Yuan Bowei
Machine Learning
October 11, 2015

Machine Learning Report 2

2. d) There is a method to create such features with 48 one-dimensional distributions. As is noted by the covariance matrix, we can see that these 48 dimensions are independent distributions with the same mean, only varies in standard variance. So we can use 48 one dimensional draws, providing their respective mean and standard variance, to generate each dimension of the features.

4.d) This solution is basically the same as that one of 2. d). The only difference is that this time these 48 dimensions are independent identical distributions. Thus we can generate 48 draws with one dimensional normal distribution with mean equals 0 and standard variance equals 1.

6. d)

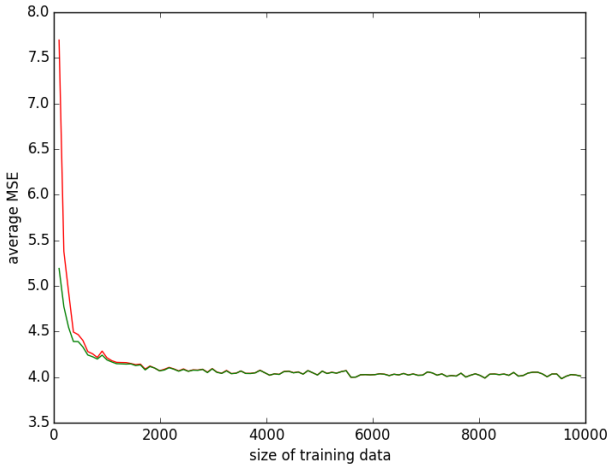
lambda=1e-30, training=100, result: 7.25772390108

lambda=1e-30, training=500, result: 4.4891754703

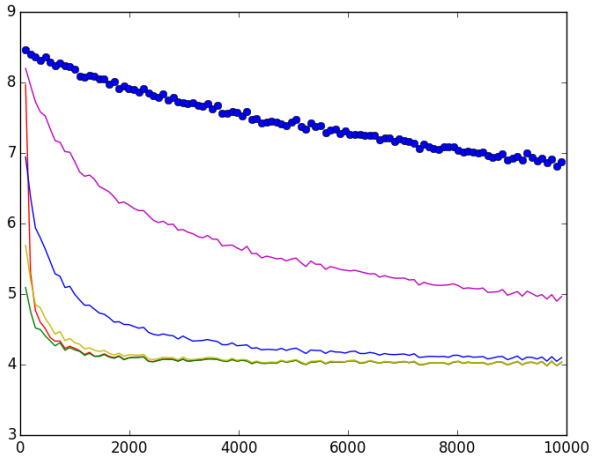
lambda=5, training=100, result: 5.19992531064

lambda=5, training=500, result: 4.41918140686

e&f&g) The following picture shows how the result converge as the training data grows:



The red curve is lambda = 1e-30, the green one is lambda = 5. Therefore we can see as the training data grow, the result presents the same. So when we have a large number of training data, we should choose a smaller regulariser as it would give a good result and won't result in over-regularisation.



When I try different lambda values, it is presented that if over-regularisation occurs, the MSE would stay high and converge very little. So if we have sufficient large scaled training data. We can choose a relatively smaller lambda.

7. d) The result is:

Method	MSE	Variance
Holdout	4.34934458846	0.41895747949
Cross Validation	4.70017009677	0.0445847373941

The variance is tremendously reduce by cross validation.

Cross validation gives a better estimation of MSE, as the variance is smaller, the parameter is more possible to be a good one.