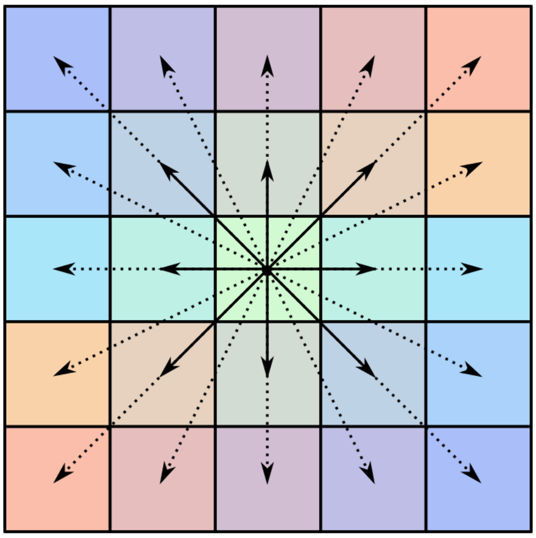
Gymnázium Evolution Jižní Město



Intro-ish To Linear Algebra

Madam Pekáč

March 20, 2025

Preface

This text covers selected topics from the curriculum of a typical undergraduate linear algebra course. Almost no pre-existing knowledge is strictly required save a superficial understanding of propositional logic and set theory. A reasonably good ability to manipulate algebraic expressions should prove advantageous, too.

Mathematics is an exact and rigorous language. Words and symbols have singular, precisely defined, meaning. Many students fail to grasp that intuition and imagination are paramount, but they serve as a *starting point*, with formal logical expression being the end. For example, an intuitive understanding of a *line* as an infinite flat 1D object is pretty much correct but not *formal*. It is indeed the formality of mathematics which puts many students off. Whereas high school mathematics is mostly algorithmic and non-argumentative, higher level maths tends to be the exact opposite – full of concepts and relations between those, which one is expected to be capable of grasping and formally describing. Owing to this, I wish this text would be a kind of synthesis of the formal and the conceptual. On one hand, rigorous definitions and proofs are given; on the other, illustrations, examples and applications serve as hopefully efficient conveyors of the former's geometric nature.

Linear algebra is a mathematical discipline which studies – as its name rightly suggests – the *linear*. Nevertheless, the word *linear* (as in 'line-like') is slightly misplaced. The correct term would perhaps be *flat* or, nigh equivalently, *not curved*. It isn't hard to imagine why curved objects (as in *geometric* objects, say) are more difficult to describe and manipulate than objects flat. For instance, the formula for the volume of a cube is just the product of the lengths of its sides. Contrast this with the volume of a still 'simple', yet curved, object – the ball. Its volume cannot even be *precisely* determined; its calculation involves approximating an irrational constant and the derivation of its formula is starkly unintuitive without basic knowledge of measure theory.

As such, linear algebra is a highly 'geometric' discipline and opportunities for visual interpretations abound. This is also a drawback in a certain sense. One should not dwell on visualisations alone as they tend to lead astray where imagination falls short. Symbolic representation of the geometry at hand is key.

The word *linear* however dons a broader sense in modern mathematics. It can be rephrased as reading, 'related by addition and multiplication by a scalar'. We trust kind readers have been acquainted with the notion of a *linear function*. A linear function is (rightly) called *linear* for it receives a number as input and outputs its *constant* multiple plus another *constant* number. Therefore, the output is in a *linear* relation to the input – it is multiplied by some fixed number and added to another. This understanding of the word is going to prove crucial already in the first chapter, where we study *linear systems*. Following are *vector spaces* and *linear maps*, concepts whose depth shall occupy the span of this text. Each chapter is further endowed with an *applications* section where I try to draw a simile between mathematics and common sense.

Contents

Chapter 1

Abstract Vector Spaces

The geometric interpretation of vectors developed in the previous chapter leaves a lot of tones unsung. We defined the space the vectors occupy as $\mathbb{R}^n$. This definition works well for building intuition but can't get us very far in the theory of linear systems. You see, the solution sets to linear systems are rarely equal to $\mathbb{R}^n$ for any n. But they *are*, in a very literal sense, 'spaces of vectors'.

To illustrate what we mean, consider again the linear equation

$$x - y + z = 4$$

from section ??. Its solution set is the set of vectors

$$\left\{ \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid y, z \in \mathbb{R} \right\}.$$

The vector $\left(\begin{smallmatrix} 4 \\ 0 \\ 0 \end{smallmatrix}\right)$ is just a single shift but the set

$$S \coloneqq \left\{ y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid y, z \in \mathbb{R} \right\}$$

indeed is an entire 'space' of vectors, in the sense that it behaves essentially the same as $\mathbb{R}^2$. To highlight certain qualities:

- It's a plane in $\mathbb{R}^3$; a two-dimensional object, just like $\mathbb{R}^2$ is a two-dimensional space.
- It contains the origin – the vector $\left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right)$.
- Adding vectors in S gives a vector in S. We cannot ever leave the plane just by moving along the vectors in the same plane.
- Multiplying vectors from S by real numbers also gives a vector in S. Enlarging or shortening a vector also doesn't allow us to leave the plane.

In this chapter, we shall endeavour to formalize the just outlined concept of a 'set of vectors which behaves just like a space does'. We're going to call these spaces *abstract vector spaces* or just *vector spaces* for short.

The most important idea behind the definition of a vector space (or a space in general) is 'closedness'. A space ought to be a universe in itself, interactions between elements cannot ever lead to the creation of an element which is not present. Mathematicians tend to call these interactions, *operations*, and label sets which meet this criterion as *closed*. There are a few other formal requirements we must enforce (e.g. commutativity and associativity of operations which we take for granted in the real numbers) but the primary aim remains to define a *closed* set, a space moving along whose vectors allows one not to escape it.

We thus proceed to define an abstract vector space as a set of vectors which can be added together and scaled by real numbers by listing axioms (basically enforced rules of behaviour) the set must satisfy.

Definition 1.0.1 (Abstract vector space)

An *(abstract) vector space* over $\mathbb{R}$ is a set V (whose elements style *vectors*) together with operations $\oplus : V \times V \longrightarrow V$ (called *vector addition*) and $\odot : \mathbb{R} \times V \longrightarrow V$ (called *scalar multiplication*) satisfying the following axioms.

(1) The operation $\oplus$ is *commutative*, i.e. $\mathbf{v} \oplus \mathbf{w} = \mathbf{w} \oplus \mathbf{v}$ for every $\mathbf{v}, \mathbf{w} \in V$.

(2) The operation $\oplus$ is *associative*, i.e. $\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w}$ for every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(3) The set V is *closed* under the operation $\oplus$, i.e. $\mathbf{v} \oplus \mathbf{w} \in V$ whenever $\mathbf{v}, \mathbf{w} \in V$.

(4) There exists a *zero vector*, i.e. a vector $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for every $\mathbf{v} \in V$.

(5) Each vector $\mathbf{v} \in V$ has an *additive inverse*, i.e. a vector $\mathbf{w} \in V$ such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$.

(6) The operation $\odot$ distributes over $+$, that is, $(r + s) \odot \mathbf{v} = r \odot \mathbf{v} \oplus s \odot \mathbf{v}$ for all $r, s \in \mathbb{R}$ and $\mathbf{v} \in V$.

(7) The operation $\odot$ distributes over $\oplus$, i.e. $r \odot (\mathbf{v} \oplus \mathbf{w}) = r \odot \mathbf{v} \oplus r \odot \mathbf{w}$ for every $r \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in V$.

(8) Ordinary multiplication (of real numbers) associates with $\odot$, i.e. $(rs) \odot \mathbf{v} = r \odot (s \odot \mathbf{v})$ for every $r, s \in \mathbb{R}$ and $\mathbf{v} \in V$.

(9) The set V is *closed* under $\odot$, that is, $r \odot \mathbf{v} \in V$ whenever $r \in \mathbb{R}$ and $\mathbf{v} \in V$.

(10) Scalar multiplication by 1 acts as the *identity operation*, that is, $1 \odot \mathbf{v} = \mathbf{v}$ for every $\mathbf{v} \in V$.

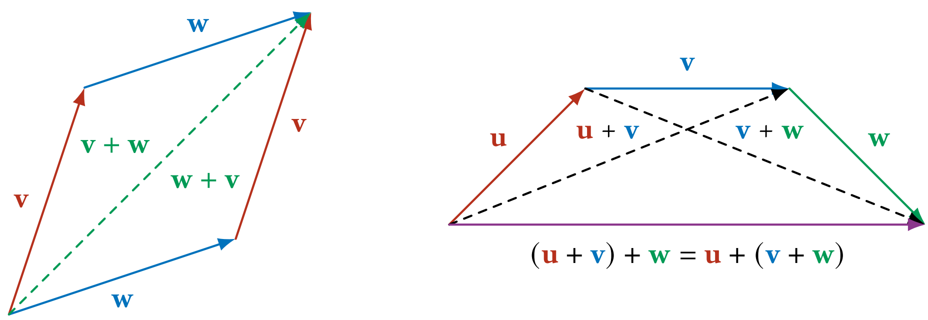


Figure 1.1: The visualization of commutativity and associativity of vector addition.

Remark 1.0.2

The definition of vector space hosts a plethora of axioms; some for different 'mathematical' reasons than others.

The axioms (3) and (9) are the 'most important' in a sense. They ascertain that the resulting set is indeed a space, in the sense discussed in the introduction to this chapter.

The axioms (1), (2), (6), (7), (8) and (10) are technical. Their role is to make vectors and numbers interact in a way we find intuitive hailing from n-dimensional real spaces. There, we take for granted the sum of two vectors is invariant under change of the order of summation but, on the other hand, one can easily design sets of elements with an addition operation not commutative. The gist of it is that we still wish to think of elements of vector spaces as … well … vectors, and once a vector, you should behave like a vector.

Finally, axioms (4) and (5) are there so that we can 'reverse arrows', in a sense. Again, we tend to treat vectors as measures of length and direction so it makes sense to be able to travel the same distance in a direction opposite. The condition of being able to reverse brings with it the necessity to have an 'original point' since otherwise the sum of a vector with its additive inverse would send us flying out of the space we're in. That ought not to happen.

Remark 1.0.3

Diligent readers have surely noticed that we denoted the operations $\oplus$ and $\odot$ on an abstract vector space differently than we would before. For predominantly didactic reasons. When we first defined vector addition, we didn't feel the need to distinguish adding two real numbers from adding two vectors because vector addition is just component-wise addition of numbers anyway. However, elements of an abstract vector space can (as we shall soon see) actually be somewhat distant from the intuitive image of vectors we harbour. It seemed apt to fully convey the perception of difference between vector addition and 'normal' addition. Scalar multiplication falls under the same argument.

Nonetheless, it is common in literature to write the vector addition operation $\oplus$ the same way as the addition of real numbers and we purport to adhere to the norm. However, we shall at least make the distinction between scalar multiplication and real multiplication by using the symbol $\cdot$ for the former and lack of a symbol for the latter. The operation $\cdot$ of scalar multiplication ought not to be confused with dot product which we have not defined for abstract vectors.

We fare ahead and list quite a few examples of abstract vector spaces. Some should come as no surprise, some as quite it.

Example 1.0.4 (n-dimensional real space)

An obvious example of a vector space is the n-dimensional real space $\mathbb{R}^n$. Veracity of the ten axioms is trivially checked given that their conception is based on the $\mathbb{R}^n$ archetype.

Example 1.0.5 (Solution sets of homogeneous linear systems)

The motivating example behind vector spaces have exactly been the sets of vectors like

$$\{r_1 \cdot \mathbf{v}_1 + r_2 \cdot \mathbf{v}_2 + \ldots + r_k \cdot \mathbf{v}_k\},$$

which should be 'correctly' written as

$$\{r_1 \odot \mathbf{v}_1 \oplus r_2 \odot \mathbf{v}_2 \oplus \ldots \oplus r_k \odot \mathbf{v}_k\},$$

for some $r_1, \ldots, r_n \in \mathbb{R}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$. These are not *exactly* equal to $\mathbb{R}^k$ although they do describe a space with k different directions of movement. Since we defined vector spaces with the primary aim of accommodating such examples, it would be quite the sorry situation should they fail to be them. Luckily, all such sets indeed are vector spaces. Naturally, by proposition **??**, these sets arise as sets of solutions of homogeneous linear systems.

Since they are sets of vectors in $\mathbb{R}^n$, the technical axioms (1), (2), (6), (7), (8) and (10) are trivially satisfied. We are going to check axioms (3) and (9) first. They say that the sum of solutions of a homogeneous linear system is also a solution of the same system and so are multiples of solutions. Suppose thus that the vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are solutions of the homogeneous linear system

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \ldots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \ldots & + & a_{2,n}x_n & = & 0 \\ & & & & & & & \vdots & \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \ldots & + & a_{m,n}x_n & = & 0 \end{array}$$

Then, for every $i \leq n$, the equality

$$\begin{aligned} a_{i,1}(v_1 + w_1) + a_{i,2}(v_2 + w_2) + \ldots + a_{i,n}(v_n + w_n) &= a_{i,1}v_1 + a_{i,2}v_2 + \ldots + a_{i,n}v_n \\ &\quad + a_{i,1}w_1 + a_{i,2}w_2 + \ldots + a_{i,n}w_n \\ &= 0 + 0 = 0 \end{aligned}$$

is satisfied, which proves (3). Similarly, if $\mathbf{v}$ is a solution, then for any $r \in \mathbb{R}$ and all $i \leq n$ we have

$$\begin{aligned} a_{i,1}(rv_1) + a_{i,2}(rv_2) + \ldots + a_{i,n}(rv_n) &= ra_{i,1}v_1 + ra_{i,2}v_2 + \ldots + ra_{i,n}v_n \\ &= r(a_{i,1}v_1 + a_{i,2}v_2 + \ldots + a_{i,n}v_n) = 0, \end{aligned}$$

which means that $r \cdot \mathbf{v}$ is also a solution, and thus (9) holds.

Finally, as far as axioms (4) and (5) are concerned, the vector $\mathbf{0}$ is always a solution of a homogeneous linear system and the inverse to a solution $\mathbf{v}$ is, of course, the solution $-1 \cdot \mathbf{v}$ (which is indeed a solution by the preceding paragraph).

Remark 1.0.6

Do note that by example 1.0.5, only the solutions of **homogeneous** linear systems are vector spaces. Solutions to non-homogeneous linear systems **always** fail to be vector spaces; they do not contain the vector $\mathbf{0}$ for example. They are so-called *affine* spaces which we shan't study in this text.

Example 1.0.7 (Sets of polynomials)

Sets of polynomials in one variable of a given degree make an interesting example of an abstract vector space. To recall, a real polynomial p in one variable of degree n is the expression

$$p(x) = r_0 + r_1 x + r_2 x^2 + \ldots + r_n x^n$$

for some $r_0, \ldots, r_n \in \mathbb{R}$. We claim that the set

$$\{p \mid p \text{ is a real polynomial of degree } n\},$$

with the addition operation being the typical addition of polynomials and scalar multiplication being simple multiplication of polynomials by real numbers, is a vector space.

Similarly to the previous example, axioms (1), (2), (6), (7), (8) and (10) are easily verified to hold. Clearly, the sum of two polynomials of degree n is a polynomial of degree n and the multiple of a polynomial of degree n is again a polynomial degree n. The zero vector is the polynomial $0 + 0x + 0x^2 + \ldots + 0x^n$ and the inverse to the polynomial p is of course $-p$.

Do note however that we **forbid polynomial multiplication** since the product of two polynomials of degree n is no longer a polynomial of degree n. We may only add polynomials and multiply them by a real number.

As a matter of fact, polynomials of degree n are in a sense ‘the same’ as vectors over $\mathbb{R}^{n+1}$. The correspondence is easy to forge – we simply encode the coefficients of the polynomial into a vector like so:

$$r_0 + r_1 x + r_2 x^2 + \ldots + r_n x^n \mapsto \begin{pmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{R}^{n+1}.$$

We can now rewrite polynomial addition and scalar multiplication as the same operations on vectors in $\mathbb{R}^{n+1}$. For example, on the set of polynomials of degree 3, the addition

$$(2x + x^2 + 3x^3) + (-1 + 3x - 5x^3) = -1 + 5x + x^2 - 2x^3$$

can be written in the following vector form.

$$\begin{pmatrix} 0 \\ 2 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \\ 0 \\ -5 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \\ 1 \\ -2 \end{pmatrix}$$

Example 1.0.8 (Matrices)

The set

$$\left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mid a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\}$$

of 2×2 matrices with real entries and entry-wise addition and scalar multiplication is a vector space over $\mathbb{R}$. This space is often denoted as $\mathbb{R}^{2\times 2}$. To give an example, the addition of matrices behaves like this:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 1+2 & 2+1 \\ 3+4 & 4+3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 7 & 7 \end{pmatrix};$$

and scalar multiplication like this:

$$2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}.$$

Upon closer inspection, 2×2 matrices behave exactly like 4-component real vectors. Indeed, we can find a correspondence

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \longleftrightarrow \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{pmatrix}$$

and treat the space $\mathbb{R}^{2\times 2}$ exactly the same as $\mathbb{R}^4$ with vector addition and multiplication. This correspondence also shows that all the defining axioms are satisfied.

This example can naturally be scaled to spaces $\mathbb{R}^{m\times n}$ of real matrices with m rows and n columns. One always gets a correspondence between $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{mn}$.

Example 1.0.9 (The trivial space)

The set $\{\mathbf{0}\}$ containing only the zero vector is a vector space. It is actually the smallest (by element count) vector space that can exist. The empty set is not a vector space precisely due to its lack of the zero vector.

Example 1.0.10 (Natural functions)

The set

$$\{f \mid f \text{ is a function } \mathbb{N} \to \mathbb{R}\},$$

with addition defined by $(f+g)(n) = f(n) + g(n)$ and scalar multiplication by $(r \cdot f)(n) = rf(n)$, is a vector space. It differs from previous examples by its *dimension*. Without proper means to define the dimension of a vector space just yet, we just vaguely state that this vector space has *infinite* dimension.

Each function $f : \mathbb{N} \to \mathbb{R}$ *can* actually be represented as a vector with real entries, but with an infinite number of them. For example, the function $f(n) = n^2 + 1$ can be represented as a vector

$$\begin{pmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \end{pmatrix} = \begin{pmatrix} 0^2+1 \\ 1^2+1 \\ 2^2+1 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \\ \vdots \end{pmatrix}$$

whose i-th entry is $f(i)$. Since there are infinitely many natural numbers, these vectors themselves must have infinite entries. We leave the verification of the vector space axioms to our hard-working readers.

Example 1.0.11 (Real functions)

Our final example features the set of all functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ with addition and scalar multiplication defined as in example 1.0.10. This is again a vector space (exercise) which has not just infinite, but *uncountable* dimension. The consequence of this fact is that such functions can no longer be reasonably represented as vectors. As real numbers cannot be *enumerated*, we have no clear idea what the i-th entry of such a vector should be.

Exercise 1.0.12

Prove (by checking the axioms) that the sets of functions mentioned in examples 1.0.10 and 1.0.11 are indeed vector spaces by definition.

The string of examples hopefully shed some light on the power of *abstraction* or *generalisation* in mathematics. It is not that (most) mathematicians enjoy working with abstract and unintuitive concepts (most of them refer to them as 'abstract nonsense' actually) but that this approach yields results about many structures at once. Everything we henceforth prove about vector spaces is going to be valid for all the sets mentioned here as well as absolutely any set that fits the definition of a vector space. To give a few examples:

- the theory of differential equations from calculus relies heavily on the description of solutions of linear systems;
- multivariable real calculus (differentiation and integration of function $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$) uses theory of matrices and their determinants (to be introduced way later) as derivatives of multivarible functions are matrices;
- vectors and matrices whose entries are module homomorphisms are a regular occurrence in my own branch of representation theory of algebras.

Had we kept working only in the spaces $\mathbb{R}^n$, we would have had to be constantly doubting whether any one result wasn't particular to this scenario.

We close this introductory passage with some 'intuitively obvious' statements about qualities of vectors that are nevertheless not mentioned as axioms and, as such, must be proven.

Lemma 1.0.13 (Abstract nonsense)

Let V be a vector space over $\mathbb{R}$. Then, for any $r \in \mathbb{R}$ and $\mathbf{v} \in V$, we have

(a) $0 \cdot \mathbf{v} = \mathbf{0}$;

(b) $(-1 \cdot \mathbf{v}) + \mathbf{v} = \mathbf{0}$;

(c) $r \cdot \mathbf{0} = \mathbf{0}$.

Proof. As for (a), observe that

$$\mathbf{v} \overset{(10)}{=} 1 \cdot \mathbf{v} = (1+0) \cdot \mathbf{v} \overset{(6)}{=} 1 \cdot \mathbf{v} + 0 \cdot \mathbf{v} \overset{(10)}{=} \mathbf{v} + 0 \cdot \mathbf{v}, \tag{1.1}$$

where the numbers above equal signs refer to axioms in the definition of a vector space. Let now $\mathbf{w}$ be the *additive inverse* of $\mathbf{v}$, i.e. $\mathbf{v} + \mathbf{w} = \mathbf{0}$; such exists by axiom (5). Adding $\mathbf{w}$ to both sides of the equation (1.1) gives

$$\begin{aligned}\mathbf{w}+\mathbf{v} &= \mathbf{w}+\mathbf{v}+0\cdot\mathbf{v}\\ \mathbf{0} &= \mathbf{0}+0\cdot\mathbf{v}\end{aligned}$$

By axiom (4), the right side equals $0\cdot\mathbf{v}$ and thus we have proven that $\mathbf{0}=0\cdot\mathbf{v}$.

The calculation

$$(-1\cdot\mathbf{v})+\mathbf{v} \overset{(6)}{=} (-1+1)\cdot\mathbf{v} = 0\cdot\mathbf{v} \overset{(a)}{=} \mathbf{0}$$

proves point (b).

As for (c), we have

$$r\cdot\mathbf{0} \overset{(a)}{=} r\cdot(0\cdot\mathbf{0}) \overset{(8)}{=} (r0)\cdot\mathbf{0} = 0\cdot\mathbf{0} \overset{(a)}{=} \mathbf{0},$$

which proves the statement. ■

1.1 Subspaces And Spans

In this section, we set out to formalise two concepts. First, in section ??, we interpreted the set

$$\left\{ y\cdot\begin{pmatrix}1\\1\\0\end{pmatrix}+z\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix} \mid y,z\in\mathbb{R}\right\}$$

as a plane in $\mathbb{R}^3$. In the introduction to this chapter, we explained the 'plane' part. However, we have not yet made it clear what we mean by 'in $\mathbb{R}^3$'. What does it *mean* for a vector space to *lie inside* another vector space? Second, we've often used the phrase 'vectors define dimensions or directions of movement'. We shall return to this idea very soon as well.

Now, a vector space which is wholly contained in another vector space is called a subspace and its definition is quite simple and natural.

Definition 1.1.1 (Subspace)

Let V be a vector space. A set S is a *subspace* of V if $S\subseteq V$ and, additionally, S is a vector space in its own right. We write $S\leq V$ to indicate that S is a subspace of V.

Simply put, subspaces of V are its subsets that are also closed under vector addition and scalar multiplication. Before we proceed to illustrate the concept on a few examples, we prove a lemma which allows us to check whether a given subset is also a subspace without having to go through all the axioms in the definition of vector space.

Lemma 1.1.2 (Characterisation of subspaces)

Let V be a vector space and S a subset of V with inherited operations of vector addition and scalar multiplication. Then, the following statements are equivalent.

(a) S is a subspace of V.

(b) *S is closed under linear combinations of pairs of vectors: for any $\mathbf{s}_1, \mathbf{s}_2 \in S$ and $r_1, r_2 \in \mathbb{R}$, the vector $r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2$ lies in S.*

(c) *S is closed under linear combinations of any number of vectors: for any $n \in \mathbb{N}$, vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n \in S$ and $r_1, r_2, \ldots, r_n \in \mathbb{R}$, the vector $r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 + \ldots + r_n \cdot \mathbf{s}_n$ lies in S.*

Proof. Instead of proving $(a) \Leftrightarrow (b)$, $(a) \Leftrightarrow (c)$ and $(b) \Leftrightarrow (c)$, it is simpler to establish the claim by proving

$$(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a).$$

As for $(a) \Rightarrow (c)$, since S is a vector space, we need only invoke the axioms. Given $\mathbf{s}_1 \in S$ and $r_1 \in \mathbb{R}$, we use the axiom (9) to ascertain that $r_1 \cdot \mathbf{s}_1 \in S$. Similarly for the other vectors $\mathbf{s}_2, \ldots, \mathbf{s}_n \in S$ and numbers $r_2, \ldots, r_n \in \mathbb{R}$. Then, by repeatedly using axiom (3), we prove that the linear combination $r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 + \ldots + r_n \cdot \mathbf{s}_n$ lies in S.

The implication $(c) \Rightarrow (b)$ is obvious, simply substitute $n = 2$.

The proof of $(b) \Rightarrow (a)$ takes the most work. We must check that S satisfies all the axioms of a vector space. The technical axioms (1), (2), (6), (7), (8) and (10) hold in S because they hold in V. We shall illustrate this on axiom (1), the rest is proven in a very similar manner. Given $\mathbf{s}_1, \mathbf{s}_2 \in S$, we know by (b) that $\mathbf{s}_1 + \mathbf{s}_2 \in S$. However, since $+$ is commutative in V and $\mathbf{s}_1 + \mathbf{s}_2 \in V$, we have the equality $\mathbf{s}_1 + \mathbf{s}_2 = \mathbf{s}_2 + \mathbf{s}_1$ in the vector space V. Because (again by (b)) $\mathbf{s}_2 + \mathbf{s}_1 \in S$, the same equality must also hold in S as it is a subset of V. This proves that vector addition is commutative in S. Axiom (4) is proven by setting $r_1 = r_2 = 0$ and using (b) together with lemma 1.0.13 (a). Similarly, axiom (5) follows (for a given $\mathbf{s} \in S$) by setting $r_1 = 0, \mathbf{s}_1 = \mathbf{0}$ and $r_2 = -1, \mathbf{s}_2 = \mathbf{s}$ while using lemma 1.0.13 (b). Finally, given $\mathbf{s}_1, \mathbf{s}_2 \in S$, axiom (3) follows by $r_1 = r_2 = 1$ and given $r \in \mathbb{R}, \mathbf{s} \in S$, axiom (9) is verified by $r_1 = r, \mathbf{s}_1 = \mathbf{s}, r_2 = 0, \mathbf{s}_2 = \mathbf{0}$. S is thus a vector space and a subset of V so (a) holds. ∎

Warning 1.1.3

By definition, $\mathbb{R}^m$ is **never** a subspace of $\mathbb{R}^n$ as long as $m \neq n$. This is because it is not a subset. The space $\mathbb{R}^n$ contains vectors with *exactly* n entries and no vector with m entries.

This issue is alleviated by 'filling' the vectors in $\mathbb{R}^m$ with zeroes so that they have n entries if $m \leq n$. Formally, the set

$$\left\{ \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n \mid v_1, \ldots, v_m \in \mathbb{R} \right\}$$

is a subspace of $\mathbb{R}^n$ and is all but formally equivalent to $\mathbb{R}^m$.

Example 1.1.4

By example 1.0.5, the solution set of a homogeneous linear system with n variables is a subspace of $\mathbb{R}^n$. To give a concrete example, the solution set of the system

$$\begin{aligned} x + 3y - z &= 0 \\ 2x + y - 2z &= 0 \end{aligned}$$

is the set

$$S \coloneqq \left\{ z \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid z \in \mathbb{R} \right\}$$

and it is a subspace of $\mathbb{R}^3$. We can prove this using lemma 1.1.2 (b). Given $\mathbf{s}_1, \mathbf{s}_2 \in S$, there exist $z_1, z_2 \in \mathbb{R}$ such that

$$\mathbf{s}_1 = z_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{s}_2 = z_2 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Then, for any $r_1, r_2 \in \mathbb{R}$, we have

$$r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 = (r_1 z_1) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + (r_2 z_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = (r_1 z_1 + r_2 z_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

and so $r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 \in S$.

Example 1.1.5

The set of polynomials of degree at most 2 is a subspace of the set of polynomials of degree at most 4. Indeed, given two polynomials $a_0 + a_1 x + a_2 x^2$ and $b_0 + b_1 x + b_2 x^2$ and numbers $r_1, r_2 \in \mathbb{R}$, we have

$$r_1 \cdot (a_0 + a_1 x + a_2 x^2) + r_2 \cdot (b_0 + b_1 x + b_2 x^2) = (r_1 a_0 + r_2 b_0) + (r_1 a_1 + r_2 b_1) x + (r_1 a_2 + r_2 b_2) x^2$$

which is a polynomial of degree at most 2. Clearly, every polynomial of degree at most 2 is a polynomial of degree at most 4.

Example 1.1.6

The set of 2×2 real matrices

$$L \coloneqq \left\{ \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \mid a + b + c = 0 \right\}$$

is a subspace of $\mathbb{R}^{2\times 2}$. Indeed, this is because we can substitute $a = -b - c$ and write

$$\begin{pmatrix} -b-c & 0 \\ b & c \end{pmatrix} = \begin{pmatrix} -b & 0 \\ b & 0 \end{pmatrix} + \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix} = b \cdot \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

and so every matrix in L is a linear combination of the matrices

$$\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

which proves that L is a subspace by lemma 1.1.2 (b).

Example 1.1.7

The set of real functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ which have a finite derivative at 0 is a subspace of the set of all real functions. Indeed, if $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ are real functions and $f'(0), g'(0) \in \mathbb{R}$, then for any $r_1, r_2 \in \mathbb{R}$ we can compute

$$(r_1 \cdot f + r_2 \cdot g)'(0) = r_1 f'(0) + r_2 g'(0) \in \mathbb{R}$$

so the derivative of $r_1 \cdot f + r_2 \cdot g$ at 0 is finite.

We now return to the idea of vectors as 'directions of movement'. The overused poor little set

$$S := \left\{ y \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid y, z \in \mathbb{R} \right\}$$

is a subspace of $\mathbb{R}^3$ and its description tells us that every vector in that space is a linear combination of

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

If such is the case, we say that S is *generated* by these vectors and write

$$S = \operatorname{span}\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right).$$

More generally, if a vector space V is the *span* of a set of vectors S, it means that every vector in V is a linear combination of vectors from S. Intuitively, this formalises the idea that every point in the space V is reachable just by moving along vectors from S.

Definition 1.1.8 (Span)

Let V be a vector space and S a subset of V (**not** necessarily a subspace). The *span of* S is the set of all linear combinations of vectors from S. Symbolically,

$$\operatorname{span} S := \left\{ \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} \mid r_{\mathbf{s}} \in \mathbb{R} \ \forall \mathbf{s} \in S \right\}$$

To avoid technical details, we only consider **finite** sums of the form above to form the set $\operatorname{span} S$. If $S = \emptyset$, we define $\operatorname{span} S := \{\mathbf{0}\}$.

Remark 1.1.9

Proposition ?? states that the set of solutions of a homogeneous linear system is of the form

$$\{r_1 \cdot \mathbf{v}_1 + r_2 \cdot \mathbf{v}_2 + \ldots + r_k \cdot \mathbf{v}_k\}.$$

We can now write the same set more succinctly as

$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k).$$

In light of lemma 1.1.2, it should come as no surprise that spans of sets are always vector spaces. The verification of this fact essentially boils down to 'a linear combination of linear combinations is a linear combination'.

Lemma 1.1.10

Let V be a vector space. Then, for every subset $S \subseteq V$, the set $\operatorname{span} S$ is a subspace of V.

Proof. By lemma 1.1.2 (b), we need only prove that $a \cdot \mathbf{s}_1 + b \cdot \mathbf{s}_2 \in \operatorname{span} S$ for $a, b \in \mathbb{R}$ and $\mathbf{s}_1, \mathbf{s}_2 \in \operatorname{span} S$. By definition, we can write

$$\mathbf{s}_1 = \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s}, \quad \mathbf{s}_2 = \sum_{\mathbf{s} \in S} t_{\mathbf{s}} \cdot \mathbf{s}$$

for adequate numbers $r_{\mathbf{s}}, t_{\mathbf{s}} \in \mathbb{R}$. Then,

$$\begin{aligned} a \cdot \mathbf{s}_1 + b \cdot \mathbf{s}_2 &= a \cdot \left(\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s}\right) + b \cdot \left(\sum_{\mathbf{s} \in S} t_{\mathbf{s}} \cdot \mathbf{s}\right) = \sum_{\mathbf{s} \in S} a \cdot (r_{\mathbf{s}} \cdot \mathbf{s}) + \sum_{\mathbf{s} \in S} b \cdot (t_{\mathbf{s}} \cdot \mathbf{s}) \\ &= \sum_{\mathbf{s} \in S} (a r_{\mathbf{s}}) \cdot \mathbf{s} + \sum_{\mathbf{s} \in S} (b t_{\mathbf{s}}) \cdot \mathbf{s} = \sum_{\mathbf{s} \in S} (a r_{\mathbf{s}} + b t_{\mathbf{s}}) \cdot \mathbf{s}. \end{aligned}$$

Relabelling $p_{\mathbf{s}} := a r_{\mathbf{s}} + b t_{\mathbf{s}}$, we get $p_{\mathbf{s}} \in \mathbb{R}$ and

$$a \cdot \mathbf{s}_1 + b \cdot \mathbf{s}_2 = \sum_{\mathbf{s} \in S} p_{\mathbf{s}} \cdot \mathbf{s},$$

hence $a \cdot \mathbf{s}_1 + b \cdot \mathbf{s}_2 \in \operatorname{span} S$ which proves that $\operatorname{span} S$ is a vector space.

By lemma 1.1.2 (c), the space V is closed under linear combinations of finite numbers of vectors. Since $\operatorname{span} S$ is defined as the set of finite linear combinations of vectors of S (which also lie in V), it is clear that $\operatorname{span} S \leq V$. ■

Quite trivially, the converse of the previous lemma also holds: any subspace of V is a span of some set of vectors from V. This is because any subspace of V is obviously its own span.

Example 1.1.11 (Line)

In any vector space V, the set $\operatorname{span}(\mathbf{v})$ is a subspace for any $\mathbf{v} \in V$. It is a line passing through the origin in the direction of $\mathbf{v}$.

Example 1.1.12

The space
$$\operatorname{span}\left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right)$$
is the entirety of $\mathbb{R}^2$. How do we prove this? By definition of a span, every vector $\mathbf{v} \in \mathbb{R}^2$ should be a linear combination of the vectors $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$. That is, there should exist numbers $a, b \in \mathbb{R}$ such that
$$\mathbf{v} = \begin{pmatrix}v_1\\v_2\end{pmatrix} = a \cdot \begin{pmatrix}1\\1\end{pmatrix} + b \cdot \begin{pmatrix}1\\-1\end{pmatrix}.$$
Rewriting the equation in a more instructive manner:
$$\begin{aligned} a + b &= v_1 \\ a - b &= v_2 \end{aligned}$$
This bespeaks the fact that $\mathbf{v} \in \operatorname{span}(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix})$ if and only if the system above has a solution in variables a and b. Gauss-Jordan elimination gives
$$\begin{aligned} a + \quad b &= v_1 \\ -\ 2b &= v_2 - v_1 \end{aligned}$$
and so $b = (v_1 - v_2)/2$ and $a = (v_1 + v_2)/2$. This proves that $\operatorname{span}(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}) = \mathbb{R}^2$ since we can express any vector in $\mathbb{R}^2$ as a linear combination of $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$.

Example 1.1.13

Consider the set
$$S \coloneqq \{3x - x^2, 2x\}$$
as a subset of the set of polynomials of degree at most 2. We wish to describe polynomials lying in $\operatorname{span} S$. This entails looking at all polynomials of the form
$$a \cdot (3x - x^2) + b \cdot (2x)$$
for $a, b \in \mathbb{R}$. A polynomial of degree at most 2 has the general form $r_0 + r_1x + r_2x^2$. We can compare coefficients on either side of the equation
$$\begin{aligned} r_0 + r_1x + r_2x^2 &= a \cdot (3x - x^2) + b \cdot (2x) \\ r_0 + r_1x + r_2x^2 &= 0 + (3a + 2b)x + (-a)x^2 \end{aligned}$$
to reach the linear system
$$\begin{aligned} -a &= r_2 \\ 3a + 2b &= r_1 \\ 0 &= r_0 \end{aligned}$$
whose solution is $(a, b) = (-r_2, (3/2)r_2 + (1/2)r_1)$ but only in the case that $r_0 = 0$, otherwise it has none. It follows that $\operatorname{span}(3x - x^2, 2x)$ is the subspace of all polynomials $r_1x + r_2x^2$ for $r_1, r_2 \in \mathbb{R}$.

The preceding two examples hint at a general way to check whether a given vector $\mathbf{v} \in V$ lies in $\operatorname{span} S$ for some $S \subseteq V$. If S is finite, say, $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$, then $\mathbf{v}$ lies in $\operatorname{span} S$ if and only if there exist numbers $r_1, \ldots, r_k \in \mathbb{R}$ such that

$$\mathbf{v} = r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 + \ldots + r_k \cdot \mathbf{s}_k.$$

If V is further a subspace of $\mathbb{R}^n$, that is, if we can write the vector $\mathbf{v}$ as

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

and the vectors $\mathbf{s}_1, \ldots, \mathbf{s}_k$ as

$$\mathbf{s}_1 = \begin{pmatrix} s_{11} \\ s_{12} \\ \vdots \\ s_{1n} \end{pmatrix}, \mathbf{s}_2 = \begin{pmatrix} s_{21} \\ s_{22} \\ \vdots \\ s_{2n} \end{pmatrix}, \ldots, \mathbf{s}_k = \begin{pmatrix} s_{k1} \\ s_{k2} \\ \vdots \\ s_{kn} \end{pmatrix},$$

then $\mathbf{v} \in \operatorname{span} S$ if and only if the linear system

$$\left(\begin{array}{cccc|c} s_{11} & s_{21} & \cdots & s_{k1} & v_1 \\ s_{12} & s_{22} & \cdots & s_{k2} & v_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{1n} & s_{2n} & \cdots & s_{kn} & v_n \end{array}\right)$$

has **at least one** solution. Any one solution of this system is then naturally the n-tuple $(r_1, \ldots, r_n)$ of the coefficients of the linear combination which gives rise to the vector $\mathbf{v}$. We can also write the matrix above succinctly as

$$\left(\begin{array}{cccc|c} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_k & \mathbf{v} \end{array}\right).$$

Put into words, a vector $\mathbf{v}$ lies in $\operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_k)$ if and only if the system with columns the vectors $\mathbf{s}_1, \ldots, \mathbf{s}_k$ and right side the vector $\mathbf{v}$ has at least one solution.

Problem 1.1.14

Determine whether the vector

$$\mathbf{v} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$$

lies in

$$\operatorname{span}\left(\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}\right).$$

If so, find $x_1, x_2, x_3, x_4 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + x_4 \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Solution. We simply attempt to solve the system

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 1 \\ 4 & -2 & 1 & 1 & 3 \\ 0 & 1 & 2 & 1 & -2 \end{array}\right).$$

Gauss-Jordan elimination gives

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 1 \\ 0 & -6 & 1 & -7 & -1 \\ 0 & 0 & 13 & -1 & -13 \end{array}\right).$$

For we're only interested in a single solution and x_4 is a free variable, we substitute $x_4 = 0$. Then, from $13x_3 - 1 \cdot 0 = -13$ follows $x_3 = -1$. Further computation gives $x_2 = 0$ and $x_1 = 1$, which means that

$$\begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

♠

Problem 1.1.14 shows that there may often be multiple ways to express a vector as a linear combination of the spanning vectors. This naturally leads to a question computational in nature: 'Given a set of vectors S, can we find another set T such that $\operatorname{span} S = \operatorname{span} T$ but T has as few vectors as possible?' The reason we'd wish to do something like this is simple; in checking whether a vector lies in a given span, free variables in the resulting system are *redundancies* that increase both the computational time and required memory. The next section is dedicated to answering the question at hand.

1.2 Linear Independence, Basis And Dimension

Given a set $S \subseteq V$ of vectors, we first answer the question of *which vectors can be removed from S while not altering its span*. That is, given a vector $\mathbf{s} \in S$, how do we find out whether $\operatorname{span}(S \setminus \{\mathbf{s}\}) = \operatorname{span} S$? Vaguely speaking, provided that $\operatorname{span} S$ is a set of linear combinations of vectors from S, should some vector in S already *be a linear combination* of the other vectors in S, it wouldn't be needed. Turns out this statement is not so vague after all, as we proceed to demonstrate.

Lemma 1.2.1

Let V be a vector space, $S \subseteq V$ and $\mathbf{v} \in V$. Then, $\operatorname{span}(S \cup \{\mathbf{v}\}) = \operatorname{span} S$ *if and only if* $\mathbf{v} \in \operatorname{span} S$.

Proof. We must prove two implications.

As for the implication ($\Rightarrow$), it is simpler to prove it in contrapositive form. If $\mathbf{v} \notin \operatorname{span} S$, then clearly $\operatorname{span} S \neq \operatorname{span}(S \cup \{\mathbf{v}\})$ simply because the latter contains the vector $\mathbf{v}$ while the former does not.

In proving ($\Leftarrow$), assume that $\mathbf{v} \in \operatorname{span} S$. Clearly, $\operatorname{span} S \subseteq \operatorname{span}(S \cup \{\mathbf{v}\})$ as the latter set contains every linear combination of the vectors in S. We must show that also $\operatorname{span}(S \cup \{\mathbf{v}\}) \subseteq$

span S. To this end, choose a vector $\mathbf{w} \in \operatorname{span}(S \cup \{\mathbf{v}\})$. This vector $\mathbf{w}$ is a linear combination of vectors from $S \cup \{\mathbf{v}\}$, i.e.

$$\mathbf{w} = \sum_{\mathbf{s} \in S \cup \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s}$$

for some $r_{\mathbf{s}} \in \mathbb{R}$. We can break this linear combination into two parts like so:

$$\mathbf{w} = \sum_{\mathbf{s} \in S \cup \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s} = \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} + r_{\mathbf{v}} \cdot \mathbf{v}. \tag{1.2}$$

Now, $\mathbf{v} \in \operatorname{span} S$ by assumption so there also exist numbers $t_{\mathbf{s}} \in \mathbb{R}$ such that

$$\mathbf{v} = \sum_{\mathbf{s} \in S} t_{\mathbf{s}} \cdot \mathbf{s}.$$

Substituting this into the equation (1.2) gives

$$\mathbf{w} = \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} + r_{\mathbf{v}} \cdot \left(\sum_{\mathbf{s} \in S} t_{\mathbf{s}} \cdot \mathbf{s} \right) = \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} + \sum_{\mathbf{s} \in S} (r_{\mathbf{v}} t_{\mathbf{s}}) \cdot \mathbf{s} = \sum_{\mathbf{s} \in S} (r_{\mathbf{s}} + r_{\mathbf{v}} t_{\mathbf{s}}) \cdot \mathbf{s}.$$

The last sum is a linear combination of vectors from S and thus $\mathbf{w} \in \operatorname{span} S$, as desired. ■

As a corollary, we get a formalisation of the idea from the introductory paragraph.

Corollary 1.2.2

Given $\mathbf{s} \in S$, *it holds that* $\operatorname{span} S = \operatorname{span}(S \setminus \{\mathbf{s}\})$ *if and only if* $\mathbf{s} \in \operatorname{span}(S \setminus \{\mathbf{s}\})$.

Proof. Follows immediately from lemma 1.2.1. Simply substitute $\mathbf{v} \coloneqq \mathbf{s}$ and $S \coloneqq S \setminus \{\mathbf{s}\}$. ■

The just uttered corollary has algorithmic vibes. Can't we just keep removing vectors from S which are linearly dependent on other vectors until there are no longer any? Indeed, we can. First however, we should devise a computationally sound way to determine which vectors we may omit. As we now stand, the best we can do is guess at random. Pick a vector $\mathbf{s} \in S$ and check if it lies in $\operatorname{span}(S \setminus \{\mathbf{s}\})$ (as we learnt to do in the the previous section). If it doesn't, tough luck, try again. We might potentially have to go through *every* vector in S before we find one that can be left out, if there even were one to begin with. This is about as algorithmic as cooking a soup by mixing random ingredients until we stumble upon a combination which is reasonably non-lethal.

Fortunately, there is an algorithmic approach to the problem and we are unveiling it promptly. Before that however, we should label sets with no 'unnecessary' vectors somehow.

Definition 1.2.3 (Linear independence)

Let V be a vector space and $S \subseteq V$. If no vector $\mathbf{s} \in S$ can be written as a linear combination of vectors from $S \setminus \{\mathbf{s}\}$, we call S *linearly independent*. If such is not the case, it is called *linearly dependent*.

There lies just a simple observation between us and a feasible algorithm for determining linear independence of a given set of vectors. Suppose $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ and may the vector $\mathbf{s}_i$ be a linear combination of the other vectors, that is to say, there are numbers $r_1, \ldots, r_{i-1}, r_{i+1}, \ldots, r_n \in \mathbb{R}$

satisfying the equation

$$\mathbf{s}_i = r_1 \cdot \mathbf{s}_1 + r_2 \cdot \mathbf{s}_2 + \ldots + r_{i-1} \cdot \mathbf{s}_{i-1} + r_{i+1} \cdot \mathbf{s}_{i-1} + \ldots + r_n \cdot \mathbf{s}_n.$$

We can naturally put $\mathbf{s}_i$ to the right hand side and set $r_i := -1$ to arrive at the equality

$$\mathbf{0} = r_1 \cdot \mathbf{s}_1 + \ldots + r_{i-1} \cdot \mathbf{s}_{i-1} + r_i \cdot \mathbf{s}_i + r_{i+1} \cdot \mathbf{s}_{i+1} + \ldots + r_n \mathbf{s}_n.$$

To express this equality in words: we have found a linear combination (with non-zero coefficients) of vectors from S that gives the zero vector. Could this happen were S linearly independent? Of course it couldn't! If it did, then we could just rearrange the last equality to the first one and get $\mathbf{s}_i$ as a linear combination of the other vectors from S, proving thus that S, in fact, had *not* been linearly independent. Let us dock this train of thought in the following, computationally indispensable, proposition.

Proposition 1.2.4

Let V be a vector space and $S \subseteq V$. Then S is linearly independent if and only if the equality

$$\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot s = \mathbf{0}$$

enforces $r_{\mathbf{s}} = 0$ for every $\mathbf{s} \in S$. In other words, the only linear combination that gives the zero vector has all coefficients equal to 0.

Proof. The paragraph preceding this proposition already illustrates the idea of the proof.

To prove the implication $(\Leftarrow)$, suppose that S is linearly dependent. That is, there exists $\mathbf{v} \in S$ such that $\mathbf{v} \neq \mathbf{0}$ and

$$\mathbf{v} = \sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s}.$$

If we put $\mathbf{v}$ to the right hand side and set $r_{\mathbf{v}} := -1$, we get

$$\mathbf{0} = \sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s} + (-1) \cdot \mathbf{v} = \sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s} + r_{\mathbf{v}} \cdot \mathbf{v} = \sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s}.$$

We've thus wrought a linear combination of vectors from S which has non-zero coefficients but equals the zero vector.

As for $(\Rightarrow)$, assume that there exists a linear combination

$$\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} = \mathbf{0}$$

with at least one $r_{\mathbf{v}} \neq 0$. This means that we can rearrange

$$\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} = \mathbf{0}$$

$$\sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s} + r_{\mathbf{v}} \cdot \mathbf{v} = \mathbf{0}$$

$$\sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} r_{\mathbf{s}} \cdot \mathbf{s} = -r_{\mathbf{v}} \cdot \mathbf{v}$$

$$-\frac{1}{r_{\mathbf{v}}} \cdot \left(\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} \right) = \sum_{\mathbf{s} \in S \setminus \{\mathbf{v}\}} -\frac{r_{\mathbf{s}}}{r_{\mathbf{v}}} \cdot \mathbf{s} = \mathbf{v}$$

and thus $\mathbf{v} \in \operatorname{span}(S \setminus \{\mathbf{v}\})$ which shows that S is linearly dependent. ■

Corollary 1.2.5 (Computing linear independence)

Let $V \leq \mathbb{R}^n$ be a subspace of $\mathbb{R}^n$ for some $n \in \mathbb{N}$ and $S = \{\mathbf{s}_1, \dots, \mathbf{s}_k\} \subseteq V$. Then, S is linearly independent if and only if the linear system

$$\left(\mathbf{s}_1 \ \mathbf{s}_2 \ \cdots \ \mathbf{s}_k \,\middle|\, \mathbf{0}\right)$$

has the unique solution $\mathbf{0}$.

Proof. The proof just amounts to rewriting the equality

$$\sum_{\mathbf{s} \in S} r_{\mathbf{s}} \cdot \mathbf{s} = \sum_{i=1}^{k} r_{\mathbf{s}_i} \cdot \mathbf{s}_i = \mathbf{0}$$

into a linear system and applying proposition 1.2.4. ■

Example 1.2.6

The set

$$S := \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right\}$$

is linearly dependent in $\mathbb{R}^2$. Indeed, the system

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 \end{array}\right)$$

is solved by $(z, -z, -z)$ for any $z \in \mathbb{R}$. Corollary 1.2.5 states that S is linearly dependent.

Example 1.2.7

The set $S := \{1-x, 1+x\}$ is linearly independent in the vector space of quadratic polynomials. To see this, consider a linear combination

$$\begin{aligned} r_1 \cdot (1-x) + r_2 \cdot (1+x) &= 0 + 0x + 0x^2 \\ (r_1 + r_2) + (-r_1 + r_2)x + 0x^2 &= 0 + 0x + 0x^2. \end{aligned}$$

Comparing coefficients gives

$$\begin{aligned} r_1 + r_2 &= 0 \\ -r_1 + r_2 &= 0 \\ 0 &= 0. \end{aligned}$$

This system has the unique solution $(0, 0)$, hence the only way to linearly combine the polynomials $1-x$ and $1+x$ into the zero polynomial requires multiplying them both by 0. Proposition 1.2.4 takes the reins.

Now that we have an algorithmic way of determining whether a given set is linearly independent or not, we should tackle the problem of which vectors can be removed from the set without shrinking its span. Before we do that, let us first ascertain that indeed every (at least **finite**) linearly dependent set can be made linearly independent by successive removal of redundant vectors.

Lemma 1.2.8 (Linearly independent subset)

Given a vector space V and a ***finite subset*** *$S \subseteq V$, there exists a set $T \subseteq S$ that is linearly independent and* $\operatorname{span} S = \operatorname{span} T$.

PROOF. Label $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$. If S is linearly independent, we're done. Otherwise, set $S_0 := S$ and find $i \in \{1, \ldots, k\}$ such that

$$\mathbf{s}_i = \sum_{j \neq i} r_j \cdot \mathbf{s}_j$$

for some $r_j \in \mathbb{R}$. Set $S_1 := S_0 \setminus \{\mathbf{s}_i\}$. By corollary 1.2.2, $\operatorname{span} S_1 = \operatorname{span} S_0$.

Repeat this process until S_m is linearly independent for some $m \in \mathbb{N}$. Such m necessarily exists because S has a finite number of elements and a one-vector set is always linearly independent. Again, by corollary 1.2.2 (applied m times), we have $\operatorname{span} S_m = \operatorname{span} S$ and thus we have found a linearly independent subset of S with the same span as S. ■

Recall from section **??** that some variables of linear systems are pivots and some are free. Pivots have their value expressed as a linear combination of free variables. If we put vectors from a given finite set $S \subseteq V$ into columns of a matrix (as in corollary 1.2.2), we claim that columns hosting free variables mark vectors that can be removed without shrinking the span of S. Why is it so? The answer is actually quite easy. We show it on an example.

Consider the set

$$S = \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right\} \subseteq \mathbb{R}^3.$$

Upon organizing the vectors of S into the matrix

$$\begin{pmatrix} 1 & 1 & 2 & -1 & 1 \\ 3 & -1 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 & 3 \end{pmatrix}$$

and performing Gauss-Jordan elimination, we get

$$\begin{pmatrix} 1 & 1 & 2 & -1 & 1 \\ 0 & -4 & -6 & 4 & -2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

It follows that the variables x_3 and x_5 are free. Why does it mean that the third and fifth vector of S are redundant? Well, the solution of the homogeneous linear system with this matrix is

$$\left\{ x_3 \cdot \begin{pmatrix} 1 \\ 0 \\ -5 \\ -6 \\ 3 \end{pmatrix} + x_5 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \\ -1 \end{pmatrix} \;\middle|\; x_3, x_5 \in \mathbb{R} \right\}. \tag{1.3}$$

By corollary 1.2.2, the set S is linearly independent if and only if the set above contains only the vector $\mathbf{0}$. However, that happens if and only if we force $x_3 = x_5 = 0$. Next, every linear combination of vectors from S is of the form

$$x_1 \cdot \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + x_4 \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + x_5 \cdot \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

If, to ensure linear independence, we must require that x_3 and x_5 both be always equal to 0, it is completely pointless that we include the vectors $\begin{pmatrix}2\\0\\1\end{pmatrix}$ and $\begin{pmatrix}1\\1\\3\end{pmatrix}$ in the linear combination in the first place.

Furthermore, observe that the vectors in the set (1.3) also hint at how we can express the third and fifth vectors as linear combinations of the other three. Indeed, the vector

$$\begin{pmatrix}1\\0\\-5\\-6\\3\end{pmatrix}$$

in fact contains the coefficients of the linear combination of vectors in S that gives the zero vector (as it is the solution of the corresponding homogeneous system). This means that

$$1\cdot\begin{pmatrix}1\\3\\2\end{pmatrix}+0\cdot\begin{pmatrix}1\\-1\\0\end{pmatrix}+(-5)\cdot\begin{pmatrix}2\\0\\1\end{pmatrix}+(-6)\cdot\begin{pmatrix}-1\\1\\1\end{pmatrix}+3\cdot\begin{pmatrix}1\\1\\3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Rearranging (and dividing by -3) gives

$$-\frac{1}{3}\cdot\begin{pmatrix}1\\3\\2\end{pmatrix}+\frac{5}{3}\cdot\begin{pmatrix}2\\0\\1\end{pmatrix}+2\cdot\begin{pmatrix}-1\\1\\1\end{pmatrix}=\begin{pmatrix}1\\1\\3\end{pmatrix}$$

and thus we have expressed the vector $\begin{pmatrix}1\\1\\3\end{pmatrix}$ as a linear combination of the other four vectors. We could do the same for the vector $\begin{pmatrix}2\\0\\1\end{pmatrix}$ which we also know to be redundant. Note, however, that we would have to in addition substitute the vector $\begin{pmatrix}1\\1\\3\end{pmatrix}$ in the resulting linear combination as it should have been already removed from S.

To breathe some clarity into the concluded discussion, we shall show the described procedure in a more algorithmic way.

Problem 1.2.9

Prove that the set

$$S := \left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}-2\\1\end{pmatrix},\begin{pmatrix}0\\-3\end{pmatrix},\begin{pmatrix}3\\0\end{pmatrix}\right\}\subseteq\mathbb{R}^2$$

is linearly dependent and reduce it to a linearly independent set $S' \subseteq S$ with $\operatorname{span} S' = \operatorname{span} S$. *In addition, express the removed vectors as linear combinations of the remaining ones.*

Solution. We compute the solution of the homogeneous linear system

$$\left(\begin{array}{cccc|c}1&-2&0&3&0\\3&1&-3&0&0\end{array}\right).$$

After Gauss-Jordan elimination, we're left with

$$\left(\begin{array}{cccc|c}1&-2&0&3&0\\0&7&-3&-9&0\end{array}\right).$$

This means that columns 1 and 2 host pivots and columns 3 and 4 the free variables. We shall thus remove the third and the fourth vector from S. To finish the calculation, back-substitute and arrive at the set

$$\left\{ x_3 \cdot \begin{pmatrix} 3 \\ 0 \\ 3 \\ -1 \end{pmatrix} + x_4 \cdot \begin{pmatrix} 0 \\ 3 \\ 1 \\ 2 \end{pmatrix} \mid x_3, x_4 \in \mathbb{R} \right\}.$$

We get rid of the fourth vector first. From the shape of the just computed solution, we infer that

$$0 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 3 \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ -3 \end{pmatrix} + 2 \cdot \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and so

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} = -\frac{3}{2} \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot \begin{pmatrix} 0 \\ -3 \end{pmatrix}, \tag{1.4}$$

which by corollary 1.2.2 proves that

$$\operatorname{span}\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \end{pmatrix} \right) = \operatorname{span} S.$$

We now proceed to further remove $\left(\begin{smallmatrix} 0 \\ -3 \end{smallmatrix}\right)$ and express it as a linear combination of the remaining two vectors. The second vector in the computed solution gives the equality

$$3 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 0 \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ -3 \end{pmatrix} - 1 \cdot \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

hence

$$\begin{pmatrix} 0 \\ -3 \end{pmatrix} = -\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \frac{1}{3} \cdot \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Substituting for $\left(\begin{smallmatrix} 3 \\ 0 \end{smallmatrix}\right)$ the linear combination in (1.4) and merging the coefficients yields

$$\begin{pmatrix} 0 \\ -3 \end{pmatrix} = -\frac{6}{7} \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \frac{3}{7} \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Finally, the desired linearly independent set S' with $\operatorname{span} S' = \operatorname{span} S$ is

$$S' = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}.$$

♠

Remark 1.2.10

Observe that the solution of problem 1.2.9 basically mimics the proof of lemma 1.2.8 with an algorithmic approach to the selection of redundant vectors.

Warning 1.2.11

The indices of columns with pivots vs. free variables only point at the vectors of the original set which **are sure not to** shrink the span but they maken't the choice of vectors unique in any way. As a matter of fact, in many cases any of the present vectors can be removed without altering the span of the original set.

To give one trivial example, consider the set

$$S \coloneqq \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}.$$

Any one of the vectors in S is linearly dependent on the other two (on either of them actually); all the vectors in S are redundant.

On the other hand, in the set

$$S \coloneqq \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\},$$

the first vector **is not** redundant. Only either of the second and third vectors may be mercilessly cut down without shrinking the span of S. It is in cases like these that the procedure outlined in problem 1.2.9 has its merit.

We conclude the introduction to the concept of linear independence with one last brief little inconsequential unimportant and just barely appealing discussion. We've proven that removing a vector from a (finite) linearly dependent set can make it independent. Adding a vector to a linearly dependent set on the other hand cannot fix linear dependence. We shall summarise the link between subsets and linear independence of the original set in table 1.1.

	$\hat{S} \subseteq S$	$\tilde{S} \supseteq S$
S linearly independent	$\hat{S}$ also linearly independent	$\tilde{S}$ can be either
S linearly dependent	$\hat{S}$ can be either	$\tilde{S}$ also linearly dependent

Table 1.1: Linear dependence/independence of subsets.

1.2.1 Basis Of A Vector Space

The study of linearly independent sets in section 1.2 carries on its back yet another question: 'Can *every* vector space be expressed as the span of a linearly independent set?' The answer this time is *almost*. As we've made customary, before we proceed to elucidate the given answer, we establish some nomenclature to achieve a manageable level of brevity.

Definition 1.2.12 (Basis)

Let V be a vector space. An ordered n-tuple $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, where $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, which is both linearly independent and spans V is called the *basis* of V.

Warning 1.2.13

We've defined the basis of a vector space specifically to be an **ordered tuple** and **not just a set**. The reason for this will be given later in the chapter when we discuss representation of vectors with respect to distinct bases. Practically, this means that a basis, for example,

$$\left(\begin{pmatrix}69\\0\end{pmatrix},\begin{pmatrix}0\\420\end{pmatrix}\right)$$

is **different** from

$$\left(\begin{pmatrix}0\\420\end{pmatrix},\begin{pmatrix}69\\0\end{pmatrix}\right).$$

Example 1.2.14

The pair

$$B \coloneqq \left(\begin{pmatrix}2\\4\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right)$$

is a basis of $\mathbb{R}^2$. Verification of this statement entails making sure that B is linearly independent and that the linear system

$$\left(\begin{array}{cc|c}2 & 1 & v_1\\4 & 1 & v_2\end{array}\right).$$

has a solution for every $v_1, v_2 \in \mathbb{R}$.

Every n-dimensional space has a basis – many of them in fact. One particular basis is considered the 'most natural', for chiefly geometric reasons. It is the basis whose vectors have directions of the coordinate axes; it bears many names, e.g. *standard*, *canonical* or *natural*.

Definition 1.2.15 (Standard basis)

The n-tuple

$$\mathcal{E}_n \coloneqq \left(\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix},\dots,\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}\right)$$

is a basis of $\mathbb{R}^n$ and is called the *standard* (or *canonical* or *natural*) basis. We denote the vectors of $\mathcal{E}_n$ (in order) $\mathbf{e}_1$ up to $\mathbf{e}_n$.

Example 1.2.16

The natural basis of the vector space of cubic polynomials is $(1, x, x^2, x^3)$. Other bases of the same space include $(x^3, 3x^2, 6x, 6)$ or $(1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3)$.

Example 1.2.17

The trivial space $\{\mathbf{0}\}$ has only one basis – the empty set $\emptyset$.

Example 1.2.18

The vector spaces of functions $f : \mathbb{N} \to \mathbb{R}$ and of functions $f : \mathbb{R} \to \mathbb{R}$ **do not** have bases because there is no reasonable way to enumerate functions with outputs in the real numbers.

Example 1.2.19

We have met bases before when studying sets of solutions of homogeneous systems; we only wouldn't call them such. The solution set of the linear system

$$\begin{aligned} x_1 + x_2 \quad\quad - x_4 &= 0 \\ x_3 + x_4 &= 0 \end{aligned}$$

is

$$\left\{ x_2 \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \mid x_2, x_4 \in \mathbb{R} \right\}.$$

Notice that the set is written as a span of two linearly independent vectors. In other words, its basis is the pair

$$\left(\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right).$$

Before we return to the original question of *existence* of a basis, we merge our current knowledge into a very important theorem which has both theoretical and computational consequences.

Theorem 1.2.20 (Characterisation of a basis)

Given a vector space V, an n-tuple $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is a basis of V if and only if every vector in V can be written as a linear combination of vectors in B in a ***unique*** *way.*

Proof. By definition of a basis, $\operatorname{span} B = V$ so indeed every vector in V must be expressible as a linear combination of vectors in B.

We now prove that this expression need be unique. For contradiction, assume that there exists a vector $\mathbf{v} \in V$ such that

$$\mathbf{v} = \sum_{i=1}^{n} r_i \cdot \mathbf{b}_i \quad \text{and also} \quad \mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{b}_i$$

with $r_j \neq t_j$ for at least one index $j \leq n$. We may rearrange

$$\begin{aligned} \sum_{i=1}^{n} r_i \cdot \mathbf{b}_i &= \sum_{i=1}^{n} t_i \cdot \mathbf{b}_i \\ \sum_{i=1}^{n} r_i \cdot \mathbf{b}_i - \sum_{i=1}^{n} t_i \cdot \mathbf{b}_i &= \mathbf{0} \\ \sum_{i=1}^{n} (r_i - t_i) \cdot \mathbf{b}_i &= \mathbf{0}. \end{aligned}$$

Since $r_j \neq t_j$ and thus $r_j - t_j \neq 0$, the linear combination on the left hand side has non-zero coefficients. By proposition 1.2.4, this means that B is linearly dependent. That's a contradiction with the assumption that it is a basis, hence such a vector $\mathbf{v}$ can't exist and the theorem is proven. ■

Unfortunately, we lack the theoretical background to fully answer the question of which vector spaces have bases and which don't. We can only define a class of vector spaces that **always** do have bases. Nevertheless, we can't prove that vector spaces outside of this class do not have bases – perhaps because it is not true ...

What we can say is that vector spaces which can be written as spans of vectors have a basis. This is actually a trivial consequence of lemma 1.2.8. Suppose a vector space V is spanned by a finite set of vectors $S \subseteq V$. We can keep removing vectors from S until we reach a set $S' \subseteq S$ which is linearly independent and $\operatorname{span} S' = \operatorname{span} S$. Any ordering of the set S' is now a basis of V. Indeed, it spans V and every vector in V can be written as a linear combination of vectors from S' in a unique way. The former statement is clear (by corollary 1.2.2) and the latter follows from the proof of theorem **??**. Should a vector $\mathbf{v} \in V$ have two different expressions in terms of vectors of S', we could subtract one from the other and get a non-trivial linear combination giving the zero vector – a contradiction with the linear independence of S' by proposition 1.2.4.

There is a point relevant to bases we should address. Intuitively, a basis of a space is the set of all possible *unique* directions of movement in that space. Wouldn't it be weird were we able to move in n possible ways in $\mathbb{R}^n$ when using the standard basis and, say, $n + 2$ ways when using some different basis? We tend to think of the dimension (or the total number of distinct directions of travel) of a space as something *fixed*, something inherent to the space itself, unrelated to any specific choice of vectors representing the directions.

As is thankfully often the case in linear algebra, our geometric intuition is correct. The formal way to express it is to say that all bases of a space should have the same number of elements. This is indeed the case and the number of elements in a basis is then called the *dimension* of said vector space.

First, we classify the vector spaces whereof we know they have a basis.

Definition 1.2.21 (Finitely generated vector space)

A vector space V is called *finitely generated* if it has a basis with finite number of vectors.

Remark 1.2.22

In the definition above, we specifically said 'has **a** basis' because we have not yet proven that all bases of a vector space have the same number of elements. Once we do so, finitely generated vector spaces can be seen as vector spaces of finite dimension.

Example 1.2.23

Every n-dimensional real space is finitely generated (take its standard basis for example) while the space of all functions $f : \mathbb{R} \to \mathbb{R}$ is not.

We prove the statement of equal number of elements across all bases in a somewhat roundabout way. This has the advantage of introducing a method of – somewhat algorithmically – transforming one basis of a space into another vector by vector.

Lemma 1.2.24 (Exchange lemma)

Assume V is a finitely generated vector space with basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and pick a vector $\mathbf{v} \in V$ given by the linear combination

$$\mathbf{v} = r_1 \cdot \mathbf{b}_1 + r_2 \cdot \mathbf{b}_2 + \ldots + r_n \cdot \mathbf{b}_n$$

with $r_i \neq 0$ for some $i \leq n$. Then, $\hat{B} \coloneqq \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{i-1}, \mathbf{v}, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n\}$ is also a basis of V.

Proof. We need to show that

(a) $\hat{B}$ is linearly independent.

(b) $\hat{B}$ spans V.

As for (a), assume we have a linear combination

$$t_1 \cdot \mathbf{b}_1 + \ldots + t_{i-1} \cdot \mathbf{b}_{i-1} + t_i \cdot \mathbf{v} + t_{i+1} \cdot \mathbf{b}_{i+1} + \ldots + t_n \cdot \mathbf{b}_n = \mathbf{0} \tag{1.5}$$

for some $t_1, \ldots, t_n \in \mathbb{R}$. Substituting for $\mathbf{v}$ gives

$$t_1 \cdot \mathbf{b}_1 + \ldots + t_{i-1} \cdot \mathbf{b}_{i-1} + t_i \cdot (r_1\mathbf{b}_1 + \ldots + r_n\mathbf{b}_n) + t_{i+1} \cdot \mathbf{b}_{i+1} + \ldots + t_n \cdot \mathbf{b}_n = \mathbf{0}.$$

Rearranging then

$$\begin{aligned}(t_1 + t_i r_1) \cdot \mathbf{b}_1 + \ldots + (t_{i-1} + t_i r_{i-1}) \cdot \mathbf{b}_{i-1} + t_i r_i \cdot \mathbf{b}_i \\ + (t_{i+1} + t_i r_{i+1}) \cdot \mathbf{b}_{i+1} + \ldots + (t_n + t_i r_n) \cdot \mathbf{b}_n = \mathbf{0}.\end{aligned} \tag{1.6}$$

This is a linear combination of vectors from the linearly independent basis B and thus by proposition 1.2.4, every coefficient of this combination is equal to 0. In particular, this means that $t_i r_i = 0$ and, since we've assumed $r_i \neq 0$, necessarily $t_i = 0$. However, in the wake of this, the combination (1.5) becomes

$$t_1 \cdot \mathbf{b}_1 + \ldots + t_{i-1} \cdot \mathbf{b}_{i-1} + t_{i+1} \cdot \mathbf{b}_{i+1} + \ldots + t_n\mathbf{b}_n = \mathbf{0},$$

i.e. a linear combination of vectors from B. Using proposition 1.2.4 again gives $t_j = 0$ for all $j \leq n$ since we already knew that $t_i = 0$. It follows that also $t_j + t_i r_j = 0$ for every $j \leq n$ and the linear combination in (1.6) has all coefficients equal to 0. This proves that $\hat{B}$ is linearly independent.

To prove (b), we check that $\operatorname{span} \hat{B} \subseteq \operatorname{span} B$ and $\operatorname{span} B \subseteq \operatorname{span} \hat{B}$. The inclusion $\operatorname{span} \hat{B} \subseteq \operatorname{span} B$ is obvious as $\mathbf{v}$ lies in $\operatorname{span} B$ (and so do all the vectors $\mathbf{b}_i$ of course). For the reverse inclusion to hold, it is enough to represent the exchanged vector $\mathbf{b}_i$ as linear combination of vectors from $\hat{B}$ because B and $\hat{B}$ share all the other vectors besides $\mathbf{b}_i$. In the linear combination

$$\mathbf{v} = r_1 \cdot \mathbf{b}_1 + \ldots + r_i \cdot \mathbf{b}_i + \ldots + r_n \cdot \mathbf{b}_n,$$

we assumed that $r_i \neq 0$. We can thus rearrange

$$\begin{aligned}\mathbf{v} &= r_1 \cdot \mathbf{b}_1 + \ldots + r_i \cdot \mathbf{b}_i + \ldots + r_n \cdot \mathbf{b}_n \\ -r_i \cdot \mathbf{b}_i &= r_1 \cdot \mathbf{b}_1 + \ldots + r_{i-1} \cdot \mathbf{b}_{i-1} + (-1) \cdot \mathbf{v} + r_{i+1} \cdot \mathbf{b}_{i+1} + \ldots + r_n \cdot \mathbf{b}_n \\ \mathbf{b}_i &= -\frac{r_1}{r_i} \cdot \mathbf{b}_1 + \ldots + \left(-\frac{r_{i-1}}{r_i}\right) \cdot \mathbf{b}_{i-1} + \left(-\frac{1}{r_i}\right) \cdot \mathbf{v} + \left(-\frac{r_{i+1}}{r_i}\right) \cdot \mathbf{b}_{i+1} + \ldots + \left(-\frac{r_n}{r_i}\right) \cdot \mathbf{b}_n\end{aligned}$$

which proves that $\mathbf{b}_i \in \operatorname{span} \hat{B}$ and with it, the lemma. ■

We intend to use the exchange lemma to prove that all bases of a finitely generated vector space have the same number of vectors by inductively exchanging the vectors of one basis for the vectors of another.

Theorem 1.2.25 (The dimension theorem)

All bases of a finitely generated vector space have the same number of elements.

Proof. Fix a vector space V and its basis $B := (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ with minimal number of elements. Given another basis $D = (\mathbf{d}_1, \ldots, \mathbf{d}_m)$, necessarily $n \leq m$ because the number of elements of B is assumed to be minimal. We shall prove that $m \leq n$.

The idea of the proof is to exchange all vectors in B for vectors in D until we get a basis of V consisting of only n vectors of D.

We proceed by induction on the number of exchanged vectors. Set $B_0 := B$. So far no vectors have been exchanged. Since B spans V and $\mathbf{d}_1 \in V$, there exists a linear combination

$$\mathbf{d}_1 = d_{1,1} \cdot \mathbf{b}_1 + d_{1,2} \cdot \mathbf{b}_2 + \ldots + d_{1,n} \cdot \mathbf{b}_n$$

with at least one $d_{1,i}$ non-zero (as the zero vector is always linearly dependent on others, thus $\mathbf{d}_1 \neq \mathbf{0}$). By the exchange lemma, we may exchange $\mathbf{d}_1$ for $\mathbf{b}_i$ and get the basis

$$B_1 := (\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}, \mathbf{d}_1, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n)$$

of V.

For the induction step, suppose the basis B_k has been formed by exchanging the vectors $\mathbf{d}_1, \ldots, \mathbf{d}_k \in D$ for exactly k vectors from B. Let us denote the set of indices of the remaining original vectors as $I \subseteq \{1, \ldots, n\}$. That is, $\mathbf{b}_i \in B_k$ if and only if $i \in I$. Pick $\mathbf{d}_{k+1} \in D$ and write

$$\mathbf{d}_{k+1} = \sum_{i=1}^{k} d_{k+1,i} \cdot \mathbf{d}_i + \sum_{i \in I} d_{k+1,i} \cdot \mathbf{b}_i$$

as a linear combination of vectors from B_k. The important observation to make is that at least one of the coefficients $d_{k+1,i}$, $i \in I$, must be non-zero. To see why, assume we have $d_{k+1,i} = 0$ for all $i \in I$. Then, the linear combination above assumes the form

$$\mathbf{d}_{k+1} = \sum_{i=1}^{k} d_{k+1,i} \cdot \mathbf{d}_i.$$

But, this means that $\mathbf{d}_{k+1}$ is a linear combination of other vectors from D. This contradicts the assumption that D is linearly independent and so this situation cannot arise.

Now that we know that there exists an index $i \in I$ such that $d_{k+1,i} \neq 0$, we may (again by the exchange lemma) exchange the vector $\mathbf{b}_i$ for $\mathbf{d}_{k+1}$ and form the basis B_{k+1}.

Upon having exchanged the last remaining vector $\mathbf{b}_i$ for $\mathbf{d}_n$, we have constructed the basis

$$B_n = (\mathbf{d}_1, \ldots, \mathbf{d}_n)$$

of the space V. Since B_n is linearly independent and spans V, it follows that $\mathbf{d}_{n+1}, \ldots, \mathbf{d}_m \in \operatorname{span} B_n$ which is a contradiction because B_n is a subset of D and D is assumed to be linearly independent. Thus, there must be no more vectors in D after $\mathbf{d}_n$ which proves that $m \leq n$ and with that also that $m = n$, as desired. ■

The dimension theorem has a few immediate consequences. For instance, we can finally define the dimension of any finitely generated vector space.

Definition 1.2.26 (Dimension)

Given a finitely generated vector space V, its *dimension* is the number of elements of any of its bases. We label it $\dim V$.

Example 1.2.27

The n-dimensional real space has dimension n. The testifying basis is $\mathcal{E}_n$, for example.

Example 1.2.28

The space of polynomials of degree at most n has dimension $n+1$. As we've partially observed, its standard basis is $(1, x, x^2, \ldots, x^n)$ which has $n+1$ elements.

Corollary 1.2.29

In a finitely generated vector space V, no linearly independent set $S \subseteq V$ can have more elements than the dimension of V.

Proof. Follows from the proof of the dimension theorem. Observe that in the proof we have never used the assumption that D spans V, only that it is linearly independent. ∎

Corollary 1.2.30

Any linearly independent set $S \subseteq V$ in a finitely generated vector space V can be expanded to a basis of V.

Proof. If $\operatorname{span} S \neq V$, then there exists a vector $\mathbf{v} \in V$ such that $\mathbf{v} \notin \operatorname{span} S$. By lemma 1.2.1, $S \subsetneq S \cup \{\mathbf{v}\}$ and $S \cup \{\mathbf{v}\}$ is linearly independent because $\mathbf{v} \notin \operatorname{span} S$. Hence, we simply keep adding vectors to S until $\operatorname{span} S = V$ and $\#S = \dim V$. ∎

Corollary 1.2.31

Any set $S \subseteq V$ with $\operatorname{span} S = V$ can be shrunk to a basis of the finitely generated vector space V.

Proof. If S is empty, then it spans the space $\{\mathbf{0}\}$ and is already a basis of it. If $S = \{\mathbf{0}\}$, then it also spans just the space $\{\mathbf{0}\}$ and we can remove the vector $\mathbf{0}$ from it, keeping its span.

Otherwise, S contains a vector $\mathbf{s}_1 \neq \mathbf{0}$. We form a basis $B_1 \coloneqq (\mathbf{s}_1)$. If $\operatorname{span} B_1 = \operatorname{span} S$, we're done. Otherwise, there exists a vector $\mathbf{s}_2 \in S$ such that $\mathbf{s}_2 \notin \operatorname{span} B_1$. Form $B_2 \coloneqq (\mathbf{s}_1, \mathbf{s}_2)$. This pair is linearly independent by the same argument as in the proof of corollary 1.2.30. We repeat this process until $\operatorname{span} B_n = \operatorname{span} S$ which takes exactly $\dim V$ steps. ∎

Corollary 1.2.32

In a vector space V with $\dim V = n$, an n-element set is linearly independent if and only if it spans V.

Proof. As for ($\Rightarrow$), any linearly independent set S can be expanded to a basis of V by corollary 1.2.30. Since a basis of V has n elements and so does S, there is no expansion to be done and any ordering of S is already a basis of V; in particular $\operatorname{span} S = V$.

The implication ($\Leftarrow$) is also immediate. If $\operatorname{span} S = V$, then by corollary 1.2.31, it can be shrunk to a basis of V, which has n elements. Since S also has n elements, no shrinking takes place and any ordering of S is again a basis of V and is thus linearly independent. ■

1.2.2 Representation With Respect To A Basis

Theorem 1.2.20 leads to a corollary of mainly computational importance: **every** vector in a vector space V with basis B corresponds to **exactly one** sequence of real coefficients of the linear combination of vectors from B that equals this vector.

To put this symbolically, denote $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ and consider a vector $\mathbf{v} \in V$. By the mentioned theorem 1.2.20, there exists exactly one n-tuple $(r_1, \ldots, r_n) \in \mathbb{R}^n$ such that

$$\mathbf{v} = r_1 \cdot \mathbf{b_1} + r_2 \cdot \mathbf{b}_2 + \ldots + r_n \cdot \mathbf{b}_n.$$

However, in chapter **??**, we observed that elements of $\mathbb{R}^n$ are really just n-dimensional vectors with entries in $\mathbb{R}$. These two facts brought together beget an important idea we shall formalise in due time – *vector spaces of dimension n are 'equivalent' to* $\mathbb{R}^n$. The last sentence should be read as such: in every vector space V, we can choose a basis B and write every vector in V as a linear combination of vectors from B. The coefficients of this linear combination (that are unique for every vector) can be assembled into a vector in $\mathbb{R}^n$. This forges a two-way relationship (a correspondence, if you will) between vectors in V and vectors in $\mathbb{R}^n$. We call this relationship a *representation* of the vector $v \in V$ for the reason that it gives a concrete form to an abstract vector.

Definition 1.2.33 (Representation of a vector)

Let V be a vector space with basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and $\mathbf{v} \in V$. We call the vector

$$[\mathbf{v}]_B \coloneqq \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{R}^n$$

a *representation of* $\mathbf{v}$ *with respect to* B if $\mathbf{v} = r_1 \cdot \mathbf{b}_1 + r_2 \cdot \mathbf{b}_2 + \ldots + r_n \cdot \mathbf{b}_n$.

Remark 1.2.34

The preceding definition underlines the necessity of defining a basis as a **sequence**, not just a set. A permutation of the elements of a basis changes the representation of many vectors with respect to it.

The notion of *representation* formalises the approach we've taken many times ere of 'writing' polynomials or matrices as vectors of coefficients. Confront the following example.

Example 1.2.35

In the space of cubic polynomials, the representation of the polynomial $x + x^2$ with respect to the basis $B = (1, 2x, 2x^2, 2x^3)$ is given by

$$[x + x^2]_B = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \\ 0 \end{pmatrix}.$$

With respect to a different basis $C = (1 + x, 1 - x, x + x^2, x + x^3)$, it instead looks like this:

$$[x + x^2]_C = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Problem 1.2.36

Find the representation of the vector

$$\mathbf{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

with respect to

$$B = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right).$$

Solution. We need to find real scalars $r_1, r_2 \in \mathbb{R}$ such that

$$r_1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + r_2 \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

This is tantamount to solving the linear system

$$\begin{array}{rcrcl} r_1 & & & = & 3 \\ r_1 & + & 2r_2 & = & 2 \end{array}$$

with obvious solution $r_1 = 3$ and $r_2 = -1/2$. With this, we've affirmed the equality

$$\left[\begin{pmatrix} 3 \\ 2 \end{pmatrix} \right]_B = \begin{pmatrix} 3 \\ -1/2 \end{pmatrix}.$$

♠

Example 1.2.37 (Representation with respect to canonical basis)

Since every vector $\mathbf{v} \in \mathbb{R}^n$ can be trivially broken into a linear combination of canonical basis vectors, its representation with respect to this basis are exactly its coordinates.

Expressed symbolically,

$$[\mathbf{v}]_{\mathcal{E}_n} = \left[\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}\right]_{\mathcal{E}_n} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

for every $\mathbf{v} \in \mathbb{R}^n$ because

$$\mathbf{v} = v_1 \cdot \mathbf{e}_1 + v_2 \cdot \mathbf{e}_2 + \ldots + v_n \cdot \mathbf{e}_n.$$

We intend not to dwell on the idea of representation any longer for now. It shall emerge again when we discuss linear transformations known as *changes of basis.* We close with a result concerning a link between linear independence and vector representation. In fact, linear independence of vectors is equivalent to the linear independence of their representations with respect to any basis.

Lemma 1.2.38

Let V be a vector space of dimension $n \in \mathbb{N}$ with basis B, $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and $r_1, \ldots, r_k \in \mathbb{R}$. Then,

$$r_1 \cdot \mathbf{v}_1 + r_2 \cdot \mathbf{v}_2 + \ldots + r_k \cdot \mathbf{v}_k = \mathbf{0}_V$$

if and only if

$$r_1 \cdot [\mathbf{v}_1]_B + r_2 \cdot [\mathbf{v}_2]_B + \ldots + r_k \cdot [\mathbf{v}]_k = \mathbf{0}_{\mathbb{R}^n}$$

where $\mathbf{0}_V$ is the zero vector of the space V and $\mathbf{0}_{\mathbb{R}^n}$ that of $\mathbb{R}^n$.

Proof. Write $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ and also denote

$$[\mathbf{v}_1]_B = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, [\mathbf{v}_2]_B = \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \ldots, [\mathbf{v}_k]_B = \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{n,k} \end{pmatrix}.$$

Then, the condition

$$r_1 \cdot \mathbf{v}_1 + r_2 \cdot \mathbf{v}_2 + \ldots + r_k \cdot \mathbf{v}_k = \mathbf{0}_V$$

is equivalent to

$$\begin{aligned} &r_1 \cdot (a_{1,1} \cdot \mathbf{b}_1 + \ldots + a_{n,1} \cdot \mathbf{b}_n) + \\ &r_2 \cdot (a_{1,2} \cdot \mathbf{b}_1 + \ldots + a_{n,2} \cdot \mathbf{b}_n) + \\ &\ldots + \\ &r_k \cdot (a_{1,k} \cdot \mathbf{b}_1 + \ldots + a_{n,k} \cdot \mathbf{b}_n) = \mathbf{0}_V. \end{aligned}$$

Grouping together coefficients of the basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ gives

$$\begin{aligned} &(r_1 \cdot a_{1,1} + r_2 \cdot a_{1,2} + \ldots + r_k \cdot a_{1,k}) \cdot \mathbf{b}_1 + \\ &(r_1 \cdot a_{2,1} + r_2 \cdot a_{2,2} + \ldots + r_k \cdot a_{2,k}) \cdot \mathbf{b}_2 + \\ &\ldots + \\ &(r_1 \cdot a_{n,1} + r_2 \cdot a_{n,2} + \ldots + r_k \cdot a_{n,k}) \cdot \mathbf{b}_n = \mathbf{0}_V. \end{aligned}$$

By proposition 1.2.4, this equality is satisfied if and only if each of the coefficients is equal to 0. On the horizon there glitters the homogeneous linear system

$$\begin{aligned} r_1 \cdot a_{1,1} + r_2 \cdot a_{1,2} + \ldots + r_k \cdot a_{1,k} &= 0 \\ r_1 \cdot a_{2,1} + r_2 \cdot a_{2,2} + \ldots + r_k \cdot a_{2,k} &= 0 \\ &\vdots \\ r_1 \cdot a_{n,1} + r_2 \cdot a_{n,2} + \ldots + r_k \cdot a_{n,k} &= 0, \end{aligned}$$

which can be rewritten (as we've done many times before) into vector form as

$$r_1 \cdot \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix} + r_2 \cdot \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{pmatrix} + \ldots + r_k \cdot \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{n,k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Shown vectors are of course just representations of the vectors $\mathbf{v}_1$ up to $\mathbf{v}_n$ with respect to B and so the result is proven. ∎

Exercise 1.2.39

Decide which of the following sets are linearly independent.

(a) $\left\{ \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ -4 \\ 14 \end{pmatrix} \right\}$;

(b) $\left\{ \begin{pmatrix} 1 \\ 7 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \\ 7 \end{pmatrix} \right\}$;

(c) $\left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \right\}$;

(d) $\left\{ \begin{pmatrix} 9 \\ 9 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ -4 \end{pmatrix}, \begin{pmatrix} 12 \\ 12 \\ -1 \end{pmatrix} \right\}$.

Exercise 1.2.40

Determine which of the sets are linearly independent in the space of quadratic polynomials.

(a) $\{3 - x + 9x^2, 5 - 6x + 3x^2, 1 + 1x - 5x^2\}$;

(b) $\{-x^2, 1 + 4x^2\}$;

(c) $\{2 + x + 7x^2, 3 - x + 2x^2, 4 - 3x^2\}$.

Exercise 1.2.41

Prove that each of the following sets is linearly independent in the vector space of all functions $f : (0, \infty) \to \mathbb{R}$.

(a) $\{x \mapsto x, x \mapsto \frac{1}{x}\}$;

(b) $\{x \mapsto \cos x, x \mapsto \sin x\}$;

(c) $\{x \mapsto \exp x, x \mapsto \log x\}$.

Exercise 1.2.42

Prove that the rows of a real-valued matrix in echelon form are a linearly independent set.

Exercise 1.2.43

Prove that if $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is a linearly independent set, then so are all its proper subsets: $\{\mathbf{x}, \mathbf{y}\}$, $\{\mathbf{x}, \mathbf{z}\}$, $\{\mathbf{y}, \mathbf{z}\}$, $\{\mathbf{x}\}$, $\{\mathbf{y}\}$, $\{\mathbf{z}\}$ and $\emptyset$. Is the converse also true?

Exercise 1.2.44

Is there a set of four vectors in $\mathbb{R}^3$ such that any three of them form a linearly independent set?

Exercise 1.2.45

Prove that a set of two perpendicular non-zero vectors in $\mathbb{R}^n$ is always linearly independent as long as $n > 1$. Generalise the result to more than two vectors.

Exercise 1.2.46

Decide whether $\{x^2 - x + 1, 2x + 1, 2x - 1\}$ and $\{x + x^2, x - x^2\}$ are bases of the space of quadratic polynomials.

Exercise 1.2.47

Find a basis for the solution set of the linear system

$$\begin{array}{rcrcrcrcl} x_1 & - & 4x_2 & + & 3x_3 & - & x_4 & = & 0 \\ 2x_1 & - & 8x_2 & + & 6x_3 & - & 2x_4 & = & 0. \end{array}$$

Exercise 1.2.48

Find a basis for $\mathbb{R}^{2\times 2}$, the space of 2×2 real matrices.

Exercise 1.2.49

Let $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ be a basis.

(a) Show that $(r_1 \cdot \mathbf{b}_1, r_2 \cdot \mathbf{b}_2, r_3 \cdot \mathbf{b}_3)$ is also a basis as long $r_1, r_2, r_3 \neq 0$. What happens if at least one of r_i **is** zero?

(b) Prove that $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ is also a basis where $\mathbf{a}_i = \mathbf{b}_1 + \mathbf{b}_i$, $i \in \{1, 2, 3\}$.

Exercise 1.2.50

Theorem **??** shows that, with respect to a basis, every linear combination is unique. If a subset is not a basis, can linear combinations be not unique? If so, must they be?

Exercise 1.2.51

Represent the polynomials

a) $2 + 4x^2$ b) $1 + 3x^2$ c) $1 + 5x^2$

with respect to the basis $B = (1 - x, 1 + x, x^2)$ of the space of quadratic polynomials. Use these representations to show that the three featured polynomials are linearly dependent.

Exercise 1.2.52

Represent the vector

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

with respect to the basis

$$B = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right)$$

of $\mathbb{R}^2$.

1.3 Linear Systems As Vector Spaces

We've defined an abstract vector space with the primary goal of accommodating the structure of sets of solutions of homogeneous linear systems. It might come as a surprise that linear systems *themselves* – not just their solutions – exhibit vector space structure. Understanding of said structure brings to light many properties of their solution sets, even.

Looking at the left hand side of linear systems (assembled into matrices), one immediately sees two sets of vectors – the **rows** of the matrix, and the **columns**. Enlightened as we have been by the late section 1.1 dealing with spans, we begin to study the vector spaces given by spans of these two sets. Believe it or not, they're actually closely related. Choosing a 'random' table of numbers, you cannot prevent its rows bearing a similar structure as the columns. Isn't that weird?

Definition 1.3.1 (Row space)

The *row space* of a matrix is the span of its rows.

Example 1.3.2

The row space of the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

is the vector space

$$\operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\right).$$

The row space of a matrix is tied to the solution of the corresponding linear system in a tight manner. Observe that the three 'row transformations' defined just prior to theorem **??** really just replace rows with linear combinations of other rows. That said, they do not alter the row space of a matrix. We shall formulate this observation as a lemma.

Lemma 1.3.3

Row transformations (1) - (3) defined above theorem **??** *do not change the row space of a matrix. I.e. if matrix B is a matrix derived from A by a series of row operations, then the row space of A equals the row space of B.*

Proof. We go through the row transformations one by one and check that they indeed do not shrink or enlarge the row space.

The operation of swapping two rows obviously doesn't change the row space as the span of a set of vectors is independent of their order.

Multiplying a vector of a set by a non-zero constant also clearly doesn't affect the span of the set.

Finally, assume $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the rows of A. The third row transformation amounts to replacing the row $\mathbf{a}_i$ with the row $\mathbf{a}_i + c \cdot \mathbf{a}_j$ for some $j \neq i$ and a constant $c \in \mathbb{R}$. Clearly, $\mathbf{a}_i + c \cdot \mathbf{a}_j$ lies in the row space of A. On the other hand, if B is the resulting matrix with rows $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_i + c \cdot \mathbf{a}_j, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n$, then $\mathbf{a}_i$ lies in the row space of B as

$$\mathbf{a}_i = 1 \cdot (\mathbf{a}_i + c \cdot \mathbf{a}_j) - c \cdot \mathbf{a}_j$$

and $\mathbf{a}_i$ is thus a linear combination of rows of B. This concludes the proof. ∎

In light of lemma 1.3.3, we may wish to formalise the intuition that Gauss-Jordan elimination in fact finds a **basis** of the row space of a matrix as it procedurally nullifies rows that can be expressed as linear combinations of preceding rows. The following lemma is an ingredient to that dish.

Lemma 1.3.4

The non-zero rows of a matrix in echelon form are linearly independent.

Proof. Each row of a matrix in echelon form has at least one more leading zero than the preceding row. That is, labelling the non-zero rows of the eliminated matrix A with n rows as

$\mathbf{a}_1, \ldots, \mathbf{a}_k$, consider the linear combination

$$r_1 \cdot \mathbf{a}_1 + r_2 \cdot \mathbf{a}_2 + \ldots + r_k \cdot \mathbf{a}_k = 0.$$

Rewriting this system in the form of a matrix gives

$$\left(\begin{array}{ccccc|c} a_{1,1} & 0 & 0 & \cdots & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,k} & 0 \end{array}\right).$$

We're certain that $a_{i,j} = 0$ if $j > i$ because the i-th row $\mathbf{a}_i$ (which is now the i-th column in the matrix) must have at least $i - 1$ leading zeroes.

Simple back substitution (starting on the first row) immediately yields $r_1 = r_2 = \ldots = r_k = 0$ and thus the row vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are linearly independent by proposition 1.2.4. ■

Let us now take a look at the vector space spanned by the columns of a matrix. We shall uncover interesting links to the row space.

Definition 1.3.5 (Column space)

The column space of a matrix is the span of its columns.

Example 1.3.6

The column space of the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

is the vector space

$$\operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}\right).$$

There is an obvious reason we care about studying the column space. Picking a matrix $A = (\tilde{\mathbf{a}}_1 | \cdots | \tilde{\mathbf{a}}_n)$ (i.e. we label its columns as $\tilde{\mathbf{a}}_1$ up to $\tilde{\mathbf{a}}_n$), every solution of the corresponding homogeneous linear system assumes the form

$$r_1 \cdot \tilde{\mathbf{a}}_1 + r_2 \cdot \tilde{\mathbf{a}}_2 + \ldots + r_n \cdot \tilde{\mathbf{a}}_n.$$

Consequently, the column space is **exactly** the vector space of solutions of the homogeneous linear system with matrix A.

In order to find a *basis* of a vector space given as a span of some set, we would assemble the spanning vectors into rows of a matrix and then put that into echelon form. Should we thus wish to find the basis for the column space, we would 'assemble the columns of a matrix into rows'. This matrix operation is called the *transpose*.

Definition 1.3.7 (Transpose of a matrix)

Given matrix $A = (\tilde{\mathbf{a}}_1 | \tilde{\mathbf{a}}_2 | \cdots | \tilde{\mathbf{a}}_n)$, its *transpose* is the matrix A^T with rows $\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \ldots, \tilde{\mathbf{a}}_n$.

Example 1.3.8

The transpose of the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

is the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}.$$

The circle has gone closed with the following lemma which establishes that row operations do not alter the column space. Restated, row operations do not change the solution set of a linear system. That is, in fact, the first theorem of the book – theorem ??.

Lemma 1.3.9

Row transformations (1) - (3) defined above theorem ?? *do not change the column space of a matrix.*

Proof. See the mentioned theorem ??. ■

We are now ready to present an important observation, one that ties together the dimension of row space to that of the column space. The crux of the matter is that Gauss-Jordan elimination actually doesn't find only the basis of the row space, it also finds the basis of the **column** space. We first illustrate why this is the case on an example.

Example 1.3.10

Let us put the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 3 & 3 & 1 & 0 \\ 7 & 8 & 2 & 4 \end{pmatrix}$$

into echelon form. Following the algorithm of Gauss-Jordan elimination gives

$$\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & -3 & 1 & -12 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

By lemma 1.3.4, the first and second rows of the eliminated matrix form the basis of the row space of the original matrix. That is,

$$\text{the row space of } A = \operatorname{span}\left(\begin{pmatrix} 1 \\ 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 1 \\ -12 \end{pmatrix}\right).$$

However, by the preceding lemma 1.3.9, the performed operations also left the column space intact. Can we see the basis for the column space in the eliminated matrix? Why, of course we can! The columns that correspond to **free variables** (in this case the third and fourth columns) are necessarily linearly dependent on previous columns. The reason for that is simple – their last non-zero entry is a pivot in the same row and some previous column, their penultimate non-zero entry is again a pivot in the same row as that entry and some previous column, etc. Therefore, for every non-zero entry of a free variable column, there exists some previous **pivot** column which has a non-zero entry at the same coordinate.

It follows that

$$\text{the column space of } A = \operatorname{span}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\-3\\0\end{pmatrix}\right).$$

Observe that the dimension of the column space is equal to that of the row space. This is not a coincidence – the echelon form of A has as many zero rows as there are free variables. But that is also exactly the number of columns that are linearly dependent on other columns. In other words, the dimension of **both** the row space **and** the column space is the number of pivots.

Let us summarise our findings in the following proposition.

Proposition 1.3.11

For every matrix A, the dimension of the row space of A is equal to the dimension of the column space of A.

Proof. The idea of the proof is pretty much contained in example 1.3.10. The number of non-zero rows in the echelon form of A equals the number of pivots (by definition of a *pivot*), and, by lemma 1.3.4, it also equals the dimension of its row space.

As we've observed in example 1.3.10, the number of pivots also equals the number of linearly independent columns and thus the dimension of the column space of A. ■

Definition 1.3.12 (Rank)

The *rank* of a matrix A equals the dimension of its row space or its column space and is denoted $\operatorname{rank} A$.

We finish the section strong by explicitly stating the relation between the rank of a matrix and the solution set of its associated homogeneous linear system.

Theorem 1.3.13

Let A be an $m \times n$ matrix. Then, the following claims are equivalent.

(1) $\operatorname{rank} M = r$.

(2) The vector space of solutions of the associated homogeneous linear system has dimension $n - r$.

Proof. By proposition 1.3.11 and the preceding example, $\operatorname{rank} M = r$ if and only if Gauss-Jordan elimination process of the matrix A ends with r non-zero rows. This in turn happens if and only if the number of pivots is exactly r. Finally, the number of pivots is r if and only if the number of free variables is $n - r$. The number of free variables is of course precisely the dimension of the set of solutions of the homogeneous linear system with matrix A. ■

Definition 1.3.14 (Regular matrix)

An $m \times n$ matrix is called *regular* if $\operatorname{rank} A = \min(m, n)$ (that is, the maximum possible). If A is not regular, it is called *singular*.

Remark 1.3.15

By theorem 1.3.13, a matrix A is *singular* if and only if the associated homogeneous linear system has infinitely many solutions.

Corollary 1.3.16

For a ***square*** *matrix A with n rows and n columns, the following claims are equivalent.*

(1) $\operatorname{rank} A = n$ *(A is regular).*

(2) The rows of A are linearly independent.

(3) The columns of A are linearly independent.

(4) ***Any*** *linear system (that is, not just homogeneous) with left side A has exactly one solution.*

Proof. The equivalences $(1) \Leftrightarrow (2)$ and $(1) \Leftrightarrow (3)$ follow from the fact that A is regular if and only if the row and column spaces of A both have dimension n. Since A has n rows and n columns, this means that both its rows and its columns must be linearly independent.

It remains to prove $(3) \Leftrightarrow (4)$. The columns of A (labelled $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n$) are linearly independent if and only if they form a basis of $\mathbb{R}^n$. Moreover, for any $\mathbf{b} \in \mathbb{R}^n$, the system

$$(\tilde{\mathbf{a}}_1 | \cdots | \tilde{\mathbf{a}}_n | \mathbf{b})$$

has a **unique** solution if and only if $\mathbf{b}$ can be represented as a linear combination of $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n$ in a **unique** way, that is, if and only if the columns of A form a basis of $\mathbb{R}^n$. ∎

Exercise 1.3.17

Decide if the vector

(a) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ lies in the row space of the matrix $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ lies in the row space of the matrix $\begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 1 \\ -1 & 2 & 7 \end{pmatrix}$.

Exercise 1.3.18

Decide if the vector

(a) $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ lies in the column space of the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ lies in the column space of the matrix $\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \\ 1 & -3 & -3 \end{pmatrix}$.

Exercise 1.3.19

Find the basis of both the row space and column space of the matrix

$$\begin{pmatrix} 2 & 0 & 3 & 4 \\ 0 & 1 & 1 & -1 \\ 3 & 1 & 0 & 2 \\ 1 & 0 & -4 & -1 \end{pmatrix}.$$

Exercise 1.3.20

Given $a, b, c \in \mathbb{R}$ what choice of $d \in \mathbb{R}$ will cause the following matrix to have rank one?

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Exercise 1.3.21

Find the column rank of the matrix

$$\begin{pmatrix} 1 & 3 & -1 & 5 & 0 & 4 \\ 2 & 0 & 1 & 0 & 4 & 1 \end{pmatrix}$$

Exercise 1.3.22

An $m \times n$ has *full row rank* if its row rank is m and has *full column rank* if its column rank is n.

(a) Show that a matrix can have both full row rank and full column rank only if it is square (that is, $m = n$).

(b) Prove that a linear system with matrix of coefficients A has a solution for **any** right side if and only if A has full row rank.

(c) Prove that a homogeneous system has a unique solution if and only if its matrix of coefficients A has full column rank.

(d) Prove that the statement '*if a system with matrix of coefficients A has any solution, then it has a unique solution*' holds if and only if A has full column rank.

Exercise 1.3.23

What is the relationship (if any) between

(a) $\operatorname{rank} A$ and $\operatorname{rank}(-A)$?

(b) $\operatorname{rank} A$ and $\operatorname{rank}(kA)$ for $k \neq 0$?

(c) $\operatorname{rank} A + \operatorname{rank} B$ and $\operatorname{rank}(A + B)$?

Chapter 2

Homomorphisms (chybějí obrázky)

In this chapter, our aim is to study and understand maps between vector spaces. Not just any kind of maps, however, but maps that *preserve structure.*

Most of modern mathematics is dedicated to the study of *structures* – basically prescribed rules of interaction between elements of a set. We call these rules, *operations*, and when moving from a set with structure to a set with structure by a map, we tend to require that said map somehow respects the structures of both sets. Such maps are often called *homomorphisms*, from Greek ὁμός („same") and μορφή („form, shape").

The only structure we consider in this book is that of a vector space given by two operations: scalar multiplication and vector addition. A *homomorphism between vector spaces* V *and* W (also called a *linear map*) is thus a map which respects both operations; in practice, this means that the image of a scalar multiple should be the same scalar multiple of the image and that the image of a sum of vectors should be the sum of the images.

One last note: we ought to be careful when comparing two structures. We labelled the operations on a vector space by symbols $\cdot$ and $+$ but these two symbols **mean different things in different vector spaces**! To keep the text tidy, we shan't resort to using yet another distinct pair of symbols. However, we *are* going to distinguish the structure in a small number of ensuing lemmata and definitions, to drive the point home.

Definition 2.0.1 (Homomorphism)

Let V and W be vector spaces over the field $\mathbb{F}$. We denote the operations of scalar multiplication and vector addition on V by $\cdot_V$ and $+_V$ and those on W by $\cdot_W$ and $+_W$. A map $f : V \to W$ is a *homomorphism* (or a *linear map*) if

(1) $f(\mathbf{v}_1 +_V \mathbf{v}_2) = f(\mathbf{v}_1) +_W f(\mathbf{v}_2)$ for every two vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$.

(2) $f(t \cdot_V \mathbf{v}) = t \cdot_W f(\mathbf{v})$ for every $t \in \mathbb{F}$ and $\mathbf{v} \in V$.

Example 2.0.2

The following maps are homomorphisms:

(a) the map $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by $f(\mathbf{v}) = 2 \cdot \mathbf{v}$;

(b) the map $f : \mathcal{P}_3(\mathbb{F}) \longrightarrow \mathbb{F}^4$ given by

$$f(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix},$$

where $\mathcal{P}_3(\mathbb{F})$ denotes the space of polynomials of degree 3 with coefficients in the field $\mathbb{F}$;

(c) the map $\pi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ given by

$$\pi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix};$$

maps that 'forget coordinates' are often called *projections*.

The following maps are **not** homomorphisms:

(a) the map $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ given by

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix};$$

(b) the map $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = x^2 + y^3 - 6.$$

(c) the map $f : \mathbb{R}^{2\times 2} \longrightarrow \mathbb{R}^2$ given by

$$f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a \cdot b + c \cdot d \\ a \cdot d - b \cdot c \end{pmatrix}.$$

In the previous example, we claimed that certain maps were homomorphisms without giving a proof. We did so because we first want to provide a characterisation of homomorphisms which makes checking whether a given map is a homomorphism somewhat easier. Hence, we now collect two qualities only homomorphisms possess.

Lemma 2.0.3 (Zero to zero)

Let $f : V \longrightarrow W$ be a homomorphism and label the zero vector of V by $\mathbf{0}_V$ and the zero vector of W by $\mathbf{0}_W$. Then,

$$f(\mathbf{0}_V) = \mathbf{0}_W.$$

Proof. Exploiting axiom (2) in the definition of homomorphism, we get

$$f(\mathbf{0}_V) = f(0 \cdot_V \mathbf{0}_V) \overset{(2)}{=} 0 \cdot_W f(\mathbf{0}_V) = \mathbf{0}_W$$

as required. ■

Lemma 2.0.4

For two vector spaces V, W over $\mathbb{F}$ and a map $f : V \longrightarrow W$, the following statements are equivalent.

(a) The map f is a homomorphism.

(b) For any two vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ and two numbers $t_1, t_2 \in \mathbb{F}$, we have

$$f(t_1 \cdot_V \mathbf{v}_1 +_V t_2 \cdot_V \mathbf{v}_2) = t_1 \cdot_W f(\mathbf{v}_1) +_W t_2 \cdot_W f(\mathbf{v}_2).$$

(c) For any vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and numbers $t_1, \ldots, t_n \in \mathbb{F}$, we have

$$f(t_1 \cdot_V \mathbf{v}_1 +_V \ \ldots \ +_V t_n \cdot_V \mathbf{v}_n) = t_1 \cdot_W f(\mathbf{v}_1) +_W \ \ldots \ +_W t_n \cdot_W f(\mathbf{v}_n).$$

Proof. In the proof (as well as the following text), we stop distinguishing between $\cdot_V$, $\cdot_W$ and $+_W$, $+_W$ for the sake of clarity. The readers should do well to keep in mind that V and W host different structures, though.

We prove $(a) \Leftrightarrow (b)$ and $(b) \Leftrightarrow (c)$.

As for $(a) \Longrightarrow (b)$, we shall, naturally, invoke the axioms (1) and (2) of the definition of homomorphism. We compute

$$f(t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2) \overset{(1)}{=} f(t_1 \cdot \mathbf{v}_1) + f(t_2 \cdot \mathbf{v}_2) \overset{(2)}{=} t_1 \cdot f(\mathbf{v}_1) + t_2 \cdot f(\mathbf{v}_2).$$

The implication $(b) \Longrightarrow (a)$ we simply substitute $t_1 = t_2 = 1$ for axiom (1) and $\mathbf{v}_2 = \mathbf{0}$ for axiom (2).

Similarly, the implication $(c) \Longrightarrow (b)$ follows trivially by $n = 2$. We prove the last implication $(b) \Longrightarrow (c)$ by induction on n. The base case $n = 2$ is covered completely by statement (b). For the induction step, label $\mathbf{w} = t_1 \cdot \mathbf{v}_1 + \ldots + t_n \cdot \mathbf{v}_n$. Then,

$$f(t_1 \cdot \mathbf{v}_1 + \ldots + t_n \cdot \mathbf{v}_n + t_{n+1} \cdot \mathbf{v}_{n+1}) = f(\mathbf{w} + t_{n+1} \cdot \mathbf{v}_{n+1}).$$

Using statement (b) again, we get

$$f(\mathbf{w} + t_{n+1} \cdot \mathbf{v}_{n+1}) = f(\mathbf{w}) + t_{n+1} \cdot f(\mathbf{v}_{n+1}).$$

By the induction hypothesis,

$$f(\mathbf{w}) = f(t_1 \cdot \mathbf{v}_1 + \ldots + t_n \cdot \mathbf{v}_n) = t_1 \cdot f(\mathbf{v}_1) + \ldots + t_n \cdot f(\mathbf{v}_n).$$

And thus,

$$\begin{aligned} f(\mathbf{w}) + t_{n+1} \cdot f(\mathbf{v}_{n+1}) &= f(t_1 \cdot \mathbf{v}_1 + \ldots + t_n \cdot \mathbf{v}_n) + t_{n+1} \cdot \mathbf{v}_{n+1} \\ &= t_1 \cdot f(\mathbf{v}_1) + \ldots + t_n \cdot f(\mathbf{v}_n) + t_{n+1} \cdot f(\mathbf{v}_{n+1}) \end{aligned}$$

and we're done. ■

Remark 2.0.5

The statement (b) in lemma 2.0.4 can be geometrically interpreted as saying that a homomorphism 'transforms parallelepipeds into parallelepipeds'. Let's see this on an example.

Any two linearly independent vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ define a parallelogram as the set of all linear combinations of $\mathbf{u}$ and $\mathbf{v}$ with coefficients between 0 and 1, i.e.

$$\mathbf{P}(\mathbf{u}, \mathbf{v}) \coloneqq \{a \cdot \mathbf{u} + b \cdot \mathbf{v} \mid a, b \in [0, 1]\}.$$

TODO obrazek

Then, the mentioned statement (b) says the following for a homomorphism $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(a \cdot \mathbf{u} + b \cdot \mathbf{v}) = a \cdot f(\mathbf{u}) + b \cdot f(\mathbf{v}).$$

However, this can be read to say that the image of every point in the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ is a point in the parallelogram determined by $f(\mathbf{u})$ and $f(\mathbf{v})$. Symbolically,

$$f(\mathbf{P}(\mathbf{u}, \mathbf{v})) = \mathbf{P}(f(\mathbf{u}), f(\mathbf{v})).$$

Let us return to example 2.0.2. There, we claimed that certain maps were homomorphisms without proof. For some of them, we're providing the proof now.

It is easily checked that the projection

$$\pi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}$$

is a homomorphism. Indeed, we can calculate

$$\begin{aligned}
\pi\left(a \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + b \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) &= \pi\left(\begin{pmatrix} ax_1 + bx_2 \\ ay_1 + by_2 \\ az_1 + bz_2 \end{pmatrix}\right) = \begin{pmatrix} ax_1 + bx_2 \\ ay_1 + by_2 \end{pmatrix} \\
&= a \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + b \cdot \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = a \cdot \pi\left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}\right) + b \cdot \pi\left(\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right).
\end{aligned}$$

The rest is up to lemma 2.0.4.

In a very similar vein, the map $f(\mathbf{v}) = 2 \cdot \mathbf{v}$ is a homomorphism for any vector space V over $\mathbb{R}$. Indeed, we may compute

$$f(a \cdot \mathbf{u} + b \cdot \mathbf{v}) = 2 \cdot (a \cdot \mathbf{u} + b \cdot \mathbf{v}) = a \cdot (2 \cdot \mathbf{u}) + b \cdot (2 \cdot \mathbf{v}) = a \cdot f(\mathbf{u}) = b \cdot f(\mathbf{v}).$$

This last homomorphism is an example of an *automorphism* – a bijective homomorphism from a space to itself. Automorphisms are just one interesting class of homomorphisms we shall present shortly.

The map

$$f\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right)=\begin{pmatrix}x\\y\\z\end{pmatrix}+\begin{pmatrix}1\\2\\-3\end{pmatrix}$$

is not a homomorphism because it doesn't send $\mathbf{0}$ to $\mathbf{0}$ (and thus contradicts lemma 2.0.3. It serves as a good example of a more general notion – homomorphisms *cannot* translate vectors. This is quite a stark restriction, yet it allows the images of homomorphisms to be vector spaces.

Also, the map

$$f\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=x^2+y^2-6$$

is equally **not** a homomorphism as it 'curves' the space. We may easily check that it breaks both axioms (1) and (2) in the definition. As we've observed by virtue of remark 2.0.5, homomorphisms send 'flat objects' to 'flat objects', not to misshapen ellipses.

The quality of maps we haven't fully commented on is left for kind readers to determine.

Exercise 2.0.6

Prove that the map $f:\mathcal{P}_3(\mathbb{F})\longrightarrow\mathbb{F}^4$ from example 2.0.2 is a homomorphism.

Exercise 2.0.7

Prove that the map $f:\mathbb{R}^{2\times 2}\longrightarrow\mathbb{R}^2$ from example 2.0.2 is **not** an isomorphism.

Now, onto learning some new words, kids, shall we?

Definition 2.0.8 (Classes of homomorphisms)

A homomorphism $f:V\longrightarrow W$ is called

(1) an *isomorphism*, if it is bijective,

(2) an *endomorphism*, if $W=V$,

(3) an *automorphism*, if it is both an *isomorphism* and an *endomorphism*.

If there exists an *isomorphism* between two vector spaces V and W, we call these spaces *isomorphic*. Intuitively, this means that the two spaces behave exactly the same, we have only chosen to represent the vectors of one a little differently than the other.

One immediate example is the correspondence between polynomials of degree n and vectors with $n+1$ entries we have mentioned many times throughout the text.

The map

$$\begin{aligned} f:\mathcal{P}_3(\mathbb{F})&\longrightarrow\mathbb{F}^4\\ a_0+a_1x+a_2x^2+a_3x^3&\mapsto\begin{pmatrix}a_0\\a_1\\a_2\\a_3\end{pmatrix}\end{aligned}$$

from example 2.0.2 is an isomorphism between $\mathcal{P}_3(\mathbb{F})$ and $\mathbb{F}^4$. The fact that it's both *injective* and *surjective* is almost obvious since it just sends a polynomial to its vector of coefficients.

We shan't spend more time discussing different classes of homomorphisms now but we certainly shall later. The readers are encouraged to make up examples themselves.

Exercise 2.0.9

Find an example of a homomorphism $f : V \to W$ which is

(a) an *isomorphism* but **not** an automorphism,

(b) an *endomorphism* but **not** an automorphism,

(c) an *automorphism*.

Now, there are quite a few special vector spaces tied to a homomorphism $f : V \to W$. As we've mentioned earlier, its image is a subspace of W, the preimage via f of a vector subspace of W is a subspace of V, and, finally, the set of homomorphisms themselves is a subspace of the vector space of all maps from $V \to W$. We are proving all these assertions now.

Before that however, we just briefly recall the important definitions. Given a map $f : V \to W$, its *image* is the set

$$\operatorname{im} f = f(V) = \{f(v) \mid v \in V\} \subseteq W.$$

The *preimage* of a subset $S \subseteq W$ via f is the set

$$f^{-1}(S) \coloneqq \{v \in V \mid f(v) \in S\} \subseteq V.$$

Finally, the set of all homomorphisms from V to W is denoted $\mathrm{Hom}(V, W)$.

Lemma 2.0.10

Let $f : V \to W$ be a homomorphism. Then,

(a) $f(V)$ is a subspace of W.

(b) $f^{-1}(U)$ is a subspace of V whenever U is a subspace of W.

(c) $Hom(V, W)$ is a subspace of the vector space of all maps $V \to W$.

Proof. To prove (a), pick two vectors $\mathbf{w}_1, \mathbf{w}_2 \in f(V)$. We shall prove that $t_1 \cdot \mathbf{w}_1 + t_2 \cdot \mathbf{w}_2 \in f(V)$. This will mean that $f(V)$ is a subspace by lemma 1.1.2. Since $\mathbf{w}_1, \mathbf{w}_2 \in f(V)$, there exist by definition vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $f(\mathbf{v}_1) = \mathbf{w}_1$ and $f(\mathbf{v}_2) = \mathbf{w}_2$. We thus rewrite

$$t_1 \cdot \mathbf{w}_1 + t_2 \cdot \mathbf{w}_2 = t_1 \cdot f(\mathbf{v}_1) + t_2 \cdot f(\mathbf{v}_2).$$

Since f is a homomorphism by assumption, the right side of the above equality can by virtue of lemma 2.0.4 be reshaped as such:

$$t_1 \cdot f(\mathbf{v}_1) + t_2 \cdot f(\mathbf{v}_2) = f(t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2)$$

and thus f sends the vector $t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2$ to $t_1 \cdot \mathbf{w}_1 + t_2 \cdot \mathbf{w}_2$. This, in particular, ascertains that $t_1 \cdot \mathbf{w}_1 + t_2 \cdot \mathbf{w}_2 \in f(V)$, as desired.

We continue with statement (b). Pick $\mathbf{v}_1, \mathbf{v}_2 \in f^{-1}(U)$. By definition of $f^{-1}(U)$, there exist $\mathbf{u}_1, \mathbf{u}_2 \in U$ such that $f(\mathbf{v}_1) = \mathbf{u}_1$ and $f(\mathbf{v}_2) = \mathbf{u}_2$. We thus compute

$$f(t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2) = t_1 \cdot f(\mathbf{v}_1) + t_2 \cdot f(\mathbf{v}_2) = t_1 \cdot \mathbf{u}_1 + t_2 \cdot \mathbf{u}_2.$$

The last linear combination lies in U as it is a subspace by assumption. It follows that $t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2 \in f^{-1}(U)$ since $f(t_1 \cdot \mathbf{v}_1 + t_2 \cdot \mathbf{v}_2) \in U$.

As for (c), we are given two homomorphisms $f, g \in \mathrm{Hom}(V, W)$. We must prove that $a \cdot f + b \cdot g$ is also a homomorphism for any $a, b \in \mathbb{F}$. For a change, we prove the axioms (1) and (2) in the definition of homomorphism. Taking $\mathbf{v}_1, \mathbf{v}_2 \in V$, we compute

$$\begin{aligned}
(a \cdot f + b \cdot g)(\mathbf{v}_1 + \mathbf{v}_2) &= (a \cdot f)(\mathbf{v}_1 + \mathbf{v}_2) + (b \cdot g)(\mathbf{v}_1 + \mathbf{v}_2) \\
&= f(a \cdot \mathbf{v}_1 + a \cdot \mathbf{v}_2) + g(b \cdot \mathbf{v}_1 + b \cdot \mathbf{v}_2) \\
&= a \cdot f(\mathbf{v}_1) + a \cdot f(\mathbf{v}_2) + b \cdot g(\mathbf{v}_1) + b \cdot g(\mathbf{v}_2) \\
&= (a \cdot f(\mathbf{v}_1) + b \cdot g(\mathbf{v}_1)) + (a \cdot f(\mathbf{v}_2) + b \cdot g(\mathbf{v}_2)) \\
&= (a \cdot f + b \cdot g)(\mathbf{v}_1) + (a \cdot f + b \cdot g)(\mathbf{v}_2),
\end{aligned}$$

hence (1) holds. The proof of (2) is left as an exercise. The validity of both axioms ascertains that $\mathrm{Hom}(V, W)$ is really a vector space and thus a subspace of the space of all maps $V \longrightarrow W$. ■

Exercise 2.0.11

Prove that for two homomorphisms $f, g \in \mathrm{Hom}(V, W)$, $a, b \in \mathbb{F}$, $\mathbf{v} \in V$ and $t \in \mathbb{F}$, we have

$$(a \cdot f + b \cdot g)(t \cdot \mathbf{v}) = t \cdot (a \cdot f + b \cdot g)(\mathbf{v}).$$

The fact that images and preimages of homomorphisms are vector spaces have geometric consequences. We now illustrate those on a few examples.

Example 2.0.12

Consider the projection $\pi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ prescribed as

$$\pi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}.$$

The image of $\pi(\mathbb{R}^3)$ is of course the entirety of $\mathbb{R}^2$. One may think of this projection as the 'squishing' of an entire room onto its floor. If we're allowed to move only along one of the walls, we're confined to just one edge of the floor. This is formalised by the fact that the image of the set of vectors with $y = 0$ is the subspace of $\mathbb{R}^2$ of vectors with $y = 0$, and similarly for x.

TODO obrazky

The preimage of a given vector $\mathbf{w} \in \mathbb{R}^2$ via π is the set of all vectors whose tips sit on the vertical line rooted at the tip of $\mathbf{w}$. This is because any vector

$$\begin{pmatrix} w_1 \\ w_2 \\ z \end{pmatrix}$$

gets mapped onto

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Of course, a set consisting of a single non-zero vector is not a subspace, just as the vertical line rooted at the tip of such vector is not.

On a similar note, the preimage of a line $\{t \cdot \mathbf{v} \mid t \in \mathbb{R}\}$ for a given $\mathbf{v} \in \mathbb{R}^2$ is the entire vertical plane containing said line in $\mathbb{R}^3$.

Example 2.0.13

The image of the homomorphism

$$h : \mathbb{R}^2 \to \mathbb{R}$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + y$$

is the entirety of $\mathbb{R}$. Fixing a number $r \in \mathbb{R}$, its preimage via h is the set of vectors whose coordinates add up to r. Their tips form a line in $\mathbb{R}^2$ given by the equation $x + y = r$.

The fact that h is a homomorphism can be expressed in a neat geometric manner – fix now two numbers $r_1, r_2 \in \mathbb{R}$. Think of the vectors whose tips lie on the line $x+y = r_1$ as 'r_1' vectors. Analogously, vectors whose tips form the line $x + y = r_2$ are regarded as 'r_2' vectors. Axiom (1) in the definition of homomorphism can be, in this particular case, restated as 'r_1' vectors plus 'r_2' vectors equal '$r_1 + r_2$' vectors. That is, $\mathbf{v}_1$ is an 'r_1' vector if $h(\mathbf{v}_1) = r_1$ and $\mathbf{v}_2$ is an 'r_2' vector if $h(\mathbf{v}_2)$; the previous sentence thus signifies exactly that $h(\mathbf{v}_1 + \mathbf{v}_2) = h(\mathbf{v}_1) + h(\mathbf{v}_2)$ since $\mathbf{v}_1 + \mathbf{v}_2$ is clearly a '$r_1 + r_2$' vector.

TODO obrazek

Example 2.0.14

Define the 'derivative' homomorphism

$$\frac{\partial}{\partial x} : \mathcal{P}_3 \to \mathcal{P}_3$$
$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mapsto a_1 + 2a_2 x + 3a_3 x^2.$$

Its image is the subspace $\mathcal{P}_2$ (polynomials of degree 2 over $\mathbb{R}$) of $\mathcal{P}_3$. The inverse image of set of polynomials of degree 2 is the 'integral' homomorphism (with constant 0). Said formally, there exists a homomorphism

$$\int : \mathcal{P}_2 \to \mathcal{P}_3$$
$$a_0 + a_1 x + a_2 x^2 \mapsto a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3$$

such that $\frac{\partial}{\partial x} \circ \int$ is the identity map on $\mathcal{P}_2$. Notice however that it is **not** the case that $\int \circ \frac{\partial}{\partial x}$ is the identity map on $\mathcal{P}_3$. The derivative $\frac{\partial}{\partial x}$ is **not** an injective map and as such it doesn't have a 'two-sided' inverse, it only has a 'right' inverse in the form of $\int$.

The last examples provokes a question: 'When does a homomorphism have an inverse which is also a homomorphism?' We remind dear readers that an inverse to a map $f : V \to W$ is a map $f^{-1} : W \to V$ such that $(f \circ f^{-1})(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$ and $(f^{-1} \circ f)(\mathbf{v}) = \mathbf{v}$ for every $\mathbf{v} \in V$. As seen in the mentioned example 2.0.14 **just one** of these equalities **is not enough**. A homomorphism may be invertible from just one side.

The rest of the introductory section to homomorphisms is dedicated to answering this question. First, we must establish a strong connection between homomorphisms and bases of their domains. We illustrate this connection first before conjuring a proof.

Consider a homomorphism f with domain $\mathbb{R}^3$. By lemma 2.0.4, the image of a vector

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

via this homomorphism can be broken into the linear combination

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = f\left(\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}\right) + f\left(\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}\right) + f\left(\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}\right) = x \cdot f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + y \cdot f\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + z \cdot f\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right).$$

This immediately suggests that the image of any vector via a homomorphism is determined purely by the images of standard basis vectors. Indeed, this statement is true in general, for any homomorphism and any basis.

Theorem 2.0.15

*Let V, W be vector spaces over $\mathbb{F}$, $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ be a basis of V. Given (not necessarily distinct) vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in W$, there exists **just a single** homomorphism $f \in \mathrm{Hom}(V, W)$ such that $f(\mathbf{b}_1) = \mathbf{w}_1, f(\mathbf{b}_2) = \mathbf{w}_2, \ldots, f(\mathbf{b}_n) = \mathbf{w}_n$.*

Proof. First, given $\mathbf{v} \in V$, it can be written uniquely as a linear combination of vectors from B, that is, there exist unique coefficients $t_1, \ldots, t_n \in \mathbb{F}$ such that

$$\mathbf{v} = t_1 \cdot \mathbf{b}_1 + \ldots + t_n \cdot \mathbf{b}_n.$$

We define $f(\mathbf{v})$ by

$$f(\mathbf{v}) = t_1 \cdot \mathbf{w}_1 + \ldots + t_n \cdot \mathbf{w}_n.$$

We must prove three statements:

(1) f is well-defined (i.e. it's *actually* a map $V \to W$);

(2) f is a homomorphism;

(3) f is unique.

The statement (1) follows from the uniqueness of representation of a vector with respect to a basis – content of theorem 1.2.20. Consequently, as the vector $\mathbf{v}$ can only ever be represented by the same n-tuple of coefficients, the image of $\mathbf{v}$ under f is always the same.

Ad (2), choose $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $t_1, t_2 \in \mathbb{F}$. We calculate