Michael Lee

CSE 446

HW1

1. Let T mean “Tested Positive”

Let H mean “Actually have the disease”

* 1. a)

=

=

b)

In our case,

To account for the constraint , we use the method of LaGrange multipliers. Hence, we subtract the function to account for our original constraint and solve by taking partial derivatives.

Since

So,

Plugging in:

3. 1) To estimate we need and

2) To estimate we need

3)

By symmetry of the definitions between x and y we attain the expression for y:

4)

=

] -factoring terms out of sum

5) Two examples where it would be beneficial to do in bath versus: 30

4.1 Lasso Regression

1.

1. The error on the training set will decrease because the focus of the algorithm will be to minimize the residual sum of squares.
2. The error on the testing set will be large because the model will over fit to the training set.
3. will have large coefficients because there is no penalty for the magnitude of coefficients.
4. All the elements of will be non-zero

2.

a) The error on the training set will be large because the

b)

c) The coefficients of will be small because the function will penalize large coefficients.

d) There will be many non-zero elements of

4.2 Ridge Regression

3. When ridge regression pushes further away and when lasso pushes further away.

5 Programming Question

Code and plots are all part of the zip, each figure is labelled appropriately.

6. How to selecting lamda

7. Max coefficient: 0.07741645, with label ‘PctIlleg’

Least coefficient: -0.064852304, with label ‘PctKids2Par’