

SUPPORT VECTOR MACHINES

Scholastic Video Book Series

Part 2

Non-Linear Support Vector Machines

(with English Narrations)

Scholastic

home.video.tutor

@gmail.com

<http://scholastictutors.webs.com>

(<http://scholastictutors.webs.com/Scholastic-Book-SupportVectorM-Part02-2014-06-03.pdf>)

Scholastic

home.video.tutor

@gmail.com

©Scholastic Tutors June, 2014)

AVAILABLE AT:

Onebyzero Edu - Organized Learning, 9110-20-140603-1
The Comprehensive Academic Study Platform for University Students in Bangladesh (www.onebyzeroedu.com)

Scholastic
home.video.tutor
@gmail.com

Support Vector Machines - #2

Non-Linear Support Vector Machines

Click here to see the video

<http://youtu.be/6cJoCCn4wuU>

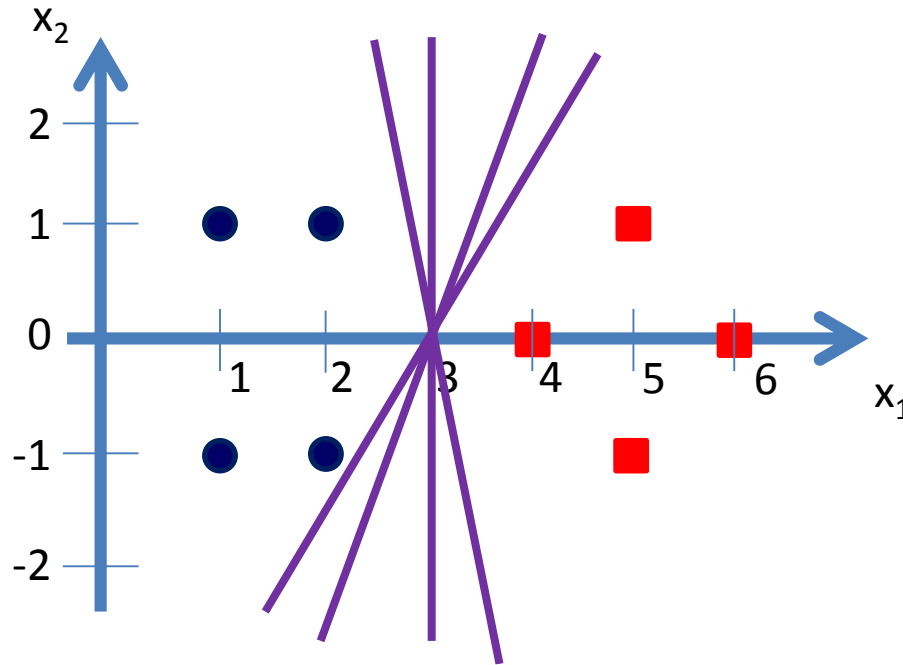
<http://scholastictutors.webs.com>

AVAILABLE AT:

(SVM-002)
Onebyzero Edu - Organized Learning, Smooth Career
The Comprehensive Academic Study Platform for University Students in Bangladesh (www.onebyzeroedu.com)

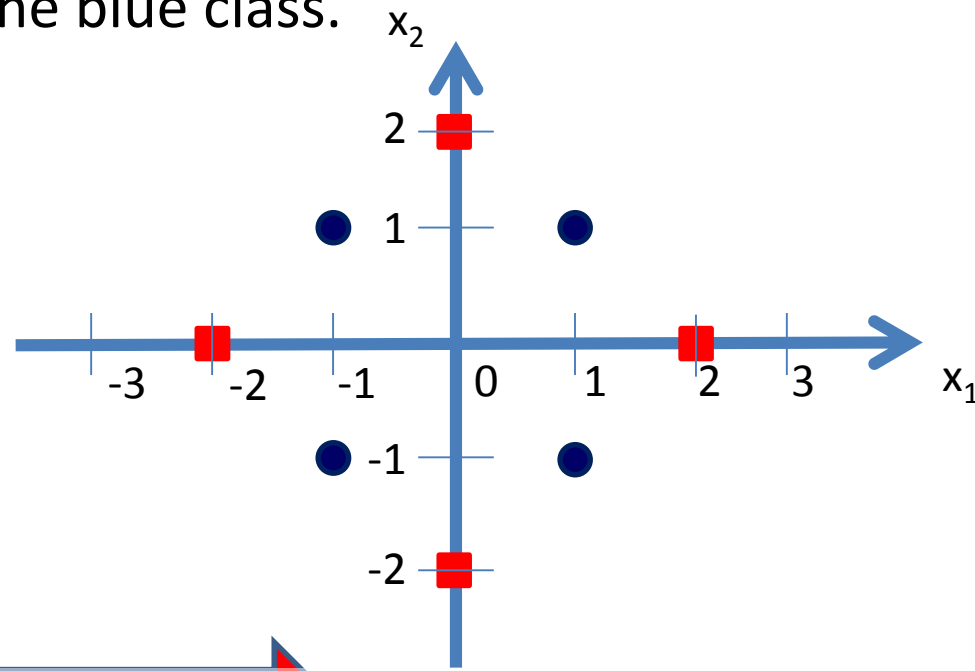
Support Vector Machines

- Support Vector Machine (SVM) algorithms are used in Classification.
- Classification can be viewed as the task of separating classes in feature space.

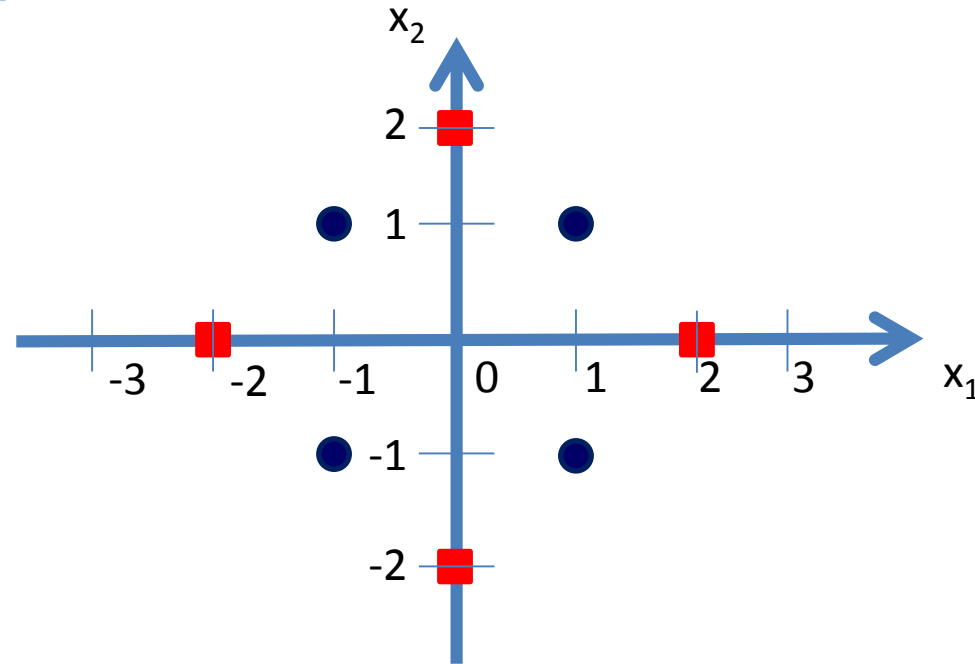


Support Vector Machines

- We looked at the linear type SVM in Part I of our SVM lesson series. Here we will look at an example like the one given below and find out how to carry out the classification.
- Obviously there is no clear separating hyperplane between the red class and the blue class.



Support Vector Machines

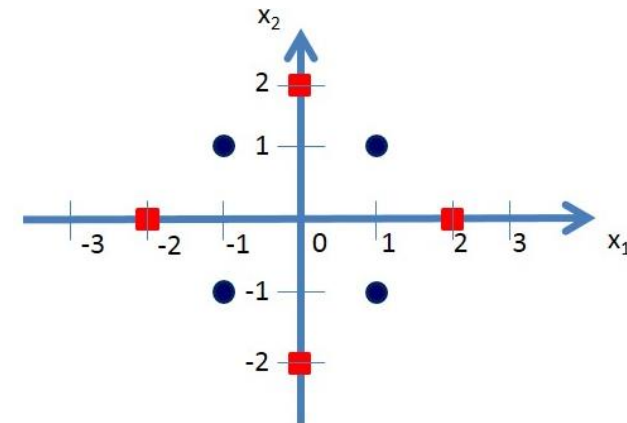


- Blue class vectors are: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
- Red class vectors are: $\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}$

Support Vector Machines

- Here we need to find a non-linear mapping function Φ which can transform these data in to a new feature space where a separating hyperplane can be found.
- Let us consider the following mapping function.

$$\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$



Support Vector Machines

- Now let us transform the blue and red class vectors using the non-linear mapping function Φ .

$$\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$

- Blue class vectors are: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ no change since $\sqrt{x_1^2 + x_2^2} < 2$ for all the vectors

Support Vector Machines

$$\bullet \quad \Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$

• Let us take Red class vectors : $\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}$

$$\bullet \quad \Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 - 2 + (2 - 0)^2 \\ 6 - 0 + (2 - 0)^2 \end{pmatrix} = \begin{pmatrix} 8 \\ 10 \end{pmatrix}$$

$$\bullet \quad \Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 - 0 + (0 - 2)^2 \\ 6 - 2 + (0 - 2)^2 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \end{pmatrix}$$

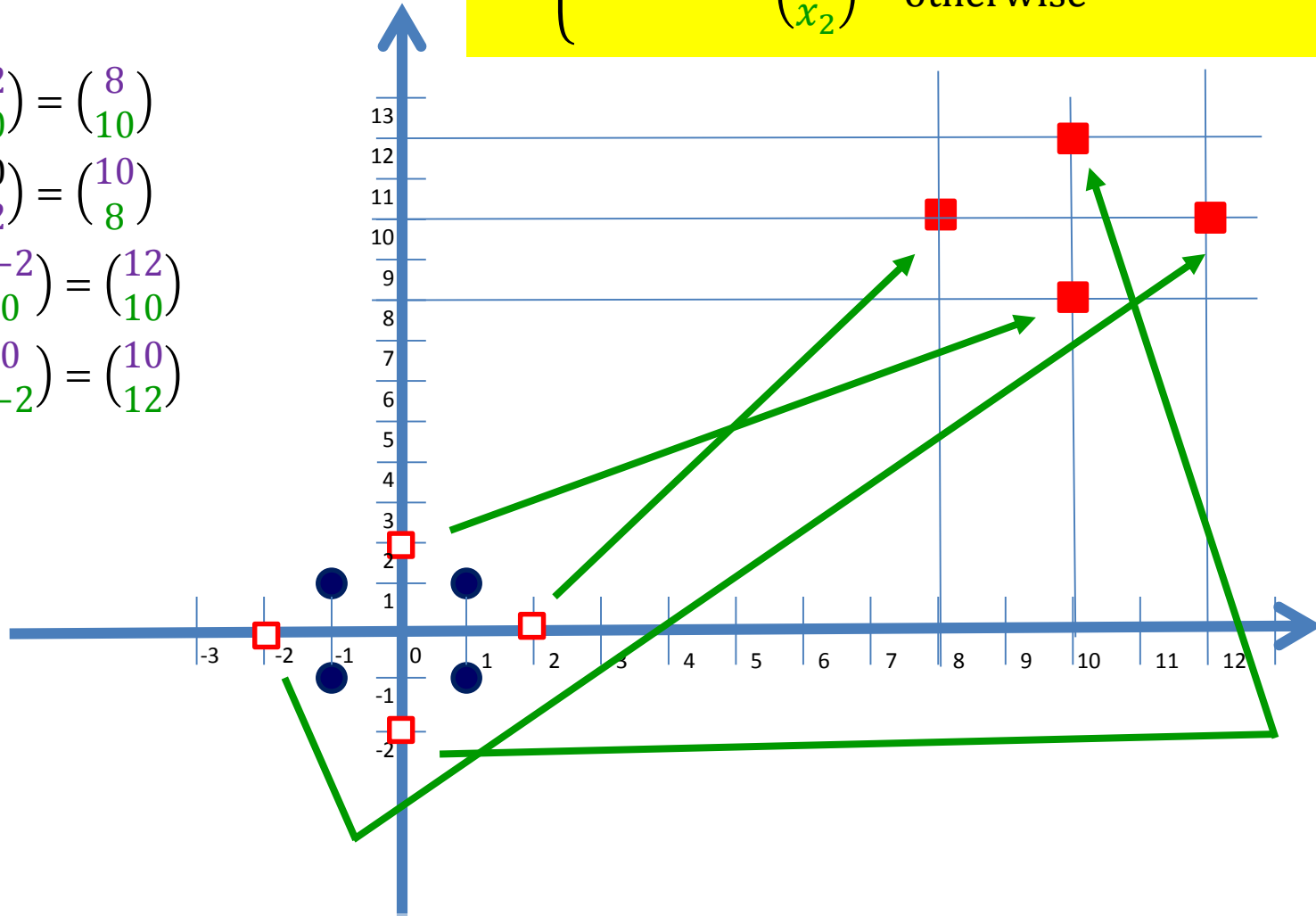
$$\bullet \quad \Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 + 2 + (-2 - 0)^2 \\ 6 - 0 + (-2 - 0)^2 \end{pmatrix} = \begin{pmatrix} 12 \\ 10 \end{pmatrix}$$

$$\bullet \quad \Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 6 - 0 + (0 + 2)^2 \\ 6 + 2 + (0 + 2)^2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$$

Vectors After Transforming

$$\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$

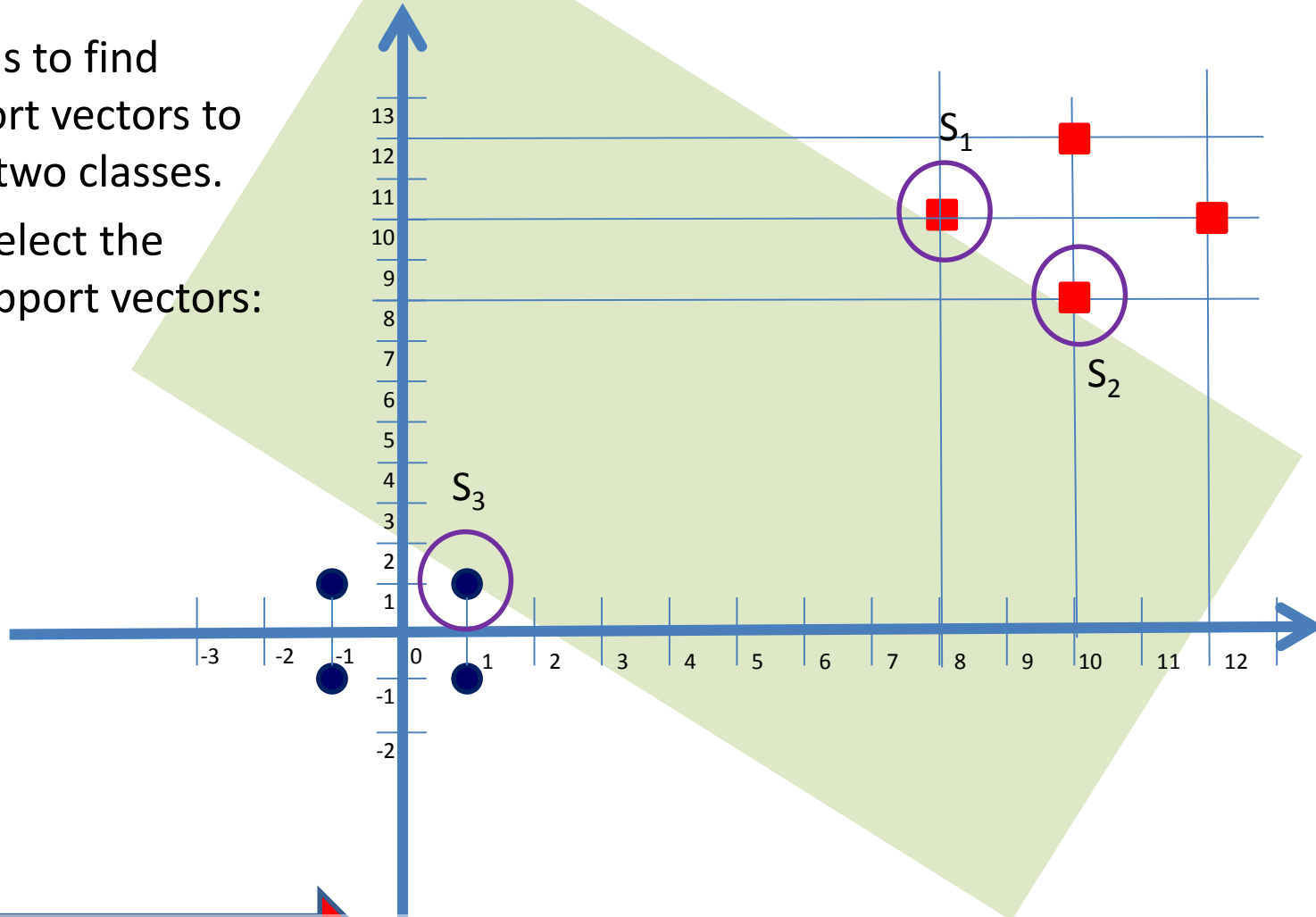
- $\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 10 \end{pmatrix}$
- $\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \end{pmatrix}$
- $\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 12 \\ 10 \end{pmatrix}$
- $\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$



AVAILABLE AT:

Selection of Support Vectors

- Now our task is to find suitable support vectors to classify these two classes.
- Here we will select the following 3 support vectors:
- $S_1 = \begin{pmatrix} 8 \\ 10 \end{pmatrix}$,
- $S_2 = \begin{pmatrix} 10 \\ 8 \end{pmatrix}$,
- and $S_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$



Augmented Support Vectors

- Here we will use vectors augmented with a 1 as a bias input, and for clarity we will differentiate these with an over-tilde. That is:

$$s_1 = \begin{pmatrix} 8 \\ 10 \end{pmatrix}$$

$$s_2 = \begin{pmatrix} 10 \\ 8 \end{pmatrix}$$

$$s_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\widetilde{s}_1 = \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix}$$

$$\widetilde{s}_2 = \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix}$$

$$\widetilde{s}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Finding Parameters

- Now we need to find 3 parameters α_1 , α_2 , and α_3 based on the following 3 linear equations:

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_1 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_1 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_1 = +1 \text{ (+ve class)}$$

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_2 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_2 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_2 = +1 \text{ (+ve class)}$$

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_3 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_3 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_3 = -1 \text{ (-ve class)}$$

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_1 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_1 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_1 = +1 \text{ (+ve class)}$$

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_2 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_2 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_2 = +1 \text{ (+ve class)}$$

$$\alpha_1 \widetilde{S}_1 \cdot \widetilde{S}_3 + \alpha_2 \widetilde{S}_2 \cdot \widetilde{S}_3 + \alpha_3 \widetilde{S}_3 \cdot \widetilde{S}_3 = -1 \text{ (-ve class)}$$

- Let's substitute the values for $\widetilde{S}_1$, $\widetilde{S}_2$ and $\widetilde{S}_3$ in the above equations.

$$\widetilde{S}_1 = \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \quad \widetilde{S}_2 = \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \quad \widetilde{S}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} = +1$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} = +1$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -1$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} = +1$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} = +1$$

$$\alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -1$$

- After multiplication we get:

$$165 \alpha_1 + 161 \alpha_2 + 19 \alpha_3 = +1$$

$$161 \alpha_1 + 165 \alpha_2 + 19 \alpha_3 = +1$$

$$19 \alpha_1 + 19 \alpha_2 + 3 \alpha_3 = -1$$

- Simplifying the above 3 simultaneous equations we get: $\alpha_1 = \alpha_2 = 0.859$ and $\alpha_3 = -1.4219$.

Discriminating Hyper Plane

- The hyper plane that discriminates the positive class from the negative class is given by:

$$\tilde{w} = \sum_i \alpha_i \tilde{S}_i$$

- Substituting the values we get:

$$\begin{aligned}\tilde{w} &= \alpha_1 \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ \tilde{w} &= (0.0859) \cdot \begin{pmatrix} 8 \\ 10 \\ 1 \end{pmatrix} + (0.0859) \cdot \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} + (-1.4219) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.1243 \\ 0.1243 \\ -1.2501 \end{pmatrix}\end{aligned}$$

Equation for Discriminating Hyper Plane

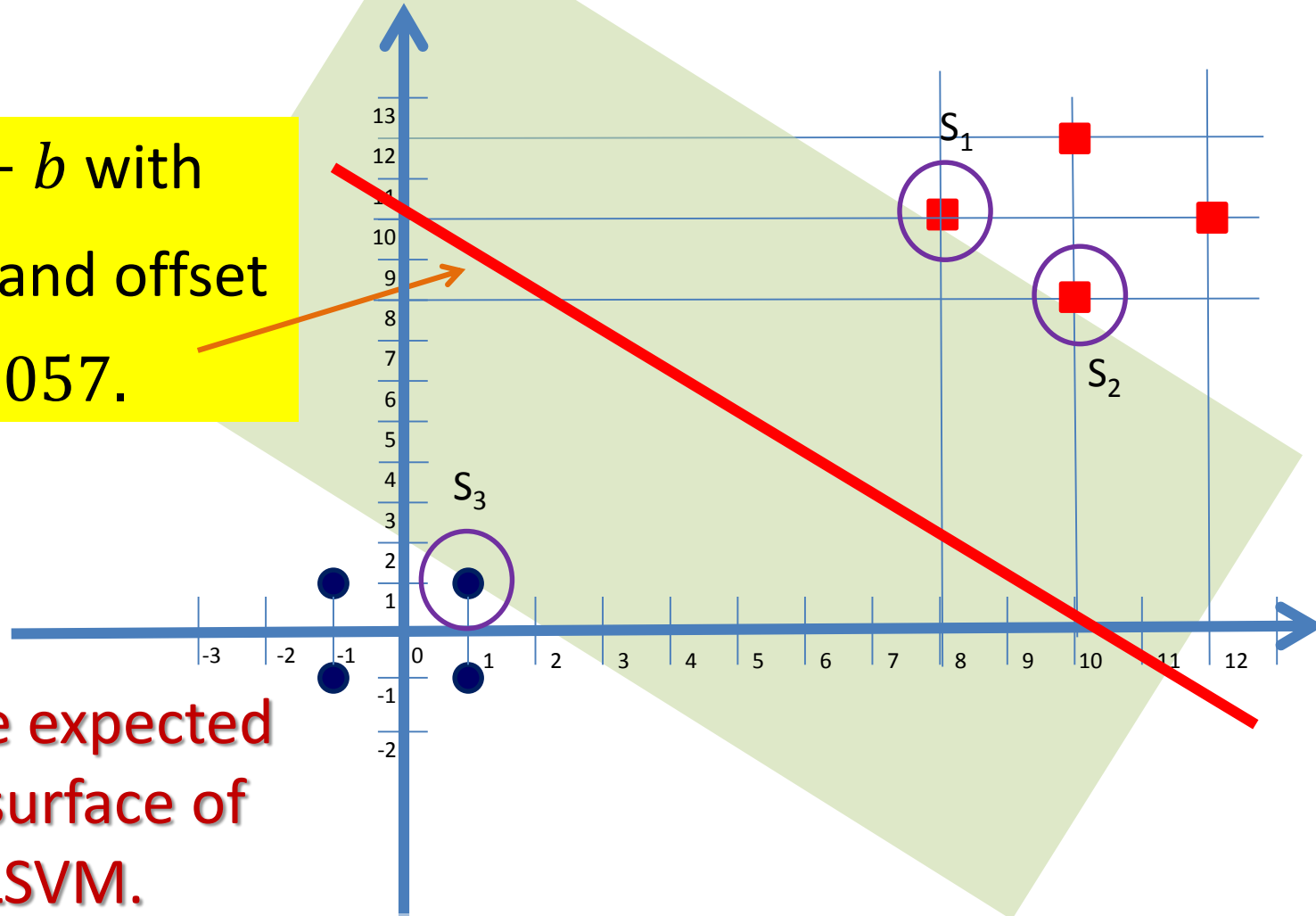
- Our vectors are augmented with a bias.
- Hence we can equate the entry in $\tilde{w}$ as the hyper plane with an offset b .
- Therefore the separating hyper plane equation

$$y = wx + b \text{ with } w = \begin{pmatrix} 0.1243/0.1243 \\ 0.1243/0.1243 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\text{and an offset } b = -\frac{1.2501}{0.1243} = -10.057.$$

Hyper Plane Seperating the Two Classes

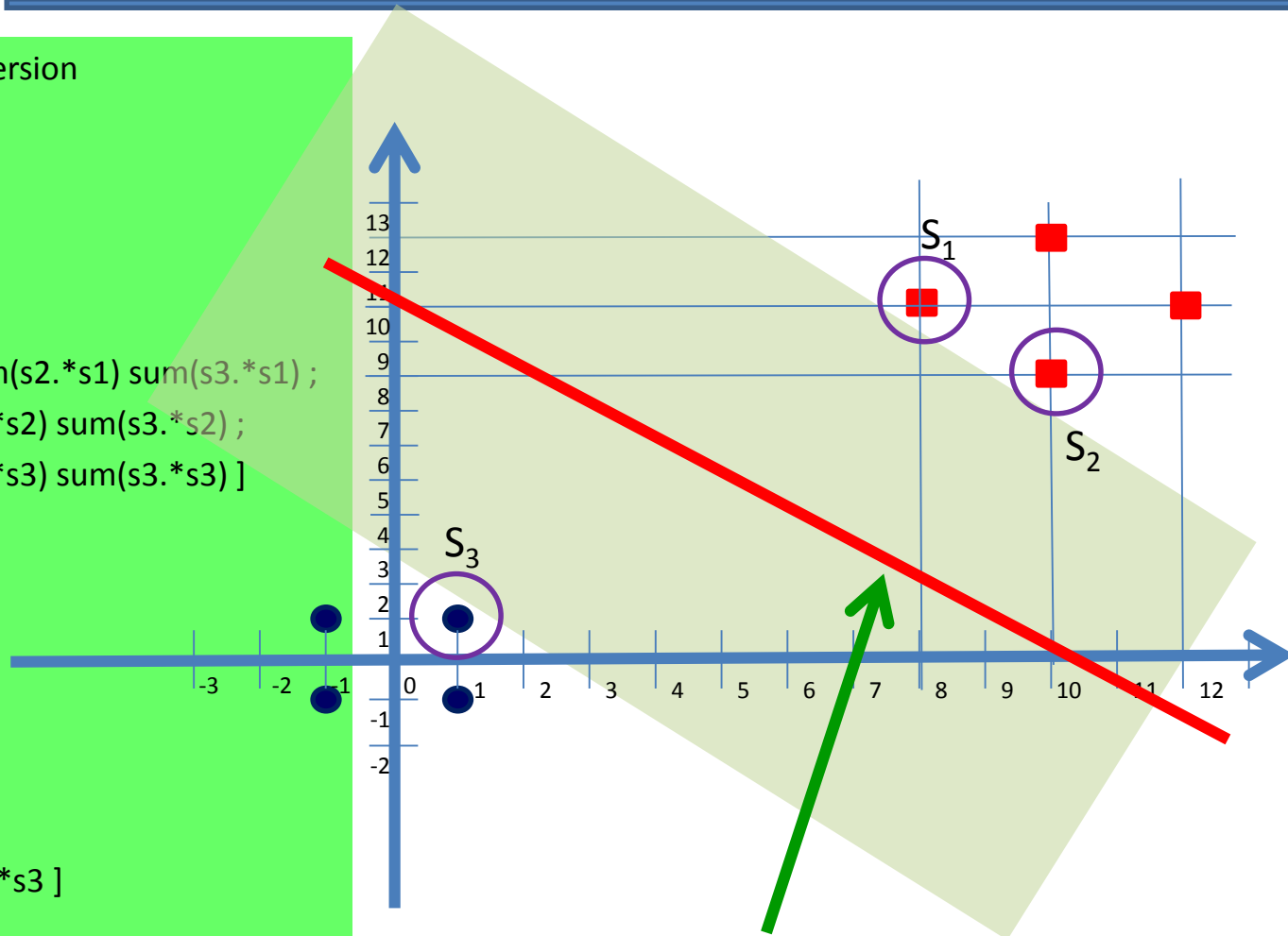
- $y = wx + b$ with $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and offset $b = -10.057$.



- This is the expected decision surface of the Non LSVM.

Support Vector Machines – Programme Code Example 1

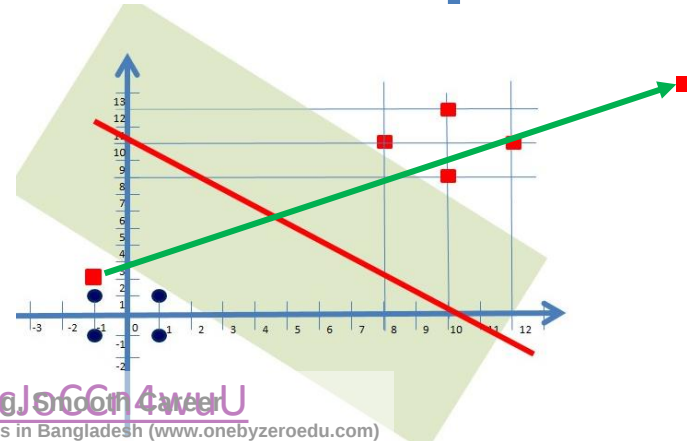
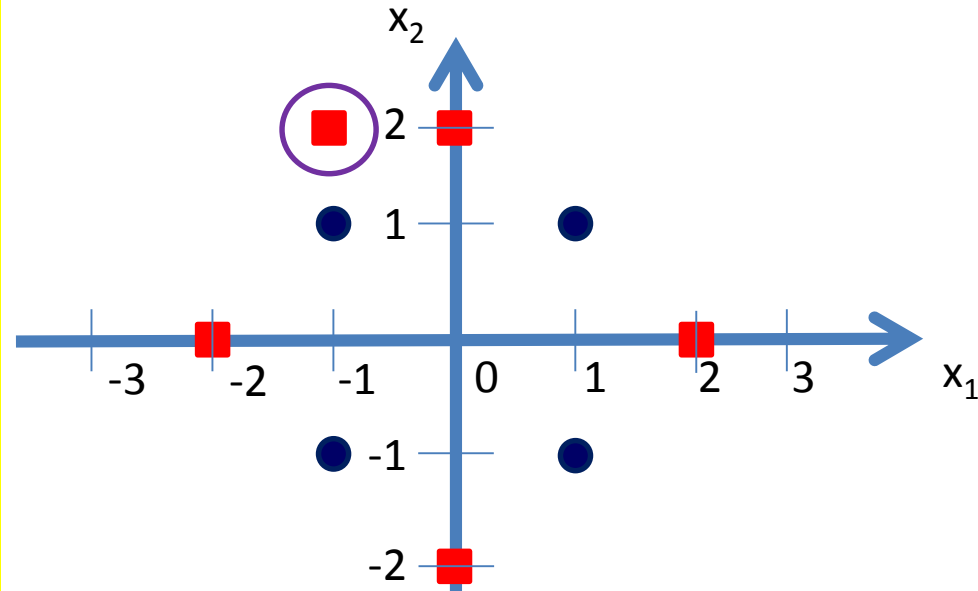
```
% 3 support vector version
•
•
• s1 = [ 8 10 1 ];
• s2 = [ 10 8 1 ];
• s3 = [ 1 1 1 ];
•
• A = [ sum(s1.*s1) sum(s2.*s1) sum(s3.*s1) ;
•       sum(s1.*s2) sum(s2.*s2) sum(s3.*s2) ;
•       sum(s1.*s3) sum(s2.*s3) sum(s3.*s3) ]
•
• Y = [ -1 -1 +1 ]
• X = Y/A
•
• p = X(1)
• q = X(2)
• r = X(3)
•
• W = [ p*s1 + q*s2 + r*s3 ]
```



When you run you should get: $\tilde{W} = [0.125 \ 0.125 \ -1.25]$. This is the line shown passing through $x_1 = 1.25 / 0.125 = 10$ with a gradient $-0.125 / 0.125 = -1$. (Note the values here are approximated)

- Let's consider a classification example here.
- Let's classify the point $(x_1, x_2) = (-1, 2)$.
- $\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 + 1 + (-1 - 2)^2 \\ 6 - 2 + (-1 - 2)^2 \end{pmatrix} = \begin{pmatrix} 16 \\ 13 \end{pmatrix}$
- $w \cdot \Phi(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 16 \\ 13 \end{pmatrix} = 29 > 10$
- Hence this point belongs to the red class.

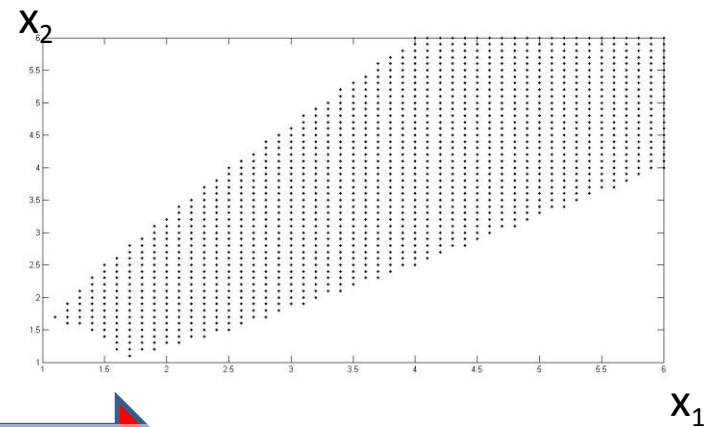
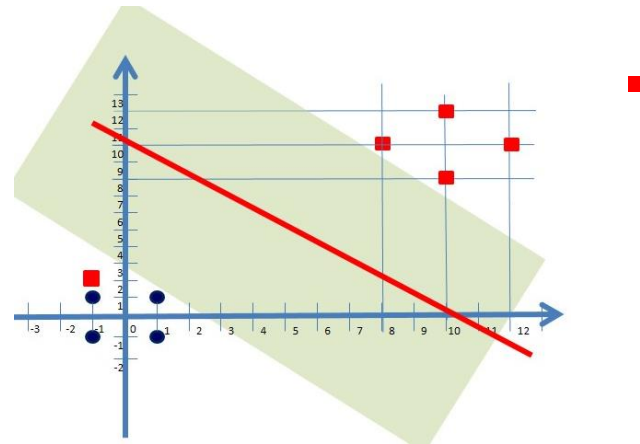
$$\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$



Important Note

- Please note that due to the nature of the non linear transfer function and the selected support vectors used, here not all the points with $\sqrt{x_1^2 + x_2^2} \geq 2$ will transform above the current classification boundary.
- The Figure shows (in dots) the region that will **not correctly** classify using the current mapping function.
- Feature vectors that lie in all the other regions including the vectors in negative quadrants can be classified correctly using this non-linear SVM approach.

$$\Phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 6 - x_1 + (x_1 - x_2)^2 \\ 6 - x_2 + (x_1 - x_2)^2 \end{pmatrix} & \text{if } \sqrt{x_1^2 + x_2^2} \geq 2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{otherwise} \end{cases}$$



Scholastic
home.video.tutor
@gmail.com

Support Vector Machines - #2

Non-Linear Support Vector Machines

Click here to see the video

<http://youtu.be/6cJoCCn4wuU>

END of the Book

If you like to see similar solutions to any Mathematics problems please
contact us at: home.video.tutor@gmail.com with your request.

<http://scholastic-videos.com>

Scholastic
home.video.tutor
@gmail.com

Scholastic Video Book Series
Support Vector Machines (SVM)
Part 2
(Non Linear SVM)
(with English Narrations)
(END)

Videos at: <http://www.youtube.com/homevideotutor>

(<http://scholastic tutors.webs.com/Scholastic-Book-SupportVectorM-Part02-2014-06-03.pdf>)

Scholastic

home.video.tutor
Onebyzero.Edu
@gmail.com

©Scholastic Tutors (Jan, 2014)

AVAILABLE AT:

Organized Learning Site On 2014-06-03-1
A Comprehensive Academic Study Platform for University Students in Bangladesh (www.onebyzeroedu.com)