EE2703: Assignment 3

Sagar (ee20b115)

February 18, 2022

1 AIM

In this assignment we aim to :

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

2 PROCEDURE

The function to be fitted is:

$$f(t) = 1.05J_2(t) - 0.105t \tag{1}$$

where $J_2(t)$ is the 'Bessel Function of the first kind and of Order 2'. The true data is obtained using this equation.

2.1 Creating data with random noise

To create data with random noise, we add random noise to $f(t)$. This random noise, denoted by $n(t)$, is given by the standard normal probability distribution:

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

The resulting data will be of the form:

$$f(t) = 1.05J_2(t) - 0.105t + n_{\sigma_i}(t) \tag{3}$$

where, $n_{\sigma_i}(t)$ is the noisy data function with $\sigma = \sigma_i$ in (2). Thus for 9 values of σ, the noisy data is created and stored in 'fitting.dat' file.

2.2 Analyzing the data with noise included

The output result looks as follows:

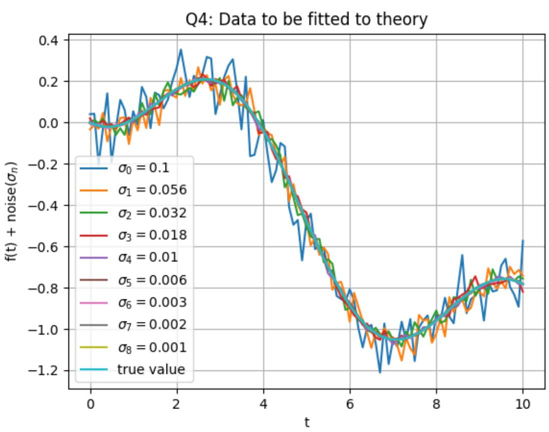

Figure 1: Noisy Data with True Data

As can be observed, the inaccuracy of the data increases with increasing σ. We can also observe this using 'errorbar' plot in python:

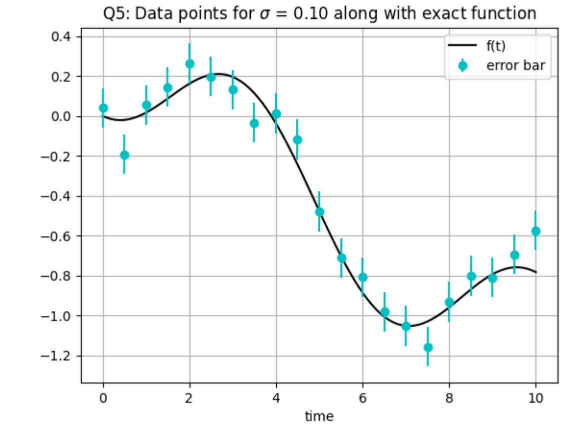

Figure 2: Noisy Data with Errorbar

The cyan lines (error bar) indicate the standard deviation of the noisy data from the original data. It is plotted for every 5^{th} point.

2.3 Finding an approximation for data with noise included

From the data, we can conclude that the data can be fitted into a function of the form:

$$g(t, A, B) = AJ_2(t) + Bt \tag{4}$$

where A and B are constants that we need to find.

To find the coefficients A and B, we first try to find the mean square error between the function and the data for a range of values of A and B, which is given by:

$$\epsilon_{ij} = \frac{1}{101}\sum_{k=0}^{101}(f(t_k) - g(t_k, A_i, B_j))^2 \tag{5}$$

where ϵ_{ij} is the error for (A_i, B_j). The contour plot of the error is shown below:

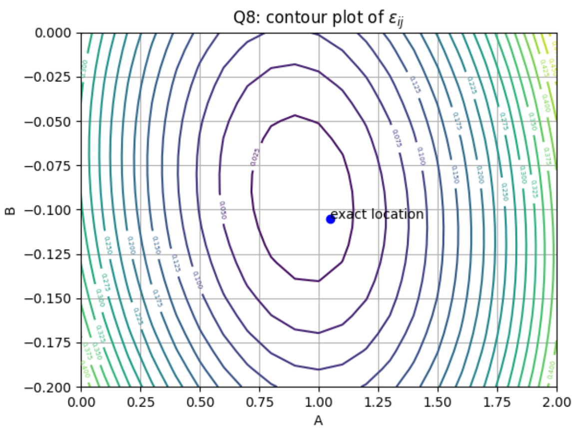

Figure 3: Contour Plot of ϵ_{ij}

We can see that there is a minimum and it is located approximately near the orignal function coefficients.

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

$$M.p = D \tag{6}$$

where

$$M = \begin{bmatrix} J_2(t_1) & t_1 \\ \dots & \dots \\ J_2(t_m) & t_m \end{bmatrix}, \; p = \begin{bmatrix} A_{fit} \\ B_{fit} \end{bmatrix} \text{ and } D = \begin{bmatrix} f(t_1) \\ \dots \\ f(t_m) \end{bmatrix} \tag{7}$$

Thus, we solve for p and then find the mean square error of the values of A_{fit} and B_{fit} found using 'lstsq' and the original values $(1.05, \; -0.105)$.

2.4 Finding out the variation of ϵ with σ_n

We solve (6) for different values of σ_n, by changing matrix D to different columns of 'fitting.dat'. We see that the variation of the MS error of values A_{fit} and B_{fit} is as follows:

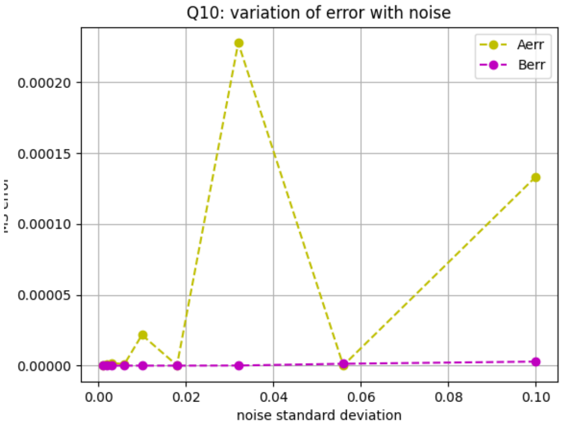

Figure 4: MS Error vs Standard Deviation

Except for one anomalous point near 0.03, it is approximately linear. To see that more clearly, we do loglog plot below:

Figure 5: Error vs Standard Deviation loglog Plot

3 CONCLUSION

we can conclude that the logarithm of the standard deviation of the noise linearly affects the logarithm of the error in the calcuation of the least error fit for a given set of data.