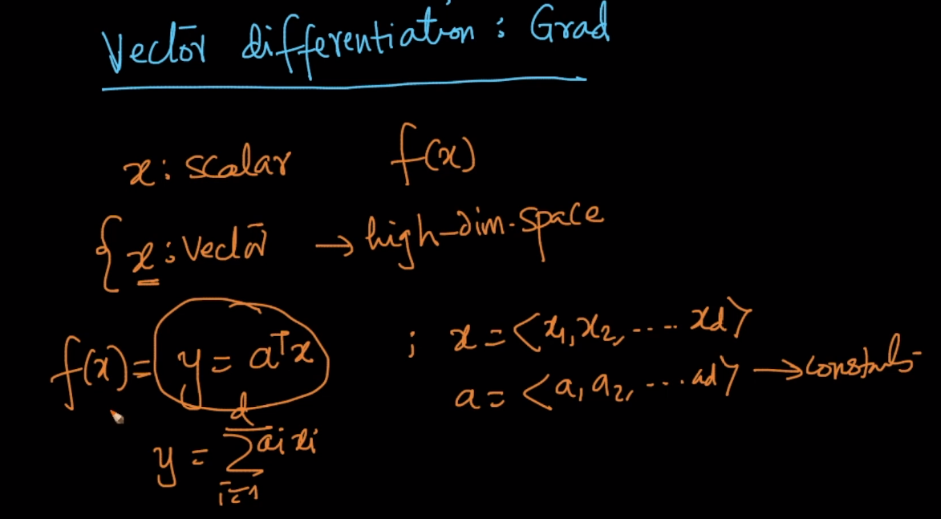
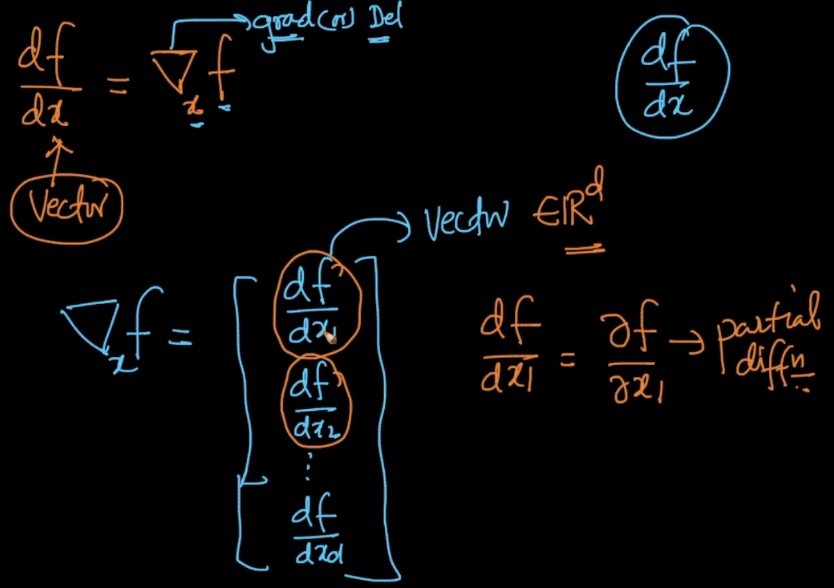
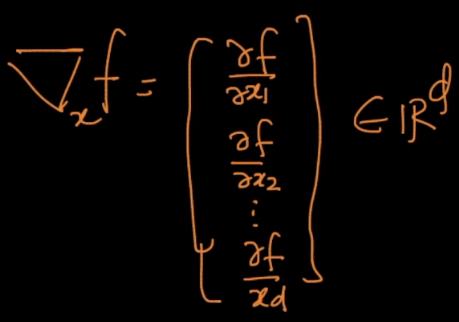
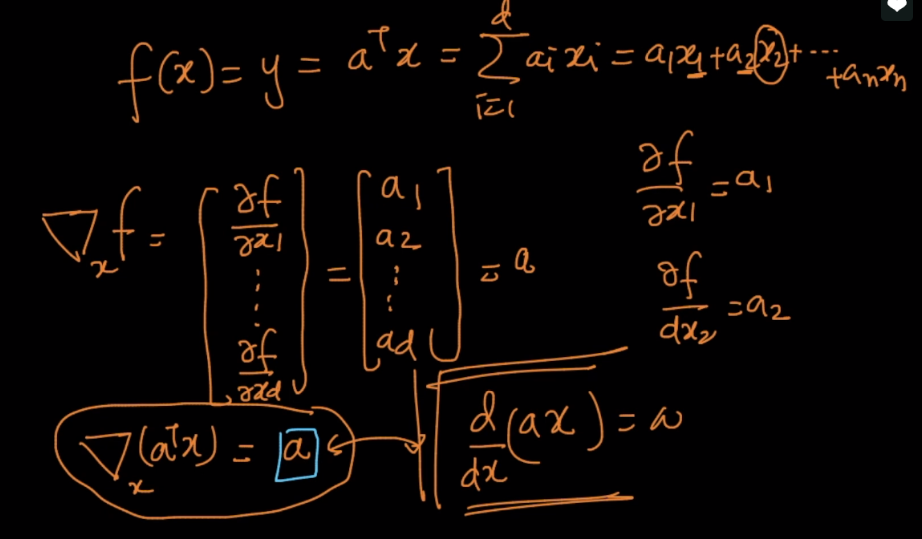
Till Now we’ve seen differentiation performed on scalar, but in ML we have vectors. And performing differentiation is very difficult that’s why we use techniques like gradient descent to find maxima and minima.

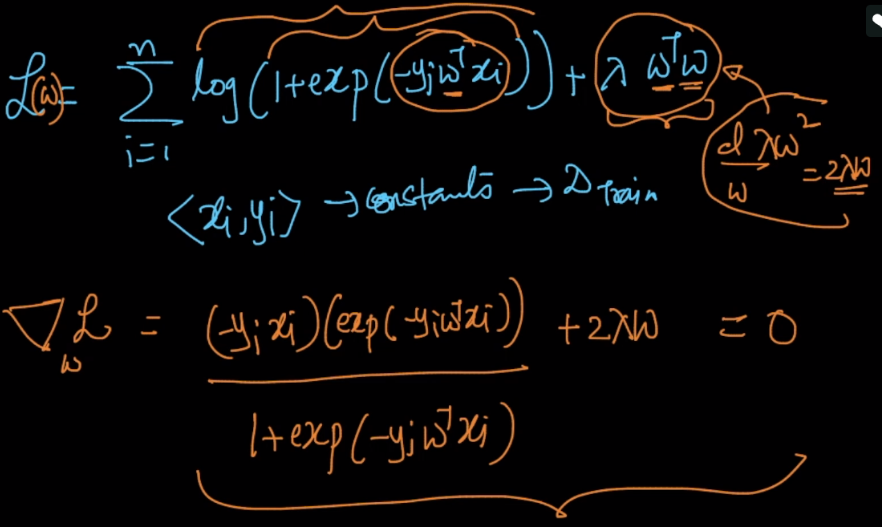


If we are differentiating with respect to a vector, then result will also be a vector in which each value will be the partial differentiation with respect to each value of vector.









Now differentiating the equation we obtained above with respect to W which is vector is very difficult, therefore we use gradient descent.