Homework 2

Collaborators:
Name: Shen Zijin
Student ID: 3160104734

Problem 2-1. A Walk Through Linear Models

(a) Perceptron

Answer:

1. When training set is 10, the training error rate is 0, testing error rate is 10.5700%.
   When training set is 100, the training error rate is 0, testing error rate is 1.2600%.
2. When training set is 10, the average time of iteration is 5.
   When training set is 100, the average time of iteration is 31.
3. Since the training data is not linearly separable, the algorithm will loop endless.

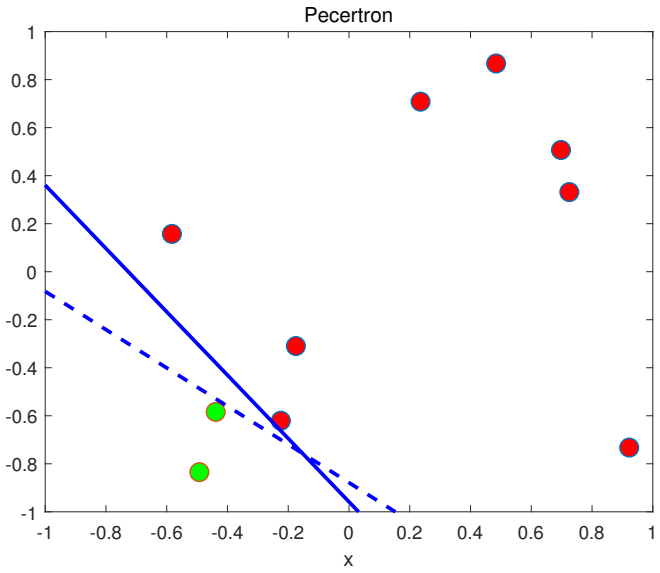

Figure 1: The plotting result for perceptron when nTest = 10.

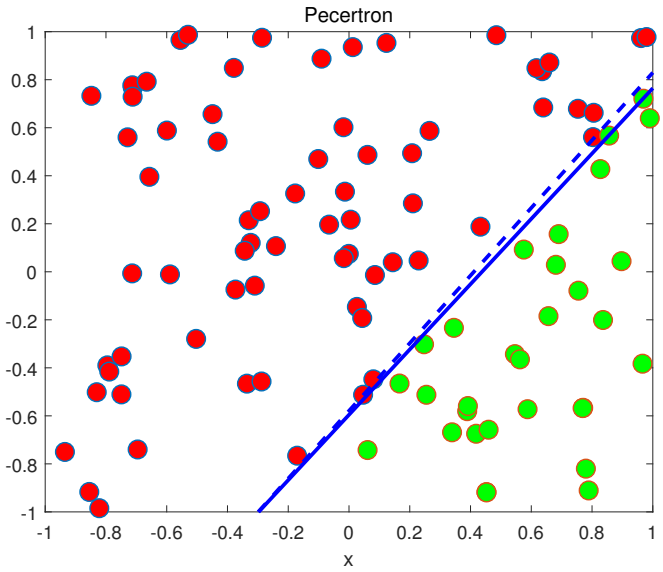

Figure 2: The plotting result for perceptron when nTest = 100.

(b) Linear Regression

Answer:

1. When training set is 100, the training error rate is 3.9630%, testing error rate is 4.8720%.

2. If the training data is noisy and not linearly separable(nTrain=100), the training error rate is 13.6700%, testing error rate is 14.9010%.

3. For poly-case WITHOUT transformation, the training error rate is 49.0000%, testing error rate is 54.9600%.

4. For poly-case WITH transformation, the training error rate is 5.0000%, testing error rate is 6.6000%.

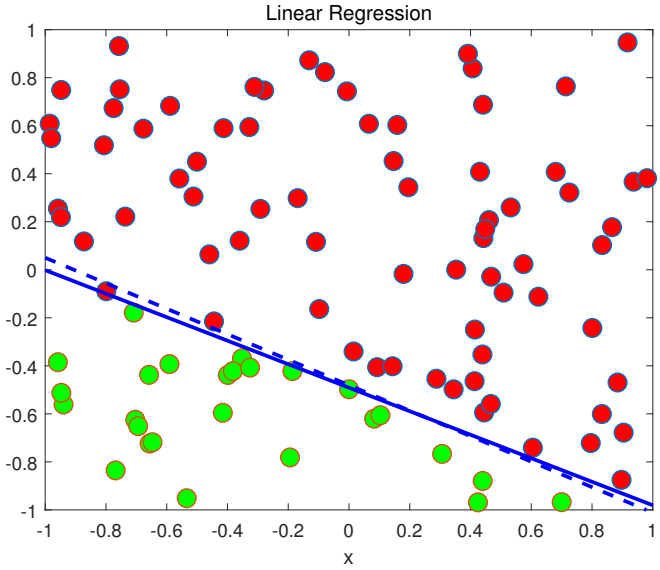

Figure 3: The plotting result for linear regression.

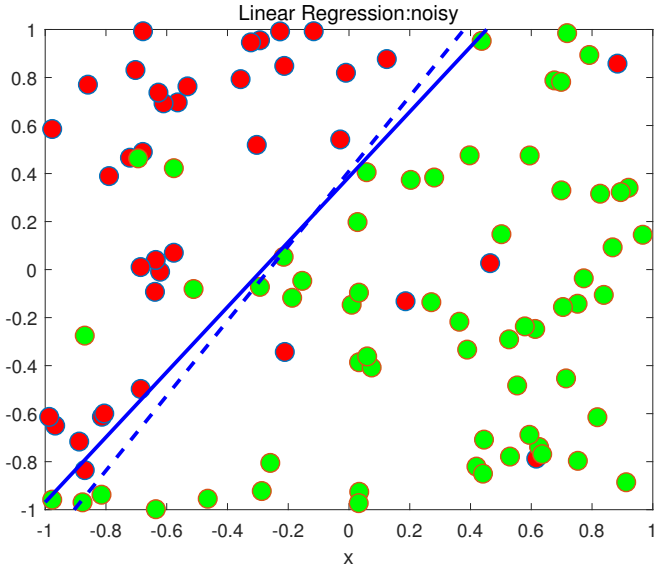

Figure 4: The plotting result for linear regression when training data is not linearly seperable.

(c) Logistic Regression

Answer:

1. When training set is 100, the training error rate is 0.3300%, testing error rate is 1.2300%.

2. If the training data is noisy and not linearly separable(nTrain=100), the training error rate is 12.7600%, testing error rate is 13.9067%.

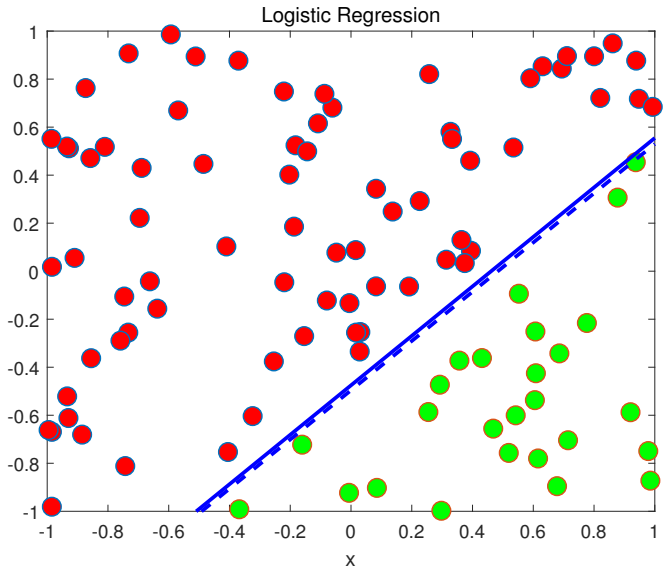

Figure 5: The plotting result for logistic regression.

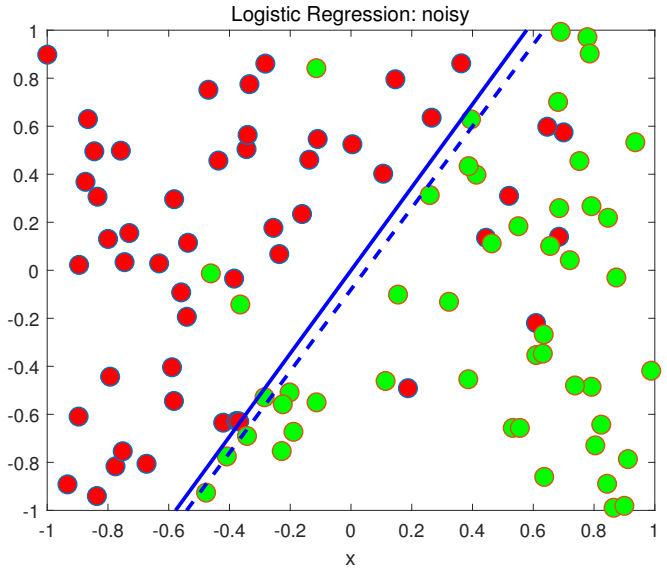

Figure 6: The plotting result for logistic regression when training data is not linearly seperable.

(d) Support Vector Machine

Answer:

1. When training set is 30, the training error rate is 0, testing error rate is 3.3867%.
2. When training set is 100, the training error rate is 0.0060%, testing error rate is 0.9880%.
3. When training set is 100, the average time of iteration is 3.

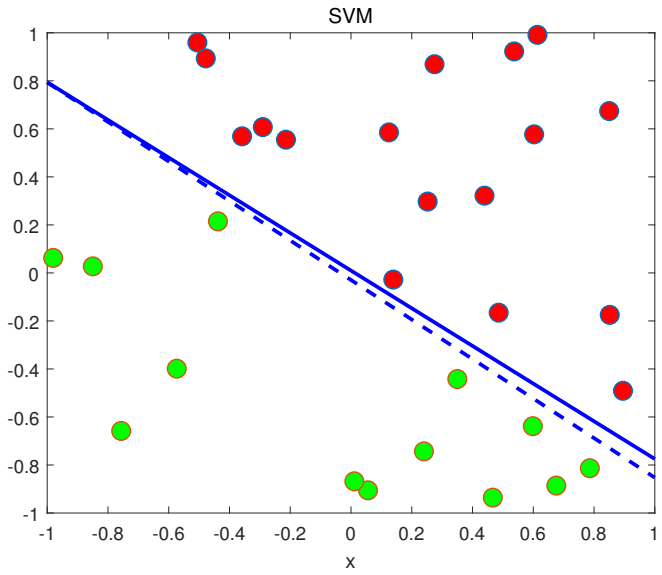

Figure 7: The plotting result for SVM when nTrain is 30.

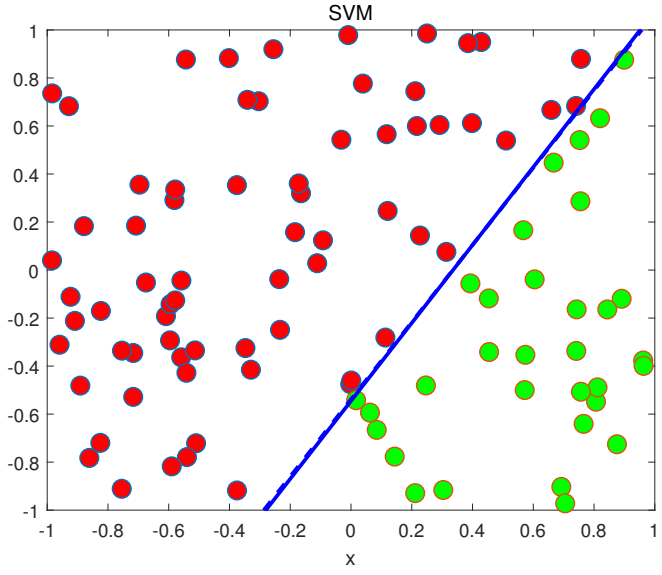

Figure 8: The plotting result for SVM when nTrain is 100.

Problem 2-2. Regularization and Cross-Validation

(a) Implement Ridge Regrssion, and use LOOCV to tune the regularization parameter λ.

Answer:

1. The λ chosen by LOOCV is 100.
2. With regularzation, 0.1332, Without regularzation, 1.0256.
3. With regularization, ETrain: 0,ETest, 5.9769%.
   Without regularization, ETrain: 0,ETest, 12.6067%.

(b) Implement Logistic Regrssion, and use LOOCV to tune the regularization parameter λ.

Answer:

1. The λ chosen by LOOCV is 0.1.
2. With regularization: ETrain: 0, ETest: 6.0773%.
   Without regularization:ETrain: 0, ETest: 6.7303%.

Problem 2-3. Bias Variance Trade-off

Let's review the bias-variance decomposition first. Now please answer the following questions:

(a) True of False

Answer:

1. False. Bigger the number of training examples will improve the test error of those models with high variance.
2. False. Because the training data set is small, models with high variance are more likely to perform worse.
3. True. The more complex the model, the higher the variance.
4. False. If the regularization parameter λ reaches a certain level, the model tends to be linear and its performance is reduced.
5. False.