# Q1)

## Part a)

HHH

## Part b)

HHH

## Part c)

HHH

# Q2)

## Part a)

In order to create less complex (parsimonious) model when you have a large number of features in your dataset, some of the Regularization techniques used to address over-fitting and feature selection are:

1. L1 Regularization (Lasso Regression)
2. L2 Regularization (Ridge Regression)

**L1** adds “*absolute value of magnitude*” of coefficient as penalty term to the loss function.

Cost Function:

It shrinks the less important feature’s coefficient to zero.

**L2** adds “squared magnitude” of coefficient as penalty term to the loss function.

Cost Function:

The difference between these techniques is that L2, unlike L1, does not reduce the coefficients to zero. Also, by using L2 we will have a more complicated model in comparison to L1.

## Part b)

Initial guess:

Second xxx:

# Q3)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| n |  |  |  |  |  |
| 10 | 189 | 561 | 12521 | 4173 | 35721 |

## Part a)

## Part b)

## Part c)

MLE

## Part d)

MLE

## Part e)

MLE