

As in classification, output was either +1 or -1, but in Regression output is a real valued number.

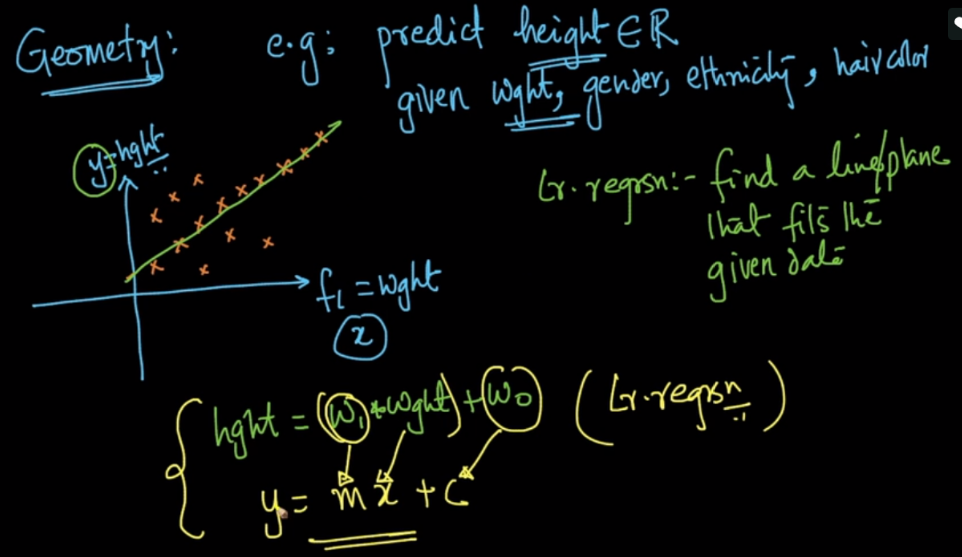
So let’s say we’ve have feature weight and we’ve to predict height. Now in training data we would find a line which would best fit according to the relationship which minimizes the error between the predicted value and the actual output value.

The At any given weight we can find it’s corresponding height using equation of line.

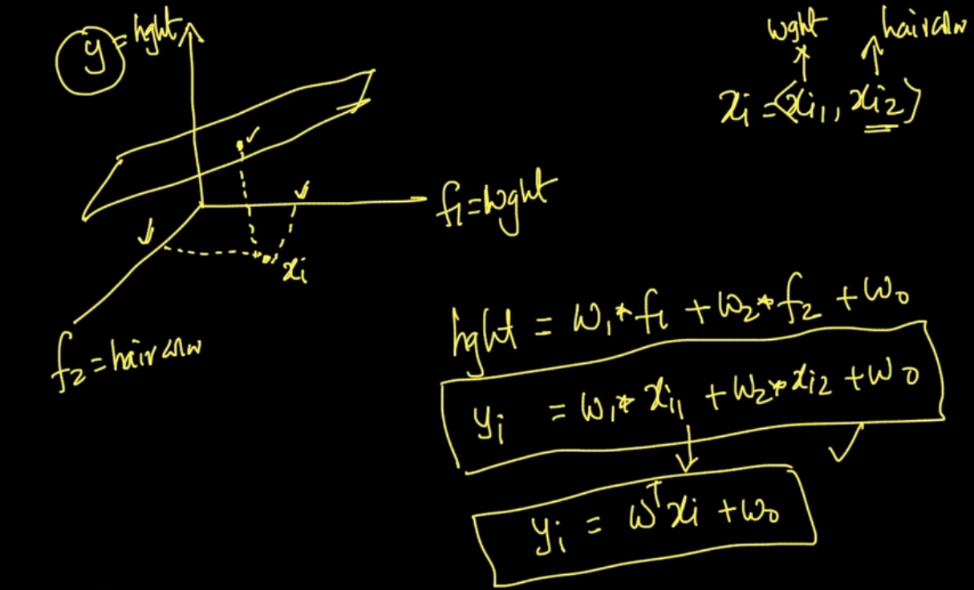
Height = w1 \* wght + w0.

Which is equivalent to y = mx + c.

Here w1 is slope and w0 is intercept over y(height) axis;



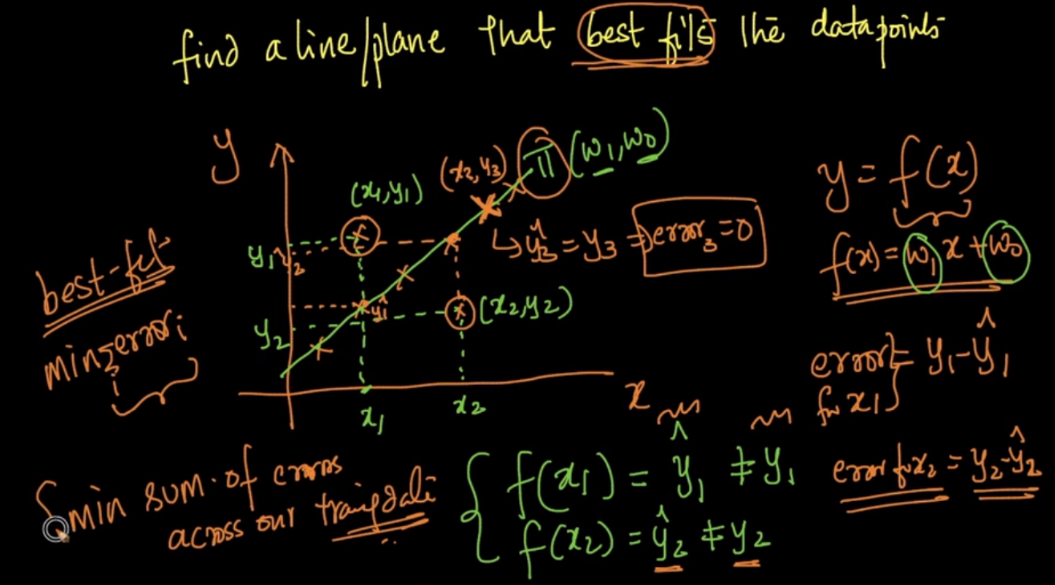
Now what if we have more than 1 independent features. In such case there would be a hyperplane, which would give the value of y. Equation for such hyperplane is given in below image.



Since our ultimate aim is to find the line/plane which best fits the datapoints. What does best fit means ?

Let’s say actual o/p value is y1 and predicted is y1^. Then difference between them will be error.

So here we will find the line which will minimizes the error for all point or we can say which minimizes the sum of errors across training data.



**Comments:**

