Linear Algebra
Week 6

G-07

31 X 2024

1 Vector Spaces

Vectors and matrices are all we talk about but we didn't have any proper definition so far. We also used the term 'space' intuitively and without formal understanding. In the following we are going to formally define vector spaces and their properties.

1.1 Definition and Examples

Before we see the definition of a vector space, let's have the definition of a vector:

A vector is an element of a **vector space**. Vector spaces are characterized by the presence of two operations on their elements: vector addition and scalar multiplication.

This definition might seem simple. It is. However, this definition tells us a lot about vectors. The reason for that is because we define vector spaces very carefully.

Definition 4.1 (Vector space). *A* vector space *is a triple* $(V, +, \cdot)$ *where* V *is a set (the vectors), and*

$$\begin{aligned} + &: V \times V \to V \quad \textit{is a function (vector addition),} \\ \cdot &: \mathbb{R} \times V \to V \quad \textit{is a function (scalar multiplication),} \end{aligned}$$

satisfying the following axioms of a vector space *for all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *and all* $\lambda, \mu \in \mathbb{R}$.

1.	$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$	*commutativity*
2.	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	*associativity*
3.	*There is a vector* $\mathbf{0}$ *such that* $\mathbf{v} + \mathbf{0} = \mathbf{v}$ *for all* $\mathbf{v}$	*zero vector*
4.	*There is a vector* $-\mathbf{v}$ *such that* $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$	*negative vector*
5.	$1 \cdot \mathbf{v} = \mathbf{v}$	*identity element*
6.	$(\lambda \cdot \mu)\mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v})$	*compatibility of* $\cdot$ *and* $\cdot$ *in* $\mathbb{R}$
7.	$\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$	*distributivity over* $+$
8.	$(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$	*distributivity over* $+$ *in* $\mathbb{R}$

An important note is that we are doing operation overloading here. That is, we assign two tasks to the symbols "+" and "$\cdot$" We use the same symbol "+"to represent vector-vector addition and the addition in real numbers (scalars) as well as we use the "$\cdot$" for scalar-vector multiplication and scalar-scalar multiplication. This should not cause any confusion so long as you know your variables. General advice: know the type of your variable is it a vector, is it a complex number, or a real number, or maybe a matrix...

The definition has only 8 axioms which you do not have to memorize. Out of these 8 axioms one can derive many features of vector spaces. As you will see in the 5th chapter of the lecture Discrete Mathematics, Algebra, many of the properties you take for granted are actually not among the axioms but they are implied by these axioms. Some of those properties that might seem trivial but actually need a proof are:

- There is only one zero vector $\mathbf{0}$
- There is only one additive inverse (Vector -$\mathbf{v}$ in the 4th axiom is unique for all $\mathbf{v}$)
- $0\mathbf{v} = \mathbf{0}$ *i.e.* 0 (the scalar) times a vector $\mathbf{v}$ equals the zero vector $\mathbf{0}$

However, since we are not as rigorous as in Discrete Maths, you can go on and assume these seemingly trivial properties hold in general in this course. Now we can wholeheartedly make the following observation:

Observation 4.2. $(\mathbb{R}^m, +, \cdot)$, with "+" as in Definition 1.1 (vector addition) and "$\cdot$" as in Definition 1.3 (scalar mutliplication), is a vector space.

Notice how we provided the addition and multiplication operations with the set of vectors. If we would say $\mathbb{R}^n$ is a vector space and stopped there, we would be wrong. The operations and the set of vectors define a vector space together.

You might wonder, what other vector spaces are there, except for $\mathbb{R}^n$? There are plenty and I would recommend taking a look at my website for an interesting list. But there are several important and fundamental vector spaces handled in the lecture and in the lecture notes. Two of them are polynomials and matrices.

Polynomials

Note: we are going to be talking about the real polynomials but you could have defined them with complex coefficients as well.

Definition 4.3 (Polynomial)**.** *A polynomial* $\mathbf{p}$ *is a sum of the form*

$$\mathbf{p} = \sum_{i=0}^{m} p_i x^i,$$

for some $m \in \mathbb{N}$*. Here* x *is a variable, and the numbers* $p_0, p_1, \ldots, p_m \in \mathbb{R}$ *are the* coefficients *of* $\mathbf{p}$*. The largest* i *such that* $p_i \neq 0$ *is the* degree *of* $\mathbf{p}$*. If all* p_i *are* 0*, we have the* zero polynomial $\mathbf{0} = 0$ *whose degree we define to be* -1*.*

Theorem 4.4. *Let* $\mathbb{R}[x]$ *be the set of polynomials in one variable* x*. Given polynomials* $\mathbf{p} = \sum_{i=0}^{m} p_i x^i$ *and* $\mathbf{q} = \sum_{i=0}^{n} q_i x^i$*, we define* $\mathbf{p} + \mathbf{q}$ *to be the polynomial*

$$\mathbf{p} + \mathbf{q} = \sum_{i=0}^{\max(m,n)} (p_i + q_i) x^i,$$

where we set $p_i = 0$ *for* $i > m$ *and* $q_i = 0$ *for* $i > n$*. For a scalar* $\lambda \in \mathbb{R}$*, we further define* $\lambda\mathbf{p}$ *as the polynomial*

$$\lambda\mathbf{p} = \sum_{i=0}^{m} (\lambda p_i) x^i.$$

Then $(\mathbb{R}[x], +, \cdot)$ *is a vector space.*

Matrices

Theorem 4.5. Let $\mathbb{R}^{m\times n}$ be the set of $m \times n$ matrices, with addition $A+B$ and scalar multiplication λA defined in the usual way, see Definition 2.2. from the lecture notes. Then $(\mathbb{R}^{m\times n}, +, \cdot)$ is a vector space.

After introducing this formality about vector spaces and understanding that a vector space is a triple $(V, +, \cdot)$ and not just a set V, we are still going to write V for the vector space, with the understanding that vector addition and scalar multiplication are clear from the context. Under the hood, be aware that this is an abuse of notation.

2 Subspaces

Since we now have a definition for vector spaces, it is only natural to ask for smaller components of it. For example we say that matrices in $\mathbb{R}^{m\times n}$ are a vector space (see the abuse of notation?). But it seems like symmetric matrices in $\mathbb{R}^{n\times n}$ also form a vector space (I did it again). If you add two symmetric matrices the sum is symmetric. Multiplying with a scalar also results in a symmetric matrix. This resembles be a vector space ***"inside"*** the vector space of $\mathbb{R}^{m\times n}$ matrices. We call such spaces the subspaces of a vector space:

Definition 4.8 (Subspace)**.** *Let V be a vector space. A nonempty subset $U \subseteq V$ is called a* subspace *of V if the following two* axioms of a subspace *are true for all $\mathbf{v}, \mathbf{w} \in U$ and all $\lambda \in \mathbb{R}$.*

(i) $\mathbf{v} + \mathbf{w} \in U$*;*

(ii) $\lambda\mathbf{v} \in U$*.*

A very important keyword is ***nonempty***. At some point in your life, you will be asked to prove that some vector space is a subspace of some other vector space. Then you must prove **3** aspects:

1. The subspace is not empty. Usually the easiest way is to show $\mathbf{0}$ is an element of the subspace because this is always the case, see lemma below.

2. The subset is closed under addition: Adding two vectors can't possibly get us out of the subspace.

3. The subset is closed under scalar multiplication: Multiplying a vector with a scalar cannot get us out of the vector space.

There are two important lemmas about the subspaces. Let's look at them and to some examples of subspaces of the spaces we mentioned above.

Lemma 4.9. Let $U \subseteq V$ be a subspace of a vector space V. Then $\mathbf{0} \in V$.

Proof Take any $\mathbf{u} \in U$ (U is nonempty), By subspace axiom *(ii)*, $0\mathbf{u} = \mathbf{0} \in U$.

□

Note that we use the fact $0\mathbf{u} = \mathbf{0}$ here which was listed under seemingly trivial facts before. As a result of this lemma, if you are drawing any subspace, the shape that represents it must go through $\mathbf{0}$.

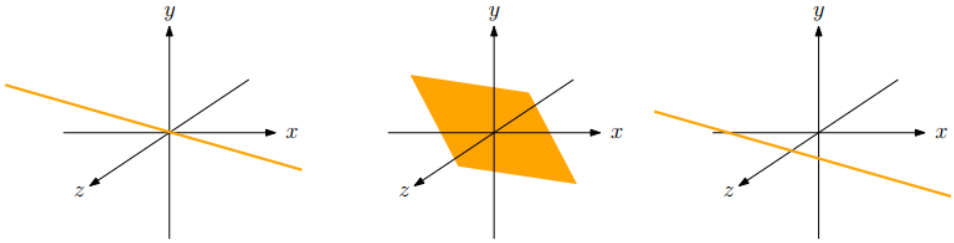

Figure 4.1: Subspaces of $\mathbb{R}^3$: a line through $\mathbf{0}$ (all scalar multiples of one vector); a plane through $\mathbf{0}$ (all linear combinations of two linearly independent vectors); the right subset is not a subspace, since it misses $\mathbf{0}$.

Lemma 4.12. Let V be a vector space and U a subspace. Then U is also a vector space (with the same "+" and "$\cdot$" as V).

Proof. Formally, to turn "+" and "$\cdot$" into functions that work for U, we have to restrict their domains to $U \times U$ and $\mathbb{R} \times U$, respectively. The subspace axioms (i) and (ii) then also restrict their ranges to U.

Next, we need to check the 8 axioms. All but axiom 4 are true for all vectors in V, since V is a vector space; in particular, they hold for all vectors in U, so there is nothing to check. In Case of axiom 3, we are also using that $\mathbf{0} \in U$ (Lemma 4.9). What remains is axiom 4: we need to make sure that for all $\mathbf{u} \in U$, $-\mathbf{u}$ is actually in U; so far we only know that it is in V. But this holds, since $(-1)\mathbf{u} \in U$ by subspace axiom (ii), and "obviously" $(-1)\mathbf{u} = -\mathbf{u}$. If you are up for it, you can prove the obvious, otherwise, you can safely believe it. □

Examples of Subspaces

1 - The Column Space of a Matrix

Lemma 4.11. Let A be an $m \times n$ matrix. Then $\mathbf{C}(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$ is a subspace of $\mathbb{R}^m$.

2 - Polynomials without a Constant Term

Polynomials without a constant term: $p = \sum_{i=0}^{m} p_i x^i$ where $p_0 = 0$ form a subspace of the vector space of polynomials $\mathbb{R}[x]$.

3 - Polynomials of Degree at Most n

Polynomials with degree at most n for a fixed $n \in \mathbb{N}$: $p = \sum_{i=0}^{n} p_i x^i$ (any p_i can be zero) form a subspace of the vector space of polynomials $\mathbb{R}[x]$. (in lecture notes there is the example of quadratic polynomials for $n = 2$)

4 - Symmetric Matrices

Symmetric matrices form a subspace of $\mathbb{R}^{m \times n}$ for a fixed n.

5 - Matrices with Trace[1] 0

Square matrices with trace 0, as in two dimensional example $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, a+d = 0$ form a subspace of $\mathbb{R}^{m \times n}$.

If you want to practice you can proof yourself that the mentioned subspaces are actually subspaces of the given vector spaces (with the same "+" and "$\cdot$"). You can also find most of the proofs in the lecture notes.

3 Bases and Dimension

Now that we defined vector spaces, we can talk about the bases and dimension of a vector space. A natural way to think of a basis is to think of it as the essential part of a vector space. Once we know a basis, then we can generate every vector in a vector space. The dimension tells us how big the basis is.

[1] Trace of a square matrix is the sum of the diagonal elements of a matrix.

3.1 Bases

Since we work with general vector spaces, it is more practical to work with *sets of vectors* instead of *sequences of vectors* as we did so far. The formality and definitions of linear combination, independence, and span are slightly different when you do it over *sets of vectors*. The intuition stays the same but it is still necessary to define these concepts for sets of vectors.

An important keyword in the following definition is the word "*finite*". We do not allow infinite linear combinations because then things get out of hand. An infinite sum might throw you out of the vector space which we do not want to happen. For examples see the lecture notes.

Definition 4.13 (Linear combination of a set of vectors)**.** *Let V be a vector space, $G \subseteq V$ a (possibly infinite) subset of vectors. A* linear combination *of G is a sum of the form*

$$\sum_{\mathbf{v} \in F} \lambda_{\mathbf{v}} \mathbf{v},$$

where $F \subseteq G$ is a finite subset of G and $\lambda_{\mathbf{v}} \in \mathbb{R}$ for all $\mathbf{v} \in F$.

With linear combinations as in Definition 4.13, we can now define span and linear independence in the usual way.

Definition 4.15 (Span and linear independence of sets of vectors)**.** *Let V be a vector space, $G \subseteq V$ a (possibly infinite) subset of vectors.*

The span *of G, written as $\mathbf{Span}(G)$, is the set of all linear combinations of G. The set G is called* linearly independent *if no vector $\mathbf{v} \in G$ is a linear combination of $G \setminus \{\mathbf{v}\}$.*

An important fact about vector spaces is that the linear combination of vectors from vector spaces keep us in the vector space, *i.e.* all linear combinations are again in the vector space. As intuitive as this sounds, it is not trivial and needs a proof.

Lemma 4.14. *Let V be a vector space, $G \subseteq V$. Every linear combination of G is again in V.*

Proof. Let $\sum_{\mathbf{v} \in F} \lambda_{\mathbf{v}} \mathbf{v}$ be a linear combination, $|F| = n$. Enumerating the elements of F in arbitrary order $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, we can write the combination as $\sum_{j=1}^{n} \lambda_j \mathbf{v}_j$.

Since V is a vector space, we have $\mathbf{w}_j := \lambda_j \mathbf{v}_j \in V$ for all j, by definition of the scalar multiplication (function $\cdot : \mathbb{R} \times V \to V$). By definition of vector addition (function $+ : V \times V \to V$), we also have $\mathbf{w}_1 + \mathbf{w}_2 \in V$. Applying this again yields $(\mathbf{w}_1 + \mathbf{w}_2) + \mathbf{w}_3 \in V$, and so on, until we get the desired conclusion $\mathbf{w}_1 + \mathbf{w}_2 + \cdots + \mathbf{w}_n \in V$. (Under the hood, this is a proof by induction, and it uses the "obvious" fact that brackets can be omitted in writing down a sum of vectors). □

Now we have the tools to formally define what a basis is.

Definition 4.16 (Basis). Let V be a vector space. A subset $B \subseteq V$ of vectors is called a basis of V if B is linearly independent and $Span(B) = V$.

One example is the set of unit vectors for $\mathbb{R}^n$, also called *the canonical basis*. $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$ for $\mathbb{R}^2$. In two dimensions it might be clear that the unit vectors are linearly independent. In higher dimensions we can argue with the *private nonzero argument*. every standard unit vector has a nonzero entry (a 1-entry, actually) at a coordinate where all other standard unit vectors have 0-entries. We call such an entry a private nonzero. A vector with a private nonzero cannot be a linear combination of the other ones, and if every vector has a private nonzero, the vectors are linearly independent.

Another example is the independent columns of a matrix A and its column space $\mathbf{C}(A)$:

Lemma 4.17. *Let A be an $m \times n$ matrix. The set of independent columns of A (Definition 2.9) is a basis of the column space $\mathbf{C}(A)$.*

Proof. $\mathbf{C}(A)$ is a subspace by Lemma 4.11. The independent columns are in the column space and linearly independent: by definition, no independent column is a linear combination of the previous columns, and this means that the independent columns are in fact linearly independent; see Corollary 1.20. Furthermore, the independent columns span the column space, as we have shown in Lemma 2.10. □

Question 6.1 Try to find bases for the vector spaces mentioned above as examples of subspaces. (The answers are in the lecture notes.)

What is the basis of the vector space $\{\mathbf{0}\}$? It is the empty set! Indeed, this is linearly independent by Definition 4.15: in the empty set, no vector is a linear combination of the others. And $\mathbf{Span}(\emptyset) = \mathbf{0}$ because the empty sum yields $\mathbf{0}$.

There are typically many bases. This is why we mention *a* basis of a vector space and not *the* basis of a vector space. In the following there is a crucial observation:

Observation 4.18. *Every set $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of m linearly independent vectors is a basis of $\mathbb{R}^m$.*

Proof. B is linearly independent, so we only need to show that $\mathbf{Span}(B) = \mathbb{R}^m$, meaning that every vector $\mathbf{v} \in \mathbb{R}^m$ is a linear combination of B. For this, let A be the $m \times m$ matrix with (linearly independent) columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. By Theorem 3.11, $A\mathbf{x} = \mathbf{v}$ has a unique solution $\mathbf{x}$, and

$$\mathbf{v} = \underbrace{\sum_{j=1}^{m} x_j \mathbf{v}_j}_{A\mathbf{x}}$$

is the desired linear combination of B. □

There might be infinitely many bases but they all have the same number of elements. All bases of a vector space consist of the same number of vectors. But to prove this, we need another tool, called the *Steinitz Exchange Lemma.* Here I provide it without a proof or an explanation because the proof and explanation in the lecture notes is already really good. Read it once, read it twice, and then a couple more times. Sometimes understanding a proof takes time.

The Steinitz Exchange Lemma

Lemma 4.19 (Steinitz exchange lemma)**.** *Let V be a vector space, $F \subseteq V$ a finite set of linearly independent vectors, and $G \subseteq V$ a finite set of vectors with $\mathbf{Span}(G) = V$. Then the following two statements hold.*

(i) $|F| \leq |G|$.

(ii) There exists a subset $E \subseteq G$ of size $|G| - |F|$ such that $\mathbf{Span}(F \cup E) = V$.

The Steinitz Exchange Lemma enables us to proof that all bases of a vector space has the same size.

Theorem 4.20. *Let V be a vector space and $B, B' \subseteq V$ two finite bases of V. Then $|B| = |B'|$.*

Proof. By Definition 4.16 of a basis, B and B' are linearly independent, and $\mathbf{Span}(B) = \mathbf{Span}(B') = V$. Then, statement (i) of the Steinitz exchange lemma with $F = B, G = B'$ yields $|B| \leq |B'|$; with $F = B', G = B$, we get $|B'| \leq |B|$. □

And yes, every vector space has a basis, even the infinite case. Here is the proof for the finite case from the lecture notes:

Definition 4.21 (Finitely generated vector space). *A vector space V is called* finitely generated *if there exists a finite subset $G \subseteq V$ with* $\mathbf{Span}(G) = V$.

For example, $\mathbb{R}^m$ is finitely generated (by $G = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$) but $\mathbb{R}[x]$, the vector space of polynomials, is not.

Theorem 4.22. *Let V be a finitely generated vector space, and let $G \subseteq V$ be a finite subset with $\mathbf{Span}(G) = V$. Then V has a basis $B \subseteq G$.*

Proof. This is what we call an "algorithmic proof'." It constructs B by an algorithm. Here is how it goes.

If G is linearly independent, $B = G$ is a basis by Definition 4.16. Otherwise, there is a "line 1" vector $\mathbf{v} \in G$ that is a linear combination of the other vectors, so we have $\mathbf{Span}(G \setminus \{\mathbf{v}\}) = \mathbf{Span}(G) = V$ via Lemma 1.23. Then we replace G with $G \setminus \{\mathbf{v}\}$ (which still spans V) and go back to line 1. As G gets smaller in every step, this must eventually stop and produce a basis. □

3.2 Dimensions

Now that we showed all bases have the same number of vectors, we can define the dimension of a vector space as the number of vectors in a basis. *E.g.* $dim(\mathbb{R}^n) = n$ Three subspaces of $\mathbb{R}^3$ are

Definition 4.23 (Dimension). Let V be a finitely generated vector space. Then $dim(V)$, the dimension of V , is the size of any basis B of V .

And if the dimension d is known, we only have to check the one of the basis axioms in definition 4.16. In other words:

Lemma 4.24. *Let V be a vector space with $\dim(V) = d$.*

(i) Let $F \subseteq V$ be a set of d linearly independent vectors. Then F is a basis of V.

(ii) Let $G \subseteq V$ be a set of d vectors with $\mathbf{Span}(G) = V$. Then G is a basis of V.

Proof. For (i), let G be a basis of V. Since $\mathbf{Span}(G) = V$, the Steinitz exchange Lemma 4.19 applies with F and G, but since $|F| = |G| = d$, the set E in part (ii) can only be the empty set. Hence, $\mathbf{Span}(F) = \mathbf{Span}(F \cup E) = V$, and F is a basis according to Definition 4.16.

For (ii), we use Theorem 4.22 which guarantees a basis $B \subseteq G$ of size d. But since $|B| = |G| = d$, we must have $B = G$, so G itself is a basis. □

4 Hints

1. In Class Exercise: See pdf on website for solution.

2. Bonus, no hints!

3. One direction is relatively easy to show. For the other direction assume that $W \nsubseteq U$. Choose two vectors cleverly. You know that their sum must be in $W \cup U$. What does this imply?

4. Seeing a concrete example in lower dimensions might help. You should "guess" a basis by generalizing the idea in lower dimensions. Then you should prove that your chosen basis is really a basis, i.e. the vectors are linearly independent and they span the whole space.

5. For **a)** don't forget you have to show 3 things to prove that a given set forms a subspace. For **c)** Think about the functions $f(x)$ and $f(-x)$. If you take the average, flipping the sign of x should not have an effect. And if you take the middle point of $f(x)$ and $f(-x)$ (*i.e.* their difference/2) then flipping the sign of x also flips the sign of the middle point.

6. Are $\mathbf{p}$,$\mathbf{q}$,$\mathbf{r}$ linearly independent?

7. No hints.

mkilic