Admission prediction model

Problem statement:

To try and estimate the chances of admission into universities on the basis of given parameters:

- GRE Score
- TOEFL Score
- University Rating
- Statement of Purpose
- Letter of Recommendation
- CGPA
- Research

Approach

The approach is to design a model that fits the given chance of admission with the given parameters. Then find the estimated chance with that model and finally find the average of the error found. We try to minimize the average error so that the model predicts the chances as closely as possible.

Model

Linear model

This model takes the relation between the parameters and the chance of admission to be linear:

The function used is:

```
def func(fdata, A1, A2, A3, A4, A5, A6, A7):
   fgre = fdata[0] / 340
   ftoefl = fdata[1] / 120
   funi_rat = fdata[2] / 5
   fsop = fdata[3] / 5
   flor = fdata[4] / 5
   fcgpa = fdata[5] / 10
   frsch = fdata[6]
   return A1*fgre + A2*ftoefl + A3*funi_rat + A4*fsop + A5*flor + A6*fcgpa + A7*frsch
```

The estimated values found were:

- A1 = 5.3746049657144935e-17
- A2 = 2.7948158718101004e-15

- A3 = 0.1291412246831608
- A4 = 0.0924330560639533
- A5 = 0.1127690166388799
- A6 = 0.5459748244712413
- A7 = 0.0609979487532703

The estimated error is : 8.805409933343526 %

Polynomial model

The model is a polynomial function with the coefficients as well as the powers of the parameters are estimated using curvefit.

The function used is :

```
def func(fdata, A1, A2, A3, A4, A5, A6, A7, P1, P2, P3, P4, P5, P6, P7):
    fgre = fdata[0] / 340
    ftoefl = fdata[1] / 120
    funi_rat = fdata[2] / 5
    fsop = fdata[3] / 5
    flor = fdata[4] / 5
    fcgpa = fdata[5] / 10
    frsch = fdata[6]
    return A1*fgre**P1 + A2*ftoefl**P2 + A3*funi_rat**P3 + A4*fsop**P4 + A5*flor**P5 + A6*fcgpa**
```

The estimated values we found were:

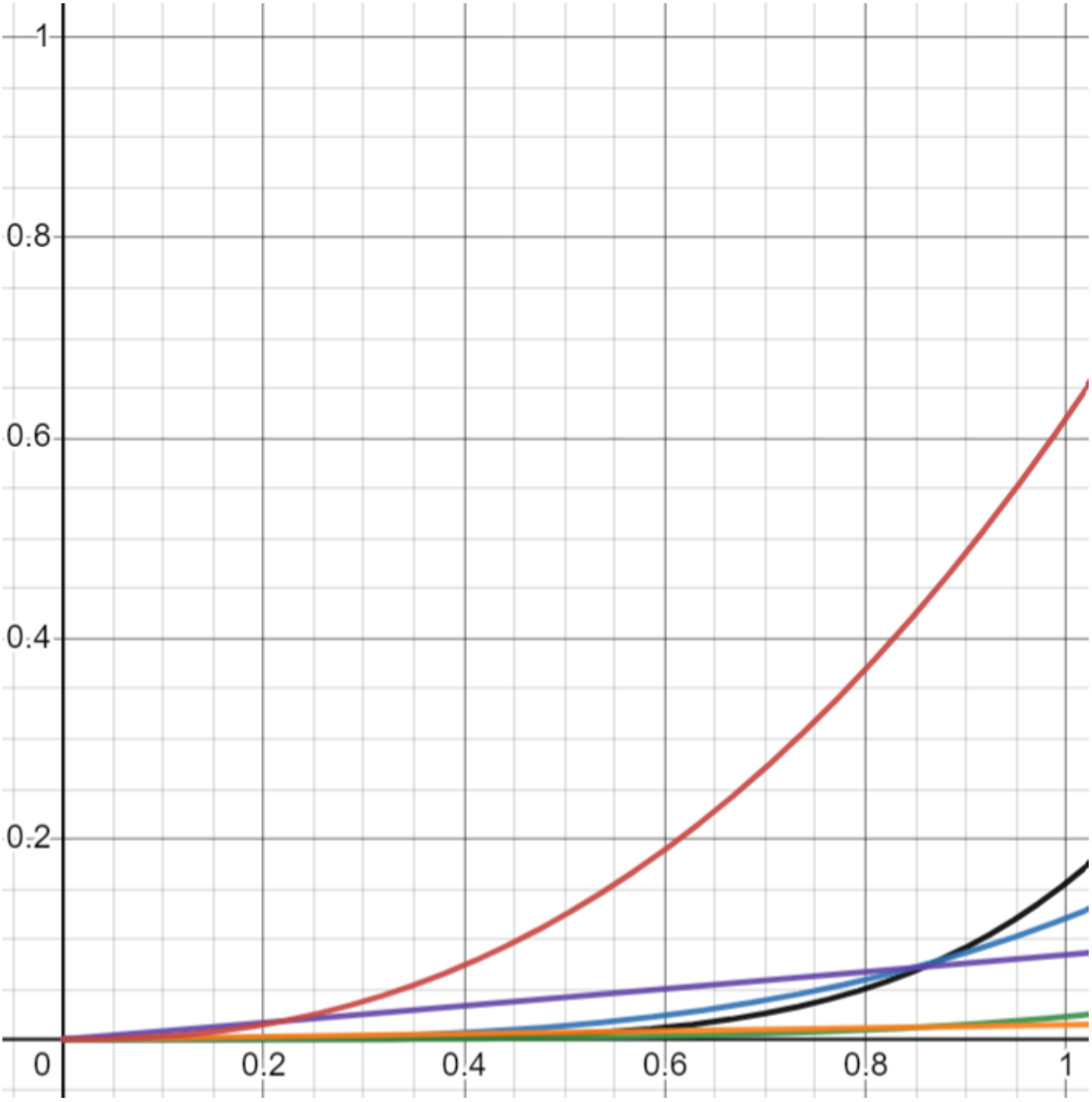
- A1 = 0.1561405188755400
- A2 = 0.1209068548723192
- A3 = 0.0226842870767161
- A4 = 0.0144438164821258
- A5 = 0.0842272781540858
- A6 = 0.6201563759289992
- A7 = 0.0256727728580029

- p1 = 5.039594899882378
- p2 = 3.173849737479317
- p3 = 3.974647398691436
- p4 = 1.068632363440998
- p5 = 1.011094671881939
- p6 = 2.321597238695864
- p7 = 5.0

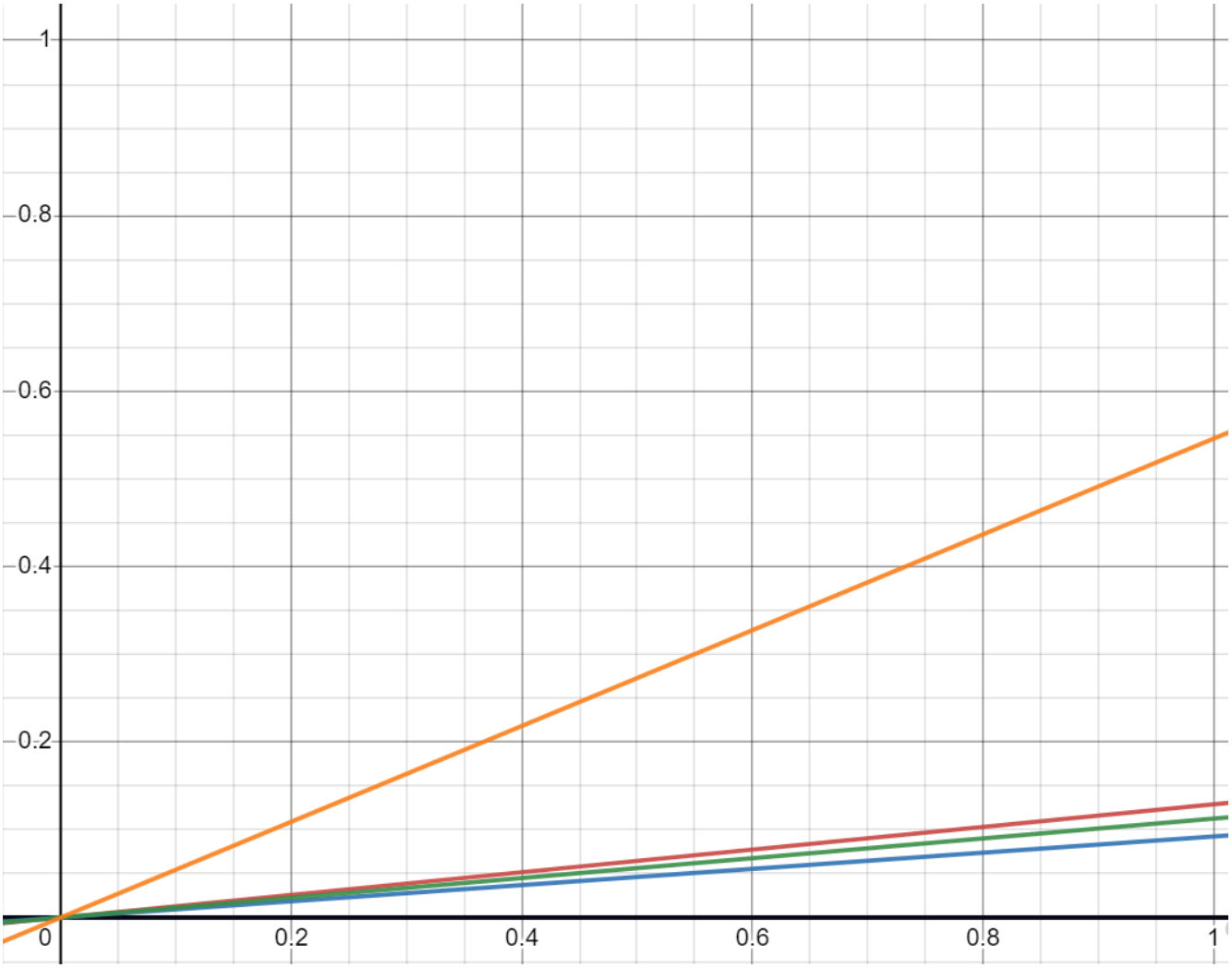
The error I got was: 6.993848763331926 %

Inferences

All the values given for each parameter were first normalized and then operated upon. From the image below, we can make the following conclusions:



- The image depicts the individual contributions of each of the parameters in the `Polynimial Model` .
- We can see that University Rating, SOP and LOR are the parameters on which there is the least dependency.
- CGPA, as expected, matters the most.



- The image depicts the individual contributions of each of the parameters in the `Linear Model` .
- For the linear model, we can directly see the contribution by the slope.
- It is clear that CGPA has the highest contribution. Whereas, GRE and TOEFL scores contribute negligibly.