L^n $\quad y^n_1 \quad y^n_2 \quad \cdots \quad y^n_{k_n}$

$\vdots$

L^2 $\quad y^2_1 \quad y^2_2 \quad y^2_3 \quad \cdots \quad y^2_{k_2}$

L_1 $\quad y^1_1 \quad y^1_2 \quad y^1_3 \quad \cdots \quad y^1_{k_1}$

in one Layer output:

say from $y^1 \longrightarrow y^2$

$y^2_i = f(Z^2_i)$

$C_i^{(2)}$ = cost fn at 2nd layer ith neuron

$W_{ij}^{n,n-1}$ = coeff at transition between $(n-1) \longrightarrow n$ th layer and neurons

y_i^n, y_j^{n-1}

$$\frac{\partial z_i^{(2)}}{\partial y_1^1} = W_{i1}^{2,1}$$

$$z_i^{(2)} = W_{i1}^{2,1} y_1^1 + W_{i2}^{2,1} y_2^1 + W_{i3}^{2,1} y_3^1 + \cdots + W_{iK_1}^{2,1} y_{K_1}^1$$

$$= \sum_{j=1}^{K_1} W_{ij}^{2,1} \cdot y_j^1$$

$$y_i^{(2)} = f\left(z_i^{(2)}\right) \longrightarrow \text{from the code}$$

The true fn: $F\left(z_i^{(2)}\right)$

The cost fn:

$$C_i^{(2)} = \frac{1}{2} \left\langle \left(f\left(z_i^{(2)}\right) - F\left(z_i^2\right) \right)^2 \right\rangle$$

$$= \frac{1}{2} \left\langle \left(f\left(\sum_{j=1}^{K_1} W_{ij}^{2,1} y_j^1 \right) - F\left(z_i^2\right) \right)^2 \right\rangle$$

$$\frac{\partial z_i^n}{\partial y_k^{n-1}} = W_{ik}^{n,n-1} \qquad \frac{\partial z_i^n}{\partial W_{ij}^{n,n-1}} = y_j^{n-1}$$

Scanned with CamScanner

i.e $$\frac{\partial C_i^{(2)}}{\partial W_{i1}^{2,1}} = \left\langle \left(f\left(\sum w_{ij}^{2,1} y_j^1\right) - F\left(z_i^2\right)\right) \times f'\left(z_i^{(2)}\right) \times \frac{\partial z_i^{(2)}}{\partial W_{i1}^{2,1}} \right\rangle$$

$$= \left\langle \left(f\left(z_i^2\right) - F\left(z_i^2\right)\right) \times f'\left(z_i^{(2)}\right) \times y_i^1 \right\rangle$$

General:

$$\frac{\partial C_i^{n,n-1}}{\partial W_{ij}^{n,n-1}} = \left\langle \left(f\left(z_i^n\right) - F\left(z_i^{(n)}\right)\right) \times f'\left(z_i^n\right) \times y_j^{n-1} \right\rangle$$

Now for multilayer transition:

~~Say:~~ $\cancel{\frac{\partial C_i^{n,n-1}}{\partial W}}$ The cost fn is:

$$C_i^{n,1} = \left\langle \left(f(z_i^n) - F(z_i^{(n)}) \right)^2 \right\rangle \frac{1}{2}$$

$$\frac{\partial C_i^{n,1}}{\partial W_{11}^{21}} = \left\langle \left(\cancel{2} f(z_i^n) - F(z_i^n) \right) \times f'(z_i^n) \times \frac{\partial z_i^n}{\partial W_{11}^{21}} \right\rangle$$

Now, $$\frac{\partial z_i^n}{\partial W_{11}^{21}} = \sum_k \frac{\partial z_i^n}{\partial y_k^{n-1}} \frac{\partial y_k^{n-1}}{\partial W_{11}^{21}}$$

$$= \sum_k W_{ik}^{n,n-1} f'(z_k^{n-1}) \times \frac{\partial z_k^{n-1}}{\partial W_{11}^{21}}$$

$$\left(\because y_k^{n-1} = f(z_k^{n-1}) \right)$$

Again apply the ith.

$$\frac{\partial Z_K^{n-1}}{\partial W_{11}^{21}} = \sum_{m} W_{Km}^{n-1,n-2} f'\left(Z_m^{n-2}\right) \cdot \frac{\partial Z_m^{n-2}}{\partial W_{11}^{21}}$$

$\vdots$

And so on.

i.e:

$$\frac{\partial C_i^{n,n-1}}{\partial W^{2,1}_{1,1}} = \Big\langle \Big(f(Z_i^n) - f(Z_i^n) \Big) \times f'(Z_i^n) \times$$

$$\sum_{K_1} \underbrace{W^{n,n-1}_{iK_1} f'(Z^{n-1}_{K_1})}_{M_{iK_1}} \times$$

$$\sum_{K_2} \underbrace{W^{n-1,n-2}_{K_1K_2} f'(Z^{(n-2)}_{K_2})}_{M_{K_1K_2}} \times$$

$$\sum_{K_3} \underbrace{W^{n-2,n-3}_{K_2K_3} f'(Z^{n-3}_{K_3})}_{M_{K_2K_3}} \times$$

$$\cdots\cdots$$

$$\sum W^{3,2}_{\alpha\beta} f'(Z^2_{K\beta}) \times \underbrace{\frac{\partial Z^2_\beta}{\partial W^{21}_{11}}}_{= y_1'} \Big\rangle$$

i.e $\dfrac{\partial C_i^{n,n-1}}{\partial W_{11}^{21}} =$

$$\left\langle \left(f(Z_i^n) - F(Z_i^n)\right) f'(Z_i^n) \times \right.$$

$$\sum_{K_{n-1}}^{L_{n-1}} \cdots \sum_{K_2}^{L_2} M_{iK_{n-1}} M_{K_{n-1}K_{n-2}} \cdots M_{K_3K_2} \times J_{11}$$

With $M_{K_iK_j} = W_{K_iK_j}^{ij} \times f'(Z_{K_j}^{j})$

L_n y_1^n y_2^n y_3^n ... y_i^n $y_{k_n}^n$

$w_{i\cdot}^{n,n-1}$

L_{n-1} y_1^{n-1} y_2^{n-2} ... $y_{\cdot}^{n-1}$ ... $y_{k_{n-1}}^{n-2}$

⋮

L_m y_1^m y_2^m y_3^m ... y_j^m ... $y_{k_m}^m$

$w_{jk}^{m,m-1}$

L_{m-1} y_1^{m-1} y_2^{m-1} y_3^{m-1} ... y_k^{m-1} ... $y_{k_{m-1}}^{m-1}$

⋮

L_2 y_1^2 y_2^2 ... $y_{k_2}^2$

L_1 y_1^1 y_2^1 y_3^1 ... $y_{k_1}^1$

in general:

$$\frac{\partial C_i^{m,m-1}}{\partial W_{jk}^{m,m-1}} = \left\langle \left(f(z_i^n) - F(z_i^n) \right) \cdot f'(z_i^n) \right.$$

$$\times \sum_{p_{m-1}}^{L_{m-1}} \sum_{p_{m-2}}^{L_{m-2}} \cdots \sum_{p_m}^{L_m} \left(M_{i p_{m-1}} M_{p_{m-1} p_{m-2}} \cdots \right.$$

$$M_{p_{m-2} p_{m-3}} \cdots M_{p_{m+1} p_m} \times y_k^{m-1}$$