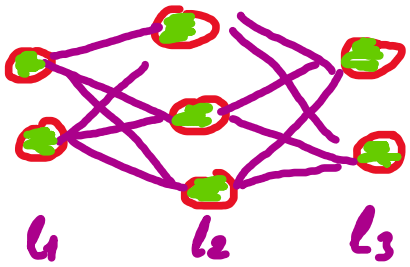
Forward and Backward Propagation

Friday, 19 February 2021 19:45

example neural network:



forward propagation:

$\sigma = activation\ function$

$$\sigma\left(w^{l3-l2}_{(2,3)} \cdot \sigma\left(w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}\right) + b^{l3}_{(2,1)}\right)$$

backward propagation:

$$l2 - l1 \Rightarrow z^{l2}_{(3,1)} = w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}$$

$$l2 - l1 \Rightarrow a^{l2}_{(3,1)} = \sigma\left(z^{l2}_{(3,1)}\right)$$

$$l3 - l2 \Rightarrow z^{l3}_{(2,1)} = w^{l3-l2}_{(2,3)} \cdot a^{l2}_{(3,1)} + b^{l3}_{(2,1)}$$

$$l3 - l2 \Rightarrow a^{l3}_{(2,1)} = \sigma\left(z^{l3}_{(2,1)}\right)$$

$$C\left(a^{l3}_{(2,1)}\right) = loss$$

$$\frac{d\,C}{d\,a^{l3}_{(2,1)}} \cdot \frac{d\,a^{l3}_{(2,1)}}{d\,z^{l3}_{(2,1)}} \cdot \frac{d\,z^{l3}_{(2,1)}}{d\,a^{l2}_{(3,1)}} \cdot \frac{d\,a^{l2}_{(3,1)}}{d\,z^{l2}_{(3,1)}} \cdot \frac{d\,z^{l2}_{(3,1)}}{d\,v^{l1}_{(2,1)}}$$

$$\delta^3 = \frac{d\,C}{d\,a^{l3}_{(2,1)}} \cdot \frac{d\,a^{l3}_{(2,1)}}{d\,z^{l3}_{(2,1)}} = \frac{d\,C}{d\,z^{l3}_{(2,1)}}$$

$$\delta^2 = \delta^3 \cdot \frac{d\,z^{l3}_{(2,1)}}{d\,a^{l2}_{(3,1)}} \cdot \frac{d\,a^{l2}_{(3,1)}}{d\,z^{l2}_{(3,1)}} = \frac{d\,C}{d\,z^{l2}_{(3,1)}}$$

$$\begin{aligned}\frac{d\,C}{d\,w^{l3-l2}_{(2,3)}} &= \frac{d\,C}{d\ z^{l3}_{(2,1)}} \cdot \frac{d\ z^{l3}_{(2,1)}}{d\,w^{l3-l2}_{(2,3)}} \\ &= \delta^3 \cdot \frac{w^{l3-l2}_{(2,3)} \cdot a^{l2}_{(3,1)} + b^{l3}_{(2,1)}}{d\,w^{l3-l2}_{(2,3)}} \\ &= \delta^3 \cdot \frac{w^{l3-l2}_{(2,3)} \cdot a^{l2}_{(3,1)} + b^{l3}_{(2,1)}}{d\,w^{l3-l2}_{(2,3)}} \\ &= \delta^3 \cdot w^{l3-l2}_{(2,3)}\end{aligned}$$

$$w^{l3-l2}_{(2,3)} \leftarrow w^{l3-l2}_{(2,3)} - \eta \frac{d\,C}{d\,w^{l3-l2}_{(2,3)}}$$

$$
\begin{aligned}
\frac{d\,C}{d\,b^{l3}_{(2,1)}} &= \frac{d\,C}{d\ \ z^{l3}_{(2,1)}} \cdot \frac{d\ \ z^{l3}_{(2,1)}}{d\,b^{l3}_{(2,1)}} \\
&= \delta^3 \cdot \frac{w^{l3-l2}_{(2,3)} \cdot a^{l2}_{(3,1)} + b^{l3}_{(2,1)}}{d\,b^{l3}_{(2,1)}} \\
&= \delta^3 \cdot \frac{w^{l3-l2}_{(2,3)} \cdot a^{l2}_{(3,1)} + b^{l3}_{(2,1)}}{d\,b^{l3}_{(2,1)}} \\
&= \delta^3
\end{aligned}
$$

$$
b^{l3}_{(2,1)} \leftarrow b^{l3}_{(2,1)} - \eta \frac{d\,C}{d\,b^{l3}_{(2,1)}}
$$

$$
\begin{aligned}
\frac{d\,C}{d\,w^{l2-l1}_{(3,2)}} &= \frac{d\,C}{d\ \ z^{l2}_{(3,1)}} \cdot \frac{d\ \ z^{l2}_{(3,1)}}{d\,w^{l2-l1}_{(3,2)}} \\
&= \delta^2 \cdot \frac{w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}}{d\,w^{l2-l1}_{(3,2)}} \\
&= \delta^2 \cdot \frac{w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}}{d\,w^{l2-l1}_{(3,2)}} \\
&= \delta^2 \cdot w^{l2-l1}_{(3,2)}
\end{aligned}
$$

$$
w^{l2-l1}_{(3,2)} \leftarrow w^{l2-l1}_{(3,2)} - \eta \frac{d\,C}{d\,w^{l2-l1}_{(3,2)}}
$$

$$
\begin{aligned}
\frac{d\,C}{d\,b^{l2}_{(3,1)}} &= \frac{d\,C}{d\ \ z^{l2}_{(3,1)}} \cdot \frac{d\ \ z^{l2}_{(3,1)}}{d\,b^{l2}_{(3,1)}} \\
&= \delta^2 \cdot \frac{w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}}{d\,b^{l2}_{(3,1)}} \\
&= \delta^2 \cdot \frac{w^{l2-l1}_{(3,2)} \cdot x^{l1}_{(2,1)} + b^{l2}_{(3,1)}}{d\,b^{l2}_{(3,1)}} \\
&= \delta^2
\end{aligned}
$$

$$
b^{l2}_{(3,1)} \leftarrow b^{l2}_{(3,1)} - \eta \frac{d\,C}{d\,b^{l2}_{(3,1)}}
$$