$$L(w,b) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \left(y_i \left(w^T x_i - b\right) - 1\right)$$

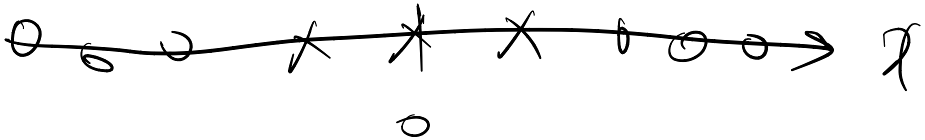
$$\begin{cases} \alpha_i \geq 0 \\ w = \sum \alpha_i y_i x_i \\ \sum \alpha_i y_i = 0 . \end{cases}$$

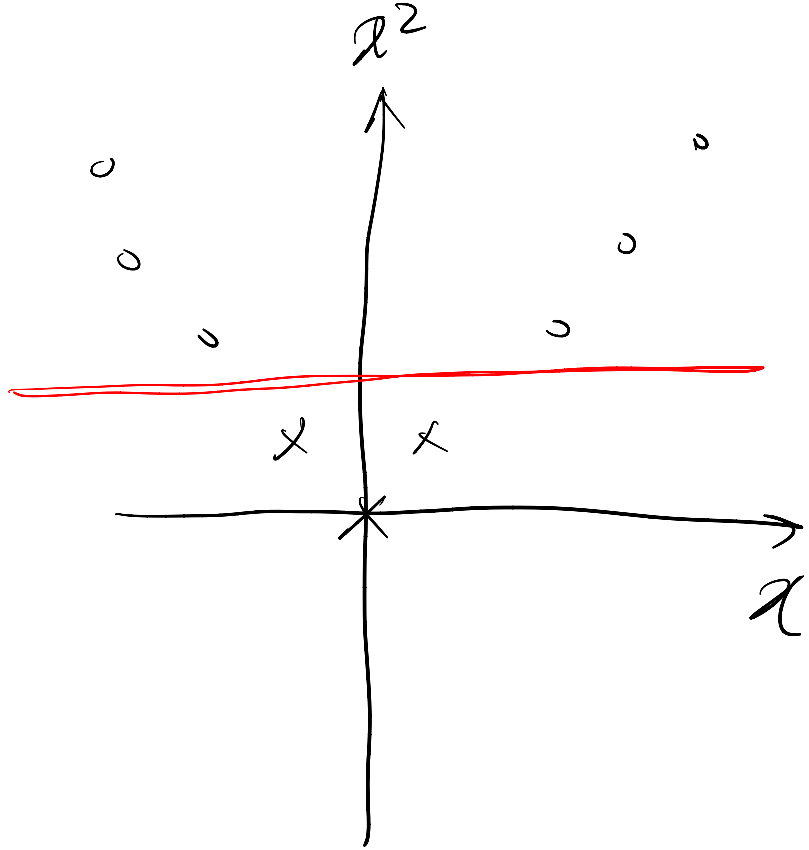
$$= \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \boxed{x_i^T x_j}$$

= 〃 − 〃 〃 〃 $\phi(x_i)^T \phi(x_j)$

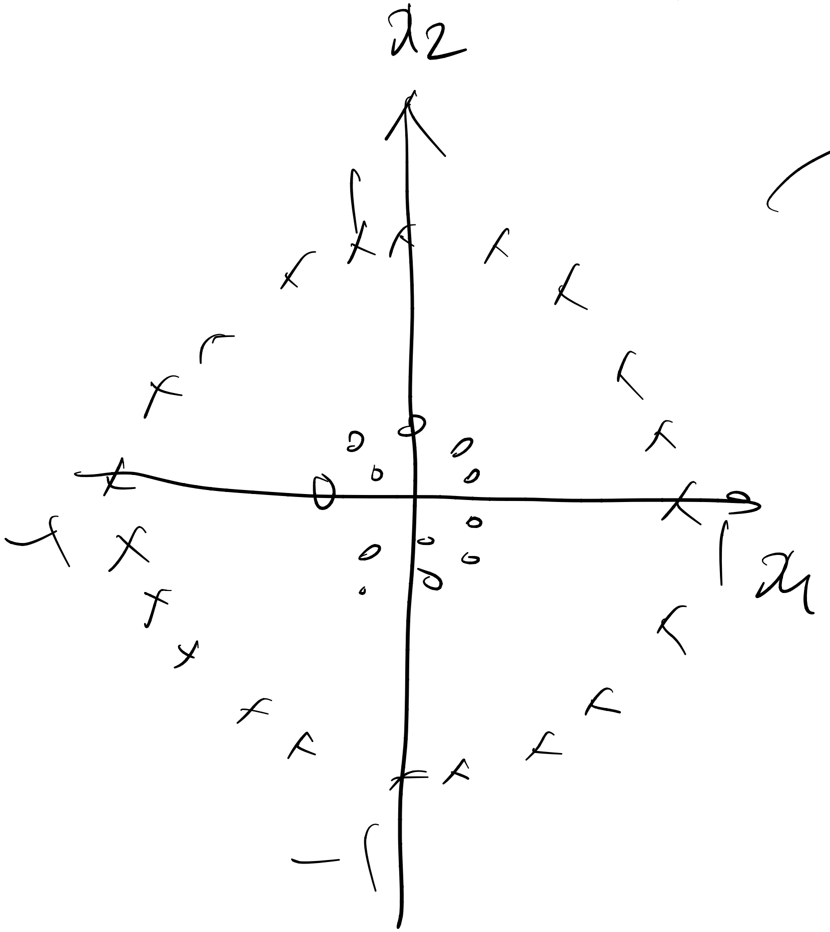
$$K(x_i, x_j)$$

$$x \to \{x, x^2\}$$



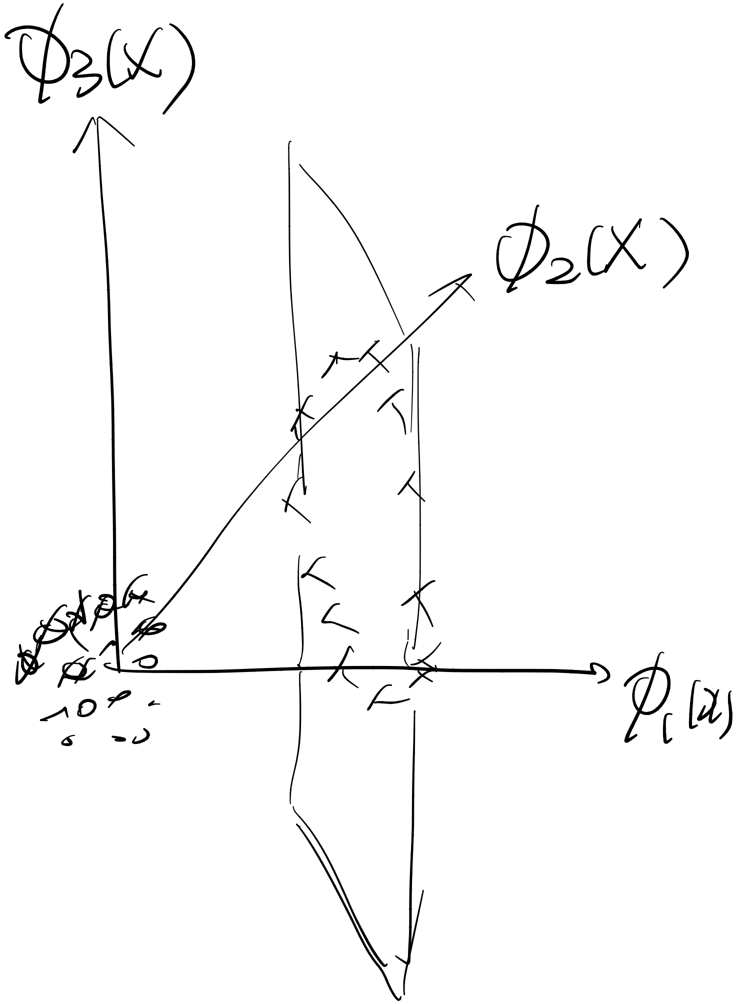


$$\phi(x) \in \mathbb{R}^3$$



$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\phi(x) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}\, x_1 x_2 \end{pmatrix}$$



Cover's theorem.

If D increases, the power of linear method increases.

Kernel Method.

$$\phi(x_1, x_2) = (x_1^2, x_2^2, \sqrt{2}x_1 x_2).$$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\phi(X)^T \phi(y) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}x_1 x_2 \end{pmatrix}^T \begin{pmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2}y_1 y_2 \end{pmatrix} = \begin{matrix} x_1^2 y_1^2 + x_2^2 y_2^2 \\ + 2x_1 x_2 y_1 y_2 \end{matrix}$$

$$= (x_1 y_1 + x_2 y_2)^2$$

$$= (X^T y)^2.$$

$$K(X, Y) = (X^T Y)^2 \in \mathbb{R}.$$

$$k(x,y) = (x^T y + 1)^d \in \mathbb{R}$$

ex) $d=2$. $x, y \in \mathbb{R}^2$.

$$k(x,y) = (x_1 y_1 + x_2 y_2 + 1)^2$$

$$= x_1^2 y_1^2 + x_2^2 y_2^2 + 1 + 2x_1 y_1 + 2x_2 y_2 + 2x_1 x_2 y_1 y_2$$

$$= \begin{pmatrix} x_1^2 \\ x_2^2 \\ 1 \\ \sqrt{2}\, x_1 x_2 \\ \sqrt{2}\, x_1 \\ \sqrt{2}\, x_2 \end{pmatrix}^T \begin{pmatrix} y_1^2 \\ y_2^2 \\ 1 \\ \sqrt{2}\, y_1 y_2 \\ \sqrt{2}\, y_1 \\ \sqrt{2}\, y_2 \end{pmatrix} = \phi(x)^T \phi(y)$$

Mercer's Theorem.

$k(x_i, x_j)$: Positive Definite Function.

$\Leftrightarrow \exists \phi(x)$ s.t. $\phi(x_i)^T \phi(x_j) = k(x_i, x_j)$

---

* Positive Definiteness of a function $k(x_i, x_j)$

For any $x_1, \cdots, x_N \in \mathbb{R}^p$

$$K = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_N) \\ & \ddots & & \\ & & \ddots & \\ & & & k(x_N, x_N) \end{pmatrix} \in \mathbb{R}^{N \times N}$$

if K is positive definite, then $k(x_i, x_j)$ is p.d.

* P.D. of a matrix M.

M is p.d $\iff x^T M x > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$

$\iff$ all of its eigenvalues are positive.

Support Vector Machines

$$L(\alpha_i) = \sum \alpha_i - \frac{1}{2} \sum \sum \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$w^T x' - b = \boxed{\sum_{i=1}^{N} \alpha_i y_i x_i^T x' - b \gtrless 0.}$$

$$= \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum \sum \alpha_i \alpha_j y_i y_j \underbrace{\phi^T(x_i)\,\phi(x_j)}_{= k(x_i, x_j)}.$$

$$\boxed{\sum \alpha_i y_i k(x_i, x') - b \gtrless 0}$$