Homework 6: Gradient Descent + Linear Regression

Complete the following exercises. Remember to explain your answers.

1. Consider the (very small) dataset:

x_i	0	10	20	30
y_i	10	19	25	34

Work through three iterations of gradient descent, using initial values $\beta_0 = 0$ and $\beta_1 = 1$ and a learning rate $l = 0.1$, showing your work for setting up each computation. Note that the MSE cost function for this dataset is

$$C(\beta_0, \beta_1) = \beta_0^2 + 30\beta_0\beta_1 - 44\beta_0 + 350\beta_1^2 - 855\beta_1 + 560.5,$$

and the gradient is

$$\nabla C(\beta_0, \beta_1) = (2\beta_0 + 30\beta_1 - 44, 30\beta_0 + 700\beta_1 - 855).$$

It may be helpful to use the code from the workbook during lecture to check your work.

2. Suppose that you are using gradient descent to train a linear regression model on a dataset. Tracking your progress through gradient descent, you get the following results for the first few iterations:

Iteration	Gradient	MSE
1	(10, 15.2)	20.5
2	(-30.2, -46.6)	115.2
3	(96.3, 159.2)	1567.3
4	(-359.1, -589)	5848.9
5	(1094.3, 3823)	238942.2

   (a) Explain what is happening, and how it affects the training of your model.

   (b) Is this a problem? If yes, what can you do to (possibly) fix this issue?

3. Suppose that you are using gradient descent to train a linear regression model on a dataset. Tracking your progress through gradient descent, you get the following results for the first few iterations:

Iteration	Gradient	MSE
1	(10, 15.2)	20.5
2	(-30.2, -46.6)	115.2
3	(10, 15.2)	20.5
4	(-30.2, -46.6)	115.2
5	(10, 15.2)	20.5
6	(-30.2, -46.6)	115.2

   (a) Explain what is happening, and how it affects the training of your model.

   (b) Is this a problem? If yes, what can you do to (possibly) fix this issue?

4. Suppose that you are using gradient descent to train a linear regression model on a dataset. Tracking your progress through gradient descent, you get the following results for the first few iterations:

Iteration	Gradient	MSE
1	(1, 1.5)	3984.5
2	(.5, 1.2)	3843.2
3	(.3, .98)	3722.4
4	(.12, .76)	3699.2
5	(.05, .56)	3681.8
6	(.015, .21)	3674

   (a) Explain what is happening, and how it affects the training of your model.

   (b) Is this a problem? If yes, what can you do to (possibly) fix this issue?