CS 534: Homework #0

Submission Instructions: The homework is due on Sept 6th at 11:59 PM EDT on Gradescope. The entry code for Gradescope is ZZN4JW. A part of your homework will be automatically graded by a Python autograder. The autograder will support Python 3.10. Additional packages and their versions can be found in the `requirements.txt`. Please be aware that the use of other packages and/or versions outside of those in the file may cause your homework to fail some test cases due to incompatible method calls or the inability to import the module. We have split homework 1 into 2 parts on Gradescope, the autograded portion and the written answer portion. If *either of the two parts is late, then your homework is late.*

1. **Enter answers for (online) into HW0-Written**: Enter your answers for the (Written) questions directly in the HW0-Written assignment. You should be able to save as you go.

2. **Submit code to the HW0-Code Assignment**: Your submitted code must contain the following two files: 'hw0.py', 'README.txt'. You are welcome to submit other files but the autograder will only copy these files when running the test cases so make sure they are self-contained (i.e., capable of running standalone). Make sure you always upload *ALL* files when you (re)submit. The `README.txt` file *must contain a signed honor statement* that contains the following words:

   ```
   /* THIS CODE IS MY OWN WORK, IT WAS WRITTEN WITHOUT CONSULTING CODE
   WRITTEN BY OTHER STUDENTS OR LARGE LANGUAGE MODELS LIKE CHATGPT.
   Your_Name_Here */

   I collaborated with the following classmates for this homework:
   <names of classmates>
   ```

1. **(Written)** Have you read through the course website, noted the important dates, and the class policies?

2. **(Written)** Which of the following courses have you taken?

   (i) Have you taken any courses on Probability/Statistics? If yes, please write down the institution name, course department, course name, and course grade.

   (ii) Have you taken any courses on Linear Algebra? If yes, please write down the institution name, course department, course name, and course grade.

   (iii) Have you taken any courses on Optimization? If yes, please write down the institution name, course department, course name, and course grade.

   (iv) Have you taken any courses on Data Mining/Pattern Recognition/Machine Learning? If yes, please write down the institution name, course department(s), course name(s), and your course grade(s).

3. **(Written)** An urn has 3 red balls, 4 blue balls, and 5 green balls. Alice draws a ball from the urn, and then Bob draws a ball. What is the probability that Bob got a green ball?

4. **(Written)** Let $\mathbf{X} \in \mathbb{R}^{n \times p}$ and $\mathbf{y} \in \mathbb{R}^n$ be given.

(i) Find the optimal vector $\mathbf{w}^* \in \mathbb{R}^p$ which solves the following problem:

$$\min_{\mathbf{w} \in \mathbb{R}^p} \ \frac{1}{2} \|\mathbf{y} - \mathbf{Xw}\|^2$$

(Hint) Consult the *Matrix Cookbook* if you want to look up expressions for derivatives in matrix/vector form.

(ii) Does your solution above work if $\mathbf{X}$ is not full rank? If not, name one way to compute $\mathbf{w}^*$.

5. **(Written)** Honor Code Acknowledgement. The CSI/CS graduate programs, and indeed all of Emory University, take academic integrity and honesty very seriously. Any breach or violation of ethical academic conduct will be considered a serious offense. The CSI/CS programs will pursue the severest possible penalties.

All students are expected to know and follow the LGS honor code or Emory College honor code. Specific rules and expectations are additionally imposed by specific programs and courses. In CS534 F23, the following practices are effective:

(a) During the in-class midterm, all electronic devices will be powered off. Possession of an active device shall be considered an honor code violation.

(b) Programming assignments will be checked for similarity and any suspected plagiarism will be dealt with via a full investigation and honor code processes.

(c) The instructor reserves the right to assign seats to students in class, including the option of re-assignment during a class or exam.

(d) Participation in class and appearing for exams constitute implicit acknowledgment and affirmation of abiding by the honor code.

Thank you for upholding the principles and values of academic integrity in our programs.

I hereby affirm by writing my name that I will abide by the honor code set forth in CS534.

6. **(Code)** Consider a 3-valued random variable X such that $P(X = 1) = 0.35$, $P(X = 0) = 0.45$ and $P(X = -1) = 0.2$. Write a Python function `draw_samples` that takes in an integer n and draws n random samples from X.

7. **(Code)** Given a `numpy` 1-d array (e.g., array([1,2,3,4])), write a Python function `sum_squares` that takes the array and computes the sum of squares of this array using `numpy.dot`.

8. **(Code)** Your 3-year-old daughter and 7-year-old son are troublemakers. While you are busy grading assignments, they came up with this game of mixing her apple juice and his oreo milkshake. She first pours 35% of her apple juice into the cup that has his milkshake, then he pours 20% of whatever-it-is-now in the milkshake into the apple juice cup. They started with exactly 1 cup of apple juice and 1 cup of oreo milkshake. Write a python function `troublemakers` that takes in an integer n denoting the number of cycles and returns a 1-d numpy array of 2 values with the first value containing the liquid in the apple juice cup and the second value containing the liquid in the milkshake cup.