Labwork 2: Linear Regression

April 21, 2024

1 Introduction

Linear regression is a form of regression analysis that assumes a linear relationship between the dependent variable and the independent variable. The goal is to find the best-fitting straight line that describes the relationship between the variables.

2 Implementation

First, we implement a load_data method to load the .csv file. Then, we implement the compute_cost method, which has the input parameter x, y, w1 and w0, to calculate the mean squared error. Next, we implement the gradient_descent method to minimize the cost function through iterative updates to the slope and the intercept, with the tolerance value is $1 * 10^{-6}$. Finally, for the linear regression, we set the learning rate is 0.0001, iterations is 10000.

3 Evaluation

Test with the housing price data, ater 10000 iterations, we find that the result equation is 15.32x + 0.13.

4 Conclusion

In this labwork, we implemented the load_data, compute_cost and gradient_descent methods.
An interesting finding is that after some iterations, the result does not improve. We can terminate the process early by increasing the tolerance value, from $1 * 10^{-6}$ to $1 * 10^{-3}$