HW 1-3

Step.1 定義模型參數

- $\hat{y}$ 為模型的預測值
- $z00, z01, z10, z11$為捲積層送入激勵函數前的數值

Step.2 $W11$ 梯度公式

則 $w11$ 梯度可如下表示：

$$
\begin{aligned}
\frac{\partial MSE}{\partial w_{11}} = \frac{\partial MSE}{\partial \hat{y}}[\frac{\partial \hat{y}}{\partial o00}\frac{\partial o00}{\partial z00}\frac{\partial z00}{\partial w11} \\
+ \frac{\partial \hat{y}}{\partial o01}\frac{\partial o01}{\partial z01}\frac{\partial z01}{\partial w11} \\
+ \frac{\partial \hat{y}}{\partial o10}\frac{\partial o10}{\partial z10}\frac{\partial z10}{\partial w11} \\
+ \frac{\partial \hat{y}}{\partial o11}\frac{\partial o11}{\partial z11}\frac{\partial z11}{\partial w11}]
\end{aligned}
$$

Step.3 計算偏微分

計算各部份的偏微分：

$$
\begin{aligned}
MSE &= \frac{1}{n}(yy-\hat{y})^2, n=1 \\
\Rightarrow \frac{\partial MSE}{\partial \hat{y}} &= \frac{\partial}{\partial (yy-\hat{y})}\frac{1}{n}(yy-\hat{y})^2 \times \frac{\partial}{\partial \hat{y}}(yy-\hat{y}) \\
&= \frac{2}{n}(\hat{y}-y) \times -1 = -2(yy-\hat{y})
\end{aligned}
$$

$$
\hat{y} = wa \times o00 + wb \times o01 + wc \times o10 + wd \times o11
$$

$$
\Rightarrow \frac{\partial \hat{y}}{\partial o00} = wa
$$

$$
\Rightarrow \frac{\partial \hat{y}}{\partial o01} = wb
$$

$$
\Rightarrow \frac{\partial \hat{y}}{\partial o10} = wc
$$

$$
\Rightarrow \frac{\partial \hat{y}}{\partial o11} = wd
$$

$$
oii = sigmod(zii)
$$

$$
\Rightarrow \frac{\partial oii}{\partial zii} = \frac{\partial}{\partial zii}\frac{1}{1+e^{-zii}} = \frac{1}{1+e^{-zii}}(1-\frac{1}{1+e^{-zii}})
$$

$$
\begin{aligned}
z00 = a_{11} \times w11 + \ldots \Rightarrow \frac{\partial z00}{\partial w11} = a_{11} \\
z01 = a_{12} \times w11 + \ldots \Rightarrow \frac{\partial z01}{\partial w11} = a_{12} \\
z10 = a_{21} \times w11 + \ldots \Rightarrow \frac{\partial z10}{\partial w11} = a_{21} \\
z11 = a_{22} \times w11 + \ldots \Rightarrow \frac{\partial z11}{\partial w11} = a_{22}
\end{aligned}
$$

將以上結果代入 $w11$ 的梯度公式，得：

$$
\begin{aligned}
\frac{\partial MSE}{\partial w11} = -2(yy - \hat{y}) \quad & [wa \times \frac{1}{1+e^{-z00}}(1 - \frac{1}{1+e^{-z00}}) \times a11 \\
& + wb \times \frac{1}{1+e^{-z01}}(1 - \frac{1}{1+e^{-z01}}) \times a12 \\
& + wc \times \frac{1}{1+e^{-z10}}(1 - \frac{1}{1+e^{-z10}}) \times a21 \\
& + wd \times \frac{1}{1+e^{-z11}}(1 - \frac{1}{1+e^{-z11}}) \times a22]
\end{aligned}
$$