# Assignment - II

CS19BTECH11039
SRI HARI . M

1. SVM derivation:

Initially the margin boundaries are defined as

$w \cdot x + b = \pm 1$

Now let's replace 1 with constant $\gamma$, where $\gamma > 0$

The margin boundaries looks as $w \cdot x + b = \pm \gamma$

Intuition:

1, Firstly, SVM creates a hyperplane.

2, Now to construct this hyperplane, it takes help of other support vectors.

3, These are parallel and equidistant from each other. Now, the hyperplane becomes optimum hypuplane.

So, now that optimum hypuplane is created, we need to maximise the distance between support vectors to create max-margin hypesplane.

clearly, in either case, $w \cdot x + b = 0$ will be our hyperplane.

Margin distance = $\max\left(\frac{2\gamma}{|w|}\right)$

as $\gamma$ is constant $\rightarrow$ max(margin distance) $\Rightarrow \min(|w|)$

Which is same as in case of $w \cdot x + b = \pm 1$

For mathematical convinience we'd write

max (Margin Distance) = $\min\left(\frac{1}{2}|w|^2\right)$

which is same for both, $w \cdot x + b = \pm 1$ and $w \cdot x + b = \pm \gamma$

but only the inequality contraint changes to

$y_i(w \cdot x_i + b) \geq \gamma$ where $\gamma > 0$

2, Given $\rho = \frac{1}{|W|}$

Required to show : $\frac{1}{\rho^2} = \sum_{i=1}^{N} \alpha_i$

We have seen the optimal conditions in class that,

$$W = \sum_j \alpha_j y_j x_j \quad \text{and}$$

$$\sum_j a_j y_j = 0 .$$

The Lagrangian addition to the primal is,

$$\Rightarrow \alpha_i - \alpha_i y_i (w \cdot x_i + b)$$

$$\Rightarrow \alpha_i - b\alpha_i y_i - \alpha_i y_i (w \cdot x_i)$$

For margins, we know that $y_i (w \cdot x_i + b) = 1$

$$\Rightarrow (w \cdot x_i) \cdot \alpha_i y_i = \alpha_i - b\alpha_i y_i$$

Summing upon $i$,

$$\Rightarrow \sum_{i=1}^{N} (w \cdot x_i) \cdot \alpha_i y_i = \sum_{i=1}^{N} \alpha_i - b \sum_{i=1}^{N} \alpha_i y_i$$

Since $W = \sum_j \alpha_j y_j x_j$

$$(W \cdot W) = \sum_{i=1}^{N} \alpha_i - b \sum_{i=1}^{N} \alpha_i y_i$$

Since $\sum \alpha_i y_i = 0$

$$|W|^2 = \sum_{i=1}^{N} \alpha_i$$

We have $|W| = \frac{1}{\rho} \Rightarrow |W|^2 = \frac{1}{\rho^2}$

$$\therefore \frac{1}{\rho^2} = \sum_{i=1}^{N} \alpha_i$$

3, Given $K_1, K_2$ are valid kernel functions.

Required to check validity of some kernels.

a, $K(x,z) = K_1(x,z) + K_2(x,z)$

We can write $K_1(x,z) = \langle \phi_1(x), \phi_1(z) \rangle$

Similarly $K_2(x,z) = \langle \phi_2(x), \phi_2(z) \rangle$

$K_1(x,z) + K_2(x,z) = \langle \phi_1(x), \phi_1(z) \rangle + \langle \phi_2(x), \phi_2(z) \rangle$

$= \langle [\phi_1(x)\ \phi_2(x)], [\phi_1(z)\ \phi_2(z)] \rangle$

It is in the form of inner product.

Let's say $\phi' = \phi_1 \cdot \phi_2$ and it can replace the equation of $x, z$.

So it's a valid kernel.

b, $K(x,z) = K_1(x,z)\ K_2(x,z)$

It is trivial that the matrix $K$ is product of matrices $K_1$ and $K_2$.

Since $K_1$ and $K_2$ are valid, $K$ is also valid.

So it's a valid kernel.

c, $K(x,z) = h(K_1(x,z))$

Given $h$ has all positive co-efficients, and is a polynomial function.

from (b), multiplication gives valid kernel.

from (a), addition also gives valid kernel.

So the polynomial function also gives valid kernels.

So, it's a valid kernel.

d, $K(x,z) = \exp(K_1(x,z))$

We know from taylor's theorem that,

$e^x = 1 + x + \frac{x^2}{2!} + \cdots$

Replacing $x$ with $K_1(x,z)$

the equation turns out to be polynomial.

from c,, polynomial function yields valid, kernels,

$\exp(K_1(x,z))$ is a valid Kernel.

e, $K(x,z) = \exp\left(-\frac{\|x-z\|^2}{\sigma^2}\right)$

$= \exp\left(-\frac{|x|^2 - |z|^2 + 2x^Tz}{\sigma^2}\right)$

$= \exp\left(-\frac{|x|^2}{\sigma^2}\right) \cdot \exp\left(-\frac{|z|^2}{\sigma^2}\right) \cdot \exp\left(\frac{2 \cdot x^T \cdot z}{\sigma^2}\right)$

$= K \cdot \exp\left(\frac{2 \cdot x^T \cdot z}{\sigma^2}\right)$

From b, multiplication yields valid kernel,

From d, exponentiation yields valid kernel.

So it's a valid kernel.