Sm21mtech14001 - Theory Questions

1. Let the equation of the margin boundary be $\Big\}$ $w^T x + b = \gamma$

$$w^T x + b = \gamma$$

$$\Rightarrow \frac{w^T}{\gamma} x + \frac{b}{\gamma} = 1$$

Let $\frac{1}{\gamma}$ be a constant 'c'

$$\Rightarrow c w^T x + cb = 1$$

In general, Now, distance of a point to margin $w^T x + b = 1$ is given as $\Big] = \frac{y(w^T x + b)}{\|w\|}$

Distance of a point to margin $c w^T x + cb = 1$ is given as,

$$d = \frac{y(c w^T x + cb)}{\|cw\|}$$

$$d = \frac{c\, y(w^T x + b)}{c\|w\|}$$

$$\Rightarrow \boxed{d = \frac{y(w^T x + b)}{\|w\|}}$$

Hence irrespective of value of γ, the solution for max. margin hyper plane is unchanged

2.

From dual equation,

$$\max_{\alpha \geq 0} \ \min \ \frac{1}{2} \|W\|^2 - \sum_j \alpha_j \left[(\vec{w} \cdot \vec{x_j} + b) y_j - 1 \right]$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow W = \sum \alpha_j y_j x_j \rightarrow ①$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_j \alpha_j y_j = 0 \longrightarrow ②$$

Now, according to constraint,

$$y_i (w x_i + b) - 1 \geq 0$$

at the margin $y_i (w x_i + b) = 1$

$$(w x_i + b) = \frac{1}{y_i} = y_i \quad \left[\because \frac{1}{y_i} = y_i \text{ as } y_i = \pm 1 \right]$$

$$\Rightarrow b = y - w x_i$$

$$\Rightarrow b = y_i - \sum_j \alpha_j y_j x_j^T x_i$$

Multiply by $\sum \alpha_i y_i$ on both sides

$$\sum_i \alpha_i y_i b = \sum \alpha_i y_i \left(y_i - \sum_j \alpha_j y_j x_j^T x_i \right)$$

$$\boxed{\sum \alpha_i y_i b = \sum \alpha_i y_i^2 - \sum_{ij} \alpha_i \alpha_j y_i y_j x_j^T x_i}$$

$$y_i^2 = 1 \text{ as } y_i = \pm 1 \rightarrow ③$$

from ② $\sum \alpha_i y_i = 0 \rightarrow ④$

and WKT $W = \sum \alpha_i y_i x_i \rightarrow (5)$

sub (3), (4) & (5) in (I).

$$0 = \sum \alpha_i - \cancel{\tfrac{1}{2}} \|W\|^2$$

$$\sum_i \alpha_i = \|W\|^2$$

WKT $\rho^2 = \frac{1}{\|W\|^2}$

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

Q3.) check if valid

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

K is a valid kernel if $X^T K X \geq 0$

$$X^T [K(X, Z)] X = X^T (K_1(X, Z) + K_2(x, z)) X$$

$$= X^T K_1 (x, z) X + X^T K_2(x, z) X$$

$$\geq 0$$

Hence $K(x, z) = K_1(x, z) + K_2(x, z)$ is also

(I) a kernel $\rightarrow$ (1)

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

$$X^T K X = X^T K_1(x,z)\, K_2(x,z) \cdot X$$

$$= K_1(x,z)\, X^T K_2(x,z)\, X$$

$$= K_1(x,z) \cdot K_3'(x,z)$$

$$= \text{Trace}(K_1 K_3') \text{ which is } \geq 0$$

$\therefore K(x,z) = K_1(x,z) \cdot \cancel{K}_2(x,z)$ is also a valid kernel $\rightarrow$ 2

(c) $K(x,z) = h(K_1(x,z))$ where h is a polynomial function with positive co-efficients

Since h is a polynomial function, it could result in sum and product of kernels which all are greater than zero (from above two proofs) $\because$ Hence

$K(x,z) = h(K_1(x,z))$ is also a kernel

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

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$\cancel{\exp}\ K_1(x,z))$ includes sum and product of kernels.

which are valid kernels from 1 and 2. Hence

$\exp(K_1(x,z))$ is also a valid kernel $\longrightarrow$ 3

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

RHS

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

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

$$= g(x)\, g(z) \cdot \exp(K_1(x, z))$$

$g(x) \cdot g(z)$ is a kernel from 2

$\exp(K_1(x, z))$ is kernel from 3

$\therefore$ Product of 2 & 3 is also kernel, $\therefore \exp\left(-\frac{\|x - z\|^2}{\sigma^2}\right)$

is also a valid kernel