

## Homework 2

### Collaborators:

Name: Youchao Zhang      Student ID: 3170100125

---

### Problem 2-1. A Walk Through Linear Models

#### (a) Perceptron

##### Answer:

1. When the train set = 10 :

The train error = 0.0%  
The test error = 10.63%

When the train set = 100 :

The train error = 0.02%  
The test error = 1.36%

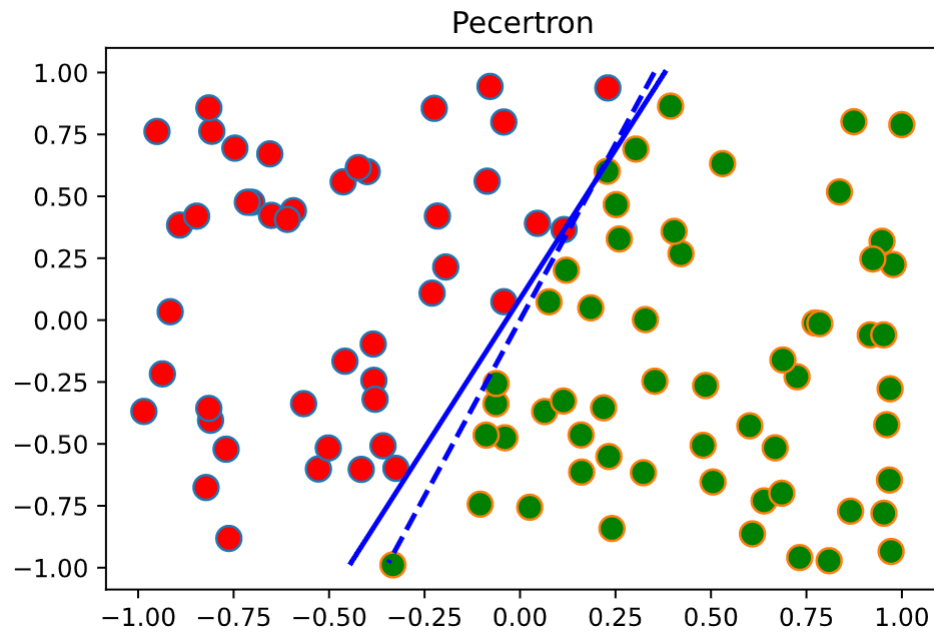
2. When the train set = 10 :

The average number of iterations = 6.39

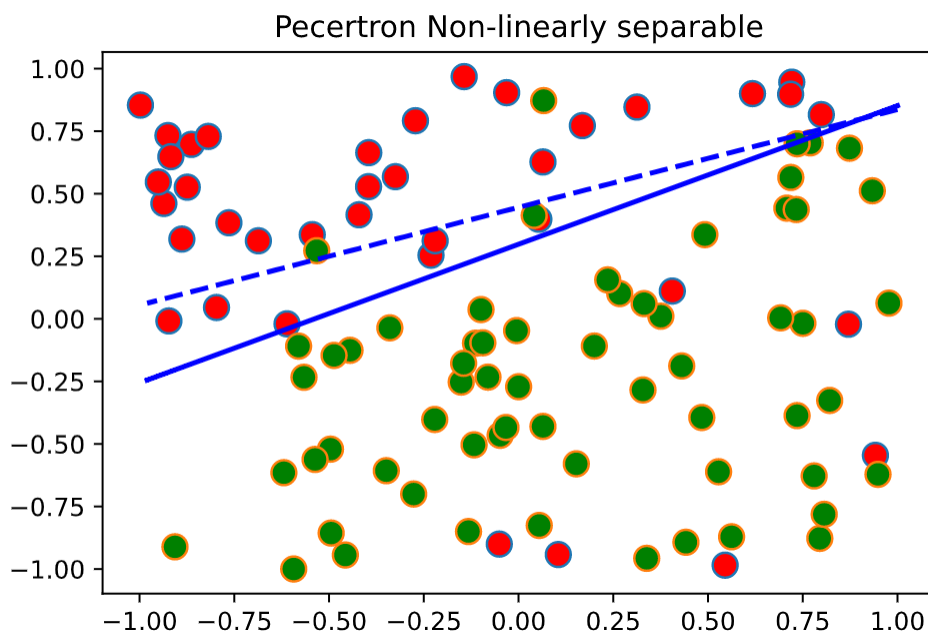
When the train set = 100:

The average number of iterations = 29.41

3. Perceptron is a linear model , so if the training data is not linearly separable ,the algorithm will perform poorly.The error rate will increase significantly.



**Figure 1:** The plotting result for perceptron when  $n_{\text{Train}} = 100$ .



**Figure 2:** The plotting result for perceptron when training data is not linearly separable.

**(b) Linear Regression**

**Answer:**

1. When the train set = 100 :

The train error = 3.91%

The test error = 4.79%

2. if the training data is noisy and not linearly separable:

The train error = 13.26%

The test error = 14.45%

3.

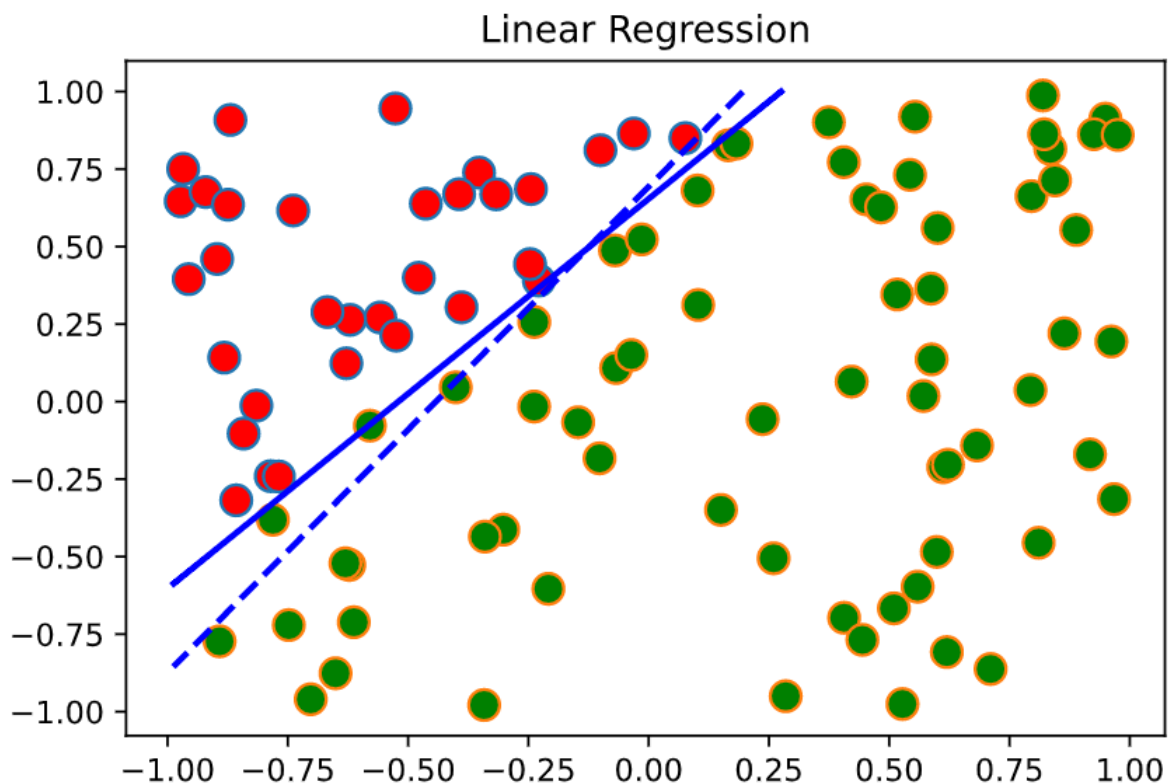
The train error = 49.00%

The test error = 54.96%

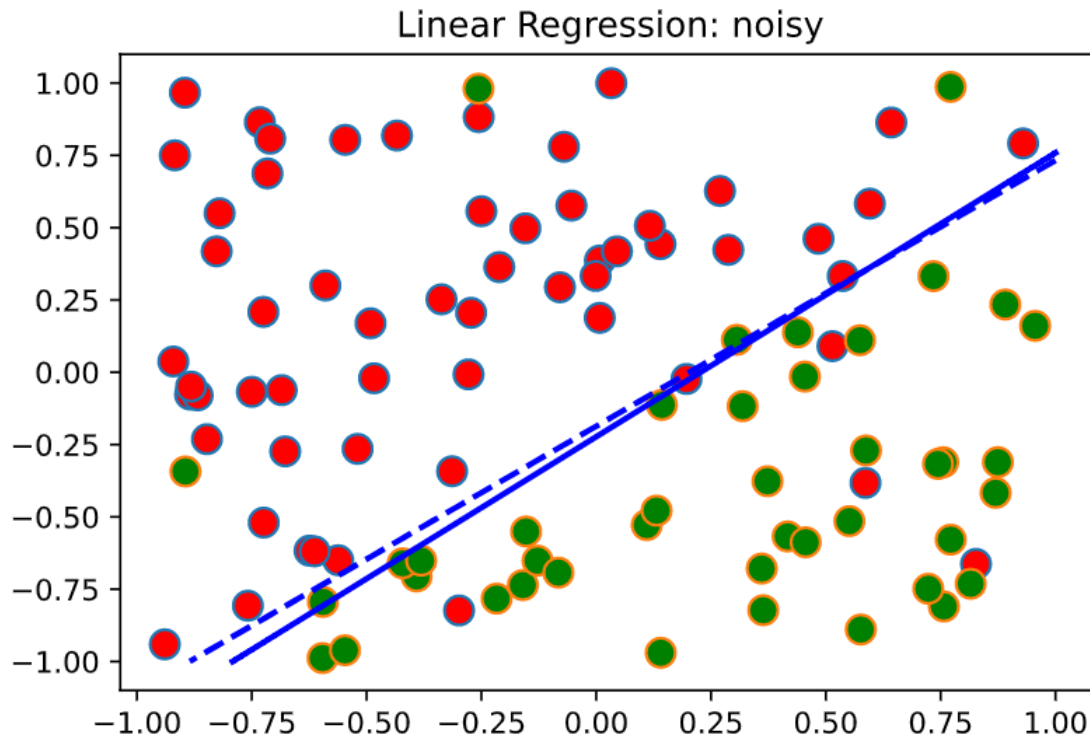
4. After transformation:

The train error = 5.00%

The test error = 6.60%



**Figure 3:** The plotting result for linear regression.



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

**(c) Logistic Regression**

**Answer:**

1.

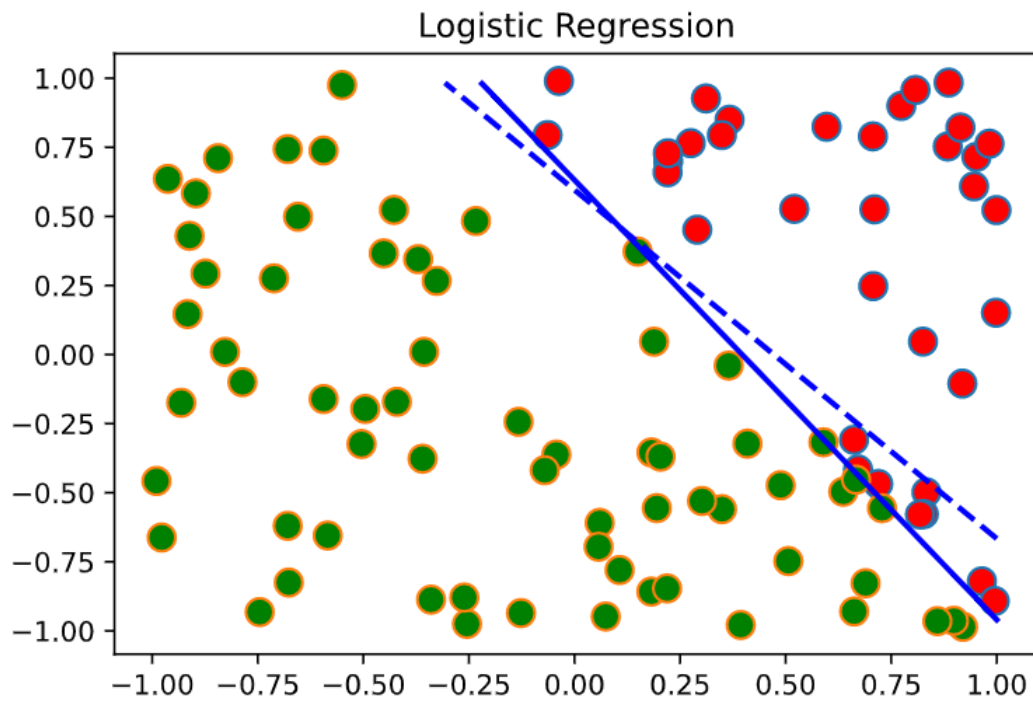
The train error = 6.23%

The test error = 7.07%

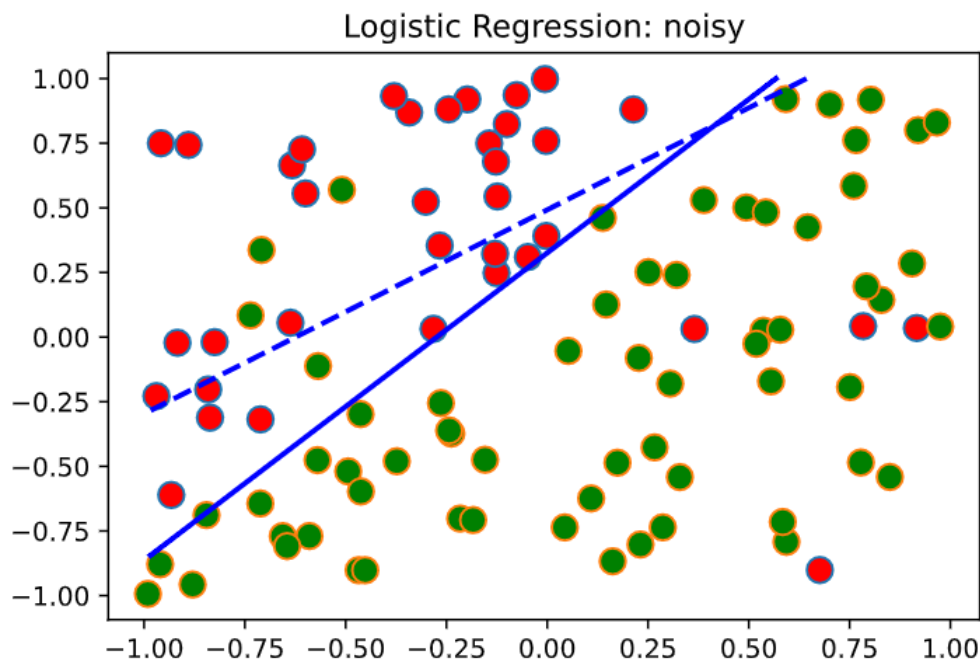
2.

The train error = 14.05%

The test error = 15.61%



**Figure 5:** The plotting result for logistic regression.



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

**(d) Support Vector Machine****Answer:**

1. When the train data set is 30:

The train error = 0.00%

The test error = 3.45%

2. When the train data set is 100:

The train error = 0.00%

The test error = 1.11%

3. For the case  $n_{\text{Train}} = 100$ , average number of support vectors in my trained SVM models is 3.215.

4. SVM with noisy training data(bonus)

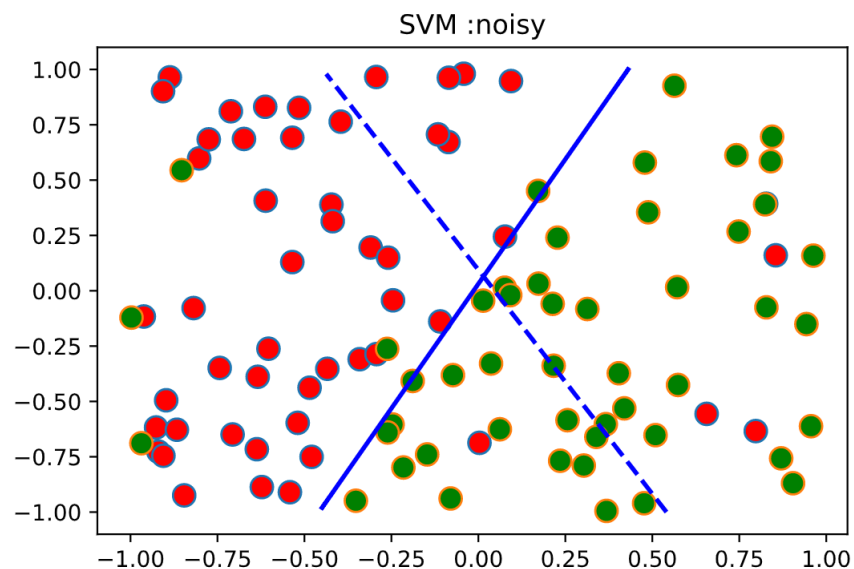
The train error = 32.0%

The test error = 27.6%

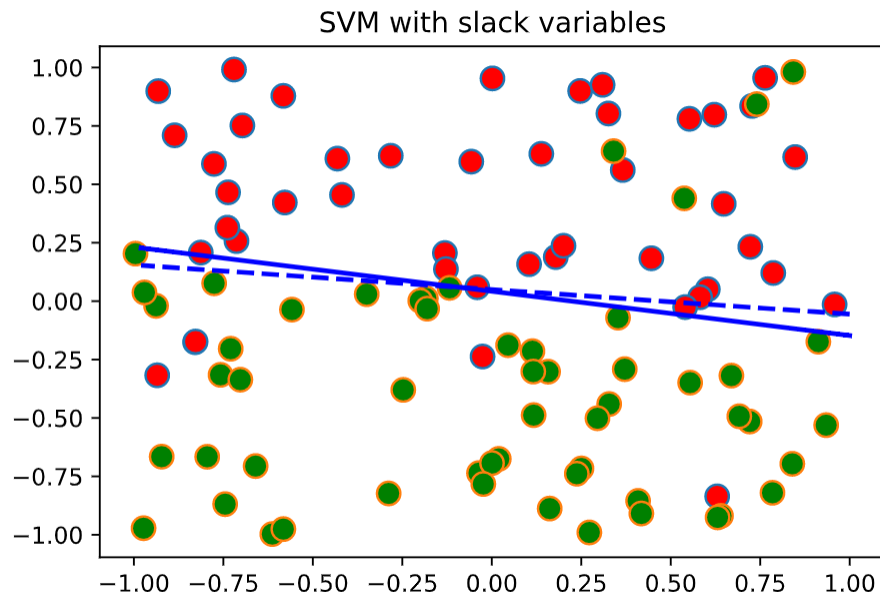
Use slack variables

The train error = 11.0%

The test error = 11.1%



**Figure 7:** SVM with noisy data.



**Figure 8:** SVM with Slack variables.

### Problem 2-2. Regularization and Cross-Validation

(a) Implement Ridge Regression, and use LOOCV to tune the regularization parameter  $\lambda$ .

**Answer:**

1. The  $\lambda$  chosen by LOOCV = 100.
2.  $\lambda = 0$

$$\sum_{n=1}^m \omega_i^2 = 0.892$$

$$\lambda = 100$$

$$\sum_{n=1}^m \omega_i^2 = 0.113$$

3. No regularization

The train error = 0.00%

The test error = 9.04%

No regularization

The train error = 0.00%

The test error = 6.47%

- (b) Implement Logistic Regression, and use LOOCV to tune the regularization parameter  $\lambda$ .

**Answer:**

1. The  $\lambda$  chosen by LOOCV = 0.001.

No regularization

The train error = 0.00%

The test error = 5.07%

Use regularization

The train error = 0.00%

The test error = 5.07%

### **Problem 2-3. Bias Variance Trade-off**

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

- (a) True or False

**Answer:**

1. F
2. T
3. T
4. F
5. F