

$$\textcircled{1} \quad \hat{y}^{(i)} = \sigma(z^{(i)}) \quad z^{(i)} = \underline{x}^{(i)T} \underline{w}$$

$$J(\underline{w}) = \frac{1}{N} \sum_{i=1}^N L(\hat{y}^{(i)}(\underline{w}), y^{(i)})$$

$$\frac{\partial J}{\partial \underline{w}} = \frac{1}{N} \sum_{i=1}^N \frac{\partial L}{\partial \underline{w}} = \frac{1}{N} \sum_{i=1}^N \frac{\partial L}{\partial \hat{y}^{(i)}} \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \frac{\partial z^{(i)}}{\partial \underline{w}}$$

$$L(\hat{y}^{(i)}, y^{(i)}) = -[y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{y}^{(i)})]$$

$$\frac{\partial L}{\partial \hat{y}^{(i)}} = -\frac{y^{(i)}}{\hat{y}^{(i)}} + \frac{1 - y^{(i)}}{1 - \hat{y}^{(i)}}$$

$$\frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} = \hat{y}^{(i)} (1 - \hat{y}^{(i)})$$

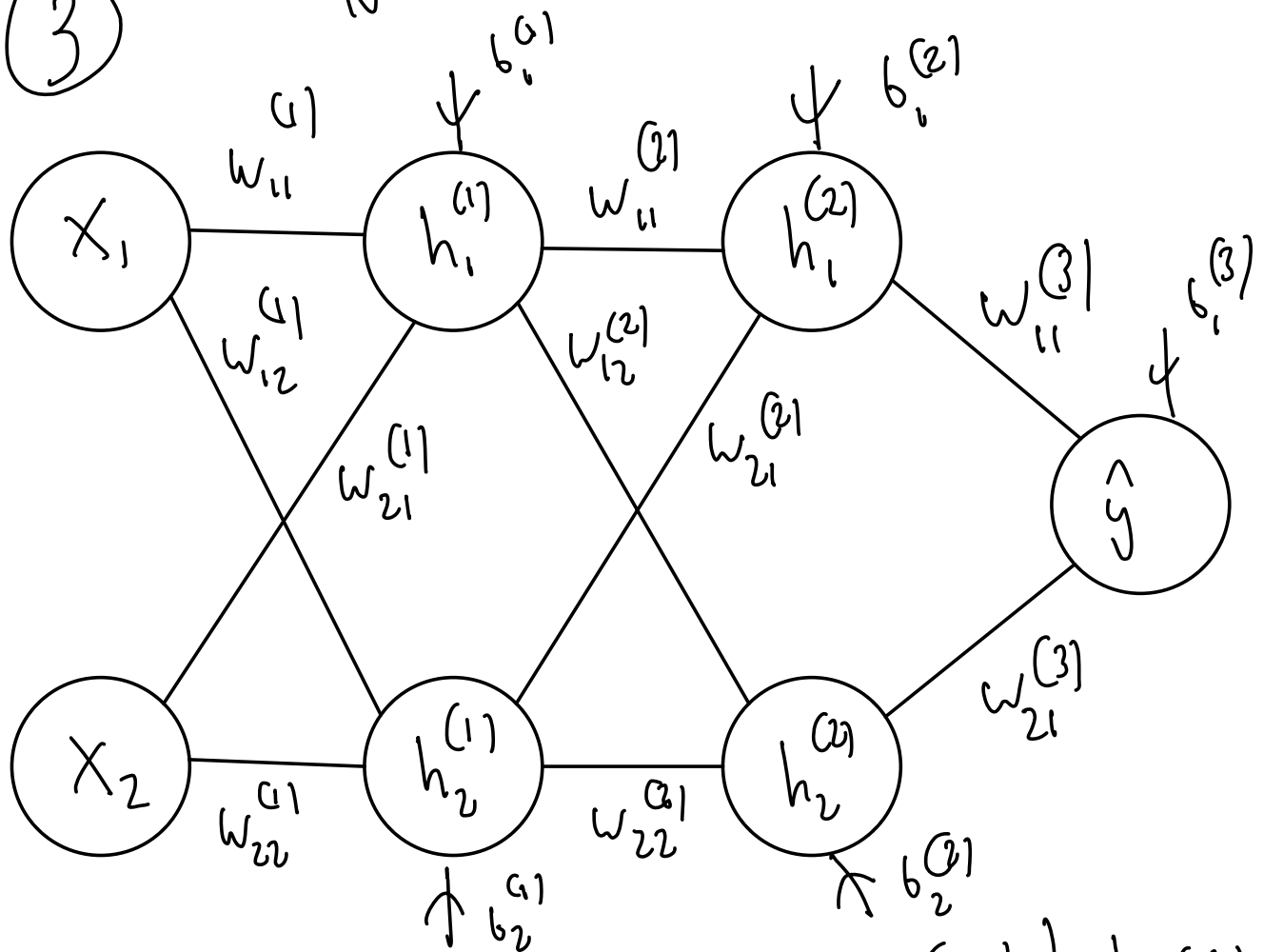
$$\begin{aligned} \frac{\partial L}{\partial \hat{y}^{(i)}} \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} &= -y^{(i)} (1 - \hat{y}^{(i)}) + \hat{y}^{(i)} (1 - y^{(i)}) \\ &= \hat{y}^{(i)} - y^{(i)} \end{aligned}$$

$$\frac{\partial J}{\partial \underline{w}} = \frac{1}{N} \sum_{i=1}^N \underline{x}^{(i)} (\hat{y}^{(i)} - y^{(i)})$$

$$= \frac{1}{N} \begin{pmatrix} 1 & 1 & \dots \\ \underline{x}^{(1)} & \underline{x}^{(2)} & \dots \\ 1 & 1 & \dots \end{pmatrix} \begin{pmatrix} \hat{y}^{(1)} - y^{(1)} \\ \hat{y}^{(2)} - y^{(2)} \\ \vdots \end{pmatrix}$$

$$= \frac{1}{N} \underline{X}^T (\underline{\hat{Y}} - \underline{Y})$$

(3)



$$\Delta_i^{(3)} = \frac{\partial \bar{J}}{\partial a_i^{(3)}}$$

$$\Delta_j^{(l)} = \left(\sum_k \Delta_k^{(l+1)} w_{jk}^{(l+1)} \right) \phi'(z_j^{(l)})$$

$$\Delta_1^{(2)} = \Delta_1^{(3)} w_{11}^{(3)} \phi'(z_1^{(2)})$$

$$\Delta_2^{(2)} = \Delta_1^{(3)} w_{21}^{(3)} \phi'(z_2^{(2)})$$

$$\Delta_1^{(1)} = \Delta_1^{(2)} w_{11}^{(2)} \Theta(z_1^{(1)}) \\ + \Delta_2^{(2)} w_{12}^{(2)} \Theta(z_1^{(1)})$$

$$\Delta_2^{(1)} = \Delta_1^{(2)} w_{21}^{(2)} \Theta(z_2^{(1)}) \\ + \Delta_2^{(2)} w_{22}^{(2)} \Theta(z_2^{(1)})$$

$$\frac{\partial \bar{J}}{\partial w_{kj}^{(1)}} = \Delta_j^{(1)} a_k^{(l-1)}$$

$$\frac{\partial \bar{J}}{\partial w_{11}^{(3)}} = \Delta_1^{(3)} a_1^{(2)}$$

$$\frac{\partial \bar{J}}{\partial w_{21}^{(3)}} = \Delta_1^{(3)} a_2^{(2)}$$

$$\frac{\partial \bar{J}}{\partial w_{11}^{(2)}} = \Delta_1^{(2)} a_1^{(1)}$$

$$\frac{\partial \bar{J}}{\partial w_{12}^{(2)}} = \Delta_2^{(2)} a_1^{(1)}$$

$$\frac{\partial \bar{J}}{\partial w_{21}^{(2)}} = \Delta_1^{(2)} a_2^{(1)}$$

$$\frac{\partial \bar{J}}{\partial w_{22}^{(2)}} = \Delta_2^{(2)} a_2^{(1)}$$

$$\frac{\partial \bar{J}}{\partial w_{11}^{(1)}} = \Delta_1^{(1)} x_1$$

$$\frac{\partial \bar{J}}{\partial w_{12}^{(1)}} = \Delta_2^{(1)} x_1$$

$$\frac{\partial \bar{J}}{\partial w_{21}^{(1)}} = \Delta_1^{(1)} x_2$$

$$\frac{\partial \bar{J}}{\partial w_{22}^{(1)}} = \Delta_2^{(1)} x_2$$