HW #4 - Regularization and Validation

By Marco

A. Comparison of errors 'in-sample' and 'out-sample' for each value of lambda (0.001, 0.01, 0.1, 0.25, 0.5, 0.75, 1):

```
Lambada Parameter: 0.001
Error in of sample: 0.023072%
Error out of sample: 0.0232%
Lambada Parameter: 0.01
Error in of sample: 0.022983%
Error out of sample: 0.0232%
Lambada Parameter: 0.1
Error in of sample: 0.023072%
Error out of sample: 0.0224%
Lambada Parameter: 0.25
Error in of sample: 0.023116%
Error out of sample: 0.0236%
Lambada Parameter: 0.5
Error in of sample: 0.023427%
Error out of sample: 0.0236%
Lambada Parameter: 0.75
Error in of sample: 0.023427%
Error out of sample: 0.0236%
Lambada Parameter: 1
Error in of sample: 0.023605%
Error out of sample: 0.0236%
```

B. Choice of best value of lambda:

```
The minimum Error out of sample is: 0.0224%
The Error in of sample is: 0.023072%
The best value of parameter is: 0.1%
```

C. Discuss effect of lambda on the non-linear transform weights

When lambda is decreasing, the risk of overfitting is increasing. when the lambda is increasing, the chance of underfitting happen is increasing. underfitting occurs when lambda is too large, because the learning algorithm has too little flexibility to fit the data. So, according to the results I have gotten, when the value of lambda which is between 0.001~0.1 is the best choice for regulation. In addition, as you decrease the optimization pays less attention to the penalty term and more to Ein, and so Ein will decrease

D. Plot of final classification curve using best lambda and corresponding average weight vector:

Blue represent the digit 1, and green represent the digit from 0 to 9 without 1. Red represent the points which is misclassified. In this case, the best lambda is 0.1 and I took average of weight vector to find classification of data points.