

What is Multivariate Regression?

- Multivariate Regression helps use to measure the angle of more than one independent variable and more than one dependent variable. It finds the relation between the variables (Linearly related).
- It used to predict the behavior of the outcome variable and the association of predictor variables and how the predictor variables are changing.
- It can be applied to many practical fields like politics, economics, medical, research works and many different kinds of businesses.
- Multivariate regression is a simple extension of multiple regression.
- Multiple regression is used to predicting and exchange the values of one variable based on the collective value of more than one value of predictor variables.
- First, we will take an example to understand the use of multivariate regression after that we will look for the solution to that issue.

Examples of Multivariate Regression

- If E-commerce Company has collected the data of its customers such as Age, purchased history of a customer, gender and company want to find the relationship between these different dependents and independent variables.
- A gym trainer has collected the data of his client that are coming to his gym and want to observe some things of client that are health, eating habits (which kind of product client is consuming every week), the weight of the client. This wants to find a relation between these variables.

As you have seen in the above two examples that in both of the situations there is more than one variable some are dependent and some are independent, so single regression is not enough to analyze this kind of data.

Here is the multivariate regression that comes into the picture.

1. Feature selection

The selection of features plays the most important role in multivariate regression.

Finding the feature that is needed for finding which variable is dependent on this feature.

2. Normalizing Features

For better analysis features are need to be scaled to get them into a specific range. We can also change the value of each feature.

3. Select Loss function and Hypothesis

The loss function calculates the loss when the hypothesis predicts the wrong value.

And hypothesis means predicted value from the feature variable.

4. Set Hypothesis Parameters

Set the hypothesis parameter that can reduce the loss function and can predict.

5. Minimize the Loss Function

Minimizing the loss by using some lose minimization algorithm and use it over the dataset which can help to adjust the hypothesis parameters. Once the loss is minimized then it can be used for prediction.

There are many algorithms that can be used for reducing the loss such as gradient descent.

6. Test the hypothesis function

Check the hypothesis function how correct it predicting values, test it on test data.

Steps to follow archive Multivariate Regression

- 1) Import the necessary common libraries such as numpy, pandas
- 2) Read the dataset using the pandas' library
- 3) As we have discussed above that we have to normalize the data for getting better results. Why normalization because every feature has a different range of values.
- 4) Create a model that can archive regression if you are using linear regression use equation

$$Y = mx + c$$

In which x is given input, m is a slop line, c is constant, y is the output variable.

5) Train the model using hyperparameter. Understand the hyperparameter set it according to the model. Such as learning rate, epochs, iterations.

6) As discussed above how the hypothesis plays an important role in analysis, checks the hypothesis and measure the loss/cost function.

7) The loss/ Cost function will help us to measure how hypothesis value is true and accurate.

8) Minimize the loss/cost function will help the model to improve prediction.

9) The loss equation can be defined as a sum of the squared difference between the predicted value and actual value divided by twice the size of the dataset.

10) To minimize the Lose/cost function use gradient descent, it starts with a random value and finds the point their loss function is least.

By following the above we can implement Multivariate regression

Advantages of Multivariate Regression

- The multivariate technique allows finding a relationship between variables or features
- It helps to find a correlation between independent and dependent variables.

Disadvantages of Multivariate Regression

- Multivariate techniques are a little complex and high-level mathematical calculation
- The multivariate regression model's output is not easily interpretable and sometimes because some loss and error output are not identical.
- It cannot be applied to a small dataset because results are more straightforward in larger datasets.

Conclusion- Multivariate Regression

- The main purpose to use multivariate regression is when you have more than one variables are available and in that case, single linear regression will not work.
- Mainly real world has multiple variables or features when multiple variables/features come into play multivariate regression are used.