C. Saurav
ES16BTECH11007

1A) Given, k_1, k_2 are valid kernels

$$\phi_1(x) = [\phi_1^1(x) \cdot \phi_2^1(x) - \phi_n^1(x)] \rightarrow \text{feature representation of kernel 1}$$

$$\phi_2(x) = [\phi_1^2(x)\ \phi_2^2(x)\ - \phi_n^2(x)] \rightarrow \text{feature representation of kernel 2}$$

a)

$$k_1(x,z) = \langle \phi_1(x), \phi_1(z) \rangle$$

$$k_2(x,z) = \langle \phi_2(x), \phi_2(z) \rangle$$

$$k(x,z) = k_1(x,z) + k_2(x,z)$$

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

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

we can see that the above representation of the kernel function is a valid dot product and hence it is a valid kernel

b) $k_1(x,z) = \langle \phi_1(x), \phi_1(z) \rangle = \sum_{i=1}^{k} \phi_{1i}(x) \cdot \phi_{1i}(z)$

$k_2(x,z) = \langle \phi_2(x), \phi_2(z) \rangle = \sum_{j=1}^{k} \phi_{2j}(x) \cdot \phi_{2j}(z)$

$k(x,z) = k_1(x,z) \cdot k_2(x,z)$

$$= \left[\sum_{i=1}^{k} \phi_{1i}(x) \cdot \phi_{1i}(z)\right]\left[\sum_{j=1}^{k} \phi_{2j}(x)\ \phi_{2j}(z)\right]$$

$$= \langle \phi_1(x) \times \phi_2(x), \phi_1(z) \times \phi_2(z) \rangle + -$$

$$- + \langle \phi_{1k}(x) \times \phi_{2k}(x), \phi_{1k}(z) \times \phi_{2k}(z) \rangle$$

$$= \langle \left[\left(\sum_{i=1}^{k} \phi_{1i}(x)\right) \cdot \phi_2(x) \right], \left[\sum_{j=1}^{k} \phi_{1j}(z)\ \phi_{2j}(z) \right] \rangle$$

Since the product of kernel functions can be expressed as an inner product, it is a valid kernel function

c) $k(x,z) = h(k_1(x,z))$

$h \to$ polynomial function

Since it will be a linear sum of different powers of kernels [the linear sum of valid kernels is a valid kernel and product of valid kernels is also a valid kernel] it will be a valid kernel.

d) $k(x,z) = \exp(k_1(x,z))$

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{\lfloor 1} + \frac{x^2}{\lfloor 2} + \cdots + \frac{x^n}{\lfloor n}\right)$$

Sub $k_1(x,z)$ in x.

$$k(x,z) = \lim_{n\to\infty} \left(1 + k_1(x,z) + \frac{(k_1(x,z))^2}{2} + \cdots + \frac{(k_1(x,z))^n}{\lfloor n}\right)$$

we know that the polynomial expression inside the exponential is a valid kernel from the 'c' part of the question.

So as the no. of terms increase, it will remain one more polynomial function and polynomial of a valid kernel function is a valid kernel function

hence it is a valid kernel

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

by expanding the squared term inside exponential

$$k(x,z) = \exp\left(-\frac{1}{\sigma^2} x^T x\right) . \exp(x^T z) . \exp\left(-\frac{1}{\sigma^2} . z^T z\right)$$

sub $k_1(x,z)$ in place of 'x'

from 'd' subpart of this question, we know that exponential of a valid kernel function is a valid kernel function

and product of kernel functions is a valid kernel function and $k(x,z)$ is an inner product -

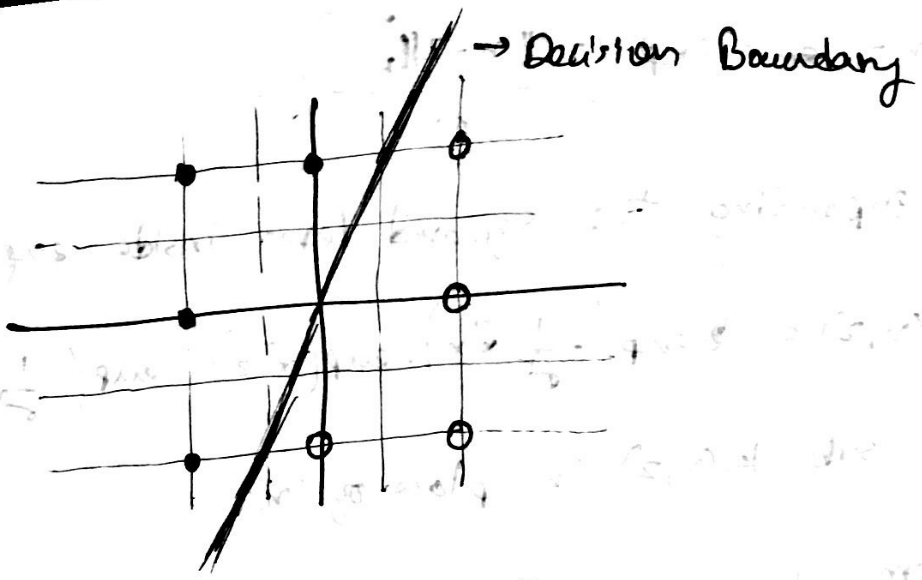
hence the given kernel function is a valid kernel function

$$k(x,z) = \underbrace{\exp\left(-\frac{1}{\sigma^2} x^T x\right)}_{\underset{+ve}{f(x)}} . \underbrace{\exp(x^T z)}_{\exp(x^T z)} . \underbrace{\exp\left(-\frac{1}{\sigma^2} z^T z\right)}_{\underset{+ve}{f(z)}}$$

$k(x,z)$ is a valid kernel function

2A)

a)



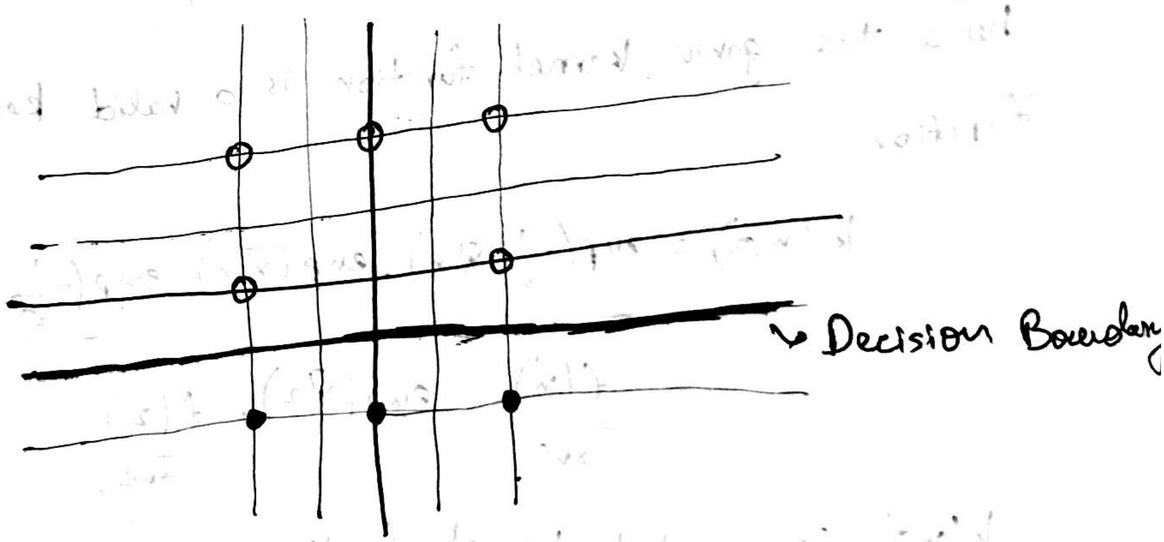
$w_1 x + w_2 y + c = 0$

line passing through origin $c = 0$
bias term

$w_1 = -2, \; w_2 = 1$

weights are $[-2, 1]$

b)



decision boundary ⇒ $y = -1$.

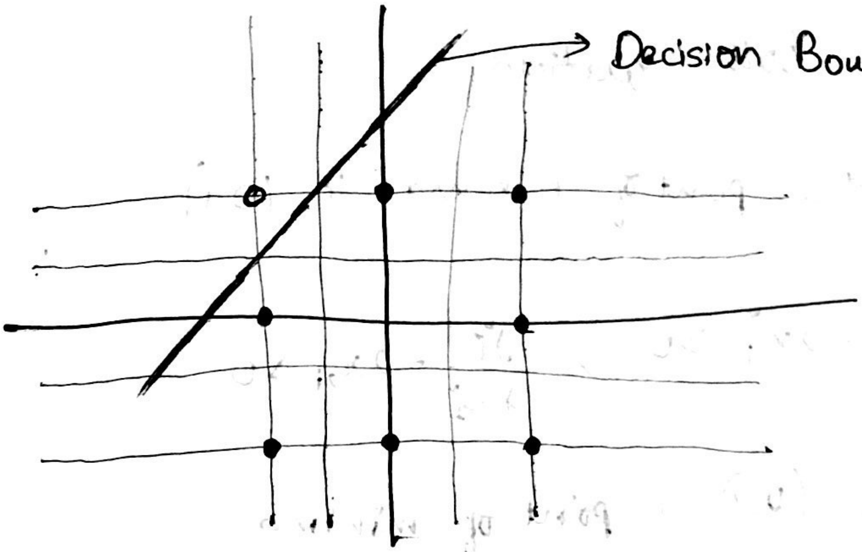
bias $= 1$.

$w_1 = 0$

$w_2 = 1$ weights $= [0 \; 1]$

Scanned with CamScanner

c)



$w_1 x + w_2 y + c = 0$

$x - y + 3 = 0 \qquad w_1 = -1, \; w_2 = 1$

bias $= 3$ weights $= [-1, 1]$

3A) given $x_1 = [1, 1) \qquad x_2 = [1, -1]$

$y_1 = 1 \qquad y_2 = -1$

$\hat{y}(x_i) = w_1 x_{1i} + w_2 x_{2i}$ and $\hat{y}$ is the predicted class

error (E) $= \frac{1}{2} \sum_{i=1}^{2} \left(y_i - (w_1 x_{1i} + w_2 x_{2i})\right)^2$

$\frac{\partial E}{\partial w_1} = 0 \quad , \quad \frac{\partial E}{\partial w_2} = 0$

$\Rightarrow \sum_{i=1}^{2} (y_i - w_1 x_{1i} - w_2 x_{2i})(-x_{1i}) = 0$

and

$\sum_{i=1}^{2} (y_i - w_1 x_{1i} - w_2 x_{2i})(-x_{2i}) = 0$

$\Rightarrow (1 - w_1 - w_2) \times 1 + (-1 - w_1 \times 1 + w_2) \times 1 = 0$

$(1 - w_1 - w_2) \times 1 + (-1 - w_1 + w_2)(-1) = 0$

Solving both equations

the point of minima is $(0,1)$

$\frac{\partial^2 E}{\partial w_1^2} = 2x_{1i}^2 > 0$, $\frac{\partial^2 E}{\partial w_2^2} = 2x_{2i}^2 > 0$

$\therefore$ $(0,1)$ is point of minima

and the curvature is upward and maybe elliptic paraboloid

(b) $$\text{Hessian} = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} \\ \frac{\partial^2 E}{\partial w_2 \partial w_2} & \frac{\partial^2 E}{\partial w_2^2} \end{bmatrix}$$

$$E = \frac{1}{2} \sum_{i=1}^{2} (y_i - w_1 x_{1i} - w_2 x_{2i})^2$$

$$\frac{\partial^2 E}{\partial w_1^2} = \sum_{i=1}^{2} x_{1i}^2 = 1+1 = 2$$

$$\frac{\partial^2 E}{\partial w_2^2} = \sum_{i=1}^{2} x_{2i}^2 = 1+1 = 2$$

$$\frac{\partial^2 E}{\partial w_1 \partial w_2} = \sum_{i=1}^{2} x_{1i} x_{2i} = 1-1 = 0$$

$$\frac{\partial^2 E}{\partial w_2 \partial w_1} = \sum_{i=1}^{2} x_{1i} x_{2i} = 1-1 = 0.$$

$$\text{Hessian} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$ where $\lambda_1 = \lambda_2 = 2$ are the eigenvalues of the hessian