Utah State University
ECE 6010
Stochastic Processes
Programming Exercise # 1
Due Friday September 9.

Introduction

This exercise will provide an opportunity to do some calculations and plots with actual data. The intent is to make some of the abstract concepts a little more concrete. This exercise is to be done using Matlab.

Background

`randn` Every call to the Matlab function `rand` generates an independent instance of a standard Gaussian random variable. That is, `randn` is a Gaussian random number generator: $X \sim \mathcal{N}(0, 1)$. Random column vectors of length `n` are generated by `randn(n,1)`. Row vectors are generated with `randn(1,n)`. A matrix of random numbers is generated with `randn(n,m)`. For more information, type `help randn` in Matlab.

`rand` Similarly, the Matlab function `rand` generates independent uniform $\mathcal{U}(0, 1)$ random numbers. Column and row vectors and matrices of random numbers are generated using `rand(n,1)`, `rand(1,n)`, `rand(n,m)`. For more information type `help rand` in Matlab.

Histograms The Matlab `hist` command produces a histogram. This is a representation of the empirical density function. In a histogram, a sequence of bins is established. For each value in a set of data, the number of times that data points fall in a bin is counted. In the Matlab `hist` command, the histogram is plotted automatically. To see an example of how the histogram works, type the following in Matlab:

```
x = randn(1,1000);   % create a vector of 1000 Gaussian random number
hist(x,20);  % plot the histogram with 20 bins
hist(x,100); % plot the histogram with 100 bins
```

Estimating mean and covariance Given a sequence of vector observations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, where each vector is a column vector of length n drawn independently and identically distributed according to some distribution, the sample mean of the distribution is

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i .$$

The sample covariance is

$$\hat{\Sigma} = \frac{1}{N-1} \sum_{i=1}^{N} (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T .$$

Exercises

1. Generate 1000 points of a $\mathcal{N}(2, 5)$ random variable. Plot the histogram of the data. Estimate the mean and the variance from the data you generate. How closely do the estimates correspond to actual parameters?

2. Generate 1000 points of a $U(-2, 5)$ random variable. Plot the histogram of the data. Estimate the mean and the variance from the data you generate. How closely do the estimate correspond to the actual parameters?

3. Write a Matlab function that will generate N points of Gaussian $\mathcal{N}(\hat{\boldsymbol{\mu}}, \Sigma)$ data, where $\boldsymbol{\mu}$ is a vector of length n. The function should have the "declaration"

   ```
   function X = gengauss(mu,Nigma,N)
   ```

4. The file `prog1dat.mat` (on the class website) contains $N = 1000$ data points representing measurements x_1, x_2, x_3, x_4 from a four-dimensional physical system. Load the data into Matlab using the command

   ```
   load prog1dat
   ```

   A 4×1000 variable `X` will be created with the data in it. Suppose that $x_1 = 5$ and $x_3 = 7$ is measured. Determine the best estimate of the variables x_2 and x_4. Explicitly state all the appropriate covariance and mean vectors, and how you obtain your estimates.

5. Continuing the previous problem, suppose that the variables x_1, x_2, and x_3 are available. Write a function `predictx4` which will estimate the corresponding value of x_4. The function should have the "declaration"

```
function x4hat = predictx4(x1,x2,x3)
```

(with possibly some other arguments as well). If $x_1 = 3$, $x_2 = 3.6$, $x_3 = 5.2$, what is the estimate of x_4?

6. Continuing the previous problem, let $\mathbf{X}^{(1)}$ be obtained from the first two components of the four-dimensional data. Estimate Σ_{11}, the covariance matrix of $\mathbf{X}^{(1)}$, and $\boldsymbol{\mu}^{(1)}$, the mean vector.

   Plot contours of the pdf of $\mathbf{X}^{(1)}$. The function `plotellipse.m` (on the class website) may be helpful. Compare the axes of the ellipses with the eigenvectors of Σ_{11}. What is the relationship?