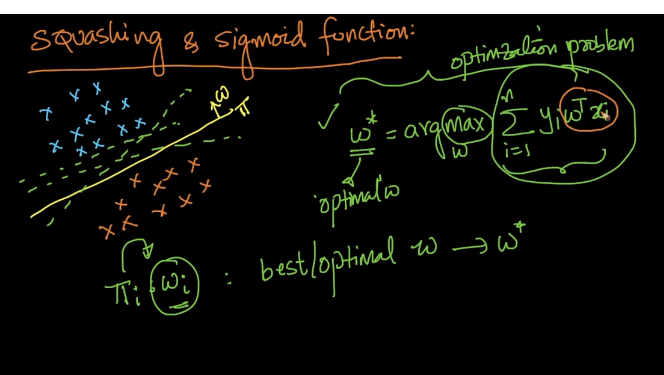
**Sigmoidal Function:**

Till now we have basic equation for Logistic regression and now we will see what are the issues it has and how can we avoid those issues.

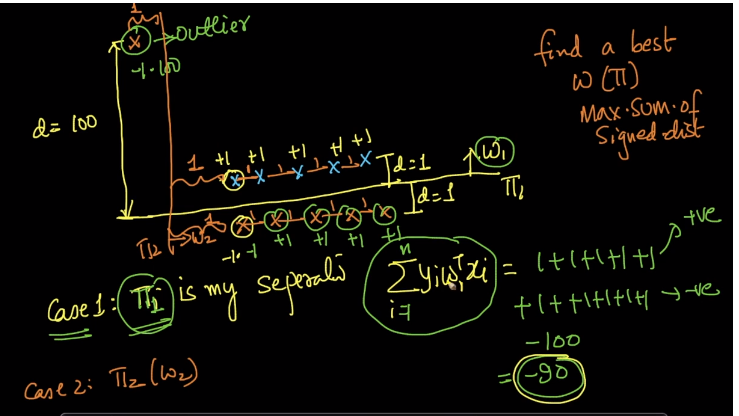
So our motto is to find optimal value for w so as to maximize the sum of signed values.(Yi \* Wt\* Xi)



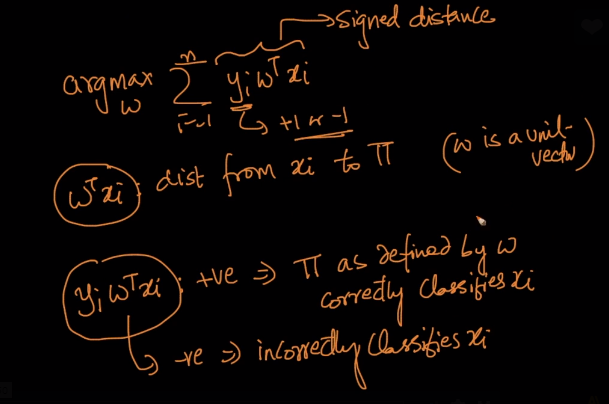
Now lets take an example and see different cases for the same example with different w.

So in case 1 as we can see in below image when we take w1 it is correctly separating 5 positive points and 5 negative points but wrongly classifying one negative point but since it is an outlier the distance of the point from w1 is very high

And so when we compute sum for signed values the value we are getting as final output is an highly negative value.



We are calling the Yi \* Wt \* Xi as signed values because sign for the multiplied value shows us if the value is correctly classified or misclassified.



Now let’s see case 2 :

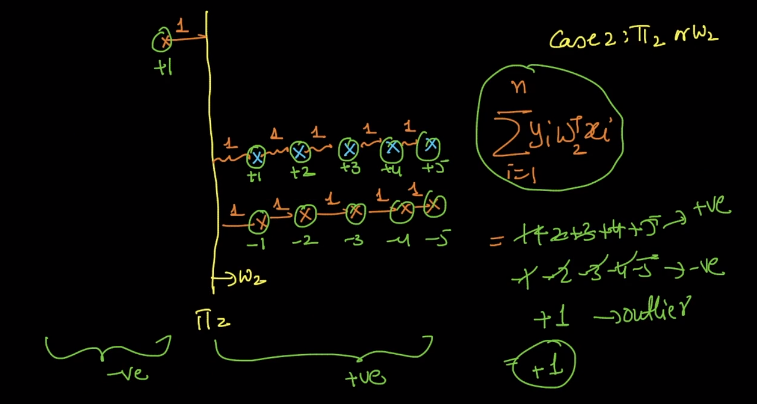
As we can see in below image when we take w2 it is correctly classifying 5 positive values and 1 negative value but misclassifying 5 negative points

But still the sum of all the signed values is an positive value.

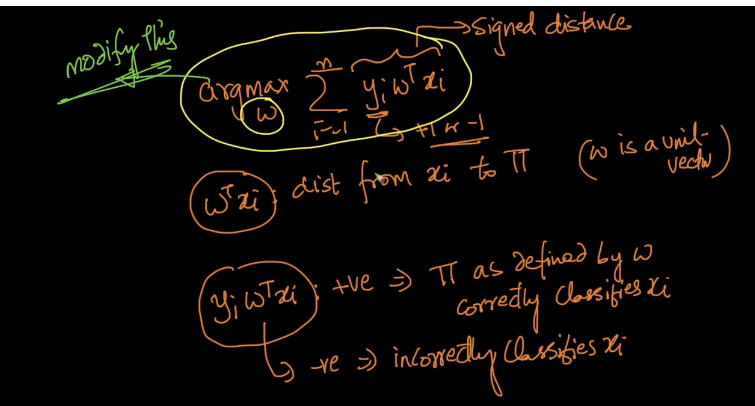
Now if have to choose between w1 and w2 we will say w2 is better classifier whereas it is totally wrong because accuracy for w2 is way less than w1.

And all this is because of the outlier value.

SO the question is how we handle such outliers?

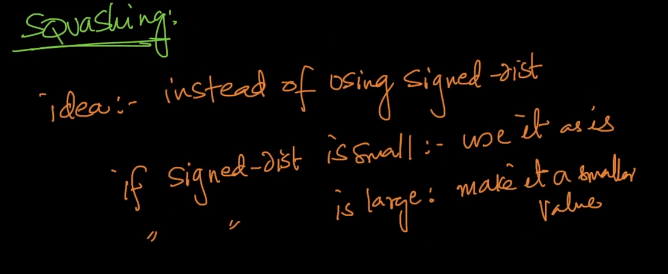


So we need to modify our formula in such a manner that it automatically handles such outlier and gives us better results.



So the technique we use to do this is squashing.

Squashing means instead of using actual signed distance for all the values lets set a benchmark or limit above which we won’t use actual distance but use minimized distance.



So to perform Squashing we take help of sigmoidal function.

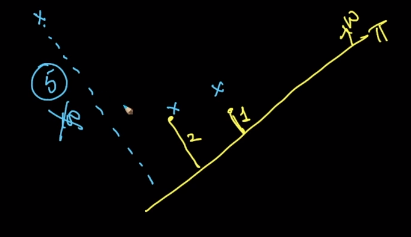
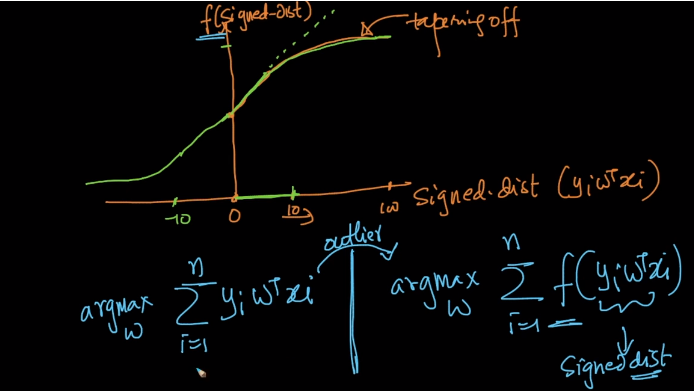
What is sigmoidal function?

It is a function which actually performs very linear till an extent and after which it just taper the value for the coordinate.

SO what we can see in below image is that we don’t want don’t want high distance values as it is but we want squashed or tapered values.

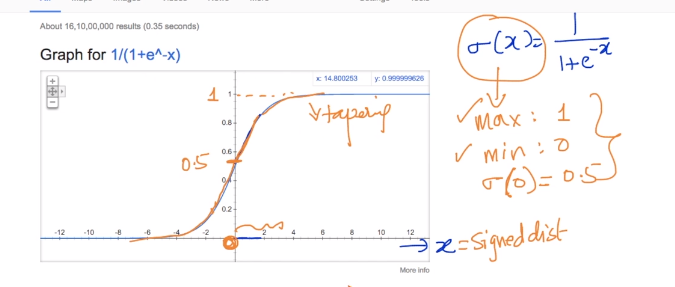
So instead 5 or above values say we will use 4 as distance.

So for this we need to make some changes in our actual formula as shown in image below.



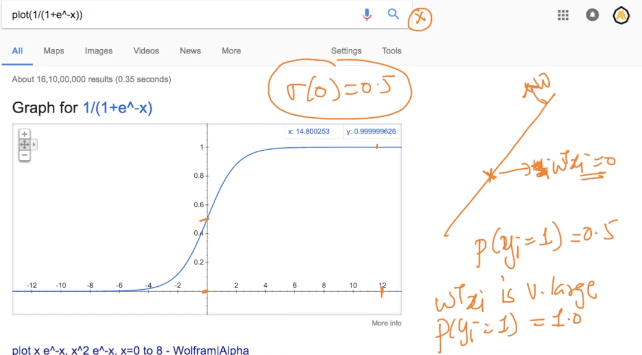
So below is the image of an sigmoidal function and it clearly shows a trend what we need.

At an extent i.e. below Y =1 it is behaving as linearly increasing function but after it, it almost gets constant and that’s what our need is.

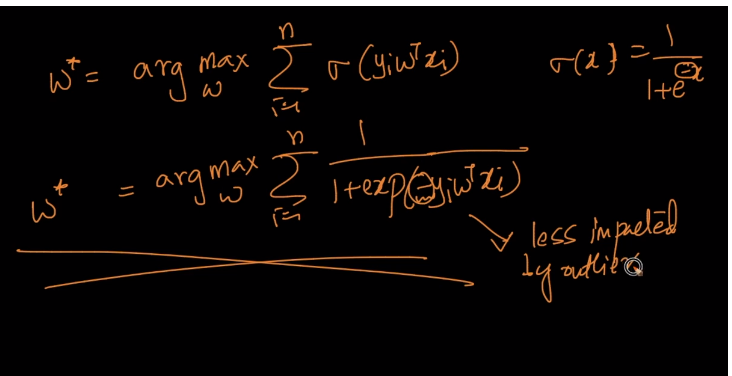


Sigmoidal function is said as (**Sigma)** and Sigma(0) = 0.5, i.e. say we have an query point exactly on our plane as shown in below image than what is its probability that it is positive or negative,

It is 1/2 of course and that’s what sigmoidal function is giving us.

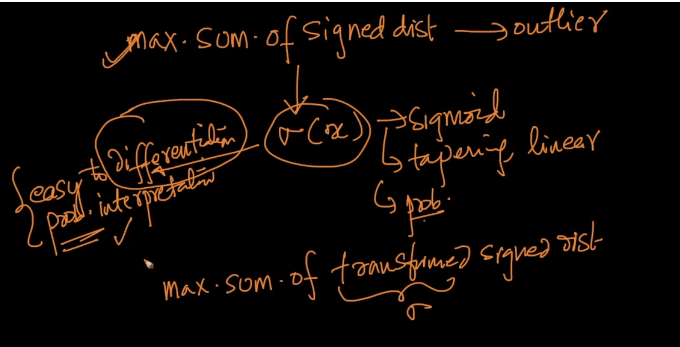


So by getting satisfied by all the constraint we get our corrected formula for Linear regression as:



So why are we using Sigmoidal function when there are other function also which can satisfy our requirements.

SO it is because it is easily differentiable in first place and why differentiation is important in real case studies, we will see this in further chapters.



Comments:

