Midterm 1 - DSC 40A, Winter 2024

Instructions

- This is a 50-minute exam consisting of 5 questions worth a total of 40 points.
- The only allowed resource is the provided reference sheet.
- No calculators.
- Please write neatly and stay within the provided boxes.
- You may fill out the **front page only** until you are instructed to start.

Statement of Academic Integrity

By submitting your exam, you are attesting to the following statement of academic integrity.

I will act with honesty and integrity during this exam.

Name:	Solutions
PID:	A12345678
Seat you are in:	

Version - A

1. (10 points) Consider a dataset D with 5 data points $\{7, 5, 1, 2, a\}$, where a is a positive real number. Note that a is not necessarily an integer.

a) (2 points) Express the mean of D as a function of a, simplify the expression as much as possible.

$$Mean_D = \boxed{3 + \frac{a}{5}}$$

b) (3 points) Depending on the range of a, the median of D could assume one of three possible values. Write out all possible median of D along with the corresponding range of a for each case. Express the ranges using double inequalities, e.g., i.e. $3 < a \leq 8$:

$Median_D = \boxed{2}$ if a is in the range of $\boxed{0 < a \leq 2}$

$Median_D = \boxed{a}$ if a is in the range of $\boxed{2 < a < 5}$

$Median_D = \boxed{5}$ if a is in the range of $\boxed{a \geq 5}$

c) (5 points) Given that $Mean_d < Median_D$, determine the range of a that satisfies this condition. make sure to show your work

Range of a:

Supporting Work:

Solution: Since there are 3 possible median values, we will have to discuss each situation separately. In case 1, when $0 < a \leq 2$, $Median_D = 2$, therefore we have:

$$3 + \frac{a}{5} < 2$$
$$a < -5$$

But $a < -5$ is in conflict with the condition $0 < a \leq 2$, therefore there is no solution in this situation, and $Median_D = 2$ is impossible.

In case 2 when $2 < a < 5$, $Median_D = a$, therefore we have:

$$3 + \frac{a}{5} < a$$
$$3 < \frac{4}{5}a$$
$$a > \frac{15}{4}$$

So a has to be larger than $\frac{15}{4}$. But remember from the prerequisite condition that $2 < a < 5$. To satisfy both conditions, we must have $\frac{15}{4} < a < 5$.

In case 3 when $a \geq 5$, $Median_D = 5$, therefore we have:

$$3 + \frac{a}{5} < 5$$
$$a < 10$$

combining with the prerequisite condition, we have $5 \leq a < 10$

Combining the range of case 2 and 3, we have $\frac{14}{5} < a < 10$ as our final answer.

2. (4 points) Let $R_{sq}(h)$ represent the mean squared error of a constant prediction h for a given dataset. For the dataset $\{3, y_1\}$, the graph of $R_{sq}(h)$ has its minimum at the point $(5, r_1)$. Find out the value of y_1 and r_1

$y_1 =$ $\boxed{7}$ and $r_1 =$ $\boxed{4}$

Solution: The mean square error is written as:

$$R_{sq}(h) = \frac{1}{n}\sum_{i=0}^{n}(y_i - h)^2$$

Since we only have two data points ($n = 2$), the equation simplifies to:

$$R_{sq}(h) = \frac{1}{2}((y_0 - h)^2 + (y_1 - h)^2)$$

Taking derivative with respect to h, we have:

$$\frac{dR_{sq}(h)}{dh} = -(y_0 - h) - (y_1 - h)$$

We know that the derivative has to be 0 at the local minima, therefore at $h = 5$, we have:

$$\begin{aligned}\frac{dR_{sq}(h)}{dh} = -(3-5) - (y_1 - 5) &= 0 \\ y_1 &= 7\end{aligned}$$

So we know that the dataset is $3, 7$. Given all these information, we can calculate r_1 with:

$$\begin{aligned}R_{sq}(5) &= \frac{1}{2}((y_0-5)^2 + (y_1-5)^2) \\ &= \frac{1}{2}((3-5)^2 + (7-5)^2) \\ &= \frac{1}{2}(4+4) = 4\end{aligned}$$

3. (10 points) The hyperbolic cosine function is defined as $cosh(x) = \frac{1}{2}(e^x + e^{-x})$. In this problem, we aim to prove the convexity of this function using power series expansion.

a) (3 points) Prove that $f(x) = x^n$ is convex if n is an even integer.

Proof:

Solution: Take the second derivative of f:

$$\begin{aligned}f'(x) &= nx^{n-1} \\ f''(x) &= n(n-1)x^{n-2}\end{aligned}$$

If n is even, then n-2 must also be even, therefore $f''(x) = n(n-1)x^{n-2}$ will always be a positive number. This means the second derivative of $f(x)$ is always larger than 0 and therefore passes the second derivative test.

b) (2 points) Power series expansion is a powerful tool to analyze complicated functions. In power series expansion, a function can be written as an infinite sum of polynomial functions with certain coefficients. For example, the exponential function can be written as:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + ... \tag{1}$$

where $n!$ denotes the factorial of n, defined as the product of all positive integers up to n, i.e. $n! = 1 \times 2 \times 3 \times ... \times (n-1) \times n$. Given the power series expansion of e^x above, write the power series expansion of e^{-x} and explicitly specify the first 5 terms, i.e., similar to the format of Equation 1:

$e^{-x} = \sum_{n=0}^{\infty} \qquad =$

Solution: $e^{-x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{24} + ...$

c) (5 points) Using the conclusions you reached in **a)** and **b)**, prove that $cosh(x) = \frac{1}{2}(e^x + e^{-x})$ is convex.

Proof:

Solution: Given that:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + ...$$

$$e^{-x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{24} + ...$$

We can add their power series expansion together, and we will obtain:

$$\begin{aligned} e^x + e^{-x} &= \sum_{n=0}^{\infty} \frac{x^n}{n!} + \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(x)^n + (-x)^n}{n!} \end{aligned}$$

Within this infinite sum, if n is even, then the negative sign in $(-x)^n$ will disappear; if n is odd, then the negative sign in $(-x)^n$ will be kept and travel out of the parenthesis. Therefore we have:

$$\begin{aligned} e^x + e^{-x} &= \sum_{n=0}^{\infty} \frac{x^n + x^n}{n!}(\text{for even n}) + \sum_{n=0}^{\infty} \frac{x^n - x^n}{n!}(\text{for odd n}) \\ &= \sum_{n=0}^{\infty} \frac{2x^n}{n!}(\text{for even n}) \end{aligned}$$

Therefore, $cosh(x) = \frac{e^x + e^{-x}}{2}$ is a sum of x^n where n is even. Since we have already proved in a) that x^n are always convex for even n, $cosh(x)$ is an infinite sum of convex function and therefore also convex.

4. (10pt) Note that we have two simplified closed form expressions for the estimated slope w in simple linear regression that you have already seen in discussions and lectures:

$$w = \frac{\sum_i (x_i - \overline{x}) y_i}{\sum_i (x_i - \overline{x})^2} \tag{1}$$

$$w = \frac{\sum_i (y_i - \overline{y}) x_i}{\sum_i (x_i - \overline{x})^2} \tag{2}$$

where we have dataset $D = [(x_1, y_1), \ldots, (x_n, y_n)]$, sample means $\overline{x} = \frac{1}{n} \sum_i x_i, \quad \overline{y} = \frac{1}{n} \sum_i y_i$. Without further explanation, $\sum_i$ means $\sum_{i=1}^{n}$

a) (6pt) Are (1) and (2) equivalent? That is, is the following equality true? Prove or disprove it.

$$\sum_i (x_i - \overline{x}) y_i = \sum_i (y_i - \overline{y}) x_i$$

Proof:

Solution: True.

$$\begin{aligned}
&\sum_i (x_i - \overline{x}) y_i = \sum_i (y_i - \overline{y}) x_i \\
&\Leftrightarrow \sum_i x_i y_i - \overline{x} \sum_i y_i = \sum_i x_i y_i - \overline{y} \sum_i x_i \\
&\Leftrightarrow \overline{x} \sum_i y_i = \overline{y} \sum_i x_i \\
&\Leftrightarrow \frac{1}{n} \sum_i x_i \sum_i y_i = \frac{1}{n} \sum_i y_i \sum_i x_i
\end{aligned}$$

In fact, the least square estimator for slope is unique.

b) (2pt) True or False: If the dataset shifted right by a constant distance a, that is, we have the new dataset $D_a = (x_1 + a, y_1), \ldots, (x_n + a, y_n)$, then will the estimated slope w change or not?

◯ True ◯ False

Solution: False. By (1), the only term affecting w is $x_i - \overline{x}$, which is unchanged after shifting. Therefore, w is unchanged.

c) (2pt) True or False: If the dataset shifted up by a constant distance b, that is, we have the new dataset $D_b = [(x_1, y_1 + b), \ldots, (x_n, y_n + b)]$, then will the estimated slope w change or not?

◯ True ◯ False

Solution: False. By (2).

5. (6 points)

Suppose the following information is given for a linear regression:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} a \\ b \end{bmatrix} \qquad \vec{w}^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Where X is the design matrix, $\vec{y}$ is the observation vector, and $\vec{w}^*$ is the optimal parameter vector. Solve for parameter a and b using the normal equation, show your work.

Answer:

Supporting Work:

Solution: Since $\vec{w}^*$ is the optimal parameter vector, it must satisfy the Normal Equation:

$$X^T X\vec{w} = X^T\vec{y}$$

The left hand side of the equation will read:

$$X^T X\vec{w} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

The right hand side of the equation is given by:

$$X^T\vec{y} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ 2a-b \end{bmatrix}$$

By setting the left hand side and right hand side equal to each other, we will obatin the following system of equations:

$$\begin{bmatrix} 4 \\ 11 \end{bmatrix} = \begin{bmatrix} a+b \\ 2a-b \end{bmatrix}$$

So we obtained this set of equations:

$$\begin{aligned} 4 &= a+b \\ 11 &= 2a-b \end{aligned}$$

To sole this equation set, we can add them together:

$$\begin{aligned} 4+11 &= a+b+2a-b \\ 3a &= 15 \\ a &= 5 \\ b &= -1 \end{aligned}$$