Vectors and Linear Spaces

Wang Xiao

Shanghai University of Finance and Economics

March 4, 2026

What is linear algebra about?

Many objects such as buildings, furniture, etc. in our physical world are delicately constituted by counterparts of almost straight and flat shapes which, in geometrical terminology, are portions of straight lines or planes. When traveling abroad, to know the value of a foreign currency in term of one's own, we need to solve the equation like $y = ax$ and $y = ax + b$ if trading cost is considered.
In geometry, the most fundamental and essential ideas are

- (directed) line segment
- parallelogram
- angle between segments.

The algebraic equivalence is linear equations such as

$$a_{11}x_1 + a_{21}x_2 = b_1.$$

The prominent goal of "linear algebra" is how to determine whether such linear equations have a solution or solutions, and if so, how to find them in an effective way. It is common that one needs to say something about the "shape" of the solutions of certain equations. Recall that in secondary school, we already have learnt the solution set of $x^2 + y^2 = r^2$ is a circle and $ax^2 + by^2 = r^2$ is an ellipse or hyperbola. Portion of linear algebra is to establish a theory to describe the solution set of linear equations.

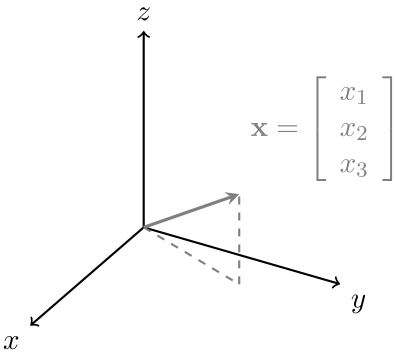
Linear algebra provides a beautiful example of the interplay between two branches of mathematics, geometry and algebra. Moreover, it provides the foundations for all of our parallel work with calculus, which is based on the idea of approximating the general function locally by a linear one. In this chapter, we introduce the basic language of vectors, linear functions, and matrices.

Vectors in $\mathbb{R}^n$

A point in $\mathbb{R}^n$ is an order n-tuple of real numbers, written $(x_1, ..., x_n)$. To it we may associate the vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

which we visualize geometrically as the arrow pointing from the origin to the point. We denote by $\mathbf{0}$ the vector all of whose coordinates are 0, called the zero vector.



More generally, any two points A and B in space determine the arrow pointing from A to B, we can denote $\overrightarrow{AB}$. If $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ and $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$

then $\overrightarrow{AB}$ is equal to the vector $v = \begin{bmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{bmatrix}$, whose tail is at the origin.

The Pythagorean Theorem tells us that when $n = 2$ the length of the vector $\mathbf{v}$ is $\sqrt{x_1^2 + x_2^2}$. A repeated application of the Pythagorean Theorem leads to the concept of "length" in higher dimension.

Definition

We define the length of the vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ to be

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

We say $\mathbf{x}$ is a unit vector if it has length 1.

There are two crucial algebraic operations we can perform on vectors, both of which have clear geometric interpretations.

Scalar multiplication:

If c is a real number and $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a vector, then we define $c\mathbf{x}$ to be the vector $\begin{bmatrix} cx_1 \\ \vdots \\ cx_n \end{bmatrix}$. Note that $c\mathbf{x}$ points in either the same direction as $\mathbf{x}$ or the opposite direction, depending on whether $c > 0$ or $c < 0$, respectively. Thus, multiplication by the real number c simply stretches or shrinks the vector by the factor of $|c|$ and reverses its direction when c is negative. Since this is a geometric "change of scale", we refer to the real number c as a *scalar* and the multiplication $c\mathbf{x}$ as *scalar multiplication*.

Note that whenever $\mathbf{x} \neq \mathbf{0}$ we can find a unit vector with the same direction by taking

$$\frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{1}{\|\mathbf{x}\|}\mathbf{x}$$

this is called *normalization* of $\mathbf{x}$. The following definition of *parallel* based on scalar multiplication extends the concept of parallelism from Euclidean geometry.

Definition

We say two vectors $\mathbf{x}$ and $\mathbf{y}$ are *parallel* if one is a scalar multiple of the other, i.e., if there is a scalar c so that $\mathbf{y} = c\mathbf{x}$ or $\mathbf{x} = c\mathbf{y}$. We say $\mathbf{x}$ and $\mathbf{y}$ are *nonparallel* if they are not parallel.

Vector addition:

If $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, the we define

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

In $\mathbb{R}^2$, we can move $\mathbf{y}$ so that its tail is at the head of $\mathbf{x}$, and draw the arrow from the origin to its head. This is so-called parallelogram law for vector addition. Geometrically, $\mathbf{x} + \mathbf{y}$ is the diagonal of the parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$. Furthermore, Euclidean geometry makes it clear that vector addition is commutative, i.e.,

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}.$$

Subtraction of one vector from another is easy to define algebraically. If $\mathbf{x}$ and $\mathbf{y}$ are as above, then we set

$$\mathbf{x} - \mathbf{y} = \begin{bmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{bmatrix}$$

As in the case with real numbers, we have the following interpretation of the difference $\mathbf{x} - \mathbf{y}$: It is the vector we add to $\mathbf{y}$ in order to obtain $\mathbf{x}$, i.e.,

$$(\mathbf{x} - \mathbf{y}) + \mathbf{y} = \mathbf{x}.$$

Exercises

1. Given $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, calculate the following operations.

- $\mathbf{x} + \mathbf{y}$
- $\mathbf{x} - \mathbf{y}$
- $\mathbf{x} + 2\mathbf{y}$
- $\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}$
- $\mathbf{y} - \mathbf{x}$
- $2\mathbf{x} - \mathbf{y}$
- $\|\mathbf{x}\|$
- $\frac{\mathbf{x}}{\|\mathbf{x}\|}$

2. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$. Describe the vectors $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$, where $s + t = 1$. What can you say about the location of $\mathbf{x}$ when $s \geq 0$ and $\mathbf{t} \geq 0$.

3. Verify both algebraically and geometrically that the following properties of vector arithmetic hold.

(a) For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

(b) For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.

(c) $\mathbf{0} + \mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.

(d) For each $\mathbf{x} \in \mathbb{R}^n$, there is a vector $-\mathbf{x}$ so that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

(e) For all $c, d \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, $c(d\mathbf{x}) = (cd)\mathbf{x}$.

(f) For all $c \in \mathbb{R}^n$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$.

(g) For all $c, d \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$.

(h) For all $\mathbf{x} \in \mathbb{R}^n$, $1\mathbf{x} = \mathbf{x}$.

Linear Spaces

Definition

V is called a linear space over the real numbers $\mathbb{R}$ (or over a field $\mathbb{F}$) if there is an *addition* $+$ on V, a *zero element* $\mathbf{0}$ in V and a *scalar multiplication* $x \to \lambda x$ with $1x = x$ in V. Additionally, we want that every x in V can be *negated*; this additive inverse element $-x$ satisfying $x + (-x) = \mathbf{0}$. The zero element $\mathbf{0}$ is required to satisfy $x + \mathbf{0} = x$ for all x. With *addition*, we mean an operation which satisfies the *associativity* law

$$(x + y) + z = x + (y + z)$$

the *commutativity* laws $x + y = y + x$ $\lambda x = x\lambda$ and the *distributivity* laws $\lambda(x + y) = \lambda x + \lambda y$, $\lambda\mu(x + y) = \lambda(\mu x + \mu y)$.

Example

- $\mathbb{R}$ is a linear space over the field of real numbers $\mathbb{R}$.
- The set of real number arrays of length n, $\mathbb{R}^n$, is a linear space over the field of real numbers $\mathbb{R}$.

Field

In abstract algebra, a *field* is a fundamental algebraic structure that generalizes the arithmetic properties of familiar number systems like the rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$, and complex numbers $\mathbb{C}$.

Definition

A *field* is a set $\mathbb{F}$ equipped with two binary operations, typically called *addition* $(+)$ and *multiplication* $(\cdot)$, satisfying the following axioms:

- Addition Axioms: $(\mathbb{F}, +)$ is an abelian (commutative) group
    1. Closure: for all $a, b \in \mathbb{F}$, $a + b \in \mathbb{F}$
    2. Associativity: for all $a, b, c \in \mathbb{F}$, $(a + b) + c = a + (b + c)$
    3. Commutativity: for all $a, b \in \mathbb{F}$, $a + b = b + a$
    4. Identity element: there exists $0 \in \mathbb{F}$ s.t. for all $a \in \mathbb{F}$, $a + 0 = a$
    5. Inverse element: for all $a \in \mathbb{F}$, there exists $-a \in \mathbb{F}$ s.t. $a + (-a) = 0$

- Multiplication Axioms, $(\mathbb{F} \setminus \{0\}, \cdot)$ is an abelian group
    1. Closure: for all $a, b \in \mathbb{F}$, $a \cdot b \in \mathbb{F}$
    2. Associativity: for all $a, b, c \in \mathbb{F}$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
    3. Commutativity: for all $a, b \in \mathbb{F}$, $a \cdot b = b \cdot a$
    4. Identity element: there exists $1 \in \mathbb{F}$ with $1 \neq 0$ such that for all $a \in \mathbb{F}$, $a \cdot 1 = a$
    5. Inverse element: for all $a \in \mathbb{F}$, $a \neq 0$, there exists $a^{-1} \in \mathbb{F}$ such that $a \cdot a^{-1} = 1$
- Distributivity between Addition and Multiplication
    1. Distributive law: for all $a, b, c \in \mathbb{F}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

Example

- Rational numbers $\mathbb{Q}$
- Real numbers $\mathbb{R}$
- Complex numbers $\mathbb{C}$
- Integer modulo a prime p: $\mathbb{Z}_p$

Dot Product

We discuss one of the crucial constructions in linear algebra, the dot product $\mathbf{x} \cdot \mathbf{y}$ of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. By definition the dot product of two vectors are given as

Definition

Given vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define their dot product is

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Proposition (Proof as an exercise)

The dot product has the following properties:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *(dot product is commutative);*
2. $\mathbf{x} \cdot \mathbf{x} = \|x\|^2 \geq 0$ *and* $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$*;*
3. $c\mathbf{x} \cdot \mathbf{y} = c(\mathbf{x} \cdot \mathbf{y})$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$*;*
4. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ *for all* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ *(the distributive property).*

Corollary

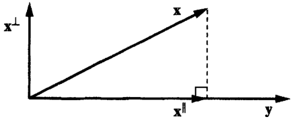
$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x}\cdot\mathbf{y} + \|\mathbf{y}\|^2.$$

Proof.

... □

The geometric meaning of this result comes from the Pythagorean Theorem: when $\mathbf{x}$ and $\mathbf{y}$ are perpendicular vectors in $\mathbb{R}^2$, then we have $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, and so, by the corollary, it must be the case that $\mathbf{x}\cdot\mathbf{y} = 0$. For this reason, we say $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are *orthogonal* if $\mathbf{x}\cdot\mathbf{y} = 0$.

With these definitions, we proceed to a construction that will be important in much of our future work. Starting with two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, where $\mathbf{y} \neq \mathbf{0}$. Following figure suggests that we are able to write $\mathbf{x}$ as the sum of a vector, $\mathbf{x}^{\|}$, that is parallel to $\mathbf{y}$ and a vector, $\mathbf{x}^{\perp}$, that is orthogonal to $\mathbf{y}$.



Let's suppose we have such an equation:

$$\mathbf{x} = \mathbf{x}^{\|} + \mathbf{x}^{\perp},$$

where $\mathbf{x}^{\|}$ is a scalar multiple of $\mathbf{y}$ and $\mathbf{x}^{\perp}$ is orthogonal to $\mathbf{y}$. To say that $\mathbf{x}^{\|}$ is a scalar multiple of $\mathbf{y}$ means that we can write $\mathbf{x}^{\|} = c\mathbf{y}$ for some scalar c.

If such an expression exists, we can determine c by taking the dot product of both sides of the equation with y:

$$\mathbf{x} \cdot \mathbf{y} = (\mathbf{x}^{\parallel} + \mathbf{x}^{\perp}) \cdot \mathbf{y} = (\mathbf{x}^{\parallel} \cdot \mathbf{y}) + (\mathbf{x}^{\perp} \cdot \mathbf{y}) = \mathbf{x}^{\parallel} \cdot \mathbf{y} = (c\mathbf{x}) \cdot \mathbf{y} = c \left\| \mathbf{y} \right\|^2.$$

This means that

$$c = \frac{\mathbf{x} \cdot \mathbf{y}}{\left\| y \right\|^2}, \quad \text{and so } \mathbf{x}^{\parallel} = \frac{\mathbf{x} \cdot \mathbf{y}}{\left\| y \right\|^2} \mathbf{y}.$$

The vector $\mathbf{x}^{\parallel}$ is called the *projection* of $\mathbf{x}$ *onto* $\mathbf{y}$, written $\text{proj}_{\mathbf{y}} \mathbf{x}$.

Definition

Let $\mathbf{x}$ and $\mathbf{y}$ be nonzero vectors in $\mathbb{R}^n$. We define the *angle* between them to be the unique θ satisfying $0 \leq \theta \leq \pi$ so that

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\left\| \mathbf{x} \right\| \left\| \mathbf{y} \right\|}.$$

In order to convince ourselves that the geometric intuition in higher dimension is actually correct, we should check algebraically that this definition make sense. Since we know by common sense that if θ really represents an "angle", then it must hold that $|\cos\theta| \leq 1$, the following result gives us what is needed.

Proposition

Cauchy-Schwarz Inequality If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *then*

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| \, .$$

Moreover, equality holds if and only if one of the vectors is a scalar multiple of the other.

One of the most useful applications of this result is the famed *triangle inequality*, which tells us that the sum of the lengths of two sides of a triangle cannot be less than the length of the third.

Corollary

For any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *we have* $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \, .$

Subspaces of $\mathbb{R}^n$

We now proceed in our study of "linear objects" that generalize lines and planes through the origin in low dimensional space such as $\mathbb{R}^2$ and $\mathbb{R}^3$.

Definition

A set $V \subset \mathbb{R}^n$ is called a *subspace* of $\mathbb{R}^n$ if it satisfies the following properties:

1. $\mathbf{0} \in V$ (the zero vector belongs to V);
2. whenever $\mathbf{v} \in V$ and $c \in \mathbb{R}$, we have $c\mathbf{v} \in V$ (V is closed under scalar multiplication);
3. whenever $\mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{v} + \mathbf{w} \in V$ (V is closed under addition).

Example

Fix a nonzero vector $A \in \mathbb{R}^n$, and consider

$$V = \{\mathbf{x} \in \mathbb{R}^n : A \cdot \mathbf{x} = 0\}.$$

Check that V is a subspace of $\mathbb{R}^n$.

Example

Let $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, explain why the following sets S are not subspaces.

$$S = \{\mathbf{x} : x_2 = 2x_1 + 1\} \quad S = \{\mathbf{x} : x_1 x_2 = 0\} \quad S = \{\mathbf{x} : x_2 \geq 0\}.$$

Given a collection of vectors in $\mathbb{R}^n$, it is natural to try to "build" a subspace from them.

Definition (Linear combination)

Let $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$. If $c_1, ..., c_k \in \mathbb{R}$, the vector

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k$$

is called a *linear combination* of $\mathbf{v}_1, ..., \mathbf{v}_k$. The set of all linear combinations of $\mathbf{v}_1, ..., \mathbf{v}_k$ is called their *span*, denoted $\text{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$.

The following result shows that the span has a subspace structure.

Proposition

Let $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$*. Then* $V = \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$ *is a subspace of* $\mathbb{R}^n$*.*

Proof.

Exercise. □

Definition

Let V and W be subspaces of $\mathbb{R}^n$. We say they are *orthogonal subspaces* if every element of V is orthogonal to every element of W, i.e., if

$$\mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for every } \mathbf{v} \in V \text{ and every } \mathbf{w} \in W.$$

Given a subspace $V \subset \mathbb{R}^n$, define

$$V^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{v} = 0 \text{ for every } \mathbf{v} \in V\}.$$

$V^\perp$ is called the *orthogonal complement* of V. Show the following result.

Proposition

$V^\perp$ is also a subspace of $\mathbb{R}^n$.