4)

a) RMSE for your training set: 1.8097

RMSE for the Test set: 2.3618

b)

|  |  |
| --- | --- |
| lambda | RMSE |
| 1 | 1.926127 |
| 0.75 | 1.914525 |
| 0.5625 | 1.90463 |
| 0.421875 | 1.896471 |
| 0.316406 | 1.889996 |
| 0.237305 | 1.885051 |
| 0.177979 | 1.881402 |
| 0.133484 | 1.878776 |
| 0.100113 | 1.876901 |
| 0.075085 | 1.875534 |
| 0.056314 | 1.87448 |
| 0.042235 | 1.873593 |
| 0.031676 | 1.872793 |
| 0.023757 | 1.872057 |
| 0.017818 | 1.871413 |
| 0.013363 | 1.870919 |
| 0.010023 | 1.870638 |
| 0.007517 | 1.870625 |
| 0.005638 | 1.870916 |
| 0.004228 | 1.871536 |

The optimal Lambda is highlighted in yellow. It’s the most optimal because it minimizes the test error. With Lambda = 0.0071517 is got the RMSE to be 2.354

c)

|  |  |
| --- | --- |
| 1 | 1.319325 |
| 0.75 | 1.311439 |
| 0.5625 | 1.304873 |
| 0.421875 | 1.299626 |
| 0.316406 | 1.295619 |
| 0.237305 | 1.292697 |
| 0.177979 | 1.290664 |
| 0.133484 | 1.289315 |
| 0.100113 | 1.28847 |
| 0.075085 | 1.287987 |
| 0.056314 | 1.287772 |
| 0.042235 | 1.287769 |
| 0.031676 | 1.287958 |
| 0.023757 | 1.28834 |
| 0.017818 | 1.288923 |
| 0.013363 | 1.289706 |
| 0.010023 | 1.290677 |
| 0.007517 | 1.291812 |
| 0.005638 | 1.293083 |
| 0.004228 | 1.294463 |

The optimal Lambda is highlighted in yellow. It’s the most optimal because it minimizes the test error. With Lambda = 0.042235 is got the RMSE to be 2.3518

4)

|  |  |
| --- | --- |
| 1 | 1.58556 |
| 0.75 | 1.577639 |
| 0.5625 | 1.573707 |
| 0.421875 | 1.573529 |
| 0.316406 | 1.576583 |
| 0.237305 | 1.582054 |
| 0.177979 | 1.588963 |
| 0.133484 | 1.596335 |
| 0.100113 | 1.603366 |
| 0.075085 | 1.609516 |
| 0.056314 | 1.614518 |
| 0.042235 | 1.618332 |
| 0.031676 | 1.621073 |
| 0.023757 | 1.622936 |
| 0.017818 | 1.624132 |
| 0.013363 | 1.624842 |
| 0.010023 | 1.625198 |
| 0.007517 | 1.625286 |
| 0.005638 | 1.62517 |
| 0.004228 | 1.624903 |

As the k fold increase the RMSE tend to decrease. 10 fold 2.352 < 5 fold 2.354 < 1 fold 2.372. Validation set if effectively k-fold with k = 1. So as the number of fold increases the error decrease a little.

The optimal Lambda is highlighted in yellow. It’s the most optimal because it minimizes the test error. With Lambda = 0.421875 is got the RMSE to be 2.372