Mehmet Mert Bezirgan

Engr 421 – homework 1

Problem: classification multivariate random variables.

Solution:

For testing accuracy of our data we prepared our data using numpy with parameters. Our data formed a shape as given

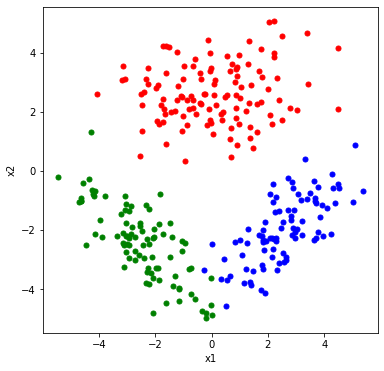


Figure 1 dataset that generated for testing

We got three different classes (1,2,3) and colored them to different colors for visuality. (class 1 : red, class 2 : green, class 3 : blue). After generating data we estimated sample means, sample covariances and prior probabilities.

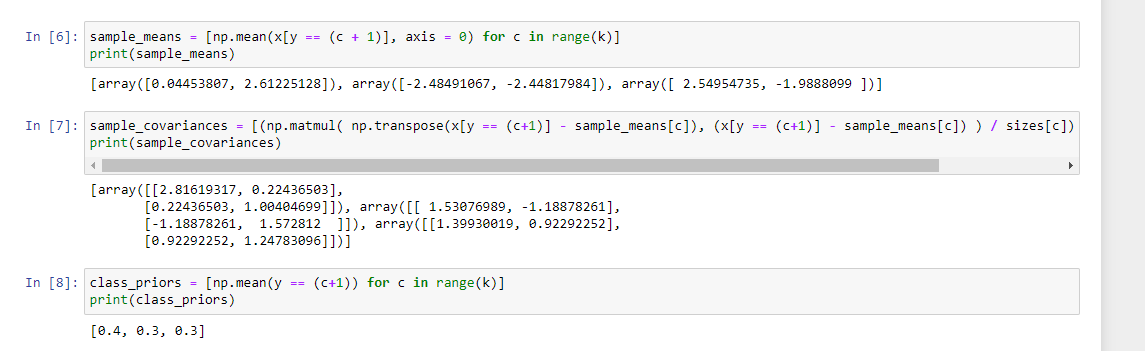


Figure 2 codes and results of calculations of sample means, sample covariances and prior probabilities.

The results of this step show us our data is formed exactly like we have provided the means and covariances before. After calculating sample means, covariances and prior probabilities we need to calculate value of score function for all data points calculated for all 3 classes.

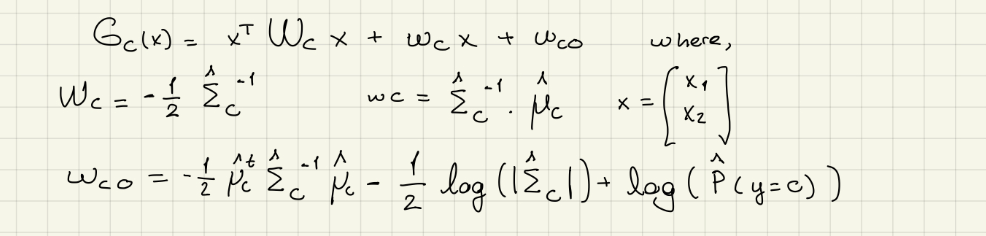


Figure 3 formula used for score function

After calculating scores, we need to pick maximum score for labeling points according to our algorithm’s prediction. After prediction we got our confusion matrix as follows.

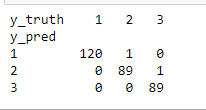


Figure 4 confusion matrix of predicted classes

As we can see from confusion matrix our algorithm made only two false classification out of 300. For visualizing our decision boundaries, we used same approach that we used for lab4. For score function we replaces Wc with 2x2 matrix with entries [Wc11 Wc12 Wc21 Wc22] and formulated it as:

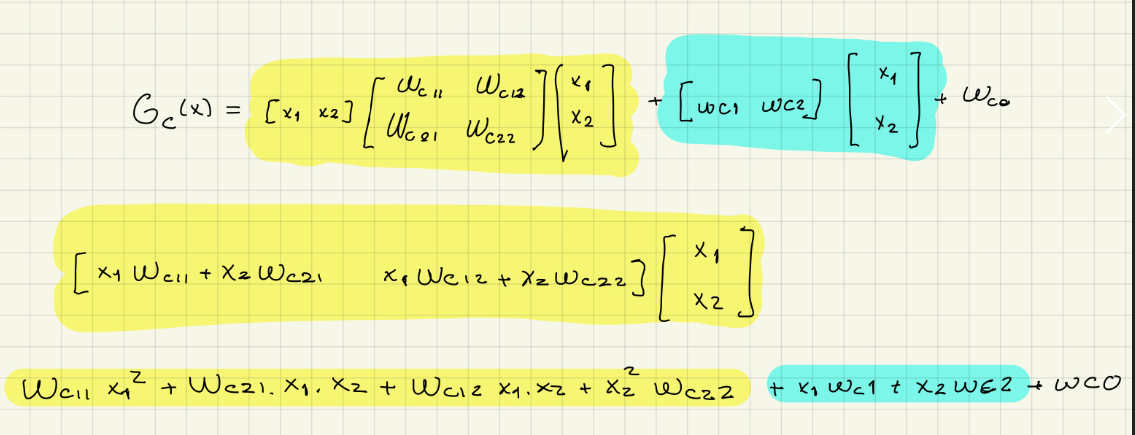


Figure 5 formula used at section 5 (drawing decision boundaries)

Conclusion:

The results of our classification method were accurate on a random generated multivariate random variable. After visualizing decision boundaries, we saw that the separation was accurate visually and the false results was marked.

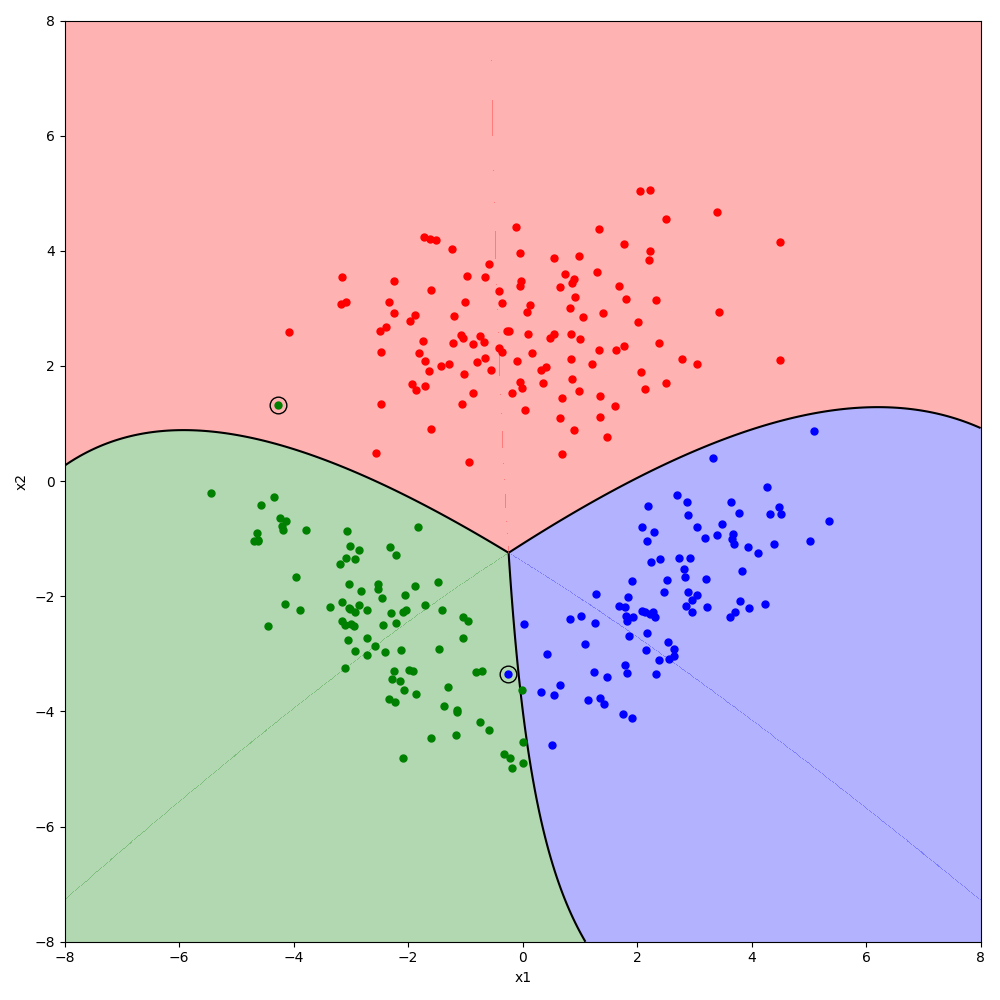


Figure 6 decision boundaries (Note: the colorful mid lines are because of the attribute alpha that we gave to contour function for coloring areas. It is not about any prediction or data that we used, it is just visual)