Linear Regression

Jubilee

1 Simple Case

1.1 Problem

Given a data set, $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, we want to find a function

$$f(x) = w \cdot x$$

to minimize the following error function :

$$E(w) = \frac{1}{2}\sum_{i=1}^{n}(f(x_i) - y_i)^2 = \frac{1}{2}\sum_{i=1}^{n}(w \cdot x_i - y_i)^2.$$

1.2 A solution

We can find a solution using $\frac{dE(w)}{dw} = 0$.

$$\frac{dE(w)}{dw} = \frac{1}{2}\frac{d(\sum_{i=1}^{n}(f(x_i) - y_i)^2)}{dw} = \frac{1}{2}\frac{d(\sum_{i=1}^{n}(w \cdot x_i - y_i)^2)}{dw}$$

$$= \sum_{i=1}^{n}(w \cdot x_i - y_i) \cdot x_i = w\sum_{i=1}^{n}x_i^2 - \sum_{i=1}^{n}y_i \cdot x_i = 0$$

$$\therefore w = \frac{\sum x_i y_i}{\sum x_i^2}.$$

2 Another Method

2.1 Newton's Method

How can we find a root of the function $f(w)$? In other words, we want to find w such that $f(w) = 0$.

We discussed the following equation:

$$\frac{f(w_1)}{w_2 - w_1} = f'(w_1).$$

That is,

$$w_2 = w_1 - \frac{f(w_1)}{f'(w_1)}.$$

Generally,

$$w_{n+1} = w_n - \frac{f(w_n)}{f'(w_n)}.$$

2.2 Back to Simple Case of Linear Regression

We solved linear regression using $\frac{dE(w)}{dw} = 0$. Newton's method says that finding w that satisfies $f(w) = 0$ can be solved by iterative method. When we substitute $E'(w)$ for $f(w)$,

$$w_{n+1} = w_n - \frac{E'(w_n)}{E''(w_n)}.$$

Simply, we get the following update equation:

$$w_{n+1} = w_n - \eta E'(w_n).$$

2.3 Update Equation for Simple Case

Recall that $E'(w) = \frac{dE(w)}{dw} = \frac{1}{2}\frac{d(\sum_{i=1}^{n}(f(x_i)-y_i)^2)}{dw} = \frac{1}{2}\frac{d(\sum_{i=1}^{n}(w \cdot x_i - y_i)^2)}{dw} = \sum_{i=1}^{n}(w \cdot x_i - y_i) \cdot x_i$. Therefore,

$$w_{n+1} = w_n - \eta E'(w_n) = w_n - \eta \sum_{i=1}^{n}(w_n \cdot x_i - y_i) \cdot x_i.$$

3 Linear Regression with Y intercept

3.1 Problem

Given a data set, $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, we want to find a function

$$f(x) = w \cdot x + \mathbf{b}$$

to minimize the following error function :

$$E(w, \mathbf{b}) = \frac{1}{2}\sum_{i=1}^{n}(f(x_i) - y_i)^2 = \frac{1}{2}\sum_{i=1}^{n}(w \cdot x_i + \mathbf{b} - y_i)^2.$$

3.2 Update Equation

Since

$$\frac{\partial E(w,b)}{\partial w}|_{w=w_n} = \sum_{i=1}^{n}(w_n \cdot x_i + b - y_i) \cdot x_i$$

and

$$\frac{\partial E(w,b)}{\partial b}|_{b=b_n} = \sum_{i=1}^{n}(w_n \cdot x_i + b - y_i) \cdot 1,$$

the update equation is as follows.

$$w_{n+1} = w_n - \eta\frac{\partial E(w,b)}{\partial w}|_{w=w_n} = w_n - \eta\sum_{i=1}^{n}(w_n \cdot x_i + b - y_i) \cdot x_i.$$

$$b_{n+1} = b_n - \eta\frac{\partial E(w,b)}{\partial b}|_{b=b_n} = b_n - \eta\sum_{i=1}^{n}(w_n \cdot x_i + b - y_i) \cdot 1.$$