Pattern Recognition - CS669
Assignment I

Avinav Sanyal, S Naga Koti Reddy, Archit Kumar

November 5, 2017

Discriminant Functions

Pattern classification is a part of machine learning in which we classify objects accordinng to their characteristics. For classfication, we need to make some function which gives the class in which the objects lie. Such functions are called discriminant functions. Discriminant functions are used to find the minimum probability of error in decision making problems. These are also known as decision surfaces.

The univariate normal density is completely specified by two parameters; its mean μ and variance σ^2. The function f_x can be written as $\mathcal{N}(\mu,\sigma^2)$ which says that **x** is distributed normally with mean μ and variance σ^2. Samples from normal distributions tend to cluster about the mean with a spread related to the standard deviation σ. For the multivariate normal density in d dimensions, f_x is written as

$$f_x = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Figure 1:

where:
x is a d-component column vector
μ is the d-component mean vector
Σ is the dxd covariance matrix
$|\Sigma|$ and Σ^{-1} are its determinant and inverse respectively.
$(x-\mu)^t$ denotes the transpose of (x-μ) and

$$\mu = \mathcal{E}[x] = \int_{-\infty}^{\infty} x p(x)\, dx$$

Figure 2:

where the expected value of a vector or a matrix is found by taking the expected value of the individual components.

The covariance matrix Σ is always symmetric and positive definite which means that the determinant of Σ is strictly positive. The diagonal elements σ_{ii} are the variances of the respective x_i (i.e., σ^2), and the off-diagonal elements σij are the covariances of x_i and xj. If x_i and x_j are statistically independent, then $\sigma_{ij} = 0$. If all off-diagonal elements are zero, $p(x)$ reduces to the product of the univariate normal densities for the components of x.

Discriminant functions are used to find the minimum probability of error in decision making problems. In a problem with feature vector $\mathbf{y}$ and state of nature variable w, we can represent the discriminant function as: where $p(\mathbf{Y}|w_i)$ is the conditional probability density function for $\mathbf{Y}$ with w_i being the state of nature, and $P(w_j)$ is the prior probability that nature is in state wj. If we take $p(\mathbf{Y}|w_i)$ as multivariate normal distributions. That is if p($\mathbf{Y}|wi$) $= \mathcal{N}(\mu,\sigma)$, then the discriminant changes to

$$f_x = \frac{1}{(2\pi)^{\frac{d}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

Figure 3: The discriminant function

Linearly Separable Data

A. Scaled identity covariance matrices

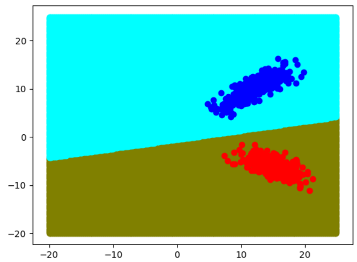


Figure 4: Decision region between class 1,2

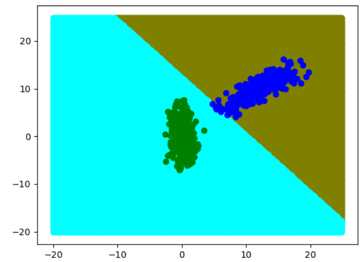


Figure 5: Decision region between class 2,3

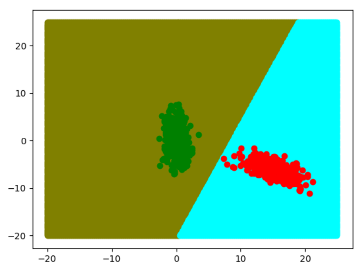


Figure 6: Decision region between class 1,3

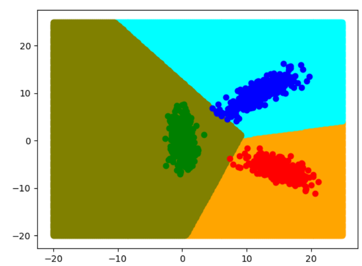


Figure 7: Combined Decision region using scaled covariance matrix

The decision boundaries are linear in nature.

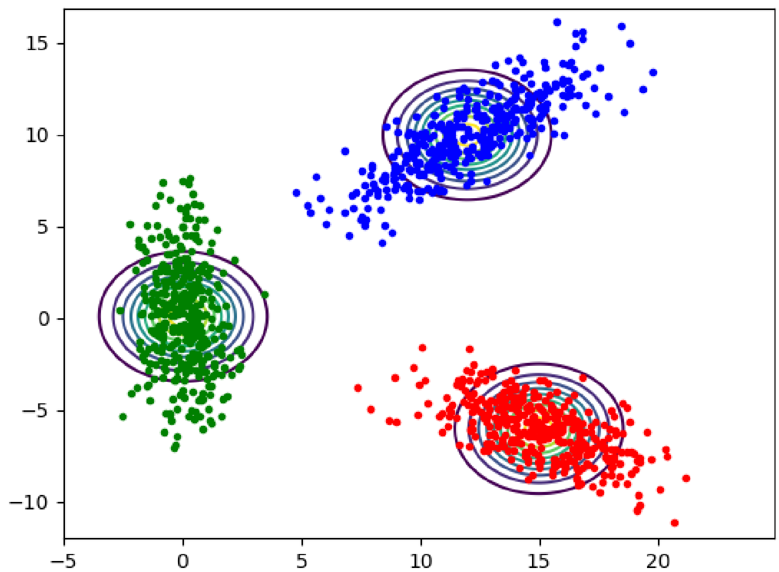


Figure 8: Contour plot

Observations

The covariance matrix is:

$$\begin{bmatrix} 2.77 & 0 \\ 0 & 2.77 \end{bmatrix}$$

As the matrix is scaled I we get circles as contour plots.

Confusion matrix:

$$\begin{bmatrix} 124 & 0 & 1 \\ 0 & 124 & 1 \\ 0 & 0 & 125 \end{bmatrix}$$

Table 1: Results

	Precision	Recall	F-measure
Class 1	0.992	0.976	0.984
Class 2	0.992	0.976	0.984
Class 3	1	0.984	0.992
Mean	0.994	0.979	0.987

Classification accuracy: 0.9946666666666667

B. Same full covariance matrices

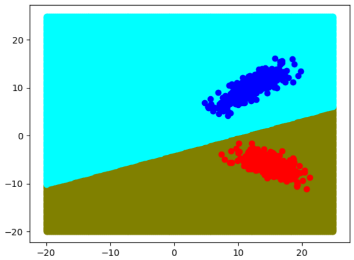


Figure 9: Decision region using same full covariance matrices

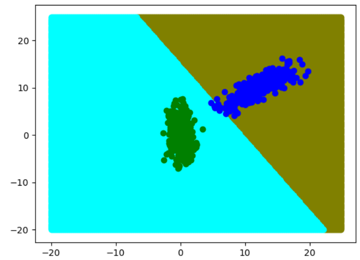


Figure 10: Decision region using same full covariance matrices

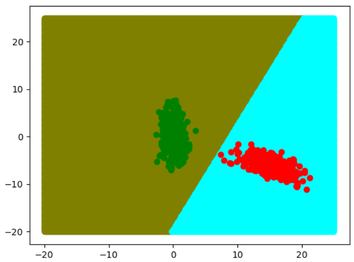


Figure 11: Decision region using same full covariance matrices

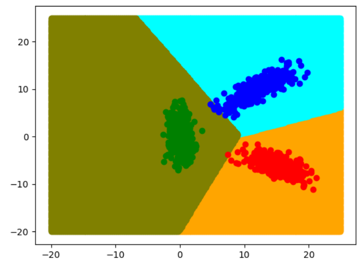


Figure 12: All classes with same full covariance matrices

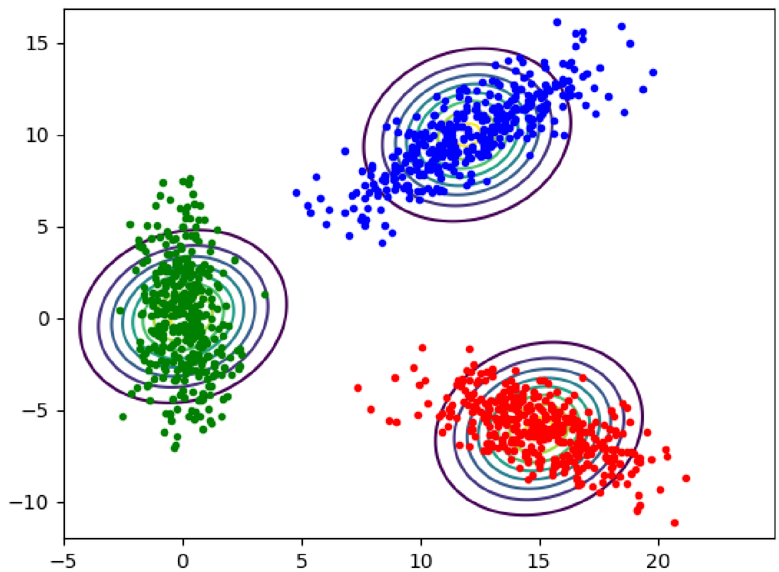


Figure 13: Contour using same full covariance matrices

Observations

The covariance matrix is:

$$\begin{bmatrix} 4.476 & 0.677 \\ 0.677 & 5.25 \end{bmatrix}$$

Variance in Y is some what more than X direction so the data is some what more spread in Y than in X direction.
As the non diagonal elements are positive that is $Cov(XY)$ and very small, so the plots are slightly bent towards right.

Confusion matrix:

$$\begin{bmatrix} 124 & 0 & 1 \\ 0 & 124 & 1 \\ 0 & 0 & 125 \end{bmatrix}$$

Table 2: Results

	Precision	Recall	F-measure
Class 1	0.992	0.976	0.984
Class 2	0.992	0.976	0.984
Class 3	1	0.984	0.992
Mean	0.995	0.979	0.987

Classification accuracy: 0.9946666666666667

C. All diagonal covariance matrices

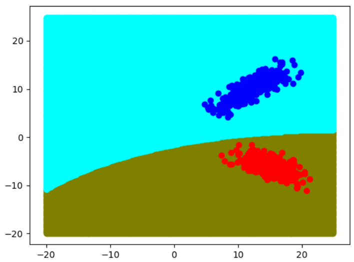


Figure 14: Decision region using all diagonal covariance matrices - class 1

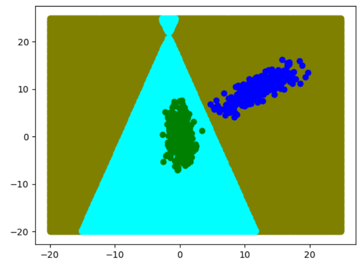


Figure 15: Decision region using all diagonal covariance matrices - class 2

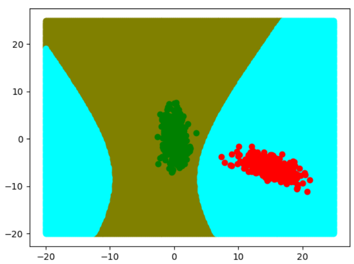
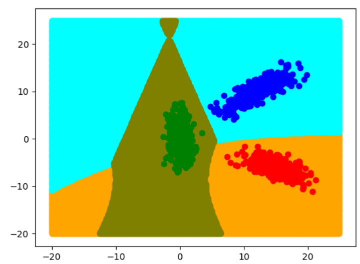


Figure 16: Decision region using all diagonal covariance matrices



0

Figure 17: For both diagonal covariance matrices

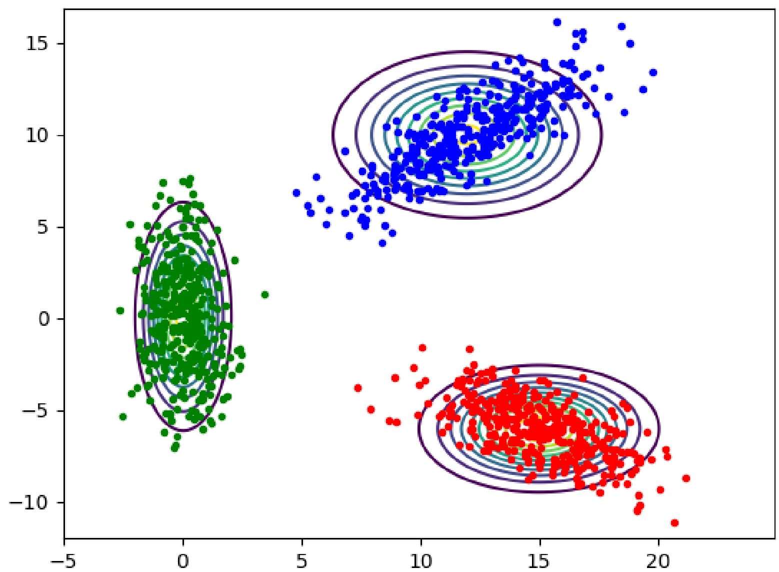


Figure 18: Contour using all diagonal covariance matrices

Observations

Covariance for class 1:

$$\begin{bmatrix} 5.351 & 0 \\ 0 & 2.5148 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.

Covariance for class 2:

$$\begin{bmatrix} 7.162 & 0 \\ 0 & 4.602 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.

Covariance for class 3:

$$\begin{bmatrix} 0.915 & 0 \\ 0 & 8.632 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is much more spread in Y than in X direction.

Confusion matrix:

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Table 3: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Mean	1	1	1

Classification accuracy: 1.0

D. All covariance matrices

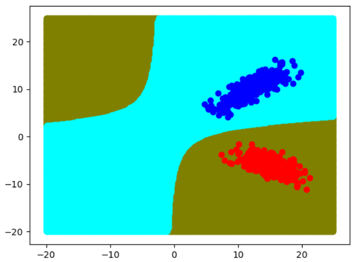


Figure 19: Decision region using all full covariance matrices

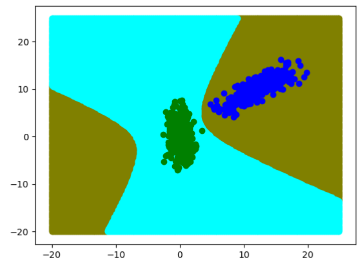


Figure 20: Decision region using full covariance matrices

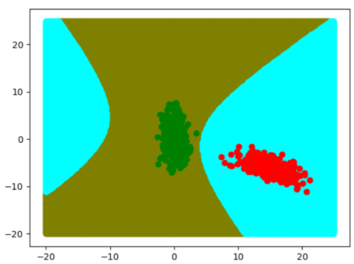


Figure 21: Decision region using full covariance matrices

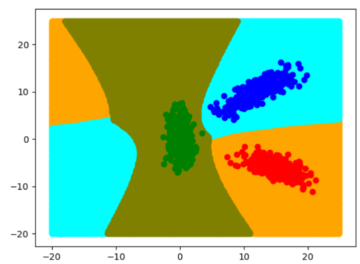


Figure 22: Decision region using full covariance matrices

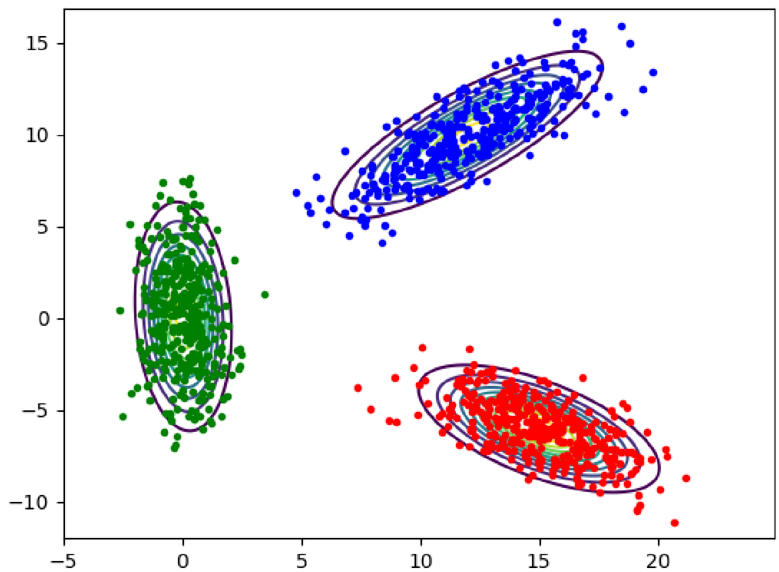


Figure 23: Contour

Observations

Covariance for Class 1:

$$\begin{bmatrix} 5.3513 & -2.254 \\ -2.254 & 2.515 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.
Here as the non diagonal elements are negative that is COV(XY) and very small so the plots axis is bent towards left.

Covariance for Class 2:

$$\begin{bmatrix} 7.162 & 4.668 \\ 4.668 & 4.602 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.
As the non diagonal elements are positive that is COV(XY) and so the plots are slightly bent towards right.

Covariance for Class 3:

$$\begin{bmatrix} 0.915 & -0.381 \\ -0.381 & 8.632 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is more spread in Y than in X direction.
Here as the non diagonal elements are negative that is COV(XY) and very small so the plots axis is bent towards left.

Confusion matrix:

$$\begin{bmatrix} 124 & 0 & 1 \\ 0 & 124 & 1 \\ 0 & 0 & 125 \end{bmatrix}$$

Table 4: Results

	Precision	Recall	F-measure
Class 1	0.992	0.976	0.984
Class 2	0.992	0.976	0.984
Class 3	1	0.984	0.992
Mean	0.994	0.979	0.987

Classification accuracy: 0.9946666666666667

Non Linearly Separable Data

A. All scaled identity covariance matrices

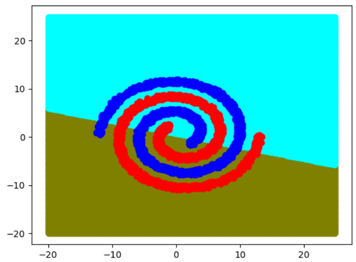


Figure 24: Decision region between classes 1 and 2

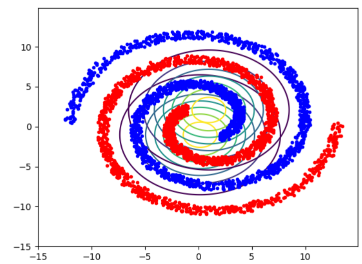


Figure 25: Contour

Observations

Covariance matrix:

$$\begin{bmatrix} 13.849 & 0 \\ 0 & 13.849 \end{bmatrix}$$

As the covariance matrix is scaled I, we get circles as contour plots.

Confusion matrix:

$$\begin{bmatrix} 395 & 217 \\ 236 & 376 \end{bmatrix}$$

Table 5: Results

	Precision	Recall	F-measure
Class 1	0.6454	0.6661	0.6556
Class 2	0.6144	0.6341	0.6241
Mean	0.6299	0.65	0.6398

Classification accuracy: 0.6299019607843137

B. Same full covariance matrices

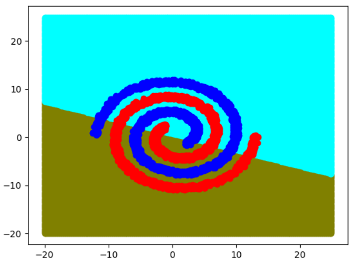


Figure 26: Decision region between classes 1 and 2

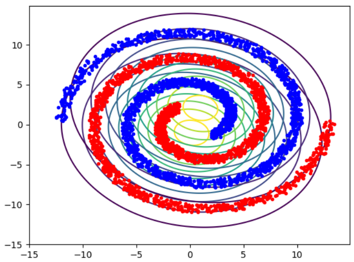


Figure 27: Using same full covariance matrices

Observations

Covariance matrix:

$$\begin{bmatrix} 31.507 & -3.000 \\ -3.000 & 29.888 \end{bmatrix}$$

Here the diagonal elements are almost same so the contour has same width in both x and y axis. The non diagonal elements are negative so the contour plots axis are bent left side.

Confusion matrix:

$$\begin{bmatrix} 395 & 217 \\ 236 & 376 \end{bmatrix}$$

Table 6: Results

	Precision	Recall	F-measure
Class 1	0.645	0.666	0.655
Class 2	0.614	0.634	0.624
Mean	0.63	0.65	0.64

Classification accuracy: 0.6299019607843137

C. All diagonal covariance matrices

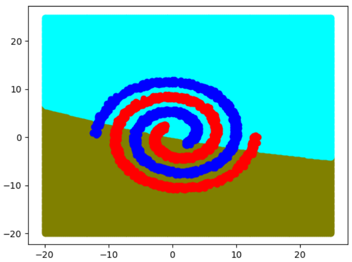


Figure 28: Decision region between classes 1 and 2

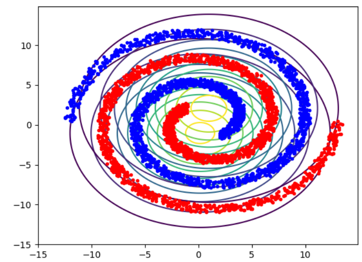


Figure 29: Contour

Observations

Covariance matrix for class 1:

$$\begin{bmatrix} 31.759 & 0 \\ 0 & 30.112 \end{bmatrix}$$

Here the diagonal elements are almost same so the contour has same width in both x and y axis.
Covariance matrix for class 2:

$$\begin{bmatrix} 31.255 & 0 \\ 0 & 29.664 \end{bmatrix}$$

Here the diagonal elements are almost same so the contour has same width in both x and y axis.

Confusion matrix:

$$\begin{bmatrix} 394 & 218 \\ 232 & 380 \end{bmatrix}$$

Table 7: Results

	Precision	Recall	F-measure
Class 1	0.644	0.659	0.651
Class 2	0.62	0.635	0.628
Mean	0.632	0.647	0.64

Classification accuracy: 0.6323529411764706

D. All full covariance matrices

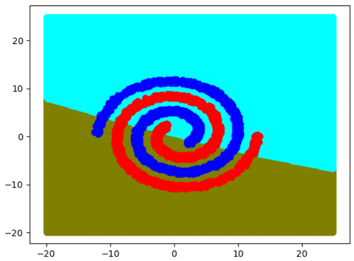


Figure 30: Decision region between classes 1 and 2

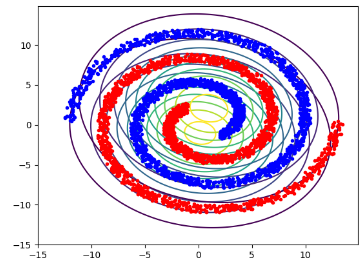


Figure 31: Contour

Observations

Covariance matrix for class 1:

$$\begin{bmatrix} 31.759 & -2.863 \\ -2.863 & 30.112 \end{bmatrix}$$

Here the diagonal elements are almost same so the contour has same width in both x and y axis. The non diagonal elements are negative so the contour plots axes are bent left side. The spread of data is the same along the axes.

Covariance matrix for class 2:

$$\begin{bmatrix} 31.255 & -3.138 \\ -3.138 & 29.664 \end{bmatrix}$$

Here the diagonal elements are almost same so the contour has same width in both x and y axis. The non diagonal elements are negative so the contour plots axes are bent left side.The spread of data is the same along the axes.

Confusion matrix:

$$\begin{bmatrix} 395 & 217 \\ 234 & 378 \end{bmatrix}$$

Table 8: Results

	Precision	Recall	F-measure
Class 1	0.645	0.663	0.654
Class 2	0.617	0.635	0.626
Mean	0.631	0.65	0.64

Classification accuracy: 0.6315359477124183

Real world data

A. All scaled identity matrices

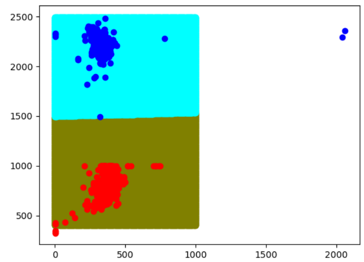


Figure 32: Decision region between classes 1,2

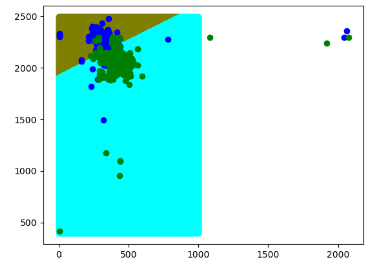


Figure 33: Decision region between classes 2,3

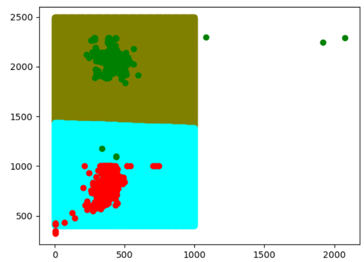


Figure 34: Decision region between classes 1,3

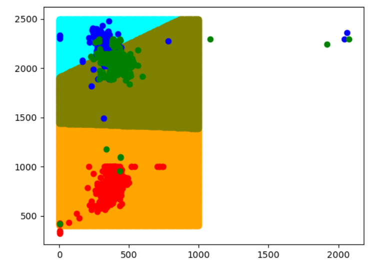


Figure 35: Decision region using same diagonal covariance matrices

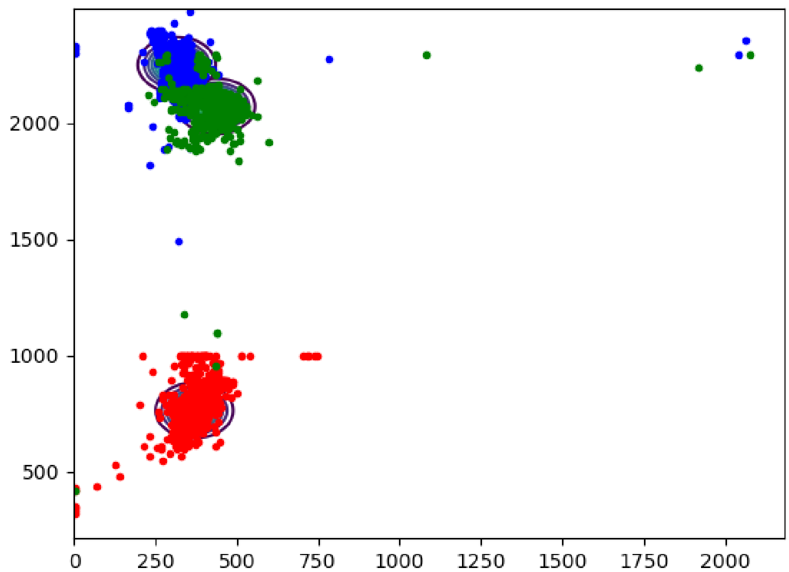


Figure 36: Contour using same diagonal covariance matrices

Observations

The covariance matrix is :

$$\begin{bmatrix} 3025.024 & 0 \\ 0 & 3025.024 \end{bmatrix}$$

As the covariance matrix is scaled I, we get circles as contour plots .

Confusion matrix:

$$\begin{bmatrix} 622 & 0 & 0 \\ 0 & 447 & 1 \\ 18 & 350 & 205 \end{bmatrix}$$

Table 9: Results

	Precision	Recall	F-measure
Class 1	1	0.971	0.985
Class 2	0.997	0.56	0.717
Class 3	0.357	0.995	0.525
Mean	0.785	0.842	0.742

Classification accuracy: 0.7754108338405356

B. All same full covariance matrices

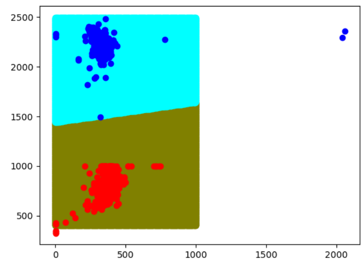


Figure 37: Decision region between classes 1,2

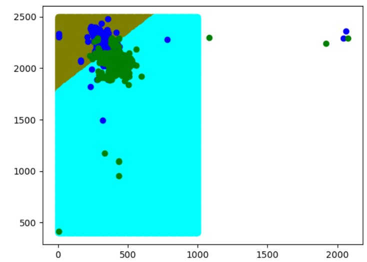


Figure 38: Decision region between classes 2,3

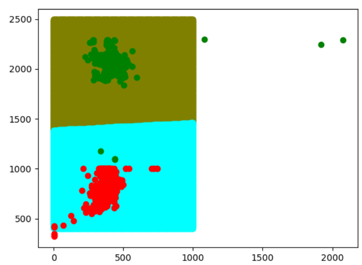


Figure 39: Decision region between classes 1,3

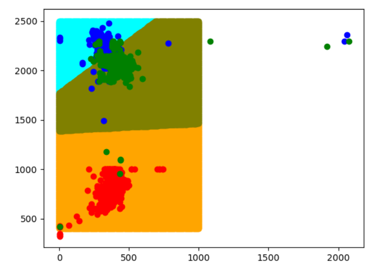


Figure 40: Combined Decision region

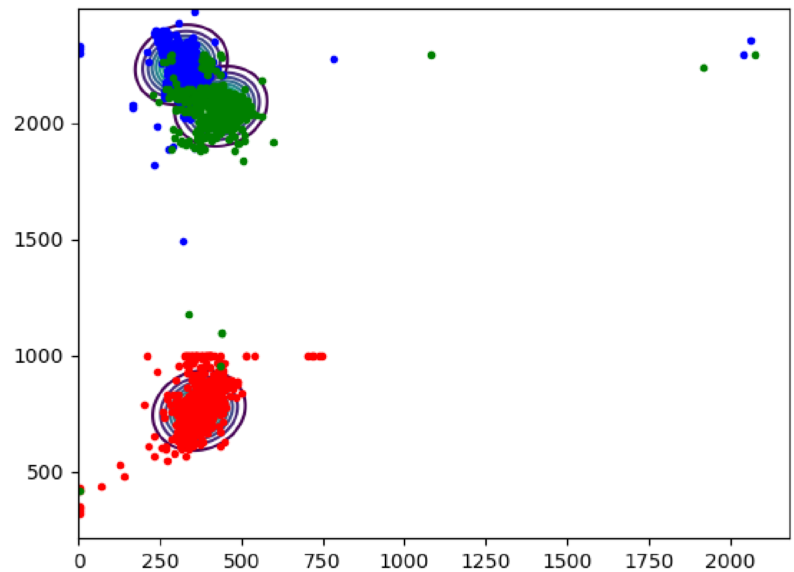


Figure 41: Contour

Observations

The covariance matrix is:

$$\begin{bmatrix} 4348.124 & 647.745 \\ 647.7459 & 6456.4811 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.
As the non diagonal elements($Cov(XY)$)are positive and so the plots are slightly bent towards right.

Confusion matrix:

$$\begin{bmatrix} 622 & 0 & 0 \\ 0 & 447 & 1 \\ 18 & 346 & 209 \end{bmatrix}$$

Table 10: Results

	Precision	Recall	F-measure
Class 1	1	0.971	0.985
Class 2	0.997	0.436	0.606
Class 3	0.364	0.995	0.533
Mean	0.787	0.807	0.708

Classification accuracy: 0.7778454047474133

C. All diagonal covariance matrices

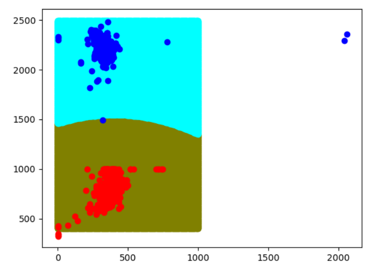


Figure 42: Decision region between classes 1,2

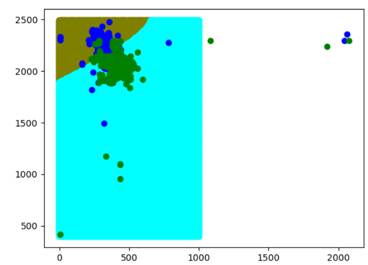


Figure 43: Decision region between classes 2,3

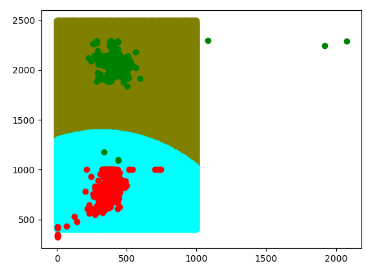


Figure 44: Decision region between classes 1,3

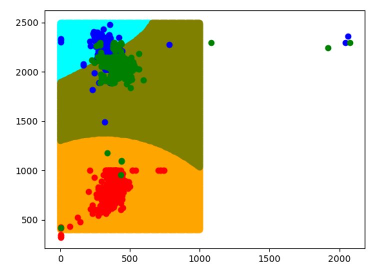


Figure 45: Decision region using diagonal covariance matrices

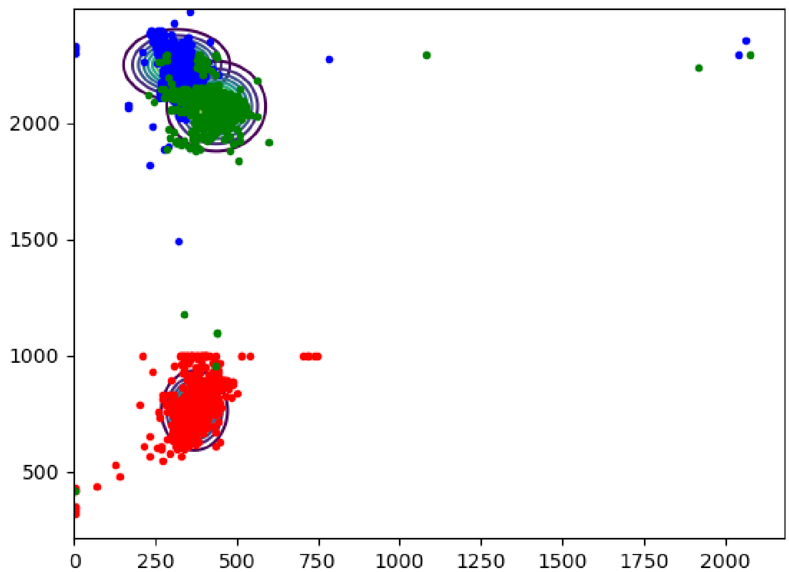


Figure 46: Contour

Observations

Covariance matrix for class 1:

$$\begin{bmatrix} 2145.311 & 0 \\ 0 & 6128.158 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is more spread in Y than in X direction.

Covariance matrix for class 2:

$$\begin{bmatrix} 5868.504 & 0 \\ 0 & 5105.129 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y.

Covariance matrix for class 3:

$$\begin{bmatrix} 5030.557 & 0 \\ 0 & 8136.156 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is more spread in Y than in X direction.

Confusion matrix:

$$\begin{bmatrix} 622 & 0 & 0 \\ 0 & 447 & 1 \\ 17 & 347 & 209 \end{bmatrix}$$

Table 11: Results

	Precision	Recall	F-measure
Class 1	1	0.973	0.987
Class 2	0.997	0.563	0.719
Class 3	0.364	0.995	0.533
Mean	0.787	0.847	0.746

Classification accuracy: 0.7778454047474133

D. All full covariance matrices

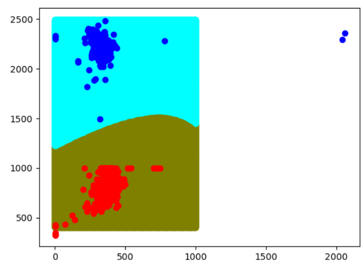


Figure 47: Decision region between classes 1,2

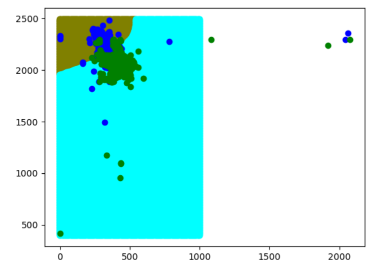


Figure 48: Decision region between classes 2,3

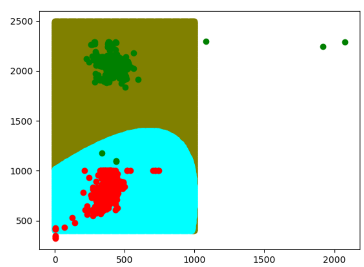


Figure 49: Decision region between classes 1,3

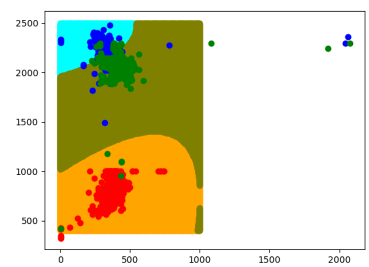


Figure 50: Combined decision region

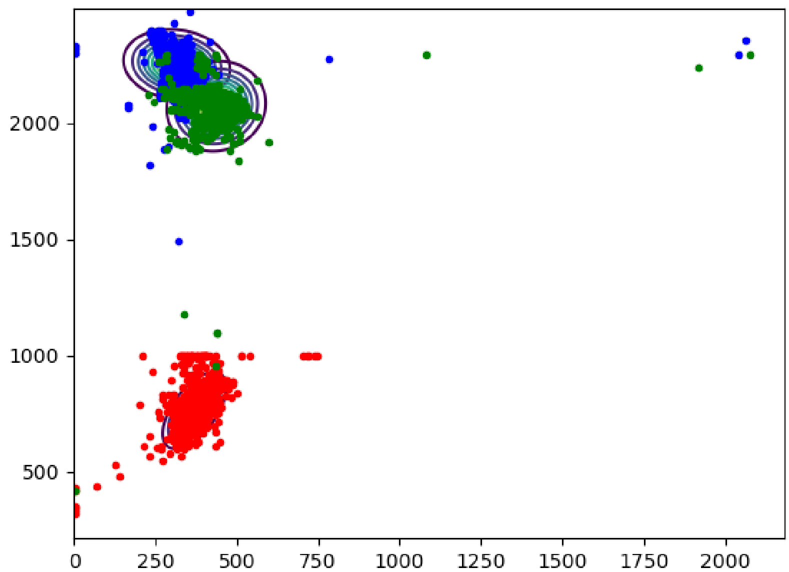


Figure 51: Contour

Observations

Covariance matrix for class 1:

$$\begin{bmatrix} 2145.311 & 2194.074 \\ 2194.074 & 6128.158 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is more spread in Y than in X direction. The non-diagonal elements are positive that is $Cov(XY)$ and so the plots are bent towards right.

Covariance matrix for class 2:

$$\begin{bmatrix} 5868.504 & -689.229 \\ -689.229 & 5105.129 \end{bmatrix}$$

Variance in X is some what more than Y direction so the data is some what more spread in X than in Y direction.
The non diagonal elements are negative so the contour plots axis are bent on left side. The spread of data is the same along the axis.

Covariance matrix for class 3:

$$\begin{bmatrix} 5030.557 & 438.389 \\ 438.389 & 8136.156 \end{bmatrix}$$

Variance in Y is much more than X direction so the data is more spread in Y than in X direction. As the non diagonal elements are positive that is $Cov(XY)$ and so the plots are bent towards right.

Confusion matrix:

$$\begin{bmatrix} 622 & 0 & 0 \\ 0 & 447 & 1 \\ 18 & 345 & 210 \end{bmatrix}$$

Table 12: Results

	Precision	Recall	F-measure
Class 1	1	0.971	0.985
Class 2	0.997	0.564	0.72
Class 3	0.366	0.995	0.535
Mean	0.787	0.843	0.746

Classification accuracy: 0.7784540474741327