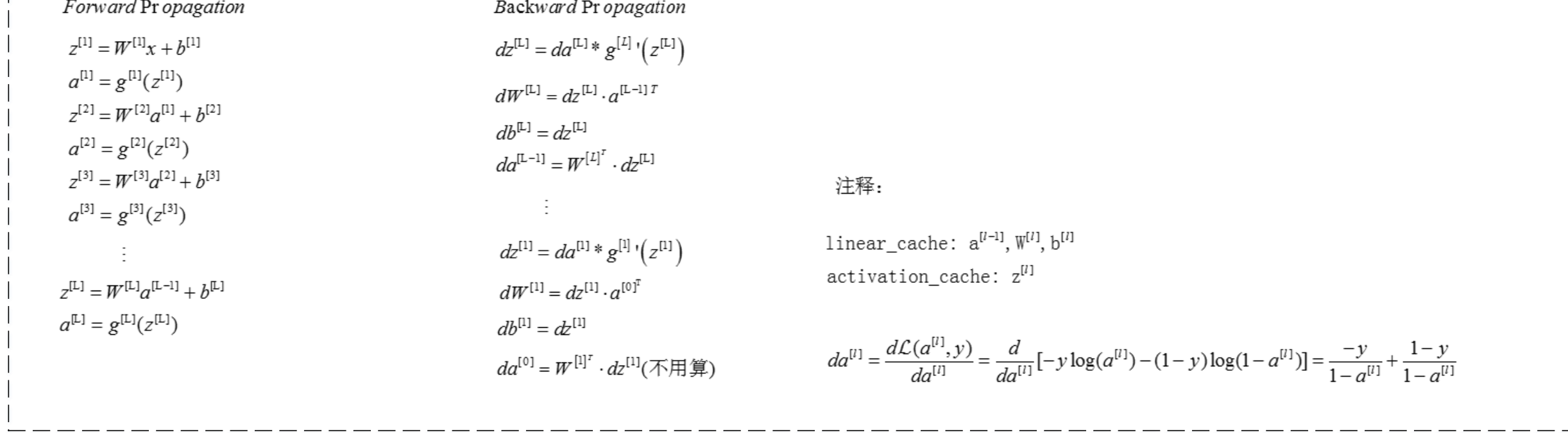
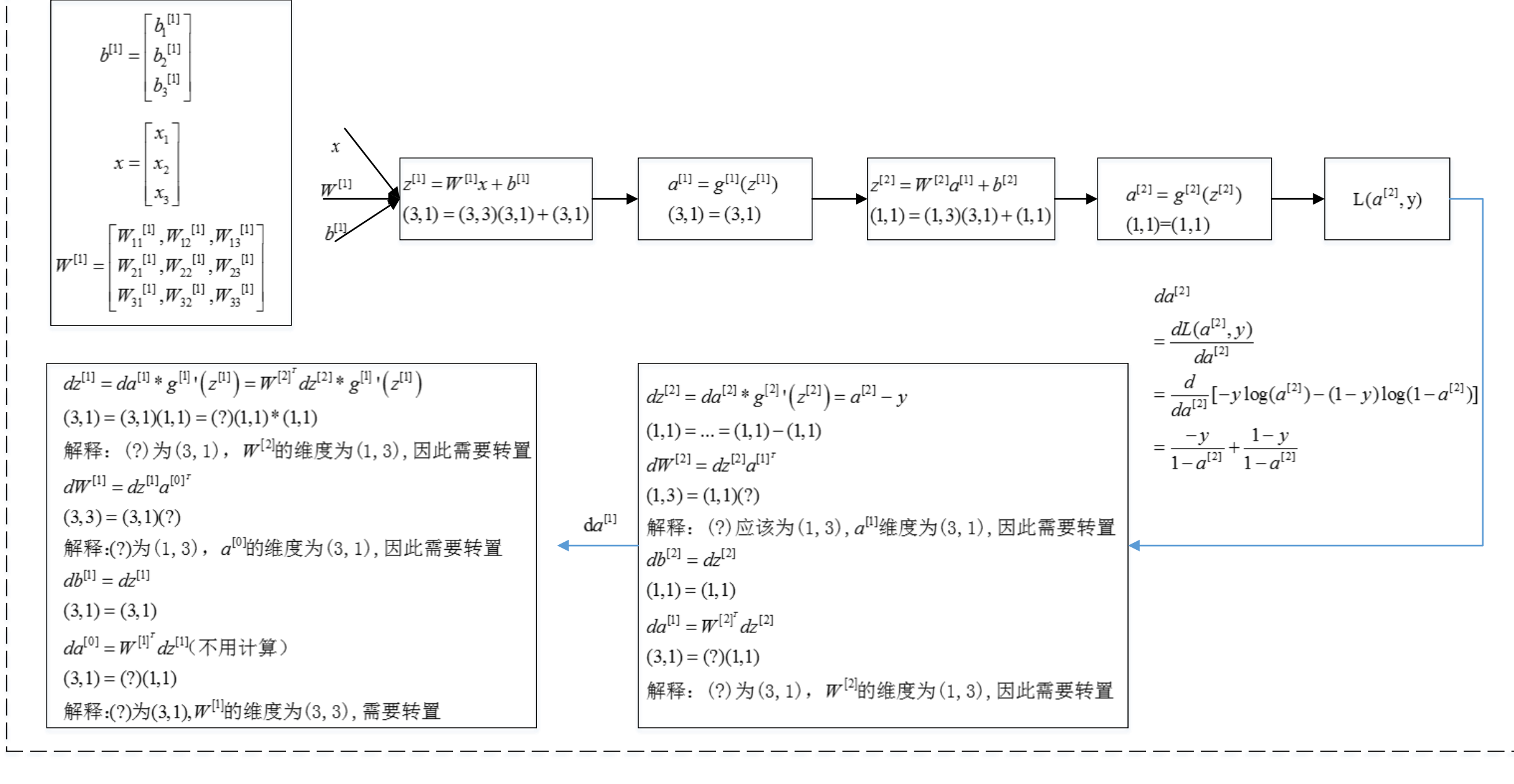


Author: 贺志尧  
Email:2282815808@qq.com

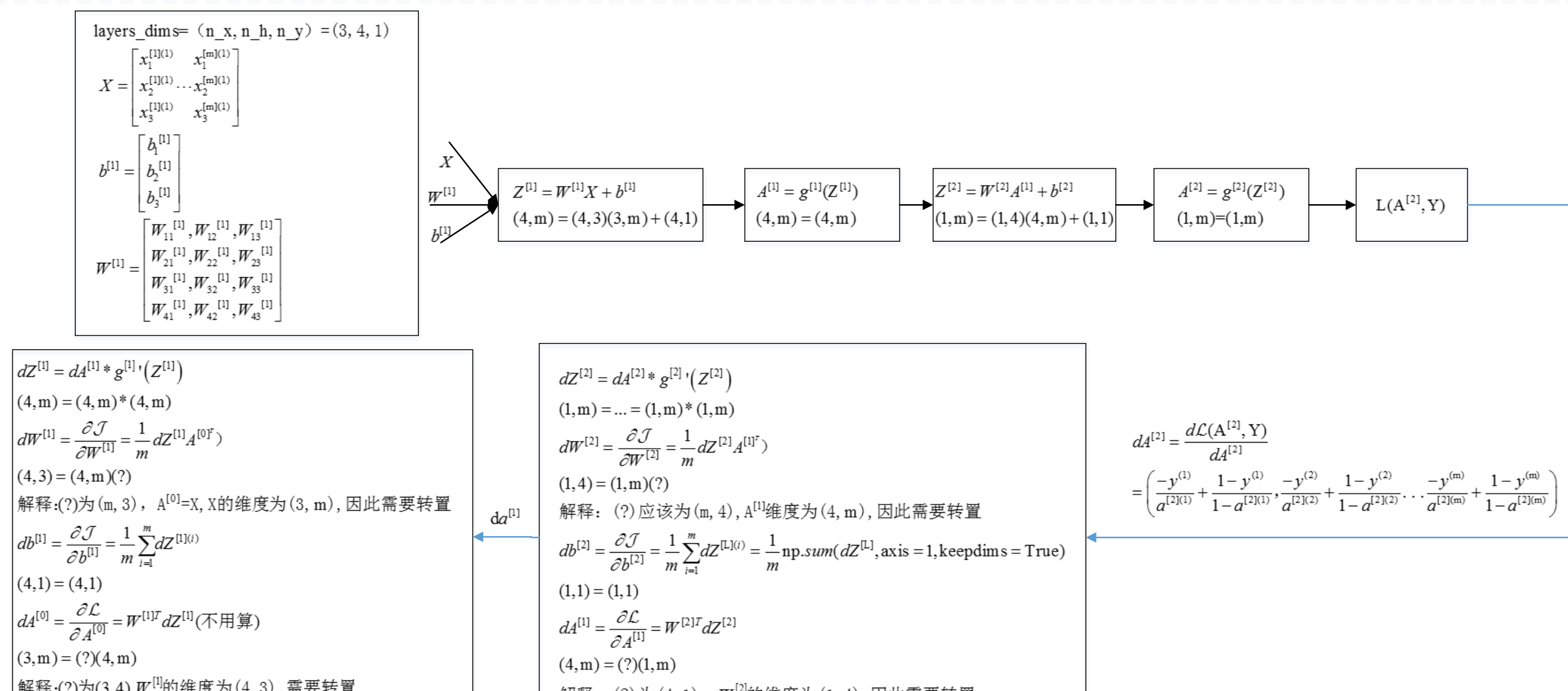


Author: 贺志尧  
Email:2282815808@qq.com

解释公式 $dW^{[L]} = dz^{[L]} \cdot a^{[L-1]T}$ 与 $da^{[L]} = W^{[L]T} \cdot dz^{[L]}$ 中， $a^{[L-1]T}$ 与 $W^{[L]T}$ 转置的原因



Author: 贺志尧  
Email:2282815808@qq.com



Author: 贺志尧  
Email:2282815808@qq.com

解释公式 $dW^{[1]} = \frac{\partial \mathcal{J}}{\partial W^{[1]}} = \frac{1}{m} dZ^{[1]} A^{[0]T}$ 与 $db^{[1]} = \frac{\partial \mathcal{J}}{\partial b^{[1]}} = \frac{1}{m} \sum_{i=1}^m dZ^{[1](i)}$ 中乘以 $\frac{1}{m}$ 的原因

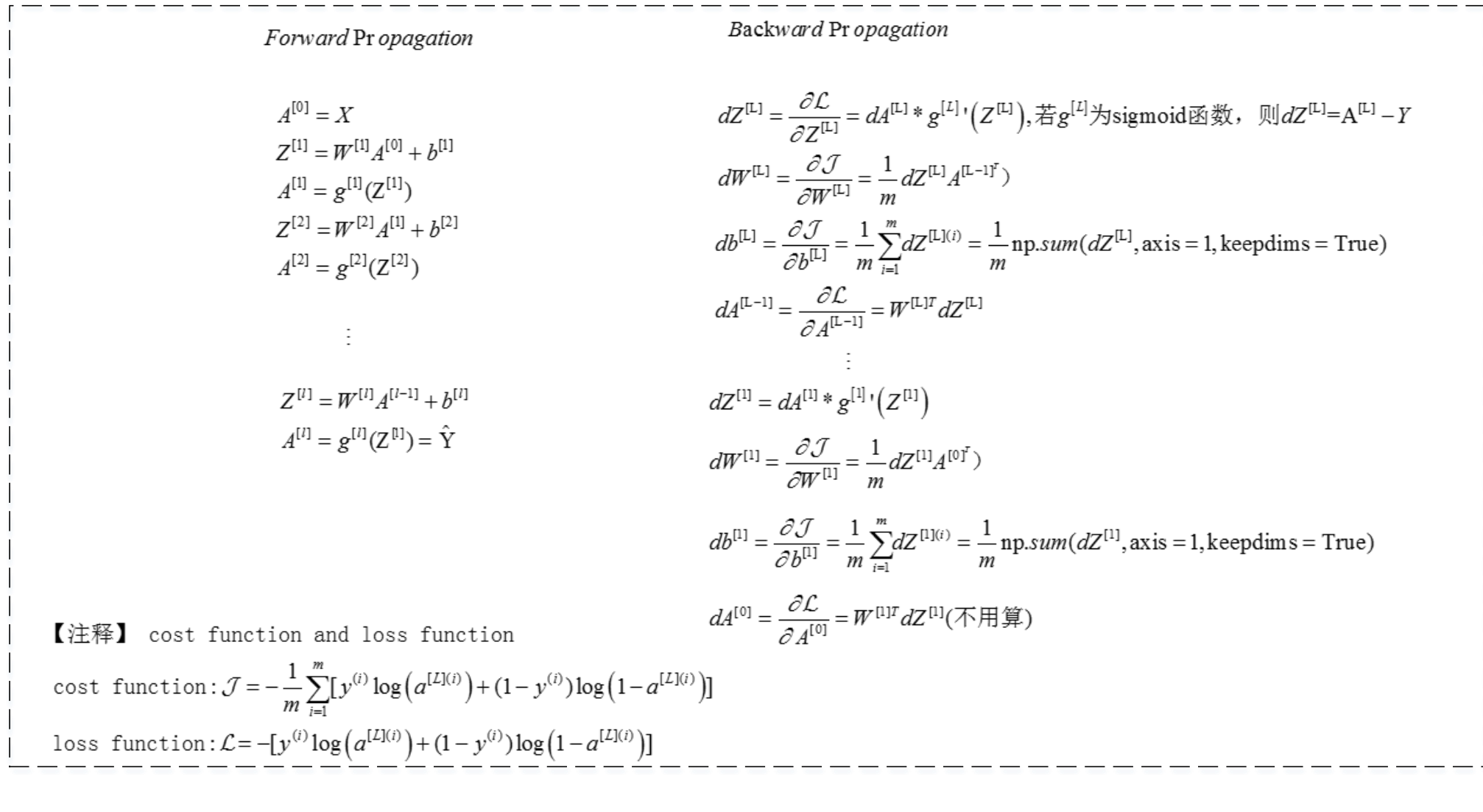
# 【证明1】

$$\begin{aligned} \mathcal{J} &= \frac{1}{m} \sum_{i=1}^m \mathcal{L}^{(i)} \\ \frac{\partial \mathcal{L}}{\partial W^{[1]}} &= \left[ \frac{\partial \mathcal{L}^{(1)}}{\partial W^{[1]}} \quad \frac{\partial \mathcal{L}^{(2)}}{\partial W^{[1]}} \quad \dots \quad \frac{\partial \mathcal{L}^{(m)}}{\partial W^{[1]}} \right] = \left[ dz^{[1](1)} \cdot a^{[0](1)T} \quad dz^{[1](2)} \cdot a^{[0](2)T} \quad \dots \quad dz^{[1](m)} \cdot a^{[0](m)T} \right] \\ \frac{\partial \mathcal{J}}{\partial W^{[1]}} &= \frac{\partial \mathcal{J}}{\partial W^{[1]}} \\ &= \frac{1}{m} \left( dz^{[1](1)} \cdot a^{[0](1)T} + dz^{[1](2)} \cdot a^{[0](2)T} + \dots + dz^{[1](m)} \cdot a^{[0](m)T} \right) \\ &= \frac{1}{m} \begin{bmatrix} dz^{[1](1)} & dz^{[1](2)} & \dots & dz^{[1](m)} \\ a_{1,1}^{[0](1)} & a_{1,2}^{[0](1)} & \dots & a_{1,n}^{[0](1)} \\ a_{2,1}^{[0](1)} & a_{2,2}^{[0](1)} & \dots & a_{2,n}^{[0](1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}^{[0](1)} & a_{m,2}^{[0](1)} & \dots & a_{m,n}^{[0](1)} \end{bmatrix}^T \\ &= \frac{1}{m} dZ^{[1]} A^{[0]T} \end{aligned}$$

# 【证明2】

$$\begin{aligned} \mathcal{J} &= \frac{1}{m} \sum_{i=1}^m \mathcal{L}^{(i)} \\ \frac{\partial \mathcal{L}}{\partial b^{[1]}} &= \left[ \frac{\partial \mathcal{L}^{(1)}}{\partial b^{[1]}} \quad \frac{\partial \mathcal{L}^{(2)}}{\partial b^{[1]}} \quad \dots \quad \frac{\partial \mathcal{L}^{(m)}}{\partial b^{[1]}} \right] = \left[ dz^{[1](1)} \quad dz^{[1](2)} \quad \dots \quad dz^{[1](m)} \right] \\ \frac{\partial \mathcal{J}}{\partial b^{[1]}} &= \frac{\partial \mathcal{J}}{\partial b^{[1]}} \\ &= \frac{1}{m} \left( dz^{[1](1)} + dz^{[1](2)} + \dots + dz^{[1](m)} \right) \\ &= \frac{1}{m} \sum_{i=1}^m dZ^{[1]} \end{aligned}$$

Author: 贺志尧  
Email:2282815808@qq.com



Author: 贺志尧  
Email:2282815808@qq.com