Logistic Regression

Đạt Nguyễn Ngọc

October 2022

1 Problem 1:

$L = -[ylog(\hat{y}) + (1-y)log(1-\hat{y})]$ in which $\hat{y} = \alpha(w_0 + w_1x_1 + w_2x_2)$

$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w}$

$$\begin{aligned}
\frac{\partial L}{\partial \hat{y}} &= \frac{\partial[-ylog(\hat{y}) + (1-y)log(1-\hat{y})]}{\partial \hat{y}} \\
&= -[y\frac{1}{\hat{y}} - (1-y)\frac{1}{1-\hat{y}}] \\
&= -\frac{y-\hat{y}}{\hat{y}(1-\hat{y})}
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial \hat{y}}{\partial w} &= \frac{\partial \hat{y}}{\partial z}\frac{\partial z}{\partial w} \\
&= \frac{\partial \frac{1}{1+z}}{\partial z}\frac{\partial z}{\partial w} \\
&= -\frac{1}{(1+z)^2}(-zw) \\
&= \frac{zX}{(1+z)^2} \\
&= X\hat{y}(1-\hat{y})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L}{\partial w} &= -\frac{y-\hat{y}}{\hat{y}(1-\hat{y})}X\hat{y}(1-\hat{y}) \\
&= X(\hat{y}-y)
\end{aligned}$$

Biểu diễn dưới dạng ma trận: $\frac{\partial L}{\partial w} = X^T(\hat{Y} - Y)$

2 Problem 2:

Chứng minh với model logistic thì loss binary crossentropy là convex function với W, loss mean square error không là convex function với W

a, Binary-crossentropy: Từ bài 1 ta có:

Do:

$$\frac{\partial L}{\partial w} = X(\hat{y} - y)$$

$$\frac{\partial \hat{y}}{\partial w} = X\hat{y}(1 - \hat{y})$$

Nên:

$$\begin{aligned} \frac{\partial^2 L}{\partial w^2} &= X\frac{\partial \hat{y}}{\partial w} \\ &= X^2\hat{y}(1 - \hat{y}) \geq 0 \end{aligned}$$

$\Rightarrow$ **The loss binary-crossentropy with logistic model is convex**

b,MSE:

$$L = \frac{1}{N}\sum_{i=1}^{N}(\hat{y} - y)^2$$

$$\begin{aligned} \frac{\partial L}{\partial w} &= \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w} \\ &= -2(y - \hat{y})x\hat{y}(1 - \hat{y}) \\ &= -2x(y\hat{y} - \hat{y}^2)(1 - \hat{y}) \\ &= -2x(y\hat{y} - y\hat{y}^2 - \hat{y}^2 + \hat{y}^3) \end{aligned}$$

Second derivate:

$$\begin{aligned} \frac{\partial^2 L}{\partial w^2} &= -2x(y\frac{\partial \hat{y}}{\partial w} - y\frac{\partial \hat{y}^2}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w} - \frac{\partial \hat{y}^2}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w} + \frac{\partial \hat{y}^3}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w}) \\ &= -2x(y.x.\hat{y}(1 - \hat{y}) - y.2.\hat{y}.x.\hat{y}(1 - \hat{y}) - 2.\hat{y}.x.\hat{y}(1 - \hat{y}) + 3.\hat{y}^2.x.\hat{y}(1 - \hat{y})) \\ &= -2.x^2.\hat{y}.(1 - \hat{y}).(y - 2.y.\hat{y} - 2\hat{y} + 3\hat{y}^2) \end{aligned}$$

Vì $x^2y(1 - \hat{y}) \geq 0$

Khi y = 0: $\frac{\partial^2 L}{\partial w^2} = 4\hat{y} - 6\hat{y}^2 = 2\hat{y}(2\hat{y} - 3\hat{y}) \geq 0 \forall \hat{y} \in [\frac{2}{3}, 1]$

Khi y = 1: $\frac{\partial^2 L}{\partial w^2} = -2(3\hat{y} - 1)(\hat{y} - 1) \geq 0 \forall \hat{y} \in [0, \frac{1}{3}]$

$\Rightarrow$ **The loss MSE with logistic model is convex**