APL(EE2703): Assignment 3

Santosh G (EE19B055)

March 5, 2021

1 Aim of the Assignment:

- Observe the error in fitting the *Least Error Fit* function to a given set of data.
- Finding the relation between the error observed and the noise in the data.

2 Introduction

Assuming the model to be a linear one, we can model the data as following:

$$f(t; p_1, p_2, .., p_n) = \sum_{i=0}^{n} p_i F_i(t) \tag{1}$$

is the function to be estimated, $F_i(t)$ are arbitrary functions of the variable t and p_i are constant parameters.

We can then find have to a function $F(t)$, such that

$$F.\vec{p} = \vec{a_0} \tag{2}$$

while the actual value is 'a'.

Te difference in these values can be termed as noise(n).

$$\vec{a_0} + \vec{n} = \vec{a} \tag{3}$$

The final function $F(t)$, can be chosen through various methods, but here we shall choose the one with least sum of squares of the error.
the Error is given by:

$$\epsilon = F.\vec{p} - \vec{a} \tag{4}$$

We shall chose a functions such that, total error is minimum where,

$$TotalError = \sum_{i} \epsilon_i^2 \tag{5}$$

Note: Noise is assumed to have zero mean and a standard deviation of σ.

we shall be trying to fit the following function:

$$f(t) = 1.05 J_2(t) - 0.105t \tag{6}$$

where $J_2(t)$ is the *Bessel Function of the first kind of Order 2*. This equation is used to obtain the actual data.

3 Creating noisy data

Random noise is added to $f(t)$, to create data. This random noise, denoted by $n(t)$, is generated using the normal probability distribution:

$$P(n(t)|\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n(t)^2}{2\sigma^2}} \tag{7}$$

The addition of noise will result in data taking the following form:

$$f(t) = 1.05J_2(t) - 0.105t + n_{\sigma_i}(t) \tag{8}$$

where, $n_{\sigma_i}(t)$ is the noise function with $\sigma = \sigma_i$ in 10. Thus for 9 different values of sigma (in a log scale from 0.001 to 0.1), the noise is created and stored in the `fitting.dat` file.

4 Questions 3 and 4

The data generated from the noise and the true data are plotted using PyPlot. The output result looks as follows:

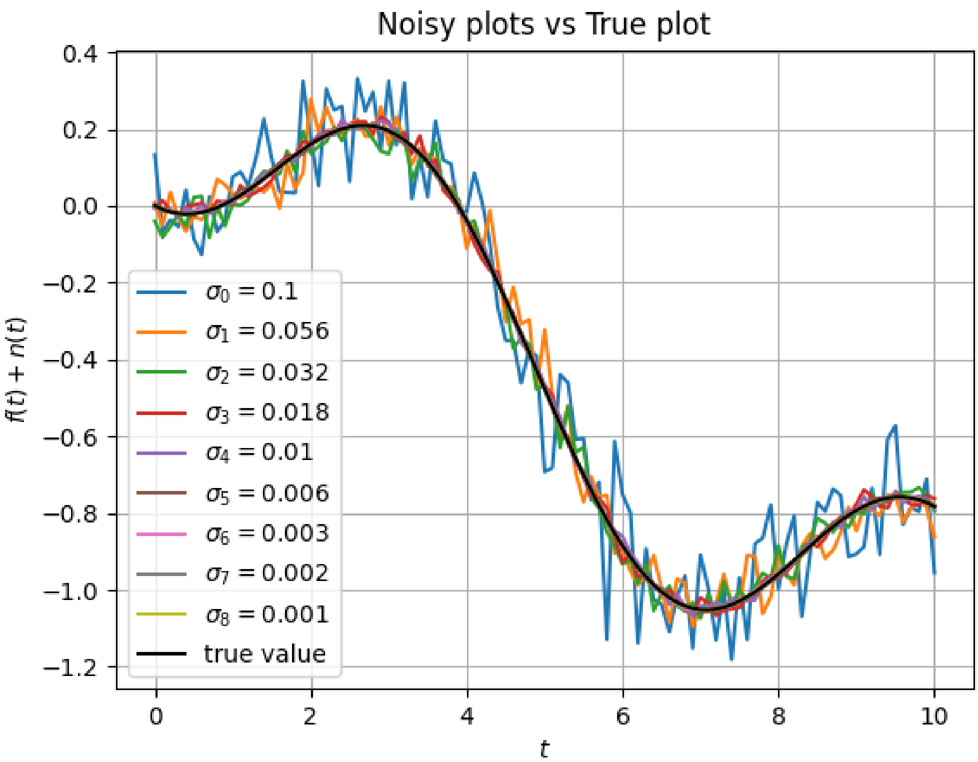

Figure 1: Noisy Data and True Data (Plot for Q3,Q4)

From the above figure it is clear that the noise(deviation from true value) in increasing with increasing standard deviation σ.

5 Question 5

The graph is plotted using the true data and sampling a few points from the noisy data with fixed standard deviation $\sigma = 0.10$.

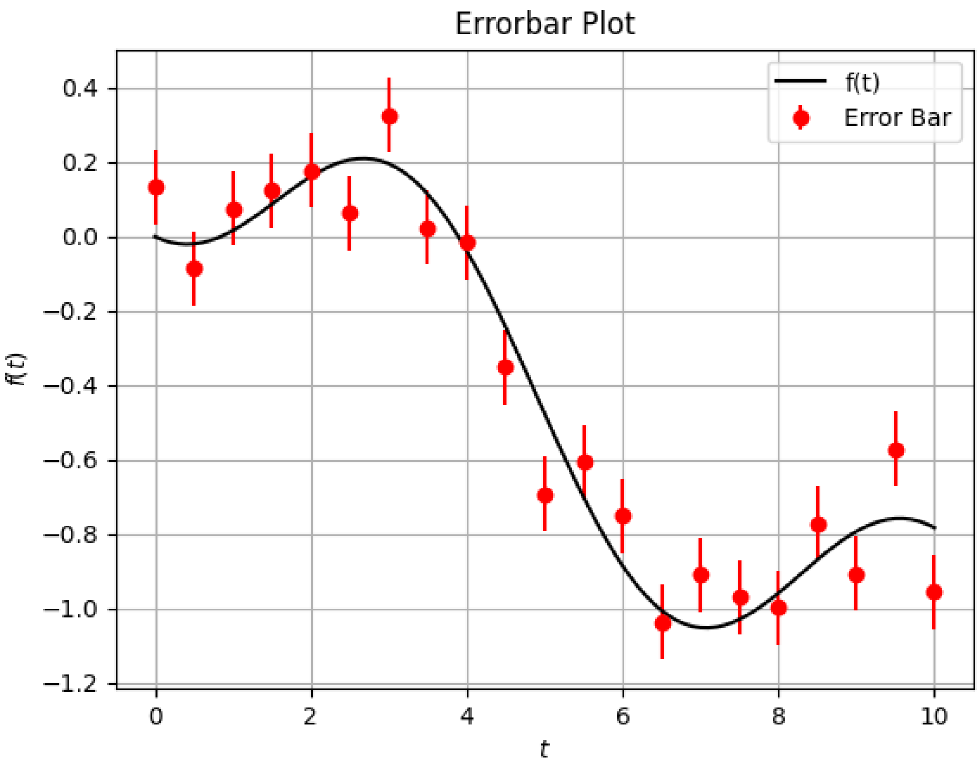

Figure 2: True Data and sampled Noisy Data (Plot for Q5)

6 Question 6

The function array_equal, which returns a boolean value can be used to check for the Equivalence of both the vectors.
Matrix M can be generated by the following:

$$M = np.c\_[sp.jn(2,t),t] \tag{9}$$

The result of follwing will let us know about the equivalence of two vectors:

$$np.array\_equal(g(t, 1.05, 0.105), M.dot(np.array([1.05, 0.105]))) \tag{10}$$

7 Question 8

The contour plot of ϵ_{ij} is plotted in the following Figure 3. There exists **only one** minima.

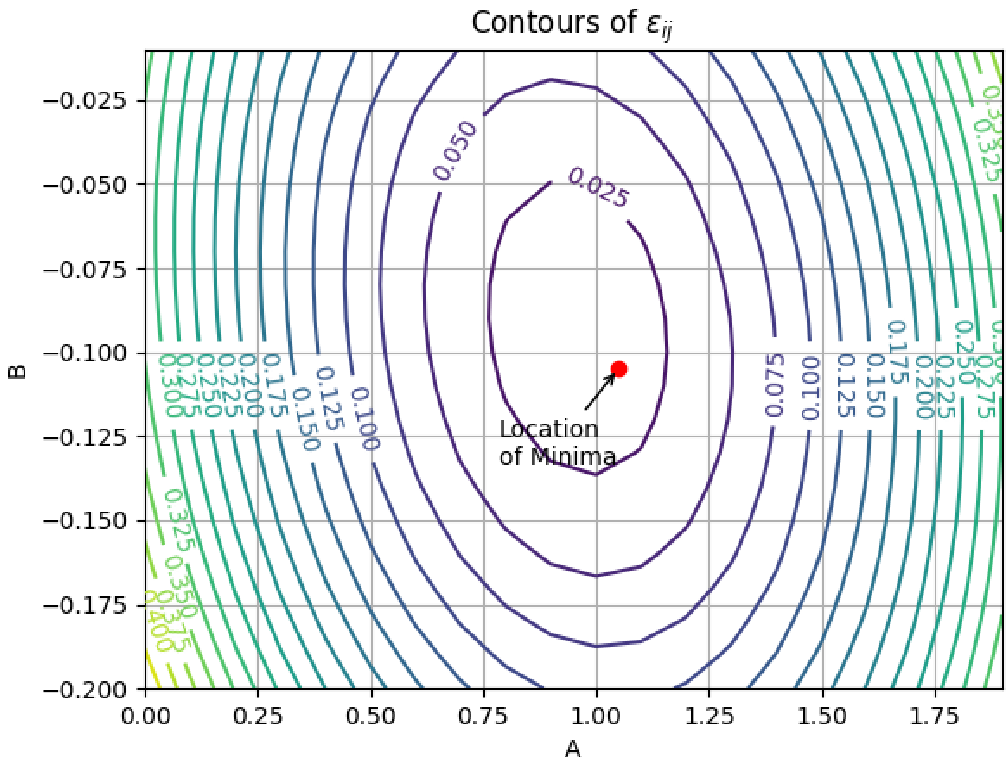

Figure 3: Contour Plot of the mean squared error versus the parameters A and B(Plot for Q8)

8 Question 10

The error in the estimate of A and B vs different noise σ have been plotted.

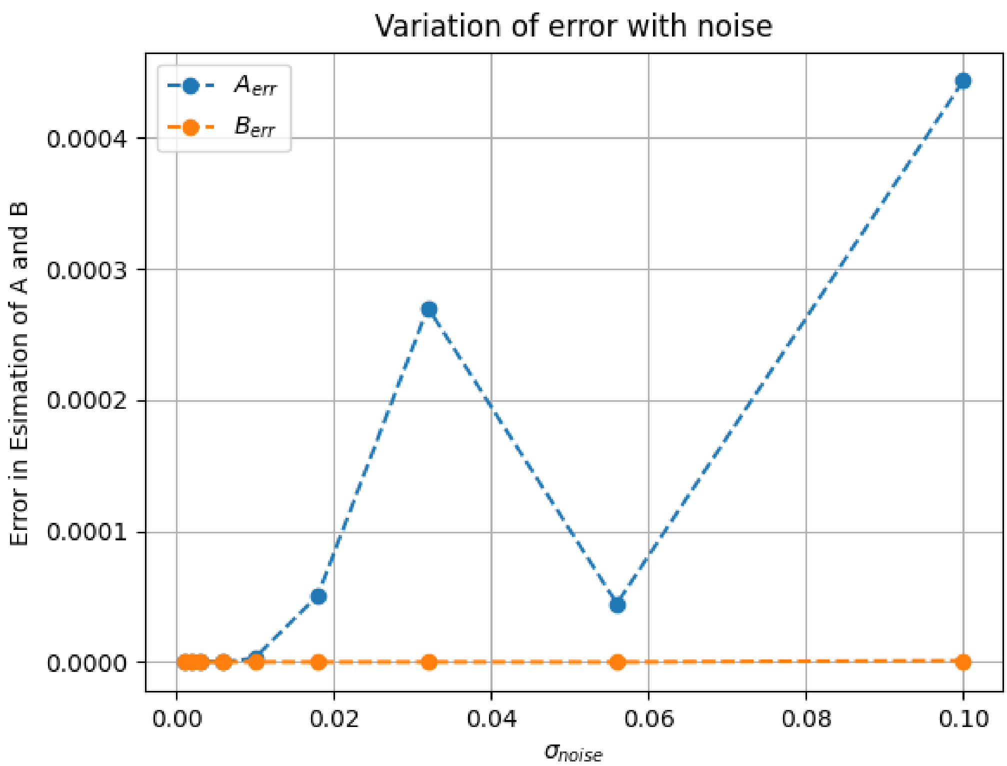

Figure 4: Error in the estimate of A and B vs different noise σ(Plot for Q10)

The error isn't linear, as it is clear from the graph the increase in error is not uniform with the uniform change with σ.

9 Question 11

The error in the estimate of A and B vs different noise σ have been plotted using a log-log plot,

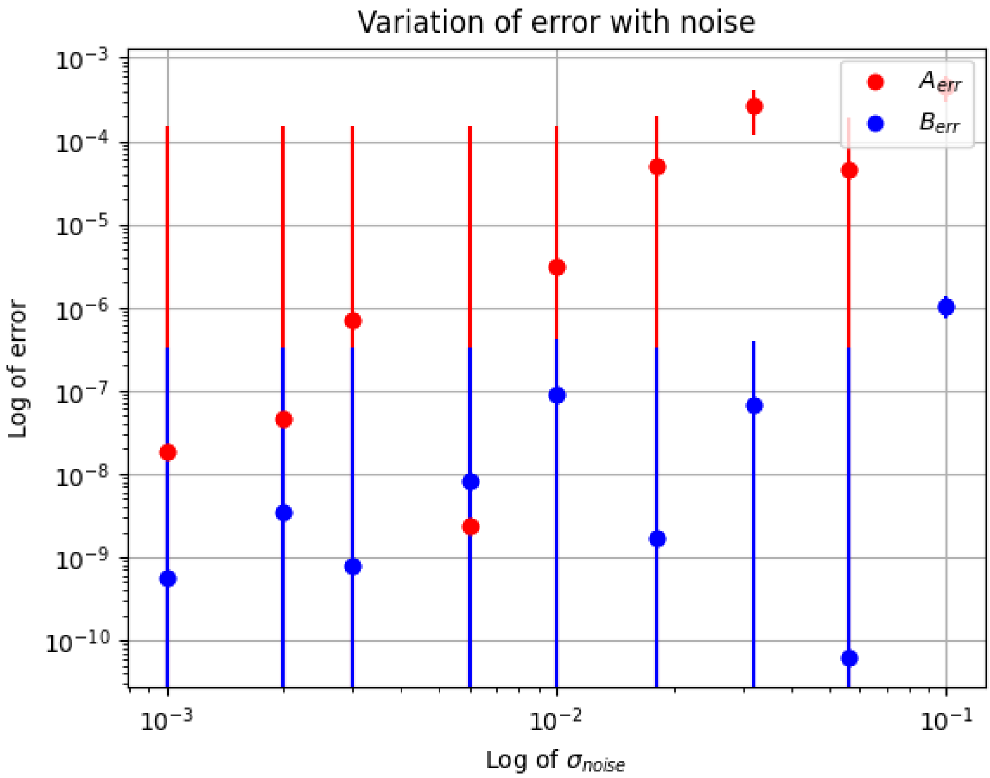

Figure 5: Log-Log plot of Error in estimate of A and B vs Different noise σ(Plot for Q10)

We can conclude, the error is a better linear fit.

10 Conclusion

From the above graphs, we can conclude that **the log of the** σ_noise is linearly related to **the log of the error** in the calcuation of the least error fit for a given data.