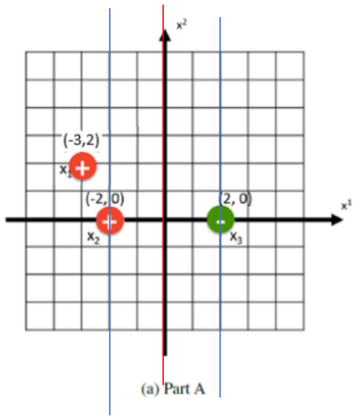
Name – **Nischay Girish Gowda**
Subject – **CSE 575 SML Fall 2022**

Q1. Support Vector Machine.

1.1 For each figure 1(a) ,the support vectors are x2 and x3.

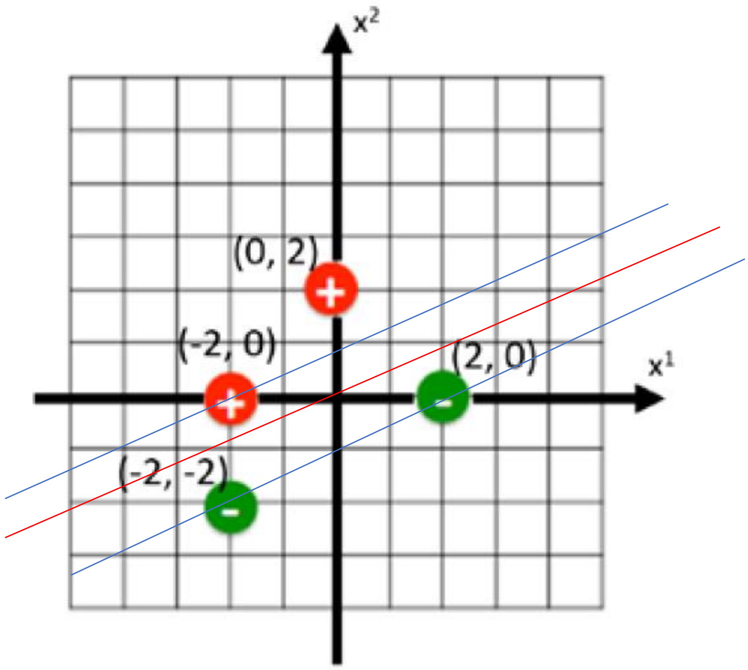
The margin is 2.



(a) Part A

For the case of figure 1(b), the support vectors are the two negative data points and one positive data point at (-2,0).

The margin is $\frac{2}{\sqrt{5}}$.



(b) Part B

1.2 The primal optimization solution is given by,

$W = [w1, w2]^T$, putting the coordinates of training data to the equation.

min $\frac{1}{2}||W||^2 s.ttn(w^T xn + b) \geq 1; n = 1,2,3 \ldots, N$

with $x1 = [-2,2]^T, x2 = [-2,0]^T,$

$and\ x3 = [2,0]^T. We\ have\ the\ primal\ problem$

min $\frac{1}{2}(W1 + W2)^2$

s.t **-2w1 + 2w2 + b >= 1**

-2w1 + b >= 1

-2w1 – b >= 1

The Lagrangian function is given by,

L(w, b, a) = $\frac{1}{2}(W1^2 + W2^2) - a1(-2w1 + 2w2 + b - 1) - a2(-2w1 + b - 1) - a3(-2w1 - b - 1)$

Now substitute training coordinates to general formula of dual problem.

$$L(\boldsymbol{a}) = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m t_n t_m \phi^{\mathrm{T}}(\mathbf{x}_n)\phi(\mathbf{x}_m)$$

s.t an > 0; n = 1,2,3,.....N

$$\sum_{n=1}^{N} a_n t_n = 0$$

We have the dual optimization for this problem.

max value is

$$-\frac{1}{2}(8a1^2 + 4a2^2 + 4a3^2 + 8a1a2 + 8a1a2 + 8a2a3) + a1 + a2 + a3$$

s.t a1 >= 0, a2>= 0, a3>= 0

-a1 - a2 + a3 = 0

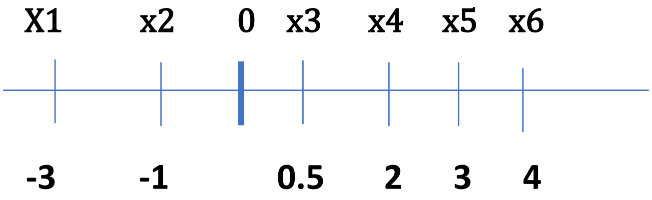
1.3 The decision boundary is **x1 = 1**. **(x1, x2) will have label 1** if x1 < 1 and will have label -1 if x1 > 1.

The decision boundary will not change due to the addition of new point, because it is on the left side of margin.

Q2. K-means

2.1 The given six data points $x1=-3, x2=-1, x3=0.5, x4=2,$ $x5=3, x6 = 4.$

Plot of six data points in 1D space,



2.2 For K = 2. Initial cluster centres $\mu1 = -1$ and $\mu2 = 3$.

After Step 1, cluster 1($\mu1 = -1$) will contain points x1, x2, x3.

cluster 2($\mu2 = 3$) will contain points x4, x5, x6.

After Step 2,

$$\mu_k = \frac{\sum_n r_{nk}\mathbf{x}_n}{\sum_n r_{nk}}.$$

Applying the above formula we get,

$$\mu1 = \frac{-3-1+0.5}{3} = \frac{-3.5}{3} = -1.1667$$

$$\mu2 = \frac{2+3+4}{3} = \frac{9}{3} = 3$$

2.3 For K = 3, initial clusters $\mu1 = -3$, $\mu2 = 0$, $\mu3 = 3$

After step1, cluster 1($\mu1$ = -3) contains points x1.

Cluster 2($\mu2$ = 0) contains points x2, x3.

Cluster 3($\mu3$ = 3) contains x4, x5, x6.

After Step 2,

$\mu 1 = -3$

$\mu 2 = \frac{-1 + 0.05}{2} = -0.25$

$\mu 3 = \frac{2 + 3 + 4}{3} = 3$

The loss function J is calculated as follows,

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2$$

$$\boldsymbol{J} = (\boldsymbol{-3} + \boldsymbol{3})^2 + (\boldsymbol{-1} + \boldsymbol{0.25})^2 + (\boldsymbol{0.05} + \boldsymbol{0.25})^2 + (\boldsymbol{-2} + \boldsymbol{3})^2 + (\boldsymbol{-3} + \boldsymbol{3})^2 + (\boldsymbol{-3} - \boldsymbol{4})^2$$

J = 3.125

2.4 For K = 6, each point will form its own cluster.

J = 0

Q3. Unconstrained Optimization

The function is given by,

$$f(x) = x1^2 + 5.x2^2$$

3.1 The gradient at point x' = $[5,1]^T$

Taking the differentiate of $f(x)$.

$$\nabla \mathrm{f(x)} = \begin{vmatrix} 2.x1 \\ 10.x2 \end{vmatrix}$$

$$\nabla \mathrm{f(x')} = \begin{vmatrix} 10 \\ 10 \end{vmatrix}$$

3.2 Steepest descent direction at point x',

$$-\nabla \mathrm{f(x')} = \begin{vmatrix} -10 \\ -10 \end{vmatrix}$$

3.3 For backtrack line search at point x' along direction $-\nabla \mathrm{f(x')}$

Using α = 0.4 and β = 0.9.

To find the largest value of t such that,

$$\mathbf{f}\left(\mathbf{x}' - \mathbf{t}\nabla\mathbf{f}(\mathbf{x}')\right) <= \mathbf{f}(\mathbf{x}') - \mathbf{0.4t}\ ||\ \nabla\mathbf{f}(\mathbf{x}')||^2$$

Solving the above equation for multiple values of t,

Start with t = 1, then **t -> tβ**

t $in\ \{1,\ 0.9, 0.9^2, 0.9^3, 0.9^4 \ldots\ldots\}$

The computed step size is $\mathbf{0.9^{16}} \approx \mathbf{0.185}$

3.4 To perform exact line search at point x’,

along the direction $-\nabla f(x')$,

the optimal value of t by solving the problem is given by,

$min\ \mathbf{f}\left(\mathbf{x}' - \mathbf{t}\nabla\mathbf{f}(\mathbf{x}')\right)$

Which is given by,

$$\boldsymbol{min\ 600\ t^2 - 200\ t + 30}$$

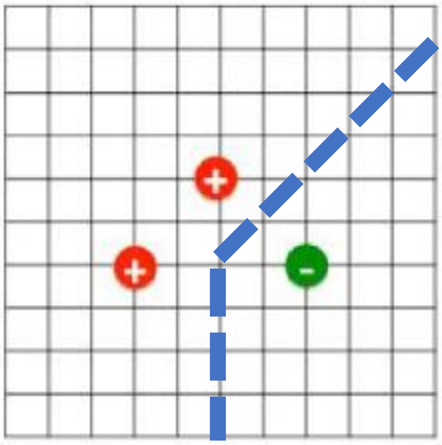
The computed step size is,

$$= \frac{\mathbf{1}}{\mathbf{6}}$$

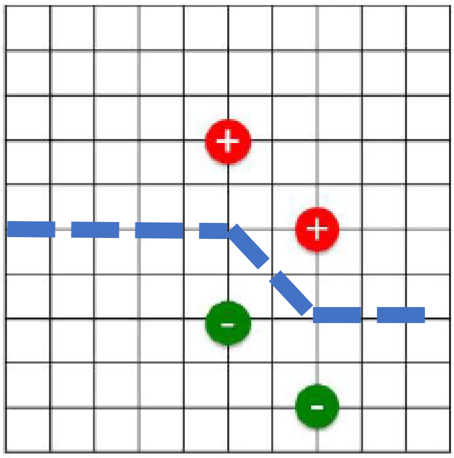
Q4. The Decision Boundary for 1NN.

Following the below steps for drawing Perpendicular bisector to draw the decision boundary.

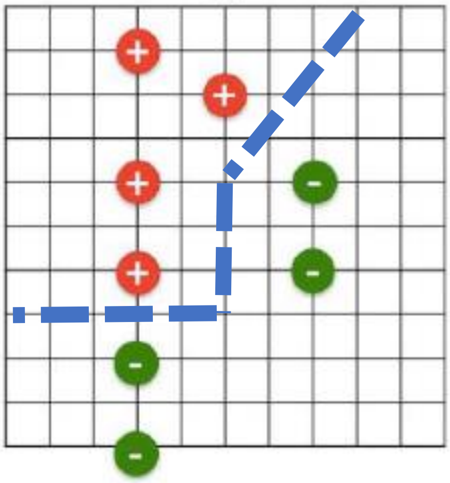
- Identify pairs of training points from different classes. ('+' or '-')
- Draw the perpendicular bisector for every segment connecting the points in a pair in step 1.
- Connect the perpendicular bisector to form a decision boundary.



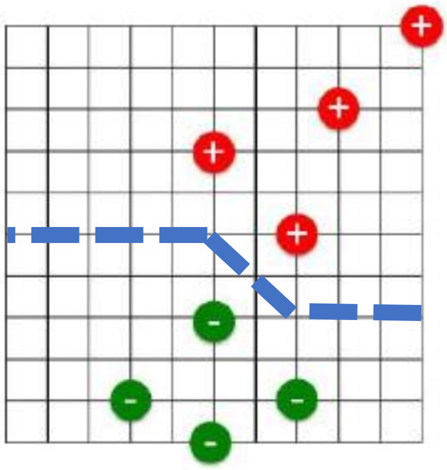
(a)



(b)



(c)



(d)