

SOC ASSIGNMENT-1

GROUP:

Sai Charan Teja : IIT2016039
Anil Gupta Somisetty : IIT2016131
Nitin Raj Singh : IIT2016132
Utpal : IIT2016133
Kiran : ITM2016009

PREDICTION:

We need to predict HOUSE PRICE.

ANALYSIS:

1.As the given data consists of features with classification problem i.e data with “yes” or “no” we remove them as we cannot use them in linear regression and we continue with the remaining data.

2.As we have irregular data meaning one feature dominates the other we NORMALISE the data with $\text{data} = (\text{data} - \text{data.mean()}) / \text{data.std()}$.

3.We will be having 6 features including bias and weights would be
6. $W = [W_0, W_1, W_2, W_3, W_4, W_5]$

CONCLUSION:

NORMAL EQUATION:

Calculated the weights using the equation $W = ((X.T * X)^{-1}) * (X.T * Y)$
Y=ORIGINAL HOUSING PRICE
X=FEATURE MATRIX
W=WEIGHTS

RESULT:

$W = [[-7.58775430e-17]$
[3.84337540e-01]
[6.59575210e-02]
[3.03672825e-01]

[2.56189746e-01]
[1.73274748e-01]]

COST for above $W=0.21893272895870813$

GRADIENT DESCENT:

Initial W would be $[0,0,0,0,0,0]$ and we would be changing W after every epoch and decreasing the cost.

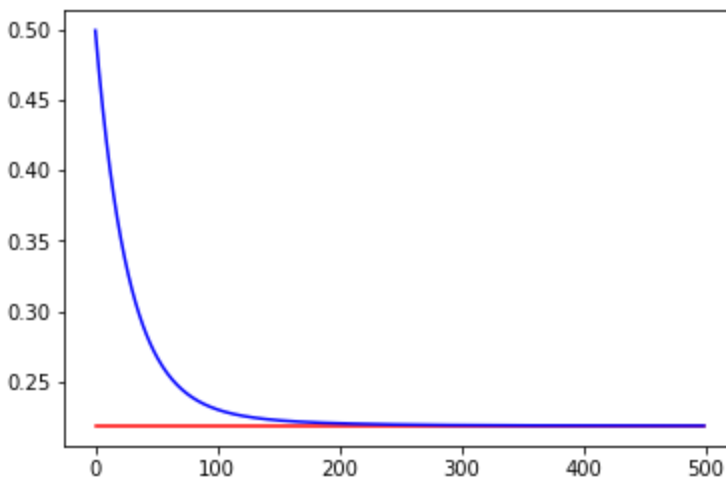
RESULTS:

For different values of learning rate the output graph convergence would be different.

Blue curve=Gradient Descent, Red curve=Normal Equation

Learning rate=0.01

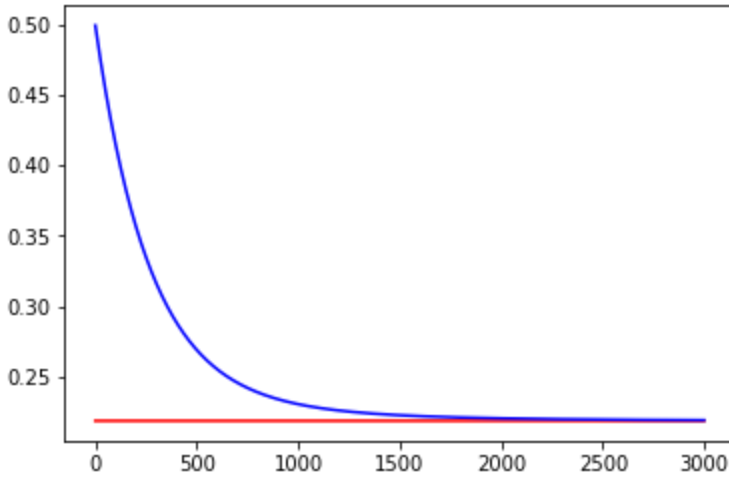
Cost vs epoch graph:



Final cost after 500 epochs:0.21897188825763936

Learning rate=0.001

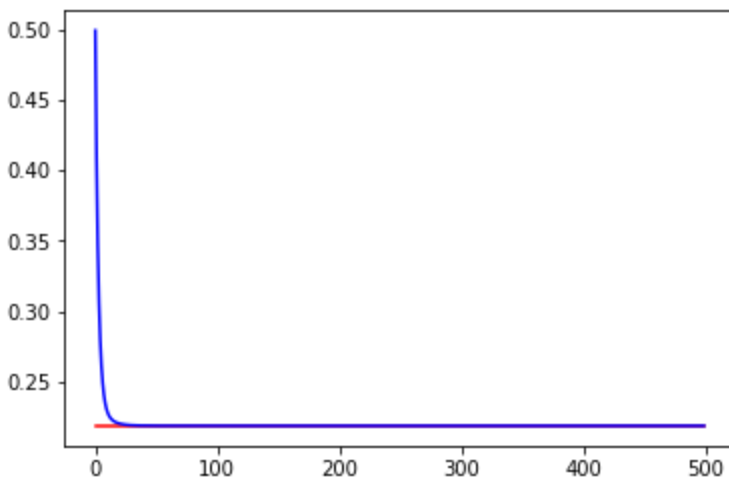
Cost vs epoch graph:



Final cost after 3000 epochs:0.2193917616073436

Learning rate=0.1

Cost vs epoch graph:



Final cost after 500 epochs:0.21893272895870836

COMPARISION:

NORMAL EQUATION:

Final cost is 0.21893272895870813

Gradient Descent:

Learning rate=0.1 took 30 epochs to converge the cost.

Learning rate=0.01 took 300 epochs to converge the cost.

Learning rate=0.001 took 3000 epochs to converge the cost.

We choose the convergence learning rate 0.1 for our algorithm with more number of epochs after observing the final cost of convergence.

If we observe the of 3 learning rates 0.1 fastly converges 0.01 moderately converges and 0.001 converges very lately.