Problem 1. (Support Vector Machines)

a) Suppose that a training dataset is composed of vectors \( x_i \in \mathbb{R}^3 \) belonging to two classes. The class membership is indicated by the labels \( y_i \in \{-1, +1\} \). Suppose that the dataset is separable. A Support vector machine is used to find the maximum-margin hyperplane \( a^T x + b = 0 \). The primal optimization problem gives the optimal \( a^* \) as \((1 \\ 3 \\ 0)^T\). Two support vectors with different labels are given as:

\[
    x_1^T = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad x_2^T = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}
\]

Find the optimal value \( b^* \).

Consider another training dataset that is non-separable. The following dual problem is solved for a support vector machine

\[
    \max_{\lambda} \sum_{i=1}^{6} \lambda_i - \frac{1}{2} \sum_{i,j} y_i y_j \lambda_i \lambda_j x_i^T x_j \quad \text{s.t.} \quad 0 \leq \lambda_i \leq 5 \quad \text{and} \quad \sum_{i=1}^{6} \lambda_i y_i = 0.
\]

The dataset with the outputs of the optimization problem are given in the following table.

<table>
<thead>
<tr>
<th>Data</th>
<th>Label</th>
<th>Solution</th>
<th>Data</th>
<th>Label</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 = \begin{pmatrix} 1 \ 1 \end{pmatrix} )</td>
<td>( y_1 = -1 )</td>
<td>( \lambda_1^* = 0 )</td>
<td>( x_4 = \begin{pmatrix} 0.5 \ -0.5 \end{pmatrix} )</td>
<td>( y_4 = 1 )</td>
<td>( \lambda_4^* = 4.73 )</td>
</tr>
<tr>
<td>( x_2 = \begin{pmatrix} 2 \ 0 \end{pmatrix} )</td>
<td>( y_2 = -1 )</td>
<td>( \lambda_2^* = 0.67 )</td>
<td>( x_5 = \begin{pmatrix} -2 \ 1 \end{pmatrix} )</td>
<td>( y_5 = 1 )</td>
<td>( \lambda_5^* = 0.94 )</td>
</tr>
<tr>
<td>( x_3 = \begin{pmatrix} 0 \ 0 \end{pmatrix} )</td>
<td>( y_3 = -1 )</td>
<td>( \lambda_3^* = 5 )</td>
<td>( x_6 = \begin{pmatrix} 0 \ -1 \end{pmatrix} )</td>
<td>( y_6 = 1 )</td>
<td>( \lambda_6^* = 0 )</td>
</tr>
</tbody>
</table>

b) Determine the support vectors.

c) Find the maximum-margin hyperplane by finding \( a^* \) and \( b^* \).
Problem 2. *(SVM Kernel)*
Let \( K_1 : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R} \) and \( K_2 : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R} \) be valid kernels for a support vector machine. Show that

a) \( K(x, y) = \alpha K_1(x, y) \), where \( \alpha > 0 \), is also a valid kernel.

b) \( K(x, y) = K_1(x, y) + K_2(x, y) \) is also a valid kernel.

c) \( K(x, y) = K_1(x, y)K_2(x, y) \) is also a valid kernel.

Problem 3. *(Polynomial Kernel)*
Suppose that a Kernel is given by \( K(x, z) = (x^Tz + c)^d \) where \( x, z \in \mathbb{R}^p \), \( c \in \mathbb{R} \), \( d \in \mathbb{N}, d \geq 2 \). Suppose that the feature space is of dimension \( \binom{p+d}{d} \) and it contains all monomials of degree less than or equal to \( d \). Determine \( \phi(x) \).