Deep Salunkhe
21102A0014

Assignment 5

Task 1

A liner model : $\boxed{y = mx + c}$

$$m = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = 2.$$

$$c = \frac{\sum y - m \sum x}{n} = 1$$

Interpretation :

- slope ($m = 2$) : For each additional hr , the number of units produced increases by 2.

- intercept ($c = 1$) : Jeut ast ohr we have 1 output

2] Task 2

The objective is to reduce MSE cn

$$\boxed{Loss = \frac{1}{2m} \sum (y_i - \hat{y})^2.}$$

For that we we update w and b (c)

$$\boxed{\begin{aligned} w &= w - \eta \frac{\partial L}{\partial w} \\ b &= b - \eta \frac{\partial L}{\partial b} \end{aligned}}$$

Stochastic Gradient Descent

$$\boxed{y = wx + b}$$

x	y
1	3
2	5
3	7
4	9
5	11

initial: $w = 0.5$, $b = 0.01$

$\eta = 0.01$

For first tuple.

$x = 1 \quad y = 3$

$$\hat{y} = 0.5 \times 1 + 0.1 = 0.6$$

Compute the gradient.

$$\frac{\partial J}{\partial w} = \cancel{-1}\times(y_i - \hat{y}_i) \times x$$

$$= (3 - 0.6) \times 1 = 2.4$$

$$\frac{\partial J}{\partial b} = \cancel{-1}(y_i - \hat{y}) = (3 - 0.6) = 2.4$$

update w

$$w = w_0 - \eta \frac{\partial J}{\partial w}$$

$$= 0.5 - 0.01 \times 2.4$$

$$= \cancel{0.26}\ 0.476$$

$$b = b_0 - \eta \frac{\partial J}{\partial b}$$

$$= 0.1 - 0.01 \times 2.4$$

$$= \cancel{-0.14}\ 0.\cancel{0976}\ 0.076$$

stochastic GD

$w = 0.5 \quad b = 0.1$

$\eta = 0.01$

$\hat{y} = [\ 0.6 \quad 1.1 \quad 1.6 \quad 2.1 \quad 2.6\ ]$

$$MSE = \frac{1}{2m} \times \left((3-0.6)^2 + (5-1.1)^2 + (7-1.6)^2 + (7-2.1)^2 + (6.6-1.7)^2 \right)$$

$$= \boxed{18.83}$$

Tuple 1

$$\frac{\partial J}{\partial w} = -1\,(y_i - \hat{y})\,x_i$$
$$= -1\,(3 - 0.6)\,1$$
$$= -2.4$$

$$\frac{\partial J}{\partial b} = -1\,(y_i - \hat{y})$$
$$= -1 \times (3 - 0.6)$$
$$= -2.4$$

updates

$$w = w_0 - \eta \frac{\partial J}{\partial w}$$
$$= 0.5 - 0.01 \times (-2.4)$$
$$= 0.524$$

$$b = 0.1 - 0.01\,(-2.4)$$
$$= 0.124$$

Tuple . 2 $\quad w = 0.524 \quad b = 0.124 \quad \eta = 0.01$

$\frac{\partial J}{\partial w} = -1 \cancel{z}$

$\hat{y}_2 = 0.524 \times 2 + 0.124$

$= 1.172$

$\frac{\partial J}{\partial w} = -1 \times (5 - 1.172) \times 2$

$= -7.652$

$\frac{\partial J}{\partial b} = -1 \times (5 - 1.172) = -3.824$

Update

$w = 0.524 - 0.01 \times (-7.652)$

$= 0.60052$

$b = 0.124 - 0.01(-3.824)$

$= 0.16224$

Table 3 $\quad w = 0.60052 \quad b = 0.16224$

$\hat{y}_3 = 0.60052 \times 3 + 0.16224 = 1.9638$

$\frac{\partial J}{\partial w} = -1 \times (7 - 1.96380) \times 3 \qquad \frac{\partial J}{\partial b} = -5.0362$

$= -15.1086$

updates

$w = 0.60052 - 0.01 \times (-15.1056)$

$= 0.7515$

$b = 0.16224 - 0.01 \times (-5.0302)$

$= 0.21254$

Tuple 4. $w = 0.7515$ $b = 0.21254$

$\hat{y} = 0.7515 \times 4 + 0.21254 = 3.21854$

$\frac{\partial J}{\partial w} = -1 \times (9 - 3.21854) \times 4 = -23.12583$

$\frac{\partial J}{\partial b} = -1 \times (9 - 3.21853) = -5.78146$

update

$w = 0.7515 - 0.01 \times (-23.12584)$

$= 0.9827$

$b = 0.21254 - 0.01 \times (-5.78146)$

$= 0.2703$

Tuple 5

$\hat{y} = 0.9827 \times 5 + 0.2703 = 5.1838$

$\frac{\partial J}{\partial w} = -1\,(11 - 5.1838) \times 5 = -29.081$

$\frac{\partial J}{\partial b} = -1\,(\text{——}\ 11\text{—}) = -5.8162$

$w = 0.9827 - 0.01 \times (-29.081)$

$= 1.27351$

$b = 0.2703 - 0.01 \times (-5.8162)$

$= 0.3284$

find weights after 1 iter

$w = 1.2737 \qquad b = 0.3287.$

$\therefore \hat{y} = [1.602 \quad 2.873 \quad 4.146 \quad 5.419 \quad 6.692]$

$$MSE = \frac{1}{2M} \sum (y_i - \hat{y})^2$$

$$= \frac{1}{2\times5} [1.954 + 4.524 + 8.145 + 12.880 + 18.504]$$

$$= \frac{1}{10} [46.005]$$

$$= 4.6005$$

#BGD

a (cong to ...)

α

$\boxed{y = 2x + 1}$

$\boxed{w = 0.5 \quad b = 0.1}$

(lets asm)

$\therefore \hat{y} = [\ 0.6 \quad 1.1 \quad 1.6 \quad 2.1 \quad 2.6\ ]$

$$MSE = \frac{1}{2m} \Sigma (y' - \hat{y})^2$$

$$= 16.83$$

$$w_{new} = w_{old} - \eta \times \frac{\partial L}{\partial w}$$

$$= 0.692$$

$$b_{new} = b_{old} - \eta \times \frac{\partial L}{\partial b}$$

$$= 0.154$$

$$MSE \Rightarrow \frac{1}{2m} \Sigma [y_i - (w_{new} \times x + b_{new})]^2$$

$$= 13.085$$

Final inference

Linear Regression ⇒ + ⇒ Simple and fast

− ⇒ Assumes linearity, sensitive to to outliers

BGD ⇒ + ⇒ stable convergence, good for small Datasets

− ⇒ slow for larger data set.

SGD ⇒ + ⇒ Effective for real time updates

− ⇒ Noisy updates.