Support Vector Machine (SVM): Mathematical Notes

Let us consider a very small linearly separable dataset to solve a Hard-Margin SVM problem manually.

Dataset

Point	x_1	x_2	Class (y)
1	1	0	+1
2	2	0	+1
3	4	0	-1
4	5	0	-1

Goal: Find the optimal separating hyperplane using Hard-Margin SVM.

I. Hard-Margin SVM Formulation

Model Equation

The hyperplane is defined as:

$$\mathbf{w}^T\mathbf{x} + b = 0$$

For Hard-Margin SVM:

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$

Subject to:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$$

II. Step-by-Step Numerical Solution

Step 1: Identify Support Vectors

The closest positive and negative points are:

- Positive support vector: $(2, 0)$
- Negative support vector: $(4, 0)$

For support vectors:

$$\mathbf{w}^T\mathbf{x} + b = +1 \quad \text{(positive)}$$

$$\mathbf{w}^T\mathbf{x} + b = -1 \quad \text{(negative)}$$

Step 2: Assume Weight Vector Form

Since all points lie on the x_1 axis:

$$\mathbf{w} = (w_1, 0)$$

Step 3: Form Equations

From support vector $(2, 0)$:

$$w_1(2) + b = 1$$

From support vector $(4, 0)$:

$$w_1(4) + b = -1$$

Step 4: Solve the System

Subtract first equation from second:

$$(4w_1 + b) - (2w_1 + b) = -1 - 1$$

$$2w_1 = -2$$

$$w_1 = -1$$

Substitute into first equation:

$$-2 + b = 1$$

$$b = 3$$

Thus,

$$\mathbf{w} = (-1, 0), \quad b = 3$$

Step 5: Final Decision Boundary

Hyperplane:

$$-x_1 + 3 = 0$$

$$x_1 = 3$$

Decision rule:

$$\hat{y} = \text{sign}(-x_1 + 3)$$

If $x_1 < 3 \rightarrow$ class +1 If $x_1 > 3 \rightarrow$ class -1

Step 6: Verify Constraints

Compute $y_i(\mathbf{w}^T x_i + b)$:

- For $(1, 0)$: $(-1)(1) + 3 = 2 \Rightarrow 2 \geq 1$
- For $(2, 0)$: $(-1)(2) + 3 = 1 \Rightarrow 1 \geq 1$
- For $(4, 0)$: $(-1)(4) + 3 = -1 \Rightarrow (-1)(-1) = 1$
- For $(5, 0)$: $(-1)(5) + 3 = -2 \Rightarrow (-1)(-2) = 2$

All constraints satisfied.

III. Margin Calculation

Margin width:

$$\text{Margin} = \frac{2}{\|\mathbf{w}\|}$$

$$\|\mathbf{w}\| = \sqrt{(-1)^2 + 0^2} = 1$$

$$\text{Margin} = 2$$

Margin boundaries:

$$\mathbf{w}^T x + b = 1 \Rightarrow x_1 = 2$$

$$\mathbf{w}^T x + b = -1 \Rightarrow x_1 = 4$$

Distance between boundaries = 2 (maximum margin).

Margin Width Derivation Between Two SVM Hyperplanes

In Support Vector Machine, the optimal separating hyperplane is surrounded by two parallel supporting hyperplanes:

$$H_1 : \mathbf{w}^T\mathbf{x} + b = 1, \qquad H_2 : \mathbf{w}^T\mathbf{x} + b = -1,$$

where $\mathbf{w} \neq 0$ is the normal vector to both hyperplanes.

Goal: Derive the perpendicular distance (margin width) between H_1 and H_2.

Step 1: Distance Formula for Parallel Hyperplanes

Consider two parallel hyperplanes of the form:

$$\mathbf{w}^T\mathbf{x} + b_1 = 0 \quad \text{and} \quad \mathbf{w}^T\mathbf{x} + b_2 = 0.$$

The perpendicular distance between them is:

$$d = \frac{|b_2 - b_1|}{\|\mathbf{w}\|}.$$

This formula follows from projecting the vector between planes onto the unit normal vector.

Step 2: Rewrite H_1 and H_2 in Standard Form

Rearrange both hyperplanes:

$$H_1 : \mathbf{w}^T\mathbf{x} + b - 1 = 0 \quad \Rightarrow \quad b_1 = b - 1$$

$$H_2 : \mathbf{w}^T\mathbf{x} + b + 1 = 0 \quad \Rightarrow \quad b_2 = b + 1$$

Step 3: Substitute into Distance Formula

$$d = \frac{|(b+1) - (b-1)|}{\|\mathbf{w}\|}$$

$$d = \frac{|2|}{\|\mathbf{w}\|}$$

$$d = \frac{2}{\|\mathbf{w}\|}$$

Final Result

$$\boxed{\text{Margin Width} = \frac{2}{\|\mathbf{w}\|}}$$

Geometric Interpretation

- The vector $\mathbf{w}$ is perpendicular (normal) to both hyperplanes.
- The shortest distance between parallel planes lies along this normal direction.
- Therefore, the margin equals the projection of the separation onto the unit normal vector.
- Since the margin is $\frac{2}{\|\mathbf{w}\|}$, maximizing the margin is equivalent to minimizing $\|\mathbf{w}\|$.
- This leads to the SVM optimization problem:

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$

Discussion

- Only the closest points determine the hyperplane (Support Vectors).
- Margin maximization is equivalent to minimizing $\|\mathbf{w}\|^2$.
- Hard-margin works only when data is perfectly separable.
- For overlapping data, Soft-Margin SVM must be used.

Prepared By:
Md. Atikuzzaman
Lecturer
Department of Computer Science and Engineering
Green University of Bangladesh
Email: atik@cse.green.edu.bd