Problem 7.5

Faisal Memon

May 25, 2024

1 Problem

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

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

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

2 Answer

$$l_i = -1(1 - y_i)\log[1 - sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]] - y_i \log[sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]]$$
$$l_i = (y_i - 1)\log[1 - sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]] - y_i \log[sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]]$$
$$l_i = a \cdot b - c \cdot d$$

Where

$$a = (y_i - 1)$$
$$b = \log[1 - sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]]$$
$$c = y_i$$
$$d = \log[sig[f[\boldsymbol{x}_i, \boldsymbol{\phi}]]]$$

Using Chain Rule twice

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

```
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

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

$$\frac{\partial l_i}{\partial f[\boldsymbol{x}_i, \boldsymbol{\phi}]} = (y_i - 1) \cdot - \frac{e^{-x}}{\left(1 - \frac{1}{1+e^{-x}}\right)\left(1 + e^{-x}\right)^2} - y_i \cdot \frac{e^{-x}}{1 + e^{-x}} \quad \square$$