Vector spaces

We first define the notion of a field, examples of which are the fields of real numbers and the field of complex number.

Definition 1. A field is a triple $(F, +, \cdot)$ consisting of a set F and two maps $+\colon F \times F \to F$ and $\cdot\colon F \times F \to F$ that satisfy the following axioms.

(A1) For all $a, b, c \in F$, $a + (b + c) = (a + b) + c$.
(A2) There exists an element $0 \in F$ such that for all $a \in F$, $a + 0 = a = 0 + a$.
(A3) For every $a \in F$, there exists $b \in F$ such that $a + b = 0 = b + a$.
(A4) For all $a, b \in F$, $a + b = b + a$.
(P1) For all $a, b, c \in F$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
(P2) There exists an element $1 \in F \smallsetminus \{0\}$ such that for all $a \in F$, $a \cdot 1 = a = 1 \cdot a$.
(P3) For every $a \in F \smallsetminus \{0\}$, there exists $b \in F$ such that $a \cdot b = 1 = b \cdot a$.
(P4) For all $a, b \in F$, $a \cdot b = b \cdot a$.
(D) For all $a, b, c \in F$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

Examples of fields are the fields of rational numbers $(\mathbb{Q}, +, \cdot)$, real numbers $(\mathbb{R}, +, \cdot)$, and complex numbers $(\mathbb{C}, +, \cdot)$ with the sum "$+$" and multiplication "$\cdot$" defined as usual. In the following, we will employ the standard practice to abuse notation and simply write F to indicate $(F, +, \cdot)$. We also often suppress $\cdot$ and write ab instead of $a \cdot b$.

Remark 2. More generally, a triple $(F, +, \cdot)$ as in Definition 1 which satisfies the axioms (A1)–(A4), (P1)–(P3), and (D), but not necessarily axiom (P4), is called a skewfield. An example of a skewfield that is not a field is Hamilton's skewfield of quaternions $(\mathbb{H}, +, \cdot)$, where

$$\mathbb{H} = \{a + ib + jc + kd \mid a, b, c, d \in \mathbb{R}\}$$

with the addition $+$ and scalar multiplication $\cdot$ defined by

$$\begin{aligned}
&(a + ib + jc + kd) + (a' + ib' + jc' + kd') \\
&\qquad = (a + a') + i(b + b') + j(c + c') + k(d + d') \\
&(a + ib + jc + kd) \cdot (a' + ib' + jc' + kd') \\
&\qquad = (aa' - bb' - cc' - dd') + i(ab' + a'b + cd' - dc') \\
&\qquad\quad + j(ac' + a'c + db' - bd') + k(ad' + a'd + bc' - b'c).
\end{aligned}$$

In the following, we will not use axiom (P4), so all definitions and theorems hold for skewfields as well as for fields.

We note that the zero element $0 \in F$ which exist by axiom (A2) is unique. Indeed, if both 0 and $0'$ satisfy (A2), then

$$0' = 0 + 0' = 0.$$

Moreover, for a given $a \in F$, the element $b \in F$ such that $a + b = 0 = b + a$ which exists by (A3) is unique. Indeed, if both b and b' satisfy (A3), then

$$b = b + 0 = b + (a + b') = (b + a) + b' = 0 + b' = b'.$$

We write $-a$ instead of b for this element. Similarly, the element $1 \in F$ which exists by axiom (P2) is unique, and for $a \in F \smallsetminus \{0\}$, the element $b \in F$ such that $a \cdot b = 1 = b \cdot a$ which exists by (P3) is unique. We write a^{-1} for this element.

`www.math.ku.dk/~larsh/teaching/F2015_LA`

Definition 3. Let F be a field. A right F-vector space is a triple $(V, +, \cdot\,)$ of a set V and two maps $+\colon V \times V \to V$ and $\cdot\colon V \times F \to V$ such that $(V, +)$ satisfies the axioms (A1)–(A4) and such that the following additional axioms hold.

(V1) For all $x \in V$ and $a, b \in F$, $(x \cdot a) \cdot b = x \cdot (a \cdot b)$.
(V2) For all $x, y \in V$ and $a \in F$, $(x + y) \cdot a = (x \cdot a) + (y \cdot a)$.
(V3) For all $x \in V$ and $a, b \in F$, $x \cdot (a + b) = (x \cdot a) + (x \cdot b)$.
(V4) For all $x \in V$, $x \cdot 1 = x$.

The notion of a left F-vector space, in which scalars multiply from the left, is defined analogously.

Example 4. (1) The field $(F, +, \cdot\,)$ both is a right F-vector space and a left F-vector space. It is a 1-dimensional right F-vector space; see Definition 14 below for the definition of dimension.

(2) The set $M_{n,1}(F)$ of $n \times 1$-matrices with entries in F admits a right F-vector space structure with sum $+\colon M_{n,1}(F) \times M_{n,1}(F) \to M_{n,1}(F)$ defined to be matrix addition and scalar multiplication $\cdot\colon M_{n,1}(F) \times F \to M_{n,1}(F)$ defined to be matrix multiplication. Here we identify $M_{1,1}(F) = F$. We write F^n for this right F-vector space. Its dimension is n.

(3) The set $\mathbb{C}$ of complex numbers admits a structure of right $\mathbb{R}$-vector space with sum and scalar multiplication, respectively, defined by

$$(x_1 + ix_2) + (y_1 + iy_2) = (x_1 + y_1) + i(x_2 + y_2),$$
$$(x_1 + ix_2) \cdot a = x_1 a + ix_2 a.$$

This right $\mathbb{R}$-vector space is 2-dimensional.

(4) The set $\mathbb{C}$ of complex numbers also admits a structure of right $\mathbb{Q}$-vector space with sum and scalar multiplication given by the same formulas as in (3), but where we now only allow $a \in \mathbb{Q}$. The dimension of the resulting right $\mathbb{Q}$-vector space is equal to the cardinality of the real numbers.

We will only consider right vector spaces. We abuse notation and write V to indicate the F-vector space $(V, +, \cdot\,)$, and we abbreviate $x \cdot a$ by xa.

We will say, synonymously, that a map $x\colon I \to X$ from a set I to a set X is a family of elements in X indexed by I and write it $(x_i)_{i \in I}$ with $x_i = x(i)$. We call the set I the index set of the family $(x_i)_{i \in I}$.

Example 5. (1) For every set X, there is a unique family of elements in X indexed by the empty set. We call it the empty family and write it $(\ )$.

(2) For every set X, the identity map $\mathrm{id}_X\colon X \to X$ is a family of elements in X indexed by X. We call it the identity family and write it $(x)_{x \in X}$.

(3) A family of elements in X indexed by the set $I = \{1, 2, \dots, n\}$ is also called an n-tuple of elements in X and written $(x_1, x_2, \dots, x_n)$ instead of $(x_i)_{i \in \{1,2,\dots,n\}}$.

The families (x) and (x, x) of elements in X are different, since their indexing sets are different. By contrast, the subsets $\{x\}$ and $\{x, x\}$ of X are equal.

If $(a_i)_{i \in I}$ a family of scalars in a field F, then we define its support to be

$$\mathrm{supp}(a) = \{i \in I \mid a_i \neq 0\} \subset I.$$

We now let V be an F-vector space and consider a family $(v_i)_{i \in I}$ of vectors in V and a family $(a_i)_{i \in I}$ of scalars in F indexed by the same set I. We assume that the family of scalars $(a_i)_{i\in I}$ has finite support. In this situation, we define

$$\sum_{i\in I} v_i a_i = \sum_{i\in \mathrm{supp}(a)} v_i a_i \in V$$

and call it a linear combination of the family $(v_i)_{i\in I}$. The following three properties of a family of vectors in a vector space are fundamental.

Definition 6. Let F be a field, let V an F-vector space, and let $(v_i)_{i\in I}$ be a family of vectors in V.

(1) The family of vectors $(v_i)_{i\in I}$ is linearly independent if the only family of scalars $(a_i)_{i\in I}$ of finite support such that

$$\sum_{i\in I} v_i a_i = 0$$

is the family $(a_i)_{i\in I}$ with $a_i = 0$ for all $i \in I$.

(2) The family of vectors $(v_i)_{i\in I}$ generates V if for every $v \in V$, there exists a family of scalars $(a_i)_{i\in I}$ of finite support such that

$$\sum_{i\in I} v_i a_i = v.$$

(3) The family $(v_i)_{i\in I}$ is a basis of V if it is both linearly independent and generates V.

Example 7. (1) The empty family $(\ )$ is linearly independent. Indeed, for the empty family, the requirement necessary to be linearly independent is vacuous.

(2) The identity family $(v)_{v\in V}$ generates V. For given $w \in V$, the family $(a_v)_{v\in V}$, where a_v is 1 if $v = w$ and 0 otherwise, is of finite support and $\sum_{v\in V} v a_v = w$.

(3) The standard basis of F^n is the family of vectors $(e_1, \dots, e_n)$, where

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

It is a basis of F^n, since we have

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = e_1 x_1 + e_2 x_2 + \cdots + e_n x_n,$$

and since this expression of the left-hand side as a linear combination of the standard basis is unique.

(4) A family of vectors $(v_i)_{i\in I}$ for which there exists $h \in I$ with $v_h = 0$ is linearly dependent. Indeed, the family of scalars $(a_i)_{i\in I}$ with a_i equal to 1 for $i = h$ and 0 otherwise has finite support and $\sum_{i\in I} v_i a_i = 0$.

Proposition 8. *Let $(v_i)_{i\in I}$ be a basis of an F-vector space V. For every vector $v \in V$, there exists a unique family of scalars $(a_i)_{i\in I}$ of finite support such that*

$$\sum_{i\in I} v_i a_i = v.$$

Proof. Since $(v_i)_{i\in I}$ generates V, there exists a family of scalars $(a_i)_{i\in I}$ of finite support such that $\sum_{i\in I} v_i a_i = v$. To prove that the family of scalars $(a_i)_{i\in I}$ is unique with this property, we suppose that also $(b_i)_{i\in I}$ is a family of scalars of finite support such that $\sum_{i\in I} v_i b_i = v$. The family of scalars $(a_i - b_i)_{i\in I}$ again is of finite support, and moreover,

$$\sum_{i\in I} v_i(a_i - b_i) = (\sum_{i\in I} v_i a_i) - (\sum_{i\in I} v_i b_i) = v - v = 0.$$

Since $(v_i)_{i\in I}$ is linearly independent, we find that $a_i - b_i = 0$ for all $i \in I$, proving the desired uniqueness statement. □

Definition 9. Let $(v_i)_{i\in I}$ be a basis of an F-vector space V. The coordinates of a vector $v \in V$ with respect to the basis $(v_i)_{i\in I}$ is the unique family of scalars of finite support $(a_i)_{i\in I}$ with the property that

$$\sum_{i\in I} v_i a_i = v.$$

Example 10. In $V = F^2$, the coordinates of the vector

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with respect to the standard basis (e_1, e_2) are (x_1, x_2). Indeed, we have

$$x = e_1 x_1 + e_2 x_2.$$

Similarly, the coordinates of x with respect to the basis (v_1, v_2) with

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are $(-x_1 + x_2, 2x_1 - x_2)$. Indeed, we have

$$x = v_1(-x_1 + x_2) + v_2(2x_1 - x_2).$$

Given a family $(x_i)_{i\in I}$ of elements in a set X and a subset $J \subset I$ of the index set, we say that the family $(x_i)_{i\in J}$ is a subfamily of the family $(x_i)_{i\in I}$. The following result is the main theorem of linear algebra.

Theorem 11. *Let F be a field and let V be an F-vector space. Suppose that $(v_i)_{i\in I}$ is a family of vectors that generates V and that $(v_i)_{i\in K}$ is a linearly independent subfamily. In this situation, there exists $K \subset J \subset I$ such that the family $(v_i)_{i\in J}$ is a basis of V.*

Proof. Let S be the set that consists of all subsets $K \subset M \subset I$ such that the family $(v_i)_{i\in M}$ is linearly independent. The inclusion relation $M \subset M'$ is a partial order on the set S. We will use Zorn's lemma, which states that S has a maximal element with respect to the inclusion relation, provided that the following hold:

(i) The set S is non-empty.
(ii) Every subset $T \subset S$ which is totally ordered with respect to the inclusion relation has an upper bound in S.

First, since $K \in S$, we conclude that (i) holds. Second, given totally ordered subset $T \subset S$, we set $M_T = \bigcup_{M\in T} M$. The family $(v_i)_{i\in M_T}$ again is linearly independent, so we have $M_T \in S$. Moreover, for every $M \in T$, we have $M \subset M_T$, which proves (ii). By Zorn's lemma, the partially ordered set S has a maximal element

J. By definition, it satisfies that $K \subset J \subset I$ and that the family $(v_i)_{i\in J}$ is linearly independent. It remains to prove the family $(v_i)_{i\in J}$ generates V. So we assume that $(v_i)_{i\in J}$ does not generate V and derive a contradiction. By this assumption, there exists an $h \in I$ such that $h \notin J$ and such that v_h is not a linear combination of $(v_i)_{i\in J}$. We claim that $J' = J \cup \{h\}$ also is an element of S. We have $K \subset J' \subset I$ by definition and must show that $(x_i)_{i\in J'}$ is linearly independent. So let $(a_i)_{i\in J'}$ be a family of scalars of finite support such that

$$\sum_{i\in J'} v_i a_i = 0.$$

First, by rewriting this equation as

$$v_h a_h + \sum_{i\in J} v_i a_i = 0,$$

we find that if $a_h = 0$. For if not, then

$$v_h = (\sum_{i\in J} v_i a_i)\cdot(-a_h^{-1}) = \sum_{i\in J} v_i(-a_i a_h^{-1}),$$

which contradicts that v_h is a not a linear combination of $(v_i)_{i\in J}$. Next, since the family $(x_i)_{i\in J}$ is linearly independent, we conclude from the equality

$$\sum_{i\in J} v_i a_i = v_h a_h + \sum_{i\in J} v_i a_i = 0$$

that also $a_i = 0$ for all $i \in J$. This proves that $(x_i)_{i\in J'}$ is linearly independent, and hence, we have proved the claim that $J' \in S$. But J is strictly contained in J', contradicting the maximality of $J \in S$, so the assumption that $(x_i)_{i\in J}$ does not generate V is false. Therefore, we conclude that $(x_i)_{i\in J}$ generates V, and hence, is a basis of V, as desired. $\square$

We will show that, given two bases of the same vector space, the cardinality of their index sets always are equal. In preparation, we prove the following lemma, which is very useful in its own right.

Lemma 12. *Let F be a field, let V be an F-vector space, and let $W \subset V$ be a subspace generated by a family $(w_1,\dots,w_m)$ of m vectors in V. If $(v_1,\dots,v_n)$ is a linearly independent family of n vectors in W, then necessarily $n \leqslant m$.*

Proof. We prove the statement by induction on m. If $m = 0$, then $W = \{0\}$ is the zero space in which the only linearly independent family of vectors is the empty family. So also $n = 0$, as desired. To prove the induction step, we assume that the statement has been proved for $m = r-1$ and prove it for $m = r$. We write

$$\begin{aligned}
v_1 &= w_1 a_{11} + w_2 a_{21} + \cdots + w_r a_{r1}\\
v_2 &= w_1 a_{12} + w_2 a_{22} + \cdots + w_r a_{r2}\\
&\vdots\\
v_n &= w_1 a_{1n} + w_2 a_{2n} + \cdots + w_r a_{rn}
\end{aligned}$$

as linear combinations of the family $(w_1,\dots,w_r)$. If the coefficients a_{rj} are zero for all $1 \leqslant j \leqslant n$, then $(v_1,\dots,v_n)$ is a linearly independent family of n vectors in the subspace $W' \subset V$ generated by the smaller family $(w_1,\dots,w_{r-1})$ of $r-1$ vectors. Hence, by the inductive hypothesis, we have $n \leqslant r-1$, and so in particular, we have $n \leqslant r$, as desired. Finally, suppose that one of the coefficients a_{rj} is nonzero. By reindexing the family $(v_1, v_2, \dots, v_n)$, if necessary, we may assume that a_{rn} is nonzero. We now consider the family $(v'_1, \dots, v'_{n-1})$ with

$$v'_j = v_j - v_n a_{rn}^{-1} a_{jn}.$$

By construction, this is a family of $n-1$ vectors in the subspace $W' \subset V$. We claim that it is also linearly independent. Granting this, we conclude from the inductive hypothesis that $n-1 \leqslant r-1$, and hence, that $n \leqslant r$, as desired. This will complete the proof of the induction step. Finally, to prove the claim, suppose that

$$v'_1 b_1 + v'_2 b_2 + \cdots + v'_{n-1} b_{n-1} = 0.$$

This equation is is equivalent to the equation

$$v_1 b_1 + v_2 b_2 + \cdots + v_{n-1} b_{n-1} - v_n a_{rn}^{-1}(a_{1n} b_1 + a_{2n} b_2 + \cdots + a_{n-1,n} b_{n-1}) = 0,$$

and since the family $(v_1, v_2, \dots, v_n)$ is linearly independent, it follows that all of the coeffients $b_1, b_2, \dots, b_{n-1}$ are zero, as required. □

In the following theorem, the case of infinite bases requires some understanding of the notion of the cardinality of a set. We will not discuss this notion here, except to say that a set α is defined, following von Neumann, to be an ordinal if it is hereditarily transitive with respect to $\in$ and that the cardinality of a set X is defined to be the smallest ordinal α for which there exists a bijection $f \colon \alpha \to X$.

Theorem 13. *Let F be a field and let V be an F-vector space. If both $(v_i)_{i \in I}$ and $(w_j)_{j \in J}$ are bases of V, then the index sets I and J have the same cardinality.*

Proof. We will show that $\text{card}(I) \leqslant \text{card}(J)$. The same argument will show that also $\text{card}(I) \geqslant \text{card}(J)$, and we may then conclude that $\text{card}(I) = \text{card}(J)$ from the Schröder-Bernstein theorem.

To show that $\text{card}(I) \leqslant \text{card}(J)$, we first assume that I is finite. If J is infinite, then there is nothing to prove, and if J is finite, then the statement follows from Lemma 12. Suppose next that I is infinite. We assume that $\text{card}(I) > \text{card}(J)$ and proceed to derive a contradiction. For every $j \in J$, we define $S_j \subset I$ to be the support of the family $(a_{i,j})_{i \in I}$ of coordinates of w_j with respect to the basis $(v_i)_{i \in I}$ and define $S = \bigcup_{j \in J} S_j \subset I$. The subsets $S_j \subset I$ are finite, by the definition of linear combination. Therefore, since I is infinite and $\text{card}(I) > \text{card}(J)$, we may conclude that also $\text{card}(I) > \text{card}(S)$. In particular, the subset $S \subset I$ is a proper subset, so there exists $h \in I$ such that $h \notin S$. We let $T_h \subset J$ be the support of the family $(b_j)_{j \in J}$ of coordinates of v_h with respect to the basis $(w_j)_{j \in J}$. Now

$$v_h = \sum_{j \in T_h} w_j b_j = \sum_{j \in T_h} (\sum_{i \in S_j} v_i a_{i,j}) b_j,$$

and therefore, the vector v_h is a linear combination of the family $(v_i)_{i \in S}$, which contradicts that $(v_i)_{i \in I}$ is a linearly independent family. □

Definition 14. Let F be a field. The dimension of an F-vector space V is the cardinality of the index set of any basis $(x_i)_{i \in I}$ of V as is written $\dim_F(V)$.

The reader may now verify the dimension statements in Example 4.