COMS 4771 S18 Homework-0
(Due Monday 11:59 PM January 22)

Instructions

This is a calibration assignment (HW0). The goal of this assignment is for you to recall basic concepts, and get familiarized with the homework submission system (gradescope). Everyone enrolled or on waitlist (intending to enroll) must submit this assignment by the due date. Anyone who does not submit HW0 by the due date will get a score of zero. Students will be cleared from the waitlist to register for the course based on their HW0 score (higher the score, better the chances). The score received on this assignment will not count towards your final grade in this course. You must show your work to receive full credit.

All homeworks (including this one) should be typesetted properly in pdf format. Handwritten solutions will not be accepted. You must include your name and UNI in your homework submission. It is **highly recommended** that the students who are comfortable with using latex use this .tex file and this .cls file

You do **not** need to submit any source code with your write-up.

Problem 1

Let X and Y be discrete random variables with joint probability distribution given by the following table:

	$Y = 1$	$Y = 2$
$X = 1$	0.1	0.2
$X = 2$	0.2	0.1
$X = 3$	0.3	0.1

(a) What is the marginal distribution of X?
(b) What is $\mathbb{P}(Y = 1 \mid X = 2)$?
(c) What is the variance of Y?
(d) Let $f\colon \mathbb{R} \to \mathbb{R}$ be the function defined by

$$f(x) = x^2, \quad x \in \mathbb{R}.$$

What is $\mathbb{E}[f(X) \mid Y = 1]$?
(e) Continuing from part (d), what is the expected value of $\mathbb{E}[f(X) \mid Y]$?

Problem 2

Let $\theta > 0$ be an arbitrary positive number, and consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{1}{Z}(e^{-\theta x} + e^{-2\theta x}) & \text{if } x \geq 0, \end{cases}$$

where $Z > 0$ is some positive constant that depends only on θ. Let X be a continuous random variable with probability density proportional to f.

(a) For what value of Z is f a probability density?
(b) What is the mean of X?
(c) What is the variance of X?

(In each case, give your answer in terms of θ.)

Problem 3

A researcher develops a blood test intended to detect if a person has a genetic disease; the result of the test is either "yes" or "no". Unfortunately, the test is not perfect, and the test confuses having the disease with having red hair. The test result is "yes" with probability 99% if the tested person has the genetic disease or has red hair; the test result is "no" with probability 99% if the tested person does not have genetic disease and does not have red hair.

The probability of a person having the genetic disease is 1%, and the probability of a person having red hair is also 1%.

If having the genetic disease is independent of having red hair, what is the probability that a tested person has the genetic disease if the test returns "yes"?

Problem 4

Let A be the 6×6 symmetric matrix defined as follows, as the product of three 6×6 matrices:

$$A := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

(The first and third matrices in this product are identical, and the second matrix is diagonal.)

(a) Let R be the row space of A. What is the dimension of R?
(b) What is the largest eigenvalue of A^3?
(c) Let V be the span of the eigenvectors of A with eigenvalue $1/2$. What is the orthogonal projection operator for V? Give your answer as a 6×6 matrix.

Problem 5

Let $J\colon \mathbb{R}^6 \to \mathbb{R}$ be the function defined by

$$J(x) := \|Ax - b\|_2^2 + \frac{1}{4}\|x\|_2^2, \quad x \in \mathbb{R}^6,$$

where

- $A \in \mathbb{R}^{6\times 6}$ is the matrix defined in Problem 4,
- $b := (1, \dots, 1) \in \mathbb{R}^6$, and
- $\|\cdot\|_2$ denotes the Euclidean norm (i.e, $\|v\|_2 = \sqrt{v_1^2 + \cdots + v_6^2}$ for $v = (v_1, \dots, v_6) \in \mathbb{R}^6$).

(a) What is $\nabla J(x)$ (i.e., the gradient of J) evaluated at $x = (0, \dots, 0)$?
(b) Find a value of x that minimizes J (i.e., find a value in $\arg\min_{x\in\mathbb{R}^6} J(x)$). *Hint*: It suffices to find a critical point of f.
(c) Write a program in a scientific programming language of your choice that implements the Richardson iteration to (approximately) find a minimizer of the function $g\colon \mathbb{R}^6 \to \mathbb{R}$, where

$$g(x) := \|Ax - b\|_2^2, \quad x \in \mathbb{R}^6.$$

The Richardson iteration starts with an initial vector $x^{(0)} \in \mathbb{R}^6$, and computes subsequent vectors using the following recursion:

$$x^{(k)} := x^{(k-1)} + \eta A^\top (b - Ax^{(k-1)}), \quad k = 1, 2, \dots,$$

where $\eta > 0$ is a "step size" parameter. Run your program starting with $x^{(0)} := (0, \dots, 0)$ and $\eta := 1/2$ for up to $k = 1000$ steps. Save the resulting 1000 vectors, as you will need them for the remaining parts of this problem. What are the vectors $x^{(100)}$ and $x^{(1000)}$ produced by your program?
(d) Continuing from part (c), let $v := (0, 0, 0, 1, -1, 0)$ and $\tilde{x} := x^{(1000)} + v$. Compute $\|Ax^{(1000)} - b\|_2^2$ and $\|A\tilde{x} - b\|_2^2$. How do these values compare? Also compute $\|x^{(1000)}\|_2^2$ and $\|\tilde{x}\|_2^2$. How do these values compare?
(e) Run your program again starting with $x^{(0)} := (0, \dots, 0)$, but this time with $\eta := 3/4$. Save the resulting 1000 vectors separately from those obtained in part (c); we'll call them $y^{(1)}, \dots, y^{(1000)}$. Make a plot with two curves. The first curve is $\|Ax^{(k)} - b\|_2^2$ as a function of k, for $1 \le k \le 100$. The second curve is $\|Ay^{(k)} - b\|_2^2$ as a function of k, for $1 \le k \le 100$. Use different line styles and different colors to distinguish the two curves. Add a legend to the plot that identifies each curve. Add an appropriate title to the plot, and add appropriate labels for the horizontal and vertical axes.