Final Latex Sheet

Cameron Brenner

May 15 2019

1 Problem 1

In question 1, we are asked to run the given code, then submit our best CV score to a Google survey. My best hyper-parameter combination was

```
C 0.03125, epsilon 0.0, gamma 2.0. Testing set CV score: 0.051664
```

Then we are asked to explain our results. The hyper-parameter C is the weight given to the "punishment" of classification, also called the soft margin. A high C parameter will mean that the program will want to stay as close to the given point as possible, within the gamma parameter. There is great penalty when increasing C. The relationship between C and gamma is that gamma is the "range" around a certain point where the data is being tested. If the point is classified by the SVM outside of the given gamma, then the C will react negatively to that based on its value. Low C values mean that if it is outside the gamma, the CV will not increase as harshly if it was a high C value. The epsilon parameter represents the how sensitive the support vectors are in classifiying the points. It's helpful in trying to find a smooth solution when noise is accounted for. Higher epsilon helps with more noise. C is the cost of going outside the given gamma, and epsilon makes the C less sensitive to the misclassified points.

2 Problem 2

In question 2 we are asked to do a similar CV test on a different dataset. We are asked, like above, to give our best CV output, and explain the results. The best CV result was

```
C 64.0, epsilon 2.5, gamma 0.25. Testing set CV score: 0.282228
```

My interpretation of this is that the large gamma "range" around the classification decreased the CV, making it better. Because the range was increased, it was more likely that our training would hit the points more effectively, and there would not be as much need for the soft margin C to provide a different fit to the data. The epsilon and C affect the number of support vectors used to make a regression model, while epsilon is forgiving to noise, C "punishes" the noise. Both of them together make a more complex model.

As of right now, it is 4:38 PM on the Wednesday this is due. My code just started running the test on C 512. I want to hand this in on time, so I am going to leave out the plot. I will just interpret the CV results I have seen, please forgive me. High epsilon and gammas have usually given the best CVs that I have seen. This is because the gamma increases the range and the epsilon increases insensitivity to noise, thus the vector is smoother. But the C punishes misclassified points, so a higher C value means a more accurate classification vector. Therefore, the high C value, with the high epsilon and gamma parameters give the best CV output.