Machine learning and vision laboratory

Dr H.Mohammad Zadeh

department: Electrical Engineering

Amirreza Velae	400102222
github	repository

Pre-Lab 2

October 20, 2024

Machine learning and vision laboratory

Pre-Lab 2

Amirreza Velae	400102222
github	repository

RSS

In a linear regression problem, the RSS is defined as follows:

$$RSS = \sum_{i=1}^{n}(y^i - \beta_0 - \beta_1 x^i)^2 \tag{1}$$

β_0^* and β_1^* are the solutions to this problem. Identify which of the following equations are correct.

1. $\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)y^i = 0$
2. $\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)(y^i - \bar{y}) = 0$
3. $\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)(x^i - \bar{x}) = 0$
4. $\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)(\beta_0 + \beta_1 x^i) = 0$

Soloution

From the equation (1) we have:

$$RSS = \sum_{i=1}^{n}(y^i - \beta_0 - \beta_1 x^i)^2 \tag{2}$$

If we differentiate the above equation with respect to β_0, β_1, x^i and y^i we have:

$$\frac{\partial RSS}{\partial \beta_0} = -2\sum_{i=1}^{n}(y^i - \beta_0 - \beta_1 x^i) \rightarrow \sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i) = 0 \tag{3}$$

$$\frac{\partial RSS}{\partial \beta_1} = -2\sum_{i=1}^{n}(y^i - \beta_0 - \beta_1 x^i)x^i \rightarrow \sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)x^i = 0 \tag{4}$$

Now, if we rewrite equation 3 and 4 we have:

$$\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)(x^i - \bar{x}) = \sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)x^i - \bar{x}\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i) \tag{5}$$

First term is zero, because of equation (4), also the second term is zero because of the equation (3), which is beeing multiplied by $\bar{x}$. For 4th equation we have:

$$\sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)(\beta_0 + \beta_1 x^i) = \beta_0^* \sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i) + \beta_1^* \sum_{i=1}^{n}(y^i - \beta_0^* - \beta_1^* x^i)x^i \quad (6)$$

Again, the first term is zero because of equation (1) and the second term is zero because of equation (2). So, the correct answer is the third and fourth equations.

underfitting

You will become familiar with the concept of overfitting in the next part. Research underfitting and briefly explain this concept.

Soloution

Imagine that we have a dataset that is not linearly separable. If we use a linear model to fit this dataset, the model will not be able to capture the underlying pattern of the data. Or, another example is that we have a dataset that is linearly separable, but we use a linear model with a low degree or we train the model with a small number of epochs. In this case, the model will not be able to capture the underlying pattern of the data. This is called underfitting. In other words, underfitting occurs when the model is too simple to capture the underlying pattern of the data.