

# Assignment 7

Vishal Kumar, MIT2019090

## Question

**What is linear SVM? Solve the problem of SVM using Quadratic Programming.**

**Ans. -**

Support Vector Machine(SVM) is a supervised machine learning algorithm which can be used for both classification or regression problems. However, it is mostly used in classification problems. In the Linear SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.

Let's assume our dataset consists N data, which is of two categories. Let's draw a linear hyper-plane that can separate them.

$$y_n = w^T x_n + b, \quad y_n \in \{1, -1\}$$

For the best hyperplane, we maximise the margin between two support vectors that's from different category.

let's assume the two points are  $x_0$  and  $x_n$ ,  
the perpendicular distance between these two will be:

$$d = \frac{|w^T x_n + b - w^T x_0 + b|}{||w||}$$

We know that for  $x_0$  that lies on the hyperplane, its corresponding prediction  $y_0$  is zero. So,

$$d = \frac{|w^T x_n + b|}{||w||} = \frac{1}{||w||}$$

Now, we have to maximize the margin to find best hyperplane.

$$d = \max\left(\frac{1}{||w||}\right) \quad \text{subject to} \quad \min_n |w^T x_n + b| = 1$$

Let's turn it to a minimization problem:

$$d = \min\left(\frac{1}{2}||w||^2\right)$$

Let's find multiplication of actual and predicted y:

actual  $y = y_n$

predicted  $y = y'_n$

$$y_n * y'_n \geq 1$$

Now, Modified equation:

$$y_n(w^T x_n + b) \geq 1$$

The primal problem using Lagrangian:

$$L(w, b) = \frac{1}{2} w^T w - \sum_{n=1}^N \alpha_n [y_n(w^T x_n + b) - 1]$$

Let's find dependent of one parameter:

$$\frac{\partial L}{\partial w} = w - \sum_{n=1}^N \alpha_n y_n x_n, \quad \frac{\partial L}{\partial b} = - \sum_{n=1}^N \alpha_n y_n$$

From these two equation, we get:

$$w = \sum_{i=1}^N \alpha_n y_n x_n, \quad \sum_{i=1}^N \alpha_n y_n = 0$$

Replace these two equations in Lagrangian:

$$L(\alpha) = \frac{1}{2} \left( \sum_{m=1}^N \alpha_m y_m x_m \right) \left( \sum_{n=1}^N \alpha_n y_n x_n \right) - \left( \sum_{n=1}^N \alpha_n y_n x_n \right) \left( \sum_{m=1}^N \alpha_m y_m x_m \right) + \sum_{n=1}^N \alpha_n$$

or,

$$L(\alpha) = \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \alpha_m \alpha_n y_m y_n x_m x_n$$

$$\text{subject to } \alpha_n \geq 0, \quad \sum_{n=1}^N \alpha_n y_n = 0$$

This is the dual problem of Linear SVM. We can also see this equation is quadratic problem that only depends on Lagrangian multiplier  $\alpha$ .