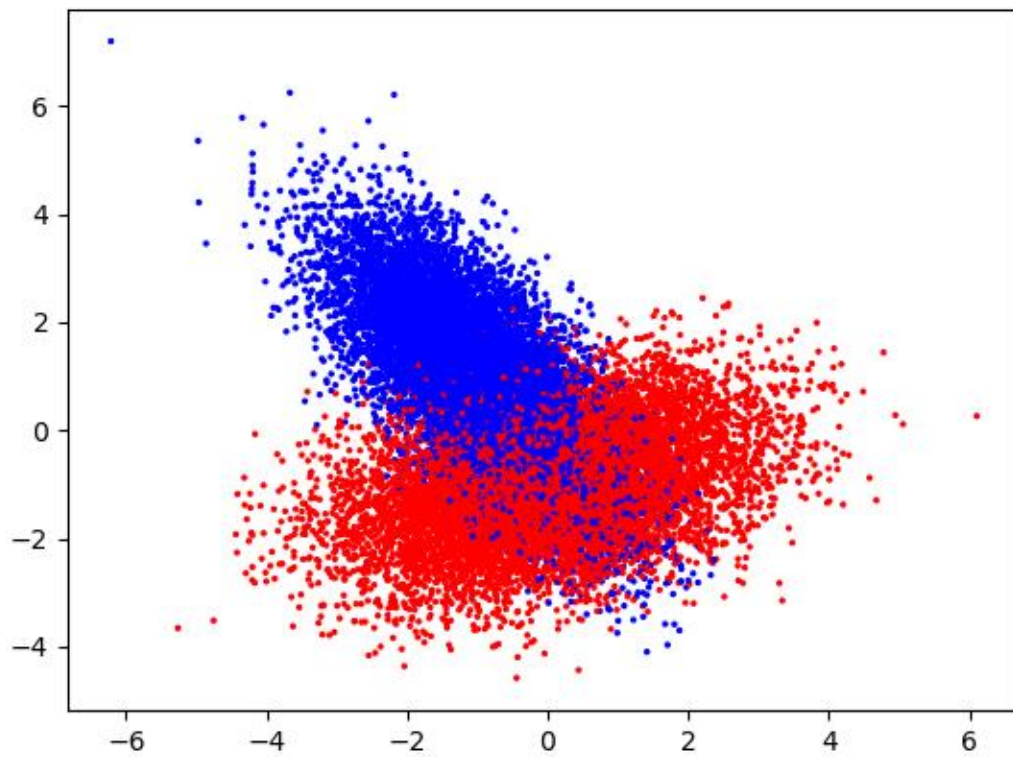
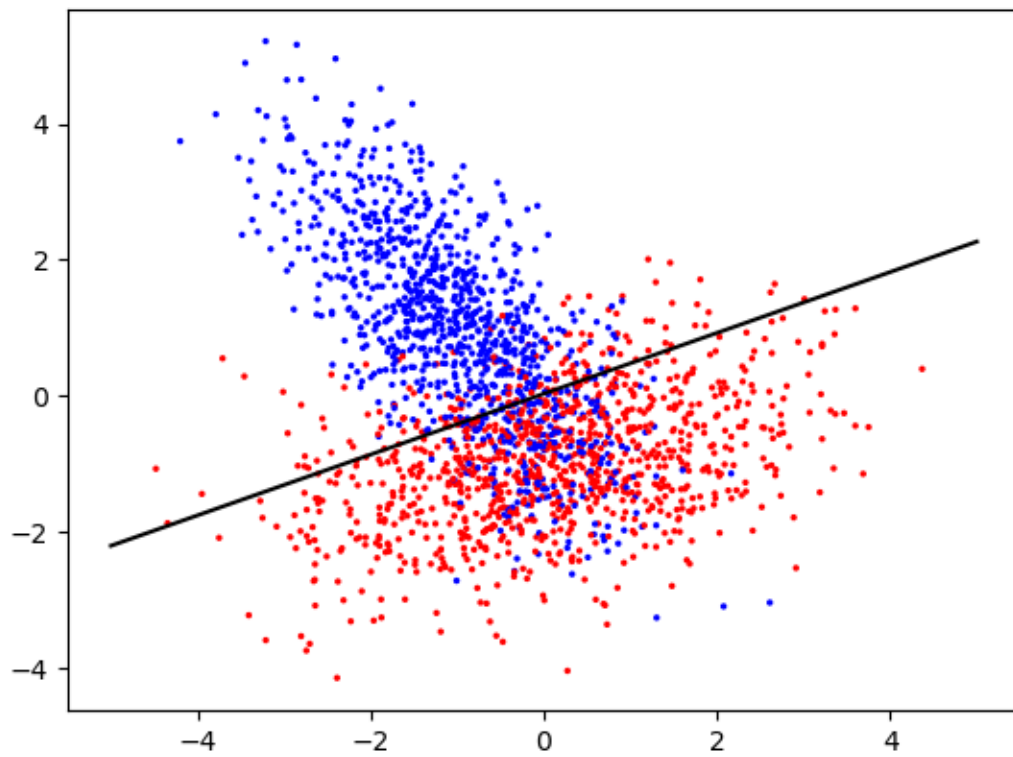


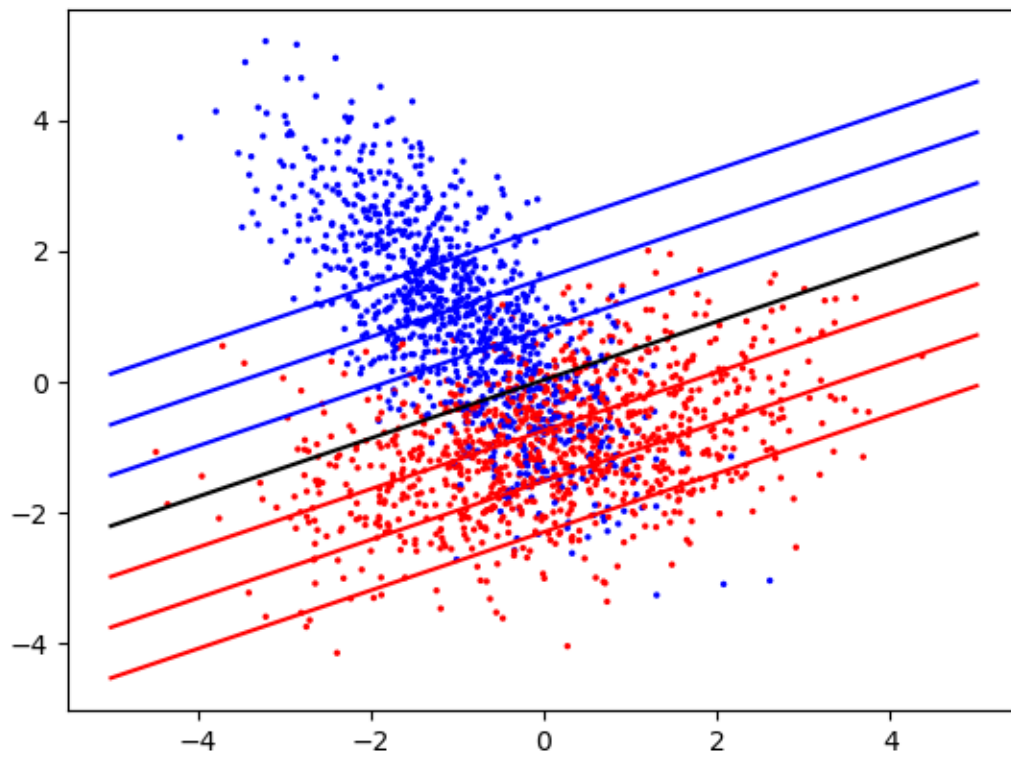
Question 1(c): sample cluster data (10,000 points per cluster)



Question 2(c): training data and decision boundary



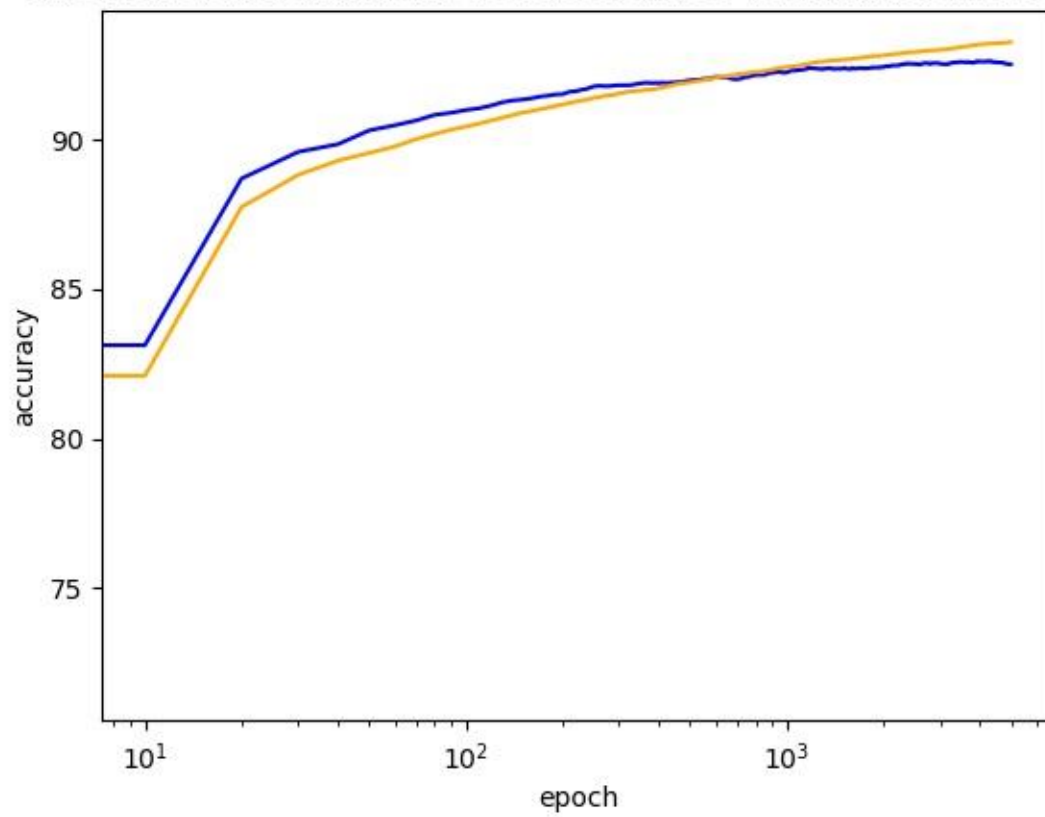
Question 2(d): decision boundaries for seven thresholds



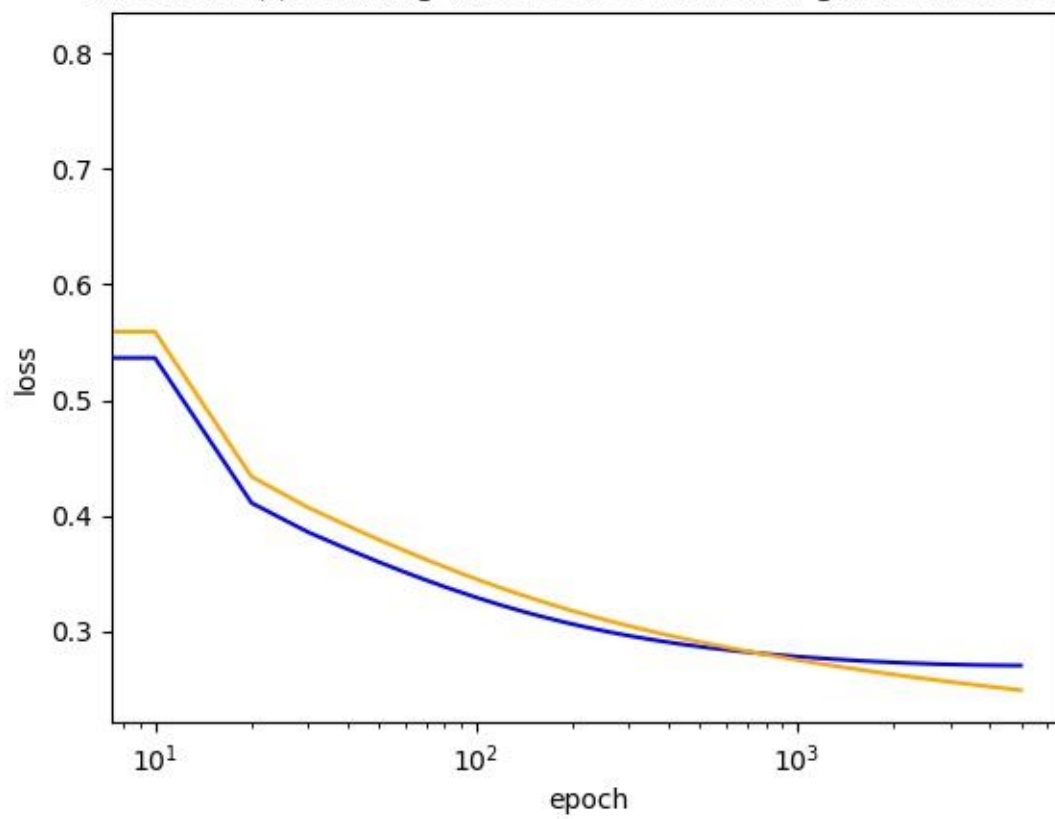
Question 3(a):36 random MNIST images

6	6	2	1	1	3
9	1	0	1	0	8
9	3	/	4	9	/
9	6	5	5	7	4
5	9	9	1	9	6
2	0	9	2	4	5

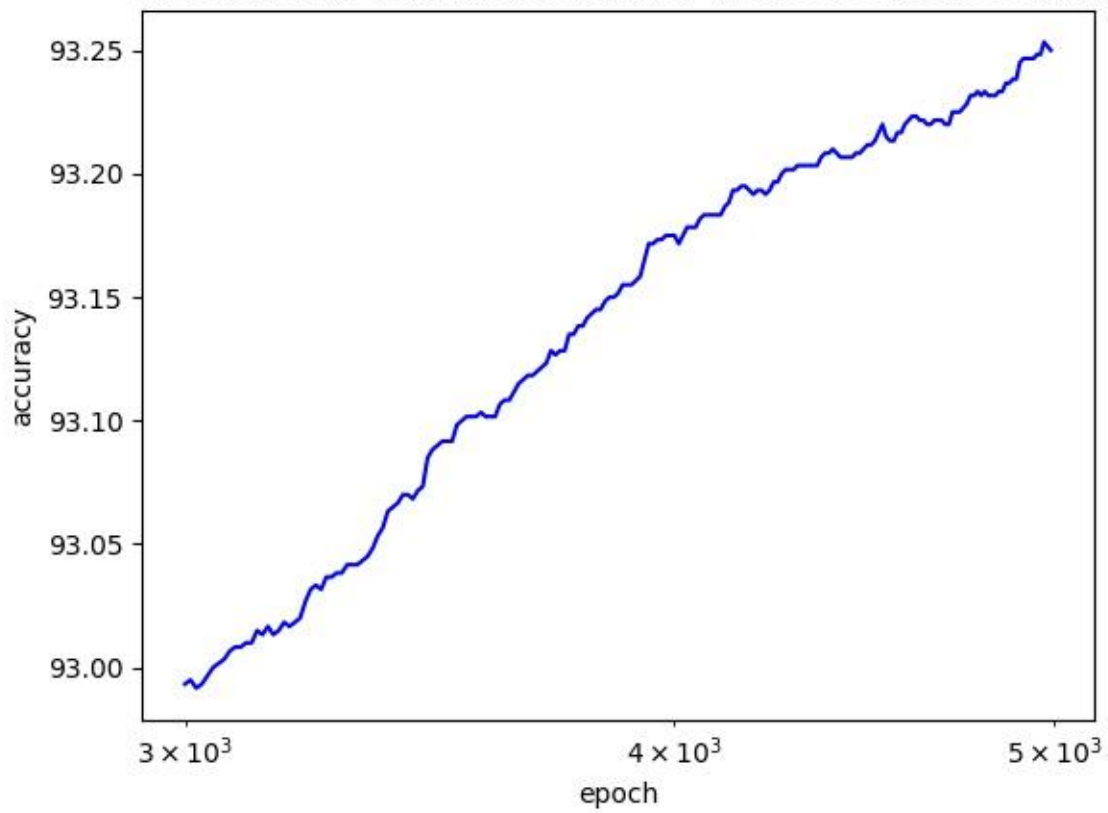
Question 6(e): training and test accuracy for batch gradient descent.



Question 6(f): training and test loss for batch gradient descent.



Question 6(g): training accuracy for last 2000 epochs of bgd.



Question 6(h): training loss for last 2000 epochs of bgd.

