Problem Set 3*

Tasha Pais†

Problem 1

Given the dual formulation of the SVM problem:

$$\max_{\alpha}\left(\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i, x_j)\right) \tag{1}$$

subject to the constraints

$$0 \leq \alpha_i \leq C, \quad \text{and} \quad \sum_{i=1}^{N}\alpha_i y_i = 0. \tag{2}$$

We can define the following to transform it into the canonical form of a quadratic program:

The objective vector v is:

$$v = -1_N, \tag{3}$$

where 1_N is an N-dimensional vector of ones.

The Hessian matrix H is:

$$H_{ij} = y_i y_j K(x_i, x_j), \tag{4}$$

which is an $N \times N$ symmetric matrix.

The inequality constraints matrix A and vector a are:

$$A = \begin{pmatrix} I_N \\ -I_N \end{pmatrix}, \quad a = \begin{pmatrix} C_N \\ 0_N \end{pmatrix}, \tag{5}$$

where I_N is the $N \times N$ identity matrix, C_N is an N-dimensional vector with all elements equal to C, and 0_N is an N-dimensional vector of zeros.

The equality constraints matrix B and vector b are:

$$B = 1_N^T y, \quad b = 0, \tag{6}$$

where y is the vector of labels y_i.

The solution u^* of the quadratic program:

$$\min_{u}\left(\frac{1}{2}u^T H u + v^T u\right) \tag{7}$$

subject to

$$Au \leq a, \quad Bu = b \tag{8}$$

will be equivalent to the solution α^* in the SVM dual problem.

*Due: November 6, 2023, Student(s) worked with: None

†NetID: tdp74, Email: tdp74@rutgers.edu

Problem 2

Given a solved Support Vector Machine (SVM) dual problem with α^* obtained, the bias term b^* can be computed as follows.

For any support vector x_k with $0 < \alpha_k^* < C$, the following condition holds:

$$y_k(\langle w^*, \phi(x_k)\rangle + b^*) = 1. \tag{9}$$

By substituting the expression for $w^* = \sum_{i=1}^{N} \alpha_i^* y_i \phi(x_i)$, we get:

$$y_k \left(\sum_{i=1}^{N} \alpha_i^* y_i \langle \phi(x_i), \phi(x_k) \rangle + b^* \right) = 1. \tag{10}$$

Given that $K(x_i, x_k) = \langle \phi(x_i), \phi(x_k) \rangle$, it simplifies to:

$$y_k \left(\sum_{i=1}^{N} \alpha_i^* y_i K(x_i, x_k) + b^* \right) = 1. \tag{11}$$

Solving for b^* gives us the exact formula:

$$b^* = y_k - \sum_{i=1}^{N} \alpha_i^* y_i K(x_i, x_k). \tag{12}$$

This formula can be used to compute b^* using any support vector x_k.