Machine Learning: Assignment 2

Gong Lixue 21721093

October 22, 2017

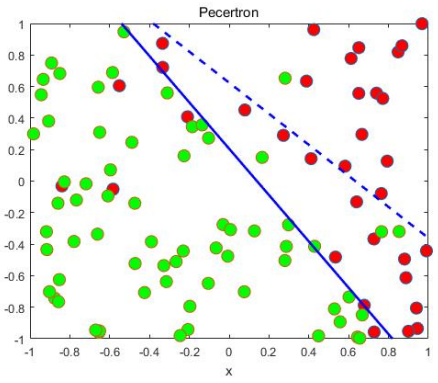
1 A Walk Through Linear Models

1.1 Perceptron

(i) I limit the maximum times of iteration(in perceptron.m) within 1000 to prevent the situation that the dataset is not linearly separable. And each iteration, I generate a large dataset with sizeof 1000 to estimate test error. When nTrain=10, the training error rate is 0.0024, and the test error rate is 0.1437. When nTrain = 100, the training error rate is 0.000010, and the test error rate is 0.0180.

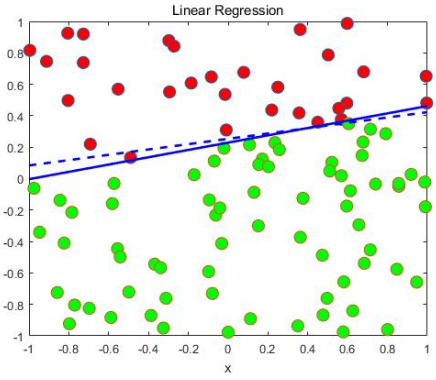
(ii) When nTrain=10, the average number of iterations is 284. When nTrain = 100, the average number of iterations is 114.

(iii) When the training data is not exactly linearly separable, we can not guarantee that our model can classify all the training data correctly. And due to the limitation of the maximum times of iteration(1000), the model we trained may be not the optimal. As the following figure, when the training data is not strictly linearly separable, the algorithm can not give a optimal solution.

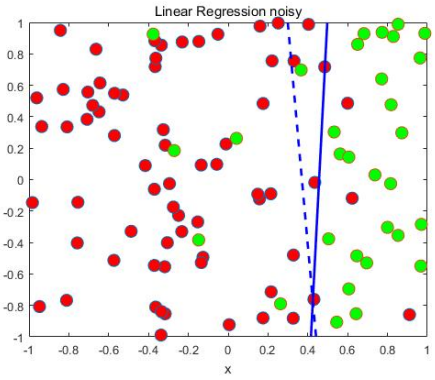


1.2 Linear Regression

(i) the the trainning set is 100, trainning error rate is 0.0384, and the test error rate is 0.04809. The decision boundary is shown below.



(ii) When the dataset is not linearly separable, the trainning error rate is 0.1375, and the test error rate is 0.0620. The decision boundary is shown below.



(iii) For the dataset *poly_train.mat*, the trainint error rate is 0.4900, the testing error rate is 0.5496.

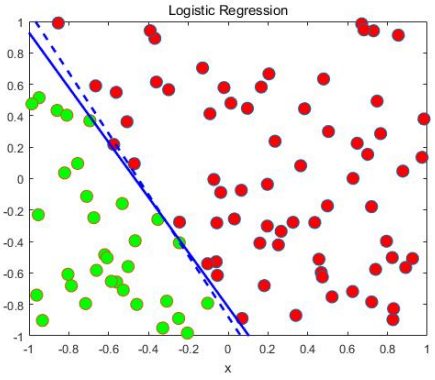
(iv) After transforming the input data, the training error rate is 0.0500, the testing error rate is 0.0660.

1.3 Logistic Regression

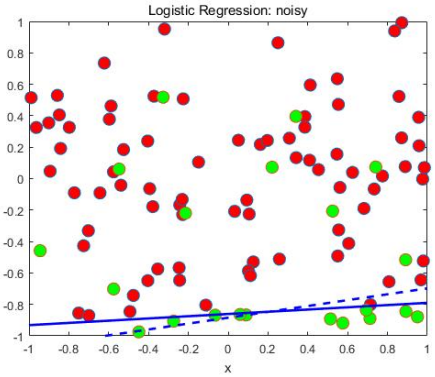
I use following parameters in *logistic.m*:

$learning\_rate = 0.02$ and $max\_iter = 5000$

(i) If the size of training set is 100, the trainning error rate is 0.0191, and the testing error rate is 0.0292. The decision boundary is shown below.

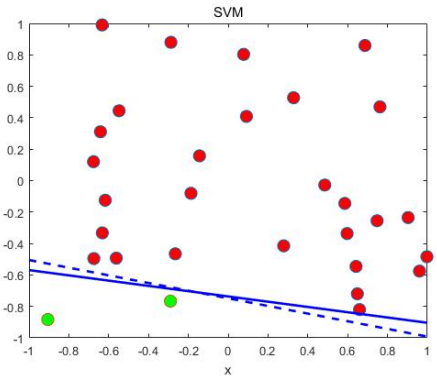


(ii) When the training data is not linearly separable, the training error rate is 0.1284, and the testing error rate is 0.0493. The decision boundary is shown below.

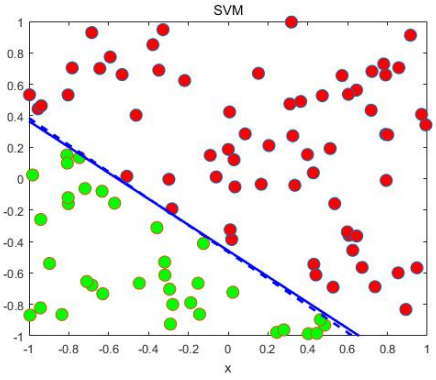


1.4 SVM

(i) If the size of training set is 30, the training error rate is 0.00, and the tesing error rate is 0.0376. The decision boundary is shown below.



(ii) If the size of training set is 100, the training error rate is 0.00, and the tesing error rate is 0.0111. The decision boundary is shown below.



(iii) For the cane nTrain = 100, the average number of support vectors is 2.8440

2 Regularization and Cross-Validation

2.1 Ridge Regression Cross Validation

(i) $\lambda = 1e3$

(ii) When $\lambda = 1e3$, $\Sigma_{i=1}^{m} \omega_i^t = 0.0340$

When $\lambda = 0$, $\Sigma_{i=1}^{m} \omega_i^t = 1.0256$

(iii) When $\lambda = 1e3$, training error rate is , testing error rate is When $\lambda = 0$, training error rate is , testing error rate is