Machine Learning: Assignment #2

Name:杨朔柳 Stu ID:21721078

**1. A Walk Through Linear Models**

(a) Perceptron

(i)

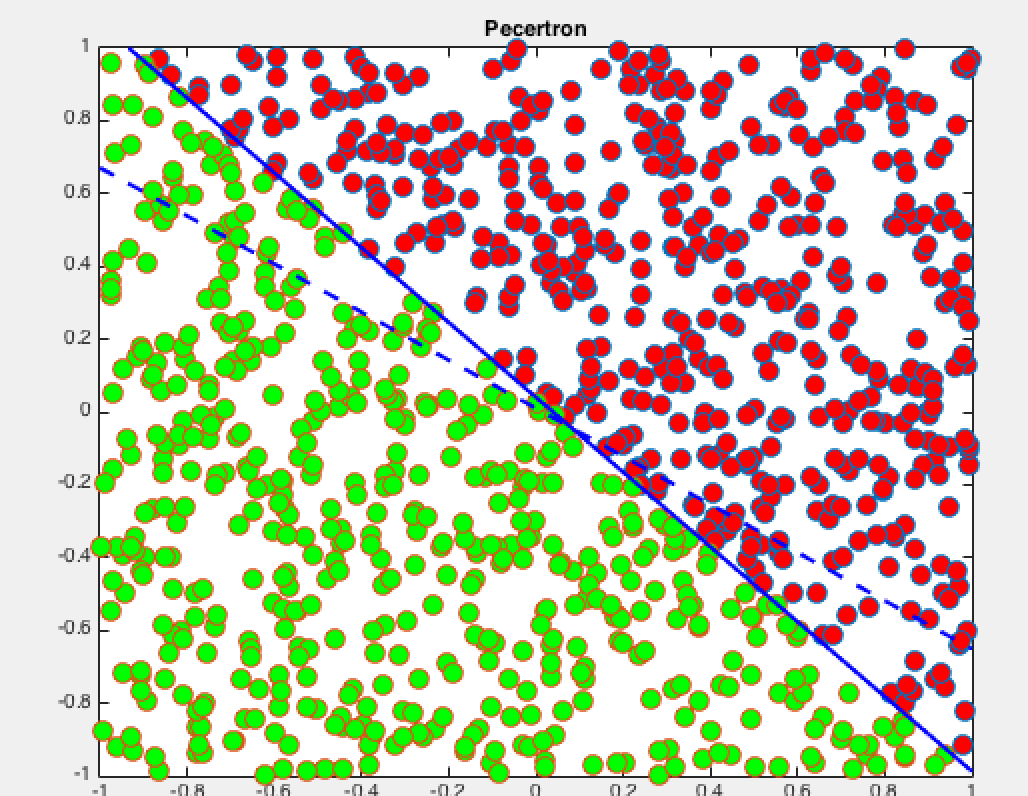
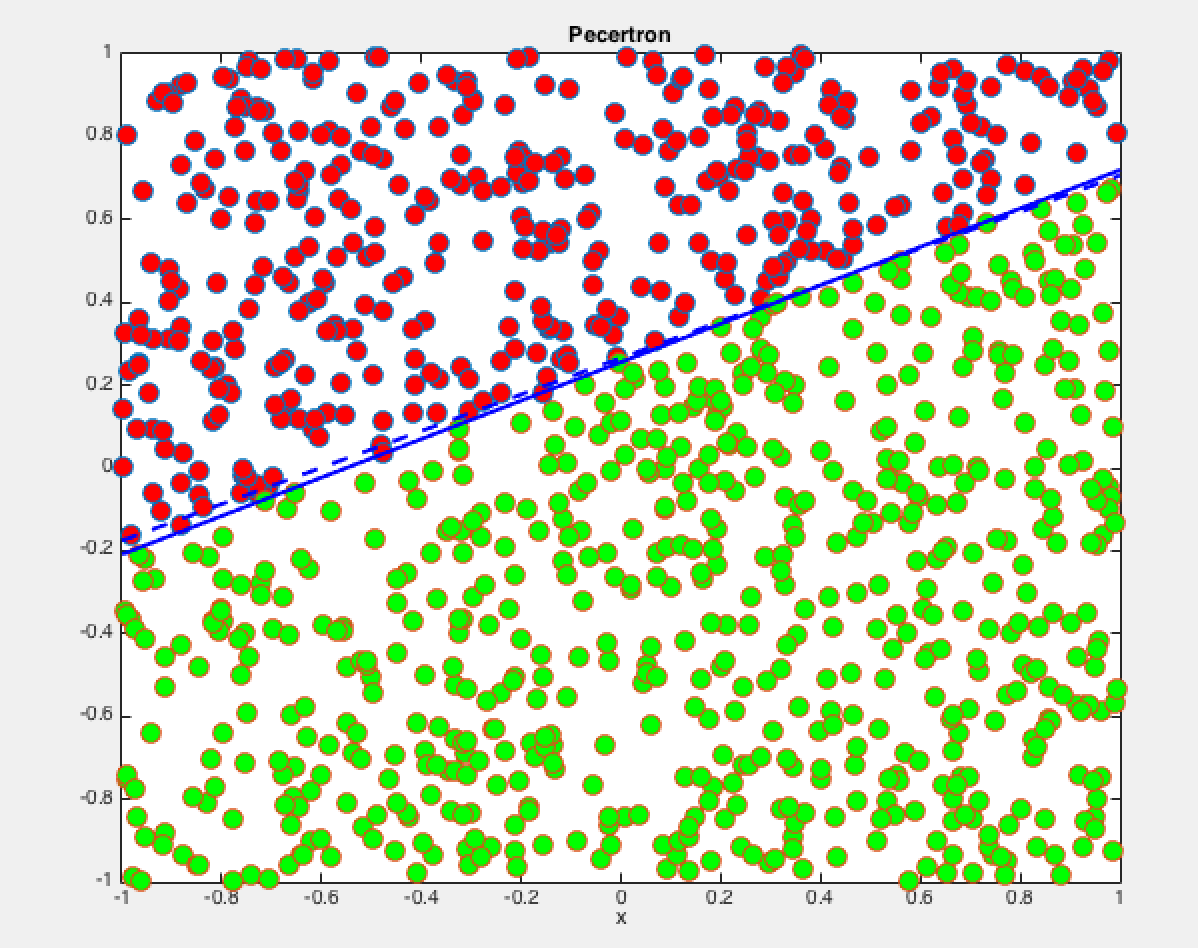
|  |  |  |
| --- | --- | --- |
|  | Train error | Test error |
| nTrain= 10, nTest = 1000 | 0 | 0.111 |
| nTrain = 100 ,nTest =1000 | 0 | 0.014 |

(ii)

|  |  |
| --- | --- |
|  | Average Iterations |
| nTrain = 10 | 10 |
| nTrain = 100 | 452 |

(iii)

The algorithm won’t stop.

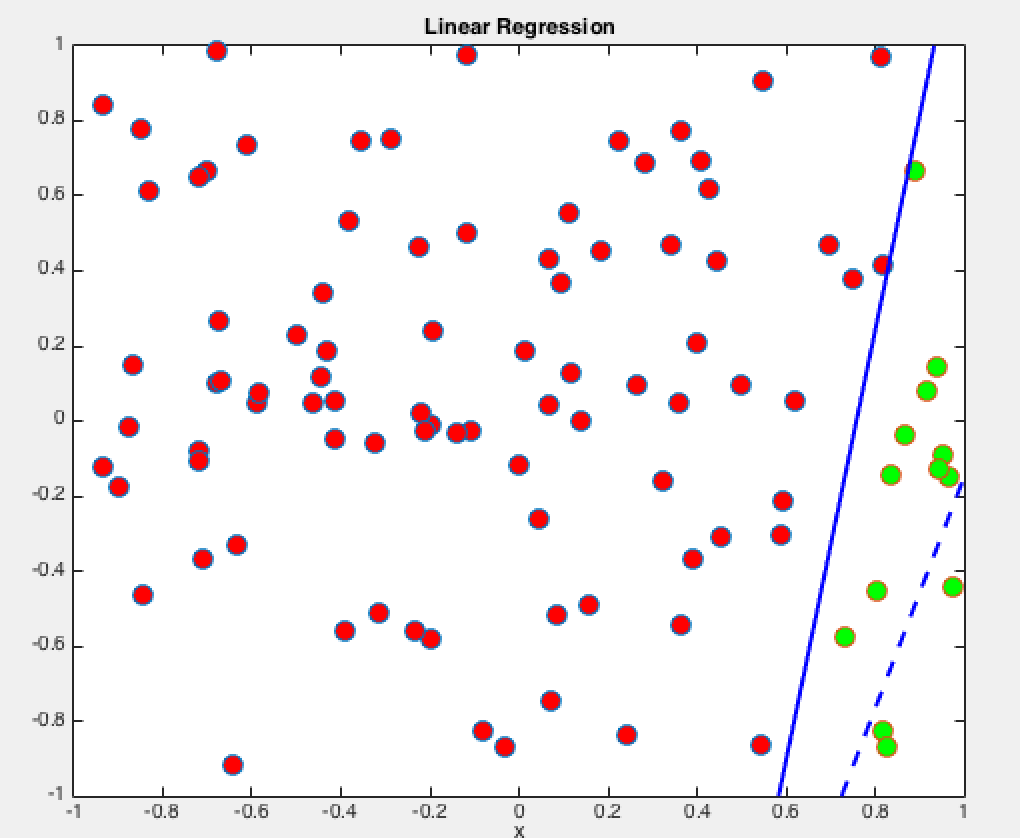


(1)100 train samples and 1000 test sample (2)10 train samples and 1000 test sample

(b) Linear Regreesion

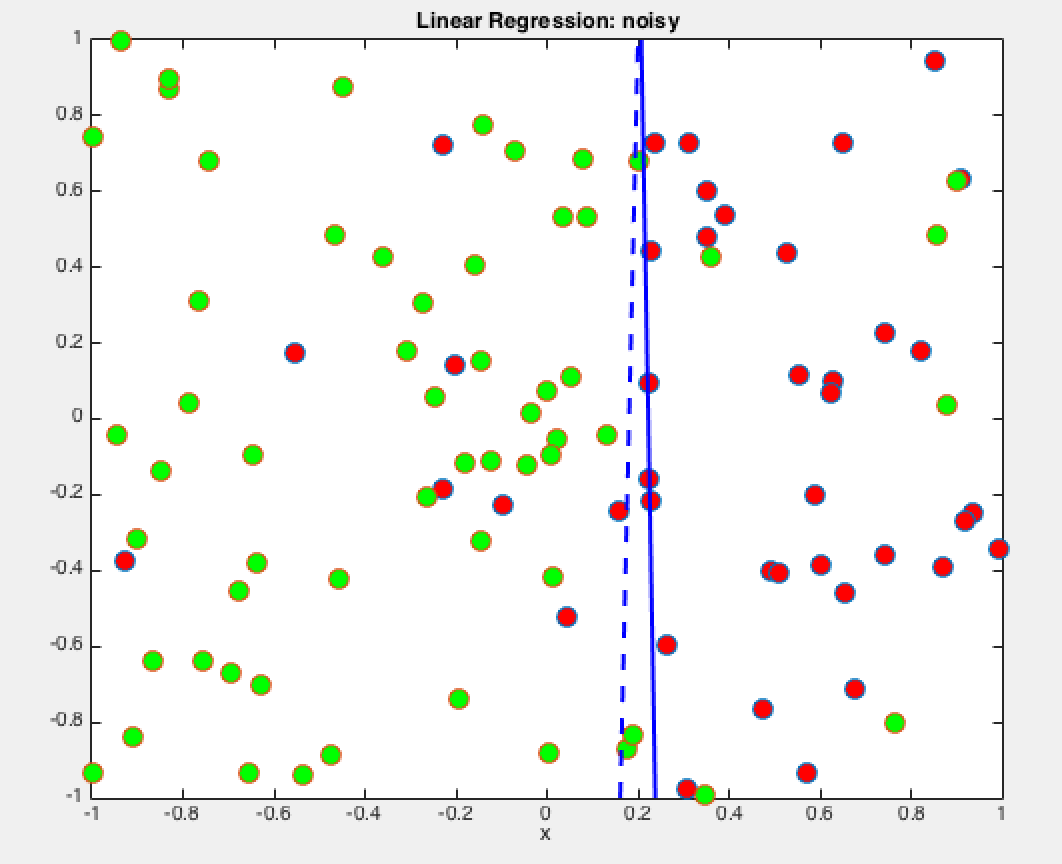
(i)

|  |  |
| --- | --- |
|  | Test error |
| nTrain = 100 nTest = 100 | 0.048 |



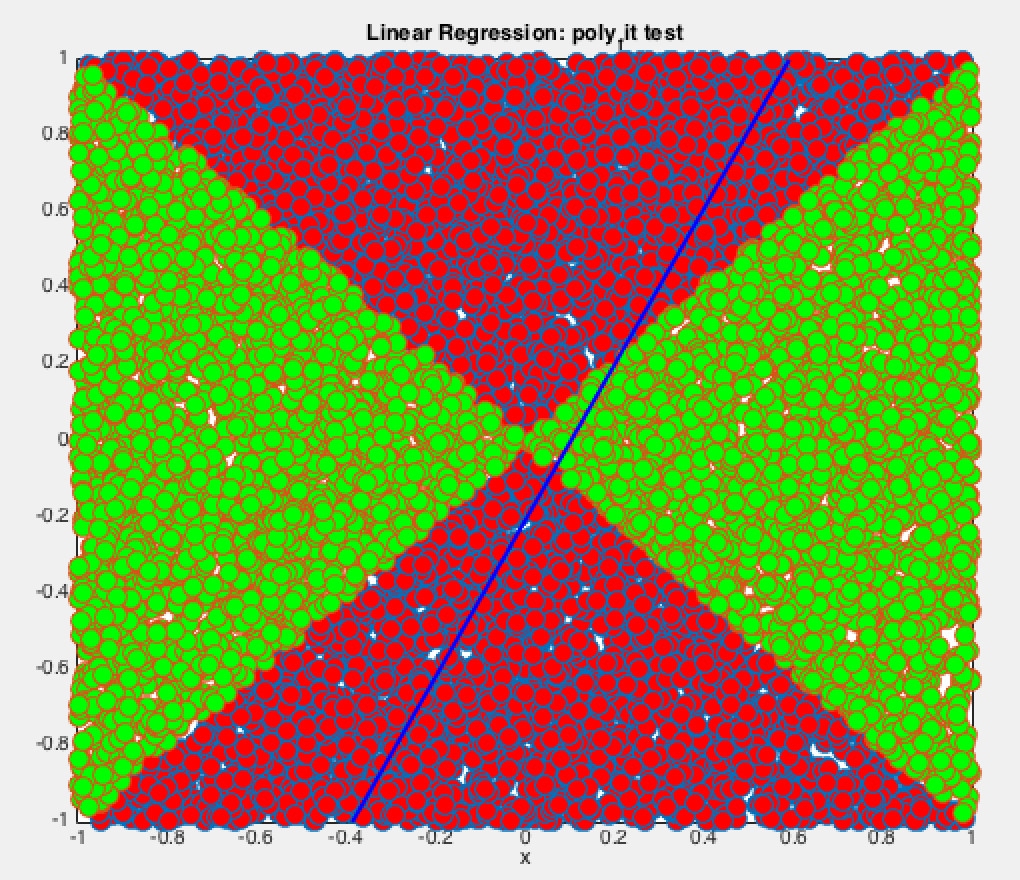
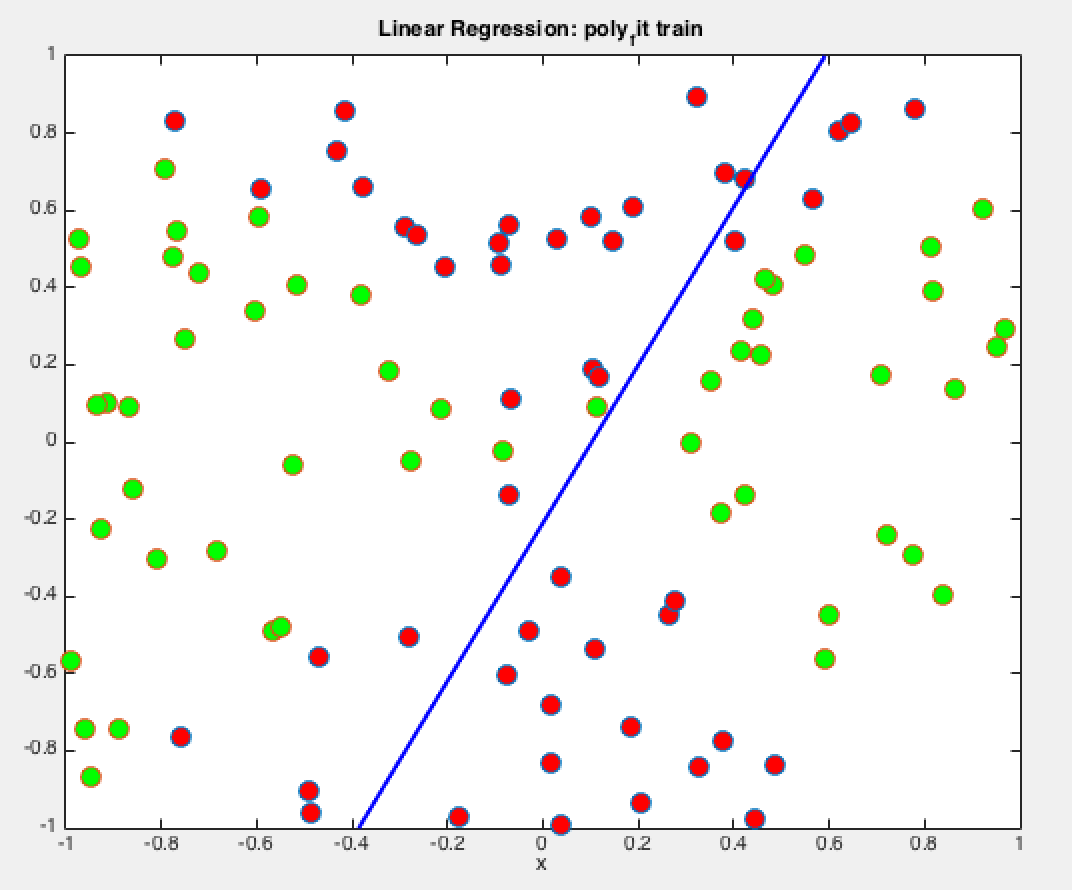
(ii)

|  |  |  |
| --- | --- | --- |
|  | Training error | Test error |
| nTrain = 100 nTest = 100 | 0.136 | 0.149 |



(iii)

|  |  |  |
| --- | --- | --- |
|  | Training error | Test error |
| nTrain = poly\_train,nTest = poly\_tset | 0.49 | 0.5496 |



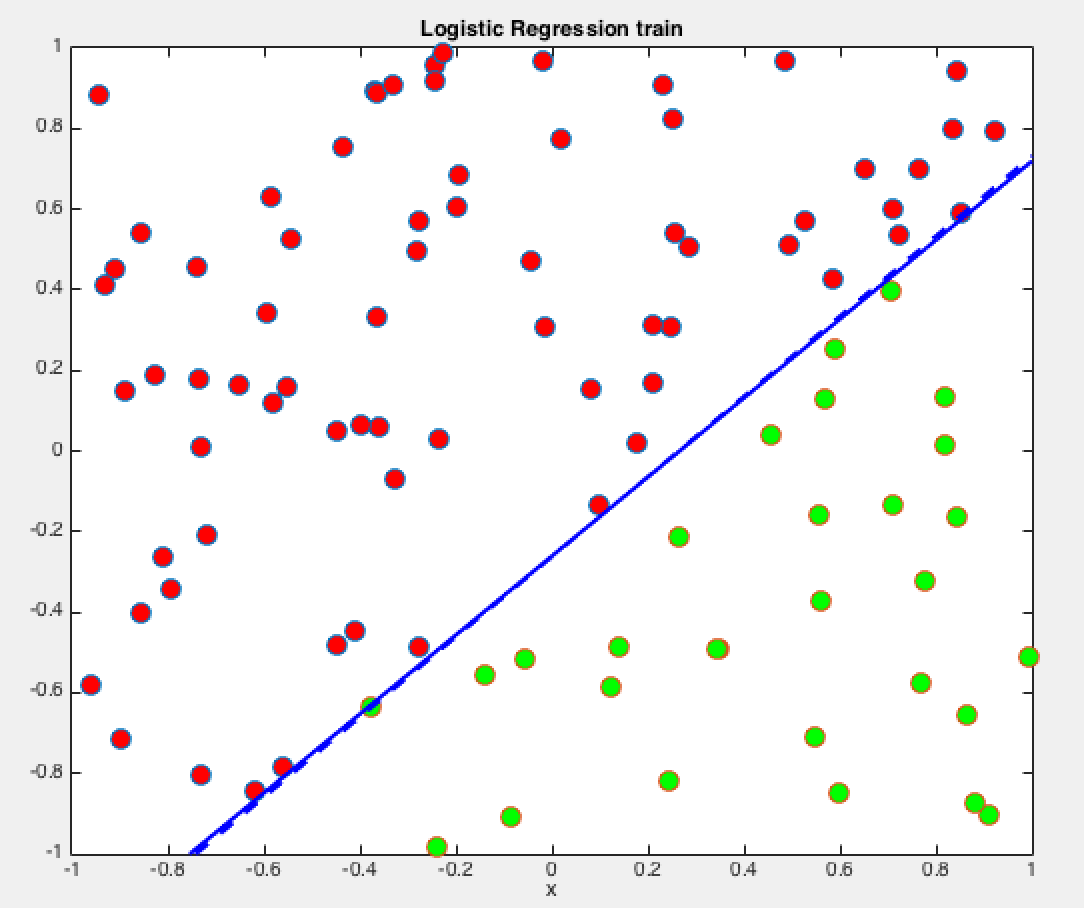
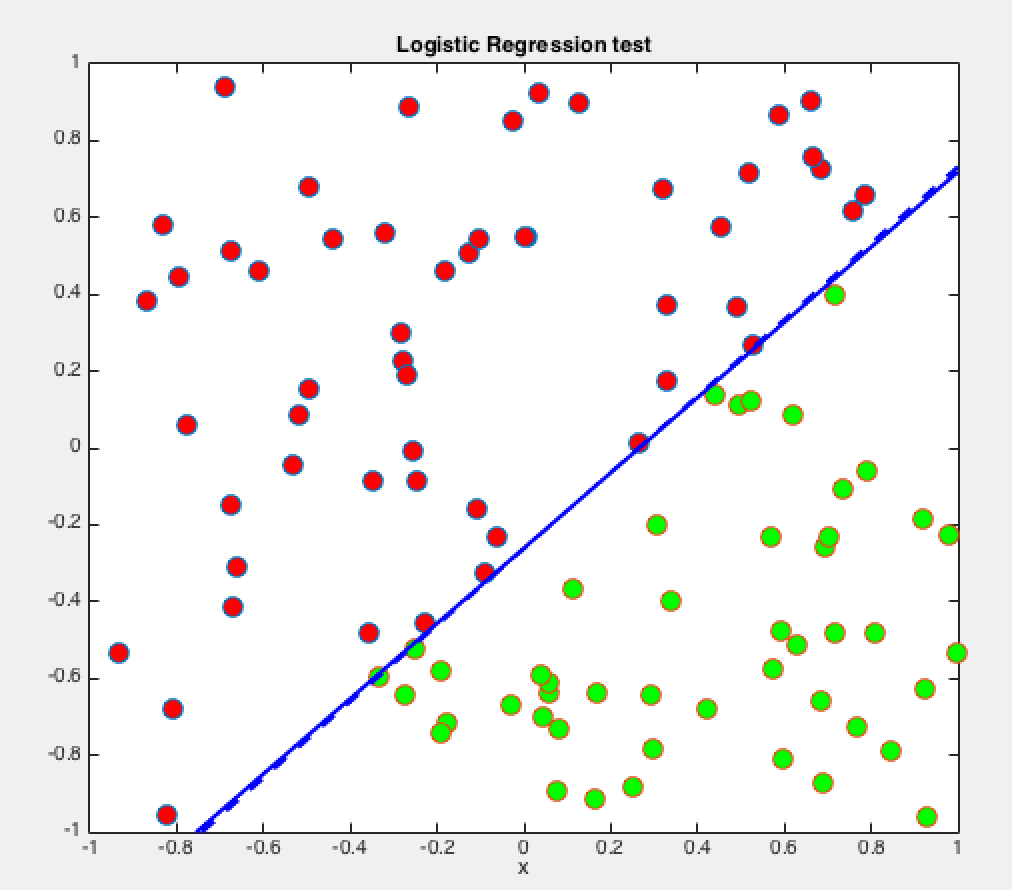
(iv)After trainsformation

|  |  |  |
| --- | --- | --- |
|  | Training error | Test error |
| nTrain = poly\_train,nTest = poly\_tset | 0.05 | 0.066 |

(C)Logistic Regression

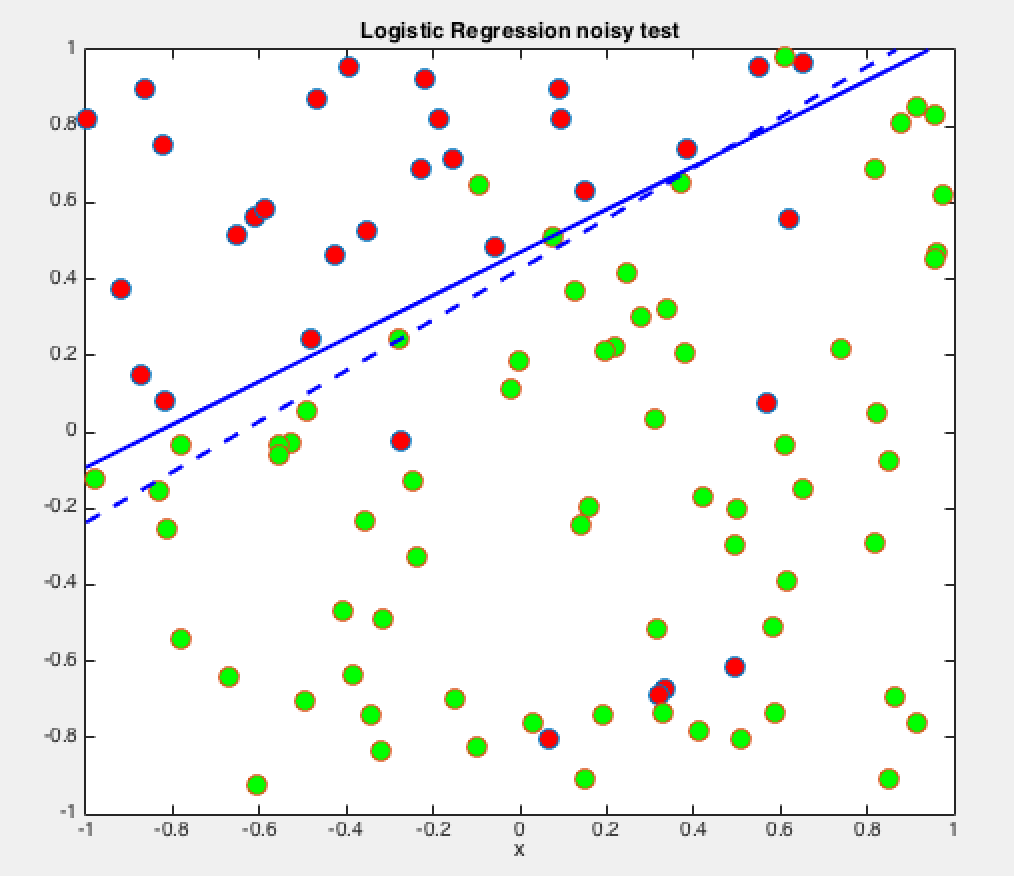
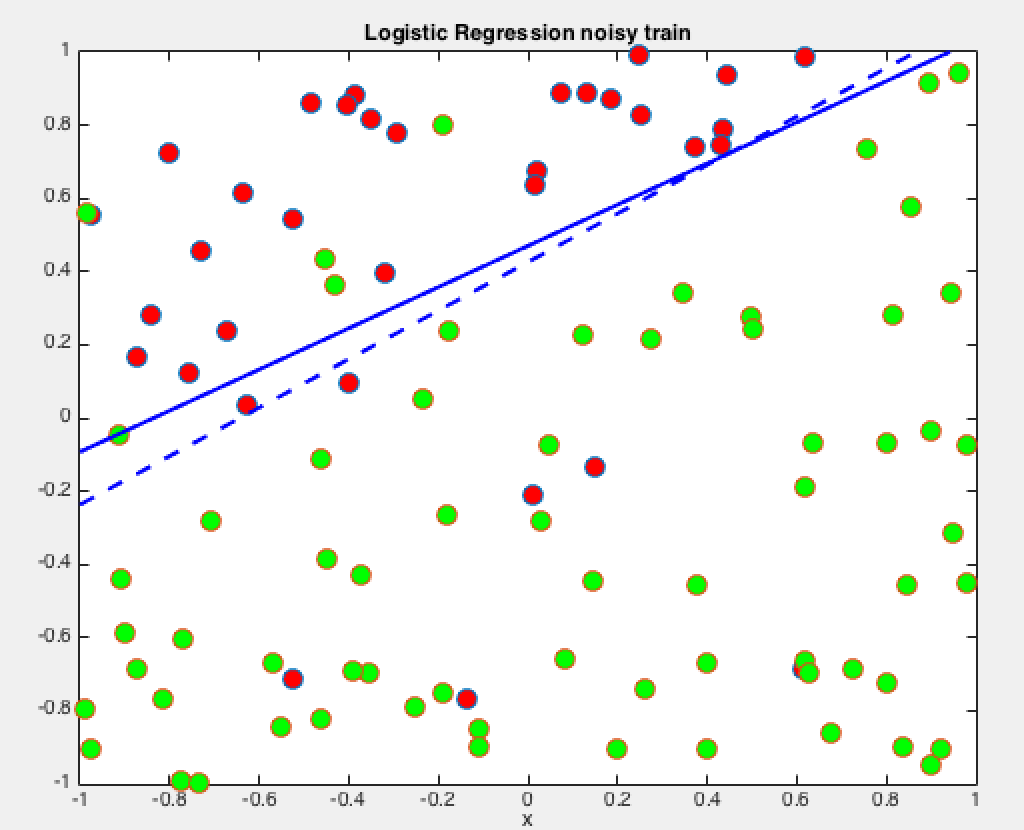
(i)

|  |  |  |
| --- | --- | --- |
|  | Training error | Test error |
| nTrain = 100 nTest = 100 | 0.0071 | 0.0164 |



(ii)

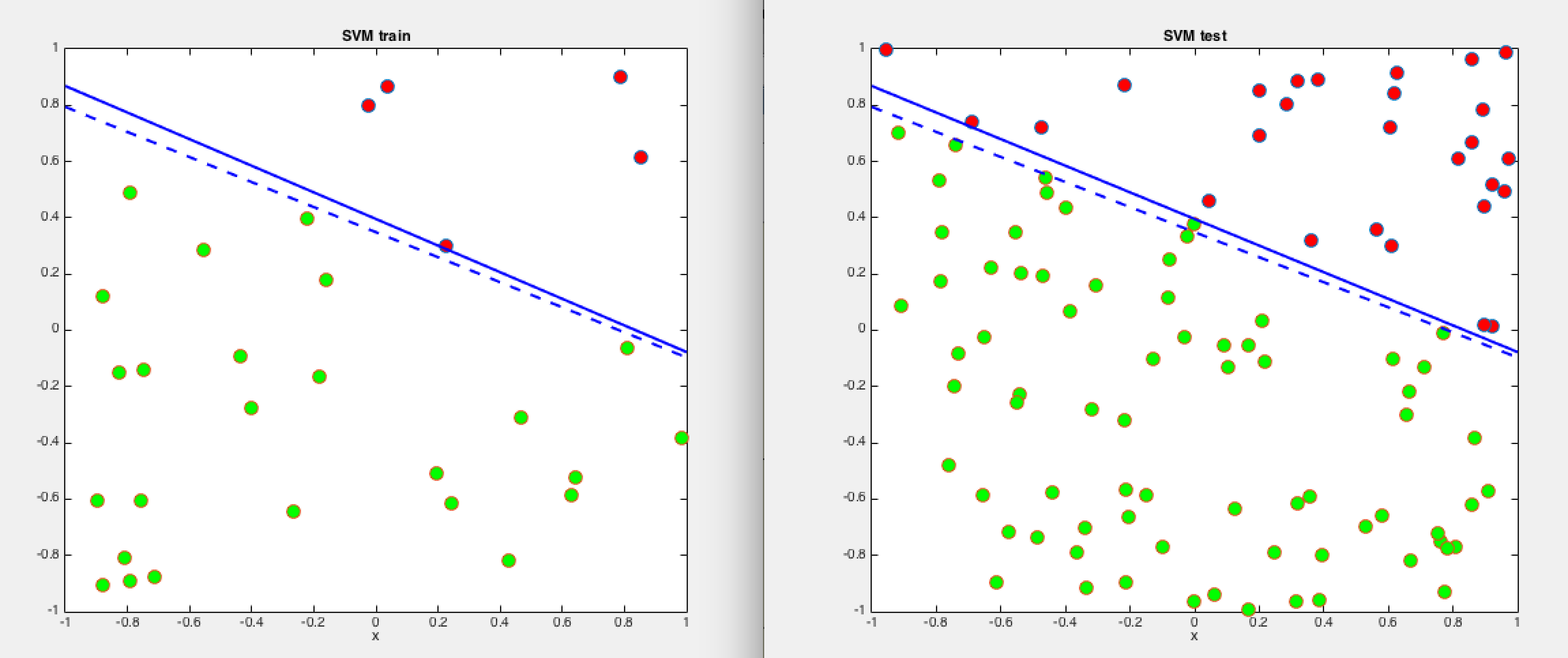
|  |  |  |
| --- | --- | --- |
|  | Training error | Test error |
| nTrain = 100 nTest = 100 | 0.228 | 0.2464 |



(d)Support Vector Machine

(i)

|  |  |  |
| --- | --- | --- |
|  | Train error | Test error |
| nTrain = 30 ,nTest = 100 | 0 | 0.03187 |



(ii)

|  |  |  |
| --- | --- | --- |
|  | Train error | Test error |
| nTrain = 100,nTest = 1000 | 0 | 0.010676 |

(iii)

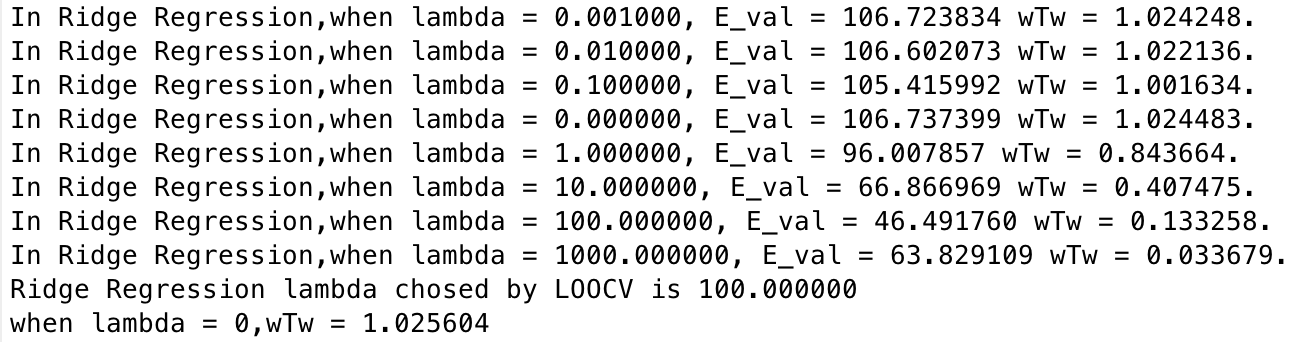
|  |  |
| --- | --- |
|  | Number of average support vectors |
| nTrain = 100, nTest =1000 | 2.676 |

**2. Regularization and Cross-Validation**

(a)

(i)lambda =100

(ii)||w|| = 1.025064 when lambda = 0

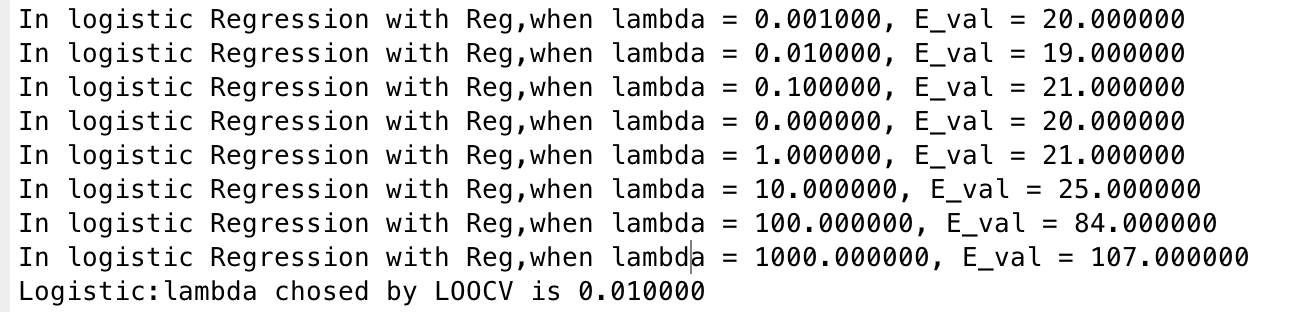


(iii)

|  |  |  |
| --- | --- | --- |
|  | Train error | Test error |
| Lambda = 100,nTrain = 200,nTest = 1991 | 0 | 0.059 |
| Lambda = 0 nTrain = 200,nTest = 1991 | 0 | 0.126067 |

(b)

|  |  |  |
| --- | --- | --- |
|  | Train error | Test error |
| Lambda=0.01,nTrain=200,nTest=1991 | 0.025 | 0.1160 |
| Lambda=0,nTrain=200,nTest=1991 | 0.51 | 0.1145 |

Lambda = 0.01

(best lambda floats among 0.01~0.0)

**3. Bias Variance Trade-off**

(a)

(i)False. If adding more training example, the test error will decrease first and then increase.

(ii)False. Model with high variance is overfitting. It works badly on test samples, which is more important for our task.

(iii)True. Model with more parameters could be a very complicated, and it is more prone to overfitting

(iv)False. Regularization is aim at working well on test dataset, not training dataset.

(v)False. Large λ will lead to underfitting. On the contrary, Small λ will lead to overfitting,

(vi)False. Too Small or too large λ won’t benefit our model.