# Chapter 4

# Linear Transformations

Linear transformations (also frequently called linear mappings) are functions between vector spaces that have particularly nice properties in relation to our two operations. They actually occur very widely in mathematics and its applications. In this chapter we shall investigate the general properties of linear transformations, in particular noting the close connection between them and matrices.

**Important note:** up till now, we have been quite relaxed about whether we write our $n$-tuples from $\mathbb{R}^n$ (or $F^n$) as rows or columns, usually preferring rows as this fits better with our work on row-operations. However, as we know, rows and columns are not conceptually "interchangeable" in a matrix setting. In what follows, we will need a matrix $A$ to act on a vector $\boldsymbol{v} \in F^n$ via $A\boldsymbol{v}$, and for this it is essential to write the vector as a column. We write all $n$-tuples as columns in this section.

**Definition 4.1** Let $V$ and $W$ be two vector spaces over the same field $F$ of scalars. Consider a mapping $T\colon V \to W$; that is, to each vector $v \in V$ the mapping associates a vector $T(v)$ in $W$. We shall call $T$ a *linear transformation* or *linear mapping* if the following two conditions hold:

(i) $T(u + v) = T(u) + T(v)$ for all $u, v \in V$, and

(ii) $T(\alpha v) = \alpha T(v)$ for all $v \in V$ and all scalars $\alpha \in F$.

**Example 4.2** *Define a function $T\colon M_{2\times 2}(F) \to M_{2\times 1}(F)$ by*

$$T\colon \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \begin{pmatrix} a \\ c \end{pmatrix}.$$

*Show that $T$ is linear.*

**Solution:** We check the conditions:

$$\begin{aligned} T\left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\right) &= T\begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1 + a_2 \\ c_1 + c_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ c_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ c_2 \end{pmatrix} = T\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + T\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \end{aligned}$$

and

$$T\left(\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = T\begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix} = \begin{pmatrix} \alpha a \\ \alpha c \end{pmatrix} = \alpha \begin{pmatrix} a \\ c \end{pmatrix} = \alpha \cdot T\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Hence $T$ is a linear transformation.

**Example 4.3** *Recall that $\mathcal{P}_n$ denotes the space of real polynomials of degree at most $n$. Define $T\colon \mathcal{P}_n \to \mathcal{P}_{n-1}$ by*

$$T\colon p \mapsto \frac{\mathrm{d}p}{\mathrm{d}x}.$$

*Show that $T$ is a linear transformation.*

**Solution:** We simply use two standard properties of differentiation. If $p$ and $q$ are polynomials (indeed, any differentiable functions) we know that

$$\frac{\mathrm{d}}{\mathrm{d}x}(p+q) = \frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\mathrm{d}q}{\mathrm{d}x} \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\alpha p) = \alpha\frac{\mathrm{d}p}{\mathrm{d}x};$$

that is,

$$T(p+q) = T(p) + T(q) \qquad \text{and} \qquad T(\alpha p) = \alpha T(p)$$

for any $p, q \in \mathcal{P}_n$ and any $\alpha \in \mathbb{R}$, as required.

**Example 4.4** *Consider the mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by*

$$T(x, y) = (x, -y).$$

*Geometrically, this reflects each point in the plane in the $x$-axis.*

*If we represent each point $(x, y)$ by the column vector (matrix) $\begin{pmatrix} x \\ y \end{pmatrix}$, then*

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

*We can write this as*

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

**Solution:** We can verify that $T$ is linear directly:

$$T\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = T\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -(y_1 + y_2) \end{pmatrix} = \begin{pmatrix} x_1 \\ -y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ -y_2 \end{pmatrix} = T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

and

$$T\left(\alpha\begin{pmatrix} x \\ y \end{pmatrix}\right) = T\begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix} = \begin{pmatrix} \alpha x \\ -\alpha y \end{pmatrix} = \alpha\begin{pmatrix} x \\ -y \end{pmatrix} = \alpha \cdot T\begin{pmatrix} x \\ y \end{pmatrix}.$$

This also follows from the basic properties of matrices: for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^2$,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}(\boldsymbol{v} + \boldsymbol{w}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\boldsymbol{v} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\boldsymbol{w}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}(\alpha\boldsymbol{v}) = \alpha\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\boldsymbol{v}.$$

**Example 4.5** *Let $A$ be an $m \times n$ matrix with entries from $\mathbb{R}$. Define a mapping $m_A : \mathbb{R}^n \to \mathbb{R}^m$ by multiplying by $A$:*

$$m_A(\boldsymbol{v}) = A\boldsymbol{v}.$$

*For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ and any $\alpha \in \mathbb{R}$,*

$$m_A(\boldsymbol{u} + \boldsymbol{v}) = A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v} = m_A(\boldsymbol{u}) + m_A(\boldsymbol{v})$$

*and*

$$m_A(\alpha\boldsymbol{v}) = A(\alpha\boldsymbol{v}) = \alpha \cdot A\boldsymbol{v} = \alpha \cdot m_A(\boldsymbol{v})$$

*Hence multiplication by $A$ is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$. (Sometimes, we shall simply call the mapping $A$, rather than $m_A$.)*

**Example 4.6** *Define $T : \mathbb{R} \to \mathbb{R}$ by*

$$T(x) = x^2.$$

*Show that $T$ is not a linear transformation.*

**Solution:** Take $x = 1$ and $x = 2$. Then $T(1) = 1$ and $T(2) = 4$, so $T(1) + T(2) = 5$, but $T(1 + 2) = T(3) = 9 \neq 5$. So $T$ does not respect addition, and hence is not linear.

The following describes the basic properties of a linear transformation. They are all direct consequences of the definition.

**Proposition 4.7** *Let $T : V \to W$ be a linear transformation between two vector spaces over the field $F$. Then*

(i) *if $v_1, v_2, \ldots, v_k \in V$ and $\alpha_1, \alpha_2, \ldots, \alpha_k$ are scalars in $F$, then*

$$T(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 T(v_1) + \cdots + \alpha_k T(v_k)$$

(ii) *$T(\mathbf{0}) = \mathbf{0}$; that is, the zero vector in $V$ is mapped to the zero vector in $W$ by $T$*

(iii) *$T(-v) = -T(v)$ for all $v \in V$*

Proof: (i) Expanding using repeated application of the two conditions of Definition 4.1, we see

$$T(\alpha_1 v_1 + \cdots + \alpha_k v_k) = T(\alpha_1 v_1) + \cdots + T(\alpha_k v_k) = \alpha_1 T(v_1) + \cdots + \alpha_k T(v_k).$$

(ii) We know that $0 \cdot \mathbf{0} = \mathbf{0}$ (by Proposition 2.12), so applying $T$ gives

$$T(\mathbf{0}) = T(0 \cdot \mathbf{0}) = 0 \cdot T(\mathbf{0}) = \mathbf{0}$$

(using Proposition 2.12(ii) again).

(iii) Proposition 2.12(iv) tells us that $(-1)v = -1v = -v$, so

$$T(-v) = T((-1)v) = (-1)T(v) = -T(v).$$

$\square$

We now have an unsurprising but important result, which tells us that the action of a linear map on the (infinitely many) elements of a vector space is completely determined by how it behaves on the (usually finite) basis of the space.

**Proposition 4.8** *Let $V$ and $W$ be vector spaces over a field $F$. Then any linear transformation $T\colon V \to W$ is uniquely determined by its effect on a basis for $V$.*

PROOF: Suppose that $\mathscr{B} = \{v_1, v_2, \dots, v_n\}$ is a basis for $V$. We shall show that if $w_i = T(v_i)$ is known for each $i$, then the effect of $T$ on any vector in $V$ is uniquely determined.

If $v \in V$, then as $\mathscr{B}$ is a basis, there exist unique scalars $\alpha_i$ such that

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

(see Theorem 3.20). Consequently, when we use Proposition 4.7(i), we see

$$\begin{aligned} T(v) &= T(\alpha_1 v_1 + \dots + \alpha_n v_n) \\ &= \alpha_1 T(v_1) + \dots + \alpha_n T(v_n) \\ &= \alpha_1 w_1 + \dots + \alpha_n w_n. \end{aligned}$$

Hence the effect of $T$ on every vector in $V$ is uniquely specified. □

**Example 4.9** *Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that*

$$T\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix} \qquad \textit{and} \qquad T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\0\\1\end{pmatrix}.$$

*What is the value*

$$T\begin{pmatrix}a\\b\end{pmatrix}?$$

**Solution:** Note that

$$\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix},$$

so by linearity

$$T\begin{pmatrix}1\\0\end{pmatrix} = T\begin{pmatrix}1\\1\end{pmatrix} - T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\\-1\end{pmatrix}.$$

We are now able to calculate the effect of $T$ on an arbitrary vector in $\mathbb{R}^2$:

$$\begin{aligned} T\begin{pmatrix}a\\b\end{pmatrix} &= T\left(a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix}\right) \\ &= aT\begin{pmatrix}1\\0\end{pmatrix} + bT\begin{pmatrix}0\\1\end{pmatrix} \\ &= a\begin{pmatrix}-1\\1\\-1\end{pmatrix} + b\begin{pmatrix}1\\0\\1\end{pmatrix} \\ &= \begin{pmatrix}b-a\\a\\b-a\end{pmatrix}. \end{aligned}$$

If we are told either a linear transformation, or its action on a basis, then we can derive one from the other.

**Example 4.10** *The* trace $Tr(A)$ *of a square* $n \times n$ *matrix* $A$ *is defined to be the sum of the entries on its main diagonal.*
*(i)Show that the trace mapping from* $M_{n\times n}(F)$ *to* $F$ *is linear.*
*(ii) Find the effect of the trace mapping* $Tr : M_{2\times 2}(F) \to F$ *on the standard basis of* $M_{2\times 2}(F)$*, and then show how the general rule can be reconstructed.*

**Solution:** (i) Let $A, B \in M_{n\times n}(F)$; say $A = [a_{ij}]$ and $B = [b_{ij}]$. Then

$$Tr(A) + Tr(B) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \sum_{i=1}^{n} (a_{ii} + b_{ii}) = Tr(A + B).$$

For any $\alpha \in F$,

$$\alpha Tr(A) = \alpha \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \alpha a_{ii} = Tr(\alpha A).$$

(ii) For $n = 2$, the trace map is $Tr : M_{2\times 2}(F) \to F$, $Tr \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d$.
Its effect on the standard basis of $M_{2\times 2}(F)$ is:

$$Tr \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1, Tr \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0, Tr \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 0, Tr \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

Conversely, given this action, we could exploit linearity to reconstruct the general rule:

$$\begin{aligned} Tr \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= T(a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}) \\ &= aT \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + bT \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + cT \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= a.1 + b.0 + c.0 + d.1 \\ &= a + d. \end{aligned}$$

## The matrix of a transformation

We saw, in Example 4.4, a case where a linear mapping could be expressed in terms of the action of an appropriate matrix. In fact, it turns out that in finite dimensional vector spaces, all linear mappings are expressible in terms of matrices — or as one textbook puts it, "linear transformations are nothing more than matrices in disguise". In what follows, we show how to associate a given linear mapping with a matrix.

Let $V$ and $W$ be finite-dimensional vector spaces over the same field $F$. Suppose that $\mathscr{B} = \{v_1, v_2, \ldots, v_n\}$ and $\mathscr{C} = \{w_1, w_2, \ldots, w_m\}$ are bases for $V$ and $W$, respectively, so that $\dim V = n$ and $\dim W = m$. (Strictly speaking we are working with *ordered bases* here, since the indexing of the basis vectors matters and reordering the vectors will change the matrix we produce.)

Now let $T \colon V \to W$ be a linear transformation. We know that $T$ is uniquely determined by the images $T(v_1)$, $T(v_2)$, ..., $T(v_n)$ of the basis vectors. These images are vectors in $W$, so we can uniquely express each one in terms of the basis $\mathscr{C}$ for $W$:

$$\begin{aligned}
T(v_1) &= \alpha_{11}w_1 + \alpha_{21}w_2 + \cdots + \alpha_{m1}w_m \\
T(v_2) &= \alpha_{12}w_1 + \alpha_{22}w_2 + \cdots + \alpha_{m2}w_m \\
\vdots \quad & \quad \vdots \qquad\qquad\qquad\qquad\quad \vdots \\
T(v_n) &= \alpha_{1n}w_1 + \alpha_{2n}w_2 + \cdots + \alpha_{mn}w_m
\end{aligned}$$

That is, the general formula is

$$T(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i \qquad \text{for } j = 1, 2, \dots, n.$$

**Definition 4.11** The $m \times n$ matrix $[\alpha_{ij}]$ whose $(i,j)$th entry is the coefficient $\alpha_{ij}$ appearing above is called the *matrix of $T$ with respect to the bases $\mathscr{B}$ and $\mathscr{C}$*.

There is no standard notation for this matrix, but we shall choose to denote it by $\mathrm{Mat}(T)$ or, when we wish to be explicit about the dependence on the bases $\mathscr{B}$ and $\mathscr{C}$, by $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(T)$.

Note that this matrix depends on the choice of bases for the two vector spaces concerned. The crucial point to remember when constructing this matrix is that the first *column* consists of those coefficients that arise when we write $T(v_1)$ in terms of the basis $\mathscr{C}$. More generally, the scalars appearing in the $j$th *column* are the coefficients that arise when we write $T(v_j)$ in terms of the basis $\mathscr{C}$.

**Example 4.12** *Consider the linear transformation $T \colon \mathbb{R}^3 \to \mathbb{R}^4$ given by*

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x+2y+z \\ y+3z \end{pmatrix}.$$

*What is the matrix of $T$ with respect to the standard bases for the vector spaces?*

**Solution:** We calculate the effect of $T$ on each of the basis vectors for $\mathbb{R}^3$ and find the coefficients when these images are written in terms of the basis for $\mathbb{R}^4$:

$$T(\mathbf{e}_1) = T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_2) = T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_3) = T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

We now write the coefficients appearing down the column of the matrix. Thus the matrix of $T$ with respect to the standard bases is

$$\operatorname{Mat}(T) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}.$$

So what does the matrix of a linear transformation do? If $T\colon V \to W$ is a linear transformation, $n = \dim V$, $m = \dim W$ and $A = \operatorname{Mat}(T)$, then $A$ gives us a linear transformation $A\colon F^n \to F^m$ (by multiplying vectors by the matrix). Moreover, $A$ has the same effect on the standard bases for $F^n$ and $F^m$ as $T$ does on the bases for $V$ and $W$ that we are considering. In particular, when we calculate the matrix $A$ of a linear transformation $T\colon F^n \to F^m$ with respect to the standard bases, then $T$ is the same as multiplication by the matrix $A$.

We can see this in the context of the above example. For the matrix

$$A = \operatorname{Mat}(T) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix},$$

we calculate

$$A\boldsymbol{v} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x+2y+z \\ y+3z \end{pmatrix} = T(\boldsymbol{v}).$$

**Example 4.13** *Recall that $\mathcal{P}_n$ denotes the space of polynomials of degree at most $n$ with real coefficients. Let $T\colon \mathcal{P}_n \to \mathcal{P}_{n-2}$ be the linear transformation given by*

$$T\colon p \mapsto \frac{\mathrm{d}^2 p}{\mathrm{d}x^2} \qquad \text{for a polynomial } p.$$

*Find the matrix of $T$ relative to some bases for $\mathcal{P}_n$ and $\mathcal{P}_{n-2}$.*

**Solution:** A natural basis for $\mathcal{P}_n$ is $\{1, x, x^2, x^3, \dots, x^n\}$ consisting of all *monomials*. A similar basis exists for $\mathcal{P}_{n-2}$. We calculate the effect of $T$ on each element of the basis for $\mathcal{P}_n$ and write it in terms of the basis for $\mathcal{P}_{n-2}$:

$$\begin{aligned} T(1) &= 0 \\ T(x) &= 0 \\ T(x^2) &= 2 \\ T(x^3) &= 0 + 6x \\ T(x^4) &= 0 + 0x + 12x^2 \\ &\vdots \\ T(x^n) &= 0 + 0x + \dots + 0x^{n-1} + n(n-1)x^{n-2} \end{aligned}$$

We now write the coefficients appearing down the columns of our matrix:

$$\mathrm{Mat}(T) = \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 6 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 12 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & n(n-1) \end{pmatrix}$$

So we can now differentiate by matrix multiplication! As an example, take $n = 3$ and the polynomial $x^3 + 2x^2 - 6x + 5$. Its first derivative is $3x^2 + 4x - 6$ and its second derivative is $6x + 4$. Forming a coefficient vector and pre-multiplying it by the matrix, we get:

$$\begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ -6 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

**Example 4.14** *Consider the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by*

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ x+y \end{pmatrix}.$$

*What is the matrix of $T$ with respect to the standard basis $\{\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ for $\mathbb{R}^2$, and the basis $\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\}$ for $\mathbb{R}^3$?*

**Solution:**

$$T(\mathbf{e}_1) = T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_2) = T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + (-2)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Thus the matrix of $T$ with respect to these bases is

$$\mathrm{Mat}(T) = \begin{pmatrix} 1 & 3 \\ 0 & -2 \\ 1 & 1 \end{pmatrix}.$$

## Rank and nullity

**Definition 4.15** Let $T\colon V \to W$ be a linear transformation.

(i) The *image* of $T$ consists of all images $T(v)$ of vectors in $V$ under $T$. It is denoted by $\operatorname{im} T$ or by $T(V)$:

$$\operatorname{im} T = T(V) = \{\, T(v) \mid v \in V \,\}.$$

(ii) The *kernel* or *nullspace* of $T$ consists of all vectors in $V$ which are mapped to the zero vector of $W$ by $T$. It is denoted by $\ker T$:

$$\ker T = \{\, v \in V \mid T(v) = \mathbf{0} \,\}.$$

**Proposition 4.16** *Let $T\colon V \to W$ be a linear transformation, where $V$ and $W$ are vector spaces over a field $F$.*

(i) *If $U$ is a subspace of $V$, then $T(U) = \{\, T(u) \mid u \in U \,\}$ is a subspace of $W$.*

(ii) *The image of $T$ is a subspace of $W$.*

(iii) *The kernel of $T$ is a subspace of $V$.*

PROOF: (i) Certainly $T(U)$ is non-empty since $U$ is non-empty. Now let $w_1, w_2 \in T(U)$. Then $w_1 = T(u_1)$ and $w_2 = T(u_2)$ for some $u_1, u_2 \in U$. Hence

$$w_1 + w_2 = T(u_1) + T(u_2) = T(u_1 + u_2) \in T(U),$$

since $U$ is a subspace of $V$ so $u_1 + u_2 \in U$. Similarly if $\alpha \in F$ and $w \in T(U)$, then $w = T(u)$ for some $u \in U$. We calculate that

$$\alpha w = \alpha T(u) = T(\alpha u) \in T(U),$$

since $U$ is a subspace of $V$ so $\alpha u \in U$. We have shown than $T(U)$ is closed under addition and scalar multiplication. We conclude that $T(U)$ is a subspace of $W$.

(ii) This follows immediately from (i) since it is the special case that $U = V$.

(iii) We know that $T(\mathbf{0}) = \mathbf{0}$ and therefore $\mathbf{0} \in \ker T$. We are at least therefore dealing with a non-empty subset of $V$. Now let $v_1, v_2 \in \ker T$. Then

$$T(v_1 + v_2) = T(v_1) + T(v_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

and we deduce that $v_1 + v_2 \in \ker T$. Now let $\alpha \in F$ and $v \in \ker T$. Then

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

and we deduce that $\alpha v \in \ker T$. Hence $\ker T$ is closed under addition and scalar multiplication, so it is a subspace of $V$. $\square$

**Example 4.17** *Define the linear map $T\colon \mathbb{R} \to \mathbb{R}^3$ by*

$$T(x) = \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix}.$$

*Find the kernel and image of $T$, and obtain bases for each.*

**Solution:** The image of $T$ is

$$\begin{aligned}\operatorname{im} T &= \{\, T(x) \mid x \in \mathbb{R} \,\}\\ &= \left\{ \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix} \middle|\, x \in \mathbb{R} \right\}\\ &= \left\{ x \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \middle|\, x \in \mathbb{R} \right\},\end{aligned}$$

the set of all scalar multiples of the vector

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Hence $\operatorname{im} T$ is the subspace spanned by the set consisting of this single non-zero vector. This set is linearly independent (a set consisting of a single non-zero vector is always linearly independent) and so we conclude that it is a basis for $\operatorname{im} T$.

The kernel of $T$ consists of those $x \in \mathbb{R}$ such that $T(x) = \mathbf{0}$. Now

$$T(x) = \mathbf{0} \qquad \Longleftrightarrow \qquad \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \Longleftrightarrow \qquad x = 0.$$

Hence $\ker T = \{0\}$, which is a zero dimensional space.

**Example 4.18** *Let $T$ be the linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^3$ given by*

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + 2z \\ 2x + y + z \\ 3x - y - 6z \end{pmatrix}.$$

*Find the image and kernel of $T$, and obtain bases for both of these.*

**Solution:** A vector in the image of $T$ has the form

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + 2z \\ 2x + y + z \\ 3x - y - 6z \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + z \begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix}.$$

Hence the set

$$\mathscr{A} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix} \right\}$$

is a spanning set for $\operatorname{im} T$. It is not clear by inspection whether or not this set is LI. To obtain an LI set which spans the same space, we can form a matrix from our three spanning vectors, and use echelon reduction:

$$A \to \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 1 & -6 \end{pmatrix} \qquad \begin{aligned} r_2 &\mapsto r_2 - r_1 \\ r_3 &\mapsto r_3 + -2r_1 \end{aligned}$$

$$\to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -4 \\ 0 & -3 & -12 \end{pmatrix} \qquad \begin{matrix} r_3 \mapsto r_3 - 3r_2 \\ r_2 \mapsto (-1)r_2 \end{matrix}$$

$$\to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

There are two non-zero vectors, which form an LI set. Hence $\mathrm{im}(T)$ has dimension 2 and is spanned by the linearly independent vectors $\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \}$, which therefore form a basis for $\mathrm{im}(T)$.

The kernel of $T$ consists of those vectors $\boldsymbol{v}$ for which $T(\boldsymbol{v}) = \mathbf{0}$; i.e.

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+2z \\ 2x+y+z \\ 3x-y-6z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This is the set of solutions to the system

$$\begin{aligned} x+y+2z &= 0 \\ 2x+y+z &= 0 \\ 3x-y-6z &= 0. \end{aligned}$$

Reducing this system of equations to reduced echelon form gives

$$\begin{aligned} x-z &= 0 \\ y+3z &= 0, \end{aligned}$$

which yields the solution $z = \alpha$, $y = -3\alpha$, $x = \alpha$ for $\alpha \in \mathbb{R}$. Thus

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \ker(T) \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}.$$

Hence $\ker(T)$ has dimension 1 and its basis consists of the single vector $\begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}$.

The above example suggests a general recipe for finding a basis of the image of a linear mapping $T : \mathbb{R}^m \to \mathbb{R}^m$.

- Take a typical element of the image, and from it derive a spanning set $\mathcal{A}$ for the image.
- If this set is visibly LI, you are done. Otherwise, proceed to find an LI set which spans the same space.
- Write the vectors of $\mathcal{A}$ as rows of a matrix, and row-reduce to echelon form.
- This will give you a set of LI vectors spanning $\mathrm{im}\, T$, i.e. a basis for $\mathrm{im}\, T$.

For the kernel,

- write the expression $T(\boldsymbol{v}) = \mathbf{0}$ as a system of equations;
- apply Gaussian elimination to obtain a solution (possibly involving parameters);
- from the resulting expression for a typical element of $\ker T$, derive a spanning set;
- check this spanning set is LI, or obtain an LI set which spans the same space, by proceeding as above.

Now that we know that the image and the kernel of $T \colon V \to W$ are subspaces of $W$ and $V$ respectively, it makes sense to talk about their dimensions. Consequently, we can make the following definition:

**Definition 4.19** Let $T \colon V \to W$ be a linear transformation.

(i) The *rank* of $T$ is the dimension of the image $\operatorname{im} T$ of $T$. We shall denote this by $\operatorname{rank} T$.

(ii) The *nullity* of $T$ is the dimension of the kernel $\ker T$ of $T$. We shall denote this by $\operatorname{null} T$.

**Comment:** Many authors use different notations or, more commonly, no specific notation for these two concepts.

We determine the rank and nullity in the important case of the action of a matrix.

**Theorem 4.20** *Let $A$ be a matrix in $M_{m\times n}(F)$. Consider the linear transformation $m_A : \boldsymbol{v} \mapsto A\boldsymbol{v}$. Then*

- $\operatorname{im}(m_A)$ *is the column-space of $A$;*
- $\ker(m_A)$ *is the solution set of $A\boldsymbol{v} = \mathbf{0}$;*
- *the rank of $m_A$ is the column-rank of $A$ (i.e. simply the rank of $A$);*
- *the nullity of $m_A$ is the dimension of the solution space of $A\boldsymbol{v} = \mathbf{0}$.*

PROOF: For (i), we show that the column-space of $A$ and the image of the linear transformation $m_A$ are the same subspace of $F^m$. Let $\boldsymbol{C}_1$, $\boldsymbol{C}_2$, ..., $\boldsymbol{C}_n$ denote the columns of $A$, so $A$ has the form

$$A = \begin{pmatrix} \boldsymbol{C}_1 & \boldsymbol{C}_2 & \dots & \boldsymbol{C}_n \end{pmatrix}.$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be the standard basis for $F^n$; that is,

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

with the 1 in the $i$th entry. Thus an arbitrary vector in $F^n$ has the form

$$\boldsymbol{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n.$$

Hence

$$\begin{aligned} A\boldsymbol{v} &= A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1\,A\mathbf{e}_1 + x_2\,A\mathbf{e}_2 + \cdots + x_n\,A\mathbf{e}_n \end{aligned}$$

and

$$A\mathbf{e}_i = \begin{pmatrix}\boldsymbol{C}_1 & \boldsymbol{C}_2 & \ldots & \boldsymbol{C}_n\end{pmatrix}\mathbf{e}_i = (\boldsymbol{C}_1 \times 0) + \cdots + (\boldsymbol{C}_i \times 1) + \cdots + (\boldsymbol{C}_n \times 0) = \boldsymbol{C}_i.$$

So

$$A\boldsymbol{v} = x_1\boldsymbol{C}_1 + x_2\boldsymbol{C}_2 + \cdots + x_n\boldsymbol{C}_n.$$

Therefore

$$\begin{aligned} \operatorname{im} A &= \{\, A\boldsymbol{v} \mid \boldsymbol{v} \in F^n \,\} \\ &= \{\, x_1\boldsymbol{C}_1 + x_2\boldsymbol{C}_2 + \cdots + x_n\boldsymbol{C}_n \mid x_1, x_2, \ldots, x_n \in F \,\} \\ &= \operatorname{Span}(\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_n). \end{aligned}$$

Hence the image of $A$ and the column-space are equal. Part (iii) follows by taking dimensions: the dimension of the image is therefore the column-rank of $A$, and we may now refer to this as simply the rank of $A$. Part (ii) is immediate from the definition, and part (iv) follows from it by taking dimensions. □

We consider the rank and nullity of our previous linear transformations.

**Example 4.21**

- *In Example 4.17,* $\operatorname{rank}(T) = 1$ *and* $\operatorname{null}(T) = 0$.
- *In Example 4.18,* $\operatorname{rank}(T) = 2$ *and* $\operatorname{null}(T) = 1$.

**Example 4.22** *Find the rank and nullity of the linear transformation from* $\mathbb{R}^2$ *to* $\mathbb{R}^3$ *given by*

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix}.$$

**Solution:** A vector in the image of $T$ has the form

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Hence the set

$$\mathscr{A} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

is a spanning set for $\operatorname{im} T$. We note that the matrix form of $T$, with respect to the standard bases, is:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

so the image is the column space of $\operatorname{Mat}(T)$. Moreover, this set is also linearly independent, for if

$$\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

then

$$\begin{pmatrix} \alpha \\ \alpha+\beta \\ \alpha-\beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and we deduce $\alpha = \beta = 0$. (In fact, to check linear independence for just two vectors, it suffices to verify that neither is a scalar multiple of the other.) Hence $\mathscr{A}$ is a basis for $\operatorname{im} T$ and so

$$\operatorname{rank} T = \dim \operatorname{im} T = 2.$$

The kernel of $T$ consists of those vectors $\boldsymbol{v}$ for which $T(\boldsymbol{v}) = \mathbf{0}$; that is, we solve

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and we deduce that $x = y = 0$. Hence

$$\ker T = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \{\mathbf{0}\},$$

which is a zero-dimensional space. Therefore $\operatorname{null} T = 0$.

Notice in Example 4.17 that $T \colon \mathbb{R} \to \mathbb{R}^3$ satisfies

$$\operatorname{rank} T + \operatorname{null} T = 1 + 0 = 1 = \dim \mathbb{R},$$

in Example 4.18 that $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ satisfies

$$\operatorname{rank} T + \operatorname{null} T = 2 + 1 = 3 = \dim \mathbb{R}^3$$

while in Example 4.22 our transformation $T \colon \mathbb{R}^2 \to \mathbb{R}^3$ satisfies

$$\operatorname{rank} T + \operatorname{null} T = 2 + 0 = 2 = \dim \mathbb{R}^2.$$

These are actually examples of a general phenomenon.

**Theorem 4.23 (Rank-Nullity Theorem)** *Let $T \colon V \to W$ be a linear transformation between finite-dimensional vector spaces over a field $F$. Then*

$$\operatorname{rank} T + \operatorname{null} T = \dim V.$$

PROOF: Let $\mathscr{B} = \{v_1, v_2, \dots, v_n\}$ be a basis for $\ker T$ (so that $n = \operatorname{null} T$). Extend this to a basis $\mathscr{C} = \{v_1, v_2, \dots, v_n, v_{n+1}, \dots, v_{n+k}\}$ for $V$ (so that $\dim V = n + k$). We will show that $\{T(v_{n+1}), \dots, T(v_{n+k})\}$ is a basis for $\operatorname{im} T$.

*Spanning:* Let $w \in \operatorname{im} T$. So $w = T(v)$ for some $v \in V$. We can write $v$ as a linear combination of the vectors in the basis $\mathscr{C}$, say

$$v = \alpha_1 v_1 + \dots + \alpha_{n+k} v_{n+k}$$

for some scalars $\alpha_1, \dots, \alpha_{n+k} \in F$. Then, applying $T$ and using linearity,

$$\begin{aligned} w = T(v) &= T(\alpha_1 v_1 + \dots + \alpha_{n+k} v_{n+k}) \\ &= \alpha_1 T(v_1) + \dots + \alpha_{n+k} T(v_{n+k}) \\ &= \alpha_{n+1} T(v_{n+1}) + \dots + \alpha_{n+k} T(v_{n+k}), \end{aligned}$$

since $T(v_1) = \dots = T(v_n) = \mathbf{0}$ as $v_1, \dots, v_n \in \ker T$. This shows that the set $\mathscr{D} = \{T(v_{n+1}), \dots, T(v_{n+k})\}$ spans $\operatorname{im} T$.

*Linear independence:* Suppose that

$$\beta_1 T(v_{n+1}) + \dots + \beta_k T(v_{n+k}) = \mathbf{0};$$

that is

$$T(\beta_1 v_{n+1} + \dots + \beta_k v_{n+k}) = \mathbf{0}.$$

Hence the vector

$$\beta_1 v_{n+1} + \dots + \beta_k v_{n+k}$$

belongs to the kernel of $T$. We know that $\mathscr{B}$ is a basis for $\ker T$ and hence

$$\beta_1 v_{n+1} + \dots + \beta_k v_{n+k} = \gamma_1 v_1 + \dots + \gamma_n v_n$$

for some $\gamma_1, \dots, \gamma_n \in F$. Rearranging we obtain the equation

$$(-\gamma_1) v_1 + \dots + (-\gamma_n) v_n + \beta_1 v_{n+1} + \dots + \beta_k v_{n+k} = \mathbf{0}.$$

This equation involves the vectors in the basis $\mathscr{C}$ for $V$. Therefore, since $\mathscr{C}$ is linearly independent, we conclude that all the coefficients involved are zero. In particular,

$$\beta_1 = \beta_2 = \dots = \beta_k = 0,$$

which is what we needed to deduce that $\mathscr{D}$ is linearly independent.

Hence $\mathscr{D} = \{T(v_{n+1}), \dots, T(v_{n+k})\}$ is a basis for $\operatorname{im} T$ and so

$$\operatorname{rank} T = \dim \operatorname{im} T = k = (n + k) - n = \dim V - \operatorname{null} T.$$

Thus

$$\operatorname{rank} T + \operatorname{null} T = \dim V.$$

$\square$

The advantage of this theorem is that it tells us that once we know either the rank or the nullity of a linear transformation, then we can (pretty much immediately) deduce the other.

During the previous chapters, we have met the definition of row-space, column-space, row-rank and column-rank of a matrix. We have seen that row-rank and column-rank are in fact equal, and that we may talk simply of the rank of a matrix. Moreover, we have seen that for a given matrix $A$, the non-zero rows in any echelon form $E$ of $A$ constitute a basis for the row-space of $A$, and hence that the rank of $A$ equals the number of non-zero rows in $E$.

If, for any linear transformation, we are interested in obtaining its rank and nullity directly, we therefore have the following recipe to do so:

- Let $T\colon V \to W$ be a linear transformation.
- Determine $A = \mathrm{Mat}(T)$, with respect to some bases.
- Apply row operations directly to $A$, to produce a matrix $E$ in echelon form that is row-equivalent to $A$.
- The rank of $T$ is equal to the rank of $A$ and this equals the number of non-zero rows in $E$.
- Determine the nullity of $T$ via the Rank-Nullity Theorem (that is, Theorem 4.23).

This yields the numbers for rank and nullity fairly quickly, but we would still need to do additional work in order to obtain bases for image and kernel.

**Example 4.24** *Consider the linear transformation $m_A : \mathbb{R}^4 \to \mathbb{R}^3$ given by $m_A(\boldsymbol{v}) = A\boldsymbol{v}$ where*

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

*(a) Find its rank and nullity.*

*(b) Obtain a basis for the image and kernel.*

**Solution:** (a) The matrix of this linear transformation with respect to the standard basis of the vector spaces is the original matrix $A$. We apply row operations to reduce it to echelon form:

$$A \to \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & 0 & 5 & -3 \\ 0 & 0 & 5 & -3 \end{pmatrix} \qquad \begin{array}{l} r_2 \mapsto r_2 - 2r_1 \\ r_3 \mapsto r_3 + r_1 \end{array}$$

$$\to \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & 0 & 5 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad r_3 \mapsto r_3 - r_2$$

This last matrix is in echelon form and has two non-zero rows. We conclude that our original matrix has rank 2. Then by the Rank-Nullity Theorem, $\operatorname{rank} A + \operatorname{null} A = \dim \mathbb{R}^4 = 4$. Hence the nullity of $A$ equals 2.

(b) Next, we are asked to obtain a basis for the image and kernel. We will begin with the kernel, as we can utilise the row-reduction of $A$ from part (i).

To find the kernel, we solve $A\boldsymbol{v} = \mathbf{0}$; that is,

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 4 & 3 & 5 \\ -1 & -2 & 6 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a system of linear equations which we solve by applying Gaussian elimination to:

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & 0 \\ 2 & 4 & 3 & 5 & 0 \\ -1 & -2 & 6 & -7 & 0 \end{array}\right)$$

Since the Gaussian elimination is performed by precisely the row operations we used above in part (a), we can immediately reduce to

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & 0 \\ 0 & 0 & 5 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Hence we solve

$$x + 2y - z + 4t = 0 \quad \text{and} \quad 5z - 3t = 0.$$

The solutions to these equations are parametrized by $y$ and $t$ and are given by $z = \frac{3}{5}t$ and $x = -2y - \frac{17}{5}t$. Therefore

$$\ker A = \left\{ \begin{pmatrix} -2y - \frac{17}{5}t \\ y \\ \frac{3}{5}t \\ t \end{pmatrix} \middle| \, y, t \in \mathbb{R} \right\}.$$

An arbitrary vector in the kernel has the form

$$\begin{pmatrix} -2y - \frac{17}{5}t \\ y \\ \frac{3}{5}t \\ t \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -17/5 \\ 0 \\ 3/5 \\ 1 \end{pmatrix}.$$

We conclude that every vector in the kernel is a linear combination of the two vectors appearing on the right-hand side here and they are clearly linearly independent since neither is a multiple of the other. Hence

$$\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -17/5 \\ 0 \\ 3/5 \\ 1 \end{pmatrix} \right\}$$

is a basis for the kernel of $A$.

The image of $m_A$ is the column-space of $A$, and a basis of this will consist of a set of linearly independent vectors which span the column-space. Putting the columns of $A$ into a matrix and applying row-reduction (this is clearly equivalent to applying column-operations to $A$ itself), we obtain:

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & 3 & 6 \\ 4 & 5 & -7 \end{pmatrix} \qquad \begin{array}{l} r_2 \mapsto r_2 - 2r_1 \\ r_3 \mapsto r_3 + r_1 \\ r_4 \mapsto r_4 - 4r_1 \end{array}$$

$$\to \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 5 & 5 \\ 0 & -3 & -3 \end{pmatrix} \qquad \begin{array}{l} r_2 \leftrightarrow r_4 \\ r_2 \mapsto -\frac{1}{3}r_2 \\ r_3 \mapsto \frac{1}{5}r_3 - r_2 \end{array}$$

$$\to \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So a basis for the image is given by $\{\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\}$; these are clearly LI as neither is a multiple of the other.

We observe that, having proved the Rank-Nullity theorem, we now have a nice way of understanding our result, from very early in the course, about existence of solutions to $A\boldsymbol{x} = \mathbf{0}$ — and a new elegant proof.

**Theorem 4.25** *Let $A \in M_{n\times n}(F)$. There exists a non-zero column vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \mathbf{0}$ if and only if $A$ is non-invertible.*

Proof: Recall $m_A : F^n \to F^n$ where $m_A(\boldsymbol{x}) = A\boldsymbol{x}$. There exists a non-zero $\boldsymbol{x}$ such that $A\boldsymbol{x} = \mathbf{0}$
$\Leftrightarrow$ $\ker(m_A)$ contains a non-zero vector
$\Leftrightarrow$ $\mathrm{null}(m_A) > 0$
$\Leftrightarrow$ $\mathrm{rank}(m_A) < n$
$\Leftrightarrow$ $A$ is non-invertible. □

In fact,the following result describes the main features of the situation:

**Theorem 4.26** *Let $A \in M_{n\times n}(F)$. For the homogeneous system $A\boldsymbol{x} = \mathbf{0}$,*

- *the solution space has dimension $n - \mathrm{rank}(A)$;*
- *there is a non-trivial solution if and only if $\mathrm{rank}(A) < n$.*

Proof: Since the solution space is the kernel of $m_A$, its dimension is $\mathrm{null}(m_A)$ and the result follows by the Rank-Nullity Theorem. There will be a non-trivial solution to the system precisely if the dimension of the solution space is positive, i.e. if $\mathrm{rank}(m_A) = \mathrm{rank}(A)$ is less than maximum. □