ECE 485:
Assignment #6

For the associated Matlab data file develop and submit a Matlab Live-Script to perform the following analysis:

1. Apply PCA to reduce the associated provided data set from its original 25 dimensions down to 2-dimensions.

2. Apply the expectation maximization (EM) algorithm to provide an estimated pdf $\hat{p}(\boldsymbol{x})$ of the data of (1) where $\hat{p}(\boldsymbol{x})$ is assumed to be a kernel density estimate formed as the sum of 5 Gaussian distributions,

$$\hat{p}(\boldsymbol{x}) = \sum_{k=1}^{5} \frac{1}{C_k} \hat{p}_k(\boldsymbol{x})$$

   where $\hat{p}_k(\boldsymbol{x}) \sim N(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_j)$ and C_k is an appropriate normalization constant, i.e., you need to find the combination of the best 5 estimated means and covariances which *together* provide the best estimate of the given data's probability distribution.

3. Separately, produce a Parzen window estimate of $\hat{p}'(\boldsymbol{x})$ of the data's probability distribution.

4. Compare the estimated $\hat{p}(\boldsymbol{x})$ of (1) with an estimated $\hat{p}'(\boldsymbol{x})$ of (2). Discuss the differences and distinctions between the two estimation methods and which estimate is better for this given data set and why?

5. For both 2 and 3 provide separate 3D plots of the computed estimated $\hat{p}_k(\boldsymbol{x})$ and $\hat{p}'_k(\boldsymbol{x})$ distributions along with separate plots of their associate 2D contour plots.

For 1-4 you must develop your own Matlab code to do the analysis, i.e, not simply call/incorporate code found on the Internet for doing expectation maximization and/or Parzen window estimation.

Your submitted Matlab LiveScript must be able to correct execute and be clearly commented, i.e., make use of both the code and text sections that are available within LiveScripts.

All plots must be properly labeled in terms of their x and y axes and titles.

Marking will be done through executing and reading through your submitted LiveScript.