Kenneth Kitahata & Spencer Tu

February 26, 2020
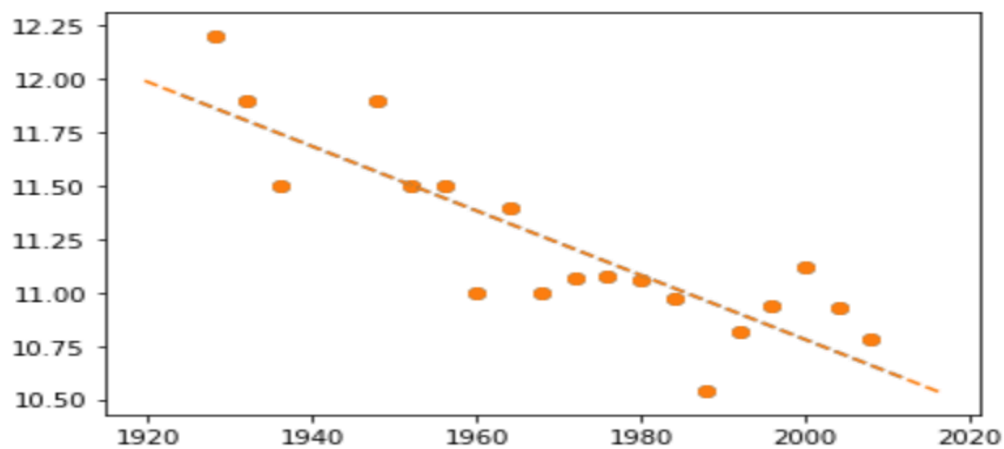
STAT 339

HW1b Write Up

1. b.)

```
Our model: [ 4.09241546e+01 -1.50718122e-02]
Package reg model: [ 4.09241546e+01 -1.50718122e-02]
```


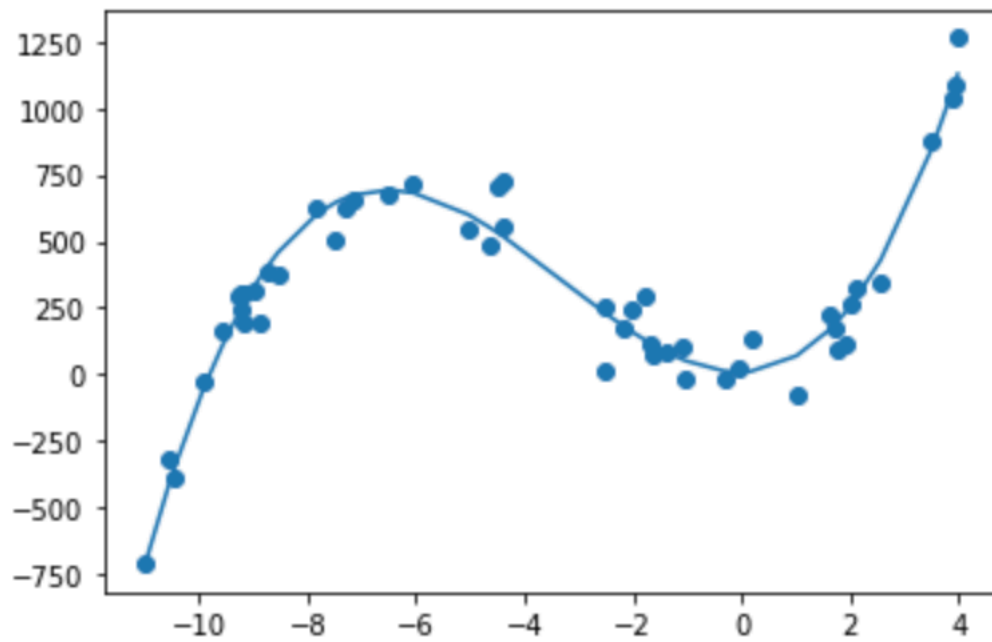The coefficients returned by our model and by package regression model are the same.
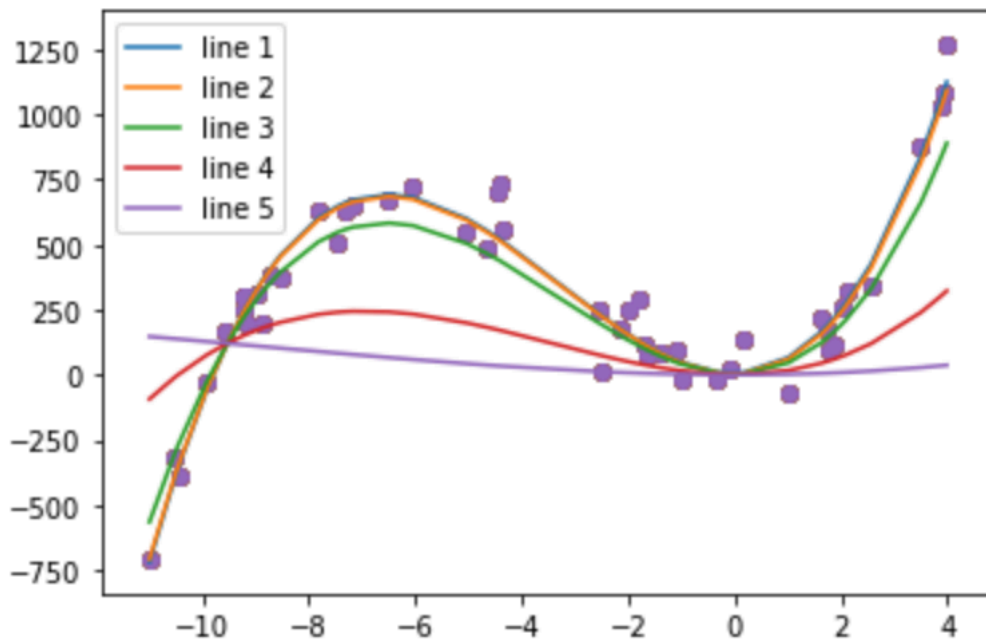
1. c.)

```
2012time: 10.599668379261583
2016time: 10.539381130312453
squared prediction error 2012: 0.022599596193839146
squared prediction error 2016: 0.029110798693456456
```

1. g.)

1. h.)

As Lamda becomes larger, the graph becomes flatter and does not represent the data set well.

Please take a look at our code.

2. d.)

```
In [44]: #2d)
         #apply model to synth data / womens data
         womens100data = np.loadtxt(fname = "http://colindawson.net/data/womens100.csv", delimiter = (","))
         x = np.reshape(womens100data[:, 0], (-1, 1))
         x = x.reshape(-1, 1)
         y = np.reshape(womens100data[:, 1], (-1, 1))


         synthdata2016 = np.loadtxt(fname = "http://colindawson.net/data/synthdata2016.csv", delimiter = (","))
         synthdata2016x = np.reshape(synthdata2016[:, 0], (-1, 1))
         synthdata2016y = np.reshape(synthdata2016[:, 1], (-1, 1))


         print(bestpolynomialorder(y, x, D = 20, k = 10, seed = 1))
         #found optimal polynomial is 1 for k = 10
         print(bestpolynomialorder(y, x, D = 20, k = 19, seed = 1))
         #found optimal polynomial is 1 for k = N
```

```
         [1.        2.42297678]
         [1.        0.34801739]
```

```
In [45]: print(bestpolynomialorder(synthdata2016y, synthdata2016x, D = 10, k = 10, seed = 1))
         #found optimal polynomial is 2 for k = 10
         print(bestpolynomialorder(synthdata2016y, synthdata2016x, D = 10, k = 50, seed = 1))
         #found optimal polynomial is 2 for k = N
```

```
         [2.        5.47387477]
         [2.        0.99045119]
```

For Womens100 data, we found that the optimal polynomial order is 1 for both K=10 and K = N.

For synthdata2016 data, we found that the optimal polynomial order is 2 for both K=10 and

K=N. We were unable to graph the errors since we couldn't figure out how to do it.

We tried but were not able to complete 2.e and 2.f, but all of our functions works just fine. Please

take a look at our code.