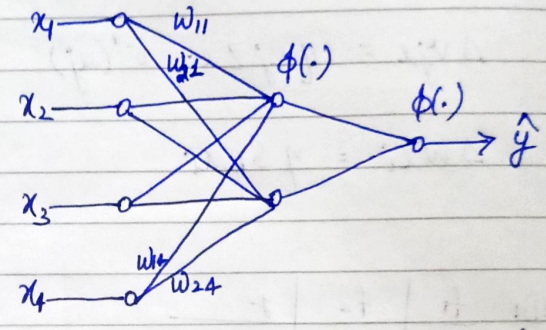
take initial weights :

$$W = \begin{bmatrix} 0.1 & 0.2 \\ 0.2 & 0.1 \\ 0.5 & 0.1 \\ 0.1 & 0.1 \end{bmatrix} \qquad V = \begin{bmatrix} 0.5 \\ 0.3 \end{bmatrix}$$

$\eta : 0.1$

Sol^n

	f_1	f_2	f_3	f_4	y
Training	0.5	0.1	0.5	11	1.1
	0.1	0.3	0.7	13	2.3
	0.3	0.7	0.9	10	1.5
	0.0	0.8	0.1	11	1.4
	0.5	0.1	0.3	16	3.3
Testing	0.9	0.3	0.2	10	2.7
	1.1	0.3	0.6	13	1.7



Activation function → sigmoid

for $\hat{y}$

$$\hat{y} = [0.5 \quad 0.1 \quad 0.5 \quad 11] \begin{bmatrix} 0.1 \\ 0.2 \\ \vdots \end{bmatrix}$$

for calc $\hat{y}_1$

$$z = [0.5 \quad 0.1 \quad 0.5 \quad 11] \begin{bmatrix} 0.1 & 0.2 \\ 0.2 & 0.1 \\ 0.5 & 0.1 \\ 0.1 & 0.1 \end{bmatrix} = [1.42 \quad 1.26]$$

$$\sigma(z) = \left[\frac{1}{1+e^{-1.42}} \quad \frac{1}{1+e^{-1.26}} \right] = [0.81 \quad 0.78]$$

$$[0.81 \quad 0.78] \begin{bmatrix} 0.5 \\ 0.3 \end{bmatrix} = 0.639$$

$$\hat{y}_1 = \sigma(0.639) = \frac{1}{1+e^{-0.639}} = 0.654$$

error = $0.654 - 1.1$

$$\delta^{(2)} = (\hat{y} - y) \cdot \hat{y}(1-\hat{y})$$
$$= (0.654 - 1.1) \cdot (0.654) \cdot (1 - 0.654) = -0.1009$$

$$\Delta v_{11} = -0.1009 \times 0.81 = -0.0813 = -0.08 \text{ (approx)}$$
$$\Delta v_{12} = -0.1009 \times 0.78 = -0.0786 = -0.078 \text{ (approx)}$$

$$\text{updated } V = \begin{bmatrix} v_{11} - \eta \Delta v_{11} \\ v_{12} - \eta \Delta v_{12} \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 - (0.1)(-0.0813) \\ 0.3 - (0.1)(-0.0786) \end{bmatrix} = \begin{bmatrix} 0.5081 \\ 0.3079 \end{bmatrix}$$

for hidden layer

$$\delta_j^{(1)} = \delta^{(2)} \cdot v_j \cdot a_j (1 - a_j)$$

$$\delta_1^{(1)} = (-0.1009) \cdot (0.5) \cdot (0.81) \cdot (1 - 0.81)$$
$$= -0.079 - 0.0079$$

$$\delta_2^{(1)} = (-0.1009) \cdot (0.3) \cdot (0.788) \cdot (1 - 0.78)$$
$$= -0.0052$$

$$\Delta W_{ij} = \delta_j^{(1)} \cdot x_i$$

update $W_{ij} = W_{ij} - \eta \Delta W_{ij}$

updating W:

$$W = \begin{bmatrix} 0.1 - (0.1)(-0.0079)\times 0.5 & 0.2 - (0.1)(-0.0052)(0.5) \\ 0.2 - (0.1)(-0.0079)\times 0.1 & 0.1 - (0.1)(-0.0052)(0.1) \\ 0.5 - (0.1)(-0.079)\times 0.5 & 0.1 - (0.1)(-0.0052)(0.5) \\ 0.1 - (0.1)(-0.079)\times 11 & 0.1 - (0.1)(-0.0052)(11 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.1004 & 0.2003 \\ 0.20008 & 0.10005 \\ 0.5004 & 0.1003 \\ 0.1087 & 0.1057 \end{bmatrix}$$

for training second data point:

$$z = [0.1 \quad 0.3 \quad 0.7 \quad 13] \begin{bmatrix} 0.1004 & 0.2003 \\ 0.20008 & 0.10005 \\ 0.5004 & 0.1003 \\ 0.1087 & 0.1057 \end{bmatrix}$$

$$z = [1.833 \quad 1.494]$$

$$\sigma(z) = \left[\frac{1}{1+e^{-1.833}} \quad \frac{1}{1+e^{-1.494}}\right]$$

$$[0.8621 \quad 0.8167]$$

$$[0.8621 \quad 0.8167] \begin{bmatrix} 0.5081 \\ 0.3079 \end{bmatrix} = 0.6895$$

$$\hat{y}_2 = \frac{1}{1+e^{-0.6895}} = 0.6659$$

$$\delta^{(2)} = (0.6659 - 2.3) \cdot (0.6659)(0.3341)$$
$$= -0.3636$$

updating V:

$$V = \begin{bmatrix} 0.5081 - (0.1)(-0.3636)(0.8621) \\ 0.3079 - (0.1)(-0.3636)(0.8167) \end{bmatrix}$$

$$V = \begin{bmatrix} 0.5395 \\ 0.3376 \end{bmatrix}$$

$$\delta_1^{(1)} = (-0.3636) \cdot (0.5081) \cdot (0.8621) \cdot (1 - 0.8621)$$
$$= -0.0220$$

$$\delta_2^{(1)} = (-0.3636)(0.3079)(0.8167)(1 - 0.8167)$$
$$= -0.0167$$

updating W:

$$W = \begin{bmatrix} 0.1004 - (0.1)(-0.022) & 0.1 & 0.2003 - (0.1)(-0.0167) & 0.1 \\ 0.20008 - (0.1)(-0.022) & 0.3 & 0.10005 - (0.1)(-0.167) & 0.3 \\ 0.5004 - (0.1)(-0.022) & 0.7 & 0.1003 - (0.1)(-0.0167) & 0.7 \\ 0.1087 - (0.1)(-0.022) & 13 & 0.1057 - (0.1)(0.0167) & 13 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.10062 & 0.20047 \\ 0.20074 & 0.10055 \\ 0.50194 & 0.10147 \\ 0.1373 & 0.1274 \end{bmatrix}$$

for 3rd data point:

~~$x = [0.2 \quad 0.1 \quad 0.9 \quad 10]$ $\quad y = 1$~~

$x = [0.3 \quad 0.7 \quad 0.9 \quad 10]$ $\quad y = 1.5$

$$z = \begin{bmatrix} 0.3 & 0.7 & 0.9 & 1.0 \end{bmatrix} \begin{bmatrix} 0.10062 & 0.20047 \\ 0.20074 & 0.10055 \\ 0.50194 & 0.10147 \\ 0.1373 & 0.1274 \end{bmatrix}$$

$$z = \begin{pmatrix} 1.995 & 1.495 \end{pmatrix}$$

$$\sigma(z) = \left(\frac{1}{1+e^{-1.995}} \quad \frac{1}{1+e^{-1.495}} \right)$$

$$\sigma(z) = \begin{pmatrix} 0.880 & 0.816 \end{pmatrix}$$

$$\begin{bmatrix} 0.880 & 0.816 \end{bmatrix} \begin{bmatrix} 0.5395 \\ 0.3376 \end{bmatrix} = 0.7505$$

$$\sigma(0.7505) = \frac{1}{1+e^{-0.7505}} = 0.6793$$

$$\text{error} = 0.6793 - 1.5$$

$$\delta^{(2)} = (0.6793 - 1.5)(0.6793)(1 - 0.6793)$$
$$= -0.1788$$

updating V,

$$V = \begin{bmatrix} 0.5395 - (0.1)(-0.1788)(0.880) \\ 0.3376 - (0.1)(-0.1788)(0.816) \end{bmatrix}$$

$$V = \begin{pmatrix} 0.5395 - (-0.01574) \\ 0.3376 - (-0.01461) \end{pmatrix}$$

$$V = \begin{bmatrix} 0.5552 \\ 0.3522 \end{bmatrix}$$

$$\delta_1^{(1)} = (-0.1788)(0.5395)(0.880)(1-0.880)$$
$$= -0.0102$$

$$\delta_2^{(1)} = (-0.1788)(0.3376)(0.8168)(1-0.8168)$$
$$= -0.0090$$

updating W,

$$W = \begin{bmatrix} 0.10062-(0.1)(-0.0102)\ 0.3 & 0.20047-(0.1)(-0.009)\ 0.3 \\ 0.20074-(0.1)(-0.0102)\ 0.7 & 0.10055-(0.1)(-0.009)\ 0.7 \\ 0.50194-(0.1)(-0.0102)\ 0.9 & 0.10147-(0.1)(-0.009)\ 0.9 \\ 0.1373-(0.1)(-0.0102)\ 10 & 0.1274-(0.1)(-0.009)\ 10 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.10062-(-0.000306) & 0.20047-(-0.00027) \\ 0.20074-(-0.000714) & 0.10055-(-0.00063) \\ 0.50194-(-0.000918) & 0.10147-(-0.00081) \\ 0.1373-(-0.0102) & 0.1274-(-0.0090) \end{bmatrix}$$

$$W = \begin{bmatrix} 0.100926 & 0.20074 \\ 0.201454 & 0.10118 \\ 0.502858 & 0.10228 \\ 0.1475 & 0.1364 \end{bmatrix}$$

for 4th data point:

$$Z = \begin{bmatrix} 0.0 & 0.8 & 0.1 & 11 \end{bmatrix} \begin{bmatrix} 0.100926 & 0.20074 \\ 0.201454 & 0.10118 \\ 0.502858 & 0.10228 \\ 0.1475 & 0.1364 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1.83395 & 1.59157 \end{bmatrix}$$

$$\sigma(z) = \left(\frac{1}{1+e^{-1.83395}} \quad \frac{1}{1+e^{-1.9157}} \right)$$

$$\sigma(z) = \begin{pmatrix} 0.8621 & 0.8309 \end{pmatrix}$$

$$\begin{pmatrix} 0.8621 & 0.8309 \end{pmatrix} \begin{pmatrix} 0.5552 \\ 0.3522 \end{pmatrix} = 0.7713$$

$$\hat{y}_4 = \sigma(7713) = \frac{1}{1+e^{-7713}} = 0.6838$$

$$\text{error} = \hat{y}_4 - y_4$$
$$0.6838 - 1.4$$

$$\delta^{(2)} = (0.6838 - 1.4)(0.6838)(1 - 0.3162)$$
$$= -0.1548$$

updating V,

$$V = \begin{bmatrix} 0.5552 - (0.1)(-0.1548) & 0.8621 \\ 0.3522 - (0.1)(-0.1548) & 0.8309 \end{bmatrix}$$

$$V = \begin{pmatrix} 0.5685 \\ 0.3651 \end{pmatrix}$$

$$\delta_1^{(1)} = (-0.1548)(0.5552)(0.8621)(0.1379)$$
$$= -0.0102$$

$$\delta_2^{(1)} = (-0.1548)(0.3522)(0.8309)(0.1691)$$
$$= -0.0078$$

updating W,

$$W = \begin{bmatrix} 0.100926 - (0.1)(-0.0102) & 0.0 & 0.20074 - (0.1)(-0.0076) & 0.0 \\ 0.201454 - (0.1)(-0.0102) & 0.8 & 0.10118 - (0.1)(-0.0076) & 0.8 \\ 0.502858 - (0.1)(-0.0102) & 0.1 & 0.10228 - (0.1)(-0.0076) & 0.1 \\ 0.1475 - (0.1)(-0.0102) & 11 & 0.1364 - (0.1)(-0.0076) & 01 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.100926 & 0.20074 \\ 0.20227 & 0.101788 \\ 0.50296 & 0.10236 \\ 0.15872 & 0.14476 \end{bmatrix}$$

for 5th data point:

$$z = \begin{bmatrix} 0.5 & 0.1 & 0.3 & 16 \end{bmatrix} \begin{pmatrix} 0.100926 & 0.20074 \\ 0.20227 & 0.101788 \\ 0.50296 & 0.10236 \\ 0.15872 & 0.14476 \end{pmatrix}$$

$$z = \begin{pmatrix} 2.7611 & 2.4575 \end{pmatrix}$$

$$\sigma(z) = \begin{pmatrix} \frac{1}{1+e^{-2.7611}} & \frac{1}{1+e^{-2.4575}} \end{pmatrix}$$

$$\sigma(z) = \begin{pmatrix} 0.9406 & 0.9211 \end{pmatrix}$$

$$\begin{pmatrix} 0.9406 & 0.9211 \end{pmatrix} \begin{pmatrix} 0.5685 \\ 0.3651 \end{pmatrix} = 0.8714$$

$$\hat{y}_5 = \frac{1}{1+e^{-0.8714}} = 0.7050$$

error = $\hat{y} - y = 0.7050 - 3.3$

$\delta^{(2)} = (0.7050 - 3.3)(0.7050)(1 - 0.7050)$

$= -0.5397$

updating V,

$$V = \begin{bmatrix} 0.5685 - (0.1)(-0.5397)(0.9406) \\ 0.3651 - (0.1)(-0.5397)(0.9211) \end{bmatrix}$$

$$V = \begin{pmatrix} 0.6193 \\ 0.4148 \end{pmatrix}$$

$\delta_1^{(1)} = (-0.5397)(0.5685)(0.9406)(1 - 0.9406)$

$= -0.0172$

$\delta_2^{(1)} = (-.5397)(0.3651)(0.9211)(1 - 0.9211)$

$= -0.0143$

updating W,

$$W = \begin{bmatrix} 0.100926 - (0.1)(-0.0172)0.5 & 0.200774 - (0.1)(-0.0143)0.5 \\ 0.202227 - (0.1)(-0.0172)0.1 & 0.101788 - (0.1)(-0.0143)0.1 \\ 0.50296 - (0.1)(-0.0172)0.3 & 0.10236 - (0.1)(-0.0143)0.3 \\ 0.15872 - (0.1)(-0.0172)16 & 0.14476 - (0.1)(-0.0143)16 \end{bmatrix}$$

$$W = \begin{pmatrix} 0.101786 & 0.201455 \\ 0.202442 & 0.101931 \\ 0.503476 & 0.102789 \\ 0.18624 & 0.16764 \end{pmatrix}$$

Testing data point 1: [0.9 0.3 0.2 1.0]

$$z = [0.9 \quad 0.3 \quad 0.2 \quad 1.0] \begin{bmatrix} 0.101786 & 0.201455 \\ 0.202442 & 0.101931 \\ 0.503476 & 0.102789 \\ 0.18624 & 0.16764 \end{bmatrix}$$

$$z = (2.1154 \quad 1.9089)$$

$$\sigma(z) = \left(\frac{1}{1+e^{-2.1154}} \quad \frac{1}{1+e^{-1.9089}} \right)$$

$$\sigma(z) = (0.8923 \quad 0.8712)$$

$$(0.8923 \quad 0.8712) \begin{pmatrix} 0.6193 \\ 0.4148 \end{pmatrix} = 0.9137$$

$$\hat{y} = \sigma(0.9137) = \frac{1}{1+e^{-0.9137}} = 0.7138$$

$$\text{Squared error} = \frac{1}{2}(2.7 - 0.7138)^2 = 1.972$$

Testing data point 2: [1.1 0.3 0.6 1.3]

$$z = [1.1 \quad 0.3 \quad 0.6 \quad 1.3] \begin{bmatrix} 0.101786 & 0.201455 \\ 0.202442 & 0.101931 \\ 0.503476 & 0.102789 \\ 0.18624 & 0.16764 \end{bmatrix}$$

$$z = (2.8959 \quad 2.4931)$$

$$\sigma(z) = \left(\frac{1}{1+e^{-2.8959}} \quad \frac{1}{1+e^{-2.4931}} \right) = (0.9475 \quad 0.9238)$$

$$\begin{pmatrix} 0.9475 & 0.9238 \end{pmatrix} \begin{pmatrix} 0.6193 \\ 0.4148 \end{pmatrix} = 0.9701$$

$$\hat{y} = \sigma(0.9701) = \frac{1}{1 + e^{-0.9701}} = 0.7251$$

$$\text{Squared error} = \frac{1}{2}(1.7 - 0.7251)^2$$

$$= 0.475$$