

CSCI 567 Machine Learning

Homework #3

Name : Yi Zhao

Question 1.1 Answer :

Define a matrix $\mathbf{K} \in R^{N \times N}$ which each element \mathbf{K}_{ij} is $k(\mathbf{x}_i, \mathbf{x}_j)$, according to the assume, we know \mathbf{K} is the Identity Maxtrix. According to the Mercer's theorem, we only need to prove that the Identity Maxtrix \mathbf{K} is positive semidefinite:

For any vector $\mathbf{a} \in R^N$, we can get,

$$\mathbf{a}^T \mathbf{K} \mathbf{a} = \mathbf{a}^T \mathbf{a} \geq 0$$

So Identity Maxtrix \mathbf{K} is positive semidefinite, and $k(\mathbf{x}, \mathbf{x}')$ is a valid kernel.

Question 1.2 Answer :

$$\text{when } \lambda = 0, \mathbf{K} = \mathbf{I}, \text{ then } J(\boldsymbol{\alpha}) = \frac{1}{2} \boldsymbol{\alpha}^T \boldsymbol{\alpha} - \mathbf{y}^T \boldsymbol{\alpha} + \frac{1}{2} \mathbf{y}^T \mathbf{y}$$

$$\frac{\partial J(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \boldsymbol{\alpha} - \mathbf{y} = 0$$

$$\boldsymbol{\alpha}^* = \mathbf{y}$$

$$\text{so, } J(\boldsymbol{\alpha}^*) = 0$$

Question 1.3 Answer :

$$k(\mathbf{x}, \mathbf{x}_n) = 0 \text{ with } \mathbf{x} \neq \mathbf{x}_n, \forall n = 1, 2, \dots, N$$

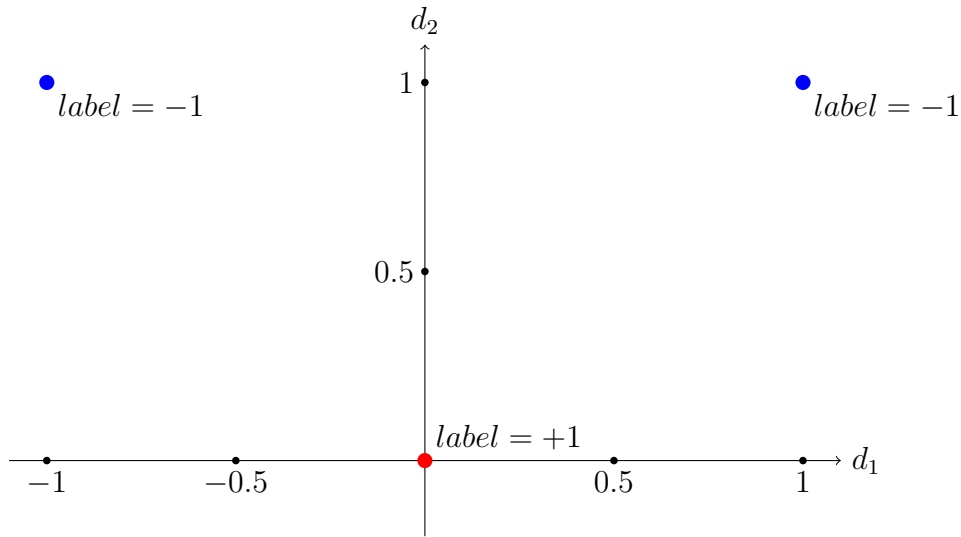
$$f(\mathbf{x}) = [0, 0, \dots, 0] \boldsymbol{\alpha}^*$$

$$f(\mathbf{x}) = \mathbf{0} \boldsymbol{\alpha}^* = 0$$

Question 2.1 Answer :

No. Because in one-dimensional feature space, assuming the linear separator is $ax + b$ for x_1, x_2, x_3 , so we can get $-1 \times (-a + b) \geq 0$, $-1 \times (a + b) \geq 0$, $1 \times b \geq 0$, then $a = 0$ and $b = 0$.

Question 2.2 Answer :



Yes, we can assume this linear decision boundary is $y[a_1d_1 + a_2d_2 + a_3] \geq 0$, substitute with three point and we can easily get one linear decision boundary $d_2 = 0.5$.

Question 2.3 Answer :

$$k(x, x') = \phi(x)^T \phi(x') = xx' + (xx')^2$$

$$\mathbf{K} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Given an non-zero column vector $\mathbf{z} = [z_1, z_2, z_3]^T$, $\forall z_1, z_2, z_3 \in R$

$$\mathbf{z}^T \mathbf{K} \mathbf{z} = [2z_1, 2z_2, 0]^T [z_1, z_2, z_3]$$

$$\mathbf{z}^T \mathbf{K} \mathbf{z} = 2z_1^2 + 2z_2^2 \geq 0, \forall z_1, z_2 \in R$$

So, K is a positive semi-definite (PSD) matrix.

Question 2.4 Answer :

Primal formulations:

$$\min_{w_1, w_2, b, \{\xi_n\}} C \sum_n \xi_n + \frac{1}{2}(w_1^2 + w_2^2)$$

$$\text{s.t. } 1 + [-w_1 + w_2 + b] \leq \xi_1$$

$$1 + [w_1 + w_2 + b] \leq \xi_2$$

$$1 - b \leq \xi_3$$

$$\xi_n \geq 0, \quad n = 1, 2, 3$$

Dual formulations:

$$\max_{\alpha} \sum_n \alpha_n - \frac{1}{2} \sum_{m,n} y_m y_n \alpha_m \alpha_n k(x_m, x_n)$$

$$\text{So, } \max_{\alpha} (\alpha_1 + \alpha_2 + \alpha_3) - \frac{1}{2}(2y_1^2 \alpha_1^2 + 2y_2^2 \alpha_2^2)$$

$$\text{So, } \max_{\alpha} (\alpha_1 + \alpha_2 + \alpha_3) - (\alpha_1^2 + \alpha_2^2)$$

$$\text{s.t. } 0 \leq \alpha_n \leq C, \quad n = 1, 2, 3$$

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

Question 2.5 Answer :

$$\min_{\alpha} \frac{1}{2} \sum_{m,n} y_m y_n \alpha_m \alpha_n k(x_m, x_n) - \sum_n \alpha_n$$

$$\min_{\alpha} \alpha_1^2 + \alpha_2^2 - (\alpha_1 + \alpha_2 + \alpha_3)$$

$$\text{s.t. } 0 \leq \alpha_n \leq C, \quad n = 1, 2, 3$$

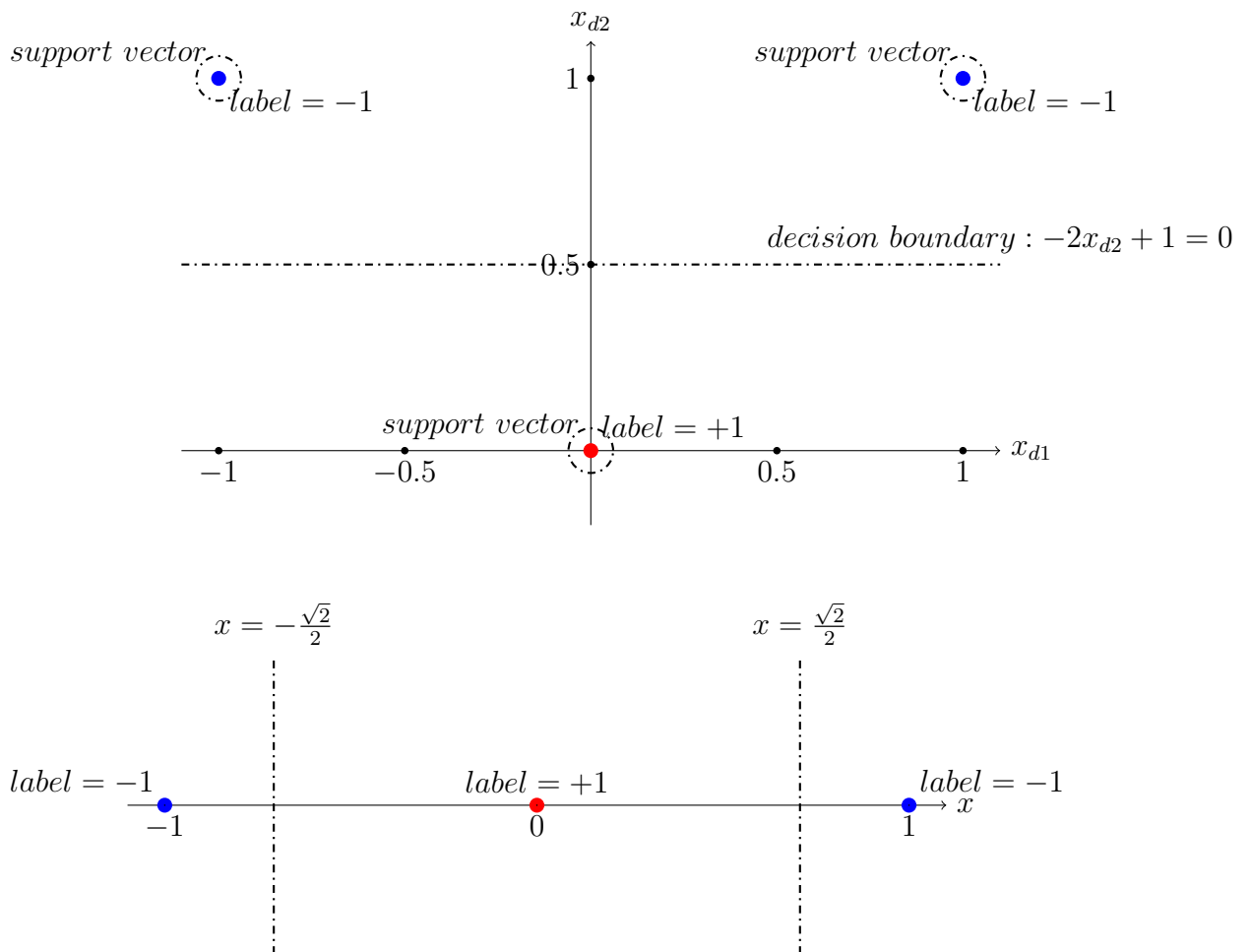
$$\alpha_1 + \alpha_2 - \alpha_3 = 0$$

$$\text{due to symmetry, } \alpha_1 = \alpha_2 = \frac{1}{2}\alpha_3$$

in order to minimize objective function, $\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 2$

$$\begin{aligned}\boldsymbol{\alpha} &= (1 \quad 1 \quad 2) \\ \boldsymbol{w} &= \sum_n \alpha_n y_n \phi(x_n) = (0 \quad -2)^T \\ b &= [y_n - \boldsymbol{w}^T \phi(x_n)] = 1\end{aligned}$$

Question 2.6 Answer :



Question 3.1 Answer :

$$\begin{aligned}f_{1(+1,-2,1)} &= \begin{cases} +1, & \text{if } x_1 > -2 \\ -1, & \text{otherwise.} \end{cases} \\ \epsilon_1 &= 0.50\end{aligned}$$

$$\beta_1 = 0.00$$

Question 3.2 Answer :

$$w_2(1) = 0.25$$

$$w_2(2) = 0.25$$

$$w_2(3) = 0.25$$

$$w_2(4) = 0.25$$

Question 3.3 Answer :

$$f_{1(+1,-0.5,1)} = \begin{cases} +1, & \text{if } x_1 > -0.5 \\ -1, & \text{otherwise.} \end{cases}$$

$$\epsilon_1 = 0.25$$

$$\beta_1 = \frac{1}{2} \ln 3 = 0.55$$

Question 3.4 Answer :

$$w_2(1) = \frac{1}{4\sqrt{3}}$$

$$w_2(2) = \frac{1}{4\sqrt{3}}$$

$$w_2(3) = \frac{1}{4\sqrt{3}}$$

$$w_2(4) = \frac{\sqrt{3}}{4}$$

Normalize them, then we can get $w_2(1) = w_2(2) = w_2(3) = \frac{1}{6} = 0.17$

$$w_2(4) = 0.50$$

$$f_{2(-1,0.5,1)} = \begin{cases} -1, & \text{if } x_1 > 0.5 \\ +1, & \text{otherwise.} \end{cases}$$

$$\epsilon_2 = w_2(2) = \frac{1}{6} = 0.17$$

$$\beta_2 = \frac{1}{2}\ln 5 = 0.80$$

Question 3.5 Answer :

$$w_3(1) = \frac{1}{6\sqrt{5}}$$

$$w_3(2) = \frac{\sqrt{5}}{6}$$

$$w_3(3) = \frac{1}{6\sqrt{5}}$$

$$w_3(4) = \frac{1}{2\sqrt{5}}$$

Normalize them, then we can get $w_3(1) = 0.10$, $w_3(2) = 0.50$, $w_3(3) = 0.10$, $w_3(4) = 0.30$

$$f_{3(-1,-0.5,2)} = \begin{cases} -1, & \text{if } x_2 > -0.5 \\ +1, & \text{otherwise.} \end{cases}$$

$$\epsilon_3 = w_3(1) = 0.1$$

$$\beta_3 = \frac{1}{2}\ln 9 = 1.10$$

Question 3.6 Answer :

$$F(\mathbf{x}) = \text{sign}[0.55h_{(+1,-0.5,1)}(\mathbf{x}) + 0.80h_{(-1,0.5,1)}(\mathbf{x}) + 1.10h_{(-1,-0.5,2)}(\mathbf{x})]$$

$$F(\mathbf{x}_1) = \text{sign}[0.55 + 0.80 - 1.10] = +1$$

$$F(\mathbf{x}_2) = \text{sign}[-0.55 + 0.80 - 1.10] = -1$$

$$F(\mathbf{x}_3) = \text{sign}[0.55 + 0.80 + 1.10] = +1$$

$$F(\mathbf{x}_4) = \text{sign}[0.55 - 0.80 - 1.10] = -1$$

All four labeled examples are correct.