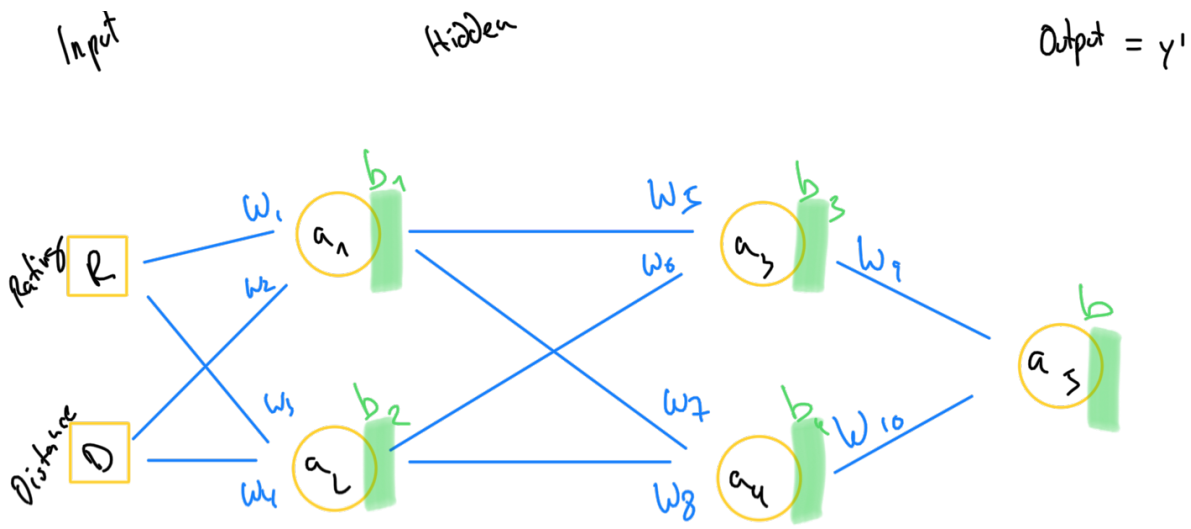
Neural network back propagation

Exercise: calculate a simple net by hand

Data

Rating	Distance	Price
3.5	0.2	157
2.4	6.5	92
3.6	19.8	207



feed forward

$$N_5(N_3(N_1(R,D), N_2(R,D)), N_4(N_1(R,D), N_2($$

$$a_1 = \text{ReLu}(z_1) \qquad z_1 = w_1 \cdot R + w_2 \cdot D + b_1$$

$$a_2 = \text{ReLu}(z_2) \qquad z_2 = w_3 \cdot R + w_4 \cdot D + b_2$$

$$a_3 = \text{ReLu}(z_3) \qquad z_3 = w_5 \cdot a_1 + w_6 \cdot a_2 + b_3$$

$$a_4 = \text{ReLu}(z_4) \qquad z_4 = w_7 \cdot a_1 + w_8 \cdot a_2 + b_4$$

$$a_5 = \text{ReLu}(z_5) \qquad z_5 = w_9 \cdot a_3 + w_{10} \cdot a_4 + b_5$$

Rating	Distance	Price

3.5	0.2	157
2.4	6.5	92
3.6	19.8	207

Epoch 1

W1	W2	W3	W4	W5	W6	W7	W8	W9	W10
0.3	0.4	0.5	2	0.8	4	0.1	0.4	0.9	3

B1		B2		B3		B4		B5	
2		6		0		0.5		1	
Z1	A1	Z2	A2	Z3	A3	Z4	A4	Z5	A5
3.13	3.13	8.15	8.15	35.104	35.104	4.073	4.073	44.8126	44.8126

Cost function sum of squares (A5 - Price)^2

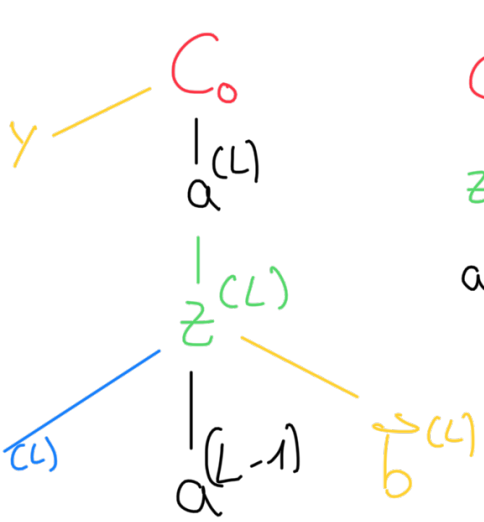
Index	C	dC/dW1
1	12586.0127	

$$\frac{\partial C}{\partial w_9} = \frac{\partial C}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_5} \cdot \frac{\partial z_5}{\partial w_9}$$

$$\frac{\partial C}{\partial w_{10}} = \frac{\partial C}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_5} \cdot \frac{\partial z_5}{\partial w_{10}}$$

$$\frac{\partial C}{\partial b_5} = \frac{\partial c}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_5} \cdot \frac{\partial z_5}{\partial b_5}$$

$$\frac{\partial c}{\partial w_8} = \frac{\partial c}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_5} \cdot \frac{\partial z_5}{\partial w_8}$$



$$C_0 = (a^{(L)} - y)^2$$

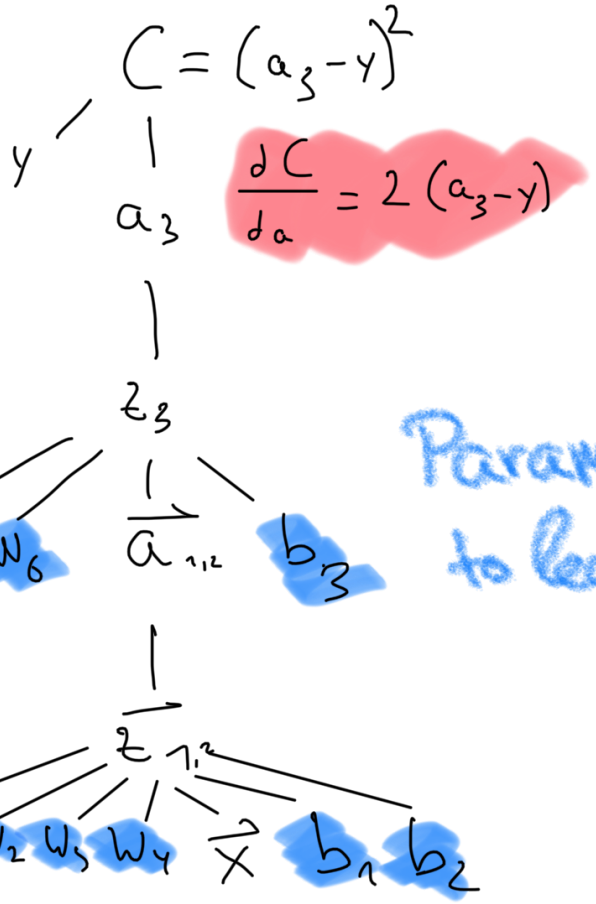
$$z^{(L)} = \vec{w} \cdot a^{(L-1)} + \vec{b}$$

$$a^{(L)} = \sigma(z^{(L)})$$

$$\sigma\left(\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \vec{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \vec{a}$$

$$\sigma\left(\begin{bmatrix} w_5 & w_6 \end{bmatrix} \cdot \vec{a} + b_3\right) = a_3$$

$$\frac{\partial C_0}{\partial a^{(L)}} = 2(a^{(L)} - y)$$



$$\frac{\partial C}{\partial a} = 2(a_3 - y)$$

$$\frac{\partial a_3}{\partial z_3} = \sigma'(z_3)$$

$$\frac{\partial z_3}{\partial w_5} = w_5 \cdot a_1 \cdot w_6 \cdot a_2 + b_3 = a_1$$

$$\frac{\partial z_3}{\partial w_6} = w_5 \cdot a_1 \cdot w_6 \cdot a_2 + b_3 = a_2$$

$$\frac{\partial z_3}{\partial b_3} = w_5 \cdot a_1 \cdot w_6 \cdot a_2 + b_3 = 1$$

Parameters to learn

$$\frac{\partial C}{\partial w_6} = \frac{\partial C}{\partial a_3}\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial w_6}$$

$$\frac{\partial C}{\partial w_5} = \frac{\partial C}{\partial a_3}\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial w_5}$$

$$\frac{\partial C}{\partial b_3} = \frac{\partial C}{\partial a_3}\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial b_3}$$

$$\frac{\partial z_3}{\partial a_1} = w_5 \qquad \frac{\partial z_3}{\partial a_2} = w_6$$

$$\frac{\partial a_1}{\partial z_1} = \sigma'(z_1) \qquad \frac{\partial a_1}{\partial z_2} = \sigma'(z_2)$$

$$\frac{\partial z_1}{\partial w_2} = w_1 \cdot x_1 + w_2 x_2 + b_1 = x_2$$

$$\frac{\partial z_1}{\partial w_1} = x_1 \qquad \frac{\partial z_2}{w_3} = x_1 \qquad \frac{\partial z_2}{\partial w_4} = x_2$$

$$\frac{\partial z_1}{\partial w_4} =$$

$$\frac{\partial C}{\partial w_4} = \frac{\partial C}{\partial a_3}\frac{\partial a_3}{\partial z_3}\left(\frac{\partial z_3}{\partial a_1}\frac{\partial a_1}{\partial z_2}\frac{\partial z_2}{\partial w_4} + \frac{\partial z_3}{\partial a_2}\frac{\partial a_1}{\partial z_2}\frac{\partial z_2}{\partial w_4}\right)$$

$$= 2(a_3 - y)\,\sigma'(z_3)\left(w_5\sigma'(z_1)x_2 + w_6\sigma'(z_2)x_2\right)$$