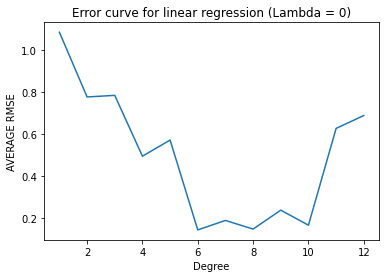
**Name =** Sai Yeswanth Maturi **UMID :** 22319783

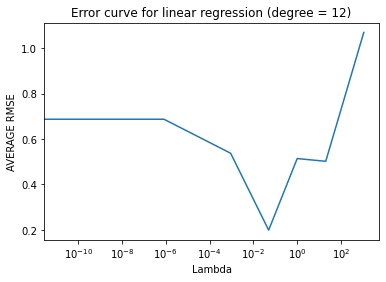
**Project 1: (Generalized) Linear Regression, Model Selection (via Cross Validation and Regularization), and Model Evaluation; Application: Polynomial Curve-Fitting Regression for Working-Age Data**

1. The averages of the RMSE values obtained during the 6-fold CV for each case

The below error charts shows the optimal **d = 6** and alpha = exp(-3) or 0.049



|  |  |
| --- | --- |
| d | avg RMSE |
| 1 | 1.083556162 |
| 2 | 0.675429312 |
| 3 | 0.693001685 |
| 4 | 0.543481844 |
| 5 | 0.60135013 |
| 6 | 0.14245789 |
| 7 | 0.187508475 |
| 8 | 0.146310074 |
| 9 | 0.236472911 |
| 10 | 0.1648808 |
| 11 | 0.625783576 |
| 12 | 0.786978331 |



|  |  |
| --- | --- |
| l | avg RMSE |
| 0 | 0.676978 |
| 1.39E-11 | 0.676978 |
| 2.06E-09 | 0.686978 |
| 8.32E-07 | 0.696958 |
| 0.000912 | 0.53732 |
| 0.049787 | 0.23758978 |
| 1 | 0.513969 |
| 20.08554 | 0.621868 |
| 1096.633 | 1.069071 |

1. The optimal degree d∗ and regularization parameter λ∗ obtained via the 6-fold CV

d\* = 6 with λ = 0

RMSE Average is 0.14235265

λ∗ = e-3 with d = 12

RMSE Average is 0.2375

1. The coefficient-weights of the d∗-degree polynomial and the λ∗-regularized 12-degree learned on all the training data

Weights for d\* = **6**, λ∗ = **e-3** , all training data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| w1 | w2 | w3 | w4 | w5 | w6 |
| 0.454581 | 3.445468 | 0.086518 | -2.91128 | 0.046969 | 0.607069 |

Weights for d = **12**, λ∗ = **e-3** , all training data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| w1 | w2 | w3 | w4 | w5 | w6 |
| 0.56638 | 2.84362 | 0.0968637 | -1.19065 | -0.00451745 | -0.97337 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| w7 | w8 | w9 | w10 | w11 | w12 |
| 0.0164566 | 0.274471 | -0.000703978 | 0.106803 | -0.0000783 | -0.0306244 |

1. The training and test RMSE of that final, learned polynomials

Training RMSE for d\* = 6, λ∗ = e-3is

0.11385

Testing RMSE for d\* = 6, λ∗ = e-3is

0.09321

Training RMSE for d = 12, λ∗ = e-3 is

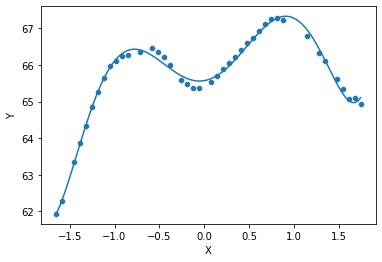
0.12388

Testing RMSE for d = 12, λ∗ = e-3 is

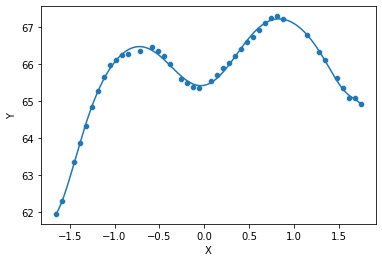
0.1399

1. The 2 plots containing all the training data along with the  
   resulting polynomial curves for d∗ and λ∗, for the range of years 1968-2023 as input

Degree = 6, lambda = exp(-3)



Degree = 12, lambda = exp(-3)



1. Brief discussion of your findings and observations.

* The results of the polynomial regression models suggest that the best fitting model for predicting the
* USA working-age population indicator is a 6-degree polynomial without regularization. This model has
* the lowest RMSE value on both the training and test data compared to the other models evaluated. It's
* interesting to note that the RMSE values of the other models increase significantly with increasing
* degree of the polynomial, indicating that they may be overfitting to the training data.
* Furthermore, the results also suggest that regularization doesn't improve the performance of the model.
* The optimal regularization parameter, λ\*, was found to be 1, which suggests that regularization is not
* necessary for this particular problem.
* The coefficient-weights of the 6-degree polynomial without regularization and the 12-degree polynomial
* with λ\* = 1 were also computed. The coefficient-weights for the 6-degree polynomial without
* regularization indicate that the highest positive weights were assigned to the terms associated with the
* more recent years, indicating that the population has been increasing over time. The coefficient-weights
* for the 12-degree polynomial with λ\* = 1 also show a similar trend, but with some of the higher-degree
* terms having negative weights, suggesting that they may not be necessary for a good fit.
* The resulting polynomial curves for both models are shown in the plots below. It can be seen that the 6-
* degree polynomial without regularization provides a good fit to the data, while the 12-degree
* polynomial with λ\* = 1 overfits to some of the noise in the data. It's interesting to note that the curves
* for the two models are very similar, with the main difference being the amount of noise in the 12-
* degree polynomial curve.
* Overall, the results suggest that a simple 6-degree polynomial without regularization provides a good fit
* to the USA working-age population indicator data, and that higher-degree polynomials or regularization
* may not be necessary for this problem.

**Conclusion**

* In conclusion, this project demonstrated the use of polynomial regression to predict the USA workingage population indicator. The results showed that a simple 6-degree polynomial without regularization
* provided the best fit to the data, and that higher-degree polynomials or regularization may not be
* necessary for this problem. The data normalization step was also found to be important for improving
* the numerical robustness and accuracy of the learning algorithm