Machine Learning (DS4023) Assignment 3

Deadline: Nov. 11, 2024.

Problem 1: Hard-margin SVM. (18 pts)

You are given the following two sets of data points, each belonging to one of the two classes (class 1 and class -1):

- Class 1 (labeled as +1):
$$(1, 2), (2, 3)$$
- Class -1 (labeled as -1):
$$(2, 1), (3, 2)$$

Please find the optimal separating hyperplane using a linear SVM and derive the equation of the hyperplane. Assume the hard-margin SVM.

1. Write down the formulation of SVM, including the separation hyperplane, the constraints and the final optimization problem with parameters. **(4 pts)**
2. Write down the Lagrangian form for this problem using the parameters and Lagrange multipliers. Please also write out its dual form. **(10 pts)**
3. Assume that the Lagrangian multipliers $\boldsymbol{\alpha}_i$'s are all 0.5 and that the point $(1, 2)$ is a support vector for ease of calculation. Please calculate the values of weight vector $\boldsymbol{w}$ and bias b. Write out the explicit form of the hyperplane. **(4 pts)**

Problem 2: Soft-margin SVM. (20 pts)

Suppose we have the data points $\boldsymbol{x} \in \mathbb{R}^{n \times d}$ with corresponding labels $\boldsymbol{y} \in \mathbb{R}^n$. We want to use a soft-margin SVM to classify these data points with a regularization parameter $C = 1$.

1. Write down the formulation of soft-margin SVM for this problem using $\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}, b$ and $\boldsymbol{\xi}$. Write out explicitly their dimensions. **(3 pts)**
2. Write down the Lagrangian form and derive the dual for the problem. Write down the detailed derivation steps. **(12 pts)**
3. Obtain the decision boundary. **(3 pts)**
4. Explain why $\boldsymbol{\xi}$ disappears in the dual. **(2 pts)**

Problem 3: Kernel SVM. (17 pts)

Consider the following 2D dataset with four training points:

$$\mathbf{x}_1 = (1, 2), \quad y_1 = 1$$

$$\mathbf{x}_2 = (2, 3), \quad y_2 = 1$$

$$\mathbf{x}_3 = (3, 1), \quad y_3 = -1$$

$$\mathbf{x}_4 = (4, 3), \quad y_4 = -1$$

We want to use the **polynomial kernel** $k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^\top \mathbf{x}_j + 1)^2$ to classify these points with a soft-margin SVM. The regularization parameter $C = 1$.

1. Compute the kernel matrix K. **(6 pts)**
2. Set up the dual optimization problem. You can use the results from Problem 2. **(4 pts)**
3. Suppose the Lagrange multipliers α's are

   $$\alpha_1 = 0.0182, \quad \alpha_2 = 0.0068, \quad \alpha_3 = 0.0250, \quad \alpha_4 = 0,$$

   and $\boldsymbol{x}_3$ is a support vector. Please compute the bias term b. **(2 pts)**
4. Classify a new point $\mathbf{x}_5 = (2, 1)$ using the learned kernel SVM model. **(5 pts)**

Problem 4: Programming (45 pts)

Complete the jupyter notebook attached on programming for ensemble learning and SVM. Submit the completed file.

To submit:

1. A file containing written answers to the Problems 1 – 3.
2. The Jupyter notebook with solutions.