Mathematical Intuition: Margins, Support Vectors, and Hinge Loss

x: Input feature vector

y: Class label +1 or −1

w: weight vector / Omega

b: bias term

Decision score: $f(x) = w.x + b$

Margins:

$$w.x + b = 0$$

SVM constructs two margin boundaries:

$$w.x + b = +1$$
$$w.x + b = -1$$

$$\text{margin width} = \frac{2}{\|w\|} \qquad \frac{2}{1} = 2 \qquad \frac{2}{0.1} = 20$$

Let assume,

the boundary is: $-2x + 6 = 0$

Here, $w = -2$, then $|w| = 2$

$$\text{Margin} = \frac{2}{2} = 1$$

* Support vectors:

Mathematically,

$$y_i(w.x_i + b) = 1 \Rightarrow$$

Let,

We have a total 100 points

3 points sit at the boundary

$100 - 3 = 97$

* Hinge loss:

The formula:

the hige loss for a point (x_i, y_i) is:

$$\ldots = \max(0, 1 - y_i(w.x_i + b))$$

The su.

the hige loss for a

$$L_{hinge}(x_i) = max(0, 1 - y_i\langle \underline{w.x_i + b}\rangle)$$

$$= max(0, 1 - 1\times 2)$$

Numerical Example:

Let, $w = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b = -3$

Example 1: Safe point

$x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $y = +1$

Score: $f(x) = 2.2 + 1.1 - 3 = 2$

Loss: $L = max(0, -1) = 0$

Interpretations:

i) If point is correctly classified and far from the margin, loss is zero

ii) if point is correctly classified but close to the margin, loss is positive

iii) if it is misclassified, loss becomes large

* Example 2: Misclassified point

$x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $y = +1$

$score = 2\times 0 + 1\times 1 - 3 = -2$

$$Loss = max(0, 1 - (1)\times(-2))$$
$$= max(0, 3)$$
$$= 3$$