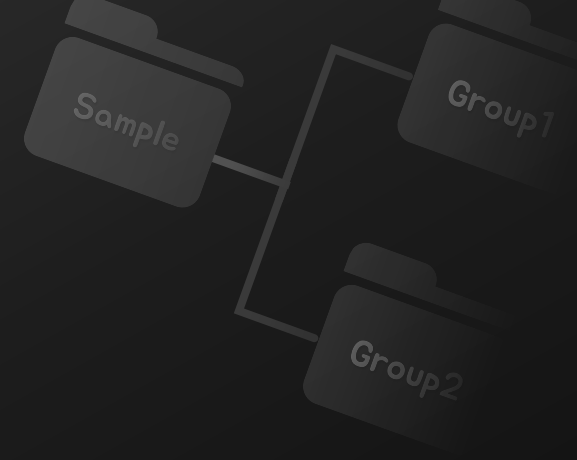
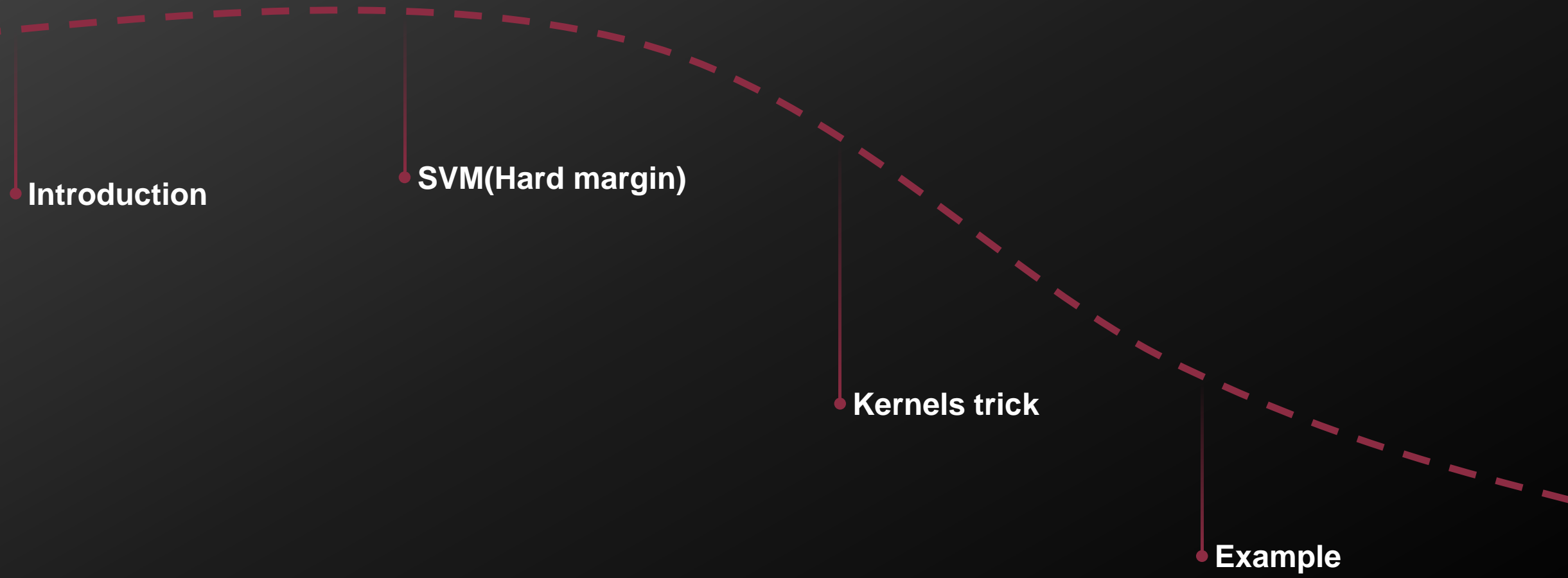


Support Vector Machine

Han Sol Kang



Contents



Introduction



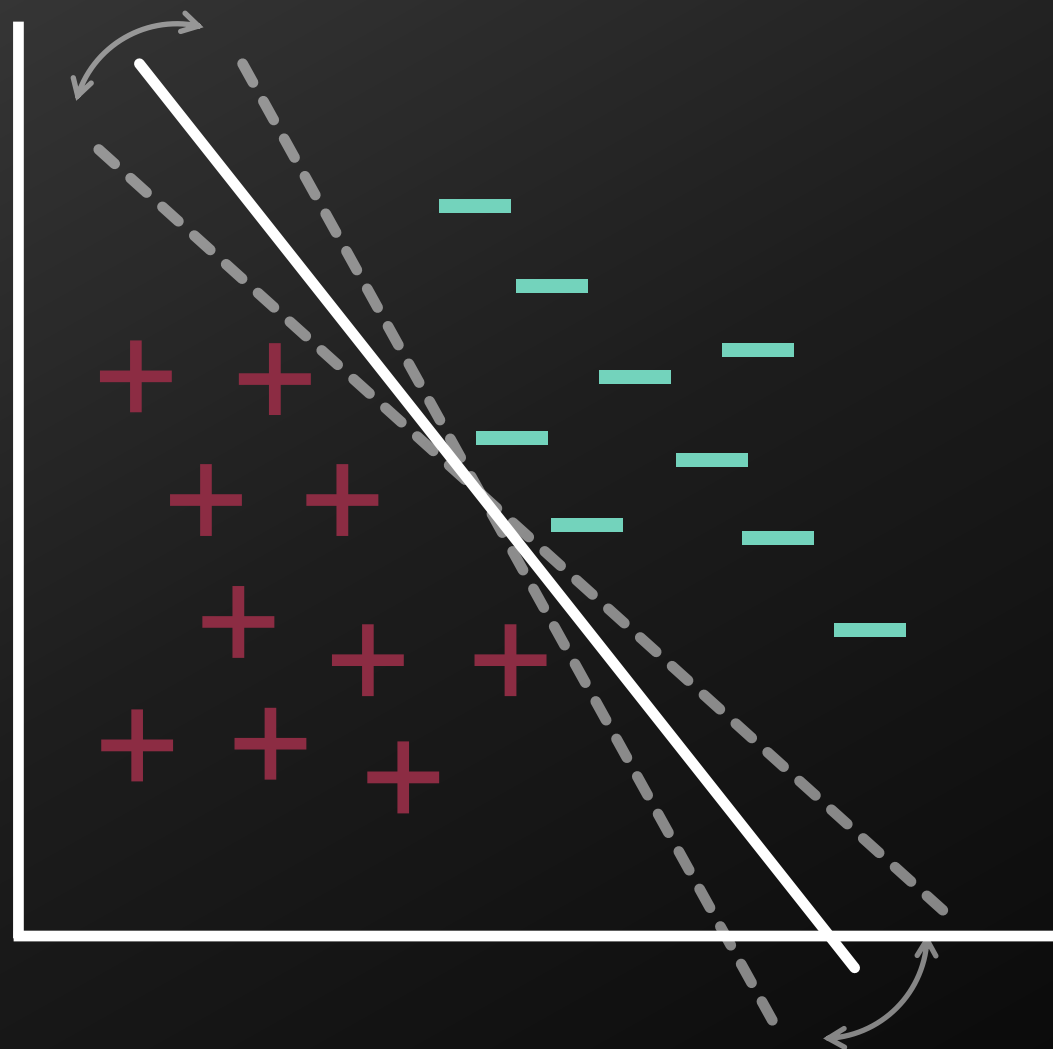
[Vladimir N. Vapnik]

Vladimir Naumovich Vapnik is one of the main developers of the Vapnik–Chervonenkis theory of statistical learning, and the **co-inventor of the support vector machine method**.

Vladimir Vapnik was born in the Soviet Union.(Dec 6, 1936)

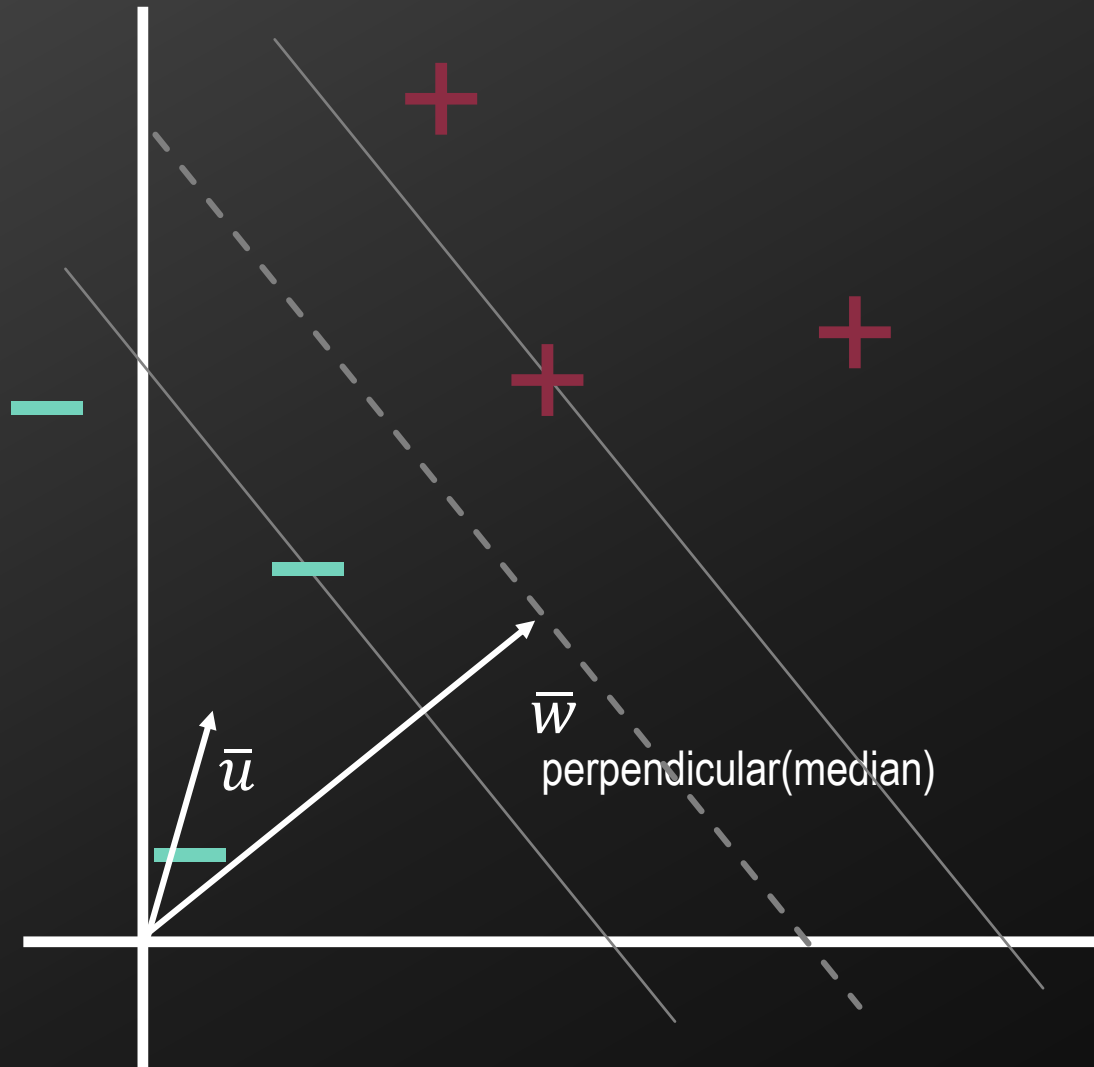
He received Ph.D in statistics at the Institute of Control Sciences, Moscow in 1964.

Introduction



SVM (Hard margin)

❖ Decision rule & Constraint.



$$\bar{w} \cdot \bar{u} \geq C$$

$$\bar{w} \cdot \bar{u} + b \geq 0 \quad \text{then } + \quad \text{Decision Rule}$$

$$\bar{w} \cdot \bar{x}_+ + b \geq 1$$

$$\bar{w} \cdot \bar{x}_- + b \leq -1$$

$$y_i = \begin{cases} +1 & \text{for } + \text{ samples} \\ -1 & \text{for } - \text{ samples} \end{cases}$$

$$y_i(\bar{w} \cdot \bar{x}_i + b) \geq 1$$

$$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 \geq 0 \quad \text{Constraint}$$

$$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 = 0 \quad \text{Gutter}$$

The diagram illustrates the geometric interpretation of the soft margin maximization problem. It shows a 2D coordinate system with a solid line representing the decision boundary and a dashed line representing the margin boundary. A unit normal vector is shown perpendicular to the margin boundary. The width of the margin is indicated by a dashed arrow. The distance from the origin to the margin boundary is labeled \bar{x}_+ .

Use the gutter condition.

- samples : $-\bar{w} \cdot \bar{x}_- - b - 1 = 0$

$$= \frac{-b+1}{\|\bar{w}\|} + \frac{b+1}{\|\bar{w}\|} = \frac{2}{\|\bar{w}\|}$$

SVM (Hard margin)

❖ Maximize(Minimize) Problem.

Width : $\frac{2}{\|\bar{w}\|}$

Maximize

Margin : $\frac{1}{\|\bar{w}\|}$

Maximize



$\|\bar{w}\|$

Minimize

$\frac{1}{2} \|\bar{w}\|^2$

Minimize



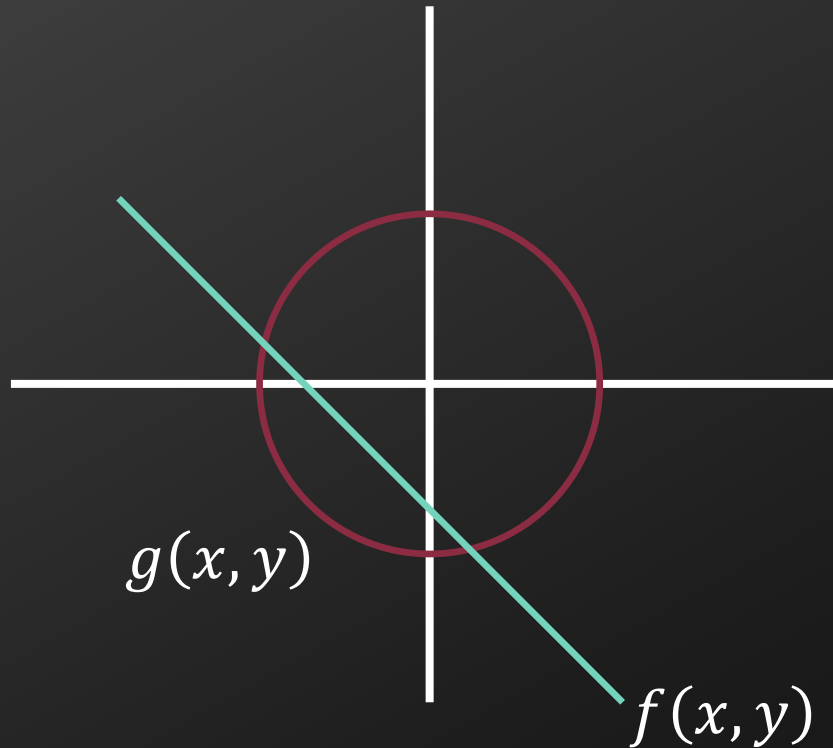
$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 \geq 0$

Constraint

Solver : Lagrange multiplier method

SVM (Hard margin)

❖ Lagrange multiplier method.



$$\nabla f(x, y) = \alpha \nabla(g(x, y))$$

$$\nabla f(x, y) - \alpha \nabla(g(x, y)) = 0$$

$$L(x, y, \alpha) = f(x, y) - \alpha(g(x, y))$$

$$L(x, y, \alpha) = f(x, y) - \sum_{i=1}^n \alpha_i(g_i(x, y))$$

$$\nabla L(x, y, \alpha) = 0$$

SVM (Hard margin)

❖ KKT(Karush–Kuhn–Tucker) Condition

$$L(\theta, \alpha) = f(\theta) - \sum_{i=1}^n \alpha_i (g_i(\theta)) \quad g_i(\theta) \geq 0$$

KKT Condition 1

$$\frac{\partial L}{\partial \theta_i} = 0, \quad i = 1, \dots, k$$

KKT Condition 2

$$\alpha_i \geq 0, \quad i = 1, \dots, k$$

KKT Condition 3

$$\alpha_i g_i(\theta) = 0, \quad i = 1, \dots, k$$

SVM (Hard margin)

❖ Finding Optimal Hyperplane

$$L = \frac{1}{2} \|\bar{w}\|^2 - \sum_{i=1}^N \alpha_i [y_i(\bar{w} \cdot \bar{x}_i + b) - 1]$$

$$= \frac{1}{2} \|\bar{w}\|^2 - \sum \alpha_i [y_i(\bar{w} \cdot \bar{x}_i + b) - 1]$$

KKT Condition 1

$$\frac{\partial L}{\partial \bar{w}} = \bar{w} - \sum \alpha_i y_i \bar{x}_i = 0$$

$$\bar{w} = \sum \alpha_i y_i \bar{x}_i$$

$$\frac{\partial L}{\partial b} = - \sum \alpha_i y_i = 0$$

$$\sum \alpha_i y_i = 0$$

KKT Condition 2

$$\alpha_i \geq 0$$

KKT Condition 3

$$\alpha_i (y_i(\bar{w} \cdot \bar{x}_i + b) - 1) = 0$$

SVM (Hard margin)

❖ Finding Optimal Hyperplane

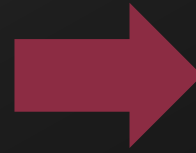
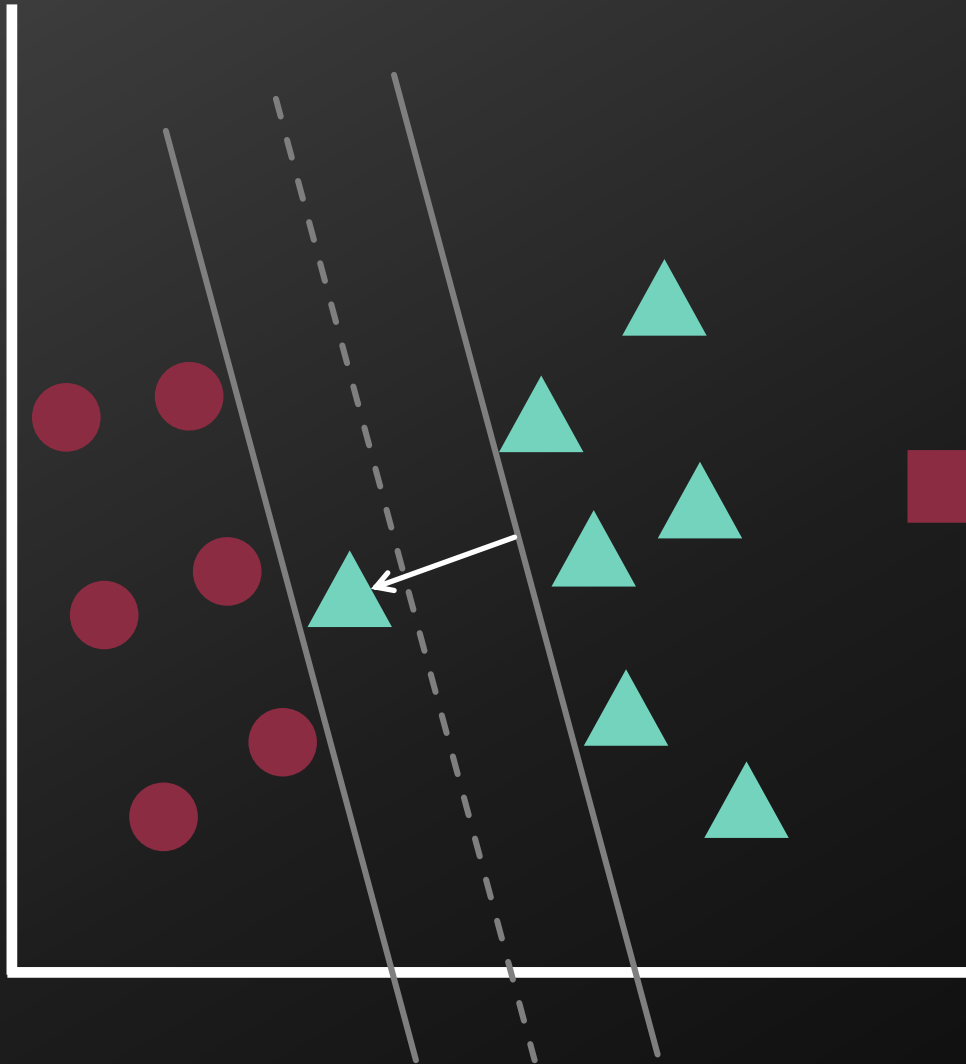
$$\begin{aligned}
 L &= \frac{1}{2} \left(\sum \alpha_i y_i \bar{x}_i \right) \left(\sum \alpha_j y_j \bar{x}_j \right) - \sum \alpha_i y_i \bar{x}_i \left(\sum \alpha_j y_j \bar{x}_j \right) - b \left(\sum \alpha_i y_i \right) + \sum \alpha_i \\
 &= \sum \alpha_i - \frac{1}{2} \left(\sum \alpha_i y_i \bar{x}_i \right) \left(\sum \alpha_j y_j \bar{x}_j \right) \\
 &= \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (\bar{x}_i \cdot \bar{x}_j) \quad \sum \alpha_i y_i = 0 \quad \alpha_i \geq 0
 \end{aligned}$$

Quadratic programming

$$\bar{w}^* = \sum \alpha_i^* y_i \bar{x}_i \quad \bar{w}^* \cdot \bar{x}_{s+} + b^* = +1$$

SVM (Hard margin)

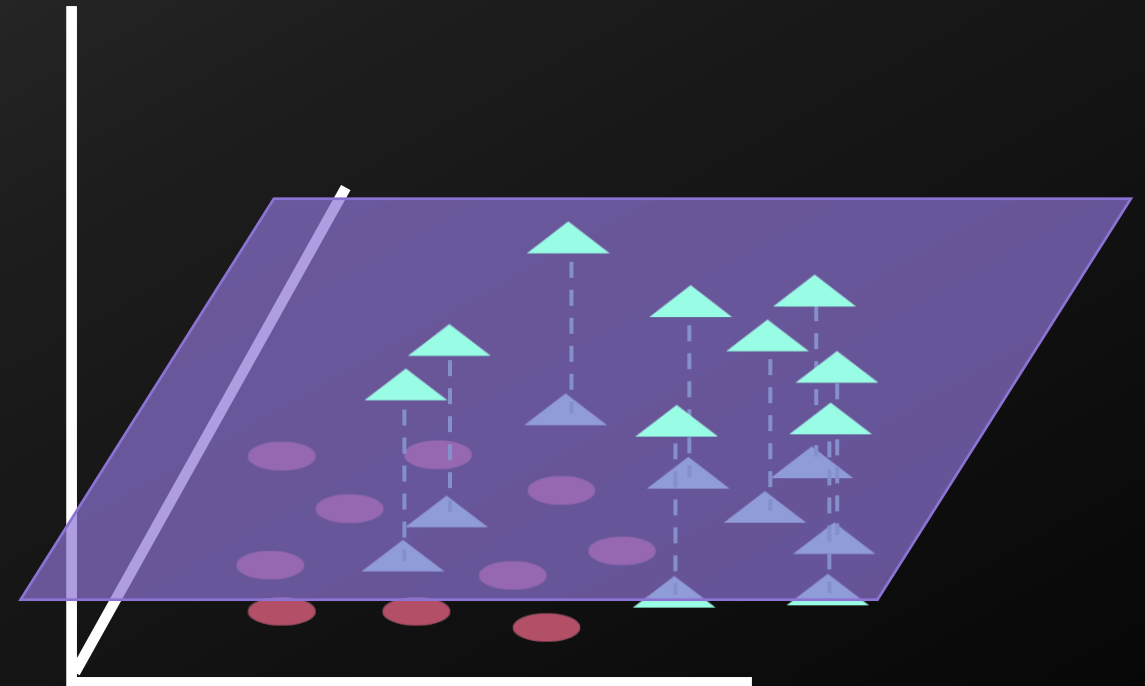
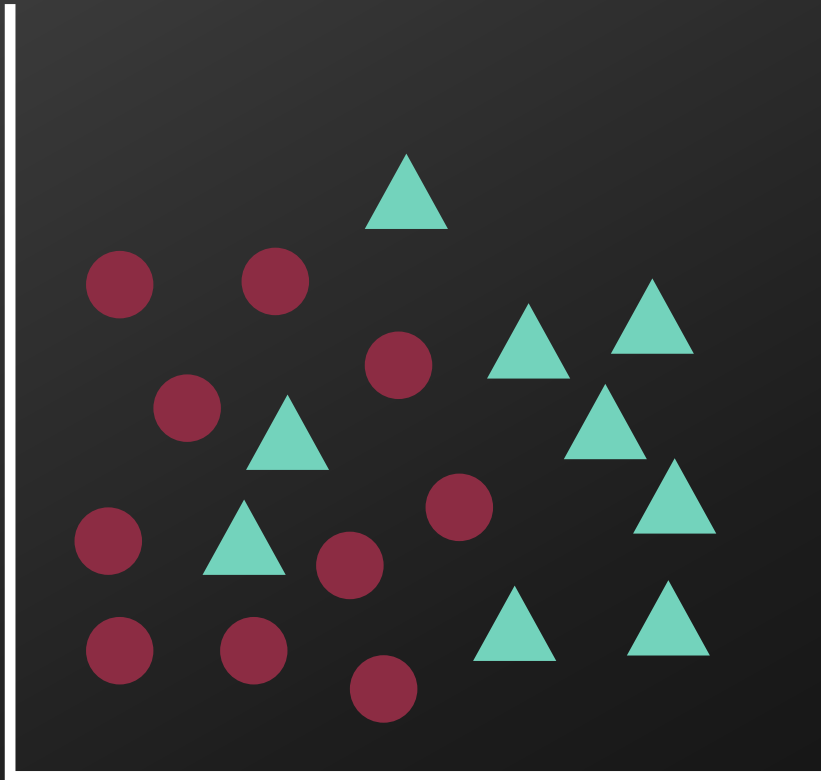
❖ Problem of hard margin method



- 1) Soft Margin with slack variable
- 2) Kernel trick

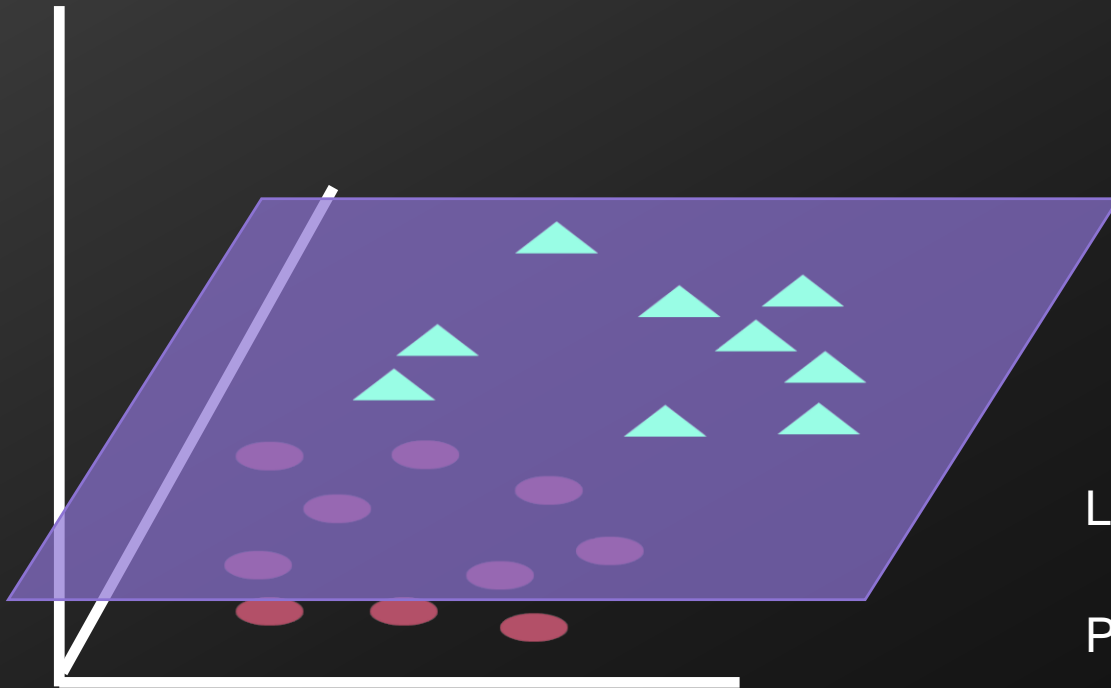
Kernel Trick

❖ Concept of Kernel trick



Kernel Trick

❖ Concept of Kernel trick



$$L = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (\bar{x}_i \cdot \bar{x}_j)$$

$$\Phi(\bar{x}_i) \cdot \Phi(\bar{x}_j)$$

$$k(\bar{x}_i, \bar{x}_j) = (\Phi(\bar{x}_i) \cdot \Phi(\bar{x}_j))$$

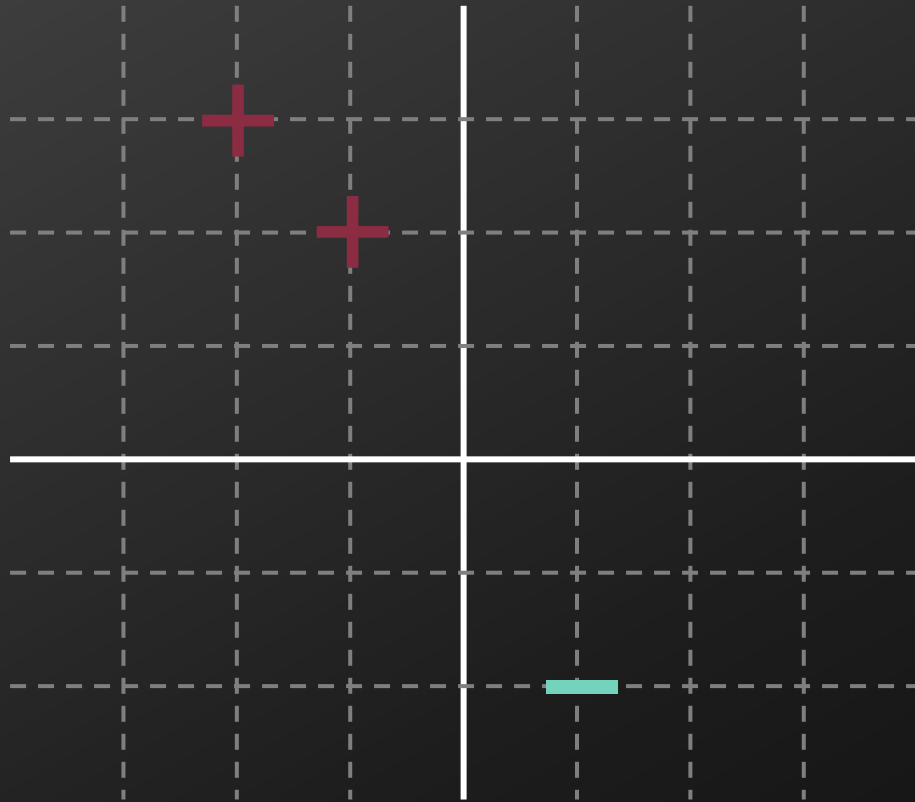
Linear : $k(\bar{x}_i, \bar{x}_j) = \bar{x}_i \cdot \bar{x}_j$

Polynomial : $k(\bar{x}_i, \bar{x}_j) = (1 + \bar{x}_i \cdot \bar{x}_j)^n$

Gaussian : $k(\bar{x}_i, \bar{x}_j) = e^{\frac{-\|\bar{x}_i - \bar{x}_j\|^2}{2\sigma^2}}$

Example

❖ SVM(Hard margin)



$$x_1 = (-1, 2), \quad x_2 = (-3, 3) \\ x_3 = (1, -2)$$

$$L = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (\bar{x}_i \cdot \bar{x}_j)$$

5	9	5
9	18	9
5	9	5

$$= \alpha_1 + \alpha_2 + \alpha_3 - \frac{5}{2} \alpha_1^2 - 9 \alpha_2^2 - \frac{5}{2} \alpha_3^2 \\ - 9 \alpha_1 \alpha_2 - 5 \alpha_1 \alpha_3 - 9 \alpha_2 \alpha_3$$

Example

❖ SVM(Hard margin)

$$L = \alpha_1 + \alpha_2 + \alpha_3 - \frac{5}{2}\alpha_1^2 - 9\alpha_2^2 - \frac{5}{2}\alpha_3^2 - 9\alpha_1\alpha_2 - 5\alpha_1\alpha_3 - 9\alpha_2\alpha_3$$

$$\sum \alpha_i y_i = \alpha_1 + \alpha_2 - \alpha_3 = 0 \quad \alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_3 \geq 0$$

from KKT condition 3

$$\alpha_i (y_i (\bar{w} \cdot \bar{x}_i + b - 1)) = 0$$

case 1) $\alpha_1 = 0$, case 2) $\alpha_2 = 0$, case 3) $\alpha_3 = 0$, case 4) $\alpha_1 \neq 0, \alpha_2 \neq 0, \alpha_3 \neq 0$

Example

❖ SVM(Hard margin)

Case 1) $\alpha_1 = 0$

$$\alpha_2 - \alpha_3 = 0, \quad \alpha_2 = \alpha_3 \quad L = \alpha_1 + \alpha_2 + \alpha_3 - \frac{5}{2}\alpha_1^2 - 9\alpha_2^2 - \frac{5}{2}\alpha_3^2 - 9\alpha_1\alpha_2 - 5\alpha_1\alpha_3 - 9\alpha_2\alpha_3$$

$$L = 2\alpha_2 - \frac{41}{2}\alpha_2^2$$

$$\frac{\partial L}{\partial \alpha_2} = 2 - 41\alpha_2 = 0 \quad \alpha_2 = 0.049, \alpha_3 = 0.049, \alpha_1 = 0$$

$$\bar{w} = \sum \alpha_i y_i \bar{x}_i = 0.049[(-3, 3) - (1, -2)] = 0.049(-4, 5) = (-0.196, 0.245)$$

$$\alpha_i[y_i(\bar{w} \cdot \bar{x}_i + b) - 1] = 0.049[(-0.196, 0.245) \cdot (-3, 3) + b - 1] = 0$$

$$b = 1 - (-0.196, 0.245) \cdot (-3, 3) = 1 - 1.323 = -0.323$$

Example

❖ SVM(Hard margin)

$$\bar{w} = (-0.196, 0.245) \quad b = -0.323$$

$$\bar{w} \cdot x_1 + b \geq 1 \quad (-0.196, 0.245) \cdot (-1, 2) - 0.323 = 0.363 \quad \text{Not satisfied}$$

$$\bar{w} \cdot x_2 + b \geq 1 \quad (-0.196, 0.245) \cdot (-3, 3) - 0.323 = 1 \quad \text{SV}$$

$$\bar{w} \cdot x_3 + b \leq -1 \quad (-0.196, 0.245) \cdot (1, -2) - 0.323 = -1 \quad \text{SV}$$

Example

❖ SVM(Hard margin)

Case 2) $\alpha_2 = 0$

$$\alpha_1 - \alpha_3 = 0, \quad \alpha_1 = \alpha_3 \quad L = \alpha_1 + \alpha_2 + \alpha_3 - \frac{5}{2}\alpha_1^2 - 9\alpha_2^2 - \frac{5}{2}\alpha_3^2 - 9\alpha_1\alpha_2 - 5\alpha_1\alpha_3 - 9\alpha_2\alpha_3$$

$$L = 2\alpha_1 - 10\alpha_1^2$$

$$\frac{\partial L}{\partial \alpha_1} = 2 - 20\alpha_1 = 0 \quad \alpha_1 = 0.1, \alpha_3 = 0.1, \alpha_2 = 0$$

$$\bar{w} = \sum \alpha_i y_i \bar{x}_i = 0.1[(-1, 2) - (1, -2)] = 0.1(-2, 4) = (-0.2, 0.4)$$

$$\alpha_i [y_i (\bar{w} \cdot \bar{x}_i + b) - 1] = 0.1[(-0.2, 0.4) \cdot (-1, 2) + b - 1] = 0$$

$$b = 1 - (-0.2, 0.4) \cdot (-1, 2) = 1 - 1 = 0$$

Example

❖ SVM(Hard margin)

$$\bar{w} = (-0.2, 0.4) \quad b = 0$$

$$\bar{w} \cdot x_1 + b \geq 1 \quad (-0.2, 0.4) \cdot (-1, 2) = 1$$

SV

$$\bar{w} \cdot x_2 + b \geq 1 \quad (-0.2, 0.4) \cdot (-3, 3) = 1.8$$

Satisfied

$$\bar{w} \cdot x_3 + b \leq -1 \quad (-0.2, 0.4) \cdot (1, -2) = -1$$

SV

Case 3) $\alpha_3 = 0$

$\alpha_1 + \alpha_2 = 0$, **Impossible case**

Example

❖ SVM(Hard margin)

Case 4) $\alpha_1 \neq 0, \alpha_2 \neq 0, \alpha_3 \neq 0$

$$L = \alpha_1 + \alpha_2 + \alpha_3 - \frac{5}{2}\alpha_1^2 - 9\alpha_2^2 - \frac{5}{2}\alpha_3^2 - 9\alpha_1\alpha_2 - 5\alpha_1\alpha_3 - 9\alpha_2\alpha_3$$

$$\frac{\partial L}{\partial \alpha_1} = 1 - 5\alpha_1 - 9\alpha_2 - 5\alpha_3 = 0$$

$$\frac{\partial L}{\partial \alpha_2} = 1 - 18\alpha_2 - 9\alpha_1 - 9\alpha_3 = 0$$

$$\frac{\partial L}{\partial \alpha_3} = 1 - 5\alpha_3 - 5\alpha_1 - 9\alpha_2 = 0$$

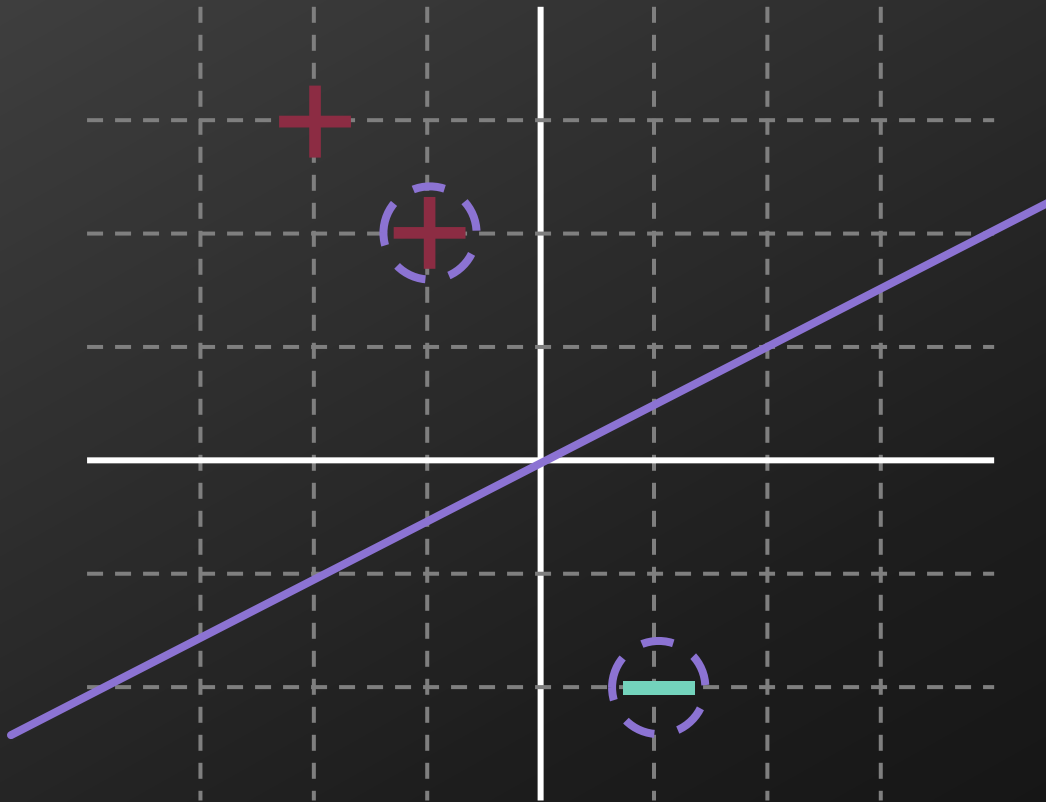
$$\begin{bmatrix} 5 & 9 & 5 \\ 9 & 18 & 9 \\ 5 & 9 & 5 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 5 & 9 & 5 \\ 9 & 18 & 9 \\ 5 & 9 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.444 \\ 0.5 \end{bmatrix}$$

Impossible case

Example

❖ SVM(Hard margin)



$$\bar{w} = (-0.2, 0.4) \quad b = 0$$

$$\bar{w} \cdot (x, y) + b = 0$$

$$-0.2x + 0.4y = 0 \quad y = 0.5x$$

Example

❖ SVM(Hard margin)

```

1 - X=[-1 2;-3 3;1 -2];
2 - Y=['p':'p':'n'];
3 -
4 - SVMModel=fitcsvm(X,Y);
5 -
6 - sv = SVMModel.SupportVectors;
7 - figure
8 - gscatter(X(:,1),X(:,2),Y)
9 - hold on
10 - plot(sv(:,1),sv(:,2),'ko','MarkerSize',10)
11 - legend('+','-','Support Vector')
12 - grid on;
13 - xlim([-5 5])
14 - ylim([-5 5])
15 - line([-5 5],[0 0])
16 - line([0 0],[-5 5])
17 -
18 - w=SVMModel.Beta
19 - b=SVMModel.Bias
20 -
21 - x=linspace(-5,5,100);
22 - y=-1*(w(1)/w(2))*x+b;
23 - plot(x,y)

```

```

>> svm_test1

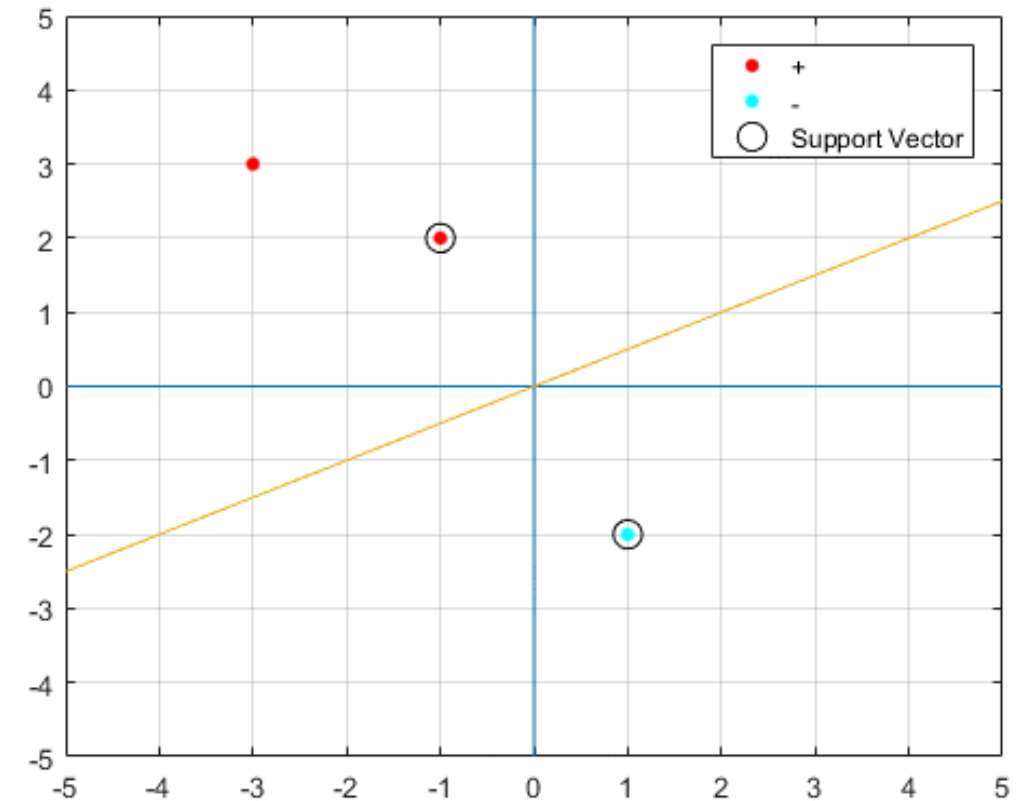
w =

-0.2000
0.4000

b =

0

```



Q & A *Thank You!!!*