

Homework 1

Name: Omar Morsy

OAM15@pitt.edu

Question 1: Regression Problem - Plant Growth Prediction

A) Features (x)

Temperature

Plant spacing

Humidity percentage

Soil pH level

Daily hours of sunlight exposure

Nutrient levels

CO2 levels

B) Labels (y)

Plant growth rate in centimeters per week

C) Data Collection Method

Gather historical growth data

Record controlled experiments with varying parameters

Deploy soil testing probes for real-time soil composition data

Use automated imaging systems for regular height measurements

Install sensors to monitor environmental conditions

D) Potential Challenges

Seasonal variations affecting growth patterns

Plants responding differently at various growth stages

Disease or pest interference

Complex interactions between environmental factors

Time lag between condition changes and growth response

Question 2: Classification Problem - Vehicle Maintenance Prediction

A) Features (x)

Oil pressure readings

Tire pressure readings

Fuel consumption rate

Engine vibration patterns

Vehicle age and mileage

Battery voltage patterns

Maintenance history

B) Labels (y)

Three categories:

0: No maintenance needed

1: Routine maintenance needed

2: Urgent repair required

C) Data Collection Method

Track performance data before/after repairs

Collect historical maintenance records

Record engine audio data

Gather fleet management system data

Monitor vehicles through maintenance cycles

Use mechanic inspection reports

D) Potential Challenges

High sensor installation costs

Environmental effects on sensor readings

Balancing early detection vs. false alarms

Varying baseline readings across vehicle models

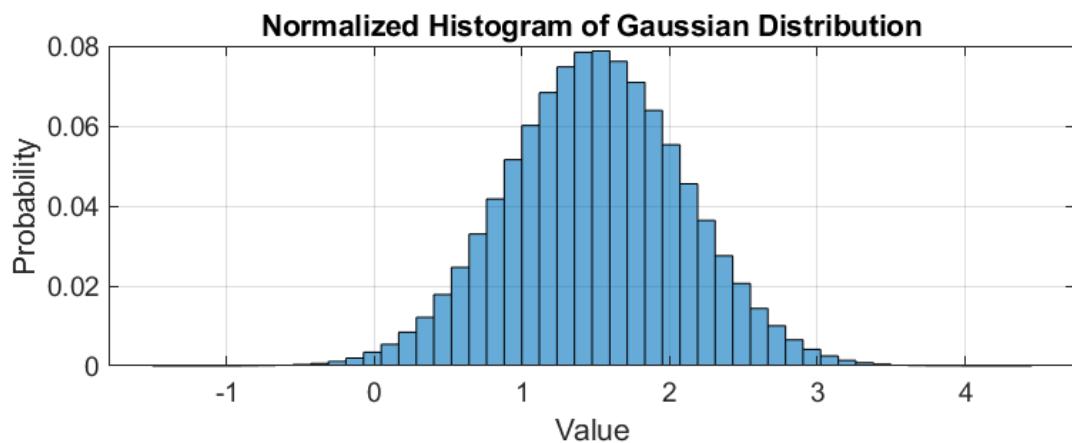
Need for extensive labeled data across vehicle types

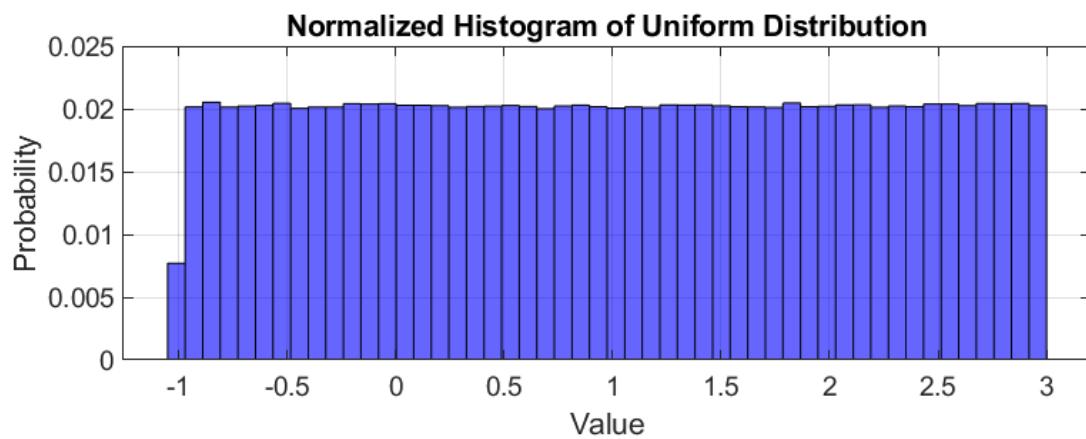
Question 3: Basic Operations

Part A and Part B

It is impossible to take a picture because it will represent a lot of random numbers

Part C - Normalized Histograms





Text response: The histogram for x shows a Gaussian distribution because it's symmetric about the mean and data near the mean occurs more frequently, creating a characteristic bell curve. The histogram for z demonstrates a uniform distribution as it shows equal likelihood of occurrence across all possible outcomes within its range.

Part D - Loop Execution Time

Text response: The execution time for adding 1 using a loop was approximately 0.0154 seconds.

```
0.2703
1.2611
1.8398
The time was taken by the loop: 0.0154 secs
The was taken without the loop: 0.0015 secs
```

Part E - Vectorized Operation

Text response: The execution time is 0.0015 seconds. The best way to add a constant to a long vector is not using a loop because the execution time is 0.0015 which is less than the execution time with a loop which is 0.0154 that means Matlab is faster without using a loop.

```
0.2703
1.2611
1.8398
The time was taken by the loop: 0.0154 secs
The was taken without the loop: 0.0015 secs
```

Part F - Vector Elements Analysis

Text response: I reran it twice (374795 and 374605), there is a small difference because z is randomly %generated each time, and theoretically, 37.5% of values should fall %between 0 and 1.5.

```
The was taken without the loop: 0.0012 secs  
The number of retrieved elements: 374605  
A =
```

```
The was taken without the loop: 0.0015 secs  
The number of retrieved elements: 374795  
fx >> _____  
Zoom: 100% | UTF-8
```

Question 4: Linear Algebra

Part A - Matrix Operations

Given matrix $A = [2 \ 1 \ 3; 5 \ 4 \ 8; 6 \ 3 \ 10]$

Results: [Next Page](#)

A =

2	1	3
5	4	8
6	3	10

min_A_rows =

1
4
3

max_A_columns =

6	4	10
---	---	----

A_min =

1

1
sum_A_column =
13 8 21

sum_A =
42

B =
4 1 9
25 16 64
36 9 100

Minimum values in each row: 1, 4 and 3

Maximum values in each column: 6, 4 and 10

Smallest value in A: 1

Sum of each column: 13, 8 and 21

Sum of all elements: 42

Matrix B (squared elements): it is equal exactly to A^2

Part B - Linear Equations Solution

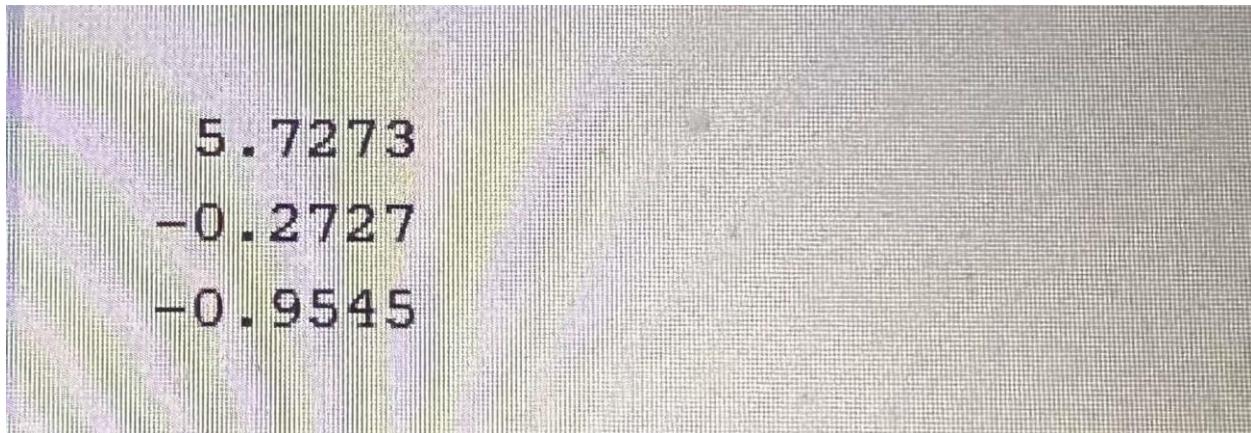
System of equations:

$$2x + 5y - 2z = 12$$

$$2x + 6y + 4z = 6$$

$$6x + 8y + 18z = 15$$

Text response: The result is : $x = 5.7273$, $y = -0.2727$ and $z = -0.9545$



Part C - Vector Norms

Vectors:

$$x_1 = [-5 \ 0 \ 2]$$

$$x_2 = [-1 \ -1 \ 0]$$

Hand calculations: L1-norm for x_1 : L2-norm for x_1 : L1-norm for x_2 : L2-norm for x_2 :

$$x_1 = \begin{bmatrix} -5 & 0 & 2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 & -1 & 0 \end{bmatrix}$$

$$L_1 \text{ norm } x_1 = |-5| + |0| + |2| = 7$$

$$L_1 \text{ norm } x_2 = |-1| + |-1| + |0| = 2$$

$$\begin{aligned} L_2 \text{ norm } x_1 &= \sqrt{(-5)^2 + (0)^2 + (2)^2} \\ &= \sqrt{25 + 4} = \sqrt{29} \approx 5.38 \end{aligned}$$

$$\begin{aligned} L_2 \text{ norm } x_2 &= \sqrt{(-1)^2 + (-1)^2 + 0^2} \\ &= \sqrt{2} \approx 1.4142 \end{aligned}$$

Analytical calculations and MATLAB verification: Next Page

```
L1_Norm_x1 =  
7  
]  
L1_Norm_x2 =  
2  
L2_Norm_x1 =  
5.3852  
L2_Norm_x2 =  
1.4142
```

3,10]
4,36,9,100]

Question 5: Splitting Data

Part A - Matrix Creation

Screenshot of matrix X and vector y:

Y =

1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

The matrix X (10*3) is:

1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10

Part B, C - Data Splitting

Three iterations of splitting results:

```
Command Window
    10   10   10
    .
    .
    .
the submatrix X_train is:
    3   3   3
    9   9   9
    6   6   6
    7   7   7
    4   4   4
    5   5   5
    1   1   1
    10  10  10

The submatrix X_test is:
    8   8   8
    2   2   2

The submatrix Y_train is:
    3   9   6   7   4   5   1   10
]

The submatrix Y_test is:
    8   2

the submatrix X_train is:
    1   1   1
    4   4   4
    9   9   9
    7   7   7
```

Editor - C:\myfiles\ECE1395\HW1\ps1_matlab_Morsy_Omar.m

Command Window

4	4	4
9	9	9
7	7	7
3	3	3
2	2	2
5	5	5
8	8	8

The submatrix X_test is:

6	6	6
10	10	10

The submatrix Y_train is:

1	4	9	7	3	2	5	8
---	---	---	---	---	---	---	---

The submatrix Y_test is:

6	10
---	----

the submatrix X_train is:

2	2	2
8	8	8
6	6	6
4	4	4
10	10	10
9	9	9
7	7	7
5	5	5

Zoom: 100% UTF-8 CRLF scrip

The screenshot shows the MATLAB Command Window with the following text output:

```
Command Window
The submatrix Y_train is:
    1     4     9     7     3     2     5     8

The submatrix Y_test is:
    6     10

the submatrix X_train is:
    2     2     2
    8     8     8
    6     6     6
    4     4     4
    10    10    10
    9     9     9
    7     7     7
    5     5     5

The submatrix X_test is:
    3     3     3
    1     1     1

The submatrix Y_train is:
    2     8     6     4     10    9     7     5

The submatrix Y_test is:
    3     1
```

Part D - Analysis of Results

Text response: I did not get the same submatrices for X and Y each time I split X because I shuffle the indices each time using randperm that is why I get different result every time.

The photos are above

