CHAPTER 2

Linear Algebra

1. Introduction

DEFINITION 1.1. A **vector space** is a set V equipped with two operations, **addition** and **scalar multiplication**, satisfying the following 8 axioms:

$$
\begin{array}{lll}
\textbf{A1:} & \forall\, v, w \in V & v + w = w + v \\
 & & \text{i.e. addition is commutative} \\
\textbf{A2:} & \forall\, v, w, x \in V & (v + w) + x = v + (w + x) \\
 & & \text{i.e. addition is associative} \\
\textbf{A3:} & \exists \text{ zero vector } 0_V \in V \text{ s.t.} & \forall\, v \in V,\ v + 0_V = 0_V + v = v \\
\textbf{A4:} & \forall\, v \in V,\ \exists\, -v \in V \text{ s.t.} & v + (-v) = (-v) + v = 0_V \\
 & & \text{i.e. additive inverse exists } \forall\, v \\
\textbf{S1:} & \forall\, v, w \in V,\ \forall\, \lambda \in \mathbb{R} & \lambda(v + w) = \lambda v + \lambda w \\
\textbf{S2:} & \forall\, v \in V,\ \forall\, \lambda, \mu \in \mathbb{R} & (\lambda + \mu)v = \lambda v + \mu v \\
\textbf{S3:} & \forall\, v \in V,\ \forall\, \lambda, \mu \in \mathbb{R} & (\lambda\mu)v = \lambda(\mu v) \\
\textbf{S4:} & \forall\, v \in V & 1 \cdot v = v \\
 & & \;\nwarrow \text{ the real number } 1
\end{array}
$$

An element of a vector space is called a **vector**. Real numbers are also referred to as **scalars**. The first 4 axioms are called the **addition axioms** and the last 4 axioms are called the **scalar multiplication axioms**.

EXAMPLE 1.2. A common example of a vector space is $\mathbb{R}^n$. This is the vector space of all ordered sets of n numbers (hence a vector in $\mathbb{R}^n$ is distinguished not only by its elements but also by the order in which they appear). Such a vector may be written as either

$$x = (x_1, \ldots, x_n) \quad \text{(this is called a } \textbf{row vector}\text{)}$$

or as

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{(this is called a } \textbf{column vector}\text{).}$$

and we may use either notation, whichever is more convenient or appropriate for the situation.

The numbers $x_1,\ x_2, \ldots, x_n$ are called the **components** of the vector. The number n is called the **dimension** of the vector (we will define the dimension of a general vector space later, and in this special case it will turn out to be the number n).

Two vectors are **equal** if and only if corresponding components are equal, so e.g., $(x_1, x_2) = (y_1, y_2)$ if and only if $x_1 = y_1$ and $x_2 = y_2$.

The **zero vector** in $\mathbb{R}^n$ is the vector with all components equal to 0,

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

We define addition and scalar multiplication in $\mathbb{R}^n$ as follows.

The **sum** of two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ is

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

(hence the components of the sum vector are obtained by adding each component of the first vector to the corresponding component of the second vector).

For a vector $x = (x_1, x_2, \dots, x_n)$ and a scalar $\alpha \in \mathbb{R}$ we define

$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$

(hence, to multiply a vector by a real number we multiply each component by that real number). This is the motivation for calling the real number α a **scalar**: it "scales" the vector.

Note: If two vectors do not have the same dimension then we cannot do any operations with them. For example, the sum $(1, 2) + (2, 5, 6, 7)$ makes no sense. ■

EXAMPLE 1.3. Let $\mathcal{P}$ be the set of all polynomials in the variable (or indeterminate) t, so a typical element of $\mathcal{P}$ is of the form

$$p = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$$

where n can be any natural number. The real numbers $a_0, a_1, \dots, a_n$ are called **coefficients**. a_0 is the constant coefficient, a_2 is the coefficient of t^2, and so on. Then $\mathcal{P}$ is a vector space and the zero vector is the polynomial all of whose coefficients are zero.

We define addition in $\mathcal{P}$ by adding corresponding coefficients, so, for example, if

$$p = 3 - 2t + 4t^2 + 2t^3$$

and

$$q = -5 + 2t^2 + 5t^3 - 7t^4,$$

then

$$\begin{aligned} p + q &= (3 - 5) + (2 + 0)t + (4 + 2)t^2 + (2 + 5)t^3 + (0 - 7)t^4 \\ &= -2 + 2t + 6t^2 + 7t^3 - 7t^4. \end{aligned}$$

We define scalar multiplication in $\mathcal{P}$ by multiplying each coefficient by the scalar, so, for example, if p is as above and $\alpha \in \mathbb{R}$, then

$$\alpha p = 3\alpha - 2\alpha t + 4\alpha t^2 + 2\alpha t^3.$$

Thus for $\alpha = 5$, say, we have

$$5p = 15 - 10t + 20t^2 + 10t^3.$$

■

DEFINITION 1.4. We say a subset W of V is a **subspace** of V if

(1) $$\text{for all } x, y \in W \quad x + y \in W$$

(2) $$\text{for all } x \in V, \text{ for all } \alpha \in \mathbb{R} \quad \alpha x \in W$$

Thus a vector subspace is a subset *closed* under the two operations of addition and scalar multiplication.

DEFINITION 1.5. If $v_1, \dots, v_n$ are vectors and $\lambda_1, \dots, \lambda_n$ are scalars then the vector $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ is called a **linear combination** of $v_1, \dots, v_n$.

DEFINITION 1.6. Let $v_1, \dots, v_n$ be vectors and $\lambda_1, \dots, \lambda_n$ be scalars. We say that the vectors $v_1, \dots, v_n$ are **linearly independent** if the only way to make the linear combination $\lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0}$ is to choose $\lambda_1 = 0, \dots, \lambda_n = 0$, i.e., the only linear combination of $v_1, \dots, v_n$ which gives $\mathbf{0}$ is the *trivial* linear combination where each $\lambda_i = 0$.

REMARK 1.7. If $v_1, \dots, v_n$ are linearly independent vectors, then none of the v_i can be written as a linear combination of the others. For, if this were possible, say (for example)

$$v_1 = \alpha_2 v_2 + \cdots + \alpha_n v_n,$$

then we could bring all the terms to one side to get

$$v_1 - (\alpha_2 v_2 + \cdots + \alpha_n v_n) = \mathbf{0}$$

which is a non-trivial linear combination of $v_1, \dots, v_n$ (certainly the coefficient of v_1 is nonzero), and this violates the definition of linearly independent.

DEFINITION 1.8. We say that the vectors $v_1, \dots, v_n$ **span** the vector space V if every vector in V can be written as a linear combination of $v_1, \dots, v_n$, i.e.,

$$\text{for all } v \in V, \qquad v = \lambda_1 v_1 + \cdots + \lambda_n v_n \text{ for some } \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

DEFINITION 1.9. We say the vectors $v_1, \dots, v_n$ are a **basis** for the vector space V if these *two* conditions hold:

$$v_1, \dots, v_n \text{ are linearly independent;} \tag{3}$$

$$v_1, \dots, v_n \text{ span } V. \tag{4}$$

A vector space may have many different bases.

REMARK 1.10. The usual basis we take for $\mathbb{R}^n$ is $\{e_1, \dots, e_n\}$, where

$$e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$$

with 1 in the i^{th} position, 0 in every other position. This basis $\{e_1, \dots, e_n\}$ is called the **standard basis** of $\mathbb{R}^n$.

Let us check that the two defining properties 3 and 4 in Definition 1.9 are true for $\{e_1, \dots, e_n\}$, i.e., it's a basis for $\mathbb{R}^n$. We use column vector notation for convenience.

For property 3 (linear independence), suppose $\lambda_1 e_1 + \cdots + \lambda_n e_n = \mathbf{0}$, i.e.,

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \lambda_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \mathbf{0}.$$

Then

$$\begin{pmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_3 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Since two vectors are equal if and only if corresponding components are equal, we see $\lambda_1 = 0, \lambda_2 = 0, \dots \lambda_n = 0$, i.e., $\{e_1, \dots, e_n\}$ is a linearly independent set.

For property 4 (spanning $\mathbb{R}^n$), let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Then

$$\begin{aligned} x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} x_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x_3 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n \end{aligned}$$

i.e., any vector $x \in \mathbb{R}^n$ can be written as a linear combination of $e_1, \dots, e_n$.

For example, in $\mathbb{R}^3$ a basis is $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

REMARK 1.11. Axiom 3 in Definition 1.9 above is a *smallness* condition — if we choose too many vectors they won't be linearly independent. Axiom 4 above is a *bigness* condition — if we choose too few vectors they won't span V. Thus a basis is a minimal set of vectors in terms of which we can write any vector in V.

2. The scalar product

DEFINITION 2.1. The **scalar product** of two vectors x and y in $\mathbb{R}^n$ is defined as follows

$$x \cdot y = (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Note that the scalar product is not a vector any longer but a real number (a scalar). This product is also called the **dot product**.

THEOREM 2.2. *If x, y and z are n-vectors and α is a scalar then*

$$\begin{array}{ll} x \cdot x > 0 \quad \textit{if } x \neq \mathbf{0} & \textit{(}\text{Positive definite}\textit{);} \\ x \cdot x = 0 \quad \textit{if } x = \mathbf{0}; & \\ x \cdot y = y \cdot x & \textit{(}\text{Symmetry}\textit{);} \\ x \cdot (y+z) = x \cdot y + x \cdot z & \textit{(}\text{Distributive Law}\textit{);} \\ (\alpha x) \cdot y = x \cdot (\alpha y) = \alpha(x \cdot y). & \end{array}$$

DEFINITION 2.3. Two vectors are called **orthogonal** (or **perpendicular**) if $x \cdot y = 0$.

EXAMPLE 2.4. The vectors $(-2, 3)$ and $(3, 2)$ are perpendicular, since $(-2, 3) \cdot (3, 2) = (-2)(3) + (3)(2) = -6 + 6 = 0$. Also, by drawing these two vectors we see that they form an angle of $90°$. ■

EXERCISE 2.5. Check which of these pairs of vectors are orthogonal. (a) $(1, 2)$ and $(-2, 1)$

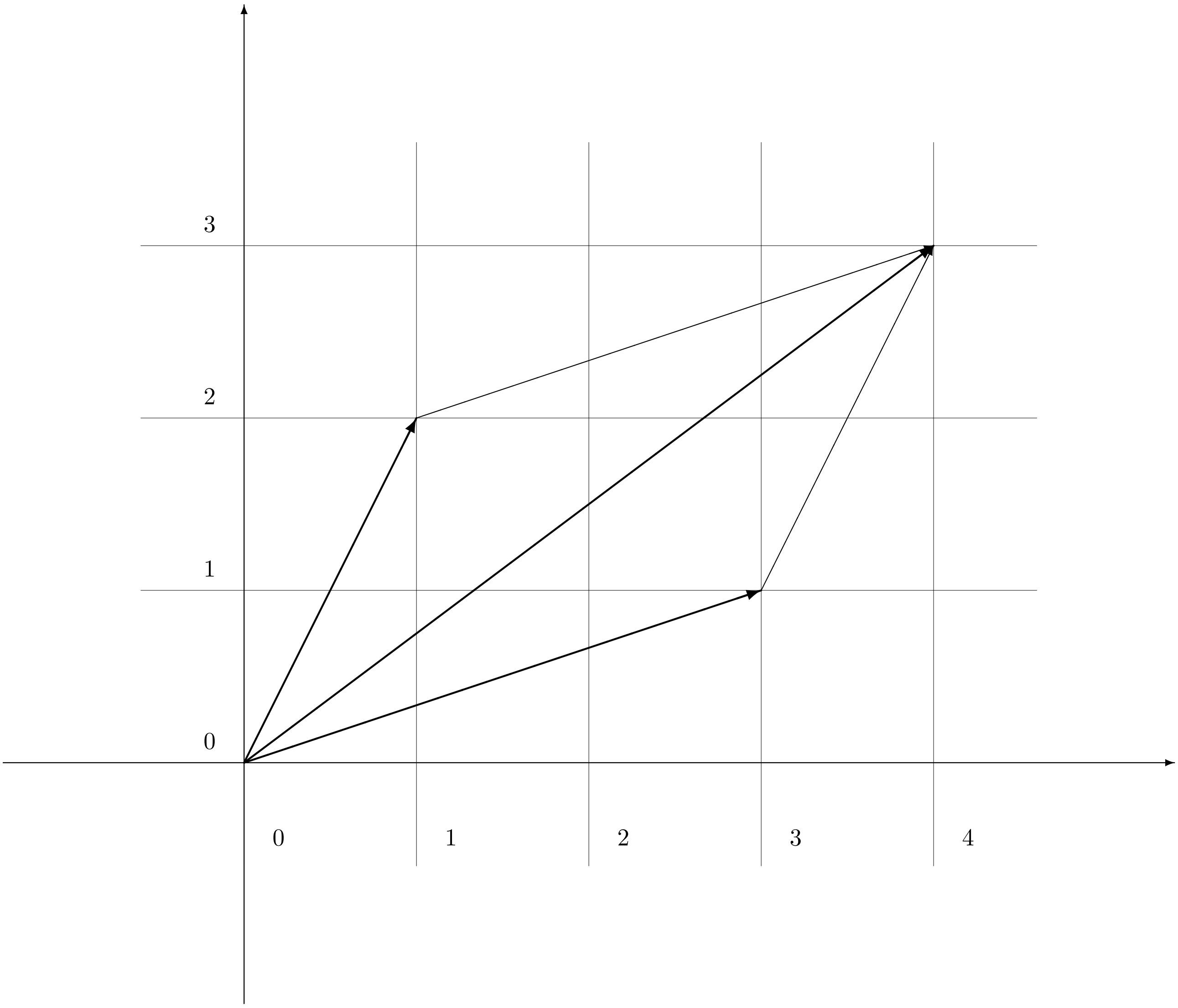
(b) $(a, -b, 1)$ and $(b, a, 0)$

(c) $(1, -1, 1)$ and $(-1, 1, -1)$.

3. Geometric interpretation of vectors

There are some properties possessed by the vector space $\mathbb{R}^n$ that are not possessed by all vector spaces. We describe these now.

By the vector $(1, 2)$ we may denote the point in the plane obtained by moving 1 unit to the right and 2 units up from the origin. We define the **geometric vector** $(1, 2) \in \mathbb{R}^2$ as the directed line segment (or arrow) joining the origin $(0, 0)$ to the point $(1, 2)$. Similarly in $\mathbb{R}^3$, the vector (a, b, c) may be viewed as the directed line segment joining the origin $(0, 0, 0)$ to the point (a, b, c) and so on for other $\mathbb{R}^n$. In the diagram below we show the geometric vectors $(1, 2)$ and $(3, 1)$, as well as their sum, $(4, 3)$.



The *tail* of a vector is where it starts from (usually the origin), the *head* is where it goes to (the point of the arrow). Notice that, geometrically, the sum vector may be obtained by *the Parallelogram Law* (i.e., complete the parallelogram two sides of which are given by the initial two vectors; then the diagonal from the origin is the sum vector) or *the Head-to-Tail Rule* (i.e., move the tail of one vector from the origin to the head of the other while keeping the direction of the moving vector fixed; then the new head of the moved vector is the head of the sum vector).

4. Dimension

DEFINITION 4.1. A vector space V is **finite-dimensional** if there exists a finite set of vectors in V which forms a basis for V.

EXAMPLE 4.2. $\mathbb{R}^n$ is finite-dimensional since $\{e_1, \ldots, e_n\}$ is a basis by Example 1.10. ■

REMARK 4.3. Not all vector spaces are finite-dimensional.

EXAMPLE 4.4. $V = \mathcal{C}(\mathbb{R})$, set of continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Define addition of functions by

$$\underbrace{(f+g)}_{\text{sum}}(x) = f(x) + g(x) \quad \forall\, x \in \mathbb{R} \quad \text{(i.e., add values)}$$

and scalar mult. by

$$\underbrace{(\lambda f)}_{\substack{\text{scalar}\\ \text{mult.}}}(x) = \lambda f(x) \quad \forall\, x \in \mathbb{R}.$$

All 8 vector space axioms hold.

The zero vector is the zero function (i.e., $g(x) = 0\ \forall\, x \in \mathbb{R}$).

The additive inverse of a function f is $-f$, where

$$(-f)(x) := -f(x) \ \forall\, x \in \mathbb{R}.$$

$\mathcal{C}(\mathbb{R})$ is *not* finite-dimensional.

Vector spaces of functions are usually *not* finite-dimensional. ■

EXAMPLE 4.5. Let $V = \mathcal{D}(\mathbb{R})$, the set of all differentiable functions $f : \mathbb{R} \longrightarrow \mathbb{R}$, with the same + and SM as in Example 4.4.

Then $\mathcal{D}(\mathbb{R})$ is a real vector space. Note $\mathcal{D}(\mathbb{R}) \subset \mathcal{C}(\mathbb{R})$, since f differentiable $\Rightarrow f$ continuous. The converse is false, e.g., absolute value function $|x|$ is continuous but not differentiable at 0. So $\mathcal{D}(\mathbb{R}) \subsetneqq \mathcal{C}(\mathbb{R})$. Thus $\mathcal{D}(\mathbb{R})$ is a proper subspace of $\mathcal{C}(\mathbb{R})$. ■

THEOREM 4.6. *Any two bases of a finite-dimensional vector space V have the same number of elements.*

DEFINITION 4.7. The number of elements in a basis for a finite-dimensional vector space V is called the **dimension** of V, written $\dim V$. This number is well-defined by Theorem 4.6.

EXAMPLE 4.8. By Example 1.10 there are n basis vectors $e_1, \ldots, e_n$ in a basis for $\mathbb{R}^n$. Thus by Definition 4.7 we have that $\dim(\mathbb{R}^n) = n$, which agrees with Example 1.2. ■

EXERCISE 4.9. Show that the set

$$M_{m,n}(\mathbb{R}) := \{\text{all } m \times n \text{ matrices with real entries}\}$$

is a vector space with addition and scalar multiplication defined as above. Also, show that the subset $D_{n,n}(\mathbb{R}) \subset M_{n,n}(\mathbb{R})$ defined by

$$D_{n,n}(\mathbb{R}) := \{\text{all diagonal } n \times n \text{ matrices with real entries}\}$$

is a vector subspace of $M_{n,n}(\mathbb{R})$ with the same addition and scalar multiplication.

Show that $M_{m,n}(\mathbb{R})$ has dimension mn and that $D_{m,n}(\mathbb{R})$ has dimension n. [*Hint:* to get a basis of $M_{m,n}(\mathbb{R})$, consider the "elementary" matrices E_{ij} which have a 1 in the (i, j) position and 0 elsewhere.]

Note that if $m = n$ we often just write $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ for the space of all real $n \times n$ (square) matrices.

5. Linear maps

(or **linear operators** or **linear transformations**)

DEFINITION 5.1. Let V, W be vector spaces over $\mathbb{R}$, and $f : V \longrightarrow W$ a function. We say f is a **linear map** if:

$$\begin{array}{llcll} (i) & f(v_1 + v_2) & = & f(v_1) + f(v_2), & \forall\, v_1, v_2 \in V; \\ (ii) & f(\lambda v) & = & \lambda f(v), & \forall\, \lambda \in \mathbb{R},\ \forall\, v \in V. \end{array}$$

i.e., f preserves the two basic vector space operations (+ and SM).

EXAMPLE 5.2. The map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by $f(x) = 3x$ is a linear map since

$$f(x + y) = 3(x + y) = 3x + 3y = f(x) + f(y)$$

and

$$f(\alpha x) = 3(\alpha x) = (3\alpha)x = \alpha(3x) = \alpha f(x)$$

i.e., f satisfies (i) and (ii) in Definition 5.1 above. This f is an expansion by a factor of 3. ■

EXAMPLE 5.3. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ by

$$f(x, y) = (6x + y, x + y, 7x - 3y).$$

Write $A = \begin{pmatrix} 6 & 1 \\ 1 & 1 \\ 7 & -3 \end{pmatrix}$. If we write vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ as *column vectors* we see that

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6x + y \\ x + y \\ 7x - 3y \end{pmatrix} = \begin{pmatrix} 6 & 1 \\ 1 & 1 \\ 7 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}.$$

Writing f this way shows f is a linear map because

$$A(v_1 + v_2) = Av_1 + Av_2$$

and

$$A(\lambda v) = \lambda A v.$$

(properties of matrix multiplication.) ■

EXAMPLE 5.4. Let $f : M_n\mathbb{R} \longrightarrow \mathbb{R}$, $f(A) = \operatorname{Trace} A$, where Trace means the sum of the diagonal entries. f is a linear map since

$$\begin{aligned} \operatorname{Trace}(A+B) &= \operatorname{Trace} A + \operatorname{Trace} B \\ \text{and} \qquad \operatorname{Trace}(\lambda A) &= \lambda \operatorname{Trace} A. \end{aligned}$$

■

EXAMPLE 5.5. Let $f : \mathcal{D}(\mathbb{R}) \longrightarrow \mathcal{C}(\mathbb{R})$ be the derivative operator:

$$f(p(x)) = p'(x) = \frac{dp}{dx}.$$

f is linear by the properties of the derivative:

$$\begin{aligned} \frac{d}{dx}(p+q) &= \frac{dp}{dx} + \frac{dq}{dx}; \\ \frac{d}{dx}(\lambda p) &= \lambda \frac{dp}{dx}. \end{aligned}$$

■

Some properties of linear maps.

LEMMA 5.6. *If $f : V \longrightarrow W$ is a linear map, then*

$$f(0_V) = 0_W.$$

*(So if f does not map zero to zero it **cannot** be linear.)*

NOTE 5.7. Translation in the plane is *not* a linear map. For the map $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $f(v) = v + v_0$ for some fixed non-zero vector v_0, we have $f(0) = v_0 \neq 0_{\mathbb{R}^2}$.

LEMMA 5.8. *If $f : V \longrightarrow W$ is a linear map, then*

$$f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f(v_i) \ \forall\, \alpha_1, \dots, \alpha_n \in \mathbb{R}, \forall\, v_1, \dots, v_n \in V.$$

PROOF. Use properties (i) and (ii) in Definition 5.1. □

DEFINITION 5.9. A map $f : V \longrightarrow W$ of real vector spaces V and W is called an **isomorphism** if f is both linear and bijective. We say V and W are **isomorphic** if there exists an isomorphism $f : V \longrightarrow W$, we write $V \cong W$.

REMARK 5.10. $V \cong W$ means there is a one-one correspondence between the sets V and W, which preserves the vector space structure.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ v & \longleftrightarrow & f(v) \end{array}$$

V and W are like two different copies of the same space.

EXAMPLE 5.11. Let $V = \mathbb{R}^2$, $W = \mathbb{C}$, complex numbers. $\mathbb{C}$ can be viewed as a real vector space via addition in $\mathbb{C}$ and multiplication by real numbers in the usual way.

$\mathbb{C}$ as a set can be identified with $\mathbb{R}^2$ (Argand diagram), so the map $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$ given by $f(x, y) = x + iy$ is bijective. Easy exercise to show f is linear, so f is an isomorphism of **real** vector spaces.

$$\mathbb{R}^2 \cong \mathbb{C} \quad \text{as } \textit{real} \text{ vector spaces.}$$

(Forget any other structure on $\mathbb{C}$ e.g., complex multiplication). ■

THEOREM 5.12 (Structure Theorem). *Let V be a real vector space of dimension n. Then V is isomorphic to $\mathbb{R}^n$.*

PROOF. Choose a basis $v_1, \dots, v_n$ for V. Let $e_1, \dots, e_n$ be the standard basis for $\mathbb{R}^n$, $e_1 = (1, 0, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0)$, etc.

Define $f : V \longrightarrow \mathbb{R}^n$ by $f(v_i) = e_i$ for each i, and

$$f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i e_i \text{ for each } \alpha_1, \dots, \alpha_n \in \mathbb{R}$$

f is well-defined, linear and bijective. Thus f is an isomorphism: $V \cong \mathbb{R}^n$. □

REMARK 5.13. In spite of our abstract definition of vector space, we have that any finite-dimensional space is just like $\mathbb{R}^n$. Our definition of dimension coincides with the usual geometric intuitive notion of dimension. There are two reasons for our abstract definition of vector space:

(a) Not all vector spaces are finite-dimensional;
(b) Finite-dimensional vector spaces may appear as $\mathbb{R}^n$ in disguise (e.g., $M_{r,s}\mathbb{R}$). It may not be obvious at first sight that you have a copy of $\mathbb{R}^n$.

6. Kernel and image

DEFINITION 6.1. Let V and W be real vector spaces and $f : V \longrightarrow W$ be a linear map.

The **kernel** of f, denoted $\ker f$, is

$$\ker f = \{v \in V : f(v) = 0_W\} \subseteq V$$

i.e., $\ker f$ = set of vectors in V which f sends to zero (the set of vectors killed by f).

Recall that the **image** of f, denoted $\operatorname{Im} f$ or $f(V)$, is

$$\operatorname{Im} f = \{w \in W : w = f(v) \text{ for some } v \in V\} \subseteq W$$

i.e., $\operatorname{Im} f$ = set of values in W taken by f.

EXERCISE 6.2. Show that $\ker f$ is a subspace of V and $\operatorname{Im} f$ is a subspace of W.

DEFINITION 6.3. Now suppose V and W are finite-dimensional.

The **nullity** of f is the dimension of $\ker f$.

The **rank** of f is the dimension of $\operatorname{Im} f$.

THEOREM 6.4 (Rank-Nullity Theorem — RNT). *Let V and W be finite-dimensional vector spaces and let $f : V \longrightarrow W$ be a linear map. Then*

$$\operatorname{rank} f + \operatorname{nullity} f = \dim V$$

i.e., $\dim(\ker f) + \dim(\operatorname{Im} f) = \dim V$.

EXAMPLE 6.5. Find the rank and nullity of the linear map $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$

$$f(w, x, y, z) = (w + x, x + y, y + z).$$

First we calculate the nullity.

$\ker f$ is the solution set of

$$\begin{aligned} w + x &= 0 \\ x + y &= 0 \\ y + z &= 0. \end{aligned}$$

Easy to see that $w = y = -x = -z$, i.e.,

$$\begin{aligned} \ker f &= \{(\alpha, -\alpha, \alpha, -\alpha) : \alpha \in \mathbb{R}\} \\ \ker f &= \{\alpha(1, -1, 1, -1) : \alpha \in \mathbb{R}\} \end{aligned}$$

so $\{(1, -1, 1, -1)\}$ is a basis, dimension is 1, i.e., nullity $f = 1$. Then, by RNT,

$$\begin{aligned} \operatorname{rank} f &= \dim \mathbb{R}^4 - \operatorname{nullity} f \\ &= 4 - 1 = 3. \end{aligned}$$

■

Some useful facts about linear maps.

LEMMA 6.6. *A linear map $f : V \longrightarrow W$ is injective iff* $\ker f = 0$*, i.e.,*

$$f(v) = 0_W \quad \Rightarrow \quad v = 0_V.$$

PROOF. Since f is linear

$$\begin{aligned} f(v_1) = f(v_2) \quad &\iff \quad f(v_1) - f(v_2) = 0_W \\ &\iff \quad f(v_1 - v_2) = 0_W. \end{aligned}$$

Hence if $\ker f = 0_W$ and $f(v_1) = f(v_2)$ then $v_1 - v_2 = 0_W$ i.e., $v_1 = v_2$ i.e., f injective.

So we've shown that $\ker f = 0 \Rightarrow f$ injective.

Conversely, if f is injective, then f can map at most one element to 0_W. We know $f(0_V) = 0_W$ (true for any linear map by Lemma 5.6), so f injective $\Rightarrow \ker f = 0$. □

REMARK 6.7. This is a useful way of checking for injectivity of a linear map, i.e., check that the kernel is zero. (We often write $\ker f = 0$ for $\ker f = \{0_V\}$.)

LEMMA 6.8. *Suppose $f : V \longrightarrow W$ is a linear map, and that* $\dim V = \dim W$ *(includes the special case of $V = W$). Then*

$$f \textit{ is injective} \iff f \textit{ is surjective.}$$

REMARK 6.9. Thus to show that a linear map $f : V \longrightarrow W$ with $\dim V = \dim W$ is an isomorphism, it suffices to show that f is injective (or to show that f is surjective).

Note that the *only* situation when a linear map $f : V \longrightarrow W$ can be bijective (i.e., an isomorphism) is when $\dim V = \dim W$.

EXAMPLE 6.10. $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $f(x, y, z) = (x + 2y, y + 2z, z + 2x)$.

f is linear, given by matrix $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$.

$V = W = \mathbb{R}^3$ so $\dim V = \dim W$.

We will show f is injective: $\ker f$ is solution set of

$$\begin{aligned} x + 2y &= 0 \\ y + 2z &= 0 \\ z + 2x &= 0 \end{aligned}$$

Easy exercise to check that $x = y = z = 0$ is the only solution of these equations.

$\therefore \ker f = 0$ and f is injective, so f is an isomorphism. ■

7. Matrices

7.1. Let A be an $m \times n$ matrix with real entries

$$A = (a_{ij}) = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} & \begin{matrix} m \\ \text{rows} \end{matrix} \\ n \text{ columns} & \end{array}$$

A gives rise to a linear map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ as follows:

Write elements of $\mathbb{R}^n$ as column vectors, $v \in \mathbb{R}^n$. Think of a column vector in $\mathbb{R}^n$ as an $n \times 1$ matrix (a row vector in $\mathbb{R}^n$ is a $1 \times n$ matrix).

Define $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by

$$f(v) = \underbrace{A \cdot v}_{(m \times n)(n \times 1)}, \ \forall\, v \in \mathbb{R}^n.$$

This matrix product Av is an $m \times 1$ matrix i.e., Av is a column vector in $\mathbb{R}^m$ (see 5.3).

f is a linear map because of the properties of matrix multiplication:

$$\begin{aligned} f(v_1 + v_2) &= A(v_1 + v_2) \\ &= Av_1 + Av_2 \\ &= f(v_1) + f(v_2) \end{aligned}$$

and

$$\begin{aligned} f(\lambda v) &= A(\lambda v) \\ &= \lambda Av \\ &= \lambda f(v). \end{aligned}$$

We'll now see that every linear map $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ arises in this way from some $m \times n$ matrix.

NOTE 7.2. Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$, written as columns. For an $m \times n$ matrix A we have

$$\begin{aligned} Ae_1 &= \text{first column of } A \\ Ae_2 &= \text{second column of } A \\ \vdots & \qquad\qquad \vdots \\ Ae_n &= n^{\text{th}} \text{ column of } A \end{aligned}$$

since

$$Ae_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}$$

etc., etc.

PROPOSITION 7.3. *There is a one-one correspondence between the set of all linear maps* $: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *and the set of all* $m \times n$ *matrices with real entries.*

EXAMPLE 7.4. Consider the linear map

$$T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} 3x_1 - x_2 \\ x_1 + 2x_2 + x_3 \end{pmatrix}.$$

To find its matrix with respect to the standard bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, we calculate

$$\begin{aligned} T(e_1) &= T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3(1) - 0 \\ 1 + 2(0) + 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \\ T(e_2) &= T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3(0) - 1 \\ 0 + 2(1) + 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \qquad \text{and} \\ T(e_3) &= T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3(0) - 0 \\ 0 + 2(0) + 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Thus the corresponding matrix is (you could just read off the entries from the description of T)

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

■

PROPOSITION 7.5. *Let V be a real vector space of dimension n and W be a real vector space of dimension m. There is a one-one correspondence between the set of all linear maps $V \longrightarrow W$ and the set of all $m \times n$ matrices with real entries.*

PROOF. By our earlier Structure Theorem, $V \cong \mathbb{R}^n$.

The isomorphism came from choosing a basis $v_1, \ldots, v_n$ for V and mapping

$$\begin{array}{rcl} V & \longrightarrow & \mathbb{R}^n \\ v_i & \longmapsto & e_i \ \forall\, i \\ & & \nwarrow \text{standard basis} \end{array}$$

Similarly $W \cong \mathbb{R}^m$ via choice of basis $w_1, \ldots, w_m$ and mapping $w_i \longmapsto e_i \ \forall\, i$.

Identify V with $\mathbb{R}^n$ and W with $\mathbb{R}^m$ by the above isomorphisms — then this propn. follows at once from the previous propn. □

REMARK 7.6. The matrix corresponding to $f : V \longrightarrow W$ depends on the choice of bases for V and W.

8. The determinant and trace

8.1. Recall that the **determinant** of a square matrix A is a real number (denoted by $\det A$ or $|A|$) associated to the matrix as follows.

The determinant of a 1×1 matrix (a) is just a itself. For larger matrices, compute the determinant recursively in terms of smaller determinants, by evaluating along a chosen row or column: For each entry a_{ij} in the chosen row (or column):

- work out the determinant of the submatrix of A got by deleting the row and column of A containing a_{ij} (write this as $\det(\cdots)$),
- multiply $\det(\cdots)$ by $(-1)^{i+j}$ to get $(-1)^{i+j}\det(\cdots)$, this is called the **cofactor** of the entry a_{ij} and is usually written C_{ij},
- then multiply $(-1)^{i+j}\det(\cdots)$ (i.e., C_{ij}) by a_{ij} to get $a_{ij}C_{ij}$;
- finally, add up the $a_{ij}C_{ij}$ for each a_{ij} in the chosen row or column.

For a 4×4 matrix you would work out four determinants of size 3×3, each of which would be worked out by reducing to three 2×2 determinants as above. The computations become very tedious for such larger matrices and are generally done by computer.

We usually choose the row or column containing the most 0's, as this gives us less work: $0 \times (-1)^{i+j}\det(\cdots)$ will of course be 0, e.g., the next example.

THEOREM 8.2. *Let A be an $n \times n$ square matrix.*

(*a*) *The determinant of the transpose is the same as the determinant of the matrix itself, $|A| = |A^T|$.*
(*b*) *If we multiply a row (or column) of A by a number α then the determinant of the new matrix is equal to $\alpha|A|$.*
(*c*) *If we interchange two rows (or columns) of A then the determinant of A changes its sign but keeps its absolute value.*
(*d*) *If a multiple of one row (or column) is added to another row (or column) then the determinant stays the same.*
(*e*) *If the rows (or columns) of A are linearly independent considered as n-vectors then* $\det A \neq 0$. *Otherwise* $\det A = 0$. *In particular, if A has one or more zero rows or columns (i.e., all the entries in that row or column are equal to 0) then* $\det A = 0$.
(*f*) *The determinant of a diagonal matrix is the product of the diagonal entries.*

THEOREM 8.3. *Let A and B be two $n \times n$ matrices. Then*

$$\det(AB) = \det A \det B.$$

DEFINITION 8.4. The **trace** of a square $n \times n$ matrix A, $\mathrm{Trace}(A)$, is defined to be the sum of the diagonal entries of A:

$$\mathrm{Trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

REMARK 8.5. In a sense, the trace is to addition what the determinant is to multiplication. For example, part (*f*) of Theorem 8.2 is clearly analogous to the definition of trace. Also, corresponding to Theorem 8.3 there is the following

Theorem 8.6. *Let A and B be two $n \times n$ matrices. Then*

$$\operatorname{Trace}(A+B) = \operatorname{Trace} A + \operatorname{Trace} B.$$

We will meet further properties of the determinant and the trace in connection with eigenvalues.

9. The inverse of a matrix

Let A be a $n \times n$ (square) matrix. If there exists an $n \times n$ matrix X such that

$$AX = XA = I$$

then we say that X is the **inverse** of A. The matrix A is said to be **invertible**. (So is X, of course.) Equivalently, the linear map f associated to A is invertible (i.e., bijective, an isomorphism).

The inverse of the matrix A is usually denoted by A^{-1}.

Proposition 9.1 (Properties of the inverse). *Let A and B be two invertible matrices. Then*

(a) *A^{-1} is invertible and $(A^{-1})^{-1} = A$ (the inverse of the inverse is the matrix itself).*
(b) *AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$ (the inverse of the product is the product of inverses, taken in the reverse order — similarly for transposes).*
(c) *The transpose of an invertible matrix is invertible: $(A^T)^{-1} = (A^{-1})^T$ (the inverse of the transpose is the transpose of the inverse).*

Theorem 9.2. *A matrix has an inverse if and only if its determinant is not equal to 0.*

Remark 9.3. We can now solve problems of the following type: If A and B are two given matrices with A invertible, find a matrix X such that $AX = B$. By multiplying both sides of $AX = B$ by the matrix A^{-1} we get

$$A^{-1}AX = A^{-1}B$$

or

$$IX = X = A^{-1}B.$$

10. The rank of a matrix

Associated with each $m \times n$ matrix A is an important number called its rank, denoted by $\operatorname{rank}(A)$.

Definition 10.1. A **minor** of order k of the matrix A is obtained by deleting all but k rows and k columns of the matrix and taking the determinant of this $k \times k$ matrix.

Definition 10.2. The **rank** of an $m \times n$ matrix A, $\operatorname{rank}(A)$, is defined to be the rank (dimension of the image) of the associated linear map.

Equivalently, $\operatorname{rank}(A)$ is the order of the largest minor of A that is different from 0.

Equivalently, $\operatorname{rank}(A)$ is the number of non-zero rows when A is reduced to echelon form.

11. Eigenvalues and eigenvectors

Definition 11.1. Let A be an $n \times n$ matrix. We say that a real number λ is an **eigenvalue** of A if there is a vector $v \neq \mathbf{0}$ such that

$$Av = \lambda v.$$

The vector v is called an **eigenvector of A corresponding to λ**, or a **λ-eigenvector**. The set of all λ-eigenvectors is called the **λ-eigenspace**.

Geometrically, this means the eigenvectors are directions *fixed* by A, i.e., all A does is to expand or contract vectors in these “special” directions, and the expansion or contraction is by a fixed amount, namely, the eigenvalue for that direction.

Note that 0 is an eigenvalue if and only if the associated linear map of the matrix is not injective (that is, some non-zero vector is killed off by the matrix: a dimension is collapsed to nothing). This is equivalent to the determinant being 0 (see Theorem 11.5).

EXERCISE 11.2. Let A be an $n \times n$ matrix and λ be an eigenvalue of A. Show that the λ-eigenspace is in fact a subspace of $\mathbb{R}^n$, and thus a vector space in its own right.

THEOREM 11.3. *Suppose that A is an $n \times n$ matrix and λ is an eigenvalue of A. Then*

$$\det(A - \lambda I) = 0 \qquad (**)$$

PROOF. Suppose λ is an eigenvalue of A and v is a λ-eigenvector, i.e., $v \neq \mathbf{0}$ and $Av = \lambda v$. Thus

$$\begin{aligned} & Av = \lambda v \\ \Rightarrow \quad & Av - \lambda v = \mathbf{0} \\ \Rightarrow \quad & Av - \lambda I v = \mathbf{0} \qquad \text{where } I \text{ is the identity matrix} \\ \Rightarrow \quad & (A - \lambda I)v = \mathbf{0} \qquad (*) \end{aligned}$$

Thus $A - \lambda I$ sends the nonzero vector v to $\mathbf{0}$. This means the matrix $A - \lambda I$ is not invertible.

For, suppose $A - \lambda I$ has an inverse, B. Then

$$v = Iv = (B(A - \lambda I))v = B((A - \lambda I)v) = B\mathbf{0} = \mathbf{0}$$

but this is impossible as $v \neq \mathbf{0}$.

This contradiction shows no such B can exist, i.e., $A - \lambda I$ is not invertible. Thus since a non-invertible matrix has determinant zero by Theorem 9.2,

$$\det(A - \lambda I) = 0$$

which is what we wished to prove. □

REMARK 11.4. Equation $(**)$ in Theorem 11.3 is so important it is given a name, the **characteristic equation**. The left hand side of the characteristic equation $(**)$ is a polynomial of degree n in the variable λ; it is called the **characteristic polynomial**. The values of λ which give 0 on the right hand side of $(**)$ are the eigenvalues, i.e.,

The eigenvalues of A are the roots of the characteristic equation of A.

Since a polynomial of degree n has n roots, we thus see that an $n \times n$ matrix has n eigenvalues (but note they need not be distinct, e.g., see Example 11.6 below).

Recall that a polynomial with real coefficients will in general have complex roots: that is, the eigenvalues of a real square matrix are in general *complex* numbers.

THEOREM 11.5. *Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ (not necessarily distinct). Then*

$$\det A = \lambda_1 \cdots \lambda_n$$

and

$$\operatorname{Trace} A = \lambda_1 + \cdots + \lambda_n$$

i.e., the determinant of A is the product of the eigenvalues, and the trace of A is the sum of the eigenvalues.

EXAMPLE 11.6. Let

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}.$$

We find the eigenvalues of A as follows. Firstly,

$$I = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{so} \qquad \lambda I = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

and thus

$$A - \lambda I = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 3-\lambda & 1 & 1 \\ 2 & 4-\lambda & 2 \\ 1 & 1 & 3-\lambda \end{pmatrix}.$$

Then

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

which is

$$\begin{aligned}
&= (3-\lambda)\det\begin{pmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{pmatrix} - \det\begin{pmatrix} 2 & 2 \\ 1 & 3-\lambda \end{pmatrix} + \det\begin{pmatrix} 2 & 4-\lambda \\ 1 & 1 \end{pmatrix} \\
&= (3-\lambda)\Big((4-\lambda)(3-\lambda)-(2)(1)\Big) - \Big(2(3-\lambda)-(2)(1)\Big) \\
&\quad + \Big((2)(1)-(4-\lambda)(1)\Big) \\
&= (3-\lambda)(12-7\lambda+\lambda^2-2)+2-6+2\lambda\ +2-4+\lambda \\
&= 30-21\lambda+3\lambda^2-10\lambda+7\lambda^2-\lambda^3-6+3\lambda \\
&= -\lambda^3+10\lambda^2-28\lambda+24
\end{aligned}$$

This is the characteristic polynomial and its roots are 2,2 and 6 (not distinct!).

To see this, note that by the last Theorem, the sum of the eigenvalues is Trace $A = 3+4+3 = 10$, and their product is $\det A$. We could work out $\det A$ from scratch but this is not necessary since $\det A$ is just the constant term in the characteristic polynomial, i.e., 24. (Why? Because the characteristic polynomial is $-\lambda^3+10\lambda^2-28\lambda+24 = \det(A-\lambda I)$ and if we set $\lambda = 0$ on both sides we see $24 = \det(A-0I) = \det A$.)

■

REMARK 11.7. The last example had integer eigenvalues for illustration but this will not usually be the case—in general you would get the roots of the characteristic polynomial using a software package.

Having found the eigenvalues we may then (for each eigenvalue, one by one) solve the vector-matrix equation

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

($(*)$ in the proof of Theorem 11.3 above) to find the eigenvectors. To do this, we need to be able to solve systems of simultaneous linear equations (see Appendix to this chapter).

DEFINITION 11.8. The non-negative (real) number

$$r_\sigma(A) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$$

is called the **spectral radius** of A: it is the magnitude of the largest eigenvalue of A in the complex plane $\mathbb{C}$. Later, we will see it is a measure of the "norm" of a matrix.

The Perron Frobenius theorem, proved by Oskar Perron (1907) and Georg Frobenius (1912), asserts that a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector has strictly positive components, and also asserts a similar statement for certain classes of nonnegative matrices.

THEOREM 11.9. *Let $A = (a_{ij})$ be an $n \times n$ positive matrix: $a_{ij} > 0$ for $1 \le i,\ j \le n$. Then the following statements hold.*

- (a) *There is a positive real number r, called the Perron root or the Perron Frobenius eigenvalue, such that r is an eigenvalue of A and any other eigenvalue λ (possibly, complex) is strictly smaller than r in absolute value, $|\lambda| < r$. Thus, the spectral radius $r_\sigma(A)$ is equal to r.*
- (b) *The Perron Frobenius eigenvalue is simple: r is a simple root of the characteristic polynomial of A. Consequently, eigenspace associated to r is one-dimensional.*
- (c) *There exists an eigenvector $v = (v_1, \ldots, v_n)$ of A with eigenvalue r such that all components of v are positive: $Av = rv$, $v_i > 0$ for $1 \le i \le n$.*

This theorem has important applications to probability theory (ergodicity of Markov chains); to economics (Leontief's input-output model); to demography (Leslie population age distribution model); to mathematical background of the internet search engines; and even to ranking of football teams.