EE2703 : Applied Programming Lab
Assignment 3
Fitting Data to Models

Adityan CG//EE19B003

March 3, 2021

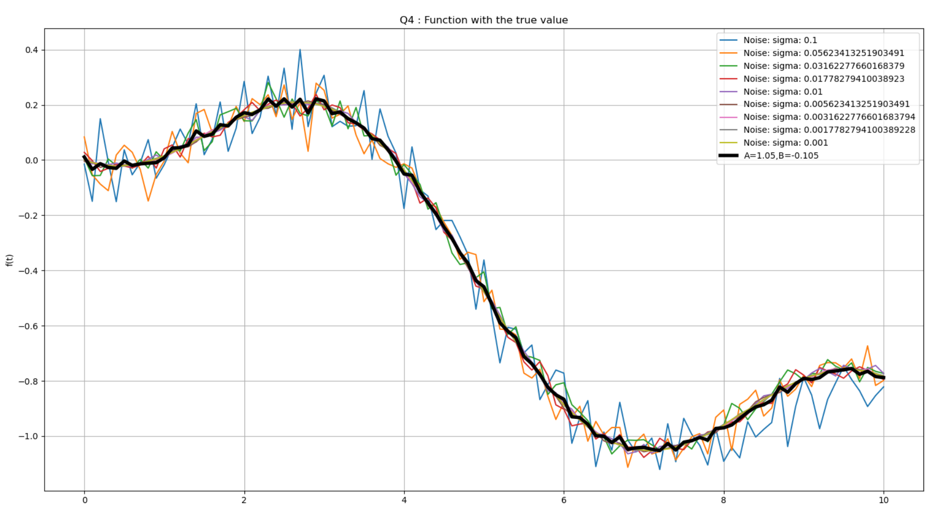
1 Introduction

This Assignment is about Organising data and Looking into the Errors of a bessel Function From "SCIPY" Lab module. A Data file with different Noise added to the original Function. The Function is Given by,

$$f(t) = A * J2(t) + B * t + n(t)$$

2 Generated Data plot and Original Value

The given data is of 10 columns where 9 correspond to the functions with different "sigma" (Standard Deviation of Noise)of noise. As expected in Question 4 Plot of the all the data points is given below and the True value plot is obtained by using the **mean** of all data points.

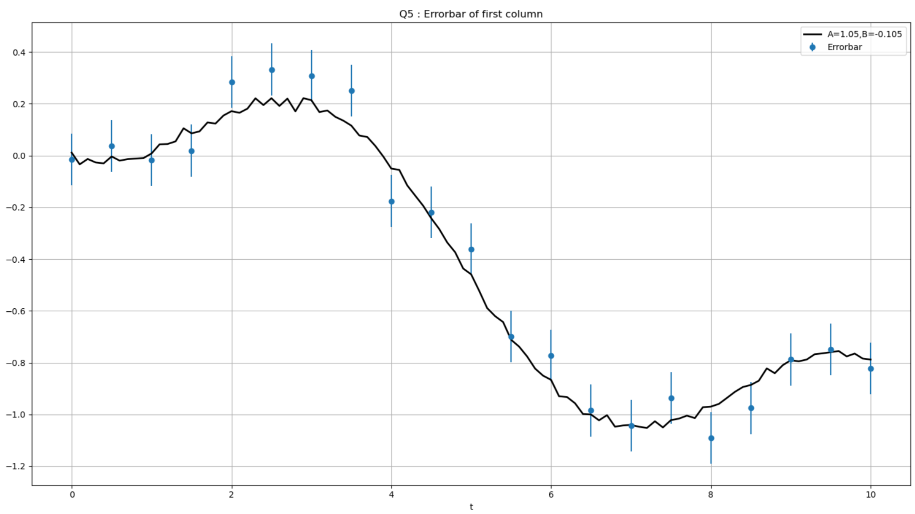


3 Error bar plot of data in 1st Column

The error bar is plotted with x as time in 1st column of data frame is plotted along with the True value plot to check how much the noise (sigma[0]) deviated the True value plot. $n(t)$ is the normally distributed noise given by

$$\mathrm{P}(n(t)|\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{n(t)^2}{2\sigma^2}\right)$$

where σ is generated using `sigma=logspace(-1,-3,9)`



4 Matrix Approach and Mean squared error

The problem at hand can also be converted to a matrix form, where the coefficients is the Matrix **p** and the two different functions of t such as **J2 and t** is defined into other matrix **M**

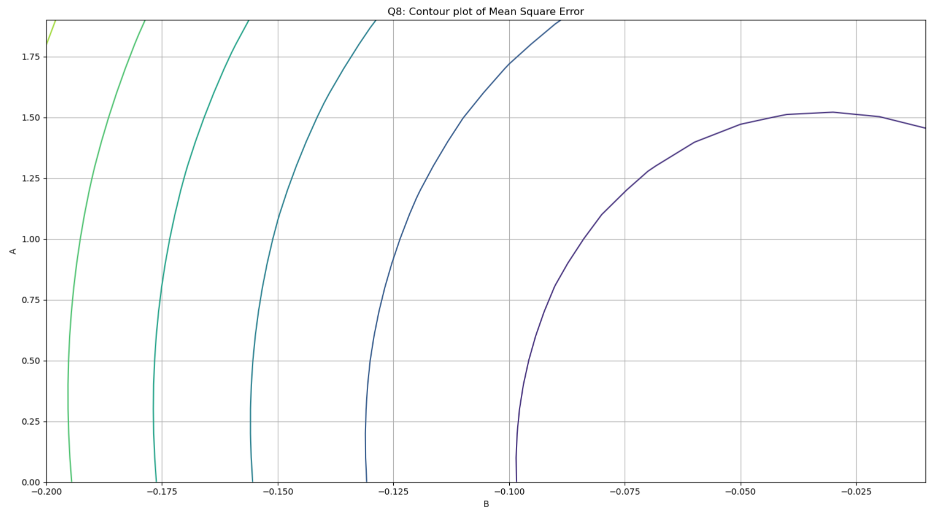
$$M = \begin{pmatrix} J_2(t_1) & t_1 \\ \dots & \dots \\ J_2(t_m) & t_m \end{pmatrix}$$

$$p = \begin{pmatrix} A \\ B \end{pmatrix}$$

The product of the Matrices Mp is defined as a function g and as we traverse on with $A = 0, 0.1, ..., 2$ and $B = -0.2, -0.19, ...0$, the **mean squared error** is calculated between the data f_k and the assumed model using the formula

$$\epsilon_{ij} = \frac{1}{101}\sum_{k=0}^{101}\left(f_k - g(t; A, B)\right)^2 \tag{1}$$

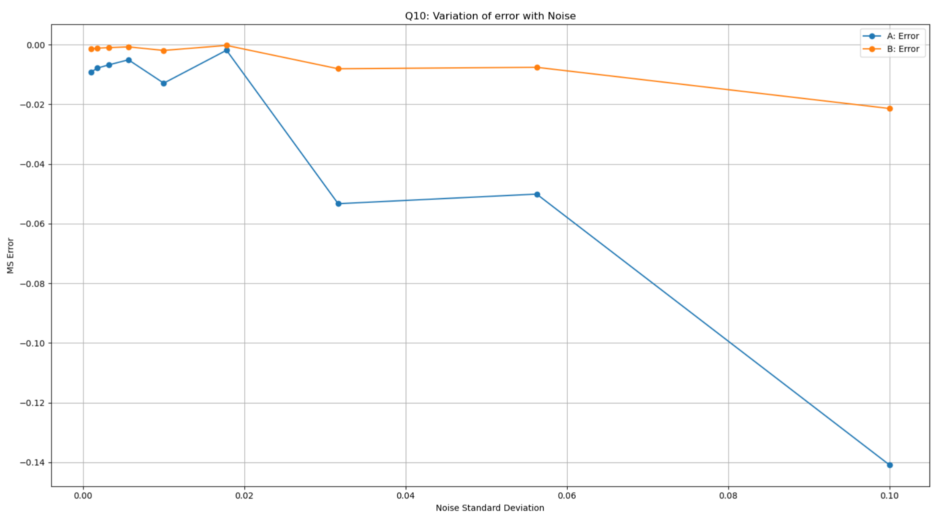
The contour plot by varying A and B is provided, the code also prints the minimum.



Also the best estimate of A and B are also calculated by using the function `np.linalg.lstsq` as in,

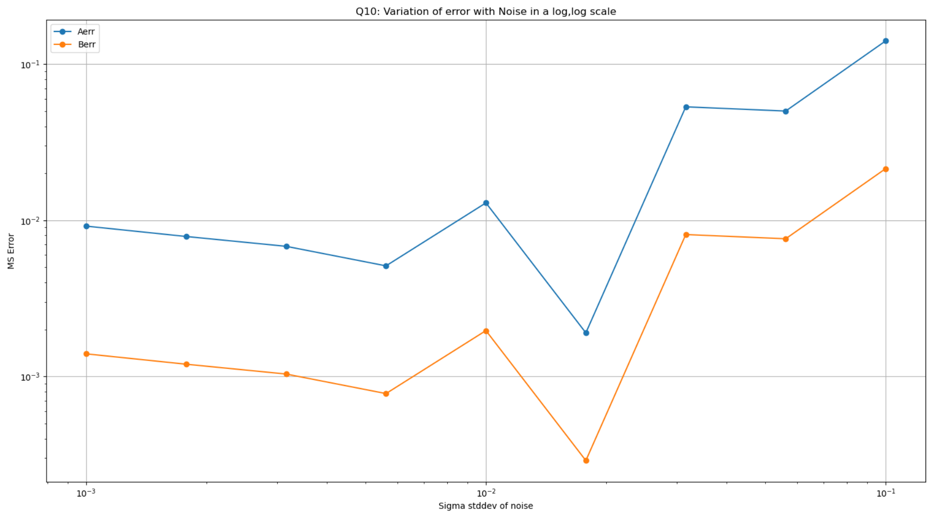
`np.linalg.lstsq(M, cols[:,i],rcond=-1)`

Now we can see the relation **Mean squared error vs Standard Deviation of Noise**. We notice that it is not a linear relation



But we also notice that the error has a linear relation with sigma on a loglog scale.

- The error has an exponential behaviour with the change in noise distribution



5 Conclusion

A best possible fit is observed for the data obtained from the file with noises by minimizing the least mean square error. It is observed that the mean square error is directly proportional to the σ of the noise in the log scale