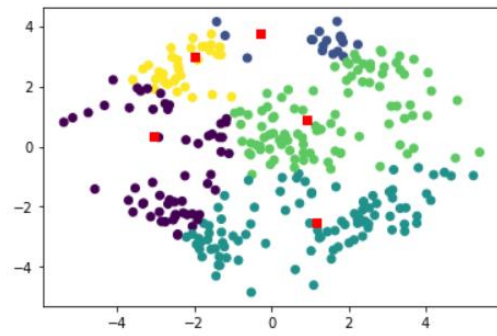


ENGR 421 / DASC 521: Introduction to Machine Learning
Homework 05: Expectation-Maximization Clustering

İzel Yazıcı-May 25, 2021

1. You are given a two-dimensional data set in the file named `hw05_data_set.csv`, which contains 300 data points generated randomly from five bivariate Gaussian densities with the following parameters.



2. To initialize your EM algorithm, you should take the centroids given in the file named `hw05_initial_centroids.csv` as the initial values for the mean vectors. By assigning the data points to the nearest center, estimate the initial covariance matrices and prior probabilities in your EM algorithm.

Intial Covariances:

Inde	Type	Size	
0	Array of float64	(2, 2)	[[1.02266303 -0.30081551] [-0.30081551 2.5368769]]
1	Array of float64	(2, 2)	[[1.16271542 -0.08591596] [-0.08591596 0.12053158]]
2	Array of float64	(2, 2)	[[3.99588953 0.9653253] [0.9653253 0.84648858]]
3	Array of float64	(2, 2)	[[2.39175488 0.72739619] [0.72739619 1.34677879]]
4	Array of float64	(2, 2)	[[0.40934714 0.13160001] [0.13160001 0.34698903]]

Prior Probabilities:

priors_probs - List (5 elements)

Inde	Type	Size	
0	float	1	0.19
1	float	1	0.06
2	float	1	0.31333333333333335
3	float	1	0.32
4	float	1	0.11666666666666667

3. After the initialization step, run your EM algorithm for 100 iterations. Report the mean vectors your EM algorithm finds. Your results should be similar to the following matrix.

	0	1			
0	-2.04419	-2.69777		[,1]	[,2]
1	2.48874	2.67687	##	[1,]	-2.0441920 -2.69776844
2	2.66222	-2.30911	##	[2,]	2.6622246 -2.30911081
3	0.155352	0.0577383	##	[3,]	2.4887435 2.67687075
4	-2.67592	2.44659	##	[4,]	-2.6759195 2.44658904
			##	[5,]	0.1553517 0.05773829

I got the same results but in different order of the clusters.

4. Draw the clustering result obtained by your EM algorithm by coloring each cluster with a different color. You should also draw the original Gaussian densities you use to generate data points and the Gaussian densities your EM algorithm finds with dashed and solid lines, respectively. Draw these Gaussian densities where their values are equal to 0.05.

