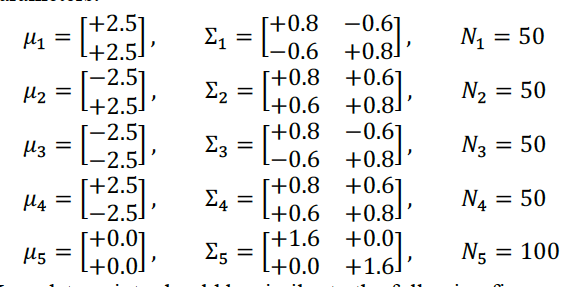
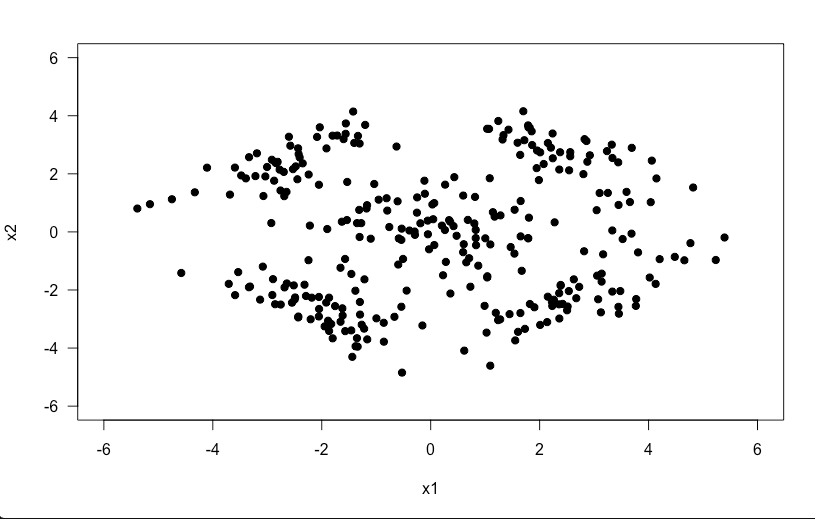
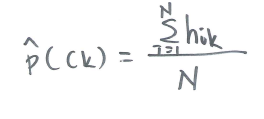
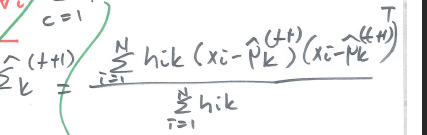
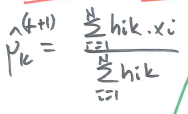
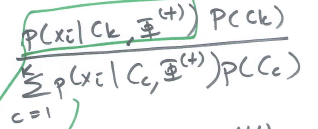
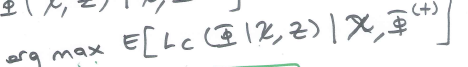
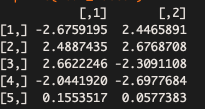
# COMP 421: Introduction to Machine Learning Homework #6

 Initially, I have generated random data points from 5 different bivariate Gaussian densities. These densities were given as in the figure. After generating the data points, I have plotted these data points and I got the results in the figure. To initialize my EM algorithm from a good enough position I have run two iterations of k-means clustering with k being 5 for our case as there are 5 clusters that can be observed from the plot. The 2 iterations are same with the lab codes where we are using the minimum euclidian distances to get our centroids, in this case these centroids will be our means. After the 2 iterations I have defined my new initial covariance matrices and gaussian matrices for each different cluster and also I have calculated the prior probabilities for the EM algorithm. After these initializations, I have run my algorithm for 100 times as requested in the homework and throughout these iterations new mean vectors, new prior probabilities and new covariance matrices as well as new gaussian distributions are calculated according to the given figures.

Then, I have calculated the expectation Step with these given parameters and gaussian as given in the figure and stored them in a 300x5 matrices so that in the end I could find the maximum of these probabilities to get my predictions about the data as the algorithm suggests.



Afterwards, I have printed out my mean vectors however, the my row orders of mean vectors are different from the homework description yet the values that I am getting are the same with the description. This is because my first row of mean vectors indicating the top left cluster but the description’s first row indicates the bottom left cluster. Therefore, this difference holds nothing but only indicating my ordering of the mean vectors are different but the values I get are same. Later on, I have draw the colored the clusters accordingly and draw the contour diagrams for these clusters. For that I have used the lab 3’s drawing codes which helped me to get the diagram in the end. The solid lines are indicating the gaussian densities that I got intially, before the algorithm, and the dashed lines are indicating the gaussian densities that are resulted from my EM algorithm. Although, the whole code is running slower compared to my other homeworks, due to lots of gaussian density calculations throughout the 100 iterations and for the contour diagrams, it works correctly. The end figure from the contour is given below.

