MAT211
Linear Algebra

Contents

Chapter 1

Vector Spaces

1.1 General vector spaces

Mathematicians like to study objects in terms of their structure, rather than the symbols, names of elements, etc. Any properties obtained from the structure, will hold for all objects that have that structure. In this way do not have to to prove the same results in different settings. In our context, linear algebra is a study of all objects that have **vector space** structure.

1.2 Real vector space

In this section we define the notion of a vector space, motivated what we have in $\mathbb{R}^n$. Thus we need a set and two operations defined on it:
addition and scalar multiplication.

A (real) vector space is a nonempty set of V(of vectors) together with two operations defined on V : addition and multiplication by scalars.
By **addition** we mean a rule which associates each two vectors $\mathbf{u}, \mathbf{v} \in V$, an element $\mathbf{u}+\mathbf{v}$, called the **sum** of $\mathbf{u}$ and $\mathbf{v}$; and by **scalar multiplication** we mean a rule which associates each element $\mathbf{u} \in V$ and each scalar (real) k, an element $k\mathbf{u}$, called the **scalar multiple** of $\mathbf{u}$ by k.

Definition 1.2.1. A vector space is a set V together with two operations of "addition" and "scalar multiplication" defined on it satisfying the following conditions:

(1) For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u}+\mathbf{v} \in V$, this means that V is closed under addition;

(2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$, $\forall \mathbf{u}, \mathbf{v} \in V$ (commutative);

(3) $\mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w}$, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ (associative);

(4) $\exists\, \mathbf{0} \in V$ satisfying $\mathbf{0}+\mathbf{u} = \mathbf{u}+\mathbf{0}$, $\forall \mathbf{u} \in V$ (zero vector);

(5) For each $\mathbf{u} \in V$, exists an element $-\mathbf{u}$ (negative of $\mathbf{u}$) such that $\mathbf{u}+(-\mathbf{u}) = \mathbf{0} = (-\mathbf{u})+\mathbf{u}$;

(6) If k is any scalar and $\mathbf{u} \in V$, then $k\mathbf{u} \in V$, this means that V is closed under scalar multiplication;

(7) $k(\mathbf{u}+\mathbf{v}) = k\mathbf{u} + k\mathbf{v}$, $\forall \mathbf{u}, \mathbf{v} \in V$ and scalar k;

(8) $(k+l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$, $\forall \mathbf{u} \in V$ and scalars $k,\ l$;

(9) $k(l\mathbf{u}) = (kl)\mathbf{u}$, $\forall \mathbf{u} \in V$ and scalars $k,\ l$;

(10) $\mathbf{1u} = \mathbf{u}$, $\forall \mathbf{u} \in V$.

Note 1.2.2. (1) In Definition 1.2.1, the actual elements are not specified and the operations are not defined.

(2) Note also that when we see addition and scalar multiplication, we do not necessarily refer to addition and scalar multiplication in $\mathbb{R}$.

(3) Some books use $\oplus$ and $\odot$ for addition and scalar multiplication (respectively) so that we are not confused by addition and scalar multiplication in $\mathbb{R}$.

We will consider a number of examples of vector spaces in what follows. We will specify the set and the operations of addition and scalar multiplication and show that conditions (1) – (10) of Definition 1.2.1 are satisfied.

Example 1.2.3. Let $V = \mathbb{R}^n$ and the standard addition and scalar multiplication defined

$$\underbrace{(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)}_{\text{standard addition}} \quad \text{and}$$

$$\underbrace{k(u_1, u_2, \ldots, u_n) = (ku_1, ku_2, \ldots, ku_n)}_{\text{standard scalar multiplication}}$$

for $\mathbf{u}$ and $\mathbf{u}$ in V and any real number k. Verify that $V = \mathbb{R}^n$ is a vector space under these operations.

Solution. Let $\mathbf{u},\ \mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. To prove Axiom 1, we must show that $\mathbf{u} + \mathbf{v} \in V$; that is, we must show that $\mathbf{u} + \mathbf{v}$ is a vector in $\mathbb{R}^n$. But this follows from the definition of addition of vectors, since

$$\mathbf{u} + \mathbf{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, u_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

has the form of a vector in $\mathbb{R}^n$. Axiom 2 follows since

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, u_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \\ &= (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n) = (v_1, v_2, \ldots, u_n) + (u_1, u_2, \ldots, u_n) \\ &= \mathbf{v} + \mathbf{u}. \end{aligned}$$

Similarly, Axiom 3 follows; To prove Axiom 4, we must find an element $\mathbf{0}$ in V such that $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u}$ in V. This can be done by defining $\mathbf{0}$ to be

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

With this definition

$$\mathbf{0} + \mathbf{u} = (0, 0, \ldots, 0) + (u_1, u_2, \ldots, u_n) = (u_1, u_2, \ldots, u_n) = \mathbf{u}$$

and similarly $\mathbf{u}+\mathbf{0}=\mathbf{u}$. To prove Axiom 5 we must show that each element $\mathbf{u}$ in V has a negative $-\mathbf{u}$ such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$ and $(-\mathbf{u})+\mathbf{u}=\mathbf{0}$. This can be achieved by defining the negative of $\mathbf{u}$ to be

$$-\mathbf{u}=(-u_1,-u_2,\ldots,-u_n).$$

With this definition

$$\mathbf{u}+(-\mathbf{u})=(u_1,u_2,\ldots,u_n)+(-u_1,-u_2,\ldots,-u_n)=(0,0,\ldots,0)=\mathbf{0}$$

and similarly $(-\mathbf{u})+\mathbf{u}=\mathbf{0}$. To prove Axiom 6, we must show that $k\mathbf{u}\in V$; that is, we must show that $k\mathbf{u}$ is a vector in $\mathbb{R}^n$. But this follows from the definition of scalar multiplication, since

$$k\mathbf{u}=k(u_1,u_2,\ldots,u_n)=(ku_1,ku_2,\ldots,ku_n)$$

has the form of a vector in $\mathbb{R}^n$, i.e., it is an n-tuple. Axioms 7, 8, 9 and 10 can be easily verified. □

Example 1.2.4. Let V the set of all 2×2 matrices with matrix addition and scalar multiplication. Show that this gives a vector space.

Solution. For convenience, in this example we will verify the axioms in the following order: 1, 6, 2, 3, 7, 8, 9, 4, 5 and 10. Let

$$\mathbf{u}=\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \quad \text{and } \mathbf{v}=\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}.$$

To prove Axiom 1, we must show that $\mathbf{u}+\mathbf{v}\in V$; that is, we must show that $\mathbf{u}+\mathbf{v}$ is a 2×2 matrix. But this follows from the definition of matrix addition, since

$$\mathbf{u}+\mathbf{v}=\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}+\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}=\begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix}.$$

Similarly, Axiom 6 holds since for any real number k we have

$$k\mathbf{u}=k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix}$$

so that $k\mathbf{u}$ is a 2×2 matrix a so it is an object in V.

Axiom 2 follows since

$$\mathbf{u}+\mathbf{v}=\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}+\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}=\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}+\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\mathbf{v}+\mathbf{u}.$$

Similarly, Axiom 3 follows; and Axioms 7, 8, and 9 can be easily verified.

To prove Axiom 4, we must find an element $\mathbf{0}$ in V such that $\mathbf{0}+\mathbf{u}=\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}$ in V. This can be done by defining $\mathbf{0}$ to be

$$\mathbf{0}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

With this definition

$$\mathbf{0}+\mathbf{u}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}+\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\mathbf{u}$$

and similarly $\mathbf{u} + \mathbf{0} = \mathbf{u}$. To prove Axiom 5 we must show that each element $\mathbf{u}$ in V has a negative $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ and $(-\mathbf{u}) + \mathbf{u} = \mathbf{0}$. This can be achieved by defining the negative of $\mathbf{u}$ to be

$$-\mathbf{u} = \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix}.$$

With this definition

$$\mathbf{u} + (-\mathbf{u}) = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$$

and similarly $(-\mathbf{u}) + \mathbf{u} = \mathbf{0}$. Finally, Axiom 10 is a simple computation:

$$1\mathbf{u} = 1 \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \mathbf{u}.$$

□

Remark 1.2.5. To verify condition (1) of Definition 1.2.1, suppose that we took:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 6 & 3 \end{bmatrix} \in V.$$

Can we conclude that V is closed under scalar multiplication?

Example 1.2.4 is a special case of the following example.

Example 1.2.6. Let V is the set of all $m \times n$ matrices with real entries with matrix addition and scalar multiplication. This is clearly a vector space, with zero vector ($\mathbf{0}$) the $m \times n$ zero matrix, and the negative of $\mathbf{u}$ is $-\mathbf{u}$ obtained by negating each entry of the matrix. This vector space is denoted by $M_{m\times n}$.

Example 1.2.7. Let V be the set of all real-valued functions defined on $\mathbb{R}$ (i.e., functions $f : \mathbb{R} \longrightarrow \mathbb{R}$) What do we mean by $\mathbf{f} + \mathbf{g}$ and $k\mathbf{f}$ where $\mathbf{f}$, $\mathbf{g} \in V$ and k is a scalar?
If $\mathbf{f} = f(x)$, and $\mathbf{g} = g(x)$ are two functions and k is any real number, define the sum function $\mathbf{f} + \mathbf{g}$ and the scalar multiple $k\mathbf{f}$ by

$$\begin{aligned} (\mathbf{f} + \mathbf{g})(x) &= f(x) + g(x) \text{ and} \\ k\mathbf{f}(x) &= kf(x). \end{aligned}$$

Show that V is a vector space. We denote this space by $F(-\infty, \infty)$.

Solution. If $\mathbf{f}$, and $\mathbf{g}$ are two functions in $F(-\infty, \infty)$ to say that $\mathbf{f} = \mathbf{g}$ is equivalent to saying that $f(x) = g(x)$ for all $x \in (-\infty, \infty)$. The vector $\mathbf{0} \in (-\infty, \infty)$ is the constant function that is identically zero for all values of x. The negative vector $-\mathbf{f} = -f(x)$. The reader should verify that the remaining vector space axioms are satisfied. □

In Example 1.2.7, we looked at functions from $\mathbb{R} \longrightarrow \mathbb{R}$. If we look at functions from $[a,\, b] \longrightarrow \mathbb{R}$ or $(a,\, b) \longrightarrow \mathbb{R}$, then we also get vector spaces. We denote these vector spaces by $F[a,\, b]$ and $F(a,\, b)$, respectively.

In the next example we will define a "scalar multiplication" in $\mathbb{R}^2$ different from the standard operation.

Example 1.2.8. Let $V = \mathbb{R}^2$ and define the operations of addition and scalar multiplication as follows:

$$\underbrace{(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)}_{\text{standard operation}} \quad \text{and} \quad \underbrace{k(u_1, u_2) = (ku_1, 0)}_{\text{non standard operation}}.$$

For example, if $\mathbf{u} = (1, 2)$ and $\mathbf{v} = (-1, 3)$ and $k = 5$, then

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (0, 5), \text{ and} \\ 5\mathbf{u} &= (5, 0). \end{aligned}$$

Does V together with these operations form a vector space?

Solution. The reader is required to verify that the first nine vector space axioms are satisfied; however there are values of $\mathbf{u}$ for which Axiom 10 fails to hold. For example, if $\mathbf{u} = (u_1, u_2)$ is such that $u_2 \neq 0$, then

$$1\mathbf{u} = 1(u_1, u_2) = (1 \cdot u_1, 0) = (u_1, 0) \neq \mathbf{u}.$$

Hence, V is not a vector space with the given operations. □

Example 1.2.9. Let V be any plane through the origin in $\mathbb{R}^3$ and use the standard operations of addition and scalar multiplication. Is this a vector space? Recall that an equation of a plane in $\mathbb{R}^3$ is given by $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = d$.

Solution. From Example 1.2.3 we have that $\mathbb{R}^3$ is a vector space under the given operations. Thus, Axioms 2, 3, 7, 8, 9, and 10 hold for all points in $\mathbb{R}^3$, and in particular for all points in the plane V. We therefore need only show that Axioms 1, 4, 5 and 6 are satisfied.
Since the plane V passes through the origin, it has an equation of the form

$$a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0} \tag{1.1}$$

Thus, if $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are points in V, $au_1 + bu_2 + cu_3 = \mathbf{0}$ and $av_1 + bv_2 + cv_3 = \mathbf{0}$. Adding these equations gives

$$a(u_1 + v_1) + b(u_2 + v_2) + c(u_3 + v_3) = \mathbf{0}.$$

This equality indicates that the coordinates of the point

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$$

satisfy Equation (1.1); thus, $\mathbf{u} + \mathbf{v}$ lies in the plane V. This proves Axiom 1. The verifications of Axioms 4 and 6 are left as exercise to the reader; we shall prove Axiom 5. Multiplying $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0}$ through by -1 gives

$$a(-u_1) + b(-u_2) + c(-u_3) = \mathbf{0}.$$

Thus, $-\mathbf{u} = (-u_1, -u_2, -u_3)$ lies in V and so Axiom 5 is established. □

Example 1.2.10. Let V consist of a single object, call it $\mathbf{u}$. Define addition and scalar multiplication as follows:

$$\begin{cases} \mathbf{u} + \mathbf{u} = \mathbf{u}; \\ k\mathbf{u} = \mathbf{u}. \end{cases}$$

Is this a vector space?

Solution. It is easy to check that all the vector space axioms are satisfied. We leave that task to the reader. $\square$

The vector space of Example 1.2.10 is called the **zero vector space**.

Example 1.2.11. Consider $\mathbb{R}^2$ with usual scalar multiplication, but define addition as follows:

$$(u_1, u_2) \oplus (v_1, v_2) = (u_1 + v_1, 2u_2 + v_2).$$

Is this a vector space?

Solution. The reader is required to verify if the other vector space axioms are satisfied; however there are values of $\mathbf{u}, \mathbf{v}$ for which Axiom 2 fails to hold, that is, in general

$$\mathbf{u}+\mathbf{v} = (u_1, u_2)\oplus(v_1, v_2) = (u_1+v_1, 2u_2+v_2) \neq \mathbf{v}+\mathbf{u} = (v_1, v_2)\oplus(u_1, u_2) = (v_1+u_1, 2v_2+u_2).$$

Take for example $\mathbf{u} = (1, 3)$ and $\mathbf{v} = (4, 7)$. We obtain

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (1, 3) \oplus (4, 7) = (1 + 4, 2 \times 3 + 7) = (5, 13)\\ &\neq \mathbf{v} + \mathbf{u} = (4, 7) \oplus (1, 3) = (4 + 1, 2 \times 7 + 3) = (5, 17).\end{aligned}$$

Hence, V is not a vector space with the given operations. $\square$

We have given a number of examples to get a feel for what a vector space is. Now let us look at some useful properties of vector spaces.

Theorem 1.2.12. *Let V be a vector space, $\mathbf{u} \in V$ and k a scalar. Then*

(a) the zero vector in V is unique;

(b) the negative of a vector in V is unique;

(c) $0\mathbf{u} = \mathbf{0}$, $\forall \mathbf{u} \in V$;

(d) $k\mathbf{0} = \mathbf{0}$, for all scalars k;

(e) $(-1)\mathbf{u} = -\mathbf{u}$;

(f) if $k\mathbf{u} = \mathbf{0}$, then $k = 0$ or $\mathbf{u} = \mathbf{0}$.

Proof. We shall proof parts (c) and (e) and leave the remaining parts as exercises for the reader.

(c) We can write

$$\begin{aligned}0\mathbf{u} + 0\mathbf{u} &= (0 + 0)\mathbf{u} && \text{[Axiom 8]}\\ &= 0\mathbf{u}. && \text{[property of the number 0]}\end{aligned}$$

By Axiom 5 the vector $0\mathbf{u}$ has a negative, $-0\mathbf{u}$. Adding this negative to both sides above yields

$$[0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u}) = 0\mathbf{u} + (-0\mathbf{u})$$

or

$$\begin{aligned}0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})] &= 0\mathbf{u} + (-0\mathbf{u}) && \text{[Axiom 3]}\\ &= 0\mathbf{u} + \mathbf{0} && \text{[Axiom 5]}\\ &= 0\mathbf{u} = \mathbf{0}. && \text{[Axiom 4]}\end{aligned}$$

(e) To show that $(-1)\mathbf{u} = -\mathbf{u}$, we must demonstrate that $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. To see this, notice that

$$\begin{aligned}
\mathbf{u} + (-1)\mathbf{u} &= 1\mathbf{u} + (-1)\mathbf{u} && \text{[Axiom 10]} \\
&= (1 + (-1))\mathbf{u} && \text{[Axiom 8]} \\
&= 0\mathbf{u} && \text{[property of numbers]} \\
&= \mathbf{0}. && \text{[by part (c) above]}
\end{aligned}$$

□

1.3 Subspaces

We now look at the sets that are subsets of a bigger vector space. We would like to know whether these smaller sets are also vector spaces using the same operations of addition and scalar multiplication of the the bigger vector space.

Definition 1.3.1. A subset W of a vector space V is called a **subspace** of V if W is itself a vector space under the operations of addition and scalar multiplication defined on V.

The notion established in Definition 1.3.1 means that W must satisfy all 10 conditions that characterize a vector space using operations of V. But some of these conditions can be gotten rid of. For example, if $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, $\forall \mathbf{u}, \mathbf{v} \in V$, then it must also be true for all $\mathbf{u}, \mathbf{v} \in W$. The following theorem shows that *only two* of the vector space conditions need to be verified to obtain a subspace.

Theorem 1.3.2. *If W is a subset of one or more vectors from a vector space V, then W is a subspace of V if and only if the following conditions hold:*

(a) if $\mathbf{u}, \mathbf{v} \in W$, then $\mathbf{u} + \mathbf{v} \in W$.

(b) if $\mathbf{u} \in W$ and k is a scalar, then $k\mathbf{u} \in W$.

Proof. If W is a subspace of V, then all the vector space axioms are satisfied; in particular, Axioms 1 and 6 hold. But these are precisely conditions (a) and (b).
Conversely, assume that conditions (a) and (b) hold. Since these are vector space Axioms 1 and 6, we need only show that W satisfies the remaining eight axioms. Axioms 2, 3, 7, 8, 9 and 10 are automatically satisfied by the vectors in W since they are satisfied by all vectors in V (why?). Thus, to complete the proof, we need only verify that Axioms 4 and 5 are satisfied by vectors in W. Let $\mathbf{u}$ be any vector in W. By condition (b), $k\mathbf{u}$ is in W for every scalar k. Setting $k = 0$, it follows from Theorem 1.2.12 that $0\mathbf{u} = \mathbf{0}$ is in W, and setting $k = -1$ it follows that $(-1)\mathbf{u} = -\mathbf{u}$ is in W. □

Example 1.3.3. In Example 1.2.9, we showed that any plane through the origin in $\mathbb{R}^3$, using the standard operations of addition and scalar multiplication is a vector space. This means that any plane through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$, under the standard operations. In light of Theorem 1.3.2 we can see that much of that work was unnecessary; it would have been sufficient to verify that the plane is closed under addition and scalar multiplication. (Axioms 1 and 6). Alternatively, we can prove this geometrically: Let W be any plane through the origin and let $\mathbf{u}$ and $\mathbf{v}$ be any vectors in W. Then $\mathbf{u}+\mathbf{v} \in W$ since it is the diagonal of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$, and $k\mathbf{u} \in W$ for any scalar k because $k\mathbf{u}$ lies on a line through $\mathbf{u}$. Thus, W is closed under addition and scalar multiplication, so it is a subspace of $\mathbb{R}^3$.

Example 1.3.4. Any line through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$.

Solution. Let W be a line through the origin of $\mathbb{R}^3$. It is evident geometrically that the sum of two vectors on this line also lies on the line and that a scalar multiple of a vector on the line lies on the line. Thus, W is closed under addition and scalar multiplication, so it is a subspace of $\mathbb{R}^3$. $\square$

Remark 1.3.5. Every nonzero vector space V has at least two subspaces: V itself is a subspace, and the set $\{\mathbf{0}\}$ consisting of the zero vector in V is a subspace called the **zero space**. Combining this with Examples 1.3.3 and 1.3.4, we obtain the following list of subspaces of $\mathbb{R}^2$ and $\mathbb{R}^3$:

Subspaces of $\mathbb{R}^2$

- $\{\mathbf{0}\}$
- Lines through the origin
- $\mathbb{R}^2$

Subspaces of $\mathbb{R}^3$

- $\{\mathbf{0}\}$
- Lines through the origin
- Planes through the origin
- $\mathbb{R}^3$

Example 1.3.6. W is the set of all $n \times n$ symmetric matrices (i.e. $A^t = A$). Is W a subspace of M_n?

Solution. (a) Since $\mathbf{O}_n^t = \mathbf{O}_n$, where $\mathbf{O}_n$ is the zero $n \times n$ matrix, we have that $\mathbf{O}_n \in W$, so $W \neq \varnothing$.
(b) Let $A, B \in W$ we need to show that $A+B \in W$. But $A \in W$ implies that $A^t = A$, $(*)$ and $B \in W$ implies that $B^t = B$ $(**)$. Therefore, showing that $A+B \in W$ translates to showing that $(A+B)^t = A+B$. Now

$$\begin{aligned} (A+B)^t &= A^t + B^t, \text{ properties of transpose of a sum} \\ &= A+B, \text{ by (*) and (**)} . \end{aligned}$$

So, we have $A+B \in W$.
(c) Finally let k be a scalar and $A \in W$ we need to show that $kA \in W$. As above $A \in W$ implies that $A^t = A$ and showing that $kA \in W$ implies proving that $(kA)^t = kA$. But $(kA)^t = kA^t = kA$, so that $kA \in W$. Hence, W is a subspace of M_n. $\square$

Example 1.3.7. Let W be the set of all upper triangular matrices Is W a subspace of M_n?

Solution. Recall that an $n \times n$ matrix A is called upper-triangular if all entries lying below the diagonal are zero, that is, if $a_{ij} = 0$ whenever $i > j$. You need to verify that the upper triangular matrices form a subspace of M_n. $\square$

Example 1.3.8. Let V be the set of all functions from $\mathbb{R} \to \mathbb{R} \quad (V = F(-\infty, \infty))$
W- the set of all continuous functions from $\mathbb{R} \to \mathbb{R}$
U- the set of all differentiable functions from $\mathbb{R} \to \mathbb{R}$.
Then U is a subspace of W and W a subspace of V. (Why?)

Solution. We need to show that U is a subspace of W and W is a subspace of V.
We first show that W is a subspace of V.
The zero constant function $\mathbf{0}$ is a continuous function, so $W \neq \varnothing$.
Recall from Calculus that if $\mathbf{f}$ and $\mathbf{g}$ are continuous functions in $(-\infty, \infty)$ and k is a scalar, then $\mathbf{f} + \mathbf{g}$ and $k\mathbf{f}$ are also continuous functions in $(-\infty, \infty)$. Hence we have that W is a subspace of V.
We now show that U is a subspace of W. It follows from Calculus that any differentiable function is a continuous function, so we have that U is a subspace of W. $\square$

Example 1.3.9. Let P_n denote set of all polynomials of degree less than or equal n. Is P_n a subspace of $V = F(-\infty, \infty)$?

Solution. Recall that $P_n = \{\mathbf{p} \,|\, p(x) = a_0 + a_1x + \ldots + a_nx^n\}$ we need to verify that P_n is a subspace of V. Certainly $P_n \neq \varnothing$, since $\mathbf{0} = 0(x) = 0 + 0x + \ldots + 0x^n \in P_n$.

Now, let $\mathbf{p}, \mathbf{q} \in P_n$. Then

$$
\begin{aligned}
p(x) &= a_0 + a_1x + \ldots + a_nx^n \\
q(x) &= b_0 + b_1x + \ldots + b_nx^n.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
(\mathbf{p} + \mathbf{q})(x) = p(x) + q(x) &= (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n \qquad \text{and} \\
(k\mathbf{p})(x) = kp(x) &= (ka_0) + (ka_1)x + \ldots + (ka_n)x^n
\end{aligned}
$$

These functions have the form of the elements of P_n, so $\mathbf{p} + \mathbf{q}$ and $k\mathbf{p}$ lie in P_n. Hence P_n is a subspace of V. $\square$

Example 1.3.10. Let $W = \{(x, y) \in \mathbb{R}^2 \,|\, x \geq 0, y \geq 0\}$ (i.e., the first quadrant of $\mathbb{R}^2$.) Is W a subspace of $\mathbb{R}^2$?

Solution. W is not a subspace of $\mathbb{R}^2$, since in general if k is any scalar and $\mathbf{u} \in W$ we have that $k\mathbf{u} \notin W$. In particular take $\mathbf{u} = (1, 1)$ and $k = -1$, then $\mathbf{u} \in W$, but $k\mathbf{u} = -1(1, 1) = (-1, -1) \notin W$. $\square$

Example 1.3.11. Let $W = \{(x, y) \in \mathbb{R}^2 \,|\, xy \geq 0\}$ (i.e., the first and the third quadrants of $\mathbb{R}^2$.) Is W a subspace of $\mathbb{R}^2$?

Solution. W is not a subspace of $\mathbb{R}^2$, since in general if $\mathbf{u}, \mathbf{v} \in W$, then $\mathbf{u} + \mathbf{v}$ need not be in W. Take for example $\mathbf{u} = (2, 3)$ and $\mathbf{v} = (-4, -1)$ then both $\mathbf{u}$ and $\mathbf{v}$ belong to W(Why?). However $\mathbf{u} + \mathbf{v} = (2, 3) + (-4, -1) = (-2, 1) \notin W$. (Why?) $\square$

Example 1.3.12. Let $V = M_{nn}$ denoted by M_n, and W be the set of all $n \times n$ invertible matrices. Note that $W \subseteq V$. Is W a subspace of V?

Solution. Exercise for the reader. $\square$

The following theorem shows that the set of all solution vectors of a homogeneous system of linear equations forms a vector space.

Theorem 1.3.13. *If $A\mathbf{x} = \mathbf{0}$ is a homogeneous system of m equations in n unknowns, then the set of solution vectors is a subspace of $\mathbb{R}^n$. (This subspace is called the* solution space *of the system of linear equations.)*

Proof. Let $W = \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} = \mathbf{0}\}$. We are required to show that W is a subspace of $\mathbb{R}^n$. Notice that $\mathbf{0}$ satisfy the equation $A\mathbf{x} = \mathbf{0}$, so $\mathbf{0} \in W$ and hence $W \neq \varnothing$.
Let $\mathbf{x}$, $\mathbf{y} \in W$ then $A\mathbf{x} = \mathbf{0}$, $A\mathbf{y} = \mathbf{0}$. We need to show that $\mathbf{x}+\mathbf{y} \in W$, that is, $A(\mathbf{x}+\mathbf{y}) = \mathbf{0}$. But

$$\begin{aligned} A(\mathbf{x}+\mathbf{y}) &= A\mathbf{x} + A\mathbf{y} \\ &= \mathbf{0} + \mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

Also for k any scalar, and $\mathbf{x} \in W$ we have that $A(k\mathbf{x}) = kA\mathbf{x} = k\mathbf{0} = \mathbf{0}$, so that $k\mathbf{x} \in W$. Hence W is a subspace of $\mathbb{R}^n$. □

Example 1.3.14. Consider the following systems of linear equations:

$$(a)\ \begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ (b)\ \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$(c)\ \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ (d)\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

All the solution spaces will be subspaces of $\mathbb{R}^3$. Describe each subspace geometrically.

Solution. Each of these systems has three unknowns, so the solutions form subspaces of $\mathbb{R}^3$. Geometrically, this means that each solution space must be a line through the origin, a plane through the origin, the origin only, or all $\mathbb{R}^3$. We shall verify that this is so, leaving to the reader the task of solving the systems.
(a) The solutions are $x = 2s - 3t,\ y = s,\ z = t$ from which we deduce that

$$x = 2y - 3z \ \text{ or } \ x - 2y + 3z = 0.$$

This is the equation of a plane through the origin with $\mathbf{n} = (1, -2, 3)$ as a normal vector.
(b) The solutions are $x = -5t,\ y = -t,\ z = t$ which are parametric equations for the line through the origin parallel to the vector $\mathbf{v} = (-5, -1, 1)$.
(c) The solution is $x = 0,\ y = 0,\ z = 0$. The solution space is the origin only, that is, $\{\mathbf{0}\}$.
(d) The solutions are $x = r,\ y = s,\ z = t$ where r, s and t have arbitrary values, so the solution space is the all $\mathbb{R}^3$. □

1.4 Linear combinations of vectors

Definition 1.4.1. A vector $\mathbf{v}$ is said to be a **linear combination** of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ if we can find scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that

$$\mathbf{v} = \lambda_1\mathbf{v_1} + \lambda_2\mathbf{v_2} + \ldots + \lambda_k\mathbf{v_k}.$$

Example 1.4.2. Consider the vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$ in $\mathbb{R}^3$. Then

$$\begin{aligned}(3, -1, -4) &= 3(1, 0, 0) + (-1)(0, 1, 0) + (-4)(0, 0, 1) \\ &= 3\mathbf{i} - \mathbf{j} - 4\mathbf{k}; \\ (a, b, c) &= a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) \\ &= a\mathbf{i} + b\mathbf{j} + c\mathbf{k}.\end{aligned}$$

Example 1.4.3. Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in $\mathbb{R}^3$. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$, but $(4, -1, 8)$ is not.

Solution. In order for $\mathbf{w}$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(9, 2, 7) = k_1(1, 2, -1) + k_2(6, 4, 2) \text{ or } (9, 2, 7) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2).$$

Equating corresponding components gives

$$\begin{aligned}k_1 + 6k_2 &= 9 \\ 2k_1 + 4k_2 &= 2 \\ -k_1 + 2k_2 &= 7.\end{aligned}$$

Solving this system yields $k_1 = -3$, $k_2 = 2$, so

$$\mathbf{w} = -3\mathbf{u} + 2\mathbf{v}.$$

Similarly, for $\mathbf{w}'$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(4, -1, 8) = k_1(1, 2, -1) + k_2(6, 4, 2) \text{ or } (4, -1, 8) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2).$$

Equating corresponding components gives

$$\begin{aligned}k_1 + 6k_2 &= 4 \\ 2k_1 + 4k_2 &= -1 \\ -k_1 + 2k_2 &= 8.\end{aligned}$$

The system of equations is inconsistent (i.e., it has no solution) (verify), so no such scalars k_1 and k_2 exist. Hence, $\mathbf{w}'$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$. □

As we have seen in Example 1.4.3, some vectors in $\mathbb{R}^3$ are a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and others are not. If we take all possible linear combinations of $\mathbf{u}$ and $\mathbf{v}$, then we obtain a subspace of $\mathbb{R}^3$. This is expressed, more generally, in the following theorem.

Theorem 1.4.4. *Let $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ be vectors in a vector space V. Then:*

(a) the set W of all linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ is a subspace of V.

(b) W is the smallest subspace of V that contains $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$; i.e., if W' is any subspace of V that contains $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, then $W \subseteq W'$.

Proof. (a) To show that W is a subspace of V, we must prove it is closed under addition and scalar multiplication. There is at least one vector in W, namely the zero vector, since $\mathbf{0} = 0\mathbf{v_1} + 0\mathbf{v_2} + \ldots + 0\mathbf{v_n}$. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then

$$\begin{aligned} \mathbf{u} &= k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n}, \text{ and} \\ \mathbf{v} &= l_1\mathbf{v_1} + l_2\mathbf{v_2} + \ldots + l_n\mathbf{v_n}, \end{aligned}$$

where $k_1, k_2, \ldots, k_n,\ l_1, l_2, \ldots, l_n$ are scalars. Therefore,

$$\mathbf{u} + \mathbf{v} = (k_1 + l_1)\mathbf{v_1} + (k_2 + l_2)\mathbf{v_2} + \ldots + (k_n + l_n)\mathbf{v_n}$$

and, for any scalar k,

$$k\mathbf{u} = (kk_1)\mathbf{v_1} + (kk_2)\mathbf{v_2} + \ldots + (kk_n)\mathbf{v_n}.$$

Thus, $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ are linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, and consequently lie in W. Therefore, W is closed under addition and scalar multiplication.

(b) Each $\mathbf{v}_i$ is a linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, since we can write

$$\mathbf{v}_i = 0\mathbf{v_1} + 0\mathbf{v_2} + \ldots + 1\mathbf{v}_i + \ldots + 0\mathbf{v_n}.$$

Therefore, the subspace W contains each of the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$. Let W' be any other subspace of V that contains $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$. Since W' is closed under addition and scalar multiplication, it must contain all linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$. Thus W' contains each vector of W.

□

We make the following definitions

Definition 1.4.5. Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be a set of vectors in a vector space V and let W be the subspace of V consisting of all linear combinations of vectors in S. Then W is said to be **spanned** by $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ (or W is spanned by S) or S **spans** W. We write $W = \mathrm{spans}(S)$ or $W = \mathrm{span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$.

From Theorem 1.4.4 we can conclude that span(S) is a subspace of V.

Example 1.4.6. (a) If $\mathbf{v}$ is a nonzero vector in $\mathbb{R}^2$ or $\mathbb{R}^3$. What is $\mathrm{span}\{\mathbf{v}\}$?

(b) Let $\mathbf{v_1}$ and $\mathbf{v_2}$ be two nonzero vectors in $\mathbb{R}^3$ originating at the origin that are not collinear (i.e., do not lie on the same line). What is $\mathrm{span}\{\mathbf{v_1}, \mathbf{v_2}\}$?

Solution. (a) $\mathrm{span}(\mathbf{v})$, is the set of all scalar multiples $k\mathbf{v}$, is the line determined by $\mathbf{v}$.
(b) If $\mathbf{v_1}$ and $\mathbf{v_2}$ are two collinear vectors in $\mathbb{R}^3$ with initial points at the origin, then $\mathrm{span}(\mathbf{v_1},\ \mathbf{v_2})$, which consists of all linear $k_1\mathrm{v}_1 + k_2\mathrm{v}_2$, is the plane determined by $\mathbf{v_1}$ and $\mathbf{v_2}$. □

Example 1.4.7. The polynomials $\mathbf{1}, \mathbf{x}, \mathbf{x^2}, \ldots, \mathbf{x^n}$ span P_n (the vector of all polynomials of degree less than or equal n) since each polynomial $\mathbf{p}$ in P_n can be written a

$$\mathbf{p} = a_0 + a_1 x a_2 x^2 + \ldots + a_n x^n$$

which is a linear combination of $\mathbf{1}, \mathbf{x}, \mathbf{x^2}, \ldots, \mathbf{x^n}$. We denote this by writing

$$P_n = \mathrm{span}(\mathbf{1}, \mathbf{x}, \mathbf{x^2}, \ldots, \mathbf{x^n}).$$

Example 1.4.8. Determine whether $\mathbf{v_1} = (1, 1, 2)$, $\mathbf{v_2} = (1, 0, 1)$ and $\mathbf{v_3} = (2, 1, 3)$ span the vector space $\mathbb{R}^3$.

Solution. We must determine whether an arbitrary vector $\mathbf{b} = (b_1, b_2, b_3)$ can be expressed as a linear combination

$$\mathbf{b} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$. Expressing this equation in terms of components gives

$$(b_1, b_2, b_3) = k_1(1, 1, 2) + k_2(1, 0, 1) + k_3(2, 1, 3)$$

or

$$(b_1, b_2, b_3) = (k_1 + 2k_2 + 2k_3, k_1 + k_3, 2k_1 + k_2 + 3k_3)$$

or on equating corresponding components

$$\begin{aligned} k_1 + 2k_2 + 2k_3 &= b_1 \\ k_1 + k_3 &= b_2 \\ 2k_1 + k_2 + 3k_3 &= b_3. \end{aligned} \tag{1.2}$$

The problem thus reduces to determining whether the system (1.2) is consistent (has solution) for all choices of $\mathbf{b} = (b_1, b_2, b_3)$. This is equivalent to showing that the matrix of coefficients

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

is invertible (has nonzero determinant). But $\det(A) = 0$ (verify), so that A is not invertible; and so $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ do not span $\mathbb{R}^3$. $\square$

Example 1.4.9. Determine whether $\mathbf{v_1} = (1, 1, 1)$, $\mathbf{v_2} = (1, 0, 1)$ and $\mathbf{v_3} = (2, 1, 0)$ span the vector space $\mathbb{R}^3$.

Solution. Left to the reader. $\square$

Sets that span $\mathbb{R}^3$, for example, are not unique. In Example 1.4.9 we have one set that spans $\mathbb{R}^2$ and in Example 1.4.2 we have another set that spans $\mathbb{R}^3$. The following theorem tells us what happens when we have two sets that span the same vector space.

Theorem 1.4.10. *Let V be a vector space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ and $S' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_r}\}$ be two sets in V. Then $\operatorname{span}(S) = \operatorname{span}(S')$ if and only if each vector in S is a linear combination of the vectors in S', and conversely each vector in S' is a linear combination of the vectors in S.*

Proof. The proof is left as an exercise for the reader. $\square$

1.5 Linear independence

We have just learnt what a spanning set is. But it may turn out that there is more than one way of expressing a vector as a linear combination of a spanning set. We are now interested in when each vector in V can be expressed as a linear combination of a spanning set in **exactly one way**. These types of spanning sets are important.

Note that $k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n} = \mathbf{0}$, then there is at least one solution, i. e., $k_1 = k_2 = \ldots = k_n = 0$. We are interested in the case where this is the only solution, as is stated in the following definition:

Definition 1.5.1. Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be a nonempty set of vectors. The set S is said to be **linearly independent** if:

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n} = \mathbf{0} \Rightarrow k_1 = k_2 = \ldots = k_n = 0 \tag{1.3}$$

(i.e. there exists only one solution to Equation (1.3)).

If Equation (1.3) has more than one solution, then S is said to be **linearly dependent**. A set S is linearly independent if a linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ is $\mathbf{0}$, then all coefficients must be zero.

Example 1.5.2. Let $\mathbf{v_1} = (2, -1, 0, 3)$, $\mathbf{v_2} = (1, 2, 5, -1)$ and $\mathbf{v_3} = (7, -1, 5, 8)$, and let $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$. Is S linearly independent?

Solution. The set S is linearly dependent since $3\mathbf{v_1} + \mathbf{v_2} - \mathbf{v_3} = \mathbf{0}$. □

Example 1.5.3. Consider the polynomials $\mathbf{p_1} = 1 - x$, $\mathbf{p_2} = 5 + 3x - 2x^2$, $\mathbf{p_3} = 1 + 3x - x^2$ in P_2 (polynomials of degree less than or equal 2.) Is $S = \{\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}\}$ linearly independent?

Solution. The set S is linearly dependent since $3\mathbf{p}_1 - \mathbf{p}_2 + 2\mathbf{p}_3 = \mathbf{0}$. □

Example 1.5.4. Consider the following vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$. Is $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ a linearly independent set?

Solution. In terms of components the vector equation

$$a\mathbf{i} + b\mathbf{j} + c\mathbf{k} = \mathbf{0}$$

becomes

$$a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = (0, 0, 0)$$

or

$$(a, b, c) = (0, 0, 0).$$

From here we deduce that $a = b = c = 0$. So the set $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is linearly independent. Similarly we can show that the vectors $\{\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, \ldots, 0, 1)\}$ form a linearly independent set in $\mathbb{R}^n$. □

Example 1.5.5. Determine whether the following vectors are linearly independent in $\mathbb{R}^3$:

$$\mathbf{v_1} = (1, -2, 3),\ \mathbf{v_2} = (5, 6, -1) \quad \text{and} \mathbf{v_3} = (3, 2, 1).$$

Solution. In terms of components the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0}$$

becomes

$$k_1(1,-2,3) + k_2(5,6,-1) + k_3(3,2,1) = (0,0,0)$$

or equivalently

$$(k_1 + 5k_2 + 3k_3, -2k_1 + 6k_2 + 2k_3, 3k_1 - k_2 + k_3) = (0,0,0).$$

Equating corresponding components

$$\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0. \end{aligned}$$

Thus $\mathbf{v_1}, \mathbf{v_2}$, and $\mathbf{v_3}$ form a linearly dependent set if and only if this system has a nontrivial solution, or a linearly independent set if and only if it has only the trivial solution. Solving this system yields

$$k_1 = -\frac{1}{2}t,\ k_2 = -\frac{1}{2}t,\ k_3 = t.$$

Thus, the system has nontrivial solutions and $\mathbf{v_1}, \mathbf{v_2}$, and $\mathbf{v_3}$ form a linearly dependent set. Alternatively, we could show the existence of nontrivial solutions without solving the system by showing that the coefficient matrix has determinant zero and consequently is not invertible (verify). □

Example 1.5.6. Let $\mathbf{p_0} = 1, \mathbf{p_1} = x, \mathbf{p_2} = x^2, \ldots, \mathbf{p_n} = x^n$ in P_n. Show that $S = \{\mathbf{p_0}, \mathbf{p_1}, \mathbf{p_3}, \ldots, \mathbf{p_n}\}$ forms a linearly independent set in P_n.

Solution. Assume that some linear combination of the polynomials $\mathbf{p_0} = 1, \mathbf{p_1} = x, \mathbf{p_2} = x^2, \ldots, \mathbf{p_n} = x^n$ is zero, say

$$a_0\mathbf{p_0} + a_1\mathbf{p_1} + a_2\mathbf{p_2} + \ldots + a_n\mathbf{p_n} = \mathbf{0}$$

or equivalently

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n = \mathbf{0} \qquad \text{for all } x \text{ in } (-\infty, \infty). \tag{1.4}$$

We must show that

$$a_0 = a_1 = a_2 = \ldots = a_n = 0.$$

To see that this is so, recall that a nonzero polynomial of degree n has at most n distinct roots. But this implies that $a_0 = a_1 = a_2 = \ldots = a_n = 0$; otherwise, it would follow from Equation (1.4) that $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ is a nonzero polynomial with infinitely many roots. □

Why do we use the words linear dependent and linear independent? The term "linear dependent" suggests that the vectors "dependent" on each other in some way. This is stated in the following theorem.

Theorem 1.5.7. *A set S of two or more vectors is:*

(a) Linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.

(b) Linearly independent if and only if no vectors in S is expressible as a linear combination of the other vectors in S.

Proof. We shall prove part (a) and leave the proof of part (b) as an exercise. (a) Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be a set with two or more vectors. If we assume that S is linearly dependent, then there are scalars $k_1, k_2, \ldots, k_n$, not all zero, such that

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n} = \mathbf{0}. \tag{1.5}$$

In particular, suppose that $k_1 \neq 0$. Then Equation (1.5) can be rewritten as

$$\mathbf{v_1} = \left(-\frac{k_2}{k_1}\right)\mathbf{v_2} + \left(-\frac{k_3}{k_1}\right)\mathbf{v_3} + \ldots + \left(-\frac{k_n}{k_1}\right)\mathbf{v_n}$$

which expresses $\mathbf{v}_1$ as a linear combination of the other vectors in S. Similarly, if $k_j \neq 0$ in Equation (1.5) for some $j = 2, 3, \ldots, n$, then $\mathbf{v}_j$ is expressible as a linear combination of the other vectors in S.
Conversely, assume that at least one vector in S is expressible as a linear combination of the other vectors. In particular, suppose that

$$\mathbf{v}_1 = c_2\mathbf{v_2} + c_3\mathbf{v_3} + \ldots + c_n\mathbf{v_n}, \text{ so } \mathbf{v}_1 - c_2\mathbf{v_2} - c_3\mathbf{v_3} - \ldots - c_n\mathbf{v_n} = \mathbf{0}.$$

It follows that S is linearly dependent since Equation (1.5) is satisfied by

$$k_1 = 1,\ k_2 = -c_2, \ldots, k_n = -c_n$$

which are not all zero. The proof in the case where some vector other than $\mathbf{v}_1$ is expressible as a linear combination of the other vectors in S is similar. □

Example 1.5.8. In Example 1.5.2 we saw that the vectors $\mathbf{v_1} = (2, -1, 0, 3)$, $\mathbf{v_2} = (1, 2, 5, -1)$ and $\mathbf{v_3} = (7, -1, 5, 8)$, are linearly dependent. It follows from Theorem 1.5.7 that at least one these vectors is expressible as a linear combination of the other two. In this example each vector is expressible as a linear combination of the other two since it follows from the equation $3\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$ (see Example 1.5.2) that

$$\begin{aligned} \mathbf{v}_1 &= -\frac{1}{3}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_3 \\ \mathbf{v}_2 &= -3\mathbf{v}_1 + \mathbf{v}_3 \\ \mathbf{v}_3 &= 3\mathbf{v}_1 + \mathbf{v}_2. \end{aligned}$$

Example 1.5.9. In Example 1.5.4 we saw that the vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$ form a linearly independent set. Thus, it follows from Theorem 1.5.7 that at least one these vectors is expressible as a linear combination of the other two. To see directly that this is so, suppose that $\mathbf{k}$ is expressible as

$$\mathbf{k} = k_1\mathbf{i} + k_2\mathbf{j}.$$

Then, in terms of components,

$$(0, 0, 1) = k_1(1, 0, 0) + k_2(0, 1, 0)$$

or

$$(0, 0, 1) = (k_1, k_2, 0).$$

But this equation is not satisfied by any values of k_1 and k_2, so $\mathbf{k}$ cannot be expressed as a linear combination of $\mathbf{i}$ and $\mathbf{j}$. Similarly, $\mathbf{i}$ is not expressible as a linear combination of $\mathbf{j}$ and $\mathbf{k}$, and $\mathbf{j}$ is not expressible as a linear combination of $\mathbf{i}$ and $\mathbf{k}$.

The following theorem gives two simple facts about linear independence that are important to know

Theorem 1.5.10. *(a) A finite set of vectors that contains the zero vector is linearly dependent.*

(b) A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.

Proof. We shall prove part (a) and leave the proof of part (b) as an exercise.
(a) For any vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$, the set $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}, \mathbf{0}\}$ is linearly dependent since the equation

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_r + 1(\mathbf{0}) = \mathbf{0}$$

expresses the zero vector $\mathbf{0}$ as a linear combination of the vectors in S with coefficients that are not all zero. □

Example 1.5.11. The functions $\mathbf{f}_1 = x$ and $\mathbf{f}_2 = \sin x$ form a linearly independent set of vectors in $F(-\infty, \infty)$, since neither function is a constant multiple of the other.

Linear independence has some useful geometric interpretations in $\mathbb{R}^2$ and $\mathbb{R}^3$:

- In $\mathbb{R}^2$ or $\mathbb{R}^3$, a set of two vectors is linearly independent if and only if the vectors do not lie on the same line when they are placed with their initial points at the origin.
- In $\mathbb{R}^3$, a set of three vectors is linearly independent if and only if the vectors are no coplanar (i.e., do not lie on the same plane, or no vector is a linear combination of the other two vectors.)

The next theorem looks at linear independence in the vector space $\mathbb{R}^n$. To prove the theorem we need the following result about homogeneous systems of linear equations:

Result 1.5.12. In a homogeneous system of n linear equations in r unknowns, if $r > n$ then the system has infinitely many solutions.

Theorem 1.5.13. *Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ be a set of vectors in $\mathbb{R}^n$. If $r > n$, then S is linearly dependent.*

Proof. Suppose that

$$\begin{aligned} \mathbf{v}_1 &= (v_{11}, v_{12}, \ldots, v_{1n}) \\ \mathbf{v}_2 &= (v_{21}, v_{22}, \ldots, v_{2n}), \\ &\vdots \\ \mathbf{v}_r &= (v_{r1}, v_{r2}, \ldots, v_{rn}). \end{aligned} \tag{1.6}$$

Consider the equation

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n} = \mathbf{0}.$$

As illustrated in Example 1.5.5, we express both sides of the system 1.6 in terms of components and then equate corresponding components, we obtain the system

$$\begin{array}{rcl}
v_{11}k_1 + v_{21}k_2 + v_{31}k_3 + \cdots + v_{r1}k_r & = & 0 \\
v_{12}k_1 + v_{22}k_2 + v_{32}k_3 + \cdots + v_{r2}k_r & = & 0 \\
 & \vdots & \\
v_{1n}v_1 + v_{2n}k_2 + v_{3n}k_3 + \cdots + v_{rn}k_r & = & 0.
\end{array}$$

This is a homogeneous system of n equations in the r unknowns $k_1, k_2, \ldots, k_r$. Since $r > n$, it follows from Result 1.5.12 that the system has nontrivial solutions. Therefore, $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ is a linearly dependent set. $\square$

Theorem 1.5.13 tells us that a set in $\mathbb{R}^2$ with more than two vectors is linearly dependent, and a set in $\mathbb{R}^3$ with more than three vectors is linearly dependent.

1.6 Linear independence of functions

Sometimes linear dependence of functions can be deduced from known identities. For example, the functions

$$\mathbf{f}_1 = \sin^2 x,\ \mathbf{f}_2 = \cos^2 x,\ \text{ and } \mathbf{f}_3 = 5$$

form a linearly dependent set in $F(-\infty, \infty)$, since the equation

$$5\mathbf{f}_1 + 5\mathbf{f}_2 - \mathbf{f}_3 = 5\sin^2 x + 5\cos^2 x - 5 = 5(\sin^2 x + \cos^2 x) - 5 = 5 - 5 = \mathbf{0}$$

expresses $\mathbf{0}$ as a linear combination of $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$ with coefficients that are not all zero. Although there is no general method that can be used to establish linear independence or linear dependence of functions in $F(-\infty, \infty)$, we shall develop a theorem that can sometimes be used show that a given set of functions is linearly independent. If $\mathbf{f}_1 = f_1(x)$, $\mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are $n-1$ differentiable functions on the interval $(-\infty, \infty)$, then the determinant

$$W(x) = \begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \ldots & f_n^{n-1}(x) \end{bmatrix}$$

is called **Wronskian** of $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$. Suppose, that $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$ are linearly dependent, then there exist scalars $k_1, k_2, \ldots, k_n$ *not all zero*, such that

$$k_1 f_1(x) + k_2 f_2(x) + \ldots + k_n f_n(x) = \mathbf{0}$$

for all x in the interval $(-\infty, \infty)$. Combining this equation obtained by $n-1$ successive differentiations yields

$$\begin{aligned} k_1f_1(x) + k_2f_2(x) + \cdots + k_nf_n(x) &= \mathbf{0} \\ k_1f_1'(x) + k_2f_2'(x) + \ldots + k_nf_n'(x) &= \mathbf{0} \\ &\vdots \\ k_1f_1^{n-1}(x) + k_2f_2^{n-1}(x) + \ldots + k_nf_n^{n-1}(x) &= \mathbf{0}. \end{aligned}$$

Thus, the linear dependence of $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ implies that the linear system

$$\begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \ldots & f_n^{n-1}(x) \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

has a nontrivial solution for every x in the interval $(-\infty, \infty)$. Thus the determinant of the given matrix is zero (not invertible). This information has shown:

- If $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ are linearly dependent then the Wronskian is zero.

  An alternate way of stating this is:

- If the Wronskian is nonzero then $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ are linearly independent.

This is stated in the following theorem:

Theorem 1.6.1. *If the functions $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ have $n-1$ continuous derivatives on the interval $(-\infty, \infty)$, and if the Wronskian of the functions is not identically zero, then these functions form a linearly independent set of vectors in $C'(-\infty, \infty)$.*

Example 1.6.2. Show that the functions $\mathbf{f}_1 = x$ and $\mathbf{f}_2 = \sin x$ form a linearly independent set of vectors in $C'(-\infty, \infty)$.

Solution. In Example 1.5.8 we showed that these vectors form a linearly independent set of vectors in $F(-\infty, \infty)$, by noting that neither vector is a constant multiple of the other. Now, for illustrative reasons, we obtain this result using Theorem 1.6.1. The Wronskian is

$$W(x) = \begin{vmatrix} x & \sin x \\ 1 & \cos x \end{vmatrix} = x\cos x - \sin x.$$

This function does not have a value zero for all x in the interval $(-\infty, \infty)$, so $\mathbf{f}_1$ and $\mathbf{f}_2$ form a linearly independent set.

□

Example 1.6.3. Show that the functions $\mathbf{f}_1 = 1$ and $\mathbf{f}_2 = e^x$ and $\mathbf{f}_3 = e^{2x}$ form a linearly independent set of vectors in $C'(-\infty, \infty)$.

Solution. The Wronskian

$$W(x) = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x} \neq 0.$$

The Wronskian does not have a value zero for all x (in fact, for any x) in the interval $(-\infty, \infty)$, so $\mathbf{f}_1$, $\mathbf{f}_2$, and $\mathbf{f}_3$ form a linearly independent set. □

Note 1.6.4. The converse of Theorem 1.6.1 is false. If the Wronskian of $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ is identically zero on $(-\infty, \infty)$, then no conclusion can be reached about the linear independence of $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$. This set of vectors may be linearly dependent or linearly independent.

1.7 Basis and dimension

We usually think of a line as being 1-dimensional, a plane as 2-dimensional and the space around us as 3-dimensional. It is the primary purpose of this section to make this intuitive notion of "dimension" precise.

In a plane, any vector (a, b) in $\mathbb{R}^2$ is a linear combination of $\mathbf{i}$ and $\mathbf{j}$:

$$(a, b) = a(1, 0) + b(0, 1) = a\mathbf{i} + b\mathbf{j}.$$

Take any two (non-collinear) vectors w_1 and w_2 in $\mathbb{R}^2$. Any vector $\mathbf{w}$ in $\mathbb{R}^2$ is a unique linear combination of w_1 and w_2. The following is an important definition which will make the preceding ideas more precise and enable us to extend the concept of a coordinate system to general vector spaces.

Definition 1.7.1. If V is any vector space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a set of vectors in V, then S is called a **basis** for V if the following two conditions hold:

(a) S is linearly independent.

(b) S spans V.

A basis is the vector space generalization of coordinate system in 2-space and 3-space. The following theorem will help us see why this is so.

Theorem 1.7.2. *If $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for a vector space V, then every vector $\mathbf{v}$ in V can be expressed in the form $\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n}$ in exactly one way.*

Proof. Since S spans V, it follows from the definition of a spanning set that every vector in V is expressible as a linear combination of the vectors in S. To see that there is only one way to express a vector as a linear combination of the vectors in S, suppose that some vector $\mathbf{v}$ can be written as

$$\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \ldots + k_n\mathbf{v_n}$$

and also as

$$\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}.$$

Subtracting the second equation from the first gives

$$\mathbf{0} = (k_1 - c_1)\mathbf{v_1} + (k_2 - c_2)\mathbf{v_2} + \ldots + (k_n - c_n)\mathbf{v_n}.$$

Since the right side of this equation is a linear combination of vectors in S, the linear independence of S implies that

$$k_1 - c_1 = 0,\ k_2 - c_2 = 0,\ \dots,\ k_n - c_n = 0,$$

that is

$$k_1 = c_1,\ k_2 = c_2,\ \dots,\ k_n = c_n.$$

Thus, the two expressions for $\mathbf{v}$ are the same. $\square$

Definition 1.7.3. If $S = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is a basis for a vector space V, and

$$\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \dots + k_n\mathbf{v_n}$$

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $k_1,\ k_2,\ \dots,\ k_n$ are called the **coordinates** of $\mathbf{v}$ relative to the basis S. The vector $(k_1,\ k_2,\ \dots,\ k_n)$ in $\mathbb{R}^n$ constructed from these coordinates is called the **coordinate vector of $\mathbf{v}$ relative to the basis** S; it is denoted

$$(\mathbf{v})_S = (k_1,\ k_2,\ \dots,\ k_n).$$

Example 1.7.4. Let $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, where $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k} = (0, 0, 1)$. Show that S is a basis for $\mathbb{R}^3$ and for any $\mathbf{v} \in \mathbb{R}^3$, find $(\mathbf{v})_S$.

Solution. In Example 1.5.4 we showed that if

$$\mathbf{i} = (1, 0, 0),\ \mathbf{j} = (0, 1, 0),\ \text{ and } \mathbf{k} = (0, 0, 1)$$

then $S = \{\mathbf{i},\ \mathbf{j},\ \mathbf{k}\}$ is a linearly independent set in $\mathbb{R}^3$. This set also spans $\mathbb{R}^3$ since any vector $\mathbf{v}$ can be written as

$$\mathbf{v} = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}. \tag{1.7}$$

Thus, S is a basis for $\mathbb{R}^3$; it is called the **standard basis** for $\mathbb{R}^3$. Looking at the coefficients of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in Equation 1.7, it follows that the coordinate of $\mathbf{v}$ relative to the standard basis are a, b, and c, so

$$(\mathbf{v})_S = (a, b, c).$$

$\square$

Example 1.7.5. Let $S = \{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}\}$, where $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, 0, \dots, 1)$ in $\mathbb{R}^n$. Show that S is a basis for $\mathbb{R}^n$ and for any $\mathbf{v} \in \mathbb{R}^n$, find $(\mathbf{v})_S$.

Solution. In Example 1.5.4 we showed that if

$$\mathbf{e}_1 = (1, 0, \dots, 0),\ \mathbf{e}_2 = (0, 1, 0, \dots, 0),\ \dots,\ \mathbf{e}_n = (0, \dots, 0, 1)$$

then $S = S = \{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}\}$ is a linearly independent set in $\mathbb{R}^n$. This set also spans $\mathbb{R}^n$ since any vector $\mathbf{v}$ can be written as

$$\mathbf{v} = (v_1, v_2, \dots, v_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n. \tag{1.8}$$

Thus, S is a basis for $\mathbb{R}^n$; it is called the **standard basis** for $\mathbb{R}^n$. It follows from Equation 1.8, that the coordinate of $\mathbf{v}$ relative to the standard basis are $v_1, v_2, \dots, v_n$, so

$$(\mathbf{v})_S = (v_1, v_2, \dots, v_n).$$

Thus a vector $\mathbf{v}$ and its coordinates relative to the standard basis for $\mathbb{R}^n$ are the same. $\square$

Remark 1.7.6. In $\mathbb{R}^2$ and $\mathbb{R}^3$, the standard basis vectors are denoted by $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ rather than $\mathbf{e_1}$, $\mathbf{e_2}$ and $\mathbf{e_3}$.

It is not always the case that $(\mathbf{v})_S$ and $\mathbf{v}$ are the same, as is seen in the following example.

Example 1.7.7. Let $v_1 = (1, 2, 1)$, $v_2 = (2, 9, 0)$ and $v_3 = (3, 3, 4)$. Show that $S = \{v_1, v_2, v_3\}$ is a basis for $\mathbb{R}^3$. For any $\mathbf{v} \in \mathbb{R}^3$, find $(\mathbf{v})_S$.

Solution. To show that the set S spans $\mathbb{R}^3$, we must show that an arbitrary vector $\mathbf{b} = (b_1, b_2, b_3)$ can be expressed as a linear combination

$$\mathbf{b} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

of the vectors in S. Expressing this equation in terms of components gives

$$(b_1, b_2, b_3) = k_1(1, 2, 1) + k_2(2, 9, 0) + k_3(3, 3, 4)$$

or

$$(b_1, b_2, b_3) = (k_1 + 2k_2 + 3k_3, 2k_1 + 9k_2 + 3k_3, k_1 + 4k_3)$$

or on equating corresponding components

$$\begin{aligned} k_1 + 2k_2 + 3k_3 &= b_1 \\ 2k_1 + 9k_2 + 3k_3 &= b_2 \\ k_1 + 4k_3 &= b_3. \end{aligned} \tag{1.9}$$

Thus, to show that S spans $\mathbb{R}^3$, we must demonstrate that the system (1.9) has a solution for all choices of $\mathbf{b} = (b_1, b_2, b_3)$. To prove that S is linearly independent, we must show that the only solution of

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0} \tag{1.10}$$

is $k_1 = k_2 = k_3 = 0$. As above, if (1.10) is expressed in terms of components, the verification of independence reduces to showing that the homogeneous system

$$\begin{aligned} k_1 + 2k_2 + 3k_3 &= 0 \\ 2k_1 + 9k_2 + 3k_3 &= 0 \\ k_1 + 4k_3 &= 0 \end{aligned} \tag{1.11}$$

has only the trivial solution. Observe that the systems (1.9) and (1.11) have the same coefficient matrix. Thus, we can simultaneously prove that S is linearly independent and spans $\mathbb{R}^3$ by showing that in systems (1.9) and (1.11) the matrix of coefficients

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix}$$

has nonzero determinant. But

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{vmatrix} = -1.$$

So S is a basis for $\mathbb{R}^3$. $\square$

Example 1.7.8. Let S be the basis in Example 1.7.7.

(a) Find the coordinate vector of $\mathbf{v} = (5, -1, 9)$ with respect to S.

(b) Find the vector $\mathbf{v}$ in $\mathbb{R}^3$ whose coordinate vector with respect to the basis S is $(\mathbf{v})_S = (-1, 3, 2)$.

Solution. (a) We must find scalars $k_1 k_2, k_3$ such that

$$\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

or, in terms of components,

$$(5, -1, 9) = k_1(1, 2, 1) + k_2(2, 9, 0) + k_3(3, 3, 4).$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 2k_2 + 3k_3 &= 5 \\ 2k_1 + 9k_2 + 3k_3 &= -1 \\ k_1 + 4k_3 &= 9. \end{aligned}$$

Solving this system we obtain $k_1 = 1,\ k_2 = -1,\ k_3 = 2$ (verify). Therefore, $(\mathbf{v})_S = (1, -1, 2)$.
(b) Using the definition of coordinate vector $(\mathbf{v})_S$, we obtain

$$\begin{aligned} \mathbf{v} &= (-1)\mathbf{v}_1 + 3\mathbf{v}_2 + 2\mathbf{v}_3 \\ &= (-1)(1, 2, 1) + 3(2, 9, 0) + 2(3, 3, 4) = (11, 31, 7). \end{aligned}$$

□

Example 1.7.9. (a) Show that $S = \{\mathbf{p_0} = 1, \mathbf{p_1} = x, \mathbf{p_2} = x^2, \ldots, \mathbf{p_n} = x^n\}$ is a basis for the vector space P_n of polynomials of the form $a_0 + a_1 x + \ldots + a_n x^n$.

(b) Find the coordinate vector of the polynomial $\mathbf{p} = a_0 + a_1 x + a_2 x^2$ with respect to the basis $S = \{1, x, x^2\}$ of P_2.

Solution. (a) We showed in Example 1.4.7 that S spans P_n and in Example 1.5.6 that S is a linearly independent set. Thus, S is a basis for the vector space P_n; it is called the **standard basis for** P_n.
(b) The coordinates of $\mathbf{p} = a_0 + a_1 x + a_2 x^2$ are the scalar coefficients of the basis vectors $1, x$, and x^2, so $(\mathbf{p})_S = (a_0, a_1, a_2)$. □

Example 1.7.10. Let

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The set $S = \{M_1, M_2, M_3, M_4\}$ is a basis for the vector space M_{22} of 2×2 matrices. To see that S spans M_{22} note that an arbitrary vector (matrix)

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

can be written as

$$\begin{aligned}\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= aM_1 + bM_2 + cM_3 + dM_4.\end{aligned}$$

To see that S is linearly independent, assume that

$$aM_1 + bM_2 + cM_3 + dM_4 = \mathbf{0}.$$

That is,

$$a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

It follows that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus, $a = b = c = d = 0$, so S is linearly independent. The basis S in this example is called **standard basis for** M_{22}. More generally, the **standard basis for** M_{mn} consists of the mn different matrices with a single 1 and zeros for the remaining entries.

Example 1.7.11. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a linearly independent set in a vector space V, then S is a basis for the subspace $\mathrm{span}(S)$ since the set S spans $\mathrm{span}(S)$ by definition of $\mathrm{span}(S)$.

Definition 1.7.12. A nonzero vector space V is called **finite-dimensional** if it contains a finite set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ that forms a basis. If no such set exists, V is called **infinite-dimensional**. In addition, we shall regard the zero vector space to be finite-dimensional.

Example 1.7.13. By Examples 1.7.5, 1.7.9 and 1.7.10, the vector spaces $\mathbb{R}^n$, P_n and M_{mn} are finite-dimensional. The vector spaces $F(-\infty, \infty)$, $C(-\infty, \infty)$, $C^m(-\infty, \infty)$ and $C^\infty(-\infty, \infty)$ are infinite-dimensional.

Theorem 1.7.14. *Let V be a finite dimensional vector space with basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Then:*
(a) Every set of vectors with more than n vectors is linearly dependent.
(b) No set with fewer than n vectors spans V.

Proof. (a) Let $S' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ be an arbitrary set of m vectors of V, with $m > n$. We need to show that S' is a linearly dependent set. Since S is a basis for V, each w_i can be expressed as a linear combination of the vectors of S, i.e.,

$$\begin{aligned} \mathbf{w}_1 &= a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + a_{13}\mathbf{v}_3 + \cdots + a_{1n}\mathbf{v}_n, \\ \mathbf{w}_2 &= a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + a_{23}\mathbf{v}_3 + \cdots + a_{2n}\mathbf{v}_n, \\ &\vdots \\ \mathbf{w}_n &= a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + a_{m3}\mathbf{v}_3 + \cdots + a_{mn}\mathbf{v}_n. \end{aligned} \tag{1.12}$$

To show that S' is linearly independent we must find scalars not all zero, $k_1, k_2, \ldots, k_m$ such that,

$$k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \ldots + k_{n-1}\mathbf{w}_{n-1} + k_n\mathbf{w}_n = 0. \tag{1.13}$$

Using system (1.12) we re-write Equation (1.13) as

$$\begin{aligned} & (k_1a_{11}+k_2a_{12}+\ldots+k_ma_{1m})\mathbf{v}_1+(k_1a_{21}+k_2a_{22}+\ldots+k_ma_{2m})\mathbf{v}_2 \\ + \quad & \ldots+(k_1a_{n1}+k_2a_{n2}+\ldots+k_ma_{nm})\mathbf{v}_n=0. \end{aligned}$$

Since S is a linearly independent set (recall S is a basis), we have that the only solution to (1.13) is the trivial, and so all is reduced to showing that there are scalars $k_1,\, k_2,\, \ldots, k_m$ not all zero which satisfy the system

$$\begin{aligned} k_1a_{11}+k_2a_{12}x_2+k_3a_{13}x_3+\cdots+k_ma_{1m}&=0 \\ k_1a_{21}+k_2a_{22}x_2+k_3a_{23}x_3+\cdots+k_ma_{2m}&=0 \\ &\vdots \\ k_1a_{m1}+k_2a_{m2}x_2+k_3a_{m3}x_3+\cdots+k_ma_{nm}&=0. \end{aligned} \tag{1.14}$$

Since the system of equations (1.14), has more unknowns than equations, it follows from Result 1.5.12 that the system has nontrivial solutions. Hence S' is linearly dependent.

(b) Let $S'=\{\mathbf{w}_1,\, \mathbf{w}_2,\, \ldots, \mathbf{w}_m\}$ be an arbitrary linearly independent set of m vectors of V, with $m<n$. We need to show that S' does not span V. We prove by contradiction. Suppose that S' spans V we must get a contradiction with the fact that S is linearly independent. If S' spans V then each vector of V can be can be expressed as a linear combination of the vectors of S'. In particular, each vector v_i of S (recall that S is a basis V, and so it spans V) is a linear combination of the vectors of S'. So we have

$$\begin{aligned} \mathbf{v}_1&=a_{11}\mathbf{w}_1+a_{21}\mathbf{w}_2+a_{31}\mathbf{w}_3+\cdots+a_{m1}\mathbf{w}_m \\ \mathbf{v}_2&=a_{12}\mathbf{w}_1+a_{22}\mathbf{w}_2+a_{32}\mathbf{w}_3+\cdots+a_{m2}\mathbf{w}_m \\ &\vdots \\ \mathbf{v}_n&=a_{1n}\mathbf{w}_1+a_{2n}\mathbf{w}_2+a_{3n}\mathbf{w}_3+\cdots+a_{mn}\mathbf{w}_m. \end{aligned} \tag{1.15}$$

To find the contradiction we will show that there are scalars not all zero, $c_1,\, c_2,\, \ldots, c_n$ such that,

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_{n-1}\mathbf{v}_{n-1}+c_n\mathbf{v}_n=\mathbf{0}. \tag{1.16}$$

Using system (1.15) we re-write Equation (1.16) as

$$\begin{aligned} & (c_1a_{11}+c_2a_{12}+\ldots+c_na_{1n})\mathbf{w}_1+(c_1a_{21}+c_2a_{22}+\ldots+c_na_{2n})\mathbf{w}_2 \\ + \quad & \ldots+(c_1a_{m1}+c_2a_{m2}+\ldots+c_na_{mn})\mathbf{w}_m=0. \end{aligned}$$

Since S' is a linearly independent set, we have that the the only solution to (1.16) is the trivial, and so all is reduced to showing that there are scalars $c_1,\, c_2,\, \ldots, c_n$ not all zero which satisfy the system

$$\begin{aligned} c_1a_{11}+c_2a_{12}+c_3a_{13}+\cdots+c_na_{1n}&=0 \\ c_1a_{21}+c_2a_{22}+c_3a_{23}+\cdots+c_na_{2n}&=0 \\ &\vdots \\ c_1a_{m1}+c_2a_{m2}+c_3a_{m3}+\cdots+c_na_{mn}&=0. \end{aligned} \tag{1.17}$$

Since the system of equations (1.17), has more unknowns than equations, it follows from Result 1.5.12 that the system has nontrivial solutions. Hence S is linearly dependent, and so we get a contradiction with the fact that S is linearly independent. So our supposition was false and the result is true. $\square$

Theorem 1.7.14 tells us that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for the vector space V, then any other basis S' cannot have more elements than S, and cannot have fewer elements than S. Hence S' must have the same number of elements as S. This important result is stated in the following:

Theorem 1.7.15. *All bases for a finite-dimensional vector space have the same number of vectors.*

Since all bases for a vector space have the same number of elements, we can make the following definition related to dimension.

Definition 1.7.16. The **dimension** of a finite-dimensional vector space V, denoted $\dim(V)$, is defined to be the number of vectors in a basis for V. In addition, we define the zero vector space to have dimension zero.

Example 1.7.17.

$$\begin{aligned}
\dim(\mathbb{R}^n) &= n && \text{[the standard basis has } n \text{ vectors (Example 1.7.5)]}\\
\dim(P_n) &= n+1 && \text{[the standard basis has } n+1 \text{ vectors (Example 1.7.9(a))]}\\
\dim(M_{mn}) &= mn && \text{[(the standard basis has } mn \text{ vectors (Example 1.7.10)]}
\end{aligned}$$

Example 1.7.18. Determine a basis and the dimension of the solution space of the homogeneous system of equations:

$$\begin{aligned}
2x_1 + 2x_2 - x_3 \qquad\quad + x_5 &= 0,\\
-x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0\\
-x_1 - x_2 + 2x_3 \qquad\quad - x_5 &= 0\\
x_3 \;+ x_4 \;+ x_5 &= 0.
\end{aligned}$$

Solution. Writing the system of equations in matrix and performing Gauss elimination we take it to row echelon form as follows:

$$\begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ -1 & -1 & 2 & -3 & 1\\ 1 & 1 & -2 & 0 & -1\\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ 0 & 0 & 3 & -6 & 3\\ 0 & 0 & 3 & 0 & 3\\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ 0 & 0 & 1 & -2 & 1\\ 0 & 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ 0 & 0 & 1 & -2 & 1\\ 0 & 0 & 0 & -2 & 0\\ 0 & 0 & 0 & -3 & 0 \end{bmatrix}$$
$$\sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ 0 & 0 & 1 & -2 & 1\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1\\ 0 & 0 & 1 & -2 & 1\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last matrix shows that the system has two free variable, say x_2 and x_5. Now we write all the other variables in terms of these two, so that $x_1 = -x_2 - x_5$, $x_2 = x_2$, $x_3 = -x_5$, $x_4 = 0$, and $x_5 = x_5$. Thus the solution space is

$$\begin{aligned} S &= \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \,|\, x_1 = -x_2 - x_5,\ x_2 = x_2,\ x_3 = -x_5,\ x_4 = 0, x_5 = x_5\} \\ &= \{(-x_2 - x_5,\ x_2, -x_5, 0, x_5) \,|x_2,\ x_5 \in \mathbb{R}\} \\ &= x_2(-1, 1, 0, 0, 0) + x_5(-1, 0, -1, 0, 1) \,|\, x_2, x_5 \in \mathbb{R}\} \end{aligned}$$

Now consider $v_1 = (-1, 1, 0, 0, 0)$ and $v_2 = (-1, 0, -1, 0, 1)$ it can be easily noticed that the set $B = \{v_1,\ v_2\}$ is linearly independent. Since any vector of the form $(-x_2 - x_5,\ x_2, -x_5, 0, x_5)$ can be expressed as a linear combination of v_1 and v_2 we have that B is a spanning set. Hence B is a basis for the solution space of the given homogeneous system of equations. Since this set has two vectors, we have that the dimension of the solution space is 2. $\square$

1.8 Some fundamental theorems

The theorems that follow give results concerning relationships among the concepts of span, linear independence, basis and dimension.

Theorem 1.8.1. *(*Plus/Minus Theorem*) Let S be a nonempty set of vectors in a vector space V.*
(a) If S is a linearly independent set, and if $\mathbf{v}$ is a vector in V such that $\mathbf{v} \notin \operatorname{span}(S)$, then $S \cup \{\mathbf{v}\}$ is still linearly independent.
(b) If $\mathbf{v}$ is a vector in S that is a linear combination of the other vectors of S, then $S - \{\mathbf{v}\}$ or $S \setminus \{\mathbf{v}\}$ and S span the same space, i.e., $\operatorname{span}(S) = \operatorname{span}(S - \{\mathbf{v}\})$.

We have seen that for a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, v_n\}$ to be a basis for a vector space V, we need to show that it is linearly independent and that it spans V. But if we know that dim$(V) = n$, then we need to check just one condition, stated in the next theorem.

Theorem 1.8.2. *Let V be a vector space with $\dim(V) = n$, and let S be a set in V having exactly n vectors. Then S is a basis for V if either S is linearly independent or* span$(S) = V$.

Example 1.8.3. Show by inspection that
(a) $\mathbf{v}_1 = (1, 2)$ and $\mathbf{v}_2 = (4, 4)$ for a basis for $\mathbb{R}^2$.
(b) $\mathbf{v}_1 = (2, 0, -1)$, $\mathbf{v}_2 = (4, 0, 7)$ and $\mathbf{v}_3 = (-1, 1, 4)$ form a basis for $\mathbb{R}^3$.

Solution. (a) Since neither vector is a multiple of the other, the two vectors form a linearly independent set of vectors. Now by Theorem 1.8.2 it follows that they form a basis for $\mathbb{R}^2$.
(b) The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a set of linearly independent vectors in the plane xz(why?). The vector $\mathbf{v}_3$ is not in the plane xz. Therefore the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. Hence by Theorem 1.8.2 it follows that they form a basis for $\mathbb{R}^3$.

$\square$

Theorem 1.8.4. *Let S be a set of vectors in a finite dimensional vector space V.*
(a) If $\operatorname{span}(S) = V$, then S can be reduced to a basis for V by removing certain vectors from V.
(b) If S is a linearly independent set that is not a basis, then S can be enlarged to a basis for V by inserting certain vectors from V into S.

Proof. Follows easily from the Plus/minus Theorem. □

Theorem 1.8.5. *If W is a subspace of a finite dimensional vector space V, then* dim$(W) \leq \dim V$. *Moreover, if* $\dim(W) = \dim(V)$, *then* $V = W$.

Proof. Since V is a finite dimensional vector space, W is also finite dimensional. Suppose that $S = \{w_1, w_2, \ldots, w_m\}$ is a basis for W. Then either S is also a basis for V or not. If S is a basis for V then $\dim(W) = \dim(V) = m$. If S is not a basis for V then by Theorem 1.8.4(b) we can enlarge the linearly independent S to a basis for V so that $\dim(W) < \dim(V)$. In either cases we have $\dim(W) \leq \dim(V)$. If $\dim(W) = \dim(V)$, then S is a set of m linearly independent vectors in a vector space V of dimension m, and by Theorem 1.8.2 we have that S forms a basis for V. Hence $W = V$ (why?). □

Chapter 2

Euclidean n-space

2.1 Row space, column space, and nullspace

In this final section of this rather long chapter on vector spaces, we shall briefly discuss a number of naturally arising vector spaces associated with matrices. We have already encountered one such space, the nullspace of a matrix. Let us start with some definitions.

Definition 2.1.1. Let $A = [a_{ij}]$ be an $m \times n$ matrix.

1. The space spanned by the rows of A is called the **row space of** A(subspace of $\mathbb{R}^n$).
   The space spanned by the columns of A is called the **column space of** A(subspace of $\mathbb{R}^m$).

2. The **row rank of** A is the dimension of the row space of A.
   The **column rank of** A is the dimension of the column space of A.

3. The **nullspace of** A is the solution space of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$.

The following theorems for which we will not provide a proof are important.

Theorem 2.1.2. *Elementary row operations do not change the nullspace of a matrix.*

Theorem 2.1.3. *Elementary row operations do not change the row space of a matrix.*

Theorem 2.1.4. *Suppose A and B are row equivalent matrices. Then*
(a) A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.
(b) A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B forms a basis for the column space of B.

The next theorem helps to find bases for the row space and column spaces.

Theorem 2.1.5. *If a matrix R is in row echelon form, then the row vectors with leading 1's form a basis for the row space of R and the column vectors with leading 1's form a basis for the column space.*

Example 2.1.6. Find a basis for the row space and column space of
(a)

$$R = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(b)

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$$

Solution. (a) The matrix R is in row echelon form. From Theorem 2.1.5 it follows that the vectors

$$\begin{aligned} r_1 &= \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \end{bmatrix} \\ r_2 &= \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \end{bmatrix} \\ r_3 &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

form a basis for the row space of R and the vectors $c_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $c_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $c_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the column space of R.

(b) Since elementary row operations do no change the row space of a matrix, we can find a basis for the row space of A by determining a basis for the row space of any matrix in row echelon form obtained from A. Reducing A to row echelon form by elementary row operations we have $R = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. It follows by Theorem 2.1.5 that the non-zero row vectors of R form a basis for the row space of R and thus a basis for the row space of A. These vectors are

$$\begin{aligned} r_1 &= \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \end{bmatrix} \\ r_2 &= \begin{bmatrix} 0 & 0 & 1 & 3 & -2 & -6 \end{bmatrix} \\ r_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 5 \end{bmatrix}. \end{aligned}$$

Notice that A and R may have have different column spaces, thus we cannot obtain a basis for the column space of A directly from the column vectors of R. However, it follows from Theorem 2.1.4(b) that if we can find a set of column vectors of R which form a basis for R then the *corresponding* column vectors of A will form a basis for the column space of A. Since the first, third and fifth columns of R possess the leading 1's of the row vectors of R, it follows that these form a basis for the column space of R; hence the corresponding column vectors of A, namely $c_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix}$, $c_3 = \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix}$, $c_5 = \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix}$ form a basis for the column space of A.

$\square$

Example 2.1.7. Find a basis for the space spanned by the vectors
$v_1 = (1, -2, 0, 0, 3)$, $v_2 = (2, -5, -3, -2, 6)$, $v_3 = (0, 5, 15, 10, 0)$, $v_4 = (2, 6, 18, 8, 6)$.

Solution. The space spanned by the vectors is the row space of the matrix

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}.$$

Reducing this matrix to row echelon form we obtain

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The non-zero row vectors of this matrix are $w_1 = (1, -2, 0, 0, 3)$, $w_2 = (0, 1, 3, 2, 0)$, $w_3 = (0, 0, 1, 1, 0)$. These vectors form a basis of the row space and thus a basis for the subspace of $\mathbb{R}^5$ spanned by the vectors $v_1 = (1, -2, 0, 0, 3)$, $v_2 = (2, -5, -3, -2, 6)$, $v_3 = (0, 5, 15, 10, 0)$, $v_4 = (2, 6, 18, 8, 6)$.

$\square$

Remark 2.1.8. Note that in Example 2.6 the basis vectors for the column space of A consisted of columns of A but the basis vectors for the row space of A may not all be row vectors of A. The next example illustrates a method to find a basis for the row space of a matrix A which consists entirely of row vectors of A.

Example 2.1.9. Find a basis for the row space of A consisting entirely of row vectors from A :

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}.$$

Solution. We start by transposing A and so convert the row space of A into the column space of A^t. We then use the method of Example 2.6(b) to find a basis for the column space of A^t. Finally we transpose again to convert the column vectors to row vectors. Transposing A we get

$$A^t = \begin{bmatrix} 1 & -2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 8 \end{bmatrix}.$$

Now reducing this matrix to row echelon form we obtain

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first, second and fourth columns of this matrix possess the leading 1's, thus the corresponding vectors in A^t form a basis for A^t. These vectors are: $c_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix}$, $c_2 = \begin{bmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{bmatrix}$, $c_4 = \begin{bmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{bmatrix}$. Transposing these last vectors we have

$$\begin{aligned} r_1 &= \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \end{bmatrix} \\ r_2 &= \begin{bmatrix} 2 & -5 & -3 & -2 & 6 \end{bmatrix} \\ r_4 &= \begin{bmatrix} 2 & 6 & 18 & 8 & 6 \end{bmatrix} \end{aligned}$$

and they form a basis for the row space of A. □

Example 2.1.10. (a) Find a subset of the vectors
$v_1 = (1, -2, 0, 3)$, $v_2 = (2, -5, -3, 6)$, $v_3 = (0, 1, 3, 0)$, $v_4 = (2, -1, 4, -7)$, $(v_5 = (5, -8, 1, 2)$
that forms a basis for the space spanned by these vectors.
(b) Express a vector not in the basis as a linear combination of the basis vectors.

Solution. (a) We start by reducing the matrix

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ v_1 & v_2 & v_3 & v_4 & v_5 \end{bmatrix} \tag{2.1}$$

to row echelon form

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ w_1 & w_2 & w_3 & w_4 & w_5 \end{bmatrix}. \tag{2.2}$$

The leading 1's occur in columns 1, 2 and 4, so by Theorem 2.5 we have that the vectors $w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $w_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $w_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the column space of the matrix in row echelon form in (2.2), and so the vectors v_1, v_2 and v_4 form a basis for the starting matrix in (2.1).

(b) We start by expressing w_3 and w_5 as a linear combinations of the vectors w_1, w_2 and w_4. Inspecting (2.2) we have

$$\begin{aligned} w_3 &= 2w_1 - w_2 \\ w_5 &= w_1 + w_2 + w_4 \end{aligned}$$

which are named dependency relations. The corresponding relations for (2.1) are

$$\begin{aligned} v_3 &= 2v_1 - v_2 \\ v_5 &= v_1 + v_2 + v_4. \end{aligned}$$

□

Remark 2.1.11. Let $S = \{v_1, v_2, \ldots, v_n\}$ be a set of vectors in $\mathbb{R}^n$. The following procedure shows how to produce a subset of these vectors that forms a basis for $\mathrm{span}(S)$ and express those vectors in S not in the basis as a linear combination of the basis vectors.

Step 1	Form a matrix A having $v_1, v_2, \ldots, v_n$ as its column vectors.
Step 2	Reduce the matrix A to **reduced row-echelon form** R, and let $w_1, w_2, \ldots, w_k$ be the column vectors of R.
Step 3	Identify the columns in R that have the leading 1's. The corresponding columns of A are the basis vectors for $\mathrm{span}(S)$.
Step 4	Express those column vectors of R that do not belong to the basis as a linear combination (by inspection). Get corresponding equations for the column vectors of A.

2.2 Rank and nullity

Since the column space and row space are related to the leading 1's in the row-reduced matrix, we have the following theorem.

Theorem 2.2.1. *If A is any matrix, then the row space and the column space have the same dimension (i.e.,* dim*(row space) =*dim*(column space).)*

Proof. Follows from Theorem 2.1.5. □

Definition 2.2.2. The **rank** of a matrix A is the dimension of the column space (and dimension of row space). We denote this by **rank**(A). The dimension of the nullspace of A is called the nullity of A. We denote this by **nullity**(A).

Example 2.2.3. Find the rank and nullity of the matrix

$$A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}.$$

Solution. The reduced echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since there are two nonzero rows, or equivalently two leading 1's the row space and the column space of A are 2-dimensional subspaces, that is, rank$(A) = 2$.
Note: You may find it difficult to understand why the first row of the previous matrix in echelon form has changed to those numbers. That was done so that the variable x_2 could be written as a dependent variable. If you multiply the second row in your reduced row echelon form by 2 and add it to the the first you get precisely that the first row has those numbers. The other way to achieve that is by substituting x_2 into the equation for x_1.

To find the nullity of A we must determine the dimension of the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$. Using the matrix in reduced row echelon form we have

$$\begin{array}{llllllllll} x_1 & & - & 4x_3 & - & 28x_4 & - & 37x_5 & + & 13x_6 & = 0 \\ & x_2 & - & 2x_3 & - & 12x_4 & - & 16x_5 & + & 5x_6 & = 0 \end{array}$$

Solving this in terms of the free variables we have

$$\begin{array}{lllllllll} x_1 & & = 4x_3 & + & 28x_4 & + & 37x_5 & - & 13x_6 \\ & x_2 & = 2x_3 & + & 12x_4 & + & 16x_5 & - & 5x_6 \end{array}$$

A general solution to the system is

$$\begin{array}{lllllllll} x_1 & = 4r & + & 28s & + & 37t & - & 13u \\ x_2 & = 2r & + & 12s & + & 16t & - & 5u \\ x_3 & = r & & & & & & \\ x_4 & = & & s & & & & \\ x_5 & = & & & & t & & \\ x_6 & = & & & & & & u. \end{array}$$

Equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r \begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{2.3}$$

The four vectors on the right hand side of Equation (2.3) form a basis for the solution space of A, so that nullity$(A) = 4$. □

Theorem 2.2.4. *If A is any matrix, then* rank $(A) = \mathrm{rank}(A^t)$.

Proof. rank$(A) = \dim(\text{row space of}(A)) = \dim(\text{column space of}(A^t)) =$ rank(A^t). □

The following theorem establishes a relationship between the rank and nullity of a matrix.

Theorem 2.2.5. *(Dimension Theorem for matrices) If A is an $m \times n$ matrix, then rank (A) + nullity$(A) = n(n =$ no. of columns).*

Example 2.2.6. Let

$$A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}.$$

Find rank(A), nullity(A), rank(A) + nullity(A).

Solution. It follows from Example 2.2.3 that rank(A)=2 and nullity(A)=4, and so rank(A) + nullity$(A) = 6$ □

Let A be an $m \times n$ matrix of rank r.

Space	Dimension
Col space of A	r
Row space of A	r
Nullspace of A	$n - r$
Nullspace of A^t	$m - r$

2.3 Eigenvalues and eigenvectors

If A is an $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$, there is usually no relationship between $A\mathbf{x}$ and $\mathbf{x}$. However, there are often certain nonzero vectors $\mathbf{x}$ such that $\mathbf{x}$ and $A\mathbf{x}$ are scalar multiples of each other. Such vectors arise naturally in the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, economics, etc. In this section, we discuss how we find these vectors.

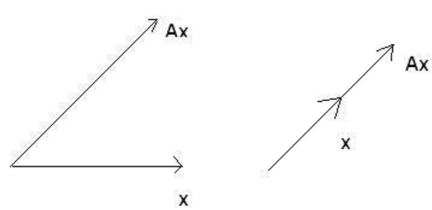

Fig. 1

Definition 2.3.1. Let A be an $n \times n$ matrix. A nonzero vector $\mathbf{x} \in \mathbb{R}^n$ is called an **eigenvector** of A if $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. The scalar λ is called an **eigenvalue** of A and $\mathbf{x}$ is the eigenvector of A corresponding to λ.

Note 2.3.2. If $\mathbf{x}$ is an eigenvector of A, then $A\mathbf{x} = \lambda\mathbf{x}$ for some λ. Thus the operator A compresses or stretches the vector $\mathbf{x}$ by a factor λ, and reverses $\mathbf{x}$ if $\lambda < 0$.

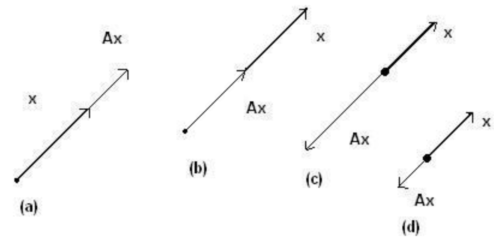

Fig. 2

In Fig.2 (a) $\lambda \geq 1$, (b) $0 \leq \lambda \leq 1$, (c) $\lambda \leq -1$, (d) $-1 \leq \lambda \leq 0$.

Example 2.3.3. The vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$:

Solution. Notice that

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3\mathbf{x}.$$

□

To find the eigenvector of an $n \times n$ matrix A we write

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} = \lambda I\mathbf{x} \\ &\Leftrightarrow \lambda I\mathbf{x} - A\mathbf{x} = 0 \\ &\Leftrightarrow (\lambda I - A)\mathbf{x} = 0 \quad (*) \end{aligned}$$

For λ to be an eigenvalue of A, we must find a nonzero vector $\mathbf{x}$ satisfying (*). But $\mathbf{x} = 0$ is already a solution to (*). Thus we need another solution. In this case $(\lambda I - A)$ will not be invertible (since (*) has more than one solution). Hence λ is an eigenvalue of A if and only if $\det(\lambda I - A) = 0$.

Definition 2.3.4. $\det(\lambda I - A) = 0$ is called the **characteristic equation** of A. The scalars satisfying this equation are the eigenvalues of A. $\det(\lambda I - A)$ is a polynomial in λ and hence called **characteristic polynomial** of A.

It can be shown that if A is an $n \times n$ matrix, then $\det(\lambda I - A)$ has degree n, and the coefficient of λ^n is 1; i.e., of the form $\lambda^n + c_1\lambda^{n-1} + \ldots + c_n$. By the Fundamental Theorem of Algebra, $\det(\lambda I - A) = 0$ has at most n roots; i.e., at most n eigenvalues.

Example 2.3.5. Find the eigenvalues of

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}:$$

Solution. The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{bmatrix} = (\lambda - 1)(\lambda - 2) - 12.$$

In addition the characteristic equation is $(\lambda - 1)(\lambda - 2) - 12 = 0$ and the eigenvalues of A are $\lambda = 5$ and $\lambda = -2$. $\square$

Example 2.3.6. Find the eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}:$$

Solution. The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4.$$

The eigenvalues of A must satisfy the equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0. \tag{2.4}$$

Equation (2.4) can be rewritten as

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1 = 0.$$

Finally the solution to Equation (2.4) is $\lambda = 4$, $\lambda = 2 + \sqrt{3}$, and $\lambda = 2 - \sqrt{3}$ which are in turn the eigenvalues of A. $\square$

Example 2.3.7. Find the eigenvalues of the upper triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}:$$

Solution. Recall that the determinant of a triangular matrix equals the product of the entries in the main diagonal. Thus we have

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & -a_{14} \\ 0 & \lambda - a_{22} & -a_{23} & -a_{24} \\ 0 & 0 & \lambda - a_{33} & -a_{34} \\ 0 & 0 & 0 & \lambda - a_{44} \end{bmatrix} = (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}).$$

In addition the characteristic equation is

$$(\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) = 0,$$

so that the eigenvalues are $\lambda = a_{11}$, $\lambda = a_{22}$, $\lambda = a_{33}$, and $\lambda = a_{44}$. These are precisely the entries in the main diagonal of A. $\square$

Note 2.3.8. This example gives us the following general theorem.

Theorem 2.3.9. *If A is an $n \times n$ triangular matrix (upper/lower) then the eigenvalues are the entries on the diagonal of A.*

Example 2.3.10. Find the eigenvalues of the lower triangular matrix

$$A = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -1 & \frac{2}{3} & 0 \\ 5 & -8 & -\frac{1}{4} \end{bmatrix}:$$

Solution. By inspection, the eigenvalues of A are :$\lambda = \frac{1}{2}$, $\lambda = \frac{2}{3}$, and $\lambda = -\frac{1}{4}$. $\square$

Remark 2.3.11. In practical problems, the matrix A is often so large that computing the characteristic polynomial is not practical. In this cases, various approximation techniques are used.

Finding bases for eigenspaces

We want to answer the following question: How do we find eigenvectors? The eigenvectors corresponding to λ are the nonzero vectors $\mathbf{x}$ satisfying $A\mathbf{x} = \lambda\mathbf{x}$. Therefore, the eigenvectors corresponding to λ are the nonzero vectors in the solution space of

$$(\lambda I - A)\mathbf{x} = \mathbf{0}.$$

This solution space is called the **eigenspace** of A corresponding to λ.

Example 2.3.12. Find the bases for the eigenspaces of

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}:$$

Solution. The characteristic polynomial of A is

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{vmatrix} = (\lambda - 2)\begin{vmatrix} \lambda & 2 \\ -1 & \lambda - 3 \end{vmatrix} \\ &= (\lambda - 2)[\lambda(\lambda - 3) + 2] = (\lambda - 2)^2(\lambda - 1). \end{aligned}$$

Hence the characteristic equation of A is

$$\det(\lambda I - A) = (\lambda - 2)^2(\lambda - 1) = \mathbf{0}.$$

The eigenvalues of A are the roots of the characteristic equation, namely $\lambda_1 = 2$ and $\lambda_2 = 1$.

By definition a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is an eigenvector of A corresponding to λ if and only if, $\mathbf{x}$ is a non-trivial solution for $(\lambda I - A)\mathbf{x} = \mathbf{0}$, that is, $\mathbf{x}$ is a non-trivial solution of

$$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We denote $\mathrm{E}\lambda_i$ the eigenspace corresponding to the eigenvalue λ_i.

I: To find E_2, we must find the nullspace of

$$\lambda_1 I - A = 2I - A = \begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence,

$$(2I - A)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Let $x_3 = s$. Then $x_1 = -s$ and $x_2 = t$. Thus,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Let

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

and

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then a basis for the nullspace of E_2 is $\{\mathbf{v}_1, \mathbf{v}_2\}$.

II : To find E_1, we must find the nullspace of

$$\lambda_2 I - A = I - A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & -1 \\ -1 & 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence,

$$(I - A)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Let $x_3 = s$. Then $x_2 = s$ and $x_1 = -2s$. Thus,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Let

$$\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Then a basis for the nullspace of E_1 is $\{\mathbf{v}_3\}$.

$\square$

Example 2.3.13. Find the bases for the eigenspaces of

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}:$$

Solution. From Example 2.3.5 we have that the eigenvalues of A are $\lambda_1 = 5$ and $\lambda_2 = -2$.

I : To find E_5, we must find the nullspace of

$$\lambda_1 I - A = 5I - A = \begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix} \sim \begin{bmatrix} 4 & -3 \\ 0 & 0 \end{bmatrix}.$$

Hence,

$$(5I - A)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 4 & -3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let $x_2 = s$. Then $x_1 = \frac{3}{4}s$. Thus,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = s \begin{bmatrix} \frac{3}{4} \\ 1 \end{bmatrix}.$$

Let

$$\mathbf{v}_1 = \begin{bmatrix} \frac{3}{4} \\ 1 \end{bmatrix}.$$

Then a basis for the nullspace of E_5 is $\{\mathbf{v}_1\}$.

II : To find E_{-2}, we must find the nullspace of

$$\lambda_2 I - A = -2I - A = \begin{bmatrix} -3 & -3 \\ -4 & -4 \end{bmatrix} \sim \begin{bmatrix} -3 & -3 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Hence,

$$(-2I - A)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let $x_2 = s$. Then $x_1 = -s$. Thus,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Let

$$\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then a basis for the nullspace of E_{-2} is $\{\mathbf{v}_2\}$.

$\square$

Eigenvalues of the powers of a matrix

Suppose that λ is an eigenvalue of A and $\mathbf{x}$ is an eigenvector of A corresponding to λ. Then

$$\begin{aligned} A^2\mathbf{x} &= A(A\mathbf{x}) = A(\lambda\mathbf{x}) \\ &= \lambda(A\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}. \end{aligned}$$

So $A^2\mathbf{x} = \lambda^2\mathbf{x}$ i.e., λ^2 is an eigenvalue for A^2 with corresponding eigenvector $\mathbf{x}$. Hence we have the following general theorem.

Theorem 2.3.14. *Let k be a positive integer, λ be an eigenvalue of A and $\mathbf{x}$ be an eigenvector of A corresponding to λ. Then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.*

Example 2.3.15. In Example 2.3.12 we showed that the eigenvalues of

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

are $\lambda = 2$ and $\lambda = 1$. So $\lambda = 2^7 = 128$ and $\lambda = 1^7 = 1$ are eigenvalues of A^7. Hence $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ are eigenvectors of A^7 corresponding to $\lambda = 2^7 = 128$ and $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of A^7 corresponding to $\lambda = 1^7 = 1$.

Eigenvalues and invertibility

The following theorem establishes a relationship between eigenvalues and the invertibility of a matrix.

Theorem 2.3.16. *A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.*

Proof. Suppose that A is an $n \times n$ matrix. Now observe that $\lambda = 0$ is a solution of the characteristic equation $\lambda^n + c_1\lambda^{n-1} + \ldots + c_n = 0$ if and only the constant term c_n equals zero. Thus, it suffices to prove that A is invertible if and only if $c_n \neq 0$. But,

$$\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \ldots + c_n. \tag{2.5}$$

Therefore letting $\lambda = 0$ in Equation (2.5) it follows that $\det(-A) = c_n$ or $(-1)^n\det(A) = c_n$. From this last equation it follows that $\det(A) = 0$ if and only if $c_n = 0$, which implies that A is invertible if and only if $c_n \neq 0$. □

Example 2.3.17. Is $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ invertible? (Do not use the determinant.)

Solution. From Example 2.3.12 we have that the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$. Since none of the eigenvalues is zero, it follows that A is invertible. □

2.4 Diagonalisation

In this section we are concerned with the problem of finding a basis for $\mathbb{R}^n$ that consists of eigenvectors of an $n \times n$ matrix A. Consider the following two, seemingly different, problems:

The Eigenvector Problem:

Given an $n \times n$ matrix A, does there exists a basis of $\mathbb{R}^n$ consisting of the eigenvectors of A?

The Diagonalisation Problem:

Given an $n \times n$ matrix A, does there exists an invertible matrix P, such that $P^{-1}AP$ is a diagonal matrix?

Let us first make some terminology available:

Definition 2.4.1. A square matrix A is said to be **diagonalisable** if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. P is said to **diagonalise** A.

The following theorem shows that the above two problems are in fact the same.

Theorem 2.4.2. *Let A be an $n \times n$ matrix. Then the following are equivalent:*

(a) A is diagonalisable.

(b) A has n linearly independent eigenvectors.

Proof. $(a) \Rightarrow (b)$: Suppose A is diagonalisable. Then there exists an invertible matrix $P = [p_{ij}]$ such that

$$P^{-1}AP = D, \tag{2.6}$$

where $D = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$ is a diagonal matrix. Multiplying by P on the left of both sides of Equation (2.6) we have $AP = PD$. But

$$\begin{aligned} AP &= PD \\ &= \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \dots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \lambda_3 p_{13} & \dots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \lambda_3 p_{23} & \dots & \lambda_n p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \lambda_3 p_{n3} & \dots & \lambda_n p_{nn} \end{bmatrix}. \end{aligned}$$

Let the columns of P be p_i where $1 \le i \le n$, then $AP = [Ap_1, Ap_2, \dots, Ap_n]$ and $PD = [\lambda_1 p_1, \lambda_1 p_2, \dots, \lambda_1 p_n]$. Equating we obtain $Ap_i = \lambda_i p_i$ for $1 \le i \le n$. Since P is an invertible matrix, each column of P is nonzero, because p_i where $1 \leq i \leq n$ are eigenvectors of A corresponding to eigenvalues λ_i, with $1 \leq i \leq n$. In addition, since P is invertible, all columns of P are linearly independent. Hence A has n linearly independent eigenvectors.

$(b) \Rightarrow (a)$: Now assume that A has n linearly independent eigenvectors p_i where $1 \leq i \leq n$ with eigenvalues λ_i for $1 \leq i \leq n$. Let

$$P = [p_1, p_2, \ldots, p_n] = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \ldots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \ldots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \ldots & p_{nn} \\ \uparrow & \uparrow & \uparrow & \ldots & \uparrow \\ p_1 & p_2 & p_3 & \ldots & p_n \end{bmatrix}.$$

Now $AP = [Ap_1, Ap_2, \ldots, Ap_n]$ by multiplication. But $Ap_i = \lambda_i p_i$, where $1 \leq i \leq n$. Hence

$$\begin{aligned} AP &= [Ap_1, Ap_2, \ldots, Ap_n] \\ &= [\lambda_1 p_1, \lambda_2 p_2, \ldots, \lambda_n p_n] \\ &= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \lambda_3 p_{13} & \ldots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \lambda_3 p_{23} & \ldots & \lambda_n p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \lambda_3 p_{n3} & \ldots & \lambda_n p_{nn} \end{bmatrix} \\ &= \begin{bmatrix} p_{11} & p_{12} & p_{13} & \ldots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \ldots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \ldots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix} \\ &= PD, \end{aligned}$$

where D is the diagonal matrix having eigenvalues λ_i, with $1 \leq i \leq n$ on the main diagonal. Since the columns of P are linearly independent, we have that P is invertible. Now from $AP = PD$ implies that $P^{-1}AP = D$ and hence A is diagonalisable. □

The proof of Theorem 2.4.2 provides a method for diagonalising A which we describe below:

Procedure for diagonalising a matrix Let A be an $n \times n$ matrix

Step 1 Find n linearly independent eigenvectors of A say $p_1, p_2, \ldots, p_n$ (wit eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively)

Step 2 Form the matrix P having $p_1, p_2, \ldots, p_n$ as its column vectors. And also form matrix D with main diagonal entries as $\lambda_1, \lambda_2, \ldots, \lambda_n$ and zeros elsewhere.

Step 3 Then $P^{-1}AP = D$, i.e., A is diagonalisable.

Example 2.4.3. Find a matrix that diagonalises

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

Solution. From Example 2.3.12 we have that

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

is the characteristic equation of A and found the following bases for the eigenspaces:

For $\lambda = 2:\ E_2 = \mathrm{span}(\mathbf{v}_1, \mathbf{v}_2)$ where

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

and

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

$\lambda = 1:\ E_1 = \mathrm{span}(\mathbf{v}_3)$ where

$$\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Since these three vectors are linearly independent in $\mathbb{R}^3$ they form a basis. Thus A is diagonalisable and

$$P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

diagonalises A. You may verify that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

Remark 2.4.4. There is no preferential way in which the columns of P must be presented. Since the i-th entry of the diagonal matrix $P^{-1}AP$ is an eigenvalue of the i-th column vector of P, changing the ordering of the columns of P simply changes the order of the eigenvalues in the main diagonal of $P^{-1}AP$. If we had written

$$P = \begin{bmatrix} -1 & -2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

in Example 2.4.3 we would have obtained

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Example 2.4.5. Find a matrix P that diagonalises

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

Solution. The characteristic polynomial of A is

$$\det(\lambda I - A) \;=\; \begin{vmatrix} \lambda-1 & 0 & 0 \\ -1 & \lambda-2 & 0 \\ 3 & -5 & \lambda-2 \end{vmatrix} = (\lambda-1)(\lambda-2)^2.$$

Hence the characteristic equation of A is

$$\det(\lambda I - A) = (\lambda-2)^2(\lambda-1) = \mathbf{0}.$$

The eigenvalues of A are the roots of the characteristic equation, namely $\lambda_1 = 2$ and $\lambda_2 = 1$.

The bases for the eigenspaces are :

For $\lambda = 2 :\ E_2 = \mathrm{span}(\mathbf{v}_1)$ where $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

For $\lambda = 1 :\ E_1 = \mathrm{span}(\mathbf{v}_2)$ where $\mathbf{v}_2 = \begin{bmatrix} \frac{1}{8} \\ -\frac{1}{8} \\ 1 \end{bmatrix}$.

Since A is a 3×3 matrix and there are in total only two vectors in a basis for A we have that A is not diagonalisable. So there is no matrix P that diagonalises A. □

Note 2.4.6. We have assumed that eigenvectors from different eigenspaces are linearly independent. We prove this in the following theorem.

Theorem 2.4.7. *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_v$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.*

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be the set of eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. We will suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is linearly dependent and obtain a contradiction. Thus proving that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is linearly independent. Notice that $\{\mathbf{v}_1\}$ is linearly independent, since $\mathbf{v}_1$ is nonzero (it is an eigenvector). Let r be the largest integer such that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is linearly independent. Since we are assuming that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly dependent, $1 \le r < k$, by definition of r we have that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{r+1}\}$ is linearly dependent. There exist scalars $c_1, c_2, \ldots, c_{r+1}$ not all zero such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{r+1}\mathbf{v}_{r+1} = \mathbf{0}. \tag{2.7}$$

Multiplying both sides of Equation (2.7) by A we obtain $A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{r+1}\mathbf{v}_{r+1}) = A\mathbf{0}$ which implies $c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \ldots + c_{r+1}A\mathbf{v}_{r+1} = \mathbf{0}$. But $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{r+1}$ are eigenvectors of A corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$ respectively. Thus we have

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \ldots + c_{r+1}\lambda_{r+1}\mathbf{v}_{r+1} = \mathbf{0}. \tag{2.8}$$

Multiplying Equation (2.7) by λ_{r+1} and subtracting Equation (2.8) from it, we obtain $c_1(\lambda_{r+1} - \lambda_1)\mathbf{v}_1 + c_2(\lambda_{r+1} - \lambda_2)\mathbf{v}_2 + \ldots + c_r(\lambda_{r+1} - \lambda_r)\mathbf{v}_r = \mathbf{0}$. Now from the fact that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$

is linearly independent it follows that $c_1(\lambda_{r+1}-\lambda_1) = c_2(\lambda_{r+1}-\lambda_2) = \ldots = c_r(\lambda_{r+1}-\lambda_r) = 0$; and since all eigenvalues are distinct we have $c_1 = c_2 = \ldots = c_r = 0$. Now substituting the values of $c_1 = c_2 = \ldots = c_r = 0$ into Equation (2.7) we obtain

$$c_{\mathbf{v}_{r+1}}\mathbf{v}_{r+1} = \mathbf{0}, \tag{2.9}$$

and since the eigenvector $\mathbf{v}_{r+1}$ is nonzero, it follows that

$$c_{r+1} = 0. \tag{2.10}$$

Now Equations (2.9) and (2.10) contradict the fact that $c_1, c_2, \ldots, c_r, c_{r+1}$ are not all zero; and so the result follows. □

Remark 2.4.8. Theorem 2.4.7 is a special case of the result that can be illustrated by an example: if we have 3 linearly independent eigenvectors corresponding to one eigenvalue and two linearly independent vectors corresponding to another (distinct) eigenvalue, them all five vectors will be linearly independent. We now prove

Theorem 2.4.9. *If A is an $n \times n$ matrix consisting of n distinct eigenvalues, then A is diagonalisable.*

Proof. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then by Theorem 2.4.7 we have that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Thus it follows by Theorem 2.4.2 that A is diagonalisable. □

Example 2.4.10. Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}.$$

Is A diagonalisable?

Solution. Refer to Example 2.3.6. A has three distinct eigenvalues, namely $\lambda = 4$, $\lambda = 2+\sqrt{3}$ and $\lambda = 2-\sqrt{3}$. Therefore A is diagonalisable. Moreover

$$P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2+\sqrt{3} & 0 \\ 0 & 0 & 2-\sqrt{3} \end{bmatrix}$$

for some invertible matrix P. We can obtain P using the same method that was used in Example 2.4.3. □

Example 2.4.11. Is

$$A = \begin{bmatrix} -1 & 2 & 4 & 0 \\ 0 & 3 & 1 & 7 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

diagonalisable?

Solution. By Theorem 2.3.9 we have that the eigenvalues of a triangular matrix are the entries on is main diagonal. Thus a triangular matrix whose entries on the main diagonal are all distinct is diagonalisable. Hence A is diagonalisable. □

Remark 2.4.12. It is possible that an $n \times n$ matrix is diagonalisable without having n distinct eigenvalues (see Example 2.4.3). Then Theorem 2.4.9 will not help. What is important for diagonalisability is that the dimensions of all eigenspaces (all dimensions) must add up to n, for an $n \times n$ matrix to be diagonalisable. Observe that in Example 2.4.3 there are 2 eigenspaces associated with the matrix, having dimension 2 and 1 respectively.. i.e., 2 + 1 =3 and so A is diagonalisable (3×3). However in Example 2.4.5 we have two eigenspaces with dimensions 1 each, so that 1 + 1 =2; thus A is not diagonalisable (3×3).

Definition 2.4.13. Let A be an $n \times n$ matrix and λ_0 be an eigenvalue of A. The dimension of the eigenspace of A corresponding to λ_0 is called the **geometric multiplicity** of λ_0. The number of times the expression $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the **algebraic multiplicity** of λ_0.

For example if $\det(\lambda I - A) = (\lambda - 1)^2(\lambda - 2)$, then the algebraic multiplicity of $\lambda = 1$ is 2, and that of $\lambda = 2$ is 1. On the other hand the dimension of the eigenspace corresponding to $\lambda = 1$ cannot exceed 2. This is stated in the following theorem which we will not prove.

Theorem 2.4.14. *Let A be a square matrix. Then*
(a) For every eigenvalue of A the geometric multiplicity is less than or equal the algebraic multiplicity.
(b) A is diagonalisable if and only if the geometric multiplicity equals the algebraic multiplicity.

Remark 2.4.15. There are many problems in applied mathematics where it is required to calculate high powers of a square matrix. If A is an $n \times n$ matrix and P is an invertible matrix, then

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P.$$

In general we have

$$(P^{-1}AP)^k = P^{-1}A^kP.$$

If in addition A is a diagonalisable matrix and $P^{-1}AP = D \longleftarrow$ diagonal matrix , then $(P^{-1}AP)^k = D^k$. Hence

$$A^k = PD^kP^{-1}. \tag{2.11}$$

Now in order to find A^k we need to be able to find D^k and below we sketch how that can be done. If $D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & d_n \end{bmatrix}$ then $D^k = \begin{bmatrix} {d_1}^k & 0 & \dots & 0 \\ 0 & {d_2}^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & {d_n}^k \end{bmatrix}$. Consider the example below:

Example 2.4.16. Let

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Find A^{13}.

Solution. We have shown in Example 2.4.3 that A is diagonalisable and that

$$P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

diagonalises A. In addition we showed that

$$D = P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So from Equation 2.11 we have that

$$\begin{aligned} A^{13} &= PD^{13}P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix}. \end{aligned}$$

□

2.5 Application to differential equations

Many laws of Physics, Chemistry, Biology, and Economics are described in terms of Differential Equations (DE's). These are equations involving functions and their derivatives. In this section we will show how Linear Algebra can be used to solve certain systems of DE's. Consider the following simple differential equation:

$$y' = ay. \tag{2.12}$$

This equation would have been encountered in Calculus when exponential growth/decay was discussed. This equation can be solved by "separation of variables" and the solution is of the form

$$y = ce^{ax} \quad \text{check!}$$

where c is an arbitrary constant. This solution is called a **general solution** to Equation 2.12.

If we are given, in addition, that $y(0) = 3$, then we can solve for c :

$$3 = y(0) = ce^0 = c \Rightarrow c = 3,$$

and the solution is $y = 3e^{ax}$. This solution is called a **particular solution** and $y(0) = 3$ is called **initial condition**.

In this section we want to solve DE's of the form:

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3 + \cdots + a_{1n}y_n, \\ y_2' &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3 + \cdots + a_{2n}y_n, \\ &\vdots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + a_{n3}y_3 + \cdots + a_{nn}y_n \end{aligned}$$

where $y_i = f_i(x)$, for $1 \le i \le n$. In matrix notation we get

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

i.e., $Y' = AY$. The following example will help us get an idea of how this process works.

Example 2.5.1. Consider the following system of differential equations:

$$\begin{aligned} y_1' &= 3y_1 \\ y_2' &= -2y_2 \\ y_3' &= 5y_3. \end{aligned}$$

(a) Write in matrix form.

(b) Solve the system.

(c) Find a solution that satisfies the initial conditions $y_1(0) = 1$, $y_2(0) = 4$, and $y_3(0) = -2$.

Solution. (a)

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} c_1e^{3x} \\ c_2e^{-2x} \\ c_3e^{5x} \end{bmatrix}.$$

(c)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} e^{3x} \\ 4e^{-2x} \\ -2e^{5x} \end{bmatrix}.$$

□

The system in Example 2.5.1 was easy to solve because each equation involved **one** unknown function and hence the matrix of coefficients is **diagonal**.

Question. How do we handle the case $Y' = AY$, where A is not a diagonal matrix?

The way to handle this case is to make a substitution for Y, so that we get a new system whose coefficient matrix is diagonal. We solve this new system, as in Example 2.5.1 and then we determine the solution of the original system. We will describe the procedure below:

1. Make a substitution

$$Y = \underbrace{P}_{n \times n \text{ matrix}} \quad \underbrace{U}_{\text{new unknowns}}$$

2. Differentiate to get

$$Y' = PU',$$

but

$$Y' = AY.$$

So

$$AY = PU'.$$

But

$$Y = PU.$$

So

$$\begin{aligned} APU = PU' \Rightarrow U' &= P^{-1}APU \\ &= \underbrace{(P^{-1}AP)}_{\text{needs to be diagonal}} \quad U \\ &= DU. \end{aligned}$$

Thus we choose a matrix P that diagonalises A.

Question. How do we solve a system of DE's $Y' = AY$ where the coefficient matrix A is diagonalisable?

Procedure to solve $Y' = AY$.

Step 1 Find a matrix P that diagonalises A.

Step 2 Make a substitution $Y = PU$ and $Y' = PU'$ to obtain $U' = DU$ where $D = P^{-1}AP$ is diagonal.

Step 3 Solve $U' = DU$ as in Example 2.5.1.

Step 4 Determine Y from $Y = PU$.

Example 2.5.2. Solve

$$\begin{aligned} y_1' &= y_1 + y_2 \\ y_2' &= 4y_1 - 2y_2. \end{aligned}$$

Find a solution that satisfies the initial conditions $y_1(0) = 1$, $y_2(0) = 6$.

Solution. We first write the system in matrix form as follows:

$$\underbrace{\begin{bmatrix} y_1' \\ y_2' \end{bmatrix}}_{Y'} = \underbrace{\begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{Y}.$$

We now find a matrix P that diagonalises A. The characteristic polynomial of A is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -1 \\ -4 & \lambda + 2 \end{vmatrix} = (\lambda + 3)(\lambda - 2).$$

Hence the characteristic equation of A is

$$\det(\lambda I - A) = (\lambda + 3)(\lambda - 2) = \mathbf{0}.$$

The eigenvalues of A are the roots of the characteristic equation, namely $\lambda_1 = -3$ and $\lambda_2 = 2$. We now have the following bases for the eigenspaces:

For $\lambda = -3$: $E_{-3} = \mathrm{Span}(\mathbf{v}_1)$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

and

for $\lambda = 2$: $E_2 = \mathrm{Span}(\mathbf{v}_2)$ where

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Since these two vectors are linearly independent, it follows that A is diagonalisable and

$$P = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix}$$

diagonalises A.

We now make $Y = PU$ i.e.,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

and $Y' = PU'$, i.e.,

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \end{bmatrix}$$

so that

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Hence

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{-3x} \\ c_2 e^{2x} \end{bmatrix}.$$

Now we determine Y from $Y = PU$, i.e.,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} c_1 e^{-3x} \\ c_2 e^{2x} \end{bmatrix} = \begin{bmatrix} c_1 e^{-3x} + c_2 e^{2x} \\ -4c_1 e^{-3x} + c_2 e^{2x} \end{bmatrix}.$$

Finally, using the initial conditions $y_1(0) = 1$, $y_2(0) = 6$ we obtain the solution:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -e^{-3x} + 2e^{2x} \\ 4e^{-3x} + 2e^{2x} \end{bmatrix}.$$

$\square$

Chapter 3

Linear Transformations

3.1 Linear transformations

You study functions in most of mathematics to understand the structure of sets and the relationships between sets. You have come across the notation $y = f(x)$, where the function f acts on x to produce y.

In this section of linear algebra we study functions which act on a **vector** $\mathbf{v}$ in a vector space V to produce a vector $\mathbf{w}$ in a vector space W.

Definition 3.1.1. A function $T: V \longrightarrow W$ (where V and W are vector spaces) is called a **linear transformation** if it satisfies the following two conditions:

(1.) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, for all $\mathbf{u}$, $\mathbf{v} \in V$.

(2.) $T(k\mathbf{u}) = kT(\mathbf{u})$, for all $\mathbf{u} \in V$ and scalars k.

In (1) and (2) of Definition 3.1.1, T is said to "preserve" addition and scalar multiplication, respectively.

Definition 3.1.1 is equivalent to the following

Exercise. Prove that a transformation T is linear if and only if $T(k\mathbf{u}+l\mathbf{v}) = kT(\mathbf{u}) + lT(\mathbf{v})$ for all $\mathbf{u}$, $\mathbf{v} \in V$ and all scalars k and l.
The previous exercise allows us to say that T "preserves" linear combinations.

Consider the equation:

$$T(\underbrace{\mathbf{u}+\mathbf{v}}_{\text{addition in } V}) = \underbrace{T(\mathbf{u}) + T(\mathbf{v})}_{\text{addition in } W}.$$

Remark 3.1.2. An immediate consequence of the definition is the following theorem:

Theorem 3.1.3. *Let V and W be vector spaces and $T: V \longrightarrow W$ be a linear transformation. Then*

(a) $T(\mathbf{0}) = \mathbf{0}$.

(b) $T(-\mathbf{u}) = -T(\mathbf{u})$ *for all* $\mathbf{u} \in V$.

(c) $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ *for all* $\mathbf{u}, \mathbf{v} \in V$ *and* $\mathbf{u}-\mathbf{v}$ *means* $\mathbf{u} + (-\mathbf{v})$.

Proof. Let $\mathbf{v}$ be any vector in V. Since $0\mathbf{v} = \mathbf{0}$, we have

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0},$$

since T is linear, which proves (a).
In addition $T(-\mathbf{v}) = T((-1)\mathbf{v})) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$, so part (b) is proved. Finally, $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1)\mathbf{v}$; thus

$$\begin{aligned} T(\mathbf{u} - \mathbf{v}) &= T(\mathbf{u} + (-1)\mathbf{v}) \\ &= T(\mathbf{u}) + T((-1)\mathbf{v}) \\ &= T(\mathbf{u}) - T(\mathbf{v}) \end{aligned}$$

which proves (c). $\square$

In words, part (a) of Theorem 3.1.3 states that a linear transformation maps $\mathbf{0}$ into $\mathbf{0}$. This property is useful when identifying transformations that are *not* linear. Notice that

$$T(\underbrace{\mathbf{0}}_{\text{the zero in } V}) = \underbrace{\mathbf{0}}_{\text{the zero in } W}.$$

Example 3.1.4. Consider a basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ for $\mathbb{R}^2$ where $\mathbf{v}_1 = (-2, 1)$ and $\mathbf{v}_2 = (1, 3)$, and let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation such that $T(\mathbf{v}_1) = (-1, 2, 0)$ and $T(\mathbf{v}_2) = (0, -3, 5)$. Find a formula for $T(x_1, x_2)$ and find $T(2, -3)$.

Solution. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ form a basis for $\mathbb{R}^2$, we can express any arbitrary vector of $\mathbb{R}^2$ as a linear combination of these bases vectors. Take $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$. If we write

$$(x_1, x_2) = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 = k_1(-2, 1) + k_2(1, 3), \tag{3.1}$$

then on equating corresponding components we obtain

$$\begin{aligned} -2k_1 + k_2 &= x_1 \\ k_1 + 3k_2 &= x_2, \end{aligned}$$

which yields $k_1 = \frac{-3x_1 + x_2}{7}$ and $k_2 = \frac{x_1 + 2x_2}{7}$, so that

$$(x_1, x_2) = \frac{-3x_1 + x_2}{7}\mathbf{v}_1 + \frac{x_1 + 2x_2}{7}\mathbf{v}_2.$$

Applying T on both sides of Equation (3.1) and using Definition 3.1.1 we obtain

$$\begin{aligned} T[(x_1, x_2)] &= \frac{-3x_1 + x_2}{7}T(\mathbf{v}_1) + \frac{\mathbf{v}_1 + 2x_2}{7}T(\mathbf{v}_2) \\ &= \frac{-3x_1 + x_2}{7}(-1, 2, 0) + \frac{x_1 + 2x_2}{7}(0, -3, 5) \\ &= \left(\frac{3x_1 - x_2}{7}, \frac{-9x_1 - 4x_2}{7}, \frac{5x_1 + 10x_2}{7}\right) \\ &= \frac{1}{7}(3x_1 - x_2, -9x_1 - 4x_2, 5x_1 + 10x_2). \end{aligned}$$

Finally $T(2, -3) = \dfrac{1}{7}(9, -6, -20)$.

$\square$

Example 3.1.5. Let $T:\mathbb{R} \longrightarrow \mathbb{R}$ be the map defined by $T(\mathbf{x}) = \sin \mathbf{x}$. Is T linear?

Solution. T is not a linear transformation since $\sin(\mathbf{x}+\mathbf{y}) \neq \sin \mathbf{x} + \sin \mathbf{y}$, for $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}$. □

Example 3.1.6. Let A be an $m \times n$ matrix and define a map $T:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $T(\mathbf{x}) = A\mathbf{x}$. Show that T is a linear transformation.

Solution. Notice first that $\mathbf{x}$ and $A\mathbf{x}$ are regarded as column vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. In general we will regard vectors in $\mathbb{R}^m$ as either columns or rows vectors whichever is appropriate in the context. We need to show that T satisfies the conditions of Definition 3.1.1. Observe that $T(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(k\mathbf{x}) = A(k\mathbf{x}) = (kA)\mathbf{x} = kA\mathbf{x} = kT(\mathbf{x})$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and scalar k. Hence T is a linear transformation. □

Example 3.1.6 is important in that it tells us that any $m \times n$ induces a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. We will denote this map by

$$T_A:\mathbb{R}^n \longrightarrow \mathbb{R}^m$$

where $T_A(\mathbf{x}) = A\mathbf{x}$, i.e., T_A is multiplication by A.

In other words, every $m \times n$ matrix gives rise to a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. We shall see shortly that, conversely, every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ arises from an $m \times n$ matrix.

Example 3.1.7. Is $T:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ given by $T([x_1, x_2]) = [x_2, x_1 - x_2, 2x_1 + x_2]$ a linear transformation?

Solution. Let $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ be vectors in $\mathbb{R}^2$ we need to show that $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(k\mathbf{x}) = kT(\mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and some scalar k. Notice that $\mathbf{x}+\mathbf{y} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $k\mathbf{x} = k(x_1, x_2) = (kx_1, kx_2)$ and that

$$\begin{aligned}
T(\mathbf{x}+\mathbf{y}) &= T(x_1+y_1, x_2+y_2) \\
&= (x_2+y_2, x_1+y_1-(x_2+y_2), 2(x_1+y_1)+x_2+y_2) \\
&= (x_2, x_1-x_2, 2x_1+x_2) + (y_2, y_1-y_2, 2y_1+y_2) \\
&= T(x_1, x_2) + T(y_1, y_2) \\
&= T(\mathbf{x}) + T(\mathbf{y}), \qquad \text{and} \\
T(k\mathbf{x}) &= T(kx_1, kx_2) \\
&= (kx_2, kx_1 - kx_2, 2(kx_1) + kx_2) \\
&= k(x_2, x_1 - x_2, 2x_1 + x_2) \\
&= kT(x_1, x_2) = kT(\mathbf{x}).
\end{aligned}$$

Hence T is a linear transformation.

□

Remark 3.1.8. We now know that an $m \times n$ matrix induces a linear transformation.

Question. Does every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ arise from an $m \times n$ matrix?

To answer this question consider Example 3.1.7 again. $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ defined by $T([x_1, x_2]) = [x_2, x_1 - x_2, 2x_1 + x_2]$. Writing in column vector form notation, we obtain

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= \begin{bmatrix} x_2 \\ x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix} \qquad (*) \\ &= \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

Let us see how A is obtained. Consider the standard basis vectors in $\mathbb{R}^2$. Substituting the components of these vectors into (*) we obtain

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) &= \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \\ T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}. \end{aligned}$$

These happen to be the columns of A. So

$$A = \left[\; T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \middle| \, T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \; \right].$$

We can generalise this to any linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, as is stated in the following theorem. Recall, in general, $\mathbf{e}_i$ is the vector in $\mathbb{R}^n$ that has 1 in the i-th position and zeros elsewhere. The vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard basis vectors in $\mathbb{R}^n$.

Theorem 3.1.9. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation and let A be the $m \times n$ matrix whose j-th column is $T(\mathbf{e}_j)$: i.e.,*

$$A = [\; T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n) \;].$$

Then $T(\mathbf{x}) = A\mathbf{x}$.

The matrix A is Theorem 3.1.9 is called the *standard matrix* of the linear transformation T and is defined by $[T]$. So

$$[T] = [\; T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n) \;],$$

and $T(\mathbf{x}) = [T]\,\mathbf{x}$.

Example 3.1.10. Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the projection mapping onto the $\mathbf{xy}$-plane, i.e., $T(x, y, z) = (x, y, 0)$. Show that T is linear. What is the standard matrix for T?

Solution. We need to show that $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(k\mathbf{u}) = kT(\mathbf{u})$ for all $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^3$ and k a scalar. So let $\mathbf{u} = (x, y, z)$ and $\mathbf{v} = (x', y', z')$ be vectors in $\mathbb{R}^3$. Notice first that $\mathbf{u}+\mathbf{v} = (x, y, z) + (x', y', z') = (x + x', y + y', z + z')$. Now

$$\begin{aligned} T(\mathbf{u}+\mathbf{v}) &= T(x + x', y + y', z + z') \\ &= (x + x', y + y', 0) \\ &= (x, y, 0) + (x', y', 0) \\ &= T(x, y, z) + T(x', y', z'). \end{aligned}$$

In addition

$$\begin{aligned} T(k\mathbf{u}) &= T(kx, ky, kz) \\ &= (kx, ky, 0) \\ &= k(x, y, 0) \\ &= kT(x, y, z). \end{aligned}$$

Hence T is a linear transformation. Finally the standard matrix for T is

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

□

Exercise. Find the standard matrix of the projection mapping onto the $\mathbf{yz}$-plane, i.e., $T(x, y, z) = (0, y, z)$.

Solution. The standard matrix for T is

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

□

Example 3.1.11. Show that the following mappings T are linear:

(a) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $T([x, y]) = (x + y, x)$.

(b) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}$ defined by $T([x, y, z]) = 2x - 3y + 4z$.

Solution. (a) The mapping T is a linear transformation, since for $\mathbf{x} = [x, y]$ and $\mathbf{x}' = [x', y']$ be vectors in $\mathbb{R}^2$ and k a scalar we have

$$\begin{aligned} T(\mathbf{x}+\mathbf{x}') &= T([x + x', y + y']) = (x + x' + y + y', x + x') \\ &= (x + y + x' + y', x + x') = (x + y, x) + (x' + y', x') \\ &= T([x, y]) + T([x', y']) = T(\mathbf{x}) + T(\mathbf{x}') \\ \text{and} \\ T(k\mathbf{x}) &= T(k[x, y]) = T([kx, ky]) \\ &= (kx + ky, kx) = (k(x + y), kx) \\ &= k(x + y, x) = kT([x, y]) \\ &= kT(\mathbf{x}). \end{aligned}$$

(b) Let $\mathbf{x} = [x, y, z]$ and $\mathbf{x}' = [x', y', z']$ be vectors in $\mathbb{R}^3$ and k a scalar. Notice that $\mathbf{x}+\mathbf{x}' = [x, y, z]+[x', y', z'] = [x+x', y+y', z+z']$ and $k\mathbf{x} = k[x, y, z] = [kx, ky, kz]$. Now

$$\begin{aligned}
T(\mathbf{x}+\mathbf{x}') &= T([x+x', y+y', z+z']) = 2(x+x') - 3(y+y') + 4(z+z') \\
&= 2x + 2x' - 3y - 3y' + 4z + 4z' = (2x - 3y + 4z) + (2x' - 3y' + 4z') \\
&= T([x, y, z]) + T([x', y', z']) = T(\mathbf{x}) + T(\mathbf{x}') \\
\text{and} & \\
T(k\mathbf{x}) &= T([kx, ky, kz]) \\
&= 2(kx) - 3(ky) + 4(kz) = k(2x - 3y + 4z) \\
&= kT([x, y, z]) \\
&= kT(\mathbf{x}).
\end{aligned}$$

Hence T is a linear transformation. □

Example 3.1.12. Show that the following mappings T are not linear:

(a) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $T([x, y]) = xy$.

(b) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ defined by $T([x, y]) = (x+1, 2y, x+y)$.

(c) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ defined by $T([x, y, z]) = (|x|, 0)$.

Solution. (a) This transformation does not satisfy either of the properties required for a linear transformation. To see that, let $\mathbf{x} = [x, y]$ and $\mathbf{x}' = [x', y']$ be vectors in $\mathbb{R}^2$ and k a scalar. Notice that $\mathbf{x}+\mathbf{x}' = [x, y]+[x', y'] = [x+x', y+y']$ and $k\mathbf{x} = k[x, y] = [kx, ky]$. Observe that

$$\begin{aligned}
T(\mathbf{x}+\mathbf{x}') &= T([x+x', y+y']) = (x+x')(y+y') = xy + xy' + x'y + x'y', \quad \text{and} \\
T(\mathbf{x})+T(\mathbf{x}') &= T([x, y]) + T([x', y']) = xy + x'y',
\end{aligned}$$

from where we deduce that $T(\mathbf{x}+\mathbf{x}') \neq T(\mathbf{x}) + T(\mathbf{x}')$.
Moreover,

$$\begin{aligned}
T(k\mathbf{x}) &= T([kx, ky]) = kxky = k^2xy, \quad \text{and} \\
kT(\mathbf{x}) &= kT([x, y]) = k(xy) = kxy.
\end{aligned}$$

So $T(k\mathbf{x}) \neq kT(\mathbf{x})$ and thus T is not a linear transformation.
(b) It suffices to verify whether $T(\mathbf{0}) = \mathbf{0}$, where the first $\mathbf{0} \in \mathbb{R}^2$ while the second $\mathbf{0} \in \mathbb{R}^3$. Notice that

$$T(\mathbf{0}) = T([0, 0]) = (1, 0, 0) \neq (0, 0, 0).$$

Hence T is not a linear transformation.
(c) As in part (a), this transformation also does not satisfy either of the properties required for a linear transformation. For, if $\mathbf{x} = [x, y]$ and $\mathbf{x}' = [x', y']$ are vectors in $\mathbb{R}^2$ and k is a scalar, notice that $\mathbf{x}+\mathbf{x}' = [x, y]+[x', y'] = [x+x', y+y']$ and $k\mathbf{x} = k[x, y] = [kx, ky]$. Now

$$\begin{aligned}
T(\mathbf{x}+\mathbf{x}') &= T([x+x', y+y']) = (|x+x'|, 0), \quad \text{and} \\
T(\mathbf{x})+T(\mathbf{x}') &= T([x, y]) + T([x', y']) = (|x|, 0) + (|x'|, 0) = (|x| + |x'|, 0)
\end{aligned}$$

so that $T(\mathbf{x}+\mathbf{x}') \neq T(\mathbf{x})+T(\mathbf{x}')$ in general.
In addition,

$$\begin{aligned} T(k\mathbf{x}) &= T([kx,ky]) = (|kx|,0) = (|k||x|,0,) \quad \text{and} \\ kT(\mathbf{x}) &= kT([x,y]) = k(|x|,0) = (k|x|,0). \end{aligned}$$

So $T(k\mathbf{x}) \neq kT(\mathbf{x})$ in general. Thus, T is not a linear transformation. $\square$

Example 3.1.13. Let $\mathbf{p} = p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ be a polynomial in $\mathcal{P}_n$. Define the map $T : \mathcal{P}_n \longrightarrow \mathcal{P}_{n+1}$ by $T(\mathbf{p}) = T(p(x)) = x \cdot p(x) = a_0x + a_1x^2 + \ldots + a_nx^{n+1}$. Show that T is a linear transformation.

Solution. The function T is a linear transformation, since for any scalar k and any polynomials $\mathbf{p}$ and $\mathbf{q}$ in $\mathcal{P}_n$ we have

$$\begin{aligned} T(\mathbf{p}+\mathbf{q}) &= T(p(x)+q(x)) \\ &= x(p(x)+q(x)) \\ &= xp(x)+xq(x) \\ &= T(p(x))+T(q(x)) \\ &= T(\mathbf{p})+T(\mathbf{q}) \text{ and} \\ T(k\mathbf{p}) &= T(k(p(x)) = x(kp(x)) \\ &= k(xp(x)) \\ &= T(p(x)) \\ &= kT(\mathbf{p}). \end{aligned}$$

$\square$

Example 3.1.14. Let $\mathbf{p} = p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ be a polynomial in $\mathcal{P}_n$. Define the map $T : \mathcal{P}_n \longrightarrow \mathcal{P}_{n+1}$ by $T(\mathbf{p}) = T(p(x)) = p'(x)$. Show that T is a linear transformation.

Solution. The function T is a linear transformation, since for any scalar k and any polynomials $\mathbf{p}$ and $\mathbf{q}$ in $\mathcal{P}_n$ we have

$$\begin{aligned} T(\mathbf{p}+\mathbf{q}) &= T(p(x)+q(x)) \\ &= (p(x)+q(x))' \\ &= p'(x)+q'(x) \\ &= T(p(x))+T(q(x)) \\ &= T(\mathbf{p})+T(\mathbf{q}) \text{ and} \\ T(k\mathbf{p}) &= T(k(p(x)) = (kp(x))' \\ &= k(p'(x)) \\ &= T(p(x)) \\ &= kT(\mathbf{p}). \end{aligned}$$

$\square$

Example 3.1.15. Let M be a fixed arbitrary matrix in M_{nn} and define $T : M_{nn} \longrightarrow M_{nn}$ by $T(A) = AM + MA$. Show that T is linear.

Solution. Let $A, B \in M_{nn}$ and k a scalar. We must show that $T(A+B) = T(A) + T(B)$ and $T(kA) = kT(A)$. Now

$$\begin{aligned} T(A+B) &= (A+B)M + M(A+B) = AM + BM + MA + MB \\ &= (AM + MA) + (BM + MB) \\ &= T(A) + T(B) \quad \text{and} \\ T(kA) &= (kA)M + M(kA) = k(AM) + k(MA) \\ &= k(AM + MA) = kT(A). \end{aligned}$$

Hence T is a linear transformation. □

Example 3.1.16 (Rotation operator)**.** Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the mapping that rotates every vector counterclockwise about the origin through an angle θ. Show that T is linear and find the standard matrix.

Solution. Let φ be the angle from the positive x-axis to the point $\mathbf{x} = (x, y)$ and $\mathbf{r}$ be the common length of $\mathbf{x}$ and $\mathbf{w}$ where $\mathbf{w} = (w_1, w_2) = T(\mathbf{x})$. Notice that T rotates $\mathbf{r}$ counterclockwise about the origin through an angle θ. To find equations relating $\mathbf{x}$ and $\mathbf{w} = T(\mathbf{x})$ note from trigonometry that

$$\begin{aligned} x &= \mathbf{r}\cos\varphi, \ y = \mathbf{r}\sin\varphi \quad (1) \\ w_1 &= \mathbf{r}\cos(\varphi+\theta), \ w_2 = \mathbf{r}\sin(\varphi+\theta) \quad (2) \end{aligned}$$

So (2) results into

$$\begin{aligned} w_1 &= \mathbf{r}\cos\varphi\cos\theta - \mathbf{r}\sin\varphi\sin\theta, \\ w_2 &= \mathbf{r}\sin\varphi\cos\theta + \mathbf{r}\cos\varphi\sin\theta, \end{aligned}$$

and substituting (1) into the latter equations we obtain

$$\begin{aligned} w_1 &= x\cos\theta - y\sin\theta, \quad (3) \\ w_2 &= x\sin\theta + y\cos\theta, \end{aligned}$$

and so

$$\mathbf{w} = T(\mathbf{x}) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta),$$

ie

$$\mathbf{w} = T(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

The equations in (3) are linear. Finally

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) &= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\ T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) &= \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}. \end{aligned}$$

Hence $[T] = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{A}$. □

Theorem 3.1.17. *Let V and W be vector spaces and $T : V \longrightarrow W$ be a linear transformation. Also let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be vectors in V so that $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ is linearly independent. Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent.*

Proof. Suppose that, for scalars $k_1, k_2, \ldots, k_n$, $k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n = \mathbf{0}$. Then

$$\begin{aligned} \mathbf{0} &= T(\mathbf{0}) = T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n) \\ &= k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2) + \ldots + k_nT(\mathbf{v}_n). \end{aligned}$$

Since the $T(\mathbf{v}_i)$ are linearly independent, we have that all $k_i = 0$. Thus the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. □

Exercise. Let V and W be vector spaces and $T : V \longrightarrow W$ be a linear transformation. Prove that if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly dependent in V, then $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ is also linearly dependent.

3.2 Composition of linear transformations

Recall the following

Definition 3.2.1. If $T_1 : U \longrightarrow V$ and $T_2 : V \longrightarrow W$ are linear transformations (where U, V and W are vector spaces), then the **composition of** T_1 with T_2 denoted by $T_2 \circ T_1$) is defined by $(T_2 \circ T_1) : U \longrightarrow W$, $(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u}))$ where $\mathbf{u}$ is a vector in U. ($\mathbf{u}$ must be in the domain of $T_2 \circ T_1$.)

Observe that

$$\textbf{Domain}(T_2 \circ T_1) = \{\mathbf{u} \in U \mid \mathbf{u} \in \mathrm{Domain}(T_1) \text{ and } T_1(\mathbf{u}) \in \mathrm{Domain}(T_2)\}.$$

Theorem 3.2.2. *If $T_1 : U \longrightarrow V$ and $T_2 : V \longrightarrow W$ are linear transformations (where U, V and W are vector spaces), then $T_2 \circ T_1 : U \longrightarrow W$, is also a linear transformation.*

Proof. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in U and k is a scalar, then from Definition 3.2.1 and the linearity of T_i $i = 1, 2$ follows that

$$\begin{aligned} (T_2 \circ T_1)(\mathbf{u} + \mathbf{v}) &= T_2(T_1(\mathbf{u} + \mathbf{v})) = T_2(T_1(\mathbf{u}) + T_1(\mathbf{v})), \quad \text{since } T_1 \text{ is linear} \\ &= T_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})), \quad \text{since } T_2 \text{ is linear} \\ &= (T_2 \circ T_1)(\mathbf{u}) + (T_2 \circ T_1)(\mathbf{v}), \quad \text{by Definition 3.2.1} \end{aligned}$$

and

$$\begin{aligned} (T_2 \circ T_1)(k\mathbf{u}) &= T_2(T_1(k\mathbf{u})), \quad \text{by Definition 3.2.1} \\ &= T_2(kT_1(\mathbf{u})), \quad \text{since } T_1 \text{ is linear} \\ &= kT_2(T_1(\mathbf{u})), \quad \text{since } T_2 \text{ is linear} \\ &= k(T_2 \circ T_1)(\mathbf{u}), \quad \text{by Definition 3.2.1.} \end{aligned}$$

□

3.3 Kernel and range of a linear transformations

We shall now discuss two useful notions namely that of "kernel" and that of "range" of a linear transformation, which help generalise two notions that you have encountered in connection with matrices.

Definition 3.3.1. Let $T : V \longrightarrow W$ be a linear transformation.

1. The **kernel** of T written ker(T) is the set of all elements of V that map onto the zero vector in W; i.e., ker(T) $= \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$
2. The **range** (denoted by $R(T)$) is the set of all vectors in W that is the image of some element of V; i.e., $R(T) = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V\}$. We can also write $R(T) = \{T(\mathbf{v}) \mid \mathbf{v} \in V\}$.

Theorem 3.3.2. *Let $T : V \longrightarrow W$ be a linear transformation. Then*

(a) ker*(T) is a subspace of V.*

(b) $R(T)$ is a subspace of W.

Proof. (a) Since $T(\mathbf{0}) = \mathbf{0}$ we have that $\mathbf{0} \in \ker(T)$, and so $\ker(T) \neq \varnothing$. Let $\mathbf{u}, \mathbf{v} \in \ker(T)$ and k a scalar, we need to show that $\mathbf{u}+\mathbf{v} \in \ker(T)$ and $k\mathbf{u} \in \ker(T)$. But $\mathbf{u} \in \ker(T) \iff T(\mathbf{u}) = \mathbf{0}$ and $\mathbf{v} \in \ker(T) \iff T(\mathbf{v}) = \mathbf{0}$. Now

$$\begin{aligned} T(\mathbf{u}+\mathbf{v}) &= T(\mathbf{u}) + T(\mathbf{v}), \text{ since } T \text{ is linear} \\ &= \mathbf{0} + \mathbf{0}, \text{ since } \mathbf{u}, \mathbf{v} \in \ker(T) \\ &= \mathbf{0}, \end{aligned}$$

so that $\mathbf{u}+\mathbf{v} \in \ker(T)$.
Finally

$$\begin{aligned} T(k\mathbf{u}) &= kT(\mathbf{u}), \text{ since } T \text{ is linear} \\ &= k\mathbf{0}, \text{ since } \mathbf{u} \in \ker(T) \\ &= \mathbf{0}, \end{aligned}$$

and so $k\mathbf{u} \in \ker(T)$. Thus $\ker(T)$ is a subspace of V.
(b) Since $T(\mathbf{0}) = \mathbf{0}$ we have that $\mathbf{0} \in R(T)$, and so $R(T) \neq \varnothing$. Let $\mathbf{w}_1, \mathbf{w}_2 \in R(T)$, then there exist vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $\mathbf{w}_1 = T(\mathbf{v}_1)$ and $\mathbf{w}_2 = T(\mathbf{v}_2)$. We need to show that $\mathbf{w}_1 + \mathbf{w}_2 \in R(T)$ and $k\mathbf{w}_1 \in R(T)$ for some scalar k. Now

$$\begin{aligned} \mathbf{w}_1 + \mathbf{w}_2 &= T(\mathbf{v}_1) + T(\mathbf{v}_2), \text{ since } \mathbf{w}_1, \mathbf{w}_2 \in R(T) \\ &= T(\mathbf{v}_1 + \mathbf{v}_2), \text{ since } T \text{ is linear} \end{aligned}$$

and so $\mathbf{w}_1 + \mathbf{w}_2 \in R(T)$.
Also

$$k\mathbf{w}_1 = kT(\mathbf{w}_1) = T(k\mathbf{w}_1), \text{ since } T \text{ is linear}$$

and so $k\mathbf{w}_1 \in R(T)$. Thus $R(T)$ is a subspace of W.

$\square$

Example 3.3.3. Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear operator given by the formula

$$T(x, y) = (2x - y, -8x + 4y)$$

(a) Which of the following vectors are in ker(T)? $(5, 10)?\,(3, 2)?\,(1, 1)?$
(b) Which of the following vectors are in $R(T)$? $(1, -4)?\,(5, 0)?\,(-3, 12)?$

Solution. (a) The vector $(5, 10) \in \ker(T)$ since

$$T(5, 10) = (2(5) - 10, -8(5) + 4(10)) = (0, 0);$$

the vector $(3, 2) \notin \ker(T)$ since

$$T(3, 2) = (4, -16) \neq (0, 0);$$

the vector $(1, 1) \notin \ker(T)$ since

$$T(1, 1) = (1, -4) \neq (0, 0).$$

(b) If $(1, -4) \in R(T)$, then there exists a vector (x, y) such that $T(x, y) = (1, -4)$, i.e., $T(x, y) = (2x - y, -8x + 4y) = (1, -4)$. If we equate components we obtain $2x - y = 1$ or $y = t$ and $x = \frac{1+t}{2}$. Thus T maps infinitely many vectors into $(1, -4)$.
Proceeding as above, we obtain the system of equations

$$\begin{aligned} 2x - y &= 5 \\ -8x + 4y &= 0. \end{aligned}$$

Since $2x - y = 5$ implies that $-8x + 4y = -20$, this system has no solution. Hence $(5, 0) \notin R(T)$.

If $(-3, 12) \in R(T)$, then there exists a vector (x, y) such that $T(x, y) = (-3, 12)$, i.e., $T(x, y) = (2x - y, -8x + 4y) = (-3, 12)$. If we equate components we obtain $2x - y = -3$ or $y = t$ and $x = \frac{t-3}{2}$, and so T maps infinitely many vectors into $(-3, 12)$.

□

Example 3.3.4. Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the orthogonal projection onto the xy-plane. Find ker(T) and $R(T)$.

Solution. Recall that $T(x, y, z) = (x, y, 0)$ and so

$$\begin{aligned} \ker(T) &= \{(x, y, z) \in \mathbb{R}^3 \mid T(x, y, z) = (0, 0, 0)\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, 0) = (0, 0, 0)\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid x = 0,\ y = 0\} \\ &= \{(0, 0, z) \in \mathbb{R}^3 \mid z \in \mathbb{R}\} \end{aligned}$$

and

$$\begin{aligned} R(T) &= \{T(x, y, z) \mid (x, y, z) \in \mathbb{R}^3\} \\ &= \{(x, y, 0) \mid x, y \in \mathbb{R}\} \\ &= \{(1, 0, 0)x + (0, 1, 0)y \mid x, y \in \mathbb{R}\}. \end{aligned}$$

□

Example 3.3.5. If $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is multiplication by the $m \times n$ matrix A, then Find ker(T) and $R(T)$.

Solution.

$$\begin{aligned}\ker(T_A) &= \{\mathbf{x} \in \mathbb{R}^n \,|\, T_A(\mathbf{x}) = \mathbf{0}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} = \mathbf{0}\} = \text{ nullspace}(A)\end{aligned}$$

To see that $R(T_A) =$ column space of A, we denote the rows of A by $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ and the columns of A by $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$. So

$$\begin{aligned}\text{Columnspace}(A) &= \{k_1\mathbf{c}_1 + k_2\mathbf{c}_2 + \ldots + k_n\mathbf{c}_m \,|\, k_1, k_2, \ldots, k_n \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{k} \\ \mathbf{r}_2 \cdot \mathbf{k} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{k} \end{pmatrix} \,\middle|\, \mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \in \mathbb{R}^n \right\} \\ &= \{A\mathbf{k} \,|\, \mathbf{k} \in \mathbb{R}^n\} \\ &= \{T_A(\mathbf{k}) \,|\, \mathbf{k} \in \mathbb{R}^n\} \\ &= R(T_A).\end{aligned}$$

Hence $R(T_A)$ is the column space of A. □

Example 3.3.6. Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the mapping that rotates every vector counterclockwise about the origin through an angle θ. Find ker(T) and $R(T)$.

Solution. Every vector in the xy-plane can be obtained by rotating some vector through the angle θ, so $R(T) = \mathbb{R}^2$. Moreover, the only vector that rotates to $\mathbf{0}$ is $\mathbf{0}$, so $\ker(T) = \{\mathbf{0}\}$. □

Definition 3.3.7. Let $T : V \longrightarrow W$ be a linear transformation. Then

(a) dim($\ker(T)$) is called the **nullity** of T, denoted nullity(T).

(b) dim($R(T)$) is called the **rank** of T, denoted rank(T).

As a consequence of Example 3.3.5 we have the following result.

Theorem 3.3.8. *If A is an $m \times n$ matrix and $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is multiplication by A, then*
(a) nullity(T_A) $= \text{nullity}(A)$.
(a) rank(T_A) $= \text{rank}(A)$.

Theorem 3.3.9 (Dimension theorem for linear transformations)**.** *Let V be an n-dimensional vector space and $T : V \longrightarrow W$, be a linear transformation, then*

$$\text{rank}(T) + \text{nullity}(T) = n.$$