A1 a) k_1 & k_2 are valid kernels

$\Rightarrow$ feature representation of kernel k_1

$\phi(x) = [\phi_1(x)\ \phi_2(x) \ldots \phi_n(x)]$

feature representation of kernel k_2

$\phi'(x) = [\phi'_1(x)\ \phi'_2(x) \ldots \phi'_n(x)]$

A1, a)

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

$k_2(x,z) = \langle \phi'(x), \phi'(z) \rangle$

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

$= \langle \phi(x), \phi(z) \rangle + \langle \phi'(x), \phi'(z) \rangle$

$= \langle [\phi(x), \phi'(x)], [\phi(z)\ \phi'(z)] \rangle$

$\Rightarrow$ From above representation $k(x,z)$ is a valid dot product hence a valid kernel.

A1(b) $k(x,z) = k_1(x,z)\, k_2(x,z)$

$$k_1(x,z) = \langle \phi(x), \phi(z) \rangle$$
$$= \sum_{i=1}^{n} \phi_i(x)\, \phi_i(z)$$

$$k_2(x,z) = \langle \phi'(x), \phi'(z) \rangle$$
$$= \sum_{i=1}^{n} \phi'_i(x)\, \phi'_i(z)$$

$$k(x,z) = \left(\sum_{i=1}^{n} \phi_i(x)\phi_i(z)\right)\left(\sum_{j=1}^{n} \phi'_j(x)\phi'_j(z)\right)$$
$$= \langle \phi_1(x) \times \phi'(x), \phi_1(z) \times \phi'(z) \rangle + \dots + \langle \phi_n(x) \times \phi'(x), \phi_n(z) \times \phi'(z) \rangle$$
$$= \left\langle \left(\sum_{i=1}^{n} \phi_i(x)\right)\phi'(x), \left(\sum_{i=1}^{n} \phi_i(z)\right) \times \phi'(z) \right\rangle$$

As it can be expressed as dot product it is a valid kernel.

A1(c) $K(x,z) = h(K_1(x,z))$

$K(x,z)$ is a linear combination of powers of $K(x,z)$ with positive coefficients. Since powers of K_1 are products of K_1 by itself. so it is a valid kernel therefore their linear combination is a valid kernel.

A1(d) $K(x,z) = \exp(K_1(x,z))$

$$\exp(x) = \lim_{n\to\infty} \left(1 + x \cdots + \frac{x^n}{\lfloor n}\right) \Rightarrow$$ It is a polynomial with +ve coefficients

~~We know~~

$$K(x,z) = \lim_{n\to\infty} \left(1 + \cdots \frac{K_1^n(x,z)}{\lfloor n}\right)$$

We know:

~~Using part c~~ $K'(x,y) = \lim_{n\to\infty} K_n'(x,y)$

Using part (c) we can say $K(x,z)$ is a valid kernel.

A1(e) $K(x,z) = \exp\left(\frac{-\|x-z\|_2^2}{\sigma^2}\right)$

$$K(x,z) = \exp\left(\frac{-1}{\sigma^2}\|x-z\|_2^2\right)$$

$$= \exp\left(\frac{-1}{\sigma^2}x^Tx\right)\exp(x^Tz)\exp\left(\frac{-1}{\sigma^2}z^Tz\right)$$

From previous part $\exp(x^Tz)$ is a valid kernel as x^Tz is an inner product

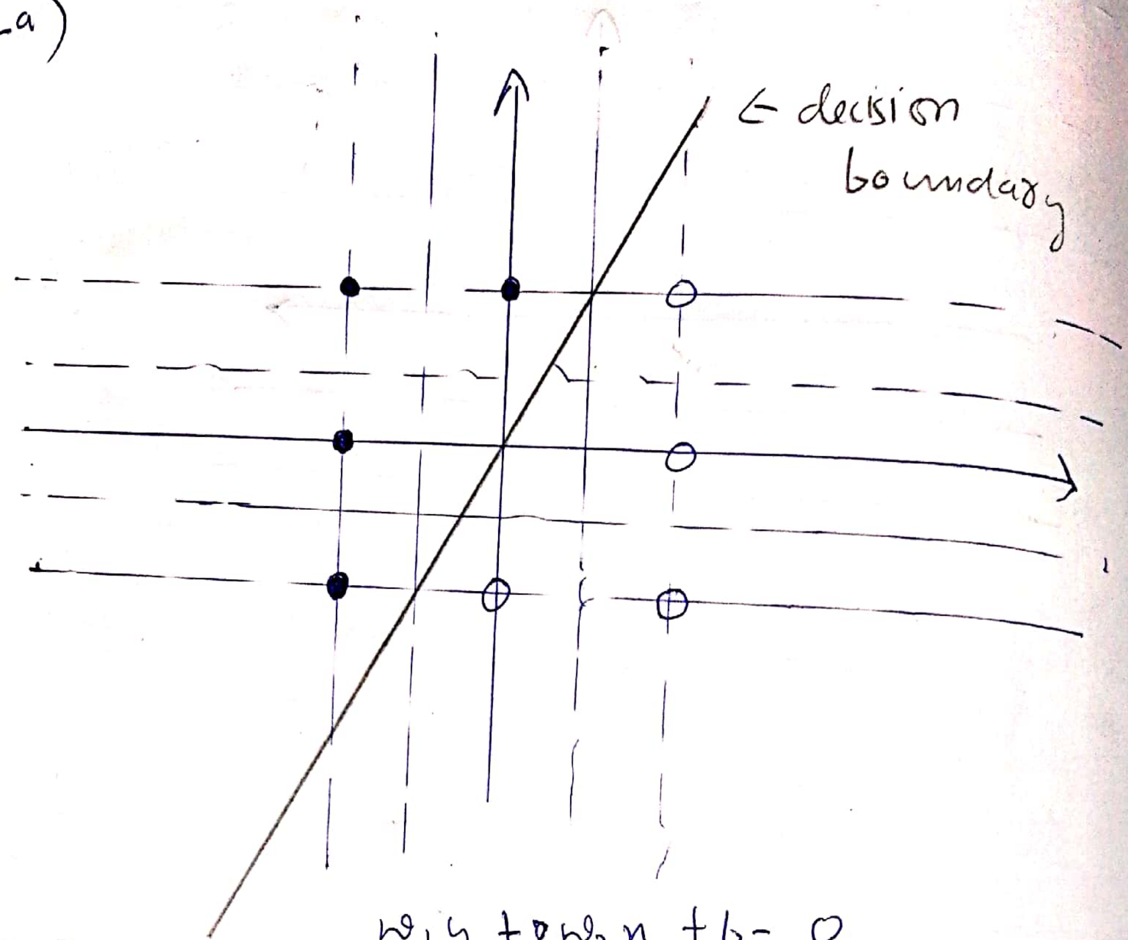
$K(x, z) = f(x) \exp(x^T z) f(z)$

+ve

+ve

$\Rightarrow$ $K(x, z)$ is a valid ~~fa~~ Kernel.

A2a)

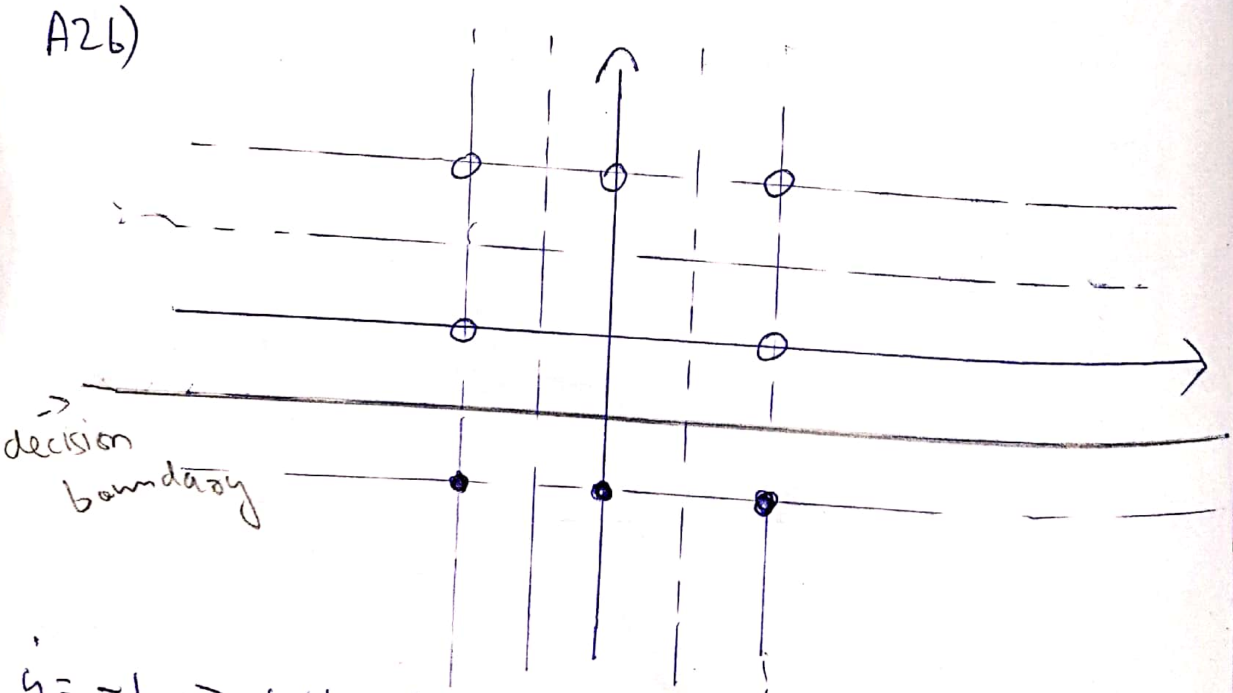


$w_1 y + w_2 x + b = 0$

$y = 2x \Rightarrow y - 2x = 0$

bias $= 0$, $w_1 = 1$, $w_2 = -2$ $\Rightarrow$ weight $= [w_1 \ w_2] = [1 \ -2]$

A2b)



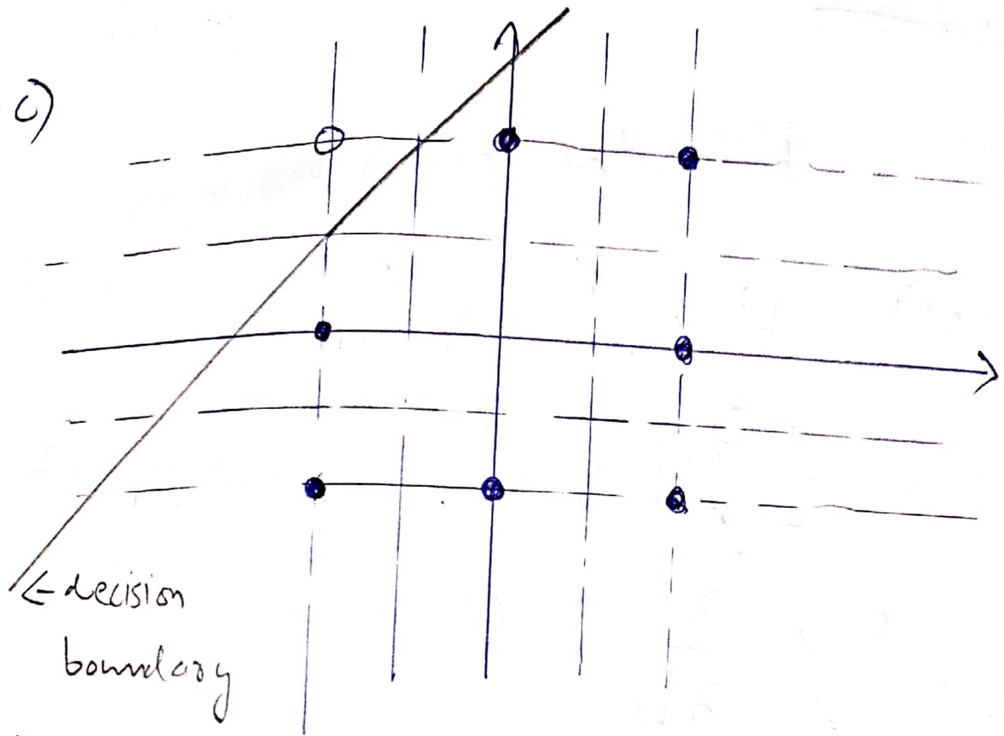
$y = -1 \Rightarrow y + 1 = 0$

~~bias = 0 ; $w_1 = 1$~~ bias $= 1$, $w_1 = 1$, $w_2 = 0$

$w_1 y + w_2 x + b = 0$

weight $= [w_1 \ \ w_2] = [1 \ \ 0]$

A2 c)



$w_1 y + w_2 x + b = 0$

So $y = x+3 \Rightarrow y - x - 3 = 0$

bias $= -3$, weight $= [w_1 \ \ w_2] = [1 \ \ -1]$

A3a) $x_1 = [1, 1]$ | $x_2 = [1, -1]$

$y_1 = 1$ | $y_2 = -1$

~~$y_i = w_1 x$~~

$\hat{y}$ is predicted class.

$$\hat{y}(x_i) = w_1 x_{1i} + w_2 x_{2i}$$

$$\underset{(E)}{\text{Error}} = \frac{1}{2} \sum_{i=1}^{2} (y_i - w_1 x_{i1} - w_2 x_{i2})^2$$

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

$$\Rightarrow \sum_{i=1}^{2} (y_i - w_1 x_{i1} - w_2 x_{i2})(-x_{i1}) = 0$$

&

$$\sum_{i=1}^{2} (y_i - w_1 x_{i1} - w_2 x_{i2})(-x_{i2}) = 0$$

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

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

$\Rightarrow$ On Solving: $w_1 = 0$ & $w_2 = 1$

$(0, 1)$ as minimum point.

$$\frac{\partial^2 E}{\partial w_1^2} = 2x_{i1}^2 > 0, \quad \frac{\partial^2 E}{\partial w_2^2} = 2x_{i2}^2 > 0$$

$\Rightarrow \therefore (0, 1)$ is the point of minima & the curvature of mean error surface is upward. It is an elliptic paraboloid.

A3 b) Hessian $= \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial E}{\partial w_1 \partial w_2} \\ \frac{\partial^2 E}{\partial w_1 \partial w_2} & \frac{\partial^2 E}{\partial^2 w_2^2} \end{bmatrix}$

$$E = \frac{1}{2} \sum_{i=1}^{2} (y_i - w_i x_{i1} - w_2 x_{i2})^2$$

$$\frac{\partial^2 E}{\partial w_1^2} = \sum x_{i1}^2 = 1 + 1 = 2$$

$$\frac{\partial^2 E}{\partial w_2^2} = \sum x_{i2}^2 = 2$$

$$\frac{\partial^2 E}{\partial w_1 \partial w_2} = \sum x_{i1} x_{i2} = 0$$

$$\frac{\partial^2 E}{\partial w_1 \partial w_2} = \sum x_{i1} x_{i2} = 0$$

$=0$

Hessian $= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$\lambda_1 = 2$, $\lambda_2 = 2$ are the eigen vectors of the Hessian matrix.

ia

ce