

Problem 7.5

Faisal Memon

May 23, 2024

1 Problem

Calculate the derivative $\frac{\partial l_i}{\partial f[\mathbf{x}_i, \phi]}$ for binary classification loss function

$$l_i = -1(1 - y_i) \log[1 - \text{sig}[f[\mathbf{x}_i, \phi]]] - y_i \log[\text{sig}[f[\mathbf{x}_i, \phi]]]$$

$$\text{With } \text{sig}[z] = \frac{1}{1 + \exp[-z]}$$

2 Answer

$$l_i = -1(1 - y_i) \log[1 - \text{sig}[f[\mathbf{x}_i, \phi]]] - y_i \log[\text{sig}[f[\mathbf{x}_i, \phi]]]$$

$$l_i = (y_i - 1) \log[1 - \text{sig}[f[\mathbf{x}_i, \phi]]] - y_i \log[\text{sig}[f[\mathbf{x}_i, \phi]]]$$

$$l_i = a \cdot b - c \cdot d$$

Where

$$a = (y_i - 1)$$

$$b = \log[1 - \text{sig}[f[\mathbf{x}_i, \phi]]]$$

$$c = y_i$$

$$d = \log[\text{sig}[f[\mathbf{x}_i, \phi]]]$$

Using Chain Rule twice

$$\frac{\partial l_i}{\partial f[\mathbf{x}_i, \phi]} = a \cdot \frac{\partial b}{\partial f[\mathbf{x}_i, \phi]} + b \cdot \frac{\partial a}{\partial f[\mathbf{x}_i, \phi]} - c \cdot \frac{\partial d}{\partial f[\mathbf{x}_i, \phi]} - d \cdot \frac{\partial c}{\partial f[\mathbf{x}_i, \phi]}$$

Using SymPy

```
sig is 1/(1 + exp(-x))
```

```
diff(sig) is \frac{e^{-x}}{\left(1 + e^{-x}\right)^2}
```

```
a is y - 1
```

```
diff(a, x) is 0
```

```
b is log(1 - 1/(1 + exp(-x)))
```

```
diff(b, x) is - \frac{e^{-x}}{\left(1 - \frac{1}{1 + e^{-x}}\right)}
```

$$\left(1 + e^{-x}\right)^2$$

c is y
diff(c, x) is 0

d is $\log(1/(1 + \exp(-x)))$
diff(d, x) is $\frac{e^{-x}}{1 + e^{-x}}$

Therefore

$$\begin{aligned} \frac{\partial l_i}{\partial f[\mathbf{x}_i, \phi]} &= (y_i - 1) \cdot -\frac{e^{-x}}{\left(1 - \frac{1}{1+e^{-x}}\right)(1 + e^{-x})^2} + \log[1 - \text{sig}[f[\mathbf{x}_i, \phi]]] \cdot 0 - y_i \cdot \frac{e^{-x}}{1 + e^{-x}} - \log[\text{sig}[f[\mathbf{x}_i, \phi]]] \cdot 0 \\ &= (y_i - 1) \cdot -\frac{e^{-x}}{\left(1 - \frac{1}{1+e^{-x}}\right)(1 + e^{-x})^2} - y_i \cdot \frac{e^{-x}}{1 + e^{-x}} \end{aligned}$$