Crash Course on Linear Algebra

Contents

Mathematics is a process of staring hard enough with enough perseverance at the fog of muddle and confusion to eventually break through to improved clarity. I'm happy when I can admit, at least to myself, that my thinking is muddled, and I try to overcome the embarrassment that I might reveal ignorance or confusion.

William Thurston

How to use these notes.

Linear algebra is the study of linearity. The main players of this story are the vectors, but instead of studying vectors individually, we study collections of vectors, called **vector spaces**, and maps between vector spaces, called **linear transformations**.

Because linearity is such a simple condition (Section 1.1), linear algebra is one of the most widely applied branches of mathematics.

These notes provide a very, very brief introduction to linear algebra. You'll learn enough to be able to continue the study of the subject on your own.

Each day you will be given a set of problems to solve in class. You should attempt as many problems as you can, and it is ok to not finish all the problems before the next class. However, you should make an honest attempt to read and understand the various definitions and theorems.[1]

[1] These notes are likely to contain ~~a few~~ several ~~typos~~ mistakes. Please do tell if you find any, thanks in advance.

1 Vector spaces

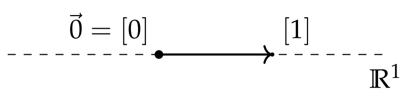
1.1 Motivation

A *scalar* is a real number. A (column) *vector* $\vec{v}$ of size n is a column of n scalars

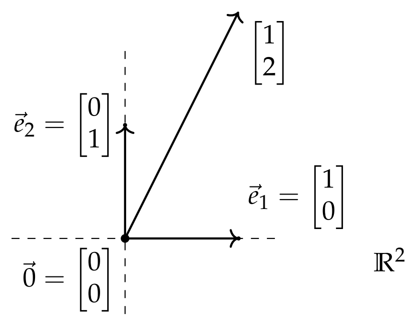
$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The scalars v_i are called the *coordinates* of $\vec{v}$. The set of all (column) vectors is called the *Euclidean space* of dimension n and is denoted by $\mathbb{R}^n$. Denote by $\vec{0}$ the vector with all coordinates 0. $\vec{e}_i$ is the vector whose i^{th} coordinate is 1 and all other coordinates are 0. The vector $\vec{e}_i$ is called the i^{th} *standard basis vector* and the collection $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ is called the *standard basis*.

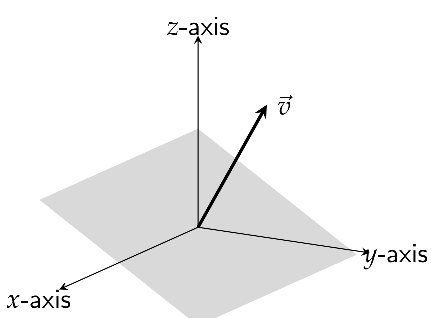
The spaces $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$ are very commonly studied objects, but unfortunately, most of us are unable to visualize the higher dimensions. Linear algebra is the language that lets us do exactly this.



$\mathbb{R}^1$ is the real line,



$\mathbb{R}^2$ is the Euclidean plane,



$\mathbb{R}^3$ is the three dimensional space,

?

What does $\mathbb{R}^n$ look like?

There are two operations we can perform on vectors in $\mathbb{R}^n$:

1. **Scalar multiplication:** Given a scalar c and a vector $\vec{v}$, the vector $c\vec{v}$ is the vector obtained by multiplying every coordinate of $\vec{v}$ by c.

   Geometrically, this is scaling the vector $\vec{v}$ by a factor of c.[1]

2. **Addition:** Given two vectors $\vec{v}$, $\vec{w}$, the sum $\vec{v} + \vec{w}$ is the vector obtained by adding the respective coordinates.

   Geometrically, vector addition is given by the Parallelogram Law.

Q. 1. Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ be vectors in $\mathbb{R}^2$. Verify that

$$\vec{0},\ \vec{v},\ \vec{v} + \vec{w},\ \vec{w}$$

form the vertices of a parallelogram.[2] This is called the *parallelogram law of vector addition*. The parallelogram law holds even for $\mathbb{R}^n$.

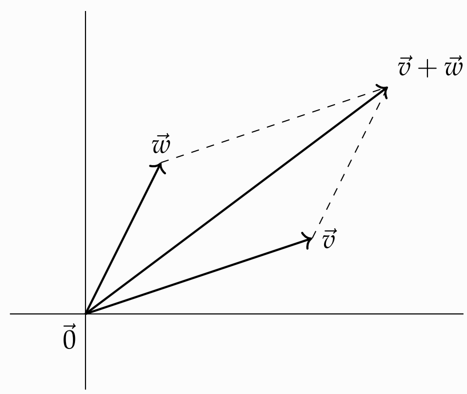


Figure 3: Parallelogram law of vector addition.

1.1.1 Lines and planes in $\mathbb{R}^3$

A line in $\mathbb{R}^2$ is given by an equation $y = mx + c$. What about a line in $\mathbb{R}^3$? One way to decribe a line in $\mathbb{R}^3$ is using vector notation.

Example 1.1. Every line L in $\mathbb{R}^3$ passing through the origin can be written as

$$L = \{c\vec{v} : c \in \mathbb{R}\}$$

[1] If c is negative, we "flip" $\vec{v}$ across the origin.

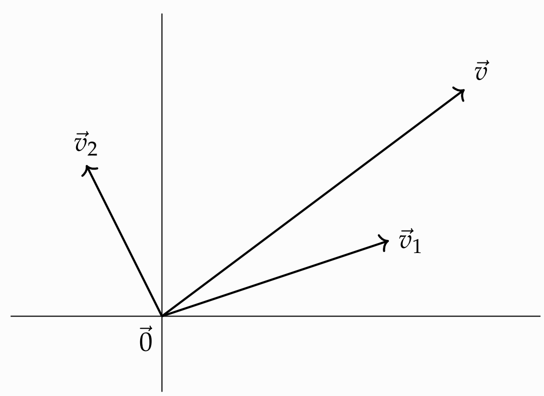
[2] Here we are thinking of vectors as points. It can be a bit confusing at first, to switch between points and arrows, but this flexibility is very useful once you get used to it.

where $\vec{v}$ is a vector in $\mathbb{R}^3$. We can then *define* a line passing through the origin in $\mathbb{R}^n$ as $\{c\vec{v} : c \in \mathbb{R}\}$ for some $\vec{v} \in \mathbb{R}^n$.

What about a plane? We needed one vector to describe a line. We will need two to describe a plane.

Q. 2. Let $\vec{v}_1$ and $\vec{v}_2$ be non-zero vectors in $\mathbb{R}^2$ which are not scalar multiples of each other. Using the Parallelogram Law from Q. 1 , argue that for every vector $\vec{v}$ in $\mathbb{R}^2$ there exist scalars c_1, c_2 such that

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2.$$



So that we can write $\mathbb{R}^2$ as

$$\mathbb{R}^2 = \{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$$

Q. 3. What is the set $\{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$, if $\vec{v}_1$ is a scalar multiple of $\vec{v}_2$, for $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$.

One can similarly check that if $\vec{v}_1$ and $\vec{v}_2$ are non-zero vectors in $\mathbb{R}^3$ which are not scalar multiples of each other, the set $\{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$ describes planes passing through the origin. We then *define* a plane in $\mathbb{R}^n$ passing through the origin, to be a set of the form $\{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$.

Q. 4. How would you define lines and planes in $\mathbb{R}^n$ not necessarily passing through the origin (using vector notation)?

Our goal then is to generalize this entire story to higher dimensions, by defining "*m dimensional subspaces of* $\mathbb{R}^n$". We'll first define a subspace, then slowly move toward defining dimension[1].

1.2 Subspaces

Definition 1.2. A *subspace* of $\mathbb{R}^n$ is a *non-empty* subset $V \subseteq \mathbb{R}^n$ satisfying the following conditions.

1. (closed under scalar multiplication) For every real number c and vector $\vec{v}$ in V, the vector $c\vec{v}$ is in V.
2. (closed under addition) For every $\vec{v}$ and $\vec{w}$ in V, the vector $\vec{v} + \vec{w}$ is in V.

Subspaces, and more generally abstract vector spaces, are the primary objects of study in linear algebra.

Q. 5.
1. Show that $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$.
2. Show that the set $V = \{\vec{0}\}$ is a subspace of $\mathbb{R}^n$.
3. For any subspace $V \subseteq \mathbb{R}^n$, show that the vector $\vec{0}$ is in V.

Q. 6. Determine, with proof, all the subspaces of $\mathbb{R}^1$.

We already know two subspaces of $\mathbb{R}^2$: $\{0\}$ and $\mathbb{R}^2$. $\mathbb{R}^2$ has one other family of subspaces given by lines.

Q. 7. When is the line in $\mathbb{R}^2$ a subspace of $\mathbb{R}^2$? What about $\mathbb{R}^3$?

Q. 8. When is a plane in $\mathbb{R}^3$ a subspace of $\mathbb{R}^3$?

Q. 9. Make a guess as to what all the subspaces of $\mathbb{R}^2$ and $\mathbb{R}^3$ are. How would you prove this?[2]

Q. 10. Let V and W be subspaces of $\mathbb{R}^n$.
1. Show that the intersection $V \cap W$ is also a subspace.
2. What can you say about the union $V \cup W$?

[1] Which is a surprisingly difficult thing to define!
[2] We will give a rigorous proof of this tomorrow.

1.3 Span, Linear Independence, and Basis

Span is a general procedure for constructing subspaces of $\mathbb{R}^n$. It generalizes the concepts in Example 1.1 and Q. 2.

Definition 1.3. For a subset S of $\mathbb{R}^n$, the span of S is defined to be the set of finite *linear combinations* of elements of S.

$$\begin{aligned}\text{Span}(S) := \{c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \quad : \quad & c_1, \ldots, c_k \text{ are real numbers,} \\ & \vec{v}_1, \ldots, \vec{v}_k \text{ are vectors in } S\}\end{aligned}$$

For the empty set $\varnothing \subseteq \mathbb{R}^n$ we define $\text{Span}(\varnothing) := \{\vec{0}\}$. We think of S as being a *generating set* of the subspace $\text{Span}(S)$.

Q. 11. Let S be a subset of $\mathbb{R}^n$. Show that $\text{Span}(S)$ is a subspace of $\mathbb{R}^n$.

Q. 12 (Practice problems)**.** Find simple descriptions of Spans of the following subsets of $\mathbb{R}^3$.

1. $\{\vec{e}_1\}$
2. $\{\vec{e}_1, \vec{e}_2\}$
3. $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$
4. $\mathbb{R}^3$
5. $\{\vec{e}_1, \vec{e}_1 + \vec{e}_2, \vec{e}_3\}$
6. $\{\vec{e}_1 - \vec{e}_2, \vec{e}_2 - \vec{e}_3, \vec{e}_3 - \vec{e}_1\}$
7. The plane $z = 0$
8. The plane $z = 1$

Q. 13. Let S be a subset of $\mathbb{R}^n$ and V be a subspace of $\mathbb{R}^n$. Show that

$$\text{if } S \subseteq V \text{ then } \text{Span}(S) \subseteq V.$$

Thus $\text{Span}(S)$ is the *smallest vector space* containing S.

The notion of Span allows for redundancies. For example, the set $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ has 3 elements and the set $\mathbb{R}^3$ has infinitely many, but both of these sets have the same Span. This redundancy is precisely captured by *linear dependence*.

Definition 1.4. A finite set of vectors $\mathcal{S} = \{\vec{v}_1, \ldots, \vec{v}_k\} \subseteq \mathbb{R}^n$ is *linearly dependent* if there exist real numbers $c_1, \ldots, c_k$, not all 0, satisfying

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = 0.$$

$\mathcal{S}$ is called *linearly independent* otherwise i.e. $\mathcal{S}$ is *linearly independent* if the only real numbers $c_1, \ldots, c_k$ for which

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = 0.$$

are $c_1 = \cdots = c_k = 0$.

An infinite set $\mathcal{S} \subseteq \mathbb{R}^n$ is said to be *linearly dependent* if it contains a finite subset which is linearly dependent, it is called *linearly independent* otherwise.

Q. 14. Let $\vec{v}_1, \vec{v}_2$ be vectors in $\mathbb{R}^n$.

1. When is the set $\mathcal{S} = \{\vec{v}_1\}$ linearly independent?
2. When is the set $\mathcal{S}' = \{\vec{v}_1, \vec{v}_2\}$ linearly independent?

Q. 15. ~~Is the empty set $\varnothing \subset \mathbb{R}^n$ linearly dependent or independent?~~

Q. 16.

1. Show that a set $\mathcal{S}$ is linearly independent if and only if every subset of $\mathcal{S}$ is linearly independent.
2. Is the statement still true if we replace linear independence with linear dependence?

Definition 1.5. For a subspace V of $\mathbb{R}^n$, a set $\mathcal{B} \subseteq V$ is said to be a *basis* of V if

1. $\mathrm{Span}(\mathcal{B}) = V$,
2. $\mathcal{B}$ is linearly independent.

Q. 17. For each of the sets $\mathcal{S}$ in Q. 12, find a basis of $\mathrm{Span}(\mathcal{S})$.

The following theorem is what makes a basis extremely useful.

Theorem 1.6. *Let V be a subspace of $\mathbb{R}^n$ with a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$. For every vector $\vec{v}$ in V there exist unique scalars $c_1, \ldots, c_k$ such that*

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k.$$

Q. 18. Prove Theorem 1.6. [1]

Definition 1.7. Let V be a subspace of $\mathbb{R}^n$ with a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$. If $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ as in Theorem 1.6, we denote

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}.$$

[1] **Hint:** First show that the scalars c_i exist. Then argue that they are unique using a proof by contradiction.

The c_i's are the *coordinates* of $\vec{v}$ in the basis $\mathcal{B}$.

Thus if the subspace V has a basis $\mathcal{B}$ of size k then V "behaves" like the Euclidean space $\mathbb{R}^k$.

Example 1.8. If $\vec{v}$ is a vector in $\mathbb{R}^n$ and $\mathcal{B}$ is the standard basis, then $[\vec{v}]_{\mathcal{B}} = \vec{v}$, hence the name *standard basis*. The standard basis is a special basis that exists only for $\mathbb{R}^n$, for other vector spaces there is usually no natural choice of a basis. But we will show that *a* basis always exists.

Q. 19. Let V be a subspace of $\mathbb{R}^n$ with a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$. Let $\vec{v}$, $\vec{w}$ be vectors in V and let α be a scalar. Show that

1. $[\alpha\vec{v}]_{\mathcal{B}} = \alpha[\vec{v}]_{\mathcal{B}}$
2. $[\vec{v} + \vec{w}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}} + [\vec{w}]_{\mathcal{B}}$
3. $\vec{v} = \vec{w}$ if and only if $[\vec{v}]_{\mathcal{B}} = [\vec{w}]_{\mathcal{B}}$.

Q. 20 (Practice problems)**.** Find $[\vec{v}]_{\mathcal{B}}$ of the vector $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \in V$ for V and $\mathcal{B}$ as below.

1. $V = \mathbb{R}^3$, $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$
2. $V = \mathbb{R}^3$, $\mathcal{B} = \{\vec{e}_1, \vec{e}_3, \vec{e}_2\}$
3. $V = \mathbb{R}^3$, $\mathcal{B} = \{\vec{e}_1 + \vec{e}_2, \vec{e}_2, \vec{e}_3\}$
4. $V = \{z = 0\}$, $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$
5. $V = \{x + y + z = 0\}$, $\mathcal{B} = \{\vec{e}_1 - \vec{e}_3, \vec{e}_2 - \vec{e}_3\}$
6. $V = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right\}$, $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}\right\}$

Generalized coordinates are *intrinsic* to the subspace and do not depend on the ambient vector space. However, they are not unique and depend upon the choice of a basis.

We will show in the next section that the *number* of generalized coordinates, on the other hand, does not depend on the basis. The dimension is then defined to equal this number.

1.4 Optional: Abstract vector spaces

We repeatedly used scalar multiplication and addition of vectors in the previous sections. Everything we did naturally carries over to any structure that has scalar multiplication and addition.

Definition 1.9. A *vector space* over the real numbers is any set V which has addition and scalar multiplication.

Example 1.10. A subspace of $\mathbb{R}^n$ is an example of a vector space.

Q. 21. Verify that the following are vector spaces.

1. Set of polynomials in a single variable x.

$$\mathrm{Poly} = \{a_0 + a_1 x + \cdots + a_n x^n \mid n \in \mathbb{Z}_{\geq 0}, a_i \in \mathbb{R}\}$$

2. Set of polynomials Poly_n in a single variable x of degree $\leq n$.

$$\mathrm{Poly}_n = \left\{a_0 + \cdots + a_k x^k \mid k \leq n, a_i \in \mathbb{R}\right\}.$$

3. Set of functions $f : \mathbb{R} \to \mathbb{R}$.

$$\mathrm{Maps}(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}\}$$

4. Set of functions $f : \mathbb{R} \to \mathbb{R}$ with $f(0) = 0$.

$$\mathrm{Maps}_0(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}\}$$

Q. 22. The set of polynomials in a single variable x of degree $= n$ is *not* a vector space. Why?

The notion of a subspace naturally extends to abstract vector spaces: Poly_n is a subspace of Poly and $\mathrm{Maps}_0(\mathbb{R}, \mathbb{R})$ is a subspace of $\mathrm{Maps}(\mathbb{R}, \mathbb{R})$. The notions of spans, linear independence, and basis also naturally generalize.

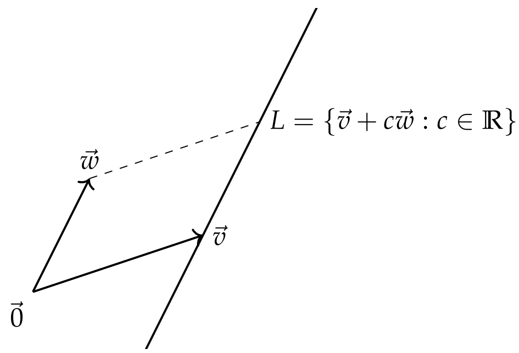
Q. 23. Find bases for Poly_n and Poly.

Q. 24. Can you find a basis for $\mathrm{Maps}(\mathbb{R}, \mathbb{R})$? Are you sure?

Solutions to selected problems - 1

Solution to Q.4. A line L not passing through the origin can be written as

$$L = \{\vec{v} + c\vec{w} : c \in \mathbb{R}\}$$



Similarly, a plane in $\mathbb{R}^3$ that does not pass through the origin is given by

$$\{\vec{v} + c_1\vec{w}_1 + c_2\vec{w}_2\}.$$

$\square$

Solution to Q.6. **Claim:** Every subspace of $\mathbb{R}^1$ is either $\{\vec{0}\}$ or $\mathbb{R}^1$.
Proof: Let V be a subspace of $\mathbb{R}^1$. We have already shown that V contains the vector $\vec{0}$. We will show that if $V \neq \{\vec{0}\}$ then $V = \mathbb{R}^1$.

Suppose V contains a non-zero vector $\vec{v} = [a]$ where $a \neq 0$. Because V is closed under scalar multiplication $c\vec{v} = c[a] = [ca]$ is also in V for every real number c. But c can be any real number, so every vector $[b]$ is in V which implies that $V = \mathbb{R}^1$. $\square$

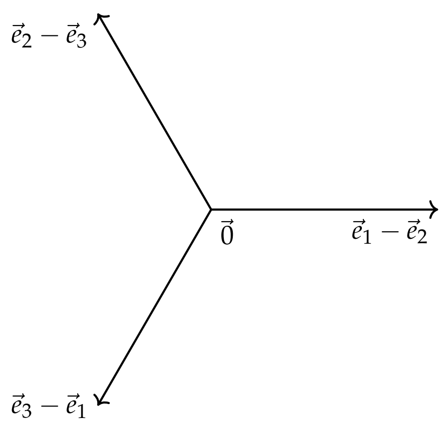
Solution to Q.7. A line in $\mathbb{R}^2$ (or $\mathbb{R}^3$) is a subspace of $\mathbb{R}^2$ (or $\mathbb{R}^3$) if and only if it passes through the origin. $\square$

Solution to Q.8. A plane in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$ if and only if it passes through the origin. $\square$

Solution Q.10 Part 2. The union of two subspaces is not always a subspace. Consider the following subspaces of $\mathbb{R}^2$. Let V be the x-axis and let W be the y-axis. Then $V \cup W$ is the union of x and y axes. But this is not a subspace as it is not closed under addition. $\square$

Solution to Q.12 Part 6. For $\mathcal{S} = \{\vec{e}_1 - \vec{e}_2, \vec{e}_2 - \vec{e}_3, \vec{e}_3 - \vec{e}_1\}$. Notice the following identity:

$$\vec{e}_3 - \vec{e}_1 = -(\vec{e}_1 - \vec{e}_2) - (\vec{e}_2 - \vec{e}_3)$$



So that we can describe a vector in $\mathrm{Span}(\mathcal{S})$ as

$$\begin{aligned}\mathrm{Span}(\mathcal{S}) \ni \vec{v} &= c_1(\vec{e}_1 - \vec{e}_2) + c_2(\vec{e}_2 - \vec{e}_3) + c_3(\vec{e}_3 - \vec{e}_1)\\ &= c_1(\vec{e}_1 - \vec{e}_2) + c_2(\vec{e}_2 - \vec{e}_3) + c_3\left(-(\vec{e}_1 - \vec{e}_2) - (\vec{e}_2 - \vec{e}_3)\right)\\ &= (c_1 - c_3)(\vec{e}_1 - \vec{e}_2) + (c_2 - c_3)(\vec{e}_2 - \vec{e}_3)\\ &= c_1'(\vec{e}_1 - \vec{e}_2) + c_2'(\vec{e}_2 - \vec{e}_3) \in \mathrm{Span}\left((\vec{e}_1 - \vec{e}_2), (\vec{e}_2 - \vec{e}_3)\right).\end{aligned}$$

So $\vec{v}$ is in $\mathrm{Span}\left((\vec{e}_1 - \vec{e}_2), (\vec{e}_2 - \vec{e}_3)\right)$.

Answer: $\mathrm{Span}(\mathcal{S})$ is a plane spanned by the vectors $\vec{e}_1 - \vec{e}_2$ and $\vec{e}_2 - \vec{e}_3$. $\square$

Summary of Section 1

- In higher dimensions,

  1. a line passing through the origin is described as
  $$L = \{c\vec{v} : c \in \mathbb{R}\}$$
  2. a plane passing through the origin is described as
  $$P = \{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$$
  when $\vec{v}_1$ is not a scalar multiple of $\vec{v}_2$.

  We generalize this using the notion of Span. L is the span of the set $\{\vec{v}\}$ and P is the span of the set $\{\vec{v}_1, \vec{v}_2\}$.

- More generally, for any subset $\mathcal{S} \subset \mathbb{R}^n$
$$\begin{aligned}\mathrm{Span}(\mathcal{S}) = \{c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \quad : \quad & c_1, \ldots, c_k \text{ are real numbers,} \\ & \vec{v}_1, \ldots, \vec{v}_k \text{ are vectors in } S\}\end{aligned}$$
  defines a subspace of $\mathbb{R}^n$. A subspace is a subset of $\mathbb{R}^n$ which is closed under scalar multiplication and addition.

  However, there might be reduncacies in $\mathcal{S}$. For example,
$$\begin{aligned}\mathrm{Span}(\vec{v}_1) &= \mathrm{Span}(\vec{v}_1, \vec{v}_2) && \text{if } \vec{v}_1 \text{ is a scalar multiple of } \vec{v}_2 \\ \mathrm{Span}(\vec{e}_1 - \vec{e}_2, \vec{e}_2 - \vec{e}_3, \vec{e}_3 - \vec{e}_1) &= \mathrm{Span}(\vec{e}_1 - \vec{e}_2, \vec{e}_2 - \vec{e}_3)\end{aligned}$$

- This is "corrected" by linear independence (Definition 1.4). For a subspace V of $\mathbb{R}^n$, a set $\mathcal{B} \subseteq V$ is said to be a *basis* of V if $\mathrm{Span}(\mathcal{B}) = V$ and $\mathcal{B}$ is linearly independent.

  In the above examples, L has basis $\mathcal{B} = \{\vec{v}\}$ and P has basis $\{\vec{v}_1, \vec{v}_2\}$.

- A basis allows us to define coordinates for a vector space.

  1. For L as above, with $\mathcal{B} = \{\vec{v}\}$
  $$[c\vec{v}]_{\mathcal{B}} = [c]$$
  2. For P as above, with $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$
  $$[c_1\vec{v}_1 + c_2\vec{v}_2]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$
  3. More generally, if V has basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ then
  $$[c_1\vec{v}_1 + \cdots + c_k\vec{v}_k]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$$

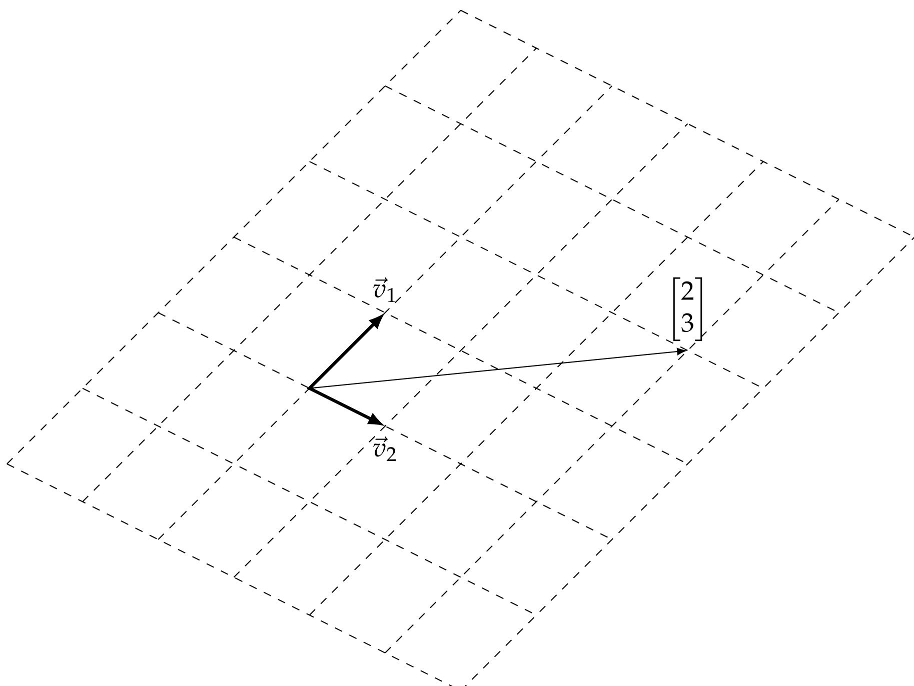


Figure 4: Generalized coordinates on a plane. In the basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$, the vector $\vec{v} = 2\vec{v}_1 + 3\vec{v}_2$ has coordinates $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix}2\\3\end{bmatrix}$.

2 Dimension

Q. 25. Consider the line $L = \{c\vec{v} : c \in \mathbb{R}\}$ where $\vec{v}$ is a non-zero vector in $\mathbb{R}^n$. One basis for L is $\mathcal{B} = \{\vec{v}\}$.

1. Argue that every basis of L has exactly one element.

Q. 26. Consider the plane $P = \{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$ where $\vec{v}_1, \vec{v}_2$ are non-zero vectors in $\mathbb{R}^n$ which are not linear multiples of each other. One basis for P is $\{\vec{v}_1, \vec{v}_2\}$.

1. Argue that every basis of P has exactly two elements.

This generalizes to arbitrary vector spaces.

Theorem 2.1. *If a subspace V of $\mathbb{R}^n$ has a basis $\mathcal{B}$ of size k then every basis $\mathcal{B}'$ of V has size k.*

This number is then defined to be the dimension of the subspace V.

The proof of Theorem 2.1 is a bit long and complicated. If you are seeing this for the first time, you can assume this theorem without proof, so the the sections below are optional. You should at least skim through them to get an idea of how the proof goes.

We start by analysing arbitrary linearly independent sets inside a vector space.

Proposition 2.2. *Let V be a subspace of $\mathbb{R}^n$ with a finite basis $\mathcal{B}$. Let $\mathcal{S}$ be an arbitrary linearly independent subset of V. Then $|\mathcal{S}| \leq |\mathcal{B}|$.*

Proof. Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be a basis of V and let $\mathcal{S}$ be a linearly independent subset of V. We want to show that $\ell \leq k$. We will instead prove the contrapositive: *If $\ell > k$ then there exist scalars $c_1, \ldots, c_\ell$, not all 0, such that*

$$c_1\vec{w}_1 + \cdots + c_\ell\vec{w}_\ell = 0. \tag{2.1}$$

In the basis $\mathcal{B}$, we can express the vectors $\vec{w}_i$'s as coordinate vectors.

$$[w_i]_{\mathcal{B}} = \begin{bmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{ki} \end{bmatrix}$$

Equation (2.1) then simplifies as

$$c_1\vec{w}_1 + \cdots + c_\ell\vec{w}_\ell = 0$$

$$
\begin{aligned}
&\implies c_1[\vec{w}_1]_{\mathcal{B}} + \cdots + c_\ell[\vec{w}_\ell]_{\mathcal{B}} = 0 \\
&\implies c_1 \begin{bmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{k1} \end{bmatrix} + \cdots + c_\ell \begin{bmatrix} A_{1\ell} \\ A_{2\ell} \\ \vdots \\ A_{k\ell} \end{bmatrix} = 0 \\
&\implies \begin{bmatrix} A_{11}c_1 \\ A_{21}c_1 \\ \vdots \\ A_{k1}c_1 \end{bmatrix} + \cdots + \begin{bmatrix} A_{1\ell}c_\ell \\ A_{2\ell}c_\ell \\ \vdots \\ A_{k\ell}c_\ell \end{bmatrix} = 0 \\
&\implies \begin{bmatrix} A_{11}c_1 + \cdots + A_{1\ell}c_\ell \\ A_{21}c_1 + \cdots + A_{2\ell}c_\ell \\ \vdots \\ A_{k1}c_1 + \cdots + A_{k\ell}c_\ell \end{bmatrix} = 0
\end{aligned}
$$

$$
\begin{aligned}
\implies A_{11}c_1 + \cdots + A_{1\ell}c_\ell &= 0 \\
A_{21}c_1 + \cdots + A_{2\ell}c_\ell &= 0 \\
&\vdots \\
A_{k1}c_1 + \cdots + A_{k\ell}c_\ell &= 0
\end{aligned}
$$

These are k equations in ℓ variables and $\ell > k$, so we have more variables than equations. Such a system is called an *under-determined* linear system. Because the right-hand is zero, such a system always has a solution (Lemma 2.8) with not all c_i equal to 0. □

Because every basis is also a linearly independent set, we can use Proposition 2.2 to say several things about bases.

Corollary 2.3. *Let V be a subspace of $\mathbb{R}^n$.*

1. *If $\mathcal{B}$ is a basis of V then $|\mathcal{B}| \leq n$.*
2. *If $\mathcal{B}$ and $\mathcal{B}'$ are two bases of V, then $|\mathcal{B}| = |\mathcal{B}'|$.*
3. *Every basis of $\mathbb{R}^n$ has n elements.*

Q. 27. Prove the above corollaries using Proposition 2.2.

Notice that the above theorems compare the sizes of linearly independent sets to the size of a basis. But we have not shown that a basis exists in the first place.

Theorem 2.4. *Let V be a subspace of $\mathbb{R}^n$. Then V has a basis.*

Proof. We provide an algorithm for constructing a basis $\mathcal{B}$.

Set $i = 0$ and let $S_0 = \varnothing$.

1. If $\operatorname{Span}(S_i) = V$ we are done.
2. If $\operatorname{Span}(S_i) \neq V$ then there exists a vector $\vec{v}_i$ such that $\vec{v}_i \in V \setminus \operatorname{Span}(S_i)$.
3. Let $S_{i+1} = S_i \cup \{\vec{v}_i\}$ and increment i by 1. Go back to Step 1.

Once this process terminates the final set S_i that we obtain is a basis for V.

Q. 28.
1. Show that the sets S_i constructed above are all linearly independent.
2. Using Proposition 2.2 argue that the above algorithm always terminates?
3. Prove that the final S_i is a basis for V?

$\square$

This allows to make the following definition.

Definition 2.5. For a subspace V of $\mathbb{R}^n$, the *dimension* of V is defined to be the size of any basis $\mathcal{B}$.

$$\dim V := |\mathcal{B}| \text{ where } \mathcal{B} \text{ is any basis of } V$$

Q. 29. (Optional) Using a similar algorithm it is possible to prove a slightly stronger statement than Theorem 2.4. Let $V \subsetneq V'$ be subspaces of $\mathbb{R}^n$. Let $\mathcal{B}$ be a basis of V. Show that $\mathcal{B}$ can be extended to a basis $\mathcal{B}'$ of V'.

Definition 2.6. If a vector space V has a finite basis $\mathcal{B}$ then we define

$$\dim V = |\mathcal{B}|$$

Otherwise we say that V is infinite dimensional.

Remark 2.7. Our proofs about subspaces of $\mathbb{R}^n$ carry over to arbitrary vector spaces verbatim, except the one about existence of basis. For arbitrary vector spaces, what do you think might go wrong with the proof of Theorem 2.4?[1]

[1]The proof of existence of basis of an arbitrary vector space relies on the Axiom of Choice.

Lemma 2.8 (Under-determined linear system)**.** *Consider the following system with k equations and ℓ variables. (A_{ij} are scalar constants and we are solving for the c_i's.)*

$$\begin{aligned} A_{11}c_1 + \cdots + A_{1\ell}c_\ell &= 0 \\ A_{21}c_1 + \cdots + A_{2\ell}c_\ell &= 0 \\ &\vdots \\ A_{k1}c_1 + \cdots + A_{k\ell}c_\ell &= 0 \end{aligned}$$

If $k < \ell$, then this system of equations always has a solution such that at least one of the c_i's is non-zero.

Proof. Proof is induction on k.

> **Q. 30.** *Base case:* Prove Lemma 2.8 for $k = 1$.

Induction hypothesis: Assume that we know Lemma 2.8 to be true when $k = m$.

Induction step: Consider now $k = m + 1$.

> **Q. 31.** Argue that if $A_{11} = \cdots = A_{1\ell} = 0$ then we are done by the induction hypothesis.

Without any loss of generality, assume that $A_{11} \neq 0$. Then the first equation gives us

$$c_1 = -\frac{c_2 A_{12} + \cdots + c_l A_{1l}}{A_{11}}$$

> **Q. 32.** Plug this in the other equations and conclude the proof using the induction hypothesis.

□

3 Linear transformations

The unknown thing to be known appeared to me as some stretch of earth or hard marl, resisting penetration... the sea advances insensibly in silence, nothing seems to happen, nothing moves, the water is so far off you hardly hear it... yet it finally surrounds the resistant substance.

Alexander Grothendieck

Notation: From now on, we will use the term *vector space* to mean a *subspace of* $\mathbb{R}^n$.[1]

The "category theory" philosophy in mathematics says that in order to understand structured objects one must study structured maps between them. The structured maps between vector spaces are the ones that preserve linearity.

Definition 3.1. A *linear transformation* is a map $\mathcal{L} : V \to W$ between vector spaces V and W satisfying

1. $\mathcal{L}(c\vec{v}) = c\mathcal{L}(\vec{v})$,
2. $\mathcal{L}(\vec{v} + \vec{w}) = \mathcal{L}(\vec{v}) + \mathcal{L}(\vec{w})$.

for all scalars $c \in \mathbb{R}$ and all vectors $\vec{v}, \vec{w} \in V$.

A linear transformation $\mathcal{L} : V \to W$ is a (vector space) *isomorphism* if it is a bijection of sets, in which case we say that the two vector spaces V and W are *isomorphic*.

Q. 33. 1. Show that if $\mathcal{L} : V \to W$ and $\mathcal{L}' : W \to U$ are linear transformations, then so is their composition $\mathcal{L}' \circ \mathcal{L} : V \to U$.

$$V \xrightarrow{\mathcal{L}} W \xrightarrow{\mathcal{L}'} U$$

2. Show that if $\mathcal{L} : V \to W$ is an isomorphism, then the set-theoretic inverse map $\mathcal{L}^{-1} : W \to V$ is also a linear transformation.

$$\begin{tikzcd} V \arrow[r, bend left, "\mathcal{L}"] & W \arrow[l, bend left, "\mathcal{L}^{-1}"] \end{tikzcd}$$

[1] This is because all of our theorems are also true for finite dimensional abstract vector spaces i.e. abstract vector spaces for which a finite basis exists.

3.1 Examples in low dimensions

Definition 3.2. A linear transformation where the source and target is the same, $\mathcal{L} : V \to V$, is called a *linear operator* on V.

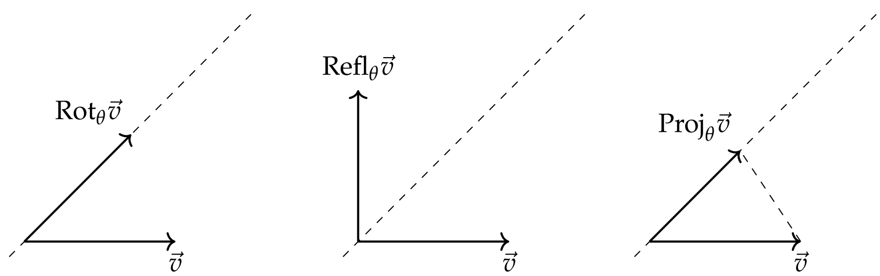
Q. 34. Use the Parallelogram Law to prove that the following maps are linear operators on $\mathbb{R}^2$.

1. $\mathrm{Rot}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, rotation by an angle θ in the counterclockwise direction.

2. $\mathrm{Refl}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, reflection about a line that forms an angle θ with the x-axis.

3. $\mathrm{Proj}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, orthogonal projection onto a line that forms an angle θ with the x-axis.

Which of the above maps are isomorphisms? For the ones that are isomorphisms, what are the inverses?



Q. 35. Is translation along the x-axis, $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+1 \\ y \end{bmatrix}$, a linear operator on $\mathbb{R}^2$?

Q. 36. What maps $\mathcal{L} : \mathbb{R}^1 \to \mathbb{R}^1$ are linear operators on $\mathbb{R}^1$? [1]

Q. 37. Show that the identity map $\mathrm{Id} : V \to V$ is a linear operator on any vector space V.

[1] **Hint:** $\mathcal{L}$ is completely determined by $\mathcal{L}([1])$.

3.2 Linear transformations and matrices

We will see in Section 3.3 that linear transformations are intricately related to bases. In this section, we will use bases and generalized coordinates to derive the formula for matrix multiplication.

Consider a linear transformation between vector spaces

$$\mathcal{L} : V \to W$$

Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ and $\mathcal{B}' = \{\vec{w}_1, \ldots, \vec{w}_\ell\}$ be bases of V and W respectively. For any $\vec{v} \in V$, as $\mathcal{L}(\vec{v})$ is a vector in W, and the set $\{\vec{w}_1, \ldots, \vec{w}_\ell\}$ is a basis of W, we can write $\mathcal{L}(\vec{v})$ as a linear combination of the $\vec{w}_i$. Suppose

$$\begin{aligned}
\mathcal{L}(\vec{v}_1) &= A_{11}\vec{w}_1 + A_{21}\vec{w}_2 + \cdots + A_{\ell 1}\vec{w}_\ell \\
\mathcal{L}(\vec{v}_2) &= A_{12}\vec{w}_1 + A_{22}\vec{w}_2 \cdots + A_{\ell 2}\vec{w}_\ell \\
&\vdots \\
\mathcal{L}(\vec{v}_k) &= A_{1k}\vec{w}_1 + +A_{2k}\vec{w}_2 \cdots + A_{\ell k}\vec{w}_\ell
\end{aligned}$$

where A_{ij} are scalars, for $1 \le i \le \ell$ and $1 \le j \le k$. In generalized coordinates, these can be written as column vectors.

$$[\mathcal{L}(\vec{v}_1)]_{\mathcal{B}'} = \begin{bmatrix} A_{11} \\ \\ \vdots \\ \\ A_{\ell 1} \end{bmatrix} \quad [\mathcal{L}(\vec{v}_2)]_{\mathcal{B}'} = \begin{bmatrix} A_{12} \\ \\ \vdots \\ \\ A_{\ell 2} \end{bmatrix} \quad \cdots \quad [\mathcal{L}(\vec{v}_k)]_{\mathcal{B}'} = \begin{bmatrix} A_{1k} \\ \\ \vdots \\ \\ A_{\ell k} \end{bmatrix}$$

We then place the coordinate vectors $[\,\mathcal{L}\vec{v}_1\,]_{\mathcal{B}'}$, $[\,\mathcal{L}\vec{v}_2\,]_{\mathcal{B}'}$, ..., $[\,\mathcal{L}\vec{v}_k\,]_{\mathcal{B}'}$ next to each other to form a grid with ℓ rows and k columns, called an $\ell \times k$ *matrix* which we will denote by $[\mathcal{L}]_{\mathcal{B} \to \mathcal{B}'}$.

$$[\mathcal{L}]_{\mathcal{B} \to \mathcal{B}'} := \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \\ \vdots & \vdots & \ddots & \vdots \\ \\ A_{\ell 1} & A_{\ell 2} & \cdots & A_{\ell k} \end{bmatrix}.$$

This is the matrix associated to/corresponding to $\mathcal{L}$ for the bases $\mathcal{B}$ and $\mathcal{B}'$.

When $V = \mathbb{R}^k$ and $W = \mathbb{R}^\ell$ and $\mathcal{B}$ and $\mathcal{B}'$ are the standard bases, we drop the basis symbols and denote the matrix by $[\mathcal{L}]$.

"Theorem." *Matrices are the same as linear transformations between Euclidean spaces. More generally, matrices are the same as linear transformations between vector spaces "once we have chosen bases for the source and target".*

Q. 38. Find matrices corresponding to the following linear operators on $\mathbb{R}^2$ in the standard bases.

1. $\mathrm{Id}_{\mathbb{R}^2}$
2. Rot_θ
3. Refl_θ
4. Proj_θ

For a general vector $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ in V by linearity, we have

$$\mathcal{L}(\vec{v}) = c_1\mathcal{L}(\vec{v}_1) + \cdots + c_k\mathcal{L}(\vec{v}_k)$$

Rewriting this equation using the bases and the generalized coordinate vectors we get

$$[\mathcal{L}]_{\mathcal{B}\to\mathcal{B}'}[\vec{v}]_{\mathcal{B}} = c_1[\mathcal{L}(\vec{v}_1)]_{\mathcal{B}'} + \cdots + c_k[\mathcal{L}(\vec{v}_k)]_{\mathcal{B}'}$$

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell 1} & A_{\ell 2} & \cdots & A_{\ell k} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = c_1 \begin{bmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{\ell 1} \end{bmatrix} + c_2 \begin{bmatrix} A_{12} \\ A_{22} \\ \vdots \\ A_{\ell 2} \end{bmatrix} + \cdots + c_k \begin{bmatrix} A_{1k} \\ A_{2k} \\ \vdots \\ A_{\ell k} \end{bmatrix}$$

And, we have derived the formula for matrix multiplication!!!?!?!???!!

Q. 39. Compute the following using the formula for matrix multiplication in the standard basis.

1. $[\mathrm{Id}_{\mathbb{R}^2}] \begin{bmatrix} x \\ y \end{bmatrix}$
2. $[\mathrm{Rot}_\theta] \begin{bmatrix} x \\ y \end{bmatrix}$
3. $[\mathrm{Refl}_\theta] \begin{bmatrix} x \\ y \end{bmatrix}$
4. $[\mathrm{Proj}_\theta] \begin{bmatrix} x \\ y \end{bmatrix}$

3.3 Linear transformations and bases

As it turns out, in order to study linear transformations between vector spaces, it suffices to study maps on the basis of the source.

Let V and W be vector spaces. Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be a basis of V.

Proposition 3.3. *An arbitrary map $\varphi : \mathcal{B} \to W$ can be extended to a linear transformation $\mathcal{L} : V \to W$.*

Proof. Consider an arbitrary map

$$\begin{aligned} \varphi : \mathcal{B} &\longrightarrow W, \\ \vec{v}_i &\longmapsto \varphi(\vec{v}_i). \end{aligned}$$

For any vector $\vec{v} \in V$, there exist unique scalars $c_1, \ldots, c_k$ such that $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$. Define a new map $\mathcal{L}_\varphi : V \to W$ as,

$$\mathcal{L}_\varphi(\vec{v}) = c_1\varphi(\vec{v}_1) + \cdots + c_k\varphi(\vec{v}_k).$$

Q. 40. Prove that $\mathcal{L}_\varphi$ is a linear transformation.

□

Proposition 3.4. *Two linear transformations $\mathcal{L}, \mathcal{L}' : V \to W$ which agree on $\mathcal{B}$ are the same. More precisely,*

$$\begin{aligned} \text{if } \ \mathcal{L}\vec{v}_i &= \mathcal{L}'\vec{v}_i \ \text{ for all } \vec{v}_i \text{ in } \mathcal{B}, \\ \text{then } \ \mathcal{L}\vec{v} &= \mathcal{L}'\vec{v} \ \text{ for all } \vec{v} \text{ in } V. \end{aligned}$$

Q. 41. Prove this using Theorem 1.6.

Combining the previous two propositions we get the following correspondence.

Theorem 3.5. *Let V and W be vector spaces. Let $\mathcal{B}$ be a basis of V. There is a 1-1 correspondence between set maps $\mathcal{B} \to W$ and linear transformations $V \to W$.*

$$\begin{aligned} \{\text{set maps } \mathcal{B} \to W\} &\longleftrightarrow \{\text{linear transformations } V \to W\} \\ \varphi &\longmapsto \mathcal{L}_\varphi \\ \mathcal{L}|_{\mathcal{B}} &\longleftarrow\!\shortmid \mathcal{L} \end{aligned}$$

3.4 Optional: Eigenvalues and eigenvectors

Definition 3.6. For a linear operator $\mathcal{L} : V \to V$ a non-zero vector $\vec{v}$ in V is called an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if

$$\mathcal{L}(\vec{v}) = \lambda\vec{v}.$$

The collection of all eigenvalues is called the *spectrum* of $\mathcal{L}$.

Q. 42. Let V_λ be the set of eigenvectors with eigenvalue λ along with the vector $\vec{0}$.

$$V_\lambda = \{\vec{v} \in V \mid \mathcal{L}(\vec{v}) = \lambda\vec{v},\, \vec{v} \neq \vec{0}\} \cup \{\vec{0}\}$$

Show that V_λ is a vector space.

V_λ is called the *eigenspace* for the eigenvalue λ. If the total dimensions of all the eigenspaces equals the dimension of V then we say that $\mathcal{L}$ is *diagonalizable*.

Q. 43. For each of the following linear operators on $\mathbb{R}^2$, find the spectrum and eigenspaces.

1. $\mathrm{Id}_{\mathbb{R}^2}$
2. Rot_θ
3. Refl_θ
4. Proj_θ

Why care about eigenvalues and eigenvectors? There are far too many answers for this, here is just one. Oftentimes, a quantum mechanical system is modelled using a linear opearator (called the Hamiltonian of the system) on a vector space of possible states. The spectrum then is the possible energy values the system can attain and the vectors in the eigenspaces are the steady states for the system i.e. states which do not change over time. If the linear operator is diagonalizable, then the system simplifies drastically and one can study each eigenspace separately.

To learn more about eigenvalues and eigenvectors consider taking Mark's class in W2 and to learn more about their use in Quantum Mechanics, consider taking Nic Ford's class in W3.

Solutions to selected problems - 2

Solution to Q.33 Part 1. For linear transformations $\mathcal{L} : V \to W$ and $\mathcal{L}' : W \to U$,

1.
$$\begin{aligned} \mathcal{L}' \circ \mathcal{L}(c\vec{v}) &= \mathcal{L}'(\mathcal{L}(c\vec{v})) \\ &= \mathcal{L}'(c\mathcal{L}(\vec{v})) && \text{as } \mathcal{L} \text{ is a linear transformation} \\ &= c\mathcal{L}'(\mathcal{L}(\vec{v})) && \text{as } \mathcal{L}' \text{ is a linear transformation} \end{aligned}$$

2.
$$\begin{aligned} \mathcal{L}' \circ \mathcal{L}(\vec{v}+\vec{w}) &= \mathcal{L}'(\mathcal{L}(\vec{v}+\vec{w})) \\ &= \mathcal{L}'(\mathcal{L}(\vec{v}) + \mathcal{L}(\vec{w})) && \text{as } \mathcal{L} \text{ is a linear transformation} \\ &= \mathcal{L}'(\mathcal{L}(\vec{v})) + \mathcal{L}'(\mathcal{L}(\vec{w})) && \text{as } \mathcal{L}' \text{ is a linear transformation} \end{aligned}$$

And so $\mathcal{L}' \circ \mathcal{L}$ is a linear transformation. □

Solution to Q.35. $\mathcal{L} : \mathbb{R}^2 \to \mathbb{R}^2$ sending $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+1 \\ y \end{bmatrix}$ is not a linear operator, as it preserves neither scalar multiplication nor vector addition.

$\mathcal{L}(\begin{bmatrix} x \\ y \end{bmatrix}) + \mathcal{L}(\begin{bmatrix} x' \\ y' \end{bmatrix}) = \begin{bmatrix} x+1 \\ y \end{bmatrix} + \begin{bmatrix} x'+1 \\ y' \end{bmatrix} = \begin{bmatrix} x+x'+2 \\ y+y' \end{bmatrix}$

$\mathcal{L}(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x' \\ y' \end{bmatrix}) = \begin{bmatrix} x+x' \\ y+y' \end{bmatrix} = \begin{bmatrix} x+x'+1 \\ y+y' \end{bmatrix}$

so $\mathcal{L}(\begin{bmatrix} x \\ y \end{bmatrix}) + \mathcal{L}(\begin{bmatrix} x' \\ y' \end{bmatrix}) \neq \mathcal{L}(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x' \\ y' \end{bmatrix})$. □

Solution to Q.36. A linear transformation $\mathcal{L} : \mathbb{R}^1 \to \mathbb{R}^1$ must preserve scalar multiplication. Hence,

$$\mathcal{L}([c]) = \mathcal{L}(c[1]) = c\mathcal{L}([1]) \tag{*}$$

The addition condition gives us

$$\begin{aligned} & \mathcal{L}([c+d]) = \mathcal{L}([c]) + \mathcal{L}([d]) \\ \implies & (c+d)\mathcal{L}([1]) = c\mathcal{L}([1]) + d\mathcal{L}([1]) && \text{by (*)} \end{aligned}$$

But this is always true. So the addition condition does not provide us any new information about $\mathcal{L}$.

There are no other conditions. Hence a linear transformation $\mathcal{L} : \mathbb{R}^1 \to \mathbb{R}^1$ is completely determined by $\mathcal{L}([1])$, which can be any real number, say $[\alpha]$. Then $\mathcal{L}([c])$ equals $[c\alpha]$, i.e. $\mathcal{L}$ is a scaling by α. □

Solution to Q.38 Part 2. In order to determine the matrix corresponding to Rot_θ we need to determine $\text{Rot}_\theta(\vec{e}_1)$ and $\text{Rot}_\theta(\vec{e}_2)$. By basic trigonometry,

$$\text{Rot}_\theta(\begin{bmatrix}1\\0\end{bmatrix}) = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$$
$$\text{Rot}_\theta(\begin{bmatrix}0\\1\end{bmatrix}) = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$$

Hence, the corresponding matrix is

$$[\text{Rot}_\theta] = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

□

Solution to Q.39 Part 2. Using the formula for matrix multiplication

$$\begin{aligned}[\text{Rot}_\theta]\begin{bmatrix}x\\y\end{bmatrix} &= \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}\\ &= x\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} + y\begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}\\ &= \begin{bmatrix}x\cos\theta - y\sin\theta\\ x\sin\theta + y\cos\theta\end{bmatrix}\end{aligned}$$

□

4 Rank & Nullity

We saw that matrices encode linear transformations between Euclidean spaces. A matrix of size $\ell \times k$ (having ℓ rows and k columns) can multiply a vector of size k to get a vector of size ℓ.

$$\begin{aligned} A : \mathbb{R}^k &\longrightarrow \mathbb{R}^\ell \\ \vec{v} &\longmapsto A\vec{v} \end{aligned}$$

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell 1} & A_{\ell 2} & \cdots & A_{\ell k} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} = v_1 \begin{bmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{\ell 1} \end{bmatrix} + v_2 \begin{bmatrix} A_{12} \\ A_{22} \\ \vdots \\ A_{\ell 2} \end{bmatrix} + \cdots + v_k \begin{bmatrix} A_{1k} \\ A_{2k} \\ \vdots \\ A_{\ell k} \end{bmatrix} \tag{4.1}$$

$$= \begin{bmatrix} A_{11}v_1 + A_{12}v_2 + \cdots + A_{1k}v_k \\ A_{21}v_1 + A_{22}v_2 + \cdots + A_{2k}v_k \\ \vdots \\ A_{\ell 1}v_1 + A_{\ell 2}v_2 + \cdots + A_{\ell k}v_k \end{bmatrix}$$

$$[\text{size } \ell \times k] \cdot [\text{size } k] = [\text{size } \ell]$$

Q. 44. Using equation (4.1), check that the i^{th} column of the matrix A is exactly the vector $A\vec{e}_i$.

Q. 45 (Practice problems)**.** Compute $A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ for each of the following matrices.

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

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

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

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

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

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

4.1 Linear systems

A *linear system* (of ℓ equations in k variables) is a collection of linear[1] equations.

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1k}x_k &= \alpha_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2k}x_k &= \alpha_2 \\ &\vdots \\ A_{\ell 1}x_1 + A_{\ell 2}x_2 + \cdots + A_{\ell k}x_k &= \alpha_\ell \end{aligned} \tag{4.2}$$

where A_{ij}'s and α_i's are real numbers, and x_i's are the variables we are solving for. By equation (4.1) this is the same as

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell 1} & A_{\ell 2} & \cdots & A_{\ell k} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_\ell \end{bmatrix}$$

$$A\vec{x} = \vec{\alpha}$$

Thus matrices encode not only the linear transformations but also linear systems. This observation allows us to study linear systems using techniques from linear algebra.

4.1.1 Image

Definition 4.1. Let A be a matrix of size $\ell \times k$. The *image* of a matrix A is simply the image of the corresponding linear transformation $A : \mathbb{R}^k \to \mathbb{R}^\ell$ i.e. it is the set of all vectors $\vec{y} \in \mathbb{R}^\ell$ such that $A\vec{x} = \vec{y}$ for some $\vec{x} \in \mathbb{R}^k$.

$$\mathrm{im}(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^k\}$$

[1]Linear = degree 1 = only addition and scalar multiplication.

Q. 46. Show that the $\mathrm{im}(A)$ is a subspace of $\mathbb{R}^\ell$.

Definition 4.2. The dimension of $\mathrm{im}(A)$ is called the *rank* of A.

$$\operatorname{rank} A := \dim \mathrm{im}(A)$$

Q. 47. Check that equation (4.2) having a solution is the same as saying that $\vec{\alpha}$ is in the image of A.

Q. 48. Using (4.1), show that $\mathrm{im}(A)$ is precisely the span of the column vectors of A.

Combining all the above statements we get,

Theorem 4.3. *The equation (4.2) has a solution if and only if the vector $\vec{\alpha}$ is in the vector space spanned by the column vectors of A.*

Finding the image i.e. the span of column vectors in general is non-trivial. However, there is a very fast algorithm called Gaussian elimination for computing it.

Q. 49. Find rank of each of the following matrices.

1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
2. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
3. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
4. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$
5. $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$
6. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Q. 50. (Do this problem if you have done the section on linear transformations.) Using geometry, find rank of each of the following matrices.

1. $[\mathrm{Id}_{\mathbb{R}^2}]$
2. $[\mathrm{Rot}_\theta]$
3. $[\mathrm{Refl}_\theta]$
4. $[\mathrm{Proj}_\theta]$

4.1.2 Kernel

Once we know a solution exists, we can ask - what is the number of solutions? In general, there can be infinitely many solutions, but the dimension of the solution space is finite. This dimension is called the nullity.

Definition 4.4. The *kernel* of a matrix A is the set of vector $\vec{x} \in \mathbb{R}^k$ such that $A\vec{x} = 0$.

$$\ker A := \{\vec{x} \mid A\vec{x} = \vec{0}\} \subseteq \mathbb{R}^k$$

Q. 51. Prove that $\ker A$ is a subspace of $\mathbb{R}^k$.

Definition 4.5. The *nullity* of A is the dimension of $\ker A$.

$$\operatorname{null} A := \dim \ker(A)$$

Q. 52. Find nullity of each of the following matrices.

1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
2. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
3. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
4. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$
5. $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$
6. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Q. 53. (Do this problem if you have done the section on linear transformations.) Using geometry, find nullity of each of the following matrices.

1. $[\mathrm{Id}_{\mathbb{R}^2}]$
2. $[\mathrm{Rot}_\theta]$
3. $[\mathrm{Refl}_\theta]$
4. $[\mathrm{Proj}_\theta]$

Q. 54. Consider a matrix A of size $\ell \times k$ and a vector $\vec{\alpha} \in \mathbb{R}^\ell$.

1. Show that if $\vec{x}_1, \vec{x}_2$ are solutions of $A\vec{x} = \vec{\alpha}$ then $\vec{x}_1 - \vec{x}_2$ is a solution of $A\vec{x} = \vec{0}$.
2. Show that if $\vec{x}_1$ is a solution of $A\vec{x} = \vec{\alpha}$ and $\vec{x}_0$ is a solution of $A\vec{x} = \vec{0}$ then $\vec{x}_1 + \vec{x}_0$ is also a solution of $A\vec{x} = \vec{\alpha}$.

Combining everything above we have the following fundamental theorem about the solution space of the system of equations.

Theorem 4.6. *Let A be a matrix of size $\ell \times k$ and let $\vec{\alpha}$ be a vector in $\mathbb{R}^{\ell}$.*

1. *The linear system $A\vec{x} = \vec{\alpha}$ has a solution if an only if $\vec{\alpha}$ is in* $\operatorname{im} A$.
2. *If a solution exists, then the space of solutions has dimension*[1] $\operatorname{null} A$.

Corollary 4.7. *If* $\operatorname{null} A = 0$ *then $A\vec{x} = \vec{\alpha}$ has either a unique solution or no solution. If* $\operatorname{null} A > 0$ *then $A\vec{x} = \vec{\alpha}$ has either infinitely many solutions or no solution.*

Once we know that we are looking for vector spaces we can use standard algorithms in linear algebra to find the solutions.

Q. 55. For your computations in Questions 49, 50, 52, and 53 what is $\operatorname{rank} A + \operatorname{null} A$?

$$\operatorname{rank} A + \operatorname{null} A = ?$$

[1] This is not strictly correct as the space of solutions of $A\vec{x} = \vec{\alpha}$ is not a vector space if $\alpha \neq 0$. Instead, we need to shift the space of solutions to the origin to make it a vector space. This is analogous to the fact that a line not passing through the origin is not a vector space, but we can shift it to the origin to make it one. Once shifted, the space of solutions becomes $\ker A$.

5 Final Remarks

Don't just read it; fight it!
Ask your own questions, look for your own examples, discover your own proofs. Is the hypothesis necessary? Is the converse true? What happens in the classical special case? What about the degenerate cases? Where does the proof use the hypothesis?

Paul Halmos

On the last day, we will just tie up some loose ends.

5.1 Rank-Nullity theorem

The following is first of the many simple and yet remarkably useful theorems in linear algebra.

Theorem 5.1 (Rank-Nullity theorem)**.** *Let A be a matrix of size $\ell \times k$ so that A defines a linear transformation $A : \mathbb{R}^k \to \mathbb{R}^\ell$. Then,*

$$\operatorname{null} A + \operatorname{rank} A = k.$$

Q. 56. Rank-nullity can be used to detect isomorphisms. Let A be a square matrix of size $k \times k$ representing a linear operator $A : \mathbb{R}^k \to \mathbb{R}^k$.

Using the Rank-Nullity theorem, show that the following are equivalent

1. A is an isomorphism,
2. $\operatorname{null} A = 0$,
3. $\operatorname{rank} A = k$.

Remark 5.2. There is one more criterion for detecting isomorphisms

4. $\det A \neq 0$,

where $\det A$ is the "determinant of A". Determinant is the signed volume of the k-dimensional cube formed by the columns of A, and can be computed very efficiently using Gaussian elimination.

Q. 57. Let a, b, c be scalars. Assume that at least one of a, b, c is non-zero.

1. Show that the matrix $A = \begin{bmatrix} a & b & c \end{bmatrix}$ has rank 1.
2. Using the Rank-Nullity Theorem, conclude that $ax + by + cz = 0$

defines a plane in $\mathbb{R}^3$.

The same proof can be used to show that a single non-trivial linear equation always defines an $n-1$ dimensional hyperplane in $\mathbb{R}^n$.

Proof of the Rank-Nullity theorem. Let $A : \mathbb{R}^k \to \mathbb{R}^\ell$. Let

$$n = \operatorname{null} A.$$

To prove the theorem, we will construct a basis of $\operatorname{im} A$ of size $k-n$.

Let

$$\mathcal{B}' = \{\vec{v}_1, \dots, \vec{v}_n\}$$

be a basis of $\ker A$.

Case 1: $n = k$. In this case, we have $\ker A = \mathbb{R}^k$. By the definition of kernel, $A\vec{v} = 0$ for all $\vec{v} \in \mathbb{R}^k$. Hence, $\operatorname{im} A = \{\vec{0}\}$ which implies that $\operatorname{rank} A = 0$.

Case 2: $n < k$. In this case, by repeatedly adding linearly independent vectors as in the proof of Theorem 2.4, we can extend $\mathcal{B}'$ to a basis $\mathcal{B}$ of $\mathbb{R}^k$.

$$\mathcal{B} = \{\underbrace{\vec{v}_1, \dots, \vec{v}_n}_{\mathcal{B}'}, \vec{w}_{n+1}, \dots, \vec{w}_k\}.$$

Claim: The set $\mathcal{B}'' = \{A\vec{w}_{n+1}, \dots, A\vec{w}_k\}$ is a basis for $\operatorname{im} A$.

We need to show that $\operatorname{Span}(\mathcal{B}'') = \operatorname{im} A$ and $\mathcal{B}''$ is linearly independent. To prove the first part, note that any vector in V can be written as

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + c_{n+1}\vec{w}_{n+1} + \cdots + c_k\vec{w}_k \tag{*}$$

for some scalars $c_1, \dots, c_k$.

Q. 58. Apply A to both sides of (*) to conclude that

$$\operatorname{im} A = \operatorname{Span}(A\vec{w}_{n+1}, \dots, A\vec{w}_k)$$

It remains to show that $\mathcal{B}''$ is linearly independent. Suppose there are scalars $d_{n+1}, \dots, d_k$ such that

$$\begin{aligned}
& d_{n+1}A\vec{w}_{n+1} + \cdots + d_kA\vec{w}_k = 0 & \\
\implies & A(d_{n+1}\vec{w}_{n+1} + \cdots + d_k\vec{w}_k) = 0 & \text{by linearity} \\
\implies & d_{n+1}\vec{w}_{n+1} + \cdots + d_k\vec{w}_k \in \ker A & \text{by definition of } \ker \\
\implies & d_{n+1}\vec{w}_{n+1} + \cdots + d_k\vec{w}_k \in \operatorname{Span}(\vec{v}_1, \dots, \vec{v}_n) & \\
\implies & d_{n+1}\vec{w}_{n+1} + \cdots + d_k\vec{w}_k = d_1\vec{v}_1 + \cdots + d_n\vec{v}_n & \text{by definition of } \operatorname{Span}
\end{aligned}$$

for some scalars $d_1, \ldots, d_n$.

But the vectors $\{\vec{v}_1, \ldots, \vec{v}_n, \vec{w}_{n+1}, \ldots, \vec{w}_k\}$ form a basis for V and hence are linearly independent. Hence, the only possible d_i's satisfying the last equation are $d_1 = \cdots = d_k = 0$ which proves the linear independence of $\mathcal{B}''$. $\square$

Remark 5.3. Rank-Nullity theorem is closely related to theorems from other areas of mathematics. For example, vector spaces are abelian groups under the vector addition operation and subspaces are the same as (normal) subgroups. The first isomorphism theorem in group theory then gives us

$$\mathbb{R}^k / \ker A \cong \operatorname{im} A.$$

The Rank-Nullity theorem is then a version of Lagrange's theorem.

Q. 59 (If you know what an exact sequence means)**.** Using the Rank-Nullity theorem, show that if

$$0 \to V_0 \to V_1 \to \cdots \to V_n \to 0$$

is an exact sequence of vector spaces then

$$\dim V_0 - \dim V_1 + \cdots + (-1)^i \dim V_i + \cdots + (-1)^n \dim V_n = 0$$

This is saying that the Euler characteristic of an exact sequence is trivial.

5.2 Composition of linear transformations

We saw in Section 3 that composition of linear transformations is itself a linear transformation. For Euclidean spaces, this gives rise to products of matrices.

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, "B"] \arrow[rr, bend right, "AB"'] & \mathbb{R}^\ell \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd}$$

By carefully expanding out $A(B(\vec{v}))$ one can find the general formula for AB, this is quite tedious to derive but easy to use. You should look it up.

Here is the formula for multiplying two 2×2 matrices.

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "B"] \arrow[rr, bend right, "AB"'] & \mathbb{R}^2 \arrow[r, "A"] & \mathbb{R}^2 \end{tikzcd}$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \tag{5.1}$$

Q. 60. In Section 3, we computed the identity, rotation, and reflection matrices

$$[\mathrm{Id}_{\mathbb{R}^2}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$[\mathrm{Rot}_\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$[\mathrm{Refl}_\theta] = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

Using (5.1) compute

1. $[\mathrm{Id}_{\mathbb{R}^2}] \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$
2. $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} [\mathrm{Id}_{\mathbb{R}^2}]$
3. $[\mathrm{Rot}_\theta][\mathrm{Rot}_\varphi]$
4. $[\mathrm{Refl}_\theta][\mathrm{Refl}_\theta]$
5. $[\mathrm{Refl}_\theta][\mathrm{Refl}_\varphi]$

See any old friends? Interpret these geometrically.

Q. 61 (Optional)**.** Derive the formula for matrix multiplication of 2×2 matrices using compositions of linear transformations.

Congratulations on making it this far!!!(Yay!) What we have "covered" in the past week is more than what you would do in a month of regular class. There is hardly an area of mathematics that does not use linear algebra, the more math you do the better you'll be able to understand and appreciate it. Here are some suggested topics to read from here:

1. Change of basis theorem,
2. Eigenvalues and eigenvectors,
3. Determinants,
4. Gaussian elimination,
5. Inner product spaces,
6. Spectral theorem,
7. Jordan canonical forms,
8. Matrix groups.