Assignment No 3

M.V.A.Suhas Kumar
EE17B109

February 12, 2019

1 Fitting Data to Models

This weeks Python assignment is mainly focused on studying the effect of noise on the fitting process.
Firstly,we generated a file fitting.dat with 10 columns with first column as time ,while the remaining columns are data
The data columns correspond to the function with different amounts of noise added.Here,noise random fluctuations in the value due to many small random effects.Noise is assumed to be normally distributed.

Each column function with given σ:

$$f(t) = 1.05J_2(t) - 0.105t + n(t)$$

with

$$Pr(n(t)|\sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp(\frac{-n(t)^2}{2\sigma^2})$$

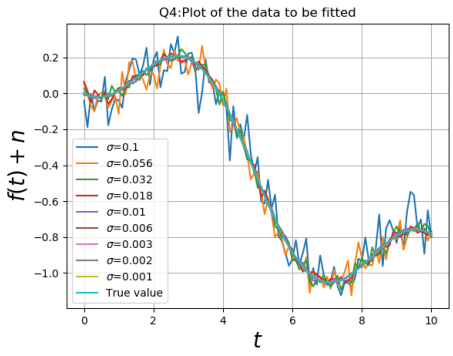

Figure 1: Plot of data to be fitted

Here I am plotting the error bar for one data column:

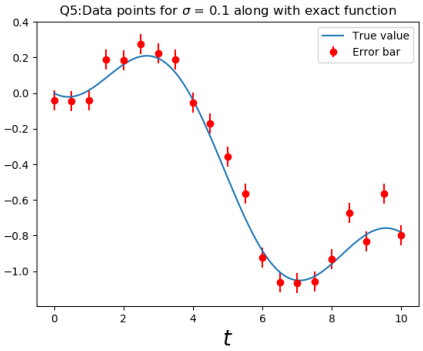

Figure 2: Data points for $\sigma = 0.1$ along with exact function

True value function is plotted by defining a function:

$$g(t, A, B) = AJ_2(t) + Bt$$

Next we assume that there exist some function which fits the noise data with general form $g(t, A, B) = AJ_2(t) + Bt$

Next we will find the (A,B) values by minimising the mean square error between the predicted values from given (A,B) and data column.

Contour Plot of $MS\ error$ for w.r.t data column1 for range of (A,B):

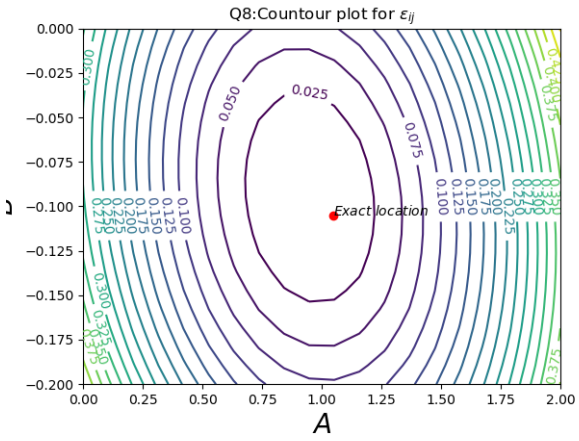

Figure 3: contour plot for ϵ_{ij}

From the above plot we can clearly see that there exist a single minima.

Using the Python function lstsq from scipy.linalg to obtain the best estimate of A and B for different data columns:

The following plots show the error in A and B for different data columns:

In the first plot B error is appearing to be constant but in reality it increases by a small amount.

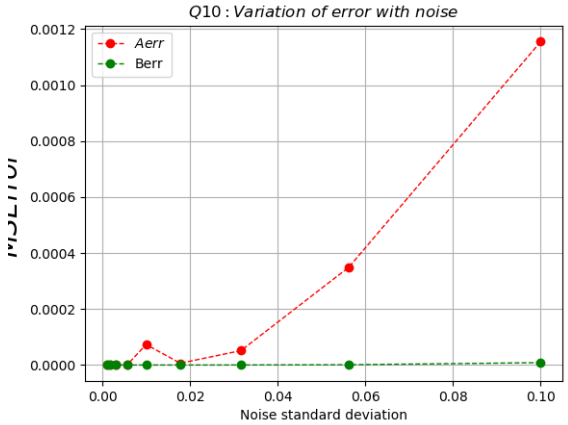

Figure 4: A and B error in linear scale

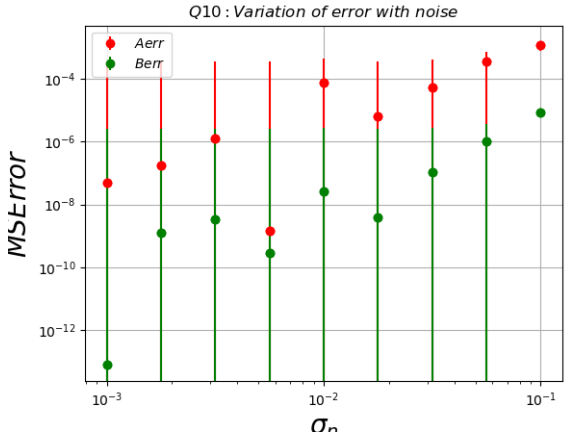

Figure 5: A and B error in log scale

Here I am plotting mean square error of predicted points w.r.t to true points versus sigma of noise.

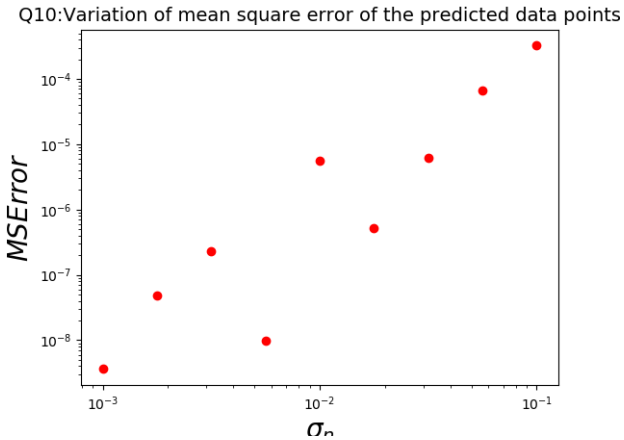

Figure 6: Mean square error for predicted points

Conclusion: If we calculate the mean square error for the predicted points w.r.t to the data points,In the log scale we observe that there will be a linear variation.