Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-11-18

Darij Grinberg

December 5, 2019

1. Vector spaces

Let us generalize what we have done so far. We have worked with matrices (including column vectors and row vectors). We have used "scalars" (i.e., numbers that we could put into a matrix) and "vectors" (i.e., things formed out of scalars, but more importantly, things that could be added to each other and scaled by scalars). We will now extend both of these notions.

1.1. Fields

Let us first generalize the notion of a "scalar".

We have mostly been using real numbers as scalars, but later said that complex numbers also work. In practice, we have been using rational numbers as much as possible, and algebraic numbers[1] as a last resort.

[1] What are algebraic numbers? **Algebraic numbers** are complex numbers that can be expressed as roots of nonzero polynomials with rational coefficients. For example,

- $\sqrt{2}$ is algebraic, since it is a root of $t^2 - 2$;
- $\sqrt{3}$ is algebraic, since it is a root of $t^2 - 3$;
- $\sqrt[3]{2}$ is algebraic, since it is a root of $t^3 - 2$;
- $\sqrt{2} + \sqrt{3}$ is algebraic, since it is a root of $t^4 - 10t^2 + 1$;
- the imaginary unit i is algebraic, since it is a root of $t^2 + 1$;
- etc.

The advantage of algebraic numbers (compared to arbitrary complex numbers) is that you can calculate with them precisely, whereas true complex numbers can only be approximated. This is why (almost) all of our examples have used algebraic numbers only. Computers don't understand real numbers either, so they also use algebraic numbers or some kind of approximation.

We also have secretly used matrices whose entries are polynomials. Indeed, our matrix $A - tI_n$ (which we used when defining the characteristic polynomial) was such a matrix. So we (silently) stretched our notion of "scalar" so that it also encompassed polynomials. What else do we want it to encompass?

The short answer is "too much to list". We don't want to restrict ourselves to a list of specific things that we allow as scalars. Instead, we will formulate some requirements that scalars should satisfy, and we shall allow any kind of objects that satisfy these requirements to be considered "scalars". Clearly, some requirements are necessary; for example, there has to be a way to add and multiply our "scalars", because we want to be able to add and multiply matrices.

Our new, generalized meaning of "scalar" (or "number", which – for our purposes – is a synonym) is going to depend on the context. You can come up with your custom type of scalars, as long as you can guarantee that they satisfy our requirements.

So what exactly do we want to require from our scalars in order to be able to put them into matrices (and have these matrices behave well)? We want them to belong to a **commutative ring**:

Definition 1.1.1. A **commutative ring** means a set $\mathbb{K}$ equipped with the following additional data:

- a binary operation called "+" (that is, a function that takes two elements $a \in \mathbb{K}$ and $b \in \mathbb{K}$ as inputs, and outputs a new element of $\mathbb{K}$ which is denoted by $a + b$);
- a binary operation called "$\cdot$" (that is, a function that takes two elements $a \in \mathbb{K}$ and $b \in \mathbb{K}$ as inputs, and outputs a new element of $\mathbb{K}$ which is denoted by $a \cdot b$ or ab);
- an element of $\mathbb{K}$ called "0" (this may or may not be the integer 0);
- an element of $\mathbb{K}$ called "1" (this may or may not be the integer 1)

satisfying the following axioms (= requirements):

- **Commutativity of addition:** We have $a + b = b + a$ for all $a, b \in \mathbb{K}$.
- **Commutativity of multiplication:** We have $ab = ba$ for all $a, b \in \mathbb{K}$.
- **Associativity of addition:** We have $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{K}$.
- **Associativity of multiplication:** We have $a(bc) = (ab)c$ for all $a, b, c \in \mathbb{K}$.
- **Neutrality of 0:** We have $a + 0 = 0 + a = a$ for all $a \in \mathbb{K}$.

- **Existence of additive inverses:** For each $a \in \mathbb{K}$, there exists an element $a' \in \mathbb{K}$ such that $a + a' = a' + a = 0$. This a' is commonly denoted by $-a$, and called the **negative** (or the **additive inverse**) of a. (It is easy to check that it is unique.)

- **Unitality (a.k.a. neutrality of 1):** We have $a \cdot 1 = 1 \cdot a = a$ for all $a \in \mathbb{K}$.

- **Annihilation:** We have $a \cdot 0 = 0 \cdot a = 0$ for all $a \in \mathbb{K}$.

- **Distributivity:** We have $a(b+c) = ab + ac$ and $(a+b)c = ac + bc$ for all $a, b, c \in \mathbb{K}$.

If $\mathbb{K}$ (equipped with $+$, $\cdot$, 0 and 1) is a commutative ring, then we refer to the operation $+$ as the **addition** of $\mathbb{K}$; we refer to the operation $\cdot$ as the **multiplication** of $\mathbb{K}$; we refer to the element 0 as the **zero** of $\mathbb{K}$; we refer to the element 1 as the **unity** (or, somewhat confusingly, **one**) of $\mathbb{K}$.

This definition was a mouthful, but its intention is rather simple: It defines a commutative ring to be a set equipped with two operations which **behave like** addition and mulitplication of numbers, and two elements which **behave like** the number 0 and the number 1. They don't **have** to be addition and multiplication of numbers, but the axioms we imposed force them to behave similarly enough that we can try to do with them whatever we can do with numbers.

As a consequence, if $\mathbb{K}$ is a commutative ring, then matrices filled with elements of $\mathbb{K}$ will behave like matrices filled with numbers, at least as far as very basic properties are concerned.

Here are some examples of commutative rings:

- Each of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ (endowed with the usual addition, the usual multiplication, the usual 0 and the usual 1) is a commutative ring.

- The set of all polynomials in a single variable x with real coefficients is also a commutative ring.

- The set of all functions from $\mathbb{R}$ to $\mathbb{R}$ (equipped with pointwise addition[2],

[2]"Pointwise addition" means that the sum $f+g$ of two functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined by

$$(f+g)(x) = f(x) + g(x) \qquad \text{for all } x \in \mathbb{R}.$$

In other words, each value of the function $f+g$ is the sum of the corresponding values of f and g.

pointwise multiplication[3], the constant-0 function[4] and the constant-1 function[5]) is a commutative ring.

- The set $\mathbb{R}^{2\times 2}$ of 2×2-matrices with real entries (equipped with matrix addition, matrix multiplication, the zero matrix $0_{2\times 2}$ and the identity matrix I_2) is **not** a commutative ring. In fact, it satisfies all of the axioms **except for** commutativity of multiplication (which it fails to satisfy because in general, $AB \neq BA$ for matrices A and B).

- If you have seen "integers modulo n": they also form a commutative ring (for any given $n \in \mathbb{Z}$).

- Consider the two-element set $\{\mathbf{0}, \mathbf{1}\}$ (consisting of two symbols $\mathbf{0}$ and $\mathbf{1}$, not to be mistaken for the numbers 0 and 1). Equip this set with the following operations and elements:

  – Addition (i.e., the binary operation "$+$") is given by

  $$\mathbf{0}+\mathbf{0}=\mathbf{0}, \qquad \mathbf{0}+\mathbf{1}=\mathbf{1}, \qquad \mathbf{1}+\mathbf{0}=\mathbf{1}, \qquad \mathbf{1}+\mathbf{1}=\mathbf{0}.$$

  (This is almost like the usual addition of numbers, except that we are setting $\mathbf{1}+\mathbf{1}=\mathbf{0}$ whereas usual numbers satisfy $1+1=2\neq 0$.)

  – Multiplication (i.e., the binary operation "$\cdot$") is given by

  $$\mathbf{0}\cdot\mathbf{0}=\mathbf{0}, \qquad \mathbf{0}\cdot\mathbf{1}=\mathbf{0}, \qquad \mathbf{1}\cdot\mathbf{0}=\mathbf{0}, \qquad \mathbf{1}\cdot\mathbf{1}=\mathbf{1}.$$

  (Unlike the addition, this multiplication is indeed the usual multiplication of numbers, with 0 and 1 replaced by $\mathbf{0}$ and $\mathbf{1}$.)

  – The elements $\mathbf{0}$ and $\mathbf{1}$ are playing the roles of 0 and 1.

  It turns out that this set $\{\mathbf{0}, \mathbf{1}\}$ (equipped with these operations and elements) is a commutative ring. It just has two elements!

I should say that, despite me using the symbols "$+$" and "$\cdot$" for the two binary operations in Definition 1.1.1 and calling them "addition" and "multiplication",

[3]"Pointwise multiplication" means that the product $f \cdot g$ of two functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined by

$$(f\cdot g)(x) = f(x)\cdot g(x) \qquad \text{for all } x \in \mathbb{R}.$$

In other words, each value of the function $f \cdot g$ is the product of the corresponding values of f and g.

[4]i.e., the function

$$\mathbb{R} \to \mathbb{R}, \qquad x \mapsto 0$$

[5]i.e., the function

$$\mathbb{R} \to \mathbb{R}, \qquad x \mapsto 1$$

they don't have to have any relation to the addition and the multiplication of numbers in the usual sense (say, real or complex numbers). This is partly because the elements of $\mathbb{K}$ are not necessarily numbers in the usual sense; but even if they are, the operations "$+$" and "$\cdot$" may be different from the usual addition and multiplication. In particular, there are many "alternative additions and alternative multiplications" that we could equip $\mathbb{Z}$ with to obtain commutative rings. We shall avoid such ambiguities, however. If we really need to define an "alternative addition" on a set that already has an addition defined on it, then we will instead create a "fresh" copy of the set. (This is why we used two symbols $\mathbf{0}$ and $\mathbf{1}$ instead of the numbers 0 and 1 in our last example above.)

What can we do if we have a commutative ring?

If $\mathbb{K}$ is a commutative ring, then some of the classical concepts of algebra make sense in $\mathbb{K}$:

- We can add and multiply elements of $\mathbb{K}$, since $\mathbb{K}$ comes with an addition and a multiplication (by Definition 1.1.1).
- We can subtract elements of $\mathbb{K}$: Indeed, if $a \in \mathbb{K}$ and $b \in \mathbb{K}$, then we define $a - b$ to be $a + \underbrace{(-b)}_{\text{additive inverse of } b}$.
- Finite sums (like $a_1 + a_2 + \cdots + a_k$ or $\sum\limits_{i=1}^{k} a_i$ or $\sum\limits_{i \in I} a_i$) of elements of $\mathbb{K}$ make sense (and are defined in the same way as for numbers). So do finite products (like $a_1 a_2 \cdots a_k$ or $\prod\limits_{i=1}^{k} a_i$ or $\prod\limits_{i \in I} a_i$). The empty sum is defined to be 0 (the zero of $\mathbb{K}$); the empty product is defined to be 1 (the unity of $\mathbb{K}$).
- We can take the power a^k for all $a \in \mathbb{K}$ and $k \in \mathbb{N}$; indeed, this power is defined by $a^k = \underbrace{aa \cdots a}_{k \text{ times}}$. (Thus, $a^0 = \underbrace{aa \cdots a}_{0 \text{ times}} = (\text{empty product}) = 1$.)
- If $n \in \mathbb{Z}$ and $a \in \mathbb{K}$, then we can define an element $na \in \mathbb{K}$. This element na is defined by
$$na = \underbrace{a + a + \cdots + a}_{n \text{ times}} \qquad \text{if } n \geq 0$$
and
$$na = -\left(\left(-n\right) a\right) \qquad \text{if } n < 0.$$
- We **cannot** (in general) divide elements of $\mathbb{K}$. That is, quotients like a/b (with $a, b \in \mathbb{K}$) may fail to exist, even if b is nonzero.
- We also **cannot** take powers a^b with both $a, b \in \mathbb{K}$. (Even for complex numbers, there is no good way to do this.)

Now, fix a commutative ring $\mathbb{K}$, and replace "real numbers", "numbers" and "scalars" by "elements of $\mathbb{K}$" throughout the linear algebra that we have seen. Which of our results remain true?

Matrices filled with elements of $\mathbb{K}$ will be called **matrices over** $\mathbb{K}$. Their basic properties are still true:

- Rules like $A(BC) = (AB)C$ and $A(B+C) = AB + AC$ and $AI_n = A$ (for matrices of the appropriate sizes) remain true.
- Concepts like "diagonal", "upper-triangular", "transpose" etc. remain valid and preserve most of their properties (for example, $(AB)^T = B^T A^T$ still holds).
- Determinants still work, since the formula

$$\det A = \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$$

  makes sense whenever the entries of A belong to a commutative ring. Basic properties of determinants still hold. (For example, $\det(AB) = \det A \cdot \det B$.)

However, commutative rings are not enough for Gaussian elimination. Indeed, some of the steps in Gaussian elimination require dividing by some entries of the matrix[6]. But if $\mathbb{K}$ is merely a commutative ring, not every nonzero element of $\mathbb{K}$ can be divided by. This has consequences: Gaussian elimination no longer works (in general), and many facts that follow from Gaussian elimination (such as the Inverse Matrix Theorem) no longer hold.

We can see this on an example with $\mathbb{K} = \mathbb{Z}$. If we perform Gaussian elimination on the matrix $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ (which is a matrix over $\mathbb{Z}$), we obtain

$$\begin{pmatrix} \boxed{2} & 1 \\ 4 & 3 \end{pmatrix} \xrightarrow{\text{scale row 1 by } 1/2} \begin{pmatrix} \boxed{1} & 1/2 \\ 4 & 3 \end{pmatrix}$$

in the first step, which already contains a non-integer entry. So we are missing a notion of division. Commutative rings that allow division are called **fields**:

Definition 1.1.2. A commutative ring $\mathbb{K}$ is called a **field** if it satisfies the following two axioms:

- **Nontriviality:** We have $0 \neq 1$. (Here, the "0" and the "1" stand not for the two integers 0 and 1, but rather for the two elements of $\mathbb{K}$ that have been designated "0" and "1" in the definition of a commutative ring. In a commutative ring $\mathbb{K}$, they can be equal, but for a field we are requiring that they are not.)

[6] Namely, in order to turn a pivot entry λ into 1, we need to scale the corresponding row by $1/\lambda$. Thus, we need $1/\lambda$ to be well-defined, i.e., we need to be able to divide by λ.

- **Existence of multiplicative inverses:** For each $a \in \mathbb{K}$, we have **either** $a = 0$, **or** there is a $b \in \mathbb{K}$ satisfying $ab = ba = 1$.

The element b in the latter axiom is unique (this is easy to see), and it is called the **inverse** of a, and is denoted by a^{-1}.

If u and v are two elements of a field $\mathbb{K}$ such that $v \neq 0$, then $\dfrac{u}{v}$ is defined to be the product uv^{-1}. Thus, any two elements of $\mathbb{K}$ can be divided by each other as long as the denominator is $\neq 0$.

For example, the commutative rings $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ and our two-element ring $\{\mathbf{0}, \mathbf{1}\}$ are fields, but $\mathbb{Z}$ is not.

Many results about matrices are based on the ability to divide by nonzero numbers; these results cannot be directly generalized to matrices over a commutative ring $\mathbb{K}$, but they can be generalized to matrices over a field $\mathbb{K}$. In particular, if $\mathbb{K}$ is a field, then

- Gaussian elimination works for matrices over $\mathbb{K}$ (so any matrix has a RREF);
- the Inverse Matrix Theorem holds;
- the notions of "linear combination", "linear (in)dependence" and "spanning" (for column vectors or row vectors over $\mathbb{K}$) can be defined (in the same way as we defined them for vectors over $\mathbb{R}$), and enjoy the same properties;
- basic properties of eigenvectors hold (e.g., eigenvectors corresponding to distinct eigenvalues are linearly independent).

There do exist some properties of matrices over $\mathbb{R}$ that cannot be generalized to matrices over an arbitrary field $\mathbb{K}$. But we have not encountered such properties so far, and will not encounter them in this class.[7]

1.2. Vector spaces

1.2.1. Definition

So we have generalized our concept of "scalars": Instead of requiring them to be real or complex numbers, we simply allow them to come from any field $\mathbb{K}$.

Now, let us generalize "vectors". We know row vectors and column vectors. What else behaves like these kinds of vectors?

We have seen (in homework set #2) that the basic concepts we defined for column vectors (such as linear combinations, relations and linear independence) can be used not only for column vectors (or row vectors), but also for matrices and for polynomials. This is because these concepts are defined entirely in terms of

[7] Such properties are often found around the concepts of **symmetric** and **orthogonal** matrices.

addition and scaling and the zero vector. We shall now agree to call any kind of objects that can be added and scaled (and that support a notion of zero) **vectors**. More precisely:

Definition 1.2.1. Let $\mathbb{K}$ be a field. The elements of $\mathbb{K}$ will be called **scalars**.

A **vector space (over $\mathbb{K}$)** means a set V equipped with two binary operations:

- a binary operation called "$+$", which takes as input two elements $v, w \in V$ and yields an element of V called $v + w$;
- a binary operation called "$\cdot$", which takes as input a scalar $\lambda \in \mathbb{K}$ (that is, an element of $\mathbb{K}$) and an element $v \in V$, and yields an element of V called $\lambda \cdot v$ or λv

as well as a chosen element of V called "$\overrightarrow{0}$", satisfying the following axioms:

- **Commutativity of addition:** We have $v + w = w + v$ for all $v, w \in V$.
- **Associativity of addition:** We have $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$.
- **Neutrality of $\overrightarrow{0}$:** We have $v + \overrightarrow{0} = \overrightarrow{0} + v = v$ for all $v \in V$.
- **Right distributivity:** We have $(\lambda + \mu) v = \lambda v + \mu v$ for all $\lambda, \mu \in \mathbb{K}$ and $v \in V$.
- **Left distributivity:** We have $\lambda (v + w) = \lambda v + \lambda w$ for all $\lambda \in \mathbb{K}$ and $v, w \in V$.
- **Associativity:** We have $(\lambda \mu) v = \lambda (\mu v)$ for all $\lambda, \mu \in \mathbb{K}$ and $v \in V$.
- **Neutrality of 1:** We have $1v = v$ for all $v \in V$.
- **Annihilation I:** We have $0v = \overrightarrow{0}$ for all $v \in V$.
- **Annihilation II:** We have $\lambda \overrightarrow{0} = \overrightarrow{0}$ for all $\lambda \in \mathbb{K}$.

The operation $+$ is called the **addition** of the vector space V; the operation $\cdot$ is called the **scaling** of V; the element $\overrightarrow{0}$ is called the **zero vector** of V. The elements of V are called **vectors**.

So our new notion of "vector" is just "element of a vector space".

We say "$\mathbb{K}$-vector space" instead of "vector space over $\mathbb{K}$"; we also omit the $\mathbb{K}$ if it is clear from the context.

If v and w are two elements of a vector space V, then we define the **difference** $v - w$ to be $v + (-1) w$. We also set $-w = (-1) w$ for any $w \in V$.

Finite sums (like $v_1 + v_2 + \cdots + v_k$ or $\sum\limits_{i=1}^{k} v_i$ or $\sum\limits_{i \in I} v_i$) of elements of any vector space V make sense; they are defined just as they are defined for numbers. Empty sums (such as $\sum\limits_{i=3}^{2} v_i$) are defined to be the zero vector $\overrightarrow{0}$.

Don't get vector spaces confused with commutative rings! Both of these objects come with an addition "$+$" and a multiplication-like operation "$\cdot$". But in a commutative ring, the latter operation takes two elements of the ring as input, whereas in a vector space, it takes a scalar and a vector as input. In other words, in a commutative ring, you can multiply two elements of the ring, but in a vector space, you can "multiply" (or, more precisely, scale) a vector by a scalar (but **you cannot multiply two vectors**[8]). This is why, for example, there is no "commutativity of multiplication" axiom for a vector space (and no "unity vector" similar to the 1 of a commutative ring).

1.2.2. Examples of vector spaces

From now on, fix a field $\mathbb{K}$.

Definition 1.2.2. Fix $n, m \in \mathbb{N}$. The set of all $n \times m$-matrices with entries in $\mathbb{K}$ shall be denoted by $\mathbb{K}^{n \times m}$. This is a $\mathbb{K}$-vector space, with the operation $+$ being addition of matrices, with the operation $\cdot$ being scaling of matrices, and with the zero vector $\overrightarrow{0}$ being the zero matrix $0_{n \times m}$.

Thus, $n \times m$-matrices are vectors (in the sense of being elements of the vector space $\mathbb{K}^{n \times m}$). This does not mean that they are column vectors or row vectors; we are just using the word "vector" in the general sense of "element of a vector space" here.

But in particular, this includes row and column vectors. Indeed, $n \times 1$-matrices are column vectors, and $1 \times m$-matrices are row vectors. Thus, row vectors and column vectors are particular cases of vectors in the new sense of this word.

Note that the vector space $\mathbb{K}^{n \times m}$ of $n \times m$-matrices incorporates matrix addition and matrix scaling, but it doesn't "know" how matrices are multiplied; matrix multiplication is not part of its structure. (And indeed, if $n \neq m$, then you cannot multiply two $n \times m$-matrices, so matrix multiplication cannot be considered to be an "internal" operation on this vector space.)

Remark 1.2.3. Not only is $\mathbb{K}^{n \times m}$ a vector space, but several useful subsets of $\mathbb{K}^{n \times m}$ are vector spaces as well. For example,

$$\left\{ A \in \mathbb{K}^{n \times m} \mid \text{each row of } A \text{ sums to } 0 \right\}$$

is a vector space (with the same addition and scaling as before). We will later understand these as examples of "subspaces".

[8]Of course, in **some** vector spaces, you can do that. But the definition of a vector space is not responsible for that!

Here are some simpler examples of vector spaces:

- The field $\mathbb{K}$ itself is a $\mathbb{K}$-vector space (with addition being addition, and scaling being multiplication, and the $\overrightarrow{0}$ being 0). This is an example where scalars and vectors are the same thing.

  Thus, in particular, $\mathbb{R}$ is an $\mathbb{R}$-vector space, and $\mathbb{C}$ is a $\mathbb{C}$-vector space.

- However, $\mathbb{C}$ is also an $\mathbb{R}$-vector space, because you can scale complex numbers by real numbers (since real numbers are complex numbers).

- The one-element set $\{0\}$ is a $\mathbb{K}$-vector space for any field $\mathbb{K}$. (Here, addition and scaling and $\overrightarrow{0}$ are defined in the only possible way: i.e., by setting $0+0=0$ and $\lambda \cdot 0 = 0$ for all $\lambda \in \mathbb{K}$, and declaring 0 to be the zero vector.)

- The set of all infinite sequences $(a_1, a_2, a_3, \ldots) \in \mathbb{K}^\infty$ is also a $\mathbb{K}$-vector space, if we define addition by

  $$(a_1, a_2, a_3, \ldots) + (b_1, b_2, b_3, \ldots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)$$

  and define scaling by

  $$\lambda (a_1, a_2, a_3, \ldots) = (\lambda a_1, \lambda a_2, \lambda a_3, \ldots)$$

  and define the zero vector $\overrightarrow{0}$ by

  $$\overrightarrow{0} = (0, 0, 0, \ldots).$$

  You can view these infinite sequences as $1 \times \infty$-matrices, i.e., as row vectors of infinite length; then, these are just particular cases of matrix addition and scaling.

- The set $\mathbb{R}[x]$ of all polynomials in one variable x with real coefficients is an $\mathbb{R}$-vector space.

- Let S be any set, and $\mathbb{K}$ be any field. Consider the set $\mathbb{K}^S$ of all maps from S to $\mathbb{K}$. Then, $\mathbb{K}^S$ becomes a $\mathbb{K}$-vector space if we define addition and scaling pointwise[9] and define $\overrightarrow{0}$ to be the constant-0 map[10].

[9]i.e., the sum $f+g$ of two maps $f, g : S \to \mathbb{K}$ is defined by

$$(f+g)(s) = f(s) + g(s) \qquad \text{for all } s \in S;$$

and the map λf (for any $\lambda \in \mathbb{K}$ and $f : S \to \mathbb{K}$) is defined by

$$(\lambda f)(s) = \lambda \cdot f(s) \qquad \text{for all } s \in S.$$

[10]i.e., the map

$$S \to \mathbb{K}, \qquad s \mapsto 0$$

- Thus, for example, the set of all maps (= functions) from $\mathbb{R}$ to $\mathbb{R}$ is an $\mathbb{R}$-vector space.
- The set $\mathcal{C}(\mathbb{R})$ of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ is an $\mathbb{R}$-vector space as well.

1.3. Linear combinations

We can now recall some notions we defined for column vectors, and extend their definitions to vectors in arbitrary vector spaces over arbitrary fields. We begin with the notions of "linear combination", "relation", "linear independence" and "linear dependence":

Definition 1.3.1. Let $\mathbb{K}$ be any field. Let $v_1, v_2, \ldots, v_k$ be some vectors in a $\mathbb{K}$-vector space V. Then, a **$\mathbb{K}$-linear combination** of $v_1, v_2, \ldots, v_k$ means a vector that can be written in the form

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \qquad \text{for some } \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}.$$

We omit the "$\mathbb{K}$" from "$\mathbb{K}$-linear combination" if it is clear what field we mean.

Definition 1.3.2. Let $\mathbb{K}$ be a field. Let $v_1, v_2, \ldots, v_k$ be some vectors in a $\mathbb{K}$-vector space V.

(a) A **relation** (more precisely: a **$\mathbb{K}$-linear relation**) between $v_1, v_2, \ldots, v_k$ means a choice of $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ satisfying

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = \overrightarrow{0}.$$

(b) The **trivial relation** between $v_1, v_2, \ldots, v_k$ is the relation obtained by choosing $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. Clearly, $v_1, v_2, \ldots, v_k$ always have this trivial relation.

(c) We say that the vectors $v_1, v_2, \ldots, v_k$ (or, more precisely, the list $(v_1, v_2, \ldots, v_k)$) are **independent** (more precisely: **$\mathbb{K}$-linearly independent**) if the only relation between $v_1, v_2, \ldots, v_k$ is the trivial relation. Otherwise, we say that these vectors are **dependent**.

This encompasses independence of column vectors, of matrices, of polynomials and of many other things – because all of these things are vectors in vector spaces.

How do we tell whether a list of vectors in a vector space is independent? There is no general method (like for column vectors); you have to know the structure of your vectors.

Example 1.3.3. **(a)** Recall that $\mathbb{C}$ is a $\mathbb{C}$-vector space. Thus, any complex number is in itself a vector over $\mathbb{C}$. Consider the two vectors 1 and i in this space. Are they $\mathbb{C}$-linearly dependent?

Yes, because

$$1 \cdot 1 + i \cdot i = 0 \qquad \text{is a nontrivial relation between 1 and } i.$$

(b) Recall that $\mathbb{C}$ is an $\mathbb{R}$-vector space. Thus, any complex number is in itself a vector over $\mathbb{R}$. Consider the two vectors 1 and i in this space. Are they $\mathbb{R}$-linearly dependent? In other words, are there **real** coefficients λ_1, λ_2 (not both 0) such that $\lambda_1 \cdot 1 + \lambda_2 \cdot i = 0$?

No, because for real numbers λ_1, λ_2, we have $\lambda_1 \cdot 1 + \lambda_2 \cdot i = (\lambda_1, \lambda_2)$, and this complex number is only 0 if both λ_1 and λ_2 are 0.

Example 1.3.4. Consider the $\mathbb{R}$-vector space $\mathcal{C}(\mathbb{R})$ of continuous functions $\mathbb{R} \to \mathbb{R}$.

(a) Are the three functions 1, sin and cos linearly dependent? (Here, 1 means the constant-1 function, i.e., the function that sends each $x \in \mathbb{R}$ to 1.)

(b) Are the three functions 1, $\sin^2$ and $\cos^2$ linearly dependent? (Here, $\sin^2$ means the function that sends each $x \in \mathbb{R}$ to $(\sin x)^2$. Similarly, $\cos^2$ means the function that sends each $x \in \mathbb{R}$ to $(\cos x)^2$.)

Answers: (a) This would mean that there exist three reals $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that

$$\lambda_1 \cdot 1 + \lambda_2 \cdot \sin + \lambda_3 \cdot \cos = 0.$$

In other words, these reals would satisfy

$$\lambda_1 + \lambda_2 \sin x + \lambda_3 \cos x = 0 \qquad \text{for all } x \in \mathbb{R}.$$

We claim that they don't exist. To see why, assume that they exist. Thus,

$$\lambda_1 + \lambda_2 \sin x + \lambda_3 \cos x = 0 \qquad \text{for all } x \in \mathbb{R}.$$

Plugging $x = 0$ into this equation yields

$$\lambda_1 + \lambda_2 \underbrace{\sin 0}_{=0} + \lambda_3 \underbrace{\cos 0}_{=1} = 0, \qquad \text{that is,}$$
$$\lambda_1 + \lambda_3 = 0.$$

Plugging $x = \pi/2$ into the same equation yields

$$\lambda_1 + \lambda_2 \underbrace{\sin(\pi/2)}_{=1} + \lambda_3 \underbrace{\cos(\pi/2)}_{=0} = 0, \qquad \text{that is,}$$
$$\lambda_1 + \lambda_2 = 0.$$

Plugging $x = \pi$ into the same equation yields

$$\lambda_1 + \lambda_2 \underbrace{\sin \pi}_{=0} + \lambda_3 \underbrace{\cos \pi}_{=-1} = 0, \qquad \text{that is,}$$

$$\lambda_1 - \lambda_3 = 0.$$

Combining these three equations yields the system $\begin{cases} \lambda_1 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 = 0 \\ \lambda_1 - \lambda_3 = 0 \end{cases}$, whose only solution is $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Thus, the only relation between $1, \sin, \cos$ is trivial. Hence, $1, \sin, \cos$ are linearly independent.

(b) This would mean that there exist three reals $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that

$$\lambda_1 \cdot 1 + \lambda_2 \cdot \sin^2 + \lambda_3 \cdot \cos^2 = 0.$$

In other words, these reals would satisfy

$$\lambda_1 + \lambda_2 \sin^2 x + \lambda_3 \cos^2 x = 0 \qquad \text{for all } x \in \mathbb{R}.$$

Such reals exist: $\lambda_1 = -1$, $\lambda_2 = 1$ and $\lambda_3 = 1$. Indeed,

$$-1 + 1 \sin^2 x + 1 \cos^2 x = -1 + \underbrace{\sin^2 x + \cos^2 x}_{=1} = -1 + 1 = 0.$$

So $1, \sin^2, \cos^2$ are linearly dependent.

There is no general method to decide whether a bunch of functions in $\mathcal{C}(\mathbb{R})$ is linearly dependent or not. But once you have some dependence/independence statements, linear algebra can help derive some useful consequences from them.