Prof. Dr. Rudolf Mathar, Dr. Arash Behboodi, Markus Rothe

Exercise 12

Friday, January 25, 2019

Problem 1. *(Support Vector Machines)*

a) Suppose that a training dataset is composed of vectors $\mathbf{x}_i \in \mathbb{R}^3$ 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 $\mathbf{a}^{\mathrm{T}}\mathbf{x} + b = 0$. The primal optimization problem gives the optimal $\mathbf{a}^{\star}$ as $\begin{pmatrix} 1 & 3 & 0 \end{pmatrix}^{\mathrm{T}}$. Two support vectors with different labels are given as :

$$\mathbf{x}_1^{\mathrm{T}} = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix}, \qquad \mathbf{x}_2^{\mathrm{T}} = \begin{pmatrix} -1 & -1 & -1 \end{pmatrix}$$

Find the optimal value $b^{\star}$.

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

$$\max_{\boldsymbol{\lambda}} \quad \sum_{i=1}^{6} \lambda_i - \frac{1}{2} \sum_{i,j} y_i y_j \lambda_i \lambda_j \mathbf{x}_i^T \mathbf{x}_j$$

$$\text{s.t.} \quad 0 \le \lambda_i \le 5 \qquad \text{and} \qquad \sum_{i=1}^{6} \lambda_i y_i = 0.$$

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

Data	Label	Solution	Data	Label	Solution
$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$y_1 = -1$	$\lambda_1^{\star} = 0$	$\mathbf{x}_4 = \begin{pmatrix} 0.5 \\ -0.5 \end{pmatrix}$	$y_4 = 1$	$\lambda_4^{\star} = 4.73$
$\mathbf{x}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$	$y_2 = -1$	$\lambda_2^{\star} = 0.67$	$\mathbf{x}_5 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$	$y_5 = 1$	$\lambda_5^{\star} = 0.94$
$\mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$y_3 = -1$	$\lambda_3^{\star} = 5$	$\mathbf{x}_6 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$y_6 = 1$	$\lambda_1^{\star} = 0$

b) Determine the support vectors.

c) Find the maximum-margin hyperplane by finding $\mathbf{a}^{\star}$ and $b^{\star}$.

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

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

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

Problem 3. *(Polynomial Kernel)*

Suppose that a Kernel is given by $K(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^T\mathbf{z} + c)^d$ where $\mathbf{x}, \mathbf{z} \in \mathbb{R}^p$, $c \in \mathbb{R}$, $d \in \mathbb{N}$, $d \geq 2$. 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 $\phi(x)$.