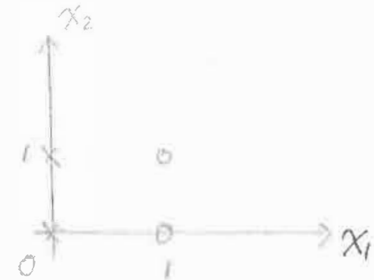
Problem 1

Class 1: $\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$

Class 2: $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}$



(a) Yes, the data points are linearly separable.

(b) Find $\underline{a}$, such that $g(y_i) = a^T y_i > 0, \quad i = 1, \cdots, 4$

$y_1 = [1\ 0\ 0]^T, \quad y_2 = [1\ 0\ 1]^T, \quad y_3 = [-1\ -1\ -1]^T, \quad y_4 = [-1\ -1\ 0]^T$

Single-sample perceptron.

Start with $a(0) = [0\ 0\ 0]$, update with $a(k+1) = a(k) + y_k$.

$a^T(0) y_1 = 0 \rightsquigarrow \quad a(1) = a(0) + y_1 = (1\ 0\ 0)^T$

$a^T(1) y_2 = 1$ ✓

$a^T(1) y_3 = -1 \rightsquigarrow a(2) = a(1) + y_3 = (0\ -1\ -1)^T$

$a^T(2) y_4 = 1$ ✓

$a^T(2) y_1 = 0 \rightsquigarrow a(3) = a(2) + y_1 = (1\ -1\ -1)^T$

$a^T(3) y_2 = 0 \rightsquigarrow a(4) = a(3) + y_2 = (2\ -1\ 0)^T$

$a^T(4) y_3 = -1 \rightsquigarrow a(5) = a(4) + y_3 = (1\ -2\ -1)^T$

$a^T(5) y_4 = 1$ ✓

$a^T(5) y_1 = 1$ ✓

$a^T(5) y_2 = 0 \rightsquigarrow a(6) = a(5) + y_2 = (2\ -2\ 0)^T$

$a^T(6) y_3 = 0 \rightsquigarrow a(7) = a(6) + y_3 = (1\ -3\ -1)^T$

$a^T(7) y_4 = 2$ ✓

$a^T(7) y_1 = 1$ ✓

$a^T(7) y_2 = 0 \rightsquigarrow a(8) = a(7) + y_2 = (2\ -3\ 0)^T$

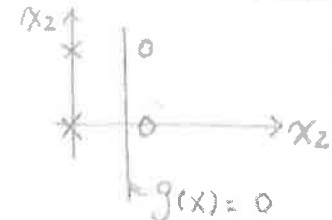
$a^T(8) y_3 = 1$ ✓, $a^T(8) y_4 = 1$ ✓, $a^T(8) y_1 = 2$ ✓, $a^T(8) y_2 = 2$ ✓

TOPS 35502

$\therefore \quad \underline{a} = \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}$

$g(Y) = a^T y \rightsquigarrow g(x) = (-2 \ -3 \ \ 0)\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} = 2 - 3x_1$

Decision boundary $g(x) = 2 - 3x_1 = 0 \rightsquigarrow \underline{x_1 = 2/3}$



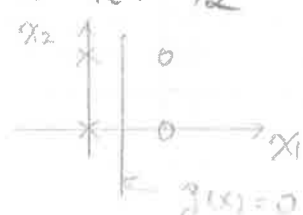
(c) MSE

$\underline{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & 0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$\underline{a} = Y^{T} b$

In MATLAB $a = Y \backslash b = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$

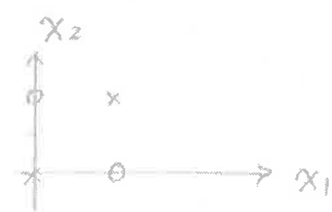
Decision boundary $g(x) = 1 - 2x_1 = 0 \rightsquigarrow x_1 = 1/2$



(d) Ho-Kashyap : See attachment for code & result.

(e) Class 1: $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Class 2: $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}$



- The data points are not linearly separable
- The fixed-increment single-sample Perceptron will not converge.

TOPS 35502

MSE: $\underline{a} = Y^+ b = 0$. The solution is not valid,

with $Y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Ho-kashyap: See code in (d)

The algorithm does not find $\underline{a}$ such that $a^T y_i > 0$ ⇝ not linearly separable.

(f) The perceptron procedure is simple, always finds a solution if the classes are linearly separable. But it does not converge if the classes are non-separable

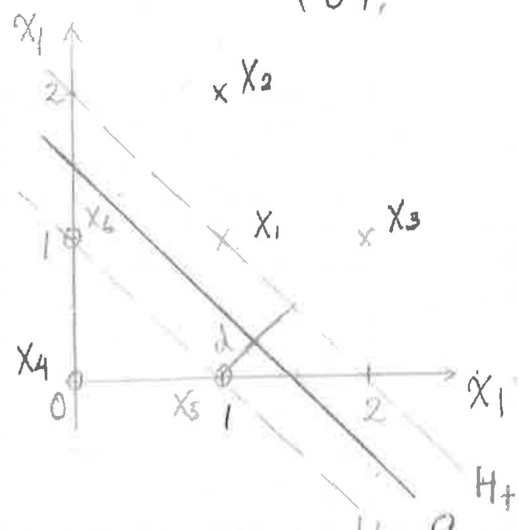
The MSE will always find a solution But the solution might not always be the separating plane ($a^T y_i > 0$) as it depends on the margin vector ($\underline{b}$) chosen.

Ho-Kashyap is more costly. It will find a separating plane if the classes are linearly separable. Again, In the non-linearly separable case, the resut is not valid.

Problem 2:

class 1: $X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $X_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $X_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

class 2: $X_4 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $X_5 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $X_6 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$



$g(x) = -3 + 2x_1 + 2x_2 = 0$

It can be seen that the support vectors are X_1, X_5 and X_6

The separating plane is: $x_1 + x_2 - 1.5 = 0$

$$\underline{a} = [-1.5 \quad 1 \quad 1]^T$$

Normalization requirement in SVM requires $a^T Y_+ = 1$

So $a^T \cdot Y_1 = [-1.5 \quad 1 \quad 1]\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0.5$, hence $c = 2$

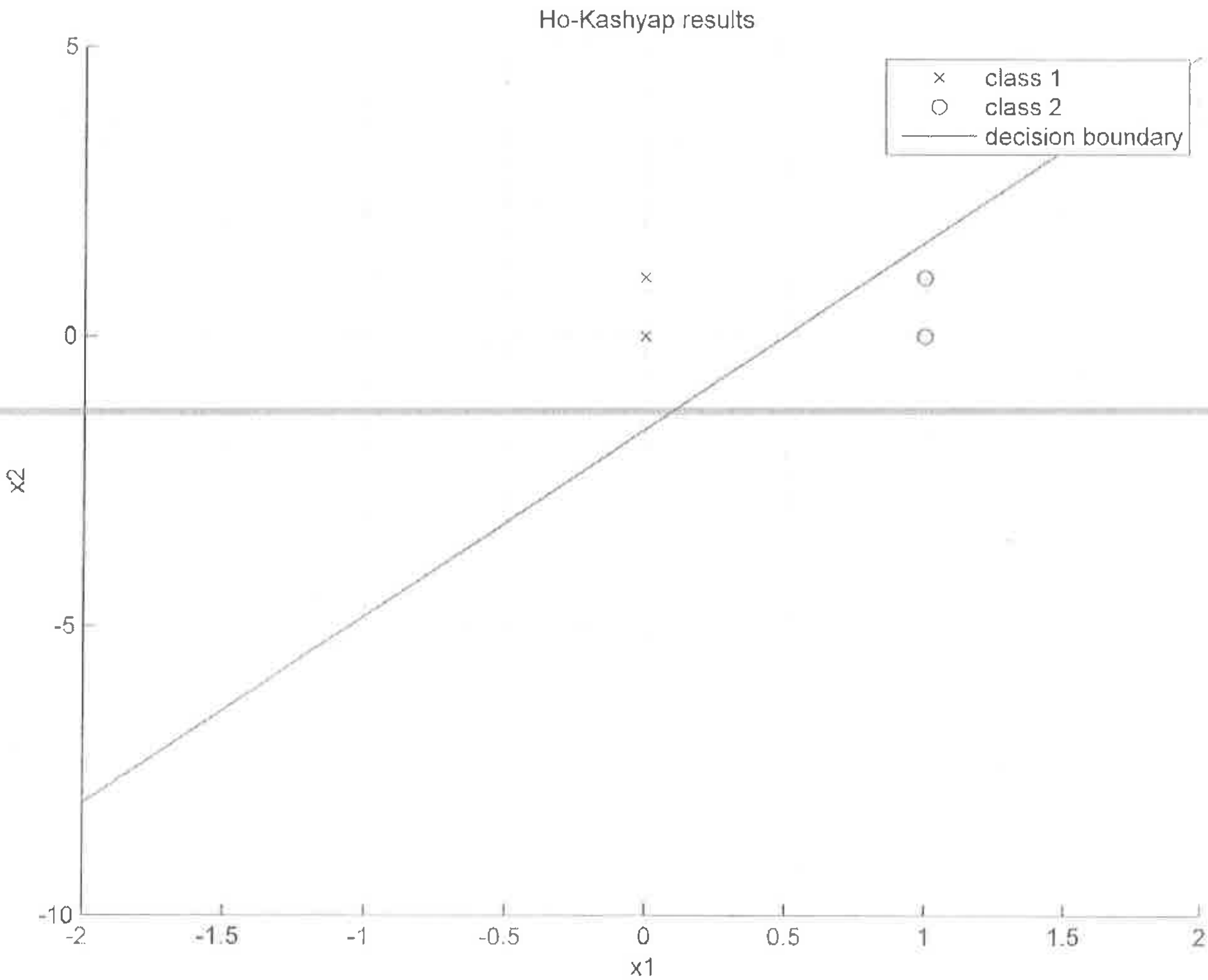
$\underline{a = [-3 \quad 2 \quad 2]}$ $\rightsquigarrow$ $\underline{g(x) = -3 + 2x_1 + 2x_2}$

(Check: $a^T Y_5 = a^T Y_6 = -1$)

Margin d = $\times \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$

TOPS 35500

```
%Ho-Kashyap procedure
%Code for Homework 3/Problem 1 (Part d and e)

%data points
X1=[0 0; 0 1]; %positive examples
X2=[1 1; 1 0]; %negative examples

if(0) %Part (e) of Problem 1
    X1=[0 0; 1 1]; X2=[0 1; 1 0];
end

%Form Y matrix
Y=[1 X1(1,:); 1 X1(2,:); -1 -X2(1,:); -1 -X2(2,:)];

[m,n]=size(Y);

%initialization
a=ones(n,1);
b=ones(m,1);
Kmax=10000;
Bmin=0.001*ones(m,1);

 for k=1:Kmax
    e=Y*a-b;
    b=b+eta*(e+abs(e));
    a=Y\b;
    if(Y*a>0 | abs(e)<Bmin)
        break
    end
 end

 if(k<Kmax)
    %Separating plane found! Plot the datapoints and the decision boundary
    close all
    scatter(X1(:,1),X1(:,2),'x')
    grid, hold on
    scatter(X2(:,1),X2(:,2),'o')
    x1=-2:0.01:2;
    x2=-1/a(3)*(a(1)+a(2)*x1);
    plot(x1,x2)
    xlabel('x1'), ylabel('x2'), title('Ho-Kashyap results')
    legend('class 1','class 2','decision boundary')
 else
      display('Not Linearly Separable')
 end
```



Ho-Kashyap results
class 1
class 2
decision boundary
5
0
-5
-10
x2
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
x1