ML Homework 3

Mehul Yesminkumar

1 Kernels

1.1

We know the given kernel putting these values in the kernel we get

$$K(x,x') = \begin{cases} 1, & \text{if } \ x = x' \\ 0, & \text{otherwise} \end{cases} \qquad \implies K = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

To get a valid kernel we should have $\alpha^T K \alpha \geq 0$

$$\alpha^T K \alpha = \begin{bmatrix} a_1 & a_2 & a_3 & . & a_n \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ a_n \end{bmatrix}$$

Hence we get $a_1^2 + a_2^2 + a_3^2 + ... + a_n^2 >= 0$. This matrix is positive and semidefinite.

Hence we can say that this is a valid kernel

1.2

$$J(\alpha) = \frac{1}{2}\alpha^T K^T K\alpha - y^T K\alpha + \frac{1}{2}y^T y \qquad since \quad \lambda = 0$$

$$\frac{\partial J(\alpha)}{\partial \alpha} = \alpha^T K^T K - y^T K = 0 \qquad \text{Differentiating w.r.t } \alpha$$

$$\implies \alpha^T K^2 - y^T K = 0$$

$$\implies \alpha^* = K^{-1}y$$

Hence $\alpha^* = y$ Since $K^{-1} = K = I$

1.3

Given, $f(x) = k(x, x_1), k(x, x_2)...k(x, x_N)\alpha^*$

We know for any $x \neq x_n k(x, x') = 0$

Hence we get, $f(x) = [0 * \ 0 * \ 0 * \ . \ . \ 0]\alpha^*$

$\implies f(x) = 0$

Thus for any $x \neq x_n$ prediction is always 0

2 Support Vector Machines

2.1

No the data is not linearly separable in this one dimensional space

$$\text{We know } y[w^T x + b] \leq 0 \implies \begin{cases} w^T - b & x = -1 \text{ and } y = -1 \\ b, & x = 0 \text{ and } y = 1 \\ -w^T - b & x = 1 \text{ and } y = -1 \end{cases}$$

From above case (1) and case (3) $w^T - b \leq 0,\ b \leq 0\ and\ -w^T - b \leq 0$ they contradict

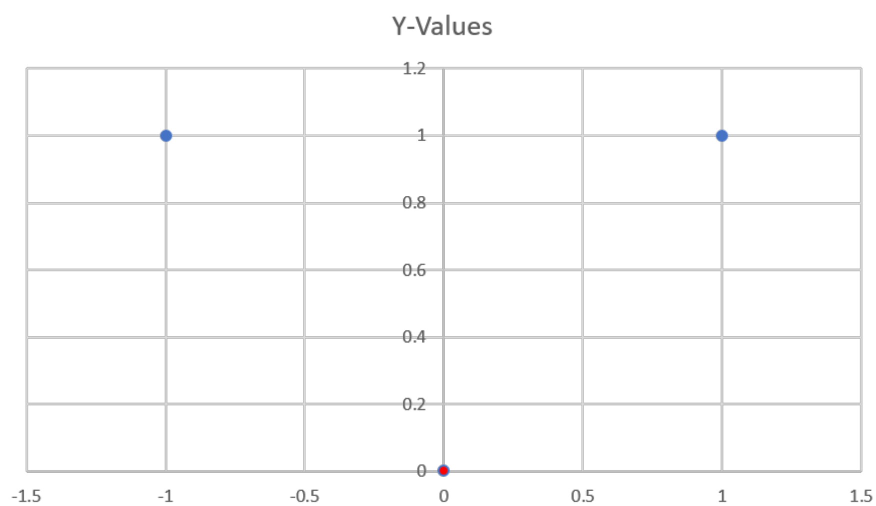
We need two decision boundaries to separate the data in 1-D feature space

Hence data is not separable

2.2

Yes there exists a boundary after transferring from 1D space to 2D space

New point in new space are $b > w, b > -w, b > 0$



2.3

$$k = \phi(x)^T \phi(x') \qquad \phi(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$$

$$\implies k = x * x' + x^2 * (x')^2$$

$$\implies K = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{let } z = \begin{bmatrix} l \\ m \\ n \end{bmatrix} \qquad z\, is\, non\, zero\, column\, vector$$

For K to be a postive semi definite, $z^T Mz > 0$

$$\begin{bmatrix} l & m & n \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l \\ m \\ n \end{bmatrix} = 2l^2 + 2m^2 > 0$$

Hence $z^T Mz > 0$

2.4

Primal formula $= min_w \frac{1}{2}||w||_2^2$

$s.t \quad w^T[-1 \quad 1]^T + b <= -1$

$w^T[1 \quad 1]^T + b <= -1$

$b >= 1$

Dual formula $= min_\alpha L(\alpha) = min_\alpha \alpha_1^2 + \alpha_2^2 - \alpha_1 - \alpha_2 - \alpha_3$

$s.t \quad \alpha_1 >= 0 \quad \alpha_2 >= 0 \quad \alpha_3 >= 0$

$-\alpha_1 - \alpha_2 + \alpha_3 = 0$

2.5

$$J(\alpha) = min\alpha\frac{1}{2}\sum_{m,n}^{3} y_m y_n \alpha_m \alpha_n \phi(x_m)\phi(x_n) - \sum_{n}^{3} \alpha a_n$$

$y_1\alpha_1 + y_2\alpha_2 + y_3\alpha_3 = 0$

We know $y_1 = -1, y_2 = -1, y_3 = 1$

$\implies$ By symmetry $\alpha_1 = \alpha_2$ and $\alpha_3 = 2\alpha_1$

$\implies min_\alpha \frac{1}{2}(2\alpha_1^2 + 2\alpha_2^2) - (\alpha_1 + \alpha_2 + \alpha_3)$ using dual formation

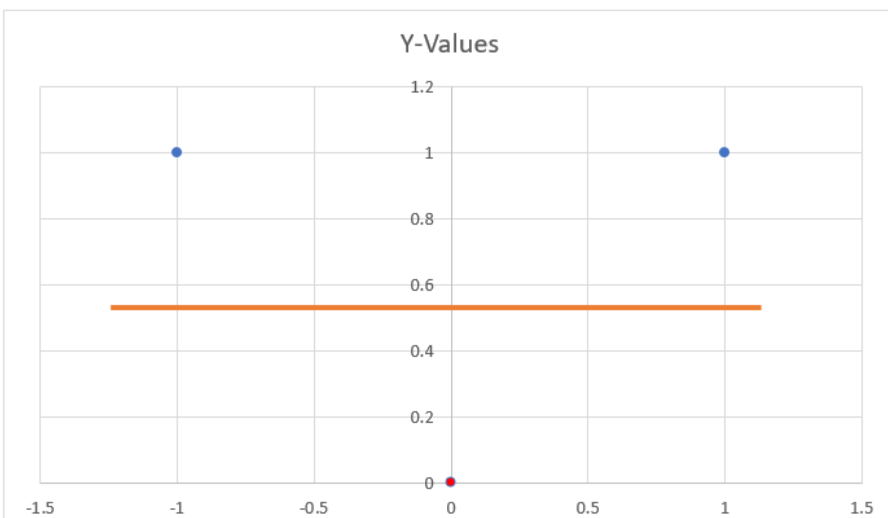
$\frac{\partial J(\alpha)}{\partial \alpha} = 4(\alpha_1 - 1) = 0$ now diff w.r.t α

$\implies \alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 2$

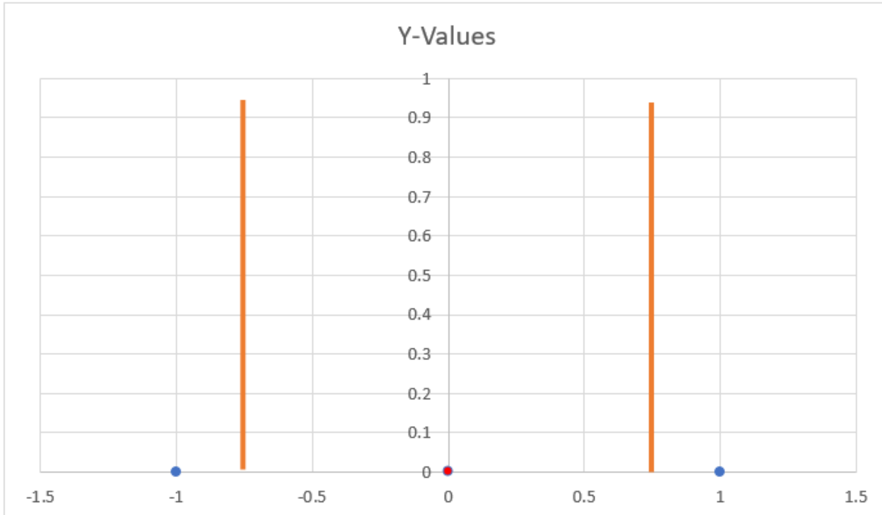
$w = \sum y_n \alpha_n \phi(x_n) = [0, -2]^T$

$b = y_n - \sum_{m=1}^{3} y_m \alpha_m k(x_m, x_n) = 1$

2.6



Two dimensional Space



One dimensional Space

The circled points are the support vectors

First plot $x^2 = 0.5$ is the decision boundary in the 2-D feature space

Second plot $x = 0.707$ and $x = -0.707$ is the the decision boundaries in 1-D feature space

3 Adaboost

3.1

$$f_1(s, b, d) = (1, 0.5, -1)$$

$$\epsilon_1 = 0.50 \qquad \beta_1 = 0$$

3.2

$$w_2(1) = 0.25 \quad w_2(2) = 0.25 \quad w_2(3) = 0.25 \quad w_2(4) = 0.25$$

$$\implies w_2(n) = [0.25 \quad 0.25 \quad 0.25 \quad 0.25]$$

3.3

$$f_1(s, b, d) = (1, -0.5, 1)$$

$$\epsilon_1 = 0.25 \qquad \beta_1 = 0.55$$

3.4

$$f_2(s, b, d) = (1, 0.5, 2)$$

$$\epsilon_2 = 0.17 \qquad \beta_2 = 0.79$$

$$w_2(n) = [0.17 \quad 0.17 \quad 0.17 \quad 0.49]$$

3.5

$$f_3(s, b, d) = (-1, -0.5, 2)$$

$$\epsilon_3 = 0.1 \qquad \beta_3 = 1.1$$

$$w_3(n) = [0.10 \quad 0.10 \quad 0.50 \quad 0.30]$$

3.6

$$F(x) = sign[0.55 * h_{(1,-0.5,1)}(x) + 0.807 * h_{(1,0.5,2)}(x) + 0.109 * h_{(-1,-0.5,2)}(x)]$$

$$F(x_1) = 1 \quad F(x_2) = -1 \quad F(x_3) = 1 \quad F(x_4) = -1$$

$$\implies F(x) = [1 \quad -1 \quad 1 \quad -1]$$

All are labeled correctly