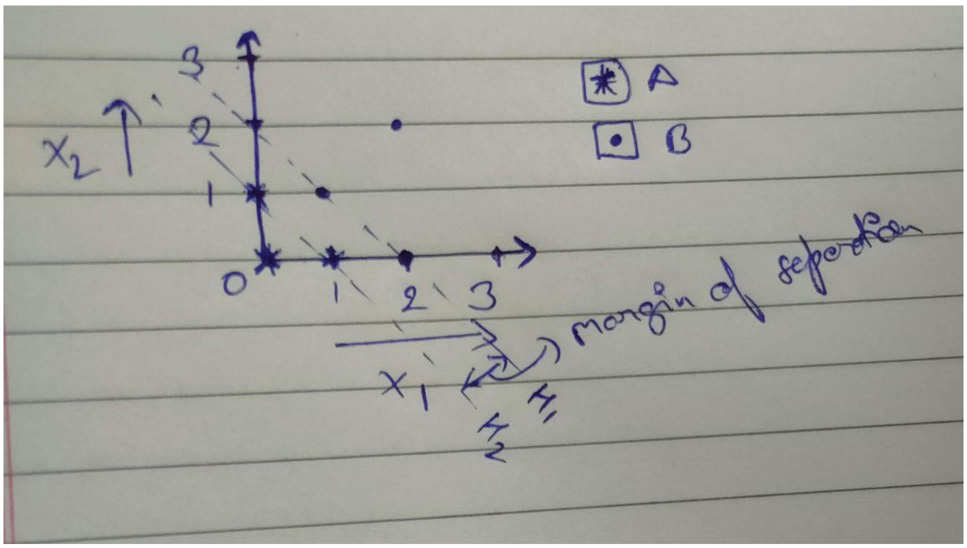
MACHINE LEARNING (CSE343/ECE343)
ASSIGNMENT-3
Sourav Goyal
2020341
SECTION-A

Ans (1)

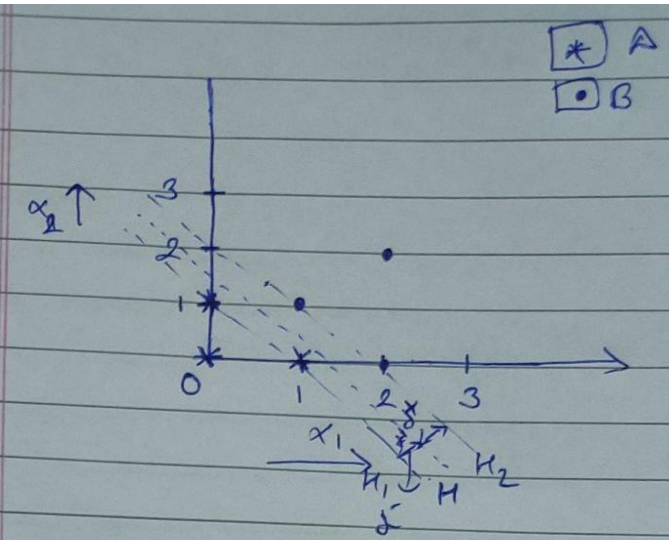
Ans 1.1 :- Yes, these points are linearly separable and plot for this dataset will be :-



By this plot we can clearly see that we have a margin of separation between class A and class B points and by this there can be multiple hyperplanes by which this dataset can be linearly separated easily.

Ans 1.2 :-

Ans 2



$\Rightarrow$ H_1 eqn $\Rightarrow$ $x_1 + x_2 = 1$

$\Rightarrow$ H_2 eqn $\Rightarrow$ $x_1 + x_2 = 2$

$\Rightarrow$ H will be our optimal Hyperplane,

$\Rightarrow$ - - d^+ will be equal to d^- for getting optimal hyperplane,

$\therefore$ H $\Rightarrow$ $x_1 + x_2 = \frac{3}{2}$ because,

$\Rightarrow$ $d^+ = \left| \frac{2+0-\frac{3}{2}}{\sqrt{2}} \right| = \frac{1}{2\sqrt{2}}$

$\Rightarrow$ $d^- = \left| \frac{1+0-\frac{3}{2}}{\sqrt{2}} \right| = \frac{1}{2\sqrt{2}}$

$\therefore$ H will be the optimal Hyperplane with eqn $\Rightarrow$ $x_1 + x_2 = \frac{3}{2}$.

Now, we have our optimal hyperplane equation :- x1 + x2 = 3/2 then our corresponding weights will be,

Let's say weight vector be w = [w1 w2 $]^T$,

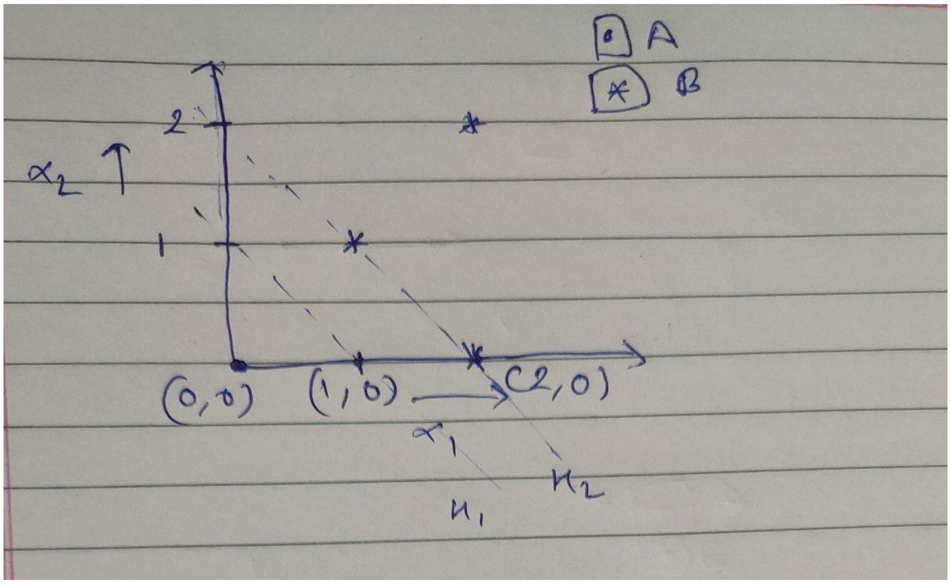
From optimal hyperplane equation we have w1 = 1 and w2 = 1 therefore,

w = [1 1 $]^T$,

Corresponding support vectors for our optimal hyperplane equation from the plot are :- (0,1), (1,0), (1,1), (2,0).

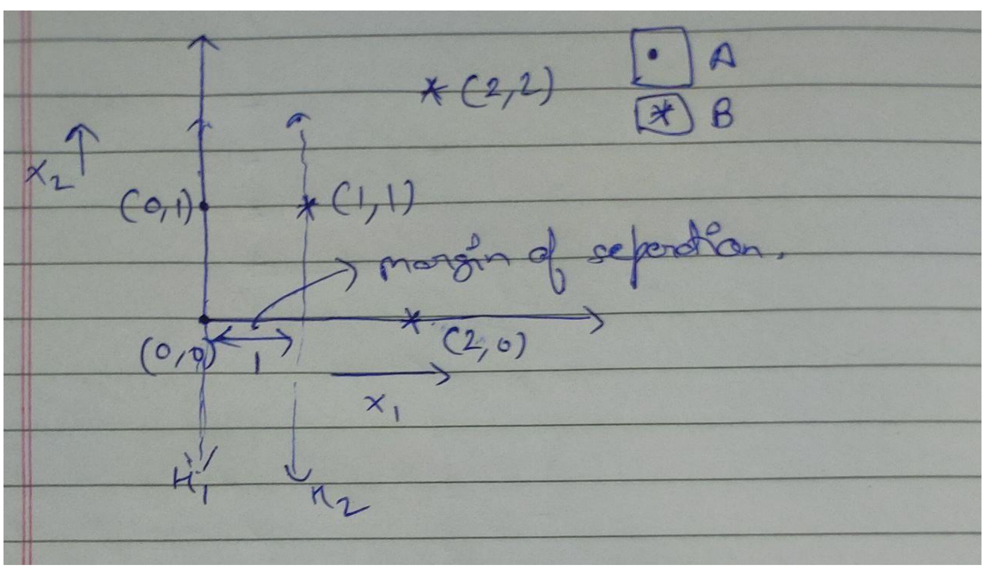
Ans 1.3 :- We have 4 support vectors so we will remove support vector one at a time and observe the changes in optimal margin, so we have corresponding 4 cases :-

Case 1 :- Remove (0,1)
Plot :-



After remove (0,1) we have only 2 sample for Class A and if we draw hyperplane using both samples then it will be parallel to X-axis and data is not separable with Class B so we will have same hyperplane for this case which we have in 2nd part and which optimal margin remains unchanged.

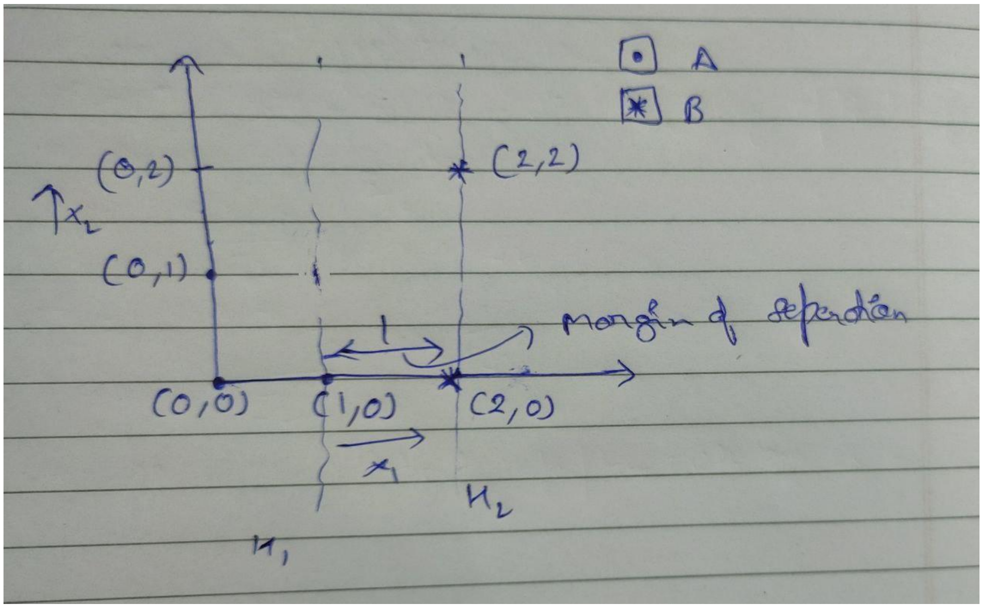
Case 2 :- Remove (1,0)
Plot :-



Now from plot we draw another hyperplane and we have margin of separation is 1 and in 1 st part our optimal margin of separation is $1/\sqrt{2}$, which means our optimal margin is increasing.

Case 3 :- Remove (1,1)

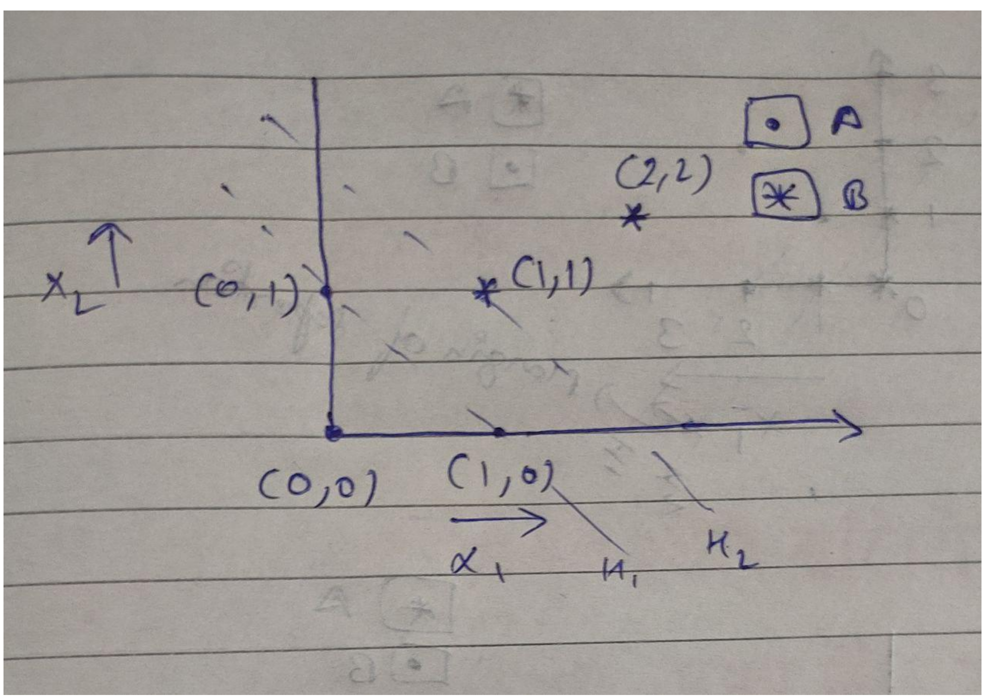
Plot :-



Again from plot we have another hyperplane and have the optimal margin separation is 1 which is greater than $1/\sqrt{2}$ which means in this case also our optimal margin separation is increasing.

Case 4 :- Remove(2,0)

Plot :-



In this case we get the same hyperplane which we have in 2nd part because if we make parallel hyperplane to line connecting the (2,2) and (1,1) it can make separable data with Class A so that's why we have no change in optimal margin separation.

Conclusion :- We have no change in 2 support vectors i.e., (2,0) and (0,1) and there is increase in optimal margin separation in other 2 support vectors i.e., (1,1) and (1,0).

Ans 1.4 :-

In SVM we are maximizing the margin separation between different training labels points from the hyperplane or in other words we find a hyperplane which separates the data at maximum margin of separation.

So when we have found an optimal hyperplane for our general dataset, call it H and have separation M. Now if we remove support vectors then one thing is fixed that this H hyperplane still exists with margin separation M which means optimal margin separation is never decreasing for any dataset on removing any support vectors. So, removing any support vector is removing a constraint from our mathematical equation. In mathematical problem if f(x) is function, so value of f(x) where x has two constraint will be smaller or equal to f(x) where x has only one constraint and same here if we remove any support vector it may be increase the separation margin if any another hyperplane we find otherwise our older hyperplane H is exist with M margin separation which means no affect happens due to remove in support vector.

Conclusion :- Optimal Margin separation can be increased or remains unchanged on removing any support vector for any general dataset but can't be decreased or less than by optimal margin separation which we have with all support vectors.