SUPPORT VECTOR MACHINE

$w^T x + b = 0$ HYPERPLANE WHICH DIVIDES TWO SETS A AND B

FIND THE PLANE WITH THE MAXIMUM MARGIN OF SEPARATION

$$(P) \begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 \\ w^T x^i + b \geq 1 & \forall x^i \in A \\ w^T x^i + b \leq -1 & \forall x^i \in B \end{cases} \equiv \begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 \\ 1 - y^i (w^T x^i + b) \leq 0 & \forall i = 1, \dots \ell \end{cases}$$

(P) IS A CONVEX QUADRATIC PROBLEM, IF (w^*, b^*, λ^*) SOLVES THE KKT SYSTEM $\Rightarrow (w^*, b^*)$ IS A GLOBAL MINIMUM

$$(D) \begin{cases} -\min_{\lambda} \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y^i y^j (x^i)^T x^j \lambda^i \lambda^j - \sum_{i=1}^{\ell} \lambda^i \\ \sum_{i=1}^{\ell} \lambda^i y^i = 0 \\ \lambda \geq 0 \end{cases} \qquad q_{ij} = y^i y^j (x^i)^T x^j$$

IF λ^* IS A KKT MULTIPLIER ASSOCIATED TO THE OPTIMUM OF (P) (w^*, b^*) $\Rightarrow \lambda^*$ IS A DUAL OPTIMUM

IF $\lambda_i^* > 0$ THEN x^i IS A SUPPORT VECTOR

IF λ^* IS A DUAL OPTIMUM THEN $w^* = \sum_{i=1}^{\ell} \lambda_i^* y^i x^i$

TAKEN $\lambda_i^* > 0 \Rightarrow b^* = \frac{1}{y^i} - (w^*)^T x^i$

IF A, B ARE NOT LINEARLY SEPARABLE $\Rightarrow$

$$\begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{\ell} \varepsilon_i \\ 1 - y^i (w^T x^i + b) \leq \varepsilon_i & \forall i = 1, \dots \ell \\ \varepsilon_i \geq 0 \end{cases} \qquad C > 0$$

(D) BECOMES

$$\begin{cases} -\min_{\lambda} \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y^i y^j (x^i)^T x^j \lambda^i \lambda^j - \sum_{i=1}^{\ell} \lambda^i \\ \sum_{i=1}^{\ell} \lambda^i y^i = 0 \\ 0 \leq \lambda_i \leq C & i = 1, \dots \ell \end{cases}$$

KERNEL FUNCTIONS

$\phi : \mathbb{R}^n \to \mathcal{H}$

$$\begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{\ell} \varepsilon_i \\ 1 - y^i (w^T \phi(x^i) + b) \leq \varepsilon_i & \forall i = 1, \dots \ell \\ \varepsilon_i \geq 0 \end{cases} \qquad C > 0$$

$$\begin{cases} -\min\limits_{\lambda} \frac{1}{2} \sum\limits_{i=1}^{\ell} \sum\limits_{j=1}^{\ell} y^i y^j \underbrace{\phi(x^i)^T \phi(w^j)}_{K(x^i, x^j)} \lambda^i \lambda^j - \sum\limits_{i=1}^{\ell} \lambda^i \\ \sum\limits_{i=1} \lambda^i y^i = 0 \\ 0 \leq \lambda_i \leq C \qquad i = 1, \dots \ell \end{cases}$$

$$b^* = \frac{1}{y^i} - \sum_{j=1}^{\ell} \lambda_j^* y^j K(x^i, x^j)$$

DECISION FUNCTION $\quad f(x) = \text{sign}\left((w^*)^T x + b^* \right)$

$$f(x) = \text{sign}\left(\sum_{i=1}^{\ell} \lambda_i^* y^i \, \text{kernel}(x^i, x) + b^* \right)$$