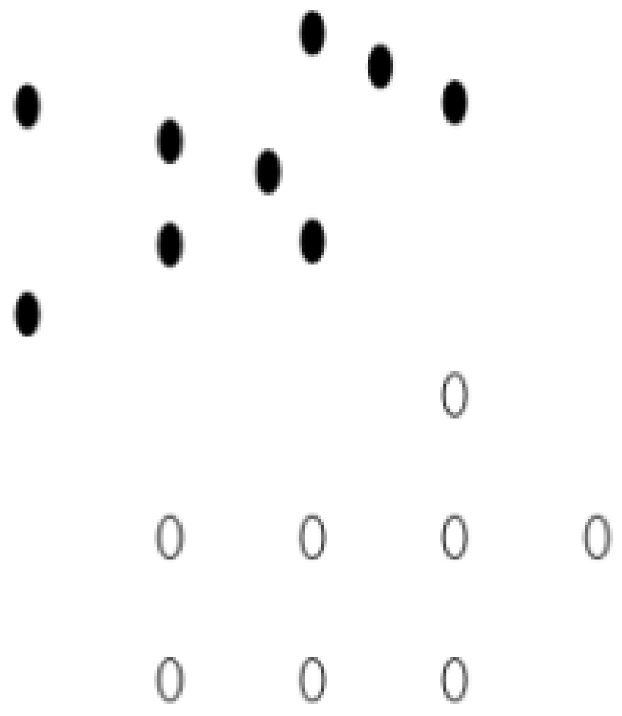
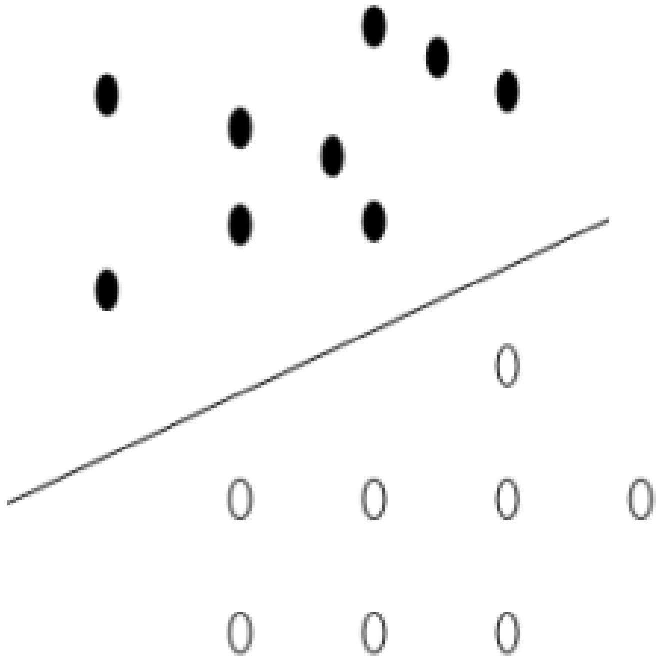
Lesson 6

Support vector machine (SVM)

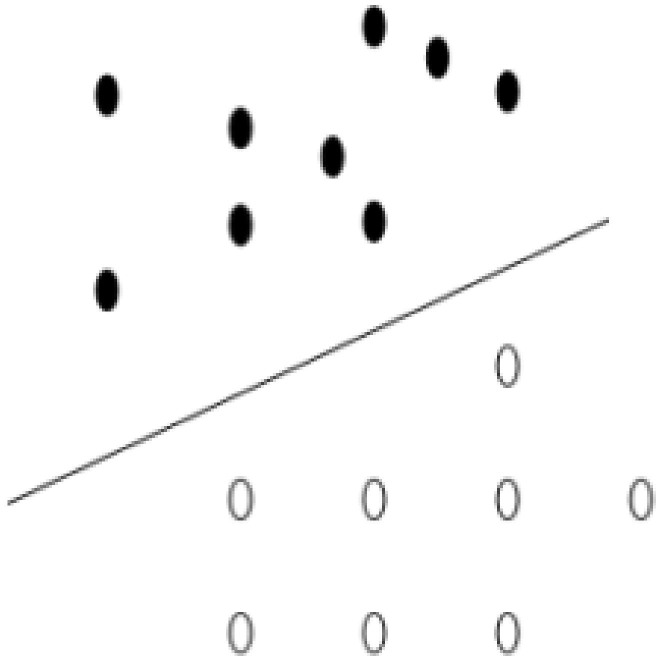
Support vector machine



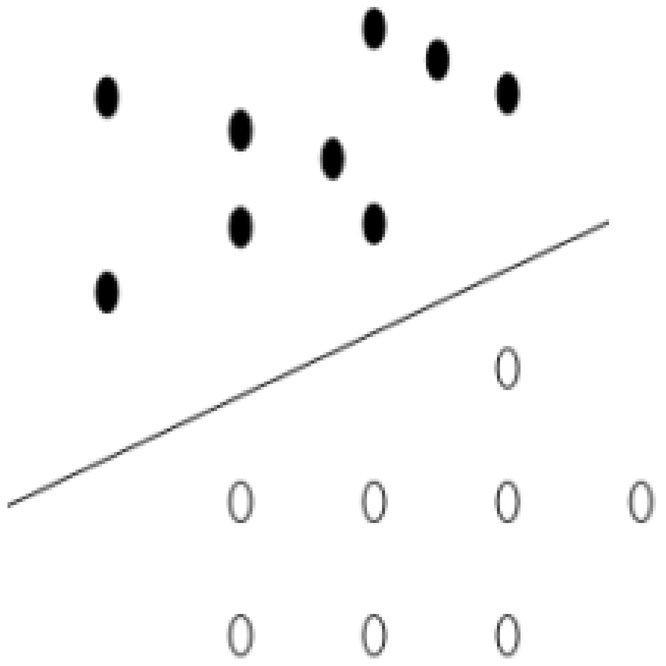
Support vector machine

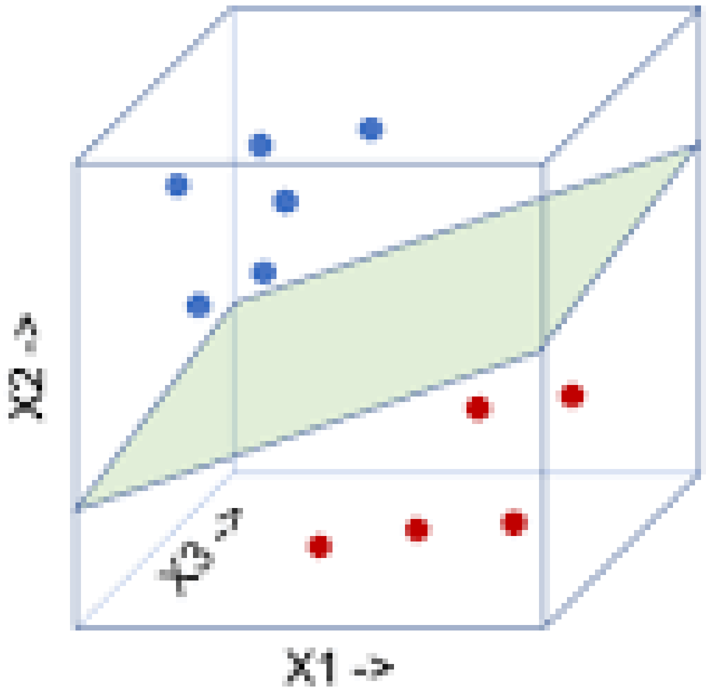


Support vector machine

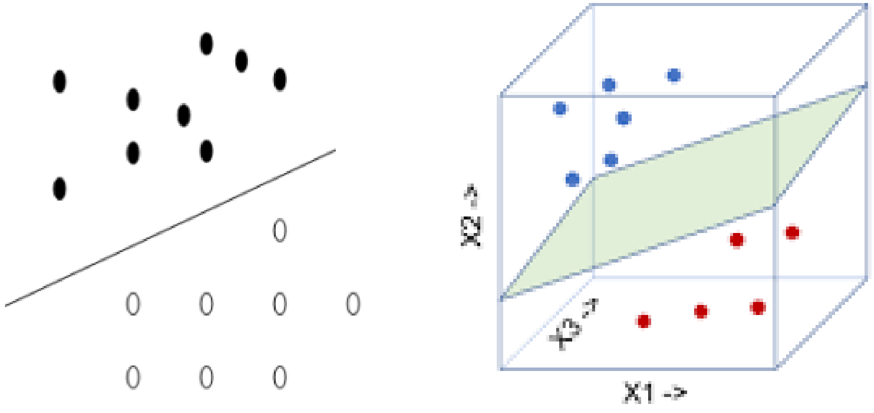


Support vector machine

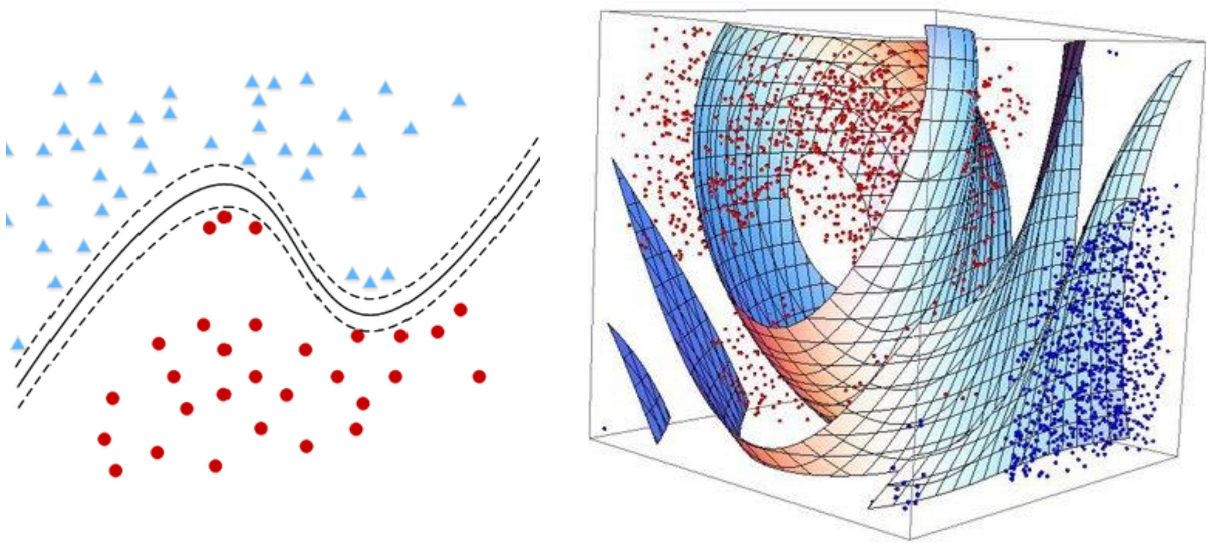




Support vector machine – Linear or Non-linear

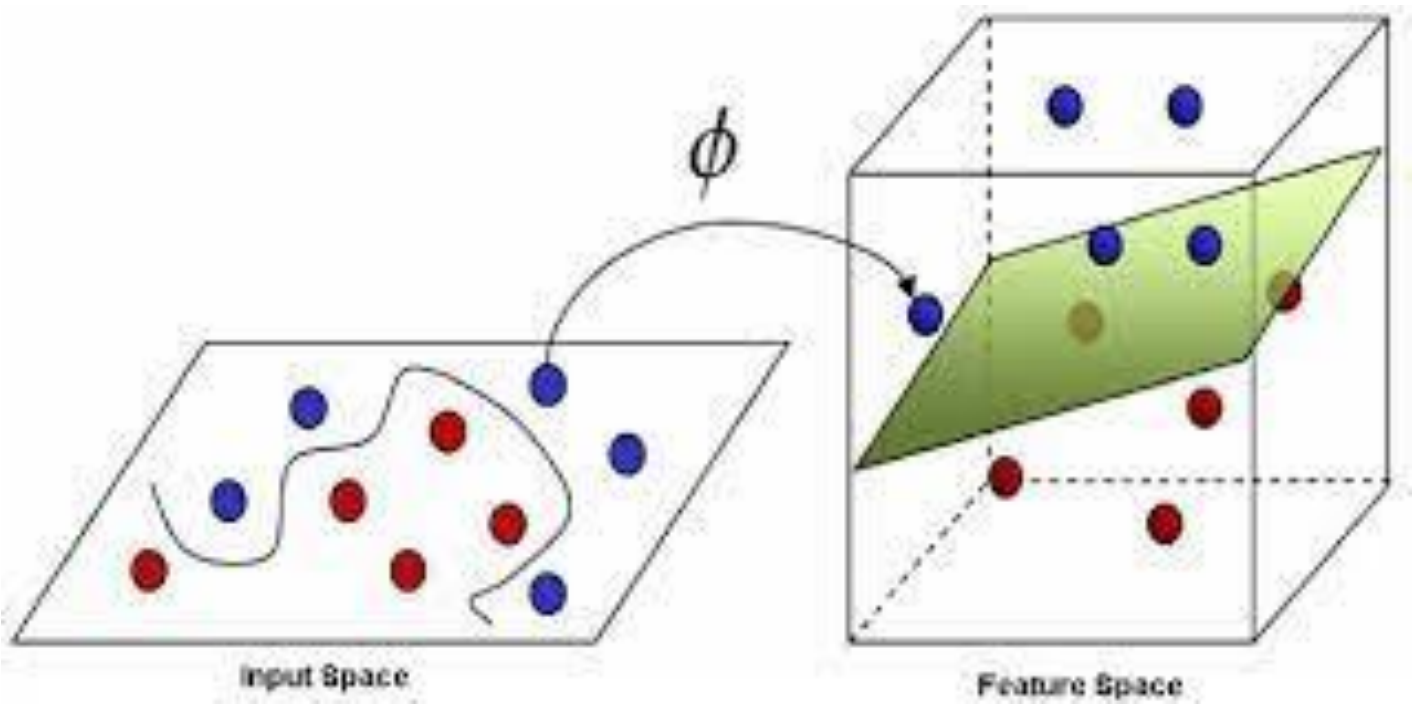


Linear SVM

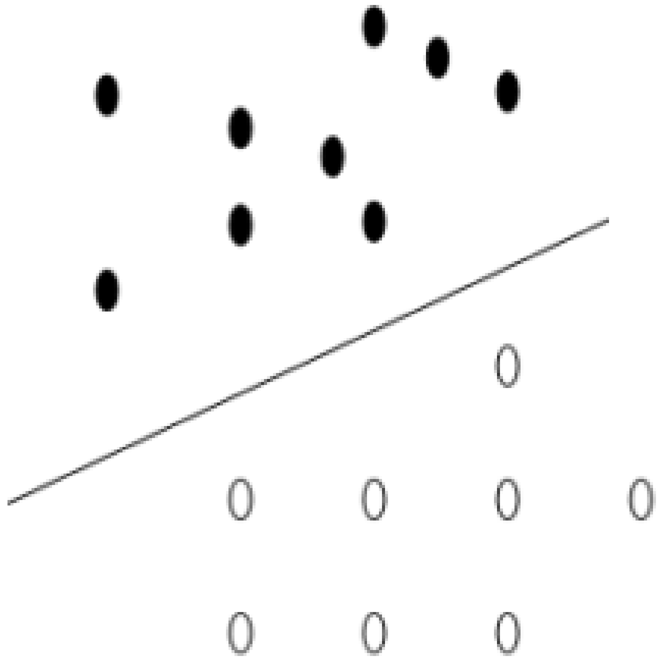


Non-Linear SVM

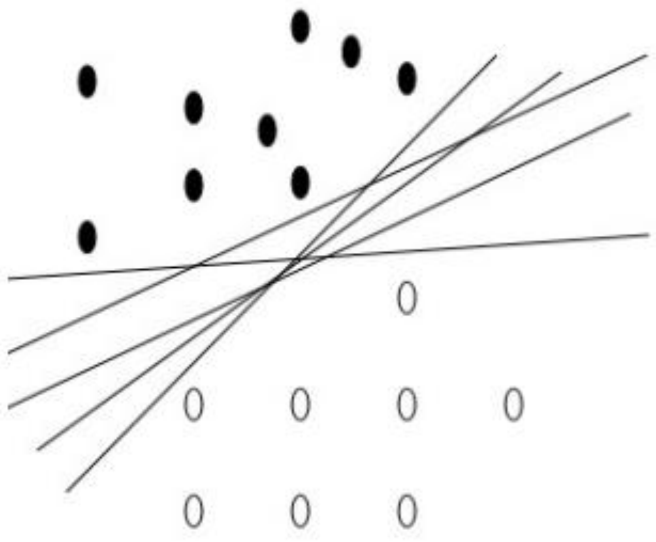
Non-Linear to Linear



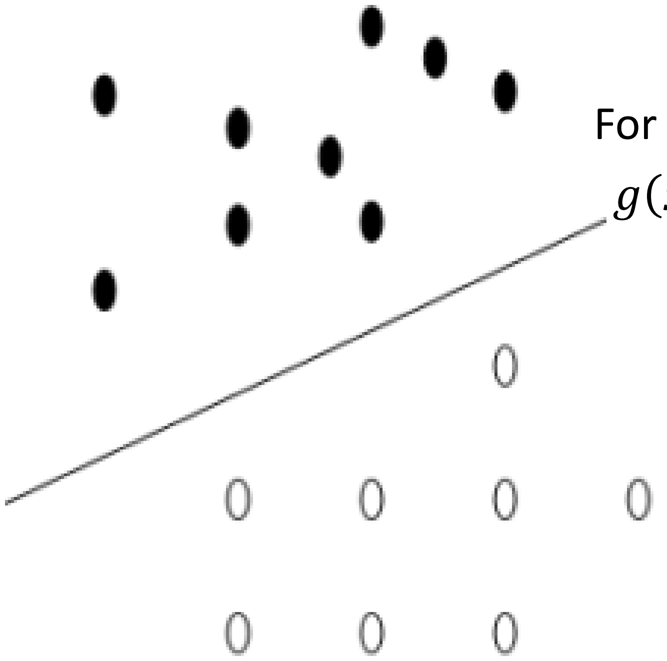
Linear Support vector machine



Linear Support vector machine



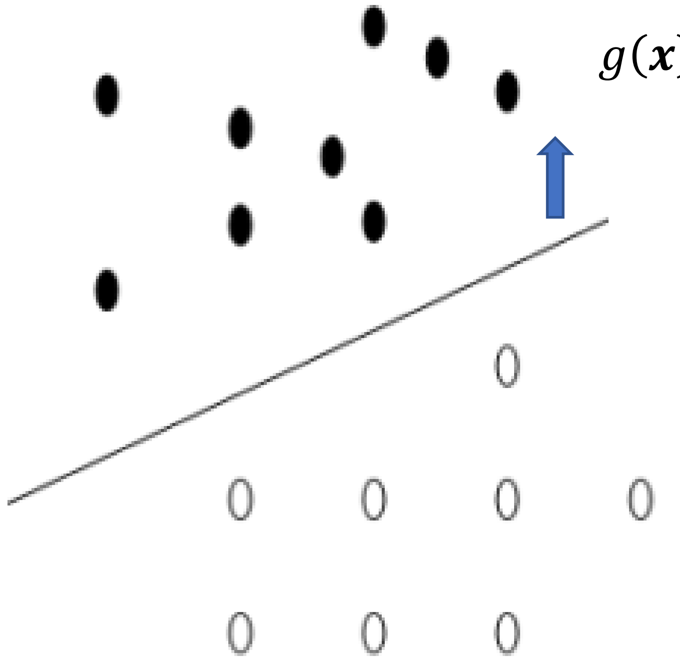
Linear Support vector machine



For all the points on the hyper plane, g(x) = 0

$$g(\boldsymbol{x}) = \boldsymbol{\omega}^T \boldsymbol{x} + \omega_0 = 0$$

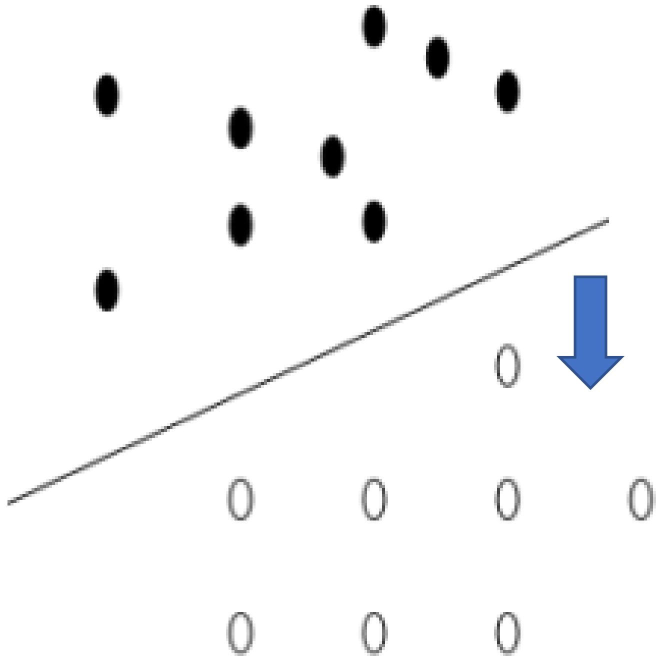
Linear Support vector machine



For all the points above the hyper plane, g(x) > 0

$$g(\boldsymbol{x}) = \boldsymbol{\omega}^T \boldsymbol{x} + \omega_0 > 0$$

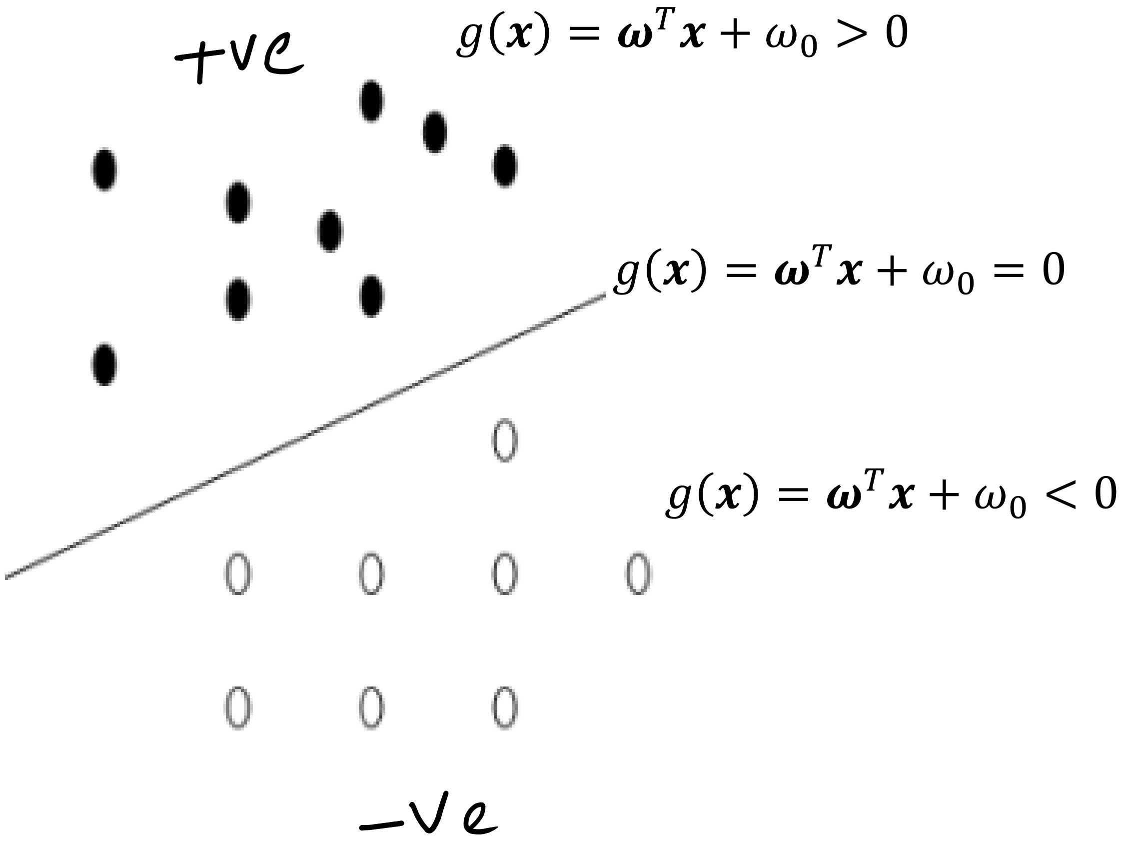
Linear Support vector machine



$g(\boldsymbol{x}) = \boldsymbol{\omega}^T \boldsymbol{x} + \omega_0 < 0$

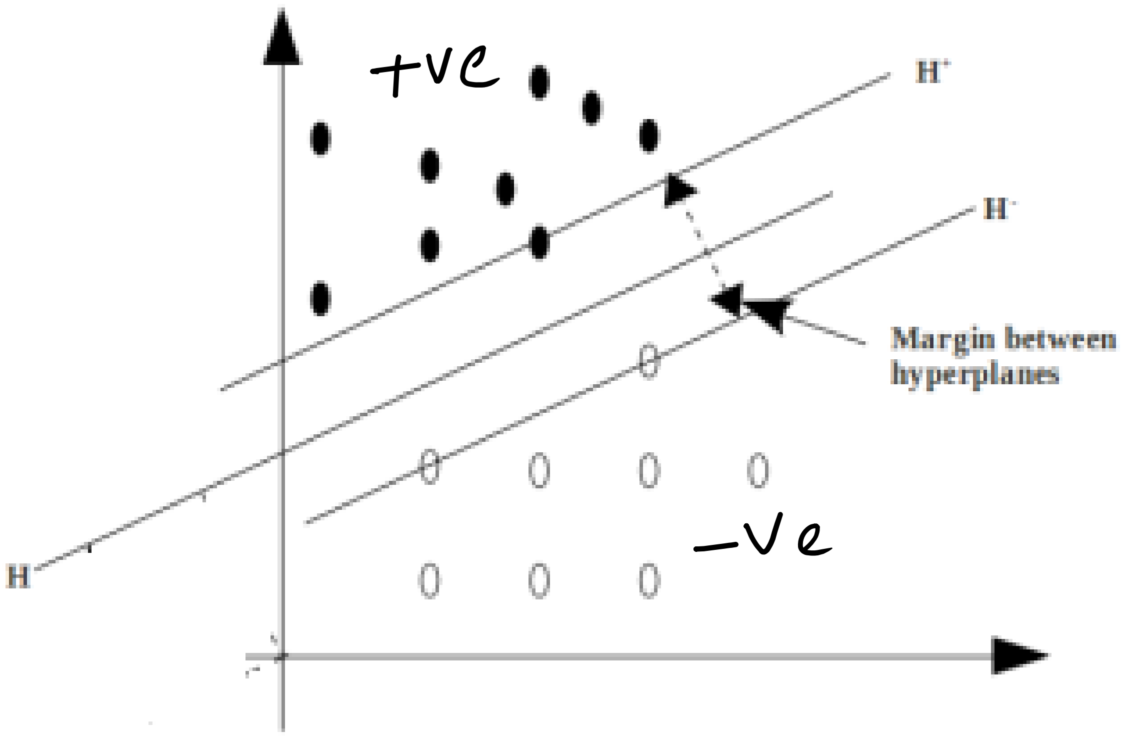
For all the points below the hyper plane, g(x) < 0

Linear Support vector machine



We choose the hyperplane that has the **maximum separating margin**

What is Separating Margin?



H is the separating hyperplane

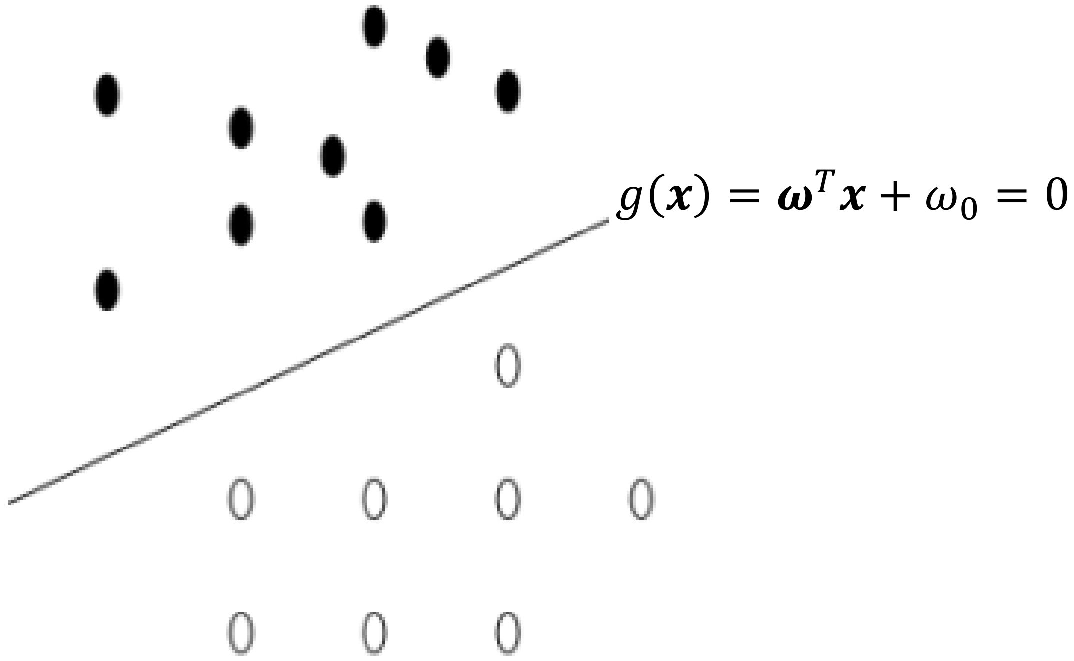
H^+ is the plane parallel to H and passing through the nearest +ve points to H

H^- is the plane parallel to H and passing through the nearest -ve points to H

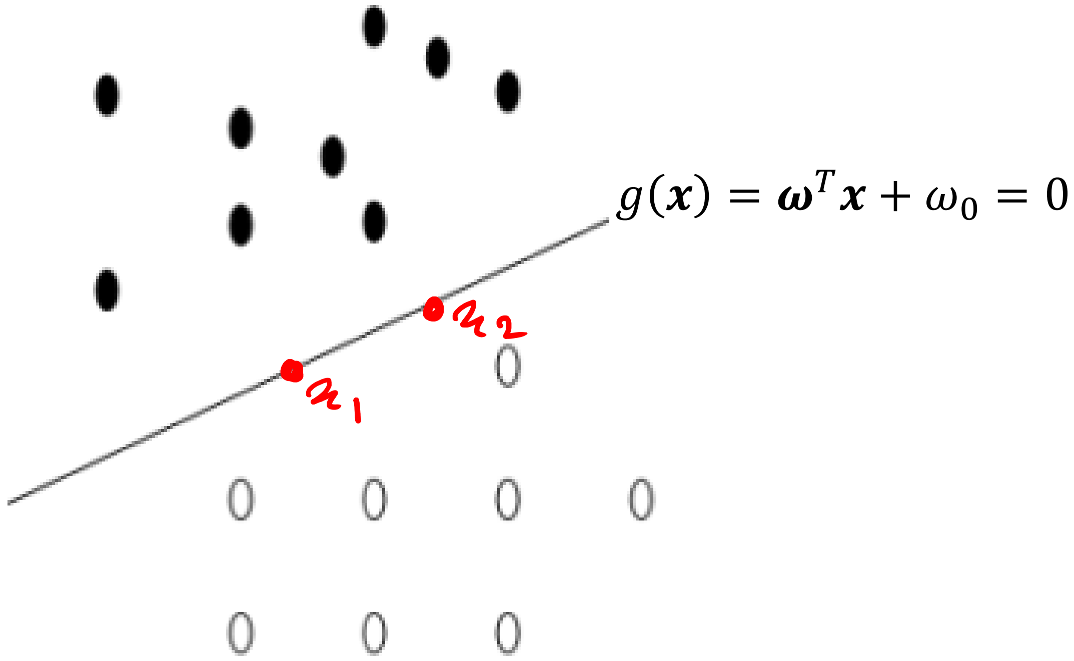
Separating margin is the distance between H^+ and H^-

We choose H that has the maximum margin.

Finding Separating Hyperplane with maximum Margin?



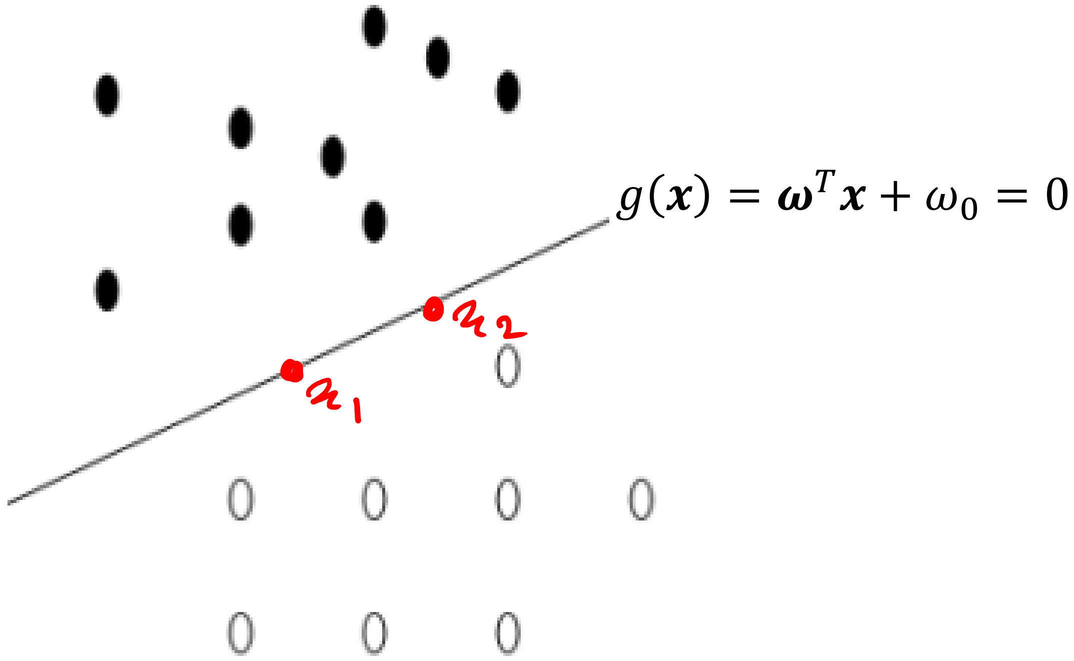
Finding Separating Hyperplane with maximum Margin?



Let us take two points ***x₁*** and ***x₂*** lying on ***g(x)***

$g(x_1) = g(x_2)$=0

Finding Separating Hyperplane with maximum Margin?



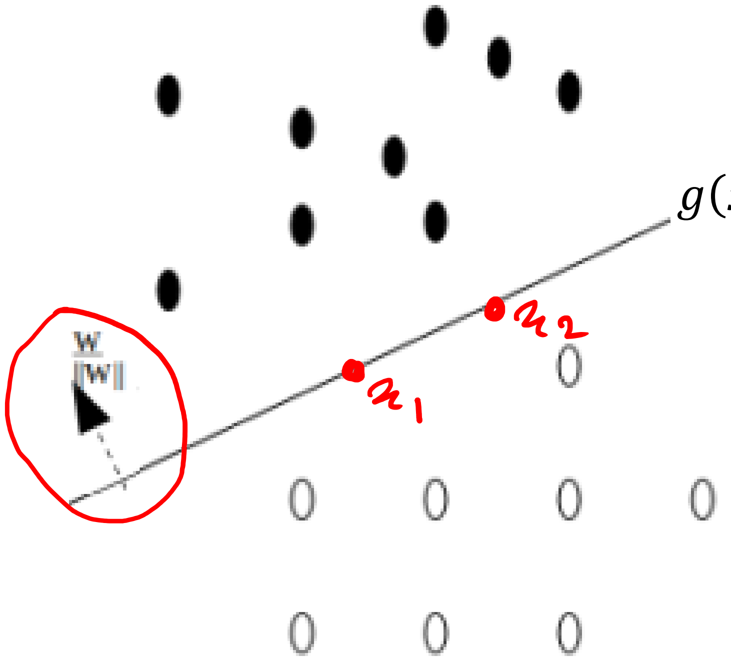
Let us take two points ***x₁*** and ***x₂*** lying on ***g(x)***

$g(x_1) = g(x_2)$=0

$\Rightarrow \boldsymbol{\omega}^T\boldsymbol{x}_1 + \omega_0 = \boldsymbol{\omega}^T\boldsymbol{x}_1 + \omega_0$

$\Rightarrow \boldsymbol{\omega}^T(\boldsymbol{x}_1 - \boldsymbol{x}_1)$ =0

Finding Separating Hyperplane with maximum Margin?



$$g(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_0 = 0$$

Let us take two points ***x_1*** and ***x_2*** lying on ***g(x)***

$$g(x_1) = g(x_2) = 0$$

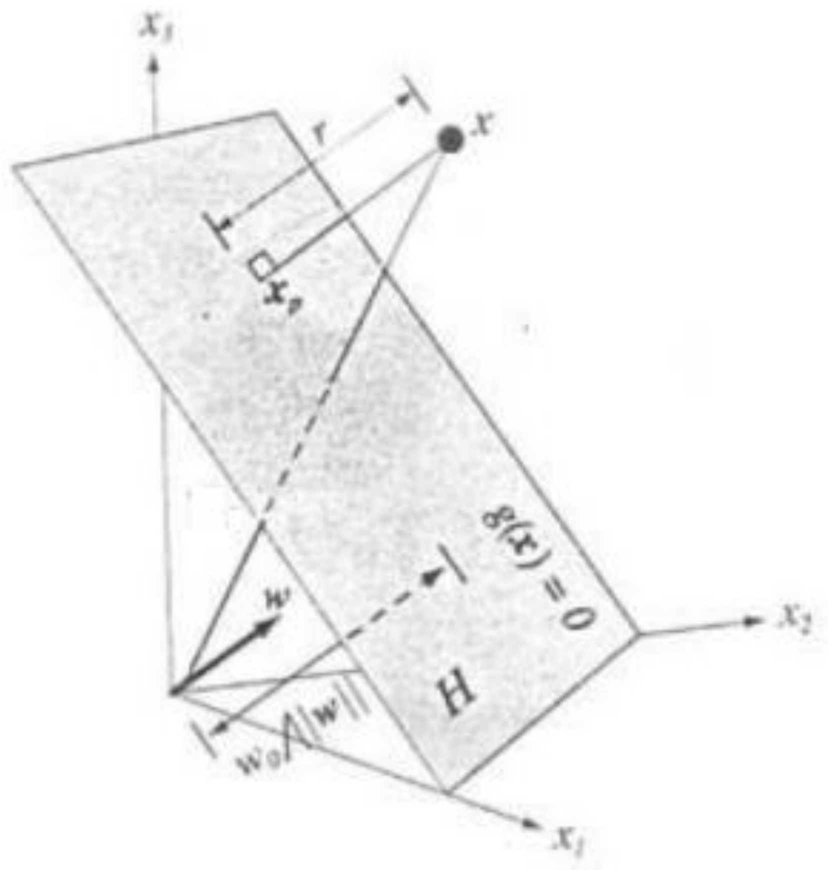
$$\Rightarrow \boldsymbol{\omega}^T\boldsymbol{x}_1 + \omega_0 = \boldsymbol{\omega}^T\boldsymbol{x}_1 + \omega_0$$

$$\Rightarrow \boldsymbol{\omega}^T(\boldsymbol{x}_1 - \boldsymbol{x}_1) = 0$$

It means $\boldsymbol{\omega}$ is orthogonal (90°/perpendicular) the vector (***x_1*** - ***x_2***).

It means $\boldsymbol{\omega}$ is orthogonal (90°/perpendicular) to ***g(x).***

Finding Separating Hyperplane with maximum Margin?



Let x be a point in space.

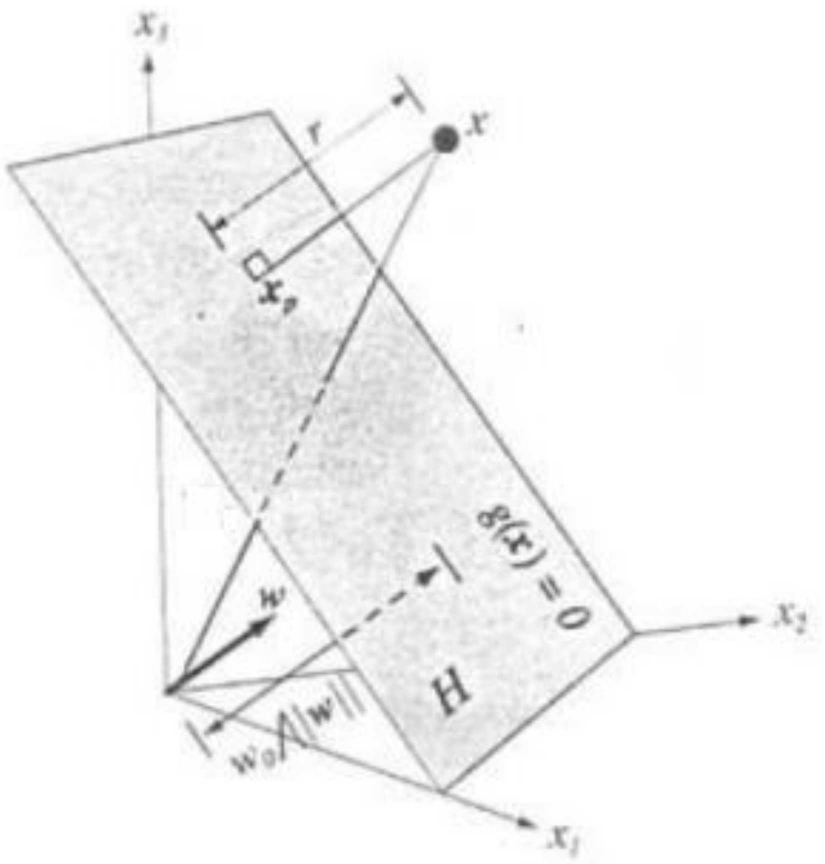
Let r be the distance of the point x from the hyperplane g(x), and x_p be the corresponding projection point of x on g(x).

Now, vector $\boldsymbol{x}$ can be defined by the sum of the vector $\boldsymbol{x_p}$ and vector $\boldsymbol{r}$.

$$\boldsymbol{x} = \boldsymbol{x_p} + \boldsymbol{r}$$

$$\Rightarrow \boldsymbol{x} = \boldsymbol{x_p} + r\frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}$$

Finding Separating Hyperplane with maximum Margin?



Let x be a point in space.

Let r be the distance of the point x from the hyperplane g(x),
and x_p be the corresponding projection point of x on g(x).

Now, vector $\boldsymbol{x}$ can be defined by the sum of the vector $\boldsymbol{x_p}$ and vector $\boldsymbol{r}$.

$$\boldsymbol{x} = \boldsymbol{x_p} + \boldsymbol{r}$$

$$\Rightarrow \boldsymbol{x} = \boldsymbol{x_p} + r\frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}$$

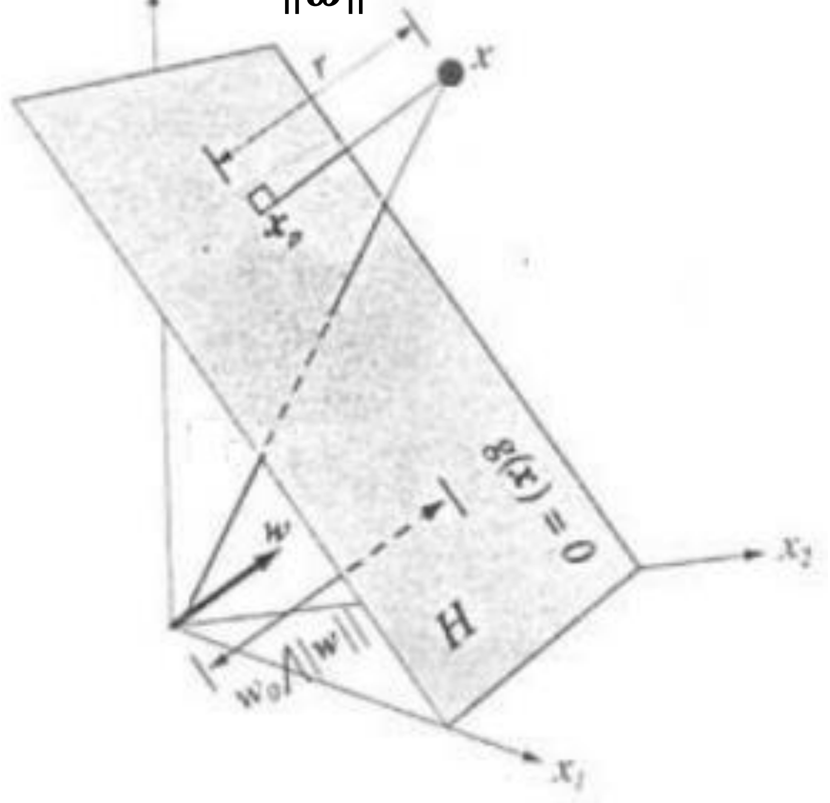
If you substitute $\boldsymbol{x}$ in g($\boldsymbol{x}$).

$$\Rightarrow \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{x_p} + r\frac{\boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|} + \boldsymbol{\omega_0}$$

$$\Rightarrow \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{x_p} + \boldsymbol{\omega_0} + r\frac{\boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|} \Rightarrow \boldsymbol{g}(\boldsymbol{x}) = r\frac{\boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|} \Rightarrow r = \frac{g(\boldsymbol{x})}{\|\boldsymbol{\omega}\|}$$

Finding Separating Hyperplane with maximum Margin?

$r = \frac{g(\boldsymbol{x})}{\|\boldsymbol{\omega}\|}$



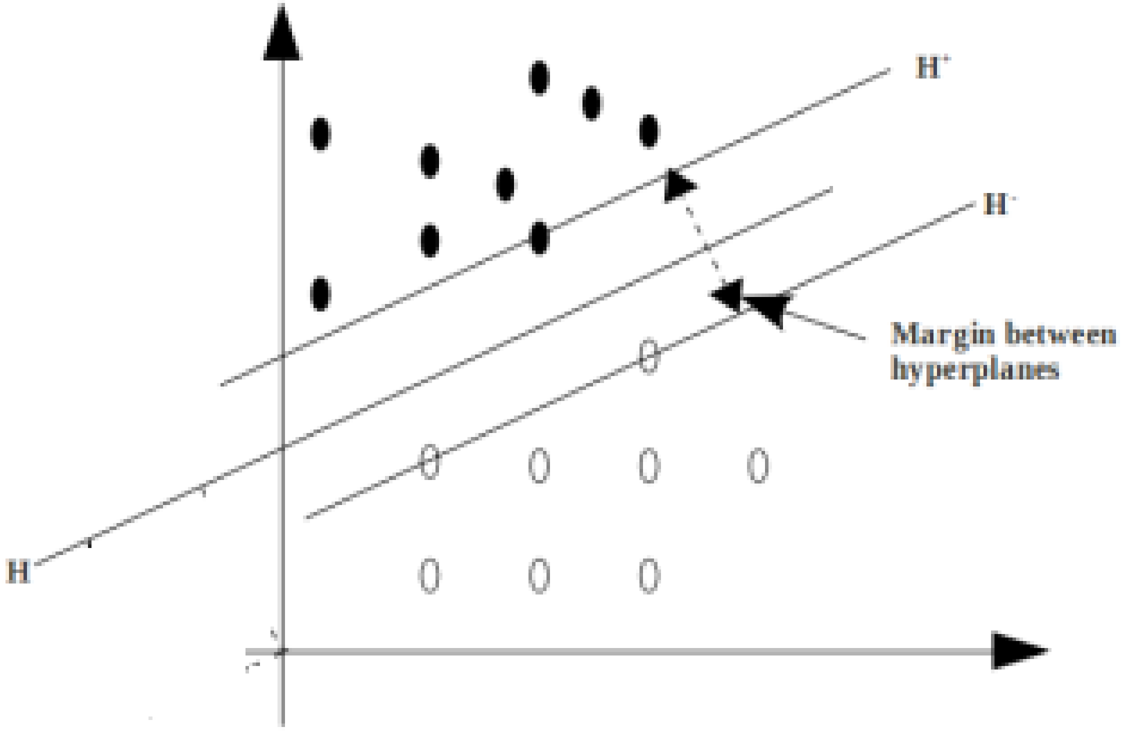
Distance of $\boldsymbol{x_p}$ from g(x)=0 is $r = \frac{g(\boldsymbol{x})}{\|\boldsymbol{\omega}\|}$

So, the distance of origin from g(x)=0 is

$$\Rightarrow r_{\mathbf{0}} = \frac{\boldsymbol{\omega}^T \mathbf{0} + \omega_0}{\|\boldsymbol{\omega}\|}$$

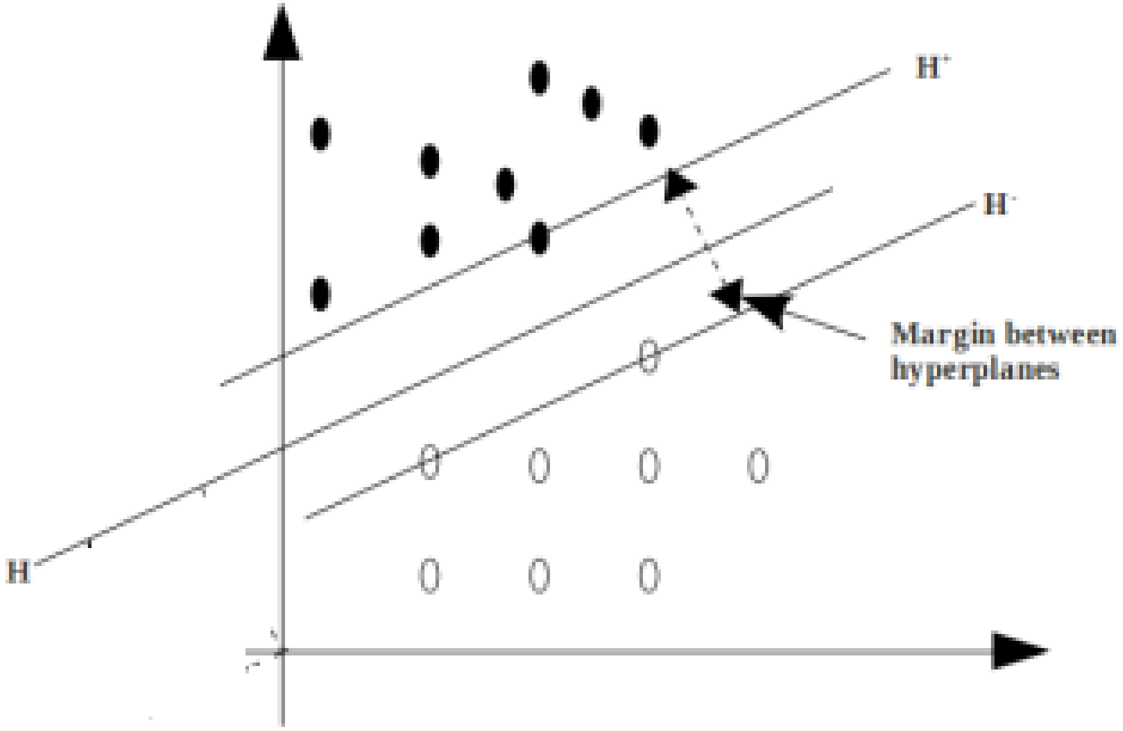
$$\Rightarrow r_{\mathbf{0}} = \frac{\omega_0}{\|\boldsymbol{\omega}\|}$$

Finding Separating Hyperplane with maximum Margin?



What is the Margin between H+ and H-?

Finding Separating Hyperplane with maximum Margin?

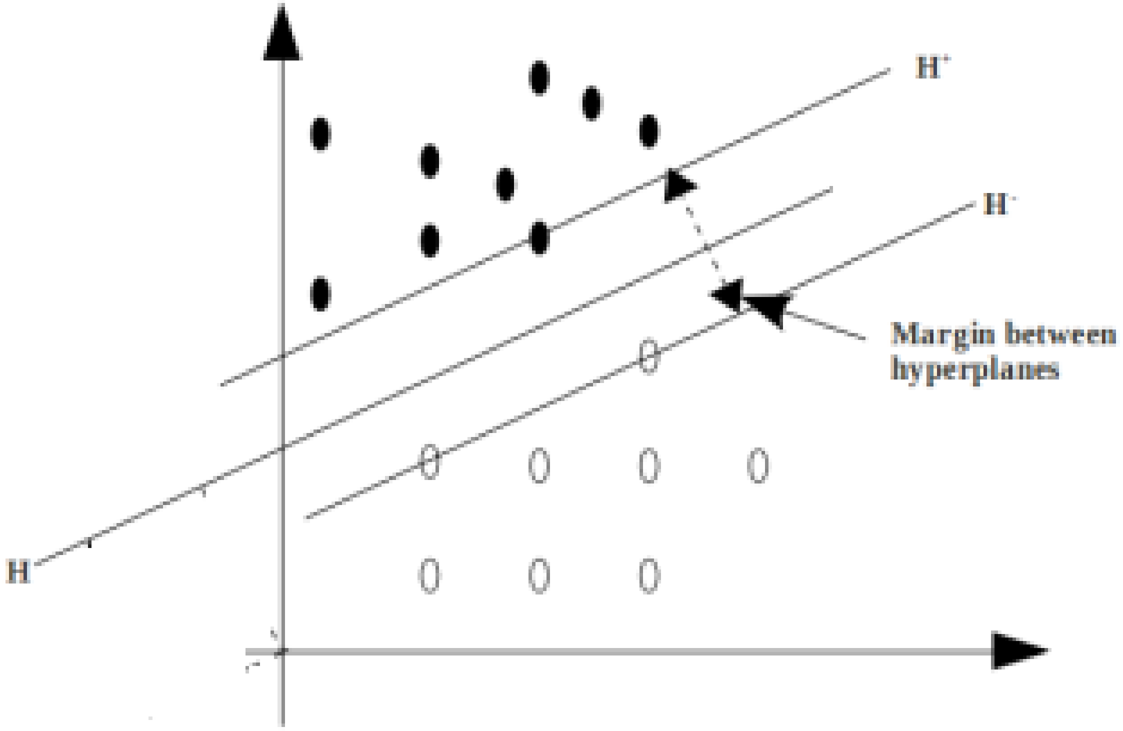


To define the expression for H+ and H-, let us make the following assumptions.

Given a datapoint < $\boldsymbol{x_i}, y_i$> where $y_i \epsilon \{+ve, -ve\}$ is the class label of $\boldsymbol{x_i}$,

- let us replace **–ve by +1**, and **–ve by -1**.

Finding Separating Hyperplane with maximum Margin?



To define the expression for H+ and H-, let us make the following assumptions.

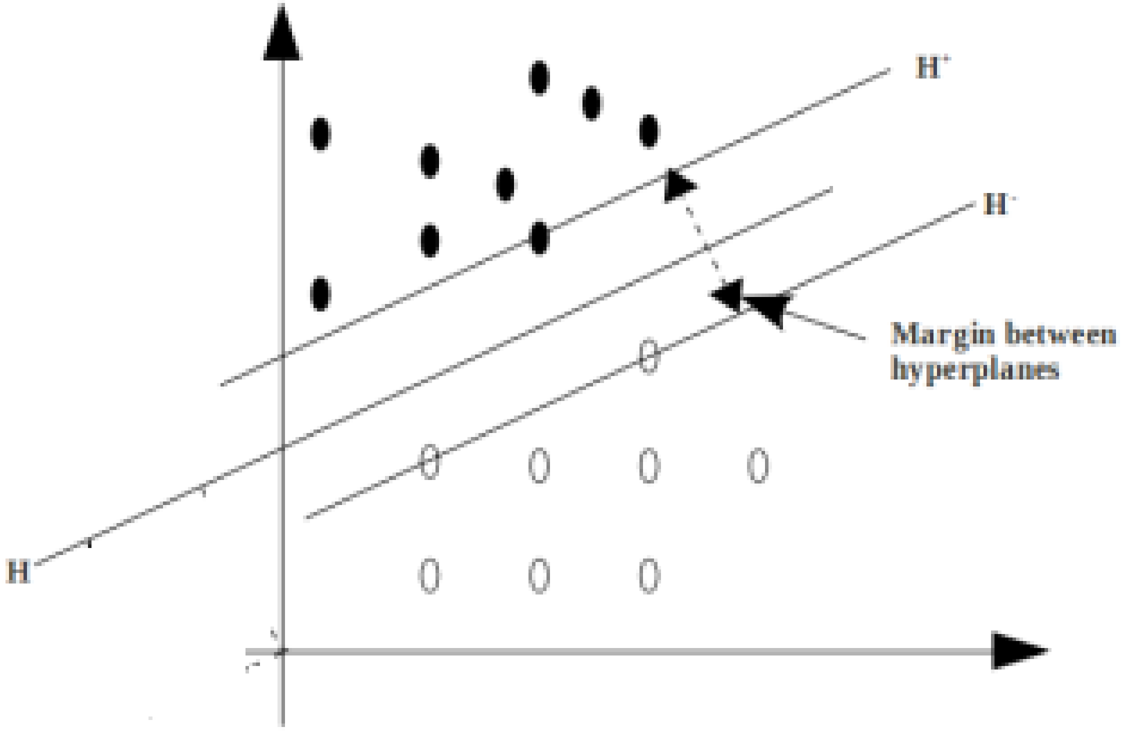
Given a datapoint < $\boldsymbol{x_i}, y_i$> where $y_i \epsilon \{+ve, -ve\}$ is the class label of $\boldsymbol{x_i}$,

- Let us replace **–ve by +1**, and **–ve by -1**.
- Now, each data point will satisfy the following

$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_{\boldsymbol{0}} \geq 1, \forall y_{\boldsymbol{i}} = +1$$

$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_{\boldsymbol{0}} \leq -1, \forall y_{\boldsymbol{i}} = -1$$

Finding Separating Hyperplane with maximum Margin?



To define the expression for H+ and H-, let us make the following assumptions.

Given a datapoint $< \boldsymbol{x_i}, y_i >$ where $y_i \epsilon \{+ve, -ve\}$ is the class label of $\boldsymbol{x_i}$,

- let us replace **–ve by +1**, and **–ve by -1**.
- Now, each data point will satisfy the following

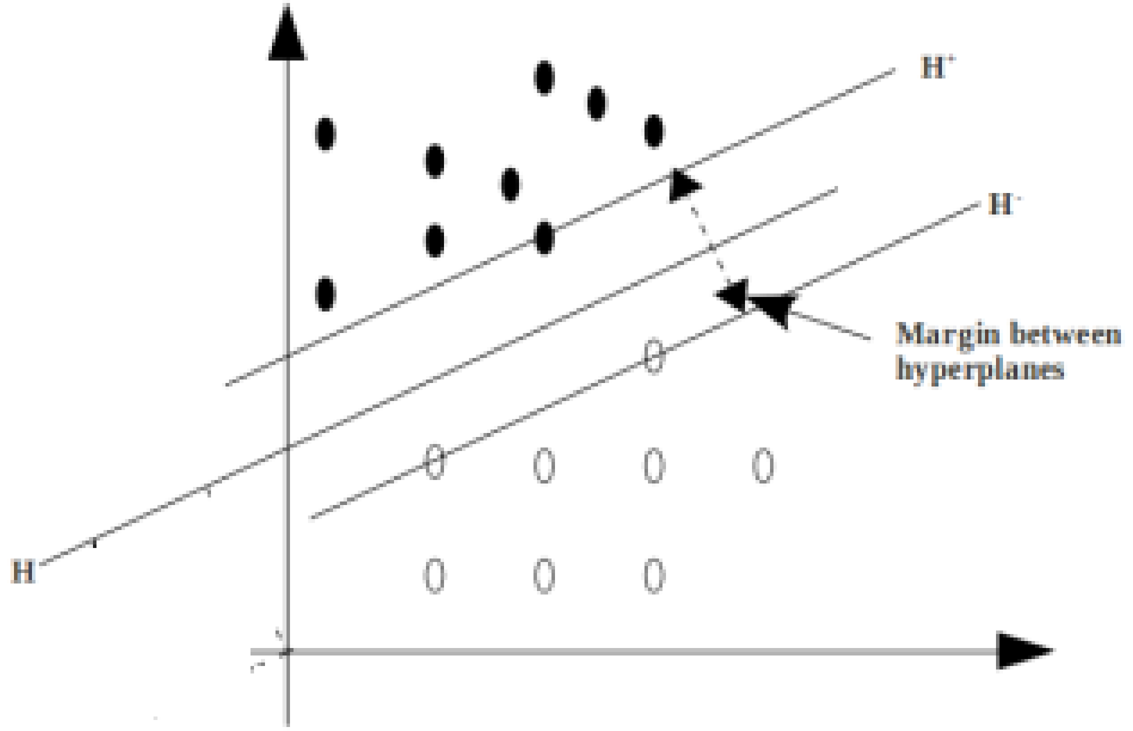
$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_{\boldsymbol{0}} \geq 1, \forall y_{\boldsymbol{i}} = +1$$

$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_{\boldsymbol{0}} \leq -1, \forall y_{\boldsymbol{i}} = -1$$

The above two expression can be merged to form a single expression

$$y_i(\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_0) \geq 1, \forall \boldsymbol{x_i}$$

Finding Separating Hyperplane with maximum Margin?



To define the expression for H+ and H-, let us make the following assumptions.

Given a datapoint < $\boldsymbol{x_i}, y_i$> where $y_i \epsilon \{+ve, -ve\}$ is the class label of $\boldsymbol{x_i}$,

- let us replace **–ve by +1**, and **–ve by -1**.
- Now, each data point will satisfy the following

$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_0 \geq 1, \forall y_i = +1$$

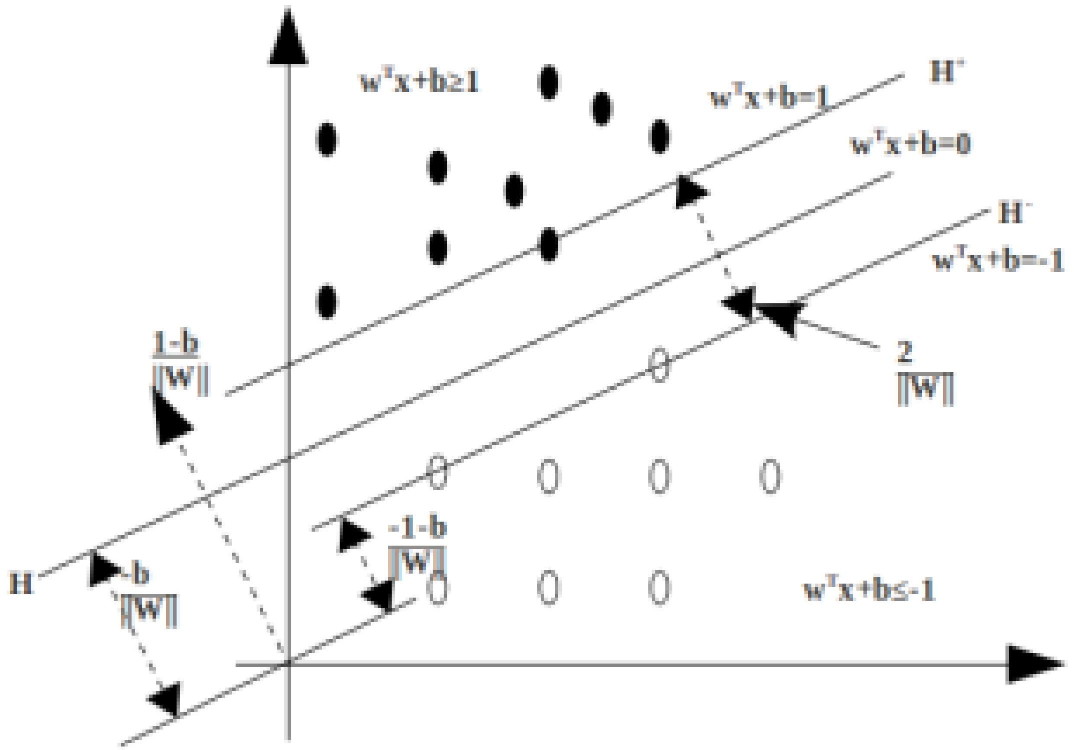
$$\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_0 \leq -1, \forall y_i = -1$$

The above two expression can be merged to form a single expression

$$y_i(\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_0) \geq 1, \forall \boldsymbol{x_i}$$

$$y_i(\boldsymbol{\omega}^T \boldsymbol{x_i} + \omega_0) - 1 \geq 0, \forall \boldsymbol{x_i}$$

Finding Separating Hyperplane with maximum Margin?

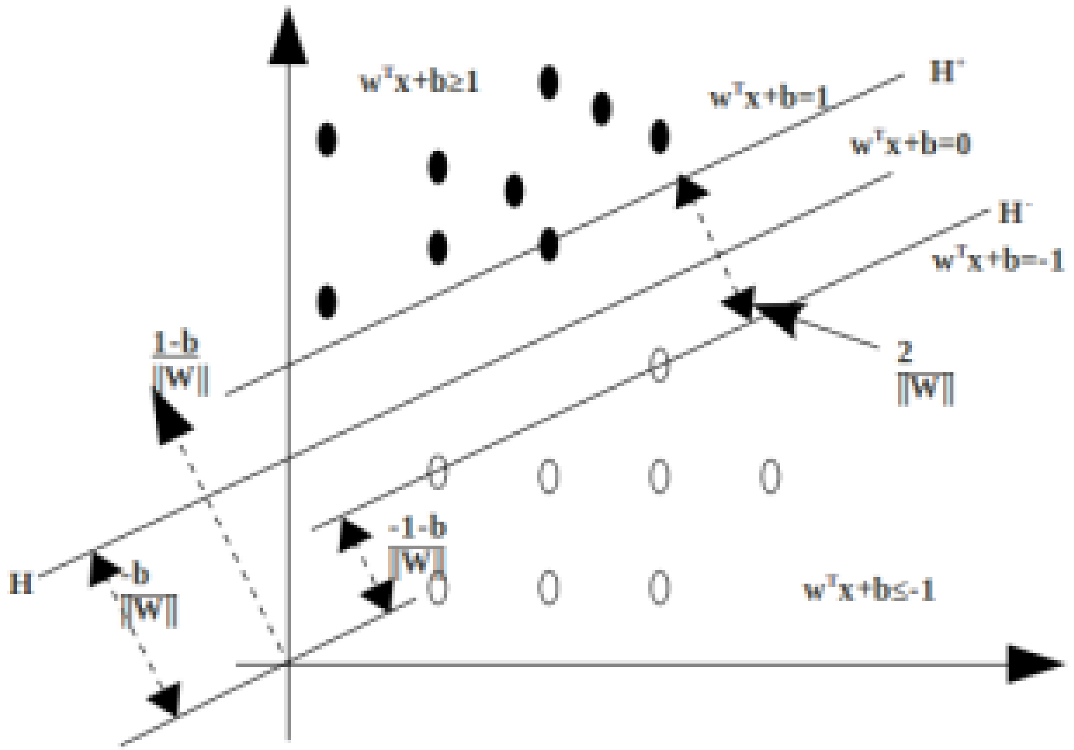


H: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 0$

H+: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 1$

H-: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = -1$

Finding Separating Hyperplane with maximum Margin?



H: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 0$

H+: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 1$

H-: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = -1$

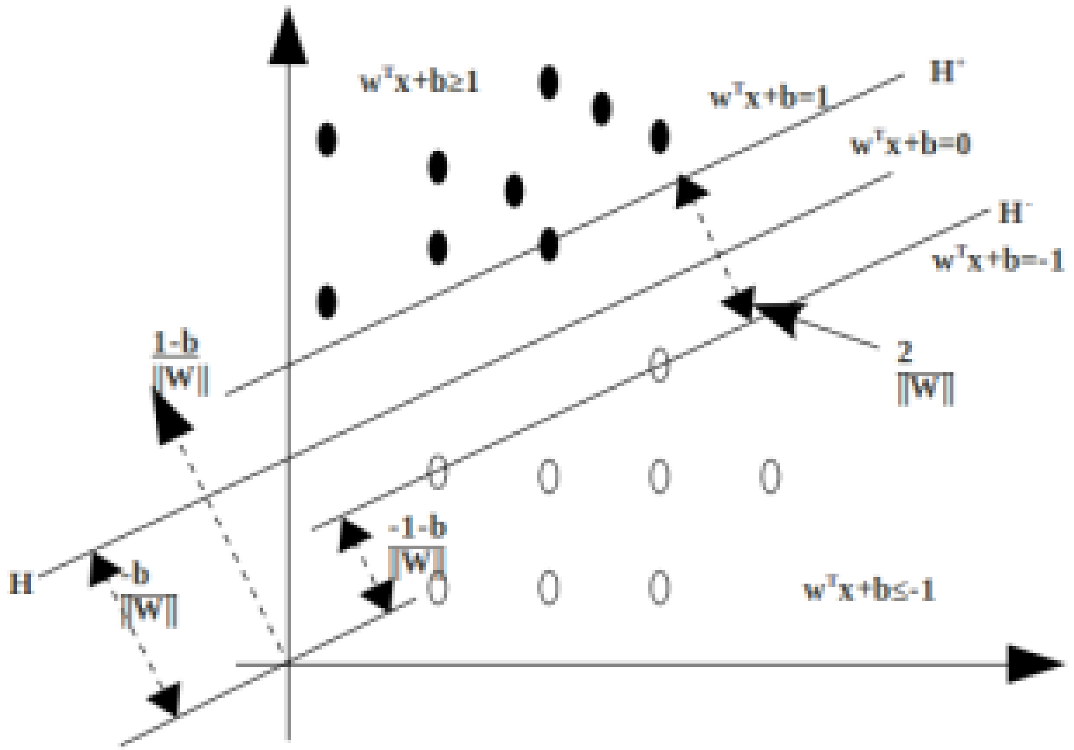
$$\text{Margin} = r_{\boldsymbol{H}^+} - r_{\boldsymbol{H}^-}$$

$$= \frac{\omega_0 - 1}{\|\boldsymbol{\omega}\|} - \frac{\omega_0 + 1}{\|\boldsymbol{\omega}\|}$$

$$= \frac{\omega_0 - 1 - \omega_0 - 1}{\|\boldsymbol{\omega}\|}$$

$$\Rightarrow Margin = \frac{-2}{\|\boldsymbol{\omega}\|}$$

Finding Separating Hyperplane with maximum Margin?



H: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 0$

H+: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = 1$

H-: $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \omega_{\mathbf{0}} = -1$

$$\text{Margin} = r_{\boldsymbol{H}^+} - r_{\boldsymbol{H}^-}$$

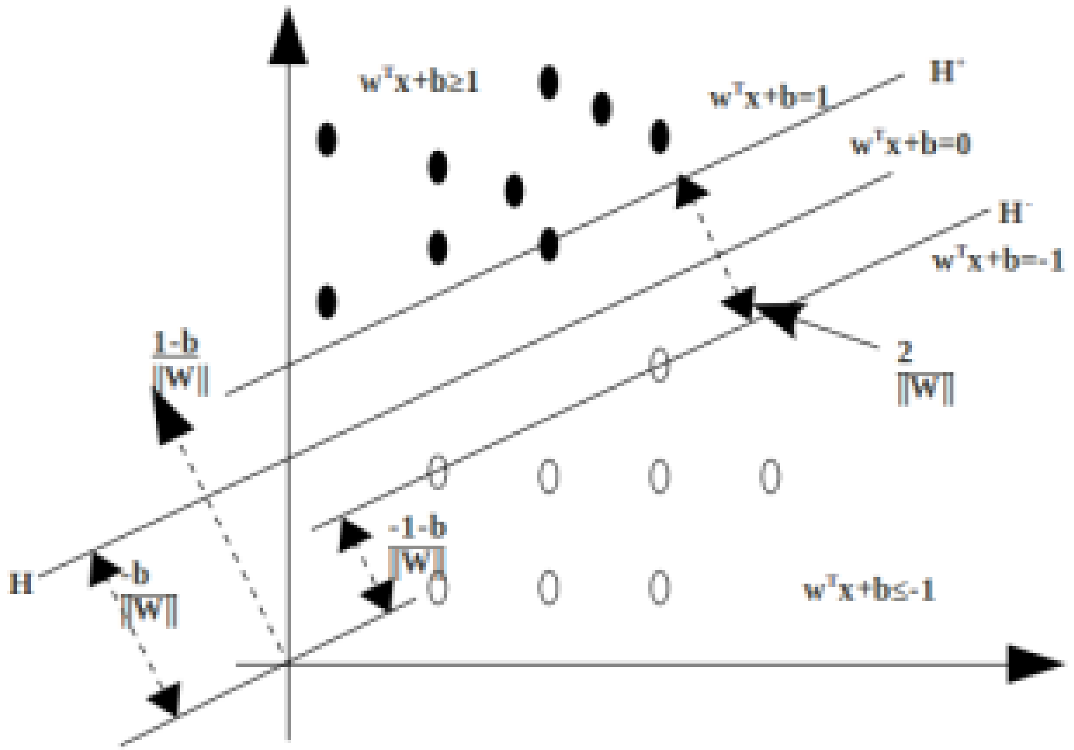
$$= \frac{\omega_0 - 1}{\|\boldsymbol{\omega}\|} - \frac{\omega_0 + 1}{\|\boldsymbol{\omega}\|}$$

$$= \frac{\omega_0 - 1 - \omega_0 - 1}{\|\boldsymbol{\omega}\|}$$

$$\Rightarrow Margin = \frac{-2}{\|\boldsymbol{\omega}\|}$$

As we are interested in the absolute value, we consider the margin as $\frac{2}{\|\boldsymbol{\omega}\|}$

Finding Separating Hyperplane with maximum Margin?

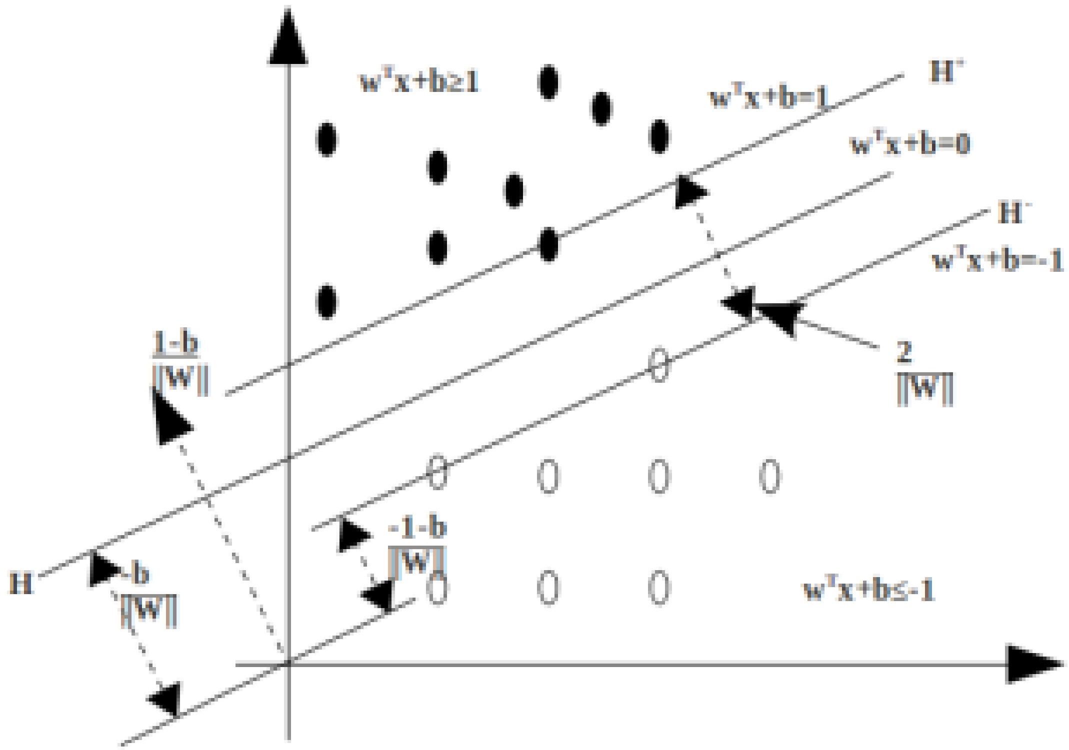


Task is to find the hyperplane (g(x)=0) that maximizes the margin $\frac{2}{\|\boldsymbol{\omega}\|}$

$$Margin = \frac{2}{\|\boldsymbol{\omega}\|}$$

$$\Rightarrow Margin = \frac{1}{\frac{\|\boldsymbol{\omega}\|}{2}}$$

Finding Separating Hyperplane with maximum Margin?



Task is to find the hyperplane (g(x)=0) that maximizes the margin $\frac{2}{\|\omega\|}$

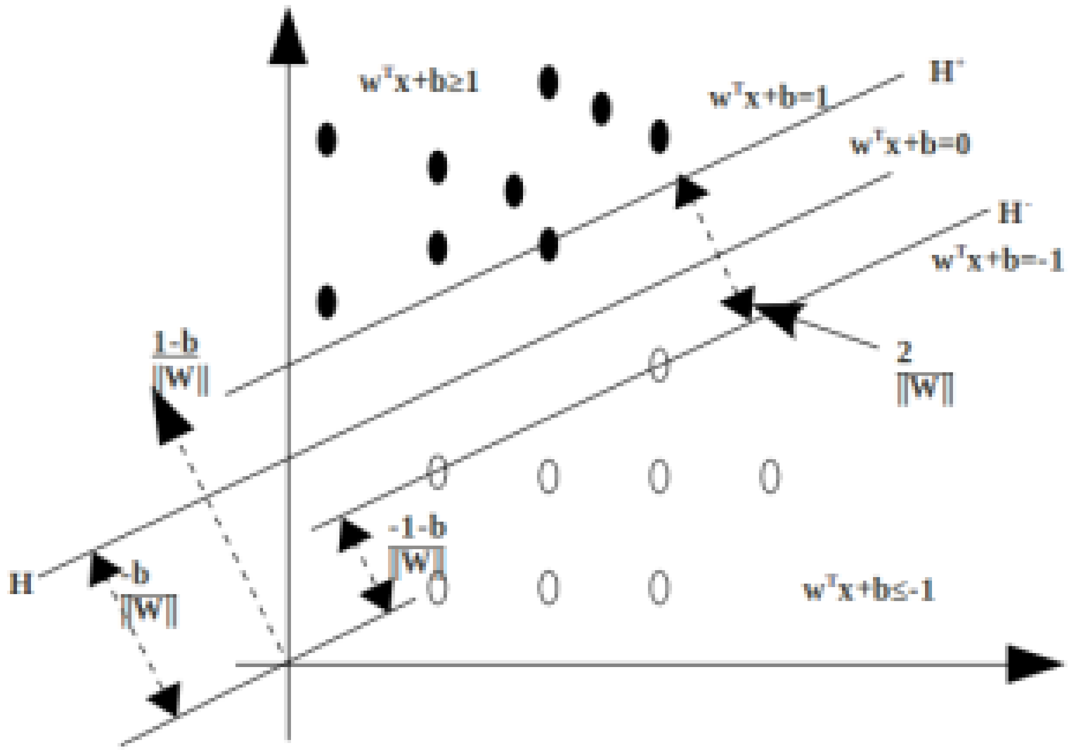
$$Margin = \frac{2}{\|\boldsymbol{\omega}\|}$$

$$\Rightarrow Margin = \frac{1}{\frac{\|\boldsymbol{\omega}\|}{2}}$$

Maximizing $\frac{2}{\|\omega\|}$ is equivalent to minimizing $\frac{\|\omega\|}{2}$

Minimizing $\frac{\|\omega\|}{2}$ is equivalent to **minimizing** $\frac{\omega^T\omega}{2}$

Finding Separating Hyperplane with maximum Margin?

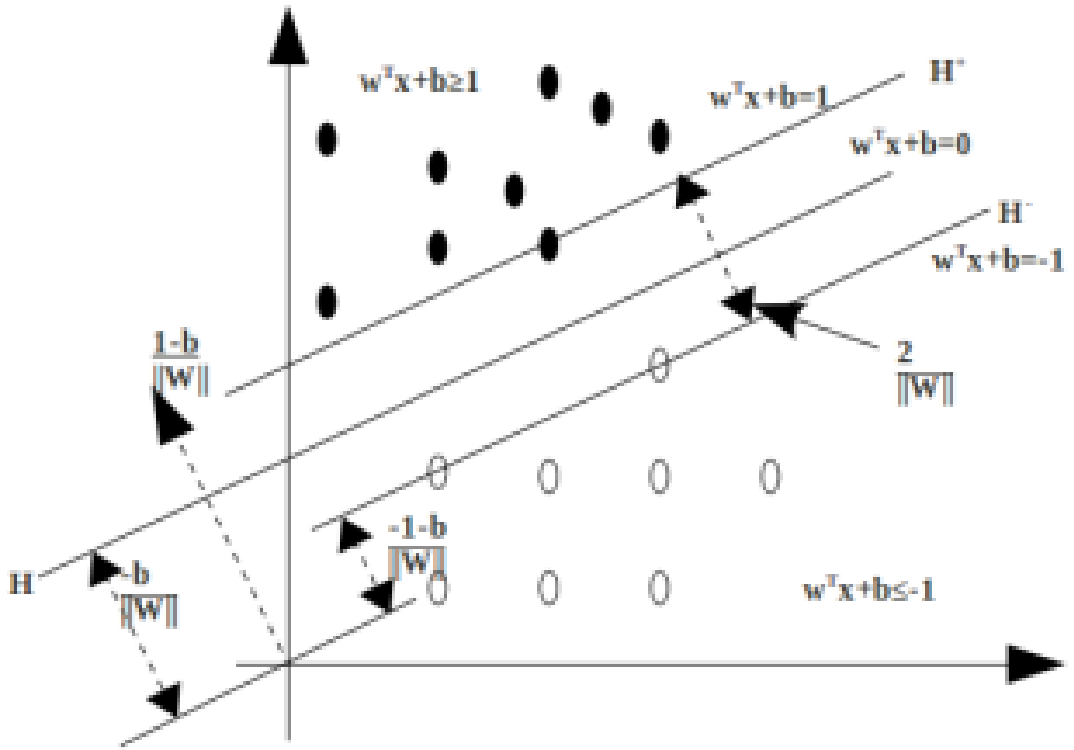


Now, we need to solve the following optimization

Minimize objective function $\frac{\boldsymbol{\omega}^T\boldsymbol{\omega}}{\mathbf{2}}$

Subject to the constraint $y_i(\boldsymbol{\omega}^T\boldsymbol{x}_i + \omega_{\mathbf{0}}) \geq 1, \forall \boldsymbol{x}_i$

Finding Separating Hyperplane with maximum Margin?



Objective function is to **minimize** $\frac{\boldsymbol{\omega}^T \boldsymbol{\omega}}{2}$

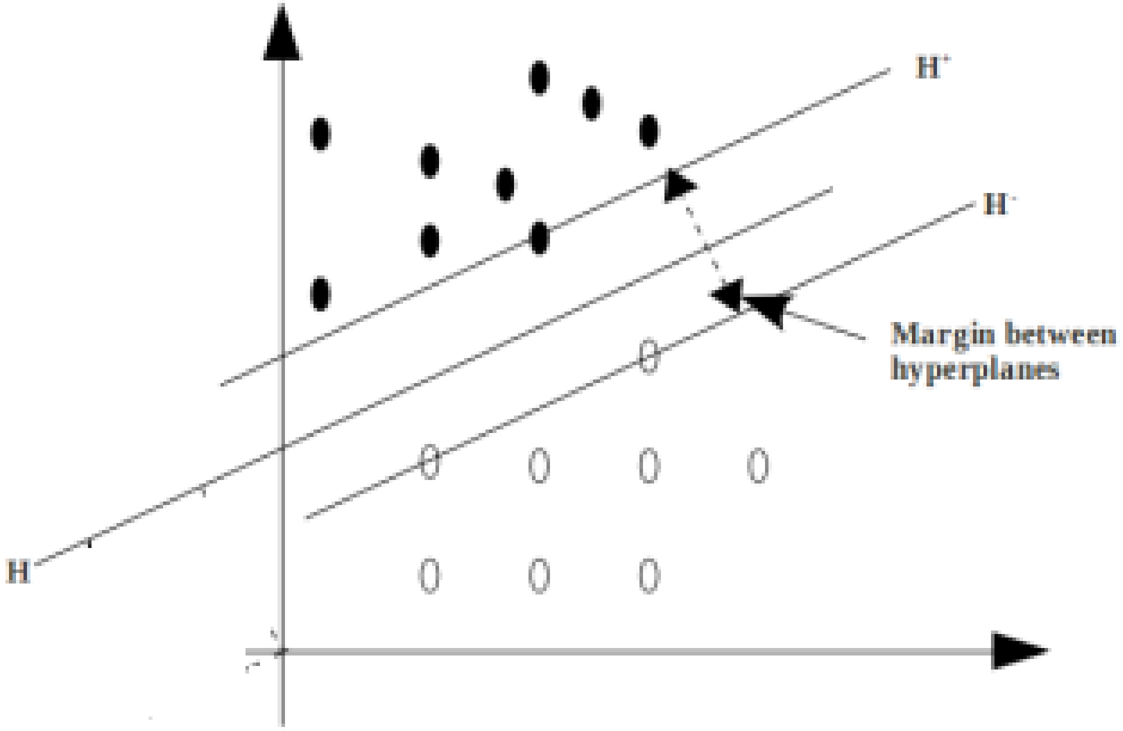
Subject to $y_i(\boldsymbol{\omega}^T \boldsymbol{x}_i + \omega_0) \geq 1, \forall \boldsymbol{x}_i$

To find the parameters $\boldsymbol{\omega}$ and where ω_0, solve the following optimization function

$$L_p = \frac{\boldsymbol{\omega}^T \boldsymbol{\omega}}{2} - \sum_{i=1}^{n} \lambda_i (y_i(\boldsymbol{\omega}^T \boldsymbol{x}_i + \omega_0) - 1)$$

where λ_i are Lagrange multipliers

What are the support vectors?



$$L_p = \frac{\boldsymbol{\omega}^T \boldsymbol{\omega}}{2} - \sum_{i=1}^{n} \lambda_i (y_i (\boldsymbol{\omega}^T \boldsymbol{x}_i + \omega_0) - 1)$$

[
In order to find the parameters, we need to solve this objective function.
]

Summary

- What is separating hyperplane?
- How to define separating hyperplane?
- What are Support Vector Machine?
- How to classify a new example using SVM