Note 9

Vector spaces

9.1 Definition and examples of vector spaces

In the previous chapters, we have worked with linear combinations of vectors from $\mathbb{F}^n$, where $\mathbb{F}$ is a field (typically $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). We have seen that elements of $\mathbb{F}^n$ can be added and multiplied with scalars, that is to say, multiplied with elements from $\mathbb{F}$. It turns out to be a great advantage to take a more abstract point of view and describe several essential properties right from the start. One says that one gives these properties as axioms. This is similar in spirit to what we did when we defined a field. Also there, several properties of the real and complex numbers were put as axioms for such a field. In case of vectors and scalars, the result is the following:

Definition 9.1

A *vector space* over a field $\mathbb{F}$ is a set V of elements called *vectors*, together with two operations satisfying eight axioms. The first operation is called addition and denoted by $+$. It takes as input two elements $\mathbf{u}, \mathbf{v} \in V$ and returns a vector in V denoted by $\mathbf{u} + \mathbf{v}$. The second operation is called scalar multiplication and denoted by $\cdot$. It takes as input an element of $c \in \mathbb{F}$, in this context often called a *scalar*, and a vector $\mathbf{u} \in V$ and returns a vector in V denoted by $c \cdot \mathbf{u}$. The eight axioms that should be satisfied are:

1. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in V$
3. There exists a vector $\mathbf{0} \in V$ called the *zero vector*, such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u} \in V$
4. For any $\mathbf{u} \in V$ there exists an element $-\mathbf{u} \in V$ called the additive inverse of $\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. $c \cdot (d \cdot \mathbf{u}) = (c \cdot d) \cdot \mathbf{u}$ for all $\mathbf{u} \in V$ and all $c, d \in \mathbb{F}$
6. $1 \cdot \mathbf{u} = \mathbf{u}$ for all $\mathbf{u} \in V$
7. $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $c \in \mathbb{F}$
8. $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$ for all $\mathbf{u} \in V$ and all $c, d \in \mathbb{F}$

Note that in item 5 in the formula $(c \cdot d) \cdot \mathbf{u}$, the first $\cdot$ (in $c \cdot d$) denotes multiplication in the field $\mathbb{F}$, while the second $\cdot$ denotes the scalar multiplication on the vector space V. Similarly in item 8, in the formula $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$, the first $+$ denotes addition in $\mathbb{F}$, while the second $+$ denotes addition in V.

|||| Example 9.2

Let us take $V = \mathbb{F}^n$ together with the addition and scalar product we have defined before in equations (7-1) and (7-2). This gives an example of a vector space. To verify this, one should check if the eight vector space axioms from Definition 9.1 are satisfied. Note that five of them were mentioned already in Theorem 7.2. The zero vector in the third axiom is simply the zero vector in $\mathbb{F}^n$, while the additive inverse of a vector required in axiom four is given as:

$$-\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} -v_1 \\ \vdots \\ -v_n \end{bmatrix}$$

This only leaves the sixth axiom, but

$$1 \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 1 \cdot v_1 \\ \vdots \\ 1 \cdot v_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \text{ for all } \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{F}^n.$$

We see that $\mathbb{F}^n$ is a vector space over the field $\mathbb{F}$.

A vector space over the field $\mathbb{R}$ is often called a *real vector space*. Similarly, a vector space over the field $\mathbb{C}$ is often called a *complex vector space*. We have in the previous chapters actually encountered examples of vector spaces already. Let us give a few.

|||| Example 9.3

Consider the set $\mathbb{C}$ of complex numbers. If we take $\mathbb{F} = \mathbb{C}$ and $n = 1$ in Example 9.2, we obtain that we can see $\mathbb{C}$ as a vector space over itself. However, we can also see $\mathbb{C}$ as a vector space over the real numbers $\mathbb{R}$. Indeed, as $+$, we simply take addition of complex numbers. Since we can multiply any two complex numbers, we can certainly multiply a real number with a complex number. This gives us the needed scalar product. That all eight axioms from Definition 9.1 are satisfied, can be deduced from Theorems 3.10 and 3.11.

|||| Example 9.4

Very similarly as in Example 9.3, one can view the set of real numbers $\mathbb{R}$ as a vector space over itself, but also as a vector space over the field of rational numbers $\mathbb{Q}$ (see Examples 2.4 and 6.2 for a description of the field $\mathbb{Q}$).

|||| Example 9.5

Consider the set $\mathbb{F}^{m\times n}$ of $m \times n$ matrices with entries in a field $\mathbb{F}$. Using addition of matrices as defined in Definition 7.14 and scalar multiplication defined by:

$$c \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} c \cdot a_{11} & \cdots & c \cdot a_{1n} \\ \vdots & & \vdots \\ c \cdot a_{m1} & \cdots & c \cdot a_{mn} \end{bmatrix},$$

the first two items of Theorem 7.15 state that the first two vector field axioms are satisfied. The $m \times n$ matrix having zero entries only, plays the role of zero vector. All other axioms can be checked as well, but we leave this to the reader.

|||| Example 9.6

Consider the set $\mathbb{C}[Z]$ of polynomials in the variable Z with coefficients in $\mathbb{C}$ as defined in Definition 4.1. On this set, we have as addition $+$, the usual addition of polynomials. Also, we can multiply any two polynomials, so we certainly can multiply a constant polynomial with another polynomial. This gives us a scalar product on $\mathbb{C}[Z]$. We will not do so here, but one can show that all eight axioms from Definition 9.1 are satisfied. Hence $\mathbb{C}[Z]$ is a vector space over $\mathbb{C}$.

|||| Example 9.7

Consider the set F of all functions with domain $\mathbb{R}$ and codomain $\mathbb{R}$. If $f : \mathbb{R} \to \mathbb{R}$ and $r \in \mathbb{R}$ are given, one can define the function $r \cdot f : \mathbb{R} \to \mathbb{R}$ as $(r \cdot f)(a) = r \cdot f(a)$ for all $a \in \mathbb{R}$. This gives a scalar multiplication on F. Addition on F is defined in a similar way: if $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are given, the function $(f + g) : \mathbb{R} \to \mathbb{R}$ is defined as $(f + g)(a) = f(a) + g(a)$ for all $a \in \mathbb{R}$. One can verify that this gives F the structure of a vector space over $\mathbb{R}$. As zero vector, one takes the zero function: $\mathbf{0} : \mathbb{R} \to \mathbb{R}$, satisfying $a \mapsto 0$ for all $a \in \mathbb{R}$.

In all examples we have given above, it holds that the product of the scalar 0 with any vector is equal to the zero vector $\mathbf{0}$. However, none of the eight vector space axioms state that $0 \cdot \mathbf{u} = \mathbf{0}$ for all $\mathbf{u} \in V$. Fortunately, the eight vector space axioms are chosen well: one can deduce quite a lot from them, for example that the formula $0 \cdot \mathbf{u} = \mathbf{0}$ indeed is true for any vector space. We prove this and another intuitive formula in the following lemma:

Lemma 9.8

Let V be a vector space. Then

$$0 \cdot \mathbf{u} = \mathbf{0} \text{ for all } \mathbf{u} \in V \tag{9-1}$$

and

$$(-1) \cdot \mathbf{u} = -\mathbf{u} \text{ for all } \mathbf{u} \in V. \tag{9-2}$$

Proof. Using that $0 = 0 + 0$ and vector space axiom eight, we see that $0 \cdot \mathbf{u} = (0 + 0) \cdot \mathbf{u} = 0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}$. Adding $-(0 \cdot \mathbf{u})$ on both sides and using vector space axioms four, one and three, we get

$$\begin{aligned}
\mathbf{0} &= 0 \cdot \mathbf{u} + (-(0 \cdot \mathbf{u})) \\
&= (0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}) + (-(0 \cdot \mathbf{u})) \\
&= 0 \cdot \mathbf{u} + (0 \cdot \mathbf{u} + (-(0 \cdot \mathbf{u}))) \\
&= 0 \cdot \mathbf{u} + \mathbf{0} \\
&= 0 \cdot \mathbf{u}.
\end{aligned}$$

This shows the first part. The second part follows similarly. Since $0 = (1 + (-1))$, we obtain that $0 \cdot \mathbf{u} = (1 + (-1)) \cdot \mathbf{u} = 1 \cdot \mathbf{u} + (-1) \cdot \mathbf{u}$. The left-hand side of this equation is equal to $\mathbf{0}$ by the first part of this lemma. Using this and vector space axiom six, we see that $\mathbf{0} = \mathbf{u} + (-1) \cdot \mathbf{u}$. Hence $(-1) \cdot \mathbf{u} = -\mathbf{u}$. □

9.2 Basis of a vector space

Very similar to what we did in Section 7.1 for vectors in $\mathbb{F}^m$, one can talk about a *linear combination* of vectors in the setting of general vector spaces. Explicitly, given a vector space V over a field $\mathbb{F}$, vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ and scalars $c_1, \ldots, c_n \in \mathbb{F}$, an expression of the form

$$c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$$

is called a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Likewise, the notion of linear (in)dependency of a finite sequence of vectors from Definition 7.4 generalizes directly to the setting of vector spaces:

Definition 9.9

Let V be a vector space over a field $\mathbb{F}$. A sequence of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ is called *linearly independent* if and only if the equation $c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0}$ with $c_1, \ldots, c_n \in \mathbb{F}$ only holds if $c_1 = \cdots = c_n = 0$.
If the sequence of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ is not linearly independent, one says that it is *linearly dependent*.

Basically, the only difference with Definition 7.4 is that $\mathbb{F}^m$ has been replaced with V. Also in the setting of general vector spaces, it is common to simply say that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly (in)dependent rather than saying that the sequence of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly (in)dependent.

There is one complication concerning linear independence of vectors in general vector spaces. In Definition 9.9, we only consider *finitely many* vectors. It turns out that sometimes, we would like to be able to state that the vectors from a possibly infinite set are linearly independent. The following definition will allow us to do that:

Definition 9.10

Let V be a vector space over a field $\mathbb{F}$. The vectors in a set S of vectors are called *linearly independent* if and only if any finite sequence of distinct vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ from S is a linearly independent sequence of vectors.
If the vectors in S are not linearly independent, one says that they are *linearly dependent*.

Basically, in Definition 9.10, the number of vectors in the set S we consider may be infinite, but when determining if they are linearly independent, we only consider finitely many at the same time. Often we will work with finite sequences of vectors only, in which case Definition 9.9 can be used.

In Examples 7.5 and 7.6 we have already given examples of linearly dependent and linearly independent vectors in the vector space $\mathbb{R}^2$. Let us consider some more examples.

Example 9.11

In Example 9.3, we considered $\mathbb{C}$ as a vector space over $\mathbb{R}$. In this example, we give examples of linearly dependent and independent vectors. First of all, consider the elements 1 and i. To determine if these are linearly independent, we consider the equation $c_1 \cdot 1 + c_2 \cdot i = 0$, where $c_1, c_2 \in \mathbb{R}$. The reason we only allow c_1 and c_2 to be real numbers, is that we in this example consider $\mathbb{C}$ as a vector space over the field $\mathbb{R}$. Hence in Definition 9.9, we have $V = \mathbb{C}$ and $\mathbb{F} = \mathbb{R}$. In particular, the scalars only come from $\mathbb{R}$ by definition.

Returning to the equation $c_1 \cdot 1 + c_2 \cdot i = 0$, where $c_1, c_2 \in \mathbb{R}$, we see that the complex number $c_1 \cdot 1 + c_2 \cdot i$ is in rectangular form. Since two complex numbers are equal if and only if they have the same real and imaginary part, the equation $c_1 \cdot 1 + c_2 \cdot i = 0$ implies that $c_1 = 0$ and $c_2 = 0$. We conclude that the complex numbers 1 and i are linearly independent over $\mathbb{R}$.

Similarly, one can show that the complex numbers 2 and $1 + i$ are linearly independent. Indeed, suppose that $c_1 \cdot 2 + c_2 \cdot (1 + i) = 0$, for some $c_1, c_2 \in \mathbb{R}$. Considering real and imaginary part, we see that this implies that $2c_1 + c_2 = 0$ and $c_2 = 0$, whence $c_1 = c_2 = 0$.

As a final example, let us consider a sequence of three complex numbers, for example $2, 1 + i$ and $2 + 3i$. Since $-(1/2) \cdot 2 + 3 \cdot (1 + i) + (-1) \cdot (2 + 3i) = 0$, we see that the three complex numbers 2, $1 + i$, and $2 + 3i$ are linearly dependent over $\mathbb{R}$.

Example 9.12

In Example 9.5, we viewed the set of matrices $\mathbb{F}^{m \times n}$ as a vector space over $\mathbb{F}$. For any pair (i, j) satisfying $1 \leq i \leq m$ and $1 \leq j \leq n$, define the matrix $\mathbf{E}^{(i,j)} \in \mathbb{F}^{m \times n}$ to be the matrix having zero entries, except for the entry (i, j), which is equal to one. For $m = n = 2$, we have for example

$$\mathbf{E}^{(1,1)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{E}^{(1,2)} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{E}^{(2,1)} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \text{ and } \mathbf{E}^{(2,2)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Continuing with $m = n = 2$, we see that the matrices $\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)}$ are linearly independent over $\mathbb{F}$. Indeed for any $c_1, c_2, c_3, c_4 \in \mathbb{F}$, one has

$$c_1 \cdot \mathbf{E}^{(1,1)} + c_2 \cdot \mathbf{E}^{(1,2)} + c_3 \cdot \mathbf{E}^{(2,1)} + c_4 \cdot \mathbf{E}^{(2,2)} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}.$$

Hence $c_1 \cdot \mathbf{E}^{(1,1)} + c_2 \cdot \mathbf{E}^{(1,2)} + c_3 \cdot \mathbf{E}^{(2,1)} + c_4 \cdot \mathbf{E}^{(2,2)} = \mathbf{0}$ implies that $c_1 = c_2 = c_3 = c_4 = 0$.

For general m and n one can show similarly that the $m \times n$ matrices $\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)}$ are linearly independent over F.

Returning to $m = n = 2$, an example of a sequence of linearly dependent matrices is:

$$\begin{bmatrix} -1 & 0 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 5 & 4 \\ 2 & 0 \end{bmatrix},$$

since

$$1 \cdot \begin{bmatrix} -1 & 0 \\ 2 & 4 \end{bmatrix} - 4 \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 5 & 4 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

|||| Example 9.13

Consider the complex vector space $\mathbb{C}[Z]$ from Example 9.6. Recall that two polynomials $p_1(Z) = a_0 + a_1 Z \cdots + a_n Z^n$ of degree n and $p_2(Z) = b_0 + b_1 Z \cdots + b_m Z^m$ of degree m are equal if and only if $n = m$ and $a_i = b_i$ for all i. This implies in particular, that a polynomial $p(Z) = c_0 + c_/1Z \cdots + c_n Z^n$ is equal to the zero polynomial if and only if $c_i = 0$ for all i. This shows that the set $\{1, Z, Z^2, \dots\}$ is a set of linearly independent polynomials over $\mathbb{C}$.

All these examples show that the notion of linear independence carries over well to the setting of general vector spaces. With this in place, we come to a very important notion in the theory of vector spaces.

|||| Definition 9.14

Let V be a vector space over a field $\mathbb{F}$. A set S of vectors is called a *basis* of V if the two following conditions are met:

1. The vectors in S are linearly independent.
2. Any $\mathbf{v} \in V$ can be written as a linear combination of vectors in S.

An *ordered basis* $(\mathbf{v}_1, \mathbf{v}_2, \dots)$ is a list of vectors, such that the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots\}$ is a basis of V.

It turns out that any vector space has a basis and we will freely use this fact. A reader who has time and motivation for a bit of extra material about this is referred to Section 9.4, but this is not required reading. If a vector space has a finite basis, i.e., if the set S containing the basis vectors, is finite, we can enumerate the elements in S and write $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Then $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ is a finite ordered basis of V. Hence any vector space with a finite basis has an ordered basis.

Before giving examples, let us give one lemma and one more definition.

Lemma 9.15

Let V be a vector space over a field $\mathbb{F}$ that has a finite ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Then any vector $\mathbf{v} \in V$ can be written in exactly one way as a linear combination of the basis vectors.

Proof. The second part of Definition 9.14 guarantees that any vector $\mathbf{v} \in V$ can be written as a linear combination of the basis vectors, say $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n$ for certain $c_1, \ldots, c_n \in \mathbb{F}$. What we need to show, is that this is the only way to write $\mathbf{v}$ as a linear combination of the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Suppose therefore that $\mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n$ for certain $d_1, \ldots, d_n \in \mathbb{F}$. We wish to show that $c_1 = d_1, \ldots, c_n = d_n$. First of all, we have

$$c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n.$$

Therefore,

$$c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n - (d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n) = \mathbf{0},$$

which in turn implies that

$$(c_1 - d_1) \cdot \mathbf{v}_1 + \cdots (c_n - d_n) \cdot \mathbf{v}_n = \mathbf{0}.$$

However, since the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent (this follows from the first part of Definition 9.14), we see that $c_1 - d_1 = 0, \ldots, c_n - d_n = 0$. But then $c_1 = d_1, \ldots, c_n = d_n$, which is what we wanted to show. □

This Lemma 9.15 motivates the following definition:

|||| Definition 9.16

Let V be a vector space over a field $\mathbb{F}$ that has a finite ordered basis $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. If for $\mathbf{v} \in V$, we have

$$\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n,$$

then we define

$$[\mathbf{v}]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$$

to be the *coordinate vector* of $\mathbf{v}$ with respect to the ordered basis β. One also says that $[\mathbf{v}]_\beta$ is the *β-coordinate vector* of $\mathbf{v}$.

The function sending a vector of V to its β-coordinate vector, has several nice properties. Two of them will be useful later on.

|||| Lemma 9.17

Let V be a vector space over a field $\mathbb{F}$ that has a finite ordered basis β. Then we have:

$$[\mathbf{u} + \mathbf{v}]_\beta = [\mathbf{u}]_\beta + [\mathbf{v}]_\beta \text{ for all } \mathbf{u}, \mathbf{v} \in V$$

and

$$[c \cdot \mathbf{v}]_\beta = c \cdot [\mathbf{v}]_\beta \text{ for all } c \in \mathbb{F} \text{ and } \mathbf{v} \in V.$$

Proof. We prove the first item only and leave the proof of the second one to the reader. Let us say that the ordered basis β is given by $\mathbf{v}_1, \ldots, \mathbf{v}_n$. If $\mathbf{u} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n$ and $\mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n$, then $\mathbf{u} + \mathbf{v} = (c_1 + d_1) \cdot \mathbf{v}_1 + \cdots (c_n + d_n) \cdot \mathbf{v}_n$. Hence

$$[\mathbf{u} + \mathbf{v}]_\beta = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_\beta + [\mathbf{v}]_\beta.$$

□

Now we will use this lemma to prove a theorem involving linear independence of vectors.

Theorem 9.18

Let V be a vector space over a field $\mathbb{F}$ that has a finite ordered basis β consisting of n vectors. Suppose we are given $\mathbf{u}_1, \ldots, \mathbf{u}_\ell \in V$ and $c_1, \ldots, c_\ell \in \mathbb{F}$. Then

$$c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell = \mathbf{0} \text{ if and only if } c_1 \cdot [\mathbf{u}_1]_\beta + \cdots + c_\ell [\cdot \mathbf{u}_\ell]_\beta = \mathbf{0}.$$

In particular, the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ in V are linearly independent if and only if the vectors $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$ in $\mathbb{F}^n$ are linearly independent.

Proof. A vector $\mathbf{v}$ in V is the zero vector if and only if its β-coordinate vector is the zero vector. Hence $c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell = \mathbf{0}$ if and only if $[c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell]_\beta = \mathbf{0}$. Using Lemma 9.17 repeatedly, we can also deduce that $[c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell]_\beta = c_1 \cdot [\mathbf{u}_1]_\beta + \cdots + c_\ell \cdot [\mathbf{u}_\ell]_\beta$. Hence the first part of the theorem follows. The second part follows directly from the first part. ☐

This theorem basically reduces the question of linear (in)dependence of vectors in V to a question of linear (in)dependence of vectors in $\mathbb{F}^n$. However, for $\mathbb{F}^n$, we already have techniques at our disposal, notably Theorem 7.8.

Example 9.19

Let $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{R}^2$. We claim that the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

form an ordered basis β for $\mathbb{R}^2$. Indeed, these vectors are linearly independent (the reader is encouraged to check this), and any vector is a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$, since

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

This means that in this case $[\mathbf{v}]_\beta = \mathbf{v}$.

Now let γ be the sequence of vectors

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$

We have seen in Example 7.6 that these two vectors are linearly independent. Further, one can show that any vector in $\mathbb{R}^2$ can be written as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. Indeed, given $v_1, v_2 \in \mathbb{R}$, the equation

$$c_1 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

gives rise to a system of two linear equations in the variables c_1, c_2. Solving this system, one can show that for any $v_1, v_2 \in \mathbb{R}$, we have

$$c_1 = -\frac{v_1}{3} + \frac{2v_2}{3} \quad \text{and} \quad c_2 = \frac{2v_1}{3} - \frac{v_2}{3}$$

so that

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \left(-\frac{v_1}{3} + \frac{2v_2}{3}\right) \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \left(\frac{2v_1}{3} - \frac{v_2}{3}\right) \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

This means that $\gamma = (\mathbf{u}, \mathbf{v})$ is an ordered basis of $\mathbb{R}^2$. Moreover, from the above we see that

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}_\gamma = \begin{bmatrix} -v_1/3 + 2v_2/3 \\ 2v_1/3 - v_2/3 \end{bmatrix}.$$

This first part of Example 9.19 can be expanded further: as in Section 7.3, let us denote the i-th column of the identity matrix $\mathbf{I}_n \in \mathbb{F}^{n \times n}$ by $\mathbf{e}_i$ for $i = 1, \ldots, n$. In other words: the vector $\mathbf{e}_i$ has 1 as its i-th coordinate and zeroes everywhere else. These vectors form an ordered basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ of the vector space $\mathbb{F}^n$ called the *standard (ordered) basis*. For the sake of completeness, let us show that they form an ordered basis:

|||| Proposition 9.20

The vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ form an ordered basis of the vector space $\mathbb{F}^n$ over $\mathbb{F}$.

Proof. According to Definition 9.14, we need to check two things:

1. The vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are linearly independent.
2. Any vector in $\mathbb{F}^n$ can be written as a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

The first item follows from the observation that

$$c_1 \cdot \mathbf{e}_1 + c_2 \cdot \mathbf{e}_2 + \cdots + c_n \cdot \mathbf{e}_n = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Indeed, this equation implies that if a linear combination is equal to the zero vector in $\mathbb{F}^n$, then all scalars $c_1, \ldots, c_n$ are zero. The second item follows, since if $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}^n$ is given, then

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_1 \cdot \mathbf{e}_1 + v_2 \cdot \mathbf{e}_2 + \cdots + v_n \cdot \mathbf{e}_n.$$

□

Similarly as in Example 9.19, if β is the standard ordered basis of $\mathbb{F}^n$, then $[\mathbf{v}]_\beta = \mathbf{v}$ for all $\mathbf{v} \in \mathbb{F}^n$. Note though that just as in Example 9.19, the vector space $\mathbb{F}^n$ has many more possible ordered bases. Now let us continue with giving examples of bases of vector spaces.

|||| Example 9.21

Continuing Examples 9.3 and 9.11, we know that the complex numbers 1 and i are linearly independent over $\mathbb{R}$. They form an ordered basis $(1, i)$, which we denote by β, since any complex number is a linear combination of 1 and i over the real numbers. More specifically, for any $a, b \in \mathbb{R}$, we have $a + bi = a \cdot 1 + b \cdot i$. Therefore, for $a, b \in \mathbb{R}$, we have

$$[a + bi]_\beta = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2.$$

Hence $[a + bi]_\beta$ is equal to the rectangular coordinates of the complex number $a + bi$.

There are many more possible bases (and hence ordered bases) for $\mathbb{C}$ when viewed as vector space over $\mathbb{R}$. For example, $(2, 1 + i)$ is a possible ordered basis. Indeed, we have already seen in Example 9.11 that the complex numbers 2 and $1 + i$ are linearly independent over $\mathbb{R}$. Also any complex number can be written as a linear combination with coefficients in $\mathbb{R}$ of 2 and $1 + i$. To see this, we need to check that for a given complex number $a + bi$, where $a, b \in \mathbb{R}$, the equation $a + bi = c_1 \cdot 2 + c_2 \cdot (1 + i)$ has a solution $c_1, c_2 \in \mathbb{R}$. Considering real and imaginary parts, we see that $a = 2c_1 + c_2$ and $b = c_2$. Hence we have as solution $c_2 = b$ and $c_1 = (a - c_2)/2 = (a - b)/2$. Denoting the ordered basis $(2, 1 + i)$ by γ, we have

$$[a + bi]_\gamma = \begin{bmatrix} (a - b)/2 \\ b \end{bmatrix} \in \mathbb{R}^2.$$

|||| Example 9.22

Continuing Examples 9.5 and 9.12, we can find an ordered basis β of the vector space $\mathbb{F}^{m\times n}$ over $\mathbb{F}$. This ordered basis is $(\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)})$. We have already seen that the matrices $\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)}$ are linearly independent, while any matrix $\mathbf{A} = (a_{ij})_{1\le i\le m;1\le j\le n}$ can be written as a linear combination of them, namely $\mathbf{A} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathbf{E}^{(i,j)}$.

Specifically for $m = n = 2$, the matrices $\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)}$ form an ordered basis $\beta = (\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)})$ and we have

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_\beta = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}.$$

|||| Example 9.23

In this example, we again consider the complex vector space $\mathbb{C}[Z]$ from Examples 9.6 and 9.13. From these examples, we already know that the set $\{1, Z, Z^2, \ldots\}$ is a set of linearly independent polynomials over $\mathbb{C}$. However, by definition of polynomials, any polynomial is a linear combination over $\mathbb{C}$ of finitely many elements from this set. Therefore the set $\{1, Z, Z^2, \ldots\}$ is in fact a basis of the complex vector space $\mathbb{C}[Z]$. This is an example of a vector space having an infinite basis.

It turns out that for a given vector space V over a field $\mathbb{F}$, the number of vectors in a basis of V is always the same. Later in this section, we will prove this in the special case where the number of vectors in a basis is finite. In general, the number of elements in a basis of V is called the *dimension* of the vector space V. A common notation for the dimension of a vector space V is: $\dim(V)$ or just $\dim V$. If one wants to make clear over which field $\mathbb{F}$ the vector space is defined, one writes $\dim_{\mathbb{F}}(V)$ or $\dim_{\mathbb{F}} V$. If the number of vectors in a basis is finite, one says that V has finite dimension, otherwise one says that V has infinite dimension, which can also be expressed in a formula as: $\dim V = \infty$.

|||| Example 9.24

Let us compute the dimensions of various examples of vector spaces that we have encountered so far. First of all from Example 9.19, we see that $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$. Much more generally one has $\dim_{\mathbb{F}}(\mathbb{F}^n) = n$, since a possible basis of $\mathbb{F}^n$ is formed by the n vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

A special case of the above is when $\mathbb{C}$ is viewed as a vector space over itself. Then it has dimension one: $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ (a possible basis is formed by the complex number 1). However, if $\mathbb{C}$ is viewed as a vector space over $\mathbb{R}$, a basis is given by $\{1, i\}$ as we have seen in Example 9.21. Hence $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

The vector space of $m \times n$ matrices $\mathbb{F}^{m \times n}$ has a basis consisting of the mn matrices $\mathbf{E}^{(i,j)}$ with $1 \le i \le m$ and $1 \le j \le n$, as we have seen in Example 9.22. Hence $\dim_{\mathbb{F}}(\mathbb{F}^{m \times n}) = mn$.

We have seen in Example 9.13 that the complex vector space $\mathbb{C}[Z]$ has a basis with infinitely many elements, namely $\{1, Z, Z^2, \dots\}$. Hence $\dim_{\mathbb{C}}(\mathbb{C}[Z]) = \infty$.

|||| Theorem 9.25

If V has a finite basis consisting of n vectors, any other set of linearly independent vectors in V has at most n elements.

Proof. Let us denote the basis vectors by $\mathbf{v}_1, \dots, \mathbf{v}_n$ and denote the resulting ordered basis by β. We will prove the theorem by contradiction. Assume therefore that there exists a set of at least $n+1$ linearly independent vectors, say $\mathbf{w}_1, \dots, \mathbf{w}_{n+1}$. Since β is an ordered basis, we can find scalars $a_{ij} \in \mathbb{F}$ such that

$$\mathbf{w}_j = a_{1j}\mathbf{v}_1 + \cdots + a_{nj}\mathbf{v}_n \text{ for } j = 1, \dots, n+1.$$

Now let $\mathbf{A} = (a_{ij}) \in \mathbb{F}^{n \times (n+1)}$ be the matrix with entries a_{ij}. Note that the j-th column in $\mathbf{A}$ is equal to $[\mathbf{w}_j]_\beta$. Since $\mathbf{A}$ has n rows, its rank $\rho(\mathbf{A})$ is at most n. Since $\mathbf{A}$ has $n+1$ columns, this implies that $\rho(\mathbf{A}) < n+1$. Then by Corollary 6.30, we see that the homogeneous system with coefficient matrix $\mathbf{A}$ has nonzero solutions. Let $(c_1, \dots, c_{n+1}) \in \mathbb{F}^{n+1}$ be such a nonzero solution. Then we have

$$c_1 \cdot \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + c_{n+1} \cdot \begin{bmatrix} a_{1\,n+1} \\ \vdots \\ a_{n\,n+1} \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Now recall that the j-th column in $\mathbf{A}$ is equal to $[\mathbf{w}_j]_\beta$. This means that we have $c_1 \cdot [\mathbf{w}_1]_\beta + c_{n+1} \cdot [\mathbf{w}_{n+1}]_\beta = \mathbf{0}$. Since from Lemma 9.17 one can deduce that $[c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1}]_\beta = c_1 \cdot [\mathbf{w}_1]_\beta + c_{n+1} \cdot [\mathbf{w}_{n+1}]_\beta$, we may conclude that $[c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1}]_\beta = \mathbf{0}$. Hence $c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1} = \mathbf{0}$. Since $(c_1, \dots, c_{n+1})$ was not the zero vector, we conclude that the vectors $\mathbf{w}_1, \dots, \mathbf{w}_{n+1}$ are not linearly independent after all. This contradiction shows that the assumption that there exists sets with at least $n+1$ linearly independent vectors was wrong. Hence the theorem is true. □

Corollary 9.26

If V has a finite basis consisting of n vectors, any other basis for V contains precisely n vectors as well.

Proof. Let S be a basis of V consisting of n vectors and T any other basis. Since the vectors T are linearly independent, Theorem 9.25 implies that the number of vectors in T is at most n. Let us denote by m, the number of vectors in T. What we have just shown is that $m \leq n$. Now applying Theorem 9.25 again, but now taking T as a basis, we can conclude that the number of elements in S is at most m, that is: $n \leq m$. Combining the inequalities $m \leq n$ and $n \leq m$, we conclude that $n = m$, which is what we wanted to show. □

This corollary justifies the definition of dimension of a vector space V as the number of basis vectors in the finite dimensional case: no matter which basis of V you pick, it will contain precisely the same number of vectors. As mentioned before, the basis vectors themselves typically will be different when comparing two possible bases. In fact, for finite dimensional vector spaces, we can characterize all possible bases:

Theorem 9.27

Let V be a vector space over a field $\mathbb{F}$ of dimension n. Then any set of n linearly independent vectors in V is a basis for V.

Proof. Let us denote the vectors in some basis of V as $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and let us write β for the corresponding ordered basis. Further, let $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be n linearly independent vectors in V. To show that these form a basis, all we need to check is item 2 in Definition 9.14. That is to say, we need to show that any $\mathbf{v} \in V$ can be written as a linear combination of $\mathbf{w}_1, \ldots, \mathbf{w}_n$. First of all, since β is a basis, we can find $a_{ij} \in \mathbb{F}$ such that

$$\mathbf{w}_j = a_{1j} \cdot \mathbf{v}_1 + \cdots + a_{nj} \cdot \mathbf{v}_n \text{ for } j = 1, \ldots, n,$$

or equivalently using the summation symbol:

$$\mathbf{w}_j = \sum_{i=1}^{n} a_{ij} \cdot \mathbf{v}_i \text{ for } j = 1, \ldots, n. \tag{9-3}$$

Now let $\mathbf{A} = (a_{ij}) \in \mathbb{F}^{n \times n}$ be the matrix with entries a_{ij}. As in the proof of Theorem 9.25, note that the j-th column in $\mathbf{A}$ is equal to $[\mathbf{w}_j]_\beta$. We claim that these columns are linearly independent vectors in $\mathbb{F}^n$. To see why, suppose that $c_1 \cdot [\mathbf{w}_1]_\beta + c_n \cdot [\mathbf{w}_n]_\beta = \mathbf{0}$ for certain $c_1, \ldots, c_n \in \mathbb{F}$. Then $[c_1 \cdot \mathbf{w}_1 + \cdots + c_n \cdot \mathbf{w}_n]_\beta = \mathbf{0}$, implying that $c_1 \cdot \mathbf{w}_1 + \cdots + c_n \cdot \mathbf{w}_n = \mathbf{0}$. Using that the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ are linearly independent, we conclude that $c_1 = 0, \ldots, c_n = 0$, which is what we wanted to show to prove our claim. Now using Theorem 7.8 and Corollary 7.25, we conclude that the matrix $\mathbf{A}$ has an inverse matrix $\mathbf{A}^{-1}$.

Now let us return to what we want to show: $\mathbf{v} \in V$ can be written as a linear combination of $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Since $\mathbf{v}$ is a linear combination of the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, it is enough to show that each of the basis vectors themselves can be written as a linear combination of $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Let us write $\mathbf{A}^{-1} = (c_{ij})_{1 \le i \le n; 1 \le j \le n}$. We claim that:

$$\mathbf{v}_j = c_{1j} \cdot \mathbf{w}_1 + \cdot + c_{nj} \cdot \mathbf{w}_n \text{ for } j = 1, \ldots, n.$$

Equivalently, using the summation symbol, we claim that:

$$\mathbf{v}_j = \sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k \text{ for } k = 1, \ldots, n.$$

To show the claim, first we use equation (9-3) to see that:

$$\begin{aligned}
\sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k &= \sum_{k=1}^{n} c_{kj} \cdot \left(\sum_{i=1}^{n} a_{ik} \cdot \mathbf{v}_i \right) \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} c_{kj} \cdot a_{ik} \cdot \mathbf{v}_i \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} \cdot c_{kj} \cdot \mathbf{v}_i \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} \cdot c_{kj} \cdot \mathbf{v}_i \\
&= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} \cdot c_{kj} \right) \cdot \mathbf{v}_i.
\end{aligned}$$

Now note that the expression $\sum_{k=1}^{n} a_{ik} \cdot c_{kj}$ is the (i, j)-th entry of the matrix product $\mathbf{A} \cdot \mathbf{A}^{-1}$. However, since $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n$, we see that $\sum_{k=1}^{n} a_{ik} \cdot c_{kj} = 1$ if $i = j$ and $\sum_{k=1}^{n} a_{ik} \cdot c_{kj} = 0$ otherwise. Hence we can conclude that $\sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k = \mathbf{v}_j$, which is exactly what we wanted to show. $\square$

9.3 Subspaces of a vector space

Given a vector space V over some field $\mathbb{F}$, it can happen that a subset W of V is closed under the scalar multiplication and the vector addition as defined on V. The word "closed" is just a way of saying that if $\mathbf{v} \in W$ and $c \in \mathbb{F}$, then $c \cdot \mathbf{v} \in W$ and if $\mathbf{u}, \mathbf{v} \in W$, then $\mathbf{u} + \mathbf{v} \in W$. Since V is a vector space, we always have $c \cdot \mathbf{v} \in V$ and $\mathbf{u} + \mathbf{v} \in V$, but if W is closed under the scalar multiplication and addition, the vectors $c \cdot \mathbf{v}$ and $\mathbf{u} + \mathbf{v}$ end up in W again. Let us consider two examples of this:

|||| Example 9.28

Let us consider the complex vector space $\mathbb{C}^2$ and consider the subset $W = \{(z, 2 \cdot z) \mid z \in \mathbb{C}\}$. Then adding two elements of W yields another element of W, since $(z, 2 \cdot z) + (w, 2 \cdot w) = (z + w, 2 \cdot (z + w))$ for all $z, w \in \mathbb{C}$. Also multiplying an element from W with a scalar $c \in \mathbb{C}$ yields an element of W, since $c \cdot (z, 2 \cdot z) = (c \cdot z, 2 \cdot (c \cdot z))$. In fact W is a vector space using this scalar multiplication and addition. For example, one has $(0, 0) \in W$, since $(0, 0) = (0, 2 \cdot 0)$. Also $-(z, 2 \cdot z) = ((-z), 2 \cdot (-z))$ for any $z \in \mathbb{C}$, which shows that if $\mathbf{v} \in W$, then also $-\mathbf{v} \in W$. The reader is encouraged to check the remaining axioms of a vector space. Note that $\dim_{\mathbb{C}}(W) = 1$ (a possible basis is given by $\{(1, 2)\}$).

|||| Example 9.29

Consider the vector space $\mathbb{R}^{2\times 2}$ of 2 by 2 matrices with coefficients in $\mathbb{R}$. As we have seen, this is a real vector space of dimension four. Now let D be the subset of $\mathbb{R}^{2\times 2}$ consisting of all diagonal matrices, that is:

$$D = \left\{ \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \right\}.$$

Then the set D is closed under scalar multiplication and matrix addition. What this means is that if $\mathbf{A}, \mathbf{B} \in D$ and $c \in \mathbb{F}$, then $c \cdot \mathbf{A} \in D$ and $\mathbf{A} + \mathbf{B} \in D$. Let us check this. If $\mathbf{A}$ has diagonal elements λ_1 and λ_2 and $\mathbf{B}$ has diagonal elements μ_1 and μ_2, then:

$$c \cdot \mathbf{A} = c \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} c \cdot \lambda_1 & 0 \\ 0 & c \cdot \lambda_2 \end{bmatrix} \in D,$$

and

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} + \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 + \mu_1 & 0 \\ 0 & \lambda_2 + \mu_2 \end{bmatrix} \in D$$

One can check that D is in fact a real vector space of dimension two: a possible ordered basis is $(\mathbf{E}^{(1,1)}, \mathbf{E}^{(2,2)})$.

To capture these type of examples, we have the following:

Definition 9.30

Let V be a vector space over a field $\mathbb{F}$. A *subspace* of V is a subset W of V that is a vector space over $\mathbb{F}$ under the scalar multiplication and vector addition defined on V.

In other words, if $W \subseteq V$ is closed under the scalar multiplication and vector addition that V has, W "inherits" these operations. If W with these operations satisfies all vector space axioms from Definition 9.1, it is called a subspace of W. Any vector space V has at least two subspace: V itself can be seen as a subspace, and also the subspace $\{\mathbf{0}\}$ containing only the zero vector of V. In general, V has many more subspaces. In all cases, however, one can say the following about the dimension of a subspace:

Lemma 9.31

Let V be a vector space over a field $\mathbb{F}$ of dimension n and W a subspace of V. Then $\dim W \leq n$.

Proof. Since V has a basis with n vectors, and W has a basis with $\dim W$ vectors. The basis vectors of W form a sequence of $\dim W$ linearly independent vectors. Hence Theorem 9.25 implies that $\dim W \leq n$. $\square$

Since V already satisfies all vector space axioms, it turns out not to be necessary to check them all when investigating if a subset W is a subspace. More precisely, we have the following lemma:

Lemma 9.32

Let V be a vector space over $\mathbb{F}$ and W a nonempty subset of V. Then W is a subspace of V if the following is satisfied:

$$\text{for all } \mathbf{u}, \mathbf{v} \in W \text{ and all } c \in \mathbb{F} \text{ it holds that } \mathbf{u} + c \cdot \mathbf{v} \in W. \tag{9-4}$$

Proof. First let us show that W is closed under the scalar multiplication and vector addition of V. First of all, since W is not empty, it contains at least one vector, say $\mathbf{w}$. Then choosing $\mathbf{u} = \mathbf{w}$ and $\mathbf{v} = \mathbf{w}$ in equation (9-4), we can conclude that the vector $\mathbf{w} + (-1) \cdot \mathbf{w}$ is also in W. Using for example equation (9-2), this implies that $\mathbf{0} \in W$. Now that we know this, we can apply equation (9-4) again, but now with $\mathbf{u} = \mathbf{0}$ and $\mathbf{v} \in W$ chosen arbitrarily. We can hence conclude that for arbitrary $\mathbf{v} \in W$, also $c \cdot \mathbf{v}$ is in W. This shows that W is closed under scalar multiplication. Applying equation (9-4) for arbitrary $\mathbf{u}, \mathbf{v} \in W$ and $c = 1$, we conclude that $\mathbf{u} + \mathbf{v}$ is in W. Hence W is closed under vector addition.

Now let us show that W is a vector space by considering the eight vector space axioms from Definition 9.1. Items 1, 2, 5, 6, 7, and 8 actually hold for all vectors in V and therefore certainly for all vectors in a subset of V. Therefore, all that remains to be checked is that items 3 and 4 are satisfied. Item 3 is fulfilled, since we already have shown that equation (9-4) implies that $\mathbf{0} \in W$. As for item 4, if $\mathbf{v} \in W$, then $(-1) \cdot \mathbf{v} \in W$, since W is closed under scalar multiplication. But by equation (9-2), $(-1) \cdot \mathbf{v} = -\mathbf{v}$, so that indeed the additive inverse $-\mathbf{v}$ is in W, for all $\mathbf{v}$ in W. □

|||| Example 9.33

Using Lemma 9.32, it is not hard to show that the subsets W and D from Examples 9.28 and 9.29 are subspaces. The reader is encouraged to check that the condition in equation (9-4) is satisfied for these examples.

|||| Example 9.34

Let C_∞ be the set of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$. It is out of scope of these notes to define very precisely what an infinitely differentiable function is, but roughly speaking this means the following: if for all $x \in \mathbb{R}$ the limit $\lim_{a\to 0}(f(x+a) - f(x))/a$ exists, we can define the derivative of f, denoted by f', to be the function $f' : \mathbb{R} \to \mathbb{R}$ with $x \mapsto \lim_{a\to 0}(f(x+a) - f(x))/a$. An infinitely differentiable function $f : \mathbb{R} \to \mathbb{R}$ has the property that one can keep on differentiating it as often as one wants. In particular, not only its derivative f' exists, but also the derivative of f' (denoted by f'' or $f^{(2)}$), the derivative of f'' (denoted by f''' or $f^{(3)}$), and so on. More generally for any positive integer n, one denotes with $f^{(n)}$ the n-th derivative of f. More precisely, one recursively defines the n-th derivative as follows:

$$f^{(n)} = \begin{cases} f & \text{if } n = 0, \\ (f^{(n-1)})' & \text{if } n > 0. \end{cases}$$

The set C_∞ is a subspace of the vector space F from Example 9.7. This amounts to showing that if $f, g \in C_\infty$ and $c \in \mathbb{R}$, then also $f + c \cdot g \in C_\infty$. In fact one can show inductively that $(f + c \cdot g)^{(n)} = f^{(n)} + c \cdot g^{(n)}$ for any $n \in \mathbb{Z}_{\geq 0}$. In particular, $f + c \cdot g$ is infinitely differentiable, which is what we needed to show.

There is one specific way to construct a subspace, which we will get in to now.

Definition 9.35

Let V be a vector space over $\mathbb{F}$ and S a set of vectors from V. Then the *span* of S, denotes by $\mathrm{Span}(S)$ is the set of all possible linear combinations of vectors from S. In particular, if $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then

$$\mathrm{Span}(S) = \{c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n \mid c_1, \ldots, c_n \in \mathbb{F}\}.$$

It is customary to define $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$. As a consequence one also says that the empty set $\emptyset$ is a basis for the vector space $\{\mathbf{0}\}$. One can verify that for any subset $S \subseteq V$, the set $\mathrm{Span}(S)$ is in fact a subspace of V, using for example Lemma 9.32. If W is a given subspace of a vector space V and $W = \mathrm{Span}(S)$, one says that the vectors in S span W. One also says in this situation that W is spanned by the vectors in S. The vectors in a basis of W will certainly span W, but in general a set of vectors spanning W need not be linearly independent.

Example 9.36

Consider the real vector space $V = \mathbb{R}^3$ and let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \text{ and } \mathbf{v}_3 = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}.$$

Question: Find a basis of the subspace W spanned by the three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Answer:

A first, but unfortunately wrong, guess could be that the three vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ themselves form a basis. Certainly any vector in W can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$. This is a direct consequence of the Definition 9.35 of the span. However, in order to be a basis, the three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ would have to be linearly independent as well. It turns out they are not. Using Theorem 7.8 this can be determined by calculating the reduced row echelon form of the 3×3 matrix $\mathbf{A}$ with columns $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$. We omit the details of this calculation, but instead encourage the reader to verify that this reduced row echelon form is:

$$\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

This shows that the three vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ are linearly dependent, but at the same time that the first two of them are linearly independent (compare to Example 7.9, where a similar approach was used for three vectors in $\mathbb{C}^3$). We can conclude that $\mathbf{v}_3$ can be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. This in turns implies that the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ span exactly the same subspace of $\mathbb{R}^3$ as the three vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$. Hence $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of W.

We have already fully answered the question, but suppose that we would like to see explicitly how to express $\mathbf{v}_3$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. To do this, we need to find a solution to the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$ of the form $(c_1, c_2, 1)$. Looking at the reduced row echelon form of $\mathbf{A}$, we see that $(-4, 1, 1)$ is such a solution. Hence $(-4) \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$, which implies that $\mathbf{v}_3 = 4 \cdot \mathbf{v}_1 - \mathbf{v}_2$.

As we saw in the previous example, saying that a subspace is spanned by certain vectors, does not mean that these vectors are linearly independent. The procedure we used in Example 9.36 to find a basis can be generalized. Let us do that in the following theorem:

Theorem 9.37

Let a subspace W of the vector space $\mathbb{F}^n$ be spanned by vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$. Further suppose that the reduced row echelon form of the matrix with columns $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ has pivots precisely in columns $j_1, \dots, j_\rho$. Then $\{\mathbf{u}_{j_1}, \dots, \mathbf{u}_{j_\rho}\}$ is a basis of W.

Proof. First of all, let us denote by $\mathbf{A}$ the matrix with columns $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ and by $\mathbf{B}$ the reduced row echelon form of $\mathbf{A}$. By definition of the reduced row echelon form of a matrix, the columns of $\mathbf{B}$ with column indices $i_1, \dots, i_\rho$ are the first ρ standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_\rho$. In particular, they are linearly independent. We claim that this implies that the columns of $\mathbf{B}$ with column indices $i_1, \dots, i_\rho$ are also linearly independent. Indeed, if $c_{j_1} \cdot \mathbf{u}_{j_1} + \cdots + c_{j_\rho} \cdot \mathbf{u}_{j_\rho} = \mathbf{0}$, then the tuple $(v_1, \dots, v_\ell) \in \mathbb{F}^\ell$ defined by $v_j = c_j$ if $j \in \{j_1, \dots, j_\rho\}$ and $v_j = 0$ otherwise, is a solution to the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$. However, we know that any such solution is also a solution to the homogeneous system with coefficient matrix $\mathbf{B}$. Since we already observed that the columns of $\mathbf{B}$ with column indices $i_1, \ldots, i_\rho$ are linearly independent, we conclude that necessarily $c_{j_1} = 0, \ldots, c_{j_\rho} = 0$. This shows that the vectors $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ are linearly independent.

Now choose any column $\mathbf{u}_j$ of $\mathbf{A}$, where $j \notin \{j_1, \ldots, j_\rho\}$. Again by definition of the reduced row echelon form, the j-th column of $\mathbf{B}$ has zeroes for its last $n - \rho$ entries. Hence it can be expressed as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_\rho$, which are just columns $j_1, \ldots, j_\ell$ of $\mathbf{B}$. This means that the homogeneous system with coefficient matrix $\mathbf{B}$ has a solution $(v_1, \ldots, v_\ell)$ such that $v_j = 1$ and $v_k = 0$ for all $k \notin \{j, j_1, \ldots, j_\rho\}$. Now using that this is also a solution to the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$, we find that the j-th column of $\mathbf{A}$ can be expressed as a linear combination of columns $j_1, \ldots, j_\rho$. This proves that the span of $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ is the same as the span of $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$.

Combining all the above, we conclude that $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ is a basis of W. □

Looking back at Theorem 6.29, we see that in that theorem the solution set to a homogeneous system of linear equations was described exactly as the span of $n - \rho$ vectors. In this case these vectors actually form a basis of the solution set and in particular they are linearly independent. Let us show this now.

|||| Corollary 9.38

Let a homogeneous system of m linear equation in n variables over a field $\mathbb{F}$ be given. Denote the coefficient matrix of this system by $\mathbf{A}$ and assume that this matrix has rank ρ. The $n - \rho$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}$ indicated in Theorem 6.29 form a basis of the solution set to the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$.

Proof. Proof sketch: we use that same notation for the vectors $\mathbf{c}_i$ and the matrix $\hat{\mathbf{A}}$ as in Theorem 6.29. Looking back at the way the vector $\mathbf{v}_i$ was defined in Theorem 6.29, one can see that $\mathbf{v}_i$ has a 1 in the coordinate j, where j satisfies that $\mathbf{c}_i$ is the j-th column in $\hat{\mathbf{A}}$. Similarly, one sees that $\mathbf{v}_i$ has coefficients equal to 0 afterwards, since $\mathbf{c}_i$ contains zeroes only after its ith coefficient. Hence the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}$ is in row echelon form. This implies that the corresponding matrix in reduced row echelon form has pivots in each column. Theorem 9.37 then implies that $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}\}$ is a basis. □

Corollary 9.39

Let V be a vector space over a field $\mathbb{F}$ of finite dimension n with ordered basis β and let $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ be vectors in V. Further suppose that the reduced row echelon form of the matrix with columns $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$ has pivots precisely in columns $j_1, \ldots, j_\rho$. Then a basis of $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell)$ is given by $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$.

Proof. We give a sketch of the proof: first of all, we see from Theorem 9.37 that a basis of the subspace of $\mathbb{F}^n$ generated by $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$ is given by $\{[\mathbf{u}_{j_1}]_\beta, \ldots, [\mathbf{u}_{j_\rho}]_\beta\}$. Now Theorem 9.18 can be used to see that $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ is a basis of $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell)$. □

9.4 Extra: why does any vector space have a basis?

This section is not required reading and can be skipped. It is meant as extra material for a student who has the time and motivation for it.

In the previous sections, we have simply used the fact that any vector space V has a basis. To prove this, we need to study the set $\mathcal{I}(V)$ consisting of all subsets of V whose elements are linearly independent vectors. For example $\emptyset \in \mathcal{I}(V)$, since the empty set contains no vectors and therefore cannot contain linearly dependent vectors. If $V \neq \{\mathbf{0}\}$ any subset of the form $\{\mathbf{v}\}$ is in $\mathcal{I}(V)$ as long as $\mathbf{v} \neq \mathbf{0}$. Intuitively, a basis B of V should be a set containing as many linearly independent vectors as possible. More precisely, this intuition would say that $B \in \mathcal{I}(V)$ and that no set of linearly independent vectors can contain B as a strict subset. This second intuitive property can be reformulated by saying that if $C \in \mathcal{I}(V)$ and $B \subseteq C$, then $B = C$. Such a B is called a maximal element of $\mathcal{I}(V)$.

The above discussion is purely to get an intuitive idea, but the following theorem shows that there is merit in that discussion.

Theorem 9.40

Let B be a maximal element of $\mathcal{I}(V)$. Then B is a basis of V.

Proof. By definition of $\mathcal{I}(V)$, the vectors in B are linearly independent. What needs to be shown is that any vector in V can be written as a linear combination of vectors in B. Suppose that this is not the case. Then there exists $\mathbf{v} \in V$ such that any linear combination of vectors in B is distinct from $\mathbf{v}$. We claim that in this case, the set $B \cup \{\mathbf{v}\}$ consists of linearly independent vectors. To show this, suppose that

$$c_0 \cdot \mathbf{v} + c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0}, \tag{9-5}$$

for some $c_0, c_1, \ldots, c_n \in \mathbb{F}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n \in B$. If $c_0 = 0$, we immediately see that $c_1 = 0, \ldots, c_n = 0$, since the vectors in B are linearly independent. However, c_0 cannot be nonzero, since if it were, equation (9-5) would imply that $\mathbf{v} = -c_0^{-1} \cdot c_1 \cdot \mathbf{v}_1 - \cdots - c_0^{-1} \cdot c_n \cdot \mathbf{v}_n$, contrary to the assumption that $\mathbf{v}$ cannot be written as a linear combination of vectors from B. Hence indeed, the set $B \cup \{\mathbf{v}\}$ consists of linearly independent vectors, just as claimed. Another way of saying this is that $B \cup \{\mathbf{v}\} \in \mathcal{I}(V)$, which in turn implies that B was not a maximal element of $\mathcal{I}(V)$, contrary to the assumption that it was. The contradiction shows that any vector in V can be written as a linear combination of vectors from B. Hence B is a basis. □

This theorem implies that in order to show that any vector space V has a basis, it is enough to show that the set $\mathcal{I}(V)$ always contains a maximal element. This is a direct consequence of a famous lemma called Zorn's lemma. Formulating and proving Zorn's lemma needs tools from foundational mathematics though that are out of scope of these notes.