ROB 101 - Fall 2021

Affine Spaces & Hyperplanes

November 22, 2021

UNIVERSITY OF MICHIGAN

Learning Objectives

- Introduce material that is assumed in UofM Computer Science courses that have Math 214 as a prerequisite.

- Provide a resource for use after you leave ROB 101.

Outcomes

- Learn how to separate $\mathbb{R}^2$ into two halves via hyperplanes.
- Orthogonal projection.
- Affine Subspace (Linear Variety).

We want to study linear structures than can be used to divide $\mathbb{R}^n$ into pieces.

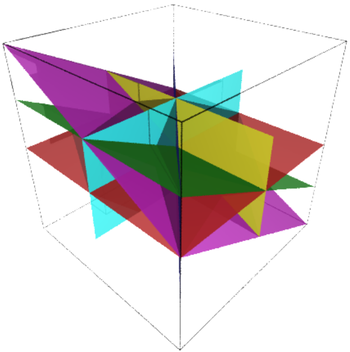


Figure: Dividing $\mathbb{R}^3$ into disjoint regions. Image from Wikimedia Commons.

Definition

Let $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ be column vectors,

$$u := \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad v := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

The dot product of u and v is defined as

$$\langle u,v \rangle = u \bullet v := \sum_{k=1}^{n} u_k v_k.$$

Remark

We note that

$$u^{\mathsf{T}} \cdot v = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{k=1}^{n} u_k v_k =: \langle u, v \rangle.$$

The dot product is also called the inner product.

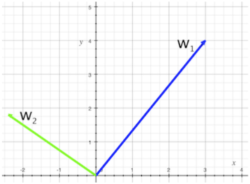
Key Use of the Inner Product of Two Vectors

The inner product will provide us a generalization of a right angle ($90\deg$ angle) between two vectors in $\mathbb{R}^n$.

$$w_1 \perp w_2 \iff \langle w_1, w_2 \rangle = 0 \iff w_1^\mathsf{T} w_2 = 0$$

(Read it as: w_1 is orthogonal to w_2 if, and only if, their inner product is zero. Orthogonal means “at right angle”)

Key Use of the Inner Product of Two Vectors



Source: https://study.com/academy/lesson/the-gram-schmidt-process-for-orthonormalizing-vectors.html

Reading the values from the graph, we have

$$w_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, w_2 = \begin{bmatrix} -\frac{7}{3} \\ \frac{7}{4} \end{bmatrix}$$

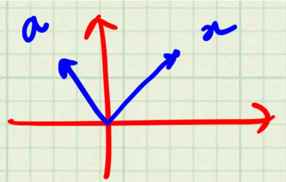
$$\implies \langle w_1, w_2 \rangle = w_1^{\mathsf{T}} w_2 = -3\frac{7}{3} + 4\frac{7}{4} = 0$$

Lines in $\mathbb{R}^2$ as Separating Hyperplanes

- Consider the set of all points, $x \in \mathbb{R}^2$ such that

$$\langle a, x \rangle = 0, \quad a \in \mathbb{R}^2.$$

- $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\langle a, x \rangle = a_1 x_1 + a_2 x_2$.



Lines in $\mathbb{R}^2$ as Separating Hyperplanes

- Consider the set of all points, $x \in \mathbb{R}^2$ such that

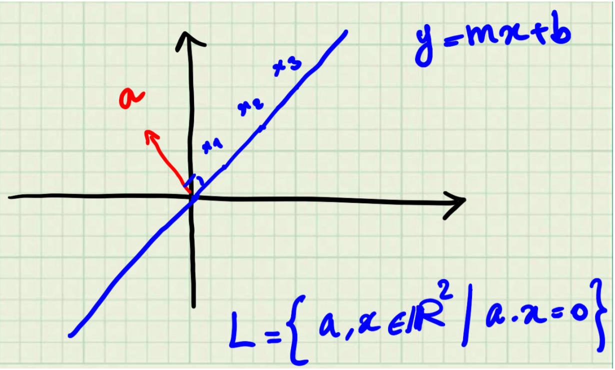
$$\langle a, x \rangle = 0, \quad a \in \mathbb{R}^2.$$

- Writing it in the set notation:

$$L = \{x \in \mathbb{R}^2 \mid \langle a, x \rangle = 0,\ a \in \mathbb{R}^2\}.$$

- L is a line that passes through the origin.

- $L = \{x \in \mathbb{R}^2 \mid \langle a, x\rangle = 0,\ a \in \mathbb{R}^2\}.$



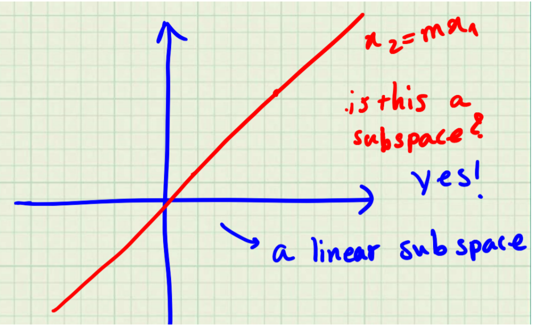
Lines in $\mathbb{R}^2$ as Separating Hyperplanes

The familiar $x_2 = mx_1 + b$ ($y = mx + b$) is called the regressor form.

- $L = \{x \in \mathbb{R}^2 \mid \langle a, x \rangle = 0,\ a \in \mathbb{R}^2\}$.
- $\langle a, x \rangle = 0 \implies m = -\frac{a_1}{a_2}$ and $b = 0$.

Lines in $\mathbb{R}^2$ as Separating Hyperplanes

- $L = \{x \in \mathbb{R}^2 \mid \langle a, x \rangle = 0,\ a \in \mathbb{R}^2\}$.
- $\langle a, x \rangle = 0 \implies m = -\frac{a_1}{a_2}$ and $b = 0$.



Let $V \subset \mathbb{R}^2$ be the set of all points that lie on a line $y = mx + b$, that is

$$V := \left\{ \begin{bmatrix} x \\ mx + b \end{bmatrix} \;\middle|\; x \in \mathbb{R} \right\}.$$

Then V is a subspace of $\mathbb{R}^2$ if, and only if, $b = 0$, that is, the line must pass through the origin.

Remark

If we view $A := a^\mathsf{T}$, then clearly $Ax = a^\mathsf{T}x = 0$ is the nullspace of A and hence is a subspace.

$$L = \{x \in \mathbb{R}^2 \mid Ax = 0\} = \mathrm{null}(A).$$

Lines in $\mathbb{R}^2$ as Separating Hyperplanes

- We can divide $\mathbb{R}^2$ into two halves.
- Indeed, we define the following half-planes.

$$H^+ := \{x \in \mathbb{R}^2 \mid \langle a, x \rangle > 0\},$$
$$H^- := \{x \in \mathbb{R}^2 \mid \langle a, x \rangle < 0\}.$$

Definition

Let $S \subset \mathbb{R}^n$ be a subset and and let $x_c \in \mathbb{R}^n$ be a vector. We define the translation of S by x_c as

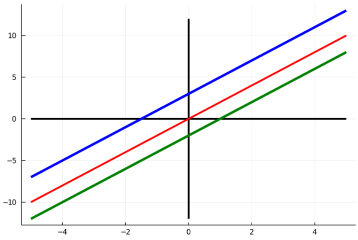
$$x_c + S := \{x_c + x \mid x \in S\}.$$

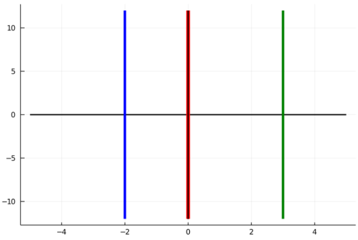
Note that, because S consists of vectors in $\mathbb{R}^n$, the addition in the above formula makes sense.

Affine Subspace (Linear Variety)

If we take a subspace and translate it by x_c, then we get an affine subspace.

- $M = x_c + L = \{x \in \mathbb{R}^2 \mid \langle a, x - x_c \rangle = 0,\ a, x_c \in \mathbb{R}^2\}$.





We can look into the translated lines as follows.

$$\langle a, x - x_c \rangle = a^\mathsf{T}(x - x_c) = 0$$
$$a^\mathsf{T} x = a^\mathsf{T} x_c =: d, \quad d \in \mathbb{R}.$$

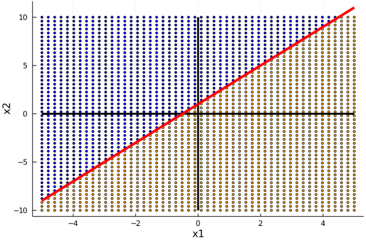
- $a^\mathsf{T} x = a_1 x_1 + a_2 x_2 = d$.
- In regressor form:

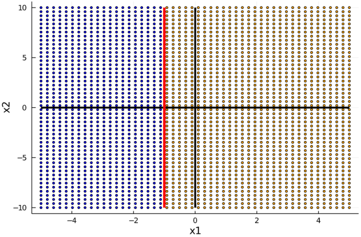
$$x_2 = -\frac{a_1}{a_2} x_1 + \frac{d}{a_2} =: m x_1 + b, \quad a_2 \neq 0.$$

Lines in $\mathbb{R}^2$ as Separating Hyperplanes

We can divide $\mathbb{R}^2$ into two halves.

$$H^+ := \{x \in \mathbb{R}^2 \mid \langle a, x - x_c \rangle > 0\},$$
$$H^- := \{x \in \mathbb{R}^2 \mid \langle a, x - x_c \rangle < 0\}.$$





Next Time

- Hyperplanes in $\mathbb{R}^n$, Quadratic Program, and Maximum Margin Classifier
- Read Chapter 13 of ROB 101 Book