Gradient Descent

Its a mathematical technique to find best weights and bias iteratively to get best least loss or correct loss

→ **Steps:**

1. Calculate loss with the current weight & bias
2. Determine the direction to move weights & bias to reduce the loss
3. Move the weight & ~~bias~~ bias with small values
4. Return ~~back~~ to step 1 process until the model cant reduce the loss any further.

Mathematical Calculation

pound in 1000's (feature)	Miles per gallon (label)
3.5	18
3.69	15
3.44	18
3.43	16
4.34	15
4.42	14
2.37	24

bias → y-intercept

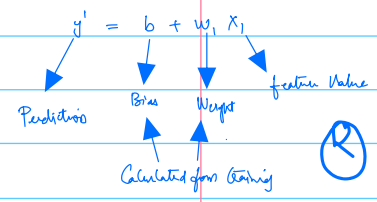
final bias

$y = mx + c$

① Initial values weight = 0, bias = 0

$$y = 0 + 0(x_1)$$

$$y' = b_1 + x_1 w_1$$



② Calculate MSE

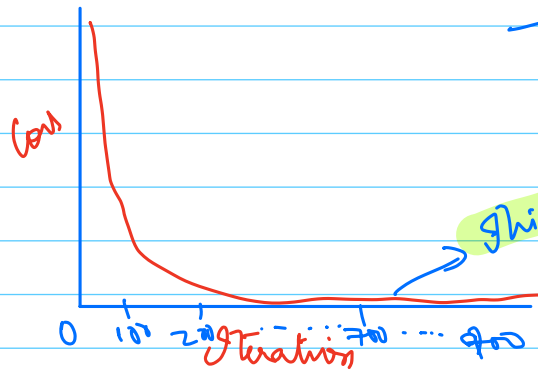
$$\text{Loss} = \frac{(18-0)^2 + (15-0)^2 + (18-0)^2 + (16-0)^2 + (15-0)^2 + (14-0)^2 + (24-0)^2}{7}$$

→ 30 3.71

③ Increases or decreases the weight & bias by small unit

④ We do this until the model converges

Converges ⇒



this state very less values change very little or not at all

Loss curve suggest convergence after 700 iterations

So Model converges when additional training won't improve the model.

When model converges basically gradient descent has found weights & biases that nearly minimize the loss.

Convergence & Convex function

↳ Loss function of linear regression always produce convex surface.

↳ Result of this property, when linear regression model converges produces least loss & correct bias & weights parameter