Mock Final Exam 1 - DSC 40A, Spring 2024

Full Name:

PID:

Seat Number:

Instructions:

- This mock exam consists of 5 questions. **Advice: Read all of the questions before starting to work, because the questions are not sorted by difficulty.**
- You have 60 minutes to complete this mock exam.
- **Note that the main purpose of this mock exam is to give you practice with answering unfamiliar questions in an exam-like environment. The questions on the real exam may be formatted differently than those here.**

By signing below, you are agreeing that you will behave honestly and fairly during and after this exam.

Signature:

Version A

Please do not open your exam until instructed to do so.

Question 1

Given a dataset $y_1 \leq y_2 \leq ... \leq y_n$, define the following functions:

$$R_{\text{abs}}(h) = \frac{1}{n}\sum_{i=1}^{n} |y_i - h| \qquad R_{\text{sq}}(h) = \frac{1}{n}\sum_{i=1}^{n} (y_i - h)^2$$

Parts (a), (b), and (c) below concern $R_{\text{abs}}(h)$, while parts (d) and (e) concern $R_{\text{sq}}(h)$.

In each of the options below, you're presented with two expressions. Determine which expression is larger, and determine by how much it is larger. If it's impossible to tell which of the two expressions is larger, write "impossible" in the provided box.

a) Let c be such that $c < c + 1 < y_1$.

Which is larger? ◯ $R_{\text{abs}}(c)$ ◯ $R_{\text{abs}}(c + 1)$

By how much?

b) Let c be such that $y_n < c < c + 2$.

Which is larger? ◯ $R_{\text{abs}}(c)$ ◯ $R_{\text{abs}}(c + 2)$

By how much?

c) Suppose $n = 10$. Let c be such that $y_3 < c < c + 1 < y_4$.

Which is larger? ◯ $R_{\text{abs}}(c)$ ◯ $R_{\text{abs}}(c + 1)$

By how much?

d) Let c be such that $c < y_1$.

Which is larger? ◯ $R_{\text{sq}}(c)$ ◯ $R_{\text{sq}}(c - 1)$

By how much?

e) Let c be such that $c < y_n$.

Which is larger? ◯ $R_{\text{sq}}(c)$ ◯ $R_{\text{sq}}(c + 1)$

By how much?

Question 2

Suppose you have a dataset of n points $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ where the least squares regression line is $H^*(x) = 3x + 7$. Suppose also that $n = 10$, the average x is $\bar{x} = 2$, and the variance of x is $\sigma_x^2 = 7$.

a) Suppose we move just one point up by 35 units: the point $(3, 2)$ becomes $(3, 37)$. Determine the slope and intercept of the new regression line. Show your work, and put a $\boxed{\text{box}}$ around your final answers.

b) Suppose instead we move all points up by 35 units. Determine the slope and intercept of the new regression line. Show your work, and put a $\boxed{\text{box}}$ around your final answers.

Question 3

Suppose $M \in \mathbb{R}^{m \times n}$ is a matrix, $\vec{v} \in \mathbb{R}^n$ is a vector, and $s \in \mathbb{R}$ is a scalar.

Determine whether each of the following quantities is a matrix, vector, scalar, or undefined. If the result is a matrix or vector, state its dimensions in the blank provided.

a) $M\vec{v}$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

b) vM

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

c) $\vec{v}^2$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

d) $M^T M$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

e) MM^T

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

f) $\vec{v}^T M\vec{v}$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

g) $(sM\vec{v}) \cdot (sM\vec{v})$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

h) $(s\vec{v}^T M^T)^T$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

i) $\vec{v}^T M^T M\vec{v}$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

j) $\vec{v}\vec{v}^T + M^T M$

◯ matrix of size ________ ◯ vector of size ________
◯ scalar ◯ undefined

Question 4

Consider the vectors $\vec{u}$ and $\vec{v}$, defined below.

$$\vec{u} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$$

a) Determine the angle between $\vec{u}$ and $\vec{v}$. Put a $\boxed{\text{box}}$ around your final answer. Leave your answer in the form $\cos^{-1}(\cdot)$.

b) Determine the projection of $\vec{u}$ onto $\text{span}(\vec{v})$. Give your answer in the form of a **vector**. Put a $\boxed{\text{box}}$ around your final answer.

Question 5

You have data on three different apartments for rent. Each apartment's data is given as a tuple $(x^{(1)}, x^{(2)}, x^{(3)}, y)$, where:

- $x^{(1)}$ is the number of bedrooms,
- $x^{(2)}$ is the floor number,
- $x^{(3)}$ is the proportion of positive Yelp reviews for the apartment complex, and
- y is the monthly rent.

Your data is:

$$(2, 9, 0.5, 500), (3, 8, 0.7, 700), (4, 10, 0.8, 900)$$

a) You want to predict an apartment's rent using a hypothesis function of the form:

$$H(x^{(1)}, x^{(2)}, x^{(3)}) = w_0 + w_1 \cdot \sqrt{x^{(1)}} + w_2 \cdot \left(x^{(2)}\right)^2 + w_3 \cdot x^{(3)}$$

Write down the design matrix, X, and the observation vector, $\vec{y}$. State how you would use these quantities to find the hypothesis function, but don't actually do any calculations.

b) Suppose you'd like to instead fit the hypothesis function:

$$H(x^{(1)}, x^{(2)}, x^{(3)}) = w_1 \cdot \sqrt{x^{(1)}} + w_2 \cdot \left(x^{(2)}\right)^2 + w_3 \cdot x^{(3)}$$

Will your answer to part (a) change? If so, how?