EE61012 – Spring Semester 2024

Class Test 4: Convex Optimization in Control and Signal Processing

Prof. Ashish Ranjan Hota
Department of Electrical Engineering, IIT Kharagpur

Q 4.1: Linear Classification with SVM

Labeled dataset $\{x^i, y^i\}_{i\in[N]}$ where each $x^i \in \mathbf{R}^4$ is associated with a label $y^i \in \{1, -1\}$ such that $y^i = 1$ if $x^i \in A$ and $y^i = -1$ if $x^i \in B$ is given. Here $N = 150$. Consider the following classification problem:

$$\begin{aligned} \min_{w\in\mathbb{R}^n, b\in\mathbb{R}} \quad & \frac{1}{2}||w||_2^2 + 10\sum_{i=1}^{N}\epsilon_i \\ \text{s.t.} \quad & 1 - y^i(w^\top\phi(x^i) + b) \le 0, \qquad \forall i \in [N]. \\ & \epsilon_i \ge 0, \qquad \forall i \in [N]. \end{aligned}$$

Using a suitable convex optimization solver, answer the following questions.

1. Let the first 120 data points be the training set and the last 30 data points be the test set.
2. Formulate and solve the dual problem and determine the optimal dual solution using the Gaussian kernel by replacing $\phi(x^i)^\top\phi(x^j)$ with

   $$k(x^i, x^j) = \exp^{-c||x^i - x^j||^2},$$

   where $c = 10$ is a constant.
3. Given the optimal dual solution λ^*, find b^*. Write a function which predicts the label of a point x given λ^*, b^* and the training data. Determine how many points from the training set are incorrectly classified. Simialarly, determine, how many points from the test set are incorrectly classified.
4. Repeat the above for $c = 1, 0.1, 0.01, 0.005, 0.001$.

Upload the code and your answers in a single pdf file.