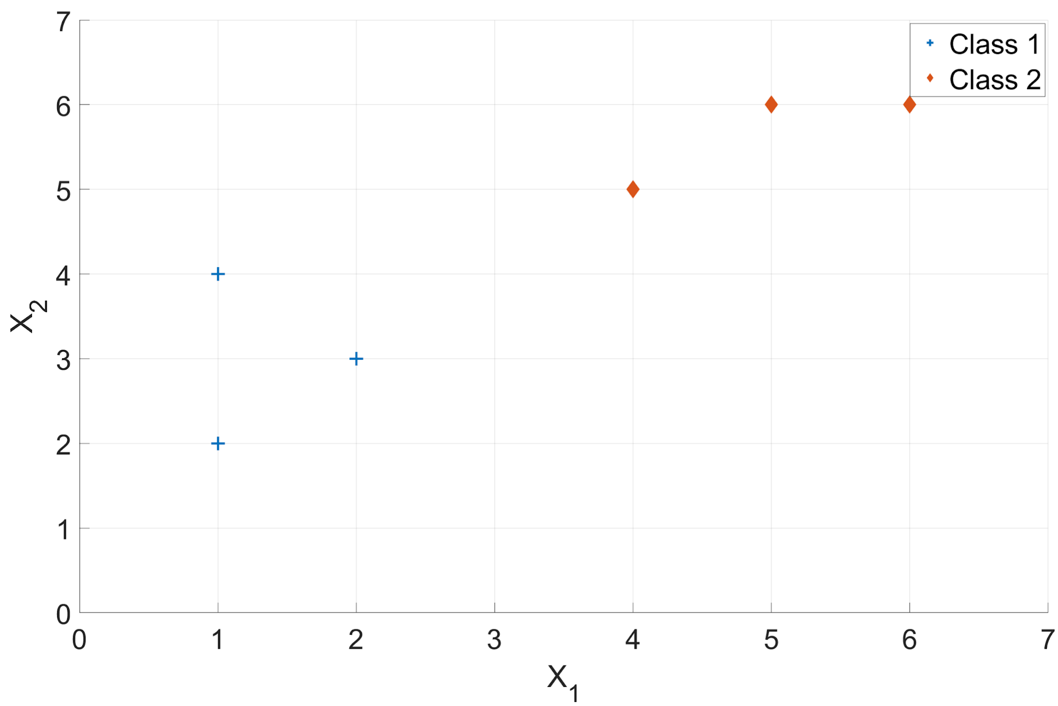
Adam Rankin
adamr14@vt.edu

1. Hard Margin SVM (5 Points)



A decision boundary which leads to the largest margin is learnt by the Support Vector Machines. Consider that SVM trains on *Dataset I*, given in the above figure consisting of 6 data points in total as shown. There are two class labels, namely - Class Label 1, which is denoted by a blue plus and Class Label 2, which is denoted by a red diamond.

(a) [1 Points] Given the *Dataset I* above and the support vectors (4, 5) and (2, 3), what is the slope of the decision boundary?

Slope is -1

(b) [3 Points] Based on the slope calculated above, find the normal vector (*w)* and bias (w_0). Based on the above findings, what is the equation to represent the decision boundary? Note - decision boundary should follow $w_0 + w_1x_1 + w_2x_2 = 0$ and the normal vector is of the form (w_1, w_2). Use the point slope form.

Normal vector: (-1, -1)

Bias: 7
Decision Boundary: $7 + -x_1 - x_2 = 0$

(c) [1 Points] Based on the decision boundary, provide any other support vectors.

(1 , 4)

2. Kernelized SVM (10 Points)

Suppose we have two sets of data points in two-dimensional space. One set represents the positive class (+1) and the second set represents the negative class (-1).

1st Class: {(2,2), (2,-2), (-2,-2), (-2,2)}
2nd Class: {(1,1), (1,-1), (-1,-1), (-1,1)}

By plotting these points on the 2-D plane, we can easily infer that these points are not linearly separable. No problem! We are giving you a Kernel or mapping function that will help you find the decision boundary.

$$\Phi_1\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{cases}\begin{pmatrix}4-x_2+|x_1-x_2|\\4-x_1+|x_1-x_2|\end{pmatrix}, & if\ \sqrt{x_1^2+x_2^2}>2\\ \begin{pmatrix}x_1\\x_2\end{pmatrix}, & otherwise\end{cases}$$

Now just follow these following steps which will enable you to classify the point of your own choice.

(a) [2 Points] Using this Kernel function, find the new feature representation of each data point from both classes.

1st Class: {(2,2), (2,-2), (-2,-2), (-2,2)}
(2,2) -> (2, 2)
(2, -2)-> (10, 6)
(-2, -2) -> (6, 6)
(-2, 2) -> (6, 10)

2nd Class: {(1,1), (1,-1), (-1,-1), (-1,1)}
(1, 1) - > (1, 1)
(1, -1) - > (1, -1)
(-1, -1) -> (-1, -1)
(-1, 1) -> (-1, 1)

(b) [2 Points] Now find the two suitable support vectors and their corresponding augmented support vectors. Given a support vector (x_1, x_2), we define its augmented support vector as (x_1, x_2, 1) by adding the third coordinate of bias = 1 in support vector.

s1: (2,2)
Aug: (2,2,1)

s2: (1,1)
Aug: (1,1,1)

(c) [2 Points] Considering *Two* augmented support vectors where s_1 is for positive class (+1) and s_2 is for negative class (-1), we can use the following two equations to learn a classification function:

$$\left(\alpha_1 \bullet \Phi_1\left(s_1\right) + \alpha_2 \bullet \Phi_1\left(s_2\right)\right)^T \bullet \Phi_1\left(s_1\right) = y_1 = 1$$

$$\left(\alpha_1 \bullet \Phi_1\left(s_1\right) + \alpha_2 \bullet \Phi_1\left(s_2\right)\right)^T \bullet \Phi_1\left(s_2\right) = y_2 =- 1$$

Based on the two augmented support vectors obtained from part (b), compute the two parameters α_1 and α_2.

$$\left(\alpha_1 \bullet (2, 2, 1) + \alpha_2 \bullet (1, 1, 1)\right)^T \bullet (2, 2, 1) = 1$$

$$(2a_1+a_2,\ 2a_1+a_2, a_1 + a_2)^T \text{ dot } (2,2,1) = 1$$

$$9a_1+5a_2 = 1 \quad \textbf{(1)}$$

$$\left(\alpha_1 \bullet (2, 2, 1) + \alpha_2 \bullet (1, 1, 1)\right)^T \bullet (1, 1, 1) =- 1$$

$$(2a_1+a_2,\ 2a_1+a_2, a_1 + a_2)^T \text{ dot } (1,1,1) = 1$$

$$5a_1 + 3a_2 = -1 \quad \textbf{(2)}$$

$a_1 = 4$,
$a_2 = -7$

(d) [2 Points] Continuing from part (c), find the hyperplane: $y = wx$ which can be used as the final classifier. Here w can be estimated based on the equation: $w = \sum_i \alpha_i \bullet s_i$, where α_i is estimated from part (c) and s_i denotes an augmented support vector. Provide the estimated parameters w.

w = s1*a1 + s2* a2 = (8, 8, 4) + (-7, -7, -7) = (1, 1, -3)

(e) [2 Points] Given a new data point (4, 5), using the kernel function and the hyperplane learned in part (d) to classify the new data point as positive or negative.

y = wx = 4*1 + 5*1 - 3 = 6 > 0

(4, 5) classified as positive