Project Assignment 1 EQ1220 Signal Theory

Anamitra Bhar, Shashang Murali

I. INTRODUCTION

In the initial problem, a one-dimensional Gaussian distribution is given which contains three sequences of different lengths. This task's goal is to demonstrate how the length of the sequence influences both the approximation of a Gaussian distribution based on the data that is provided and the recovery of the distribution's parameters. For the second task a second .mat file is used which contains two matrices and is based upon the theories of bivariate Gaussian distributions. The goal is to derive the general expression for the joint Gaussian distribution and generate a three-dimensional plot of the empirical pdfs and determine which plot matches to a particular ρ value. Once this is accomplished, we need to reason out the expected shape of the plot if we had $-\rho$ instead of ρ. The aim of the next task which makes use of the expression of joint Gaussian distribution derived in the previous task is to derive the conditional pdf of a given random variable in its closed-form provided the definition of conditional distribution pdf. In majority of the cases, a transmitted signal is usually corrupted or disturbed by some noise, and in the scope of this project we consider this noise to be Gaussian noise. In Task 4, we are required to plot the periodograms of these possible outcomes and determine which plot matches to which of the two output sequences. Additionally, we are also required to present our opinion on the accuracies of the recovered sinusoidal frequencies. For the next task of this section, the given file contains a sequence of correlated noise samples and making use of the periodograms we have to provide justifications on the impact of noise correlation on the estimation accuracy of the normalized frequencies. For the final part of the project, a white noise sequence is provided for the accomplishment of the following two tasks whose output sequence is then fed into another linear system. We are required to derive and plot the power spectra of the two given signals followed by obtaining and plotting the autocorrelation function of one of the output processes.

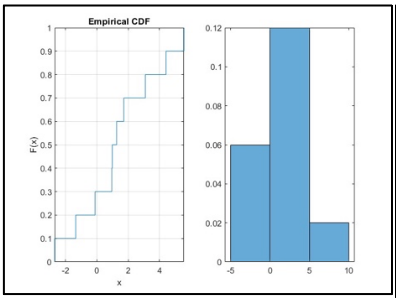
II. PROBLEM FORMULATION AND SOLUTION

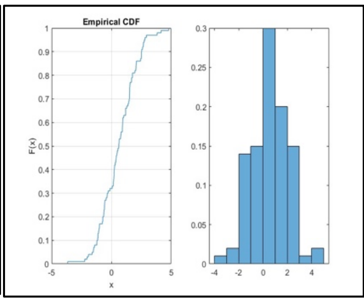
A. Task 1

The file Gaussian1D.mat was fed into the software MATLAB for the completion of this task. Using in-built functions, we were able to get the mean and variance values of the distribution of each sequence. The obtained values are mentioned below:

Sequence 1:	Sequence 2:	Sequence 3:
Mean = 1.3829	Mean = 0.6135	Mean = 0.4408
Variance = 5.6385	Variance = 2.1711	Variance = 1.9596

The empirical distributions of the given three sequences as obtained are presented below:





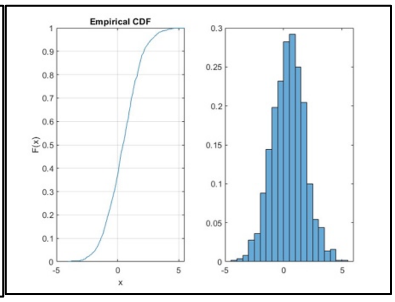


Figure 1: Empirical Distribution of Sequence 1
Figure 2: Empirical Distribution of Sequence 2
Figure 3: Empirical Distribution of Sequence 3

B. Task 2

For this task, we import another file Gaussian2D.mat which contains two matrices which have been generated from bivariate Gaussian distributions. General expression for joint Gaussian distribution:

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\} \quad \textbf{(1)}$$

In order to obtain the ρ_i values we computed the variance and covariance of the two sequences. They were found out to be as follows:

For ρ_1:		For ρ_2:	
1.0000 (variance)	0.2467 (covariance)	1.0000 (variance)	0.7518 (covariance)
0.2467 (covariance)	1.0000 (variance)	0.7518 (covariance)	1.0000 (variance)

Histograms were plotted in order to obtain the 3D plots of the empirical pdfs. They were observed as follows:

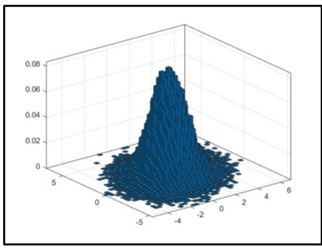
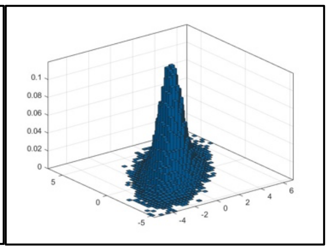



Figure 4: Empirical pdf Plot Sequence 1

Figure 5: Empirical pdf Plot Sequence 2

As mentioned in the titles of the above figures, Figure 4 corresponds to ρ_1 and Figure 5 corresponds to ρ_1. It can be understood from the above plots that as the correlation coefficient increases, the correlation (direct proportionality) between the x and y axes increases and hence the data points are scattered in the plane in Figure 4. From Figure 5, we can observe that with a lesser correlation coefficient value the coefficient between the axes becomes lesser and a starker comparison can be made between the values of both the axes (more pointed shape of the plot, less scattering of data points in the plane). If the pdf was characterized from -ρ instead of ρ, an inversely proportional relation would have existed between the two axes. So, as the value of x would have increased the value of y would have decreased or vice versa and a change in behavior of the curve would have been observed (the direction in which the data points scatter in the plane would change).

C. Task 3

When the random variables X and Y are jointly normally distributed random variables, we can say that X+Y is also normally distributed, and the mean is the sum of means. Hence,

$$\sigma_{X+Y} = \sqrt{\sigma_X{}^2 + \sigma_Y{}^2 + 2\rho\sigma_X\sigma_Y} \quad \textbf{(2)}$$

From the general expression derived in the previous task we can obtain the following:

$$f_Z(z) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \iint_{xy} \exp\{\left[-\frac{1}{2(1-\rho^2)}\right]\left(\frac{x^2}{\sigma_x{}^2} + \frac{y^2}{\sigma_x{}^2} - \frac{2\rho xy}{\sigma_x\sigma_y}\right)\}\,\delta(z-(x+y))dxdy \quad \textbf{(3)}$$

Substituting y as (z-x) and solving this integral we get the pdf of X+Y:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_+} \exp\,(-\frac{z^2}{2\sigma_+{}^2})$$

Where σ_+ is the same as σ_{X+Y} from equation (1). In the case of Z=X-Y, we can write the following expression:

$$f_Z(z) = \frac{1}{\sqrt{2\pi({\sigma_x}^2+{\sigma_y}^2-2\rho\sigma_x\sigma_y)}} \exp\left(-\frac{z^2}{2({\sigma_x}^2+{\sigma_y}^2-2\rho\sigma_x\sigma_y)}\right) \quad \textbf{(4)}$$

From equation (4) we can write that,

$$\sigma_{X-Y} = \sqrt{{\sigma_X}^2 + {\sigma_Y}^2 - 2\rho\sigma_X\sigma_Y} \quad \textbf{(5)}$$

D. Task 4

The sequences y1 and y2 from SinusInNoise1.mat is in the form of the following equations:

$$H0 : y_0(n) = w(n) \quad \textbf{(6)}$$

$$H1 : y_1(n) = \sum_{k=0}^{k=1} \sin(2\pi\nu_k n + \varphi_k) + w(n)$$

y1 and y2 are plotted in MATLAB using a built-in function
Fn : periodogram(X)
X -> Sequence y1 and y2 which are vectors of size 100x1 in our case
This function calculates PSD by taking the Fast Fourier Transform of the input sequence and further squaring the frequency domain output to get the power corresponding to the normalized frequencies. The periodogram of both the sequences are given below in Figure 6 and Figure 7.

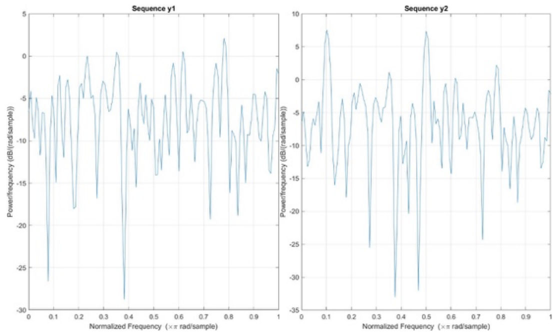


Figure 6: Periodogram of y1
Figure 7: Periodogram of y2

E. Task 5

The process of plotting the periodogram of sequence y from SinusInNoise2.mat is same of Task 4. By using the in-built periodogram function we observe Figure 8.

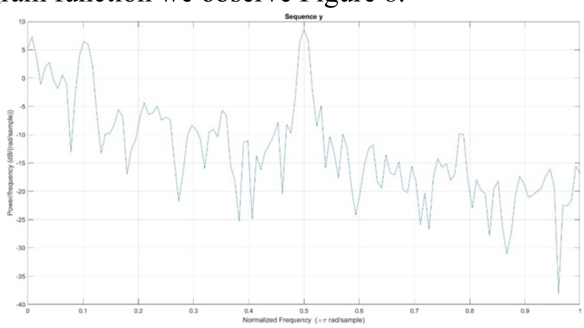


Figure 8: Periodogram of Sequence y

F. Task 6 & Task 7

The AR(1) sequence given a noise input is modelled as

$$x_1(n) = \alpha x_1(n-1) + z(n) \quad \textbf{(7)}$$

Autocorrelation function of $x_1(n)$ is calculated as the following:

$$r_{x1}(k) = E[x_1(n)x_1(n-k)]$$
$$= E[(\alpha x_1(n-1) + z(n))x_1(n-k)\} \quad \textbf{(8)}$$

Solving equation (8) by substituting k=0,1,2 and σ_z^2 we get,

$$r_{x1}(0) = \frac{\sigma_z^2}{1-\alpha^2}$$
$$r_{x1}(1) = \alpha r_{x1}(0)$$
$$r_{x1}(2) = \alpha^2 r_{x1}(0)$$

In general, we get

$$x_2(n) = \sum_{k=0}^{+\infty} h_2(k)x_1(n-k)$$

The Power Spectral Density (PSD) of $x_1(n)$ is given by taking the Fourier Transform of $r_{x1}(k)$.

$$R_{x1}(\nu) = \frac{1}{1.0625 - 0.5\cos(2\Pi\nu)}$$
$$R_{x2}(\nu) = \frac{1.778}{1.0625 - 0.5\cos(2\Pi\nu)} \quad \textbf{(9)}$$
$$r_{x2}(k) = 1.898 \times 0.25^{|k|}$$

Equation 9 is given by using the formula from standard discrete Fourier transforms. The plot of $R_1(\nu)$ is given in Figure 9.

The AR(1) sequence is fed as input to an linear system with impulse response $h_2(n) = \beta^n u(n)$. The response of the system is given by

$$x_2(n) = \sum_{k=0}^{+\infty} h_2(k)x_1(n-k) \quad \textbf{(10)}$$

The power spectra of $x_2(n)$ can be written in terms of PSD of $x_1(n)$ and the Fourier transform of the impulse function,

$$R_{x2}(\nu) = R_{x1}(\nu)\,|H_2(\nu)|^2$$
$$H_2(\nu) = F\{h_2(n)\}$$
$$H_2(\nu) = \frac{1}{1-\beta e^{-j2\Pi\nu}}$$
$$|H_2(\nu)|^2 = \frac{1}{(1-\beta)^2}$$

Hence,

$$R_{x2}(\nu) = \frac{1}{(1-\beta)^2}\frac{\sigma_z^2}{1+\alpha^2-2\alpha\cos(2\Pi\nu)} \quad \textbf{(11)}$$

The plot of $R_{x2}(\nu)$ is shown in Figure 10.

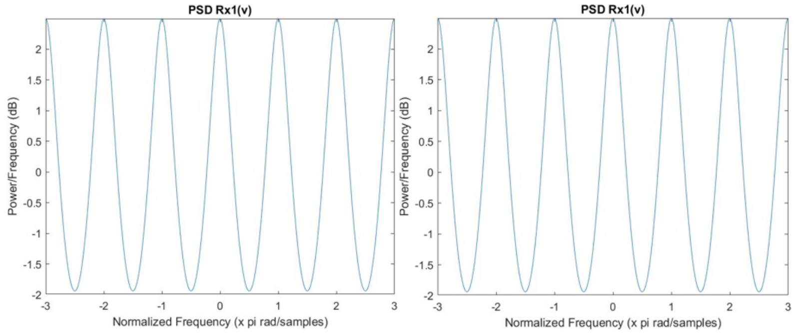


Figure 9: Plot of $R_1(\nu)$

Figure 10: Plot of $R_{x2}(\nu)$

The autocorrelation function of $x_2(n)$ is derived by taking the inverse Fourier transform of $R_{x2}(\nu)$ which is,

$$r_{x2}(k) = F^{-1}\{R_{x2}(\nu)\} = F^{-1}\{\frac{1}{(1-\beta)^2}\frac{\sigma_z^2}{1+\alpha^2-2\alpha\cos(2\Pi\nu)}\}$$

The output of the above equation is of the form, $r_{x2}(k) = c_1 a^{|k|}$ **(12)**

where c_1 is a constant.

III. CONCLUSIONS

From the computations and analysis in task 1 we observed that as the sequence length increases more Gaussian-like characteristic is seen in the shape and behavior of the empirical distribution plots of the sequences. In task 2, we saw that as the correlation coefficient increases how the axes are more correlated to each other and how the behavior of the 3D plot varies with different values of the correlation coefficient. We also analyzed the case of -ρ, and arrived at the conclusion that an inversely proportional relation would have existed between the two axes. In the next task we used the general expression for the joint Gaussian distribution to derive the pdfs for X+Y and X-Y. Coming to task 4, from Figure 6 we can notice that there is a spike in the power at 0.1π rad/sample (2*0.05*π) and 0.5π rad/sample (2*0.25*π). It is noticed that there are impulse pulses at frequencies ν_0 and ν_1 which is the Fourier transform of a sinusoidal function. This implies that sequence y2 represents H1 (second part of equation 6) and y1 represents white noise of the form of H0 (first part of equation 6). From the periodogram of y2, it can be observed that the two impulse functions can be filtered almost accurately from the noise by using appropriate bandpass filters. The white noise component is uniformly distributed over the whole frequency interval at the power of σ_w^2 and is negligible at ν_0 and ν_1. In task 5, from Figure 8 we observe that non-white noise is not uniformly distributed in whole frequency interval unlike white noise. Impulse peaks at ν_0 and ν_1 can still be seen but the power of noise below ν_0 is as much as the power of the impulse peaks which makes it difficult to design a filter to separate the signal from the noise at low frequency. Frequency ν_1 is at a higher frequency where the noise does not correlate much with the signal, hence can be filtered out without any complexities. For tasks 6 and 7 we can conclude that, $R_{x1}(\nu)$, $R_{x2}(\nu)$ and $r_{x2}(k)$ after substituting $\sigma_z^2 = 1, \alpha = 0.25, \beta = 0.25$ we get the below equations,

$$R_{x1}(\nu) = \frac{1}{1.0625 - 0.5\cos(2\Pi\nu)}$$

$$R_{x2}(\nu) = \frac{1.778}{1.0625 - 0.5\cos(2\Pi\nu)}$$

$$r_{x2}(k) = 1.898 \times 0.25^{|k|}$$

REFERENCES

[1] P. Handel, R. Ottoson, H. Hjalmarsson, Signal Theory, KTH, 2012

[2] Sum of normally distributed random variables n.d., accessed on 19 September 2022, <https://en.wikipedia.org/wiki/Sum_of_normally_distributed_random_variables#Correlated_random_variables>