

Tristan Pitts
4040825
trp42@pitt.edu
Intro to Machine Learning
Homework 1

1. You are in the market for a new guitar. You decide to save some money by buying a used guitar. Determine a fair price for a guitar based on the brand, model and condition.

To solve: Collect data for several brands, models and conditions. Once we have enough data, we can use regression to find the best fit for our data. At this point we could plug in the information about any guitar we happen to be interested to see what the approximate fair price would be based on our data.

1.a) Brand of guitar, model of guitar and the condition of the guitar

1.b) The prices that previous guitars have sold for

1.c) Data could be collected from online stores such as eBay or Letgo. Additionally you could travel to music stores and look at their used guitars.

1.d) It may be difficult to collect enough data since there are three features and all can take on many values.

2. You'd like to sort the music you have saved on your computer by genre but you don't have the time to go through all of it individually and none of the files have metadata associated with them.

To solve: Select a small subset that contains songs from all genres you wish the songs to be classified into. Repeat training and test for accuracy. Once accuracy is satisfactory allow the program to have access to the full music library.

2a) BPM, types of instruments heard, vocals/vocal type, song length (certain metal subgenres have longer average song length)

2.b) Genres of music (metal, electronic, rap, rock etc.)

2.c) You could select a smaller subset of music you have from different genres as your training data

2.d) Songs within the same genre can still be very different. You may have to use a very large subset of songs to get accurate performance.

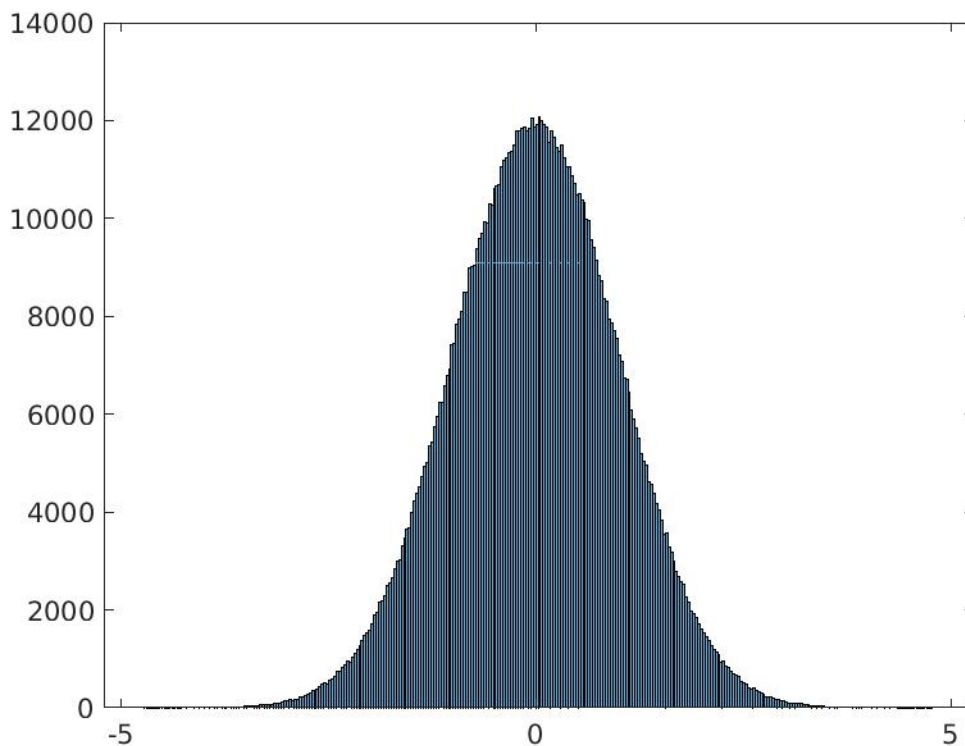
3.a)

```
The min element of a is -17.681892
The max element of a is 18.311326
The mean of the elements in a is -0.000345
The standard deviation of the elements in a is 3.703094
```

3. b)

```
The min element of a is -17.681892
The max element of a is 18.311326
The mean of the elements in a is -0.000345
The standard deviation of the elements in a is 3.703094
```

3.c) Yes, the histogram appears to be a Gaussian distribution



3.d) The elapsed time was 0.03 seconds

```
Elapsed Time: 0.030000
```

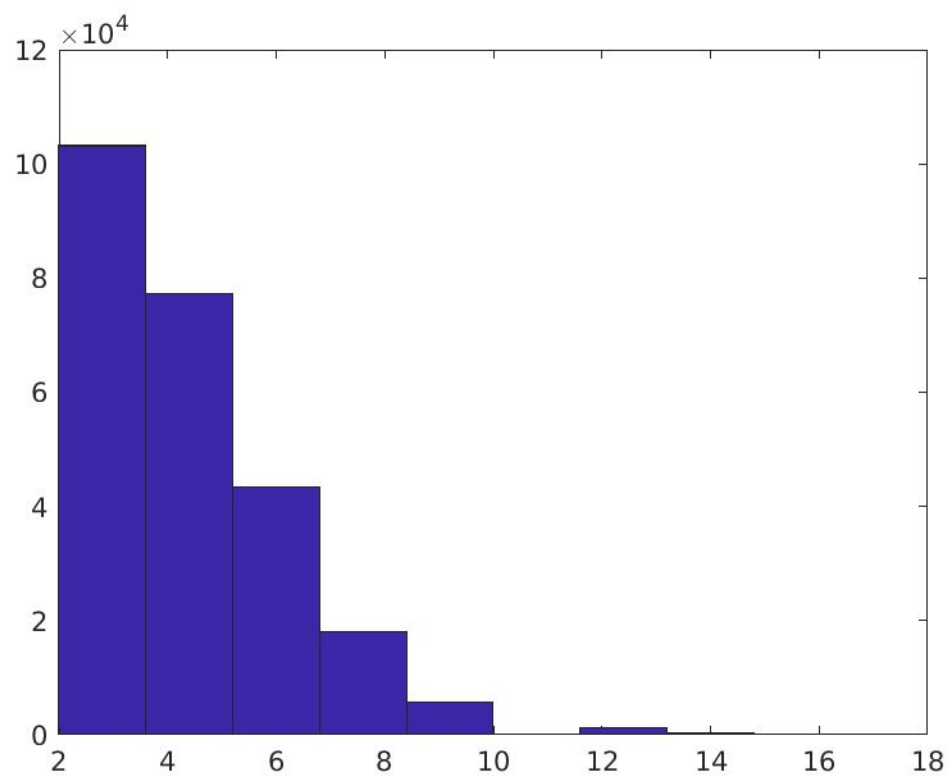
3.e) The execution time was reported as 0 seconds, possibly because the actual execution time was too small to represent using a normal matlab variable.

```
format short
```

```
ans =
```

```
0
```

3.f)



4.a)

A =

2	1	3
2	6	8
6	8	18

The minimum value in the first row of A is 1
The minimum value in the second row of A is 2
The minimum value in the third row of A is 6
The maximum value in the first column of A is 6
The maximum value in the second column of A is 8
The maximum value in the first column of A is 18
The largest value in A is 18

4.b) The resultant solution is: $x=0.3$, $y=0.4$ and $z=0$

$$F = \text{inv}(C) * D$$

C =

```
2    1    3
2    6    8
6    8   18
```

D =

```
1
3
5
```

F =

```
0.3000
0.4000
0
```

4.c) $x1 = [0.5 \ 0 \ 1.5]$

L1 Norm: $0.5 + 0 + 1.5 = 2$

L2 Norm: $(0.5^2 + 0^2 + 1.5^2)^{0.5} = 1.58$

Matlab's norm function gives the answer 1.581139

$x2 = [1 \ 1 \ 0]$

L1 Norm: $1 + 1 + 0 = 2$

L2 Norm: $(1^2 + 1^2 + 0^2)^{0.5} = 1.41$

Matlab's norm function gives the answer 1.414214

The L1 norm of x1 (the sum of all elements in x1) is: 2

The L2 norm of x1 (the square root of the sum of the squares of all elements in x1) is: 1.581139

Matlab's norm of x1: 1.581139

The L1 norm of x2 (the sum of all elements in x2) is: 2

The L2 norm of x2 (the square root of the sum of the squares of all elements in x2) is: 1.414214

Matlab's norm of x2: 1.414214

5)

A =

1	2	3
4	5	6
7	8	9

B =

0.1667	0.3333	0.5000
0.2667	0.3333	0.4000
0.2917	0.3333	0.3750

A =

10	17	21
4	2	99
12	144	3
4	10	13
1	2	44

B =

0.2083	0.3542	0.4375
0.0381	0.0190	0.9429
0.0755	0.9057	0.0189
0.1481	0.3704	0.4815
0.0213	0.0426	0.9362