E0-270 Assignment1

Jayabrata Chowdhury
SR No: 17449, Dept: RBCCPS, IISC

June 7, 2020

1 Question 1a

k(x,y) = $x^T Ay$
k(y,x) = $y^T Ax$
for k to be valid kernel k(x,y)=k(y,x). The matrix A should be symmetric and positive semi definite for k(x,y) to be a valid kernel.

2 Question 1b

k1 is not a valid kernel. k2 is a valid kernel. k3 is not a valid kernel. k4 is not a valid kernel. k5 is not a valid kernel.

3 Question 2a

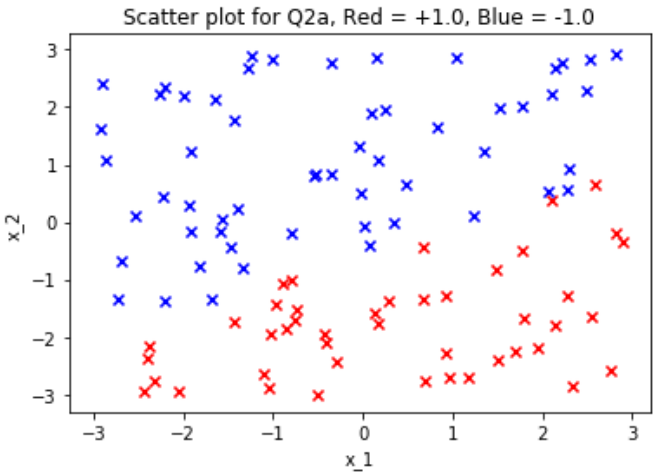

Figure 1: Scatter plot of question 2a

4 Question 2b

Training set accuracy in percentage: 94. Testing set accuracy in percentage: 52. Optimal value of hyperparameter,C: 50. Optimal value of w is: [-3.09065331, 3.33717864]. Optimal value of b is: -2.61070408051118

5 Question 2c(i)

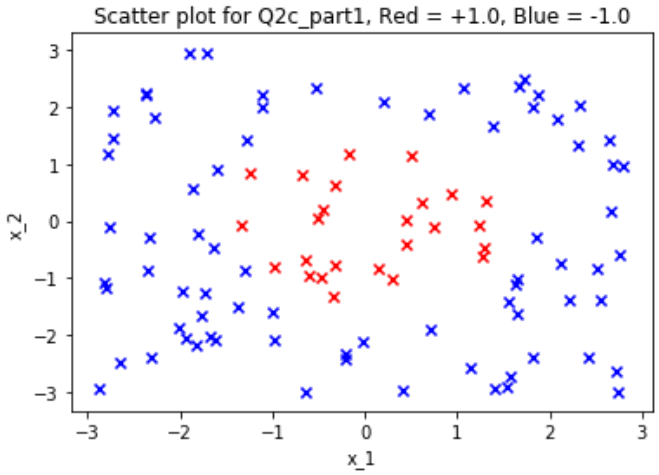

Figure 2: Scatter plot of question 2c(i)

6 Question 2c(ii)

Training set accuracy in percentage: 58. Testing set accuracy in pecentage: 52. Optimal value of w is: [-1.51867425e-08, 1.02752153e-08]. Optimal value of b is: -2.3454037848304716e-08. Otimal value of C is: 200.

7 Question 2d(i)

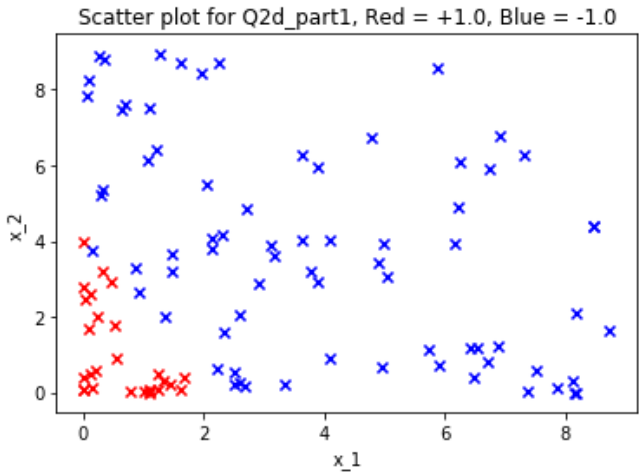

Figure 3: Scatter plot of question 2d(i)

8 Question 2d(ii)

Training set accuracy in percentage: 78. Testing set accuracy in pecentage: 76. Optimal value of w is: [-15.07796288, -7.12693605]. Optimal value of b is: 28.825748172884886. Otimal value of C is: 150.

9 Question 2d(iii)

Yes, I am getting better accuracy in 2d than 2c for both training and testing set.

10 Question 2e

kernel used for this question: k(x,y) = $(1 + x^T y)^2$

Training accuracy in percentage: 58. Testing accuracy in percentage: 56. Hyperparameter: 25.

11 Question 3a

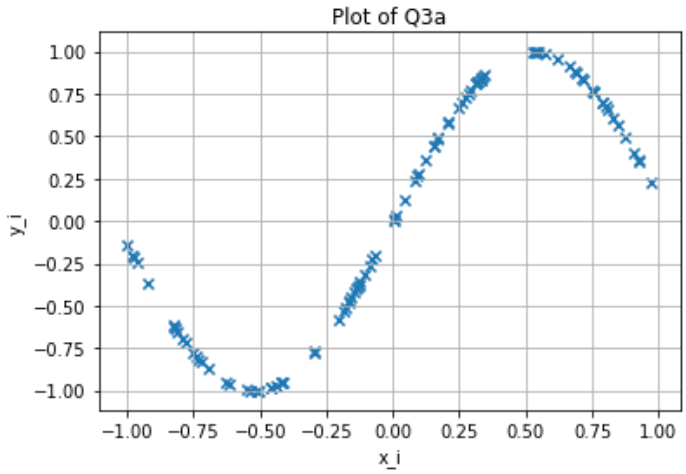

Figure 4: Scatter plot of question 3a

12 Question 3b

Mean Square Error on training set: 0.1388.

Figure 5: Scatter plot of question 3b

13 Question 3c

MSE for D1: 0.001929
MSE for D2: 0.003092
MSE for D3: 0.000001
MSE for D4: 0.000005
MSE for D5: 0.000000
MSE for D6: 0.000000

MSE for D7: 0.000000
MSE for D8: 0.000000
MSE for D9: 0.000000
MSE for D10: 0.000000

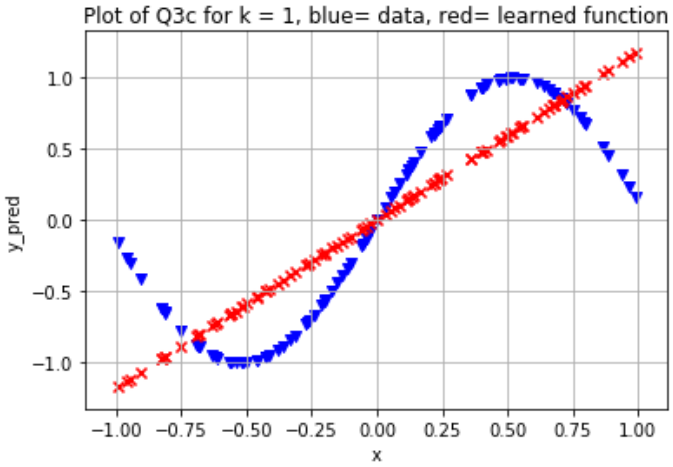

Figure 6: Scatter plot of question 3c for k=1

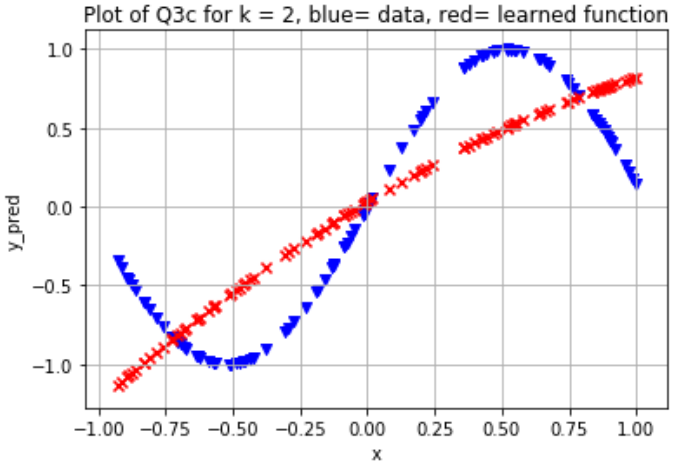

Figure 7: Scatter plot of question 3c for k=2

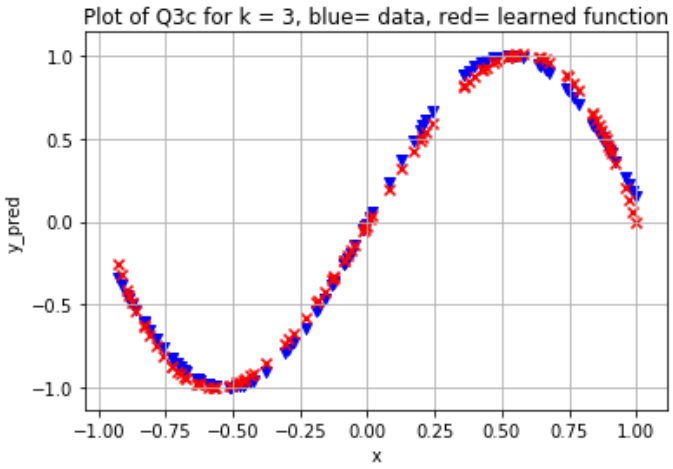

Figure 8: Scatter plot of question 3c for k=3

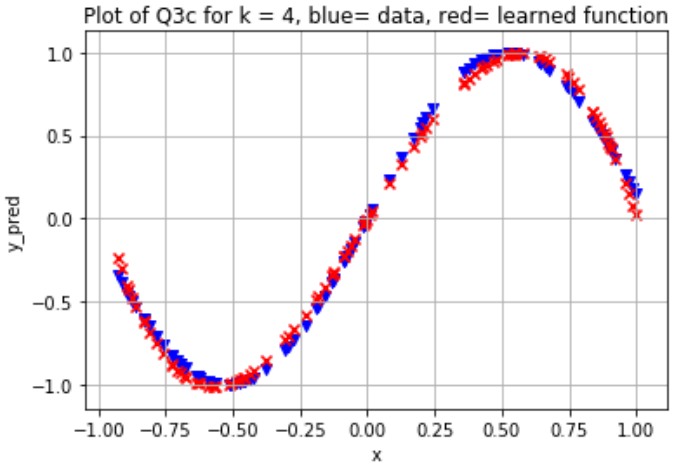

Figure 9: Scatter plot of question 3c for k=4

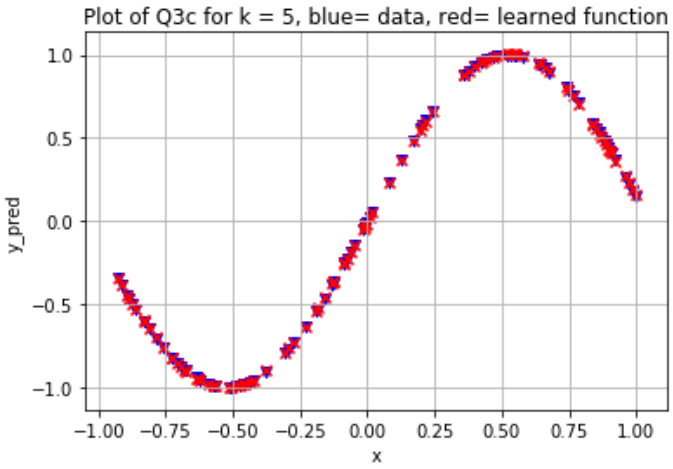

Figure 10: Scatter plot of question 3c for k=5

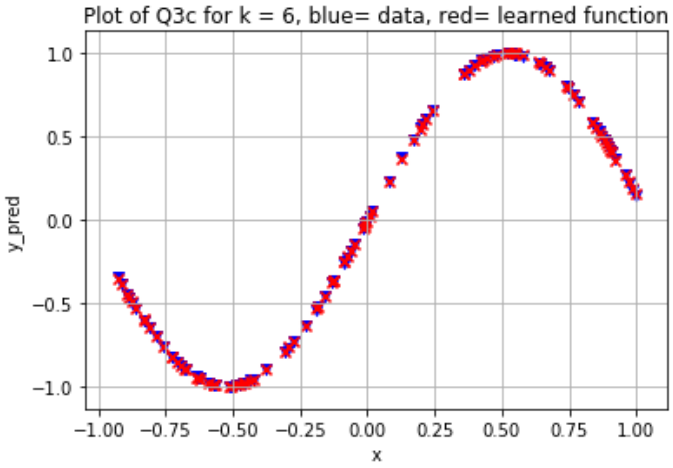

Figure 11: Scatter plot of question 3c for k=6

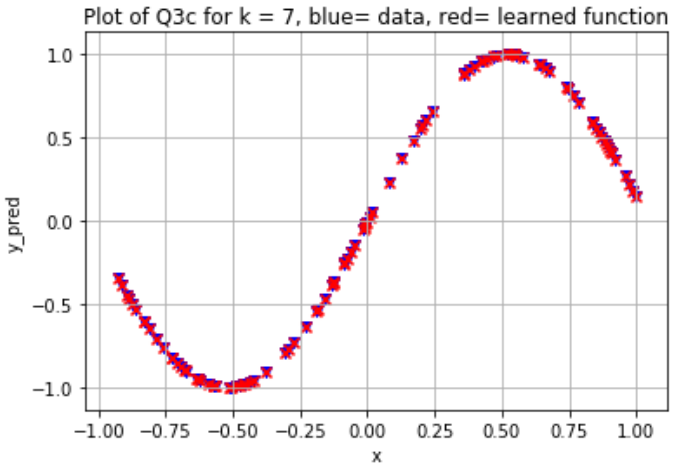

Figure 12: Scatter plot of question 3c for k=7

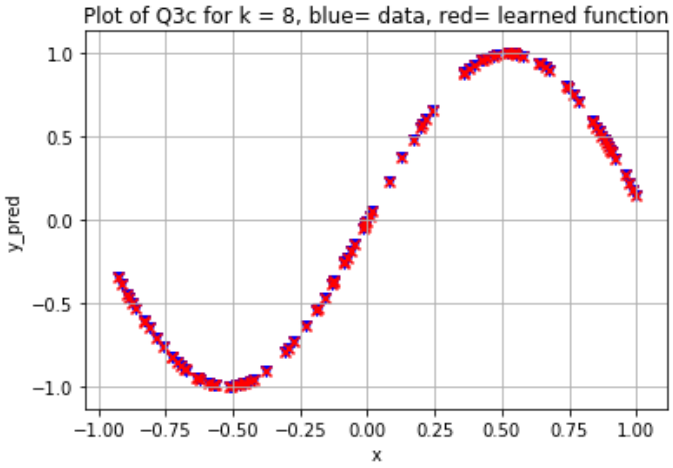

Figure 13: Scatter plot of question 3c for k=8

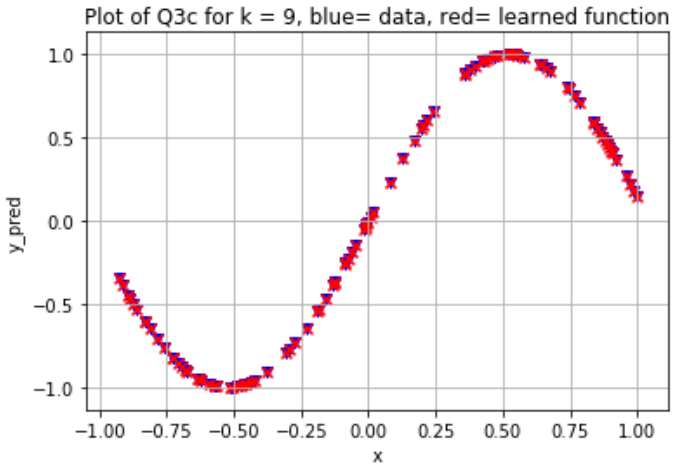

Figure 14: Scatter plot of question 3c for k=9

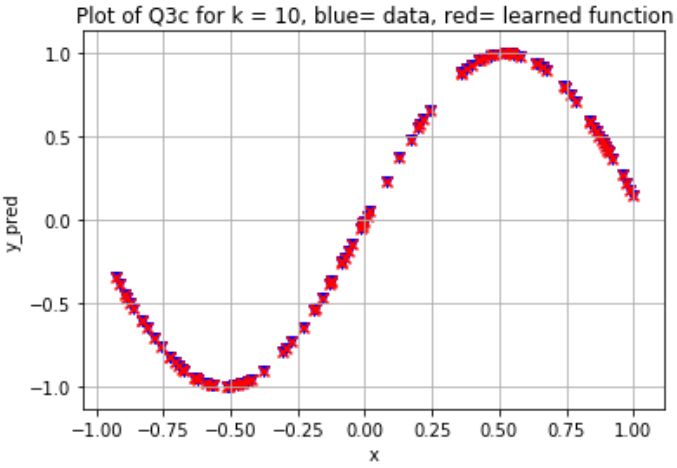

Figure 15: Scatter plot of question 3c for k=10

14 Question 3d

(i) Kernel used in Q3d is k(x,y) = $x^T y$.

(ii) Training set MSE in percentage: 16.59492363.

Testing set MSE in percentage: 18.43880403.

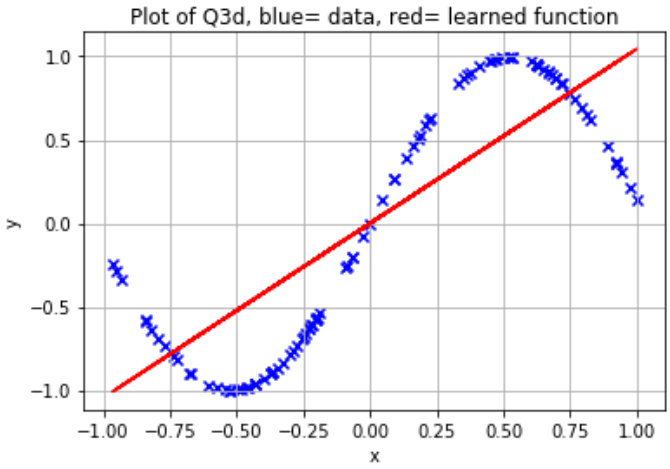

Figure 16: Plot of question 3d

15 Question 4a

The sum of all entries of D is 0.

16 Question 4b

The summation $\sum_{i=1}^{10} \mathbf{d}_i = 0$

17 Question 5a

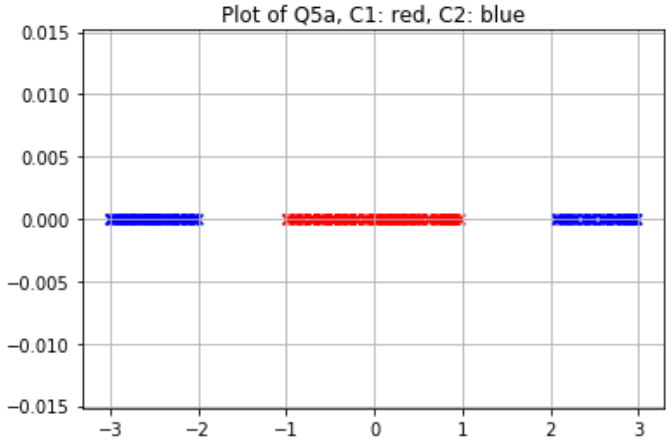

Figure 17: Scatter plot of question 5a

18 Question 5b

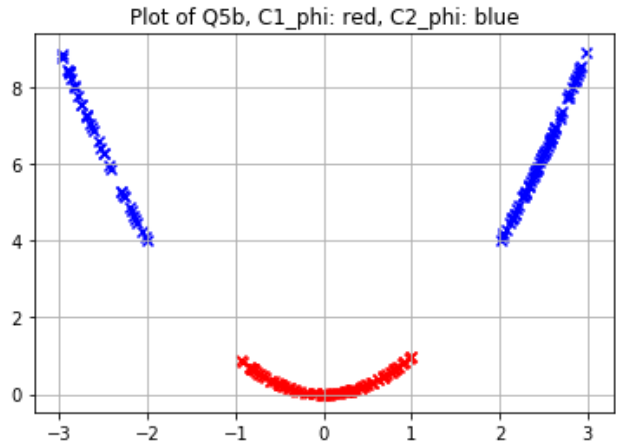

Figure 18: Scatter plot of question 5b