CSE - 5526

Homework 3

Submitted by

Utkarsh Pratap Singh Jadon (500711257)
jadon.1@osu.edu

Contents

1 Question 1 **2**

2 Question 2 **5**

1 Question 1

Given the following lineraly separable training patterns:

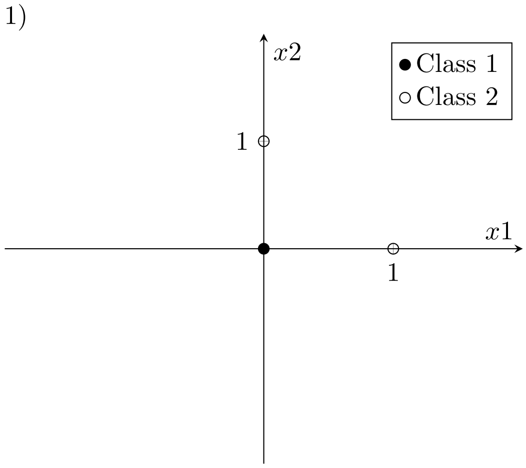
$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $d_1 = 1$

$\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $d_2 = -1$

$\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $d_3 = -1$

1. Find $\mathbf{w}_o$ and b_o for the optimal hyperpane by optimizing the Lagrangian function.
2. Write down the discriminant function.
3. Specify which of the input patterns are support vectors.

Solution
1)



The SVM Dual Problem is given by:

$$Q(\alpha) = \sum \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j x_i^T x_j \tag{1}$$

subject to:

1. $\sum_{i=1}^{N} \alpha_i d_i = 0$
2. $\alpha_i \geq 0$

Only the points x_i that lie on the supporting hyperplane have $\alpha_i > 0$. These Support Vectors determine the decision boundary.

$W_0 = \sum_{i=1}^{N_s} \alpha_i d_i x_i$

Substituting the values of x_i, x_j, d_i, and d_j in Eq (1), we get:

$Q(\alpha) = -\frac{1}{2}[\alpha_2^2 + \alpha_3^2] + \alpha_1 + \alpha_2 + \alpha_3$

such that $\alpha_1 - \alpha_2 - \alpha_3 = 0$

Now, substituting the value of α_1, we get:

$Q(\alpha) = -\frac{1}{2}[\alpha_2^2 + \alpha_3^2] + 2\alpha_2 + 2\alpha_3$

To maximize the cost function $Q(\alpha)$, we take partial derivatives with respect to α_2 and α_3, as follows:

$\frac{\mathrm{d}Q(\alpha)}{\mathrm{d}\alpha_2} = -\alpha_2 + 2 = 0$
$\implies \alpha_2 = 2$

Similarly,

$\frac{\mathrm{d}Q(\alpha)}{\mathrm{d}\alpha_3} = -\alpha_3 + 2 = 0$
$\implies \alpha_3 = 2$

We know that $\alpha_1 - \alpha_2 - \alpha_3 = 0$

$\implies \alpha_1 = 4$

We know that, $W_0 = \sum_{i=1}^{N_s} \alpha_i d_i x_i$. Substituting the values of α_i, d_i, and x_i, we get:

$W_0 = (4)(1)(0,0)^T + (2)(-1)(1,0)^T + (2)(-1)(0,1)^T$

$\implies \mathrm{W}_0 = (-2,-2)^T$

Width of the margin is given by:

$d = \frac{2}{\|W_0\|} = \frac{2}{2\sqrt{2}} = \frac{1}{\sqrt{2}}$

To get the bias b, we substitute the values in $d(W_0^T x_i + b) = 1$ for any data point. For $(1,0)^T$, $d = -1$

$(-1)((-2,-2)(1,0)^T + b) = 1$

$\implies \mathrm{b} = 1$

2) Discriminant function is given by:

$\mathrm{g(x)} = 0$

$\implies \mathrm{W}_0^T x + b = 0$

$$\implies \begin{bmatrix} -2 - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b = 0$$

3) All the input patterns with non-zero α values are support vectors. In this case, since all the obtained α values are non-zero, all three given data points $[0,0]^T, [1,0]^T$, and $[0,1]^T$ are the support vectors.

2 Question 2

Prove that the kernel matrix **K** is positive semidefinite (for a definition, see p. 283 of the textbook) for inner-product kernel functions.

Solution
Kernel matrix K for inner-product kernel function is given by:

$$K = k(x_i, x_j)_{i,j=1}^{N}$$

where inner-product kernel function k is defined as:

$$k(x_i, x_j) = x_i^T x_j$$

Therefore, for a data-set with N = 2, kernel matrix K can be written as:

$$K = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) \\ \\ K(x_2, x_1) & K(x_2, x_2) \end{bmatrix}$$

$$K = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 \\ \\ x_2^T x_1 & x_2^T x_2 \end{bmatrix}$$

We can prove K matrix to be Positive Semi-definite if following conditions are met:
1. K is symmetric
2. $a^t K a \geq 0$

Now, K^T is given by:

$$K^T = \begin{bmatrix} x_1^T x_1 & x_2^T x_1 \\ \\ x_1^T x_2 & x_2^T x_2 \end{bmatrix}$$

Here, $x_1^T x_2 = x_2^T x_1$ since dot product is commutative. Therefore, $K = K^T$, and hence K is symmetric.

Now, for $a^T K a \geq 0$, a is any real valued non-zero vector whose dimension is compatible with that of K.:

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$a^T K a = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1^T x_1 & x_1^T x_2 \\ \\ x_2^T x_1 & x_2^T x_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$= \begin{bmatrix} a_1 x_1^T x_1 + a_2 x_2^T x_1 & a_1 x_1^T x_2 + a_2 x_2^T x_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1^2 x_1^T x_1 + a_1 a_2 x_2^T x_1 + a_1 a_2 x_1^T x_2 + a_2^2 x_2^T x_2 \end{bmatrix}$$

$$= a_1^2 \|x_1\|^2 + a_1 a_2 (x_1^T x_2 + x_2^T x_1) + a_2^2 \|x_2\|^2$$

$$= a_1^2 \|x_1\|^2 + 2 a_1 a_2 x_1^T x_2 + a_2^2 \|x_2\|^2$$

$$\implies (\mathrm{a}_1 x_1 + a_2 x_2)^2 \geq 0$$

Hence, it is proved that kernel matrix K is Positive Semi-definite.