Homework 3

DSE 210

Guidelines for Homework Submission

1. Each HW is released on the day of the lecture. Students are expected to submit their solutions for the respective HW until 11:59 PM, the day before the next lecture.

2. Each HW is divided into multiple worksheets based on the concepts discussed in the class. Students are expected to attempt all the worksheets in a given HW.

3. HW has to be done individually. NO group work is allowed.

4. The solutions for the HW can be handwritten or typed. (Note: If the handwriting is illegible or if the pages are not scanned properly then the student shall receive 0 points for the respective question)

5. The solutions have to be uploaded to Gradescope in PDF format only. (Note: Do not forget to map question numbers and the pages containing the respective answers while uploading on Gradescope)

6. Some questions in the HW have the tag "***Programming Question***". Students are expected to implement the solution for the respective question in Python. PDF version of the source code has to be uploaded to Gradescope. (Note: Students should combine the theoretical solutions and Python source codes into one PDF file and then upload it to Gradescope)

DSE 210: Probability and statistics **Winter 2020**

Worksheet 7 — Linear algebra primer

1. (2 points) Find the unit vector in the same direction as $x = (1, 2, 3)$.
2. (2 points) Find all unit vectors in $\mathbb{R}^2$ that are orthogonal to $(1, 1)$.
3. (2 points) How would you describe the set of all points $x \in \mathbb{R}^d$ with $x \cdot x = 25$?
4. (2 points) The function $f(x) = 2x_1 - x_2 + 6x_3$ can be written as $w \cdot x$ for $x \in \mathbb{R}^3$. What is w?
5. (2 points) For a certain pair of matrices A, B, the product AB has dimension 10×20. If A has 30 columns, what are the dimensions of A and B?
6. (3 points, 1 each) We have n data points $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^d$ and we store them in a matrix X, one point per row.
   (a) What is the dimension of X?
   (b) What is the dimension of XX^T?
   (c) What is the (i, j) entry of XX^T, simply?
7. (2 points) Vector x has length 10. What is $x^T x x^T x x^T x$?
8. (2 points) For $x = (1, 3, 5)$ compute $x^T x$ and xx^T.
9. (2 points) Vectors $x, y \in \mathbb{R}^d$ both have length 2. If $x^T y = 2$, what is the angle between x and y?
10. (2 points) The quadratic function $f : \mathbb{R}^3 \to \mathbb{R}$ given by
$$f(x) = 3x_1^2 + 2x_1x_2 - 4x_1x_3 + 6x_3^2$$
can be written in the form $x^T M x$ for some **symmetric** matrix M. What is M?
11. (4 points, 1 each) Which of the following matrices is necessarily symmetric?
   (a) AA^T for arbitrary matrix A.
   (b) $A^T A$ for arbitrary matrix A.
   (c) $A + A^T$ for arbitrary square matrix A.
   (d) $A - A^T$ for arbitrary square matrix A.
12. (4 points, 2 each) Let $A = \text{diag}(1, 2, 3, 4, 5, 6, 7, 8)$.
   (a) What is $|A|$?
   (b) What is A^{-1}?
13. (4 points, 2 each) Vectors $u_1, \ldots, u_d \in \mathbb{R}^d$ all have unit length and are orthogonal to each other. Let U be the $d \times d$ matrix whose rows are the u_i.

(a) What is UU^T?

(b) What is U^{-1}?

14. (2 points) Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & z \end{pmatrix}$ is singular. What is z?

15. ***Programming Question:*** (10 points) Code following matrix operations in Python:

    (a) Find the vector matrix product of A and M (where A is a vector and M is a matrix). where,
    A = [1, -1, 0]
    M = [[1, 2, 3], [4, 5, 6], [7, 8, 9]]

    (b) Find the matrix product of 3 matrices A, B and C (i.e. A_dot_B_dot_C). where,
    A = [[1, 2]]
    B = [[2, 3, 4],[5, 6, 7]]
    C = [[-1, 1, -1, 1], [0, 0, 0, 0], [1, 1, 1, 1]]

DSE 210: Probability and statistics **Winter 2020**

Worksheet 5 — Classification with Generative models 1

16. (5 points) A man has two possible moods: `happy` and `sad`. The prior probabilities of these are:

$$\pi(\texttt{happy}) = \frac{3}{4}, \quad \pi(\texttt{sad}) = \frac{1}{4}.$$

His wife can usually judge his mood by how talkative he is. After much observation, she has noticed that:

- When he is happy,

$$\Pr(\text{talks a lot}) = \frac{2}{3}, \;\; \Pr(\text{talks a little}) = \frac{1}{6}, \;\; \Pr(\text{completely silent}) = \frac{1}{6}$$

- When he is sad,

$$\Pr(\text{talks a lot}) = \frac{1}{6}, \;\; \Pr(\text{talks a little}) = \frac{1}{6}, \;\; \Pr(\text{completely silent}) = \frac{2}{3}$$

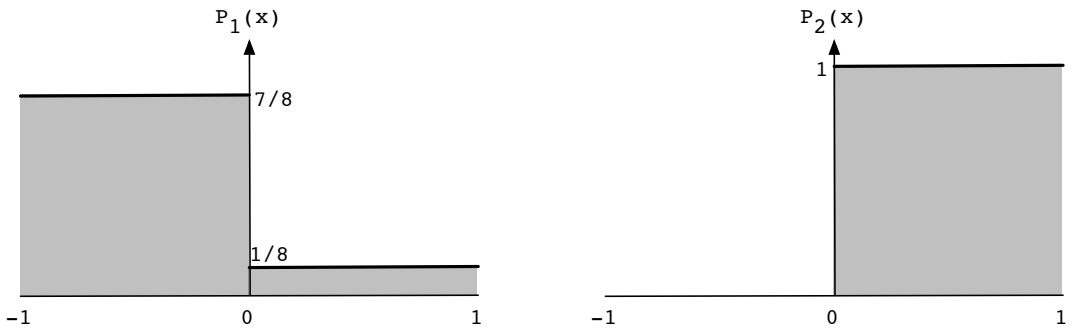
(a) Tonight, the man is just talking a little. What is his most likely mood?

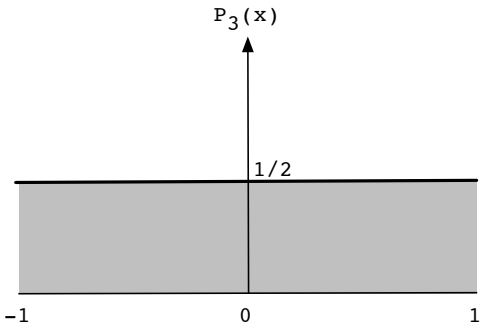
(b) What is the probability of the prediction in part (a) being incorrect?

17. (5 points) Suppose $\mathcal{X} = [-1, 1]$ and $\mathcal{Y} = \{1, 2, 3\}$, and that the individual classes have weights

$$\pi_1 = \frac{1}{3}, \;\; \pi_2 = \frac{1}{6}, \;\; \pi_3 = \frac{1}{2}$$

and densities P_1, P_2, P_3 as shown below.





What is the optimal classifier h^*? Specify it exactly, as a function from $\mathcal{X}$ to $\mathcal{Y}$.

18. ***Programming Question:*** (15 points total, 3 points each)

    (1) Create a classification problem with 3 classes, 15 features and 5000 rows.

    (2) Take the last 1000 rows to be the test set.

    (3) Run Gaussian naive bayes on this problem and report test accuracy.

    (4) Calculate class prior probabilities for each class in training data (first 4k rows).

    (5) Calculate the probability of the samples for each class in the test set.