1 Neural Network

1.1 One Step further of Logistic Regression

Model Description
$z^{[1](i)}$ means first layer, ith sample

$$z^{[1](i)} = W^{[1]}X^{(i)} + b^{[1](i)}$$

in matrix form

$$\begin{pmatrix} z_1^{[1](1)} & z_1^{[1](2)} & ... & z_1^{[1](m)} \\ ... & ... & ... & ... \\ z_4^{[1](1)} & z_4^{[1](2)} & ... & z_4^{[1](m)} \end{pmatrix} = \begin{pmatrix} w_{11}^{[1]} & w_{12}^{[1]} \\ w_{21}^{[1]} & w_{22}^{[1]} \end{pmatrix} \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & ... & x_1^{(m)} \\ x_2^{(1)} & x_2^{(2)} & ... & x_2^{(m)} \end{pmatrix}$$

$$a^{[1](i)} = tanh(z^{[1](i)})$$

In matrix notation

$$\begin{pmatrix} a_1^{[1](1)} & a_1^{[1](2)} & ... & a_1^{[1](m)} \\ ... & ... & ... & ... \\ a_4^{[1](1)} & a_4^{[1](2)} & ... & a_4^{[1](m)} \end{pmatrix} = tanh \begin{pmatrix} z_1^{[1](1)} & z_1^{[1](2)} & ... & z_1^{[1](m)} \\ ... & ... & ... & ... \\ z_4^{[1](1)} & z_4^{[1](2)} & ... & z_4^{[1](m)} \end{pmatrix}$$

$$z^{[2](i)} = W^{[2]}X^{(i)} + b^{[2](i)}$$

in matrix form

$$\begin{pmatrix} z_1^{[2](1)} & z_1^{2](2)} & ... & z_1^{[2](m)} \end{pmatrix} = \begin{pmatrix} w_1^{[2]} & ... & w_4^{[2]} \end{pmatrix} \begin{pmatrix} a_1^{[1](1)} & a_1^{[1](2)} & ... & a_1^{[1](m)} \\ ... & ... & ... & ... \\ a_4^{[1](1)} & a_4^{[1](2)} & ... & a_4^{[1](m)} \end{pmatrix}$$

$$a^{[2](i)} = sigmoid(z^{[2](i)})$$

In matrix notation

$$\begin{pmatrix} a_1^{[2](1)} & a_1^{[2](2)} & ... & a_1^{[2](m)} \end{pmatrix} = sigmoid \begin{pmatrix} z_1^{[2](1)} & z_1^{[2](2)} & ... & z_1^{[2](m)} \end{pmatrix}$$

Gradient Calculation We now derive the gradient of the model above.
1) Cost function

$$J = -\frac{1}{m}\sum_i^m [y^{(i)}log(a^{[2](i)})(1-y^{(i)})log(1-a^{[2](i)})]$$

(i) denotes sample index [j] denotes layer index The sequence of calculating the gradient is

J - $a^{[2]}_{(i)}$ - $z^{[2]}_{(i)}$ - $W^{[2]}_k b^{[2]}$ - chain rule ...
2)

$$
\begin{aligned}
\frac{\partial J}{\partial a^{[2]}_{(j)}} &= \frac{\partial}{\partial a^{[2]}_{(j)}}(-\frac{1}{m}\sum_i [y_{(i)} log(a^{[2]}_{(i)}) + (1-y_{(i)}) log(1-a^{[2]}_{(i)}]) \text{ only i = j survives} \\
&= \frac{\partial}{\partial a^{[2]}_{(j)}}(-\frac{1}{m}[y_{(j)} log(a^{[2]}_{(j)}) + (1-y_{(j)}) log(1-a^{[2]}_{(j)})] \\
&= -\frac{1}{m}[\frac{y^{(j)}}{a^{[2]}_{(j)}} - \frac{(1-y^{(j)})}{(1-a^{[2]}_{(j)})}]
\end{aligned}
$$

3)

$$
\begin{aligned}
&\frac{\partial a}{\partial z} \\
=&\frac{\partial}{\partial z}(\frac{e^z}{e^z+1}) \\
=&\frac{1}{(e^z+1)^2}(e^z(e^z+1) - e^z e^z) \\
=&\frac{1}{(e^z+1)^2} e^z \\
=&\frac{e^z}{e^z+1}\frac{1}{e^z+1} \\
=&\frac{e^z}{e^z+1}\frac{e^z+1-e^z}{e^z+1} \\
=&g(z)(1-g(z))
\end{aligned}
$$

4)

$$
\begin{aligned}
&\frac{\partial J}{\partial z^{[2]}_j} \\
=&\frac{\partial J}{\partial a^{[2]}_{(j)}}\frac{\partial a^{[2]}_{(j)}}{\partial z^{[2]}_{(j)}} \\
=&-\frac{1}{m}[\frac{y^{(j)}}{a^{[2]}_{(j)}} - \frac{(1-y^{(j)})}{(1-a^{[2]}_{(j)})}] a^{[2]}_{(j)}(1-a^{[2]}_{(j)}) \\
=&-\frac{1}{m}(y_{(j)} - y_{(j)}a^{[2]}_{(j)} - a^{[2]}_{(j)} + y_j a^{[2]}_{(j)}) \\
=&-\frac{1}{m}(y_{(j)}a^{[2]}_{(j)})
\end{aligned}
$$

matrix form

$$
\begin{aligned}
&\frac{\partial J}{\partial Z^{[2]}}\\
=&(\frac{\partial J}{\partial z_1^{[2]}}, \frac{\partial J}{\partial z_2^{[2]}}, ..., \frac{\partial J}{\partial z_m^{[2]}})\\
=&-\frac{1}{m}(y_{(1)}a_{(1)}^{[2]}, y_{(2)}a_{(2)}^{[2]}, ..., y_{(m)}a_{(m)}^{[2]})\\
=&\frac{1}{m}(A^{[2]} - Y)
\end{aligned}
$$

5)

$$
\begin{pmatrix} z^{[2](1)} & z^{[2](2)} & ... & z^{[2](m)} \end{pmatrix} = \begin{pmatrix} w_1^{[2]} & w_2^{[2]} & ... & w_4^{[2]} \end{pmatrix} \begin{pmatrix} a_1^{[1](1)} & a_1^{[1](2)} & ... & a_1^{[1](m)} \\ ... & ... & ... & ... \\ a_4^{[2](1)} & a_4^{[1](2)} & ... & a_4^{[1](m)} \end{pmatrix} + (b[2], b[2], ..., b[2])
$$

$$
z_{[2](i)} = \sum_l w_l^{[2]} a_l^{[1](i)} + b^{[2]}
$$

$$
\frac{\partial z_i^{[2]}}{\partial w_k^{[2]}} = a_l^{[1](i)}
$$

$$
\frac{\partial z_i^{[2]}}{\partial b^{[2]}} = 1
$$

6)

$$
\begin{aligned}
&\frac{\partial J}{\partial w_k^{[2]}}\\
=&\sum_{(i)} \frac{\partial J}{\partial z_{(i)}^{[2]}} \frac{\partial z_{(i)}^{[2]}}{\partial w_k^{[2]}} \text{ chain rule in higher dimension}\\
=&\sum_i^m -\frac{1}{m}(y_{(i)} - a_{(i)}^{[2]}) a_k^{[1](i)}\\
=&\frac{1}{m} \begin{pmatrix} a_{(1)}^{[2]} - y_{(1)} & a_{(2)}^{[2]} - y_{(2)} & ... & a_{(m)}^{[2]} - y_{(m)} \end{pmatrix} \begin{pmatrix} a_k^{(1)[1]} \\ a_k^{(2)[1]} \\ ... \\ a_k^{(m)[1]} \end{pmatrix}\\
&\text{in matrix form}\\
=&\frac{1}{m}[A^{[2]} - Y](a_k^{[1]})^T
\end{aligned}
$$

$$\begin{pmatrix} \frac{\partial J}{\partial w_k^{[2]}}, & \frac{\partial J}{\partial w_k^{[2]}}, & \dots & \frac{\partial J}{\partial w_k^{[2]}} \end{pmatrix}$$

$$\frac{1}{m}\begin{pmatrix} a_{(1)}^{[2]} - y_{(1)} & a_{(2)}^{[2]} - y_{(2)} & \dots & a_{(m)}^{[2]} - y_{(m)} \end{pmatrix} \begin{pmatrix} a_1^{(1)[1]} & a_2^{(1)[1]} & \dots & a_4^{(1)[1]} \\ a_1^{(2)[1]} & a_2^{(2)[1]} & \dots & a_4^{(2)[1]} \\ & \dots & \dots & \\ a_1^{(m)[1]} & a_2^{(m)[1]} & \dots & a_4^{(m)[1]} \end{pmatrix}$$

in matrix form

$$=\frac{1}{m}[A^{[2]} - Y](A^{[1]})^T$$

7)

$$\begin{aligned} \frac{\partial J}{\partial z_l^{[1](j)}} = & \qquad \frac{\partial J}{\partial z^{[2](j)}} \frac{\partial z^{[2](j)}}{\partial a_l^{[1](j)}} \frac{\partial a_l^{[1](j)}}{\partial z_l^{[1](j)}} \\ = & \frac{\partial J}{\partial z^{[2](j)}} \frac{\partial \sum_{l=1}^{4} w_l^{[2]} a_l^{[1](j)}}{\partial a_l^{[1](j)}} g^{'}(z_l^{[1](j)}) \\ = & \frac{\partial J}{\partial z^{[2](j)}} w_l^{[2]} g^{'}(z_l^{[1](j)}) \end{aligned}$$

in matrix form

$$\begin{pmatrix} \frac{\partial J}{\partial z_1^{[1](1)}}, & \dots & \dots & \frac{\partial J}{\partial z_1^{[1](m)}} \\ \dots & \dots & \dots & \\ \frac{\partial J}{\partial z_4^{[1](1)}}, & \dots & \dots & \frac{\partial J}{\partial z_4^{[1](m)}} \end{pmatrix}$$

$$= \begin{pmatrix} w_l^{[2]} \frac{\partial J}{\partial z^{[2](1)}}, & \dots & \dots & w_l^{[2]} \frac{\partial J}{\partial z^{[2](m)}} \\ \dots & \dots & \dots & \\ , & \dots & \dots & \\ w_4^{[2]} \frac{\partial J}{\partial z^{[2](1)}}, & \dots & \dots & w_4^{[2]} \frac{\partial J}{\partial z^{[2](m)}} \end{pmatrix}$$

$$\circ \begin{pmatrix} g^{'}(z_l^{[1](1)}), & \dots & \dots & g^{'}(z_l^{[1](m)}) \\ \dots & \dots & \dots & \\ g^{'}(z_4^{[1](1)}), & \dots & \dots & g^{'}(z_4^{[1](m)}) \end{pmatrix}$$