Primal SVM Formulation

For hard margin -

$$\min_{w} \frac{w^T w}{2} \quad \text{s.t. } y_i(w^T x_i + b) \geq 1$$

Dual Formulation -

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{n} \alpha_i \left(y_i \left[w^T x_i + b\right] - 1\right)$$

for every i

$$\max_{\alpha_i \geq 0} \min_{w, b} L(w, b, \alpha)$$

Stationary constraints

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$\alpha_i = 1$ only for s.v. , 0 for everything else

($\because$ w consists of only s.v)

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum \alpha_i y_i = 0$$

<u>Dual</u>

$$\max_{\alpha} J(\alpha) = \sum_{i=1}^{N} \alpha_i - \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

no. of data pts << no. of features

svm formulation with Kernels

$$\max_{\alpha} J(\alpha) = \sum_{k=1}^{N} \alpha_i - \sum_{i,j} \alpha_i \alpha_j y_i y_j \underbrace{x_i^T x_j}_{\text{Kernel}(x_i, x_j)}$$

s.t.

$$\sum \alpha_i y_i = 0$$

$$c \geq \alpha_i \geq 0$$

Case 1 : $x_1 \rightarrow \phi(x_1)$
$x_2 \rightarrow \phi(x_2)$

computationally complex to transfer every data pt. to high dim.

Case 2 :

Aim : $\phi(x_1)^T \cdot \phi(x_2)$

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

We don't transform entire data to higher dim.

$\cancel{\phi(x_i) \cdot \phi(x_j)}$: Transform first & then inner product in ϕ space

Instead calc. inner product in original space & then transform using kernel to inner product in new space.

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