Logistic Regression

G

Mahankaligayathri · Follow

7 min read · Sep 25, 2023

What is logistic regression?

A classification model called logistic regression employs a number of independent characteristics to forecast a binary-dependent result. It is a very efficient method for determining the connection between information, cues, or a specific occurrence.

Logistic regression attempts to model the likelihood of a particular result using a set of input factors. In logistic regression, the output variable is binary; it can only take one of two possible values (for example, 0 for the event not to occur or 1 for the event to occur).

Then, during model training, the logistic regression can be applied on brand-new input data that the model has never seen before.

In logistic regression, we fit a “S” shaped logistic function, which predicts two maximum values (0 or 1), rather than a regression line.

The logistic function’s curve shows the possibility of several things, including whether or not the cells are malignant, whether or not a mouse is obese depending on its weight, etc.

Why the Name Logistic Regression?

Since the methodology underlying it is relatively similar to that of linear regression, it is known as “logistic regression.” The Logit function, which is used in this categorization method, is where the term “Logistic” originates.

Logistic Regression Equation

The linear regression equation yields the logistic regression equation. The following are the mathematical steps to obtain Logistic Regression equations:

- We know the equation of the straight line can be written as:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

- In Logistic Regression y can be between 0 and 1 only, so for this let's divide the above equation by (1-y):

$$\frac{y}{1-y}\text{; 0 for y= 0, and infinity for y=1}$$

- But we need range between -[infinity] to +[infinity], then take logarithm of the equation it will become:

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

The above equation is the final equation for Logistic Regression.

Sigmoid Function

Logistic Regression can be expressed as,

$$log\left(\frac{p(X)}{1 - p(X)}\right) = \beta 0 + \beta 1X$$

When odds are defined as p(x)/(1-p(x)) and the left-hand side is referred to as the logit or log-odds function. The ratio of the chances of success to the chances of failure is known as the odds. As a result, a linear combination of inputs is converted to log(odds), with an output of 1, in Logistic Regression.

The following is the inverse of the aforementioned function

$$p(X) = \frac{e^{\beta 0 + \beta 1X}}{1 + e^{\beta 0 + \beta 1X}}$$

This produces an S-shaped curve and is known as the Sigmoid function. Always between 0 and 1, it returns a probability value. To translate expected values into probabilities, the Sigmoid function is utilised. Any real number is transformed into a number between 0 and 1 by the function. To convert predictions to probabilities in machine learning, we use sigmoid.

The mathematically sigmoid function can be,

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

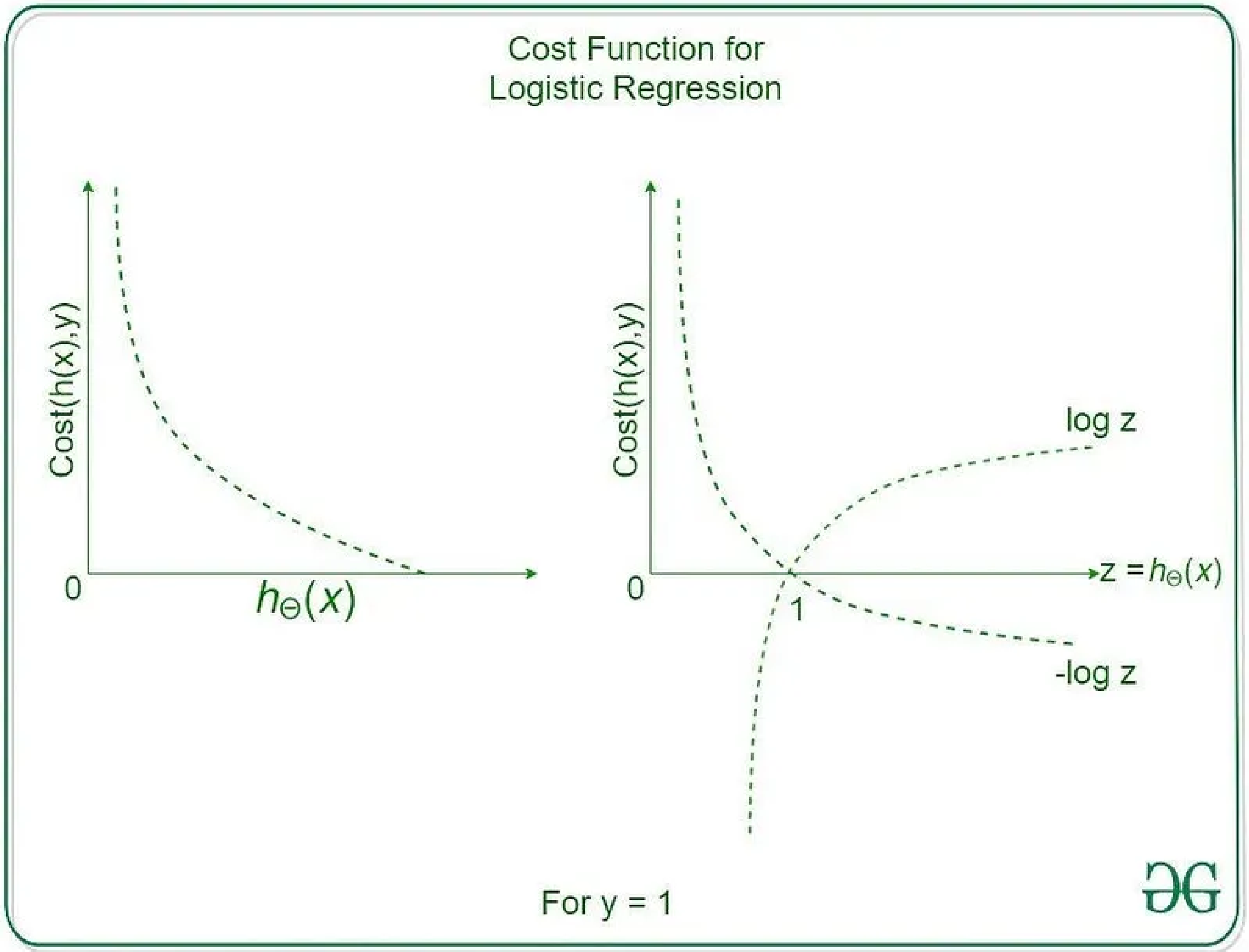
Decision Boundary

To determine which class a data set belongs to, a threshold might be defined. Based on this threshold, several classes are created for the estimated likelihood that is derived.Label the student as passing if the expected result is less than 0.5; otherwise, mark them as failing. Decision boundaries can be either linear or non-linear. The polynomial order can be raised to provide a complex decision boundary.

Cost function

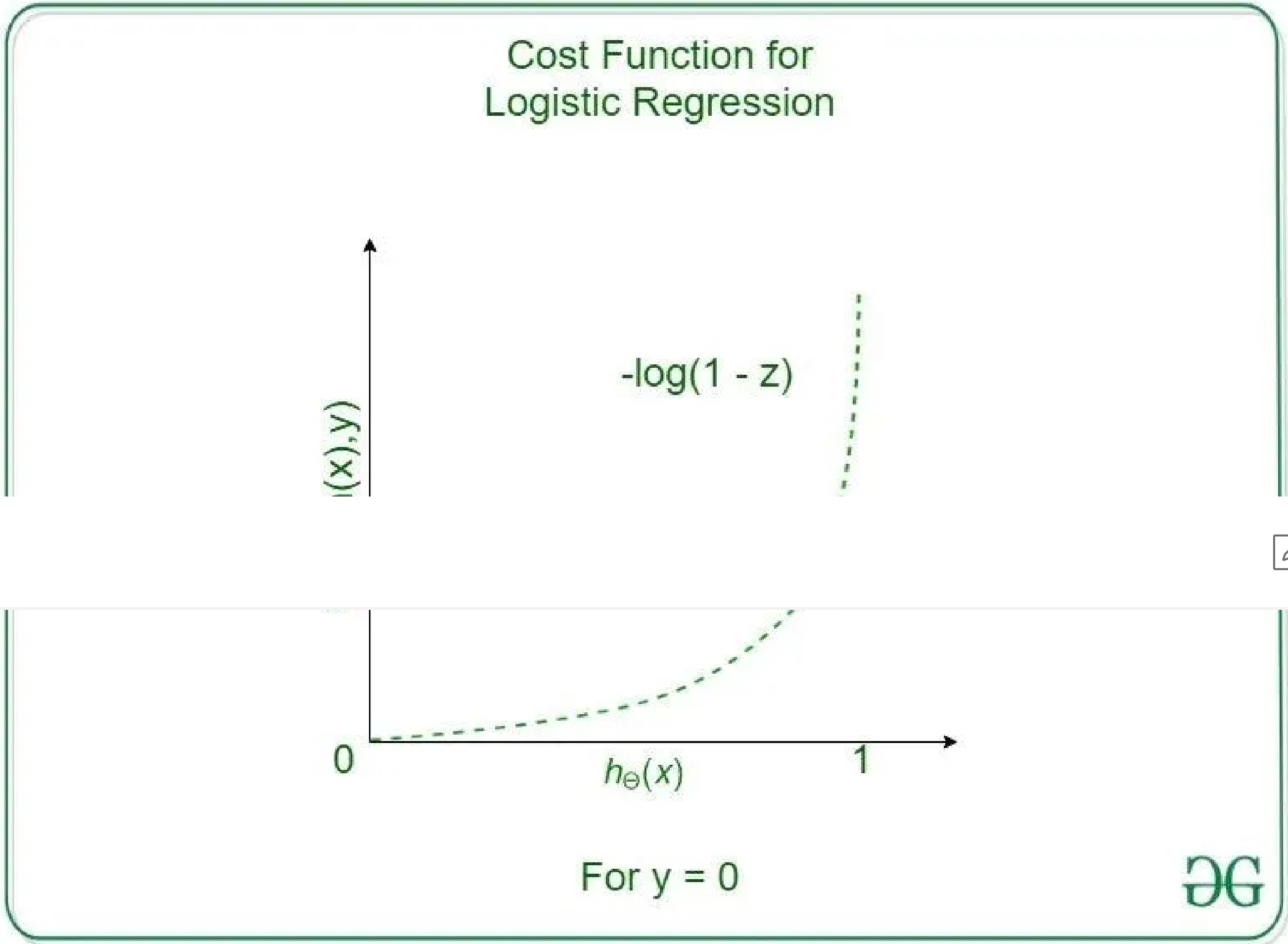
It will lead to the above-described non-convex cost function. Therefore, the cost function we choose for logistic regression is also referred to as the cross entropy or the log loss.

Case 1: If y = 1, then the class's actual label is 1. If the label's projected value is also 1, the cost is equal to 0. However, as h(x) approaches 0 and deviates from 1, the cost function grows exponentially and goes to infinity, as seen by the graph below.



Cost Function for Logistic Regression for the case y=1

Case 2: If y = 0, then the class's actual label is 0. If the anticipated value of the label is also 0, the cost is equal to 0. However, when h(x) moves away from 0 and towards 1, the cost function grows exponentially and goes to infinity, as seen by the graph below.



Cost Function for Logistic Regression for the case y=0

We were able to create a loss function by altering the cost function that penalises the model weights more severely the more the predicted value of the label deviates from the actual label.

Gradient Descent

The cost function can be decreased via gradient descent. The primary goal of gradient descent is to decrease the cost value.