Homework

January 27, 2022

HW-1) Apply Newtons method to steepest-descent algorithm to the optimal step size η, and check how many iterations are required for convergence

$$\mathbf{w}^{new} = \mathbf{w}^{old} + \eta \mathbf{X}^T(\mathbf{t} \text{ - } \mathbf{X}\mathbf{w})|_{\mathbf{W}=\mathbf{W}^{old}}$$

HW-2) Suppose you are experimenting with L_1 and L_2 regularization. Further imagine that you are running gradient descent and at some iteration your weight vector is $\mathbf{w} = [1, \epsilon] \in R^2$ where $\epsilon > 0$ is very small. With the help of this example explain why L_2 norm does not encourage sparsity i.e., it will not try to drive ϵ to 0 to produce a sparse weight vector. Give mathematical explanation.

HW-3) Till now we have been considering a scalar target t from a vector of input observations $\mathbf{x}$. How do you extend this approach for regressing a vector of targets $\mathbf{t} = (t_1, t_2, ..., t_p)$. Derive the close form solutions and write sequential update equations using SGD.