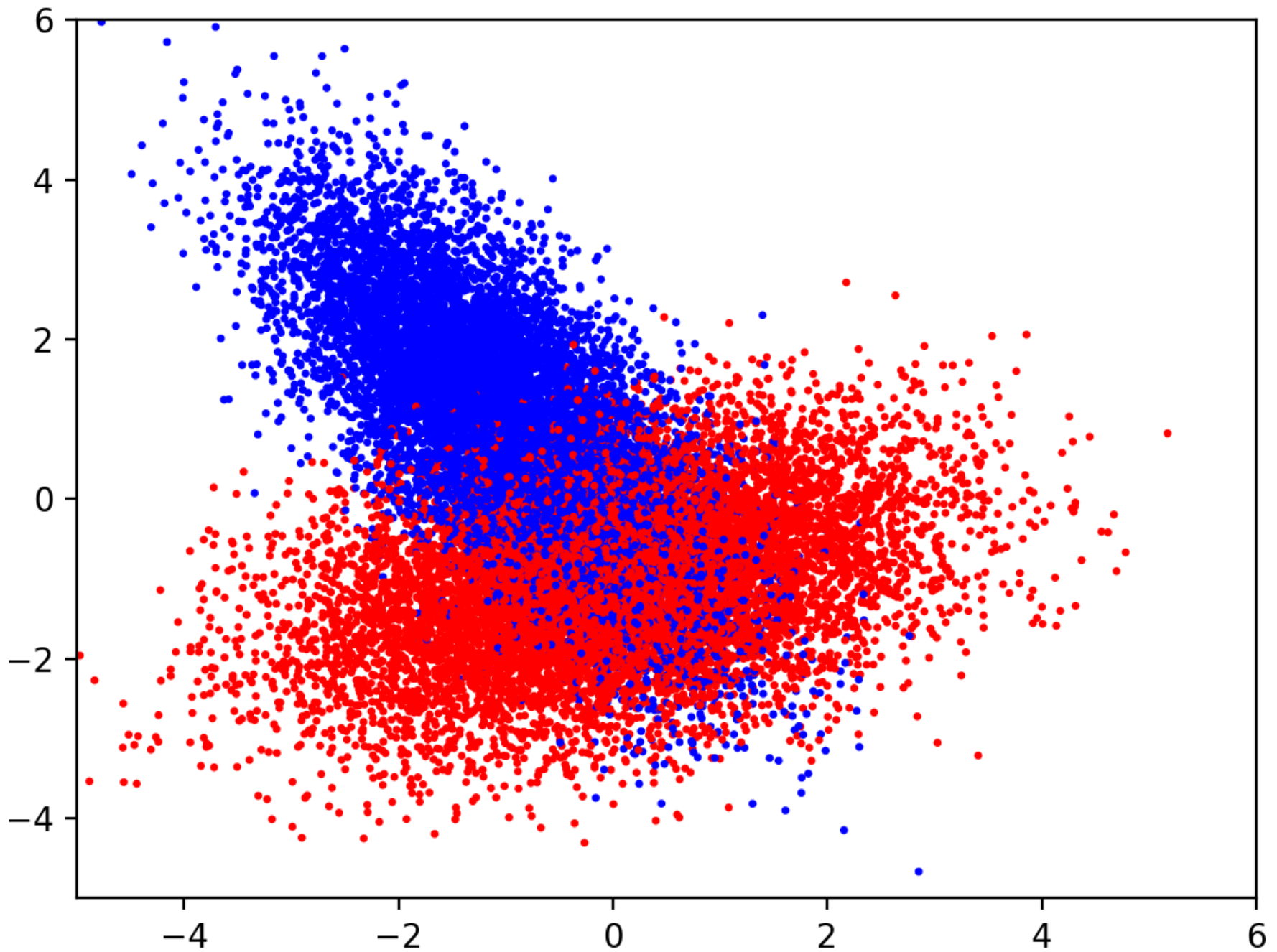
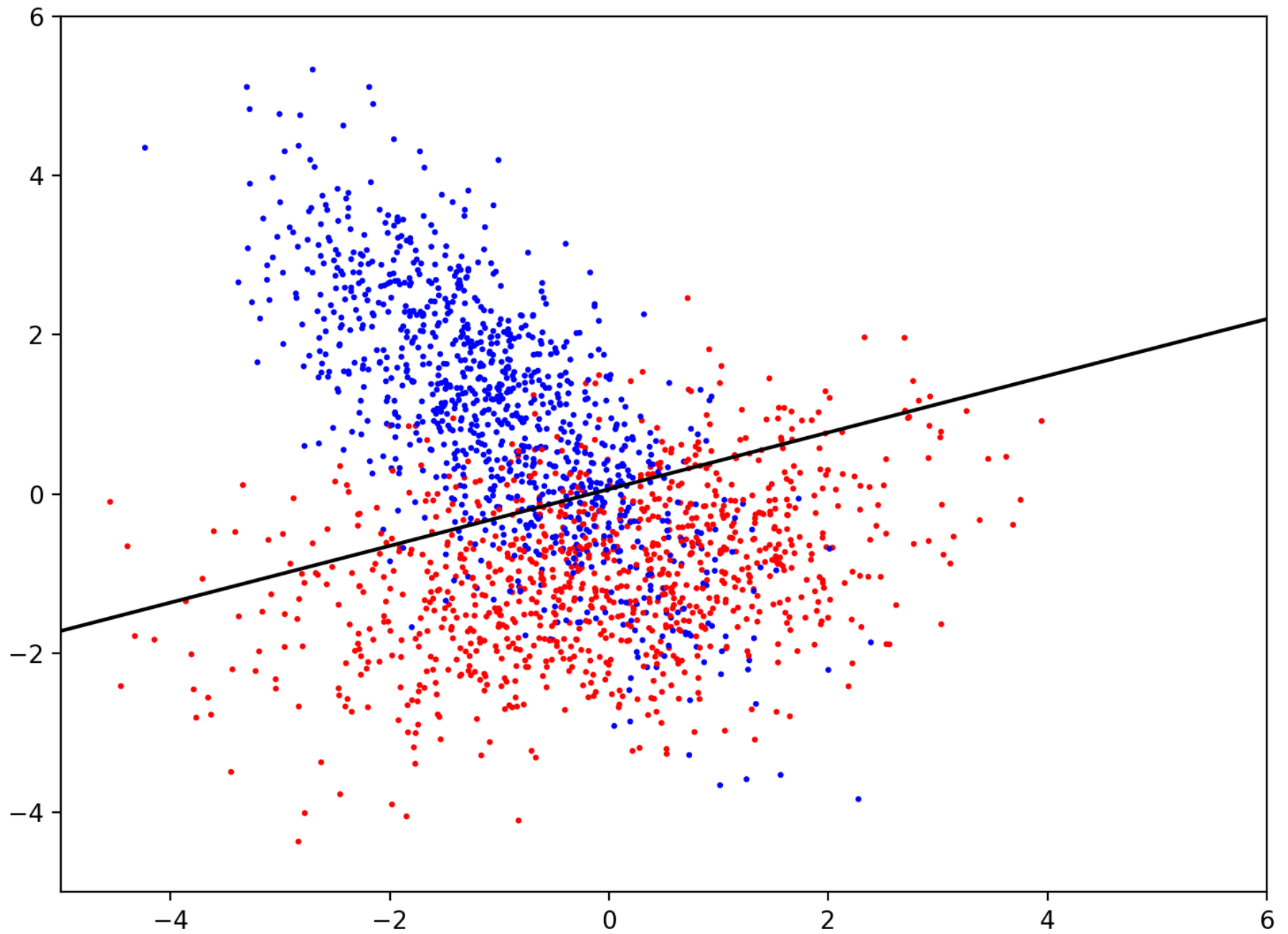


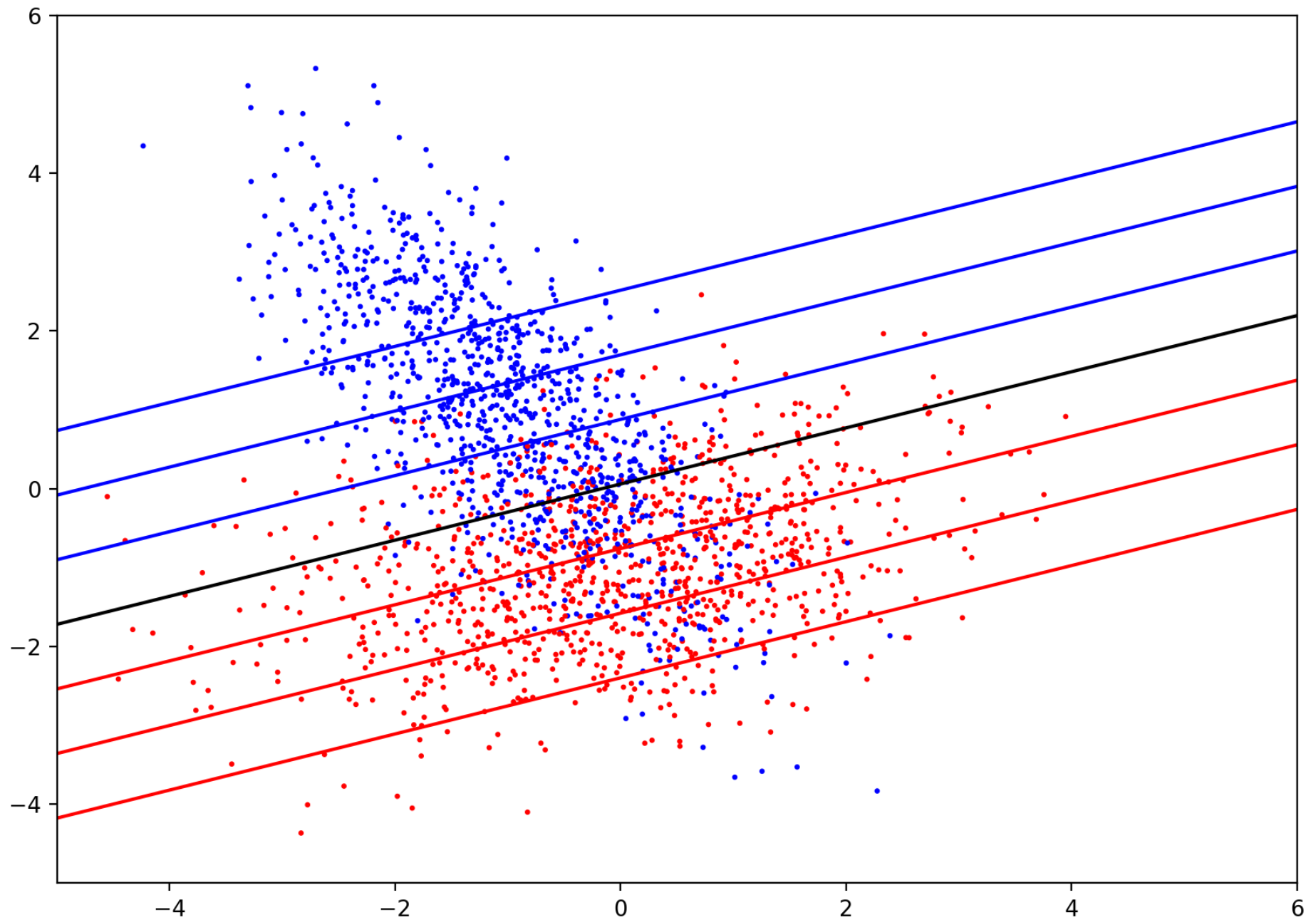
Question 1(c): sample cluster data (10,000 points per cluster)



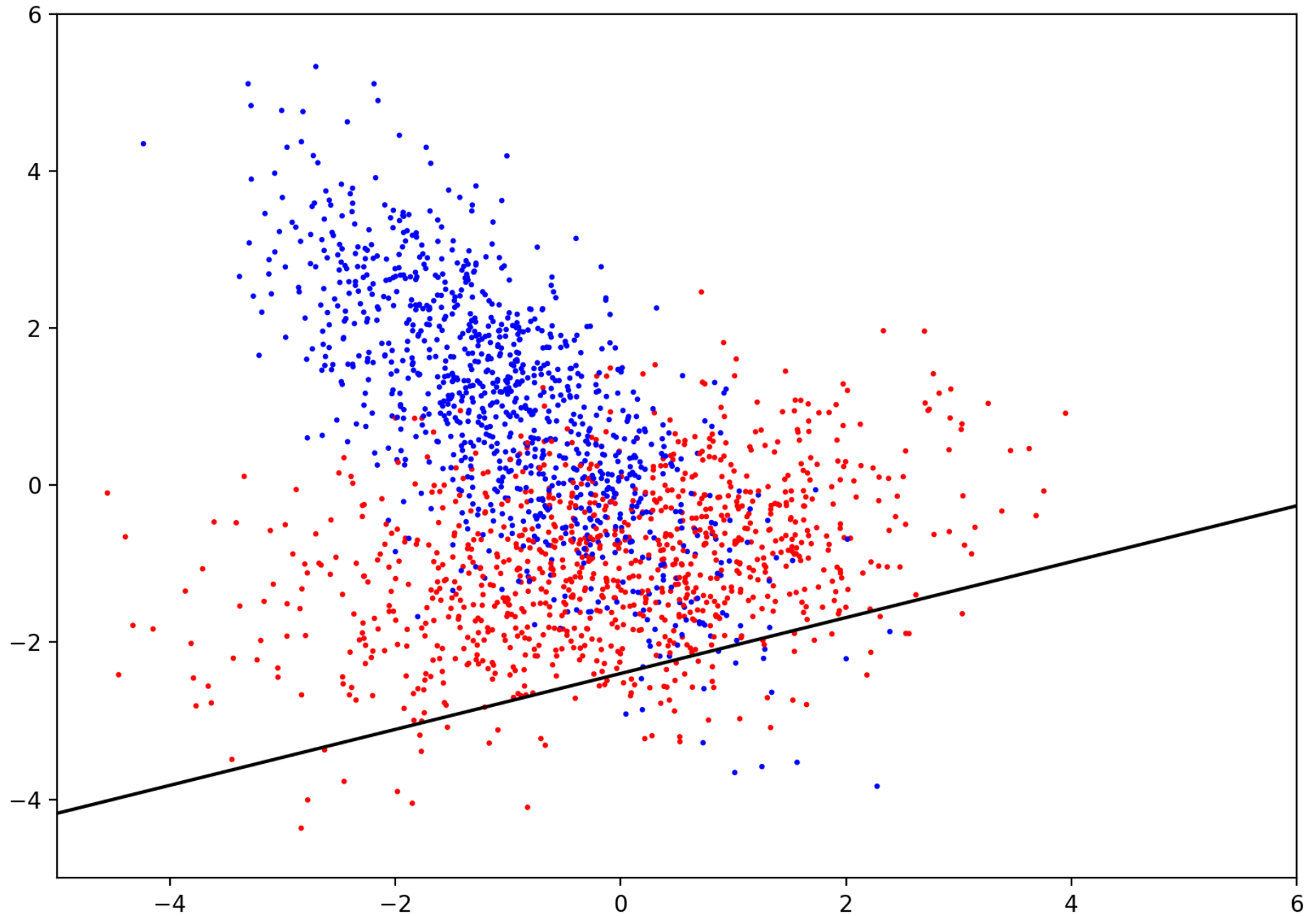
Question 2(c): training data and decision boundary



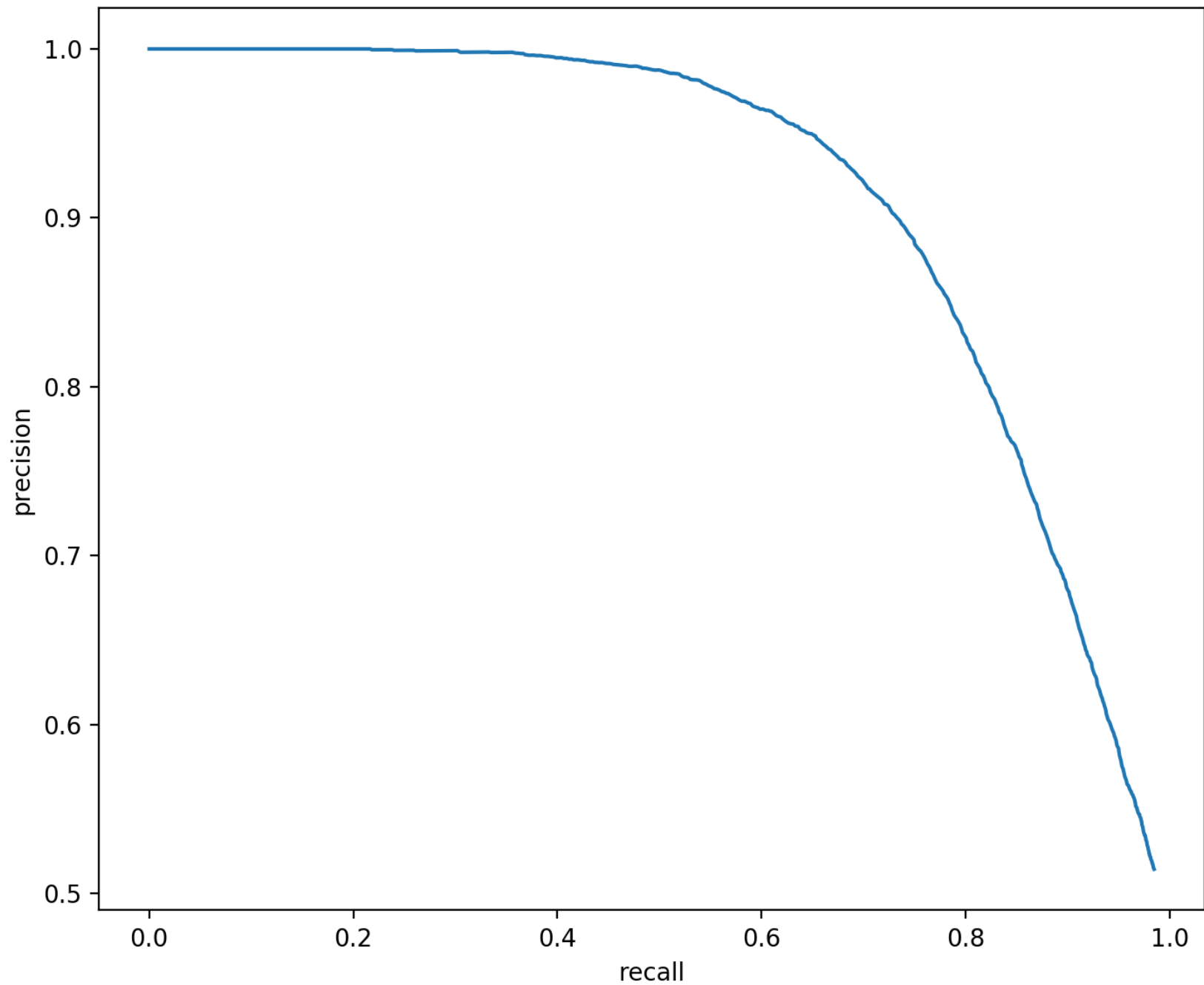
Question 2(d): decision boundaries for seven thresholds



Question 2(e):  $t=-3$  gives the greatest number of False Positives



Question 2(i) precision/recall curve



Question 3(a): 16 random MNIST images

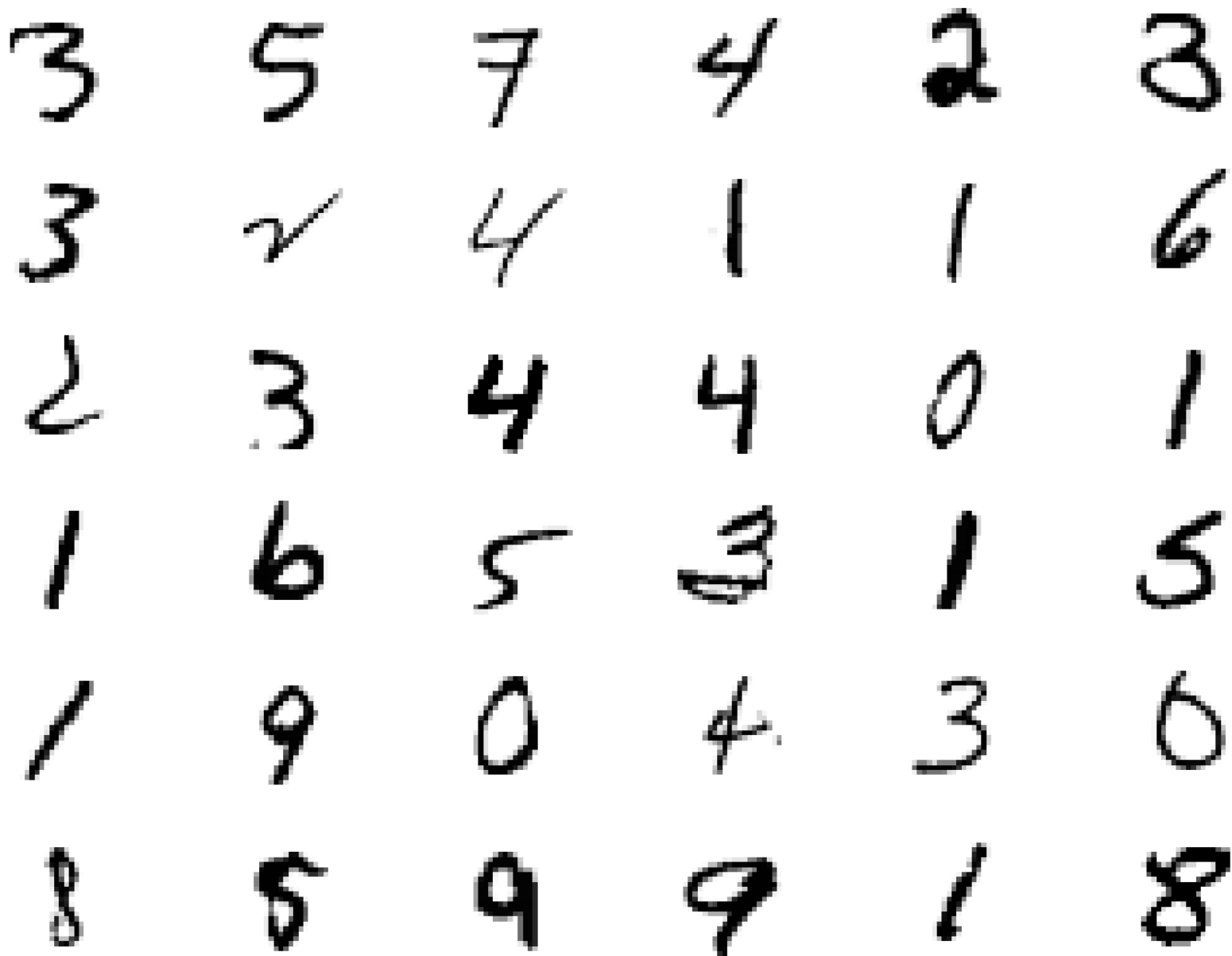
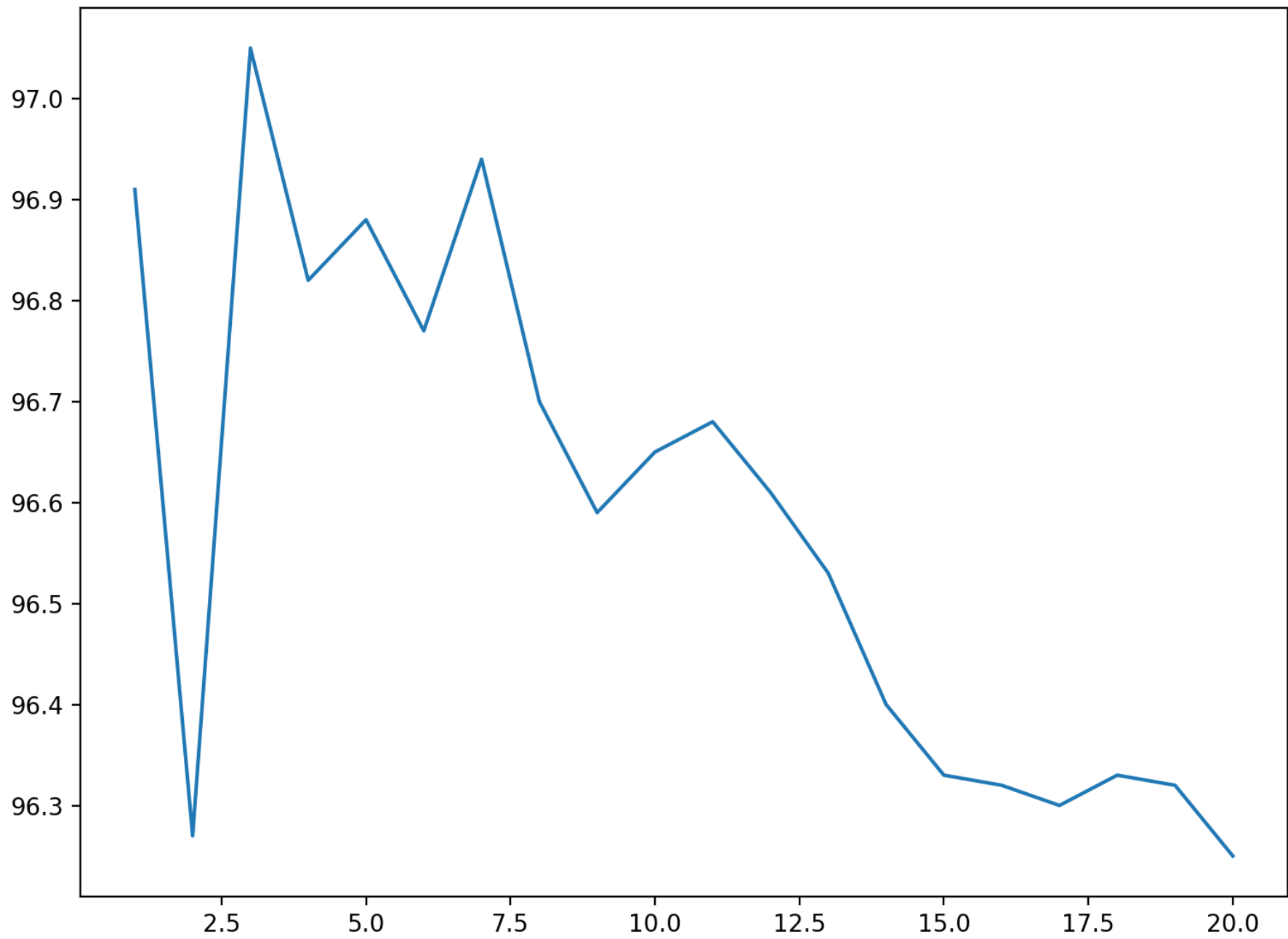
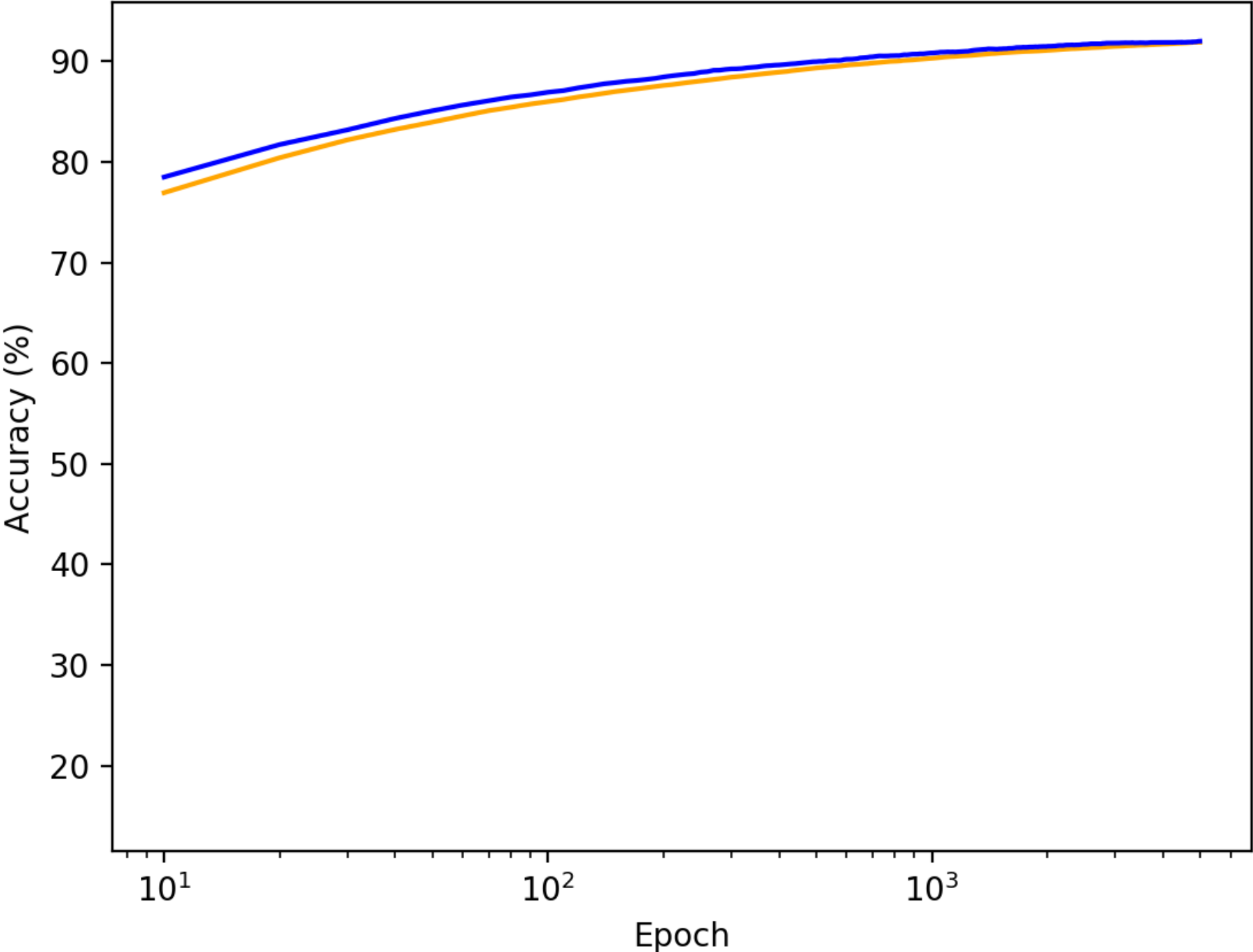


Figure 3(c): KNN test accuracy

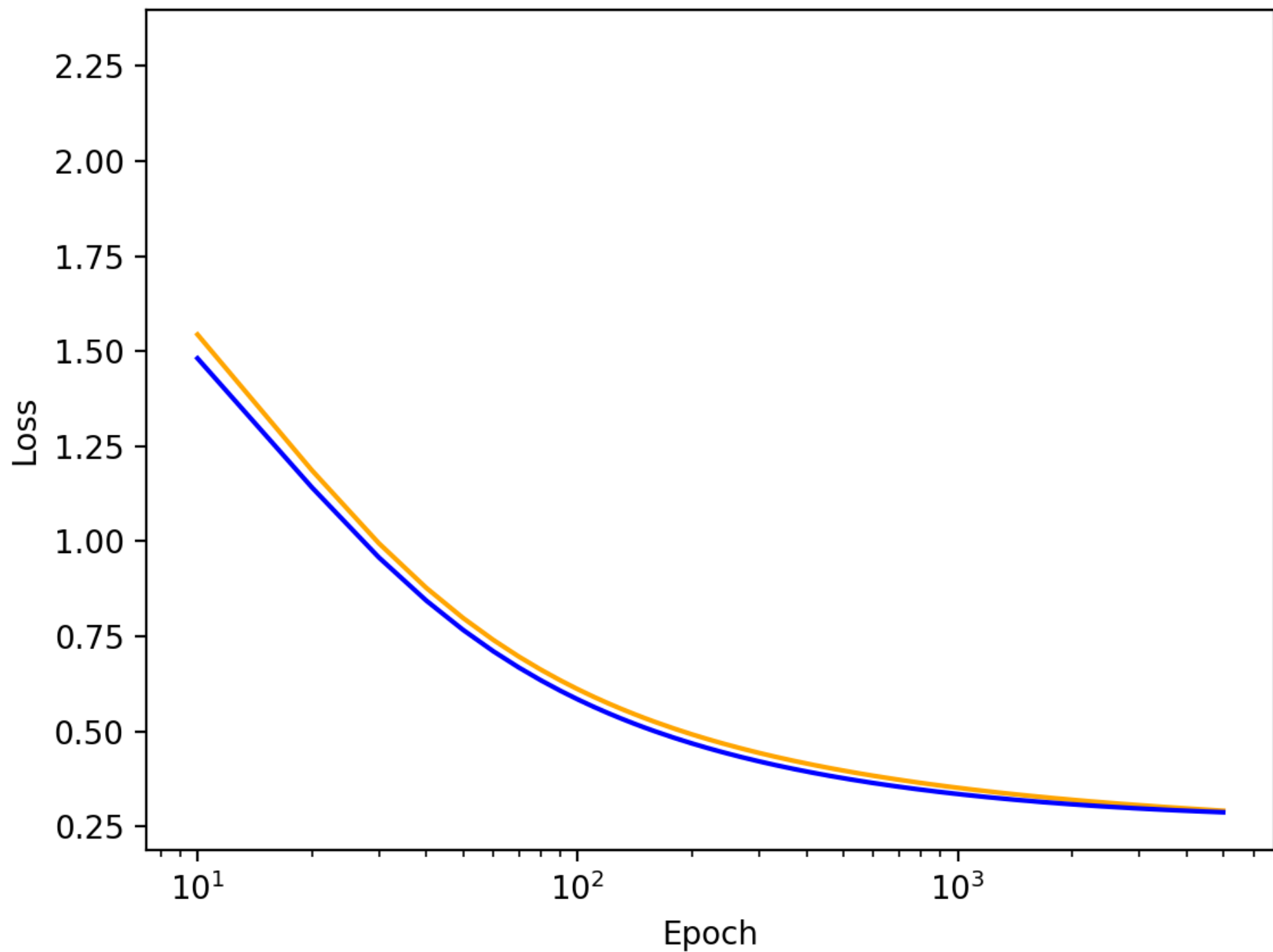


Question 6(e): training and test accuracy for batch gradient descent

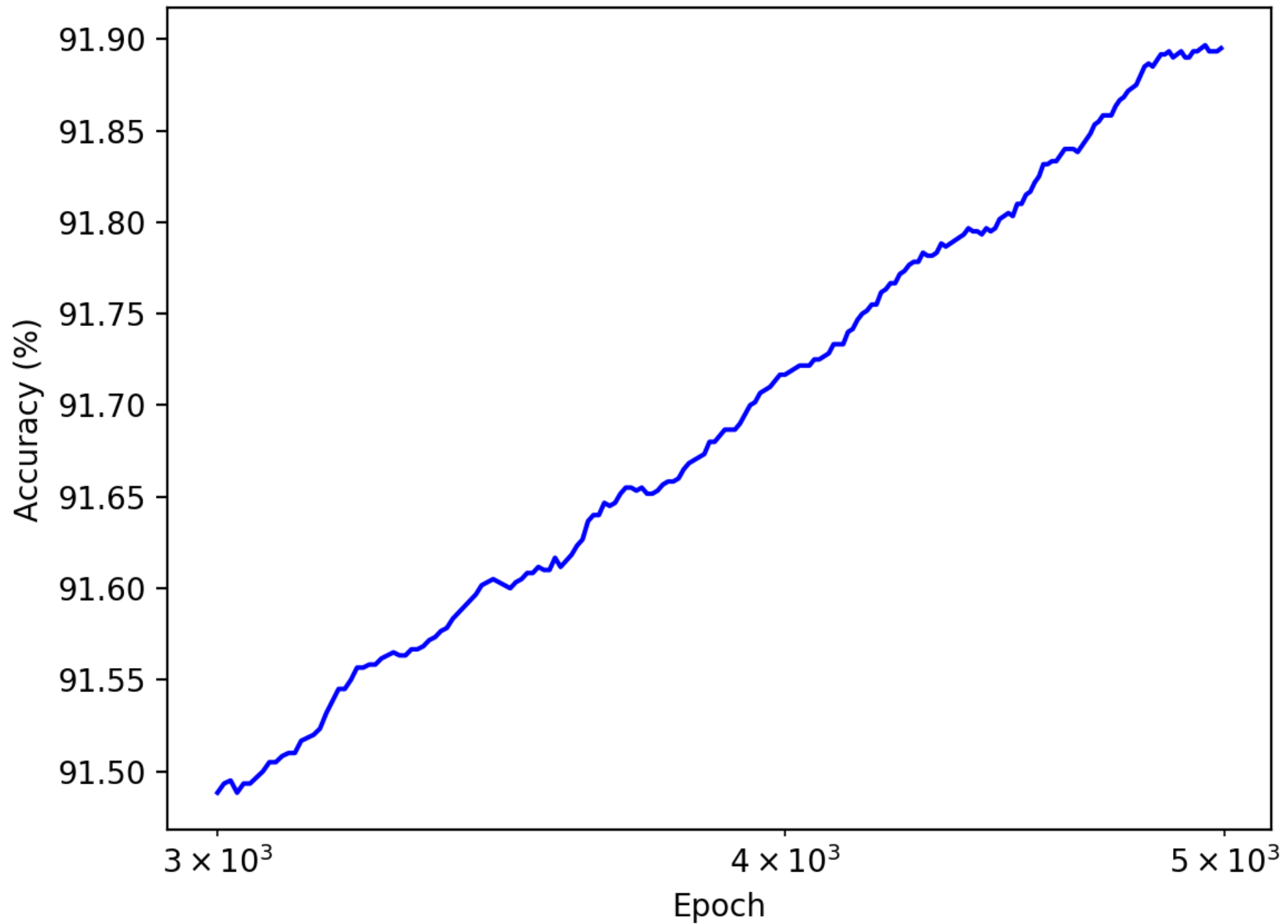




Question 6(f): training and test loss for batch gradient descent



Question 6(g): training accuracy for last 200 epochs of bgd



Question 6(h): training loss for last 200 epochs of bgd

