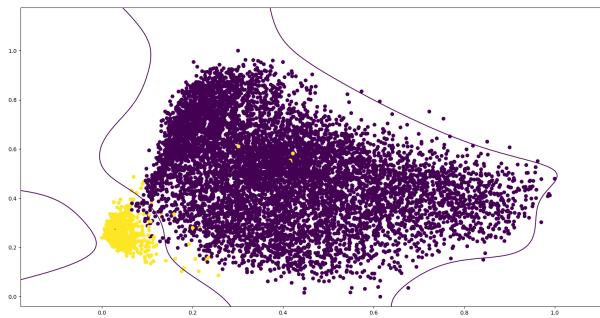# Homework 9

Karan Sarkar
sarkak2@rpi.edu
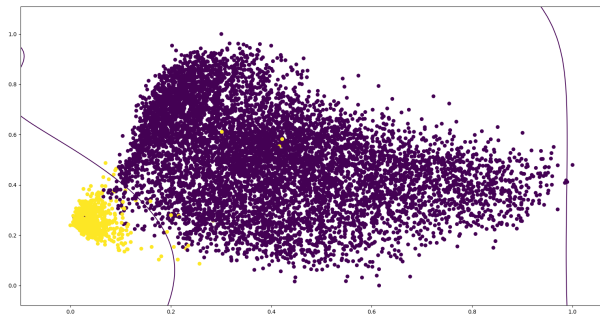
November 12, 2019

1. The dimensions of Z are 300×45.
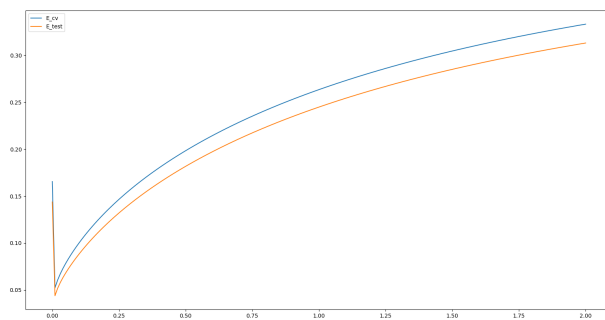
2.

The decision boundary with no regularization overfits the data.

3.

The decision boundary with regularization perhaps underfits the data. The decision boundary has little curvature. Adding some more curvature could provide a greater margin.

4.

   Both the cross-validation error and test error follow the same pattern. They decrease rapidly with a little regularization and then increase gradually. The cross validation error seems to be a little higher than the test error.
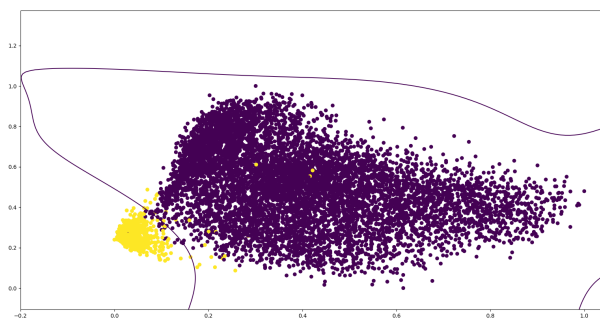
5.

   Using $\lambda^* = 0.01$ seems to be the right amount of regularization.

6. The test error seems to be $E_{\text{test}} = 0.0299$.

7. Cross validation error will tend to be slightly larger than test error because it has slightly less data to train on. However, it will not be biased to the training data.

8. Because we chose the amount of regularization that minimized our test error, we have some selection bias with regards to test error. In particular, the because we chose the minimum test error, the true error will tend to be larger than the test error. This could be fixed by using insample cross validation error to select regularization or making two test sets.