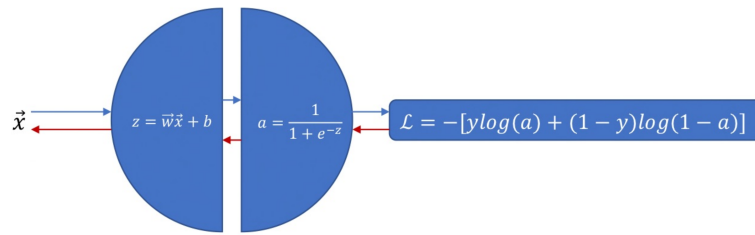


Question. 7-02

다음과 같이 Logistic Regression Model에서 Activation Function을 Sigmoid Function으로 설정하였다. 이때, 다음 물음에 답하시오.



- 1) Sigmoid Function의 Partial Derivative를 구하고 Backpropagation되는 값을 구하시오.
(단, Loss Function으로 부터의 Backpropagation 값은 $\frac{\partial \mathcal{L}}{\partial a}$ 로만 표기한다.)
- 2) Backpropagation 진행 시 Sigmoid Function의 Partial Derivative가 업데이트 시키는 크기의 최댓값과 최솟값을 찾고 그때의 input z를 구하시오.
- 3) Sigmoid Function의 Partial Derivative의 크기를 최대 혹은 최소로 만드는 input z의 의미는 무엇인가.

1) Sigmoid Function의 z에 대한 Partial Derivative를 구한다.

$$\frac{\partial \sigma(z)}{\partial z} = \frac{\partial}{\partial z} \left(\frac{1}{1+e^{-z}} \right) = \frac{0 \cdot (-1)e^{-z}}{(1+e^{-z})^2} = \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}} \left(1 - \frac{1}{1+e^{-z}} \right) = a(1-a)$$

$$\therefore \frac{\partial \sigma(z)}{\partial z} = \sigma(z)(1-\sigma(z))$$

그러므로 Activation Function으로 부터 Backpropagation되는 값은 $\frac{\partial \mathcal{L}}{\partial a} \cdot \sigma(z)(1-\sigma(z))$ 이다.

$$\begin{aligned} 2) \frac{\partial}{\partial z} \left[\frac{e^{-z}}{(e^{-z}+1)^2} \right] &= \frac{e^{-z} \cdot \frac{\partial}{\partial z} [-z] \cdot (e^{-z}+1)^2 - e^{-z} \cdot 2(e^{-z}+1) \cdot \frac{\partial}{\partial z} [e^{-z}+1]}{(e^{-z}+1)^4} \\ &= \frac{-e^{-z} \cdot 1 \cdot (e^{-z}+1)^2 - e^{-z} \cdot 2(e^{-z}+1) \cdot (e^{-z}) \cdot (-1)}{(e^{-z}+1)^4} \\ &= \frac{2e^{-2z}(e^{-z}+1) - e^{-z}(e^{-z}+1)^2}{(e^{-z}+1)^4} \\ &= \frac{(e^{-z}-1)e^{-z}}{(e^{-z}+1)^3} \end{aligned}$$

$$\therefore z = \pm\infty \text{ or } z=0, \frac{\partial}{\partial z} \left(\frac{\partial \sigma}{\partial z} \right) = 0$$

$$\text{if } z = \pm\infty, \frac{\partial \sigma}{\partial z} = 0$$

$$\text{if } z=0, \frac{\partial \sigma}{\partial z} = 0.25$$

$\therefore z = \pm\infty$ 일시 최솟값 0, $z=0$ 일시 최댓값 0.25를 찾는다.

3) $z = \pm \infty$ 은 Binary Classification의 결과가 0 혹은 1로 확실한 경우를 의미한다.

이때, 확실한 것은 재므로 update가 일어나지 않는다.

$z = 0$ 은 Binary Classification의 결과가 가장 불확실한 경우를 의미한다.

따라서, update가 가장 크게 일어난다.