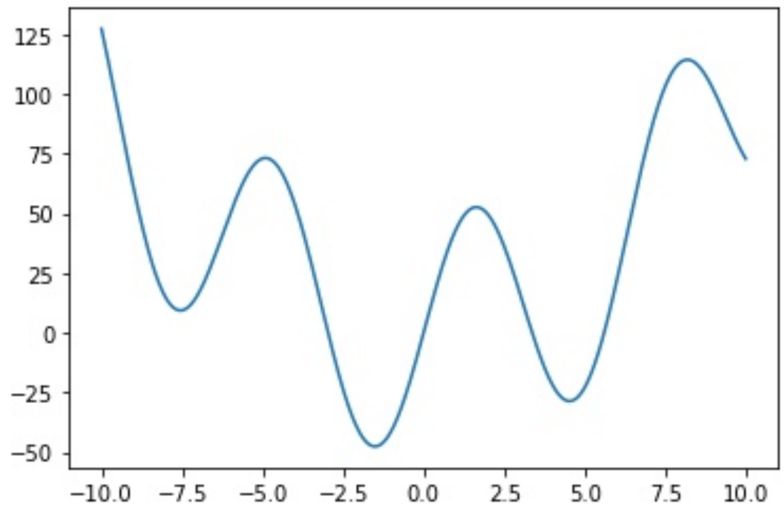
HOMEWORK 4
COSC 522

Harshvardhan
Student Id: 609162

Problem 1

(A)



The local minima points are at the following points (approx): x = -8, x = -1.5, x = 4.5. The global minima is at x = -1.5. All numbers are approximate.

(B)

Python implementation is in the appended Python notebook.

The minima with learning rate 0.01 and starting at x = 7 gives the optimal minima as x = 4.530754493687802.

The minima with learning rate 0.01 and starting at x = 1 gives the optimal minima as x = -1.509643252851762. This is local minima and global minima.

Both the algorithms find the local minima which is closest to the starting point. If we had to find the global minima, we would look at the graph and start around -2. Then, we would obtain the minima at x = -1.511035886414506. This is close to the minima obtained earlier. This is the global minima.

(C)

Every algorithm looks for the minima which is closest to the starting point. When we start at the point which is close to the local minima, gradient descent method found the local minima as that point.

The learning rate can complicate our analysis. If our function is very wobbly, then a small learning rate makes sense. However, in our case the function is very smooth. Thus, a not-so-small-not-so-large learning rate makes more sense.

Using such learning rate we get the minima faster and the algorithm converges faster. Else, the algorithm takes longer to converge (although it gives the same value as minima).

Problem 2

All codes are implemented in the Python notebook submitted.

Problem 3

<u>For AND gate:</u>

X	Y
0, 0	0
0, 1	0
1, 0	0
1, 1	1

$$\mu_1 = \frac{1}{3}\left[\begin{pmatrix}0\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\right] = \begin{bmatrix}2/3\\2/3\end{bmatrix} = \begin{bmatrix}0.66\\0.66\end{bmatrix}$$

$$\mu_2 = \begin{pmatrix}1\\1\end{pmatrix}$$

$$S_1 = \sum_{i=1}^{3} (X_i - \mu_1)(X_i - \mu_1)'$$

$$= \begin{bmatrix}-0.66\\-0.66\end{bmatrix}\begin{bmatrix}-0.66 & -0.66\end{bmatrix} + \begin{bmatrix}-2/3\\1/3\end{bmatrix}\begin{bmatrix}-2/3 & 1/3\end{bmatrix} +$$

$$\begin{bmatrix} 1/3 \\ -2/3 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 \end{bmatrix}$$

$$= \begin{bmatrix} 4/9 & 4/9 \\ 4/9 & 4/9 \end{bmatrix} + \begin{bmatrix} 4/9 & -2/9 \\ -2/9 & 1/9 \end{bmatrix} + \begin{bmatrix} 1/9 & -2/9 \\ -2/9 & 4/9 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$S_2 = \sum_{i=1}^{1} (X_i - \mu_2)(X_i - \mu_2)'$$

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

$$S_W = S_1 + S_2 = I$$

$$S_W^{-1} = I$$

$$W = S_W^{-1}(\mu_1 - \mu_2) = I * \begin{bmatrix} -0.33 \\ -0.33 \end{bmatrix} = \begin{bmatrix} -0.33 \\ -0.33 \end{bmatrix}$$

$$y_1 = \begin{bmatrix} -0.33 \\ -0.33 \end{bmatrix}' \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

$$y_2 = -0.33$$

$$y_3 = -0.33$$

$$y_4 = -0.66$$

With minimum dist classifier:

$$\mu_1 = \frac{0 + (-0.33) - 0.33}{3} = -\frac{0.66}{3} = -0.22$$

$$\mu_2 = -0.66$$

For PCA

$\mu_X = 0.5$

$\mu_Y = 0.5$

$Var_X = 0.333$

$Var_Y = 0.333$

$Cov_{X,Y} = 0$

$$\Sigma = \begin{bmatrix} 0.333 & 0 \\ 0 & 0.333 \end{bmatrix}$$

eigenvalues: 0.33 & 0.33

eigenvectors: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ & $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Projections

$y_1 = 0$

$y_2 = 1$

$y_3 = 0$

$y_4 = 1$

$\therefore$ PCA can't classiffy completely

For Perceptron

Found from code: $W = [1.83, 1.15]$, $b = 2.13$

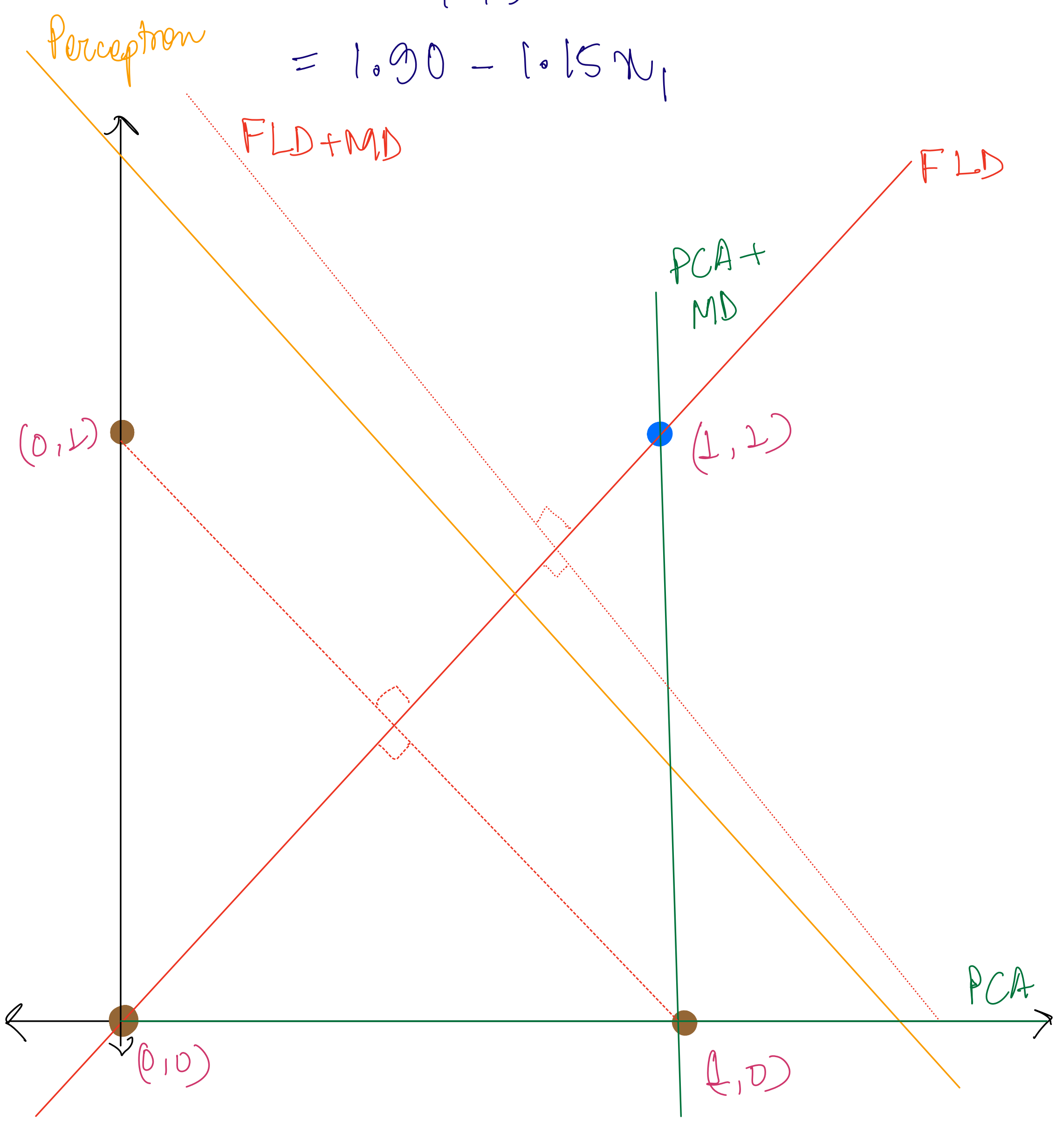
Thus the equation is :

$$w^T x - b = 0$$

$$\Rightarrow 1.33 x_1 + 1.15 x_2 - 2.19 = 0$$

$$\Rightarrow x_2 = \frac{2.19 - 1.33 x}{1.15}$$

$$= 1.90 - 1.15 x_1$$



Legend

Orange line is for the perceptron.

Green lines are for PCA & PCA+MD.

Red lines are for FLD & FLD+MD.