POLY 2

数学模型：$$z = w_0 + \sum_{j=1}^{n} w_j x_j + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} w_{jk} x_j x_k .$$

$$\hat{y} = \frac{1}{1+e^{-z}}$$

损失函数 $= J(w_0, w_1, w_2 \dots w_n, w_{12}, w_{13} \dots w_{n-1,n})$

$$= -\frac{1}{m} \sum_{i=1}^{m} [y^{(i)} \log \hat{y}^{(i)} + (1-y^{(i)}) \log (1-\hat{y}^{(i)})]$$

梯度下降：

①
$$\frac{\partial J}{\partial w_j} = -\frac{1}{m} \sum_{i=1}^{m} [y^{(i)} \frac{\partial \log \hat{y}^{(i)}}{\partial w_j} + (1-y^{(i)}) \frac{\partial \log (1-\hat{y}^{(i)})}{\partial w_j}]$$

$$= -\frac{1}{m} \sum_{i=1}^{m} [y^{(i)} \times \frac{\partial \log \hat{y}^{(i)}}{\partial \hat{y}^{(i)}} \times \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \times \frac{\partial z^{(i)}}{\partial w_j}$$

$$+ (1-y^{(i)}) \times \frac{\partial \log (1-\hat{y}^{(i)})}{\partial (1-\hat{y}^{(i)})} \times \frac{\partial (1-\hat{y}^{(i)})}{\partial \hat{y}^{(i)}} \times \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \times \frac{\partial z^{(i)}}{\partial w_j}]$$

$$= -\frac{1}{m} \sum_{i=1}^{m} [\, y^{(i)} \times (1-\hat{y}^{(i)}) + (1-y^{(i)})(-\hat{y}^{(i)})] x_j^{(i)}$$

$$= \boxed{-\frac{1}{m} \sum_{i=1}^{m} (y^{(i)} - \hat{y}^{(i)}) \, x_j^{(i)}}$$

② $\frac{\partial J}{\partial w_0} = \boxed{-\frac{1}{m}\sum_{i=1}^{m}(y^{(i)} - \hat{y}^{(i)})}$

③ $\frac{\partial J}{\partial w_{jk}} = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \times \frac{\partial \log \hat{y}^{(i)}}{\partial w_{jk}} + (1-y^{(i)}) \times \frac{\partial \log(1-\hat{y}^{(i)})}{\partial w_{jk}}$

$= -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \times \frac{\partial \log \hat{y}^{(i)}}{\partial \hat{y}^{(i)}} \times \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \times \frac{\partial z^{(i)}}{\partial w_{jk}}$

$\frac{\partial \log \hat{y}^{(i)}}{\partial \hat{y}^{(i)}} \to \frac{1}{\hat{y}^{(i)}}$ $\quad \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \to \hat{y}^{(i)}(1-\hat{y}^{(i)})$ $\quad \frac{\partial z^{(i)}}{\partial w_{jk}} \to x_j^{(i)} x_k^{(i)}$

$+ (1-y^{(i)}) \frac{\partial \log(1-\hat{y}^{(i)})}{\partial (1-\hat{y}^{(i)})} \times \frac{\partial (1-\hat{y}^{(i)})}{\partial \hat{y}^{(i)}} \times \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \times \frac{\partial z^{(i)}}{\partial w_{jk}}$

$\frac{\partial \log(1-\hat{y}^{(i)})}{\partial (1-\hat{y}^{(i)})} \to \frac{1}{1-\hat{y}^{(i)}}$ $\quad \frac{\partial (1-\hat{y}^{(i)})}{\partial \hat{y}^{(i)}} \to -1$ $\quad \frac{\partial \hat{y}^{(i)}}{\partial z^{(i)}} \to \hat{y}^{(i)} \times (1-\hat{y}^{(i)})$ $\quad \frac{\partial z^{(i)}}{\partial w_{jk}} \to x_j^{(i)} x_k^{(i)}$

$= -\frac{1}{m}\sum_{i=1}^{m} [y^{(i)} \times (1-\hat{y}^{(i)}) + (1-y^{(i)})(-\hat{y}^{(i)})] x_j^{(i)} x_k^{(i)}$

$= \boxed{-\frac{1}{m}\sum_{i=1}^{m} (y^{(i)} - \hat{y}^{(i)}) x_j^{(i)} x_k^{(i)}}$

$$\begin{cases} w_0 = w_0 + \eta \cdot \frac{1}{m}\sum_{i=1}^{m}(y^{(i)} - \hat{y}^{(i)}) \\ w_j = w_j + \eta \frac{1}{m}\sum_{i=1}^{m}(y^{(i)} - \hat{y}^{(i)}) x_j^{(i)} \\ w_{ij} = w_{ij} + \eta \frac{1}{m}\sum_{i=1}^{m}(y^{(i)} - \hat{y}^{(i)}) x_j^{(i)} x_k^{(i)} \end{cases}$$