Homework #4

Using EM to Estimate a Probability Density with a Mixture of Gaussians

EELE 578
Assignment Date: 11/4/2022
Due Date: 11/18/2022

Implement the EM algorithm (see page 6 of the paper "Using EM to Estimate a Probability Density with a Mixture of Gaussians", which is given below) to find the mixture of Gaussians that explain the data points seen in the second figure.

Initialize: all $\langle z_{im} \rangle$, π_m, $\boldsymbol{\mu}_m$, and $\boldsymbol{\Sigma}_m$

1: **repeat**
2: **for** $i = 1$ to N **do** *//The E step*
3: **for** $m = 1$ to M **do**
4:

$$p(\mathbf{x}_i | z_{im} = 1; \boldsymbol{\theta}) = (2\pi)^{-d/2} |\boldsymbol{\Sigma}_m|^{-1/2} \exp\left\{ -\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu}_m)^T \boldsymbol{\Sigma}_m^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_m) \right\}$$

$$\langle z_{im} \rangle = \frac{p(\mathbf{x}_i | z_{im} = 1; \boldsymbol{\theta}) \pi_m}{\sum_j^M p(\mathbf{x}_i | z_{ij} = 1; \boldsymbol{\theta}) \pi_j}$$

5: **end for**
6: **end for**
7: **for** $m = 1$ to M **do** *//The M step*
8:

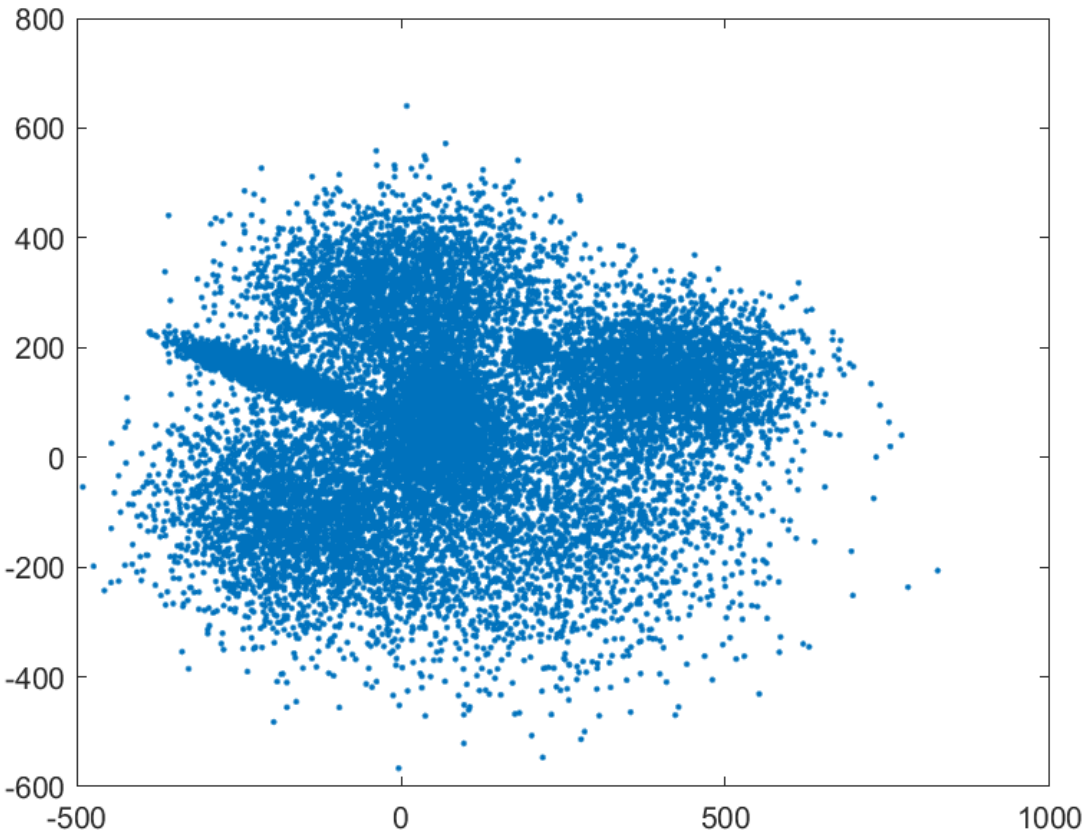
$$\boldsymbol{\Sigma}_m = \frac{\sum_{i=1}^N \langle z_{im} \rangle (\mathbf{x}_i - \boldsymbol{\mu}_m)(\mathbf{x}_i - \boldsymbol{\mu}_m)^T}{\sum_{i=1}^N \langle z_{im} \rangle}$$

$$\boldsymbol{\mu}_m = \frac{\sum_{i=1}^N \langle z_{im} \rangle \mathbf{x}_i}{\sum_{i=1}^N \langle z_{im} \rangle}$$

$$\pi_m = \frac{\sum_{i=1}^N \langle z_{im} \rangle}{N}$$

9: **end for**
10: **until** model parameters converge

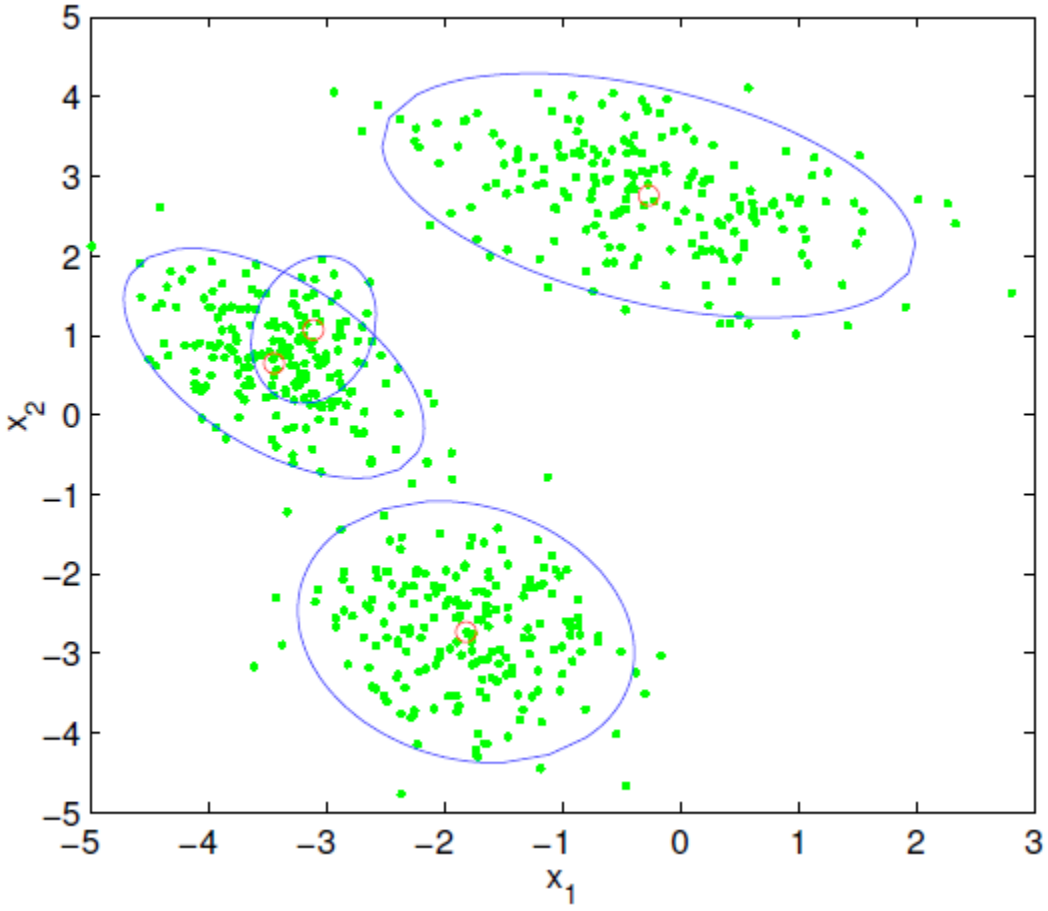
Algorithm 1: Estimate $\pi_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m$ for $1 \leq m \leq M$



The data is found in the provided Matlab file **HW04_data.mat** where each row is a data point (i.e. the data is two dimensional).

Submit your Matlab code to D2L that implements the EM algorithm along with a document that includes:

1. Your estimate of the **means and covariant matrices** of the Gaussians.
2. What is your estimate of the **number of Gaussians** that produced this data?
3. Do the EM initial conditions have an impact on your estimates?
4. Create a figure of the data where you **plot a marker point for each Gaussian mean and also plot the 2-sigma ovals** (95% confidence ovals) for each Gaussian on top of the data in a similar fashion to the figure below.



Here are a couple links that describes how to draw the 95% confidence ellipses.

https://www.visiondummy.com/2014/04/draw-error-ellipse-representing-covariance-matrix/

https://www.mathworks.com/matlabcentral/fileexchange/4705-error_ellipse

See also page 343 of the online book *Data Mining and Machine Learning* that describes the EM algorithm.

https://dataminingbook.info/book_html/chap13/book.html