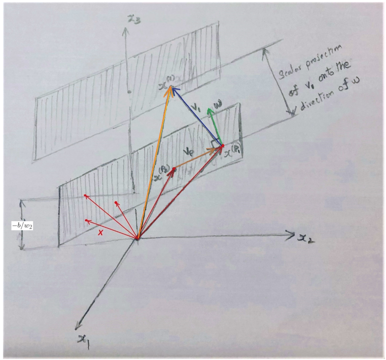
MANIPAL SCHOOL OF INFORMATION SCIENCES
MANIPAL
(A constituent unit of MAHE, Manipal)

AML5203 | Machine Learning Principles & Applications | Problem Set-1



The figure on the left hand side shows the following vectors:

$\mathbf{x}^{(1)}$: vector corresponding to the first sample in standard position,

$\mathbf{x}$: a generic vector in standard position representing a point on the bottom plane,

$\mathbf{x}^{(p_1)}$: vector corresponding to a point on the bottom plane in standard position,

$\mathbf{x}^{(p_2)}$: vector corresponding to another point on the bottom plane in standard position,

$\mathbf{v}_p$: a vector lying on the bottom plane,

$\mathbf{v}_1$: vector connecting the first point on the bottom plane and tip of $x^{(1)}$,

$\mathbf{w}$: vector from the head of $x^{(p_1)}$ such that it is perpendicular to the bottom plane.

1. Suppose $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $b = -4$. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Solve the equation $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$ for the

unknown vector $\mathbf{x}$ and fill in the missing entries below:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \left\{ x_2 \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix} + ? \begin{bmatrix} ? \\ ? \\ 1 \end{bmatrix} + \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix} \,\middle|\, ?, ? \in \mathbb{R} \right\}.$$

2. Go to the code `AML5204_MLPA_EvenSemester2024_CodingAssignment1.ipynb` and fill in the missing details in `Cell-1` for obtaining a grid of values for the components x_1, x_2, and x_3.

3. Fill in `Cell-2` for producing a scatter plot of the components created in the previous cell. The resulting points, calculated for may different values of the free variables, are, for example, represented by the head of red vectors. What is the shape of the resulting geometry formed by the points (which are solutions to the equation $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$)?

4. Using the $\mathbf{w}$ and b given above, expand the equation $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$.. There is a special relationship between the geometry of the points derived in the previous part and the vector $\mathbf{w}$. Confirm that relationship by filling in `Cell-3` and executing it several times.

5. The distance of the sample $\mathbf{x}^{(1)}$ from the plane $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$ can be seen as the magnitude of the scalar projection of the vector $\mathbf{v}_1$ onto the direction of the (normal) vector $\mathbf{w}$. Show that the distance is $\left|\mathbf{w}^{\mathrm{T}}\mathbf{x}^{(1)} + b\right| / \|\mathbf{w}\|$. This distance is also referred to as the margin of sample $\mathbf{x}^{(1)}$. Without the absolute value, it's called the directed distance of the sample.

6. Suppose we have samples $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ corresponding to output labels $y^{(1)}, y^{(2)}, \ldots, y^{(n)}$ that can be -1 or 1. In SVM, the goal is find a plane (that is, find $\mathbf{w}$ and b) that maximally separates the positive and negative samples. In other words, we want a maximum-margin separator. Suppose we decide that for positive samples $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b \geq 1$ and for negative samples $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b \leq 1$. Fill in the question marks below:

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

$$= \text{minimize } \; ?,$$

subject to the conditions $y^{(i)}\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}^{(i)} + b\right) \geq \; ? \;$ for $i = 1, \ldots, n$.

6. Consider the equation of the straight line

$$x_2 = -2x_1 + 4.$$

The straight line can also be represented as a vector

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

Write the vector $\mathbf{x}$ as follows:

$$\mathbf{x} = \boxed{?}\begin{bmatrix}\boxed{?}\\ \\ \boxed{?}\end{bmatrix} + \begin{bmatrix}\boxed{?}\\ \\ \boxed{?}\end{bmatrix}.$$

In one line, explain how the straight line can be visualized using the vector $\mathbf{x}$ above.

6. Consider the following dataset for a binary classification problem:

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

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

(a) Calculate the unit vector normal to the hyperplane.

(b) Calculate the full margin width of the hyperplane classifier.

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