# Exercise3 - 1930026143

We know that the cost function of logistic regression as:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^{(i)}\log\left(h_\theta\left(x^{(i)}\right)\right) - \left(1-y^{(i)}\right)\log\left(1-h_\theta\left(x^{(i)}\right)\right)\right] \#(1)$$

Then the gradient descent function $\theta_j \coloneqq \theta_j - \alpha\frac{\partial J(\theta)}{\partial\theta_j}$ simultaneously update all $\theta_j$ and repeat iterations. In order to facilitate the operation, we first analyze the derivative of the sigmoid function:

$$f(x) = \frac{1}{1+e^{-z}} \#(2)$$

$$f^{'}(x) = e^{-z}\left(\frac{1}{1+e^{-z}}\right)^2 = \frac{1}{1+e^{-z}}\left(1-\frac{1}{1+e^{-z}}\right) = f(x)\left(1-f(x)\right) \#(3)$$

Using this property (3) and the chain rule, we can expand the formula $\frac{\partial J(\theta)}{\partial\theta_j}$ to simplify:

$$\begin{aligned}\frac{\partial J(\theta)}{\partial\theta_j} = \frac{1}{m}\sum_{i=1}^{m}\Bigg[&-y^{(i)}\frac{1}{h_\theta\left(x^{(i)}\right)}h_\theta\left(x^{(i)}\right)\left(1-h_\theta\left(x^{(i)}\right)\right)x_j^{(i)}\\ &-\left(1-y^{(i)}\right)\frac{1}{1-h_\theta\left(x^{(i)}\right)}h_\theta\left(x^{(i)}\right)\left(1-h_\theta\left(x^{(i)}\right)\right)x_j^{(i)}\Bigg] \#(4)\end{aligned}$$

$$= \frac{1}{m}\sum_{i=1}^{m} - y^{(i)}x^{(i)} + h_\theta\left(x^{(i)}\right)x_j^{(i)}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$$

In conclusion, the gradient descent function of logistic regression is:

$repeat\{$

$$\theta_j \coloneqq \theta_j - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$$

} simultaneously update all $\theta_j$.