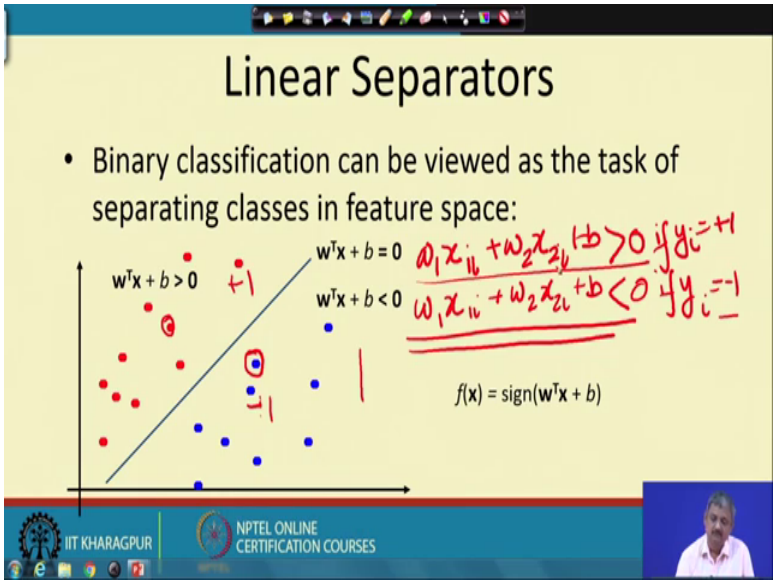
Data Mining
Prof. Pabitra Mitra
Department of Computer Science & Engineering
Indian Institute of Technology, Kharagpur

Lecture – 23
Support Vector Machine – II

Let us continue our discussions on the linear discriminate. So, let us see this picture.

(Refer Slide Time: 00:27)

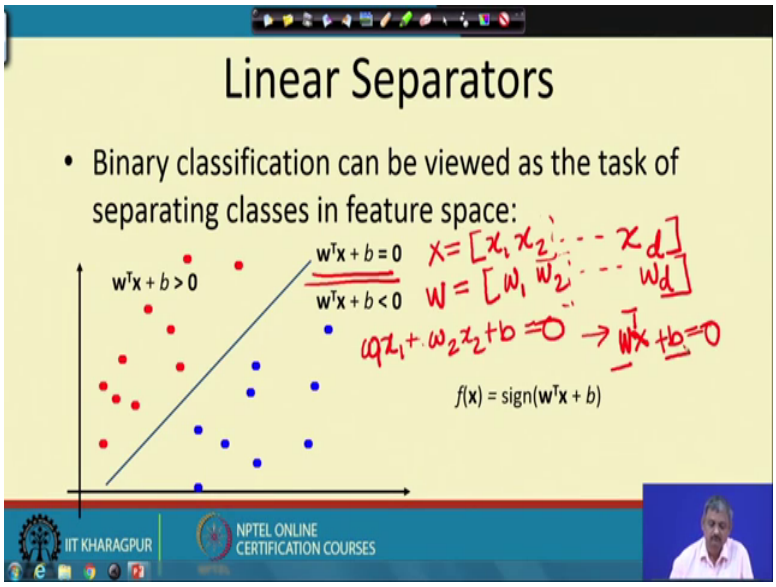


So, I have 2 classes, yeah I am marked them by red and blue points, and this blue classes have a y value, sorry blue classes have a y value of, blue classes have a y value plus one and the red classes have a y value, sorry it is the other way round. So, I have marked with other way round minus one. I will introduce, so if we recollect in our last discussion we said that, w 1 x 1 i plus w 2 into x 2 i plus b should be greater than 0, if y i is plus 1 and w 1 x 1 i x 2 i plus b should be less than 0, if y i equal to minus 1.

I think all of, all of it is clear. So, what I am basically meaning is that, I should pick up a value of w 1, w 2, b such a way that, if I pick up any point whose class level is plus 1 y is plus 1 it should be on the non origin side; that means, it should give this value to be greater than 0. So, if I pick up any of this red point and compute plug in the x 1, x 2 value here to this value of w and b, I should get positive, should lie on the positive side which will happen only when it lies on the positive side.

Similarly, if I pick up a blue point which has y equal to minus 1, it should give a negative value. So, what I actually want to say is that, if I look at the sign of this quantity, positive or negative, that should match with the class level; if I call that as f x sign up this that should match with the class level. So, in general I am writing here x 1, x 2, in general and there can be more than a component. So, there can be x 3, x 4, x 3 a d dimensional anything it can be, right. So, let me introduce a little bit of vector terminology.

(Refer Slide Time: 03:15)

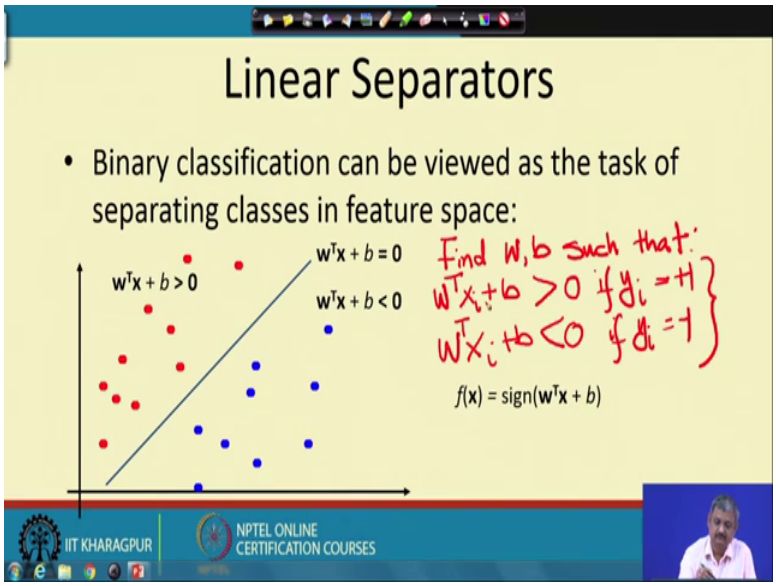


Let us say, the input vector is like this. And, I call something as the weight vector or W vector, which is nothing but each of these weights.

Then another way of writing this equation, the way of writing this equation 0 is this, you see, if you take the W vector, in this case for 2 dimension there will be only this much X will be only this much. If you take this 2 dimension then this is nothing but the transpose of the W vector into the X vector. So, I can write down this as w transpose, W vector transpose into X vector plus b equals 0. So, that is the format I have convention, I have used in denoting this line, denoting this line ok.

So, what I am supposed to do is, to find out a value of w, find out a value of b it gives you a good discriminate. So, let us see how to do that.

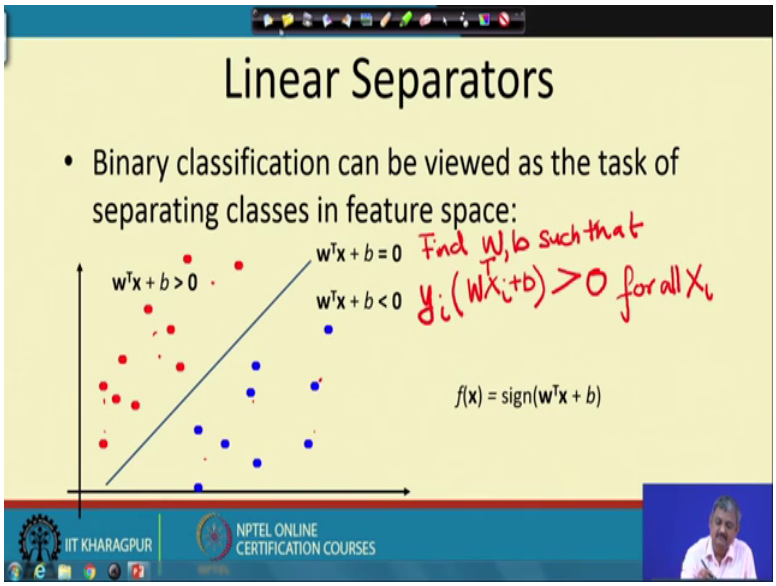
(Refer Slide Time: 05:17)



So, up to this is clear this vector notation, and origin side and non origin side. So, let me see what I do next, what I do next is the following. I sort, sorry I combine these 2 as a condition, sorry not this, not this. I combine these 2 as a condition, I say that find. So, the first condition I like down will basically tell that all plus y are in non origin side and minus y are in origin side.

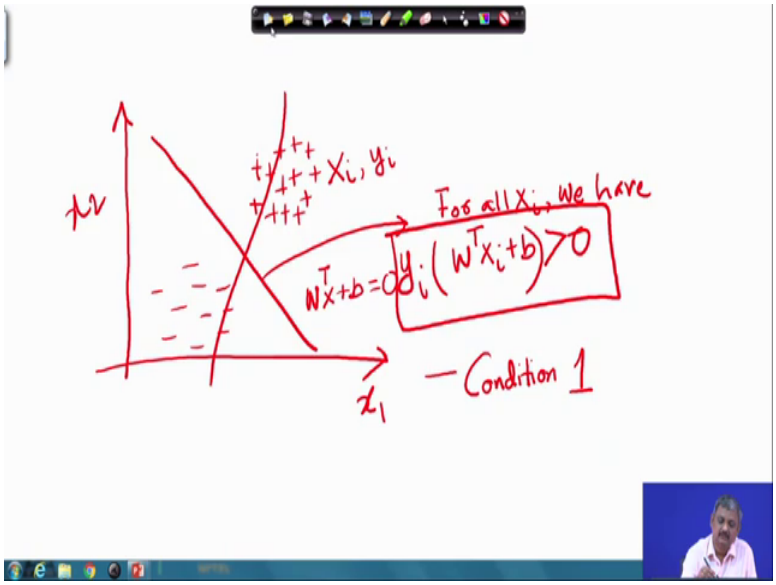
So, I tell the first condition on w and b requires that, if you just look at it carefully, I can in fact, combine these 2 into 1 single condition which says that, the sign of y and the sign of this quantity should be same should be same. Let me write that down.

(Refer Slide Time: 07:17)



And this would hold for each and every X i, for each and every X i this would hold. So, let me see what this means actually, let me draw the picture and I tell you.

(Refer Slide Time: 08:09)

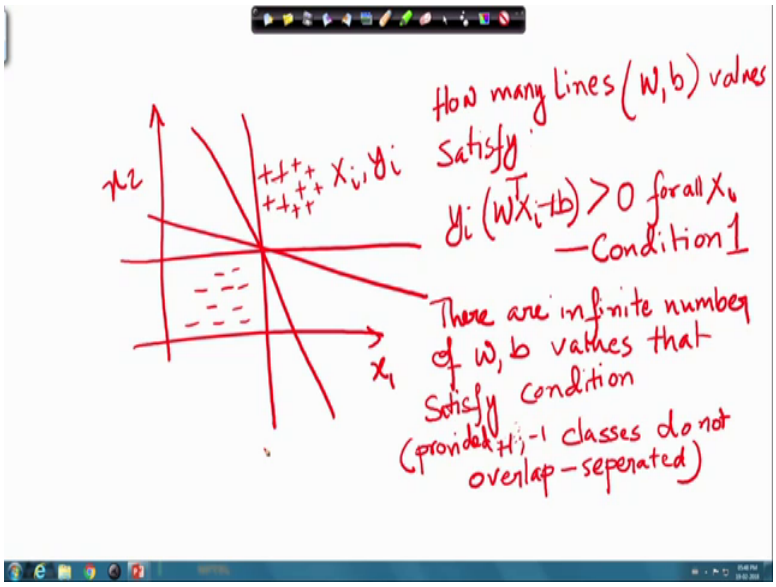


W is the weight vector, X is the data vector. This actually means that all positive points should lie on one side negative point.

So, this line satisfies this. This line, suppose I take some other bad line like this, which is not separating, this will not satisfy this, this points will fail, for all points it will not satisfy. So, this is sorry, do not copy this, right. This is the condition, sign of y i and this

quantity should match. So, this I will call as condition 1. So, what it says actually, it the lines i, the points i have plus and minus point. They put some constant on what are linear discriminates that we can choose, what are linear discriminates I can choose ok, so that it correctly classifies. So, but now let us look at 1 thing, fine, I want to satisfy that.

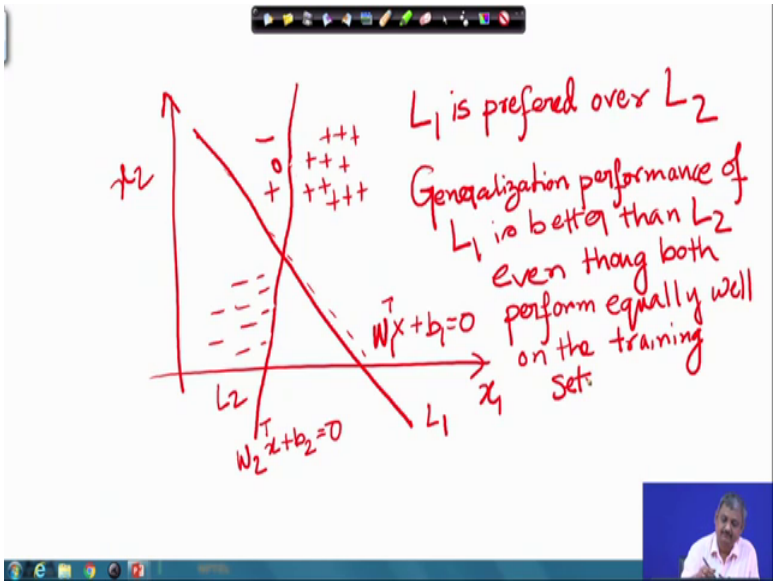
(Refer Slide Time: 10:41)



So, I basically ask you the question that how many lines you can draw? so, how many W and b values you can pick up so that this condition is satisfied for all the x axis even in that training set? let us see. So, this draws satisfies definitely, correct classification. This satisfies, this satisfies, this satisfies. There are many lines which satisfy. In fact, provided we are separated, they are separated, I can actually have infinite number of lines which satisfy this.

So, what I want to do is to sort of pick up one of this infinite line, which I consider is good. Let us see how do I pick up one among this infinite line, let me again draw it in a thin clutter free.

(Refer Slide Time: 14:00)

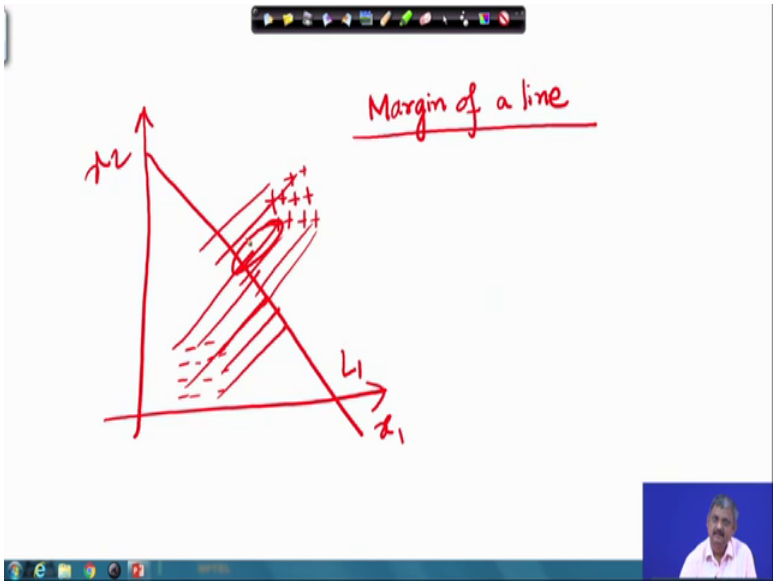


So, you can imagine there are many infinitely many line like this, like this, like this, many thing I can draw. I have to choose one. So, which one will you choose? Which one will you choose among these possibilities? Maybe your answer will be the one that passes through the middle, why let us see.

Why you are wanted to choose that. Let us compare 2 lines, one passes through the middle I call it as, W 1 transpose X, b 1 equals 0. And another just touches, call it as. This is L 1, this is L 2. I maybe, I do it slightly bad, maybe this point is not there. That does not touch, similarly these points are not. Of course, you will say that L 1 is preferred over L 2, the middle line is preferred over this just corner kind of line. Why? Both of them are correctly classifying, both of them are satisfying condition 1, if you remember condition 1, both of them are satisfying one side other side. Why this is better? Why L 1 is better than L 2? the answer is, for these given points they are fine, but our goal is to classify a new point ok.

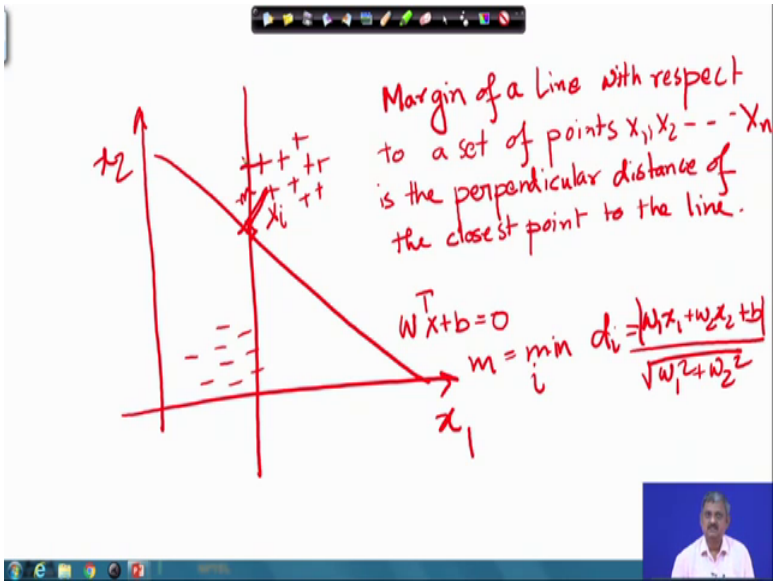
So, suppose I have a point like this, a point like this. So, L 2 would put it into minus class, which is wrong, this is most probably plus class. Whereas, L 1 will still put it in a plus class, is this correct. So, even though on the training set they are equal, on a test set L 1 is better, middle one is better. So, let me try to quantify or see which is the quantity, which L 1 is better in terms of which L 1 is better than L 2, let me draw again.

(Refer Slide Time: 18:53)



I hope I am not boring you, this repeated redrawing, but a little help the same old picture, one line L 1. I define a quantity called Margin of a line, the final quantity I will define it. What is the margin of a line? You take the closest point to that line. How do you find out the closest point? Draw perpendicular from every point to the line. See which point has the closest smallest perpendicular distance, let us say this point has the. That perpendicular distance is the margin of a line. So, you will let me redraw again.

(Refer Slide Time: 20:18)



Zero and so the margin is, take the closest point, drop a perpendicular, and this the perpendicular distance I would call margin m.

So, let me define. Clear, so if I, it is very easy to write it mathematically, if I take a point X i, can you define what is the. So, do you remember; what is the perpendicular distance from X i to that line? let me call it as d i from a point X i to a line W transpose X plus b it is w 1 x 1 plus w 2 x 2 plus b absolute value of that, positive value of that, divided by. In this is the perpendicular distance and margin is nothing, but the maximum, minimum of this, over all i. Now you see that the center line, now if we will consider the margin say, the corner line will have a much smaller line, a margin than the central line. So, basically I will, I want a line with a high margin. I want the line with the maximum margin.

So, I stop here today. In my next class, I will explain how to get the line with the highest margin which will give me the equation of the line.

Thank you.