Homework-6 Solutions

Question 1

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).) Using a linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

Answer

The Gram matrix is:

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2^2 + 2\alpha_2\alpha_4 + \alpha_3^2 + 2\alpha_3\alpha_4 + 2\alpha_4^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad -\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0 \end{aligned}$$

Question 2

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Show that with linear kernel the (dual) optimization problem that needs to be solved in terms of the α_i is:

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2 + \alpha_4)^2 + (\alpha_3 + \alpha_4)^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad \alpha_1 = \alpha_2 + \alpha_3 + \alpha_4 \end{aligned}$$

B

The solution to the above optimization problem is: $\alpha_1 = 4,\ \alpha_2 = 2,\ \alpha_3 = 2,\ \alpha_4 = 0$.

a. What are the indexes of the support vectors?

Answer: $1, 2, 3$.

b. What computation needs to be carried out to determine the classification of the point $x_5 = (-1, 0)$ by this SVM.

Answer: Using the first support vector: $b = \frac{1}{-1} - 0 = -1$. For the 3 support vectors: $K(x_j, x_5) = (0, 0, -1)$.

$$\sum_{x_j \text{ is support vector}} \alpha_j y_j K(x_j, x_5) + b = -4(0) + 2(0) + 2(-1) + (-1) = -3 < 0$$

Therefore x_5 should be classified as negative.

c. What computation needs to be carried out to determine the classification of the point $x_5 = (-1, 1)$ by this SVM.

Answer: for the 3 support vectors: $K(x_j, x_5) = (0, 1, -1)$.

$$\sum_{x_j \text{ is support vector}} \alpha_j y_j K(x_j, x_5) + b = -4(0) + 2(1) + 2(-1) + (-1) = -1$$

Therefore the classification of x_5 is negative.

d. What computation needs to be carried out to determine the classification of the point $x_5 = (1, 1)$ by this SVM.

Answer: for the 3 support vectors: $K(x_j, x_5) = (0, 1, 1)$. Compute the sign of:

$$\sum_{x_j \text{ is support vector}} \alpha_j y_j K(x_j, x_5) + (-1) = -4(0) + 2(1) + 2(1) + (-1) = 3$$

Therefore x_5 should be classified as positive.

Question 3

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	-1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).) Using a linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

Answer

The Gram matrix is:

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left(\alpha_2^2 - 2\alpha_2\alpha_4 + \alpha_3^2 - 2\alpha_3\alpha_4 + 2\alpha_4^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad -\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 0 \end{aligned}$$

Question 4

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	-1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Show that with linear kernel the (dual) optimization problem that needs to be solved in terms of the α_i is:

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2 - \alpha_4)^2 + (\alpha_3 - \alpha_4)^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad \alpha_1 = \alpha_2 + \alpha_3 - \alpha_4 \end{aligned}$$

B

Observe that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = k$ satisfies the constraints for any $k \geq 0$, and that the function to be maximized in terms of k is $4k$. Based on these observations, what is the solution to the above (dual) optimization problem?

Answer to B

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \infty$$

The problem is not linearly separable with linear kernel.

Question 5

An SVM is trained with the following data:

i	1	2	3	4	5
x_i	$(0,0)$	$(1,0)$	$(2,0)$	$(3,0)$	$(0,1)$
y_i	-1	1	1	1	-1

Let $\alpha_1, \ldots, \alpha_5$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Using the linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

Answer

$$\text{The Gram matrix for the linear kernel:} \quad G = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 2 & 4 & 6 & 0 \\ 0 & 3 & 6 & 9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Maximize:

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 - \frac{1}{2}(\alpha_2^2 + 4\alpha_2\alpha_3 + 6\alpha_2\alpha_4 + 4\alpha_3^2 + 12\alpha_3\alpha_4 + 9\alpha_4^2 + \alpha_5^2)$$

Subject to:

$$\alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0,\ \alpha_5 \geq 0, \quad -\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \alpha_5 = 0$$

B

The solution to the optimization problem is:

$$\alpha_1 = 2, \quad \alpha_2 = 2, \quad \alpha_3 = 0, \quad \alpha_4 = 0, \quad \alpha_5 = 0.$$

a. Show the computation that needs to be carried out to determine the classification of the point $x = (1, 1)$ by this SVM.

$$\text{using first support verctor:} \quad b = \frac{1}{-1} + 2 \cdot 0 - 2 \cdot 0 = -1$$

$$w'x + b = -2 \cdot 0 + 2 \cdot 1 - 1 > 0 \quad \Rightarrow \quad (1, 1) \ \text{ is positive}$$

b. Show the computation that needs to be carried out to determine the classification of the point $x = (0, -2)$ by this SVM.

$$w'x + b = -2 \cdot 0 + 2 \cdot 0 - 1 < 0 \quad \Rightarrow \quad (1, -2) \ \text{ is negative}$$

C

To obtain solution with soft margins using the l_1 norm (this is the one described in class), the term $C\sigma_i\zeta_i$ is added to the primal function. Using the value of $C = 10$, repeat parts a, b of B.

The only difference in the dual function is the added constraints $\alpha_i \leq 10$. Since the global minimum is at 2 or 0, it would not change and the computation would be identical to the previous one.