HomeWork - 2

Q1(a) Data, we will classify C1 as -1 and C2 as 1, thus we have dataset

$D = (4.8, -10.4) \to -1$
$(12.8, 5.6) \to 1$
$(3.2, -7.2) \to -1$
$(7.2, 4) \to 1$
$(0.4, -10) \to -1$
$(9.6, 8.8) \to 1$
$(1.8, -2.2) \to -1$
$(6.6, 2.6) \to 1$
$(5, -6.2) \to -1$
$(12, -4) \to 1$

→ Plot of the above given data is attached in plot page - 1

Optimization Function : $\min_{w,b} \frac{1}{2} \|w\|^2$

Constraints : $y^{(i)} \cdot (w^T x^{(i)} + b) \geq 1$

For this problem, this translates to:
Optimization function : $\min_{w,b} = \frac{1}{2}(w_1^2 + w_2^2)$

constraints -

jupyter Untitled Last Checkpoint: 4 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3 (ipykernel)

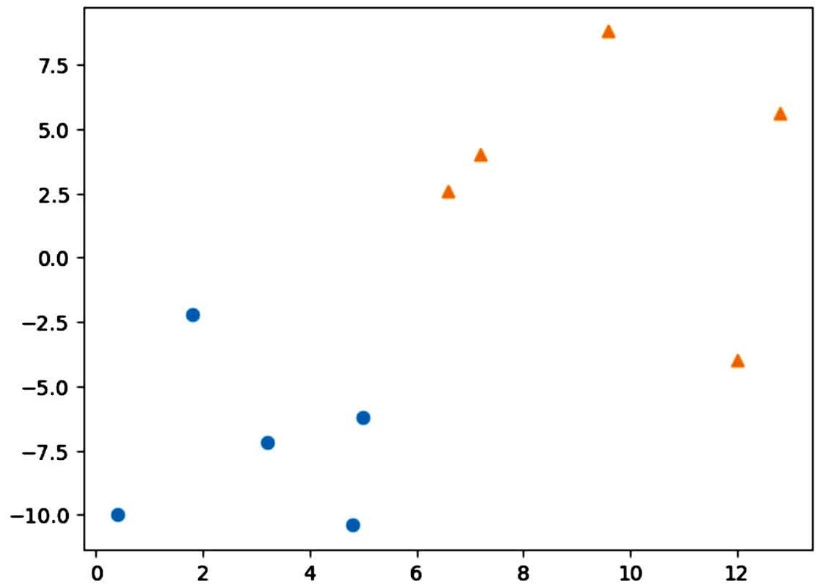
Run | Code

PLOT 1

In [26]:

```
X = [4.8,3.2,0.4,1.8,5]
Y = [-10.4,-7.2,-10,-2.2,-6.2]
plt.scatter(X,Y, marker="o")
X = [12.8,7.2,9.6,6.6,12]
Y = [5.6,4,8.8,2.6,-4]
plt.scatter(X,Y,marker = "^")
plt.show()

```



constraints - $-(4.8w_1 - 10.4w_2) \geq 1$

$(12.8w_1 + 5.6w_2) \geq 1$

$-(3.2w_1 - 7.2w_2) \geq 1.$

$(7.2w_1 + 4w_2) \geq 1$

$-(0.4w_1 - 10w_2) \geq 1$

$(9.6w_1 + 8.8w_2) \geq 1$

$-(1.8w_1 - 2.2w_2) \geq 1$

$(6.6w_1 + 2.6w_2) \geq 1$

$-(5w_1 - 6.2w_2) \geq 1$

$(12w_1 - 4w_2) \geq 1$

→ The corresponding Lagrangian in general is

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \left[y^{(i)} \left(w^T x^{(i)} + b \right) - 1 \right]$$

For this problem it translates into

$$\begin{aligned}
\mathcal{L}(w, b, \alpha) = \frac{1}{2}(w_1^2 + w_2^2) &- \alpha_1 [-(4.8w_1 - 10.4w_2 + b) - 1] \\
&- \alpha_2 [(12.8w_1 + 5.6w_2 + b) - 1] \\
&- \alpha_3 [-(3.2w_1 - 7.2w_2 + b) - 1] \\
&- \alpha_4 [(7.2w_1 + 4w_2 + b) - 1] \\
&- \alpha_5 [-(0.4w_1 - 10w_2 + b) - 1] \\
&- \alpha_6 [(9.6w_1 + 8.8w_2 + b) - 1] \\
&- \alpha_7 [-(1.8w_1 - 2.2w_2 + b) - 1] \\
&- \alpha_8 [(6.6w_1 + 2.6w_2 + b) - 1] \\
&- \alpha_9 [-(5w_1 - 6.2w_2 + b) - 1] \\
&- \alpha_{10} [(12w_1 - 4w_2 + b) - 1]
\end{aligned}$$

The Langrangian Dual problem is defined as

Optimization Function

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y^{(i)} y^{(j)} \alpha_i \alpha_j \langle x^{(i)}, x^{(j)} \rangle$$

constraint $\quad \alpha_i \geq 0$

$$\sum_{i=1}^{n} \alpha_i y^{(i)} = 0$$

For the given problem it translates to

Optimization Function

$$\max_{\alpha} \sum_{i=1}^{8} \alpha_i - \frac{1}{2} \Big(\alpha_1 (131.2\alpha_1 - 3.2\alpha_2 + 90.24\alpha_3$$

$$+ 7.04\alpha_4 + 105.92\alpha_5 + 45.44\alpha_6 + 31.52\alpha_7 - 4.64\alpha_8$$

$$+ 88.48\alpha_9 - 99.12\alpha_{10})$$

$$+ \alpha_2 (-3.2\alpha_1 + 195.2\alpha_2 - 0.64\alpha_3 + 114.56\alpha_4 + 50.88\alpha_5 + 172.16\alpha_6$$

$$- 10.72\alpha_7 + 99.04\alpha_8 - 29.28\alpha_9 + 131.2\alpha_{10})$$

$$+ \alpha_3 (90.24\alpha_1 - 0.64\alpha_2 + 62.08\alpha_3 + 5.76\alpha_4 + 73.28\alpha_5 + 32.64\alpha_6 + 21.6\alpha_7$$

$$- 2.39\alpha_8 + 60.64\alpha_9 - 67.2\alpha_{10})$$

$$+ \alpha_4 (7.04\alpha_1 + 114.56\alpha_2 + 5.76\alpha_3 + 67.84\alpha_4 + 37.12\alpha_5 + 104.32\alpha_6 - 4.16\alpha_7 + 57.92\alpha_8$$

$$- 11.2\alpha_9 + 70.4\alpha_{10}) + \alpha_5 (105.92\alpha_1 + 50.88\alpha_2 + 73.28\alpha_3 + 37.12\alpha_4 + 100.16\alpha_5 +$$

$$+ 84.16\alpha_6 + 22.72\alpha_7 + 23.36\alpha_8 + 64.0\alpha_9 - 44.8\alpha_{10}) + \alpha_6 (45.44\alpha_1 + 172.16\alpha_2 +$$

$$32.64\alpha_3 + 104.32\alpha_4 + 84.16\alpha_5 + 169.6\alpha_6 + 2.08\alpha_7 + 86.24\alpha_8 + 6.56\alpha_9$$

$$+ 80\alpha_{10}) + \alpha_7 (31.52\alpha_1 - 10.72\alpha_2 + 21.6\alpha_3 - 4.16\alpha_4 + 22.72\alpha_5 + 2.08\alpha_6 +$$

$$8.08\alpha_7 - 6.16\alpha_8 + 22.64\alpha_9 - 30.40\alpha_{10}) + \alpha_8 (-4.64\alpha_1 + 99.04\alpha_2 - 2.4\alpha_3 +$$

$$57.92\alpha_4 + 23.36\alpha_5 + 86.24\alpha_6 - 6.16\alpha_7 + 50.32\alpha_8 - 16.88\alpha_9 + 68.8\alpha_{10})$$

Contd $+\alpha_9(88.48\alpha_1 - 29.28\alpha_2 + 60.64\alpha_3 - 11.2\alpha_4 + 64\alpha_5 + 6.56\alpha_6 + 22.64\alpha_7 - 16.88\alpha_8 + 63.44\alpha_9 - 84.8\alpha_{10})$

$+\alpha_{10}(-99.2\alpha_1 + 131.2\alpha_2 - 67.2\alpha_3 + 70.4\alpha_4 - 44.8\alpha_5 + 80\alpha_6 - 30.4\alpha_7 + 68.8\alpha_8 - 84.8\alpha_9 + 160\alpha_{10}))$

Constraints

$\alpha_i \geq 0$ where $i = 1 \dots 10$

$-\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 - \alpha_5 + \alpha_6 - \alpha_7 + \alpha_8 - \alpha_9 + \alpha_{10} = 0$

(b) * Now from this we can perform SMO using a heuristic to pick pairs of α variables and optimizing for them

→ So now we have the least distance b/w the two classification points and they are α_7, α_8

α_8, α_7

→ For first iteration we will use these two ~~constraint~~ α values

Iteration - 1

→ For iteration first consider the following equations

$$w = \sum_{i=1}^{10} \alpha_i y^{(i)} x^{(i)} = 0$$

assuming all the α_i are zero

and the offset b for the primal:

$$b = -\frac{\max_{i:y^{(i)}=-1} w^T x^{(i)} + \min_{i:y^{(i)}=1} w^T (x)^{(i)}}{2} = 0$$

Iteration 1 (α_8, α_7)

$$\xi = -\sum_{i \neq 8,7} \alpha_i y^{(i)} = 0$$

$\therefore \alpha_8 = (\xi - \alpha_7 y^{(7)}) y^{(7)} = \alpha_7$

→ We can rewrite the unclipped optimization

~~$\therefore \alpha_7^{new,uc} = \text{argmax}_{\alpha_7} \, \phi(\alpha_7 + \alpha_8)$~~

$$\therefore \alpha_7^{new,uc} = \text{argmax}_{\alpha_7} \; \alpha_7 + \alpha_8 - \frac{1}{2}\left(\alpha_7(8.08\alpha_7 - 6.16\alpha_8) + \alpha_8(-6.16\alpha_7 + 50.32\alpha_8)\right)$$

Putting the equation $\alpha_7 = \alpha_8$

$$\therefore \alpha_7 = \text{argmax}_{\alpha_7} \; 2\alpha_7 - \frac{1}{2}\left[1.92\alpha_7^2 + 44.16\alpha_7^2\right]$$

$$= \text{argmax}_{\alpha_7} \; 2\alpha_7 - \frac{1}{2}\left[46.08\alpha_7^2\right]$$

To optmize this we take derivative and set it to 0

$$\therefore \frac{\partial}{\partial \alpha_7}\left(2\alpha_7 - \frac{1}{2}\left[46.08\alpha_7^2\right]\right) = 0$$

$$\therefore 2 - 46.08\alpha_7 = 0$$

$$\therefore \alpha_7 = 2/46.08 = 0.0434$$

Thus we can say

$\alpha_7^{\text{new clipped}} = 0.0434$, $\alpha_8^{\text{new clipped}} = 0.0434$

Now we can resolve

$$w = \sum_{i=1}^{n} \alpha_i y^{(i)} x^{(i)} = (0.0434)(-1)(1.8, -2.2)^T + (0.0434)(1)(6.6, 2.6)^T$$

$$\therefore w = \begin{bmatrix} -0.07812 \\ 0.09548 \end{bmatrix} + \begin{bmatrix} 0.28644 \\ 0.11284 \end{bmatrix}$$

$$= \begin{bmatrix} 0.20832 \\ 0.20832 \end{bmatrix}$$

$$b = \frac{-0.083328 + 1.667}{2}$$

$$= 0.791$$

The plot for this graph is attached as plot-2

jupyter Untitled Last Checkpoint: 4 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3 (ipykernel)

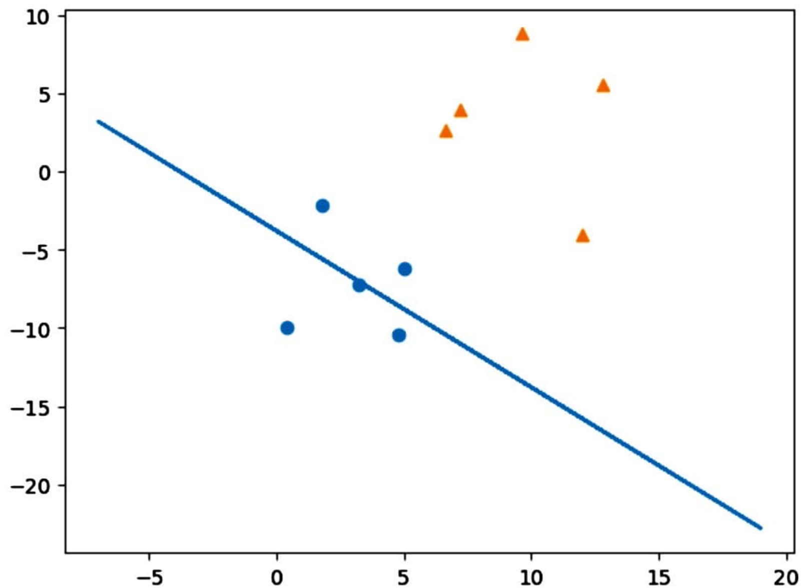
Run | Code

PLOT 2

In [27]:

```
X = [4.8,3.2,0.4,1.8,5]
Y = [-10.4,-7.2,-10,-2.2,-6.2]
plt.scatter(X,Y, marker="o")
X = [12.8,7.2,9.6,6.6,12]
Y = [5.6,4,8.8,2.6,-4]
plt.scatter(X,Y,marker = "^")
x = np.random.randint(low=-10, high=20, size=40)
y = -x-3.797
plt.plot(x,y)
plt.show()

```



Iteration 2 :: α_9 & α_7 are now defying the margin hence we again iterate

$$y = -\sum_{i \neq 9,7} \alpha_i y^{(i)} = 0$$

$\therefore$ Rewriting α_7 in terms of α_9

$\therefore \alpha_7 + \alpha_9 - \alpha_8 = 0$

$\therefore \alpha_7 = \alpha_8 - \alpha_9$

$= 0.20832 - \alpha_9$

$$\therefore \alpha_9^{new\ unclipped} = \text{argmax}_{\alpha_9}\ 0.20832 - \alpha_9 + \alpha_9 - \frac{1}{2}\big(\alpha_7(8.08\alpha_7 - 6.16\alpha_8 + 22.64\alpha_9) + \alpha_8(-6.16\alpha_7 + 50.32\alpha_8 - 16.88\alpha_9) + \alpha_9(-30.4\alpha_7 + 68.80 - 84.8\alpha_9)\big)$$

$$\therefore \alpha_9^{new} = \text{armax}_{\alpha_9} = 0.20832 - \frac{1}{2}\big[\alpha_7(8.082\alpha_7 - 6.16\times0.20832)\alpha_9 + 0.20832(-6.16\alpha_7 + 50.32(0.20832) - 16.88\alpha_9) + \alpha_9(-30.4\alpha_7 + 68.8\times0.20832 - 84.8\alpha_9)$$

$\cancel{= 0.20832}$

$\cancel{= 0.20832 - \frac{1}{2}\big[(0.20832-\alpha_9)(8.082(0.2083-\alpha_9)}$

$$= 0.20832 - \frac{1}{2}\left(-68.96\alpha_9^2 + 8.397\alpha_9 + 2\right)$$

Now taking derivative for optimization

$$\therefore \frac{d}{d\alpha_9}\left(0.20832 - \frac{1}{2}\left(-68.96\alpha_9^2 + 8.397\alpha_9 + 2\right)\right)$$

$$\therefore \quad 68.96\,\alpha_9 - 8.397 = 0$$

$$\therefore \alpha_9 = \frac{8.397}{68.96} = 0.121$$

$$\therefore \alpha_7^{new} = 0.20832 - 0.121$$

$$= 0.0865$$

$$\therefore w = \sum_{i=1}^{n} \alpha_i y^{(i)} x^{(i)} =$$

$$= (-1)(\cancel{0.0865}\ 0.0865) \begin{bmatrix} 1.8 \\ -2.2 \end{bmatrix}$$

$$+ (1)(0.20832) \begin{bmatrix} 6.6 \\ 2.6 \end{bmatrix}$$

$$+ (-1)(0.121) \begin{bmatrix} 5 \\ -6.2 \end{bmatrix}$$

$$\begin{bmatrix} -0.1557 \\ 0.1903 \end{bmatrix} + \begin{bmatrix} 1.374 \\ 0.541 \end{bmatrix} + \begin{bmatrix} -0.605 \\ 0.7502 \end{bmatrix}$$

$$= \begin{bmatrix} 0.6138 \\ 1.4815 \end{bmatrix}$$

$$b = \frac{-2.15536 + 1.4335}{2} = -0.36093$$

9 The plot is attached as plot-3

jupyter Untitled Last Checkpoint: 4 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

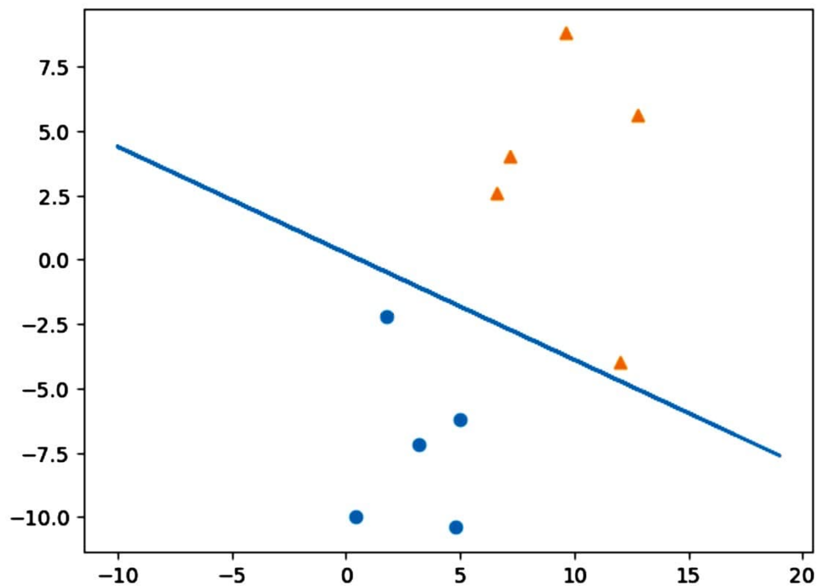
Trusted | Python 3 (ipykernel)

Run | Code

PLOT 3

In [30]:

```
X = [4.8,3.2,0.4,1.8,5]
Y = [-10.4,-7.2,-10,-2.2,-6.2]
plt.scatter(X,Y, marker="o")
X = [12.8,7.2,9.6,6.6,12]
Y = [5.6,4,8.8,2.6,-4]
plt.scatter(X,Y,marker = "^")
x = np.random.randint(low=-10, high=20, size=40)
y = -0.4139*x+0.2544
plt.plot(x,y)
plt.show()
```



Iteration 3 - Now there are no points which are being violated thus now I pick α for point $(12, -4)$ and α_1 $(4.8, -10.4)$ α_{10},

$\rightarrow \therefore -\sum \alpha_i y^{(i)} = 0$

$\therefore \cancel{\alpha_1 + \alpha_{10}}\ \alpha_1 - \alpha_{10} + 0.0865 - 0.20832 + 0.121 = 0$

$\therefore \alpha_1 - \alpha_{10} = 0.00082$

$\therefore \alpha_1 = 0.00082 + \alpha_{10}$

$\alpha_{10}^{new} =$

$\therefore \operatorname{argmax}_{\alpha_{10}} \ \alpha_1 + \alpha_{10} - \frac{1}{2}\left[\alpha_1(131.2\alpha_1 - 99.2\alpha_{10}) + \alpha_{10}(-99.2\alpha_1 + 160\alpha_{10})\right]$

$= 2\alpha_{10} + 0.00082 - \frac{1}{2}\left[0.00082 + \alpha_{10}(131.2(0.00082 + \alpha_{10}) - 99.2\alpha_{10}) + \alpha_{10}(-99.2(0.00082 + \alpha_{10})) + 160\alpha_{10})\right]$

$= 2\alpha_{10} + 0.00082 - \frac{1}{2}\left[92.8\alpha_{10}^2 + 0.0449\alpha_{10}\right]$

for optimization

$\frac{d}{d\alpha_{10}}\left(2\alpha_{10} + 0.00082 - \frac{1}{2}\left[92.8\alpha_{10}^2 + 0.0449\alpha_{10}\right]\right) = 0$

$\therefore 2 - \frac{\cancel{2} \times 92.8\alpha_{10}}{\cancel{2}} - \frac{1}{2} \times 0.0449 = 0$

$\therefore 2 - 92.8\alpha_{10} - 0.02245 = 0$

$\therefore 92.8\alpha_{10} = 1.97755$

$\therefore \alpha_{10} = 0.0213$

$\therefore \alpha_1^{new\,clipped} = 0.0222$

$\therefore \; w = \sum \alpha_i y^{(i)} x^{(i)}$

$$= \begin{bmatrix} 0.6133 \\ 1.4815 \end{bmatrix} + (-1)(0.0222)\begin{bmatrix} 4.8 \\ -10.4 \end{bmatrix} + (1)(0.0213)\begin{bmatrix} 12 \\ -4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.6133 - 0.10656 + 0.2556 \\ 1.4815 + 0.23088 - 0.0852 \end{bmatrix}$$

$$= \begin{bmatrix} 0.76234 \\ 1.62718 \end{bmatrix}$$

$$b = \frac{-2.207584 + 2.63936}{2}$$

$$= 0.21588$$

The plot for this iteration is in plot-4 attached.

jupyter Untitled Last Checkpoint: 5 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help | Trusted | Python 3 (ipykernel)

Run | Code

```
# PLOT 4
```

In [31]:

```
X = [4.8,3.2,0.4,1.8,5]
Y = [-10.4,-7.2,-10,-2.2,-6.2]
plt.scatter(X,Y, marker="o")
X = [12.8,7.2,9.6,6.6,12]
Y = [5.6,4,8.8,2.6,-4]
plt.scatter(X,Y,marker = "^")
x = np.random.randint(low=-10, high=20, size=40)
y = -0.4685*x-0.1326
plt.plot(x,y)
plt.show()
```