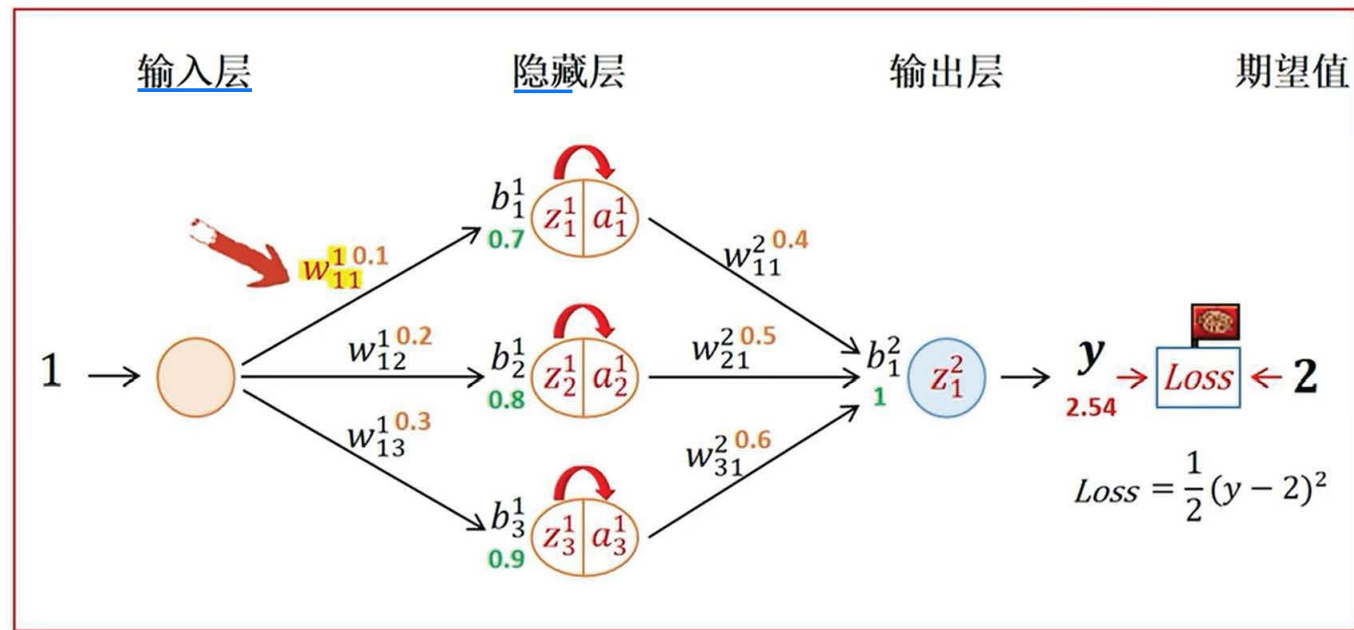
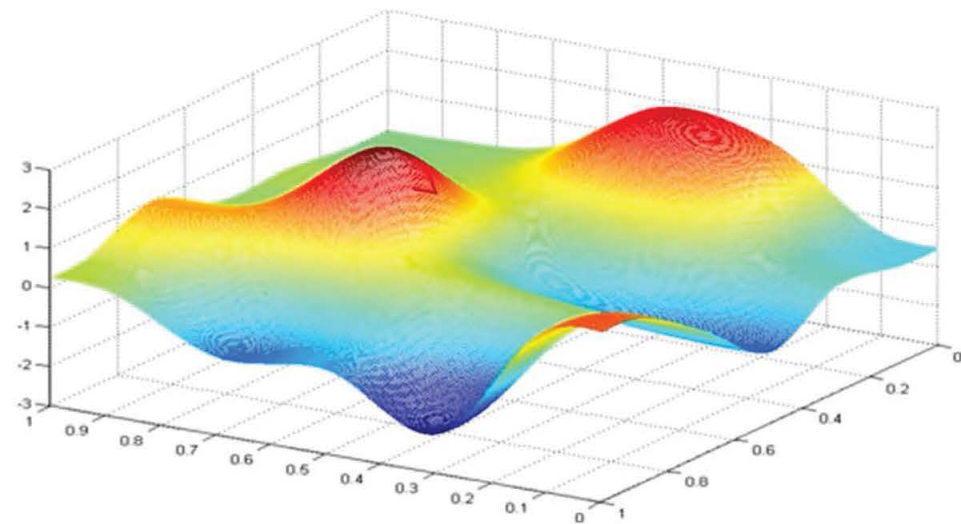


梯度消失/爆炸



$$0 = \frac{\partial Loss}{\partial y} \frac{\partial y}{\partial z_1^2} \frac{\partial z_1^2}{\partial a_1^1} \frac{\partial a_1^1}{\partial z_1^1} \frac{\partial z_1^1}{\partial w_{11}^1}$$

$$\cancel{w_{11}^1} = w_{11}^1 - \alpha 0$$



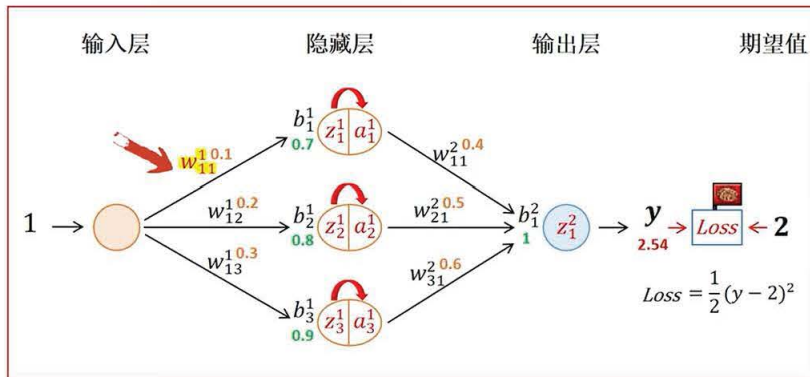
梯度消失



+ 关注

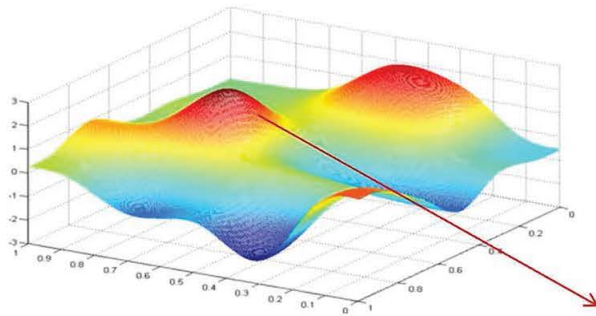


发现《Carefree》>



$$\frac{\partial Loss}{\partial w_{11}^1} = \frac{\partial Loss}{\partial y} \frac{\partial y}{\partial z_1^2} \frac{\partial z_1^2}{\partial a_1^1} \frac{\partial a_1^1}{\partial z_1^1} \frac{\partial z_1^1}{\partial w_{11}^1}$$

$$NaN = w_{11}^1 - \alpha \frac{\partial Loss}{\partial w_{11}^1}$$



梯度消失

梯度爆炸

