CS 3891/5891

Spring 2020

Assignment 1

Linear Regression using Gradient Descent

Part 1: Linear regression with one variable

Suppose you are the CEO of a restaurant franchise and are considering different cities for opening a new outlet. The chain already has food trucks in various cities and you have data for profits and populations from the cities. You'd like to figure out what the expected profit of a new food truck might be given only the population of the city that it would be placed in. Predict the profits for a food track using simple regression based on the available data.

The file `food_truck_data.txt` contains the dataset for the problem. The data consists of two columns; the first column is the population of a city and the second column is the profit of a food truck in that city. A negative value for profit indicates a loss.

- Fit a linear model to the data
    - Compute the parameters of the model using gradient descent
    - Compute and plot the cost function as a function of the gradient descent iterations
- Plot the data with the linear model
- Submit
    a. Your software and a report describing your solution or
    b. A Jupyter notebook that includes the code, results, and explanations.

Part 2: Linear regression with multiple variables

Suppose you are selling your house and you want to know what a good market price would be. You can use available data on recent housing prices sold and learn a linear model.

The file `housing_price_data.txt` contains the dataset for the problem. The data consists of three columns; the first column is the size of the house (in square feet), the second column is the number of bedrooms, and the third column is the price of the house.

- Fit a linear model to the data
    - Use feature scaling
    - Compute the parameters using gradient descent
    - Compute and plot the cost function as a function of the gradient descent iterations
- Assume features of your house and predict a good market price
- Submit
    - Your software and a report describing your solution or
    - A Jupyter notebook that includes the code, results, and explanations.