Fitting Data to Models

Assignment 3
EE2703 - Applied Programming Lab

Abhigyan Chattopadhyay
EE19B146

March 2021

1 The Problem at Hand

We are given a data file with values and a linear model to fit it, and asked to show the effect of noise on the fit.

2 The Curve Fitting Process

We use the method detailed in the Assignment document to fit a linear model to the given data.

Since the data has random noise, the best linear fit will be slightly shifted from the actual model. We will use several plots to show the effect of noise on the fitting process.

There will be some small errors in the linear fit since the data has some random noise.

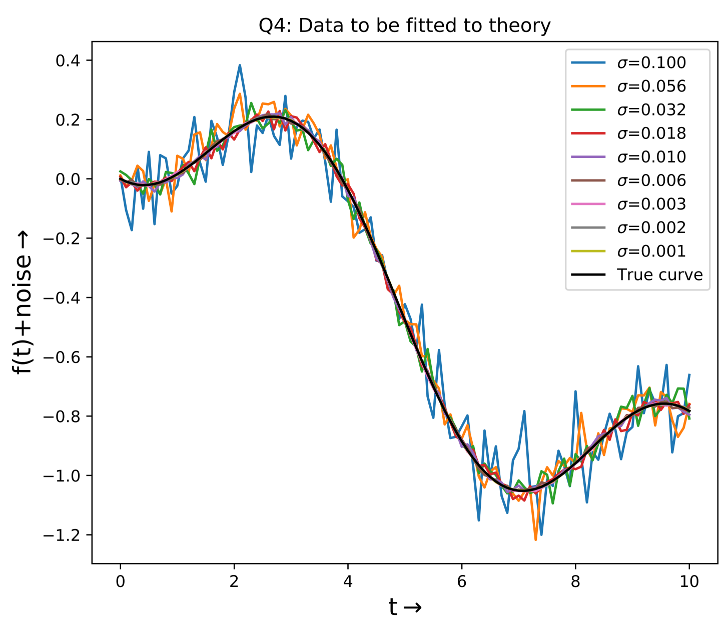
Looking at the plots below made with the python `matplotlib` library, we will look into the effects of the noise on the data given to us.

2.1 Step 1: Extracting and Visualizing the Data

The given data was obtained by running the `generate_data.py` script.

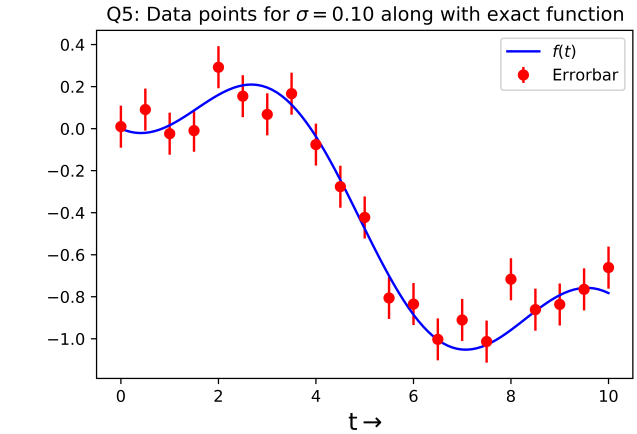
The data contained 10 columns : the first column was time, and the other 9 columns were each the data along with different amounts of noise, with standard deviation uniformly sampled from a logarithmic scale.

On plotting all the 9 columns, the following graph was observed:



2.2 Step 2: Visualizing Noise using an Error-bar Plot

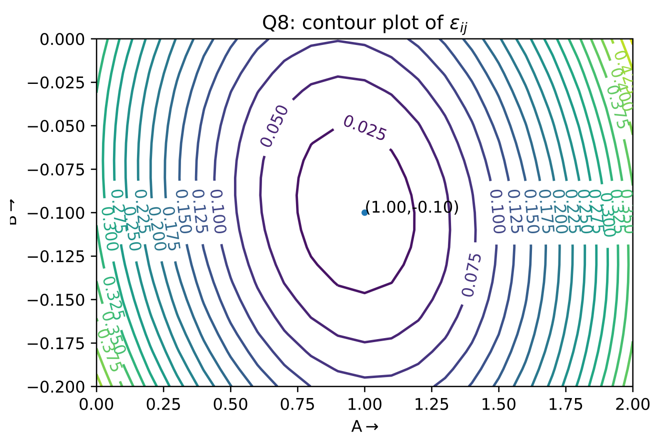
Here, the error in every 5th data point with respect to the original data was plotted using the `errorbar` function in `matplotlib`.



The points with the most error will affect the prediction of our parameters A and B the most in the final linear model.

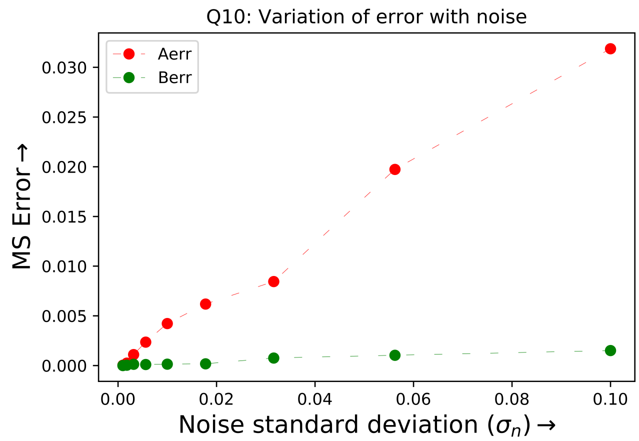
2.3 Step 3: Finding and Plotting the Error on a Contour Plot

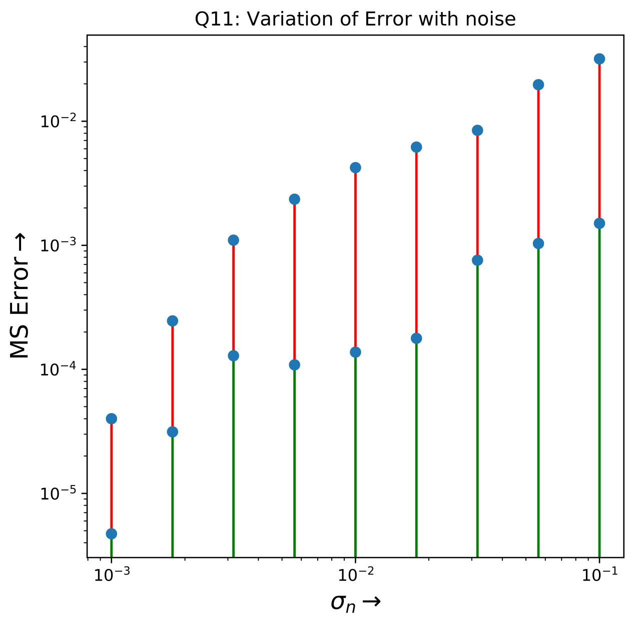
Here, the error is calculated with respect to each individual point and plotted using the `contour` function in `matplotlib`.



2.4 Step 4: Parameter Estimation

Next, the parameters are estimated and plotted using Stem Plots.





In Q11, it is asked what it means for log of error to vary linearly with the log of noise.

Since a linear relationship on a log-log plot represents a power law relationship between x and y (i.e. $y = kx^a$, $a \in \mathbb{R}$), hence this shows that our error is proportional to the noise with the power a being equal to the slope of the log-log graph.