Assignment 3
Dimensionality Reduction using Principal Component Analysis
and
Clustering using Expectation-maximization Algorithm

Rituparna Datta
1505091

November 14, 2020

1 PCA plot

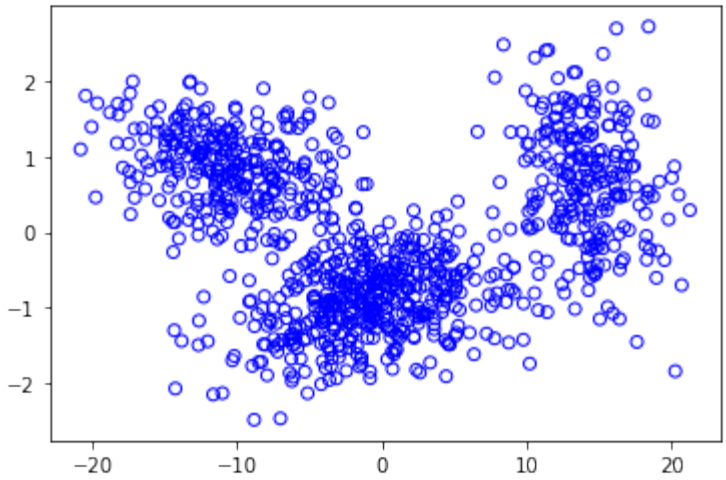

Figure 1: PCA implementation of given data

2 EM algorithm

Here, I've taken 0.0000001 as **Threshold** value and the EM stops if the absolute difference of current and previous log likelihood value is less than this. **Max Iteration** = 500 is also set. As, initialization is done by taking random means from the processed data set, the algorithm will converge at different iterations.

The final output of EM is given below.

Dimension	2
Cluster	3
Iteration	42
Log likelihood	-4596.686729426368

Mean,μ_k

-10.4117459	0.85632562
14.23274218	0.67617259
-0.65487466	-0.91102844

Co variances ,Σ_k

16.48528289	-0.61336521
-0.61336521	0.24986364

6.92264976	-0.47561616
-0.47561616	0.66883417

22.04691267	0.83026929
0.83026929	0.26001895

Mixing coefficients, w_k:

0.2996239704378329	0.24035191113329293	0.4600241184288746

3 PCA plot, using 3 clusters

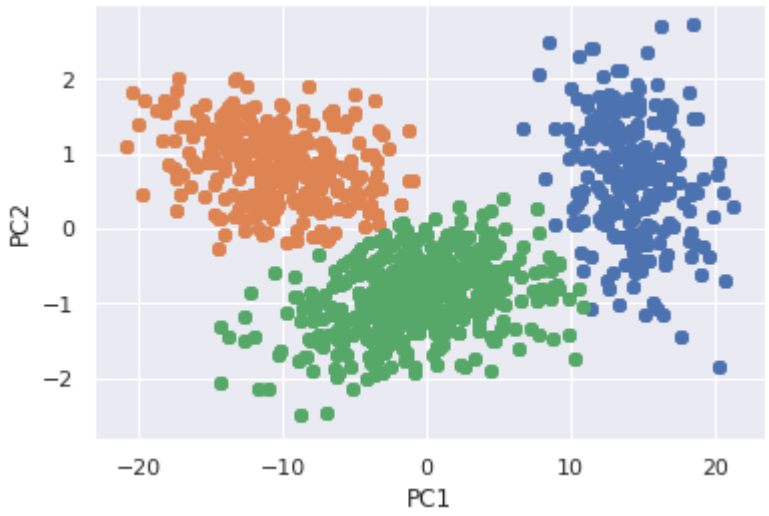

Figure 2: PCA using 3 clusters