1.i) $o = w_0 + w_1(x_1 + x^2) + \dots + w_n(x_n + x_n^2)$ —— (A)

Error function is given by

$$E(w) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2 \text{ —— (B)}$$

Gradient can be obtained by:

$$\frac{dE}{dw_i} = \frac{d}{dw_i} \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2$$

$$= \frac{1}{2} \sum_{d \in D} (2)(t_d - o_d) \frac{d}{dw_i} (t_d - o_d))$$

Substitution (A) in this, we get

$$= \frac{1}{\cancel{2}} \cancel{(2)} \sum_{d \in D} (t_d - o_d) \frac{\partial}{\partial w_i} \{t_d - (w_0 + w_1 x_1 + w_1 x_1^2 + \dots + w_n x_n + w_n x_n^2)\}$$

$$= \sum_{d \in D} (t_d - o_d)(-x_{id} - x_{id}^2) \text{ —— (C)}$$

where

$x_{id} \rightarrow$ single input component x_i for training example d.

Training rule is as follows :-

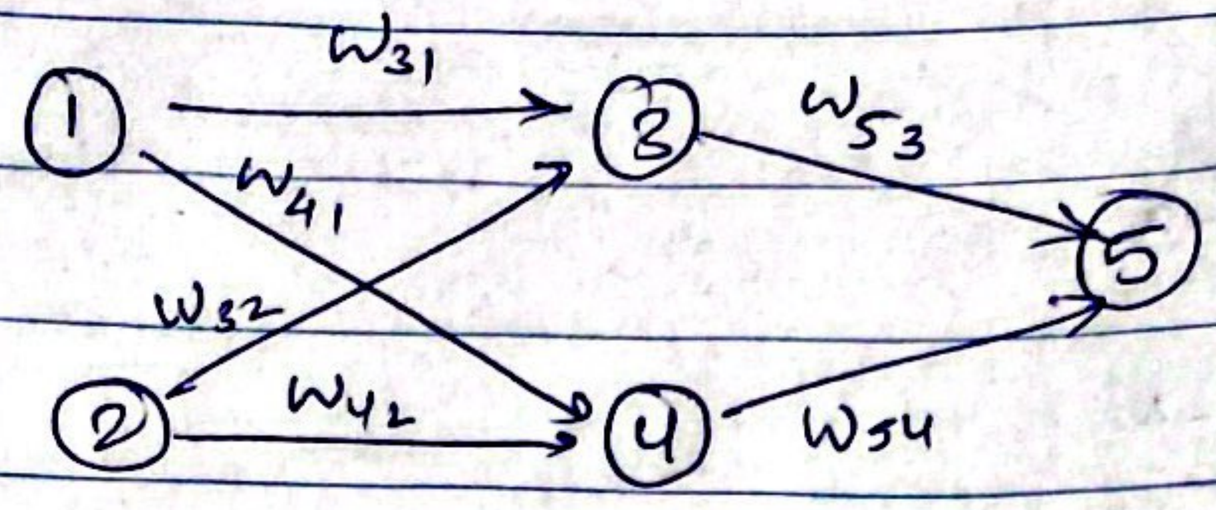
$$w_i = w_i + \Delta w_i$$

$$\Delta w_i = -\eta \frac{dE}{dw_i}$$

put this in eq (C) we get

$$\Delta w_i = (-\eta)\left(\sum_{d \in D} (t_d - o_d)\right)(-x_{id} - x_{id}^2)$$

1.(2)



(a) Input node 3 : $x_1 w_{31} + x_2 w_{32}$

Input node 4 : $x_1 w_{41} + x_2 w_{42}$

Output node 3 : $x_3 = h(x_1 w_{31} + x_2 w_{32})$

Output node 4 : $x_4 = h(x_1 w_{41} + x_2 w_{42})$

Input node 5 : $x_3 w_{53} + x_4 w_{54}$

$\therefore$ Output of node 5 $= y_5 = h(x_3 w_{53} + x_4 w_{54})$

1.2 (b) Let input for hidden layer be $H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$ where H_1 & H_2 are inputs for 3 & 4

$$H = X \cdot W^{(1)}$$

$$H = W^{(1)} \cdot X$$

$$\begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = \begin{bmatrix} w_{31} & w_{32} \\ w_{41} & w_{42} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Above is the input to the output $\hat{H}$

$$\hat{H} = h(H)$$

Output layer $= h(W^{(2)} \cdot \hat{H})$

So, Output layer $= h(W^{(2)} \cdot h(W^{(1)} \cdot X))$

(c) sigmoid − $h_s(x) = \frac{1}{1+e^{-x}}$

tanh − $h_t(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

$$h_{(t)}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$= \frac{e^x + e^{-x} - 2e^{-x}}{e^x - e^{-x}}$$ (Adding e^{-x} & $-e^{-x}$ in numerator)

$$= 1 + \frac{-2e^{-x}}{e^x + e^{-x}}$$

$$= 1 - \frac{2}{e^{2x}+1}$$

from logistic sigmoid's perspective

tanh − $h_t(x) = 1 - \frac{2}{e^{2x}+1}$

$$= 1 - 2h_s(-2x)$$

$$= 1 - 2(1 - h_s(2x))$$

$$= 1 - 2 + 2h_s(2x)$$

$$= 2h_s(2x) - 1$$

so, $h_t = 2h_s(2x) - 1$