Exercise 12
- Proposed Solution -
Friday, January 25, 2019

Solution of Problem 1

a) $b$ is given as $- \frac{1}{2} a^T (x_1 + x_2) = 3$.

b) Supporting vectors are those with $\lambda_i \neq 0$, which are $x_2, x_3, x_4, x_5$.

<table>
<thead>
<tr>
<th>$x_i$</th>
<th>$y_i$</th>
<th>$\lambda_i$</th>
<th>$x_i$</th>
<th>$y_i$</th>
<th>$\lambda_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1)$</td>
<td>$-1$</td>
<td>$0$</td>
<td>$(0.5)$</td>
<td>$1$</td>
<td>$4.73$</td>
</tr>
<tr>
<td>$(2)$</td>
<td>$-1$</td>
<td>$0.67$</td>
<td>$(-2, 1)$</td>
<td>$1$</td>
<td>$0.94$</td>
</tr>
<tr>
<td>$(0)$</td>
<td>$-1$</td>
<td>$5$</td>
<td>$(0, -1)$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

For these vectors, the normal vector of the hyperplane is obtained as:

$$a = \sum_{i=1}^{6} \lambda_i y_i x_i = \lambda_2 y_2 x_2 + \lambda_3 y_3 x_3 + \lambda_4 y_4 x_4 + \lambda_5 y_5 x_5$$

To find $b$, take two support vectors $x_k$ and $x_l$ with $y_k = 1$ and $y_l = -1$ with $0 < \lambda < 5$. For these support vectors, we have $y_i (a^T x_i + b) = 1$. Hence:

$$b^* = -\frac{1}{2} a^* (x_k + x_l) = -\frac{1}{2} \left( \begin{array}{c} -0.86 \\ -1.43 \end{array} \right) \left( \begin{array}{c} 2 \\ 0 \end{array} \right) + \left( \begin{array}{c} -2 \\ 1 \end{array} \right) = \frac{1.43}{2} = 0.715.$$  \hspace{1cm} (1)

Solution of Problem 2

From Mercer’s Theorem we know that a kernel $K : \mathbb{R}^p \rightarrow \mathbb{R}$ is a valid kernel if and only if for any $\{x_1, \ldots, x_n\}$ the kernel matrix $K = (K(x_i, x_j))_{i,j=1,\ldots,n}$ is non-negative definite.

a) For $K(x, y) = K_1(x, y)$ we have that

$$K = (K(x_i, x_j))_{i,j=1,\ldots,n} = (\alpha K_1(x_i, x_j))_{i,j=1,\ldots,n} = \alpha K_1.$$

Then, since $K_1$ is non-negative definite and $\alpha > 0$, it holds

$$z^T K z = \alpha \underbrace{z^T K z}_{\geq 0} \geq 0, \quad \forall z \in \mathbb{R}^p \quad \Rightarrow \quad K \text{is non-negative definite.}$$
b) For $K(x, y) = K_1(x, y) + K_2(x, y)$ we have that $K = K_1 + K_2$, thus

$$z^T K z = z^T (K_1 + K_2) z = z^T K_1 z + z^T K_2 z \geq 0, \quad \forall z \in \mathbb{R}^p \Rightarrow K \text{ is non-negative definite.}$$

If a Kernel is given by $K(x, z) = (x^T z + c)^d$, we have:

$$(x^T z + c)^d = \left( \sum_{i=1}^{n} x_i z_i + c \right)^d = \sum_{\alpha_1, \ldots, \alpha_n+1} \beta(\alpha_1, \ldots, \alpha_n+1) (x_1 z_1)^{\alpha_1} \ldots (x_n z_n)^{\alpha_n} c^{\alpha_{n+1}}$$

Note that for any $i, j = 1, \ldots, n$ the matrix $\rho_i \lambda_i (v_i \odot u_j) (v_i \odot u_j)^T$ is a rank-1 matrix with eigenvalue $\rho_i \lambda_i$ non-negative definite matrix. This means that $K$ is a sum of non-negative definite matrices, therefore, as shown in b), it is also non-negative definite.

Solution of Problem 3

If a Kernel is given by $K(x, z) = (x^T z + c)^d$, we have:

$$(x^T z + c)^d = \left( \sum_{i=1}^{n} x_i z_i + c \right)^d = \sum_{\alpha_1, \ldots, \alpha_n+1} \beta(\alpha_1, \ldots, \alpha_n+1) (x_1 z_1)^{\alpha_1} \ldots (x_n z_n)^{\alpha_n} c^{\alpha_{n+1}}$$

where the sum is taken over all $(\alpha_1, \ldots, \alpha_n+1) \in \mathbb{N}^{n+1}$ such that $\sum_{i=1}^{n+1} \alpha_i = d$, $\alpha_i \in \mathbb{N}$. The number of all these monomials are given by the number of answers to the above equation which is $\binom{n+d}{d}$. Therefore the feature map can be considered as:

$$\phi(x) = \left( \sqrt{c^{\alpha_{n+1}} \beta(\alpha_1, \ldots, \alpha_n+1) x_1^{\alpha_1} \ldots x_n^{\alpha_n}} \right)_{\sum_{i=1}^{n+1} \alpha_i = d, \alpha_i \in \mathbb{N}} \in \mathbb{R}^{\binom{n+d}{d}}.$$