STA 5107: Home Assignment # 2

Spring 2021/Due Date: February 2nd

1. **Fisher's Discriminant Analysis**: Write a program to perform Fisher's Discriminant Analysis for the given data sets. The dataset is given in form of a large matrix $X = [X_1 X_2 \dots X_C] \in \mathbb{R}^{n \times Ck}$, where $X_c \in \mathbb{R}^{n \times k}$ is the data for each of the classes $c = 1, 2, \dots, C$. $C = 5$ and $k = 100$ in all the examples, but n changes.

   1. Compute mean μ_c of each class, for $c = 1, 2, \dots, C$.
   2. Compute within class scatter matrix S_w and between class scatter matrix S_b
   3. Perform generalized eigen-value decomposition using `[V,D] = eig(Sb,Sw,'chol');`
   4. Reorder solution using `V = fliplr(V)` and `D = flipud(fliplr(D))`.
   5. Set `U = V(:,1:d)` and project $Z = U^T X$. As part of your solution: (1) generate a plot of the eigenvalues, (2) show a colored 2D scatter plot of the projected data. There are several datasets provided for this problem.

2. Write a program to simulate a homogeneous Poisson counting process over the interval $[0, 100]$. Generate a few (5) sample paths for intensity $\lambda = 0.1$ and display them on the same plot. Remember that the mean of exponential arrival times is $\frac{1}{\lambda}$.

3. Using the program in the previous problem, count the number of events occurring in the interval $[10, 60]$. Plot a histogram of 50 realizations of this random number.

4. Write a matlab program to simulate a random walk, for a given value of T and $s = \alpha\sqrt{T}$. Plot the sample paths of the X_t versus the real time $t \in [0, 10]$, for $\alpha = 1.0$ and $T = 1,\ 0.1,\ 0.01$, and 0.001. In each case, choose the total number of steps n to be $10/T$, so that you go up to time $t = 10$ irrespective of T. Use the "stairs" command in matlab to plot the sample paths.