problem 1

1) Given data points written in matrix form

$$\text{features} = \begin{bmatrix} 1 & 1 \\ 2 & -2 \\ -1 & -1.5 \\ -2 & -1 \\ -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \qquad \text{labels} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

let $\hat{y} = Wx + b$ be the perceptron

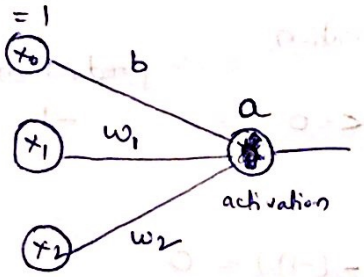
Since two features for each training sample, this can be written as

$$\hat{y} = w_1 x_1 + w_2 x_2 + b x_0 \quad \text{where} \quad x_0 = 1$$

Rewriting this in matrix form we get

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -2 \\ 1 & -1 & -1.5 \\ 1 & -2 & -1 \\ 1 & -2 & 1 \\ 1 & 1.5 & -0.5 \end{bmatrix}_{6\times 3} \begin{bmatrix} b \\ w_1 \\ w_2 \end{bmatrix}_{3\times 1} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \\ \hat{y}_5 \\ \hat{y}_6 \end{bmatrix}_{6\times 1}$$

$x_0 \quad x_1 \quad x_2$




$$= \begin{cases} 1 & \text{if } \hat{y} > 0 \quad a > 0 \\ -1 & \text{if } \hat{y} \leq 0 \quad a <= 0 \end{cases}$$

prediction = $\hat{y}$

First initialize $b = 0$, $w_1 = 0$, $w_2 = 0$. Then, for each sample update weights using below

$$b = b + (y - \hat{y})$$

$$w_1 = w_1 + \alpha (y - \hat{y}) x_1$$

$$w_2 = w_2 + \alpha (y - \hat{y}) x_2$$

$\alpha$ is learning rate let $\alpha = 0.01$

Epoch - 1

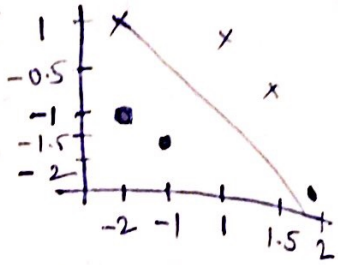
Sample 1

| | | activation | prediction | real |
|---|---|---|---|---|
| $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ | $0 <= 0$ | $-1$ | $1$ |

$b = 0 + 0.01(1 - (-1)) = 0.02$

$w_1 = 0 + 0.01(1 - (-1)) \times 1 = 0.02$

$w_2 = 0 + 0.01(1 - (-1)) \times 1 = 0.02$



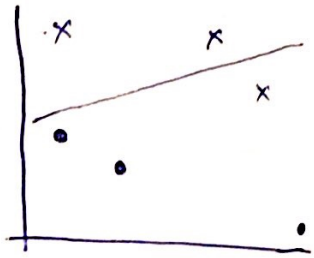

Sample 2

| | | activation | prediction | real |
|---|---|---|---|---|
| $\begin{bmatrix} 1 & 2 & -2 \end{bmatrix}$ | $\begin{bmatrix} 0.02 \\ 0.02 \\ 0.02 \end{bmatrix}$ | $0.02 > 0$ | $+1$ | $-1$ |

$b = 0.02 + 0.01(-1-1) = 0$

$w_1 = 0.02 + 0.01(-1-1) \times 2 = -0.02$

$w_2 = 0.02 + 0.01(-1-1) \times -2 = 0.06$



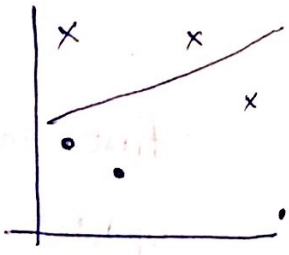
Sample 3

| | | activation | prediction | real |
|---|---|---|---|---|
| $\begin{bmatrix} 1 & -1 & -1.5 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ -0.02 \\ 0.06 \end{bmatrix}$ | $-0.07 <= 0$ | $-1$ | $-1$ |

$b = 0 + 0.01(-1 - (-1)) = 0$

$w_1 = -0.02 + 0.01(-1 - (-1)) \times -1 = -0.02$

$w_2 = 0.06 + 0.01(-1 - (-1)) \times -1.5 = 0.06$



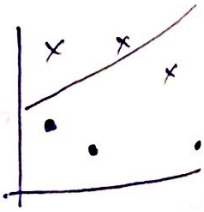
Sample 4

| | | activation | prediction | real |
|---|---|---|---|---|
| $\begin{bmatrix} 1 & -2 & -1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ -0.02 \\ 0.06 \end{bmatrix}$ | $-0.02 <= 0$ | $-1$ | $-1$ |

$b = 0 + 0.01(-1 - (-1)) = 0$

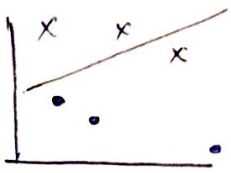
$w_1 = -0.02 + 0.01(-1 - (-1)) \times -2 = -0.02$

$w_2 = +0.06 + 0.01(-1 - (-1)) \times -1 = 0.06$



Sample 5

| | activation | prediction | real |
|---|---|---|---|
| $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -0.02 \\ 0.06 \end{bmatrix}$ | $0.1 > 0$ | 1 | 1 |



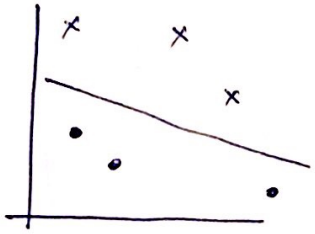
$b = 0 + 0.01(1-1) = 0$

$w_1 = -0.02 + 0.01(1-1)\times -2 = -0.02$

$w_2 = 0.06 + 0.01(1-1)(1) = 0.06$

Sample 6

| | activation | prediction | real |
|---|---|---|---|
| $\begin{bmatrix} 1 & 1.5 & -0.5 \end{bmatrix} \begin{bmatrix} 0 \\ -0.02 \\ 0.06 \end{bmatrix}$ | $0 <= 0$ | -1 | 1 |



$b = 0 + 0.01(1-(-1)) = 0.02$

$w_1 = -0.02 + 0.01(1-(-1))(1.5) = 0.01$

$w_2 = 0.06 + 0.01(1-(-1))(0.5) = 0.05$

Epoch 2

Sample 1

| | activation | prediction | real |
|---|---|---|---|
| $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.02 \\ 0.01 \\ 0.05 \end{bmatrix}$ | $0.08 > 0$ | 1 | 1 |

$b = 0.02 + 0.01(1-1) = 0.02$

$w_1 = 0.01 + 0.01(1-1)\times 1 = 0.01$

$w_2 = 0.05 + 0.01(1-1)\times 1 = 0.05$

Since, the weights didn't change the final equation is set as $\hat{y} = 0.01x_1 + 0.05x_2 + 0.02$

plot of $x_1$ vs $x_2$ and for two classes