② here, a three layer network

$h^1 = \sigma(w^1 x)$

$h^2 = \sigma(w^2 h^1)$

$f(x) = \langle w^3, h^2 \rangle$

→ here, x is a vector so, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

w^1 is a matrix so, $w^1 = \begin{bmatrix} w^1_{11} & w^1_{12} \\ w^1_{21} & w^1_{22} \end{bmatrix}$

w^2 is a matrix so, $w^2 = \begin{bmatrix} w^2_{11} & w^2_{12} \\ w^2_{21} & w^2_{22} \end{bmatrix}$

w^3 is a matrix so, $w^3 = [w^3_1, \; w^3_2]$

→ $f = \langle w^3, h^2 \rangle$

$$= [w^3_1 \;\; w^3_2]_{1\times 2} \begin{bmatrix} \sigma(w^2_{11}\sigma(w^1_{11}x_2 + w^1_{12}x_2) + w^2_{12}\sigma(w^1_{21}x_1 + w^1_{22}x_2)) \\ \sigma(w^2_{21}\sigma(w^1_{11}x_1 + w^1_{12}x_2) + w^2_{22}\sigma(w^1_{21}x_1 + w^1_{22}x_2)) \end{bmatrix}_{2\times 1}$$

$$= w^3_1[\sigma(w^2_{11}\sigma(w^1_{11}x_1 + w^1_{12}x_2) + w^2_{12}\sigma(w^1_{21}x_2 + w^1_{22}x_2))]$$
$$+ w^3_2[\sigma(w^2_{21}\sigma(w^1_{11}x_1 + w^1_{12}x_2) + w^2_{22}\sigma(w^1_{21}x_1 + w^1_{22}x_2))]$$

also we can write f different way using figure

$$f = w^3_1[\sigma(h^2_1)] + w^3_2[\sigma(h^2_2)]$$
$$= w^3_1[\sigma(w^2_{11}\sigma(h^1_1) + w^2_{12}\sigma(h^1_2))] +$$
$$w^3_2[\sigma(w^2_{21}\sigma(h^1_1) + w^2_{22}\sigma(h^1_2))]$$

now, for compute, $\frac{\partial f}{\partial w^1_{ij}}$

now looking from figure we can apply chain rule for all four derivations

Scanned with CamScanner

$$\frac{\partial f}{\partial \omega'_{11}} = \omega_1^3 6'[\omega_{11}^2 6(\omega'_{11}x_1 + \omega'_{21}x_2) + \omega_{21}^2 6(\omega'_{12}x_1 + \omega'_{22}x^2)]$$

$$* [\omega_{11}^2 6'(\omega'_{11}x_1 + \omega'_{21}x_2)\cdot x_1]$$

$$+ \omega_2^3 6'[\omega'_{12} 6(\omega'_{11}x_1 + \omega'_{21}x_2) + \omega_{22}^2 6(\omega'_{12}x_1 + \omega'_{22}x_2)] * [\omega'_{12} 6'(\omega'_{11}x_1 + \omega'_{21}x_2)\cdot x_1]$$

$$\therefore \frac{\partial f}{\partial \omega'_{11}} = \omega_1^3 6'(h_1^2)\cdot[\omega_{11}^2 6'(h'_1)\cdot x_1] + \omega_2^3 6'(h_2^2)\cdot[\omega_{12}^2 6'(h'_1)\cdot x_1] \quad \text{———} \ (1)$$

Similarly,

$$\therefore \frac{\partial f}{\partial \omega'_{12}} = \omega_1^3 6'(h_1^2)\cdot[\omega_{21}^2 6'(h'_2)\cdot x_1] + \omega_2^3 6'(h_2^2)\cdot[\omega_{22}^2 6'(h'_2)\cdot x_1] \quad \text{———} \ (2)$$

continue,

$$\therefore \frac{\partial f}{\partial \omega'_{21}} = \omega_1^3 6'(h_1^2)\cdot[\omega_{11}^2 6'(h'_1)\cdot x_2] + \omega_2^3 6'(h_2^2)\cdot[\omega_{12}^2 6'(h'_1)\cdot x_2] \quad \text{———} \ (3)$$

And,

$$\therefore \frac{\partial f}{\partial \omega'_{22}} = \omega_1^3 6'(h_1^2)[\omega_{21}^2 6'(h'_2)\cdot x_2] + \omega_2^3 6'(h_2^2)[\omega_{22}^2 6'(h'_2)\cdot x_2] \quad \text{———} \ (4)$$

This (1), (2), (3), (4) are our final equation.