

Assignment 1

problem 2

2.B)

Effect of degree on results:

As the degree of the fitting curve is increased I was able to fit the train data more accurately and the empirical risk/error was brought close to zero in all the run cases. When I used θ^* obtained from train data and used it check how well it predicted test data - it showed a greater accuracy of prediction, but after a certain degree the results on test data showed the signs of being overfitted. This is exactly in accordance with what has been discussed in class.

Effect of luck on results:

When the test data and train data were close, the empirical risk trends of test data and train data seem to match each other and the effects of over fitting was not observed in the test data (expecially for the case where $N = M$). On the other hand when test data and train data were different, it took a higher degree of curve to reduce the empirical risk on test data.

Effect of inaccuracies of operations on computer:

When I was not using pseudo inverse in computing the θ^* , I noticed the effect of skinny matrix vs fat matrix. When the matrix was skinny I was able to get a unique solution, but as the degree of p increased the matrix became fat and no longer the solution was unique. I then used psuedo inverse which gave the solution with smallest norm. This i feel is a limitation as we are picking one solution and ignoring the others which could have given us better θ^* .

Effect of varying N and M :

Case 1: $N = M = 10$: In most of the cases (except bad luck cases), the test data and train data showed almost overlapping results in empirical risk vs degree plots.

Case 2: $N = 100$ and $M = 10$:

In this case, a large set of training data was used to fit the curves, and when finding empirical risk on test data, at a lower degree of p the optimum value was achieved as compared to the next case when $N = 10$ and $M = 100$.

Case 3: $N = 10$ and $M = 100$:

In this case a small set of train data was used to fit the curves and then tested the curves fitted on our large test data. The result followed the general trend of empirical risk decreasing with degree and then attaining an optimum value. But compared to the earlier case it attained optimum at a slightly higher value of p .