```
../../ML/online/Q/svm071.stex
% q07 q08f q09f ho14s ho15 ho16 ho17 ho18 ho19 ho21
```

The training data for an SVM consists of 5 points: $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), (x_5, y_5)$, where: $y_1 = -1, y_2 = 1, y_3 = -1, y_4 = 1, y_5 = 1$. The values of the feature vectors $(x_1, \ldots, x_5)$ are not known explicitly but their Gram matrix is known:

$$G = \begin{pmatrix} 9 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & -2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix}$$

(Element i, j of the Gram matrix is the dot product of x_i and x_j.) Let $\alpha_1, \ldots, \alpha_5$ be the Lagrange multipliers associated with this data. (α_i is associated with (x_i, y_i).)

Part A

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

Maximize

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

Subject to:

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

b.

Using the polynomial kernel of order 2 , and soft margins specified by the parameter $C = 10$, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

Answer

The Gram matrix:

$$\begin{pmatrix} 100 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 4 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 25 & 1 \\ 1 & 1 & 1 & 1 & 4 \end{pmatrix}$$

Maximize

$$\begin{aligned} L(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = {} & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \\ & - \frac{1}{2}(100\alpha_1^2 - 2\alpha_1\alpha_2 + 8\alpha_1\alpha_3 - 2\alpha_1\alpha_4 - 2\alpha_1\alpha_5 + \alpha_2^2 \\ & - 2\alpha_2\alpha_3 + 2\alpha_2\alpha_4 + 2\alpha_2\alpha_5 + 4\alpha_3^2 - 2\alpha_3\alpha_4 - 2\alpha_3\alpha_5 + 25\alpha_4^2 + 2\alpha_4\alpha_5 + 4\alpha_5^2) \end{aligned}$$

Subject to:

$$\begin{aligned} & -\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 + \alpha_5 = 0 \\ & 0 \leq \alpha_1 \leq 10, \quad 0 \leq \alpha_2 \leq 10, \quad 0 \leq \alpha_3 \leq 10, \quad 0 \leq \alpha_4 \leq 10, \quad 0 \leq \alpha_5 \leq 10 \end{aligned}$$

Part B

The table below shows the results obtained for the 4 cases involving linear and second order polynomial kernels, with two different values of C.

Case	kernel	C	α_1	α_2	α_3	α_4	α_5
1	linear	∞	$6.6 \cdot 10^7$	$9.4 \cdot 10^7$	$2.0 \cdot 10^8$	$5.6 \cdot 10^7$	$1.1 \cdot 10^8$
2	2nd order polynomial	∞	0	0.666	0.666	0	0
3	linear	10	3.6	4.8	10	2.9	5.8
4	2nd order polynomial	10	0	0.666	0.666	0	0

a.

Select one of these cases and show that the SVM correctly classifies the entire training data. Show and explain your computations.

Answer

I am using Case 2 or 4. From the first support vector (x_2, y_2):

$$b = \frac{1}{1} - (0.66 \cdot 1 \cdot 1 + 0.66 \cdot (-1) \cdot 1) = 1$$

$$\begin{array}{lll}
\text{For Point 1:} & 1 + (0.666 \cdot 1 \cdot 1 + 0.666 \cdot (-1) \cdot 4) & < 0 \\
\text{For Point 2:} & 1 + (0.666 \cdot 1 \cdot 1 + 0.666 \cdot (-1) \cdot 1) & > 0 \\
\text{For Point 3:} & 1 + (0.666 \cdot 1 \cdot 1 + 0.666 \cdot (-1) \cdot 4) & < 0 \\
\text{For Point 4:} & 1 + (0.666 \cdot 1 \cdot 1 + 0.666 \cdot (-1) \cdot 1) & > 0 \\
\text{For Point 5:} & 1 + (0.666 \cdot 1 \cdot 1 + 0.666 \cdot (-1) \cdot 1) & > 0
\end{array}$$

b.

Select one of these cases and show that the SVM does not correctly classify the entire training data. Show and explain your computations.

Answer

For Case 1 the answer is not clean. It should probably not be used since the values of the alphas may be infinity.

For Case 3 the answer is clean. The second support vector (x_2, y_2) is on a hard margin. Its dot product with all vectors is 0. Therefore:

$$b = \frac{1}{1} - (0 + 0 + 0 + 0 + 0) = 1$$

Point 3 is on a soft margin, and therefore, may not be correctly classified.

$$\text{For Point 3:} \quad 1 + (3.6 \cdot 1 \cdot (-3) + 10 \cdot 1 \cdot 1) \quad > 0$$