Linear spaces

Joaquín Gómez

April 9, 2025

Contents

Chapter 1

Introduction

1.1 Definition of a linear space

Let V be a non-empty set of objects, called *elements*. The set V is called linear space if it satisfies the following ten axioms, which are stated in three groups.

Axioms of a linear space

Axiom 1. *Closure property of addition*: for every pair of elements $x, y \in V$, their sum is written as $z = x + y$ and $z \in V$.

Axiom 2. *Closure property of scalar product*: for any $x \in V$ and $a \in \mathbb{R}$, there is an element $z = ax \in V$.

Axioms for addition

There are four axioms of addition, we will use a number and a letter to refer to them. However if we are talking of addition properties we will simply use a letter to reference them. The same will be done for the axioms of scalar products.

Axiom 3.a. *Commutative law*: for any $x, y \in V$, we have $x + y = y + x$.

Axiom 4.b. *Associative law*: for any $x, y, z \in V$, we have $(x+y)+z = x+(y+z)$.

Axiom 5.c. *Existence of zero as an element*: there is a number in V, designated as O (big 'o'), that satisfies

$$x + O = x, \quad \forall x \in V$$

Axiom 6.d. *Opposite elements*: for all $x \in V$, the element $(-1)x$ has the property

$$x + (-1)x = O$$

Axioms for scalar product

Axiom 7.a. *Associative law*: for all $x \in V$ and every pair $a, b \in \mathbb{R}$, we have

$$a(bx) = (ab)x$$

Axiom 8.b. *Distributive law for addition in* V: for all $x, y \in V$ and $a \in \mathbb{R}$, it is true that

$$a(x + y) = ax + ay$$

Axiom 9.c. *Distributive law for addition in* $\mathbb{R}$: for any $x \in V$ and $a, b \in \mathbb{R}$, we have

$$(a + b)x = ax + bx$$

Axiom 10.d. *Existence of an identical element*: for all $x \in V$ theres an unique element I such that $Ix = x$ (commonly this element is 1. But, for example, the identical element in matrix spaces is called *identity matrix*, defined as $I = \text{diag}(1)$)

1.2 Examples of linear spaces

The following examples can be proven to be linear spaces

1. Real numbers
2. The vector space of real numbers $\mathbb{R}^n$
3. The set of all matrices
4. Polynomials P with $\deg P \leq n$ (in this case, if $\deg P = n$, we would have a problem with axioms of additions. We can't ensure the sum of two polynomials of degree n has degree n).
5. The set of all polynomials
6. The set of continuous functions on an interval $[a, b]$. This space is designated as $C(a, b)$.
7. The set of all integrable functions on an interval
8. The set of differentiable functions on an interval
9. A plane in $\mathbb{R}^3$ with the equation $ax + by + cz = 0$. Note that this plane must always go through the origin to be a linear space.

There are plenty of examples for linear spaces. We can "create" a linear space if we define addition and multiplication for that space.

1.3 Consequences of the axioms

The following theorems are a consequence of the axioms of linear space.

Theorem 1.1 (Uniqueness of 'O')**.** *In any linear space there is one and only one zero element*

Proof. Axiom 5 ensures that there is at least one 'O' in V. Now, suppose there are two zeroes in V. Let $x = O_1$ and $O_2 = O$, thus $x + O = x + O_2 = x = O_1$, but as O_1 is zero, $O_1 + O_2 = O_2$, this means that $O_1 = O_2 = O$ □

Theorem 1.2 (Uniqueness of opposites)**.** *In any linear space each x has one and only one opposite y such that $x + y = O$*

Proof. Axiom 6 ensures there is at least one opposite of x in V. Let $y_1, y_2 \in V$ be two different opposite elements for x. Then $x+y_1 = O$ and $x+y_2 = O$, then

$$(x + y_1) + y_2 = y_2 + O = y_2$$

and

$$y_1 + (x + y_2) = y_1 + O = y_1$$

Thus $y_1 + (x + y_2) = y_1 + (x + y_1) = y_1 + O = O + y_1$, this proves that $y_1 = y_2$. □

1.4 Subspaces of a linear space

Let V be a linear space and let S be a subset of V, if S is also a linear space, then we say that "S is a subspace of V".

A subset of a linear space if a subspace only if it satisfies the axioms of closure.

Theorem 1.3. *Let V be a linear space, if $S \subset V$ and $S \neq \emptyset$ satisfies the ten axioms of closure then S is a subspace of V.*

The proof for this theorem is easy, and so I discarded it.

Definition 1.1. *Let $S \subset V$, and $S \neq \emptyset$, where V is a linear space. If $x \in V$ and*

$$x = \sum_{i=1}^{k} c_i x_i$$

where $x_1, x_2, \ldots, x_k \in S$ and $c_1, c_2, \ldots, c_k \in \mathbb{R}$, is called a linear combination of elements in S. The set of linear combinations of the elements of S satisfies the axioms of closure, so it is also a subspace of V. We say that this subspace is generated by S and we call it the linear span of S, designated by $L(S)$. If $S = \emptyset$, we define $L(S) = \{O\}$.

1.5 Dependent and independent subsets of a linear space

In this section we introduce the concept of independence, that is important when working with systems of linear equations, matrices, and other subjects in linear algebra.

Definition 1.2. *Let S be a set of elements of a linear space V. S is dependent if there exists a finite set of distinct elements in $x_1, x_2, \ldots, x_k \in S$, and a set of scalars $c_1, c_2, \ldots, c_k$ where not all of them are zero, that satisfies*

$$\sum_{i=1}^{k} c_i x_i = 0$$

A set is independent if it is not dependent. So the following

$$\sum_{i=1}^{k} c_i x_i = 0, \quad \textit{implies } c_1 = c_2 = \cdots = c_k = 0$$

Independency and dependency are properties of sets of elements. However, we can apply the same concepts to the elements itself. For example, a set of vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ is called independent if there is **not** a linear combination of these vectors that produce the zero vector.

Example 1.1. *Let $u_k(t) = t^k$ for $k = 1, 2, \ldots, n$ and $t \in \mathbb{R}$. The set $V = \{u_1, u_2, \ldots, u_n\}$ is independent except in the subset S where $t = -1$ and n is odd.*

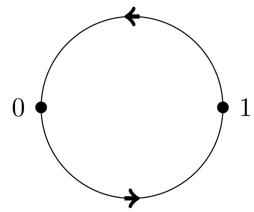
Proof. For S to be independent, there must be $c_1, c_2, \ldots, c_n$, where $c_1 = c_2 = \cdots = c_n = 0$ and

$$\sum_{k=0}^{n} c_k t^k = 0$$

To solve this, we set $c_0 = c_1 = \cdots = c_n$. If we define

$$f(t) = \sum_{k=0}^{n} c_k t^k,$$

note that $f(-1) = \begin{cases} 1 & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases}$ We can draw a picture for this problem. Imagine a circle, in which we have two points. We can travel the circumference counterclockwise starting from 1. We start from 1 because in the case $n = 0$, $f(-1) = 1$.



Now start performing counterclockwise turns, and count how many times you go from 0 to 1, and from 1 to 0.

If $C_{0\to1}$ and $C_{1\to0}$ are the counts of going from 0 to 1 and from 1 to 0 respectively, we have that if $C_{0\to1} = C_{1\to0}$, then $f(-1) = 1$. Otherwise, we must have $f(-1) = 0$.

But having $C_{0\to1} = C_{1\to0}$ and that the total count is $C = C_{0\to1} + C_{1\to0}$, means

$$C = 2C_{1\to0} = 2C_{0\to1}$$

Hence, C is an even number. To verify that S is dependent, we set $n = 2r-1$ for any integer r and $t = -1$. With the results above, we can see that

$$\sum_{k=0}^{2r-1} c_k(-1)^k = 0$$

if $c_1 = c_2 = \cdots = c_{2r-1}$, but not necessarily zero. This proves that S is dependent. $\square$

Theorem 1.4. *Let $S = \{x_1, x_2, \ldots, x_k\}$ an independent set formed by k elements of a linear space V and let $L(S)$ be the linear span of S. Then, any set of $k+1$ elements from $L(S)$ is dependent.*

What this theorem says is that, taking any set of vectors in the linear span of S, this is, formed by combining elements of S (this vectors can be of any nature), then, if we form a subset of the linear span of S and it has more elements than S itself, the set will be dependent. This is because we are not providing any new "dimension" to the new set. Say $S \in \mathbb{R}^{\nvDash}$, as we are restricted to be in $\mathbb{R}^{\nvDash}$, taking 4 vectors won't make any object in $R^{i>3}$

Proof. Let $T = \{y_1, y_2, \ldots, y_{n+1}\} \subset L(S)$, this means that each y_i is a linear combination of elements in S

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad \text{for } i = 1, 2, \ldots, n+1$$

For T to be dependent, there must be some scalar set $C = \{c_1, c_2, \ldots, c_{n+1}\}$, where not all of them are zero, that satisfies

$$\sum_{i=1}^{n+1} c_i y_i = 0$$

We now want to prove by induction that for $n-1$ elements of T, there is a linear combination that satisfies dependency. Thus, we can try to form an equation that represents T as a linear combination of $n-1$ elements. For this, we are going to take one element of T, multiply it by some scalar and subtract each element of T.

Take the 1^{st} element in T and multiply it by $c_i = \frac{a_{i1}}{a_{11}}$

$$c_i y_1 = a_{i1}x_1 + \sum_{j=2}^{n} c_i a_{1j} x_j$$

Now subtract y_1

$$\begin{aligned} c_i y_1 - y_i = a_{i1}x_1 + \sum_{j=2}^{n} c_i a_{1j} x_j - a_{i1}x1 + \sum_{j=2}^{n} a_{ij} x_j \\ = \sum_{j=2}^{n} c_i a_{1j} x_j - a_{ij} x_j \qquad (1.1) \\ = \sum_{j=2}^{n} \left(c_i a_{1j} - a_{ij}\right) x_j \end{aligned}$$

Equation (1.1) is indeed a linear combination of $n-1$ elements of S. By induction for n, we can prove that there are n scalars $t_2, t_3, \ldots, t_{n+1}$, that satisfy

$$\sum_{j=2}^{n+1} t_i \left(c_i y_1 - y_i\right) = 0 \qquad (1.2)$$

As each y_i is a linear combination of elements of S, we can write y_i in terms of y_1.

Equation (1.2) is solvable, because $y_i = c_i y_1$, this is true by the fact that $T \subset L(S)$. □

1.6 Basis and dimension

Definition 1.3. *A finite set S of elements of a linear space V is called a finite basis of V is S is independent and spans V. V is of finite dimension if it has a finite basis. Otherwise, V has infinite dimension.*

Theorem 1.5. *Let V be a linear space of finite dimension. Then any finite basis of V has the same number of elements.*

Proof. This theorem can be proved with theorem 1.4, let S and T be two finite bases for V, with k and m elements respectively. If S generates V, then V must have k elements, we know that any set of $k+1$ elements of V is dependent. Thus, T must have $m \geq k$ elements. Applying the same reasoning vice-versa yields that $k = m$. □

This does not mean that a set of $k+1$ elements of V can't span V. It states that, the number of elements for a finite basis of a linear space V of dimension k, must have the same number of elements.

Definition 1.4. *If a linear space V has a finite basis of n elements, we write $n = \dim V$.*

The following theorem will not be proven. However, it has an intuitive explanation.

Theorem 1.6. *Let V be a linear space of finite dimension, with $\dim V = n$. Then*

1. *If S is a finite basis for V, and T is a set of independent elements of V, then $T \subseteq S$.*

2. *Any set of n independent elements of V is a finite basis for V.*

1.7 Components

Let V be a linear space with $\dim V = n$, and consider an ordered basis $\{e_1, e_2, \ldots, e_n\}$. This ordered basis is considered as an n-tuple $(e_1, e_2, \ldots, e_n)$.

Definition 1.5. *An ordered basis of a linear space V is a set of elements of V that form a basis and provides information about the order of its elements.*

If $x \in V$, we can express x as a linear combination of elements of the basis

$$x = \sum_{i=1}^{n} c_i e_i \tag{1.3}$$

This ensures that there is only one representation of x, take $x = \sum_{i=1}^{n} c_i e_i$, and $x = \sum_{i=1}^{n} d_i e_i$. Then

$$\sum_{i=1}^{n} c_i e_i = \sum_{i=1}^{n} d_i e_i$$

Then $\sum_{i=1}^{n} (c_i - d_i) e_i = O$, where O is the zero vector/element of V. This means that $c_i = d_i$ for $i = 1, 2, \ldots, n$. So there is only one representation of x in V.

Chapter 2

Euclidean spaces, inner products and norms

We start the section by defining what is a Euclidean space.

Definition 2.1. *A Euclidean space is a finite-dimensional linear space that satisfies Euclidean geometry. They also are metric spaces, which are sets that have a notion of distance between its elements. Euclidean spaces are equipped with an inner product.*

Euclidean spaces have a set of properties, that were defined as axioms in *Euclid's Elements*, which are

1. If $a = b$ and $b = c$ then $a = c$ (the transitive property)
2. If $a = b$ then $a + c = b + c$ (the equal sum property)
3. If a line segment $\overline{AB}$ coincides in length and direction with $\overline{CD}$ then $\overline{AB} = \overline{CD}$.
4. The whole is greater than the part. This can be thought as: *let A and $B \subset A$ be two arbitrary sets, then A is "bigger" than B.*
5. Things that are double of the same thing are equal to each other. (This one is very obvious, consider two equal circles with radius r_1 and r_2, then we can say that $r_1 = r_2$).

2.1 Dot product and inner product

Definition 2.2. *The inner product is a function that maps two elements x and y from a linear space V to a real number. We write the inner product as (x, y).*

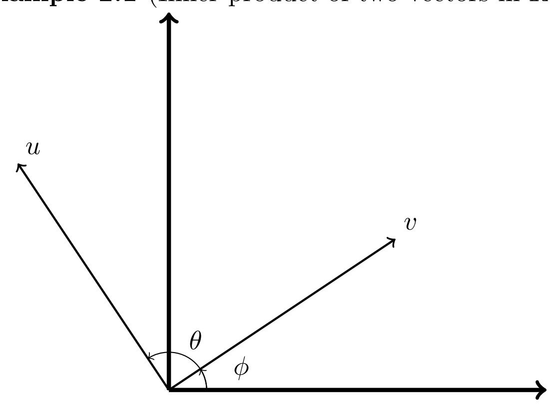
Any inner product satisfies the following properties:

1. $(x, y) = \overline{(y, x)}$ *(hermitian symmetry)*

2. $(x, y+z) = (x, y) + (x, z)$ *(linearity)*
3. $c\,(x, y) = (cx, y)$ *(homogeneity)*
4. $(x, x) \geq 0$ *(positive definite)*

Remember: *a linear space with inner product is called a Euclidean space.*

Example 2.1 (Inner product of two vectors in $\mathbb{R}^2$)**.**



Now, we have $u = (u_x, u_y)$ and $v = (v_x, v_y)$. If we define the inner product of two vectors as

$$(u, v) = u \cdot v = \sum_{i=1}^{n} u_i v_i$$

For $n = 2$, we have $u \cdot v = u_x v_x + u_y v_y$. We know the following relationships

$$\begin{aligned} \cos\theta &= \frac{u_x}{|u|}, \qquad \sin\theta = \frac{u_y}{|u|} \\ \cos\phi &= \frac{v_x}{|v|}, \qquad \sin\phi = \frac{v_y}{|v|} \end{aligned} \tag{2.1}$$

If we solve for u and v and substitute in the inner product formula, we get

$$\begin{aligned} u \cdot v = |u| \cdot |v| \cos\theta\cos\phi + |u| \cdot |v| \sin\theta\sin\phi \\ = |u||v|(\cos\theta\cos\phi + \sin\theta\sin\phi) \\ = |u||v|(\cos(\theta - \phi)) \end{aligned} \tag{2.2}$$

This means that the dot product of two vectors is the product of their length times the cosine of the angle between them. The angle between u and v is given by

$$\theta - \phi = \arccos\left(\frac{u \cdot v}{|u||v|}\right) \tag{2.3}$$

Well, this rises a question. How can you derive the inner product for a real vector space? Well, there are various points to note, but let's imagine that we want to measure the length of a vector. How can we measure distance? But also, we want to measure the distance between two vectors. Let's go with an example to make things clearer.

Example 2.2 (Distance of two vectors in $\mathbb{R}^n$)**.** *We first define two vectors* $u = (u_1, u_2, \dots, u_n)$ *and* $v = (v_1, v - 2, \dots, v_n)$

Now, a third vector, we call it $w = u - v$ *has squared length*

$$|w|^2 = (u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2 \tag{2.4}$$

Expanding one of the right-hand side terms we get $(u_i - v_i)^2 = u_i^2 - 2v_iu_i + v_i^2$*. Grouping the terms in (2.4) results in*

$$\sum_{i=1}^{n} \left(w_i^2\right) = \sum_{i=1}^{n} \left(u_i^2\right) + \sum_{i=1}^{n} \left(v_i^2\right) - 2\sum_{i=1}^{n} u_iv_i \tag{2.5}$$

Note that in (2.5) the dot product appears in the last term of the right-hand side. We can rewrite the equation as

$$w \cdot w = v \cdot v + u \cdot u - 2u \cdot v \tag{2.6}$$

And using formula (2.2)

$$|w|^2 = |v|^2 + |u|^2 - 2|u||v| \cos\theta \tag{2.7}$$

Where θ *is the angle between* u *and* v*. Note that, because the angle between a vector and itself is* $\theta = 0$*,* $\cos\theta = 1$*.*

Equation (2.7) is nothing more than the *Law of Cosines*. Now, the dot product does not follow a "natural pattern" as one would call it. Think of the exponential function, it has a very natural reasoning, for example, in the growth of populations or in differential equations. However, the dot product is present when we measure elements in Euclidean spaces, like segments or vectors.

The dot product is not "derived" in a way most things are. Instead, it is useful because it simply "appears" in measurements.

The **inner product** is a generalization of the dot product in more general spaces. Each space can have a different definition for its inner product. For example

Example 2.3 (Inner product of a functional space $C(a, b)$)**.** *Let* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *be continuous functions in an interval* $[a, b]$*, the inner product is defined as*

$$(f, g) = \int_a^b f(x)g(x)dx \tag{2.8}$$

2.2 Norms and length