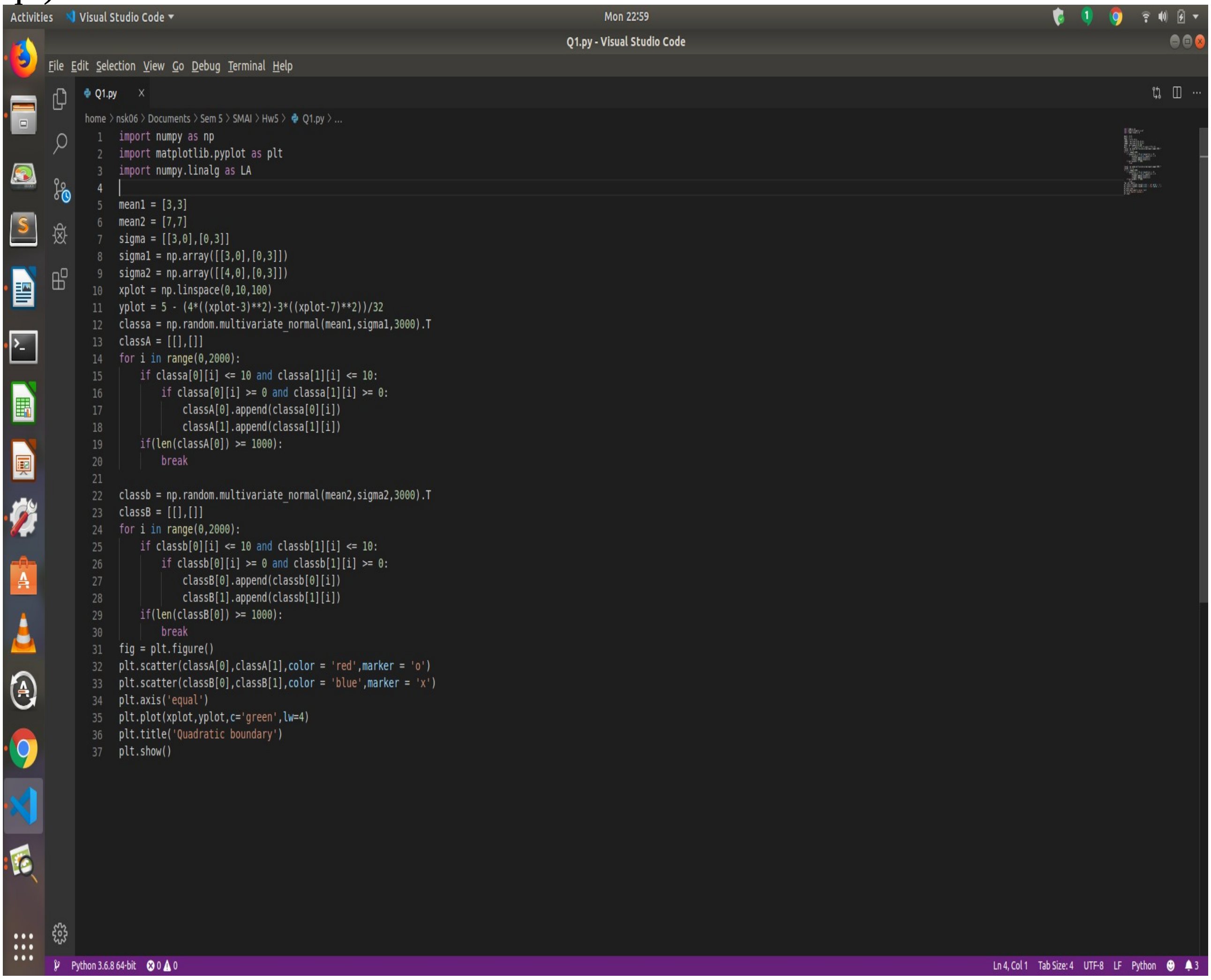
Homework – 5

q1)



This is the code to generate the plots the analytical equations are:

Ans-1 a) $\mu_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $\mu_2 = \begin{bmatrix} 7 \\ 7 \end{bmatrix}$

$$\Sigma_1 = \Sigma_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

The decision boundary here is

$$[X-\mu_1]^T \Sigma^{-1} [X-\mu_1] = [X-\mu_2]^T \Sigma^{-1} [X-\mu_2]$$

$$\begin{bmatrix} x-3 \\ y-3 \end{bmatrix}^T \frac{1}{9} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x-3 \\ y-3 \end{bmatrix} = \begin{bmatrix} x-7 \\ y-7 \end{bmatrix}^T \frac{1}{9} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x-7 \\ y-7 \end{bmatrix}$$

$$\Rightarrow (x-3)^2 + (y-3)^2 = (x-7)^2 + (y-7)^2$$

$$\Rightarrow \boxed{y = 10 - x}$$

b) $\Sigma_1 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ $\Sigma_2 = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

$$[x-3, y-3] \frac{1}{9} \begin{bmatrix} 3(x-3) \\ 3(y-3) \end{bmatrix} = [x-7, y-7] \frac{1}{12} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x-7 \\ y-7 \end{bmatrix}$$

$$\frac{1}{9}\left(3(x-3)^2 + 3(y-3)^2\right) = \frac{1}{12}\left(3(x-7)^2 + 4(y-7)^2\right)$$

$$\frac{(x-3)^2}{3} + \frac{(y-3)^2}{3} = \frac{(x-7)^2}{4} + \frac{(y-7)^2}{3}$$

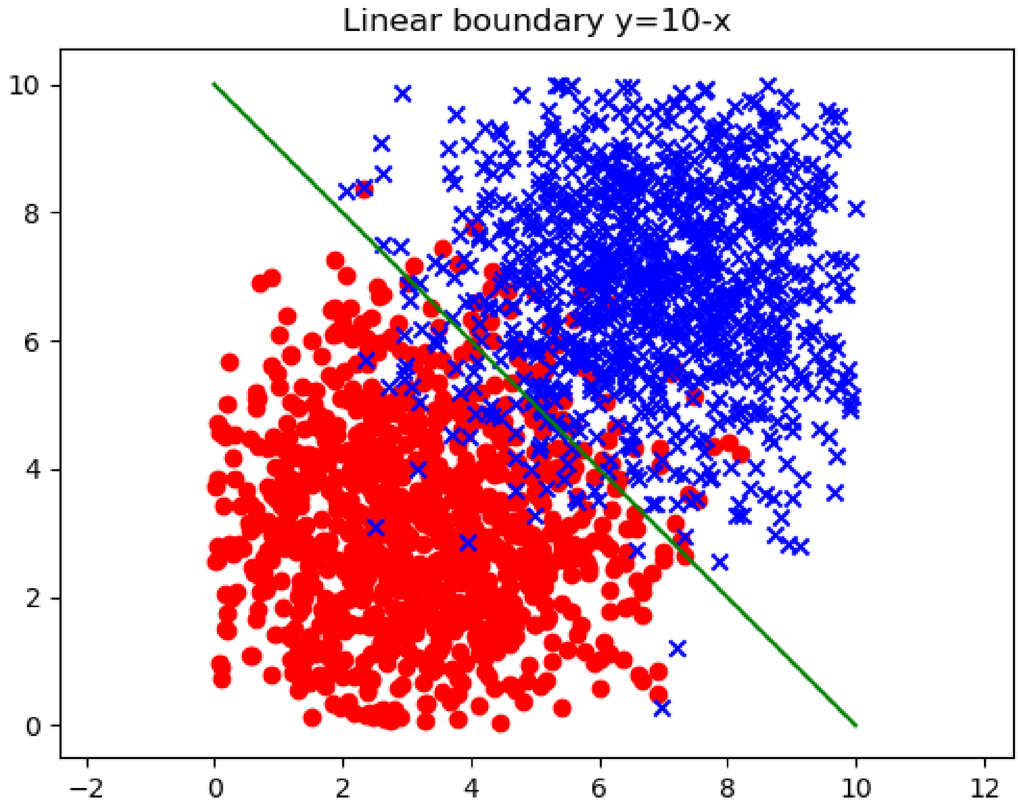
$$\Rightarrow (x-3)^2 - \frac{3}{4}(x-7)^2 = (y-7)^2 - (y-3)^2$$

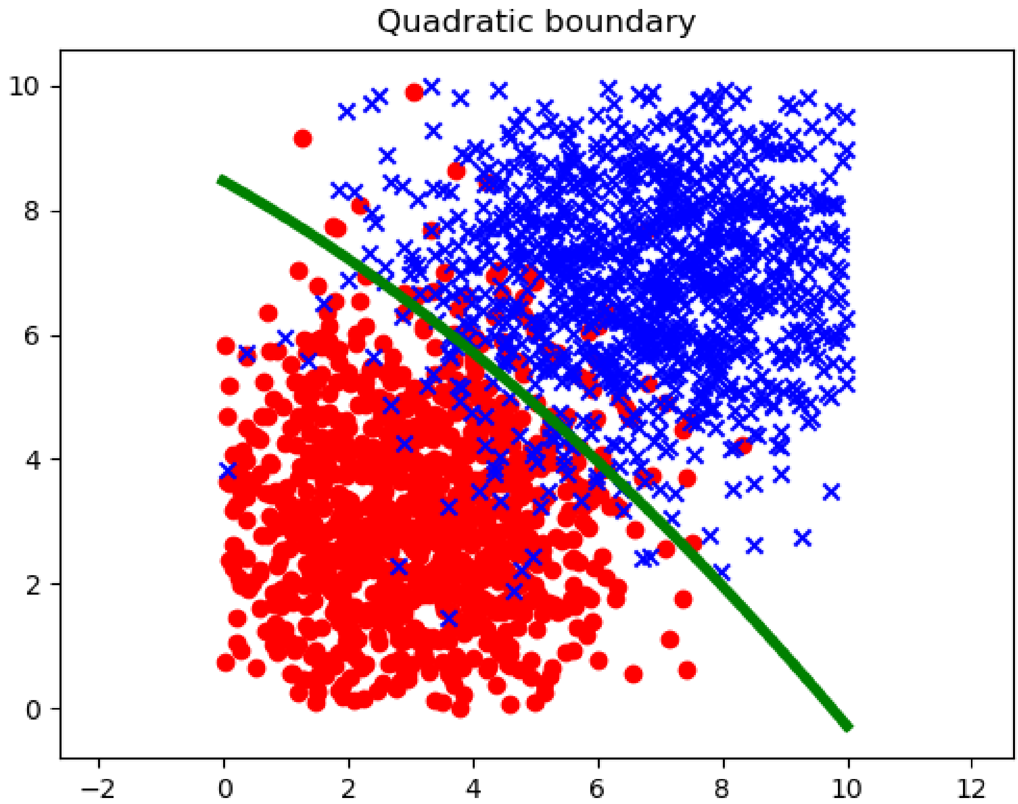
$$= (y-7-y+3)(2y-10)$$

$$\frac{4(x-3)^2 - 3(x-7)^2}{4} = -8(y-5)$$

$$\Rightarrow \boxed{y = 5 - \frac{1}{32}\left(4(x-3)^2 - 3(x-7)^2\right)}$$

The generated plots are:





Q2a)

Ans=2 a) The dimensions of R are 2000X200.
As there are about 35 different types of students so practical rank is 35.
Now if R_{ij} is the number of times student i orders from restaurant j then we can have practical rank of atmax 200. Hence by this rank increases.

Q2b)

b) Now from data we can map the info of new students like which student orders from which restaurants and how many times.
Now we will classify these new students as one of the 35 categories using k nearest neighbour algorithm.

The KNN algorithm is described as:
Calculate distance of sample point from all training points
sort the distances and select the nearest k points
Depending on the majority of a class of points in neighbourhood classify the sample point in one of the classes.

Q2c)

c) Now we have to recommend students to a restaurant. Let us a define a restaurant as d x 1 vector with d type of menu items. With SVD we will find rank of this and find the categories of restaurants. Now based on restaurant category & transaction details we can recommend students to a restaurant by using the KNN algorithm & by finding the category of restaurant & thus recommending students which order from similar restaurants.

The KNN algorithms is described as:
Calculate distance of sample point from all training points
sort the distances and select the nearest k points
Depending on the majority of a class of points in neighbourhood classify the sample point in one of the classes.

Q3)

Ans-3 By definition a positive semi-definite matrix is such $u^T \Sigma u \geq 0$ $\forall$ $n \times 1$ vectors u.

$$\Sigma = \text{Expected}\left((X-\mu)(X-\mu)^T\right) = E\left((X-\mu)(X-\mu)^T\right)$$

$$u^T \Sigma u = E\left(\left(u^T(X-\mu)\right)\left(u^T(X-\mu)\right)^T\right) = E(Y^2)$$

where $Y = u^T(X-\mu)$

Now expected value of positive quantity is positive

so $E(Y^2) \geq 0 \implies u^T \Sigma u \geq 0$

so Σ is positive semi-definite.