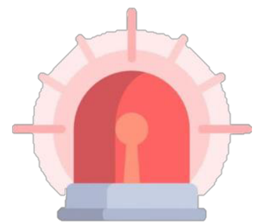
Support Vector Machine

$\boxed{\text{SVM}}$

Caution

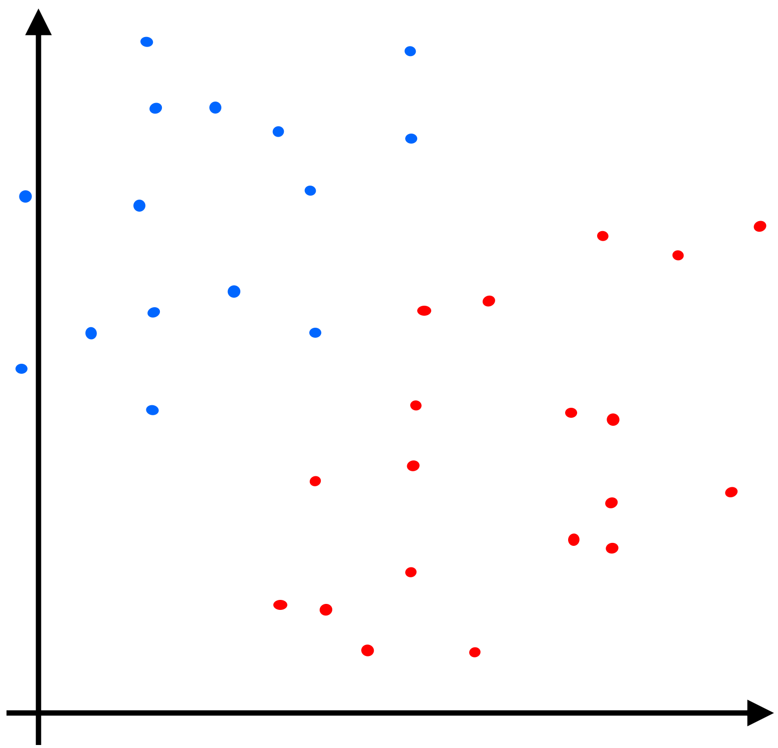


Very Complex and Difficult

But, many details will be skipped!!

Linear Support Vector Machine

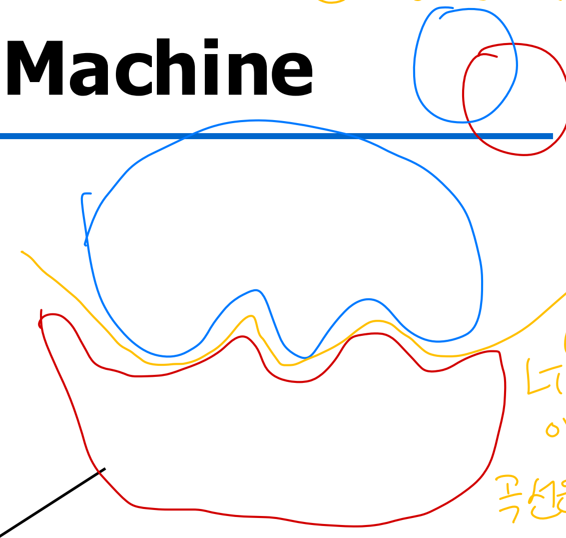
- **Find a linear boundary**
  - Two class problem



Linear Support Vector Machine

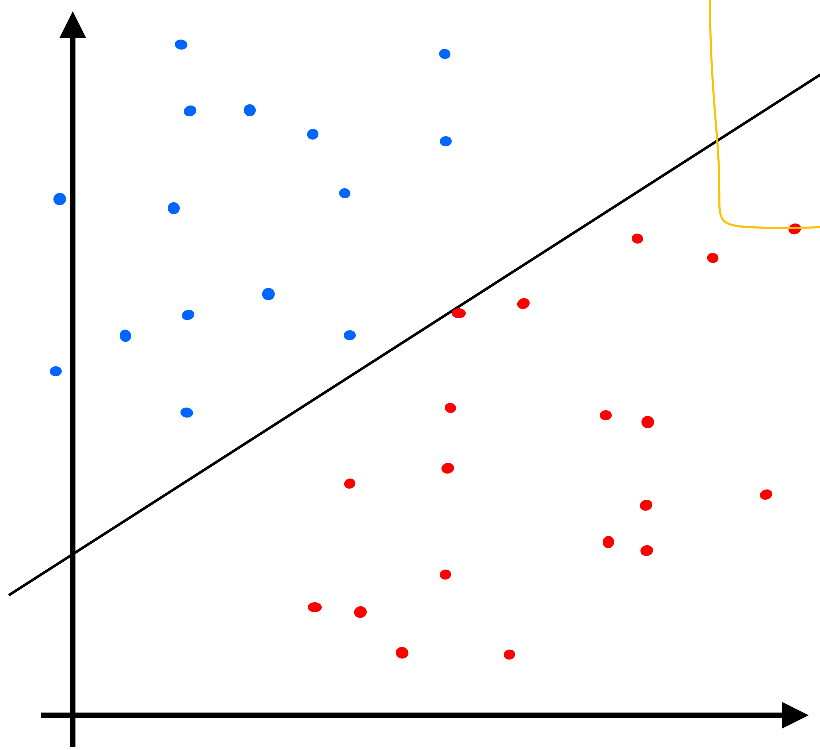
<LSVM배우고 그다음주제>

① 겹쳐진 경우 어떻게?



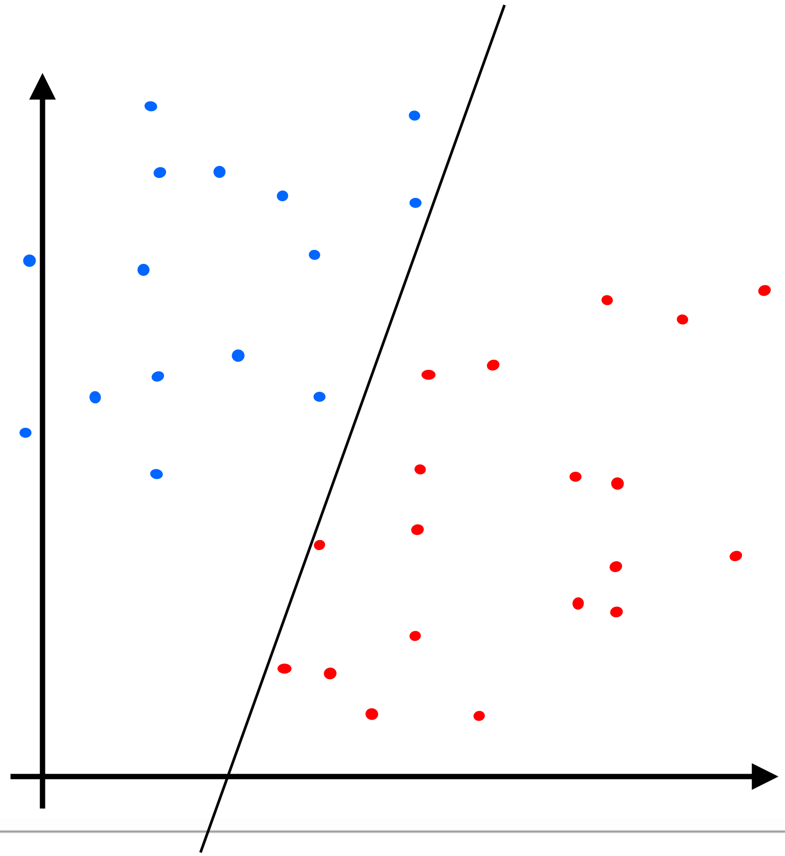
② Linear가 아니라 곡선은 어떻게?

- **How about this?**
  - Correct boundary.. Good or Bad?



Linear Support Vector Machine

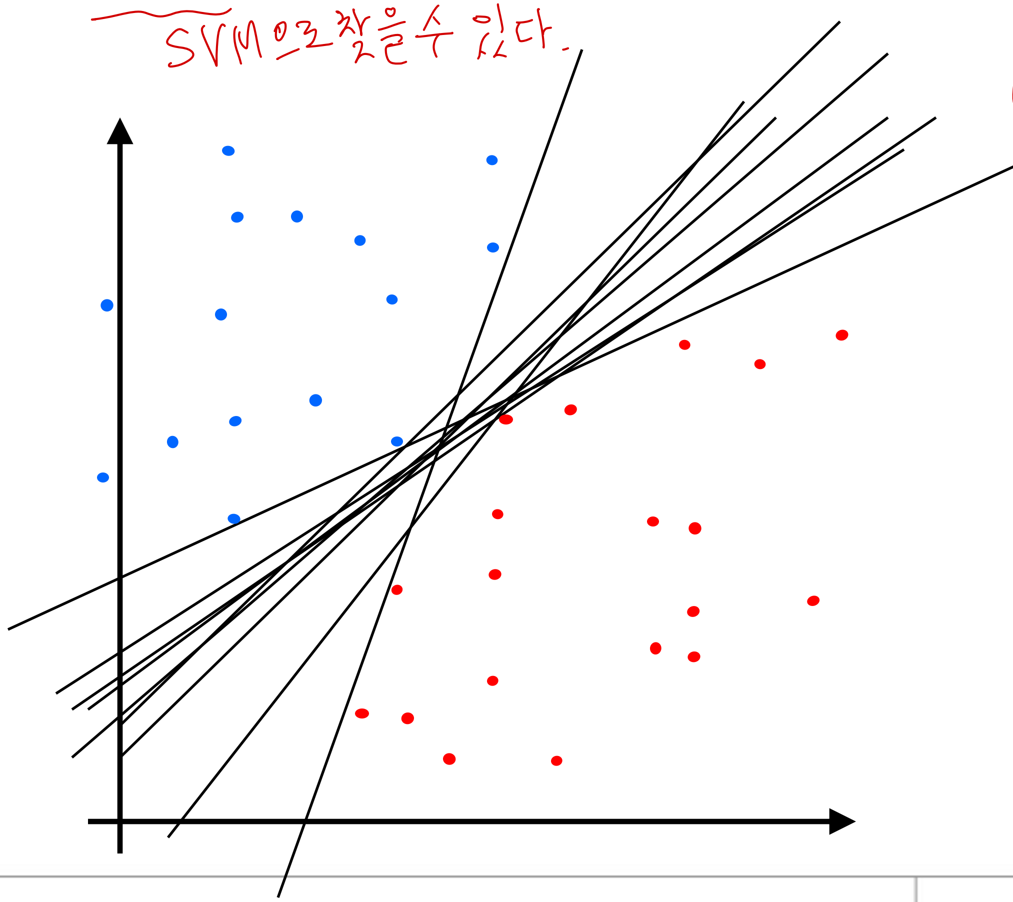
- **How about this?**
  - Correct boundary.. Good or Bad?



Linear Support Vector Machine

- **Any of these would be find.. But**
  - Which is best??

SVM으로 찾을 수 있다.

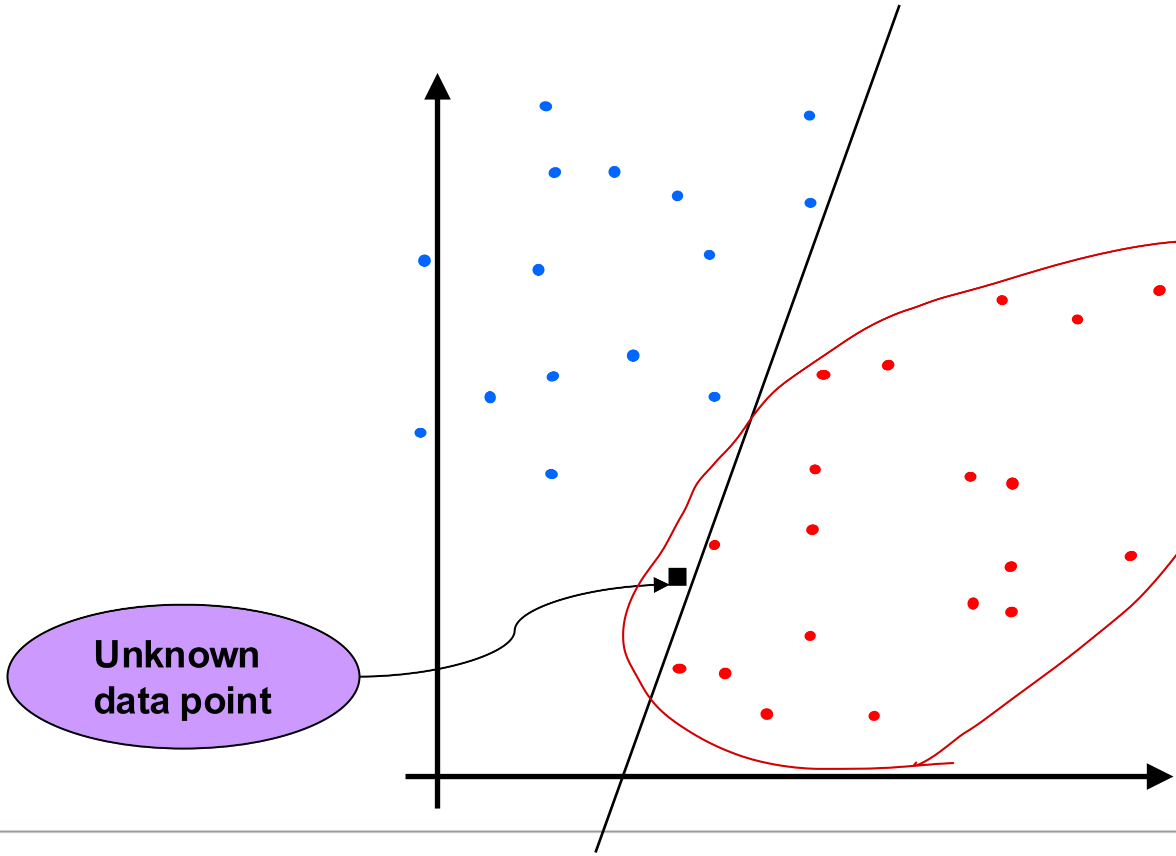


Not Bad Boundaries

→ 최고를 찾고 싶어!

Linear Support Vector Machine

- **A Given Unknown Data Point**
  - RED or Blue ?



Based on the intuition, which class may it belong to?

Red에 가깝지만

경계선에 의하면 Blue가 됨

→ 좋지 않은 경계

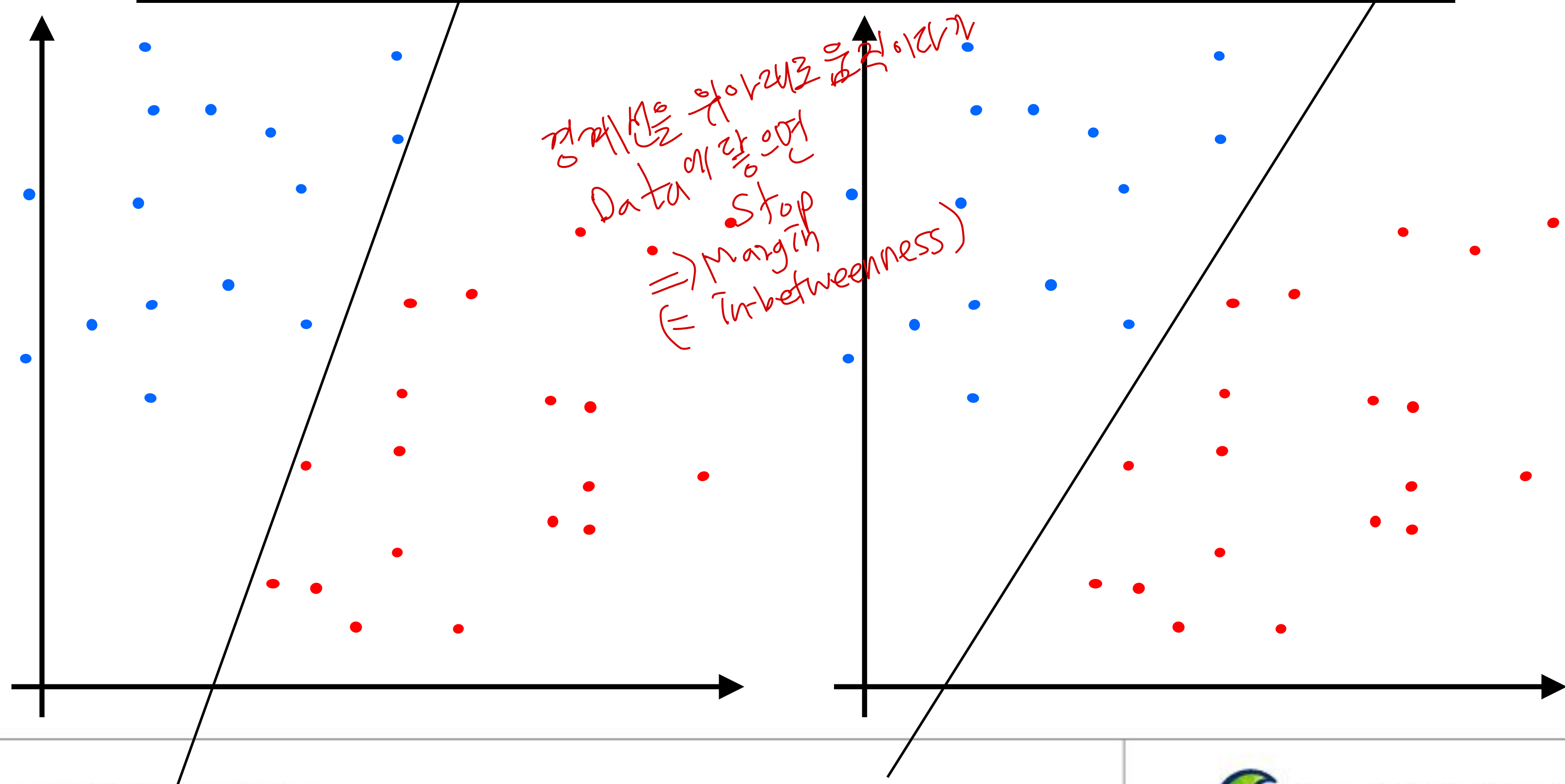
Based on the boundary, which class may it belong to?

Linear Support Vector Machine

- **Which boundary is better?**

Which boundary is more "in between" class ?

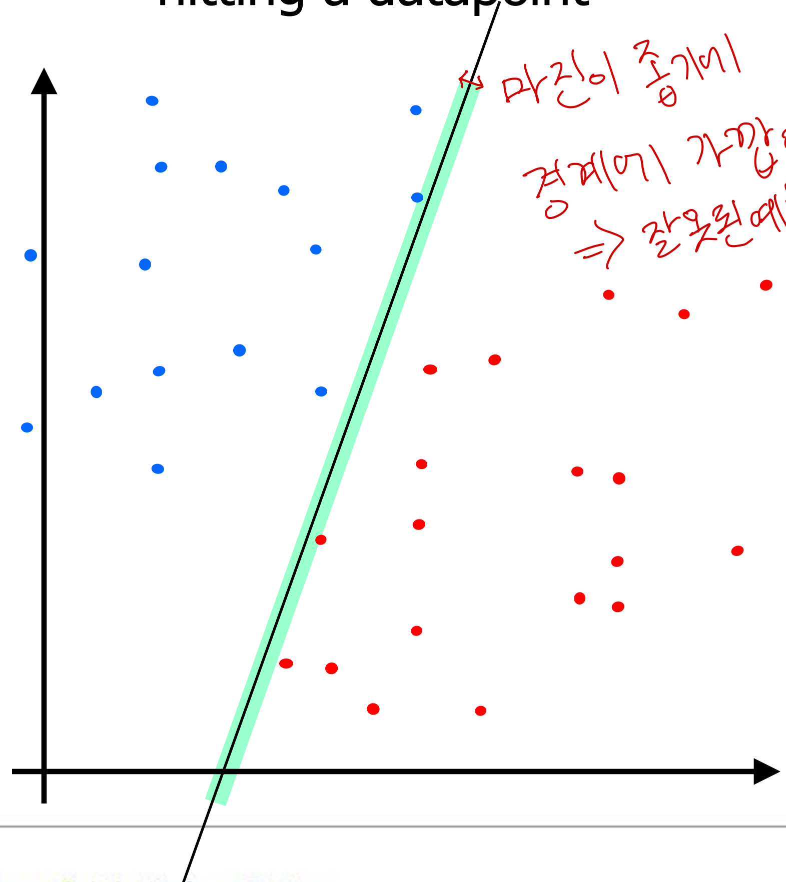
How can you evaluate the in-bewteenness ?



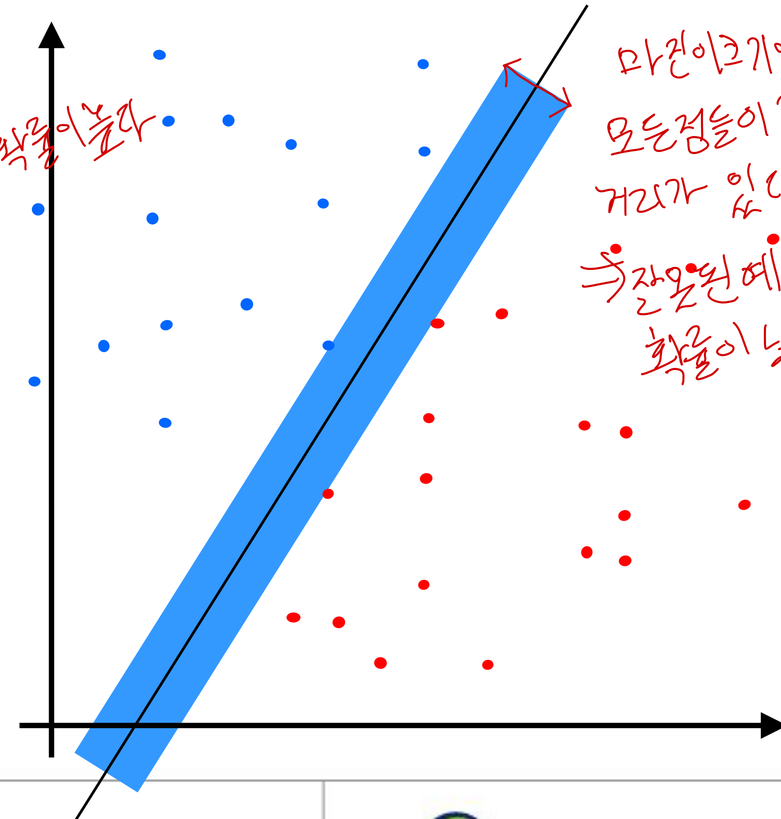
경계선을 위아래로 움직이다가
Data에 닿으면
Stop
⇒ Margin
(≡ In-betweenness)

Linear Support Vector Machine

- **Margin of a linear classifier** ∴ Best ⇒ "Largest Margin"
  - the width that the boundary could be increased by before hitting a datapoint



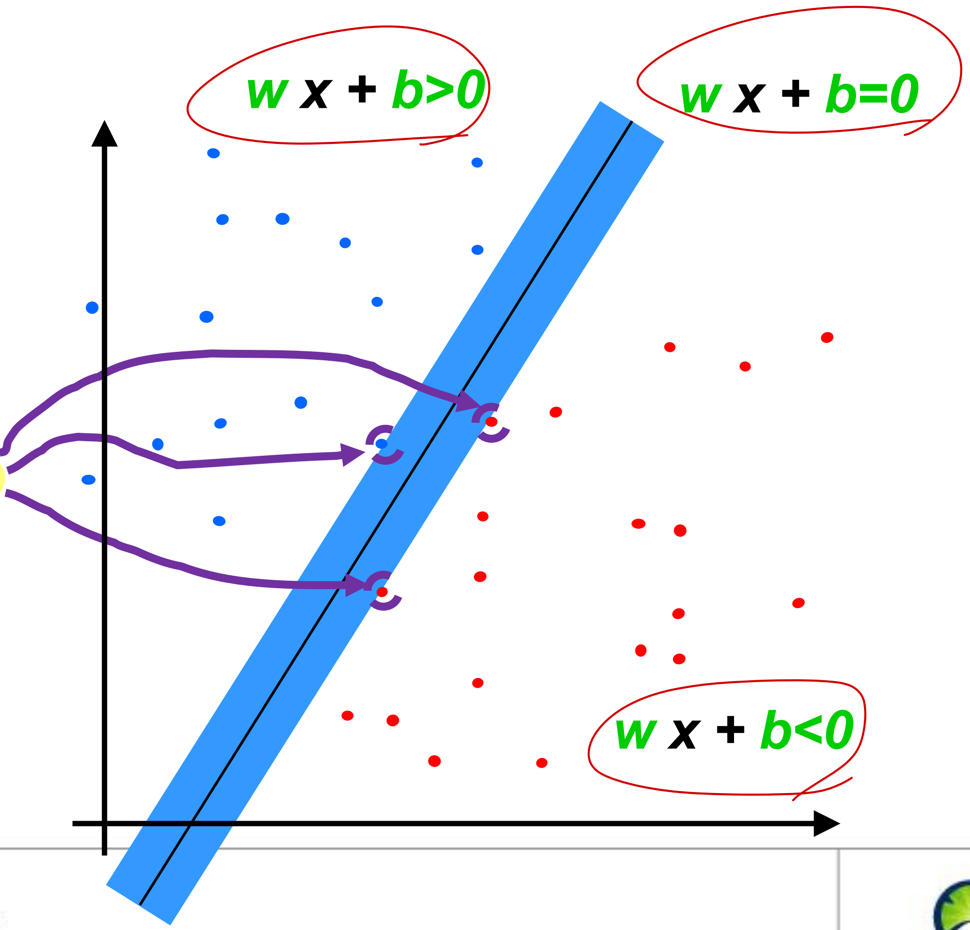
마진이 좁기에 경계에 가깝다
⇒ 잘못된 예측할 확률이 높다



마진이 크기에 모든 점들이랑 거리가 있다
⇒ 잘못된 예측할 확률이 낮다

Linear Support Vector Machine

- **Let's find the boundary with maximum margin**
  - Maximizing the margin is good according to intuition

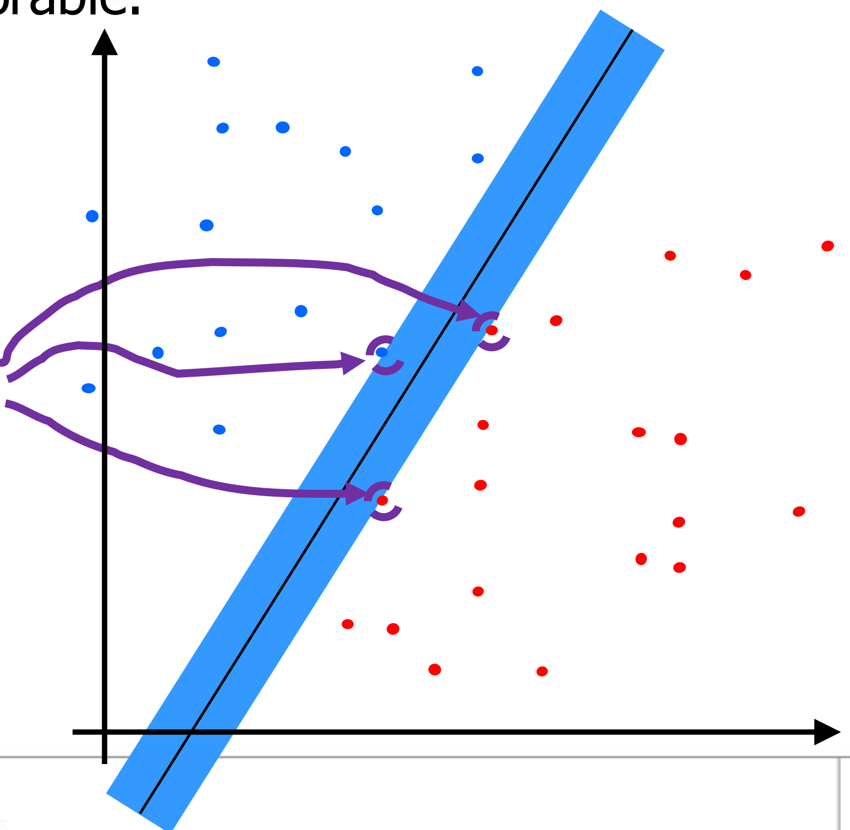


Support Vectors are those datapoints that the margin pushes up against

Linear Support Vector Machine

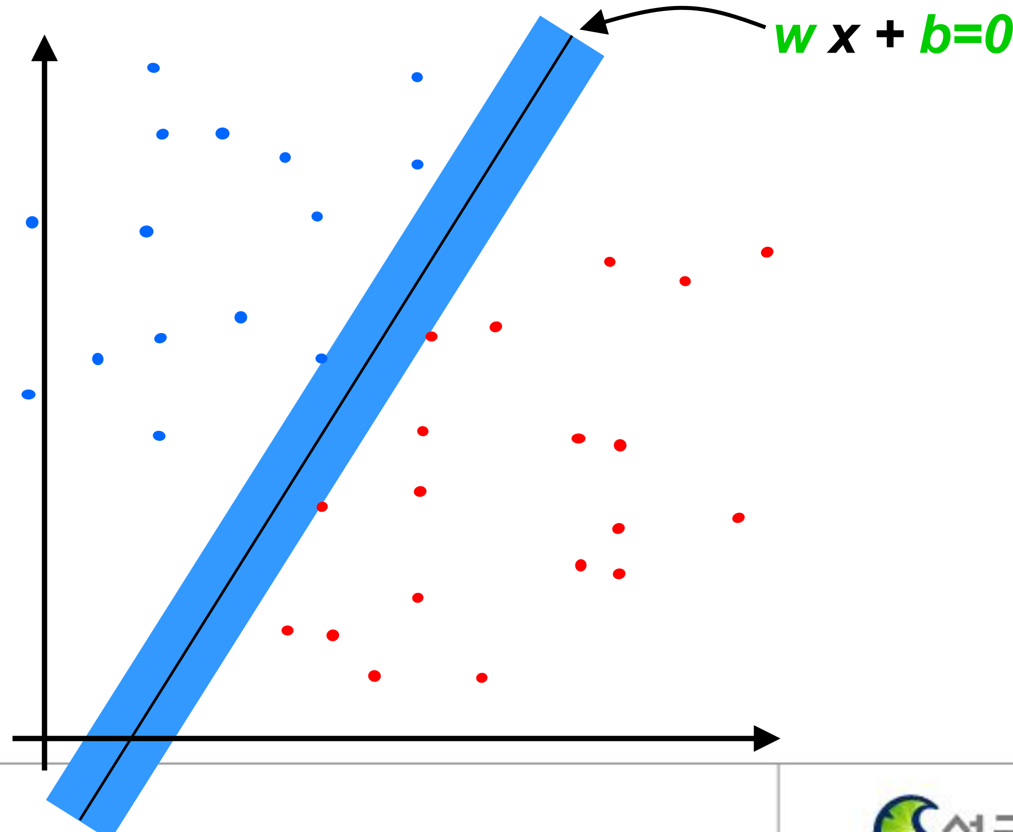
- **LSVM**
  - Empirically it works very well.
  - Only support vectors are important; other training examples are ignorable.

Support Vectors are those datapoints that the margin pushes up against



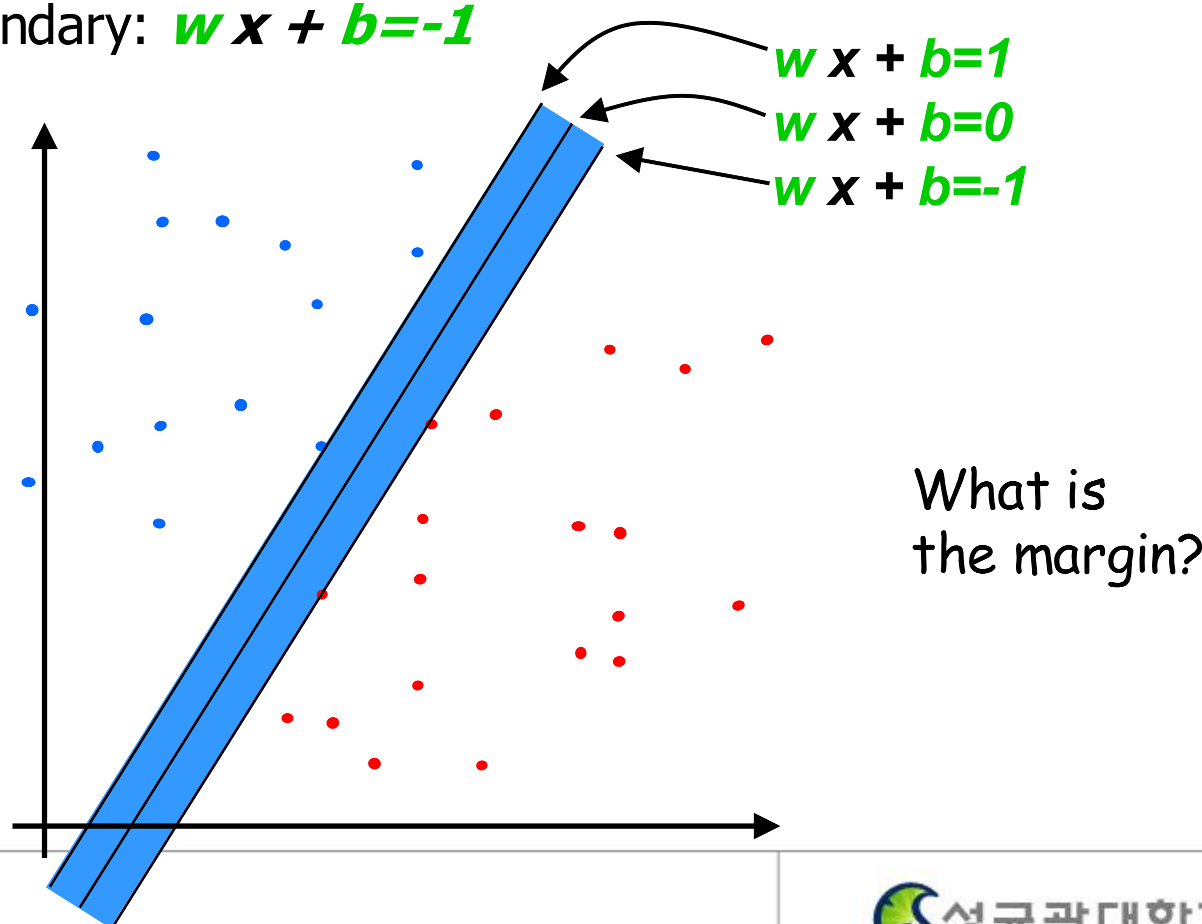
Learning Linear SVM

- **Let's find the maximum margin boundary**
  - First, let's assume that we have the boundary: $\mathbf{w}\,\mathbf{x} + b = 0$



Learning Linear SVM

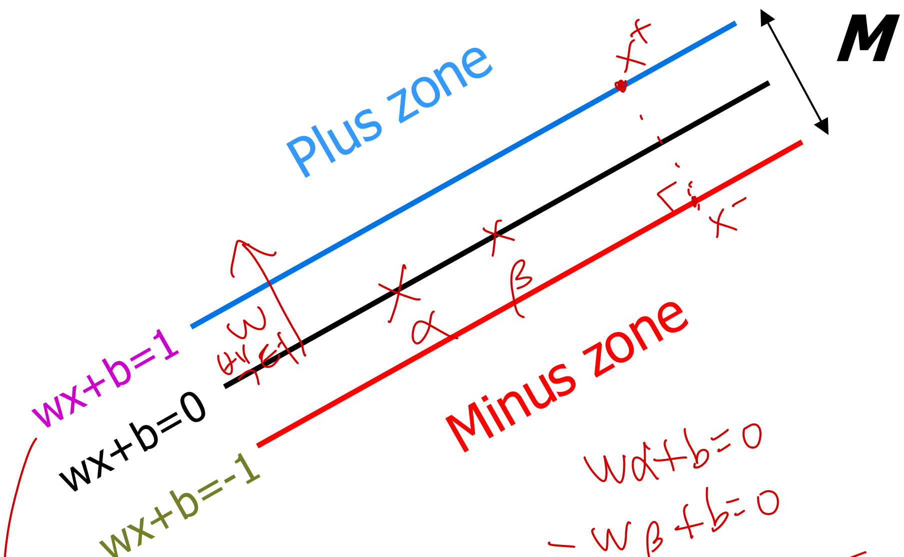
- **Let's adjust *w* and *b* so that**
  - Plus boundary: $\boldsymbol{w}\,\boldsymbol{x} + \boldsymbol{b}=1$
  - Minus boundary: $\boldsymbol{w}\,\boldsymbol{x} + \boldsymbol{b}=-1$



Learning Linear SVM

GPT 돌려서 이해하기

- **Margin**
  - Distance between Plus and Minus boundaries



wα+b=0

-wβ+b=0

(wα-β)=0

1. Vector **w** is perpendicular to the lines

   Why?
   Let's say ~~a~~ α and ~~b~~ β are on the line
   Then, $\mathbf{w}(\cancel{a-b}) = 0$
   α-β

2. Let's say $\mathbf{x}^+$ is on the plus line and $\mathbf{x}^-$ is the corresponding point on the minus line.

   $(\mathbf{x}^+ - \mathbf{x}^-)$ is perpendicular to the lines.
   Why?

3. $|\mathbf{x}^+ - \mathbf{x}^-| = M$ Why?
   vector

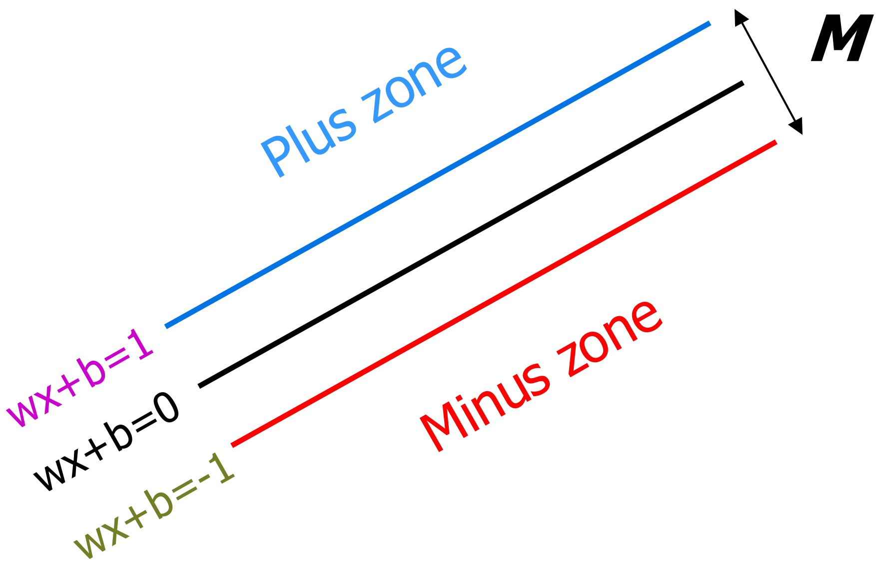
4. $\mathbf{x}^+ = \mathbf{x}^- + \lambda w$ Why?

5. $\mathbf{w}(\mathbf{x}^- + \lambda \mathbf{w}) + b = 1$ Why?

Learning Linear SVM

- **Margin**
  - Distance between Plus and Minus boundaries

Distance



3. $|\mathbf{x}^+ - \mathbf{x}^-| = M$
4. $\mathbf{x}^+ = \mathbf{x}^- + \lambda w$

$\rightarrow |\mathbf{x}^+ - \mathbf{x}^-| = M = \lambda |w|$

5. $\mathbf{w}(\mathbf{x}^- + \lambda \mathbf{w}) + b = 1 \longrightarrow \mathbf{w}\mathbf{x}^- + \lambda \mathbf{w}\cdot\mathbf{w} + b = 1$

=-1

6. $-1 + \lambda \mathbf{w}\cdot\mathbf{w} = 1 \longrightarrow \lambda = 2/\mathbf{w}\cdot\mathbf{w}$

Innerproduct

$$M = \frac{2}{|w|} \quad \text{OR} \quad \frac{2}{w \cdot w}$$

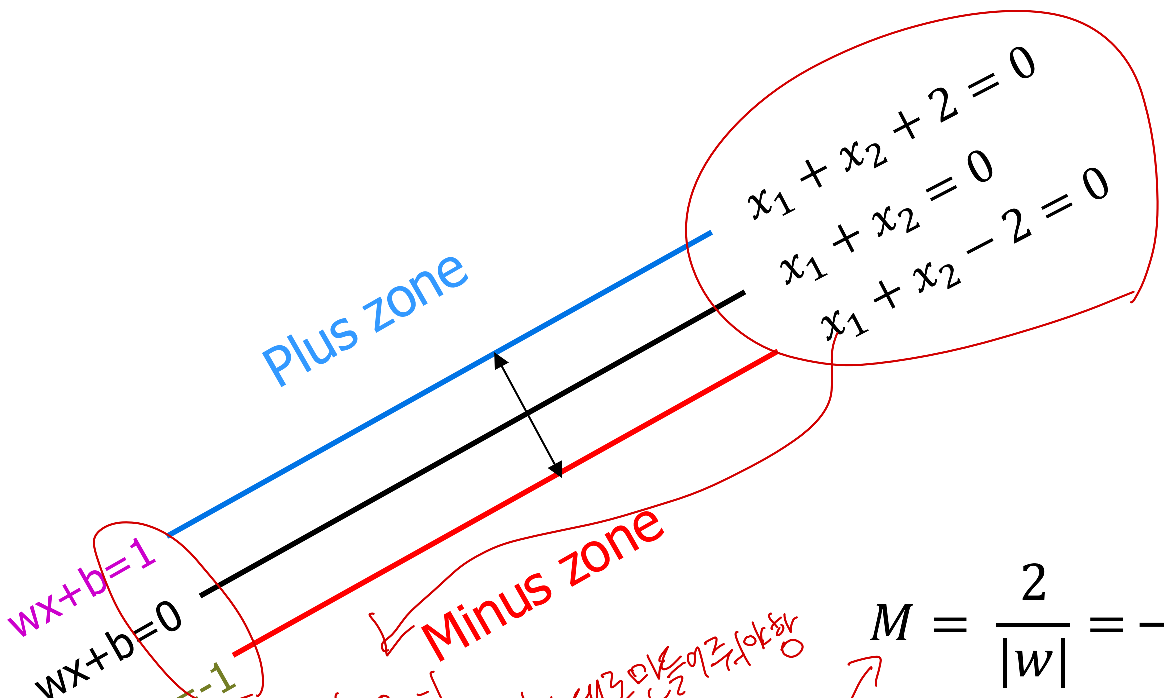
둘중 뭐가 맞는지
나중에 찾아보기

벡터 W의 길이

Learning Linear SVM

- **Example Margin**
  - Distance between Plus and Minus boundaries

$w = \left(-\frac{1}{2}, -\frac{1}{2}\right)$



$$-\frac{1}{2}x_1 - \frac{1}{2}x_2 = 1$$
$$-\frac{1}{2}x_1 - \frac{1}{2}x_2 = 0$$
$$-\frac{1}{2}x_1 - \frac{1}{2}x_2 = -1$$

우변 상수가 1,0,-1 형태가 아니면 이 형태로 만들어주어야함

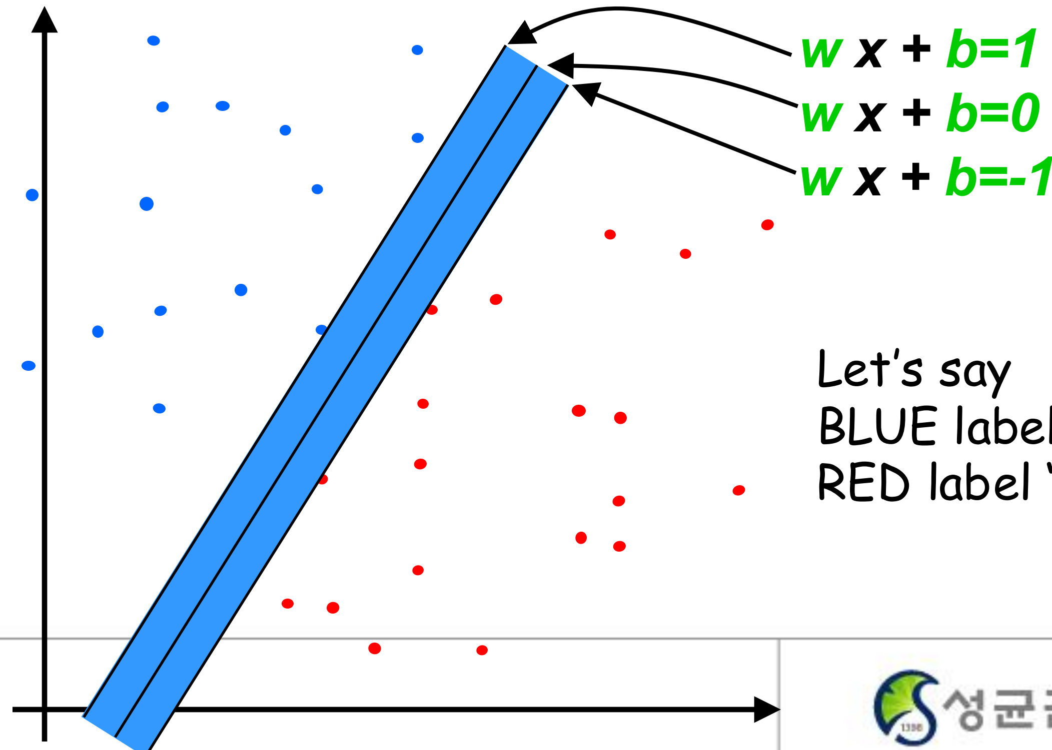
$$M = \frac{2}{|w|} = \frac{2}{\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2}} = 2\sqrt{2}$$

Learning Linear SVM

- **Goal: Maximize the Margin** $M = \dfrac{2}{|w|}$

  while the following conditions are satisfied
  - All plus points (Blue) are above the plus boundary
  - All minus points (Red) are below the minus boundary

boundary의 조건

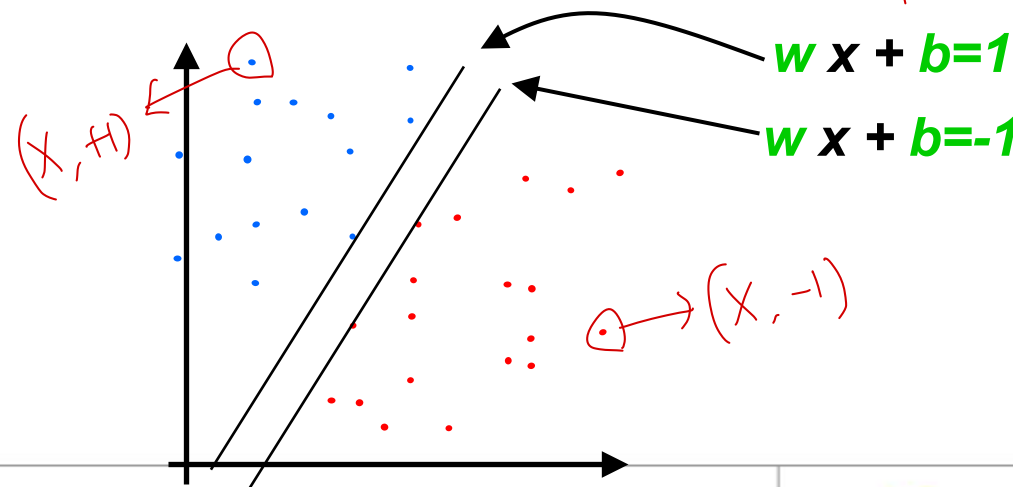


Let's say
BLUE label "+1" and
RED label "-1"

Learning Linear SVM

- **Goal: Maximize the Margin** $M = \dfrac{2}{|w|}$

  while the following conditions are satisfied
  - All plus points ($y_i = 1$) are above the plus boundary
    - ⇨ If $y_i = +1, \quad \mathbf{w} \cdot \mathbf{x}_i + b \geq 1$ (Blue 대신 +1 사용) =label (다른 값 사용해도 괜찮음, +1/-1이 제일 쉬움)
  - All minus points ($y_i = -1$) are below the minus boundary
    - ⇨ If $y_i = -1, \quad \mathbf{w} \cdot \mathbf{x}_i + b \leq -1$ (Red 대신 -1 사용) =label



Learning Linear SVM

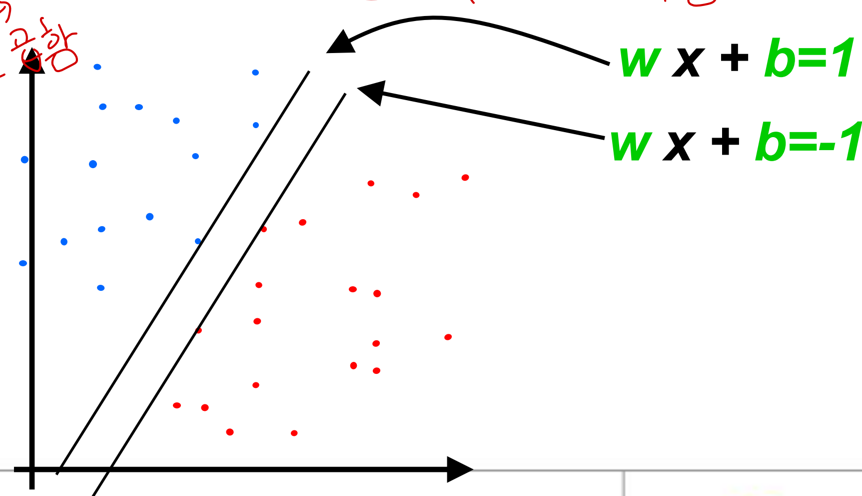
- **Goal: Maximize the Margin** $M = \dfrac{2}{|w|}$

while the following conditions are satisfied

두조건을 (1개로 합침)

- All plus points ($y_i = 1$) are above the plus boundary
  - ⇨ If $y_i = +1$, y_i 곱함→ $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$
- All minus points ($y_i = -1$) are below the minus boundary
  - ⇨ If $y_i = -1$, $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$ 음수라 ≥1 됨

y_i를 곱함



Learning Linear SVM

- **Goal: Maximize the Margin** $M = \frac{2}{|w|}$

  Find w to

  w 조건①

  w가 마진을 최대로 만드는 값

  while the following conditions are satisfied

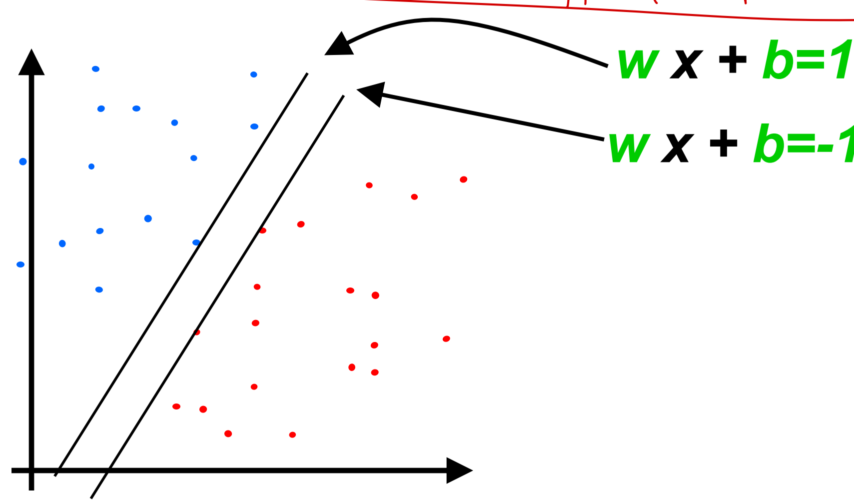
  1개의 조건 $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \text{ for } i = 1, \cdots, n$

  w조건②

  w가 B/R을 구분하는 경계선

  => wx+b=1, -1

  # of samples



조건에 의하면 일단은 Data set이

이렇게 분포된 것은 무시하고 생각하자

☆기말출제 → 손으로 푸는 문제 나옴

Learning Linear SVM

- **Formalization of LSVM**
  - When we have $D = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$ *where* $y_i \in \{-1, +1\}$

$$\underset{\mathbf{w},b}{\operatorname{argmax}} \frac{2}{|\mathbf{w}|}$$

Margin

← 앞 슬라이드 내용과 같음

$$\text{subject to} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \; i = 1, \dots, n$$

  - or

$|w|^2$ ($|w|$ 대신 사용)

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2} \mathbf{w} \cdot \mathbf{w}$$

$$\text{subject to} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \; i = 1, \dots, n$$

$w \cdot x + b = 0$ seperate Blue & Red

  - If you solve the problem, we obtain the maximum margin boundary

But, how can I solve this?? Don't worry.
We use Qudratic Programming Technique,
but you don't need to know the detail of QP

최적화 기법

python 라이브러리

SVM ⇒ 요즘에는 잘 안씀
(∵ 계산, 시간이 많이 듦, $O(n^2)$이라)

Dual Form for Learning LSVM

- **Example without QP**
  - What is the maximum margin boundary for $(0,1,1)$ and $(1,0,-1)$

x y label   x y label

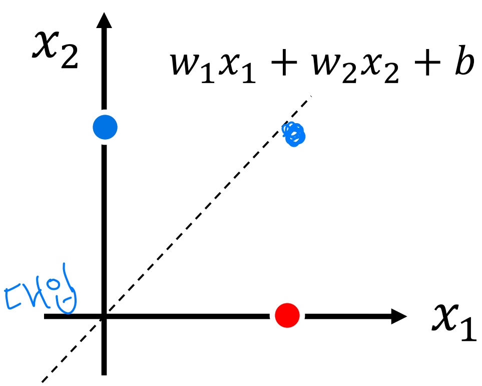
ex $(1,1,-1)$ 추가

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2}\mathbf{w}\cdot\mathbf{w} = \frac{1}{2}(w_1^2+w_2^2)$$

subject to $\quad y_i(\mathbf{w}\cdot\mathbf{x}_i + b) \geq 1 \;\; i = 1,\ldots,n$

$y_3(w_1x_1+w_2x_2+b) \geq 1$

① 대입



② solve

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2}(w_1,w_2)\cdot(w_1,w_2)$$

subject to $\quad 1\cdot(1\cdot w_2 + b) \geq 1$

$\quad -1\cdot(1\cdot w_1 + b) \geq 1$

$\Rightarrow$

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2}(w_1^2+w_2^2)$$

subject to $\quad w_1 \leq -1 - \mathrm{b}$

$\quad w_2 \geq 1 - \mathrm{b}$

$W_1 + W_2 \leq -1-b$

QP가 계산해줌

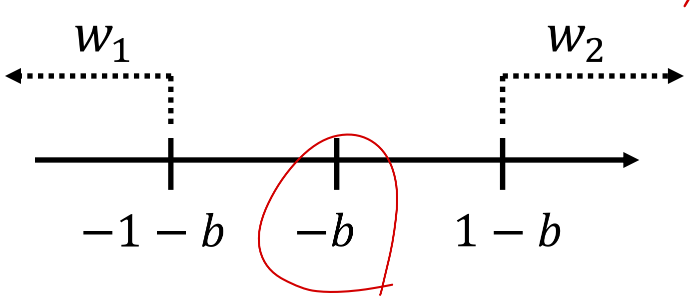
Dual Form for Learning LSVM

- **Example without QP**
  - What is the maximum margin boundary for $(0,1,1)$ and $(1,0,-1)$

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2}(w_1^2 + w_2^2)$$

subject to $w_1 \le -1 - b$

$w_2 \ge 1 - b$



3 가지 가정

$i)\ 1 - b \le 0$

최솟값이 되려면 $w_1 = -1 - b$, $w_2 = 0$

$ii)\ -1 - b \le 0 \le 1 - b$

둘 다 제곱이 최소되려면 $w_1 = -1 - b$, $w_2 = 1 - b$

$iii)\ 0 \le -1 - b$

$w_1 = 0$

$w_2 = 1 - b$

$i)\ 1 \le b$

$w_1^2 + w_2^2 = \frac{1}{2}(-1-b)^2$

$= \frac{4}{2}$ when $b = 1$

$ii)\ -1 \le b \le 1$

$w_1^2 + w_2^2 = b^2 + 1$ $\quad \frac{1}{2}(2b^2+2)$

$= 1$ when $b = 0$

minimum!

$\Rightarrow w_1 = -1,\ w_2 = 1,\ b = 0$

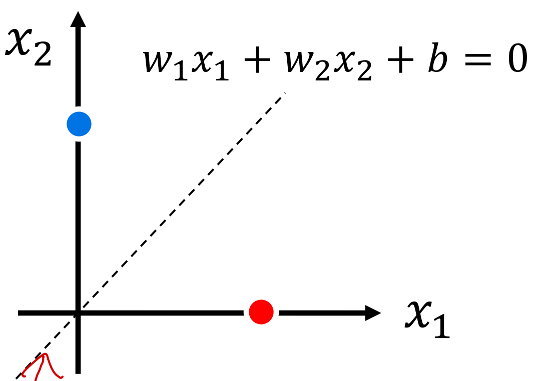
$iii)\ b \le -1$

$w_1^2 + w_2^2 = \frac{1}{2}(1-b)^2$

$= \frac{4}{2}$ when $b = -1$

Dual Form for Learning LSVM

- **Example without QP**
  - What is the maximum margin boundary for $(0,1,{\color{blue}1})$ and $(1,0,{\color{red}-1})$



$ii)\ -1-b \le 0 \le 1-b$

$w_1 = -1-b$

$w_2 = 1-b$

$ii)\ -1 \le b \le 1$

$w_1^2 + w_2^2 = b^2 + 1$

$= 1$ when $b = 0$

$w_1^2 + w_2^2$ is minimized
when $b = 0, w_1 = -1, w_2 = 1$

Boundary is
$-x_1 + x_2 = 0$

Dual Form for Learning LSVM

- **Lagrangian Dual Form of the Original Problem**
  - Instead of solving the original problem

변환가능(같은 솔루션)

$$\underset{w,b}{\operatorname{argmin}} \frac{1}{2}\mathbf{w}\cdot\mathbf{w}$$

$$\text{subject to} \quad y_i(\mathbf{w}\cdot\mathbf{x}_i + b) \geq 1 \; i = 1, \dots, n$$

변수 : w, b

  - We can find out $\mathbf{w}$ and b by solving Lagrangian Dual Form of the Original Problem

# of samples

변수 : $\alpha_1, \alpha_2, \cdots, \alpha_n$

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j)$$

$$\text{subject to} \quad \begin{cases} \alpha_i \geq 0 \; for \; i = 1, \dots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases}$$

n+1 개 condition

~~How do both give the same solution?~~ 안배움

~~Hmm.. You don't need to know the details~~

☆ 암기

Dual Form for Learning LSVM

- **Lagrangian Dual Form of the Original Problem**
  - For given $D = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\}$ where $y_i \in \{-1, +1\}$
  - For $\mathbf{x}_i$, define a variable α_i
  - Solve the followings, and obtain $\alpha_1, \alpha_2, \cdots, \alpha_n$

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j)$$

$$\text{subject to} \quad \begin{cases} \alpha_i \geq 0 \ for \ i = 1, \ldots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases}$$

① solve 하면 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 얻을수있음

② 적용

  - Then, we can obtain $\mathbf{w}$ and b for the boundary of $f(\mathbf{x}) = \mathbf{w}\mathbf{x} + b$

$$\mathbf{w} = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i$$

$$b = y_k - \mathbf{w} \cdot \mathbf{x}_k \quad \text{for any } \mathbf{x}_\mathbf{k} \text{ such that } \alpha_k > 0$$

Support vectors

③ w, b 얻기

Dual Form for Learning LSVM

- **Why Dual Form?**

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x_j})$$

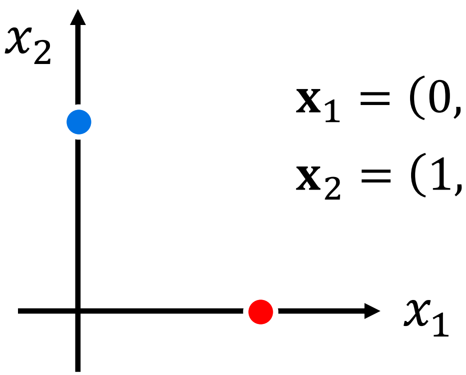
$$\text{subject to} \quad \begin{cases} \alpha_i \geq 0 \; for \; i = 1, \ldots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases}$$

– We can easily solve using "Quadratic Programming"
– There are many efficient algorithms to solve !!
– Especially, we may use **Kernel Trick** !! <- More important

I do NOT explain what QP is,
but want to say "It is NOT difficulty to solve"

Dual Form for Learning LSVM

- **Example**
  - What is the maximum margin boundary for $\overset{\alpha_1}{(0,1,1)}$ and $\overset{\alpha_2}{(1,0,-1)}$



$$\mathbf{x}_1 = (0,1) \quad y_1 = 1$$
$$\mathbf{x}_2 = (1,0) \quad y_2 = -1$$

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \overset{\text{내적}}{(\mathbf{x}_i \cdot \mathbf{x}_j)}$$

$$\text{subject to} \begin{cases} \alpha_i \geq 0 \ for\ i = 1, \dots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases}$$

$$\underset{\alpha_1\alpha_2}{\operatorname{argmax}}\ \alpha_1 + \alpha_2 - \frac{1}{2}\left(\begin{array}{l} \overset{\text{스칼라}}{\boxed{1\cdot 1\cdot \alpha_1 \cdot \alpha_1}} \cdot \overset{\text{벡터}}{(0,1)} \overset{\text{내적}}{\cdot} \overset{\text{벡터}}{(0,1)} + \\ 1\cdot -1 \cdot \alpha_1 \cdot \alpha_2 \cdot (0,1)\cdot(1,0) + \\ -1\cdot 1\cdot \alpha_2\cdot \alpha_1 \cdot (1,0)\cdot(0,1) + \\ -1\cdot -1\cdot \alpha_2 \cdot \alpha_2 \cdot (1,0)\cdot(1,0) \end{array}\right)$$

$$\text{subject to} \begin{cases} \alpha_1, \alpha_2 \geq 0 \\ 1\cdot\alpha_1 + (-1)\cdot\alpha_2 = 0 \end{cases}$$

$\left.\begin{array}{l}\alpha_1 \geq 0 \\ \alpha_2 \geq 0\end{array}\right)$ 2개

⟨One vs all⟩

D, C, H, M

하나의클래스

⇒ D, the others → SVM_D (not H)

C, the others → SVM_C (not H)

H, the others → SVM_H (H)

M, the others → SVM_M (not H)

Unknown image

2개 클래스니까

⟨비교⟩

D vs C　C vs H

D vs H　C vs M　⇒ 6번의

D vs M　H vs M　SVM 적용

SVM은 multi-class 처리 못함

⇒ 처리가능한 Many tricks이 존재한다.

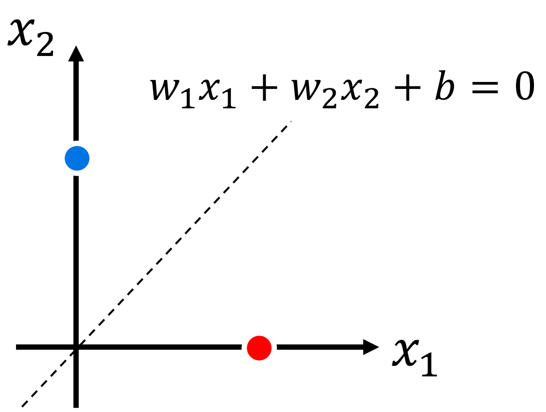
(ex one vs all 이라는 방법)

Dual Form for Learning LSVM

▪ Example

– What is the maximum margin boundary for $(0,1,1)$ and $(1,0,-1)$

$\alpha_1=\alpha_2$ 대입　$-\alpha_1^2+2\alpha_1$ ⇒ $\alpha_1=1$일때 최대



$$\underset{\alpha_1\alpha_2}{\operatorname{argmax}}\ \alpha_1+\alpha_2-\frac{1}{2}(\alpha_1^2+\alpha_2^2)$$

앞슬라이드 정리식

$$\text{subject to } \begin{cases}\alpha_1,\alpha_2\geq 0\\ \alpha_1-\alpha_2=0\end{cases}$$

Solution

⇨ $\alpha_1=\alpha_2=1$

$$\mathbf{w}=\sum_{i=1}^{n}\alpha_i y_i \mathbf{x}_i$$

choose any x_k

$b=y_k-\mathbf{w}\cdot\mathbf{x}_k$ for any $\mathbf{x_k}$ such that $\alpha_k>0$

k=1 적용

벡터 + 벡터 벡터

$$\mathbf{w}=1\cdot 1\cdot(0,1)+1\cdot(-1)\cdot(1,0)=(-1,1)$$

$$b=1-(-1,1)\cdot(0,1)=0$$

Boundary: $-x_1+x_2=0$