**Maths Behind Simple Neural Network**

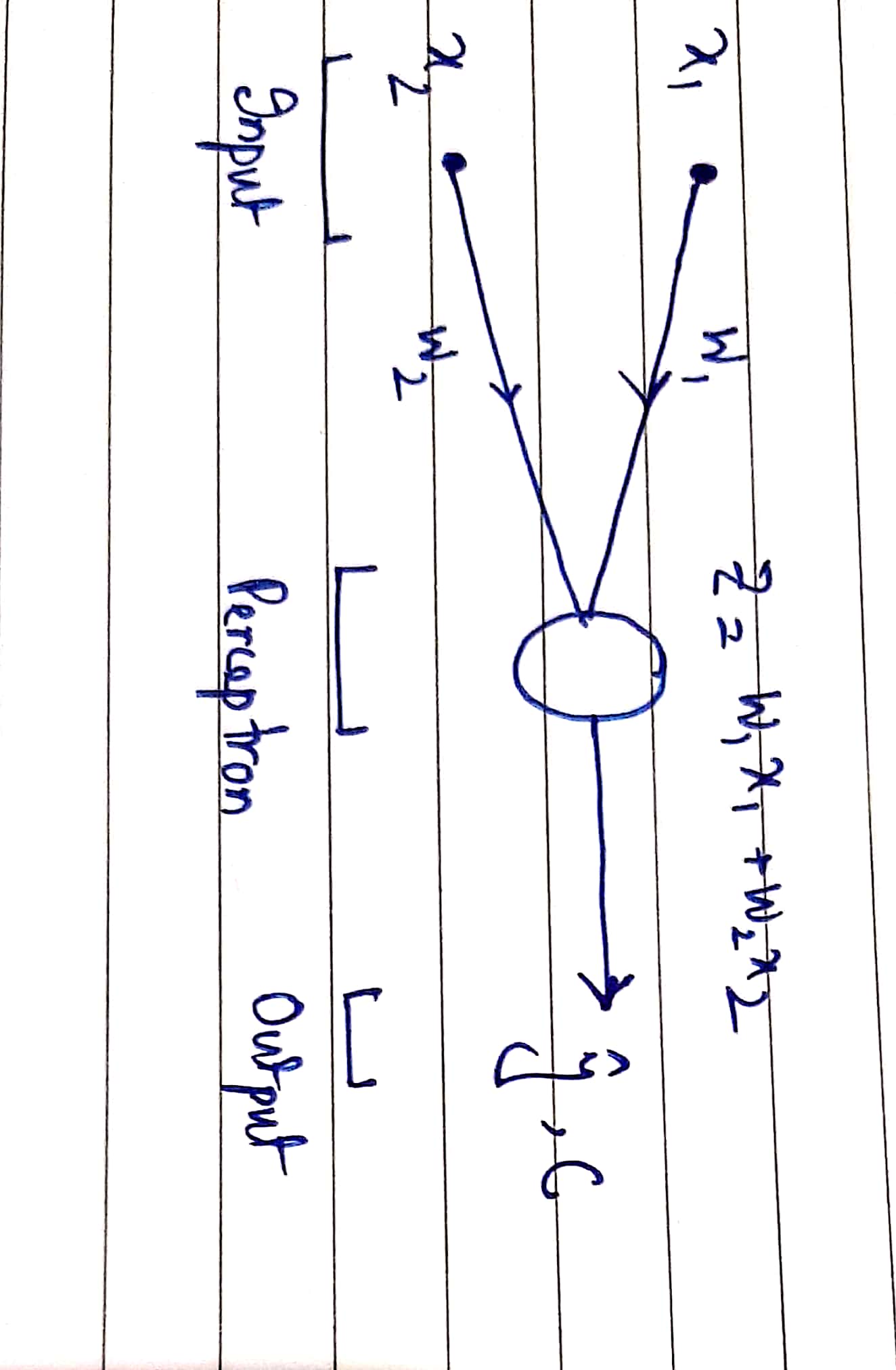
**Contents:**

*Section 1- Linear Perceptron*

*Section2 – Non-Linear Perceptron*

*Section3- Multi Layer Perceptron*

**Section 1- Linear Perceptron**

Appearance- 

**Math:**

(1)

(2)

(3)

Our aim is to reduce the error. From Eq (3) and Eq (2) shows that error is a function of as well as inputs x1 and x2.

Among these, only w1 and w2 i.e., weights and b are tuneable parameters, i.e., we can change it as we wish.

Hence, our aim is to change w1 and w2 and b such that error becomes low.

Updating weight-

for i=1,2 (4)

(5)

Let’s calculate first

But, c is not directly related to wi

Hence, apply chain rule-

(6)

1 2 3

Term 1- (7)

Term 2- (8)

Term 3- (9)

From Eq (6) and (7), (8) and (9)

(10)

Similarly,

(11)

(12)

for i=1,2 (13)

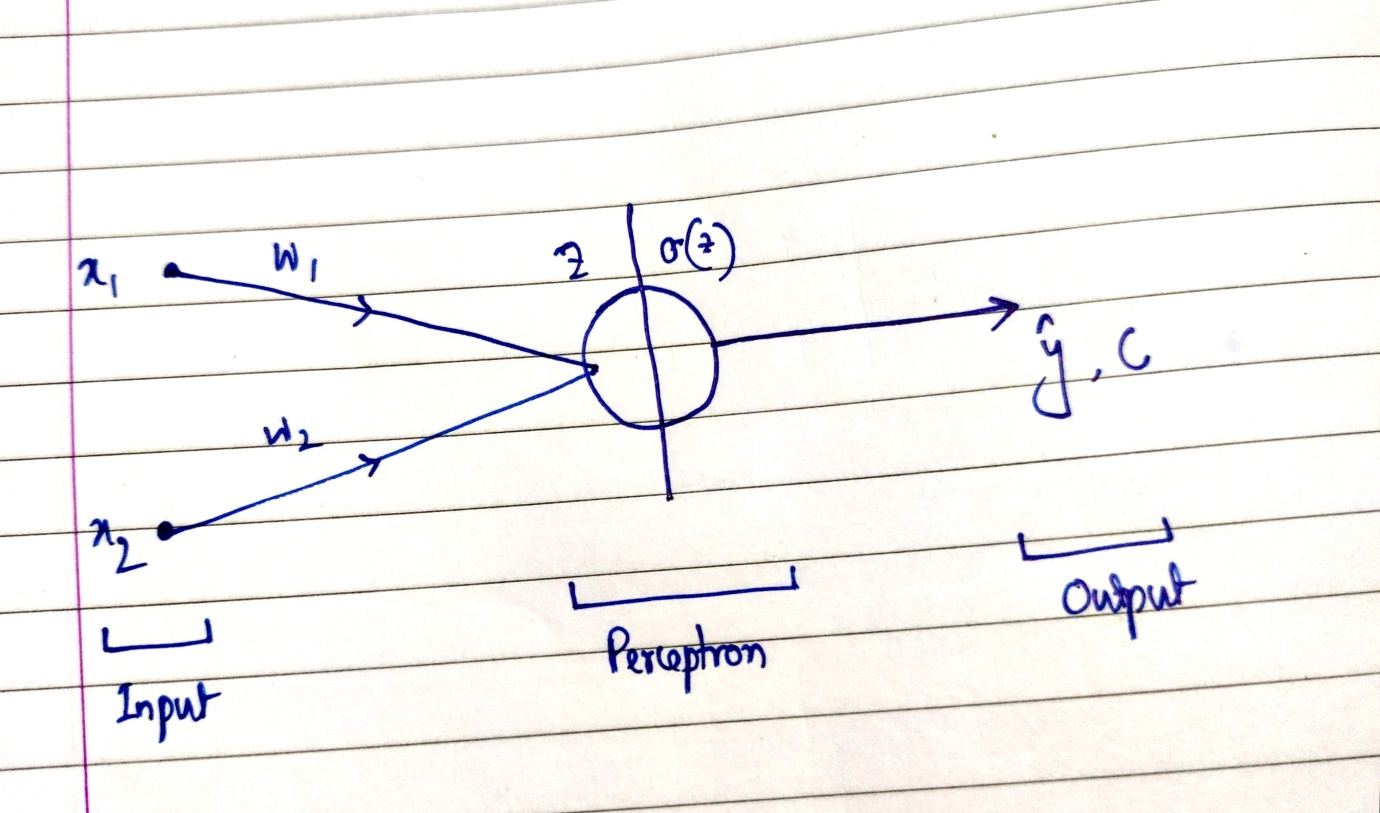
(14)

**Contd. To (Section-1) of**

***single\_linear\_nonlinear\_perceptron.ipynb***

**Section 2- Non-linear Perceptron**

Appearance:



**Math:**

(1)

Term 1- (2)

Term 2- (3)

Term 3- (4)

From Eq (1) and (2), (3) and (4)

(5)

(6)

Hence, from Eq (5) and (6)

(7)

(8)

Hence,

(9)

And,

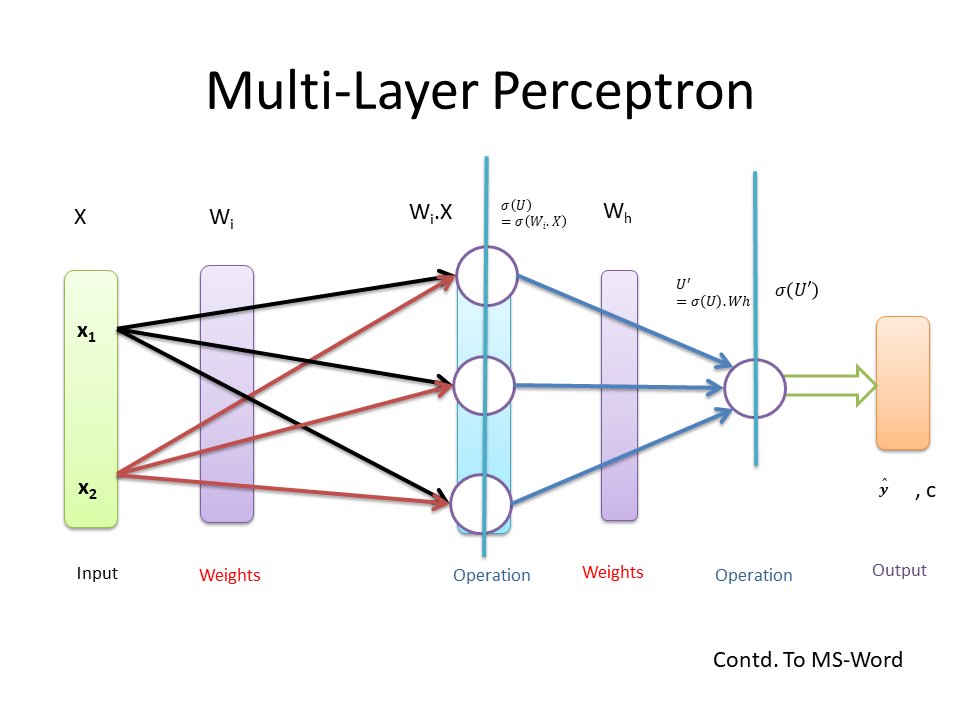
(10)

**Contd. To (Section-2) of**

***single\_linear\_nonlinear\_perceptron.ipynb***

**Section 3- Multi-layer Perceptron**

Appearance:

****

**Math:**

**(1)**

**(2)**

(3)

(4)

(5)

Now, Let’s calculate , c is not directly related to wh, so, again Chain Rule

(8)

Individual Terms,

Hence,

(9)

Now, Let’s calculate , c is not directly related to wi, so, again Chain Rule

(10)

Again, c is not directly related to h, hence again Chain Rule,

Hence, from Eq (10),

Individual Terms,

Hence,

(11)

Generalized the terms, is error or E.

is slope\_at\_output\_layer (SOL)

is slope\_at\_hidden\_layer (SHL)

X is input and wh is weight at output layer.

Hence, change factor (delta) at the output layer, dependent on the gradient of error multiplied by the slope of output layer activation= d\_out = E\*slope\_at\_output\_layer

**Final Anatomy Report**

**Output**

**Hidden**

**Input**

Error w.r.t Wi= Input X d\_hidden = SHL X Error\_hidden = Wh X d\_out = SOL X E

Error w.r.t Wh = Input to output layer X d\_out = SOL X E

Or,

(hidden layer activation)

**Contd. To *non\_linear\_mlp.ipynb***