E0 270: Machine Learning (Jan-April 2025)
Problem Sheet #2

Indian Institute of Science

Problem 1

Suppose we are given a dataset $D = \{(x_n, y_n)\}_{n=1}^N$. For solving the SVM problem on this dataset, suppose the four support vectors $x_{n_1}, x_{n_2}, x_{n_3}, x_{n_4}$ are given a priori. Is it possible to directly solve the SVM problem using these support vectors without needing to use the whole dataset? Justify.

Problem 2

Let $Z(.)$ denote the vector of monomials up to degree d, i.e.,

$$Z(x) = \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^d \end{bmatrix}. \tag{1}$$

Let Q be a symmetric positive definite matrix. Is the following function a valid kernel? Justify.

$$k(x, y) = Z(x-y)^T Q Z(x-y). \tag{2}$$

Problem 3

Let $X = [0, \pi/2]^2$. Is the following function defined on $X \times X$ a kernel? Justify.

$$k(x, y) = \cos(x_1 - y_1)\cos(x_2 - y_2), \quad \forall x = (x_1, x_2), y = (y_1, y_2) \in X. \tag{3}$$

Problem 4

ℓ_2 norm soft margin SVM:

When introducing slack variables to soft margin SVM, instead of adding ξ_n to the objective function, we can also add ξ_n^2, giving rise to the following optimization problem:

$$\min_{w,b,\xi} \quad \frac{1}{2}||w||^2 + c\sum_{n=1}^{N} \xi_n^2 \tag{4}$$

subject to

$$y_n(w^T x_n + b) \geq 1 - \xi_n, \quad n \in [N]. \tag{5}$$

(a) Compared to the standard soft margin SVM formulation, we have dropped the extra set of constraints $\xi_n \geq 0$. Show that these non-negativity constraints can be removed without affecting the optimal solution.

(b) Write the Lagrangian for the above optimization problem.

Problem 5

Consider the logistic regression model with a target variable $t \in \{-1, 1\}$. If we define $p(t = 1|y) = \sigma(y)$ where $y(x)$ is given by $y(x) = w^T \phi(x) + b$, show that the negative log-likelihood, with the addition of a quadratic regularization term, takes the form:

$$\sum_{n=1}^{N} \mathrm{E}_{LR}(y_n t_n) + \lambda \|w\|^2$$

where

$$\mathrm{E}_{LR}(y_t) = \ln\left(1 + \exp(-y_t)\right).$$