CSCI 567 Machine Learning

Homework #3

Name : Yi Zhao

Question 1.1 Answer :

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

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

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{a} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} \geq 0$$

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

Question 1.2 Answer :

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

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

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

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

Question 1.3 Answer :

$$k(\boldsymbol{x}, \boldsymbol{x}_n) = 0\ with\ \boldsymbol{x} \neq \boldsymbol{x}_n,\ \forall n = 1, 2, \cdots, N$$

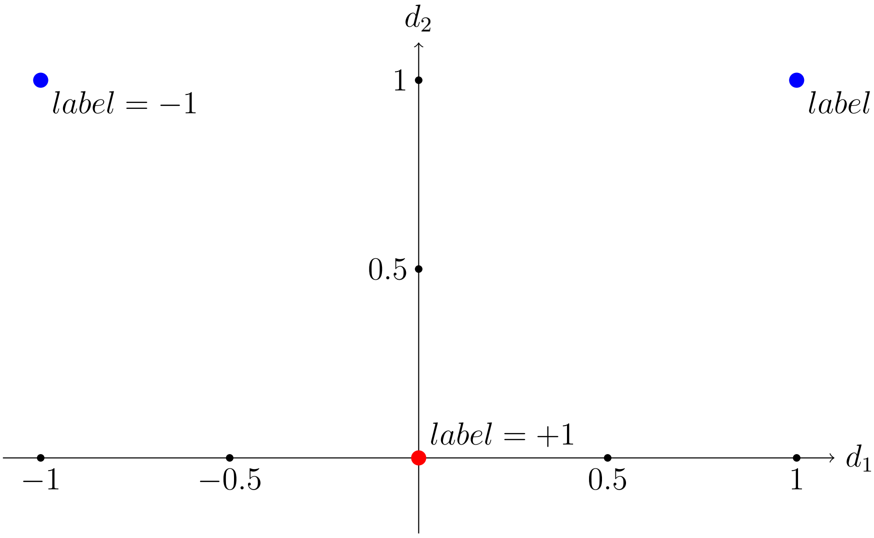
$$f(\boldsymbol{x}) = [0, 0, \cdots, 0]\boldsymbol{\alpha}^*$$

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

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 :

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

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

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

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

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

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

Question 2.4 Answer :

Primal formulations:

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

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

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

$$1 - b \leq \xi_3$$

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

Dual formulations:

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

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

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

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

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

Question 2.5 Answer :

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

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

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

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

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

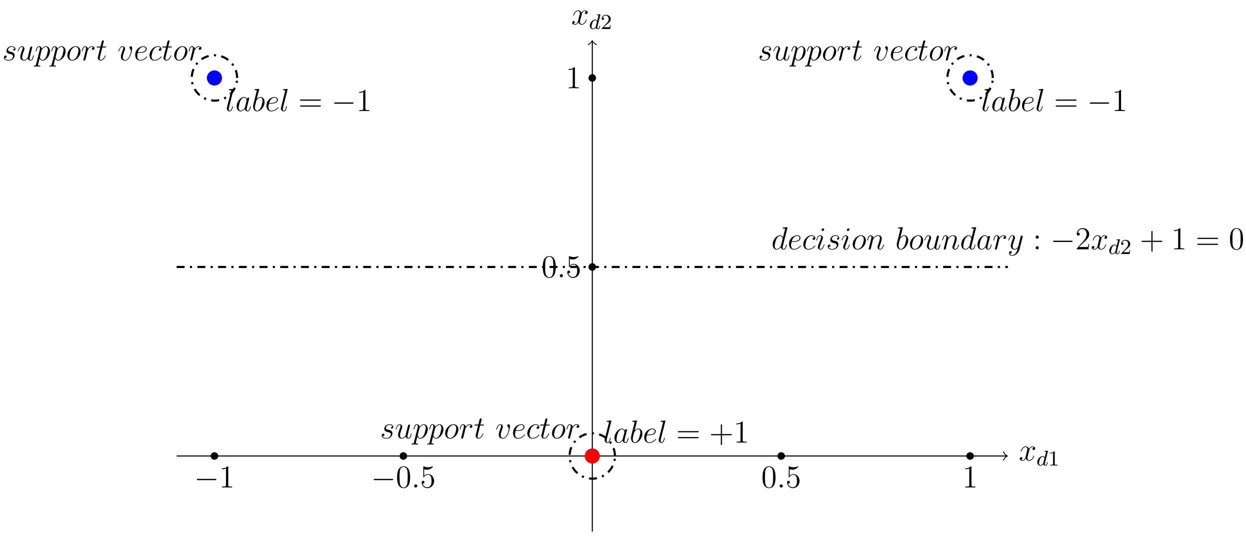
$$\text{in order to minimize objective function,} \;\; \alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 2$$

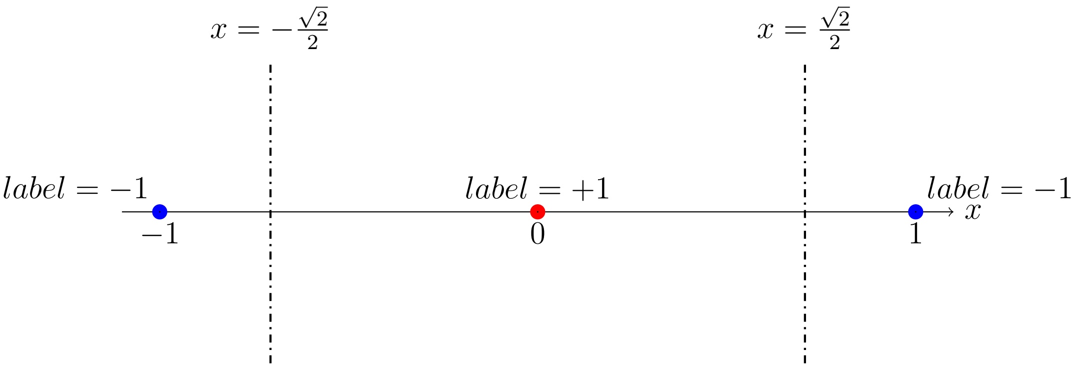
$$\boldsymbol{\alpha} = \begin{pmatrix}1 & 1 & 2\end{pmatrix}$$

$$\boldsymbol{w} = \sum_n \alpha_n y_n \phi(x_n) = \begin{pmatrix}0 & -2\end{pmatrix}^{\mathrm{T}}$$

$$b = [y_n - \boldsymbol{w}^{\mathrm{T}}\phi(x_n)] = 1$$

Question 2.6 Answer :





Question 3.1 Answer :

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

$$\epsilon_1 = 0.50$$

$$\beta_1 = 0.00$$

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 :

$$f_{1(+1,-0.5,1)} = \begin{cases} +1, & if\ x_1 > -0.5 \\ -1, & otherwise. \end{cases}$$
$$\epsilon_1 = 0.25$$
$$\beta_1 = \frac{1}{2}\ln 3 = 0.55$$

Question 3.4 Answer :

$$w_2(1) = \frac{1}{4\sqrt{3}}$$
$$w_2(2) = \frac{1}{4\sqrt{3}}$$
$$w_2(3) = \frac{1}{4\sqrt{3}}$$
$$w_2(4) = \frac{\sqrt{3}}{4}$$
$$Normalize\ them, then\ we\ can\ get\ w_2(1) = w_2(2) = w_2(3) = \frac{1}{6} = 0.17$$
$$w_2(4) = 0.50$$
$$f_{2(-1,0.5,1)} = \begin{cases} -1, & if\ x_1 > 0.5 \\ +1, & otherwise. \end{cases}$$
$$\epsilon_2 = w_2(2) = \frac{1}{6} = 0.17$$

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

Question 3.5 Answer :

$$w_3(1) = \frac{1}{6\sqrt{5}}$$
$$w_3(2) = \frac{\sqrt{5}}{6}$$
$$w_3(3) = \frac{1}{6\sqrt{5}}$$
$$w_3(4) = \frac{1}{2\sqrt{5}}$$

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

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

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

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

Question 3.6 Answer :

$$F(\boldsymbol{x}) = \mathrm{sign}[0.55h_{(+1,-0.5,1)}(\boldsymbol{x}) + 0.80h_{(-1,0.5,1)}(\boldsymbol{x}) + 1.10h_{(-1,-0.5,2)}(\boldsymbol{x})]$$
$$F(\boldsymbol{x_1}) = \mathrm{sign}[0.55 + 0.80 - 1.10] = +1$$
$$F(\boldsymbol{x_2}) = \mathrm{sign}[-0.55 + 0.80 - 1.10] = -1$$
$$F(\boldsymbol{x_3}) = \mathrm{sign}[0.55 + 0.80 + 1.10] = +1$$
$$F(\boldsymbol{x_4}) = \mathrm{sign}[0.55 - 0.80 - 1.10] = -1$$

All four labeled examples are correct.