

CS5691: Assignment 3

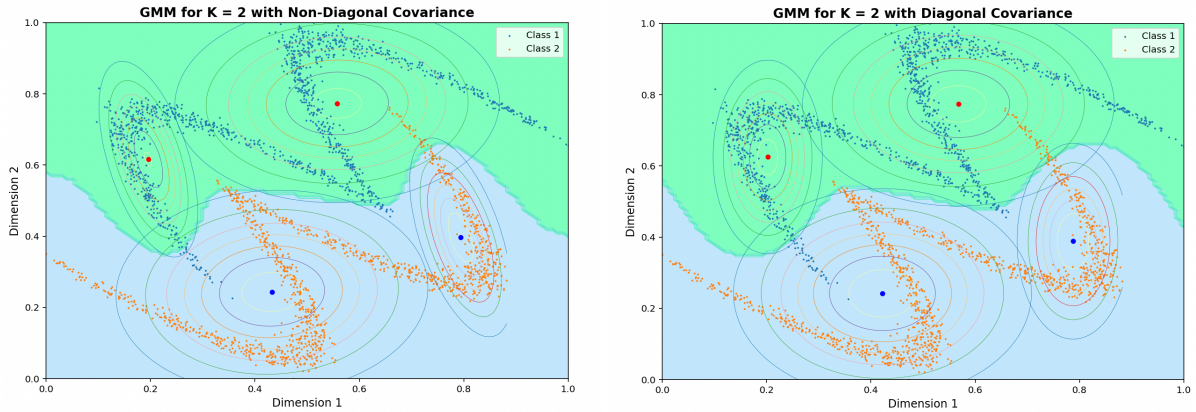
Akshat Meena (CS19B052)
Shubham Patel (ME19B170)

A. K-means, GMM

1. Synthetic Dataset

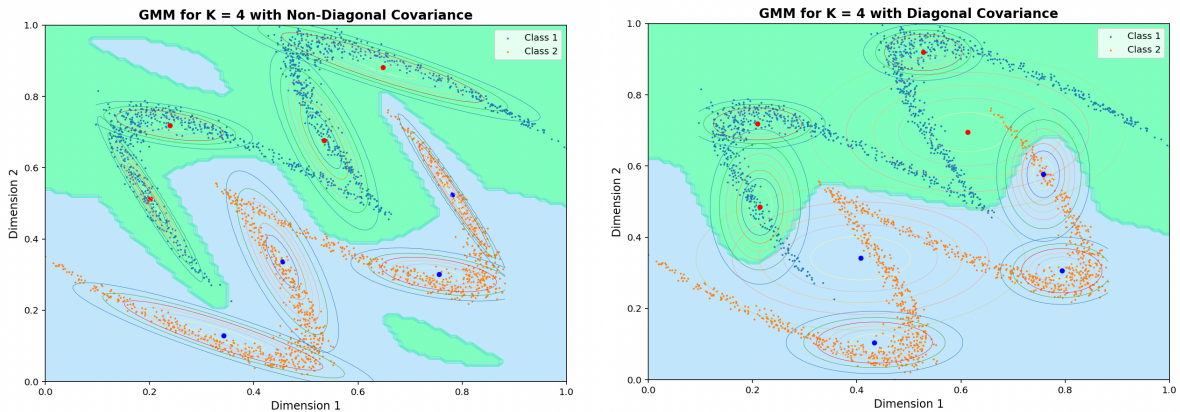
- First we do the observations for different number of clusters ($K=2,4,8,16$) for both diagonal and non diagonal covariance matrices.
- We start with plotting the GMMs for $K=2$, we observe that there is not much difference in the decision surface for both the covariance matrices. For Diagonal matrix it is a little smoother.

Figure 1: GMMs for synthetic data for $K = 2$ and Non-Diagonal and Diagonal Covariance for each class



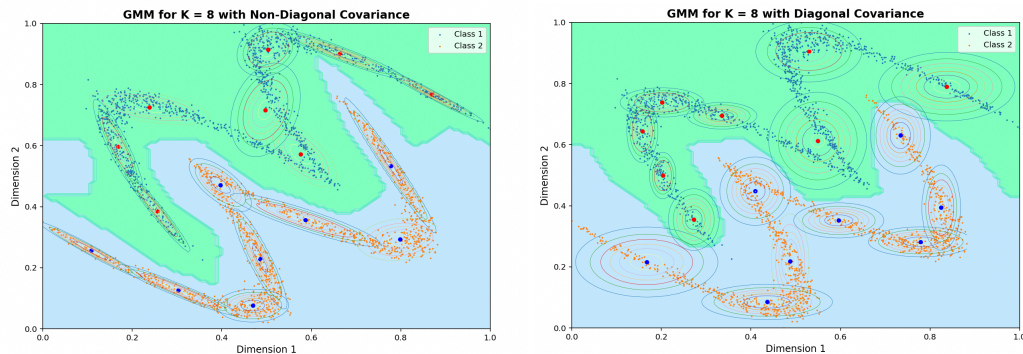
- Next we plot for $K = 4$. We observe that the decision plot for diagonal covariance matrix is better than non-diagonal.
- We observe that the decision surface starts taking better shape.

Figure 2: GMMs for synthetic data for $K = 4$ and Non-Diagonal and Diagonal Covariance for each class



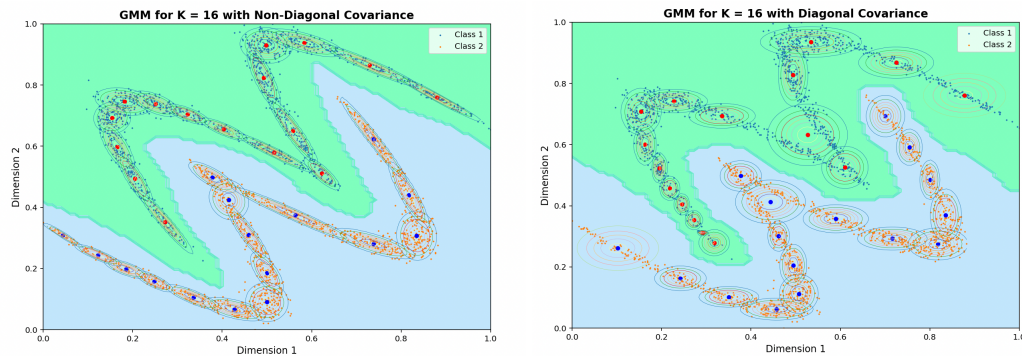
- For $K = 8$, we see the contour of the Gaussian fitting each branch in the entire model properly. The contours of the branches are aligned with the the branch's alignment and each branch completely fit in case of non-diagonal covariance matrix

Figure 3: GMMs for synthetic data for $K = 8$ and Non-Diagonal and Diagonal Covariance for each class



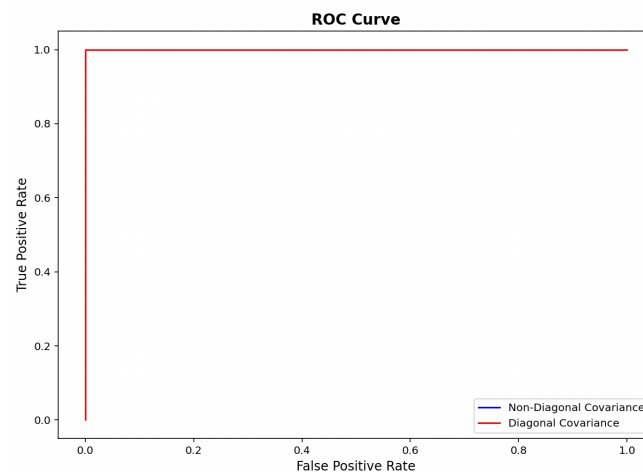
- When we choose $K = 16$, we see that the data starts getting overfitted, more in case of non-diagonal covariance matrix. We also see that the decision surface does very exact classification.

Figure 4: GMMs for synthetic data for $K = 16$ and Non-Diagonal and Diagonal Covariance for each class



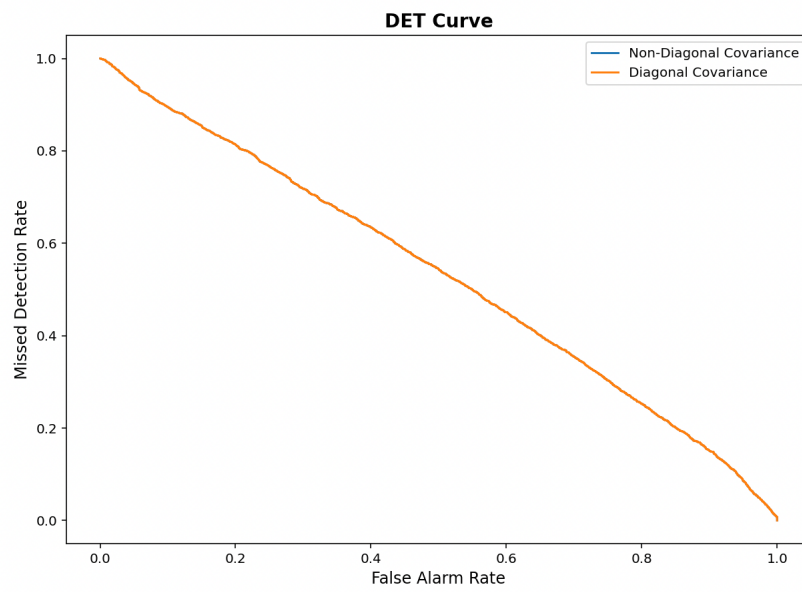
- **ROC Curve**

Figure 5: ROC Curve for Synthetic Data



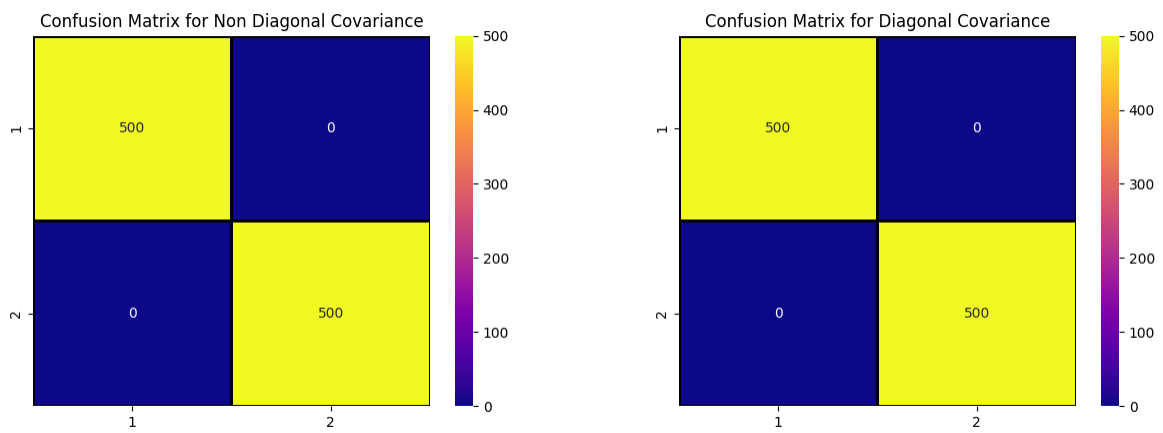
- **DET Curve**

Figure 6: DET Curve for Synthetic Data



- **Confusion Matrix**

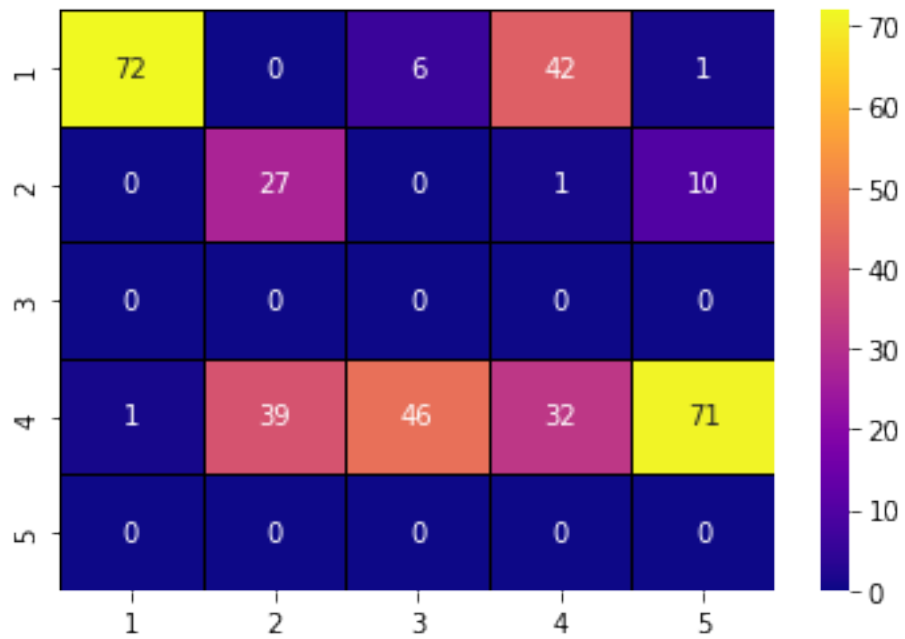
Figure 7: Confusion Matrix for both Non Diagonal and Diagonal Matrix and K = 16



2. Image Dataset

- For Diagonal Covariance Matrix we get accuracy of **37.64%**

Figure 8: Confusion Matrix for K = 4 and 10 iterations for Diagonal Covariance



- We observe that classification is better with diagonal covariance than non diagonal