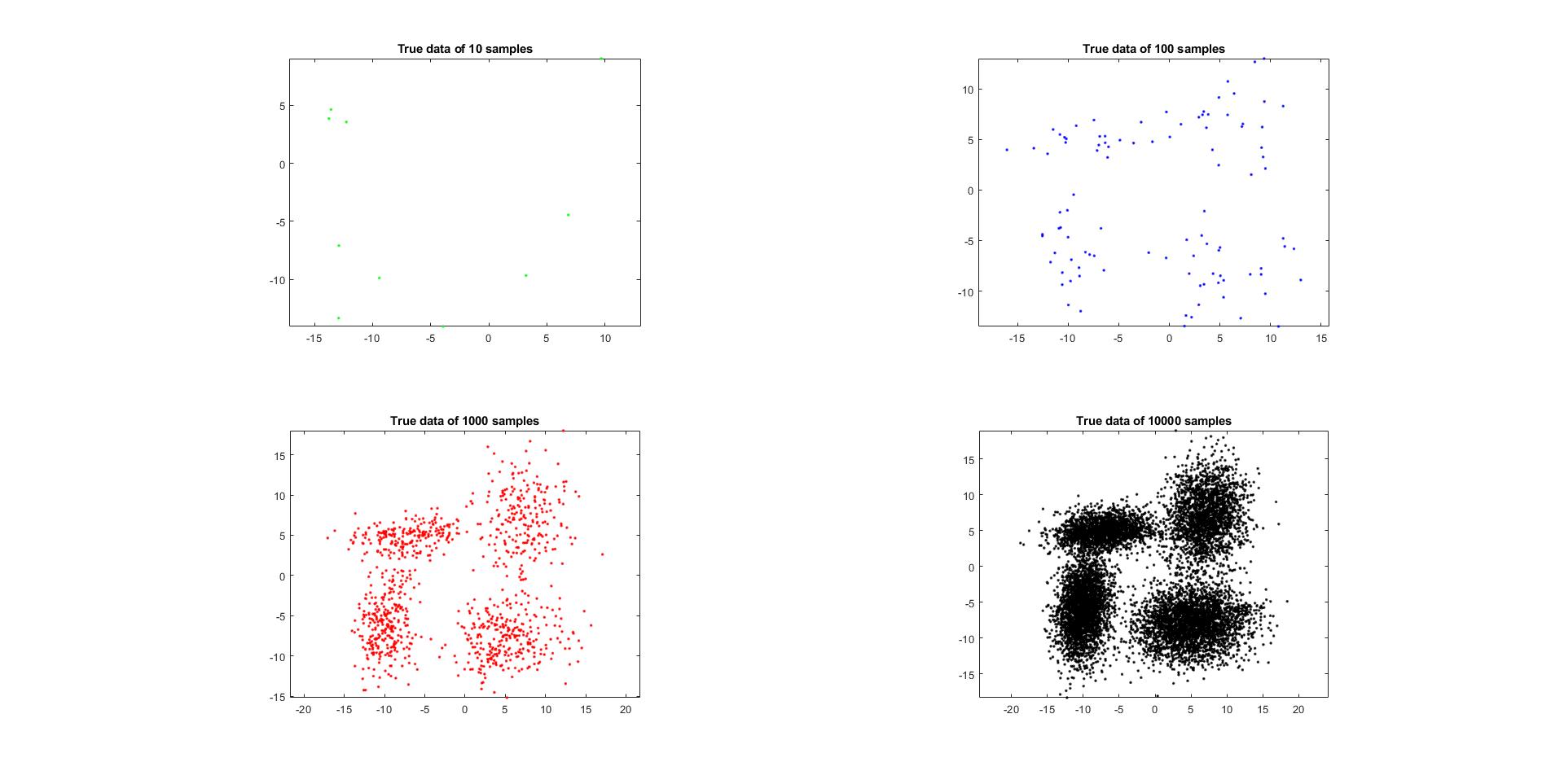
1. Conduct GMM model using 10-fold cross-validation procedure

This plots shows the true data of 10, 100, 1000, 10000 sample of the parameter

Alpha = [0.3, 0.2, 0.2, 0.3]

Mean = [-10 -7 7 5;-6 5 7 -8];

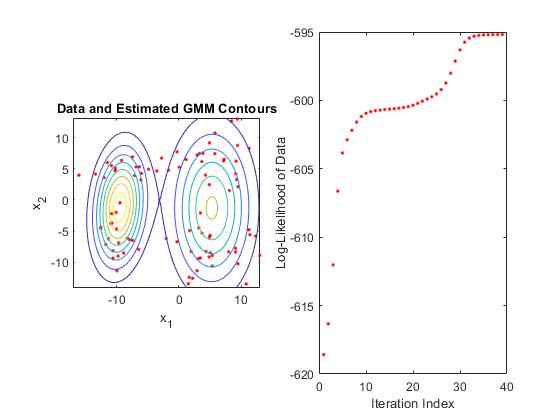
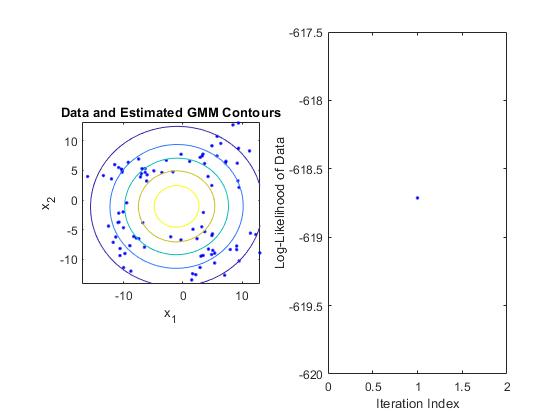
Sigma(:,:,1) = [3 1;1 12];

Sigma (:,:,2) = [12 1;1 2];

Sigma (:,:,3) = [7 1;1 13];

Sigma (:,:,4) = [15 1;1 7];

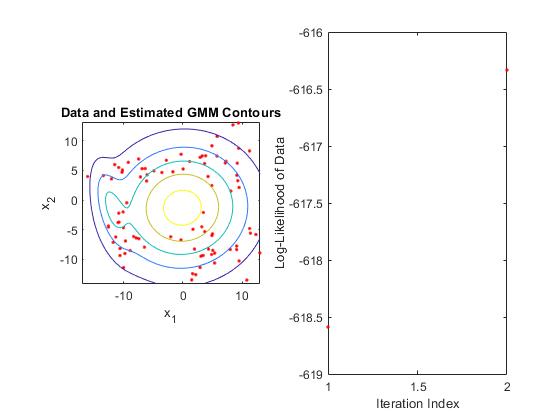
Apply the 10-fold to each sample, that the 10-sample case is not enough for 10-fold cross validation because there are not enough sample for picking the initial EM algorithm. Thus, dismiss the 10-sample case.



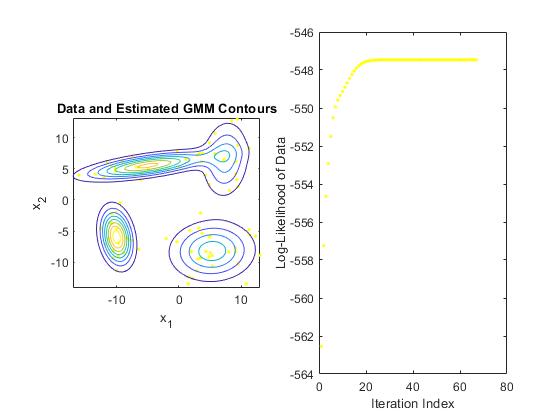
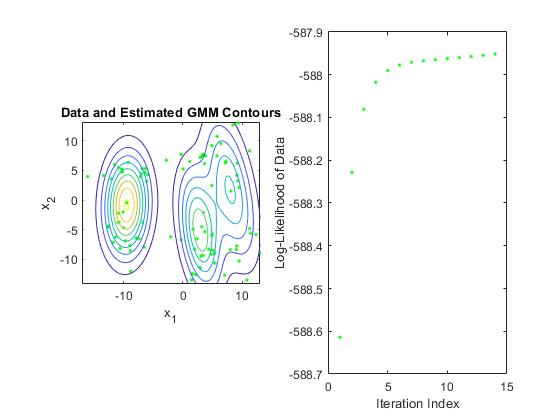
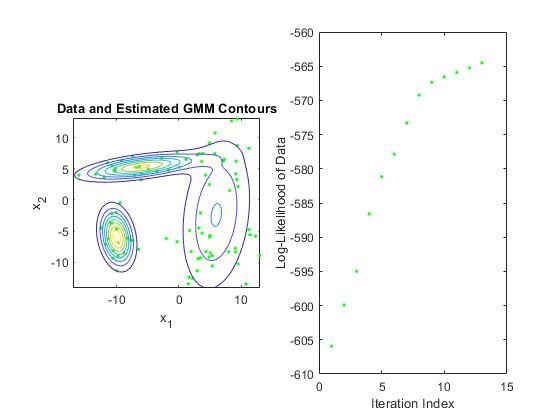
This two plots shows that using 1, 2 as the number of Gaussian training component.

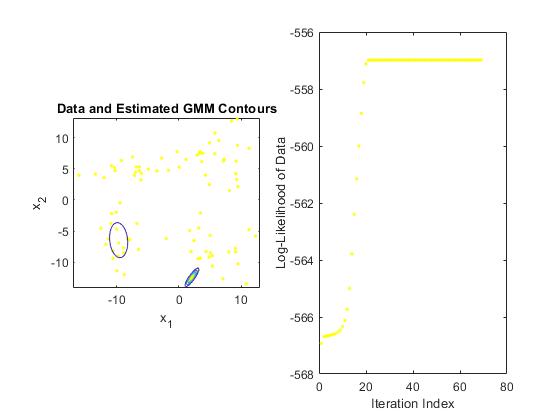
In the order 1, only one point is picked for the EM algorithm. Thus the likelihood will set up immediately to determine the new mean and sigma value.

In second plot, it shows that assuming two component of the Gaussian training.

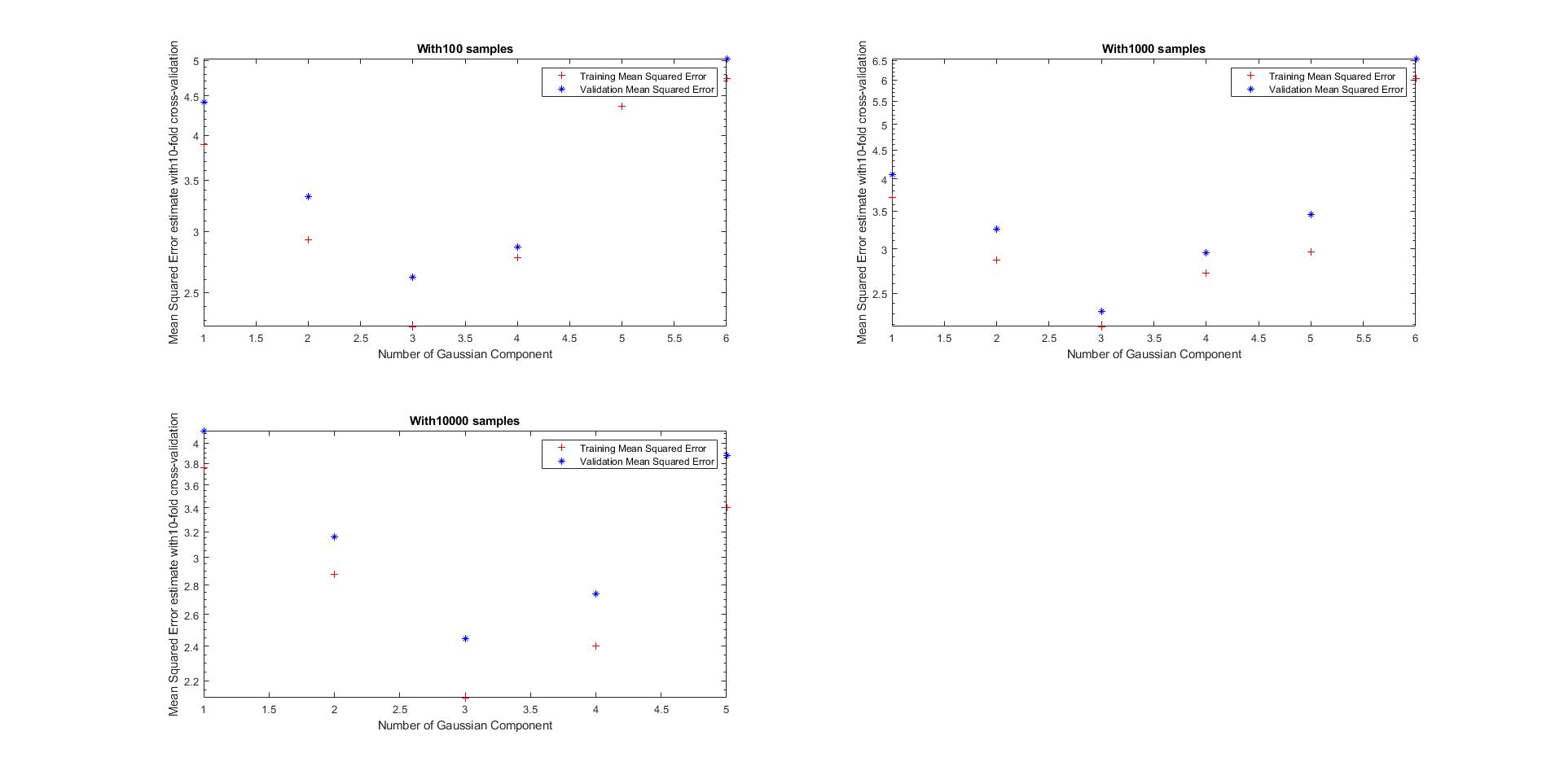
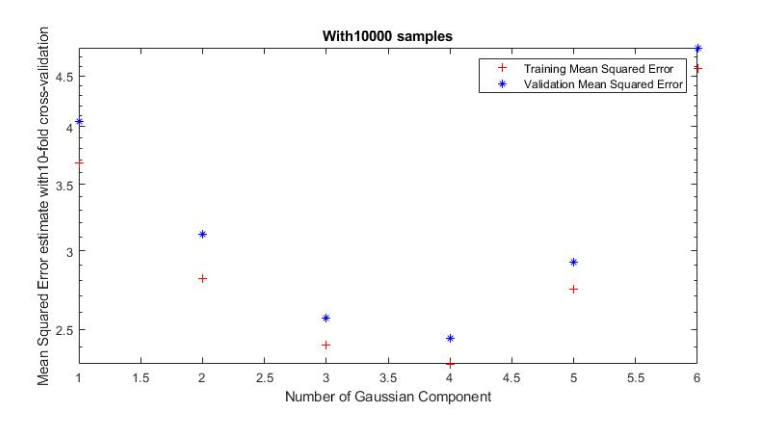


Here shows the case that in the beginning of the order 2 case in one of the 10-fold training. It will finally convergence to two gaussian contour.



Same, we can observe the processing of order 3 and order 4. However, the processing will not stop for the convergence sometimes. The following plot will show the bad case.

This plot shows the bad case of the processing will not terminate. The EM algorithm pick the numbers of the point as the assuming components. However, the picking point will gather up sometimes. This means the initial point is picked up from the same true GMM group. Thus, the processing will not finding the best parameter sometimes. Thus, setting a threshold for determine the bad choosing case to continue on the next block training.

By the observation and the testing mentioned before, we could calculate the mean squared error of each number of Gaussian guessing. The following plots show the result of the mean squared error of different samples and different order.

According to the plots showed above. When we have 100 samples. The best guessing is setting Gaussian component number as 3. 1000 samples case is the same. 10000 sample case is determined as 4 is the best number of the component. We can observe that the training result of error is decreasing when the samples number increase. Finally, we determine the best order will be around 3 to 4 because the true number of Gaussian components is 4. The result will not be more than 4 is because more parameter has determined, more error will be detected when the assuming number is large than true number.

1. Training for the classifier with MAP. Fisher LDA classifier and original labels with 999 2 dimensional smaples.