THE UNIVERSITY OF TEXAS AT DALLAS

An Report
On

Machine Learning : Linear Regression
Assignment - 1

By

Soumyadeep Choudhury
Net ID :- sxc180056
UTD ID :- 2021439916

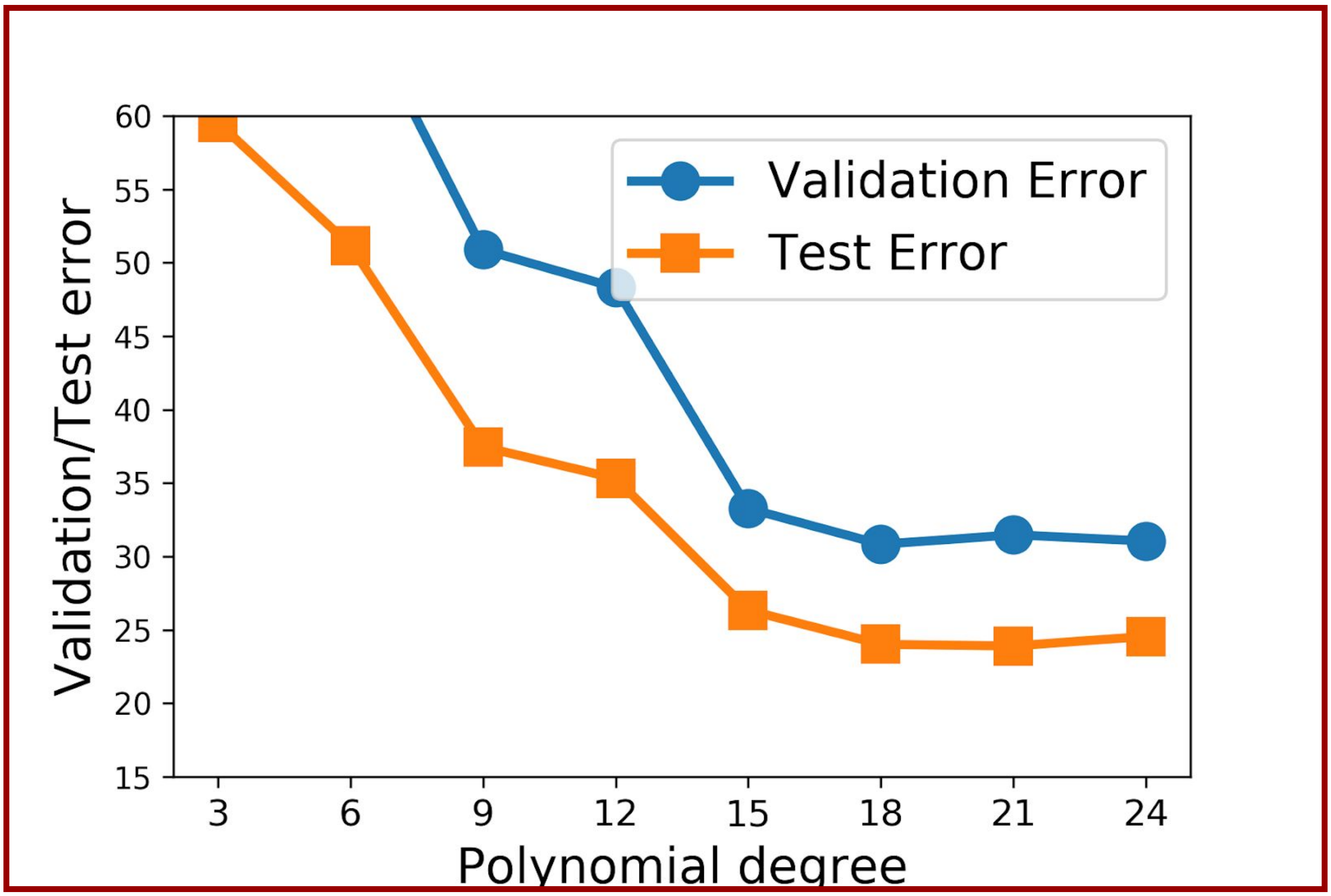
1. Regression with Polynomial Basis Functions

Output Analysis :

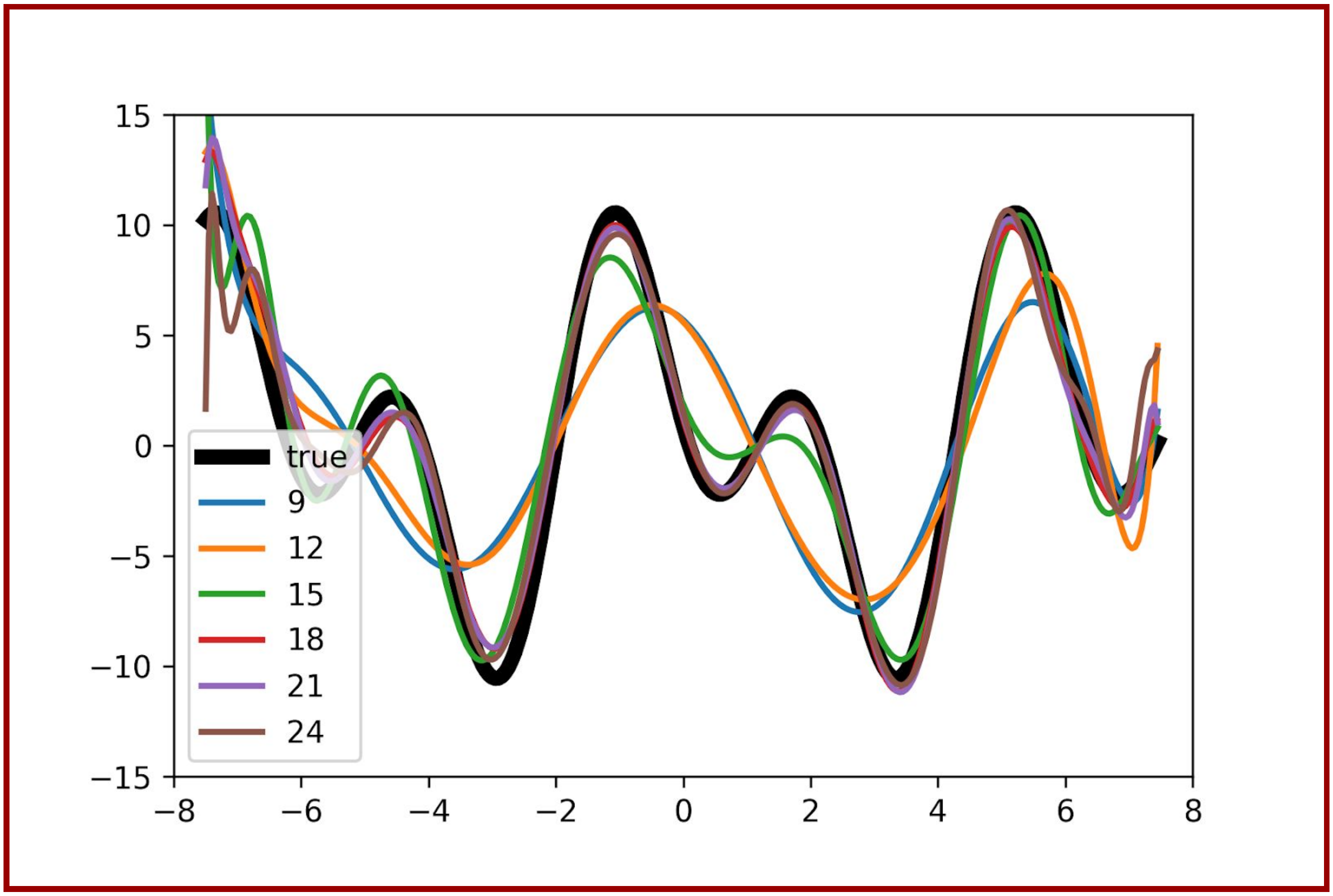
c) The result of error analysis of the polynomial basis function is shown below-

validation_error ----
{3: 86.28589095614353, 6: 68.65325584842675, 9: 50.90026583416794, 12: 48.31049770182963, 15: 33.263098007085, 18: 30.82772081782816, 21: 31.472847974615448, 24: 31.052752812007096}
test_error ----
{3: 59.569298390085244, 6: 51.18127497253844, 9: 37.466771498158536, 12: 35.3163808793748, 15: 26.341871406781042, 18: 24.018566084666176, 21: 23.906907366457396, 24: 24.557809747931902}

From the error graph or data, we can analyse that the minimum error is achieved at the polynomial degree of 18. The minimum error is not only measured by test error but also validation error. Lower the validation error, more trained is the network model due to drop in invariant bias in the weight model. Hence, I choose the value of d=18 for my best fit model.



d) From the graph, we can analyse that with the increase in the depth of polynomial degree, the data-points tries to converge more towards the true function due to the in collinearity in the data, ensuring better accuracy and low validation error. However, the model gets saturated at a certain degree of the polynomial ensuring linear -covariance relationship among the data-points, which brings more generalization in the trained model.



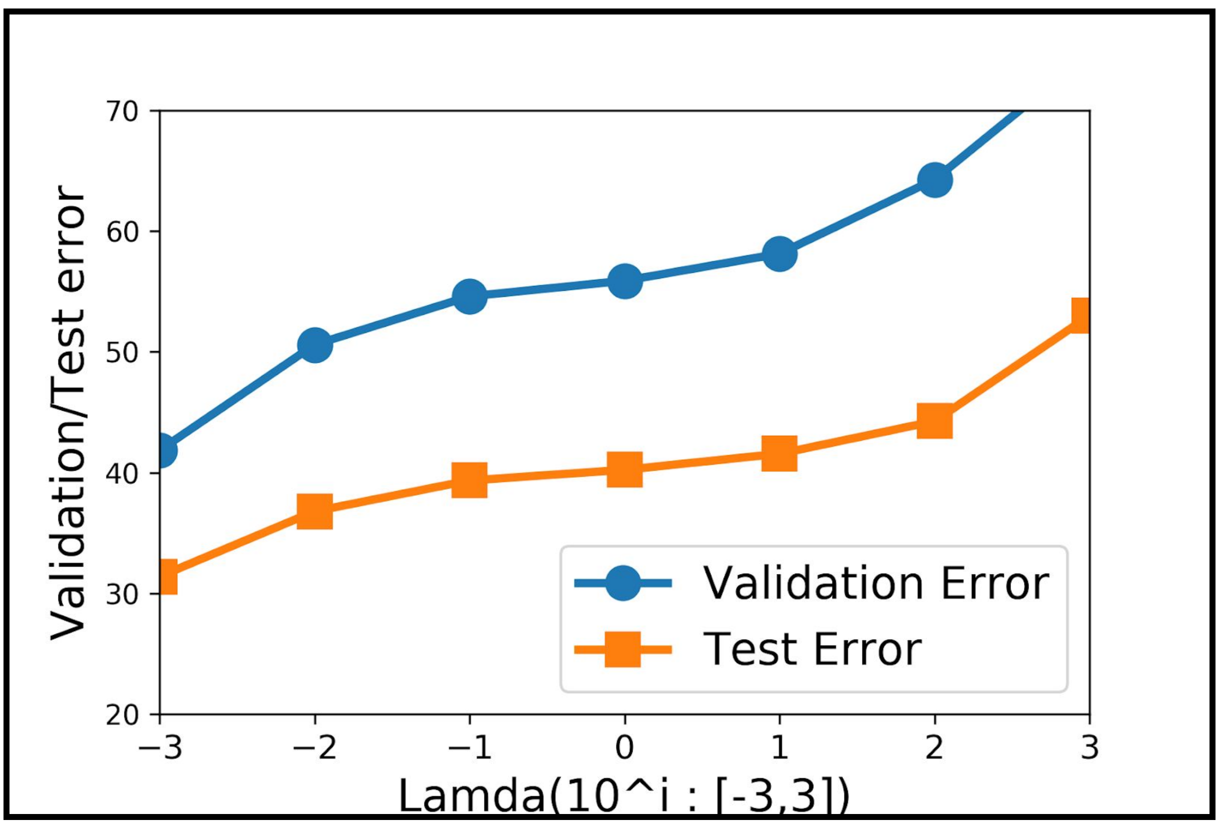
2. Regression with Radial Basis Functions

Output Analysis :

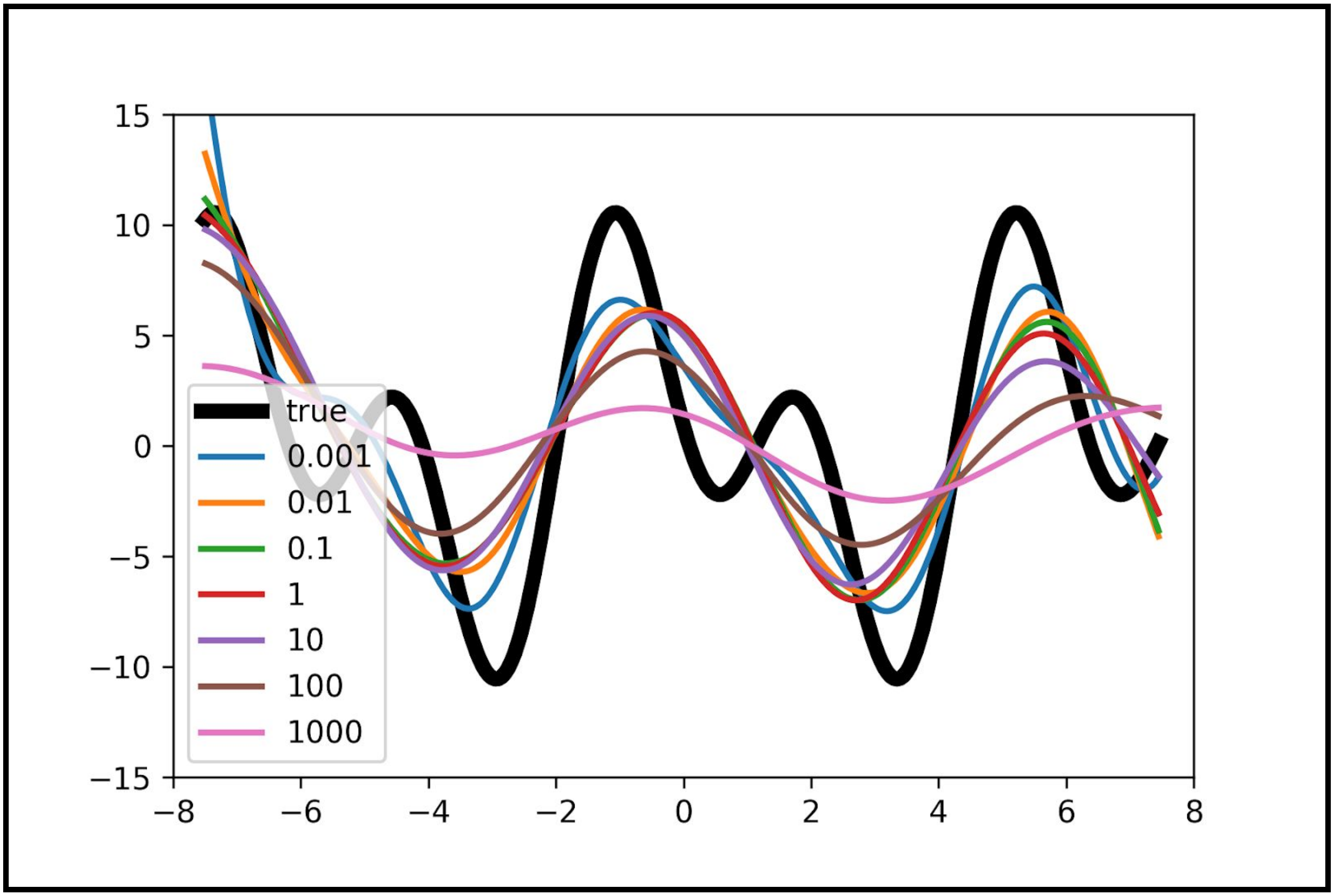
c) The error analysis results of the radial basis function is given below -

range of lambda --- [0.001, 0.01, 0.1, 1, 10, 100, 1000]
validation_error ----
{0.001: 41.83726957531996, 0.01: 50.59224116600389, 0.1: 54.60080756552955, 1: 55.89188647274824, 10: 58.135971437181354, 100: 64.26764258848266, 1000: 74.71346689715566}
test_error ----
{0.001: 31.391334356204013, 0.01: 36.79022846133964, 0.1: 39.35056826463211, 1: 40.23248184622335, 10: 41.563304037500835, 100: 44.25522059974583, 1000: 53.00567219289972}

From the error graph and result, we can analyse that the minimum validation error and test error are achieved with lambda (lm) = 0.001. It is observed that the error tends to increase with the increase in lamda value as high value of lambda adds more impurity in data-points by increasing the normalized points data-space which inherently reduces the collinear relationship between the points resulting in poor model. This may also lead to a case of underfitting due to additional bias effects.



d) From this graph, we can observe the convergence of data-points in a radial-basis function with varying lambda parameters. It is also highlighted that the regression line with lambda = 0.0001 converges most closely to the true function ensured a low error and a better trained model.



------------------------------------- END --------------------------------------