Suryank Tiwari MT19019

**REPORT**

**Question 1**

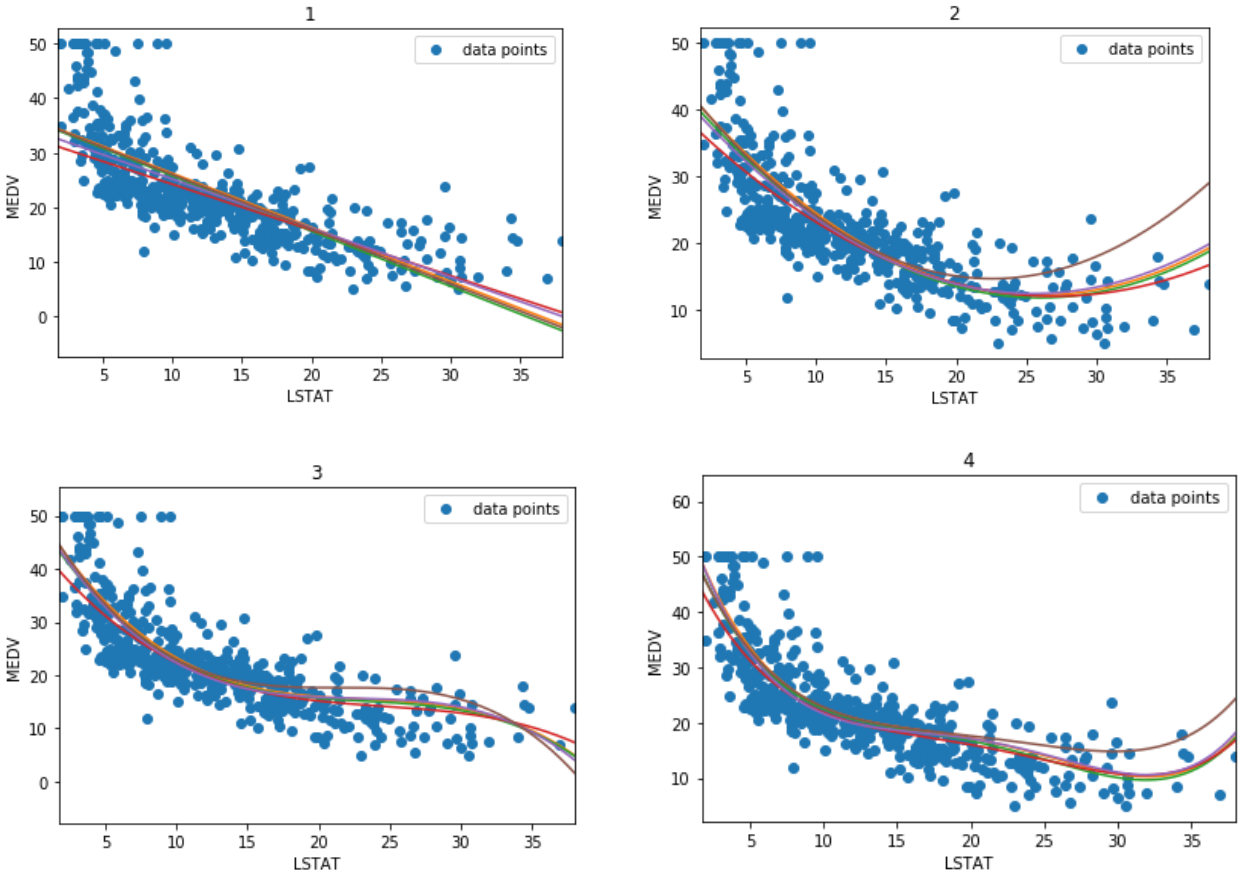
Performed Linear Regression on all features and computed the RMSE for training and testing set. Following are the data values observed.

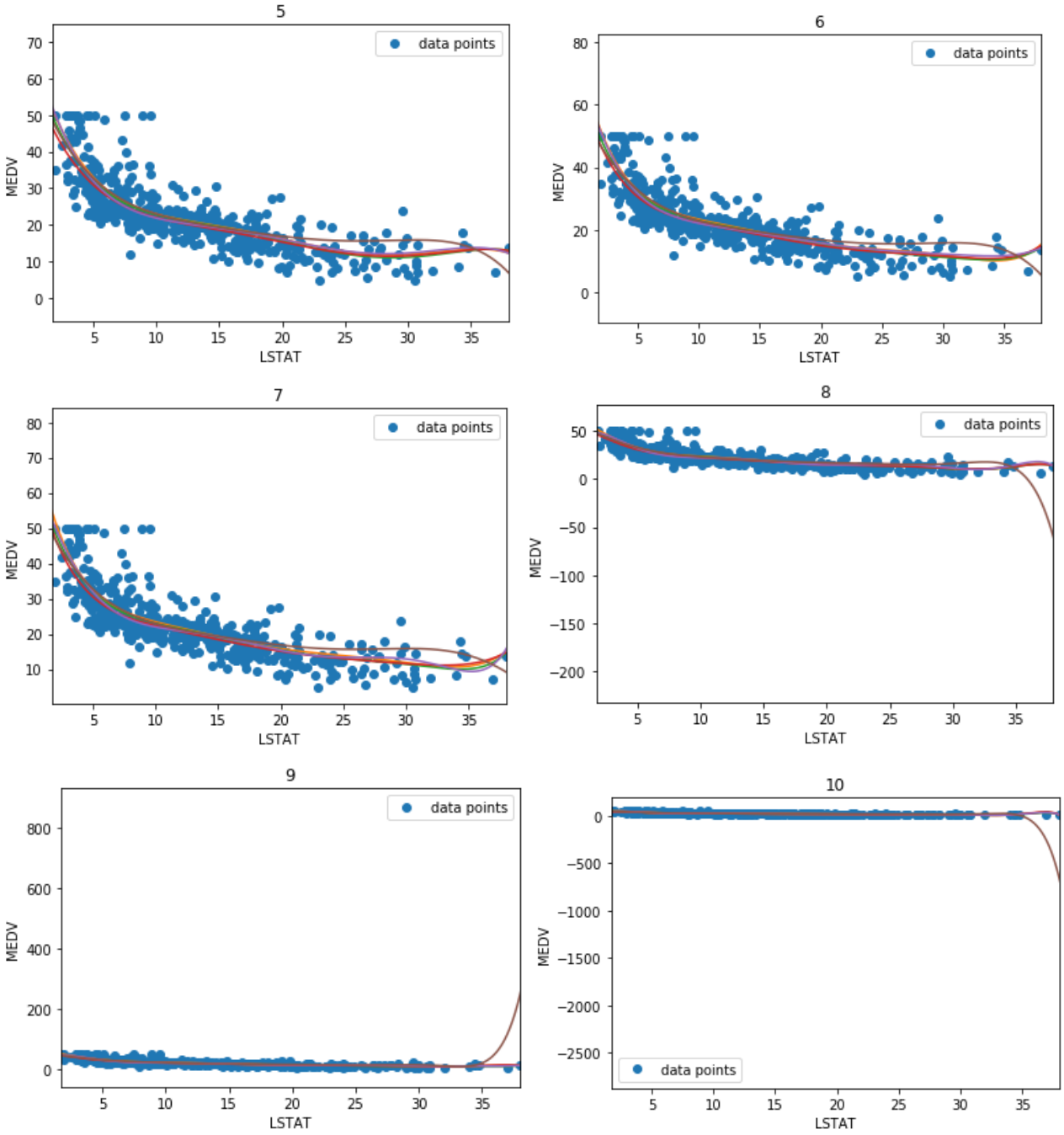
```
Weights learned:
 [[ 3.01834801e+01]
 [-1.94651664e-01]
 [ 4.40677436e-02]
 [ 5.21447706e-02]
 [ 1.88823450e+00]
 [-1.49475195e+01]
 [ 4.76119492e+00]
 [ 2.62339333e-03]
 [-1.30091291e+00]
 [ 4.60230476e-01]
 [-1.55731325e-02]
 [-8.11248033e-01]
 [-2.18154708e-03]
 [-5.31513940e-01]]
RMSE for test 5.771794050734153
RMSE for train 4.768352896282809
```

Where the weights learned are [w0, w1, …, w13] respectively.

For the feature LSTAT to predict MEDV:

Corresponding to degree

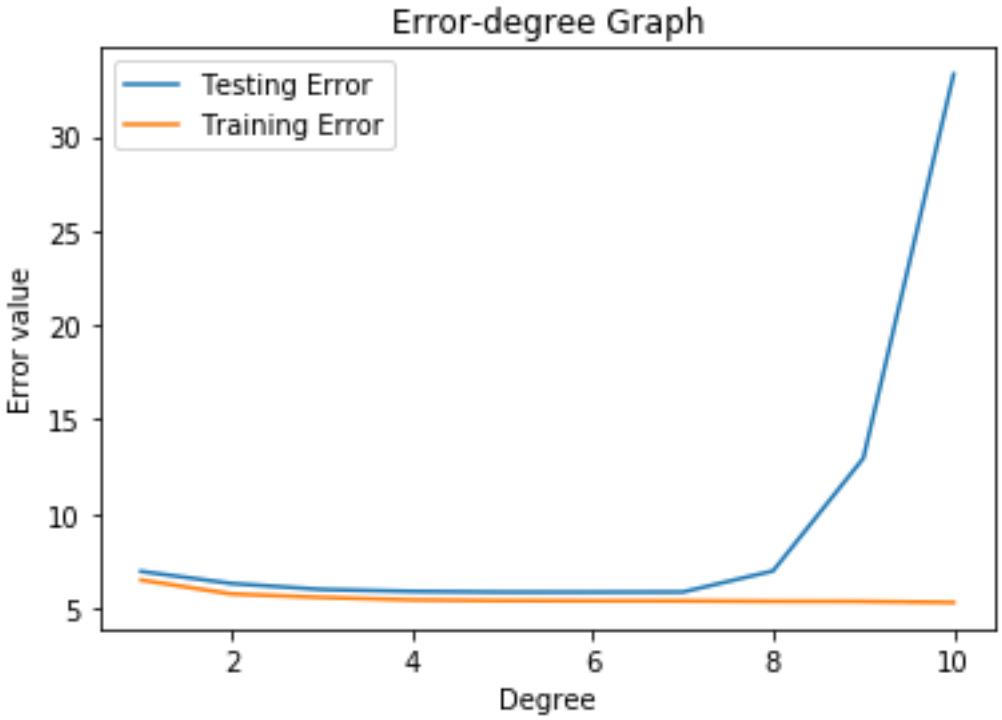





RMSE for various polynomial degrees via k fold:

```
RMSE Errors - Testing, Training, Degree
(6.939674798252203, 6.4823076757826605, 1)
(6.2987106277271545, 5.746153611018438, 2)
(5.981604012291513, 5.559194801663524, 3)
(5.8795679632337485, 5.427617578695064, 4)
(5.8426846824142284, 5.3898288184895735, 5)
(5.839740616634016, 5.377898415313588, 6)
(5.845373581803588, 5.375416063243268, 7)
(6.964162288320744, 5.35299025522938, 8)
(12.943879668163905, 5.346373550948727, 9)
(33.30702697817594, 5.2915308881885545, 10)
Best Polynomial Degree Fit: 6
```

In the RMSE for various polynomial degrees via k fold we can see that testing error decreases at first and then shoots up. This is because at higher degrees the models start to overfit and reduce the training error.




The graph for testing error vs degree of polynomial.

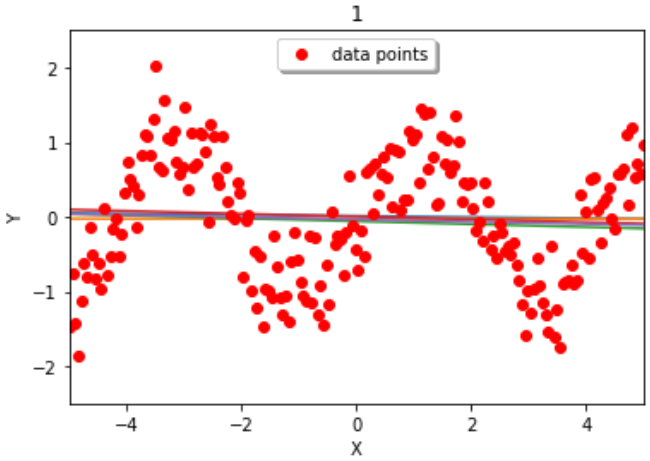
80% dataset train and then 20% test data RMSE:

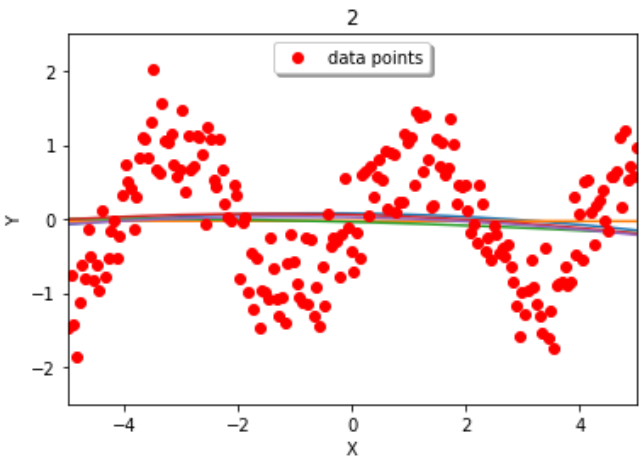
```
RMSE of test dataset:  3.893891859843845
```

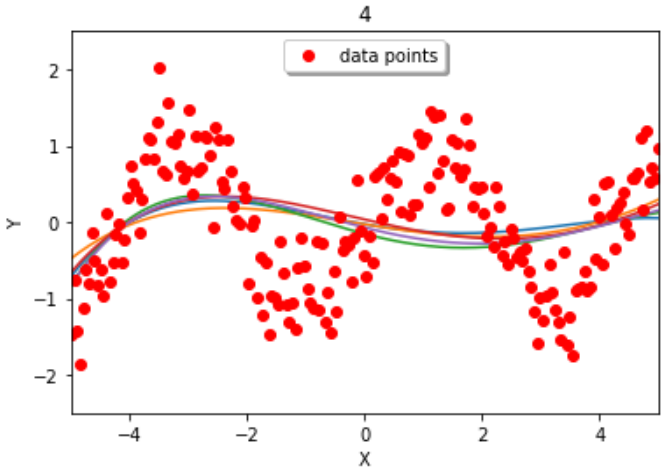
Suryank Tiwari MT19019

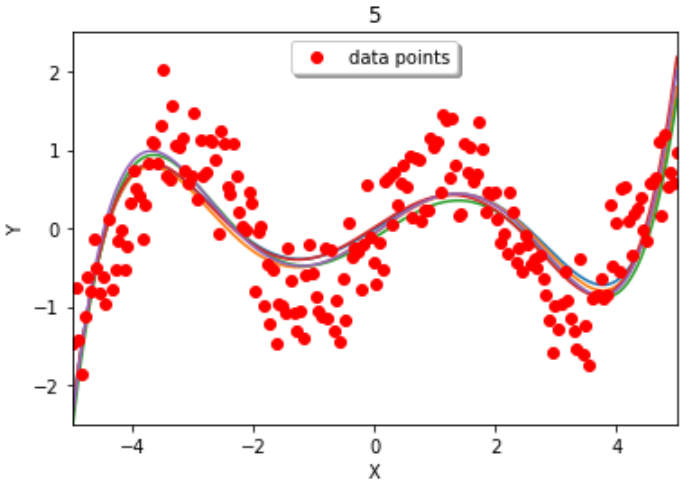
**Question 2**

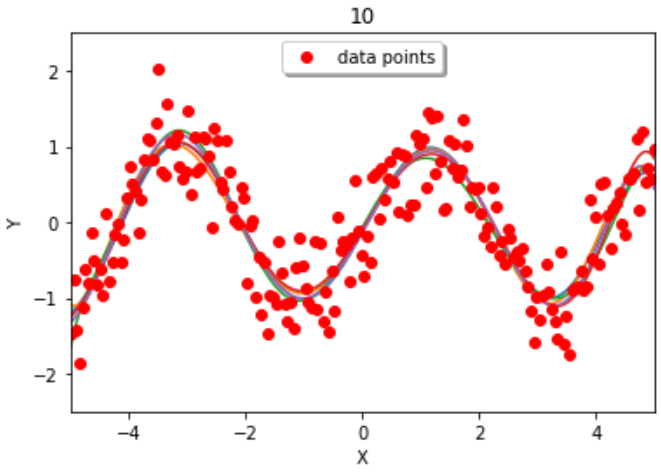
Following are the graphs for data points vs the polynomials observed during k folds. Image titles are the degrees of the polynomials

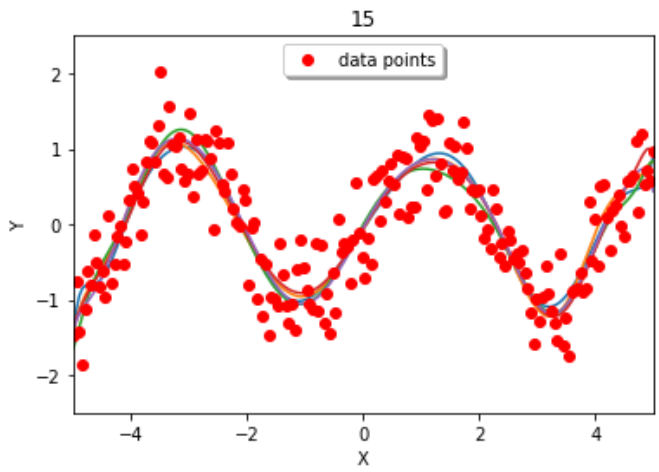


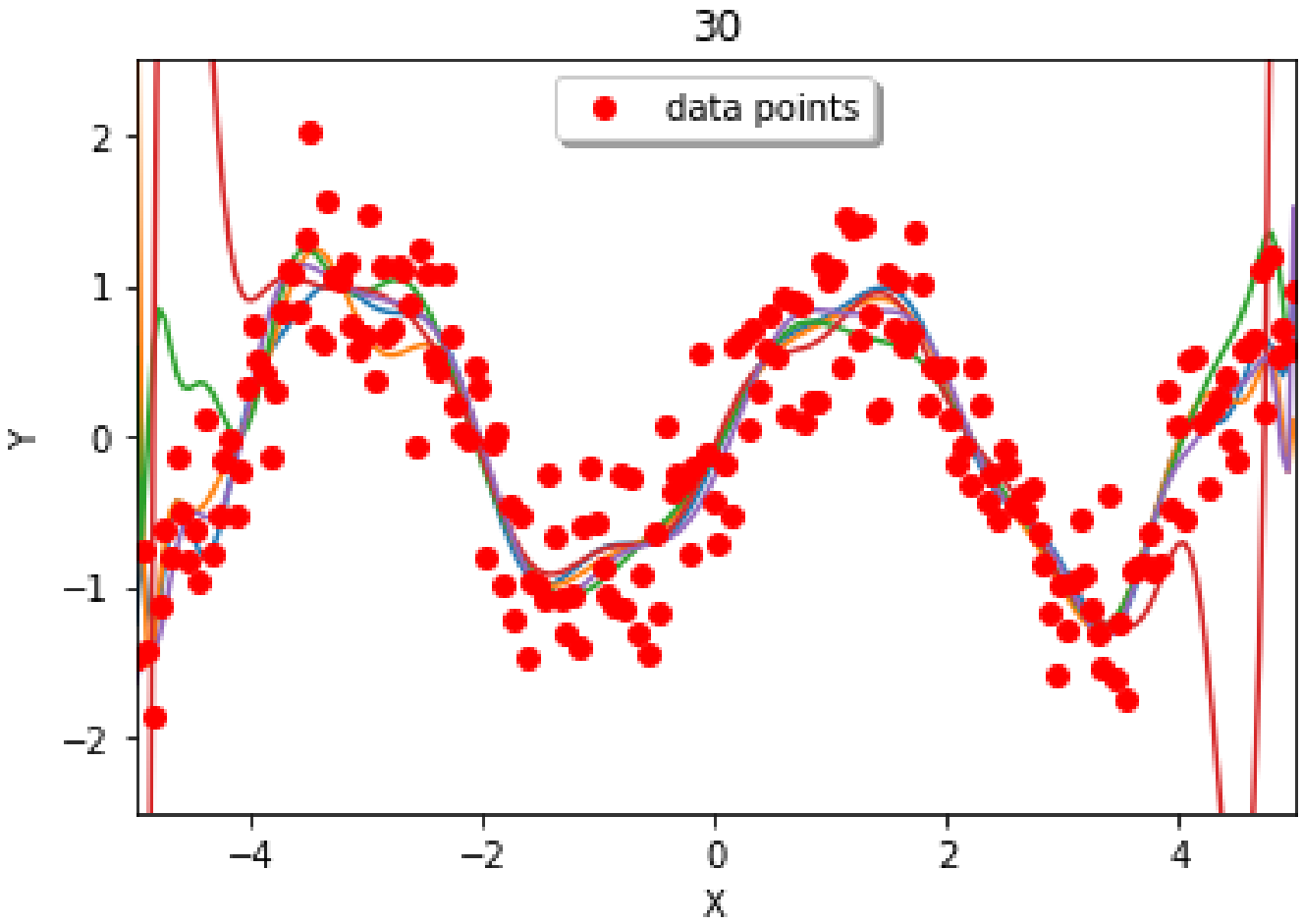






We can clearly see that degree 30 polynomial is overfitting, whereas degree 1-4 models are too simple. Degree 10 looks to be the optimal value here.

RMSE values for the given degrees are as follows:

```
(0.8246939684048611, 1)
(0.8265014922025485, 2)
(0.8100678517175516, 4)
(0.5989746623221318, 5)
(0.42607804600187166, 10)
(0.4287250609431146, 15)
(2.3883349499418682, 30)
Best Polynomial Degree Fit: 10
```

Following observations are for k fold testing error for degree 1 – 30:

```
(0.8246939684048611, 1)
(0.8265014922025485, 2)
(0.808407491657036, 3)
(0.8100678517175516, 4)
(0.5989746623221318, 5)
(0.6075029273265181, 6)
(0.420129388116409, 7)
(0.4212174271845108, 8)
(0.41868638086950843, 9)
(0.42607804600187166, 10)
(0.4224259037689761, 11)
(0.42094596882868285, 12)
(0.4209962154002931, 13)
(0.42320698079027547, 14)
(0.4287250609431146, 15)
(0.43248928691167066, 16)
(0.4487209292090131, 17)
(0.4570224603157642, 18)
(0.460219692896, 19)
(0.4772390590455179, 20)
(0.4413318734914227, 21)
(0.4543796431658548, 22)
(0.5371976050154249, 23)
(0.65151670506443, 24)
(0.7434820942977372, 25)
(1.0498060639878513, 26)
(0.6254600874590492, 27)
Best Polynomial Degree Fit: 9
```

Whereas, overfitting can be shown clearly without k fold. If all data is used for training and testing

```
2 134.05419478619604
3 133.82803214598255
4 126.29399419785213
5 125.5567933119209
6 65.49212283762637
7 65.38133807926202
8 31.86484522595084
9 31.68284702334479
10 30.641679353573306
11 30.633402535417968
12 30.182509437765955
13 30.023663514982754
14 30.02115150078822
15 29.893100991153155
16 29.875381223796875
17 29.865082751733876
18 29.21476003136461
19 28.544959381831266
20 28.536565101169316
```

i.e. error only decreases