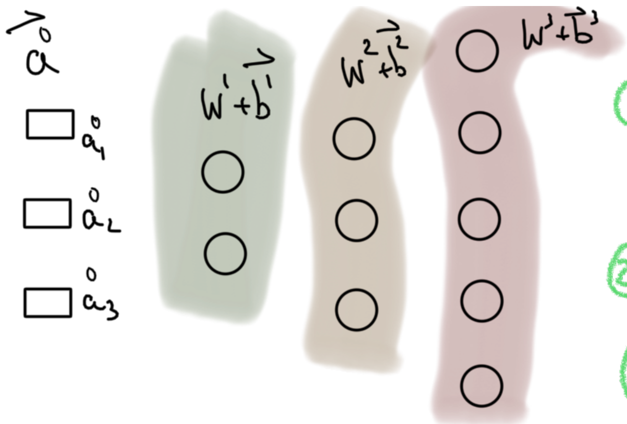
Back propagation 3

New approach



① $C = \sum_{m=0}^{3} (\vec a^3_m - \vec y_m)^2$

② $\vec a^L = \sigma(\vec z^L)$

③ $z^L = W^L a^{L-1} + b^L$

$$W^1 = \begin{bmatrix} w^1_{11} & w^1_{21} & w^1_{31} \\ w^1_{21} & w^1_{22} & w^1_{32} \end{bmatrix} \quad \vec b^1 = \begin{bmatrix} b^1_1 \\ b^1_2 \end{bmatrix} \quad \vec a^1 = \begin{bmatrix} a^1_1 \\ a^1_2 \end{bmatrix} = \sigma(z^1) \quad z^1 = W^1 a^0 + b^1$$

$$W^2 = \begin{bmatrix} w^2_{11} & w^L_{12} \\ w^L_{21} & w^L_{22} \\ w^2_{31} & w^L_{32} \end{bmatrix} \quad \vec b^2 = \begin{bmatrix} b^2_1 \\ b^2_2 \\ b^2_3 \end{bmatrix} \quad \vec a^2 = \begin{bmatrix} a^2_1 \\ a^2_2 \\ a^2_3 \end{bmatrix} = \sigma(z^L) \quad z^2 = W^2 a^1 + b^2$$

$$W^3 = \begin{bmatrix} w^3_{11} & w^3_{12} & w^3_{13} \\ w^3_{21} & w^3_{22} & w^3_{23} \\ w^3_{31} & w^3_{32} & w^3_{33} \\ w^3_{41} & w^3_{42} & w^3_{43} \\ w^3_{51} & w^3_{52} & w^3_{53} \end{bmatrix} \quad \vec b^3 = \begin{bmatrix} b^3_1 \\ b^3_2 \\ b^3_3 \\ b^3_4 \\ b^3_5 \end{bmatrix} \quad \vec a^3 = \begin{bmatrix} a^3_1 \\ a^3_2 \\ a^3_3 \\ a^3_4 \\ a^3_5 \end{bmatrix} = \sigma(z^3) \quad z^3 = W^3 a^2 + b^3$$

Each Layer must compute change to $W, \vec b, \vec a$ with respect to Error.
Last Layer is special.

$$\frac{\partial C}{\partial a^L_k} = 2(a^L_k - y_k) \qquad \frac{\partial C}{\partial w^L_{mn}} = \frac{\partial C}{\partial a^L_m} \frac{\partial a^L_m}{\partial z^L_m} \frac{\partial z^L_m}{\partial w^L_{mn}} = 2(a^L_m - y_m)\, \sigma'(z^L_m)\, a^{L-1}_n$$

$$\frac{\partial C}{\partial b^L_m} = \frac{\partial C}{\partial a^L_m} \frac{\partial a^L_m}{\partial z^L_m} \frac{\partial z^L_m}{\partial b^L_m} = 2(a^L_m - y_m)\, \sigma'(z^L_m)$$

Last Layer
$L=3$

Layer $L-1$ with $L = 2, \dots N$

CACHE LOOKUP

$$\frac{\partial C}{\cdots} = \sum^{N_L} \frac{\partial C}{\cdots} \frac{\partial a^L_m}{\cdots} \frac{\partial z^L_m}{\cdots} = \sum^{N_L} \left(\partial C\right) \sigma'(z^L_m)\, w^L \cdots$$

$$\frac{\partial C}{d a_n^{L-1}} = \sum_{m=1} \frac{\partial C}{\partial a_m^L} \frac{}{\partial z_m^L} \frac{}{\partial a_N^{L-1}} = \sum_{m=1} \boxed{\frac{}{\partial a_m^L}} \sigma'(z_m) w_{mn}$$

$$\frac{\partial C}{d W_{mn}^{L-1}} = \frac{\partial C}{d a_m^{L-1}} \frac{d a_m^{L-1}}{d z_m^{L-1}} \frac{d z_m^{L-1}}{d W_{mn}^{L-1}} = \boxed{\frac{\partial C}{d a_m^{L-1}}} \sigma'(z_m^{L-1})\, a_n^{L-2}$$

Not the same as above

$$\frac{\partial C}{d b_m^{L-1}} = \frac{\partial C}{d a_m^{L-1}} \frac{d a_m^{L-1}}{d z_m^{L-1}} \frac{d z_m^{L-1}}{d b_m^{L-1}} = \boxed{\frac{\partial C}{d a_m^{L-1}}} \sigma'(z_m^{L-1})$$