

2.7 Geometric Interpretation

As you have seen in the introductory module vectors in $\mathbb{R}^2/\mathbb{R}^3$ can be represented as points in 2 or 3 dimensional space. In this section we will look geometric interpretations of some of the things we have seen so far.

A system of linear equations in n unknowns specifies a set in n -space.

Example 2.7.1.

Consider:

$$\begin{aligned}x_1 + x_2 + x_3 &= -1 \\2x_1 + x_3 &= 1 \\3x_1 + x_2 &= -4\end{aligned}$$

Using row reduction we get $x_1 = -0.5$, $x_2 = -2.5$, $x_3 = 2$, which specifies a point. Whereas:

$$\begin{aligned}x_1 + x_2 + x_3 &= -1 \\2x_1 + x_3 &= 1\end{aligned}$$

Using row reduction we get $x_1 = -2.5 - 0.5x_3$ and $x_2 = 1.5 - 0.5x_3$ giving the line

$$\begin{pmatrix} -2.5 \\ 1.5 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -0.5 \\ -0.5 \\ 1 \end{pmatrix} \text{ for } \lambda \in \mathbb{R}$$

We have seen that we can apply matrices to vectors via matrix multiplication. So we can see a matrix $A \in M_{m \times n}(\mathbb{R})$ as a map:

$$\begin{aligned}A : \quad \mathbb{R}^n &\mapsto \mathbb{R}^m \\A(v) &= Av\end{aligned}$$

We can represent many different operations using matrices.

Example 2.7.2.

Consider $A = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$

Then $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x \\ 5y \end{pmatrix}$.

This is a stretch by a factor of 5.

Definition 2.7.3. Let T be a function from \mathbb{R}^n to \mathbb{R}^m then we say T is a *linear transformation* if for every $v_1, v_2 \in \mathbb{R}^n$ and every $\alpha, \beta \in \mathbb{R}$ we have:

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$$

Proposition 2.7.4. Let $A \in M_{n \times m}(\mathbb{R})$ then seen as a map from \mathbb{R}^n to \mathbb{R}^m A is a linear transformation.

Proof:

$$\begin{aligned} A(\alpha v_1 + \beta v_2) &= A(\alpha v_1) + A(\beta v_2) \quad \text{by distributivity of matrix multiplication} \\ &= \alpha A(v_1) + \beta A(v_2) \quad \text{by exercise} \end{aligned}$$

Proposition 2.7.5. Let $A \in M_{n \times n}(\mathbb{R})$. The following are equivalent:

- (i) A is invertible with inverse $A^{-1} = A^T$
- (ii) $A^T A = I_n = A A^T$.
- (iii) A preserves inner products (i.e. for all $x, y \in \mathbb{R}^n$ $(Px) \cdot (Py) = x \cdot y$).

Proof:

(i) \Leftrightarrow (ii) is just by definition.

(ii) \Leftrightarrow (iii) First note that for $x, y \in \mathbb{R}^n$ $x \cdot y$ as defined in the intro to maths course is just $x^T y$ as matrix multiplication. So A preserves inner products if and only if:

$$\begin{aligned} &(Px) \cdot (Py) = x \cdot y \quad \forall x, y \in \mathbb{R}^n \\ \Leftrightarrow &(Px)^T (Py) = x^T y \quad \forall x, y \in \mathbb{R}^n \\ \Leftrightarrow &x^T P^T P y = x^T I_n y \quad \forall x, y \in \mathbb{R}^n \\ \Leftrightarrow &x^T (P^T P - I_n) y = 0 \quad \forall x, y \in \mathbb{R}^n \end{aligned}$$

(ii) \Rightarrow (iii) now trivial.

(iii) \Rightarrow (ii) let $x_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ i.e. column vector with 0's everywhere except the i^{th} row where there is a 1. Then we know for each x_i $(x_i)^T (P^T P - I_n) y = 0$ so we can conclude that

$$(P^T P - I_n) y = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Similarly taking y_i to be the column vector with 0's everywhere except the i^{th} row where there is a 1 we get $(P^T P - I_n) = 0$ so $P^T P = I_n$.

Definition 2.7.6. A matrix $A \in M_{n \times n}$ is called *Orthogonal* if it is such that $A^{-1} = A^T$

Example 2.7.7.

1. Consider the matrix

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

This matrix is orthogonal as $A^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

If we apply it to $\begin{pmatrix} x \\ y \end{pmatrix}$ we get $\begin{pmatrix} -y \\ x \end{pmatrix}$. This is a rotation through $\frac{\pi}{2}$ radians anti clockwise.

2. Consider the matrix

$$B = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

This matrix is orthogonal as $A^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

If we apply it to $\begin{pmatrix} x \\ y \end{pmatrix}$ we get $\begin{pmatrix} -y \\ -x \end{pmatrix}$. This is a reflection through the line $y = -x$.

Exercise 2.7.8. Watch the linear Algebra video to help you.

1. Let R_θ be the anticlockwise rotation of \mathbb{R}^2 about the origin through θ radians. Using any school geometry or trigonometry you like, find the matrix representing R_θ .

Assuming R_θ is linear (see lectures!) the vector $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is rotated to $\begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ while $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is rotated to $\begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$ so the matrix is

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

2.8 Fields

So far, for both matrices and linear equations, we have only been using entries in \mathbb{R} . However, we could have taken entries from any field.

Every field has distinguished elements 0 (additive identity) and 1 (multiplicative identity).

Fact 2.8.1. Over any field F we can define:

1. The null matrix (i.e. the additive identity matrix) for $M_{n \times m}(F)$ as

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}$$

2. The (multiplicative) identity matrix for $M_{n \times n}(F)$ as

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

Remark 2.8.2 It is important to know what field we are working in, and that we don't say take scalars from a different field to the one matrix entries are from. (e.g. the set of matrices $M_{n \times m}(\mathbb{Q})$ is not closed under scalar multiplication by elements from \mathbb{R}).

Being able to work over a general field allows us to use finite fields.

Theorem 2.8.3. Let $\mathbb{F}_p = \{0, 1, \dots, p-1\}$, consider \mathbb{F}_p with addition defined by addition modulo p and multiplication as multiplication modulo p . Then the structure $(\mathbb{F}_p, +_{(\text{mod } p)}, \times_{(\text{mod } p)})$ is a field.

Proof:

A1-4 (Additive (commutative) group) obvious from properties of addition in \mathbb{Z} .

M1-3 (multiplicative semigroup with 1) obvious from properties of addition in \mathbb{Z} .

M4: inverses: obviously for $0 < x < p$ we have $\gcd(x, p) = 1$ by Intro to Uni Maths we have: $\exists s, t \in \mathbb{Z}$ such that $1 = sx + tp$ then take $x^{-1} = s \pmod{p}$.

D1 (distributive law) obvious from properties of addition in \mathbb{Z} .

Example 2.8.4. \mathbb{F}_6 defined as above is not a field. For example $3 \neq 0$ does not have an inverse.

3 Vector Spaces

3.1 Intro to Vector Spaces

Definition 3.1.1. Let F be a field. A **vector space** over F is a non-empty set V together with the following maps:

1. **Addition**

$$\begin{aligned}\oplus : V \times V &\mapsto V \\ (v_1, v_2) &\mapsto v_1 \oplus v_2\end{aligned}$$

2. **Scalar Multiplication**

$$\begin{aligned}\odot : F \times V &\mapsto V \\ (f, v_2) &\mapsto f \odot v_2\end{aligned}$$

\oplus and \odot must satisfy the following *Vector Space axioms*:

For Vector Addition:

A1 Associative law: $(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w})$.

A2 Commutative law: $\mathbf{v} \oplus \mathbf{w} = \mathbf{w} \oplus \mathbf{v}$.

A3 Additive identity: $0_V \oplus \mathbf{v} = \mathbf{v}$, where 0_V is called the **IDENTITY** vector (or sometimes the *zero vector*).

A4 Additive inverse: $\mathbf{v} \oplus (\ominus \mathbf{v}) = 0_V$.

For scalar multiplication:

A5 Distributive law: $r \odot (\mathbf{v} \oplus \mathbf{w}) = (r \odot \mathbf{v}) \oplus (r \odot \mathbf{w})$.

A6 Distributive law: $(r + s) \odot \mathbf{v} = (r \odot \mathbf{v}) \oplus (s \odot \mathbf{v})$.

A7 Associative law: $r \odot (s \odot \mathbf{v}) = (rs) \odot \mathbf{v}$.

A8 Identity for scalar mult: $1 \odot \mathbf{v} = \mathbf{v}$.

From now on we will drop the \oplus and \odot , and use $+$ and \cdot the point was to emphasise that these are not the same as field addition and multiplication.

Definition 3.1.2. Let V be a vector space over F we call:

- Elements of V are called *vectors*.
- Elements of F are called *scalars*.
- We sometimes refer to V as an F -vector space.

Example 3.1.3. The following are examples of vector spaces over \mathbb{R} :

- The canonical example is the set of vectors \mathbb{R}^n over \mathbb{R} , where \oplus is normal vector addition and \odot is multiplication by a scalar. The additive inverse of \mathbf{v} is simply $-\mathbf{v}$
- The set M_{mn} of all $m \times n$ matrices. This is because addition of two $m \times n$ matrices produces an $m \times n$ matrix and multiplication of an $m \times n$ matrix by a scalar also produces a $m \times n$ matrix. The zero vector is the zero matrix, and for any matrix A , the matrix $-A$ is the additive inverse. Properties of matrix arithmetic covered in Chapter 1, show that all properties in Definition 3.1.2 required of a vector space are satisfied. We will see this later in the course in detail.
- Define \mathbb{R}^X to be the set of real valued functions on X (i.e. $\mathbb{R}^X := \{f : f \text{ a function, } f : X \rightarrow \mathbb{R}\}$). Then for $f, g \in \mathbb{R}^X$ and $\alpha \in \mathbb{R}$ define:

$$\begin{aligned} f \oplus g : X &\rightarrow \mathbb{R} & (\alpha \odot f) : X &\rightarrow \mathbb{R} \\ (f \oplus g)(x) &= f(x) + g(x) & (\alpha \odot f)(x) &= \alpha(f(x)) \end{aligned}$$

Exercise 3.1.4. Which of the following examples of vector spaces over \mathbb{R} :

1. The set of vectors

$$V = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a, b \in \mathbb{Z} \right\} \text{ with the usual vector addition and multiplication}$$

$$\text{No, because } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in V, \quad , \text{ but } \sqrt{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin V$$

2. The set of vectors:

$$V = \left\{ \begin{pmatrix} a+1 \\ 2 \end{pmatrix} : a \in \mathbb{R} \right\} \text{ with the usual vector addition and multiplication}$$

$$\text{No, because } \begin{pmatrix} 0 \\ 0 \end{pmatrix} \notin V$$

3. $V = \mathbb{R}^2$ with the following addition and scalar multiplication operations:

$$\begin{pmatrix} x \\ y \end{pmatrix} \oplus \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x+a \\ y+b \end{pmatrix} \quad \text{and} \quad r \odot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ ry \end{pmatrix}$$

yes

3.2 Subspaces

Definition 3.2.1. A subset W of a vector space V is a **subspace** of V if

- S1 W is not empty (i.e. $e \in W$)
- S2 for $\mathbf{v}, \mathbf{w} \in W$, then $\mathbf{v} \oplus \mathbf{w} \in W$ *closed under vector addition*
- S3 $\mathbf{v} \in W$ and $r \in \mathbb{R}$, then $r \odot \mathbf{v} \in W$ *closed under scalar multiplication*.

N.B. Sometimes we use the notation $U \leq V$ to mean U is a subspace of V .

Remark 3.2.2 Note that V and the zero subspace, $\mathbf{0}$ are always subspaces of V . Any other subspace of V is called a **proper subspace** of V .

Proposition 3.2.3. Every subspace of an F -vector space V must contain the zero vector.

Proof:

Claim: For an F -vector space V with $0 \in F$ the field additive identity we have $0v = 0_V$ for all $v \in V$. *Proof of claim:* Exercise (note enough to show that $0v$ is the vector space additive identity).

Let $v \in V$ (V is non-empty) then $0_V = 0 \oplus v \in V$ as V is closed under scalar multiplication.

Example 3.2.4. Show that the set $X = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix}; x \in \mathbb{R} \right\}$ is a subspace of \mathbb{R}^2 .

Worked Answer

S1 The vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in X$, therefore X is non-empty.

S2 If $\mathbf{v} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} x_2 \\ 0 \end{pmatrix}$, then

$$\begin{aligned}\mathbf{v} \oplus \mathbf{w} &= \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} x_2 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} \in X\end{aligned}$$

S3 If $\mathbf{v} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ and $r \in \mathbb{R}$, then

$$\begin{aligned}r \odot \mathbf{v} &= r \odot \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} rx_1 \\ 0 \end{pmatrix} \in X\end{aligned}$$

Therefore, $X = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix}; x \in \mathbb{R} \right\}$ is a subspace of \mathbb{R}^2 .

Exercise 3.2.5. All subspaces of a vector space over F are vector spaces over F in their own right.

Theorem 3.2.6. Let U, W be subspaces of V . Then $U \cap W$ is a subspace of V . In general, the intersection of any set of subspaces of a vector space V is a subspace of V .

Proof Let C be a set of subspaces of V and T is their intersection. Then $T \neq \emptyset$ since every subspace of V (and therefore every subspace in C) contains the zero vector, and so does T .

Suppose that $x, y \in T$. Since x and y belong to every subspace W in C , so does $x \oplus y$, and therefore $x \oplus y \in T$.

If $x \in T$, then x belongs to every subspace W in the set C , and so does $r \odot x$ and so $r \odot x \in T$.

Therefore T is a subspace of V .

Example 3.2.7. Note that in general if U and W are subspaces of V , then $U \cup W$ is not a subspace of V . For example, let

$$U = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}, W = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \mathbb{R} \right\} \quad \text{and} \quad V = \mathbb{R}^2.$$

Then

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in U \cup W$$

but

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin U \cup W.$$