Linear Regression Using Normal Equation

Farnaz Golnam
MCS 19576

Introduction

Regression models are used to model a relationship between the dependant and independent variables. When data shows a **straight trend** this relationship is **linear** otherwise the relationship is **non-linear**.

In this project we are going to use python to fit a linear regression model into our data using normal equation concept

Linear Regression Equation(y) = a + bx

Slope(b) = (NΣXY - (ΣX)(ΣY)) / (NΣX^2 - (ΣX)^2)

Intercept(a) = (ΣY - b(ΣX)) / N

Data

The data we are going to use is in CSV format with below attributes::
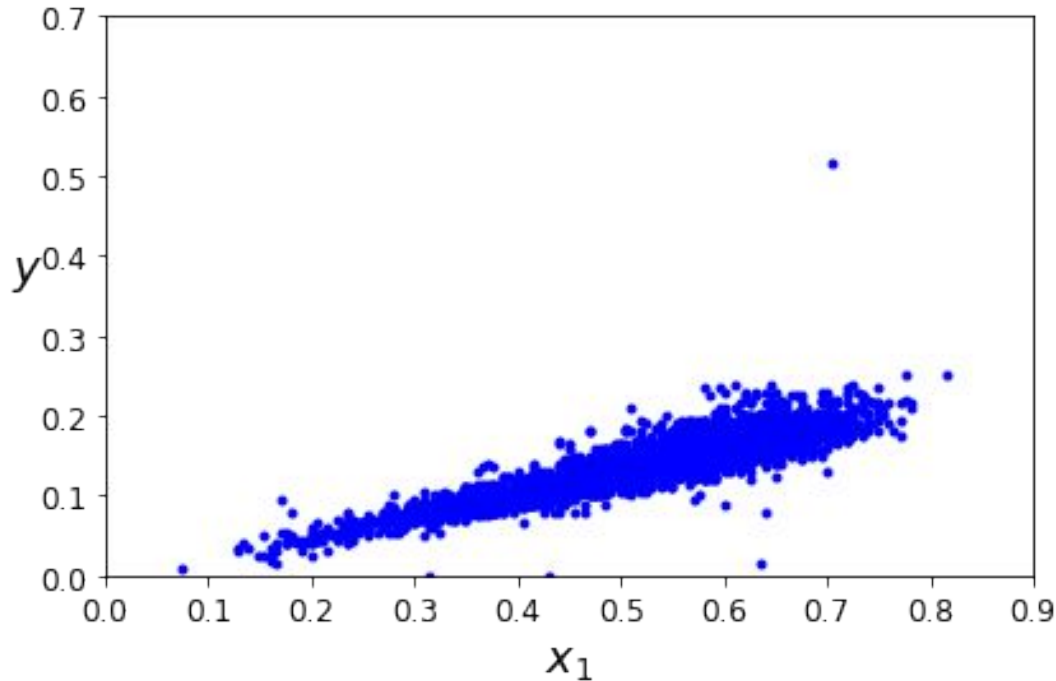
["Length", "Diameter", "Height", "Whole weight", "Shucked weight",

   "Viscera weight", "Shell weight", "Age"]

Y = Height

X = Length

Checking the linearity

To make sure that the relation between our X and Y data follows a linear trend,

we first plot the data

To implement the concept of normal equation in python, we use sklearn library

```python
from sklearn.linear_model import LinearRegression

#to find the intercept and coefficient of the linear model:
lin_reg = LinearRegression()
lin_reg.fit(X, y)
lin_reg.intercept_, lin_reg.coef_

# to predict the y for a new data(x):
lin_reg.predict(X_new)
```

The linear model created as shown below:
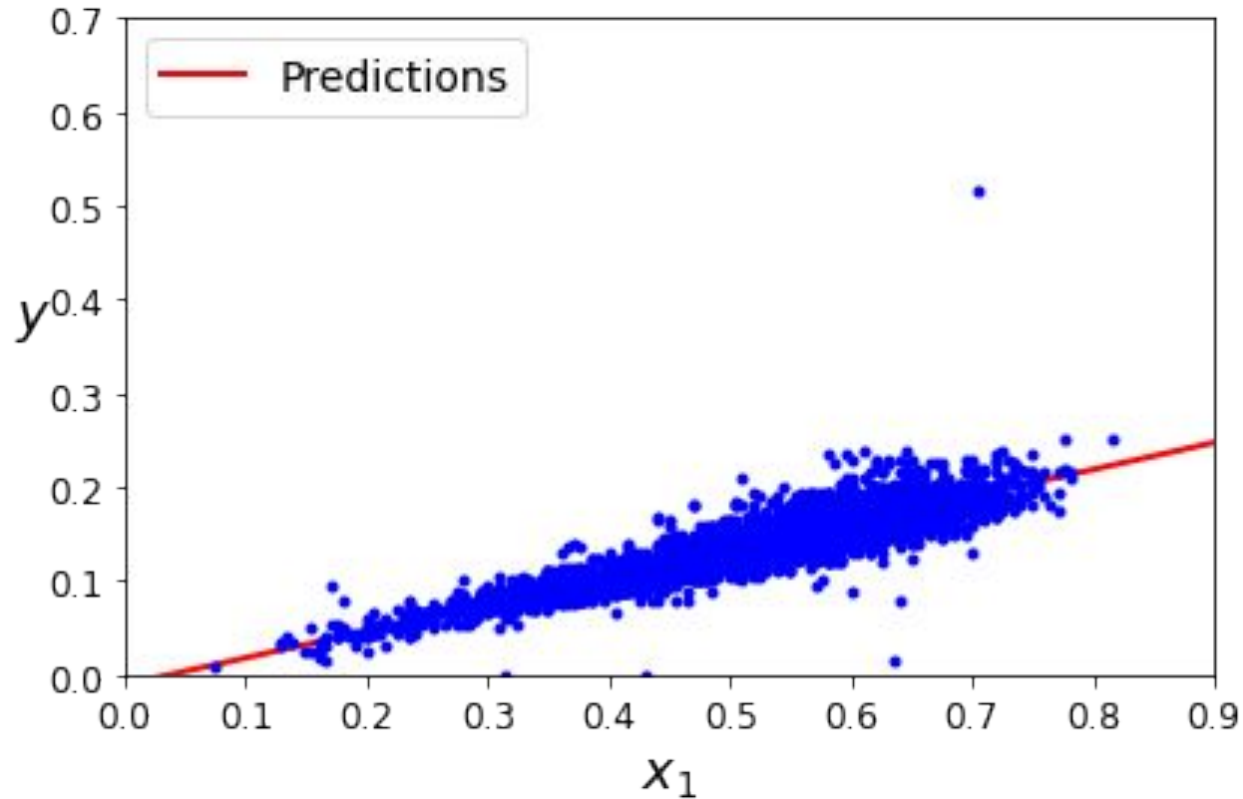
The complete version of different steps of training a linear model in Python is available in the link below

[Training linear regression using normal equation](#)