LFD Problem Set 9

John Cohen

November 11, 2024

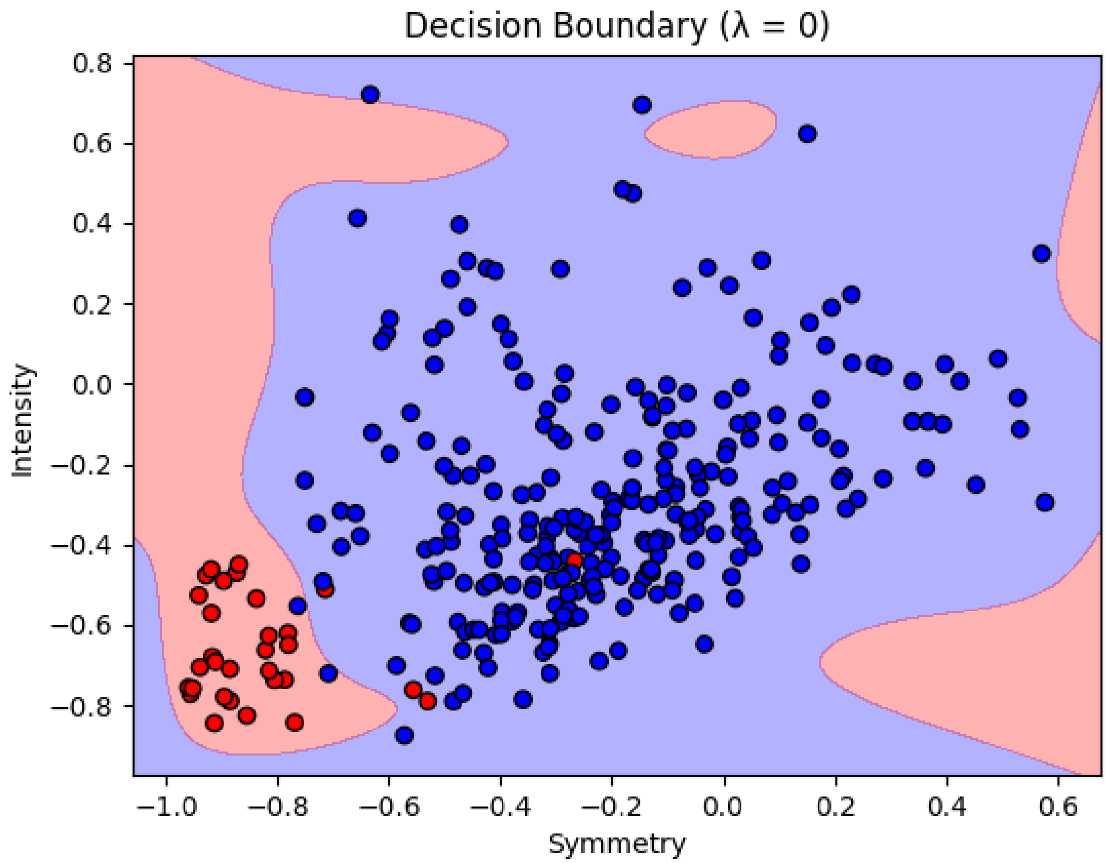
8th order Feature Transform

Use the 8th order Legendre polynomial feature transform to compute Z. What are the dimensions of Z?

The dimensions of Z are 300 by 45. N is 300 and we are working with the 8th order feature transform which gives us 45 dimensions.

Overfitting

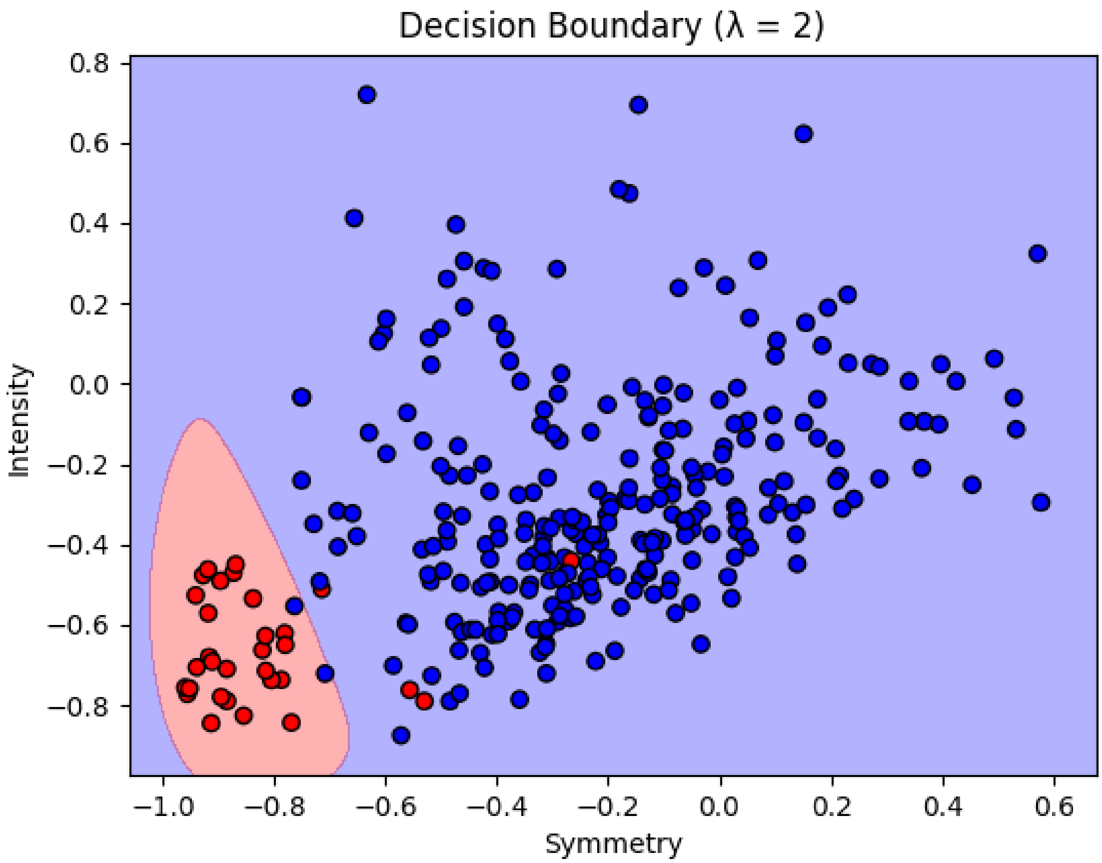
Give a plot of the decision boundary for the resulting weights produced by the linear regression algorithm without any regularization ($\lambda = 0$). Do you think there is overfitting or underfitting?



I believe there was overfitting done. Figure 1 has crazy red regions that are completely disjoint from the red points. For example there is a circle in the decision boundary.

Regularization

Give a plot of the decision boundary for the resulting weights from retularized linear regression with $\lambda = 2$. Do you think there is overfitting or underfitting?



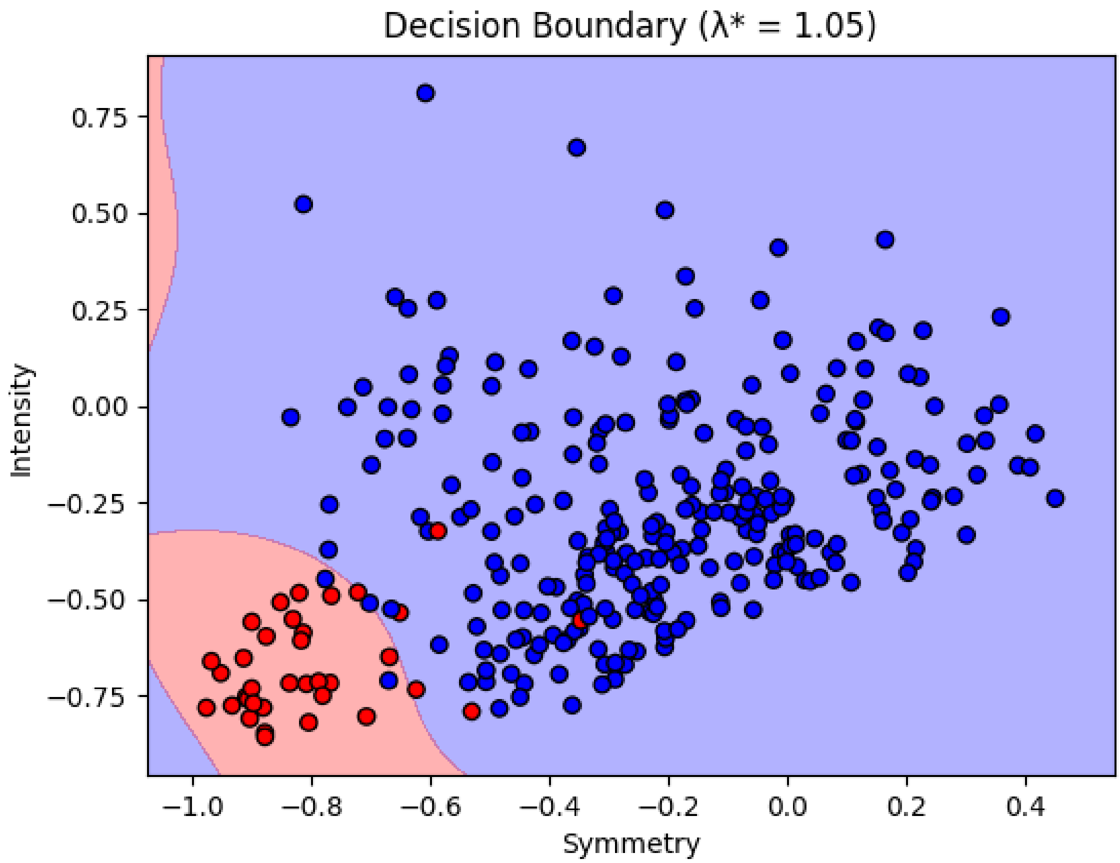
I think there might have been some underfitting, however the model is looking much better now with the regularization. Looking at the decision boundary, there is some misclassification however I believe an attempt to add more red in the correct boundary who result is gross over fitting or more misclassification. Tightening of the right side of the circle would do well to reduce the misclassification of the blue without harming the red.

Cross Validation.

Both E_{cv} and $E_{test}(w_{reg}(\lambda))$ are very high. As λ increases they fall dramatically with a strong positive acceleration. They seem to track well. I at the minimum of λ^* the two are very close. After this the values very gradually rise as the data underfits.

Pick λ

Use the cross validation error to pick the best value of λ, call it λ^*. Give a plot of the decision boundary for the weights $w_{reg}(\lambda^*$.



Estimate the Classification Error $E_{out}(w_{reg}(\lambda^*))$.

$$E_{test}(w_{reg}(\lambda^*)) = 0.030451211380306736$$

$$E_{out}(w_{reg}(\lambda^*)) \leq E_{test}(w_{reg}(\lambda^*)) + \sqrt{\frac{1}{2(N)} ln(\frac{2M}{\delta})}$$

$$E_{out}(w_{reg}(\lambda^*)) \leq E_{test}(w_{reg}(\lambda^*)) + \sqrt{\frac{1}{2(8998)} ln(\frac{2(1)}{0.01})}$$

$$E_{out}(w_{reg}(\lambda^*)) \leq 0.03045121 + 0.01715857$$

$$E_{out}(w_{reg}(\lambda^*)) \leq 0.03045121 + 0.01715857$$

$$E_{out}(w_{reg}(\lambda^*)) \leq 0.04760978$$

Is E_{cv} Bias?

Is $E_{cv}(\lambda^*)$ an unbiased estimate of $E_{test}(w_{reg}(\lambda^*))$? Explain. (Treat both as regression errors.)

No, $E_{cv}(\lambda^*)$ is not an unbiased estimate of $E_{test}(w_{reg}(\lambda^*))$. Data independence is not enough. The difference in how the errors are computed produces a bias that is not removed by random sampling or disjoint datasets. Both errors are based on models trained with different amounts of data and evaluated under different circumstances.

Data snooping.

Is $E_{test}(w_{reg}(\lambda^*))$ an unbiased estimate of $E_{out}(w_{reg}(\lambda^*))$? Explain. (Treat both as regression errors.) If it is not an unbiased estimate, how can what you did be fixed so that it is?

No, $E_{test}(w_{reg}(\lambda^*))$ is not an unbiased estimate of $E_{out}(w_{reg}(\lambda^*))$. The very data we used to create the shift and scale values was in the test data set. The test data has had influence on the training data set. There is a bias in picking the hypothesis. To fix this separate the training and the test data before normalization.