CS612: Statistical Pattern Recognition Laboratory
Programming Assignment 3

Sudip Kumar Mandal
CS24MT013

Aakash Harsh
CS24MT016

Subhradip Palit
CS24MT009

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DHARWAD
November, 2024

Contents

List of Tables

List of Figures

Chapter 1

Dataset 1: 1-dimensional (Univariate) input data

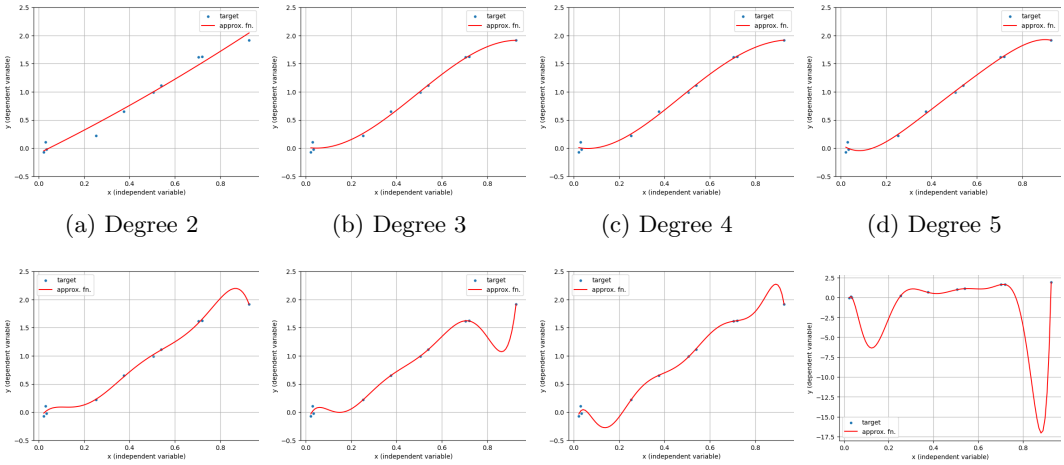
1.1 Dataset description

- Total no. of data points: 1001
- Test data size: 700
- Train data size: 301
- Dimension of independent variable: 1
- Dimension of dependent variable: 1

1.2 Result

1.2.1 Train data size 10

1.2.1.1 Plot of approximated functions



(a) Degree 2 (b) Degree 3 (c) Degree 4 (d) Degree 5

(e) Degree 6 (f) Degree 7 (g) Degree 8 (h) Degree 9

Figure 1.1: Approximate functions for degrees 2 to 9 with training data of size 10, is shown respectively from (a) to (h)

Figure 1.1, shows the plots of approximate functions for degrees 2 to 9 with training data size 10.

1.2.1.2 Train and Test MSE

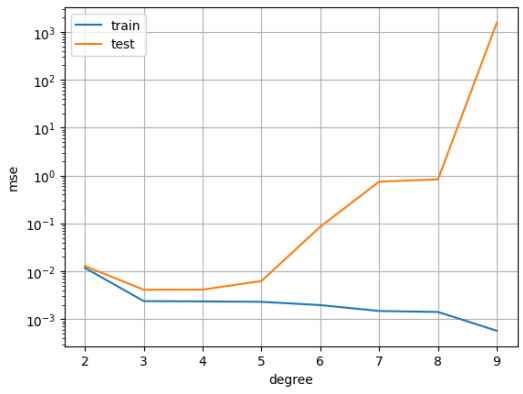


Figure 1.2: MSE for train and test data vs degree of function on training size 10

The plot in Figure 1.2 shows the MSE values for both training and test datasets across polynomial degrees 2 to 9.

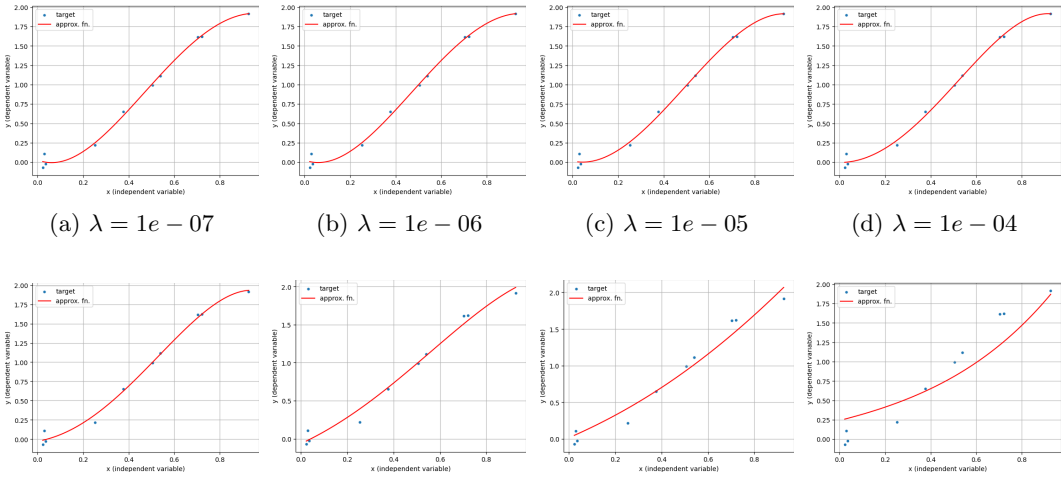
1.2.1.3 Summary of Model Weights and MSE and determining degrees with over-fitting

Table 1.1: Weights, train and test MSEs for degrees 2 to 9

Degree	Weights (w^T)	Train MSE	Test MSE
2	[-0.0937, 1.99, 0.347]	0.0118	0.0129
3	[0.017, -0.748, 8.18, -5.57]	0.00238	0.00412
4	[0.0288, -1.19, 10.3, -9.08, 1.81]	0.00235	0.00414
5	[0.0775, -3.26, 25.3, -49.6, 48.7, -19.4]	0.0023	0.00627
6	[-0.144, 7.09, -78.5, 384, -823, 816, -305]	0.00197	0.0854
7	[-0.291, 16, -238, 1.5e+03, -4.64e+03, 7.51e+03, -6.12e+03, 1.98e+03]	0.00148	0.746
8	[-0.456, 29.7, -592, 4.73e+03, -1.88e+04, 4.13e+04, -5.11e+04, 3.32e+04, -8.85e+03]	0.00141	0.832
9	[-5.01, 426, -1.15e+04, 1.17e+05, -6.06e+05, 1.8e+06, -3.21e+06, 3.4e+06, -1.97e+06, 4.83e+05]	0.000572	1.55e+03

Table 1.1 shows the weights, train and test MSEs for degrees 2 to 9. From this table and from the figure 1.2, we see that the test MSE starts increasing from degree 4 onward, indicating over-fitting. Hence regularization is applied on degrees 4 to 9.

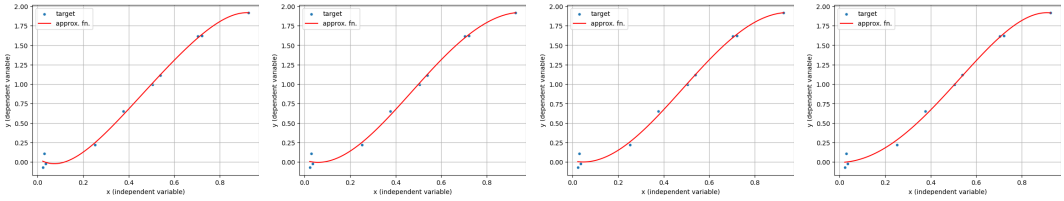
1.2.1.4 Plot of approximated functions with regularization



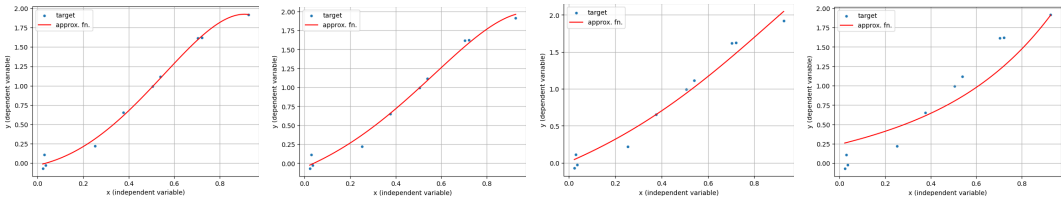
(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.3: Approximation functions for varying regularization parameter (λ) values for degree 4

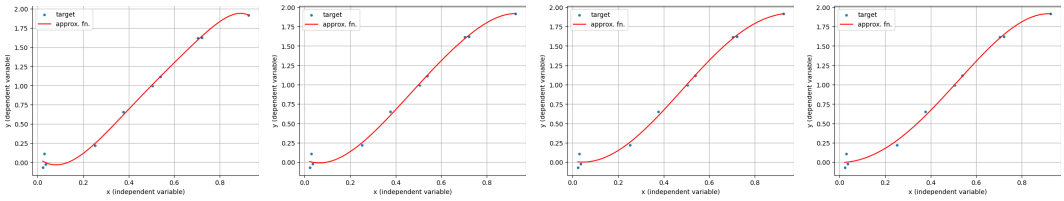


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

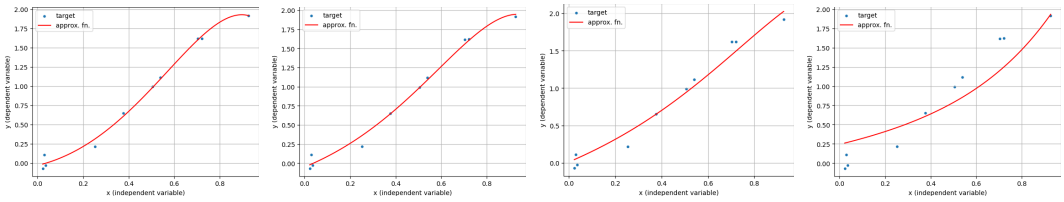


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.4: Approximation functions for varying regularization parameter (λ) values for degree 5

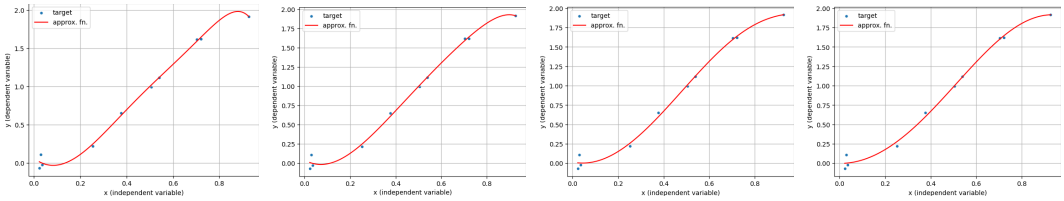


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

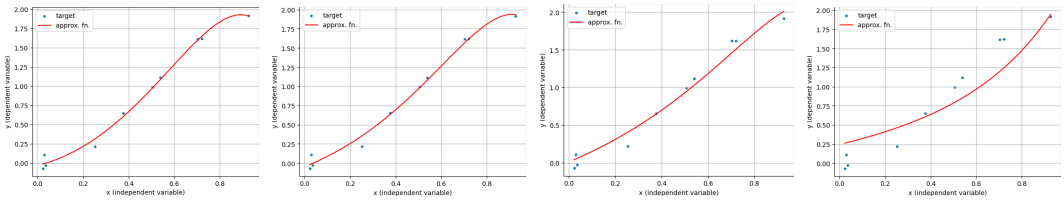


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.5: Approximation functions for varying regularization parameter (λ) values for degree 6

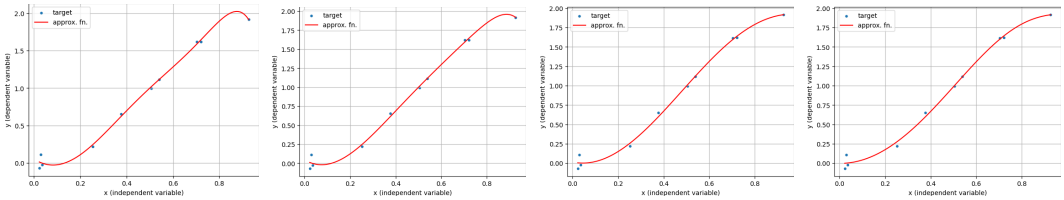


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

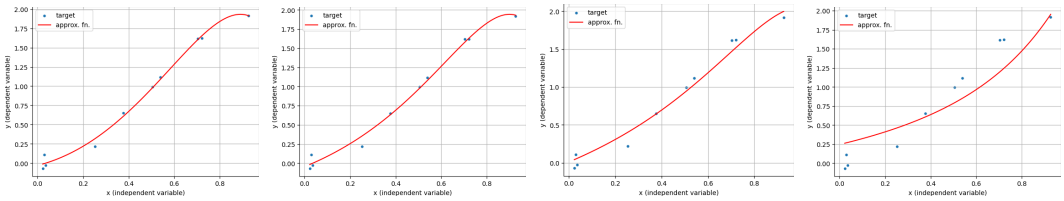


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.6: Approximation functions for varying regularization parameter (λ) values for degree 7

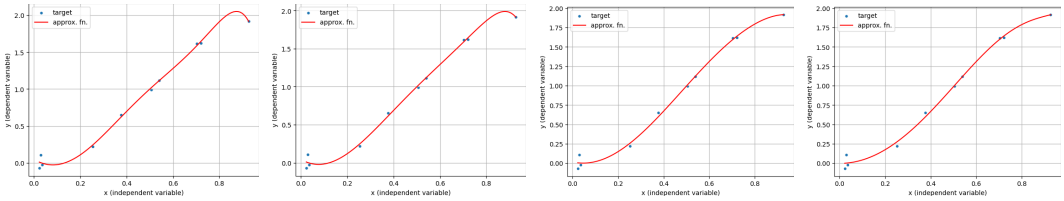


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

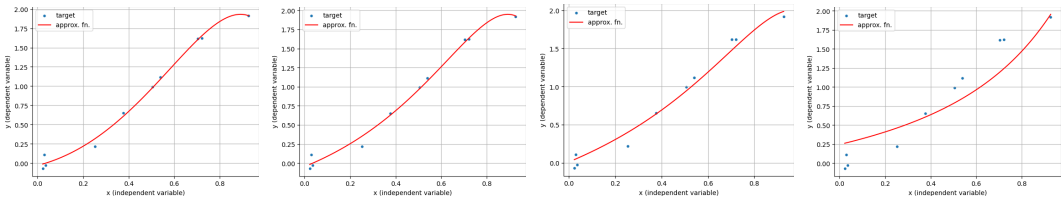


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.7: Approximation functions for varying regularization parameter (λ) values for degree 8



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

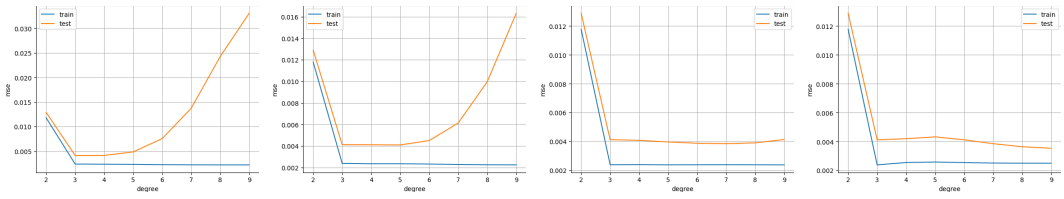


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

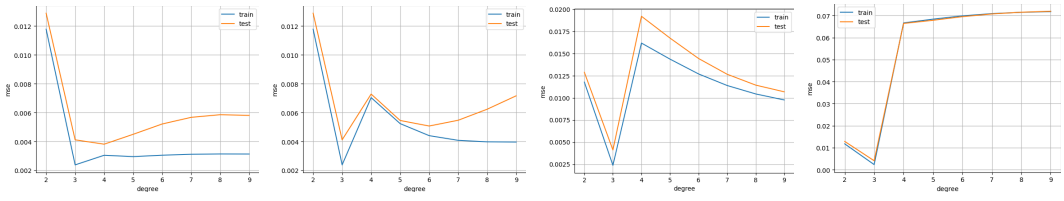
Figure 1.8: Approximation functions for varying regularization parameter (λ) values for degree 9

Figure 1.3 to 1.8 shows the Approximation functions for varying regularization parameter (λ) values respectively for degrees 4 to 9.

1.2.1.5 Train and Test MSEs with regularization



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$



(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.9: Train and test MSE vs degree of function for varying regularization parameter (λ) values

Figure 1.9 shows the plots of train and test MSE vs degree of function for varying regularization parameter (λ) values.

1.2.1.6 Summary of weights, train, test MSE with regularization

Table 1.2: Weights, Train and Test MSEs for Degrees 4 to 9 with Varying regularization parameter (λ)

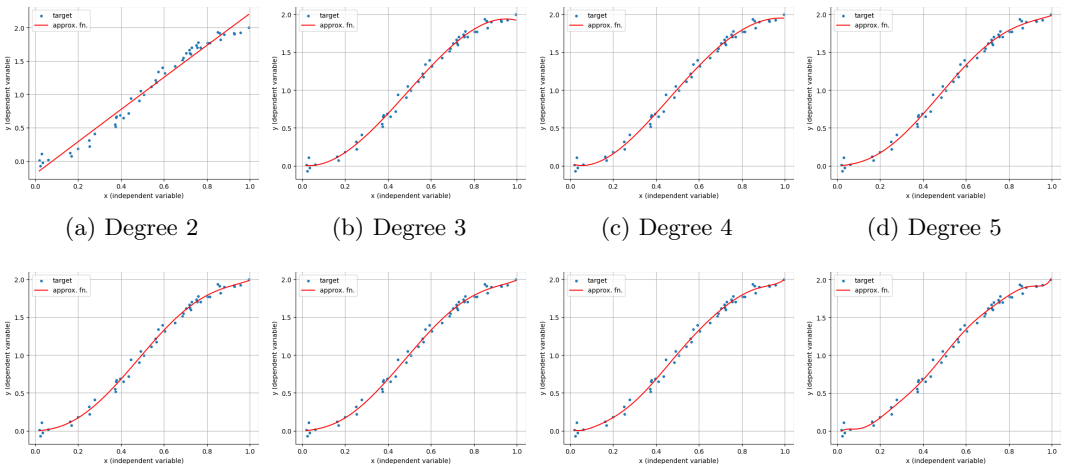
Degree	λ	Weights (w^T)	Train MSE	Test MSE
4	1e-07	[0.0285, -1.18, 10.3, -9.02, 1.78]	0.00235	0.00414
	1e-06	[0.0265, -1.1, 9.95, -8.45, 1.49]	0.00235	0.00411
	1e-05	[0.0143, -0.665, 7.82, -5.07, -0.23]	0.00239	0.00407
	0.0001	[-0.0049, -0.0143, 4.82, -0.428, -2.53]	0.00254	0.00419
	0.001	[-0.0286, 0.524, 3.36, 0.521, -2.47]	0.00305	0.00381
	0.01	[-0.0641, 1.42, 1.49, 0.223, -0.981]	0.00704	0.00729
	0.1	[0.0165, 1.37, 0.775, 0.248, -0.112]	0.0162	0.0192
	1.0	[0.241, 0.735, 0.532, 0.372, 0.263]	0.0666	0.0664
5	1e-07	[0.05, -2.09, 16.9, -27.1, 22.8, -8.73]	0.00232	0.00485
	1e-06	[0.0258, -1.07, 9.69, -7.72, 0.614, 0.374]	0.00235	0.00409
	1e-05	[0.0154, -0.681, 7.5, -3.17, -3.39, 1.63]	0.00237	0.00396
	0.0001	[-0.00651, 0.0424, 4.63, -0.465, -1.95, -0.421]	0.00258	0.00432
	0.001	[-0.0285, 0.64, 2.67, 0.881, -0.72, -1.62]	0.00295	0.0045
	0.01	[-0.0565, 1.29, 1.49, 0.566, -0.33, -0.974]	0.00524	0.00546
	0.1	[0.0135, 1.35, 0.814, 0.321, -0.0198, -0.238]	0.0144	0.0168
	1.0	[0.242, 0.721, 0.509, 0.347, 0.237, 0.163]	0.0683	0.0678
6	1e-07	[0.0621, -2.55, 18.3, -21, -8.27, 34.8, -19.7]	0.00226	0.00755
	1e-06	[0.0329, -1.36, 11.4, -10.4, -1.38, 7.78, -4.23]	0.00233	0.0045
	1e-05	[0.0141, -0.618, 7.04, -2.22, -3.19, -0.409, 1.31]	0.00238	0.00387
	0.0001	[-0.00512, 0.00183, 4.68, -0.234, -2.08, -1.18, 0.665]	0.00254	0.00412
	0.001	[-0.03, 0.722, 2.39, 0.824, -0.357, -0.846, -0.927]	0.00305	0.00521
	0.01	[-0.0514, 1.24, 1.42, 0.674, -0.0274, -0.52, -0.831]	0.00441	0.00507
	0.1	[0.0117, 1.33, 0.836, 0.376, 0.0555, -0.151, -0.278]	0.0127	0.0145
	1.0	[0.244, 0.715, 0.499, 0.334, 0.224, 0.15, 0.101]	0.0698	0.0695
7	1e-07	[0.0573, -2.29, 15.1, -7.54, -25.6, 21.7, 23.4, -23.4]	0.00222	0.0137
	1e-06	[0.0403, -1.66, 12.8, -10.9, -5.37, 7.84, 6.65, -7.68]	0.00228	0.00613
	1e-05	[0.0136, -0.598, 6.92, -2.08, -3.05, -0.519, 0.857, 0.388]	0.00238	0.00384
	0.0001	[-0.00427, -0.014, 4.61, 0.0634, -1.94, -1.55, -0.29, 1.02]	0.00251	0.00385
	0.001	[-0.0309, 0.756, 2.3, 0.766, -0.302, -0.662, -0.632, -0.45]	0.00312	0.00568
	0.01	[-0.0485, 1.22, 1.34, 0.693, 0.107, -0.29, -0.533, -0.67]	0.00408	0.00547
	0.1	[0.0108, 1.31, 0.847, 0.413, 0.11, -0.0862, -0.208, -0.28]	0.0114	0.0127
	1.0	[0.244, 0.713, 0.494, 0.328, 0.217, 0.144, 0.0944, 0.0614]	0.0709	0.0707
8	1e-07	[0.0499, -1.96, 12.2, 0.711, -30, 10.3, 24.3, 5.03, -19.5]	0.00221	0.0242
	1e-06	[0.0439, -1.78, 13.1, -9.64, -7.98, 5.16, 8.89, 2.33, -8.55]	0.00226	0.00996
	1e-05	[0.0141, -0.618, 7.02, -2.17, -3.19, -0.541, 1.06, 0.826, -0.482]	0.00238	0.0039
	0.0001	[-0.00409, -0.00827, 4.5, 0.277, -1.74, -1.63, -0.726, 0.242, 1.03]	0.0025	0.00364
	0.001	[-0.0312, 0.765, 2.28, 0.747, -0.3, -0.633, -0.576, -0.373, -0.147]	0.00314	0.00586
	0.01	[-0.0468, 1.22, 1.29, 0.684, 0.165, -0.172, -0.371, -0.478, -0.531]	0.00398	0.00624
	0.1	[0.0105, 1.29, 0.852, 0.437, 0.148, -0.0392, -0.156, -0.226, -0.265]	0.0104	0.0114
	1.0	[0.245, 0.712, 0.492, 0.325, 0.214, 0.14, 0.0911, 0.0582, 0.0359]	0.0714	0.0715
9	1e-07	[0.0465, -1.82, 11.1, 3.18, -30, 6.83, 21.7, 7.82, -8.68, -9.29]	0.00221	0.0331
	1e-06	[0.0446, -1.8, 12.8, -8.23, -9.08, 2.88, 8.25, 5.28, -1.33, -7.56]	0.00225	0.0163
	1e-05	[0.015, -0.654, 7.18, -2.25, -3.43, -0.678, 1.18, 1.28, 0.277, -1.05]	0.00237	0.00413
	0.0001	[-0.0042, 0.00478, 4.39, 0.409, -1.57, -1.6, -0.903, -0.14, 0.479, 0.927]	0.0025	0.00353
	0.001	[-0.0311, 0.764, 2.28, 0.751, -0.299, -0.636, -0.583, -0.384, -0.162, 0.0338]	0.00313	0.00581
	0.01	[-0.0459, 1.22, 1.25, 0.668, 0.19, -0.11, -0.279, -0.368, -0.408, -0.42]	0.00396	0.00717
	0.1	[0.0106, 1.28, 0.852, 0.452, 0.174, -0.00533, -0.118, -0.185, -0.224, -0.244]	0.00977	0.0107
	1.0	[0.245, 0.711, 0.491, 0.324, 0.213, 0.139, 0.0896, 0.0567, 0.0345, 0.0195]	0.0718	0.0719

Table 1.2 shows the weights, train and test MSE for different regularization parameter (λ) and for degrees 4 to 9.

From the plots of train and test MSE vs degree of function in figure 1.9, we see that introducing a regularization parameter smoothens the graph, reducing the test MSE. We see a general trend that for (λ) values 1e-07 and 1e-06, there isn't much effect, but for 1e-05, 1e-04, the problem of over-fitting decreases, but from 1e-03 onwards, the graph smoothens a lot and the MSE actually increases.

1.2.2 Train data size 50

1.2.2.1 Plot of approximated functions



(a) Degree 2 (b) Degree 3 (c) Degree 4 (d) Degree 5

(e) Degree 6 (f) Degree 7 (g) Degree 8 (h) Degree 9

Figure 1.10: Approximate functions for degrees 2 to 9 with training data size 50, is shown respectively from (a) to (h)

Figure 1.10, shows the plots of approximate functions for degrees 2 to 9 with training data size 50.

1.2.2.2 Train and Test MSE

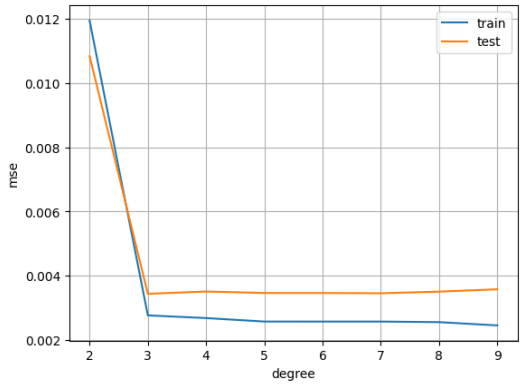


Figure 1.11: MSE for train and test data vs degree of function on training size 50

The plot in Figure 1.11 shows the MSE values for both training and test datasets across polynomial degrees 2 to 9.

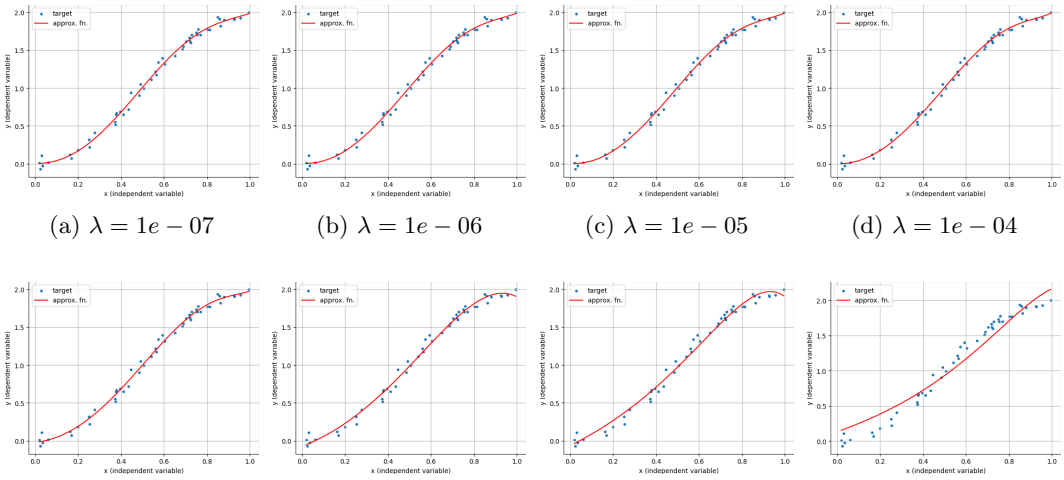
1.2.2.3 Summary of Model Weights and MSE and determining degrees with over-fitting

Table 1.3: Weights, train and test MSEs for degrees 2 to 9

Degree	Weights ($\mathbf{w}^T$)	Train MSE	Test MSE
2	[-0.195, 2.43, -0.0281]	0.0119	0.0108
3	[0.00846, -0.438, 7.35, -5]	0.00276	0.00343
4	[0.0312, -1.02, 9.99, -9.06, 2.01]	0.00268	0.00351
5	[-0.00234, 0.202, 1.52, 13.1, -22.5, 9.68]	0.00257	0.00346
6	[0.000391, 0.082, 2.64, 8.89, -15.1, 3.41, 2.02]	0.00257	0.00346
7	[-0.00444, 0.322, -0.196, 23.2, -51.6, 53.3, -32.5, 9.54]	0.00257	0.00345
8	[0.0304, -1.65, 30.3, -184, 689, -1.43e+03, 1.66e+03, -1e+03, 249]	0.00255	0.0035
9	[-0.0759, 5.33, -108, 1.02e+03, -4.86e+03, 1.34e+04, -2.21e+04, 2.16e+04, -1.14e+04, 2.53e+03]	0.00245	0.00357

Table 1.3 shows the weights, train and test MSEs for degrees 2 to 9. From this table and from the figure 1.11, we see that the test MSE starts increasing from degree 8 onward, indicating over-fitting. Hence regularization is applied on degrees 8 and 9.

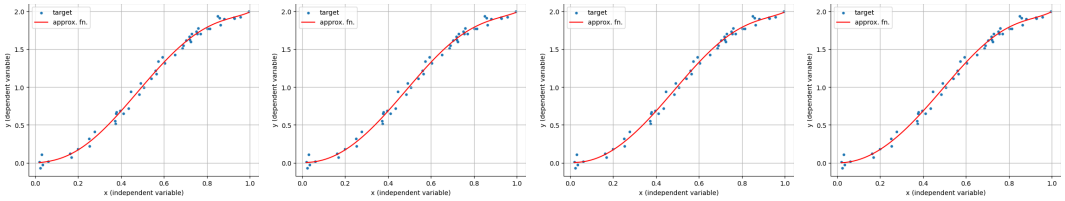
1.2.2.4 Plot of approximated functions with regularization



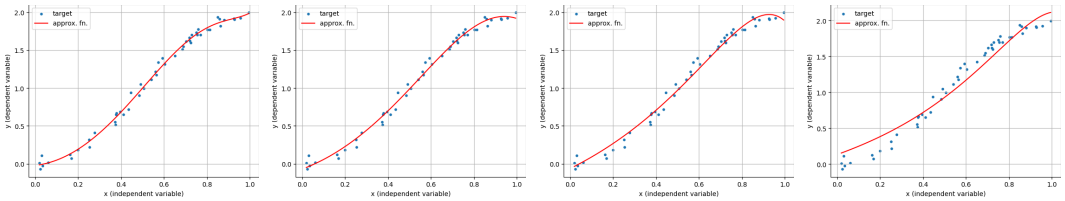
(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.12: Approximation functions for varying regularization parameter (λ) values for degree 8



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

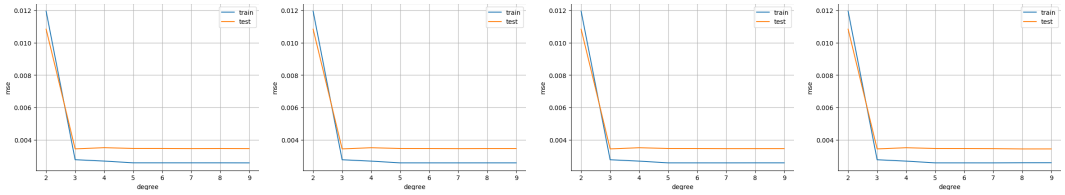


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

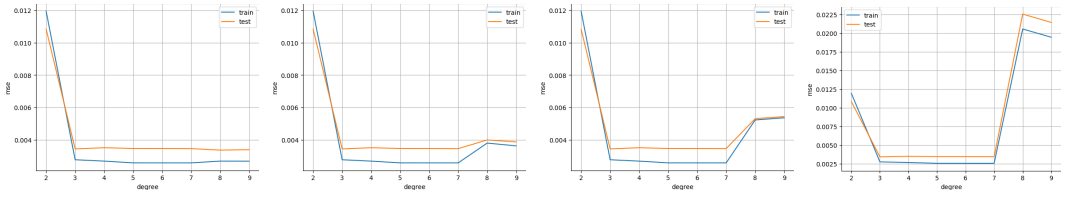
Figure 1.13: Approximation functions for varying regularization parameter (λ) values for degree 9

Figure 1.12 to 1.13 shows the Approximation functions for varying regularization parameter (λ) values respectively for degrees 8 and 9.

1.2.2.5 Train and Test MSEs with regularization



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$



(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.14: Train and test MSE for varying regularization parameter (λ) values, dataset 1

1.2.2.6 Summary of weights, train, test MSE with regularization

Table 1.4: Weights, Train and Test MSEs for Degrees 4 to 9 with Varying regularization parameter (λ)

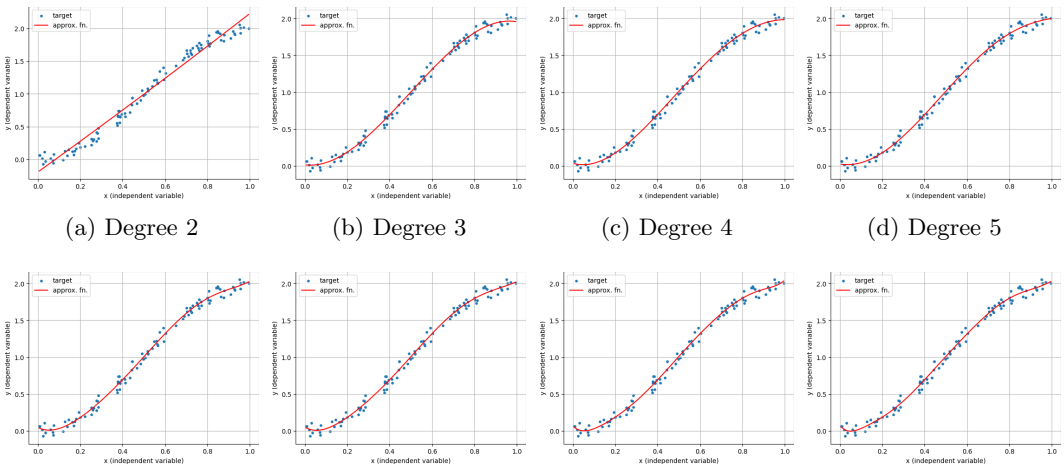
Degree	λ	Weights ($\mathbf{w}^T$)	Train MSE	Test MSE
8	1e-07	[0.00195, 0.0175, 3.19, 6.96, -11.4, -1.29, 7.17, -3.94, 1.32]	0.00257	0.00346
	1e-06	[0.00517, -0.122, 4.44, 2.56, -4.93, -3.07, 1.52, 2.07, -0.494]	0.00257	0.00346
	1e-05	[0.00721, -0.217, 5.24, 0.282, -2.93, -2.08, 0.0619, 1.15, 0.475]	0.00257	0.00345
	0.0001	[0.00124, -0.0964, 4.89, 0.194, -2.12, -1.9, -0.687, 0.492, 1.22]	0.00258	0.00343
	0.001	[-0.0216, 0.302, 3.74, 0.707, -1.36, -1.78, -1.12, 0.0697, 1.45]	0.00268	0.00336
	0.01	[-0.0728, 1.1, 1.96, 0.609, -0.393, -0.745, -0.625, -0.232, 0.297]	0.00379	0.00398
	0.1	[-0.0646, 1.42, 1.09, 0.505, 0.0783, -0.171, -0.293, -0.334, -0.326]	0.00522	0.0053
	1.0	[0.133, 1.09, 0.763, 0.434, 0.191, 0.0235, -0.0893, -0.165, -0.214]	0.0206	0.0226
9	1e-07	[-0.000729, 0.14, 2.03, 10.8, -14.9, -6.26, 13.5, 6.49, -17.5, 7.69]	0.00257	0.00345
	1e-06	[0.00519, -0.122, 4.43, 2.63, -5.05, -3.09, 1.7, 2.2, -0.919, 0.206]	0.00257	0.00346
	1e-05	[0.00702, -0.21, 5.19, 0.355, -2.88, -2.17, -0.0878, 1.15, 0.884, -0.254]	0.00257	0.00345
	0.0001	[0.000919, -0.0919, 4.9, 0.0963, -2.09, -1.75, -0.567, 0.38, 0.703, 0.409]	0.00258	0.00343
	0.001	[-0.0226, 0.347, 3.55, 0.76, -1.1, -1.54, -1.14, -0.405, 0.402, 1.15]	0.00267	0.00339
	0.01	[-0.0709, 1.08, 1.98, 0.673, -0.338, -0.748, -0.715, -0.422, 0.00407, 0.482]	0.00362	0.00387
	0.1	[-0.063, 1.42, 1.07, 0.493, 0.0833, -0.146, -0.249, -0.273, -0.25, -0.201]	0.00536	0.00543
	1.0	[0.133, 1.08, 0.763, 0.444, 0.211, 0.0511, -0.056, -0.127, -0.173, -0.203]	0.0195	0.0214

Table 1.4 shows the weights, train and test MSE for different regularization parameter (λ) and for degrees 8 and 9.

From the plots of train and test MSE vs degree of function in figure 1.14, and from the table 1.4, we see that for degree 8, the test MSE, which as 0.0035, decreased slightly with regularization parameters 1e-07 to 1e-03, but increased from 1e-02 onwards. Similar trend is observed with degree 9 function.

1.2.3 Train data size 100

1.2.3.1 Plot of approximated functions



(a) Degree 2 (b) Degree 3 (c) Degree 4 (d) Degree 5

(e) Degree 6 (f) Degree 7 (g) Degree 8 (h) Degree 9

Figure 1.15: Approximate functions for degrees 2 to 9 with training data size 100, is shown respectively from (a) to (h)

Figure 1.15, shows the plots of approximate functions for degrees 2 to 9 with training data size 100

1.2.3.2 Train and Test MSE

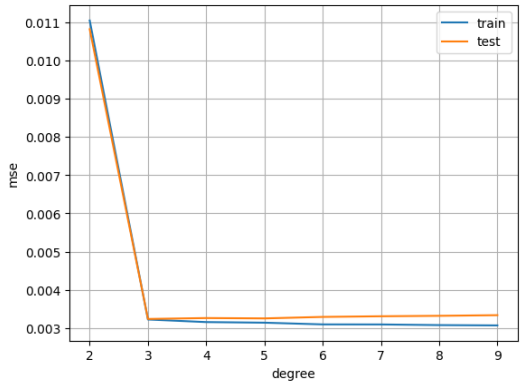


Figure 1.16: MSE for train and test data on training size 100 without regularization

The plot in Figure 1.16 shows the MSE values for both training and test datasets across polynomial degrees 2 to 9.

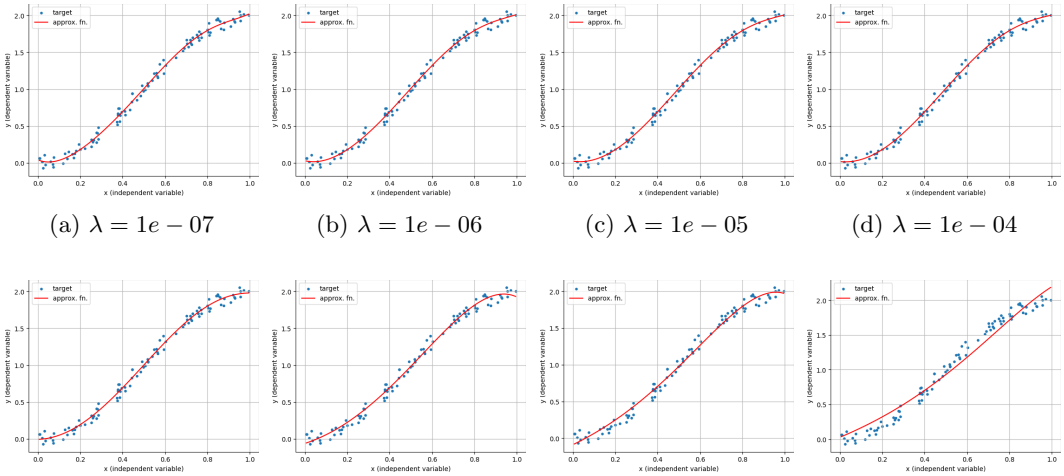
1.2.3.3 Summary of Model Weights and MSE and determining degrees with over-fitting

Table 1.5: Weights, train and test MSEs for degrees 2 to 9

Degree	Weights (w^T)	Train MSE	Test MSE
2	[-0.192, 2.33, 0.09]	0.011	0.0108
3	[0.0142, -0.341, 6.98, -4.7]	0.00323	0.00324
4	[0.0351, -0.811, 9.14, -8.08, 1.7]	0.00316	0.00327
5	[0.0243, -0.43, 6.36, -0.486, -6.98, 3.51]	0.00314	0.00326
6	[0.0455, -1.45, 16.7, -42, 70.9, -65.1, 22.9]	0.0031	0.0033
7	[0.0498, -1.71, 20.3, -61.8, 125, -143, 79.3, -16.1]	0.0031	0.00331
8	[0.0657, -2.92, 41.8, -220, 722, -1.39e+03, 1.54e+03, -912, 224]	0.00308	0.00332
9	[0.0784, -4.08, 67.3, -457, 1.87e+03, -4.56e+03, 6.79e+03, -6.02e+03, 2.92e+03, -596]	0.00307	0.00334

Table 1.5 shows the weights, train and test MSEs for degrees 2 to 9. From this table and from the figure 1.16, we see that the test MSE starts increasing from degree 6 onwards, hence there is over-fitting from degree 6 onwards. Hence regularization is applied to those cases.

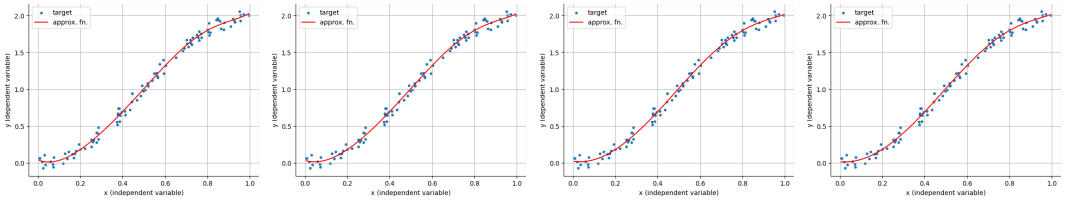
1.2.3.4 Plot of approximated functions with regularization



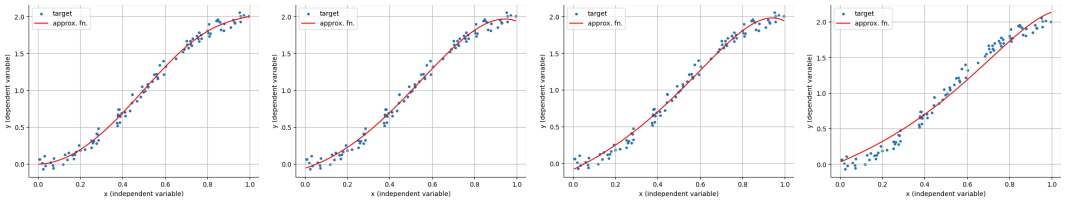
(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.17: Approximation functions for varying regularization parameter (λ) values for degree 6

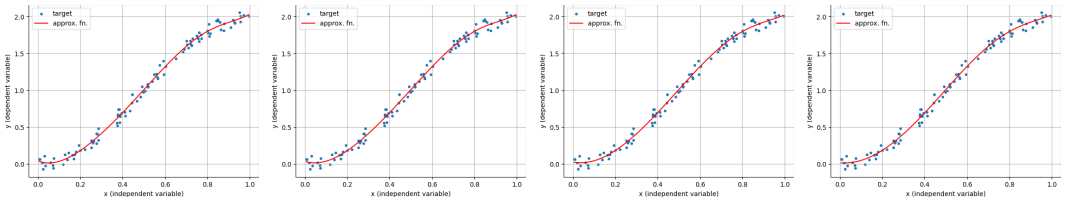


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

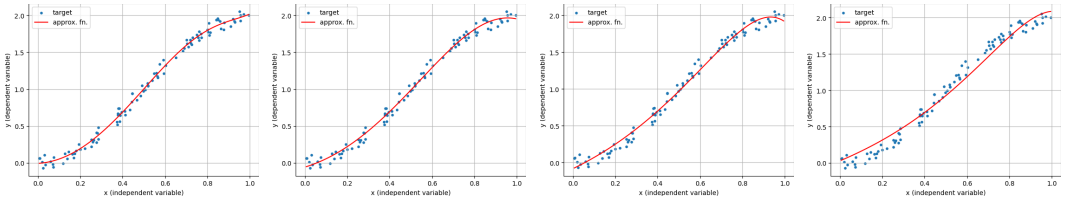


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.18: Approximation functions for varying regularization parameter (λ) values for degree 7

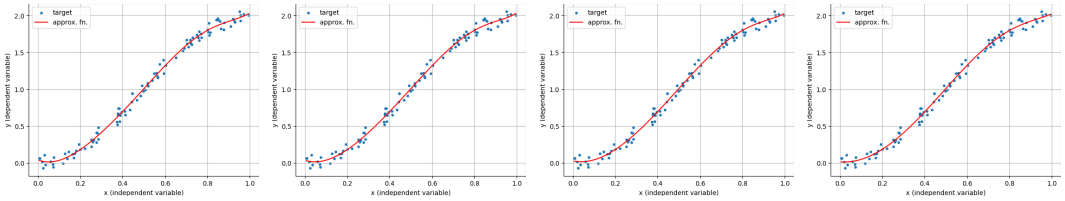


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

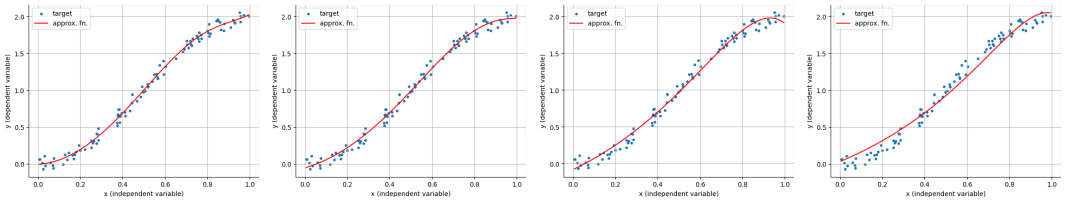


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.19: Approximation functions for varying regularization parameter (λ) values for degree 8



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

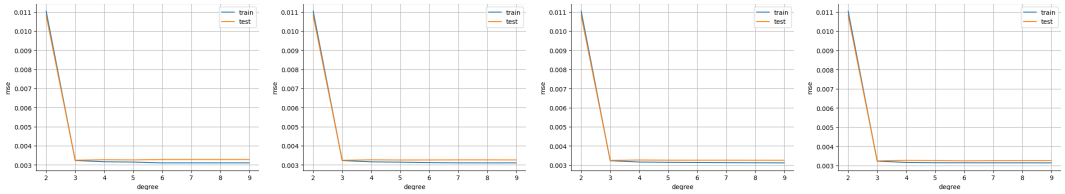


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

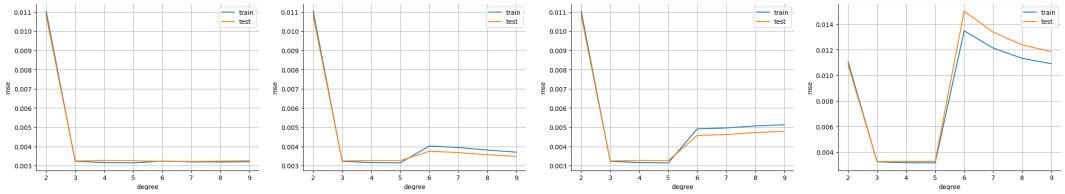
Figure 1.20: Approximation functions for varying regularization parameter (λ) values for degree 9

Figure 1.17 to 1.20 shows the Approximation functions for varying regularization parameter (λ) values respectively for degrees 6 to 9.

1.2.3.5 Train and Test MSEs with regularization



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$



(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.21: Train and test MSE vs degree of function for varying regularization parameter (λ) values

Figure 1.14 shows the plots of train and test MSE vs degree of function for varying regularization parameter (λ) values.

1.2.3.6 Summary of weights, train, test MSE with regularization

Table 1.6: Weights, Train and Test MSEs for Degrees 6 to 9 with Varying regularization parameter (λ)

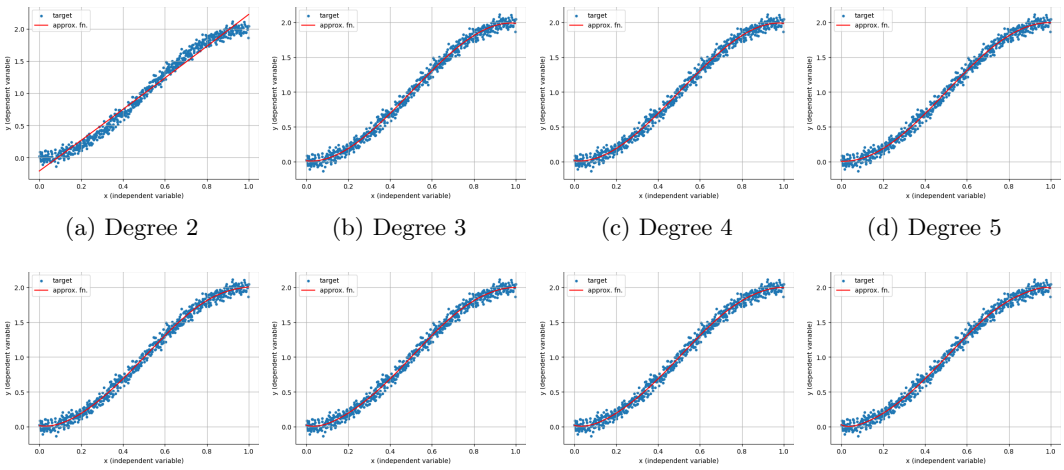
Degree	λ	Weights (w^T)	Train MSE	Test MSE
6	1e-07	[0.0402, -1.21, 14.3, -32.6, 53.7, -50.2, 18]	0.0031	0.00328
	1e-06	[0.0285, -0.663, 8.97, -11.7, 15.3, -16.9, 7.06]	0.00312	0.00326
	1e-05	[0.0228, -0.418, 6.65, -2.85, -0.822, -3.13, 2.55]	0.00314	0.00326
	0.0001	[0.0176, -0.296, 5.96, -1.15, -2.99, -1.67, 2.13]	0.00314	0.00324
	0.001	[-0.00616, 0.0879, 4.77, -0.414, -2.5, -1.43, 1.47]	0.00324	0.00322
	0.01	[-0.0614, 0.942, 2.38, 0.407, -0.715, -0.797, -0.236]	0.00403	0.00375
	0.1	[-0.0875, 1.4, 1.31, 0.54, -0.0744, -0.456, -0.663]	0.00491	0.00457
	1.0	[0.028, 1.28, 0.863, 0.405, 0.0718, -0.153, -0.302]	0.0135	0.015
7	1e-07	[0.0405, -1.18, 13.5, -26, 30.7, -11.6, -13.1, 9.69]	0.0031	0.00328
	1e-06	[0.0336, -0.882, 10.7, -15.9, 14.1, -0.266, -14.1, 8.28]	0.00311	0.00326
	1e-05	[0.0233, -0.466, 7.2, -4.81, 1.04, -0.593, -2.95, 2.57]	0.00313	0.00326
	0.0001	[0.0151, -0.233, 5.71, -1.2, -2.11, -1.4, -0.237, 1.47]	0.00314	0.00325
	0.001	[-0.0054, 0.122, 4.44, 0.0831, -1.97, -1.8, -0.41, 1.54]	0.0032	0.00322
	0.01	[-0.0598, 0.918, 2.42, 0.468, -0.716, -0.907, -0.472, 0.28]	0.00394	0.00368
	0.1	[-0.0837, 1.39, 1.25, 0.526, 0.000141, -0.291, -0.417, -0.444]	0.00496	0.00462
	1.0	[0.0282, 1.26, 0.867, 0.44, 0.131, -0.076, -0.211, -0.298]	0.0121	0.0134
8	1e-07	[0.0403, -1.17, 13.3, -25.1, 29.1, -11.6, -9.77, 5.82, 1.37]	0.0031	0.00328
	1e-06	[0.0348, -0.909, 10.7, -14.2, 8.56, 3.44, -5.61, -5.83, 5.87]	0.00311	0.00326
	1e-05	[0.025, -0.537, 7.75, -6.03, 0.917, 1.51, -1.77, -2.61, 2.76]	0.00312	0.00325
	0.0001	[0.0143, -0.223, 5.73, -1.45, -1.89, -0.944, -0.278, 0.194, 0.859]	0.00314	0.00325
	0.001	[-0.00661, 0.172, 4.19, 0.242, -1.56, -1.62, -0.854, 0.187, 1.26]	0.00319	0.00324
	0.01	[-0.0571, 0.887, 2.44, 0.57, -0.648, -0.964, -0.7, -0.134, 0.559]	0.00381	0.00357
	0.1	[-0.0824, 1.4, 1.22, 0.506, 0.0144, -0.237, -0.325, -0.318, -0.259]	0.00507	0.00472
	1.0	[0.0295, 1.24, 0.864, 0.458, 0.168, -0.0235, -0.147, -0.226, -0.274]	0.0113	0.0124
9	1e-07	[0.0404, -1.17, 13.4, -25.2, 29, -10.8, -10.3, 4.3, 3.61, -0.861]	0.0031	0.00328
	1e-06	[0.0345, -0.888, 10.4, -12.8, 6.61, 2.66, -2.83, -3.37, -0.782, 3.07]	0.00311	0.00326
	1e-05	[0.0264, -0.586, 8.03, -6.29, 0.0812, 2.11, -0.0183, -2.22, -1.74, 2.63]	0.00311	0.00325
	0.0001	[0.0142, -0.226, 5.79, -1.61, -1.88, -0.741, -0.0894, 0.078, 0.169, 0.516]	0.00314	0.00325
	0.001	[-0.00795, 0.205, 4.07, 0.238, -1.38, -1.41, -0.833, -0.161, 0.423, 0.877]	0.0032	0.00325
	0.01	[-0.0547, 0.867, 2.42, 0.656, -0.546, -0.939, -0.81, -0.406, 0.116, 0.669]	0.0037	0.00348
	0.1	[-0.0822, 1.4, 1.2, 0.495, 0.0137, -0.224, -0.299, -0.28, -0.21, -0.115]	0.00513	0.00479
	1.0	[0.0311, 1.22, 0.858, 0.466, 0.19, 0.0109, -0.103, -0.174, -0.217, -0.242]	0.0109	0.0119

Table 1.6 shows the weights, train and test MSE for different regularization parameter (λ) and for degrees 6 to 9.

From this table and from the figure 1.21, we again see that regularization slightly decreases the test MSE, but the parameters 0.001, 0.01, 0.1 and 1 smoothens the graph a lot, thus increasing the test MSE. The MSE figures clearly shows that due to extreme smoothening of the graph, the test MSE increases from degree 6 onward, because regularization was only applied on those degrees.

1.2.4 Train data size 700

1.2.4.1 Plot of approximated functions



(a) Degree 2 (b) Degree 3 (c) Degree 4 (d) Degree 5

(e) Degree 6 (f) Degree 7 (g) Degree 8 (h) Degree 9

Figure 1.22: Approximate functions for degrees 2 to 9 with complete training data set, is shown respectively from (a) to (h)

Figure 1.22, shows the plots of approximate functions for degrees 2 to 9 with complete training training data set.

1.2.4.2 Train and Test MSE

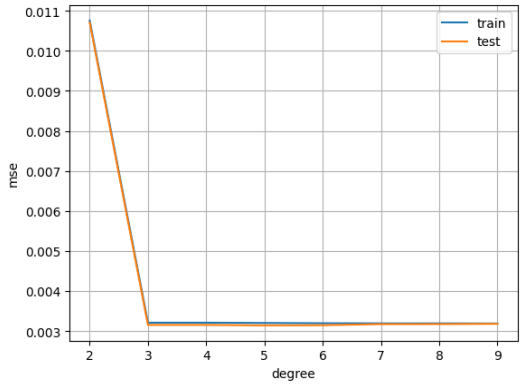


Figure 1.23: MSE for train and test data on complete training training data set without regularization

The plot in Figure 1.23 shows the MSE values for both training and test datasets across polynomial degrees 2 to 9.

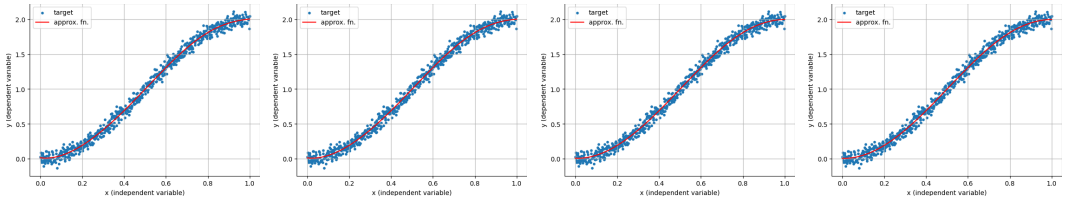
1.2.4.3 Summary of Model Weights and MSE and determining degrees with over-fitting

Table 1.7: Weights, train and test MSEs for degrees 2 to 9

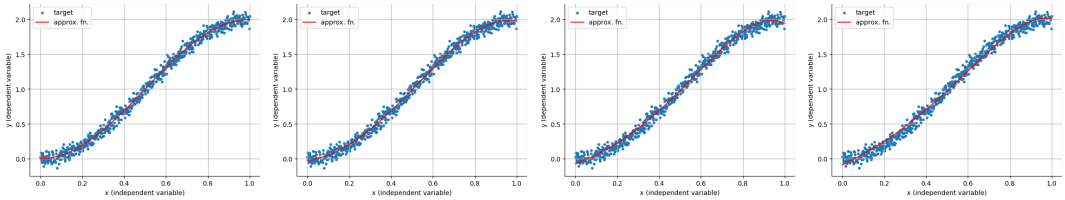
Degree	Weights (w^T)	Train MSE	Test MSE
2	[-0.206, 2.37, 0.0584]	0.0108	0.0107
3	[0.0143, -0.31, 6.83, -4.55]	0.00321	0.00315
4	[0.0145, -0.314, 6.85, -4.58, 0.0166]	0.00321	0.00315
5	[0.00707, -0.0742, 5.13, 0.0486, -5.23, 2.11]	0.0032	0.00314
6	[0.0153, -0.456, 9.04, -15.7, 24.5, -24.1, 8.73]	0.0032	0.00315
7	[0.0222, -0.879, 14.8, -47.6, 112, -150, 99.6, -25.9]	0.00319	0.00317
8	[0.0214, -0.82, 13.8, -40, 83.5, -90.6, 30.4, 16.4, -10.6]	0.00319	0.00317
9	[0.0288, -1.51, 29.1, -184, 784, -2.05e+03, 3.3e+03, -3.18e+03, 1.69e+03, -377]	0.00319	0.00318

Table 1.7 shows the weights, train and test MSEs for degrees 2 to 9. From this table and from the figure 1.23, we see that the test MSE starts increasing from degree 6 onward, indicating over-fitting. Hence regularization is applied on degrees 6 to 9.

1.2.4.4 Plot of approximated functions with regularization

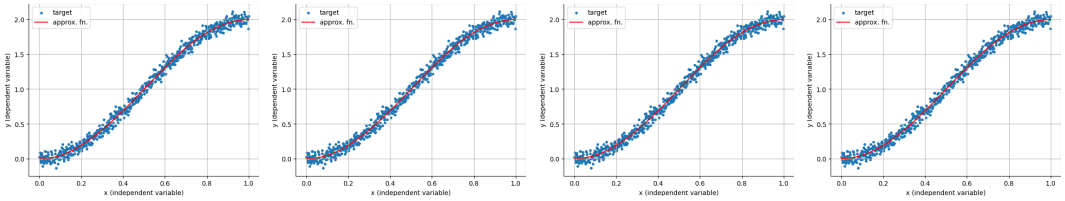


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

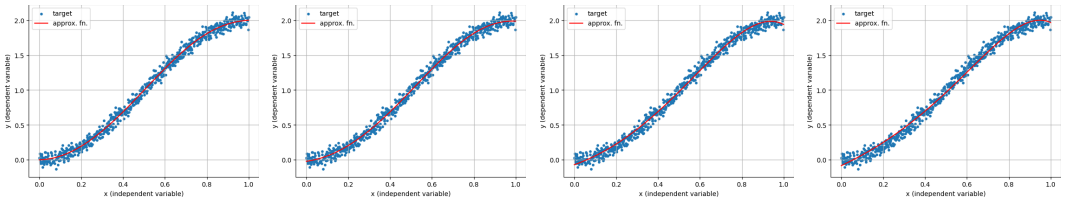


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.24: Approximation functions for varying regularization parameter (λ) values for degree 6

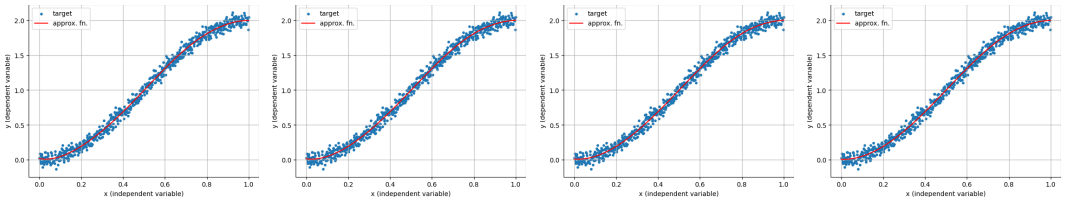


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

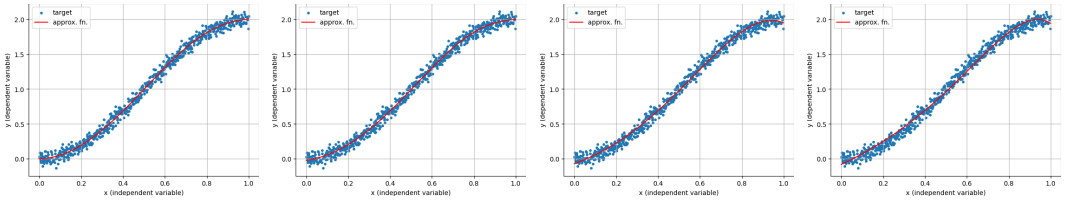


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.25: Approximation functions for varying regularization parameter (λ) values for degree 7

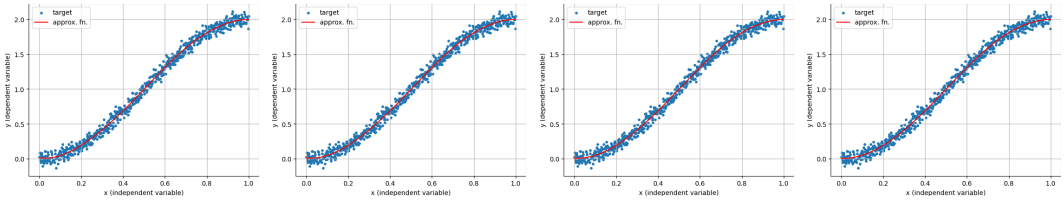


(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

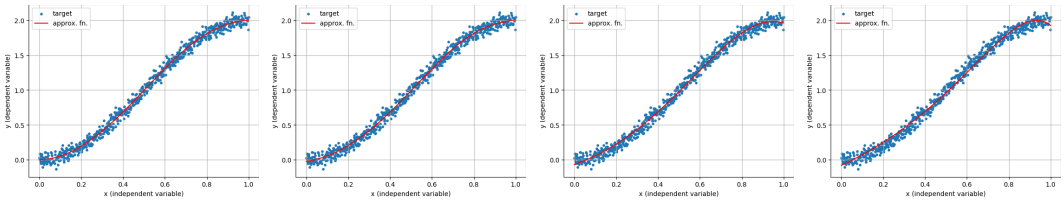


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.26: Approximation functions for varying regularization parameter (λ) values for degree 8



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$

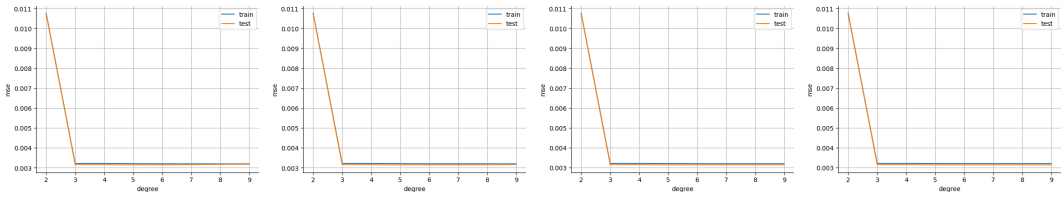


(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

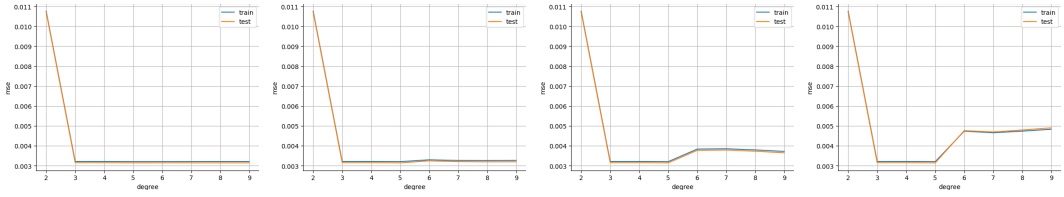
Figure 1.27: Approximation functions for varying regularization parameter (λ) values for degree 9

Figure 1.24 to 1.27 shows the approximation functions for varying regularization parameter (λ) values respectively for degrees 6 to 9.

1.2.4.5 Train and Test MSEs with regularization



(a) $\lambda = 1e-07$ (b) $\lambda = 1e-06$ (c) $\lambda = 1e-05$ (d) $\lambda = 1e-04$



(e) $\lambda = 1e-03$ (f) $\lambda = 1e-02$ (g) $\lambda = 1e-01$ (h) $\lambda = 1e+00$

Figure 1.28: Train and test MSE for varying regularization parameter (λ) values

Figure 1.28 shows the plots of train and test MSE vs degree of function for varying regularization parameter (λ) values.

1.2.4.6 Summary of weights, train, test MSE with regularization

Table 1.8: Weights, Train and Test MSEs for Degrees 6 to 9 with Varying regularization parameter (λ)

Degree	λ	Weights (w^T)	Train MSE	Test MSE
6	1e-07	[0.015, −0.442, 8.9, −15.2, 23.4, −23.1, 8.43]	0.0032	0.00315
	1e-06	[0.0129, −0.348, 7.97, −11.5, 16.7, −17.3, 6.5]	0.0032	0.00314
	1e-05	[0.00843, −0.155, 6.07, −4.03, 2.89, −5.36, 2.58]	0.0032	0.00314
	0.0001	[0.00602, −0.0694, 5.33, −1.35, −1.77, −1.49, 1.34]	0.0032	0.00314
	0.001	[0.00076, 0.0323, 4.86, −0.525, −2.36, −1.34, 1.34]	0.0032	0.00314
	0.01	[−0.0236, 0.396, 3.78, 0.0222, −1.73, −1.19, 0.718]	0.0033	0.00324
	0.1	[−0.0709, 1.06, 1.99, 0.574, −0.391, −0.688, −0.535]	0.00383	0.00377
	1.0	[−0.0897, 1.42, 1.23, 0.539, −0.0284, −0.413, −0.653]	0.00473	0.00476
7	1e-07	[0.0169, −0.569, 10.7, −25.5, 52.6, −66, 39.8, −9.09]	0.00319	0.00315
	1e-06	[0.013, −0.349, 7.88, −10.5, 12.9, −10.7, 0.973, 1.76]	0.0032	0.00314
	1e-05	[0.00999, −0.219, 6.6, −5.43, 3.09, −1.23, −3.11, 2.29]	0.0032	0.00314
	0.0001	[0.00597, −0.0791, 5.47, −1.95, −1.14, −0.739, −0.405, 0.832]	0.0032	0.00314
	0.001	[−0.00124, 0.0804, 4.66, −0.489, −1.82, −1.23, −0.125, 0.924]	0.00321	0.00314
	0.01	[−0.0223, 0.406, 3.6, 0.337, −1.42, −1.44, −0.458, 0.984]	0.00326	0.0032
	0.1	[−0.0713, 1.07, 1.98, 0.562, −0.389, −0.664, −0.485, −0.061]	0.00385	0.00379
	1.0	[−0.0839, 1.41, 1.17, 0.541, 0.06, −0.239, −0.404, −0.483]	0.00465	0.00469
8	1e-07	[0.019, −0.671, 11.7, −28.3, 48.6, −33.2, −22.3, 41.5, −15.3]	0.00319	0.00317
	1e-06	[0.0139, −0.402, 8.55, −13.7, 18.7, −11.5, −9.73, 14.7, −4.65]	0.0032	0.00315
	1e-05	[0.0102, −0.225, 6.62, −5.29, 2.49, −0.747, −2.16, 0.572, 0.736]	0.0032	0.00314
	0.0001	[0.00621, −0.0896, 5.56, −2.15, −1.13, −0.409, −0.254, 0.0336, 0.436]	0.0032	0.00314
	0.001	[−0.0018, 0.0897, 4.65, −0.578, −1.7, −1.04, −0.165, 0.36, 0.394]	0.00321	0.00314
	0.01	[−0.0229, 0.438, 3.42, 0.467, −1.13, −1.33, −0.786, 0.0381, 0.9]	0.00326	0.00319
	0.1	[−0.07, 1.05, 1.99, 0.603, −0.366, −0.692, −0.58, −0.228, 0.234]	0.00379	0.00373
	1.0	[−0.0814, 1.4, 1.13, 0.524, 0.0843, −0.169, −0.292, −0.334, −0.328]	0.00474	0.00479
9	1e-07	[0.0188, −0.648, 11.2, −23.9, 33.5, −12.1, −20.2, 2.67, 24.2, −12.8]	0.00319	0.00317
	1e-06	[0.0154, −0.475, 9.27, −15.9, 19.3, −4.56, −15.1, 1.33, 16.3, −8.27]	0.00319	0.00316
	1e-05	[0.0102, −0.231, 6.71, −5.73, 3.18, −0.532, −3.13, −0.22, 3.04, −1.09]	0.0032	0.00314
	0.0001	[0.00626, −0.0914, 5.57, −2.17, −1.15, −0.385, −0.201, 0.0403, 0.293, 0.0864]	0.0032	0.00314
	0.001	[−0.00179, 0.0897, 4.65, −0.574, −1.7, −1.05, −0.169, 0.364, 0.412, −0.014]	0.00321	0.00314
	0.01	[−0.0239, 0.466, 3.32, 0.483, −0.979, −1.18, −0.786, −0.236, 0.276, 0.667]	0.00326	0.0032
	0.1	[−0.0682, 1.04, 1.99, 0.654, −0.312, −0.684, −0.647, −0.385, −0.0158, 0.388]	0.00372	0.00365
	1.0	[−0.0805, 1.41, 1.12, 0.509, 0.0863, −0.146, −0.248, −0.27, −0.247, −0.2]	0.00483	0.0049

Table 1.8 shows the weights, train and test MSE for different regularization parameter (λ) and for degrees 6 to 9.

Similar to the previous case with train size 100, here also we observers that regularization slightly decreases the test MSE, but as the graph smoothens it increases.

1.2.5 Plots for best Model

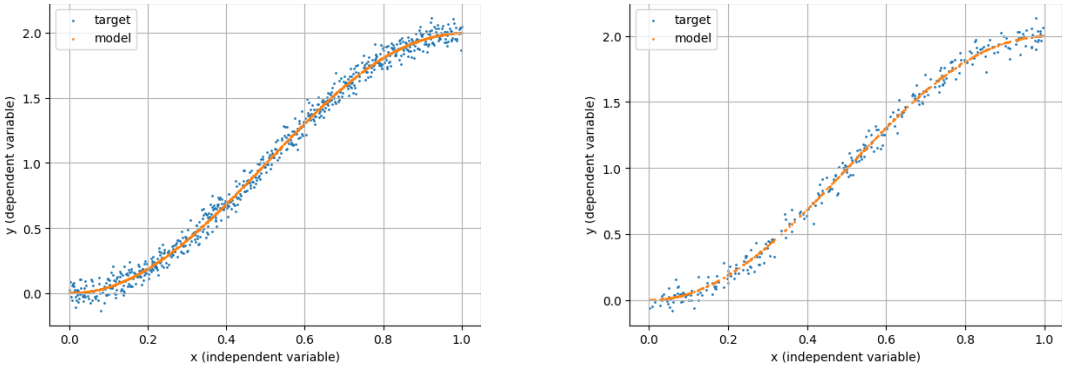
1.2.5.1 Selecting best model

Table 1.9: Weight values, train and test MSE on different train size, degree and reg. parameter (λ), sorted according to test MSE

Train Size	Degree	λ	$\mathbf{w}^T$	Train MSE	Test MSE
700	9	0.0001	[0.00626, −0.0914, 5.57, −2.17, −1.15, −0.385, −0.201, 0.0403, 0.293, 0.0864]	3.20024e-03	3.13938e-03
700	8	0.0001	[0.00621, −0.0896, 5.56, −2.15, −1.13, −0.409, −0.254, 0.0336, 0.436]	3.20026e-03	3.13957e-03
700	8	1e-05	[0.0102, −0.225, 6.62, −5.29, 2.49, −0.747, −2.16, 0.572, 0.736]	3.19839e-03	3.13957e-03
700	7	0.0001	[0.00597, −0.0791, 5.47, −1.95, −1.14, −0.739, −0.405, 0.832]	3.20055e-03	3.14023e-03
700	7	1e-05	[0.00999, −0.219, 6.6, −5.43, 3.09, −1.23, −3.11, 2.29]	3.19827e-03	3.14030e-03
700	6	0.0001	[0.00602, −0.0694, 5.33, −1.35, −1.77, −1.49, 1.34]	3.20135e-03	3.14106e-03
700	9	1e-05	[0.0102, −0.231, 6.71, −5.73, 3.18, −0.532, −3.13, −0.22, 3.04, −1.09]	3.19788e-03	3.14113e-03
700	6	1e-05	[0.00843, −0.155, 6.07, −4.03, 2.89, −5.36, 2.58]	3.19964e-03	3.14125e-03
700	5	0	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	1e-07	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	1e-06	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	1e-05	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	0.0001	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	0.001	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	0.01	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	0.1	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	5	1.0	[0.00707, −0.0742, 5.13, 0.0486, −5.23, 2.11]	3.20296e-03	3.14191e-03
700	7	1e-06	[0.013, −0.349, 7.88, −10.5, 12.9, −10.7, 0.973, 1.76]	3.19693e-03	3.14305e-03
700	7	0.001	[−0.00124, 0.0804, 4.66, −0.489, −1.82, −1.23, −0.125, 0.924]	3.20519e-03	3.14368e-03

The best model is selected based on its test MSE. The table 1.9 shows some of the weight values, train and test MSE on different train size, degree and reg. parameter (λ), sorted according to test MSE. From this table, we conclude our best model is the one having weights [0.00626, -0.0914, 5.57, -2.17, -1.15, -0.385, -0.201, 0.0403, 0.293, 0.0864]. This model has degree 9, regularization parameter value 0.0001 and was trained on complete training data. This model gives us the lowest test MSE of 3.13938e-03.

1.2.5.2 Plot of best model output and target output

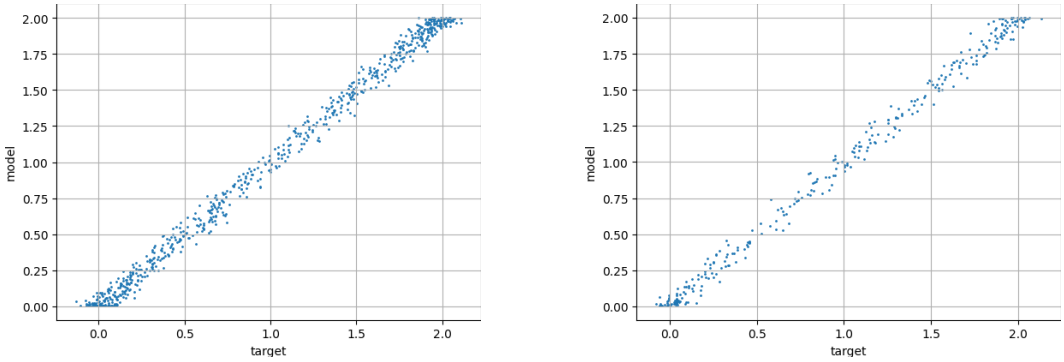


(a) Train data (b) Test data

Figure 1.29: Plot of best output and target output for training data (a) and testing data (b)

Figure 1.29 shows plots of best output and target output for training data and testing data.

1.2.5.3 Scatter plot of target output vs model output



(a) Train data

(b) Test data

Figure 1.30: scatter plot with target output on x-axis and model output on y-axis for training data (a) and testing data (b)

Figure 1.30 shows scatter plot with target output on x-axis and model output on y-axis

1.3 Inference

From all the experiments done so far, we observe the following things:

As the model complexity increases the test MSE first decreases and then starts increasing, whereas the train MSE always decreases. This indicates under-fitting in lower degrees and over-fitting in higher degrees. This trend is proved from the train and test MSE plots in figures 1.2, 1.11, 1.16 and 1.23, and can also be observed from the tables 1.1, 1.3, 1.5 and 1.7.

The degrees where the test MSE starts increasing, we say there is over-fitting in those cases, so we apply regularization on functions with those degrees with regularization parameter values 1e-07, 1e-06, 1e-05, 1e-04, 1e-03, 0.01, 0.1 and 1.

We can observer from the plot of train and test MSE in figure 1.9, that the parameters, 1e-07 and 1e-06, has slight effect on over-fitting. The parameters 1e-05, 1e-04 works very well in eliminating the over-fitting problem, but from 1e-03 onward, we see that the approximate function graph smoothens a lot and the test MSE increases, thus creating under-fitting problem. Similar trends in observed for other training sizes.

Chapter 2

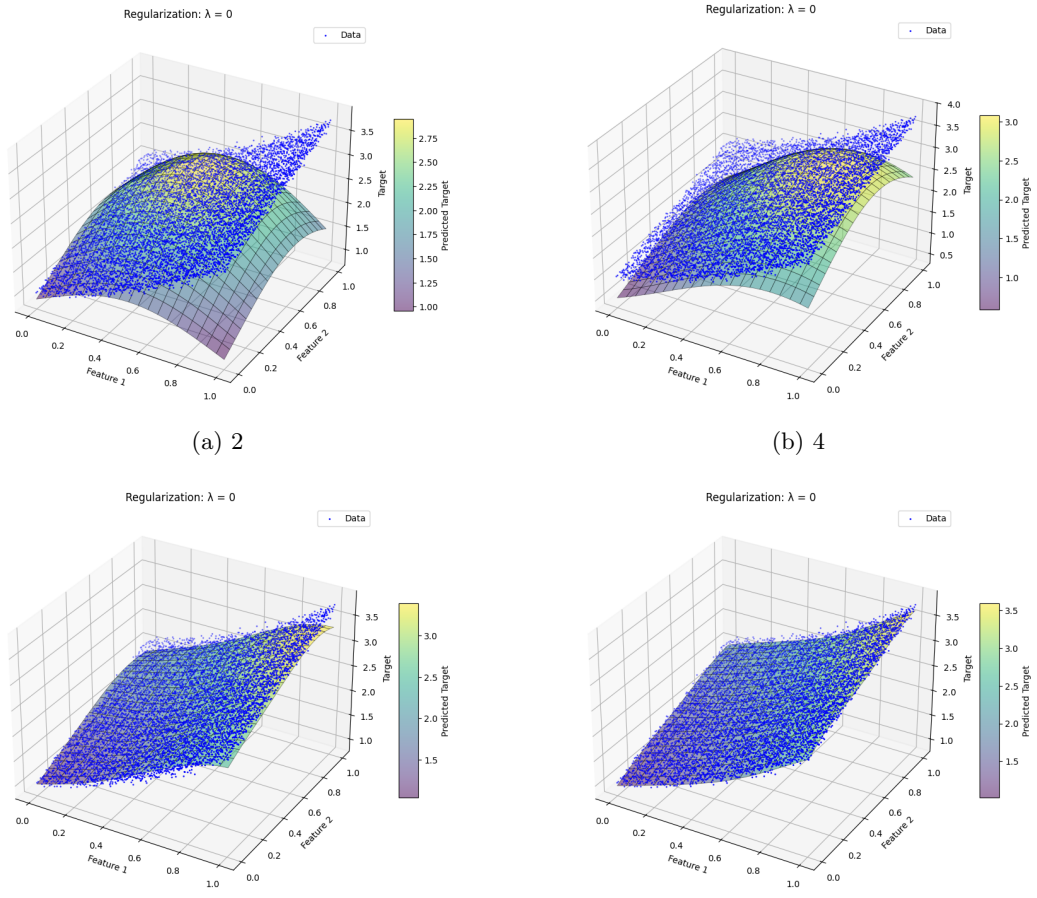
Dataset 1: 2-dimensional (Bivariate) input data

2.1 Dataset description

- Total no. of data points: 10201
- Test data size: 7140
- Train data size: 3061
- Dimension of independent variable: 2
- Dimension of dependent variable: 1

2.2 Result

2.2.1 Without regularization



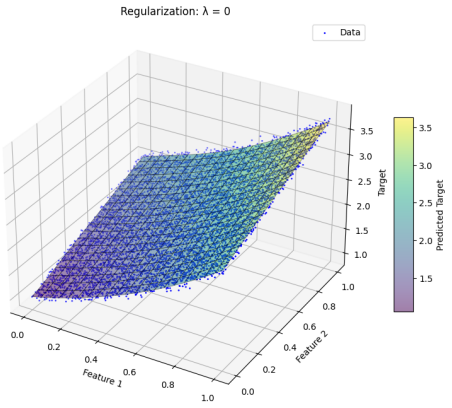
(a) 2

(b) 4

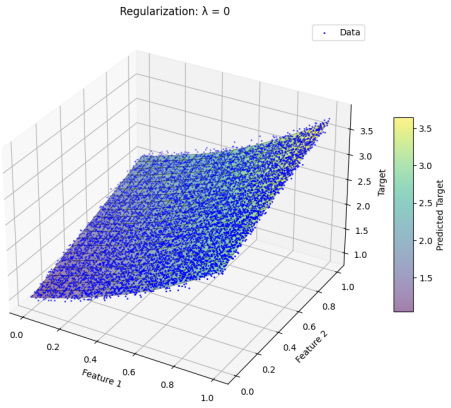
(c) 8

(d) 16

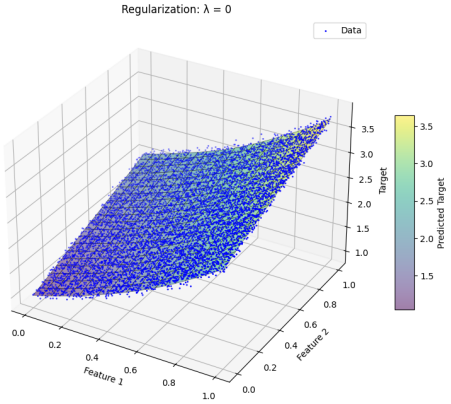
Figure 2.1: Plots of approximate functions for dataset 2 with training data size variations. Each subplot shows a different training data size.



(e) 32



(f) 128



(g) 256

Figure 2.1: (Continued) Plots of approximate functions for dataset 2 with training data size variations.

Figure 2.1 shows the plots of approximate functions with model complexities 2, 4, 8, 16, 32, 128 and 256.

2.2.2 Train and Test MSE

2.2.2.1 Plot of train and test MSE

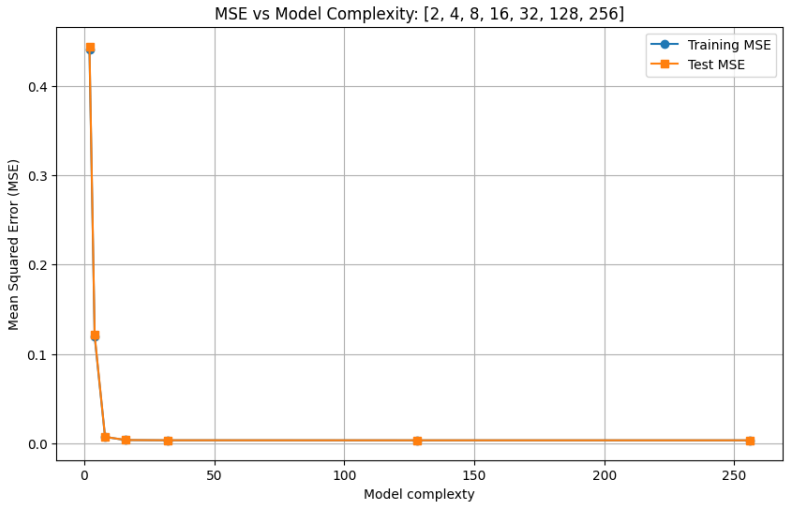


Figure 2.2: MSE for train and test data vs model complexity without regularization

The plot in Figure 2.3 shows the train and test MSE for different model complexities 2, 4, 8, 16, 32, 128 and 256 without regularization.

2.2.2.2 Train and test MSEs for different model complexities

Table 2.1: Train and test MSEs for different model complexities

Model Complexity	Train MSE	Test MSE
2	0.44063	0.44427
4	0.11926	0.12113
8	0.00700	0.00711
16	0.00348	0.00353
32	0.00325	0.00329
128	0.00320	0.00325
256	0.00320	0.00325

Table 2.1, shows the train and test MSEs for different model complexities.

2.2.2.3 MSE vs regularization parameter for model complexity 256

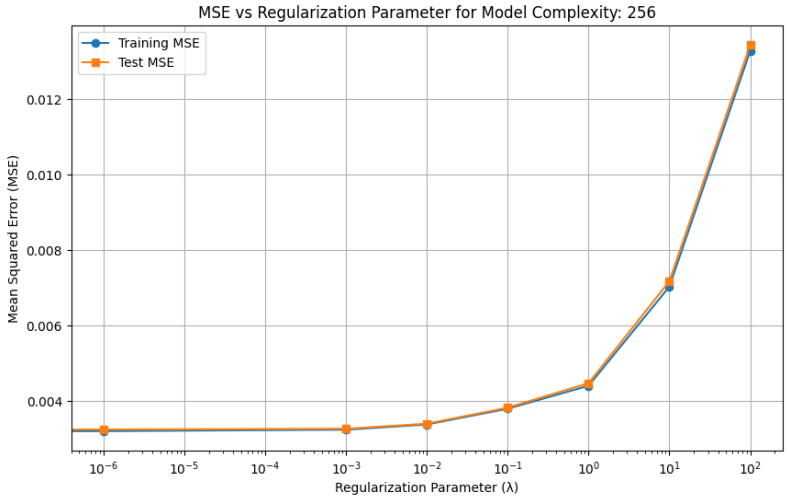
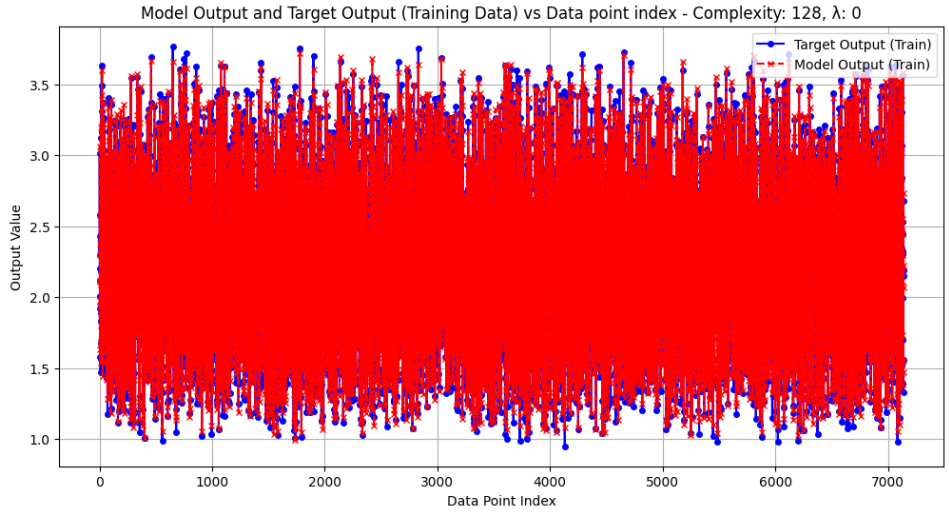


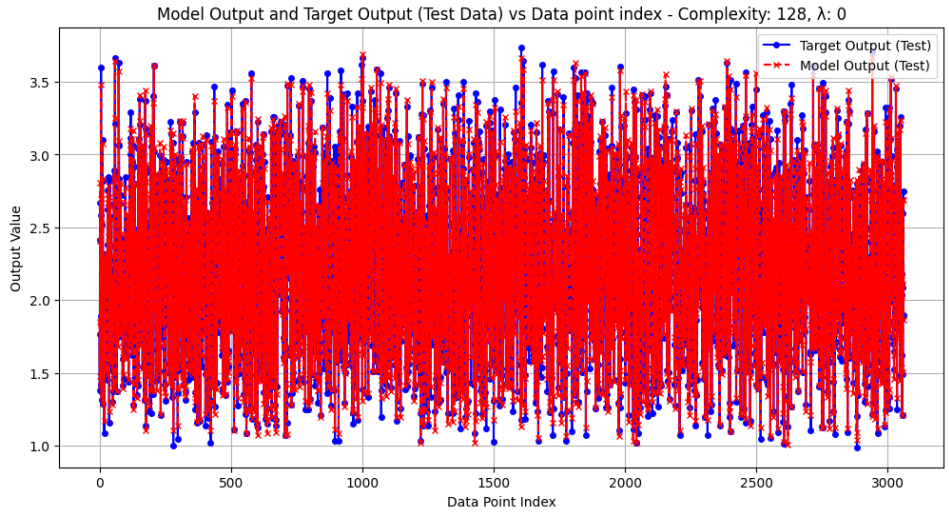
Figure 2.3: MSE for train and test data vs regularization parameter for model complexity 256

The plot in Figure 2.3 shows MSE for train and test data vs regularization parameter for model complexity 256.

2.2.3 Plots of model output and target output for best model complexity



(a) Train data

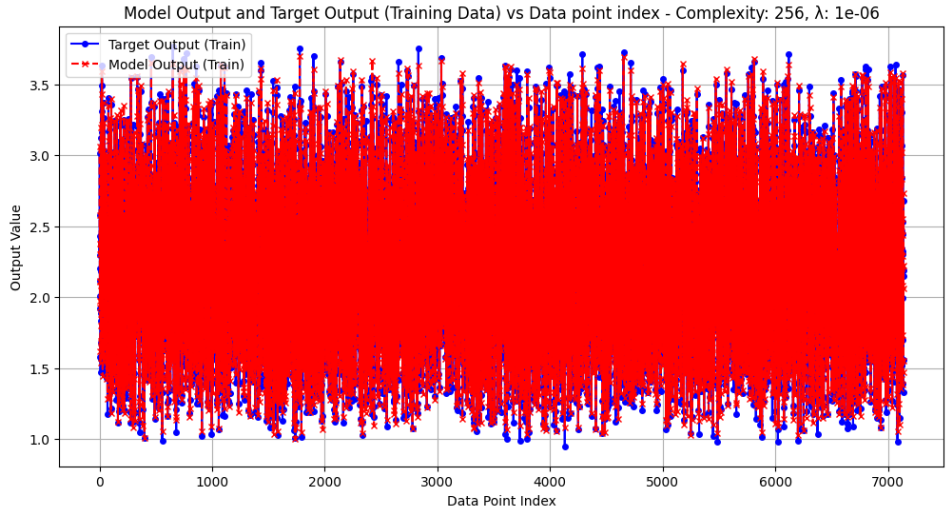


(b) Test data

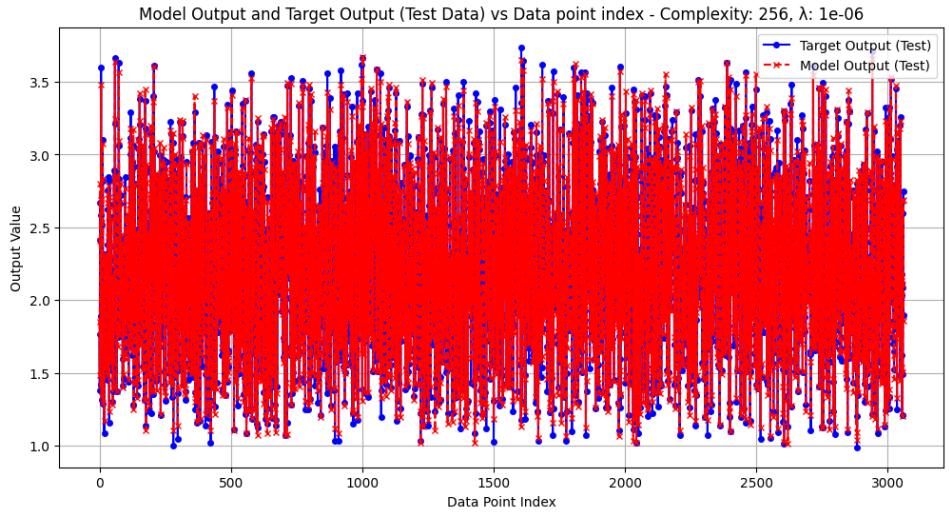
Figure 2.4: Plot of best output and target output for training data (a) and testing data (b)

Figure 2.6 shows plots of model output and target output for best model complexity.

2.2.4 Plots of model output and target output for model complexity 256



(a) Train data



(b) Test data

Figure 2.5: Plot of best output and target output for training data (a) and testing data (b)

Figure 2.6 shows plots of model output and target output for model complexity 256.

2.2.5 Scatter plot for training data, and test data for best model complexity

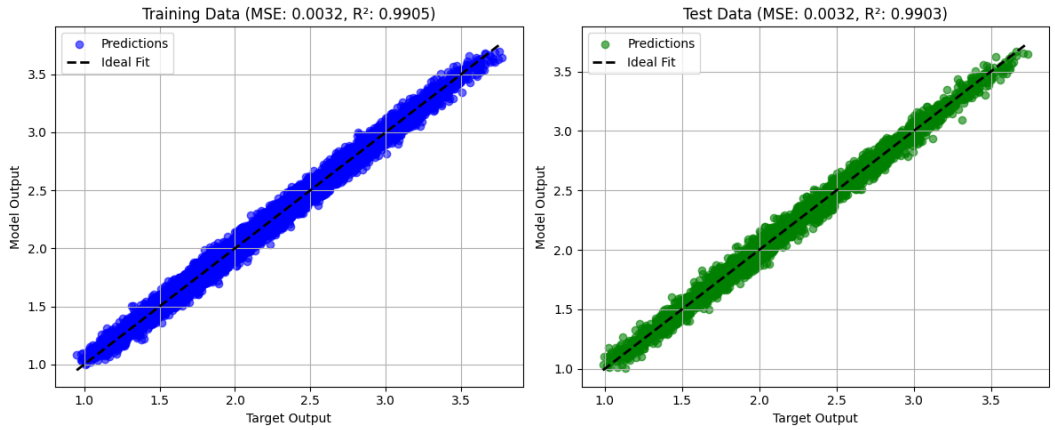


Figure 2.6: Plot of best output and target output for training data (left) and testing data (right)

Figure 2.6 shows the scatter plot with target output on x-axis and model output on y-axis, for training data, and test data for best model complexity.

2.3 Inference

We observer from the table 2.1 that the test MSE decreases with increasing model complexities. The same can be concluded from the plot of MSE for train and test data vs model complexity without regularization in figure 2.3 which shows that there is no over-fitting in the model. Hence, no regularization is needed.

However if we apply regularization for no. of basis function = 256, we observe that the MSE values without regularization and with regularization of 1e-06 are same and for higher regularizations MSE value is increasing slightly. It reflects the fact that regularization is not needed when model is not over-fitted.

One of the potential reasons of model not being over-fitted even for higher model complexities is that the no. of training data points are sufficient and diverse enough to capture all the potential noises and can tackle the complexities of higher number of basis functions.