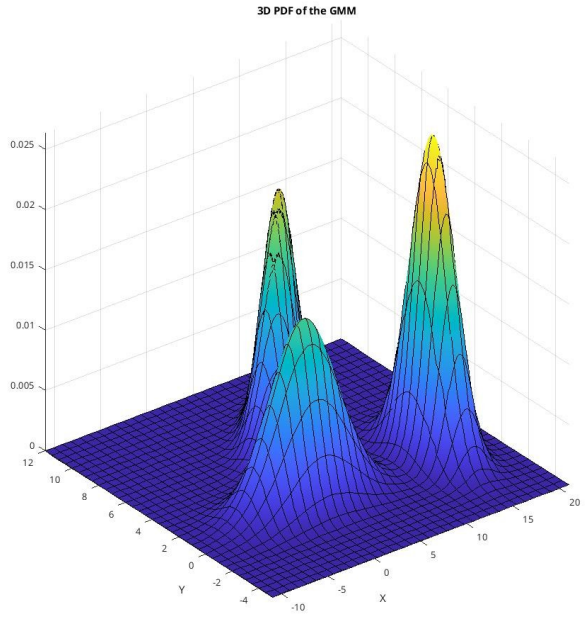
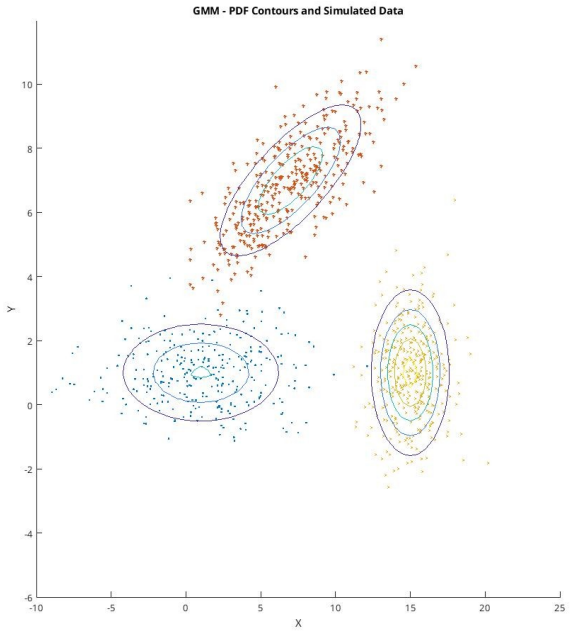
Exercise A





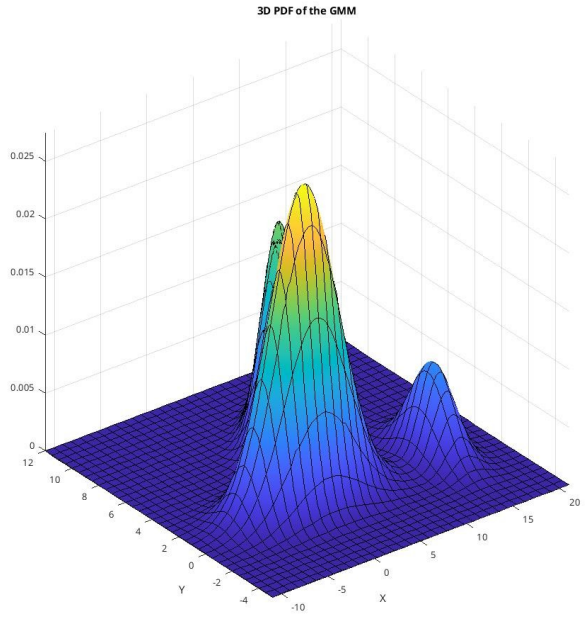
In this exercise, I have used matlab's "gmdistribution" method to create a a distribution. This method takes mean, covariance and prior probabilities(optional) as input and return a "gmdistrbution object" which contains properties of the distribution. This object can be used together with random method to generate samples.

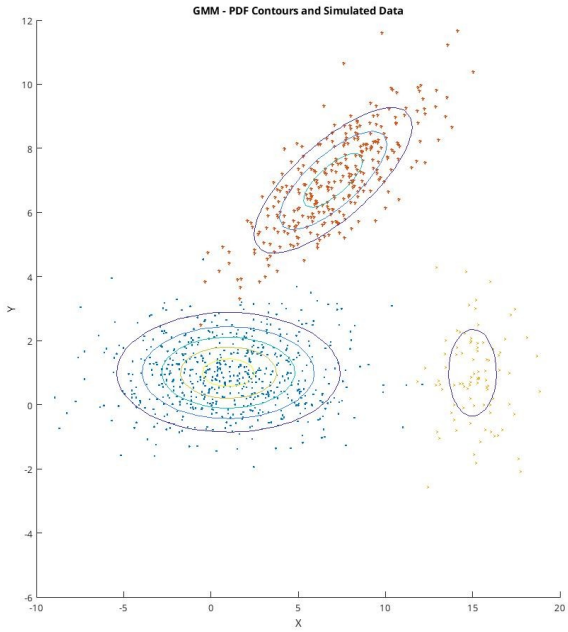
Then I created a parametric function of the pdf using "gmdistributon" object and plotted a 3d surface plot using "fsurf" method.

For second plot, I used "scatter" method multiple times to plot each class in same figure and added contours of pdf using "fcontour" method.

Code for this exercise can be found in CE_A.m

Exercise B



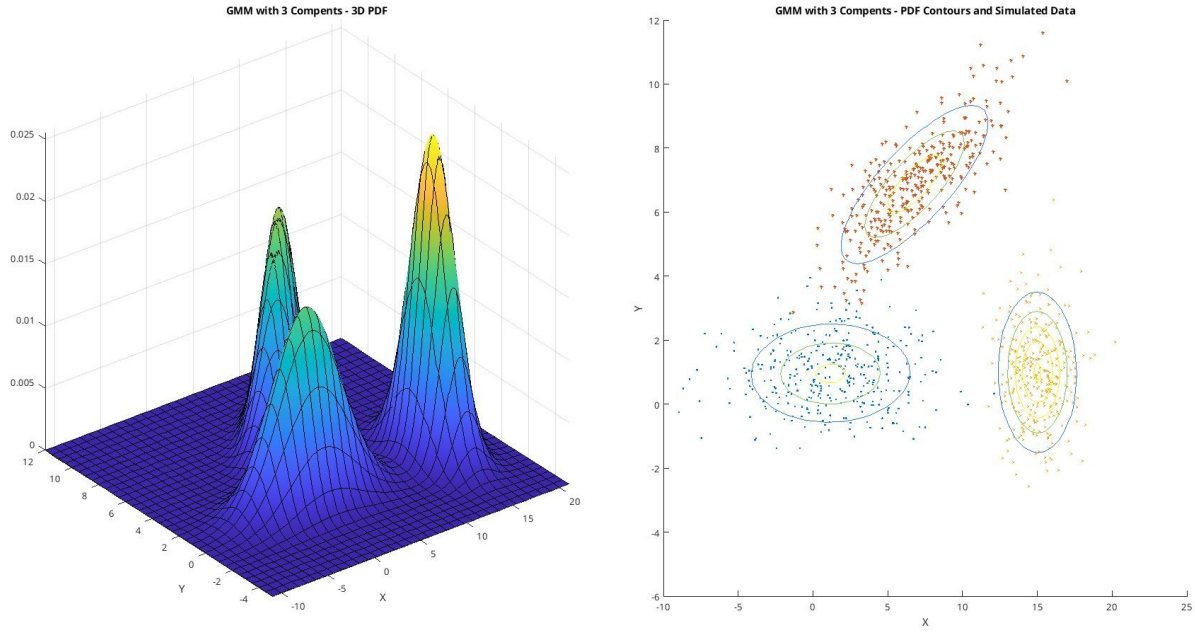


In this exercise I have followed same steps as the exercise A. Only diffrence is I added priori probabilites of classes as parameter to the “gmdistribution” method.

Code for this exercise can be found in CE_B.m

Exercise C

C1 Plot - GMM with 3 Components



GMM with 3 Components - Report

Gaussian mixture distribution with 3 components in 2 dimensions
Component 1:
Mixing proportion: 0.321691
Mean: 15.0093 0.9946
Sigma:
2.0630 -0.0293
-0.0293 1.9223

Component 2:
Mixing proportion: 0.332074
Mean: 6.7907 6.8667
Sigma:
9.4291 3.7220
3.7220 2.3811

Component 3:
Mixing proportion: 0.346236
Mean: 1.1816 0.9550
Sigma:
12.3040 0.0926
0.0926 1.0018

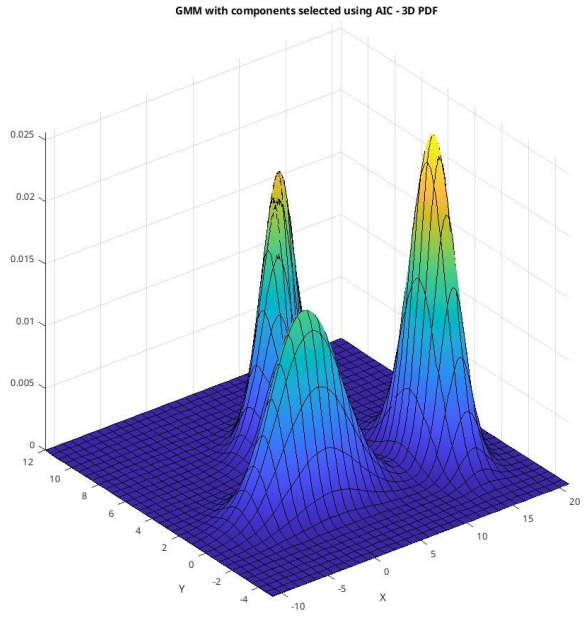
In this exercise, I used “fitgmdist” method to fit a 3 component gaussian mixture distribution to the data generated in the first exercise. This method found the means, covariances and mixin propotions for 3 classes.

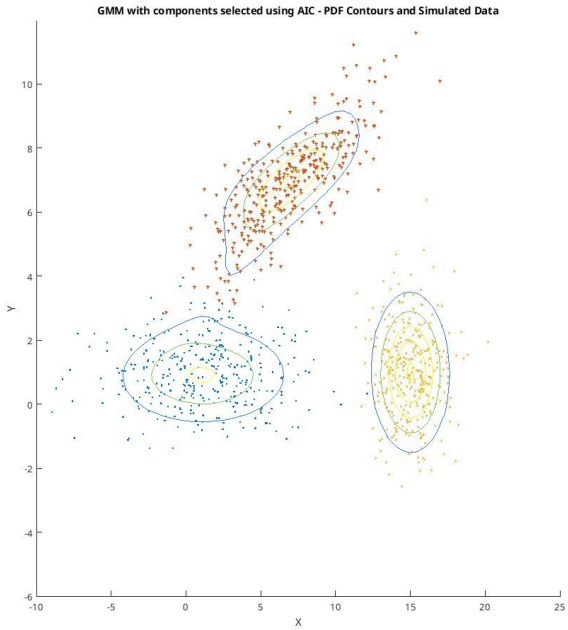
Parameters of the 1. component is very close to the parameters of the 3. class in Exercise A.

Sigma of the 2. component is a bit different than the 2. class in Exercise A but mean and mixin proportion is close to mean and priori probabilities of the 2. class.

3. component matches with the 1. class with minimal diffrence.

C2 Plot – GMM with components selected using AIC





GMM with components selected using AIC - Report

Gaussian mixture distribution with 4 components in 2 dimensions
Component 1:
Mixing proportion: 0.331350
Mean: 1.2051 0.9000
Sigma:
12.6783 0.0360
0.0360 0.9273

Component 2:
Mixing proportion: 0.321597
Mean: 15.0101 0.9948
Sigma:
2.0605 -0.0307
-0.0307 1.9216

Component 3:
Mixing proportion: 0.197295
Mean: 6.9588 7.0790
Sigma:
8.3137 2.0282
2.0282 1.0715

Component 4:
Mixing proportion: 0.149757
Mean: 5.9631 6.1170
Sigma:
13.3590 7.9509
7.9509 5.5697

In this exercise, I have used AIC to select the number of components. I ran the "fitgmdist" method 10 times and each time used component numbers from 1 to 10. Then I measured AIC for each gm distribution model and picked the one with smallest value. Model with smallest AIC value uses 4 components to fit, which is 1 more than the classes used to generate this dataset.

1. component matches with the 1. class, 2. component matches with the 3. class with some difference and similarly 3. component matches with 2. class. The extra component, 4. component, is detected somewhere between 1. class and 3. class, possibly due to proximity of these classes and datapoints created by them.

Code for this exercise can be found in CE_C.m

Exercise D