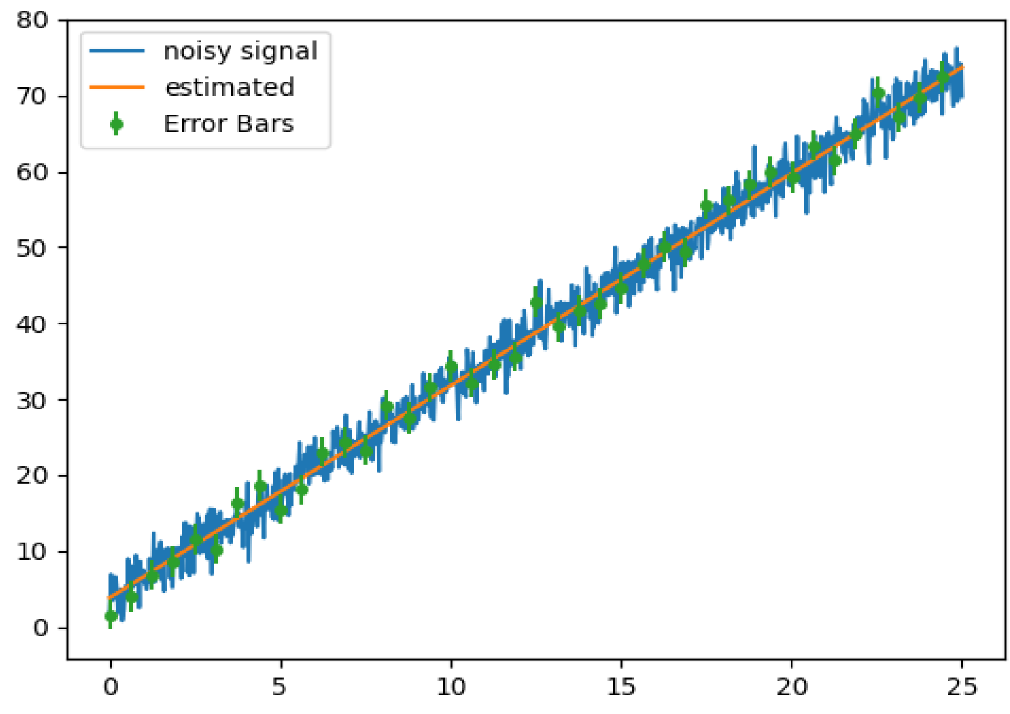
APL Assignment 3

Anurag Manoj ee22b003

October 2023

1 Dataset 1

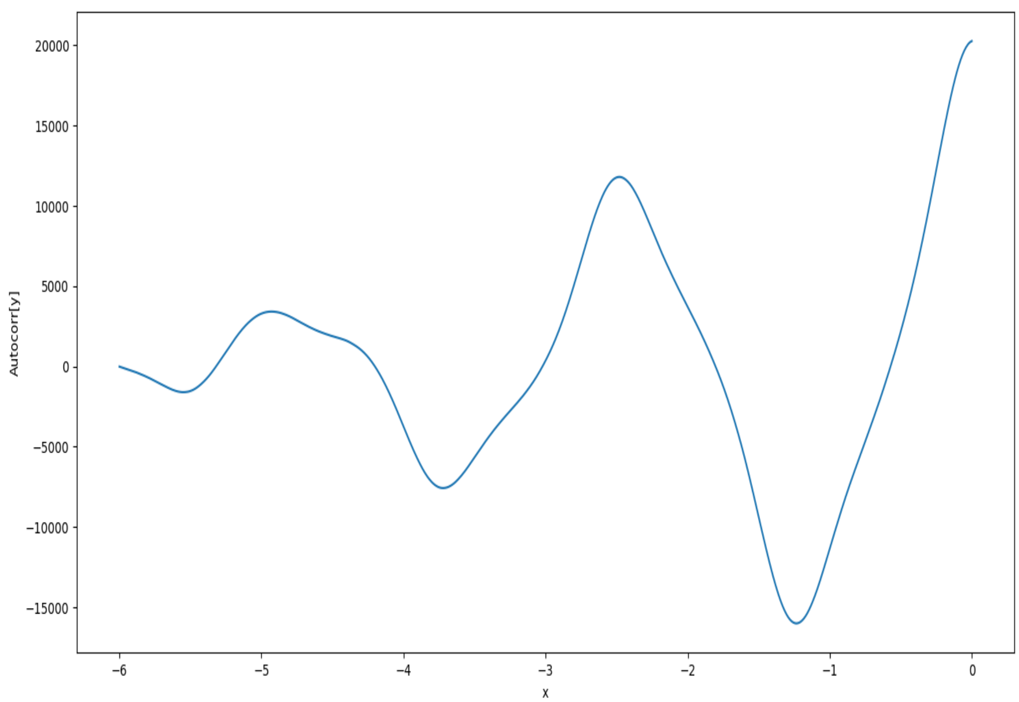
The formula to find the best fit line to the x and y vales given in arrays x and y is y=E[y]+$\frac{cov(x,y)}{var(x)}$ (x-E[x])
Slope=$\frac{cov(x,y)}{var(x)}$
Intercept-E[y]-E[x] $\frac{cov(x,y)}{var(x)}$



2 Dataset 2

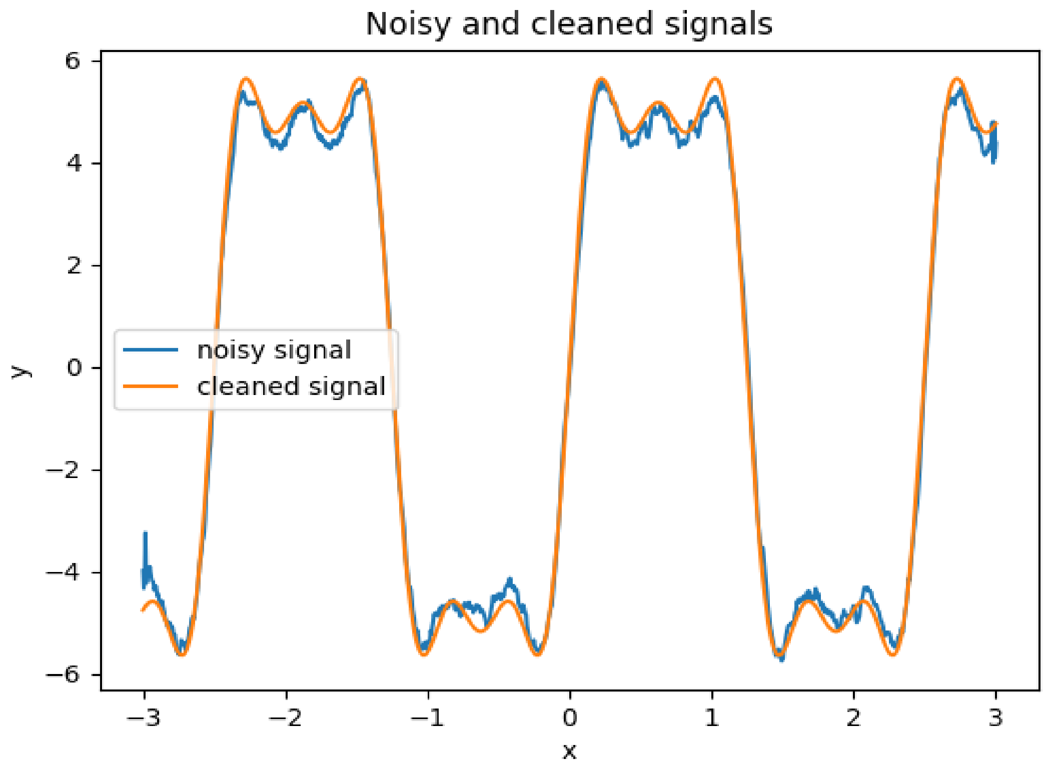
2.1 Estimating Time Period

The cross-correlation of 2 signals is given by sliding the second signal over the first left to right and calculating the sum of products of corresponding terms. The auto-correlation of two signals is the cross-correlation of the same signals. While sliding a signal over itself, the correlation will be max when the signals are displaced by a whole multiple of T.
Thus, the average Δx between 2 peaks of the auto-correlation will be the time period, which is found to be T=2.5, from the graph.



2.2 Finding best fit sum of 3 sines

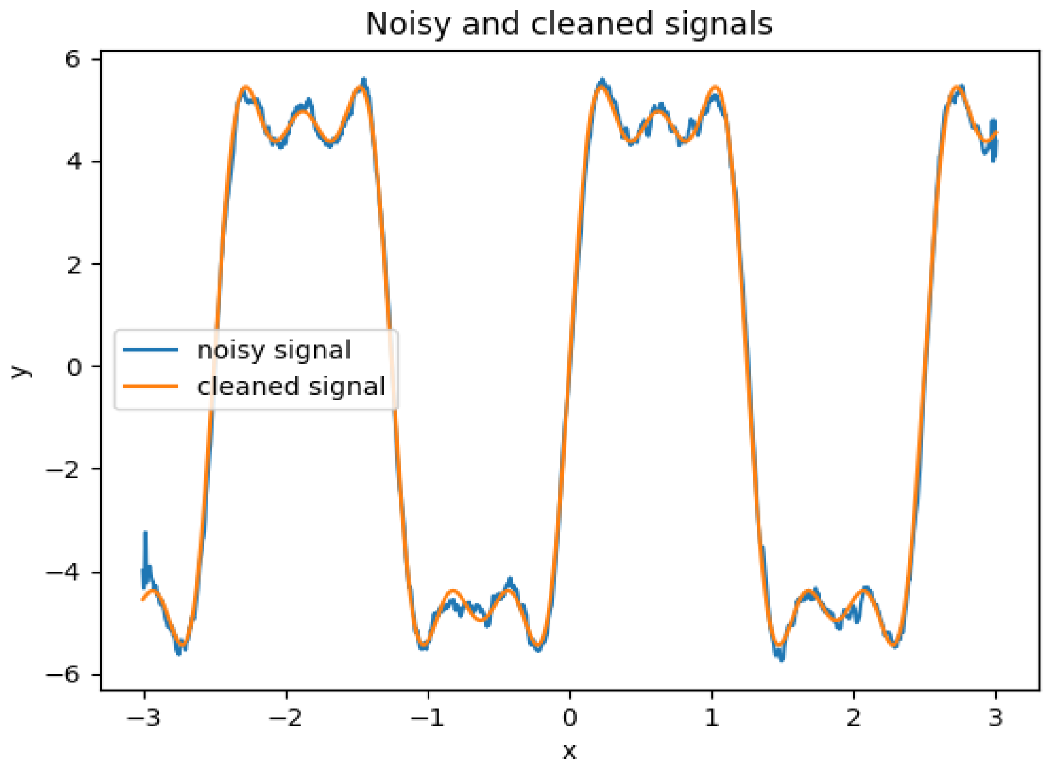
The phase of the input signal (x) can be found by finding the peak of cross-correlation of sin(2πfx) and the signal, but it is visually obvious that the input signal is zero at x=0 and it is given that it is a sum of sine functions.



Hence, a Fourier sine series can be made for the input signal.
Supposing f1,f2 and f3 are the 3 most important frequencies, then the signal can be approximated as a*sin(2πf1x)+b*sin(2πf2x)+c*sin(2πf3x) where a,b and c are chosen to minimize the mean square error. This happens when the best fit function (a vector in a space of dimensions=len(x)) is the projection of the input signal on the hyperplane spanned by the 3 vectors sin(2πf1x),sin(2πf2x) and sin(2πf3x).
But these are none other than the orthogonal basis vectors the signal is decomposed into by a fourier-sine series. Hence, the least square error coefficients a,b and c are none other than the fourier sine series coefficients. And the frequencies are those whose coefficients have highest amplitude. The coefficients are calculated with a separate function.

y(x) $\approx 4.998in(\frac{4\pi x}{5}) + 1.614sin(\frac{12\pi x}{5}) + 0.754sin(\frac{20\pi x}{5})$

2.3 Using scipy curve fit



This is the graph found by scipy.optimize.curvefit, with [1,1,1] as initial guess. The scipy curve fit seems to produce a visually more accurate fit.

3 Dataset 3

3.1 Part I

The integral from $-\infty\ to + \infty$ of Plancks distribution is
$\frac{\sigma T^4}{\pi}$, from Stefans law. From the graphing the data, it
is seen that nearly all the energy is contained in the range of values recorded.
Hence, the integral of y(x), calculated using scipy.integrate
is equated with Stefans law to calculate "T".

T=5000K

3.2 Part II

With only 2 effective parameters, 4 unknowns cannot be calculated