CARNEGIE MELLON UNIVERSITY 10-606

HOMEWORK 1

DUE: Monday, Sept. 20, 2021

https://www.cs.cmu.edu/~10606

INSTRUCTIONS

- **Format:** Use the provided LaTeX template to write your answers in the appropriate locations within the *.tex files and then compile a pdf for submission. We try to mark these areas with STUDENT SOLUTION HERE comments. Make sure that you don't change the size or location of any of the answer boxes and that your answers are within the dedicated regions for each question/part. If you do not follow this format, we may deduct points.

  On this first homework only, you may type you answer or write by hand on the digital or printed pdf. Illegible handwriting will lead to lost points. However, we suggest that try to do at least some of your work directly in LaTeX as this will be required in future assignments.

- **How to submit:** Submit to Gradescope a pdf with your answers. Again, make sure your answer boxes are aligned with the original pdf template.

- **Policy:** See the course website for homework policies, including late policy, and academic integrity policies.

Name	
Andrew ID	
Hours to complete (nearest hour)	

1 Matrix Multiplication [9 pts]

1. **[3 pts]** Given two `numpy` 2D arrays 'X' and 'Y', write a Python statement that computes their matrix multiplication. You can assume `numpy` has been imported as `np`.

   Note: Indenting is a pain in LaTeX, but that's ok, you shouldn't need to use any tabs in your code for this question. E.g., don't use any for loops.

   Note: Your code should work, so make sure you test it.

```
import numpy as np

# create a randomly initialized 10x10 matrix and assign to X
X = np.random.randn(10,10)

# create a randomly initialized 10x10 matrix and assign to Y
Y = np.random.randn(10,10)
```

2. **[6 pts]** Suppose you have a vector $\mathbf{v} \in \mathbb{R}^M$ and matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$. Write an expression for $\mathbf{v}^T e^{\mathbf{A}\mathbf{v}}$ without using any linear algebra notation such as matrices or vectors. In other words, write $\mathbf{v}^T e^{\mathbf{A}\mathbf{v}}$ in terms of the elements of $\mathbf{A}$, the elements of $\mathbf{z}$, and M.

   Notation: the exponential function applied to a matrix is simply equal to applying the exponential function to each element of the matrix. For example, for $\mathbf{B} = \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix}$,

$$e^{\mathbf{B}} = \begin{bmatrix} e^{B_{1,1}} & e^{B_{1,2}} \\ e^{B_{2,1}} & e^{B_{2,2}} \end{bmatrix}$$

   You do not need to show your work.

   You may write the element in the i-th position of $\mathbf{v}$ as v_i. You may write the element in the i-th row and j-th column of $\mathbf{A}$ as $A_{i,j}$.

   Do not include the full vector $\mathbf{v}$ or the full matrix $\mathbf{A}$ in your answer.

   You must be precise. You man not use "$\cdots$" for example. However, you may use summation and product notation, e.g., $\sum_{i=1}^{N}$ and $\prod_{i=1}^{N}$

   Hint: Write this out on paper for $M = 2$.

2 Linear Systems and Linear Algebra [10 pts]

1. True or False

   (a) **[1 pts]** A system of N linear equations with N unknowns has at least one solution. *Select the best answer.*

   ◯ True ◯ False

   (b) **[1 pts]** A system of N linear equations with N unknowns has at most one solution. *Select the best answer.*

   ◯ True ◯ False

   (c) **[1 pts]** If a square matrix is full rank, it cannot be inverted. *Select the best answer.*

   ◯ True ◯ False

   (d) **[1 pts]** Rank of matrix product $\mathbf{AB}$ can be greater than the rank of either $\mathbf{A}$ or $\mathbf{B}$. *Select the best answer.*

   ◯ True ◯ False

2. **[3 pts]** Consider the following matrix:

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & 5 \\ a & b & c \\ d & e & f \end{bmatrix}$$

   Complete the matrix such that the rank of $\mathbf{A}$ equals 1. Restrictions:

   - $a, b, c, d, e, f \in \mathbb{Z}$ (integers)
   - $a, b, c, d, e, f \notin -3, 2, 5$
   - a, b, c, d, e, f are all different

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & 5 \\ ?? & ?? & ?? \\ ?? & ?? & ?? \end{bmatrix}$$

3. **[3 pts]** Consider the following matrix:

$$\mathbf{B} = \begin{bmatrix} 2 & 5 & 2 & 5 \\ 8 & 8 & -2 & 3 \\ 0 & 0 & -2 & -1 \\ -2 & -1 & 0 & 0 \end{bmatrix}$$

Compute the matrix rank of $\mathbf{B}$. You may do it by hand or by using Python and numpy.

You must show your work. If you use Python and numpy:

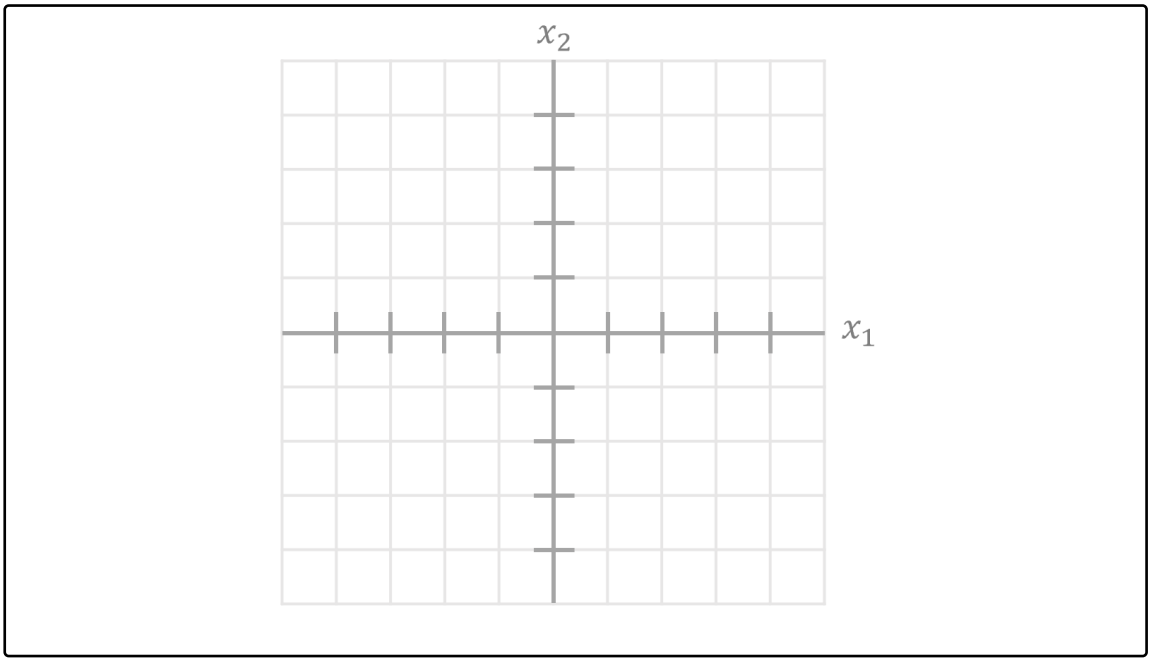
- Just paste your working code in the Work box
- If your code happens to contain LaTeX special characters, such as an '_' just put a backlash in front of that character in your code
- Don't worry if LaTeX goofs up your Python format a bit, e.g. drops tabs, etc.

Rank:

Work:

3 Plotting Linear (Affine) Functions [8 pts]

1. **[4 pts]** Given $\mathbf{w} = [4, -8]^T$ and $b = -16$, draw the line $\{\mathbf{x} \mid \mathbf{w}^T\mathbf{x} + b = 0, \ \ \mathbf{x} \in \mathbb{R}^2\}$ in the 2-D plane. Please indicate the axes and origin of the plane.



2. Suppose we want to move the line closer to the origin.

   (a) **[2 pts]** For a fixed $\mathbf{w}$, how should we change $|b|$?

   ◯ Increase ◯ Decrease ◯ Not possible

   (b) **[2 pts]** For a fixed b, how should we change $\|\mathbf{w}\|_2$?

   ◯ Increase ◯ Decrease ◯ Not possible

4 Linear Regression [18 pts]

We are given the following four-point data set, $\mathcal{D} = \{(x_1^{(i)}, x_2^{(i)}, y^{(i)}\}_{i=1}^4$:

x_1	x_2	y
2	-1	0
0	2	-1
-3	-2	0
1	3	2

We are going to use a linear model with no offset/bias term for our prediction function:

$$\hat{y} = w_1 x_1 + w_2 x_2$$

Our goal is to find the parameters that minimize a mean squared error objective function given our data:

$$J(w_1, w_2; \mathcal{D}) = \frac{1}{4} \sum_{i=1}^{4} \left(y^{(i)} - \hat{y}^{(i)} \right)^2$$

1. **[8 pts]** Prove that this objective function can fit into the following form by computing values for a, b, c, d, e, f given our dataset:

$$J(w_1, w_2; \mathcal{D}) = a w_1^2 + b w_2^2 + c w_1 w_2 + d w_1 + e w_2 + f$$

a: b: c: d: e: f:

Work (optional):

2. **[2 pts]** What shape is function $J(w_1, w_2; \mathcal{D})$ when you plug in the values for a, b, c, d, e, f that you found above? We are looking for a one word answer for the shape. Be specific, e.g. don't say rectangle if it is in fact a square.

3. **[8 pts]** Write equations for the partial derivatives of our objective with respect to the parameters.

   To avoid potential error propagation from the previous part, write the derivatives in terms of symbols a, b, c, d, e, f rather than the values you computed above.

   $\frac{\partial J}{\partial w_1}$:

   $\frac{\partial J}{\partial w_2}$:

5 Collaboration Policy

After you have completed all other components of this assignment, report your answers to the following collaboration questions.

1. Did you receive any help whatsoever from anyone in solving this assignment? If so, include full details including names of people who helped you and the exact nature of help you received.

2. Did you give any help whatsoever to anyone in solving this assignment? If so, include full details including names of people you helped and the exact nature of help you offered.

3. Did you find or come across code that implements any part of this assignment? If so, include full details including the source of the code and how you used it in the assignment.