**JHU Engineering for Professionals**
**Applied and Computational Mathematics**
**Data Mining: 625.740**

**Homework 11**

Generate 1,000 two-dimensional samples for each of two Gaussians, $p(\mathbf{x}|\omega_i) \sim N(\boldsymbol{\mu}_i, \Sigma_i)$ with

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \boldsymbol{\mu}_2 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } \Sigma_2 = PDP^T$$

where

$$P = \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \text{ and } D = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}.$$

Run $k$-means clustering, fuzzy $k$-means clustering, and Expectation-Maximization on these data. Plot the resulting membership sets for each run and briefly discuss the results.