National University of Computer and Emerging Sciences, Lahore Campus

Course Name:	Introduction to Data Science	Course Code:	DS2001
Degree Program:	BSDS	Semester:	Fall 2021
Exam Duration:	60 Minutes	Total Marks:	22
Paper Date:	18-10-2021	Weight	15 %
Section:	ALL	Page(s):	8
Exam Type:	Midterm-I		

Student : Name:_____ **Roll No.**_____ **Section:**_____

Instruction/Notes: Attempt the examination on the question paper and write concise answers. Extra pages are provided for rough work at the end. Do not attach extra sheets with the question paper. Don't fill the table titled Questions/Marks.

Question	Objective	1	2	3	Total
Marks	/ 10	/ 5	/ 3	/ 4	**/ 22**

Section 1 (Objective part) [points 10]

Clearly circle the correct options and explain your choice with reasoning.

Q 1: The line described by the regression equation attempts to

A) pass through as many points as possible. B) pass through as few points as possible.

C) minimize the number of points it touches. D) minimize the squared distance from the points.

Q 2: Suppose you have a dataset with m = 2000000 examples and n = 10 features for each example. You want to use multivariate linear Regression to fit the parameters θ to our data. Should you prefer gradient descent or the normal equation?

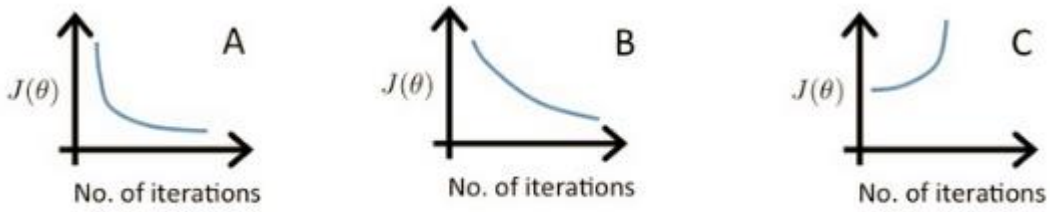
A. Gradient descent, since $(X^T X)^{-1}$ will be very slow to compute in the normal equation.
B. Gradient descent, since it will always converge to the optimal θ.
C. The normal equation, since gradient descent might be unable to find the optimal θ.
D. The normal equation, since it provides an efficient way to directly find the solution.

Q 3: Gradient Descent always finds the global optimum irrespective of the nature or shape of the cost function. Explain your choice with reasoning.

a) True b) False

Reason:

Q 4: Which of the following is true about below graphs (A,B, C left to right) between the cost function and Number of iterations?



Suppose l1, l2 and l3 are the three learning rates for A,B,C respectively. Which of the following is true about l1, l2 and l3?

A) l2 < l1 < l3 B) l1 > l2 > l3 C) l1 = l2 = l3 D) It depends on derivative

Q5. A multiple regression model has the form: $y = 2 + 7x_1 + 3x_2$. As x1 increases by 1 unit (holding x_2 constant), y will

(A) decrease by 7 units (B) increase by 7 units

(C) decrease by 3 units (D) increase by 3 units

Q6. Which of these is a reasonable definition of machine learning?

(A) Machine learning is learning from labeled data.

(B) Machine learning is the field of allowing robots to act intelligently.

(C) Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.

(D) Machine learning is the science of programming computers.

Q7. In regression analysis, the variable that is being predicted is

(A) the independent variable (B) the dependent variables

(C) usually denoted by x (D) none of the options

Q8. Let f be some function so that $f(\Theta_0, \Theta_1)$ outputs a number. For this problem, f is some arbitrary/unknown smooth function (not necessarily the cost function of linear regression, so f may have local optima). Suppose we use gradient descent to try to minimize $f(\Theta_0, \Theta_1)$ as a function of Θ_0 and Θ_1. Which of the following statements are true? (select all that apply.)

(A) If the first few iterations of gradient descent cause $f(\theta_0, \theta_1)$ to increase rather than decrease, then the most likely cause is that we have set the learning rate α to too large a value.

(B) If the learning rate α is too small, then gradient descent may take a very long time to converge.

(C) Even if the learning rate α is very large; every iteration of gradient descent will decrease the value of $f(\Theta_0, \Theta_1)$.

(D) No matter how Θ_0 and Θ_1 are initialized, so long as α is sufficiently small, we can safely expect gradient descent to converge to the same solution.

Q9. Write the list comprehension to pick out only negative integers from a give list 'myList'?

(A) [x < 0 in myList] (B) [x for x < 0 in myList] (C) [x in myList for x < 0] (C) [x for x in myList if x < 0]

Q10. What will be the output of this statement?

```
class Std_Name:
   def __init__(self, Std_firstName, Std_Phn, Std_lastName):
      self.Std_firstName = Std_firstName
      self. Std_PhnStd_Phn = Std_Phn
      self. Std_lastNameStd_lastName = Std_lastName

Std_firstName = "Wick"
name = Std_Name(Std_firstName, 'F', "Bob")
Std_firstName = "Ann"
name.lastName = "Nick"
print(name.Std_firstName, name.Std_lastName)
```

(A) Ann Bob (B) Ann Nick (C) Wick Bob (D) Wick Nick

Section 2 (Subjective part) (marks 12)

Q1. [3+1+1 Marks] Linear Regression: In the table below, training set of a small sample is given.

Assuming linear model for prediction : $h_\theta(x) = \theta_0 + \theta_1 x$

x	y
2	3
4	6
6	9

Here each row is one training example and total number of training examples are denoted by m.

where $\Theta_0 = 0$, answer the following questions.

(A). In Figure B, draw the curve for cost function ***J(Θ)***, assuming different values for Θ_1 (0, 1, 1.5, 2, 2.5 etc.).

(B) At what value of Θ_1, we get the global optimum for this cost function?

(c) At $\Theta_1 = 1$, draw the fitted line (in Figure A).

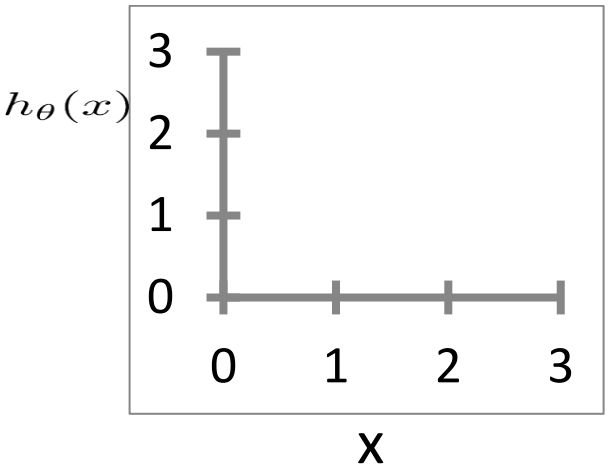


Figure A

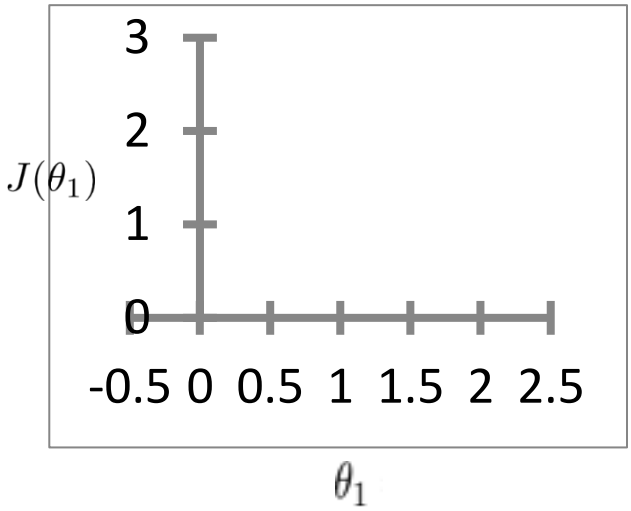


Figure B

Q2. [3 marks] What is convex and non-convex function? Explain with graph. What can be the problem for gradient descent algorithm if we have non-convex function?

Q3. Short questions (4 points)

a) [1 point] Find the local optimum (either maximum or minimum) of function $\mathbf{g(x) = x^3 - 9x^2 + 15x - 7}$.

(Note: you can use calculus to calculate the local optimum for this function)

b) [1 point] Suppose you have m=25 training examples with n=6 features (excluding the additional all-ones feature for the intercept term, which you should add). The normal equation is $\theta=(X^TX)^{-1}X^Ty$. For the given values of m and n, what are the dimensions of θ, X, and y in this equation?

Answer:

c) [1 point] Given the following python code, write down the output.

fName = "Ahmad" lName = "Ali" name = fName + lName a = name*3 b = a[-1] c = a[5:] d = a[5:8] print(a, b, c, d)	Output: ……

d) [1 point] Write a Python script to generate and print a dictionary that contains a number (between 1 and n) in the form (x, x*x). Sample Dictionary (n = 5) :

Expected Output : {1: 1, 2: 4, 3: 9, 4: 16, 5: 25}

Solution: