ECE 517 HW-6.2: DUAL FORMULATION OF A NONLINEAR MMSE CLASSIFIER

Susan Sapkota

November 2020

1 Problem 6.2

Use the functions of the previous assignment to reconstruct the example of lesson 6.1 but using a dual representation an the polynomial kernel $k(x_i, x_j) = (x_i^T x_j + 1)^3$

1. Construct a train dataset and represent them.
2. Construct a function that computes the kernel matrix $\mathbf{K}$.
3. Compute the dual weights α_i of the MMSE solution.
4. Write an estimator in dual form as a function of kernel dot products between the training and test data.
5. Plot the boundary,
6. Repeat the experiment, but using the Ridge Regression solution, this is $\alpha = (K + \gamma I)^{-1} y$
   where gamma is a small number. Show the result for different values of the parameter that are able to produce different solutions. Comment the results.

Provide a document that summarizes the theory and a graph of the result. Comment your results.

2 Solution

I used function and kernel matrix given in the problem 6.1 to make data sets. I used Volterra expansion for getting Training dataset.

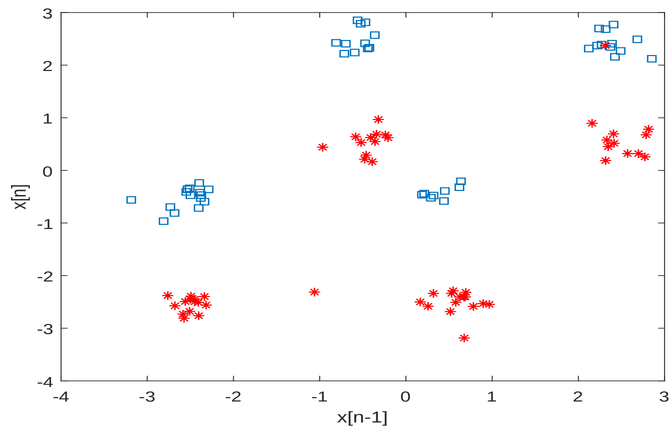

Figure 1: Training Data

2.1 Kernel Function

We can use equation given in question to construct kernel function.

$$k(x_i, x_j) = (x_i^T x_j + 1)^3 \tag{1}$$

Now, MMSE solution is simple as follows,

$$\alpha = (\Phi^T \Phi)^{-1} y \tag{2}$$

The solution of α is a 1×100 and shown in figure using MMSE and kernel function.

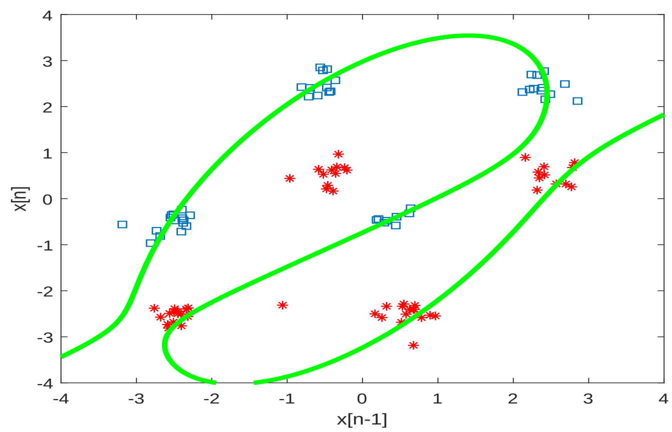

Figure 2: Kernal classification

we can clearly see that there is so many data miss classified by MMSE and kernel function. so, we have to use something better like Ridge Regression for classification.

2.2 Ridge Regression

We use same training set data then used Ridge solution given by,

$$\alpha = (K + \gamma I)^{-1} y \tag{3}$$

We draw the solution using Ridge Regression using above equation with varying parameter γ given below.we see that the lower value of γ doesn't affect so much in classification but high γ miss-classify data which agree with the theory we learnt in class.

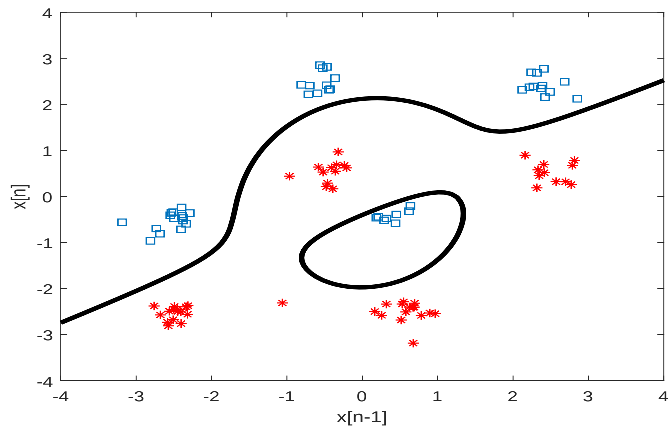

Figure 3: Ridge classification with $\gamma = 1$

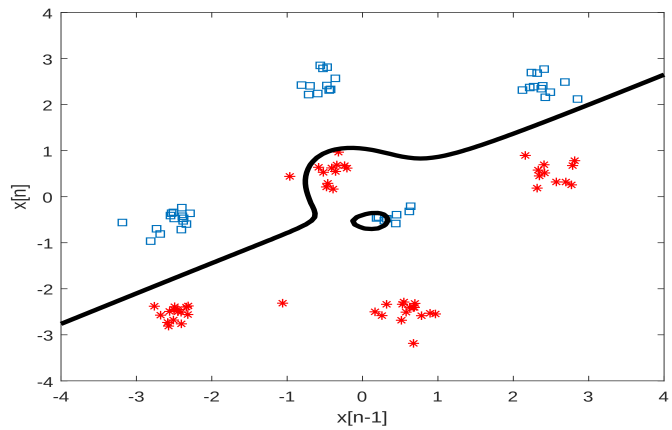

Figure 4: Ridge classification with $\gamma = 200$