simple linear regression deduction

Auegst-chen

2020.7.23

For sample set：$DD = (y_1, y_1), (x_2, y_2), ..., (x_m, y_m)$ Linear regression model is to find a model：$f(x) = wx + b$.
$\forall i \in [1, m]$,make$f(x_i)$ as close as possible to y_i .Therefore, we minimize the sum of squares of the difference between the predicted value and the real value of each sample：

$$(w^*, b^*) = arg_{w,b}, min \sum_{i=1}^{m}(f(x_i)-y_i)^2 = arg_{w,b}, min \sum_{i=1}^{m}(wx_i + b-y_i)^2 \quad (1)$$

This is actually the case for the multivariate function $h(w, b) = \sum_{i=1}^{m}(wx_i + b-y_i)^2$, then according to the knowledge of calculus,，functions$h$ ill take the partial derivatives of w 和 br espectively, and make the partial derivatives of the two values be zero, then the corresponding w^* and b^* are the minimum values of h Therefore, the partial derivative is obtained as follows:

$$\begin{aligned}\frac{\partial h(w,b)}{\partial w} &= \sum_{i=1}^{m}(2(wx_1 + b - y_i)) * x_i \\ &= 2(w\sum_{i=1}^{m}(x_i^2) - \sum_{i=1}^{m}((y_i - b)x_i)) \end{aligned} \quad (2)$$

$$\begin{aligned}\frac{\partial h(w,b)}{\partial b} &= 2(w\sum_{i=1}^{m}(2(wx_i + b - y_i))) * 1 \\ &= 2(mb - \sum_{i=1}^{m}(y_i - wx_i)) \end{aligned} \quad (3)$$

If they are equal to 0, the optimal solutions of $W$and Bcan be obtained

$$w = \frac{\sum_{i=1}^{m}(y_i(x_i - \bar{a}))}{\sum_{i=1}^{m}(x_i^2) - \frac{1}{m}(\sum_{i=1}^{m}(x_i))^2} \tag{4}$$

$$b = \frac{1}{m}(\sum_{i=1}^{m}(y_i - wx_i)) \tag{5}$$

Among them, equation (5) is easy to get from formula (3), which is not repeated here, but formula (4) is not so intuitive. Let's deduce it in detail. If formula (2) is equal to 0, then there are:

$$w\sum_{i=1}^{m}(x_i^2) = \sum_{i=1}^{m}(x_i y_i) - b\sum_{i=1}^{m}(x_i) \tag{6}$$

$$\begin{aligned} w\sum_{i=1}^{m}(x_i^2) &= \sum_{i=1}^{m}(x_i y_i) - \frac{1}{m}\sum_{i=1}^{m}(x_i)(\sum_{i=1}^{m}(y_i - wx_i)) \\ &= \sum_{i=1}^{m}(x_i y_i) - \bar{x}(\sum_{i=1}^{m}(y_i)) + \frac{1}{m}\sum_{i=1}^{m}(x_i * \sum_{i=1}^{m}(wx_i) \\ &= \sum_{i=1}^{m}(x_i y_i) - \bar{x}\sum_{i=1}^{m}(y_i) + \frac{w}{m}(\sum_{i=1}^{m}(x_i))^2 \end{aligned} \tag{7}$$

The right square term of formula (7) is shifted to the left

$$\begin{aligned} w\sum_{i=1}^{m}(x_i^2) - \frac{w}{m}(\sum_{i=1}^{m}(x_i))^2 &= \sum_{i=1}^{m}(x_i y_i) - \bar{x}\sum_{i=1}^{m}(y_i) \\ w[sum_{i=1}^{m}(x_i^2) - \frac{1}{m}(\sum_{i=1}^{m}(x_i))^2] &= \sum_{i=1}^{m}(x_i y_i) - \bar{x}\sum_{i=1}^{m}(y_i) \end{aligned} \tag{8}$$

Formula (4) can be obtained after the transformation of equation (8). So far, the expressions of W and B are only related to the observed samples, so we can estimate W and B by the observed values of samples.