Computer experiment 4

(EM algorithm for Gaussian mixtures)

1. Consider a mixture of three Gaussians $\sum_{i=1}^{3} N(\mu_i, \Sigma_i)P_i$ with mean vectors $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}4\\4\end{bmatrix}, \begin{bmatrix}10\\2\end{bmatrix}$ and covariance matrices $\begin{bmatrix}1 & 0.4\\0.4 & 1\end{bmatrix}, \begin{bmatrix}1 & -0.6\\-0.6 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$, respectively. Generate 500 samples according to the following rule. The first two samples are generated from the 2nd Gaussian, the 3rd sample from the 1st one, and the 4th sample from the last Gaussian. This rule repeats until all 500 samples are generated..
   (a) Use EM algorithms and the generated samples to estimate the unknown parameters $\boldsymbol{\mu}_i, \Sigma_i, P_i\ (i = 1,2,3)$. Please specify your experimental settings (e.g., initialization, stopping criterion) in the report.
   (b) Repeat the mixture density estimation by EM when the mean vectors are $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}6\\2\end{bmatrix}$.
   (c) Repeat the mixture density estimation by EM when the mean vectors are $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix}, \begin{bmatrix}3\\1\end{bmatrix}$.
   (d) Compare the results (in terms of confusion matrices or 2-D visualization) and draw your conclusion.

(K-means algorithm)

2. Use k-means algorithm on the data set of the previous experiment, for $k = 2,3,4$. Compare the results and draw your conclusion.

(Linear discrimination functions)

3. Consider a two-class and 2-D classification task, where the feature vectors in each class are Gaussian pdfs $N(\mu_i, \Sigma_i), (i = 1,2)$ with mean vectors $\mu_1 = \begin{bmatrix}1\\1\end{bmatrix}$, $\mu_2 = \begin{bmatrix}0\\0\end{bmatrix}$ and covariance matrices $\Sigma_1 = \Sigma_2 = \begin{bmatrix}0.2 & 1\\0 & 0.2\end{bmatrix}$. Generate 100 vectors from each class. To guarantee **linear separability** of the classes, disregard vectors with $x_1 + x_2 < 1$ for the 1st class and vectors with $x_1 + x_2 > 1$ for the 2nd class.
   (a) Apply the <u>perceptron algorithm</u> and the <u>sum-of-squared-error classifier</u> on

the data set.

(b) Plot the data set and the decision lines.

(c) You may try various initial values, apply different variants of the algorithms, or include additional constraints in your implrmentation. Compare and discuss the results.