show that the mean of the overall mixture is given by

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

11. **Expectation Maximization**

(This problem is a bit harder than we would ask on the exam, but it is a good review of the EM algorithm.)

Imagine that we have a collection of n binary images $\mathbf{x}_1, \ldots, \mathbf{x}_n$, each of which is 5×5. We treat each image as a 25-dimensional binary vector where image i is denoted as a vector $\mathbf{x}_i$ and the jth pixel is $x_{i,j}$.

We wish to estimate the parameters of a mixture model, and use a product of Bernoulli distributions for each component in the mixture:

$$p(\mathbf{x}_i \,|\, \boldsymbol{\mu}_k) = \prod_{j=1}^{25} \mu_{k,j}^{x_{i,j}} (1 - \mu_{k,j})^{1-x_{i,j}} \qquad \mathbf{x}_i \in \{0,1\}^{25} \qquad \boldsymbol{\mu}_k \in (0,1)^{25} \,.$$

Each of the C mixture components has a parameter $\boldsymbol{\mu}_k$ and each dimension $\mu_{k,j}$ is a Bernoulli distribution parameter that specifies the probability of pixel j being black. The (known) mixture weights are $\{\pi_k\}_{k=1}^{c}$. You want to learn the parameters $\{\boldsymbol{\mu}_k\}_{k=1}^{c}$, so you decide to use the expectation-maximization algorithm. (There are four parts to this, make sure you do parts (c) and (d) on the next page.)

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

$$p(\mathbf{x} \,|\, \{\boldsymbol{\mu}_k\}_{k=1}^{c}, \{\pi_k\})$$

.

(b) This is an example of a latent variable model. What is the form of the latent variables? Draw the plate diagram for this problem indicating what is known and unknown.

(c) In the E-step, you improve the estimate of the latent variables, fixing the parameters $\{\boldsymbol{\mu}_k\}_{k=1}^{c}$. Derive the update for this approximation. In particular compute the q function.

(d) {8pts} In the M-step, you update the parameters $\{\boldsymbol{\mu}_k\}_{k=1}^{c}$. Derive the update for the parameters. Assuming that you have the q function.

12. **PCA**

Consider a 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 normalized feature 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 would be the first and second principal components?

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