Problem Set 5

Edison Wang

Oct 20, 2023

Problem 1

(a) the graph is shown in Figure 1. Note that all three curves go up at first but then go down after some x values, as expected.

(b) To show that the maximum falls at $x = a-1$, we first take the derivative of the integrand:

$$\frac{d}{dx}(x^{a-1}e^{-x}) = (a-1)x^{a-2}e^{-x} - x^{a-1}e^{-x}.$$

Then we make it equal to 0 (since the slope is 0 at maximum) and solve for x:

$$\begin{aligned}(a-1)x^{a-2}e^{-x} - x^{a-1}e^{-x} &= 0\\ (a-1) - x &= 0\\ x &= a-1.\end{aligned}$$

Therefore, the maximum falls at $x = a - 1$.

(c) For the given change of variables:

$$z = \frac{x}{c+x},$$

it is clear that $z = 1/2$ when $x = c$. Since the maximum falls at $x = a-1$, the appropriate choice of c is $a - 1$ which puts the maximum at $\frac{1}{2}$.

(d) Using the given substitution, the integrand becomes

$$I = e^{(a-1)\ln x}e^{-x} = e^{(a-1)\ln x - x}.$$

This expression is a better one because it now requires three steps of calculation that do not easily lead to overflow or underflow: the first step is $(a-1)\ln x$, where the logarithm ensures that this number is neither too small nor too large; the second step is to subtract x from the previous result, which is numerically simple; the third step is to compute an exponential, but the magnitude of the exponent $(a-1)\ln x - x$ is for the most part (of the range of the integral) smaller than the old exponent $-x$ (of the e^{-x} term), so there is less chance of underflow.

(e) With the change of variables, the gamma function becomes

$$\int_0^1 e^{(a-1)\ln\frac{z(a-1)}{1-z}-\frac{z(a-1)}{1-z}}\frac{a-1}{(1-z)^2}dz.$$

Using the method of Gaussian quadrature with 100 points, I obtain

$$\Gamma\left(\frac{3}{2}\right)\approx 0.8862269613087213,$$

which is only $3.58559633095723\times 10^{-8}$ from the theoretical result $\frac{1}{2}\sqrt{\pi}$.

(f) I obtain

$$\begin{aligned}\Gamma(3) &= 2.0000000000000013;\\ \Gamma(6) &= 119.99999999999999;\\ \Gamma(10) &= 362879.99999999994.\end{aligned}$$

These results are closely equal to 2, 120, and 362880, respectively.

Problem 2

(a) The plot of the data is shown in Figure 2 (the units are not specified in the problem and therefore left out).

(b) The matrix A is defined as the matrix with the i-th column representing the time matrix (vector) whose elements are raised to the i-th power (starting at $i=0$). Using SVD, I calculate the pseudo-inverse of A, and then multiply that by the signal matrix (vector) to obtain the coefficients. Finally, the matrix of predicted signals are obtained by multiplying A by the coefficient matrix (vector).

The coefficients of the fitted polynomial are

$$[-1.96089274e+00, -1.58229162e-26, 1.18890711e-17, -9.18508556e-27],$$

the first corresponding to t^0, the second t^1, etc. The plot of this polynomial is shown in Figure 3. Visually, it does not seem to be a good fit to the data.

(c) The residuals are calculated by subtracting the predicted signals from the actual signals and are shown in Figure 3. Using the absolute values of the residuals and the given standard deviation of 2.0, I find that only 56.7% of the data lie within one standard deviation of our model (68% expected for normal distribution), and only 89.1% lie within two standard deviations (95% expected). Therefore, this model is clearly not a good fit in any sense.

(d) For this part, I find that fitting with an eighth order polynomial yields the highest percentages of data lying within both one and two standard deviations (57.3% and 89.5% respectively). The polynomial is shown in Figure 4. However, this is just slightly better than the third-order fit and is still not really a good fit to the data. The condition number is simply infinity since A is singular.

(e) I used 23 cosine terms in the form $\cos \omega_i t$ and 23 sine terms in the form $\sin \omega_i t$, starting with a period equal to half of the time span covered, and ending with one-tenth of that period. The resulting curve fit is shown in Figure 6. The percentages of data lying within one or two standard deviations are respectively 69.9% and 95.6%, suggesting that this is probably a good fit to the data. The typical periodicity is roughly 1.3×10^8.

My Github link: https://github.com/ziqiu-wang/phys-ua210.

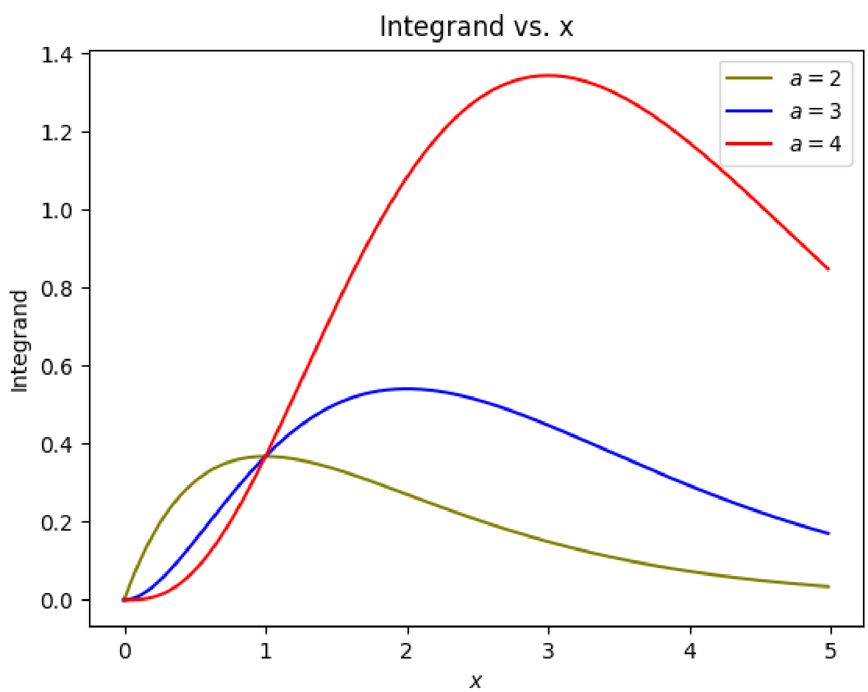

Figure 1: Integrand vs. x graph for Problem 1.

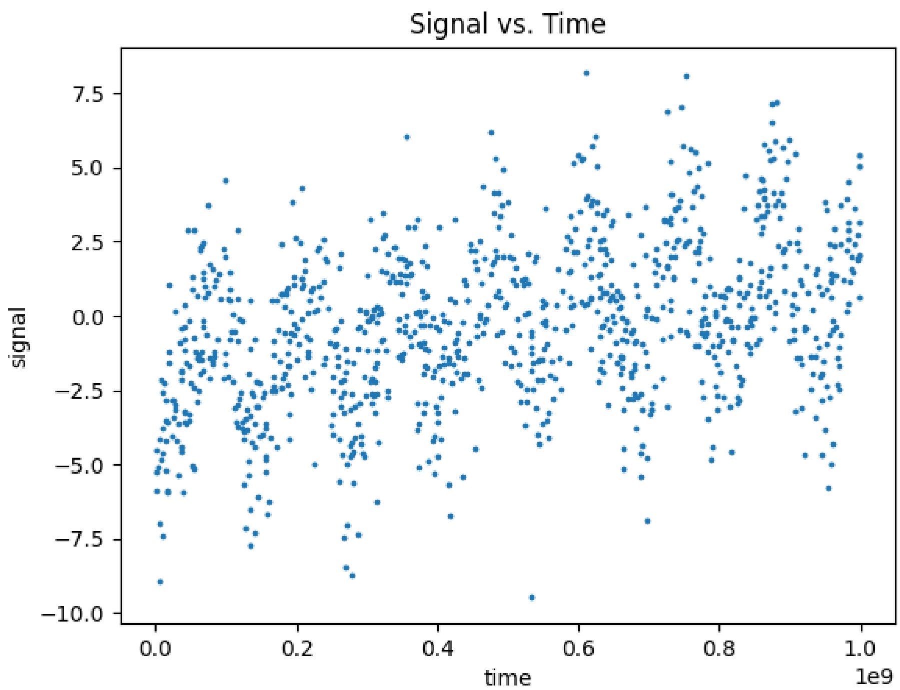

Figure 2: Problem 2 Part (a): plot of signal vs. time.

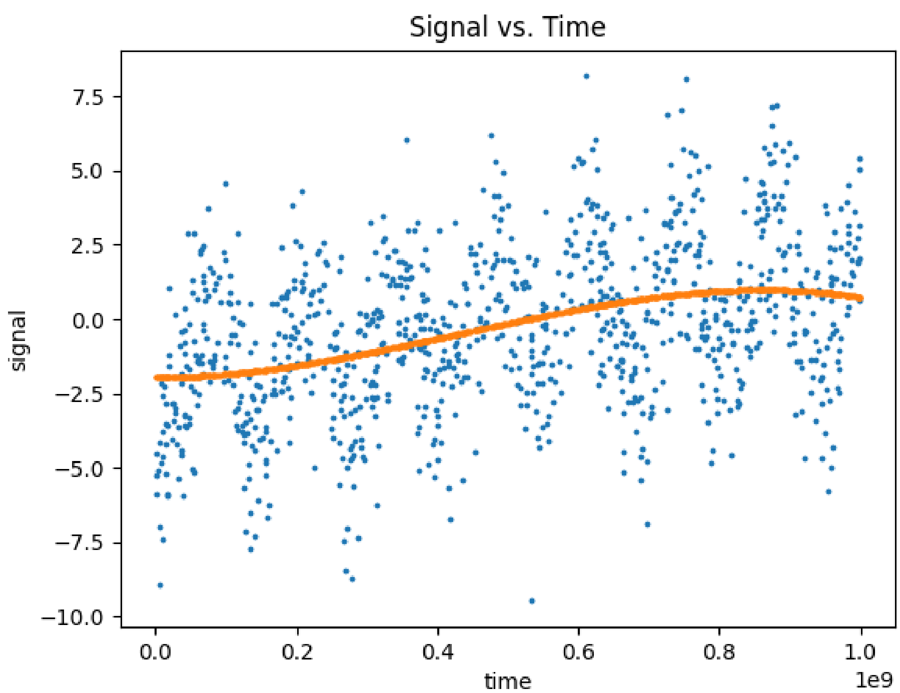

Figure 3: Problem 2 Part (b): third-order curve fit.

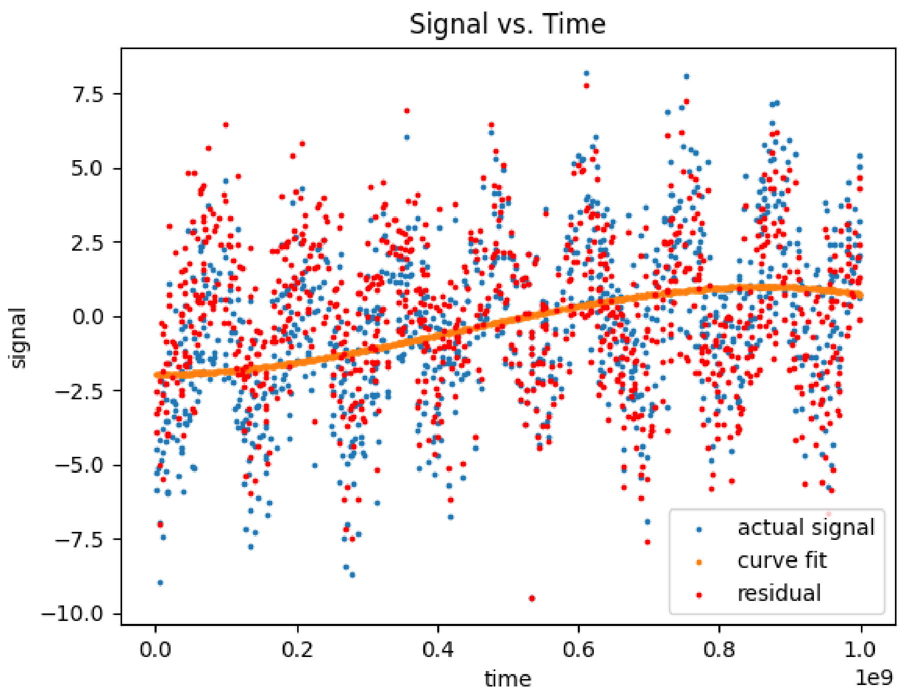

Figure 4: Problem 2 Part (c): third-order curve fit with residuals.

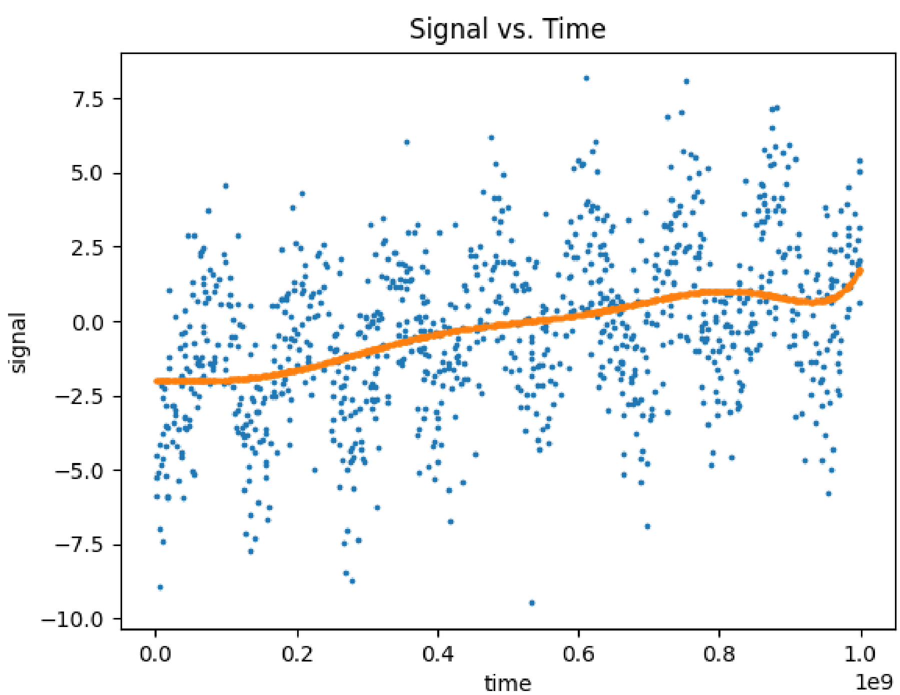

Figure 5: Problem 2 Part (d): eighth-order curve fit.

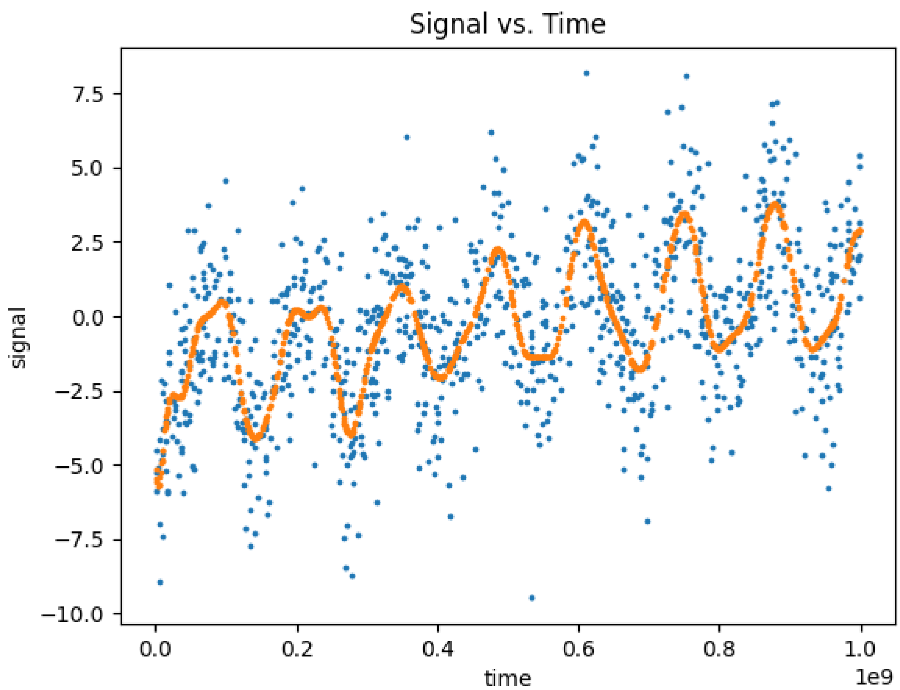

Figure 6: Problem 2 Part (e): curve fit using cosines and sines.