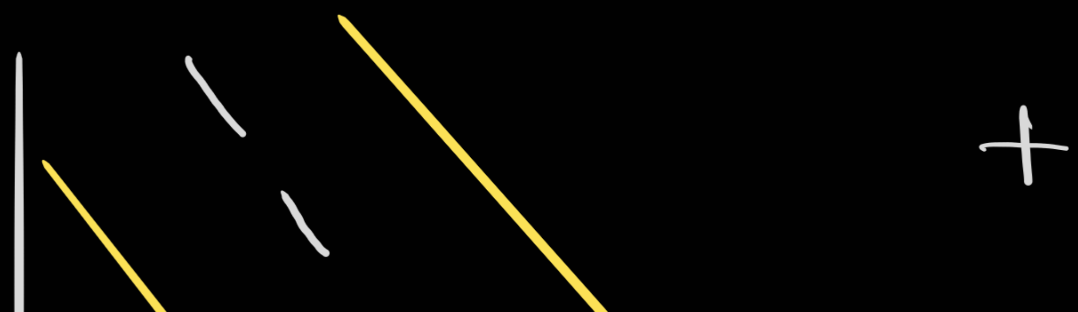
$\bar{w} \to$ Perpendicular vector to street

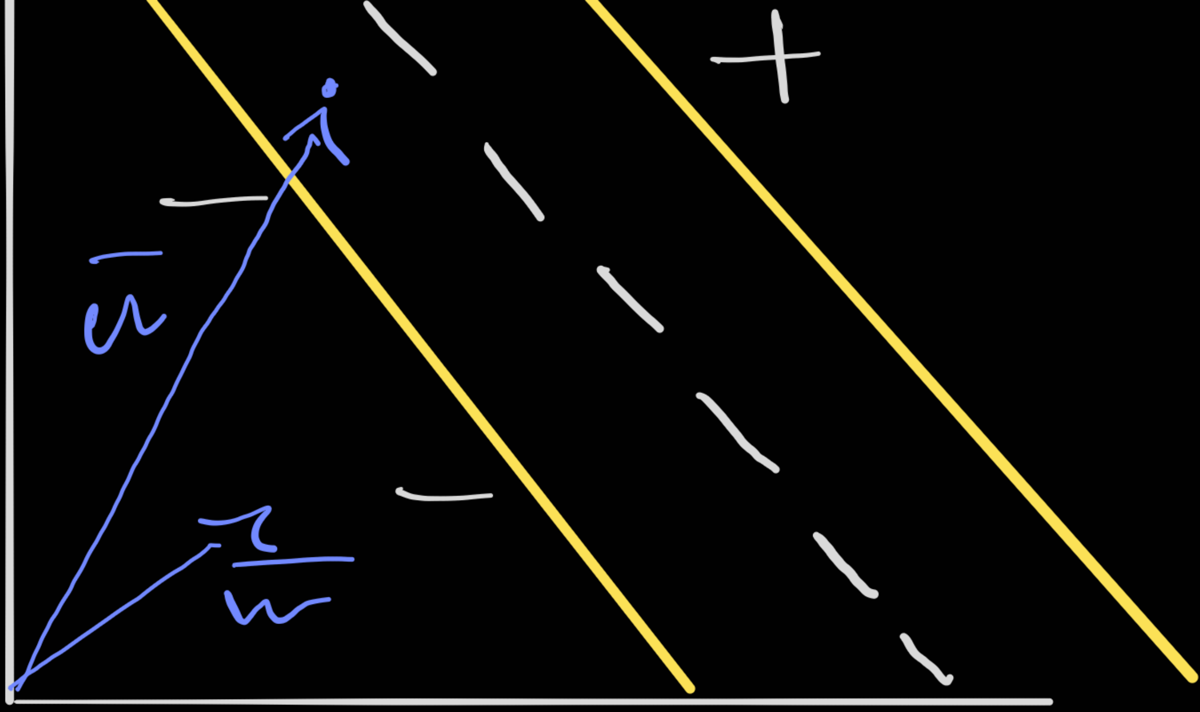
$\bar{u} \to$ vector to unknown data point

$\bar{w} \cdot \bar{u} \to$ projection of $\bar{u}$ onto $\bar{w}$

$$\boxed{\bar{w} \cdot \bar{u} + b \geq 0 \text{ then } + \text{ class}}$$

Decision Rule (1)





Finding $\bar{w}$ and b:

Positive sample

$$\left.\begin{aligned} \bar{w}\cdot\bar{x}_{+} + b &\geq 1 \\ w\cdot\bar{x}_{-} + b &\leq -1 \end{aligned}\right\} \Rightarrow *$$

Negative sample

introduce y_i:

$$y_i = \begin{cases} +1 & \text{for positive sample} \\ -1 & \text{for negative sample} \end{cases}$$

multiply by y_i

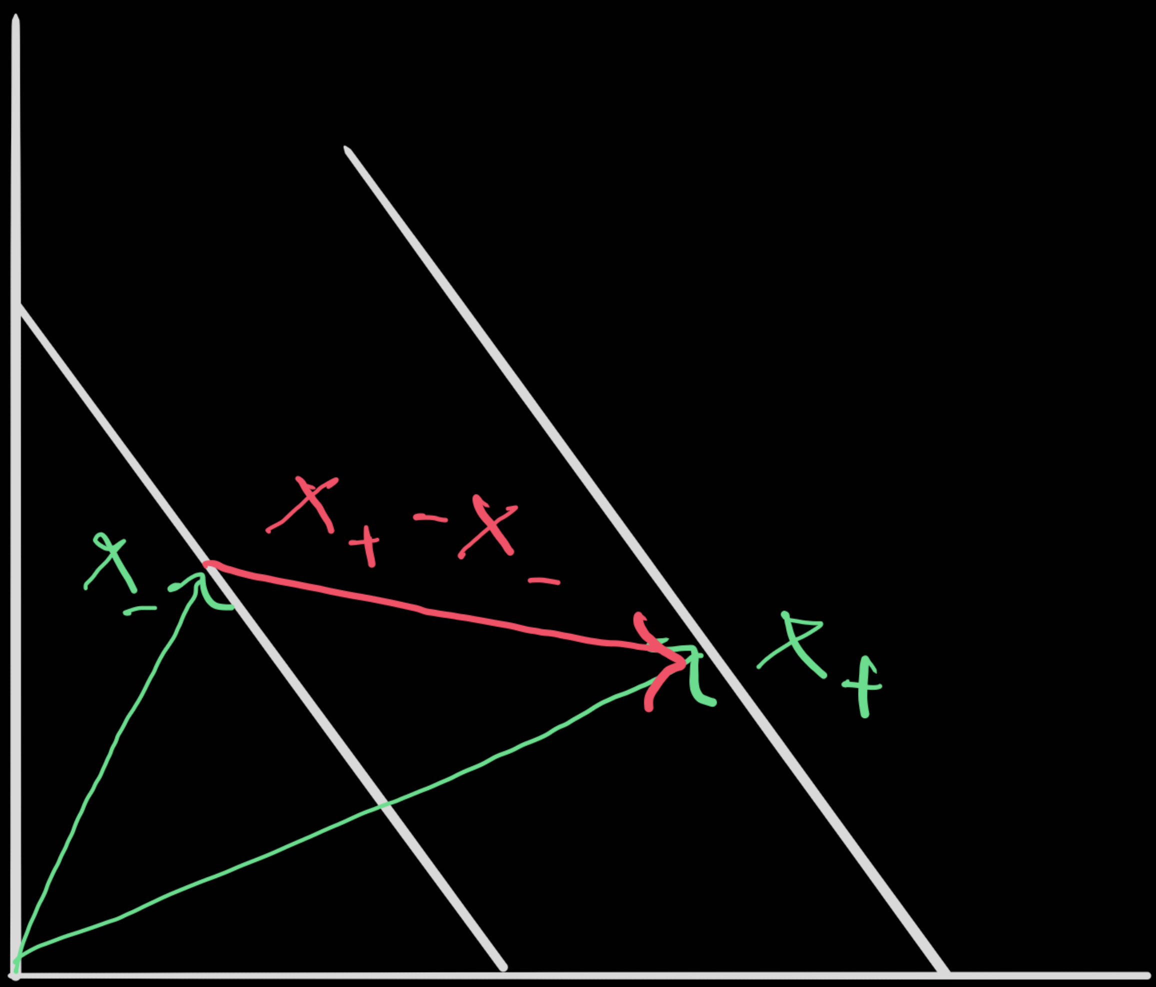
$$* \implies \begin{cases} y_i(\bar{w} \cdot \bar{x}_i + b) \geq 1 \\ \qquad \parallel \\ y_i(\bar{w} \cdot \bar{x}_i + b) \geq 1 \end{cases}$$

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

* $y_i($

$* \; y_i(w \cdot x_i + b) - 1 = 0 \quad (2)$

Exactly on the line



unit vector

$$\text{width} = (\bar{x}_+ - \bar{x}_-) \cdot \left(\frac{\bar{w}}{\|w\|} \right)$$

(2) (2) $(\|w\|)$

$1-b + 1+b$

$$= \frac{2}{\|w\|} \quad 2$$

$$\text{MAX}\left(\frac{2}{\|w\|}\right)$$

$$\iff \text{MAX}\left(\frac{1}{\|w\|}\right)$$

$$\iff \text{MIN}(\|w\|)$$

$$\iff \text{MIN}\left(\frac{1}{2}\|w\|^2\right)$$

$\Leftrightarrow \text{MIN}\left(\frac{1}{2}\|w\|\right)$

(mathematical convenience)

(easier derivative)

Using Lagrange

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

we have to find the extremum

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

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

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

$$\frac{\partial L}{\partial b} = -\sum \alpha_i y_i = 0 \qquad (3)$$

$$\Rightarrow \boxed{\sum \alpha_i y_i = 0}$$

$$\Rightarrow L = \frac{1}{2}\left(\sum \alpha_i y_i \bar{x}_i\right) \cdot \left(\sum \alpha_j y_j \bar{x}_j\right)$$

$$- \left(\sum \alpha_i y_i \bar{x}_i\right) \cdot \left(\sum \alpha_j y_j x_j\right)$$

$$- \underbrace{\cancel{\sum_i \alpha_i y_i b}}_{0} + \sum_i \alpha_i$$

$$L = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \qquad (4)$$

** Optimization depends only on the dot-product of pairs of samples.

New Decision Rule

$$\left(\sum \alpha_i y_i \bar{x}_i \cdot \bar{u}\right) + b \geq 0 \rightarrow + \text{ class}$$

---

* Optimization never gets stuck in local extrema because it creates a Convex.

If samples are not linearly separable then we bring them into another space.

↳ Transformation → $\Phi(\bar{x})$

$\Downarrow$

** $\Rightarrow$ we only need to optimize $\varphi(\bar{x}_i) \cdot \varphi(\bar{x}_j)$

with new decision rul

$$\varphi(x_i) \cdot \varphi(u)$$

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$$

$\Downarrow$

we only need

this function K

called the <u>Kernel Trick</u>

(we don't need the transformation $\phi(\bar{x})$)

$K \rightarrow$ Polynomial

$K \rightarrow$ RBF