CIS 4190/5190: Applied Machine Learning	Fall 2024

Homework 0

Handed Out: Wed August 28   *Due: Wed September 11, 7:59 p.m. ET*

- The goal of written HW0 is to give you an idea of the mathematical background expected in this course. You should have seen all this material before; if you are rusty on any of it, this homework should encourage you to revisit it. HW0 does not comprehensively test *all* the prerequisites; instead it samples from areas that are relevant to this class to remind you to **brush up on those areas broadly**.

- Your performance in this class will depend significantly on your ability to understand mathematical descriptions of ML algorithms in the lectures. **If you find that you are not familiar with more than 30% of the material in this problem set, you should strongly reconsider taking the course.** Speak to the instructor / TAs if you have concerns. If you've seen this material before but just need a refresher, the primers will help you.

- Discussion with your classmates or with course staff is permitted for clarifying the problem specification, but **do not ever share solutions / code**. Acknowledge all your discussions at the beginning of your report.

- You are encouraged to format your solutions using LaTeX. You may make a copy of the template here and start adding your solutions: https://www.overleaf.com/read/vjmrrbzpjmjv#f26f48. Handwritten & scanned solutions are permitted, but remember that **if your work is deemed not legible by your grader, they will not give you points for it**, and we will not accept any post hoc explanations. You will submit your solution manuscript for written HW0 as a single PDF file.

- We will be using Gradescope to collect your homework assignments. Submit your solution manuscript as a PDF file via Gradescope. Your solutions won't be graded if you don't **clearly annotate which page each question is answered on, on Gradescope.**

- **If you are on the waiting list** to join the class, we recommend that you still work on and submit the assignment.

- As long as you have answered more than **50%** of HW0 correctly, you will receive **full credit**.

1 Multiple Choice & Written Questions

Note: You do not need to show work for multiple choice questions. If formatting your answer in LaTeX, use our LaTeXtemplate `hw0_template.tex` (This is a read-only link. You'll need to make a copy before you can edit. Make sure you make only private copies.).

1. [Calculus] Let $f(x, y) = 3x^2 + y^2 - xy - 11x$

   a. (4pts) Find $\frac{\partial f}{\partial x}$, the partial derivative of f with respect to x. Find $\frac{\partial f}{\partial y}$. Select one of the below four options.

   A) $\frac{\partial f}{\partial x} = 6x - 6y - 10, \frac{\partial f}{\partial y} = y - x$

   B) $\frac{\partial f}{\partial x} = x - y - 11, \frac{\partial f}{\partial y} = 2y - x$

   C) $\frac{\partial f}{\partial x} = 6x - y - 11, \frac{\partial f}{\partial y} = 2y - x$

   D) $\frac{\partial f}{\partial x} = x - 6y - 11, \frac{\partial f}{\partial y} = 2y - x$

   b. (4pts) Find $(x, y) \in \mathbb{R}^2$ that minimizes f. Select one of the four options below.

   A) $(2, 1)$

   B) $(1, 1)$

   C) $(2, 2)$

   D) $(1, 2)$

2. [Probability] Assume that the probability of obtaining heads when tossing a coin is λ.

   a. (4pts) What is the probability of obtaining the first head on the $(k + 1)$-th toss? Select one of the below four options.

   A) $(1 - \lambda)^k \lambda^k$

   B) $(1 - \lambda)\lambda$

   C) $(1 - \lambda)\lambda^k$

   D) $(1 - \lambda)^k \lambda$

   b. (4pts) What is the expected number of heads if the coin is tossed k times? Select one of the below four options.

   A) $k \times \lambda$

   B) k^λ

   C) λ^k

   D) $(1 - \lambda)^k$

3. [Probability] Assume we know the following facts about COVID-19 testing:

   - 95% of infected people test positive. This is called the true positive rate.

   - 0.01% of uninfected people test positive. This is the false positive rate.

   - Based on population-level statistics in Pennsylvania, we expect that among all people who are tested, 10% are infected.

With this information, please answer the two questions below:

a. (4pts) John, who lives in Pennsylvania, tests positive. What is the probability that he is infected?

A) ≈ 0.999

B) ≈ 0.099

C) ≈ 0.5

D) 1

b. (4pts) With all other values remaining the same, what should the false positive rate change to, for John's probability of being infected to change to 0.95?

A) $\approx 0.556\%$

B) $\approx 0.056\%$

C) $\approx 0.1\%$

D) $\approx 1\%$

4. [Probability] The variance of a random variable X is defined as

$$Var(X) = E[(X - E[X])^2] .$$

a. (4pts) If $E[X] = 0$ and $E[X^2] = 1$, what is the variance of X? If $Y = a + bX$, where a and b are constants, what is the variance of Y? Select one of the below four options.

A) $Var(X) = \frac{1}{\sqrt{2}}, Var(Y) = a^2 + b^2$

B) $Var(X) = 1, Var(Y) = b^2$

C) $Var(X) = \sqrt{2}, Var(Y) = a^2$

D) $Var(X) = 2\sqrt{2}, Var(Y) = b^2$

b. (9pts) Prove that $Var(X) = E[X^2] - E[X]^2$. Please show your work in detail.

5. [Linear Algebra] Let

$$\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} 4 & 7 \\ 5 & 8 \\ 6 & 9 \end{bmatrix}$$

a. (4pts) What is the length of the vector $\mathbf{v}$, $\|\mathbf{v}\|_2$ (sometimes written as $\|\mathbf{v}\|$)? Select one of the below four options.

A) $\sqrt{3}$

B) -1

C) $\sqrt{5}$

D) 5

b. (4pts) What is the result of the multiplication $\mathbf{v^T M}$, where ^{T} indicates the transpose operation? Select one of the below four options.

   A) Undefined, sizes do not align.

   B) $\begin{bmatrix} 0 & 0 \\ 5 & 8 \\ -12 & -18 \end{bmatrix}$

   C) $\begin{bmatrix} -3 & -6 \end{bmatrix}$

   D) $\begin{bmatrix} -7 & -10 \end{bmatrix}$

c. (4pts) What is the result of the multiplication $\mathbf{M v^T}$? Select one of the below four options.

   A) Undefined, sizes do not align.

   B) $\begin{bmatrix} 0 & 0 \\ 5 & 8 \\ -12 & -18 \end{bmatrix}$

   C) $\begin{bmatrix} -3 & -6 \end{bmatrix}$

   D) $\begin{bmatrix} -7 & -10 \end{bmatrix}$

6. [Linear Algebra] Let

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$$

a. (7pts) What are the eigenvalues of the matrix $\mathbf{A}$? Please show your work in detail.

b. (4pts) Find the maximum value of $\mathbf{x^T A x}$ for any unit vector $\mathbf{x}$. Select one of the four options below.

   A) $-\frac{9}{4} + \frac{\sqrt{10}}{2}$

   B) $\frac{9}{2} - \frac{\sqrt{5}}{2}$

   C) $\frac{9}{2} + \frac{\sqrt{10}}{2}$

   D) $\frac{9}{4} + \frac{\sqrt{5}}{2}$

2 Python Programming Questions

A Google Colab notebook is linked in the "HW0 Coding" assignment on the class website. This will tell you everything you need to do, and provide starter code. See the recording from the first day of class for how to use Colab and how to submit homeworks.

- Submit the notebook as .ipynb file to the coding assignment on Gradescope.

- Remember to download and embed the figure for question 4 into the *written problem set solution pdf.*