Daniel Campos October 17st, 2013

RCS: Campod2 RIN:660996361

Machine Learning CS 4100

I worked on these problems with Zoe Konrad

Problem Set 7

1. Problem 3.16

Thus

* 1. Super market

|  |  |  |
| --- | --- | --- |
| F🡪 | +1 | -1 |
| +1 | 0 | 1 |
| -1 | 10 | 0 |

CIA

|  |  |  |
| --- | --- | --- |
| F🡪 | +1 | -1 |
| +1 | 0 | 1000 |
| -1 | 1 | 0 |

Based on this we see that for the CIA we wont accept a fingerprint as correct unless we are 99.9(1000/1001)% sure that the fingerprint is correct and when we look at the super market we will accept if we are at least 9% sure(1/11).

1. Problem 3.17
   1. At order one



 ,

🡪🡪

🡪





* 1. Based on the Lagrange method to minimize E1

Macintosh HD:Users:Rob:Desktop:Screen Shot 2013-10-17 at 12.32.52 PM 1.png





solving for when the directional derivatives are   and Emin =around 2.25 which is allt he same result as the negative gradient vector times a static constant()

* 1. 

Thus based on our previous and





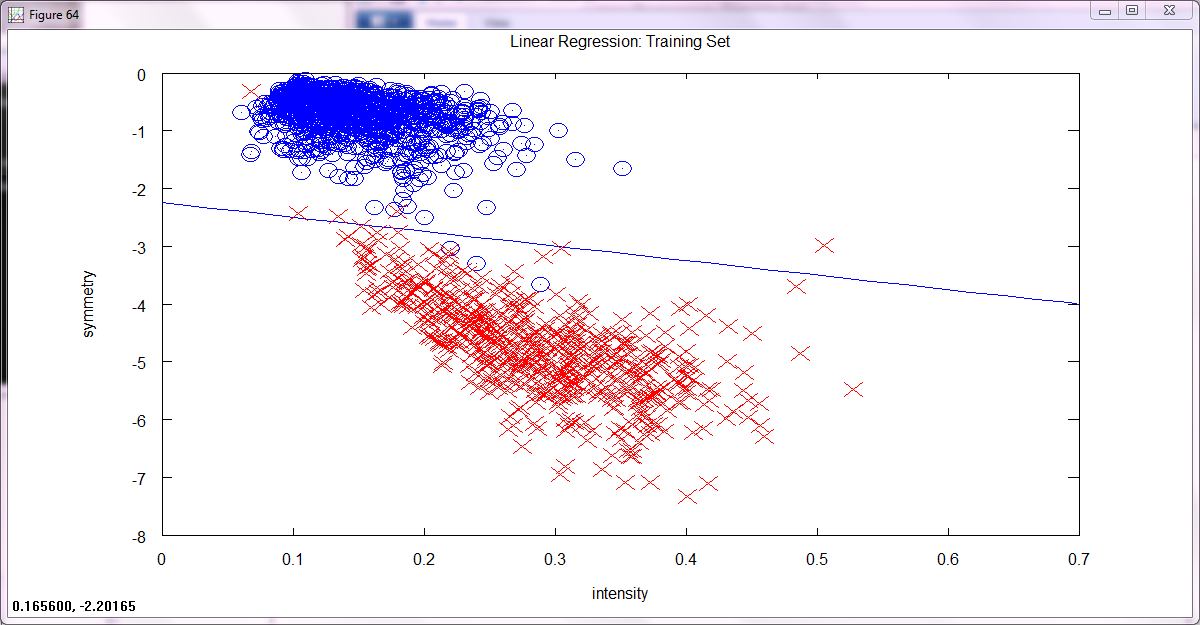




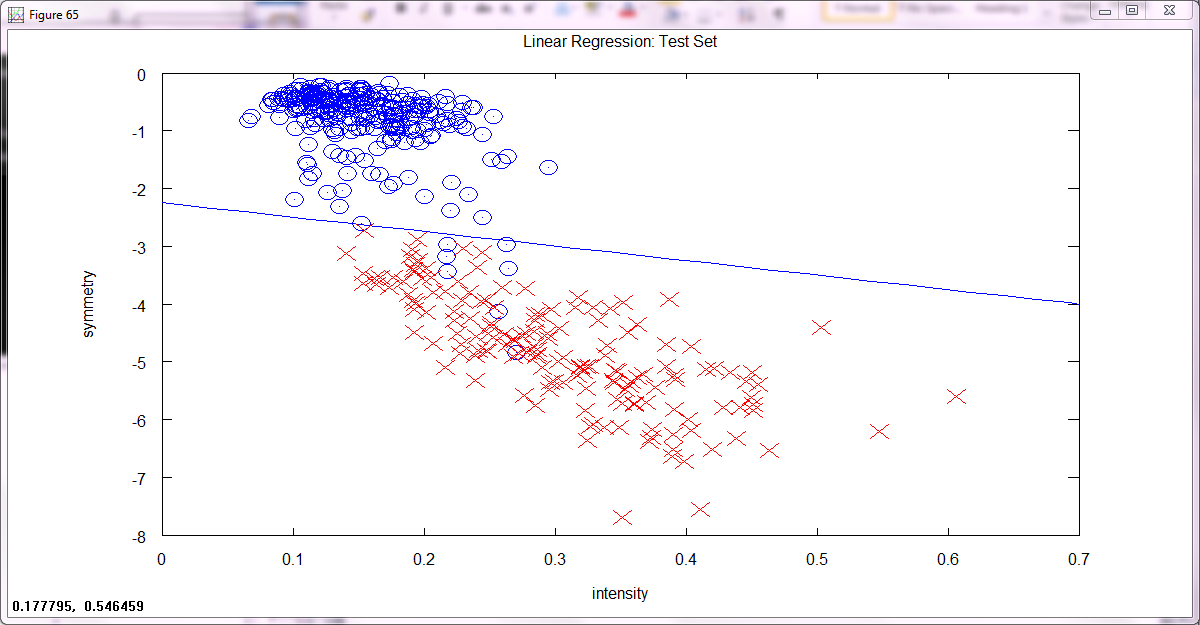
* 1. By deriving both  and then solving the related equations as a result we get at the Newtonian direction at (0,0)
  2. 1. After we normalize the newton direction to a length of 0.5 out Newtonian vector is (0.46875, 0.203125) which we get by multiplying the by 0.5 each giving an optimal value for Emin of about 1.89 when we plug in 
     2. by using and then use a root finding algorithm I found on the internet from values [0,2pi) we obtain the true point at (0.436, 0.246) which gives a Emin=1.87.

While my first 1.89 is not the exact output it is within about 2.1% of the true Emin. When we compare this to the initial Emin of 2.25 we find that using the Taylor series we can approximate better by a factor of 10

1. Classifying Handwritten Digits: 1 vs. 5
   1. Training data



Testing Data



* 1. Training n=1561 Ein=0.005125

Test n=424 Etest=0.01651

* 1. Based on theorem 2.12



thus our traning data our true out of sample error is less than 39% and as to the test data our out of sample error is les that 70% with a centantry of 95%. This is to be expeced because our training set has almost 4 times as many point as our test set.

* 1.  

so our changes in errors are

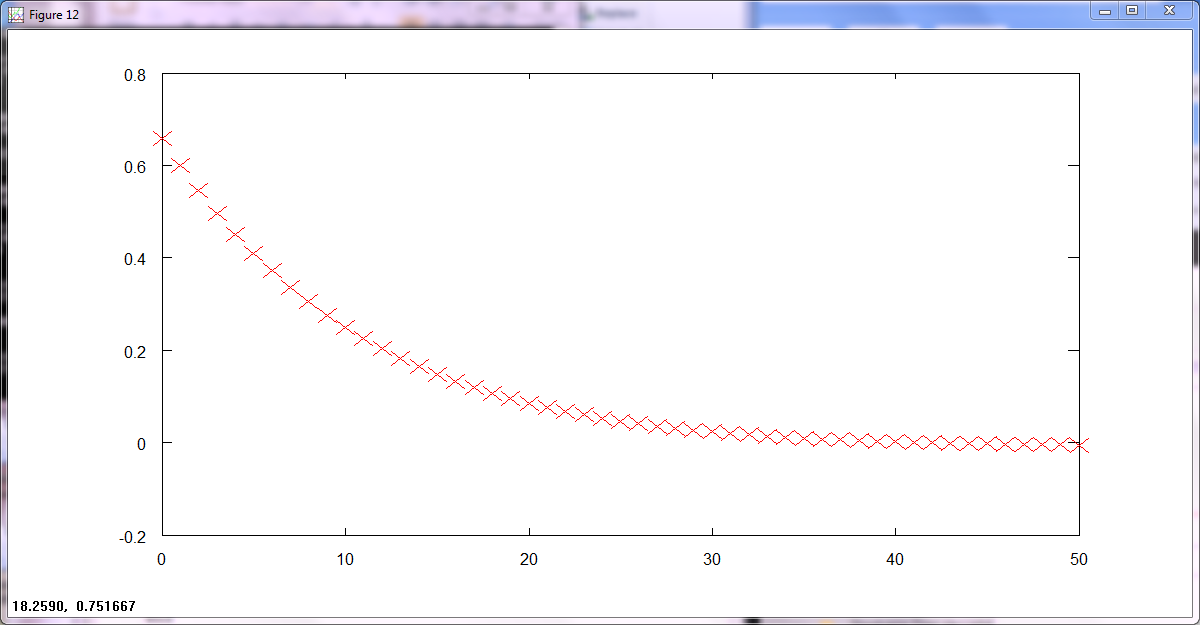
Ein =. 000241

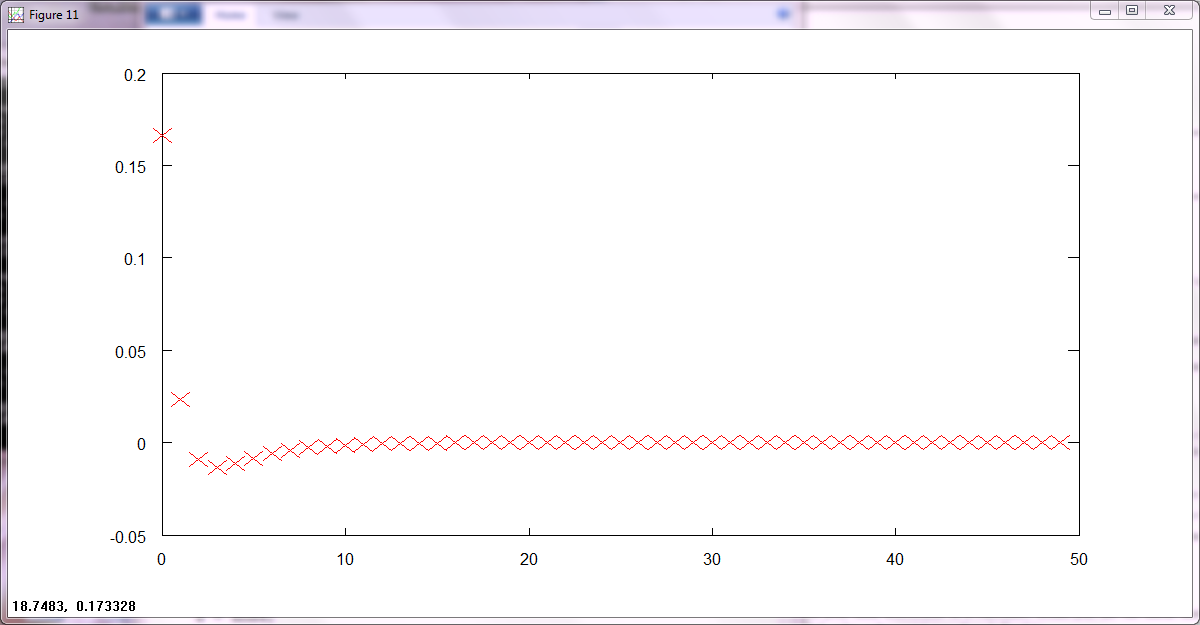
ETest=0.011

* 1. It is better to deliver a linear model since the vc bound increases greatly with the transform. Though our Ein error went down and so did our e error, out bound on certainty has gone terribly wrong . When we do this the bound for our training set is up to 70% and our test set is at 123%(Horrible.)

1. Gradient Descent on a “Simple” Function



* 1. .01

.1

* 1. based on value of 0.1 for n

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0.1,0.1 | 1,1 | −0.5, −0.5 | −1, −1 |
| Value | .720983 | 3 | .75 | 3 |
| Minimum Location | 4.4208E-006, -1.043E006 | 3.641E-005, -8.5952E-006 | -1.8205E-005,4.2976E-006 | -3.641E-005,8.5952E-006 |
| Minimum Value | -3.4255E-10 | -2.3236E-8 | -5.8090E-9 | -2.3236E-8 |

based on value of 0.01 for n

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0.1,0.1 | 1,1 | −0.5, −0.5 | −1, −1 |
| Minimum Value | .720983 | 3 | .75 | 3 |
| Minimum Location | 0.0302319,-0.0025037 | 0.0302319,-0.0025037 | -.151160, 0.012519 | 0.0302319,-0.0025037 |
| value | -.005014 | -.204 | -.1047 | -.204 |