## Assignment 3 - Curve Fitting

You are given 3 datasets, and the overall goal is to extract useful information from them. Details about the individual datasets are given below, along with the specific questions you need to answer for each dataset.

Note that in this exercise the document you submit is more important than the code. The documentation itself is expected to be about **1-2 pages per dataset** (3-8 pages total since the 3rd dataset has 2 associated questions). Long and unnecessary documentation will lead to less marks - only explain what is different in your code from the original problem statement.

Submit (in the usual .zip format):

- one PDF file containing the documentation
- one Python file per dataset containing the steps necessary to solve the problem
    - filenames should be `dataset1.py`, `dataset2.y`, `dataset3_1.py`, `dataset3_2.py`
    - note that for dataset 3 there are two parts which is why you use 2 names as above.
- README file with instructions on how to run the code

You can assume that `numpy` and `scipy` libraries are available on the target system, but should not make other assumptions. Make sure your code works properly on the Jupyter server before submitting - it will be tested with the default Python installation on that server.

### Dataset 1

This is a straight line with noise added to it. You need to set up a least squares regression and estimate the slope and intercept of the line. Your documentation for this should contain the following:

- How the M matrix for least squares is constructed
- Plot of the original noisy data with error bars - but error bars should be plotted only for 1 in every 25 data points
- On the same plot, you should also have a plot of the estimated line overlaid
- Add a legend for the plot - it should have entries for all the lines shown and should look similar to the figure below (this has been drawn with different data so your figure will look different).

### Dataset 2

This signal consists of 2 sine waves (with period $T$ and $3T$ - one is 3 times the frequency of the other) that have been added together, and then corrupted with noise. Your goal is to estimate the amplitudes
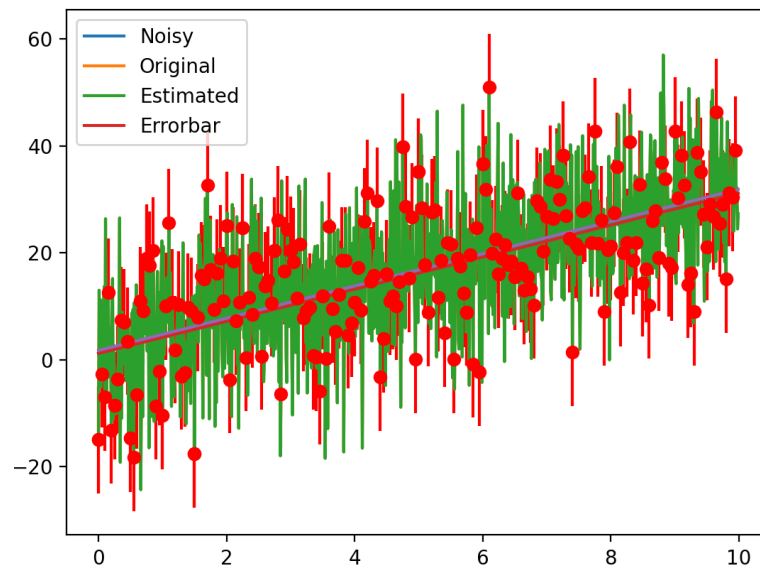
**Figure 1:** Example of `legend` use in matplotlib

of the two sine waves.

For this, you will first have to form an estimate of the periodicity of the waves. How you do that is left to you - you need to justify your answer. Your documentation should contain the following:

- How did you estimate the periodicity of the sine waves? You are free to use any method you want, but need to give some intuitive justification here. Note that the simpler the technique, the better - you are not expected to look for complicated answers here, but should be able to justify whether the approach you use is likely to give you a reasonable answer or not.
- How was the M matrix constructed and what is the equation you set up for least squares.

- Also solve the problem using `curve_fit` and report on the difference, if any.

**Dataset 3**

The data here corresponds to a simulation of the intensity of Black Body radiation emitted by an object at some temperature. The equation for Black Body radiation is of course given by the Planck equation:

$$B(f, T) = \frac{2hf^3}{c^2} \frac{1}{e^{\frac{hf}{k_B T}} - 1}$$

where $h$ is Planck's constant, $k_B$ is Boltzmann's constant, and $c$ is the speed of light in vacuum.

You need to solve this problem in two ways:

**First solution**

Look up values for $h$, $k_B$ and $c$, and use these values to estimate the temperature at which the data was simulated.

Note that you may need to provide the `curve_fit` function with suitable starting points so that it converges. This may require some trial and error - document how you approach this.

**Second solution**

Use the same data to estimate all the values: $h$, $k_B$, $c$ as well as $T$. Again, you may need to give reasonably close starting points for your estimates in order to make it converge.

Comment on the quality of the results you obtain. How close are they to the actual values, and can you think of ways to improve your estimates.