

CMSC-478 Machine Learning - Spring 2019
Homework Assignment 3

Due at the start of class on February 28th

1. The standard logistic regression algorithm, if run for many iterations, finds a weight vector that separates the instances and then keeps increasing the magnitude of the weight vector. Explain why the weight vector keeps increasing even after the instances are classified correctly.
2. The standard logistic regression algorithm finds the weight vector \mathbf{w} that maximizes $\sum_i \log(p(y_i|x_i, \mathbf{w}))$. An algorithm known as ridge regression maximizes $\sum_i \log(p(y_i|x_i, \mathbf{w})) - \gamma \|\mathbf{w}\|^2$ where $\gamma \geq 0$ and $\|\mathbf{w}\|^2$ is $\sum_j w_j^2$.

Derive an expression for the overall gradient for ridge regression that is analogous to the gradient for normal logistic regression as shown on slide number 52 in the lectures notes.

Explain how ridge regression helps overcome the problem illustrated by the first question in this homework.