Machine Learning Week 6

Logistic Regression

Bui Phuong Thao - 11215341 - DSEB 63

1 Exercise 1

a) Give the formula of sigmoid function and calculate its derivative

$$
\begin{aligned}
\sigma(z)' &= \Big(\frac{1}{1+e^{-z}}\Big)' \\
&= -\frac{(1+e^{-z})'}{(1+e^{-z})^2} \\
&= -\frac{(e^{-z})'}{(1+e^{-z})^2} \\
&= -e^{-z}.\frac{-1}{(1+e^{-z})^2} \\
&= \frac{e^{-z}}{(1+e^{-z})^2} \\
&= \frac{e^{-z}}{1+e^{-z}}.\frac{1}{1+e^{-z}} \\
&= \Big(1-\frac{1}{(1+e^{-z})^2}\Big).\frac{1}{1+e^{-z}} \\
&= \sigma(z).(1-\sigma(z))
\end{aligned}
$$

b) Give the formula for loss function in logistic regression. What type of loss function is it?
The loss function of logistic regression is

$$L = -\log p(t|w) = -\sum_{n=1}^{N}\Big(t_n \log_{y_n} + (1-t_n)\log(1-y_n)\Big)$$

This is a binary classification loss function

c) Calculate the gradient vector for loss function in logistic regression

$$
\begin{aligned}
L &= -\sum_{n=1}^{N}\Big(t_n \log_{y_n} + (1-t_n)\log(1-y_n)\Big) \\
L &= -\sum_{n=1}^{N}\Big(t_n \log(\sigma(wx_n)) + (1-t_n).\log(1-\sigma(wx_n))\Big) \\
\frac{\delta(L)}{\delta(w)} &= -\sum_{n=1}^{N}\Big(t_n \frac{1}{\sigma(wx_n)}\sigma(wx_n)(1-\sigma(wx_n)).(wx_n)' + (1-t_n)\frac{-\sigma(wx_n)(1-\sigma(x_n w))}{1-\sigma(x_n w)}.(wx_n)'\Big) \\
&= -\sum_{n=1}^{N}\Big(t_n(1-\sigma(wx_n))x_n - (1-t_n)\sigma(wx_n)x_n\Big) \\
&= -\sum_{n=1}^{N}\Big(t_n x_n - \sigma(wx_n)t_n x_n - \sigma(wx_n)x_n + t_n\sigma(wx_n)x_n\Big) \\
&= \sum_{n=1}^{N}\Big(-t_n x_n + \sigma(wx_n)t_n x_n + \sigma(wx_n)x_n - t_n\sigma(wx_n)x_n\Big) \\
&= \sum_{n=1}^{N}\Big(x_n(\sigma(wx_n) - t_n)\Big)
\end{aligned}
$$

The equation can be written in the vector form as follow

$\rightarrow \frac{\delta(L)}{\delta(W)} = X^T.(\sigma(XW) - T)$

2 Exercise 2

a) Implement logistic regression algorithm from scratch. You can use Gradient Descent or Newton Raphson as optimization method for loss function. Note that your implementation should follow an OOP form.