CSCI.635.01 - Intro to Machine Learning (CSCI63501...

Homework 1

1. Let $\boldsymbol{a}$ = [1, 4, 5]^T, $\boldsymbol{b}$ = [3, -1, 9]^T and $\boldsymbol{c}$ = [-4, -1, 3]^T. Compute the following ***by hand***:
    - a. [1 point] $\mathbf{a}^T\boldsymbol{b}$
    - b. [1 point] $\boldsymbol{ab}^T$
    - c. [1 point] $(\boldsymbol{ab}^T)\boldsymbol{c}$
    - d. [1 point] $\boldsymbol{c}^T(\boldsymbol{ab}^T)$
2. Read the exercise on projections on the Deep Learning book website. You may want to try to do it yourself before reading the solution, as you will better understand it that way (btw calculus, specifically partial differentials and integration, are very important skills to brush up on). Suppose we had a point $\boldsymbol{p}$ = (x, y) in $\mathbb{R}^2$, which we can write as the vector [x,y]^T, but all we know of it is y and its projection μ on a unit vector $\boldsymbol{v}$ = [v_x, v_y]^T ≠ [0,1]^T, where [0,1]^T is the unit vector in the direction of the y-axis, i.e., y is $\boldsymbol{p}$'s projection on [0,1]^T. Crucially, we are not given x.
    - a. [6 points] Draw a figure showing the x- and y-axes, the unit circle, the arbitrary unit vector v, and p (which you may assume is also a unit vector for simplicity's sake). Label all the variables and vectors named above.
    - b. [5 points] Give an expression for x in terms of $\boldsymbol{p}$, y, μ, $\boldsymbol{v}$, v_x, and v_y (you don't have to use all of the variables).
3.
    - a. [5 points] Write a program, called matrix_exp.py, that runs four experiments. Each experiment should run 10000 trials. In each trial, randomly generate a 10x10 numpy array and test whether it is invertible using numpy.linalg.inv.

      Each experiment should generate the random numbers to populate the arrays in a dIfferent way:
    - Experiment 1 should use numpy.random.rand(10,10)
    - Experiment 2 should use numpy.random.randint(2, size=(10,10))

  - Experiment 3 should work like numpy.random.randint(2, size=(10,10)) except that there should be a 75% chance that each entry is 0 and 25% chance that the entry is 1.
  - Experiment 4 should be like experiment 3, in that there should be a 75% chance that each entry is 0. But otherwise, instead of choose 1, you should choose uniformly at random between 0 and 1 a floating point number.

  - b. [5 points] Provide a table where the columns are experiments 1-4 and the rows contain: (1) The number of invertible matrices (2) the standard deviation (the square root of the variance) around the mean (i.e., the number of invertible matrices /10000), and (3) the 90% confidence interval (see the "Basic steps" section of this wikipedia article for one way to do this).

  - c. [5 points] What happens in each random model when the size of the array increases? Justify your answer the best way you can. Discuss the shortcomings of your approach. Try to write no more than a page or two. Use any experimental and/or mathematical analysis you see fit.