CIS 581 001 – COMPUTATIONAL LEARNING

MIDTERM PROJECT

POLYNOMIAL CURVE FITTING REGRESSION FOR WORKING-AGE DATA

NAME	:	SURYA SUBRAMANI
UMID	:	69499602
DATE OF SUBMISSION	:	03-13-2023

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

Sum of RMSE AVERAGE CV = 6	**Degree**					
Lambda	**0**	**1**	**2**	**3**	**4**	**5**
0	1.016	1.084	0.775	0.783	0.493	0.570
e^{-25}	1.016	1.084	0.775	0.783	0.493	0.570
e^{-20}	1.016	1.084	0.775	0.783	0.493	0.570
e^{-7}	1.016	1.084	0.776	0.783	0.494	0.571
e^{-14}	1.016	1.084	0.775	0.783	0.493	0.570
e^{-3}	1.035	1.097	0.795	0.812	0.550	0.642
1	2.168	2.178	2.974	3.173	3.002	3.074
e^{3}	23.982	24.287	23.427	23.609	25.006	25.303
e^{7}	63.681	63.713	61.974	62.093	58.273	58.478

Sum of RMSE AVERAGE CV = 6	**Degree**					
Lambda	**7**	**8**	**9**	**10**	**11**	**12**
0	0.188	0.146	0.236	0.165	0.626	0.687
e^{-25}	0.188	0.146	0.236	0.165	0.626	0.687
e^{-20}	0.188	0.146	0.236	0.165	0.626	0.687
e^{-7}	0.185	0.133	0.191	0.127	0.099	0.432
e^{-14}	0.142	0.188	0.146	0.236	0.165	0.624
e^{-3}	0.238	0.344	0.461	0.242	0.305	0.469
1	3.532	2.945	3.042	3.392	4.012	3.845
e^{3}	24.435	24.239	24.619	27.148	29.966	30.190
e^{7}	55.010	58.115	59.655	62.302	66.737	62.201

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

   d* = **11**, λ∗ = **e^{-7}**
   RMSE Average is **0.0987**

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

   Weights for d* = **11**, λ∗ = **e^{-7}** , all training data

w0	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11
65.469	0.500	5.303	-0.218	-6.615	0.558	3.222	-0.494	-0.880	0.224	0.107	-0.036

   Weights for d = **12**, λ∗ = **e^{-7}** , all training data

w0	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
65.472	0.510	5.202	-0.315	-6.076	0.800	2.276	-0.733	-0.166	0.326	-0.135	-0.052	0.030

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

   Weights and RMSE for optimal Model d* = **11**, λ∗ = **e^{-7}**, all training data

w0	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11
65.469	0.500	5.303	-0.218	-6.615	0.558	3.222	-0.494	-0.880	0.224	0.107	-0.036

   RMSE for Training: 0.06626
   RMSE for Test: 0.3944

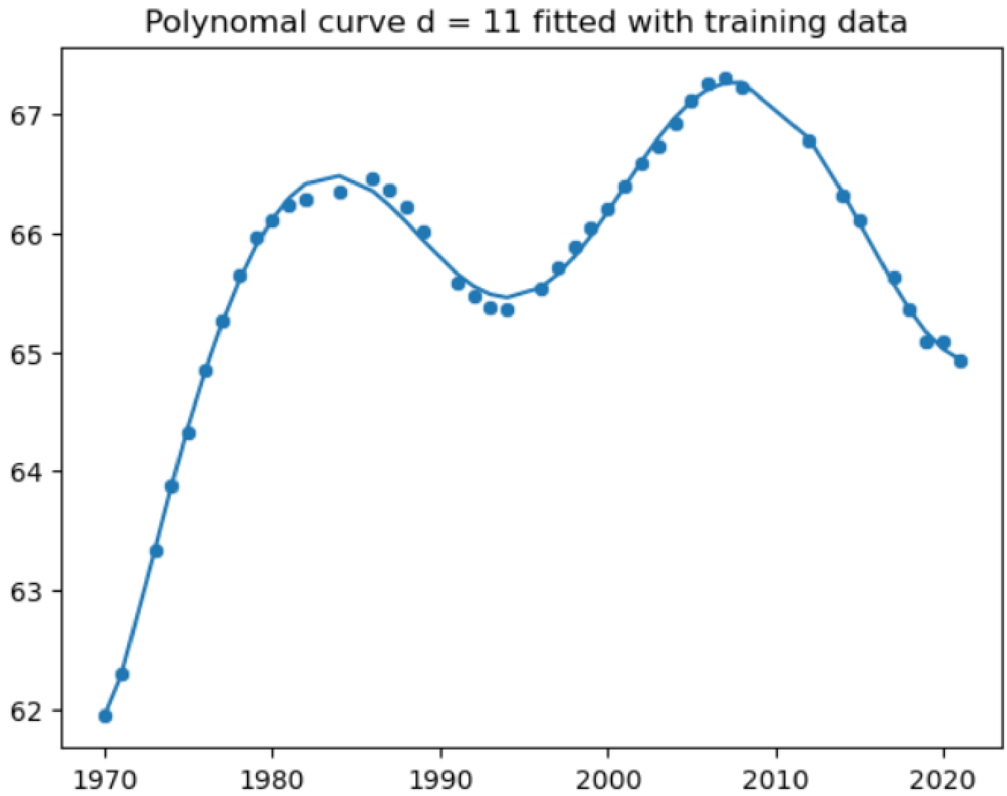
   Weights and RMSE for d = **12**, λ∗ = **e^{-7}** , all training data

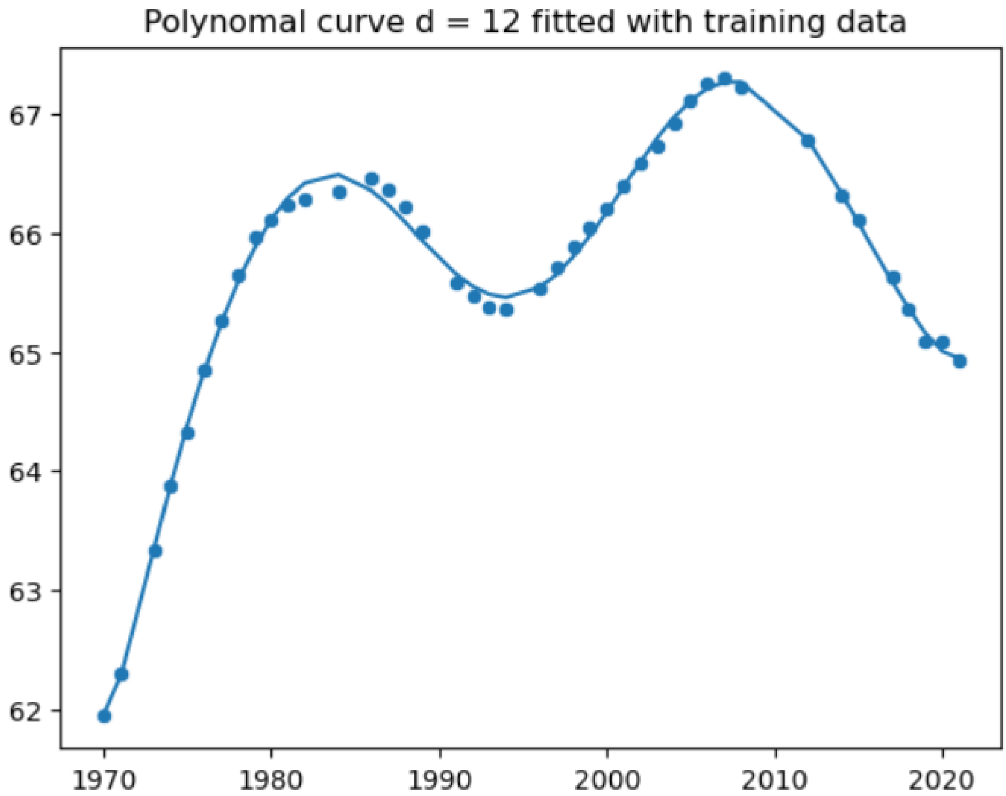
w0	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
65.472	0.510	5.202	-0.315	-6.076	0.800	2.276	-0.733	-0.166	0.326	-0.135	-0.052	0.030

   RMSE for Training: 0.06766
   RMSE for Test: 0.5594

5. 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

Combined Graph: (Scatter plot is training data and line plot is polynomial curve)





6. Brief discussion of your findings and observations.

From the data and the graph, we can able to observe that for every 20 years the age indicator goes to the peak and starts descending. This repeats for every 30 years with an increase in indicator.

Criteria	Year	True Values
count	42	42
mean	1994.80952	65.70398
std	15.209493	1.181847
min	1970	61.93886
25%	1981.25	65.35731
50%	1995	65.98761
75%	2005.75	66.36326
max	2021	67.29843

(the above shown Data is before scaling)

To reduce the standard deviation, which can affect the model training, The input matrix is scaled to minimum values using (X-mean)/std.

Criteria	Year	True Values
count	42.00	42.00
mean	0.00	65.70
std	1.00	1.18
min	-1.63	61.94
25%	-0.89	65.36
50%	0.01	65.99
75%	0.72	66.36
max	1.72	67.30

(the above shown Data is after scaling)

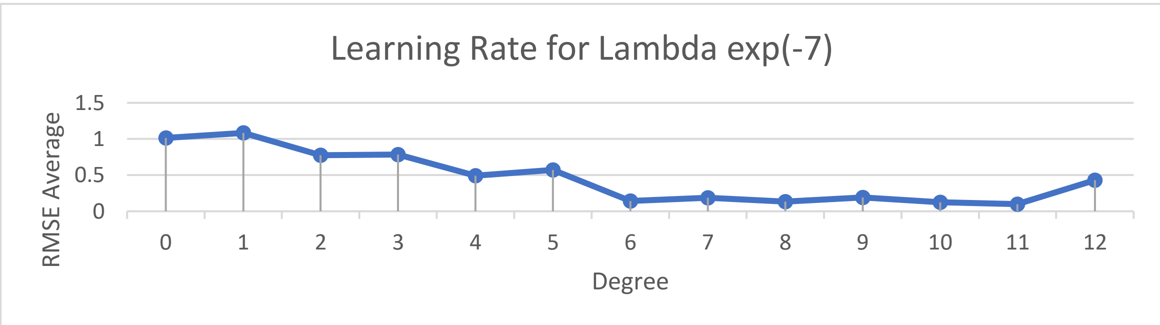
To train model and calculate weights I have used the below equation

$$W = (X^T.X + lambda*I).X^T.y$$

Which is the derivation of

$$L(\mathbf{w}) = \frac{1}{2}\sum_{l=1}^{m}\left(y^{(l)} - \sum_{i=0}^{d} w_i \left(x^{(l)}\right)^i\right)^2 + \frac{\lambda}{2}\|\mathbf{w}\|_2^2$$

Exp(7) gives higher RMSE a



As you can see from the above graph the optimal degree is 11 for the l2 penalty exp(7).