0.1 Determinants

Let $A = (a_{ij})$ be a 2×2 matrix. Recall that the determinant of A was defined by

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} \,. \tag{1}$$

Notation 0.1.1. For any $n \times n$ matrix A, let A_{ij} denote the submatrix formed by deleting the i-th row and the j-th column of A. We call A_{ij} the (i, j)-*minor* of A.

Warning: This notation differs from the one used in the course text

Example 0.1.2. If

$$A = \begin{pmatrix} 3 & 2 & 5 & -1 \\ -2 & 9 & 0 & 6 \\ 7 & -2 & -3 & 1 \\ 4 & -5 & 8 & -4 \end{pmatrix} ,$$

then

$$A_{23} = \begin{pmatrix} 3 & 2 & -1 \\ 7 & -2 & 1 \\ 4 & -5 & -4 \end{pmatrix} .$$

Definition 0.1.3. Let $A = (a_{ij})$ be an $n \times n$ matrix. The *determinant* of A, written $\det(A)$, is defined as follows:

- If $n = 1$, then $\det(A) = a_{11}$.
- If $n > 1$ then $\det(A)$ is the sum of n terms of the form $\pm a_{i1} \det(A_{i1})$, with plus and minus signs alternating, and where the entries $a_{11}, a_{21}, \ldots, a_{n1}$ are from the first column of A. In symbols:

$$\begin{aligned} \det(A) &= a_{11} \det(A_{11}) - a_{21} \det(A_{21}) + \cdots + (-1)^{n+1} a_{n1} \det(A_{n1}) \\ &= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det(A_{i1}) \,. \end{aligned}$$

Example 0.1.4. Compute the determinant of

$$A = \begin{pmatrix} 0 & 0 & 7 & -5 \\ -2 & 9 & 6 & -8 \\ 0 & 0 & -3 & 2 \\ 0 & 3 & -1 & 4 \end{pmatrix}.$$

Solution.

$$\begin{vmatrix} 0 & 0 & 7 & -5 \\ -2 & 9 & 6 & -8 \\ 0 & 0 & -3 & 2 \\ 0 & 3 & -1 & 4 \end{vmatrix} = -(-2)\begin{vmatrix} 0 & 7 & -5 \\ 0 & -3 & 2 \\ 3 & -1 & 4 \end{vmatrix} = 2\cdot 3\begin{vmatrix} 7 & -5 \\ -3 & 2 \end{vmatrix} = 2\cdot 3\cdot[7\cdot 2-(-3)\cdot(-5)] = -6\,.$$

□

Definition 0.1.5. Given a square matrix $A = (a_{ij})$, the (i, j)**-cofactor** of A is the number C_{ij} defined by

$$C_{ij} = (-1)^{i+j}\det(A_{ij})\,.$$

Thus, the definition of $\det(A)$ reads

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}.$$

This is called the **cofactor expansion down the first column of** A.

Theorem 0.1.6 (Cofactor Expansion Theorem)**.** *The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any column or row. The expansion down the j-th column is*

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

and the cofactor expansion across the i-th row is

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}\,.$$

Example 0.1.7. Use a cofactor expansion across the second row to compute $\det(A)$, where

$$A = \begin{pmatrix} 4 & -1 & 3 \\ 0 & 0 & 2 \\ 1 & 0 & 7 \end{pmatrix}.$$

Solution.

$$\begin{aligned}\det(A) &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23}\\ &= (-1)^{2+1}a_{21}\det(A_{21}) + (-1)^{2+2}a_{22}\det(A_{22}) + (-1)^{2+3}a_{23}\det(A_{23})\\ &= -0\begin{vmatrix}-1 & 3\\ 0 & 7\end{vmatrix} + 0\begin{vmatrix}4 & 3\\ 1 & 7\end{vmatrix} - 2\begin{vmatrix}4 & -1\\ 1 & 0\end{vmatrix}\\ &= -2[4\cdot 0 - 1\cdot(-1)] = -2\,.\end{aligned}$$

□

Theorem 0.1.8. *If A is either an upper or a lower triangular matrix, then* $\det(A)$ *is the product of the diagonal entries of A.*

0.2 Properties of determinants

Theorem 0.2.1. *Let A be an $n \times n$ matrix.*

(a) If two rows of A are interchanged to produce B, then $\det(B) = -\det(A)$.

(b) If one row of A is multiplied by α to produce B, then $\det(B) = \alpha\det(A)$.

(c) If a multiple of one row of A is added to another row to produce a matrix B then $\det(B) = \det(A)$.

Example 0.2.2. Compute

$$\begin{vmatrix}3 & -1 & 2 & -5\\ 0 & 5 & -3 & -6\\ -6 & 7 & -7 & 4\\ -5 & -8 & 0 & 9\end{vmatrix}.$$

Solution. Perhaps the easiest way to compute this determinant is to spot that when adding two times row 1 to row 3 we get two identical rows, which, by another application of the previous theorem, implies that the determinant is zero:

$$\begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{vmatrix} \underset{R_3+2R_1}{=} \begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{vmatrix}$$

$$\underset{R_3-R_2}{=} \begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 0 & 0 & 0 \\ -5 & -8 & 0 & 9 \end{vmatrix} = 0\,,$$

by a cofactor expansion across the third row. □

Theorem 0.2.3. *A matrix A is invertible if and only if* $\det(A) \neq 0$.

Definition 0.2.4. A square matrix A is called **singular** if $\det(A) = 0$. Otherwise it is said to be **nonsingular**.

Corollary 0.2.5. *A matrix is invertible if and only if it is nonsingular*

Theorem 0.2.6. *If A is an $n \times n$ matrix, then* $\det(A) = \det(A^T)$.

Proof. We prove by induction. The statement is clearly true for 2×2 matrices A. Suppose the statement is true for $k \times k$ matrices A. This is the Inductive Hypothesis. We show that the statement is also true for $(k+1) \times (k+1)$-matrices A.

We fix some notation first. Write $A = (a_{ij})$, $A^T = (a^t_{ij})$ with $a^t_{ij} = a_{ji}$. Observe that

$$A_{ij} = (A^T_{ji})^T.$$

Now let A be a $(k+1) \times (k+1)$-matrix. Then we have

$$\begin{aligned} \det A &= \sum_{j=1}^{k+1} a_{ij}(-1)^{i+j} \det A_{ij} = \sum_{j=1}^{k+1} a_{ij}(-1)^{i+j} \det (A^T_{ji})^T \\ &= \sum_{j=1}^{k+1} a_{ij}(-1)^{i+j} \det A^T_{ji} \quad \text{(by Inductive Hypothesis since } A^T_{ji} \text{ is } k \times k) \\ &= \sum_{j=1}^{k+1} a^t_{ji}(-1)^{j+i} \det A^T_{ji} \\ &= \det A^T. \end{aligned}$$

□

By the previous theorem, each statement of the theorem on the behaviour of determinants under row operations (Theorem 0.2.1) is also true if the word 'row' is replaced by 'column', since a row operation on A^T amounts to a column operation on A.

Theorem 0.2.7. *Let A be a square matrix.*

(a) If two columns of A are interchanged to produce B, then $\det(B) = -\det(A)$.

(b) If one column of A is multiplied by α to produce B, then $\det(B) = \alpha\det(A)$.

(c) If a multiple of one column of A is added to another column to produce a matrix B then $\det(B) = \det(A)$.

Theorem 0.2.8. *If A is an $n \times n$ matrix and E an elementary $n \times n$ matrix, then*

$$\det(EA) = \det(E)\det(A)$$

with

$$\det(E) = \begin{cases} -1 & \text{if } E \text{ is of type I (interchanging two rows)} \\ \alpha & \text{if } E \text{ is of type II (multiplying a row by } \alpha\text{)} \\ 1 & \text{if } E \text{ is of type III (adding a multiple of one row to another)} \end{cases} .$$

Theorem 0.2.9. *If A and B are square matrices of the same size, then*

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

Proof. Case I: If A is not invertible, then neither is AB, for otherwise $A(B(AB)^{-1}) = I$. Thus, by Theorem 0.2.3,

$$\det(AB) = 0 = 0 \cdot \det(B) = \det(A)\det(B)\,.$$

Case II: If A is invertible, then by the Invertible Matrix Theorem A is a product of elementary matrices, that is, there exist elementary matrices $E_1, \dots, E_k$, such that

$$A = E_k E_{k-1} \cdots E_1\,.$$

For brevity, write $|A|$ for $\det(A)$. Then, by the previous theorem,

$$\begin{aligned}|AB| &= |E_k \cdots E_1 B| = |E_k||E_{k-1} \cdots E_1 B| = \dots \\ &= |E_k| \cdots |E_1||B| = \dots = |E_k \cdots E_1||B| \\ &= |A||B|\,.\end{aligned}$$

□

Let C_{ij} be the (i, j)-cofactor of an $n \times n$ matrix A. We define the **adjugate** of A, denoted by $\operatorname{adj} A$, to be the following matrix of cofactors (note that the order of the indices is reversed!) :

$$\operatorname{adj} A = (C_{ji}) = \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix} \tag{2}$$

Theorem 0.2.10 (Inverse Formula)**.** *Let A be an $n \times n$ matrix. Then*

$$A(\operatorname{adj} A) = (\det A)I$$

where I is the identity matrix. Further, if A is invertible, then

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)\,.$$

Example 0.2.11. Find the inverse of the following matrix using the Inverse Formula

$$A = \begin{pmatrix} 1 & 3 & -1 \\ -2 & -6 & 0 \\ 1 & 4 & -3 \end{pmatrix}\,.$$

Proof. First we need to calculate the 9 cofactors of A:

$$C_{11} = +\begin{vmatrix} -6 & 0 \\ 4 & -3 \end{vmatrix} = 18\,, \quad C_{12} = -\begin{vmatrix} -2 & 0 \\ 1 & -3 \end{vmatrix} = -6\,, \quad C_{13} = +\begin{vmatrix} -2 & -6 \\ 1 & 4 \end{vmatrix} = -2\,,$$

$$C_{21} = -\begin{vmatrix} 3 & -1 \\ 4 & -3 \end{vmatrix} = 5\,, \quad C_{22} = +\begin{vmatrix} 1 & -1 \\ 1 & -3 \end{vmatrix} = -2\,, \quad C_{23} = -\begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = -1\,,$$

$$C_{31} = +\begin{vmatrix} 3 & -1 \\ -6 & 0 \end{vmatrix} = -6\,, \quad C_{32} = -\begin{vmatrix} 1 & -1 \\ -2 & 0 \end{vmatrix} = 2\,, \quad C_{33} = +\begin{vmatrix} 1 & 3 \\ -2 & -6 \end{vmatrix} = 0\,.$$

Thus

$$\operatorname{adj}(A) = \begin{pmatrix} 18 & 5 & -6 \\ -6 & -2 & 2 \\ -2 & -1 & 0 \end{pmatrix},$$

and since $\det(A) = 2$, we have

$$A^{-1} = \begin{pmatrix} 9 & \frac{5}{2} & -3 \\ -3 & -1 & 1 \\ -1 & -\frac{1}{2} & 0 \end{pmatrix}.$$

$\square$

0.3 Vector spaces

In this chapter, we will study abstract vector spaces. Roughly speaking a vector space is a mathematical structure on which an operation of addition and an operation of scalar multiplication is defined, and we require these operations to obey a number of algebraic rules. We have already encountered examples of vector spaces in this module. Recall that $\mathbb{R}^n$ is the collection of all n-vectors. On $\mathbb{R}^n$ two operations were defined:

- *addition*: if

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n\,, \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n\,,$$

  then $\mathbf{x} + \mathbf{y}$ is the n-vector given by

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}\,.$$

- *scalar multiplication*: if

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n\,, \quad \text{and} \quad \alpha \text{ is a scalar}$$

  then $\alpha\mathbf{x}$ is the n-vector given by

$$\alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}\,.$$

After these operations were defined, it turned out that they satisfy a number of rules We are now going to turn this process on its head. That is, we start from a set on which two operations are defined, we *postulate* that these operations satisfy certain rules, and we call the resulting structure a ‘vector space’:

Definition 0.3.1. A **vector space over** $\mathbb{R}$, or a **a real vector space**, is a non-empty set V, equipped with two operations which are mappings

$$(\mathbf{u}, \mathbf{v}) \in V \times V \mapsto \mathbf{u} + \mathbf{v} \in V\,, \quad (\alpha, \mathbf{u}) \in \mathbb{R} \times V \mapsto \alpha\mathbf{u} \in V,$$

called respectively *addition* and *scalar multiplication*, satisfying the following axioms:

(C1) the sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in V;

(C2) the scalar multiple of $\mathbf{u}$ by α, denoted by $\alpha\mathbf{u}$, is in V;

(A1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

(A2) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$;

(A3) there is an element $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$;

(A4) for each $\mathbf{u}$ in V there is an element $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$;

(A5) $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$;

(A6) $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$;

(A7) $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u})$;

(A8) $1\mathbf{u} = \mathbf{u}$

for all $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all $\alpha, \beta \in \mathbb{R}$.

The elements in V are called **vectors**, and we usually write them using bold letters $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, etc. The numbers in $\mathbb{R}$ are called the **scalars**.

If, in the above definition, the scalar field $\mathbb{R}$ is replaced by the complex numbers $\mathbb{C}$, then we call V a *vector space over* $\mathbb{C}$, or a *complex vector space*.

Throughout, by a vector space V, we shall mean either a **real** or a **complex** vector space V.

Example 0.3.2. Let $\mathbb{R}^{m\times n}$ denote the set of all $m \times n$ matrices. Define addition and scalar multiplication of matrices in the usual way. Then $\mathbb{R}^{n\times m}$ is a real vector space.

Example 0.3.3. Let P_n denote the set of all real polynomials with real coefficients of degree less than n. Thus, an element $\mathbf{p}$ in P_n is of the form

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \,,$$

where the coefficients $a_0, \ldots, a_n$ and the variable t are real numbers.

Define addition and scalar multiplication on P_n as follows: if $\mathbf{q} \in P_n$ is given by

$$\mathbf{q}(t) = b_0 + b_1 t + b_2 t^2 + \cdots + b_n t^n \,,$$

$\mathbf{p}$ is as above and α a scalar, then

- $\mathbf{p} + \mathbf{q}$ is the polynomial

$$(\mathbf{p} + \mathbf{q})(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n$$

- $\alpha\mathbf{p}$ is the polynomial

$$(\alpha\mathbf{p})(t) = (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_n)t^n \,.$$

Note that (C1) and (C2) clearly hold, since if $\mathbf{p}, \mathbf{q} \in P_n$ and α is a scalar, then $\mathbf{p} + \mathbf{q}$ and $\alpha\mathbf{p}$ are again polynomials of degree less than n. Axiom (A1) holds since if $\mathbf{p}$ and $\mathbf{q}$ are as above, then

$$\begin{aligned}(\mathbf{p} + \mathbf{q})(t) &= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n \\ &= (b_0 + a_0) + (b_1 + a_1)t + \cdots + (b_n + a_n)t^n \\ &= (\mathbf{q} + \mathbf{p})(t)\end{aligned}$$

so $\mathbf{p} + \mathbf{q} = \mathbf{q} + \mathbf{p}$. A similar calculation shows that (A2) holds. Axiom (A3) holds if we let $\mathbf{0}$ be the zero polynomial, that is

$$\mathbf{0}(t) = 0 + 0 \cdot t + \cdots + 0 \cdot t^n \,,$$

since then $(\mathbf{p} + \mathbf{0})(t) = \mathbf{p}(t)$, that is, $\mathbf{p} + \mathbf{0} = \mathbf{p}$. Axiom (A4) holds if, given $\mathbf{p} \in P_n$ we set $-\mathbf{p} = (-1)\mathbf{p}$, since then

$$(\mathbf{p} + (-\mathbf{p}))(t) = (a_0 - a_0) + (a_1 - a_1)t + \cdots + (a_n - a_n)t^n = \mathbf{0}(t) \,,$$

that is $\mathbf{p} + (-\mathbf{p}) = \mathbf{0}$. The remaining axioms are easily verified as well, using familiar properties of real numbers.

Example 0.3.4. Let $C[a, b]$ denote the set of all real-valued functions that are defined and continuous on the closed interval $[a, b]$. For $\mathbf{f}, \mathbf{g} \in C[a, b]$ and α a scalar, define $\mathbf{f} + \mathbf{g}$ and $\alpha\mathbf{f}$ *pointwise*, that is, by

$$(\mathbf{f} + \mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t) \quad \text{for all } t \in [a, b]$$

$$(\alpha\mathbf{f})(t) = \alpha\mathbf{f}(t) \quad \text{for all } t \in [a, b]$$

Equipped with these operations, $C[a, b]$ is a vector space. The closure axiom (C1) holds because the sum of two continuous functions on $[a, b]$ is continuous on $[a, b]$, and (C2) holds because a constant times a continuous function on $[a, b]$ is again continuous on $[a, b]$. Axiom (A1) holds as well, since for all $t \in [a, b]$

$$(\mathbf{f} + \mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t) = \mathbf{g}(t) + \mathbf{f}(t) = (\mathbf{g} + \mathbf{f})(t)\,,$$

so $\mathbf{f} + \mathbf{g} = \mathbf{g} + \mathbf{f}$. Axiom (A3) is satisfied if we let $\mathbf{0}$ be the zero function,

$$\mathbf{0}(t) = 0 \quad \text{for all } t \in [a, b]\,,$$

since then

$$(\mathbf{f} + \mathbf{0})(t) = \mathbf{f}(t) + \mathbf{0}(t) = \mathbf{f}(t) + 0 = \mathbf{f}(t)\,,$$

so $\mathbf{f}+\mathbf{0} = \mathbf{f}$. Axiom (A4) holds if, given $\mathbf{f} \in C[a, b]$, we let $-\mathbf{f}$ be the function

$$(-\mathbf{f})(t) = -\mathbf{f}(t) \quad \text{for all } t \in [a, b],$$

since then

$$(\mathbf{f} + (-\mathbf{f}))(t) = \mathbf{f}(t) + (-\mathbf{f})(t) = \mathbf{f}(t) - \mathbf{f}(t) = 0 = \mathbf{0}(t)\,,$$

that is, $\mathbf{f} + (-\mathbf{f}) = \mathbf{0}$. We leave it as an exercise to verify the remaining axioms.

We shall now derive a number of elementary properties of vector spaces.

Theorem 0.3.5. *If V is a vector space and $\mathbf{u}$ and $\mathbf{v}$ are elements in V, then*

(a) $0\mathbf{u} = \mathbf{0}$;

(b) if $\mathbf{u} + \mathbf{v} = \mathbf{0}$ *then* $\mathbf{v} = -\mathbf{u}$;[1]

[1] In the language of MTH4104 (Introduction to Algebra) this statement says that the additive inverse is unique.

(c) $(-1)\mathbf{u} = -\mathbf{u}$.

Proof. (a) We start by observing that

$$\mathbf{u} \overset{(A8)}{=} 1\mathbf{u} = (0+1)\mathbf{u} \overset{(A6)}{=} 0\mathbf{u} + 1\mathbf{u} \overset{(A8)}{=} 0\mathbf{u} + \mathbf{u}\,. \tag{3}$$

Now, by (A4), there is an element $-\mathbf{u} \in V$ such that

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}\,. \tag{4}$$

Thus

$$\mathbf{0} \overset{(4)}{=} \mathbf{u} + (-\mathbf{u}) \overset{(3)}{=} (0\mathbf{u} + \mathbf{u}) + (-\mathbf{u}) \overset{(A2)}{=} 0\mathbf{u} + (\mathbf{u} + (-\mathbf{u})) \overset{(4)}{=} 0\mathbf{u} + \mathbf{0} \overset{(A3)}{=} 0\mathbf{u}\,.$$

(b) Suppose that $\mathbf{u} + \mathbf{v} = \mathbf{0}$. Then

$$\begin{aligned} -\mathbf{u} &\overset{(A3)}{=} -\mathbf{u} + \mathbf{0} = -\mathbf{u} + (\mathbf{u} + \mathbf{v}) \overset{(A2)}{=} (-\mathbf{u} + \mathbf{u}) + \mathbf{v} \\ &\overset{(A1)}{=} (\mathbf{u} + (-\mathbf{u})) + \mathbf{v} \overset{(A4)}{=} \mathbf{0} + \mathbf{v} \overset{(A1)}{=} \mathbf{v} + \mathbf{0} \overset{(A3)}{=} \mathbf{v}\,. \end{aligned}$$

(c) Notice that

$$\mathbf{0} \overset{(a)}{=} 0\mathbf{u} = (1 + (-1))\mathbf{u} \overset{(A6)}{=} 1\mathbf{u} + (-1)\mathbf{u} \overset{(A8)}{=} \mathbf{u} + (-1)\mathbf{u}\,,$$

so, by (b), we conclude that $(-1)\mathbf{u} = -\mathbf{u}$. □

0.4 Subspaces

Definition 0.4.1. A nonempty subset H of a vector space V is called a *subspace* of V if it satisfies the following two conditions:

(i) if $\mathbf{u}, \mathbf{v} \in H$, then $\mathbf{u} + \mathbf{v} \in H$;

(ii) if $\mathbf{u} \in H$ and α is a scalar, then $\alpha\mathbf{u} \in H$.

Theorem 0.4.2. *Let H be a subspace of a vector space V. Then H with addition and scalar multiplication inherited from V is a vector space in its own right.*

Remark 0.4.3. If V is a vector space, then $\{\mathbf{0}\}$ and V are clearly subspaces of V. All other subspaces are said to be **proper subspaces** of V. We call $\{\mathbf{0}\}$ the **zero subspace** of V.

Example 0.4.4. Show that the following are subspaces of $\mathbb{R}^3$:

(a) $L = \{ (r, s, t,)^T \mid r, s, t \in \mathbb{R} \text{ and } r = s = t \}$;[2]

(b) $P = \{ (r, s, t,)^T \mid r, s, t \in \mathbb{R} \text{ and } r - s + 3t = 0 \}$.

Solution. (a) Notice that an arbitrary element in L is of the form $r(1, 1, 1)^T$ for some real number r. Thus, in particular, L is not empty, since $(0, 0, 0)^T \in L$. In order to check that L is a subspace of $\mathbb{R}^3$ we need to check that conditions (i) and (ii) of Definition 0.4.1 are satisfied.

We start with condition (i). Let $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to L. Then $\mathbf{x}_1 = r_1(1, 1, 1)^T$ and $\mathbf{x}_2 = r_2(1, 1, 1)^T$ for some real numbers r_1 and r_2, so

$$\mathbf{x}_1 + \mathbf{x}_2 = r_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + r_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (r_1 + r_2) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in L\,.$$

Thus condition (i) holds.

We now check condition (ii). Let $\mathbf{x} \in L$ and let α be a real number. Then $\mathbf{x} = r(1, 1, 1)^T$ for some real number $r \in \mathbb{R}$, so

$$\alpha\mathbf{x} = \alpha r \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in L\,.$$

Thus condition (ii) holds.

Let's summarise: the non-empty set L satisfies conditions (i) and (ii), that is, it is closed under addition and scalar multiplication, hence L is a subspace of $\mathbb{R}^3$ as claimed.

(b) In order to see that P is a subspace of $\mathbb{R}^3$ we first note that $(0, 0, 0)^T \in P$, so P is not empty.

Next we check condition (i). Let $\mathbf{x}_1 = (r_1, s_1, t_1)^T \in P$ and $\mathbf{x}_2 = (r_2, s_2, t_2)^T \in P$. Then $r_1 - s_1 + 3t_1 = 0$ and $r_2 - s_2 + 3t_2 = 0$, so

$$\mathbf{x}_1 + \mathbf{x}_2 = \begin{pmatrix} r_1 + r_2 \\ s_1 + s_2 \\ t_1 + t_2 \end{pmatrix} \in P\,,$$

[2] In order to save paper, hence trees and thus do our bit to prevent climate change, we shall sometimes write n-vectors $\mathbf{x} \in \mathbb{R}^n$ in the form $(x_1, \ldots, x_n)^T$. So, for example,

$$(2, 3, 1)^T = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}\,.$$

since $(r_1+r_2)-(s_1+s_2)+3(t_1+t_2) = (r_1-s_1+3t_1)+(r_2-s_2+3t_2) = 0+0 = 0$. Thus condition (i) holds.

We now check condition (ii). Let $\mathbf{x} = (r, s, t)^T \in P$ and let α be a scalar. Then $r - s + 3t = 0$ and

$$\alpha\mathbf{x} = \begin{pmatrix} \alpha r \\ \alpha s \\ \alpha t \end{pmatrix} \in P$$

since $\alpha r - \alpha s + 3\alpha t = \alpha(r - s + 3t) = 0$. Thus condition (ii) holds as well.

As P is closed under addition and scalar multiplication, P is a subspace of $\mathbb{R}^3$ as claimed. □

Remark 0.4.5. In the example above the two subspaces L and P of $\mathbb{R}^3$ can also be thought of as geometric objects. More precisely, L can be interpreted geometrically as a line through the origin with direction vector $(1, 1, 1)^T$, while P can be interpreted as a plane through the origin with normal vector $(1, -1, 3)^T$.

More generally, all proper subspaces of $\mathbb{R}^3$ can be interpreted geometrically as either lines or planes through the origin. Similarly, all proper subspaces of $\mathbb{R}^2$ can be interpreted geometrically as lines through the origin.

Example 0.4.6. Let $H = \{\, \mathbf{f} \in C[-2, 2] \mid \mathbf{f}(1) = 0 \,\}$. Then H is a subspace of $C[-2, 2]$. First observe that the zero function is in H, so H is not empty. Next we check that the closure properties are satisfied.

Let $\mathbf{f}, \mathbf{g} \in H$. Then $\mathbf{f}(1) = 0$ and $\mathbf{g}(1) = 0$, so

$$(\mathbf{f} + \mathbf{g})(1) = \mathbf{f}(1) + \mathbf{g}(1) = 0 + 0 = 0\,,$$

so $\mathbf{f} + \mathbf{g} \in H$. Thus H is closed under addition.

Let $\mathbf{f} \in H$ and α be a a real number. Then $\mathbf{f}(1) = 0$ and

$$(\alpha\mathbf{f})(1) = \alpha\mathbf{f}(1) = \alpha \cdot 0 = 0\,,$$

so $\alpha\mathbf{f} \in H$. Thus H is closed under scalar multiplication.

Since H is closed under addition and scalar multiplication it is a subspace of $C[-2, 2]$ as claimed.

Definition 0.4.7. Let $A \in \mathbb{R}^{m\times n}$. Then

$$N(A) = \{\, \mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0} \,\}$$

is called the *nullspace* of A.

Theorem 0.4.8. *If $A \in \mathbb{R}^{m\times n}$, then $N(A)$ is a subspace of $\mathbb{R}^n$.*

Proof. Clearly $\mathbf{0} \in N(A)$, so $N(A)$ is not empty.

If $\mathbf{x}, \mathbf{y} \in N(A)$ then $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$, so

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}\,,$$

and hence $\mathbf{x}+\mathbf{y} \in N(A)$.

Furthermore, if $\mathbf{x} \in N(A)$ and α is a real number then $A\mathbf{x} = \mathbf{0}$ and

$$A(\alpha\mathbf{x}) = \alpha(A\mathbf{x}) = \alpha\mathbf{0} = \mathbf{0}\,,$$

so $\alpha\mathbf{x} \in N(A)$.

Thus $N(A)$ is a subspace of $\mathbb{R}^n$ as claimed. $\square$

Example 0.4.9. Determine $N(A)$ for

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

Solution. We need to find the solution set of $A\mathbf{x} = \mathbf{0}$. To do this you can use your favourite method to solve linear systems. Perhaps the fastest one is to bring the augmented matrix $(A|\mathbf{0})$ to reduced row echelon form and write the leading variables in terms of the free variables. In our case, we have

$$\left(\begin{array}{ccccc|c} -3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{array}\right) \sim \cdots \sim \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The leading variables are x_1 and x_3, and the free variables are x_2, x_4 and x_5. Now setting $x_2 = \alpha$, $x_4 = \beta$ and $x_5 = \gamma$ we find $x_3 = -2x_4 + 2x_5 = -2\beta + 2\gamma$ and $x_1 = 2x_2 + x_4 - 3x_5 = 2\alpha + \beta - 3\gamma$. Thus

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2\alpha + \beta - 3\gamma \\ \alpha \\ -2\beta + 2\gamma \\ \beta \\ \gamma \end{pmatrix} = \alpha \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix},$$

hence

$$N(A) = \left\{ \alpha \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \middle| \alpha, \beta, \gamma \in \mathbb{R} \right\}.$$

□

0.5 Direct sum of subspaces

Let V be a real or complex vector space. Let W_1 and W_2 be subspaces of V. We say that V is a *direct sum* of W_1 and W_2 if

$$V = W_1 + W_2 \quad \text{and} \quad W_1 \cap W_2 = \{\mathbf{0}\}.$$

We denote this by $V = W_1 \oplus W_2$. The first condition above implies that each vector $v \in V$ can be expressed as a sum of a vector w_1 in W_1 and a vector $w_2 \in W_2$. However, the second condition above implies that there is **only one way** of writing v as a sum $w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$. Indeed, if $v = v_1 + v_2$ with $v_1 \in W_1$ and $v_2 \in W_2$, then we have $w_1 - v_1 = v_2 - w_2 \in W_1 \cap W_2 = \{\mathbf{0}\}$ which implies

$$w_1 = v_1 \quad \text{and} \quad w_2 = v_2.$$

Conversely, if each $v \in V$ can be written uniquely as a sum $w_1 = w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$, we must have $W_1 \cap W_2 = \{\mathbf{0}\}$.