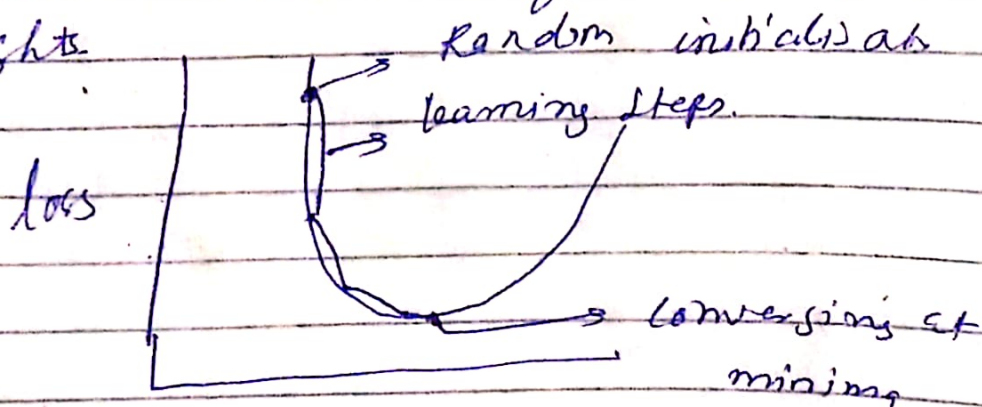


Day - 5

- Optimization algorithms are the engines that power neural network and enable them to learn patterns from data.
- Optimization algo. helps us to improve W and b .
- Most common algorithm for optimization -
→ Gradient Descent.
 - has two primary flavors:
 - standard "Vanilla"
 - Optimised "Stochastic"

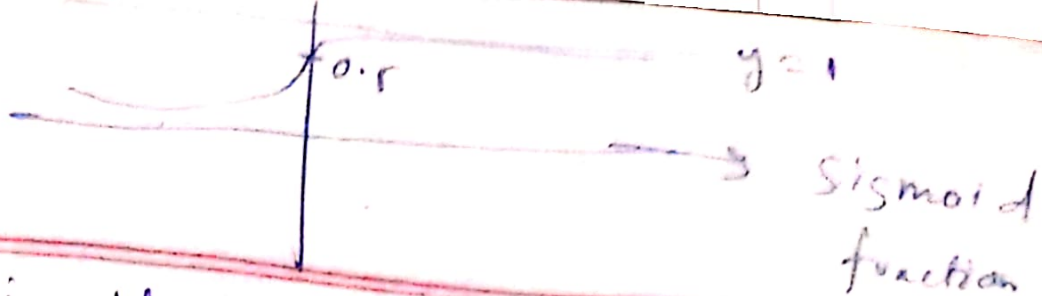
Goal of Gradient Descent is to minimize loss and find the respective weights.



Implementing G.D :-

- sigmoid function:

$$f(z) = \frac{1}{1 + e^{-z}}$$



Sigmoid functions maps every real number into the range of 0 to 1.

Hypothesis function:-

$$(Y = (X \times W) + b) \rightarrow \text{Prediction}$$

\downarrow \downarrow
 Weights bias.

Using sigmoid function

$$Y_{\hat{}} = \text{sigmoid}(Y_-)$$

Predictions:

Predictions is made on the value of $Y_{\hat{}}$ obtained from the sigmoid function.

if value is $Y_{\hat{}} = 0.5$ it is in class = 1. else it is class 0.

• Max-Iteration: No. of times our weights changes

$$W \rightarrow W - \eta \times \frac{dL}{dW} \rightarrow \text{Loss function}$$

$\eta \rightarrow$ Learning rate.

For Logistic Regression

$$\text{Loss function is } = -(Y \log(Y_-) + (1-Y_-) \log(1-Y_-))$$

$$\text{and Gradient is } = (Y_- - Y) \times X$$