ML Algorithms

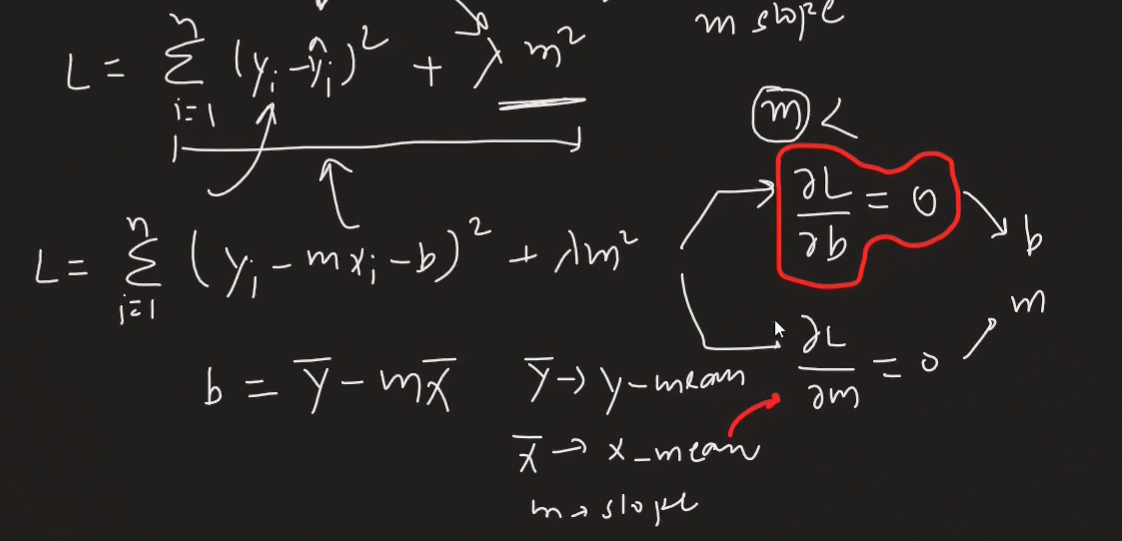
Ridge regularization proof:

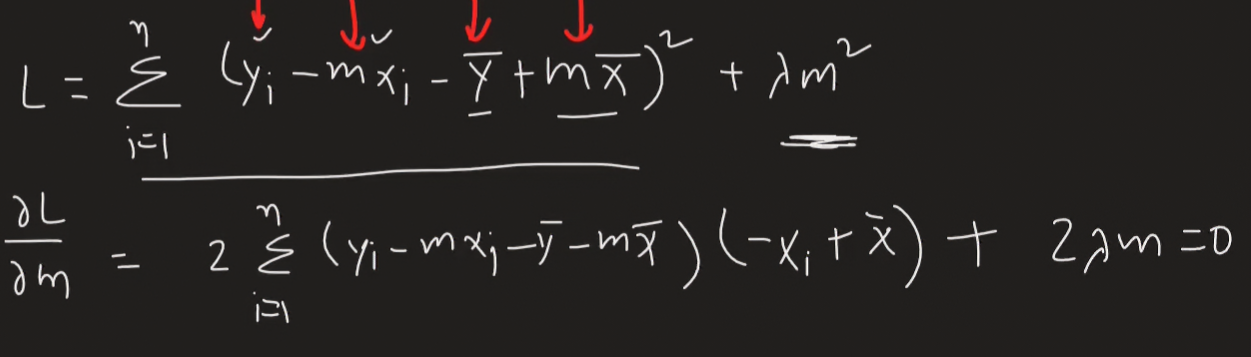
Dl/Db is same as we proved in previous algos as there is no change

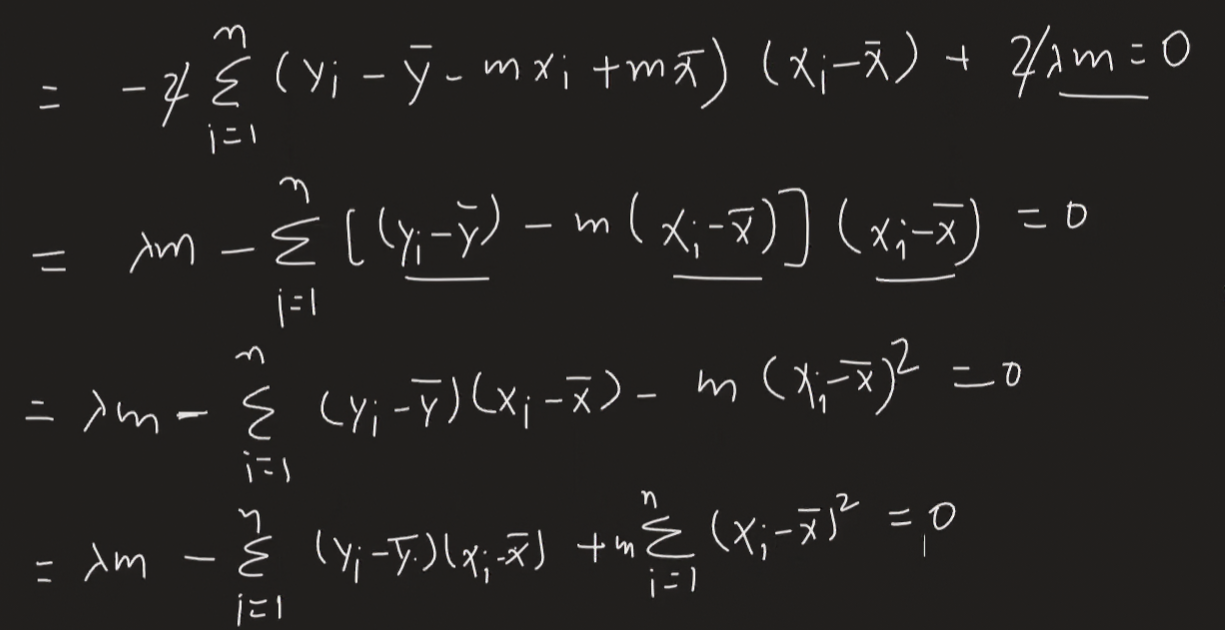
And b comes out to be:

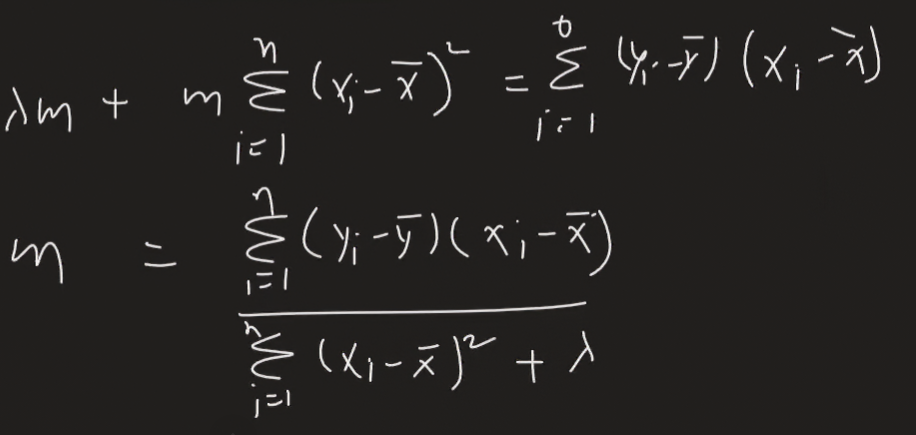
b = Ybar – m(Xbar), here Ybar = y mean and Xbar = X mean

Will focus on Dl/Dm:



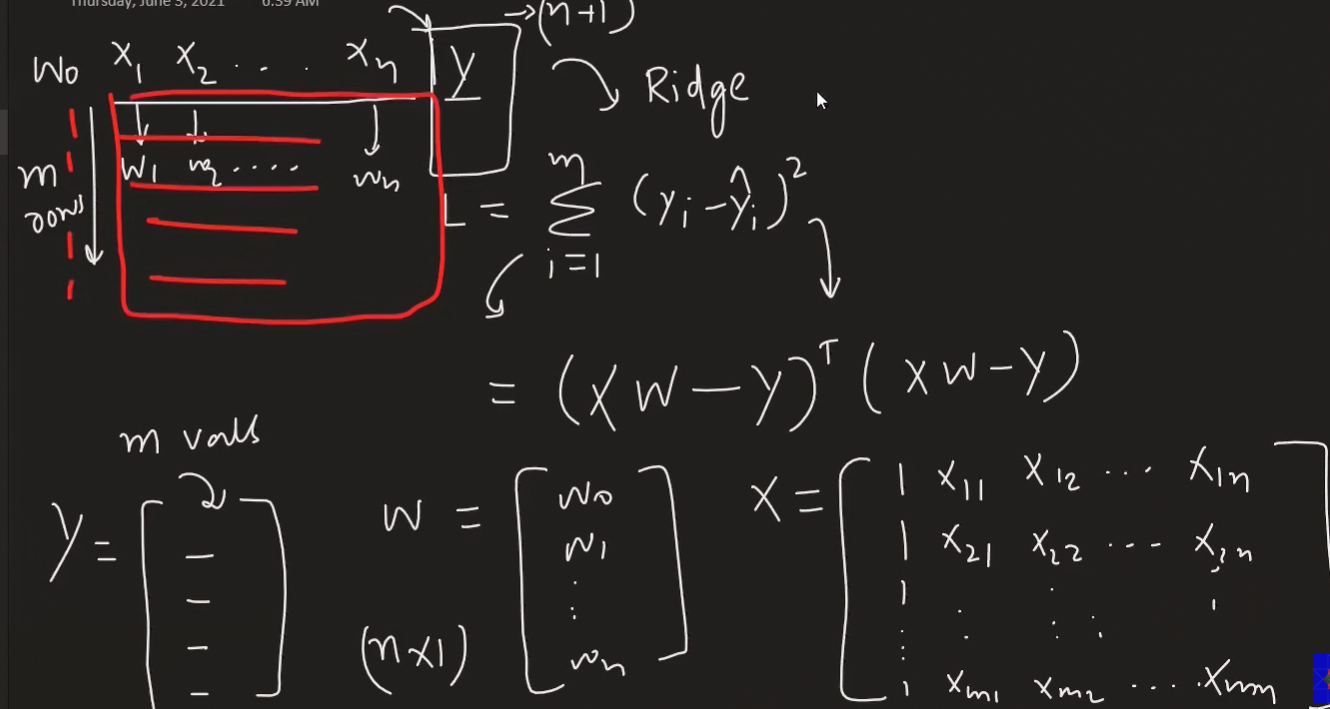






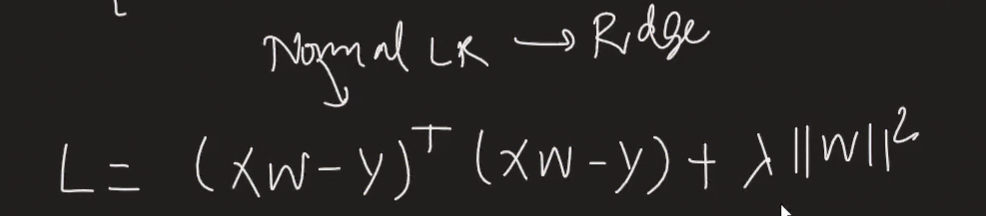
Here m is inversely proportional to lambda.

Ridge Regularization for nD data



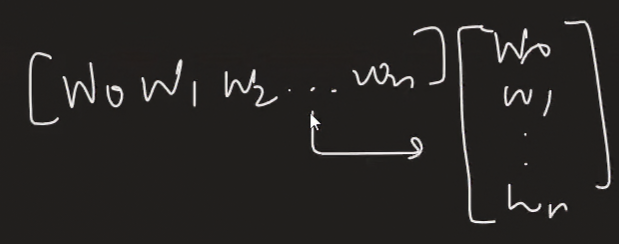
This is the formula et \* e

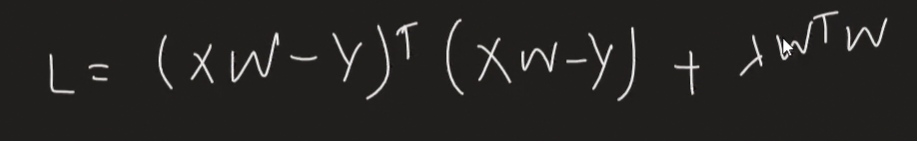
Adding Ridge regression hyperparameter to it

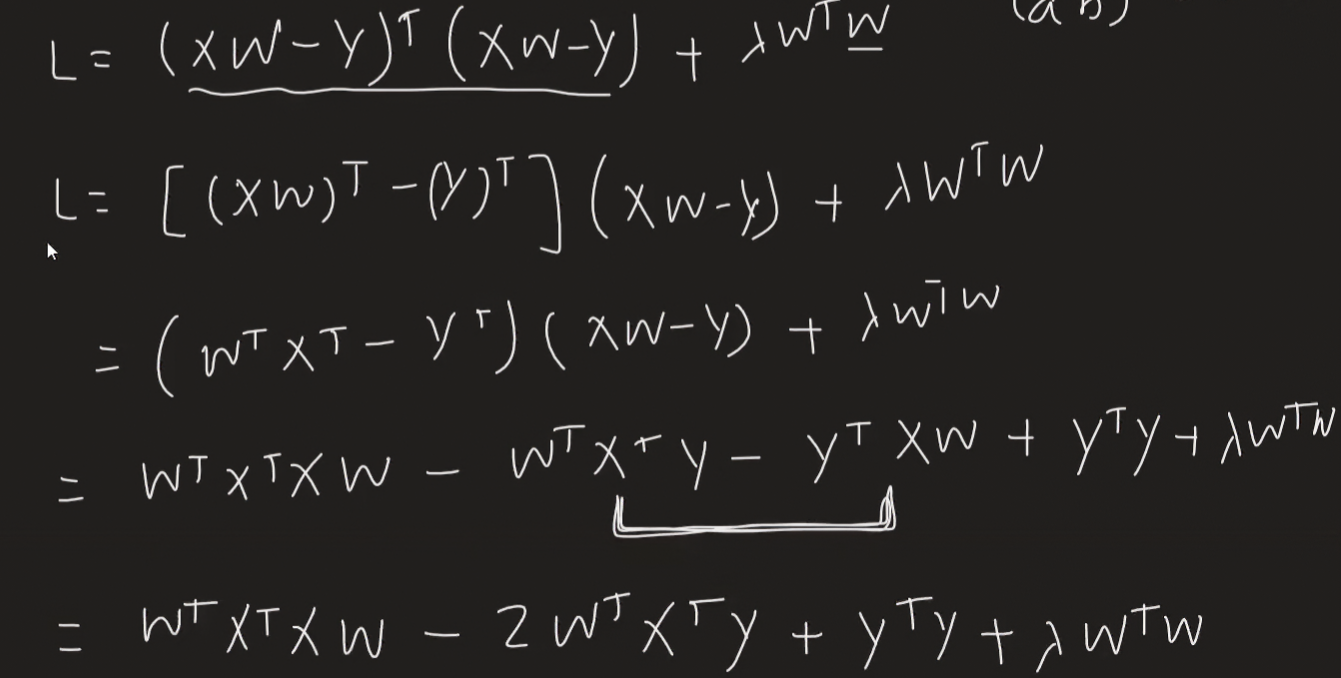


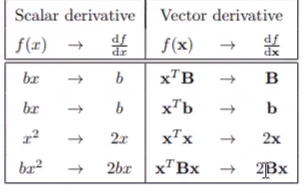
Matrix of w2 can be written as = wt \* w

As it represent the below figure



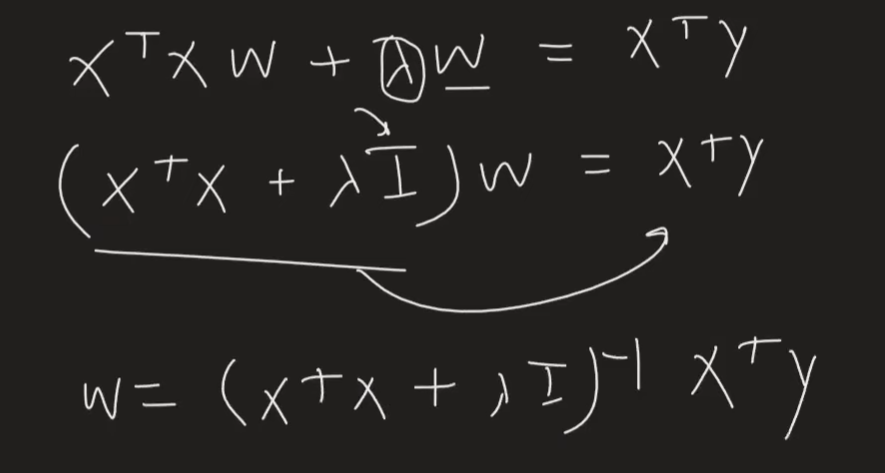




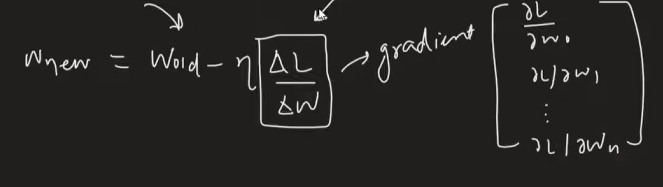


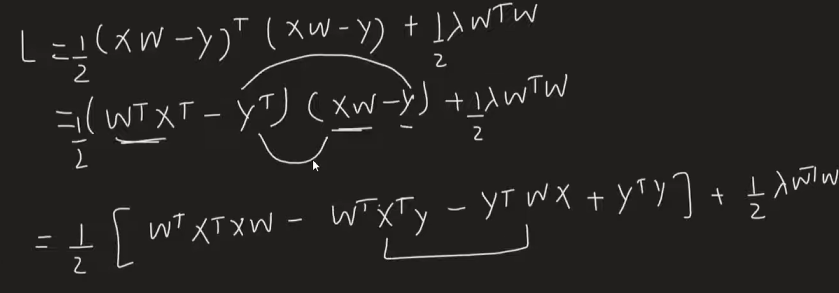
These are some formulae for vector differentiation on upcoming proof

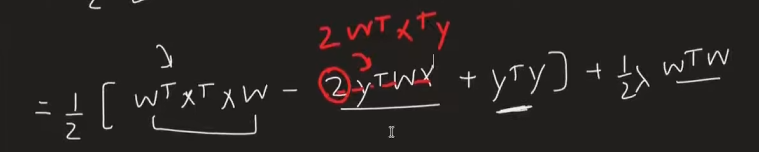


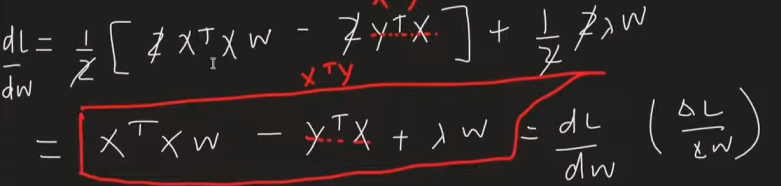


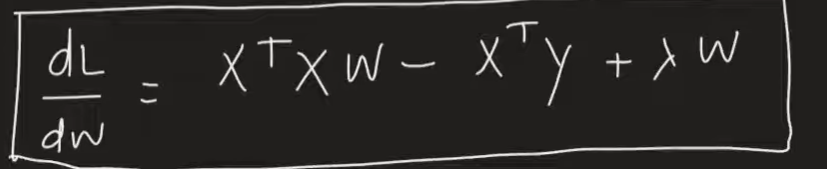
Ridge Regularization using GD:



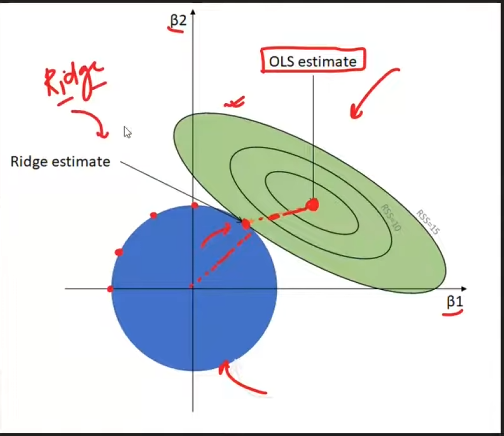








Why this is called ridge :



**Ridge regression is called “ridge” because the optimal solution lies on a smooth circular ridge formed by the L2 constraint boundary.**

**Final mental picture**

|  |  |
| --- | --- |
| **Element in image** | **Meaning** |
| Green ellipses | Same RSS (error) |
| Red dot | OLS solution |
| Blue circle | Ridge constraint (L2 norm) |
| Touching point | Ridge estimate |
| Circular boundary | Ridge |