NCTU Machine Learning Hw1

Liang Yu Pan 0486016

October 25, 2015

1 Bayesian Linear Regression

Overview For Bayesian, the data is certain. So, the **w** can be viewed as uncertain distribution, so we need to integral **w**. Therefore, the problem can be transformed to be $p(t) = \int_{\infty}^{\infty} p(t|\mathbf{w})p(\mathbf{w})$. Then **w** is calculated by **x**, **t**($= p(\mathbf{w}|\mathbf{x}, \mathbf{t})$), using **maximum posterior**. The $p(t|\mathbf{w})$ is calculated by **maximum likelihood function**($= p(t|x, \mathbf{w}, \beta)$).

Useful formula If $p(x), p(y), p(x|y), p(y|x)$, are gaussian distribution.

$$p(x) = N(x|\mu, \Lambda^{-1}) \tag{1}$$
$$p(y|x) = N(y|Ax + b, L^{-1}) \tag{2}$$
$$p(y) = N(y|A\mu + b, L^{-1} + A\Lambda^{-1}A^T) \tag{3}$$
$$p(x|y) = N(x|\Sigma A^T L(y - b) + \Lambda\mu, \Sigma), \Sigma = (\Lambda^{-1} + A^T L A^T)^{-1} \tag{4}$$

And now we see x as **w** and y as t.

Calculate $p(w)$ Because $p(t|w)$ is a likelihood function of gaussian distribution given by the problem, we can assume its prior as gaussian distribution due to **conjugacy** property. And *posterior* $\propto$ *likelihood* $\times$ *prior*. Then by the formula 1 and 4, we can get

$$\begin{aligned} p(w|t) &= N(w|m_N, S_N) \\ m_N &= S_N(S_0^{-1}m_0 + \beta\Phi^T t) \\ S_N^{-1} &= S_0^{-1} + \beta\Phi^T\Phi \end{aligned}$$

Let $m_N = 0, S_N = \alpha^{-1}$, we get

$$\begin{aligned} p(w|\alpha) &= N(w|0, \alpha^{-1}I) \\ m_N &= \beta S_N \Phi T t \\ S_N^{-1} &= \alpha I + \beta\Phi^T\Phi \end{aligned}$$

Calculate $p(t)$ By the formula 3, let $y = t, A = \Phi, b = 0, L^{-1} = \beta^{-1}, x = w, \mu = m_N, \Lambda^{-1} = S_N$. We get the following answer

$$m(x) = \beta\phi(x)^T \sum_{n=1}^{N} \phi(x_n)t_n$$
$$s^2(x) = \beta^{-1} + \phi(x)^T S\phi(x)$$
$$S^{-1} = \alpha I + \beta \sum_{n=1}^{N} \phi(x_n)\phi(x_n)^T$$

2 Polynomial Curve Fitting

Overview Because the error function is $E(w) = 1/2 \sum_{n=1}^{N} y(x_n, w) - {t_n}^2$. To minimize the function, using differential with $w_0, ..., w_M$.

(a)

Order	Train_RMS	Test_RMS
0	1.4646	0.89624
1	1.4183	1.1457
2	1.3657	1.3403
3	0.97111	2.5302
4	0.67148	1.4339
5	0.43383	1.8343
6	0.33239	1.4827
7	0.30948	2.5546
8	0.29367	4.6491
9	0.29358	4.9425

Figure 1:

(b)

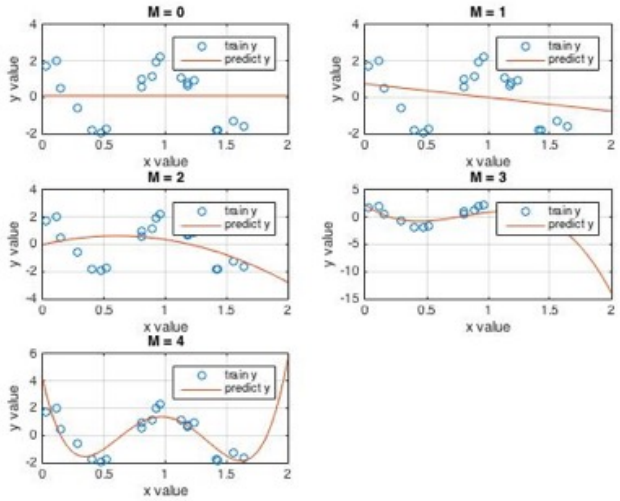

Figure 2:

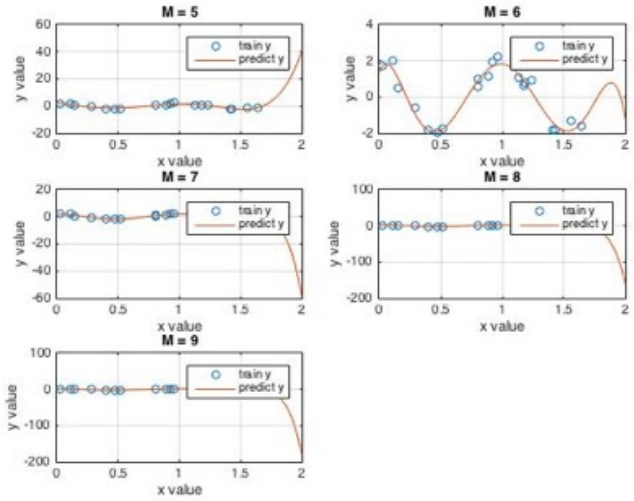

Figure 3:

(c)

Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0	0.097105	0.73961	-0.017728	2.2713	4.117	2.3642	1.3706	1.8507	1.3766	1.4319
1	NaN	-0.75019	2.129	-15.901	-40.122	-6.8095	17.624	0.84799	20.54	17.86
2	NaN	NaN	-1.7602	25.105	91.499	-48.635	-192.72	-47.056	-254.56	-220.3
3	NaN	NaN	NaN	-10.6	-72.742	148.75	490.68	-16.998	888.89	700.84
4	NaN	NaN	NaN	NaN	18.607	-128.81	-510.75	348.22	-1667.9	-1125.9
5	NaN	NaN	NaN	NaN	NaN	34.909	235.62	-516.74	1975.8	1077
6	NaN	NaN	NaN	NaN	NaN	NaN	-40.005	288.33	-1444.8	-557.27
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-56.464	577.58	62.451
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-94.919	67.278
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-21.373

Figure 4:

(d)

Lambda	Order	Train_RMS	Test_RMS
0.1	0	1.4646	0.89644
0.1	1	1.4183	1.1414
0.1	2	1.3666	1.3133
0.1	3	1.1778	1.7585
0.1	4	1.1289	2.2114
0.1	5	1.1221	2.1544
0.1	6	0.99645	1.4799
0.1	7	0.84222	1.0387
0.1	8	0.79294	2.0294
0.1	9	0.81032	2.7846

Figure 5:

(e)

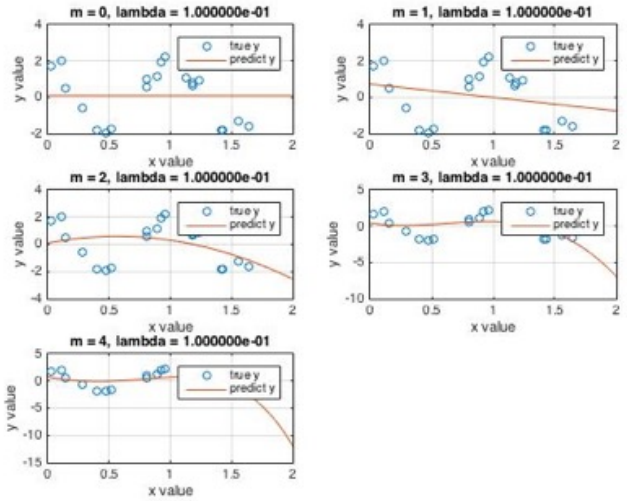

Figure 6:

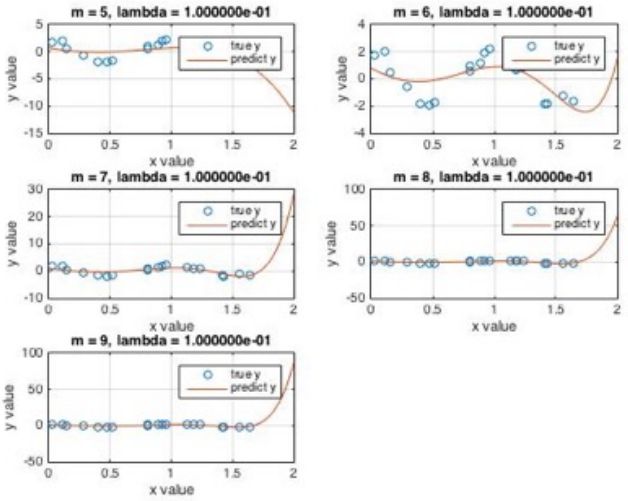

Figure 7:

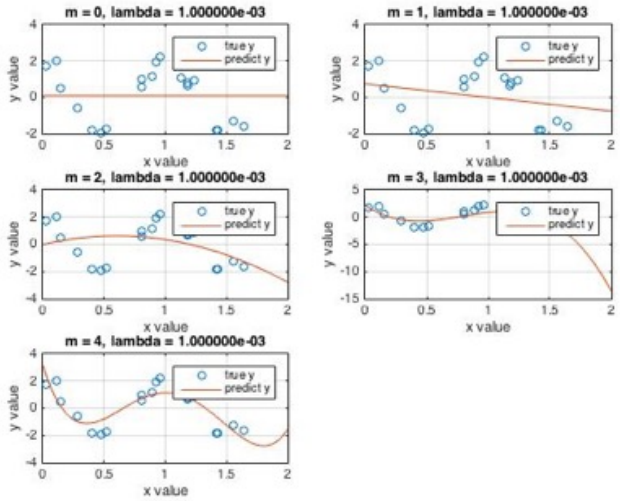

Figure 8:

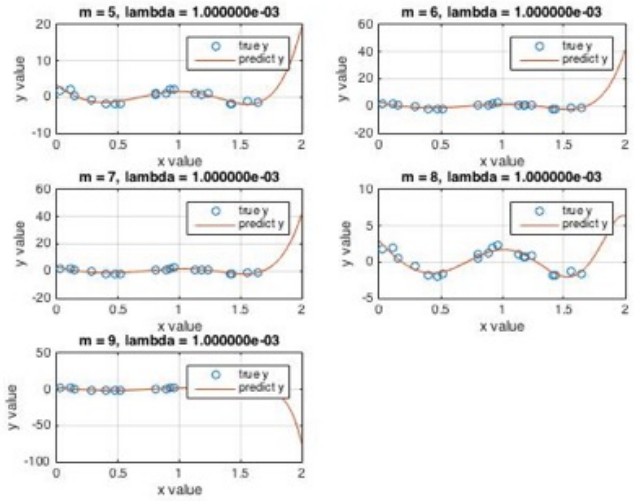

Figure 9:

(f)

Lmbda	Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0.1	0	0.096862	0.72554	0.095518	0.4378	0.72588	0.71768	0.8837	0.98754	1.0939	1.1042
0.1	1	NaN	-0.73588	1.7369	-2.6461	-3.8234	-3.8811	-4.9847	-5.9758	-6.0995	-5.9274
0.1	2	NaN	NaN	-1.5312	6.218	4.2671	4.4777	5.0274	4.1376	3.04	2.7127
0.1	3	NaN	NaN	NaN	-3.3719	1.8257	1.9882	4.212	5.3084	4.7895	4.3165
0.1	4	NaN	NaN	NaN	NaN	-2.2931	-2.8109	-1.3512	1.5984	2.5823	2.3947
0.1	5	NaN	NaN	NaN	NaN	NaN	0.22031	-4.9387	-3.8222	-0.6503	-0.22688
0.1	6	NaN	NaN	NaN	NaN	NaN	NaN	2.1331	-4.5416	-2.9225	-1.9074
0.1	7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.674	-2.5686	-1.8191
0.1	8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.0279	-0.16003
0.1	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.81509

Figure 10:

Lmbda	Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0.001	0	0.097102	0.73947	-0.016343	2.1934	3.1733	3.2537	2.9419	2.939	2.9853	2.748
0.001	1	NaN	-0.75004	2.1242	-15.346	-28.889	-25.084	-19.658	-19.632	-18.179	-15.283
0.001	2	NaN	NaN	-1.7574	24.318	62.577	30.974	15.677	15.705	8.6642	0.97603
0.001	3	NaN	NaN	NaN	-10.3	-46.794	20.643	20.867	20.649	25.83	23.853
0.001	4	NaN	NaN	NaN	NaN	11.07	-42.657	-8.9537	-8.9194	1.2006	11.365
0.001	5	NaN	NaN	NaN	NaN	NaN	14.375	-18.182	-17.717	-21.53	-11.065
0.001	6	NaN	NaN	NaN	NaN	NaN	NaN	8.9593	8.5057	-12.454	-20.049
0.001	7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.12268	20.671	-4.8725
0.001	8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-5.3931	21.202
0.001	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-7.0133

Figure 11:

(g)

a.

Order	Train_RMS	Test_RMS
0	1.347	1.1884
1	1.3394	1.2165
2	1.3073	1.3355
3	1.2929	1.3515
4	0.79734	1.3712
5	0.76153	1.3667
6	0.36607	1.5321
7	0.36149	1.5097
8	0.33584	1.4741
9	0.3355	1.4684

Figure 12:

b.

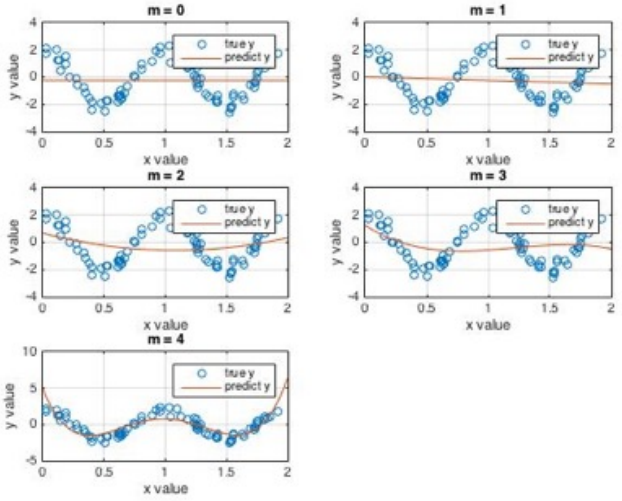

Figure 13:

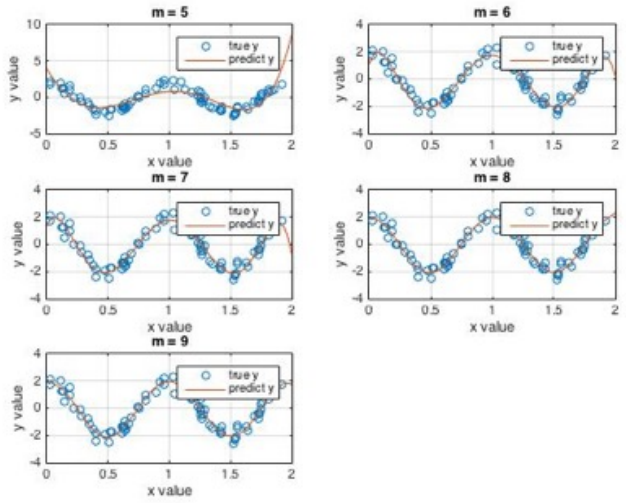

Figure 14:

c.

Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0	-0.23886	0.012256	0.6877	1.2478	4.9366	3.9507	1.087	1.3562	2.0466	2.1453
1	NaN	-0.25485	-2.3777	-5.7429	-41.968	-28.557	26.643	19.638	-3.7703	-7.9075
2	NaN	NaN	1.102	5.4094	89.228	42.816	-238.47	-190.52	15.887	60.895
3	NaN	NaN	NaN	-1.4852	-68.959	-6.1515	576.37	440.08	-332.13	-544.25
4	NaN	NaN	NaN	NaN	17.513	-18.794	-584.94	-392.68	1092.2	1619.3
5	NaN	NaN	NaN	NaN	NaN	7.515	265.64	123.36	-1464	-2219.9
6	NaN	NaN	NaN	NaN	NaN	NaN	-44.572	8.3066	961.1	1608.2
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-7.7841	-388.13	-634.33
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	38.682	127.97
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-10.232

Figure 15:

d.

Lambda	Order	Train_RMS	Test_RMS
0.1	0	1.347	1.1882
0.1	1	1.3394	1.2165
0.1	2	1.3074	1.3279
0.1	3	1.2965	1.327
0.1	4	1.2302	1.2735
0.1	5	0.98131	1.2033
0.1	6	0.90131	1.242
0.1	7	0.91373	1.2502
0.1	8	0.78342	1.2183
0.1	9	0.62329	1.2292

Figure 16:

Lambda	Order	Train_RMS	Test_RMS
0.001	0	1.347	1.1884
0.001	1	1.3394	1.2165
0.001	2	1.3073	1.3354
0.001	3	1.2929	1.3511
0.001	4	0.80149	1.3378
0.001	5	0.7623	1.3577
0.001	6	0.6991	1.3712
0.001	7	0.49002	1.4099
0.001	8	0.44247	1.4148
0.001	9	0.44568	1.4138

Figure 17:

e.

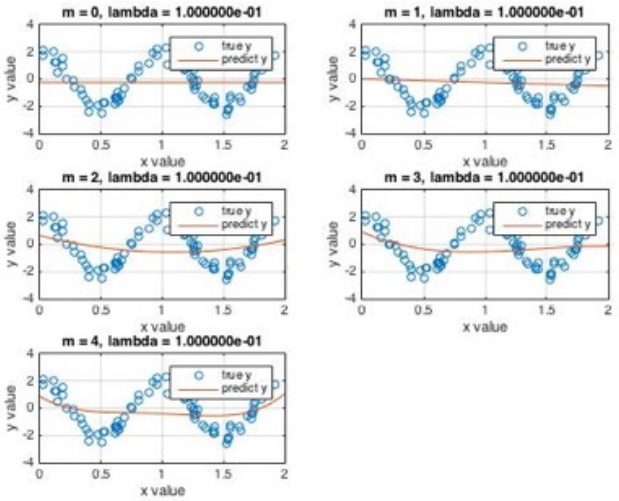

Figure 18:

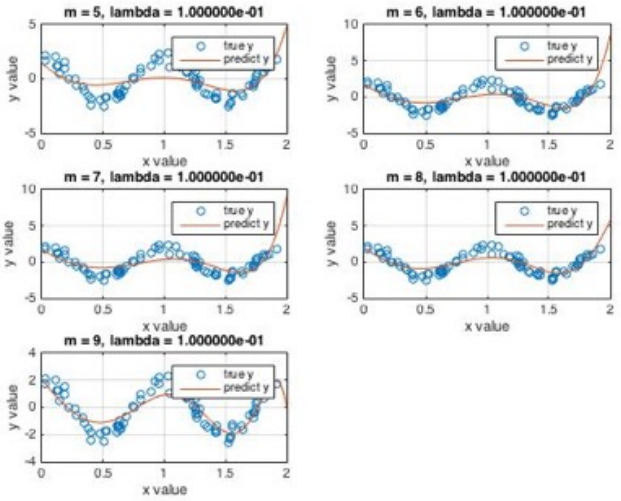

Figure 19:

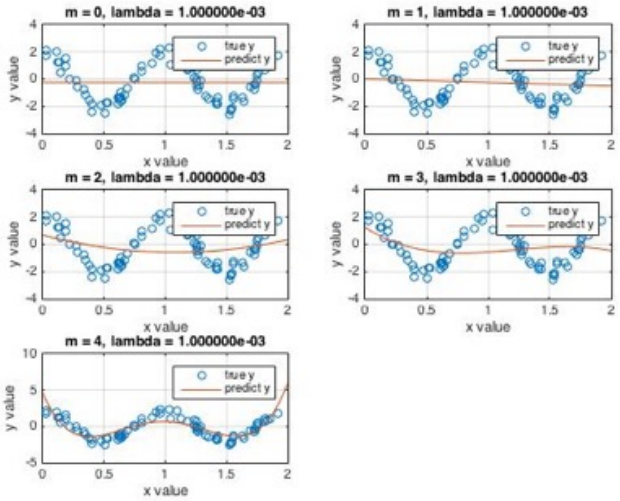

Figure 20:

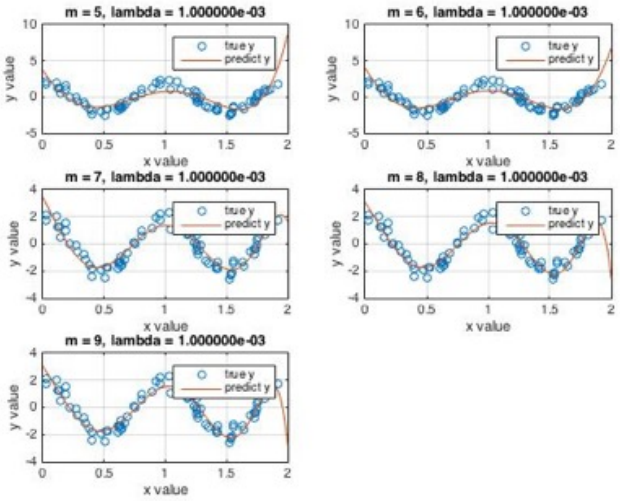

Figure 21:

f.

Lmbda	Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0.1	0	-0.23871	0.011728	0.64975	0.90675	0.87609	1.4536	1.7191	1.7038	1.9015	2.0793
0.1	1	NaN	-0.25432	-2.2719	-3.9221	-5.3958	-9.9374	-10.193	-9.9209	-11.34	-11.686
0.1	2	NaN	NaN	1.0501	3.2394	9.2545	12.258	8.19	7.7435	7.9387	5.7985
0.1	3	NaN	NaN	NaN	-0.77191	-7.0029	4.3058	6.885	6.4128	9.1124	9.6566
0.1	4	NaN	NaN	NaN	NaN	1.8751	-12.15	-2.62	-2.2278	0.10331	3.9326
0.1	5	NaN	NaN	NaN	NaN	NaN	4.194	-6.681	-5.4623	-7.6581	-4.024
0.1	6	NaN	NaN	NaN	NaN	NaN	NaN	3.0491	1.7432	-5.0979	-7.6628
0.1	7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.34257	7.3979	-2.6489
0.1	8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-1.7788	7.7909
0.1	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-2.3013

Figure 22:

Lmbda	Order	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
0.001	0	-0.23886	0.012251	0.6873	1.2425	4.5857	3.7387	3.9957	3.4456	3.0494	3.0507
0.001	1	NaN	-0.25484	-2.3766	-5.7144	-38.795	-26.19	-27.354	-17.363	-13.636	-13.763
0.001	2	NaN	NaN	1.1015	5.3753	82.276	35.826	27.98	-10.865	-12.318	-11.487
0.001	3	NaN	NaN	NaN	-1.474	-63.567	2.1769	38.071	70.413	46.093	44.966
0.001	4	NaN	NaN	NaN	NaN	16.151	-23.132	-71.828	-17.587	7.9481	7.1321
0.001	5	NaN	NaN	NaN	NaN	NaN	8.3381	35.355	-70.649	-34.017	-32.506
0.001	6	NaN	NaN	NaN	NaN	NaN	NaN	-5.3297	55.662	-16.523	-14.903
0.001	7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-11.785	28.418	24.999
0.001	8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-7.4862	-5.6324
0.001	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-0.33243

Figure 23:

3 Application for Polynomial Regression

Overview Like problem 2, using differential to get the parameter.

(a)

Order	Train_RMS	Test_RMS
1	0.22628	0.17112
2	0.18859	0.15323

Figure 24:

(b)

Dimension	Train_RMS	Test_RMS
1	0.50375	0.35805
2	0.70422	0.69297
3	0.24568	0.20435
4	0.23804	0.11363

Figure 25: