

## Homework 2

### Instructions

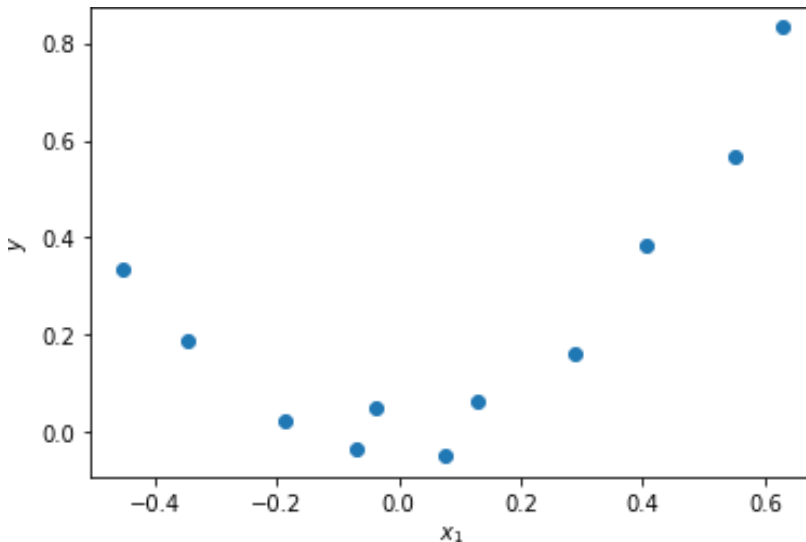
This homework contains **5** concepts and **8** programming questions.

- Concept questions (this file): Using MS word or a similar text editor, write down the problem number and your answer for each problem. Use a separate page for each new problem. Combine pages into a PDF file.
- Jupyter notebook files: Export/print the Jupyter notebook as a PDF file including the code you implemented and the outputs of the program. **Make sure all plots and outputs are visible in the PDF.**
- Combine all concept and Jupyter notebook PDF files into a single PDF named andrewID\_hw2.pdf and submit it to Gradescope.
- In Gradescope, use “Assign Questions and Pages” to pair up questions to the corresponding pages in your PDF file. **We cannot grade your assignment if you do not do this.** 10 points will be deducted if you fail to do this and ask the teaching crew for a regrade after fixing the issue.
- Refer to the syllabus and the assessment schedule for the deadlines and late submission policies.

Here is a breakdown of the points for the concept and Jupyter notebook programming questions:

Name	Points
Concept 1	4
Concept 2	2
Concept 3	2
Concept 4	2
Concept 5	2
M2_L1_p1	6
M2_L1_p2	6
M2_L1_p3	6
M2_L2_p1	6
M2_L2_p2	6
M2_L2_p3	6
M2_HW1	36
M2_HW2	36
<b>Total</b>	120
Fly Solo	20

### Problem 1



**a) Consider fitting a second order LLS regression model to the following data:**

What is the size (height and width) of the design matrix  $X$ ?

What is the size (height and width) of the parameter vector  $w$ ?

What is the size of the product  $X'X$ ?

**b) Now consider a dataset with 1000 points, for the same second order model:**

What is the size (height and width) of the design matrix  $X$ ?

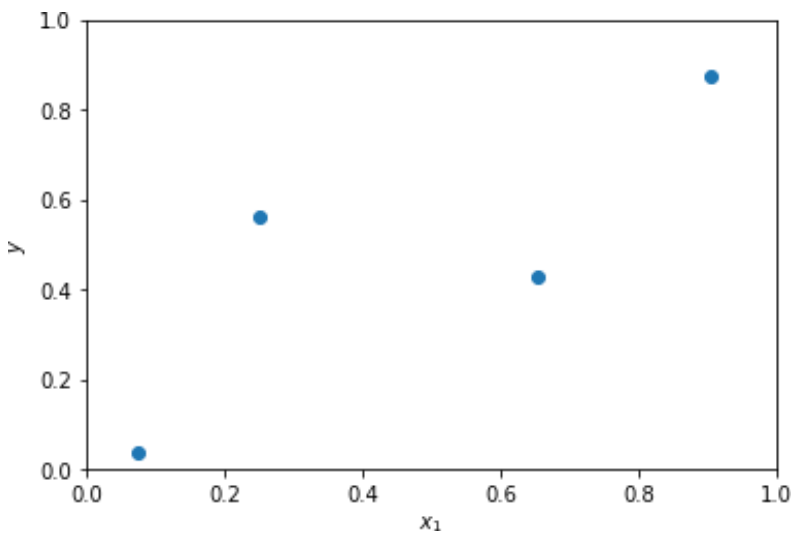
What is the size (height and width) of the parameter vector  $w$ ?

What is the size of the product  $X'X$ ?

**c) Which of the following statements is true:**

1. The model complexity is determined by the number of data points used for training
2. The model complexity is determined by the number of model parameters

## Problem 2



Consider the four points shown in the following figure:

If we use polynomial regression, what is the smallest order polynomial to perfectly interpolate these points? (multiple choice)

1. 1st order -  $w_1x_1 + w_2$
2. 2nd order -  $w_1x_1^2 + w_2x_1 + w_3$
3. 3rd order -  $w_1x_1^3 + w_2x_1^2 + w_3x_1 + w_4$
4. 4th order -  $w_1x_1^4 + w_2x_1^3 + w_3x_1^2 + w_4x_1 + w_5$

### Problem 3

Imagine you are hiking to the top of a mountain, but you get lost off trail, and it gets dark. All you have is a dim flashlight, but it is enough for you to see your feet and which way the mountain slopes. Since you took AIML for engineers you know you can use gradient descent to find the extrema of the mountain. Which of the following formulations will lead you to the top of the mountain?

1. 
$$x_{new} = x_{old} - \eta \cdot \frac{\partial \text{obj}}{\partial x}$$

2. 
$$x_{new} = x_{old} + \eta \cdot \frac{\partial \text{obj}}{\partial x}$$

#### Problem 4

You are implementing batch gradient descent but notice that your solution is taking a long time to minimize the objective function due to oscillations. Which of the following approaches are likely to help resolve this issue? (Multiple choice, select all that are true)

1. Increase the batch size
2. Switch to stochastic gradient descent
3. Reduce the learning rate

### Problem 5

The function for elastic net regularization in sklearn has a `L1_ratio` parameter which describes the weight of the L1 and L2 penalty terms relative to each other. The  $\alpha$  parameter describes the strength of the regularization terms. The formulation of the objective function follows:

$$\|Xw - y\|_2^2 + \alpha \cdot L1_{ratio} \cdot \|w\|_1 + \alpha \cdot (1 - L1_{ratio}) \cdot \|w\|_2^2$$

If the `L1_ratio` term is 0, what is the formulation equivalent to?

1. Standard LLS
2. LLS with L1 regularization
3. LLS with L2 regularization
4. LLS with L1 and L2 regularization