Math 79SI Notes

Mason Rogers

9 Vector Spaces I

9.1 Background

In the previous section, we looked at how a proof motivated by intuition in $\mathbb{R}^n$ actually applied to any vector space equipped with a suitable 'dot product'. Remarkably, we were able to extend a result that a priori seems dependent on the geometry of $\mathbb{R}^n$ to vector spaces that do not have any evident sense of 'geometry'. In particular, we saw the significance of the Cauchy-Schwarz Inequality in some vector spaces that do not have 'standard bases'.

That $\mathbb{R}^n$ has a 'standard basis' is convenient, for we can prove many theorems about $\mathbb{R}^n$ by writing out matrix equations. But those theorems rarely require the standard basis beyond computations. Additionally, many important vector spaces have no preferred basis, so we want a way to treat vector spaces without excessive reliance on solving systems of equations or using the notion of a 'standard basis'.

Here is the definition of a vector space over $\mathbb{R}$:

Definition 9.1. A *vector space* V is a set equipped with the operations of 'addition' and 'scalar multiplication by real numbers' such that

- Addition is associative and commutative (i.e. $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ and $v_1 + v_2 = v_2 + v_1$).
- There exists an element $\vec{0} \in V$ such that $v + \vec{0} = v$ for all $v \in V$.
- For all $v \in V$ there exists $-v \in V$ such that $v + (-v) = \vec{0}$.
- For all $a, b \in \mathbb{R}$ and all $v \in V$, $a(bv) = (ab)v$.
- For all $v \in V$, $1v = v$.
- Scalar multiplication distributes over addition in V (i.e. $a(v_1 + v_2) = av_1 + av_2$).
- Scalar multiplication distributes over addition in $\mathbb{R}$ (i.e. $(a_1 + a_2)v = a_1v + a_2v$).

It is convenient to think of the required properties of vector space 'arithmetic' as exactly what one might expect based on vector addition/scalar multiplication in $\mathbb{R}^n$.

Elements of such V are called *vectors*. Note that V does not necessarily have a notion of 'length' and 'direction', so the saying that "a vector is an object with a magnitude and a direction" is incorrect. Instead, vectors are simply elements of vector spaces, and 'length' and 'direction' are concepts we *define* on top of vector space structure. The important feature of a vector space is its linear structure, and any other geometric ideas we wish to consider in the context of a vector space require additional data defined independently from the linear structure.

Observe that the definition does not require V to have a 'preferred basis', or even a basis at all. In particular, consider the following examples of vector spaces.

Example 9.2. Consider the second-order differential equation $y'' + y = 0$. Observe that for any solutions y_1 and y_2, $y_1 + y_2$ is also a solution since

$$(y_1 + y_2)'' + (y_1 + y_2) = (y_1'' + y_1) + (y_2'' + y_2) = 0 + 0 = 0.$$

Additionally, for any solution y and any $c \in \mathbb{R}$, cy is also a solution because

$$(cy)'' + (cy) = c(y'' + y) = c \cdot 0 = 0.$$

Lastly, $y = 0$ is a solution, so the set of solutions to $y'' + y = 0$ is a vector space! Moreover, it is known that any solution to this differential equation has the form $a\cos(x) + b\sin(x)$ for some $a, b \in \mathbb{R}$.

Example 9.3. Consider a homogeneous third-order linear differential equation with highest order coefficient 1:

$$f''' + P(x)f'' + Q(x)f' + R(x)f = 0.$$

Solutions to this differential equation are functions $f(x)$ which satisfy the differential equation. Because $(f_1 + f_2)' = f_1' + f_2'$ (and likewise for f'' and f''') for any differentiable functions, if f_1 and f_2 are solutions of the differential equation, so is $f_1 + f_2$:

$$\begin{aligned}
&(f_1 + f_2)''' + P(x)(f_1 + f_2)'' + Q(x)(f_1 + f_2)' + R(x)(f_1 + f_2) \\
&\qquad = (f_1''' + P(x)f_1'' + Q(x)f_1' + R(x)f_1) + (f_2''' + P(x)f_2'' + Q(x)f_2' + R(x)f_2) \\
&\qquad = 0 + 0 = 0.
\end{aligned}$$

Additionally, since $(af)' = a(f')$ for all $a \in \mathbb{R}$ and all differentiable functions f (and likewise for f'' and f'''), if f is a solution and $a \in \mathbb{R}$, then af is also a solution:

$$\begin{aligned}
&(af)''' + P(x)(af)'' + Q(x)(af)' + R(x)(af) \\
&\qquad = a(f''' + P(x)f'' + Q(x)f' + R(x)f) \\
&\qquad = a \cdot 0 = 0.
\end{aligned}$$

Lastly, the zero function is a solution. Thus, the set of solutions to this differential equation is a vector space! However, it is not clear what a basis of this space would look like, and even in cases where bases can be found, there is not generally a choice of a "standard" basis. In fact, it is not clear that there are even nonzero solutions to this ODE, but we do not need to know how many solutions there are or any particular solutions to verify that the solution set is indeed a vector space.

Example 9.4. Consider the set of continuous functions $f : \mathbb{R} \to \mathbb{R}$. Because sums and scalar multiples of continuous functions are continuous, this set forms a vector space. Again, it is unclear how to choose a basis for this space, and even if it is possible, there might not be a 'best choice'.

In light of these ideas, we want to develop the theory of vector spaces in such a way that we do not rely on explicit formulas in $\mathbb{R}^n$, but instead spaces with operations having the same *properties* as in $\mathbb{R}^n$. In particular, we will focus on 'finite dimensional' vector spaces. We will look at questions that can be asked for general vector spaces and then understand what they mean in specific cases, much as we did with the Cauchy-Schwarz Inequality.

9.2 Examples

We begin by recalling a few notions from Math 51 which will be relevant in this section. Let V be a vector space.

Definition 9.5. A *linear combination* of vectors $\{v_1, \ldots, v_n\} \subset V$ is an element $v \in V$ given by $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$ for some $c_i \in \mathbb{R}$.

Definition 9.6. The *span* of a set of vectors $\{v_1, \ldots, v_n\}$ is the set of all linear combinations of the v_i. In other words, it is the set of all $v \in V$ that can be written as $c_1v_1+c_2v_2+\ldots+c_nv_n$ for some $c_i \in \mathbb{R}$.

Definition 9.7. A set of vectors $\{v_1, \ldots, v_n\}$ is *linearly independent* if the only way to write $\vec{0}$ as a linear combination of $\{v_1, \ldots, v_n\}$, i.e. to write $\vec{0} = c_1v_1 + \ldots + c_nv_n$, is for c_i to equal 0 for all $i \in \{1, \ldots, n\}$.

We will begin by examining these concepts with some examples in $\mathbb{R}^4$—beyond the range of visualization, but still in the familiar context of Euclidean space. We will then consider a more abstract setting: the vector space of continuous functions $f : [0, 1] \to \mathbb{R}$.

Example 9.8. Consider the vectors $u = (1, 0, -2, 1), v = (1, 1, 0, -1), w = (3, 2, -2, 0)$. A linear combination of these vectors is any vector of form $au + bv + cw$, where $a, b, c \in \mathbb{R}$, and $\operatorname{span}(u, v, w)$ is the set of all vectors of form $au + bv + cw$.

We now will investigate whether or not $\{u, v, w\}$ is a linearly independent set. Suppose we have $au + bv + cw = \vec{0}$ for some real constants a, b, c. For the second coordinate of $au+bv+cw$ to equal 0, we need $b+2c = 0$. Likewise, setting the third and fourth coordinates of $au + bv + cw$ to 0 gives the equations $a + c = 0$ and $a - b = 0$; these three equations have the unique solution $(a, b, c) = (0, 0, 0)$, so we conclude that $\{u, v, w\}$ is a linearly independent set (without even having to write out a fourth equation).

However, suppose we replace w with $w' = (3, 2, -2, -1)$. Repeating the above process gives a system of four equations for three unknowns, which has the (not unique) nonzero solution $(a, b, c) = (1, 2, -1)$. In other words, we have $\vec{0} = u + 2v - w'$, so $\{u, v, w'\}$ is not a linearly independent set.

Example 9.9. Let V be the vector space of continuous functions $f : [0, 1] \to \mathbb{R}$. We claim that the functions $\sin(x)$ and $\cos(x)$ are linearly independent in V.

This draws the important distinction between treating functions like $\sin(x)$ and $\cos(x)$ as objects to be evaluated on an input variable and treating functions as objects *in and of themselves*. From the former perspective, it may be tempting to find a nontrivial linear combination of $\sin(x_0)$ and $\cos(x_0)$ which comes out to 0 for some particular $x_0 \in [0,1]$;[1] however, we cannot choose a, b not both equal to 0 such that $a\sin(x) + b\cos(x)$ is *identically the zero function.* To see why, observe that $\sin(0) = 0$ and $\cos(0) = 1$, so for $a\sin(0)+b\cos(0)$ to equal 0, we need $b = 0$. Then for $a\sin(x) + 0\cos(x)$ to equal 0 at any $x \in (0,1]$, we need $a = 0$. Thus, the only way for $a\sin(x) + b\cos(x)$ to equal 0 for every $x \in [0,1]$ is to have $a = b = 0$. From this, we conclude that $\sin(x)$ and $\cos(x)$, seen *as vectors in the space of continuous functions* $f : [0,1] \to \mathbb{R}$, are linearly independent.

A piece of intuition for the definition of linear independence is that linearly independent sets have no 'redundant' elements in the sense that none of the elements can be written as a linear combination of the others:

Proposition 9.10. A set $\{v_1, \ldots, v_n\}$ is linearly independent if and only if for all $i \in \{1, \ldots, n\}$, $v_i \notin \operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$.

Proof. We will prove this biconditional statement by proving that if $\{v_1, \ldots, v_n\}$ is linearly *dependent* then there exists some $i \in \{1, \ldots, n\}$ such that $v_i \in \operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$, then proving the converse of that statement.

Suppose $\{v_1, \ldots, v_n\}$ is a linearly dependent set, so we can choose $a_1, \ldots, a_n$ not all equal to 0 such that $a_1v_1 + \ldots + a_nv_n = \vec{0}$. Without loss of generality (by relabeling indices), we may and do suppose $a_1 \neq 0$. We then have $a_1v_1 = -a_2v_2 - \ldots - a_nv_n$ and can divide through by $a_1 \neq 0$ to obtain

$$v_1 = -\frac{1}{a_1}(a_2v_2 + \ldots + a_nv_n) = -\frac{a_2}{a_1}v_2 - \ldots - \frac{a_n}{a_1}v_n,$$

thereby expressing v_1 as a linear combination of the elements of $\{v_2, \ldots, v_n\}$. Thus, $v_1 \in \operatorname{span}(v_2, \ldots, v_n)$, completing the first half of the proof.

Now suppose that for some $i \in \{1, \ldots, n\}$, $v_i \in \operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$, and without loss of generality, let $i = 1$. We can then write $v_1 = c_2v_2 + \ldots + c_nv_n$ for some $c_2, \ldots, c_n \in \mathbb{R}$. Subtracting v_1 from both sides gives

$$\vec{0} = -v_1 + c_2v_2 + \ldots + c_nv_n = (-1)v_1 + c_2v_2 + \ldots + c_nv_n,$$

which is a linear combination of the elements of $\{v_1, \ldots, v_n\}$ having v_1-coefficient $-1 \neq -0$ which adds up to $\vec{0}$. Thus, the set $\{v_1, \ldots, v_n\}$ is linearly dependent, completing the proof. $\square$

Linearly independent sets are helpful because every element of their spans can be written *uniquely* as a linear combination of elements of the set:

Proposition 9.11. Suppose $\{v_1, \ldots, v_n\}$ is a linearly independent set. Then if

$$a_1v_1 + \ldots + a_nv_n = v = c_1v_1 + \ldots + c_nv_n,$$

we have $a_i = c_i$ for all $i \in \{1, \ldots, n\}$.

[1] In fact, for *any* $x_0 \in \mathbb{R}$, $\sin(x_0)$ and $\cos(x_0)$ are linearly dependent elements of $\mathbb{R}$. Because for all $x_0 \in \mathbb{R}$, at least one of $\sin(x_0)$ and $\cos(x_0)$ is nonzero, we can choose $a = \cos(x_0)/\sin(x_0), b = -1$ when $\sin(x_0) \neq 0$ and $a = -1, b = \sin(x_0)/\cos(x_0)$ when $\cos(x_0) \neq 0$ to obtain $a\cos(x_0) = b\sin(x_0) = 0$ with at least one of a, b not equal to 0.

Proof. By the definition of linear independence, $\vec{0}$ can be represented *uniquely* as a linear combination of the elements of $\{v_1, \ldots, v_n\}$ (i.e. the linear combination with all coefficients 0). We want to extend this to say that all vectors in $\operatorname{span}(v_1, \ldots, v_n)$ are given by a *unique* linear combination of the elements of $\{v_1, \ldots, v_n\}$, which we will do by using the 'linear' structure of the space. Suppose there exist some real numbers $a_1, \ldots, a_n, c_1, \ldots, c_n$ such that

$$a_1v_1 + \ldots + a_nv_n = c_1v_1 + \ldots + c_nv_n.$$

Subtracting $c_1v_1 + \ldots + c_nv_n$ from both sides gives

$$a_1v_1 + \ldots + a_nv_n - c_1v_1 \ldots - c_nv_n = \vec{0},$$

and rearranging via the properties of vector space arithmetic gives

$$(a_1 - c_1)v_1 + \ldots + (a_n - c_n)v_n = \vec{0}.$$

By the definition of linear independence, $a_1 - c_1 = 0, \ldots, a_n - c_n = 0$, so we conclude that $a_i = c_i$ for all $i \in \{1, \ldots, n\}$. □

We next want a procedure to take a linearly *dependent* set and remove the 'redundant' vectors to obtain a linearly independent list with the same span. To do so, we will show that we can remove one 'redundant' vector from the set at a time without affecting the set's span, then induct on the process of removing 'redundant' vectors.

Lemma 9.12 (Linear Dependence Lemma)**.** If $S = \{v_1, \ldots, v_n\} \subset V$ is a linearly dependent set and $v_1 \neq \vec{0}$, there exists $j \in \{2, \ldots, n\}$ such that

- $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$ and
- $\operatorname{span}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) = \operatorname{span}(v_1, \ldots, v_n)$, i.e. removing v_j from S does not affect the span of S.

Intuitively, our 'redundant' vector is a vector that lies in the span of the previous vectors, and removing it does not affect the span of the set.

Proof. Because the v_i are linearly dependent, we can write $\vec{0} = c_1v_1 + \ldots + c_nv_n$ for some c_i not all equal to 0. Let j be the index of the largest nonzero c_i, and note $j > 1$ since otherwise $a_1v_1 = 0$ with $a_1 \neq 0$, which is impossible since $v_1 \neq \vec{0}$. Because all of the c_i vanish for $i > j$, we have $\vec{0} = c_1v_1 + \ldots + c_jv_j$, which we can rearrange to obtain $v_j = (-1/c_j)(c_1v_1 + \ldots + c_{j-1}v_{j-1})$. We have therefore expressed v_j as a linear combination of elements in the set $\{v_1, \ldots, v_{j-1}\}$, so $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$.

We are left to prove that the span of $\{v_1, \ldots, v_n\}$ is unaffected if v_j is removed. To show this, we must show that an arbitrary linear combination of the v_i can be rewritten *without* a v_j term. Indeed, given $v = a_1v_1 + \ldots + a_jv_j + \ldots + a_nv_n$, we can substitute our above equation for v_j, giving

$$v = a_1v_1 + \ldots + a_{j-1}v_{j-1} - \frac{a_j}{c_j}(c_1v_1 + \ldots + c_{j-1}v_{j-1}) + a_{j+1}v_{j+1} + \ldots + a_nv_n$$

$$= \left(a_1 - \frac{a_j}{c_j}c_1\right) + \ldots + \left(a_{j-1} - \frac{a_j}{c_j}c_{j-1}\right) v_{j-1} + a_{j+1}v_{j+1} + \ldots + a_n v_n.$$

We have written v as a linear combination of the elements of $\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots v_n\}$, so $\text{span}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) = \text{span}(v_1, \ldots, v_n)$, concluding the proof. □

Proposition 9.13. Given any finite set $S = \{v_1, \ldots, v_n\} \subset V$ with $v_1 \neq \vec{0}$, there exists a linearly independent subset $T = \{v_{i_1}, \ldots, v_{i_k}\} \subset S$ with the same span.

Proof. We will induct on the size of S using Lemma 9.12. As the base case, observe that any set with one vector (which is nonzero) is linearly independent, so we are done. Suppose that the statement in the proposition is true for sets of size $n-1$ with $n-1 \geq 1$; we will show that it is also true for a set S of size n. If S is linearly independent, we are done. Otherwise by Lemma 9.12 we can choose a vector $v_j \in S$ such that $S \setminus \{v_j\}$ has the same span as S. Then by the inductive hypothesis, $S \setminus \{v_j\}$ admits a linearly independent subset T such that $\text{span}(T) = \text{span}(S \setminus \{v_j\}) = \text{span}(S)$. □

We have now proven that given a finite set of vectors, we can choose a linearly independent subset with the same span. The approach we took is rather theoretical, avoiding row reduction in $\mathbb{R}^n$ to gain conceptual efficiency at the cost of being non-constructive. However, if you think carefully about what you do when you perform row reduction operations, you will realize that this method is the same method in disguise.

Next, we will restrict our attention to 'finite dimensional' vector spaces and develop an understanding of the concept of dimension. First, we must define our new setting:

Definition 9.14. A nonzero vector space V is *finitely generated* if it admits a finite spanning set.

Some familiar spaces are finitely generated, such as $\mathbb{R}^n$ and the set of solutions to the class of third-order differential equations in Example 9.3. This is because $\mathbb{R}^n$ can be written as the span of the n vectors e_i, and the solution space to the aforementioned differential equation is the span of three functions due to a rather serious theorem on linear ordinary differential equations:

Theorem 9.15. The space of solutions to an ordinary differential equation of form

$$f^{(n)} + \sum_{i=0}^{n} P_i(x) f^{(i)},$$

where $n \geq 1$ and $P_i(x)$ is an infinitely differentiable function for $i \in \{0, \ldots, n-1\}$, is a finitely generated vector space. More precisely, for any scalars $c_0, \ldots, c_{n-1} \in \mathbb{R}$, there is a unique function f which satisfies the differential equation and the 'initial conditions' $f(0) = c_0, f'(0) = c_1, \ldots, f^{(n-1)}(0) = c_{n-1}$. If we let f_i denote the unique solution satisfying the initial conditions $c_i = 1$, $c_j = 0$ for $j \neq i$, then the set $\{f_0, \ldots, f_{n-1}\}$ is a spanning set for the space of all solutions.

Recall Example 9.2, where we introduced the differential equation $y'' + y = 0$. In the language of Theorem 9.15, we have $f_0 = \cos(x)$, $f_1 = \sin(x)$, and $\{\cos(x), \sin(x)\}$ is a spanning set for the whole space of solutions.

Theorem 9.15 is often called an 'existence/uniqueness' theorem since it asserts that solutions to a problem exist for any initial conditions and that specified initial conditions determine a solution uniquely. The $n = 2$ case agrees with physical intuition, as it is akin to a statement that the trajectory of a particle subject to a given potential energy landscape exists and is determined uniquely by the particle's initial position and velocity.

However, the theorem does not give a formula for what solutions will be; it simply asserts that they exist and in what sense they are unique. This may seem abstract, but consider the similar level of abstraction in the intermediate value theorem for calculus, which asserts that a continuous function $f : [a, b] \to \mathbb{R}$ must take every value between $f(a)$ and $f(b)$ but does not assert where.

We now recall the concept of a basis:

Definition 9.16. For a nonzero vector space V, a set $B \subset V$ is a *basis* for V if B is a linearly independent set and $\operatorname{span}(B) = V$.

Beware that the existence of a basis of a finitely generated vector space must be proven! However, we have already done the majority of the work to prove this.

Lemma 9.17. If V is a finitely generated vector space, V has a basis.

Proof. Let V be a nonzero finitely generated vector space, which by definition admits a finite spanning set $S = \{v_1, \ldots, v_n\}$. By Proposition 9.13, S admits a linearly independent subset with the same span, i.e. a basis $B \subset S$. □

We want a stronger understanding of *dimension* in finitely generated settings as the 'number of basis vectors'. In $\mathbb{R}^n$, this idea is discussed in Math 51 using the language of coordinates. But in more flavorful vector spaces like the set of solutions to a linear ordinary differential equation, 'dimension' gives us a new perspective.

We want to prove that the number of basis vectors, which we will refer to as the *dimension* of a vector space V, is independent of any particular basis. That way, we can talk about dimension in settings where there is no obvious choice of basis.

We need to show that for finitely generated vector spaces, all bases have the same size. To do this, we will first show that any spanning set is longer than any basis, again using our 'eliminating redundancy' approach and Lemma 9.12. Then, we will be able to show that any two bases must have the same size.

Proposition 9.18. Let V be a finite dimensional vector space, let $\{v_1, \ldots, v_n\}$ be a linearly independent set, and let $\{u_1, \ldots, u_m\}$ be a spanning set. Then $m \geq n$.

Proof. We will inductively replace the u_i with the v_i to create new spanning sets, then show that we run out of v_i to add before we run out of u_i to remove.

Observe that none of the v_i equals $\vec{0}$ since $\{v_1, \ldots, v_n\}$ is linearly independent. We claim that since $\{u_1, \ldots, u_m\}$ is a spanning set, $\{v_1, u_1, \ldots, u_m\}$ must be linearly dependent. This holds because $v_1 = v_1$ and $v_1 = a_1u_1 + \ldots + a_mu_m$ for some $a_1, \ldots, a_m$ since $\operatorname{span}(u_1, \ldots, u_m) =$

V, so $a_1u_1+\ldots+a_mu_m+(-1)v_1=\vec{0}$. Thus, by Lemma 9.12, we can remove some u_j without affecting the span of the new set. Let $T_1=\{v_1,u_1,\ldots,u_{j-1},u_{j+1},\ldots,u_m\}$.

Repeating the process for $2\leq k\leq n$, we can define T_k from T_{k-1} by adding v_k to the *front* of T_{k-1} and subtracting a 'redundant' vector from T_{k-1} by Lemma 9.12. We claim that the vector we remove will never be one of the v_i. This is because $\{v_1,\ldots,v_n\}$ is linearly independent, so by Proposition 9.10, none of the v_i is in the span of the earlier v_j. However, we require that the vector we remove from T_{k-1} at each step is in the span of all previous vectors in the collection, so in particular it cannot be one of the v_i. Thus, at each step of the process, we remove one of the u_i from T_{k-1}. Accordingly, there must be at least as many u_i as there are v_i, so $m\geq n$. $\square$

Proposition 9.19. Any two bases have the same size.

Proof. Let B_1 be a basis of size n and B_2 be a basis of size m. Since B_2 spans V and B_1 is linearly independent, by Proposition 9.18, $n\leq m$. Reversing the roles of B_1 and B_2 since each is both spanning *and* linearly independent, we also have $m\leq n$. Thus, $m=n$, so any two bases have the same size. $\square$

Observe how the above proof (and the previous results it depends on) do not rely on $\mathbb{R}^n$-specific methods like row reduction. This makes for a clean and general proof which provides intuition for how all that is necessary to establish dimension is a finitely generated vector space, regardless of whether or not it comes with a 'preferred basis'.

We are now ready to define the main concept of this section: the dimension of a finitely generated vector space. Proposition 9.19 then tells us that any basis of V (which necessarily exists by Lemma 9.17) has the same size as any other basis of V, so we make the following definition:

Definition 9.20. Let V be a finitely generated nonzero vector space. The *dimension* of V is the common size of any basis of V.

We began with an intuitive criterion for finitely generated vector spaces: there should exist some finite subset that is sufficient to describe all points in the vector space by linear combinations. We then formalized this idea with the language of linear independence, allowing us to prove several useful results: that nonzero, finitely generated vector spaces have bases, and that all bases of a particular such vector space have the same size. This confirms a lot of our intuition from $\mathbb{R}^n$ in a more broadly applicable setting, and we are left to examine what 'dimension' means as defined independently from any *preferred* basis. We will do so now to see how our new *definition* informs our *intuition* for a more general context than $\mathbb{R}^n$.

Example 9.21. Let $f'''+P(x)f''+Q(x)f'+R(x)f=0$ be a linear third-order homogeneous ordinary differential equation and V be the vector space of solutions. Then $\dim(V)=3$, and we need 3 initial conditions (e.g. $f(0)=c_1, f'(0)=c_2, f''(0)=c_3$) to specify a unique solution.[2]

[2]Note that this is a serious theorem in ordinary differential equations, and that even though a three dimensional vector space of solutions exists, there is no formula for what the solutions are in terms of P, Q, and R.

More generally, for infinitely differentiable functions $P_i(x)$, the solution space to a linear n-th order ordinary differential equation $f^{(n)} + \sum_{i=0}^{n-1} P_i(x) f^{(i)}(x) = 0$ is an n-dimensional vector space, and n initial conditions are required to specify a unique solution.[3]

Example 9.22. The set P_n of real-coefficient polynomials of degree less than or equal to n is an $(n+1)$-dimensional vector space: a basis is $\{1, x, x^2, \ldots, x^n\}$.[4] However, the set P of all real-coefficient polynomials is 'infinite-dimensional'. To see this, observe that for given any finite set $\{p_1, \ldots, p_m\}$, there must be a finite maximum degree d of the p_i. We can then choose any degree-$(d+1)$ polynomial q, and q cannot be in $\text{span}(p_1, \ldots, p_m)$. Thus, P does not admit a finite spanning set, so we say P is 'infinite-dimensional'.

9.3 Exercises

1. Prove that if V is any vector space equipped with a 'dot product' and 'length' $\|v\| = \sqrt{v \cdot v}$, all vectors $v, w \in V$ satisfy the parallelogram law (see Figure 1):

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

2. For continuous functions $f : [0, 1] \to \mathbb{R}$, we can talk about 'size' in different ways. This is helpful when breaking down functions into principal components, for instance in terms of trigonometric functions as introduced in the last section. We can define the size of a function to be its biggest deviation from 0:

$$\|f\| = \max\{|f(x)| : x \in [0, 1]\}.$$

   We can also define the size of a function to depend on its integral:

$$\|f\| = \int_0^1 |f(x)|\, dx.$$

   Lastly, we can define the size of a function to depend on the integral of its square:

$$\|f\| = \left(\int_0^1 f^2\, dx \right)^{1/2},$$

   similar to length in $\mathbb{R}^n$ by the intuition that integration is a continuous analog of summation. Using the previous exercise, prove that two of these three definitions *do not* arise from an inner product.

3. A *subspace* W of a vector space V is a subset of V which satisfies three properties:

   - $0 \in W$

[3] Again, this is a theorem which requires real work to prove, and there is no explicit formula for solutions in general.

[4] Note that these functions are linearly independent as *functions*, even though they may take the same values at particular points. This follows from the fact that a nonconstant polynomial of degree at most n, i.e. a polynomial $p(x) = c_0 + c_1 x + \ldots + c_n x^n$ with $c_j \neq 0$ for some $j > 0$, has at most n different roots, so on a domain with more than n points (e.g. $\mathbb{R}$), $p(x)$ cannot be identically zero.

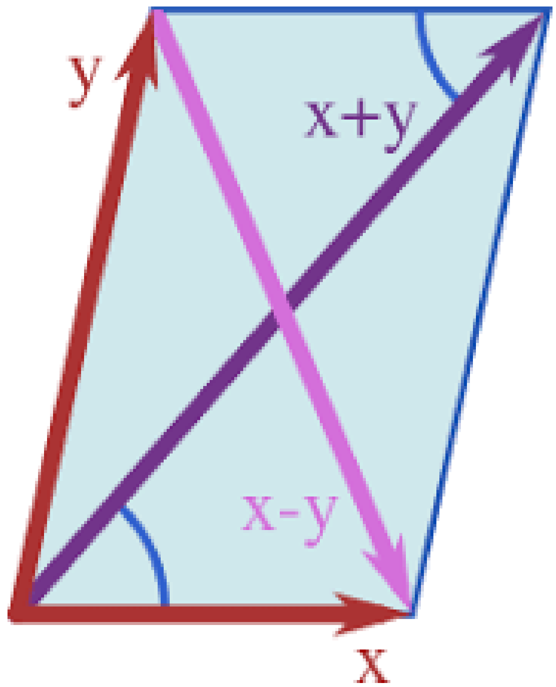

Figure 1: An illustration of the parallelogram law, which states that the sum of the squared lengths of the diagonals is twice the sum of the squared lengths of the sides.

- For all $x, y \in W$, $x + y \in W$.
- For all $x \in W$ and all $c \in \mathbb{R}$, $cx \in W$.

Prove whether or not the following sets are subspaces:

(a) The graph of the function $y = x$ as a subset of $\mathbb{R}^2$

(b) The graph of the function $y = x^2$ as a subset of $\mathbb{R}^2$

(c) The set of all polynomials of degree less than or equal to n as a subset of the vector space of all polynomials

(d) The intersection $W_1 \cap W_2$ of two subspaces $W_1, W_2 \subset V$

(e) The union $W_1 \cup W_2$ of two subspaces $W_1, W_2 \subset \mathbb{R}^3$

4. In the language of Theorem 9.15, consider the solutions $f_0, \dots, f_{n-1}$ to an n-th order homogeneous linear ordinary differential equation (i.e. the setting in which the theorem applies).

   (a) For $n = 2$, prove that f_0 and f_1 are linearly independent.

   (b) Generalize the above statement to arbitrary n to prove that $f_0, \dots, f_{n-1}$ are linearly independent.

5. Prove whether the following real vector spaces are finite dimensional or not. If finite dimensional, calculate the dimension.
   (a) $\mathbb{C}$
   (b) The set of $m \times n$ matrices
   (c) The set of functions $f : [0, 1] \to \mathbb{R}$ with vector space operations taken pointwise
   (d) The set of functions $f : \{A, B, C\} \to \mathbb{R}$ with vector space operations taken pointwise
6. Prove by induction that a subspace of a finite dimensional vector space is finite dimensional by the following steps. Suppose that $\dim(V) = n$ and $W \subset V$ is a nonzero subspace.
   (a) Observe that for any linearly independent set $S \subset W$ of size m, either $\text{span}(S) = W$ or there exists a vector $w \in W$ such that $w \notin \text{span}(S)$. If w exists, prove that the set $S \cup \{w\}$ is a linearly independent as a subset of V with $m + 1$ elements.
   (b) Argue that the above process of adding vectors in W to S to build larger linearly independent subsets of V must terminate after at most n steps.
7. Give an example of an 'infinite-dimensional' vector space V and a finite dimensional subspace $W \subset V$.
8. Give another example of a finite dimensional vector space other than the ones introduced in this section. What is its dimension (numerically), and what intuitive significance does its dimension have? (*Example: The vector space of solutions to* $f''' + P(x)f'' + Q(x)f' + R(x)f = 0$ *has dimension 3, which is the number of initial conditions needed to specify a unique solution.*)

9.4 Further Reading

Linear Algebra Done Right Chapter 2, Sheldon Axler	A more detailed overview of the concepts introduced in this section available via SpringerLink to Stanford undergraduates.
Ordinary Differential Equations, V.I. Arnold	An ordinary differential equations reference which contains a proof of Theorem 9.15. The proof is quite advanced but worthy of stopping to appreciate.