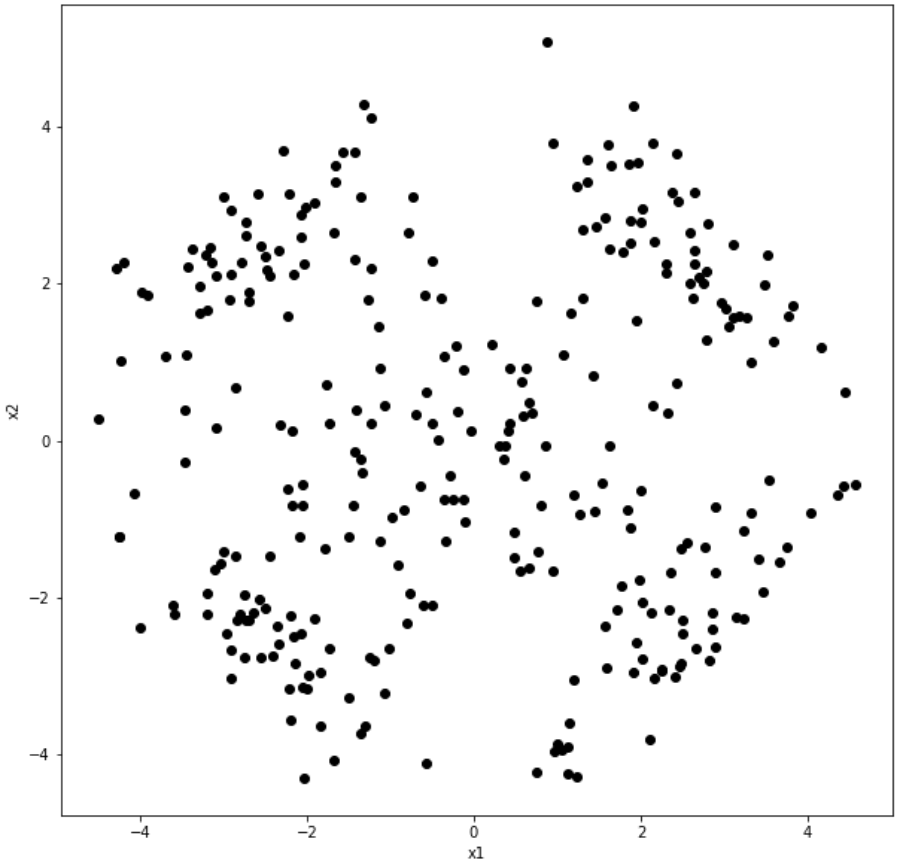
İrem Şahin

ENGR421

Mehmet Gönen

HW6-Report

For this project, I have written my code in a Jupyter Notebook .ipynb file. I have started my code by importing necessary libraries and writing our parameters down which were given in the pdf file. As requested in the pdf file,I have written down my means, covariances and sizes. Then, I have generated my points according to the given parameters. My generated data point for seed 421 was looking like this:



Then, I have defined the update_centroids, update_memberships and plot__state functions for future use. I have used the algorithms we have learned at Lab11.

After this, I have iterated these twice(k-means clustering). This way I have generated an initial centroids matrix. Centroids after 2 iterations are like this:

```
[[-0.72448243  3.11594719]
 [ 2.15067196  0.43893571]
 [ 2.01250253 -2.06593643]
 [-1.8293349  -2.95685596]
 [-3.18772687 -2.20425841]]
```

Which is not entirely correct. However , since we did only 2 iterations, this was expected.

After this, I have reinitialized the sample_covariances matrix and class priors using the centroids as the sample_means matrix. Then I have used this matrix as the sample_means matrix for the EM algorithm.

For the EM algorithm, I started by implementing the gaussian in the formulas below:

The second term is independent of π_i and using the constraint that $\sum_i \pi_i = 1$ as the Lagrangian, we solve for

$$\nabla_{\pi_i} \sum_t \sum_i h_i^t \log \pi_i - \lambda \left(\sum_i \pi_i - 1 \right) = 0$$

and get

$$\pi_i = \frac{\sum_t h_i^t}{N} \tag{7.11}$$

which is analogous to the calculation of priors in equation 7.2.

Similarly, the first term of equation 7.10 is independent of the components and can be dropped while estimating the parameters of the components. We solve for

$$\nabla_{\Phi} \sum_t \sum_i h_i^t \log p_i(\boldsymbol{x}^t|\Phi) = 0 \tag{7.12}$$

If we assume Gaussian components, $\hat{p}_i(\boldsymbol{x}^t|\Phi) \sim \mathcal{N}(\boldsymbol{m}_i, \mathrm{S}_i)$, the M-step is

$$\begin{aligned} \boldsymbol{m}_i^{l+1} &= \frac{\sum_t h_i^t \boldsymbol{x}^t}{\sum_t h_i^t} \\ \mathrm{S}_i^{l+1} &= \frac{\sum_t h_i^t (\boldsymbol{x}^t - \boldsymbol{m}_i^{l+1})(\boldsymbol{x}^t - \boldsymbol{m}_i^{l+1})^T}{\sum_t h_i^t} \end{aligned} \tag{7.13}$$

where, for Gaussian components in the E-step, we calculate

$$h_i^t = \frac{\pi_i |\mathrm{S}_i|^{-1/2} \exp[-(1/2)(\boldsymbol{x}^t - \boldsymbol{m}_i)^T \mathrm{S}_i^{-1} (\boldsymbol{x}^t - \boldsymbol{m}_i)]}{\sum_j \pi_j |\mathrm{S}_j|^{-1/2} \exp[-(1/2)(\boldsymbol{x}^t - \boldsymbol{m}_j)^T \mathrm{S}_j^{-1} (\boldsymbol{x}^t - \boldsymbol{m}_j)]} \tag{7.14}$$

However, since the online term is exhausting and every teacher gives homework like as if there is no tomorrow, I was not able to finish implementing the algorithm.