CS188, Winter 2017
Problem Set 2:
Due 2/16/2017

Shannon Phu

(a) The trend of the training data has a negative sloped curve and is fairly obvious. It looks almost like a third order polynomial. On the other hand, the relationship of the test data is not so apparent, looking like 2 clusters instead of a linear relation. Thus its harder to notice the relation between the testing data's features.

(d)

Step Size	Iterations	Coefficients	Objective Function
10^-4	10000	[1.91573585 -1.74358989]	5.49356558874
10^-3	10000	[2.4463815 -2.81630184]	3.91257640947
0.01	1487	[2.44640699 -2.81635337]	3.91257640579
0.0407	381	[2.44640705, -2.81635351]	3.91257640579

With a smaller step size the objective function's final value is larger because the small step size makes it so that gradient descent does not converge in under 10k iterations. Thus, prematurely terminating, the predicted regression line is not as close as with larger step size.

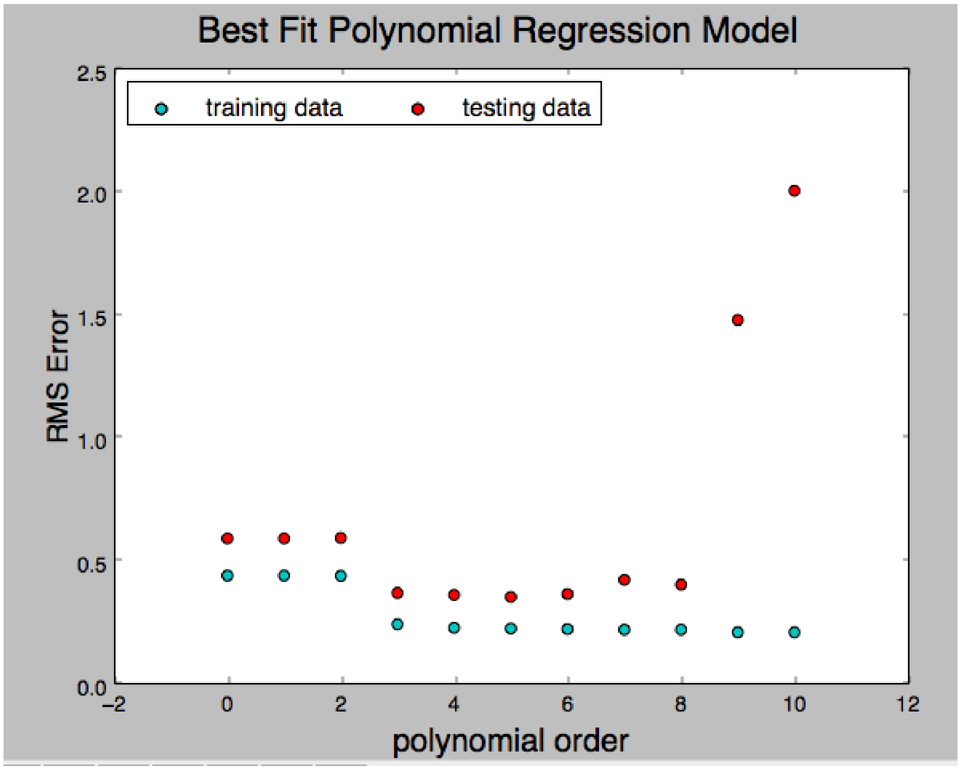
(e) The computed coefficients are: [2.44640709 -2.81635359] for theta0 and theta1 respectively. This is very close to what was computed by gradient descent when it converges. The cost is 3.91257640579. This is the same as the cost from gradient descent.

This closed form solution is faster for a smaller dataset with fewer features. But once the scale of the dataset becomes large, the computational intensity is too large to multiple and invert matrices and instead gradient descent is better for large datasets.

(f) With 100,000 iterations, gradient descent converges at 24584 iterations, with coefficients [2.4464093 -2.81635805] and final cost of 3.91257640582.

(h) RMS normalizes the error instead of computing just a magnitude that doesn't come with context.

(i) A third degree polynomial seems to fit the data the best based on the plot. This is because it seems that a minimum RMSE is reached first at polynomial order 3. There is extreme overfitting when the order of the polynomial hits 10 since the error of the test data increases dramatically.


Best Fit Polynomial Regression Model
training data
testing data
RMS Error
polynomial order