Linear Algebra 1: Lecture 6

Irena Penev

Winter 2022/2023

Notation: We denote by $\mathbb{N}$ the set of all positive integers, and we denote by $\mathbb{N}_0$ the set of all non-negative integers.

1 Fields

A *field* is an ordered triple $(\mathbb{F}, +, \cdot)$, where $\mathbb{F}$ is a set, and $+$ and $\cdot$ are binary operations on $\mathbb{F}$ (i.e. functions from $\mathbb{F} \times \mathbb{F}$ to $\mathbb{F}$), called *addition* and *multiplication*, satisfying the following axioms:

1. addition and multiplication are associative, that is, for all $a, b, c \in \mathbb{F}$, we have that $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
2. addition and multiplication are commutative, that is, for all $a, b \in \mathbb{F}$, we have that $a + b = b + a$ and $a \cdot b = b \cdot a$;
3. there exist distinct elements $0_{\mathbb{F}}, 1_{\mathbb{F}} \in \mathbb{F}$ such that for all $a \in \mathbb{F}$, $a + 0_{\mathbb{F}} = a$ and $a \cdot 1_{\mathbb{F}} = a$; $0_{\mathbb{F}}$ is called the *additive identity* of $\mathbb{F}$, and $1_{\mathbb{F}}$ is called the *multiplicative identity* of $\mathbb{F}$;
4. for every $a \in \mathbb{F}$, there exists an element in $\mathbb{F}$, denoted by $-a$ and called the *additive inverse* of a, such that $a + (-a) = 0_{\mathbb{F}}$;
5. for all $a \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$, there exists an element in $\mathbb{F}$, denoted by a^{-1} and called the *multiplicative inverse* of a, such that $a \cdot a^{-1} = 1_{\mathbb{F}}$;
6. multiplication is distributive over addition, that is, for all $a, b, c \in \mathbb{F}$, we have that $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

Notation:

- If operations $+$ and $\cdot$ are understood from context, we typically write just "field $\mathbb{F}$" instead of "field $(\mathbb{F}, +, \cdot)$."
- For $a, b \in \mathbb{F}$, we typically write ab instead of $a \cdot b$, and we typically write $a - b$ instead of $a + (-b)$.

- As usual, unless parentheses indicate otherwise, we perform multiplication before performing addition. So, for $a, b, c \in \mathbb{F}$, we write $ab + c$ instead of $(a \cdot b) + c$, and similarly, we write $a + bc$ instead of $a + (b \cdot c)$.

Remarks:

- Axioms 1. and 3. above imply that $(\mathbb{F}, +)$ and $(\mathbb{F}, \cdot)$ are monoids with identity elements $0_{\mathbb{F}}$ and $1_{\mathbb{F}}$, respectively. Proposition 1.1 from Lecture Notes 5 guarantee that $0_{\mathbb{F}}$ and $1_{\mathbb{F}}$ are unique.
  - When there is no danger of confusion, we write 0 and 1 instead of $0_{\mathbb{F}}$ and $1_{\mathbb{F}}$, respectively.
- Axioms 1., 2., 3., and 4. imply that $(\mathbb{F}, +)$ is an abelian group with identity element $0_{\mathbb{F}}$. By Proposition 2.1 from Lecture Notes 5, this implies that each element $a \in \mathbb{F}$ has a **unique** additive inverse $-a$.
- By Proposition 1.2, for any $a, b \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$, we have $ab \neq 0_{\mathbb{F}}$, i.e. $ab \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$. This, together with axioms 1. and 3. implies that $(\mathbb{F} \setminus \{0_{\mathbb{F}}\})$ is a monoid with identity element $1_{\mathbb{F}}$. Next, by Proposition 1.2, and by axioms 2. (commutativity of addition) and 3. ($0_{\mathbb{F}} \neq 1_{\mathbb{F}}$), we have that we have that $a0_{\mathbb{F}} = 0_{\mathbb{F}}a = 0_{\mathbb{F}} \neq 1_{\mathbb{F}}$. This, together with axiom 5. implies that the multiplicative inverse of any element $a \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$ also belongs to $\mathbb{F} \setminus \{0_{\mathbb{F}}\}$. So, $(\mathbb{F} \setminus \{0_{\mathbb{F}}\}, \cdot)$ is an abelian group with identity element $1_{\mathbb{F}}$. By Proposition 2.1 from Lecture Notes 5, it follows that every element $a \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$ has a **unique** multiplicative inverse a^{-1}.
- By axioms 2. and 6., for all $a, b, c \in \mathbb{F}$, we have that $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$, or written in a simplified manner, $(b + c)a = ba + ca$.[1]

Example 1.1. *All the following are fields:*

1. $(\mathbb{Q}, +, \cdot)$*;* *2.* $(\mathbb{R}, +, \cdot)$*;* *3.* $(\mathbb{C}, +, \cdot)$*.*

Note that $(\mathbb{Z}, +, \cdot)$ is **not** a field. This is because elements of $\mathbb{Z} \setminus \{-1, 0, 1\}$ do not have multiplicative inverses. As we shall see, $(\mathbb{Z}_p, +, \cdot)$ is a field for every prime p (see Proposition 1.4). First, we prove some preliminary results about fields.

Proposition 1.2. *Let* $(\mathbb{F}, +, \cdot)$ *be a field. Then all the following hold:*

(a) for all $a \in \mathbb{F}$*,* $0a = 0$*;*

(b) for all $a, b \in \mathbb{F}$*, if* $ab = 0$*, then* $a = 0$ *or* $b = 0$*;*

[1] Indeed, for $a, b, c \in \mathbb{F}$, we have that $(b + c)a \overset{\text{ax. 2.}}{=} a(b + c) \overset{\text{ax. 6.}}{=} ab + ac \overset{\text{ax. 2.}}{=} ba + ca$.

(c) for all $a \in \mathbb{F}$, $(-1)a = -a$.[2]

Proof. We first prove a. Fix $a \in \mathbb{F}$. First, note that

$$0a \overset{(*)}{=} (0+0)a \overset{(**)}{=} 0a + 0a,$$

where (*) follows from the fact that $0 + 0 = 0$ (because 0 is the additive inverse of the field), and (**) follows from axiom 3. of the definition of a field. We have now established that $0a = 0a + 0a$, and it follows that

$$\begin{aligned}
0 &= (-(0a)) + 0a && \text{because } -(0a) \text{ is the additive inverse of } 0a \\
&= (-(0a)) + (0a + 0a) && \text{because } 0a = 0a + 0a \\
&= ((-(0a)) + 0a) + 0a && \text{because } + \text{ is associative} \\
&= 0 + 0a && \text{because } -(0a) \text{ is the additive inverse of } 0a \\
&= 0a && \text{because 0 is the additive identity of the field.}
\end{aligned}$$

Thus, $0a = 0$. This proves (a).

Next, we prove (b). Fix $a, b \in \mathbb{F}$ such that $ab = 0$. We may assume that $b \neq 0$, for otherwise we are done. But now b has a multiplicative inverse b^{-1}, and we compute:

$$\begin{aligned}
a &= a \cdot 1 \\
&= a(bb^{-1}) \\
&= (ab)b^{-1} && \text{because multiplication is associative} \\
&= 0b^{-1} && \text{because } ab = 0 \\
&= 0 && \text{by (a).}
\end{aligned}$$

Thus, (b) holds.

It remains to prove (c). Fix $a \in \mathbb{F}$. First, we have that

$$0 \overset{(a)}{=} 0a = (1-1)a = 1a + (-1)a = a + (-1)a,$$

[2]This statement may require some clarification. Here, $-a$ is the additive inverse of a. On the other hand, $(-1)a$ is the product of -1 (the additive inverse of the multiplicative identity) and a. So, $-a$ is not simply shorthand for $(-1)a$. The two quantities are indeed equal, but this requires proof!

and consequently,

$$\begin{aligned} -a &= -a + 0 \\ &= -a + (a + (-1)a) && \text{because } 0 = a + (-1)a \text{ (by the above)} \\ &= (-a + a) + (-1)a && \text{because addition is associative} \\ &= 0 + (-1)a \\ &= (-1)a. \end{aligned}$$

This proves (c). □

1.1 Finite fields

Recall from Lecture Notes 0 that for all positive integers n and integers a, we defined

$$[a]_n := \{x \in \mathbb{Z} \mid x \equiv a \pmod{n}\}.$$

For instance, we have

- $[0]_2 = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$;
- $[1]_2 = \{\ldots, -3, -1, 1, 3, 5, \ldots\}$;
- $[0]_3 = \{\ldots, -6, -3, 0, 3, 6, \ldots\}$;
- $[1]_3 = \{\ldots, -5, -2, 1, 4, 7, \ldots\}$;
- $[2]_3 = \{\ldots, -4, -1, 2, 5, 8, \ldots\}$.

Moreover, we have that for all $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$, $[a]_n = [b]_n$ if and only if $a \equiv b \pmod{n}$.[3] For $n \in \mathbb{N}$, we defined

$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$

Moreover, when n is understood from context, we typically write $0, 1, \ldots, n-1$ instead of $[0]_n, [1]_n, \ldots, [n-1]_n$, respectively.

For $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$, we defined

$$[a]_n + [b]_n = [a+b]_n \quad \text{and} \quad [a]_n[b]_n = [ab]_n.$$

This is well-defined by Proposition 1.3 from Lecture Notes 0.

Proposition 1.3. *Let $n \in \mathbb{N}$. Then*

[3] This is because equivalence modulo n is a transitive relation (see Proposition 1.2(3) from Lecture Notes 0).

- $(\mathbb{Z}_n, +)$ *is an abelian group;*
- *the identity element of the abelian group* $(\mathbb{Z}_n, +)$ *is* $[0]_n$*;*
- *for all* $a \in \mathbb{Z}_n$*, the inverse element of* $[a]_n$ *in the abelian group* $(\mathbb{Z}_n, +)$ *is* $[-a]_n = [n-a]_n$*.*

Proof. This is immediate from the relevant definitions. □

Remark: Note that we can write, for example, that $-1 = 2$ in $\mathbb{Z}_3$. This simply means that, in $\mathbb{Z}_3$, the additive inverse of 1 is equal to 2, which is true, since $1 + 2 = 0$ (in $\mathbb{Z}_3$).

For $n \in \mathbb{N}$ and $a \in \mathbb{Z}_p$, we define:

- $a^0 = 1$;[4]
- for all $m \in \mathbb{N}$, $a^m := \underbrace{a \cdot \cdots \cdot a}_{m \text{ times}}$.

We now recall Fermat's Little Theorem, proven in Lecture Notes 0.

Fermat's Little Theorem. *If* $p \in \mathbb{N}$ *is a prime number and* $a \in \mathbb{Z}_p \setminus \{0\}$*, then* $a^{p-1} = 1$.[5]

Proposition 1.4. *Let* $p \in \mathbb{N}$*. Then* $(\mathbb{Z}_p, +, \cdot)$ *is a field if and only if* p *is a prime number.*

Proof. It suffices to prove the following:

(i) if p is a prime number, then $(\mathbb{Z}_p, +, \cdot)$ is a field;

(ii) if p is **not** a prime number, then $(\mathbb{Z}_p, +, \cdot)$ is **not** a field.

We first prove (ii). Suppose that p is not a prime number. If $p = 1$, then $\mathbb{Z}_p$ has exactly one element, and so $(\mathbb{Z}_p, +, \cdot)$ does not satisfy axiom 3. from the definition of a field.[6] Suppose now that $p \geq 2$. Since p is not prime, there exist integers $a, b \in \mathbb{N}$ such that $a, b \geq 2$ and $p = ab$. Then $a, b \leq p - 1$, and in particular, a and b are not multiples of p, i.e. $[a]_p \neq [0]_p$ and $[b]_p \neq 0$. But on the other hand, $[a]_p[b]_p = [ab]_p = [p]_p = [0]_p$. So, $(\mathbb{Z}_p, +, \cdot)$ does not satisfy the statement of Proposition 1.2(b), and consequently, $(\mathbb{Z}_p, +, \cdot)$ is not a field. This proves (ii).

It remains to prove (i). Suppose that p is a prime number. Then the additive and multiplicative identities of $(\mathbb{Z}_p, +, \cdot)$ are $[0]_p$ and $[1]_p$, respectively. The additive inverse of any $[a]_p$, where $a \in \mathbb{Z}$, is $[-a]_p = [p-a]_p$. Furthermore,

[4] Here, we mean $1 = [1]_n$.

[5] As usual, 0 stands for $[0]_p$, and 1 stands for $[1]_p$.

[6] By axiom 3., the additive and multiplicative identity of a field cannot be equal. So, any field has at least two elements.

for any $a \in \mathbb{Z}_p \setminus \{[0]_p\}$, Fermat's Little Theorem guarantees that a^{p-2} is the multiplicative inverse of a. The other axioms from the definition of a field are obviously satisfied, and so $(\mathbb{Z}_p, +, \cdot)$ is indeed a field. □

Theorem 1.5. *Let $n \geq 2$ be an integer. Then there exists a field of size n if and only if n is the power of a prime.*[7] *Moreover, if n is a the power of a prime, then up to "isomorphism" (i.e. up to renaming the operations and elements of the field), there is exactly one field of size n, and it is denoted by $\mathbb{F}_n$.*[8]

Proof. Omitted. □

Remark: For a prime number p, we have that $\mathbb{F}_p = \mathbb{Z}_p$. However, if $n = p^m$, where p is a prime number and $m \geq 2$ is an integer, then $\mathbb{F}_n \neq \mathbb{Z}_n$ (this is because $\mathbb{F}_n$ is a field, but by Proposition 1.4, $\mathbb{Z}_n$ is not a field).

2 Vector spaces

Let $\mathbb{F}$ be a field with additive identity 0 and multiplicative identity 1. In what follows, we shall refer to elements of $\mathbb{F}$ as *scalars*. A *vector space* (or *linear space*) over the field $\mathbb{F}$ is a set V, together with a binary operation $+$ on V (called *vector addition*) and an operation $\cdot : \mathbb{F} \times V \to V$ (called *scalar multiplication*), satisfying the following axioms:

1. $(V, +)$ is an abelian group; the identity element of $(V, +)$ is denoted by $\mathbf{0}$ ("zero vector"), and for any vector $\mathbf{v} \in V$, the inverse of $\mathbf{v}$ in $(V, +)$ is dented by $-\mathbf{v}$;
2. for all vectors $\mathbf{v} \in V$, we have $1\mathbf{v} = \mathbf{v}$;
3. for all vectors $\mathbf{v} \in V$ and scalars $\alpha, \beta \in \mathbb{F}$, we have $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$;
4. for all vectors $\mathbf{v} \in V$ and scalars $\alpha, \beta \in \mathbb{F}$, we have $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;
5. for all vectors $\mathbf{u}, \mathbf{v} \in V$ and scalars $\alpha \in \mathbb{F}$, we have $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$.

Example 2.1. *Let $\mathbb{F}$ be a field. Then the following are vector spaces over $\mathbb{F}$ (in each case, vector addition and scalar multiplication are defined in the natural way):*

1. *$\mathbb{F}^n$;*
2. *$\mathbb{F}^{n \times m}$;*
3. *the set of all functions from $\mathbb{F}$ to $\mathbb{F}$;*

[7] "n is the power of a prime" means that there exists some prime number p and a positive integer m such that $n = p^m$.

[8] Technically, the field is $(\mathbb{F}_n, +, \cdot)$, but we typically just say $\mathbb{F}_p$.

4. *the set $\mathbb{P}_{\mathbb{F}}$ of all polynomials with coefficients in $\mathbb{F}$;*[9]

5. *for a positive integer n, the set $\mathbb{P}^n_{\mathbb{F}}$ of all polynomials of degree at most n.*[10]

Note that each of the cases above, elements of our vector space are considered vectors (even if they do not "look like" vectors, i.e. even if they are matrices, functions, or polynomials).

If you have studied calculus, here is another example.

Example 2.2. *The following are vector spaces over $\mathbb{R}$ (with vector addition and scalar multiplication defined in the usual way):*

1. *the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$;*

2. *the set of differentiable functions from $\mathbb{R}$ to $\mathbb{R}$.*

Proposition 2.3. *Let V be a vector space over a field $\mathbb{F}$. Then all the following hold:*

(a) for all $\mathbf{v} \in V$, $0\mathbf{v} = \mathbf{0}$;[11]

(b) for all $\alpha \in \mathbb{F}$, $\alpha\mathbf{0} = \mathbf{0}$;

(c) for all $\mathbf{v} \in V$ and $\alpha \in \mathbb{F}$, if $\alpha\mathbf{v} = \mathbf{0}$, then $\alpha = 0$ or $\mathbf{v} = \mathbf{0}$;

(d) for all $\mathbf{v} \in V$, $(-1)\mathbf{v} = -\mathbf{v}$.[12]

Proof. The proof is similar to that of Proposition 1.2. We prove (a). The rest is left as an exercise. Fix $\mathbf{v} \in V$. Then $0\mathbf{v} = (0+0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$, and consequently,

$$\begin{aligned}
\mathbf{0} &= -(0\mathbf{v}) + 0\mathbf{v} \\
&= -(0\mathbf{v}) + 0\mathbf{v} + 0\mathbf{v} && \text{because } 0\mathbf{v} = 0\mathbf{v} + 0\mathbf{v} \\
&= \mathbf{0} + 0\mathbf{v} && \text{because } -(0\mathbf{v}) + 0\mathbf{v} = \mathbf{0} \\
&= 0\mathbf{v}.
\end{aligned}$$

This proves (a). □

[9] $\mathbb{P}_{\mathbb{F}}$ is not entirely standard notation. However, this vector space will be a reasonably frequent example in the remainder of the course, and so we will need a convenient way to refer to it.

[10] Again, $\mathbb{P}^n_{\mathbb{F}}$ is not entirely standard notation.

[11] Here, 0 is the zero of the field $\mathbb{F}$, and $\overline{0}$ is the zero vector in V.

[12] Here, -1 is the additive inverse of the multiplicative identity of the field $\mathbb{F}$, and in particular, -1 is a scalar. So, $(-1)\mathbf{v}$ is the product of the scalar -1 and the vector $\mathbf{v}$. On the other hand, $-\mathbf{v}$ is the additive inverse of the vector $\mathbf{v}$.

2.1 Vector subspaces

Let V be a vector space over a field $\mathbb{F}$. A *vector subspace* (or *linear subspace* or simply *subspace*) of V is a set $U \subset V$ such that U is itself a vector space over $\mathbb{F}$, when equipped with the vector addition and scalar multiplication operations inherited from V.[13]

Notation: If V is a vector space over a field $\mathbb{F}$, and U is a subspace of V, then we write $U \leq V$.

Remark: It is obvious that the subspace relation is transitive. More precisely, for a vector space V over a field $\mathbb{F}$, if U is a subspace of V, and W is a subspace of U, then W is a subspace of V.[14]

Example 2.4. *Let V be a vector space over a field $\mathbb{F}$. Then V is a subspace of itself, and $\{\mathbf{0}\}$ is a subspace of V.*

Example 2.5. *Let n be a positive integer, and let $\mathbb{F}$ be a field. Then $\mathbb{P}^n_{\mathbb{F}}$ is a subspace of $\mathbb{P}_{\mathbb{F}}$.*

If you have studied calculus, here is another example.

Example 2.6. *The set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ forms a subspace of the set of all functions from $\mathbb{R}$ to $\mathbb{R}$. Similarly, the set of all differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ forms a subspace of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$.*

Theorem 2.7. *Let V be a vector space over a field $\mathbb{F}$, and let $U \subseteq V$. Then U is a subspace of V if and only if the following three conditions are satisfied:*

(i) $\mathbf{0} \in U$*;*

(ii) *U is closed under vector addition, that is, for all $\mathbf{u}, \mathbf{v} \in U$, we have that $\mathbf{u} + \mathbf{v} \in U$;*

(iii) *U is closed under scalar multiplication, that is, for all $\mathbf{u} \in U$ and $\alpha \in \mathbb{F}$, we have that $\alpha\mathbf{u} \in U$.*

Proof. Suppose first that (i), (ii), and (iii) are satisfied. By (ii), the restriction of $+$ to $U \times U$ (denoted $+ \restriction (U \times U)$, or just $+$ for simplicity) is a binary operation on U,[15] and by (iii), the restriction of $\cdot$ to $\mathbb{F} \times U$ (denoted by $\cdot \restriction (\mathbb{F} \times U)$, or just $\cdot$ for simplicity) is indeed a function from $\mathbb{F} \times U$ to U. So, U is equipped with both the vector addition operation and the scalar multiplication operation. Moreover, (i) guarantees that U contains the zero vector of V, and by Proposition 2.3(d) and (ii), we have that for all $\mathbf{v} \in U$, we have that $-\mathbf{v} = (-1)\mathbf{v} \in U$. We now see that axiom 1. from the definition

[13] Note that the field $\mathbb{F}$ must remain the same for U as for V!

[14] So, if $W \leq U$ and $W \leq V$, then $W \leq V$.

[15] In other words, we have that $+ \restriction (U \times U) : U \times U \to U$.

of a vector space over $\mathbb{F}$ is satisfied for U. The remaining axioms from the definition of a vector space over $\mathbb{F}$ are satisfied for U because they are satisfied for the vector space V, and the vector addition and scalar multiplication for U are inherited from the respective operations for V. So, U is indeed a vector space over $\mathbb{F}$ under the vector addition and scalar multiplication operations inherited from $\mathbb{F}$. It follows that U is a subspace of V.

Suppose now that U is a subspace of V. Since the vector addition and scalar multiplication operations of the vector space U are inherited from the ones for V, we see that (ii) and (iii) hold. Moreover, since U is a vector space, we know that it contains the zero vector, call it $\mathbf{0}_U$.[16] We must show that $\mathbf{0}_U = \mathbf{0}$.[17] Since $\mathbf{0}_U$ is the identity element of $(U, +)$, we see that $\mathbf{0}_U + \mathbf{0}_U = \mathbf{0}_U$. Since $\mathbf{0}_U \in V$ and $\mathbf{0}$ is the identity element of V, we see that $\mathbf{0}_U + \mathbf{0} = \mathbf{0}_U$. So, $\mathbf{0}_U + \mathbf{0} = \mathbf{0}_U + \mathbf{0}_U$. By now adding $-\mathbf{0}_U$ to both sides of the equation,[18] and we obtain $\mathbf{0} = \mathbf{0}_U$. So, (i) holds. □

2.2 Linear combinations and linear span

Suppose that V is a vector space over a field $\mathbb{F}$. Given vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$, we say that a vector $\mathbf{v} \in V$ is a *linear combination* of $\mathbf{u}_1, \ldots, \mathbf{u}_k$ if there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that

$$\mathbf{v} \quad = \quad \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k$$

The *linear span* (or simply *span*) of the set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, denoted by $\mathrm{Span}(\{\mathbf{u}_1, \ldots, \mathbf{u}_k\})$ or $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, is the set of all linear combinations of $\mathbf{u}_1, \ldots, \mathbf{u}_k$, i.e.

$$\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) \quad = \quad \{\alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F}\}.$$

As a convention, we define $\mathrm{Span}(\emptyset) := \{\mathbf{0}\}$. Here, the idea is that the "empty sum" is equal to the zero vector.[19]

Given a vector space V over a field $\mathbb{F}$, and given vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$, we say that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a *spanning set* of V, or that that the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ *spans* V, or that vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ *span* V, provided that $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. (Note that $\emptyset$ is a spanning set of the trivial vector space $\{\mathbf{0}\}$ over a field $\mathbb{F}$.)

[16] Since U is a subspace of V, we know, in particular, that $(U, +)$ is an abelian group, and consequently, it contains a (unique) identity element. We call this identity element $\mathbf{0}_U$.

[17] Here, $\mathbf{0}$ is the zero vector in V, and i.e. the identity element of the abelian group $(V, +)$. Since $(U, +)$ is an abelian group, it must have an identity element, and we call this identity element $\mathbf{0}_U$. However, could it be that $\mathbf{0}_U \neq \mathbf{0}$, so that (i) potentially fails? We show that this cannot happen.

[18] Here, $-\mathbf{0}_U$ is the additive inverse of the vector $\mathbf{0}_U$ in V.

[19] Here, the "empty sum" would be the sum $\alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k$, where $k = 0$ (and so we don't actually have any $\mathbf{u}_i$'s or α_i's).

Example 2.8. *Consider vectors* $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ *and* $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ *in* $\mathbb{R}^3$. *Then* $Span(\mathbf{e}_1, \mathbf{e}_2) = \{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R}\}$. *So,* $Span(\mathbf{e}_1, \mathbf{e}_2)$ *is the* x_1x_2*-plane in the Euclidean space* $\mathbb{R}^3$.

Example 2.9. *Consider the polynomials* $1, x, x^2$ *in* $\mathbb{P}_{\mathbb{R}}$. *Then* $Span(1, x, x^2) = \{a_2x^2 + a_1x + a_0 \mid a_0, a_1, a_2 \in \mathbb{R}\} = \mathbb{P}^2_{\mathbb{R}}$.

Proposition 2.10. *Let* V *be a vector space over a field* $\mathbb{F}$, *and let* $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$. *Then* $Span(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ *is a subspace of* V. *Moreover,* $\mathbf{u}_1, \ldots, \mathbf{u}_k \in Span(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

Proof. To see that $\mathbf{u}_1, \ldots, \mathbf{u}_k \in Span(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, we simply observe that for all $i \in \{1, \ldots, k\}$, we have that $\mathbf{u}_i = 0\mathbf{u}_1 + \cdots + 0\mathbf{u}_{i-1} + 1\mathbf{u}_i + 0\mathbf{u}_{i+1} + \cdots + 0\mathbf{u}_k$, and so $\mathbf{u}_i \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

It remains to show that $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is a subspace of V. For this, it is enough to show that $Span\{\mathbf{u}_1, \ldots, \mathbf{u}_k)$ satisfies (i), (ii) and (iii) from Theorem 2.7, that is, that all the following hold:

(i) $\mathbf{0} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$;

(ii) $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$ is closed under vector addition, that is, for all $\mathbf{v}_1, \mathbf{v}_2 \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$, we have that $\mathbf{v}_1 + \mathbf{v}_2 \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$;

(iii) $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$ is closed under scalar multiplication, that is, for all $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$ and $\alpha \in \mathbb{F}$, we have that $\alpha\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$.

We first prove (i). First, note that $0\mathbf{u}_1 + \cdots + 0\mathbf{u}_k \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$. But clearly, $0\mathbf{u}_1 + \cdots + 0\mathbf{u}_k = \mathbf{0}$, and so $\mathbf{0} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$. So, (i) holds.

Next, we prove (ii). Fix $\mathbf{v}_1, \mathbf{v}_2 \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{0}_k)$. Then there exist scalars $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{F}$ such that $\mathbf{v}_1 = \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k$ and $\mathbf{v}_2 = \beta_1\mathbf{u}_1 + \cdots + \beta_k\mathbf{u}_k$. But now

$$\begin{aligned} \mathbf{v}_1 + \mathbf{v}_2 &= (\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k) + (\beta_1\mathbf{u}_1 + \cdots + \beta_k\mathbf{u}_k) \\ &= (\alpha_1 + \beta_1)\mathbf{u}_1 + \cdots + (\alpha_k + \beta_k)\mathbf{u}_k, \end{aligned}$$

and we deduce that $\mathbf{v}_1 + \mathbf{v}_2 \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. This proves (ii).

It remains to prove (iii). Fix $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ and $\alpha \in \mathbb{F}$. Since $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, we see that there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k$. But now

$$\alpha\mathbf{v} = \alpha(\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k) = (\alpha\alpha_1)\mathbf{u}_1 + (\alpha\alpha_k)\mathbf{u}_k,$$

and so $\alpha\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. This proves (iii), and we are done. □

Note that Proposition 2.10 remains true even if $k = 0$, since $\text{Span}(\emptyset) = \{\mathbf{0}\}$ is a subspace of V.

Proposition 2.11. *Let V be a vector space over a field $\mathbb{F}$, and let $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$. Then Span$(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is the intersection of all the subspaces of V that contain $\mathbf{u}_1, \ldots, \mathbf{u}_k$.*[20]

Proof. By Proposition 2.10, $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is itself a subspace of V that contains $\mathbf{u}_1, \ldots, \mathbf{u}_k$. So, the intersection of all subspaces of V that contain $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is a subset of $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.[21]

It remains to show that $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is a subset of the intersection of all subspaces of V that contain $\mathbf{u}_1, \ldots, \mathbf{u}_k$. For this, it is enough to show that $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is a subset of every subspace of V that contains $\mathbf{u}_1, \ldots, \mathbf{u}_k$.[22] So, fix any subset U of V that contains $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Now, fix any $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. Then there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k$. Since U is a subspace of V, it satisfies (ii) and (iii) from Theorem 2.7. Since $\mathbf{u}_1, \ldots, \mathbf{u}_k \in U$, (iii) guarantees that $\alpha_1\mathbf{u}_1, \ldots, \alpha_k\mathbf{u}_k \in U$; but then (ii) guarantees that $\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k \in U$, i.e. $\mathbf{v} \in U$. So, $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) \subseteq U$. □

Remark: In some texts, for a vector space V over a field $\mathbb{F}$, and for vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$, the linear span (or simply span) of $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is defined to be the intersection of all subspaces of V that contain $\mathbf{u}_1, \ldots, \mathbf{u}_k$. By Proposition 2.11, this definition is equivalent to the one that we gave at the beginning of the section.

[20] Note that at least one such subspace exists, namely the subspace V. We also note that the proposition remains true for $k = 0$. In that case, it simply states that $\text{Span}(\emptyset)$ is equal to the intersection of all subspaces of V. But this is true because $\text{Span}(\emptyset) = \{\mathbf{0}\}$, and likewise, the intersection of all subspaces of V is equal to $\{\mathbf{0}\}$ (indeed, by Theorem 2.7, every subspace of V contains $\mathbf{0}$; on the other hand, $\{\mathbf{0}\}$ is a subspace of V).

[21] Here, we are using the fact that for all non-empty collections $\mathcal{I}$ of sets, and every set $A \in \mathcal{I}$, we have that $\bigcap_{X \in \mathcal{I}} X \subseteq A$.

[22] Here, we are using the fact that if $\mathcal{I}$ is a non-empty collection of sets and A is a subset of all elements of $\mathcal{I}$, then $A \subseteq \bigcap_{X \in \mathcal{I}} X$.