Kernel

It is a function to mapp the samples into high-dimensional feature space.

Non-linear SVM

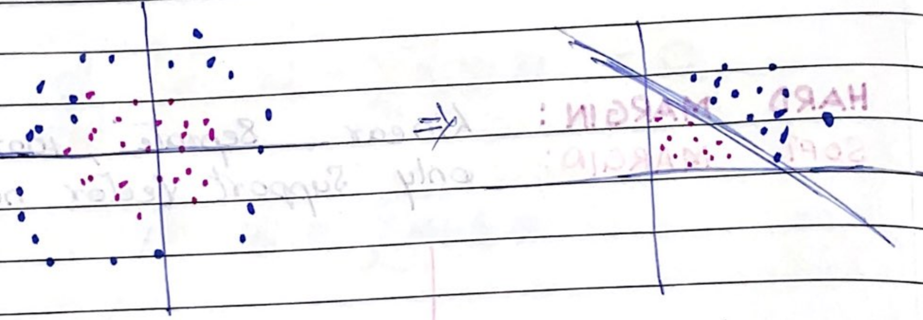
$\phi : X \rightarrow \phi(x)$ (Higher dimensional)

Computational cost :

$\sum \alpha_i \alpha_j x_i x_j y_i y_j$

↳ $\sum \alpha_i \alpha_j \underline{(\phi x_i)(\phi x_j)} y_i y_j$

⇒



⇒ $x_i \quad x_j$

$\phi(x_i) \quad \phi(x_j)$

→ $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$

Kernel functions

$g(x) = w_i \phi(x) + b$

$= \sum \underbrace{\phi(x_i)\,\phi(x_j)}_{K(x_i, x_j)} + b$

DATE

$X = [x_i, x_2]$

$K(x_i, x_j) = (1 + x_i \cdot x_j)^2$

$x_i = (x_{i1}, x_{i2})$

$x_j = (x_{j1}, x_{j2})$

$\Rightarrow 1 + x_{i1}^2 x_{j1}^2 + 2 x_{i1} x_{j1} x_{i2} x_{j2} + x_{j2}^2 x_{j1}^2$

$+ 2 x_{i1} x_{j1} + 2 x_{i2} x_{j2}$

$\Rightarrow \phi(x_i) = (1, x_{i1}^2, \sqrt{2} x_{i1} x_{i2}, x_{i2}^2, \sqrt{2} x_{i1}, \sqrt{2} x_{i2})$

$w(\alpha) = \sum_{i_2} \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j K(x_i, x_j)$

$(1 + x_i \cdot x_j)^2$