Nguyen Vu Anh Ngoc
Machine Learning 1
November 11, 2023

Homework Week 9: Support Vector Machines

Question 1. Find solution for decision hyperplane.

Consider the two-class classification problem with linear models:

$$y(x) = w^T \varphi(x) + b$$

The distance of a point x_i to the decision surface is:

$$D = \frac{t_i y(x_i)}{\|w\|}$$

We aim to maximize the margin, the perpendicular distance to the closest point x_n, by optimizing w and b:

$$\arg\max_{w,b} \left[\frac{1}{\|w\|} \min \Big(t_i (w^T \varphi(x_i) + b) \Big) \right]$$

This leads to a constrained optimization problem:

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\|w\|_2^2 \\ \text{subject to:} \quad & t_i(w^T \varphi(x_i) + b) \geq 1 \end{aligned}$$

The Lagrange formulation of this problem is:

$$L(w, b, \alpha) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{N} \alpha_i \Big(t_i(w^T \varphi(x_i) + b) - 1 \Big)$$

Taking derivatives of L with respect to w and b, we find:

$$\begin{aligned} \frac{\partial L}{\partial w} &= w - \sum_{i=1}^{N} \alpha_i t_i \varphi(x_i) = 0 \\ \frac{\partial L}{\partial b} &= -\sum_{i=1}^{N} \alpha_i t_i = 0 \end{aligned}$$

Rewriting the Lagrangian in terms of support vectors ($\alpha_i \neq 0$):

$$L = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j t_i t_j k(x_i, x_j)$$

where $k(x, x') = \varphi(x)^T \varphi(x')$.

The classifier is expressed as:

$$y(x) = \sum_{i=1}^{N} \alpha_i t_i k(x, x_i) + b$$

KKT conditions dictate:

$$\begin{aligned} \alpha_i &\geq 0 \\ t_i y(x_i) &\geq 1 \\ \alpha_i (t_i y(x_i) - 1) &= 0 \end{aligned}$$

In conclusion, the parameters are determined as:

$$w = \sum_{i=1}^{N} \alpha_i t_i \varphi(x_i)$$
$$b = \frac{1}{n_S} \sum_{i \in S} (t_i - \sum_{i \in S} \alpha_i t_i k(x, x_i))$$

with n_S being the number of elements in the set $S = \{i \mid \alpha_i \neq 0\}$.