Logistic Regression

Hà Phng

October 12, 2021

1 Exercise 1:

Assume that the probability that a data point x falls into class 1 is $f(w^T x)$ and falls into class 0 is $1 - (w^T x)$

$$P(y_i = 1|\mathbf{x_i}, \mathbf{w}) = f(w^T x)$$

$$P(y_i = 0|\mathbf{x_i}, \mathbf{w}) = 1 - f(w^T x)$$

We have $\mathbf{X} = [x_1 + x_2 + \cdots + x_n] and \mathbf{y} = [y_1 + y_2 + \cdots + y_n]$ We need to find $\mathbf{w}$ to maximize

$$P(\mathbf{y}|\mathbf{X}, \mathbf{w})$$

Assuming that the data points are randomly generated independently, we can write:

$$P(\mathbf{y}|\mathbf{X}, \mathbf{w}) = \Pi P(y_i|\mathbf{x_i}, \mathbf{w}) = \Pi z_i^{y_i}(1 - z_i)^{1-y_i}$$

We have

$$J(\mathbf{w}) = -logP(\mathbf{y}|\mathbf{X}, \mathbf{w} = -\sum(y_i log z_i + (1 - y_i)log(1 - z_i))$$

So to find $\mathbf{w}$ to maximize $P(\mathbf{y}|\mathbf{X}, \mathbf{w})$ equal to find $\mathbf{w}$ to minimize $J(\mathbf{w})$
Derivative $J(\mathbf{w})$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(J(\mathbf{w_i}; \mathbf{x_i}, y)\right) = -\left(\frac{y_i}{z_i} - \frac{1 - y_i}{1 - z_i}\right)\frac{dz_i}{d\mathbf{w}} = \frac{z_i - y_i}{z_i(1 - z_i}\frac{dz_i}{ds}\frac{ds}{d\mathbf{w}}$$

We have $z = f(\mathbf{w^T x})$ and $s = \mathbf{w^T x} and z is sigmoid function (z = \sigma(s))$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(J(\mathbf{w_i}; \mathbf{x_i}, y)\right) = \frac{z_i - y_i}{z_i(1 - z_i)} z_i(1 - z_i)\mathbf{x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(J(\mathbf{w_i}; \mathbf{x_i}, y)\right) = (z_i - y_i)\mathbf{x_i}$$

2 Exercise 2:

We have $f^{'}(x) = f(x)(1 - f(x))$

$$\frac{df(x)}{d(x} = f(x)(1 - f(x))$$

$$=> \frac{df(x)}{f(x)(1 - f(x))} = d(x$$

$$=> (\frac{1}{(1 - f(x))} + \frac{1}{(f(x))})df(x) = d(x)$$

Primitive

$$\int(\frac{1}{1 - f(x)} + \frac{1}{f(x)})df(x) = \int d(x)$$

$$=> lnf(x) - ln(1 - f(x)) = x$$

$$=> ln\frac{f(x)}{1-f(x)} = x$$

$$=> \frac{f(x)}{1-f(x)} = e^x$$

$$=> f(x) = \frac{1}{1+e^{-x}}$$