The Gradient

Benjamin S. Knight

March 6, 2017

$$
\begin{aligned}
\frac{\delta E}{\delta y} &= \frac{\delta}{\delta y}\Big(1/2\sum(y^{\mu}-\hat{y}^{\mu})^2\Big) \\
&= 2(1/2)\sum(y^{\mu}-\hat{y}^{\mu})^1 \\
&= \sum(y^{\mu}-\hat{y}^{\mu}) \qquad \rightarrow \qquad y^{\mu}-\hat{y}^{\mu} \\
&\qquad\qquad\qquad\qquad\qquad\quad \downarrow \\
&\qquad\qquad\qquad\qquad\qquad 1-0.21 \\
&\qquad\qquad\qquad\qquad\qquad\quad \downarrow \\
&\qquad\qquad\qquad\qquad\qquad\ \ 0.79
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
\frac{\delta}{\delta x}\sigma(x) = \sigma(x)&(1-\sigma(x)) \\
&\downarrow \\
\frac{\delta E}{\delta \sigma} &= (y-\hat{y})\sigma(x)(1-\sigma(x)) \\
&= (1-0.21)\Big(\frac{1}{1+e^{-(-1.28)}}\Big)\Big(1-\frac{1}{1+e^{-(-1.28)}}\Big) \\
&= (1-0.21)(0.21)(1-0.21) \\
&= (0.79)(0.21)(0.79) \\
&= 0.131
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
f(o1) = h_1(W_{h1}^{o1}) + h_2(W_{h2}^{o1}) \qquad \rightarrow \qquad \frac{\partial o_1}{\partial h_1} &= W_{h1}^{o1} = 0.7 \\
\frac{\partial o_1}{\partial h_2} &= W_{h2}^{o1} = -0.6
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
f(h1) = x_1(W_{x1}^{h1}) + x_2(W_{x2}^{h1}) \qquad \rightarrow \qquad \frac{\partial h_1}{\partial x_1} &= W_{x1}^{h1} = 0.3 \\
\frac{\partial h_1}{\partial x_2} &= W_{x2}^{h1} = 0.8
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
f(h2) = x_1(W_{x1}^{h2}) + x_2(W_{x2}^{h2}) \qquad \rightarrow \qquad \frac{\partial h_2}{\partial x_1} &= W_{x1}^{h2} = -0.1 \\
\frac{\partial h_2}{\partial x_2} &= W_{x2}^{h2} = -0.5
\end{aligned} \tag{5}
$$

Weight recalculations:

$$
\begin{aligned}
f(h2) = x_1(W_{x1}^{h2}) + x_2(W_{x2}^{h2}) \qquad \rightarrow \qquad \frac{\partial h_2}{\partial x_1} &= W_{x1}^{h2} = -0.1 \\
\frac{\partial h_2}{\partial x_2} &= W_{x2}^{h2} = -0.5
\end{aligned} \tag{6}
$$

$$1/2\sum(y^\mu - \hat{y}^\mu)^2$$

$$y^\mu - \hat{y}^\mu$$

$$\sigma(x)$$

$$\sigma(x)(1-\sigma(x))$$

$$h_1(W_{h1}^{o1}) + h_2(W_{h2}^{o1}) \to \frac{\partial o_1}{\partial h_1}$$