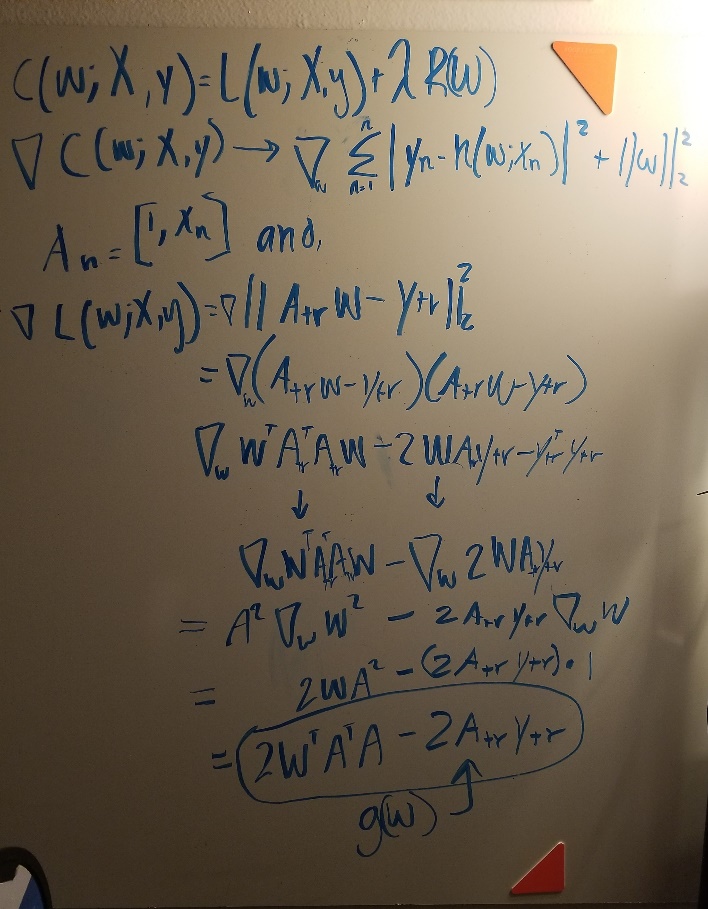
AI EXP : HW3 : Linear Regression

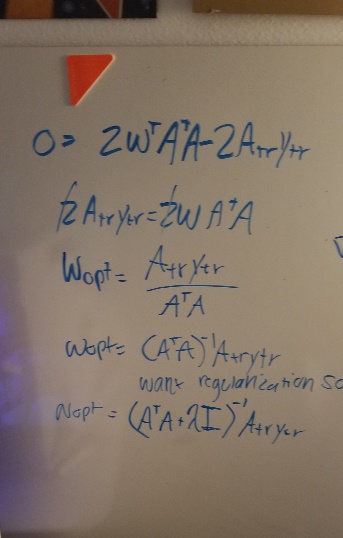
Alex Lamarche

**Problem 1: Setting up the functions**

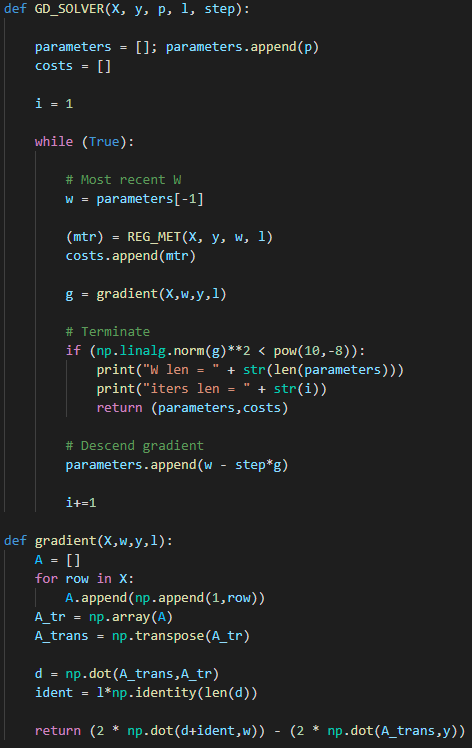
1.A - Scripts included or check the GitHub link here : <https://github.com/ajl3545/AI_HW3>

1.B – Derivation of g(w) the gradient function

1.C Derivation of g(Wopt) the gradient function whose value is: g(w) = 0



1.D – Gradient Algorithm presentation and effects of α and τ:

*Discussion –* α represents the *step* that must be taken towards the minimal value of the computed gradient. Each of the w values changes with the gradient descent iterations - towards an optimal set of parameters. By subtracting the previous set of w parameters by the current gradient “direction,” computed by gradient(W,x,y,z), the algorithm gets closer to a solution that minimizes the cost function. Step is a multiplier that is used to get closer to the minimum of the gradient function. Increasing the step could lead to overfitting and would overshoot on the gradient.

τ, a threshold per se, is used to compute how close the parameters get us to an optimal regression line. First, we step towards the minimum, then we check how close we are. τ ensures that we are close to the minimal value. If the summation of the squares of the parameters (the cost at that parameter iteration) is lower than τ, then that means the most recent w parameters have- as closely as possible and within reason – calculated a minimal cost.

\*ignore the print statements\*

**Problem 2 – Applying functions from problem 1 to testing and training data**

**2.1**