EE2703 : Applied Programming Lab
Assignment 3

S. Uttam Raj
EE20B118

February 2022

1 Abstract

This assignment will focus on the following topics:

- Reading data from files and parsing them
- Analysing the data to extract information
- Study the effect of noise on the fitting process
- Plotting graphs

2 Introduction

The first step in this task is to generate data using a linear combination of Bessel function and y = x. Noise is added in various levels to this mix.

$$f(t) = A * J_2(t) - B * t + n(t)$$

In the above equation:

$$J_2(t) = Besselfunction$$

$$n(t) = Noisefunction$$

A = 1.05 and B = -0.105

We need to study the relationship between the standard deviation of the noise introduced and the error of our A and B estimations.

3 Assignment questions

3.1 Generation and loading Of Data

Data in the form of a file *fitting.dat* is generated using the code *generate_data.py* provided. This data is loaded into our program.
Using numpy.loadtxt function we import data from *fitting.dat* into 10 different columns.

3.2 Visualizing the data

The first column was time, and the other nine columns were each riddled with varying amounts of noises, with standard deviation taken uniformly from a logarithmic scale. The plot below was created by showing the true function's value along with all 9 noise added values.

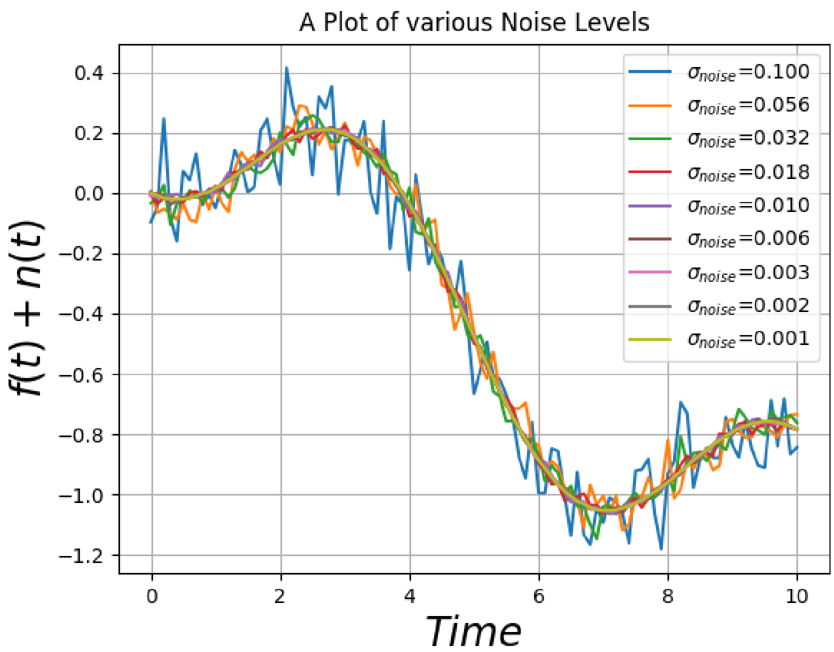

Figure 1: True and Noise plots

3.3 The Errorbar plot

Another view of how the noise affects the data can be seen using the Errorbar() function.Errorbars for the first data column of every 5^{th} data point with the original data are plotted.

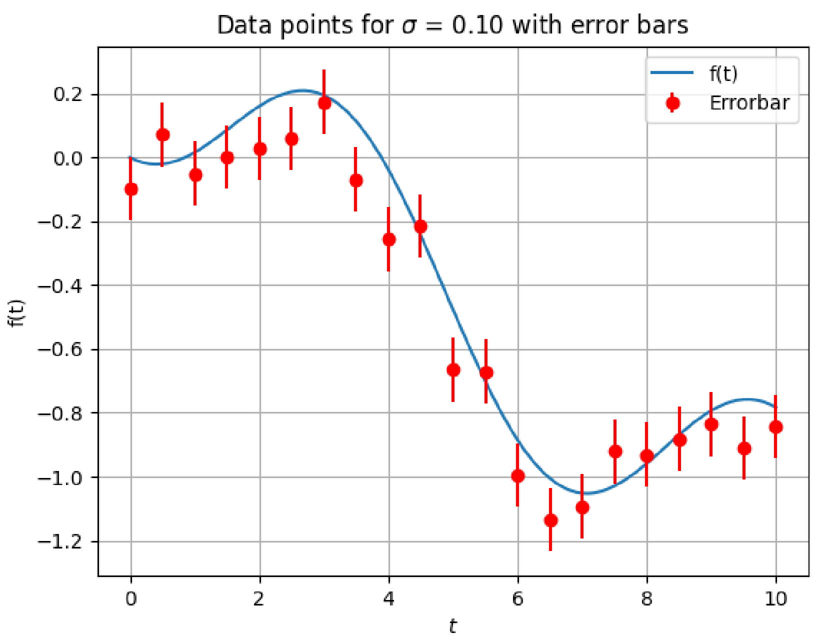

Figure 2: Noisy Data with Errorbar

3.4 Constructing Matrix

The matrix M which is $\begin{bmatrix} J_2(t_1) & t_1 \\ J_2(t_2) & t_2 \\ \dots & \dots \\ J_2(t_m) & t_m \end{bmatrix}$ created when multiplied with p that is $\begin{bmatrix} A \\ B \end{bmatrix}$ matrix will give rise to the actual function. The results of matrix multiplication and the user-defined function can be compared using numpy.array_equal() function. They are equal if this returns a True value.

Using the lstsq function in scipy package, we solve for:

$$M.p = F$$

where F is $\begin{bmatrix} f_(t_1) \\ f_(t_2) \\ \dots \\ f_(t_m) \end{bmatrix}$

3.5 Error computation

The mean squared error measures the error between noisy and correct functional data. The formula for error computation is as follows:

$$\epsilon_{ij} = \frac{1}{101}\sum_{k=0}^{101}(f_k - g(t_k, A_i, B_j))^2$$

where f_k is the $(k+1)^{th}$ column of data Contour plot of ϵ_{ij} is shown below:

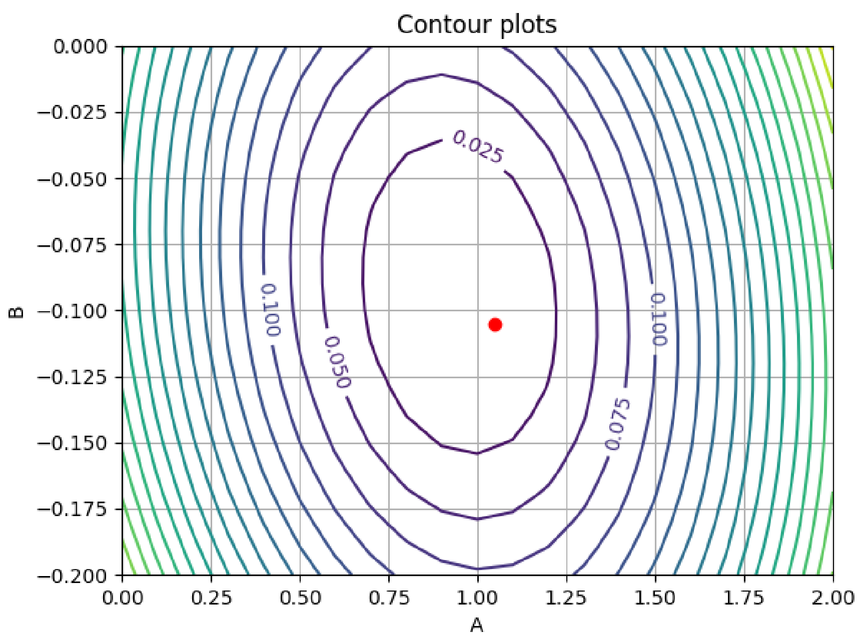

Figure 3: Contour Plot

From the plot, we can say that the contour plot has only one minimum marked in red colour and that occurs at A = 1.10 and B = -0.105

3.6 A,B parameters estimation

We solve the different values of σ_{noise} by changing D matrix to different columns of *fitting.dat*.Mean squared error in the estimate of A and B for different amounts of noises versus the standard deviation (σ_{noise}) is as follows:

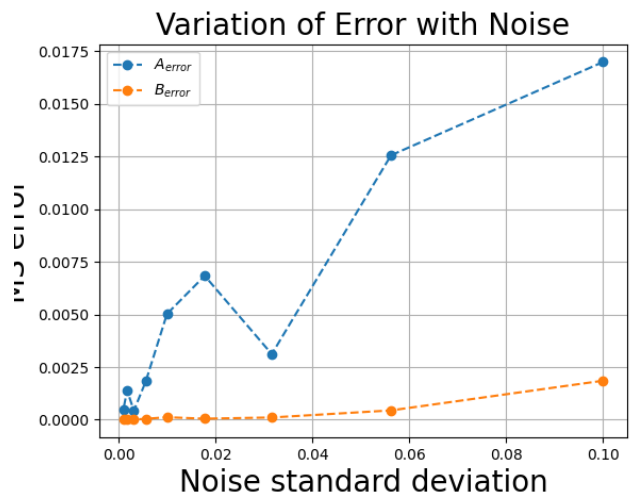

Figure 4: Mean Squared Error vs Standard Deviation

The above plot is on a linear scale. The error in the estimate is not growing linearly with the noise. However when we plot it on loglog scale, it is almost linear.

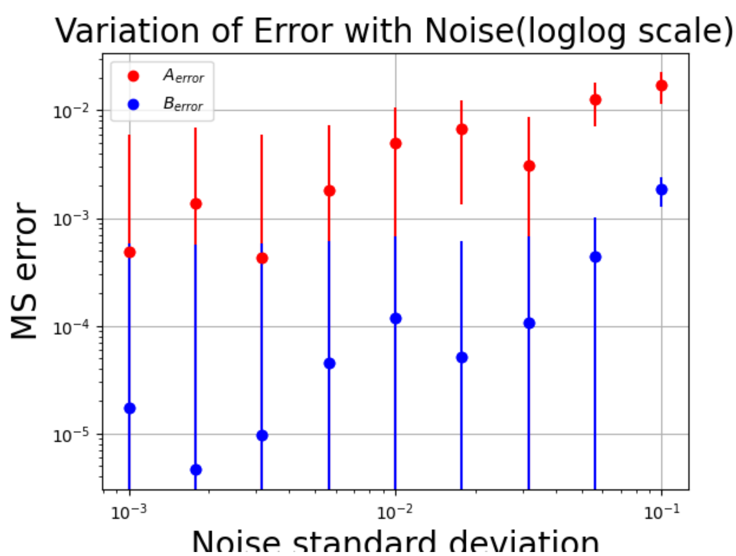

Figure 5: Error vs Standard Deviation loglog Plot

4 Conclusion

We were able to deduce from the preceding technique that as the noise in the data used to estimate the linear combination increases, the error of the prediction also increases on a linearly on a loglog scale.