Machine Learning Principles & Applications.
Problem set - 1

1) suppose $w = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $b = -4$. Let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

solve the eq^n $w^T x + b = 0$ for the unknown vector x and fill in the missing entries below.

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + (-4) = 0$$

$$x_1 + 2x_2 + 3x_3 - 4 = 0$$
$$x_1 = -2x_2 - 3x_3 + 4$$

$$x = \begin{bmatrix} 4 - 2x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$x = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

2), 3) - code
4) - code

5) distance of sample $x^{(1)}$ from plane is $w^T x + b = 0$ is $\dfrac{|w^T x^{(1)} + b|}{\|w\|}$

Scalar Projection : $\dfrac{v w}{\|w\|}$

$v \to v_1$ (vector corresponding to 1st sample $x^{(1)}$)
$w \to [1, 2, 3]$
$v \cdot w \to w^T x^{(1)}$

Scanned with ACE Scanner

→ Scaler Projection = $\frac{w^T x^{(1)}}{\|w\|}$

→ distance of $x^{(1)}$ from the plane is the Magnitude of its scalar projection

Distance = $\left| \frac{w^T x^{(1)}}{\|w\|} \right| = \frac{|w^T x^{(1)}|}{\|w\|}$

$w^T x + b = 0$, distance of sample $x^{(1)}$ from the plane is $\frac{|w^T x^{(1)} + b|}{\|w\|}$

6) maximize $\left(\text{minimum of } \frac{w^T x^{(i)} + b}{\|w\|} \right)$

$= \text{maximize} \left(\text{minimum of } \frac{|w^T x^{(i)} + b|}{\|w\|} \right)$

= ~~minimize~~

~~= maximize (minimum of $\frac{w^T x^{(i)} + b}{\|w\|}$~~

$= \left(\frac{V}{\|w\|} \right)$

$$y^{(i)} (w^T x^{(i)} + b) \geq 1$$

7) consider the equation of the straight line

$x_2 = -2x_1 + 4$

The straight line can also be represented as a vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$y = mx + c \rightarrow m = -2 \quad c = 4$

Scanned with ACE Scanner

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ -2x_1 + 4 \end{bmatrix}$$

$$x = x_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

→ line moves in the direction of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$, starting @ point $\begin{bmatrix} 0 \\ 4 \end{bmatrix}$ when $x_1 = 0$.

8) consider the following dataset for a binary classf^n problem:

$$X = [x^{(1)} \; x^{(2)} \; x^{(3)} \; x^{(4)} \; x^{(5)}] = \begin{bmatrix} 1 & -1 & 0 & 2 & -2 \\ -1 & 1 & 4 & -3 & -2 \end{bmatrix}$$

consider the hyperplane $3x_1 - 4x_2 + 1 = 0$.

→ coefficients slope $= \frac{-m}{c} \Rightarrow \frac{-4}{}$

$$-4x_2 + 1 = -3x_1$$

$$4x_2 - 1 = 3x_1$$

$$x_2 = \frac{3}{4}x_1 + \frac{1}{4}$$

~~Slope = (-m~~ $\frac{3}{4}$

Slope = m = $\frac{3}{4}$

∴ normal vector will have slope = $-\frac{4}{3}$

$Ax + By = C$

vector normal to line = $\begin{bmatrix} A \\ B \end{bmatrix}$

$\therefore$ normal vector $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$

unit normal vector = $\frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix} \rightarrow \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}$

(b) margin = $\frac{2}{\|w\|}$

$\|w\| = \sqrt{3^2 + (-4)^2} = 5$

norm of weights vector = $\frac{2}{5}$

(c) calculate the directed distance of each sample from the hyperplane. which samples have the smallest and largest margins?

$$d = \frac{Ax_0 + By_0 + C}{\sqrt{A^2 + B^2}}$$

$$d_1 = \frac{(3\times1) + (-4) + (-1) + 1}{\sqrt{3^2 + (-4^2)}} = \frac{10}{5} = 2$$

$$d_2 = \frac{(3\times-1) + (-4\times1) + 1}{5} = \frac{-5}{5} = -1$$

$$d_3 = \frac{(3\times0) + (-4\times4) + 1}{5} = -3$$

$$d_4 = \frac{(3\times2) + (-4\times-3) + 1}{5} = \frac{17}{5}$$

Scanned with ACE Scanner

$$d_5 = \frac{(3 x - 2) + (-4 x - 2) + 1}{5} = -\frac{3}{5}$$

smallest $\rightarrow x_3 \rightarrow d_3 = -3$

largest $\rightarrow x_4 \rightarrow d_4 = 17/5$

Scanned with ACE Scanner