Stochastic G.D.

DATE

for ~~vector~~ Single record

$$\hat{y} = w_1x_1 + w_2x_2 + w_3x_3 + b$$

w_1, w_2, w_3, b

$x_1 \rightarrow \hat{y}_1 \longrightarrow L_1 \rightarrow (y_1 - \hat{y}_1)^2$

x_2

x_3

$$\frac{\partial L_1}{\partial w_1} = \frac{\partial L_1}{\partial \hat{y}_1} \frac{\partial \hat{y}_1}{\partial w_1} = -2(y_1 - \hat{y}_1) \cdot x_1$$

$$\frac{\partial L_1}{\partial w_2} = \frac{\partial L_1}{\partial \hat{y}_1} \frac{\partial \hat{y}_1}{\partial w_2} = -2(y_1 - \hat{y}_1) \cdot x_2$$

$$\frac{\partial L_1}{\partial w_3} = \frac{\partial L_1}{\partial \hat{y}_1} \frac{\partial \hat{y}_1}{\partial w_3} = -2(y_1 - \hat{y}_1) \cdot x_3$$

$$\frac{\partial L_1}{\partial b} = \frac{\partial L_1}{\partial \hat{y}_1} \frac{\partial \hat{y}_1}{\partial b} = -2(y_1 - \hat{y}_1)$$

$\frac{\partial \hat{y}_1}{\partial b} \hookrightarrow 1$

$$\frac{\partial L}{\partial W} = -2(y_1 - \hat{y}_1) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \frac{\partial L}{\partial b} = -2(y_1 - \hat{y}_1)$$

$$W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

→ For Stochastic gradient calculate for each record and update weight & bias

$$w_i := w_{i-1} - lr * \frac{\partial L}{\partial w_i} \qquad b_i := b_{i-1} - lr * \frac{\partial L}{\partial b}$$