Chapter 3

Vector Spaces

The subject of this course is a theory of sets for which there is a notion of linear combinations of elements. We have already encountered linear combinations of equations and linear combinations of matrices; we are now going to formalize axiomatically such sets, which we call vector spaces. Vector spaces are abundant in mathematics (and its applications in all branches of science), and their theory is foundational to that branch of mathematics called algebra.

3.1 Definitions and examples

Definition 3.1 *Let $\mathbb{F}$ be a field. A* ***vector space*** *(מרחב וקטורי) over $\mathbb{F}$ is a non-empty set V (whose elements we call* ***vectors****) on which are defined two operations:* ***vector addition*** *(חיבור וקטורי),*

$$+ : V \times V \to V,$$

taking every $\mathbf{u}, \mathbf{v} \in V$ to an element $\mathbf{u} + \mathbf{v} \in V$, and ***scalar multiplication*** *(כפל בסקלר),*

$$\cdot : \mathbb{F} \times V \to V,$$

taking every $a \in \mathbb{F}$ and $\mathbf{u} \in V$ to an element $a\mathbf{u} \in V$.

Vector addition satisfies the following properties:

(a) Commutativity: for every $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(b) *Associativity: for every* $\mathbf{u},\mathbf{v},\mathbf{w}\in V$, $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(c) *Neutral element: there exists a vector* $0_V\in V$ *(or just* 0 *in short) such that for all* $\mathbf{u}\in V$, $\mathbf{u}+0_V=\mathbf{u}$.

(d) *Additive inverse: every* $\mathbf{u}\in V$ *has an element* $(-\mathbf{u})$, *such that* $\mathbf{u}+(-\mathbf{u})=0_V$.

Scalar multiplication satisfies the following properties:

(e) *Identity element: For every* $\mathbf{u}\in V$, $1_{\mathbb{F}}\cdot\mathbf{u}=\mathbf{u}$.

(f) *Associativity: for every* $a,b\in\mathbb{F}$ *and every* $\mathbf{u}\in V$, $a(b\mathbf{u})=(ab)\mathbf{u}$ *(note the distinction between the products* $\cdot:\mathbb{F}\times\mathbb{F}\to\mathbb{F}$ *and* $\cdot:\mathbb{F}\times V\to V$*).*

Finally, the two operations satisfy the distributive laws:

(g) *For every* $a\in\mathbb{F}$ *and* $\mathbf{u},\mathbf{v}\in V$, $a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$.

(h) *For every* $a,b\in\mathbb{F}$ *and* $\mathbf{u}\in V$, $(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u}$.

(Note the distinction between the sums $+:\mathbb{F}\times\mathbb{F}\to\mathbb{F}$ *and* $+:V\times V\to V$*.)*

Comments:

(a) A vector space hinges on two structures, a set of vectors and a field. Formally, a vector field is a four-tuple, $(V,+,\mathbb{F},\cdot)$.

(b) Vector spaces are also called **linear spaces** (מרחבים ליניאריים).

(c) Be careful not to confuse $0_{\mathbb{F}}\in\mathbb{F}$ and $0_V\in V$, although we often denote them by the same symbol, 0.

(d) There is no meaning to a product $\mathbf{u}a$, with $\mathbf{u}\in V$ and $a\in\mathbb{F}$ (even though we could have defined it by commutativity).

(e) Vector spaces don't have a canonical notion of products of vectors. For those who are acquainted with scalar and vector products, these products assume additional structure.

(f) Inductively, a vector space is closed under any finite linear combination of vectors. That is, for every $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ and $a^1,\dots,a^n\in\mathbb{F}$,

$$a^1\mathbf{v}_1+\cdots+a^n\mathbf{v}_n\in V.$$

We will often write such sums using our notation for matrix multiplication,

$$a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = \begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}a^1\\ \vdots \\ a^n\end{bmatrix}.$$

The interpretation is that the column of scalars "acts" on the row of vectors to produce a linear combination. At this stage, the role of matrices enclosed by square bracket becomes "operators" forming linear combinations. Note that we obtain products such as $\mathbf{v}_1 a^1$, which we interpret as $a^1\mathbf{v}_1$.

(g) Physicists often describe vectors as entities having a "magnitude" and a "direction"; at this stage (and throughout this course) vectors have neither magnitudes nor directions.

Example: Let $\mathbb{F}$ be any field. A set comprising just one element, $V = \{0_V\}$, is a vector space with vector addition and scalar multiplication defined the only possible way, namely

$$0_V + 0_V = 0_V \qquad \text{and} \qquad a\,0_V = 0_V.$$

Such a vector space is called the zero space (מרחב האפס), even though strictly speaking, the vector space $(\{0_V\}, +, \mathbb{F}, \cdot)$ is a different space for each field $\mathbb{F}$.
▲▲▲

Example: For any field $\mathbb{F}$ and every $n \in \mathbb{N}$, the set $V = \mathbb{F}^n$ is a vector space over $\mathbb{F}$ with respect to vector addition,

$$(u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n),$$

and scalar multiplication

$$a(u_1, \dots, u_n) = (au_1, \dots, au_n).$$

The zero vector of this space is

$$0_{\mathbb{F}^n} = (0_{\mathbb{F}}, \dots, 0_{\mathbb{F}}),$$

and the additive inverse of a vector is given by

$$-(v_1, \dots, v_n) = (-v_1, \dots, -v_n).$$

All the vector space axioms follow from the properties of the field $\mathbb{F}$ (which you should verify). Thus, $(\mathbb{F}^n, +, \mathbb{F}, \cdot)$ is a vector space. The same applies if we rather consider $\mathbb{F}^n_{\text{row}}$ or $\mathbb{F}^n_{\text{col}}$. ▲▲▲

Example: In particular, setting $n = 1$, $\mathbb{F}$ is a vector space over itself! That is, for every field $\mathbb{F}$, $(\mathbb{F}, +, \mathbb{F}, \cdot)$ is a vector space. This is quite confusing as the same set plays two different roles. ▲▲▲

Example: Consider the vector space $(\mathbb{F}^2, +, \mathbb{F}, \cdot)$ and let $\mathbf{v}_1 = (2,3)$ and $\mathbf{v}_2 = (4,5)$. The linear combination $8\mathbf{v}_1 + 9\mathbf{v}_2$ is written using the action of a matrix,

$$\begin{pmatrix} (2,3) & (4,5) \end{pmatrix} \begin{bmatrix} 8 \\ 9 \end{bmatrix}.$$

▲▲▲

Example: The space of $m \times n$ matrices with entries in $\mathbb{F}$ is a vector space over $\mathbb{F}$ with respect to vector addition

$$(A + B)^i_j = a^i_j + b^i_j$$

and scalar multiplication

$$(\lambda A)^i_j = \lambda a^i_j.$$

The zero element of this space is $0_{m \times n}$. And don't be confused: in this vector space, the vectors are matrices. ▲▲▲

Example: Let S be any non-empty set and let $V = \text{Func}(S, F)$ be the space of functions $f : S \to \mathbb{F}$ (you will learn about functions in depth in the calculus course, but let's just think of a function as a "machine" which when fed with an element in S, returns an element in $\mathbb{F}$). Then, V is a vector space over $\mathbb{F}$ with respect to vector addition

$$(f + g)(s) = f(s) + g(s)$$

and scalar multiplication

$$(a f)(s) = a f(s).$$

The zero element of this space is the function returning $0 \in \mathbb{F}$ for all $s \in S$. The additive inverse $(-f)$ of a function f is the function

$$(-f)(s) = -f(s).$$

Thus, $(\text{Func}(S, \mathbb{F}), +, \mathbb{F}, \cdot)$ is a vector space. Once again don't be confused: in this vector space, the vectors are functions. ▲▲▲

Example: Another example is that of **polynomial spaces** (מרחבי פולינומים). Let $\mathbb{F}$ be a field and let X be a symbol. We denote by $\mathbb{F}[X]$ the set of expressions of the form

$$P = p_0 + p_1 X + p_2 X^2 + p_n X^n,$$

where $p_0, \dots, p_n$ are scalars and $p_n \neq 0$. We call p_n the **leading coefficient** (מקדם מוביל) and we call $p_n X^n$ the **leading term** (איבר מוביל). To this set we also add the scalar $0_{\mathbb{F}}$. If p_n is the leading coefficient, we say that P is of **degree** (דרגה) n, and write

$$\deg P = n.$$

The degree of $P = 0_{\mathbb{F}}$ is set to be $-\infty$.

Let

$$P = \sum_{i=1}^{n} p_i X^i \qquad \text{and} \qquad Q = \sum_{i=1}^{m} q_i X^i,$$

where without loss of generality, $m \leq n$. Then, we define

$$P + Q = \sum_{i=1}^{m} (p_i + q_i) X^i + \sum_{i=m+1}^{n} p_i X^i,$$

and $P + 0_{\mathbb{F}} = P$. Likewise, we define scalar multiplication by

$$cP = \sum_{i=1}^{n} (c p_i) X^i.$$

It is readily checked that $\mathbb{F}[X]$ forms a vector space over $\mathbb{F}$ with respect to these operations. ▲▲▲

Example: The complex numbers $\mathbb{C}$ are a field, hence $\mathbb{C}$ is a vector space over $\mathbb{C}$ under the natural operations of addition and multiplication by scalars. On the other hand, $\mathbb{C}$ is also a vector space over $\mathbb{R}$, which is a totally different vector space, despite the fact that the elements of the space (i.e., the vectors) are the same. More generally, $\mathbb{C}$ is a vector space over any subfield of $\mathbb{C}$ (e.g., the complex rationals). ▲▲▲

3.2 Basic properties

Like fields, vector spaces satisfy a number of generic properties:

Proposition 3.2 *Let V be a vector space over $\mathbb{F}$. Then,*

(a) Every vector $\mathbf{v} \in V$ has a unique additive inverse.

(b) For every $a \in \mathbb{F}$, $a\,0_V = 0_V$.

(c) For every $\mathbf{u} \in V$, $0_{\mathbb{F}}\,\mathbf{u} = 0_V$.

(d) If $a \in \mathbb{F}$ and $\mathbf{u} \in V$ satisfy $a\mathbf{u} = 0_V$, then either $a = 0_{\mathbb{F}}$ or $\mathbf{u} = 0_V$.

(e) For every $\mathbf{u} \in V$, $(-1_{\mathbb{F}})\mathbf{u} = -\mathbf{u}$.

Proof:

(a) Suppose that $\mathbf{u} + \mathbf{v} = 0_V$ and $\mathbf{w} + \mathbf{v} = 0_V$. It follows from the first three properties of vector addition that

$$\mathbf{u} = 0_V + \mathbf{u} = (\mathbf{w} + \mathbf{v}) + \mathbf{u} = \mathbf{w} + (\mathbf{v} + \mathbf{u}) = \mathbf{w} + (\mathbf{u} + \mathbf{v}) = \mathbf{w} + 0_V = \mathbf{w}.$$

(b) By the properties of 0_V and distributivity,

$$a\,0_V = a(0_V + 0_V) = a\,0_V + a\,0_V.$$

Adding $-(a\,0_V)$ to both sides and using the properties of vector addition,

$$\begin{aligned} 0_V &= a\,0_V + (-(a\,0_V)) \\ &= (a\,0_V + a\,0_V) + (-(a\,0_V)) \\ &= a\,0_V + (a\,0_V + (-(a\,0_V))) \\ &= a\,0_V + 0_V \\ &= a\,0_V, \end{aligned}$$

proving that $a\,0_V = 0_V$.

(c) Similarly,

$$0_{\mathbb{F}}\,\mathbf{u} = (0_{\mathbb{F}} + 0_{\mathbb{F}})\mathbf{u} = 0_{\mathbb{F}}\,\mathbf{u} + 0_{\mathbb{F}}\,\mathbf{u}.$$

Adding $-(0_{\mathbb{F}}\,\mathbf{u})$ to both sides,

$$\begin{aligned} 0_V &= 0_{\mathbb{F}}\,\mathbf{u} + (-(0_{\mathbb{F}}\,\mathbf{u})) \\ &= (0_{\mathbb{F}}\,\mathbf{u} + 0_{\mathbb{F}}\,\mathbf{u}) + (-(0_{\mathbb{F}}\,\mathbf{u})) \\ &= 0_{\mathbb{F}}\,\mathbf{u} + (0_{\mathbb{F}}\,\mathbf{u} + (-(0_{\mathbb{F}}\,\mathbf{u}))) \\ &= 0_{\mathbb{F}}\,\mathbf{u} + 0_V \\ &= 0_{\mathbb{F}}\,\mathbf{u}, \end{aligned}$$

proving that $0_{\mathbb{F}}\,\mathbf{u} = 0_V$.

(d) Suppose that $a\mathbf{u} = 0_V$. If $a \neq 0_{\mathbb{F}}$, then using the fact that a has a multiplicative inverse,

$$\mathbf{u} = 1_{\mathbb{F}} \cdot \mathbf{u} = (a^{-1}a)\mathbf{u} = a^{-1}(a\mathbf{u}) = a^{-1}0_V = 0_V,$$

i.e., either $a = 0_{\mathbb{F}}$ or $\mathbf{u} = 0_V$.

(e) We have

$$0_V = 0_{\mathbb{F}}\,\mathbf{u} = (1_{\mathbb{F}} + (-1_{\mathbb{F}}))\mathbf{u} = \mathbf{u} + (-1_{\mathbb{F}})\mathbf{u},$$

and it follows from the uniqueness of the inverse that $(-1_{\mathbb{F}})\mathbf{u} = -\mathbf{u}$.

■

Comment: The fourth item has an important consequence: suppose that $\mathbf{v} \in V$ is non-zero and there exist $a, b \in \mathbb{F}$, such that $a\mathbf{v} = b\mathbf{v}$. Then, $(a-b)\mathbf{v} = 0$, from which we deduce that $a = b$.

We now come to the *raison d'être* of vector spaces—the formation of linear combinations:

Definition 3.3 *Let V be a vector space over a field $\mathbb{F}$ and let $(\mathbf{u}_1, \dots, \mathbf{u}_n) \subset V$ be a sequence of n vectors. A vector $\mathbf{v} \in V$ is said to be a* ***linear combination*** *of $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, if there exists a sequence of scalars $(a^1, \dots, a^n) \in \mathbb{F}^n$, such that*

$$\mathbf{v} = a^1\mathbf{u}_1 + \cdots + a^n\mathbf{u}_n,$$

or in matrix form, if there exists an $\mathbf{a} \in \mathbb{F}^n_{col}$, such that

$$\mathbf{v} = \begin{pmatrix} \mathbf{u}_1 & \dots \mathbf{u}_n \end{pmatrix} \mathbf{a}.$$

Some notations: Let $(V,+,\mathbb{F},\cdot)$ be a vector space. For every $\mathbf{v}\in V$, we denote by

$$\mathbb{F}\mathbf{v} = \{a\mathbf{v} \ : \ a\in\mathbb{F}\}$$

the set of scalar multiples of $\mathbf{v}$. For $S,T\subset V$ we denote by

$$S+T = \{\mathbf{u}+\mathbf{v} \ : \ \mathbf{u}\in S, \ \ \mathbf{v}\in T\}$$

the set of vectors obtained by sums of elements of S and T.

Exercises

(easy) 3.1 What is a vector? Let S be any non-empty set and let $x\in S$. How can we tell whether x is a vector?

(easy) 3.2 In what sense is every field a vector space? Is it true that every vector space is a field?

(easy) 3.3 Let S be any non-empty set and let $V = \mathrm{Func}(S,\mathbb{F})$. Prove that it is indeed a vector space with respect to the vector addition and scalar multiplication defined above.

(easy) 3.4 Let $V = \mathbb{R}^2$ be the set of pairs of real number and let $\mathbb{F} = \mathbb{R}$. Define

$$(x,y)+(w,z) = (x+w,0)$$
$$a(x,y) = (ax,0).$$

Is V a vector space over $\mathbb{R}$ under these operations?

(easy) 3.5 What is the smallest vector space containing more than one vector?

(easy) 3.6 Show that any vector space over $\mathbb{R}$ is either the zero space, or contains infinitely-many vectors.

(intermediate) 3.7 Let V be a vector space over $\mathbb{F}$. Prove that for every $\mathbf{v},\mathbf{w}\in V$ and $0\neq a\in\mathbb{F}$ there exists a unique $\mathbf{u}\in V$ satisfying

$$a\mathbf{u}+\mathbf{v} = \mathbf{w}.$$

Hint: you've done something very similar in the context of fields.

(intermediate) 3.8 Use the result of Exercise 3.7 to deduce the uniqueness of the additive inverse.

(intermediate) 3.9 Let

$$V = \{x \in \mathbb{R} \ : \ x > 0\}.$$

For $x, y \in V$ and $a \in \mathbb{R}$ define

$$x \oplus y = xy \qquad \text{and} \qquad a \odot x = x^a.$$

Prove that $(V, \oplus, \mathbb{R}, \odot)$ is a vector space.

(intermediate) 3.10 Consider the vector space $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$. Let

$$\mathbf{w} = (2, -1) \in \mathbb{R}^2,$$

and define on $\mathbb{R}^2$ the following two operations,

$$\mathbf{u} \boxplus \mathbf{v} = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \qquad \text{and} \qquad \mathbf{a} \boxdot \mathbf{v} = a\mathbf{v} + (a-1)\mathbf{w}.$$

(a) Is there an element in $\mathbb{R}^2$ neutral to $\boxplus$? If yes, what is it? (b) Does any element in $\mathbb{R}^2$ have an additive-inverse with respect to $\boxplus$? If yes, what is it? (c) Are the operations distributive, namely,

$$a \boxdot (\mathbf{u} \boxplus \mathbf{v}) = a \boxdot \mathbf{u} \boxplus a \boxdot \mathbf{v} \ ?$$

(intermediate) 3.11 Consider the vector space $(\mathbb{C}^3, +, \mathbb{C}, \cdot)$. Which vectors are linear combinations of the vectors $(1, 0, -1)$, $(0, 1, 1)$ and $(1, 1, 1)$?

3.3 Subspaces

3.3.1 Definitions and examples

A recurring theme in mathematics is to consider a subset of a structure, which inherits the properties of the structure it is part of. This leads us to the definition of a linear subspace of a vector space.

Let $W \subseteq V$. Since every vector in W is a vector in V, we can add together vectors in W. The restriction of the operation $+ : V \times V \to V$ to pairs of vectors in W is denoted by

$$+|_{W\times W} : W \times W \to V.$$

The sum of two vectors in W is not necessarily a vector in W, but it is necessarily a vector in V. Likewise, the restriction of the operation $\cdot : \mathbb{F}\times V \to V$ to pairs of vectors in W is denoted by

$$\cdot|_{\mathbb{F}\times W} : \mathbb{F} \times W \to V.$$

A scalar multiple of a vector in W is not necessarily a vector in W, but it is necessarily a vector in V.

Definition 3.4 *Let V be a vector space over $\mathbb{F}$. A* ***subspace*** *(or* ***linear subspace****) (תת מרחב וקטורי) of V is a non-empty subset $W \subseteq V$, which is closed under vector addition and scalar multiplication, namely, for all $\mathbf{u}, \mathbf{v} \in V$ and $a \in \mathbb{F}$,*

$$\mathbf{u} + \mathbf{v} \in W \qquad \text{and} \qquad a\mathbf{v} \in W.$$

We denote the relation of W being a linear subspace of V by $W \le V$.

The following proposition asserts that a linear subspace of a vector space is a vector space in its own right:

Proposition 3.5 *Let V be a vector space over a field $\mathbb{F}$ and let $W \le V$. Then, $(W, +|_{W\times W}, \mathbb{F}, \cdot|_{\mathbb{F}\times W})$ is a vector space.*

Proof: Since W is not empty, it contains at least one element $\mathbf{w}$. Then,

$$(-1)\mathbf{w} \in W \qquad \text{and} \qquad (-1)\mathbf{w} + \mathbf{w} \in W,$$

i.e., W includes 0_V. Likewise, for every $\mathbf{w} \in W$,

$$(-1)\mathbf{w} + 0_V \in W,$$

i.e., every element of W has its additive inverse in W. It remains to show that all eight axioms are satisfied, but this follows from the axioms in V. For example, for every $\mathbf{u}, \mathbf{v} \in W$, since $\mathbf{u}, \mathbf{v} \in V$, it follows that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. ∎

Example: Let V be a vector space over $\mathbb{F}$. The subset $\{0_V\} \subset V$ is a linear subspace of V. It is called the **zero subspace** of V, $\{0_V\} \leq V$, ▲▲▲

Example: Every vector space is a subspace of itself, $V \leq V$. We will refer to a **proper subspace** to emphasize that W is a strict subset of V, i.e., that $V \setminus W$ is not empty. We denote this relation by $W < V$. ▲▲▲

Example: Consider the vector space $(M_n(\mathbb{F}), +, \mathbb{F}, \cdot)$. A matrix $A \in M_n(\mathbb{F})$ is called **symmetric** if $a^i_j = a^j_i$ for all $i, j \in 1, \dots, n$. It is easy to see that the subset of symmetric matrices is a linear subspace of $M_n(\mathbb{F})$. ▲▲▲

Example: Consider the vector space $V = (\mathbb{F}^n_{\text{col}}, +, \mathbb{F}, \cdot)$. Let $A \in M_{m\times n}(\mathbb{F})$ and let

$$W = \{\mathbf{x} \in \mathbb{F}^n_{\text{col}} \;:\; A\mathbf{x} = 0\}$$

be the set of solutions of the corresponding homogeneous system of equations. By Theorem 2.40, $W \leq V$. ▲▲▲

Example: Let V be a vector space and let $\mathbf{w} \in V$. Consider the subset of V,

$$W = \mathbb{F}\mathbf{w}.$$

We claim that W is not just a subset of V; it is a linear subspace. Why? It is not-empty as it includes $1_{\mathbb{F}} \cdot \mathbf{w} = \mathbf{w}$. Moreover, let $\mathbf{u}, \mathbf{v} \in W$. By the definition of W, there exist $a, b \in \mathbb{F}$, such that

$$\mathbf{u} = a\,\mathbf{w} \qquad \text{and} \qquad \mathbf{v} = b\,\mathbf{w}.$$

Then,

$$\mathbf{u} + \mathbf{v} = a\,\mathbf{w} + b\,\mathbf{w} = (a+b)\,\mathbf{w} \in W.$$

Let $c \in \mathbb{F}$, then

$$c\,\mathbf{u} = c\,(a\,\mathbf{w}) = (ca)\,\mathbf{w} \in W,$$

proving W is a linear subspace of V. ▲▲▲

Exercises

(easy) 3.12 Let V be a vector space over $\mathbb{F}$. Prove that $W \leq V$ and $U \leq W$ implies that $U \leq V$.

(easy) 3.13 Consider the vector space $(V, \oplus, \mathbb{R}, \odot)$ in Exercise 3.9. Is it a linear subspace of the vector space $(\mathbb{R}, +, \mathbb{R}, \cdot)$?

(intermediate) 3.14 Consider the vector space $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.

(a) Find a subset $W \subset \mathbb{R}^2$ including the zero vector, which is closed under scalar multiplication but not closed under vector addition.

(b) Find a subset $U \subset \mathbb{R}^2$ including the zero vector, which is closed under vector addition but not closed under scalar multiplication.

(c) Does there exist a non-empty subset $V \subset \mathbb{R}^2$ which does not include the zero vector, which is closed under scalar multiplication?

(intermediate) 3.15 In each of the following items is given a subset W of a vector space $(V, +, \mathbb{F}, \cdot)$. Determine whether $W \leq V$.

(a) $V = (\mathbb{C}^2, +, \mathbb{C}, \cdot)$ and

$$W = \{(z, w) \; : \; 2z = 3w\}.$$

(b) $V = (M_{2\times 2}(\mathbb{R}), +, \mathbb{R}, \cdot)$ and

$$W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \; : \; ad = 0 \right\}.$$

(c) $V = (\mathbb{R}[X], +, \mathbb{R}, \cdot)$ and

$$W = \{p(X) \in \mathbb{R}[X] \; : \; p(0) = p(2)\}.$$

(d) $V = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$ and

$$W = \{(x, y, z) \; : \; 2x - y + z = 0, \;\; y - 2z = 0\}.$$

(e) $V = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$ and

$$W = \{(x, y, z) \; : \; xy = z\}.$$

(f) $V = \text{Func}(\mathbb{R}, \mathbb{R})$ over $\mathbb{R}$, and

$$W = \{f : \mathbb{R} \to \mathbb{R} \ : \ f(2) = f(3)\}.$$

(g) $V = \text{Func}(\mathbb{R}, \mathbb{R})$ over $\mathbb{R}$, and

$$W = \{f : \mathbb{R} \to \mathbb{R} \ : \ f(0) = f^2(1)\}.$$

(intermediate) 3.16 Let $V = \text{Func}(\mathbb{R}, \mathbb{R})$ over $\mathbb{R}$. Which of the following subsets is a linear subspace?

(a) The functions f satisfying $f(-1) + f(1) = 0$.
(b) The functions f satisfying $f(0) + f(1) = 1$.
(c) The functions f satisfying $f(0) \cdot f(1) = 0$.
(d) The functions f satisfying $f(-x) + f(x) = 0$ for all $x \in \mathbb{R}$.

(intermediate) 3.17 Consider the vector space $(\mathbb{C}^n, +, \mathbb{C}, \cdot)$. Let $W \le \mathbb{C}^n$ and consider the set $U \subseteq \mathbb{C}^n$,

$$U = \left\{ (\bar{z}^1, \dots, \bar{z}^n) \ : \ (z^1, \dots, z^n) \in W \right\},$$

where $\bar{z}$ is the complex conjugate of z. Show that $U \le \mathbb{C}^n$.

(intermediate) 3.18 Consider the vector space $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$ for some $n \ge 3$. Which of the following subsets of $\mathbb{R}^n$ is a linear subspace?

(a) All $\mathbf{x} = (x_1, \dots, x_n)$ satisfying $x_1 \ge 0$.
(b) All $\mathbf{x} = (x_1, \dots, x_n)$ satisfying $x_1 + 3x_2 = x_3$.
(c) All $\mathbf{x} = (x_1, \dots, x_n)$ satisfying $x_2 = x_1^2$ (here the superscript 2 is a square).
(d) All $\mathbf{x} = (x_1, \dots, x_n)$ satisfying $x_1 x_2 = 0$.
(e) All $\mathbf{x} = (x_1, \dots, x_n)$ such that x_2 is rational.

(harder) 3.19 Let V be a vector space over a field $\mathbb{F}$ and let

$$S = \{\mathbf{v}_\alpha \ : \ \alpha \in I\}$$

be a non-empty subset of V; here I is an index set (which could be infinite). Consider the subset $U \subset V$ comprising all linear combinations of vectors in S,

$$U = \left\{ \sum_{\alpha \in J} a^\alpha \mathbf{v}_\alpha \;:\; J \subset I \text{ is finite, } a^\alpha \in \mathbb{F},\ \mathbf{v}_\alpha \in S \right\}.$$

Prove that $U \le V$.

3.3.2 The subspace generated by a set

A vector space may have many linear subspaces. The following proposition asserts that the intersection of any collection of linear subspaces is again a linear subspace:

Proposition 3.6 *Let V be a vector space over $\mathbb{F}$. Let $\mathscr{C}$ be a (possibly infinite) collection of linear subspaces of V (i.e., $\mathscr{C}$ is a set whose elements are linear subspaces of V). Then,*

$$W = \bigcap \mathscr{C}$$

is a linear subspace of V.

Proof: First, let's interpret the statement of this proposition. There is a collection of linear subspaces of V; this collection could be finite (e.g., seven subspaces, which we could denote by $W_1, \dots, W_7$); this collection could be **countable** (בת מנייה), i.e., form a **sequence** (סדרה) (which we could denote by $W_1, W_2, \dots$); this collection could also be **uncountably infinite**. The set

$$W = \bigcap \mathscr{C}$$

comprises all those elements in V which are elements in U for *every* $U \in \mathscr{C}$, i.e., $\mathbf{w} \in W$ if and only if $\mathbf{w} \in U$ for all $U \in \mathscr{C}$. The claim is that this set is a linear subspace of V.

By definition, we need to show that W is not empty, and that for every $\mathbf{u}, \mathbf{v} \in W$ and $a \in \mathbb{F}$,

$$\mathbf{u} + \mathbf{v} \in W \qquad \text{and} \qquad a\,\mathbf{u} \in W.$$

Each of the $U \in \mathscr{C}$ is a linear subspace of V, hence

$$0_V \in U \quad \text{for every } U \in \mathscr{C},$$

from which follows that $0_V \in W$.
Let $\mathbf{u}, \mathbf{v} \in W$. By the very definition of W,

$$\mathbf{u}, \mathbf{v} \in U \quad \text{for every } U \in \mathscr{C},$$

Since every such U is a linear subspace,

$$\mathbf{u} + \mathbf{v} \in U \quad \text{for every } U \in \mathscr{C},$$

from which follows that $\mathbf{u} + \mathbf{v} \in W$.
Likewise for $a \in F$ and $\mathbf{u} \in V$,

$$a\,\mathbf{u} \in U \quad \text{for every } U \in \mathscr{C},$$

from which follows that $a\,\mathbf{u} \in W$. This concludes the proof. ■

This proposition has an important consequence, whose likes are recurring in many branches of mathematics. Let $S \subset V$ be a collection of vectors, which could be finite, countably infinite, uncountable infinite or even empty. Consider the collection of all linear subspaces of V which contain all those vectors, namely,

$$\mathscr{C} = \{W \le V \ : \ S \subseteq W\}.$$

This collection is not empty, because V itself is a linear subspace of V containing all vectors in S, i.e.,

$$V \in \mathscr{C}.$$

Whatever this collection of linear subspaces is, its intersection is a linear subspace of V. We call it the linear subspace **generated** (**תת מרחב נוצר**) by the vectors in S, and denote it by

$$\langle S \rangle = \bigcap \{W \le V \ : \ S \subseteq W\}. \tag{3.1}$$

The following two lemmas provide a useful characterization of the generated subspace:

Lemma 3.7 *Let V be a vector space over $\mathbb{F}$, and let $S \subseteq V$. If $S \subseteq W \leq V$, then $\langle S \rangle \leq W$.*

Proof: This is really a direct consequence of the definition (3.1). If $W \leq V$ contains S, i.e.,

$$W \in \{\tilde{W} \leq V \ : \ S \subset \tilde{W}\},$$

then,

$$\bigcap \{\tilde{W} \leq V \ : \ S \subset \tilde{W}\} \subset W,$$

as an intersection of any collection of sets is contained in any set in that intersection, but this is exactly what we have to prove. ■

Lemma 3.8 *Let V be a vector space over $\mathbb{F}$, and let $S \subseteq V$. If $T \subseteq V$ satisfies that $T \subseteq W$ for every $W \leq V$ containing S, then*

$$T \in \langle S \rangle .$$

Proof: Once again, this is a direct consequence of the definition of the generated subspace. If

$$T \subseteq W \quad \text{for all} \quad W \in \{\tilde{W} \leq V \ : \ S \subset \tilde{W}\},$$

then

$$T \subseteq \bigcap \{\tilde{W} \leq V \ : \ S \subset \tilde{W}\}.$$

■

Example: Let V be a vector space over $\mathbb{F}$. Let $\mathbf{w} \in V$ and let $S = \{\mathbf{w}\}$. As a matter of convenience, we write $\langle \mathbf{w} \rangle$ rather than $\langle \{\mathbf{w}\} \rangle$. We will show that

$$\langle \mathbf{w} \rangle = \mathbb{F}\mathbf{w},$$

that is, the linear subspace generated by a single vector is the subspace obtained by all multiples of that vectors by scalars. If $\mathbf{w} = 0_V$, then this subspace is the zero subspace. Otherwise, it is a *line*.

We have already seen that

$$\mathbb{F}\mathbf{w} \leq V.$$

Since $\{\mathbf{w}\} \subset \mathbb{F}\mathbf{w} \leq V$, if follows from Lemma 3.7 that

$$\langle \mathbf{w} \rangle \subseteq \mathbb{F}\mathbf{w}.$$

Conversely, every vector in $\mathbb{F}\mathbf{w}$ must be included in any linear subspace containing $\mathbf{w}$, namely

$$\mathbb{F}\mathbf{w} \subseteq W \quad \text{for all} \quad W \in \{\tilde{W} \leq V \ : \ \{\mathbf{w}\} \subset \tilde{W}\},$$

and by Lemma 3.8,

$$\mathbb{F}\mathbf{w} \subseteq \langle \mathbf{w} \rangle .$$

▲▲▲

The following properties of the generated subspace follow almost directly from the definition.

Proposition 3.9 *In every vector space*

$$\langle \varnothing \rangle = \{0_V\}.$$

Proof: Since $\{0_V\} \leq V$ contains the empty set, it follows from Lemma 3.7 that

$$\langle \varnothing \rangle \subseteq \{0_V\}.$$

Conversely, since $\{0_V\}$ is contained in every linear subspace of V, it follows that

$$\{0_V\} \subseteq W \quad \text{for all} \quad W \in \{\tilde{W} \leq V \ : \ \varnothing \subset \tilde{W}\},$$

and by Lemma 3.8,

$$\{0_V\} \subseteq \langle \varnothing \rangle .$$

■

Proposition 3.10 *Let V be a vector space over $\mathbb{F}$ and let $S \subseteq T \subseteq V$. Then,*

$$\langle S \rangle \leq \langle T \rangle .$$

(Note that we write $\langle S \rangle \leq \langle T \rangle$ rather than $\langle S \rangle \subseteq \langle T \rangle$ because these are linear subspaces.)

Proof: Since $S \subseteq T$, it follows that

$$\{W \leq V \ : \ T \subseteq W\} \subseteq \{W \leq V \ : \ S \subseteq W\}.$$

Intersecting over more sets can only reduce the intersection, hence

$$\langle T \rangle = \bigcap\{W \leq V \ : \ T \subset W\} \supseteq \bigcap\{W \leq V \ : \ S \subseteq W\} = \langle S \rangle .$$

■

More properties of generated subspaces are derived in the exercise section.

Exercises

(easy) 3.20 Let $W_1, W_2 \leq V$. Prove directly (i.e., without recurring to the general theorem proved above) that $W_1 \cap W_2 \leq V$.

(intermediate) 3.21 Let V be a vector space over a field $\mathbb{F}$ and let $W \leq V$. Show that

$$\langle W \rangle = W.$$

(intermediate) 3.22 Let V be a vector space over a field $\mathbb{F}$ and let $S \subseteq V$. Show that

$$\langle \langle S \rangle \rangle = \langle S \rangle .$$

(intermediate) 3.23 Let V be a vector space over $\mathbb{F}$. Let $S_1, S_2 \subseteq V$ be non-empty subsets. Suppose that

$$S_1 \subseteq \langle S_2 \rangle \qquad \text{and} \qquad S_2 \subseteq \langle S_1 \rangle .$$

Show that

$$\langle S_1 \rangle = \langle S_2 \rangle .$$

(intermediate) 3.24 Let V be a vector space over $\mathbb{F}$. Let $S_1, S_2 \subseteq V$ be non-empty subsets. For each of the following statements, determine whether it is true or not:

(a) If $\langle S_1 \rangle \subseteq \langle S_2 \rangle$ then $S_1 \subseteq S_2$.

(b) If $S_2 \subseteq S_1$ and $\langle S_1 \rangle \subseteq \langle S_2 \rangle$, then $\langle S_1 \rangle = \langle S_2 \rangle$.

(c) If $S_2 \subseteq S_1$ and $\langle S_2 \rangle \neq \langle S_1 \rangle$, then for every $\mathbf{v} \in S_1 \setminus S_2$ we have $\mathbf{v} \notin \langle S_2 \rangle$.

(d) If $S_2 \subseteq S_1$ and $\langle S_2 \rangle \neq \langle S_1 \rangle$, then there exists $\mathbf{v} \in S_1 \setminus S_2$ such that $\mathbf{v} \notin \langle S_2 \rangle$.

(e) If $S_1 \cap S_2 = \varnothing$, then $\langle S_1 \rangle \cap \langle S_2 \rangle = \{0\}$.

3.3.3 The linear span of a set of vectors

The definition of the linear subspace generated by a collection of vectors is quite implicit. We will now provide a more explicit characterization.

Definition 3.11 *Let V be a vector space over a field $\mathbb{F}$ and let $S \subseteq V$ be a non-empty collection of vectors. Then, the* ***linear span*** *(הפרוס הליניארי) of S is the set of all (finite) linear combinations of elements of S,*

$$\operatorname{Span} S = \left\{a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n \; : \; n \in \mathbb{N},\; a^i \in \mathbb{F},\; \mathbf{v}_i \in S\right\}. \tag{3.2}$$

In the particular case where $S = \varnothing$ we define $\operatorname{Span} S = \{0_V\}$.

Example: Let $\mathbf{w} \in V$. Then, the only linear combinations of $\{\mathbf{w}\}$ are scalar multiples of $\mathbf{w}$,

$$\operatorname{Span}\{\mathbf{w}\} = \mathbb{F}\mathbf{w}.$$

Note that $\operatorname{Span}\{\mathbf{w}\} = \langle \mathbf{w} \rangle$. We will shortly see that this is a general identity (note also that we defined the span such that $\operatorname{Span} \varnothing = \{0_V\} = \langle \varnothing \rangle$). ▲▲▲

Example: Let $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ and let

$$S = \{(1,1), (-1,1)\}.$$

Then,

$$\operatorname{Span} S = \{a(1,1) + b(-1,1) \; : \; a, b \in \mathbb{R}\} = \{(a-b, a+b) \; : \; a, b \in \mathbb{R}\}.$$

It is not hard to see that for every $(x,y) \in \mathbb{R}^2$,

$$(x,y) = (a-b, a+b),$$

where

$$a = \tfrac{1}{2}(y+x) \qquad \text{and} \qquad b = \tfrac{1}{2}(y-x),$$

proving that $\operatorname{Span} S = \mathbb{R}^2$. ▲▲▲

Proposition 3.12 *Let V be a vector space over a field $\mathbb{F}$ and let $S \subseteq V$. Then $\operatorname{Span} S \leq V$. (Note that this was already mentioned in Exercise 3.19.)*

Proof: If $S = \varnothing$, then $\operatorname{Span} S = \{0_V\} \le V$ by convention. Otherwise, it suffices to show that for every $\mathbf{u}, \mathbf{v} \in \operatorname{Span} S$ and $c \in \mathbb{F}$,

$$\mathbf{u} + \mathbf{v} \in \operatorname{Span} S \qquad \text{and} \qquad c\,\mathbf{u} \in \operatorname{Span} S.$$

If $\mathbf{u}, \mathbf{v} \in \operatorname{Span} S$, then there exist vectors $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m \in S$ and scalars $a^1, \dots, a^n, b^1, \dots, b^m \in \mathbb{F}$, such that

$$\mathbf{u} = a^1\mathbf{u}_1 + \cdots + a^n\mathbf{u}_n \qquad \text{and} \qquad \mathbf{v} = b^1\mathbf{v}_1 + \cdots + b^m\mathbf{v}_m.$$

Then

$$\mathbf{u} + \mathbf{v} = a^1\mathbf{u}_1 + \cdots + a^n\mathbf{u}_n + b^1\mathbf{v}_1 + \cdots + b^m\mathbf{v}_m \in \operatorname{Span} S.$$

Likewise,

$$c\,\mathbf{u} = c\,a^1\mathbf{u}_1 + \cdots + c\,a^n\mathbf{u}_n \in \operatorname{Span} S,$$

proving that $\operatorname{Span} S \le V$. ■

Theorem 3.13 *Let V be a vector space over a field $\mathbb{F}$ and let $S \subset V$. Then,*

$$\operatorname{Span} S = \langle S \rangle\,.$$

Proof: If $S = \varnothing$, then this holds by definition. Otherwise, since

$$\operatorname{Span} S \in \{W \le V \;:\; S \subseteq W\},$$

it follows from Lemma 3.7 that

$$\langle S \rangle \subseteq \operatorname{Span} S.$$

Conversely, since every $W \le V$ containing S must contain every vector in $\operatorname{Span} S$, i.e.,

$$\operatorname{Span} S \subseteq W \qquad \text{for all} \qquad W \in \{\tilde{W} \le V \;:\; S \subseteq \tilde{W}\},$$

it follows by Lemma 3.8 that

$$\operatorname{Span} S \subseteq \langle S \rangle\,,$$

which completes the proof. ■

Corollary 3.14 *Let $W \leq V$. Then,*

$$\operatorname{Span} W = W.$$

Proof: This corollary asserts that linear subspaces are closed under linear combinations. We can prove it directly, but we can get this as a consequence of the last theorem, recalling that $\langle W \rangle = W$ (see Exercise 3.21). ■

Example: Let $V = \mathbb{R}^5$ and let

$$\begin{aligned} \mathbf{v}_1 &= (1,2,0,3,0) \\ \mathbf{v}_2 &= (0,0,1,4,0) \\ \mathbf{v}_3 &= (0,0,0,0,1). \end{aligned}$$

A vector $\mathbf{v} \in \mathbb{R}^5$ is in the linear span of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ if and only if there exist scalar a, b, c, such that

$$\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3,$$

i.e., if there exist such scalars such that

$$\mathbf{v} = (a, 2a, b, 3a + 4b, c).$$

We can relate this to linear systems: $\operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is the set of all $\mathbf{x} \in \mathbb{R}^5$, such that

$$x^2 = 2x^1 \qquad \text{and} \qquad x^4 = 3x^1 + 4x^3.$$

▲▲▲

Exercises

(easy) 3.25 Let V be a vector space over the field $\mathbb{F}_2$ and let $\mathbf{v} \in V$ be a non-zero vector. Write explicitly all the vectors in $\operatorname{Span}\{\mathbf{v}\}$.

(easy) 3.26 Consider the vector space $(\mathbb{F}^3, +, \mathbb{F}, \cdot)$. Find two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^3$, such that

$$\operatorname{Span}\{\mathbf{u}, \mathbf{v}\} = \{(0_{\mathbb{F}}, a, b) \ : \ a, b \in \mathbb{F}\}.$$

(easy) 3.27 Consider the vector space $(\mathbb{R}^4, +, \mathbb{R}, \cdot)$. Find two different sets $S, T \subset \mathbb{R}^4$, such that

$$\operatorname{Span} S = \operatorname{Span} T = \{(a, a-b, b, a+b) \ : \ a, b \in \mathbb{R}\}.$$

(intermediate) 3.28 Consider the vector space $(\mathbb{R}^4, +, \mathbb{R}, \cdot)$, and let

$$\begin{aligned}\mathbf{v}_1 &= (2, -1, 3, 2)\\ \mathbf{v}_2 &= (-1, 1, 1, -3)\\ \mathbf{v}_3 &= (1, 1, 9, -5).\end{aligned}$$

Is

$$(3, -1, 0, -1) \in \operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}?$$

(intermediate) 3.29 Let V be a vector space over $\mathbb{R}$ and let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(a) Is $\operatorname{Span}\{\mathbf{u}-\mathbf{v}, \mathbf{v}-\mathbf{w}, \mathbf{w}\} = \operatorname{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$?

(b) Is $\operatorname{Span}\{\mathbf{u}-\mathbf{v}, \mathbf{v}-\mathbf{w}, \mathbf{w}-\mathbf{u}\} = \operatorname{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$?

(c) Is it possible that $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are distinct and $\operatorname{Span}\{\mathbf{u}-\mathbf{v}, \mathbf{v}-\mathbf{w}, \mathbf{w}-\mathbf{u}\} = \operatorname{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$?

(harder) 3.30 Let $W \subset \mathbb{R}^5$ be the set of all solutions to the linear system

$$\begin{array}{cccccl} 2X^1 & -X^2 & +\frac{4}{3}X^3 & -X^4 & & = 0\\ X^1 & & +\frac{2}{3}X^3 & & -X^5 & = 0\\ 9X^1 & -3X^2 & +6X^3 & -3X^4 & -3X^5 & = 0. \end{array}$$

Find a set of three vectors spanning W.

(intermediate) 3.31 Prove that the only linear subspaces of $\mathbb{R}$ (the field $\mathbb{R}$ as a vector space over itself) are $\mathbb{R}$ and $\{0\}$.

(intermediate) 3.32 Let V be a vector space over $\mathbb{F}$ and let $W \le V$. Let $S \subseteq V$ satisfying $\operatorname{Span} S = V$. Prove or disprove: there exists a subset $T \subseteq S$, such that $\operatorname{Span} T = W$.

(intermediate) 3.33 Let V be a vector space over $\mathbb{F}$. Let $W \le V$ and let $\mathbf{u}, \mathbf{v} \in V \setminus W$. Show that

$$\mathbf{u} \in \operatorname{Span}(W \cup \{\mathbf{v}\}) \quad \text{if and only if} \quad \mathbf{v} \in \operatorname{Span}(W \cup \{\mathbf{u}\}).$$

(harder) 3.34 Prove that the only linear subspaces of $\mathbb{R}^2$ are $\mathbb{R}^2$, $\{0\}$ or sets of the form

$$\mathbb{R}\mathbf{v}$$

for some $\mathbf{v} \in \mathbb{R}^2$.

(harder) 3.35 What are all the linear subspaces of $(\mathbb{C}, +, \mathbb{R}, \cdot)$?

(harder) 3.36 Let $W_1, W_2 \le V$. Suppose that $W_1 \cup W_2 \le V$. Prove that either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

3.3.4 The row space of a matrix

Let $A \in M_{m\times n}(\mathbb{F})$. The rows of A,

$$\{\mathrm{Row}^i(A) \;:\; i = 1, \dots, m\}$$

are a subset of $\mathbb{F}^n_{\mathrm{row}}$, which is a vector space over $\mathbb{F}$. Their linear span is called the **row space** (**מרחב השורות**) of A, denoted by

$$\mathscr{R}(A) = \mathrm{Span}\left\{\mathrm{Row}^i(A) \;:\; i = 1, \dots, m\right\}.$$

We can express linear combinations of the rows of A using matrix notation,

$$\sum_{i=1}^m c_i \,\mathrm{Row}^i(A) = \begin{bmatrix} c_1 & \cdots & c_m \end{bmatrix} \begin{bmatrix} \mathrm{Row}^1(A) \\ \vdots \\ \mathrm{Row}^m(A) \end{bmatrix}.$$

Namely,

$$\mathscr{R}(A) = \{\mathbf{c}A \;:\; \mathbf{c} \in \mathbb{F}^m_{\mathrm{row}}\}.$$

Example: Consider the case where $m < n$, and

$$A = \begin{bmatrix} 1 & & & 0 & 0 \\ & 1 & & 0 & 0 \\ & & \ddots & \vdots & \vdots \\ & & 1 & 0 & 0 \end{bmatrix}.$$

Then

$$\mathscr{R}(A) = \{\mathbf{x} \in \mathbb{F}^n_{\mathrm{row}} \;:\; x_{m+1} = \cdots = x_n = 0\}.$$

▲ ▲ ▲

Example: Consider the case where $m > n$, and

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Then,

$$\mathscr{R}(A) = \mathbb{F}^n_{\text{row}}.$$

▲▲▲

Lemma 3.15 *Let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times k}(\mathbb{F})$. Then,*

$$\mathscr{R}(AB) \le \mathscr{R}(B).$$

Proof: We can think of the product AB as

$$AB = \begin{bmatrix} \text{Row}^1(A) \\ \vdots \\ \text{Row}^m(A) \end{bmatrix} B = \begin{bmatrix} \text{Row}^1(A)B \\ \vdots \\ \text{Row}^m(A)B \end{bmatrix},$$

so that each row of AB is a linear combination of the rows of B. That is,

$$\left\{\text{Row}^i(AB) \ : \ i = 1,\dots,m\right\} \subset \mathscr{R}(B),$$

from which follows that

$$\mathscr{R}(AB) \le \mathscr{R}(B).$$

■

The following theorem connects the notion of row-equivalence to the row spaces of matrices:

Theorem 3.16 *Two matrices $A, B \in M_{m\times n}(\mathbb{F})$ are row-equivalent if and only if $\mathscr{R}(A) = \mathscr{R}(B)$. In particular, the row space of every matrix is equal to the row space of its row-reduced form.*

Proof: Recall that A and B are row-equivalent if and only if there exist matrices $P, Q \in M_m(\mathbb{F})$, such that

$$B = PA \qquad \text{and} \qquad A = QB.$$

By Lemma 3.15, if A and B are row-equivalent, then

$$\mathscr{R}(B) \leq \mathscr{R}(A) \qquad \text{and} \qquad \mathscr{R}(A) \leq \mathscr{R}(B),$$

hence $\mathscr{R}(A) = \mathscr{R}(B)$.
Conversely, if $\mathscr{R}(A) = \mathscr{R}(B)$, then every row of A is a linear combination of the rows of B and vice-versa, i.e., there exist matrices $P, Q \in M_m(\mathbb{F})$ such that

$$B = PA \qquad \text{and} \qquad A = QB,$$

hence they are row-equivalent. ∎

Exercises

(intermediate) 3.37 Consider the vector space $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ and the sets

$$S = \{(1,2,3), (2,2,1)\} \qquad \text{and} \qquad T = \{(2,3,-1), (3,0,-2)\}.$$

Is $\operatorname{Span} S = \operatorname{Span} T$?
Hint: find matrices A, B such that $\operatorname{Span} S = \mathscr{R}(A)$ and $\operatorname{Span} T = \mathscr{R}(B)$. Reduce these matrices and base your answer on those reduced representations.

3.3.5 The column space of a matrix

Let $A \in M_{m\times n}(\mathbb{F})$. The columns of A,

$$\{\mathrm{Col}_i(A) \ : \ i = 1, \dots, n\}$$

are a subset of $\mathbb{F}^m_{\text{col}}$, which is a vector space over $\mathbb{F}$. Their linear span is called the **column space** (**מרחב העמודות**) of A, denoted by

$$\mathscr{C}(A) = \operatorname{Span}\{\operatorname{Col}_i(A) \;:\; i = 1, \dots, n\}.$$

We can express linear combinations of the columns of A using matrix notation,

$$\sum_{i=1}^{n} c^i \operatorname{Col}_i(A) = \begin{bmatrix} \operatorname{Col}_1(A) & \dots & \operatorname{Col}_n(A) \end{bmatrix} \begin{bmatrix} c^1 \\ \vdots \\ c^n \end{bmatrix}.$$

Namely,

$$\mathscr{C}(A) = \{A\mathbf{c} \;:\; \mathbf{c} \in \mathbb{F}^n_{\text{col}}\}.$$

Example: Consider the case where $m < n$, and

$$A = \begin{bmatrix} 1 & & & & 0 & 0 \\ & 1 & & & 0 & 0 \\ & & \ddots & & \vdots & \vdots \\ & & & 1 & 0 & 0 \end{bmatrix}$$

Then

$$\mathscr{C}(A) = \mathbb{F}^m_{\text{col}}.$$

▲▲▲

Example: Consider the case where $m > n$, and

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Then,

$$\mathscr{C}(A) = \left\{\mathbf{x} \in \mathbb{F}^m_{\text{col}} \;:\; x^{n+1} = \dots = x^m = 0\right\}.$$

▲▲▲

Example: Let $A \in M_{2\times 2}(\mathbb{R})$ be given by

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then,

$$\mathscr{R}(A) = \operatorname{Span}\{[1,1]\} = \left\{\begin{bmatrix} c & c \end{bmatrix} \ : \ c \in \mathbb{R}\right\},$$

whereas

$$\mathscr{C}(A) = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} = \left\{\begin{bmatrix} c \\ 0 \end{bmatrix} \ : \ c \in \mathbb{R}\right\}.$$

▲▲▲

Lemma 3.17 *Let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times k}(\mathbb{F})$. Then,*

$$\mathscr{C}(AB) \le \mathscr{C}(A).$$

Proof: We can think of the product AB as

$$AB = A\begin{bmatrix} \operatorname{Col}_1(B) & \dots & \operatorname{Col}_n(B) \end{bmatrix} = \begin{bmatrix} A\operatorname{Col}_1(B) & \dots & A\operatorname{Col}_n(B) \end{bmatrix},$$

so that each column of AB is a linear combination of the columns of A, from which follows that

$$\mathscr{C}(AB) \le \mathscr{C}(A).$$

■

The following theorem connects the column space of a matrix with the consistency of associated non-homogeneous systems:

Theorem 3.18 *Let $A \in M_{m\times n}(\mathbb{F})$. Then, the non-homogenous system $A\mathbf{X} = \mathbf{b}$ is consistent if and only if*

$$\mathbf{b} \in \mathscr{C}(A).$$

Proof: If you think of it, there is nothing to prove. $\mathbf{b} \in \mathscr{C}(A)$ if and only if there exists an $\mathbf{x} \in \mathbb{F}^n_{\text{col}}$, such that $A\mathbf{x} = \mathbf{b}$, which by definition amounts to the system $A\mathbf{X} = \mathbf{b}$ being consistent. ■

3.3.6 The sum of linear subspaces

Definition 3.19 *Let V be a vector space over a field $\mathbb{F}$ and let $S_1, S_2, \dots, S_n$ be non-empty subset of V (not necessarily linear subspaces). We define their sum*

$$S_1 + S_2 + \cdots + S_n = \{\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n \ : \ \mathbf{v}_i \in S_i \ \forall i\}.$$

Be careful not to confuse $S_1 + S_2$ and $S_1 \cup S_2$.

Example: Let $V = (\mathbb{F}^2_{\text{col}}, +, \mathbb{F}, \cdot)$ and let

$$S_1 = \left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix} \right\} \subset V,$$

and

$$S_2 = \left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \end{bmatrix} \right\} \subset V,$$

Then,

$$S_1 \cup S_2 = \left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \end{bmatrix} \right\},$$

whereas

$$S_1 + S_2 = \left\{ \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 9 \\ 11 \end{bmatrix}, \begin{bmatrix} 6 \\ 8 \end{bmatrix}, \begin{bmatrix} 11 \\ 13 \end{bmatrix} \right\}.$$

▲▲▲

Proposition 3.20 *Let V be a vector space over a field $\mathbb{F}$ and let $W_1, W_2, \dots, W_n$ be linear subspaces of V. Then,*

$$W = W_1 + W_2 + \cdots + W_n$$

is a linear subspace of V. Furthermore,

$$W = \operatorname{Span}\left(\bigcup_{i=1}^{n} W_i \right).$$

Proof: We need to show that W is closed under linear combinations. Let $\mathbf{u}, \mathbf{v} \in W$. By definition, they can be written in the form

$$\mathbf{u} = a^1 \mathbf{u}_1 + \cdots + a^n \mathbf{u}_n \qquad \text{and} \qquad \mathbf{v} = b^1 \mathbf{v}_1 + \cdots + b^n \mathbf{v}_n,$$

where $\mathbf{u}_i, \mathbf{v}_i \in W_i$ for every $i = 1, \dots, n$. Then,

$$\mathbf{u} + \mathbf{v} = (a^1 \mathbf{u}_1 + b^1 \mathbf{v}_1) + \cdots + (a^n \mathbf{u}_n + b^n \mathbf{v}_n) \in W,$$

and for every $c \in \mathbb{F}$,

$$c\,\mathbf{u} = c\,a^1 \mathbf{u}_1 + \cdots + c\,a^n \mathbf{u}_n \in W,$$

thus proving that $W \le V$.

For the second part of the proposition, we observe that on the one hand, every $\mathbf{w} \in W$ is of the form

$$\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_n,$$

with $\mathbf{w}_i \in W_i$, proving that

$$W \le \operatorname{Span}\left(\bigcup_{i=1}^n W_i \right).$$

On the other hand, since $W \le V$ contains the union of the W_i's, it follows by Lemma 3.7 that

$$\left\langle \bigcup_{i=1}^n W_i \right\rangle \le W,$$

and by Theorem 3.13,

$$\operatorname{Span}\left(\bigcup_{i=1}^n W_i \right) \le W,$$

which completes the proof. ■

Exercises

(harder) 3.38 Let W_1, W_2 be linear subspaces of a vector space V, such that

$$W_1 + W_2 = V \qquad \text{and} \qquad W_1 \cap W_2 = \{0_V\}.$$

Prove that for every vector $\mathbf{v} \in V$ there exist *unique* vectors $\mathbf{w}_1 \in W_1$ and $\mathbf{w}_2 \in W_2$, such that

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2.$$

3.4 Bases and dimensions

In the previous section we considered the linear subspace generated (or equivalently, spanned) by a set of vectors. Clearly, the whole space spans itself. A question of interest is to characterize minimal sets of vector spanning the whole space, where minimality is in the sense that if any vector in the set is removed, then the set no longer spans the entire space. As we will see, if there exists a finite set of vectors spanning the space, then it is possible to define a dimension for that space, in the same sense as a line is one-dimensional and a plane is two-dimensional.

3.4.1 Linear dependence

Definition 3.21 *Let V be a vector space over $\mathbb{F}$. Let $S \subseteq V$. We say that a vector $\mathbf{v} \in V$ is* ***linearly-dependent*** *(תלוי ליניארית) on S if*

$$\mathbf{v} \in \operatorname{Span} S,$$

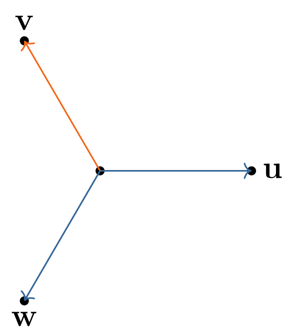
i.e., if we can compose $\mathbf{v}$ as a linear combination of vectors in S.

Example: It is always true that 0_V is linearly dependent on any set S (even empty), as 0_V is in the span of every subset. ▲▲▲

Example: Let $V = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$. Then, the vector $\mathbf{v} = (1,1,0)$ is linearly-dependent on $S = \{(1,0,0),(0,1,0)\}$; it is also linearly-dependent on $S = \{(1,0,0),(0,1,0),(0,0,1)\}$, but it not linearly-dependent on $\{(1,0,0),(0,0,1)\}$. Note also that $(1,0,0)$ is dependent on $S = \{(0,1,0),\mathbf{v}\}$. ▲▲▲

Example: Let $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ and consider the vectors

$$\mathbf{u} = (1,0) \qquad \mathbf{v} = (-1/2, \sqrt{3}/2) \qquad \text{and} \qquad \mathbf{w} = (-1/2, -\sqrt{3}/2).$$



Then, every vector is dependent on every two other. Furthermore, we note that

$$1 \cdot \mathbf{u} + 1 \cdot \mathbf{v} + 1 \cdot \mathbf{w} = 0.$$

▲▲▲

This last example motivates the following definition:

Definition 3.22 *Let V be a vector space over $\mathbb{F}$. A set (possibly infinite) $S \subseteq V$ is called **linearly-dependent** if there exists an $n \in \mathbb{N}$, a sequence of* <u>*distinct*</u> *vectors $(\mathbf{v}_1, \ldots, \mathbf{v}_n) \subset S$ and a sequence of scalars $(a^1, \ldots, a^n) \subset \mathbb{F}$* not all of which are $0_{\mathbb{F}}$, *such that*

$$a^1 \mathbf{v}_1 + \cdots + a^n \mathbf{v}_n + = 0_V.$$

*The set is called **linearly-independent** (בלתי תלוי ליניארית) if it is not linearly-dependent.*

Example: Let $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ and let

$$S = \{(1,0), (0,1), (1,1)\}.$$

This set is linearly-dependent because

$$1 \cdot (1,0) + 1 \cdot (0,1) + (-1) \cdot (1,1) = (0,0) = 0_V.$$

On the other hand, the set $\{(1,0), (0,1)\}$ is linearly-independent because

$$a^1(1,0) + a^2(0,1) = (a^1, a^2)$$

equals 0_V only if $a^1 = a^2 = 0$. ▲▲▲

Comment: Be aware of the difference between a **set** (קבוצה) and a **sequence** (סדרה). A set is a collection of elements with no notion of order among them; moreover, every element only counts once, e.g., $\{1\} \cup \{1\} = \{1\}$. A sequence, on the other hand, is an assignment of an element of a set to ordinal positions (first, second etc.). In particular, the same element may appear repeatedly in different positions of that sequence.

Comment: We can reformulate the properties of linear-dependence and linear-independence using matrix notation. Given a sequence of vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, we may express linear combinations of that sequence via multiplication by a column vector $\mathbf{c} \in \mathbb{F}^n_{\text{col}}$,

$$c^1\mathbf{v}_1 + \cdots + c^n\mathbf{v}_n = \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix} \mathbf{c}.$$

Then, a set of vectors S is linearly-independent if for every sequence of distinct elements $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ in S,

$$\begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix} \mathbf{c} = 0$$

if and only if $\mathbf{c} \in \mathbb{F}^n_{\text{col}}$ is the zero vector.

Proposition 3.23 *Let V be a vector space over $\mathbb{F}$. Let $S \subseteq V$. Then, the following statements are equivalent:*

(a) S is linearly-dependent.

(b) There exists a vector $\mathbf{v} \in S$ which is dependent on $S \setminus \{\mathbf{v}\}$.

Proof: (a) $\Rightarrow$ (b): Suppose that S is linearly-dependent. By definition, there exists a sequence of distinct vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ in S and a sequence of scalars $(a^1, \dots, a^n)$, not all zero, such that

$$a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = 0.$$

Let $j \in \{1, \dots, n\}$ be such that $a^j \neq 0$ (at least one such j exists). Then,

$$\mathbf{v}_j = -\sum_{i \neq j} (a^i/a_j)\mathbf{v}_i,$$

proving that $\mathbf{v}_j \in \operatorname{Span} S \setminus \{\mathbf{v}_j\}$.

(b) $\Rightarrow$ (a): Suppose that there exists a $\mathbf{v} \in S$, such that $\mathbf{v} \in \operatorname{Span} S \setminus \{\mathbf{v}\}$. That is, there exists a sequence of distinct vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset S \setminus \{\mathbf{v}\}$ and a sequence of scalars $(a^1, \dots, a^n)$, such that

$$\mathbf{v} = a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n.$$

Setting $\mathbf{v}_{n+1} = \mathbf{v}$ and $a^{n+1} = (-1)$, we obtain that $(\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{v}_{n+1})$ are distinct vectors in S satisfying

$$a^1\mathbf{v}_1 + \dots + a^n\mathbf{v}_n + a^{n+1}\mathbf{v}_{n+1} = 0,$$

i.e., S is linearly-dependent. ■

What makes a set $S \subseteq V$ linearly-independent? If for every sequence $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset S$ of distinct vectors and every sequence $(a^1, \dots, a^n)$ of scalars,

$$a^1\mathbf{v}_1 + \dots + a^n\mathbf{v}_n = 0$$

if and only if $a^1 = \dots = a^n = 0$.

Proposition 3.24 *Let V be a vector space over $\mathbb{F}$. Let $S \subset V$. Then, the following statements are equivalent:*

(a) S is linearly-independent.

(b) Every $\mathbf{v} \in S$ is linearly-independent of $S \setminus \{\mathbf{v}\}$.

Proof: This is just a reformulation of the previous proposition in negated form. ■

The following claims are quite immediate:

Proposition 3.25 *Let V be a vector space over $\mathbb{F}$.*

(a) A set containing a linearly-dependent subset is linearly-dependent.

(b) A subset of a linearly-independent set is linearly-independent.

(c) Any set containing 0_V is linearly-dependent.

(d) A set S is linearly-independent if and only if every finite subset of S is linearly-independent.

Proof:

(a) Suppose that $S \subset V$ contains a subset $T \subset S$ which is linearly-dependent. By definition, there exist distinct vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset T$ and scalars $(a^1, \dots, a^n) \subset \mathbb{F}$, such that

$$a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = 0.$$

Since $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset S$, it follows by definition that S is linearly-dependent.

(b) Let S be linearly-independent and let $T \subset S$ be a non-empty subset. If T was linearly-dependent, it would follows from the first item that S is linearly-dependent, which is a contradiction. Hence every non-empty subset of S is linearly-independent.

(c) Suppose that $0_V \in S$. Then, taking $n = 1$, $\mathbf{v}_1 = 0_V$ and $a^1 = 1_{\mathbb{F}}$, we obtain that

$$a^1\mathbf{v}_1 = 1_{\mathbb{F}} \cdot 0_V = 0_V,$$

hence S is linearly-dependent.

(d) By the second item, if S is linearly-independent, then *any* of its subsets, whether finite or not, is linearly-independent. Conversely, suppose that every subset of S is linearly independent. If S is linearly-dependent, then there exist distinct $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset S$ and scalars $(a^1, \dots, a^n) \subset \mathbb{F}$, not all of which are zero, such that

$$a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = 0.$$

This implies that the finite subset $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of S is linearly-dependent, in contradiction. Hence, S is linearly-independent.

■

Exercises

(easy) 3.39 Why did we insist in Definition 3.22 that the vectors $\mathbf{v}_i$ be distinct?

(easy) 3.40 Let $\mathbf{v} \in V$. Show that the set $\{\mathbf{v}\}$ is linearly-dependent if and only if $\mathbf{v} = 0_V$.

(easy) 3.41 Show that if two vectors are linearly-dependent, then one is a (scalar) multiple of the other.

(intermediate) 3.42 Find three vectors in $(\mathbb{R}^3,+,\mathbb{R},\cdot)$ that are *linearly – dependent*, but every pair of them is linearly-independent.

(intermediate) 3.43 Let $V = (\mathbb{R}^4,+,\mathbb{R},\cdot)$. Are the vectors

$$\mathbf{v}_1 = (1,1,2,4) \qquad \mathbf{v}_2 = (2,-1,-5,2)$$
$$\mathbf{v}_3 = (1,-1,-4,0) \qquad \mathbf{v}_4 = (2,1,1,6)$$

linearly-independent?

(intermediate) 3.44 Consider the set $\mathbb{C}^2$ and let

$$\mathbf{u} = (1-\imath, 3+\imath) \qquad \text{and} \qquad \mathbf{v} = (1, 1+2\imath).$$

(a) Is $\{\mathbf{u},\mathbf{v}\}$ linearly-dependent in $(\mathbb{C}^2,+,\mathbb{C},\cdot)$?
(b) Is $\{\mathbf{u},\mathbf{v}\}$ linearly-dependent in $(\mathbb{C}^2,+,\mathbb{R},\cdot)$?

(intermediate) 3.45 Let V be a vector space over $\mathbb{F}$ and let $U_1, U_2 \le V$ such that $U_1 \cap U_2 = \{0_V\}$. Let $L_1 \subseteq U_1$ and $L_2 \subseteq U_2$ be linearly-independent sets. Show that $L_1 \cup L_2$ is linearly-independent.

(intermediate) 3.46 In conjunction with the previous exercise, could we omit the condition that $U_1 \cap U_2 = \{0_V\}$?

(intermediate) 3.47 Consider the vector space $(\mathbb{F}^2_{\text{col}},+,\mathbb{F},\cdot)$. Show that the set

$$S = \left\{ \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} \right\}$$

is linearly-independent if and only if $ad - bc \ne 0$.

3.4.2 Bases

Definition 3.26 *Let V be a vector space over $\mathbb{F}$. A subset $S \subseteq V$ is called a **generating set** (קבוצה יוצרת) if*

$$\operatorname{Span} S = V.$$

*It is called a **basis** (בסיס) for V if it is a linearly-independent generating set. A vector space having a finite basis is called **finite-dimensional** (ממימד סופי), or **finitely-generated** (נוצר סופי). Otherwise, it is called **infinite-dimensional**.*

Example: Let $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$. Then,

$$S = \{(1,0), (0,1), (1,1)\} \subset V$$

is not a basis because it is linearly-dependent; on the other hand, it spans V, i.e., it is a generating set. Likewise,

$$T = \{(1,0)\} \subset V$$

is not a basis because it does not span V; for example,

$$(1,1) \notin \operatorname{Span} T.$$

▲ ▲ ▲

Example: Let $V = (\mathbb{R}^n, +, \mathbb{R}, \cdot)$ for some $n \in \mathbb{N}$. The set of vectors

$$\begin{aligned} \mathbf{e}_1 &= (1,0,0,\dots,0,0) \\ \mathbf{e}_2 &= (0,1,0,\dots,0,0) \\ \vdots &= \vdots \\ \mathbf{e}_n &= (0,0,0,\dots,0,1) \end{aligned}$$

is a basis called the **standard basis** (הבסיס הסטנדרטי). We leave it as an exercise to prove that this is indeed a basis. ▲ ▲ ▲

Example: Consider the vector space $(\mathbb{C}, +, \mathbb{C}, \cdot)$. Then,

$$S = \{1\}$$

is a basis. Why? A singleton containing a non-zero vector is always linearly-independent. On the other hand, $\operatorname{Span} S = \mathbb{C}$, as every $z \in \mathbb{C}$ can be written is

$$z = z \cdot 1,$$

where z on the left-hand side is viewed as a vector, whereas z on the right-hand side is viewed as a scalar. In fact, $\{1\}$ is a basis for every field viewed as a vector space over itself. ▲ ▲ ▲

Example: Consider the vector space $(\mathbb{C}, +, \mathbb{R}, \cdot)$. Then,

$$S = \{1\}$$

is *not* a basis because for example, $\imath \notin \operatorname{Span}\{1\}$. On the other hand,

$$T = \{1, \imath\}$$

is a basis. I recommend thinking again about the difference between the last two examples. ▲▲▲

Example: Let $A \in \mathrm{GL}_n(\mathbb{F})$ and consider

$$S = \{\mathrm{Col}_i(A) \ : \ i = 1, \dots, n\} \subset \mathbb{F}^n_{\mathrm{col}}.$$

We claim that S is a basis for $V = \mathbb{F}^n_{\mathrm{col}}$. There are two things to show: that S is a linearly-independent set and that S spans $\mathbb{F}^n_{\mathrm{col}}$ (i.e., that the column space of A is $\mathbb{F}^n_{\mathrm{col}}$).

Let $\mathbf{x} \subset \mathbb{F}^n_{\mathrm{col}}$ be such that

$$\sum_{i=1}^n x^i \,\mathrm{Col}_i(A) = 0_V.$$

Noting that for every $j = 1, \dots, n$,

$$\left(\sum_{i=1}^n x^i \,\mathrm{Col}_i(A) \right)^j = \sum_{i=1}^n a^j_i x^i = (A\mathbf{x})^j,$$

it follows that $A\mathbf{x} = 0_V$. Since A is invertible, it follows that $\mathbf{x} = 0_V$, proving that S is linearly-independent.

It remains to show that S spans V. Let $\mathbf{c} \in V$. Since A is invertible, the linear system $A\mathbf{x} = \mathbf{c}$ is solvable, i.e., there exists a linear combination of the columns of A equal to $\mathbf{c}$, proving that S spans V. ▲▲▲

Example: Thus far, all the vector spaces in this section were finitely-generated. Consider now the vector space of polynomials $\mathbb{F}[X]$. This space is infinite-dimensional. Why? Let $P_1, \dots, P_n \in \mathbb{F}[X]$ be a finite set of polynomials; we will show that it cannot span $\mathbb{F}[X]$. Let

$$N = \max_{i=1}^n \deg P_i.$$

Then, each P_i can be written as

$$P_i = \sum_{j=1}^{N} c_{ij} X^j,$$

where some of the c_{ij} may be zero. For every scalars $a^1, \dots, a^n$,

$$\sum_{i=1}^{n} a^i P_i = \sum_{j=1}^{N} \left(\sum_{i=1}^{n} a^i c_{ij} \right) X^j,$$

It follows, for example, that X^{N+1} is not in the span of $\{P_1, \dots, P_n\}$. ▲▲▲

We next provide two additional characterization to bases.

Definition 3.27 *Let V be a vector space over $\mathbb{F}$. A subset $L \subset V$ is called* ***maximally linearly-independent*** *(בלתי תלויה מקסימלית) if it is linearly-independent, and for every $\mathbf{v} \in V \setminus L$, $L \cup \{\mathbf{v}\}$ is linearly-dependent.*

Proposition 3.28 *Every maximally linearly-independent set is a basis.*

Proof: Let $L \subset V$ be maximally linearly-independent. In order to show that it is a basis, it only remains to show that it is a generating set. Suppose it wasn't a generating set. By definition, there exist a $\mathbf{v} \notin \operatorname{Span} L$. We claim that $L \cup \{\mathbf{v}\}$ is linearly-independent, in contradiction to L being maximally linearly-independent. Indeed, if $L \cup \{\mathbf{v}\}$ was linearly-dependent, there would exist $(\mathbf{v}_1, \dots, \mathbf{v}_n) \subset L$, $(a^1, \dots, a^n) \subset \mathbb{F}$ and $a \in \mathbb{F}$, such that $(a^1, \dots, a^n, a)$ are not all zero, and

$$a^1\mathbf{v}_1 + \dots + a^n\mathbf{v}_n + a\mathbf{v} = 0_V.$$

We argue that $a \neq 0$, for otherwise, it would imply that L is not linearly-independent. Thus,

$$\mathbf{v} = \sum_{i=1}^{n} (-a^i/a)\mathbf{v}_i,$$

in contradiction to $\mathbf{v} \notin \operatorname{Span} L$. We conclude that $V = \operatorname{Span} L$, hence L is a basis. ■

Definition 3.29 *Let V be a vector space over $\mathbb{F}$. A subset $G \subset V$ is called* ***minimally-generating*** *(יוצרת מינימלית) if it is generating, and for every $\mathbf{v} \in G$, $G \setminus \{\mathbf{v}\}$ is not generating.*

Proposition 3.30 *Every minimally-generating set is a basis.*

Proof: Let $G \subset V$ be minimally-generating. Since it is a generating set, it remains to prove that it is linearly-independent. Suppose it weren't linearly-independent. This implies the existence of a $\mathbf{v} \in G$, such that

$$\mathbf{v} = a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n$$

for some $(\mathbf{v}_1, \ldots, \mathbf{v}_n) \subset G \smallsetminus \{\mathbf{v}\}$ and $(a^1, \ldots, a^n) \subset \mathbb{F}$. We claim that $G \smallsetminus \{\mathbf{v}\}$ is a generating set in contradiction to the minimality of G. Indeed, since G is a generating set, every $\mathbf{u} \in V$ can be written in the form

$$\mathbf{u} = b^1\mathbf{u}_1 + \cdots + b^m\mathbf{u}_m,$$

for some $(\mathbf{u}_1, \ldots, \mathbf{u}_m) \subset G$ and $(b^1, \ldots, b^m) \subset \mathbb{F}$. Now either the $\mathbf{u}_i$ do not comprise $\mathbf{v}$, in which case

$$\mathbf{u} \in \operatorname{Span} G \smallsetminus \{\mathbf{v}\},$$

or, if one of the $\mathbf{u}_i$ equals $\mathbf{v}$, we can substitute for $\mathbf{v}$ its expression as a linear combination of element in $G \smallsetminus \{\mathbf{v}\}$, so in either case

$$\mathbf{u} \in \operatorname{Span} G \smallsetminus \{\mathbf{v}\},$$

showing that $G \smallsetminus \{\mathbf{v}\}$ is a generating set—contradiction. ■

Proposition 3.31 *Let V be a vector space over $\mathbb{F}$. Let $G \subset V$ be a finite generating set and let $L \subseteq G$ be linearly-independent. Then, there exists a basis $\mathcal{B}$ for V, such that*

$$L \subseteq \mathcal{B} \subseteq G.$$

In other words, every linearly-independent set which is partial to a generating set, can be expanded into a basis contained that generating set. Alternatively, every generating set containing a linearly-independent set can be reduced to a basis containing that linearly-independent set.

Proof: Start with L, and add to it vectors in G, as long as the set remains linearly-independent. This process must terminate, as G is a finite set. Consider the resulting set $L \subset \mathcal{B} \subset G$. By construction, $\mathcal{B}$ is linearly-independent, and for every $\mathbf{v} \in G \smallsetminus \mathcal{B}$ we obtain that $G \cup \{\mathbf{v}\}$ is linearly-dependent. It follows that every such $\mathbf{v}$ is in the span of $\mathcal{B}$, i.e.,

$$G \smallsetminus \mathcal{B} \subset \operatorname{Span} \mathcal{B},$$

from which follows at once that

$$G \subset \operatorname{Span} \mathcal{B},$$

hence

$$V = \operatorname{Span} G \subset \operatorname{Span} \mathcal{B} \subset V,$$

i.e., $\mathcal{B}$ is a generating set, hence a basis. ■

Corollary 3.32 *Every finitely-generated vector space has a basis (which in particular is finite).*

Proof: Apply the previous proposition with $L = \varnothing$. ■

As it turns out, every vector space has bases; to show it in the general case is much more involved, and relies on a fundamental axiom of set theory called the **axiom of choice** (אקסיומת הבחירה). You are welcome to read in Wikipedia about **Hamel bases**.

Exercises

(easy) 3.48 Prove that for every field $\mathbb{F}$, $\{1_{\mathbb{F}}\}$ is a basis for $(\mathbb{F}, +, \mathbb{F}, \cdot)$.

(easy) 3.49 Find a basis for $(\mathbb{C}^2, +, \mathbb{C}, \cdot)$.

(easy) 3.50 Find a basis for $(\mathbb{C}^2, +, \mathbb{R}, \cdot)$.

(intermediate) 3.51 Find a basis for the subspace of $\mathbb{R}^4$ spanned by the four vectors of Exercise 3.43.

(intermediate) 3.52 Show that the vectors

$$\mathbf{v}_1 = (1, 0, -1) \qquad \mathbf{v}_2 = (1, 2, 1) \qquad \mathbf{v}_3 = (0, -3, 2)$$

form a basis for $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$. Write each of the standard basis vectors as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

(intermediate) 3.53 Let $V = (M_{2\times 2}(\mathbb{F}), +, \mathbb{F}, \cdot)$ and consider the subsets

$$W_1 = \left\{ \begin{bmatrix} a & -a \\ b & c \end{bmatrix} \ : \ a, b, c \in \mathbb{F} \right\},$$

and

$$W_2 = \left\{ \begin{bmatrix} a & b \\ -a & c \end{bmatrix} \ : \ a, b, c \in \mathbb{F} \right\}.$$

(a) Prove that $W_1, W_2 \leq V$.

(b) Prove that $W_1 + W_2 \leq V$ (repeat the proof which was given for the general case).

(c) Find bases for W_1, W_2, $W_1 + W_2$ and $W_1 \cap W_2$.

(intermediate) 3.54 Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \in M_3(\mathbb{R}).$$

This matrix defines two subspaces of $\mathbb{R}^3_{\text{col}}$: the column space and the space of solutions to $A\mathbf{X} = 0$. Find a basis for each subspace. Show that those subspaces are different.

(intermediate) 3.55 Let S be a non-empty finite set and let $V = \operatorname{Func}(S, \mathbb{F})$ (we saw that this is a vector space over $\mathbb{F}$). For each $t \in S$, denote by $f_t : S \to \mathbb{F}$ the function defined by

$$f_t(s) = \begin{cases} 1 & s = t \\ 0 & \text{otherwise.} \end{cases}$$

Show that

$$\{f_t \ : \ t \in S\}$$

is a basis for V.

(harder) 3.56 Let V be a vector space over $\mathbb{F}$. Let $G \subseteq V$ be a generating set and let $L \subseteq V$ be linearly-independent. Show that for every $\mathbf{u} \in L \setminus G$ there exists a $\mathbf{v} \in G \setminus L$ such that

$$(G \setminus \{\mathbf{v}\}) \cup \{\mathbf{u}\}$$

is generating, and

$$(L \setminus \{\mathbf{u}\}) \cup \{\mathbf{v}\}$$

is linearly-independent. This fact is known as the exchange lemma.

(harder) 3.57 Let V be a vector space over $\mathbb{F}$. A proper subspace $W < V$ is called a **hyperplane** (**על מישור**) if for every subspace $W \leq U \leq V$ either $U = W$ or $U = V$. Show that if V is finitely-generated, $\dim_{\mathbb{F}} V \geq 2$, then there exists a maximal hyperplane $0 < W < V$. Show that for every $\mathbf{v} \in V \setminus W$,

$$V = \mathbb{F}\mathbf{v} + W.$$

3.4.3 The dimension of a vector space

We are now in measure to define the dimension of a finitely-generated vector space:

Proposition 3.33 *Let V be a finitely-generated vector space over $\mathbb{F}$. Let $G = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$ be a (finite) generating set for V. Then, any linearly-independent set of vectors has no more than n elements.*

Proof: Let $S \subset V$ have more than n elements, and let $\{\mathbf{u}_1, \dots, \mathbf{u}_{n+1}\} \subset S$ be distinct vectors. By the definition of a generating set, each of the vectors $\mathbf{u}_i$ is in the span of that G, hence there exist $(n+1) \times n$ scalars a^i_j such that

$$\mathbf{u}_j = a^1_j \mathbf{v}_1 + \dots + a^n_j \mathbf{v}_n \qquad \text{for every } j = 1, \dots, n+1.$$

(Since G is finite, we may always consider vectors in its span as a linear combination of all $\mathbf{v}_i$, with some coefficients being possibly zero.) For any sequence of scalars $(c^1, \dots, c^{n+1})$,

$$\sum_{j=1}^{n+1} c^j \mathbf{u}_j = \sum_{j=1}^{n+1} c^j \left(\sum_{i=1}^{n} a^i_j \mathbf{v}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n+1} a^i_j c^j \right) \mathbf{v}_i.$$

Consider the $n \times (n+1)$ matrix A whose entries are a^i_j. It represents a system of equations having more variables than equations. We know that for such a system, the homogeneous equation has non-trivial solutions. That is, there exists a $\mathbf{c} \neq 0_{\mathbb{F}^{n+1}_{\text{col}}}$, such that

$$a^i_1 c^1 + \cdots + a^i_{n+1} c^{n+1} = 0 \qquad \text{for all } i = 1, \ldots, n.$$

For that $\mathbf{c}$,

$$c^1 \mathbf{u}_1 + \cdots + c^{n+1} \mathbf{u}_{n+1} = 0_V,$$

proving that the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\} \subset S$ is linearly-dependent, hence so is S (which contains a linearly-dependent set). It follows that any linearly-independent set of vectors contains at most n vectors.

We may rewrite this proof in matrix form. Let the *sequence* $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an ordered n-tuple containing the vectors in G, and let $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$, $m > n$, be any sequence of m vectors. Since the $\{\mathbf{v}_i\}$ are a generating set, there exists an $n \times m$ matrix A, such that

$$\begin{pmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_m \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{pmatrix} A.$$

Since A has more columns than rows, there exists a non-zero $\mathbf{c} \in \mathbb{F}^m_{\text{col}}$ such that $A\mathbf{c} = 0$, i.e.,

$$\begin{pmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_m \end{pmatrix} \mathbf{c} = \begin{pmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{pmatrix} A\mathbf{c} = 0_V,$$

proving that the vectors $\{\mathbf{u}_i\}$ are linearly-dependent. ■

Corollary 3.34 *Let V be a finitely-generated vector space over $\mathbb{F}$. Then, every two bases have the same number of elements.*

Proof: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be two bases for V. By the previous theorem, since bases are by definition generating sets and linearly-independent, $m \leq n$ and $n \leq m$, which completes the proof. ■

This last corollary implies that we can associate with every non-zero finitely-generated vector space a natural number which is the cardinality of any of its bases. This number is called the **dimension** (מימד) of V, and is denoted

$$\dim_{\mathbb{F}} V.$$

Note the explicit mention of the field $\mathbb{F}$, as the same set of vectors may constitute a vector space of different dimension depending on the field. The zero space (which contains no independent sets of vectors) is assigned dimension zero,

$$\dim_{\mathbb{F}}\{0_V\} = 0.$$

If follows from the last theorem that if $\dim V = n$, then every set of vectors containing more than n elements is linearly-dependent, and that no set of vectors containing fewer than n elements can span V (see exercises).

Example: The vector space $(\mathbb{C}, +, \mathbb{C}, \cdot)$ has dimension 1 (because $\{1\}$ is a basis),

$$\dim_{\mathbb{C}} \mathbb{C} = 1.$$

The vector space $(\mathbb{C}, +, \mathbb{R}, \cdot)$ has dimension 2 (because $\{1, \imath\}$ is a basis),

$$\dim_{\mathbb{R}} \mathbb{C} = 2.$$

Generally, for any field $\mathbb{F}$, the vector space $(\mathbb{F}^n, +, \mathbb{F}, \cdot)$ has dimension n,

$$\dim_{\mathbb{F}} \mathbb{F}^n = n,$$

because the standard basis has n elements. ▲▲▲

Lemma 3.35 *Let V be a vector space over $\mathbb{F}$. Let $S \subset V$ be linearly-independent and let $\mathbf{v} \notin \operatorname{Span} S$. Then, $S \cup \{\mathbf{v}\}$ is linearly-independent.*

Proof: This was essentially proved in Proposition 3.28, but for the sake of completeness, we repeat the proof. Suppose, by contradiction that $S \cup \{\mathbf{v}\}$ is linearly-dependent. Then, there exist distinct vectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in S$ and scalar $c^1, \dots, c^n, c$, not all zero, such that

$$c^1\mathbf{v}_1 + \dots + c^n\mathbf{v}_n + c\mathbf{v} = 0_V.$$

If $c = 0$, then this contradicts the linear-independence of S. If on the other hand $c \neq 0$, then

$$\mathbf{v} = \sum_{i=1}^{n} (-c^i/c)\mathbf{v}_i,$$

in contradiction to $\mathbf{v}$ not being in the span of S. ■

Proposition 3.36 *Let V be a finitely-generated vector space over $\mathbb{F}$. Let $W \le V$ be a linear subspace. Then, every linearly-independent subset $S \subseteq W$ is part of a basis for W. In particular, since $V \le V$, every basis for W can be extended to a basis for V.*

Proof: If S spans W then it is a basis for W and we are done. Otherwise, there exists a vector

$$\mathbf{v}_1 \in W \smallsetminus \operatorname{Span} S.$$

By the previous lemma, $S \cup \{\mathbf{v}_1\}$ is linearly-independent. If it spans W then it is a basis for W and we are done. Otherwise, there exists a vector

$$\mathbf{v}_2 \in W \smallsetminus \operatorname{Span}(S \cup \{\mathbf{v}_1\}).$$

By the previous lemma, $S \cup \{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly-independent. We proceed inductively. Eventually, after no more than $\dim_{\mathbb{F}} V$ steps (because there exist at most $\dim_{\mathbb{F}} V$ linearly-independent vectors), we obtain a basis for W containing S. ■

Corollary 3.37 *Let V be a finitely-generated vector space over $\mathbb{F}$. Let $W < V$ be a* proper *linear subspace. Then,*

$$\dim_{\mathbb{F}} W < \dim_{\mathbb{F}} V.$$

(In particular, W is finitely-generated.)

Proof: If $W = \{0_V\}$ then $\dim_{\mathbb{F}} W = 0$ and we are done. Otherwise, there exists a non-zero $\mathbf{w} \in W$. By the previous proposition and its proof, there exists a basis S for W containing $\mathbf{w}$ and having no more than $\dim V$ elements, hence

$$\dim_{\mathbb{F}} W \le \dim_{\mathbb{F}} V.$$

Since $W < V$, there exists a vector $\mathbf{v} \in V \smallsetminus W$, hence not in the span of S. It follows that $S \cup \{\mathbf{v}\}$ is linearly-independent (as a collection of vectors in V), hence a basis for V contains at least $\dim_{\mathbb{F}} W + 1$ vectors, from which we conclude that $\dim_{\mathbb{F}} W < \dim_{\mathbb{F}} V$. ■

In fact, the following holds:

Corollary 3.38 *Let V be a finitely-generated vector space of dimension n and let $W \le V$. Then,*

$$\dim_{\mathbb{F}} W = \dim_{\mathbb{F}} V \qquad \textit{if and only if} \qquad W = V.$$

Proof: One direction is immediate. For the other direction, suppose that $\dim_{\mathbb{F}} W = \dim_{\mathbb{F}} V$. If $W < V$, then there exists a $\mathbf{v} \in V \smallsetminus W$. Let L be a maximally-independent set for W; then $L \cup \{\mathbf{v}\}$ is independent, proving that L is not maximally-independent for V, hence $\dim_{\mathbb{F}} W < \dim_{\mathbb{F}} V$, which is a contradiction. ∎

Finally, a statement reminiscent of the inclusion-exclusion principle:

Proposition 3.39 *Let W_1, W_2 be finitely-generated linear subspaces of a vector space V. Then, the linear subspace $W_1 + W_2$ is finitely-generated and*

$$\dim_{\mathbb{F}}(W_1 + W_2) = \dim_{\mathbb{F}} W_1 + \dim_{\mathbb{F}} W_2 - \dim_{\mathbb{F}}(W_1 \cap W_2).$$

Comment: The example you should have in mind is $V = \mathbb{R}^3$, W_1 being the xy-plane and W_2 being the xz-plane. Then, $W_1 + W_2 = \mathbb{R}^3$ and $W_1 \cap W_2$ is the x-axis. In this case,

$$\underbrace{\dim_{\mathbb{R}}(W_1 + W_2)}_{=3} = \underbrace{\dim_{\mathbb{R}} W_1}_{=2} + \underbrace{\dim_{\mathbb{R}} W_2}_{=2} - \underbrace{\dim_{\mathbb{R}}(W_1 \cap W_2)}_{=1}.$$

Proof: Note that V may not be finitely-generated, but since $W_1 \cap W_2 \le W_1, W_2$, it follows from Corollary 3.37 that $W_1 \cap W_2$ is finitely-generated. Let

$$\dim_{\mathbb{F}}(W_1 \cap W_2) = k \qquad \dim_{\mathbb{F}} W_1 = k + n \qquad \text{and} \qquad \dim_{\mathbb{F}} W_2 = k + m$$

(a priori, k, m and n may be zero). Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis for $W_1 \cap W_2$. By Proposition 3.36, it is part of a basis

$$\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \cup \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

for W_1 and it is part of a basis

$$\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \cup \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$$

for W_2. Clearly, the set

$$S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\} \cup \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \cup \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$$

spans $W_1 + W_2$ (convince yourself that this is true). If we show that S is also linearly-independent then, by definition, it is a basis for $W_1 + W_2$, in which case $\dim_{\mathbb{F}}(W_1 + W_2) = k + m + n$, proving the claim.

Suppose, by contradiction, that S is dependent. This implies that there exist scalars $a^1, \dots, a^k$, $b^1, \dots, b^n$ and $c^1, \dots, c^m$, not all of which are zero, such that

$$\sum_{i=1}^{k} a^i \mathbf{u}_i + \sum_{i=1}^{n} b^i \mathbf{v}_i + \sum_{i=1}^{m} c^i \mathbf{w}_i = 0_V.$$

Thus,

$$\sum_{i=1}^{m} c^i \mathbf{w}_i = -\sum_{i=1}^{k} a^i \mathbf{u}_i - \sum_{i=1}^{n} b^i \mathbf{v}_i.$$

The left-hand side is in W_2, whereas the right-hand side is in W_1. Thus, both sides are in $W_1 \cap W_2$. Moreover, they can't be zero, otherwise either the $\{\mathbf{w}_i\}$ or the $\{\mathbf{u}_i\} \cup \{\mathbf{v}_i\}$ are not linearly-independent (recall that at least one of the coefficients is non-zero). Thus, we conclude that at least one of the $\{c^i\}$ is non-zero.

Since the vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ form a basis for $W_1 \cap W_2$, it follows that there exist scalars $d^1, \dots, d^k$, such that

$$\sum_{i=1}^{m} c^i \mathbf{w}_i = \sum_{i=1}^{k} d^i \mathbf{u}_i,$$

or

$$\sum_{i=1}^{m} c^i \mathbf{w}_i - \sum_{i=1}^{k} d^i \mathbf{u}_i = 0,$$

contradicting the fact that the vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \cup \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ are linearly-independent. Hence, S is linearly-independent, and therefore a basis. ■

We end this section with a very important theorem:

Theorem 3.40 *Let $A \in M_n(\mathbb{F})$. Then, A is invertible if and only if its rows form a linearly-independent set in $\mathbb{F}^n_{row}$.*

Proof: Suppose that the rows of A form a linearly-independent set in $\mathbb{F}^n_{\text{row}}$. Since $\dim_{\mathbb{F}} \mathbb{F}^n_{\text{row}} = n$, it follows that the rows of A form a basis, and in particular are a generating set. Thus, there exists for every $i = 1, \dots, n$ a vector $[x^i_1, \dots, x^i_n]$, such that

$$\begin{bmatrix} x^i_1 & \cdots & x^i_n \end{bmatrix} \underbrace{\begin{bmatrix} \text{Row}^1(A) \\ \cdots \\ \text{Row}^n(A) \end{bmatrix}}_{A} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix},$$

where the right-hand side is a vector of zeros except for a 1 in the i-th column. Then,

$$\begin{bmatrix} x^1_1 & \cdots & x^1_n \\ \vdots & \vdots & \vdots \\ x^n_1 & \cdots & x^n_n \end{bmatrix} \underbrace{\begin{bmatrix} \text{Row}^1(A) \\ \cdots \\ \text{Row}^n(A) \end{bmatrix}}_{A} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix},$$

proving that A is invertible. Conversely, if A is invertible, let $B = A^{-1}$. Suppose that $\mathbf{c} \in \mathbb{F}^n_{\text{row}}$ satisfies $\mathbf{c}A = 0_{\mathbb{F}^n_{\text{row}}}$. Then,

$$\mathbf{c} = \mathbf{c}\, I = \mathbf{c}\,(AB) = (\mathbf{c}\, A)B = 0_{\mathbb{F}^n_{\text{row}}},$$

i.e., the only vanishing linear combination of the rows of A is the trivial one, proving that the rows of A are linearly-independent. ∎

Exercises

(easy) 3.58 Show that the vector space $(M_{2\times 2}(\mathbb{F}), +, \mathbb{F}, \cdot)$ has dimension four. More generally, show that the vector space $(M_{m\times n}(\mathbb{F}), +, \mathbb{F}, \cdot)$ has dimension mn.

(easy) 3.59 Let V be a vector space of dimension 3. Show that if $U, W \le V$ with $\dim_{\mathbb{F}} U = \dim_{\mathbb{F}} W = 2$, then $U \cap W \ne \{0_V\}$.

(intermediate) 3.60 Let $A \in M_n(\mathbb{F})$. Show that A is invertible if and only if its columns form a linearly-independent set in $\mathbb{F}^n_{\text{col}}$.

(intermediate) 3.61 Find a basis for the vector subspace

$$W = \left\{ \mathbf{x} \in \mathbb{R}^5 \quad : \quad \begin{matrix} x^1 + x^2 + x^3 + x^4 = 0 \\ x^3 + x^4 + x^5 = 0 \\ x^1 + x^2 + x^5 = 0 \\ x^1 + x^3 = 0 \end{matrix} \right\}$$

What is its dimension?

(intermediate) 3.62 Let V be a finitely-generated vector space and let $U, W \le V$. Which of the following assertions is true? Prove or find a counter example.

(a) If $2 + \dim_{\mathbb{F}} V \le \dim_{\mathbb{F}} U + \dim_{\mathbb{F}} W$ then $V = U + W$.
(b) If $2 \dim_{\mathbb{F}} V \le \dim_{\mathbb{F}} U + \dim_{\mathbb{F}} W$ then $V = U + W$.
(c) If $\dim_{\mathbb{F}} V > \dim_{\mathbb{F}} U + \dim_{\mathbb{F}} W$ then $V \ne U + W$.
(d) If $\dim_{\mathbb{F}} V > \dim_{\mathbb{F}} U + \dim_{\mathbb{F}} W$ then $U \cap W = \{0\}$.

(intermediate) 3.63 Consider the linear subspaces of $\mathbb{R}^4$,

$$U = \operatorname{Span}\{(1, 0, -1, -2), (-1, -1, 0, 2), (1, 2, 1, -1)\}$$

and

$$W = \left\{ \mathbf{x} \in \mathbb{R}^4 \quad : \quad \begin{matrix} x^1 + 3x^2 + x^3 - x^4 = 0 \\ x^2 - 3x^3 + 2x^4 = 0 \end{matrix} \right\}.$$

What is $\dim_{\mathbb{R}}(U + W)$?

(harder) 3.64 Let V be a vector space over $\mathbb{F}$, such that

$$\dim_{\mathbb{F}} V = n.$$

Show that any set of vectors containing less than n vectors does not span V.

(harder) 3.65 Consider the vector space $(\mathbb{R}, +, \mathbb{Q}, \cdot)$ (i.e. the vectors are real numbers, the scalars are rational numbers, with the operations of vector addition and scalar multiplication defined as usual in $\mathbb{R}$). Prove that this vector space is not finitely-generated. (Hint: start by convincing yourself that $\{1\}$ is not a basis for this space; the argument is based on the fact that $\mathbb{Q}$ is countable, whereas $\mathbb{R}$ is not.)

(harder) 3.66 Let V be a vector space of dimension n. What is the maximal m for which there exists linear subspaces

$$W_0 < W_1 < \cdots < W_m\,?$$

3.4.4 The rank of a matrix

Let $A \in M_{m\times n}(\mathbb{F})$. We defined for a matrix two vector spaces,

$$\begin{aligned}\mathscr{R}(A) &= \operatorname{Span}\{\operatorname{Row}^i(A) \;:\; i = 1,\dots,m\} \subseteq \mathbb{F}^n_{\text{row}}\\ \mathscr{C}(A) &= \operatorname{Span}\{\operatorname{Col}_j(A) \;:\; j = 1,\dots,n\} \subseteq \mathbb{F}^m_{\text{col}}.\end{aligned}$$

The **row-rank** (**דרגה לפי שורות**) of a matrix is the dimension of its row space, whereas its **column-rank** (**דרגה לפי עמודות**) is the dimension of its column space. Even though these two spaces are seemlingly unrelated, it turns out that

$$\dim_{\mathbb{F}} \mathscr{R}(A) = \dim_{\mathbb{F}} \mathscr{C}(A).$$

This joint dimension is called the **rank** (**דרגה**) of the matrix A.

We start with the row space:

Proposition 3.41 *Let R be the row-reduced form of A. Then,*

$$\dim_{\mathbb{F}} \mathscr{R}(A)$$

equals the number of non-zero rows in R.

Proof: Lemma 3.15 shows that

$$\mathscr{R}(A) = \mathscr{R}(R).$$

Let p be the number of non-zero rows in R. Then, the row space of R (equivalently A) is spanned by the first p rows of R. These p rows are linearly-independent because each an entry $r_{j,k_j} = 1$, while $r_{i,k_j} = 0$ for all $i \neq j$. It follows that

$$\dim_{\mathbb{F}} \mathscr{R}(R) = p.$$

■

We proceed with the column space: let $R = PA$ with $P \in \mathrm{GL}_m(\mathbb{R})$. We have seen that the non-homogeneous system $A\mathbf{X} = \mathbf{b}$ is solvable if and only if $\mathbf{b} \in \mathscr{C}(A)$. But this system is solvable if and only if the system

$$R\mathbf{X} = PA\mathbf{X} = P\mathbf{b}$$

is solvable, i.e., if and only if $P\mathbf{b} \in \mathscr{C}(R)$. That is,

$$\mathbf{b} \in \mathscr{C}(A) \qquad \text{if and only if} \qquad P\mathbf{b} \in \mathscr{C}(R).$$

Proposition 3.42 *The column-rank of a matrix equals that of its row-reduced form.*

Proof: Let $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ be an ordered basis for $\mathscr{C}(A)$. If we show that $(P\mathbf{v}_1, \ldots, P\mathbf{v}_p)$ is an ordered basis for $\mathscr{C}(R)$ then we are done. Let $\mathbf{w} \in \mathscr{C}(R)$. Then, $P^{-1}\mathbf{w} \in \mathscr{C}(A)$, and there exist scalars $(a^1, \ldots, a^p)$ such that

$$P^{-1}\mathbf{w} = a^1\mathbf{v}_1 + \cdots + a^p\mathbf{v}_p,$$

from which we deduce that

$$\mathbf{w} = P(a^1\mathbf{v}_1 + \cdots + a^p\mathbf{v}_p) = a^1P\mathbf{v}_1 + \cdots + a^pP\mathbf{v}_p,$$

proving that $(P\mathbf{v}_1, \ldots, P\mathbf{v}_p)$ generates $\mathscr{C}(R)$.
Suppose then

$$a^1P\mathbf{v}_1 + \cdots + a^pP\mathbf{v}_p = 0_{\mathbb{F}^m_{\mathrm{col}}}.$$

It follows that

$$P^{-1}(a^1P\mathbf{v}_1 + \cdots + a^pP\mathbf{v}_p) = a^1\mathbf{v}_1 + \cdots + a^p\mathbf{v}_p = 0_{\mathbb{F}^m_{\mathrm{col}}},$$

proving that $(P\mathbf{v}_1, \ldots, P\mathbf{v}_p)$ is independent, hence a basis for $\mathscr{C}(R)$. ■

Proposition 3.43 *Let R be the row-reduced form of A. Then,*

$$\dim_{\mathbb{F}} \mathscr{C}(A)$$

equals the number of non-zero rows in R.

Proof: Let p be the number of non-zero rows in R. By the previous proposition it suffices to show that $\dim_{\mathbb{F}} \mathscr{C}(R) = p$. Take the p columns $\mathrm{Col}_{k_j}(R)$, $j = 1, \dots, p$. They are independent and spanning because the form together the unit matrix. ■

Example: Consider the matrix

$$A = \begin{bmatrix} 0 & 0 & 1 & 4 \\ 2 & 4 & 2 & 6 \\ 3 & 6 & 2 & 5 \end{bmatrix}$$

You may verify that

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{R} = \underbrace{\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -2 & -3 & -2 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 0 & 0 & 1 & 4 \\ 2 & 4 & 2 & 6 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A},$$

and that

$$\underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 2 & 1 \end{bmatrix}}_{Q=P^{-1}} \underbrace{\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{R} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 4 \\ 2 & 4 & 2 & 6 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A}.$$

Consider first the row space of R. It is spanned by two non-zero rows, hence its dimension is at most 2; it is in fact equal to 2, because

$$a \begin{bmatrix} 1 & 2 & 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 0 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$$

if and only if $a = b = 0$. Consider then the column space of R. It consists of column vector of length 3 whose last entry is zero; this space has dimension at most 2. Its dimension is 2 because the first and third columns are linearly-independent. Thus,

$$\dim_{\mathbb{R}} \mathscr{R}(R) = \dim_{\mathbb{R}} \mathscr{C}(R) = 2.$$

The question is why these are also the dimensions of the row and column spaces of A. The easier part to see is the row space. The rows of A and linear combinations of the rows of R and vice-versa, hence,

$$\{\mathrm{Row}^i(A) : i = 1,2,3\} \subset \mathscr{R}(R) \qquad \text{and} \qquad \{\mathrm{Row}^i(R) : i = 1,2,3\} \subset \mathscr{R}(A),$$

from which we deduce that $\mathscr{R}(A) = \mathscr{R}(R)$, hence

$$\dim_{\mathbb{R}} \mathscr{R}(A) = 2.$$

The more surprising fact is that the column space of A has the same dimension as the column space of R, even though the two spaces are not identical.

The second column of R equals twice its first column,

$$\mathrm{Col}_2(R) = 2\ \mathrm{Col}_1(R),$$

and the same holds for the column of A,

$$\mathrm{Col}_2(A) = 2\ \mathrm{Col}_1(A).$$

Likewise,

$$\mathrm{Col}_4(R) = 4\ \mathrm{Col}_3(R) - 5\ \mathrm{Col}_1(R),$$

but also,

$$\mathrm{Col}_4(A) = 4\ \mathrm{Col}_3(A) - 5\ \mathrm{Col}_1(A).$$

In other words, the relations between the column of A are the same as the relations between the columns of R.

Look again at the identity

$$\begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 2 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{R} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 4 \\ 2 & 4 & 2 & 6 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A}.$$

It states that the first and third columns of A are the first and third columns of R, and that the other columns of A are linear combinations of those same two columns of R. Hence $\dim_{\mathbb{R}} \mathscr{C}(A) = 2$. ▲▲▲

3.5 Coordinates

3.5.1 Motivation

Consider the vector space $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$. It is not hard to verify that the set

$$S = \{(1,2),(2,1)\}$$

is a basis for $\mathbb{R}^2$. The fact that S spans $\mathbb{R}^2$ implies that every vector $(x, y) \in \mathbb{R}^2$ can be written as a linear combination

$$(x, y) = a(1, 2) + b(2, 1)$$

for some $a, b \in \mathbb{R}$. The fact that the vectors in S are independent, implies that a and b are determined uniquely, as if the pairs of scalars a, b and c, d satisfy

$$a(1, 2) + b(2, 1) = c(1, 2) + d(2, 1),$$

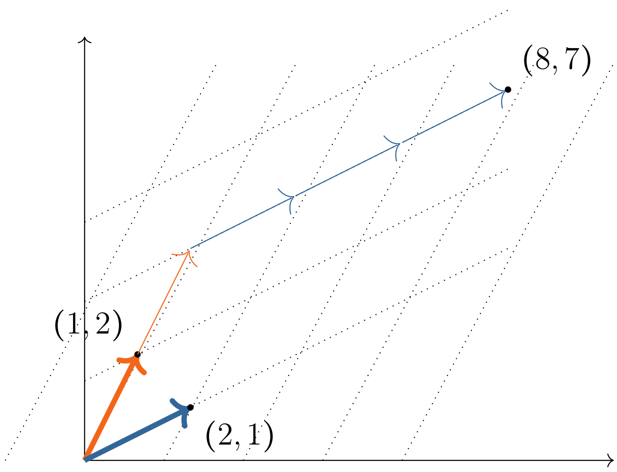
then

$$(a - c)(1, 2) + (b - d)(2, 1) = (0, 0),$$

which implies that $a = c$ and $b = d$. This means that given the basis S, every element in $\mathbb{R}^2$ can be identified with a pair of scalars, which are coefficients of the basis vectors. For example,

$$(8, 7) = 2(1, 2) + 3(2, 1).$$

This is shown in the following plot, where the vector $(8, 7)$ is shown to be twice the vector $(1, 2)$ plus three times the vector $(2, 1)$. Note also how the two basis vectors define a grid.



Given the choice of a basis, every point in $\mathbb{R}^2$ can be characterized in a unique way as a pair of scalars representing coefficients of the two basis vectors. In other words, after having chosen a basis S, we may identify the point $(8, 7)$ with the pair of scalars 2 and 3. Note however, that these coefficients cannot be viewed as an *ordered* pair unless we impose an order on the basis vectors. Thus, for example, if we decided that the basis vector $(1,2)$ is "first" and the basis vector $(2,1)$ is "second", then we could have identified the points $(8,7) \in \mathbb{R}^2$ as the ordered pair of numbers $[2,3]^T$, via

$$(8,7) = ((1,2),(2,1))\begin{bmatrix}2\\3\end{bmatrix}.$$

The column vector $[2,3]^T$ is called the coordinate matrix of $(8,7)$ with respect to the *ordered* basis $((1,2),(2,1))$.

3.5.2 Ordered bases and coordinates

We defined a basis for a vector space as a set of vectors that are both generating and linearly-independent. We already mentioned the fact that a set in not endowed with an order among its elements. If we want elements in a set to be ordered, this requires an additional structure. This leads us to the following definition:

Definition 3.44 *Let V be a finitely-generated vector space. An* ***ordered basis*** *(בסיס סדור) for V is a finite sequence $(\mathbf{v}_1,\dots,\mathbf{v}_n)$ of vectors, which is linearly-independent and spans V.*

Note that the only difference between an ordered basis and any old basis is that its elements are ordered... also, a priori, not all the elements of a sequence have to be distinct, but linear-independence implies at once that all the elements in the sequence are distinct.

Proposition 3.45 *Let V be a finitely-generated vector space, and let $\mathfrak{B} = (\mathbf{v}_1,\dots,\mathbf{v}_n)$ be an ordered basis for V. Then, to every $\mathbf{v} \in V$ there corresponds a unique $\mathbf{a} \in \mathbb{F}^n_{col}$, such that*

$$\mathbf{v} = a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = \begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}a^1\\ \vdots\\ a^n\end{bmatrix}.$$

Proof: Since a basis spans V, the existence of such scalars is guaranteed. For uniqueness, let $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n_{\text{col}}$ be such that

$$\begin{aligned} \mathbf{v} &= a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n \\ \mathbf{v} &= b^1\mathbf{v}_1 + \cdots + b^n\mathbf{v}_n. \end{aligned}$$

Thus,

$$(a^1 - b^1)\mathbf{v}_1 + \cdots + (a^n - b^n)\mathbf{v}_n = 0_V,$$

but since the vectors in $\mathfrak{B}$ are independent, it follows that $a^i = b^i$ for every $i = 1, \dots, n$, proving the uniqueness of the representation. ∎

Since, on the other hand, every $\mathbf{a} \in \mathbb{F}^n_{\text{col}}$ defines a vector in V via linear combinations of the vectors in the ordered basis, we have just discovered the following fact:

Given an ordered basis $\mathfrak{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ for a finitely-generated vector space, there exists a one-to-one correspondence between the elements of V and elements of $\mathbb{F}^n_{\text{col}}$: every element in V can be identified with a unique $\mathbf{a} \in \mathbb{F}^n_{\text{col}}$, such that

$$\mathbf{v} = a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix},$$

and vice-versa, every $\mathbf{a} \in \mathbb{F}^n_{\text{col}}$ can be identified with a unique $\mathbf{v} \in V$. The vector $\mathbf{a} \in \mathbb{F}^n_{\text{col}}$ is called the **coordinate matrix** (**מטריצת הקואורדינטות**) of $\mathbf{v}$ relative to the basis $\mathfrak{B}$. We will denote by

$$[\mathbf{v}]_{\mathfrak{B}} \in \mathbb{F}^n_{\text{col}}$$

the coordinates of $\mathbf{v}$ relative to the basis $\mathfrak{B}$, namely, for every basis $\mathfrak{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, the column vector

$$[\mathbf{v}]_{\mathfrak{B}} = \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix}$$

is the unique vector satsfying

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix} = \mathfrak{B}\,[\mathbf{v}]_{\mathfrak{B}}.$$

Example: Let $V = (\mathbb{F}^n, +, \mathbb{F}, \cdot)$ and let $\mathfrak{E} = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}$ be the standard basis. Every $\mathbf{x} = (x^1, \dots, x^n) \in \mathbb{F}^n$ can be represented as

$$\mathbf{x} = x^1\mathbf{e}_1 + \cdots + x^n\mathbf{e}_n,$$

i.e., by definition

$$[\mathbf{x}]_{\mathfrak{E}} = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}.$$

I.e., the i-th coordinate of $\mathbf{x}$ is x^i, which is really what we would expect. In other words, when we write the entries of a vector $\mathbf{x} \in \mathbb{F}^n$ as a column vector, we really write its coordinate matrix. ▲▲▲

Example: Let $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ and let

$$\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2),$$

with $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (1, -1)$. $\mathfrak{B}$ is an ordered basis for $\mathbb{R}^2$. Consider now the vector

$$\mathbf{v} = (3, 5).$$

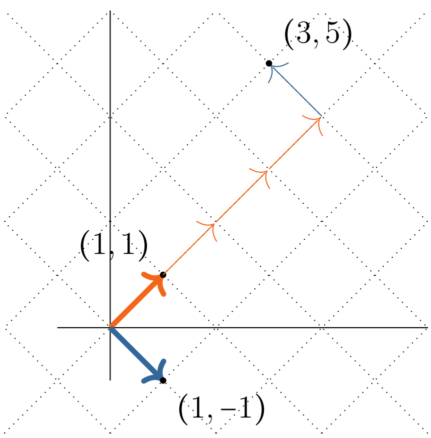
A direct calculation shows that

$$(3, 5) = \begin{pmatrix} (1,1) & (1,-1) \end{pmatrix} \begin{bmatrix} 4 \\ -1 \end{bmatrix},$$

i.e., $\mathbf{v} = \mathfrak{B}\,[\mathbf{v}]_{\mathfrak{B}}$, where

$$[\mathbf{v}]_{\mathfrak{B}} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$$

See diagram below. ▲▲▲



The following proposition shows that operations on vectors correspond to analogous operations on their coordinates:

Proposition 3.46 *Let $(V,+,\mathbb{F},\cdot)$ be a finitely-generated vector space, and let $\mathfrak{B}=(v_1,\dots,v_n)$ be an ordered basis for V. Then, for every $\mathbf{u},\mathbf{v}\in V$ and $c\in\mathbb{F}$,*

$$[\mathbf{u}+\mathbf{v}]_{\mathfrak{B}}=[\mathbf{u}]_{\mathfrak{B}}+[\mathbf{v}]_{\mathfrak{B}},$$

and

$$[c\,\mathbf{u}]_{\mathfrak{B}}=c\,[\mathbf{u}]_{\mathfrak{B}}.$$

Comment: Note that $\mathbf{u}+\mathbf{v}$ and $c\,\mathbf{u}$ are operations in $(V,+,\mathbb{F},\cdot)$, whereas $[\mathbf{u}]_{\mathfrak{B}}+[\mathbf{v}]_{\mathfrak{B}}$ and $c\,[\mathbf{u}]_{\mathfrak{B}}$ are operations in $(\mathbb{F}^n_{\text{col}},+,\mathbb{F},\cdot)$.

Proof: By definition,

$$[\mathbf{u}]_{\mathfrak{B}}=\begin{bmatrix}a^1\\ \vdots\\ a^n\end{bmatrix}\qquad\text{and}\qquad[\mathbf{v}]_{\mathfrak{B}}=\begin{bmatrix}b^1\\ \vdots\\ b^n\end{bmatrix},$$

are the unique matrices satisfying

$$\mathbf{u}=\mathfrak{B}\,[\mathbf{u}]_{\mathfrak{B}}\qquad\text{and}\qquad\mathbf{v}=\mathfrak{B}\,[\mathbf{v}]_{\mathfrak{B}}.$$

That is,

$$\mathbf{u}=a^1\mathbf{v}_1+\cdots+a^n\mathbf{v}_n=\begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}a^1\\ \vdots\\ a^n\end{bmatrix},$$

and

$$\mathbf{v}=b^1\mathbf{v}_1+\cdots+b^n\mathbf{v}_n=\begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}b^1\\ \vdots\\ b^n\end{bmatrix}.$$

Hence,

$$\mathbf{u}+\mathbf{v}=(a^1+b^1)\mathbf{v}_1+\cdots+(a^n+b^n)\mathbf{v}_n=\begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}a^1+b^1\\ \vdots\\ a^n+b^n\end{bmatrix},$$

which we may also write as

$$\mathbf{u} + \mathbf{v} = \mathfrak{B}\,([\mathbf{u}]_{\mathfrak{B}} + [\mathbf{v}]_{\mathfrak{B}}),$$

proving that $[\mathbf{u} + \mathbf{v}]_{\mathfrak{B}} = [\mathbf{u}]_{\mathfrak{B}} + [\mathbf{v}]_{\mathfrak{B}}$. Likewise,

$$c\,\mathbf{u} = c(a^1\mathbf{v}_1 + \cdots + a^n\mathbf{v}_n) = (c\,a^1)\mathbf{v}_1 + \cdots + (c\,a^n)\mathbf{v}_n,$$

which we may write as

$$c\mathbf{u} = \begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix} c\,a^1 \\ \vdots \\ c\,a^n \end{bmatrix} = \mathfrak{B}\,(c\,[\mathbf{u}]_{\mathfrak{B}}),$$

proving that $[c\,\mathbf{u}]_{\mathfrak{B}} = c\,[\mathbf{u}]_{\mathfrak{B}}$. ■

Example: Consider once again the vector space $V = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ with the ordered basis

$$\mathfrak{B} = ((1,1),(1,-1)).$$

Let

$$\mathbf{u} = (-1,0) \qquad \text{and} \qquad \mathbf{v} = (3,5) \quad \text{hence} \quad \mathbf{u} + \mathbf{v} = (2,5).$$

We proceed the calculate the coordinates,

$$\begin{aligned} \mathbf{u} &= \begin{pmatrix}(1,1) & (1,-1)\end{pmatrix}\begin{bmatrix} -1/2 \\ -1/2 \end{bmatrix} \\ \mathbf{v} &= \begin{pmatrix}(1,1) & (1,-1)\end{pmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix} \\ \mathbf{u} + \mathbf{v} &= \begin{pmatrix}(1,1) & (1,-1)\end{pmatrix}\begin{bmatrix} 7/2 \\ -3/2 \end{bmatrix}, \end{aligned}$$

so that indeed

$$[\mathbf{u} + \mathbf{v}]_{\mathfrak{B}} = [\mathbf{u}]_{\mathfrak{B}} + [\mathbf{v}]_{\mathfrak{B}}.$$

▲▲▲

Before we end this section, we emphasize its main result. The choice of an ordered basis allows us to view vectors in V as matrices of coordinates. Both V and $\mathbb{F}^n_{\text{col}}$ are vector spaces over the same field $\mathbb{F}$, but they are different spaces. What we have is an identification (which really is a one-to-one and onto function) of elements in V with elements in $\mathbb{F}^n_{\text{col}}$. What we proved is that vector addition and scalar multiplication "respect" this identification: for example, the column vector representing the sum of two vectors is the sum of the column vectors representing each vector.

Exercises

(easy) 3.67 Let $V = \mathbb{R}^2$ and let

$$\mathfrak{B} = ((1,0),(1,1)) \qquad \text{and} \qquad \mathfrak{C} = ((1,2),(2,1))$$

be ordered bases.

(a) Find $[\mathbf{v}]_{\mathfrak{B}}$ and $[\mathbf{v}]_{\mathfrak{C}}$ for $\mathbf{v} = (3,3)$.

(b) Find $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ for which $[\mathbf{v}]_{\mathfrak{B}} = [\mathbf{w}]_{\mathfrak{C}} = [3,3]^T$.

(c) What are the coordinate matrices of $(1,2)$ and $(2,1)$ relative to the basis $\mathfrak{C}$?

(easy) 3.68 Denote by $\mathbb{R}_2[X]$ the space of polynomials of degree up to 2 with real coefficients and let

$$\mathfrak{B} = (1, X, X^2) \qquad \mathfrak{C} = (X^2, X, 1) \quad \text{and} \quad \mathfrak{D} = (X+1, X^2, X-1)$$

be ordered bases.

(a) Write $[p]_{\mathfrak{B}}$, $[p]_{\mathfrak{C}}$ and $[p]_{\mathfrak{D}}$ for $p = 4 + 2X - 6X^2$.

(b) Find polynomials p_1, p_2, p_3 such that $[p_1]_{\mathfrak{B}} = [p_2]_{\mathfrak{C}} = [p_3]_{\mathfrak{D}} = [1,1,1]^T$.

(intermediate) 3.69 Show that the vectors

$$\begin{aligned} \mathbf{v}_1 &= (1,1,0,0) & \mathbf{v}_2 &= (0,0,1,1) \\ \mathbf{v}_3 &= (1,0,0,4) & \mathbf{v}_4 &= (0,0,0,2) \end{aligned}$$

form a basis for $(\mathbb{R}^4, +, \mathbb{R}, \cdot)$. What are the coordinate matrices of each of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ is the ordered basis $\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$?

(intermediate) 3.70 Let $V = (\mathbb{C}^3, +, \mathbb{C}, \cdot)$. What are the coordinates of the vector $(1,0,1)$ in the ordered basis

$$\mathfrak{B} = \big((2\imath, 1, 0) \quad (2, -1, 1) \quad (0, 1+\imath, 1-\imath)\big) \ ?$$

(intermediate) 3.71 Let

$$\mathfrak{B} = \big((1,0,-1) \quad (1,1,1) \quad (1,0,0)\big)$$

be an ordered basis for $\mathbb{R}^3$. calculate

$$[(a,b,c)]_{\mathfrak{B}}$$

for arbitrary $a, b, c \in \mathbb{R}$.

(intermediate) 3.72 Let $W \leq \mathbb{C}^3$ be the subspace generated by the vectors

$$\mathbf{v}_1 = (1, 0, \imath) \qquad \text{and} \qquad \mathbf{v}_2 = (1 + \imath, 1, -1).$$

(a) Show that $(\mathbf{v}_1, \mathbf{v}_2)$ form an ordered basis for W.

(b) Show that

$$\mathbf{u}_1 = (1, 1, 0) \qquad \text{and} \qquad \mathbf{u}_2 = (1, \imath, 1 + \imath)$$

form another basis for W.

(c) What are the coordinate matrices of $\mathbf{v}_1, \mathbf{v}_2$ in the ordered basis $(\mathbf{u}_1, \mathbf{u}_2)$?

(intermediate) 3.73 Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be vectors in $\mathbb{R}^2$ such that

$$u_1^2 + u_2^2 = v_1^2 + v_2^2 = 1 \qquad \text{and} \qquad u_1 v_1 + u_2 v_2 = 0.$$

(a) Interpret the properties of those vectors geometrically.

(b) Show that $\{\mathbf{u}, \mathbf{v}\}$ is a basis for $\mathbb{R}^2$.

(c) Find the coordinates of (x, y) in the ordered basis $(\mathbf{u}, \mathbf{v})$.

3.5.3 Transitions between bases

An ordered basis of n vectors enables us to view vectors (which are abstract entities) as n-tuples of scalars, which are more concrete entities. But bear in mind that we cannot say that a vector in a general finitely-generated vector space *is* an n-tuple of scalars. This identification relies on the choice of a basis. The same vector may have different coordinate matrices depending on the ordered basis relative to which they are defined. A natural question is the relation between coordinates of vectors relative to different bases.

Consider now a finitely-generated vector space, and let

$$\mathfrak{B} = \begin{pmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_n \end{pmatrix} \qquad \text{and} \qquad \mathfrak{C} = \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix}$$

be two ordered bases. What can be said about the relation between coordinates relative to both bases. In other words, for $\mathbf{v} \in V$, what is the relation between $[\mathbf{v}]_{\mathfrak{B}}$ and $[\mathbf{v}]_{\mathfrak{C}}$?

Since $\mathfrak{B}$ is a basis, each of the vectors $\mathbf{v}_i$ in the basis $\mathfrak{C}$ has a unique representation as a linear combination of the basis vectors $\mathbf{u}_i$. In other words, for every $i = 1\dots,n$, there exists n scalars $p_i^1,\dots,p_i^n$, such that

$$\mathbf{v}_i = p_i^1\mathbf{u}_1 + \cdots + p_i^n\mathbf{u}_n,$$

i.e.,

$$\mathbf{v}_i = \begin{pmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_n\end{pmatrix}\begin{bmatrix}p_i^1 \\ \vdots \\ p_i^n\end{bmatrix}.$$

In fact, that column vector is nothing but the coordinate matrix of $\mathbf{v}_i$ relative to the basis $\mathfrak{B}$,

$$\mathrm{Col}_i(P) = [\mathbf{v}_i]_{\mathfrak{B}},$$

where P is the $n\times n$ matrix whose entries are p_j^i. Since this hold for every $i = 1\dots,n$,

$$\begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix} = \begin{pmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_n\end{pmatrix}\begin{bmatrix}p_1^1 & \dots & p_n^1 \\ \vdots & \vdots & \vdots \\ p_1^n & \dots & p_n^n\end{bmatrix},$$

namely

$$\mathfrak{C} = \mathfrak{B}\,P.$$

Symmetrically, denoting by Q the $n\times n$ matrix such that for every $i = 1,\dots,n$,

$$\mathbf{u}_i = q_i^1\mathbf{v}_1 + \cdots + q_i^n\mathbf{v}_n,$$

namely,

$$\mathrm{Col}_i(Q) = [\mathbf{u}_i]_{\mathfrak{C}},$$

we obtain that

$$\begin{pmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_n\end{pmatrix} = \begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix}\begin{bmatrix}q_1^1 & \dots & q_n^1 \\ \vdots & \vdots & \vdots \\ q_1^n & \dots & q_n^n\end{bmatrix},$$

or

$$\mathfrak{B} = \mathfrak{C}\,Q.$$

Combining the two, for every $i = 1,\dots,n$,

$$\mathbf{v}_i = \sum_{j=1}^n p_i^j\left(\sum_{k=1}^n q_j^k\mathbf{v}_k\right) = \sum_{k=1}^n\left(\sum_{j=1}^n p_i^j q_j^k\right)\mathbf{v}_k,$$

namely,

$$\mathfrak{C} = \mathfrak{C} \cdot QP,$$

or

$$\mathfrak{C} \cdot (QP - I) = 0.$$

Since the basis vectors in $\mathfrak{C}$ are all independent, and since multiplication by $(QP - I)$ yields n linear combinations of the basis vectors $\mathfrak{C}$, these combinations vanish only if each column of $QP - I$ is identically zero, form which we deduce that

$$QP = I,$$

i.e., $P \in \mathrm{GL}_n(\mathbb{F})$ and $Q = P^{-1}$. That is, the transitions between bases is through a right-multiplication by an invertible $n \times n$ matrix. The matrices P and Q are called **transition matrices** (**מטריצות מעבר**).

Let now $\mathbf{v} \in V$. By definition,

$$\mathbf{v} = \mathfrak{B}\,[\mathbf{v}]_{\mathfrak{B}} \qquad \text{and} \qquad \mathbf{v} = \mathfrak{C}\,[\mathbf{v}]_{\mathfrak{C}}$$

Since $\mathfrak{C} = \mathfrak{B}\,P$, it follows that

$$\mathbf{v} = (\mathfrak{B}\,P)\,[\mathbf{v}]_{\mathfrak{C}} = \mathfrak{B}\,(P\,[\mathbf{v}]_{\mathfrak{C}}),$$

which implies that

$$[\mathbf{v}]_{\mathfrak{B}} = P\,[\mathbf{v}]_{\mathfrak{C}}.$$

Likewise, since $\mathfrak{B} = \mathfrak{C}\,Q$,

$$\mathbf{v} = (\mathfrak{C}\,Q)\,[\mathbf{v}]_{\mathfrak{B}} = \mathfrak{C}\,(Q\,[\mathbf{v}]_{\mathfrak{B}}),$$

from which we deduce that

$$[\mathbf{v}]_{\mathfrak{C}} = Q\,[\mathbf{v}]_{\mathfrak{B}}.$$

Let's summarize this as a theorem:

Theorem 3.47 *Let V be an n-dimensional vector space over $\mathbb{F}$. Let $\mathfrak{B} = \begin{pmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_n \end{pmatrix}$ and $\mathfrak{C} = \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix}$ be two ordered bases for V. Then the matrix $P \in M_n(\mathbb{F})$ given by*

$$\mathrm{Col}_i(P) = [\mathbf{v}_i]_{\mathfrak{B}}.$$

is invertible and $Q = P^{-1}$ is given by

$$\mathrm{Col}_i(Q) = [\mathbf{u}_i]_{\mathfrak{C}}.$$

Furthermore,

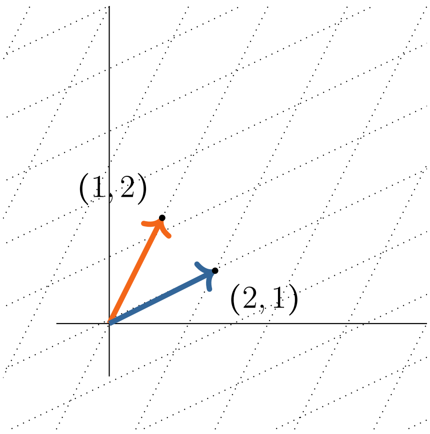
$$\mathfrak{B}\, P = \mathfrak{C} \qquad \textit{and} \qquad \mathfrak{C}\, Q = \mathfrak{B},$$

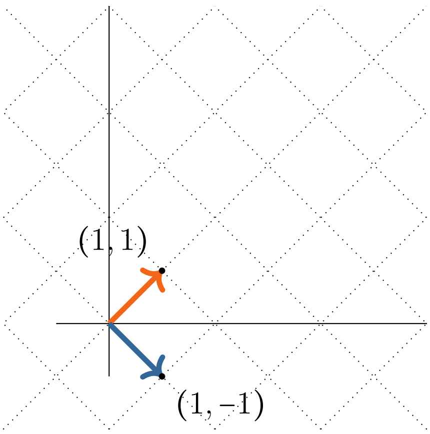
and for every $\mathbf{v} \in V$,

$$[\mathbf{v}]_{\mathfrak{B}} = P[\mathbf{v}]_{\mathfrak{C}} \qquad \textit{and} \qquad [\mathbf{v}]_{\mathfrak{C}} = Q[\mathbf{v}]_{\mathfrak{B}}.$$

Example: Let $V = \mathbb{R}^2$ and consider two bases

$$\mathfrak{B} = \big((1,2) \quad (2,1)\big) \qquad \text{and} \qquad \mathfrak{C} = \big((1,1) \quad (1,-1)\big).$$





We verify that

$$[(1,1)]_{\mathfrak{B}} = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix} \qquad \text{and} \qquad [(1,-1)]_{\mathfrak{B}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

so that

$$\underbrace{\big((1,1) \quad (1,-1)\big)}_{\mathfrak{C}} = \underbrace{\big((1,2) \quad (2,1)\big)}_{\mathfrak{B}} \underbrace{\begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix}}_{P},$$

and
$$[(1,2)]_{\mathfrak{C}} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix} \qquad \text{and} \qquad [(2,1)]_{\mathfrak{C}} = \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix},$$
so that
$$\underbrace{\begin{pmatrix} (1,2) & (2,1) \end{pmatrix}}_{\mathfrak{B}} = \underbrace{\begin{pmatrix} (1,1) & (1,-1) \end{pmatrix}}_{\mathfrak{C}} \underbrace{\begin{bmatrix} 3/2 & 3/2 \\ -1/2 & 1/2 \end{bmatrix}}_{Q}.$$
Indeed,
$$\begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 3/2 & 3/2 \\ -1/2 & 1/2 \end{bmatrix}.$$
Let $\mathbf{v} = (3,4)$. A direct calculation shows that
$$[\mathbf{v}]_{\mathfrak{B}} = \begin{bmatrix} 5/3 \\ 2/3 \end{bmatrix} \qquad \text{and} \qquad [\mathbf{v}]_{\mathfrak{C}} = \begin{bmatrix} 7/2 \\ -1/2 \end{bmatrix}.$$
You may verify that
$$\underbrace{\begin{bmatrix} 5/3 \\ 2/3 \end{bmatrix}}_{[v]_{\mathfrak{B}}} = \underbrace{\begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 7/2 \\ -1/2 \end{bmatrix}}_{[v]_{\mathfrak{C}}}.$$

▲ ▲ ▲

Example: Consider the vector space $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ and the ordered bases
$$\mathfrak{B} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{pmatrix} = \begin{pmatrix} (\cos\alpha, \sin\alpha) & (-\sin\alpha, \cos\alpha) \end{pmatrix},$$
and
$$\mathfrak{C} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} (\cos\beta, \sin\beta) & (-\sin\beta, \cos\beta) \end{pmatrix},$$
for some $\alpha, \beta \in \mathbb{R}$ (convince yourself geometrically that these are ordered bases). You may verify that for every $(x,y) \in \mathbb{R}^2$,
$$[(x,y)]_{\mathfrak{B}} = \begin{bmatrix} x\,\cos\alpha + y\,\sin\alpha \\ -x\,\sin\alpha + y\,\cos\alpha \end{bmatrix}.$$
In particular,
$$[\mathbf{v}_1]_{\mathfrak{B}} = \begin{bmatrix} \cos\beta\,\cos\alpha + \sin\beta\,\sin\alpha \\ -\cos\beta\,\sin\alpha + \sin\beta\,\cos\alpha \end{bmatrix} = \begin{bmatrix} \cos(\beta-\alpha) \\ \sin(\beta-\alpha) \end{bmatrix},$$

and

$$[\mathbf{v}_2]_{\mathfrak{B}} = \begin{bmatrix} -\sin\beta\,\cos\alpha + \cos\beta\,\sin\alpha \\ \sin\beta\,\sin\alpha + \cos\beta\,\cos\alpha \end{bmatrix} = \begin{bmatrix} -\sin(\beta-\alpha) \\ \cos(\beta-\alpha) \end{bmatrix}.$$

That is, $\mathfrak{C} = \mathfrak{B}\cdot P$, where

$$P = \begin{bmatrix} \cos(\beta-\alpha) & -\sin(\beta-\alpha) \\ \sin(\beta-\alpha) & \cos(\beta-\alpha) \end{bmatrix}.$$

We know how to invert a 2×2 matrix,

$$P^{-1} = \begin{bmatrix} \cos(\beta-\alpha) & \sin(\beta-\alpha) \\ -\sin(\beta-\alpha) & \cos(\beta-\alpha) \end{bmatrix}.$$

If follows that for every $(x,y)\in\mathbb{R}^2$,

$$[(x,y)]_{\mathfrak{C}} = \begin{bmatrix} \cos(\beta-\alpha) & \sin(\beta-\alpha) \\ -\sin(\beta-\alpha) & \cos(\beta-\alpha) \end{bmatrix} [(x,y)]_{\mathfrak{B}}.$$

▲▲▲

Exercises

(easy) 3.74 Consider Exercise 3.67.

(a) Find the matrix P whose columns are the coordinates of the vectors in $\mathfrak{C}$ relative to the basis $\mathfrak{B}$.

(b) Show directly that P is invertible and find its inverse.

(c) Find the matrix whose columns are the coordinates of the vectors in $\mathfrak{B}$ relative to the basis $\mathfrak{C}$.

(intermediate) 3.75 Let V be a vector space over $\mathbb{F}$ and let $\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ be a sequence of linearly-independent vectors.

(a) Explain why is $\mathfrak{B}$ an ordered basis for $W = \operatorname{Span}\mathfrak{B}$.

(b) Show that

$$\mathfrak{C} = (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3)$$

is also an ordered basis for W.

(c) Find the matrix P such that $\mathfrak{B} = \mathfrak{C}P$.