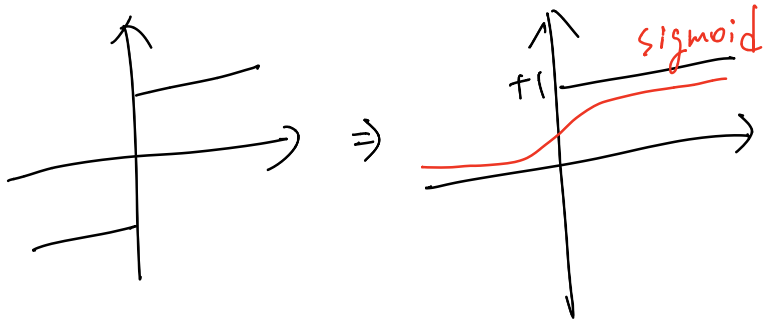
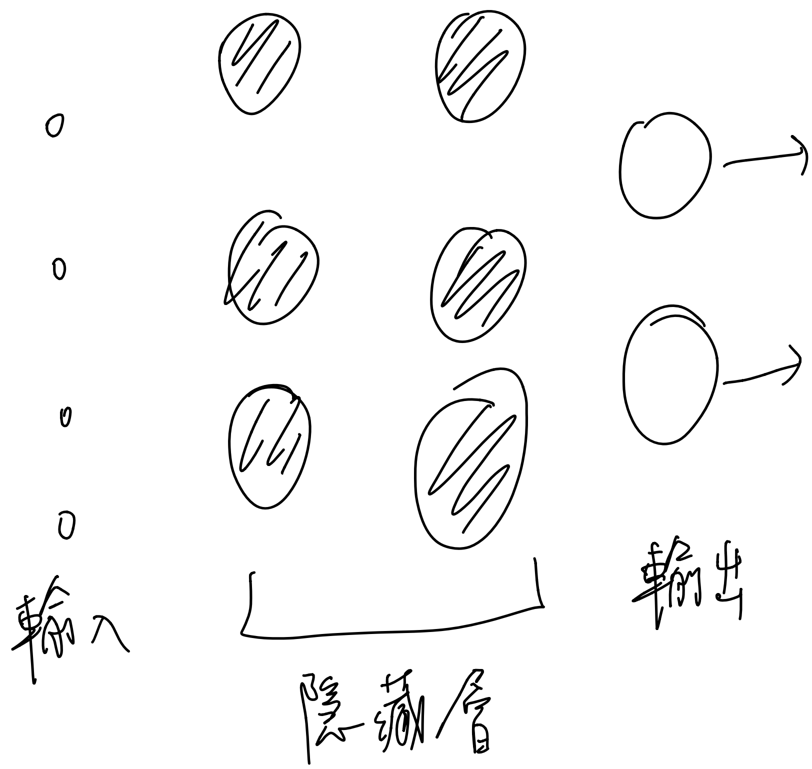
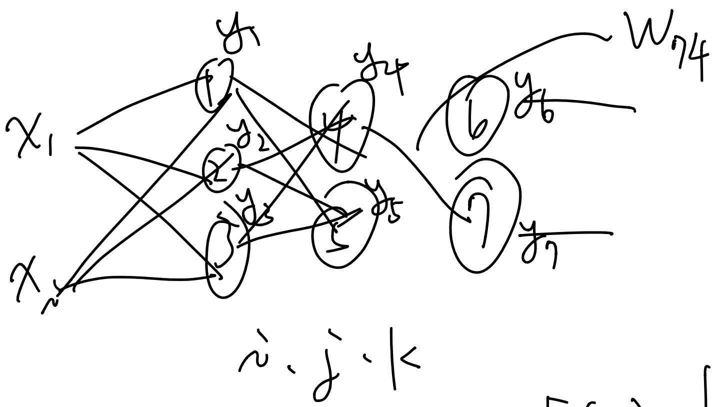
Ch3 note



$$y_j = \varphi(v_j) = \frac{1}{1+e^{-v_j}}$$



*此為3層(有計算能力才算層)



3、3(1)

$$E(n) = \frac{1}{2}\left(e_6^2 + e_7^2\right)$$

$$i \xrightarrow{W_{ji}} j$$

x_1	d_1
⋮	
x_N	d_N

$$E_{av} = \frac{E(1) + E(2) + \cdots + E(N)}{N}$$

※ 3.3(3)

$$E(n) = \frac{1}{2}\sum_{j\in C} e_j^2(n) = \frac{1}{2}\sum_{j\in C}\left(d_j(n)-y_j(n)\right)^2$$

$$V_j(n) = \sum_i W_{ji}(n)\,x_i(n)$$

$$y_j(n) = \varphi_j\left(V_j(n)\right)$$

$$\boxed{W_{ji}(n+1) = W_{ji}(n) - \eta\frac{\partial E(n)}{\partial W_{ji}(n)}} = W_{ji}(n) + \underbrace{\eta\,\delta_j(n)\cdot y_i(n)}_{}$$

$$\Delta W_{ji}(n) = \eta\,\delta_j(n)\cdot y_i(n)$$

$$\frac{\partial E(n)}{\partial W_{ji}(n)} = \frac{\partial E(n)}{\partial V_j(n)}\,\frac{\partial V_j(n)}{\partial W_{ji}(n)}$$

$$\frac{\partial E(n)}{\partial V_j(n)} = -\delta_j(n) \qquad \frac{\partial V_j(n)}{\partial W_{ji}(n)} = x_i(n)$$

$$\delta_j(n) = -\frac{\partial E(n)}{\partial V_j(n)}$$

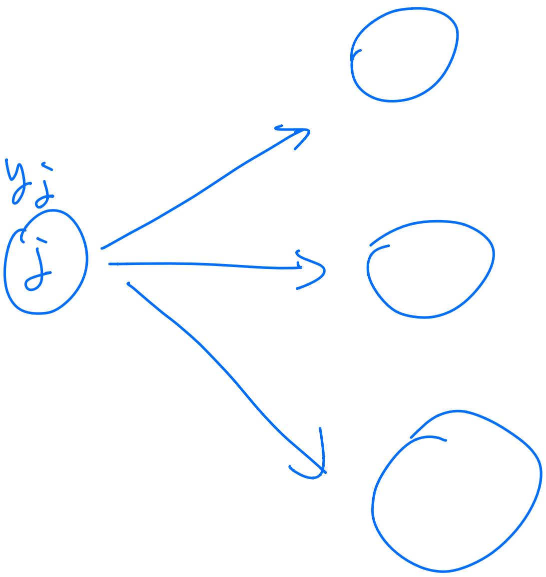
※ 依是否為輸出層
有不同算法

Case 1 $j \in C$ (輸出)

$$\delta_j(n) = -\frac{\partial E(n)}{\partial v_j(n)} = -\frac{\partial E(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)}$$

$$= -\left(d_j(n) - y_j(n)\right)(-1) \cdot \varphi_j'\left(v_j(n)\right)$$

★ Case 2 $\quad j \notin C$



ex: $f(x,y) = 3xy^2$

$$\frac{\partial f}{\partial y} = 6xy$$

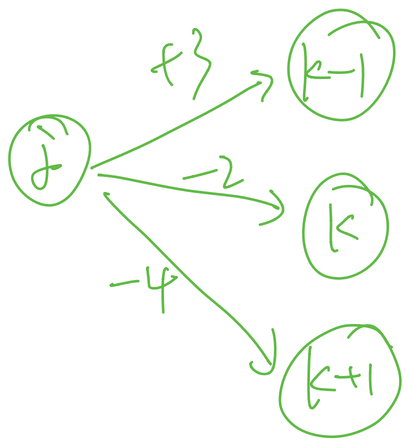
$$\Sigma$$

$$f(x,y) = 3xy^2 - 4x^2y = \underline{f_1(x,y)} + \underline{f_2(x,y)}$$

$$\delta_j(n) = -\frac{\partial E(n)}{\partial v_j(n)} = -\frac{\partial E(n)}{\partial y_j(n)} \cdot \frac{\overset{\varphi(v_j(n))}{\overset{\shortparallel}{\partial y_j(n)}}}{\partial v_j(n)}$$

*

δ：責任！



$\delta_{k-1} = 2.1$

$\delta_k = -1.2$

$\delta_{k+1} = 2.3$

$$\delta_j = \varphi_j'(v_n)\left(+3 \cdot 2.1 + (-2)(-1.2) + (-4) \cdot 2.3\right)$$

$$\delta_j(n) = -\frac{\partial E(n)}{\partial v_j(n)}$$

$$* \quad y_j(n) = \varphi_j(v_j(n)) = \frac{1}{1 + e^{-v_j(n)}}$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)) = \frac{\partial (1 + e^{-v_j(n)})^{-1}}{\partial v_j(n)}$$

$$= -(1 + e^{-v_j(n)})^{-2} \cdot (-1) e^{-v_j(n)}$$

$$= \underbrace{\frac{1}{(1 + e^{-v_j(n)})}}_{= y_j(n)} \times \frac{e^{-v_j(n)}}{(1 + e^{-v_j(n)})}$$

$$1 - \frac{1}{1 + e^{-v_j(n)}}$$

$$= 1 - y_j(n)$$

$$= y_j(n)(1 - y_j(n))$$

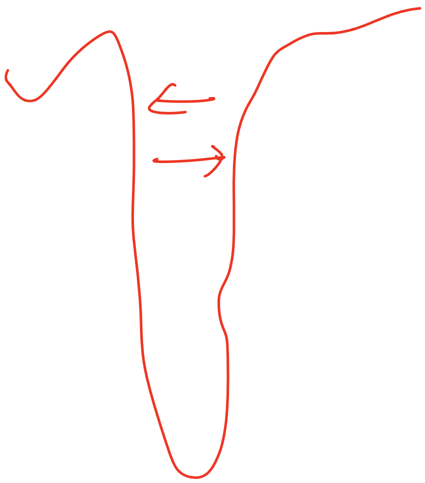
review ch3

① $j \in C$ v.s. $j \notin C$

② $g(v_j(n)) = ?$

3.4 (2)

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n)$$



$+$ $\quad$ $+$

表示同向 → 步伐可加大

$+$ $\quad$ $-$

$-$ $\quad$ $+$

表示不同向 → 步伐 ↓

☆ 3.4(3)

1、Pattern Recognition

Classfication

K class   ex: K=2

$\boxed{\theta}$ — $y_j [0,1]$

$y_j > 0.5 \to 1$

$y_j \le 0.5 \to 2$

2、Function Approximation

Regression

(1,3)

input_data = [-1, 1, 1]

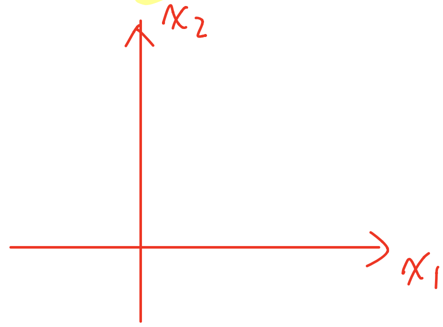
$W_1(0) = [[-1, -1]$ $= (f+1, neu)$

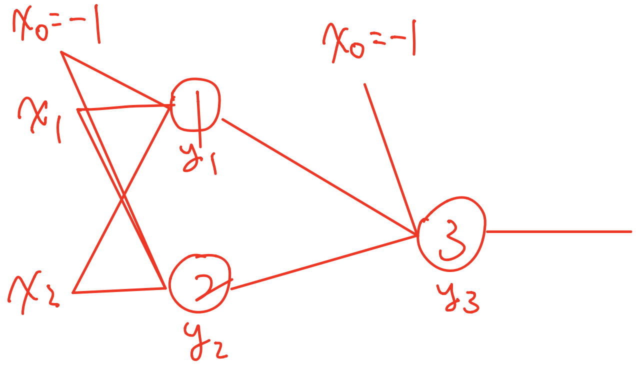
(3,2) $= [-2, -2]$

$[-3, -3]]$

* EX 3.1 XOR problem (1)

x	d
$\begin{pmatrix}0\\0\end{pmatrix}$	0
$\begin{pmatrix}1\\1\end{pmatrix}$	0
$\begin{pmatrix}1\\0\end{pmatrix}$	1
$\begin{pmatrix}0\\1\end{pmatrix}$	1





$x = \begin{pmatrix}-1\\1\\1\end{pmatrix}$

$\eta = 0.5$

$W_1(0) = \begin{pmatrix}-1.2\\1\\1\end{pmatrix}$ $W_2(0) = \begin{pmatrix}0.3\\1\\1\end{pmatrix}$ $W_3(0) = \begin{pmatrix}0.5\\0.4\\0.8\end{pmatrix}$

$V_1 = 1.2+1+1 = 3.2$ $V_2 = -0.3+1+1 = 1.7$

$y_1 = \dfrac{1}{1+e^{-3.2}} = 0.96$ $y_2 = \dfrac{1}{1+e^{-1.7}} = 0.84$

$V_3 = -0.5 + 0.4 \times 0.96 + 0.8 \times 0.84$

$y_3 = 0.63$

Case 1 : $j \in C$

$$\delta_j = y_j(1-y_j)(d_j - y_j)$$

Case 2 : $j \notin C$

$$\delta_j = y_i(1-y_i)\sum_k W_{kj}\delta_j$$

$$\delta_3 = 0.63 \times (1-0.63)(0-0.63)$$

$$= -0.147$$

$$\delta_2 = 0.84 \times (1-0.84)\cdot 0.8 \cdot \underbrace{(-0.147)}_{} = -0.0158$$

$$\delta_1 = 0.96 \times (1-0.96)\cdot 0.4 \cdot (-0.147) = -0.002$$

$$w_{ji}(n+1) = w_{ji}(n) + \eta \delta_j(n) y_j(n) \quad \boxed{3.1(4)}$$

$$W_1(1) = \begin{pmatrix} -1.2 \\ 1 \\ 1 \end{pmatrix} + 0.5 \times (-0.0002) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1.199 \\ 0.999 \\ 0.999 \end{pmatrix}$$

$W_2(1)$

- - - - - - -

$W_3(1)$

收斂成 $W_1(n) = \begin{pmatrix} -1.198 \\ 0.9121 \\ 1.179 \end{pmatrix}$ $W_3(n) = \begin{pmatrix} \\ \\ \end{pmatrix}$

$$W_2(n) = \begin{pmatrix} 0.294 \\ 0.836 \\ 0.98 \end{pmatrix}$$

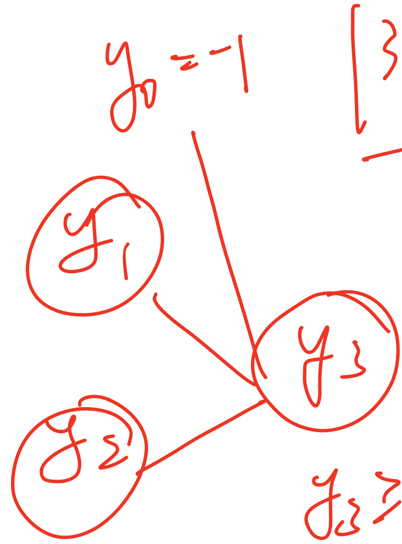
$\therefore V_1 = (-1.198, 0.9121, 1.179)\begin{pmatrix}-1\\0\\0\end{pmatrix}$

$y_1 =$

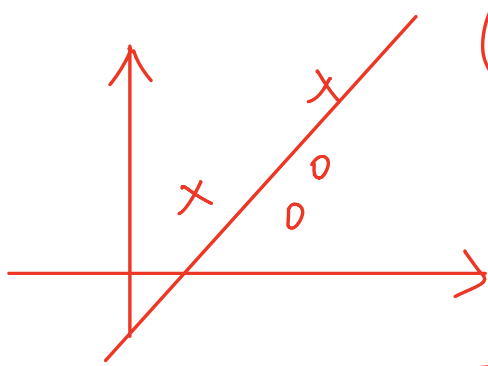
$V_2 =$

$y_2 =$

$\boxed{3.1.13}$



对 y_3 来看



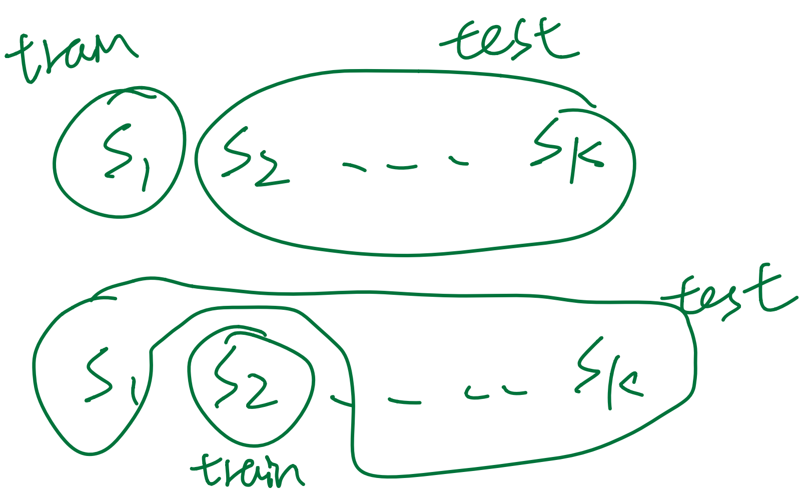
$y_3 \geq 0.5$ $y=1$

$y_3 < 0.5$ $y=0$

隱藏層將 $\boxed{\text{非線性}}$ ⇒ $\boxed{\text{線性}}$

* System Identification $\boxed{3.2}$

* K-fold 3.6(1)



將all結果加起來平均

每個等皆訓練到