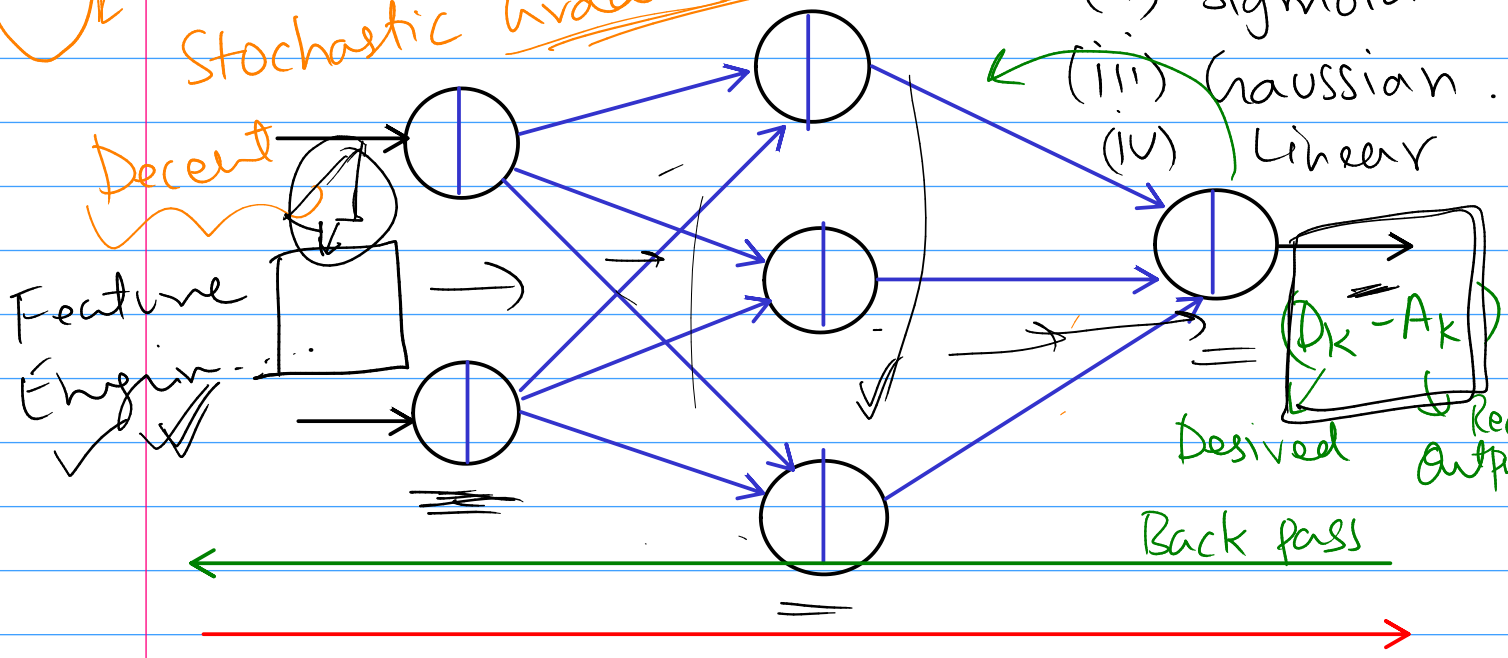
SGD (lr ≈ 1)
Stochastic Gradient Decent

(i) BTLN
(ii) Sigmoid
(iii) Gaussian
(iv) Linear



Shallow Learning

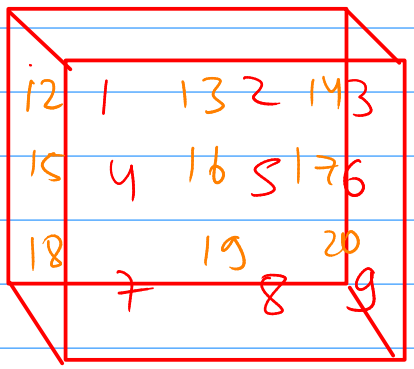
(i) Data

Scalars, Vectors, Matrices, Tensors.

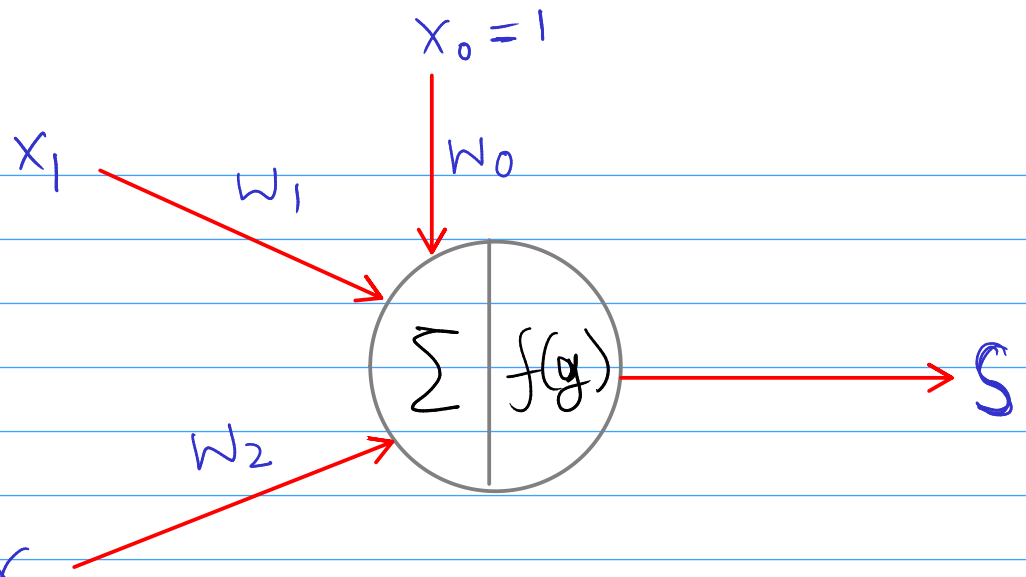
Numbers
(0, 1, 2, ...)
dim = 0

i
[0, 1, 2, 3]
dim = 1

i
j
$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix}$
n = 2



dim ≥ 3



In Vector form :

Input Vector $X = [1, x_1, x_2]$, $W = [w_0, w_1, w_2]$

$(1, 3)$ $(1, 3)$ Weight Vector.

Dot Product.

Calculate activation by X & W.

$y = ?$

$y = x_1 w_1 + x_2 w_2 + w_0$ (Matrix Dot Product Multiplication)

$$[\,]_{m \times n} \cdot [\,]_{n \times m} = [\,]_{m \times m}$$

$X = [1, x_1, x_2]$, $W = [w_0, w_1, w_2]$

$$W^T = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}$$

$\Rightarrow$ $XW^T = y$

$WX^T = y$

Activation equation...

⇒ PCA ⇒ Principal Component Analysis

⇒ Tensors :

Images ⇒ 240 X 320

RGB

240 X 320 X 3

Matrix

Tensors