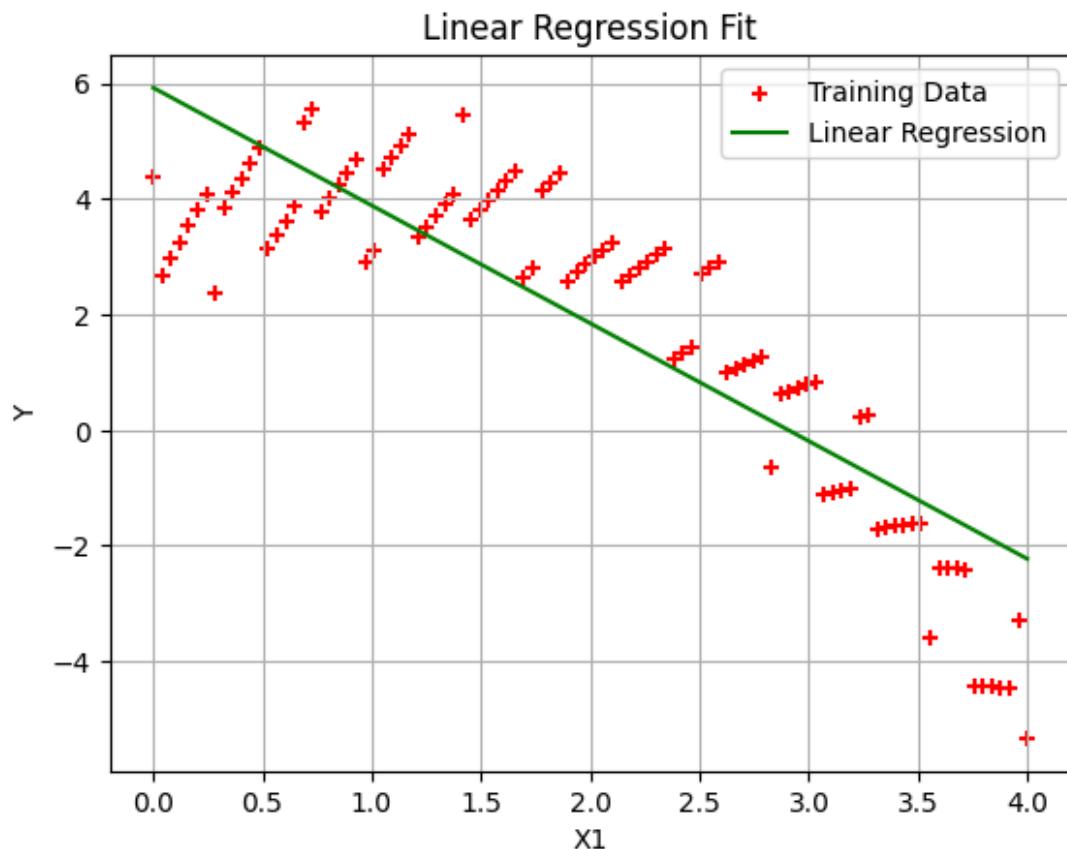


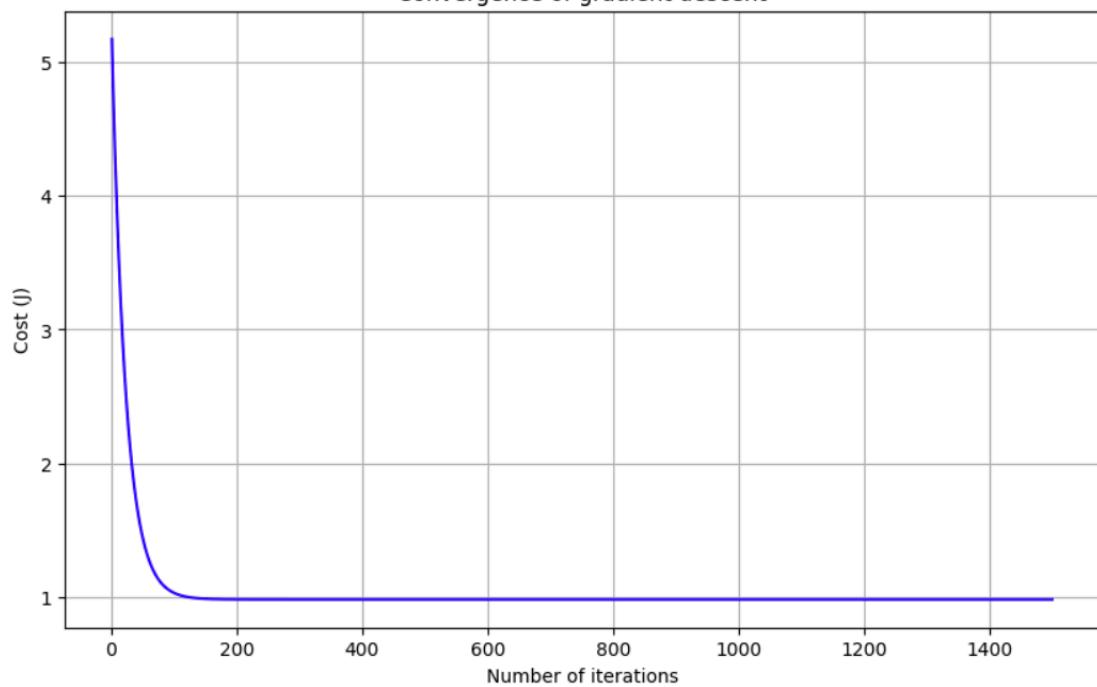
Homework 1 Report
Name: Viprav Lipare
Student ID: 801288922
Homework Number: 1
Github Repository:

Problem 1

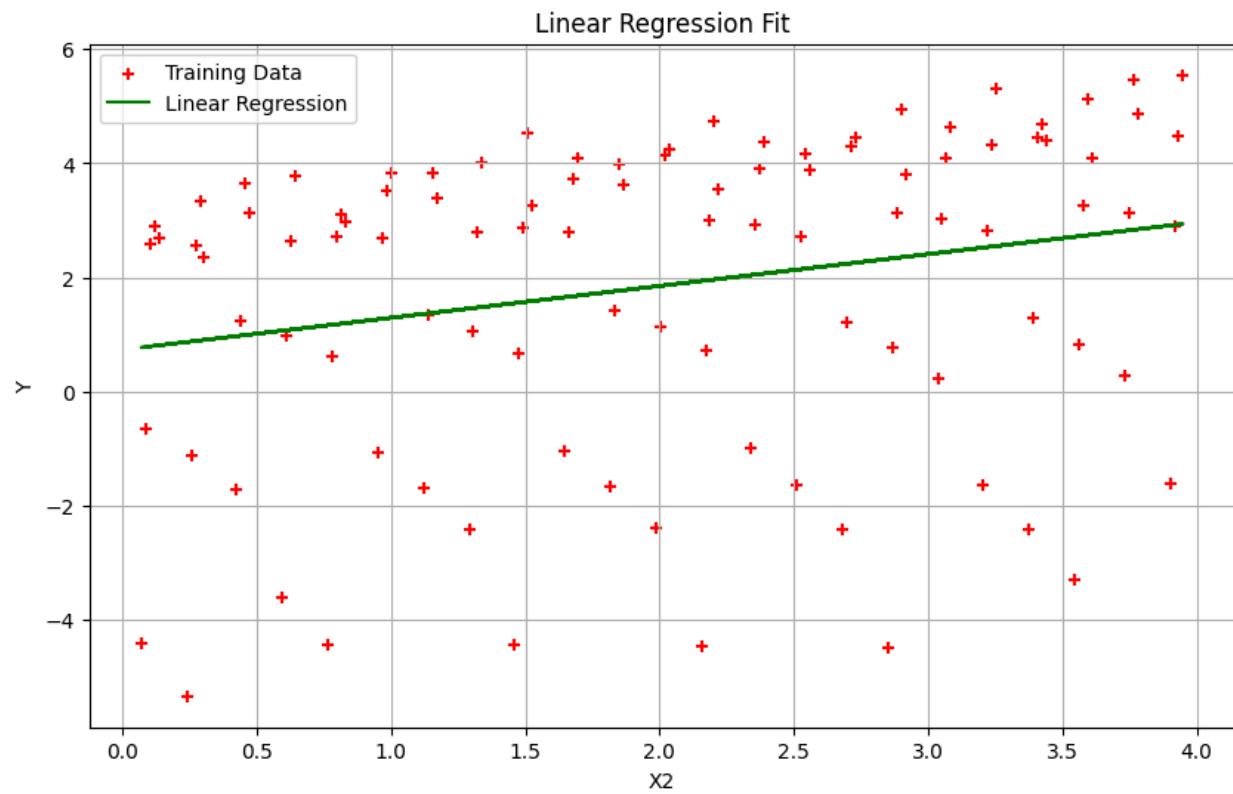
X1 - Plots and Linear Model



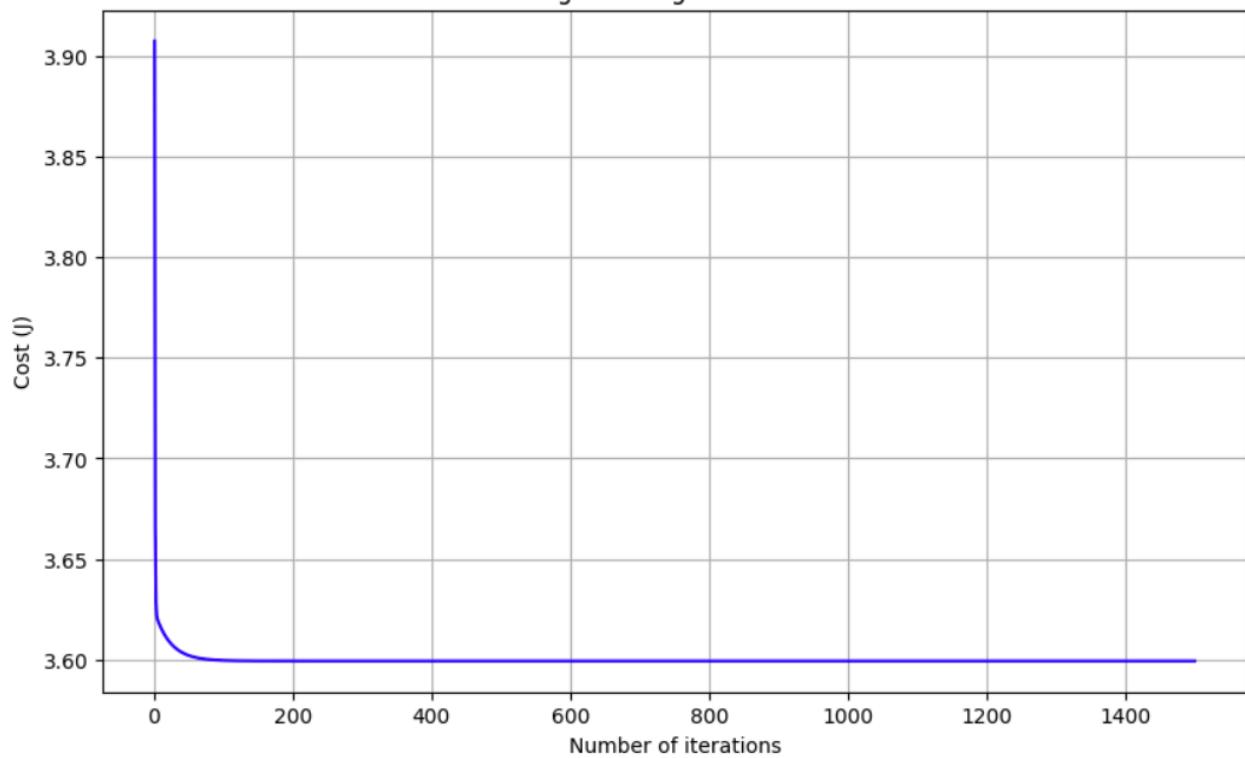
Convergence of gradient descent



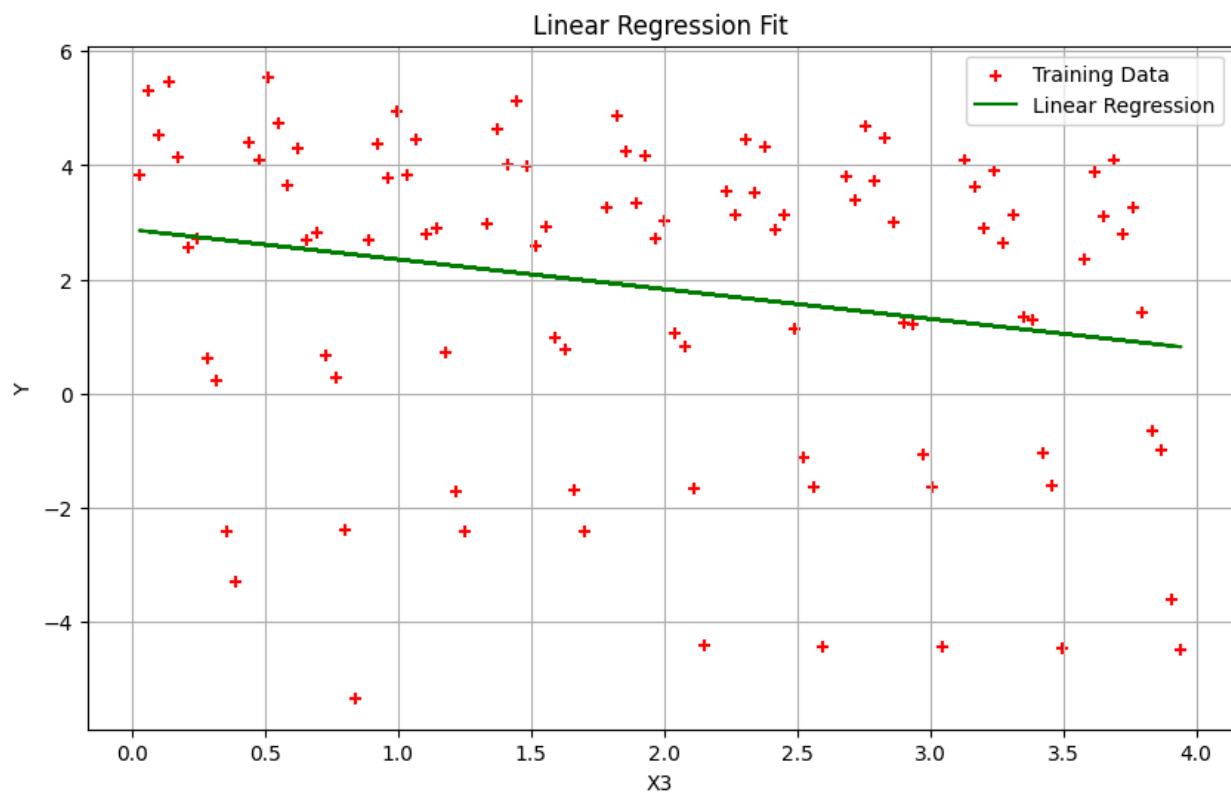
X2 - Plots and Linear Model

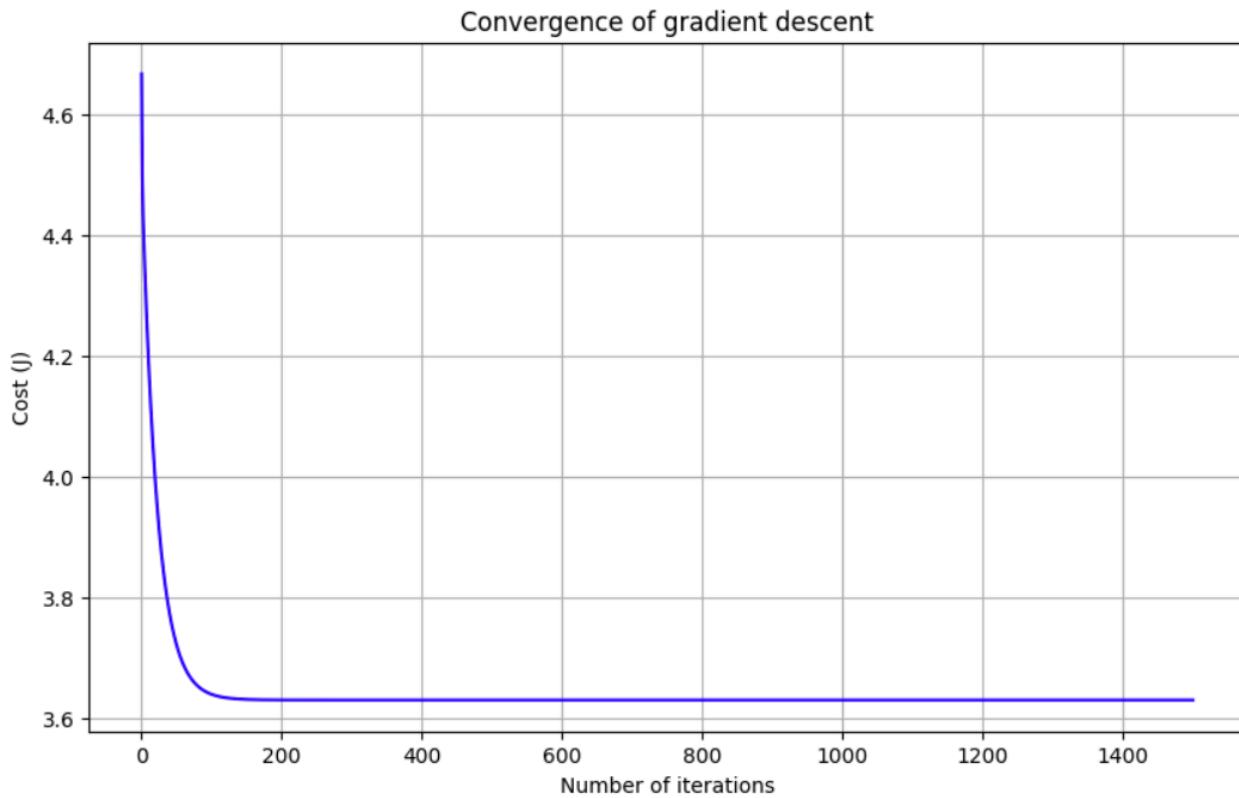


Convergence of gradient descent



X3 - Plots and Linear Model



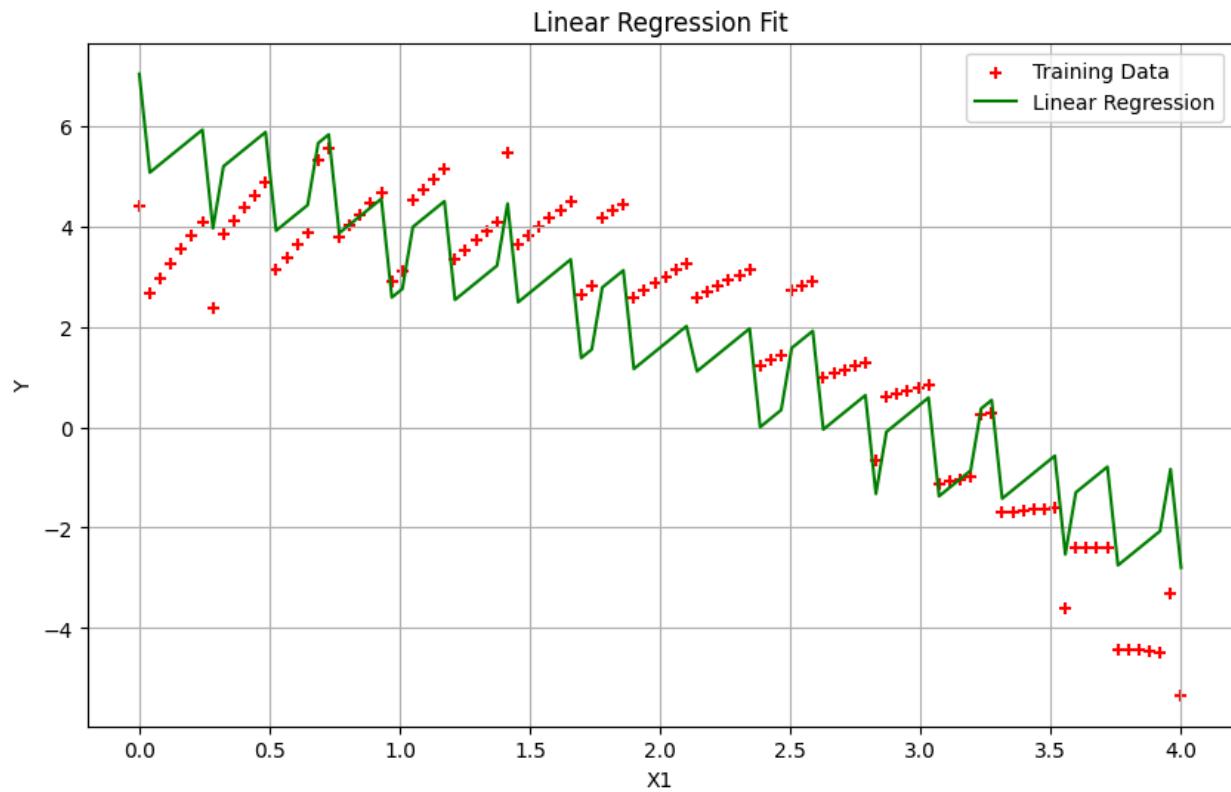


Problem 1 Questions:

1. Linear Models for each explanatory variable
 - a. X1 : Linear model: $y = 5.927948918061593 + -2.038336633650798 \cdot x_1$
 - b. X2 : Linear model: $y = 0.7360604300947937 + 0.557607610332603 \cdot x_2$
 - c. X3 : Linear model: $y = 2.871422103988201 + -0.520482884300103 \cdot x_3$
2. Each of the explanatory variables X1, X2, X3 are plotted above in relation to Y. Seeing the plots for each of the three explanatory variables only X1 goes in an order. X2 and X3 have scatterplots that have the plot points spread around the full plots so the linear regression line only fits for X1.
3. This is also supported by the loss graphs because for both X2 and X3 the minimum cost for them is around 3.5, but for X1 the minimum cost is at 1.
4. I ran the code with 0.1 learning rate and a 0.01 learning rate, the only difference was that with the higher learning rate the minimum loss was reached at a much faster rate.

Problem 2

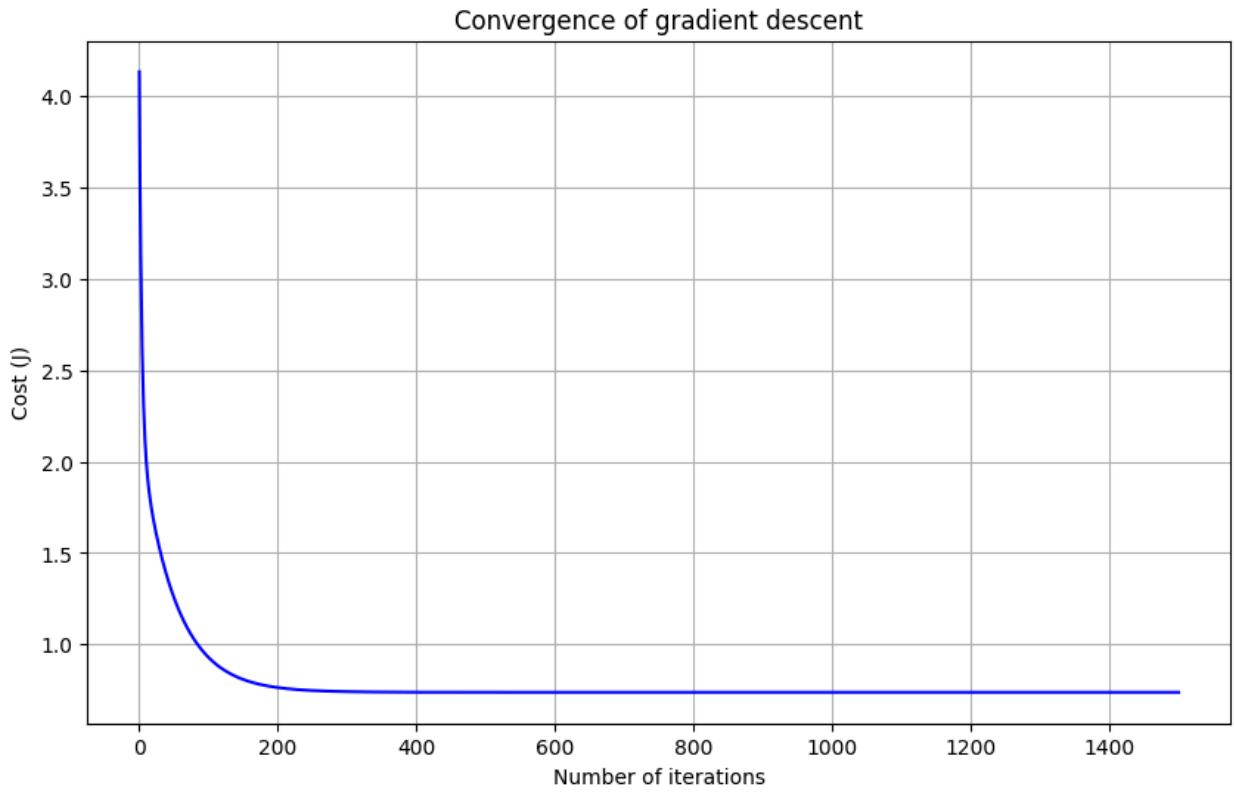
Plots for X1, X2, X3 and Y Linear Relation



Problem 2 Questions:

1. Linear Models for explanatory variables

a. `Linear model: y = 5.31416562749957 + -2.003719050517275*X1 + 0.5325635914250081*X2 + -0.2656016390618418*X3`



2.

3. The difference in the learning rates were more visible in the problem 2 test runs. With an advanced learning rate the lines followed a pattern that followed the plot points more vividly. With a lower learning rate the regression line followed a more spiked pattern that was oscillating through the points. The same thing can be said about the loss plot as the first problem, the loss plot shows a much faster decrease and a good stabilization when the learning rate is higher.

4. `Y values for new X values = [3.57740853 0.24432097 0.10253402]`