House Size	Price
2104	399900
1600	329900
2400	369000
1416	232000

A. Assuming β_0_=80000and β_1_=_100. Calculate the Cost function $J(\beta_1, \beta_0)$_.

$h(x) = 100x + 80\,000$

$h_1(x) = 100 \times 2104 + 80\,000 = 290400$

$h_2(x) = 100 \times 1600 + 80\,000 = 24\,0000$

$h_3(x) = 100 \times 2400 + 80000 = 32\,0000$

$h_4(x) = 100 \times 1416 + 80000 = 221600.$

$y_1 = 399900$

$y_2 = 329900$

$y_3 = 369\,000$

$y_4 = 232000$

$h_1(x) - y_1 = 290400 - 399900 = -109500$

$h_2(x) - y_2 = -329900 + 240000 = -89900$

$h_3(x) - y_3 = +320000 + (-369000) = -49000$

$h_4(x) - y_4 = 221600 - 232000 = -10400$

$J(\beta_1, \beta_0) = 2822677500$

$$J(\beta_1, \beta_0) = \frac{1}{2m} \sum_{i=1}^{n} \left(\left(h(x^{(i)}) - y^{(i)}\right)^2\right)$$

B. Assuming a learning rate of α=0.1 calculate the new value for β_1 and β_0 after applying gradient descent.

$\alpha = 0.1$ Sensitivity ng rate.

$$\frac{\partial}{\partial \beta_0} = \frac{1}{m} \sum_{i=1}^{m} \left(h(x^i) - y^i \right)$$

$$\cancel{87200} = -64700.$$

$$\frac{\partial}{\partial \beta_1} = \frac{1}{m} \sum_{i=1}^{m} \left(\left(h(x^i) - y^i \right) x^i \right)$$

$$= -5065544\,00 / 4 = -1266\,38600$$

$$\beta_0 = 80000 - 0.1 \times (-64700)$$

$$\boxed{\beta_0 = 86470}$$

$$\beta_1 = 100 - 0.1 \times (-126638600)$$

$$\boxed{\beta_1 = 12663960}$$

.69 in

C. Given the correct values for β_1 and β_0 once gradient descent finishes should be 132.6 and 83458 respectively; review your answer in 2 above. Are your new values for β_1 and β_0 seem reasonable? Briefly discuss your result and why it may be incorrect. Include a brief explanation of what you should have done prior to doing step 1 and 2 and how you might choose your learning rate (α) for this problem in the future.

Answer:

No, the new values of the $\beta 1$ and $\beta 0$ are not reasonable since as compared to what they should be, they are significantly higher, the case of overshooting the minimum value of the error function, this would never reach the given values, i.w., there is no value of the learning rate that would lead to values converging to the optimal values.

Since the scales for the House size and Price are different, this hampers the learning ability of the gradient descent algorithm, so to perform the better version of learning process, the values could be scaled using MinMax scaling before performing step 1 and step 2.

Before step 2 it is important that the new values do not overshoot the expected values, thus the smaller value of the learning rate should be selected before performing of this step.