Assignment - 4

SHAH DHRUVANSHI KAMLESH

EE22B143

We are given a dataset containing Chances of Admission to a particular university along with various parameters like CGPA, GRE Score, TOEFL Score affecting it. To form a function which predicts the chances given the parameters, I am using `curve_fit` to fit a polynomial function(depending on parameters).

Initially, I am normalising the input parameters to a scale of 1, as then a change in the normalised parameters represents %change in the actual parameters, and since our parameters have values on different scale, this step normalises them.

I am then defining a function that returns the polynomial function that we want to fit to the given data, then using `curve_fit`, I am obtaining the optimal parameters that will fit the curve the best and also calculating the standard deviation and correlation coefficient.

Note that since we have normalised the parameters, we can exploit the fact that any function can be represented by Taylor's polynomial near origin and can safely assume a polynomial function as the error near origin will be minimal.

After obtaining the parameters, I am evaluating the correlation coefficient(of the obtained parameter, with the given chances of admission) as well as the partial derivative of the function averaged over its values.

> Correlation coefficient of estimated function and given data = 0.906988649618902
>
> Correlation coefficient with CGPA = 0.8824125749045744
>
> Average partial derivative for CGPA = 1.1701140337363927
>
> Correlation coefficient with GRE Score = 0.8103506354632598
>
> Average partial derivative for GRE Score = 0.6450301193777527
>
> Correlation coefficient with TOEFL Score = 0.792227614305083
>
> Average partial derivative for TOEFL Score = 0.36200664700576074
>
> Correlation coefficient with University Ranking = 0.6901323687886891
>
> Average partial derivative for University Ranking = 0.011712473038934353
>
> Correlation coefficient with SOP = 0.6841365241316721

Average partial derivative for SOP = 0.03873425483284882

Correlation coefficient with LOR = 0.6453645135280108

Average partial derivative for LOR = 0.08399694300319197

Correlation coefficient with Research = 0.5458710294711376

Average partial derivative for Research = 0.0255002494270167

On comparing the the correlation coefficient and partial derivative(which gives the rate of change of a fixed parameter other parameters kept constant), we can judge the strength of effect of each parameter on the function as follows.

(1): CGPA

(2): GRE Score

(3): TOEFL Score

(4): LOR

(5): SOP

(6): Research

(7): University Ranking

I have also attempted to the problem assuming a linear function, the final result of which does not differ much from polynommial case. However, the standard deviation obtained from linear case is greater smaller than that obtained from polynomial case, which gives results as belows, but the order almost remains same like CGPA is still having highest effect:

Correlation coefficient = 0.9065874141624943

The effect of various parameters on Chances of admission are as follows:

(1) GRE Score: 0.6318918845899039

(2) TOEFL Score: 0.333356827569506

(3) University Ranking: 0.029706788113846812

(4) SOP: 0.00793071846280193

(5) LOR: 0.08429369970344631

(6) CGPA: 1.183850658336586

(7) Research: 0.024307480512715193