1

Vector spaces

> The purpose of abstraction is not to be vague, but to create a new semantic level in which one can be absolutely precise.
>
> Edsger W. Dijkstra

1.1 Vectors

1.1.1 Basic definitions

Vectors are tuples of numbers to which two types of operations can be applied: sum and scaling.

Vectors with n components will be denoted n-vectors. We will denote the set of all n-vectors as $\mathbb{R}^n$.

We will represent n-vectors either as vertical stacks of n numbers put between brackets

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

or between parentheses with comma-separated values

$$(a_1, a_2, \dots, a_n)$$

Each number in the tuple will be referred to as an entry or component of the vector

The zero-vector $(0, 0, \dots, 0)$ will be represented as $\vec{\mathbf{0}}$.

1.1.2 Sum of vectors

Two n-vectors can be summed together to give another n-vector, doing the numerical sum component by component:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}.$$

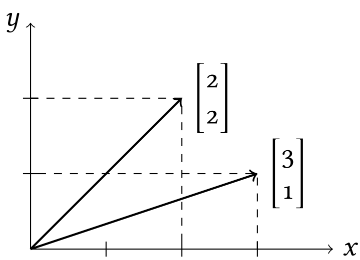
1.1.3 Scaling

Given a number $\alpha \in \mathbb{R}$, which we will often refer to as the *scaling factor*, and an n-vector, we can combine the two to produce another n-vector:

$$\alpha \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{bmatrix}$$

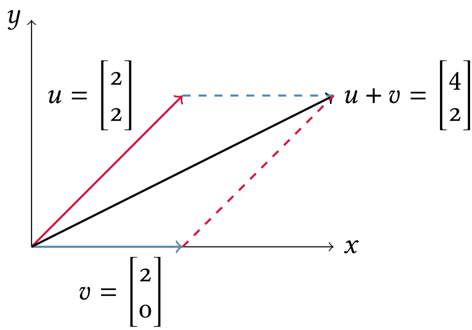
1.1.4 Graphical representation

Vectors can be represented in a Cartesian reference frame as arrows with stem at the origin of the reference frame O and tip at the point that has Cartesian coordinates given by the very entries of the vector:

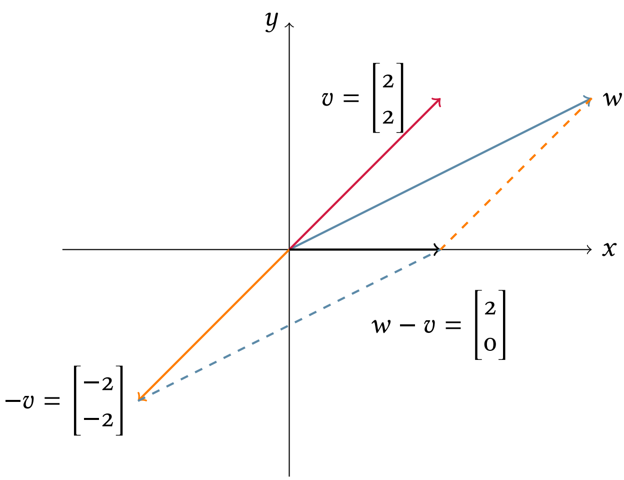


From now on you can think of vectors either as tuples or as arrows with stem at the origin in a Cartesian reference frame. We will refer to those paradigms as the *tuple perspective* and the *arrow perspective*.

Under the arrow perspective the sum of vectors follows the so-called *parallelogram rule*. Here is a graphical summary of the rule:



The same applies if we want to substract vectors. In the end of the day, substraction amounts to summation with a vector that has been scaled by -1.



Under the arrow perspective, scaling turns the input vector into an new stretched/contracted version, reversed if the scaling factor is negative.

1.1.5 Length of a vector

We define the *length* of an n-vector $u = (u_1, \ldots, u_n)$, denoted $\|u\|$, as the length of its associated arrow. Using Pythagoras theorem, this can be given as

$$\|u\| = \sqrt{u_1^2 + \ldots + u_n^2}$$

Can you see why?

1.2 Vector Spaces

We will define *vector space* as any collection of vectors $V \subset \mathbb{R}^n$ for some n that satisfies the following conditions:

1. $\vec{\mathbf{0}} \in V$
2. If $u, v \in V$ then $u + v \in V$
3. If $u \in V$ then $\lambda u \in V$ for any $\lambda \in \mathbb{R}$

1.2.1 Examples of vector spaces

- The set $\mathbb{R}^n$ of all n-vectors is a vector space for any $n \geq 1$.

- If S is a subset of $\mathbb{R}^n$ that satisfies the following linear equation:

$$x_1 + x_2 + \ldots + x_n = 0$$

then it is a vector space.

1. $\vec{0} = (0, 0, \ldots, 0) \in S$ because $0 + 0 + \ldots + 0 = 0$

2. if $a_1 + a_2 + \ldots + a_n = 0$ and $b_1 + b_2 + \ldots + b_n = 0$ then $(a_1 + b_1) + (a_2 + b_2) + \ldots + (a_n + b_n) = 0$

3. if $a_1 + a_2 + \ldots + a_n = 0$ then $\alpha a_1 + \alpha a_2 + \ldots + \alpha a_n = 0$

- An *homogenous linear equation* is a linear equation that has zero as independent term. A subset $S \subset \mathbb{R}^n$ satisfying any system of homogeneous linear equations is a vector space. For example the set of vectors $(x_1, x_2, x_3) \in \mathbb{R}^3$ that satisfy the following system of homogenous linear equations

$$\begin{cases} 2x_1 + x_2 - x_3 = 0 \\ x_2 + x_3 = 0 \\ -2x_1 + 2x_3 = 0 \end{cases}$$

is a vector space. You can try to prove it, along the same lines of the previous example, and convince yourself about the general claim.

If V is some vector space and $S \subseteq V$ is yet another vector space sitting inside V, we say that S is a *vector subspace* of V.

1.3 Linear Combinations

Given vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ and scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$, we can combine them as follows

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k$$

to produce another vector.

This way of combining the vectors by scaling them first, then summing across all the resulting vectors, is known as a *linear combination.*

1.3.1 Span

Given a collection of n-vectors $v_1, v_2, \ldots, v_k$ we can define a new vector space taking the set of all possible linear combinations of $v_1, v_2, \ldots, v_k$: we denote it $\mathbf{span}(v_1, v_2, \ldots, v_k)$.

Can you see why $\mathbf{span}(v_1, v_2, \ldots, v_k)$ is a vector space?

1.4 Linear independence

A set of vectors $v_1, v_2, \ldots, v_k$ is a *linearly independent set* – often we simply say that they are *linearly independent* – if they are not linearly dependent. In other words, none of the vectors of the set can be expressed as a linear combination of the others.

There is an equivalent definition that is often more useful for practical computation: a set of vectors $v_1, v_2, \ldots, v_k$ is a *linearly independent set* if and only if the only possible scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ that satisfy the equation

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = \overrightarrow{\mathbf{0}}$$

are $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

A set of vectors is a *linearly dependent set* – often we simply say that they are *linearly dependent* – if it is not a linearly independent set, i.e., if some vector can be expressed as a linear combination of the others.

For example, the n-vectors v_1, v_2, v_3 form a linearly dependent set if $v_2 = 2v_1 + 3v_3$.

1.5 Generating set

Given a set of n-vectors $\mathscr{V} = \{v_1, v_2, \ldots, v_k\}$, we say that they are a *generating set* of the vector space V if

$$\mathbf{span}(\mathscr{V}) = V.$$

A collection of vectors $\mathscr{U}$ of V is a *minimal generating set* of V if it meets the following two conditions:

- $\mathscr{U}$ is a generating set of V, i.e. $\mathbf{span}(\mathscr{U}) = V$
- no subset of $\mathscr{U}$ can be a generating set of V

Minimal generating sets can be thought of as generating sets where all possible redundancies have been already ruled out. This might remind us of the idea of linear independence and in fact both are very much related. It turns out that if we are given a (finite) generating set for a vector space V we can prune it repeatedly, one vector at a time, until we attain a minimal generating set that must be a linearly independent set.

Proposition 1. *If a vector u_k can be expressed as a linear combination of the vectors $u_1, \ldots, u_{k-1}$ then*

$$\boldsymbol{span}(u_1, \ldots, u_{k-1}) = \boldsymbol{span}(u_1, \ldots, u_{k-1}, u_k)$$

Let's assume that $u_k = \lambda_1 u_1 + \ldots + \lambda_{k-1} u_{k-1}$. Then any vector v in $\mathbf{span}(u_1, \ldots, u_{k-1}, u_k)$ can be expressed as a linear combination of $u_1, \ldots, u_{k-1}$. This is left as an exercise to the reader.

Proposition 2. *If $\mathscr{U}$ is a minimal generating set, it must be a linearly independent set.*

To prove this we will employ a logical trick which consists of proving the contrapositive version of the claim, i.e. we will prove that if $\mathscr{U}$ is a linearly dependent set then it cannot be a minimal generating set. If $\mathscr{U}$ is a linearly dependent set, then some of its vectors, say u_k, can be expressed as a linear combination of the others. But this means that if we remove u_k from $\mathscr{U}$ the span of the remaining smaller vector set $\mathscr{U}' = \mathscr{U} \setminus \{u_k\}$ will continue to be the same – as per the previous result.

Proposition 3. *If $\mathscr{U}$ is a maximal linearly independent set of V, then $\mathscr{U}$ spans V.*

If $\mathscr{U}$ is a linearly independent set that does not span V, it means that there is at least a vector v in V such that it cannot be expressed as a linear combination of the vectors in $\mathscr{U}$. Therefore, we can create an even larger linearly independent set $\mathscr{U}' = \mathscr{U} \cup \{v\}$, meaning that $\mathscr{U}$ was not maximal.

1.6 Basis of a vector space

Minimal generating sets are also known as *bases* (plural form of *basis*).

Taking into account everything that we have discussed so far, it is easy to check that a set of vectors $\mathscr{V} = \{v_1, v_2, \ldots, v_d\}$ is a *basis* of the vector space V when the following conditions hold:

1. $\mathscr{V}$ is a linearly independent set;

2. $\mathscr{V}$ is a generating set of V, i.e. $\mathbf{span}(\mathscr{V}) = V$.

And conversely, if the conditions 1 and 2 hold, then the set $\mathscr{V}$ must be a basis of V. So we can take this conditions as an alternative definition of *basis* of a vector space.

1.6.1 Examples

- Canonical basis of $\mathbb{R}^3$ $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$

- Canonical basis of $\mathbb{R}^n$ $(e_i)_j = \delta_{ij}$ where $\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ otherwise.

- Vector spaces have many bases (infinitely many).

Can two bases have a different number of vectors? Somewhat suprisingly, the answer is no.

Proposition 4. *Vector spaces have several (infinitely many) bases, but all the bases for a given vector space have the same number of elements.*

We will reason by contradiction. Suppose that we could find two basis of V with different number of elements, $\mathscr{U} = \{u_1, \ldots, u_n\}$ and $\mathscr{V} = \{v_1, \ldots, v_m\}$ with $n < m$. Since both sets are bases, both span V.

Given the fact that $\mathscr{U}$ spans V, v_1 can be expressed as a linear combination of the elements of $\mathscr{U}$:

$$v_1 = \lambda_1 u_1 + \ldots + \lambda_n u_n$$

Not all the coefficients can be zero, because $v_1 \neq \vec{0}$, since it is part of the linearly independent set $\mathscr{V}$. We can assume $\lambda_1 \neq 0$ – reindeixing if necessary – which means that we can express u_1 as a linear combination of

$$\mathscr{U}_1 = \{v_1, u_2, \ldots, u_n\}$$

(why?) which in turn implies that $\mathscr{U}_1$ spans V.

Now we claim that we can continue to replace elements of $\mathscr{U}$ by elements of $\mathscr{V}$ and yet have a generating set of V. Imagine that we have accomplished a generating set of V of this form:

$$\mathscr{U}_k = \{v_1, \ldots, v_k, u_{k+1}, \ldots, u_n\}.$$

Since this is a generating set of V, we can express v_{k+1} as a linear combination of the elements of $\mathscr{U}_k$:

$$v_{k+1} = \mu_1 v_1 + \ldots + \mu_k v_k + \alpha_{k+1} u_{k+1} + \ldots + \alpha_n u_n$$

Not all the coefficients α_i can be zero, because that would imply that $\mathscr{V}$ is a linearly dependent set, so we can assume $\alpha_{k+1} \neq 0$ – reindexing if necessary – which means that we can express u_{k+1} as a linear combination of

$$\mathscr{U}_{k+1} = \{v_1, \ldots, v_{k+1}, u_{k+2}, \ldots, u_n\}$$

which in turn implies that $\mathscr{U}_{k+1}$ spans V.

If we continue to replace elements of $\mathscr{U}$ by elements of $\mathscr{V}$ until we run out of elements of $\mathscr{U}$, we will reach to that conclusion that the set $\{v_1, \ldots, v_n\}$ spans V, but this is in contradicion with the fact that $\mathscr{V} = \{v_1, \ldots, v_m\}$ is a basis with $n < m$.

This common number of elements in any basis of a given vector space V is known as the *dimension* of V and it is denoted $\mathbf{dim}(V)$.

1.6.2 Coordinates of a vector in a basis

Consider a basis $\mathscr{V} = \{v_1, \ldots, v_n\}$ of a vector space V and a vector $v \in V$. Because $\mathscr{V}$ is a basis, we know two things:

- because $\mathscr{V}$ spans V, v can be expressed as a linear combination of the elements of $\mathscr{V}$;

- because $\mathscr{V}$ is a linearly independent set, there is at most one way to express a vector as a linear combination of the elements of $\mathscr{V}$.

It follows that every vector of V can be expressed, in a unique way, as a linear combination of the elements of $\mathscr{V}$. In other words, if

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

then $(\lambda_1, \ldots, \lambda_n)$ completely identifies v if we know what is the basis we refer to and the order in which we consider the elements of the basis. Can you tell why?

The scalars $\lambda_1, \ldots, \lambda_n$ are referred to as the *coordinates* of v in the basis $\mathscr{V}$.

Note that, although we did not talk about it, our default representation of vectors correspond to their coordinates in the canonical basis $\mathscr{C} = e_1, \ldots, e_n$. The entries of a vector are its coordinates in the canonical basis.

1.7 TL;DR

Vectors are tuples with numeric entries that can represent data from a tabular environment – rows or columns of some dataframe. Vectors can be scaled and vectors of equal size can summed together. Using this two operations we can define linear combinations, which gives a way to combine a finite collection of vectors into a new vector.

We like to think about vectors as making part of collections known as vector spaces. Vector spaces arise naturally in a number of common scenarios like when we try to solve a system of linear equations or when we try to assert dependencies between rows or columns of data. The dimension of a vector space is a measure of the size of the vector space, accounting for how big linearly independent sets fitting inside the vector space can be. Intuitively, the dimension of a vector space reminds us of how much information or diversity or degrees of freedom the vector space comprises.