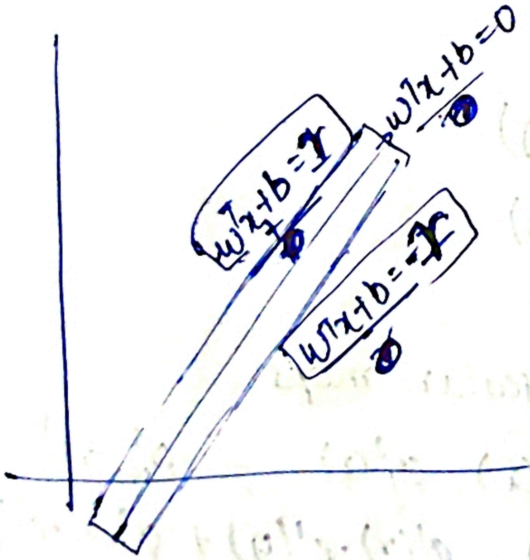

Since $w^Tx+b=0$ & $c(w^Tx+b)=0$ define the same plane, we have freedom to choose normalization.

As given in question we are taking $w^Tx+b=\gamma$ & $w^Tx+b=-\gamma$ for positive & negative support vectors respectively

⇒ Distance from example to the seperator is $r = y\frac{w^Tx+b}{|w|}$

$y=\pm 1$

$W^T x + b = \gamma$; $W^T x + b = -\gamma$.

$\Rightarrow \frac{W^T}{\gamma} x + \frac{b}{\gamma} = 1$; $\frac{W^T}{\gamma} x + \frac{b}{\gamma} = -1$.

W^T & b just gets divided by γ it will be the same equation.

now the $\frac{W^T}{\gamma} = W$ $\quad \frac{b}{\gamma} = B$.

So now the new equations are $W x_+ + B = 1$; $W x_- + B = -1$

distances from example to the seperator is $\gamma = y \frac{W x + B}{|W|}$

$$\Rightarrow +1 \frac{W x_+ + B}{|W|} + (-1) \frac{W x_- + B}{|W|} = \frac{1 + (-1)(-1)}{|W|}$$

$$= \frac{2}{|W|}$$

Since we don know the value w, b dividing it by γ just normalizes the equation it does not change the margin.

(2) * $L_P = \frac{1}{2}\|w\|^2 - \sum \alpha_i((w^T x_i + b)y_i - 1)$.

from KKT conditions.

$\frac{dL_P}{d\alpha} = 0.$ ★

$\Rightarrow \boxed{[(w^T x_i + b)y_i - 1)] = 0}$

$\frac{\partial L}{\partial w} \Rightarrow w = \sum \alpha_i y_i x_i \rightarrow (1)$

$\frac{\partial L}{\partial b} \Rightarrow \sum \alpha_i y_i = 0. \rightarrow (2)$.

Substituting (1) & (2) in L_P we get L_D.

$\Rightarrow L_D = -\frac{1}{2}\|w\|^2 - w^T \sum \alpha_i y_i x_i + b \sum \alpha_i y_i + \sum \alpha_i$.

$L_D = \frac{\|w\|^2}{2} - \|w\|^2 + b(0) + \sum \alpha_i$.

In SVM strong duality holds

For strong duality $L_P = L_D$.

From ★.

$L_P = \frac{1}{2}\|w\|^2 - \sum \alpha_i(0) = \frac{1}{2}\|w\|^2$.

$\frac{\|w\|^2}{2} = \frac{-\|w\|^2}{2} + \sum \alpha_i$

Given $\|w\| = 1/\rho$.

$\Rightarrow \sum \alpha_i = \|w\|^2$

$\Rightarrow \boxed{\sum \alpha_i = 1/\rho^2} \longrightarrow$ Hence proved.

3 a. $k(x,z) = k_1(x,z) + k_2(x,z)$

let $\phi^1(x) = (\phi_1^1(x) \; --- \; \phi^1_{N_1}(x))$

$\phi^2(x) = (\phi_1^2(x) \; --- \; \phi^2_{N_2}(x))$.

Be the feature maps for k_1 & k_2

Defining $\phi(x)$ by concatenating the feature maps.

$\phi(x) = ((\phi_1^1(x) \; --- \; \phi^1_{N_1}(x), \phi_1^2(x) \; --- \; \phi^2_{N_2}(x))$

The mapping clearly satisfies $\phi(x)\cdot\phi(y) = \phi^1(x)\cdot\phi^1(y) + \phi^2(x)\cdot\phi^2(y)$

Obey kernel trick.

---

b. $k(x,z) = k_1(x,z)\, k_2(x,z)$

$\phi^1(x) = (\phi_1^1(x) \; --- \; \phi^1_{N_1}(x))$.

$\phi^2(x) = (\phi_2^2(x) \; --- \; \phi^1_{N_2}(x))$

$\phi^1(x) \times \phi^2(x) = \phi(x)$.

Here ~~k_1 & k_2~~ we are multiplying ϕ expressions for k_1 & k_2. to see that k kernels with the space products of features from ϕ^1 & ϕ^2

→ Obeys kernel trick.

---

c. $k(x,z) = h(k_1(x,z))$ h is a polynomial function.

Since each polynomial term is product of kernels with a & b we can say that $h(k_1(x,z))$ is a positive definite kernel function.

---

d.

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

We have $\exp(x) = \lim\limits_{i \to \infty} (1 + \cdots x^i/i!)$ The proof follows from c.

& the fact that $\boxed{K(x,z) = \lim\limits_{i \to \infty} K_i(x,z)}$

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

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

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

$$= h(x)\, h(z) \exp(K_1(x,z))$$

Here there's just one feature defined by $h()$. (x,z). from.

(d) we can say $\exp(K_1(x,z))$ is kernel function.

Hence the whole function is valid.