Logistic Regression

What is Logistic Regression and how it works ?

Data science and Machine learning Online Course

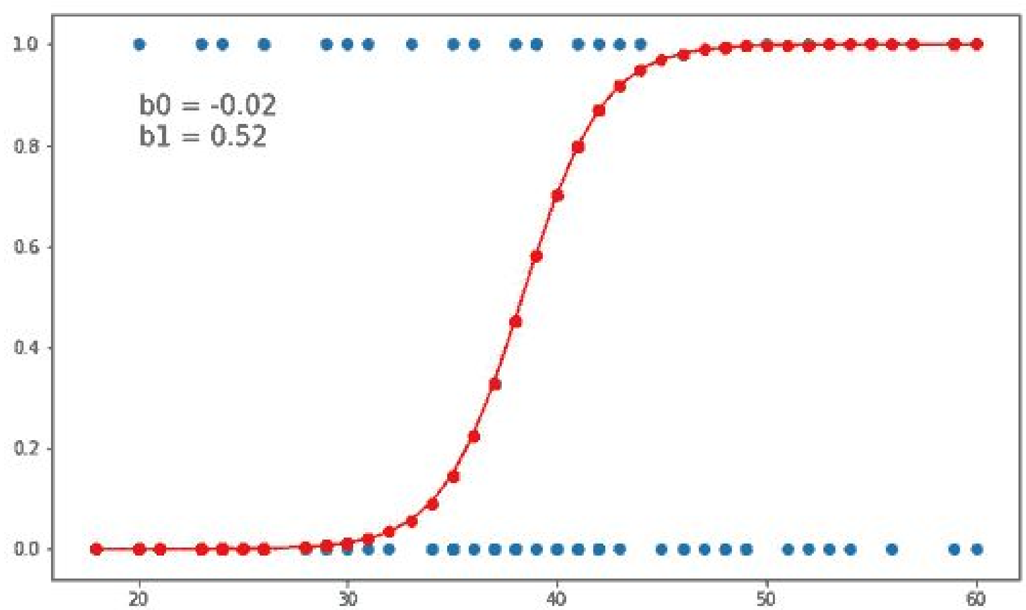
What is Linear Regression ?

Logistic regression is a classification algorithm used to assign observations to a discrete set of classes. Logistic regression transforms its output using the logistic sigmoid function to return a probability value.

What are the types of logistic regression:

1. Binary (eg. Tumor Malignant or Benign)
2. Multi-linear functions (eg. Cats, dogs or Sheep's)

Logistic Regression is a Machine Learning algorithm which is used for classification problems, it is a predictive analysis algorithm and **based on the concept of probability.**



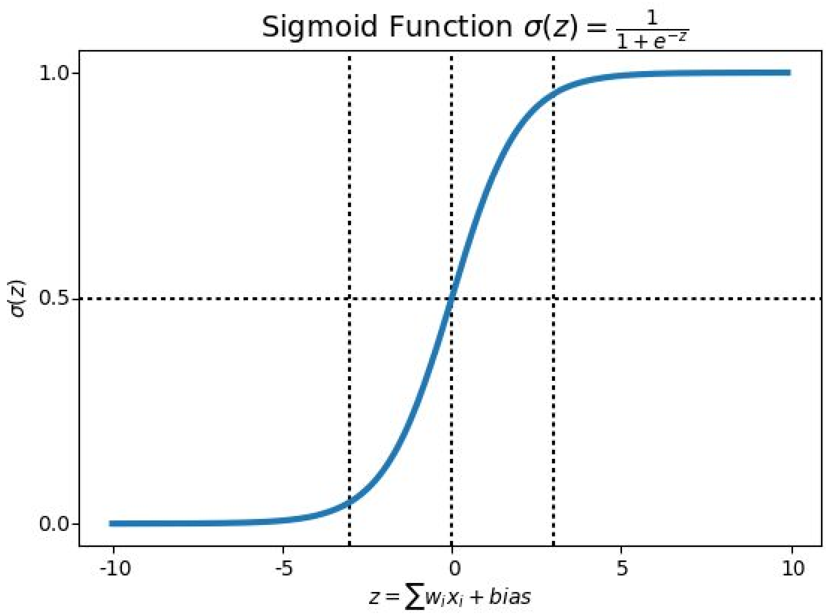
The hypothesis of logistic regression tends it to limit the cost function between 0 and 1

$$0 \leq h_\theta(x) \leq 1$$

What is the sigmoid function ?

In order to map predicted values to probabilities, we use the Sigmoid function. The function maps any real value into another value between 0 and 1.

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



Hypothesis Presentation:

When using *linear regression* we used a formula of the hypothesis i.e.

$$h\Theta(x) = \beta_0 + \beta_1 X$$

For logistic regression we are going to modify it a little bit i.e.

$$\sigma(Z) = \sigma(\beta_0 + \beta_1 X)$$

We have expected that our hypothesis will give values between 0 and 1.

Model Representation:

We have expected that our hypothesis will give values between 0 and 1.

$$Z = \beta_0 + \beta_1 X$$

$$h\Theta(x) = \text{sigmoid}(Z)$$

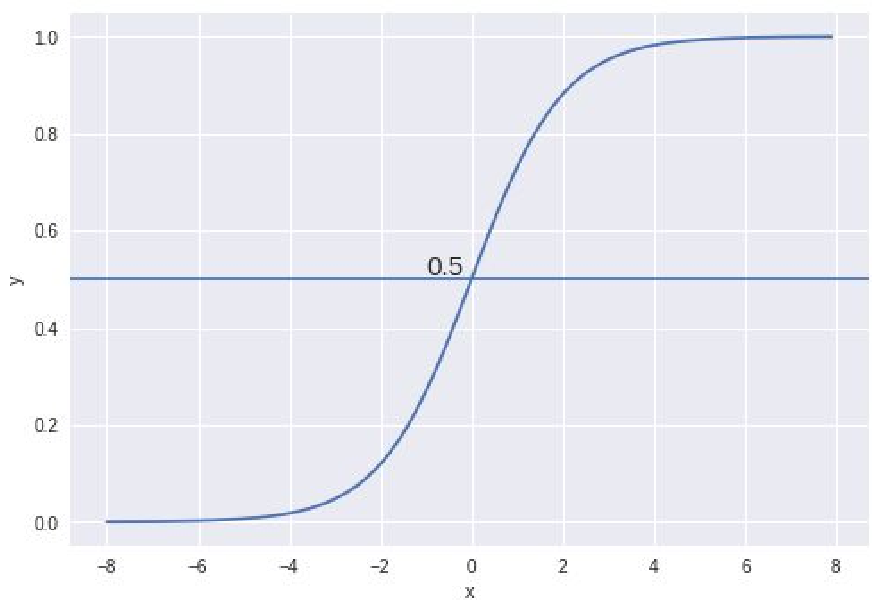
$$\text{i.e. } h\Theta(x) = 1/(1 + e^{-(\beta_0 + \beta_1 X)})$$

$$h\theta(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}}$$

Decision Boundary:

We expect our classifier to give us a set of outputs or classes based on probability when we pass the inputs through a prediction function and returns a probability score between 0 and 1.

We basically decide with a threshold value above which we classify values into Class 1 and if the value goes below the threshold then we classify it in Class 2.



As shown in the above graph we have chosen the threshold as 0.5, if the prediction function returned a value of 0.7 then we would classify this observation as Class 1. If our prediction returned a value of 0.2 then we would classify the observation as Class 2.

Cost Function:

We talked about the cost function $J(\theta)$ in the *Linear regression*, the cost function represents optimization objective i.e. we create a cost function and minimize it so that we can develop an accurate model with minimum error.

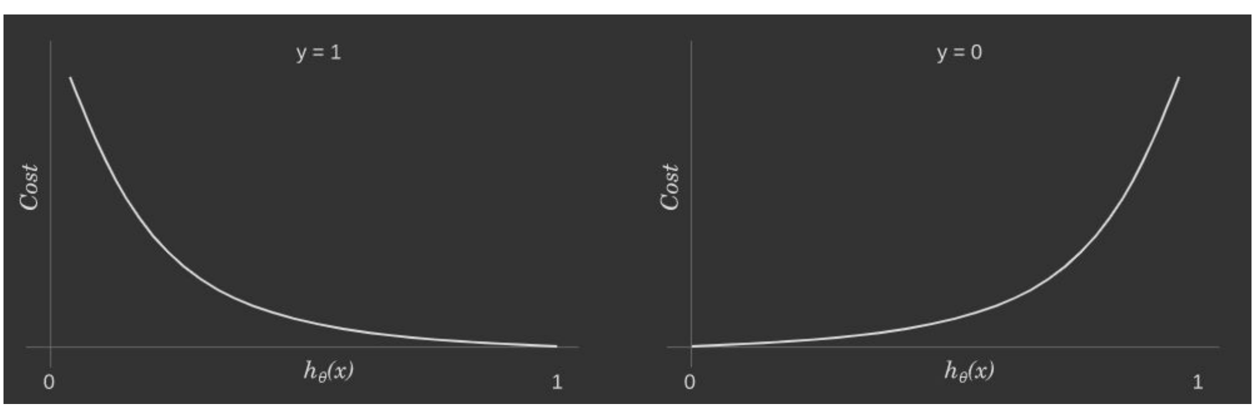
$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2.$$

The Cost function of Linear regression

If we try to use the cost function of the linear regression in 'Logistic Regression' then it would be of no use as it would end up being a non-convex function with many local minimums, in which it would be very difficult to minimize the cost value and find the global minimum.

For logistic regression, the Cost function is defined as:

$$Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{if } y = 1 \\ -log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$



The above two functions can be compressed into a single function i.e.

$$J(\theta) = -\frac{1}{m}\sum\left[y^{(i)}\log(h\theta(x(i))) + \left(1-y^{(i)}\right)\log(1-h\theta(x(i)))\right]$$

Gradient Descent:

Now the question arises, how do we reduce the cost value. Well, this can be done by using **Gradient Descent**. The main goal of Gradient descent is to minimize the cost value. i.e. min J(θ).

Now to minimize our cost function we need to run the gradient descent function on each parameter i.e.

$$\theta j := \theta j - \alpha \frac{\partial}{\partial \theta j} J(\theta)$$

Want $\min_\theta J(\theta)$:

Repeat $\{$

$$\theta_j := \theta_j - \alpha \sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)}$$

(simultaneously update all θ_j)

$\}$

Gradient descent has an analogy in which we have to imagine ourselves at the top of a mountain valley and left stranded and blindfolded, our objective is to reach the bottom of the hill