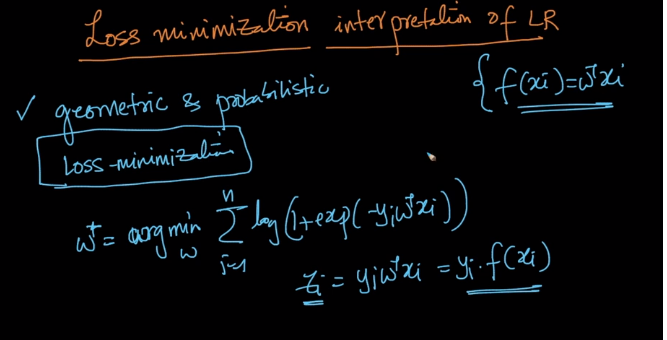
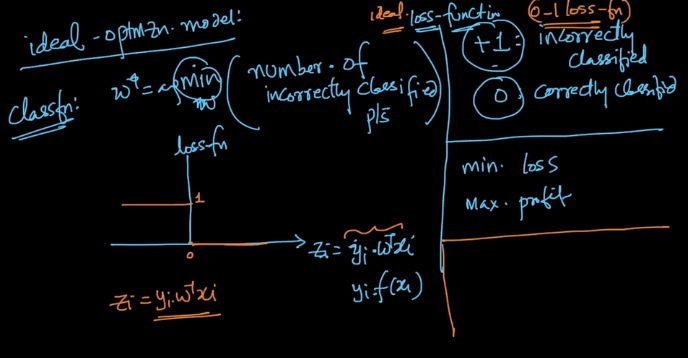
So till now we have seen two approaches to derive LR i.e. Geometric interpretation and another is Probabilistic approach.

Now lets see third approach i.e. Loss minimization interpretation.



So above image is just our optimization problem

So lets take a condition where we will see how ideal loss function behave.



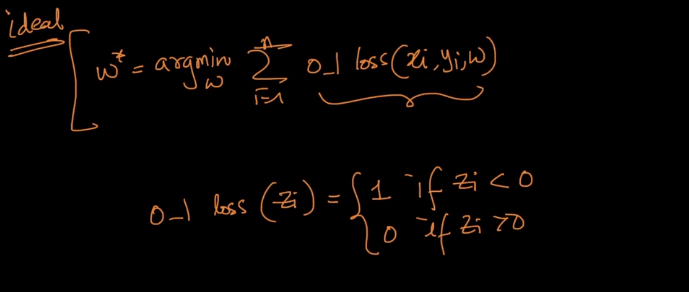
So in above image as it is shown that we assume +1 for incorrectly classified points and 0 for correctly classified points.

So our objective is to minimize the loss and maximize the profit i.e. minimize the number of incorrect classification and maximize the correctly classified point.

And so how our ideal loss function will look like in is shown in above figure.

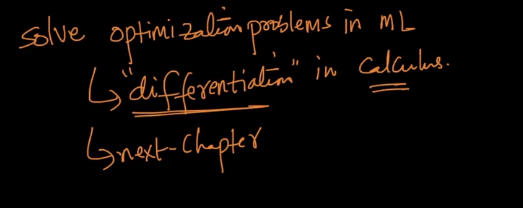
It will be 0 for all the positive Zi and for negative Zi it will be 1 (considering the case for ideal Loss Function).

And such a function is called 0-1 loss function.



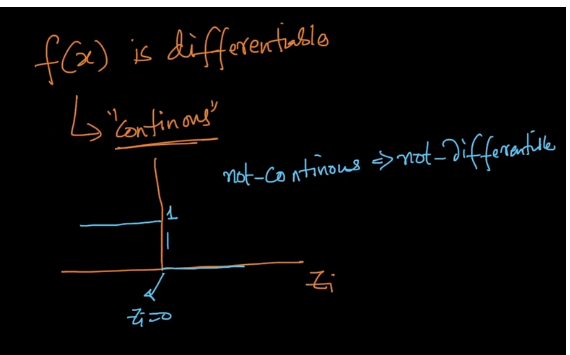
So this is how our modified problem looks like and how o-1 loss(Zi) looks like is shown in above image.

So we do much of the part in ML using differentiation in calculus and if function is not differentiable than we cant do much about the problem and why is it so? That we will learn later on in next chapter.

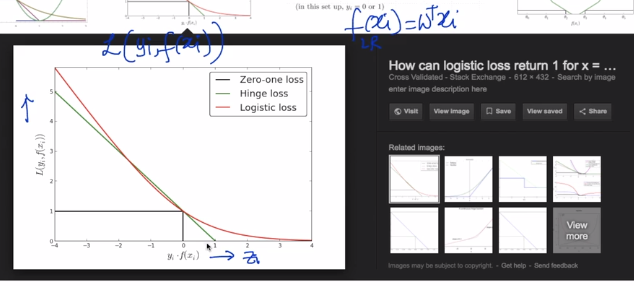


And Since our 0-1 loss function is not continuous it means that it not defined because a function to be differentiable it needs to be continuous.

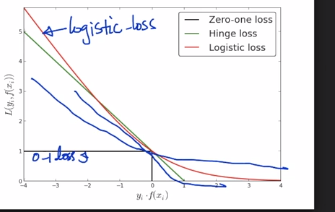
But at Zi = 0 it is not defined.



Since we cant do much with the function what we can do?

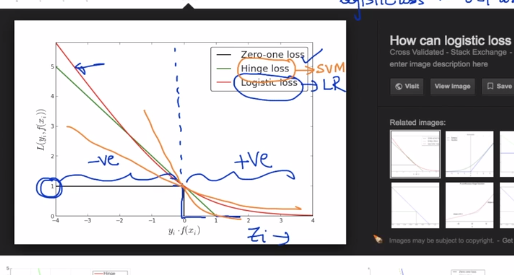


We will make approximation and one such approximation is Logistic loss.



SO the red line in above image is logistic loss.

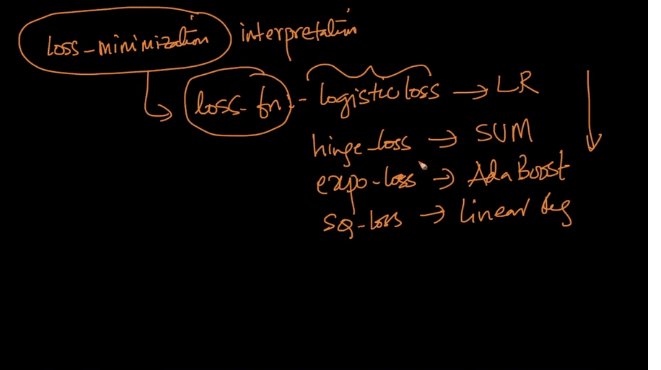
There can be other approximations also as shown with blue drawn lines but we are using Logistic loss.



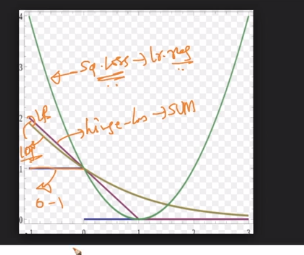
So logistic loss is not one of the best approximation as we can see in above image on the positive it is not giving the value as 0 like 0-1 loss function or on negative side either it is not giving me value of 1.

So as we said there can be hundreds of approximations and one another is shown in image above is Hinge Loss which results in another beautiful algorithm called SVM.

So this is the beautiful thing about Loss function that by changing the loss function we get the different algorithms as shown in below image.



So in one loss interpretation diagram we can represent all the algos.



Comments:

